



## New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method

Rencheng Song<sup>a</sup>, Senle Zhang<sup>a</sup>, Juan Cheng<sup>a</sup>, Chang Li<sup>a</sup>, Xun Chen<sup>b,\*</sup>

<sup>a</sup> Department of Biomedical Engineering, Hefei University of Technology, Hefei, Anhui, 230009, China

<sup>b</sup> Department of Electronic Science and Technology, University of Science and Technology of China, Hefei, Anhui, 230026, China

### ARTICLE INFO

MSC:  
00-01  
99-00

#### Keywords:

Remote photoplethysmography  
Heart rate estimation  
Semi-blind source separation  
Camera quantization noise  
Video resolution

### ABSTRACT

Remote photoplethysmography (rPPG), a non-contact technique to estimate heart rates (HR) from video recordings, has attracted much attention from researchers in recent years. It is well-known that rPPG signals can be extracted from low-resolution videos. However, the measurement quality may degrade due to camera quantization noise if only a small number of pixels are within the skin region of interest. The purpose of this paper is to comprehensively investigate the benefit of using a super-high resolution for the rPPG-based HR estimation under various shooting distances. A new semi-blind source separation (semi-BSS) rPPG method, which is proposed to combine the advantages of BSS and model-based methods, is fully tested on both the public UBFC-RPPG and self-collected video datasets. The experimental results demonstrate that the new semi-BSS method outperforms several existing techniques. A consistent and remarkable improvement on the rPPG signal quality has been observed with the super-high resolution when the shooting distance is no less than 1.0 m. This indicates that selecting an appropriate resolution based on a given shooting distance also plays a crucial role to improve the quality of rPPG measurements.

### 1. Introduction

Heart rate (HR) is one of the most important physiological parameters to evaluate an individual's health and affective state. The HR can be measured in both contacting and non-contacting ways [1]. Compared to traditional electrocardiography [2] and photoplethysmography [3] measurements, which need specific sensors to touch with the skin of a subject, remote photoplethysmography (rPPG) is a contactless HR measurement method detecting blood volume variations in the microvascular bed of tissue from facial videos. RPPG monitors the heart rate in a simple, convenient, and non-invasive way. It has many potential applications such as remote patient monitoring [4], vivo detection [5], and driver fatigue assessment [6] etc.

However, many challenges may hinder the applications of rPPG. The accuracy of this physiological measurement can be disrupted by many factors such as poor signal strength, motion artifacts [7] and illumination variations [8]. The rPPG pulsation is very weak which is not visible to naked eyes and therefore is easily disturbed by noise. The motion can cause an angle change between the camera and skin region of interest (ROI), which further leads to motion-induced changes of light source intensity. The illumination variation may change the

source light spectrum or intensity in the ROI. The latter two factors have been studied intensively in existing rPPG literature. Relevant reviews can be found in [9].

On the other hand, camera parameters such as the video compression and video resolution also have a salient interference on the rPPG measurement accuracy. McDuff et al. [10] studied the impact of video compression and indicated the compression degraded the quality of non-contacting pulse signal remarkably. Zhao et al. [11] reviewed papers considering the impact of video compression and introduced a novel rPPG signal preserving video compression algorithm to overcome this difficulty.

Besides the impact of video compression, the influence of video resolution is also complex. Most existing rPPG studies have been conducted with relatively low video resolutions because high-resolution videos require more storage space, transmission bandwidth and computational burden. However, if the video resolution is too low, the rPPG pulse can be easily interfered by the quantization noise of image sensors, which is difficult to be eliminated through a spatial averaging in video frames. Even if the resolution is high, pixels within a ROI may still not be adequate at a far shooting distance. Meanwhile, the signal strength of a rPPG pulse also decreases along with distances due

\* Corresponding author.

E-mail addresses: [rcsong@hfut.edu.cn](mailto:rcsong@hfut.edu.cn) (R. Song), [zhangsenle@mail.hfut.edu.cn](mailto:zhangsenle@mail.hfut.edu.cn) (S. Zhang), [chengjuan@hfut.edu.cn](mailto:chengjuan@hfut.edu.cn) (J. Cheng), [changli@hfut.edu.cn](mailto:changli@hfut.edu.cn) (C. Li), [xunchen@ustc.edu.cn](mailto:xunchen@ustc.edu.cn) (X. Chen).

<https://doi.org/10.1016/j.complbiomed.2019.103535>

Received 24 July 2019; Received in revised form 6 November 2019; Accepted 7 November 2019

Available online 16 November 2019

0010-4825/© 2019 Elsevier Ltd. All rights reserved.

to a dropping amount of light received by the camera sensor [12]. Therefore, it is necessary to investigate the effect of video resolution, especially the rarely used super-high resolution, under the consideration of varying shooting distances.

To evaluate the resolution effect, a new rPPG-based HR extraction method was tested on a self-collected dataset covering various resolution and distance configurations. The proposed method applies a fast kernel density independent component analysis (KDICA) [13] on the chrominance signals defined from an optical reflection model [14]. As known, the CHROM method [14] extracts pulse explicitly from chrominance signals which is considered to be robust against motion artifacts. However, the alpha tuning in CHROM may fail if the strengths of pulse and specular signals are at a similar level [15]. On the other hand, the KDICA is a nonparametric blind source separation (BSS) algorithm that does not require a prior assumption on distributions of hidden sources. It has been proved in general to be more accurate and robust than the classical parametric-based independent component analysis (ICA) algorithms such as FastICA [16]. Our semi-BSS method employs the KDICA instead of alpha tuning to extract pulse from chrominance signals which combines the advantages of BSS and model-based approaches. We verify the superiority of the proposed method over several conventional methods on a public video dataset UBFC-RPPG [17]. The proposed rPPG method was then tested thoroughly on the self-collected dataset to demonstrate the effect of super-high resolution for rPPG under different distances.

In summary, the contributions of this paper are twofold. First, we evaluated the effect of video resolutions, especially the super-high 2.7K one, to rPPG measurement quality considering various distances. The results reveal the benefit of using high resolution for rPPG-based HR extraction even for short shooting distances. This breaks the misperception that high resolution is only necessary at long distances. Second, we proposed a new semi-BSS rPPG algorithm which combined the benefits of the model-based method and the blind source separation method. The experimental results have verified the best performance of our new method.

## 2. Related work

### 2.1. Typical rPPG methods

Many methods have been proposed to extract rPPG-based pulse signals. We will list some typical methods below such as the blind source separation (BSS) based methods, the skin optical reflection model based methods and deep learning based rPPG methods. For more details, please refer to some recent reviews like [9] and [18].

The BSS-based rPPG methods usually assume the target signal to satisfy some statistical nature, such as independence or correlation. Poh et al. [19] firstly employed the ICA to extract a pulse signal from raw RGB traces. Lam [20] further used ICA on multiple random patches and finally selected the optimal one through a histogram analysis. Wei et al. [21] introduced a second-order BSS to estimate HR with RGB signals from dual facial ROIs. Qi et al. [22] proposed to use a joint blind source separation (JBSS) for rPPG measurement. The JBSS considers the intrinsic correlations of HR among different facial regions. Recently, Al-Naji et al. [23] measured the rPPG signal through a canonical correlation analysis (CCA) combined with the ensemble empirical mode decomposition (EEMD).

Different from the BSS-based methods, the model-based methods assume the pulse signal to satisfy a skin optical reflection model [15]. De Haan et al. [14] firstly proposed a chrominance-based signal processing method (CHROM) to explicitly extract pulse signal against specular and motion artifacts. The RGB channels were projected into a chrominance subspace where the motion component was greatly eliminated. Later, they introduced a PBV [24] method, which defined an optimal projection assuming the knowledge of the blood volume pulse signature. Wang et al. [15] introduced a POS method to use a different projection

**Table 1**

Papers considering the influence of video resolution for rPPG.

Paper	Distance (m)	Resolution
Sun et al. [29]	0.35 (0.2)	640 × 512 (320 × 240)
Han et al. [30]	0.3~1.8 m with a 0.3 m step	640 × 480
Blackford et al. [31]	1.5	658 × 492 and 329 × 246
Blackford et al. [12]	25,50,100	1920 × 1080
Blackford et al. [32]	25,50,100,150	1920 × 1080
Ibrahim et al. [33]	1.0,3.0,5.0	1440 × 1080
McDuff et al. [34]	1.5	658 × 492,164 × 120 41 × 30, DRCN <sup>a</sup>
Our paper	0.5, 1.0, 1.5 2.0, 2.5, 3.0	2704 × 1520,1920 × 1080 1280 × 720, 640 × 480, 320 × 240

<sup>a</sup>DRCN indicates a deeply-recursive convolutional network model for image super resolution.

orthogonal to the skin tone compared to CHROM. The POS method was considered to be more robust in complex illumination scenarios. The relation and differences of all those model-based methods were discussed in paper [15].

Recently, inspired by the progress of deep learning methods in computer vision, more and more deep learning based rPPG methods have been proposed to estimate heart rate. In 2018, an end-to-end system with a soft-attention mask was proposed by Chen et al. [25] to establish a mapping between the contrast of video frames and the corresponding pulse derivative. Later, Niu et al. [26] employed a convolutional neural network (CNN) to map the spatiotemporal features of the cardiac information to its heart rate value. Qiu et al. [27] proposed a different method to construct the spatiotemporal feature map. They also used the Eulerian Video Magnification (EVM) [28] to enhance the signal-to-noise ratio through magnifying facial color changes. Since all these deep learning based methods are data-driven, a large amount of training data are required to fit the network. The diversity and quality of the training data may affect the performance significantly.

In summary, the deep learning based rPPG methods have great potentials. But they are still in an early developing stage. Currently, the well-known BSS- and model-based methods are still the mainstream methods in rPPG research. ICA and CHROM are two representative methods that are used very frequently. However, the commonly used FastICA is a parametric BSS method which requires a prior assumption on the distributions of hidden sources. It may not be realistic for practical rPPG applications. The CHROM method assumes a standardized skin-tone vector. It explicitly extracts the pulse signal using an alpha tuning. But the prior information used by alpha tuning may not match with the actual situation, which can lead to a failure of this method.

Considering the above limitations, we will try to design a new approach that combines the advantages of BSS- and model-based methods while overcoming their main drawbacks.

### 2.2. Studies considering video resolution

The influence of video resolution has already been considered in some existing papers from different perspectives. However, to our best knowledge, those works mainly take into account two situations: (1) A fixed resolution with varying distances; (2) Multiple resolutions with a fixed distance. The related literature are summarized in Table 1.

In 2012, Sun et al. [29] compared rPPG measurements using two cameras with different resolutions. No apparent difference was observed from a quite near distance around 0.2 ~ 0.3 m. In 2015, Han et al. [30] explored an optimal skin-camera distance using several different cameras capturing videos at a 640x480 resolution. The shooting distances varied from 0.3 m to 1.8 m with a 0.3 m step. Blackford et al. [31] considered the influence of video resolution at a fixed distance of 1.5 m. Two resolutions (658x492 and 329x246) were explored and little difference was observed in results. Later in

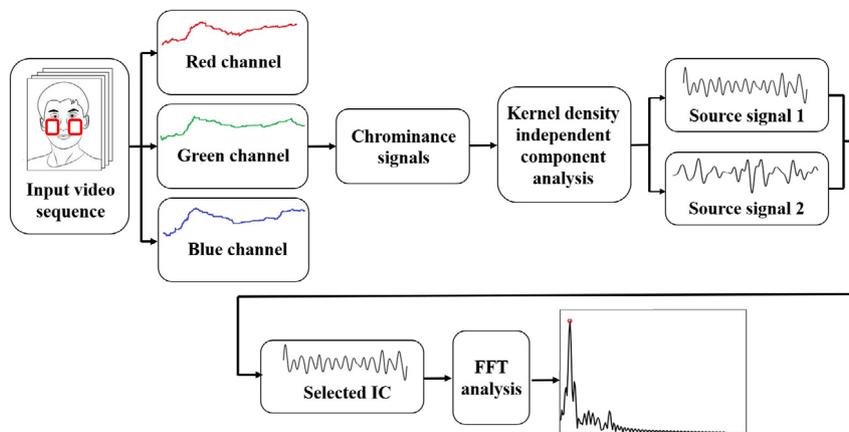


Fig. 1. Video-based non-contact physiological parameter extraction flowchart.

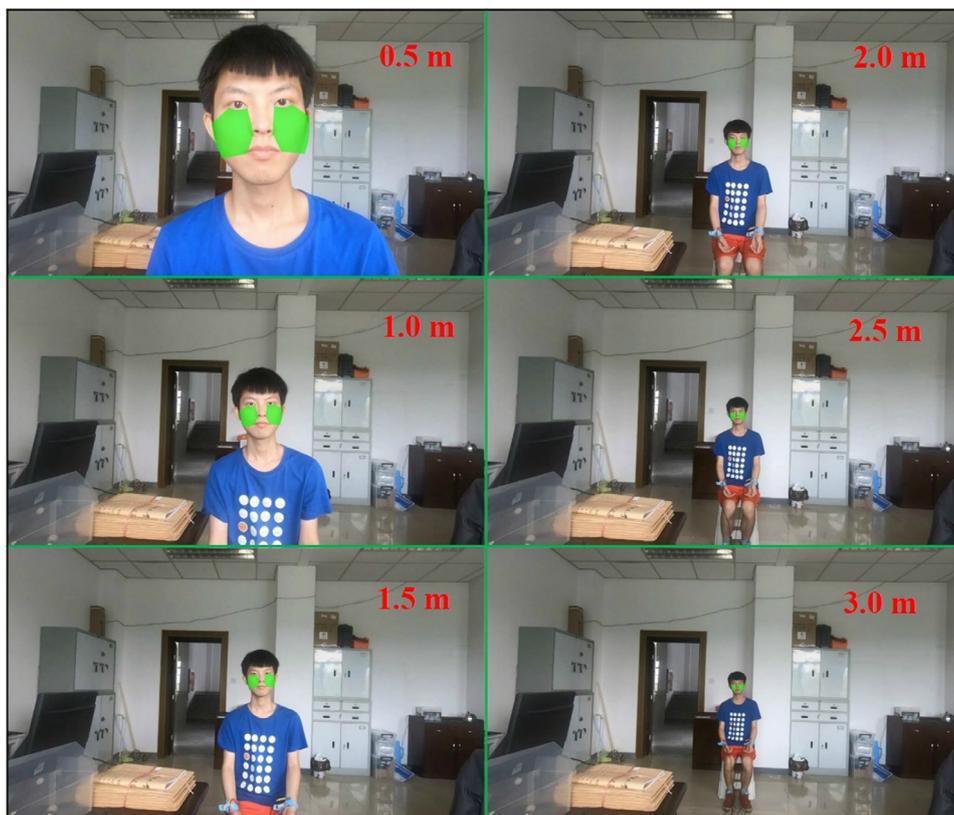


Fig. 2. Region of interest at different distances.

2016 and 2017, they further [12,32] investigated the rPPG-based HR extraction at 25 m, 50 m, 100 m and 150 m ultra-long distances with a 1080p high resolution. The results showed that high resolution could greatly reduce quantization noise when the distance was far. In 2018, Ibrahim et al. [33] explored the influence of ROI selection on rPPG solutions at different distances (1 m, 3 m, 5 m) with a fixed resolution at  $1440 \times 1080$ . McDuff [34] also demonstrated the effectiveness of image super-resolution to improve the signal-to-noise ratio of rPPG measurement at a 1.5 m distance.

As discussed above, it is well known that rPPG can work with relatively low resolutions [29,31] at near shooting distances. However, high resolution is still necessary if the distance is too far or pixels within ROI are inadequate [12,32,34]. It remains to be unknown about the influence of resolution for the intermediate state. For this purpose,

this paper attempts to take a full investigation of the reciprocity relation of resolution and distance on the rPPG accuracy considering various configurations. Particularly, we will explore the benefit of using a super-high resolution for rPPG measurement under several near shooting distances.

### 3. Method

The framework of the proposed rPPG-based HR extraction method is as shown in Fig. 1. Firstly, a video covering face region is recorded using a color camera. Then the facial landmarks are detected and tracked in each video frame. A ROI on the cheek area is determined by corresponding facial landmarks. Pixels within the ROI are averaged in each frame to get time series. The RGB sequences are further projected to generate chrominance signals. The fast kernel density independent

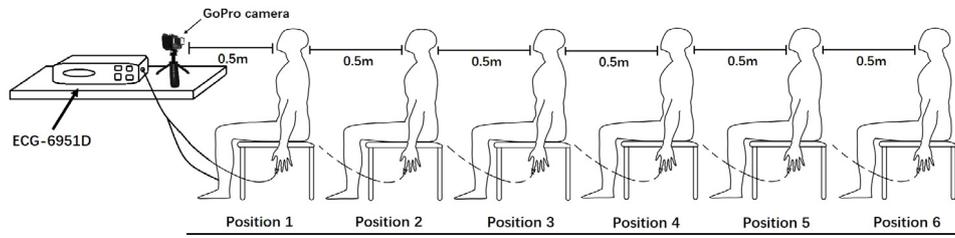


Fig. 3. Experimental setup.



Fig. 4. Region of interest for videos in UBFC-RPPG dataset: an example from subject 30.

analysis (KDICA) algorithm [13] is then used to extract the desired pulse signal. Finally, the dominant frequency of the candidate signal is calculated by a Fast Fourier Transform (FFT) based spectrum analysis to obtain the HR value of the human body.

### 3.1. Region of interest

In this paper, we adopt a 68-point facial landmark detection algorithm [35] based on a multi-task learning to determine the landmarks. A standard facial tracker such as the Kanade–Lucas–Tomasi feature tracker is employed to accelerate the landmark detections in video frames. This ensures an accurate and fast facial landmark detection which also helps to compensate for the interference of motion artifacts. The ROI is defined as the left and right cheek regions with corresponding landmarks. For different distances, the same regions are determined as ROIs, as shown in Fig. 2. This is to make sure the evaluation of the resolution effect is against a uniform baseline.

### 3.2. The proposed rPPG algorithm

Wang et al. [15] explain the ICA-based and the model-based rPPG methods in a unified mathematical model which describes the skin optical reflection mechanism. Under the roof of this model, the chrominance signals  $S = \begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix}$  in CHROM method can be written as

$$\begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \begin{pmatrix} 3R_f - 2G_f \\ 1.5R_f + G_f - 1.5B_f \end{pmatrix} = A \cdot \begin{pmatrix} i(t) \\ p(t) \end{pmatrix} \quad (1)$$

where  $A$  is a  $2 \times 2$  mixing matrix,  $i(t)$  indicates the light intensity variation,  $p(t)$  is the rPPG pulse signal,  $R_f$ ,  $G_f$  and  $B_f$  are the bandpass filtered version of normalized  $R$ ,  $G$  and  $B$  channels respectively.

An alpha tuning technique is used to further separate  $i(t)$  and  $p(t)$  as

$$\hat{p}(t) = S_1(t) - \alpha \cdot S_2(t) \quad \text{with } \alpha = \frac{\sigma(S_1)}{\sigma(S_2)}. \quad (2)$$

Here  $\sigma(\cdot)$  indicates the standard deviation of a given signal ( $\cdot$ ). It is concluded that  $\hat{p}(t)$  is proportional to  $p(t)$  only when  $i(t)$  or  $p(t)$  dominates. However, if  $i(t)$  and  $p(t)$  have a similar magnitude,  $\hat{p}(t)$  is suboptimal [15].

Due to the drawback of alpha tuning described above, here we take a KDICA [13] algorithm to further separate the pulse signal from chrominance signals. The proposed chrominance based KDICA algorithm is abbreviated as CK, where the full algorithm is listed in Algorithm 1. Different from the fastICA used in existing rPPG papers [19, 20, 36], the CK method only separates two sources since chrominance signals are supposed to remove the motion artifacts already. More importantly, it does not need to assume the probability distribution of unknown sources, which is more practical.

### Algorithm 1 CK

- 1: Define chrominance signals  $S_1(t)$  and  $S_2(t)$  ;
- 2: Prewhiten:  $\hat{X}(t) = \hat{\Sigma}_S^{-1/2} S(t)$  for  $t = 1, 2, \dots, n$ , where  $\hat{\Sigma}_S$  is the sample variance-covariance matrix of  $S(t)$  ;
- 3: Optimize  $\hat{O} = \arg \max F(O)$  to get a rotation matrix using the gradient algorithm, where  $F(O)$  is mutual information [13] defined with kernel density estimator using Laplacian kernel function  $K(x) = \frac{1}{2} e^{-|x|}$  ;
- 4: Output the demixing matrix as :  $\bar{W} = \hat{O} \hat{\Sigma}_S^{-1/2}$  .

The spectrum of each extracted source signal is calculated with FFT, where the dominant frequency is determined as the HR candidate. We select the target source with the highest power (normalized by a total power) among all candidates.

## 4. Results and discussion

In this section, we will first take a full evaluation of the performance of the proposed semi-blind source separation method. The effect of resolution under various distances is then investigated by the proposed method together with two other conventional methods.

### 4.1. Experimental setup

#### A. The UBFC-RPPG dataset

The public dataset UBFC-RPPG [17] is taken here to verify the performance of the proposed method. The UBFC-RPPG is specifically designed for the remote HR measurement task. It contains 42 videos from 42 different subjects. The videos were recorded by a Logitech C920HD Pro camera with a resolution of 640x480 in an uncompressed 8-bit RGB format. Each subject sits in front of a camera about 1 meter away. The participant is required to play a time-sensitive mathematical game to keep their heart rate varied. The video records natural movements of subjects, including rigid and non-rigid motions.

For each video, we process the first minute and set the processing window as 10 s with a 5-seconds overlap between neighboring windows.

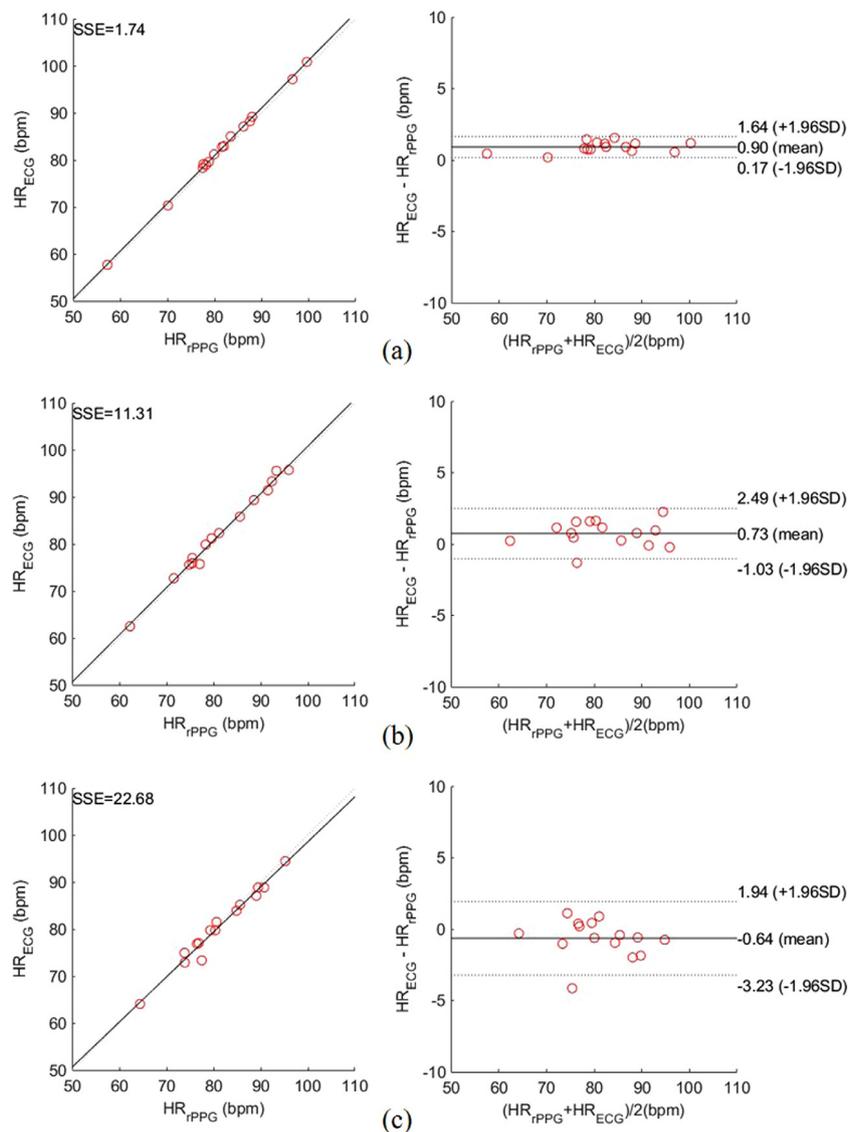


Fig. 5. The correlation and Bland–Altman plots of the CK method with a 2.7K resolution at three different distances: (a) 1.0 m; (b) 2.0 m; (c) 3.0 m, where SSE indicates the sum of squared errors ( $SSE = \sum_{i=1}^n (HR_{ECG}(i) - HR_{rPPG}(i))^2$  with  $n = 15$ ).

### B. The self-collected dataset

Considering there is a lack of public datasets for resolution–distance evaluation, we collect the videos by ourselves. The experimental setup is shown in Fig. 3. A high-resolution action camera GoPro HERO6 Black (GoPro Inc., San Mateo, California, U.S.) was employed to collect video sequences. It was mounted on a mini tripod put on a table, with a distance of 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 m, respectively away from the subject. Under each distance, videos were recorded with a 2.7K (2704x1520) resolution at 30 frames per second. The videos were recorded in an H.265 compression standard. And lower-resolution videos including 240p (320x240), 480p (640x480), 720p (1280x720), and 1080p (1920x1080) were directly downsampled from the same 2.7K video using a nearest neighbor interpolation method. It is to ensure that the comparison is against the same baseline except for the resolution. The illumination is natural sunlight from an indoor environment, which is considered to be uniform. It ensures a minimum illumination variation in recorded videos.

With the approval of the Ethics Review Committee of Hefei University of Technology, 15 voluntary Asian-skin-color subjects (two females and thirteen males), with the ages ranging from 22 to 25 years old ( $23.47 \pm 1.55$ ), participated in the experiment. During the video

recording, the subjects sit stationarily in front of the camera with their faces visible. Meanwhile, the ECG acquisition system ECG6951D (Nihon Kohden Co., Shinjuku-ku, Tokyo, Japan) was utilized to acquire the HR ground truth, which was synchronized with the recorded videos. For each subject, a total of six videos were captured with a 2.7K resolution at six different distances. Then each video was downsampled to four other lower resolutions. Consequently, a total of 450 synchronized videos and ECG recordings (for 15 subjects) were collected to build a database, with each video lasting at least 60 s. The processing window is 10 seconds, with a 5-second overlap between neighboring windows.

### 4.2. Evaluation of the performance of the proposed method

To evaluate the performance of the proposed method, we also compare it with some other rPPG algorithms. We employed the MATLAB toolbox ‘iPhys’ [37] for the implementation of five typical non-contact HR measurement methods including the GREEN [38], ICA [19], CHROM [14], POS [15] and Ballistocardiography (BCG) [39]. It is known that the input of BSS methods has a clear impact on the quality of source extraction [40]. To demonstrate the superiority of using chrominance signals instead of the conventional RGB traces in

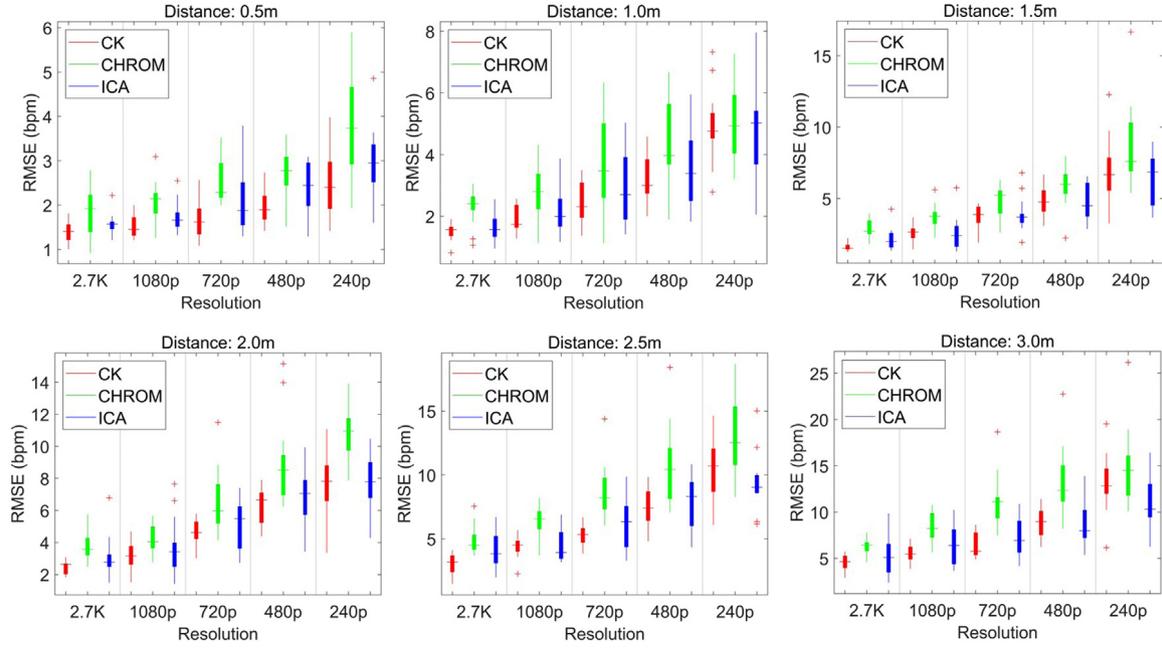


Fig. 6. The RMSE box plots of rPPG methods with different resolutions at a fixed distance.

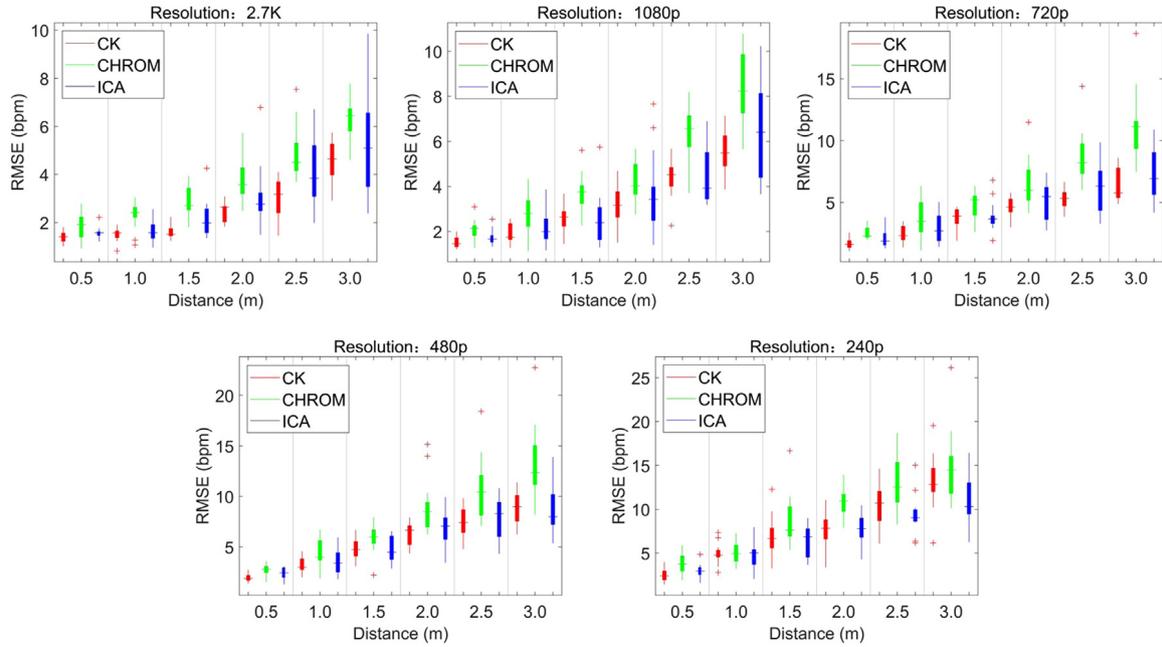


Fig. 7. The RMSE box plots of rPPG methods at different distances with a fixed resolution.

BSS methods, we also test the KDICA with these two different types of inputs. Similarly, to demonstrate the better performance of KDICA compared to ICA in extracting heart rate signals, we also compared the results of ICA with the same inputs. The ICA with chrominance inputs is abbreviated as CI corresponding to the CK for KDICA.

All those methods were tested on the public UBFC-RPPG dataset. For each algorithm, we choose the same region of interest as in Section 3.1, as shown in Fig. 4. A bandpass filter with [0.7,3] Hz was performed on all RGB channels defined in the ROI. The filtered RGB signal was fed into each algorithm for further processing. The results are shown in Table 2 below. Several quality metrics are compared including the mean absolute error (MAE,  $MAE = \frac{1}{n} \sum_{i=1}^n |HR_{pd}^{(i)} - HR_{gt}^{(i)}|$ ), the standard deviation (SD), the root mean squared error (RMSE, where  $RMSE = \frac{1}{n} \sum_{i=1}^n (HR_{pd}^{(i)} - HR_{gt}^{(i)})^2$ ), and the Pearson correlation coefficient

( $r$ ). Here  $HR_{pd}^{(i)}$  and  $HR_{gt}^{(i)}$  indicate the predicted heart rate value and the ground truth of the  $i$ th sample, respectively. The best results are highlighted in bold.

It is observed that the CK method outperforms all the other methods. The results of GREEN and BCG methods overall are worse than the other ones. This may be because both methods are susceptible to noise, especially the motion noise. The model-based methods (CHROM and POS) were slightly better than the BSS-based methods (ICA and KDICA). But if we use chrominance signals as the inputs of BSS methods, the results of the semi-BSS methods (CI and CK) are better than either the pure BSS-based or model-based methods. It verifies that the use of chrominance signals for the BSS algorithm can effectively improve the accuracy of heart rate signal extraction.

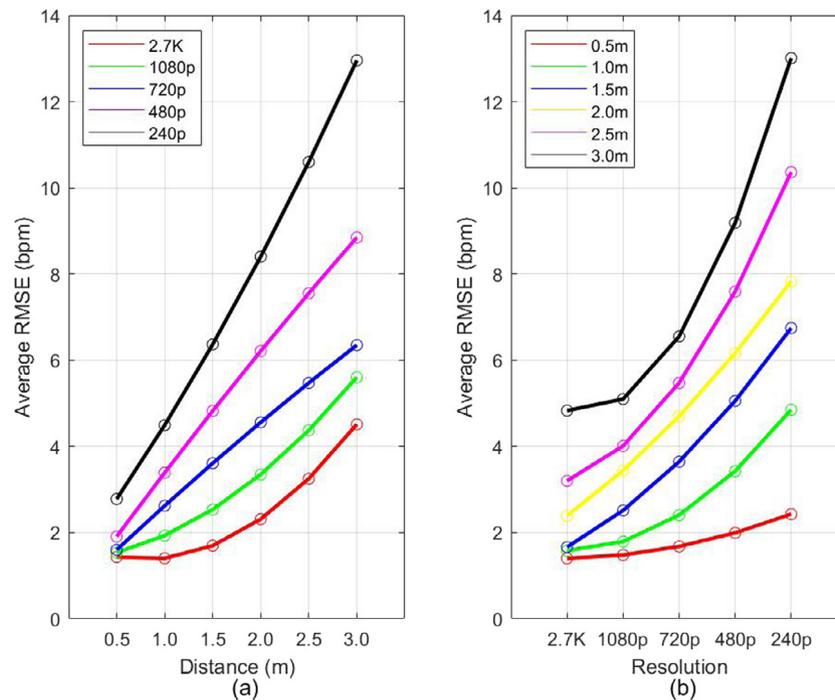


Fig. 8. The averaged RMSE curves of the CK method: (a) Varying distances with a fixed resolution; (b) Varying resolutions at a fixed distance.

Table 2

Performance of different rPPG methods with the UBFC-RPPG dataset.

Method	SD (bpm)	MAE (bpm)	RMSE (bpm)	$r$
GREEN [38]	11.090	4.469	11.598	0.842
BCG [39]	26.113	27.922	37.962	0.249
CHROM [14]	4.454	3.435	4.614	0.968
POS [15]	6.501	2.436	6.608	0.936
ICA [19]	8.258	3.507	8.635	0.908
KDICA	7.720	3.417	8.094	0.918
CI	3.723	2.519	4.074	0.979
CK	3.525	2.292	3.803	0.981

The improvement of the proposed semi-BSS method comes from two aspects. Firstly, the chrominance signals in Eq. (1) eliminate most of the motion noise through a weighted subtraction among different RGB channels. It is equivalent to take a projection on the coefficient vector of motion signals in the optical reflection model. Therefore, the generated chrominance signals are considered to be motion-robust. Secondly, the chrominance signals only have two channels. It eases the difficulty of the BSS since one source is already explicitly removed compared to the original three-channel RGB signals. Since KDICA is taken to further extract the pulse signal, it overcomes the drawbacks of alpha tuning in both CHROM and POS. The alpha tuning requires the sources to have different magnitudes. This may be easily disrupted in practice.

The performance of KDICA is slightly better than that of ICA. This has been verified with both the RGB inputs and the chrominance signal inputs. It is considered that the KDICA does not require a prior assumption on the distribution of hidden sources, which is more practical for rPPG applications than the commonly used parametric-based FastICA algorithm.

#### 4.3. Evaluation of the influences of resolutions

We take a full experimental study on the self-collected dataset to investigate the influences of resolution under different distances. To eliminate the bias from a single method, we select three methods including the CHROM, ICA, and CK for the evaluation. The experimental results are as follows. First, the performance of rPPG solutions

with 2.7K, 1080p and 720p resolutions is shown in Table 3. It is observed that the proposed CK method still gets overall the best performance. Consistent in all the three methods, the measurement accuracy decreases as distance increases for a given resolution.

The above results indicate that the rPPG measurement quality gets an apparent degradation when the distances increase. This is further verified in the correlation and Bland-Altman plots of the CK method in Fig. 5, where the videos were captured with a super-high 2.7K resolution at 1.0 m, 2.0 m, and 3.0 m, respectively. The  $HR_{rPPG}$  here indicates the averaged HR value of the predicted pulse while the  $HR_{ECG}$  represents the averaged HR value of the corresponding ECG signal. We can see that they match well with each other under all three distances. However, the sum of squared errors (SSE) grows from 1.74 to 22.68 when distances increase, which indicates a signal quality degradation. The reason behind it should be the remarkable camera quantization noise effect at a considerable distance.

The box plots indicating RMSE of all solutions are illustrated in Fig. 6. It shows that the RMSE increases while video resolution decreases for each fixed shooting distance. This is consistently observed in all the three rPPG methods. It verifies the benefit of using a super-high resolution even at a commonly used near shooting distance such as 1.0 m.

Some other RMSE box plots are shown in Fig. 7, where the resolution is fixed while the distance changes. In detail, with a given resolution at 240p, the median RMSE of the CK method varies from 2.4 to 12.8 bpm along with distances. However, for the super-high 2.7k resolution, the RMSE only grows from 1.4 to 4.6 bpm. It shows that super-high resolution is a better choice if the computational cost is not a consideration.

Finally, the averaged RMSE of all CK solutions is summarized in Fig. 8. It is observed in Fig. 8(a) that the curve for a super-high 2.7K resolution is much flatter compared to other resolutions. Meanwhile, there is not much difference among resolutions at a 0.5 m distance in Fig. 8(b). However, the curves quickly become steep if distances are longer than 1 m. It once again proves the benefit of using high resolution in rPPG.

**Table 3**  
Performance of three rPPG methods with different distance–resolution configurations.

Distance (m)	Resolution	Method	MAE (bpm)	SD (bpm)	RMSE (bpm)	r
1.0	2.7K	ICA	1.316	1.410	1.627	0.991
		CHROM	1.816	2.245	2.313	0.975
		CK	<b>1.260</b>	<b>1.234</b>	<b>1.503</b>	<b>0.993</b>
2.0	2.7K	ICA	2.131	3.225	3.037	0.944
		CHROM	2.892	3.702	3.729	0.931
		CK	<b>1.838</b>	<b>2.406</b>	<b>2.475</b>	<b>0.970</b>
3.0	2.7K	ICA	5.197	5.530	1.627	0.809
		CHROM	5.024	6.233	6.361	0.766
		CK	<b>3.420</b>	<b>4.634</b>	<b>4.595</b>	<b>0.863</b>
1.0	1080p	ICA	<b>1.557</b>	2.162	2.123	0.978
		CHROM	2.123	2.810	2.767	0.961
		CK	1.564	<b>1.808</b>	<b>1.922</b>	<b>0.984</b>
2.0	1080p	ICA	<b>2.501</b>	4.038	3.649	0.917
		CHROM	3.348	4.293	4.239	0.914
		CK	2.603	<b>3.373</b>	<b>3.273</b>	<b>0.940</b>
3.0	1080p	ICA	4.517	6.654	6.480	0.704
		CHROM	6.670	8.430	8.461	0.770
		CK	<b>4.259</b>	<b>5.630</b>	<b>5.616</b>	<b>0.805</b>
1.0	720p	ICA	2.067	3.149	2.955	0.953
		CHROM	2.983	3.950	3.689	0.921
		CK	<b>1.910</b>	<b>2.514</b>	<b>2.466</b>	<b>0.969</b>
2.0	720p	ICA	<b>3.214</b>	5.123	5.046	0.876
		CHROM	5.105	6.770	6.539	0.774
		CK	3.569	<b>4.729</b>	<b>4.659</b>	<b>0.897</b>
3.0	720p	ICA	<b>4.602</b>	7.226	7.172	0.715
		CHROM	9.112	11.544	11.197	0.488
		CK	5.091	<b>6.181</b>	<b>6.406</b>	<b>0.766</b>

#### 4.4. Discussion

The current study shows that the resolution has a remarkable influence on rPPG solutions, especially when the camera-subject distance is over 1 meter. The use of a super-high resolution such as 2.7K can improve the signal quality definitely. Therefore, although the rPPG signal can be extracted from low-resolution videos, it is better to use high resolution if computational and storage resources are affordable. Alternatively, choosing an appropriate resolution needs always to be considered with the shooting distance and the desired accuracy of the rPPG pulse.

On the other hand, existing rPPG methods usually only take a pixel averaging on video frames. Then the generated time series is denoised in the temporal domain. This may not be adequate in rPPG signal extraction. As proved in the works of McDuff [34] and Fukunishi [41], the attempts of spatial signal enhancement or denoising also demonstrate potentials to improve the quality of rPPG pulse signal. Therefore, it is worth to taking further investigations on the spatial domain in addition to existing temporal methods.

Finally, the proposed method is proved to get overall the best performance compared to some other related methods. This inspires further explorations to find new ways of combining BSS-based and model-based methods. The merit lies in enforcing physical principles to simplify or restrict the source extraction process. Consequently, it can lead to some new approach which combines the advantage of existing ones.

#### 5. Conclusion

In this paper, we have investigated the influence of video resolutions on the quality of rPPG solutions under various distances. A newly proposed rPPG method has been tested comprehensively. The new method is considered to combine the advantages of blind source separation and model-based methods. A dataset has been prepared which includes synchronized videos and ECG measurements with various resolution and distance configurations. The experimental results verify the best performance of the proposed method. It shows that a high resolution,

particularly the super-high 2.7K resolution used for the first time to our best knowledge, improves the signal quality remarkably when the shooting distance is no less than 1 meter. Consequently, a careful selection of resolution is necessary to achieve a desired accuracy of rPPG measurement under some given shooting distance. Furthermore, this study also gives some inspirations on the spatial-based denoising study for future rPPG research.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants 61922075, 81571760 and 41901350), the Fundamental Research Funds for the Central Universities (Grants JZ2019HGTA0049 and JZ2019HGBZ0151), and the Young Elite Scientists Sponsorship Program by CAST (Grant 2017QNRC001).

#### References

- [1] Y. Sun, N. Thakor, Photoplethysmography revisited: From contact to noncontact, from point to imaging, *IEEE Trans. Biomed. Eng.* 63 (3) (2016) 463–477.
- [2] A. Singh, L. Sharma, S. Dandapat, Multi-channel ECG data compression using compressed sensing in eigenspace, *Comput. Biol. Med.* 73 (2016) 24–37.
- [3] F. Fan, Y. Yan, Y. Tang, H. Zhang, A motion-tolerant approach for monitoring SpO<sub>2</sub> and heart rate using photoplethysmography signal with dual frame length processing and multi-classifier fusion, *Comput. Biol. Med.* 91 (2017) 291–305.
- [4] L.P. Malasinghe, N. Ramzan, K. Dahal, Remote patient monitoring: a comprehensive study, *J. Ambient Intell. Human. Comput.* 10 (1) (2017) 57–76.
- [5] W. Wang, S. Stuijk, G. de Haan, Unsupervised subject detection via remote PPG, *IEEE Trans. Biomed. Eng.* 62 (11) (2015) 2629–2637.
- [6] E.M. Nowara, T.K. Marks, H. Mansour, A. Veeraghavany, SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near-infrared, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2018.

- [7] F. Peng, Z. Zhang, X. Gou, H. Liu, W. Wang, Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter, *Biomed. Eng. Online* 13 (1) (2014) 50.
- [8] J. Cheng, X. Chen, L. Xu, Z.J. Wang, Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition, *IEEE J. Biomed. Health Inf.* 21 (5) (2016) 1422–1433.
- [9] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, Z.J. Wang, Video-based heart rate measurement: Recent advances and future prospects, *IEEE Trans. Instrum. Meas.* (2018) 1–16., <http://dx.doi.org/10.1109/tim.2018.2879706>.
- [10] D.J. McDuff, E.B. Blackford, J.R. Estep, The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography, in: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition, FG 2017, 2017, pp. 63–70.
- [11] C. Zhao, W. Chen, C.-L. Lin, X. Wu, Physiological signal preserving video compression for remote photoplethysmography, *IEEE Sensors J.* 19 (12) (2019) 4537–4548.
- [12] E.B. Blackford, J.R. Estep, A.M. Piasecki, M.A. Bowers, S.L. Klosterman, Long-range non-contact imaging photoplethysmography: cardiac pulse wave sensing at a distance, in: *Optical Diagnostics and Sensing XVI: Toward Point-of-Care Diagnostics*, vol. 9715, 2016, p. 971512.
- [13] A. Chen, Fast kernel density independent component analysis, in: *International Conference on Independent Component Analysis and Signal Separation*, 2006, pp. 24–31.
- [14] G. De Haan, V. Jeanne, Robust pulse rate from chrominance-based rPPG, *IEEE Trans. Biomed. Eng.* 60 (10) (2013) 2878–2886.
- [15] W. Wang, A.C. den Brinker, S. Stuijk, G. de Haan, Algorithmic principles of remote PPG, *IEEE Trans. Biomed. Eng.* 64 (7) (2017) 1479–1491.
- [16] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, *Neural Netw.* 13 (4–5) (2000) 411–430.
- [17] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, Unsupervised skin tissue segmentation for remote photoplethysmography, *Pattern Recognit. Lett.* 124 (2017) 82–90.
- [18] S. Zauneder, A. Trumpp, D. Wedekind, H. Malberg, Cardiovascular assessment by imaging photoplethysmography – a review, *Biomed. Eng. / Biomed. Techn.* 63 (5) (2018) 617–634.
- [19] M.-Z. Poh, D.J. McDuff, R.W. Picard, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation, *Opt. Express* 18 (10) (2010) 10762.
- [20] A. Lam, Y. Kuno, Robust heart rate measurement from video using select random patches, in: *Proceedings of the IEEE International Conference on Computer Vision, ICCV*, 2015, pp. 3640–3648.
- [21] B. Wei, X. He, C. Zhang, X. Wu, Non-contact, synchronous dynamic measurement of respiratory rate and heart rate based on dual sensitive regions, *Biomed. Eng. Online* 16 (1) (2017) 17.
- [22] H. Qi, Z. Guo, X. Chen, Z. Shen, Z.J. Wang, Video-based human heart rate measurement using joint blind source separation, *Biomed. Signal Process. Control* 31 (2017) 309–320.
- [23] A. Al-Naji, A.G. Perera, J. Chahl, Remote monitoring of cardiorespiratory signals from a hovering unmanned aerial vehicle, *Biomed. Eng. Online* 16 (1) (2017) 101.
- [24] G. De Haan, A. Van Leest, Improved motion robustness of remote-PPG by using the blood volume pulse signature, *Physiol. Meas.* 35 (9) (2014) 1913.
- [25] W. Chen, D. McDuff, Deepphys: Video-based physiological measurement using convolutional attention networks, in: *Proceedings of the European Conference on Computer Vision, ECCV*, 2018, pp. 349–365.
- [26] X. Niu, H. Han, S. Shan, X. Chen, Synrhythm: Learning a deep heart rate estimator from general to specific, 2018 24th International Conference on Pattern Recognition, ICPR, 2018, pp. 3580–3585.
- [27] Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, A. El Saddik, EVM-CNN: Real-time contactless heart rate estimation from facial video, *IEEE Trans. Multimed.* 21 (7) (2018) 1778–1787.
- [28] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W.T. Freeman, Eulerian video magnification for revealing subtle changes in the world, *ACM Trans. Graph.* 31 (4) (2012) 1–8, (Proc. SIGGRAPH 2012).
- [29] Y. Sun, V. Azorin-Peris, R. Kalawsky, S. Hu, C. Papin, S.E. Greenwald, Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam, *J. Biomed. Opt.* 17 (3) (2012) 037005.
- [30] B. Han, K. Ivanov, L. Wang, Y. Yan, Exploration of the optimal skin-camera distance for facial photoplethysmographic imaging measurement using cameras of different types, in: *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare*, 2015, pp. 186–189.
- [31] E.B. Blackford, J.R. Estep, Effects of frame rate and image resolution on pulse rate measured using multiple camera imaging photoplethysmography, in: *Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 9417, 2015, p. 94172D.
- [32] E.B. Blackford, J.R. Estep, Measurements of pulse rate using long-range imaging photoplethysmography and sunlight illumination outdoors, in: *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics*, vol. 10072, 2017, p. 100720S.
- [33] N. Ibrahim, R. Tomari, W.N.W. Zakaria, N. Othman, Analysis of non-invasive video based heart rate monitoring system obtained from various distances and different facial spot, in: *J. Phys.: Conf. Ser.*, 1049 (1) (2018) 012003.
- [34] D. McDuff, Deep super resolution for recovering physiological information from videos, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2018, pp. 1367–1374.
- [35] Z. Zhang, P. Luo, C.C. Loy, X. Tang, Facial Landmark detection by deep multi-task learning, in: *Proceedings of the European Conference on Computer Vision, ECCV*, 2014, pp. 94–108.
- [36] K.M. van der Kooij, M. Naber, An open-source remote heart rate imaging method with practical apparatus and algorithms, *Behav. Res. Methods* (2019) 1–14.
- [37] D. McDuff, E. Blackford, Iphys: An open non-contact imaging-based physiological measurement toolbox, 2019, pp. 1–4, arXiv preprint [arXiv:1901.04366](https://arxiv.org/abs/1901.04366).
- [38] W. Verkrusye, L.O. Svaasand, J.S. Nelson, Remote plethysmographic imaging using ambient light, *Opt. Express* 16 (26) (2008) 21434–21445.
- [39] G. Balakrishnan, F. Durand, J. Guttag, Detecting pulse from head motions in video, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2013, pp. 3430–3437.
- [40] D. Wedekind, A. Trumpp, F. Gaetjen, S. Rasche, K. Matschke, H. Malberg, S. Zauneder, Assessment of blind source separation techniques for video-based cardiac pulse extraction, *J. Biomed. Opt.* 22 (3) (2017) 035002.
- [41] M. Fukunishi, K. Kurita, S. Yamamoto, N. Tsumura, Video based measurement of heart rate and heart rate variability spectrogram from estimated hemoglobin information, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2018.