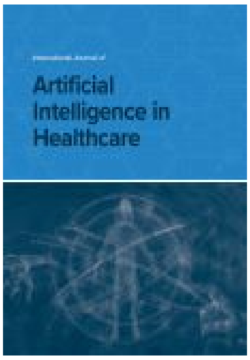


International Journal of Artificial Intelligence in Healthcare

ISSN online: 3050-2470 - ISSN print: 3050-2462
https://www.inderscience.com/ijaih

Robust heart rate estimation during intensive physical training using AI-enhanced particle filter

Kokila Bharti Jaiswal

DOI: 10.1504/IJAIH.2025.10073501

Article History:

Received:	08 April 2025
Last revised:	15 July 2025
Accepted:	10 August 2025
Published online:	20 October 2025

Robust heart rate estimation during intensive physical training using AI-enhanced particle filter

Kokila Bharti Jaiswal

Department of ECE,
Bhilai Institute of Technology,
Durg, 491001, India
Email: kokila.bharti@bitdurg.ac.in

Abstract: The remote photoplethysmography (rPPG) signal provides an essential data for the estimation of heart rate (HR). However, rPPG signal is often corrupted by noise due to motion, driving the conventional denoising techniques to failure. Perhaps, the motion artefact is of more concern when it is falsely captured as a real pulse signal. In this article, a novel methodology is proposed leveraging particle filters (PF) for the robust HR measurement in the presence of motion artefact. The proposed method improves the measurement accuracy of HR specifically for the cases of physical exercise, when the subject is severely corrupted by motion artefacts. Proposed approach yields better results in terms of estimation of HR on benchmark datasets such as UBFC-rPPG and PURE. Moreover, the proposed method is a stand-alone technique, and can be easily associated with the existing algorithms.

Keywords: remote health monitoring; heart rate; remote photoplethysmography; rPPG; particle filter; denoising; motion artefact.

Reference to this paper should be made as follows: Jaiswal, K.B. (2025) 'Robust heart rate estimation during intensive physical training using AI-enhanced particle filter', *Int. J. Artificial Intelligence in Healthcare*, Vol. 1, No. 1, pp.92–108.

Biographical notes: Kokila Bharti Jaiswal received her BE in Electronics and Telecommunication from CSVTU University, India, MS (by Research) in VLSI System from National Institute of Technology, Tiruchirappalli, India, and PhD from NIT Raipur. Her research interests include computer vision, machine learning and pattern recognition.

1 Introduction

It is evident from the epidemiological studies that exercise boost the immune system and has great benefits in cardiovascular system. At the same time it becomes hazardous and leads to morbidity if not monitored properly. Heart rate (HR) monitoring is an important phenomena in rehabilitation centres and for a person performing intensive exercise (Nauman et al., 2011). Present market is flooded with numerous devices to fulfill the purpose. Available devices can be worn at different body parts such as chest, abdomen and wrist. Wearable devices need to snugly fitted to the body part, so as to get the accurate measurements.This exhibits resistance to the person exercising. Compared to all other devices wrist worn devices may imparts little discomfort, but remote photoplethysmography (rPPG) techniques shows its superiority needing no contact at all. The basic phenomena of rPPG is that, the subtle variation in reflection of light through the skin can be captured by a camera. This subtle change occurs due to change in blood volume underneath skin in synchronisation with the heart beat.

Feasibility of the HR extraction from video was first demonstrated by Verkruysse et al. (2008). Since then many advancements have been done limited to strict movements and controlled environmental conditions. Large body movements may cause some loss of information in tracking of subjects face which may hampers the measurement of HR due to motion contamination. The motion artefact (MA) generally arises due to natural movements of subjects face while monitoring through camera. In addition, for real-time measurement, the subject are allowed to freely operate daily routine while monitoring of HR. Thus if the HR measurement is corrupted with noise due to motion then other clinical analysis relying on HR measurement would be affected. In recent years several methods have been proposed to improve the robustness of rPPG in realistic conditions. Blind source separation (BSS)-based methods such as PCA (Poh et al., 2010a) and ICA (Poh et al., 2010b) decomposes the signal in linear fashion, and the signal associated with the heart beat can be extracted. Signal filtering methods such as bandpass filtering (Poh et al., 2010a), adaptive filtering (Li et al., 2014) and homomorphic filtering (Liu et al., 2020) isolates the pulse signal from the mixture of noisy signals. Various denoising methods proposed earlier extracts the PPG signal first from conventional methods than denoising is performed using filters.

The existing filters for denoising used in the state-of -the-art methods suffers from several limitations. LMS filters due to dependence on input size parameters suffers from slow convergence (Li et al., 2014). DWT is also useful in denoising in Das et al. (2022) but it is always uncertain to choose appropriate subbands, as the frequency of noise is usually unknown. Bounded Kalman Filter approach used in Prakash and Tucker (2018) is restricted to be used for Gaussian noise. Due to the low SNR of HR signals, conventional filtering approaches do not give satisfactory results. Particle filter (PF) (Gustafsson et al., 2002) is suitable for HR estimation due to its ability to handle time series data and handle limited variations within a small range. PF is a popular choice in tracking and estimation problems, especially when dealing with nonlinear and non-Gaussian systems. It work by representing the underlying probability distribution of the system with a set of randomly generated particles and then updating the particles based on the observed data. This allows PFs to accurately track and estimate the HR in real-time, even in the presence of measurement noise and other uncertainties.

In this paper, a new framework particle filter-based heart rate estimation (PAHRE) is proposed. The proposed model utilises the PF to estimate the HR under the influence of strong motion.

The contribution of the proposed method are as follows:

- A novel method is proposed for estimation of HR from signal contaminated with wide variety of MAs specifically in the case of physical exercise.
- The proposed method utilises the Bayesian approach to rectify false HR estimates.

- The method uses the signal acquired by tracking nose movement to update the weight of particles in a PF, which helps to reduce the impact of measurement noise caused by MAs.

This article is structured as follows. Previous works related to the MA removal is discussed in Section 2. The design of motion robust HR estimation method using PF is presented in Section 3. Experimental results are discussed in Section 4. Finally discussion on work is presented in Section 5 followed by conclusion in Section 6.

2 Previous works done for MA removal from rPPG signals

Traditional HR estimation using PPG signals use BSS methods. In these demixing of signal is done without any prior knowledge of mixing criteria. These methods use PCA and ICA for separating signals. However they have assumed the signals to be statistically independent and non-Gaussian. The limitation of these approach is the requirement of long duration signals. Poh et al. (2010a) uses traditional Viola Jones algorithm for face detection. After averaging all the channels of R, G and B independent component analysis (ICA) is used to decompose the signal into various frequency components and finally the component with highest frequency is selected as HR. Results showcased by the above methods gives more accurate HR estimation than using Green channel alone as proposed by Verkruysse et al. (2008). But the decomposed signal is highly affected due to motion.

Lewandowska et al. (2011) employs principal component analysis (PCA), resulting in much faster detection of HR. However, the subject is constrained to be in the motionless state.

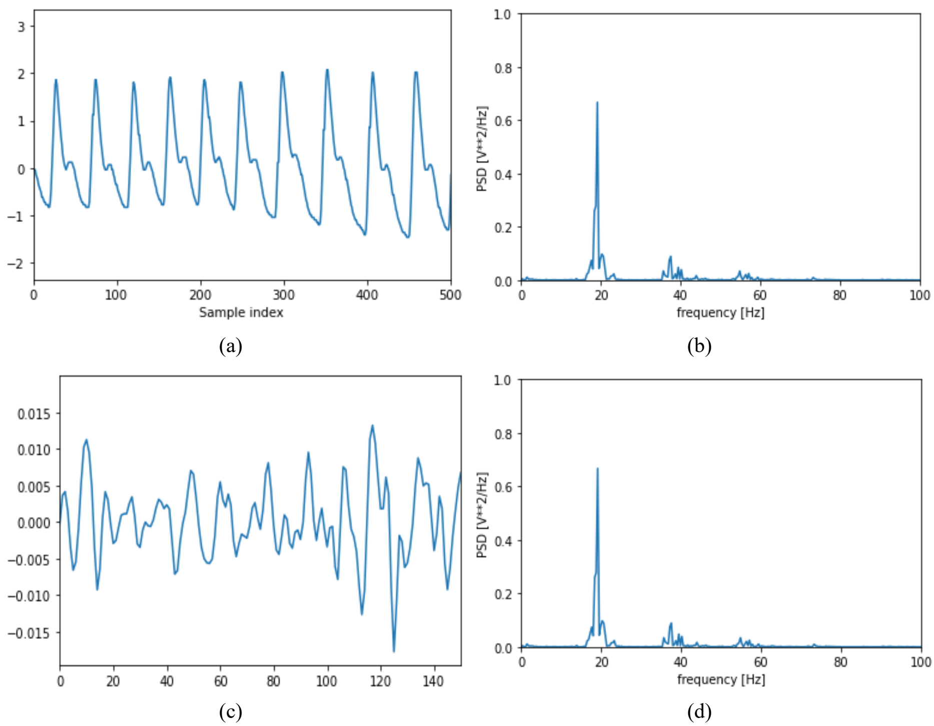
To measure the HR in a non-contact manner, facial region such as forehead and cheeks, called as ROI needs to be tracked. Eventually, the rPPG signals are obtained by applying different filters to the average pixels values of this ROIs. But in real life case, the significant natural movements of face makes it difficult to get the correct HR measurement from rPPG signal. For such cases motion robust methods designed to counters the disruption caused due to motion, posses a significant reason to study. Model-based methods like CHROM and POS were introduced by De Haan and Jeanne (2013) and Wang et al. (2016). These methods utilises the optical properties of skin and assumes standardised skin-tone. The RGB is divided into three components, which makes the model more robust to motion.

For tracking the face, we need to first detect the face of a subject, which is majorly done by Viola-Jones face detector in all the state-of-the-art methods (Asthana et al., 2013; Fiaz et al., 2019; Lam and Kuno, 2015). Once the face is detected, the most commonly used object tracking method such as Kanade-Lucas-Tomasi (KLT) (Bourel et al., 2000) is used to track the desired face region. KLT is based on optical flow of good features between two subsequent frames, results in the fast face tracking. However the KLT fails to track the object when there is a large motion and also when the face is partially occluded. To improve this a discriminative correlation filtering (DCF) (Xu et al., 2019) and deep trackerFiaz et al. (2019) is proposed which provides better accuracy at the cost of computational overhead.

Loosing a track of facial region is highly possible for large facial movements. Researchers have developed many robust tracking mechanism. Li et al. (2014)

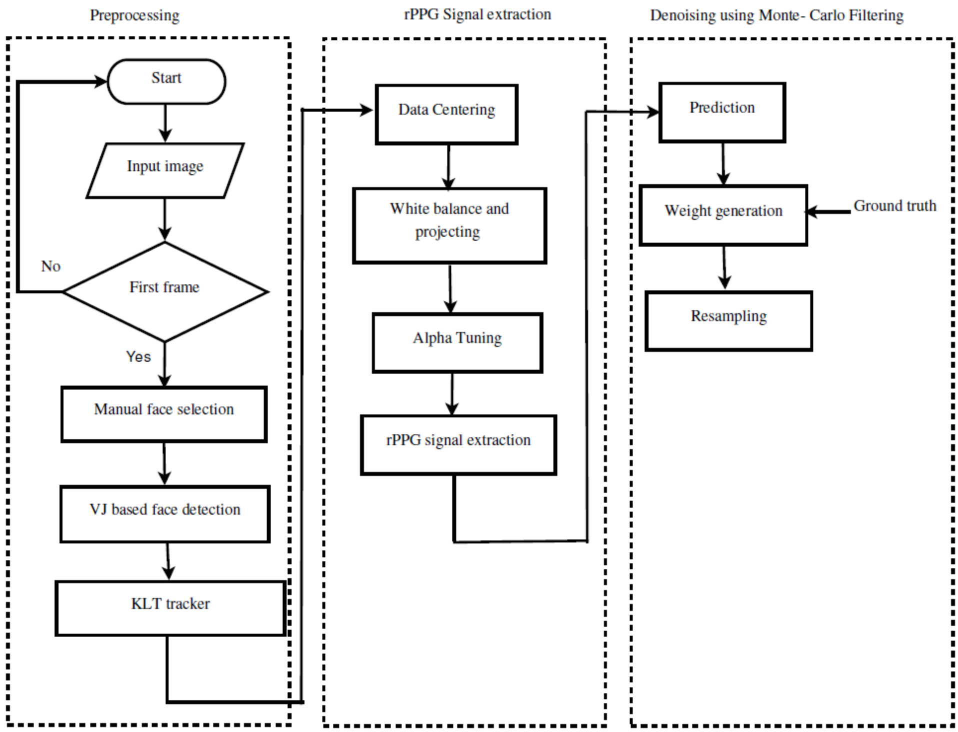
incorporated normalised least mean square (NLMS) adaptive filter to counter the effect of slight movements in the subjects's face. Galli et al. (2017) uses the landmark position of the face, then this landmark is tracked. PFLD method is proposed to track landmark position, but in condition of large facial movement and occlusion, even this method failed. Kumar et al. (2015) utilises signals from different regions of skin. Each region has given different weights for HR estimation. Prakash and Tucker (2018) used a bounded Kalman filter to estimate and track motion signals. To handle the blurring effect occured due to uncontrolled head movements, blur identification and removal is done in each frame, but that posses some limitations too.

Figure 1 Spectral analysis of rPPG signals with and without MA, (a) rPPG without motion (b) periodogram of rPPG without motion (c) rPPG with motion (d) periodogram of rPPG with motion (see online version for colours)



Tremendous performance growth of deep learning techniques in all areas, intrigued researchers to use those techniques for rPPG also. Researchers have started using deep learning techniques in rPPG to get the more robust measurements (Qiu et al., 2018; Niu et al., 2019; Lokendra and Puneet, 2022; Spetlik et al., 2018; Yu et al., 2020). Deephys (Chen and McDuff, 2018) is the intrusion of deep learning for HR measurement. This is the first end-to-end method for HR estimation using convolutional neural networks. Attention mechanism which takes the motion features as input guides the model for correct HR measurement. Niu et al. (2018) uses the spatiotemporal representation of input image to transform image into different representation, more informative and learnable by CNN network to detect HR. However, the above methods do not perform well under unconstrained scenarios (condition when the person is engaged in physical activity).

Figure 2 Flow diagram of Monte Carlo simulation



3 Methodology

The proposed method PAHRE detects the HR using two signals, one from cheeks and another from motion signal of nose movements. The two signals are independently extracted from the captured facial video. PAHRE consists of three important modules:

3.1 ROI detection and tracking

ROI detection is done using classical face detection algorithm, openCV Haar classifier based on Viola Jones algorithm. The height and width of the face is fixed in the ratio of 60:40. After face detection a region is curated consisting of skin region between eyes and mouth. Exclusion of eyes and mouth is done to get the ROI free from any natural movements such as blink or lip movements. Moreover, the eyes and mouth region has no significant contribution to rppg signal, hence its absence does not results in a loss. This patching strategy reduces the number of pixels to be processed which drastically reduces the computational complexity while enhancing the SNR. ROI is tracked using famous KLT algorithm (Bourel et al., 2000). RGB signals from the selected ROI is averaged and used for further processing.

Once the ROI is detected in each frame, mean value of all the pixels, belongs to each channel is calculated. For noisy rPPG signal, CHROM method (De Haan and Jeanne, 2013) is implemented by projection of R, G, B channels in different subspace as follows:

$$X_n = 3\mathcal{E}[R_n]-2\mathcal{E}[G_n] \tag{1}$$

$$Y_n = 1.5\mathcal{E}[R_n] + \mathcal{E}[G_n]-1.5\mathcal{E}[B_n] \tag{2}$$

where $\mathcal{E}[.]$ denotes the mean operator and n = 1, 2, 3, ..., represents frame numbers; X_n and Y_n represent two orthogonal signals. The original chrominance signal, denoted as C_n is the subtraction of two orthogonal signals and illustrated as follows:

$$C_n = X_n-Y_n \tag{3}$$

For further reduction of noise, a finite impulse response (FIR) bandpass filter (BPF) with hamming window is applied on the original signal C_n. The clinical measurement range of HR lies between 30 to 240 beats per minute. Therefore the cutoff frequencies of the band pass filter is set to be 0.50 Hz and 4.04 Hz.

Secondly, the trajectory of nose provides motion signal. To select the nose rectangle we have used Viola Jones nose detector. In the first frame, central point of the nose m_1 is calculated, and Euclidean distance is calculated between the central points of first frame and subsequent frames $m_1 = \Delta m_1 + \Delta m_2 + \Delta m_3 + ... + \Delta m_n$. These Euclidean distance values formed the motion signal.

The waveform of motion signal and trajectory of nose is different but their frequency spectrum will be same. We have calculated periodogram of the PPG signal and motion signal obtained from tracking nose. As we can see from Figure 1 the peaks due to motion is overlapping with the actual HR peaks, hence the periodogram is unable to provide distinction between real pulse and fake pulse. This fails the traditional methods, which depends on the spectral analysis of signal for larger motion associated with physical training. For accurate tracking of pulse signal in the presence of large uncontrolled motion we need a robust tracking system discussed further.

3.2 Stage-2 PF

Let us represent the HR value at time t by i^{th} particle x_i^t. the range of x is 60–240 BPM, which represents a normal HR of a healthy person during exercise. The PF algorithm is explained in following steps:

3.2.1 Monte Carlo filter

The Monte Carlo Simulation process (Gustafsson et al., 2002) is designed to deal with non-Gaussian and nonlinear noises. Any nonlinear and non-Gaussian time series data can be achieved using the following state space model

$$\begin{aligned} p_k &= f(p_{k-1}, v_k) \\ q_k &= h(p_k, w_k) \end{aligned} \tag{4}$$

where $f(.)$ and $h(.)$ are two functions named as state transition and state observation respectively. v_k and w_k are called system noise and observation noise, respectively. The state estimation function can be converted to probability density function represented as shown in equation. The main challenge with state estimation problem is that we need to estimate the p using observations of q $(Q_1, Q_2, Q_3, ..., Q_k)$. The state estimation problem is classified into three cases, based on observation states. The stages are as follows:

3.2.1.1 Initialisation

Initial distribution of the particles is determined. In most of the database, during inital few seconds the HR values are relatively low. This distribution is uniform and lies between range 60 bpm and 170 bpm. The algorithm for generating particles is described as follows:

1 We generate particles $X_0 = x_0(i)$ where $i = 1, ..., N$ from $p(x_0)$.

2 Calculate the weights of particle using

$$W(t)^i = p(y_t | x^i_{t|t-1}) \tag{5}$$

where $t = 1, ..., T$.

3 Normalise weight

$$W_n(t) = W(t)^i \Big/ \sum_{i=1}^{N} W(t)^i \tag{6}$$

3.2.1.2 Prediction

Probability density function is computed for the current time using all the past observations using equation

$$\mathcal{P}(p_k | Q_{1:k-1}) = \int \mathcal{P}(p_k | p_{k-1})(\mathcal{P}(p_{k-1} | Q_{1:k-1}))dp_{k-1}$$

HR aware prediction model is designed using the dataset of UBFC-rPPG values. A histogram of the dataset is plotted, as shown in Figure 3, it is quite evident from the figure that distribution forms a bell shaped curve and the change in bpm observed in the window of 5 s. As we can see from figure that the range of bpm is mostly ± 4 bpm. these observations proves that the PF designed can be used to track any change occurs in the 4 s window. The bell shaped spread of the HR values enables the PF to track instantaneous HR values and also HR values that falls beyond the range also. Better spread of the particles benefits the calculation of correct HR in few steps, even when the previous measurement is largely deviated from true value.

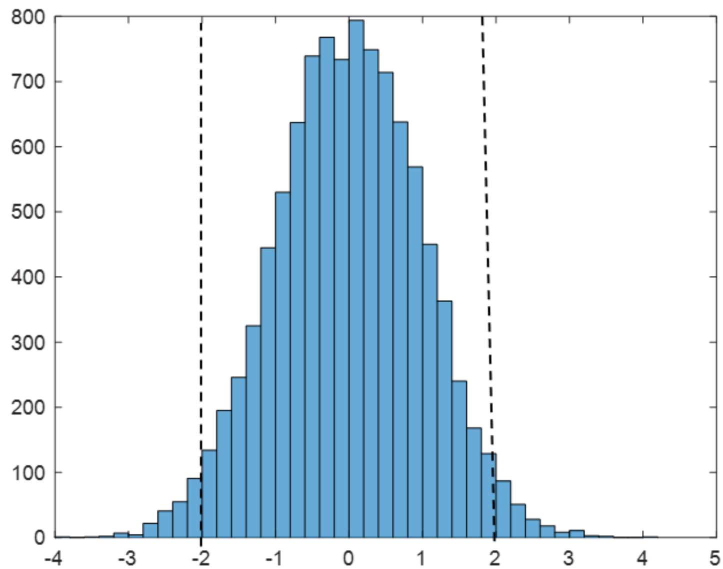
3.2.1.3 Smoothning/weight calculation

To calculate the weights of the particle we use periodgram of rPPG signal obtained from CHROM and periodgram of motion signal derived using trajectory of nose. The unwanted peaks obtained in the periodgram of PPG signal due to motion can be suppressed by using bandpass filtering. The frequency range for bandpass filtering can be obtained by motion signal. But this process is unable to completely eliminate the MA in cases, where the frequency of noise due to motion overlaps the HR frequency. Weight is calculated using equation:

$$W(t)^i = p(y_t|x^i_{t|t-1}) \tag{7}$$

Weight normalisation is done to ensure that the sum of all weights sums up to 1.

Figure 3 HR differences of UBFC-rPPG dataset (see online version for colours)



3.2.1.4 HR estimation and update

The PF output and the periodgram are used to determine the most likely BPM in this stage. The estimation of the BPM's likelihood of being the actual HR is made by

$$p(\beta) = \sum_{x_{t|t-1}=\beta} w_t(x_{t|t-1}) \tag{8}$$

$\beta_{\max}$ is the maximum point and is likely to be the value of HR. In cases, when difference of the estimated HR through PF and the peak position of the periodgram of true HR, i.e., x_n is less than threshold 't', then x_n is considered to be the HR value.

4 Experimental setup

The proposed algorithm is tested on two publicly available databases PURE (Stricker et al., 2014) and UBFC-rPPG (Macwan et al., 2019) dataset. The databases are chosen such that the criteria of application of the proposed method is feasible. In order to test the proposed algorithm the database must contain ground truth signal. In case of unavailability of ground truth signal, first five seconds PPG signal can be considered as ground truth. As this initial phase of data acquisition does not involve any voluntary movements of a subject.

4.1 Datasets used

- *PURE (Stricker et al., 2014):* A benchmark video dataset, recruited ten healthy subjects to acquire total 59 videos. The subjects are allowed to perform various types of head motions, they are categorised as

  1. no movement
  2. talking
  3. slow speed movement in parallel to the camera
  4. fast speed parallel to the camera
  5. small rotation
  6. medium rotation.

  The videos were captured with a camera by at a frame rate of 30 Hz and resolution of 640 × 480 pixels. Ground truth data is collected using pulse oximeter having a sampling rate of 60 Hz.

- *UBFC-rPPG (Macwan et al., 2019):* This dataset was created using a simple low-cost camera with 30 fps and a resolution of 640 × 480 in eight-bit RGB format. Ground truth is collected using pulse oximeter consists of PPG waveforms as well as PPG HRs. All the recordings are done in indoor environment with changing amount of sunlight.

Table 1 Comparision result of proposed method with state-of-the-art method on PURE dataset

	MAE						*RMSE*					
	Steady	*Talking*	*ST*	*FT*	*SR*	*MR*	*Steady*	*Talking*	*ST*	*FT*	*SR*	*MR*
CHROM (De Haan and Jeanne, 2013)	1.17	4.35	1.17	3.78	2.66	7.69	2.8	8.5	2.2	5.7	5.3	11.09
Our CHROM	1.16	4.03	0.8	2.67	2.5	3.5	2.8	7.5	1.2	1.9	1.01	9.2
PBV (De Haan and Van Leest, 2014)	1.20	3.4	0.6	1.4	0.6	1.18	3.03	7.2	1.2	2.73	0.89	2.11
Our PBV	1.15	3.6	0.6	0.9	0.6	0.8	2.9	7.9	1.14	1.5	0.85	0.8
POS (Wang et al., 2016)	1.17	2.02	0.7	2.5	0.9	2.29	2.8	4.5	1.31	4.16	1.69	5.28
Our POS	1.14	2.5	0.6	2.02	0.8	1.32	2.7	4.8	1.11	1.56	1.03	3.70
Phys-Net (Yu et al., 2019)	1.76	6.2	0.9	1.5	2.5	3.2	3.89	10.7	1.9	3.09	4.9	6.4
Our Phys-Net	1.25	5.8	0.8	0.6	1.9	1.9	3.0	10.2	1.16	5.32	2.2	3.67

Notes: ST = slow translation, FT = fast translation, SR = small rotation, and MR = medium rotation.

Table 2 Performance analysis of 10 subjects from UBFC-rPPG database

Subject	*1*	*3*	*4*	*5*	*8*	*9*	*10*
MAE	3.16	1.4	0.23	2.7	2.79	0.2	0.82
RMSE	22.12	9.8	1.61	18.9	19.53	1.4	5.77

Table 3 Comparision result of proposed method with state-of-the-art method on UBFC-rPPG dataset

	MAE	*RMSE*
CHROM (De Haan and Jeanne, 2013)	3.53	6.5
Ours-CHROM	3.2	4.3
PBV (De Haan and Van Leest, 2014)	4.6	9.39
Ours-PBV	3.9	7.4
POS (Wang et al., 2016)	3.36	6.50
Ours-POS	2.9	5.4
PhysNet (Yu et al., 2019)	0.73	1.87
Ours-PhysNet	0.6	1.56

4.2 Evaluation metrics

The commonly used evaluation metrics such as mean absolute error (MAE) and root mean square error (RMSE) are used for evaluation of performance of the proposed method.

1 MAE:

$$HR_{\text{mae}} = \frac{1}{n}\sum_{i=1}^{n} HR^i_{estimated} - HR^i_{gnd} \tag{9}$$

2 RMSE: RMSE is calculated using formula given in equation (10)

$$HR_{\text{rmse}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(HR^i_{estimated} - HR^i_{gnd})^2} \tag{10}$$

Lower RMSE values is an indication of strong correlation of data.

5 Results

The proposed algorithm is basically performing a denoising process, which can be embedded with any core rPPG signal extraction algorithm such as CHROM, PBV, POS and also with deep learning-based rPPG extraction methods like PhysNet. The comparison is done between original rPPG signal obtained from core algorithms and denoised rPPG obtained after embedding proposed algorithm. The experiment are conducted on two datasets i.e, PURE and UBFC-rPPG. PURE dataset provides a challenging environment by variations in motion, which helps to check the efficiency of the proposed method for motion robustness.

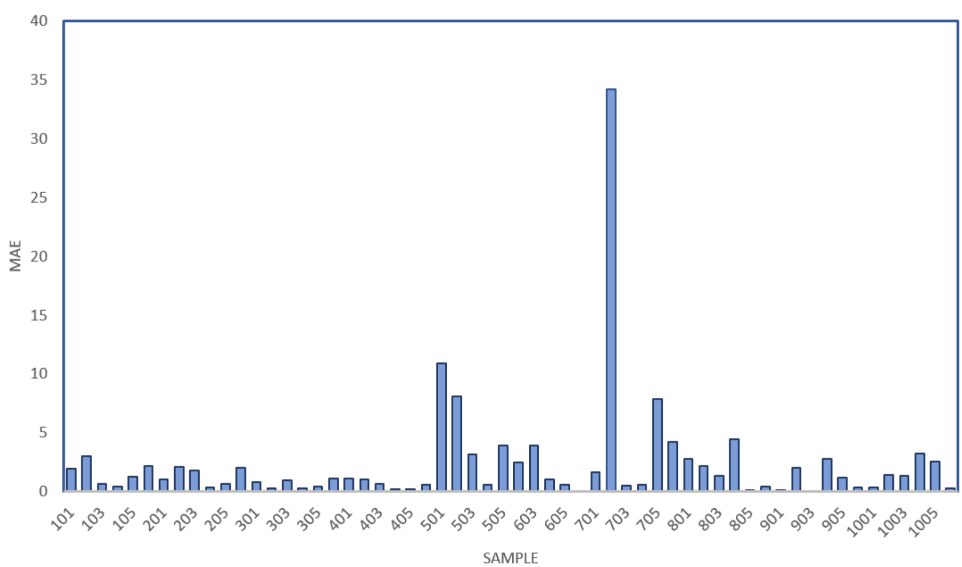
Table 4 Performance analysis of ten subjects from PURE database

Subject	*Case*	*MAE*	*RMSE*
1	Steady	4.207218	13.30439
	Talking	3.763329	11.90069
	ST	0.543268	1.717965
	FT	1.791255	5.664447
	SR	2.325384	7.353509
	MR	2.62114	8.288772
2	Steady	3.550839	11.22874
	Talking	0.596635	1.886726
	ST	1.861992	5.888137
	FT	2.019778	6.387098
	SR	0.784785	2.481707
	MR	3.955538	12.50851
3	Steady	0.084094	0.26593
	Talking	0.141372	0.447057
	ST	0.425972	1.347042
	FT	0.497418	1.572974
	SR	0.438704	1.387303
	MR	3.921608	12.40121
4	Steady	1.268748	4.012133
	Talking	0.734054	2.321283
	ST	2.083063	6.587223
	FT	1.442243	4.560771
	SR	1.989507	6.291373
	MR	1.05473	3.335348
5	Steady	4.555777	14.40663
	Talking	4.011605	12.68581
	ST	0.278721	0.881393
	FT	1.808068	5.717614
	SR	1.201539	3.7996
	MR	4.361422	13.79203
6	Steady	1.745322	5.519193
	ST	6.031565	19.07348
	FT	1.525263	4.823304
	SR	3.26859	10.33619
	MR	0.058266	0.184254
7	Steady	1.295781	4.097621
	Talking	0.756145	2.39114
	ST	0.508196	1.607056
	FT	0.761427	2.407845
	SR	3.906207	12.35251
	MR	1.868391	5.90837

Table 4 Performance analysis of ten subjects from PURE database (continued)

Subject	*Case*	*MAE*	*RMSE*
8	Steady	2.912074	9.208788
	Talking	2.251683	7.120448
	ST	0.894561	2.828852
	FT	5.049209	15.967
	SR	0.212805	0.67295
	MR	0.333977	1.056129
9	Steady	4.99497	15.79
	Talking	3.676976	11.62762
	ST	0.410569	1.298333
	FT	0.06386	0.201944
	SR	0.504851	1.59648
	MR	0.875864	2.769725
10	Steady	4.380654	13.85284
	Talking	5.721565	18.09318
	ST	3.443686	10.88989
	FT	2.782313	8.798446
	SR	3.478123	10.99879
	MR	0.631111	1.995748

Figure 4 MAE of each samples of PURE database using proposed tracking algorithm integrated with CHROM (see online version for colours)



In Table 1, comparison is done for PURE dataset, the result shows that there is a significant improvement in MAE and RMSE then the original method, when PF is applied for tracking. Specifically for the case, fast translation and rotation for angle of 35 degree there is a decrease in MAE and RMSE by large margin. The proposed algorithm exhibit an average improvement of 29.3% improvement in MAE over CHROM, 8.07% improvement in MAE over PBV, 2.08% improvement in MAE over POS and 23.7 % improvement in MAE over Phys-Net.

Figure 5 MAE of each subject of UBFC-rPPG database using proposed tracking algorithm integrated with CHROM (see online version for colours)

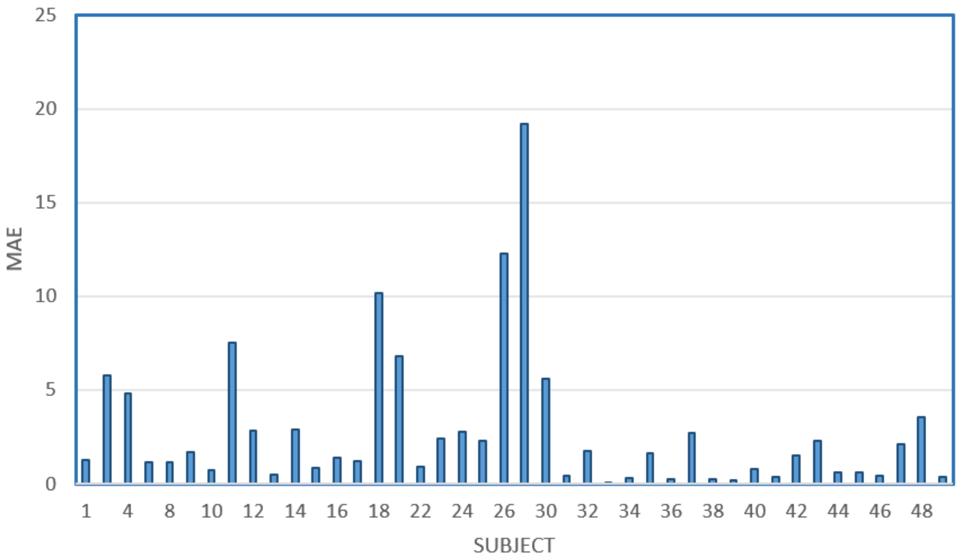


Figure 6 Predicted HR of subject-23 of UBFC-rPPG dataset compared with ground truth (see online version for colours)

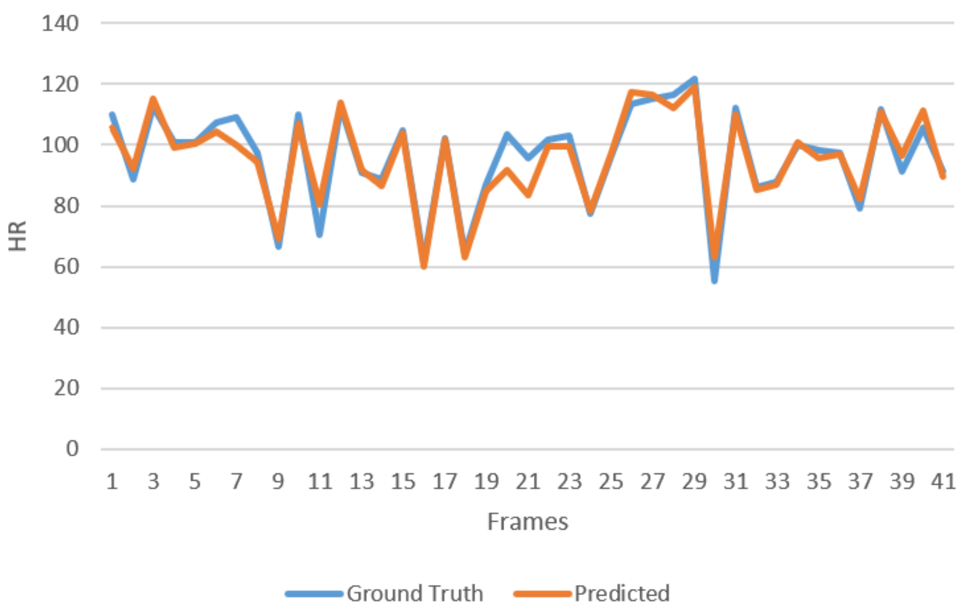


Figure 7 Comparision of ground truth with state-of-the-art methods and our method (see online version for colours)

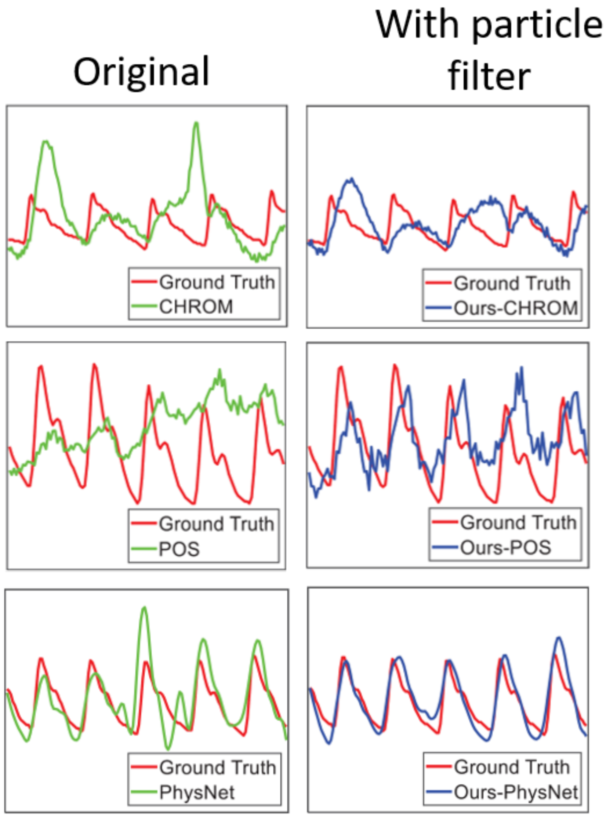


Table 3 depicts the performance of proposed method in UBFC-rPPG dataset. It can be observed that the the MAE improvement of 9.3%, 15.2%, 13.6% and 17.78% is gained on CHROM, PBV, POS and Phys-Net respectively.

Overall from the results of both the datasets, it can be seen that the improvements are more in motion scenarios rather than stationary conditions. This is obvious because the algorithm is designed to work on fast motions.

Figure 7 shows the comparison of ground truth with core rPPG algorithm and proposed algorithm. It can be seen from the waveform that the original methods exhibits sudden high peaks. These anomaly in waveform is due to the abrupt motions. Whereas, after embedding PF method to the original method, the rPPG signal follows the ground truth signal with marginal error. These is in line with the hypotheses that the motion robustness of the core rPPG algorithms would improve with tracking using PF.

6 Discussion

It can be inferred from results, that the proposed model embedded with the core HR estimation algorithms provides robustness against motion. In this section we will analyse the key components behind improved robustness.

The PF-based HR estimation performs best for wide angle motions occurs during physical excercise. In this paper we aimed to build a model with robustness towards motion but without any compromise in performance. In a physical training process, facial movements intends to occur, which interrupts HR measurements. In such cases, landmarks which needs to be tracked may be lost, and thus the conventional method fails. In the proposed method we have used the PF which allows the HR estimates to quickly converge back to the original while tracking.

We have also evaluated our proposed method on yen subjects from UBFC-rPPG dataset as shown in Table 5 by varying the range of particles. It can be observed that the subject-9, the MAE is minimum as the particles is tracking the subject HR very efficiently. Whereas, for subject-1 the MAE is large because there is no large motion occur in the subject.

The performance analysis of the proposed algorithm for even more challenging dataset, i.e., PURE database which offers wide variety of motion. The reason behind choosing this database is the close similarity of motion occurred when a person is performing any kind of exercise and the participants motion in the dataset. From the results shown in Table 4 it can be depicted that the proposed method performs better in case of large motions because the PF reacted quickly to unforeseen changes in HR, which ultimately reduced a delay issue.

7 Conclusions

Our approach addresses the persistent challenge of MAs in remote HR measurement by dynamically tracking the rPPG signal trajectory rather than merely filtering out motion noise. Leveraging PFs, we ensure that the estimated rPPG signal closely follows the ground truth, preserving signal integrity even under extensive motion. A key advantage of this method is its independence from specific rPPG extraction techniques, allowing it to be seamlessly integrated with existing methods to significantly enhance accuracy and robustness. This tracking-based framework represents a versatile and effective advancement in non-contact physiological monitoring.

Declarations

All authors declare that they have no conflicts of interest.

References

Asthana, A., Zafeiriou, S., Cheng, S. and Pantic, M. (2013) 'Robust discriminative response map fitting with constrained local models', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.3444–3451.

Bourel, F., Chibelushi, C.C. and Low, A.A. (2000) 'Robust facial feature tracking', in *BMVC*, pp.1–10.

Chen, W. and McDuff, D. (2018) 'DeepPHYS: video-based physiological measurement using convolutional attention networks', in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp.349–365.

Das, M., Choudhary, T., Bhuyan, M., Sharma, L. et al. (2022) 'A multiresolution method for non-contact heart rate estimation using facial video frames', in *2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, IEEE, pp.115–119.

De Haan, G. and Jeanne, V. (2013) 'Robust pulse rate from chrominance-based rPPG', *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 10, pp.2878–2886.

De Haan, G. and Van Leest, A. (2014) 'Improved motion robustness of remote-PPG by using the blood volume pulse signature', *Physiological Measurement*, Vol. 35, No. 9, p.1913.

Fiaz, M., Mahmood, A., Javed, S. and Jung, S.K. (2019) 'Handcrafted and deep trackers: recent visual object tracking approaches and trends', *ACM Computing Surveys (CSUR)*, Vol. 52, No. 2, pp.1–44.

Galli, A., Narduzzi, C. and Giorgi, G. (2017) 'Measuring heart rate during physical exercise by subspace decomposition and Kalman smoothing', *IEEE Transactions on Instrumentation and Measurement*, Vol. 67, No. 5, pp.1102–1110.

Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R. and Nordlund, P-J. (2002) 'Particle filters for positioning, navigation, and tracking', *IEEE Transactions on signal processing*, Vol. 50, No. 2, pp.425–437.

Kumar, M., Veeraraghavan, A. and Sabharwal, A. (2015) 'DistancePPG: robust non-contact vital signs monitoring using a camera', *Biomedical Optics Express*, Vol. 6, No. 5, pp.1565–1588.

Lam, A. and Kuno, Y. (2015) 'Robust heart rate measurement from video using select random patches', in *Proceedings of the IEEE International Conference on Computer Vision*, pp.3640–3648.

Lewandowska, M., Rumiński, J., Kocejko, T. and Nowak, J. (2011) 'Measuring pulse rate with a webcam – a non-contact method for evaluating cardiac activity', in *2011 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE, pp.405–410.

Li, X., Chen, J., Zhao, G. and Pietikainen, M. (2014) 'Remote heart rate measurement from face videos under realistic situations', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.4264–4271.

Liu, X., Yang, X., Wang, D. and Wong, A. (2020) 'Detecting pulse rates from facial videos recorded in unstable lighting conditions: an adaptive spatiotemporal homomorphic filtering algorithm', *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, pp.1–15.

Lokendra, B. and Puneet, G. (2022) 'AND-rPPG: a novel denoising-rppg network for improving remote heart rate estimation', *Computers in Biology and Medicine*, Vol. 141, p.105146.

Macwan, R., Benezeth, Y. and Mansouri, A. (2019) 'Heart rate estimation using remote photoplethysmography with multi-objective optimization', *Biomedical Signal Processing and Control*, Vol. 49, pp.24–33.

Nauman, J., Janszky, I., Vatten, L.J. and Wisløff, U. (2011) 'Temporal changes in resting heart rate and deaths from ischemic heart disease', *JAMA*, December, Vol. 306, pp.2579–2587.

Niu, X., Han, H., Shan, S. and Chen, X. (2018) 'SynRhythm: learning a deep heart rate estimator from general to specific', in *2018 24th International Conference on Pattern Recognition (ICPR)*, pp.3580–3585.

Niu, X., Shan, S., Han, H. and Chen, X. (2019) 'RhythmNet: end-to-end heart rate estimation from face via spatial-temporal representation', *IEEE Transactions on Image Processing*, Vol. 29, pp.2409–2423.

Poh, M-Z., McDuff, D.J. and Picard, R.W. (2010a) 'Advancements in noncontact, multiparameter physiological measurements using a webcam', *IEEE Transactions on Biomedical Engineering*, Vol. 58, No. 1, pp.7–11.

Poh, M-Z., McDuff, D.J. and Picard, R.W. (2010b) 'Non-contact, automated cardiac pulse measurements using video imaging and blind source separation', *Optics Express*, Vol. 18, No. 10, pp.10762–10774.

Prakash, S.K.A. and Tucker, C.S. (2018) 'Bounded kalman filter method for motion-robust, non-contact heart rate estimation', *Biomedical Optics Express*, Vol. 9, No. 2, pp.873–897.

Qiu, Y., Liu, Y., Arteaga-Falconi, J., Dong, H. and El Saddik, A. (2018) 'EVM-CNN: real-time contactless heart rate estimation from facial video', *IEEE Transactions on Multimedia*, Vol. 21, No. 7, pp.1778–1787.

Spetlik, R., Cech, J., Franc, V. and Matas, J. (2018) 'Visual heart rate estimation with convolutional neural network', in *Proceedings of the British Machine 293 Vision Conference*, Newcastle, UK, pp.3–6.

Stricker, R., Müller, S. and Gross, H-M. (2014) 'Non-contact video-based pulse rate measurement on a mobile service robot', in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, IEEE, pp.1056–1062.

Verkruysse, W., Svaasand, L.O. and Nelson, J.S. (2008) 'Remote plethysmographic imaging using ambient light.', *Optics Express*, Vol. 16, No. 26, pp.21434–21445.

Wang, W., Den Brinker, A.C., Stuijk, S. and De Haan, G. (2016) 'Algorithmic principles of remote PPG', *IEEE Transactions on Biomedical Engineering*, Vol. 64, No. 7, pp.1479–1491.

Xu, T., Feng, Z-H., Wu, X-J. and Kittler, J. (2019) 'Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking', *IEEE Transactions on Image Processing*, Vol. 28, No. 11, pp.5596–5609.

Yu, Z., Li, X. and Zhao, G. (2019) *Remote Photoplethysmograph Signal Measurement from Facial Videos using Spatio-Temporal Networks*, arXiv preprint arXiv:1905.02419.

Yu, Z., Li, X., Niu, X., Shi, J. and Zhao, G. (2020) 'AutoHT: a strong end-to-end baseline for remote heart rate measurement with neural searching', *IEEE Signal Processing Letters*, Vol. 27, pp.1245–1249.