Monte Carlo Filter-Based Motion Artifact Removal From Electrocardiogram Signal for Real-Time Telecardiology System

Soumyendu Banerjee, *Graduate Student Member, IEEE*, and Girish Kumar Singh

***Abstract*— Motion artifact (MA) contamination with electrocardiogram (ECG) signal is a common issue caused by body movement or sensor loosening, resulting in distortion of clinical features of ECG. In this work, the Monte Carlo filter (MCF)-based MA removal from single-channel ECG signal is proposed, assisting in real-time telecardiology systems. Initially, after R-peak detection and beat extraction, principal component (PC) analysis was performed upon clean ECG beats, and PC, with the highest energy, was assumed to be the feature beat. Using this feature beat, MA corrupted beats were denoised successively to achieve a clean pattern of ECG using MCF. A new approach of weight calculation and resampling was also proposed for better performance of the MCF. Performance of the proposed algorithm was tested on the IEEE Signal Processing Cup Challenge 2015 ECG database and MIT-BIH arrythmia records, with an improvement of signal-to-noise ratio between 10 and 15 dB, after MA removal. The proposed work was also tested on real-time ECG data collected from ten healthy volunteers using the AD8232 ECG module and Raspberry Pi, resulting in correlation coefficient higher than 0.99, between the original and denoised signals. The proposed algorithm was able to remove MA from any single-channel MA corrupted ECG signal, irrespective of lead category, using features of clean beats. A comparative study of the obtained result with previously published works ensured the superior performance on MA removal from ECG in the proposed work, along with real-time data collection, processing, and transmission.**

***Index Terms*— Electrocardiogram (ECG), Monte Carlo filter (MCF), motion artifact (MA), Raspberry Pi (Rs-Pi), signal-to-noise ratio improvement.**

I. Introduction

A. Overview and Motivation

ELECTROCARDIOGRAM (ECG) signal plays a vital role while detecting cardiac activities, caused by simultaneous contraction and expansion of human heart [1]. ECG is measured by computing potential differences between electrodes, which are placed at specific positions on human body. A clean structure of ECG beat is always necessary for the analysis of various characteristic features, which are related to medical diagnosis (e.g., P and T-waves and QRS complex), and it might be hampered due to the contamination of noisy signals and motion artifacts (MAs). The noisy data can be eliminated by using various bandpass filters [2], notch filter, but the frequency of MA lies within the same frequency range of ECG that contains clinical attributes. Thus, normal filtering techniques are not the successful solution to eliminate those MAs [3]. The MA is generally caused by body movement or sensor loosening, hence hampering the clinical features. In addition, for real-time personal healthcare system, patient undergoes through daily routine while monitoring the ECG signal; thus, if the ECG morphology is hampered due to MA contamination, clinical analysis procedure would be affected.

Manuscript received March 7, 2021; revised June 17, 2021; accepted July 26, 2021. Date of publication August 5, 2021; date of current version August 13, 2021. The Associate Editor coordinating the review process was Dr. Bruno Ando. *(Corresponding author: Soumyendu Banerjee.)*

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand 247667, India (e-mail: banerjeesoumyendu@gmail.com).

Digital Object Identifier 10.1109/TIM.2021.3102737

B. Related Works on MA Removal From ECG

To remove MA, the feature of clean ECG signals is always necessary to realize the general pattern of denoised signal. The feature extraction procedure can be performed by principal component analysis (PCA) [4] and also by independent component analysis (ICA) [5], providing a supervised learning approach to detect the original pattern of the corrupted beat [6]. After feature extraction, various filtering techniques can be successfully implanted to separate noisy data. In this aspect, the adaptive filtering approach has provided better result amongst common filtering techniques, by controlling the filter parameters using the optimization process [7]. The least mean square (LMS) adaptive filtering is a popular filtering tool, to remove MA from various physiological signals [8]. Apart from adaptive filtering, discrete wavelet transform (DWT) is a useful approach to decompose any time-domain signal into time-frequency subbands [9]. Thus, using proper filtering logic, the noisy data can be easily eliminated from the original signal. Xie *et al.* [10] proposed both high- and low-pass filtering approaches, on wavelet decomposed signal, to remove MA. Over DWT, the stationary wavelet transform (SWT) is a better approach, which overcomes the lack of translation invariance problem [11]. Berwal *et al.* [12] proposed two-stage SWT-based level thresholding, to remove MA from raw ECG signals, which resulted in high correlation coefficients (CORR) and low mean squared error (MSE) between the corrupted and denoised signals.

The particle filter [13] is another approach, in which a set of particles update their likelihood weights within a probability distribution function, so that they can coincide to represent a posterior distribution of a random process. The particle filter was also used to [14] remove MA, assuming the clean ECG features as ground truth. In real-time operation, Kalman filtering is a useful tool that predicts and updates future measurements using current state variables, in a recursive way, to filter out the noisy data [15].

C. Contributions and Article Organization

The LMS adaptive filter possesses a slow convergence rate and it is sensitive to input step-size parameters. Although DWT or SWT is very useful tools for denoising, it is difficult to choose the best wavelet category and decomposition level because the noise frequency is usually unknown. On the other hand, the implementation of Kalman filter is restricted in the Gaussian space. In the area of particle filter, the Monte Carlo filter (MCF) [16] provides an approach of forecasting, filtering, and smoothing for any nonlinear and non-Gaussian state-space systems. In this work, MCF was utilized to remove MA from noisy ECG beats, assuming the PCA extracted features of clean ECG beats as ground truth during computation. The main contributions of this work are given in the following:

1) proposal of implementing MCF to eliminate MA from corrupted ECG beat;
2) to introduce a new weight updating technique in MCF while predicting the new values of particle.

In addition, for real-time operation, ECG was collected from human body using the AD8232 module, the entire algorithm was tested using Raspberry Pi (Rs-Pi), and real-time data transmission to cloud storage or smartphone devices was done using an active internet facility.

The rest of this article is organized as follows. Section II describes the hardware setup and methodology, proposed in this work. Section III shows the experimental result on MA removal, implemented on various single-channel ECG signals along with detailed comparison of obtained result with previously published works, and finally, Section IV concludes the inferences drawn from this work.

II. Methodology

Fig. 1 shows the basic block diagram of the proposed work. For real-time data collection, the single-channel ECG was obtained from human body using the AD8232 ECG module, which was preprocessed in a signal conditioning unit (SCU) and an analog-to-digital converter (ADC) before processed in Rs-Pi, to remove MA. Finally, the clean ECG signal was either stored or transmitted to cloud server or smartphone device using an active internet facility.

A. Hardware Setup and Connection Scheme

The entire algorithm was implemented on Rs-Pi [17], which is a single-board computer module, providing standalone operation of data processing. This device operates using a 64-bit quad-core processor with 1.5-GHz clock speed and built-in 4-GB random access memory (RAM). The ECG signal was acquired from the human body using the AD8232 ECG module, which operates in a 3.3-V dc voltage source and is capable of providing single-channel ECG by placing three electrodes (named RA, LA, and LL) on the human chest at specific position [1] (RA: placed under right clavicle within rib cage; LA: placed under left clavicle within rib cage; and LL: placed on lower edge of left rib cage, below pectoral muscles). Rs-Pi can deal with only digital signals, but AD8232 provides analog output voltage; therefore, to convert the output of ECG module into digital signal, an ADC chip MCP3008 was used,

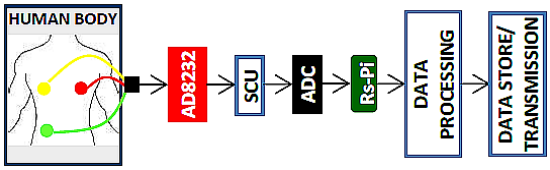

Fig. 1. Block diagram of the proposed work.

and the equivalent digital signals were fed into Rs-Pi. Now, the entire algorithm was processed within the Rs-Pi module, and finally, after MA removal, the clean ECG was either stored or wirelessly transmitted, as desired. Noise contamination into original ECG signal is a common factor, and thus, filtering was a necessary step along with amplification. For this purpose, an SCU was designed, which is mainly composed of four circuit modules (shown in Fig. 2) as described next.

1) *Low-Pass Filter:* To eliminate various high-frequency noises, a second-order zero phase active low-pass filter was designed, with cutoff frequency limited to 120 Hz [18]. In Fig. 2, the low-pass filter was designed using two resistors, viz., R_1 and R_2 and two capacitors, viz., C_1 and C_2, so that the following equation satisfies:

$$1/\left(2\pi\sqrt{R_1R_2C_1C_2}\right) \approx 120. \tag{1}$$

2) *Notch Filter:* This is basically a band-reject filter with narrow bandwidth. It was used to remove power-line interference noise, which is nearly about 50 Hz in India. The values of used resistors (R_3 and R_4) and capacitors (C_3 and C_4) were chosen so that the following equation satisfies:

$$1/(2\pi R_4C_3) \approx 50. \tag{2}$$

3) *High-Pass Filter:* Noise signals, with too low frequency, cause baseline artifact to the ECG signal [19]. Hence, the values of resistance and capacitance were so selected that the cutoff frequency can be limited within 0.5 Hz [20], shown in the following equation:

$$1/\left(2\pi\sqrt{R_5R_6C_5C_6}\right) \approx 0.5. \tag{3}$$

4) Instrumentation Amplifier (IA): This amplifier falls in the differential amplifier category because it amplifies the difference signals of two inputs. The main advantage of using IA is that it possesses a high common-mode rejection ratio and thus can eliminate unnecessary noises, generated by the circuit. In Fig. 2, the IA was constituted using three Op-amps and five resistors (R_7–R_{13}) with gain as follows:

$$\text{Gain} = [\{1 + (R_7 + R_9)/R_8\} \times (R_{13}/R_{10})]. \tag{4}$$

The value of the resistors was so chosen that the amplitude of ECG was limited to 0.1 and 5 mV. For ease of operation, in this work, $R_{13} = R_{12}$, $R_{10} = R_{11}$, and $R_7 = R_8 = R_9$ were chosen.

As center frequency of bandpass filter was less than half of the sampling frequency (SF), which is 125 Hz, antialiasing filter was avoided in the filter circuit module. After processing through SCU, the analog output voltage was converted

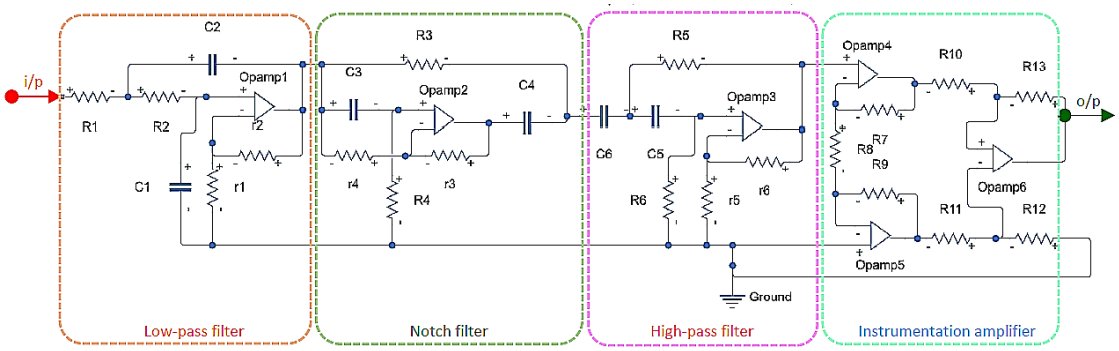

Fig. 2. Circuit diagram of SCU, proposed in this work.

to an equivalent digital signal using the MCP3008 ADC module. The data collection was performed in 125-Hz SF with 10-bit per sample resolution. The digital signal was finally transferred to Rs-Pi through the SPI serial connection protocol. The pseudocode for data collection procedure is shown in Algorithm 1.

Algorithm 1 Pseudocode for Data Collection in Rs-Pi From ADC

```
Initialize Rs-Pi
Initialize GPIO ports ## define input–output ports of Rs-Pi
Initialize SPI configuration
Define pin configuration ## CLK, MISO, MOSI, CS)
Define hardware configuration ## SPI_PORT, SPI_DEVICE
Define variable:
    port = 0 ## select CH0 as analog input pin
Define readadc(adcnm) ## adcnm refers to input port
      number
    if adcnm >7 or adcnm < 0, then return -1 ## define port
       0 to 7
    p = spi.xfer2([1, 8 + adcnm ≪ 4, 0])
    val = ((p[1] & 3) ≪ 8) + p[2]
    return val
while true:
    port_value = readadc(port)
    x = [x;port_value] ## store adc values into variable, 'x'
delay
```

B. Peak Detection and Beat Extraction in Rs-Pi

The initial task in Rs-Pi before MA removal was to detect R-peaks in ECG signal followed by the beat extraction. In this work, the derivative-based R-peak detection procedure was implemented based on the Pan–Tompkins algorithm [21]. In this process, the sample-to-sample differentiation was computed and the position of R-peaks was easily detected containing comparatively higher energy zones. After detecting R-peaks, the corresponding beat was estimated using the following equation [22]:

$$\begin{aligned} R_n^{\text{on}} &= R_n - 0.33 \times ({}_{n-1}RR_n) \\ R_n^{\text{off}} &= R_n + 0.67 \times ({}_{n}RR_{n+1}) \end{aligned} \tag{5}$$

where R_n^{on} and R_n^{off} represent the onset and offset, respectively, of estimated beat of the nth R-peak, R_n; ${}_{n1}RR_{n2}$ represent the RR interval between the $n1$ and the $n2$th R-peak.

C. Definition of MCF

The Monte Carlo simulation [23] process has the ability to handle general structures of non-Gaussian and nonlinear noises. Any time series data can be achieved by the following state-space model, which is nonlinear and non-Gaussian in nature:

$$\begin{aligned} x_k &= f(x_{k-1}, v_k) \\ y_k &= h(x_k, w_k) \end{aligned} \tag{6}$$

where x_k is an l-dimensional state vector, which can be estimated using the r-dimensional observations state y_k at time instance, k. The two nonlinear functions $f(\cdot)$ and $h(\cdot)$ are named state transition function and state observation function, respectively. v_k and w_k are called system noise and observation noise, respectively.

The main problem of these state estimation equations mentioned in (6) is that it can be formulated as an evaluation of the probability density function (PDF) as $p(x_t|Y_{1:k})$, where $Y_{1:k}$ represents the observations $\{Y_l, Y_2, \ldots, Y_k\}$. Now, the important task of state-space problem is to find x_k using the observation of y_k. Based on the observation states, the state estimation problem can be classified into three cases as, $k > t$, $k = t$, and $k < t$ that are called predictor, filter, and smoother, respectively. These stages are explained as follows.

1) *Prediction:* PDF of the state vector x_k is computed at the present time instance k using all the past observations from time instances 1 to $k-1$ by the following equation:

$$p(x_k|Y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Y_{1:k-1})dx_{k-1}. \tag{7}$$

2) *Filter:* The density-based filtering is performed by weight updating and resampling, based on the observation y_k. For any particle denoted as p_k^i, the PDF is computed as

$$\delta_k^i = p\left(y_k|p_k^i\right) = \phi\left(g\left(y_k, p_k^i\right)\right)\left|{}^{\partial g}/_{\partial y_k}\right| \tag{8}$$

where $i = 1, 2, \ldots, r$ and $g(\cdot)$ is the inverse of function $h(\cdot)$ and $\phi(\cdot)$ denotes the density of conservation noise. In this way, r particles $[q_k^1, q_k^2, \ldots, q_k^r]$ are by resampling the particles $[p_k^1, p_k^2, \ldots, p_k^r]$, along with the PDF proportional to $[\delta_k^i, \delta_k^i, \ldots, \delta_k^r]$. $[q_k^1, q_k^2, \ldots, q_k^r]$ is realized as filter, i.e., $p(x_k|Y_{1:k})$.

3) *Updating:* For initial time step, i.e., $k = 0$, total r particles were randomly considered from the initial

PDF, $p(x_0)$, which is denoted as, $x_0^i \sim p(x_0)$, where $i = 1, 2, \ldots, r$.

Now, the prediction was performed for each particle as $x_{k|k-1}$ at time step k, using the previous time step $x_{k-1|k-1}$, and the associated weights were updated as $w_k^i = p(y_k|X_{0|k}^i, Y_{0|k-1})$. Finally, the weighted sum was calculated as

$$W_k^i = w_k^i \Big/ \left(\sum_{l=1}^{k} w_k^j \right). \tag{9}$$

4) *Resampling:* The resampling was performed (if necessary) using the following equation:

$$x_{k+1|k}^i \sim p\left(\mathrm{x}_{k+1|k}|X_k^i, Y_k\right). \tag{10}$$

The resampling was necessary to generate the new set of particles in each iteration, giving preference to particles with larger weights in the previous stage.

D. Proposed Approach of Filtering

To forecast a clean pattern of the corrupted signal, it was always necessary to analyze a comparison with a ground truth of general pattern of ECG. Therefore, after beat extraction from ECG signals, PCA extracted feature of five consecutive clean beats, occurred before the corrupted beat was computed. For this purpose, a beat matrix B with order $(m \times l)$, containing five successive beats, was created by the last sample padding (where l is the normalized beat length after padding) [24]. Now, PCA was implemented on this beat matrix as follows [25].

At first, the corresponding mean subtracted matrix $\bar{B}$ of matrix B was computed as

$$\begin{aligned} \mathrm{BM}_{1,l} &= \sum_{i=1}^{l} B_{i,m}/l \\ \bar{B}_{i,m} &= B_{i,m} - \mathrm{BM}_{1,m}. \end{aligned} \tag{11}$$

Now, the eigenvalue–eigenvector decomposition was performed as

$$\begin{aligned} \varepsilon &= \left[B \times B^{\mathrm{T}}\right] \\ \varepsilon \times \nu &= \sigma \times \nu \\ \phi &= \nu^{\mathrm{T}} \times \bar{B} \end{aligned} \tag{12}$$

where ν and σ represent the eigenvector and eigenvalue matrix, respectively, and Φ represents the principal component (PC) matrix with all PCs, arranged in decreasing order of variance. From this matrix, PC with the highest variance was assumed to be the ground truth of successive MA contaminated beats, from which the noise signal was eliminated using MCF [26] as discussed next.

Let us consider that, for any time instant t, M_t^i and N_t^i represent the ith sample or particle of PCA extracted beat M and MA contaminated beat N, respectively. The length of both signals was made the same by the last sample padding so that the R-peaks of both beats lie on the same position. Now, the initial distribution of each particle of signal N was made bound between $N_t^i \pm \{\alpha(|N_t^i - M_t^i|)\}$, where α is a positive integer such as $1.2 \leq \alpha \leq 0.98$. The limits of α were chosen based on the hit and trial method, and it was observed that the accuracy of filtering was very high within this range. Along with this, three parameters, denoted as $F1$, $F2$, and $F3$, were computed for classifying these two signals as mentioned in the following:

$$\begin{aligned} \overline{Z} &= \left(\sum_{i=1}^{h} z_i \right) \Big/ h \\ \sigma(Z) &= \sqrt{\sum_{i=1}^{h} \left(z_i - \overline{Z}\right)^2 / h} \\ \xi_1(Z) &= \left(\overline{Z} - Z_{\min}\right)/\sigma(Z) \\ \xi_2(Z) &= \left(Z_{\max} - \overline{Z}\right)/\sigma(Z) \\ F1 &= [\{\sigma(M) - \sigma(N)\}/\sigma(M)] \\ F2 &= [\{\xi_1(M) - \xi_1(N)\}/\xi_1(M)] \\ F3 &= [\{\xi_2(M) - \xi_2(N)\}/\xi_2(M)] \end{aligned} \tag{13}$$

where z_i denotes the ith sample of signal Z and $\sigma(z)$ represents the standard deviation (STD) of signal Z with length h. Using these equations, the STD of M and N was computed, which were mentioned as $\sigma(M)$ and $\sigma(N)$, from which the three features were computed as $F1$, $F2$, and $F3$, as shown in (13). Now, within the random distribution mentioned earlier, the filtering procedure was initiated as follows initialized by updating the likelihood weights. The weights of each particle were a measurement of deviation from the ground truth. For this purpose, three features were used to update each particle weight as

$$\begin{aligned} w_k\left(\mathrm{x}_{k|k-1}^i\right) &= N\left(\mathrm{x}_{k|k-1}^i\right) + \tanh(\rho) \\ \rho &= \{(c1 \times F1 + c2 \times F2 + c3 \times F3)/3\} \end{aligned} \tag{14}$$

where $c1$, $c2$, and, $c3$ are the random numbers between 0.5 and 0.9. Finally, the weights were normalized so that the addition of all weights becomes one.

For resampling, Kitagawa [27] proposed "random sampling" and "two-stage filter" methods, which are computationally complex. For this purpose, in this work, the sampling was made quite easier way by taking a moving window of five samples in each iteration. Thus, each sample passed through five stages of filtering, and the weights were updated accordingly so that the probability is

$$p\left(x_{k|k-1}^i\right) = \{0.9 - (h \times 0.05)\} \sum_{x_{k|k-1}^i} w_k\left(x_{k|k-1}^i\right) \tag{15}$$

where $h = 1, 2, \ldots, 5$ for each sample. Now, if the particle number was P, then the following steps were followed for resampling as given.

1) $\tilde{p} = P \times p(x_{k|k-1}^i)$.
2) Round off $\tilde{p}$ to its nearest integer.
3) Generate samples within the new range $\tilde{p}$.

The computational complication was lower in this proposed method. As in this direct method of resampling, convergence time might be faster resulting in lower accuracy of filtering, and the moving window of five samples was used so that the accuracy of filtering be increased step by step. The three

TABLE I
INFORMATION ABOUT VOLUNTEERS

Sl. No.	Information about Volunteer				
	Weight (kg)	Gender	Height (cm)	Age	Diagnosis
1	53	M	156	18	Normal[a]
2	61	M	158	43	Low BP[b]
3	66	M	149	32	Normal
4	68	M	152	35	Normal
5	52	F	157	27	Normal
6	54	F	157	24	Normal
7	77	M	154	58	High BP
8	51	M	172	26	Normal
9	61	M	168	25	Normal
10	78	F	154	27	Normal

[b]BP: blood pressure; [a]Normal: normal BP and no other specific abnormality

features $F1$, $F2$, and $F3$ were utilized to update the noisy beat to its clean pattern using the features of M; the $\tanh(\cdot)$ operator helped to update the weights of each particle within a specific limit. Random variables were utilized to prevent the noisy beat, "N" from attaining the exact pattern of feature beat, M. The entire process of MCF is shown in Algorithm 2.

Algorithm 2 Algorithm for Monte Carlo Filter

```
Initialize time step t
Generate new sample x_0^i ~ p(x_0)
For i = 1:N
    Define: p(x_t|Y_{1:k}) using {Y_1, Y_2, ..., Y_{k-1}}
    Calculate features F1, F2, and F3
    Compute the weights
    Normalize weights
    Update the weights
    Resample
Return
```

III. RESULTS AND DISCUSSION

The entire algorithm was first tested IEEE Signal Processing Cup Challenge (SPCC) 2015 database [28] and MIT-BIH arrhythmia records. In addition, for real-time operation purpose, the algorithm was implemented on single-channel ECG records, obtained from ten healthy volunteers with their informed consent in the Biomedical Research Laboratory, Indian Institute of Technology Roorkee, Roorkee, India. The details of volunteers are provided in Table I. The number of volunteers was chosen arbitrarily to examine the on-device performance of the proposed algorithm. The diagnosis was performed using the standard digital blood pressure (BP) measuring instrument. The hardware setup and connection scheme of various components are discussed earlier. Hence, the quality measure metrics, which were used to analyze the performance of the proposed algorithm, are mentioned below followed by the detailed result analysis.

A. Quality Measure Indices for R-Peak Detection

The performance of R-peak detection was measured using positive predictivity (+P) and sensitivity (Se), as follows:

$$\begin{aligned} +\mathrm{P} &= \mathrm{TP}/(\mathrm{TP}+\mathrm{FP}) \\ \mathrm{Se} &= \mathrm{TP}/(\mathrm{TP}+\mathrm{FN}) \end{aligned} \tag{16}$$

where true positive (TP) represents the number of perfectly detected peaks, false positive (FP) represents the number of miss-detected peaks, and false negative (FN) represents the number of wrong detected peaks.

B. Quality Measure Indices for MA Elimination

The performance of the proposed algorithm was measured using the following quality measure metrics as follows:

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\sum_{i=1}^{\lambda}(U_i - V_i)^2/\lambda} \\ \mathrm{SNR} &= 10\log_{10}\left\{\sum_{i=1}^{\lambda}U_i^2 \Big/ \sum_{i=1}^{\lambda}(U_i - V_i)^2\right\} \\ \mathrm{SNR_{imp}} &= 10\log_{10}\left[\left\{\sum_{i=1}^{\lambda}(U_i - V_i)^2\right\} \Big/ \left\{\sum_{i=1}^{\lambda}(U_i - T_i)^2\right\}\right] \\ \mathrm{CORR} &= \text{Correlation Coefficeint} \end{aligned} \tag{17}$$

where U_i, V_i, and T_i represent the ith sample of original signal U, corrupted signal V, and filtered signal T, respectively, with the same signal length, λ. The RMSE, SNR, and $\mathrm{SNR_{imp}}$ represent the root-mean-square error, signal-to-noise ratio, and SNR improvement, respectively. Another quality index was measured, mentioned as fractional distortion error (FDE) to head-to-dead analysis of various ECG features, viz., P-, Q-, R-, S-, and T-peak amplitude within denoised beat with respect to that of original beat. The FDE was measured as

$$^{X}\mathrm{FDE}_{Y} = (\zeta_X - \zeta_Y)/\zeta_Y \tag{18}$$

where $^{X}\mathrm{FDE}_{Y}$ represents the variation of various features of signal X, i.e., ζ_X with respect to the same feature within signal Y, i.e., ζ_Y.

C. Experimental Result on R-Peak Detection

The R-peaks detection procedure was performed by a derivative-based approach inspired by the Pan–Tompkins algorithm. This process provides low execution time to detect the positions of R-peaks with less computational complexions. The average Se and +P were achieved of values 99.90 and 99.92 for this database, respectively. A pictorial representation of peak detection is shown in Fig. 3. The peak detection procedure was necessary to generate the respective beats, which were used to extract features using PCA.

D. Experimental Result on Real-Time Data

The real-time data were obtained from ten healthy volunteers, using Rs-Pi, AD8232 ECG module, and MCP3008 ADC chip. For noise elimination and amplification, an SCU was designed as mentioned earlier. The ECG module contains three electrodes, which were placed on the human chest so that "Einthoven's triangle" law can be satisfied. Each signal was recorded in 125-Hz SF with 10-bit/sample data resolution. Now, to achieve MA corrupted signal, each volunteer was undergone through two basic body-movement stages, and hence, data collection was performed in three stages as follows.

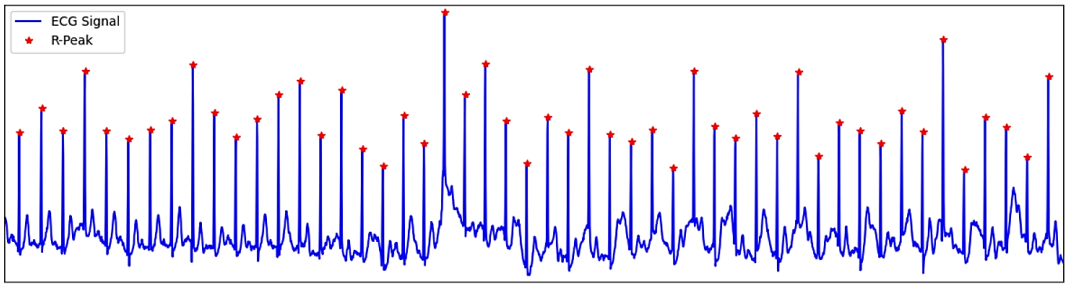

Fig. 3. R-peak detection performance on single-channel ECG marked in "blue" with detected peaks marked in "*" with "red."

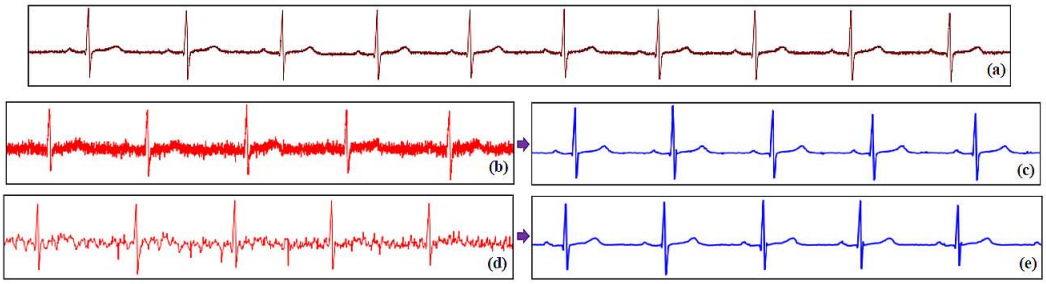

Fig. 4. Experimental result on real-time data of volunteer-01: (a) original pattern on ECG in Stage-1 i.e., resting period, (b) and (c) MA corrupted and respective denoised signal, respectively, in Stage-2, and (d) and (e) MA corrupted and respective denoised signal, respectively, in Stage-3.

1) *Stage-1:* Resting period for 60 s. This period was necessary to obtain the clean ECG signal pattern for each volunteer, from which the feature beat was extracted using PCA.
2) *Stage-2:* Body shaking caused by moderate jumping for 60 s, which resulted in contamination of high MA into the ECG signal.
3) *Stage-3:* In the third stage, to generate high MA, 10–20 push-ups were performed by every volunteer and data recording was done for 30-s duration.

Fig. 4 shows the experimental result on MA removal using the proposed algorithm on the volunteer data-1. It can be easily observed that the MA corrupted signals (marked in "red") are denoised, thus resulting in clean ECG signal (marked in "blue"). The pattern of ECG in the resting period is shown in "marron." However, the main disadvantage of real-time data is the unavailability of clean ECG pattern on the MA corrupted signal. In addition, result comparison can also not be performed. Therefore, the algorithm was tested on globally accepted MA corrupted signals as follows.

E. Experimental Result on IEEE SPCC Data

The IEEE SPCC 2015 database contains six-channel measurement that includes one channel ECG, two-channel PPG, and three-channel accelerometer data, among which only the ECG signal was used for computation. The signals were recorded from 24 patients who were put at a physical resting period followed by the systematic physical exercise, with fixed duration. The dataset contains two types of records named (TYPE01) and TYPE02. Table II shows the speed and time step for both these records. In Table II, "A|B" (second row of each column) represents the running speed for two types of records with A and B km/h, respectively, along with the duration of running in the first row of respective column. In this work, TYPE02 was used for computation.

Due to the initial resting period (i.e., 0 km/h), the clean ECG beats were available, from which PCA extracted feature beats were computed, which were assumed to be ground truth measurement. The main task was to filter the ECG signals recorded during the running period using MCF. To remove MA from noisy ECG beat, a features beat was computed using five clean ECG beats, generated before the noisy beat. In this way, the new denoised beat was used for further feature extraction procurement and removing MA from the further beats. Fig. 5 shows a pictorial representation of MA removal from noisy ECG signal, tested on record id: DATA_05_TYPE02.mat. The clean ECG signal generated during the resting period is represented in "maroon" hue in Fig. 5. The MA corrupted signal, generated further (in running period, marked in "red"), was denoised using the feature of these clean ECG beats, and the resulted signal is shown in "blue." Because the clean and noisy ECG signals were created at different time spans, the denoised ECG did not have the same clean ECG pattern. However, following a visual inspection by medical professionals, it was determined that the properties of the denoised signal were nearly identical to those of the clean signal. The main disadvantage of this database is that the original clean signal pattern was not available during the time period when the MA distorted signal was recorded. As a result, the denoise signal characteristics could not be compared with its original and clean pattern. For this purpose, noisy signal (the noise signal was extracted from the noisy beats after getting it denoised) was manually added to a clean ECG beat, and then, the noisy beat was further filtered. Now, the quality of denoised beat was compared with respect to the original beat using various quality indices mentioned earlier.

The accuracy of MA removal was tested using the following steps.

Due to the unavailability of the original signal during running period, the accuracy of the proposed algorithm

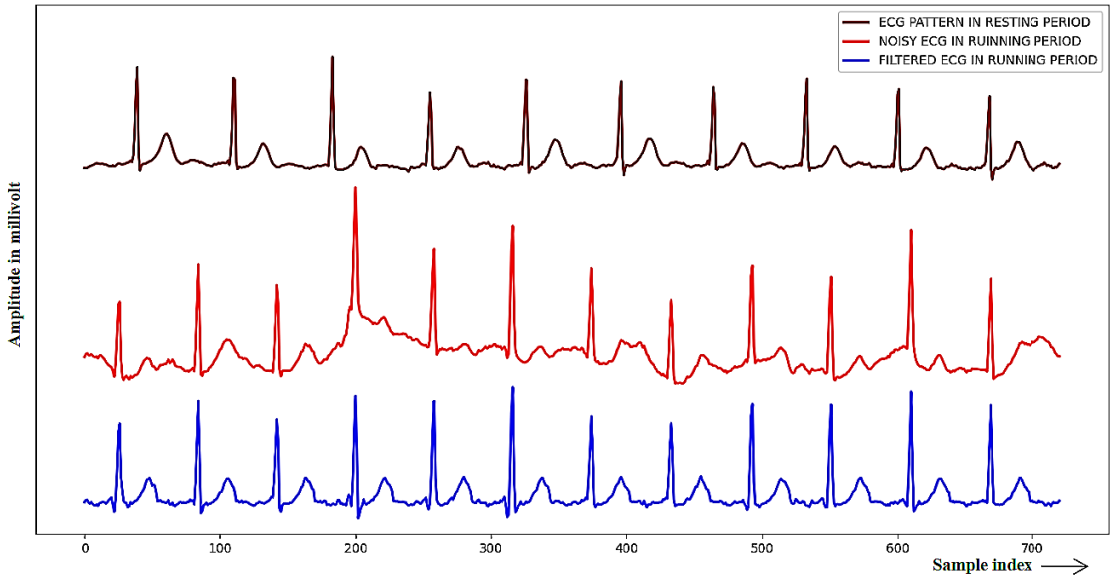

Fig. 5. Experimental result on the IEEE Signal Processing Cup2015 database with records id: DATA_05_TYPE02.mat.

TABLE II
DETAILS ABOUT THE IEEE SPCC 2015 DATABASE

Duration (s)	30	60	60	60	60	30
Speed (km/h)	0|0	8|6	15|12	8|6	15|12	0|0

(i.e., comparison between the original and the MA eliminated signal) was measured in an indirect manner using the following stages.

1) For the Kth corrupted beat, feature beat ${}^{K-1}M_{K-5}$ was computed using five successive beats (i.e., $K-5$, $K-4, \ldots, K-1$.) occurred before the Kth beat.
2) Using this feature beat, the corrupted beat (i.e., Kth beat) was denoised, resulting in a new MA eliminated beat, ${}^{F}K$.
3) Then, the noise signal was computed by taking intersample differentiation between K and ${}^{F}K$, and it was added to $(K-1)$th clean beat, to generate an MA corrupted beat, ${}^{C}(K-1)$.
4) Now, using the feature beat, ${}^{K-2}M_{K-6}$, the noise-contaminated beat ${}^{C}(K-1)$ was filtered using MCF to obtain the filtered beat ${}^{F}(K-1)$.

The accuracy of algorithm was tested between the original beat $(K-1)$ and the obtained beat ${}^{F}(K-1)$ using various parameters mentioned earlier.

A pictorial representation of this denoising procedure is shown in Fig. 6. In Fig. 6(a), the original ECG beat (i.e., $(K-1)$th clean beat, mentioned in the algorithm, shown above) is shown. After noise contamination (obtained by intersample differentiation between K and ${}^{F}K$, mentioned above) with variable noise power, the respective CORR is shown in Fig. 6(b), (d), and (f), and after denoising, the respective beat structure is shown in Fig. 6(c), (e), and (g), respectively, with improved CORR. Table III shows the improvement of RMSE and CORR for these signals, and the comparison was made for both after noise contamination and noise removal. Table IV shows the head-to-head analysis of improvements of various beat features after denoising. The improvements factor (IF)

TABLE III
EXPERIMENTAL RESULT USING IEEE SPCC 2015 RECORDS

Noise power	**Quality metrics**			
	Noisy beat		**Denoised beat**	
SNR (in dB)	**RMSE**	**CORR**	**RMSE**	**CORR**
2.6	0.785	0.8199	0.2	0.9820
9.77	0.653	0.9558	0.098	0.9909
14.06	0.441	0.9818	0.1	0.9989

TABLE IV
IMPROVEMENT OF FDE AFTER MA REMOVAL WITH DIFFERENT NOISE LEVELS

Noise power	**Improvement factor of Beat features**				
SNR	**P-peak**	**Q-peak**	**R-peak**	**S-peak**	**T-peak**
2.6	1.23	0.86	0.61	0.48	0.72
9.77	0.95	0.76	0.14	0.11	1.01
14.06	1.17	0.59	4.89	0.24	0.93

was measured using the following formulas:

$$\text{IF} = \left({}^{N}\text{FDE}_O - {}^{D}\text{FDE}_O\right)/{}^{D}\text{FDE}_O \tag{19}$$

where the superscripts "N" and "D" represent the MA contaminated beat and MA eliminated beat, respectively, the subscript "O" represents the original beat. It can be observed that the improvement was almost nearer to 1, and hence, it can be observed that the various features were almost restored after denoised. It was also observed that the positions of R-peaks were the same after denoising the MA corrupted signal, and thus, the RR intervals were intact after denoising.

F. Experimental Result on MIT-BIH Records

The MIT-BIH arrhythmia database contains 48 dual-channel ECG measurements with 360-Hz sampling frequency and 11-bit/sample data resolution. The proposed algorithm was tested on these records (Lead II), which were treated as clean ECG signals. The noise signal was obtained from the "em" signal, which is available in the MIT-BIH noise stress test (NST) database. The noise signal was uniformly scaled to obtain an input SNR of 5, 10, and 15 dB for each record

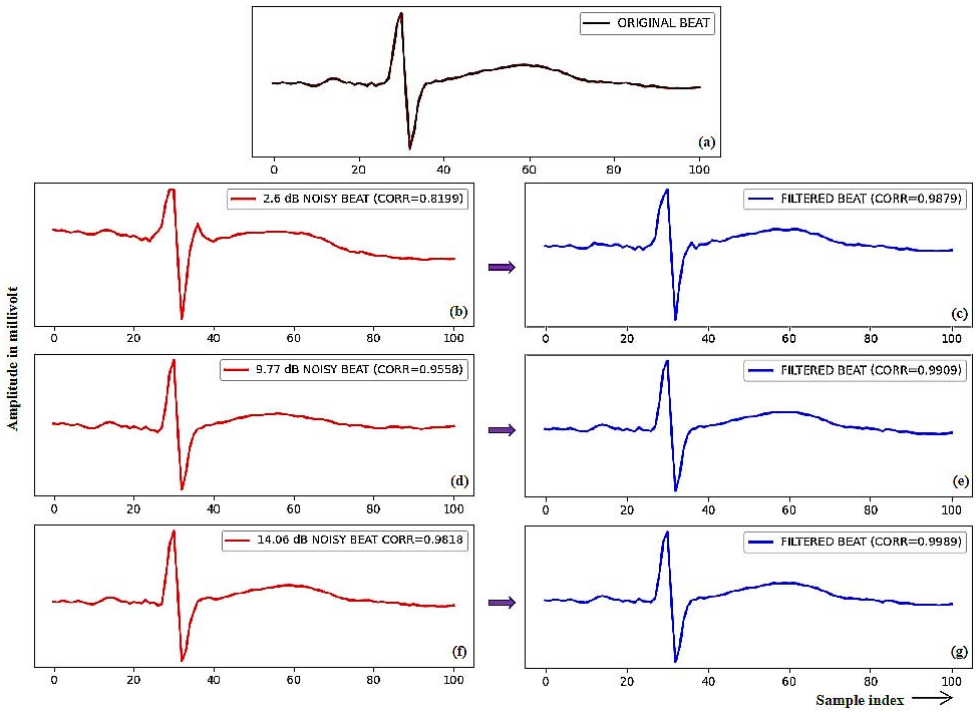

Fig. 6. Experimental result on MA removal using IEEE SPCC 2015 record with CORR improvement: (a) reference beat; 2.6 dB noisy beat and denoised beat shown in (b) and (c), respectively; 9.77 dB noisy beat and denoised beat shown in (d) and (e), respectively; 14.07 dB noisy beat and denoised beat shown in (f) and (g), respectively.

TABLE V
EXPERIMENTAL RESULT USING MIT-BIH RECORDS

Noise power	Quality metrics			
	Noisy signal		Denoised signal	
SNR (in dB)	RMSE	CORR	RMSE	CORR
5	0.1969	0.8152	0.051	0.9812
10	0.1107	0.9265	0.043	0.9892
15	0.0623	0.9774	0.033	0.9914

separately and added to each clean signal to obtain MA corrupted ECG. Fig. 7 shows the effect of noise contamination in MIT-BIH record-id 100, with various input SNRs.

Table V shows the average experimental result on MIT-BIH records on MA removal. The average improvement of CORR and RMSE is shown for various input noise powers. The significant improvement of these parameters ensures the success on MA removal from MIT-BIH records using the NST database as a noise signal. Fig. 8 shows a pictorial representation of MA removal using these records.

G. Comparison With Existing Techniques

Table VI shows a comparative study of experimental results of previously published works on MA removal from noisy ECG signals, with the proposed algorithm. Lin *et al.* [13] proposed the marginalized particle filter to remove MA from ECG. On implementation of their algorithm on MIT-BIH Normal Sinus Rhythm database (128-Hz SF), the improvement of SNR with variable SNR levels is mentioned in Table VI. Nagai *et al.* [11] used SWT on various ECG signals after being corrupted by MA forcefully. The improvement of CORR was from 0.71 to 0.88, achieved for noised and denoised signal, respectively. Xiong and Chen [6] proposed PCA for MA removal from clean ECG signal, available in the MIT-BIH database. The improvement of CORR was much better as proposed in their paper. Wavelet transform-based MA removal was proposed in [10] in recent works. In addition to noise signal (NST database) with various noise powers in MIT-BIH records, the improvement of SNR was far better than mentioned in other works. Compared to published works, in this work, MCF was proposed to remove MA from the corrupted signal. After addition of noise signal with different noise powers, the improvement of SNR along with CORR is shown in Table VI. As the noise signal was arbitrarily added to IEEE SPCC records, it is difficult to perform a similar comparison using this database with previously published works. The improvement of performance metrics outperformed the other works using MIT-BIH records. From the experimental result, it can be concluded that the proposed work provided superior results on MA removal from ECG signal.

H. Discussion

The real-time data collection and processing was performed to check the on-device performance. The physical movements were made arbitrarily only to generate MA corrupted signal [4], [12]. The main advantage of using the IEEE SPCC 2015 database is that each record consists of clean ECG along with MA corrupted ECG. In all the previous works, noisy signal was added to clean ECG artificially, and then, it was denoised. In this work, the MA signal was not generated manually. To analyze the accuracy of MA removal algorithm, various noise signals (obtained from noisy signal after getting it denoised) were added to clean beats, and the improvement

TABLE VI
COMPARISON WITH PREVIOUSLY PUBLISHED WORKS

Related works	Tool	Databases	SNR in dB	SNR improvement	MSE	CORR improvement	
						Corrupted	Denoised
C. Lin, [13], 2011	Particle filter	MIT-BIH Normal Sinus Rhythm database	5	~10-15	-	-	-
			10	~9-12	-	-	-
			15	~5-10	-	-	-
S. Nagai, [11], 2017	SWT	-	-	-	-	0.71	0.88
F. Xiong, [6], 2020	PCA	MIT-BIH arrythmia	-	-	-	0.11	0.608
X. Xie, [10], 2021	DWT	MIT-BIH arrythmia	5	14.1447	0.0040	-	0.9770
Proposed	**Monte Carlo Filter**	**MIT-BIH arrythmia**	**5**	**15.013**	**0.0026**	**0.8152**	**0.9812**
			10	**11.245**	**0.0019**	**0.9265**	**0.9892**
			15	**8.523**	**0.0011**	**0.9774**	**0.9914**
		IEEE Signal Processing Cup Challenge 2015	**2.06**	**11.93**	**0.04**	**0.8199**	**0.9820**
			9.77	**11.8**	**0.009**	**0.9558**	**0.9909**
			14.06	**10.04**	**0.01**	**0.9818**	**0.9889**

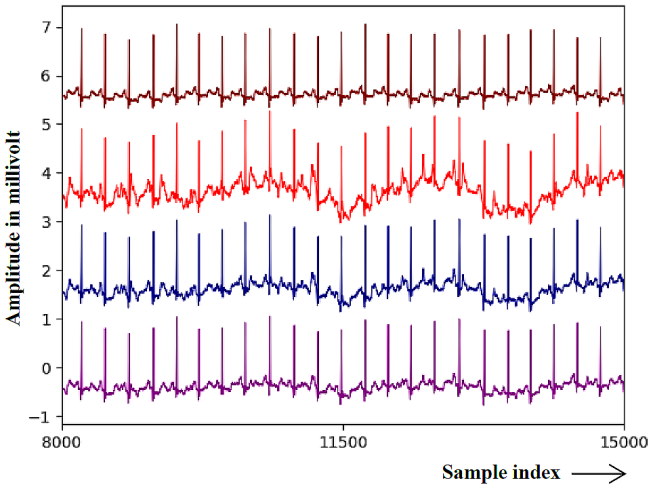

Fig. 7. Original MIT-BIH record id. 100, shown in "marron" (6 mV biased), and MA contaminated signals with an input SNR of 5, 10, and 15 dB, shown in "red," "blue," and "purple," biased by 4, 2, and 0 mV, respectively.

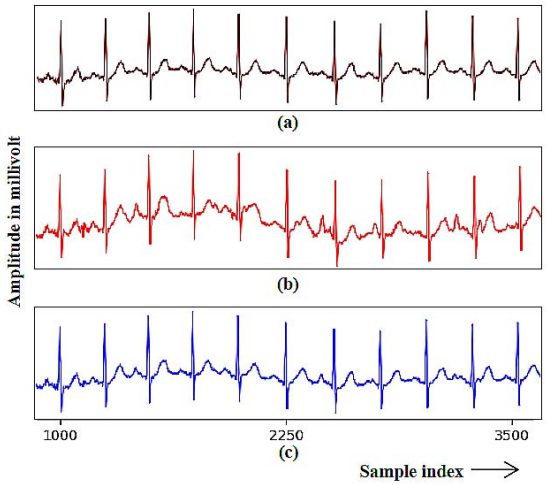

Fig. 8. MA removal from the MIT-BIH record id 209. (a) Original ECG. (b) Noise-contaminated ECG using the first lead of "em" signal from the MIT-BIH NST database with 5-dB SNR. (c) Filtered ECG.

of SNR and CORR is shown in Table V. The improvement of various ECG features after MA removal ensured successful performance of the proposed algorithm. The experimental result was also tested on MIT-BIH arrhythmia ECG records by adding the MIT-BIH NST database ("em" signal) as noise signal, to obtain noisy ECG with different noise powers. In addition, the entire algorithm was performed in a hardware module for real-time operation, thus providing standalone operation. The value of each component used in SCU was chosen based on the availability to meet the desired response. The success of the proposed work is truly dependent on successful R-peak detection. It was also observed that, with the increment of number of MA corrupted beats, the performance of denoising process gradually decayed. The main reason was that each successive corrupted beat was denoised by assuming the PCA extracted feature beat of the previous five clean beats, which might be previously denoised. The algorithm took 214-μs per sample computation in the Rs-Pi. The time of computation excludes the data collection and transmission procedure. For cloud storage, the denoised signal was transmitted and stored to personal online drive (such as Google drive) of patient or the doctor, directly from the Rs-Pi, using an active internet facility. This file can be further downloaded from the drive, on smartphone device, computer, or the Rs-Pi itself. Based on the comparison table (refer to Table VI), it can be concluded that the proposed method is able to remove MA from the corrupted ECG signal, with improved performance.

IV. CONCLUSION

In this article, MCF-based MA removal from single-channel ECG signal is proposed, irrespective of lead category. The clean features of original beat were used to extract the original pattern of corrupted beats. The MCF utilizes the Bayesian filtering approach to provide the filtered state estimated model of any time-domain dynamic model, which is typically non-linear and non-Gaussian in nature, using downsampling of randomly chosen particles in the posterior distribution. The entire algorithm was performed in the IEEE SPCC 2015 ECG database, so that for any MA corrupted signal, the clean feature can be achieved from the same patient. In addition, the proposed work was tested on MIT-BIH records for result analysis and better comparison. For real-time data collection and transmission, hardware implementation was also proposed. A detailed comparison results of the proposed work with

previously published works ensured the superiority of quality of MA eliminated signal assisting in real-time telecardiology applications.

ACKNOWLEDGMENT

The authors would like to thank Dr. Subhasis Mahato, R. G. Kar Medical College and Hospital, Kolkata, for clinical annotations of the electrocardiogram (ECG) data records and helpful guidance and valuable suggestions.

REFERENCES

[1] R. Gupta, M. Mitra, and J. Bera, *ECG Acquisition and Automated Remote Processing*. New Delhi, India: Springer, 2014.
[2] U. Satija, B. Ramkumar, and M. S. Manikandan, "Automated ECG noise detection and classification system for unsupervised healthcare monitoring," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 3, pp. 722–732, May 2018.
[3] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, Dec. 2007.
[4] T. Pawar, S. Chaudhuri, and S. P. Duttagupta, "Body movement activity recognition for ambulatory cardiac monitoring," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 874–882, May 2007.
[5] M. B. Dembrani, K. B. Khanchandani, and A. Zurani, "Accurate detection of ECG signals in ECG monitoring systems by eliminating the motion artifacts and improving the signal quality using SSG filter with DBE," *J. Circuits, Syst. Comput.*, vol. 29, no. 2, Feb. 2020, Art. no. 2050024.
[6] F. Xiong and D. Chen, "CEEMDAN-IMFx-PCA-CICA: An improved single-channel blind source separation in multimedia environment for motion artifact reduction in ambulatory ECG," *Complex Intell. Syst.*, 2020, doi: 10.1007/s40747-020-00188-7.
[7] H. Kim *et al.*, "Motion artifact removal using cascade adaptive filtering for ambulatory ECG monitoring system," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Hsinchu, Taiwan, Nov. 2012, pp. 160–163.
[8] C. Beach, M. Li, E. Balaban, and A. J. Casson, "Motion artefact removal in electroencephalography and electrocardiography by using multichannel inertial measurement units and adaptive filtering," *Healthcare Technol. Lett.*, pp. 1–11, 2021, doi: 10.1049/htl2.12016.
[9] F. R. Hashim, L. Petropoulakis, J. Soraghan, and S. I. Safie, "Wavelet based motion artifact removal for ECG signals," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci.*, Langkawi, Malaysia, Dec. 2012, pp. 339–342.
[10] X. Xie *et al.*, "A multi-stage denoising framework for ambulatory ECG signal based on domain knowledge and motion artifact detection," *Future Gener. Comput. Syst.*, vol. 116, pp. 103–116, Mar. 2021.
[11] S. Nagai, D. Anzai, and J. Wang, "Motion artefact removals for wearable ECG using stationary wavelet transform," *Healthcare Technol. Lett.*, vol. 4, no. 4, pp. 138–141, Jun. 2017.
[12] D. Berwal, C. R. Vandana, S. Dewan, C. V. Jiji, and M. S. Baghini, "Motion artifact removal in ambulatory ECG signal for heart rate variability analysis," *IEEE Sensors J.*, vol. 19, no. 24, pp. 12432–12442, Dec. 2019.
[13] C. Lin, M. Bugallo, C. Mailhes, and J.-Y. Tourneret, "ECG denoising using a dynamical model and a marginalized particle filter," in *Proc. Conf. Rec. 45th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Pacific Grove, CA, USA, Nov. 2011, pp. 1679–1683.
[14] V. Nathan, I. Akkaya, and R. Jafari, "A particle filter framework for the estimation of heart rate from ECG signals corrupted by motion artifacts," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Milan, Italy, Aug. 2015, pp. 6560–6565.
[15] J. Oster, J. Behar, O. Sayadi, S. Nemati, A. E. W. Johnson, and G. D. Clifford, "Semisupervised ECG ventricular beat classification with novelty detection based on switching Kalman filters," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 9, pp. 2125–2134, Sep. 2015.
[16] C. Andrieu, A. Doucet, and E. Punskaya, "Sequential Monte Carlo methods for optimal filtering," in *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer-Verlag, 2001, pp. 79–95.
[17] A. Hafid, S. Benouar, M. Kedir-Talha, F. Abtahi, M. Attari, and F. Seoane, "Full impedance cardiography measurement device using raspberry PI3 and system-on-chip biomedical instrumentation solutions," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 6, pp. 1883–1894, Nov. 2018.
[18] P. R. Rijnbeek, J. A. Kors, and M. Witsenburg, "Minimum bandwidth requirements for recording of pediatric electrocardiograms," *Circulation*, vol. 104, no. 25, pp. 3087–3090, 2001.
[19] J. W. Mason, E. W. Hancock, and L. S. Gettes, "Recommendations for the standardization and interpretation of the electrocardiogram," *Heart Rhythm*, vol. 4, no. 3, pp. 413–419, Mar. 2007.
[20] M. Singha Roy, B. Roy, R. Gupta, and K. Das Sharma, "On-device reliability assessment and prediction of missing photoplethysmographic data using deep neural networks," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 6, pp. 1323–1332, Dec. 2020.
[21] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, Mar. 1985.
[22] S. Banerjee, R. Gupta, and J. Saha, "Compression of multilead electrocardiogram using principal component analysis and machine learning approach," in *Proc. IEEE Appl. Signal Process. Conf. (ASPCON)*, Kolkata, India, Dec. 2018, pp. 23–27.
[23] E. Bølviken, P. J. Acklam, N. Christophersen, and J.-M. Størdal, "Monte Carlo filters for non-linear state estimation," *Automatica*, vol. 37, no. 2, pp. 177–183, Feb. 2001.
[24] S. Banerjee and G. K. Singh, "Quality aware compression of multilead electrocardiogram signal using 2-mode tucker decomposition and steganography," *Biomed. Signal Process. Control*, vol. 64, Feb. 2020, Art. no. 102230.
[25] S. Banerjee and G. K. Singh, "A new approach of ECG steganography and prediction using deep learning," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102151.
[26] C. S. Manohar and D. Roy, "Monte Carlo filters for identification of nonlinear structural dynamical systems," *Sadhana*, vol. 31, no. 4, pp. 399–427, Aug. 2006.
[27] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *J. Comput. Graph. Statist.*, vol. 5, no. 1, pp. 1–25, 1996.
[28] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 522–531, Feb. 2015.

Soumyendu Banerjee (Graduate Student Member, IEEE) received the B.Tech. degree in electrical engineering from the Government College of Engineering and Technology, Berhampore, India, in 2016, and the M.Tech. degree in electrical engineering from the University of Calcutta, Kolkata, India, in 2018.

He was an Assistant Professor at the Department of Electrical Engineering, University of Engineering and Management, Kolkata, from 2018 to 2019. He has been working as a Research Scholar at the Department of Electrical Engineering, IIT Roorkee, India, since 2019. His research interests include data steganography, digital signal compression, and noise cancellation.

Girish Kumar Singh received the B.Tech. degree in electrical engineering from G. B. Pant University of Agriculture and Technology, Pantnagar, India, in 1981, and the Ph.D. degree in electrical engineering from Banaras Hindu University, Varanasi, India, in 1991.

He worked in the industry for nearly five-and-a-half years. In 1991, he became a Lecturer at M. N. R. Engineering College, Allahabad, India. In 1996, he moved to the University of Roorkee, Roorkee, India. He is currently a Professor at the Department of Electrical Engineering, IIT Roorkee, Roorkee. He has been involved in the design and analysis of electrical machines in general and high-phase-order ac machines in particular as well as power system harmonics and power quality. He has coordinated a number of research projects sponsored by the Council of Scientific and Industrial Research (CSIR) and University Grants Commission (UGC), Government of India. He was a Visiting Associate Professor at the Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, South Korea, and a Visiting Professor at the Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara, Turkey.

Dr. Singh received the Pt. Madan Mohan Malaviya Memorial Medal and the Certificate of Merit Award 2001–2002 at The Institution of Engineers, India.