Automated Identification of Myocardial Infarction Using Harmonic Phase Distribution Pattern of ECG Data

Deboleena Sadhukhan, *Student Member, IEEE*, Saurabh Pal, and Madhuchhanda Mitra, *Member, IEEE*

***Abstract*— Incorporation of automated electrocardiogram (ECG) analysis techniques in home monitoring applications can ensure early detection of myocardial infarction (MI), thus reducing the risk of mortality. Most of the published techniques use advanced signal processing tools, a huge number of ECG features, and complex classifiers, which make their hardware implementation difficult. This paper proposes the use of harmonic phase distribution pattern of the ECG data for MI identification. The morphological and temporal changes of the ECG waveform caused by the presence of MI are reflected in the phase distribution pattern of the Fourier harmonics. Two discriminative features, clearly reflecting these variations, are identified for each of the three standard ECG leads (II, III, and V2). Classification of the healthy and MI data is performed using a threshold-based classification rule and logistic regression. The proposed technique has achieved an average detection accuracy of 95.6% with sensitivity and specificity of 96.5% and 92.7%, respectively, for classifying all types of MI data from the Physionet Physikalisch-Technische Bundesanstalt diagnostic ECG database. The robustness of the algorithm is confirmed with real data as well. The algorithm is also implemented and validated on a microcontroller-based Arduino board, which can serve as a prototype ECG analysis device. Apart from providing comparable performance to other reported techniques, the proposed technique provides distinct advantages in terms of computational simplicity of the features, significantly reduced feature dimension, and use of simple linear classifiers which ensure faster and easier MI identification.**

***Index Terms*— Classification algorithms, computer-aided analysis, discrete Fourier transforms (DFTs), electrocardiogram, harmonic analysis, medical expert systems, myocardial infarction (MI), phase detection.**

I. Introduction

MYOCARDIAL infarction (MI), more commonly known as heart attack, remains a primary health concern all over the world contributing to 15.53% of the total mortality rate in 2015 [1]. An electrocardiogram (ECG) is the most preferred tool for preliminary diagnosis of MI, due to its ease of acquisition using low cost and noninvasive techniques [2]. ECG is the recording of the electrical activity of the heart muscle fibers which causes the alternate contraction and relaxation of the heart chambers (atria and ventricles). An ECG wave, corresponding to one cardiac cycle, is characterized by the sequence of the P, QRS, and T waves. The P wave represents the atrial contraction, whereas, the QRS and T waves denote the ventricular contraction and relaxation, respectively [3]. Changes in the morphological and temporal wave features from a standard 12-lead ECG system are used by cardiologists to diagnose MI development in different regions of the heart.

The automated ECG analysis tools use different signal processing techniques to extract discriminating features from the digital ECG data to identify the presence of MI. Integration of such algorithms with the portable health care devices [4], [5] can incorporate the necessary diagnostic intelligence to ensure early detection. This, in turn, can expedite the treatment process and reduce the risk of mortality [6]. Such devices are imposed with some inherent limitations which include the following:

1) availability of limited number of ECG leads for analysis;
2) noise contamination of the ECG data;
3) limited computational ability of the embedded processors;
4) limited memory capacity.

Hence, the automated ECG analysis algorithms used for such devices must use fewer ECG leads and have to be less computationally intense to ensure easy implementation.

Although several recent publications [7]–[19] have focused on the development of computerized ECG analysis techniques for MI detection, a majority of them are not ideal for implementation in the portable devices mainly due to their high computational burden and memory requirement. Approaches based on the use of morphological ECG features (e.g., Q-wave, QT interval, ST elevation/depression, and T-wave) [7]–[11] rely on accurate identification of the ECG characteristic points. This is often difficult to achieve due to the wide pathological variation in the waveform morphologies and noise contamination of the data. A wide range of advanced signal processing tools such as the Wavelet transform has been applied to extract different frequency and time–frequency features [12]–[15] from the ECG data to improve the diagnostic ability. But the computation of such features is still difficult to implement in hardware platforms owing to their computational complexity. Moreover, the high-dimensional feature space [7], [9], [10], [13], [14], [15] and the use of advanced nonlinear classifiers [11], [15]–[17] also have an adverse effect on the computational performance and storage requirement. Most of the reported techniques aim at both detection and localization of MI. They, thus, utilize the entire 12-lead ECG data [7], [9], [10], [13], [15], which is difficult to record using

Manuscript received September 22, 2017; revised February 27, 2018; accepted March 4, 2018. The work of D. Sadhukhan was supported by the Department of Science and Technology, Government of India, through the DST INSPIRE Fellowship. The Associate Editor coordinating the review process was Dr. V. R. Singh. *(Corresponding author: Madhuchhanda Mitra.)*

The authors are with the Instrumentation Engineering Section, Department of Applied Physics, University of Calcutta, Kolkata 700073, India (e-mail: mmaphy@caluniv.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2018.2816458

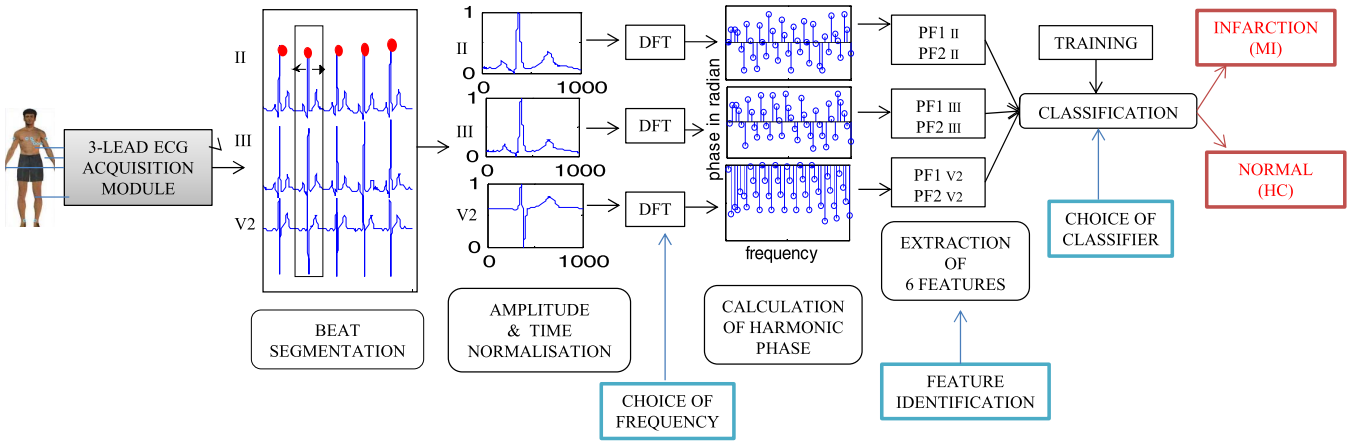


Fig. 1. Detailed block diagram of the proposed MI classification technique. The black boxes represent the processing steps, and the blue boxes represent the design considerations. Representative beats extracted from three selected ECG leads are normalized in amplitude and time. The relevant harmonic phase values are computed using DFT. The 2-D feature vector (PF1 and PF2) is extracted for each lead. Finally, linear decision boundary-based classifiers are used to identify the presence or the absence of MI.

the portable devices. The use of vector-cardiogram (VCG), recorded using three orthogonal leads, has also been proposed in [18] and [19] to utilize the spatiotemporal distribution of heart's electrical activity for MI detection. But the standard ECG monitors do not generally record these orthogonal VCG leads.

The main motivation behind this paper is to develop an automated ECG analysis algorithm and its hardware implementation to meet the requirements of the portable health care devices. For this purpose, we explore the utility of the harmonic phase distribution pattern of the ECG data. In the Fourier representation of a signal, the phase values are known to capture more information about the location and morphology of events than the harmonic magnitudes [20]. The presence of MI causes different morphological and temporal changes in ECG waveform [2], which results in significant alteration of the harmonic phase distribution. Mitra *et al.* [21] have pointed out few differences in the phase property of normal and infarction data but have not utilized them to develop a suitable classification technique. Earlier, we have proposed the preliminarily level utility of the Fourier harmonic phase distribution for identification on one type of MI using single-lead ECG data [22]. This paper extends the work to effectively utilize this feature space to develop a complete classification tool to identify the presence of all types of infarction. The key contributions of the proposed technique are as follows.

1) The use of harmonic phase values as features eliminates the need of extracting explicit time plane features. Moreover, the computation of the phase values using discrete Fourier transform (DFT) of ECG beats does not involve operations requiring high computational overhead (Section II-D).
2) Only two discriminating features reflecting the changes in the phase distribution pattern are identified from each ECG lead for MI classification (Section II-E). The reduced feature dimension ensures faster classification.
3) The high discriminability of the feature space enables the use of linear decision boundaries developed using heuristic threshold-based classification rules and logistic regression (LR) (Section II-F).
4) ECG data from only three standard leads (II, III, and V2) are used to detect the presence of MI with all possible localizations as obtained from the Physikalisch-Technische Bundesanstalt (PTB) diagnostic ECG database (Section II-B and Table IV).
5) The algorithm has been implemented and validated on an embedded platform with limited computational ability and memory, which can serve as a prototype ECG analysis device for MI monitoring (Section III-E).

The detailed performance evaluation of the proposed technique and its comparison with the other reported literature have been provided in Section III. Section IV elaborates the major considerations for the proposed approach. Finally, Section V concludes this paper highlighting the main contributions and future scopes.

II. Material and Methods

Fig. 1 shows the detailed block diagram of the proposed MI identification technique. The key concept is to compute the harmonic phase values of representative ECG beats from three chosen ECG leads and identify the discriminative changes to classify the presence of MI by the use of suitable classifiers. The processing steps including the design considerations are explained subsequently.

A. Database Used

The benchmark PTB diagnostic ECG database [23] available under Physionet [24] is used for the development and validation of the proposed technique. It is a collection of digitized 12-lead ECG records collected from healthy control volunteers (HC) and patients having different heart diseases at the Department of Cardiology, University Clinic Benjamin Franklin, Berlin, Germany. The data are sampled at 1000 Hz with 16-bit resolution. Among a total of 290 individual

TABLE I
LEAD-WISE ECG CHANGES CAUSED BY DIFFERENT MI TYPES

	I	II	III	aV_R	aV_L	aV_F	V1	V2	V3	V4	V5	V6
AWMI	#				#		*	*	*	*	*	*
ASWMI							*	*	*	#		
ALWMI	#				#					*	*	#
IWMI		*	*			*						
ILWMI		*	*			*					#	#
LWMI	*		*								#	#
PWMI							$	$				

'*' denotes ST elevation, deep Q, T wave inversion,
'#' denotes reciprocal changes like ST depression, symmetrical pointed T,
'$' denotes mirrored changes like symmetrical pointed T, elongated R wave

subjects, 52 are HC and 148 are MI patients contributing to 79 HC records and 369 MI records. Based on the affected heart region, the MI records are classified into 10 different types including the inferior wall (IWMI), inferior–lateral wall, inferior–posterior–lateral wall, anterior wall, anterior–lateral wall, anterior–septal wall (ASWMI), lateral wall (LWMI), posterior wall (PWMI), and posterior–lateral wall.

B. Choice of the ECG Leads

The 12-lead ECG system [25] reflects the electrical activity of the heart from different perspectives. The presence of infarction in different zones of the heart causes distinctive morphological changes in specific leads and reciprocal changes in the connected leads, as detailed in Table I.

In standard clinical assessment, the entire 12-lead ECG data are assessed to detect and localize MI [3]. However, the system contains highly redundant information as some leads are derived from the others. Moreover, for a portable health care setup recording of all the 12 leads involves high instrumentation cost and it also hinders the patient mobility and comfort. Earlier, Fayn [8] and Rubel *et al.* [26] have suggested the use of the reduced lead subset for MI detection. For this paper, the standard leads II, III, and V2 are chosen according to the suggestions from experienced cardiologists. Leads II and III reflect changes caused by all groups of inferior MI, lead III reflects reciprocal changes due to lateral MI. Changes due to posterior, anterior, and septal MI are reflected in lead V2 data. Hence, the combination of these leads is sufficient to reflect morphological changes caused by infarction in any of the heart regions.

C. Beat Segmentation and Normalization

Unlike other arrhythmic diseases, presence of infarction causes permanent damage to the heart muscles and all ECG beats are uniformly affected. Therefore, instead of beat-by-beat analysis, a single representative beat (comprising of the P, QRS, and T waves) from each ECG lead can be analyzed for MI identification. The location of the R peak of lead II is used for beat segmentation, as it is the most prominent feature. Due to the computational simplicity, the method proposed in [27] is employed to register the R peaks. The double difference of the ECG data array is used to localize the high-frequency QRS regions and then the peaks are detected using slope- and magnitude-based criteria.

According to the standard ECG beat morphology for nonarrhythmic cases [3], the T wave lies within the first 2/3rd and the P-wave within the next 1/3rd length of the R–R interval duration. The use of fixed window length around R peaks, as used in [8], [10], and [14], cannot ensure inclusion of the full cardiac cycle due to the variation in the heart rate. To account for it, the beats are extracted according to the following criteria based on the R-to-R duration so as to include the P-wave before a detected peak and the T wave after it:

$$\text{Start Point} = R_n - \frac{1}{3}[R_n - R_{n-1}] \tag{1}$$

$$\text{End point} = R_n + \frac{2}{3}[R_{n+1} - R_n] \tag{2}$$

where R_n denotes the current R peak location, while R_{n-1} and R_{n+1} denote the locations of the previous and the subsequent R peaks, respectively. These start and end points are used to extract time-aligned ECG beats from the three leads.

The frequency resolution of the DFT spectra depends on the data length. For 1-kHz sampling frequency, the extracted beat length varies from 600–950 data points due to variation in the heart rate. To ensure parity of spectral resolution between the analyzed beats, all extracted beats are time normalized to contain 1000 data points each by the use of fast Fourier transform (FFT)-based interpolation technique. To eliminate the effect of any added baseline, all the beats are amplitude normalized in the range of 0–1.

D. Computation of Harmonic Phase

Fourier transformation of a signal involves translation of a signal from the time domain to its frequency components [28] using constituent sinusoids of a fundamental frequency and its harmonics. General DFT of a signal uses complex exponential basis functions to generate the complex coefficient values which represent the harmonic magnitude and phase. However, for the extraction of the exact phase relationship and the ease of data handling, an ECG beat $x(n)$ of length N is expanded using the sine and cosine basis functions to represent the DFT as

$$x(n) = \sum_{k=1}^{N}\left[A_k \cos\left(\frac{2\pi \text{nk}}{N}\right) + B_k \sin\left(\frac{2\pi \text{nk}}{N}\right)\right] \tag{3}$$

where n is the time index and k is the coefficient index corresponding to the frequency harmonics.The frequency resolution of the harmonic components is given by F_S/N Hz, where F_S denotes the sampling frequency.

The coefficient values representing the magnitude of the harmonic components are computed using the following equations:

$$A_k = \frac{2}{N}\sum_{n=1}^{N}\left[x(n) \times \cos\left(\frac{2\pi nk}{N}\right)\right] \tag{4}$$

$$B_k = \frac{2}{N}\sum_{n=1}^{N}\left[x(n) \times \sin\left(\frac{2\pi nk}{N}\right)\right]. \tag{5}$$

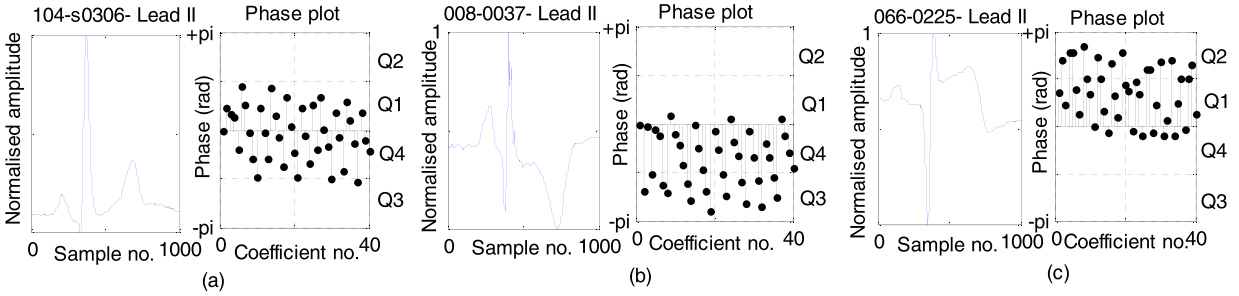


Fig. 2. Lead II ECG beats and their corresponding phase plots for HC and IWMI records. Only the first 40 phase values are shown, and different quadrants (Q1–Q4) are marked. (a) Healthy data (patient ID 104 record no. 306) having absolute normal morphology. The phase values are evenly distributed between the first and last quadrants, but majority of them lie in the range of $[-\pi/2, +\pi/2]$. (b) Inferior MI data (patient ID 008 record no. 037) showing deep Q and inverted T-wave. Majority of the phase values are negative, and significant numbers are less than $-\pi/2$. (c) Data (patient ID 066 record no. 225) showing significant ST elevation. Majority of the phases are positive, and significant numbers of them exceed the range of $+\pi/2$.

The phase of each harmonic component is then given by

$$\theta_k = \tan^{-1} \frac{B_k}{A_k}. \tag{6}$$

The relative phases of different harmonic components, computed by subtracting the fundamental harmonic phase from the others, capture the temporal information about the signal [20]. A translation in position (time or space) of a signal event will affect only the corresponding harmonic phase and not the magnitude.

Presence of infarction alters the morphological pattern of the ST segment, T-waves, and the QRS complexes of the ECG wave. The ST segment being nearly isoelectric corresponds to the low-frequency zone of the ECG spectra. The slowly varying T-wave corresponds to mid frequencies below 10 Hz, whereas, the fast varying QRS complex region comprises of high frequencies from 15–40 Hz [3], [13], [24]. For 1-kHz sampling frequency and fixed beat length of 1000, the frequency resolution of the obtained Fourier spectra is 1 Hz. Therefore, the first 40 harmonics are sufficient to capture the MI-induced ECG changes in the frequency band of 1–40 Hz. The choice of this frequency band also eliminates the effect of high frequency noises and power-line interference of 50 Hz, which can significantly corrupt the data in time plane.

The changes in ECG waveform morphology (ST elevation or depression, T-wave peaking or inversion, and deepening of Q waves) are clearly manifested as changes in the relative phases of the harmonic components. Figs. 2–4 depict the relative phase distribution pattern of lead II, lead III, and lead V2, respectively, for the HC and the prominent MI cases.

E. Identification and Extraction of Phase Features

To effectively utilize the phase variations for classification, it is required to identify discriminating parameters that depict the changes. The morphological changes induced in a particular ECG lead by all the types of MI which affect that lead are similar and hence cause similar changes in the phase distribution as well. Thus, for feature identification and classifier development it is sufficient to consider any one type of MI which is most prominent in that lead. Features for leads II and III are identified using HC and IWMI cases, while those for lead V2 are identified using ASWMI cases. It is evident from Figs. 2–4 that the MI-induced ECG morphological changes cause variation in distribution of the phases between different quadrants. Hence, a quadrant-based analysis of the phase values for each of the three leads is performed. The number of phases in each quadrant P1 (no. of phases in Q1 $[0, \pi/2]$), P2 (no. of phases in Q2 $[\pi/2, \pi]$), P3 (no. of phases in Q3 $[-\pi/2, -\pi]$), P4 (no. of phases in Q4 $[-\pi/2, 0]$), and their combinations are evaluated as features. Visual inspection of the phase plots of lead II and lead III, as illustrated in Figs. 2 and 3, for the HC and IWMI cases displays the following distinctive changes.

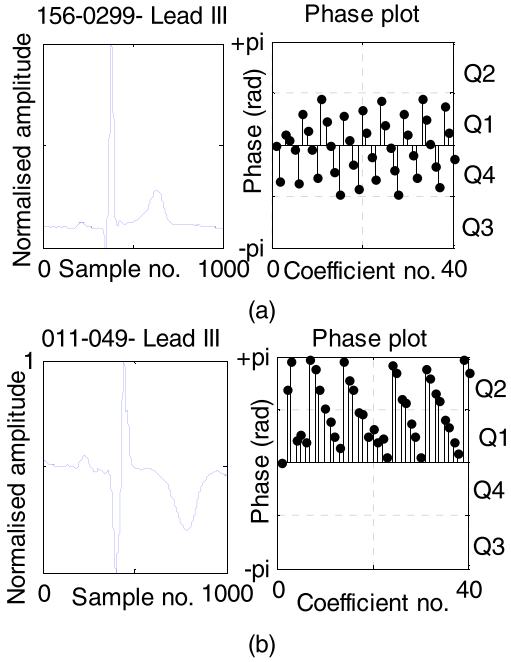


Fig. 3. Lead III ECG beats and their corresponding phase plots for HC and IWMI records. (a) Healthy data (patient ID 156 record no. 299). The phase values are evenly distributed between the first and last quadrants, and majority of them lie in the range of $[-\pi/2, +\pi/2]$. (b) Inferior MI data (patient ID 011 record no. 049) showing deep Q and inverted T-wave. All phase values are positive, and significant numbers are greater than $+\pi/2$.

1) For the HC beats, the harmonic phases are evenly distributed between the first two (positive values) and last two quadrants (negative values). However, for the infarction data, majority of the phases are either positive or negative.
2) Also for the HC beats, most of the phases are contained in the quadrants Q1 and Q4, i.e., in the range of $[-\pi/2, +\pi/2]$. For the MI cases, a significant number of phases lie outside this range.

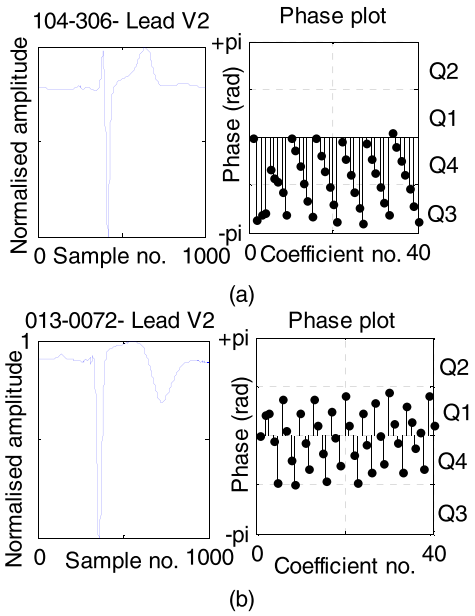


Fig. 4. Lead V2 ECG beats and their corresponding phase plots for HC and ASWMI records. (a) Healthy data (patient ID 104 record no.306) showing absolute normal V2 morphology. Most of the phases are negative, and significant numbers of them are lower than $-\pi/2$. (b) Antero-septal MI data (patient ID 013 record no. 072) showing ST elevation and T-wave inversion. The phase values are evenly distributed between the positive and negative quadrants, but all of them are restricted in the range of $[-\pi/2, +\pi/2]$.

Waveform morphology of lead V2 data is significantly different from that leads II or III. Visual analysis of the phase distribution pattern for lead V2 data reveals that majority of the phases of HC beats are clustered either in Q1 and Q2 or Q3 and Q4, but more evenly distributed inside and outside the range of $[-\pi/2, +\pi/2]$. The presence of MI causes distinct deviation from this pattern, as shown in Fig. 4.

In accordance with these observations, two discriminative features are identified for each of the three leads.

1 Absolute difference between the number of positive and negative phases (PF1II, PF1III, and PF1V2)

$$\mathrm{PF1} = \mathrm{abs}[(P1 + P2) - (P3 + P4)]. \tag{7}$$

2 Difference between the number of phases in and outside the range of $[-\pi/2, +\pi/]$ ($\mathrm{PF2_{II}}$, $\mathrm{PF2_{III}}$, and $\mathrm{PF2_{V2}}$)

$$\mathrm{PF2} = (P1 + P4) - (P2 + P3). \tag{8}$$

The p values of the features are evaluated using the student T-test [29] to prove the existence of distinct distribution of the feature values between the HC and MI classes. Low "p" values of the order of 10^{-5} ensure more prominent separation. Table II tabulates the mean, standard deviation (SD), and the p values of the extracted features evaluated from 500 representative beats each from the HC and IWMI records (for leads II and III) and ASWMI records (for lead V2).

For lead V2, the presence of MI affects the ST segment and the T-waves more prominently than the QRS zone. Accordingly, the changes in the phase distribution pattern should be more prominent for the first 20 harmonics only. The mean, SD, and the p values of lead V2 phase features evaluated for both 20 and 40 harmonics are displayed in Table II. It can be seen that the deviation in the mean value of the features is more prominent for the 20 harmonics and it generates lower p value as well. Hence, for classification purpose phase features of lead V2 are evaluated for first 20 harmonics.

TABLE II
EXTRACTED FEATURE VALUES FOR HC AND MI CASES

		Healthy(HC)		Infarction(MI)		p value
		Mean	SD	Mean	SD	
Lead II	$\mathbf{PF1_{II}}$	8.7	5.1	26.9	13.2	6.7×10^{-18}
	$\mathbf{PF2_{II}}$	29.1	5.6	7.8	7.9	8.0×10^{-12}
Lead III	$\mathbf{PF1_{III}}$	10.2	4.9	26.5	12.7	3.3×10^{-14}
	$\mathbf{PF2_{III}}$	29.4	5.3	8.3	7.6	5.5×10^{-15}
Lead V2	$\mathbf{PF1_{V2}(40)}$	26.9	8.6	19.8	5.3	5.3×10^{-3}
	$\mathbf{PF2_{V2}(40)}$	12.8	15.9	20.7	9.6	6.9×10^{-5}
	$\mathbf{PF1_{V2}(20)}$	17.9	3.5	8.5	3.6	1.1×10^{-8}
	$\mathbf{PF2_{V2}(20)}$	2.8	3.6	11.5	4.5	2.3×10^{-12}

Fig. 5(a)–(c) shows the 2-D scatter plots for visualization of the intraclass variation of the features for HC and MI data for the three leads. Table II and Fig. 5 show that the two features extracted from each of the three leads are adequately discriminatory for classification of HC and MI cases.

F. Classification

Classification involves designing the decision boundary for separating the two classes of HC and MI. Two types of classification techniques are applied as follows.

1) Linear Decision Boundary Using Logistic Regression: Fig. 5 clearly depicts the linear separability between the classes for all three leads. Hence, any basic linear classifier aimed at modeling the linear decision boundary can be employed. Both linear discriminant analysis (LDA) and LR [30] provide the advantages of easy implementation, less training time, and computational simplicity. But, LDA makes an assumption that the data from the two classes are normally distributed with equal variance [31], which is inappropriate for the current application. Hence, LR proves to be a more suitable and robust option. Binary LR classifier models the probability that a data instance defined by the feature set x $[x1, x2]$ will belong to a particular class (0 or 1) by the use of sigmoid function as follows:

$$h(\mathbf{x}) = \frac{1}{1 - e^{-(a_0 + a_1 x1 + a_2 x2)}}. \tag{9}$$

A data instance is classified to a particular class (1 for MI or 0 for HC) if the corresponding probability is greater than a predefined threshold. The model parameters (a_0, a_1, a_2) are determined by the use of Maximum-likelihood estimation [30] from a set of training data.

2) Threshold-Based Classifier: The distinct separation between the feature values of the HC and MI classes enables the use of threshold-based classification which is much easier to implement and takes lesser time for training and classification than that of LR (proved in Section III-C). The threshold values of the features for each of the three leads are estimated to be the boundary feature values so as to give 99% coverage to the MI class.

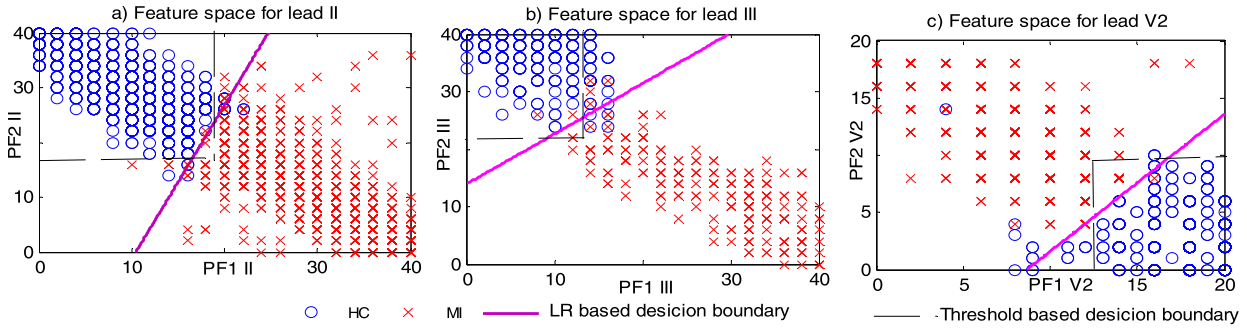


Fig. 5. 2-D scatter plots of the feature space for 500 beats each from HC, IWMI, and ASWMI cases, and the modeled classifier boundaries for each lead. The MI subjects are visualized with the red crosses and HC with the blue circles. Magenta solid lines show the linear decision boundary modeled by LR. Black dotted lines show the selected threshold feature values. (a) 2-D scatter plot of features $PF1_{II}$ and $PF2_{II}$ extracted from lead II ECG beats for healthy and inferior MI records. (b) 2-D scatter plot of lead III features ($PF1_{III}$ and $PF2_{III}$) for healthy and inferior MI records. (c) 2-D scatter plot of lead V2 features ($PF1_{V2}$ and $PF2_{V2}$) for healthy and antero-septal MI records.

TABLE III
RESULTS OF CLASSIFIER MODELING

Lead	True class	Logistic Regression Classifier						Threshold Based Classification					
		Predicted Class		Acc (%)	Se (%)	Sp (%)	Parameters (a_0,a_1,a_2)	Predicted Class		Acc (%)	Se (%)	Sp (%)	Threshold ($PF1_{TH}$, $PF2_{TH}$)
		MI	HC					MI	HC				
II	MI (IWMI)	489	11	98.6	97.8	99.4	-17.6, 1.7, -0.6	496	3	99.1	99.2	99	19, 17
	HC	3	497					5	495				
III	MI (IWMI)	487	13	98.1	97.4	98.8	6.6,0.4, -0.5	495	5	97.9	99.0	96.8	13,22
	HC	6	494					16	484				
V2	MI (ASWMI)	495	5	98.2	99.0	97.4	13.0, -1.5, 1.3	493	7	97.5	98.6	96.4	13,10
	HC	13	487					18	482				

The classification rule for the individual leads are then designed in the form

$$\text{“If } [x1 \geq (or \leq)\, PF1_{TH}] \text{ AND } [x2 \leq (\text{or} \geq)\, PF2_{TH}] \text{ then class} = 0(HC), \text{else class} = 1(MI)''$$

where $x1$ and $x2$ denote the two features for the individual leads and $PF1_{TH}$ and $PF2_{TH}$ denote the corresponding threshold values for the two features.

As the MI-induced changes are manifested only in specific ECG leads depending on the location of infarction, all three leads need to be analyzed to detect the presence of all kinds of MI. Three separate classifiers are developed for each of leads II, III, and V2 using the feature sets ($PF1_{II}$ and $PF2_{II}$), ($PF1_{III}$ and $PF2_{III}$), and ($PF1_{V2}$ and $PF2_{V2}$), respectively. A record is classified as MI if any one or more of the lead data are detected as MI.

III. Experimental Results and Validation

The three different classifiers for each lead are modeled and tested using a specific data set. The predesigned decision boundaries are then employed to classify the other records to assess the overall performance and generalizability. Performance evaluation is done using the standard statistical parameters—accuracy (Acc), sensitivity (Se), and specificity (Sp) as defined in [32].

A. Results of Classifier Modeling

Classifiers for leads II and III are trained using HC and IWMI records and that for lead V2 using HC and ASWMI records. ECG morphology shows a wide variation among the HC variant and may sometimes exhibit some pathological features such as inverted T-wave [3]. Among the 79 HC records, 50 records having absolute normal ECG features are chosen for training the classifiers in accordance with cardiologist's opinion. Among the MI records in the data set, 89 were IWMI, and 78 ASWMI. Although it is sufficient to use single representative beat from each record for MI analysis, to avoid classifier bias problem (due to shortage of data) and class-imbalance problem (due to unequal training data from each class), 500 representative beats from 50 records of each of the variant are used for training the classifiers. To obtain generalized classification boundaries for each lead, a fivefold cross-validation technique is adopted. The entire data set of 1000 instances (500 HC and 500 MI) is partitioned into five disjoint and equal sets. Among these, four sets are used to generate the classifier model and the remaining one set is used to test it. The process is repeated for 5 times to cover all the sets as test set. The data partitioning is manually modulated so that beats from the same record are not used both in the training and test sets, which may have led to model over fitting. The average model parameters obtained from the five test folds are used as the final decision boundary for classifying all other types of MI. Fig. 5 shows the final modeled classifier boundaries for the three leads both using threshold selection rule and LR. Table III tabulates the average performance measures and the final confusion matrix of the fivefold validation obtained using both the classifier models. The obtained model parameters are also shown.

TABLE IV
CLASSIFICATION PERFORMANCE TESTED WITH ENTIRE DATASET

Data group	No. of Records	LR classifier		Threshold based	
		Mis-classification	Acc (%)	Mis-classification	Acc (%)
IWMI	85	4	95.2	6	92.9
ILWMI	54	2	96.2	2	96.2
LWMI	2	1	50.0	2	0.0
PWMI	4	2	50.0	1	75.0
PLWMI	5	1	80.0	1	80.0
AWMI	43	3	93.1	6	86.0
ALWMI	40	3	92.5	6	85.0
ASWMI	75	6	92.0	10	86.6
HC	65	5	92.3	8	87.6

The threshold values for the threshold-based classification are selected giving 99% coverage to the MI data points. As a consequence, it resulted in higher detection sensitivity but lower specificity and overall accuracy.

B. Classification Performance Evaluation

In order to establish the robustness of the classifier models, it must be tested with unseen data. The modeled decision boundaries are employed to classify records from the entire data set including all other types of MI. Classification results of the three leads are combined to designate the final class. MI-induced morphological changes are equally visible in all ECG beats in a record and hence, cause no significant change in the phase features as well. However, to reduce the effect of sudden noise contamination, the average feature values extracted from five randomly selected beats at suitable time intervals (depending on the data length) are used as the generalized feature values for each record. For the HC, IWMI, and the ASWMI records, beats used in the training phase are excluded. 14 HC and 30 MI records with significant discrepancies in the time plane waveform morphologies are eliminated from the study based on cardiologist's opinion. Table IV tabulates the classification performance for all different types of MI and HC records.

The algorithm achieves a better detection accuracy of nearly 95% for identification of all the inferior groups of MI. Detection accuracy falls to 92% for the anterior groups due to the consideration of only one anterior lead V2. For the PWMI and LWMI classes, the detection accuracy is restricted to only 50%. However, the numbers of MI records available in these groups are insufficient to properly justify the performance. The detection accuracy for the HC actually provides the measure of specificity.

To assess the overall performance and enable a fair comparison with other reported techniques, the algorithm is tested with 20 000 randomly selected beats from the database (5000 HC beats and 15 000 MI beats) and the results are displayed in Table V. A fairly high detection sensitivity of 96.5% is achieved with a specificity of 92%. The receiver operating characteristic curves and the area under the curves (AUCs) are shown in Fig. 6. The misclassified records are visually cross-verified by trained medical practitioners as well. The misclassified MI records mainly correspond to the data recorded from patients at their recovery stage, when MI-induced morphological changes are less distinct. This further justifies the performance of the proposed technique.

TABLE V
OVERALL CLASSIFICATION PERFORMANCE (BEAT WISE)

Classification method	True class	Predicted Class		Acc (%)	Se (%)	Sp (%)
		MI	HC			
Logistic Regression	MI	14485	515	95.6	96.5	92.7
	HC	361	4639			
Threshold Based	MI	14293	707	91.1	93.9	89.9
	HC	502	4498			

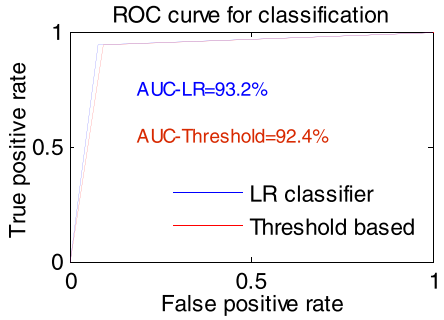


Fig. 6. Receiver operating characteristic curve for classifier results shown in Table V. The AUCs are marked.

To verify the robustness of the proposed technique with respect to variation in the data set and clinical setting, the algorithm is also tested with real lead II ECG records collected using BIOPAC MP-45 system at a sampling frequency of 1 kHz. Due to limitations of the data acquisition hardware, all the three leads of the ECG could not be acquired. Hence, only one type of MI is considered for the present evaluation. Data are collected from 30 healthy subjects in the age group of 22–55 with no known cardiac disorders in the biomedical laboratory of the Department of Applied Physics, University of Calcutta.

For the MI records, data are collected from 30 patients (in the age group of 35–65) admitted to the cardiac unit of the Calcutta Medical College and Hospital, Kolkata, India, reported with acute IWMI. Written informed consent has been obtained from all subjects in accordance with institutional policy and operational principles of the Helsinki Statement. Table VI tabulates the test results with the 60 lead II ECG records. The proposed technique achieves a high detection sensitivity of 96.7% and 93.3% using the LR classifier and the threshold-based classifier, respectively.

C. Computational Complexity Analysis

The computational overhead and execution time are important parameters for the implementation of algorithms in portable devices. The proposed algorithm consists of four distinct subparts—R peak identification and beat segmentation, interpolation of beat length, phase feature extraction, and finally classification.

The R peak detection algorithm [27] executed on a data length of M includes—double differencing which requires

TABLE VI
CLASSIFICATION PERFORMANCE ON REAL DATA (LEAD II)

Classification method	True class	Predicted Class		Acc (%)	Se (%)	Sp (%)
		MI	HC			
Logistic Regression	MI	29	1	95.0	96.7	93.3
	HC	2	28			
Threshold Based	MI	28	2	90.0	93.3	86.6
	HC	4	26			

M number of computations, sorting which has an overall time complexity $\mathbf{O}$ (M log M) and finally peak detection which also has a time complexity $\mathbf{O}$ (M log M). Therefore, the overall complexity of the peak detection phase is $\mathbf{O}$ (M log M). If the sample number between two consecutive R peaks is D, the beat extraction process will have a time complexity $\mathbf{O}$ (D). The FFT interpolation technique used for beat length normalization has a time complexity $\mathbf{O}$ (D log D). The phase feature extraction part includes the implementation of DFT algorithm which has a complexity of $\mathbf{O}(N^2)$ for a signal of length N. But for this paper, all beats have a fixed length of 1000 and only 40 coefficients of each beat are computed for each lead. This involves ($3 \times 2 \times 1000 \times 40$) computations for coefficient computation, (3×40) divisions for computation of phase values, ($3 \times 4 \times 40$) computations for the quadrant-based analysis of the phase values, and finally (3×6) computations for the extraction of the two-phase features per lead. Hence, the time complexity of the phase feature extraction is constant, i.e., $\mathbf{O}(1)$.

Classification of a sample using LR with two features for each of the three leads involves (3×2) multiplications, (3×2) additions, three subtractions, and three divisions. The exponential power is evaluated from a lookup table. Hence, time complexity of the LR classification phase is also constant. The threshold-based classification involves only (3×2) comparisons, which is again a constant. Hence, the overall time complexity of the algorithm to label a new record is $\mathbf{O}$ (M log M) + $\mathbf{O}$ (D) + $\mathbf{O}$ (D log D)+ $\mathbf{O}$ (1)+ $\mathbf{O}$ (1)$\sim$$\mathbf{O}$ (M log M)**,** because $D \ll M$. Thus, considering small data segments at the time of beat extraction can significantly reduce the computation time. Time complexity of the training phase of the LR classifier depends on the convergence of the optimization algorithm and generally varies linearly with the size of the training set [33].

The algorithm is developed using MATLAB 2007 software on a desktop with Intel Celeron processor (2.4 GHz) and 4-GB RAM. The average time for training the LR classifiers (using the 1000 training beats for each classifier) is 0.062 s. The computation time for classification of an ECG record is 0.95 s (beat extraction time = 0.93 s, time for feature extraction from the three leads = 0.0129 s, and final classification time = 0.01163 s).

Though the performance of the threshold-based classification is lower as compared to the LR classifier, it has few inherent advantages. For instance, it employs median-order statistics to find the threshold values which requires lower computation time (0.01 s) than the optimization algorithms used to determine the LR classifier parameters (0.06 s). For classifying a new sample, the threshold-based classifier involves only ($3 \times 2 = 6$) number of comparisons, whereas, the LR classifier requires ($3 \times 4 = 12$) number of computations for solving the linear equations. It also needs to refer to the lookup table, which further increases the prediction time. Hence, the threshold-based technique is easier and faster to implement.

D. Performance Comparison

In order to verify the applicability of any MI identification technique for the portable health care devices, performance comparison should consider not only the performance metrices, but also the other important parameters, such as the number of leads used, number of features, complexity of the features and classifiers. Table VII gives a detailed performance comparison of the proposed technique with the other recently reported literature. The relative advantages and disadvantages are also highlighted. The proposed algorithm outperforms some of the reported techniques and provides comparable performance to most of the others.

The use of conventional time plane features used in [7]–[11] needs explicit identification of the characteristic waves which is difficult to achieve in noise contaminated data. These algorithms, thus, rely on initial data preprocessing for noise removal, which adds to the computational burden. However, the features used for the proposed MI identification technique involves the phase distribution of only the first 40 harmonics. The presence of high frequency noises, including power-line interference, does not alter these features. This eliminates the need for initial data denoising. The implementation of Rough set for rule mining using the 144 ECG features from all 12 leads by Mitra *et al.* [7] requires a significantly high execution time of nearly 98 s (when implemented on the same workstation used for the proposed technique). The cross-wavelet transform-based features [12] rely on implementation of continuous wavelet transform, which requires an execution time of 0.17 s for a single-lead ECG signal of length 1000 samples. Execution of a six level DWT decomposition on the same data length used for the extraction of the subband-based features in [13]–[15] requires 0.035 s. Extraction of the features from the obtained wavelet coefficients further increases the computation time. However, the execution of DFT used in this paper for computing the phase features from the first 40 harmonics of a single-lead ECG beat of length 1000 samples requires a considerably low time of 0.0094 s. This suggests the computational simplicity of the proposed phase-based features.

The k-nearest neighbor (kNN) classifier used in [9], [13], and [14] needs to store all the training data which requires significant amount of memory depending on the training data size [30]. Whereas, for the LR classifier used in the proposed technique, only the model parameters ($3 \times 3 = 9$) need to be stored. Also, for classification of a sample kNN needs to compute "k" number distance matrices which increases the prediction time (typically of the order of 0.1 s for two features using a minimum "k" value of 5). The use of nonlinear decision boundaries modeled using neural network [11]

TABLE VII
DETAILED PERFORMANCE COMPARISON WITH OTHER REPORTED METHODS

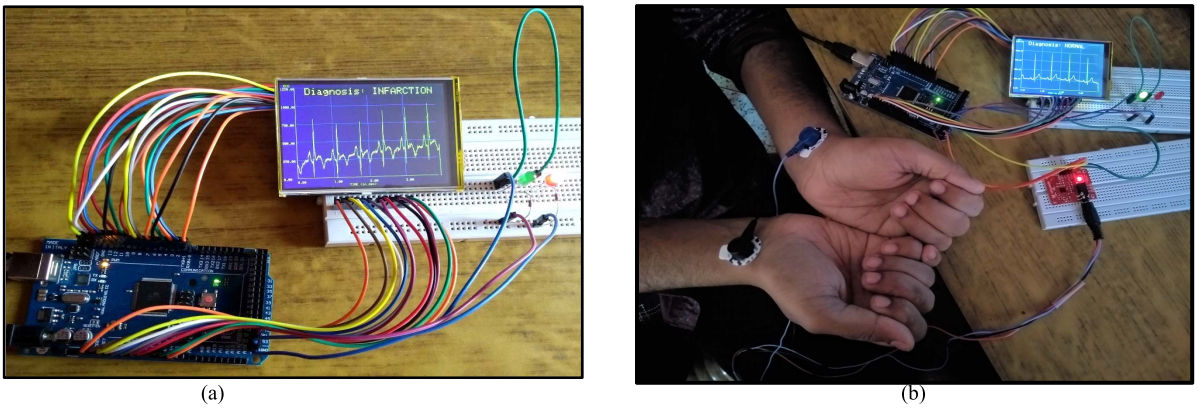
Ref. (Year)	Dataset	No. of leads	Features used		Classifier	Localisation	Performance		
			Number	Type			Acc (%)	Se (%)	Sp (%)
[7] (2007)	Unknown (200 records)	12	12/lead 144 (Total)	Time plane *(noise sensitive)*	Rough set *(high training time)*	No	-	95.8	*100*
[8] (2011)	Unknown (180 records)	3	5 /lead 15 (Total)	Time plane *(noise sensitive)*	Classification tree	No	97.9	97.8	98
[9] (2012)	PTB database	12	36	Time plane features extracted using DWT	KNN *(high memory and classification time)*	Yes	-	99.97	99.9
[10] (2012)	PTB database	12	6 /lead 72 (total)	ST segment polynomial features *(high computational overhead)*	latent topic multiple instance learning *(high computational overhead)*	Yes	-	92.5	89.1
[16] (2012)	Unknown (2 MI types)	4	4	Hidden Markov model and log likelihood	Gaussian mixture model and SVM	No	82.50	85.71	79.82
[11] (2014)	PTB database	1	2	Total integral , T wave integral *(noise sensitive)*	Probabilistic Neural net *(high training and classification time)*	Yes (few types)	94	-	-
[12] (2014)	PTB database (one MI type)	3	2 /lead 6 (total)	Cross WT, cross spectrum coherence *(high computational burden)*	heuristically determined mathematical formula *(easy to implement , lacks robustness)*	No	97.6	97.3	*98.8*
[13] (2015)	PTB database	12	5 /lead 60 (total)	Energies and eigen values of multi-scale DWT *(high computational overhead)*	KNN *(memory requirement)* Nonlinear SVM *(high training time)*	Yes	96	93	99
[14] (2016)	PTB database (partial)	1	47	Non-linear DWT based	KNN	Yes	98.8	99.45	*96.3*
[15] (2017)	PTB database	12	35	SVD and wavelet energy *(high computational overhead)*	SVM	Yes	95.30	94.6	96.0
Proposed	PTB database (all MI types)	3	2/lead 6 (total)	Phase of DFT *(lesser computation)*	LR Threshold based *(easier implementation, less time)*	No	95.6 91.1	96.5 93.6	92.7 89.9

or support vector machine [13], [16], [15] require greater number of parameters compared to that of linear ones [30]. This increases the training time due to the execution of greater number of optimization algorithms (one for each parameter). With the increase in number of feature dimension, both the training time and prediction time increases. The use of linear decision boundary modeled with LR using two features used in the proposed technique considerably reduces the training time, classification time, and also the memory requirement.

E. Hardware Implementation and Validation

The algorithm is implemented on the commercially available Arduino Mega 2560 board based on ATmega2560 microcontroller with a processing speed of 16 MHz. Its low power consumption, simple and fully static design along with its small size and cost-effectiveness make it highly suitable for the portable home monitoring devices. The board includes 256 kB of flash memory for code storage, 8 kB of SRAM and 4 kB of EEPROM, and feeds on 5-V supply. A 3.5-in thin-film-transistor liquid-crystal display is interfaced with a board to display a single-lead ECG record (lead II) and also the results of classification. Two light emitting diodes (LEDs) are also interfaced to provide a quick visual display. The green LED indicates an HC case, and red LED indicates MI. Fig. 7 shows the hardware setup. The firmware for the proposed algorithm is developed in the Arduino Software Integrated Development Environment (IDE) written on Java platform. The firmware uses the pretrained LR classifier for final classification.

For performance evaluation, the firmware is tested with the 308 MI records and 65 HC records from the PTB diagnostic ECG database that has been used for testing the software version of the proposed technique, as described in Table IV. ECG data of 4-s duration from the three leads (II, III, and V2) of each record are down-sampled to 250 Hz and loaded to the internal RAM of the Arduino board. The data duration and sampling rate are chosen to optimize the computational overload and memory requirement. These data are then processed by the firmware of the Arduino board to generate the classification output for each record. The 291 MI records and 59 HC records are correctly identified while six HC records were misclassified as MI. Hence, the firmware achieves an overall



(a) (b)

Fig. 7. Hardware setup with Arduino Mega 2560 board and 3.5-in TFT-LCD display. (a) System is displaying result for processing an MI record from PTB database patient ID 008 record no. 036. The monitor displays the lead II ECG data and the result of diagnosis as "Diagnosis- INFARCTION." The red LED also glows to indicate MI diagnosis. (b) Hardware setup with interfaced AD8232-based ECG acquisition board. The sensor lead cables are connected to the board using a 3.5-mm jack. The on-board red LED indicates that the acquisition board is powered ON. The system is acquiring and processing data from a healthy volunteer. The monitor displays the lead II ECG data and the result of diagnosis as "Diagnosis-NORMAL." The green LED of the system glows to indicate HC.

detection accuracy of 93.3% with sensitivity and specificity of 94.5% and 90.7%, respectively. The processing time to generate the final classification output is nearly 55 s. Fig. 7(a) shows the system operation for an MI record taken from the PTB database.

For validating the system with real data, the Arduino board is interfaced with an AD8232 [34]-based integrated board for single-lead ECG recording. It is a cost-effective, fully integrated signal conditioning block for single-lead ECG acquisition. Lead II ECG signals from 20 healthy volunteers from the department and laboratory and 12 IWMI patients from the Cardiac Unit of the Calcutta Medical College and Hospital, Kolkata, India, are acquired by the system at 250-Hz sampling rate for 4-s duration and processed with the firmware after seeking written informed consent from them in accordance with the institutional policy and operational principles of the Helsinki Statement. Data acquisition from the IWMI patients was also approved by the ethics committee of the hospital and was carried on in close vigilance of the trained medical practitioners. The developed system correctly identified 12 MI records and 18 HC records, thus, achieving a detection accuracy of 93.75%, detection sensitivity of 100%, and specificity of 85.7%. Fig. 7(b) shows the system interfaced with the AD8232-based ECG acquisition board while data processing from a healthy volunteer. These results are quiet in accordance with the ones obtained by the software simulation.

IV. Discussion

For the proposed technique, localization (identification of the zone of infarction) has not been considered as it would require a greater number of leads to be analyzed. The use of reduced lead set limits the detection accuracy for some types of MI specially the anterior group. The choice of the classification technique is governed by the need for low computational overhead and memory requirement [32] for its implementation in portable devices. However, use of advanced classification techniques can be considered to improve the detection accuracy at the cost of increased computation time.

All the experimentation involving the development and validation phases are performed with expert suggestion from trained medical practitioners. The algorithm is tested on records having prevalent MI. Hence, only single representative beat from each record was considered for analysis. For real implementation in portable health care devices, recorded ECG beats at suitable intervals need to be analyzed to detect the onset of MI.

V. Conclusion

The phase values of the Fourier harmonics carry significant information about the waveform morphology and timings. MI-induced morphological changes in the ECG waveform are clearly reflected in the harmonic phase distribution pattern, but are largely unexplored till the date for MI identification. This paper utilizes this feature space for generating an automated ECG analysis technique for identification of MI, which is primary health concern all over the world. The proposed technique may not outperform all other earlier reported techniques, but provides significant advantages as-computational simplicity of the features, noise robustness, reduced feature dimension, and use of simple linear decision boundary-based classifier. All of these enable easier implementation and faster identification. Moreover, only three ECG leads are employed for identification of infarction in different heart regions. The algorithm performs equally well on the embedded platform with limited processing speed and memory capacity. Moreover, the developed system along with the ECG acquisition module is not only compact in size, but also uses low power and is cost effective. These advantages justify the utility of the proposed system for the development of low cost and portable ECG analysis device to ensure faster and reliable MI diagnosis.

This paper is focused on extending the system to develop a prototype hand-held device with multilead ECG acquisition for identification of different cardiac abnormalities. The diagnostic system can also be incorporated with the state-of-the-art smart phones. During the experimentations, study of the harmonic phase values of different ECG beats from single subjects revealed clear similarity in the distribution pattern. This finding

provides scope for further research in exploring the utility of this feature space for biometric identification.

ACKNOWLEDGMENT

The authors would like to thank Dr. S. Guha, MD, DM (Professor and HOD), and other members from the Department of Cardiology, Medical College and Hospital, Kolkata, India, for their valuable clinical feedback and assistance.

REFERENCES

[1] WHO Fact Sheet. (Jan. 2017). *The Top Ten Causes of Death, Fact Sheet-310*. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs310/en/

[2] K. Thygesen *et al.*, "Third universal definition of myocardial infarction," *Eur. Heart J.*, vol. 33, no. 20, pp. 2551–2567, 2012.

[3] C. C. Schamroth, *An Introduction to Electrocardiography*, 7th ed. New York, NY, USA: Wiley, 2009.

[4] J. Fayn and P. Rubel, "Toward a personal health society in cardiology," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 401–409, Mar. 2010.

[5] C. De Capua, A. Meduri, and R. Morello, "A smart ECG measurement system based on Web-service-oriented architecture for telemedicine applications," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2530–2538, Oct. 2010.

[6] R. V. Luepker, "Delay in acute myocardial infarction: Why don't they come to the hospital more quickly and what can we do to reduce delay?" *Amer. Heart J.*, vol. 150, no. 3, pp. 368–370, 2005.

[7] S. Mitra, M. Mitra, and B. B. Chaudhuri, "A rough-set-based inference engine for ECG classification," *IEEE Trans. Instrum. Meas*, vol. 55, no. 6, pp. 2198–2206, Dec. 2006.

[8] J. Fayn, "A classification tree approach for cardiac ischemia detection using spatiotemporal information from three standard ECG leads," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 95–102, Jan. 2011.

[9] M. Arif, I. A. Malagore, and F. A. Afsar, "Detection and localization of myocardial infarction using k-nearest neighbor classifier," *J. Med. Syst.*, vol. 36, no. 1, pp. 279–289, 2012.

[10] L. Sun, Y. Lu, K. Yang, and S. Li, "ECG analysis using multiple instance learning for myocardial infarction detection," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3348–3356, Dec. 2012.

[11] N. Safdarian, N. J. Dabanloo, and G. Attarodi, "A new pattern recognition method for detection and localization of myocardial infarction using T-wave integral and total integral as extracted features from one cycle of ECG signal," *J. Biomed. Sci. Eng.*, vol. 7, no. 10, pp. 818–824, 2014.

[12] S. Banerjee and M. Mitra, "Application of cross wavelet transform for ECG pattern analysis and classification," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 326–333, Feb. 2014.

[13] L. N. Sharma, R. K. Tripathy, and S. Dandapat, "Multiscale energy and eigenspace approach to detection and localization of myocardial infarction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1827–1837, Jul. 2015.

[14] U. R. Acharya *et al.*, "Automated detection and localization of myocardial infarction using electrocardiogram: A comparative study of different leads," *Knowl. Based Syst.*, vol. 99, pp. 146–156, May 2016.

[15] S. Padhy and S. Dandapat, "Third-order tensor based analysis of multilead ECG for classification of myocardial infarction," *Biomed. Signal Process. Control*, vol. 31, pp. 71–78, Jan. 2017.

[16] P.-C. Chang, J.-J. Lin, J.-C. Hsieh, and J. Weng, "Myocardial infarction classification with multi-lead ECG using hidden Markov models and Gaussian mixture models," *Appl. Soft Comput.*, vol. 12, no. 10, pp. 3165–3175, Oct. 2012.

[17] B. Liu *et al.*, "A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection," *Comput. Biol. Med.*, vol. 61, pp. 178–184, Jun. 2015.

[18] H. Yang, S. T. Bukkapatnam, T. Le, and R. Komanduri, "Identification of myocardial infarction (MI) using spatio-temporal heart dynamics," *Med. Eng. Phys.*, vol. 34, no. 4, pp. 485–497, 2012.

[19] T. Q. Le, S. T. S. Bukkapatnam, B. A. Benjamin, B. A. Wilkins, and R. Komanduri, "Topology and random-walk network representation of cardiac dynamics for localization of myocardial infarction," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 8, pp. 2325–2331, Aug. 2013.

[20] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, May 1981.

[21] S. Mitra, M. Mitra, and B. B. Chaudhuri, "Frequency-plane analysis of normal and pathological ECG signals for disease identification," *J. Med. Eng. Technol.*, vol. 29, no. 5, pp. 219–227, Sep. 2005.

[22] D. Sadhukhan, S. Pal, and M. Mitra, "Automated ECG analysis using-Fourier harmonic phase," in *Proc. IEEE Region 10 Symp. (TENSYMP)*, Cochin, India, Jul. 2017, pp. 1–5.

[23] A. L. Golberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Component of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[24] *The PTB Diagnostic ECG Database*. Accessed: Sep. 29, 2015. [Online]. Available: https://physionet.org/physiobank/database/

[25] A. L. Goldberg, *Clinical Electrocardiography*, 7th ed. Amsterdam, The Netherlands: Elsevier, 2010.

[26] P. Rubel *et al.*, "Toward personal eHealth in cardiology. Results from the EPI-MEDICS telemedicine project," *J. Electrocardiol.*, vol. 38, no. 4, pp. 100–106, 2005.

[27] D. Sadhukhan and M. Mitra, "Detection of ECG characteristic features using slope thresholding and relative magnitude comparison," in *Proc. EAIT*, Kolkata, India, Nov./Dec. 2012, pp. 122–126.

[28] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, Apr. 2006.

[29] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 145–171, 2015.

[30] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY, USA: Wiley, 2007.

[31] S. J. Press and S. Wilson, "Choosing between logistic regression and discriminant analysis," *J. Amer. Statist. Assoc.*, vol. 73, no. 364, pp. 699–705, 1978.

[32] D. G. Altman and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity," *Brit. Med. J.*, vol. 308, p. 1552, Jun. 1994.

[33] P. Komarek, "Logistic regression for data mining and high-dimensional classification," Ph.D. dissertation, Robot. Inst. School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, 2004.

[34] *AD8232-Single Lead, Heart Rate Monitor Front End*. Accessed: Aug. 2, 2017. [Online]. Available: http://www.analog

Deboleena Sadhukhan (S'16) received the B.Tech. and M.Tech. degrees (First Class Rank-First) in Instrumentation engineering from the Department of Applied Physics, University of Calcutta, Kolkata, India, in 2010 and 2012, respectively, where she is currently pursuing her Ph.D. degree. She received the S.P Bhattacharya merit award in 2010 and Kanoria research scholar award in 2010 from the University of Calcutta.

Her current research interests include biomedical signal processing and pattern recognition.

Saurabh Pal (M'13) received the B.Tech., M.Tech., and Ph.D. (Tech) degrees from the University of Calcutta, Kolkata, India, in 2000, 2006, and 2013, respectively.

He has teaching and research experience of more than 15 years. He is currently an Associate Professor with the Department of Applied Physics, University of Calcutta. His current research interests include biomedical instrumentation and signal and image processing.

Madhuchhanda Mitra (M'11) received the B.Tech., M.Tech., and Ph.D. (Tech) degrees from the University of Calcutta, Kolkata, India, in 1987, 1989, and 1998, respectively.

She is currently a Professor with the Department of Applied Physics, University of Calcutta. Her current research interests include biomedical signal processing, machine fault analyses, and material science.

Dr. Mitra received the "Griffith Memorial Award" from the University of Calcutta.