This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

A Review of Signal Processing Techniques for Electrocardiogram Signal Quality Assessment

Udit Satija, *Student Member, IEEE*, Barathram.Ramkumar, and M. Sabarimalai Manikandan, *Member, IEEE*

***Abstract*—Electrocardiogram (ECG) signal quality assessment (SQA) plays a vital role in significantly improving the diagnostic accuracy and reliability of unsupervised ECG analysis systems. In practice, the ECG signal is often corrupted with different kinds of noises and artifacts. Therefore, numerous SQA methods were presented based on the ECG signal and/or noise features and the machine learning classifiers and/or heuristic decision rules. This paper presents an overview of current state-of-the-art SQA methods and highlights the practical limitations of the existing SQA methods. Based upon past and our studies, it is noticed that a lightweight ECG noise analysis framework is highly demanded for real-time detection, localization and classification of single and combined ECG noises within the context of wearable ECG monitoring devices which are often resource constrained.**

***Keywords*—Electrocardiogram, Signal Quality Assessment, Signal Quality Index, ECG Noises.**

I. Introduction

Electrocardiogram (ECG) signals are widely acquired for many applications such as cardiovascular disease diagnosis, arrhythmias recognition, physiological feedback, sleep apnea detection, chronic patient surveillance, sudden cardiac arrest prediction, biometric, emotional and physical activity recognition systems [1]- [6]. All of these ECG application systems highly demand the exact determination of fiducial points (or characteristic points) of the ECG signal for accurate and reliable measurements of morphological features (including, amplitude, duration, polarity and shape) of local waves such as P-wave, QRS-complex, T-wave and U-wave, and the interval features (including, PR-segment, QT-segment, ST-segment, RR interval) [4]–[6]. Most ECG analysis systems require relatively noise-free ECG signals for obtaining the ECG measurements more accurately and reliably [6]. In practice, the ECG signals are often corrupted with different types of noises such as baseline wander (electrode contact noise and electrode motion artifacts), power line interference (PLI), electromyogram (EMG) noise, and instrumentation (IN) noise, making it almost impossible to perform a morphological analysis of such contaminated ECG beats [6]- [11]. Therefore, automatic assessment of ECG signal is highly demanded in reducing the false alarms due to presence of unacceptable level of noises.

A. Characteristics of ECG Noises

A brief description of temporal and spectral characteristics of ECG noise sources is summarized below.

Udit Satija, Barathram.Ramkumar, and M. Sabarimalai Manikandan are with School of Electrical Sciences, Indian Institute of Technology Bhubaneswar, Odisha India. e-mail: us11@iitbbs.ac.in, barathram@iitbbs.ac.in, msm@iitbbs.ac.in

1) Baseline Wander and Abrupt Drift: The baseline wander and abrupt drift noises are caused due to the respiration, body movements, poor electrode contact and skin-electrode impedance [8], [9]. The peak amplitude and duration of the BW depend on the various factors such as the electrode properties, electrolyte properties, skin impedance and the subject movement [8]. The frequency spectrum of the baseline wandering ranges between 0.05 Hz and 1 Hz. The amplitude of the respiration baseline wander varies by about 15 percent of peak-to-peak ECG amplitude at frequencies between 0.15 Hz and 0.3 Hz [8]. Under stress ECG recordings, the frequency of the baseline wandering increases due to the increased breathing rate. The motion artifacts are transient baseline changes (rapid drifts) caused by changes in the electrode-skin impedance with electrode motion [9]. The magnitude of the baseline wander may exceed the amplitude of the QRS complex by several times. The amplitude and duration of the motion artifact are 500% of peak-peak ECG amplitude and 300-500 ms, respectively [9]. The severe baseline wanders or motion artifacts can distort the ST segment and other low-frequency components of ECG signal. The ST segment distortion may lead to wrong diagnosis of myocardial infarction, Brugada syndrome and other ST-segment related abnormalities.

2) Power-line Interference: The power-line interference (PLI) noises are caused by inductive and capacitive couplings of ubiquitous power lines in the ECG signal acquisition circuitry. The PLI is a narrow-band noise centered at 50/60 Hz with a bandwidth of less than 1 Hz and an amplitude of up to 50 percent of the peak-to-peak ECG amplitude [10]. The lower frequency noise components of the PLI are mixed with the frequency content of the ECG signal. The severe structured noises can distort the morphological features such as amplitude, duration and shape of low-amplitude local waves of the ECG signal. In particularly, the P-wave distortions can lead to wrong diagnosis of atrial arrhythmias such as atrial enlargement and fibrillation.

3) Muscle Artifacts: The muscle artifacts (or electromyographic (EMG)) are caused by the electrical activity of muscles during periods of contraction or due to a sudden body movement. In general, the amplitude of an EMG noise is about 10% of the ECG amplitude with a bandwidth between 20 and 1000 Hz. Some studies show that the frequency range is between DC and 10,000 Hz with an average amplitude of 10% level [8]. Previous studies showed that the muscle artifacts can significantly alter the shapes of local waves of the ECG signal because the frequency of EMG noise considerably overlaps with that of the ECG signals, in the range of 0.01 Hz to 100 Hz [9]. Thus, removal of muscle artifacts is quite challenging without distorting the clinical features which are most essential

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

TABLE I: Description of Test ECG Databases

ECG Databases	No. of Records	Duration	Sampling Frequency	Bits/ sample	No. of Leads	Heartbeat Types	Noise Types	Remarks
MITBIHA [22]	48	30 min	360 Hz	11	2	NSR, PVC, APC, LBBB, RBBB, P, VF, F, j, f, x and Q	BW, MA, PLI, EM, AB	Mostly records are noise-free
PICC 2011 [23]	2000	10 sec	500 Hz	16	12	-	FL, Spk, AB, BW, MA, PLI, EM	Contains both acceptable and unacceptable category of data
MIT-BIHSTC [24]	28	24 min	360 Hz	12	2	Myocardial hypertrophy, Valvular heart disease, Myocarditis, Miscellaneous, Healthy controls	MA, PLI, BW	Mostly noise-free
Fantasia [25]	40	120 min	250 Hz	16/12	1	NSR	MA, PLI, BW	Mostly noise-free
MACE [26]	27	8 sec	500 Hz	16	1	NSR	Motion artifact	Contains motion artifact introduced in different activities
MIHBIHNST [27]	2	30 min	360	11	2	NSR, RBBB, APC, PVC, Blocked APC	MA, EM and BW	Noise is added at different SNRs
MIMIC-II [28]	>800	-	125	10	-	Multiple expert-annotations for alarms for asystole, extreme bradycardia, extreme tachycardia, and ventricular tachycardia	-	-
TELE [29], [30]	250	48 min	500	12	1	Pathological segments	MA, EM	Contains large set of bad quality signals

Normal sinus rhythm, NSR; premature ventricular contraction, PVC; atrial premature contraction, APC; left bundle branch block, LBBB; right bundle branch block, RBBB; paced beat, P; ventricular fibrillation, VF; fusion of ventricular and normal beat, F; nodal escape (junctional) beat, j; fusion of paced and normal beat, f; non-conducted P-wave, x; unclassified beat, Q; baseline wander, BW; muscle artifact, MA; electrode motion, EM; powerline interference, PLI; Flatline, FL; Spike, Spk;

TABLE II: Fiducial Features and Heuristics Rules (FF-HR) Based Signal Quality Assessment Methods

Method	Signal	Event Detection	Features Based Rules	Classes	Database	Performance
Orphanidou,2014 [31]	ECG & PPG	R-peak	Rules based on HR, R-R intervals, ratio of R-R intervals and template matching (TM)	Good/Bad	PICC, JR-HA	Sp =97%, Se = 93%
Quesnel,2013 [32]	ECG	QRS Complex	PQRST averaging of segmented beats	Good/Bad	PLTSTD, MIT-BIHNST	Pcc= 0.89
Quesnel,2014 [33]	ECG	PQRST shape	Ensemble averaging of PQRST complexes	Clean/ Contaminated	PLTSTD, MIT-BIHNST	-
Johannesen,2012 [34]	ECG	QRS complex	QRS onset, offset points for detecting the macroscopic ((MS) issues missing leads (ML) and SQA	ROC Analysis	PICC	Ac(Tr)= 92.3% Ac(Te)= 90.0%
Tat,2011 [35]	ECG	QRS Complex	Thresholding	Good/Bad	-	A = 92 %
Xie,2010 [36]	ECG	QRS Complex	Simple thresholding and the average peak-to-peak amplitude, locations of the QRS complexes to detect artifact such as rail contact, HF noise, low-power mask, sudden amplitude variation	Trend detection	Real time ECGs using TMC-HM	-
Vaglio,2010 [37]	ECG	QRS Complex	HR, QRS Regularity cross-correlation criteria, median beats, amplitude and interval parameters, QRS onsets after filtering, T-wave offset position,	Good/Bad	-	-
Wang, 2002 [38]	ECG	QRS Complex	Area differences between successive QRS complexes	Histogram and CHP	MITBIHA	-
Hayn,2012 [39]	ECG	QRS amplitude, regularity of RR intervals	Basic signal properties (amplitude, spikes, constant signal portions), number of CP in among various leads, and QRS amplitude vs. noise-amplitude ratio	Good/ Bad	PICC	Ac(Tr)= 93.3% Ac(Te)= 91.6% and 0.834(Event 2) 0.873(Event 3)
Liu,2011 [40]	ECG	RR interval	Four flags for detecting misplaced electrode, huge impulse, strong Gaussian noise, error of R-wave peaks by the TM after calculating Std and TH	Good/Bad	PICC	Se = 90.67% Sp = 89.78%
Hayn,2011 [41]	ECG	QRS Complex	QRS amplitude vs. noise-amplitude ratio ECG amplitude, spikes, constant signal portions, number of CP in between different leads	Good/ Bad	PICC	score=0.916 (Event 1), score=0.834 (Event 2) and score=0.873
Castiglioni,2011 [42]	ECG	QRS Complex	Ratio between the power of the main cepstral peak and the total cepstral power between 0.05 and 3 s, RR interval	Classification in spectrum	Recorded using wearable MagIC device	99.8%(subject 1) and 60% (subject 2)
Johannesen,2011 [43]	ECG	QRS onset, average RR interval	Constant derivative, amplitude TH	Excellent, Good, Adequate, Poor or Unacceptable	PICC	Se = 91% and Sp = 85%

Note: CHP: cumulative histogram plots; TMC-HM: teleMedCare health monitor; Pcc: probability of correct classification; CP:crossing points; Tr:training set; Te:test set; TH:threshold; Std: standard deviation

and ambulatory ECG records available from John Radcliffe (JR) hospital in Oxford. The method had a Sp of 97% and a Se of 93%. It is noted that the performance of the SQA method highly relies on accurate and reliable QRS complex detection under ambulatory and exercise ECG recording conditions.

2) Real-time Ambulatory ECG Quality Analysis: P.X.Quesnel et al. [32], [33] presented a real-time biosignal quality analysis of ambulatory ECG signals for detection of myocardial ischemia. The SQI based on the ensemble averaging of PQRST complexes was presented to detect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

TABLE III: Fiducial Features and Machine Learning (FF-ML) Based Signal Quality Assessment Methods

SQI Method	Signal	Event Detection	Features	Classifier	Classes	Database	Performance
Li,2012 [44]	ECG, PPG and ABP	QRS complex	Beat by beat onset points, amplitude of PPG, HR systolic, diastolic, mean and pulse BP, difference of beats	RVM	Excellent, Acceptable, Unacceptable	PICC, MIMIC II	FA suppression = 86.4% (A), 100% (EB) and 27.8% (ET) with no TA suppression
Redmond, 2012 [45]	ECG	QRS complex	QRS-detection-based features like MR of VEN to mean CVN, MR of VEN to median QRS height, MR of VEN to mean QRS CVN, MC of entire beat with CBS, median correlation of entire beat with clustered beat shape, MC of QRS with QRS from CBS, median correlation of QRS with QRS from CBS	PWC	Good, Average, Bad	Real ECG recordings (collected using TeleMedCareHealth Monitor)	Ac= 78.7%
Clifford,2012 [46]	ECG	QRS complex	SED, HOS and inter-channel and inter-algorithm agreement, 7 SQI including the percentage of beats detected, relative power in the QRS, skewness, kurtosis, percentage of the signal to be flat line, relative power in the baseline	MLP, SVM	Acceptable, Indeterminate, Unacceptable	PICC, MITBIHNST	Ac(Tr)=98% and Ac(Te)=97%
Naseri,2014 [47]	ECG	RR-interval	Energy-concavity index, correlation-based examination subroutine	NN	Good/ Bad	PICC	-
Behar,2013 [48]	ECG	QRS complex (using eplimited and wqrs)	Kurtosis, skewness, relative power in the QRS complex, relative power in the baseline, fraction and ratio of beats detected, sum of the eigenvalues associated with the five principal components over the sum of all eigenvalues obtained	SVM with a Gaussian kernel	Good/ Bad	PICC, MITBIHA, MIMIC II	Ac = 99% (Tr and Te on NSR and upto 95% for arrhythmias
Behar,2012 [49]	ECG	QRS complex and onset points,	pSQI, kSQI, basSQI, bSQI, rSQI, pcaSQI	SVM	Classification of noise	PICC and MITBIHA, MITBIHNST	Ac(Tr)=97.1% Ac(Te)= 93%
Kuzilek, 2011 [50]	ECG	QRS complex	TH on variance, covariance, maxima of ECG, maxima of dynamic range and average values, time lagged covariance matrices, mean, kurtosis and number of QRS complexes	SVM	Good/Poor	PICC	Score = 0.999 (Tr), 0.836 (Te)

Note: ABP:atrial blood pressure; MR:mean ratio; VEN:vector error norm; CVN:cluster vector norm; CBS:clustered beat shape; MC: mean correlation; PWC:parzen window classifier; FA:false alarm; RVM: relevance vector machine; A:asystole; EB:extreme bradycardia; ET:extreme tachycardia; Ac:accuracy; Tr:training set; Te:test set; NSR:normal sinus rhythm; GSF:Gaussian smoothing function; ASF:adaptive smoothing filtering; HOS:higher-order statistics

difference of beats etc. The method achieved the maximum FA reduction rate of 30.5% with a true alarm (TA) suppression rate below 1% on the MIMIC II database.

2) Parzen Window Supervised Classifier Based Method: S.J.Redmond et al. [45] presented ECG signal quality measures for unsupervised telehealth environments. This SQA method consists of QRS complex detector, heartbeat waveform feature extraction and Parzen window supervised statistical classifier. The method was evaluated on the single-lead ECG recordings. The method had an accuracy of 78.7%.

3) Machine Learning Approaches for ECG Quality: G.D.Clifford et al. [46] studied SQIs and data fusion for determining clinical acceptability of ECGs using the six SQIs (72 features for twelve leads) and four classifiers such as linear discriminant analysis (LDA), Naive Bayes (NB), support vector machine (SVM) and multi-layer perceptron (MLP) neural network to classify the ECGs as "acceptable (1)" or "unacceptable (-1)". Results on the PICC database showed that the SVM and MLP achieved the best accuracies of 99% on the training data (Set-a) and 95% on the test data (Set-b).

4) Correlation and Neural Network Based Method: H.Naseri, and M.R.Homaeinezhad [47] proposed a method based on three major stages: preprocessing, energy-concavity index (ECI) analysis and correlation-based examination with a suitably trained neural network. The method achieved the best Se of 100% and PPV of 98.92% for detecting high-energy noise and the Se of 92.36% and PPV of 94.77% for recognizing any other kind of disturbances using the PICC database.

5) Support Vector Machine Based Method: J.Behar et al. [48] presented a method to assess the ECG quality for both normal and abnormal rhythms for false arrhythmia alarm reduction of ICU monitors. The method employed seven SQIs introduced in previously published works and the SVM classifier with a Gaussian kernel. The signal quality indices (SQIs) are: the relative power in the QRS complex ($pSQI$); the relative power in the baseline ($basSQI$); the skewness of the distribution ($sSQI$); the kurtosis of the distribution ($kSQI$); the percentage of matching of detected beats by the *eplimited* and *wqrs* algorithms ($bSQI$); the ratio of number of beats detected by the *eplimited* and *wqrs* algorithms ($rSQI$); and the ratio of the sum of the eigenvalues associated with five principal components to the sum of all eigenvalues $pcaSQI$ [48], [49], [46]. The method was evaluated using three databases: the PICC database, the MITBIHA database, and the MIMIC II database. By using the SVM classifier with a Gaussian kernel function, the method achieved a classification accuracy of 99% for ECGs with normal sinus rhythm and accuracy upto 95% for ECGs with arrhythmia. In [49], the authors presented a single channel ECG quality metric using six SQIs and SVM classifier with a Gaussian kernel. The best results achieved on

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

TABLE IV: Non-Fiducial Features and Heuristics Rules Based Signal Quality Assessment Methods

Method	Signal	Feature Based Rules	Classes	Database	Performance
Zhang,2015 [51]	ECG	Multiscale entropy	Good/Bad	MITBIHA and MITBIHNST	-
Mann,2014 [52]	ECG	Overall range between 0.2 and 15mV, ratio of power in frequency range 5-20 Hz to total 0- 62.5Hz, SF of beat by beat, ILC among leads	Good/Bad	PICC	Confidence (major movement to rest)= 0.58 to 0.90
Tabares,2012 [53]	ECG	Autocorrelation and cross correlation among different channels and diversity due to channels	ROC	MITBIHA	Se = 86%, Sp = 91%, Ac =90% (ROC curve)
Maan,2011 [54]	ECG	Correlation and Kors matrix	Good/Bad	PICC	Pc= 92.2%, Se=97%, Sp=75.1%
Noponen,2011 [55]	ECG	Block wise l_1-norm of the residual of linear prediction and filtered signal	Good/Bad	PICC	Ac=93.2%, Se= 96.9%, Sp=80.4%, Pp=94.5%, -P=88.3%
Moody,2011 [56]	ECG	TH by tracking consecutive sample, overall range	Good/Bad	PICC	Score(Tr)= 0.913 Te=0.896
Chudcek,2011 [57]	ECG	Detecting missing lead (sum over the lead is less than zero), poor contact(amplitude change of limited time duration), high amplitude artefact(high peak or absolute value of the maxima), baseline drift, noisy lead (by threshold)	Good/Bad	PICC	0.903 for training data set A, 0.833 for testing data set B and 0.872 for evaluation data set C.
Langley,2011 [58]	ECG	Amplitude TH to detect the flat baseline (FB); saturation (SA); baseline drift (BD); low amplitude (LA); high amplitude (HA); steep slope (SS);	Acceptable/ Unacceptable	PICC	Score = 85.7% (Te)
Aboukhalil,2008 [59]	ECG	Some parameters like intervals, amplitude and gradient	Good/Bad	MIMIC II	Arrhythmia FAR (42.7% $\rightarrow$ 17.2%)
Vallejo,2014 [60]	ECG	Difference in rate-of-change of ECG and noise sources in spectral components	Good/Bad	PICC, MITBIHA	Pearson and Spearman rank correlations.
Xia,2011 [61]	ECG	Temporal-Spectral analysis, joint temporal-spectral analysis, self correlation, cross correlation, and entropy analysis	Acceptable, Unacceptable and Intermediate	PICC	Ac(Tr)= 93.5% and Ac(Te)=90.0%
Lee,2012 [62]	ECG	empirical mode decomposition (EMD) and statistical measures namely Shannon entropy, mean, and variance,to calculate threshold on first intrinsic mode function (IMF)	ROC	Recorded using ScottCare RZ153 series	Se = 96.63% and Sp = 94.73% (both Tr and Te set), Sp= 73.66%$\rightarrow$85.04% with Se (74.48%-74.62%) for AF
Oh,2004 [63]	ECG	KLT based RMS error and residual error percent ratio of the sum of largest principal components (KL coefficients) and the trace of the eigenvalue matrix	MOS test	Europian ST-T	Se = 99.57%, Sp = 99.98%

Note: SF:self correlation; ILC:inter lead correlation; QA:quality assessment; ROC:receiver operating characteristics; Tr:training set; Te:testing set; BPF: band pass filtering; Pp:positive predictivity; Np, negative predictivity; TH:threshold; FAR:false alarm rate; KLT:Karhunen-Love transform; RMS:root mean square; and MOS:mean opinion score

6) Rules-Based ECG Quality Evaluation: B. E. Moody [56] explored a number of heuristic rules that are used to detect the most common problems in ECG recordings in real time on a mobile phone. The *acceptable* signals occasionally have an extremely small or extremely large amplitude range. If the range was smaller than 0.2 mV, or larger than 15 mV, the signal was more likely to be *unacceptable*. Five rules were developed for detecting missing lead (sum over the lead is less than zero), poor contact (amplitude change of limited time duration), high amplitude artifact (high peak or absolute value of the maxima), isoline drift, and noisy lead (by threshold) in [57]. Results show that a combination of several rules is able to correctly detect a majority of poor-quality ECG signals as demonstrated using the PICC database. The method achieved a score of 0.913 and 0.896 for training and test sets, respectively.

7) Mobile ECG Quality Assessment: P. Langley et al. [58] presented an algorithm based on a set of thresholds to assess the ECG signal quality on the mobile phone. The method detects bad quality signals including flat baseline, saturation, baseline drift, low amplitude, high amplitude and steep slope. The method was evaluated using 500 ECGs taken from the PICC database. The method had an accuracy of 85.7%.

8) ECG False Alarm Reduction Using ABP Waveform: A.Aboukhalil et al. [59] described a false alarm (FA) reduction algorithm using only one arterial blood pressure (ABP) waveform. By using the simultaneous ECG and ABP recordings, the overall FA rate was reduced from 42.7% to 17.2%.

9) Modulation Spectrum Based ECG Quality Index: D. T. Vallejo et al. [60] presented a modulation spectrum based ECG quality index for telehealth applications. The modulation spectral signal representation (MSSR) is used to quantify the difference between spectral variations of ECG waveforms and noise sources. The method was evaluated on twelve-lead ECGs and two-lead ambulatory ECGs from the PICC and MITBIHA databases respectively, the synthetic ECGs with varying noise levels and the real-time ECGs obtained using the Hexoskin garment during three activities (sitting, walking, running). Results showed that the method outperforms the other quality metrics such as ECG sample kurtosis and the QRS complex 'in-band to out-of-band' spectral power ratio.

10)Computer Algorithms for Evaluating ECG Quality: H.Xia et al. [61] explored various techniques including the time domain analysis, frequency domain analysis, time-frequency analysis, self correlation, cross correlation, and entropy analysis for evaluating ECG quality. The ECG quality was measured by the spectral radius of the Matrix of Regularity. Evaluation on the PICC database showed that the algorithms had an accuracy of 93.5% in the train-set and of 90% in the test-set.

11)Motion and Noise Artifact Detection in Holter ECGs: J. Lee et al. [62] presented an automatic motion and noise artifact detection in Holter ECGs using empirical mode decomposition (EMD) and statistical features such as Shannon entropy, mean, and variance of the first intrinsic mode function (IMF). The method achieved a Se and Sp of 96.63% and 94.73%, respec-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

TABLE V: Non-Fiducial Features and Machine Learning Based Signal Quality Assessment Methods

SQI Method	Waveform	Preprocessing	Features	Classifier	Classes	Database	Performance
Tobon,2015 [64]	ECG	-	Modulation spectral signal representation	SVM and LDA	Good/Bad	MITBIHA, and MITBIHNST	-
Morgado,2015 [65]	ECG	-	Cross-covariance matrix among different leads	DT and SVM	-	-	89.8%
Schumm, 2010 [66]	EDA and A-ECG	-	Crosscorrelograms, cumulative frequency plots, and the Kolmogorov-Smirnov test	SVM	Classification	MITBIHA	Se=99.69% and Pp= 99.77%
Clifford,2011 [67]	ECG	HPF to remove BW and LF noise	Energy distribution, higher order moments and inter-channel and inter algorithm agreement six SQIs: iSQI; bSQI; fSQI; sSQI; kSQI; pSQI	LDA,SVM, MLP, ANN	Acceptable, Indeterminate, Unacceptable	PICC	Ac=99% (Set-a), 95% (Set-b) and 92.6% (test data)
Zaunseder,2011 [68]	ECG	-	Powers in different frequency bands	Decision tree	Good/bad	PICC	Ac = 90.4%
Yu,2006 [69]	ECG and PPG	-	HFE,LFE,LFE/HFE, PV, SNR power, entropy	SVM	Good/bad	-	92% correct classification

Note: A-ECG: ambulatory ECG; DT:decision tree; PPG:photoplethsmogram; HFE:high-frequency energy; LFE:low-frequency energy; and PV:pulse-wave variability

TABLE VI: Filtering Based ECG Signal Quality Evaluation Methods

SQI Method	Waveform	Features	Filtering	Classes	Database	Performance
Nemati,2010 [70]	SQI and data fusion of ECG	Nth order spectral moment, prominent peaks, spectral purity waveform	Modified KF	Good/Poor	-	-
Li, 2009 [71]	ECG and BP	SBP, DBP, and MBP, maximum positive and negative pressure slope, maximum up-slope duration, maximum duration above threshold pulse-to-pulse interval (T), pulse pressure and ECG-ABP delay time	Modified KF	Classification	MIMIC II	Se= 99.8%, Pp= 99.3% (true BP alarms) and Se=98.2%, Pp=99.4% (FA)
Li,2008 [72]	ECG and ABP waveform	Kurtosis and spectral distribution for ECG and beat by beat fuzzy logic and heuristic TH for ABP and KF for tracking after weighting each estimate	KF	Good/Bad	MIMIC II, MITBIHNST	99.9% (TTE), and 35% (FTE)
Aldecoa,2011 [73]	ECG	Kurtosis and temporal Dispersion, correlation coefficient, beat detection, time resolution and dynamic range	LMS adaptive filtering	Good/Bad	ECG from g.Tec (g.USBamp)	high correlation 0.95 ± 0.00, Pp = 0.97 ± 0.00
Silva,2012 [74]	ECG, ABP, PPG, Resp	Point by point SQI calculation and masking	Multichannal adaptive Filtering	Good, Bad, and May Be based on ROC	MIMIC II	-
Jekova,2012 [75]	ECG	Adjustable TH for different noise detection (LF, HF and PLI)	Derivative and other filters with different cutoff frequency	ROC	PICC	Se = 98.7%, Sp = 80.9% (InPr), Sp = 97.8%, Se = 81.8% (TrPr)
Jekova,2011 [76]	ECG	Amplitude dependent thresholding and rules	LPF, HPF and BPF with corresponding cutoff frequency	Acceptable/ Unacceptable	PICC	Score=0.908 (Event 1), Se=96.8%
Redmond,2008 [77]	ECG	Four mask including rail contact mask, HF mask, low power mask and final combined mask to detect amplifier saturation, HF artifact, low signal power and the maximum continuous segment of usable ECG	Notch filer and HF noise by 5th order elliptic HF filter	Good/bad	Recorded from TeleMedCare Health Monitor	Pp=98% & 10% (Mean & Std), -P=97% & 19% (Mean & Std), Se=89% & 20% (Mean & STD), Sp=98% & 14% (Mean & Std)

Note: ABP:atrial blood pressure; TH:thresholding; InPr, incidental problems; TrPr: transient problems; KF:Kalman filter; SBP:systolic blood pressure; DBP: diastolic BP; MBP:mean BP; TTE:true tachycardiac episodes; FTE:false tachycardiac episodes; Std: standard deviation; LPF:low pass filter; HPF:high pass filter; and BPF:band pass filter

(IIR) bandpass filter of pass-band $0.7 - 33$ Hz then square and smooth the signal by normalized 0.05 second hamming window FIR filter followed by square root operation) and the final combined signal mask (all three masks). The method was evaluated on real-time ECG signals. The method achieved a Se of 89%, Pp of 98%, and Sp of 98%. These masks do not exhibit the characteristic traits of artifact.

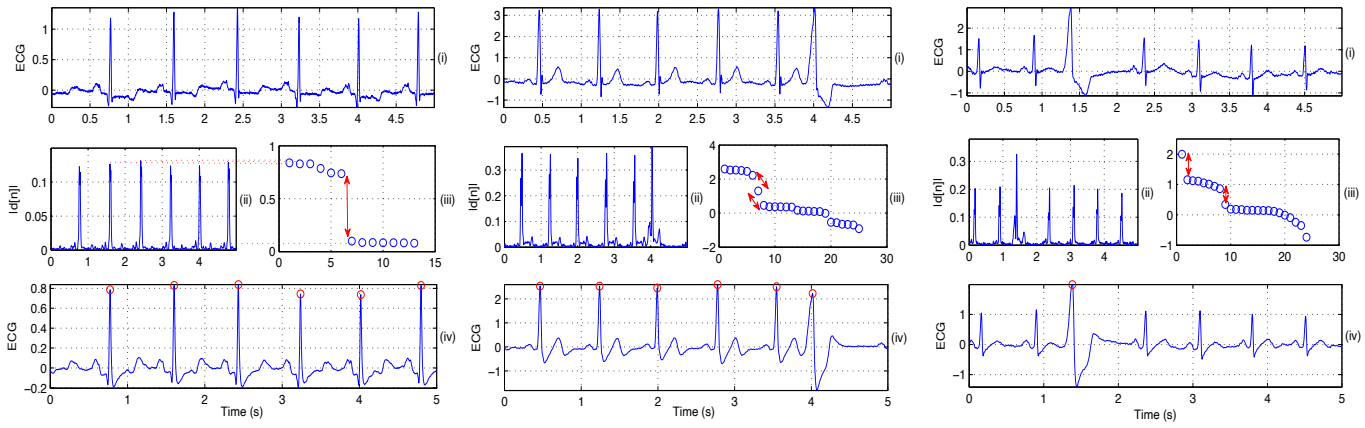
IV. Evaluation of Existing SQA Methods

One of the main objectives of this review paper is to highlight the practical limitations of the signal quality assessment methods. Therefore, we implemented some of the SQA methods using the well-known Pan-Tomkins R-peak detector, which was widely used in SQA methods, the QRS detection based SQA method [39], the template matching based SQA methods [31]–[33], higher order statistics (HOS) based SQA method [81], [48] for validating the performances using the pathological and non-pathological ECG signals taken from the standard ECG databases.

A. QRS Detector Based SQA Methods

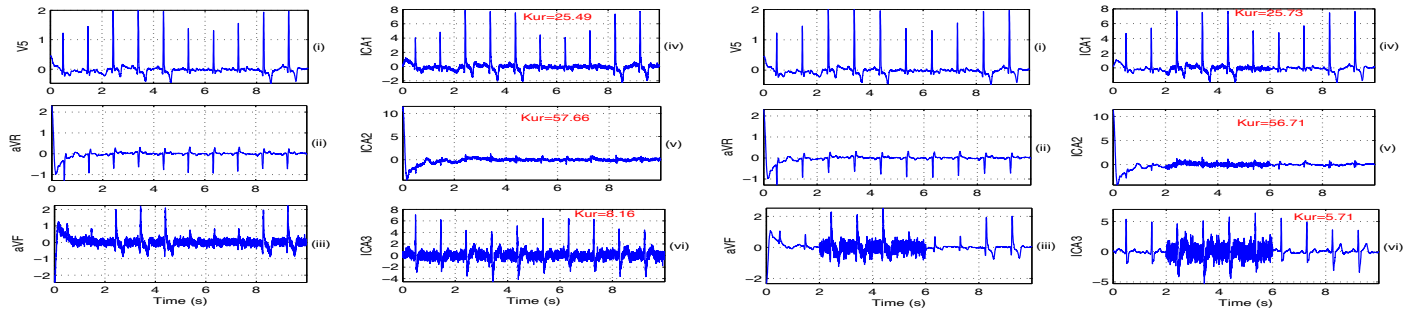
Many fiducial features based SQA methods employ traditional Pan-Tomkins QRS complex detector and its variants for automatically detection of R-peaks in the ECG signals [36], [45], [46], [49]. The Pan-Tomkins algorithm uses sets of amplitude-dependent, duration-dependent, interval-dependent thresholds to reject or include the noise peaks and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering



(a) (b) (c)

Fig. 3: Illustrates the threshold based R-peak detection stages [39]: (i) Original ECG signal; (ii) Forward difference; (iii) Mapping of maximas using blanking function; (iv) R-peak detection results for the ECG signals as shown in (a)-(c) with different types of heartbeat patterns.



(a) (b)

Fig. 6: Illustrates the kurtosis feature for 3-lead ECG separation using the ICA [81]: (a) the noisy ECG signals, (b) the noisy ECG signals with localized noise portions in the third lead.

missed peaks, respectively. The Pan-Tomkins QRS detector can achieve better detection rates for normal ECG with regular heart rates. It is difficult to find optimal threshold parameters under time-varying PQRST morphologies and heart rates and various noises. Under time-varying heart rates and long pauses, the search-back mechanism may produce more false positives since the algorithm does not have any information about a number of true R-peaks present in a real-time ECG signal. The performance of the Pan-Tomkins QRS detector is shown in Fig. 2 for the ECG signals corrupted with muscle artifacts. This detector had more false positives under severe noises and more false negatives for ECG signals with sudden changes in QRS complex amplitudes and shapes. The detection performance may be degraded under ambulatory and exercise ECG recordings. One can argue that if the conventional QRS detectors can detect all types of R-peaks more accurately and reliably, it can resolve the problems of some of ECG signal application systems. The results of the R-peak detection method reported in [39] are shown in Fig. 3. Although the method can accurately detect all R-peaks in the ECG signals as shown in Fig. 3 (a) and (b), the method results in more false negatives (FNs) for an ECG signal as shown in Fig. 3 (c). For different types of noise-free and noisy ECG signals, the detection performance of the method is shown in Fig. 4. Results show that the method had poor R-peak detection performance under time-varying PQRST morphologies. The limitations of the existing QRS detectors were highlighted in our previously published R-peak detection methods [78]–[80]. From our studies, it is noted that the R-peak detection errors can lead to inaccurate measurements of PQRST morphological and RR interval features which can result in an incorrect classification of ECG quality.

B. Template Matching Based SQA Method

Some of SQA methods assess the signal quality using the similarity metric between the extracted beats and the template

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2810957, IEEE Reviews in Biomedical Engineering

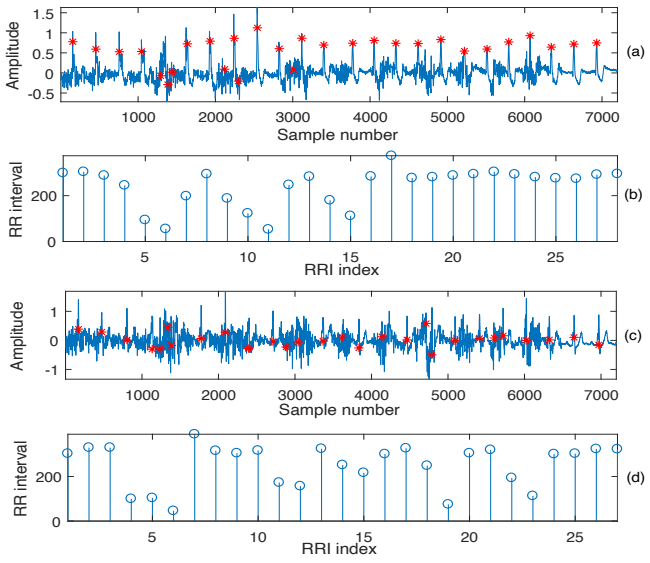


Fig. 2: Performance of Pan-Tomkins QRS complex detection for the noisy ECG signals. In some cases, the noisy RR intervals may be within the normal RR interval range which are used in the existing SQA methods.

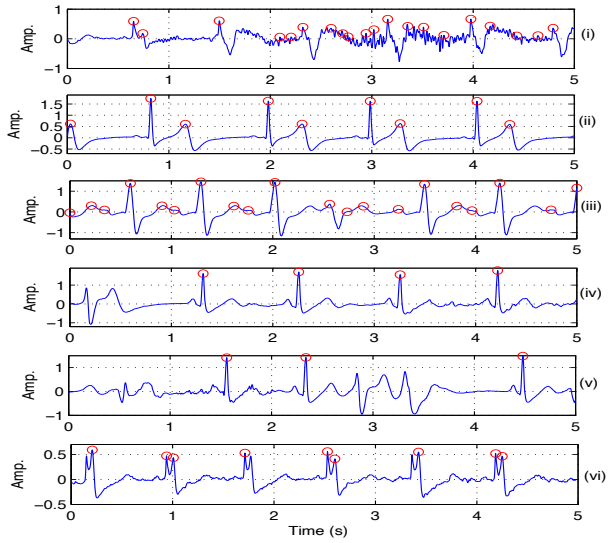


Fig. 4: Performance of the R-peak detector [39] for the ECG signals with different PQRST complexes and artifact.

beat [31]–[33]. The R-peaks are first identified by using the conventional Pan-Tomkins algorithm. Then, the ECG beats are extracted by using the locations of detected R-peaks. The SNR calibration [32], [33] and correlation [31] were performed between the individual beats and averaged PQRST beat. Fig. 5 illustrates the extracted beats and the estimated ensemble beat for 10 second ECG segment as reported in [31]. From the results, it is noted that the ensemble heartbeat and all the

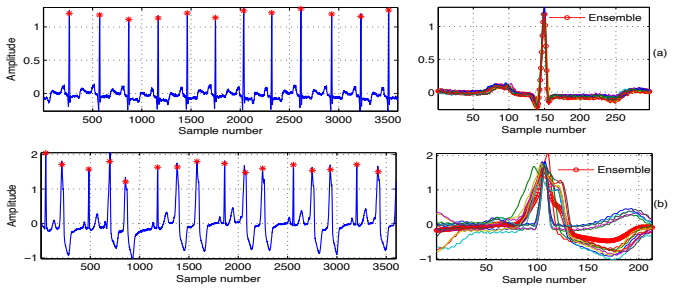


Fig. 5: Peformance of the correlation based SQA method [31]: (a) ECG signal (MITBIHA record 100); and (b) ECG signal (MITBIHA record 208).

extracted heartbeats exhibit a high correlation for the ECG signal with regular heartbeat patterns. The correlation between the ensemble heartbeat and other heartbeats is poor in case of ECG signals with different types of PQRST complexes. Therefore, these SQA methods [31]–[33] can provide good SQA performance for similar types of heartbeats but results in poor SQA performance for the dissimilar heartbeats within a quality decision making interval.

C. Higher-Order Statistics Based SQA Method

Some of the methods used kurtosis and variance features to detect the presence of Gaussian and non-Gaussian noises for assessing the signal quality [48], [81]. We implemented the independent component analysis (ICA) based method which consists of two stages: the ICA of three-lead ECGs and the kurtosis computation for all leads [81]. Fig. 6 demonstrates failure cases of ICA and kurtosis based method in identifying the noises. In [48], kurtosis and skewness features were used for discrimination of good quality and bad quality signals. Effectiveness of these features is demonstrated for the noise-free and noisy ECG signals with muscle artifacts, power-line and additive white Gaussian noises in Fig. 7 (a)-(c). Results show that the kurtosis and skewness values are higher for the good quality ECG signals as compared to the bad quality. It is noted that the kurtosis and skewness features may result in better classification accuracy for the ECG signals corrupted with AWGN and muscle artifacts and had poor result for PLI noises. However, the performance of the high-order statistics based SQA methods needs to be validated using different kinds of pathological ECG signals with various kinds of noise levels.

V. Challenging Issues and Future Directions

Based upon our past and current studies on ECG signal processing (such as R-peak detection [78]- [80], ECG compression [82], ECG denoising [16], [12], and the ECG noise detection and classification [6], [7], [83]–[86]), we highlight key challenging practical issues of existing SQA methods. We further present future research directions for the development of effective and efficient ECG signal analysis systems by significantly reducing false alarm rates under resting, ambulatory and exercise ECG recording conditions.