

Chapter 2

Quality Assessment for the Electrocardiogram (ECG)

Abstract In this chapter, we review a variety of signal quality assessment (SQA) techniques that robustly generate automated signal quality indices (SQIs) for the Electrocardiogram (ECG). Because of the many kinds of noise that may be present in the ECG, quantifying noise is not straightforward. However, several proposed techniques have been shown to provide efficient, accurate and robust quality assessment. In this chapter, emphasis will be given on currently proposed techniques for dealing with noise resulting from motion artifact focusing on data recorded from wearable sensors. The aim of this chapter is not to review currently proposed systems as complete solutions but rather to provide the reader with a basic understanding of the general framework for designing SQA systems for the ECG, in order to trigger further research.

Keywords ECG • Signal quality assessment • Feature extraction • Decision rules

In this chapter, we review a variety of signal quality assessment (SQA) techniques that robustly generate automated signal quality indices (SQIs) for the ECG. Because of the many kinds of noise that may be present in the ECG, quantifying noise is not straightforward. However, several proposed techniques have been shown to provide efficient, accurate and robust quality assessment. Standard approaches for measuring noise in the ECG assume stationarity in the dynamics of the noise; many of these techniques are covered in [1]. In this chapter, emphasis will be given on currently proposed techniques for dealing with noise resulting from motion artifact focusing on data recorded from wearable sensors. The aim of this chapter is not to review currently proposed systems as complete solutions but rather to provide the reader with a basic understanding of the general framework for designing SQA systems for the ECG, in order to trigger further research.

2.1 The Electrocardiogram (ECG)

The ECG is one of the most commonly recorded signals and can be found in even the most basic clinical environments. The distinctive morphology of the ECG signal can provide any clinically-trained observer a lot of information about the condition of the recorded individual's heart; it is the major visual record that cardiologists use for identifying a wide range of different heart disorders.

The ECG records electrical potential differences between prescribed locations on the surface of the body using skin electrodes; these differences in potential occur during the heart cycle, as the heart muscles contract and expand to pump blood around the circulatory system. The standard ECG set-up consists of 12 leads but often 3- and 5-lead setups are used as they are easier to structure. ECG can now also be measured using chest bands, skin patches and monitoring garments.

2.1.1 ECG Morphology

The ECG waveform is composed of three distinct waveforms, the P, QRS and T waves, each one representing an independent event related to the heart cycle. An example is shown in Fig. 2.1 with the three waveforms marked.

The most dominant wave, which should be present in all ECG recordings is the QRS wave, commonly known as the *QRS complex*. The QRS complex occurs as a result of the depolarization activity in the right and left ventricles [2]. Correct identification of the R-peaks in a segment of ECG is sufficient for measuring a reliable HR. The P-wave, which results from the depolarization of the right and left

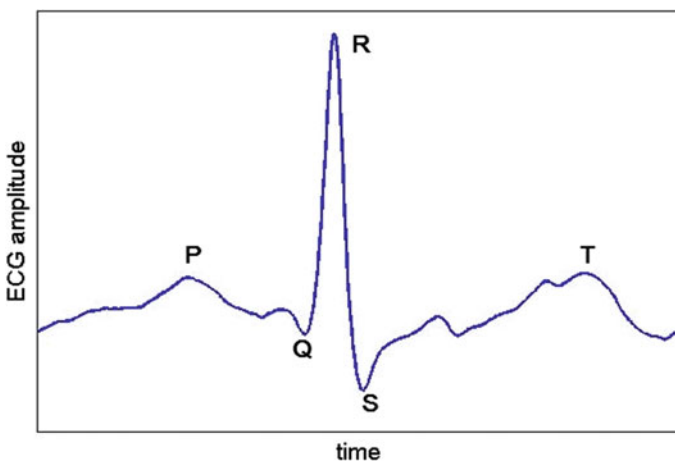


Fig. 2.1 Example PQRST waveform with P, QRS and T waves marked

atria [2] should be present in a healthy ECG recording, however it is absent in the presence of certain conditions, such as atrial fibrillation [3]. The amplitude of the T-wave, which represents the recovery of the ventricles, is also an important indicator of certain heart conditions; T-wave alternans, for example, which is the periodic variation in the peak amplitude and morphology of the T-wave, has been linked to sudden cardiac death [4]. The presence of T-wave abnormalities has also been associated with many heart conditions such as myocardial ischemia [5] and coronary heart disease [6].

When an ECG exhibits clearly the different waveforms and a clean and regular heart beat can be observed without any abnormalities, the heart is said to be in *sinus rhythm* [1]. To correctly identify any deviations from sinus rhythm or any abnormalities in the functioning of the heart, identification of all distinct waveforms present in an ECG recording is important.

In addition to morphological characteristics of single PQRST waveforms, the variation in the time-intervals between successive heart beats (the *RR-intervals*) within a longer ECG segment, known as Heart Rate Variability (HRV) is an important index of the activity of the Autonomic Nervous System (ANS) [7] and should be correctly measured from an ECG segment.

Correctly identifying the different waveforms of the PQRST wave is challenging at the best of times, because of the morphological variability of the signal for different individuals; even for the same individual, different recordings may differ substantially because of electrode type or placement. In the presence of noise, QRS

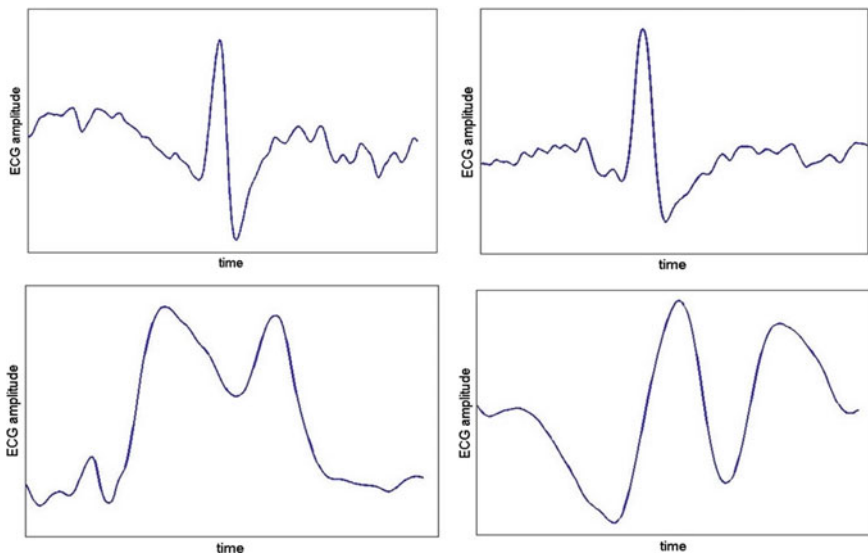


Fig. 2.2 PQRST waves containing noise. The *top* two waveforms contain high-frequency noise. While the R-peak is clean, the P-wave and the T-wave are contaminated with noise and cannot be correctly identified. In the *bottom* two PQRST waves the motion artefact makes determining the R-peak extremely challenging

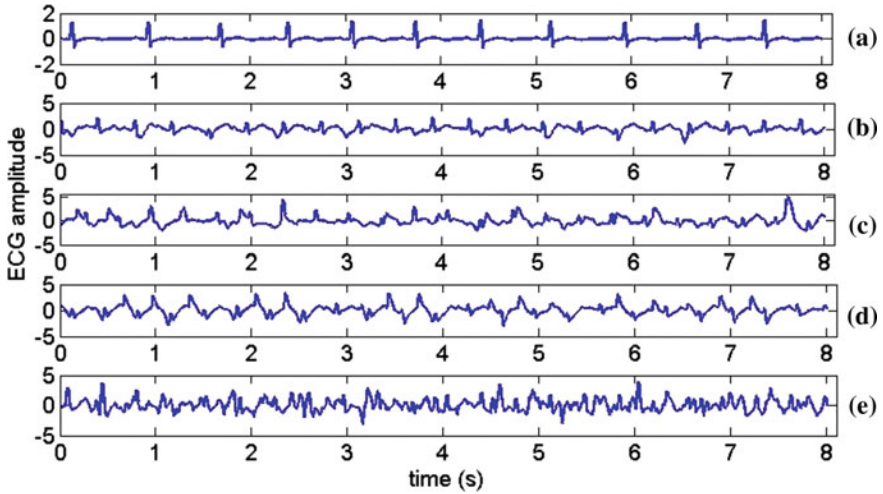


Fig. 2.3 Eight-second ECG segments with distinct kinds of noise. **a** Clean ECG segment. **b** Borderline case; this segment could result in estimating a reliable HR but would be unsuitable for diagnostic purposes. **c–e** Progressively noisier ECG segments which would result in unreliable vital sign estimates. As the segments get noisier, the overall morphology becomes more varied

(or R-peak) detection, for example, becomes even more challenging. According to Pan and Tompkins [8], the types of noise commonly found in the ECG which could affect the robustness of QRS detection are “muscle noise, artifacts due to electrode motion, power-line interference, baseline wander, and T-waves with high-frequency characteristics similar to QRS complexes.” P-wave and T-wave detection are also similarly affected by noise. The way that noise can interfere with waveform morphology is illustrated in Fig. 2.2 which shows examples of PQRST waveforms containing artefact. As we will see later, a good SQA approach is to identify and utilize morphological features of individual QRS or PQRST waveforms which are present in clean segments of ECG but are altered in the presence of noise. Other promising approaches build on the variability of QRS/PQRST morphology or R-R intervals rather than looking at individual beat waveforms; a clean segment of ECG is expected to have regular morphology in individual beat waveforms whereas a noisy one will be irregular [9]. When setting criteria for regularity/irregularity of QRS complexes in an ECG segment an important consideration is any irregularity which may be attributed to physiology. For example, R-R intervals are expected to show variability because of Respiratory Sinus Arrhythmia (RSA), the cyclic variation of heart rate associated with respiration [10]. The presence of arrhythmias will also add to this variability in the R-R intervals [11]. It is, thus, important for quality assessment approaches to set decision rules which do not result in a high number of false positives because of incorrectly identifying physiological abnormality as noise. Figure 2.3 shows examples of 8-second segments of ECG with different instances of noise showing the variability in beat waveforms which can form the basis for a SQA approach.

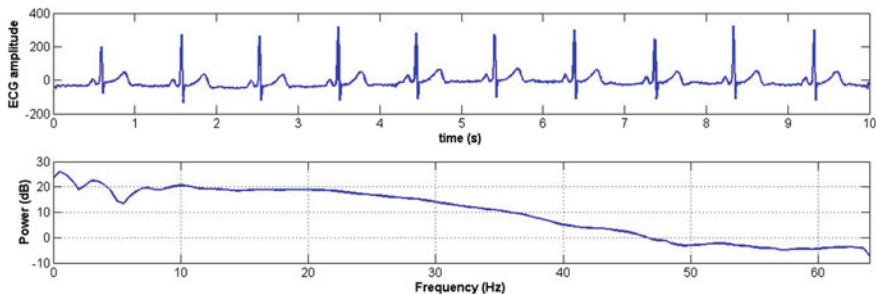


Fig. 2.4 Ten seconds of 128-Hz ECG in sinus rhythm (taken from the MIT-BIH Normal Sinus Rhythm Database) and associated 128-point Welch periodogram calculated with a hamming window and a 64-point overlap. The periodogram calculates the power spectral density (PSD) of the signal, i.e., the power (in dB, vertical axis) contained in each frequency (in Hz, horizontal axis). We note the peak power at approximately 1 Hz which is the HR (which can be verified from the ECG plot) and secondary peaks around 4 Hz (typical frequency of the T-wave), 7 Hz (typical frequency of the P-wave) and 10 Hz (typical frequency of the QRS complex) [1]

2.1.2 Spectral Characteristics of the ECG

The spectrum of a clean ECG in sinus rhythm has a distinct distribution of power; the accepted range of diagnostic ECG is 0.05–100 Hz (even though micro-fluctuations are said to be found in frequencies up to 500 Hz [1]). QRS energy is typically contained in frequencies below 30 Hz with peak power occurring in the range of 4–12 Hz [12]. The presence of notches in the QRS complexes (associated with some conditions) may broaden the power spectrum of the ECG but most power is still contained below 30 Hz. When noise is present in the ECG signal, the distribution of power is altered; an accumulation of power occurs in frequency bands which are either outside the physiologically-attributable range for the ECG or the overall distribution of power which would be considered normal is altered (Fig. 2.4).

Later, in Sect. 2.3.1.4 we will discuss methods for quantifying these alterations in the spectra of noisy signals to extract robust features for differentiating between *acceptable* and *unacceptable* ECG signals.

2.2 ECG Quality Considerations

As discussed in Chap. 1, the definition of signal quality differs depending on the application. For the ECG, we may define signal quality in two ways:

- *Basic quality*: R-peaks are clearly identifiable. In this case, a reliable HR can be extracted as well as some types of arrhythmias and HRV information.

- *Diagnostic quality*: P (if present), QRS and T waveforms are clearly identifiable. In this case, the signal can also be used for clinical diagnosis of more subtle conditions such as myocardial ischemia and coronary heart disease.

The first definition is less strict; as illustrated in Fig. 2.3, in the presence of noise it is often still possible to obtain a reliable HR. Baseline wander, for example, which is often considered to be the result of motion, does not necessarily hinder the measurement of HR. For assessing the quality of data obtained from wearable health and fitness sensors (which would be more likely to contain noise), most often the requirement of *basic quality* is sufficient. Diagnostic applications from the ECG are rarely used when the user is in motion; an exception is wearable systems intended for identifying, for example, episodes of Atrial Fibrillation (AF) [3, 11]. In the case of AF, the absence of the P-wave is a common indicator for the presence of this arrhythmia, so being able to identify it amongst the motion artefact is important.

For the remainder of this chapter, we will take a close look at different approaches proposed in literature for assessing the quality of ECG signals. We present current past and current techniques for feature extraction and for determining associated decision rules. Most approaches focus on obtaining *basic quality* of ECG, rather than *diagnostic quality*. We focus mostly on the former definition since for most diagnostic applications ECG assessment is done while the patient is static so the likelihood of motion artefact occurring is minimal. We separate approaches to single-channel and multi-channel ones. Single-channel approaches rely on extracting morphological or spectral characteristics from a single ECG waveform and applying heuristic, empirically determined, or machine learning-based decision rules. Multi-channel approaches, on the other hand, mostly rely on the level of agreement between features and estimations calculated from multiple simultaneously recorded channels of ECG.

2.3 Single-Channel Approaches

The advantage of single-channel (or single-lead) approaches is that the system is independent of the number of channels recorded [13]. Single-lead approaches are useful for applications requiring the real-time monitoring using wearable sensors and can be used in the most basic clinical environments where the option of multi-channel acquisition is not always available.

2.3.1 ECG Feature Extraction

Most signal quality assessment algorithms propose applying a series of different steps for accepting or rejecting segments of ECG. Often, simple feasibility rules are applied first, to quickly detect common distorting factors such as disconnected

electrodes, poor skin-electrode contact and external electrical interference [14]. These disturbances are successfully tracked via simple rules and result in the quick identification of a significant percentage of unusable segments, often delivering high classification accuracy. More sophisticated approaches are required for questionable segments which pass these initial checks and may produce a physiologically viable measurement of HR which, however, is likely to be erroneous.

We will begin by considering simple feasibility checks and then proceed to discuss ECG features which have been shown to successfully differentiate between “acceptable” and “unacceptable” segments. These features may be based on ECG morphology, spectral characteristics of the ECG or trend-based measures within a window of signal. For most time-domain techniques, the first step is applying a beat detector to the ECG.

2.3.1.1 ECG Beat Detection

Time-domain signal quality assessment algorithms require the use of a robust beat detection algorithm. In fact, most quality assessment schemes are intrinsically linked to the beat detector used since beat detectors themselves have different noise sensitivities. Two popular beat detectors used in many current approaches are the ones proposed by Hamilton and Tompkins [15] and Zong et al. [16]. The former, relies on the application of an elaborate linear and non-linear digital filtering scheme followed by differentiation, squaring, and averaging of the signal before rule-based peak detection, to identify the fiducial point. The latter is based on digital filtering followed by application of the length transform (LT) followed by application of a decision rule. Both methods are well-documented and available in open-source on Physionet [17].

2.3.1.2 Feasibility Checks

Following beat detection, the first step to a quality assessment approach is to apply simple feasibility checks which are often applied in a cascade of decision steps. Even though such rules cannot fully assess signal quality, they can be useful for quickly rejecting segments of low quality, or identifying flat lines and segments of saturated signal.

Flat Line Detection

Real ECG is never constant; even in segments that may appear flat, there is always some low-level noise. Flat segments of ECG, signifying a missing lead [13] are simply identified by tracking the number of consecutive sample points with the same value and setting a threshold to the time-period of unchanged signal value, which is considered to indicate that the signal is unusable [18]. Indicative thresholds

proposed in literature vary from 0.2 s [18] to 1 s [19]. An equivalent technique is searching for a constant derivative of zero where the derivative is obtained by taking differences between consecutive values [20]. A missing lead may not only result in a flat line but also a straight line. Searching for a constant non-zero derivative is also a good approach for identifying such instances of unusable signal.

Amplitude Limits and Variability

Within a given segment of ECG, the range of values that the ECG takes may be used as a simple quality indicator. A small range might indicate that an electrode is not correctly attached rendering the signal useless [18]. A broad range might indicate the presence of high-frequency noise in the form of spikes or low-frequency noise in the form of baseline wander [18]. Common quality indicators include thresholds on the acceptable range (min-max) of ECG values within a segment; Moody [18] proposes a minimum cutoff of 0.2 mV and maximum cutoff of 15 mV. Lead saturation, i.e. amplitude exceeding a threshold indicating the presence of spikes [21] or excessive amplitude which is maintained for a specific time-period (e.g., 200 ms [19]), indicating signal saturation, are also useful quality metrics for identifying unacceptable segments of signal [14, 19], as is the detection of a very small maximum amplitude, steep slopes [19] and sudden amplitude changes [22]. Often, these checks have to satisfy percentage portions of the signal rather than the whole signal for a segment to be rejected.

Baseline Wander

Baseline wander (BW) (or baseline drift), caused by patient movement, can also be quickly identified after the application of a 1 Hz low-pass filter and the assessment of the resulting signal. If the BW exceeds an empirically-set threshold, the signal is classified as unacceptable. As mentioned earlier, the presence of BW does not always render the signal unusable so for many applications the presence of BW is permitted.

Noise-Power Measurements

Standard noise-power measurements may also be applied to the ECG signal which can quantify the artefact present. According to Clifford et al. [1] these include:

- *Root mean square (RMS) power in the isoelectric¹ region;*
- *Ratio of the R-peak amplitude to the noise amplitude in the isoelectric region;*
- *Ratio of the peak value of a signal to its RMS value;*
- *Ratio between in-band (5–40 Hz) and out-of-band spectral power;*
- *Power in the residual after a filtering process.*

While these measures have been shown to correctly identify specific instances of noise, they rely on the assumption that the signal is stationary and that the noise we are trying to detect is Gaussian. This is not always the case for the ECG since often noises

¹The isoelectric region is a range around the isoelectric level which is the region between the P-wave and QRS complex, considered to be the most stable marker for 0 V [Clifford]. It records the short pause between the atrial and ventricular depolarization of the heart.

and artifacts are transient and their onset and duration are unknown. Additionally, some kinds of noise which can occur (e.g. movement artefact) are not Gaussian. The use of these indices is thus limited and they should be used in combination with other indices more suitable for identifying a wider range of noises and artefacts.

Physiological feasibility

The natural limits of cardiovascular physiology which drive the morphology of the ECG signal can be used as signal quality indicators either as individual viability checks or as a cascade of checks which can quickly identify very-low-quality segments and label them as unacceptable. Following beat detection, the following limits have been used to check for physiological feasibility:

- *Average HR*: the average HR measured within a segment of signal needs to satisfy a physiologically valid rate. Proposed limits in literature were 40–180 beats per minute (bpm) (0.67–3 Hz) [9, 23]. If the users are exercising the limits need to be adjusted to allow for increased valid HRs (e.g. up to 300 bpm (5 Hz)). Any estimated HR outside these limits is not physiologically viable and the segment can be automatically rejected.
- *Maximum R-R interval*²: based on a minimum viable HR of 40 bpm, the maximum R-R interval within a segment of signal cannot exceed $60/40 \text{ bpm} = 1.5 \text{ s}$. Allowing for a single missed beat, a maximum R-R interval limit of 3 s has been proposed as a simple feasibility check [23].
- *Maximum R-R interval/Minimum R-R interval*: under normal monitoring conditions, HR is not expected to change rapidly within a short time-segment. As a result, when considering short segments of ECG, R-R intervals are not expected to change substantially. In a short segment of signal (e.g. 10 s), the ratio of the maximum to the minimum R-R interval should not exceed 1.1, since a change in HR of over 10% within 10 s is improbable. Allowing for a single missed beat, a maximum ratio of 2.2 has been proposed as a feasibility rule for acceptable ECG [23].

Particularly for feasibility rules which are based on the distribution of R-R intervals, it is important for thresholds to consider the possible presence of arrhythmias, as they would affect the distribution of R-R intervals. Additionally, the last criterion only applies when considering short segments of signal and must be adapted for larger segments since in larger time-intervals it is possible to have greater changes in the values of instantaneous HR measured via the R-R interval.

2.3.1.3 Trend-Based Approaches

We now discuss trend-based techniques for SQA which search for regularity in a short segment of ECG. In a typical clean ECG recording in sinus rhythm,

²The R-R interval is defined as the time difference between successive R-peaks in an ECG segment.

irrespective of the specific QRS morphology of the recording, QRS complexes appear almost periodically and are very similar in morphology. The presence of artefact will distort this regularity. Therefore, the more regular a segment of ECG is, the more reliable it is presumed to be. The techniques we discuss next, search for this regularity using different measures within a short segment of ECG.

High order statistics (HOS)-skewness and kurtosis of the ECG

Skewness and Kurtosis are defined as the third and fourth standardized moments of a probability distribution, respectively; skewness measures the symmetry of the distribution and can take positive or negative values depending on whether the skew is on the right or the left tail of the distribution. Kurtosis, on the other hand, is a measure of how sharp the peak of the distribution is. A distribution with many outliers will have a high (absolute) value of skewness since the outliers will cause asymmetry in the distribution; it is also expected to have a low value of kurtosis since its probability distribution will be flatter and will approach zero slower. A distribution with no outliers will be more symmetric, and will have a sharp peak, thus approach zero faster and, consequently, have high values of skewness and kurtosis. As a result, skewness and kurtosis have been proposed as indices of the presence of outliers in an ECG segment which is equivalent to the presence of noise. The kurtosis of the normal distribution is 3; clean ECG in sinus rhythm has been shown to have a kurtosis larger than 5 [24]. Additionally, Clifford [1] has shown that muscle artifact has a kurtosis of around 5 and baseline wander and power-line interference have kurtosis lower than 5. The measurements of skewness (S) and kurtosis (K) of the distribution of an ECG segment can, thus, be used as an index of signal quality [13, 25] and can be calculated for 10 s segments from

$$S = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^3}{\hat{\sigma}^3} \quad (2.1)$$

$$K = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^4}{\hat{\sigma}^4} \quad (2.2)$$

where x_i is the discrete signal, $\hat{\mu}$ and $\hat{\sigma}$ are the empirical estimates of the mean and standard deviation of the distribution of x_i and N the number of data points in the signal. Li et al. [25] used the kurtosis-based quality index to assign segments with a kurtosis value greater than 5 as *acceptable* and a value of kurtosis lower than 5 as *unacceptable*. Skewness has not been used as a rule-based quality index but both skewness and kurtosis were used by Clifford et al. [13] as inputs into machine-learning-based quality indices. Figure 2.5 shows examples of acceptable and unacceptable ECG segments with their respective distributions and skewness and kurtosis values.

Variability in the R-R Interval

The regularity in the R-R intervals, calculated as the time-intervals between successive R-peaks, after R-peak detection, is another indicator of quality in an ECG segment since heart rhythm is expected to be regular. This measure relies on the reduced

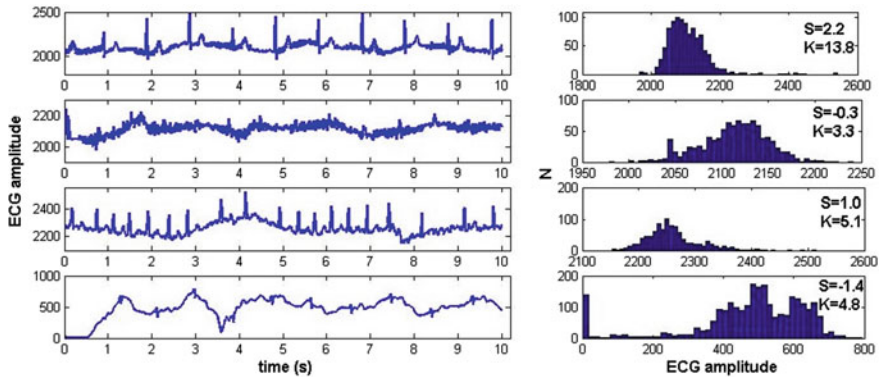


Fig. 2.5 Examples of 10 s ECG segments with corresponding discrete distributions (histograms) and skewness (S) and kurtosis (K) values. The kurtosis values are smaller for artefact-corrupted signals but skewness values are less consistent. For example, the second example from the top contains a large amount of high-frequency noise and the kurtosis value is very low but because of the nature of the noise, the distribution is approximately symmetric giving a low skewness value

performance of the QRS detector in the presence of noise: when artefact is present, the QRS detector underperforms by either missing R-peaks or erroneously identifying noisy peaks as R-peaks. This results in a high degree of variability in the distribution of R-R intervals and a resulting high value in the standard deviation of R-R intervals. Hayn et al. [21], proposed calculating the variability in the R-R intervals of an ECG segment using the coefficient of variation of the R-R intervals, given by

$$C_v = \frac{\widehat{\sigma_{RR}}}{\widehat{\mu_{RR}}}, \quad (2.3)$$

where $\widehat{\sigma_{RR}}$ and $\widehat{\mu_{RR}}$ are the empirical estimates of the mean and standard deviation of the distribution of the R-R intervals within a segment of ECG. The RR-variability quality index was then determined by assigning segments with a value of C_v smaller or equal to 0.64 as *acceptable* and a value of C_v greater than 0.64 as *potentially unacceptable* [21]. The threshold of 0.64 was determined empirically. Figure 2.6 shows examples of *acceptable* and *unacceptable* ECG segments with their respective distributions of R-R intervals and coefficient of variation values.

Template Matching

Another trend-based approach assigning a measure of regularity in an ECG segment relies on the similarities in the morphology of the QRS complexes within a single segment of signal. The template matching approach, used in [23], estimates the average QRS complex in a segment of ECG by performing beat detection and then averaging over all QRS complexes. Each QRS complex is extracted in a window around each detected heartbeat equal to the median R-R interval. Averaging all

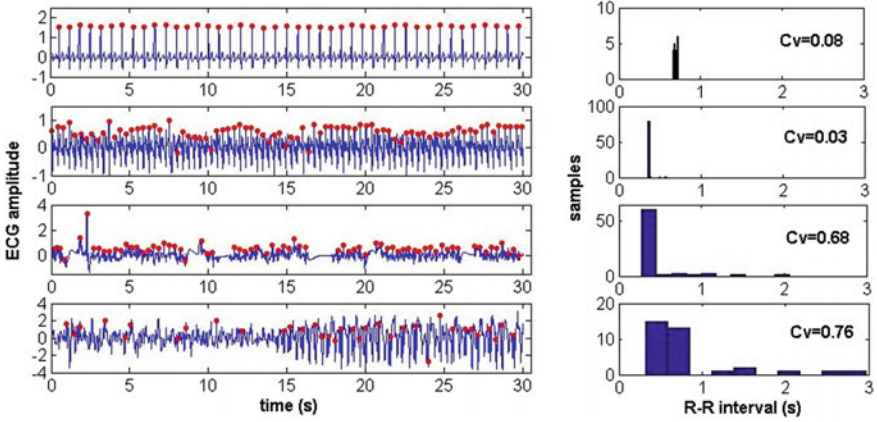


Fig. 2.6 Examples of 30 s ECG segments with *red circles* indicating R-peaks identified by application of the Hamilton and Tompkins QRS detector [15]. On the *right hand side*, we show the discrete distribution of the R-R intervals calculated from the R-peak detection and the corresponding coefficient of variation of the R-R intervals. The *top two figures* show examples of clean ECG in sinus rhythm with very small spread in the distribution of R-R intervals (i.e., no outliers) and a small value of C_v . The third and fourth examples are examples of noisy ECG either with flat regions (third example from the *top*) or an excessive amount of high-frequency noise which results in erroneous QRS detection. For both two cases, the spread in the distribution is large and the corresponding value of C_v is greater than the acceptability cutoff of 0.64. These examples illustrate the dependence of this quality index on the robustness of the QRS detector

QRS complexes provides a QRS template, which is adapted for every different ECG segment. The procedure is illustrated in Fig. 2.7 for a 10 s segment of ECG.

The Pearson's correlation coefficient of the template with each individual QRS complex is then calculated and averaged over all beats in the segment. *Pearson's correlation coefficient*, usually denoted by r for discrete time series analysis, measures the similarity between two time-series x_i and y_i using the following formula:

$$r = \frac{\sum_{i=1}^N (x_i - \hat{x})(y_i - \hat{y})}{\sqrt{\sum_{i=1}^N (x_i - \hat{x})^2} \sqrt{\sum_{i=1}^N (y_i - \hat{y})^2}}, \quad (2.4)$$

where $\hat{x}$ and $\hat{y}$ are the empirically determined means of x_i and y_i , respectively, and N is the number of samples in the two time-series.

This approach relies on the fact that in the presence of artefact, some or all QRS complexes will be corrupted; this will distort the QRS template (which is averaged over all QRS complexes) which will then have low correlation with individual QRS complexes. The template-matching quality index is assigned by labeling segments with an average correlation ≥ 0.66 as *acceptable* and a value of average correlation < 0.66 as *unacceptable* [23], where the average correlation threshold was determined empirically.

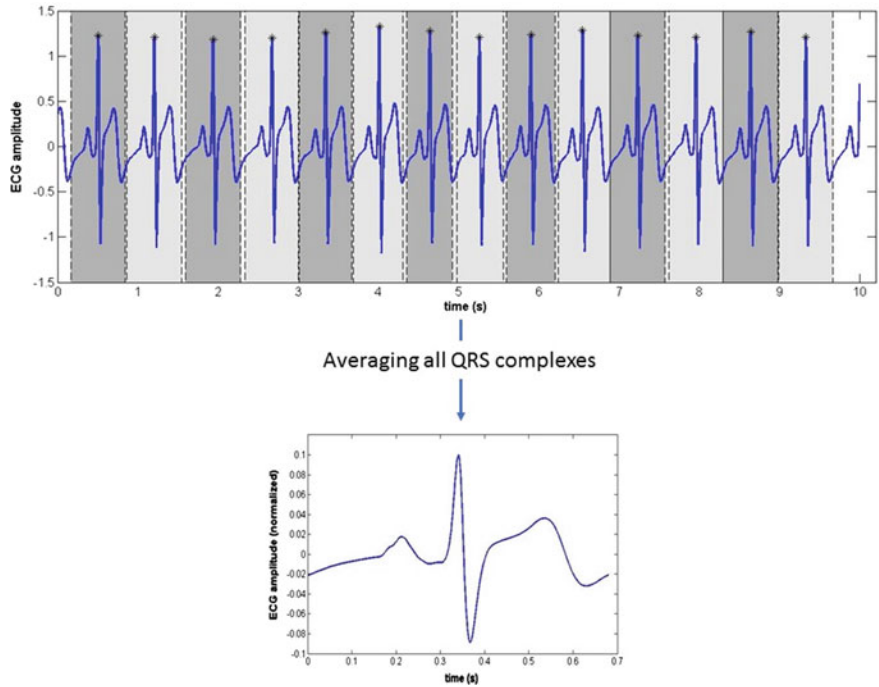


Fig. 2.7 Process for obtaining the average QRS template from a 10 s segment of ECG. The top figure shows individual QRS complexes (in alternate shading for ease of interpretation). The duration of each QRS complex is taken as the median R-R interval in the segment and it is centered around each beat detected (marked in *black stars*). The QRS complexes are then averaged to obtain the QRS template shown in the *bottom plot*

Figure 2.8 shows examples of *acceptable* and *unacceptable* ECG segments with their respective average QRS templates and average correlation coefficient values.

2.3.1.4 Frequency Domain Features

As discussed in Sect. 2.2, noise manifests itself on the spectrum of the ECG as increased power in frequencies outside the physiologically-attributed limits of the ECG or as a change in the distribution of power. We will now discuss some approaches for assessing the quality of the ECG which search for a distortion in the spectrum of the ECG. One approach presented, considers the spectral distribution of the Heart Rate Variability (HRV) signal, a physiologically rich signal which can be estimated from the ECG signal.

Frequency Content Across Different Bandwidths

The distribution of power in different bandwidths of the ECG signal has been investigated as a feature to differentiate between clean and noisy signals initially by

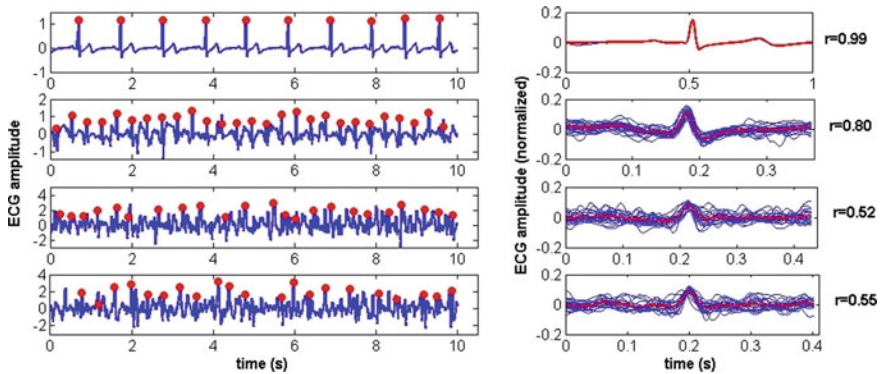


Fig. 2.8 Examples of 10 s ECG segments and extracted QRS templates. The *right-hand plots* show individual QRS complexes (in *blue*) and the QRS template in *red* with corresponding average correlation coefficient. The *top* ECG segment is a clean segment in sinus rhythm. As a result, QRS complexes are almost identical with an average correlation coefficient almost equal to 1. Despite the presence of some noise in the second segment, the robustness of the QRS detector results in a high enough correlation coefficient. The last two segments have a lot of variability in QRS morphology, resulting in an average correlation coefficient below the acceptable threshold of 0.66 and are thus labeled as unacceptable

Allen [26]. Since most physiologically relevant information in the ECG is contained in frequencies below 30 Hz, unusually high power in the frequencies outside this bandwidth will likely signify the presence of noise. The presence of baseline wander (BW) will also distort the distribution of power in the lower frequencies. Allen's seminal work [26] was an investigation on the bandwidths which show the most difference in frequency content between good quality and bad quality ECGs. Using a set of filters, the ECG was split into six different bandwidths such that the signal strength can be measured in each bandwidth. The six bandwidths considered were: low frequency (LF, 0.05–0.25 Hz), lower ECG bandwidth (ECG1, 0.25–10 Hz), higher ECG bandwidth (ECG2, 10–20 Hz), medium frequency (MG, 20–48 Hz), mains powerline noise (50, 48–52 Hz) and high-frequency (HF, 52–100 Hz). The signal strength in each bandwidth was then measured using a root-mean-square (RMS) detector. The system also considers a seventh feature, called the out-of-range event (ORE) feature, defined as the number of times the ECG exceeded a pre-set limit of ± 4 mV (this essentially represents instances of gross movement of the electrodes). The ECG1 and ECG2 bandwidths represented the typical monitoring bandwidth of the ECG and the rest of the bandwidths represented noise [26]. After establishing a baseline of good quality ECG, the authors studied differences in the frequency content features and OREs between good quality ECGs and ECGs collected at night (which were expected to be of lower quality due to uncontrolled user motion). All seven features investigated were found to be higher during the night compared with good quality ECGs with the most significant differences occurring in the LF and ECG1 bandwidths and the OREs feature. These features are thus considered promising for evaluating signal quality.

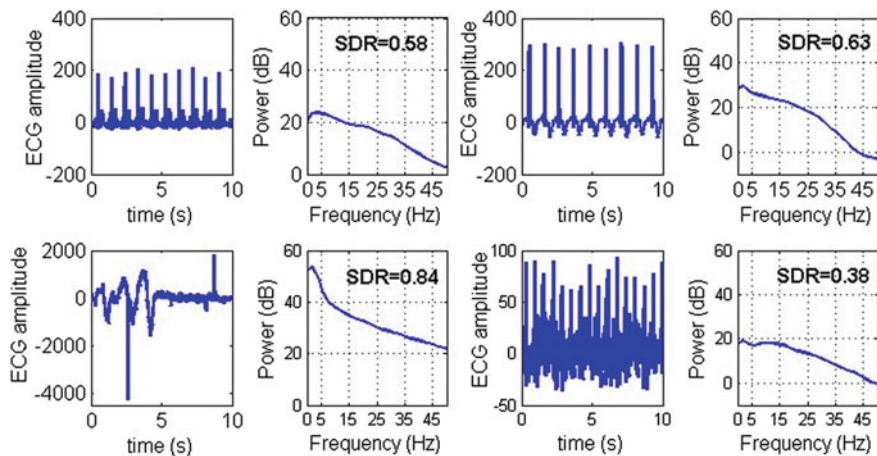


Fig. 2.9 Good quality (*top*) and bad quality (*bottom*) ECG segments with PSD plots and SDR values. The *top* two segments have an SDR value within the limits of acceptable signal whereas the *bottom* two have SDR values outside and are thus labeled unacceptable

Based on the same principle as Allen's investigation, another quality index proposed by Li et al. [25] calculates the ratio between the power spectral density (PSD) (using the Fast Fourier Transform (FFT)) between the 5–14 Hz frequency band and the 5–50 Hz frequency band producing the so-called spectral distribution ratio (SDR) of the ECG. In the presence of high-frequency noise, the SDR is likely to be low. When the SDR is high, it may be the case that excessive QRS-like artifact is present. The SDR-based quality index is then assigned by labeling segments with an SDR value between 0.5 and 0.8 as *acceptable* and signals with SDR values outside this range as *unacceptable* [25], where the thresholds were determined empirically.

Figure 2.9 shows examples of clean and noisy ECG segments with associated PSD plots and SDR values.

Spectral analysis of the HRV signal

It is possible to derive frequency-based features differentiating between *acceptable* and *unacceptable* segments of ECG from the heart rate variability (HRV) signal derived from the ECG. HRV reflects the variation in the time-intervals between consecutive heart beats. HRV is a physiological index, which reflects the interaction of the parasympathetic nervous system with the heart rate, when the heart is in normal sinus rhythm [7]. The classic way to obtain the HRV signal is to firstly calculate the R-R intervals; because the time point of each R-R interval can be either the preceding or succeeding R-peak, and is thus nonperiodic, the next step is to interpolate (using either linear or cubic spline interpolation) in order to obtain a smooth, periodically-sampled signal. HRV has been studied extensively since it reflects the underlying activity of the sympathetic and parasympathetic systems of the Autonomic Nervous System (ANS) and has been found to provide useful

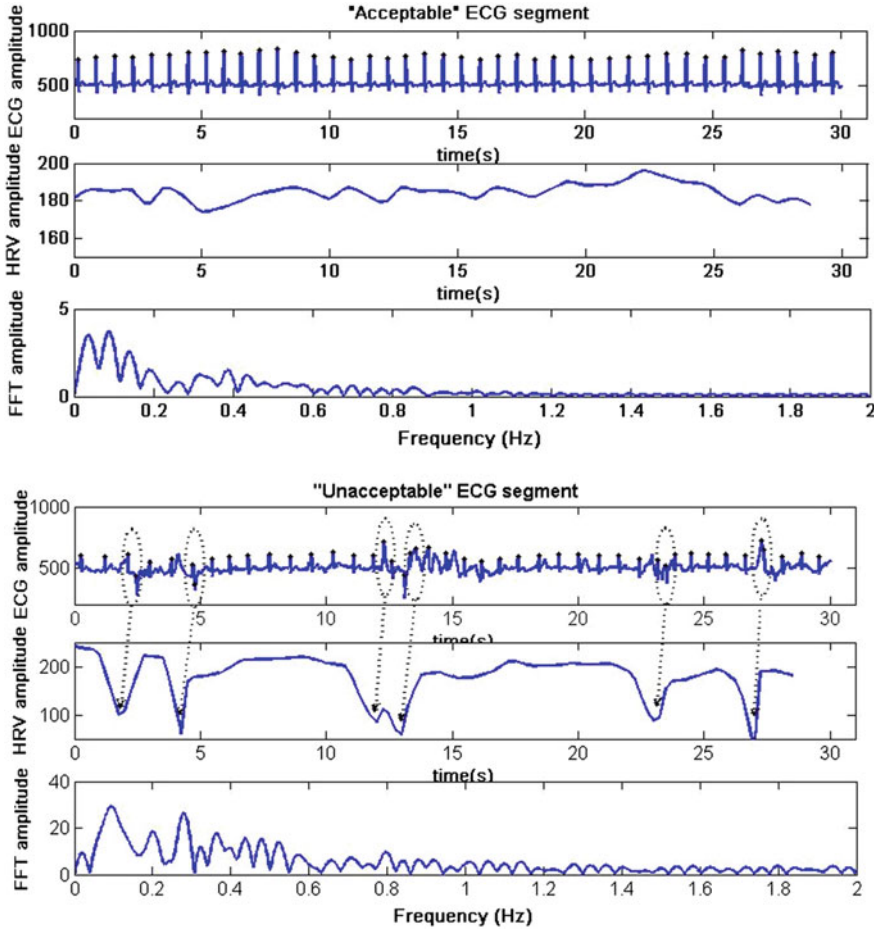


Fig. 2.10 Examples of “acceptable” and “unacceptable” ECG segments with extracted HRV signal and associated FFT spectrum. The HRV signal was derived from the R-R intervals of the ECG using cubic spline interpolation at a frequency of 4 Hz. The presence of noise in the unacceptable segment creates “dips” (*circled*) in the HRV signal which alter the spectrum and increase the strength of frequencies over the physiologically attributable upper limit of 0.4 Hz. Image adapted from [9]

clinical information [9]. The frequency content of the HRV signal is the result of different physiological processes with an upper limit of physiologically-attributed frequencies of 0.4 Hz. Since the calculation of the HRV signal relies on accurate beat detection, in the presence of noise, the HRV signal is distorted. The distorted HRV signal would also have a distorted distribution of energy in the different frequency bands since there would be energy in the non-physiologically-relevant bands, e.g., in frequencies greater than 0.4 Hz [9] (Fig. 2.10).

To extract the energy in different frequency bands wavelet entropy measurements were proposed. Wavelet decomposition allows the representation of the temporal features of a signal at different resolutions [20]. In practice, application of the discrete wavelet transform is the successive application of a two-channel, perfect reconstruction filter bank comprising of a low-pass and a high-pass filter, following by decimation by a factor of 2. The frequency bands associated with each decomposition level depend on the sampling frequency of the signal; at every step of the wavelet decomposition scheme the signal is split into its bottom half bandwidth (because of the application of a high-pass filter) producing the *approximation* wavelet coefficients and its top half bandwidth (as a result of application of a low-pass filter) producing the *detail* wavelet coefficients. The procedure is repeated, dyadically splitting the approximation levels for a specified number of decomposition levels (depending on the characteristics of the signal under study). For the HRV signal, which was interpolated from the R-R intervals time series at a frequency of 4 Hz, 5 levels were used which resulted in 6 sets of wavelet coefficients in the following bandwidths: 2–4 Hz (D1), 1–2 Hz (D2), 0.5–1 Hz (D3), (D4), 0.125–0.25 Hz (D5) and 0–0.125 Hz (A5). The *Shannon entropy* [27] of each decomposition level was then calculated, giving a measure of the disorder at each level. Comparison of the wavelet entropy measurements between acceptable and unacceptable ECG segments showed that all entropy measurements showed statistically significant differences between the two groups and that building a classification scheme using only level D2 (1–2 Hz) as a feature had a high degree of accuracy which was only marginally improved with the addition of D4 (0.25–0.5 Hz) and D5 (0.125–0.25 Hz). This is presumably because the artefact-corrupted segments of ECG had a high frequency content in the 1–2 Hz range whereas the clean segments didn't.

2.3.2 Decision Rules

In this section, we review the types of decision rules which may be used for making assessments of signal quality based on extracted time-domain or frequency-domain features from the ECG. Far better results may be obtained when combinations of features are used. When multiple features are being used together, the decision can be made based on a set of independent rules that need to be satisfied either concurrently or in a cascade of decision steps. In addition, machine learning can provide a “black box” strategy for building robust classifiers using an assortment of different features.

2.3.2.1 Thresholds and Combinations of Rules

The most common approaches for setting thresholds on features extracted from the ECG are based on:

- *Physiological limitations*: in human subjects, there are natural restrictions to the values that HR can have as well as the statistics of the distribution in the R-R intervals. It is, thus, possible to quickly identify signal segments whose derived measurements fall outside these physiologically-attributable restrictions. An important consideration when setting thresholds based on the R-R interval is the frequency and duration of the signal being analyzed, as well as the possible presence of an arrhythmia. A signal with a higher sampling rate will show more variability in the R-R intervals compared to a signal with a lower sampling rate. Furthermore, the “permitted” change in HR or variation in R-R intervals in a 60 s window of signal is greater than in a 10 s window. With respect to variability in R-R intervals an important but tricky task is to account for the variability in R-R intervals which might be the result of the presence of arrhythmias, rather than the presence of noise in the data [28]. This is important so that the system doesn’t end up producing false alerts when an arrhythmia occurs.
- *Expected limits of signal characteristics, such as amplitude, rate of change and frequency content, determined heuristically*: these rules are set either heuristically, based on the experience of the expected behavior of the signal itself, rather than the physiology the signal is measuring. In the absence of noise, ECG amplitude, for example, is not expected to change greatly, so many proposed algorithms have set thresholds on the maximum permitted amplitude (measuring it either from the isoelectric region or from the preceding or succeeding minimum within a specific time-window (for example around the QRS complex)). Often thresholds are set heuristically, and differ in different approaches and for different datasets. Hayn et al. [21] for example, has set a quality criterion which classifies a signal as “unacceptable” when “*the portion of samples that showed amplitudes of more than ± 2 mV was higher than 40% of the analyzed signal*”. Moody, on the other hand proposes labelling as “unacceptable” signal segments with a range (maximum-minimum within a 10 s segment) smaller than 0.2 mV, or larger than 1.5 mV [18]. Interestingly, both these cutoffs were set for the same dataset. Most works proposing decision rules based on expected signal characteristics do not contain detailed justifications of the thresholds set and they often seem arbitrary. It is assumed that most authors decide on the thresholds using combinations of experience on expected behavior, knowledge about equipment limits and trial and error using data available for training of the proposed algorithms, as discussed next.
- *Empirical evidence via the use of available data*: another way of setting quality cutoffs in extracted features is learning them on datasets available to researchers and developers. The most common approach is using a representative portion of the labelled data as the *training set*, trying different thresholds, and picking the optimal threshold by drawing the Receiver Operating Characteristics (ROC) curve which is a plot of 1-specificity (horizontal axis) against the sensitivity of the system (vertical axis) for different system parameters. If the same weight is given to the two statistical measures then the optimal threshold is the one which minimizes the distance to the point (0,1). When choosing thresholds empirically, the dataset used is important; if the aim is to create a system that can

generalize, parameters need to be set using data from multiple monitors, populations and clinical scenarios [23]. If parameters are tailored to a specific monitor, population and clinical scenario, it is possible that when used on data from a different monitor and in a different clinical setting, the SQA system might underperform (since the type of noise might differ).

As mentioned earlier, most of proposed SQA systems employ combinations of rules based on multiple features which need to be satisfied concurrently or in a cascade of decision steps. The order and combination of steps is mostly chosen either heuristically or to optimize system performance on training data in the same way as explained earlier. Often, the more basic checks such as flat line detection, physiological feasibility checks and checking for extreme signal behaviors are performed first; this way obviously low-quality segments are rejected quickly before more sophisticated and computationally intensive techniques are used for the ambiguous cases.

2.3.2.2 Machine Learning Models

In many cases, machine learning approaches are suitable for use in SQA systems. The more traditional rule-based decision-making approaches we have discussed so far, require prior information about the process to be modelled (i.e., information about how certain features differ in acceptable and unacceptable signal segments). Especially in the more ambiguous cases, the differences between acceptable and unacceptable signals are not clearly understood. Furthermore, the underlying processes which result in an unreliable signal are often very complex and multidimensional. Machine learning favors a *black box* approach: the relationship between features and labels does not need to be fully understood [29]; for a robust signal quality assessment system to be built, the system does not need to understand the underlying labeling process, it just needs to learn how to replicate it. The inclusion of aggregate data in machine learning models may reveal new information that is not seen by the individual. Machine learning models also offer the possibility to model the labeling process in a non-linear fashion which would not be possible, using traditional rule-based approaches. For robust machine learning models to be built, large amounts of labelled training data need to be available. In the way that the human expert, the gold standard of clinical decision-making, gains clinical acumen through experience, via machine learning, the system is taught how to perform a task (in this case classifying signal in terms of quality) using a lot of examples of how it should be done. The more information the system receives, the “experience” grows and the decision making is, thus, improved.

An overview of the use of machine learning in applications using multidimensional clinical data can be found in [29]. In the context of signal quality, Support Vector Machines (SVMs) have been used for building signal quality classifiers [9, 13] as well as Neural Networks via the Multi-Layer Perceptron (MLP) [13]. In Ref.

[9], the quality assessment system used wavelet entropy measurements from the HRV signal, derived from the ECG. A classification model was learnt using expertly annotated data from ambulatory patients, using SVMs with a radial basis function (RBF) kernel. Using only three features, a classification performance of 96% was achieved with a sensitivity of 92.3% and specificity of 99.1%. Despite no prior information being input into the classifier, the performance of the classifier using different combinations of features (which in this case were the entropies in different frequency bands, extracted via wavelet analysis, as explained in Sect. 2.1.3.4) gave an insight into the bandwidths where the differences between the two classes occurred (i.e., where the noise occurred).

In Ref. [13], the authors extracted 7 morphological, spectral and statistical features from 12 channels of simultaneously recorded ECG (see next section for an overview of multi-lead approaches), resulting in 84 features in total. A classifier was then trained using a Support Vector Machine (SVM) also with an RBF kernel and a standard feed-forward Multi-Layer Perceptron Neural Network (MLPNN). Classifiers were firstly tested using all 84 features and then using only the 7 features from a single lead. Like the approach by [9], explained earlier, different combinations of the seven features were tested to find the best. For the single-lead classifiers the results from the MLP and SVMs were almost the same, with the SVM approach slightly outperforming the MLP, with a classification accuracy of 96.5%, Sensitivity of 97.2% and Specificity of 95.8% on the test set using only four features. For the 12-lead case, the best classification accuracy occurred when using five features using the MLP, with an accuracy of 95.9% on the test set.

As discussed in the earlier discussion of empirical determination of thresholds, also in the case of machine learning models, the dataset used for training the system needs to be varied if the aim is for the system to be able to generalize to data from various monitors. Often, quality assessment systems are designed for on-board processing and in that case parameters can be tailored to that specific scenario. If the aim is for the system to be used independently of the monitor used, it needs to be trained on data from a variety of sources.

2.4 Multi-lead Approaches

So far, we have considered approaches extracting features and setting decision rules which can be used for the quality assessment of single-lead ECGs. We will now consider the situation where multiple channels of simultaneously recorded ECG is available. The presence of multiple channels of signal can be exploited for producing more reliable vital sign measurements, by employing methods based on the fusion of information from the various channels. Compared to single-channel approaches, when multiple channels of ECG are available system performance may be improved via a) the extraction of new features derived from the relationship between the different signals b) the use of an increased number of features, compared to single-lead cases, extracted from the different available signals. The 2011

Physionet Challenge addressed this exact problem: the development of an algorithm for assessing the quality of 12-lead ECG recordings collected via a mobile phone in real-time [17, 30]. Researchers had access to 1500 expertly annotated, 10 s in duration 12-lead ECG segments, sampled at 500 Hz, two thirds of which were to be used for training and the rest for testing. Next, we discuss approaches proposed by researchers which exploit the relationship between the multiple leads of ECG to obtain robust quality assessment.

2.4.1 Features Specific to Multi-lead ECGs

Correlation between different leads

According to Mogardo et al. [31] “the 12-lead ECG signals are different projections of the same electrical activation process of the heart”. As a result, it is expected that in a clean recording, the 12 leads should look very similar. In a 12-lead ECG contaminated with noise, the similarity between the different leads is presumably reduced (since noise may be manifested differently in the various leads). Correlation, defined in Sect. 2.3.1.3, measures the similarity between two time-series and has been proposed by Kranstein [32] as a quality feature measuring the similarity between different leads. While the way Kranstein measured correlation between the twelve leads is unclear, the potential of the technique may be demonstrated by calculating the correlation coefficient between each possible pair of leads in a 12-lead ECG signal (66 combinations) and taking the mean *absolute* correlation (to account for inverted leads). Such examples for acceptable and unacceptable signals are illustrated in Figs. 2.11 and 2.12, respectively, along with the average inter-lead correlation coefficient. For the acceptable signal, the average absolute correlation coefficient was 0.67, while for the unacceptable signal it was 0.31. In proposed methods, correlation measures were used as inputs into machine learning classifiers, however, it may be possible for the average absolute correlation coefficient to be used as part of a rule-based system, where a threshold of acceptance is determined on available data.

The concept of measuring cross-correlation between the different leads was taken a step further by Morgado [31] who, using the data from the 2011 Physionet Challenge, calculated the covariance matrix of the leads (a matrix in which each element z, k is the covariance value between leads z and k). Like correlation, covariance is also a measure of similarity between two time-series x_i and y_i , given by

$$\text{cov}(x, y) = \sum_{i=1}^N (x_i - \hat{x})(y_i - \hat{y}), \quad (2.5)$$

where $\hat{x}$ and $\hat{y}$ are the empirically determined means of x_i and y_i , respectively, and N is the number of samples in the two time-series. (Comparing with the formula for

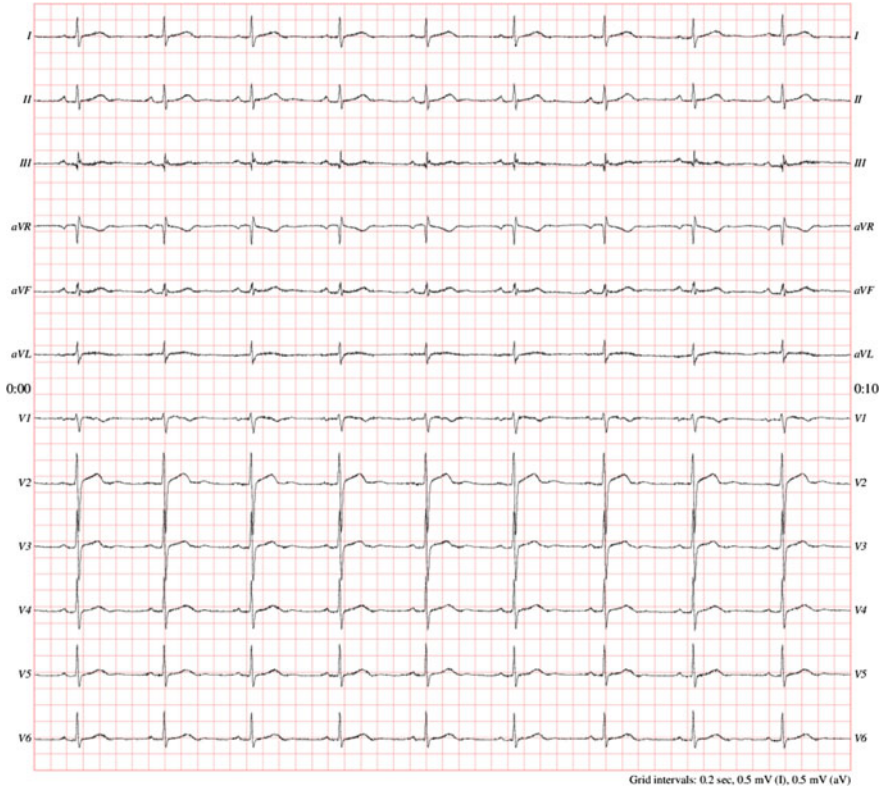


Fig. 2.11 Example of acceptable 12-lead ECG from the 2011 Physionet Challenge (image taken from Physiobank ATM [17]). The average inter-lead correlation coefficient was 0.67

the correlation coefficient (3), we can see that the covariance is the unnormalized equivalent of the correlation coefficient.). The authors used 8 out of the 12 leads of ECG and, thus, extracted an 8×8 covariance matrix which they then analyzed further to extract the eigenvalues. The eigenvalues were then used as features, fed into three binary classifiers, trained on the training data available. The highest accuracy obtained using this novel technique was 92.7% with a sensitivity of 83.1% and a specificity of 95.5%.

Agreement in QRS Detection Between Leads

In a clean 12-lead ECG beats should be detected successfully in all leads at the same time-points. Failure for this to happen signifies the presence of artefact. On this basis, the agreement in QRS detection between leads was proposed as a quality feature by [13] and used as one of the seven features fed into a machine learning-based classifier. This quality index was defined as “the percentage of beats detected on each lead which were detected on all leads” and was measured separately for each lead.

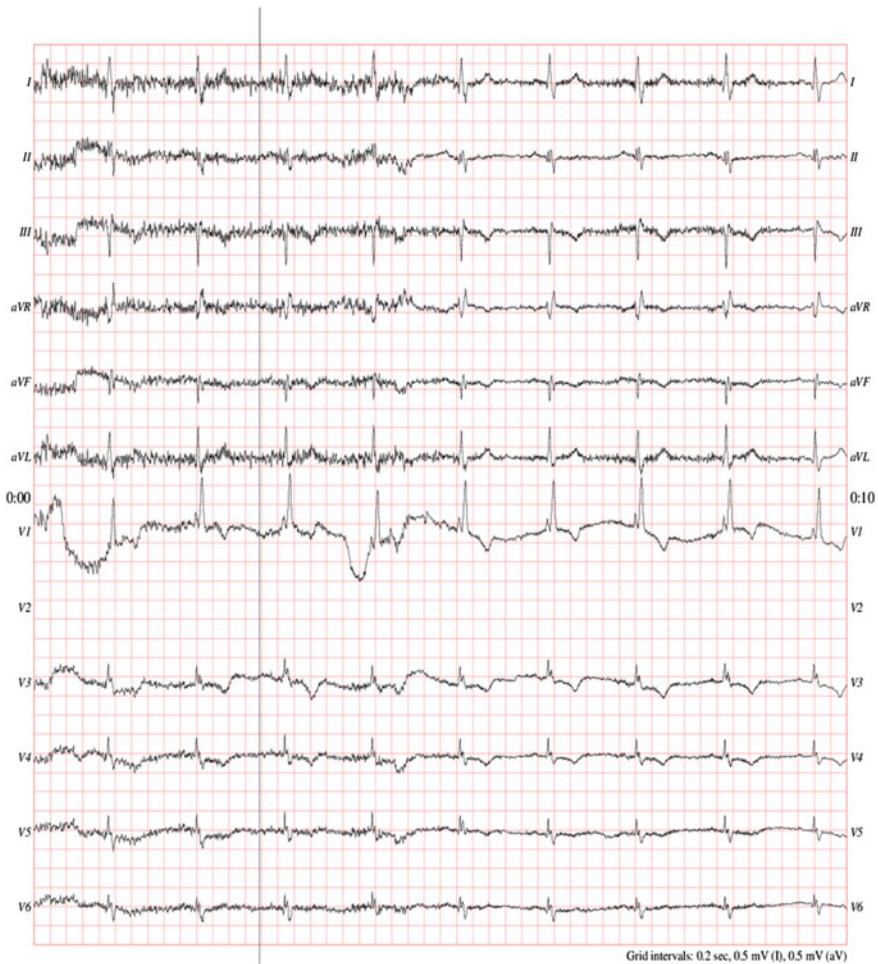


Fig. 2.12 Example of unacceptable 12-lead ECG from the 2011 Physionet Challenge (image taken from Physiobank ATM [17]). The average inter-lead correlation coefficient was 0.31. (Note that one lead is missing)

Lead Crossing Point Detection

This measure, proposed by Hayn et al. [21], was based on the visualization method used by Physionet's ATM [17], which plots all 12 leads in succession. When artefact was present, often one or more of the leads was plotted "over" other leads, obscuring them. Since human annotators used the same view for labeling the ECG records, the authors proposed a quality measure which would penalize records where one or more leads drifted onto another lead. They proposed a measure where the measured the number of crossing points of one lead with any other leads and if the maximum number exceeded 4.9 per second (this was presumably determined

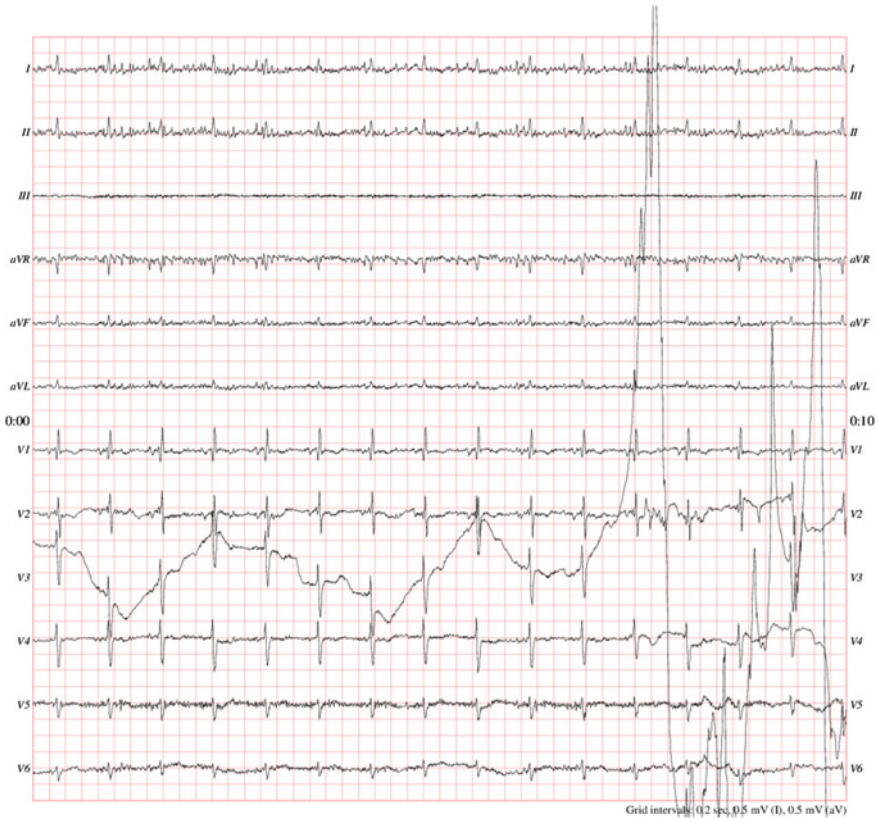


Fig. 2.13 Example of 12-lead ECG from the 2011 Physionet Challenge (image taken from Physiobank ATM [17]) with drifting leads interfering with other leads' plots on the right-hand side

empirically), the signal was classified as “unacceptable”. Figure 2.13 shows an example of an ECG record from the Physionet ATM with drifting leads on the right-hand side.

In addition to new features exploiting the relationship between different leads, most researchers proposed techniques which relied on the extraction of features from every lead and using them together either in a rule-based or machine-learning-based scheme for assessing the quality of the ECG records.

2.5 Summary

This chapter introduced the ECG and provided an overview of proposed approaches for feature extraction and of the decision rules employed for designing SQA systems for the ECG. Single-lead approaches have more practical applications but

where available, multi-lead approaches may significantly improve robustness. Not all past and currently-proposed techniques can be presented. Rather, this chapter is intended to introduce the reader to a variety of proposed approaches which have proved successful, and the underlying principles underpinning them, such that further research can be triggered, to improve the robustness of current approaches.

References

1. Clifford, G. D., Azuaje, F., & McSharry, P. E. (2006). *Advanced methods for tools for ECG data analysis*. Norwood, MA: Artech House.
2. Hurst, J. W. (1998). *Naming of the waves in the ECG, with a brief account of their genesis*. *Circulation*, 98, 1937–1942 (originally published November 3, 1998).
3. Romero, I., Penders, J., & Kriatselis, C. (2010). P-wave analysis for atrial fibrillation detection in ambulatory recordings. *Journal of Electrocardiology*, 43(6), 647.
4. Monasterio, V., Laguna, P., & Martínez, J. P. (2009). Multilead analysis of T-wave alternans in the ECG using principal component analysis. *IEEE Transactions in Biomedical Engineering*, 56(7), 1880–1890.
5. Channer, K., & Morris, F. (2002). Myocardial ischaemia. *BMJ: British Medical Journal*, 324 (7344), 1023–1026.
6. Jacqueline, M., Dekker, J. M., Schouten, E. G., Klootwijk, P., Pool, J., & Kromhout, D. (1995). ST segment and T wave characteristics as indicators of coronary heart disease risk: The Zutphen study. *Journal of the American College of Cardiology*, 25(6), 1321–1326.
7. Pieper, S. J., & Hammill, S. C. (1996). Heart rate variability: Technique and investigational applications in cardiovascular medicine. *Mayo Clinic Proceedings*, 70(10), 354–381.
8. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME 32(3), 230–236.
9. Orphanidou, C., & Drobnyak, I. (2017). Quality assessment of ambulatory ECG using wavelet entropy of the HRV signal. *IEEE Journal of Biomedical and Health Informatics*, 21(5), 1216–1223.
10. Orphanidou, C., Fleming, S., Shah, S. A., & Tarassenko, L. (2013). Data fusion for estimating respiratory rate from a single-lead ECG. *Biomedical Signal Processing and Control*, 8(1), 98–105.
11. Lian, J., Wang, L., & Muessig, D. (2011). A simple method to detect atrial fibrillation using RR intervals. *The American Journal of Cardiology*, 107(10), 1494–1497.
12. Murthy, V. K., Grove, T. M., Harvey, G. A., & Haywood, L. J. (1978). Clinical usefulness of ECG frequency spectrum analysis. In *Proceedings of the Annual Symposium on Computer Application in Medical Care* (pp. 610–612), November 1978.
13. Clifford, G. D., Behar, J., Li, Q., & Rezek, I. (2012). Signal quality and data fusion for determining the clinical acceptability of electrocardiograms. *Physiological Measurement*, 33, 1419–1433.
14. Jekova, I., Krasteva, V., Christov, I., & Abacherli, R. (2012). Threshold-based system for noise detection in multilead ECG recordings. *Physiological Measurement*, 33, 1463–1477.
15. Hamilton, P. S., & Tompkins, W. J. (1986). Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Transactions on Biomedical Engineering*, 33 (12), 1157–1165.
16. Zong, W., Moody, G. B., & Jiang, D. (2003). A robust open-source algorithm to detect the onset and duration of QRS complexes. *Computing in Cardiology Conference*, 30, 737–740.
17. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ch, Ivanov P., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220.

18. Moody, B. E. (2011). Rule-based methods for ECG quality control. In *Computing in Cardiology Conference 2011* (Vol. 38, pp. 361–363).
19. Langley, P., Di Marco, L. Y., King, S., Duncan, D., Di Maria, C., Duan, W., et al. (2011). An algorithm for assessment of quality of ECGs acquired via mobile phones. *Computing in Cardiology Conference*, 38, 281–284.
20. Johannesen, L. (2011). Assessment of ECG quality on an android platform. *Computing in Cardiology Conference*, 38, 433–436.
21. Hayn, D., Jammerbund, B., & Schreier, G. (2012). QRS detection based ECG quality assessment. *Physiological Measurement*, 33, 1449–1461.
22. Ho Chee Tat, T., Xiang, X., & Thiam, L. (2011). Physionet challenge 2011: Improving the quality of electrocardiography data collected using real-time QRS-complex and T-wave detection. *Computing in Cardiology*, 38, 441–444.
23. Orphanidou, C., Bonnici, T., Charlton, P., Clifton, D., Vallance, D., & Tarassenko, L. (2015). Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring. *IEEE Journal of Biomedical and Health Informatics*, 19(3), 832–838.
24. He, T., Clifford, G. D., & Tarassenko, L. (2006). Application of independent component analysis in removing artefacts from the electrocardiogram. *Neural Computing and Applications*, 15, 105–116.
25. Li, Q., Mark, R. G., & Clifford, G. D. (2008). Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiological Measurement*, 29(1), 15–32.
26. Allen, J., & Murray, A. (1996). Assessing ECG signal quality on a coronary care unit. *Physiological Measurement*, 17, 249–258.
27. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423, 623–656.
28. Daluwatte, C., Johannesen, L., Galeotti, L., Vicente, J., Strauss, D. G., & Scully, C. G. (2016). Assessing ECG signal quality indices to discriminate ECGs with artefacts from pathologically different arrhythmic ECGs. *Physiological Measurement*, 37, 1370–1382.
29. Orphanidou, C., & Wong, D. (2017). Machine learning models for multidimensional clinical data. In S. U. Khan, A. Y. Zomaya, & A. Assad (Eds.), *Handbook of large-scale distributed computing in smart healthcare, scalable computing and communications* (pp. 177–216). Cham: Springer.
30. Silva, I., Moody, G. B., & Celi, L. (2011). Improving the quality of ECGs collected using mobile phones: The PhysioNet/Computing in Cardiology Challenge 2011. *Computing in Cardiology Conference*, 38, 272–276.
31. Morgado, E., Alonso-Atienza, F., Santiago-Mozos, R., Barquero-Pérez, Ó., Silva, I., Ramos, J., et al. (2015). Quality estimation of the electrocardiogram using cross-correlation among leads. *Biomedical Engineering Online*, 14, 59.
32. Kalkstein, N., Kinar, Y., Na'aman, M., Neumark, N., & Akiva, P. (2011). Using machine learning to detect problems in ECG data collection. *Computing in Cardiology Conference*, 38, 437–440.

Signal Quality Assessment in Physiological Monitoring

State of the Art and Practical Considerations

Orphanidou, C.

2018, VI, 63 p. 23 illus. in color., Softcover

ISBN: 978-3-319-68414-7