# Event Detection and Signal Compression in Digital Electrocardiograms

by

Travis P. Blanchett

A Thesis Submitted to the
Faculty of Engineering
in Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Major Subject: Engineering Mathematics

DALHOUSIE UNIVERSITY — DALTECH

Halifax, Nova Scotia 1998

*To my sweetheart Terry–Lynne, my pal Ryan, and the new friend I will meet in a very few months.*

# Table of Contents

# List of Tables

# List of Figures

# List of Symbols and Abbreviations

| | |
|---|---|
| $\downarrow n$: | Downsample by $n$ operator |
| $\alpha$: | Lipshitz exponent |
| $\delta$: | Fraction of variance captured by approximation |
| $\epsilon_k^j$: | Quality control error measure for block $j$ and beat $k$ |
| $\psi$: | Wavelet function in discrete wavelet transform |
| $\phi$: | Alternate symbol for scaling function ( see $w$ ) |
| $\mu_j^m$: | Degree of membership of coefficient $m$ of subband $j$ in the fuzzy set *large* |
| $\mu_{in}$: | Inferred membership in the consequent fuzzy set interest |
| $\theta$: | Scaling function in discrete wavelet transform |
| $\sigma^2$: | Sample variance |
| $a_n$: | Approximation coefficients of discrete wavelet transform |
| **A**: | Beat matrix |
| $b$: | Quality controlled compression threshold for $\epsilon_k^j$ |
| $b^{\widehat{PQ}}$: | $\widehat{PQ}$ block row vector |
| $b^{\widehat{QRS}}$: | $\widehat{QRS}$ block row vector |
| $b^{\widehat{ST}}$: | $\widehat{ST}$ block row vector |
| $B_k$: | Beat row vector |
| $B_k^r$: | Resampled beat row vector |
| $C_{QC}$: | Compression rate with quality controlled based method |
| $C_V$: | Compression rate with variance based method |
| $d_j$: | Detail coefficients of wavelet transform |
| DISPO: | Digital smoothing polynomial filter |
| $\bar{e}^j$: | Centroid of $j$th membership event envelope |
| $e_f^j$: | Point at which a membership event envelope is finished |
| $e_s^j$: | Point at which a membership event envelope starts |
| FN: | False positive error |
| FP: | False negative error |

$\vec{E}_w$: Resultant vector of detail coefficients
ECG: Electrocardiogram
$g(k)$: Low pass filter in discrete wavelet transform
$h(k)$: High pass filter in discrete wavelet transform
KLT: Karhunen–Loeve transform
$L^2$: The 2-norm
$m_k$: The $k$th moment of a signal
$2M$: Order of DISPO filter
MAD: Mean absolute deviation
$n_k$: Integer address of $k$th singularity in a signal
$N_{PQ}$: Number of points in $\widehat{PQ}$ block
$N_{QRS}$: Number of points in $\widehat{QRS}$ block
$N_{ST}$: Number of point in $\widehat{ST}$ block
$\widehat{PQ}$: Block of ECG roughly associated with a PQ interval
$\widehat{QRS}$: Block of ECG roughly associated with QRS complex
$R^2$: Coefficient of determination
$R(k)$: List of integer R-wave addresses
RSS: Residual sum of squares
$S_j^n$: Raw data vector $j$ from record $n$ of MIT–BIH database
SDF: Estimated spectral density function
SVD: Singular value decomposition
$\widehat{ST}$: Block of ECG roughly associated with ST segment
SYY: Sample sum of squares
$T_{PQ}$: Sampling period in $\widehat{PQ}$ block
$T_{QRS}$: Sampling period in $\widehat{QRS}$ block
$T_{ST}$: Sampling period in $\widehat{ST}$ block
$w$: Alternate symbol for wavelet function ( see $\phi$ )
$W_s$: Wavelet transform operator at scale s
$W_{2^j}$: Wavelet transform operator at scale $2^j$
$y$: Signal
$\mathbf{y}$: Signal column vector
$\hat{y}$: model of signal vector over an interval
$\mathbb{Z}$: The set of integers

# ACKNOWLEDGMENTS

I am most fortunate for having had the opportunity to explore in the company of the DalTech Engineering Mathematics community . I am indebted to several individuals to such a conspicuous degree that their contributions must be recognized distinctly: Mike Lee for cultivating my interest in biological problems and for the realization that one of the difficulties in engineering is that there are too many engineers in it, Bill Phillips for having the patience to put up with the examples and ideas I would continually bombard him with, Peter Gregson for lively discussions and for including me in his lab group for a while (where there always seemed to be something exciting going on), and Gordon Fenton for his insight and pragmatism without which the quality controlled compression portion of this work would have suffered tremendously. Most importantly, I must express my gratitude to Guy Kember for taking the time to be involved with this work at a fundamental level, and for the congenial and highly productive coffee shop banter. Your influence and my appreciation can not be overstated.

Financial assistance for this work was provided by the Natural Sciences and Engineering Research Council of Canada's (NSERC) postgraduate scholarship program..

# ABSTRACT

Novel methods for event detection and signal compression are presented for the electrocardiogram (ECG). The event detection method uses a simple and fast algorithm which is highly effective. The method is based on the local and fuzzy evaluation of the size of Haar wavelet transform coefficients of the signal. The average error rate of the method was 0.68 % over the entire MIT–BIH database which is of the same order as previously published results. The method achieves this level of performance without operator interference, and with globally fixed parameters which is unique in the literature. The compression method is based on active error control, and is the first method which has successfully applied a local error measure to the electrocardiogram compression problem. A resampling strategy based on the physiology of the signal is used to achieve performance improvement and computationally feasible access to singular value decomposition. The method is simultaneously capable of higher compression and higher fidelity of the reconstructed approximation than previously reported results. The average compression rates achieved by the method are 27.1:1, and 15.4:1 for the 100 and 200 series of the MIT–BIH database respectively.

# Chapter 1

# Introduction

The action of the heart is associated with an electric field. An electrocardiogram (ECG) is a recording of the potential between two points within this field measured as a function of time, usually, from the surface of the body. Scalar values of voltage are generally recorded from several different positions on the surface of the body simultaneously. The points on the body where the measurements are taken from have been standardized and are referred to as lead placements [6]. Since the placements of the electrodes are known and fixed, the directional vectors associated with the voltages are also known. Using this recording, a great deal of information can be inferred about the cardiovascular health of a patient.

Electrocardiograms are measured in a variety of ways in order to evaluate different aspects of a patients cardiovascular health. Since the early 1960s ambulatory ECGs (or Holter recordings) have been used to measure two or three lead ECGs over a period of about 24 hours, and it is this type of ECG with which we are principally concerned. The use of a Holter monitor is the only method which allows the measurement of the ECG while a patient carries on with their normal daily activities and it has therefore become a widely used and important diagnostic tool [4]. A Holter monitor records the measured ECG signal on a tape or other storage medium. The interpretation of the recorded ECG typically requires that the signal be quantized and discretely sampled. The digital version of the signal can be displayed on a computer monitor, transmitted for remote interpretation, stored for future reference, or subjected to further processing using a multitude of digital signal processing techniques. A Holter monitor typically has other capabilities, such as a clock to add a time stamp to the recording and an event marker which allows a patient to record a flag at the onset of any perceived symptoms. Ambulatory ECGs remain relatively expensive and time consuming due to the dependence on skilled technicians for the interpretation of the

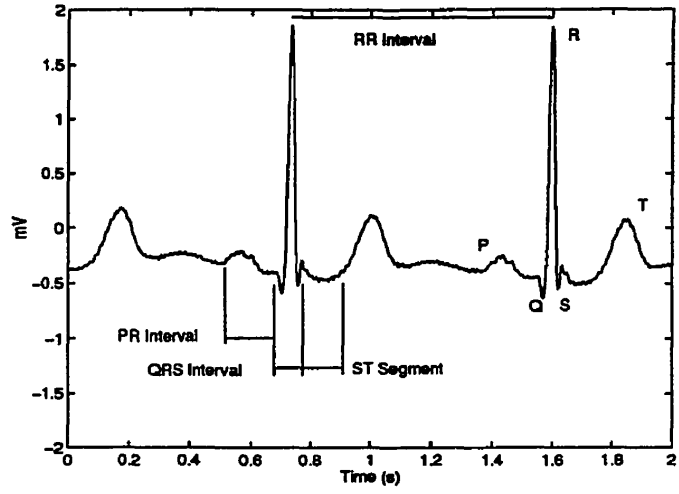


Figure 1.1: An example of the normal ECG extracted from record 103 of the MIT-BIH database showing important features and intervals with standard labels.

recordings and the unreliability of automated methods [4].

This study and others, including [2], [15], [7], and [9], have employed the MIT-BIH Arrhythmia database [14]. The MIT-BIH database is a collection of digitized Holter recordings which were produced for developing and testing automated techniques for ECG analysis. The ECG recording has been annotated with markers that describe the position and type of each beat, important changes in signal quality, changes in the prevailing heart rhythm, and comments. The recordings were initially annotated by two cardiologists independently. Discrepancies were subsequently resolved by consensus [12]. In figure 1.1 an excerpt from channel 1 of record 103 of the MIT-BIH database is shown. For reference, each visible feature in the cardiac cycle is shown with its standard clinical label. The time intervals of particular interest to this study are also displayed with their standard labels [16]. The cardiac cycle has three main constituents, atrial activation, ventricular activation and ventricular repolarization which are associated with the P-wave, QRS complex and the T-wave respectively. The time evolution of these signal features is, under normal circumstances, approximately fixed with respect to the QRS complex. That is, the order, the shape and the relative positions of these events are approximately constant within a normal beat's ECG.

## 1.1 Compression of the ECG

Our immediate concern is to determine an estimate for the amount of information that requires storage. In order to maintain the fidelity of the signal, it should be quantized with at least 10 bit precision and sampled at a rate on the order of 500 Hz [8]. This yields a minimum storage requirement of about 50 Mb for each lead per day of ECG recording. The large amount of data inherent in this kind of recording has generated a great deal of interest in means of compressing the signal. If a suitable compression technique could be developed, a great savings in storage costs could be realized. In addition, the potential convenience of storing the digitized signal directly in the Holter recorder presents many difficulties with respect to weight, size, storage capacity, and power consumption which could be reduced through effective compression.

The topic of ECG compression has been actively pursued for 40 years [8] and has generated a large literature. Of particular interest is irreversible or "lossy" compression due to the potential for much higher compression rates. The cost of a higher compression rate is the loss of some information in the signal. If this loss were restricted to noise and artifact, the cost would obviously be acceptable. Quantitative evaluation of the quality of reconstructed ECG signals has been almost exclusively based on ensemble averages of the mean squared error and related measures [2] [8] [19]. Ensemble averages obviously do not ensure clinically adequate, beat-by-beat reconstruction since a number of beats must exceed the error criterion. The persistence of this kind of evaluation is quite remarkable given the following:

1. An ECG is interpreted based on the relative character of features which are localized in time and which make widely differing contributions to the variance of the signal. Disturbances of these features which may have small influence on signal variance can be of clinical significance.

2. The localization of ECG features in time has lead to the application of localized bases such as wavelets to the ECG compression problem [19] [2].

Thus, research has been actively pursued in compressing a signal which is interpreted based on the relative changes of temporally localized features with a basis which, when truncated, generates compactly supported errors. However, the quality of the reconstructed approximation to the signal, and, therefore, the effectiveness of the algorithm, are evaluated using a globally averaged error measure. The use of such error measures for evaluating the quality of compressed ECG signals has previously

been recognized as being of no practical value [8].

Algorithms such as AZTEC, CORTES, Fan/SAPA, and Turning Point have been studied in [8] and [5] and have been shown to provide compression ratios of between 2:1 and 8:1 on data from the MIT-BIH database while failing to provide a clinically adequate representation of the signal. The wavelet based algorithm MULTIWAVE, introduced in [20], generated highly localized, and oscillatory errors in the ST-segment of a normal ECG. In addition, the QRS complex was visibly broadened and phase distortion was introduced in the signal through the use of a non-linear phase filter bank. These characteristics were expressed at a compression ratio of 16:1. A greater degree of success has been achieved using a combination of wavelets and a Karhunen–Loeve transform (KLT) [2], where the author reports errors based on a fraction of the original signals variance which was captured by the reconstruction. A compression rate of 21.4:1 is achieved on a subset of the MIT-BIH database and visually better reconstructions are claimed, however, the minimum amount of the variance captured by the reconstruction was only 96.5 % and intervention was frequently required by the author, in the form of threshold adjustments, in order to maintain an acceptable result [2].

## 1.2 Event Detection in ECG

In order to take advantage of transform based compression methods, the detection of individual beats is desirable. Great gains in compression can be obtained by taking advantage of beat to beat correlations [1] [2]. Detection of beats generally implies detection of the QRS complex. The QRS complex is the highest energy event in the cardiac cycle, and therefore is the easiest feature to detect. QRS detection is a fundamental step towards transform based compression methods, and also pattern recognition methods for automated ECG interpretation, because other events in the ECG are relatively fixed with respect to the QRS complex. As such, a large literature has been generated in this field.

Three key papers which have appeared in the literature concerned with the topic of QRS detection are [15], [7], and [9]. The authors report success rates of 99.33%, 99.46% and 99.85% respectively using the MIT-BIH database. At a fundamental level the algorithms in these papers are closely related. In [15] an algorithm is presented which uses a band-pass filter to improve the signal to noise ratio, followed by differentiating and squaring of the signal to enhance regions of relatively high slope. QRS complexes are then detected based on an adaptive threshold. In their algorithm a

strategy called *refractory blanking* is employed, where detections that occur during a fixed interval after a previously detected beat are ignored. This facet of the algorithm was justified based on the refractory period that the heart muscle experiences, after ventricular depolarization, during which no activation pulse can be conducted [16]. In addition, if no beat is detected for a length of time greater than a constant times the local average RR interval, a *search back* technique is used where a lower threshold is reapplied to the same interval in order that a supposedly missed beat can be discovered. If an event is detected between the end of the refractory blanking interval and 360 ms after a previously detected beat it is assumed to be a T-wave, unless a criterion is met with respect to its slope. In [7] a virtually identical algorithm is applied, however parameter optimization is performed to reduce the error rate.

In [9] a strategy based on adaptive thresholding of the detail coefficients of the signal's wavelet transform was introduced which is a modified version of the adaptive technique used in [15]. In this approach maximum–minimum pairs within a short window at each level of the decomposition were found. The existence of a zero crossing in the detail coefficients of the wavelet transform implies that there was a slope change in the event being examined, as would be expected of a QRS complex, and as would likely be absent from a rapid base line shift. In addition, an attempt was made to estimate the singular degree of the signal so as to exclude noise and artifact. The refractory blanking and search back techniques optimized in [7] for the MIT-BIH database were retained. Let $W_s$ be a continuous wavelet transform operator at scale $s$. Let our wavelet function, $\psi(x)$, be the first derivative of a smoothing function $\theta(x)$, that is:

$$\psi(x) = \frac{d\theta}{dx}. \tag{1.1}$$

Let $f(x)$ represent a function in the $L^2$ space. In [10] it was shown how if 1.1 holds that the zero crossings of the wavelet transform, $W_s f(x)$, correspond to the extrema in $f(x)$. The decay of the wavelet coefficients across small scales is a function of the Lipschitz regularity of the signal. Given the wavelet transform of $f(x)$, the following remarks are made:

1. When the scale $s$ is small enough, the maxima of $|W_s f(x)|$ indicate the location of sharp variation signal points.

2. The function $f(x)$ is Lipschitz $\alpha \in (0, 1)$ over $[a, b]$ iff there exists a constant $A$ such that $\forall x \in [a, b]$,

$$|W_{2^j} f(x)| \leq A 2^{j\alpha}. \tag{1.2}$$

In addition, these notions were recast in a discrete framework and a fast transform algorithm was provided [10].

Let $W_{2^j}$ be a dyadic wavelet transform operator at scale $2^j$, such that the wavelet transform of a signal, say $f(n)$, can be written in the following manner:

$$S_{2^j} f(n) = \sum_k h_k S_{2^{j-1}} f(n - 2^{j-1}k) \tag{1.3}$$

$$W_{2^j} f(n) = \sum_k g_k S_{2^{j-1}} f(n - 2^{j-1}k) \tag{1.4}$$

The wavelet coefficients were computed in [9] using equations 1.3 and 1.4 at levels $j = 1, 2, 3, 4$. The magnitude of the wavelet coefficients in each subband were evaluated using four separate adaptive thresholds. A list of modulus maxima which exceeded the threshold was generated for each subband and was denoted $n_k^j$ for levels $j = 1, 2, 3, 4$ where $n_k$ is the index of the $k$th event in the ECG signal. The R-waves in the ECG signal were then identified by searching through these lists of detected singular events, $n_k^j$, and finding points where local modulus maxima occurred in different subbands at approximately the same time. In [9] an estimate of the upper bound on the Lipschitz exponent, $\alpha'$, was then computed based on the decay of the modulus maxima of the wavelet coefficients, $|W_{2^j} f(x)|$, across subbands. Let $a_j(n_k) = |W_{2^j} f(n_k)|$. The estimate of $\alpha'$ was computed in [9] using:

$$\alpha' = \frac{\alpha_1 + \alpha_2}{2}, \tag{1.5}$$

where,

$$\alpha_j = \log_2 a_{j+1}(n^{j+1}) - \log_2 a_j(n^j). \tag{1.6}$$

A *sharp reduction* in $\alpha'$ from a singularity at $n_k$ to the next, at $n_{k+1}$, was used as evidence that the event at $n_{k+1}$ was noise or artifact. The mechanism by which a "*sharp reduction*" is evaluated is not specified in [9], and therefore the algorithm cannot be reproduced. In addition, because the wavelet coefficients are primarily dependent on the slope of the signal due to nature of the wavelet basis, the decay of the wavelet coefficients is dependent on signal regularity and is insensitive to its magnitude. Thus, in order to apply the Lipschitz analysis *significant* singularities must be preselected based on the size of the coefficients, which can be seen to be a source of a great deal of complexity.

Several serious flaws exist in the three methods. Most importantly, in each of the methods the operator initializes the thresholds, and the RR interval based on

data extracted form each record manually before analysis is commenced. Given that the ECG is only constrained within an individual, this intervention applies a critical constraint on the range of parameters to apply during the detection. In essence, each of the algorithms must be supplied with a local definition of a normal beat by a human interpreter in order to function. The use of this initialization procedure implies that these techniques are unstable and cannot maintain their reported accuracies in the presence of sudden changes in signal character. Because changes in signal character are common in ECG recordings [4] the requirement of this kind of manipulation is inappropriate.

All three methods employ search back techniques which are a fundamentally flawed approach to QRS detection. A search back technique can only be applied to improve the detection rate in a situation where the investigator knows *a priori* that each beat is closely followed by another, such as for example, in a fixed database. Obviously, in practice no such constraint exists.

In [15] and [7] the strategy of rejecting T-waves based on the occurrence of a detected event at a fixed interval beyond a detected beat has no basis in physiology. These algorithms used values of 200 ms and 360 ms for this strategy. Since the QT interval is a function of heart rate [4] a fixed interval can only be used on a restricted set of data.

In addition to fundamental flaws in the algorithms, the following specific points should be considered:

1. The parameters for search back and refractory blanking used in [9] and [7] were optimized over the entire MIT-BIH database. The database was subsequently used as the testing set.

2. Two minutes and 24 seconds of data were censored from record 207 of the MIT-BIH database in [7] in order to reduce the error rate of the method. During this time ventricular flutter waves were recorded on the ECG, which are not associated with a distinct QRS complex [4]. Ventricular flutter is considered a fatal arrhythmia [4] [16].

3. Records 214 and 215, representing 5% of the MIT-BIH database, are excluded from the analysis in [9] without explanation.

## 1.3 Goals of this research

The purpose of this research was to develop a QRS detector and a ECG compression algorithm. Since a transform based method was adopted for compression, a reliable QRS detector would be needed, if the compression algorithm was to be implemented in a practical setting.

In order to address some of the issues raised with respect to the prior art, the following constraints were imposed:

1. The QRS detection strategy was to be physiologically based, and autonomous with a high detection rate.

2. The ECG compression algorithm was to use a physiologically based error measure that forced the reconstructed signal to maintain its clinical usefulness while providing a high rate of compression.

Using an error measure which has clinical significance allows the use of *active* error management. That is, the error in the reconstruction is measured locally as the signal is compressed and is controlled so that a clinically acceptable representation is achieved. In [1] the use of such a measure is investigated. This work is extended to include the case of high fidelity reconstructions, which in many cases are indistinguishable from the original signal, while maintaining low bit rates over the MIT database.

The QRS detection and ECG compression algorithms discussed hereafter are both highly effective. Many of the capabilities of the algorithms had previously not appeared in the literature. For example, the compression algorithm uses an error measure which is capable of detecting localized departures from typical signal behavior and the QRS detector uses only 6 globally fixed parameters which were deduced from a small sampling of the database. Despite the radical departure from the prior art that is represented by these algorithms, they remain quite simple. For example, the QRS detection algorithm has only five steps.

The most critical aspect of this work is that together the two algorithms represent the first step to the solution of the much larger ECG pattern recognition problem. That is, the development of an algorithm capable of detecting pathological beats in an ECG automatically. Consider that the output of the ECG compression algorithm contains information related to localized departures from typical signal behavior. By definition, many pathological ECG signals fall into this category. Therefore, the algorithm provides critical information for the classification of each beat.

# Chapter 2

# Event Detection

The event we are interested in detecting during the cardiac cycle is the QRS complex. This problem is difficult because a number of confounding influences exist in the signal:

1. QRS complexes, although normally constrained within an individual, are realized in a extraordinary variety of shapes globally [see [16] for examples]. When pathological beats are considered it becomes clear that no precise general statements can be made with respect to the shape of QRS complexes.

2. Due to the variability in QRS complex shape, its spectral content can not be precisely specified.

3. When considering a single lead of the ECG, it should be remembered that the ECG is a result of a three-dimensional electric field projected on the axis of the ECG electrodes. The resulting null space of this projection means some pathological beats may not be well represented in a single lead recording.

4. Motion artifact and muscle noise can not be removed from the ECG by filtering because these constituents of the signal occupy the same regions of the spectrum as the features of interest — that is, QRS complexes, P-waves and T-waves [9].

## 2.1 Detection system

A method of detecting singularities has been developed based on a local rescaling of wavelet coefficients and a voting strategy that employs a fuzzy logic system. The envelopes of various events in the signal are detected, and the average behavior of

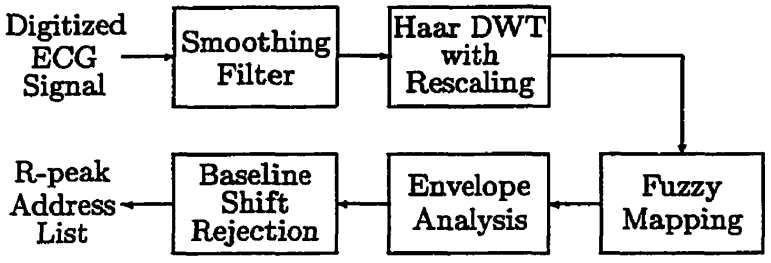


Figure 2.1: Block diagram of QRS detection system.

the event over the envelope is used to exclude artifact. Tracing through the block diagram in figure 2.1 we see the various operations that together constitute the QRS detection system. The signal is processed in segments of about 3s in duration. The input to the system is the raw ECG data extracted from the MIT-BIH CD-ROM, and the output is a list of integer addresses of the QRS complexes.

### 2.1.1 Digital Smoothing Polynomial filters

We will adopt two noise models for the ECG signal. First we use a white noise model for which we assume stationarity over a short time scale, and second is a correlated noise model for events such as rapid base line shifts. In this section we will restrict our attention to white noise. The purpose of this processing stage is to reduce the noise intensity so as to improve event resolution in the wavelet coefficients.

In order to reduce the influence of white noise on future processing, the ECG signal is convolved with a digital smoothing polynomial (DISPO) filter. This particular type of filter is equivalent to performing a least–squares fit of a polynomial of any desired order, say $2M$, to the data. The impulse and frequency responses of the fourth order 29 tap filter applied in this analysis are shown in figure 2.2. The filter has a symmetric, and finite, impulse response and therefore has zero phase. A DISPO filter has the following properties [3]:

1. A filter of degree $2M$, conserves every polynomial up to degree $2M + 1$.

2. For general filters of a given length satisfying item 1, the DISPO filter is optimal with respect to noise reduction, if the noise is stationary and white.

3. A DISPO filter of degree $2M$ will exactly conserve all existing moments in a signal up to $2M + 1$.

Let $m_k$ be the $k$th moment of a signal $x(n)$ of length $N$. We can compute $m_k$ using:

$$m_k = \sum_{n=0}^{N-1} x(n)n^k \tag{2.1}$$

By inspection we see that the preservation of $m_0$ requires the integrated area of the signal to be conserved, and therefore, a peak may be shortened if it is concurrently broadened. Similarly, we see that the preservation of $m_1$ requires the centroid of the signal to remain unchanged, which implies the filter must have zero phase. However, in order to preserve $m_2$ the width of features must also be preserved in the signal, and thus, the time scale of features must be conserved. Since it is our object to discover singular features in the signal, this characteristic is an enormous benefit. The key to the filtering operation is the reduction of noise while preserving the majority of signal information. In order to accomplish this result, the length and the order of the DISPO filter must be tuned to the underlying signal's features. In [17] it is recommended that for a filter of degree 4, the number of coefficients in the filter should be within 1 to 2 times the width of the features that are to be preserved at half the feature's maximum height. In this application the features to be preserved are QRS complexes which have a mean duration of about 80 ms [4] for normal beats. If we assume that the shape of a QRS complex can be grossly approximated with a triangle, a filter width of 360Hz $\times$ 0.080$s$ $\approx$ 29 samples. The normal QRS complex is at the lower bound of duration of these events, and the DISPO filter will have a smaller influence, that is, produce less of a smoothing effect, on broader features in the signal [17].

The effect of filtering record 108 of the database is shown in figure 2.3. This record is difficult for QRS detection due to the low amplitude of the R-waves, and the approximately 250 ms PR interval. An annotation is provided in figure 2.3 to allow distinguishing between the P-waves and R-waves. Inspection of the residuals in figure 2.3 shows the signal has been well preserved, while the high frequency noise has been attenuated. The widths of the QRS complexes and P-waves have also been preserved as expected. Given the initially low energy of the QRS complexes, preservation of the signal was particularly important in this case in order to avoid rendering them undetectable. In figure 2.4 the effect of the filter on the reconstructed level 1 detail subband of the signal's Haar wavelet transform is shown. This subband is sensitive to high frequency noise due to the small scale over which it operates, as shown in the lower left plot in the figure. In the lower right plot, the resolution of features in the

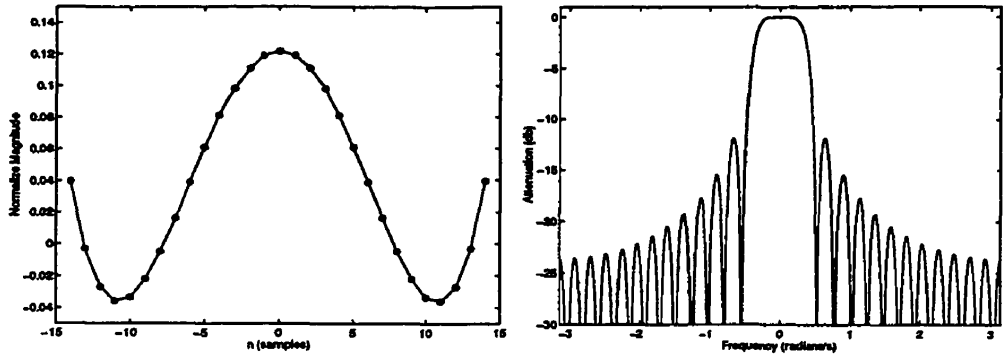

Figure 2.2: Impulse and frequency responses of the fourth order 29 tap DISPO filter.

signal is dramatically improved because of the action of the filter. The Haar wavelet transform is discussed further in section 2.1.2.

### 2.1.2 Haar wavelet transform

The DISPO filtered ECG signal is subjected to a 3-level Haar discrete wavelet transform (DWT). The transform is implemented as a filter bank and uses a downsampling operator at each level — that is, the odd numbered coefficients are discarded, reducing the total number of coefficients in each channel by a factor of 2. Tracing through the filter bank arrangement shown in figure 2.6, we see the filters are applied at each stage followed by downsampling. If the filters are chosen correctly the signal can be exactly recovered from these coefficients [19]. Each stage of decomposition involves the computation of two sets of coefficients, which are simply the output of the filtering operation. One set of coefficients represents the high frequency information in the signal at the scale being considered, and is associated with the high pass filter $g(k)$. These coefficients are referred to as the detail coefficients. The other channel is associated with the low pass filter $h(k)$ and produces a set of approximation coefficients related to the low frequency signal behavior at the scale being considered. We can see that these operations are applied to the approximation coefficients from the previous level, where at level 1, the sampled signal is assumed to be equivalent to the approximation coefficients at level 0. The action of the filter bank is equivalent to projecting the signal on the scaling function and wavelet at dyadic time scales [19]. The scaling function and wavelet from the Haar filter bank are shown in figure 2.5. This implementation is different from that expressed in equations 1.3 and 1.4 where no downsampling is used [10]. Let the transform coefficients for an $n$ level transform

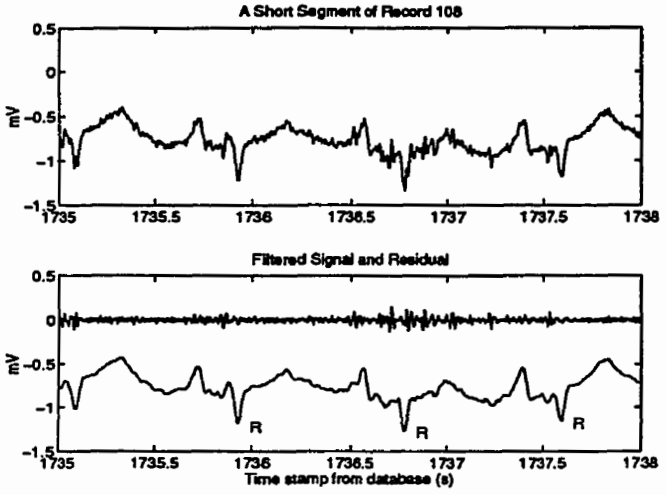


Figure 2.3: The effect of the fourth order, 29 tap, DISPO filter on record 108 of the MIT–BIH database.

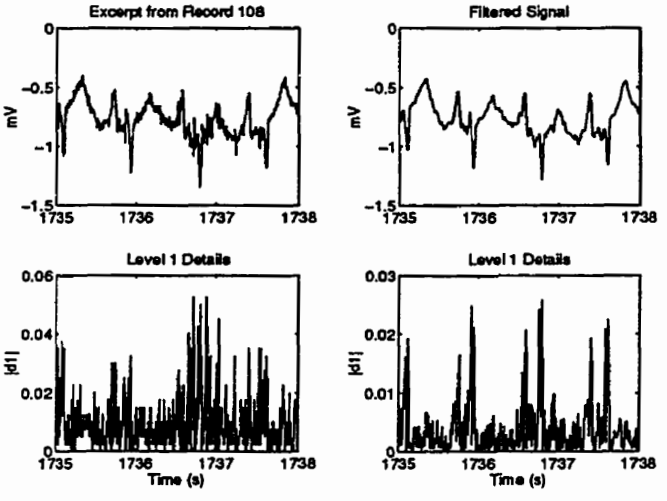


Figure 2.4: The effect of DISPO filtering on the reconstructed level 1 detail subband of record 108 of the MIT–BIH database.

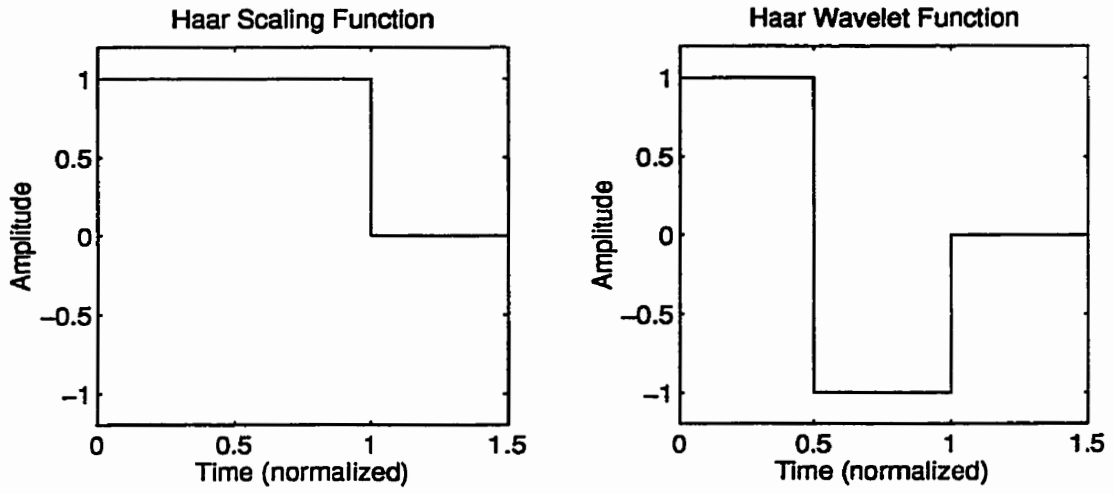


Figure 2.5: The Haar wavelet and scaling function are shown at unity scale. For the first level of decomposition in the filter bank, this scale is simply the sampling rate of the function. The scale increases in a dyadic fashion as higher levels of the decomposition are considered as a result of the downsampling operators.

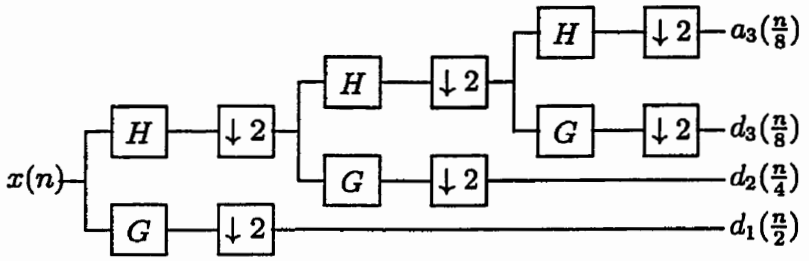


Figure 2.6: Three level acausal filter bank representation of the DWT.

be $\{a_n, d_j\}$, where $j = 1, 2, \ldots, n$, as shown in figure 2.6. In the Haar case $g(k)$ is orthogonal to $h(k)$ and simply represent the normalized averages and differences of adjacent points, specifically, the filter coefficients are:

$$h(k) = \left\{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\} \tag{2.2}$$

$$g(k) = \left\{\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right\} \tag{2.3}$$

The accuracy to which the signal can be recovered from the approximation coefficients, $a_j$, is dependent on the smoothness of the signal and the approximation order, typically denoted $p$ in the literature, of the scaling function [19]. That is, a high approximation order and a smooth signal leads to an accurate approximation and rapidly decreasing transform coefficients. Around singular points in the signal we expect the detail coefficients to play a greater role. For the Haar scaling function $p = 1$ [19] and we expect the accuracy of the approximation to be poor. More precisely, the approximation order is the lowest order polynomial which cannot be exactly represented by the scaling function [19]. Thus, the Haar scaling function cannot exactly represent a linear function, but it can represent a constant over any interval.

A numerical example of the properties of the Haar wavelet transform is displayed in figure 2.7. In the figure, the subbands for each level have been reconstructed by first zeroing all coefficients except those of the subband in question and then inverting the transform. Inspection of each subband shows how the energy in the detail subbands is related to the slope of the signal. In this example, normal and pathological beat types are present which allows demonstration that the detail coefficients of a QRS complex generally stand out against the local background at these scales, and how this fact must be evaluated locally due to non-stationarity in the signal.

In figure 2.8 a similar analysis is performed using a Daubechies 10 wavelet. The scaling function of this wavelet has an approximation order of $p = 10$. The figure shows that the accuracy of $a_3$, the approximation at level 3, is quite irregular due to its dependence on the smoothness of the signal through the QRS complex. Hence, in order to detect QRS complexes using a wavelet which is a good approximator of the signal being considered, we must include higher levels of decomposition in the analysis to ensure that representation of each QRS event exists in the detail coefficients. In figure 2.8 we see that where this representation occurs will remain uncertain due to changes in signal regularity. In figures 2.8 and 2.7 we see that up to level 3 the details are free of information from P-waves or T-waves. At higher levels of decomposition these features will begin to influence the detail coefficients. Therefore,

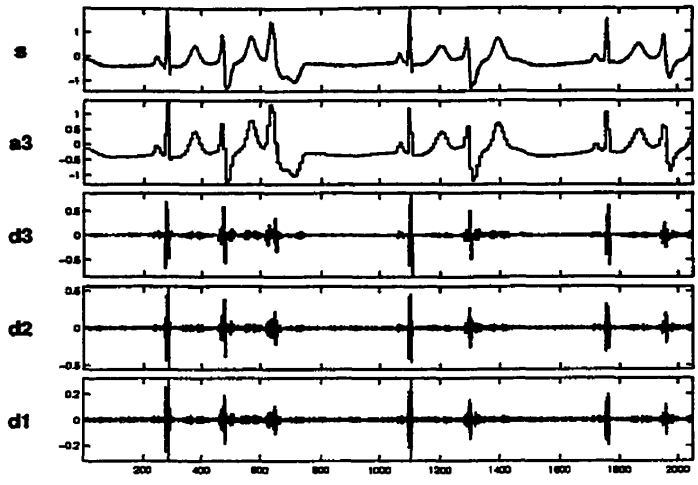


Figure 2.7: Three level Haar wavelet decomposition of an ECG signal from record 106 of the MIT-BIH database.

in order to employ higher levels of decomposition, the detection strategy must be able to distinguish between various types of deterministic information in order to correctly identify the QRS complexes. By reducing the accuracy of the approximation we force more signal information into the small scale detail coefficients around singular points, while maintaining insensitivity to P-waves and T-waves. This decoupling of the signals information allows the use of a simple detection strategy.

The representation of noise in the detail coefficients is of great importance. It was recognized in [9] that the use of level 4 detail coefficients provided a degree of isolation of the features from white noise. Inspection of the frequency response of the Haar filters in figure 2.9 reveals that the small scales cover the majority of the the spectrum, and therefore contain the majority of the white noise power. If we assume the detail coefficients collect half of the white noise power at each level of decomposition, then the details from levels 1 through 3 contain $\frac{7}{8}$ of the white noise power of the raw signal. We can retain the attractive capability of the small scale details to decouple the signals features and decrease the influence of white noise through the use of a prefilter as discussed in section 2.1.1.

### 2.1.3 Local Rescaling of the Wavelet Coefficients

In the previous section we identified that it is the local size of the coefficients which is of interest due to noise and signal non-stationarity. Consider the detail coefficients as

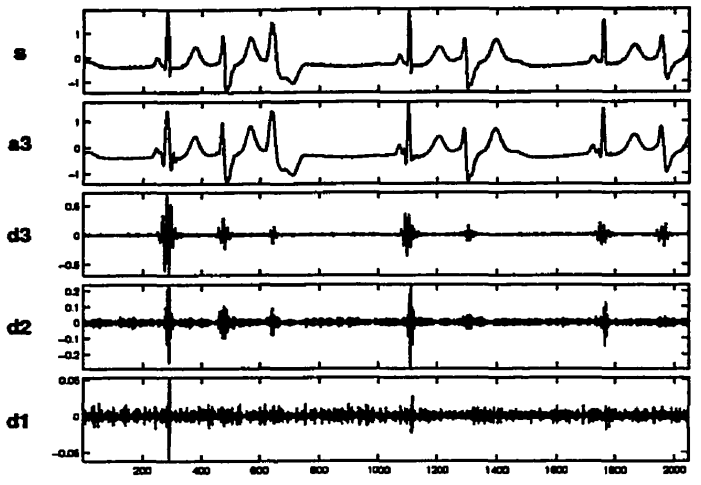


Figure 2.8: Three level wavelet decomposition of an ECG signal from record 106 of the MIT-BIH database. Here a Daubechies 10 wavelet has been applied. The scaling function associated with this wavelet has an approximation order of 10. That is, the scaling function can exactly represent a polynomial of order less than 10.

a set of rectangular coordinates. Let $\vec{E}_w$ be the resultant vector of these coordinates, that is:

$$\vec{E}_w(n) = \langle d_3(\frac{n}{8}),\ d_2(\frac{n}{4}),\ d_1(\frac{n}{2}) \rangle \tag{2.4}$$

The time series $\vec{E}_w(n)$ can now be considered a trajectory through a three dimensional space. Given a white noise input, the trajectory will be confined to a zero mean ellipsoid with major axis lengths and orientation determined by the statistical properties of the data. For a correlated disturbance, such as a QRS complex, relatively large coefficients occur across the subbands, as seen in figure 2.7. In order to simplify the analysis of this ellipsoid, the coordinates of the vector $\vec{E}_w(n)$ are rescaled by their local mean absolute deviation (MAD), given by:

$$\text{MAD}_j = \frac{2^{j-1}}{k} \sum_m |d_j(m) - \overline{d}_j| \tag{2.5}$$

where: $m \in \mathbb{Z}$, $0 \leq m < \frac{k}{2^{j-1}}$, $\overline{d}_j$ is the sample mean, and $k$ is the length of the original signal vector, we have,

$$d_j^*(m) = \frac{1}{\text{MAD}_j} d_j(m),\ j \in \{1,2,3\} \tag{2.6}$$

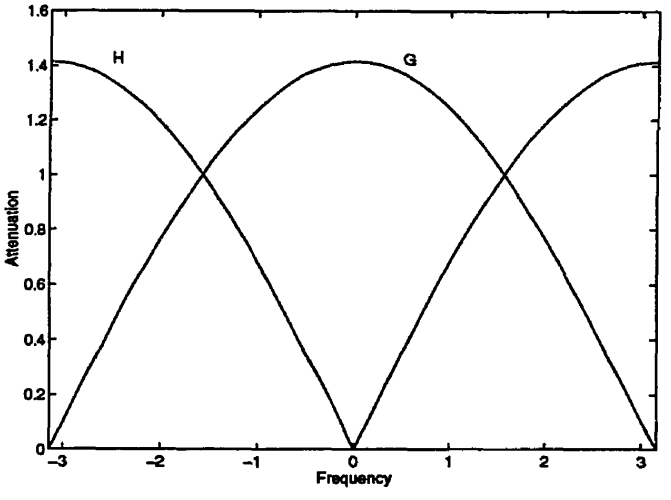


Figure 2.9: Frequency response of the filters $g(k)$ and $h(k)$, which are applied at each level of decomposition. If we assume the white noise power is split evenly between each subband at each level of decomposition, then the majority of white noise power can be seen to be captured in the small scale subbands, $d_j$ $j = 1, 2, 3$, because these subbands span the majority of the spectrum. This characteristic imposes the requirement of prefiltering in order to increase the resolution of signal features as discussed in section 2.1.1.

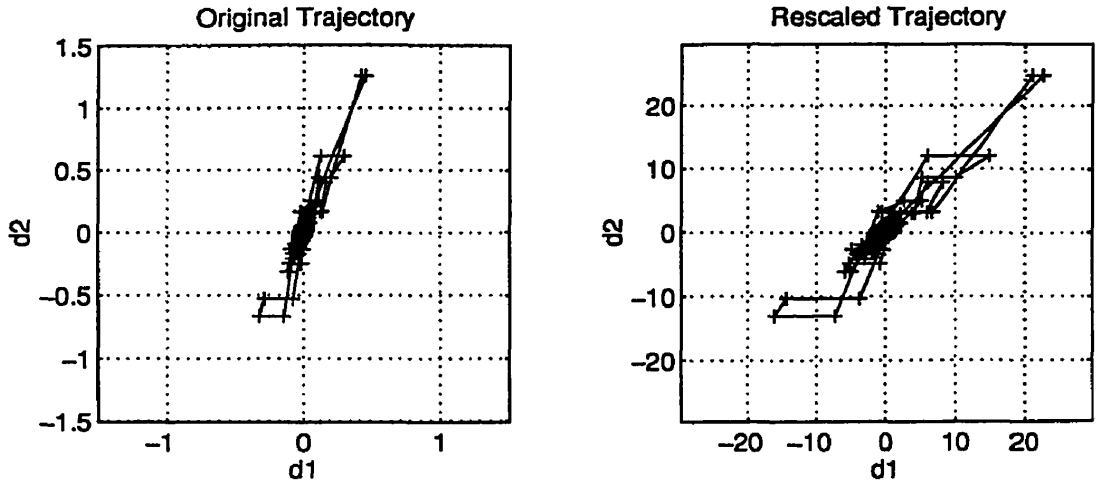


Figure 2.10: The trajectory of $\vec{E}_w(n)$ projected on the $d_1$–$d_2$ plane for a 1024 point segment of record 106 of the MIT–BIH database before and after rescaling of each coordinate by its local mean absolute deviation.

Because the majority of data points in the signal are not in QRS complexes, and due to the relative insensitivity of the MAD to outliers, the noise ellipsoid is made more spherical, and the axes of the entire distribution are scaled to lines of unity slope. An example of this process is shown in figure 2.10 where $d_1(\frac{n}{2})$ is plotted against $d_2(\frac{n}{4})$ before and after rescaling. This plot represents the projection of the trajectory $\vec{E}_w(n)$ on the $d_1$–$d_2$ plane. Before rescaling we see that the coefficients from the $d_1$ subband are smaller than the coefficients from the $d_2$ subband by a factor of about 3. In order to distinguish between signal and noise in this situation, each subband would have to be evaluated with separate adaptive thresholds, as has been proposed in [9]. Rescaling permits treating the magnitude of coefficients from different subbands uniformly and eliminates the need for adaptive thresholding.

### 2.1.4 Fuzzy Mapping

The persistence of physiological features over time is a critical piece of evidence which will allow us to separate these features from noise. If the signal is white noise then each sampled point will be a statistically independent event. Because the filters are of length 2 and are succeeded by a downsampling operator, and the filters are orthogonal, the transform coefficients will remain uncorrelated in this situation. Therefore an event which produces locally large coefficients ( that is, outliers ) across more than

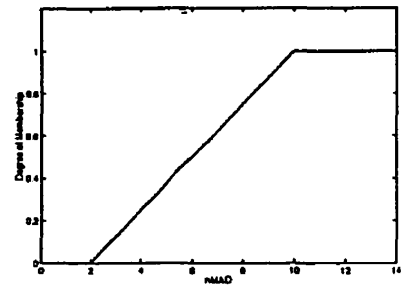

Figure 2.11: The global fuzzy set *large* used to measure coefficient relative sizes.

one subband and which furthermore is of long duration would have a vanishingly small probability of being associated with white noise.

Given our previous arguments, we will accept a coefficient as being of interest if it is concurrent with at least one other large coefficient in a different subband. We do not specify which, of the three subbands, these coefficients need be in. Since the coefficients are previously rescaled by their local mean absolute deviation, we can evaluate how big they are by fuzzifying them with a single fuzzy set *large*. Inspection of figure 2.11 shows that coefficients larger than twice the MAD of the sample being considered are to some degree *large*. Coefficients above 10 × MAD are considered *large* to degree 1.0. The fuzzifying operation will return a degree of membership in the fuzzy set on the interval $[0, 1]$. Let $\mu_j^m$ be the degree of membership of a coefficient from the subband at level $j$. Let $m = \frac{n}{2^j - 1}$ where $n \in \mathbb{Z}$ is a sample index over the segment of the ECG signal being considered and $j \in 1, 2, 3$. Using fuzzy logic operations, we can express our rule as:

**IF** the median coefficient is *large* **THEN** the singular point is of *interest*

Let the degree of membership in the fuzzy set *interest* be $\mu_{int}$, then:

$$\mu_{int}(n) = \text{median}(\mu_1^m, \mu_2^m, \mu_3^m) \tag{2.7}$$

The median operator itself is not a fuzzy operator, but it can be shown to be constructible from a combination of *min* and *max* operators, which in the case of three numbers reduces to the minimum of the two largest coefficient membership values. In this context the median is a *fuzzy voting strategy*. Each coefficient gets one "vote". A single large coefficient is insufficient evidence to show interest in a region, and at the same time a single zero coefficient is insufficient evidence to show disinterest.

The results of applying this measure to the first two seconds of records 103 and 203 are shown in figure 2.12. The segments of both records shown in figure 2.12

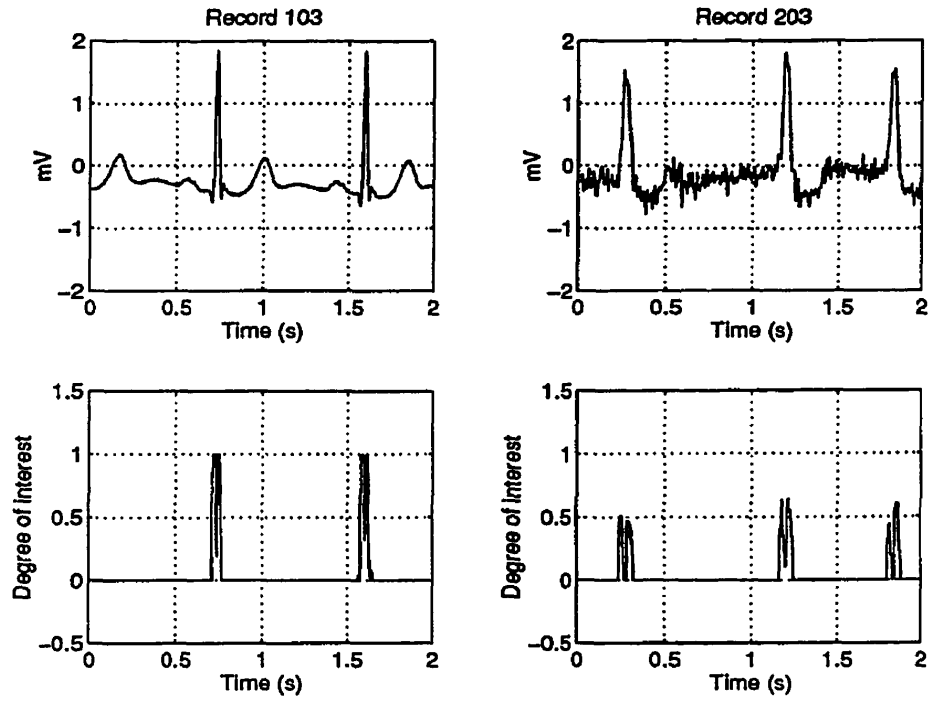


Figure 2.12: Magnitude of $\vec{E}_w(n)$ or, equivalently, the degree of interest $\mu_{int}(n)$, measured using the fuzzy voting strategy is shown for the first two seconds of records 103 and 203 of the MIT-BIH database with the corresponding ECGs.

have QRS complexes with physiologically normal duration. The segment of record 203 is noisy which results in a lower degree of interest in the QRS complexes. This phenomenon is an expression of the uncertainty that has been created by inferring our interest in a region contaminated with noise. Inspection of figure 2.12 reveals that the response of the measure $\mu_{int}(n)$ is zero for all features unrelated to QRS complexes in these examples.

### 2.1.5 Membership Envelope Analysis

The membership time series $\mu_{int}(n)$ generated by equation 2.7 tends towards unity around singular regions in the ECG signal. The kinds of singularities that cause this behavior are context dependent due to the local rescaling and can include deterministic information from the ECG signal such as P-waves and T-waves, or artifacts such as rapid baseline shifts. Hence, it is not possible to pick a threshold which will allow the global discrimination of these events due to the variable nature of the ECG sig-

nal. Our approach is to use a low threshold (typically 0.05) and to use the following algorithm to mark out the envelope of each singular region in the signal.

1. Search forward through the signal until the threshold is exceeded.
2. Set marker $e_s^j$ at the beginning of the singular region of event $j$.
3. Evaluate the length of the interval over which the time series $\mu_{int}(n)$ exceeds the threshold. Because the envelopes can be notched due to the slope change of a QRS complex (see figure 2.12). Therefore, based on the visual inspection of a few seconds each of data from several records of the MIT-BIH database we will accept a brief gap in the envelope of 30 ms during this evaluation.
4. Set marker $e_f^j$ at the end of the singular region.
5. If the region is less than 3 samples in length (about 8 ms) it is rejected as not being of significance. This constraint is imposed by the use of a linear model of the data in future processing (see section 2.1.6).
6. Compute the centroid of the event's envelope, thus for event $j$ we have:

$$\overline{e}^j = e_s^j + \frac{\sum_{n=e_s^j}^{e_f^j} n\mu_{int}(n)}{\sum_{n=e_s^j}^{e_f^j} \mu_{int}(n)}. \tag{2.8}$$

7. Test the distance between the current event's centroid and that of the previous event's. If the distance falls below 200 ms, join the two envelopes and recompute the centroid. That is, set $e_s^j = e_s^{j-1}$ and repeat step 6. This approach is justified by the absolute refractory period that the heart experiences after a beat. We accept that by definition this threshold rejects any heart rate above 300 beats per minute.
8. Store the final values of the centroid $\overline{e}^j$ and of the envelope brackets $e_s^j$ and $e_f^j$.
9. If not at end of signal return to step 1.

In figure 2.13 the analysis of a challenging section of record 105 of the MIT–BIH database is shown. This section was chosen in order to display the behavior of the fuzzy voting strategy in the presence of simultaneous artifact, variable QRS complex amplitude and duration, noise, and baseline drift. The labeled artifact at $t = 1202.5$ s is conspicuously difficult and is eventually declared a QRS complex resulting in a

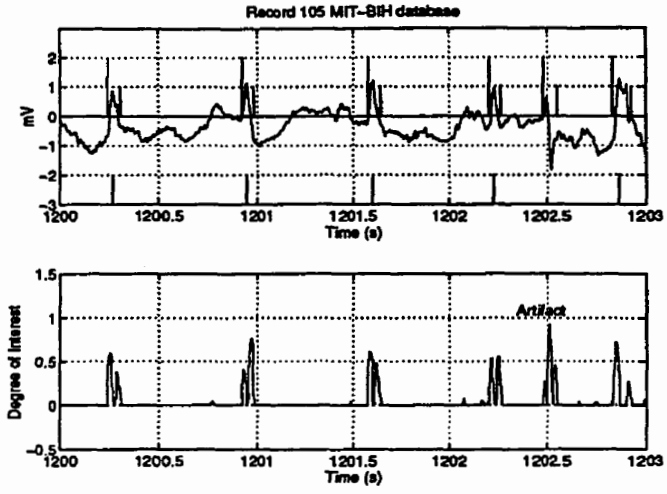


Figure 2.13: The top graph shows record 105 ECG, with the envelope markers $e_s^j = 2.0$ and $e_f^j = 1.0$, and R-peak markers generated using the MIT-BIH database annotations. The lower graph shows the associated measure $\mu_{int}(n)$.

false positive error. The envelope of each singular region of interest has been detected successfully in these examples.

### 2.1.6 The Detection of Rapid Baseline Shifts

Rapid baseline shifts are a frequent artifact in ambulatory ECG signals. The difficulty with these artifacts is that their representation in the time series $\mu_{int}(n)$ can easily equal that of a QRS complex. If we consider the wide variation in magnitude of event envelopes, particularly if pathological beats are included, which may not be well represented in the studied lead, it becomes clear that more information is required to reject these artifacts.

Examination of the rapid base line shifts shown in figure 2.14 at $t = 1205.0$ s and $t = 1204.4$ s over their event envelopes suggests a feature which distinguishes them from a QRS complex. Specifically, a rapid base line shift has a *linear* character over its event envelope. In contrast a QRS complex has a grossly *triangular* structure. We are, therefore, concerned with the average behavior of the signal over an event envelope. Previously, we wished to detect the singular regions using the detail coefficients. This allowed the algorithm to operate in an environment that was independent of the local mean signal intensity, that is, the strategy was insensitive to whether a QRS complex occurred at a baseline of -2 mV or 3 mV. Once the singular regions have been

isolated, more information is required to evaluate the character of each singularity. We assert that this characterization should be accomplished by the use of a model of the singularity. In the case of a QRS complex, the model cannot be fixed, however, it is possible to model what is not a QRS complex in the important case of rapid baseline shift. Consideration of this discussion yields the following heuristics:

1. In order for a region to correspond to a detectable event envelope it must to some extent be singular.

2. If a region is singular and is not linear over its event envelope, it cannot be rejected as artifact or noise due to the wide variety of morphologies of QRS complexes.

In order to incur a tremendous computational convenience we simply assume that a rapid baseline shift will be very near a straight line that passes through the two points where the event brackets intersect the filtered ECG signal. Such an event occurs at $t$ =1205.25 s in figure 2.14. We evaluate the validity of the model by measuring the fraction of the variance of the data that it explains using the *coefficient of determination*, $R^2$ [21]. It should be remembered that the model applied here does not minimize the $L^2$-norm of the residuals, which would be typical if a least squares model was fit to each envelope. Let $y(n)$ be the measured ECG signal, and $\hat{y}^j(n)$ be the linear base line shift model through the event envelope $j$ which is defined over the domain $e_s^j \leq n \leq e_f^j$. Specifically, we compute:

$$RSS_j = \sum_{e_s^j}^{e_f^j} (y(n) - \hat{y}^j(n))^2 \tag{2.9}$$

$$SYY_j = \sum_{e_s^j}^{e_f^j} (y(n) - \overline{y}^j)^2 \tag{2.10}$$

where $\overline{y}^j$ is the sample mean over event envelope $j$, then:

$$R_j^2 = 1 - \frac{RSS_j}{SYY_j}. \tag{2.11}$$

If over any event envelope $R_j^2$ exceeds the threshold of 0.9, we say the event is a rapid linear base line shift and exclude that event from the analysis.

The result of applying the discussed QRS detection strategy is shown in figure 2.15 for a short segment of record 105 of the MIT–BIH database. This segment has a small

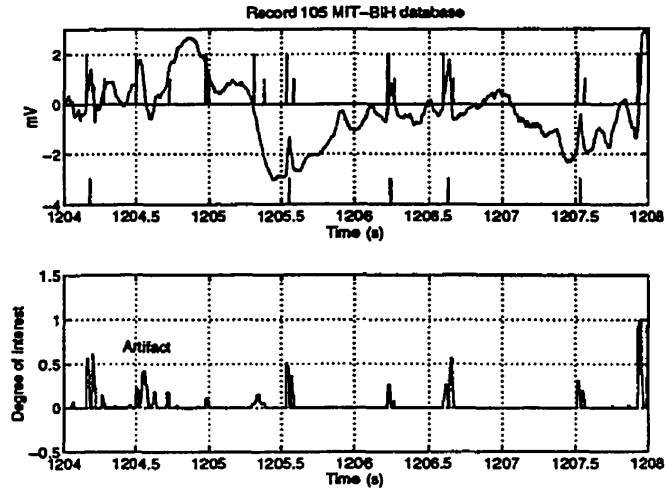


Figure 2.14: The top graph shows record 105 ECG, with the envelope markers $e_s^j = 2.0$ and $e_f^j = 1.0$, and R-peak markers generated using the MIT-BIH database annotations. The lower graph shows the associated measure $\mu_{int}(n)$. Of particular interest is the influence of rapid base line shifts in the analysis.

amount of noise and an irregular baseline. The top most trace of the figure shows a marker at each R-wave location taken from the database annotation. The lower most trace shows the R-wave locations computed using the QRS detection algorithm. Of interest is the extremely accurate localization of the R-waves.

## 2.2 Results

The algorithm used a total of 6 globally fixed parameters for the study performed on the MIT-BIH database. These parameters are summarized in the table 2.1. Because the algorithm is highly effective at separating regions of the signal associated with QRS complexes from those that are not, it was possible to deduce the values of the parameters based on the performance of the algorithm over several short segments of data extracted from the data base. The segments of data were either two or three minutes long and were extracted from records 103, 105, 108, 113, 201, 203, 232, and 234 of the MIT–BIH database and had a total length of less than 20 minutes. These examples represent less than 1.5 % of the total amount of data in the database.

The MIT-BIH database contains 109986 annotated QRS complexes. Channel 1 of all 48 records of the MIT-BIH database were analyzed. The annotations associated

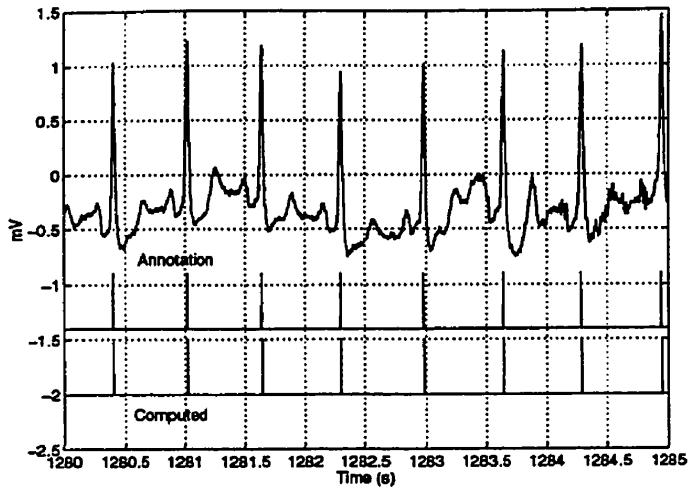


Figure 2.15: An excerpt from record 105 of the MIT-BIH database is shown. The trace labeled 'Annotation' was constructed from the R-wave locations supplied with the database. The lower trace represents the R-wave locations generated using the discussed QRS detection algorithm and is labeled 'Computed'.

Table 2.1: Summary of the parameters used in the QRS detection algorithm.

| Parameter Name | Value |
|---|---|
| Joining Interval | 200 ms |
| Maximum notch width | 30 ms |
| Low saturation for fuzzy set | 2 × MAD |
| High saturation for fuzzy set | 10 × MAD |
| Global threshold | 0.05 |
| Minimum $R^2$ for linear baseline shift rejection | 0.90 |

with non-conducted P-waves, and ventricular flutter waves were not counted as beats, because no QRS complex is associated with these events. That is, if the algorithm detected one of these excluded events, that detection would be counted as a false positive error.

The program was executed in approximately 27 minutes on a DEC Alpha 3000 computer. This time includes all file handling overhead. Each record contains 650000 samples for each lead quantized at 11b precision. The total size of the database can be seen to be 48 records $\times$ 650000 samples $\times$ 11 bits / 8 bits/byte / 1024 bytes/kbyte or about 42000 kbytes of data per lead. The file handling overhead for this amount of data is obviously non-trivial. The time required to load the data from the CD-ROM was measured for the first 10 records of the database. The mean loading time for the records was about 17 s with the minimum time at about 15 s and the maximum about 18 s. The value of 17 s for each record translates to about 13.5 minutes, or half, of processing time spent on loading data. In addition, optimizing the program for speed remains possible. For example, the prefiltering stage is implemented as a time domain convolution, rather than the potentially faster FFT based overlap and save strategy [18].

The evaluation of errors was based on the annotated QRS positions supplied with the data. Each detected beat was assigned to the closest neighboring beat in the annotated list. A false positive was said to have occurred in the event that more than one beat was assigned to an annotation. The number of false positives is equal to the total number of beats assigned to an annotated position less 1. A false negative was said to have occurred in the event that no detected beat was assigned to an annotated position. The total number of false positives recorded was 528, and the total number of false negatives was 224. This yields an average success rate of 99.32%.

Comparison of the artifact in figure 2.14 at $t = 1204.5$ s and the ventricular contractions present in the top trace of figure 2.7 at samples 500 and 1300 demonstrates why the error rate using the MIT-BIH database, and in general, cannot be zero using a single lead ECG. The artifact which is not an annotated beat in the database is virtually identical in shape, amplitude, and duration to the annotated ventricular contractions in figure 2.7 (note, for convenience 100 samples at 360 Hz is about 275 ms). It can be seen in figure 2.14 that no evidence is present in the single lead ECG which would allow the rejection of this artifact.

Table 2.2: Results of the QRS detection study on the MIT–BIH database.

| Record | No. QRS complexes | FP (beats) | FN (beats) | FP+FN (beats) | Total Errors (%) |
|---|---|---|---|---|---|
| 100 | 2273 | 0 | 0 | 0 | 0 |
| 101 | 1865 | 4 | 0 | 4 | 0.21 |
| 102 | 2187 | 0 | 0 | 0 | 0 |
| 103 | 2084 | 0 | 0 | 0 | 0 |
| 104 | 2230 | 15 | 5 | 20 | 0.90 |
| 105 | 2572 | 13 | 13 | 26 | 1.01 |
| 106 | 2027 | 8 | 1 | 9 | 0.44 |
| 107 | 2137 | 0 | 8 | 8 | 0.37 |
| 108 | 1763 | 58 | 3 | 61 | 3.46 |
| 109 | 2532 | 0 | 5 | 5 | 0.20 |
| 111 | 2124 | 2 | 1 | 3 | 0.14 |
| 112 | 2539 | 0 | 0 | 0 | 0 |
| 113 | 1795 | 2 | 1 | 3 | 0.18 |
| 114 | 1879 | 8 | 2 | 10 | 0.53 |
| 115 | 1953 | 0 | 0 | 0 | 0 |
| 116 | 2412 | 1 | 16 | 17 | 0.70 |
| 117 | 1535 | 0 | 0 | 0 | 0 |
| 118 | 2275 | 2 | 0 | 2 | 0.09 |
| 119 | 1987 | 76 | 0 | 76 | 3.82 |
| 121 | 1863 | 2 | 2 | 4 | 0.21 |
| 122 | 2476 | 0 | 0 | 0 | 0 |
| 123 | 1518 | 0 | 0 | 0 | 0 |
| 124 | 1619 | 1 | 0 | 1 | 0.06 |
| 200 | 2601 | 8 | 3 | 11 | 0.42 |
| 201 | 1963 | 136 | 2 | 138 | 7.03 |
| 202 | 2136 | 4 | 2 | 6 | 0.28 |
| 203 | 2980 | 5 | 69 | 74 | 2.48 |
| 205 | 2656 | 0 | 22 | 22 | 0.83 |
| 207 | 1860 | 68 | 5 | 73 | 3.92 |
| 208 | 2955 | 5 | 13 | 18 | 0.61 |
| 209 | 3004 | 5 | 0 | 5 | 0.17 |
| 210 | 2650 | 4 | 10 | 14 | 0.53 |
| 212 | 2748 | 2 | 0 | 2 | 0.07 |
| 213 | 3251 | 0 | 5 | 5 | 0.15 |
| 214 | 2261 | 4 | 3 | 7 | 0.31 |
| 215 | 3363 | 0 | 0 | 0 | 0 |
| 217 | 2208 | 1 | 7 | 8 | 0.36 |
| 219 | 2154 | 0 | 0 | 0 | 0 |
| 220 | 2048 | 0 | 0 | 0 | 0 |
| 221 | 2427 | 0 | 4 | 4 | 0.16 |
| 222 | 2483 | 6 | 1 | 7 | 0.28 |
| 223 | 2605 | 0 | 13 | 13 | 0.50 |
| 228 | 2053 | 23 | 4 | 27 | 1.32 |
| 230 | 2256 | 2 | 0 | 2 | 0.89 |
| 231 | 1571 | 1 | 0 | 1 | 0.06 |
| 232 | 1780 | 62 | 1 | 63 | 3.60 |
| 233 | 3079 | 0 | 3 | 3 | 0.10 |
| 234 | 2753 | 0 | 0 | 0 | 0 |
| Totals: | 109986 | 528 | 224 | 752 | 0.68 |

# Chapter 3

# Signal Compression

An ECG is generally sampled at a uniform rate. The ECG sampling frequency must be high enough to resolve the QRS complex and is set to 360 Hz in the MIT-BIH database. However, since the QRS complex operates at a comparatively short time scale, the majority of each beat tends to be oversampled. This situation allows the application of a variable rate sampling strategy to reduce the bit rate of the signal, and reduce the computational burden of future processing. To assess an appropriate sampling rate within a beat, the frequency content of each phase of the cardiac cycle must be understood. Each beat is therefore partitioned into three consecutive, non-overlapping time windows, or *blocks*. These blocks are labeled $\widehat{PQ}$, $\widehat{QRS}$, and $\widehat{ST}$, and approximate the standard clinically defined intervals. The function of the blocks is to divide a beat into the three distinct phases of the cardiac cycle, specifically, atrial activation, ventricular activation and ventricular repolarization.

## 3.1 Block definitions

For convenience, and in order to avoid biasing the results due to interaction with the QRS detector, the MIT–BIH database annotations were used to locate the R-peak of each beat. Let $\mathbf{y}$ be a column vector holding the ECG signal loaded from the CD-ROM, and let $R(k)$ be a list of R-peak addresses in $\mathbf{y}$, numbered $k = 0, 2, \ldots, K$. Each R-peak is centered in a $\widehat{QRS}$ block, denoted $b^{\widehat{QRS}}$, and the block's duration is set to 1/4 second, or 90 samples at 360 Hz (where the typical normal QRS complex is 50-110 ms [6]), that is for $0 < k < K$:

$$\mathbf{b}_k^{\widehat{QRS}} = \mathbf{y}_m^T,\ R(k) - 45 < m \leq R(k) + 45 \tag{3.1}$$

where $T$ represents the transpose operation, and $m$ is an index over partition $k$ in the measured signal. The PR interval is generally less than 40 percent of its corresponding RR interval [6]. The duration of the $\widehat{PQ}$ block is therefore set to 40 percent of its corresponding RR-interval less one-half of the duration of the $\widehat{QRS}$ block, that is:

$$\mathbf{b}_k^{\widehat{PQ}} = \mathbf{y}_m^T,\ \mathrm{Q}\left(\frac{4R(k)+6R(k-1)}{10}\right) < m \leq R(k) - 45 \tag{3.2}$$

where $\mathrm{Q}(\cdot)$ indicates that the quotient is retained and the remainder is discarded. Similarly, the duration of the $\widehat{ST}$ block is set to 60 percent of its corresponding RR-interval less one-half the $\widehat{QRS}$ block, that is:

$$\mathbf{b}_k^{\widehat{ST}} = \mathbf{y}_m^T,\ R(k) + 45 < m \leq \mathrm{Q}\left(\frac{4R(k)+6R(k+1)}{10}\right) \tag{3.3}$$

This partitioning strategy generally results in each phase of the cardiac cycle being associated with its correct block, that is, QRS complexes are contained in the $\widehat{QRS}$ blocks, P–waves in the $\widehat{PQ}$ block and T–waves in the $\widehat{ST}$ block. Let the beat vector associated with $R(k)$ be $B_k$. Each beat vector can be assembled from its constituent blocks according to the following relationship:

$$B_k = \left[\mathbf{b}_k^{\widehat{PQ}} \mid \mathbf{b}_k^{\widehat{QRS}} \mid \mathbf{b}_k^{\widehat{ST}}\right] \tag{3.4}$$

Furthermore, the signal vector $\mathbf{y}$, less the first and last beat, denoted $\mathbf{y}^*$ can be constructed from:

$$\mathbf{y}^* = \left[B_1 \mid B_2 \mid \cdots \mid B_{K-1}\right]^T \tag{3.5}$$

Due to variability in the RR interval, the dimension of each beat vector $B(k)$ is not the same. Special consideration must be made for the end effects if the first and last beats are of importance. In this study these end effects were ignored as a matter of convenience, however, the difficulty could be resolved by simply padding the blocks to an appropriate length with the first or last recorded sample as required.

The physiology of the ECG restricts the bandwidth of the signal because of the specialized conduction tissues which carry the heart's activation pulse. That is, the frequency content of the signal is a function of the conduction rate in the heart. Normally, each QRS complex is preceded by a P-wave, and followed by a T-wave. Thus, given the known relationships between RR interval and the P-R and S-T intervals, the normal time evolution of the cardiac cycle can be considered approximately constant with respect to the QRS complex. Because the QRS complex has been shown to be detectable in the majority of cases, it can be used as a robust datum for partitioning and resampling the signal, greatly simplifying any variable rate sampling scheme.

## 3.2 Specification of the Sampling Strategy

An appropriate sampling rate for each block can be determined from its spectral density function (SDF). The SDF was estimated by randomly sampling 200 normal beats from the MIT-BIH population, computing the SDF within each block of each beat using a zero-padded fast Fourier transform, and averaging across the ensemble. This approach is conservative because normal beats exist at the lower bound of QRS complex duration [16], and therefore operate at higher frequencies than pathologically conducted beats.

In [1] it was shown that the $\widehat{PQ}$ and $\widehat{ST}$ blocks have negligible power over about 10 Hz, while the $\widehat{QRS}$ block operates at up to about 30 Hz. The $\widehat{QRS}$ block can thus be conservatively resampled at a rate of 180 Hz, which yields a Nyquist frequency of 90 Hz. The average SDFs for the $\widehat{PQ}$ and $\widehat{ST}$ blocks indicate that these regions can be resampled at a rate of 72 Hz (every fifth point), yielding a Nyquist frequency of 36 Hz. These sampling rates were considered adequate to preserve details of clinical interest in the original signal. To minimize aliasing and noise effects, a computationally simple 9-point centered local average is computed in the original signal around each retained sample point in the $\widehat{PQ}$ and $\widehat{ST}$ blocks and a 3 point centered local average is similarly found around each retained sample point in the $\widehat{QRS}$ block. The centered average acts as a zero-phase, low-pass filter which attenuates noise (10 dB attenuation or better) above about $\frac{\pi}{6}$ in the $\widehat{PQ}$ and $\widehat{ST}$ blocks and above $\frac{\pi}{2}$ in the $\widehat{QRS}$ block, as shown by figure 3.1. Since both of these values approximate the Nyquist frequencies corresponding to the resampling rates there will be negligible aliasing taking place nor is the signal of interest compromised due to its band limited nature.

The positions of the retained points in $B_k$ must be chosen carefully in order to avoid difficulty in the region of a sampling rate shift. Let $\downarrow n$ represent a downsample by $n$ operator and let $B_k^r$ be a resampled beat. The sampling scheme is defined as:

1. For beat $B_k$ we begin in the $\widehat{QRS}$ block, $\mathbf{b}_k^{\widehat{QRS}}$. The total number of points in $\downarrow 2\ \mathbf{b}_k^{\widehat{QRS}}$ is 45, and the R-peak, $R(k)$, is fixed at $\downarrow 2\ \mathbf{b}_k^{\widehat{QRS}}(23)$.

2. The $\widehat{QRS}$ block is resampled and concurrently filtered moving outward until 22 samples have been extracted from each side of $R(k)$ in $\mathbf{y}$.

3. The sampling rate is shifted to $\downarrow 5$ .

4. The $\widehat{PQ}$ block is resampled in time reversed order, that is every fifth point starting at the boundary with the $\widehat{QRS}$ block and moving backward until 40 % of its corresponding RR interval is sampled.

5. The $\widehat{ST}$ block is resampled starting at the $\widehat{QRS}$ block boundary and moving forward until 60 % of its corresponding RR interval is accounted for.

By adopting this method the fact that the number of points in the $\widehat{PQ}$ and $\widehat{ST}$ blocks will not generally have an integer multiple of 5 samples in them is recognized. Because the R-peaks are at known locations, the RR-intervals are also known, and the signal can be reconstructed by letting the interval separating each beat vary. In this way, the temporal alignment of all features in the signal is preserved while minimizing overhead storage cost. The filtered and resampled signal can now be constructed using:

$$B_k^r = \left[\downarrow 5\ \mathbf{b}_k^{\widehat{PQ}} \mid \downarrow 2\ \mathbf{b}_k^{\widehat{QRS}} \mid \downarrow 5\ \mathbf{b}_k^{\widehat{ST}}\right] \tag{3.6}$$

and the signal vector becomes:

$$\mathbf{y}^r = \left[B_1^r \mid B_2^r \mid \cdots \mid B_{K-1}^r\right]^T \tag{3.7}$$

Due to the finite impulse response of the filters, and the fact that the filters are low pass, no discontinuity can be introduced at the block boundaries, that is, an average value used as the resampled data point is bounded by the maximum and minimum values used in its computation. Therefore, a discontinuity can only be attenuated by the filtering and resampling process.

## 3.3 Quality Controlled Compression

The aim of quality controlled compression is to ensure clinically adequate reconstruction on a beat-by-beat basis, preserving relevant features within each beat. To implement quality controlled compression, a criterion is proposed herein in which the number of retained transform coefficients is such that the relative mean absolute deviation, $\epsilon_k^j$, for each block $j = 1, 2$, or 3, corresponding to blocks $\downarrow 5\ \mathbf{b}_k^{\widehat{PQ}}$, $\downarrow 2\ \mathbf{b}_k^{\widehat{QRS}}$, and $\downarrow 5\ \mathbf{b}_k^{\widehat{ST}}$ respectively, and for each beat $0 < k < K$ is less than or equal to a prespecified tolerance. Specifically, the relative error over block $j$ for beat $k$ is defined as

$$\epsilon_k^j = \frac{\| \downarrow n^j \widehat{\mathbf{b}}_k^j - \downarrow n^j \mathbf{b}_k^j \|_1}{\| \downarrow n^j \mathbf{b}_k^j \|_1}, \qquad j = 1, 2, 3 \tag{3.8}$$

where $\| \cdot \|_1$ is the 1-norm, and $\downarrow n^j \widehat{\mathbf{b}}_k^j$ is the reconstruction approximation of block $\downarrow n^j \mathbf{b}_k^j$ from the retained transform coefficients. The number of retained coefficients

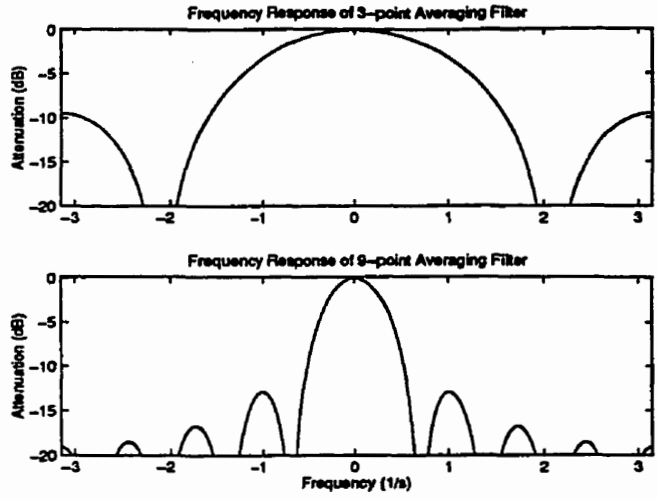


Figure 3.1: Frequency response of the 3-point and 9-point averaging filters. The original sampling frequency of the MIT–BIH database, and therefore the sampling rate which the filters operate at in the algorithm, is 360 Hz.

for the entire beat, $B_k^r$ is determined to be the minimum integer such that $\max_j \epsilon_k^j \leq b$ for some tolerance $b$. This criterion is iterative since $\epsilon_k^j$ depends on the number of retained coefficients. However, because of the orthogonality of the KLT basis, $\downarrow n^j \widehat{\mathbf{b}}_k^j$ is easily constructed incrementally and the minimum number of retained coefficients is found efficiently. For example, assuming that the eigenvalues have been sorted in descending order, if $\epsilon_k^1$, for a particular beat $k$, was found to be in excess of the relative error tolerance $b$, the reconstruction would be improved by additionally retaining the transform coefficient associated with the next largest unused eigenvalue, and its eigenvector. Due to orthogonality of the KLT basis, the residual vector can be modified directly, thus avoiding reconstructing $\downarrow n^j \widehat{\mathbf{b}}_k^j$ at each iteration.

## 3.4 Implementation Details

Each of the 48 records of the MIT-BIH database are analyzed in three segments of approximately 10 minutes of data. Let $S_j^n$ be such a segment where $j = 1, 2, 3$ and $n = 1, 2, \ldots, 48$ is an integer counter which represents each record of the database in sequential order. Baseline drift in each segment is removed by subtracting a one second wide centered moving average. Each segment is then separated into $N$ beats based on the supplied database annotations. Each beat is then partitioned into blocks

and resampled. The beats are loaded sequentially as a row vectors of dimension $M$ into a *beat matrix*, denoted **A**. The details of the blocking of a segment are defined as follows:

1. The maximum RR-interval $R_{max}$ is identified.
2. The number of data points in the $\widehat{PQ}$, $\widehat{QRS}$, and $\widehat{ST}$ blocks, $N_{pq}$, $N_{qrs}$, and $N_{st}$ respectively, are computed using the following expressions:

$$N_{pq} = \frac{0.4R_{max} - \frac{1}{8}}{T_{pq}}, \tag{3.9}$$

$$N_{qrs} = \frac{1}{4T_{qrs}}, \tag{3.10}$$

$$N_{st} = \frac{0.6R_{max} - \frac{1}{8}}{T_{st}}. \tag{3.11}$$

   Where $T_{pq}$, $T_{qrs}$ and $T_{st}$ are the sampling periods in the $\widehat{PQ}$, $\widehat{QRS}$ and $\widehat{ST}$ blocks respectively after each block has been resampled.
3. The row dimension, $M$ of the beat matrix is set to $M = N_{pq} + N_{qrs} + N_{st}$.
4. Since all beats have RR-intervals less than or equal to $R_{max}$, they must be aligned within the matrix. To do this, the R-peak for each beat vector is placed at the index $N_{pq} + (N_{qrs} - 1)/2$ (for $N_{qrs}$ odd) and the vector is padded to dimension $M$ using the first sample in the $\widehat{PQ}$ block at the beginning of the vector and the last sample in the $\widehat{ST}$ block at the end.
5. The beat vectors are then loaded into the $N \times M$ beat matrix **A**.

The KLT is a linear transformation of a matrix **A**. The basis vectors used for the transformation are the eigenvectors of the covariance matrix of **A**. The KLT is computed by finding the singular value decomposition (SVD) of **A**. It can be shown that any $N \times M$ matrix **A** where $N \geq M$, can be described as a product of an $N \times M$ column orthogonal matrix **U**, an $M \times M$ diagonal matrix $\Sigma$ and the transpose of an $M \times M$ orthogonal matrix **V** [17], that is:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T \tag{3.12}$$

Multiplying 3.12 through by $\mathbf{A}^T$, and recalling the properties of orthogonal and diagonal matrices, yields:

$$\mathbf{A}^T\mathbf{A}\mathbf{V} = \mathbf{V}\Sigma^2. \tag{3.13}$$

Immediately we recognize equation 3.13 as the eigenvalue problem. Thus the SVD of **A** yields the KLT basis vectors in the matrix $\mathbf{V}^T$ and the transform coefficients in the matrix **U**.

## 3.5 Results

The following results are based on the channel 1 lead of the MIT-BIH database. To obtain meaningful compression measures, the total storage consumed by all of the information required to reconstruct the approximation to the signal must be considered. We will assume that the precision with which the required coefficients are stored is the same as that for the original signal. If the dimension of the beat matrix is $N > M$ (true if no beat lasts more than about 5 seconds throughout the 10-minute segment), then the KLT produces an $M \times M$ matrix of eigenvectors, an $N \times M$ matrix of transform coefficients, and the $M$ eigenvalues $\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2$.

Some rudimentary analysis is useful for estimating the order of the compression rate. For the variance criterion, $m$ eigenvalues (arranged in descending order of magnitude) and eigenvectors, along with $m$ transform coefficients for each of $N$ beats are retained with $m$ chosen so that $\sum_{i=1}^{m} \sigma_i^2 / \sum_{i=1}^{M} \sigma_i^2 \geq \delta_m$. Including the $N$ RR-interval values needed to reconstruct each beat, the variance control criterion requires the storage of $m(M+1)+N(m+1)$ values in total. A 10 minute recording originally containing $(10)(60)(360) = 216,000$ data values typically has $N = 750$ (corresponding to a heart rate of 75 beats per minute) and $M = 100$ after downsampling. For $m = 3$ retained transform coefficients, the storage requirement totals $3(100+1)+750(3+1) = 3303$ values yielding a compression ratio $C = 65{:}1$. If $m = 10$, $C$ drops to about 23:1.

In contrast to the variance criterion, quality controlled compression, requires the storage of:

1. A variable number of transform coefficients $m_i$, $i \subset \{1, \ldots, N\}$.
2. $m_{\max} = \max m_i$, eigenvalues and eigenvectors.
3. The RR-interval for each of the $N$ beats.

The total storage requirement for segment of $N$ beats is then $m_{max}(M+1)+N(\bar{m}+1)$ where $\bar{m}$ is the average number of retained transform coefficients over the $N$ beats. The compression ratio for a 10 minute segment (as above) with $N = 750$, $M = 100$, $\bar{m} = 3$ and $m_{max} = 20$ is then $C = 43{:}1$. While most beats are stored using only 2 or 3 transform coefficients this approach allows for the occasional (pathological)

beat to be accurately represented using a larger number of transform coefficients, maintaining, overall, an excellent compression ratio.

The MIT–BIH database is divided into two series, the first will be denoted the 100 series and the second the 200 series to reflect their labeling in the database. The first series contains 23 records which were randomly selected from a population of recordings, while the second series, which contains 25 records, was hand picked to represent unusual recordings that would not be well represented in a random sample due to their infrequency [14]. Because of this distinction, the average compression ratios will be reported separately. The compression ratio on the 100 series should reflect an estimate of the algorithms global performance. The compression performance on the 200 series is not as important as the algorithm's ability to preserve rare events of clinical interest. Ideally, an algorithm will be capable of high compression and fidelity on the 100 series and as a minimum high fidelity on the 200 series, although maintaining a good compression rate is, of course, desirable.

The average compression ratio over the 100 series of the MIT-BIH database for the variance criterion with $\delta_m = 0.995$ is $C_V \approx 33$ while the same for the quality control criterion with tolerance $b = 0.25$ is $C_{QC} \approx 27$. For the 200 series the compression rates are $C_V \approx 21$ and $C_{QC} \approx 15$. These numbers for the variance and quality control criteria are base compression ratios. A complete compression system generally includes a stage of lossless compression which takes advantage of the fact that the coefficients of the transform are not uniformly distributed random variables [19]. For example, since normal beats closely resemble one another, the transform coefficients associated with each eigenvector should be close to the same size over a segment. Thus if the $N$ transform coefficients associated with the first eigenvector were grouped, the magnitude of the first transform coefficient could be stored followed by the differences between transform coefficients. If the coefficients are correlated, the differences can be stored at a lower bit rate than the transform coefficients without any loss of precision.

In table 3.1, the average compression achieved on each of the 100 series records of the MIT-BIH database is reported. Both the compression ratio, and the average bit rate are shown. Typically, the quality control criterion provides a lower average compression ratio. This implies, however, that there are a significant number of beats that meet the variance criterion but go on to fail the quality control criterion (so that $m$ must be increased). Due to the nature of the variance measure we expect these errors to appear in regions of of low amplitude, as these regions make smaller contributions to the signal's variance.

Quality control does not exclude the possibility of a higher compression ratio than

the variance based method. For example record 112 of the MIT–BIH database was compressed at $C = 27.7$:1 using the variance method and $C = 51.3$:1 using the quality control method. In the occurrence of a localized region in the signal which is difficult to represent and is also a significant disturbance of the residual vector, the higher number of transform coefficients required to represent this region are applied globally in the variance based method. Thus many beats will be over–represented, increasing the storage cost for the record. By varying the number of coefficients used to represent each beat, the quality controlled method can maintain a high fidelity representation which is uniform across the record. In the second 10 minute segment of record 213, see figures 3.4 and 3.5, the maximum number of retained transform coefficients was $m_{max} = 30$, however, the average number of coefficients was only $\overline{m} = 3.1$ in order to achieve $b \leq 0.25$. For the variance based method $m = 10$ transform coefficients were retained to meet the $\delta = 0.995$ minimum. Clearly, 10 transform coefficients is insufficient locally while simultaneously being globally extravagant.

Our first example is taken from record 103 of the MIT–BIH database. The initial few seconds of the record are shown in figure 3.2, with the quality controlled compression and variance based compression plotted simultaneously in the top and bottom graphs respectively. These charts are intended to represent the average behavior of the compression methods. The quality controlled reconstruction was compressed at $C = 30.7$:1 and the variance based method was compressed at $C = 48.5$:1. Inspection of figure 3.2 shows that the quality controlled method's reconstruction is superior to that of the variance based method and except for a small departure from the T-wave at $t \approx 3.25s$ the two representations are indistinguishable. The variance based method suffers from some localized departures from the signal, although they are not likely of clinical significance. In figure 3.3 several individual beats from various locations in the record have been extracted from the reconstruction because the number of retained transform coefficients, $m_i$, is greater than 15 in order that the quality control criterion is satisfied. We see that the use of a globally averaged error measure in the variance based reconstruction allows the existence of large localized errors, whereas the quality controlled compression is of uniformly high fidelity.

The second example is taken from record 213 of the MIT–BIH database. In this case the compression rate using the variance criterion was $C = 17.4$:1 and the compression rate using the quality control method was $C = 35.4$:1 or approximately double that of the variance base method. Inspection of figure 3.4 shows the continuous reconstructions based on both the variance, and the quality controlled methods. We see that the quality controlled method makes errors of similar sizes to those in

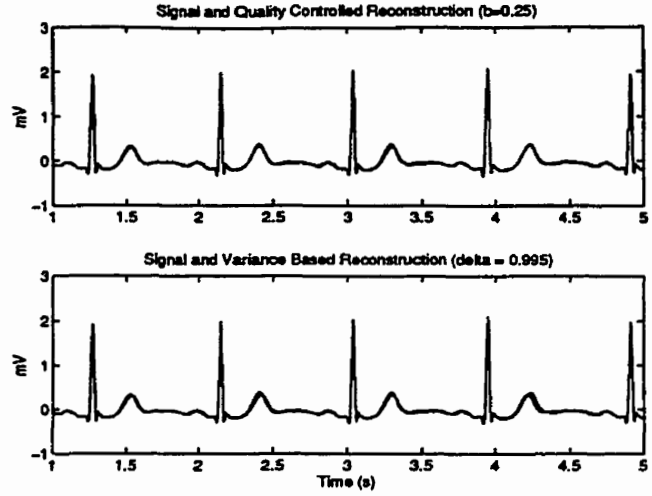


Figure 3.2: Continuous reconstruction of the first few seconds of record 103 of the MIT–BIH database. The top graph shows the original signal and the quality controlled reconstruction ($C = 30.7$:1), and the bottom graph shows the original signal and the variance based reconstruction ($C = 48.5$:1).

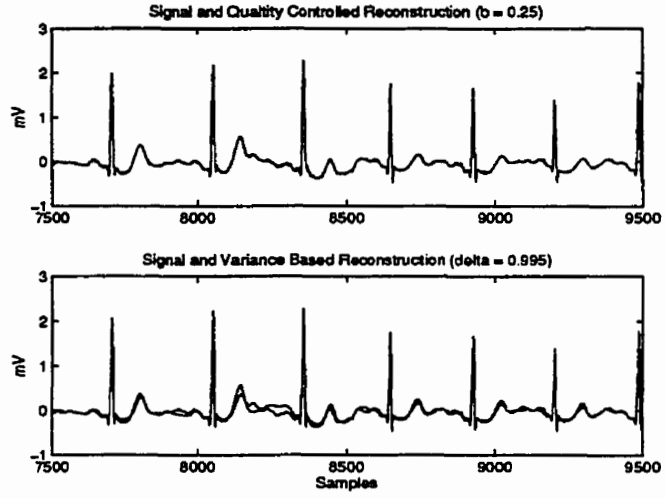


Figure 3.3: Several individual beats from record 103 of the MIT–BIH database are displayed where for the quality controlled method ($C = 30.7$:1), shown in the top graph, $m > 15$. In the bottom graph, the variance based method ($C = 48.5$:1), where $m_j = \{3, 6, 6\}$, for 10 minute segments $j = 1, 2, 3$, is a constant, allows localized departures from the signal due to global averaging of the reconstruction error.

the variance based method, however the errors are of a more uniform magnitude over the reconstruction. In the variance based reconstruction the errors are much smaller in some beats than in others resulting in increased storage costs. In figure 3.5 we see that in spite of the much lower compression rate, the variance based reconstruction is still subject to localized departures from the signal. We see, while maintaining two times the compression rate, the quality controlled reconstruction remains remarkably accurate. In summary, an algorithm employing KLT, along with its attractive properties relating to error measures and optimal bases, combined with resampling and a quality control criterion has been found to lead to efficient and accurate compression of ECG recordings. By careful consideration of the bandwidth containing the signal of interest, resampling can be performed without loss of pertinent information. Resampling, in turn, leads to efficient computation of the KLT, significant eigenvector storage savings, and compression ratios which typically exceed those achieved via wavelet packet-based algorithms [2]. The KLT also naturally lends itself to noise model extensions relevant to signal identification. Finally, the quality control criterion allows clinically acceptable reconstructions while maintaining high compression ratios.

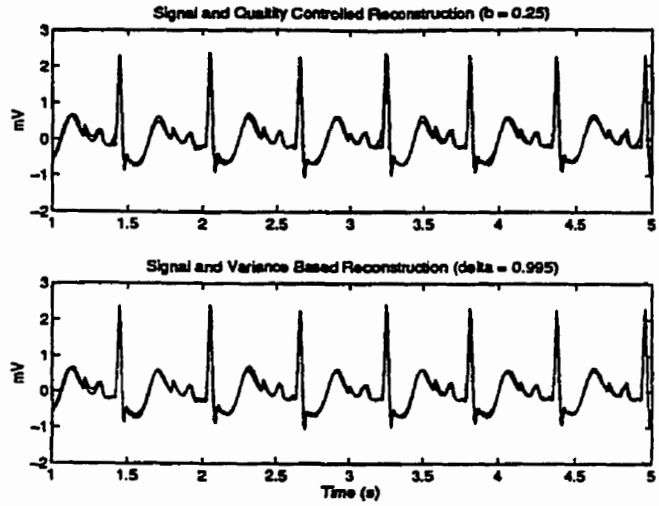


Figure 3.4: Continuous reconstruction of the first few seconds of record 213 of the MIT–BIH database. The top graph shows the original signal and the quality controlled reconstruction ($C = 35.4$:1), and the bottom graph shows the original signal and the variance based reconstruction ($C = 17.4$:1).

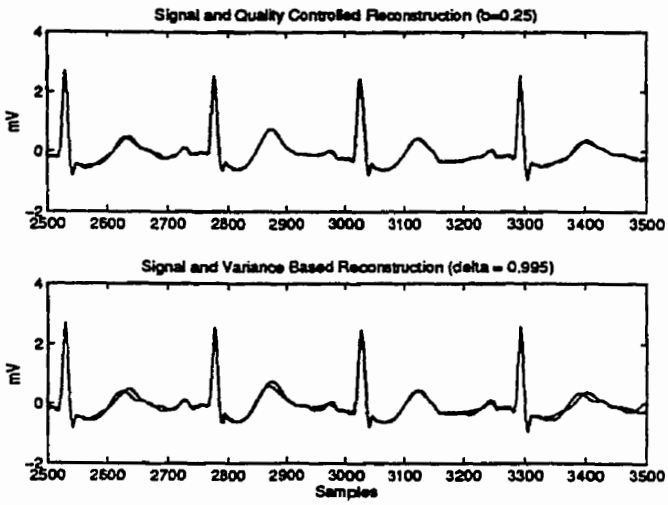


Figure 3.5: Several individual beats from record 213 of the MIT–BIH database are displayed where for the quality controlled method ($C = 35.4$:1), shown in the top graph, $m > 15$. In the bottom graph, the variance based method ($C = 17.4$:1), where $m_j = \{9, 10, 10\}$, for 10 minute segments $j = 1, 2, 3$, is a constant. Thus, with a lower compression rate, the variance based reconstruction still suffers from localized errors.

Table 3.1: Results of the ECG compression study using the 100 series of the MIT–BIH database.

| Record | Variance method ($\delta_m = 0.995$) | | Quality controlled method ($\epsilon_j \leq 0.25$) | |
|---|---|---|---|---|
| | Ratio | bps | Ratio | bps |
| 100 | 33.0 | 120 | 26.6 | 149 |
| 101 | 36.3 | 109 | 14.3 | 278 |
| 102 | 18.6 | 213 | 17.3 | 228 |
| 103 | 48.5 | 81.7 | 30.7 | 129 |
| 104 | 11.5 | 345 | 9.3 | 428 |
| 105 | 17.9 | 221 | 16.1 | 246 |
| 106 | 20.3 | 195 | 17.1 | 232 |
| 107 | 23.7 | 167 | 18.5 | 214 |
| 108 | 13.6 | 290 | 11.2 | 353 |
| 109 | 45.7 | 86.6 | 27.8 | 143 |
| 111 | 21.3 | 186 | 21.5 | 184 |
| 112 | 27.7 | 143 | 51.3 | 77.2 |
| 113 | 41.7 | 95.0 | 52.0 | 76.2 |
| 114 | 15.4 | 256 | 10.5 | 378 |
| 115 | 55.9 | 70.9 | 40.1 | 98.8 |
| 116 | 31.2 | 127 | 30.3 | 131 |
| 117 | 41.9 | 94.6 | 39.9 | 99.3 |
| 118 | 20.7 | 191 | 16.5 | 239 |
| 119 | 31.7 | 125 | 25.9 | 153 |
| 121 | 34.4 | 115 | 33.7 | 117 |
| 122 | 72.2 | 54.9 | 54.9 | 72.2 |
| 123 | 54.7 | 72.5 | 31.7 | 125 |
| 124 | 48.3 | 81.9 | 25.3 | 157 |
| Average | 33.3 | 150 | 27.1 | 187 |

Table 3.2: Results of the ECG compression study using the 200 series of the MIT–BIH database.

| Record | Variance method ($\delta_m = 0.995$) | | Quality controlled method ($\epsilon_i \leq 0.25$) | |
|---|---|---|---|---|
| | Ratio | bps | Ratio | bps |
| 200 | 13.3 | 299 | 6.8 | 583 |
| 201 | 16.6 | 239 | 7.3 | 541 |
| 202 | 30.1 | 131 | 18.3 | 217 |
| 203 | 11.2 | 352 | 6.6 | 603 |
| 205 | 26.4 | 150 | 23.0 | 172 |
| 207 | 26.1 | 152 | 6.9 | 575 |
| 208 | 14.7 | 270 | 11.4 | 347 |
| 209 | 12.3 | 321 | 15.3 | 260 |
| 210 | 16.7 | 237 | 8.3 | 479 |
| 212 | 18.4 | 216 | 20.4 | 194 |
| 213 | 17.4 | 228 | 35.4 | 112 |
| 214 | 23.0 | 173 | 11.7 | 338 |
| 215 | 13.9 | 285 | 19.9 | 199 |
| 217 | 21.2 | 186 | 22.5 | 176 |
| 219 | 23.1 | 171 | 8.3 | 476 |
| 220 | 39.1 | 101 | 25.1 | 157 |
| 221 | 22.5 | 176 | 10.9 | 364 |
| 222 | 10.4 | 381 | 8.7 | 456 |
| 223 | 23.7 | 167 | 19.1 | 207 |
| 228 | 15.1 | 263 | 12.9 | 307 |
| 230 | 26.2 | 151 | 19.4 | 204 |
| 231 | 33.2 | 119 | 18.6 | 213 |
| 232 | 11.8 | 226 | 5.3 | 753 |
| 233 | 20.0 | 198 | 18.1 | 218 |
| 234 | 44.1 | 89.7 | 25.1 | 158 |
| Average | 21.2 | 216 | 15.4 | 332 |

# Chapter 4

# Conclusion

The QRS detection, and ECG compression problems share the remarkable characteristics of having long histories and large literatures which serve as confounding influences in the solution of these problems. For example, despite the identification of the inadequacy of globally averaged error measures [8], these measures have continued to be applied to ECG compression results. The significance of this fact is that the use of such measures allows the compression algorithm to cause errors which may change a diagnosis, as was clearly pointed out in [1].

In the case of the QRS detection problem, extraordinarily high success rates have been reported using algorithms which are optimized over their entire testing data set, and require frequent intervention of the user. The difficulty with this approach is that the requirement of this intervention implies that the algorithms cannot deal with changes in signal character, where in general, any change in signal may be of clinical interest. In addition, the recordings used as testing data are 30 minutes in length. This length excludes any of the daily variation in the ECG which is known to be extreme [4]. In addition, these algorithms include tactics based on constraints which are database dependent, solely to improve their performance over the testing set.

In the case of QRS detection research an algorithm has been presented with the following properties:

1. The algorithm is simple and executes rapidly enough for use in the analysis of Holter recordings. In addition, the algorithm is fully disclosed.

2. The algorithm uses only fixed global parameters.

3. There are only six parameters in the algorithm. All of these parameters were

deduced from the performance of the algorithm over a small set of example recordings.

4. The algorithm is free of database dependent tactics.

5. The algorithm maintains an error rate of the same order as those presented in the previous literature.

Although the QRS detection algorithm presented here is a radical departure from the previous literature, there still exists a significant amount of research to be done before this problem can be considered solved. Several questions of interest are, for example:

1. What is the sensitivity of the result to the parameters?

2. Can the error rate, and in particular the lower bound on the error rate, be improved by including information from other leads?

3. What kind of strategies would be required in order that this information can be included, realizing, that the signal quality changes between leads, and the representation of a QRS complex is different in each of the leads?

4. What are the effects of complex arrythmias such as ventricular flutter ( and others ) on the performance of the algorithm? Because the database is not exhaustive, these kinds of rhythms tend to be underrepresented.

Considering the results of the compression work, we see that the quality control measure is the first meaningful local error measure to be applied to the ECG compression problem. The success of the measure is directly linked to the use of resampling to make SVD computationally accessible. This compression technique has the following properties:

1. The quality controlled compression method provides a higher average compression rate than previously reported strategies.

2. The method uses a physiologically based error measure.

3. The method uses active error control.

4. The method provides a reconstructed approximation of higher accuracy than previously reported methods.

5. The method uses globally fixed parameters, and does not require intervention of a user.

6. The method is not predisposed to making large localized errors.

Areas of further interest include:

1. The interaction of the compression method with the automatic QRS detection strategy.

2. A clinical study of the effect of the method on the diagnostic quality of the signal.

3. The relationship between the transform coefficients and pathology. That is, can the fact that the quality control measure is capable of detecting localized departures from typical signal behavior be used to advantage in the the automated detection of pathological ECG signals.

4. The inclusion of lossless coding strategies.

This study represents the first step towards the solution of the ECG pattern recognition problem. That is, the identification of individual hearts beats must, by definition, precede their classification, and the quality control criterion, far beyond being a simple compression alternative, represents a significant step towards applying the SVD transform coefficients to the pattern recognition problem. Consider that one approach to the automated diagnosis of ECG would be to express each beat as a small number of coefficients and identify patterns in these coefficients associated with the various types of beat which are known to exist. This action represents *compression* of the signal. If the compression does not preserve the local information in the signal, the pattern recognition problem *cannot* be solved due to the potential for projection of information associated with pathology into the null space of the transform. Thus, quality control transcends compression, and in turn is impossible without reliable QRS detection. The literature has failed to provide a suitable starting point for the pattern recognition problem. In the minimum, this study forces the reconsideration of these results. In another light, this study is a foundation on which to build meaningful solutions.

# Bibliography


[1] Blanchett, T., Kember, G.C., Fenton, G. A., "KLT–based quality controlled compression of single–lead ecg," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 942–945, 1998.

[2] Bradie, Brian, "Wavelet packet based compression of single lead ECG", *IEEE Trans. Biomed. Eng.*, vol. 43, pp.493–501, 1996.

[3] Bromba, Manfred U.A., Ziegler, Horst, "Application hints for Savitzky–Golay digital smoothing filters," *Anal. Chem.*, vol. 53, pp. 1583–1586, 1981.

[4] Chou, Te-Chaun, Knilans, Timothy K., *Electrocardiography in Clinical Practice—4th ed.*, Philadelphia, W. B. Saunders, 1979.

[5] El-Sherief, H., Pham, S., El-Sherief, N., Caref, E., "Clinical evaluation of ECG data compression techniques for ambulatory recording," *Proc. 16th IEEE Conf. Engineering in Medicine and Biology.* Baltimore, MD, USA, 1994. pp. 1306–1307.

[6] Goldman, M. J., *Principles of Clinical Electrocardiography 8th ed*, Lange Medical Publications, Los Altos, CA, 1973.

[7] Hamilton, Patrick S., Tompkins Willis J., "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," *IEEE Trans. Biomed. Eng.*, vol. BME-33, 1986, pp. 1157–1165.

[8] Jalaleddine, Sateh M.S., Hutchens, Chriswell G., Strattan, Robert D.,Cobberly, William A., "ECG data compression techniques — a unified approach," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 329–343, 1990.

[9] Li, Cuiwei, Zheng, Chongxun, Tai, Changfeng, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 21–28, 1995.

[10] Mallat, Stephane, Zhong, Sifen, "Characterization of signals from multiscale edges," *IEEE Trans. Patt. Anal. Machine Intel.*, vol. 14, pp. 710–732. 1992.

[11] Mendel, Jerry M., "Fuzzy logic systems for engineering: a tutorial," *Proc. IEEE*, vol. 83, pp. 345–377, 1995.

[12] *The MIT-BIH Arrhythmia Database Directory -2nd ed.*, BMEC TR010 (Revised), Massachusetts Institute of Technology Biomedical Engineering Center, 1988.

[13] Moody, George B., Mark, Roger G., "QRS morphology representation and noise estimation using the Karhunen–Loeve transform," in *Proc. Computers in Cardiology 1989.* Los Alamitos, CA, USA: IEEE Comput. Soc. Press, 1990, pp. 269–272.

[14] Moody, George B., Mark, Roger G., "The MIT–BIH arrhythmia database on CD–ROM and software for use with it," in *Proc. Computers in Cardiology 1990.* Los Alamitos, CA, USA: IEEE Comput. Soc. Press, 1991, pp. 185–188.

[15] Pan, Jiapu, Tompkins, Willis J., "A real–time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol BME–32, 1985, pp. 231–236.

[16] Phibbs, Brendan, *The Cardiac Arrhythmias*, Mosby, Saint Louis, 1973.

[17] Press, William H., Vetterling, William T., Teukolsky, Saul A., Flannery, Brian P., *Numerical Recipies in C*, Cambridge, Cambridge University Press, 1992.

[18] Proakis, John G., Manolakis, Dimitris G., *Digital Signal Processing*, Upper Saddle River, New Jersey, Prentice Hall, 1996.

[19] Strang, Gilbert, Nguyen, Truong *Wavelets and Filter Banks*, Wellesley MA, Wellesley–Cambridge Press, 1996.

[20] Thakor, N. V., Sun, Y., Rix, H., "MULTIWAVE: A multiresolution wavelet–based ECG data compression algorithm," in *Proc. Computers in Cardiology 1993.* Los Alamitos, CA, USA: IEEE Comput. Soc. Press, 1993, pp. 393–396.

[21] Weisberg, Sanford, *Applied Linear Regression*, New York, Wiley, 1985.