An Efficient Method for Handling Ectopic Beats Using the Heart Timing Signal

Kristian Solem, Pablo Laguna, *Member, IEEE*, and Leif Sörnmo*, *Senior Member, IEEE*

***Abstract*—The problem of analyzing heart rate variability in the presence of ectopic beats is revisited. Based on the integral pulse frequency modulation model and the closely related heart timing signal, a new technique is introduced which corrects for the occasional presence of ectopic beats. The correction technique, which involves the occurrence times of a certain number of beats preceding the ectopic beat, is computationally very efficient. From actual heart rate data, the results show that the new technique is associated with a much lower computational complexity (flops reduced by a factor of about 3000) than the original heart timing technique, while producing similar performance. It is also shown that the power spectrum and related clinical indices obtained by the new technique are more accurately estimated than by other methods.**

***Index Terms*—ECG, ectopic beat correction, heart timing signal, HRV, IPFM model.**

I. Introduction

THE presence of ectopic beats perturbs the impulse pattern initiated by the sinoatrial node, and implies that RR intervals adjacent to an ectopic beat cannot be used for heart rate variability (HRV) analysis. In such cases, autonomic modulation of the sinoatrial node is temporarily lost, and an ectopic focus instead initiates the next beat prematurely. The location of the ectopic focus gives rise to different types of RR interval perturbation; a beat of ventricular origin inhibits the next sinus beat so that a compensatory pause is introduced after the ectopic beat, whereas a beat of supraventricular origin discharges the sinoatrial node ahead of schedule and causes the following sinus beat to also occur ahead of schedule. Other perturbations of physiological origin are those related to an interpolated ectopic beat, manifested by two short RR intervals adjacent to the ectopic beat, or an escape beat, manifested by a prolonged RR interval.

Since ectopic beats may occur in both normal subjects and patients with heart disease, their presence represents an important error source which must be dealt with before spectral analysis can be performed. If not dealt with, the analysis of an RR interval series containing ectopic beats results in a power spectrum with spurious frequency components.

Manuscript received October 18, 2004; revised July 4, 2005. This work was supported by a grant from Gambro AB. The work of P. Laguna was supported by the the Spanish CICYT agency and FEDER under Grant TEC2004-05263-C02-02. *Asterisk indicates corresponding author.*

K. Solem is with the Signal Processing Group, Department of Electroscience, Lund University, SE-221 00 Lund, Sweden.

P. Laguna is with the Aragon Inst. for Engineering Research, Zaragoza University, E-50009 Zaragoza, Spain.

*L. Sörnmo is with the Signal Processing Group, Department of Electroscience, Lund University, P.O. Box 118, SE-221 00 Lund, Sweden (e-mail: leif.sornmo@es.lth.se).

Digital Object Identifier 10.1109/TBME.2005.859780

A number of techniques have been developed which deal with the presence of ectopic beats, all techniques conforming to the restriction that only ECG segments with occasional ectopic beats should be processed [1], [2]. Segments containing frequent ectopic beats or, worse, runs of ectopic beats, perturb the underlying sinus rhythm and must, therefore, be excluded from further analysis [3]. A simplistic approach to the correction of an occasional ectopic beat is to delete the aberrant RR intervals from the series of RR intervals. However, interval deletion does not try to fill in the interval variation that should have been present, had no ectopic beat occurred, and, as a result, the "corrected" interval series remains unsuitable for HRV analysis.

A popular correction technique performs interpolation over the gap caused by the ectopic beat in order to obtain values that align with the adjacent NN intervals, see, e.g., [4]–[6]; low-order interpolation is usually employed. When the aim is to specifically analyze HRV with a nonparametric, spectral approach, the correlation function estimator required for computing the periodogram can be modified to account for ectopic beats [7].

The heart timing (HT) signal was recently suggested for the characterization of HRV [8]. This signal is based on the well-known integral pulse frequency modulation (IPFM) model for the generation of normal sinus beats [9], characterizing HRV in terms of a modulation function $m(t)$. From the unevenly sampled HT signal, the function $m(t)$ can be estimated by analyzing the deviations of the event times t_k from the expected occurrence times, defined by the mean RR interval length. The definition of the HT signal has later been extended to also account for the presence of occasional ectopic beats [10], [11]; see also [12] for a similar IPFM-based approach. In terms of spectral distortion, the results showed that the HT-based correction produced one order of magnitude lower error than did interpolation-based correction techniques when applied to other HRV representations than the HT signal. While producing excellent results, the HT-based correction is associated with heavy computations which, for example, in the analysis of Holter recordings, may become prohibitive.

The present paper introduces a correction method which drastically reduces the computational demands of the method presented in [11], while introducing no significant reduction in performance. Section II starts with a brief overview of the IPFM model and the HT signal, followed by a description of the present correction method. The performance is evaluated on a set of ECG recordings containing ectopic beats, obtained from 132 subjects [13], and compared to the performance of the original HT-based method (Sections III and IV). Finally, certain properties of the new method are discussed in Section V.

II. Methods

A. IPFM Model

The IPFM model generates a series of occurrence times for normal sinus beats ("events") with known rate variability, and reflects basic electrophysiological properties of the sinoatrial node [9]. The input signal to the IPFM model consists of the sum of a dc level, related to the average heart rate (HR), and a modulating signal, $m(t)$, related to the variability due to parasympathetic and sympathetic activity. The input signal to the IPFM model is integrated until a threshold, T_0 (representing the mean interval length between successive events), is reached. Then, an event is created at time t_k as the output of the model, and the integrator is reset to zero. As a result, the output signal of the IPFM model becomes an event series which represents the beat occurrence times. In mathematical terms, the following equation defines the series of event times

$$\int_0^{t_k} (1+m(\tau))\, d\tau = kT_0 \quad k=0,\ldots,K \tag{1}$$

where k is an integer that indexes the k^{th} beat following the initial event, and the initial event assumed to occur at $t_0 = 0$. The function in (1) can be generalized to a continuous-time function by introducing the following definition:

$$\int_0^{t} (1+m(\tau))\, d\tau = \kappa(t)T_0. \tag{2}$$

The integral can now be calculated up to any time t, and is proportional to an index function $\kappa(t)$ whose value at t_k is identical to the integer beat index k, i.e., $\kappa(t_k) = k$.

B. Heart Timing Representation

The HT signal $d_{\text{HT}}(t)$ is at $t = t_k$ defined as the difference between t_k and the expected occurrence time at the mean HR, kT_0 [8]. The HT signal is closely related to the IPFM model and its modulating signal $m(t)$. Using the HT signal, the modulating signal $m(t)$ can be estimated in order to produce the HRV power spectrum. In order to see how $d_{\text{HT}}(t)$ and the modulating signal $m(t)$ are related, the model equation in (1), for a particular time t_k, is rewritten according to

$$\int_0^{t_k} m(\tau)d\tau = kT_0 - t_k \equiv d_{\text{HT}}(t_k). \tag{3}$$

The mean RR interval length T_0 must be estimated from the available data before $d_{\text{HT}}(t_k)$ can be computed. This can be done by simply dividing the time of the last event with the number of events, i.e.,

$$\hat{T}_0 = \frac{t_K}{K}. \tag{4}$$

Using the generalized IPFM model in (2), the HT signal can be expressed in continuous-time as

$$d_{\text{HT}}(t) = \int_0^{t} m(\tau)d\tau = \int_{-\infty}^{t} m(\tau)d\tau \tag{5}$$

where the integration interval is extended to $-\infty$ since $m(t)$ is assumed to be a causal function. If the Fourier transforms of $m(t)$ and $d_{\text{HT}}(t)$ are denoted with $D_m(\Omega)$ and $D_{\text{HT}}(\Omega)$ respectively, we have from (5) that

$$\begin{aligned} D_{\text{HT}}(\Omega) &= \int_{-\infty}^{\infty} d_{\text{HT}}(t)e^{-j\Omega t}dt \\ &= \frac{D_m(\Omega)}{j\Omega} + \pi D_m(0)\delta(\Omega) \\ &= \frac{D_m(\Omega)}{j\Omega} \end{aligned} \tag{6}$$

where $\Omega = 2\pi F$ and $D_m(0) = 0$, since $m(t)$ has a dc component equal to zero. Once the Fourier transform of the HT signal, $D_{\text{HT}}(\Omega)$, is known, the desired spectrum $D_m(\Omega)$ can be computed according to

$$D_m(\Omega) = j\Omega D_{\text{HT}}(\Omega). \tag{7}$$

The spectrum $D_{\text{HT}}(\Omega)$ is obtained either by a technique for unevenly sampled signals, or interpolation and resampling followed by use of the discrete Fourier transform.

C. Dealing With Ectopic Beats

In this section, we briefly summarize the recently presented technique [11] which compensates for the presence of ectopic beats using $d_{\text{HT}}(t)$, and then continue with the new approach. In the description below, we assume that sinus beats occur at the times $t_0, t_1, \ldots, t_K$, and that one ectopic beat occurs at time t_e. The time t_e is not included in the series $t_0, t_1, \ldots, t_K$, and the sinus beat immediately preceding the ectopic beat occurs at t_{k_e} and the sinus beat immediately following at t_{k_e+1}.

1) Heart Timing Representation: In order to compensate for the presence of an ectopic beat, the above definition of $d_{\text{HT}}(t_k)$ is modified by the introduction of a parameter s according to [11]

$$d_{\text{HT}}(t_k) = \begin{cases} kT_0 - t_k & k = 0,\ldots,k_e, \\ (k+s)T_0 - t_k & k = k_e+1,\ldots,K \end{cases}. \tag{8}$$

The parameter s can be viewed as a jump in the resetting of the integral in the IPFM model. If the value of s is close to zero it indicates that an artifact probably has caused the event at t_e. A value close to one usually indicates that the event is a premature ectopic beat followed by a compensatory pause. From the modified HT signal in (8), the IPFM generalization yields

$$d_{\text{HT}}(t) = \kappa(t)T_0 - t \tag{9}$$

where

$$\kappa(t) = \frac{1}{T_0}\int_0^{t} (1+m(\tau))\, d\tau \tag{10}$$

and, for the case with an ectopic beat, the indexing function $\kappa(t)$ is at t_k given by

$$\kappa(t_k) = \begin{cases} k & k = 0,\ldots,k_e, \\ k+s & k = k_e+1,\ldots,K \end{cases}. \tag{11}$$

If the modified formulation of $d_{\text{HT}}(t_k)$ in (8) is to be useful, we need to estimate the parameter s and update our estimator of T_0 so that it accounts for the presence of an ectopic beat. Several steps are required to derive an estimator of s, but once known it is straightforward to determine an estimator of T_0.

The indexing function $\kappa(t)$ can be estimated from the occurrence times which precede $(t_k, \kappa(t_k) = k)$ and follow an

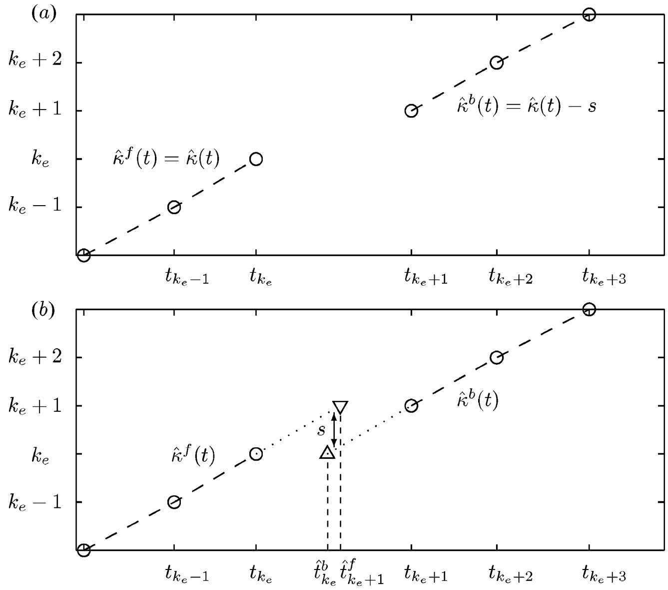


Fig. 1. (a) The indexing functions $\hat{\kappa}(t)$, based on occurrence times up to t_{k_e}, and $\hat{\kappa}(t)-s$, based on occurrence times starting from t_{k_e+1}. (b) The indexing functions shown in (a) extrapolated forward and backward to produce the occurrence times $\hat{t}^f_{k_e+1}$ and $\hat{t}^b_{k_e}$.

ectopic beat $(t_k, \kappa(t_k) = k + s)$. Thus, two different estimators of $\kappa(t)$ can be obtained: one that will be "forward-extending" based on the occurrences at $(t_0, 0), \ldots, (t_{k_e}, k_e)$ and denoted with $\hat{\kappa}^f(t)$, and another that will be "backward-extending" based on $(t_{k_e+1}, k_e+1), \ldots, (t_K, K)$ and denoted with $\hat{\kappa}^b(t)$ being an offset version of the original $\kappa(t)$ (i.e., $\kappa^b(t) = \kappa(t) - s$), see Fig. 1(a). Note that $\kappa(t_k) = k$, $k = 0, \ldots, k_e - 1, k_e$, and $\kappa(t_k) = k + s$, $k = k_e + 1, \ldots, K$.

Since the resulting two indexing functions would differ by an offset equal to the desired parameter s, these two functions can be extrapolated forward and backward in time, respectively, to such an extent that they overlap and thereby allow for estimation of s, see Fig. 1(b). This is first done by forwardly extending the occurrence times $t_0, \ldots, t_{k_e}$ with a new time $\hat{t}^f_{k_e+1}$ under the assumption that the sinus rhythm continues. Similarly, the occurrence times $t_{k_e+1}, \ldots, t_K$ are backwardly extended with a new time $\hat{t}^b_{k_e}$ under the assumption that the sinus rhythm preceded t_{k_e+1}. This procedure continues until the desired overlap exists, i.e., $\hat{t}^f_{k_e+1} > \hat{t}^b_{k_e}$. The computation of these two occurrence times is defined by

$$\hat{t}^f_{k_e+1} = t_{k_e} + d^i_{IF}\left(\hat{t}^f_{k_e+1}\right) \quad (12)$$

$$\hat{t}^b_{k_e} = t_{k_e+1} - d^i_{IF}\left(t_{k_e+1}\right) \quad (13)$$

where $d^i_{IF}(t)$ denotes the interpolated interval function. Hence, in order to determine $\hat{t}^f_{k_e+1}$ and $\hat{t}^b_{k_e}$ one has to first interpolate the interval function, given by

$$d_{IF}(t_k) = t_k - t_{k-1} \quad k \leq k_e \quad \text{or} \quad k \geq k_e + 2 \quad (14)$$

where the intervals adjacent to the ectopic beat have been excluded from the computation of $d_{IF}(t_k)$. The value of $\hat{t}^f_{k_e+1}$ is obtained by solving (12) recursively; it is straightforward to obtain the value of $\hat{t}^b_{k_e}$ from (13). Once the values of $\hat{t}^f_{k_e+1}$ and $\hat{t}^b_{k_e}$ have been obtained, we need to interpolate the two indexing functions $\hat{\kappa}^f(t)$ and $\hat{\kappa}^b(t)$ in the interval $\hat{t}^b_{k_e} \leq t \leq \hat{t}^f_{k_e+1}$. The interpolation is unproblematic since $\hat{\kappa}^f(t)$ is known in $(t_0, 0), \ldots, (t_{k_e}, k_e), (\hat{t}^f_{k_e+1}, k_e+1)$ and $\hat{\kappa}^b(t)$ is known in $(\hat{t}^b_{k_e}, k_e), (t_{k_e+1}, k_e+1), \ldots, (t_K, K)$.

Adopting the LS criterion, estimation of s can be obtained by minimizing

$$\mathcal{E}(s) = \int_{\hat{t}^b_{k_e}}^{\hat{t}^f_{k_e+1}} \left(\hat{\kappa}^f(t) - \left(\hat{\kappa}^b(t) + s\right)\right)^2 dt. \quad (15)$$

Differentiation of $\mathcal{E}(s)$ with respect to s, and setting the result equal to zero, yields the value of s that minimizes $\mathcal{E}(s)$

$$\hat{s} = \frac{1}{\hat{t}^f_{k_e+1} - \hat{t}^b_{k_e}} \int_{\hat{t}^b_{k_e}}^{\hat{t}^f_{k_e+1}} \left(\hat{\kappa}^f(t) - \hat{\kappa}^b(t)\right) dt. \quad (16)$$

Hence, the estimator computes the area between the two indexing functions in the overlap interval, normalized with the length of the overlap interval. In practice, the computation of the integral in (16) is approximated with a sum over a set of discretized times.

With $\hat{s}$ available, it is possible to estimate the mean RR interval length T_0 according to

$$\hat{T}_0 = \frac{t_K}{K + \hat{s}}. \quad (17)$$

This expression is almost identical to that in (4), except for $\hat{s}$ which accounts for the delay in time due to the ectopic beat. Note that $\hat{s} = 0$, corresponding to the absence of an ectopic beat, results in (4).

2) A Computationally Efficient Method of the Heart Timing Signal: A different approach to deal with ectopic beats is to observe that an ectopic beat shifts the occurrence times of the following normal heartbeats by the time δ, and that we have from the definition of the IPFM model

$$\int_{t_{k-1}}^{t_k} (1 + m(\tau))\, d\tau = T_0 \quad k \neq k_e + 1. \quad (18)$$

Since the variations of $m(t)$ are unknown between t_{k_e} and t_{k_e+1} due to the ectopic beat, certain assumptions on $m(t)$ must be done in this interval to obtain a value of (18) for $k = k_e + 1$. From t_{k_e} to $t_{k_e+1} - \delta$, i.e., to the occurrence time that would follow t_{k_e} had no ectopic beat been present, $t^f_{k_e+1} = t_{k_e+1} - \delta$, it is assumed that the variations of $m(t)$ are the same as if no ectopic beat is present. The value of (18) for $k = k_e + 1$ is obtained from

$$\begin{aligned}\int_{t_{k_e}}^{t_{k_e+1}} &(1 + m(\tau))\, d\tau \\ &= \int_{t_{k_e}}^{t_{k_e+1}-\delta} (1 + m(\tau))\, d\tau + \int_{t_{k_e+1}-\delta}^{t_{k_e+1}} (1 + m(\tau))\, d\tau \\ &= T_0 + \delta \end{aligned} \quad (19)$$

where it has been assumed that the integral of $m(t)$ is zero in the remaining time interval $(t_{k_e+1} - \delta, t_{k_e+1})$ as if no ectopic beat had been present. Note that (19) becomes (18) when no ectopic beat is present, i.e., $\delta = 0$. Thus, if $t_e < t_k$, (1) becomes

$$\int_0^{t_k} (1 + m(\tau))\, d\tau = kT_0 + \delta \quad (20)$$

and (3) becomes

$$\int_0^{t_k} m(\tau)d\tau = kT_0 - t_k + \delta \equiv d_{HT_\delta}(t_k). \tag{21}$$

Hence, by estimating the time shift δ the presence of an ectopic beat can be accounted for by

$$d_{HT_\delta}(t_k) = \begin{cases} kT_0 - t_k & k = 0, \dots, k_e, \\ kT_0 - t_k + \delta & k = k_e + 1, \dots, K \end{cases} \tag{22}$$

where $d_{HT_\delta}(t_k)$ is the HT signal, $d_{\mathrm{HT}}(t_k)$, when the assumptions used in (19) about $m(t)$ around the ectopic interval are incorporated. Note that the functions in (8) and (22) are not identical, hence, $\delta \neq sT_0$ (see Section V). In order to estimate δ, we make use of (19) such that

$$\begin{aligned} T_0 &= \int_{t_{k_e}}^{t_{k_e+1}-\delta} (1 + m(\tau))\, d\tau \\ &= t_{k_e+1} - t_{k_e} - \delta + \int_{t_{k_e}}^{t_{k_e+1}-\delta} m(\tau)d\tau \end{aligned} \tag{23}$$

and

$$\delta = t_{k_e+1} - t_{k_e} - T_0 + \int_{t_{k_e}}^{t_{k_e+1}-\delta} m(\tau)d\tau. \tag{24}$$

We now introduce a new parameter, $\overline{m}_k$, crucial to the estimation of δ, defined by

$$\overline{m}_k = \begin{cases} \int_{t_k}^{t_{k+1}} m(\tau)d\tau & k \neq k_e, \\ \int_{t_{k_e}}^{t_{k_e+1}-\delta} m(\tau)d\tau & k = k_e \end{cases} \tag{25}$$

and, thus using (24) we can write

$$\delta = t_{k_e+1} - t_{k_e} - T_0 + \overline{m}_{k_e}. \tag{26}$$

For the special case of a constant HR, i.e., $\kappa(t)$ is linear, or, equivalently, $m(t) = 0$ and $\overline{m}_k = 0$, we obtain an estimator of δ according to

$$\hat{\delta}_0 = t_{k_e+1} - t_{k_e} - T_0 \tag{27}$$

referred to as the zero order estimator of δ. Although this equation cannot be used as an estimator of δ since T_0 is unknown, $\hat{\delta}_0$ is later used to derive a useful estimator. The corresponding estimator of $\overline{m}_k$ is denoted $\hat{\overline{m}}_{k,0} = 0$.

If we assume that the variations of $m(t)$ are small within the integration interval, the beat-to-beat variations in $\overline{m}_k$ are also small. Hence, an improved estimator of $\overline{m}_k$ would be the value corresponding to the previous occurrence time. This estimator, denoted $\hat{\overline{m}}_{k,1}$, is the first order estimator of $\overline{m}_k$, and can be calculated as the sum of $\hat{\overline{m}}_{k,0}$ and a first order difference of $\overline{m}_k$, denoted $\Delta\overline{m}_{k,1}$ (which is set to $\overline{m}_k$), according to

$$\begin{aligned} \hat{\overline{m}}_{k_e,1} &= \hat{\overline{m}}_{k_e,0} + \Delta\overline{m}_{k_e-1,1} \\ &= \overline{m}_{k_e-1} \\ &= \int_{t_{k_e-1}}^{t_{k_e}} m(\tau)d\tau \\ &= d_{\mathrm{HT}}(t_{k_e}) - d_{\mathrm{HT}}(t_{k_e-1}) \\ &= k_eT_0 - t_{k_e} - (k_e - 1)T_0 + t_{k_e-1} \\ &= t_{k_e-1} - t_{k_e} + T_0. \end{aligned} \tag{28}$$

Combining $\hat{\overline{m}}_{k_e,1}$ with (26) the first order estimator of δ is given by

$$\hat{\delta}_1 = t_{k_e+1} - 2t_{k_e} + t_{k_e-1}. \tag{29}$$

Note the similarity between (27) and (29), since (29) can be written as

$$\begin{aligned} \hat{\delta}_1 &= t_{k_e+1} - t_{k_e} - (t_{k_e} - t_{k_e-1}) \\ &= t_{k_e+1} - t_{k_e} - T_0 - (t_{k_e} - t_{k_e-1} - T_0) \\ &= \hat{\delta}_0 - \hat{d}_{k_e-1,0} \end{aligned} \tag{30}$$

where $\hat{d}_{k_e-1,0}$ is the zero order estimator of d_{k_e-1} ($\overline{m}_k = 0$), with d_k defined as

$$d_k = t_{k+1} - t_k - T_0 + \overline{m}_k = 0 \quad k \neq k_e. \tag{31}$$

Note also the close relationship between (26) and (31), since (31) becomes (26) when $k = k_e$. In order to better understand what $\hat{\delta}_1$ implies on the RR interval at the ectopic beat, (29) is rewritten according to

$$\left(t_{k_e+1} - \hat{\delta}_1\right) - t_{k_e} = t_{k_e} - t_{k_e-1} \tag{32}$$

or, equivalently

$$d_{IF}^i\left(\hat{t}_{k_e+1}^f\right) = d_{IF}(t_{k_e}). \tag{33}$$

Thus, $\hat{\delta}_1$ maintains continuity in the RR interval by replacing the RR interval at the ectopic beat with the previous RR interval, which implies a constant approximation of the HR variations during ectopy.

A higher order estimator of $\overline{m}_k$ would be to include variations in $\overline{m}_k$, which can be done in the second-order estimator denoted $\hat{\overline{m}}_{k,2}$. This estimator is obtained from $\hat{\overline{m}}_{k,1}$ by adding a second-order difference of $\overline{m}_k$, defined by

$$\begin{aligned} \Delta\overline{m}_{k,2} &= \Delta\overline{m}_{k,1} - \Delta\overline{m}_{k-1,1} \\ &= \overline{m}_k - \overline{m}_{k-1}. \end{aligned} \tag{34}$$

Thus, a second-order estimator of $\overline{m}_{k_e}$ is given by

$$\begin{aligned} \hat{\overline{m}}_{k_e,2} &= \hat{\overline{m}}_{k_e,1} + \Delta\overline{m}_{k_e-1,2} \\ &= \overline{m}_{k_e-1} + (\overline{m}_{k_e-1} - \overline{m}_{k_e-2}) \\ &= 2\int_{t_{k_e-1}}^{t_{k_e}} m(\tau)d\tau - \int_{t_{k_e-2}}^{t_{k_e-1}} m(\tau)d\tau \\ &= 2\left(d_{\mathrm{HT}}(t_{k_e}) - d_{\mathrm{HT}}(t_{k_e-1})\right) \\ &\quad - \left(d_{\mathrm{HT}}(t_{k_e-1}) - d_{\mathrm{HT}}(t_{k_e-2})\right) \\ &= 2\left(k_eT_0 - t_{k_e}\right) - 3\left((k_e - 1)T_0 - t_{k_e-1}\right) \\ &\quad + (k_e - 2)T_0 - t_{k_e-2} \\ &= 3t_{k_e-1} - 2t_{k_e} - t_{k_e-2} + T_0. \end{aligned} \tag{35}$$

The accuracy of $\hat{\overline{m}}_{k_e,2}$ may be sufficient when the variations in $\overline{m}_k$ are small. Combining $\hat{\overline{m}}_{k_e,2}$ with (26) will give us a second-order estimator of δ according to

$$\hat{\delta}_2 = t_{k_e+1} - 3t_{k_e} + 3t_{k_e-1} - t_{k_e-2}. \tag{36}$$

Note the similarity between (29) and (36), since (36) can be written as

$$\begin{aligned} \hat{\delta}_2 &= t_{k_e+1} - 2t_{k_e} + t_{k_e-1} - (t_{k_e} - 2t_{k_e-1} + t_{k_e-2}) \\ &= \hat{\delta}_1 - \hat{d}_{k_e-1,1}. \end{aligned} \tag{37}$$

Both (30) and (37) show that higher order estimators of δ can be obtained from the difference between lower order estimators of δ and d_{k_e-1}. The correct value of d_{k_e-1} is known and equals zero, e.g., (31). Since both the lower order estimators of δ and

d_{k_e-1} contain the same approximations, the estimator of d_{k_e-1} can be viewed as the error, made in the approximations, of δ. If (36) is rewritten in a similar way as in (32), we obtain

$$\left(t_{k_e+1} - \hat{\delta}_2\right) - t_{k_e} = t_{k_e} - t_{k_e-1} + ((t_{k_e} - t_{k_e-1}) - (t_{k_e-1} - t_{k_e-2})) \quad (38)$$

or, equivalently

$$d_{IF}^{i}\left(\hat{t}_{k_e+1}^{f}\right) = d_{IF}(t_{k_e}) + (d_{IF}(t_{k_e}) - d_{IF}(t_{k_e-1})). \quad (39)$$

Thus, $\hat{\delta}_2$ replaces the RR interval at the ectopic beat using linear interpolation from the previous RR interval, which implies a linear approximation of the HR variations during ectopy.

Generalization of the Method: A generalization of higher order estimators of δ is obtained when variations in $\overline{m}_k$ are included. If the estimator of $\overline{m}_k$ is updated according to

$$\hat{\overline{m}}_{k,p} = \hat{\overline{m}}_{k,p-1} + \Delta\overline{m}_{k-1,p} \quad (40)$$

where $\Delta\overline{m}_{k-1,p}$ is the p^{th} order difference of $\overline{m}_{k-1}$

$$\Delta\overline{m}_{k-1,p} = \Delta\overline{m}_{k-1,p-1} - \Delta\overline{m}_{k-2,p-1}. \quad (41)$$

It is shown in Appendix that the N^{th} order estimator of δ is given by the following recursive equation:

$$\hat{\delta}_N = \hat{\delta}_{N-1} - \hat{d}_{k_e-1,N-1} \quad N = 1, 2\ldots \quad (42)$$

where

$$\hat{\delta}_0 = t_{k_e+1} - t_{k_e} - T_0. \quad (43)$$

A similar recursive equation for the N^{th} order estimator of d_k is given by

$$\hat{d}_{k,N} = \hat{d}_{k,N-1} - \hat{d}_{k-1,N-1} \quad k \neq k_e \quad N = 1, 2, \ldots \quad (44)$$

and can be proven in the same way as (42).

Using Pascal's triangle and the expressions in (29), (31), (42), and (44), we can express $\hat{\delta}_N$ directly in terms of the occurrence times according to

$$\hat{\delta}_N = \sum_{l=0}^{N+1} (-1)^l \binom{N+1}{l} t_{k_e+1-l} \quad N = 1, 2, \ldots. \quad (45)$$

Recall that for $N = 0$, $\hat{\delta}_N$ is given by (43), but not useful as an estimator since T_0 is unknown. As N increases, higher order approximations of the HR variations are implied during ectopy in a similar way as in (33) and (39). Once $\hat{\delta}_N$ is obtained from (45), it is straightforward to estimate T_0 by

$$\hat{T}_0 = \frac{t_K - \hat{\delta}_N}{K}. \quad (46)$$

This expression is almost identical to that in (4), except for $\hat{\delta}_N$ which accounts for the delay in time due to the ectopic beat. The HT signal, $d_{HT_\delta}(t_k)$, in (22) can now be calculated since the required parameter values are available.

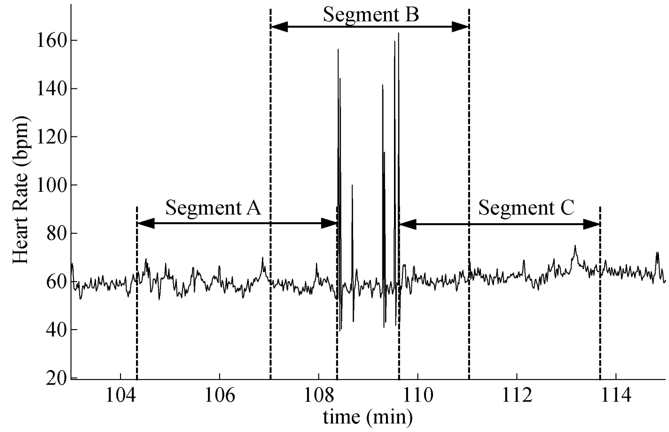


Fig. 2. The three different overlapping segments A, B, and C.

If more than one ectopic beat is present in the ECG, $\hat{\delta}_N$ has to be calculated for each ectopic beat. The δ_N estimator used in (46) is then the sum of all the different $\hat{\delta}_N$'s and the δ used in (22) should be the sum of the $\hat{\delta}_N$'s corresponding to the ectopic events prior to event time t_k. It should be pointed out that if two ectopic beats are next to each other then they can be treated as one. Furthermore, if two ectopic beats are too close to each other (the closeness depends on the order of the δ_N estimator in (45)), then $\hat{\delta}_N$ corresponding to the second ectopic beat is influenced, since $\hat{\delta}_N$ involves nearby occurrence times. If this is the case, the first ectopic beat is within the interval of the occurrence times used to estimate the second beat. In order to obtain a correct estimate of the second $\hat{\delta}_N$, these occurrence times must be corrected with the help of the previous estimated $\hat{\delta}_N$ (corresponding to the first ectopic beat). In this case, an appropriately estimated $\hat{\delta}_N$ is obtained by simply adding the previous estimated $\hat{\delta}_N$ to the occurrence times prior to the first ectopic beat, and applying (45) as before.

III. Database

The database consists of 132 ECG episodes selected from the European ST-T database, previously studied in [11] and [13]. The ectopic beat composition of the 132 episodes is as follows: 91 episodes containing one ectopic beat, 28 containing two, 5 containing three, 4 containing four, 2 containing eight, and 2 containing ten ectopic beats. Each ECG episode is divided into three overlapping, 4-min segments: A, B, and C, see Fig. 2. Segments A and C are ectopic-free, whereas segment B contains the ectopic beat(s). Segment A contains the 4 min preceding the ectopic beat(s), segment B is centered around the ectopic beat(s), and segment C contains the 4 min following the ectopic beat(s).

The database is studied using the evaluation approach introduced in [11], where it was suggested that the spectral characteristics of the ectopic-free segments A and C can be compared to the corrected segment B assuming that the HR is stationary once ectopy has been removed. Three parameters, ΔAC, ΔAB, and ΔBC, are defined, where ΔAC is the difference in spectral power between segment A and C, and so on. Moreover, the power is divided into two subbands: a low-frequency (LF) band (0.04–0.15 Hz) and a high-frequency (HF) band (0.15–0.40 Hz). Assuming stationarity during the

TABLE I
HRV POWER SPECTRAL DIFFERENCES OF THE s ESTIMATOR AND THE δ_1 ESTIMATOR WHEN USING THE HT SIGNAL FOR ALL THE 132 ECG EPISODES. VALUES ARE GIVEN IN MEAN ± STD IN THE UNIT ms^{-2}

	ΔAC		ΔAB		ΔBC	
Estimator	*LF*	*HF*	*LF*	*HF*	*LF*	*HF*
s	49 ± 808	–22 ± 190	–32 ± 644	–26 ± 153	80 ± 498	4 ± 158
δ_1			–32 ± 651	–30 ± 136	81 ± 501	8 ± 145

TABLE II
HRV POWER SPECTRAL DIFFERENCES OF THE HEART TIMING SIGNAL, $d_{\rm HT}(t)$, USING THE s ESTIMATOR AND THE HEART RATE SIGNAL, $d_{\rm HR}(t)$, USING SPLINE INTERPOLATION. THE VALUES HAVE BEEN IMPORTED FROM [11]. VALUES ARE GIVEN IN MEAN ± STD IN THE UNIT ms^{-2}

	ΔAC		ΔAB		ΔBC	
Signal	*LF*	*HF*	*LF*	*HF*	*LF*	*HF*
$d_{HT}(t)$	32 ± 670	–20 ± 186	–71 ± 575	–58 ± 213	103 ± 415	38 ± 177
$d_{HR}(t)$	31 ± 658	–17 ± 140	–1256 ± 11767	–331 ± 2323	1288 ± 11657	314 ± 2292

segments A, B, and C, ΔAC is expected to be close to zero in both frequency bands, and following correction of segment B, ΔAB and ΔBC should be close to that of ΔAC.

IV. RESULTS

The performance of the δ_N estimator in (45), based on $d_{HT_\delta}(t_k)$, is compared to that of the s estimator in (16), based on $d_{\rm HT}(t_k)$, in terms of HRV power spectral differences. The HRV power spectra of the ectopic-free segments A and C are computed using (3), whereas the power spectrum of segment B requires that either the δ_N or s estimator is used.

Table I presents the results when all 132 ECG episodes are analyzed. The performance of the two different estimators is almost identical for both the LF and HF bands of ΔAB and ΔBC, and is comparable to the variation of ΔAC. Table II is included to compare the results of Table I with those presented in [11], obtained using two different signals, namely, the HT signal, $d_{\rm HT}(t)$, and the HR signal, $d_{\rm HR}(t)$ (the latter signal is obtained by interpolation of the HR at each event time and ectopic beat correction using spline interpolation). These two different signals produce similar variation as reflected by the ΔAC column of Table II, and is similar to that obtained in the present paper, see Table I. The results of Tables I and II for $d_{\rm HT}(t)$, using the s estimator, are also similar, the slight difference being due to different implementations. When comparing the performance of the two different methods in Table II, it is obvious that $d_{\rm HT}(t)$ has better performance.

In order to compare complexity of the two estimators, the number of floating point operations (flops) was studied, see Table III. The results show that the s estimator requires almost 3000 times more flops than the δ_1 estimator. It is also noted that the number of flops used by the δ_N estimator is deterministic since it is, in contrast to the s estimator, independent of where the ectopic beat occurs.

The performance of the δ_N estimator for different orders was also compared in both ECG episodes with only one ectopic beat and all 132 ECG episodes, see Table IV. In episodes with only one ectopic beat, the performance of the δ_1 estimator was found to be the best, although δ_2 and δ_3 have similar performance. However, the performance of the δ_2 and δ_3 estimators differ from that of δ_1 when episodes with more than one ectopic beat are also included. Thus, the first-order estimator δ_1 was found to produce the best performance. For higher order estimators, the performance deteriorates.

TABLE III
FLOP STATISTICS FOR THE s ESTIMATOR AND THE δ_1 ESTIMATOR, WHEN USING THE HT SIGNAL FOR THE 132 ECG EPISODES

Estimator	*Mean*	*Std*
s	22514	26359
δ_1	8	7

TABLE IV
HRV POWER SPECTRAL DIFFERENCES WHEN USING DIFFERENT ORDERS OF THE δ_N ESTIMATOR IN COMBINATION WITH $d_{HT_\delta}(t)$. (A) HRV POWER SPECTRAL DIFFERENCES FOR ECG EPISODES CONTAINING *one* ECTOPIC BEAT. (B) HRV POWER SPECTRAL DIFFERENCES FOR *all* THE 132 ECG EPISODES. VALUES ARE GIVEN IN MEAN ± STD IN THE UNIT ms^{-2}

(a)

One	ΔAB		ΔBC	
Estimator	*LF*	*HF*	*LF*	*HF*
δ_1	–17 ± 539	–16 ± 105	90 ± 482	3 ± 118
δ_2	–18 ± 537	–16 ± 103	91 ± 481	3 ± 118
δ_3	–19 ± 535	–21 ± 102	93 ± 480	7 ± 120

(b)

All	ΔAB		ΔBC	
Estimator	*LF*	*HF*	*LF*	*HF*
δ_1	–32 ± 651	–30 ± 136	81 ± 501	8 ± 145
δ_2	–41 ± 663	–31 ± 139	90 ± 516	9 ± 150
δ_3	–88 ± 771	–57 ± 210	137 ± 662	35 ± 184

V. DISCUSSION AND CONCLUSION

The present paper sheds new light on the problem of ectopic beat correction in HRV analysis by introducing a new HT-based method. The performance was compared to the original method [11] and was found to be the same when performance is measured in power spectral terms. However, the new method requires a dramatically smaller amount of computations and is, therefore, much better suited for implementation in systems for long-term ECG analysis. Although the implementation of the s estimator can be further optimized than what was done in the present study, the difference in estimator complexity will nevertheless persist since the δ_1 estimator merely requires two additions.

Higher order estimators did not produce results better than those of the first-order estimator δ_1 (see Table IV). Such a result may, at a first glance, appear to be unexpected since a higher order estimator provides, from a modeling viewpoint, a more accurate description of $m(t)$. However, a number of aspects must be taken into account when processing real ECG data which explains the performance relationship between estimators of different orders.

An important aspect is that performance is expressed in terms of differences in spectral power between the intervals A and C (assumed to be stationary both within and across each other) as compared to differences between B and either A or C. The differences between A and C given in Table I represent a bound of the performance results; results obtained from interval B and A or C being lower than these values do not improve performance since they are accompanied by an increase with respect to the complementary interval, i.e., C or A. Thus, since a first-order estimator already produce results which are close to the bound given in Table I, improvements associated with higher order estimators remain obscured by the natural variability from A to C. It is also worthwhile to mention that even if the intervals A and C would be perfectly stationary, differences can still arise due to discrepancies between the IPFM model and the underlying physiology.

Another aspect to be considered when using higher order estimators is their use of higher order differentiation. Since beat locations are estimated with a certain error, which at best is lower bounded by the time quantization of the sampling rate, such errors show up as noise in the HT signal. It is well-known that noise becomes amplified as the order of differentiation increases.

Finally, the finding that the δ_1 estimator exhibits much better performance in Table IV(b) for the multiectopic case is explained by the fact that several ectopic beats are present and, sometimes, occurring closely in time. Since higher order estimators rely on occurrence times farther away they are more influenced by adjacent ectopic beats than is the first-order estimator, the results of Table IV(b) are inferior to those of IV(a).

By comparison of (8) and (22) one may conclude that $\delta = sT_0$, a statement which is incorrect. The two functions $d_{\mathrm{HT}}(t_k)$ and $d_{HT_\delta}(t_k)$ model the same events, but are not equal since they involve different assumptions on the variations of $m(t)$ around the ectopic beat and, thus, the corresponding estimators of δ and sT_0 differ. It remains to be established whether an approach based on the assumption of $\delta = sT_0$ offers particular advantages over the present approach.

Appendix

In this Appendix, the proof of the following recursion

$$\hat{\delta}_N = \hat{\delta}_{N-1} - \hat{d}_{k_e-1,N-1} \quad N = 1, 2, \ldots \tag{47}$$

being the N^{th} order estimator of δ, is given. The validity of (47) for $N =$ 1,2 has already been proven in connection with (30) and (37). However, following induction proof shows the validity of (47) for all $N = 1, 2, \ldots$. Thus, we must prove that:

1) (47) is valid for $N = 1$;
2) if (47) is valid for $N = p - 1$, then (47) also is valid for $N = p$.

The first statement has already been proven according to (30), and the second statement is proven by first concluding [e.g., with the help of (26), (27), (31), and (40)] that

$$\begin{cases} \hat{\delta}_p = \hat{\delta}_{p-1} + \Delta\overline{m}_{k_e-1,p} \\ \hat{d}_{k,p} = \hat{d}_{k,p-1} + \Delta\overline{m}_{k-1,p} \quad k \neq k_e \end{cases} \tag{48}$$

and

$$\begin{aligned} \Delta\overline{m}_{k_e-1,p} &= \Delta\overline{m}_{k_e-1,p-1} - \Delta\overline{m}_{k_e-2,p-1} \\ &= \hat{\delta}_{p-1} - \hat{\delta}_{p-2} - \left(\hat{d}_{k_e-1,p-1} - \hat{d}_{k_e-1,p-2}\right) \\ &= [(47) \text{ valid for } N = p-1] \\ &= -\,\hat{d}_{k_e-1,p-1}. \end{aligned} \tag{49}$$

Then, (47) is valid for $N = p$, since

$$\hat{\delta}_p = \hat{\delta}_{p-1} - \hat{d}_{k_e-1,p-1} \tag{50}$$

and (47) is thereby valid for all $N = 1, 2, \ldots$

Acknowledgment

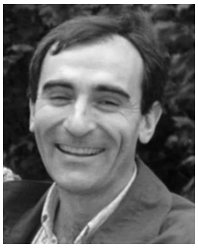
The authors would like to thank Dr. J. Mateo for generously providing them with the ECG data for performance evaluation.

References

[1] M. N. Cheung, "Detection of recovery from errors in the cardiac interbeat intervals," *Psychophysiol.*, vol. 18, pp. 341–346, 1981.
[2] M. V. Kamath and E. L. Fallen, "Correction of the heart rate variability signal for ectopics and missing beats," in *Heart Rate Variability*, M. Malik and A. J. Camm, Eds. Armonk, NY: Futura, 1995, ch. 6, pp. 75–86.
[3] G. A. Myers, G. J. Martin, N. M. Magid, P. S. Barnett, J. W. Schaad, J. S. Weiss, M. Lesch, and D. H. Singer, "Power spectral analysis of heart rate variability in sudden cardiac death: comparison to other methods," *IEEE Trans. Biomed. Eng.*, vol. BME-33, pp. 1149–1156, 1986.
[4] C. L. Birkett, M. G. Kienzle, and G. A. Myers, "Interpolation of ectopics increases low frequency power in heart rate variability spectra," *Comput. Cardiol.*, pp. 257–259, 1991.
[5] L. J. M. Mulder, "Measurement and analysis methods of heart rate and respiration for use in applied environments," *Biol. Psychol.*, vol. 34, pp. 205–236, 1992.
[6] N. Lippman, K. M. Stein, and B. B. Lerman, "Comparison of methods for removal of ectopy in measurements of heart rate variability," *Am. J. Physiol.*, vol. 267, pp. H411–H418, 1994.
[7] P. Albrecht and R. J. Cohen, "Estimation of heart rate power spectrum bands from real world data: dealing with ectopic beats and noise data," *Comput. Cardiol.*, pp. 311–314, 1988.
[8] J. Mateo and P. Laguna, "Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 8, pp. 997–1009, Aug. 2000.
[9] B. W. Hyndman and R. K. Mohn, "A model of the cardiac pacemaker and its use in decoding the information content of cardiac intervals," *Automedica*, vol. 1, pp. 239–252, 1975.
[10] J. Mateo and P. Laguna, "Extension of the heart timing signal HRV analysis in the presence of ectopic beats," *Comput. Cardiol.*, pp. 813–816, 2000.
[11] ——, "Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 334–343, Mar. 2003.
[12] M. Brennan, M. Palaniswami, and P. Kamen, "A new model-based ectopic beat correction algorithm for heart rate variability," in *Proc. 23rd Annu. Conf. IEEE EMBS* , vol. 1, Oct. 2001, pp. 567–570.
[13] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi, "The European ST-T database: standards for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography," *Eur. Heart J.*, vol. 13, pp. 1164–1172, 1992.

Kristian Solem received the M.Sc. degree in electrical engineering from Lund University, Lund, Sweden, in 2003. He is currently working towards the Ph.D. degree in signal processing and its application to biomedical signals.

His main research activity is in the field of ECG signals and their significance in heart rate variability analysis and hemodialysis. He is also involved in teaching basic courses in Signal Processing.



Pablo Laguna (M'92) was born in Jaca (Huesca), Spain, in 1962. He received the M.S. degree in physics and the Ph.D. degree from the Science Faculty at the University of Zaragoza, Zaragoza, Spain, in 1985 and 1990, respectively. His Ph.D. thesis was developed at the Biomedical Engineering Division of the Institute of Cybernetics (U.P.C.-C.S.I.C.) under the direction of P. Caminal.

He is Full Professor of Signal Processing and Communications in the Department of Electrical Engineering at the Engineering School, and a Researcher at the Aragn Institute for Engineering Research (I3A), both at University of Zaragoza, Spain. From 1992 to 2005, he was Associate Professor at same university and from 1987 to 1992 he worked as Assistant Professor of Automatic control in the Department of Control Engineering at the Politecnic University of Catalonia (U.P.C.), Catalonia, Spain, and as a Researcher at the Biomedical Engineering Division of the Institute of Cybernetics (U.P.C.-C.S.I.C.). His professional research interests are in signal processing, in particular applied to biomedical applications. He is, together with L. Sörnmo, the author of *Bioelectrical Signal Processing in Cardiac and Neurological Applications* (Elsevier, 2005).

Leif Sörnmo (S'80–M'85–SM'02) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1978 and 1984, respectively.

He held a research position with the Department of Clinical Physiology, Lund University, from 1983 to 1995, where he worked on computer-based ECG analysis. Since 1990, he has been with the Signal Processing Group, Department of Electroscience, Lund University, where he now holds a position as a Professor in biomedical signal processing. His main research interests include statistical signal processing and modeling of biomedical signals. His current research projects include methods in ischemia monitoring, time-frequency analysis of atrial fibrillation, power efficient signal processing in pacemakers, hemodialysis, and detection of otoacoustic emissions. He is, together with P. Laguna, the author of *Bioelectrical Signal Processing in Cardiac and Neurological Applications* (Elsevier, 2005).

Dr. Sörnmo has been an Associate Editor of *Computers in Biomedical Research* (1997–2000). He is currently on the editorial boards of the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and the *Journal of Electrocardiology*.