16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

A TIME DELAY ESTIMATION TECHNIQUE FOR OVERLAPPING SIGNALS IN ELECTROCARDIOGRAMS

Aline Cabasson and Olivier Meste

Laboratoire I3S - CNRS - Université de Nice-Sophia Antipolis
2000, route des Lucioles - 06903 Sophia Antipolis, France
phone: + (33) 492942741, fax: + (33) 492942898
email: {cabasson,meste}@i3s.unice.fr

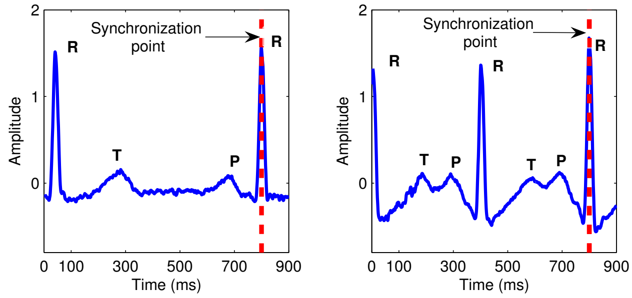
ABSTRACT

This paper presents a new technique for the estimation of the PR intervals from the electrocardiograms (ECG) taking into account the influence of the T wave overlapping the P wave during high heart rate. The decreasing part of the T wave is modelled by a piecewise linear function, in a maximum likelihood approach. This leads to improve the estimation of the PR intervals in exercise and recovery. Experimental results show that this new method gives more accurate estimation of the PR intervals both in synthetic data and in real data. Also, we show that measurements of PR and RR intervals permit us to classify athletes in function of their training degree.

1. INTRODUCTION

Estimating the heart periods variability during exercice and recovery is a real challenge in biomedical engineering. One reason is that the global understanding of the interaction between the neural activity and the cardiac outputs would be relevant to improve the performances of the future pacemakers. For instance, the analysis of the PR interval trend could be performed in order to evaluate the sympathetic-parasympathetic balance but also to reveal the atrioventricular conduction properties [1].
However, the analysis of PR intervals during exercice and recovery has been rarely addressed to date [2]. The main reason is that the estimation of these intervals is particularly biased at high heart rates, since the T wave overlaps the P one (see figure 1). Consequently, new ad-hoc time delay estimators have to be designed.

Among the well-known time delay estimators, the techniques based on the detection of the maximum of the cross correlation is not relevant to estimate the PR intervals at high heart rates [3, 4]. Indeed, since the overlapping of the P and T waves is not considered, the latter introduces a bias when estimating the PR intervals. Finally, one solution to efficiently estimate the PR intervals should be to model this overlapping in order to reduce the influence of the T wave presence. In parallel, when the signal under consideration is unknown, the Woody technique [5], is a candidate that belongs to the cross correlation family. This method, that has been recently put in light [6], has to be extended to account for the mentioned overlapping. In addition, it has been shown that the Woody technique is suboptimal [7, 8]. Moreover, for time delay estimation problems, several other techniques have been used mainly working in the frequency domain [6, 9, 10]. Unfortunately, the introduction of *a priori*



(a) At rest. (b) During exercise.

Figure 1: Example of a real ECG where the T and P waves are disjointed at rest (a), and overlapped during exercise (b).

concerning the T wave is not an easy task in the frequency domain. Conversely, it is well described in the time domain, i.e. the monophasic behaviour. Also, the wavelet transform is a good candidate for bioelectrical signals analysis [11]. Although it has been shown to be a promising tool, these typical methods do not overcome our problem because of the high level of overlapping between the T and the P waves in exercise (see section 2).

We therefore propose in this paper the development of an original approach which improves significantly the estimation of the PR intervals, during exercice or at rest. Our method combines: i) an efficient modelling of the T wave which allows the cancellation of an expected overlapping with the P wave; ii) an improved version of the Woody technique using an iterative Maximum-Likelihood Estimator (MLE) [7, 8]. As it will be presented in the sequel, the global modelling of the observations will involve the unknowns in a non-linear manner. Unlike the similar models introduced by [9, 10], we will assume that a ML estimator exists. In [6, 12], it has been shown that, when this model identification is addressed with a ML approach, saddle point singularities appear. Thus, this theoretical assessment will not be in the scope of this paper.

The rest of the paper is organized as follows. Section 2 introduces the problem statement and the proposed approach. Section 3 deals with the complete description of the proposed T wave modelling. Section 4 gives the validation of the proposed T wave modelling on synthetic data of ECG, and one application on real data: the PR intervals estimation on ECG recorded in exercise test. Finally, we conclude and present some future works in section 5.

2. PROBLEM STATEMENT AND CONTRIBUTION

The goal of our current work is to estimate the PR intervals on ECG recorded in exercise test. Since the T and P waves are overlapping at high heart rates, it introduces a bias in the estimation of the PR intervals.

We emphasize that the proposed method of time delay estimation is processed in the time domain [5, 7, 8]. However among the reference methods, the most recent exploit the frequency domain because the observation noise covariance is unknown, and the real delays are not integer [9, 10]. Note that this technique assumes that a set of observations is available, where the unknowns are the amplitude jitter, the time delay and, the reference signal itself [10]. A more general model is provided by [6], where the reference signal is in fact a mixture of several unknown signals.
When *a priori* of the signal shape is available, the wavelet transform (WT) provides a good estimation of intervals based on the detection of singularity in the time scale domain [11]. Unfortunately, it does not overcome the problem of T-P waves fusion in exercise. This statement is supported by the observation that the two waves share a common frequency range. Then, the zero-crossings of the WT, indicating the location of the signal shape variation points, fails. Similarly to the WT assumption, our *a priori* information will be expressed in the time domain; we will assume a white gaussian noise for the error measurement and a positive monophasic shape for the T wave.
The T-P waves fusion analysis could be simplified by using a known constant T wave, chosen as a template. This template could be optimally segmented from a real T wave in the resting period. Unfortunately, the shape of the T wave varies as long as the effort increases, discarding this simple approach. Moreover, methods based on spline interpolation can not be applied because the anchor points of the T wave are hidden when the fusion occurs.

These previous comparisons justify the extension of the T wave modelling techniques using a straight line [13], or a constrained 3^{rd} order polynomial function [14]. The technique proposed in this paper is a refined modelling of the T wave by a piecewise linear function imposing additional constraints. These constraints will account for the decreasing behaviour of the ending part of the T wave (see figure 2).

The outline of our PR intervals estimator algorithm is:

1. modelling of the decreasing part of the T wave by a piecewise linear function imposing additional constraints;
2. inclusion of the T wave modelling in our global observations model;
3. estimation of the PR intervals with our Woody Improved method [7, 8]: an iterative estimation technique based on Maximum Likelihood which includes Least Square problem with linear inequality constraints.

3. GLOBAL MODEL ESTIMATION

In this section, we explain the proposed modelling to take into account the T-P waves fusion. The main improvement is due to the addition of a decreasing piecewise linear function, assumed to fit the ending part of the T wave, in the observations model (see figure 2).

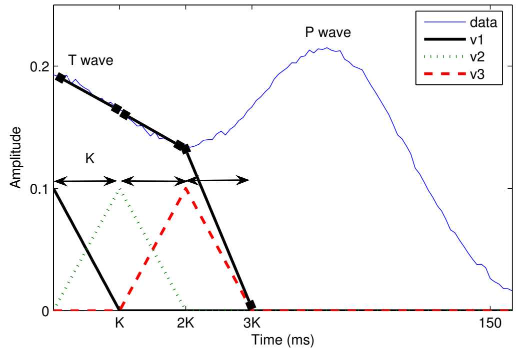


Figure 2: Example of a real T wave modelled as a decreasing piecewise linear function based on 3 basic functions.

Before the PR estimation, two pre-processing methods provide us the position of the R waves and rough P wave localizations [3]. A threshold technique applied on the high-pass filtered and demodulated ECG, refines the estimation of the time occurrences of the R waves. Then, we obtain segments including each P wave, and synchronized with its next R wave. The length of the segments is fixed for all beats. Whatever the heart rate, the left boundary of the segment is adjusted in order to get only the decreasing part of the T wave, and to ensure that the whole P wave is included (see figure 2). In a real case, this condition is readily fulfilled and the T wave should not be present in our observation window for low heart rates.

We introduce a model where the observation $x_i(n)$ represents the sample in the segment i (for all n), of the considered i^{th} PR interval (with $i = 1..I$, I the number of realizations). Each observation contains $s_{d_i}(n)$, considered unknown here, assumed to be the P wave delayed by d_i as $s_{d_i}(n) = s(n-d_i)$, plus $e_i(n)$ an observation's noise. Since the T-P fusion occurs during exercise tests, we consider the T wave represented by a function $f(n;\theta_i)$ linearly parameterized. Finally, our observations model is expressed as:

$$x_i(n) = \alpha_i . s_{d_i}(n) + \alpha_i . f_{d_i}(n;\theta_i) + e_i(n). \qquad (1)$$

Without loss of generality, the identifiability of the model will impose a constraint such as $\sum_i d_i = 0$. Then the d_i, consequently the i^{th} PR interval, will be estimated up to an unknown constant.
The T wave is considered as a piecewise linear function $f(n;\theta_i)$, defined as a weighted sum of basic functions $\mathbf{v}_l$:

$$f(n;\theta_i) = \sum_{l=1}^{L} \theta_i[l].v_l[n]. \qquad (2)$$

We build a collection of L basic functions which defines L intervals of width K. L and K are chosen arbitrarily such as the length $L \times K$ corresponds to the expected maximal width of the segmented decreasing part of the T wave. The accuracy of this knowledge is not crucial. But it has to be chosen in accordance with the trade-off between the good approximation of the T wave and the variance of the estimated weights of f (2). Given the estimation process, increasing the number of basic functions will reduce the approximation error while the variance of the estimated weights increases.

As it will be shown in the sequel, this modelling will provide us a tractable solution that accounts for our *a priori* information concerning the T wave. So, it is expected that this knowledge reduces the bias avoiding the nonexistence of an unique solution.

In the figure 2, we choose basic functions $\mathbf{v}_l$ as piecewise linear functions. In order to be consistent with the observations, some constraints are added:

- in each interval, a negative slope is imposed for any basic functions combination;
- in order to keep the continuity of the modelled T wave, the joining points between two consecutive intervals must respect the following constraint: the last point of the l^{th} interval must be identical to the first point of the $(l+1)^{th}$ interval.

The aim is to build a collection of L basic functions. We choose arbitrarily $L=3$ as in the figure 2 (Note that L does not influence a lot the bias of the estimator on simulated data.). Therefore, we consider three intervals.

Choosing basic functions as in figure 2, on each interval, for $n \in [k\times K : (k+1)\times K]$ (with $k = 0\dots2$), we model the T wave by a linear function that is a weighted sum of two non zero basic functions:

$$\begin{cases} f[n] &= \theta_1.v_1[n] + \theta_2.v_2[n];\ n \in [0:K], \\ f[n] &= \theta_2.v_2[n] + \theta_3.v_3[n];\ n \in [K:2K], \\ f[n] &= \theta_3.v_3[n];\ n \in [2K:3K]. \end{cases}$$

Moreover, we want to model the T wave by a decreasing linear function, so we need to impose the following conditions on each interval:

$$\begin{cases} f'[n] &= \theta_1.v_1'[n] + \theta_2.v_2'[n] \le 0;\ n \in [0:K], \\ f'[n] &= \theta_2.v_2'[n] + \theta_3.v_3'[n] \le 0;\ n \in [K:2K], \\ f'[n] &= \theta_3.v_3'[n] \le 0;\ n \in [2K:3K], \end{cases}$$

where f' stands for the temporal derivative of f.

In order to obtain a tractable relation linking the coefficients θ_i, we choose arbitrarily the 3 basic functions as:

$$\begin{cases} v_1'[n] < 0;\ n \in [0:K], \\ v_1'[n] = -v_2'[n];\ n \in [0:K], \\ v_2'[n] = -v_3'[n];\ n \in [K:2K]. \end{cases} \tag{3}$$

These relations imply that v_1 and v_3 are decreasing respectively on the intervals $[0:K]$ and $[2K:3K]$.

Imposing these properties to the basic functions, we need to check the conditions of continuity at the joining points ($n=K$ and $n=2K$) between two consecutive intervals. Thus, for example when $n=K$, we get:

$$\theta_1.v_1[K] + \theta_2.v_2[K] = \theta_2.v_2[K] + \theta_3.v_3[K]. \tag{4}$$

However, using (3), on each interval we obtain the relations:

$$\begin{cases} v_1[n] = -v_2[n] + C_1;\ n \in [0:K], \\ v_2[n] = -v_3[n] + C_2;\ n \in [K:2K], \end{cases} \tag{5}$$

where C_1 and C_2 are constant values.

By replacing in (4), the condition of continuity in K becomes:

$$(\theta_2-\theta_1).v_2[K] + \theta_1.C_1 = (\theta_3-\theta_2).v_3[K] + \theta_2.C_2. \tag{6}$$

We impose that $v_1[K] = v_3[K] = 0$, which implies given the relations (5):

$$\begin{cases} v_2[K] = C_1, \\ v_2[K] = C_2. \end{cases} \tag{7}$$

The condition of continuity (6) for $n=K$ becomes:

$$\theta_2.v_2[K] = (\theta_2-\theta_3).v_2[K] + \theta_3.C_2.$$

Thanks to this relation and (7) we have $C_1 = C_2$ for all θ_k, so the continuity for $n=K$ is ensured.

Finally, when building a collection of L basic functions such as in figure 2, we can apply these rules:

- the first basic function is decreasing on the interval $[0:K]$ and is null after,
- the last basic function is null for $n \in [0:(L-2)K]$ and is decreasing on the interval $[(L-1)K:LK]$.

This implies that θ_L must be positive in order to keep the decreasing property of the modelled T wave. Besides, thanks to the hypotheses (3) and (5), the constraints on the θ_i's are:

$$\forall\ l \in [1:L-1],\ \theta_i[l] > \theta_i[l+1] > 0. \tag{8}$$

Note that the previous development has been given without lack of generality since it is valid for any number of basic functions, L.

Thanks to the proposed method, we can solve the problem of estimation of the PR intervals, i.e. the d_i's, considering the observations model as (1). Indeed, we combine the presented modelling of the decreasing part of the T wave with our Woody Improved method [7, 8]. For the estimation of the d_i's in this case, we modify our Woody Improved method based on an iterative Maximum Likelihood (ML) in order to transform it as a sum of Least Squares (LS) problems. Also, one advantage of the presented approach is that the inequality constraints introduced here (8) can be easily included in the LS solution. For this purpose, the LS with linear-Inequality constraints scheme (LSI problem) converted in a Least Distance Programming (LDP problem), is applied [15].

4. EXPERIMENTAL RESULTS

In this section we present some experimental results relative to the technique proposed in this paper: a new T wave modelling combined to our time delay estimation method. In a first time, we valid our modelling for the T wave on synthetic data of ECG, and in a second time we propose to apply it on ECG.

4.1 Validation of the T wave modelling

The synthetic data of ECG in exercise are presented in figure 3. This ECG has constant PR intervals, and a time-varying T-P distance. The overlapping ratio is higher as the beat number increases. On figure 3, the extreme left-hand side and the extreme right-hand side T waves correspond respectively to the 50^{th} and the 200^{th} beat number.

Figure 4 shows the evolution of the bias between the real PR intervals and the estimated ones (in function of the beat number) obtained with our technique (thick blue curve). This figure also shows the bias obtained with two other time delay estimators: the green curve and the red dashed one are using a T wave modelling based on a constrained 3^{rd} order polynomial function [14] or a decreasing single straight line [13], respectively. The black dotted curve corresponds to an estimation of the PR intervals without any modelling of the overlapping effect.

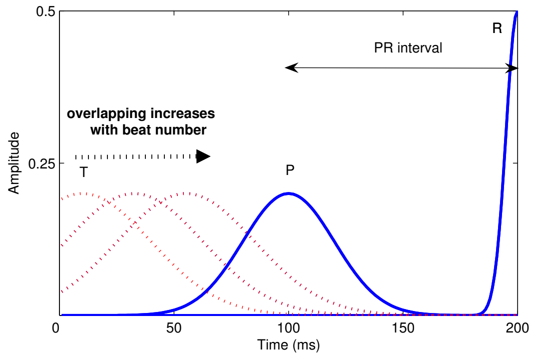


Figure 3: Synthetic data of ECG in exercise. During exercise, the T wave overlaps more and more the P wave.

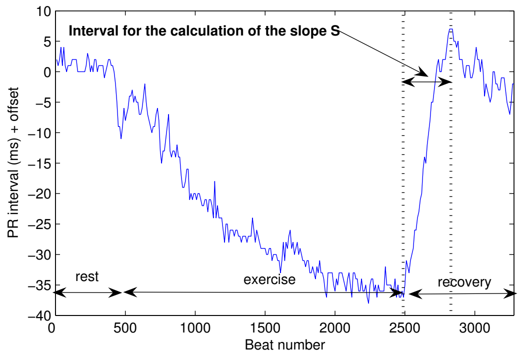


Figure 5: PR intervals for one subject. The interval used for the calculation of the slope S, indicative of the "recovery rate", is delimited by the two dotted vertical lines.

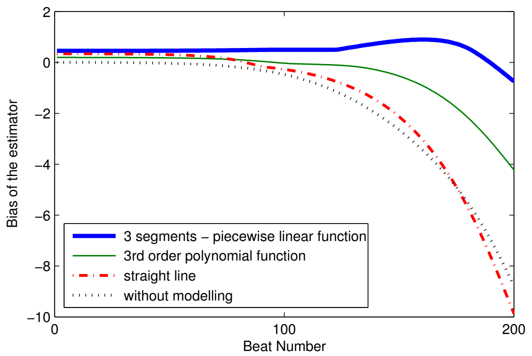


Figure 4: Bias of the time delay estimator for the three considered models.

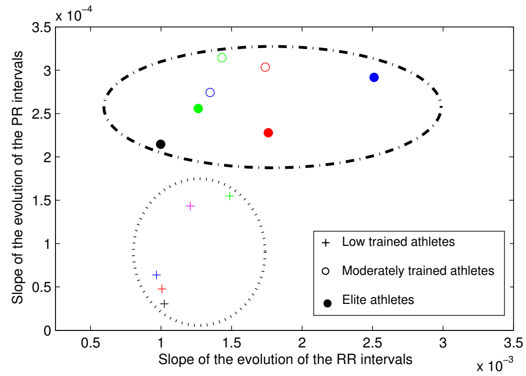


Figure 6: Relationship between the slope of the evolutions of the PR and RR intervals for the 12 subjects. Two groups can be clustered according to the PR slopes: one constituted by the low trained athletes, and the second by the remainder.

We observe that whatever the technique used, the bias is low when the T-P fusion does not exist. But when the beat number increases, we observe that our technique outperforms the others time delay estimators. It shows that, at least in simulation, our T wave modelling is more accurate than the previous ones when the overlapping appears.

	LTA	MTA	EA
Hours of training/week	< 5	$5 <$ and < 10	> 10
V_{O_2}max [ml/min/kg]	< 55	$55 <$ and < 65	> 70

Table 1: Definition of the differently trained athletes.

4.2 Application to real ECG

Now we present some results on data acquired from real exercise tests. The real ECG recordings are from 12 healthy men: 5 Low Trained Athletes (LTA), 3 Moderately Trained Athletes (MTA) and 4 Elite Athletes (EA). The definition of these different groups of athletes is presented in the table 1. The subjects performed a graded and maximal exercise test on a cycle.

For all subjects, in order to estimate the PR intervals, we compute the technique based on our Woody Improved method combined with a T wave modelling by a decreasing piecewise linear function ($L = 3$), presented in section 3. The PR intervals estimation is computed on 10 iterations assuming the algorithm convergence at the end. We obtain for each subject a result similar to figure 5 which presents the evolution of the PR intervals up to an unknown offset.

As noted in a previous work [13], it exists during recovery, an abrupt change of slope of the PR intervals which is significantly correlated with the RR's one. The location of this change of slope is specific for each subject. Thus, for all subjects, we calculate the slopes of the PR and RR intervals on a specific time interval. This latter is delimited by the dotted vertical lines on figure 5 between the end of the exercise and the abrupt change.

Figure 6 relates the slopes of the PR and RR intervals for the 12 subjects of our study. The slopes of the PR intervals are lower than those of RR intervals, which is expected since the total variation of the PR intervals is lower. Besides, we observe that for the low trained athletes, the slopes of the PR evolution are significantly lower.
It is clear that from the figure 6, two groups can be clustered according to the PR slopes: one for the low trained subjects,

and another one for the remainder. In order to confirm this observation, we have used a k-means clustering algorithm on our data set. This algorithm confirms these clusters according to the slope on PR intervals. Indeed, while exploiting data of both PR and RR intervals, we obtain 33% of misclassification, 60% considering the data of the RR intervals only and 0% considering the data of the PR intervals only. This confirms the clustering but also the difficulty to distinguish the moderately trained and elite athletes.

5. CONCLUSION AND FUTURE WORKS

In this paper, we aim to determine the PR intervals from ECG recorded in exercise test. But the localization of the P wave is a real challenge when the T wave overlapping appears at high heart rates. In this study, we propose an extension of the previous models [13, 14]: we fully describe the modelling by a decreasing piecewise linear function for the presence of the T wave merging in the P one. We include it in our Woody Improved method based on an iterative Maximum-Likelihood approach for the PR intervals estimation [7, 8].
On synthetic data, we conclude that our method combined with the new modelling exhibits better performances than the previous ones according to the bias of the estimators, especially when the overlapping ratio is high.
Furthermore, the exploitation of the PR intervals estimation on real data exhibits a previously found phenomena [14]: for the low trained athletes, the slope of the PR intervals in the early phase of the recovery is lower than for moderately trained and elite athletes. It is known that despite similar resting heart rate and stroke volume, athletes compared to sedentary men significantly enhance stroke volume, ventricular filling, and cardiac contractility during incremental exercise [16]. This suggests that it could exist a mechanical effect on the atrioventricular node more important for trained athletes which could explain the difference of the PR slopes between the low trained and the high trained athletes.

This work presents several perspectives. For instance, the method assumed a fixed segmentation window synchronized with the R wave. A refinement to this definition would be to adapt its position using the Bazett correction or any Q-T predictor [17]. Also, thanks to the estimation of the derivative when the signal is embedded in noise [18], we could find the position of the maximum of the T wave and build a matched window.

Note also that this method could be easily applied for the problem of the QT or ST intervals estimation, recorded in exercise test, accounting for a similar P wave modelling.

Acknowledgment

The authors would like to thank Gregory Blain and Stéphane Bermon for providing the data of ECG recorded in exercise.

REFERENCES

[1] R. Shouldice, C. Heneghan, P. Nolan, and P.G. Nolan, "PR and PP ECG Intervals as Indicators of Autonomic Nervous Innervation of the Cardiac Sinoatrial and Atrioventricular Nodes," *IEEE EMBS Conference in Neural Engineering*, pp. 261–264, 2003.

[2] Task Force of the European Society of Cardiology the North American Society of Pacing and Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, pp. 1043–1065, 1996.

[3] O. Meste, G. Blain, and S. Bermon, "Hysteresis Analysis of the PR-PP relation under Exercise Conditions," in *Computers In Cardiology*, 2004, vol. 31, pp. 461–464.

[4] A. Cabasson, O. Meste, G. Blain, and S. Bermon, "A New Method for the PP-PR Hysteresis Phenomenon Enhancement under Exercise Conditions," in *Computers In Cardiology*, 2005, vol. 32, pp. 723–726.

[5] C. D. Woody, "Characterization of an Adaptative Filter for the Analysis of Variable Latency Neuroelectric Signals," *Med. & biol. Eng. Comp.*, vol. 5, pp. 539–553, 1967.

[6] L. Xu, P. Stoica, J. Li, S. L. Bressler, X. Shao, and M. Ding, "ASEO: A method for the simultaneous estimation of single-trial event-related potentials and ongoing brain activities," *IEEE Trans. Biomed. Eng.*, to appear.

[7] A. Cabasson, O. Meste, G. Blain, and S. Bermon, "Optimality Statement of the Woody's Method and Improvement," *Research Report*, vol. ISRN I3S/RR-2006-28-FR, 2006, `http://www.i3s.unice.fr/%7Emh/RR/2006/liste-2006.html`.

[8] A. Cabasson and O. Meste, "Time Delay Estimation: A New Insight Into the Woody's Method," *IEEE SIgnal Processing Letters*, to appear.

[9] D. T. Pham, J. Möcks, W. Köhler, and T. Gasser, "Variable latencies of noisy signals: Estimation and testing in brain potential data," *Biometrika*, vol. 74, pp. 525–533, 1987.

[10] P. Jaśkowski and R. Verleger, "Amplitudes and Latencies of Single-Trial ERP's Estimated by a Maximum-Likelihood Method," *IEEE Trans. Biomed. Eng.*, vol. 46 (no.8), pp. 987–993, 1999.

[11] J.S. Sahambi, S.N. Tandon, and R.K.P. Bhatt, "Using Wavelet Transforms for ECG Characterization - An On-line Digital Signal Processing System," *IEEE Eng. Med. Biol. Mag.*, pp. 77–83, 1997.

[12] P. Stoica and J. Li, "On nonexistence of the maximum likelihood estimate on blind multichannel identification," *IEEE Sig. Proc. Mag.*, vol. 21, pp. 99–101, 2005.

[13] A. Cabasson, O. Meste, G. Blain, and S. Bermon, "Estimation, Analysis and Comparison of the PR and RR Intervals under Exercise Conditions and Recovery," in *Computers In Cardiology*, 2006, vol. 33, pp. 497–500.

[14] A. Cabasson, O. Meste, G. Blain, and S. Bermon, "A New Modeling of the Overlapping T wave for the efficient Estimation of the P-R Intervals during Exercise and Recovery," in *IEEE Eng. Med. Biol. Conf.*, 2007, vol. 29, pp. 616–619.

[15] Charles L. Lawson and Richard J. Hanson, *Solving Least Squares Problems*, Prentice Hall, 1974.

[16] S.H. Boutcher, P.F. McLaren, Y. Cotton, and Y. Boutcher, "Stroke Volume Response to Incremental Submaximal Exercise in Aerobically Trained, Active, and Sedentary Men.," *Appl. Physiol. Nutr. Metab.*, vol. 28 (1), pp. 12–26, 2003.

[17] K. Aytemir, N. Maarouf, M.M. Gallagher, Y.G. Yap, J.E.P. Waktare, and M. Malik, "Comparison of Formulae for Heart Rate Correction of QT Interval in Exercise Electrocardiograms.," *Pacing Clin. Electrophysiol.*, vol. 22 (9), 1999.

[18] T. Gasser, W. Köhler, C. Jennen-Steinmetz, and L. Sroka, "The analysis of noisy signals by nonparametric smoothing and differentiation," *IEEE Trans. Biomed. Eng.*, vol. 33, no. 12, 1986.