Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), Larnaca, Cyprus, 11-13 November 2012

Optimal Anisotropic Lead Scaling of Multichannel ECG to Reduce Magnitude Signal Variability

Kai Noponen and Tapio Seppänen
Department of Computer Science and Engineering
University of Oulu
Oulu, Finland

Abstract— **A method for selecting the best functional to nonlinearly project multilead electrocardiogram (ECG) measurements into a specific type of single channel signal is presented. The functional is restricted to a family of time-invariant quadratic functionals parameterized with lead-wise weights. This way, the projected signals are useful in multilead ECG delineation. The method determines the optimal weights in the sense of least beat-to-beat variability, eliminating much of the extra-cardiac influence, which in its turn results in a stable signal. According to the results obtained, the multilead approach is better than using any single lead alone as signal variability is reduced in 80 % of the cases even when using a suboptimal uniform weighting scheme. With the presented optimal lead scaling method, the variability is further reduced in all cases compared to individual leads, and in 92 % of cases, compared to the uniform weighting scheme. The results also show that there is no single set of weights suitable for all situations due to notable variation between the test cases.**

Keywords—vectorcardiogram, noise reduction, artifact elimination, segmentation, delineation.

I. Introduction

In this work, we find the optimal functional belonging to a family of quadratic functionals in order to form a stable representation of the electrical activity of the heart. The representation is meant to be free of extra-cardiac events to the extent this is possible under our beat-to-beat variability measure definition. To motivate our approach, we will first discuss the most obvious application of such a signal, the delineation of the characteristic points of ECG, to show how it relates to the family of functionals at hand.

A number of studies have observed that the approach outlined below to multilead ECG delineation has proven to be effective: the number of inspected leads and necessary decision rules is reduced by first combining the information from all the leads into a new single channel signal, and then suitable single-channel detection methods are used on the derived signal. This way, much of the information in the measured leads can be retained, which can lead to more accurate and stable detection of the characteristic points of the ECG compared to methods based on any individual measured lead only. Two similar examples of this approach are the root-mean-square (RMS) [1, 2] and magnitude [3] signals that differ only in a normalizing scaling constant. Wavelet-based delineators working on these signals have been shown [1, 4] to yield state-of-the art results. Another approach to multichannel delineation is to utilize intelligent lead switching/selection algorithms [2, 4]. Since the morphological variation between individual leads is quite significant, this approach requires an array of lead-specific delineation-methods. In addition, other techniques, such as wavelet based projections [5], have been used to form suitable signals for delineation purposes, and have been shown [6] to be more immune to the effects of respiration in localizing the end of the T-wave compared to single-lead based approaches.

Especially with vectorcardiogram (VCG), the aforementioned magnitude signal is sometimes referred to as the spatial amplitude or spatial magnitude as it can be viewed as the changing length of the electric heart vector (EHV). It is worth to mention that the magnitude is closely related to another surrogate, also used in delineation, namely the spatial velocity that is the time derivative of the EHV. With simple vector calculus, it can be shown that the time-derivative of the magnitude signal corresponds to the projection of the spatial velocity on the radial unit vector, i.e. the radial part of the (spatial) speed of the tip of the EHV. Continuing with the equivalent dipole source point of view, a similar interpretation can be given for multichannel ECG where the individual leads can be seen as projections of the EHV onto lead vectors. Therefore, in this case, the magnitude signal can be seen again as the length of the EHV, but now its components have been scaled to give larger values in directions with more leads placed, and smaller values in directions having fewer leads.

The RMS and the magnitude signal based approaches are similar in that the new derived signal is proportional to the L^2-norm of the measured leads at any time instant. The signals represent the instantaneous totality of the electrical activity of the heart. They are, by definition non-negative, and enable the detection of individual waveforms of the ECG as positive pulses having characteristic properties such as amplitude, duration, and pattern of occurrence. Thus, it can be seen that the RMS and magnitude signals share a major advantage with the signal envelope that has been found useful [7] in single-channel QRS fiducial point detection. That advantage is the continuous dependence on morphological changes. For instance, the magnitude signal will only show a gradual change with respect to minor changes in the electrical axis of the heart, whereas the effects in individual leads can be much more

Research supported by Walter Ahlström Foundation and Tauno Tönning Foundation.

pronounced; especially, near the on- and off-sets of the individual waves [1,4]. Consequently, a major advantage of the RMS/magnitude type approach is being able to operate on a rather fixed set of assumptions about the signal morphology.

An important shortcoming of the RMS and magnitude signals is that they combine information from all the available leads regardless of the validity of that information. While the combining of leads into a new signal can reduce the amount of unwanted signal variability in itself, e.g. when forming linear combinations of multiple closely placed leads with non-correlated measurement noise (spatial filtering), there is still room for improvement. It is known that the individual leads can be affected by body position, respiration, movement, perspiration, inhomogeneity of the thorax, lead placement, and electrode contact quality, to name a few [3,8]. Typically, the leads are also exposed to varying amounts of noise such as electromyographic (EMG) noise, and external electromagnetic interference (EMI). By taking these sources of variation into account, the combining of the leads can be improved on.

In this paper, we address the problem and present an optimal method to assign weights to each channel to provide the most stable magnitude signal in the sense of least beat-to-beat variability. The rest of the paper is organized as follows: the materials and methods are described in Section II, the results are presented in Section III, conclusions are drawn in Section IV, and finally, the work is discussed in Section V.

II. MATERIALS AND METHODS

A. Data

We have made ECG measurements with the help of healthy volunteers who were 19-53 years old. All the participants gave their informed consent, and the Ethical Committee approved the work. Out of the 25 volunteers, 15 were male and 10 female. In the measurements, we have used a Medilog AR12 Digital ECG Recorder (Oxford Instruments, Eynsham, UK) with the RAFE lead system [9]. The device stores the measurements as the three Frank leads (X, Y, Z) [10]. We have used 16 bit amplitude resolution together with 1024 Hz (X), and 512 Hz (Y, Z) sampling rates, but the X channel has been resampled to use the same rate of 512 Hz afterwards to match the sampling rate of the other leads.

The measurements have been made according to a controlled protocol [3] that includes different body positions and respiration depths. For each subject, the protocol has lasted 18 minutes during which an instructor has been giving orders to change the body position, and the respiration depth at certain time instants guided by a timer. Afterwards, we have segmented the recordings to extract individual beats using our implementation of the shift invariant wavelet transform based delineation algorithms described in [11, 12]. Finally, we have removed the linear trend from each of the Frank leads beat-by-beat to reduce baseline wander since it causes unwanted changes in the magnitude and RMS signals.

B. Mean Shape of the Magnitude Signal

Let us consider a multichannel ECG measurement with an arbitrary lead set. First, we collect all the signals in the M measured leads at the time instant t into the vector $\mathbf{x}(t) = [x_1(t) \ldots x_M(t)]^{\mathrm{T}}$, and denote the corresponding set of real-valued weights by the vector $\boldsymbol{\alpha} = [\alpha_1 \ldots \alpha_M]^{\mathrm{T}}$. Using this notation, we may now define the anisotropically scaled magnitude signal as

$$r_{\boldsymbol{\alpha}}(t) = \sqrt{\mathbf{x}^{\mathrm{T}}(t)\mathbf{Q}_{\boldsymbol{\alpha}}\mathbf{x}(t)}, \tag{1}$$

where $\mathbf{Q}_{\boldsymbol{\alpha}} = \mathrm{diag}(\boldsymbol{\alpha})$ is the diagonal M-by-M matrix of scaling weights related to the quadratic form within the square-root. This operation defines a non-linear but time-invariant functional that maps the actual measured multilead signals into a weighted magnitude type single channel signal. When all the weights are one, $\mathbf{Q}_{\boldsymbol{\alpha}}$ is the identity matrix $\mathbf{I}$, and (1) becomes the plain magnitude signal. Accordingly, the RMS signal is obtained by setting $\boldsymbol{\alpha} = [1/M \ldots 1/M]^{\mathrm{T}}$. Moreover, using just one nonzero weight makes (1) to act as lead selection from absolute valued leads.

Let us next consider that there are N beats in total within a given data record, and that for each beat i we collect the samples of the magnitude signal (1) during the beat $t \in [a_i, b_i]$ into a L-dimensional vector $\mathbf{y}_{\boldsymbol{\alpha},i}$ with zero padding when the beat is shorter than the predetermined duration of L samples. The beats can be prealigned according to their detected R-peak position keeping it at a fixed location in the vector, but this is not necessary. We define the mean shape of the magnitude signal as

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\mu}} \frac{1}{N}\sum_{i=1}^{N} d^2(\mathbf{y}_{\boldsymbol{\alpha},i}, \boldsymbol{\mu}), \tag{2}$$

where the distance between the ith beat of the magnitude signal and the mean shape candidate $\boldsymbol{\mu}$ is

$$d^2(\mathbf{y}_{\boldsymbol{\alpha},i}, \boldsymbol{\mu}) = \min_{\tau} \left\| \tau \circ \mathbf{y}_{\boldsymbol{\alpha},i} - \boldsymbol{\mu} \right\| \tag{3}$$

when they are time-aligned as close to each other as possible. Here, τ denotes the cyclic interpolation operator that operates on the magnitude signal sifting it to match the mean shape candidate.

In essence, (2) and (3) together define the mean shape of the magnitude signal as a representation that has the shortest average squared distance to all the observed beats when they are aligned in time to each other as closely as possible with respect to the Euclidean distance. Given a fixed weight vector $\boldsymbol{\alpha}$, the mean shape can be estimated using the following algorithm:

- First, initialize the mean shape candidate vector $\boldsymbol{\mu}$ as the average of magnitude signal presentations $\mathbf{y}_{\boldsymbol{\alpha},i}$ for all the beats. Instead of the average, any single $\mathbf{y}_{\boldsymbol{\alpha},i}$ could be used as well.
- Second, align all the magnitude signals $\mathbf{y}_{\boldsymbol{\alpha},i}$ in time to the current mean shape candidate $\boldsymbol{\mu}$. The amount of time sifting can be found via cyclic cross-correlation, and subsample accuracy can be obtained using interpolation.

- Third, compute a new mean shape candidate as the average of all time-aligned magnitude signals obtained in the previous step.
- Finally, if the mean shape candidate changed significantly in the previous step, return to the second step. Otherwise, quit as the algorithm has found a suitable mean shape estimate.

Typically, the algorithm takes only a few iterations for convergence. For more details about the underlying theory, we refer the reader to a more detailed exposition [3] of these concepts in a more general setting.

C. Choosing optimal weights

Next, we show how to obtain the best set of weights. Consider the effects of the weights on (1). If, for example, there are multiple leads with redundant information, but some with more noise or other unwanted disturbances, lowering the corresponding weights towards zero will reduce the effects on (1), while retaining information via leads with more substantial weights. Now, given the previous definitions, it is straight forward to measure how much the magnitude signal representations $\mathbf{y}_{\boldsymbol{\alpha},i}$ of each beat differ from the mean shape estimate we have obtained. This, in turn, leads us to define a measure for the relative amount of beat-to-beat variation in the signal. By denoting the average AC-power of the magnitude signal during the ith beat with $\mathrm{var}(\mathbf{y}_{\boldsymbol{\alpha},i})$, we define the objective function

$$J(\boldsymbol{\alpha}) = \frac{\sum_{i=1}^{N} d^2\left(\mathbf{y}_{\boldsymbol{\alpha},i}, \hat{\boldsymbol{\mu}}_{\boldsymbol{\alpha}}\right)}{\sum_{i=1}^{N} \mathrm{var}(\mathbf{y}_{\boldsymbol{\alpha},i})} \tag{4}$$

that can be used to find the optimal weight vector $\boldsymbol{\alpha}$. It should be noted that for each choice of $\boldsymbol{\alpha}$, the associated mean shape estimate changes as well.

In (4), the numerator describes the cumulative amount of beat-to-beat variation in the signal measured by the discrepancy between the mean shape and the observed beats. Naturally, noise and disturbances will increase this variation. The denominator describes the "total size" of the signal and makes (4) scale-invariant. This way, (4) favors – in relative terms – a stable signal with a smaller amplitude range and little beat-to-beat variation over a larger amplitude signal with more beat-to-beat variation. It is evident that inherent temporal variability in the source, e.g. due to chaotic behavior, cannot be overcome completely and will remain in the derived signal, even after weight changes suggesting a lower bound for (4). However, some source change dependent effects can be compensated for to some extent. These include situations such as a change in the orientation of the electrical axis due to respiration causing typical magnitude signal level "pumping". This can be overcome by setting suitable weights to correct differences in tissue conductivities and electrical contact qualities of the electrodes, among other things. The reader is also referred to [13] for a discussion on the properties of the objective function.

To obtain the weights that give as small as possible beat-to-beat variation, but at the same time, as "strong" as possible magnitude signal, we minimize (4) with respect to $\boldsymbol{\alpha}$ such that $\|\boldsymbol{\alpha}\| = 1$, and $0 \le \alpha_i \le 1$ for all $i = 1,\ldots, M$. These constraints represent the feasible region for the problem. It should be noted that $\boldsymbol{\alpha}$ can be constrained to be a unit vector because the objective function (4) is scale-invariant. The minimization problem is well-defined for signals that are not constant valued. But since the approach requires prior segmentation of beats, those signals are automatically excluded. Strictly speaking, however, if there are multiple linearly dependent leads, the objective function (4) obtains the same value for all the linear combinations of those leads. In practice, this situation is unlikely to occur with real measurements as even the slightest differences in noise levels usually make a lead more suitable than the other leads, i.e. results in a lower value for (4).

D. Problem dimensionality reduction for VCG recordings

The weight vector $\boldsymbol{\alpha}$ has M components, but the restriction to a unit vector reduces the degrees of freedom by one. In the specific case of the three-dimensional Frank lead system used in our data set, $M = 3$, and we may parameterize the unit weight vector $\boldsymbol{\alpha}$ in a spherical coordinate system as $\boldsymbol{\alpha}(\theta, \phi)$ with just two independent angles: the azimuth $\theta \in [0, \pi/2]$, and the elevation $\phi \in [0, \pi/2]$. This allows for easy visualization and effective solving of the problem. The angle intervals correspond to the constraints of the original minimization problem as we are using a right-handed coordinate frame in which the Frank lead X corresponds to (0, 0), Y to $(\pi/2, 0)$, and Z to $(0, \pi/2)$ in the spherical coordinate system. Moreover, the special case of equal weights, i.e. $\alpha_i = 1/\sqrt{3}$, $i = 1, 2, 3$, corresponds to $(\pi/4, \operatorname{arccot}\sqrt{2})$ after the spherical coordinate transformation.

We can finally formulate our specific minimization problem as

$$\left(\hat{\theta}, \hat{\phi}\right) = \underset{0 \le \phi, \theta \le \pi/2}{\arg\min}\, J\left(\boldsymbol{\alpha}(\theta, \phi)\right). \tag{5}$$

To solve (5), we use an active-set algorithm [14] that utilizes a sequential quadratic programming method, and makes quasi-Newton approximations to the Hessian of the Lagrangian.

III. RESULTS

Fig. 1 shows the objective function (4) using a set of equipotential curves for each of the 25 volunteers in the spherical coordinates enumerated in the conventional first left-to-right and then top-down fashion. The curves are interpreted like the contour lines denoting the elevation in a map. The cases of specific interest (cf. Section II D) in which the weights either coincide with the single leads, or in the special case of equal weights, are marked in each plot. In addition, the locations of the minima (5) are depicted. From the separation, orientation, and shape of the contours, we can see that there is significant subject-to-subject or recording-to-recording variation in the objective functions. We can also note that the location of the minimum is relatively near to the equal weight distribution for some subjects, but much further for the others.

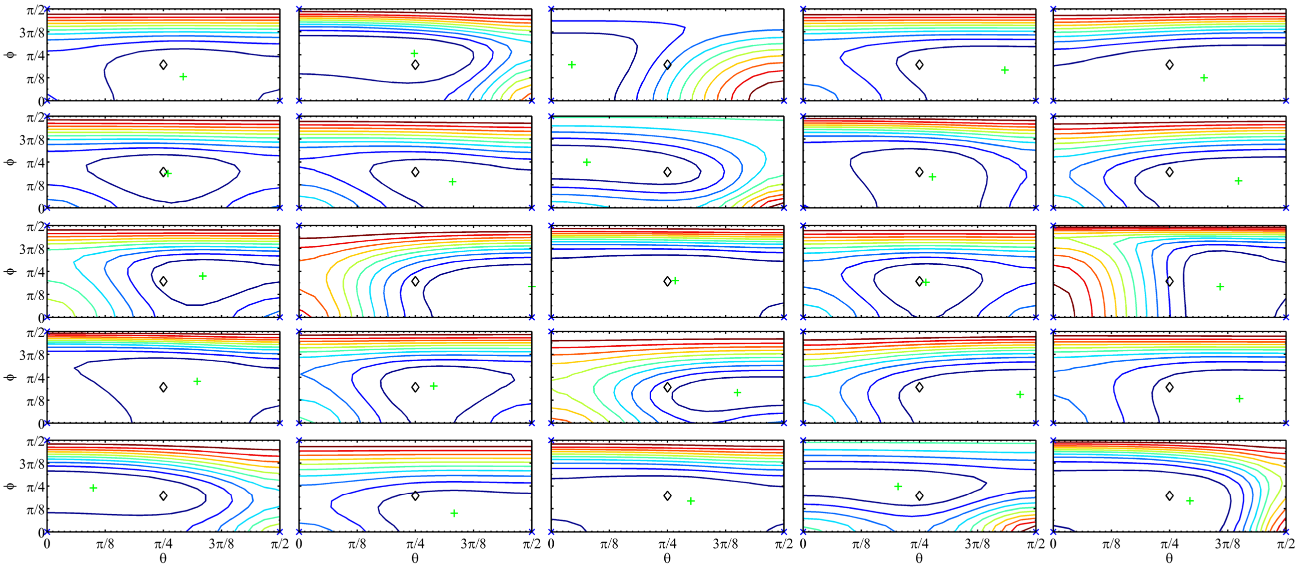

Figure 1. A contour plot of the objective function (4) for each subject with respect to the azimuth (θ) and the elevation (ϕ) in the spherical coordinate system. The green plus-sign (+) signifies the location of the minimum (5), the black diamond (◊) the case of equal weights, and the blue crosses (x) on the axes the cases of individual leads in which the weight for the lead in question is 1, and the other weights are 0.

In Table I, we enumerate the relative value of the objective function

$$k(\theta,\phi)=\frac{J(\boldsymbol{\alpha}(\theta,\phi))}{J(\boldsymbol{\alpha}(\hat{\theta},\hat{\phi}))} \qquad (6)$$

in percent units for the aforementioned weight combinations. The numbers are rounded to the nearest integer. On average, the non-weighted (equi-weighted) magnitude signal gives a 14 percent units higher value for the objective function (4) than the optimal weight combination (5), but the results vary subject-to-subject in the range from 0 to 149 percent units (cf. the last column in Table I). In the case of individual leads, i.e. a nonzero weight with one of the leads only, the increase in unwanted variation ranges from 7 percent units (subject 20) to 633 percent units (subject 5) as measured by (6). It is worth to note that for five subjects, an individual lead would actually offer better performance than the non-weighted approach, but for the majority (80 % out of 25), the non-weighted approach gives better results than any individual lead. However, the optimal weights of (5) always give the lowest variability as (6) yields 100 % by design. Only in two cases (subjects 6 and 14) is this minimum reached using the non-weighted approach.

TABLE I. RELATIVE VALUES OF THE OBJECTIVE FUNTION

Subject #	$k(\theta, \phi) \times 100$ %			
	(0, 0)	(π/2, 0)	(0, π/2)	(π/4, arccot √2)
1	161	135	426	106[a]
2	128	242	282	104
3	123	677	244	249
4	200	117[b]	424	117[b]
5	157	111	733	109
6	150	145	260	100
7	185	118	292	105
8	165	257	173	104
9	154	167	391	101
10	223	113	382	112
11	189	144	247	109
12	241	108	266	122
13	128	144	499	101
14	181	161	296	100
15	296	121	305	134
16	153	141	438	107
17	261	188	439	104
18	142	116	155	107
19	226	110	385	117
20	223	107	496	124
21	140	250	337	105
22	139	110	220	108
23	143	142	507	102
24	258	424	255	108
25	112	221	227	101
Average	179,1	182,8	347,2	114,2

a. The underlining denotes the minimum on each row.
b. The values are the same within the rounding precision.

IV. CONCLUSION

We have presented a novel approach to anisotropic ECG lead scaling with the aim of decreasing the beat-to-beat variability of the magnitude signal, while keeping as high as possible variance at the same time. Overall, the results show that even the sub-optimal uniform weights give the combined signal (1) a smaller beat-to-beat variation compared to the individual leads. What is more, the presented optimal weight selection can further reduce the variation in almost all the cases, but no fixed set of weights suits all the situations due to notable variation between the test cases.

V. DISCUSSION

This paper extends our previous work (cf. [13]) to the class of quadratic functionals that are known to be important in ECG delineation, for instance. In effect, the method assigns less weight to leads with a higher amount of beat-to-beat variability regardless of its origin; be it noise or other unwanted phenomena. Simultaneously, the method combines information from several leads together in such a way that the changes in a lead can be counteracted by the changes in other leads, which in its turn results in a more stable magnitude signal.

Compared to the traditional time-invariant or adaptive linear-filtering based preprocessing, a major advantage of the presented method is that it can also reduce in-band noise such as EMG by reducing the coefficients of the leads exhibiting noisy behavior. In this regard, the method is similar to preprocessing the signal, e.g. with a spatial source consistency filter [15], or denoising it with independent or principal component analysis based methods [16].

Although the method requires at least a rough segmentation of beats to operate in the first place, this is not a major drawback because R-peaks have a high signal-to-noise ratio enabling them to be detected, even when the other characteristic features are buried under a significant amount of noise [3]. After the rough segmentation around the R-peaks, the optimal weights can then be determined to yield a more stable magnitude signal overall. This may improve QT-interval measurements, for example. It should also be noted that depending on the specific aims, the focus of the beat-to-beat variability reduction can be directed on any particular portion of the beat, such as the T-wave, instead of the full P-QRS-T cycle used in this work.

An online version of the method could be used, for example, with wearable sensors incorporating a number of electrodes or even a large matrix of electrodes with redundant information. The weights could then be adapted beat-by-beat to yield a stable magnitude signal with a low beat-to-beat variability. This approach would be especially advantageous in unsupervised measurement situations where one cannot guarantee the signal quality in all the electrode positions or in situations where one does not want to halt the measurement for any minor issue. These preliminary results also open up new possibilities to expand the methodology to an even broader family of functionals by requiring the matrix Q of (1) only to be positive semidefinite instead of purely diagonal.

REFERENCES

[1] M. L. Soria, J. P. Martínez, and P. Laguna, "A Multilead Wavelet-based ECG Delineator based on the RMS Signal," Comp. Cardiol. 2006, vol. 33, pp. 153−156.

[2] F. Rincon, N. Boichat, V. Barbero, N. Khaled, and D. Atienza, "Multi-lead wavelet-based ECG delineation on a wearable embedded sensor platform," Comp. Cardiol. 2009, vol. 36, pp. 289−292.

[3] K. Noponen, J. Kortelainen, and T. Seppänen, "Invariant trajectory classification of dynamical systems with a case study on ECG," Pattern Recog., vol. 42, pp. 1832–1844, September 2009.

[4] F. Rincón, J. Recas, N. Khaled, and D. Atienza, "Development and Evaluation of Multilead Wavelet-Based ECG Delineation Algorithms for Embedded Wireless Sensor Nodes," IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 6, pp. 854–863, November 2011.

[5] R. Almeida, J. P. Martinez, A. P. Rocha, and P. Laguna, "Multilead ECG Delineation Using Spatially Projected Leads From Wavelet Transform Loops," IEEE Trans. Biomed. Eng., vol. 56, no. 8, pp.1996–2005, August 2009.

[6] M. Noriega, J. P. Martínez, P. Laguna, D. Romero, R. Bailón, and R. Almeida, "Respiration effect on single and multi lead ECG delineation strategies," Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp. 3575–3578.

[7] M. Nygårds, and L. Sörnmo, "Delineation of the QRS complex using the envelope of the e.c.g.," Med. Biol. Eng. Comput., vol. 21, no. 5, pp. 538–547, September 1983.

[8] J. Malmivuo, and R. Plonsey, Bioelectromagnetism — Principles and Applications of Bioelectric and Biomagnetic Fields. New York: Oxford University Press, 1995.

[9] G. E. Dower, and J. A. Osborne, "A clinical comparison of three VCG lead systems using resistance-combining networks." Am. Heart J., vol. 55, no. 4, pp. 523–534, April 1958.

[10] E. Frank, "An accurate, clinically practical system for spatial vectorcardiography," Circulation, vol. 13, no. 5, pp. 737–749, May 1956.

[11] S. Kadambe, R. Murray, and G. F. Boudreaux-Bartels, "Wavelet transform-based QRS complex detector," IEEE Trans. Biomed. Eng., vol. 46, no.7, pp. 838–848, July 1999.

[12] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: evaluation on standard databases," IEEE Trans. Biomed. Eng., vol. 51, no. 4, pp. 570–581, April 2004.

[13] K. Noponen, and T. Seppänen, "Calculating Optimal Virtual Lead from Multichannel ECG by Minimizing Morphological Beat-to-Beat Variability," Comp. Cardiol. 2009, vol. 36, pp. 717–720.

[14] M. J. D. Powell, "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations," in Numerical Analysis, Lecture Notes in Mathematics, vol. 630, G. A. Watson Ed. Springer Verlag, 1978.

[15] D. W. Mortara, "Source consistency filtering-a new tool for ECG noise reduction," Comp. Cardiol. 1991, vol. 19, pp. 125–128.

[16] I. Romero, "PCA and ICA applied to noise reduction in multi-lead ECG," Comp. Cardiol. 2011, vol. 38, pp. 613–616.