

# New Algorithms for Continuous Analysis of Long Term ECG Recordings Using Symplectic Geometry and Fuzzy Pattern Recognition

M Arzi

Inserm ERM107 - MTIC, Cardiology Hospital of Lyon, Bron, France

## Abstract

*In this paper, we describe how to analyze continuously ECG of any duration, and we describe a group of new and original algorithms which can do this analysis almost on-line. These algorithms use symplectic geometric calculus as well as an original fuzzy clustering method for localization of various waves and calculation of different parameters. The programs calculate the isoelectric curve of the ECG, then localize all ECG complexes and delineate the P, QRS and T waves within each complex. Then they calculate various parameters such as spatial area of the QRS wave, QRS-T angle etc. Finally, according to predefined criterions, every complex of ECG is classified into one of several predefined classes and an average prototype of every class is calculated and recorded. Spurious parts and outliers in continuous signal or in calculated parameters are automatically detected. The calculation of the isoelectric curve is performed by using an original fuzzy pattern recognition method operating on a detection function.*

## 1. Introduction

Continuous analysis of long duration ECG is useful in various investigations. For example exercise stress test helps to evaluate dynamic characteristics of heart performance such as depolarization dynamics as a response to heart rate variability. With the development of wireless transmission and high capacity compact wearable ECG recording systems, it is possible to record ECG of people during activity for relatively long periods of time. In some investigations, it is necessary to analyze ECG continuously and gaps between long periods of heart activity is not permitted. For example in stress test, it is necessary to analyze the whole ECG from before the beginning of exercise until sufficiently after the end of effort protocol when the subject has recovered his rest heart rate. This is necessary for the evaluation of dynamic characteristics of the heart during acceleration and deceleration. In some other investigations it is also necessary to analyze ECG in synchronization to external events. In these cases also an

analysis of the entire ECG during and between these events is necessary.

One of the major problems in continuous analysis of ECG is the fact that long term curves recorded from surface electrodes are subject to various, non predictable events. These events include variations in electrode-skin contact, spurious additive signal, and variations of body position. Sometimes, some of the electrodes produce non informative signals in which electrical activity of heart may completely be hidden. To overcome these problems, our analysis, to be described in this paper, is based on global information contained in the whole set of ECG curves obtained from all leads. In this way, if some electrodes give non interpretable signals during some periods of time, information from other leads help correct errors and fill in gaps. Consequently more leads means more robustness. For statistical independence of leads, orthogonal or near orthogonal ones are used. A minimum of two leads is necessary for our programs to give satisfactory results. But a standard three-lead ECG or eight-lead system will work even better. As will be explained later, global analysis is also done in time domain. This means that local errors are corrected by information obtained from appropriate temporal context. In the following we will describe the theoretical basis of our method and illustrate some results of ECG analysis during stress test.

## 2. Methods

The following sequence of data processing is undertaken:

1. A detection function based on global information contained in the whole ECG data from all leads is computed.
2. The isoelectric pattern of ECG curves is constructed.
3. All QRS complexes are isolated.
4. A classification is done on a complete cycle of heart activity (from the beginning of a QRS to the beginning of the next one). To do this, all cycles are isolated and classified in a number of categories based on some criterions.
5. P and T waves are localized.
6. According to these results, RR interval as well as time duration of other significant events such as QRS, P, and T

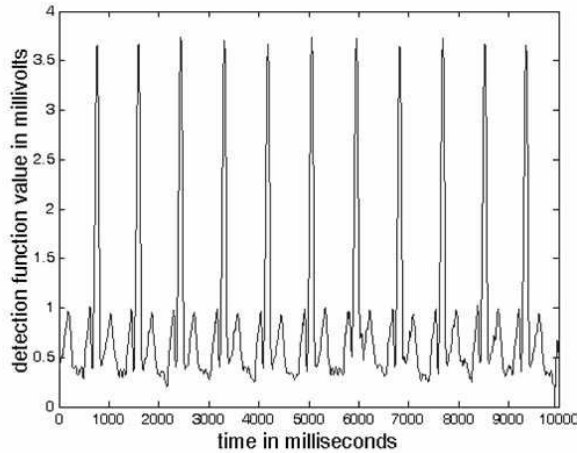


Figure 1. An example of detection function representing overall electric activity of heart. The level of this function is extremely high during QRS waves. So in order to permit visibility of heart electric activity during P and T waves, the 8th root of the function is plotted in this figure.

waves are calculated for the whole ECG.

7. A global pattern recognition on the results of analysis is done and so data outliers are discarded.

In the following, a brief description of each of these processes is given.

## 2.1. Detection function

To construct the detection function, all ECG curves together with their first and second derivatives are used. To understand the theoretical basis of this calculus, the reader should refer to advanced calculus texts on differential forms and multi-vector algebra (Grassmann algebra). For example [1] and [2] are recommended. Here we sketch an outline of this calculus: Consider a sample of ECG at some time instant  $t$ , obtained from all leads. This represents, for each time instant, an  $n$ -dimensional vector, where  $n$  is the number of leads. The same is also true for the first and second derivatives of the ECG curves. We shall denote the ECG vector by  $E(t)$ , its first and second derivative, by  $EP(t)$  and  $EPP(t)$  respectively. Their components are denoted by  $E_i(t)$ ,  $EP_i(t)$  and  $EPP_i(t)$  respectively. Where  $(i=1, \dots, n)$  denotes lead number and  $t$  denotes time. From these three vectors the following multi-vectors are constructed by exterior (outer) product:

- (1)  $E(t) \wedge EP(t)$ ,
- (2)  $E(t) \wedge EPP(t)$ ,
- (3)  $EP(t) \wedge EPP(t)$  and
- (4)  $E(t) \wedge EP(t) \wedge EPP(t)$ .

The wedge represents outer product of vectors. The first three are bi-vectors. Their dimension is  $C(n,2)$ , the number of ways of choosing two elements from a set of  $n$  elements ( $n$  should be at least 2). For example for  $n=8$  (8-leads

system), they are 28 dimensional, and for 3-leads system ( $n=3$ ) they are 3 dimensional. Bi-vectors represent simplicial contents of a surface element which, for the case of (1), represents the outer product of ECG vector by the vector representing its derivative. Similar thing is true for cases (2) and (3). Case (4) is a 3-vector whose number of components is  $C(n,3)$ . In this case the minimum number of leads needed is 3. The simplicial content of this vector is a volume element representing the outer product of the ECG value by its first and second derivative. Each of these multi-vectors has its advantages and applications which will not be detailed here. It will only be noted that each component of these multi-vectors has a specific signification. For every data point, each component represents the volume element (2- or 3-dimensional) consisting of the simplex formed by the value of individual leads at time  $t$ , simultaneously with the velocity and/or acceleration of each lead. In this way, all the dynamics of the signal obtained from various leads are encoded in the components of these multi-vectors. For signals whose analysis are given here, the bi-vector (3), e.g.  $EP(t) \wedge EPP(t)$ , has been used. The modulus of this bi-vector:  $|EP(t) \wedge EPP(t)|$  gives our detection function. An example of detection function is given in figure 1.

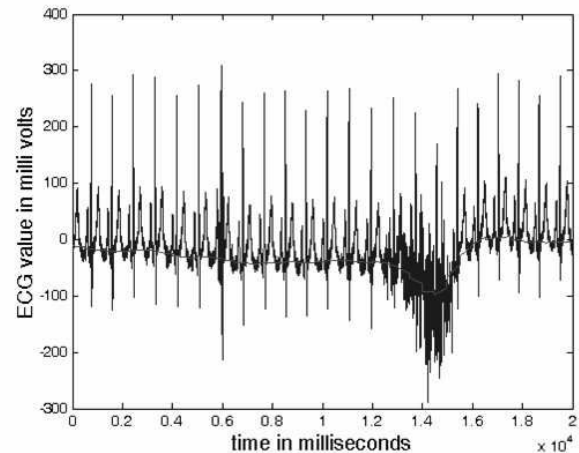


Figure 2. An example of isoelectric curve localization of an ECG curve. Superimposed on the ECG curve is the isoelectric curve. As can be seen, even when the data are highly corrupted, the isoelectric curve is robustly localized.

## 2.2. Isoelectric curve calculation

Considering the fact that long duration ECG recordings almost always contain non predictable variations and many artifacts, it is necessary to do a pattern recognition on the detection signal for every analysis. For this we have devel-

oped a fuzzy auto-adaptive pattern recognition algorithm based on a method developed earlier by the author and described in [3], [4] and [5]. This method has been considerably improved for this study. What this algorithm does is essentially the following: The distribution of data points in each level is calculated. This is done in a moving window of a suitable duration to include a little more than one ECG cycle. Data in different levels are related to each other through a coherence function which is constructed in an iterative process. The coherence function represents a fuzzy membership function which measures the degree of coherence of each point with some set of other points of data. The coherence being defined through their degree of fit with a piecewise polynomial function. This latter function defines the global shape of the parts of data which we seek and that we know by a priori information they should have a certain pattern which can be defined very approximately. In the case of isoelectric curve, the global dominant form should be a piecewise linear or quadratic polynomial. The algorithm begins by a roughly defined pattern, then a relaxation process between this pattern and the coherence function (initialized to a constant one) refines them simultaneously. After convergence of the algorithm, the final global dominant form gives the isoelectric curve. Details of this algorithm are too long to be described here. The interested reader can read references given above. An example of isoelectric curve localization is given in figure 2.

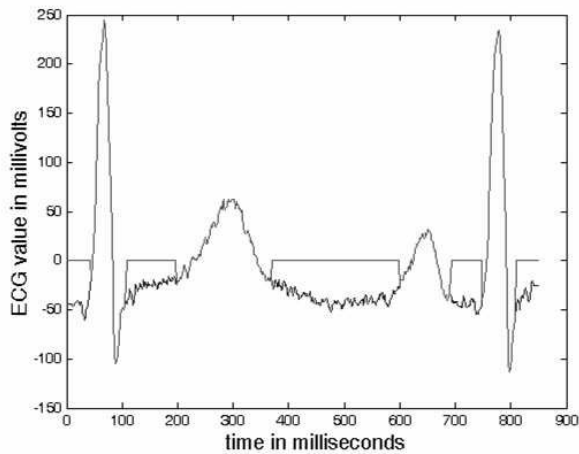


Figure 3. Extraction of QRS, T and P waves. In this figure a curve representing QRS, T and P waves is superimposed on the ECG curve. The value of the former curve is equal to the ECG curve during QRS, T and P wave and its value is zero elsewhere. This analysis is done on every ECG cycle.

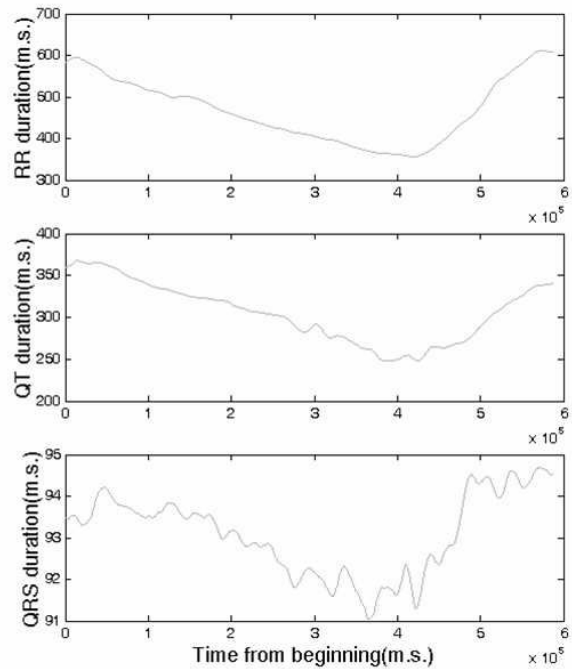


Figure 4. Some partial results of ECG analysis. Here stress test is done during approximately 10 minutes. At start, the subject is at rest. Then the bicycle charge is gradually raised until a certain level and then dropped. ECG recording is continued some time after, until the subject has recovered his rest heart rate. Experiment was performed in Cardiology Hospital of Lyon.

### 2.3. Event extraction

Based on the isoelectric curve and the detection function, significant events are extracted. First, QRS complexes are delineated and a set of all individual QRS signals for further analysis is constructed. Then P and T waves are similarly extracted (see figure 3). Each cycle of heart activity (a cycle means ECG curve from the beginning of one QRS to the beginning of the next one) is classified into a number of categories according to its pattern. Template patterns are calculated from the ECG and are adapted to each patient. Detection of QRS waves is easier and more reliable than the detection for P and T waves and is done with very high degree of reliability. This is because heart electric activity is considerably higher during QRS signal. For the detection of P and T waves information from QRS location is used and a non linear peak detection algorithm is used. Localization of these events permits to compute different well known parameters such as RR intervals, ST segment, QT interval duration and so on.

### 3. Results

Various signals have been analyzed successfully by these programs. In this paper some examples are given which concern a simple stress test. In this test the subject rides on a stationary laboratory bicycle. Effort intensity is augmented gradually from zero to some level then stopped. In figure 4, acceleration and deceleration of heart rate is illustrated in the top curve. The middle curve shows how QT intervals follow heart rate variations and the bottom one that of QRS durations. Dynamics of this QT follow up allow to quantify the depolarization-repolarization adaptability of heart.

### 4. Discussion and conclusions

Necessity of continuous, long term ECG analysis is becoming more and more evident. New technology, high capacity holter recordings or other newly developed systems permit a variety of novel investigations which were not possible before. Home care and permanent survey of patients becomes possible now. A robust analysis of long term ECG necessitates sophisticated algorithms. Here we propose new, original algorithms that perform this analysis with high degree of success.

Various noisy signals which were corrupted by many artifacts have been successfully analyzed by this method. Some preliminary results are given which show that reliable analysis is possible and partially illustrate what can be obtained from such analysis.

### References

- [1] Edelen D. Applied exterior calculus. Dover Publications Inc, Mineola, NY, 11501, 2005.
- [2] Hestenes D, Sobczyk G. Clifford algebra to geometric calculus. Kluwer Academic Publishers, 101 Philip Drive, Norwell, MA 02061, USA, 1984.
- [3] Arzi M. Traitement automatique des signaux vestibulo-oculaires et optocintique. Ph.D. thesis, 1987.
- [4] Arzi M, Magnin M. fuzzy set theoretical approach to automatic analysis of nystagmic eye movements. IEEE Trans Biomed Eng September 1989;36(9):954–63.
- [5] Arzi M, Laissard G, Prablanc C. A self calibration method using a soft clustering procedure for eye movement recordings. Neurosci Methods March 1998;80(1):1–11.

Address for correspondence:

Mohammad Arzi  
ERM 107, 28 avenue du doyen Ipine, 69500, Bron, France.  
E-mail address : arzi@lyon.inserm.fr