

Degree in Mathematics

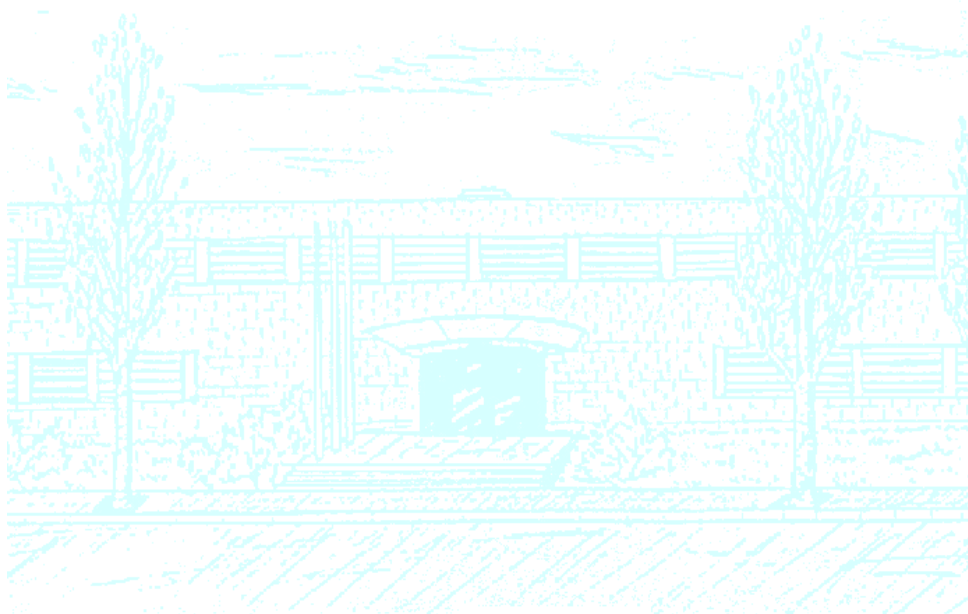
Title: PLAR: Concept and Applications. A new approach to Electrocardiography

Author: Martí Fernández-Real Girona

Advisor: Sonia Fernandez Mendez

Department: Enginyeria Civil i Ambiental

Academic year: 2017/2018



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya

Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

PLAR: Concept and Applications
A new approach to
electrocardiography

Martí Fernández-Real Girona

Supervised by Kumaran Kolandaivelu and tutored by Sonia Fernández

June, 2018

I would like to thank everyone that has been involved professionally or personally in this project.

First, thanks to Prof. Kumaran Kolandaivelu for welcoming me in his research group and giving me the opportunity to be involved in a project of this caliber. Thanks for sharing your passion for science.

Thanks specially to CFIS for giving me the support and help through the mobility scholarship, without which this international experience would not have been possible.

Thank you also to my father and brother, for showing interest and support receiving nothing in return.

To finish, special thanks to Adrià for sharing part of his project with such enthusiasm, and Sonia Fernández for tutoring me part of it.

Contents

1	Introduction	1
2	Objectives	2
3	Previous Concepts	3
3.1	Electrocardiography	4
3.2	Vectorcardiogram	6
3.3	Attractor	8
4	PLAR : Our New Space	11
4.1	First Component of PLAR	12
4.2	Second Component of PLAR	13
4.3	Third Component of PLAR	14
4.4	Obsevation	15
5	Data Acquisition for First PLAR Development	19
6	Signal Filtering and Smoothing	22
7	ST Segment Variations	26
7.1	Methodology	26
7.2	Results and Observations	28
8	Progressive Disease Tracking	31
8.1	Methodology	32
8.2	TDR: Today-Related Coefficient	36
8.3	Results and Observations	38
9	Future Work	42
10	Conclusions	45

References	46
List of Figures	47
A APPENDIX A	50
A.1 Orthogonalization matrix	50
B APPENDIX B (Matlab code)	51

1. Introduction

The characterization of heart with a view to distinguish abnormal from normal behaviour is an interesting topic in clinical sciences. Currently there exist several medical tests for checking the state of the heart at a given moment. However there are 2 limitations when using these tests that are basically the cost and the level of expertise needed to perform and analyse them. The cheapest and most common one is the Electrocardiogram (ECG), which has been used for many years, but this recording of electrical outputs of the heart fails to gather all physical and physiological information and more tests are required for a proper diagnosis.

Apart from that, in general the medical community is cursed with immobilism. Outdated methods are still used because of their relatively good performance instead of adopting new ones that may perform better, based on the technology advancements being made in the recent years. The vast increment in computational power available for a single individual and new visualization techniques are the key for new treatments and patient follow-ups, and for improving healthcare in general.

Current techniques treat all the patients equally, there is not patient specificity. It is clear that, in medical terms, two patients will never behave the same and there is now the possibility of recording tons of incredibly complex data for each patient, while processing it according to the needs.

Also, as many data has been recorded over time, applying new algorithms as deep learning techniques may be able to, for each patient, perform an anticipated diagnosis.

With evolving electronical and computing devices it is possible to track continuously the individual and continuously giving automatic diagnosis. That would change the medical care for good. For many years medical community has been treating with unhealthy individual, when the patient feels ill goes to the hospital. Nevertheless the concept of anticipated continuous diagnosis would change that drastically, triaging better the potential patients and prioritizing the ones at most risk in the queue of medical care.

In conclusion, new algorithms and visualization techniques are the first step to improve medical care using the evolving technology. However, these new implementations would not be useful without human supervision. This is why new tools are welcomed when it comes to improving diagnosis, but should never be totally automatized.

2. Objectives

The aim of this thesis is to provide a line of investigation consisting of a new visualization tool which helps solving the problems mentioned in the current medical community. Specifically, we pretend to develop a tool to better understand the underlying mechanisms of heart functioning, so the tool will be applied to cardiology.

We will focus on the information coming from an electrocardiogram and try to extract the relevant information that it is containing. First we are introducing certain concepts that will be useful to understand how we pretend to process the data to provide patient specificity in our visualization tool. Then, after defining our space, two types

of heart irregularities are being addressed: sudden irregularities and progressive ones. The results represent an example of how our tool can be applied to identify sudden and track progressive anomalies over time.

The process of defining our visualization space and testing it is developed using the power of applied mathematics. We will be using the Matlab mathematical calculation tool for all the process, like signal acquisition, signal processing, plotting, results analysis and final visualization.

Understanding the electrocardiogram as electrical output signals of the heart, our algorithm processes that information and reshapes it to obtain a personalized three dimensional representation. At some stage of this thesis we will test how a variant of our tool can relate to physical and physiological properties of the heart only through the processing of this electrical information.

Note how we do not pretend to develop and automatic diagnose algorithm. We will provide with a new visualization tool trying to capture and optimize the information from electrocardiograms that in the right hands could lead to better diagnosis.

This thesis represents the first step of a new approach to electrocardiography, and we pretend to make a tiny contribution to medicine.

3. Previous Concepts

In this section we pretend to introduce and explain the basic concepts that are being used in the development of our new visualisation tool.

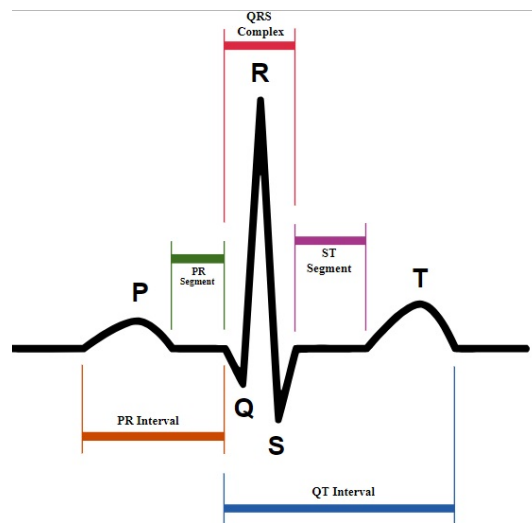


Figure 1: Scheme of the main part of the standard sinus rhythm

3.1 Electrocardiography

One of the most commonly performed medical test is Electrocardiogram (ECG). This is the process of recording electrical activity of the heart through electrodes placed in different parts of the human body. These electrodes can detect the tiny changes in electrical activity travelling through the cells and muscles of the heart, the movement of which are controlled by the repolarizing and depolarizing patterns at every heartbeat. That is, the physical movement of the heart is controlled by circulation of electrical pulses (order of few mV), which are captured in standard 12-lead ECGs (an example of standard ECG record is shown on top of Figure 6 in a later Section). Each lead represents a measurement of electric potential of the heart from different angles in space, during a fixed time period (usually 10 seconds).

The electrical signal of a heartbeat can be divided into different parts. Each

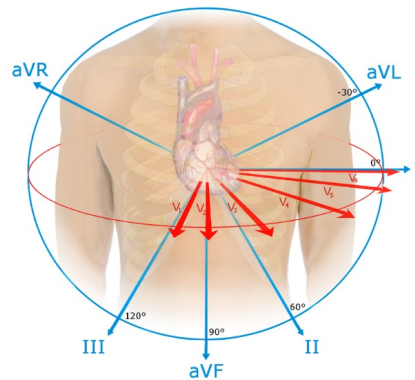


Figure 2: Directions of the 12 leads of standard ECG

part can be visualized in a different way from each lead. For the purpose of this work, it is only necessary to understand that different heart diseases can be diagnosed thanks to different common alterations in different parts of the signal of the heartbeat in different leads, and depending on the current state of the patient.

Though, the main parts of the heartbeat can be visualized in the Figure 1 and are the P-wave, the QRS-complex, the ST-segment and the T-wave, that represent electrical pulses in different regions of the heart. The peak with the greatest amplitude in the standard lead, that is the R-peak, represents for instance the electrical signal that makes the ventricles contract and pump a great amount of blood to the body.

As explained, each lead of the 12-lead in an ECG represents the electrical activity of the heart in certain direction over time. Each lead has a given name: the 3 limb leads (I, II, III), the 3 augmented limb leads (aVR, aVL, aVF), and the 6 precordial leads (V1, V2, V3, V4, V5, V6), which directions can be checked in Figure 2. The limb leads and the augmented limb leads lie in the body plane (frontal plane), while the precordial leads lie in the horizontal plane, perpendicular to the frontal and containing the heart.

3.2 Vectorcardiogram

After knowing of the existence of the standard 12-lead ECG, one may wonder if the information that each lead is gathering from the heart may be redundant. That is, leads close to one another may capture information that does not differ too much. So, assuming the human body is a three-dimensional structure, it should be possible to represent the evolution of electrical signals coming from the heart in a $3D$ space. Doing that, the electrical signal can be represented in an understandable space and also be captured but with less computational space: instead of a $1D$ vector for each lead (voltage over time) for a total of 12 vectors of length N (depending on the sample frequency) it could be stored in a single $3D$ vector, or a $N \times 3$ matrix, that could be able to capture the most important information avoiding redundant data.

This exact idea led to the development of Vectorcardiography, by E. Frank in 1950. He wanted to construct three orthogonal leads containing all the electric information: leads X , Y , Z . This idea of expressing the electric information in Vectorcardiograms (VCG) started to popularize in late 90s, with the advance of computational methods and power. The X , Y and Z vectors are just the result of projecting the standard 12-lead of the ECG in 3 orthogonal directions in space. Different matrixes can be used to do this linear transformation, depending on which 3 directions we want to project the leads and which parts of the heartbeat want to be maximized. The dimensions of the matrixes are 3×8 , as the augmented limb leads and the limb lead III are not used. The most commonly used is called Dower matrix. However the paper [1]

studies different matrix transformations for different purposes, and concludes that the best one for maximizing information in QRS complex is the QLSV matrix. This will be useful for a certain study in our thesis, and along with other matrixes can be found in the Appendix A.

Then, given a fixed matrix M , the X , Y , Z values are computed through the function f :

$$f: \mathbb{R}^8 \longrightarrow \mathbb{R}^3$$

$$\begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_4(t) \\ V_5(t) \\ V_6(t) \\ I(t) \\ II(t) \end{bmatrix} \mapsto \begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix} = f \left(\begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_4(t) \\ V_5(t) \\ V_6(t) \\ I(t) \\ II(t) \end{bmatrix} \right) = M \cdot \begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_4(t) \\ V_5(t) \\ V_6(t) \\ I(t) \\ II(t) \end{bmatrix}$$

The representation of the heartbeat in the VCG is a 3D vector loop, starting and finishing at the same point. The different parts of the standard ECG (P-wave, QRS-complex, etc.) are observed in different characteristic zones of the loop.

Several studies have been performed about the improvements in diagnostics of certain cardiologic diseases using this way of concentrating and representing the data.

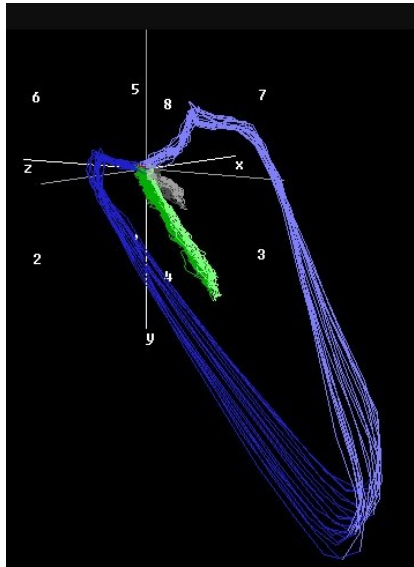


Figure 3: Example of a VCG representation over time

However, the authors of the current work think that by forcing the 3 directions over which to project the 12 leads to be orthogonal (in fact we use 8 leads instead of 12, and we can do that) it is hard to incorporate patient specificity in the visualization, and thus this representation can be improved by taking into account the variability between different patients, extracting the most relevant information from the current subject. Also introducing concepts related to dynamic systems could be interesting, like the reconstruction of the systems attractor.

3.3 Attractor

An attractor is a mathematical concept in the field of Dynamic Systems. It is a set of numerical values toward which a system tends to evolve. System values getting close enough to the attractor values will remain close even if slightly disturbed.

When applying this concept to our case, treating the heart as a dynamic system, it is impossible to know the exact attractor as the equations leading the system are unknown. However, Takens theorem [2], also known as embedding theorem, allows us to reconstruct it.

The theorem establishes that, when there is a measured quantity from a dynamical system (electrical activity from the heart in our case), it is possible to reconstruct a state space that is equivalent to the original (but unknown) state space composed of all the dynamical variables. These dynamical variables should have the key for perfect diagnosis, and reproducing them is the aim of [3], which among others articles motivated this thesis. It is possible to see in it how the reconstruction of the attractor and other parameters are used in order to triage successfully healthy and unhealthy patients.

Here, for a given output signal $x(t)$, we produce n -dimensional objects $y(t)$ as following:

$$y(t) = [x(t), x(t + T), \dots, x(t + (n - 1)T)] \quad (1)$$

Where n refers to the embedding dimension and T is a given time delay parameter. Takens theorem allows us to take $n \geq 3$ for an embedding, though dimension of the phase space is unknown. However, there is much literature about how to set the value of T . As established in [4], T is set to be the first minimum of the function I , called mutual information, understanding the formula in the discrete form, since we

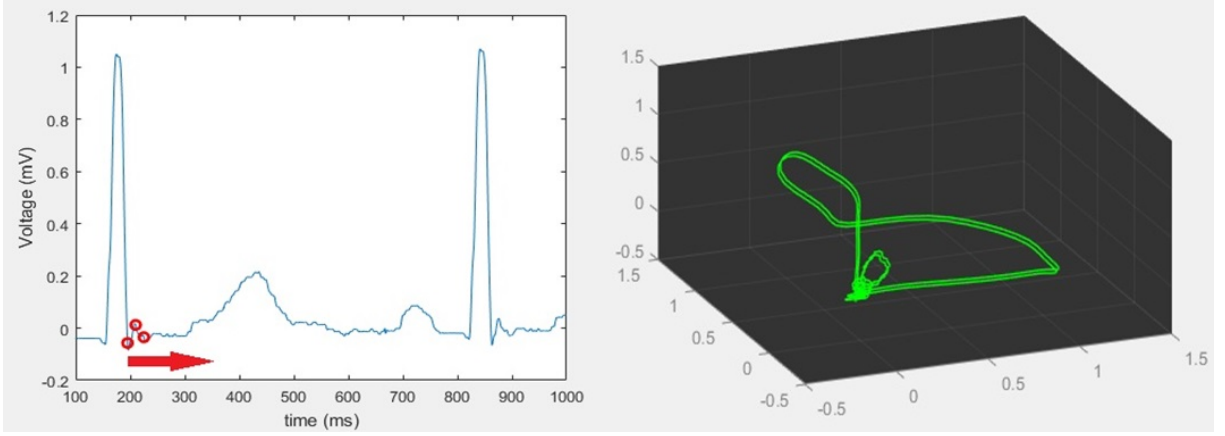


Figure 4: Representation of how a 3D attractor of a single ECG signal lead is constructed for a given $T = 15ms$ time delay. We simply plot the evolution of the 3 red dots over time.

will work with a fixed sampling frequency f_s :

$$I(a, b) = \sum_{j,k} P(a_j, b_k) \log_2 \left[\frac{P(a_j, b_k)}{P(a_j)P(b_k)} \right] \quad (2)$$

Where $a = x[n]$, $b = x[n + Tf_s]$ and $x[n] = x(t_0 + \frac{n}{f_s})$. Note how $d[n]$ with $n \in \mathbb{N}$ refers to d being a discrete signal; and $P(d_j)$ refers to the probability of measurement d_j (d_j are the points in the d signal), and practically we will compute them using the normalized histograms, as we are working with deterministic signals. Ultimately T would be chosen as the first minimum of $I(T) = I(x[n], x[n + Tf_s])$.

4. PLAR : Our New Space

After having introduced the main concepts that will be applied in this work, it is time to focus in our approach.

We wanted to create a new three-dimensional visualization tool that captures and concentrates the most relevant electric information for each patient, relatively easy to interpret for non-experts in the cardiologic field and offering the possibility of optimum patient-triaging and anticipated diagnosis in further steps of the tool development, with the final objective of speeding up the health system and focusing resources on most critic patients.

Taking as a starting point the concept of VCG, we wanted to create a three dimensional space in which the representation of the electric signal of the heart would be able to adapt to every different patient. This patient specificity is accomplished through choosing the directions over which project the standard 12-leads from the ECG. Or what is equivalent, learning from the X , Y and Z vectors of the VCG and projecting them to certain directions depending on how they behave in order to broaden and maximize the electric information. However, we did not want only to deal with the electrical signal at every time step, but also to take into account the previous electric behaviour of the heart in order to apply the attractor concept introduced in the previous section. So, understanding the heart as a dynamic system, it does not sound wild to add a component of attractor in our space, that is a time-delayed signal, so that our tool may be able to grasp part of the underlying mechanisms of the system

allowing us to capture subtle changes over short or long time periods.

Our new visualization tool will be referred in this thesis as PLAR : Patient-Learning Attractor Representation.

4.1 First Component of PLAR

Maximizing the electric information of the heart may sound ambiguous. There is a point in a heartbeat of every 12-lead ECG where the absolute value of the voltage is maximum. This corresponds to the ventricular contraction, that is the moment the ventricles contract and pump blood to the body and lungs: the R-peak in the QRS-complex. In the VCG vector loop, the QRS-complex is represented as the wider turn in the heartbeat representation. So, analogously to what happens in every 12-lead of an ECG, there is a direction in the 3D space of the VCG in which projecting the loop over that direction, the maximum of the absolute projected signal is maximum. Or formally, we want the vector $\vec{R} \in \mathbb{R}^3$ such that:

$$\vec{R} = \arg \max_{\vec{Q}, \|\vec{Q}\|=1} \left(\max_t \left[\left\| [X(t), Y(t), Z(t)] \cdot \vec{Q} \right\| \right] \right) \quad (3)$$

So, that is the direction over which, intuitively, the electric potential measured from the heart is maximum. For our visualization tool then, it will be useful to have the projection of the VCG over the \vec{R} direction as the first component of our new

representation:

$$\gamma_1(t) = \vec{R} \cdot \begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix} \quad (4)$$

The Matlab code implemented for the computation of this \vec{R} direction can be found in Appendix B.

4.2 Second Component of PLAR

Similarly to what has been done for the first component, a vector \vec{V} over which project the signal coming from VCG is also desired to define the PLAR's second component. In this case, to avoid redundant electric information, we will choose \vec{V} such that, if $L_{\vec{V}}$ refers to the VCG signal projected over \vec{V} direction:

$$\vec{V} = \arg \min_{\vec{Q}, \|\vec{Q}\|=1} I(L_{\vec{R}}, L_{\vec{Q}}) \quad (5)$$

Where I refers to the mutual information function, defined previously in expression (2). That is, we want the second component to be the projection of the VCG signal over the direction \vec{V} that has the least information in common with the VCG signal projected over \vec{R} , which is the first PLAR component. Doing so, we pretend to broaden the space, and capture as much information as possible with only 2 projections of the

VCG signal. Then we obtain the second component of our PLAR as:

$$\gamma_2(t) = \vec{V} \cdot \begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix} \quad (6)$$

Details about how this second component is computed are explained in a project that was being developed in parallel to this thesis (see [7]).

4.3 Third Component of PLAR

To incorporate the concept of attractor, the third component will be the same as the first one but delayed a certain amount of time T , or a certain time samples:

$$\gamma_3(t) = \gamma_1(t + T) \quad (7)$$

Note how choosing this signal $\gamma_1(t)$ to be delayed can seem arbitrary. But we definitely decided on this after testing several other options.

The delay term T can be computed every time as stated in [4], or set to a fixed value. We have observed that taking $T = 15ms$ as delay has a good behaviour for all the study in this thesis, or 15 time samples since the frequency sample of our ECGs is $f_s = 1000Hz$. So, after some tuning of the parameters, the standard PLAR transformation is set to be the three-dimensional curve over time (note we delayed

the second component as well for visualization purposes):

$$\gamma(t) = [\gamma_1(t), \gamma_2(t + T), \gamma_3(t)] \in \mathbb{R}^3 \quad (8)$$

Having defined the 3 components of our PLAR, given 8 leads of information of a heartbeat, L , the 3D curve obtained as PLAR is totally defined knowing the matrix M (D or $QLSV$, depending on the intentions in this thesis), the 2 vectors over which we project the VCG curve \vec{R} and \vec{V} , and the time delay T from which we are defining the attractor component. That means the 3D curve S obtained in PLAR can be understood as a function of those parameters:

$$S = S(L, M, \vec{R}, \vec{V}, T)$$

Note we will use S to refer to the set of three-dimensional points obtained in the PLAR space.

An example of a healthy patient PLAR is presented in Figure 5. See how we compute the 3 components of our PLAR curve for a given ECG signal in Appendix B.

4.4 Obseervations

From now on we will use the notation L_{HB} to denote the electric signal from the 8 leads of standard ECG corresponding to a single heartbeat used in the VCG transformation. Note how in practice there is a sample frequency, we know the electric

output of the heart not as a continuous function of time but sampled, though the notation does not take into account this fact. Supposing the sample frequency is the same for all the leads, and let T_{HB} and f_s be the duration in seconds of the heartbeat and the sample frequency, respectively, then, if $p = \lfloor T_{HB} \cdot f_s \rfloor$, we practically have $L_{HB} \in \mathbb{R}^{8 \times p}$. We will use $S(L_{HB})$ to denote our PLAR transformation explained previously, $S(L_{HB}) \in \mathbb{R}^{3 \times \lfloor (p - T \cdot f_s) \rfloor}$ (see how the fact of using an attractor results in having a few less points in PLAR space, if only known the electric information of a single heartbeat).

At this point, having defined our space in order to broaden the information received from electric heart impulses, it would be the time to start verifying it. However, when doing so, one must be aware of certain aspects regarding the PLAR.

To start, as we are performing a three dimensional representation of the electric signal from 8 different signals, it is clear that modifications in any of the 8 ECG signals will have an impact on the $3D$ representation. However, if we wished to modify the three dimensional representation to see what changes in \mathbb{R}^8 space (changes in ECG signal) would cause the modification, it would not be feasible since the exact inverse transformation cannot be defined. Even if we were able to reproduce the three dimensional VCG figure from our PLAR curve, the fact of going linearly from an 8 dimensional space to a 3 dimensional one makes the inverse application clearly non well-defined. It is like projecting a $3D$ vector over a plane, and then trying to reproduce the original vector **only** through the plane's projection. Then the process of understanding

carefully our new space can only be performed unidirectionally.

Another aspect to take into account is that PLAR is insensitive to rotations in the VCG space. That is, suppose there are two different 8-lead ECG signals, $L_1 = L_{HB,1}(t) \in \mathbb{R}^8$ and $L_2 = L_{HB,2}(t) \in \mathbb{R}^8$, and $L_1^{VCG} = L_{HB,1}^{VCG}(t) \in \mathbb{R}^3$ and $L_2^{VCG} = L_{HB,2}^{VCG}(t) \in \mathbb{R}^3$ being their VCG transformations through a linear transformation. If L_2^{VCG} can be obtained rotating L_1^{VCG} around the 3 principal VCG axes (the same angles for all t), then the PLAR curves obtained are the same for both signals ($S(L_1) = S(L_2)$).

Demonstration: Supposing the representation in the VCG space starts at the origin, and let γ, β, α be the successive rotations around z , y and x VCG axis respectively.

That is, the rotation matrixes are:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, B = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}, C = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Such that $L_2^{VCG}(t) = A \cdot B \cdot C \cdot L_1^{VCG}(t)$ for every $t \in [0, T_{hb}]$, where T_{hb} is the duration in seconds of the heartbeat being processed.

If $\vec{R}^1 = (R_1^1, R_2^1, R_3^1)^T$ is the direction over which the maximum of projecting L_1^{VCG} is achieved, of value $max \in \mathbb{R}$, that is, it exists a point $p_{max}^1 = (a, b, c)^T \in L_1^{VCG}$ such that $\vec{R}^1 \cdot p_{max}^1 = max$, then the rotation of \vec{R}^1 , $\vec{R}^2 = A \cdot B \cdot C \cdot \vec{R}^1$, is the direction over which the maximum of projecting L_2^{VCG} is achieved, also giving value

max , with the point p_{max}^2 being the rotation of p_{max}^1 :

$$\vec{R}^1{}^T \cdot p_{max}^1 = max \implies \vec{R}^2{}^T \cdot p_{max}^2 = (A \cdot B \cdot C \cdot \vec{R}^1)^T \cdot (A \cdot B \cdot C \cdot p_{max}^1) = \quad (9)$$

$$= \vec{R}^1{}^T \cdot C^T \cdot B^T \cdot A^T \cdot A \cdot B \cdot C \cdot p_{max}^1 = \vec{R}^1{}^T \cdot p_{max}^1 = max \quad (10)$$

Where it has been used that the inverse matrix of rotation matrixes are their own transposed. Note how if there exists another direction over which projecting L_2^{VCG} is maximum and greater than max , then the rotation of the negative angles γ, β, α of that direction would give a direction over which projecting L_1^{VCG} would be greater than max , following an analogous procedure, coming into contradiction.

Then, as the direction \vec{R}^2 defined is the one over which projecting L_2^{VCG} is maximum, then it is clear that, following an analogous procedure that one in (9) and (10) changing p_{max} for each point in L_2^{VCG} and expressing it in terms of rotation of L_1^{VCG} , the first and third components of our PLAR, $\gamma_1(t)$ and $\gamma_3(t)$, are the same for both L_1 and L_2 .

Considering now the minimization of mutual signal information as defined in (5), the vector \vec{V}^2 over which projecting the signal gives the minimum information when compared to the signal projected over \vec{R}^2 is the rotation of the one that does so when compared to \vec{R}^1 (by following again an analogous procedure of (9) and (10)). So the second components $\gamma_2(t)$ are also equal for L_1 and L_2 . \square

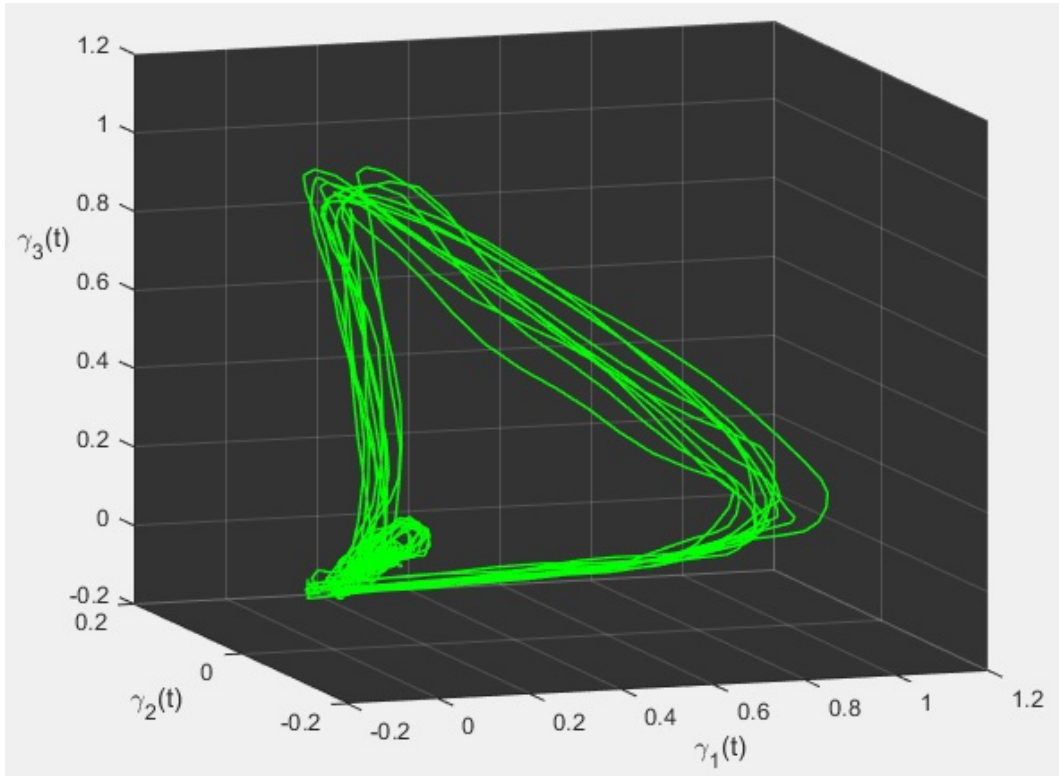


Figure 5: Example of PLAR from a 10 seconds ECG signal of a cardiologically healthy individual

5. Data Acquisition for First PLAR Development

After developing theoretically the idea of the new visualization cardiologic tool we wanted to introduce different sets of real ECGs with simplified most common anomalies in order to set the value to certain model parameters, like time delay T , to an optimum value and identifying how these changes are expressed in our PLAR space. To do so we disposed from a large amount of ECG records of different patients and different years. However, as strange as it may sound, current ECG records are still

stored in the hospital database only in *.pdf* format, that is no voltage vectors linked to ECG leads are stored for any patient. Then we had to work on an algorithm that enabled us extract the electrical signals of the beats from those standardized ECGs records with the objective of analyzing and modifying them at our will once extracted.

The main parts in which the algorithm can be split are:

1. From *.pdf* to *.png*: In order to extract numerical data from the ECG records we convert the *pdf* file to *png*, so that as an image *.png* format we can work now with a matrix (size $n \times m$, where n and m are determined by the resolution of the files conversion) of 0 and 1, where 1 means there is electrical signal in certain zones of the image file and 0 means no electrical signal recorded.
2. From *.png* to electrical 12-lead signals: Knowing the resolution of the matrix (eg. of the image file) and supposing each lead of the 12 lead signals is located approximately at the same region of the standardized ECG paper, it is possible to extract the time vectors and voltage vectors of each lead.
3. Extracting a single beat: Using the *Pan and Tompkins* algorithm ([5]) to detect automatically the peak of the R-wave in the reference lead, and an averaged-window correlation method to detect when a P-wave starts, it is possible to extract a single 12-lead heartbeat for each patient ECG record.
4. Extra modifications at ST segment: alterations introduced artificially on the ST segment, in different leads, have been made in order to observe how they are shown in our space.

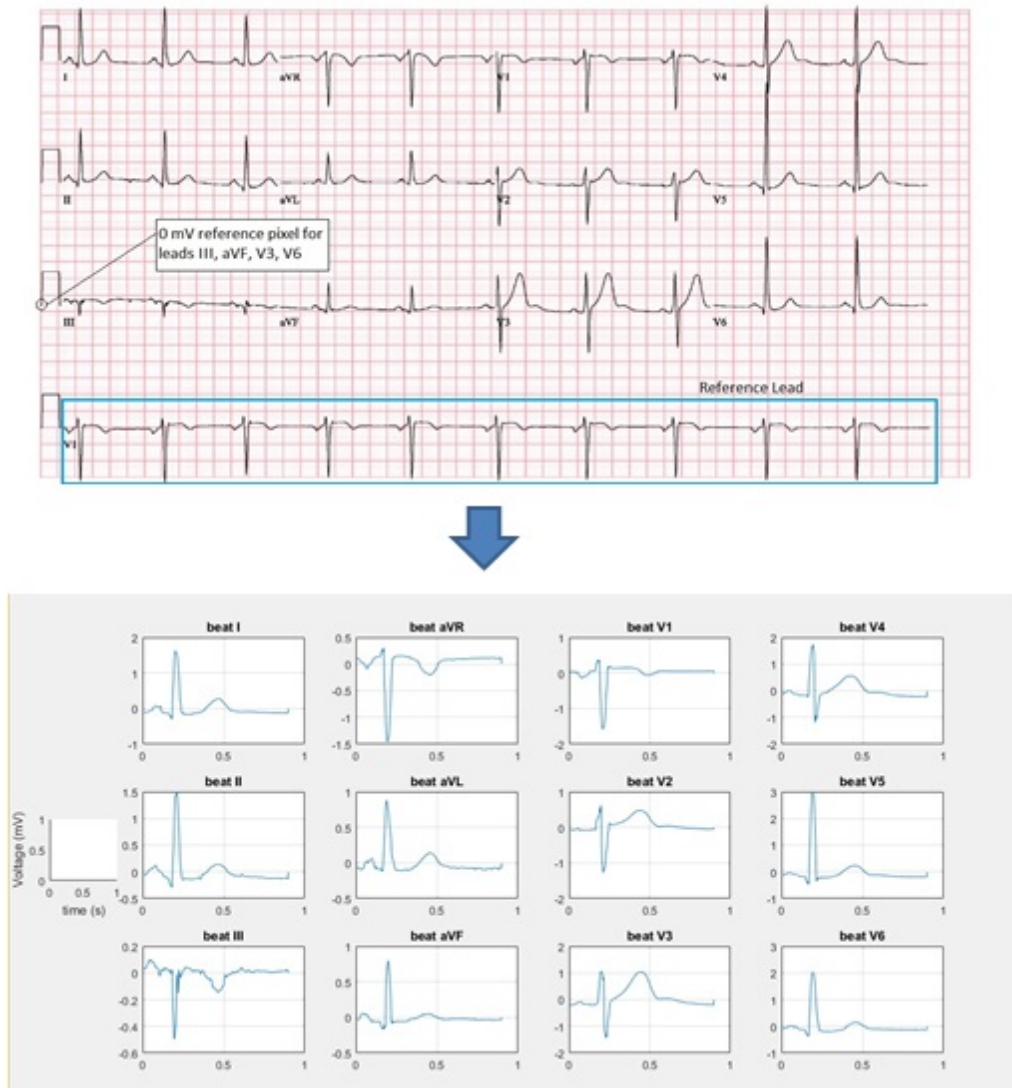


Figure 6: The implementation of the algorithm allows a conversion from paper ECG records to digital data

The main idea behind the algorithm is that, given the 0 and 1 matrix representing the image of the ECG, knowing the vertical pixel associated to 0 mV for each lead, the vertical range of pixels where a lead signal starts and each lead duration (all these parameters are fixed for a standardized ECG), to follow the adjacent contiguous 1s pixels (meaning signal has been recorded) over each horizontal pixel. Doing that with resolution enough it is feasible to collect the electric voltage for each lead of a

single heartbeat (under the assumption of low variability of the heartbeat signal over small periods of time) from only a *pdf* record of the ECG (see Figure 6).

It is important to understand the limitation of our study using this method to extract the electrical data from the ECG record. Errors and noise in the process of getting the *pdf* file from the ECG recording machine, added to errors due to resolution, filtering of the signal and R and P wave algorithm detectors can result in a total error that is far from being null. Thus, for further testing of our visualization tool these errors are not taken into account, but for future stages of the technology it is necessary to perform further research in order to diminish these errors.

6. Signal Filtering and Smoothing

Once the signal has been extracted from the ECG record, a process of filtering and smoothing the signal needs to be done in order to get rid of non-desired perturbations and to make the 3D PLAR curve visually cooler. The difficulty is in maintaining the electric information outputted while eliminating other electric perturbations.

Like other electrical signals, ECG signals can be corrupted by various kinds of noise. These can be presented as low-frequency noise, where the baseline of the signal is perturbed (known as baseline wander), and high-frequency noise. They are caused mainly by imperfections in recording devices or sensors, and the wrong placement of the electrodes or some shaking of the patient during the recording of the data.

For the baseline wander perturbation, it is obvious that by default the signal should

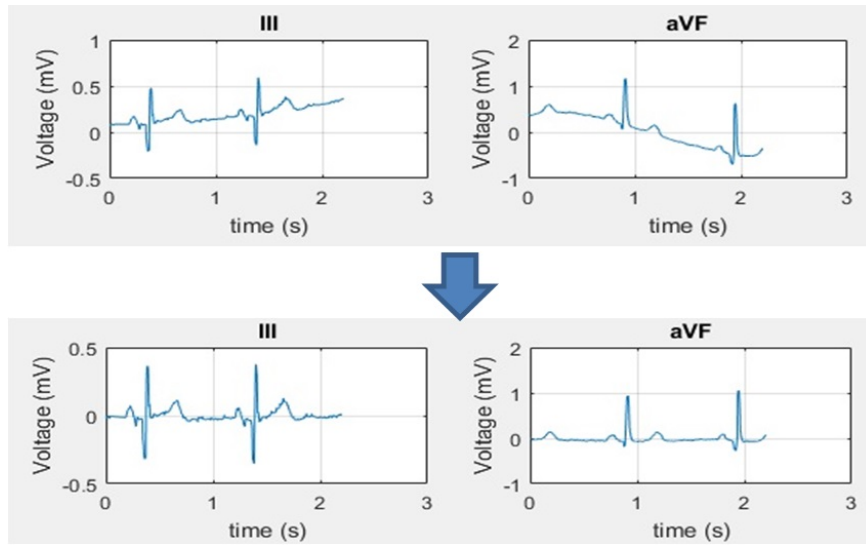


Figure 7: Example of low frequency noise filtering using a 3rd degree polynomial fitting

be zero centered. The signal in each lead should be 0 when there is no electrical response from a depolarization in the heart cells. For large ECG records, of the order of minutes, there is a lot of bibliography of ECG baseline filtering, and a very common treatment is a *Chebyshev Type II filter*. However for a 2.5 seconds lead signal baseline wander that cannot be used, and simply fitting a 3rd degree polynomial by linear least squares and subtracting it from the original signal has resulted in the fastest and best method.

Our algorithm is robust to high-frequency noise, as we will see in later chapters. However in order to smooth the signal for visualization purposes and to soften the ECG signal extracted through our data acquisition algorithm we have implemented the *Savitzky-Golay filter*. This filter is designed to increase the signal-to-noise ratio, through a convolution process, by fitting successive sub-sets of adjacent data points with a low-degree polynomial again by linear least squares method, and its applica-

bility in the electrocardiography field is studied in [6]. Though it can be expressed in terms of successive coefficients, the idea of this filter is, for each point of the signal y_i , if n is the degree of the polynomial and m is the frame length (must be odd):

- Find the coefficients of the n -degree polynomial $P_{i,n}$ that bests fits the m points

$$\left(y_{i-\frac{m-1}{2}}, y_{i-\frac{m-1}{2}+1}, \dots, y_{i-1}, y_i, y_{i+1}, \dots, y_{i+\frac{m-1}{2}} \right).$$

- Assign to the new signal the value $Y_i = P_{i,n}(x_i)$ instead of y_i .

Note how for the first $\frac{m-1}{2}$ points the filter is not defined, and the points of the end of the signal are used. The opposite happens for the last $\frac{m-1}{2}$ last points.

After some tuning of the parameters, it results that for our data the best combination is $n = 3$ and $m = 17$, the same as established in [6]. This combination offers a compromise between smoothing of the signal and relatively low error when comparing the resulting signal to the signal non-smoothed.

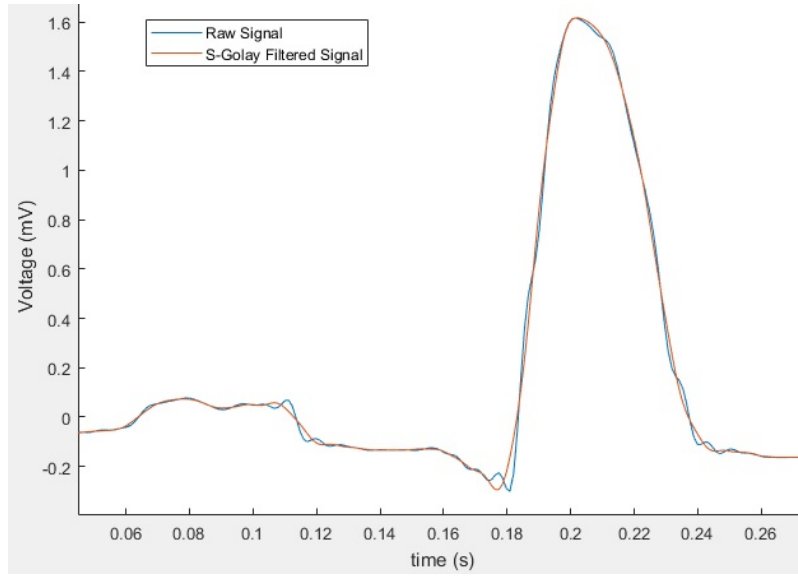


Figure 8: Zoomed image of a lead I signal of a heartbeat for the signal unfiltered and filtered using S-Golay filter against high frequency noise

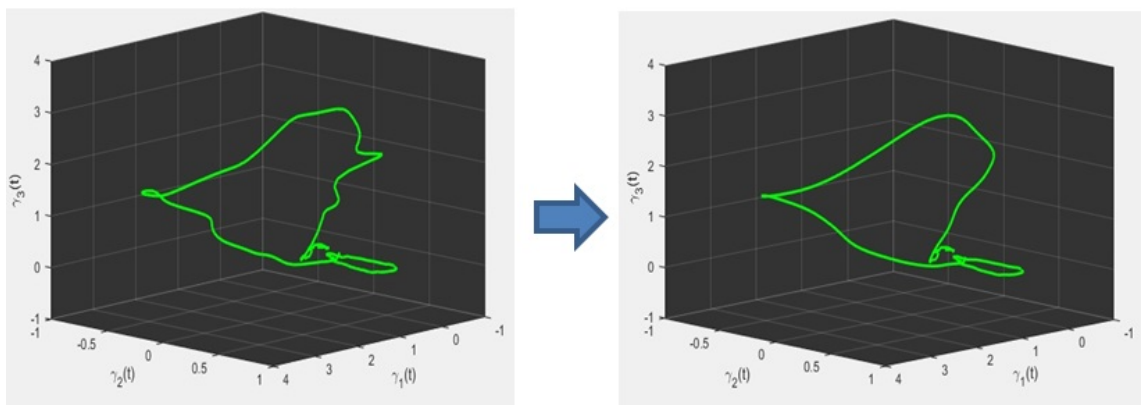


Figure 9: PLAR representation before and after smoothing the signal extracted from the ECGs

7. ST Segment Variations

We want to test the behaviour of the PLAR new space defined to certain cardiologic anomalies. As the space has been created theoretically and intuitively to capture and synthesize heart electric information, though no theoretical results can be obtained, we can start to view how experimentally it behaves. There are many ways to do so, introducing different ECGs with anomalies. However in this step we have focused on one common anomaly, ST segment variations. As its name suggests, this anomaly is defined by a variation of the ST segment above or below 0.1 mV of the signal baseline, which means 1 mm above or below it in the paper ECG. This variation can be present in one or more leads of ECG, and it is associated to several conditions, the most important being heart attack. Because of that, being able to observe and understand these variations in our PLAR might enable our tool to work as a good support in diagnosis.

7.1 Methodology

To study how different S-T segment variations are shown in our PLAR space different artificial and manual modifications have been made to some leads. Specifically we have introduced, separately, alterations of -2 mm, -1 mm, 1 mm and 2 mm in the ST segment of the precordial leads V1, V2, V3, V4, V5 and V6. This alteration has been performed by adding a clean pulse of different amplitude in the specific lead

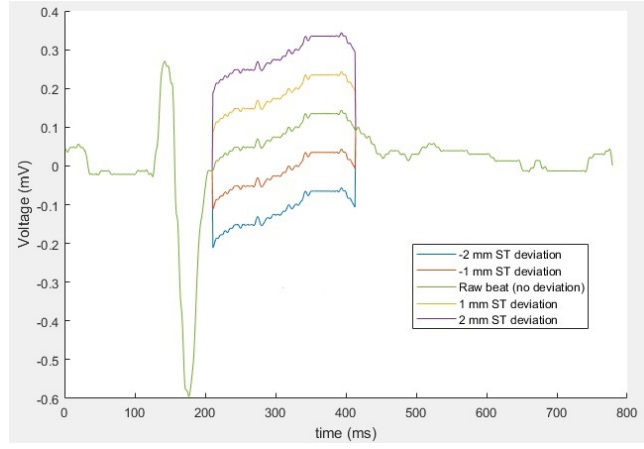


Figure 10: Lead V3 with fake ST deviations added before the high-frequency noise filtering

location (ST segment) for the considered ECG. An example of a lead modified is shown in Figure 10.

We have chosen the precordial leads because, as seen in previous chapters, they are all located in the frontal plane and we want to see how this rotation between the leads is transformed into the PLAR space.

After that, to study the consistence of our representation to high-frequency noise, some controlled noise (Gaussian white noise) has been added to the raw signal. Note how despite having developed a filter against high frequency noise, we pretend to study how our PLAR space is really affected by this noise in the ECG. The amount of noise added is quantified through the signal to noise ratio (STNR), measured in dB. If the signal is expressed as $s(t) = e(t) + n(t)$, where $e(t)$ is the unperturbed signal, without noise, and $n(t)$ represents the noise, then we have

$$STNR = 20 \cdot \log_{10} \left(\frac{A_{e(t)}}{A_{n(t)}} \right) \quad (11)$$

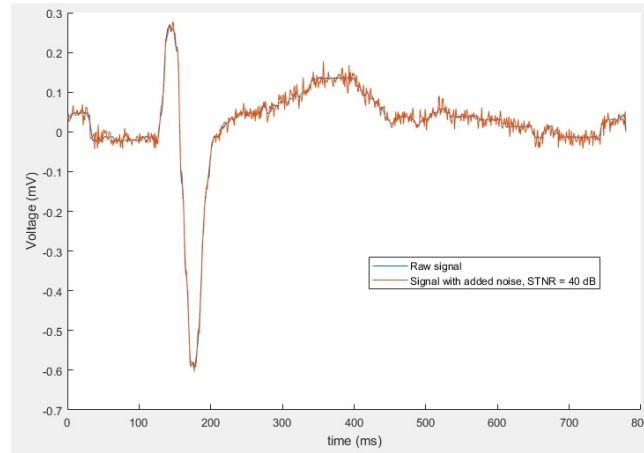


Figure 11: Heartbeat in V3 lead, raw signal and signal with added noise. Signal To Noise Ratio is 40 dB

Where A_i refers to the root mean square (RMS) amplitude of the signal i (RMS is defined as the square root of the mean over time of the squared signal). Note how the higher the $STNR$ the cleaner is the signal.

7.2 Results and Observations

As expected, ST elevations in different leads are presented in different ways in our space. In Figure 12 we can observe that all PLAR remains the same for all leads deviations except the zone associated to the ST segment. In the zoomed image we can appreciate how 1 mm ST deviation in different precordial leads make the resulting PLAR differ for each lead. When comparing the PLAR curves with ST deviations to the raw or healthy PLAR curve it has been observed that each lead ST deviation has a direction of propagation. What is more, given a lead, in the zone of the curves associated to the ST segment, distances in the PLAR space to the raw PLAR curve

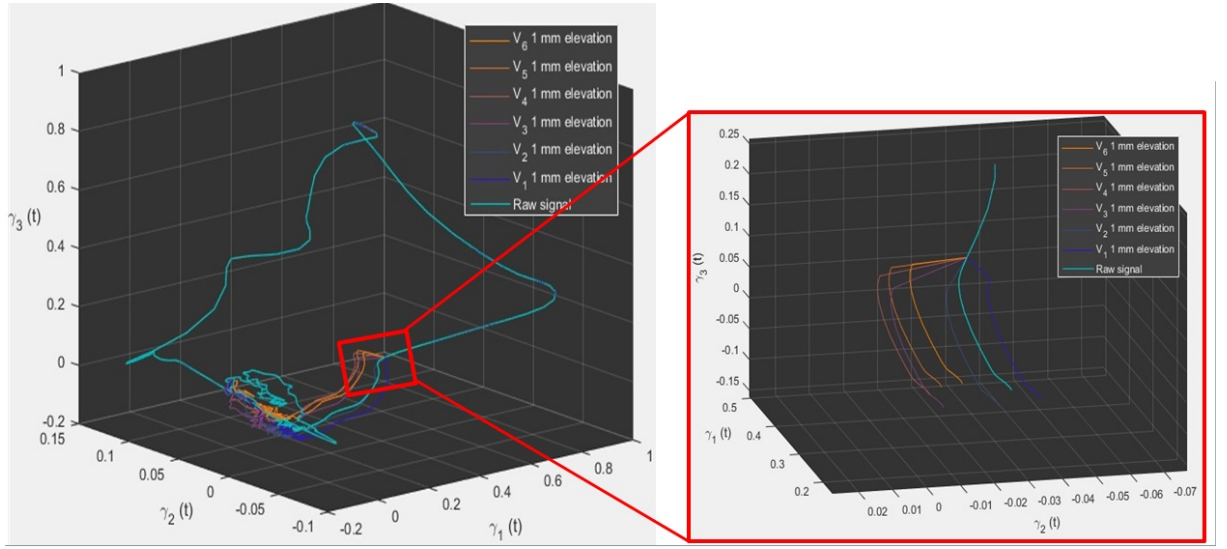


Figure 12: PLARs of the same beat with 1 mm ST elevations in each precordial lead

are proportional to the ST deviation in the lead. This fact is shown in Figure 13, in which PLARs resulting from deviations in lead V3 ST segment lie in the same line, and the distance between the curves are the same for the given deviations. Note how a safety criterion could be established for a given patient whose standard beat is recorded, and S-T segment deviations from it in any lead could be rapidly detected.

To better understand the rotation of the curves depending on the lead, we present in Figure 14, for 2 independent ECG, the rotation of each curve relative to the curve with ST elevation on lead V1.

The same angle evolution but with added noise in the signals is also presented. Then, watching these results, it can be stated that the relative rotation between the leads is conserved in our visualization. Notice how the perturbations in different leads only manifest in the zone linked to the ST segment. The closest lead to V1 in angle is V2, with a 30 angle separating them approximately, and then come

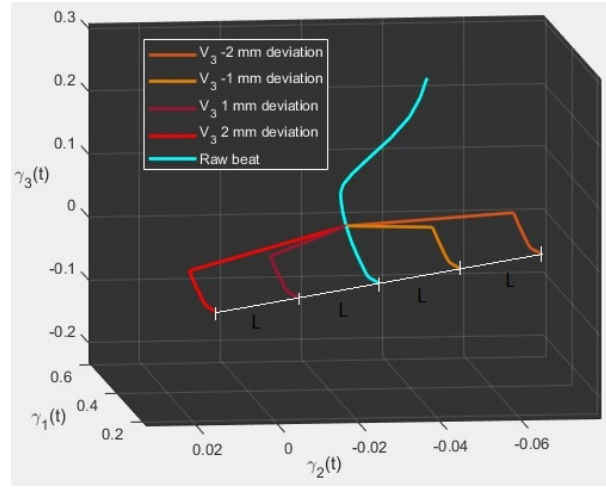


Figure 13: Zoomed image of how deviations in lead V3 are observed in PLAR

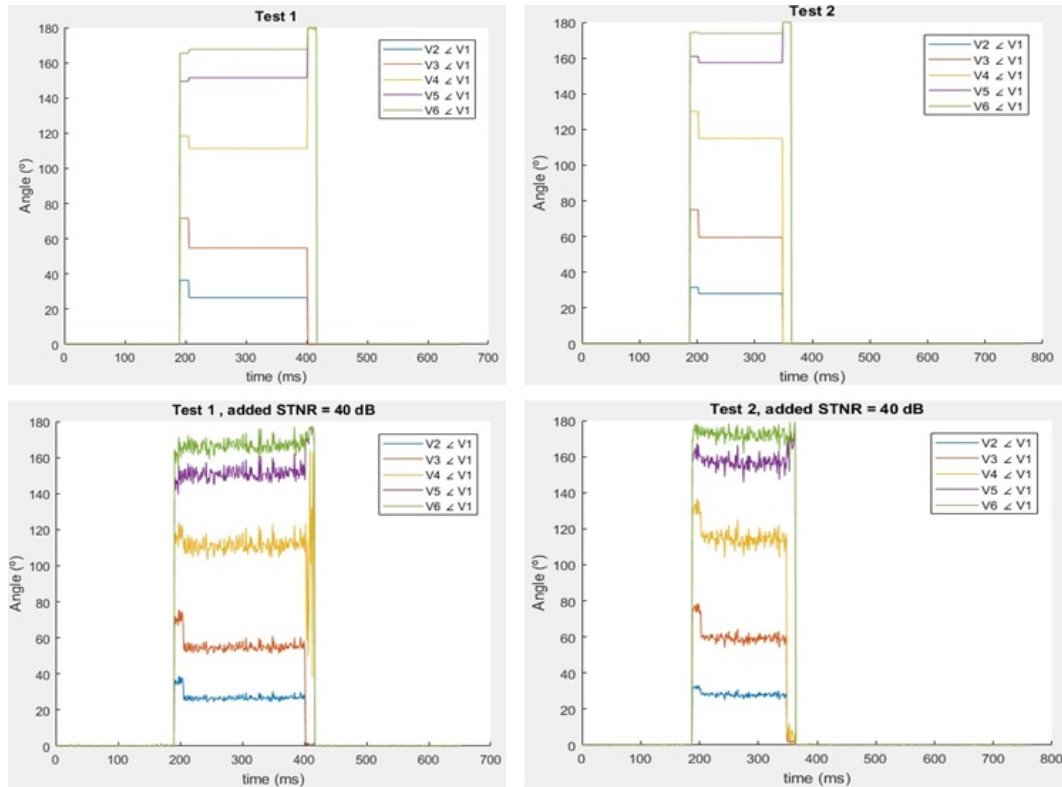


Figure 14: Angle evolution for 2 different ECGs when a 1 mm increment in ST segment is incorporated , without and with added noise. Note how the effect of the delay T is clearly observed in the first 15 and last 15 milliseconds of the variations

V3, V4, V5 and V6 just in order, coinciding with the vector leads distribution in space (see Section 3.2).

Also, note how the addition of noise is clearly observed in the graphs, but the overall distribution of angles between the leads remains the same. However, the relative angles of V6 and V5 leads start to overlap when further noise is applied and it would not be possible to totally distinguish them in certain zones if the signal had a STNR greater than 40 dB.

These results constitute an example of how our tool can be used for a fast and understandable diagnosis tool. Compared to the VCG representation, the results perform much better in PLAR space when it comes to isolate the anomaly and distinguish ST-segment deviations in different leads.

8. Progressive Disease Tracking

Apart from testing how sudden changes in the signal from the heart are represented in PLAR we would like our tool to be also useful for tracking progressive conditions. For instance, evolving physical and physiological changes in the heart must have an impact in the electric signal measured over years, which adequately processed through our new technology can lead to better tracking and control of the patient. Here we will test our hypothesis and develop our tracking methods applied to Hypertrophic Cardiomyopathy (HCM). This is a condition in which a portion of the heart becomes thickened without an obvious cause, which causes difficulty to pump blood. It is the

main cause of death in young athletes. The region of the standard ECG that seems of interest is the QRS complex. In standard ECGs, when the disease is not yet fully developed it is hard to tell when a patient has this condition, and further medical and more expensive tests are required (mainly Echocardiograms, EEG). Fully describing the evolution of the condition only through ECG data would mean a decrease in the cost of treatment and patient follow-up.

8.1 Methodology

For the theoretical development of our method we require of a standard record of patient suffering from HCM. We disposed of 10 standard ECGs, from years 2011 to 2018, of the same patient, as his heart developed the condition. If we represent the curves of the PLAR or the VCG representation for each year ECG, we are not able to see any trend that make us think that the heart is getting thicker and unhealthier, as it can be observed in Figure 15. Here we are using the *QLSV* matrix for generating the VCG from the ECG signal and then obtaining also our PLAR, because as mentioned in [1] this is the matrix that maximizes the information in the QRS complex.

In this case neither PLAR nor VCG curves evolution over time show a clear advance of the disease or have a behaviour that expresses a trend over time.

However, if we consider that in the present (last 2018 ECG) the \vec{R} and \vec{V} directions presented in 4.1 and 4.2 are the ones telling more information about the heart at the present moment, if we project the VCGs over years over these present directions

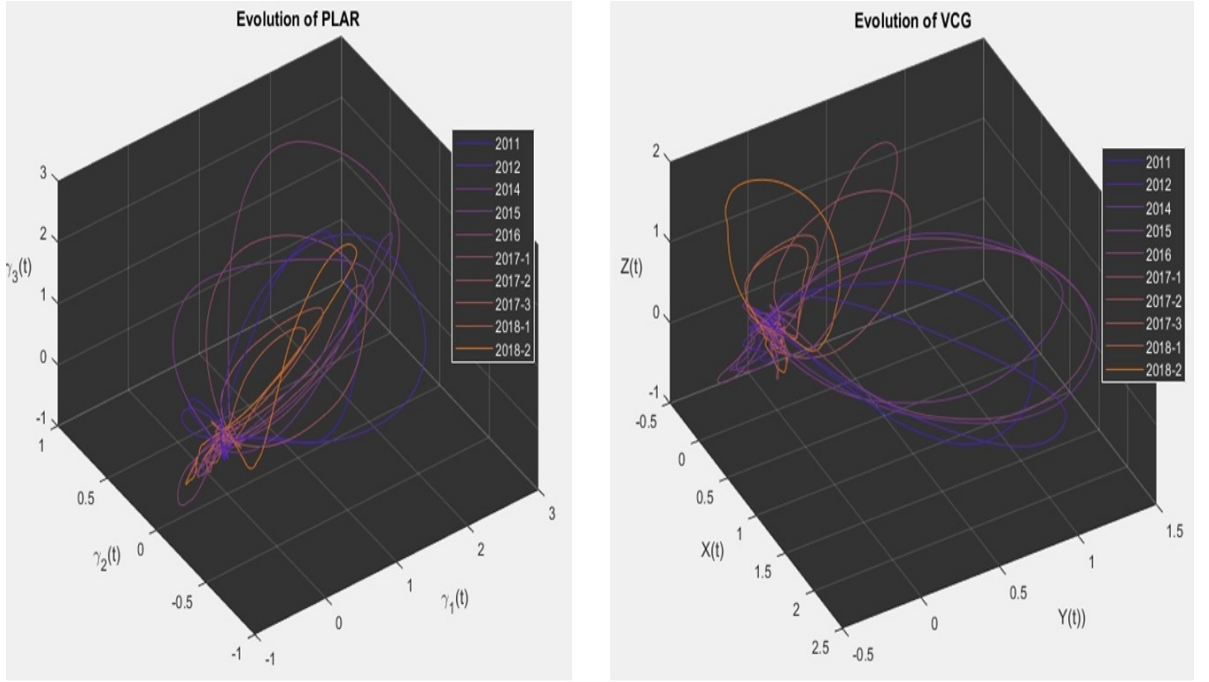


Figure 15: Evolution over years of PLAR and VCG representation for a HCM patient

instead of over each years \vec{R} and \vec{V} , we should be able to observe a trend in time. Formally, if we call \vec{R}_p and \vec{V}_p to the \vec{R} and \vec{V} directions of the present (or last year ECG), and \vec{R}_i and \vec{V}_i to those directions at i th year, how we were obtaining our PLAR curves before was, synthetically:

$$S_i = S(L_i, QLSV, \vec{R}_i, \vec{V}_i, T)$$

And now we pretend to obtain a retuned S_i as follows:

$$S_i = S(L_i, QLSV, \vec{R}_p, \vec{V}_p, T)$$

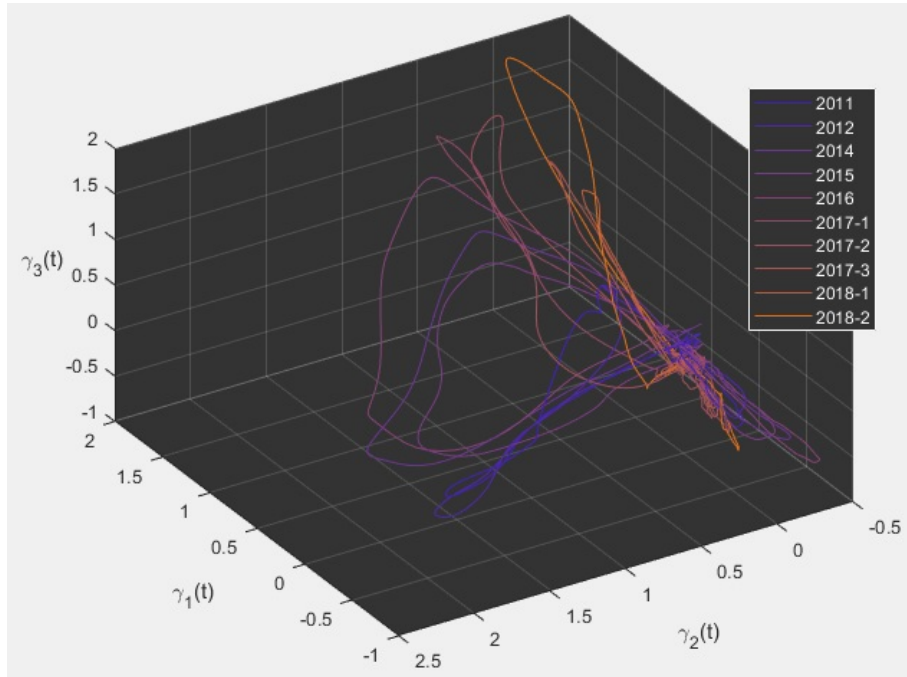


Figure 16: Retuned PLAR over years

Doing so, we can observe this retuned PLARs curves in the Figures 16 and 17.

In this case, if we focus on the main loop, one can observe that the visual representations over the years are getting closer and closer to the last 2018 figure. It is like the curves were unfolding over time until reaching the present state. In Figure 17 we represented only the evolution of the main QRS-complex loop of each PLAR, as medically it should be the part that tells us more information about the heart in HCM progressive disease.

For the automatic process of isolating the QRS complex loop in each PLAR curve, an algorithm has been implemented. It is based in the fact that the points corresponding to the QRS-complex, in the PLAR curve have more distance (Euclidian) from one another than the rest of the points of the heartbeat. Or from another point of view,

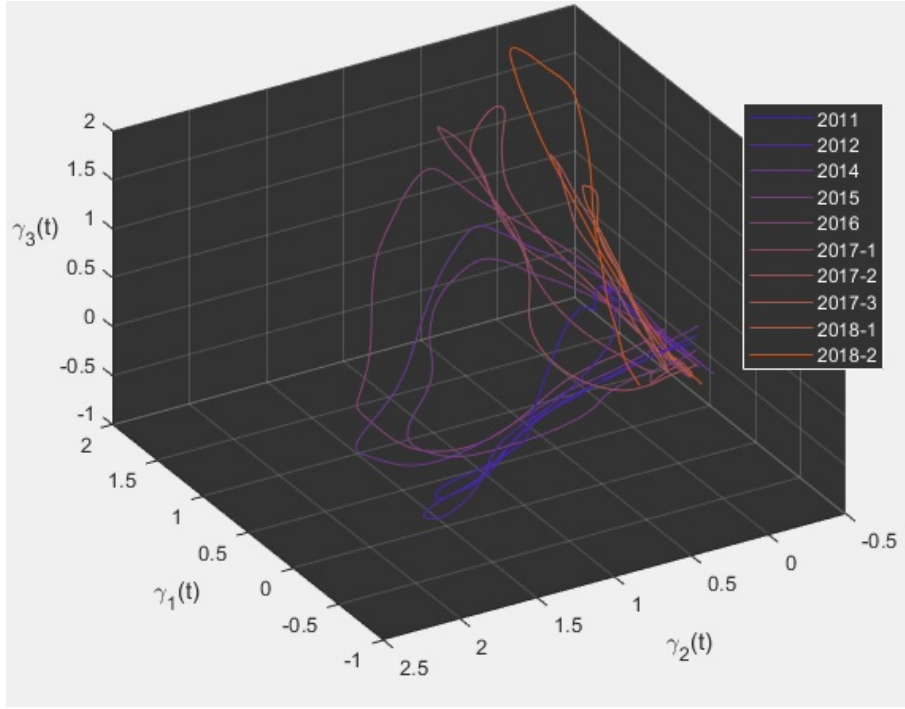


Figure 17: Main QRS loops of retuned PLAR over years

for a given heartbeat signal, it exists $\epsilon > 0$ such that:

$$|B_\epsilon(S(r)) \cap S(L_{HB})| = 1 \quad \forall r \in [T_{QRS,1}, T_{QRS,2}] \quad (12)$$

$$|B_\epsilon(S(s)) \cap S(L_{HB})| > 1 \quad \forall s \in [0, T_{HB}] \setminus [T_{QRS,1}, T_{QRS,2}] \quad (13)$$

Where $B_\epsilon(S(t))$ denotes the ball of radius ϵ in the PLAR \mathbb{R}^3 space, understanding that $S(t)$ is the point in the PLAR curve of the L_{HB} ECG linked to instant of time t ; and the interval $[T_{QRS,1}, T_{QRS,2}]$ denotes the time interval in which the QRS complex is taking place in the standard ECG. Again, consider the notation in the context of discrete temporal space, and see also how for ϵ small enough, the expression (12) holds $\forall t \in [0, T_{HB}]$.

What the relations in (12) is saying is that there is no other $S(L_{HB})$ points at a lower

distance than ϵ of the point $S(r)$ than the own $S(r)$, while this does not hold for the same ϵ when the points do not belong to the QRS complex. Applying a weaker variant of this property (code presented in Appendix B) it is possible to isolate the PLAR main loops as represented in Figure 17 and study their tendency over the years.

8.2 TDR: Today-Related Coefficient

Once we have observed the graphical tendency of the main loops of the PLAR over the years to get closer and unfolding to the final or present PLAR curve, it is time to define a metric to assess these differences. We do this with the objective of quantifying how different is each i th curve to the final or present curve, and see if there is effectively a trend over years due to the HCM condition. Also we will use the metric to check if our new representation is capable of capturing the physiological and physical changes that are taking place in the heart, which we will relate to parameters extracted from EEG over years.

Different metrics have been considered and studied. We would like to have a coefficient associated to each year PLAR curve to measure how close is the figure to the final PLAR (we call it TRC: Today-Related Coefficient). The metric that seems to best capture the evolution of the curves and that best correlates to changes in EEG parameters is the following:

$$TRC_i = \frac{1}{T_i} \int_i \overrightarrow{S_i(t)} \cdot \overrightarrow{S_p(t)} dt \quad (14)$$

This is a theoretical formula, where each i years PLAR main loop position vector is represented by $\overrightarrow{S_i(t)}$; $\overrightarrow{S_p(t)}$ is the present or last years PLAR main loop position vector; I refers to the integration domain of the curves, referring to the PLAR main QRS loop, and T_I is the duration of the PLAR main QRS loop. However, as each years heartbeat time duration does not have to be equal, and the PLAR main loops are not continuous but discrete curves, for practical purposes we use an analogous expression:

$$TRC_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \overrightarrow{S_i(j)} \cdot \overrightarrow{S_p(j)} \quad (15)$$

Where now N_i refers to the minimum number of points between the S_i curve and S_p curve, and j refers to the j -th point of the S_p curve.

Generally the number of points in S_i and S_p is different, as they are heartbeats taken at different years. Suppose N_i is the number of points in S_i and let M be the points in S_p , such that $M > N_i$ (if its the opposite the process described is analogous). In order to synchronize the position vectors multiplied at each term of the summation in the previous expression, we will approximate or *relocate* the terms $\overrightarrow{S_p(j)}$ present inside the sum of expression (15) as:

$$\overrightarrow{S_p(j)} \approx \left(\frac{j}{N_i} M - j_1 \right) \cdot \overrightarrow{S_p(j_1)} + \left(j_2 - \frac{j}{N_i} M \right) \cdot \overrightarrow{S_p(j_2)} \quad (16)$$

Where $j_1 = \lfloor \frac{j}{N_i} M \rfloor$ and $j_2 = \lceil \frac{j}{N_i} M \rceil$. See how the terms multiplying $\overrightarrow{S_p(j_1)}$ and $\overrightarrow{S_p(j_2)}$ are between 0 and 1, and adding them equals 1.

Note again how we proposed this metric with the intention of capturing the tendency

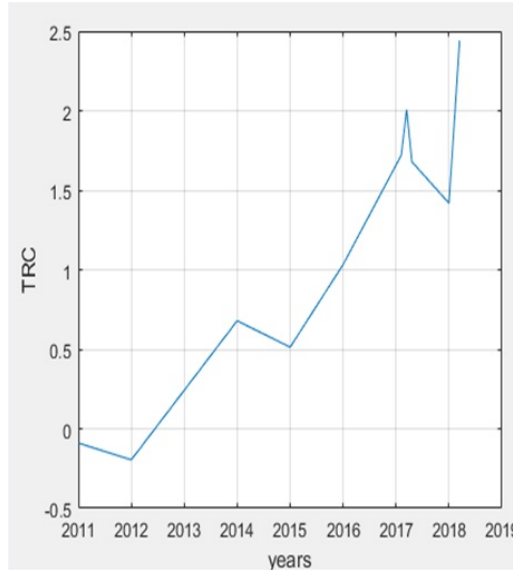


Figure 18: TRC evolution over the disease period

of the retuned PLAR curves over years after knowing how was their behaviour.

Having defined our metrics as in (15), it is obvious that TRC_i will have a greater value when the i curve is very close to the final p curve and a lower (even negative) value when the shapes of the 2 curves differ (supposing $\overrightarrow{S_i(j)}$ and $\overrightarrow{S_p(j)}$ have more or less the same modulus for each j and i).

8.3 Results and Observations

Plotting the evolution of the TRC over the years we see how, in Figure 18, there is a trend of our parameter increasing, meaning the curves are getting closer to the last PLAR curve. Note how there are 3 points associated to 2017 (there are 3 ECGs from that year) and 2 to 2018 (the last being the maximum value). Globally TRC is increasing over time, though in 2017 something changes. However, we do not want

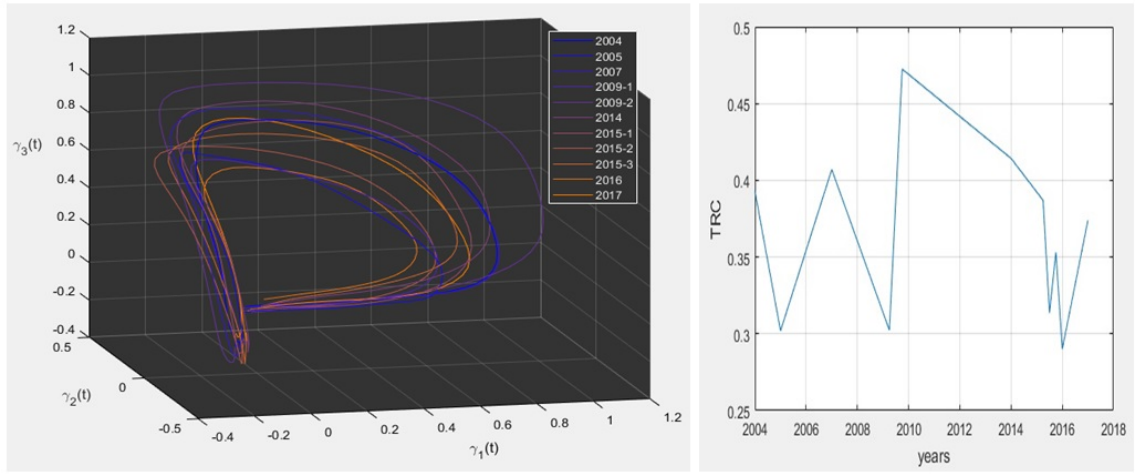


Figure 19: Retuned PLARs and TRC evolution over years for a healthy individual.

our space and parameter to express solely time progression, but to be able to track physical and physiological changes of the heart as it suffers the disease.

The evolution of TRC in Figure 18 is significant and we hypothesize that is due to the heart disease HCM. For instance, if we study the evolution of TRC as defined for a healthy individual over years (without any noticeable heart conditions) as well as the retuned PLAR QRS-complex loop curves, we observe that there is not a progression or a behaviour that could make us think that the heart is getting worse (see Figure 19). The PLAR curves are close to one another and very similar for all the years in this case, which did not happen in the case of the HCM patient, and the evolution of the TRC looks like random noise around a baseline signal, which depends on the amplitude of the PLAR curve. Apart from that note how the TRC values are between 0.3 and 0.5, while in the HCM case these values fluctuated from -0.2 to 2.5 . For healthy individuals that would be the expected result over years, no abnormalities observed in the ECG follow-up over time.

Physiological Tracking

Going back to the HCM case, a commonly used parameter to check the functioning of the heart is the Ejection Fraction (EF), obtained from EEG test. This is the volumetric fraction of blood ejected from the heart, usually referring to the left ventricle, and its expression is: $EF(\%) = \frac{SV}{EDV} \cdot 100$, where SV refers to stroke volume (difference of volumes just before and after the contraction) and EDV to end-diastolic volume (volume of ventricle just before contraction). It is a measure of pumping efficiency. Healthy individuals have EF between 50% and 65%. Over the years EF is changing accordingly to the progression of the HCM condition. To see if there is statistical correlation between EF and TRC for each year, a commonly used technique is to normalize the features (divide by the maximum value achieved in all the years) and see if there is a linear correlation between them. The results of doing so are shown in Figure 20. We should mention that the value of EF normalized does not equal 1 in any represented point because it is considered that the maximum of EF takes place in 2010, and it is used this value to normalize the feature.

In this case the statistical correlation R^2 of the 2 magnitudes EF normalized and TRC normalized is 0.8368. Albeit there is not a clear linear relation, which may be due to human variability, the R^2 coefficient is high enough to justify further studies in this field, with several patients, while tuning certain model parameters (like T , the matrix M or modify the metric) for the optimum physiological tracking through TRC .

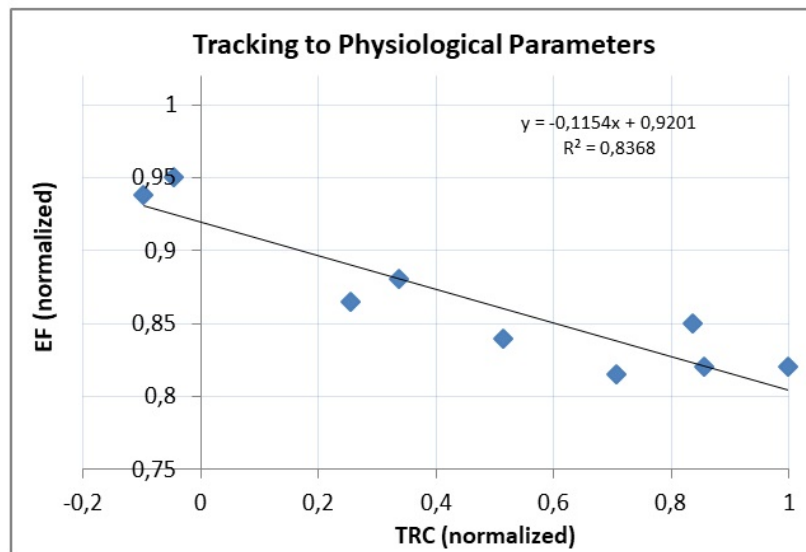


Figure 20: Normalized *EF* vs. normalized *TRC*

Physical Tracking

Another commonly used parameter to check hearts functioning is Posterior Wall Thickness (*PWT*). It refers to the average left ventricular wall thickness (measured in *mm*). In this case, this parameter is also of interest because as heart suffers from HCM, the ventricular wall is getting thicker. In average adults a *PWT* of 11 mm is normal, reaching 13 mm in certain athletes, while it being higher means hypertrophy of certain part of the heart. Again to see if there is statistical correlation between *PWT* and *TRC* an analogous procedure than the one in *Physiological Tracking* has been followed, resulting in Figure 21. See again that the R^2 coefficient could be higher; there is not an obvious linear relation between our *TRC* and the physical parameter studied. However, also the R^2 is high enough to justify further studies in this field and development of the new tool.

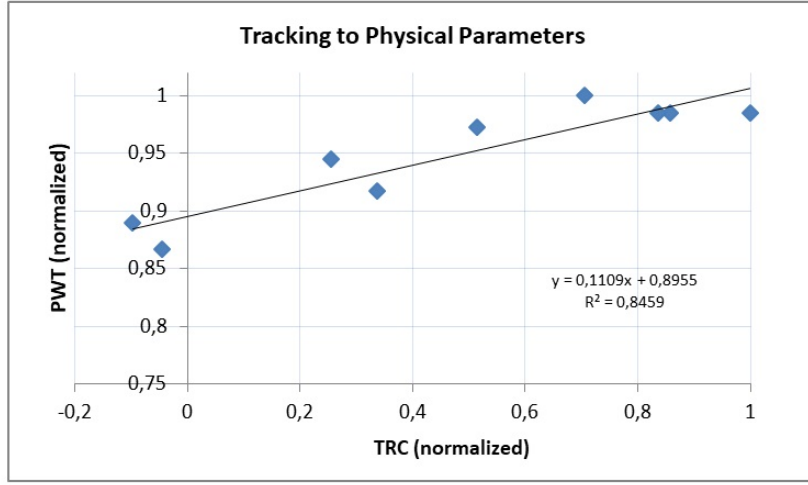


Figure 21: Normalized *PWT* vs. Normalized *TRC*

9. Future Work

There is still much work to do in order to understand the complexity of PLAR and how we can use it to detect other heart anomalies. It would be senseless to believe at this point that the same parameters for generating the PLAR curve would work for helping in diagnose of several different diseases. So now would be the time to test our visualization tool in some more patient ECGs with different conditions, while tuning the parameters in **black** used to obtain the PLAR curve:

$$S = S(L, \mathbf{M}, \vec{R}, \vec{V}, \mathbf{T})$$

However, even the \vec{R} and \vec{V} directions could be modified in the future in order to improve patient-specificity of our representation depending on the condition studied. Also, about following the physical and physiological state of the heart over time

as we tried to do in Section 8, other metrics or methods can be applied with the objective of tracking the heart progression using only electrical data. A possible future step could be to study the PLAR curves obtained over time through different curve parameters. During this project we tried to work with local curve parameters to try to find a better tracking metric, through the numerical computation of the local Frenet Trihedral, curvature and torsion for each point of the PLAR curve (wellknown concepts in the Differential Geometry field used to characterize the behaviour of a given $3D$ curve). Even though there is not a theoretical evidence that this might be a useful path, we hypothesized from Figure 17 that a metric able to better capture the evolution of the PLAR main QRS-complex loops over time would have to take more complex parameters of the PLAR curves compared into account.

We developped an algorithm that calculates the local Frenet Trihedral, curvature and torsion for each point of the QRS-complex loop of the PLAR from the whole PLAR curve, as shown in Figure 22. Note how it is evident here how the PLAR curve is composed of several discrete points, and that computing numerically the curve parameters may result at some point in not having a smooth evolution of the vectors of the Frenet Trihedral (due to errors or noise that are magnified in the computation of the derivatives of the curves). So for future work the signal acquisition and processing needs to be accurate.

The fact that we have not been able to find a consistent metric using these parameters does not mean that they are useless. Further investigation and development needs to be carried out to relate better the PLAR curve to heart state using

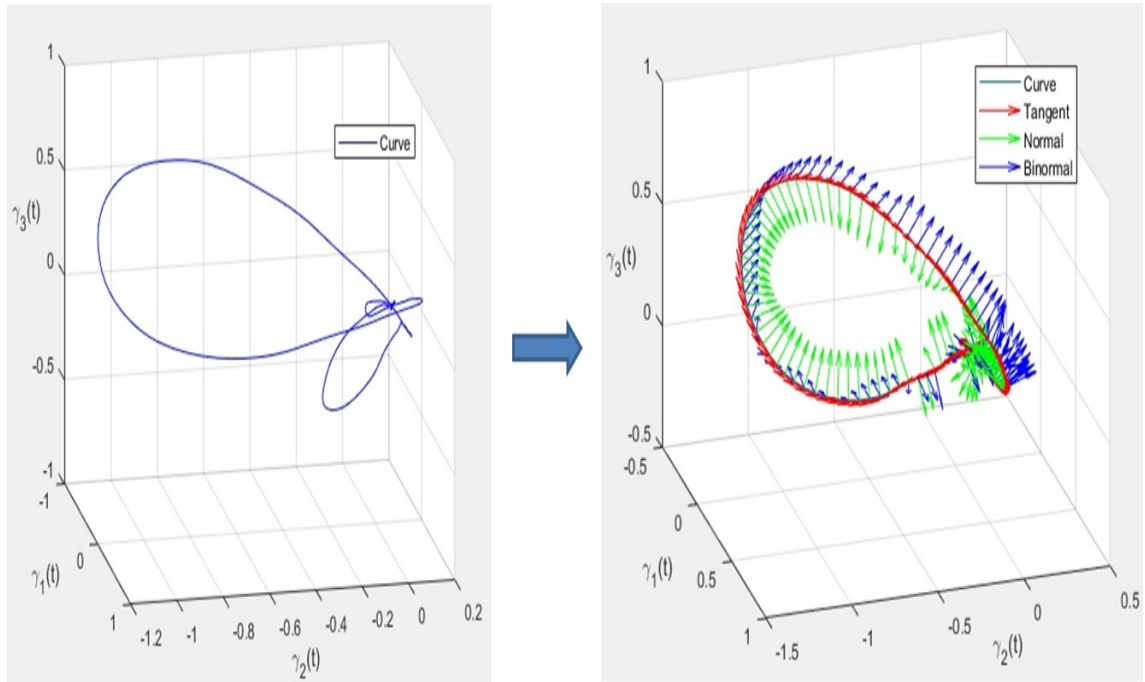


Figure 22: From a PLAR curve we are able to compute the Frenet Trihedral, curvature and torsion for each point of the QRS-complex loop.

the concepts introduced here or others.

10. Conclusions

This thesis constitutes the first development steps of the PLAR new visualization tool.

We have been able to define a three-dimensional space, using the power of applied mathematics, where subtle changes in ECGs are represented. As a first example, ST segment deviations can be characterized when they take place in different precordial leads. Also the defined space seems to be consistent to noise in the ECG data.

We also wanted our tool to be able to help tracking progressive heart diseases. We developed a method which a priori can distinguish between healthy and HCM affected individuals after studying their ECGs over years, and we obtained relatively good results when tracking physical and physiological properties evolution to the HCM process over time. Because of that many specialized and expensive tests performed only to know the current physical and physiological state of the heart, like echocardiograms or X-ray analysis, could be avoided, and this fact could mean a change in medical care for good.

Albeit we did not solve all the problems presented in the introduction, we proposed a line of investigation that may be of interest. This thesis finishes here, but it opens a door to long term projects and further work that could end up in what we initially expected from the tool, that is becoming a good support tool for medical assessment in the complex field of cardiology.

References

- [1] Guillem, M. S., A. V. Sahakian, and S. Syiryn. *Derivation of orthogonal leads from 12-lead ECG. Accuracy of a single transform for the derivation of atrial and ventricular waves*. Computers in Cardiology, 2006. IEEE, 2006.
- [2] F. Takens. *Dynamical Systems and Turbulence*. Warwick 1980, Lecture Notes in Mathematics 898, ed. D. Rand and L. S. Young, Berlin: Springer-Verlag, 1981.
- [3] S. M. Shekatkar, Y. Kotriwar, K. P. Harikrishnan, G. Ambika. *Detecting abnormality in heart dynamics from multifractal analysis of ECG signals*. Scientific reports, 2017, 7(1), 15127.
- [4] H. D. Abarbanel, T. W. Frison, L. S. Tsimring. *Obtaining order in a world of chaos. Time-Domain Analysis of Nonlinear and Chaotic Signals*. IEEE Signal Processing Magazine, 1998, 15(3), 59-65.
- [5] J. Pan, WJ. Tompkins. *A real-time QRS detection algorithm*. IEEE Trans Biomed Eng, vol. 32, pp. 230-6, 1985.
- [6] S. Hargittai *Savitzky-Golay least-squares polynomial filters in ECG signal processing*. Computers in Cardiology, 2005, Lyon, 2005, pp. 763-766.
- [7] A. Barja *A new approach to electrocardiography: patient specific spatio-temporal ECG*. Mathematics Bachelor's Degree Thesis, 2018.

- [8] The Figure 1 and Figure 2 were extracted from Wikipedia contributors. *Electrocardiography*. Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 29 May. 2018. Web. 4 Jun. 2018.
- [9] The Figure 3 was extracted from Wikipedia contributors. *Vectorcardiography*. Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 29 May. 2018. Web. 4 Jun. 2018.

List of Figures

1	Scheme of the main part of the standard sinus rythm	4
2	Directions of the 12 leads of standard ECG	5
3	Example of a VCG representation over time	8
4	Representation of how a 3D attractor of a single <i>ECG</i> signal lead is constructed for a given $T = 15ms$ time delay. We simply plot the evolution of the 3 red dots over time.	10
5	Example of PLAR from a 10 seconds ECG signal of a cardiologically healthy individual	19
6	The implementation of the algorithm allows a conversion from paper ECG records to digital data	21

7	Example of low frequency noise filtering using a 3rd degree polynomial fitting	23
8	Zoomed image of a lead I signal of a heartbeat for the signal unfiltered and filtered using S-Golay filter against high frequency noise	25
9	PLAR representation before and after smoothing the signal extracted from the ECGs	25
10	Lead V3 with fake ST deviations added before the high-frequency noise filtering	27
11	Heartbeat in V3 lead, raw signal and signal with added noise. Signal To Noise Ratio is 40 dB	28
12	PLARs of the same beat with 1 mm ST elevations in each precordial lead	29
13	Zoomed image of how deviations in lead V3 are observed in PLAR . .	30
14	Angle evolution for 2 different ECGs when a 1 mm increment in ST segment is incorporated , without and with added noise. Note how the effect of the delay T is clearly observed in the first 15 and last 15 milliseconds of the variations	30
15	Evolution over years of PLAR and VCG representation for a HCM patient	33
16	Retuned PLAR over years	34

17	Main QRS loops of retuned PLAR over years	35
18	TRC evolution over the disease period	38
19	Retuned PLARs and TRC evolution over years for a healthy individual.	39
20	Normalized EF vs. normalized TRC	41
21	Normalized PWT vs. Normalized TRC	42
22	From a PLAR curve we are able to compute the Frenet Trihedral, curvature and torsion for each point of the QRS-complex loop. . . .	44

A. APPENDIX A

A.1 Orthogonalization matrix

In the Section 3.2 we mention the existence of different M orthogonalization matrix used for going from a \mathbb{R}^8 space to \mathbb{R}^3 , that is for creating the VCG space. Through the thesis we use D and $QLSV$ matrix to reconstruct better different parts of the ECG, though $PLSV$ is also used in literature.

$$D = \begin{bmatrix} -0.172 & -0.074 & 0.122 & 0.231 & 0.239 & 0.194 & 0.156 & -0.010 \\ 0.057 & -0.019 & -0.106 & -0.022 & 0.040 & 0.048 & -0.227 & 0.887 \\ -0.228 & -0.310 & -0.245 & -0.063 & -0.054 & 0.108 & 0.021 & 0.102 \end{bmatrix}$$

$$QLSV = \begin{bmatrix} -0.147 & -0.058 & 0.037 & 0.139 & 0.232 & 0.226 & 0.199 & -0.018 \\ 0.023 & -0.085 & -0.003 & 0.033 & 0.060 & 0.104 & -0.146 & 0.503 \\ -0.184 & -0.163 & -0.190 & -0.119 & -0.023 & 0.043 & 0.085 & -0.130 \end{bmatrix}$$

$$PLSV = \begin{bmatrix} -0.266 & 0.027 & 0.065 & 0.131 & 0.203 & 0.220 & 0.370 & -0.154 \\ 0.088 & -0.088 & 0.003 & 0.042 & 0.047 & 0.067 & -0.131 & 0.717 \\ -0.319 & -0.198 & -0.167 & -0.099 & -0.009 & 0.060 & 0.184 & -0.114 \end{bmatrix}$$

B. APPENDIX B (Matlab code)

In this section we present the most fundamental code implemented in the project.

The different functions here allow us to:

- *ecg_to_vcg.m* : Calculate the resulting VCG of a given ECG data
- *findRvect.m* : Compute the vector R over which projecting the VCG data gives a maximum value
- *vcg_to_plar.m* : From a three-dimensional representation of electrical information in form of VCG, we follow the methodology explained to get to our PLAR curve
- *get_QRS_loop.m* : Isolate the main QRS loop from a PLAR curve
- *metric_curves.m* : Given 2 curves, it computes the metric that tells how much different are from each other. It is the implementation of the expression in (15).

Finally we present an example of how all these functions can be concatenated to, from an ECG signal in a *pdf* format, obtain a PLAR representation, and compute the difference between 2 different ECGs in the PLAR space.

Matlab Code

```
1 function [ Port ] = ecg_to_vcg( P, m )
2     % Transforms the 12 lead ECG in P to the 3 lead VCG in
   Port
3     % m = 1 for using Dower matrix;
4     % m = 2 for using PLSV matrix;
5     % m = 3 for using QLSV matrix;
6     if m == 1
7         %Dower
8         D = [ -0.172, -0.073, 0.122, 0.231, 0.239, 0.193,
9               0.156, -0.009;
10              0.057, -0.019, -0.106, -0.022, 0.040, 0.048,
11              -0.227, 0.886;
12              -0.228, -0.310, -0.245, -0.063, 0.054, 0.108,
13              0.021, 0.102];
14     elseif m == 2
15         %PLSV
16         D = [ -0.266, 0.027, 0.065, 0.131, 0.203, 0.220,
17               0.370, -0.154;
18               0.088, -0.088, 0.003, 0.042, 0.047, 0.067, -0.131,
19               0.717;
20               -0.319, -0.198, -0.167, -0.099, -0.009, 0.060,
21               0.184, -0.114];
22     elseif m == 3
23         %QLSV
24         D = [ -0.147, -0.058, 0.037, 0.139, 0.232, 0.226,
25               0.199, -0.018;
26               0.023, -0.085, -0.003, 0.033, 0.060, 0.104,
27               -0.146, 0.503;
28               -0.184, -0.163, -0.190, -0.119, -0.023, 0.043,
29               0.085, -0.130];
30     else
31         printf( 'Only 0, 1, 2 or 3' )
32     end
33     leads = [7,8,9,10,11,12,1,2];
34     Port = (D*P(:,leads)')';
35 end
```

```

1 function [ R ] = findRvect( VCG )
2     % Finds the R vector maximizing the value of the
3     % projecting VCG over that
4     % direction (note is the same as maximizing absolute
5     % value of the projection)
6     fobj = @(v) -max(project(VCG,v)); %We define the
7     % function we want to optimize
8     % These lines are to start the optimization of the
9     % function close enough to
10    % the optimum R vector
11    THETA = 0:0.01*pi:2*pi;
12    PHI = 0:0.01*pi:pi;
13    [THETA,PHI] = meshgrid(THETA,PHI);
14    X = cos(THETA).*sin(PHI);
15    Y = sin(THETA).*sin(PHI);
16    Z = cos(PHI);
17    F = zeros(size(Z,1),size(Z,2));
18    for i = 1:size(Z,1)
19        for j = 1:size(Z,2)
20            F(i,j) = fobj([X(i,j);Y(i,j);Z(i,j)]);
21        end
22    end
23    [M,J] = min(F);
24    [anything,i] = min(M);
25    j = J(i);
26    theta = THETA(1,i);
27    phi = PHI(j,1);
28    %Now we know the cilindric coordinates of the vector
29    %starting the
30    %iteration to find R
31    v0x = cos(theta).*sin(phi);
32    v0y = sin(theta).*sin(phi);
33    v0z = cos(phi);
34    v0 = [v0x;v0y;v0z];
35    R = fminunc(fobj,v0); % We use the Matlab non-linear
36    % optimization tool given a first value for iterating
37    R = R/norm(R); % We make sure the vector obtained is
38    % normalized
39 end

```

```

1 function [plar] = vcg_to_plar(VCG,T)
2     % gets the PLAR representation of the VCG signal, with
    delay T
3     % find R direction
4     R = findRvect(VCG);
5     R = R(:);
6     % find V direction
7     [ THETA, PHI, I ] = correlationsurface(VCG,R);
8     [M,J] = min(I);
9     [anything,i] = min(M);
10    j = J(i);
11    theta = THETA(1,i);
12    phi = PHI(j,1);
13    V = [cos(theta).*sin(phi), sin(theta).*sin(phi), cos(phi)
        )]';
14    e1 = VCG*R; % Project the VCG to R direction: First
    component of PLAR
15    e2 = VCG*V; % Project the VCG to V direction: Second
    component of PLAR without delay
16    gamma1 = e1(1:end); % Final First component
17    gamma2 = e2(T+1:end); % Final Second component
18    gamma3 = e1(T+1:end); % Final third component
19    k = min([length(gamma1),length(gamma2),length(gamma3)
        ]); % To make components have the same length
20    gamma1 = gamma1(1:k);
21    gamma2 = gamma2(1:k);
22    gamma3 = gamma3(1:k);
23    plar = [gamma1(:),gamma2(:),gamma3(:)];
24 end

```

```

1 function [QRS_loop] = get_QRS_loop(plar)
2     % Isolates the main QRS loop from the PLAR curve in
   plar
3     x = plar(:,1); y = plar(:,2); z = plar(:,3);
4     n = length(plar);
5     long = 0; % The longitude of the QRS_loop at each
   iteration
6     eps = 0.008; % We start iterating with epsilon really
   small
7     deps = 0.002; % Parameter we will increment eps at
   each iteration
8     perc = 0.08; % To ensure the loop obtained is long
   enough; We set 8% of the length of all PLAR;
9     % We compute the distances between each point and the
   subsequent of the
   % plar
10    p0 = [x(1);y(1);z(1)];
11    vect_dist = zeros(n-1,1);
12    for i=2:n
13        p1 = [x(i);y(i);z(i)];
14        vect_dist(i-1) = norm(p1-p0);
15        p0 = p1;
16    end
17    vect_dist = vect_dist(:);
18    vect_dist = sgolayfilt(vect_dist,3,17); % We smooth
   the distances vector to avoid noisy errors
19    while (long<perc*n)
20        vect_ind = [];
21        entrat = 0; % Parameter that controls when to stop
   after getting all QRS-loop
22        eps = eps+deps;
23        for i=2:n
24            if(vect_dist(i-1)>eps)
25                entrat = 1;
26                vect_ind = [vect_ind; i]; % We add the
   indexes that fulfill the stablished
   criteria
27            else
28                if(entrat==1)
29                    break
30                end
31            end
32        end
33    end
34    long = length(vect_ind);
35    QRS_loop = [x(vect_ind),y(vect_ind),z(vect_ind)];
36 end
37 end

```

```

1 function [ my_metr ] = metric_curves( Si, Sp, term )
2     % Computes the metric in my_metr of Si and Sp, as
3     % established in 'term'
4     % term = 1 for calculating the 'distance' through
5     % metric 1 (used in thesis)
6     % term = 2 for calculating the 'distance' through
7     % metric 2
8     % term = 3 for calculating the 'distance' through
9     % metric 3
10    n = length(Si); m = length(Sp);
11    N = min(n,m);
12    M = max(n,m);
13    if (N==n)
14        c1 = Si; c2 = Sp;
15    else
16        c1 = Sp; c2 = Si;
17    end
18    my_metr = 0;
19    %c1 is the shortest curve
20    %c2 is the longest curve
21    %We compute the sum as established in metric parameter
22    %'term'
23    for i=1:N
24        j = i*M/N; % We approximate the location of the
25        % longest
26        % curve related to the index i of the
27        % shortest curve
28        j1 = floor(j); j2 = ceil(j);
29        if (term == 1)
30            xyz_here = (c2(j1,:) + c2(j2,:))/2;
31            my_metr = my_metr + norm(xyz_here - c1(i,:));
32        elseif (term == 2)
33            xyz_here = (c2(j1,:) + c2(j2,:))/2;
34            xyz_here = xyz_here([1 3]);
35            my_metr = my_metr + norm(xyz_here - c1(i,[1
36            3])));
37        elseif (term == 3)
38            xyz_here = (c2(j1,:) + c2(j2,:))/2;
39            my_metr = my_metr + xyz_here*(c1(i,:)');
40        else
41            fprintf('Wrong term number')
42        end
43    end
44    my_metr = my_metr/N;
45 end

```

All the functions can be concatenated in a simple *Matlab Script* . Note how the function *extract_signals* uses the algorithm described in Section 5 to get the data from an ECG in *pdf* format to numerical.

```

1 %% Compute and represent our PLAR curve of a pdf paper
2 ECG = extract_signals('ECG_example.pdf');
3 % Here we obtain a 12 x n matrix in ECG (12 leads of
  normal ECG)
4
5 matr = 1;
6 VCG = ecg_to_vcg(ECG,matr);
7 T = 15; % Time delay , in time samples
8 PLAR = vcg_to_plar(VCG, T);
9
10 figure()
11 plot3(PLAR(:,1), PLAR(:,2), PLAR(:,3)) % To show our PLAR
  representation
12
13 %% Compute the difference of 2 different PLAR main loops
  through an established metric
14 ECG2 = extract_signals('ECG_example2.pdf');
15 VCG2 = ecg_to_vcg(ECG2,matr);
16 PLAR2 = vcg_to_plar(VCG2,T);
17
18 S1 = get_QRS_loop(PLAR);
19 S2 = get_QRS_loop(PLAR2);
20 term = 1;
21 my_metric = metric_curves(PLAR, PLAR2,term)

```

In addition we include the code implemented to compute, given a $3D$ curve, the Frenet concepts (orthogonal vectors at each point: Tangent, Normal and Binormal, along with curvature and torsion) introduced in *Future Work* Section.

```

1 function [T,N,B,k,t] = get_frenet_trihedral(x,y,z)
2 % Given the x, y, z components of the points in a curve,
   this function
3 % calculates the Tangent, Normal and Binormal vectors
   orthonormal base at
4 % each point, along with the curvature k and torsion t
5 % First we want all components as column vectors
6 x = x(:);
7 y = y(:);
8 z = z(:);
9 % We calculate the derivatives
10 dx = gradient(x);
11 dy = gradient(y);
12 dz = gradient(z);
13 dr = [dx dy dz];
14 % We calculate the 2nd derivatives
15 ddx = gradient(dx);
16 ddy = gradient(dy);
17 ddz = gradient(dz);
18 ddr = [ddx ddy ddz];
19 % The tangent vectors will be (each coordinate refers to
   the respective point in the curve)
20 T = dr./mag(dr,3);
21 % From T we get also its derivatives
22 dTx = gradient(T(:,1));
23 dTy = gradient(T(:,2));
24 dTz = gradient(T(:,3));
25 dT = [dTx dTy dTz];
26 % We calculate the normal vector (see mag function at the
   end)
27 N = dT./mag(dT,3);
28 % And the binormal vector
29 B = cross(T,N);
30 % Now the curvature
31 k = mag(cross(dr,ddr),1)/((mag(dr,1)).^3);
32 % We compute the 3rd derivatives
33 dddx = gradient(ddx);
34 dddy = gradient(ddy);

```

```

35 dddz = gradient(ddz);
36 dddr = [dddx dddy dddz];
37 % And finally the torsion
38 t = vdot(cross(dr, ddr), dddr) ./ mag(cross(dr, ddr),1)
    .^2;
39
40 % Subfunctions used:
41 function V = vdot(A, B),
42 V=zeros(size(A,1),1);
43 for i=1:size(A,1)
44 V(i) = dot(A(i,:), B(i,:));
45 end
46
47 function N = mag(T,n),
48 N = sum(abs(T).^2,2).^(1/2);
49 d = find(N==0);
50 N(d) = eps*ones(size(d));
51 N = N(:,ones(n,1));

```