Reconstruction of Premature Atrial Contraction and Premature Ventricular Contraction on ECG Traces by Applying PLA as Segmentation Process

Joel A. Gonzalez, Roberto L. Avitia, Nellyzeth Flores, Miguel E. Bravo, and Marco A. Reyna
Department of Bioengineering and Environmental Health, Autonomous University of Baja California, Mexico
Email: {ravitia, ravitia, mbravo, mreyna}@uabc.edu.mx, {antonio_30-91, nellyzeth}@hotmail.com

Lucio A. Cetto
Computational Biology Department, MathworksInc, Natick MA, United States
Email: lucio.cetto@mathworks.com

***Abstract*—Premature Atrial Contraction (PAC) and Premature Ventricular Contraction (PVC) are two types of ECG arrhythmias that are identified for presenting anomalies in the normal cardiac rhythm, generating alterations in the heart rate that disrupts his mechanic and electric heart activity due to these delayed contractions (premature). Also when a delayed contraction occurs (i.e. PVC or PAC), it generates alterations in heart rate that can produce ventricular tachycardia and eventually sudden cardiac arrest. Piece Wise Linear Algorithm (PLA) is a widely used segmentation process that represents each signal as linear segments with different scopes. In this work we are presenting an evaluation of PLA and the optimal number of segments when this type of segmentation is applied on normal ECG traces and when these ECG traces present PAC or PVC. We established that optimal number of segments is related with the coefficient of determination (R^2), before and after the reconstruction of an ECG trace. Results show that number of segments taken on normal ECG traces differs radically compared when this method was applied on traces with PVC and PAC events.**

***Index Terms*—premature atrial contraction, premature ventricular contraction, piecewise linear algorithm**

I. Introduction

Electrocardiogram or simply ECG trace is still an important pre-diagnosis tool that is continuously studied. Some researchers have developed many techniques and methods to interpret ECG as normal or abnormal based on professional experience [1]. To classify an ECG as normal or abnormal, professional analyzes carefully characteristic waveforms called P wave, QRS complex, ST segment and T wave. A Normal Sinus Rhythm is reflective of a normally functioning conduction system, and an experimented physician can determine it looking the ECG trace and some parameters such as Rhythm, Rate, P wave configuration, PR interval and QRS duration and configuration [2]. Thus, a Normal Sinus Rhythm has a heart rate of 60-100 bpm (beats per minute), a regular P-P interval, a regular R-R interval, there is always a P wave and the ratio between P and QRS is always 1, which means there is always a P wave associated with a QRS, and the QRS duration is less than 100 msec.

An arrhythmia is defined as a non-normal rhythm and some of them can produce ventricular tachycardia and eventually a sudden cardiac arrest [3].

Two types of arrhythmias are the Premature Ventricular Contractions (PVCs) and the Premature Atrial Contractions (PACs). PVCs are caused by an ectopic cardiac pacemaker located in the ventricle. On the ECG they are characterized by premature and bizarrely shaped QRS complexes, usually wider than 120 msec. These complexes are not preceded by a P wave, and the T wave is usually larger, and its direction is opposite the major deflection of the QRS. In a PVC ECG record there is irregular P-P interval and an R-R interval, and there is no a P-QRS ratio associated.

On the other hand PACs are amongst the most common forms of arrhythmias. They are due to the premature discharge of an electrical impulse in the atrium, causing a premature contraction. On the ECG, they are characterized by an abnormally shaped P wave [4]. In a ECG record with PAC the P-P interval and the R-R interval are irregular too, but the P-QRS ratio 1:1 associated and the QRS duration is considered as normal because is always less than 100 msec.

These two events can be seen in healthy people and people with cardiac disorders, normally asymptomatic. The main reason why they are very dangerous is because of their characteristic to be asymptomatic and normally silently can increase the risk, and eventually may lead to a sudden cardiac arrest.

Signal segmentation process is a very common technique applied in many biomedical digital systems, as wells as analysis of long-term ECG traces [5], [6]. Every application where signal processing is been used, a segmentation process is included. An example is the ECG segmentation process, where the ECG trace is been transmitted by a different kind of digital protocols [7]. One of the most applied segmentation process is the Piecewise

Manuscript received April 5, 2015; revised August 31, 2015.

doi: 10.18178/joig.3.2.112-116

Linear Algorithm, where signals are represented using segments of lines with different scopes and tendency. It is a technique of importance not only in many digital computer applications, but is widely employed in analog computer embodiments for function representation.

In this work we are presenting an evaluation of PLA and the optimal number of segments taken in the ECG reconstruction, when this type of segmentation is applied on normal ECG traces, and when these ECG traces present PAC or PVC. We established that optimal number of segments is related with the coefficient of determination (R^2), before and after the reconstruction of an ECG trace. Results show that number of segments taken on normal ECG traces differs radically compared when this method was applied on traces with PVC and PAC events.

II. Materials and Methods

A. Signals and Filters

First, we got 30 ECG records with 10 presenting PAC, 10 with PVC and 10 Healthy Control from the standard database PTB ECG Database (or simply PTBED) [8]. The PTBDED is part of Physionet databases that offers free access via the web to large collections of recorded physiologic signals. We selected 10 ECG traces considered as Normal ECG, or simply classified by them as Normal Sinus Rhythm. Then we selected 10 ECG records where we appreciated PAC anomalies and 10 records considered with the presence of PVC anomalies (Fig. 1).

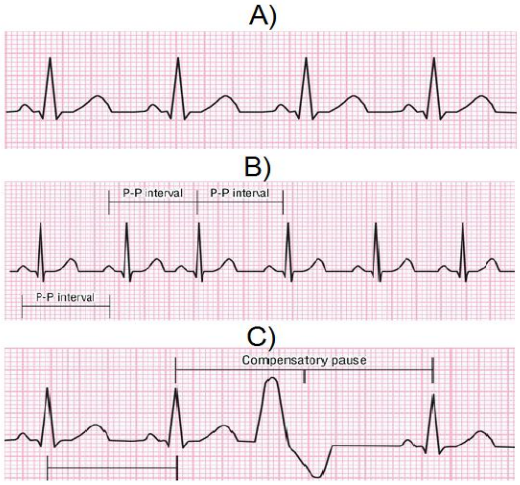

Figure 1. ECG waveforms for A) normal sinus rhythm (healthy control ECG record), B) premature atrial contraction and C) premature ventricular contraction.

After we selected ECG records, each of them were loaded in MATLAB software. Since the signals were downloaded in *.dat* extension, it was necessary to extract the ECG leads with the *fopen*, *fread* and *fclose* commands. Then we applied a low-pass filter using the coefficient values as described in (1) and a moving averaged filter in order to reduce the high frequency noise of the records and not to affect the ECG bandwidth.

$$b1 = [0.24523\ 0.24523]; \quad (1)$$
$$a1 = [1\ -0.509525];$$
$$filteredECG = filter(b1, a1, ECGrecord)$$

B. Isoelectric Baseline Adjustment

Isoelectric baseline is one of the most important problems on ECG records. In order to have our ECG trace centered and eliminate all kind on signal fluctuations, we adjusted the baseline applying the *msbackadj* function. As we can see in Fig. 2, all ECG records were very long for the *msbackadj* command, and in order to determine the number of intensity points for each record, we divided every ECG record (about 10000 heartbeats) in approximately 10 heartbeats each, and then we applied the *msbackadj* with a variable window from 100 to 900 samples per window depending on Healthy Controls, PACs or PVCs records, as we can see in Table I-Table III.

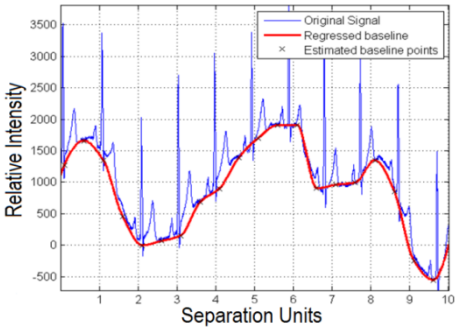

Figure 2. An ECG record (blue) that presents a non-zero isoelectric baseline (red) must be adjusted by applying *msbackadj* function. Then the isoelectric tendency should be find it using different intensity points.

TABLE I. Different Window Size Applied to ECG Records

ECG record	Window Size for Healthy Controls
1	200
2	500
3	800
4	600
5	400
6	800
7	500
8	100
9	600
10	700

TABLE II. Different Window Size Applied to ECG Records

ECG record	Window Size for Records with PACs
1	300
2	100
3	300
4	200
5	100
6	100
7	100
8	300
9	300
10	200

TABLE III. Different Window Size Applied to ECG Records

ECG record	Window Size for Record with PVCs
1	600
2	700
3	800
4	600
5	600
6	900
7	600
8	900
9	700
10	600

C. Heartbeats Extraction from an ECG Record

In order to calculate the beginning and the ending of the QRS complex (QRSon and QRSoff respectively), a slightly modified of classical Pan and Tompkins algorithm was applied. This algorithm involves two filters that act as a band-pass filter (5-15 Hz) and a differentiation process that enhance the QRS complex and provides information of the slope. After this differentiation process, the QRS is squared to heighten much more the higher frequencies in the complete ECG trace. A moving integration window gives information to detect the QRS_{on} and QRS_{off}. Once we found QRS_{on} and QRS_{off} the difference between them was defined as QRS complex duration in ms or simply QRS [9]. After the QRS complex detection, we chose 200 samples and 150 samples after and before the R-peak, respectively. In this manner we had about 300 heartbeats per ECG record.

D. PLA and Reconstruction of ECG Records

After the signal was filtered, a little modified algorithm of PLA was applied. In this case we applied PLA on the original heartbeats of the ECG traces defined as *Y*, by increasing the number of samples in the line segment *Seg* and by increasing the number of samples *D* in the original ECG record, in order to reconstruct the lines or segments per heartbeat. The number of segments decreases when more samples *D* in the original record *Y* are added. On the other hand, when the number of samples is reduced, the number of segments is increased. Knowing this, we go through 1 to 60 samples for each line during the segmentation process. We defined the points (X_1, Y_1) and (X_2, Y_2) as the start and ending points of one segment, then they were used in the linear equation 2 to obtain the slope *M* and a constant *B*.

$$M = (Y_2 - Y_1)/(X_2 - X_1) \quad (2)$$

And:

$$B = (X_1.Y_2) - (X_2.Y_1)/(X_1 - X_2)$$

Once we got these two characteristics values, the linear equation was utilized to reconstruct the segmented record *Y*, defined as Y_s. As we can appreciate in (3), this equation contains the number of segments *(Seg)*.

$$Y_S = \sum_{s=1}^{n} M_s \,.\, X_i + B_s \quad (3)$$

Afterwards the last point (X_2, Y_2) becomes (X_1, Y_1) and the number of samples (*D*) was added in a new variable, in order to have the length of the next segment.

If:

$$Y(X_1, Y_1) = (X_2, Y_2) \quad (4)$$

Then $Y(X_2, Y_2) = (X_2 + D, Y_2 + D)$.

This is continuously applied until the last sample point of the record Y is reached. With the parameters B, M and segmented record Ys, now it is possible to reconstruct the original record Y, as we can see in (5), where ln represents the length of the line n.

$$Seg_n = (M_n.l_n) + B_n \quad (5)$$

And $Y_s = \sum_1^n Seg_n$.

E. Evaluation of Reconstruction

In order to evaluate the reconstruction and the number of segments taken in consideration on the PLA segmentation process, we used the coefficient of determination R^2 as the ratio and the relationship with the original record *Y* and the reconstructed ECG record Y_s. This measured value is explained by the least-squares regression line, and it is a value between 0 and 1, where:

If $R^2 = 0$ the matching between *Y* and Y_s it is null.

If $R^2 = 1$ means that two records are exactly alike.

We calculated this coefficient of determination after the reconstruction, using original records such as healthy controls, PACs records and PVCs records in order to determine the best number of segments to be used for this kind of arrhythmias.

The sequence of functions in MATLAB to determine the coefficient of determination is as follow:

$$p = polifit(Y, Y_s, 1)$$

$$yfit = polyval(p, Y)$$

$$Yresid = Y_s - yfit$$

$$Y_sResid = sum(Yresid^2)$$

$$Y_sTotal = (length(Y_s) - 1) * var(Y_s)$$

And:

$$R^2 = 1 - (Y_sResid/Y_sTotal)$$

III. Results

As we described before in algorithm implementation, all ECG traces were smoothed, adjusted to an isoelectric baseline and normalized to 1 as the highest value. After we eliminated this moving isoelectric baseline, we applied the PLA as segmentation process as we can see in Fig. 3. The number of segments per heartbeat it is an averaged of the total segments divided by the number of heartbeats in the ECG record.

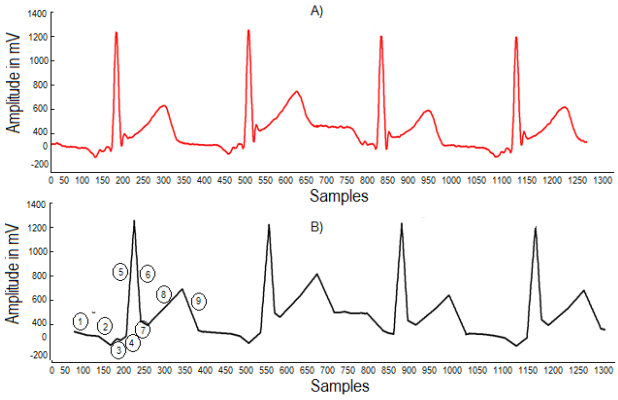

Figure 3. PLA applied to a healthy control ECG record (A), the number of segments necessary is about 9 segments per heartbeat (B) and a R^2 higher than 0.96.

On the other hand Fig. 4 represents the reconstruction of 10 Healthy Control ECG records taken from the standard database using PLA as segmentation process. The coefficient of determination depends on the number of segments per heartbeat.

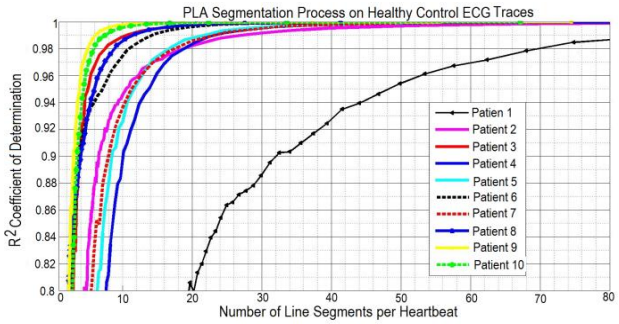

Figure 4. PLA applied to 10 healthy control ECG traces. A good reconstruction could be appreciated on less than 15 segments per heartbeat (mean).

We labeled each segmented trace as patient trace number. Thus, we can see that the number of segments for patient 1 (Normal ECG trace 1 or Record 1), the coefficient of determination reached an acceptable value (R^2=0.98) about 70 segments per heartbeat in the segmentation process. On the other hand, the patient trace 2 to 10 reached an acceptable value using less than 20 segments per heartbeat.

Now, as we can see in Fig. 5, we analyzed 10 ECG records with PACs and the number of segments to reached an acceptable coefficient of determination for ECG trace 1 (labeled as patient 1) was40 segments, for patient 2, patient 3 patient 4, patient 7 and patient 8 about 25 segments, for patient 5 more than 80 segments, patient 6 about 40 and for patient 9 and 10 about 40 and 60 segments respectively.

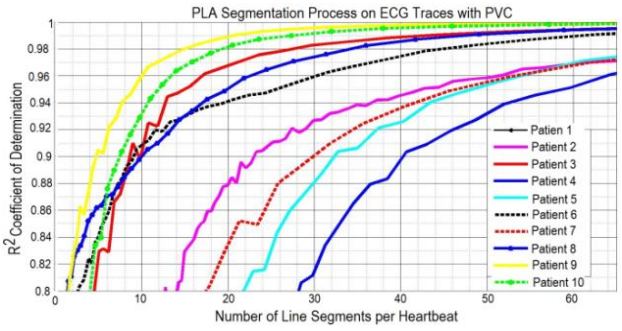

Figure 5. PLA applied to 10 ECG traces with PACs. The number of segments to reach an acceptable R2 is at least 30 segments per heartbeat (mean).

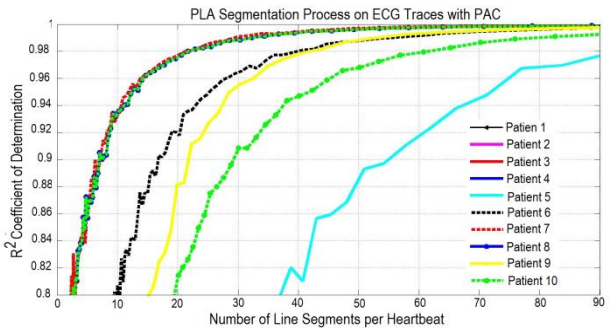

Figure 6. PLA applied to 10 ECG traces with PVCs. A good reconstruction is appreciated on 38 segments per heartbeat (mean).

In Fig. 6, we are representing PLA as segmentation process with ECG traces with PVCs. Results show that the number of segments to reach an acceptable value of R^2 increased significantly. The most of the ECG traces with PVCs required 45 segments for patient 1, more than 50 segments for patient 2, 5 and 7, less than 30 segments for patient 3, more than 60 for patient 4, more than 40 for patient 6, about 25 for patient 8 and in the best cases were 17 and 19 segments for patient 9 and 10, respectively.

IV. Conclusion

In this work we presented one of the most applied segmentation process on signals, but in this case on ECG traces. Before the application of PLA process, it was necessary to applied an isoelectric reducer due to records had high levels of voltage added with the signal of interest (ECG record). After that, two different types of arrhythmias were evaluated when they appeared on ECG traces for 10 different patients. Results show that PAC and PVC present a mean of 30% and 60% respectively more segments per heartbeat in the reconstruction of the ECG traces (i.e. when we considered a R^2 of 0.98 as an acceptable value).

This significant difference is due to the period of the heartbeat that is affected in the presence of PAC and PVC. More specifically, the changes in the ventricle electrical activity (QRS complex) appeared suddenly between two normal R-R intervals. With these results, we can conclude that the segmentation process applied on ECG traces should be evaluated before his application with this very important type of arrhythmias. For example on digital communication systems where an ECG trace is going to be transmitted, cardiac disorders such as PAC and PVC will not appear if the segmentation process is as simple as that applied on Healthy Control ECG traces.

Acknowledgment

This work was supported by Seventeenth Internal Research Call for Research Projects from Autonomous University of Baja California under Grant 105/6/C/CC/17.

References

[1] F. G. Yanowitz, *Introduction to ECG Interpretation V8.0*, LDS Hospital & Intermountain Medical Center, July 2012, pp. 1-83.

[2] M. B. Conover, *Understanding Electrocardiography*, 8th ed., St Louis, MO: Mosby, 2003.

[3] R. L. Avitia, M. A. Reyna, M. E. Bravo-Zanoguera, and L. A. Cetto. "QRS complex duration enhancement as ventricular late potential indicator bysignal-averaged ECG using time-amplitude alignments," *Biomedical Engineering-Biomedizinische Technik*, vol. 58, no. 2, pp. 179-186, 2013.

[4] D. Conen, *et al.*, "Premature atrial contractions in the general population: frequency and risk factors," *Circulation*, vol. 126, pp. 2302-2308, 2012.

[5] H. Azami, K. Mohammadi, and H. Hassanpour, "An improved signal segmentation method using genetic algorithm," *International Journal of Computer Applications*, vol. 29, no. 8, pp. 5-9, 2011.

[6] H. Krim and D. H. Brooks, "Feature-Based segmentation of ECG signal," in *Proc. IEEE SP International Symposium on Time-Frequency and Time-Scale Analysis*, April 1996.

[7] Y. Ungson, M. AReyna, M. E. Bravo-Zanogueraand, and R. L. Avitia, "Validation of a wireless ambulatory ECG," in *Proc. Pan American Health Care Exchanges Conference*, April 2015, pp. 129-134.

[8] A. L Goldberger, *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, pp. E215-E220, 2000.
[9] J. Pan and W. J. Tompkins, "A real time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 230-235, 1985.

Joel A. Gonzalez received a B.S. degree in bioengineering and biomedical instrumentation specialization from the Autonomous University of Baja California (UABC), Mexico in 2014. He won an award on scientific innovation in 2014 before finishing his major. His past employment includes assistant researcher at Benemérita Universidad Autonoma de Puebla (BUAP).
Mr. Gonzalez is currently employed in PepsiCo. as Quality Technician. His interests include development of all kind of instrumentation and especially biomedical instrumentation.

Roberto L. Avitia received the B.S degree in biomedical electronic engineering from National Technological of Mexico in 2004, the M.S. degree in bioelectronics from the Center for Research and Advanced Studies of the National Poly technic Institute, Mexico in 2006 and the Ph.D. in biomedical engineering from Autonomous University of Baja California (UABC), Mexico in 2013.His past employment includes industrial at Hirata Engineering and hospital experience at National Institute of Cardiology.
He joined the Department of Bioengineering and Environmental Health at UABC in 2007 by applying dynamic programming and machine learning algorithms in high resolution electrocardiography (HRECG). He was creator of the bioengineering undergraduate program at UABC.
Dr. L. Avitia is currently full professor of the bioengineering program of the UABC and member of the National System of Researchers. His research interests include nonlinear modeling of HRECGs, development of medical instrumentation, and pattern recognition in biomedical images.

Nellyzeth Flores is currently on the last year to complete her B.S degree in bioengineering and biomedical instrumentation specialization from Autonomous University of Baja California (UABC), Mexico. Her past employment includes assistant researcher in University of Guanajuato, and an internship as biomedical engineer at hospital ISSSTECALI Mexico. Her research interests include development of medical instrumentation, biomedical signals and biomedical images.

Lucio A. Cetto received the B.S. degree in electrical engineering from Autonomous University of Baja California (UABC), Mexico, in 1992. He received the M.S. degree in electronics and telecommunications from Center Scientific Research and Higher Education of Ensenada, Mexico, in 19994, and the Ph.D. degree in electrical engineering from Northeastern University, U.S.A, in 2007. His past employment includes lecturer at CETYS University from 1994 to 1996. Dr. Cetto currently is senior engineer at The Math Works, with more than 12 years of experience as software developer. His interests include Signal Processing, Machine Learning, and Algorithms for Computational Biology.

Miguel E. Bravo received his B.S degree in electromechanical engineering in 1983 from the Autonomous University of Baja California. He received his M.S. degree in electrical engineering from San Diego State University in 1991. And He received a Ph.D. in bioengineering by the University of California, San Diego, in 2001. His past employment includes technical support engineer in a computer manufacturing company, research assistant in the Bioengineering Department at UCSD and professor at the Faculty of Engineering UABC since 1986.
Dr. Bravo has received multiple recognition such as "Space Act Award", awarded by the US Agency Aeronautics and Space Administration (NASA), April 19, 2007, related to the research work on multiparallel three dimensional optical microscopy and the CONACyT recognition for achievements in conducting original research (2004-2010).
He is one of the creators of the undergraduate program of Bioengineering that started in 2009, and now is responsible for leading efforts to obtain approval from the accreditation board. His research interest and teaching relates to digital signal processing, biomedical instrumentation and automated vision systems.

Marco A. Reyna received the B.S. degree in electrical engineering from Autonomous University of Baja California (UABC)-Mexicali, Mexico, in 1990. He Received the M.S. degree in biomedical engineering from Metropolitan Autonomous University-Mexico D.F., in 1994, and the Ph.D. degree in bioengineering from the Polytechnic University of Catalu ña, Barcelona, Spain, in 1998. From 1988 to 1989, he was an electrical technician at the International Circuits enterprise in Mexicali, B.C., Mexico, where he performed the electrical test of different electronic products.
In 1989 he was hired by the UABC, in Baja California, Mexico, where he has taught undergraduate and graduate courses, managed several research projects, published scientific articles in various journals, published book chapters, and a book, all related to bioengineering and environmental health. He was the creator of the bioengineering and environmental health laboratory and one of the creators of the bioengineering undergraduate degree program.
Dr. Reyna currently is full researcher at UABC, leader of the bioengineering and environmental health research group and member of the National System of Researchers.