COMPLEXITY MEASURE REVISITED: A NEW ALGORITHM FOR CLASSIFYING CARDIAC ARRHYTHMIAS

U. Ayesta, L. Serrano, I. Romero
Department of Electrical and Electronic Engineering, Public University of Navarra
Campus de Arrosadía, 31006 Pamplona, Spain
lserrano@unavarra.es

Abstract- **This paper deals to the suitability of Complexity Measure of ECG signals for the classification of cardiac arrhythmia's. Applying this algorithm to data from the MIT-BIH database a very poor performance, especially for SR signals, and an overall error rate of 20% is obtained. In this study a novel measure, named SPDR, Sample Percentage in the Dynamic Range, to be used in combination with the Complexity Measure algorithm, is proposed. Using this novel proposal the result of the classification is improved decreasing the overall error rate until to 9%. The algorithm has been implemented in a computer using LabView and C++ software.**
Keywords **– Biomedical Signal Processing, Arrhythmia Classification.**

I. INTRODUCTION

Reduction of mortality from Ventricular Fibrillation, VF, Ventricular Tachycardia, VT, and others cardiac causes depends mainly on rapid detection and accurate classification of these arrhythmia's. Conventional algorithms used in both surface ECG monitors and in implantable cardiovertor/defibrillators rely on simple heart rate for detection-classification even though the rate range of VF overlaps with that of VF. Due to this big efforts are devoted to looking for a reliable and robust algorithm to classify these signal in real time [1].

Regarding Time domain analysis, few proposals and algorithm have been published deal to ECG classification mainly because of the poor results obtained. However, one of the advantages of these algorithms is its simple and computationally efficient implementation providing a faster classification than the algorithms based on Frequency, Time-Frequency analysis or on other more exotic transforms as Wavelet [1].

The Complexity Measure, CM, proposed by Lempel and Ziv, is one of the most interesting time analysis algorithms used for classifying ECG [2]. X-S Zhang states in [3] to have obtained a correct classification using it. In the CM algorithm described in [3], for a specific window length, the algorithm generates a 0-1 string comparing the raw electrocardiogram data to a selected suitable threshold. The Complexity Measure is obtained from the 0-1 string.

Using the normal sinus rhythm and malignant ventricular arrhythmias of the MIT-BIH database with the CM algorithm we got an overall error rate of 22%. The worst result was with SR signals, where approximately only 33% of the signals were classified correctly. Yet the result for VT and VF was satisfactory.

The CM algorithm was implemented with more levels, in order to keep more information than with two symbols, but the final result still remained unsatisfactory.

Due to the above results of the CM algorithm we studied the possibility to split the process up in two steps. In the first one, using the novel measure Sample Percentage in the Dynamic Range, SPDR, the rhythm was classified like SR or VT/VF. In the second one and using the CM algorithm the signal was classified either as VT or VF.

Next section describes the material and the methods, i.e., the Complexity Measure, CM, in combination with the Samples Percentage in the Dynamic Range, SPDR, used for classifying ECG. Section III shows the results obtained and describes the computationally efficient implementation. In Section IV discussion and open issues are addressed. Finally, in Section V conclusions and future work are explained.

II. MATERIAL AND METHODS

A. ECG Data

A set of ECG records obtained from the MIT-BIH database is used for testing the proposed methods. In this particular case, 21 SR records from Normal Sinus Rhythm subset of the MIT-BIH, and 56 VT and 61 VF signals obtained from the Malignant Arrhythmia subset of MIT-BIH, were selected and used for both development and evaluation stages. For the verification of the SPDR method, an additional set of 20 SR signals from the arrhythmia database subset of MIT-BIH were selected.

All the data segments were 10sec length and the sampling frequency was set to 250Hz. In order to remove baseline drift and high frequency noise, a band-pass filter with band-pass 0,5-20 Hz. was applied [3].

B. Complexity Measure.

The Complexity Measure, proposed by Lempel and Ziv [2], was implemented as it is explained in [3]. We evaluate the complexity of the ECG sequence from the point of view of a simple learning machine which, as it scans a given n-digit sequence, $S=s_1\ s_2 ...\ s_n$, from left to right, adds a new word to its memory every time it discovers a sub-string not previously encountered. The size of the compiled vocabulary and the rate at which new words are encountered along S, are the basic ingredients in the proposed evaluation of the complexity of S.

Studying the Complexity Measure obtained with VT and VF signals, a threshold was established, $C_{VT/VF}=0.18$. From now on, we will measure the complexity of any signal making a comparison with this threshold so that we will decide that the signal is VT whether $C(ECG) < C_{VT/VF}$, otherwise, the signal is VF.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Complexity Measure Revisited: A New Algorithm for Classifying Cardiac Arrhythmias	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department of Electrical and Electronic Engineering Public University of Navarra Campus de Arrosadia, 31006 Pamplona, Spain	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
3

C. SPDR.

Because of the morphological differences between SR signals and the two other rhythms, we propose a novel measure named, Sample Percentage in the Dynamic Range, SPDR. Given a set of ECG segments, we have observed that measuring the percentage of samples in the range *Peak_Positive_Value(PPV) - 10% PPV*, we are able to separate the rhythms into two groups, SR signals and VT-VF records. Lot of proofs was made with different percentages but the range *PPV-10%PPV* was chosen because it shows to have the most appropriate characteristic and the biggest capability to classify the signals.

Because SPDR is a novel measure, we checked its result with an additional set of 20 SR ECG records obtained, as we mentioned in section II.A, from the arrhythmia database.

In the same way as the Complexity Measure algorithm a threshold value, $m_{SPRD,}$ was defined. The values of this new parameter obtained for SR signals, ranged from 2% to 30%, and for VT/VF signals from 33% to 60%. With these results, we established a threshold, m_{SPRD}=30%, that we will use to classify a signal either as SR when $SPDR(ECG) < m_{SPRD}$ or VT/VF in case of $SPDR(ECG) > m_{SPRD}$.

III. Classification Algorithm

As it was mentioned in the introduction, we do the classification of a signal in two different steps, such as depicts Fig. 1. Firstly, the ECG signal is preprocessed, i.e., windowing and filtering. Once the signal is spurious free, the SPDR algorithm is used in order to classify it either as SR signals or VT/VF. In case of the signal has been classified as SR the process is finished. Otherwise, the ECG will be further analyzed using the CM algorithm in order to classify it either as VT or VF.

We worked with ECG segments of different length in order to observe the behavior of the algorithms and choose the optimal window size. We corroborated that the performance we got, remains approximately similar for any windowed longer than 3 sec. The detection results using our ECG database are shown in Table I.

The classification SPDR-CM algorithm has been implemented in a computer using Lab-View and C++ software. The tool is capable to acquire signal either from a patient using a commercial system for acquisition of ECG signals or from a commercial DataBase. The user has the possibility to customize the Acquisition front-end panel setting up the sample frequency, channels, window length, etc, Fig. 2.

IV. Discussion and Open Issues

Even though X-S Zhang states in [3] that has achieved a perfect classification with the CM algorithm, we got a very poor result testing the algorithm with signals from MIT-BIH database. The reason may be that in [3] the ECG records are own-recorded and it's known that a database derived from defibrillator implantation studies is usually much more stable and rhythm specific in comparison to a general ECG (like MIT-BIH) database [4]. In table II we represent the results obtained from different authors, and it is obvious that authors who use own-recorded signal achieve a higher performance.

TABLE I
Performances of the proposed method for detecting SR, VT and VF arrhythmia's for different window lengths.
*Sensitivity=TP/(TP+FN), where TP=true positive, FN=false negative [5]

Window Length (sec)	SENSITIVITY (%)			TOTAL CORRECT
	SR	VT	VF	
3	21/21=100	44/61=72.1	54/56=96.4	119/138=86.2
4	21/21=100	46/61=75.4	53/56=94.6	120/138=86.9
5	21/21=100	47/61=77	53/56=94.6	121/138=87.7
6	21/21=100	47/61=77	53/56=94.6	121/138=87.7
7	21/21=100	50/61=81.9	52/56=92.8	123/138=89.1
8	21/21=100	50/61=81.9	53/56=94.6	124/138=89.8
9	21/21=100	49/61=80.3	53/56=94.6	123/138=89.1
10	21/21=100	48/61=78.7	51/56=91	120/138=86.9

We may speculate that the MIT-BIH database includes a broad range of VT, consisting of both monomorphic and polymorphic types, which has been known as a precursor to lethal VF [4]. Therefore it is sometimes very difficult to distinguish multifocal VT from VF and this can be one of the reasons of the error obtained in our study.

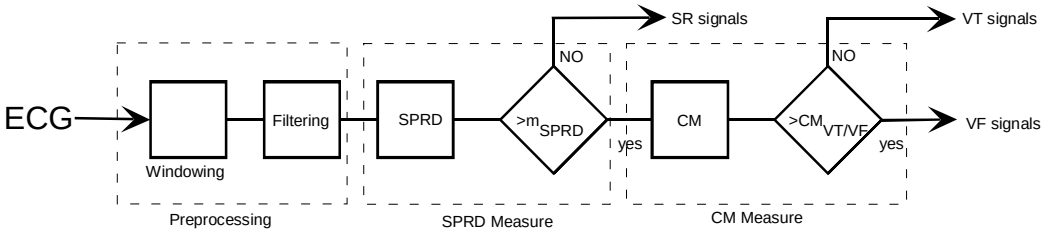

Fig. 1 SPDR-CM Algorithm Flow graph

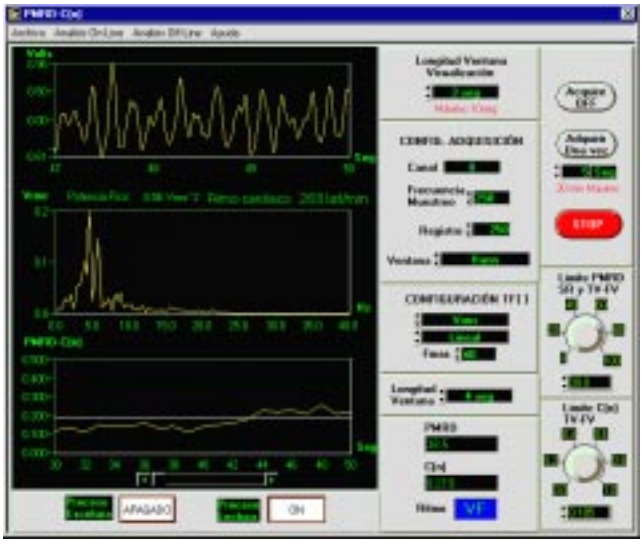

Fig.2 SPDR-CM Algorithm front-end

One limitation of this current study, as in previous studies [3][4][6], is that the algorithm performance was tested on the same database used in the development of the method. We need a larger database to test the performance of our method and verify the validity of SPDR measure.

Both thresholds established in this study, $C_{VT/VF}$ and m_{SPRD}, could be modified in order to be adaptable to the patient. The utility of this method falls in this very important issue.

Finally, the most common approach to classify arrhythmias is based in the detection of the QRS complex. Recently new works have appeared suggesting another approaches, where in spite of detecting the QRS complex, the ECG signals are windowed with relatively large window [4]. This works belong to this philosophy. Therefore it will be interesting to compare the result of these both algorithms using the same ECG database.

V. CONCLUSION

Based on experimental results using the MIT Data Base, the Complexity Measure algorithm proposed by X-S Zhang is not enough for the correct classification of SR, VT and VF signals in the time domain. Intensive experimental studies using SR signals have demonstrated that a new feature named Sample Percentage in the Dynamic Range, SPDR, is capable to separate SR signals from VT/VF signals. The combination of both, Complexity Measure and Sample Percentage in the Dynamic Range has given interesting performances obtained an overall error rate SR classification of 9%. The SPDR-CM algorithm has been implemented in a computer providing the user with the possibility to set-up the acquisition and processing characteristics of ECG signal.

REFERENCES

[1] Special Issue on Electrical Therapy of Cardiac Arrhythmias, *Proceedings of the IEEE*, Vol. 84, No. 3, March 1996.

[2] A. Lempel, J. Ziv, "On the Complexity of Finite Sequences", *IEEE Trans. Inform. Theory*, Vol. 22, pp 75-81, January 1976.

[3] Xu-Sheng Zhang, Yi-Sheng Zhu. "Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure", *IEEE Trans. Biomed. Eng.*, Vol. 46. pp 548-555, May 1999.

[4] S-W. Chen, P.M. Clarkson, and Q.Fan, "A robust sequential detection algorithm for cardiac arrhythmia classification", *IEEE Trans. Biomed. Eng.*, Vol. 43. pp 1120-1125, 1996.

[5] N.V. Thakor, "From Holter monitors to automatic defibrillators: Developments in ambulatory arrhythmia monitoring", *IEEE Trans. Biomed Eng.*, Vol. BME-31, pp 770-779, 1984.

[6] S. Chen, N.V. Thakor y M.M. Mover. "Ventricular fibrillation detection by a regression test on the autocorrelation function", *Mec. Biol, Eng. Comput.*, Vol. 25, pp 241-249, 1987.

[7] CJ Finelli, JM Jenkins, LA DiCarlo. "Detection and Identification of Cardiac Arrhythmias Using an adaptive, Linear-Predictive Filter", *IEEE 1993 Computers in Cardiology*, pp 177-180.

TABLE II
COMPARATIVE RESULTS OF SENSITIVITY AMONG DIFFERENT METHODS

Paper/Method		SR	VT	VF	DATABASE
[3]/C		100	100	100	Own-recorded from body surface 34 SR, 85 VT monomor. and VF 85
[6]/Regression Test (RT)		NA	100	100	Own-recorded from body surface Not available quantity
[4]	RT, same [6]	NA	81	90	MIT-BIH Malignant Arrhythmia Dat. 30 VT and 70 VF
	SPRT		93	96	
[7]	CWA	NA	100	50	Own-recorded and intracardiac 11 VT and 8 VF
	ALPF		91	75	
[4]/ANN		99.3	59.1	91.2	Own-recorded and intracardiac. 9050 SR, 1249 VT and 2297 VF
Here	C	23.8	81.9	94.6	MIT-BIH Malignant Arrhythmia Data. MIT-BIH Normal Synus Rhythm 21 SR, 61 VT y 56 VF
	SPRD-C	100	81.9	94.6	