# ECG CLASSIFICATION USING AFFINE INVARIANT CHARACTERIZATION OF PHASE SPACE

Heba Afify, Manal A. Wahed, Yasser M. Kadah
Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt
E-mail: ymk@k-space.org

*Abstract*-**Methods of classification of cardiac arrhythmias have decisive influence on performance of all electrocardiography (ECG) signal processing systems. This work presents a new framework for nonlinear pattern recognition /classification of ECG based on the n-dimensional moment invariants recordings of patients suffering from ventricular arrhythmias. A novel approach of applying the theory of moment invariants to analyze ECG used to improve the accuracy and reliability of the outcome predictions. The performance measures of the sensitivity and specificity of these algorithms will also be presented using data sets from the MIT-BIH Arrhythmia database.**
*Keywords* - **Signal processing, Arrhythmia detection, moment invariants.**

## I. INTRODUCTION

Application of nonlinear methods (fractal analysis, order statistics based methods) was used in detecting of various diseases and abnormal patient states from physiological signals like ECG. The main interest is focused on detecting states leading to sudden cardiac death from the ECG signals.

In the last two decades, there has been an increasing interest in applying techniques from the chaotic dynamical system theory in studying ECG signals, several features can be used to describe system dynamics including correlation dimension, Lyapunov exponents, approximate entropy, etc. These features have been used to explain ECG signal behavior by several studies. Nevertheless, these studies applied such techniques only to a few sample ECG signals. Due to the stochastic of such signals, such studies did not allow the extraction of a general statistical description of the dynamics of different arrhythmia types.

A nonlinear, state space based method to quickly and accurately identify Life-threatening arrhythmia was proposed. By means of state space reconstruction, a time series of interest can be embedded to a high-dimensional space, namely state space. The trajectory constructed from the time series using state space reconstruction reveals more straightforwardly the dynamical behaviors of the system that governs the generation of the time series. The high-dimensional trajectory contains useful information for nonlinear signal classification. So far, most nonlinear features in the current literature are global measurements of the trajectories in the sense of dynamics and geometry [1], among which Lyapunov exponent and fractal dimension are the dominant means [2]. The global measurements are limited in that shape details of trajectories do not reflect in the measurements at all. In case that the global measurements are not distinct enough among different classes, shape details of trajectories becomes the only clues to separate them. However, we have been suffering from the lack of available tools to characterize shape details of the trajectories. To fill this gape, a new framework based on moment invariants has been shown to be effective for classification of ECG signals. The ECG recordings were transformed into phase space, and some features based on the n-dimensional moment invariants were computed. A significant test was applied on the computed features to be used in the classification process. Finally, statistical classification techniques were used to assess the possibility of detecting and classifying arrhythmia using such features.

The ECG signals were obtained from the MIT-BIH arrhythmia database; the data set used in the classification experiments contains the signals from five different types including normal ECG, ventricular couplet (VC), ventricular bigeminy (VB), ventricular tachycardia (VT), and ventricular fibrillation (VF). The data set was divided into learning and testing data set, 64 independent signals for the learning set of each type and 32 independent signals for the test set of each type. All the signals were sampled at 360 samples/s except VF signals that were sampled at 250 sample/sec.

## II. METHODOLOGY

The implementation of the proposed scheme is outlined as follows.

1) *State space reconstruction*: State space reconstruction is the fundamental for analyzing nonlinear signals, by which a time series can be embedded to n-dimensional space. The state space reconstruction performed on a given time series $[S_i]$:$i$=1,2,…,$N_T$ is as follows. First, two parameters, the delay time $J$ and embedding dimension $N$, should be computed. Then, a $N$-dimensional vector sequence $[X_j]$:$j$=1,2,…,$M$ can be constructed from the time series by letting $X_j=(S_j,S_{j+J},S_{j+2J},\ldots,S_{j+(N-1)J})^T$, where $M=N_T-(N-1)J$. The determination of $J$ and $N$ are open problems yet. Here, the computation of $J$ follows [3]. Let $A(k)$ represent the autocorrelation function of the time series, where $k$ denotes the discrete-time step. Once $A(k)$ drops below $A(0)/e$, let $J=k$.

The $n$-dimensional phase space trajectory of the ECG signals was reconstructed using the delay time embedding method, where the embedding dimension $n$ was calculated using the false nearest neighborhood (*FNN*) algorithm ($n$ = 8), using the *TSTOOL* package[4].

Examples of reconstructed state spaces of the arrhythmias under study are shown in figure1.

2) *High-dimensional moment invariants*: Moment invariants are properties of trajectories in ECG signal. They are useful because they define a simply calculated set of region properties that can be used for classification and they describe signal independent of translation, scale and rotation.

The history of moment invariants begun many years before the appearance of first computers, in the 19th century under the framework of the theory of algebraic invariants. The theory of algebraic invariants probably originates from famous German mathematician David Hilbert [5]. Moment invariants were firstly introduced to the pattern recognition community in 1962 by Hu [6], who employed the results of the theory of algebraic invariants and derived his seven famous invariants to rotation of 2-D objects. Since that time, numerous works have been devoted to various improvements and generalizations of Hu's invariants and also to its use in many application areas. Dudani [7] and Belkasim [8] described their application to aircraft silhouette recognition, Wong and Hall [9], Goshtasby [10] and Flusser and Suk [11] employed moment invariants in template matching and registration of satellite images, and many other authors used moment invariants for character recognition. Moreover, no one of them paid attention to use moment invariants in signal analysis.

The $n$-dimensional moments of order $p$ of a function of intensity $\rho(x_1, ..., x_n) = \rho(x)$ are defined in terms of Rieman integral as:

$$m_{P1\ldots Pn} = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_1^{P1}\ldots x_n^{Pn}\ \rho(x)\mathrm{d}x_1\ldots\mathrm{d}x_n \quad (1)$$

Where $p_i + ... p_n = p$, $0 < p < \infty$. It is assumed that $\rho(x)$ is piecewise continuous and therefore bounded function, and it can have nonzero values only in a finite part of the $R^n$; then the moments of all order exist.

The central moments are defined as

$$\mu_{P1\ldots Pn} = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (x_1 - \bar{x}_1)^{P1}\ldots(x_n - \bar{x}_n)^{Pn}\ \rho(x)\mathrm{d}x_1\ldots\mathrm{d}x_n \quad (2)$$

$$\text{Where}\quad \bar{x}_1 = \frac{m_{1\ldots0}}{m_{0\ldots0}}, \ldots\ldots \bar{x}_n = \frac{m_{0\ldots1}}{m_{0\ldots0}} \quad (3)$$

From this, we get $\mu_{0\ldots0} = m_{0\ldots0} = \mu$.

Where $\mu$ represents the zeroth order moment, the central first moment is zero, and the second central moment is the variance. In this work the moment invariants were reconstructed with P=2, and n = 8.

We select from the following absolute moment invariants [12] seven features for the ECG classification.

$$\varphi_1 = \frac{1}{\mu^{n+2}}\begin{vmatrix}\mu_{2\ldots0} & \cdots & \mu_{1\ldots1}\\ \cdots & \cdots & \cdots\\ \mu_{1\ldots1} & \cdots & \mu_{0\ldots2}\end{vmatrix} \quad (4)$$

$$\varphi_2 = \frac{1}{\mu^4}(\mu_{20}\mu_{02} - \mu^2_{11}). \quad (5)$$

$$\varphi_3 = \frac{1}{\mu^{10}}((\mu_{30}\mu_{03} - \mu_{21}\mu_{12})^2 - 4(\mu_{30}\mu_{12} - \mu^2_{21})(\mu_{21}\mu_{03} - \mu^2_{12})). \quad (6)$$

$$\varphi_4 = \frac{1}{\mu^6}(\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu^2_{22}). \quad (7)$$

$$\varphi_5 = \frac{1}{\mu^9}(\mu_{40}\mu_{22}\mu_{04} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}\mu^2_{13} - \mu^2_{31}\mu_{04} - \mu^3_{22}). \quad (8)$$

$$\varphi_7 = \frac{1}{\mu^5}(\mu_{200}\mu_{020}\mu_{002} + 2\mu_{110}\mu_{101}\mu_{011} - \mu_{200}\mu^2_{011} - \mu^2_{110}\mu_{002} - \mu^2_{101}\mu_{020}). \quad (9)$$

$$\varphi_8 = \frac{1}{\mu^7}(\mu^2_{20}\mu_{04} - 4\mu_{20}\mu_{11}\mu_{13} + 2\mu_{20}\mu_{02}\mu_{22} + 4\mu^2_{11}\mu_{22} - 4\mu_{11}\mu_{02}\mu_{31} + \mu^2_{02}\mu_{40}). \quad (10)$$

*3) Features due to transform every trajectory to a new coordinates:* The three features used to control the transformation have been selected to be the θ, α are angular displacements in radians measured from the x-y plane, and the x-z plane, respectively, and R is the distance from the center of gravity to points in trajectory. All features for five different ECG types serve as input to a set of statistical classifiers.

## III. Results and Discussion

The seven features ($\varphi_1$, $\varphi_2$, $\varphi_3$, $\varphi_4$, $\varphi_5$, $\varphi_7$, $\varphi_8$), and the three features (R, θ, α) were extracted to form the features vectors. In this work, the T-test was used to test the significance of each feature to be used in classifying different arrhythmia types. Results of significance test for each feature was shown in tables I-XI.

TABLE I

P-values of t-test for $\varphi_1$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 6.6829e-5 | 0.0344 | 0.0881 | **0.3413** |
| **VC** | | 0.0708 | **0.1438** | **0.4132** |
| **VT** | | | **0.1300** | **0.3840** |
| **VB** | | | | **0.5245** |

TABLE II

P-values of t-test for $\varphi_2$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 5.8459e-13 | 1.426e-4 | 6.9459e-8 | 0.0135 |
| **VC** | | 0.0149 | 4.5932e-5 | 0.0441 |
| **VT** | | | **0.1119** | **0.0575** |
| **VB** | | | | **0.2859** |

TABLE III

P-values of t-test for $\varphi_3$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 0.0192 | **0.1505** | 3.0668e-4 | **0.1506** |
| **VC** | | **0.5259** | 0.0015 | **0.2121** |
| **VT** | | | 1.256e-4 | **0.1340** |
| **VB** | | | | 5.328e-4 |

TABLE IV

P-values of t-test for $\varphi_4$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 3.4168e-5 | 4.3813e-4 | 1.0644e-5 | 0.0182 |
| **VC** | | 0.0285 | 2.765e-4 | **0.0574** |
| **VT** | | | 1.7405e-4 | **0.0668** |
| **VB** | | | | **0.1636** |

TABLE V

P-values of t-test for $\varphi_5$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 7.8312e-6 | **0.2607** | 0.0026 | 0.0377 |
| **VC** | | **0.3718** | 0.0165 | 0.0967 |
| **VT** | | | **0.7309** | **0.4369** |
| **VB** | | | | **0.6860** |

TABLE VI

P-values of t-test for $\varphi_7$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 2.5071e-10 | 1.9405e-5 | 1.2443e-4 | **0.1994** |
| **VC** | | 0.0083 | 0.0043 | **0.3185** |
| **VT** | | | 0.0130 | **0.2752** |
| **VB** | | | | **0.4572** |

TABLE VII

P-values of t-test for $\varphi_8$

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 1.1110e-4 | 2.3489e-4 | 1.7371e-8 | 0.0184 |
| **VC** | | 0.0160 | 1.4495e-6 | 0.0412 |
| **VT** | | | 2.367e-5 | 0.0387 |
| **VB** | | | | **0.3582** |

TABLE VIII

P-values of t-test for R

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | 1.410e-12 | 3.534e-4 | 3.9276e-3 | 0 |
| **VC** | | **0.0593** | 4.237e-3 | 0 |
| **VT** | | | 1.1176e-4 | 1.560e-4 |
| **VB** | | | | 3.1814e-4 |

TABLE IX

P-values of t-test for θ

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | **0.5455** | **0.0610** | **0.1213** | 0.0040 |
| **VC** | | **0.2217** | 0.0283 | 9.233e-4 |
| **VT** | | | 3.2251e-4 | 1.2225e-5 |
| **VB** | | | | 0.0400 |

.

TABLE X

P-values of t-test for α

| | VC | VT | VB | VF |
|---|---|---|---|---|
| **Normal** | **0.1919** | **0.3001** | **0.4849** | 6.3963e-4 |
| **VC** | | **0.8002** | 0.0272 | 6.034e-6 |
| **VT** | | | 0.0499 | 2.1702e-5 |
| **VB** | | | | 4.825e-4 |

The results confirm that the normal ECG signals can be statistically differentiated from VC by the features ($\varphi_1$, $\varphi_2$, $\varphi_3$, $\varphi_4$, $\varphi_5$, $\varphi_7$, $\varphi_8$, θ, α), from VT by the features ($\varphi_1$, $\varphi_2$, $\varphi_4$, $\varphi_7$, $\varphi_8$, R), from VB by the features ($\varphi_1$, $\varphi_2$, $\varphi_3$, $\varphi_4$, $\varphi_5$, $\varphi_7$, $\varphi_8$, R), and from VF by the features ($\varphi_2$, $\varphi_4$, $\varphi_5$, $\varphi_8$, R, θ, α). We can also deduce that features (R, θ, α) can discriminate between VF and other abnormal signals, and features ($\varphi_3$, R) can discriminate between VB and other abnormal signals. Moreover, features ($\varphi_1$, $\varphi_2$, $\varphi_4$, $\varphi_7$, $\varphi_8$) can discriminate between VC and VT.

The features were fed into classification process based on minimum distance classifier; *K-NN* classifier. Results showed that the features in one vector were applicable to classification of different classes. Classification results are listed in tables XI-XII.

The K-NN classifier provides better results in both the detection and classification process, especially at k=1. But Min.distance classifier shows the lowest sensitivity.

## IV. CONCLUSION

This paper advocates that moment invariants can be used in signal analysis as features for description and recognition of ECG signal. Moreover, this new framework strategy can be used for enhancing the detection purpose. Invariant-based approach is a significant step towards robust and reliable recognition methods. It has a deep practical impact because it provides very useful information for cardiac arrhythmias. Based on the high levels of accuracy obtained by the proposed method, future work should extend to the analysis and classification of other types of signal by other statistical classification techniques.

## REFERENCES

[1] P. E. Rapp, T. A. A. Watanabe, P. Faure, and C. J. Cellucci, "Nonlinear signal classification," International Journal of Bifurcation and Chaos, vol. 12, pp. 1273-1293, 2002.

[2] M. I. Owis, A. H. Abou-Zied, A. M. Youssef, and Y. M. Kadah, "Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification," IEEE Transactions on Biomedical Engineering, vol. 49, pp. 733-736, 2003.

[3] T. S. Parker, and L. O. Chua, Practical numerical algorithms for chaotic systems, New York: Springer-Verlag, 1989

[4] TSTOOL package Ver.1.1 web site http://www.physik3.gwdg.de/tstool.

[5] D. Hilbert, *Theory of Algebraic Invariants*. Cambridge University Press,1993.

[6] M. K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. Information Theory*, vol. 8, pp. 179–187, 1962.

[7] S. A. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. Computers*, vol. 26, pp. 39–45, 1977.

[8] S. O. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern recognition with moment invariants: a comparative study and new results," *Pattern Recognition*, vol. 24, pp. 1117–1138, 1991.

[9] R. Y. Wong and E. L. Hall, "Scene matching with invariant moments", *Computer Graphics and Image Processing*, vol. 8, pp. 16–24, 1978.

[10] A. Goshtasby, "Template matching in rotated images". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, pp. 338–344, 1985.

[11] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion", *IEEE Trans. Geoscience and Remote Sensing*, vol. 32, pp. 382–38, 1994.

[12] A. G .Mamistvalov ", n-Dimensional moment invariants and conceptual mathematical theory of recognition n-dimensional solids"IEEE Transactions on pattern analysis and machine intelligence, vol. 20, no. 8, pp. 819-831, 1998.

TABLE XI
Results for detection process using different classifiers

| Classifier | Specificity | Sensitivity |
|---|---|---|
| Min.Dis. | 84.375% | 60.156% |
| **KNN(k=1)** | **87.50%** | **86.718%** |
| KNN(k=2) | 85.906% | 81.615% |
| KNN(k=3) | 85.781% | 81.615% |
| KNN(k=4) | 81.093% | 81% |
| KNN(k=5) | 81.093% | 81% |
| KNN(k=6) | 81.875% | 75.15% |

TABLE XII
Results for classification process using different classifiers

| classifier | specificity | Sensitivity for VC | Sensitivity for VT | Sensitivity for VB | Sensitivity for VF |
|---|---|---|---|---|---|
| Min.Dis. | 84.375% | 46.875% | 53.125% | 43.750% | 96.875% |
| **KNN(k=1)** | **87.50%** | **93.75%** | **81.25%** | **84.375%** | **87.50%** |
| KNN(k=2) | 84.375% | 84.375% | 78.125% | 75.00% | 78.125% |
| KNN(k=3) | 84.375% | 78.125% | 78.125% | 68.75% | 78.125% |
| KNN(k=4) | 81.25% | 75.00% | 68.75% | 62.50% | 78.125% |
| KNN(k=5) | 81.25% | 71.875% | 71.875% | 62.50% | 78.125% |
| KNN(k=6) | 78.125% | 68.75% | 78.125% | 62.50% | 78.125% |

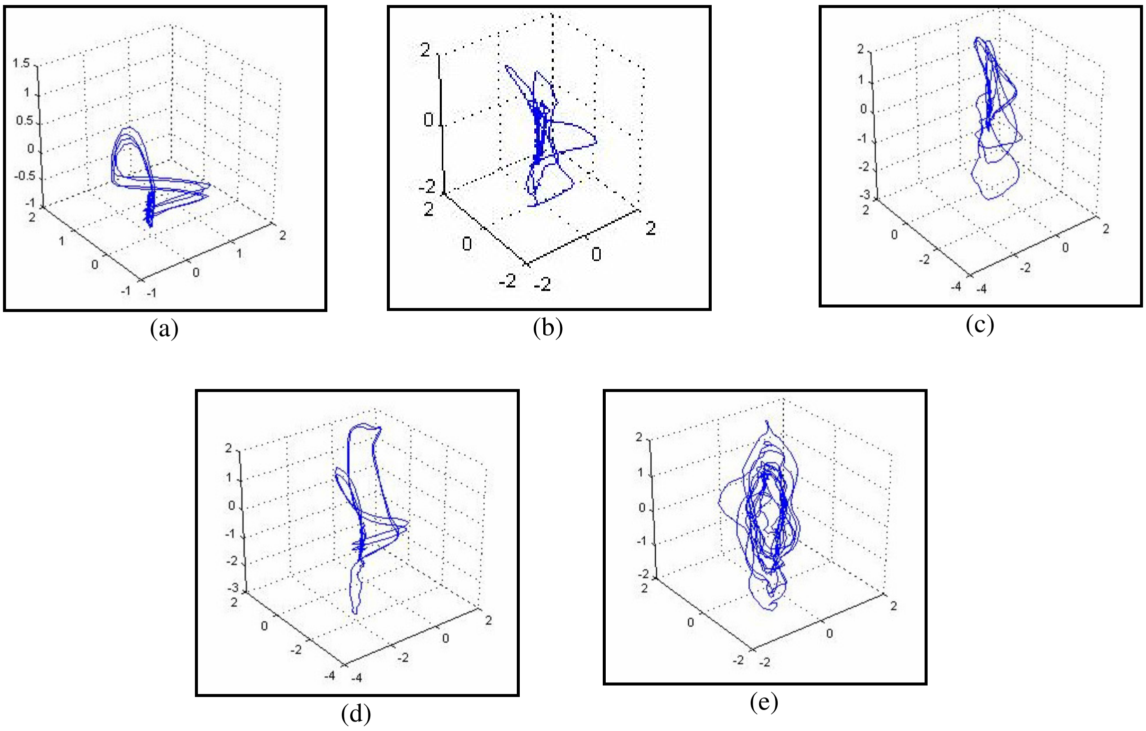

Fig.1. The 3-D plots of the first 3 vectors of the reconstructed phase space of
(a) normal, (b) VC, (c) VT, (d) VB, (e) VF