Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Morphological Heart Arrhythmia Detection Using Hermitian Basis Functions and kNN Classifier

S. Karimifard, A. Ahmadian*, *Senior Member, IEEE,* M. Khoshnevisan, M. S. Nambakhsh

***Abstract* - This paper presents the results of morphological heart arrhythmia detection based on features of electrocardiography, ECG, signal. These signals are obtained from MIT/BIH arrhythmia database. The ECG beats were first modeled using Hermitian basis functions, (HBF). In this step, the width parameter, σ, of HBF was optimized to minimize the model error. Then, the feature vector which consists of the parameters of the model is used as an input to k-nearest neighbor, kNN, classifier to examine the efficiency of the model. In our experiments, seven different types of arrhythmias have been considered. We achieved the sensitivity of 99.00% and specificity of 99.84% which are comparable to previous works. These results were obtained in less than 0.6 second which is suitable for real-time diagnosis of heart arrhythmias.**

***Index Term* - Hermitian basis function, kNN classifier, ECG beat, Morphological arrhythmia.**

I. Introduction

ECG signal shows electrical activities of the heart. Any difference in the main rhythm of the ECG is called "Arrhythmia". Heart arrhythmias are the most lethal reason in heart disease. To avoid these deaths, it is necessary to diagnose these arrhythmias accurately and fast. Common methods for arrhythmia diagnosis are based on identifying a feature in ECG by a human. According to the number of patients which need attention and also the requirement of continuous observations in such condition, it is felt that development of non-invasive and automatic techniques to detect arrhythmias are needed.

Methods of automated arrhythmia detections have different aspects. In some methods which can be categorized as signal analysis, some features of the signal are extracted such as amplitudes and position of the PQRST peaks of ECG, either in time or transform domain [1]-[4]. In these methods for arrhythmia detection, some suitable results are obtained, but they are not useful for morphological arrhythmias, because their approach has been mainly based on feature extraction rather than shape analysis. The most disadvantages of these methods are their weakness in diagnosis of morphological arrhythmias.

To solve the above problem, it is needed to use methods, which can detect the morphology of the ECG signal [5]-[7]. This goal can be achieved by modeling the signal. The main objective of modeling is to obtain some parameters which can best describe the signal. One way to reach this goal is to obtain parameters of a mathematical model for the signal which can be used to approximate the signal. By these parameters, it is also possible to obtain features like amplitude and position of the peaks, which are obtained by previous methods.

In this paper Hermitian basis functions are used to model the ECG beat, and the parameters obtained by this model are used in a kNN classifier to separate seven different arrhythmias. These ECG beats are: Normal sinus rhythm (N), Atrial premature beat (A), Ventricular premature beat (V), Right bundle branch block (R), Left bundle branch block (L), Ventricular escape beat (E) and Paced beat (P). We have used ECG beats from MIT/BIH arrhythmia database [8].

To measure the accuracy of the method, two parameters, Sensitivity and Specificity are calculated.

II. Hermitian Basis Functions

Hermite polynomials are orthogonal polynomials which are defined in the range of $(-\infty,+\infty)$ [9]. Hermite polynomials can be obtained by using the following recursive formula:

$$H_n(x) = 2x\,H_{n-1}(x) - 2(n-1)\,H_{n-2}(x) \tag{1}$$

With initial values:

$$H_0(x) = 1 \tag{2}$$

$$H_1(x) = 2x \tag{3}$$

By using Hermite polynomials, the Hermitian Basis Functions can be obtained by the following equation:

$$\phi_n(t,\sigma) = \frac{1}{\sqrt{\sigma 2^n\, n!\sqrt{\pi}}}\, e^{\frac{-t^2}{2\sigma^2}}\, H_n(t/\sigma) \tag{4}$$

S. Karimifard is with the Department of Biomedical Systems & Medical Physics, Tehran University of Medical Sciences, and Research Center for Science and Technology in Medicine, TUMS, Tehran, Iran.
(e-mail: karimifard@razi.tums.ac.ir)

* Corresponding Author: A. Ahmadian is an Associate Professor at Tehran University of Medical Sciences, TUMS, Dept. of Biomedical Systems & Medical Physics and Research Center for Science and Technology in Medicine, TUMS, Tehran, Iran.
(e-mail: ahmadian@sina.tums.ac.ir)

M. Khoshnevisan is an Associated Professor at the Business School, Griffith University, Gold Coast, Queens Land, 9726, Australia
(e-mail: m.khoshnevisan@griffith.edu.au)

M. S. Nambakhsh is with the Department of Biomedical Systems & Medical Physics, Tehran University of Medical Sciences, and Research Center for Science and Technology in Medicine, TUMS, Tehran, Iran.
(e-mail: nambakhsh@razi.tums.ac.ir)

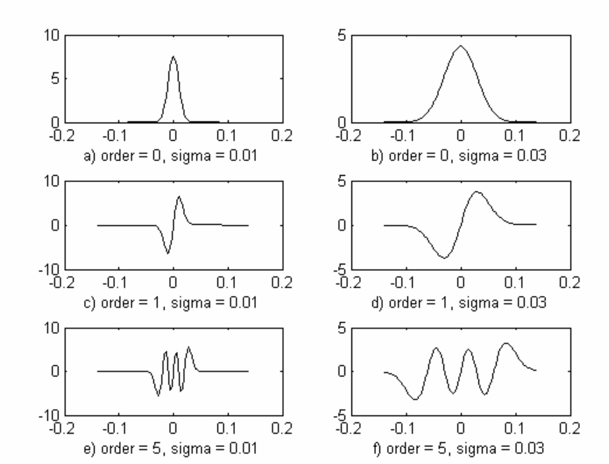

Fig. 1 Graphs of some Hermitian basis functions

In the above equation ϕ_n 's are orthogonal functions, and $H_n(t/\sigma)$ are the Hermitian polynomials. According to similarity between ECG signal and Hermitian basis functions, it is expected that an ECG beat can be expressed by a linear combination of HBF's as follows:

$$x(t)=\sum_{n=0}^{N-1} a_n\,\phi_n(t,\sigma) \tag{5}$$

$a_n's\ (n=0,1,...,N-1)$ are the coefficients of this linear combination. $\phi_n(t,\sigma)$'s are the Hermitian basis functions. There are two main parameters in this model which have impact on the final approximation error named as Hermitian order, N, and width parameter, σ .

Increasing the order of the Hermitian basis functions will lead to more oscillation on these functions which enable the model to track fast variations of the ECG signal. Another parameter in the HBF's is the width parameter, σ , which controls the dilation and contraction of the basis functions. In Fig.1, some basis functions and the effect of changing σ are shown.

This linear combination is similar to a multi-resolution approximation of the signals by considering the HBFs as the basis functions which are being dilated or contracted by width parameter to reach the approximation of the signal at the specific resolution defined by N.

In this paper, we used 201 samples from each beat of ECG signal; a sample on R peak and 100 samples before and after that. Samples of each beat are shifted by the value of mean of the first and last samples of beat. These samples are approximated by linear combinations of Hermitian basis functions according to Eq. 5. As it is shown in Fig. 2, a good approximation of the signal with an error of less than 10% can be achieved by a combination of 20 basis functions. The coefficients, a_n 's, are calculated based on minimizing the least square error as defined below:

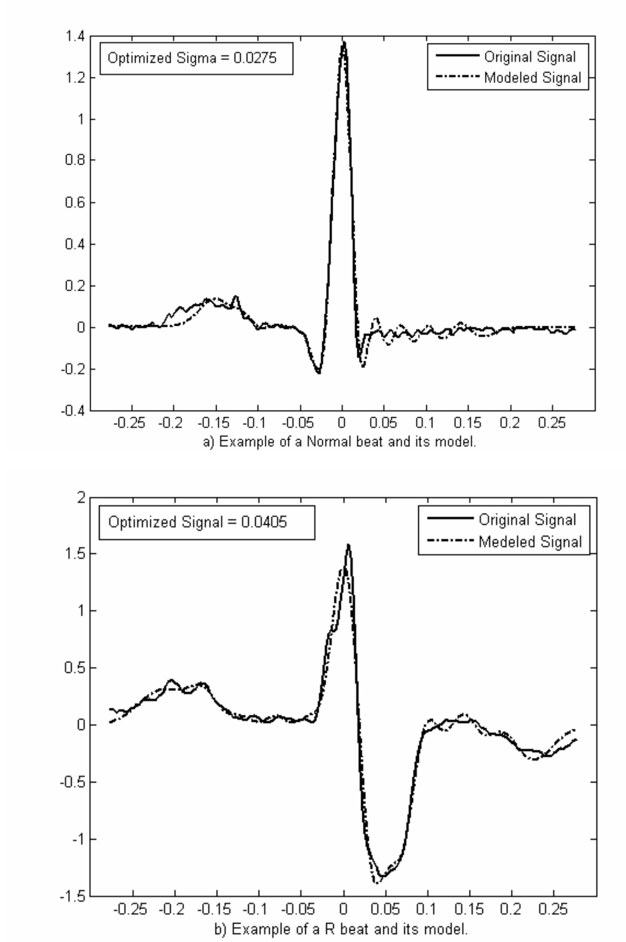

Fig. 2 Two examples of ECG beats and their models
a) A normal sinus rhythm beat
b) A right bundle branch block beat

$$E=\sqrt{\sum_i\left(x(t_i)-\sum_{n=0}^{N-1} a_n\,\phi_n(t_i,\sigma)\right)^2} \tag{6}$$

In here, the pseudo-inverse method is used to obtain a_n which is based on minimizing least square error [9]. Assume that Eq. 5 is rewritten as in a matrix form as below:

$$\Phi A=X \tag{7}$$

According to the fact that matrix Φ is non-square, thus it is not reversible. In pseudo-inverse method a matrix Φ' is found to reach the equality $\Phi'\Phi\cong I$, using least square estimation[10]. After that the coefficient matrix A can be calculated as Eq. (8).

$$A=\Phi' X \tag{8}$$

In order to find an optimized model, with a predefined N, the width parameter was also considered. For each ECG beat, the σ parameter is also optimized to minimize the approximation error defined in Eq. 6. This parameter has an important role to reach the minimum approximation error. As we experienced, changing this parameter has a large impact on the final error as shown in Fig.3.

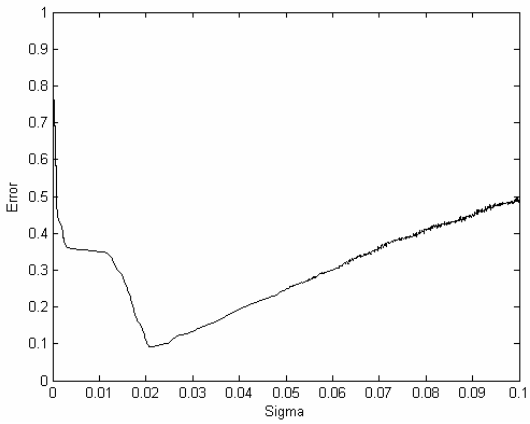

Fig. 3 Variation of normalized approximation error versus width parameter

III. kNN CLASSIFIER

In this work kNN classifier was used to separate different types of ECG beat according to the following steps [11]:

- Out of N training vectors, identifying k nearest neighbors, irrespective of class label.
- Among these k samples, identifying the number of vectors, k_i, that belongs to a arrhythmia class (ω_i).
- Assigning the input sample x to a arrhythmia class with the maximum number, k_i, of samples.

Determining the nearest neighbor can be calculated using various distance measures like Euclidean distance, sum of absolute differences and correlation measure. These distances are calculated according as below [9]:

Euclidean distance:

$$d = \sqrt{\sum_i (x_i - y_i)^2} \tag{9}$$

Correlation distance:

$$d = 1 - \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})^2}} \tag{10}$$

Sum of absolute differences:

$$d = \sum_i |x_i - y_i| \tag{11}$$

Three above measures with two different values of k (1 and 3) are used in this experiment to classify the beats.

IV. RESULTS

In our experiment, the results were obtained by using 11549 different ECG beats of seven different types. The total number of training and test beats of each type is shown in Table I.

TABLE I
Total number of ECG beats which is used as training and test data.

	Total beats	Training beats	Test beats
N	2000	1200	800
A	722	433	289
V	2938	1763	1175
R	2251	1351	900
L	1456	874	582
P	2077	1246	831
E	105	63	42
Total	**11549**	**6930**	**4619**

Sensitivity and specificity are two parameters which are used to examine the efficiency of the model and classifier. They are computed with the following equations:

$$Sp = \frac{TN}{TN + FP} \tag{12}$$

$$Se = \frac{TP}{TP + FN} \tag{13}$$

Where TP represents the True Positive, TN represents the True Negative, FN represents the False Negative and FP represents the False Positive detections.

According to Table II, it is decided experimentally which feature vector of length 20, including coefficients of linear combinations of HBF, is used as an input vector to kNN classifier. As can be seen in Table II, using feature vector of length more than 20, decreases sensitivity and specificity of the algorithm. This is due to the fact that some elements which have less importance in modeling are added to the feature vector while having the same weight factor in classifier. Another problem in using larger feature vector is increasing the time needed to model the signal.

The results of classification using 1NN with correlation measure are shown in Table III, in detail. Other results for 1NN and 3NN with Euclidean, Sum of absolute difference and Correlation measures, with feature vector of length 20 are shown in Table IV.

TABLE II
Results of Classifying by 1NN with correlation distance for different orders of model

Order	Sensitivity	Specificity	Time Consuming
2	27.76	83.33	0.17 sec
5	93.59	98.83	0.23 sec
10	96.82	99.45	0.32 sec
15	98.05	99.65	0.45 sec
20	98.85	99.78	0.56 sec
25	98.21	99.67	0.90 sec
30	98.40	99.71	0.97 sec

TABLE III
Results of classifying by 1NN classifier and Correlation distance measure

	TOTAL	TP	TN	FP	FN	SE	SP
N	800	799	3809	10	1	**99.87**	**99.74**
A	289	270	4320	11	19	**93.42**	**99.75**
V	1175	1159	4142	7	16	**98.64**	**99.83**
R	900	895	4077	10	5	**99.44**	**99.75**
L	582	580	4031	6	2	**99.65**	**99.85**
E	42	39	4575	2	3	**92.86**	**99.96**
P	831	831	3788	0	0	**100**	**100**
	4619	4573	28742	46	46	**99.00**	**99.84**

TABLE IV
Classification results of the length 20 feature vector using kNN classifier.

	1NN Euclidean Distance		3NN Euclidean Distance		1NN Sum of absolute difference		3NN Sum of absolute difference		1NN Correlation		3NN Correlation	
	SE	**SP**	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**	**SE**	**SP**
N	99.87	99.50	100	99.37	99.87	99.55	100	99.45	99.87	99.74	100	99.66
A	90.31	99.68	88.58	99.65	90.31	99.68	88.23	99.70	93.42	99.75	92.39	99.65
V	98.13	99.71	97.96	99.76	98.30	99.71	97.79	99.73	98.64	99.83	97.96	99.76
R	99.33	99.68	99.22	99.66	99.33	99.71	99.44	99.66	99.44	99.75	99.33	99.68
L	98.97	99.83	99.14	99.80	98.80	99.80	99.14	99.73	99.65	99.85	99.14	99.88
E	92.86	99.98	92.86	99.98	92.86	99.96	92.86	99.93	92.86	99.96	92.85	99.91
P	100	100	100	100	100	100	100	100	100	100	100	100
	98.57	99.77	98.44	99.75	98.59	99.77	98.42	99.75	99.00	99.84	98.70	99.79

V. CONCLUSION

The experiments show that the proposed feature vector in this paper is quite suitable for detecting morphological arrhythmias. Also kNN classifier is quite powerful for separating different types of arrhythmias. Some important points of this experiment are listed below:

- Time consumption for extracting model parameters and classification process by 1NN classifier with correlation distance measure (on an AMD 2200+, 512MB RAM computer) takes almost 0.56 second for each beat. This value is less than a normal ECG beat duration, which makes the algorithm suitable for real-time diagnosis of heart arrhythmias.
- Sensitivity and Specificity obtained in this experiment are comparable to the previous works such as the experiment in [7] which was based on Autoregressive modeling.
- Feature vector proposed in here has reasonable length which leads to fast classification.
- According to the results of applying different kNN classifier and those obtained by Support Vector Machine classifier in [6], it can be concluded that this feature vector has very good stability in terms of different classifier.

REFERENCES

[1] Afonoso, Tompkins, “Detecting ventricular fibrillation: Selecting the appropriate time-frequency analysis tool for the application”, IEEE Trans. Biomed. Eng., No. 14, pp. 879-882, 1995.

[2] K. Minami, H. Nakajima, T. Yoyoshima, “Real time discrimination of the ventricular tachyarrhythmia with Fourier transform neural network”, IEEE Trans. Biomed. Eng., vol. 46, pp. 179-185, 1999.

[3] M. Lagerholm, C. Peterson, G. Braccini, L. Edhenbrandt, L. Sornmo, “Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps”, IEEE Trans. Biomed. Eng., vol. 47, No. 7, pp. 838-848, July 2000.

[4] S. Z. Mahmoodabadi, A. Ahmadian, M. D. Abolhasani, M. Eslam, J. H. Bidgoli, “ECG Feature Extraction Based on Multiresolution Wavelet Transform”, Prceeding of the 2005 IEEE, Engineering in Medicine and Biology.

[5] S. Osowski, T. H. Linh, “ECG Beat Recognition Using Fuzzy Hybrid Neural Network”, IEEE Trans. Biomed. Eng., vol. 48, No. 11, pp. 1265-1271, Nov. 2001.

[6] S. Osowski, T. H. Linh, T. Markiewicz, “Support Vector Machine based expert system for reliable heart beat recognition”, IEEE Trans. Biomed. Eng., 2003.

[7] D. Ge, N. Srinivasan, S. M Krishnan, “Cardiac arrhythmia classification using autoregressive modeling”, Biomedical Engineering Online, 2002. http://www.biomedical-engineering-online.com/content/l/l/5.

[8] http://www.physionet.org/

[9] http://www.mathworks.com/

[10] Kaare Brandt Petersen, Michael Syskind Pedersen, “The Matrix CookBook”, pp. 17-18, Feb. 2006.

[11] S. Theodoridis, K. Koutroumbas, “Pattern Recognition”, Academic Press, 2nd edition, 2003.