International Journal of Electronics Engineering, 3 (1), 2011, pp. 55– 58

ECG Beat Recognition using Principal Components Analysis and Artificial Neural Network

Amitabh Sharma[1], and Tanushree Sharma[2]

[1]*Department of Electronics Engineering, University College of Engg., Kota, INDIA*

[2]*Department of Electronics Engineering, University College of Engg., Kota, INDIA*

E-mail: amitabh_sharma22@yahoo.com, and tanushreesharma2003@yahoo.co.in

Abstract: The analysis of heart beat cycles in an ECG (electrocardiogram) signal is essential for long-term monitoring of heart patients. However, it is a very tedious and time-consuming task to analyze the ECG recording beat by beat in a long-term monitoring. This is because the abnormal heart beats can occur randomly and a long-term ECG record, say 24 hours, may contain hundreds of thousands of beats. Hence, it is highly desirable to automate the entire process of heart beat classification. The present work proposes a technique for heart beat recognition. First, the individual beats belonging to each category were extracted from the MIT-BIH arrhythmia database using an R-peak detection algorithm and after preprocessing, features are extracted from the beats using Principal Component Analysis (PCA) .This process drastically reduces the dimensionality of the vectors to be classified. The feature vectors thus obtained are used to train a neural network (NN) classifier. After the network is trained, its performance in terms of its generalizing ability is tested on a separate test dataset which was not used during training.

Keywords: ECG, PCA, Neural Network, PVC, LVQ

1. INTRODUCTION

Arrhythmia is the irregularity of the heart function and is the most common cardiac disease. There are numerous types of arrhythmia disorders, out of which premature ventricular contraction for three or more successive cycles and an elevated heart rate is categorized as Ventricular Tachycardia, which might lead to the life threatening Ventricular Fibrillation. The PVC is characterized by premature occurrence of a QRS complex which is bizarre in shape and has a duration usually exceeding the normal QRS complex. For clinical detection, the PVC arrhythmia is most important to be detected. Bundle branch block (BBB) is a relatively frequent finding on the electrocardiogram (ECG). Sometimes BBB itself needs to be treated; sometimes it indicates significant underlying cardiac disease that needs to be treated. The chief effect of a bundle branch block is to disrupt the normal, coordinated and simultaneous distribution of the electrical signal to the two ventricles. Because with BBB the ventricles receive the electrical impulse one after another instead of at the same time, it takes longer to form the QRS complex on the ECG. The QRS complex is said to "widen", as shown in Fig. 1. The characteristic shapes of the QRS complex allow doctors to determine whether the right or the left bundle branch is blocked.

ECG beat classification typically requires extraction of beat templates about a fiducial point and then feature extraction from the beats. The approach of classifying the output of a feature detector offers reduced computational complexity and greater accuracy than that of using a neural network directly upon an ECG beat. The extracted features may comprise morphological features, such as the width and height of the QRS complex, QRS complex area, positions of *P*, *Q*, *S* and *T* waves etc. [1-3] heart beat temporal intervals such as R-R interval, P-R interval [4,5] etc.

Several techniques have been used for classification of ECG beats, such as linear regression [6], support vector machines (SVMs) [7, 8], neural networks [9], mixture-of-experts approach [10] etc. The neural networks most commonly used for the purpose of classification include backpropagation nets, self organizing map, learning vector quantization, ART and radial basis function nets.

2. PRESENT WORK

The various steps used in the present beat recognition system are discussed below.

2.1 Training and Test Dataset

The MIT-BIH Arrhythmia Database provided by MIT and Boston's Beth Israel Hospital has been used in the present work. It consists of 48 half-hour recordings of annotated ECG with a sampling rate of 360 Hz and 11-bit resolution over a 10-mV range. Each record is a continuous waveform and we need to extract individual beats. This is done by considering the *R* peak and extracting 90 samples on either side of this *R* peak, thus taking 181 samples (nearly 500 ms duration) corresponding to each beat. Since the focus in the present work is to classify a heart beat into one of the four categories: Normal, PVC, LBBB (Left Bundle Branch Block) and RBBB (Right Bundle Branch Block), as depicted in Fig. 1, therefore records containing these types of beats (100, 103, 115, 123, 106, 109, 111, 118, 119, 124, 214, 221, and 231) were included in the study.

First, the continuous ECG signal is passed through a high pass filter. The output of the HPF is further inputted into a nonlinear LPF constructed by a cascade of a simple point-by-point squaring operation and a moving window integration system. The signals obtained at various stages are shown in Fig. 2. Finally, a threshold is used for decision making to detect the R-peaks. Since the MIT/BIH database comes with annotations for each heartbeat marked by expert cardiologists, a training set may be prepared consisting of beats belonging to the four categories, viz., Normal, PVC, LBBB and RBBB. The signals are extracted with 181 samples (corresponding to nearly 500ms), where the *R* wave was the 91s t sample.

For the ANN classification purpose, we require three datasets: training, validation and test. The training and test dataset consisted of 800 beats, with equal quantity of each beat type. A separate validation set was also prepared for "early stopping" while training the neural network in order to ensure that it does not begin to memorize the training dataset. The beats used in the test dataset were not used during training in order to test the generalization capability of the net.

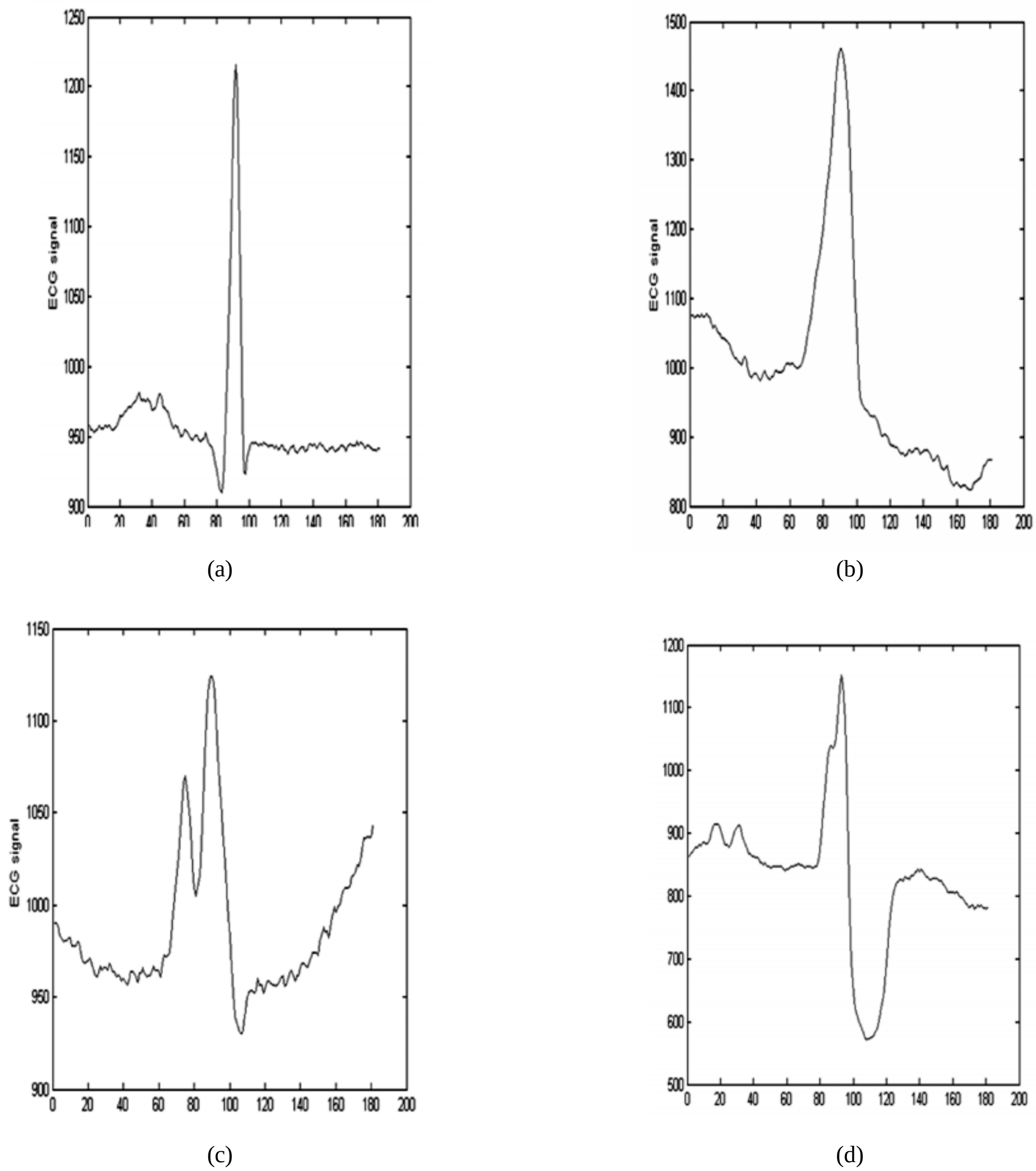
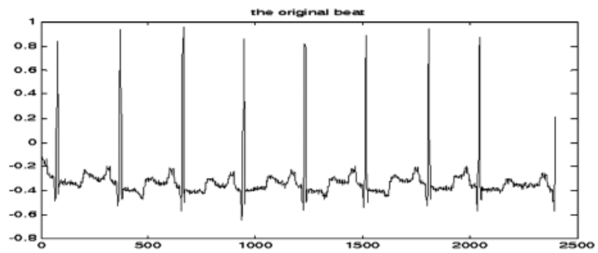


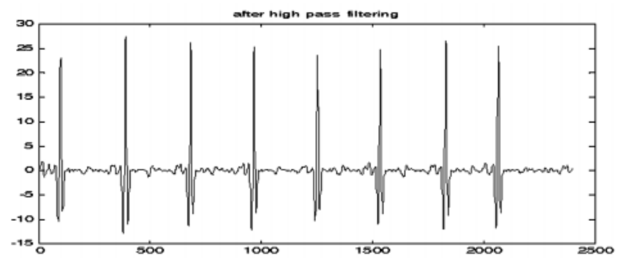
Fig. 1: Beat Types in Consideration: (a) Normal
(b) PVC (c) LBBB (d) RBBB

2.2 Feature Extraction

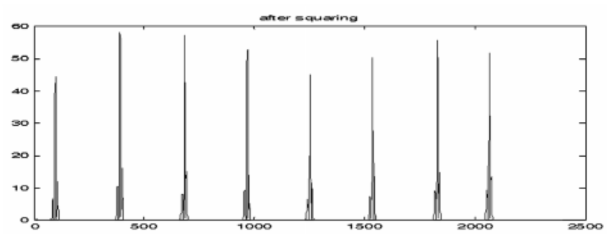
Automated ECG beat classification was traditionally performed using various time domain features extracted from an ECG beat. The features used include the width and height of QRS complex, RR interval, QRS complex area, etc. One of the difficulties is that these features are susceptible to variations of ECG beat morphology and temporal characteristics. Here, Principal Components Analysis (PCA) also known as Karhunen-Loève (KL) transform has been used for feature extraction from the beats. Principal components analysis (PCA) is a useful technique for the compression and classification of data. The purpose is to reduce the dimensionality of a data set (sample) by finding a new set of variables, smaller than the original set of variables that still retains most of the sample's information. By information we mean the variation present in the sample given by the correlations between the original variables. The new variables, called principal components (PCs), are uncorrelated, and are ordered by the fraction of the total information each retains. It can be viewed as a rotation of the existing axes to new positions in the space defined by original variables. New axes are orthogonal and represent the directions with maximum variability. This technique has 3 effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other); it orders the resulting orthogonal components (principal components) so that those with the largest variation come first; and it eliminates those components that contribute the least to the variation in the data set. Here, we have eliminated those principal components that contribute less than 1% to the total variation in the data set. Finally, the transformed input vectors and the principal component transformation matrix are obtained. After the network has been trained, this matrix should be used to transform any future inputs that are applied to the neural network classifier. It effectively becomes a part of the network, just like the network weights and biases. If we multiply the normalized input vectors by the transformation matrix, we obtain the transformed input vectors.



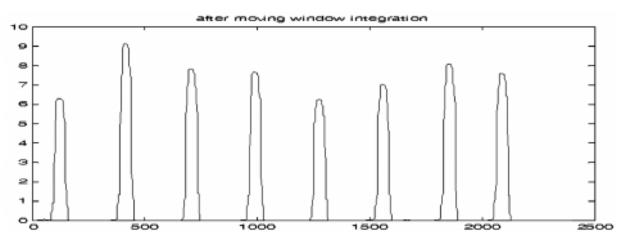
(a)



(b)



(c)



(d)

Fig. 3: (a) The original Beats (b) After High Pass Filtering (c) After Squaring (d) After Moving Window Integration

2.3 Classification using Neural Network

After the aforementioned feature extraction process, the feature vectors containing five components were obtained. Thus the dimensionality of vectors to be input to the neural network has been drastically reduced from 181 to 5. This indicates that there was significant redundancy present in the dataset. These feature vectors are now applied at the input of a neural network classifier. In the present work, an LVQ net has been used. LVQ is a supervised clustering based classification technique which classifies a feature vector according to the label of the cluster prototype (code word) into which is clustered. Classification error occurs when the

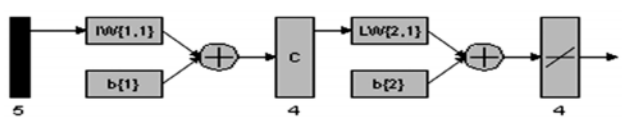


Fig. 3: Architecture of the LVQ net used

feature vectors within the same cluster are actually drawn from different classes. To minimize classification error, the LVQ algorithm fine tunes the clustering boundary between clusters of different class labels by modifying the position of the clustering center (prototype or code word). This method is called “learning vector quantization”. In this study, the optimized learning-rate LVQ1 and LVQ2.1 algorithms proposed by Kohonen were used for the training and fine-tuning of the code book respectively. The network implementation was carried out in MATLAB.

Table 1
Results Obtained Using the Proposed Classifier

Number of class	*Beat type extracted*	*Number of beats*		*Code of class (LVQ)*	*% Accuracy*
		training	*test*		
0	Normal	400	400	1000	100
1	LBBB	400	400	0100	98
2	RBBB	400	400	0010	96.4
3	PVC	400	400	0001	98.7

3. RESULT AND DISCUSSION

On simulating the resultant network with the test dataset, the results are summarized in Table 1. A detection accuracy of 100% was obtained for Normal beats, 98% for LBBB beats, 96.4% for RBBB beats and 98.7 % for PVC beats.

4. CONCLUSION

It can be concluded that PCA provides a powerful feature extraction tool, which leads to a great reduction in the dimensionality of the input vectors to be presented to the neural network classifier. When the obtained feature vectors are applied to a suitable neural network, high values of sensitivity and positive predictivity are obtained. In the future, it is intended to develop a classifier that is able to recognize an even larger no. of beats, while still maintaining the accuracy.

REFERENCES

[1] M. O'Dwyer, P. De Chazal, R. B. Reilly, "Beat Classification for use in Arrhythmia Analysis", *Comput. Cardiol*. 27, 395-398, 2000.

[2] E., Pietka, "Expert Systems in Parameter Extraction of the ECG Signal", *In: Annu. Intl. Conf. IEEE Engineering Medicine and Biology Society*, pp. 165-166, 1988.

[3] T.H. Yeap, F. Johnson, "Rachniowski, ECG Beat Classification by a Neural Network", *In: Proc. Annu. Intl. Conf. IEEE Engineering Medicine and Biology Society*. 12, pp. 1457-1458, 1990.

[4] B. Karlik, Y. Ozbay, "A new Approach for Arrhythmia Classification", *In: Annu. Intl. Conf. IEEE Engineering in Medicine and Biology Society*. pp. 1646-1647, 1996.

[5] P. de Chazal, M. O'Dwyer, R.B. Reilly, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features", *IEEE Trans. Biomed. Eng.* **51,** 1196-1206, 2004.

[6] Young, T.Y., Huggins, W.H., "Computer Analysis of Electrocardiograms Using a Linear Regression Technique", *IEEE Trans. Biomed. Eng*. 11(3), 60-67, 1964.

[7] Zellmer, E., Shang, F., Zhang, H., "Highly Accurate ECG Beat Classification Based on Continuous Wavelet Transformation and Multiple Support Vector Machine Classifiers", *In: 2nd Int. Conf. on Biomedical Engineering and Informatics*'09, pp. 1-5, 2009.

[8] Zhang, H., Zhang, L.Q., "ECG Analysis Based on PCA and Support Vector Machines", *Int. Conf. on Neural Networks and Brain*'05, 2005.

[9] Y.H. Hu, W. Tompkins, J.L. Urrusti, V.X. Alfonso, "Applications of Artificial Neural Networks for ECG Signal Detection and Classification", *J. Electrocardiol*. 26 Suppl., 66-73, 1 994.

[10] Y.H. Hu, S. Palreddy, W.J. Tompkins, "A Patient Adaptable ECG Beat Classifier Using a Mixture of Experts Approach", *IEEE Transactions on Neural Networks*. 44, 891-899, 1997.