Manifold Learning for Premature Ventricular Contraction Detection

BR Ribeiro[1], JH Henirques[1], AM Marques[1], MA Antunes[2]

[1]CISUC, Department of Informatics Engineering, University of Coimbra, Coimbra, Portugal
[2]Department of Cardiothoracic Surgery, University of Coimbra, Coimbra, Portugal

Abstract

Prompt diagnosis of abnormally shaped wave forms in ECG signal is an important component in the early diagnosis of cardiac arrhythmias, improving the quality of life of patients. Meanwhile, detection models for Premature Ventricular Contractions (PVC) are widely investigated, a less studied problem is data analysis and visualization. In this paper, we propose an approach for PVC detection and data visualization by exploiting the intrinsic geometry of the high-dimensional data using manifold learning and Support Vector Machines (SVM). ISOMAP forms a neighborhood-preserving projection which allows to uncover the low-dimensional manifold and is used here as a pre-processing step. Then by incorporating training labels the method is capable of recognizing PVC patterns with comparable accuracy of kernel learning machines.

1. Introduction

Cardiovascular diseases (CVD) are the leading cause of death in developed countries. In this context Ventricular Arrhythmias (VA) assume a very important role, since their incidence in population can lead to critical situations and severe risk. Moreover, VA evolve from simple premature ventricular contractions (PVC) which are in most situations benign, to ventricular tachycardia (VT) and finally to critical ventricular fibrillation episodes which are potentially fatal and the main cause of sudden cardiac death.

The detection of PVCs is thus of major importance, since they are associated with an increased risk of adverse cardiac events. Resulting from an ectopic depolarization on the ventricles (which replace the normal start of the cardiac beat), a wider and abnormally shaped QRS complex occurs. Additionally, typically QRS complexes are not preceded by P waves, and T waves are usually larger and with opposite deflection to the QRS complex.

The automatic detection of PVC has been an active research during the last years. Basically the fundamental differences lie in the morphology of the ECG. Typically, a two step approach is then considered: i) first, measurements of average wave amplitudes, time duration and wave areas are used to extract a set of characteristic ECG parameters [1]; ii) second, a specified technique for classification is then applied ranging from probabilistic approaches, to knowledge-based systems or neural networks [2]. This motivates research authors either for investigating the most appropriate classifier, or for finding the most discriminative features and feature extraction methods of PVC characteristics. In [3] an algorithm for PVC detection based on QRS complex morphological characteristics is presented. In [4] the mean-square value of QRS was proposed as PVC discriminative features, together with two linear prediction coding coefficients. In [2] the proposed feature extraction methods are based on the spectral content of the signal, linear prediction coding and Principal Component Analysis (PCA). Wang et al. [5] focused on the extraction of template families of features from QRS morphology characteristics. In [6] PCA is also applied for feature extraction along with a Hebbian neural network classifier. For PVC classification, numerous algorithms such as decision trees and fuzzy ruled based networks have been proposed. Also, in [7], a wavelet transform method is introduced for feature extraction which uses a multilayer perceptron neural network.

In this paper, we look at finding meaningful low dimensional structure from the high dimensional PVC features dataset for visualization and detection purposes. Instead of working with points in the high-dimensional space a reduced space is found in the embedded learning process. While linear methods such as Principal Component Analysis (PCA) and Multidimensional Scaling (MDS) break down with the problem nonlinearities, manifold learning attempts to uncover the low-dimensional manifold by nonlinear methods. One of the most used algorithms [8] for nonlinear dimension reduction is ISOMAP (Isometric Feature mapping) which can be viewed as an extension of the Multidimensional Scaling [9]. While ISOMAP is used in unsupervised learning our approach combines this algorithm with information of class labels for PVC automatic detection.

The approach is validated using data sets from MIT-BIH Arrhythmia Data Base. We show that the accuracy of the proposed approach is comparable to benchmarked Support Vector Machines (SVM) [10] and Relevance Vector Machines (RVM) [11] despite the fewer dimensions used resulting from embedding learning.

The paper is organized as follows. In section 2 manifold learning and Isomap are briefly described. In section 3 the experimental setup is introduced and results are presented and discussed. The conclusions and future work are presented in section 4.

2. Methods

Manifold is an abstract mathematical space in which every point has a neighborhood which resembles the spaces described by Euclidean geometry. Given data points $x_1, x_2, ..., x_n \in \mathbb{R}^D$ we assume that the data lies on a d-dimensional M manifold embedded into $\mathbb{R}^D$, where $d < D$. Moreover, we assume the manifold M is described by a single coordinate chart $f: M \to \mathbb{R}^d$. The manifold learning consists of finding $y_1, ..., x_n \in \mathbb{R}^d$, where $y_i = f(x_i)$. ISOMAP is an algorithm [8] for nonlinear dimension reduction which can be viewed as an extension of the MDS. ISOMAP consists of 3 main steps:

1) Estimates which points are neighbors on the manifold M, based on the distances $d_X(i,j)$ between pairs of points i, j in the input space X by computing the weighted graph G of neighborhood relations given by the edges of $d_X(i,j)$;

2) Estimates the geodesic distances between all pairs of points in the manifold M by computing the shortest path distance on the k's nearest neighbor graph built on the data;

3) Applies classical MDS to the matrix of graph distances $D_G = \{d_G(i,j)\}$, constructing an embedding of the data in a d-dimensional Euclidean space Y that best preserves the manifold's estimated intrinsic geometry.

ISOMAP assumes that there is an isometric chart that preserves distances between points. If x_i and x_j are two points in the manifold M embedded into $\mathbb{R}^D$ and the geodesic distance between them is $d_G(x_i, x_j)$, then there is a chart $f: M \to \mathbb{R}^d$ such that $\|f(x_i) - f(x_j)\| = d_G(x_i, x_j)$.

The algorithm presupposes that for nearby points in the high dimensional space the Euclidean distance is a good approximation of the geodesic distance whereas for distant points this is not true. In this case, a shortest path computation algorithm such as, Dijkstra's or Floyd's will approximate the remainder geodesic distances. This algorithm takes as input a weighted graph of the k's nearest neighbors where its edges are weighted by the Euclidean distances between nearby data points. MDS [9] is then used to estimate the points whose Euclidean distance equals the geodesic distances. Given a matrix $D \to \mathbb{R}^{n \times n}$ of dissimilarities, MDS constructs a set of points whose interpoint Euclidean distances match those in D closely.

2.1. Supervised Isomap

In [12] an approach, S-ISOMAP, is developed for supervised nonlinear dimensionality reduction in visualization and classification problems. In S-ISOMAP, the information provided by the training class labels are used to refine the distances between inputs guiding the procedure of dimensionality reduction. The rationale is that both classification and visualization can benefit when the inter-class dissimilarity is larger than the intra-class dissimilarity. To achieve this purpose, the Euclidean distance $d(x_i, x_j)$ between two given observations x_i and x_j, labeled y_i and y_j respectively, is replaced by a dissimilarity measure [12]:

$$D(x_i, x_j) = \begin{cases} \sqrt{1 - e^{\frac{-d^2(x_i, x_j)}{\beta}}} & y_i = y_j, \\ \sqrt{e^{\frac{-d^2(x_i, x_j)}{\beta}}} - \alpha & y_i \neq y_j \end{cases} \tag{1}$$

Note that the Euclidean distance $d(x_i, x_j)$ is in the exponent and the parameter β is used to avoid that $D(x_i, x_j)$ increases too rapidly when $d(x_i, x_j)$ is relatively large. Hence, β depends on the *density* of the data set and is usually set to the average Euclidean distance between all pairs of data points. On the other hand, α gives a certain possibility to points in different classes to be *closer*, i.e. to be more similar, than those in the same class. This procedure allows a better determination of the relevant features and will definitely improve visualization. However, as the learning mapping is not explicitly given by the previous procedure, our approach includes a Generalized Regression Neural Network (GRNN) for learning the mapping, prior to applying any of the classifiers: 1) KNN (KNearest Neighbor) 2) Fisher Linear Discriminant (FLD) and 3) SVM. The steps of our procedure are illustrated in Figure 1 where only the SVM classifier is shown.

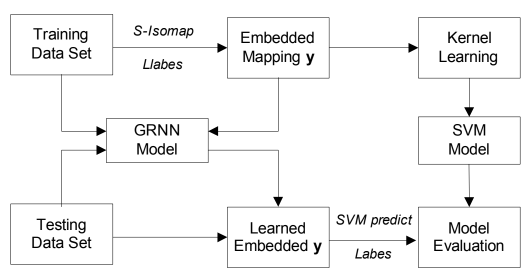

Figure 1. S-ISOMAP supervised approach.

2.2. Kernel- based methods

SVM [10] were introduced by Vapnik, based on Statistical Learning Theory, and follow the principle of maximal margin separation, i.e find an optimal separating hyperplane between the two classes in the binary problem. RVM [11] are also a kernel-based approach which use a Bayesian probabilistic learning framework. Both yield sparse models able to efficiently classify binary patterns.

3. Experimental results

Experimental data used to test the proposed approaches were taken from MIT-BIH Arrhythmia Database[1]. Extensive pre-processing of the data files allowed the construction of the training, test and validation data sets each one consisting of 19391 sample data points [13]. Sensitivity and Specificity measures were chosen for algorithms comparison. They were preferred to simple accuracy metrics since they circumvent misinterpretations, when the datasets are unbalanced, i.e the metrics are reliable even for skewed datasets.

3.1. Feature extraction

Table 1 shows the 18 extracted features, computed simultaneously with the QRS complex detector (see [13]). Figure 2 illustrates the QRS complex with the extracted feature set.

The size of the descriptive feature vector was then substantially reduced with respect to the original ECG in the Arrhythmia data base. However, for matters of data visualization further reduction is essential.

Table 1. Extracted features from ECG signal

Feature	Description
RRav	RR mean interval
RR_0	Last RR interval
SN	Signal/Noise estimation
Ql	Q-wave length
(Qc_x, Qc_Y)	Q-wave mass centre (x,y) coordinates
(Qp_x, Qp_Y)	Q-wave peak (x,y) coordinates
Rl	R-wave length
(Rc_x, Rc_Y)	R-wave mass centre (x,y) coordinates
(Rp_x, Rp_Y)	R-wave peak (x,y) coordinates
Sl	S-wave length
(Sc_x, Sc_Y)	S-wave mass centre (x,y) coordinates
(Sp_x, Sp_y)	S-wave (x,y) coordinates

[1] http://www.physionet.org/physiobank/database/html/mitdbdir/

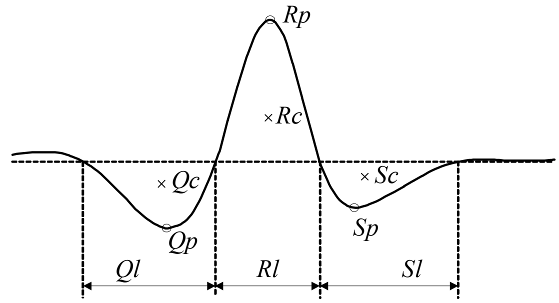

Figure 2. QRS Complex

3.2. Discussion

In Figure 3 the supervised approach, S-ISOMAP, is used for dimension reduction of d=3 with Euclidean distance for computing the dissimilarity matrix of data points. For graph computation we used K=5 to K=200 nearest neighbors. The red dots in Figure correspond to abnormally shaped QRS while blue dots to normal ones.

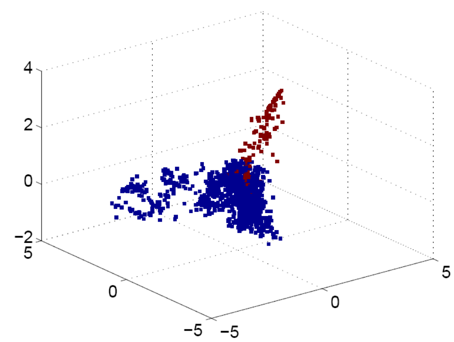

Figure 3. PVC Data Visualization with d=3 and K = 20.

Table 2 presents validation results from running the algorithm with different K nearest neighbors followed by classification for model evaluation (see Figure 1). We have used 10 runs of 10-fold cross validation resulting in 100 evaluations of each learning machine. The mean and standard deviations for the specified measures are evaluated and compared. The best result is indicated in bold and corresponds to **89.39±4.23** for sensitivity and **98.28±0.65** for specificity with S-ISOMAP using SVM. While in Table 2 results allow to test the method with respect to trustworthiness of the neighborhood projection, therefore in visualization, in Table 3 we compare our approach with the state-of-the-art kernel machines SVM [10] and RVM [11]. We observe that, despite the fewer dimensions used, our approach is better than RVM by 5% in sensitivity and by 9% in specificity over testing data. While SVM performs better by 4% in sensitivity and only 1% in specificity.

Table 2. Supervised ISOMAP

S-Isomap	KNN		FLD		SVM	
K	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
5	78.29±1.00	99.39±1.00	71.26±1.00	86.79±1.00	87.92±1.00	97.90±1.00
7	77.01±5.69	99.10±0.27	72.12±10.83	87.36±6.19	83.65±9.41	97.75±0.47
10	76.65±6.60	99.31±0.49	74.46±10.25	85.68±10.37	84.11±6.93	98.23±1.26
15	79.12±7.33	99.35±0.47	74.90±14.34	94.31±4.08	86.34±6.04	98.59±0.94
20	77.66±2.31	99.78±0.42	80.85±5.65	98.23±3.01	89.36±5.12	98.68±0.39
40	85.37±2.47	99.39±0.49	88.75±3.06	96.81±0.94	**89.39±4.23**	**98.28±0.65**
60	75.48±1.11	99.56±0.19	84.37±1.92	96.24±0.47	81.24±2.63	98.58±0.11
80	74.22±6.47	99.46±0.22	86.67±2.58	96.18±1.78	81.60±2.66	98.87±0.55
100	80.00±3.55	99.14±0.19	85.70±0.99	95.90±0.31	84.21±1.95	98.42±0.28
150	76.74±2.21	99.78±0.10	86.05±0.29	96.72±0.32	84.88±2.54	99.02±0.24
200	75.21±0.37	99.31±0.06	83.92±0.51	98.66±0.12	86.95±0.19	98.59±0.37

Table 3. Algorithm's Comparison (K = 20)

	Testing		Training	
	Sensitivity	Specificity	Sensitivity	Specificity
S-ISOMAP-KNN	77.66±0.32	99.78±0.25	88.02±3.16	99.95±0.08
S-ISOMAP-FLD	80.85±4.80	98.23±0.93	81.50±9.46	96.38±5.55
S-ISOMAP-SVM	**89.36±1.51**	**98.68±0.07**	**98.72±1.01**	**98.77±1.15**
SVM	**94.63±0.38**	**99.74±0.02**	**94.84±0.75**	**99.75±0.08**
RVM	85.78±3.48	89.53±2.89	92.28±1.16	93.53±0.81

4. Conclusions

The supervised ISOMAP combined with SVM applied to MIT-BH Arrhythmia data base were able to distinguish normal from abnormal PVC beats. Despite the reduced space used, S-ISOMAP-SVM performs better than RVM while single SVM is slightly better. For this reason, our approach takes advantage over other methods for the purpose of real-time PVC detection. Further work is foreseen to improve learning in the embedded procedure.

Acknowledgements

MyHeart Project IST-2002-507816 is gratefully acknowledged

References

[1] Hu YH, Palreddy S, Tompkins WJ. A patient-adaptable ECG beat classifier using a mixture of experts approach. IEEE Trans Biomed Eng 1997;44:891 900.

[2] ChikhM, N B, Chikh A, Reguig F. The use of artificial neural network to detect the premature ventricular contraction (PVC) beats. Elect Journal Technical Acoustics 2004.

[3] Moreas J, Seixas M, Vilani F, Costa E. A real time QRS complex classification method using mahalanobis distance. In IEEE Comp in Cardiology, volume 29. 2002; 201–204.

[4] Ham M, Han S. Classification of cardiac arrhythmias using fuzzy artmap. IEEE Trans on Biomedical Engineering 1996;43(4):425–430.

[5] Wang J, Yeo C, Aguire A. The design and evaluation of a new multi-lead arrhythmia monitoring algorithm. In IEEE Comp in Cardiology, volume 26. 1999; 675–678.

[6] Al-Nashash H. Cardiac arrhythmia classification using neural neural networks. Technology and Health Care 2000;8(6):363–372.

[7] Omer T, Giovangrandi L, Kovacs G. Robust neuralnetwork-based classification of premature ventricular contraction using wavelet transform and timing interval features. IEEE Trans on Biomedical Engineering 2006; 53(12):2507 2515.

[8] Tenenbaum JB, de Silva annd J. C. Langford V. A global geometric framework for nonlinear dimensionality reduction. Science 2000;290(5500):2319–2323.

[9] Cox TF, Cox MA. Multidimensional Scaling. 2nd edition. Chapman and hall/CRC, 2001.

[10] Vapnik VN. The nature of statistical learning theory. New York: Springer Verlag, 1995.

[11] Tipping ME. Sparse bayesian learning and the relevance vector machine. Journal of Machine Learning Research 2001;1:211–244.

[12] Geng X, Zhan DG, Zhou ZH. Supervised nonlinear dimensionality reduction in visualization and classification. IEEE Trans on Systems Man and Cybernetics Part B Cybernetics 2005;35(6):1098–1107.

[13] Ribeiro B, Marques A, Henriques J, Antunes M. Choosing real time predictors for ventricular arrhythmia detection. Int Journal of Pattern Recognition and Artificial Intelligence, 2007;21(8):1–15.

Address for correspondence:
Jorge Henriques, jh@dei.uc.pt
Department of Informatics Engineering, DEI-FCTUC, Polo II, 3030-290 Coimbra, Portugal