A Hybrid Model of Maximum Margin Clustering Method and Support Vector Regression for Solving the Inverse ECG Problem

Mingfeng Jiang[1], Jiafu Lv[1], Chengqun Wang[1], Wenqing Huang[1], Ling Xia[2], Guofa Shou[2]

[1] The College of Electronics and Informatics, Zhejiang Sci-Tech University, Hangzhou, China
[2] Department of Biomedical Engineering, Zhejiang University, Hangzhou, China

Abstract

Compared to body surface potentials (BSPs) recordings, myocardial transmembrane potentials (TMPs) can provide more detailed and complicated electrophysiological information. Noninvasively reconstructing the TMPs from BSPs constitutes one form of the inverse problem of ECG. In this study, the inverse ECG problem is treated as a regression problem with multi-inputs (BSPs) and multi-outputs (TMPs), which will be solved by the support vector regression (SVR) method. In this paper, the Maximum Margin Clustering (MMC) approach is adopted to cluster the training samples (different time instant BSPs), and the individual SVR model for each cluster is then constructed. For each testing sample, find the cluster to which it belongs, and then use the corresponding SVR model to reconstruct the TMPs. When reconstructing the TMPs over the testing samples, the experiment results show that SVR method combined with maximum margin clustering method can perform better than the single SVR method in solving the inverse ECG problem, leading to a more accurate reconstruction of the TMPs.

1. Introduction

The problem of noninvasively imaging the heart's electrical activity from body surface potentials (BSPs) constitutes one form of the inverse problem of ECG [1]. Approaches to solving the inverse ECG problem have generally relied on either an activation-based model or a potential-based model (such as epicardial, endocardial, or trans-membrane potentials). Activation-based models are used to investigate the arrival time of the propagation wavefront within the myocardium [2]. The potential-based models are used to evaluate the potential values on the cardiac surface [3] or within the myocardium at certain time instants. In this study, we have focused on the potential-based inverse solutions.

Support Vector Regression (SVR) [4] is an alternative, more robust approach to solve the inverse ECG problem. During the solution procedure, the inverse ECG problem will be treated as a regression problem with multi-inputs (BSPs) and multi-outputs (Transmembrane Potentials, TMPs). Compared with conventional regularization methods (e.g., zero order Tikhonov and LSQR), the SVR method can produce more accurate results in terms of reconstruction of the transmembrane potential distributions on epi- and endocardial surface. In addition, the SVR method with feature extraction (PCA-SVR and KPCA-SVR) outperforms that without the extract feature extraction (single SVR) in terms of the reconstruction of the TMPs [5].

Xu et al [6] proposed the Maximum Margin Clustering (MMC) method, which can perform clustering by simultaneously finding the large margin separating hyperplane between clusters. The MMC method has been demonstrated that it is very successful in many clustering problem. Recently, Zhang et al proposed [7] an efficient approach for solving the MMC via alternating optimization, which implemented by using the SVR method with the Laplacian loss in the inner optimization subproblem. The modified MMC algorithm is more accurate, much faster and more practical. In this paper, the hybrid model of MMC method and SVR is proposed to solve the inverse ECG Problem, which is referred to as MCC-SVR method.

The main purpose of this study is to investigate the reconstruction capability of TMPs from the BSPs by comparing between MCC-SVR model and a single SVR model for solving the inverse ECG problem. Based on our previously developed realistic heart-torso model, the EDL source model method is applied to generate the dataset for training and testing the SVR model.

2. Theory and methodology

The framework of the proposed MCC-SVR method is shown in the Fig. 1. The MCC method is used to classify the input data; the SVR is then applied to construct the regression model of each cluster. The detailed introduction and literature review of MCC method and SVR method can be seemed in the following sections.

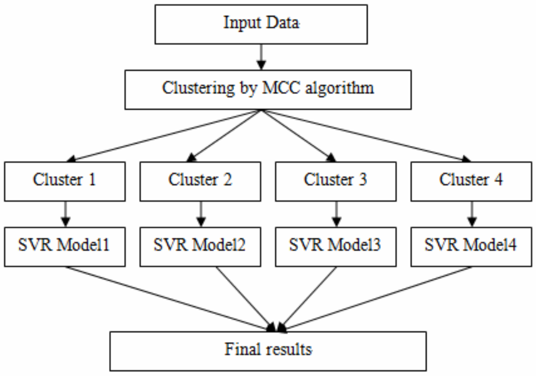


Figure 1 The framework of proposed MCC-SVR method

2.1. Maximum margin clustering method

MMC aims at extending large margin methods to assign the input data points to different classes, so that the separation between the different classes is as wide as possible. Here, we consider the case when there are only two clusters. Since one could simply assign all the data points to the same class and obtain an unbounded margin, some sort of constraint on the class balance need to be imposed. Xu et al. [6] introduced a class constraint that requires y to satisfy:

$$-\ell \le e^T y \le \ell \tag{1}$$

where $\ell \ge 0$ is a user-defined constant controlling the class imbalance. Then the margin is maximized with respect to both unknown y and unknown SVM parameter (ω, b) as follows:

$$\min_y \min_{\omega,b,\xi} \quad \|\omega\|^2 + 2C\xi^T e$$
$$\text{subject to} \quad \begin{cases} y_i(\omega\varphi(x_i)+b) \ge 1-\xi_i \\ \xi_i \ge 0,\ y_i \in \{\pm1\} \quad i=1,\cdots,n \\ -\ell \le e^T y \le \ell \end{cases} \tag{2}$$

where $\xi = [\xi_1,\cdots,\xi_n]^T$ is the vector of slack variable for the errors, and $C > 0$ is the tradeoff parameter between the smoothness $\|\omega\|^2$ and the fitness ($\xi^T e$) of the decision function $f(x)$. Here, $\varphi(x)$ denotes the high-dimensional feature space, which is non-linearly mapped from the input space x by the kernel function k. The origin nonconvex MMC problem in Eq. (2) can be formulated as a sequence of QPs which can be solved by many efficient QP solvers. However, it suffers from premature convergence and easily gets stuck in poor local optima. Zhang et al [7] proposed to replace the SVM by SVR with Laplacian loss, which can lead to a significant improvement in the clustering performance compared to that of iterative SVM procedure. The primal problem of SVR with Laplacian loss can be formulated as:

$$\min_{\omega,b,\xi_i,\xi_i^*} \quad \|\omega\|^2 + 2C\sum_{i=1}^{n}(\xi_i+\xi_i^*)$$
$$\text{subject to} \begin{cases} y_i-(\omega^T\varphi(x_i)+b) \le \xi_i \\ (\omega^T\varphi(x_i)+b)-y_i \le \xi_i^* \quad for\ i=1,\cdots,n \\ \xi_i \ge 0,\ \xi_i^* \ge 0 \end{cases} \tag{3}$$

where ξ_i and ξ_i^* are slack variables. With the labels being unknown, the MMC problem based on iterative SVR with Laplacian loss become

$$\min_{\omega,b,\xi_i,\xi_i^*} \quad \|\omega\|^2 + 2C\sum_{i=1}^{n}(\xi_i+\xi_i^*)$$
$$\text{subject to} \begin{cases} y_i-(\omega^T\varphi(x_i)+b) \le \xi_i \\ (\omega^T\varphi(x_i)+b)-y_i \le \xi_i^* \\ \xi_i \ge 0,\ \xi_i^* \ge 0 \qquad for\ i=1,\cdots,n \\ y_i \in \{\pm1\} \\ -\ell \le e^T y \le \ell \end{cases} \tag{4}$$

After ω is obtained from the optimization of SVR, the problem in Eq.(4) reduced to the form:

$$\min_{y,b} \quad \sum_{i=1}^{n}\left|(\omega^T\varphi(x_i)+b)-y_i\right|$$
$$\text{subject to} \begin{cases} y_i \in \{\pm1\} \quad i=1,\cdots,n \\ -\ell \le e^T y \le \ell \end{cases} \tag{5}$$

According to Zhang's proposition [7], for a fixed *b*, the optimal strategy to determine the y_i's in (5) is to assign all y_i's as -1 for those with $\omega^T\varphi(x_i)+b<0$, and assign y_i's as 1 for those with $\omega^T\varphi(x_i)+b>0$.

The bias *b* can be determined as follows: ①we sort the $\omega^T\varphi(x_i)$'s and use the set of midpoints between any two consecutive sorted values as the candidates of b; ② From these sorted *b*'s, the first and the last $(n-\ell)/2$ of them can be dropped ,and the middle ℓ can be remained; ③For each remaining candidate, we determine the y_i's according to above proposition and compute the corresponding objective value in Eq. (5); ④Finally, we choose the ***b*** that has the smallest objective.

2.2. Support vector regression model

The SVR algorithm was initially developed by Vapnik [8], a brief description of the algorithm is given here, for details, see references [4, 8]. As a linear regression model, the SVR algorithm relies on an estimation of a linear regression function:

$$f(x) = <\omega, x> + b, \qquad (\omega, x \in \Re) \tag{6}$$

where ω and b are the slope and offset of the regression linear, and <·, · >denotes the dot product in $\Re$. The above regression problem can be written as a convex optimization problem:

$$\begin{aligned} &\min \quad \frac{1}{2}\|\omega\|^2 \\ &subject\ to \begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon \end{cases} \end{aligned} \quad (7)$$

In Eq. (7), an implicit assumption is that a function f essentially approximates all pairs (x_i, y_i) with ε precision, but sometimes this may not be the case. Therefore, one can introduce two additional positive slack variables ξ_i, ξ_i^* to refine the estimation of variables ω and b. Now Eq. (7) can be re-formulated [8] as

$$\begin{aligned} &\min \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &subject\ to \begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \quad \ge 0 \end{cases} \end{aligned} \quad (8)$$

where the constant C is a trade-off parameter and n denotes the number of samples; ξ_i represents the upper training error, and ξ_i^* is the lower training error subject to ε intensive tube. According to the strategy outlined by Vapnik [8], using Lagrange multipliers, the constrained optimization problem shown in Eq. (8) can be further restated as the following equation

$$\begin{aligned} &f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b \\ &subject\ to\ \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0\ and\ 0 \le \alpha_i, \alpha_i^* \le C \end{aligned} \quad (9)$$

where α_i and α_i^* are the Lagrange multipliers. The term $K(x_i, x_j)$ in Eq. (9) is defined as the kernel function, whose values are the inner product of two vectors x_i and x_j in the feature space $\varphi(x_i)$ and $\varphi(x_j)$. And bias b can be computed as follows:

$$b = \begin{cases} y_i - \sum_{j=1}^{n}(\alpha_i - \alpha_i^*)K(x_j, x_i) - \varepsilon & for\ \alpha_i \in (0, C) \\ y_i - \sum_{j=1}^{n}(\alpha_i - \alpha_i^*)K(x_j, x_i) + \varepsilon & for\ \alpha_i^* \in (0, C) \end{cases} \quad (10)$$

The kernel function has to satisfy Mercer's condition, and is intended to handle any dimension feature space without the need to explicitly calculate $\varphi(x)$. In this study, the Gaussian kernel function is chosen as the SVR's application mapping in this study.

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \quad (11)$$

where x_i and x_j are input vector space; σ^2 is the bandwidth of the kernel function.

In this study, an accurate and fast approach based on the GA and the simplex search techniques is presented to determine the optimal hyper-parameters of the SVR model. The implementation of the GA algorithm is realized with a GA toolbox developed by Chipperfield et al., and the simplex optimization method can be easily found from the Matlab optimization toolbox. The software LIBSVM was used to train and validate the developed SVR model.

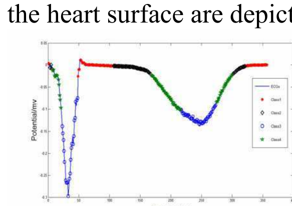
2.3. Simulation protocol and data set

The SVR model is tested with our previously developed realistic heart-torso model [3, 5]. In this investigation, a normal ventricular excitation is illustrated as an example to calculate the data set for the SVR model. The considered ventricular excitation period from first breakthrough to the end is 357 *ms* and the time step is 1*ms*, and thus, 358 BSPs φ_B and TMPs φm temporal data sets are calculated; In addition, the BSPs φ_B are added with 30dB simulated Gaussian white noise to mimic the clinical measurement noises. 60 data sets at times of 3*ms*, 9*ms*, 15*ms*,…, 357*ms* after the first ventricular breakthrough are used as testing samples to evaluate the generalization capacity of the proposed SVR model. The rest 298 in 358 data sets are employed as the training samples for building the SVR model.

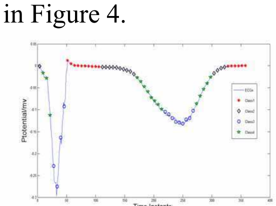
3. Results

According to MCC method, the 298traing samples are classified 4 clusters as shown in the Figure2(a), the number of the four clusters is 80, 74, 70, and 74 respectively. Then we can train the individual SVR model for each cluster, and the hyper-parameters are determined using the GA-Simplex method. For each testing sample in testing data, find the cluster to which it belongs to as shown in the Figure 2(b).

In this study, one sequential testing time points (27ms respectively after ventricle excitation) are presented to illustrate the performances of the reconstructed TMPs. The inverse ECG solutions are shown in Figure 3, it can be seen that the MCC-SVR method offers superior performances than the single SVR method, whose solution are more close to the simulated TMPs distributions. The time courses of the simulated TMPs and reconstructions for one representative source point on the heart surface are depicted in Figure 4.



(*a*)



(*b*)

Figure 2 on one epicardial point, (a) the 298 training samples are classified into four clusters by using MCC method; (b) for the 60 testing samples, find the cluster which it belongs to.

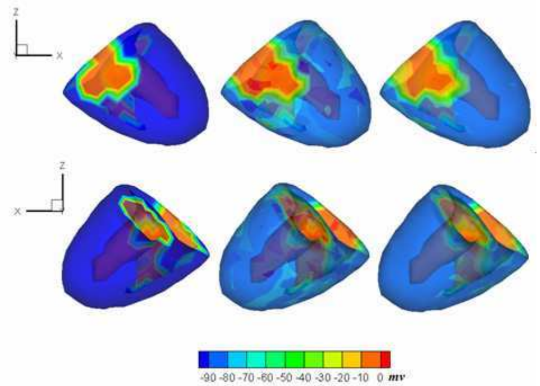


Figure 3 The TMPs distribution on the ventricular surface at 27 ms after the first ventricular breakthrough. The upper row shows the TMPs distribution from an anterior view and the lower from a posterior view. (a) the simulated TMPs; (b) the reconstruction with SVR method; (c) the reconstruction with the MCC-SVR method.

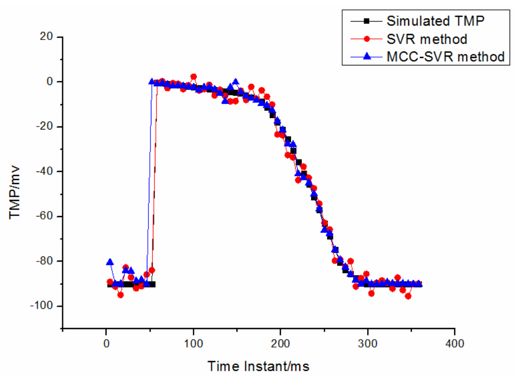


Figure 4 The time courses of the TMPs for one representative source point on the heart surface. The reconstruction TMPs over the 60 testing times with SVR method, and MCC-SVR method are all compared with those simulated TMPs.

4. Discussion and conclusion

The SVR method is a powerful technique to solve the nonlinear regression problem, and can server as a promising tool for performing the inverse reconstruction of the TMPs. Maximum margin clustering (MMC) is a recent large margin unsupervised learning approach that has often outperformed conventional clustering methods, which can be implemented by the iterative SVR method. In this paper, the MMC approach is adopted to cluster the training samples (different time instant BSPs), and the individual SVR model for each cluster is then constructed. For each testing sample, find the cluster to which it belongs, and then use the corresponding SVR model to reconstruct the TMPs. The reconstructed TMPs obtained by using the MCC-SVR and single SVR method are given in figures 3 and 4. When reconstructing the TMPs over the 60 testing samples, the results from the MCC-SVR method are superior to those from the single SVR method, which are close to the simulated TMPs.

In this paper, the MCC-SVR method is proposed for the inverse solutions of the ECG problem. The new algorithm was tested and compared with single SVR schemes using a realistic heart-torso model. The experimental results show that the MCC-SVR can improve the generalization performance of the single SVR in reconstructing the TMPs, leading to a more accurate reconstruction of the TMPs. According to these results, the MCC-SVR method can server as a promising tool for solving the nonlinear regression problem of the inverse ECG problem.

Acknowledgements

This work is supported in part by the 973 National Key Basic Research & Development Program (2010CB732502), China Postdoctoral Science Foundation (20090461376), and by the National Natural Science Foundation of China (30900322, 50875245 and 61101046). This work is also funded by Zhejiang Provincial Natural Science Foundation of China (Y2090398).

References

[1] Ramanathan C., Ghamen R., Jia P., Ryu K. and Rudy Y. Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. Nature Med. 2004; 10:422-8.

[2] van Dam P. M., et al. Non-invasive imaging of cardiac activation and recovery. Ann Biomed Eng. 2009; 37:1739-56.

[3] Jiang M., Xia L., Shou G., Wei Q., Liu F., and Crozier S. Effect of Cardiac Motion on Solution of the Electrocardiography Inverse Problem. IEEE Trans. Biomed. Eng. 2009; 56:923-931.

[4] Smola A J and Sch¨olkopf B A. tutorial on support vector regression Stat. Comput. 2004; 14:199–222

[5] Jiang M. F., Zhu L., Wang Y., Xia L., Shou G., Liu F. and Crozier S. Application of kernel principal component analysis and support vector regression for reconstruction of cardiac transmembrane potential. Phys. Med. Biol., 2011; 56:1727-1742.

[6] Xu L., Neufeld J., Larson B., and Schuurmans D. Maximum margin clustering. Advances in Neural Information Processing Systems. 2005, vol. 17, Cambridge, MA: MIT Press.

[7] Zhang K., Tsang I. W., and Kwok J. T. Maximum Margin Clustering Made Practical. IEEE Trans. Neural Networks, 2009; 20(4):583-596.

[8] Vapnik V. Statistical Learning Theory. John Wiley & Sons, NewYork, USA, 1998.

Address for correspondence.

Name: Mingfeng Jiang
Full postal address: No. 588 mailbox, Zhejiang Sci-Tech University, Hangzhou, China (310018)
E-mail address: jiang.mingfeng@hotmail.com