KKU ENGINEERING JOURNAL 2016;43(S3):494-498 Research Article

KKU Engineering Journal

https://www.tci-thaijo.org/index.php/kkuenj/index

Support vector regression-based synthesis of 12-lead ECG system from the standard 5 electrode system using lead V1

Piroon Kaewfoongrungsi and Daranee Hormdee*

Embedded System Research and Development Group, Department of Computer Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand.

Received April 2016
Accepted June 2016

Abstract

The standard 12-lead electrocardiogram (ECG) is a fundamental but very efficient clinical method for heart disease diagnosis. Measuring all 12 leads is often cumbersome and impractical especially on a long term monitoring. There have been ways to reduce the number of electrodes in ECG system also from 10 down to 5 or 6 electrodes and various regression methods were applied to derive back those 12-lead ECG. This paper presents how Support Vector Regression (SVR) was used to find a set of transfer function for deriving the 12-lead ECG from the standard 5-electrode setting using lead V1 system. All dataset used in this work has been obtained from PhysioNet database consisting of 4,810 samples. Five-fold cross-validation was applied to find the best parameter of SVR. Two kernel functions, RBF and ERBF, have been explored and evaluated in this work. The experiments strongly presented that SVR methodology was worth considerate for synthesizing the 12-lead ECG signals from the standard 5-lead electrode system using V1. The results also showed ERBF kernel function gave better RMSE (Root Mean Square Error) than RBF kernel function in the case here.

Keywords: 12-lead ECG system, 5-electrode system with lead V1, Support vector regression, RBF kernel function, ERBF kernel function, PhysioNet database

1. Introduction

Typically for measuring the standard 12-lead Electrocardiogram (ECG) for diagnosis of physician requires 9 electrodes to be placed strategically on the body and one electrode to be connected to ground as shown in Figure 1 [1-2].

Shown in Figure 1, the standard 12-lead ECG signals are call lead I, lead II, lead III, lead aVR, lead aVL, lead aVF, lead V1, lead V2, lead V3, lead V4, lead V5 and lead V6 signals.

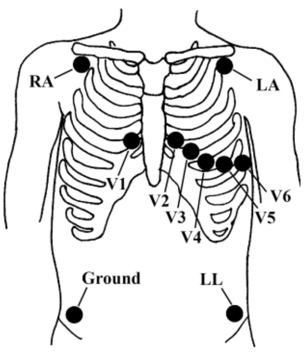


Figure 1 Standard 12-lead ECG

In 2002, D. Wei [3] has presented a method of minimizing number of electrodes which used for measuring the ECG signals by adding V1 and V6 signals on the top of the standard 4 leads. This can reduce the number of electrodes from 10 down to 6 as in Figure 2 [4].

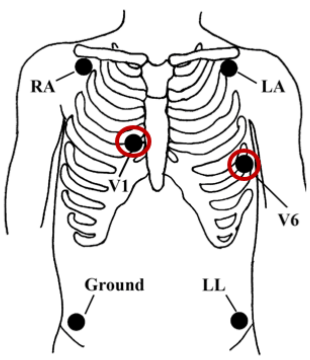


Figure 2 The 5-lead system for deriving 12-lead ECG [4]

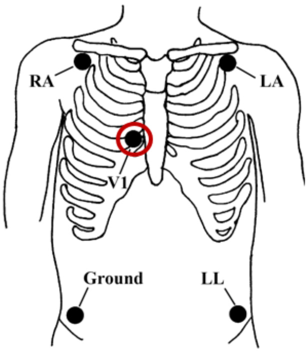
Likewise, there was another way to reduce the number of electrodes in ECG system also from 10 to 5 electrodes. Then regression method was applied to derive back those 12-lead ECG. This cuts the cost and increase mobility for long term

*Corresponding author. Tel.: +6683 360 3232
Email address: darhor@kku.ac.th; darhor@gmail.com
doi: 10.14456/kkuenj.2016.165

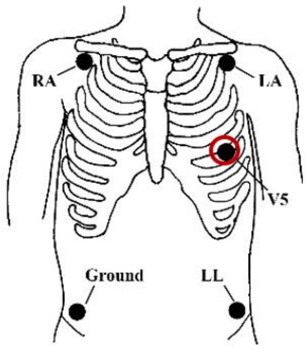
patients monitoring. There are two standard approaches shown in Figure 3 for placing electrodes for 5-lead ECG [5]. Lead V1 is used in Figure 3(a), while lead V5 is used in Figure 3(b). The choice of choosing either of these two standards is arbitrary.

The objective of this paper is to present how to apply support vector regression (SVR) method on standard 5-lead ECG signals according to Figure 3(a) to complete all 12-lead ECG signals.

In order to reduce the number of electrodes, the missing leads, lead V2 to lead V5, had to be derived from other existing leads. This is done by exploiting lead I, II, V1 and V6 signals as input data for the regression using the linear combination and least mean square method to find the coefficients of regression [4].



(a) 5-lead system with V1



(b) 5-lead system with V5

Figure 3 The lead system for deriving 12-lead ECG [4]

2. Literature reviews

2.1 Support vector regression method

Support Vector Regression (SVR) [6-8] in the past it has been used to solve nonlinear problems. Basic idea behind support vector regression is to map input data into higher dimensional space to map nonlinearity in original data as to perform linear in higher dimensional space using a kernel function and construct the separated hyper plane. The SVR function is shown in Equation (1).

$$f(X) = \langle W \cdot \mathrm{K}(X) \rangle + b \qquad (1)$$

Where W is the weight vector, X is the input column vector, K is the kernel function for mapping data to higher dimension, b is the bias value. The dataset used to train with SVR is $\{(X_i, Y_i)\}_{i=1}^{l}$, $X \in \mathbb{R}^n$, $Y \in \mathbb{R}$ where X_i is the input data vector, Y_i is desired output vector, X is input space, Y is output space. In the function $f(X)$ has the deviation value (ε) called "loss function" and all input data X_i that give value of $f(X)$ within $\pm\varepsilon$ interval are called "support vector" as shown in Figure 4.

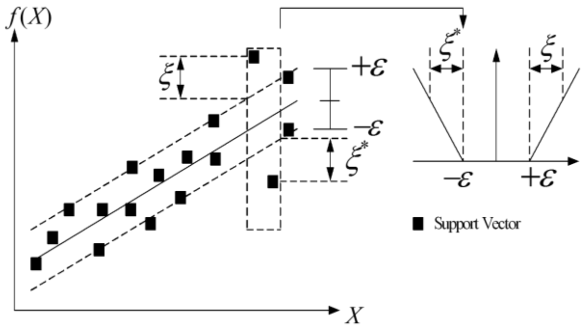


Figure 4 Soft margin $\varepsilon -$ insensitive in linear SVR

From Figure 4, the optimization was used to find weight vector (W) as in Equation (2-4).

$$\min_{W,\xi_i,\xi_i^*} \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{l}(\xi_i, +\xi_i^*) \qquad (2)$$

Subject to

$$\left.\begin{array}{c} Y_i - \langle W \cdot K(X, X_i)\rangle - b \le \varepsilon + \xi_i \\ \langle W \cdot K(X, X_i)\rangle + b - Y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{array}\right\}$$

Where C is a constant variable, W is weight vector obtained by solving with optimization problem as in Equation (3).

$$W = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X_i) \qquad (3)$$

Substitute Equation (3) into (1), the function $f(X)$ can be written as in Equation (4).

$$f(X) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X, X_i) + b \qquad (4)$$

Where $K(X, X_i)$ is a kernel mapping function between X and X_i.

The performance of SVR is majorly dependent on kernel function being used. In this paper, two kernels were used for mapping function to map input data to a higher dimension as in Equation (5-6). The parameter ε was set to 0.001 and parameter C was set to 5,000.

RBF kernel

$$K_{RBF}(X, X_i) = \exp(-\|X - X_i\|^2/2\sigma^2) \qquad (5)$$

ERBF kernel

$$K_{ERBF}(X, X_i) = \exp(-\|X - X_i\|/2\sigma^2) \qquad (6)$$

Where σ is the bandwidth of the kernel function.

2.2 Experimental methodology

The experiments were conducted to use synthesis methodologies for deriving the 12-lead ECG from 5-electrode system using lead V1. All dataset used in this work are obtained from PhysioNet database [9] consisting of 4,810 samples for each signal to shuffle data sets in order to prevent over fitting and using five-fold cross-validation, to find the best parameter of SVR.

These dataset are used as the training dataset and the other as the testing dataset. At the training process; all 5 lead signals are used in order to derive the transfer function. Then at the testing process; 7 lead signals (lead I, lead II, lead III, lead aVR, lead aVL, lead aVF and lead V1) represent the measured data. By substituting these signals into the derived function, all missing 5 lead signals (lead V2, lead V3, lead V4, lead V5 and lead V6) are obtained.

The following steps present how to derive the transfer function;

1) The total dataset from PhysioNet has been into two parts (90:10). The first '90%' part was used to find kernel parameters for SVR while the last '10%' part was used for blind test.

2) As five-fold cross-validation was utilized in this work, the first 90% dataset was then divided into 5 equal parts/folds. Each round, a single fold is used for testing, leaving the other 4 folds for training. In the nth round, fold#n is used for testing while the remaining folds are used for training. For instance, in the 2th round, fold#2 is used for testing while folds#1 and folds#3-5 are used for training. In total 5 rounds are processed. To find the average errors in the regression of each fold, the Root Mean Squared Error (RMSE) in the Equation (7) is used.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2} \tag{7}$$

Where A_t is the actual value in time t, F_t is the forecast value in time t and n is sample of testing set in each fold.

3) From all 5 folds, the RMSE values of the lead V2, lead V3, lead V4, lead V5 and lead V6 signals are considered. In order to find the transfer function of each signal, the fold that provides the minimum RMSE value of that signal must be identified. Then parameter σ from that fold was substituted into the equation of kernel function in Equation (5-6).

4) After obtaining them, the transfer function models of each signal was tested with blind test data of 10% to find RMSE value.

5) Finally in order to evaluate these transfer functions, the big test can then be started. By feeding the whole data set from those 4,810 data samples into these 5 transfer functions to get the calculated lead n signal, the RMSE and correlation coefficient values of each lead signal can be determined from the calculated signals and the ones from the PhysioNet dataset.

3. Results

The testing results with 5-fold cross-validation to find RMSE and correlation coefficient value of SVR using RBF and ERBF kernel function for deriving 5 signals from 5-electrode system using lead V1 are listed in Table1-2.

Table 1 RMSE (mV) with SVR using RBF kernel function

Signals	RBF kernel function				
	Root mean squared error (*RMSE*) mv				
	Fold#				
	1	*2*	*3*	*4*	*5*
Lead V2	22.695	19.560	18.427	22.195	20.034
Lead V3	37.649	29.058	27.212	30.159	30.817
Lead V4	33.312	29.199	32.020	31.788	29.328
Lead V5	22.893	21.511	20.078	20.357	22.144
Lead V6	17.675	21.640	16.117	17.768	20.986

Table 2 RMSE (mV) with SVR using ERBF kernel function

Signals	ERBF kernel function				
	Root mean squared error (*RMSE*) mv				
	Fold#				
	1	*2*	*3*	*4*	*5*
Lead V2	13.948	8.291	9.188	10.909	10.676
Lead V3	23.497	11.470	11.252	18.561	20.050
Lead V4	19.093	10.273	12.843	14.677	17.942
Lead V5	13.940	14.271	12.581	10.299	14.444
Lead V6	13.964	20.165	13.244	12.090	18.329

From Table 1 and 2, the minimum RMSE values of lead V2, lead V3, lead V4, lead V5 and lead V6 are highlighted of each fold. The parameter σ value of those folds with the minimum of RMSE value used for derived ECG of 5 signals with RBF and ERBF kernel function are shown in Table 3.

Table 3 The parameter σ of RBF and ERBF kernel function

Signals	RBF		ERBF	
	Fold#	σ	Fold#	σ
Lead V2	3	0.6	2	3
Lead V3	3	0.6	3	2
Lead V4	2	0.4	2	3
Lead V5	3	0.6	4	3
Lead V6	3	0.6	4	3

Then, the test using the transfer function models for each signal is evaluated with blind test data (the 10% part) and all 4,810 data samples to find RMSE and correlation coefficient values. The results are shown and compared in Table 4-5.

Table 4 RMSE (mV) tested with blind test data of 10% and 4,810 data samples

Signals	RBF kernel function		ERBF kernel function	
	RMSE (mV)		RMSE (mV)	
	Blind test data	All 4,810 data samples	Blind test data	All 4,810 data samples
Lead V2	20.359	18.987	8.817	4.663
Lead V3	28.671	27.816	12.515	6.253
Lead V4	29.684	27.962	11.911	6.082
Lead V5	21.824	19.503	13.170	6.189
Lead V6	16.536	14.777	15.984	7.300
Average	**23.415**	**21.809**	**12.479**	**6.097**

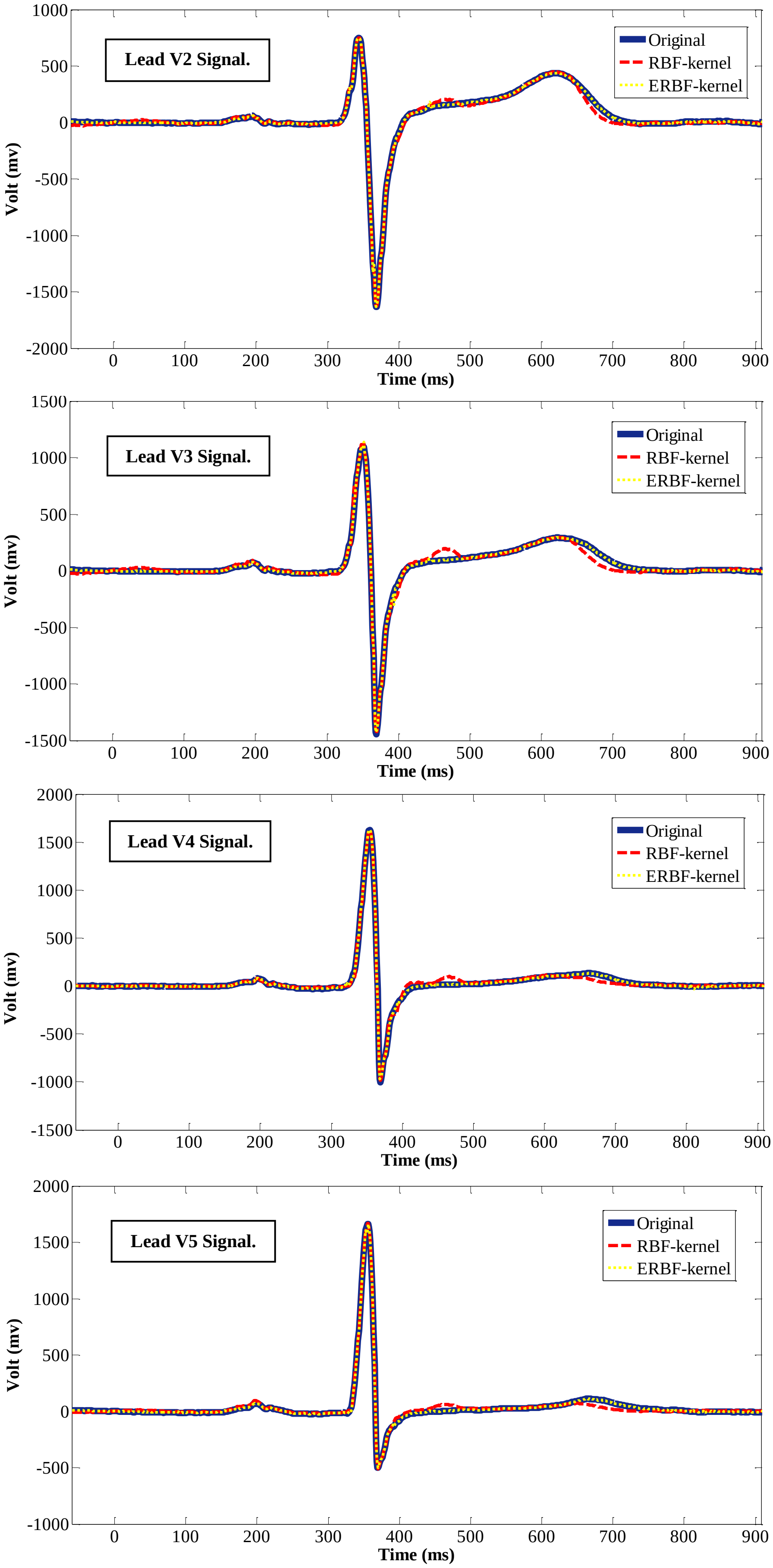


Figure 5 Derived vs original values of lead V2, V3, V4, V5 and V6 signals

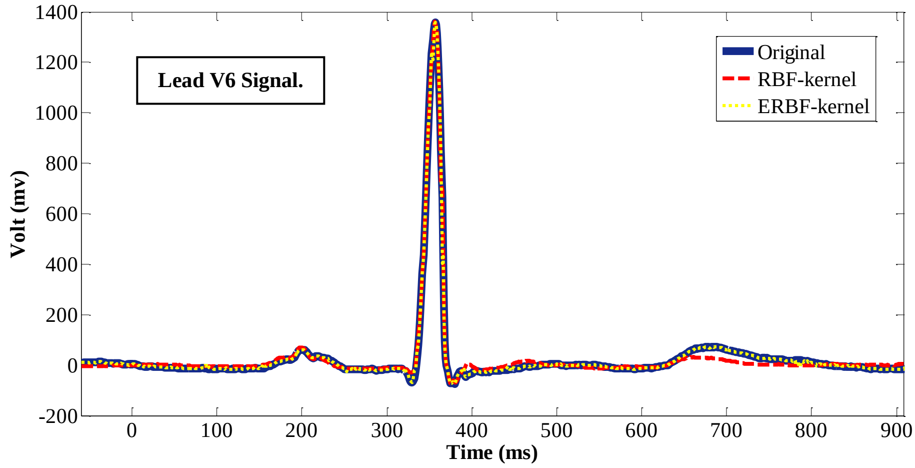


Figure 5 Derived vs original values of lead V2, V3, V4, V5 and V6 signals (Cont.)

Table 5 The correlation coefficient value tested with 4,810 data samples

Signals	RBF kernel function	ERBF kernel function
Lead V2	0.9966	0.9998
Lead V3	0.9911	0.9996
Lead V4	0.9928	0.9997
Lead V5	0.9948	0.9995
Lead V6	0.9948	0.9987

Lastly, the result graphs of lead V2, lead V3, lead V4, lead V5 and lead V6 signals measured using standard 12-lead ECG method versus derived from 5-electrode system using lead V1 by SVR with RBF and ERBF kernel function are shown in Figure 5.

4. Conclusions

This paper has presented a finding of transfer function models for deriving 5 missing ECG signals (lead V2, lead V3, lead V4, lead V5, and lead V6) by SVR method from 7 measured signals (lead I, lead II, lead III, lead aVR, lead aVL, lead aVF and lead V1) with standard 5 lead ECG system using lead V1.

Two kernel functions, RBF and ERBF, have been evaluated here. The experimental results from Table 4 and Figure 5 showed the best performance in this work, was obtained from the SVR using ERBF kernel function method for all derived signal cases.

As for correlation coefficients from Table 5, the highest values were also obtained from SVR using ERBF kernel function.

Therefore, it is obvious to conclude that nonlinear regression with SVR using ERBF kernel function is worth chosen for deriving 12-lead ECG system from the standard 5 electrode system using lead V1. However, these five transfer functions from the SVR using ERBF kernel function method obtained here are suitable with the ECG data from PhysioNet database only. In order to apply this method on some other practical data, the new dataset from that actual machine is needed. Then the same old processes must be repeated including training process with 12-lead ECG signals for deriving new transfer functions.

5. References

[1] Berbari EJ. Principles of Electrocardiography. In: Bronzino JD, editor. The Biomedical Engineering Handbook 2nd ed. Boca Raton: CRC Press LLC; 2000. p. 231-240.

[2] www.bem.fi [Internet]. New York: Bioelectromagnetism - principles and Applications of Bioelectric and Biomagnetic Fields, [updated 2014; cited 2016 Mar 26]. Available from: http://www.bem.fi/book/.

[3] Wei D. Deriving The 12-lead Electrocardiogram From Four Standard Leads using Information Redundancy in The 12-lead System. International Journal of Bioelectromagnetism 2002;4:127-128.

[4] Frank E. The image surface of a homogeneous torso. Am Heart J 1954;47:757-768.

[5] www.nursecom.com [Internet]. Canada: The 5 Lead Wire ECG System [updated 2003; cited 2016 Mar 26]. Available from: http://www.nursecom.com/ECGprimer.pdf.

[6] Smola AJ, Scholkopf B. A tutorial on support vector regression. Technical report. ESPRIT working group in Neural and Computational Leaning; 1998 Oct. No.: NC2-TR-1998-030.

[7] Theera-Umpon N, Boonprasert U. Leeway Prediction of Oceanic Disastrous Target via Support Vector Regression. JACIII 2004;8:606-612.

[8] Yodjaiphet A, Theera-Umpon N, Auephanwiriyakul S. Electrocardiogram Reconstruction using Support Vector Regression. IEEE International Symposium on Signal Processing and Information Technology (2012); 2012 Dec 12-15; Ho Chi Minh, Vietnam. New Jersey: IEEE; 2013. p. 269-273.

[9] PhysioNet. PhysioNet Database. [Internet]. 2007 [cited 2016 Mar 26]. Available from: http://www.Physionet.org/challenge/2007/data/.