Can the Regression Trees Be Used to Model Relation Between ECG Leads?

Ivan Tomasic[1(✉)], Roman Trobec[2], and Maria Lindén[1]

[1] Biomedical Engineering Research Group, Mälardalen University, Box 883 721 23 Västerås, Sweden
{ivan.tomasic,maria.linden}@mdh.se

[2] Department of Communication Systems, Jožef Stefan Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia
roman.trobec@ijs.si

Abstract. Presented is a preliminary study that investigates regression trees application for the purpose of mapping relationship between three differential ECG leads and leads of the 12-lead ECG. The approach was evaluated on a single ECG measurement on which it was superior to two syntheses performed by universal and personalized linear transformations, in terms of correlation coefficients between the synthesized and measured leads. A prominent imperfection however is that the regression trees can output only a limited number of values equal to the number of leaf nodes. The paper indicates some ideas on how to overcome this deficiency.

Keywords: ECG · Leads synthesis · Derived electrocardiograms · Regression trees · Differential leads · Electrocardiography · Wireless electrodes

1 Introduction

Both linear and nonlinear methods have been used before to model relation between electrocardiographic (ECG) leads for the purpose of leads' syntheses [1, 2]. Among the nonlinear methods mostly neural networks were used [3], whereas for the linear mapping, linear regression is the most often used method. Recent publications in addition to multiple linear regression [4], report using state-space model [5], support vector regression [6], multi-scale linear regression [7], and combination of methods [8, 9] for the ECG leads synthesis. In this work we will investigate the usage of regression trees for the same purpose.

In our previous studies, we have shown that it is possible to synthesize high quality 12-lead ECGs from three differential leads (DLs) [10]. Differential leads are bipolar leads that measure the potential between two closely placed body-surface electrodes. DL measurements can be obtained by so called wireless body electrodes (WBEs) – novel devices that enable minimal obtrusion and wireless transmission of recorded signals [11].

By using the algorithm for selection of optimal DLs [10], we have identified optimal universal positions of three DLs (see Fig. 1) from which 12-lead ECG can be synthesized

B. Mandler et al. (Eds.) IoT 360° 2015, Part I, LNICST 169, pp. 467–472, 2016.
DOI: 10.1007/978-3-319-47063-4_50

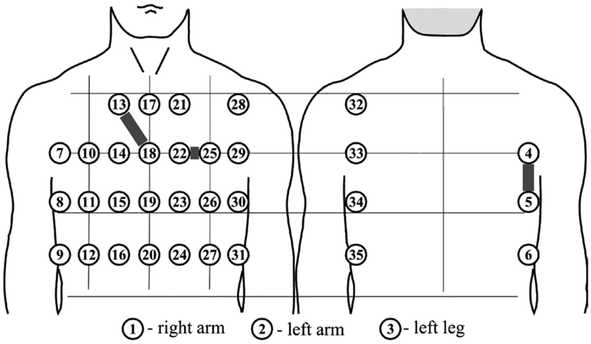

Fig. 1. Schematic locations of multichannel ECG (MECG) electrodes on the chest (left) and the back (right). The *gray lines* represent the calculated optimal universal leads for the 12-lead ECG synthesis: {(13, 18), (22, 25), (4, 5)}.

by employing universal or personalized linear transformations [12] that map the three DLs to the leads of the 12-lead ECG. In this study we investigate how regression trees can be used to map the same relation.

2 Methods

2.1 Studied Data

A single MECG measurement was obtained from a patient scheduled for a coronary artery bypass surgery. The data recording device and data acquisition procedure is described in [13]. The measurement was obtained during our previous study [14]. The positions of MECG electrodes are specified in Fig. 1.

The length of the measurement was 360 s. The measurement was processed using MATLAB (MathWorks, Inc.) where first the baseline wander was removed by interpolating a cubic spline through the isoelectric points of each MECG lead and subtracting it from the corresponding lead.

Subsequently, 10-second interval was randomly extracted from the MECG measurement and associated target 12-lead ECG obtained simultaneously. The extracted interval was filtered by a low-pass filter with a cutoff frequency of 40 Hz, an attenuation of 60 dB, and a stop frequency of 100 Hz. First seven seconds subinterval of the extracted interval was used for building the regression trees, whereas the remaining three seconds were used for the evaluation.

2.2 Regression Trees

Regression trees are simple but effective method of fitting a set of numeric input variables to a single numeric output variable. The general concept of regression trees is to partition the space defined by input variables and fit a simple model (usually a constant) in each one of the partitions [15]. In addition to regression trees, which give numeric responses, there are also classification trees, which give responses that are nominal, such as *true* or *false*.

To create regression trees between the three optimal DLs and the 12-lead ECG we have used MATLAB's Statistics and Machine Learning Toolbox™ trees which are binary. The details of how MATLAB creates trees can be found in the MATLAB's documentation [16]. All the parameter of function that creates trees were left on default values.

2.3 Experimental Procedure

The three optimal differential leads were calculated from the input MECG data (10 s intervals) by taking differences between appropriate MECG unipolar leads, e.g. DL (13, 18) is obtained as a difference between MECG leads 13 and 18. A regression tree was then built between the three DLs and each of the leads of the 12-lead ECG which makes 12 trees all together. The trees were examined for the number of nodes created.

To evaluate the performance of the created trees they were used to synthesize the 12-lead ECG (in regression trees terminology: "predict the response") on the evaluation three second subintervals. The linear correlation coefficients (CCs) between the 12-lead ECG synthesized by using trees, and the measured target 12-lead ECG, were compared to the CCs obtained by the universal and personalized transformation matrix. The synthesized 12-lead ECG was also compared to the target measured 12-lead ECG visually.

3 Results

Table 1. shows the CCs between the synthesized and measured (target) leads for the tree synthesis methods on the evaluation interval. Additionally, the synthesized 12-lead ECG obtained from the regression trees, are compared with the measured 12-lead ECG, graphically in Fig. 2. whereas a closer look on a segment from lead I is presented in Fig. 3.

Table 1. Correlation between the synthesized leads and the measured leads for universal linear transformation, for personalized linear transformation, and for regression trees.

Lead	Universal transformation	Personalized transformation	Regression trees
I	0.9543	0.9940	0.9920
II	0.9446	0.9485	0.9610
III	0.8716	0.9885	0.9889
aVR	0.9644	0.9918	0.9903
aVL	0.9332	0.9933	0.9919
aVF	0.6133	0.9473	0.9686
V1	0.9668	0.9796	0.9918
V2	0.9604	0.9943	0.9946
V3	0.9087	0.9876	0.9968
V4	0.8447	0.9718	0.9939
V5	0.8967	0.9681	0.9912
V6	0.9752	0.9823	0.9894

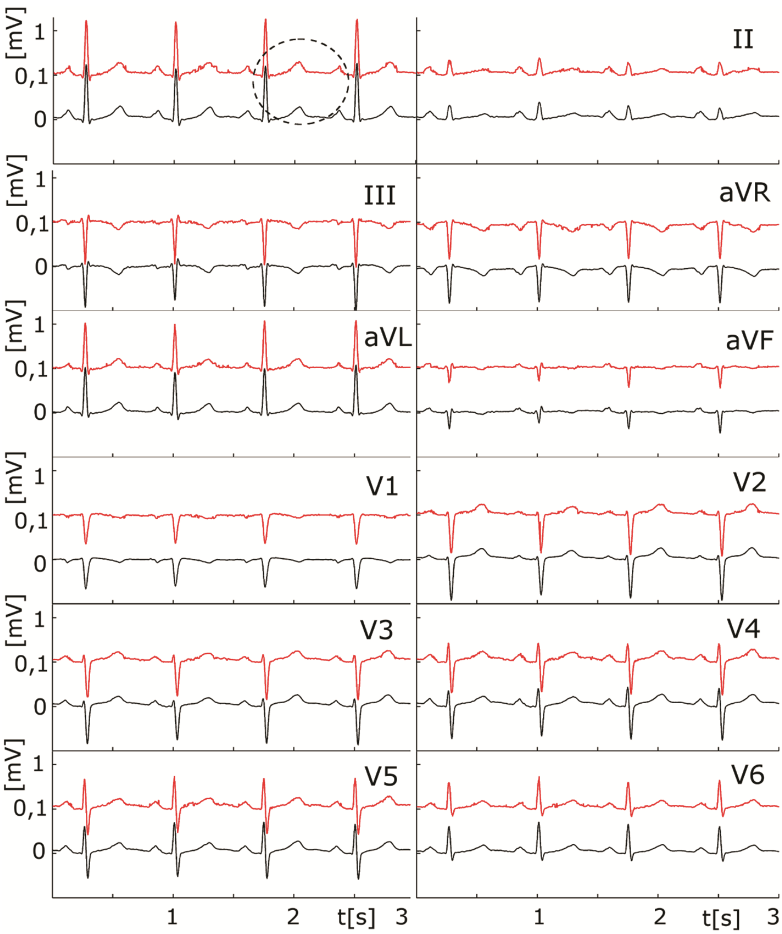

Fig. 2. Target (below) and synthesized (above) 12-lead ECG. On lead I, a T-wave that is zoomed in on Fig. 3., is approximately marked with a dashed ellipse.

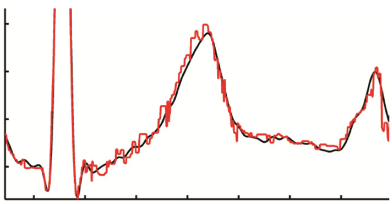

Fig. 3. Target (smooth) and synthesized (coarse) signals on a segment form lead I approximately marked in Fig. 2. (The two signals were moved one on top the other for easier comparison.).

4 Discussion

Table 1. reveals that even though the CCs between the synthesized and the measured leads are high for all the leads and all three synthesis methods, the synthesis with the regression trees is still superior to the other two methods. In only three leads has personalized transformation outperformed the regression trees (leads I, aVR, aVL) but in those situations the CCs for both methods are larger than 0.99.

In Fig. 2 the leads synthesized by regression trees seem almost identical to the measured leads of the 12-lead ECG, but a closer inspection (Fig. 3) reveals that the synthesized leads are coarser that the measured leads. This is because the output from regression trees can take only limited number of values, i.e. the number of output values is equal to the number of leafs a tree has. In our experiment the trees had number of leafs in a range from 1302 to 1374.

The possibilities for improving the synthesis output form the regression trees would be to increase the number of output leafs which will in turn increase the number of levels the predicted value can obtain, and/or to use a filter to smooth the output of the synthesis.

5 Conclusion

We have used regression trees to map relations between three DLs and the leads of the 12-lead ECG, for the purpose to synthesize the 12-lead ECG. The approach was evaluated by using one MECG measurement and has shown to be superior to the universal and personalized linear transformations used in previous investigations.

This is a preliminary research which shows that the regression trees have a potential in modeling relations between ECG leads. Further work is needed to investigate their performance on more measurements from different subjects, and to find satisfying solution for smoothing the output of regression trees.

The proposed synthesis of the 12-lead ECG from the 3 DLs enables new applications of the WBEs in Internet of Things (IoT) supported long-term remote health monitoring, because it enables a significant data reduction since only three leads have to be communicated, instead of eight independent leads in the standard 12-lead ECG.

Acknowledgments. This work was supported by the Swedish Knowledge Foundation (KKS) financed research profile Embedded sensor systems for health (ESS-H), and by the distributed environment Ecare@Home (2015-2019) funded also by KKS.

References

1. Tomasic, I., Trobec, R.: Electrocardiographic systems with reduced numbers of leads-synthesis of the 12-Lead ECG. IEEE Rev. Biomed. Eng. **7**, 126–142 (2014)
2. Vozda, M., Cerny, M.: Methods for derivation of orthogonal leads from 12-lead electrocardiogram: a review. Biomed. Sig. Process. Control **19**, 23–34 (2015)

3. Atoui, H., Fayn, J., Rubel, P.: A novel neural-network model for deriving standard 12-lead ECGs from serial three-lead ECGs: application to self-care. IEEE Trans. Inf. Technol. Biomed. **14**, 883–890 (2010)
4. Figueiredo, C.P., Mendes, P.M.: Towards wearable and continuous 12-lead electrocardiogram monitoring: synthesis of the 12-lead electrocardiogram using 3 wireless single-lead sensors. In: International Conference on Biomedical Electronics and Devices (BIODEVICES), pp. 329–332 (2012)
5. Lee, J., Kim, M., Kim, J.: Reconstruction of precordial lead electrocardiogram from limb leads using state-space model. IEEE J. Biomed. Health Inform. **20**, 818–828 (2015)
6. Yodjaiphet, A., Theera-Umpon, N., Auephanwiriyakul, S.: Electrocardiogram reconstruction using support vector regression. In: IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 269–272. IEEE (2012)
7. Nallikuzhy, J.J., Dandapat, S.: Enhancement of the spatial resolution of ECG using multi-scale Linear Regression. In: Twenty First National Conference on Communications (NCC), pp. 1–6. IEEE (2015)
8. Acharyya, A., Schiariti, M., Puddu, P.E.: Personalized reduced 3-lead system formation methodology for remote health monitoring applications and reconstruction of standard 12-lead system. Int. Arch. Med. **8**, 1–15 (2015)
9. Hsu, C., Wu, S.: Robust signal synthesis of the 12-lead ECG using 3-lead wireless ECG systems. In: First IEEE International Conference on Communications (ICC), pp. 3517–3522 (2014)
10. Trobec, R., Tomašić, I.: Synthesis of the 12-lead electrocardiogram from differential leads. IEEE Trans. Inf. Technol. Biomed. **15**, 615–621 (2011)
11. Tomasic, I., Avbelj, V., Trobec, R.: Smart wireless sensor for physiological monitoring. In: Studies in Health Technology and Informatics, pHealth 2015, pp. 295–301. IOS Press (2015)
12. Tomasic, I., Frljak, S., Trobec, R.: Estimating the universal positions of wireless body electrodes for measuring cardiac electrical activity. IEEE Trans. Biomed. Eng. **60**, 3368–3374 (2013)
13. Avbelj, V., Trobec, R., Gersak, B., Vokac, D.: Multichannel ECG measurement system. In: Proceedings of the 10th IEEE Symposium on Computer Based Medical Systems, pp. 81–84 (1997)
14. Frljak, S., Avbelj, V., Trobec, R., Meglic, B., Ujiie, T., Gersak, B.: Beat-to-beat QT interval variability before and after cardiac surgery. Comput. Biol. Med. **33**, 267–276 (2003)
15. Hastie, T., Tibshirani, R., Friedman, J.: Tree-based methods. The Elements of StatisticalLearning: Data Mining, Inference, and Prediction, 2nd edn., pp. 305–317. Springer Science & Business Media (2009). doi:10.1007/b94608. ISBN: 978-0-387-84857-0
16. ©The MathWorks Inc.: Classification Trees and Regression Trees. In: Statistics and Machine Learning Toolbox™ Documentation (R2014a) (2015)