Prior Model Selection in Bayesian MAP Estimation-Based ECG Reconstruction

Ege Ozkoc, [1]Elifnur Sunger, Kutay Ugurlu, Yesim Serinagaoglu Dogrusoz

Electrical and Electronics Engineering Department, Middle East Technical University, Ankara, Turkey

Email: ege.ozkoc@metu.edu.tr, elifnur.sunger@metu.edu.tr, kutay.ugurlu@metu.edu.tr, yserin@metu.edu.tr

Abstract. *The inverse problem of electrocardiography (ECG) aims to reconstruct cardiac electrical activity using body surface potential measurements and a mathematical model of the body. However, this problem is ill-posed; therefore, it is essential to use prior information and regularize the solution to get an accurate solution. A statistical estimation has been applied to the inverse ECG problem with success, but a "good" a priori probability model is required. In this study, the Bayesian Maximum A Posteriori (MAP) estimation method is applied for solving the inverse ECG problem. Several prior models (training sets) are constructed, and the corresponding results are evaluated in terms of electrogram reconstruction, activation time estimation and pacing site localization accuracy. Our results showed that the training data consisting of beats from the 1st or 2nd neighbors of the test beat pacing nodes resulted in more successful results, implying that the prior models, including moderate amount and coverage of training data, might lead to an improved reconstruction of electrograms.*

Keywords: Electrocardiography, Inverse Problem, Bayesian MAP Estimation

1. Introduction

The objective of electrocardiographic imaging (ECGI) is to reconstruct the electrical activity of the heart using the body surface potentials and a mathematical model of the torso [1]. It provides high-resolution functional images of cardiac electrical activity. However, this is an ill-posed problem and to get a meaningful solution, the solution should be regularized. One of the most widely used regularization methods is Tikhonov regularization, which makes a trade-off between fidelity to the measurements and a good fit to an a priori constraint [2].

Alternatively, statistical estimation methods could also be used where the solution is represented in terms of probability distributions [3],[4]. The most widely used statistical methods are Bayesian Maximum A Posteriori (MAP) estimation and Kalman filtering. Kalman filtering uses an algorithm where the measurements observed over time are used to estimate the solution by representing the solution as state-space formulation [3]. In Bayesian MAP estimation, the solution is chosen to maximize the posterior probability density function (pdf) of the electrograms [4]. However, choosing a "good" prior pdf is still a challenge in the application of the MAP approach to ECGI.

In this study, we compare different prior models based on different training sets in terms of their performance on MAP-based ECGI. Bayesian MAP estimation is applied to 3 different test beats, and solutions are obtained for five different training sets for each test beat. These solutions are examined and compared with each other. Then, how to choose the training set to get the best result is discussed.

ISBN 978-80-972629-5-2

2. Subject and Methods

Problem Definition

The problem can be described by the equation $y_i = Ax_i + n_i$ where y_i, x_i and n_i denote the BSP measurement vector, epicardial surface data to be reconstructed, and noise vector, respectively, at time instance i, and A is the forward transfer matrix. The forward matrix A is found using the Boundary Element Method (BEM) in a homogeneous torso [1].

Bayesian MAP Estimation

Solving the problem at each time instant separately and dropping the time index, the ECGI solution is chosen to maximize the posterior pdf of the sources, which can be expressed as:

$$\hat{x}_{MAP} = \underset{x}{argmax}\, \boldsymbol{p}(x|y) = \underset{x}{argmax} \frac{\boldsymbol{p}(y|x)p(x)}{\boldsymbol{p}(y)} . \quad (1)$$

In this study, the noise is assumed to be Gaussian zero-mean independent identically distributed ($n \sim N(0,\ C_n)$). The epicardial potentials are also assumed to have Gaussian distribution ($x \sim N(\mu_x,\ C_x)$). Thus, the following expression can be obtained as the MAP solution:

$$\hat{x}_{MAP} = \left(A^T {C_n}^{-1} A + {C_x}^{-1}\right)^{-1} \left({C_x}^{-1}\mu_x + A^T {C_n}^{-1} y\right), \quad (2)$$

The prior mean (μ_x) and the covariance matrix (C_x) are estimated from a training dataset consisting of previously available epicardial potentials [4].

Experimental Data

Both the training and test data constitute measurements taken from Utah tank experiments [5]. Sock electrodes with 490 nodes on the epicardial surface were utilized to measure the electrograms (EGM), while BSPs were simulated from these EGMs at 192 electrodes on the torso surface at 30 dB SNR. Three different test beats from two different datasets were used to compare the reconstruction accuracy of the MAP approach. These three test beats, whose pacing nodes have a higher number of neighboring pacing locations, were used. The QRS segment of the test beat was reconstructed using the generated training data.

For each of the three test beats, five different training datasets were generated. Each of the training datasets includes training beats paced from the neighbors of the test beat pacing node up to the respective order given by the training set name. To illustrate, Training Set 3 includes beats initiated from the pacing locations up to the third-order neighborhood of the test beat pacing site. Using the heart surface mesh, first-order neighbors have direct connectivity to the node of interest, second-order neighbors constitute the first order neighbors of the node of interest combined with their first order neighbors, etc. The pacing site of the test beat and those of the training beats are illustrated in Fig. 1. The metrics describing the proximity of the training data pacing nodes to the test data pacing location and the coverage do not remarkably change among the test datasets. The mean distances to the pacing node for Training Set 1 to 5 are about 5 mm, 7.8mm, 10mm, 13mm, and 15.7mm. The numbers of beats in the training sets generated for Test Beat-1 and Test Beat-3 are close to each other, whereas, for Test Beat-2, the number of training beats is significantly larger.

3. Results

Reconstructed potentials were evaluated using Pearson's correlation; spatial correlation coefficients (sCC) for each time instant over all leads, and temporal correlation coefficients (tCC) for each lead over all time instants were computed. Activation times (ATs) were estimated using a spatio-temporal approach [6]. The accuracy of AT was evaluated by using

ISBN 978-80-972629-5-2

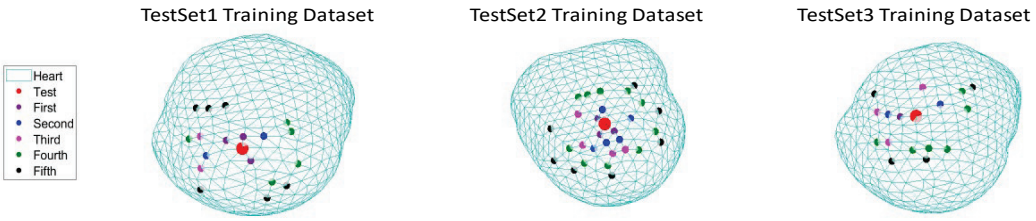

Fig. 1. Neighborhood map of pacing nodes regarding the training and test data.

Pearson's CC. The pacing site was assigned as the earliest activated node. Localization error (LE) was computed as the Euclidean distance between the true and estimated pacing sites.

Fig. 2 presents the box plots for sCC and tCC results for the three test beats and their respective five training data. Test Beat-1 has the lowest median sCC and tCC values for Training Data-1. Other training datasets have similar median values and interquartile ranges (IQRs). All the five training datasets used in the reconstruction of potentials of Test Beat-2 give similar tCC results. However, Training Data-1 and 2 provide slightly smaller IQR for sCC. Therefore, Training Data-1 and 2 outperform other training datasets. For Test Beat-3, median tCC and sCC decrease going from Training Data-1 to Training Data-5. Training Data-1 has a wider IQR than Training Data-2 for sCC results.

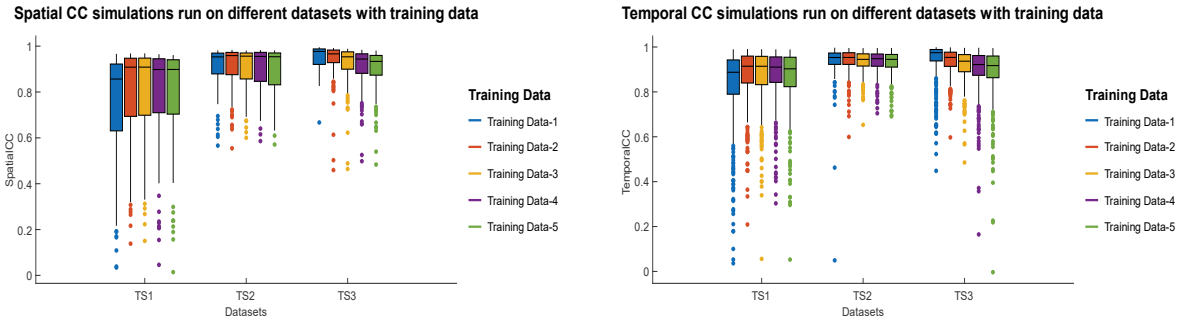

Fig. 2. Box plots for spatial and temporal correlation coefficients between three test beats (TS1, TS2 and TS3) and their Bayesian MAP estimations with five training datasets.

Table 1 and 2 present the CC values for the reconstructed ATs and the LE (mm), respectively. CC results for AT estimation are not significantly affected by choice of training datasets. Unlike the performance in reconstructing the epicardial potential maps, the localization error was not observed to be strictly correlated with the training set choice. To illustrate, the localization error obtained using Training Data-1 is approximately half of that of Training Data-2, although they are equally successful considering sCC and tCC as the evaluation metric.

Table 1. CC values for ATs for three test beats with five training datasets.

Test Beats	Training Data-1	Training Data-2	Training Data-3	Training Data-4	Training Data-5
Test Beat-1	0.95	0.96	0.95	0.95	0.95
Test Beat-2	0.93	0.94	0.93	0.94	0.94
Test Beat-3	0.97	0.98	0.98	0.97	0.96

ISBN 978-80-972629-5-2

Table 2. Localization errors (in mm) for three test beats with five training datasets.

Test Beats	Training Data-1	Training Data-2	Training Data-3	Training Data-4	Training Data-5
Test Beat-1	12.98	9.50	9.50	9.41	9.50
Test Beat-2	4.27	9.00	8.67	8.67	8.32
Test Beat-3	24.94	9.56	13.38	13.36	16.74

4. Discussion & Conclusion

In this study, we applied the Bayesian MAP estimation to ECGI to evaluate the best-case scenario where we know the pacing node of the test beats from the measured epicardial potentials. We observed that the selection of the training set is essential to get a good result. Including the nodes up to the first or second neighbors of the training set is generally enough to obtain a good estimation. As further nodes are included, the result slightly gets worse but are still adequate in terms of localization. This can be attributed to the fact that the prior pdf may get less successful at representing the local features when the training beats are taken further away from the pacing location. However, the decrease in the performance is very small even with Training Data-5. We did not observe any contradiction in the behavior of spatial and temporal CCs, so considering only one of them to check the results was enough. The seemingly contradictory relation observed in some cases between LE and the other metrics can be attributed to the inaccurate performance of the AT estimation algorithm. The accuracy of the algorithm needs to be improved for a more reliable evaluation of the ECGI methods. Nevertheless, Bayesian MAP estimation has some limitations in clinical tasks since the epicardial covariance matrices cannot be estimated in the absence of epicardial measurements. Tikhonov regularization, which does not require any prior knowledge [2], has an advantage in this sense with a trade-off of reconstruction success. The success of the prior model constructed with simulated training data will be investigated to overcome this problem.

Acknowledgements

This work was supported by the grant: TUBITAK-118E244. The authors thank Dr. Robert S. MacLeod and his colleagues from the University of Utah for the data used in this study. All authors have contributed equally to this study.

References

[1] Gulrajani, R.M. (1998). The forward and inverse problems of electrocardiography. *IEEE Engineering in Medicine and Biology Magazine*, 17(5), 84–101.

[2] Cluitmans, M.J.M., *et al.* (2017). In vivo validation of electrocardiographic imaging, *JACC: Clinical Electrophysiology*, 3(3), 232–242.

[3] Erenler, T., Serinagaoglu Dogrusoz, Y. (2019). ML and MAP estimation of parameters for the Kalman filter and smoother applied to electrocardiographic imaging. *Medical & Biological Engineering & Computing*, 57(10), 2093–2113.

[4] Serinagaoglu Y., *et al.* (2005). Bayesian solutions and performance analysis in bioelectric inverse problems, *IEEE Transactions on Biomedical Engineering*, 52(6), 1009–1020.

[5] MacLeod R. S., *et al.* (1995). Electrocardiographic mapping in a realistic torso tank preparation. In *Proceedings of 17th International Conference of the EMBS*, 45–246.

[6] Erem B., *et al.* (2014). Using transmural regularization and dynamic modeling for noninvasive cardiac potential imaging of endocardial pacing with imprecise thoracic geometry, *IEEE Transactions on Medical Imaging*, 33(3), 726-738.