

Computational Reconstruction of Electrocardiogram Lead Placement

Alexander D Wissner-Gross^{1,2}, Suraj Kapa³, James Lee⁴, Desmond B Keenan⁴, Natasha Drapeau⁴,
Kenneth Londoner⁴

¹Institute for Applied Computational Science, Harvard University, Cambridge, Massachusetts, USA

²Reified, Cambridge, Massachusetts, USA

³Mayo Clinic, Rochester, Minnesota, USA

⁴BioSig Technologies, Westport, Connecticut, USA

Abstract

We present a method for computationally reconstructing the spatial placement of electrocardiogram (ECG) leads using only correlations between their recorded signals and without requiring external calibration or other prior knowledge. We then apply our method to 12-lead ECGs obtained from the training dataset of the PhysioNet 2020 Challenge and examine the association of various cardiac abnormalities with the reconstructed geometries. Finally, we review potential clinical applications of our method, including automated recommendation of optimal lead placement, simplified visual summarization of ECG recordings, and improved automated classification of patient conditions.

1. Introduction

Machine learning approaches to analysis of biomedical data have been of increasing interest for the purposes of both streamlining human interpretation and attaining novel insights from large and often complex data sets [1]. Many of the principles underlying bioelectrical signals, in particular, contain an inherent logic. For example, the pattern of electrical deflections seen on a 12-lead electrocardiogram (ECG), assuming an ellipsoid heart with a single wavefront of activation, will contain within it a logical pattern defined by the locations of individual patches. During acquisition of the 12-lead ECG, misplacement of patches may similarly be identified by expert human interpreters due to illogical deviations from standard activation patterns. Thus, the inherent logic contained within the ECG may be one from which machine learning-based approaches may be applied to predict, without supervision, where leads may have been individually placed relative to the heart. Such a process may be valuable in that it may facilitate automated approaches to recognizing lead misplacement or offer further value in broadly identifying specific cardiac

conditions leading to deviations from normal activation patterns beyond rhythm and interval analysis (e.g., myocardial fibrosis, malrotation syndromes, etc.). Here, we present the use of a computational approach to reconstruct spatial placement of ECG leads without requiring prior human annotation or prior definitions.

2. Methods

Our analysis was focused on 12-lead ECGs obtained from the training dataset of the PhysioNet 2020 Challenge. Each ECG was sampled at 500 Hz and had a duration of 5-20 sec, and the set spanned a range of diagnosed cardiac conditions. We focused, in particular, on channels corresponding to precordial leads (V1-V6) given their distinctive geometric placement along a simple curve around the chest.

For our preliminary analysis, initial 6x6 cross-correlation matrices were obtained by examining individual ECGs in their entirety and plotting the trajectory in potential-potential phase space of each pair of precordial channels during the recording. Initial analysis focused on the first patient in the training dataset with one of nine isolated diagnosed conditions from the set: Normal, Right Bundle Branch Block (RBBB), Atrial Fibrillation (AF), Premature Ventricular Contraction (PVC), ST Depression (STD), Left Bundle Branch Block (LBBB), ST Elevation (STE), Atrioventricular Block (AVB), and Premature Atrial Contraction (PAC). No patients with more than one diagnosed condition were considered at this stage.

For the next stage in our analytical pipeline, cross-correlative lag analysis, we computed the normalized cross-correlation $c_{a,b}$ between channels a and b over the course of an entire ECG recording:

$$c_{a,b}[k] = \frac{\sum_n a[n+k]b[n]}{\sum_n 1_a[n+k]1_b[n]},$$

where $1_a, 1_b$ denote one-valued time series with the same respective durations as a, b . We then extracted the lag k

between -30 ms and +30 ms for which the absolute value of this cross-correlation was maximized.

For the third stage of our analysis pipeline, we treated the computed pairwise lags between precordial channels as the edge weights on a graph whose nodes corresponded to the six precordial channels (V1-V6). We then applied the NEATO graph layout solver [2] to that weighted graph to attempt to estimate a 2-D geometric arrangement of leads such that the distance between nodes was proportional to the corresponding edge weights.

For the fourth and final stage of our analysis pipeline, we prepared 36-dimensional vectors for the first 4,000 patients in the training set, each consisting of the pairwise lags computed earlier. We then trained linear support vector machine classifiers on balanced sets of patients to predict whether or not a given patient had been diagnosed with that condition.

3. Results

3.1. Cross-correlations between leads

We first show correlation plot matrices across 6 precordial leads for the first patient with each isolated condition in the PhysioNet 2020 Challenge Training Dataset (see Figure 1). Shown along the diagonal is a channel crossed with itself, resulting in a linear scatterplot. Note that as distance from the diagonal increases, the distance between electrodes increases, and the phase diagrams tend to become more circular, which is indicative of an increased phase difference between the channel signals. Also of note is that the normal patient had more circular and less noisy phase loops than the patients with cardiac abnormalities.

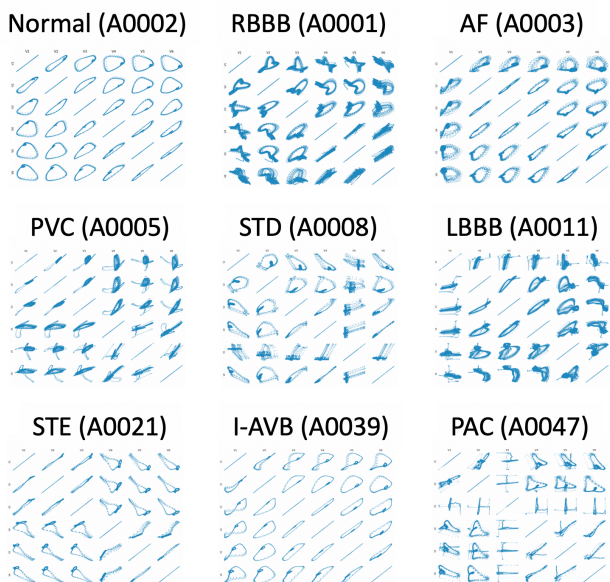


Figure 1. Correlation plots between precordial leads.

We next show the correlograms between the leads for each of the patients, centered on zero time lag (see Figure 2). Again, along the diagonals, the peak is centered on zero time lag, and the peak tends to move away from zero time lag as we move away from the diagonals. Off-diagonal signal lags of 2-16 ms were physiologically consistent with known right atrial nerve conduction speeds [3] of 13-20 m/s and lead separations of 2.6-32 cm.

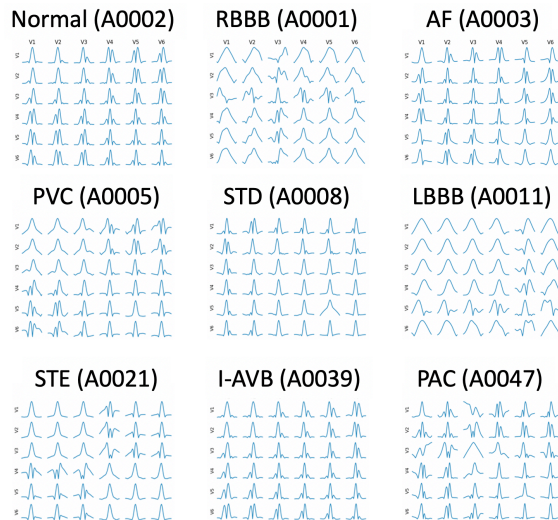


Figure 2. Correlograms between precordial leads.

3.2. Lead geometry reconstruction

We show the reconstructed lead geometries for the same set of patients (see Figure 3). Of note, the normal and AF patients displayed the most linear reconstructions, while other conditions corresponded to reconstructions with V1 and V6 unphysically close to each other and/or crossings of edges between physically adjacent leads.

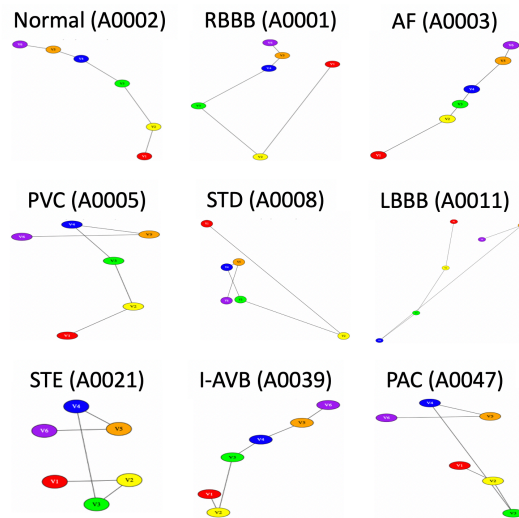


Figure 3. Reconstructed lead geometries, using AHA color codes to indicate lead identity.

3.3. Inferring conditions from geometries

We inferred individual conditions from those geometries using support vector classification (see Table 1). Left and right bundle branch block conditions were most accurately distinguishable from the normal geometries, while first-degree AV block and premature atrial complex were least accurately distinguishable from normal geometries.

Table 1. Prediction accuracies from geometries.

Condition	Training/validation set size	Prediction accuracy
LBBB	130	76.2%
RBBB	1068	73.6%
STD	500	63.0%
PVC	398	61.6%
AF	700	58.6%
STE	130	58.5%
I-AVB	432	55.1%
PAC	358	46.4%

4. Discussion

We have demonstrated that multi-lead ECG recordings from the PhysioNet 2020 Challenge Training Dataset contain sufficient information to partially or fully reconstruct the relative positions of precordial leads without significant prior information. Moreover, we have shown that these reconstructed geometries encode nontribal diagnostic information about underlying cardiac abnormalities.

This ability to distill high-dimensional ECG recordings into simple geometric arrangements opens the door to multiple applications of potential clinical value. First, such geometries might highlight any misplacement of electrodes as deviations from a linear reconstruction, and any mistaken swapping of electrodes as corresponding node swaps. Second, our method offers a novel means for visually summarizing an entire ECG recording at a glance with a simple diagram for a physician. Third, our method could potentially improve automated classification of patient conditions with less training data by reducing ECG time series from thousands of dimensions down to 36 dimensions with demonstrated clinical salience.

Acknowledgments

We thank J. Stephenson and O. Chaudoir for helpful discussions.

References

- [1] E. Sejdic and T. H. Falk, eds. *Signal Processing and Machine Learning for Biomedical Big Data*. CRC Press, 2018.
- [2] S. C. North, “NEATO user’s guide,” *Murray Hill, NJ: AT&T Bell Laboratories* (1992).
- [3] A. S. Paintal, “The conduction velocities of respiratory and cardiovascular afferent fibres in the vagus nerve,” *J. Physiol.*, vol. 121, pp. 341-359 (1953).

Address for correspondence:

Alexander D. Wissner-Gross, Ph.D.
PO Box 390452, Cambridge, MA 02139 USA
alexwg@post.harvard.edu