Synchronization of Conventional Electrocardiogram Recordings for Accurate Vectorcardiography Reconstruction

Elisa Ramírez[1,2], Samuel Riupérez-Campillo[1,3,4], Francisco Castells[1], Rubén Casado-Arroyo[5], José Millet[1,6]

[1]ITACA Institute, Universitat Politècnica de València, Valencia, Spain
[2]Universidad Carlos III de Madrid, Madrid, Spain
[3] D-ITET, Swiss Federal Institute of Technology (ETHz), Zürich, Switzerland
[4] School of Medicine, Stanford University, Palo Alto, CA, USA Cardiac
[5]Electrophysiology Lab, Hôpital Erasme, Brussels, Belgium
[6]Centro de Investigación Biomédica en Red Enfermedades Cardiovascular, Madrid, Spain.

Abstract

The vectorcardiogram (VCG) provides a comprehensive representation of the heart's electrical activity in 3D aiding in the diagnosis and treatment of cardiovascular diseases. The conventional electrocardiogram (ECG) records twelve leads intermittently at intervals of 2.5 seconds, with lead II typically recorded continuously, which poses a challenge for reconstructing the VCG, as each lead's beats belong to different time instances. The purpose of this research is to propose and validate a methodology for accurately synchronizing the recording beats to reconstruct the VCG. To achieve this goal, a phantom was created to mimic the standard 12-lead ECG setup. The temporal offset of each beat from the first is calculated using cross-correlation utilizing the continuous lead and the same offset is applied to all leads, and finally reconstructing the VCG. The results demonstrate precise synchronization, as evidenced by Pearson correlation values of 0.9959 ± 0.0034, an MAE of 0.0077 ± 0.0024 mV, and an RMSE of 0.0119 ± 0.0038 mV in the VCG reconstruction. This technique is essential for the accurate diagnosis and treatment of cardiovascular diseases and can be applied to conventional ECG recordings taken on paper to obtain VCG.

1. Introduction

Vectorcardiography (VCG) is a powerful tool used to visualize the heart's electrical activity, providing information on the orientation and magnitude of the cardiac vector in 3D [1]. This aids in the diagnosis and treatment of cardiovascular diseases with high sensitivity, including myocardial infarction, ischemia, and hypertrophy [2]. However, the use of VCG in clinical practice is limited due to the need for additional electrodes on the patient's body. Alternative methods have been developed to derive VCG from standard 12-lead ECGs [2-4].

Nowadays it is common to have the standard 10-second ECG digitized, which allows for different types of analysis, including reconstruction of the VCG. But this has not always been the case, and for a long time only the ECG printed on paper, or more recently in *pdf* format, was available. Printed ECG recordings are commonly obtained by acquiring three leads at a time, at intervals of 2.5 seconds, along with a continuous 10-second lead, usually lead II. Nonetheless, this conventional approach poses a challenge when constructing the VCG since the beats of each lead occur at different times. Thus, achieving accurate VCG construction necessitates the synchronization of beats. The process from the ECG printout to the representation of the VCG consists of two steps: digitization and synchronization. Although there are many studies addressing the digitization of printed recordings [5-7], only the work from Morales *et al.* [8] focuses on the synchronization of leads for the purpose of VCG reconstruction.

We propose in this study an ECG synchronization tool that enables the accurate reconstruction of the vectorcardiogram (VCG). Proper reconstruction of the VCG requires synchronization of beats that accounts for the lags between peaks in the leads, resulting from the propagation of the heart's electrical activity. Therefore, addressing this issue at the lead level may not produce the best outcomes, as the actual temporal relationship between leads may not be preserved in certain beats. To the authors' knowledge, this proposed method represents the first effort to resolve the synchronization problem at the beat level while maintaining the inter-lead relationship and lags.

2. Materials

2.1. Database

We used the PTB-XL ECG dataset [9] available in Physionet [10], because it contains a large number of signals, 21799 out of 18869 patients. In addition, the signals have all 12 leads in 10-second format and are available with a sampling rate of 100 Hz and 500 Hz. Cardiologists have manually annotated the data, labeling it according to the following categories: normal ECG, myocardial infraction, ST/T changes, conduction disturbance and hypertrophy.

All signals presenting any irregularity caused by noise or an abnormal heartbeat were excluded from the study. In addition, phantoms that did not contain a complete beat in any lead were also discarded as it would not be possible to reconstruct the VCG correctly. Finally, 11787 records were evaluated as meeting the above conditions.

2.2. Phantom

A phantom was created to simulate traditional paper ECG recordings. It recreates the lead sequence and timing of the original recordings using complete ECG signals (see Figure 1). However, the digital signals were used directly, evaluating in this study only the synchronization and not the digitization process.

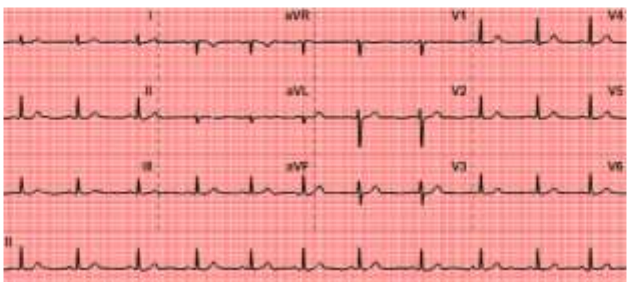

Figure 1. Phantom representation of a traditional layout 12-lead ECG recording.

3. Methods

3.1. Ground Truth Generation

The proposed algorithm uses beats from different time instants corresponding to each lead, assuming invariance over the 10-second recording period. For determining the ground truth for VCG reconstruction, this assumption must be considered.

By averaging beats the variability can be captured but the individual beat morphological details worsen. Thus, the ground truth must choose the average of the minimum number of beats that capture enough variability. For determining that, the Pearson correlation between the mean of all beats and the mean of *N* beats was calculated for all leads and it is shown in Figure 2.

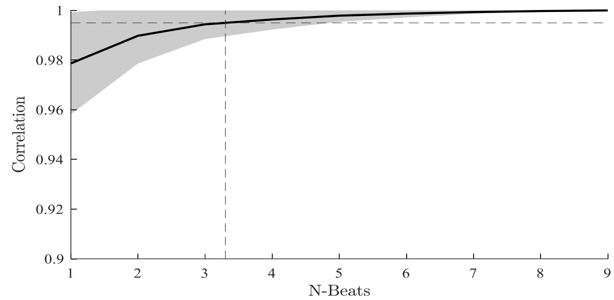

Figure 2. Correlation between the mean of *N* beats and all beats.

A non-linear increase in correlation it is observed as the number of beats used in the average increases, and by establishing a threshold of 99.5% of variability, three beats are enough for computing the ground truth.

Additionally, performance evaluation was conducted using two other ground truth scenarios, one consisting of a single beat and the other of the average of all beats. The results obtained from all three ground truth conditions were highly similar.

3.2. Synchronization and Reconstruction Algorithm

The algorithm for reconstructing the VCG from the phantom consists of an ECG synchronization and a VCG reconstruction method. Overall, the methodology can be divided into four steps as shown in Figure 3.

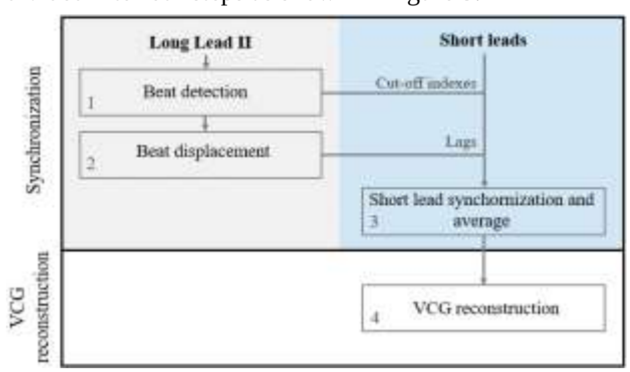

Figure 3. Synchronization and reconstruction algorithm proposed in this study.

3.2.1. Beat Detection

The first step of the proposed pipeline consists of the beat detection. For this purpose, the following steps have been designed: (i) R-peak detection with the R-DECO software [11], (ii) calculation of the start and end of each beat. In ECG, the ratio between the PR and QT intervals is known to be approximately 1/3 under normal conditions. Therefore, to accurately determine the start and end of each beat, a method was based on dividing each RR interval into 2/3 for the preceding beat and 1/3 for the following beat (see Figure 4.A).

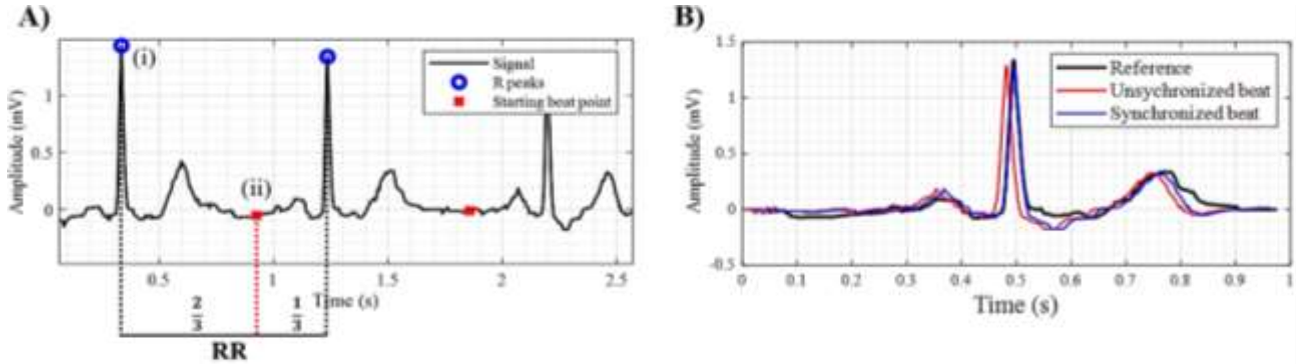

Figure 4. A) Beat selection procedure based on: (i) detection of R peaks and (ii) establishment of cut-off points. B) Representation of an unsynchronized and synchronized beat with the reference beat by cross-correlation, thus calculating the lag applied to that beat.

3.2.2. Beat Displacement

The subsequent step in the proposed method involves determining the displacement of each beat in lead II relative to the first beat. To accomplish this, cross-correlation was utilized, selecting the point of maximum similarity as the displacement (see Figure 4.B).

3.2.3. Short lead synchronization and average

Following the determination of the cut-off indices and displacements of each beat, it became possible to synchronize all phantom beats by applying them to the corresponding leads. Subsequently, the heartbeats from each lead were averaged to produce a single heartbeat per lead.

3.2.4. VCG reconstruction

The last step of the algorithm corresponds to the VCG reconstruction. The inverse Dower transform [12] is a commonly used technique in VCG reconstruction from ECG signals. It involves a linear transformation that reduces the information from the 12 leads to three space coordinates (see Figure 5).

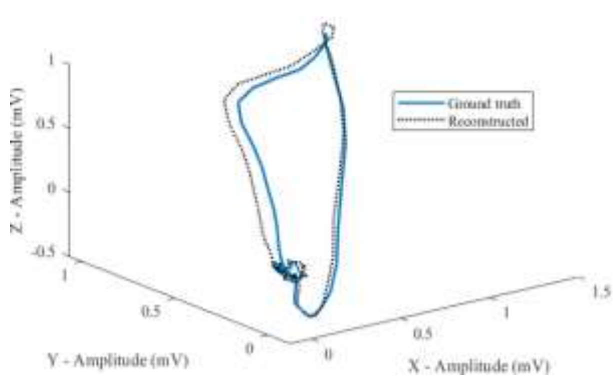

Figure 5. Three-dimensional representation of the VCG for the ground truth and the phantom reconstruction.

4. Results and Discussions

4.1. ECG level assessment

In order to assess the correct synchronization of the beat-to-beat leads, the similarity between the phantom ECG leads and the ground truths was evaluated. For each lead, three metrics have been extracted. The correlation indicates the similarity in the shape and the Mean Average Error (MAE) and Root Mean Squared Error (RMSE) differences in terms of magnitude.

Table 1. Metrics for evaluating the synchronization at ECG level ($\times 10^{-2}$).

Lead	Correlation	MAE	RMSE
I	99.00±1.44	1.41±0.71	1.05±0.54
II	99.35±0.96	1.16±0.61	0.87±0.47
III	97.97±3.48	1.29±0.67	0.90±0.48
aVR	98.99±1.45	1.45±0.59	1.11±0.47
aVF	96.86±4.75	1.60±0.64	1.15±0.47
aVL	98.06±3.09	1.35±0.55	0.97±0.41
V1	99.21±1.30	1.54±0.65	1.08±0.45
V2	99.53±0.66	1.96±0.87	1.29±0.52
V3	99.50±0.65	1.99±0.88	1.29±0.51
V4	99.36±0.81	2.31±1.03	1.43±0.54
V5	99.36±0.82	2.10±0.90	1.35±0.51
V6	99.14±2.47	1.84±0.86	1.27±0.56

The extracted metrics presented in Table 1, show a high similarity between ground truth, and beat synchronization in all leads.

In [8] the synchronization is tackled at lead level and to report the results they extract the NRMSE. The overall results they expose is 12.66%. The results of the method we propose provides an average of 0.67% NRMSE, being the maximum and minimum value 0.92% and 0.56% for lead V4 and II respectively. Although this comparison is not direct as it is not the same database, the proposed method results reduce the error by a factor of more than 10. This characteristic renders this method the most superior among the current state-of-the-art synchronization methods.

4.2. VCG level assessment

In the inverse Dower transform reconstruction matrix, each of the leads is not weighted equally, so it is necessary to perform a second evaluation based on the similarity of the phantom VCG reconstruction and the references.

Specifically, this transformation eliminates the contribution of leads III and augmented leads, which were the worst performers in the previous section. In addition, the inverse Dower transform gives a high weighting to lead II, especially in the Y coordinate, so the results provided in this coordinate will be determined by how accurate the synchronization of lead II is.

The metrics have been extracted for the three spatial coordinates (X, Y and Z), the average (Avg) metric of those coordinates, and also the metric in the 3D matrix.

Table 2. Metrics for evaluating the synchronization at VCG level ($\times 10^{-2}$).

	Correlation	MAE	RMSE
X	99.66±0.45	0.82±0.31	1.30±0.56
Y	98.89±1.82	0.74±0.38	0.98±0.48
Z	99.58±0.57	8.09±3.43	1.13±0.47
Avg	99.38±0.70	3.22±1.16	1.14±0.35
3D	99.59±0.34	0.77±0.24	1.19±0.38

The results in Table 2, reflect the correct synchronization of the beats when performing VCG reconstruction. The slight differences with respect to the ground truth are being due to random noise in the signal but are not clinically significant since it will not affect the interpretation of the VCG waveform or the diagnosis of cardiac conditions.

5. Conclusions

The proposed algorithm has demonstrated promising outcomes in the reconstruction of VCGs obtained from non-simultaneous ECG recordings. The results indicate that synchronizing and reconstructing the VCG at the beat level yields superior outcomes compared to the synchronization method at the lead level. By doing so, the physiological relationship between the delays occurring in the waves of different leads is retained. This approach holds significant clinical relevance and may have important implications for the evaluation of VCG in centers where paper records remain in use, as well as retrospective studies of cardiac pathology where a significant portion of data is in this format.

Acknowledgments

This work was supported by PID2019-109547RB-I00 (National Research Program, Ministerio de Ciencia e Innovación, Spanish Government) and CIBERCV CB16/11/00486 (Instituto de Salud Carlos III).

References

[1] H.C. Burger. Heart and Vector: Physical Basis of Electro-cardiography. Gordon and Breach; 1968.

[2] R. Jaros, R. Martinek, and L. Danys. Comparison of Different Electrocardiography with Vectorcardiography Transformations. Sensors (Basel), vol. 19, no. 14, p. 3072, Jul. 2019.

[3] L. Edenbrandt, A. Houston, and P. W. Macfarlane. Vectorcardiograms synthesized from 12-lead ECGs: A new method applied in 1792 healthy children. Pediatr. Cardiol., vol. 15, pp. 21–26, 1994.

[4] M. Lingman et al. Transient repolarization alterations dominate the initial phase of an acute anterior infarction—A vectorcardiography study," J. Electrocardiol., vol. 47, pp. 478–485, 2014.

[5] J. Millet-Roig et al. A method capable of digitising every type of biosignal paper recordings as a tool for database improvement. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286), Hong Kong, China, 1998, pp. 1238-1241 vol.3.

[6] P. Gautam, R. K. Sunkaria, and L. D. Sharma. Digitisation of paper-ECG using column wise-median approach. Int. J. Comput. Appl. Technol., vol. 66, no. 2, 2022.

[7] M. Baydoun et al. High Precision Digitization of Paper-Based ECG Records: A Step Toward Machine Learning. IEEE J Transl Eng Health Med, vol. 7, 1900808, Nov. 2019.

[8] Morales, E., Sevilla, D., Pierluissi, J.H., Nazeran, H. (2005). Digitization and Synchronization Method for Electrocardiogram Printouts. Conference Proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society Conference, 2, 1588-1591.

[9] Wagner, P., Strodthoff, N., Bousseljot, R., Samek, W., & Schaeffter, T. (2022). PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3). PhysioNet, Jan, 2022.

[10] Goldberger, A., Amaral, L., et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online], vol. 101, no. 23, pp. e215-e220, Jun, 2000.

[11] Moeyersons, J., Amoni, M., et al. R-DECO: An open-source Matlab based graphical user interface for the detection and correction of R-peaks (version 1.0.0). PhysioNet.

[12] Lars Edenbrandt, Olle Pahlm. Vectorcardiogram synthesized from a 12-lead ECG: Superiority of the inverse Dower matrix. Journal of Electrocardiology, vol. 21, no. 4, pp. 361-367, Apr. 1988. ISSN 0022-0736.

Address for correspondence:
Elisa Ramírez
ITACA Institute Edificio 8G, Acceso B, 1a planta, Universitat Politècnica de València, Camino de Vera S/N 46022 Valencia España.
eliramir@pa.uc3m.es