Novel Rank-based Features of atrial potentials for the Classification Between Paroxysmal and Persistent Atrial Fibrillation

Hanie Moghaddasi[1], Richard C. Hendriks[1], Alle-Jan van der Veen[1], Natasja M.S. de Groot[1,2], Borbála Hunyadi[1]
[1] Circuits and Systems, Delft University of Technology, Delft, The Netherlands
[2] Department of Cardiology, Erasmus University Medical Center, Rotterdam, The Netherlands

Abstract

Atrial fibrillation (AF) is the most common arrhythmia. Although the exact cause is unclear, electropathology of atrial tissue is one contributing factor. Electropathological characteristics derived from intra-operative epicardial measurements, such as conduction block (CB), can be used to assess the severity of AF. In sinus rhythm, however, this parameter does not indicate significant difference between different development stages of AF, such as paroxysmal and persistent AF. Therefore, we propose a methodology to improve AF severity detection using intra-operative electrograms. We propose a model that describes the spatial diversity of atrial potential waveforms during a single beat on the multi-channel electrograms. Based on this model, we derive two novel features. During sinus rhythm, we used 334 beats from patients with a history of PAF or PsAF. Using a random forest classifier, we achieved 80.38% classification accuracy, while classification based on the CB leads to an accuracy of 57.22%.

1. Introduction

Atrial fibrillation (AF) is the most sustained arrhythmia. Typically, AF is classified into four categories based on the duration of an episode: paroxysmal, persistent, long-standing persistent, and permanent AF [1]. In this paper, we focus, in particular, on the classification between paroxysmal AF (PAF) and persistent AF (PsAF) to investigate the development stages of AF. Moreover, AF is a progressive disease and the success rate of the most common treatment, ablation therapy, highly depends on the severity of AF [2, 3]. Understanding how rank-related features of atrial potentials can differ between different stages of AF is important.

Previously, we developed a method which classifies PAF and PsAF using a multi-channel electrocardiogram (ECG) [4]. In this work, we go one step beyond from the surface of the body (i.e., from ECGs) to the surface of the heart (i.e., to electrograms) to gain more insights into the electropathology of AF. Typically, classification based on intra-operative epicardial measurements is done using features like conduction block (CB). However, this parameter does not clearly differentiate PAF and PsAF in sinus rhythm (SR) [5,6]. Van der Does et al. showed that among all locations, only at Bachmann's bundle (BB) there is a higher number of continues conduction delay and block (cCDCB) line (p-value=0.040), though [5]. However, the CB and cCDCB rely on an estimate of the local activation time (LAT). On the other hand, from the potential morphological viewpoint, Ye et al. demonstrated that the R/S ratio of single potentials (SPs) might be a useful characteristic for investigating the development stages of AF [7]. Therefore, we propose in this paper a complimentary feature, related to the morphology of the signal, based on the underlying idea that AF progression can be related to variations in the atrial potential waveforms (APW).

2. Materials and Method

2.1. Data and pre-processing

In this study, patients with normal sinus rhythm (NSR) who had a history of PAF or PsAF were included. We used multi-site high-density electrograms (EGMs) measured during open-heart surgery, at a sample rate of 1 kHz. One ECG is measured simultaneously with the EGMs (188 electrodes) to detect the R peak location for heartbeat segmentation. Using a wavelet-based ECG delineator technique [8], we segmented the atrial activity of each EGM to concentrate on the atrial activity (AA). We used a fixed window with a length of 180 ms, between 240 ms and 60 ms before the R peak. We used 3-6 beats per patient from 9 PAF patients and 14 PsAF patients from three atrial regions. Altogether we included 110 beats from PAF patients and 224 beats from PsAF patients. More details on the mapping scheme and atrial regions are explained in [6].

2.2. Signal model

The heart's electrical activity is initiated by the sinoatrial (SA) node and propagates across the atrium. We assume

that in NSR, all cells generate the same action potential. In that case, each electrode measures the attenuated-delayed version of the same reference AA. For the mth electrode this can be modelled as

$$y_m(t) = a_m s(t) * \delta(t - \tau_m) \tag{1}$$

where $s(t)$ is the reference AA, a_m models the positive real attenuation and $\delta(t - \tau_m)$ is the delay of the AA at electrode m compared to the reference AA, and $*$ is the convolution operator. In the frequency domain, Eq. (1) can be written as

$$\tilde{y}_m(\omega) = \int_{-\infty}^{+\infty} y_m(t) e^{-j\omega t} dt = a_m \tilde{s}(\omega) e^{-j\omega\tau_m} \tag{2}$$

where $\tilde{(\cdot)}$ denotes the frequency domain. The matrix $\tilde{\boldsymbol{Y}} \in \mathbb{C}^{M\times\Omega}$ is constructed by stacking all the AAs for all electrodes and frequencies, where $m \in \{1, 2, \cdots, M\}$ is the number of electrodes and $\omega \in \{1, 2, \cdots, \Omega\}$ denotes the angular frequency index. Let us denote the absolute value of $\tilde{y}_m$ by

$$|\tilde{y}_m| = a_m |\tilde{s}| \tag{3}$$

then we define $\tilde{\boldsymbol{B}} \in \mathbb{R}^{M\times\Omega}_{\rho}$ as a matrix with the element-wise absolute values of $\tilde{\boldsymbol{y}}_{\boldsymbol{m}} = [\tilde{y}_m(1), \tilde{y}_m(2), \cdots, \tilde{y}_m(\Omega)]$, where ρ is the rank of matrix $\tilde{\boldsymbol{B}}$. In the case that the AA shows little variation, i.e., the action potentials are the same up to a delay and each electrode only measures the AA from the atrial site beneath it, $\tilde{\boldsymbol{B}}$ is a rank-1 matrix. In case there is more variation on the atrial APWs, the AA will be different on some electrodes and the rank of $\tilde{\boldsymbol{B}}$ will be higher. Thus, when there is variation across the atrial regions resulting in $\rho \leq M$ different APWs, $\tilde{\boldsymbol{B}}$ will be a rank-ρ matrix. In this work, we assume that there are a limited number of different APWs. Therefore, $\tilde{\boldsymbol{B}}$ is a low-rank matrix and we can decompose it into a sum of rank-1 matrices. Using the singular value decomposition (SVD), matrix $\tilde{\boldsymbol{B}}$ is factorized as

$$\tilde{\boldsymbol{B}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^* = \boldsymbol{U}\left[\begin{array}{ccc|ccc} \sigma_1 & 0 & \cdots & & \cdots & 0 \\ 0 & \sigma_2 & & & & \\ \vdots & & \ddots & & & \\ & & & \sigma_\rho & & \\ \hline & & & 0 & & \\ \vdots & & & & \ddots & \\ 0 & & & & & 0 \end{array}\right]\boldsymbol{V}^* \tag{4}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices, $\boldsymbol{\Sigma}$ is a non-negative diagonal matrix that contains the singular values of matrix $\tilde{\boldsymbol{B}}$ in a descending order,$(^*)$ denotes the Hermition operator. The rank is determined by the number of non-zero singular values of $\tilde{\boldsymbol{B}}$. However, in practice, the measurements are rather noisy and we need to divide the data matrix into signal subspace and noise subspace and determine the rank of the signal subspace. Given ρ, we can truncate the SVD to its first ρ terms and estimate the least square optimal $\tilde{\boldsymbol{B}}$ matrix that is

$$\hat{\boldsymbol{B}} = \boldsymbol{U}_{\boldsymbol{\mathcal{R}}}\boldsymbol{\Sigma}_{\boldsymbol{\rho\times\rho}}\boldsymbol{V}^*_{\boldsymbol{\mathcal{R}}}. \tag{5}$$

where matrices $\boldsymbol{U}_{\boldsymbol{\mathcal{R}}}$ and $\boldsymbol{V}_{\boldsymbol{\mathcal{R}}}$ are the range spaces of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, and $\boldsymbol{\Sigma}_{\boldsymbol{\rho\times\rho}}$ is the truncated version of $\boldsymbol{\Sigma}$ to its first ρ singular values . The rank of $\tilde{\boldsymbol{B}}$ thus directly says something about the variation in AA and thus about the severity of AF. However, rank estimation with noisy and limited data records is rather challenging. Therefore, we propose two rank-related features that can be used to discriminate between PAF and PsAF.

2.3. Proposed features

We can measure the amount of variance captured by each singular value by computing the norm of the rank-1 reconstruction divided by the norm of the whole data matrix. This ratio shows the importance of each singular value in the reconstruction. Thus, we propose feature I_1 based on the overall variance captured by the first ρ singular values such that these first ρ singular values capture about 80% data variance. We choose the variance threshold of 80%, as this value achieved the highest classification accuracy on our dataset. Feature I_1 is then given by

$$I_1 = arg\min_{\rho_1} \left| \frac{\sum_{i=1}^{\rho_1} ||\boldsymbol{u}_i s_i \boldsymbol{v}_i^*||^2}{\sum_{j=1}^{M} ||\boldsymbol{u}_j s_j \boldsymbol{v}_j^*||^2} - 0.8 \right|$$

$$= \text{argmin}_{\rho_1} \left| \frac{\sum_{i=1}^{\rho_1} s_i^2}{\sum_{j=1}^{M} s_j^2} - 0.8 \right| \tag{6}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the ith left and right singular vectors of $\tilde{\boldsymbol{B}}$. The higher I_1, the more diverse the atrial potential morphologies in the $\tilde{\boldsymbol{B}}$ matrix.

With the second feature, we focus on the relative importance of the singular values in the data matrix reconstruction. We define the ratio between the consecutive singular values σ_i and σ_{i+1} as

$$\varrho_i = \frac{\sigma_i}{\sigma_{i+1}}. \tag{7}$$

The rate of change as a function of i in the ϱ ratio is related to the relative importance of the singular values. It means that when the rate of change is low, increasing the rank for the reconstruction will not change the variance of the data considerably. Therefore, the singular value can be considered to belong to the noise subspace. In Fig 1,

the ϱ ratio is shown. The ϱ ratio for PAF patients shows a monotonic descending pattern but not in PsAF for the first components. Instead, there is a ratio increase on the third ratio (σ_3 to σ_4) demonstrating that the relative importance of σ_3 is much higher than that of σ_4. Moreover, looking at the ϱ ratios from the third ratio onward, PAF patients and PsAF patients show the same pattern with different rates of change. Taking these two points into account, we propose to take the first three singular value ratios (ϱ_1, ϱ_2 and ϱ_3) as features to extract the dominant components. Furthermore, to quantify the rate of change of ϱ_i as a function of i, we propose a feature that captures the bending point around the plateau as

$$I_2 = arg\min_i \quad |(|\log(\varrho_{i+1}) - \log(\varrho_i)| - \epsilon)| \\ \text{s.t.} \quad 1 \leq \varrho_{i+1} \leq 1.5. \tag{8}$$

where ϵ was set to 0.1. Similarly to the variance threshold parameter, this parameter was optimized based on the whole dataset in order to achieve maximal accuracy. I_2 approximates the second derivative between the consecutive singular values in the log domain. Indeed, a higher I_2 demonstrates the lower rate of change in the ϱ ratio which shows a higher number of important subspaces, i.e higher number of different APWs in the signal subspace.

3. Results

This section presents a comparison of the performance of the proposed features with the reference features (CB). From a total of 334 heartbeats, we used the five proposed features, namely $I_1, I_2, \varrho_1, \varrho_2, \varrho_3$. Fig 2.A shows the box plots of the proposed I_1 and I_2 features. Comparing I_1 and I_2 between PAF and PsAF, I_1 and I_2 are larger for PsAF than for PAF. It shows that a higher variation of APWs across the atrial regions is present in PsAF than in

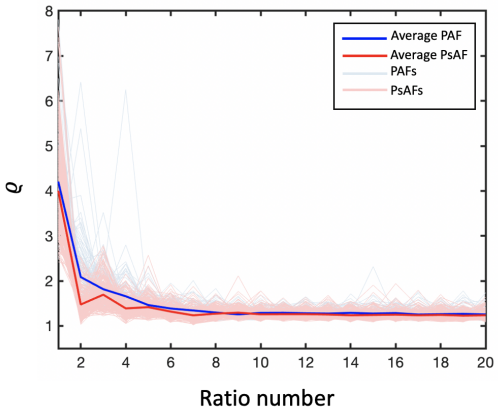

Figure 1. The average of the first 20 singular value ratios across all patients and all beats in PAF and PsAF

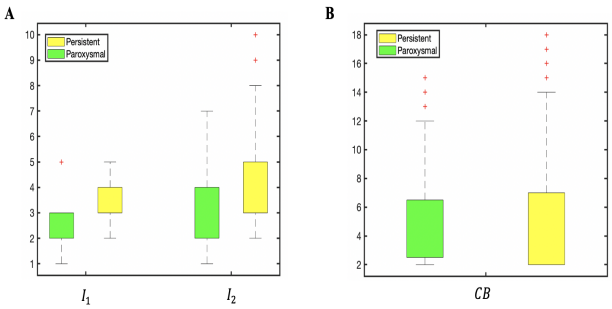

Figure 2. Box plots of the features, A) proposed features, B) state-of-the-art features.

PAF. We compared the classification performance of the proposed features with the electropathological characteristics based on the CB. These parameters are calculated based on the LAT. The LAT is defined as the steepest deflection of the electrogram. The CB is defined as ΔLAT of adjacent electrodes $\geq$ 12ms. Comparing CB between PAF and PsAF in Fig 2.B , there is no considerable difference between these two groups. Moreover, we use two classifiers to evaluate the performance of the proposed features. First, a support vector machine (SVM) with radial basis function (RBF) kernel function and kernel parameter $\sigma = 1.5$ has been employed. The parameter has been chosen empirically based on our dataset. Second, a random forest (RF), an ensemble method, has been employed with a bagged ensemble of 30 classification trees. For validation, we used a 5-fold cross-validation approach on 334 heartbeats. Table 1(green rows) shows the performance of the classifiers on the proposed features. To investigate the importance of each feature, we trained the classifiers using features separately. Using all the proposed features (i.e., $I_1, I_2, \varrho_1, \varrho_2, \varrho_3$), we achieved 80.38% accuracy with the RF classifier, while CB could reach the accuracy of 57.22% with an SVM classifier. The proposed features improve the classification accuracy between PAF and PsAF. Furthermore, using all the proposed features, RF slightly

Table 1. The performance of the SVM and RF classifiers on the features

Features \ Classifiers	RF	SVM
CB	53.97	57.22
$\boldsymbol{I_1}$	76.91	76.11
$\boldsymbol{I_2}$	73.16	74.03
$\boldsymbol{I_2, \varrho_1, \varrho_2, \varrho_3}$	77.56	76.88
$\boldsymbol{I_1, I_2, \varrho_1, \varrho_2, \varrho_3}$	80.38	78.97

outperforms SVM in PAF and PsAF classification.

4. Discussion

This paper proposes rank-related features to discriminate between PAF and PsAF. The rank estimation problem has been addressed in many studies, Minimum description length (MDL) [9] and Aikaike's information criterion (AIC) [10] being two well-known methods in this context. These techniques, however, assume that the variance of the sensor self-noise for all sensors are the same, which is not the case in our database. Therefore, we compared our results with two relevant methods developed by van der Veen et al. and Koutrouvelis et al. in [11] and [12], respectively. Koutrouvelis et al. found the scree point based on an approximation of the second derivative between the consecutive eigenvalues. Van der Veen et al. proposed a method to determine a rank threshold for counting the number of sources using the covariance matrix. Our preliminary comparison show that the I_2 developed in this paper slightly outperforms the methods introduced in [11] and[12]. Note that a few parameters that were used to calculate our features, such as the variance threshold in Eq. (6), were optimized on the full available dataset. Future work should investigate whether our parameter settings generalize well to an independent dataset and hence can maintain classification accuracy.

5. Conclusion

This study proposed a method for the classification between patients with PAF or PsAF using intra-operative EGMs during the sinus rhythm. The features reflect the differences in the degree of electropathology between the PAF and PsAF patients. We quantified the degree of electropathology using features related to the rank of the matrix containing the absolute frequency spectrum of the EGMs measured at each electrode. Feeding an RF classifier with the proposed features, we achieved 80.38% accuracy for the classification between PAF and PsAF. Our future work will aim to further investigate the relationship between APW morphology and the different stages of atrial fibrillation.

Acknowledgments

This research was funded in part by the Medical Delta Cardiac Arrhythmia Lab (CAL), The Netherlands. The authors would like to thank Paul Knops for his assistance in data acquisition.

References

[1] January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, Ellinor PT, Ezekowitz MD, Field ME, Furie KL, et al. 2019 aha/acc/hrs focused update of the 2014 aha/acc/hrs guideline for the management of patients with atrial fibrillation: a report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. Journal of the American College of Cardiology 2019;74(1):104–132.

[2] Ortigosa N, Galbis A, Fernández C, Cano O. Gabor frames for classification of paroxysmal and persistent atrial fibrillation episodes. Medical Engineering Physics 2017;39.

[3] Moghaddasi H, Hendriks RC, van der Veen AJ, de Groot NM, Hunyadi B. Classification of de novo post-operative and persistent atrial fibrillation using multi-channel ecg recordings. Computers in Biology and Medicine 2022; 143:105270.

[4] Moghaddasi H, van der Veen AJ, de Groot NM, Hunyadi B. Tensor-based detection of paroxysmal and persistent atrial fibrillation from multi-channel ecg. In 2020 28th European Signal Processing Conference (EUSIPCO). IEEE, 2021; 1155–1159.

[5] van der Does WF, Heida A, van der Does LJ, Bogers AJ, de Groot N. Conduction disorders during sinus rhythm in relation to atrial fibrillation persistence. Journal of clinical medicine 2021;10(13):2846.

[6] Lanters EA, Yaksh A, Teuwen CP, van der Does LJ, Kik C, Knops P, van Marion DM, Brundel BJ, Bogers AJ, Allessie MA, et al. Spatial distribution of conduction disorders during sinus rhythm. International journal of cardiology 2017; 249:220–225.

[7] Ye Z, van Schie MS, de Groot NM. Signal fingerprinting as a novel diagnostic tool to identify conduction inhomogeneity. Frontiers in physiology 2021;12:652128.

[8] Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ecg delineator: evaluation on standard databases. IEEE Transactions on biomedical engineering 2004;51(4):570–581.

[9] Rissanen J. Modeling by shortest data description. Automatica 1978;14(5):465–471.

[10] Akaike H. A new look at the statistical model identification. IEEE transactions on automatic control 1974;19(6):716–723.

[11] Koutrouvelis AI, Hendriks RC, Heusdens R, Jensen J. Estimation of sensor array signal model parameters using factor analysis. In 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019; 1–5.

[12] van der Veen AJ, Romme J, Cui Y. Rank detection thresholds for hankel or toeplitz data matrices. In 2020 28th European Signal Processing Conference (EUSIPCO). IEEE, 2021; 1911–1915.

Address for correspondence:

Hanie Moghaddasi
Circuits and Systems, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD, Delft, The Netherlands.
H.moghaddasi@tudelft.nl