

# Estimation of f-wave Dominant Frequency Using a Voting Scheme

Shany Biton<sup>1</sup>, Mahmoud Suleiman<sup>2</sup>, Noam Ben Moshe<sup>1,3</sup>, Leif Sörnmo<sup>4</sup>, Joachim A Behar<sup>1</sup>

<sup>1</sup>Faculty of Biomedical Engineering, Technion-IIT, Haifa, Israel

<sup>2</sup>Department of Cardiology, Rambam Medical Center and Technion The Ruth and Bruce Rappaport Faculty of Medicine, Haifa, Israel

<sup>3</sup>Faculty of Computer Science, Technion-IIT, Haifa, Israel

<sup>4</sup>Department of Biomedical Engineering, Lund University, Lund, Sweden

## Abstract

**Introduction:** Atrial fibrillation (AF) is the most common heart arrhythmia, characterized by the presence of fibrillatory waves (f-waves) in the ECG. We introduce a voting scheme to estimate the dominant atrial frequency (DAF) of f-waves. **Methods:** We analysed a subset of Holter recordings obtained from the University of Virginia AF Database. 100 Holter recordings with manually annotated AF events, resulting in a total 363 AF events lasting more than 1 min. The f-waves were extracted using four different template subtraction (TS) algorithms and the DAF was estimated from the first 1-min window of each AF event. A random forest classifier was used. We hypothesized that better extraction of the f-wave meant better AF/non-AF classification using the DAF as the single input feature of the RF model. **Results:** Performance on the test set, expressed in terms of AF/non-AF classification, was the best when the DAF was computed using the three best-performing extraction methods. Using these three algorithms in a voting scheme, the classifier obtained AUC=0.60 and the DAFs were mostly spread around 6 Hz, 5.66 (4.83-7.47). **Conclusions:** This study has two novel contributions: (1) a method for assessing the performance of f-wave extraction algorithms, and (2) a voting scheme for improved DAF estimation.

## 1. Introduction

Atrial fibrillation (AF) is the most common arrhythmia with a prevalence that increases with age [1]. AF diagnosis using the surface ECG is mainly based on whether the RR interval pattern is irregular, P-wave are absent, and fibrillatory waves (f-waves) present, where RR interval irregularity stands out as the most prominent. However, other arrhythmias are also manifested by irregular RR interval patterns which resemble AF. Therefore, f-wave information is important to consider when distinguishing AF from

other arrhythmias. However, f-wave extraction is complex since the atrial and ventricular activities overlap in both the spectral and the temporal domain. Several methods have been proposed to address this problem. When the number of leads is limited such as in Holter systems, average beat subtraction, also known as template subtraction (TS), is a widely used technique [2]. A disadvantage with this approach is its inability to handle non-stationary situations, leading to that the extracted f-wave signal is distorted and contains QRST residuals.

The performance of f-wave extraction algorithms has been benchmarked by means of simulated ECG signals using local performance measures, including the root mean square error between the ground truth and the extracted f-wave signal [3]. Other types of local performance measures quantify the amount of QRST-related residuals and the spectral concentration [4]. Another approach to assessing performance is to employ global measures, pursued in the present study together with the hypothesis that better f-wave extraction means better AF/non-AF classification. Classification is performed using the dominant atrial frequency (DAF) as the single input feature.

## 2. Methods

### 2.1. Dataset

The University of Virginia Atrial Fibrillation Database (UVAFDB) [5, 6] consists of 3-lead ECG Holter recordings acquired during the period 2004–2010 by the University of Virginia health system. Recordings were digitized at 200 Hz. The database contains 2,891 recording of patients, with rhythm annotation automatically generated by the Philips Holter software. In the present study, we use a subset of 100 recordings from patients older than 18 years, each recording manually reviewed and annotated with respect to AF events by an expert cardiologist as described in [7]. The median and interquartile patient age were 69.0 (58.8-76.2), each was 24h long. A total of 457

non-overlapping AF events lasting 128.0 (68–718) seconds was analyzed.

## 2.2. Preprocessing

Similar to the work in [8], we computed the signal quality index bSQI [9–11] and discarded segments when below 0.8. ECG recordings were prefiltered using a zero-phase second-order infinite impulse response bandpass filter, with a passband of 0.67–100 Hz to remove baseline wander and high-frequency noise [12]. A notch filter at 60 Hz was used to remove power-line interference. Channel 1, approximating a modified lead V1 was used as it is the most informative lead for expert annotation.

## 2.3. F-wave extraction

In this work, we evaluate four TS techniques for single-lead f-wave extraction. The algorithms were originally implemented in the context of fetal ECG extraction from the abdominal fetal-maternal mixture [13]. These algorithms are similar to those used in the context of f-wave extraction as they aim to cancel out the pseudo-periodic components in phase with the (maternal) cardiac cycle, while leaving the fetal ECG (in case of fetal ECG) or the f-waves (in case of AF) in the residual signal.

The first algorithm is the basic TS adapted from [14] denoted as  $TS_B$ : A QRST template is obtained by averaging temporal QRST windows with respect to the duration of the whole cardiac cycle. The QRST template is then subtracted from the original ECG signal, resulting in a residual signal that contains the f-wave and residual noise. The other three TS variants offer different degrees of adaptability of the template: in  $TS_{CE}$  the QRST template is scaled by a gain before subtraction [15]; in  $TS_{SU}$  the scaling procedure is performed for the P-wave, the QRS-complex and the T-wave individually [16]; in  $TS_{PCA}$  (principal component analysis, PCA) the almost periodic characteristics of the ECG is used for selective separation of the QRST and the f-waves [17].

The algorithms are listed in Table 2.4. Following the preprocessing step, AF events were extracted, then f-wave extraction algorithms were run only on the first minute of each event. We used 1-min windows similar to the work in [18]. Matlab R2021a and Python 3.8 were used. An example of the extracted f-wave is shown in Fig 1.

## 2.4. Machine learning

A random forest (RF) classifier was used. We extracted similar number of non-AF and AF windows for which the DAF was computed. A train-test split of 80%–20% was performed with equal number of AF/non-AF windows in the training set. To assess the performance of the trained RF

in classifying AF and non-AF windows correctly, the following statistics were computed: the harmonic mean between the sensitivity and positive predictive value, denoted  $F_1$ , and the area under the receiver operating characteristic (AUROC). The default decision threshold was used for classification in training and testing.

Table 1. Summary of the f-wave extraction algorithms experimented in this research. They were implemented in the previous work from Behar, J. et al. [9].

Method	Reference
$TS_B$	[19]
$TS_{CE}$	[19]
$TS_{SU}$	[16]
$TS_{PCA}$	[17]

## 3. Frequency estimation

Spectral analysis of the extracted f-wave signals was applied to estimate the DAF, defined as the position of the largest peak in the interval of 4–12 Hz of the power spectrum of the extracted f-wave signal [2]. The power spectrum was computed using Welch’s method with a Hamming window and 50% segment overlap. An example of power spectrum is shown in Fig. 2. After extracting the DAF using the different algorithms, a voting scheme was applied to determine the final DAF, i.e., the median value was taken as the DAF across the best performing methods for AF classification. Since real data were analyzed, a comparison of the resulting DAF to that of a reference f-wave signal could not be done.

## 4. Results

No recordings were excluded from the dataset due to low quality. 476 AF event were extracted with length of 30 seconds defined to be the minimal length of AF episode as in [8]. Out of those, 94 (20.57%) events shorter than 1-min were excluded, as windows of 1-min length were used. Results are shown for total of 363 AF windows belonging to 72 recordings.

The DAFs were mostly spread within the interval 6–8 Hz:  $TS_B$  6.1 (5.04–8.59),  $TS_{CE}$  5.81 (4.88–7.91),  $TS_{SU}$  5.96 (4.88–8.20) and  $TS_{PCA}$  5.62 (4.83–7.56). See the distributions in Fig. 3.

The test set classification performances for the DAF estimation using  $TS_B$ ,  $TS_{SU}$  or  $TS_{CE}$  were AUROC of 0.59, 0.59 and 0.59, respectively (Table 2). The voting scheme involving the three best performing techniques obtained an  $F_1$  of 0.63 and an AUROC equal to 0.60. Using the voting scheme, the DAF was mostly spread around median,

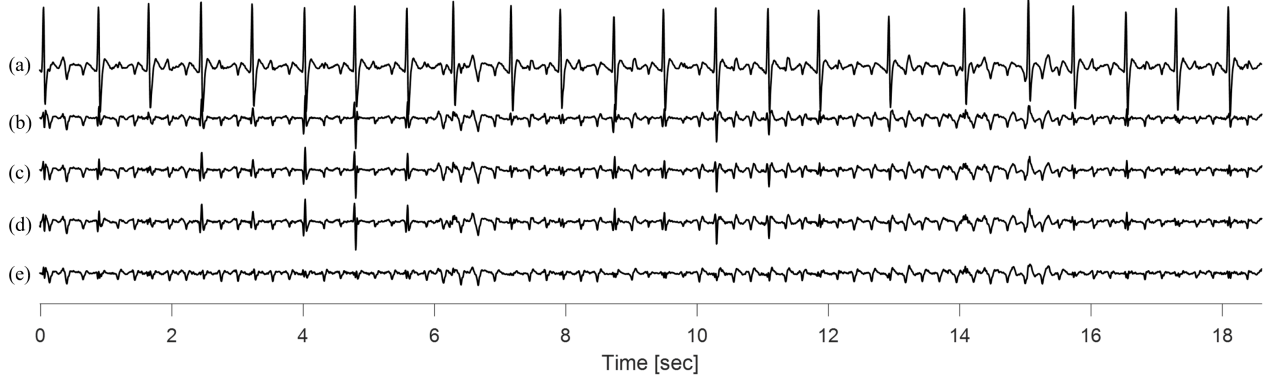


Figure 1. Example of f-wave extraction. (a) A single-lead ECG in AF and related extracted f-waves obtained using (b)  $TS_B$ , (c)  $TS_{CE}$ , (d)  $TS_{SU}$  and (e)  $TS_{PCA}$ .

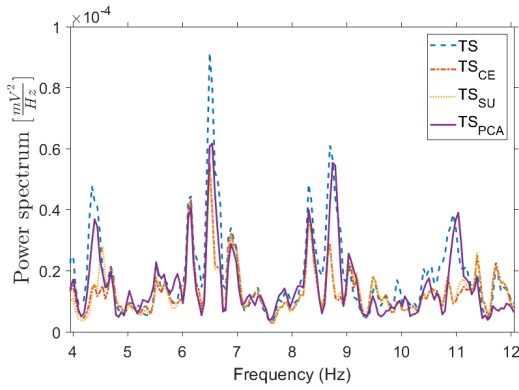


Figure 2. Example of power spectra for extracted f-wave signals obtained by the four different algorithms.

(Q1-Q3) 6 Hz (4.87–7.47).

Table 2. Performance statistics for different RF models.

Method	$F_1$	AUROC
$TS_B$	0.62	0.59
$TS_{CE}$	0.61	0.59
$TS_{SU}$	0.61	0.59
$TS_{PCA}$	0.56	0.53
Voting scheme	<b>0.63</b>	<b>0.60</b>

## 5. Discussion

The main contribution of this research is the introduction of a novel method for assessing f-wave extraction performance, whereas, in previous research, the emphasis has been on assessing local performance of f-wave extraction. An interesting approach was to use simulated ECG sig-

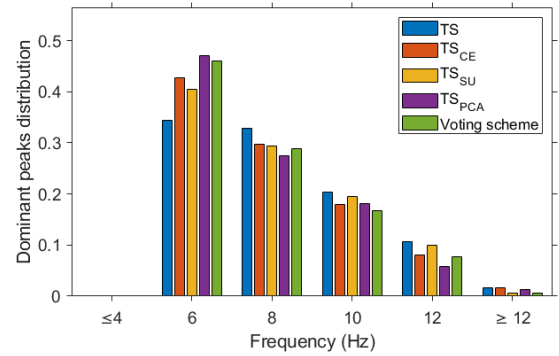


Figure 3. Distribution of the DAF obtained using four different f-wave extraction algorithms for AF examples.

nals as it provides access to the ground truth, however, simulation has the disadvantage of not fully account for the physiological variability across patients. As a consequence, we seek to develop a new method for assessing the performance of f-wave extraction algorithms based on real data experiments where the ground truth DAF is unknown.

Another important contribution of this study is the introduction of a voting scheme for DAF estimation based on the three best performing algorithms, i.e.,  $TS_B$ ,  $TS_{SU}$  and  $TS_{CE}$ .

This work has important limitations. First, the dataset is limited to 100 patients all from a single cardiac center. Future work needs to evaluate the performance of the algorithms on datasets from multiple centers, across ethnicity, age and gender. A second limitation is that we focused on the extraction of the DAF as the sole feature for assessing the performance of f-wave extraction. The usage of this single feature may not be enough to best rank the f-wave extraction algorithms using our data-driven method. Other f-wave features such as amplitude may also be taken

into account to supplement the DAF in training the RF classifier. In that respect the ranking of the algorithms could change. Finally, other extraction algorithms could be included [2], e.g., noise-dependent weighted averaging, adaptive filtering and Kalman filter based methods [13].

## Acknowledgments

This research was supported for SB, NBM and JAB by a grant (3-17550) from the Ministry of Science & Technology, Israel & Ministry of Europe and Foreign Affairs (MEAE) and the Ministry of Higher Education, Research and Innovation (MESRI) of France. SB, NBM, MS and JAB acknowledge the support of the Technion-Rambam Initiative in Artificial Intelligence in Medicine.

## References

- [1] Chugh SS, Havmöller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, Gillum RF, Kim YH, McAnulty Jr JH, Zheng ZJ, et al. Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study. *Circulation* 2014;129(8):837–847.
- [2] Sörnmo L, Petrénas A, Laguna P, Marozas V. Extraction of f waves. In Sörnmo L (ed.), *Atrial Fibrillation from an Engineering Perspective*. Springer Nature, 2018; 137–220.
- [3] Petrénas A, Marozas V, Sörnmo L, Lukosevičius A. An echo state neural network for QRST cancellation during atrial fibrillation. *IEEE Trans Biomed Eng* 2012; 59:2950–2957.
- [4] Sörnmo L. *Atrial Fibrillation from an Engineering Perspective*. Springer, 2018.
- [5] Carrara M, Carozzi L, Moss TJ, De Pasquale M, Cerutti S, Ferrario M, Lake DE, Moorman JR. Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy. *Physiol Meas* 2015; 36(9):1873.
- [6] Moss TJ, Lake DE, Moorman JR. Local dynamics of heart rate: detection and prognostic implications. *Physiol Meas* 2014;35(10):1929–1942.
- [7] Biton S, Aldhafeeri M, Marcusohn E, Tsutsui K, Szwagier T, Elias A, Oster J, Sellal JM, Suleiman M, Behar JA. Generalizable and robust deep learning algorithm for atrial fibrillation diagnosis across ethnicities, ages and sexes. *arXiv preprint arXiv:2207.09667* 2022;.
- [8] Chocron A, Oster J, Biton S, Mandel F, Elbaz M, Zeevi YY, Behar JA. Remote atrial fibrillation burden estimation using deep recurrent neural network. *IEEE Trans Biomed Eng* 2020;68(8):2447–2455.
- [9] Behar J, Oster J, Li Q, Clifford GD. ECG signal quality during arrhythmia and its application to false alarm reduction. *IEEE Trans Biomed Eng* 2013;60(6):1660–1666.
- [10] Li Q, Mark RG, Clifford GD. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiol Meas* 2007; 29(1):15–32.
- [11] Gendelman S, Biton S, Derman R, Zvuloni E, Levy J, Luggassy S, Alexandrovich A, Behar JA. Physiozoo ECG: Digital electrocardiography biomarkers to assess cardiac conduction. In *Computing in Cardiology*, volume 48. 2021; 1–4.
- [12] Kligfield P, Gettes LS, Bailey JJ, Childers R, Deal BJ, Hancock EW, Van Herpen G, Kors JA, Macfarlane P, Mirvis DM, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol* 2007;49(10):1109–1127.
- [13] Behar J, Oster J, Clifford GD. Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data. *Physiol Meas* 2014;35(8):1569–1589.
- [14] Slocum J, Sahakian A, Swiryn S. Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity. *J Electrocardiol* 1992;25(1):1–8.
- [15] Beckers F, Anne W, Verheyden B, de Kestergat CvdD, Van Herk E, Janssens L, Willems R, Heidebuchel H, Aubert A. Determination of atrial fibrillation frequency using QRST-cancellation with QRS-scaling in standard electrocardiogram leads. In *Computers in Cardiology*. 2005; 339–342.
- [16] Martens SM, Rabotti C, Mischi M, Sluijter RJ. A robust fetal ECG detection method for abdominal recordings. *Physiol Meas* 2007;28(4):373–388.
- [17] Kanjilal PP, Palit S, Saha G. Fetal ECG extraction from single-channel maternal ECG using singular value decomposition. *IEEE Trans Biomed Eng* 1997;44(1):51–59.
- [18] Alcaraz R, Rieta JJ. Non-invasive organization variation assessment in the onset and termination of paroxysmal atrial fibrillation. *Comput Meth Programs Biomed* 2009; 93(2):148–154.
- [19] Cerutti S, Baselli G, Civardi S, Ferrazzi E, Marconi AM, Pagani M, Pardi G. Variability analysis of fetal heart rate signals as obtained from abdominal electrocardiographic recordings. *J Perinat Med* 1986;14(6):445–452.

Address for correspondence:

Dr. Joachim Behar  
 jbehar@technion.ac.il  
 Faculty of Biomedical Engineering  
 Technion - Israel Institute of Technology  
 Haifa 3200003, Israel