

TIME-FREQUENCY ANALYSIS COMPENSATING MISSING DATA FOR ATRIAL FIBRILLATION ECG ASSESSMENT

Sandun Kodituwakku, Rodney A. Kennedy, Thushara D. Abhayapala

School of Engineering, The Australian National University, Canberra, Australia.

ABSTRACT

We propose a novel algorithm for temporal tracking of the fibrillatory frequency for Atrial Fibrillation (AF) ECG. Both Atrial activity extraction and fibrillatory frequency tracking are combined into a single algorithm and analytical expressions are derived for final time-frequency distributions, thus the proposed method is computationally efficient. It can also serve the general purpose of time-frequency analysis of a non-stationary signal with missing or corrupted data segments. The method uses short-time expansion of an orthogonal basis set, and regularized least squares solution used to compute a stable coefficients for the basis. A simplified and computationally efficient time-frequency distribution is derived based only on the coefficient vector. The algorithm is successfully applied to simulated as well as real ECG with AF, and its advantages over conventional average beat subtraction method are discussed.

Index Terms— Time-frequency analysis, Missing data, Orthogonal Basis, Electrocardiogram, Atrial Fibrillation

1. INTRODUCTION

Atrial Fibrillation (AF) is the most common sustained cardiac arrhythmia, increasing in prevalence with age, accounting for approximately one third of hospitalizations for cardiac rhythm disturbances [1]. AF is characterized by the replacement of consistent P-waves on the ECG by rapid oscillations (fibrillatory waves) that vary in amplitude, frequency, and shape, associated with an irregular ventricular response. AF affects approximately 10% of the population over age of 75 and is associated with an increased risk of stroke [2, 1]. Noninvasive analysis of AF and other supraventricular tachyarrhythmia conditions require cancellation of the ventricular activity (QRS complex and T-wave) from the ECG [3, 4]. Due to the overlapping spectral contents of atrial and ventricular activities, linear filtering is not possible, and also the higher dominance of ventricular activity in the ECG makes the extraction of atrial activity non-trivial [5, 6].

First type of methods proposed in literature, known as average beat subtraction (ABS) methods work on the assumptions that time invariance of wave morphology, and short-time stationarity of the signal. These assumptions often break in the presence of ectopic beats, and movement of electrodes,

thus significant amount of residual QRS-T can be present in the extracted atrial signal [7]. The performance could be further limited for signals with short durations, as constructing a good QRS-T template become difficult [8]. Second type, known as blind source separation methods, explore the spatial diversity of multi-lead ECG, and techniques such as independent component analysis, and principal component analysis are employed [6, 9]. Performance of source separation techniques significantly reduces for fewer number of leads, i.e., Holter recordings with few electrodes available [10]. Following atrial activity extraction, the time-frequency analysis is usually performed to identify the dominant AF frequency and its time variation [3, 5], but neither of above methods are capable of integrating time-frequency analysis into the original algorithm and it has been performed as a separate step.

The proposed method in this paper integrate atrial activity extraction and time-frequency analysis into a single algorithm, thus more efficient real-time analysis of AF from ECG is viable. Atrial activity during the QRS-T duration was estimated by the information from immediately preceding and following segments of AF signal. A mode limited short-time expansion of an orthogonal basis function set was used to reconstruct the atrial activity from the ECG. An analytical expression for the time-frequency distribution was derived solely based on coefficients of the basis set, thus time-frequency plots can be constructed with less computations to observe time variation of dominant AF frequency. Additionally, proposed algorithm can be generally applicable to time-frequency analysis of other non-stationary signals with missing or corrupted data segments.

2. METHODS

In this section, we introduce a method to reconstruct the missing or corrupted data segments of a signal based on an expansion of orthogonal basis set with Tikhonov regularization. The observed ECG was considered as the AF signal which has been corrupted by the ventricular activity. Then, we apply the reconstructed signal in the standard time-frequency distributions to derive a simplified analytical form of the TFD in terms of basis coefficients.

Let $s(t)$ be the observed signal with corrupted data segments. By replacing the corrupted segments by zeros, the known signal can be written as,

$$\tilde{s}(t) = \sum_{i=1}^N \mathcal{X}_{[a_i, b_i]}(t) s(t), \quad (1)$$

where $\mathcal{X}_{[\cdot, \cdot]}$ is the indicator function, i.e.,

$$\mathcal{X}_{[a, b]}(t) = \begin{cases} 1 & t \in [a, b] \\ 0 & t \notin [a, b]. \end{cases} \quad (2)$$

Here, without loss of generality, we assume that signal is correctly known in the intervals $t \in \{[a_i, b_i]_{i=1}^N\}$, and corrupted in between. Now consider the interval $[a_i, b_{i+1}]$. We use the information available in the intervals $[a_i, b_i]$ and $[a_{i+1}, b_{i+1}]$ to reconstruct the corrupted segment (b_i, a_{i+1}) . This construction based on localized information preserves the good localization in the time-frequency domain.

Let $f_i(t)$ be the unknown signal (AF signal) over the interval $[a_i, b_{i+1}]$. In general, most signals have a frequency band of interest, as in the case of AF signal fibrillatory frequencies are limited to 15Hz with possible harmonics, thus a mode limited orthonormal basis set can be used to represent $f_i(t)$. Let $\{\phi_i^n(t)\}$ be a orthonormal basis set in the interval $T_i = b_{i+1} - a_i$. Assuming that $f_i(t)$ is mode limited to N with respect to $\{\phi_i^n(t)\}$, then

$$f_i(t) = \sum_{n=-N}^N c_i^n \phi_i^n(t) \quad a_i \leq t \leq b_{i+1}, \quad (3)$$

where c_i^n are the localized coefficient set of the basis functions in the interval i . The best number of modes is a trade off between capturing underlying signal and improving the noise robustness, and can be optimized empirically.

Coefficients c_i^n can be estimated via regularized least square solution given below. Using the known information in the intervals $[a_i, b_i]$ and $[a_{i+1}, b_{i+1}]$, we get

$$\begin{pmatrix} \tilde{s}(a_i) \\ \vdots \\ \tilde{s}(b_i) \\ \tilde{s}(a_{i+1}) \\ \vdots \\ \tilde{s}(b_{i+1}) \end{pmatrix} = \begin{pmatrix} \phi_i^{-N}(a_i) & \dots & \phi_i^N(a_i) \\ \vdots & \ddots & \vdots \\ \phi_i^{-N}(b_i) & \dots & \phi_i^N(b_i) \\ \phi_i^{-N}(a_{i+1}) & \dots & \phi_i^N(a_{i+1}) \\ \vdots & \ddots & \vdots \\ \phi_i^{-N}(b_{i+1}) & \dots & \phi_i^N(b_{i+1}) \end{pmatrix} \begin{pmatrix} c_i^{-N} \\ \vdots \\ c_i^N \end{pmatrix}, \quad (4)$$

or $\tilde{\mathbf{s}} = \Phi \mathbf{c}$ in short.

The above system will be over-determined given $(b_i - a_i) + (b_{i+1} - a_{i+1}) > 2N + 1$, and will be the usual case provided observed signal is sufficiently sampled. Otherwise the given system will be under-determined. Regardless of the undeterministic nature, regularization can be applied to overcome the associated problems. It also avoids the modeling of noise, and guarantees a bounded solution. We use Tikhonov regularization [11] in solving above system, and the solution can be written in the form

$$\mathbf{c} = (\Phi^T \Phi + \lambda^2 \mathbf{I})^{-1} \Phi^T \mathbf{s}, \quad (5)$$

where \mathbf{I} is the identity matrix, λ is the regularization parameter, and \top stands for the transpose of the matrix.

We use estimated coefficients to find an analytical expression for the time-frequency distributions. Wigner-Ville distribution was used here as it achieves high time and frequency resolutions. Although it is susceptible to cross-terms for multi-component signals, due to the presence of a dominant AF spectral component, interference terms are not significant. Nevertheless, it should be emphasized that any time-frequency distribution can be applied here, for example members of Cohen family [12], but simplified analytical expressions may not always found depending on the kernel function used in the distribution. Wigner-Ville distribution is given by

$$WV(t, \omega) = \frac{1}{2\pi} \int e^{-j\omega\tau} f(t + \tau/2) f^*(t - \tau/2) d\tau. \quad (6)$$

Replacing $f(\cdot)$ by the reconstructed signal in (3) with the coefficient set found as in (5), we get

$$WV_i(t, \omega) = \frac{1}{2\pi} \sum_{n=-N}^N \sum_{m=-N}^N c_i^n c_i^{m*} \int e^{-j\omega\tau} \phi_i^n(t + \tau/2) \times \phi_i^{m*}(t - \tau/2) d\tau. \quad (7)$$

The integral in (7) can be further simplified as $\{\phi_i^n(t)\}$ is only dependent on the interval i . Due to the frequency modulated harmonic nature of the AF signal, the natural choice for the basis set was considered to be the complex exponential functions

$$\phi_i^n(t) = e^{j2\pi n t / T_i}, \quad (8)$$

where $T_i = b_{i+1} - a_i$. Nevertheless, any valid orthonormal basis such as Fourier-Bessel, Fourier-Legendre can be used here depending on the application domain. Given that the complex exponentials are used as the basis, (7) further simplifies to

$$WV_i(t, \omega) = \frac{1}{2\pi} \sum_{n=-N}^N \sum_{m=-N}^N c_i^n c_i^{m*} e^{-j\frac{2\pi}{T_i}(n-m)t} \times \delta[\omega - \frac{\pi}{T_i}(n+m)]. \quad (9)$$

Thus, expression (9) provides a computationally efficient way of constructing the Wigner-Ville distribution using estimated coefficient vectors without explicitly reconstructing the signal itself.

3. RESULTS

Algorithm was first applied to simulated ECG with AF episodes. ECGs were generated by dynamical model of motion [13] which was modified by removing the P-wave, and adding an AF signal. The AF signal was modeled by a sum of frequency modulated sinusoidals with time varying amplitude and its harmonics [7, 14],

$$g(t) = \sum_{k=1}^{M+1} a_k(t) \cos[k\omega_0 t + \frac{\Delta\omega}{\omega_f} \sin(\omega_f t)] + n(t) \quad (10)$$

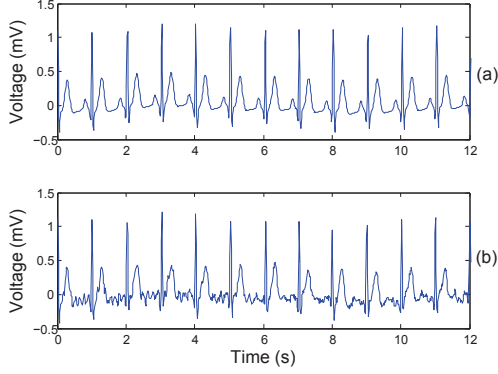


Fig. 1. (a) Simulated ECG signal for normal sinus rhythm generated by the model in [13] (b) P-wave of ECG replaced by the simulated AF signal (10).

where $a_k(t) = e^{-\gamma(k-1)}(a + \Delta a \sin(\omega_a t))$, ω_0 is the dominant AF frequency, ω_f is the frequency of frequency modulation, $\Delta\omega$ is the maximum frequency deviation, M is the number of harmonics excluding the fundamental, γ is the decaying factor of harmonics, a is the average amplitude of the fundamental, ω_a is the frequency of amplitude modulation, and Δa is the maximum amplitude deviation. $n(t)$ represents white Gaussian noise, and other ECG artifacts. AF frequency is given by

$$\omega_{AF}(t) = \omega_0 + \Delta\omega \cos(\omega_f t). \quad (11)$$

For the simulations following parameters were used: $\omega_0 = 2\pi 5$ rad/s, $\omega_f = 2\pi 0.05$ rad/s, $\Delta\omega = 3\pi 2$ rad/s, $M = 3$, $\gamma = 0.9$, $a = 0.05$ mV, $\omega_a = 2\pi 0.1$ rad/s, $\Delta a = 0.01$ mV. Figure 1 shows a simulated ECG signal for sinus rhythm and AF signal superimposed onto ECG after P-wave removal.

Number of pre-processing steps were conducted before applying the proposed algorithm which included baseline wander removal, R-peak detection, and fiducial point identification. Accurate identification of fiducial points were important in constructing intervals $[a_i, b_i]$, and a_i was marked as the offset of T-wave while b_i was marked as the onset of QRS complex. All the pre-processing steps were automated, so that real-time processing is viable.

Empirical optimization with the criterion of maximum correlation between reconstructed AF signal from ECG and original AF signal yields to number of modes $N = 18$, and the regularization parameter $\lambda = 1.7$. The ABS method was applied to the same ECG data, and performances were compared. ABS was performed by constructing a mean QRS-T template, and template subtraction from each beat after spatiotemporal alignment.

Figure 2 shows the time-frequency distributions for original and reconstructed AF signals. It is evident that proposed scheme closely resembles the original AF time-frequency plot compared to the ABS based reconstructed. Deviations in the ABS reconstructed time-frequency plot are mainly due to the

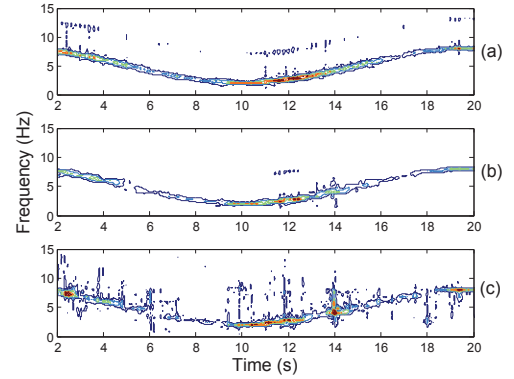


Fig. 2. Time-frequency distributions (9) for (a) simulated AF signal (10) (b) proposed AF reconstruction (c) ABS reconstruction.

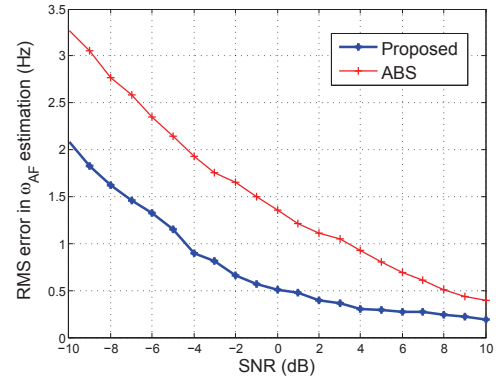


Fig. 3. Comparison between proposed method and ABS method in terms of RMS error of the estimated dominant AF frequency.

insufficient QRS-T cancellation. Also it should be emphasized that the ABS method required reconstructing the signal and performing the time-frequency analysis as a two step process, where as the proposed method has constructed the time-frequency plot straight from the basis coefficients vectors, thus less processing time was required.

We computed the RMS errors in dominant AF frequency estimation for a range of SNR values to quantify the performance, and similar analysis was done for ABS method as well for the comparison. Results are shown in Figure 3. Dominant AF frequency was estimated as the frequency where maximum energy is observed on the time-frequency plot. For all considered SNRs proposed method performs better than ABS method, and at low SNRs, which is of high interest especially for Holter monitors, proposed method outperforms the ABS method considerably. Proposed scheme reduces the error to be within 0.5Hz for all SNRs greater than 0dB, where as in order to achieve the same error limits ABS requires the SNR to be greater than 8dB.

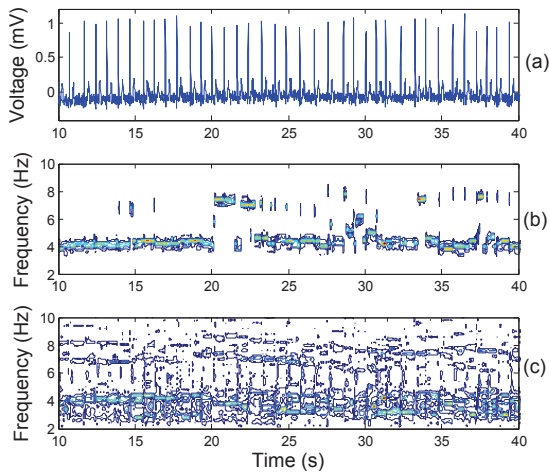


Fig. 4. (a) Physionet Atrial Fibrillation Termination Database record n09 (b) TFD by proposed method (c) TFD via ABS reconstruction.

We show the applicability of the proposed method in analyzing real ECG with AF episodes by applying the algorithm in ECG records in Physionet Atrial Fibrillation Termination Database. The example shown in Figure 4 shows that proposed scheme estimates the dominant AF frequency to a higher degree of resolution compared to the standard ABS method.

4. CONCLUSION

A novel algorithm was introduced to perform time-frequency analysis of Atrial Fibrillation ECG, and to identify temporal evolution of dominant AF frequency. Superiority of the algorithm was shown by comparing it to the standard ABS method. The algorithm can be easily adaptable to time-frequency analysis of any non-stationary signal with missing or corrupted data segments.

5. REFERENCES

- [1] V. Fuster, L. E. Ryden, and et al., "Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation," *Amer. J. Cardiol.*, vol. 48(4), pp. 854–906, Aug 2006.
- [2] S. N. Singh, X. C. Tang, and et al., "Quality of life and exercise performance in patients in sinus rhythm versus persistent atrial fibrillation," *Amer. J. Cardiol.*, vol. 48(4), pp. 721–730, Aug 2006.
- [3] A. Bollmann, N. A. Kanuru, K. K. McTeague, P. F. Walter, D. B. DeLurgio, and J. J. Langberg, "Frequency analysis of human atrial fibrillation using the surface electrocardiogram and its response to ibutilide," *Amer. J. Cardiol.*, vol. 81, pp. 1439–1445, Jan 1998.
- [4] S. Petrutiu, J. Ng, G. M. Nijm, H. Al-Angari, S. Swiryn, and A. V. Sahakian, "Atrial fibrillation and waveform characterization: A time domain perspective in the surface ECG," *IEEE Eng. in Med. and Bio. Mag.*, vol. 25, no. 6, pp. 24–30, 2006.
- [5] A. Bollmann, D. Husser, L. Mainardi, F. Lombardi, P. Langley, A. Murray, J. J. Rieta, J. Millet, S. B. Olsson, M. Stridh, and L. Sornmo, "Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications," *Europace*, vol. 8, no. 11, pp. 911–926, 2006.
- [6] J.J. Rieta, F. Castells, C. Sanchez, V. Zarzoso, and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1176–1186, 2004.
- [7] M. Stridh and L. Sornmo, "Spatiotemporal QRST cancellation techniques for analysis of atrial fibrillation," *IEEE Trans. Biomed. Eng.*, vol. 48(1), pp. 105–111, Jan 2001.
- [8] M. Lemay, J. M. Vesin, A. van Oosterom, V. Jacquemet, and L. Kappenberger, "Cancellation of ventricular activity in the ECG: Evaluation of novel and existing methods," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 542–546, 2007.
- [9] F. Castells, J.J. Rieta, J. Millet, and V. Zarzoso, "Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 258–267, 2005.
- [10] R. Alcaraz and J. J. Rieta, "Adaptive Singular Value QRST Cancellation for the Analysis of Short Single Lead Atrial Fibrillation Electrocardiograms," *Proc. Computers in Cardiology*, vol. 34, pp. 513–516, 2007.
- [11] A. Ribes and F. Schmitt, "Linear inverse problems in imaging," *IEEE Sig. Proc. Mag.*, vol. 25, no. 4, pp. 84–99, 2008.
- [12] L. Cohen, "Time-frequency distributions-a review," *IEEE Proc.*, vol. 77, no. 7, pp. 941–981, Jul 1989.
- [13] P.E. McSharry, G.D. Clifford, L. Tarassenko, and L.A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, 2003.
- [14] S. Kodituwakku, T. D. Abhayapala, and R. A. Kennedy, "Atrial fibrillation analysis using Bessel kernel based time frequency distribution technique," *Proc. Computers in Cardiology*, vol. 35, pp. 837–840, Sep 2008.