Detecting Flutter Waves in the Electrocardiogram Using Generalized Likelihood Ratio Test

Muhammad Haziq Kamarul Azman[1][2], Olivier Meste[1], Kushsairy Kadir[2]

[1] Université Côte d'Azur, CNRS, I3S, France
[2] Universiti Kuala Lumpur, Malaysia

Abstract

*Accurate detection of f waves during atrial flutter is difficult. However, f waves contain information on the pathology and are useful for non-invasive diagnosis. The setup and difficulties of f wave detection lends itself to the use of statistical signal detection techniques. Real-life constraints can be modeled in the signal observation using several parameters in order to produce signal detectors with good performance. Several detectors were developed and tested using real 12-lead ECG recordings with manually annotated f wave markers available. At the end, a simple detector is obtained with relatively good detection performance (*AUC = 0.89, (Se, Sp) = (0.84, 0.81)*) and a threshold is available for use in automatic detection of f waves.*

1. Introduction

In many cases, it is necessary to detect and delineate different deviations (P, Q, R, S and T waves) visible on an electrocardiogram (ECG) to obtain diagnostic information on cardiac pathologies in a non-invasive manner. This allows us to obtain different timing and morphology-related measures which are closely related to cardiac functionality.

In supra-ventricular arrhythmia such as atrial flutter (AFL), the P wave, commonly referred to as f waves is of interest. They represent the continuous pathological activation of the atrium, driven by depolarizing wavefronts rotating around both atria using a specific circuit. A beat-to-beat study of these f waves have shown to provide useful information for localizing AFL circuits [1].

Accurate detection of these f waves is difficult as their amplitudes are small compared to noise and other deviations. During low atrio-ventricular conduction blocks ($<$ 3:1), they can be overlapped by T waves. Furthermore, ECG recordings are usually embedded in noise that are non-Gaussian. Remarkably, the description above can be represented as mathematical signal models. This allows the use of statistical signal detection techniques [2]. They allow us to obtain detectors which can be parameterized to account for e.g. weak-amplitude signals, T wave overlaps and non-Gaussian noise distribution.

In this paper, we develop several detectors using the generalized likelihood ratio test (GLRT) procedure [2]. The development accounts for several variations of the observed signal model, which allow us to better model real-life constraints. We then test these detectors using real 12-lead AFL recordings with known f wave onset markers and evaluate their performance using standard metrics. Finally, we show how the detection threshold, used to decide if a signal is present or not is learned using a simple machine learning approach. The threshold value can be then be used for automatic detection of f waves.

2. Methodology

2.1. Dataset and workbench

25 recordings of 12-lead ECG from the Centre Hospitalier Princesse Grace in Monaco were acquired from patients with AFL during ablation procedures using an acquisition system (Bard, USA). All signals were filtered and downsampled to 250 Hz. The recordings contain manual annotations of f wave onsets (2926 manual annotations in total). Algorithms were implemented in MATLAB R2014b (MathWorks, USA) on a machine equipped with an Intel Core i7 6500U processor running at 3GHz and 8GB of memory.

2.2. Detection setup

f waves manifest on the ECG as a continuous waveform with a relatively stable beat-to-beat morphology. It is often possible to find a single f wave not overlapped on top of QRS or T waves, usually just before the onset of a QRS complex. This single f wave $\mathbf{s}$ of length N was manually segmented and serves as a reference for the signal of interest. Throughout the ECG recording, we observe multiple f waves which can be thought of as a version of the reference modulated by a factor A and delayed by an integer

n_1.

ECG recordings are typically perturbed by noise $\mathbf{w}$ originating from various sources. These noise are randomly distributed and has an associated probability density function (PDF) $p(\mathbf{w})$. In this paper, we consider two distributions: normal Gaussian and Laplacian. The latter is particularly common with electrophysiological signals, where the noise tends to be spiky and have a heavy-tailed distribution. For simplicity, we assume that the noise is independent and identically distributed from sample to sample.

Given a sample $\mathbf{x}$ of the recording data from a single lead, we aim to decide between a signal-present ($\mathcal{H}_1$) and signal-absent ($\mathcal{H}_0$) hypothesis, formulated as follows:

$$\mathcal{H}_0: \qquad \mathbf{x} = \mathbf{w}$$
$$\mathcal{H}_1: \quad \mathbf{x} = A\mathbf{s}_{n_1} + \mathbf{w}$$

It is shown that the detector which obtains optimum detection performance in this setup is found by calculating the likelihood ratio [2]:

$$L(\mathbf{x}) = \frac{p(\mathbf{x};\, A,\, n_1,\, \sigma_{w_1}^2,\, \mathcal{H}_1)}{p(\mathbf{x};\, \sigma_{w_0}^2,\, \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \qquad (1)$$

where the PDFs are parameterized by A, n_1 and σ_w^2 under each hypotheses, and γ is the detection threshold. When the likelihood ratio value is above γ, we decide $\mathcal{H}_1$ is true, and if it is below γ, we decide $\mathcal{H}_0$ is true.

2.3. Parameter estimation

The parameterized PDF depends on the values of the parameters and must be resolved to obtain a usable detector. Values are assumed to be either known or unknown. This allows the full use of any prior available information, and in the latter case, allows them to be estimated.

σ_w^2 was estimated using the maximum likelihood (ML) technique. The logarithm of (1) was differentiated with respect to σ_w^2 and set equal to 0. This can be thought of as estimating the value of σ_w^2 which minimizes the slope of the PDF curve, or which maximizes the PDF value for a given $\mathbf{x}$.

The value of A should be strictly positive; ML estimators are by default unconstrained. To obtain a constrained version, we used a non-negative least squares algorithm to estimate A under Gaussian noise [3]. Under Laplacian noise, the ML estimation of A consists of minimizing the sum of absolute values, and is solved using the method of steepest descent [4].

The estimator of n_1 is $\hat{n}_1 = \arg\max_{n_1} L(\mathbf{x})$ [2]. This is done by sliding the reference signal across the whole recording, and calculating the detector output at each time instant. The estimate for n_1 is then the peak of the output.

2.4. Detector expressions

The detector is obtained by developing (1), after resolving all parameter values. The top half of Table 1 summarizes the different single-lead detector expressions assuming the combination of known or unknown parameter values. The lead of choice was selected as the one with the largest energy ratio between $\mathbf{s}$ and its corresponding QRS complex. $\hat{A}$ indicates an estimate of A using the technique discussed previously. Detector D_5 is known as an approximate to the original GLRT [2], with sgn the operator that outputs the sign of its argument. The symbol $^\mathrm{T}$ indicates a transpose.

2.5. T wave model and multilead extensions

The signal model can be modified to account for T waves that may overlap certain f waves. The hypotheses are reformulated as follows:

$$\mathcal{H}_0: \qquad \mathbf{x} = \mathbf{H}\mathbf{b}_0 + \mathbf{w}$$
$$\mathcal{H}_1: \quad \mathbf{x} = A\mathbf{s}_{n_1} + \mathbf{H}\mathbf{b}_1 + \mathbf{w}$$

where $\mathbf{H}$ is a matrix of basis functions and $\mathbf{b}$ the weights of each function. In this paper, we used polynomials of degrees 0 to 3 as a set of basis functions to estimate T wave shapes, denoted as $\mathbf{T} = \mathbf{H}\mathbf{b}$. The parameters $\mathbf{b}_0$ and $\mathbf{b}_1$ were estimated using computational methods under both noise distribution. The bottom half of Table 1 shows the estimators obtained using this model. $\mathbf{Q} = \mathbf{I} - \mathbf{H}(\mathbf{H}^\mathrm{T}\mathbf{H})^{-1}\mathbf{H}^\mathrm{T}$ is the orthogonal projector to the null space of $\mathbf{H}$.

Standard ECG recordings contain 12 leads. It is natural then to consider using all leads in hopes to obtain a better detector. The observations of each lead can be grouped into a matrix $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_{12}]$. The PDF of $\mathbf{X}$ requires information on the covariance structure between each lead. In this paper, we consider that the observations are uncorrelated amongst the leads and have the same variance. This results in the multilead PDF becoming the product of all single-lead PDFs, which translates into a sum of detector outputs when σ_w^2 is known and a product of detector outputs when σ_w^2 is unknown. This also applies when the T wave is modeled in a multilead setting.

Considering all possible configurations, there were 32 detectors in total (16 single-lead detectors and 16 multilead detectors).

2.6. Determination of the decision threshold and performance

An example of a detector output is shown in (Fig 1). The threshold γ separates true and false detection and in this

Table 1. Single-lead detector expressions

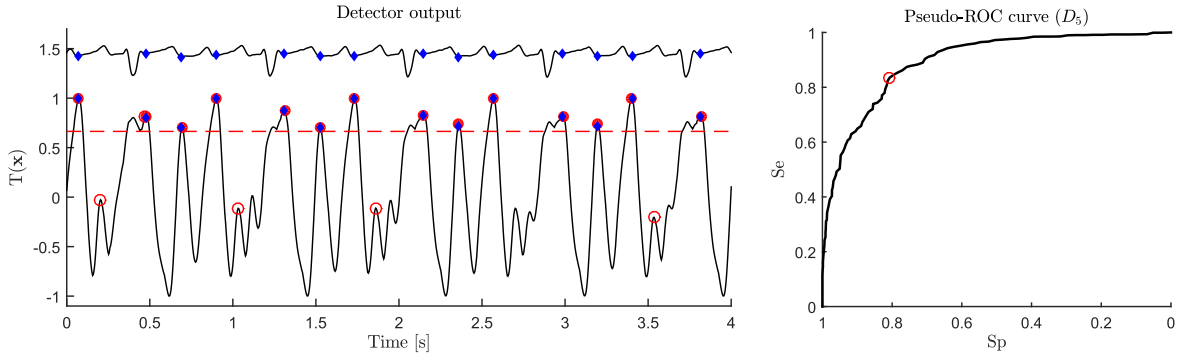| Parameter values	$\mathbf{w} \sim$ Gauss		$\mathbf{w} \sim$ Laplace	
A and σ_w^2 known	$D_1(\mathbf{x}) =$	$\mathbf{x}^{\mathrm{T}}\mathbf{s}$	$D_5(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} \mathrm{sgn}\,(x[n])s[n]$
A unknown	$D_2(\mathbf{x}) =$	$\hat{A}(2\mathbf{x}^{\mathrm{T}}\mathbf{s} - \hat{A}\mathbf{s}^{\mathrm{T}}\mathbf{s})$	$D_6(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n]| - |x[n] - \hat{A}s[n]|$
σ_w^2 unknown	$D_3(\mathbf{x}) =$	$\dfrac{\mathbf{x}^{\mathrm{T}}\mathbf{x}}{(\mathbf{x}-\mathbf{s})^{\mathrm{T}}(\mathbf{x}-\mathbf{s})}$	$D_7(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n]| \Big/ \sum\limits_{n=0}^{N-1} |x[n] - As[n]|$
A and σ_w^2 unknown	$D_4(\mathbf{x}) =$	$\dfrac{\mathbf{x}^{\mathrm{T}}\mathbf{x}}{(\mathbf{x}-\hat{A}\mathbf{s})^{\mathrm{T}}(\mathbf{x}-\hat{A}\mathbf{s})}$	$D_8(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n]| \Big/ \sum\limits_{n=0}^{N-1} |x[n] - \hat{A}s[n]|$
A and σ_w^2 known	$D_9(\mathbf{x}) =$	$\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{s}$	$D_{13}(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n] - \hat{T}_0[n]| - |x[n] - As[n] - \hat{T}_1[n]|$
A unknown	$D_{10}(\mathbf{x}) =$	$\hat{A}(2\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{s} - \hat{A}\mathbf{s}^{\mathrm{T}}\mathbf{Q}\mathbf{s})$	$D_{14}(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n] - \hat{T}_0[n]| - |x[n] - \hat{A}s[n] - \hat{T}_1[n]|$
σ_w^2 unknown	$D_{11}(\mathbf{x}) =$	$\dfrac{\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}}{(\mathbf{x}-\mathbf{s})^{\mathrm{T}}\mathbf{Q}(\mathbf{x}-\mathbf{s})}$	$D_{15}(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n] - \hat{T}_0[n]| \Big/ \sum\limits_{n=0}^{N-1} |x[n] - As[n] - \hat{T}_1[n]|$
A and σ_w^2 unknown	$D_{12}(\mathbf{x}) =$	$\dfrac{\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}}{(\mathbf{x}-\hat{A}\mathbf{s})^{\mathrm{T}}\mathbf{Q}(\mathbf{x}-\hat{A}\mathbf{s})}$	$D_{16}(\mathbf{x}) =$	$\sum\limits_{n=0}^{N-1} |x[n] - \hat{T}_0[n]| \Big/ \sum\limits_{n=0}^{N-1} |x[n] - \hat{A}s[n] - \hat{T}_1[n]|$

Figure 1. Example of a detector output (bottom trace; D_5 using a single lead) and its corresponding ECG record (top trace; lead V1) and pROC curve (right). Red circles indicate detection peaks: filled ones being true detection. Blue diamonds indicate manual annotations. The dashed line indicates the best threshold selected using the method described in Section 2.6. There is a good agreement with the detections and manual annotations. The red circle on the pROC curve marks the point of best threshold.

paper, machine learning methods are applied to determine an optimal threshold γ_{opt} that performs the best separation. This threshold can then be used with new recordings to perform automatic detection of f waves. An approach based on leave-one-out (LOO) cross-validation was used to obtain this threshold.

The peaks for each recording were classed as either true or false detection by comparing the instant of a detector peak to manual annotations. We used a limit of ±25 ms from a true annotation and selection of highest peak as criteria to decide a true detection. The remaining peaks are classed as false detection. Peaks within the interval of QRS complexes were discarded. To standardize the detection output across the whole dataset of recordings, the peaks for each recording were normalized by its largest-valued peak, producing maximum values that are always 1 for all recordings.

For a fixed threshold γ', we applied a LOO cross-validation algorithm. First, the peaks from the detector output of a single recording $\{y_{\text{test}}^{(1)} \cdots y_{\text{test}}^{(M)}\}$ with labels $\{l_{\text{test}}^{(1)} \cdots l_{\text{test}}^{(M)}\}$ were held out and the peaks from the remaining recordings were grouped together. Using the remaining recordings, we calculate $\mathrm{P}_{above}^{\mathrm{True}}$, the probability that the remaining labels are true knowing that their associated peak value is above γ', and likewise for false ones, noted $\mathrm{P}_{above}^{\mathrm{False}}$. The same is also done for samples below the threshold, producing $\mathrm{P}_{below}^{\mathrm{True}}$ and $\mathrm{P}_{below}^{\mathrm{False}}$.

Validation is then performed by predicting each test la-

bel $l_{\text{test}}^{(m)}$ through a comparison of two probabilities (P_{above}^{True} and P_{above}^{False} or P_{below}^{True} and P_{below}^{False}) depending on the location of $y_{\text{test}}^{(m)}$ with regards to γ', and deciding the class based on which probability is larger. The predicted label is tested against $l_{\text{test}}^{(m)}$: if matching, then the counter C_+^{True} or C_-^{True} is incremented based on whether $l_{\text{test}}^{(m)}$ is true or false. If not matching, then the counter C_+^{False} or C_-^{False} is incremented likewise. The algorithm then repeats by putting back the held out recording and selecting a different test recording, until all recordings have been used. At the end, the counters indicate the total number of true positives, true negatives, false positives and false negatives respectively for γ'. After all values of γ have been evaluated, standard performance metrics such as sensitivity Se, specificity Sp, accuracy Acc can be calculated for each value of γ:

$$\text{Se} = C_+^{\text{True}}/(C_+^{\text{True}} + C_-^{\text{False}})$$
$$\text{Sp} = C_-^{\text{True}}/(C_-^{\text{True}} + C_+^{\text{False}})$$
$$\text{Acc} = \frac{C_+^{\text{True}} + C_-^{\text{True}}}{C_+^{\text{True}} + C_-^{\text{False}} + C_+^{\text{False}} + C_-^{\text{True}}}$$

and a pseudo-receiver-operating characteristic (pROC) curve can be derived, allowing us to calculate AUC: the area under the pROC curve. All metrics are valued between 0 and 1. The threshold that maximizes the Youden statistic $J = \text{Se} + \text{Sp} - 1$ is deemed the optimal threshold γ_{opt}.

3.   Results and Discussion

The overall performance of the detectors can be quantified by the AUC. Of all 32 detectors, the single-lead version of D_5 and D_{10} have the highest AUC value (AUC = 0.89).

Table 2. Best detector performances and threshold

	1-lead D_5		1-lead D_{10}		M-lead D_{15}	
AUC	0.89		0.89		0.88	
Se, Sp	0.84	0.81	0.91	0.79	0.74	0.87
Acc	0.82		0.80		0.85	
γ_{opt}	0.66		2.98×10^{-4}		8.01×10^{-7}	

The thresholds issued by the algorithm and their corresponding values of Se and Sp are shown in Table 2 for the 3 best detectors in terms of AUC. We observe similar performance for D_5 and D_{10} in this regard, therefore a choice of detector must be made. In an automatic detection setting for use in a beat-to-beat analysis, it is preferable to have less false alarms. The performance of the two detectors are comparable in specificity, however, D_{10} generates much more false alarms (35455 false alarms) than D_5

(1352 false alarms) over all the recordings. This makes D_5 a better choice despite a lower sensitivity. The timing error for detector D_5 is 1.26 ± 4.25 ms, given as the mean $\pm$ the standard deviation.

Although the multilead version of D_{15} has a much lower rate of false alarms, the estimation of the different parameters of the signal model required massive computation time (~ 13 hours for a single record). This is partly due to the requirement of the detector expression and makes it unpractical for use. The expression may be simplified further, or other non-GLRT strategies that do not require parameter estimation may be considered [2].

4.   Conclusion

In this paper, we developed several f wave detectors based on the principles of generalized likelihood ratio test. The detectors are adapted to real-life conditions by accounting for various effects (modulation, T wave overlap, non-Gaussian noise) by using various signal models and parameterization of the likelihood ratio. The detection threshold–useful for automatic detection of f waves–is learned using an original technique based on leave-one-out cross-validation. The detector of choice shows good performance with a small timing error.

Acknowledgements

References

[1] Kamarul Azman MH, Meste O, Kadir K, Latcu DG. Localizing atrial flutter circuit using variability in the vectorcardiographic loop parameters. In Computing in Cardiology, volume 45. September 2018; .

[2] Kay S. Fundamentals of statistical signal processing: detection theory. Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1998.

[3] Lawson C, Hanson R. Solving least squares problems. Society for Industrial and Applied Mathematics, 1995.

[4] Singleton RR. A method for minimizing the sum of absolute values of deviations. Ann Math Statist 09 1940;11(3):301–310.

Address for correspondence:

Muhammad Haziq Kamarul Azman
Bureau 206
Laboratoire I3S
2000 route des Lucioles
06900 Sophia Antipolis, France
mhaziq@unikl.edu.my