Available online at www.sciencedirect.com
ScienceDirect
Journal of Electrocardiology 42 (2009) 522–526
JOURNAL OF Electrocardiology
www.jecgonline.com

Improvements in atrial fibrillation detection for real-time monitoring

Saeed Babaeizadeh, PhD,* Richard E. Gregg, MS, Eric D. Helfenbein, MS, James M. Lindauer, MD, Sophia H. Zhou, PhD

Advanced Algorithm Research Center, Philips Healthcare, Thousand Oaks, CA, USA

Received 13 May 2009

Abstract Electrocardiographic (ECG) monitoring plays an important role in the management of patients with atrial fibrillation (AF). Automated real-time AF detection algorithm is an integral part of ECG monitoring during AF therapy. Before and after antiarrhythmic drug therapy and surgical procedures require ECG monitoring to ensure the success of AF therapy. This article reports our experience in developing a real-time AF monitoring algorithm and techniques to eliminate false-positive AF alarms. We start by designing an algorithm based on R-R intervals. This algorithm uses a Markov modeling approach to calculate an R-R Markov score. This score reflects the relative likelihood of observing a sequence of R-R intervals in AF episodes versus making the same observation outside AF episodes. Enhancement of the AF algorithm is achieved by adding atrial activity analysis. P-R interval variability and a P wave morphology similarity measure are used in addition to R-R Markov score in classification. A hysteresis counter is applied to eliminate short AF segments to reduce false AF alarms for better suitability in a monitoring environment. A large ambulatory Holter database (n = 633) was used for algorithm development and the publicly available MIT-BIH AF database (n = 23) was used for algorithm validation. This validation database allowed us to compare our algorithm performance with previously published algorithms. Although R-R irregularity is the main characteristic and strongest discriminator of AF rhythm, by adding atrial activity analysis and techniques to eliminate very short AF episodes, we have achieved 92% sensitivity and 97% positive predictive value in detecting AF episodes, and 93% sensitivity and 98% positive predictive value in quantifying AF segment duration.

Keywords: Electrocardiogram; Atrial fibrillation; Patient ECG monitoring; AF monitoring algorithm

Introduction

Atrial fibrillation (AF) is the most common arrhythmia in clinical practice worldwide. It has been estimated that 2.2 million Americans have paroxysmal or persistent AF. Hospital admissions for patients with AF have increased by 66% in the last 20 years. Atrial fibrillation is associated with an increased risk of stroke, heart failure, and all-cause mortality.[1] The prevalence of AF increases with age, reaching as high as 9% in octogenerians.[2] A high incidence of AF is also observed in patients who have undergone a pulmonary vein isolation procedure. Atrial fibrillation is responsible for 15% to 20% of all strokes. Recent findings indicated that acute onset of AF may contribute to the hypercoagulable state.[3] The mortality rate of patients with AF is higher than that of patients in normal sinus rhythm and is linked to severity of underlying heart disease.[2] Reoccurrence of AF is very common, and 30% to 50% of patients will experience recurrent AF within a year after conversion therapy.[4] Patients with enlarged atria and heart failure have increased risk of recurrent AF. Electrocardiographic (ECG) rhythm monitoring is highly recommended for efficient evaluation of AF surgical[5] or medical therapy whether the patient has symptom or not. It is known that continuous ECG monitoring has an advantage over intermittent ECG monitoring.[6] Therefore, a continuous automated AF detection algorithm is clinically desirable.

Electrocardiographic presentation of AF is continuous and rapid chaotic electrical activity of the atria and absence of P wave. Ventricular response is poorly coupled with atrial activity because the atrial rate exceeds the conducting capability of the atrioventricular node. The ventricular rate in

* Corresponding author. Advanced Algorithm Research Center, Philips Healthcare, 1525 Rancho Conejo Blvd., Suite 100, Thousand Oaks, CA 91320, USA.

E-mail address: saeed.babaeizadeh@philips.com

0022-0736/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jelectrocard.2009.06.006

Table 1
The database used for algorithm development was a large collection of ambulatory Holter ECGs (n = 633)

Records (n)				AF		AFL		AF/AFL		Non-AF
Total	AF	AFL	AF/AFL	Episode (n)	Duration (min)	Episode (n)	Duration (min)	Episode (n)	Duration (min)	Duration (min)
633	162	17	4	274	6929	60	543	4	927	47 737

This database was gathered from multiple sources with sampling rates ranging from 128 to 1000 sps and resolutions from 0.31 to 10 μV/lsb. The ECGs were annotated for AF, AFL, AF/AFL, and non-AF episodes by 2 experienced human readers independently. The onset and offset of the episodes were marked for calculation of episode duration. AFL was separated from AF when either was distinct. AF/AFL was used when a distinction between the 2 could not be made. AF/AFL was counted as AF in the development process, but AFL was not. Of the total 633 records, 162 had AF episodes, 17 had AFL episodes, and 4 had mixed AF/AFL episodes. Some of the ECG records contained more than 1 type of arrhythmia. The total number of episodes and duration of AF, AFL, AF/AFL, and non-AF in the development database is listed.

AF is highly irregular and sometimes rapid due to the random series of consecutive ventricular cycle lengths. The atrial waveform called F wave is also highly irregular due to simultaneous presence of multiple reentrant pathways with different rates. These hallmark characteristics of AF make ECG monitoring the most convenient tool to assist in AF diagnosis and therapy and to detect relapse of AF in the hospital or home.

Systems and algorithms have been developed for identifying AF from the ECG waveform of patients with symptoms. These algorithms typically rely on 1 or more of general characteristics of the waveform including irregularly irregular rhythm, high-frequency chaotic atrial waveform, and absence of P waves. Existing AF detection algorithms are sometimes unable to specify AF over some other arrhythmias and may misclassify other irregular rhythms as AF, resulting in false alarms. Development of an algorithm to detect AF with high accuracy and low false alarm rate is a major challenge. For the purpose of calculating AF burden, knowing the exact duration when the patient is in AF is also important.[7] The AF algorithm should be able to detect the onset and offset of AF episodes. Arrhythmias with similar characteristics, such as frequent premature atrial contractions, often confuse the AF detection algorithms. This study reports our experience with the development of an automated AF detection algorithm, the impact of accurate P wave detection in preventing misclassifying non-AF irregularly regular rhythms, and the techniques applied to eliminate false AF alarms.

Materials and methods

ECG Recordings

Two databases were used in this study. For algorithm development, we used a large collection of ambulatory Holter ECGs (n = 633). The ECGs were annotated by 2 expert human readers independently for episodes of AF, atrial flutter (AFL), AF/AFL and non-AF episodes. Atrial flutter was separated from AF when either was distinct. Atrial fibrillation/atrial flutter was used when a distinction between the two could not be made. The non-AF segments were usually sinus rhythm or arrhythmias other than AF and AFL. The onset and offset of the episodes were marked for calculation of episode duration and, ultimately, the AF burden. In the development process, AF/AFL was counted as AF, but AFL was not. Of the total 633 Holter ECG records, 162 had AF episodes, 17 had AFL episodes, and 4 had mixed AF/AFL episodes. Some ECG records contained more than 1 type of arrhythmia. The total number of episodes and duration of AF, AFL, AF/AFL, and non-AF in the development database is listed in Table 1. The development database was gathered from multiple sources, with sampling rates ranging from 128 sample-per-second (sps) to 1000 sps and resolutions from 0.31 to 10 μV/least significant bit (lsb).

For the purpose of algorithms validation, we used the MIT-BIH AF database from publicly accessible ECG database PhysioNet.[8] This AF database includes 23 long-term 2-channel ECG recordings as summarized in Table 2. Each record is approximately10 hours in duration and contains 2 ECG channels sampled at 250 sps with 12-bit resolution over a range of ±10 mV. All 23 records have AF episodes. Of the 23 records, 8 also have AFL episodes. Using this database allows us to compare our results with previously published algorithms.

Statistical approach of AF detection algorithms

Real-time ECG processing involves signal conditioning, beat detection, and beat classification. This study will not cover those steps and will only focus on AF detection.

We developed 2 statistical classifiers for detecting AF. The first classifier relies only on R-R interval. R-R interval is usually highly irregular in AF, unless the patient responds to rate-controlling medication. The second classifier is based on a combination of R-R interval irregularity and absence of a P

Table 2
The database used for AF algorithm testing is the publicly available MIT-BIH AF database with a total of 23 long-term ECGs

Records			AF		AFL		Non-AF
Total	AF	AFL	Episode (n)	Duration (min)	Episode (n)	Duration (min)	Duration (min)
23	23	8	291	5604	14	98	8355

All records have AF episodes, mostly paroxysmal. Each record is approximately 10 hours in duration and contains 2 ECG channels sampled at 250 samples per second with 12-bit resolution over a range of ±10 mV. Of the 23 records, 8 of them also have AFL episodes.

Table 3
Performance results in the evaluation database (n = 23) are listed below for both algorithms, R-R interval–based algorithm, and combined R-R interval/atrial activity algorithm

Algorithm		Duration statistics					Episode statistics		
		True positive (min)	False positive (min)	Sensitivity (%)	Specificity (%)	PPV (%)	False positive (n)	Sensitivity (%)	PPV (%)
R-R only	Gross	5236	179	94	98	97	48	91	90
	Average			91	96	86		97	86
R-R and P wave	Gross	5219	118	94	99	98	26	91	94
	Average			89	96	88		96	90

The results are reported in 2 ways, gross and average. Gross is to calculate the statistics by combining all records into one. Average is to calculate the statistics for each individual record separately and then report the mean of all the individual statistics. Episodes shorter than 1 minute were ignored.
PPV indicates positive predictive value.

wave, or in other words, inconsistency in location and morphology of an automatically detected atrial activity peak considered as a P wave candidate.

Markov process modeling was used to analyze R-R interval irregularity. The calculated Markov score reflects the relative likelihood of observing a sequence of R-R intervals in AF episode versus making the same observation outside AF segment in recorded ECGs.[8] In this method, the sequence of R-R interval is assumed to be controlled by a stationary first-order Markov process characterized by a transition probability matrix. The matrix represents the probability of R-R interval transition from one state to another. In this matrix, elements lower than a threshold represent transitions that are relatively more likely to occur in AF than non-AF. Therefore, the first classifier, which uses this R-R irregularity only, compares the Markov score to a fixed threshold to classify the rhythm.

The second classifier includes 2 additional parameters associated with atrial activity, namely, P wave location and P wave morphology. As P wave location measurement, we use P-R interval variation, where P-R interval is defined as the time between the onset of P wave to onset of QRS, and its variation is the deviation from the average P-R interval. In sinus rhythm, the P-R interval variation is relatively small, and in atrial arrhythmia such as AF, due to the absence of P wave, this variation is either nonmeasurable if there is no measurable atrial activity, or very large if the algorithm detects some P wave–like atrial activity as a P wave candidate. As P wave morphology measurement, we used a similarity measure between 2 consecutive P waves. In sinus rhythm, the P waves usually match well when the signal is not too noisy, and in AF, the match is usually poor. After combining R-R interval Markov score with the 2 P wave measurements, we used a statistical approach to classify this group of measurements as either AF or non-AF. Having tried different classification approaches, such as Mahalanobis distance discriminant,[9] first principal component classifier,[10] and decision tree,[11] we found the decision tree advantageous over the other 2 for our purpose. The commonly used Mahalanobis distance discriminant classifier assumes that

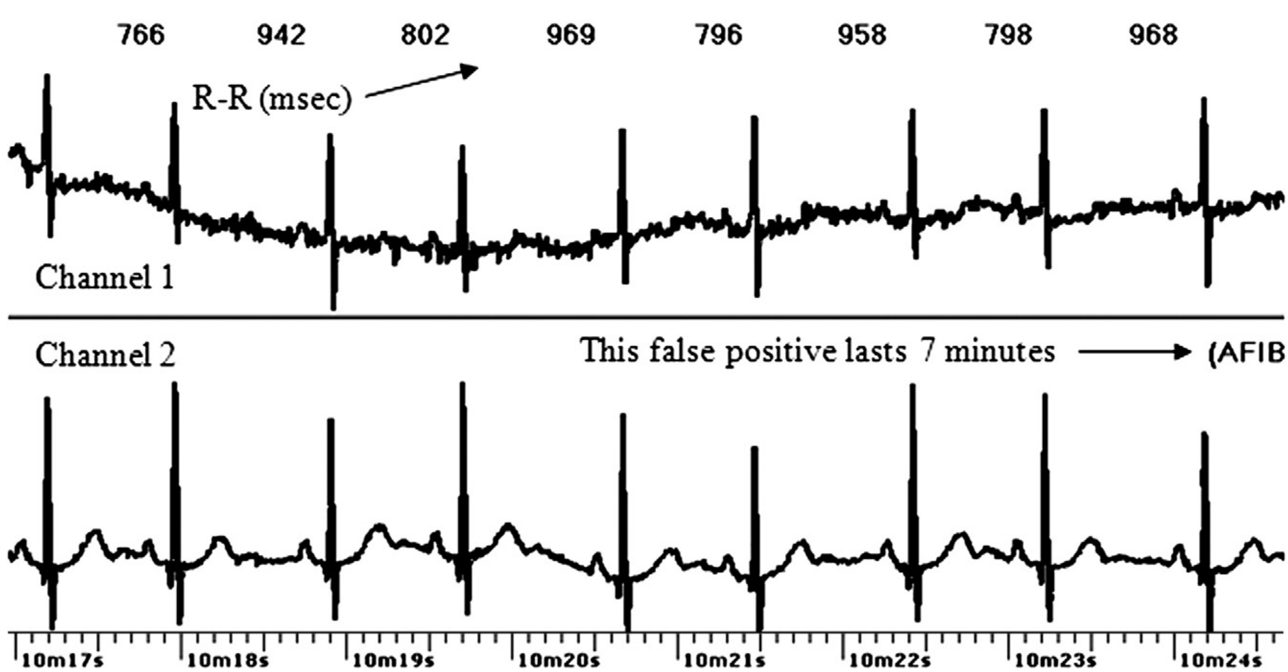

Fig. 1. A 2-channel ECG with irregular R-R intervals and clear P waves. Because of the irregular R-R interval, the R-R interval–based algorithm erroneously reports an AF episode that lasts 7 minutes. The beginning of this false episode of atrial fibrillation is shown as (AFIB) on channel 2. The algorithm with combined R-R and P wave analysis does not misclassify the irregular rhythm as AF because the P waves are correctly detected.

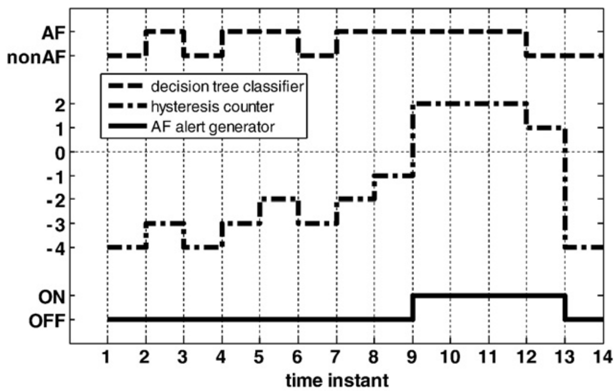

Fig. 2. An example of a nonsymmetric hysteresis counter with minimum of 4, threshold of 0, and maximum of 2 is shown. The dashed line on the top indicates the output of the decision tree classifier. The dot-dashed line in the middle illustrates the outcome of the hysteresis counter that counts up with every AF episode and counts down with any non-AF segment. The solid line at the bottom represents the behavior of AF alarm generator. The algorithm turns the AF alarm on if the hysteresis count rises above the predefined threshold of 0 and turns the AF alarm off if the hysteresis count drops below the defined threshold.

the measurements have a multivariate normal distribution when this assumption is invalid for AF classification data. The decision tree, on the other hand, is a set of simple rules that makes no prior assumption about the distribution of the measurements in either AF or non-AF groups.

Results

The performance of the R-R interval–based algorithm and the combined R-R interval/atrial activity algorithm is reported in Table 3. The method used for arrhythmia episode-by-episode comparison is the algorithm specified by Association for the Advancement of Medical Instrumentation (AAMI) standards for ambulatory ECG analyzers.[12] Atrial flutter episodes were treated as non-AF. All episodes shorter than 1 minute were excluded from episode and duration statistics. The results are reported in 2 ways, gross and average. The gross method calculates the statistics by combining all records into one. The average method calculates the statistics for each individual record separately and then reports the mean of all the individual statistics. Looking at the duration gross statistics, we have achieved 94% sensitivity and 98% positive predictive value using R-R interval and P wave information.

As seen in Table 3, analyzing atrial activity significantly reduces false positives and hence improves positive predictive value for both duration and episode statistics. This is achieved mainly by detecting valid P waves on an ECG recording with an irregularly irregular rhythm other than AF, which prevents the algorithm from misclassifying it as AF solely based on its rhythm. The difference between the 2 algorithms is well illustrated in Fig. 1. This example shows an irregular rhythm with clear P waves in channel 2. Because of the irregular R-R interval, the R-R interval–based algorithm erroneously reports an AF episode that lasts 7 minutes. The algorithm with combined R-R and P wave analysis does not misclassify the irregular rhythm as AF because the P waves are correctly detected.

AF alarm enhancement

In a real-time ECG monitoring environment, AF episodes are often announced by means of a soft alarm on the monitoring device. False alarms in ECG monitoring are very common because of variety of reasons. We found short segments of R-R irregularity, noise, and difficulties in P wave detection as the main sources of false alarms. Extra effort has been made to eliminate false alarms in our development. Further enhancement of the AF algorithm is achieved by adding a postclassification corrector after the decision tree classifier to modify the rules in special cases. For example, if a very good P wave is present, no matter how irregular the R-R intervals are, the rhythm should be classified as non-AF. To reduce the number of short false-positive episodes, a hysteresis counter may be used to begin (or end) an episode if only few consecutive sets of parameters have been classified as AF (or non-AF). This counter may be nonsymmetric to make it less sensitive to begin an episode than to end it.

Fig. 2 shows an example of such a counter. The dashed line on the top shows the output of the decision tree classifier. The dot-dashed line in the middle shows the hysteresis counter that counts up with every AF classification and counts down with any non-AF classification. As seen in the bottom of the plot, the algorithm turns the AF alarm on only if the hysteresis counter goes above the predefined threshold of 0 and turns the alarm off when the counter goes below that threshold.

Discussion

We report the performance of the AF detection algorithm with combined R-R interval and atrial activity analysis, which includes a corrector and a hysteresis counter, labeled as Philips algorithm, in Table 4. A comparison of the Philips AF algorithm performance with other previously published

Table 4
Performance statistics in the evaluation database (n = 23) for Philips AF monitoring algorithm

Algorithm		Duration Statistics					Episode Statistics		
		True Positive (minute)	False Positive (minute)	Sensitivity (%)	Specificity (%)	PPV (%)	False Positive (n)	Sensitivity (%)	PPV (%)
Philips algorithm	Gross	5,193	129	93	98	98	11	92	97
	Average			91	96	89		97	94

This algorithm analyzes R-R interval and atrial activity and includes the nonsymmetric hysteresis counter shown in Fig. 2.
PPV indicates positive predictive value.

Table 5
Performance statistics of Philips AF monitoring algorithm compared with previously published algorithm performance results on MIT-BIH AF database

Method	Parameters used for classification	Gross duration sensitivity (%)	Gross duration specificity (%)	Gross duration PPV (%)
Philips algorithm	R-R and P wave	93	98	98
Logan and Healey[13]	R-R	96	89	Not available
Moody and Mark[8]	R-R	93.58	Not available	85.92
Novák[14]	R-R & P wave	93.83	90.12	Not available
Wild[15]	P wave	88.98	87.99	Not available
Young et al[16]	R-R	94.75	Not available	91.38

Philips AF algorithm has the highest specificity (98%) and PPV (98%) in quantifying AF duration.
PPV indicates positive predictive value.

algorithms on the same evaluation database is reported in Table 5. The accuracy in detection of episodes by duration is particularly important for AF burden calculation. Visual validation of duration is labor intensive. Our AF algorithm has outperformed all other previously published results, with 93% sensitivity and 98% specificity and positive predictive value in the quantification of AF duration.

It is noteworthy to point out the limitations of the algorithm. When AF is present and R-R intervals are regular, for example, because of complete atrioventricular block, ventricular pacing, or rate-controlling medication, an AF episode may be misclassified as non-AF. Furthermore, P wave detection is often difficult when extensive noise is present. If a patient has an irregularly irregular but non-AF rhythm with exceptional noise in P wave regions, the rhythm may erroneously be classified as AF.

Conclusions

R-R interval irregularity is the most accessible ECG characteristic for AF detection. Analysis of atrial activity leads to improved performance, particularly in specificity and positive predictive value. This reduction of false positives is mainly achieved by detecting true P wave, which prevents misclassifying non-AF arrhythmias as AF. The 2 parameters derived from atrial activity, the P wave similarity measure and the P-R interval variation, are effective classifiers in an AF detection algorithm.

The approach we have taken for postclassification correction with an added hysteresis counter further enhances the AF alarm behavior and makes the AF algorithm more user-friendly in the clinical environment, whether in the hospital or home.

References

1. Fuster V, Rydén LE, Cannom DS, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation. JACC 2006;48:854.
2. Estes NAM, Halperin JL, Calkins H, et al. ACC/AHA/Physician Consortium 2008 Clinical Performance Measures for Adults With Nonvalvular Atrial Fibrillation or Atrial Flutter: a report of the American College of Cardiology/American Heart Association Task Force on performance measures and the physician consortium for performance improvement. JACC 2008;51:865.
3. Akar JG, Jeske W, Wilber DJ. Acute onset human atrial fibrillation is associated with local cardiac platelet activation and endothelial dysfunction. JACC 2008;51:1790.
4. Nilsson B, Chen X, Pehrson S, et al. Recurrence of pulmonary vein conduction and atrial fibrillation after pulmonary vein isolation for atrial fibrillation: a randomized trial of the ostial versus the extraostial ablation strategy. AHJ 2006;152:537.
5. Janse PA, van Belle YL, Theuns DA, et al. Symptoms versus objective rhythm monitoring in patients with paroxysmal atrial fibrillation undergoing pulmonary vein isolation. Eur J Cardiovasc Nurs 2008;7:147.
6. Ziegler PD, Koehler JL, Mehra R. Comparison of continuous versus intermittent monitoring of atrial arrhythmias. Heart Rhythm 2006;39:1445.
7. Perez FJ, Lung TH, Ellenbogen KA, Wood MA. Is time to first recurrence of atrial fibrillation correlated with atrial fibrillation burden? AJC 2006;97:1343.
8. Moody GB, Mark RG. A new method for detecting atrial fibrillation using R-R intervals. IEEE Comput Cardiol 1983;10:227.
9. Mahalanobis P. On the generalized distance in statistics. Proc Indian Natl Inst Sci 1936;2:49.
10. Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer-Verlag; 2002.
11. Breiman L, Friedman J, Olshen RA, Stone CJ. Classification and regression trees. Wadsworth; 1984.
12. American National Standard ANSI/AAMI EC38:2007. Medical electrical equipment—Part 2-47: particular requirements for the safety, including essential performance, of ambulatory electrocardiographic systems. Association for the Advancement of Medical Instrumentation. Arlington, Virginia.
13. Logan BT, Healey J. Robust detection of atrial fibrillation for a long term telemonitoring system. IEEE Comput Cardiol 2005;32:619.
14. Novák D. Homepage. Project in atrial fibrillation. https://nit.felk.cvut.cz/~xnovakd1/Projects/Atrial/HongYing.pdf [Accessed on 6 July 2009].
15. Wild J. Bachelor thesis. Detection of atrial fibrillation using P-wave template matching. Czech Technical University, Prague; 2006.
16. Young B, Brodnick D, Spaulding R. A comparative study of a hidden Markov model detector for atrial fibrillation. Neural Netw Signal Process IX 1999:468.