Atrial Fibrillation Analysis Based on Blind Source Separation in 12-lead ECG

Pei-Chann Chang, Jui-Chien Hsieh, Jyun-Jie Lin and Feng-Ming Yeh

Abstract—Atrial Fibrillation is the most common sustained rhythmia encountered by clinicians. Because of the invisible aveform of atrial fibrillation in atrial activation for human, it is ecessary to develop an automatic diagnosis system. 12-Lead ECG ow is available in hospital and is appropriate for using Independent Component Analysis to estimate the AA period. In this research, we lso adopt a second-order blind identification approach to transform he sources extracted by ICA to more precise signal and then we use frequency domain algorithm to do the classification. In experiment, we gather a significant result of clinical data.

Keywords—12-Lead ECG, Atrial Fibrillation, Blind Source Separation, Kurtosis

I. Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia encountered by clinicians [1] and occurs less than 1% of those under 60 years of age and more than 6% in those aged over 60 [2]. The interest in the research and understanding of AF has considerably increased during the last years [3]. The analysis and characterization of AF and other atrial tachyarrhythmias from noninvasive techniques requires the previous estimation of the atrial activity (AA) signal from the standard electrocardiogram (ECG). During AF, the electric waveform arising from the atria appear in a disordered way. In normal situation, steady heart rhythm typically beats 60-80 times a minute, but in cases of atrial fibrillation, the rate of atrial impulses can range from 300 to 600 beats per minute. That is, in ECG, AF is often described by those consistent P-waves that are replaced with fast, irregular fibrillatory waves (F-waves) of variable sizes, shapes and timing [4].

Recently, researches indicate that the dynamics of atrial activity play an important role to characterize and classify atrial tachyarrhythmias. To extract the feature, some approaches base on either QRST cancellation through spatiotemporal method [5], average beat subtraction (ABS) [6] or blind source separation (BSS), e.g. independent component analysis (ICA) [7], Hilbert-Huang transform [8] and time frequency analysis approaches, like wavelet decomposition [9]. Regarding the ECG, these researches observed that the AA signal typically exhibits a narrowband spectrum with a main frequency of between 3.5-9 Hz[10]. This is an important information we use to classify the AF from AA period in frequency domain.

Pei-Chann Chang is with Yuan Ze University, Taoyuan, Taiwan. He is now with the Department of Information Management. (e-mail: iepchang@saturn.yzu.edu.tw).

Jui-Chien Hsieh is with Yuan Ze University, Taoyuan, Taiwan. He is now with the Department of Information Management. (e-mail: jchsieh@saturn.yzu.edu.tw).

Jyun-Jie Lin is with Yuan Ze University, Taoyuan, Taiwan. He is now with the Department of Information Management. (corresponding author to provide e-mail: s969203@mail.yzu.edu.tw).

Feng-Ming Yeh is with Yuan Ze University, Taoyuan, Taiwan. He is now with the Department of Information Management. (corresponding author to provide e-mail: yaya10802000@yahoo.com.tw).

II. Methodology

A. 12-Lead ECG data

In this research, we proposed a blind source separation based approach to classify the atrial fibrillation in 12-lead ECG data. The data in this research is collected from Taoyuan Armed Forces General Hospital in Taiwan. All rhythms were diagnosed and verified by a cardiologist, Doctor Yeh. There are totally 98 12-lead ECG data selected and recorded by Philps TC50. The sampling rate is 500Hz with a 10-second record. According to the result diagnosed by the doctor, there are 25 normal rhythms while the number of AF is 73 and these are the fundamental data we used to compare the accuracy.

Clinical 12-lead ECG data is now available in most hospitals and it includes more detailed information about cardiac disease. The standard 12-lead ECG is composed of six horizontal plane leads called the limb leads, and six horizontal leads, which were also called chest leads [11]. Those leads offer 12 different angles for visualizing the activities of the heart and are named lead I, II, III, aVL, aVF, aVR, V1, V2, V3, V4, V5 and V6, respectively. It is worth noting that ECG complex does not look the same in all the leads of the standard 12-lead system and the shape of the ECG constituent waves may vary depending on the lead

Atrial tachyarrhythmias are cardiac arrhythmias in normal atrial electrical activation is substituted by continuous activation with multiple wavelets depolarizing the atria simultaneously [4]. AF is one of the most frequent atrial tachyarrhythmias and is characterized by apparently chaotic atrial activation with the irregular and rapid waves, so called fibrillation. Figure 1 shows an example of normal rhythm and AF ECGs

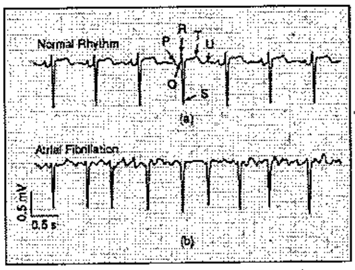

Fig. 1 A Comparison of the Normal Rhythm and AF

In this research, we follow Castells and Rieta's research[7, 12] imarily to process our ECG data but with some different odifications according to our experiment. One of the fferences is this study reconstructs the signals before ecuting Second-Order Blind Identification (SOBI) and nother one is we use the Power Spectral Density (PSD) iagram to classify the ECG data belongs to AF or not. The etail flowchart in this research is shown in figure 2.

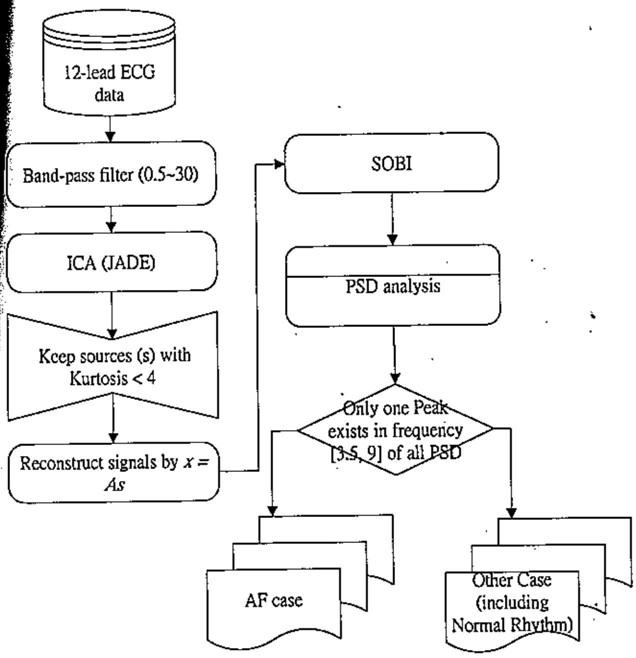

Fig. 2. The overall framework

B. Independent Component Analysis

Independent component analysis is considered because it has precise and successful work in biomedical engineering. In multichannel signal process such as 12-lead ECG data, ICA especially shows the effective utilization to extract AA and ventricular activity (VA) signal [12].

ICA is a statistical method that seeks nonsingular linear transformations of multivariate data so that the transformed variables have the minimal dependence between each other statistically [13].

The basic ICA approach uses the following linear model:

$$x = As$$

where s represents d independent sources and A is a nonsingular $d \times d$ linear mixing matrix. The x means the composed result of d observed signals. The goal of ICA is to recover "unknown" original source. In order to get the result, ICA calculates the inversed matrix, $W = A^{-1}$ and the formula can be rewritten as

$$s = A^{-1}x = Wx$$

General speaking, the key to estimating the ICA model is nongaussianity [14]. A classical measure of nongaussianity is kurtosis. For a signal x, it is classically defined as

$$kurt(x) = E(x^4) - 3\left[E\left(x^2\right)\right]^2$$

the kurtosis value is zero for Gaussian densities and for most nongaussian random distribution, kurtosis is nonzero value.

In ECG data, we can assume that AA and VA are independent and do not present random Gaussian distributions. The assumption makes ICA is possible to reconstruct the AA signals from VA or other, like breath effect and muscular noise in ECG. Before applying ICA on the ECG data, this study uses a low-pass filter to reduce the noise and the results of normal rhythm and AF case by applying low-pass filter and ICA in steps.

C. Kurtosis measurement

ICA can extract all non-Gaussian sources based on the kurtosis measurement. For all the data applied by ICA, the result is resorted by kurtosis value in ascending order. By the experiments, the ventricular sources should show the high kurtosis value while AA is quasi-Gaussian and has smaller kurtosis. This study keeps the source signals with kurtosis value small than four; because of those signals contain more AA information than others.

D. Second-Order Blind Identification (SOBI)

The SOBI algorithm is to separating a mixture of uncorrelated sources with different spectral content through a second-order statistical analysis which also takes into consideration the source temporal information[15]. SOBI seeks a transformation that simultaneously diagonalizes several correlation matrices at different lags. SOBI shows a number of attractive features including (1) SOBI only relies on second-order statistics of the received signals. (2) It allows the separation of Gaussian sources that cannot be separated by ICA. (3) The use of several covariance matrices makes the algorithm more robust and unlikely indeterminacies so that this study adopts SOBI to do the second times transformation after applying ICA.

The goal of separating the AA from other sources of interference is equivalent to finding an orthogonal transformation from the observation whitened by ICA. Because AA has a narrowband spectrum, SOBI algorithm is appropriate for estimating the AA.

E. Power Spectral Density Based Classification

SOBI can separate the mixed uncorrelated sources. An important assumption in this study is that the sources contain vibrated waves in AF cases while normal rhythm should have ılar signals. This assumption leads us to use frequency nain technique to classify the cases are AF or not. The cedure is estimated by using Welch's averaged modified iodogram and the window size is 1000. In classification ge, we only consider the bound of frequency is between 0.5 to Hz because the rest frequencies have a low contribution. The ults of AF and normal rhythm after PSD analyses the nsformed sources from SOBI.

For classification, for all PSD diagrams, if there exists only e peak in one or more PSD diagrams, then the input case will e considered as AF; if there is no peak in any PSD diagrams or PSD diagram has more than one peak in itself, then the case hould be classified as non-AF. The figure 8 to figure 11 shows he PSD and fast Fourier transform (FFT) results of AF and ormal rhythm respectively.

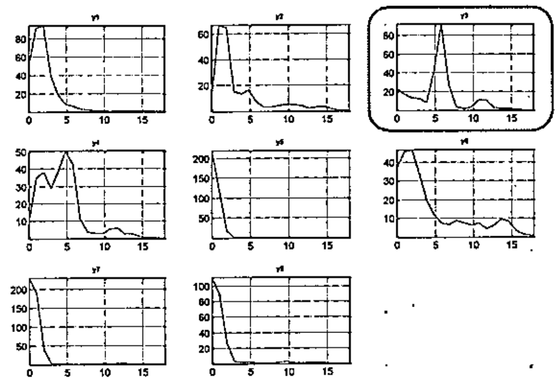

Fig. 3. PSD diagram for AF

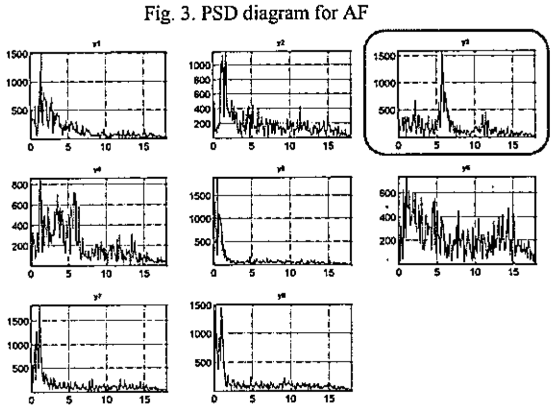

Fig. 4. FFT diagram for AF

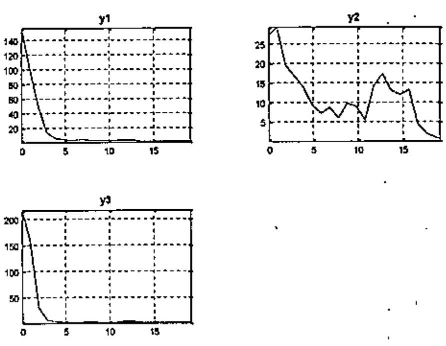

Fig. 5. PSD diagram for Normal Rhythm

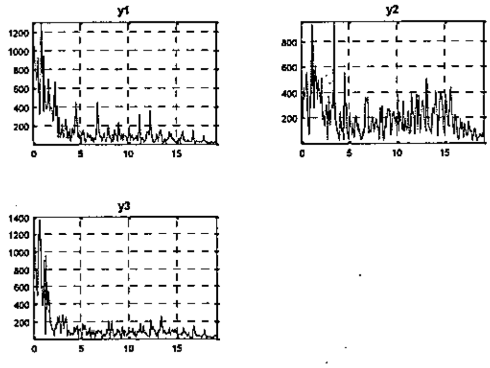

Fig. 6. FFT diagram for Normal Rhythm

In figure 3, because of there exists a diagram with only one peak in frequency 3.5 to 9, hence the case will be regarded as AF, in contrast, the normal rhythm should not have such phenomenon.

III. Experimental Result

According to the literature, it is believed that ICA and SOBI can separate the AF signal in AA period if it exists and if the ECG is normal rhythm, than the PSD should not have any peak information. In the future, we will focus on how to classify AF or not by some novel two-class classifier.

The performance for this experiment is measured by three criterions: accuracy, sensibility (SE), giving the ability of true event detection and specificity (SP), and reflecting the probability of a negative test among patients without disease. These performance measures are defined as following.

$$accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$

$$SE = \frac{TP}{TP+FN}$$

$$SP = \frac{TN}{TN+FP}$$

where TP (True Positive) is the number of matched events and FN (False Negative) is the number of events that are not detected by this approach. FP (False Positive) is the number of events detected by this approach and non-matched to the detector annotations. TN (True Negative) presents as the percentage of events truly identified as not defectives, or normal. In this study, the positive means AF. Table 1 lists the final result.

Table I. The performance of this study

	Sensibility	Specificity	Accuracy
our approach	68.49%	96.00%	75.51%
Philps TC50	91.78%	20.00%	73.47%

IV. CONCLUSION

ccording to the literature, it is believed that ICA and SOBI separate the existed AF signals from AA period and if the G is normal rhythm, the PSD morphology should not have peak information. This study adopts several approaches for ferent situations. For cancelling the noise simply, the nd-pass filter is adopted, for removing VA, a threshold value four (because of the different definition about kurtosis, the lue in Rieta et.al 's research is 1.5, equals to 4.5 in this study) chosen, as shown in figure 4, the waveform with the larger lue of kurtosis seems contains more VA information. Because f the limitation about Gaussian sources separation in higher rder statistics based independent source analysis, the second rder statistics based approach can be applied, hence, SOBI is sed to separate the Gaussian-like signals from AA. According o the literature and our experiment, the PSD diagram can be sed to classify the input case is AF or not because of the peak hall exist in frequency 3.5 to 9 when the AA period contains the ibratory waves.

As shown in Table 1, this study's sensibility is lower than Philips TC's diagnosis result while specificity is much higher than Philips TC50. In the future, we will focus on two directions, one is the problem mentioned above, the possible solution including adopting high-pass filter to reduce the lower frequency's contribution, or applying the third signal separation approaches, like wavelet decomposition or Hilbert-Huang algorithm. These two directions should be increase the accuracy in the future research.

REFERENCES

[1] Y. J. Lin, Y. B. Liu and C. C. Chu, "Incremental changes in QRS duration predict mortality in patients with atrial fibrillation," *PACE - Pacing and Clinical Electrophysiology,* vol. 32, pp. 1388-1394, 2009.

[2] S. A. Hunt, *et al.*, "ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: Executive summary. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure)," *Journal of the American College of Cardiology,* vol. 38, pp. 2101-2113, 2001.

[3] R. Alcaraz and J. J. Rieta, "A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms," *Biomedical Signal Processing and Control,* 2009.

[4] V. Fuster, *et al.*, "ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the European Society of Cardiology Committee for practice guidelines (Writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation)," *Circulation,* vol. 114, pp. e257-e354, 2006.

[5] M. Stridh and L. Sörnmo, "Spatiotemporal QRST cancellation techniques for analysis of atrial fibrillation," *IEEE Transactions on Biomedical Engineering,* vol. 48, pp. 105-111, 2001.

[6] N. V. Thakor and Y. S. Zhu, "Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE Transactions on Biomedical Engineering,* vol. 38, pp. 785-794, 1991.

[7] F. Castells, J. J. Rieta, J. Millet and V. Zarzoso, "Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias," *IEEE Transactions on Biomedical Engineering,* vol. 52, pp. 258-267, 2005.

[8] Z. Huang, Y. Chen and M. Pan, "Time-frequency characterization of atrial fibrillation from surface ECG based on Hilbert-Huang transform," *Journal of Medical Engineering and Technology,* vol. 31, pp. 381-389, 2007.

[9] P. S. Addison, J. N. Watson, G. R. Clegg, P. A. Steen and C. E. Robertson, "Finding coordinated atrial activity during ventricular fibrillation using wavelet decomposition," *IEEE Engineering in Medicine and Biology Magazine,* vol. 21, pp. 58-61+65, 2002.

[10] P. Langley, J. P. Bourke and A. Murray, "Frequency analysis of atrial fibrillation," in *Computers in Cardiology,* 2000, pp. 65-68.

[11] T. B. Garcia and N. E. Holtz, *Introduction to 12-Lead ECG: The Art of Interpretation*: Jones and Bartlett Publishers, 2001.

[12] J. J. Rieta, V. Zarzoso, J. Millet-Roig, R. García-Civera and R. Ruiz-Granell, "Atrial activity extraction based on blind source separation as an alternative to QRST cancellation for atrial fibrillation analysis," 2000, pp. 69-72.

[13] P. Comon, "Independent component analysis, A new concept?," *Signal Processing,* vol. 36, pp. 287-314, 1994.

[14] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks,* vol. 13, pp. 411-430, 2000.

[15] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Transactions on Signal Processing,* vol. 45, pp. 434-444, 1997.