

ATRIAL FIBRILLATION ANALYSIS BASED ON ICA INCLUDING STATISTICAL AND TEMPORAL SOURCE INFORMATION

F.Castells¹, J. Igual², J.J. Rieta¹, C. Sanchez¹, J. Millet¹

1 Bioengineering, Electronics and Telemedicine Research Group, Universidad Politecnica de Valencia
2 Communications Department, Universidad Politecnica de Valencia

ABSTRACT

The analysis and characterization of atrial fibrillation requires the previous extraction of the atrial activity from the electrocardiogram, where the independent atrial and ventricular activities are combined in addition to noise. An independent component analysis method is proposed where additional knowledge about the time and statistical structure of the sources is incorporated. Finally, a combined method based on maximum likelihood and second order blind identification is obtained and validated with results that improve those obtained with traditional ICA algorithms.

1. INTRODUCTION

Characterization of cardiac arrhythmias from electrocardiogram recordings (ECG) has become over the years an important field of research in biomedical engineering. Recently, interest has focused mainly on the study and comprehension of atrial fibrillation (AF) [1] for several reasons: firstly, AF is the most frequent cardiac arrhythmia and has a prevalence of 10% in population over 70. Secondly, existing therapies do not provide satisfactory solutions and, thirdly, the latest advances in biomedical signal processing allow us to develop new techniques for the estimation of atrial activity (AA) from the analysis of 12-lead ECG [2].

Indeed, it has been demonstrated that the problem of AA extraction free from any ventricular activity (VA), which has higher amplitude (i.e. QRS complexes and T waves), admits a blind source separation (BSS) model [3], achieving high performances in the estimation of AA. However, a statistical analysis of AA and VA sources uncovers the statistical behavior of each source, providing interesting prior knowledge that can be used in order to improve source separation.

This paper has two goals: to describe the statistical characteristics of AA and VA sources, and to present a suitable AA separation method that exploits such information. In order to measure performance, synthesized AF ECGs with known AA content have been

created so that the estimated AA can be directly compared to the real AA. The results provided with a prior knowledge technique are compared to those obtained with a generic independent component analysis (ICA) approach with no other assumption than the statistical independence of the sources [4,5].

2. STATISTICAL SOURCE ANALYSIS

In order to design a suitable separation algorithm, a previous statistical analysis of sources is carried out. Actually, the sources contained in an ECG recording can be divided into three classes of a different nature. VA sources are the ECG components with highest energy. These components have high amplitude during ventricular depolarization and repolarization (QRS complex and T wave respectively), but the rest of the time have low values close to zero due to the non-activity period of the ventricular cells. Thus, VA sources are supergaussian random variables. In AF episodes, AA consists of small and continuous wavelets with a cycle around 160ms. Indeed, AA can be modeled by a saw-tooth signal [2], or, if it is considered as a random variable with a pdf described by its histogram, a subgaussian signal. Finally, noise and other artifacts present similar amplitudes than atrial sources, but their behavior is well-modeled by AWGN.

3. METHODS

A two step strategy was designed for the estimation of the AA source. In a first step the goal is to eliminate as much as possible any ventricular component. A maximum likelihood (ML) approach will be employed in order to force supergaussian behavior to VA sources. As a result, 12 sources are obtained where two subspaces can be distinguished, one of them composed by ventricular sources and the other one by any other non-ventricular component, i.e. AA, ECG artifacts and noise. In the second step the VA subspace is not considered, since the goal is to separate the AA from any noise source and all of them are contained in the second subspace. In order to achieve this objective, the spectral information of the AA

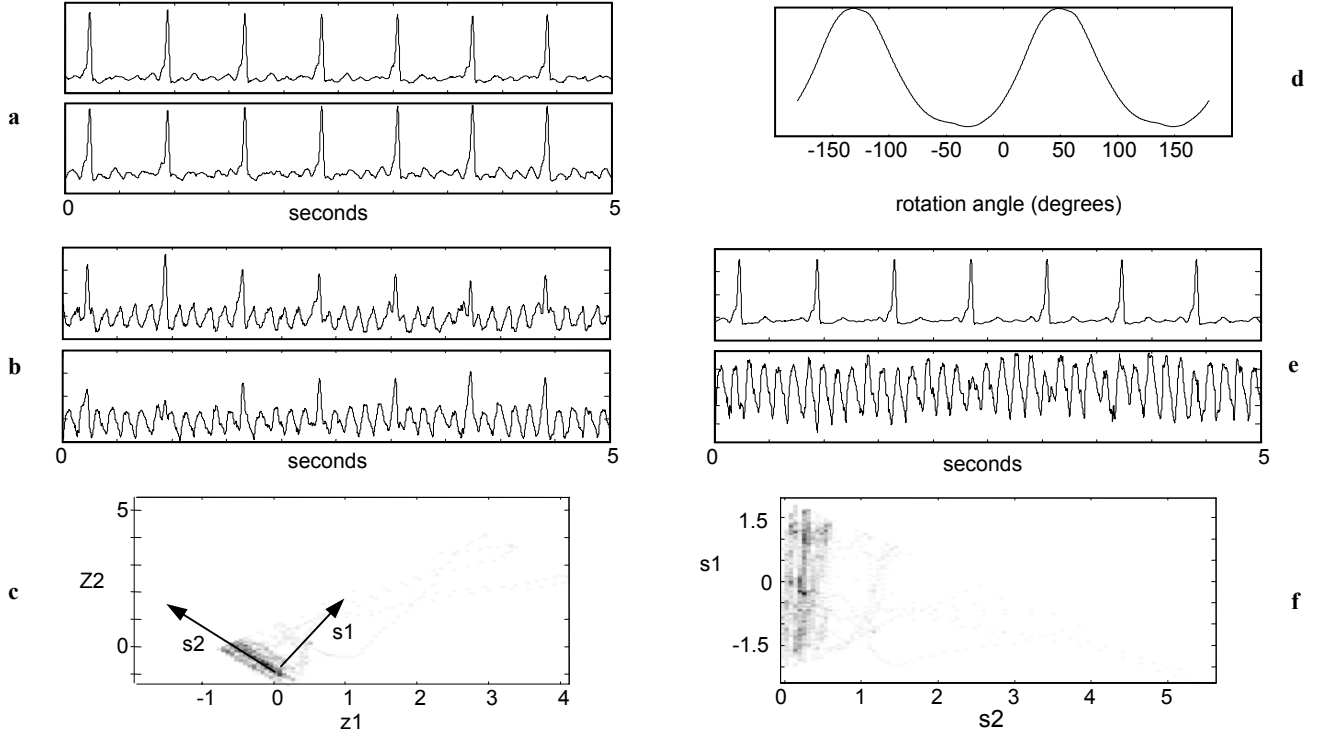


Fig. 1: Illustration of a Maximum Likelihood approach with synthetic data created by a combination of one VA and one AA signals Fig. 1.a: Mixtures. Fig. 1.b: whitened mixtures. Fig. 1.c: scatter-plot of the whitened signals. Fig. 1.d: likelihood function vs rotation angle. Notice the indeterminacy of 180° in the estimation of θ , since the order of the sources is forced by the previous estimations of r_1 and r_2 (ICA without prior knowledge has an indeterminacy of 90°). Fig. 1.e: recovered signals. Fig 1.f: scatter-plot of the recovered signals.

is used. The quasi-periodic behavior of the AA with a specific autocorrelation function allows us to employ a Second Order Blind Identification (SOBI) algorithm which is able to separate independent sources with different spectral content [6].

3.1. Maximum Likelihood

Considering a general approach, the independent sources can be modeled as generalized gaussian random variables [7], whose pdf reads:

$$f(x_i) = \alpha_1(r_i) \cdot \exp\left[-\alpha_2(r_i) \cdot |x_i|^{r_i}\right] \quad (1)$$

where α_1 and α_2 can be expressed in terms of the Gamma function. Particular distributions where $r=1$, $r=2$ and $r=\text{Inf}$ correspond for laplacian, gaussian and uniform pdfs respectively. Notice that this expression corresponds for zero mean, unit variance random variables as a result of a previous whitening processing. Indeed, before applying ML, a preprocessing step based on principal component analysis (PCA) is recommended, since after pre-whitening, the real sources and the whitened vectors are just related through an orthogonal transformation \mathbf{Q} . Assuming a simplified scenario of two sources (one VA

and one AA) such transformation becomes a Givens rotation:

$$\mathbf{Q}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (2)$$

Hence the joint sources pdf becomes

$$\begin{aligned} f(x_1, x_2) &= \\ &= \alpha_1(r_1)\alpha_2(r_2) \cdot \exp\left[-\alpha_2(r_1)|x_1|^{r_1} + \alpha_2(r_2)|x_2|^{r_2}\right] \quad (3) \\ x_1 &= z_1 \cos\theta - z_2 \sin\theta \\ x_2 &= z_1 \sin\theta + z_2 \cos\theta \end{aligned}$$

where the parameters $[r_1, r_2, \theta]$ are estimated by the ML algorithm. Experimental results show that the parameters corresponding to the pdf of the sources are $\hat{r}_1 = 0.5 \pm 0.15$ and $\hat{r}_2 = 4.5 \pm 2$, corresponding to the AA and VA signals respectively, i.e., a subgaussian and supergaussian random variable, confirming our assumption about its statistical characteristics.

3.2. Laplacian and Uniform Approximations

In order to reduce computational complexity the pdf of the sources can be parameterized according to traditional and

well-known pdfs. Hence, the VA and AA pdfs will be approximated by a laplacian and uniform pdf respectively, so that the unique parameter to estimate is the rotation angle $\hat{\theta}$. The rotation angle will be then estimated as the angle that maximizes the log-likelihood function. The joint pdf for independent VA and AA sources is:

$$f(x_{VA}, x_{AA}) = \frac{1}{\sqrt{2}} \exp[-\sqrt{2} \cdot |x_{VA}|] \cdot \frac{1}{2\sqrt{3}} |x_{AA}| < \sqrt{3} \quad (4)$$

Taking logarithms and expressing x_{VA} as a function of θ , the rotation angle can be estimated as:

$$\hat{\theta} = \max_{\theta} \sum_i -|z_1^i \cos \theta - z_2^i \sin \theta| \quad (5)$$

In Fig.1 we can see an example of an AA extraction based on these approximations of the pdfs. The results obtained in many simulations allow us to conclude that the general modeling of the AA and VA is correct, so it can be included in the general ICA approach to the real problem starting from the ECG.

3.3. Second Order Blind Identification

SOBI techniques consist of separating a mixture of independent sources with different spectral content through second order analysis considering also temporal information of the sources. For this purpose, SOBI methods aim to find a transformation that diagonalizes the cross-correlation matrix at several lags simultaneously. Since there may exist no transformation which accomplishes that condition, a function that measures the diagonalization at different lags must be defined in order to maximize the joint diagonalization criteria.

Concerning our specific problem for AA estimation, the number of matrices to jointly diagonalize and the time lags must be properly designed. Since the autocorrelation of the AA source is quasi-periodic with a period around 160ms –i.e. 160 samples due to a sampling frequency of 1Khz– and the autocorrelation of noise has values close to zero for lags distinct than null, cross-correlation matrix with time lags involving two cycles –i.e. 320 samples– are chosen. In order to achieve best results, a total of 17 cross-correlation matrices at equispaced lags of 20 samples will be employed for joint diagonalization.

4. DATABASE: AF RECORDINGS

The fact that the AA is unknown in real AF recordings hinders an in-depth study of the performance. Therefore suitable simulated AF ECGs must be designed in order to evaluate the performance of a ML SOBI approach and to compare other ICA methods. Following this formulation, AF ECGs have been generated just with by adding known AA waves to a normal sinus rhythm (NSR) ECG, being

the amplitude of AA signals 15dB under VA. This process was repeated over a set of 8 AF signals created by combination of known AA and real NSR databases. Fig. 2.a shows an example of simulated AF signals.

4.1. Performance Index

The fact that the AA is known in simulated AF ECGs allows us to carry out an accurate analysis of the performance. Following an ICA formulation, the sources \mathbf{s} and the channel-parameter matrix \mathbf{M} are related such that

$$\begin{aligned} \mathbf{x} &= \mathbf{M} \cdot \mathbf{s} \\ \mathbf{s} &= \mathbf{W} \cdot \mathbf{x} \end{aligned} \quad (6)$$

being $\mathbf{W}=\mathbf{M}^{-1}$ the separation matrix. Since the observations \mathbf{x} are a combination of a NSR ECG \mathbf{x}_{NSR} and simulated AA waves \mathbf{x}_{AA} , the original sources can be decomposed as it follows

$$\mathbf{s} = \mathbf{W} \cdot \mathbf{x}_{NSR} + \mathbf{W} \cdot \mathbf{x}_{AA} \quad (7)$$

This is to say that the source i is recovered from a linear combination of the leads given by the i -th row coefficients of the \mathbf{W} matrix. Our interest is focused on the row corresponding to the extraction of the AA signal, \mathbf{w}_{AA} ($1 \times q$ dimensional), yielding to a linear combination which cancels the contribution of the QRS complexes whereas maximizes the contribution of the AA:

$$\begin{aligned} \hat{s}_{AA} &= \mathbf{w}_{AA} \cdot \mathbf{x} \quad , \quad s_{AA} = \mathbf{w}_{AA} \cdot \mathbf{x}_{AA} \\ I &= \mathbf{w}_{AA} \cdot \mathbf{x}_{NSR} = \hat{s}_{AA} - s_{AA} \end{aligned} \quad (8)$$

where s_{AA} and \hat{s}_{AA} are the real and estimated atrial activity sources respectively and I is the interference caused by the VA, mainly by QRS complexes. Thus the SIR can be computed as

$$SIR = \frac{E[s_{AA}^2]}{E[(\hat{s}_{AA} - s_{AA})^2]} \quad (9)$$

Other interesting performance indicator parameters are also the cross correlation coefficients of s_{AA} and \hat{s}_{AA} .

5. RESULTS

ML and SOBI were applied over a set of 8 simulated AF recordings with known AA content. In all cases, one AA source was identified among the whole set of 12 separated sources. Performance evaluation was then computed as it has been explained above. Results have been compared with those obtained with other ICA methods. Several approaches included in the ICALAB toolbox [8] have been applied (JADE, Fixed-point, etc.), obtaining equivalent solutions. Table 1 indicates mean values and standard deviation of SIR and correlation indexes. An improvement of 2dB in the SIR of the estimated AA via ML SOBI versus traditional ICA methods is appreciated.

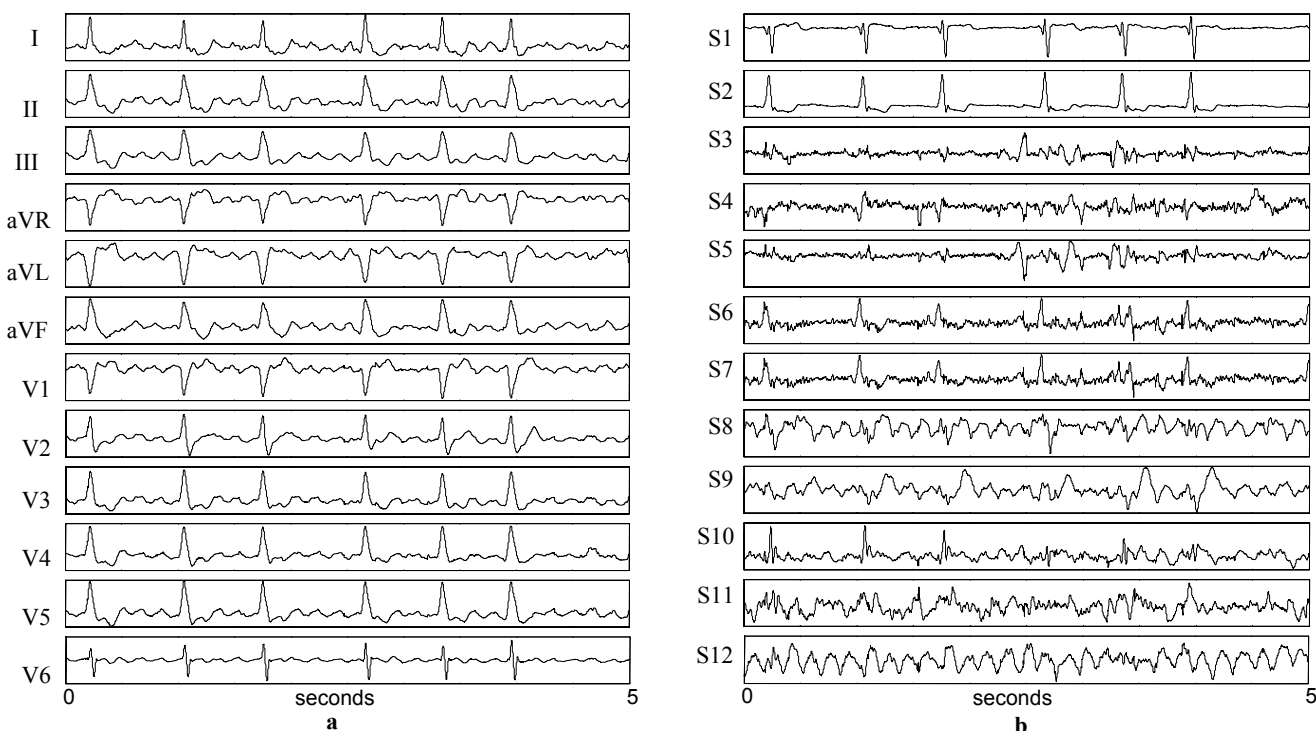


Fig. 2: a) Simulated AF ECG from combination of a NSR ECG and AA waves. b) Sources obtained after ML and SOBI processing. Sources S1 and S2 correspond to VA, whereas AA can be identified in S12

Figure 2.b shows an example of VA-AA separation where AA appears in source 12 free from QRST waves.

Table 1. SIR and Correlation Measurements

	SIR (dB)	Correlation
ICA	10.57 ± 1.56	0.959 ± 0.018
ML SOBI	12.46 ± 1.46	0.971 ± 0.010

6. CONCLUSION

BSS-ICA is a novel technique for AA extraction in ECGs with AF episodes. Due to a prior knowledge of the sources, an ICA approach that combines Maximum Likelihood and Second Order Blind Identification is proposed, so the problem is no longer blind. As long as the imposed conditions are fulfilled by the real data, this method provides better performance than traditional ICA techniques because the new information is included in the algorithm. The results obtained in this study show the suitability of the described model for the AA estimation from 12-lead standard ECGs.

The contribution here presented offers a solution to an important step in AF analysis. The AA extraction allows us to carry out new challenges as AF characterization, pattern recognition, time-frequency parameter extraction, etc. These advances in the clinical knowledge of AF will provide new and customized AF treatments.

ACKNOWLEDGEMENT

This work has been supported by BSS-CEDIC ‘Programa Incentivo a la Investigación UPV’, TIC2002-00957 and Motorola Foundation.

REFERENCES

- [1] M.Allessie, K.Konings, M.Wijffels. Atrial Arrhythmias – State of the Art: Electrophysiological Mechanism of Atrial Fibrillation, *J.P. DiMarco and E.N. Prystowsky, Eds. Armonk, NY: Futura Publ. Co., 1995.*
- [2] M. Stridh, L. Sörnmo. “Spatiotemporal QRST Cancellation Techniques for Analysis of Atrial Fibrillation”, *IEEE Transactions on Biomedical Engineering*, Vol. 48, No. 1, January 2001.
- [3] 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", in: *IEEE Computers in Cardiology*, Vol. 27, Boston, MA, September 24-27, 2000, pp. 69-72.
- [4] J.F.Cardoso, A.Souloumiac. Blind beamforming for non Gaussian signals. *IEE Proceedings-F*, 140(6):362-370, 1993.
- [5] A.Hyvärinen, J.Karhunen, E.Oja. Independent Component Analysis. *John Willey & Sons, Inc., Ed. 2001.*
- [6] 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, No. 2 pp. 434-444, Feb.1997
- [7] R. Everson, S. Roberts. Blind source separation for non-stationary sources. *Kluwer Academic Publisher, Boston.*
- [8] A. Cichocki, S. Amari, K. Siwek et al., ICALAB Toolboxes, 2002 <http://www.bsp.brain.riken.go.jp/ICALAB>