New Methods for Atrial Activity Extraction in Atrial Tachyarrhythmias

Raúl Llinares and Jorge Igual
Universidad Politécnica de Valencia
Spain

1. Introduction

Atrial fibrillation (AF) is the most common human arrhythmia. The analysis of the associated atrial activity (AA) provides features of clinical relevance. In particular, the fibrillatory rate has primary importance in AF spontaneous behavior (Asano et al., 1992), response to therapy (Stambler et al., 1997) or cadioversion (Manios et al., 2000). Previously, the atrial signal must be obtained. Hence, the identification of the AA is necessary. AA is embedded into the surface electrocardiogram (ECG), including the rest of the signals, such as the ventricular rhythm, breathing or noise. The cancellation of the ventricular signal can be done in different ways: from template matching subtraction to source separation approaches. One of these AA extraction methods is implemented on a commercially ECG system available on the market. It features ECG recording and signal processing for non-invasive assessment of atrial fibrillatory activity (Grubitzsch et al., 2008).
This chapter presents further developed algorithms to extract the AA in the frequency domain. The methods presented are extensions of classical ICA methods based on second and higher order statistics. The algorithms exploit the prior assumption about the sources in order to obtain the source extraction algorithms that are focused on the extraction of the atrial component.
The chapter is organized as follows. Section 2 describes the AF. Section 3 reviews the methods of AA extraction existent in the literature. Section 4 outlines the new methods in detail. In Section 5, the quality parameters of AA extraction are discussed. Numerical simulations and comparisons are provided in Section 6. Finally, concluding remarks are given in Section 7.

2. Atrial Fibrillation

The importance of atrial tachyarrhythmias in humans, such as atrial fibrillation (AF) or atrial flutter (AFL), is revealed by statistics: AF affects 0.4% of the general population, but the probability of developing it rises with age, less than 1% for people under 60 years of age and greater than 6% in those over 80 years (Fuster et al., 2001).
During AF, the atria beat chaotically and irregularly, out of coordination with the ventricles, increasing the risk of stroke and death. There is no unique theory about the mechanisms of

AF, but some characteristics of AF in the ECG are well established: the atrial activity is irregular in timing and shape; there is a substitution of the P-waves by an oscillating baseline that consists of low amplitude fibrillatory F-waves (Petrutiu et al., 2006). The shape, amplitude and frequency of the F-waves depend obviously on the patient, being more regular in the AFL case. Atrial rates are usually in the range 240-540 waves per minute in AF and 240-320 in the case of AFL (Stridh et al., 2006). In addition, the ventricular response during AF episodes becomes irregular, with higher average rate (shorter RR intervals). Figure 1 shows an episode of AF (a) and a normal sinus rhythm (b). Note the fibrillatory waves between the R peaks, the absence of the P waves and the irregular ventricular rhythm in (a). The normal sinus rhythm (b) contains the P waves and it is more regular.

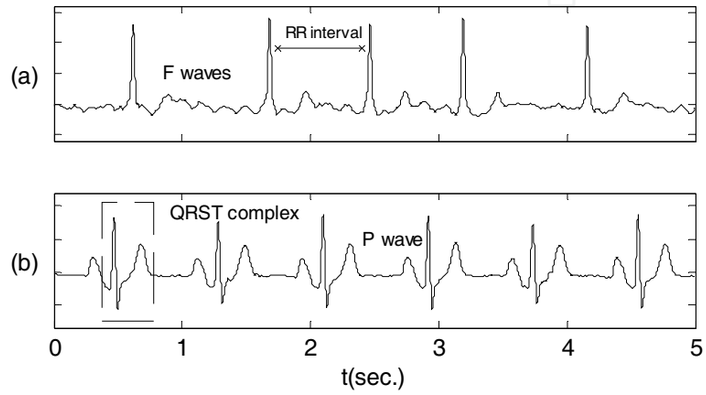

Fig. 1. Comparison between Atrial Fibrillation (a) and Normal Sinus Rhythm (b)

From the signal processing point of view, AA shows a power spectral density concentrated around a main peak in a frequency band (narrowband signal), with slight variations depending on the authors; for example, 4-9 Hz. (Petrutiu et al. 2006, Stridh et al. 2006), 5-10 Hz. (Langley et al. 2000) or 3.5-9 Hz. (Castells et al., 2005a). This spectrum is a key feature to distinguish between AF and other non fibrillatory rhythms. Figure 2 represents two examples of atrial activity: AF in (a) and AFL in (c); (b) and (d) plot their spectra and the corresponding peak frequencies.

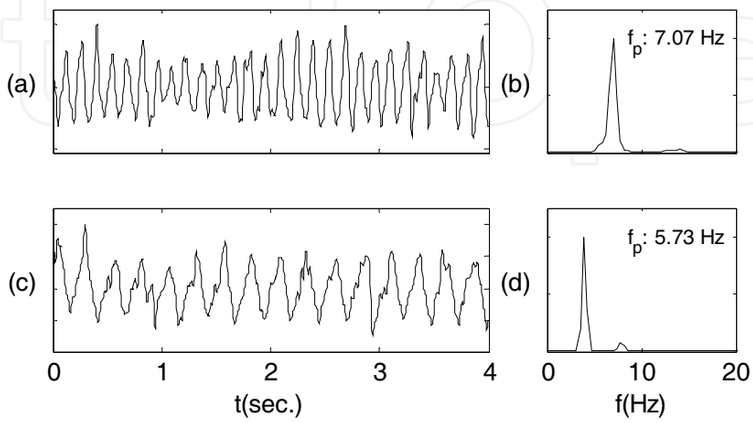

Fig. 2. Two examples of AA, their spectra and the corresponding peak frequencies

Analyzing the AA from a statistical point of view, the AA shows a Gaussian or subgaussian distribution (zero or negative kurtosis value, depending on the patient and the stage of the disease). From a time series point of view, it can be modeled as a sawtooth signal consisting of a sinusoid with several harmonics (Stridh & Sornmo 2001). In this case, the kurtosis values are close to zero. In Figure 3 the histograms of the atrial activities of Figure 2 are represented together with their associated kurtosis values. The continuous solid lines on the plots represent the closest Gaussian approximations to the observed distributions.

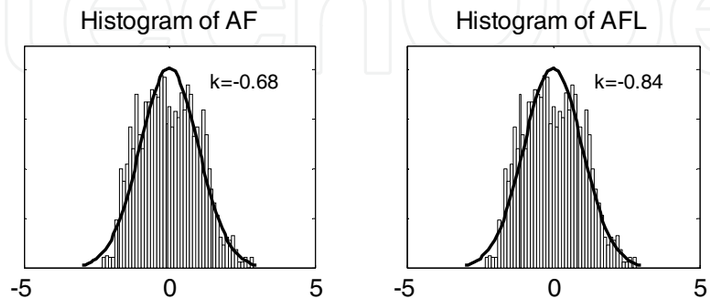

Fig. 3. Histogram of the atrial activities in Fig. 2

3. Methods of Atrial Activity Extraction in the Time Domain

3.1 Strategies

Two main approaches are used to extract the AA: techniques based on exploiting the spatial diversity of multi-lead ECG recordings and single-lead techniques. The first group includes algorithms based on averaged beat subtraction (ABS) and algorithms based on source separation performing an independent component analysis (ICA) or a principal component analysis (PCA) of the recorded ECG. Similar results were obtained when these algorithms were compared (Langley et al., 2006). The single-lead methods to extract the AA are mainly based on averaged beat subtraction. Single-lead algorithms are not benefited by the information present in all the leads. However, single-lead algorithms permit the analysis of early stages of AF with Holter systems where there is no more than two or three available leads, that are not sufficient to exploit the spatial diversity.

3.2 Averaged Beat Subtraction

ABS methods are the most widespread techiques of atrial signal extraction. These methods are based on QRST cancellation in the time domain. They assume two premises: (i) the atrial and ventricular activity are decoupled during AF episodes, and (ii), each individual beat can be represented approximately by an average (template) beat. Once the template is created, it is used to subtract the ventricular activity (VA) from each individual beat, obtaining a remainder or residual ECG containing only the F-waves.

ABS methods can be applied directly to single leads (Slocum et al., 1992, Shkurovich et al., 1998). They differ in the clustering of the different beat morphologies and the estimation of the template. One key point in ABS methods is the time alignment of the average beat and the QRST complex before the subtraction. The alignment can be carried out directly from the R wave timings or maximizing the cross-correlation between the template and the processed beat for different time shifts.

Other ABS methods work in a multi-lead ECG environment. Spatiotemporal QRST cancellation (Stridh & Sornmo, 2001) takes advantage of the spatial diversity to compensate for variations in the electrical axis, variations in the tissue conductivity and heart position. In this case, the ventricular activity is modeled by:

$$\mathbf{X} = \mathbf{J}_\tau \bar{\mathbf{X}} \mathbf{D} \mathbf{Q} \tag{1}$$

where $\mathbf{J}_\tau = \left[\mathbf{0}_{Nx(\Delta+\tau)} \quad \mathbf{I}_{NxN} \quad \mathbf{0}_{Nx(\Delta-\tau)} \right]$ is the shift matrix used for time alignment (with τ an integer time shift and $\pm\Delta$ the maximum corrected alignment error), $\bar{\mathbf{X}}$ is the average beat, $\mathbf{D}$ is a diagonal amplitude scaling matrix and $\mathbf{Q}$ is a rotation matrix.

The optimization of the parameters $\mathbf{D}$, $\mathbf{Q}$ and $\mathbf{J}_\tau$ is solved by means of a minimization problem:

$$\varepsilon^2_{\min} = \min_{\mathbf{D},\mathbf{Q},\tau} \left\| \mathbf{Y} - \mathbf{J}_\tau \bar{\mathbf{X}} \mathbf{D} \mathbf{Q} \right\|^2_F \tag{2}$$

where $\mathbf{Y} = \mathbf{X} - \tilde{\mathbf{X}}_A$ perfoms the AF reduction step with $\mathbf{X}$ the beat being processed and $\tilde{\mathbf{X}}_A$ the TQ-based fibrillation signal (Stridh & Sornmo, 2001). Note that when $\mathbf{Y} = \mathbf{X}$ (no AF reduction step) and $\mathbf{DQ} = \mathbf{I}$, the algorithm corresponds to a traditional ABS algorithm applied to a single-lead.

Lemay et al. (2007) propose a method that processes the QRS complexes separated from the T waves basing on the different nature of the repolarization and depolarization waves.

The main problem of these methods is the reduction of the performance when a high quality QRST cancellation template is difficult to obtain. This is the case of clinical practice where there is only available no more than 10 seconds (Lemay et al. 2007). Other limitations are their high sensitiveness to variations in QRST morphology or the difficulty of finding the optimal selection of the complexes to generate the template (Alcaraz & Rieta, 2008).

3.3 Independent Component Analysis

ICA is a signal processing tool for estimating individual source components from mixtures of them recorded at sensors. The estimation is carried out only with the statistical independence of the sources as assumption. The most basic formulation of ICA is the linear noiseless instantaneous mixture model for real-valued sources and mixtures:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{3}$$

where $\mathbf{x}$ (*Mx1*) is the observed vector that is a linear transformation (mixing matrix $\mathbf{A}$ (*MxK*)) of a source vector $\mathbf{s}$ (*Kx1*) whose components are statistically independent, i.e., the joint probability is the product of the marginal densities $p(\mathbf{s}) = \prod_i p(s_i)$.

The source distributions are not available, so the independence condition cannot be enforced and a measure of the independence is required. ICA algorithms differ in the way that they approximate the independence condition (Hyvarinen et al., 2001).

ICA based on higher order statistics (HOS) has been applied to AA extraction problem (Rieta et al., 2004). The approach satisfies the basic conditions of ICA: independence of the sources, non-gaussianity and generation of observations by instantaneous linear mixing of the sources. The identification of the AA among the set of separated sources was carried out using a kurtosis-based reordering of the separated signals followed by spectral analysis of the subgaussian sources.

Taking advantage of the time structure of the ECG recordings, ICA based on second order statistics (SOS) has also been applied successfully to this problem (Llinares et al., 2006). The sources in the ECG have different spectra allowing the application of SOS-based algorithms. In this case, the identification of the atrial activity was carried out using spectral analysis and kurtosis values.

Castells et al. (2005) proposed a two-step solution based on HOS-ICA (first stage) and SOS-ICA (second stage) to extract the AA. The identification of the atrial activity was done in frequency domain searching for the source with a peak in the range of 3-10 Hz.

Regarding the approach, ICA is a multi-lead technique that suffers a decrement of the quality of the extraction as the number of leads is reduced. Recent studies focus on the optimization of the location of the leads to apply blind source separation (BSS) techniques with a reduced number of leads (Igual et al., 2006). Starting from 64-leads recordings (body surface potential mapping), the AA was extracted using only two leads and an ICA technique.

The methods based on source separation waste computational load in the separation of non interesting sources. In addition, they need an additional step to choose the atrial signal among the recovered sources.

3.4 Principal Components Analysis

PCA is a statistical technique that applies a linear transformation $\mathbf{V}$ (MxK) to the observation data $\mathbf{x}$ ($Mx1$) obtaining a vector of uncorrelated variables $\mathbf{z}$ ($Kx1$):

$$\mathbf{z} = \mathbf{V}\mathbf{x} \tag{4}$$

where the elements of $\mathbf{z}$ are called the principal components. The first principal components will retain most of the variation present in all of the original variables.

PCA has been applied to multi-lead ECG for extracting the AA (Raine et al., 2004). The first components are related to ventricular activity and its variability since this activity presents the largest energy. Among the rest of principal components, the AA is identified in the frequency domain since it exhibits a narrowband spectrum. PCA applied to multi-lead ECG provides the optimal solution for orthogonal mixtures. However, the mixing matrix may have an arbitrary structure, obtaining non-satisfactory results. In addition, if the signals are Gaussian, decorrelation means independence, so in this case PCA and ICA are the same transformation.

PCA has also been applied to single-lead ECG for extracting the AA successfully exploiting the interbeat redundancy. When PCA is applied to several consecutive beats from the same lead, it outputs the principal components and their projections on each beat. The first principal component is related to the main QRST waveform, several components are related to AA and the rest of the components correspond to noise. In the case of several QRST morphologies in the lead or non-regular QRST waveforms, other principal components will

appear representing the different patterns or the dynamics of the QRST waveforms. Castells et al. (2005) estimate the AA reconstructing the atrial subspace from the projections of the non-ventricular components:

$$\mathbf{s}_{AA} = \mathbf{V}_{AA}\mathbf{z} \tag{5}$$

Other application of PCA to extract the AA in single-lead ECGs is presented by Alcaraz & Rieta (2008). They calculate the singular value decomposition of a matrix containing *N* QRST complexes of the analyzed lead with *L* samples each one. They use the first non-normalized principal component as a QRST template to cancel out VA with a previous amplitude adaptation between each QRST complex and the template. The method also avoids sudden transitions at the beginning or the end of each QRST segment provoked by the subtraction of the template to each individual complex.

4. Atrial Fibrillation Extraction in the Frequency Domain

Considering the aforementioned approaches to extract the atrial activity, the specific characteristics of the target signal, i.e., the F-waves, are not included in the separation process. In fact, the spectrum of the F-waves that characterizes this signal is a key feature to distinguish between AF and other non fibrillatory rhythms. Llinares & Igual (2009) presented two algorithms that exploits this discriminative frequency information of the atrial rhythm, focusing only on the extraction of the atrial component in a more effective and simpler manner.

The blind source extraction of the atrial component $y_A(t)$ is:

$$y_A(t) = \mathbf{b}^T\mathbf{x}(t) \tag{6}$$

The recorded signals are first whitened by PCA in order to reduce the dimensions of the problem, to exhaust the use of the second order statistics and to assure that all the sources have the same variance: $\mathbf{z}(t) = \mathbf{V}\mathbf{x}(t)$, where $\mathbf{V}$ is the *NxM* whitening matrix, so $E\{\mathbf{z}(t)\mathbf{z}(t)^T\} = \mathbf{I}$. The new problem reads:

$$y_A(t) = \mathbf{w}^T\mathbf{z}(t) \tag{7}$$

with the restriction $\mathbf{w}^T\mathbf{w} = 1$ (to assure unit power signals). The aim is to estimate $y_A(t)$, i.e., the recovering unit norm vector $\mathbf{w}$, maximizing the independence using HOS-based or SOS-based ICA algorithms and at the same time looking for a signal that is as close as possible to the known spectral properties of the atrial signal. This implies the addition to the independence criterion (HOS or SOS) of a new term that models prior information. The new combined cost function to be maximized can be expressed as $J(\mathbf{w}) = J_1(\mathbf{w}) + \lambda J_2(\mathbf{w})$. The first term, $J_1(\mathbf{w})$, is the ICA only contrast function; it is maximum when a (independent) source is extracted. The second term, $J_2(\mathbf{w})$, models the prior information about the spectral content of the atrial signal. It is maximum when the extracted signal corresponds to the F-

waves that define the atrial signal. The tradeoff between the ICA term and the prior is controlled by the parameter λ that is experimentally adjusted.

The mathematical formulation of the power spectral density information about the atrial waveform corresponds to the integral of the power spectrum in the proper range of frequencies:

$$J_2(\mathbf{w}) = \int_{f_1}^{f_2} \phi_y(f) df \tag{8}$$

where $\phi_y(f)$ is the power spectrum of the recovered signal $y(t)$. It is maximum for the atrial activity $y_A(t)$; the interval of integration $[f_1, f_2]$ depends on prior information about the patient and the criterion to fix the bandwidth. This interval is calculated from a central frequency f_c (maximum of the power spectral density function) and a bandwidth $2B$: $f_1 = f_c - B$ and $f_2 = f_c + B$.

The estimation of the central frequency must be first addressed. Llinares & Igual (2009) propose a simple procedure for the estimation of the peak frequency. First, starting from leads V1 or II (they have the largest ratio of atrial to ventricular signal amplitude (Petrutiu et al., 2006)), the time intervals where there is no QRST complex are extracted, i.e., the signal corresponding only to the T-Q intervals is obtained; second, these T-Q segments are joined smoothly; third, the spectrum of the new signal composed of only T-Q fragments is calculated; fourth, the peak frequency is obtained as the maximum of the FFT in the interval of frequencies 3-9 Hz. Another option for estimating the initial peak frequency is to follow the procedure described by Sassi et al. (2007). From leads V1 or II, the time intervals of the QRST complex (QT intervals) are eliminated and a spectral analysis of the rest of the signal (T-Q interval) is made with a method appropriate for unevenly sampled data, the Lomb-Scargle periodogram (Flannery et al., 1988). Although the automated measurement of the QT interval in atrial fibrillation episodes is complex (Pai & Rawles, 1989), satisfactory results were obtained applying a simple rule based on Bazzet formula (Bazzet, 1997) and fixing $QT_c = 550$ ms (Sassi et al. 2007).

4.1 Atrial Activity Extraction based on HOS

The HOS-ICA cost function $J_1(\mathbf{w})$ chosen is the approximation of the negentropy given by Hyvarinen (1999):

$$J_{HOS}(u) \alpha \left(E[g(u)] - E[g(u_G)]\right)^2 \tag{9}$$

where E is the expectation operator, u is a random variable, u_G is a Gaussian with the same variance of u and g is the non linear function $g(u) = \frac{1}{a_1} \log\cosh(a_1 u)$ with $1 \le a_1 \le 2$.

This non-linear function is a good general-purpose contrast function that avoids possible problems of outliers associated with the kurtosis based contrast function. The final formulation of the optimization problem is:

$$\arg\max_{\mathbf{w}}\left(\left(E[g(y)]-E[g(y_G)]\right)^2+\lambda\int_{f_1}^{f_2}\phi_y(f)df\right)\quad subject\ to\ \mathbf{w}^T\mathbf{w}=1 \tag{10}$$

where the spectrum $\phi_y(f)$ is estimated by means of the periodogram.

The final updating rule consists of two terms. The first one corresponds to the independence criterion, i.e., the negentropy term (Hyvarinen, 1999). The second one implements the optimization of the cost function corresponding to the prior information. See (Llinares & Igual, 2009) for details in the optimization process:

$$\begin{aligned}&\mathbf{w}\leftarrow E\left[\mathbf{z}\tanh\left(a_1\mathbf{w}^T\mathbf{z}\right)\right]-E\left[a_1\left(1-\tanh^2\left(a_1\mathbf{w}^T\mathbf{z}\right)\right)\right]\mathbf{w}+\lambda(\mathbf{C}\mathbf{w}+\mathbf{C}^T\mathbf{w}),\quad \mathbf{C}=\frac{1}{N}\sum_{i=I_1}^{I_2}\mathbf{Z}\boldsymbol{\omega}_i\boldsymbol{\omega}_i^{*T}\mathbf{Z}^T\\&\mathbf{w}\leftarrow\mathbf{w}/\|\mathbf{w}\|\end{aligned} \tag{11}$$

where $[I_1,I_2]$ is the interval of digital frequencies corresponding to $[f_1,f_2]$ and $\mathbf{Z}=[\mathbf{z}(0),\mathbf{z}(1),\ldots,\mathbf{z}(N-1)]$ is the matrix of whitened observations. The second step of the algorithm enforces that $\mathbf{w}$ is a unit norm vector, a necessary condition in order to maintain the unit variance of the recovered source.

4.2 Atrial Activity Extraction based on SOS

The SOS-ICA cost function $J_1(\mathbf{w})$ chosen consists in a "sequential diagonalization" of the autocovariance matrices (Li et al., 2007). The cost function to be maximized is:

$$J_{SOS}\left(\mathbf{w},\mathbf{t},d_0,d_1,\ldots,d_Q\right)=-\sum_{\tau=0}^{Q}\left\|\mathbf{R}_\tau\mathbf{w}-d_\tau\mathbf{t}\right\|^2 \tag{12}$$

where $R_\tau=E\left[\mathbf{z}(t)\mathbf{z}^T(t-\tau)\right],\tau=0,1,\ldots,Q$ are $Q+1$ autocorrelation matrices of the whitened mixtures, $d_0,d_1,\ldots,d_Q$ are $Q+1$ unknown scalars, and $\|\cdot\|$ denotes the Euclidean length of vectors. In order to avoid the trivial solution, the constraints $\|\mathbf{t}\|=1$ and $\left\|\left[d_0,d_1,\ldots,d_Q\right]\right\|=1$ are imposed. One source is perfectly extracted when $\mathbf{R}_\tau\mathbf{w}=d_\tau\mathbf{t}$, since $\mathbf{t}$ is collinear with one column vector in $\mathbf{A}$, and $\mathbf{w}$ is orthogonal to the other $K-1$ column vectors in $\mathbf{A}$. The final formulation of the optimization problem is:

$$\arg\max_{\mathbf{w}}\left(-\sum_{\tau=0}^{Q}\left\|\mathbf{R}_\tau\mathbf{w}-d_\tau\mathbf{t}\right\|^2+\lambda\int_{f_1}^{f_2}\phi_y(f)df\right),\quad \text{subject to}\ \mathbf{w}^T\mathbf{w}=1 \tag{13}$$

The final updating rule consists of three steps. The first one adjusts the independence criterion together with the implementation of the prior information. The second one adjusts the value of $\mathbf{d}$ and the third one adjusts the value of $\mathbf{t}$. The final algorithm reads (see (Llinares & Igual, 2009) for details in the optimization process):

$$
\begin{aligned}
&(i)\quad \mathbf{w} \leftarrow \mathbf{T}\left(\sum_{\tau=0}^{Q} d_\tau \mathbf{R}_\tau\right)\mathbf{t} + \lambda\left(\mathbf{C}\mathbf{w} + \mathbf{C}^T\mathbf{w}\right),\quad \mathbf{T} = \left[\sum\nolimits_{\tau=0}^{Q} \mathbf{R}_\tau^2\right]^{-1};\ \mathbf{C} = \frac{1}{N}\sum_{i=I_1}^{I_2} \mathbf{Z}\boldsymbol{\omega}_i \boldsymbol{\omega}_i^{*T}\mathbf{Z}^T \\
&\qquad \mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\| \\
&(ii)\quad \mathbf{d} = \left[\left(\mathbf{R}_0\mathbf{w}\right)^T\mathbf{t}, \left(\mathbf{R}_1\mathbf{w}\right)^T\mathbf{t}, \ldots, \left(\mathbf{R}_Q\mathbf{w}\right)^T\mathbf{t}\right]^T \\
&\qquad \left[d_0, d_1, \ldots, d_Q\right] \leftarrow \mathbf{d} / \|\mathbf{d}\| \\
&(iii)\quad \mathbf{v} = \sum\nolimits_{\tau=0}^{Q} d_\tau \mathbf{R}_\tau \mathbf{w} \\
&\qquad \mathbf{t} \leftarrow \mathbf{v} / \|\mathbf{v}\|
\end{aligned}
\tag{14}
$$

5. Quality Parameters in Atrial Activity Extraction

To quantify the quality of the extraction with real ECG recordings is not a simple task. In fact, the problem is unsolvable: the true atrial signal is an unobserved random process that is only measured mixed with not only the QRST complexes, but other biologic and non biologic signals, such as 60 (50) Hz powerline interference, instrumentation noise, baseline drift and wander or electromyogram. It makes impossible to compare the residual ECG with the real one, so it is not possible to obtain the typical measures of error in signal processing such as the mean squared error (MSE) or correlation coefficients. As a consequence, the quality of the extraction (bad-good) cannot be quantified with a distance measure from the real solution. This section reviews some of the most extended measures found in the literature and proposes two new quality indexes.

5.1 Parameters Defined in the Frequency Domain

Two parameters in the AA extraction are defined in the frequency domain: the peak frequency and the spectral concentration around the peak frequency.

The peak frequency is defined as the frequency at which the power spectral density is the highest in the 3 - 10 Hz band. It is not really a quality measurement by itself, but a parameter to characterize the AA. The dispersion of its value when we apply the different algorithms can be used to assure the goodness of the extraction.

The Spectral Concentration (SC) is defined as (Castells et al., 2005a):

$$
SC = \frac{\int_{0.82 f_p}^{1.17 f_p} P_A(f)\,df}{\int_0^{\infty} P_A(f)\,df} \tag{15}
$$

where $P_A(f)$ is the power spectrum of the extracted atrial signal and f_p the main peak frequency. SC is a measure of the relative power contained in the narrow band around the peak frequency. This means that when we compare different algorithms applied to the same patient, a higher SC value indicates a better extraction of the atrial F-waves.

We propose another index that explains the kind of noise present in the estimated signal. For every patient, the true atrial signal has an unknown peak frequency and SC value. This

is equivalent to say that, assuming the bandwidth of the atrial signal from 3 to 10 Hz, every patient has a centroid frequency, defined such as:

$$f_c = \frac{\int_3^{10} f \cdot S_{y_A}(f)df}{\int_3^{10} S_{y_A}(f)df} \tag{16}$$

The centroid frequency measures the form of the distribution of the spectrum over the bandwidth of the atrial signal. If the SC is very large, it should be close to the peak frequency. If not, the centroid frequency will move in the frequency direction where there is more power. Remember that the spectrum of the ventricular activity includes frequencies above the 10 Hz and below 3 Hz. With respect to other non atrial signals, their spectrum is also known in advance, e.g., the high frequency powerline interference or the low frequency breathing. We propose an index that measures the centroid frequency displacement when we calculate the centroid frequency of the estimated atrial rhythm considering the atrial bandwidth or the full range of frequencies. It will give us an idea about the kind of noise remaining:

$$\hat{I}_l = \frac{\int_3^{10} f \cdot S_{y_A}(f)df}{\int_0^{10} f \cdot S_{y_A}(f)df} \tag{17}$$

$$\hat{I}_h = \frac{\int_3^{10} f \cdot S_{y_A}(f)df}{\int_3^{\infty} f \cdot S_{y_A}(f)df} \tag{18}$$

In the case of perfect atrial estimation, both indexes equals to one. If not, the indexes are lower than one. A value close to one means a better estimation. In addition, it gives us information about the kind of noise. For example, $\hat{I}_h$ will be large for an algorithm that does not cancel the powerline interference.

5.2 Parameters Defined in the Time Domain

Four parameters in the AA extraction are defined in the time domain: the ventricular residue, the ventricular depolarization reduction, the percentage of significant QRS residua detected and the relations between amplitudes of the AA and the different segments of a reference lead in QT and TQ intervals.

The ventricular residue (VR) proposed by Alcaraz & Rieta (2008) permits to estimate ventricular residua in the extracted AA. VR is defined as:

$$VR_i = \frac{1}{\frac{1}{Q}\sum_{n=1}^{Q} y_A^2(n)} \sqrt{\frac{1}{2H+1} \sum_{k=r_i-H}^{r_i+H} y_A^2(k)} \cdot \max_{k=r_i-H \cdots k=r_i+H} \left(\left| y_A(k) \right| \right) \tag{19}$$

where $2H+1$ is the duration of the QRS interval in samples for a value of H corresponding to 50 ms. In this case, high values of VR indicate greater QRS residua, i.e., a poor AA extraction.

The ventricular depolarization reduction (VDR) measures the reduction of the R-peak amplitude achieved by the algorithm under evaluation (Rieta & Hornero, 2007). VDR is defined as:

$$VDR(dB) = 10\log\left(R_{RL} / R_{VR}\right) \tag{20}$$

where R_{RL} is the R-peak amplitude of a reference lead and R_{VR} is the R-peak amplitude of the recovered AA. High values of VDR indicate good performance of the algorithm.

Other parameter used to evaluate the quality of the atrial activity is the percentage of significant QRS residua detected (Lemay et al., 2007). The method identifies the significant ventricular residua comparing the absolute values of the QRS intervals with a threshold $\varsigma = m + 2\sigma$ where m is the median of the amplitude distribution of $y_A(t)$ and $\sigma = qr / 0.6745$ is the standard deviation estimated from the interquartile range qr of the amplitude distribution of $y_A(t)$. Those QRS complexes with absolute values above the threshold are considered as significant. Finally, the parameter is presented showing the percentage of significant QRS residua (PSQRS).

Langley et al. (2006) compared the amplitudes of the extracted atrial signal with a reference lead, V1, in QRST segments. In addition, they defined a reference atrial signal as the segments of V1 containing no ventricular activity and compared this reference atrial signal with noQRST and QRST segments of the extracted AA. The amplitude of each segment was the peak-to-peak amplitude.

The above parameters (VR, VDR and PSQRS) are designed for ABS methods. They only take into account the recovered atrial signal in QRST segments, ignoring the form of the signal in noQRST segments. Respect to the comparison of amplitudes, it can be carried out only with methods that do not modify the amplitude of the original signal. ICA and PCA returns signals with unit power, distorting the results when comparing the amplitudes of the signals.

We propose the kurtosis as quality measure of the atrial signal extraction. It considers the whole signal (QRST and noQRST segments) and it is valid for signals with unit power. The kurtosis of a subgaussian (supergaussian) random variable is negative (positive), allowing distinguish between atrial and ventricular activities. In the cases where the kurtosis of the atrial activity is marginally positive, it will be always lower than the ventricular activity. This property can also be extended to measure the quality of the atrial activity extraction. Since the atrial activity is near the Gaussian distribution, a high absolute value of the kurtosis means a poor extraction.

6. Results

In this section, we demonstrate the use of the new algorithms in the AA extraction and the accuracy of the proposed quality parameters compared with other quality parameters found in the literature.

We have used two different databases for a total amount of 50 patients. Ten belong to the ''PTB Diagnostic ECG Database'' from the MIT database Physionet (Goldberger et al., 2000). Each record includes 15 simultaneously measured signals: the conventional 12 leads together with the 3 Frank lead ECG. The other forty belong to a clinical database recorded at the Clinical University Hospital, Valencia, Spain, and were taken with a commercial recording system with 12 leads (Prucka Engineering Cardiolab system). For both databases, the signals are digitized at 1000 samples per second with a resolution of 16 bits.

In our experiments, we have used all the available leads for a period of 10 s for every patient. The signals were preprocessed in order to reduce the baseline wander, high frequency noise and power line interference for the later signal processing. The recordings were filtered with an 8 coefficients highpass Chebyshev filter and with a 3 coefficients lowpass Butterworth filter to select the bandwidth of interest: 0.5—40 Hz. In order to reduce the computational load, the data were downsampled to 200 samples per second with no significant changes in the quality of the results.

The detection of the R peak wave event is carried out by means of a parabolic fitting algorithm (Illanes Manriquez & Zhang, 2007). The power spectrum is obtained with the modified periodogram using the Welch-WOSA method with a Hamming window of 4096 points length, a 50% overlapping between adjacent windowed sections, and an 8192-point fast Fourier transform (FFT).

We use the following acronyms for the algorithms: cFastICA (constrained FastICA) corresponding to Eq. (11) and cSAD (constrained Sequential Approximate Diagonalization) to Eq. (14). The parameter λ is 4 for cFastICA and 2 for cSAD and the bandwidth used in both cases is ±0.85 Hz.

In Table 1 the results of the quality indexes in the frequency domain are shown. Note the agreement between the SC values and the centroid frequency indexes: cSAD obtains slightly better results in SC and I_l and I_h.

	cFastICA	*cSAD*
$\mathbf{f_p}$ **(Hz)**	5.543 ± 1.233	5.556 ± 1.247
SC	0.541 ± 0.149	0.555 ± 0.145
$\mathbf{f_c}$ **(Hz)**	5.604 ± 0.855	5.622 ± 0.885
$\mathbf{I_l}$	0.945 ± 0.040	0.950 ± 0.045
$\mathbf{I_h}$	0.565 ± 0.132	0.581 ± 0.125

Table 1. Quality indexes in the frequency domain

Figure 4 shows two examples of AA recovered from different patients by cFastICA (a) and cSAD (b) and their respective spectra (c) and (d) with the associated value of the peak frequency of each AA.

Figure 5 represents one example of AA with low value of I_l (a) and another one with low value of I_h (b) and their respective spectra (c) and (d) together with the values of the indexes. The first case presents low frequency content and the second case shows high frequency content.

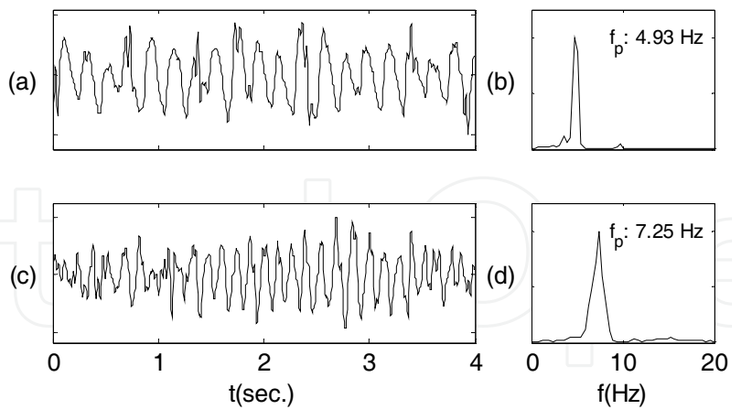

Fig. 4. Two examples of AA recovered by cFastICA (a) and cSAD (c), their spectra and the corresponding peak frequencies (b) and (d)

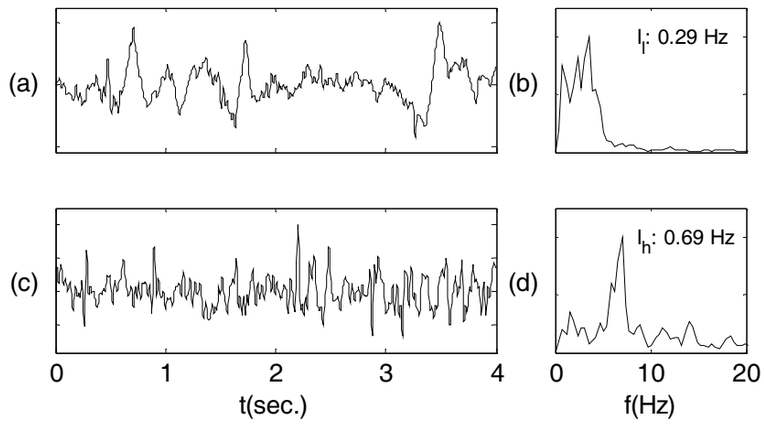

Fig. 5. (a) Atrial signal with low I_l value and its spectrum (b); (c) atrial signal with low I_h value and its spectrum (d)

Finally, we show a scatter plot of SC values and the indexes I_l (a) and I_h (b) for both methods in the Figure 6. Clearly, there exists a positive correlation between these parameters, i.e., both parameters increase together.

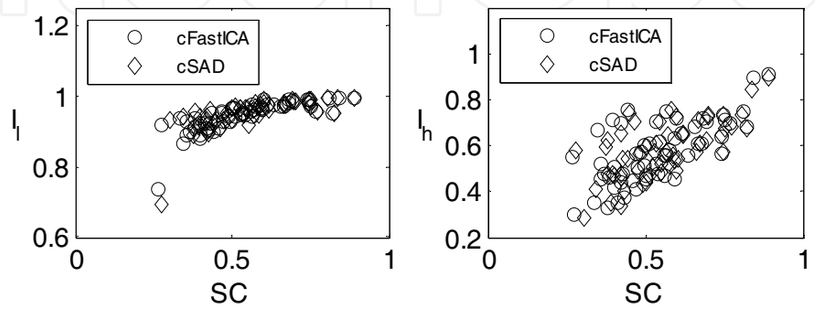

Fig. 6. Scatter plot between SC and the indexes I_l (left) and I_h (right) values

In the case of time domain quality indexes, for the sake of simplicity, we only use the parameter VR and the kurtosis. Table 2 summarizes the results obtained. Note the agreement between the VR values and the kurtosis: low values of kurtosis imply low values of VR.

	cFastICA	*cSAD*
VR	2.893 ± 1.051	3.129 ± 1.169
Kurtosis	-0.030 ± 0.648	0.219 ± 0.794

Table 2. Quality indexes in the time domain

In Figure 7 we analyze the effect of an extraction where the kurtosis of the estimated atrial signal is high (3.80), due to the presence of QRST residua near 0.6 seconds, 2 seconds and 3.5 seconds.

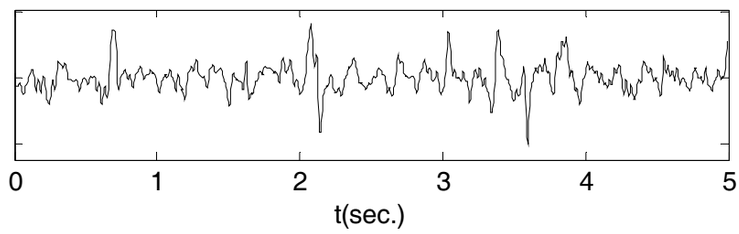

Fig. 7. Example of AA with high kurtosis value

Finally, Figure 8 represents the scatter plot between VR and kurtosis for both methods. The figure confirms the relation between the correct cancellation of the QRST complexes and the low values of the associated kurtosis.

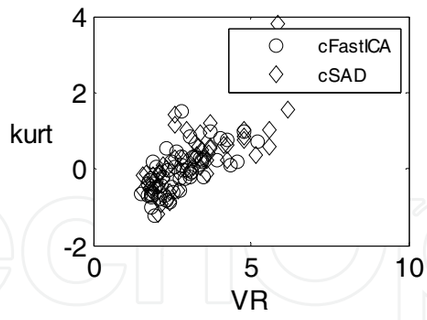

Fig. 8. Scatter plot between VR and kurtosis values

7. Conclusion

The extraction of the atrial signal in atrial fibrillation has attracted the attention of many researchers in the last years. We have reviewed the statement of the problem and the different approaches to solve it. We have shown how the traditional multilead solutions can be adapted considering the characteristics of the signals involved in the problem. In addition, we have paid attention to the difficult goal of how to measure the quality of the

extraction since the solution (the true atrial rhythm) is unknown. We have presented some indexes that can help in order to analyze the goodness of the extraction.

8. Acknowledgments

The authors would like to thank Francisco Castells and Jose Millet for sharing the real AF database provided by the cardiologists Ricardo Ruiz and Roberto Garcia-Civera during the project TIC2002-00957.

9. References

Alcaraz, R. & Rieta, J. (2008). Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. *Physiological Measurement*, Vol. 29, No. 12, 1351–1369, ISSN 0967-3334

Asano, Y.; Saito, J.; Matsumoto, K.; Kaneko, K.; Yamamoto, T. & Uchida, M. (1992). On the mechanism of termination and perpetuation of atrial fibrillation. *Am J Cardiol*, Vol. 69, No. 12, 1033–1038, ISSN 0002-9149

Bazzet, H. (1997). An analysis of the time-relations of electrocardiograms. *Annals of Noninvasive Electrocardiology*, Vol. 2, 177–194, ISSN 1082-720X

Castells, F.; Mora, C.; Rieta, J.; Moratal-Perez, D. & Millet, J. (2005). Estimation of atrial fibrillatory wave from single-lead atrial fibrillation electrocardiograms using principal component analysis concepts. *Medical and Biological Engineering and Computing*, Vol. 43, No. 5, 557– 560, ISSN 0140-0118

Castells, F.; Rieta, J.; Millet, J. & Zarzoso, V. (2005a). Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias. *IEEE Transactions on Biomedical Engineering*, Vol. 52, No. 2, 258– 267, ISSN 0018-9294

Flannery, B.; Press, W.; Teukolsky, S. & Vetterling, W. (1988). Numerical Recipes in C: The Art of Scientific Computing, *Cambridge University Press*, ISBN 0-521-43108-5

Fuster, V.; Ryden, L.E.; Asinger, R.W.; Cannom, D.S.; Crijns, H.J.; Frye, R.L.; et al (2001). ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation: executive summary, *Circulation*, Vol. 104, 2118-2150, ISSN 0009-7322

Goldberger, A.L.; Amaral, L.A.N.; Glass, L, Hausdorff, J.M.; Ivanov, P.Ch.; Mark, R.G. et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, *Circulation*, Vol. 101, 215-220, ISSN 0009-7322

Grubitzsch, H.; Modersohn, D.; Leuthold, T. & Konertz, W. (2008). Analysis of atrial fibrillatory activity from high-resolution surface electrocardiograms: Evaluation and application of a new system, *Experimental & Clinical Cardiology*, Vol. 13, No. 1, 29-35, ISSN 1205-6626

Hyvarinen, A. (1999). Fast and robust fixed point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, 626–634, ISSN 1045-9227

Hyvarinen, A.; Karhunen, J. & Oja, E. (2001). Independent Component Analysis, *Wiley*, ISBN 0-471-22131-7

Igual, J.; Llinares, R.; Guillem, M. & Millet, J. (2006). Optimal Localization of Leads in Atrial Fibrillation Episodes. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2,* pp. 1192-1195, ISBN 1-4244-0469-X

Illanes Manriquez, A. & Zhang, Q. (2007). An algorithm for QRS onset and offset detection in single lead electrocardiogram records. *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 541–544, ISBN 0-14244-0788-5

Langley, P.; Bourke, J. & Murray, A. (2000). Frequency analysis of atrial fibrillation. *Proceedings of Computers in Cardiology*, pp. 65-68, ISBN 0-7803-6557-7, Cambridge, MA, USA

Langley, P.; Rieta, J.; Stridh, M.; Millet, J.; Sornmo, L. & Murray, A. (2006). Comparison of atrial signal extraction algorithms in 12-lead ECGs with atrial fibrillation. *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 2, 343–346, ISSN 0018-9294

Lemay, M.; Vesin, J.; van Oosterom, A.; Jacquemet, V. & Kappenberger, L. (2007). Cancellation of Ventricular Activity in the ECG: Evaluation of Novel and Existing Methods. *IEEE Transactions on Biomedical Engineering*, Vol. 54, No. 3, 542–546, ISSN 0018-9294

Li, X. & Zhang, X. (2007). Sequential Blind Extraction Adopting Second-Order Statistics. *IEEE Signal Processing Letters*, Vol. 14, No. 1, 58–61, ISSN 1070-9908

Llinares, R. & Igual, J. (2009). Application of constrained independent component analysis algorithms in electrocardiogram arrhythmias. *Artificial Intelligence in Medicine,* Vol., No., (to appear) page numbers (to appear), ISSN 0933-3657

Llinares, R.; Igual, J.; Millet, J. & Guillem, M. (2006). Independent component analysis of body surface potential mapping recordings with atrial fibrillation. *Proceedings of International Joint Conference on Neural Networks,* pp. 2287-2294, ISBN 0-7803-9490-9

Manios, E.; Kanoupakis, E.; Chlouverakis, G.; Kaleboubas, M.; Mavrakis, H. & Vardas, P. (2000). Changes in atrial electrical properties following cardioversion of chronic atrial fibrillation: relation with recurrence. *Cardiovascular Research*, Vol. 47, No. 2, 244-253, ISSN 0008-6363

Petrutiu, S.; Ng, J.; Nijm, G.; Al-Angari, H.; Swiryn, S. & Sahakian, A. (2006). Atrial fibrillation and waveform characterization, *IEEE Engineering in Medicine and Biology Magazine*, Vol. 25, No. 6, 24–30, ISSN 0739-5175

Pai, G. & Rawles, J. (1989). The QT interval in atrial fibrillation, *British Medical Journal*, Vol. 61, 510–513, ISSN 0959-8138

Raine, D.; Langley, P.; Murray, A.; Dunuwille, A. & Bourke J.P. (2004). Surface atrial frequency analysis in patients with atrial fibrillation: a tool for evaluating the effects of intervention. *J Cardiovasc Electrophysiol*, Vol. 15, No. 9, 1021–1026, ISSN 1045-3873

Rieta, J.; Castells, F.; Sanchez, C.; Zarzoso, V. & Millet, J. (2004). Atrial activity extraction for atrial analysis using blind source separation. *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 7, 1176– 1186, ISSN 0018-9294

Rieta, J. & Hornero (2007). Comparative study of methods for ventricular activity cancellation in atrial electrograms of atrial fibrillation. *Physiological Measurement*, Vol. 28, 925–936, ISSN 0967-3334

Sassi, R.; Corino, V. & Maniardi, L. (2007). Analysis of surface atrial signals using spectral methods for time series with missing data. *Proceedings of Computers in Cardiology*, pp. 153-156, ISBN: 978-1-4244-2533-4

Slocum, J.; Sahakian, A. & Swiryn, S. (1992). Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity. *Journal of electrocardiology*, Vol. 25, No. 1, 1–8, ISSN 0022-0736

Shkurovich, S.; Sahakian, A. & Swiryn, S. (1998). Detection of atrial activity from high-voltage leads of implantable ventricular defibrillators using a cancellation technique, *IEEE Engineering in Medicine and Biology Magazine*, Vol. 45, No. 2, 229–234, ISSN 0739-5175

Stambler, B.; Wood, M.; Ellenbogen, K. (1997). Antiarrhythmic actions of intravenous ibutilide compared with procainamide during human atrial flutter and fibrillation electrophysiological determinants of enhanced conversion efficacy. *Circulation*, Vol. 96, No. 12, 4298–4306, ISSN 0009-7322

Stridh, M.; Bollmann, A.; Olsson, S. & Sornmo, L. (2006): Detection and feature extraction of atrial tachyarrhythmias: a three stage method of time-frequency analysis. , *IEEE Engineering in Medicine and Biology Magazine*, Vol. 25, No. 6, 31–39, ISSN 0739-5175

IntechOpen

Biomedical Engineering

Edited by Carlos Alexandre Barros de Mello

ISBN 978-953-307-013-1

Hard cover, 658 pages

Publisher InTech

Published online 01, October, 2009

Published in print edition October, 2009

Biomedical Engineering can be seen as a mix of Medicine, Engineering and Science. In fact, this is a natural connection, as the most complicated engineering masterpiece is the human body. And it is exactly to help our “body machine” that Biomedical Engineering has its niche. This book brings the state-of-the-art of some of the most important current research related to Biomedical Engineering. I am very honored to be editing such a valuable book, which has contributions of a selected group of researchers describing the best of their work. Through its 36 chapters, the reader will have access to works related to ECG, image processing, sensors, artificial intelligence, and several other exciting fields.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen