

# Estimation of Atrial Fibrillatory Waves from One-Lead ECGs Using Principal Component Analysis Concepts

F Castells<sup>1</sup>, JJ Rieta<sup>1</sup>, C Mora<sup>1</sup>, J Millet<sup>1</sup>, C Sánchez<sup>2</sup>

<sup>1</sup>Universidad Politécnica de Valencia, Valencia, Spain

<sup>2</sup>Universidad de Castilla la Mancha, Cuenca, Spain

## Abstract

*In this contribution is presented a new method for estimating the atrial activity (AA) from one-lead atrial fibrillation (AF) ECGs. This methodology is appropriate for holter signals, where the reduced number of leads is insufficient to exploit the spatial information of the ECG.*

*The proposed approach is based on principal component analysis (PCA) concepts, taking advantage from the time dependence of ventricular and atrial components. The principal components are extracted from the analysis of the ECG at successive cardiac beats, obtaining ventricular and non-ventricular related components. The AA is then reconstructed by mapping back all non-ventricular components at each cardiac beat. This methodology has been tested and validated using a significative database with simulated and real AF recordings. Its main advantage respect to adaptive template subtraction is its robustness to variations in the QRST shape, thus minimizing the QRST residua in the estimated AA.*

## 1. Introduction

Signal processing techniques have been widely employed in biomedical applications. In particular, the analysis of electrocardiograms (ECG) has provided important advances in the understanding, characterisation and diagnosis of cardiac arrhythmias. One of them is atrial fibrillation (AF), which consists of a malfunction of the atrium characterised by a modification of the normal atrial activity (AA) pattern on the ECG [1].

The proper characterisation of AF from non-invasive techniques (e.g. the ECG) requires the analysis of the atrial fibrillatory signal. However, the signal recorded at the surface skin level is a mixture of the ventricular activity (VA) and AA, and a previous step that cancels the VA, i.e. the QRS complex and the T wave, is essential [2].

Multichannel signal processing methods are applicable in this context to persistent AF [3][4], where the signals are usually recorded at an electrophysiology lab using the 12-lead standard ECG. However, in the case of early stages of AF, i.e. paroxysmal AF, these techniques are no longer valid, since the signals are usually recorded by a holter system with no more than two or three electrodes. Such a reduced number of available leads is insufficient for exploiting spatial information satisfactorily, being those techniques based on adaptive template subtraction the unique alternatives [1]. However, these techniques are very sensitive to QRST wave variations, and any QRS complex residue may hinder an appropriate analysis of the AA, e.g. when tracking the main frequency in the time-frequency domain.

In order to offer an alternative, we propose a method based on PCA concepts that is more robust to the inherent limitations to adaptive template subtraction. Following this consideration, different segments of the ECG signal can be regarded as several observations with a high degree of mutual information. The main goal of this contribution is to model the separation of VA and AA from a single lead as a PCA problem. The proposed methodology will be validated with a significant database composed of simulated AF ECGs and paroxysmal AF episodes obtained from holter recordings.

## 2. Methods

An in-depth study of VA and AA properties revealed that both components present certain time dependence: Successive QRST waves are consequence of the same bioelectrical activity, and hence the corresponding waveforms present a high degree of redundancy. AA presents also certain time dependence, since it consists of continuous wavelets whose spectrum presents a main frequency peak (typically around 6Hz) [2][5]. Therefore, its autocorrelation function exhibits significant values at non-zero lags, being equal to null only at the following time lags:

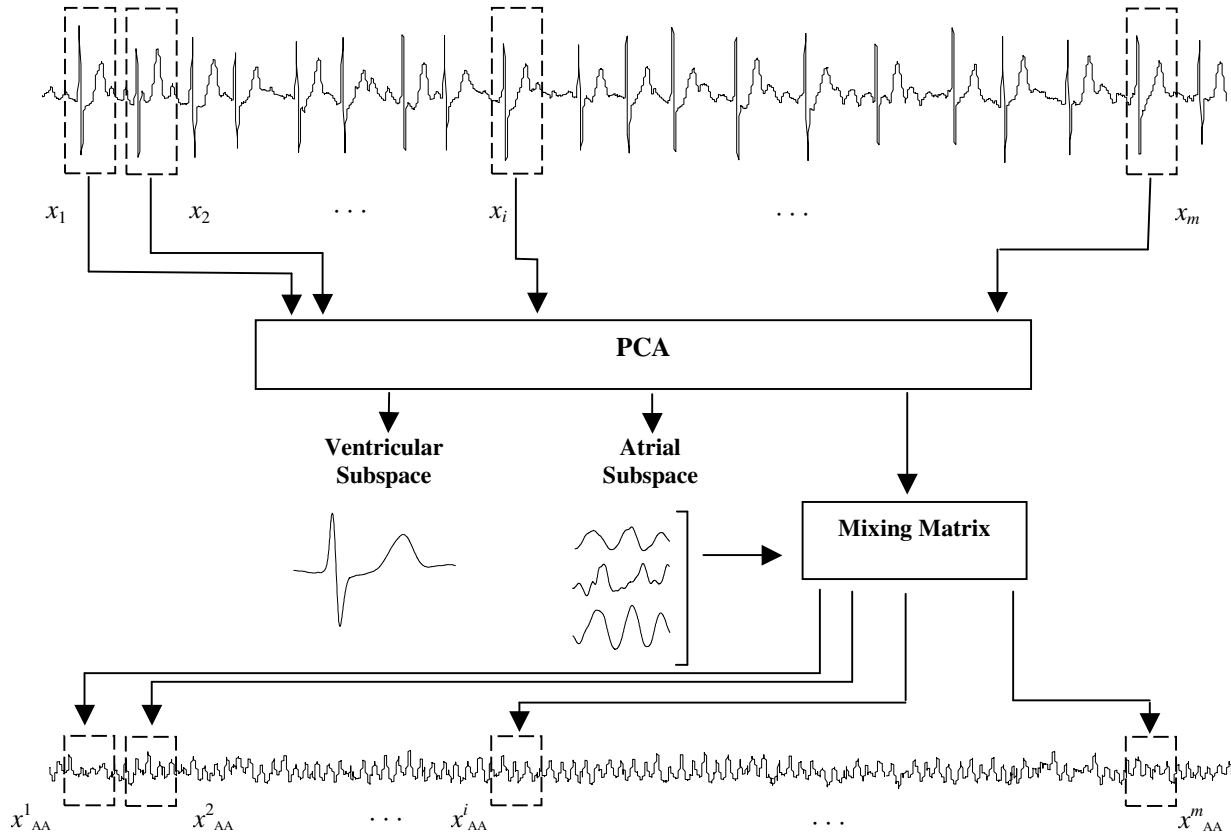


Figure 1. Illustration of the multidimensional PCA approach for the estimation of the atrial fibrillatory wave.

$$\tau_n = \frac{1 + 2k}{4} T, \quad (1)$$

being  $T$  the cycle length and  $k$  an integer index.

Motivated by this observation, we propose to extract different components corresponding to different biological activities by minimising the mutual information contained in the ECG at different cardiac beats.

The QRST waveforms can be obtained from the ECG by using an R-detector that identifies the position of the cardiac beats [6] and employing an  $n$ -length window that covers the whole Q-T interval. Considering a total number of  $m$  cardiac beats, the observations can be rewritten as an  $m$ -length vector  $\mathbf{x}(t)$  which is indeed a combination of the independent components  $\mathbf{s}(t)$ :

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (2)$$

where  $\mathbf{A}$  is the mixing matrix. The problem of recovering different components separately follows the basis of a PCA model [7]. These components mainly consist of an important ventricular component related to the QRST wave  $s_{VA}(t)$ , several components related to the AA subspace  $s_{AA}(t)$  and other nuisance sources that conform the noise subspace  $s_n(t)$ . In those cases where there is

more than one unique shape for the QRST wave, those waveforms are considered as new independent components related to the VA subspace  $\mathbf{s}_{VA}(t)$ . The mixing matrix can be decomposed in three matrices of sizes  $m \times m_{VA}$ ,  $m \times m_{AA}$  and  $m \times m_n$ , such that  $\mathbf{A} = [\mathbf{A}_{VA}, \mathbf{A}_{AA}, \mathbf{A}_n]$  is full column rank.

Hence, by identifying the AA subspace from the whole set of principal components, the AA at each observation interval  $x_{AA}^i(t)$  can be recomposed using the mixing submatrix  $\mathbf{A}_{AA}$ . The estimated AA wave in the ECG can be finally obtained by mapping back the AA content in the observations  $\mathbf{x}_{AA}(t)$  to the corresponding time intervals. Fig. 1 illustrates the methodology employed for the estimation of the AA source.

The identification of the ventricular and atrial subspaces can be carried out automatically from the covariance matrix. The ventricular components are those that contribute with much more power to the ECG signal. As a result, the eigenvalue sequence of the covariance matrix is sharply decreasing after the components related to the ventricular subspace, and a limit that separate both ventricular and atrial subspaces can always be estimated.

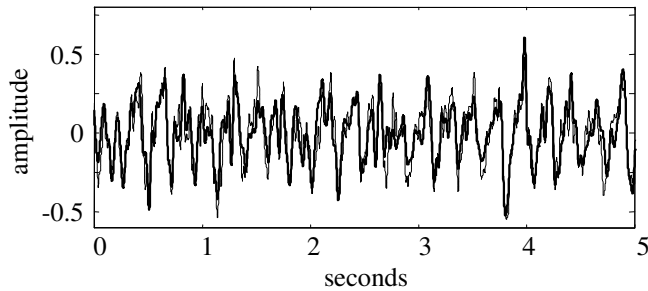


Figure 2. Comparison of the estimated AA (solid line) with the real AA (dashed line) in a typical case (patient 6).

### 3. Database

The fact that the AA is unknown in real recordings hinders an in-depth performance evaluation of the proposed methodology. Hence, suitable simulated AF ECGs with known AA content must be designed, which allows us to compare the estimated and the original AA. Ultimately the method is to be applied over actual AF episodes, and thus a database of such recordings is also employed to demonstrate the suitability of the algorithm in real scenarios.

The first database is composed of 10 simulated AF ECGs, which have been generated by adding VA and AA extracted from real AF patients. All recordings are 30 seconds in length and digitised at a sampling rate of 1 KHz and an amplitude resolution of 16 bits. The second database consists of 10 paroxysmal AF recordings obtained from Holter systems, and hence, digitised at a sampling rate of 250Hz and an amplitude resolution of 12 bits. All recordings were resampled at 1 KHz.

### 4. Results

Firstly, the proposed methodology was applied to the database of simulated AF recordings. In all cases, it was possible to remove the QRS complex and the T-wave. The estimated AA was compared with the original AA in terms of Pearson correlation indices, being of  $0.774 \pm 0.106$  in average. The performance measurement obtained for each patient is detailed in Table 1.

Table 1. Correlation values of the estimated AA and original AA for the simulated AF ECGs.

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
0.717	0.744	0.699	0.854	0.888	0.762	0.906	0.861	0.879	0.697

Table 2. Main frequency of the AA signal for each AF patient.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
$f_p$ (Hz)	5.7	4.8	4.7	3.2	6.9	6.8	3.8	4.6	5.0	5.4

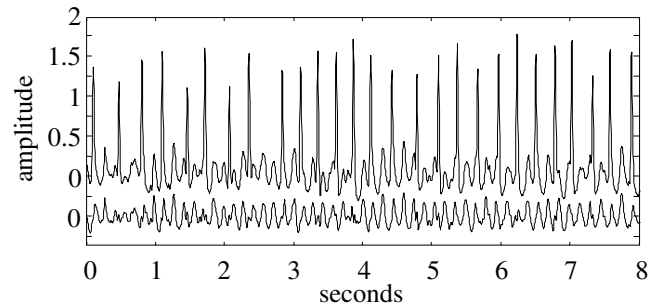


Figure 3. AA estimation from a patient that suffered from paroxysmal AF (below). The corresponding ECG signal has been plotted above in order to facilitate visual comparison.

Taking into account that the AA presents much less amplitude than the VA, these results indicate that the proposed approach is able to estimate the AA component free from QRST residua. To illustrate the quality of the AA extraction, the estimated AA corresponding to patient 6 is compared with the original AA in fig. 2.

Secondly, the same methodology was applied to the database of recordings that have been obtained from patients that suffer from paroxysmal AF, which is the final aim of the method. The number of beats for each ECG varied from 36 to 49 depending on heart rate of the patient. The number of beats available determines the number of observations. Therefore, the length of the ECG signal has a direct impact in the performance of the AA extraction. From our experiments, we suggest that the ECG length should be between 30 and 60 seconds. In all patients the QRST complex was successfully cancelled. In 6 out of 10 patients the VA subspace was conformed just by one component due to the regularity of the QRST waveform. In the remaining cases two or three VA components were identified. The number of components that corresponded to the AA subspace varied from 4 to 10. The rest of components scarcely contributed to the ECG signal and could be considered as the nuisance subspace. The fact that so few components appear in VA and AA subspaces confirms the initial assumption that VA and AA were highly dependent at different intervals.

The parameter of the AA signal that has showed major clinical importance has been the detection of the main frequency peak  $f_p$ , and the presence or absence of harmonics, which includes some information regarding the organization of the electrical activation of the atria. In all cases, the main frequency could be detected, which is specified in Table 2 for each patient.

The ECG signal of patient 1 and the corresponding AA estimation is represented in fig. 3. Notice that the VA has been completely cancelled, whereas the AA has been preserved.

## 5. Discussion and conclusions

The estimation of AA in paroxysmal AF episodes requires the implementation of QRST cancellation techniques for only one lead. Therefore, existing AA estimation techniques for multilead ECGs are suitable for persistent AF arrhythmias, but are no longer valid in the case of paroxysmal AF. However, the fact that electrical cardiac activities show a high degree of dependency at different time intervals allows us to take profit from this information and develop an AA estimation technique based on PCA concepts.

The methodology proposed in this study avoids the inherent limitations in adaptive template subtraction due to variation in the QRST shape. As a result, the QRST residue in the estimated AA is minimized. This technique has been validated in both synthesized and real AF recordings, and has been employed as a previous AA estimation stage for the prediction of spontaneous AF termination in the Computers in Cardiology Challenge 2004 [8].

## Acknowledgements

The authors would like to acknowledge the helpful support received from Servicio de Hemodinámica of the Hospital Clínico Universitario the Valencia, and specially from Ricardo Ruiz, Salvador Morell and Roberto García Civera, for providing signals and for the high quality of their clinical advice. This study has been partly funded by TIC2002-00957 and the Universidad Politécnica de Valencia (UPV).

## References

- [1] Alessie M, Konings K, Wijffels M. Atrial Arrhythmias – State of the Art: Electrophysiological Mechanism of Atrial Fibrillation. J.P. DiMarco and E.N. Prystowsky, Eds. Arrmonk, NY: Futura Publ. Co., 1995.
- [2] Bollmann A, Sonne K, Esperer HD, Toepffer I, Langberg JJ, Klein HU. Non-invasive assessment of fibrillatory activity in patients with paroxysmal and persistent atrial fibrillation using the Holter ECG. *Cardiovasc Res.* 1999;44:60-6.
- [3] Rieta JJ, Castells F, Sanchez C, Zarzoso V, Millet J. Atrial activity extraction for atrial fibrillation analysis using blind source separation. *IEEE Trans. Biomed. Eng.* 2004; 51:1176-86.
- [4] Stridh M, Sörnmo L. Spatiotemporal QRST Cancellation Techniques for Analysis of Atrial Fibrillation. *IEEE Trans. Biomed. Eng.* 2001;48:105-11.
- [5] Stridh M, Sörnmo L, Meurling C, Olsson B. Time-dependent spectral properties in atrial fibrillation signals. *IEEE Trans. Biomed. Eng.* 2001;48:19-27.
- [6] Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* 1985;32:230-6.
- [7] Jolliffe IT. *Principal Component Analysis*, Springer Verlag 2002.
- [8] Castells F, Mora C, Ruiz R, Rieta JJ, Millet J, Sanchez C, Morell S. Prediction of Spontaneous Termination of Atrial Fibrillation Using Time-Frequency Analysis of the Atrial Fibrillatory Wave. *IEEE Computers in Cardiology*, 2004.

Address for correspondence

Francisco Castells Ramon  
Escuela Politécnica Superior de Gandía – UPV  
Ctra. Nazaret-Oliva s/n  
46730 Gandía - SPAIN  
fcastells@eln.upv.es