Estimation of atrial fibrillatory wave from single-lead atrial fibrillation electrocardiograms using principal component analysis concepts

F. Castells C. Mora J. J. Rieta D. Moratal-Pérez J. Millet

Abstract—*A new method for the assessment of the atrial fibrillatory wave (AFW) from the ECG is presented. This methodology is suitable for signals registered from Holter systems, where the reduced number of leads is insufficient to exploit the spatial information of the ECG. The temporal dependence of the bio-electrical activity were exploited using principal component analysis. The main features of ventricular and atrial activity were extracted, and several basis signals for each subspace were determined. Hence, the estimated (AFW) are reconstructed exclusively from the basis signals that formed the atrial subspace. Its main advantage with respect to adaptive template subtraction techniques was its robustness to variations in the QRST morphology, which thus minimised QRST residua. The proposed approach was first validated using a database of simulated recordings with known atrial activity content. The estimated AFW was compared with the original AFW, obtaining correlation indices of 0.774 $\pm$ 0.106. The suitability of this methodology for real recordings was also proven, though its application to a set of paroxysmal AF ECGs. In all cases, it was possible to detect the main frequency peak, which was between 4.6 Hz and 6.9 Hz for the patients under study.*

Keywords—*Atrial fibrillation, ECG, Principal component analysis*

Med. Biol. Eng. Comput., 2005, **43**, 557–560

1 Introduction

THE PROPER characterisation of atrial fibrillation (AF) from non-invasive techniques requires the correct estimation of the atrial fibrillatory wave (AFW) as a previous step (SLOCUM *et al.*, 1992; BOLLMANN 1999). Several methods have already been proposed for this purpose, the most powerful techniques being those that exploit the spatial diversity of the multilead ECG (STRIDH and SÖRNMO 2001; LANGLEY *et al.*, 2000; RIETA *et al.*, 2004).

Multichannel signal processing techniques are mainly applicable in this context to persistent AF, where the signals are usually recorded using 12-lead electrocardiogram (ECG) equipment. However, the performance of these techniques is seriously reduced when they analyse the early stages of AF, i.e. paroxysmal AF, as most recordings are obtained from a Holter system with no more than two or three electrodes. Such a reduced number of leads is not sufficient to exploit the spatial information of the ECG, rendering those techniques based on averaged beat subtraction the main alternatives. However, these techniques are very sensitive to QRST wave variations, and the estimated AFW can be affected by some QRST residua, which may be important owing to the low amplitude of the AFW.

Hence, a correct AFW estimation method is required for an appropriate analysis of AF, e.g. when tracking the evolution of the main frequency peak in time–frequency analysis (STRIDH *et al.*, 2003).

The analysis of ventricular activity (VA) and atrial activity (AA) reveals that both signals present a certain time dependence. Successive QRS complexes contain redundant information and can be represented by a small number of pattern vectors. VA feature extraction using the Karhunen–Loève transform has already been employed for data reduction, when information related to noise and artifacts is discarded (MOODY and MARK 1989). The AFW is also highly autocorrelated, as its waveform presents a main cycle length, with a spectrum characterised by a main frequency peak (LANGLEY *et al.*, 2000; HOLM *et al.*, 1998).

In this work, an alternative AFW estimation method based on principal component analysis (PCA) concepts is derived. This method has been designed to overcome the inherent limitations to adaptive template subtraction, being more robust to changes in the QRST morphology. Following this consideration, different segments of the ECG signal can be regarded as several observations with a high degree of mutual information.

After the estimation of the AFW has been modelled as a PCA problem, several basis functions corresponding to the atrial activity are determined, which are finally employed for the reconstruction of the estimated AFW. The proposed metho-

Correspondence should be addressed to Mr Francisco Castells; email: fcastells@eln.upv.es

Paper received 30 December 2004 and in final form 12 April 2005

MBEC online number: 20054023

dology was validated with a significant database composed of simulated and real AF recordings.

2 Methods and materials

2.1 *AFW estimation algorithm*

Motivated by the observation that VA and AA present a certain time dependence, we proposed to extract different basis signals corresponding to VA and AA by exploiting the mutual information contained in the ECG at different time intervals. After windowing m cardiac beats employing a fixed n-length window, the different observations can be rewritten as an m-length vector $\boldsymbol{x}(t)$, which is indeed a linear combination of the basis functions $\boldsymbol{s}(t)$

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) \tag{1}$$

where $\boldsymbol{A}$ is the mixing matrix. The problem of extracting different features corresponding to different bio-electrical activities follows a PCA model (JOLIFFE, 2002). The key point in PCA is that the different elements of $\boldsymbol{x}(t)$ should be mutually correlated, which is accomplished in this formulation. After PCA processing, the principal components can be grouped in three orthogonal subspaces: the components related to the VA s_{VA} (t), the components related to the AA $s_{AA}(t)$ and other nuisance components that form the noise subspace $s_n(t)$. The mixing matrix $\boldsymbol{A}$ can be decomposed into three matrices of sizes $m \times m_{VA}$, $m \times m_{AA}$ and $m \times m_n$, such that $\boldsymbol{A} = [\boldsymbol{A}_{VA}\boldsymbol{A}_{AA}\boldsymbol{A}_n]$ is full column rank.

The identification of VA and AA subspaces can be carried out from the study of the eigenvalue sequence corresponding to the covariance matrix. These values describe a decreasing sequence where the first eigenvalue corresponds to the component with higher variance, i.e. the component with highest contribution to the windowed segments, and the last eigenvalue corresponds to the component with lower variance. The VA subspace contributes with much more variance than the AA subspace does, and the eigenvalue sequence of the covariance matrix is sharply decreasing in the edge that divides both subspaces. Hence, the ventricular components can be identified at the first locations of the principal component sequence. In addition, the components in the noise subspace contribute with much lower variance (HYVÄRINEN *et al.*, 2001) and can be found at the end positions.

With these observations, it is generally possible to identify the principal components corresponding to the VA and noise subspaces, being the rest of the components considered as the atrial subspace. The AFW at each observation interval x^i_{AA} (t) can be reconstructed from $\boldsymbol{A}_{AA}$ and $\boldsymbol{s}_{AA}$ (t)

$$\boldsymbol{x}_{AA}(t) = \boldsymbol{A}_{AA}\boldsymbol{s}_{AA}(t) \tag{2}$$

The estimated AFW in the ECG can be finally obtained by mapping back the AA content at each observation x^i_{AA} (t) to its corresponding time interval. Finally, bandpass filtering within the frequency band of the ECG is needed to overcome any undesirable edge effects when the signal segments are matched. The complete procedure that estimates the AFW from the original ECG is illustrated in Fig. 1.

2.2 *Database*

The methodology presented was validated using a database composed of ten simulated AF recordings and ten paroxysmal AF ECG. Simulated recordings allowed us to compare the estimated and original AFW, as it was known *a priori*. Synthesised AF signals were created from the combination of AA and VA, which were synthesised separately. The AFW

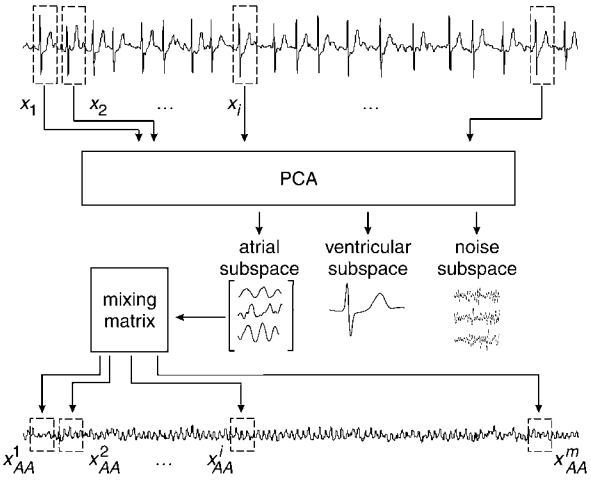

Fig. 1 *Illustration of proposed approach for AA estimation*

was generated from the smooth concatenation of successive TQ segments extracted from AF ECGs. The VA was synthesised from normal sinus rhythm ECGs, after P-wave cancellation. A detailed description of the simulation model can be found in CASTELLS *et al.* (2005).

In addition, actual AF recordings allowed us to evaluate the suitability of the algorithm to be applied over real scenarios, which was the final purpose. All these recordings were obtained from Holter systems of two leads, digitised at a sampling frequency of 128 Hz and a resolution of 12 bits. The second of these leads was the input signal of the AFW estimation approach, as it was the signal with higher AA content.

2.3 *Performance assessment*

The performance of the AFW estimation in simulated recordings was computed by comparing the estimated and original AFWs ($\hat{s}_{AA}$ and s_{AA}, respectively) in terms of Pearson correlation indices ρ. Assuming zero mean signals,

$$\rho = \frac{E[s_{AA}\hat{s}_{AA}]}{\sigma_{AA}\hat{\sigma}_{AA}} \tag{3}$$

where $E[\cdot]$ is the expectation operator, and σ_{AA}, $\hat{\sigma}_{AA}$ are the standard deviations of s_{AA} and $\hat{s}_{AA}$, respectively. The cross-correlation coefficient is a well-known parameter that is usually employed in signal processing applications to measure the similarities between two signals. This parameter becomes 1 in the case of perfect matching and 0 in the case of completely different and non-dependent signals, thus being an indicator of how well the estimated AFW reflects the original AFW.

Regarding real AF recordings, the fibrillatory frequency peak f_p is a parameter that has been shown to have major clinical importance (BOLLMANN *et al.*, 1999; PEHRSON *et al.*, 1998). The correct assessment of performance in simulated and real AF recordings is discussed in detail in CASTELLS *et al.* (2005) and LANGLEY *et al.* (2003).

3 Results

The proposed methodology was first applied to the simulated AF recordings. In all cases, it was possible to remove the QRS and T-wave. The performance measurement obtained for each recording is detailed in Table 1, being 0.774 ± 0.106 an average. The quality of the estimated AFW is illustrated in Fig. 2, which corresponds to signal **S6**. As can be appreciated, the estimated AFW nearly matches the original AFW, without any QRST residue.

Table 1 Correlation values of estimated and original AFWs for simulated AF ECGs

S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
0.717	0.744	0.699	0.854	0.888	0.762	0.906	0.861	0.879	0.697

The same methodology was also applied to ECGs from paroxysmal AF patients. The number of beats considered for each patient varied from 36 to 49, depending on the heart rate. The QRST was successfully removed in all patients. In six out of ten patients, the VA subspace was composed just of one component, owing to the regularity of the QRST waveform. In the remaining cases, two or three VA components were identified, owing to the higher variability of the QRST shape.

The number of components that corresponded to the atrial subspace varied from four to ten, depending on the patient. The rest of the components scarcely contributed to the ECG signal and could be considered as the noise subspace. The fact that so few components arise at the VA and AA subspaces in comparison with the number of beats considered for the analysis confirms the initial assumption that the observations contain a high degree of mutual information.

The main frequency peak of the AFW could be determined in all cases, which is specified in Table 2 for each patient. The frequency peak was always within the expected frequency range for AF, which is usually between 4 and 8 Hz. The ECG signal of patient 1 and the corresponding estimated AFW are represented in Fig. 3, which illustrates successful QRST cancellation, whereas the AFW has been preserved. The spectrum of the AFW is also shown in Fig. 4 which indicates a high concentration of spectral power around the main frequency peak.

4 Discussion and conclusions

The estimation of the AFW in paroxysmal AF episodes requires the implementation of QRST cancellation techniques for single-lead ECGs, rendering adaptive template subtraction

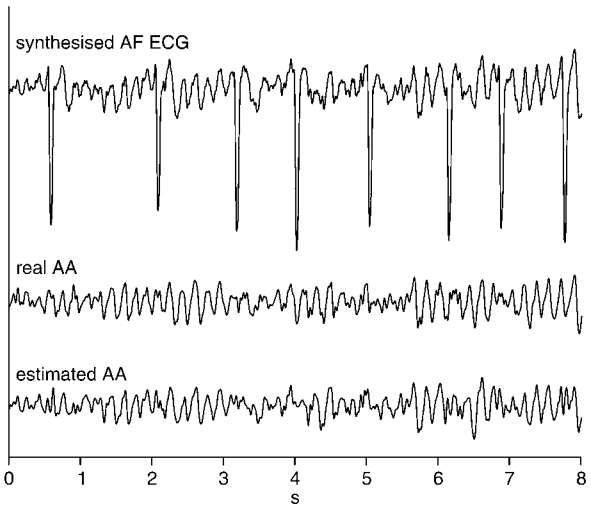

Fig. 2 *Comparison of estimated AA with real AA in typical case (patient 6)*

Table 2 Main AFW frequency for each paroxysmal AF patient

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
f_p, (Hz)	5.7	4.8	4.7	6.2	6.9	6.8	5.3	4.6	5.0	5.4

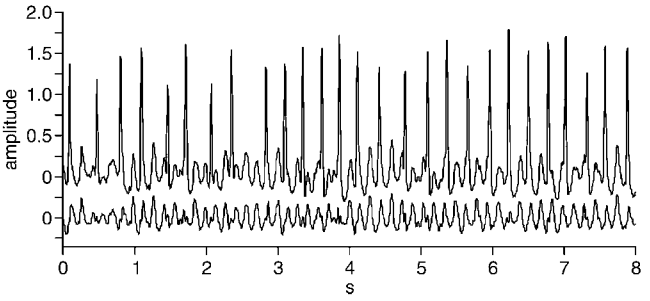

Fig. 3 *AFW estimation from paroxysmal AF ECG (below). Corresponding ECG signal has been plotted above to facilitate visual comparison*

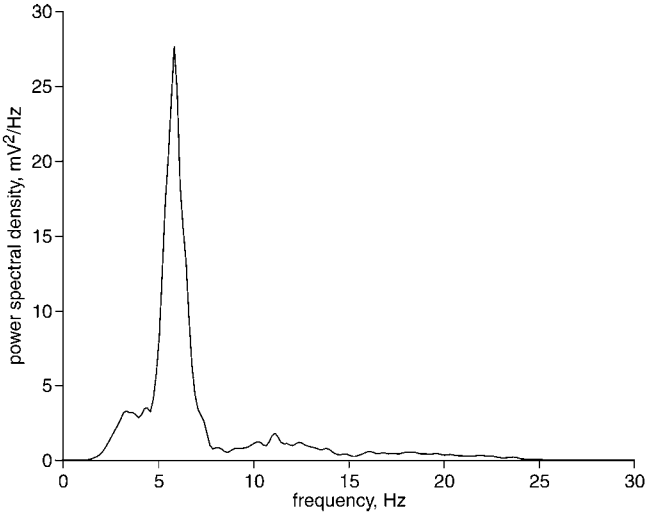

Fig. 4 *Spectrum of AFW with main frequency peak at 5.7 Hz*

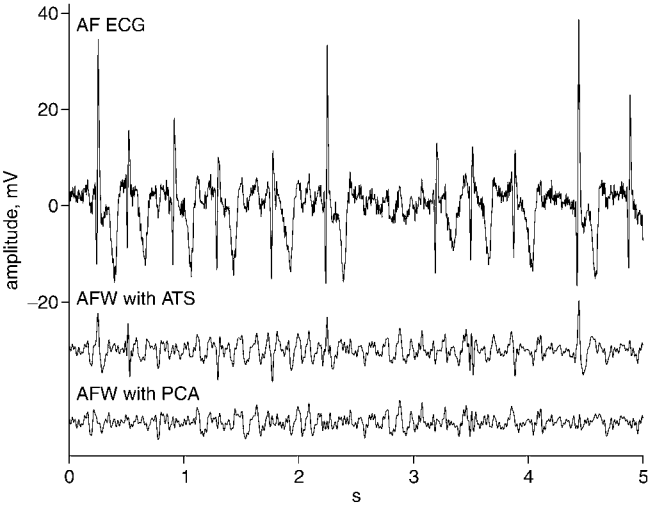

Fig. 5 *Example of AFW estimation in case with irregular QRST shape. Comparison between ATS and PCA*

(ATS) the unique existing solution. In this work, we have proposed an alternative technique based on PCA concepts that is intended to overcome the limitations of ATS, which is highly sensitive to QRST morphology variations.

The proposed approach has been validated using simulated and real AF recordings, showing its suitability for the correct estimation of the AFW. This technique presents added value with respect to ATS, as the QRST can be cancelled from a set of basis functions, instead of cancellation from one single template. This is especially useful in ECGs with variable QRST morphology. An example is shown in Fig. 5. As can

be observed, the estimated AFW obtained using ATS is highly contaminated with QRST residua, whereas PCA enhances AFW estimation, minimising any QRST content.

A model based on PCA concepts was employed to determine several basis signals for the VA and AA, from which the QRST waveforms and the AFW could be reconstructed. This approach is a more general case than the ATS algorithm, where the QRST content is cancelled from only one basis signal. As a consequence, any QRST residua are minimised. In addition, this algorithm can also remove some noise, which can be important owing to the low amplitude of the AFW.

This solution takes us one step forward in the characterisation and treatment of the most frequent cardiac arrhythmia in its earlier stages. The correct estimation of the AFW is highly important in clinical practice involving AF and could have, in the future, a direct impact on decisions about the most suitable treatment strategy.

Acknowledgments—The authors would like to acknowledge the helpful support received from Servicio de Hemodinàmica of the Hospital Clínico Universitario de Valencia and especially from Ricardo Ruiz, Salvador Morell, Roberto García Civera and Vicente Ruiz, for providing signals and for the high quality of their clinical advice.

This study has been partially funded by TIC2002-00957 and the Universidad Politécnica de Valencia (UPV) under its research incentive program.

References

BOLLMANN, A., SONNE, K., ESPERER., H., TOEPFFER, I., LANGBERG, J., and KLEIN, H. (1999): 'Non-invasive assessment of fibrillatory activity in patients with paroxysmal and persistent atrial fibrillation using the holter ECG', *Cardiovasc. Res.*, **44**, pp. 60–66

CASTELLS, F., RIETA, J., MILLET, J., and ZARZOSO, V. (2005): 'Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias', *IEEE Trans. Biomed. Eng.*, **52**, pp. 258–267

HOLM, M., PEHRSON, S., INGEMANSSON, M., SÖRNMO, L., JAHANSSON, R., SANDHALL, L., SUNEMARK, M., SMIDEBERG, B., OLSSON, C., and OLSSON, S. (1998): 'Non-invasive assessment of the atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method', *Cardiovasc. Res.*, **38**, pp. 69–81

HYVÄRINEN, A., KARHUNEN, J., and OJA, E. (2001): 'Independent component analysis' (John Wiley & Sons, New York, 2001)

JOLIFFE, I (2002): 'Principal component analysis' (Springer Verlag, Berlin, 2002)

LANGLEY, P., BOURKE, J., and MURRAY, A. (2000): 'Frequency analysis of atrial filbrillation'. Proc. Int. Conf. IEEE Computers in Cardiology 2000, IEEE Computer Society Press, pp. 65–68

LANGLEY, P., RIETA, J., STRIDH, M., MILLET, J., SÖRNMO, L., and MURRAY, A. (2003): 'Reconstruction of atrial signals derived from the 12-lead ECG using atrial signal extraction techniques'. Proc. Int. Conf. IEEE Computers in Cardiology 2003, IEEE Computer Society Press, pp. 129–132

MOODY, G., and MARK, R. (1989): 'QRS morphology representation and noise estimation using the Karhunen-Loève transform'. Proc. Int. Conf. IEEE Computers in Cardiology, 1989, IEEE Computer Society Press, pp. 269–272

PEHRSON, S., HOLM, M., MEURLING, C., INGEMANSSON, M., SMIDEBERG, B., SÖRNMO, L., and OLSSON, S. (1998): 'Non-invasive assessment of magnitude and dispersion of atrial cycle length during chronic atrial fibrillation in man', *Eur. Heart J.*, **19**, pp. 1836–1844

RIETA, J., F., C., S., V., Z., J., M. (2004): 'Atrial activity extraction for atrial fibrillation analysis using blind source separation', *IEEE Trans. Biomed. Eng.*, **51**, pp. 1176–1186

SLOCUM, J., SAHAKIAN, A., and SWIRYN, S. (1992): 'Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity', *J. Electrocardiol.*, **25**, pp. 1–8

STRIDH, M., and SÖRNMO, L. (2001): 'Spatiotemporal QRST cancellation techniques for analyses of atrial fibrillation', *IEEE Trans. Biomed. Eng.*, **51**, pp. 105–111.

STRIDH, M., MEURLING, C., OLSSON, B., and SÖRNMO, L. (2003): 'Detection of autonomic modulation in permanent atrial fibrillation', *Med. Biol. Eng. Comput.*, **41**, pp. 625–629

Authors' biographies

FRANCISCO CASTELLS received the M.Eng. degree in telecommunications engineering from the Universidad Politécnica de Valencia (UPV), Spain, in 2000, where he is currently working toward a PhD degree. After working in the telecommunications manufacturing industry for Alcatel SEL AG in Germany (2000–2001), he started his PhD studies at UPV in 2002. He is currently an Associate Lecturer with the Electronics Engineering Department, UPV, where he is also a member of the Bioengineering, Electronic and Telemedicine (BeT) research group. His research interests lie in the area of biomedical signal processing, with special emphasis on the application of blind signal processing techniques to atrial fibrillation analysis.

CIBELES MORA received the M.Eng degree in telecommunications engineering from the Universidad Politécnica de Valencia (UPV) in September 2003. She was working with the Bioengineering Electronic and Telemedicine (BeT) research group on advanced signal processing techniques, such as time-frequency transforms to study atrial tachyarrhythmias, until September 2005. She is currently working towards a PhD degree in the Department of Electronics Engineering, UPV, and she belongs to the Institute of Corpuscular Physics (IFIC), devoted to experimental research in particle and nuclear physics.

JOSÉ JOAQUÍN RIETA received the M.Eng. degree in image and sound from the Polytechnic University of Madrid, Madrid, Spain, in 1991, and the M.Sc. degree in telecommunications in 1996 and the PhD degree in Biomedical Signal Processing in 2003, both from the Universidad Politécnica de Valencia (UPV), Spain. Since 1994, he has been a Lecturer with the Electronic Engineering Department in the Valencia University of Technology, developing his teaching responsibilities at the Gandía Higher School of Technology. He has taught several subjects related to electronic and biomedical instrumentation, analog systems, data conversion systems and control engineering, and has been the author of several docent publications in these areas. He belongs to the Bioengineering, Electronic and Telemedicine (BeT) research group, where he is responsible for Advanced Signal Processing research. His research interests include statistical signal and array processing applied to biomedical signals, specially focused in cardiac signals, blind signal separation techniques, and the development of clinical applications to study and characterise the atrial activity inside the challenging problem of supraventricular arrhythmias.

JOSÉ MILLET received the M.S. degree in physics from the University of Valencia (UV), Valencia, Spain, in 1991 and the PhD degree in electrical and electronic engineering from the Universidad Politécnica de Valencia (UPV), Spain, in 1997. Since then, he has been the coordinator of the Biomedical Engineering branch of the Biomedical Engineering, Electronics, and Telemedicine (BET) research group at UPV. His professional research interests involve biomedical signal processing, biomedical signal acquisition and instrumentation, implantable devices for treatment of cardiac arrhythmias and cardiac MRI.

DAVID MORATAL-PÉREZ received the M.Sc. degree in Telecommunications Engineering from the Universidad Politécnica de Valencia (UPV), Spain in 2001. He is also Ingénieur Supélec, 2001, from the École Supérieure d'Électricité, France. He has been pursuing the Doctorate degree for the last four years in the Department of Electronics Engineering of the UPV, in close collaboration with the Radiology Department of the Emory University School of Medicine, Atlanta, USA. In 2002, he was a finalist of the Rosanna Degani Young Investigator Award, organised annually by Computers in Cardiology. His research interests are focused on image processing and magnetic resonance imaging (MRI) and, more specifically, in acceleration strategies in cardiac MRI.