An Automatic System for the Analysis and Classification of Human Atrial Fibrillation Patterns from Intracardiac Electrograms

Giandomenico Nollo*, *Member, IEEE*, Mattia Marconcini, *Student Member, IEEE*, Luca Faes, *Member, IEEE*, Francesca Bovolo, *Member, IEEE*, Flavia Ravelli, and Lorenzo Bruzzone, *Senior Member, IEEE*

***Abstract*—This paper presents an automatic system for the analysis and classification of atrial fibrillation (AF) patterns from bipolar intracardiac signals. The system is made up of: 1) a feature-extraction module that defines and extracts a set of measures potentially useful for characterizing AF types on the basis of their degree of organization; 2) a feature-selection module (based on the Jeffries–Matusita distance and a branch and bound search algorithm) identifying the best subset of features for discriminating different AF types; and 3) a support vector machine technique-based classification module that automatically discriminates the AF types according to the Wells' criteria. The automatic system was applied on 100 intracardiac AF signal strips and on a selection of 11 representative features, demonstrating: a) the possibility to properly identify the most significant features for the discrimination of AF types; b) higher accuracy (97.7% using the seven most informative features) than the traditional maximum likelihood classifier; and c) effectiveness in AF classification also with few training samples (accuracy = 88.3% with only five training signals). Finally, the system identifies a combination of indices characterizing changes of morphology of atrial activation waves and perturbation of the isoelectric line as the most effective in separating the AF types.**

***Index Terms*—Arrhythmia organization, automatic classification, feature extraction and selection, human atrial fibrillation, intracardiac electrograms, signal processing, support vector machines (SVMs).**

I. Introduction

Atrial fibrillation (AF) is a very common cardiac disorder. It is associated with an increased risk for stroke and embolic events and has an occurrence increasing with age [1]. Among the possible therapeutic approaches, the recently developed strategies based on catheter ablation targeted in the area of the pulmonary veins have provided very encouraging results in patients suffering from paroxysmal AF [2]. However, other forms of AF do not benefit out of this specific approach, and seem to require a complete evaluation of the dynamics of propagation in both atria. On that basis, the analysis of the patterns of electrical activity in different regions of the heart has been indicated as relevant to the successful ablative intervention [3], [4]. Hence, an objective and accurate characterization of the electrical activation during AF might be important for the definition of the optimal therapeutic approach.

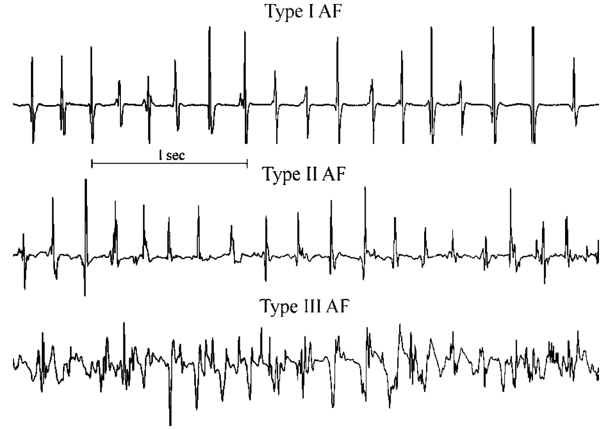

Fig. 1. Examples of bipolar intracardiac signals acquired during AF, classified into Type I, Type II, and Type III AF according to the Wells' criteria [5].

In this context, the classification of the degree of organization shown by intracardiac signals plays an important role for the definition of the complexity of AF episodes. The classification scheme currently adopted as clinical standard is that proposed by Wells *et al.* [5]. It is based on classifying single bipolar electrograms into three different types (see Fig. 1): Type I AF (AF1) shows discrete atrial electrogram complexes of variable morphology and cycle length separated by an isoelectric line free of perturbation; in Type II AF (AF2), the electrogram complexes present various perturbations and the baseline is not isoelectric; Type III AF shows highly fragmented atrial electrograms with no discrete complexes or isoelectric intervals. A major disadvantage of this approach is that the classification is subjective and time-consuming, as it is commonly executed by visual scoring of the intracardiac electrograms. Nevertheless, an analysis looking at the overall characteristics of AF electrograms such as the one proposed by Wells may have a peculiar electrophysiological relevance, as it may reflect the propagation patterns underlying the maintenance of AF [6], [7]. In addition, the Wells approach was used in several clinical and experimental studies to identify spatial organization patterns in paroxysmal and chronic AF [8]–[10], and to support the ablative treatment of AF [8], [10].

Recently, it has been demonstrated that an automated classification of bipolar intracardiac signals in accordance with the

Manuscript received August 30, 2007; revised January 30, 2008. This work was supported in part by Fondazione Cassa di Risparmio di Trento e Rovereto, Italy, under a grant. *Asterisk indicates corresponding author.*

*G. Nollo is with the Biophysics and Biosignals Laboratory, Department of Physics, University of Trento, 38050 Trento, Italy (e-mail: nollo@science.unitn.it).

L. Faes and F. Ravelli are with the Biophysics and Biosignals Laboratory, Department of Physics, University of Trento, 38050 Trento, Italy (e-mail: luca.faes@unitn.it; flavia.ravelli@unitn.it).

L. Bruzzone, M. Marconcini, and F. Bovolo are with the Remote Sensing Laboratory, Department of Information and Communication Technologies, University of Trento, 38050 Trento, Italy (e-mail: lorenzo.bruzzone@ing.unitn.it; mattia.marconcini@unitn.it; francesca.bovolo@disi.unitn.it).

Digital Object Identifier 10.1109/TBME.2008.923155

Wells' criteria is feasible [11], on the basis of methods quantifying to a different extent the organization of such signals. Indeed, several algorithms have been proposed to characterize the complexity of AF episodes starting from single-site intracardiac recordings [12]–[15]. Despite this large body of research, at present it is not clear which are the best descriptors of the complex activation patterns present during AF, and which descriptors should be integrated into an automatic classification system to obtain the best discrimination of the different AF types.

In the present study, a system for the automatic characterization of short bipolar intracardiac signals measured during AF is proposed. The system is made up of: 1) a feature-extraction module, returning a set of indices that are effective in discriminating the AF types according to the Wells' criteria; 2) a feature-selection module based on the Jeffries–Matusita (JM) distance and the branch and bound (BB) search strategy [16] aimed at identifying the features that are more informative for the classification of the AF signals; and 3) a classification module based on support vector machines (SVMs) [17]–[21] capable of providing high classification accuracy even in the presence of few training patterns. The effectiveness of the system is tested by checking the discrimination capability of each one of the extracted features, and by evaluating the classification accuracy by varying the number of selected features and the number of training patterns available for learning.

II. Data Collection and Preprocessing

A. Data Collection

The study group consisted of 11 patients with idiopathic AF, randomly chosen from among those undergoing electrophysiological tests for radiofrequency catheter ablation. In all patients, antiarrhythmic drugs were suspended for at least five half lives, and no one had received Amiodarone within the preceding six months. Electrophysiological studies were carried out using a multipolar basket catheter (Constellation catheter, Boston Scientific) placed in the right atrium via a right femoral approach. Thirty-two bipolar intracardiac recordings were acquired by coupling adjacent pairs of electrodes. The surface ECG (lead II) was also acquired. Signals were simultaneously recorded (CardioLab System, Prucka Eng., Inc.) and digitized at 1-kHz sampling rate and 12 bit precision. The typical range for the acquired signals was between -5 mV and 5 mV, corresponding to an amplitude resolution of 2.44 μV. Channels were discarded when the signal was absent or below the amplitude threshold of 70 μV (e.g., due to bad electrode–tissue contact and/or heart movement).

When not spontaneously present, AF was induced by atrial extrastimuli or atrial bursts. The duration of each considered AF episode was at least 5 min, and the first and last minutes of AF were excluded from the analysis. Each recording was carefully inspected by an experienced cardiologist and classified as normal sinus rhythm or AF of type I, II, or III. Only segments lasting at least 4 s of the same stable AF type (AF1, AF2, or AF3) were considered for the analysis. The final labeled data set consisted of 100 AF segments (35 AF1, 30 AF2, and 35 AF3), each truncated to a duration of 4 s. Examples of AF1, AF2, and AF3 signals are reported in Fig. 1. The 4 s duration was selected in accordance with the literature [6], [11], [12], as a tradeoff between the needs of favoring the consistency of organization measures that prompt for long duration, and of allowing real-time applications in the context of AF classification for clinical purposes that prompt for short duration.

B. Data Preprocessing

To minimize the effects of the ventricular interference, an adaptive template of the ventricular artifact was subtracted from the atrial recording in correspondence with the detected ventricular activation times [22]. The atrial activation times, i.e., the times representative of the passage of the propagating wave in the area under the acquiring electrode, were estimated as the local barycenters of the signal [12]. To do that, a specific procedure for atrial wave recognition, based on a specific pass-band filtering technique [12] was applied to obtain a signal with amplitude proportional to the power content of the oscillatory components typical of AF signals. The atrial waveforms were then detected from the filtered signal by threshold crossing. The barycenter of each detected wave was finally estimated as the time dividing in two equal parts the local area of the signal, and was taken as the activation time of the wave.

For a signal in which N atrial activations were detected, the activation waves (AWs), $\mathbf{x}_i,\ i = 1, \ldots, N$, were defined as signal windows lasting 90 ms (thus containing $p = 90$ points) and centered on the atrial activation times [12]. To prevent factors not related to the organization of the arrhythmia (e.g., quality of electrode contact and direction of wave propagation) from affecting the reliability of morphological indices, each AW was normalized by $\hat{\mathbf{x}}_i = \mathbf{x}_i/\|\hat{\mathbf{x}}_i\|$, where $\|\cdot\|$ indicates the Euclidean norm. As the AWs are points of the p-dimensional real space, the normalized AWs belong to the surface of the p-dimensional unitary sphere. Hence, a measure of the morphological dissimilarity between two normalized AWs $\mathbf{x}_i$ and $\mathbf{x}_j$ was taken as the standard metric of the sphere, i.e. $d(\mathbf{x}_i,\ \mathbf{x}_j) = \arccos(\mathbf{x}_i \cdot \mathbf{x}_i)$, where "$\cdot$" denotes the dot product.

III. Feature Extraction Module

The extraction of the features to be given as input to the selection module was performed after an exhaustive review of the current literature, aimed first to categorize the different approaches that can be followed to describe the complexity of single intracardiac recordings from a signal processing point of view, and then to select, within each considered approach, the measures that in previous studies were shown to better discriminate the different AF types. With this extraction criteria, 11 indices based on atrial rhythm analysis, time-domain signal processing, Fourier analysis, signal quantization, and morphological evaluation were selected as detailed next. Fig. 2 shows the distribution within the three AF classes of the 11 indices estimated for the 100 labeled signals and normalized between 0 and 1.

A. Features Based on Atrial Activation Times

After detection of the AWs as described earlier, the atrial cycle length series was calculated as the sequence of the time

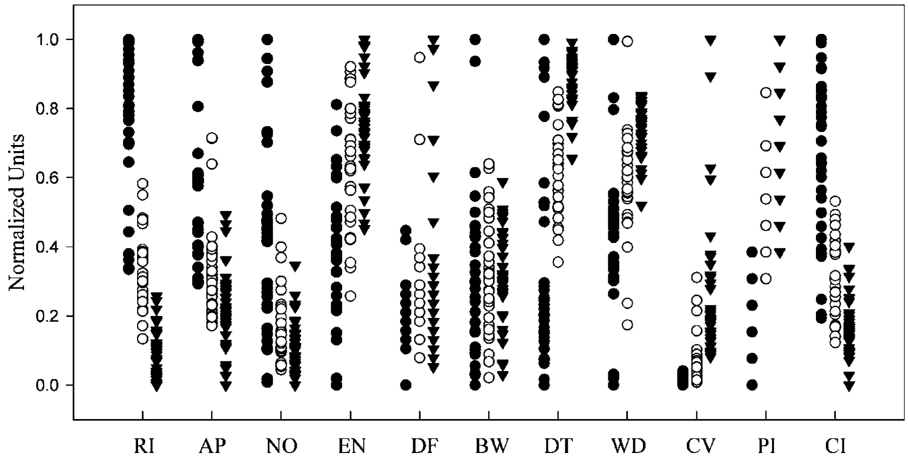

Fig. 2. Distribution of the 11 indices, extracted as features of the proposed classification system, on the three AF classes (AF1: filled circles, AF2: empty circles; AF3: triangles). From left to right: regularity index (RI), mean atrial period (AP), number of baseline points (NO), Shannon entropy (EN), dominant frequency (DF), signal bandwidth (BW), distance to a template (DT), average wave duration (WD), atrial period coefficient of variation (CV), principal component analysis index (PI), and cluster analysis index (CI).

intervals occurring between each pair of consecutive detected activation times. The mean atrial period (AP) and its coefficient of variation (CV) were then obtained by taking the mean of the time intervals and their standard deviation normalized to the mean, respectively. These two indices are commonly used as simple descriptors of AF dynamics as it was observed that episodes of increasing complexity show atrial periods of shorter duration and higher beat-to-beat variability [13].

B. Features Based on Time-Domain Analysis

The duration of each detected AW was defined as the length of the window containing 90% of the total power of the wave. The average of the wave durations (WD) contained in the analyzed signal was then taken as a time-domain feature for the classification analysis. The WD values are expected to be inversely related to the organization of AF, as signals with increasing complexity class usually present longer AWs that are the result of the interaction among a larger number of fibrillatory wavelets [6].

C. Features Based on Frequency-Domain Analysis

The power spectral density (PSD) of each signal was estimated by means of the weighted autocovariance method, i.e., by Fourier transforming the truncated and windowed autocorrelation function of the signal. The Hanning window, with a spectral bandwidth of 0.02 Hz, was used to smooth the autocorrelation during PSD estimation, and 1024 points were chosen for PSD representation. The total power of the signal was computed by integrating the PSD up to 200 Hz, and the signal bandwidth (BW) was then defined as the frequency bin below which 95% of the total power of the signal was contained. The index BW was selected as the first frequency-domain feature, upon the consideration that more complex AF signals exhibit more spread frequency spectra [11]. Another feature based on power spectrum calculation is the dominant frequency (DF) of the signal. This parameter is gaining importance for the characterization of AF organization from single intracardiac recordings, based upon the consideration that the degree of organization is related to the presence of well-defined oscillatory components in the intracardiac signals [14]. In this study, the DF was obtained as the peak frequency of the Fourier transform of the signal obtained after applying the Hanning window and bandpass filtering (3–15 Hz) the original signal.

D. Features Based on Signal Quantization

Based on the rationale that perturbations of the isoelectric line of AF signals are associated with their complexity class [5], two features resulting from the quantization of the signal amplitude were considered. Quantization was performed by normalizing the data within the analyzed signal to the average amplitude of the detected AWs, and then by dividing the amplitude range into 33 levels [11]. The first feature was the relative number of baseline points (NO), calculated as the number of points falling into the central quantization level divided by the total number of points in the signal [15]. The second feature was the estimate of the Shannon entropy (EN) of the basis of the proposed quantization

$$\mathrm{EN} = \sum_{i=1}^{33} p_i \ln p_i \tag{1}$$

where p_i is the probability density of the ith quantization level, estimated as the relative number of points falling into that level. With these definitions, NO is expected to decrease, and EN to increase, while increasing the complexity class of the analyzed signal.

E. Features Based on Morphological Analysis

Four different features measuring the morphological similarity among the AWs detected in each AF signal were extracted. The relevance of these features to the classification analysis relies on the consideration that AF signals of increasing complexity class exhibit a lower degree of similarity among their AWs [23]. Correlation waveform analysis [11] was performed using the average of the normalized AWs as a template representing the mean wave, and calculating the average distance to the template (DT) as the mean of the distances of each normalized AW to the template.

For a signal with N AWs, the regularity index (RI) was defined as the relative number of similar pairs of AWs [12]

$$\mathrm{RI} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \Theta\left(\varepsilon - d\left(\mathbf{x}_i, \mathbf{x}_j\right)\right) \tag{2}$$

where θ is the Heaviside function and the distance threshold defining the similarity between two AWs (i.e., $\mathbf{x}_i$ and $\mathbf{x}_j$ are similar if $d(\mathbf{x}_i, \mathbf{x}_j))$ was set to $\varepsilon = \pi/3$ rad [12]. This feature is an estimate of the probability of finding two similar AWs in the considered signal.

Principal component analysis was exploited to find the data representation such that the variability in morphology among the AWs was minimal [24]. Briefly, the eigenvectors of the covariance matrix of the AWs were found and sorted in decreasing order of the corresponding eigenvalues. Since the eigenvalues account for the fraction of variability among AWs, the principal components were defined as the sorted eigenvectors such that their corresponding eigenvalues encompassed at least 95% of the variability. The number of principal components (PI) was finally taken as an organization measure.

Cluster analysis was implemented to measure the tendency of the AWs to be assigned to few groups having similar characteristics [24]. The algorithm implemented was based on hierarchical agglomerative clustering, by which the AWs were grouped iteratively on the basis of the dissimilarity measure taken as the standard metric of the p-dimensional unitary sphere to which normalized AWs belong. The index based on cluster analysis (CI) measured the level of grouping of the AWs, and was inversely related to the minimum distances found during the iteration of the clustering process. Details of the algorithm are given in [24].

IV. Feature Selection Module

Given n available features obtained by feature extraction, the aim of feature selection is to identify the subset of $m < n$ features that, among all the possible subsets of m features, is more effective in discriminating the considered information classes. The optimal approach to perform feature selection would be using the same algorithm (i.e., the SVM) adopted for the subsequent classification phase. This approach needs to evaluate the classification accuracy versus all the possible combinations of features given as input to the classifier and this would require a very high computational time, particularly with the adopted SVM classifier that for each possible combination of features would require an intensive model selection phase. For this reason, we use a feature selection technique based on a simpler, but yet effective criterion function (which measures the effectiveness of each considered subset of features) and on an efficient search algorithm (which explores the solution space by evaluating explicitly only a subset of feature combinations). This choice assures a low computational load in the training phase thus improving the operational utility of the overall system.

A. Criterion Function

Feature selection identifies from the set F of the n =11 available features the subset $F_m^* \subset F$ maximizing an appropriate criterion function, $J(\cdot)$, evaluating the separability of the information classes for a given subset of features. Based on theoretical properties and experimental evidences we considered the JM distance as a criterion function [25]. The JM distance represents a measure of the average statistical distance between the conditional probability density functions $p(\mathbf{x} \mid \omega_i)$ and $p(\mathbf{x} \mid \omega_j)$ related to the information classes ω_i and ω_j. This establishes an explicit relationship between the behaviors of the feature-selection criterion and the Bayesian error probability of the classifier, providing important indications on the number of features necessary for properly discriminating classes. We calculated the JM distance by

$$J_{ij}\left(F_m^*\right) = \sqrt{2\left(1 - e^{-B_{ij}\left(F_m^*\right)}\right)} \tag{3}$$

where B_{ij} is the Bhattacharyya distance. Under the assumption that ω_i and ω_j can be modeled by a Gaussian distribution, the Bhattacharyya distance can be expressed as

$$B_{ij}\left(F_m^*\right) = \frac{1}{8}\left(\mathbf{m}_i - \mathbf{m}_j\right)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} \left(\mathbf{m}_i - \mathbf{m}_j\right) + \frac{1}{2} \ln \frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{\sqrt{|\Sigma_i|\,|\Sigma_j|}} \tag{4}$$

where $\mathbf{m}_i$ and $\mathbf{m}_j$ are the mean values of the distributions of ω_i and ω_j, respectively, and Σ_i and Σ_j are the corresponding covariance matrices.

The addressed multiclass problem is defined by a set $\Omega = \{\omega_1, \omega_2, \omega_3\}$ of three information classes, associated with the three investigated types of AF (i.e., AF1, AF2, and AF3). In order to use the JM distance as a criterion function in the problem of discriminating among $\omega_1,\ \omega_2,$ and ω_3, we exploited its multiclass extension [26], [27]

$$\mathrm{JM} = \sum_{i=1}^{3} \sum_{j>1}^{3} \sqrt{P\left(\omega_i\right) P\left(\omega_j\right)} \cdot \mathrm{JM}_{ij}^2 \tag{5}$$

where $P(\omega_i)$ represents the prior probability of the generic ith class.

B. Search Algorithm

As the number of considered features is not too large, we adopt the branch and bound (BB) algorithm, which is very efficient as it avoids exhaustive enumeration by rejecting suboptimal combinations of features without a direct evaluation of the criterion function [16], [28]. Assuming a criterion function that satisfies monotonicity, the BB algorithm selects the subset of features that optimize the criterion function (i.e., maximize the JM). The BB algorithm is independent from the ordering of the features, does not enumerate any sequence more than once (even as permutation), and considers, either explicitly or implicitly, all possible sequences. The reader is referred to [29] for more details about the algorithm.

V. Classification Module: SVM Technique

We based our classification module on SVMs [17]–[21]. SVMs perform linear separation of the patterns belonging to two information classes selecting the hyperplane that maximizes its distance from the closest training pattern of both classes (i.e., the margin) in the space where the samples are mapped.

Let $Z = \{\mathbf{z}_l\}_{l=1}^{M}, \mathbf{z}_l \in \Re^m$ be a set of M training samples, made up of m features chosen by the feature selection module from the 11 available features. As SVMs are binary classifiers, the strategy adopted to solve the addressed multiclass problem defined by the set $\Omega = \{\omega_1, \omega_2, \omega_3\}$ was the one-against-all strategy, which involves a parallel architecture of three different SVMs (one for each class). The sth SVM, $s = 1, \ldots 3$, solves the binary problem defined by the information class $\{\omega_s\}$ against all the others, $\Omega - \{\omega_s\}$. The "winner-takes-all" rule is used to make the final decision: given a pattern $\mathbf{z}$, the winning class is the one corresponding to the SVM with the highest output, i.e. $\mathbf{z} \in \omega_i \Leftrightarrow \omega_i = \arg\max\{f_s(\mathbf{z})\}$, $s = 1, 2, 3$, where $f_s(\mathbf{z})$ represents the output of the sth SVM.

For the generic sth SVM, let us define $Y_s = \{y_{sl}\}_{l=1}^{M}$ the set of labels associated with the training samples $\{\mathbf{z}_l\}_{l=1}^{M}$, where $y_{sl} = +1$ if $\mathbf{z}_l \in \omega_s$ and $y_{sl} = -1$ otherwise. To simplify the notation, in the following we will omit the subscript s. SVMs aim at linearly separating data by means of the hyperplane: $h\colon f(\mathbf{z}) = \mathbf{w} \cdot \mathbf{z} + b = 0$, where $\mathbf{z}$ is a generic sample, $\mathbf{w}$ is a vector normal to the hyperplane, b is a constant such that $b/||\mathbf{w}||^2$ represents the distance of the hyperplane from the origin, and $d(h_1\colon \mathbf{w} \cdot \mathbf{z} + b = -1,\ h_2\colon \mathbf{w} \cdot \mathbf{z} + b = +1) = 2/||\mathbf{w}||^2$ represents the margin. The concept of margin is central in the SVM algorithm as it is a measure of the generalization capability: the larger the margin is, the higher the expected generalization will be. Accordingly, maximizing the margin is equivalent to minimize $||\mathbf{w}||$; thus, SVMs solve a quadratic optimization problem with proper inequality constraints

$$\begin{cases} \min\limits_{\mathbf{w},b,\xi} \left\{ \dfrac{1}{2} \|\mathbf{w}\|^2 + C \displaystyle\sum_{l=1}^{M} \xi_l \right\} \\ y_l \left(\mathbf{w} \cdot \mathbf{z}_l + b \right) \geq 1 - \xi_l \quad \forall l = 1, \ldots, M \\ \xi_l > 0. \end{cases} \tag{6}$$

To allow the possibility for some training samples to fall within the margin band, $R = \{\mathbf{z} \,|\, \mathbf{z} \in \Re^m, -1 \leq f(\mathbf{z}) \leq 1\}$, for increasing the generalization ability of the classifier, the slack variables ξ_l and the associated *penalization parameter C* are introduced. The constraints imply a penalty of cost $C\xi_l$ for each data point that falls within the margin on the correct side of the separation hyperplane (i.e., $0 < \xi_l \leq 1$), or on its wrong side (i.e., $\xi_l > 1$). In this way, the penalty is proportional to the amount by which a given pattern is misclassified. The parameter C controls the relative weighting between the goals of making the margin large and that of minimizing the number of misclassified samples. Larger values of C involve a larger penalty for classification errors; hence, each misclassified pattern can exert a stronger influence on the boundary.

As direct handling of inequality constraints is difficult, Lagrange multipliers $\alpha_{l=1}^{M}$ are introduced for obtaining the equivalent dual representation

$$\begin{cases} \max\limits_{\boldsymbol{\alpha}} \left\{ \displaystyle\sum_{l=1}^{M} \alpha_l - \dfrac{1}{2} \sum_{l=1}^{M} \sum_{i=1}^{M} y_l y_i \alpha_l \alpha_i \mathbf{z}_l \cdot \mathbf{z}_i \right\} \\ 0 \leq \alpha_l \leq C, \quad 1 \leq l \leq M \\ \displaystyle\sum_{l=1}^{M} y_l \alpha_l = 0. \end{cases} \tag{7}$$

According to the Karush–Kuhn–Tucker conditions [19], [20], the solution is a linear combination of either mislabeled training samples or correctly labeled training samples falling into the margin band. These samples are called *support vectors* (SVs) and are the only patterns associated with nonzero Lagrangian multipliers. To make the constrained minimization process (7) efficient, quadratic optimization techniques are employed [30]. Hence, once the dual variables α_l are obtained, it is possible to determine $\mathbf{w}$ and to predict the label for a given sample $\mathbf{z}$ according to $\hat{y} = \operatorname{sgn}[f(\mathbf{z})]$. If the data in the input space cannot be linearly separated, they can be projected into a higher dimensional feature space (e.g., a Hilbert space) with a nonlinear mapping function $\Phi(\cdot)$ defined in accordance with the Cover's theorem [31]. As a consequence, the inner product between the two mapped feature vectors $\mathbf{z}_l$ and $\mathbf{z}_i$ becomes $\Phi(\mathbf{z}_l) \cdot \Phi(\mathbf{z}_i)$. In this case, due to the Mercer's theorem [32], by replacing the inner product in (7) with a kernel function $k(\mathbf{z}_l, \mathbf{z}_i) = \Phi(\mathbf{z}_l) \cdot \Phi(\mathbf{z}_i)$, it is possible to avoid representing the feature vectors explicitly. Thus, the dual representation with the constraint $0 \leq \alpha_l \leq C$ can be expressed in terms of the inner product with a kernel function as follows:

$$\begin{cases} \max\limits_{\boldsymbol{\alpha}} \left\{ \displaystyle\sum_{l=1}^{M} \alpha_l - \dfrac{1}{2} \sum_{l=1}^{M} \sum_{i=1}^{M} y_l y_i \alpha_l \alpha_i K_{li} \right\} \\ 0 \leq \alpha_l \leq C, \quad 1 \leq l \leq M \\ \displaystyle\sum_{l=1}^{M} y_l \alpha_l = 0 \end{cases} \tag{8}$$

where $K_{li} = k(\mathbf{z}_l, \mathbf{z}_i)$ is the generic element of the M-squared positive definite matrix $\mathbf{K}$ that is called *kernel matrix*. $\mathbf{K}$ is symmetric and satisfies the following condition:

$$\sum_{l=1}^{M} \sum_{i=1}^{M} \alpha_l \alpha_i K_{li} > 0. \tag{9}$$

Unlike other classification techniques, the kernel $k(\cdot,\cdot)$ ensures that the objective function is convex and accordingly, there are no local maxima in the cost function in (12). Due to their well-proved very good performances in several different frameworks, we employed Gaussian radial basis function (RBF) kernels

$$k(\mathbf{z}_l, \mathbf{z}_i) = \exp\left(-\frac{\|\mathbf{z}_l - \mathbf{z}_i\|^2}{2\sigma^2} \right)$$

where σ represents the spread parameter and tunes the generalization ability of the SVM.

TABLE I
AVERAGE PERCENTAGE OVERALL ACCURACY ± STANDARD DEVIATION (COMPUTED OVER 1000 REALIZATIONS) OBTAINED CONSIDERING EACH OF THE AVAILABLE EXTRACTED FEATURES INDIVIDUALLY

Rank	Feature	Overall Accuracy %			
		Multiclass	AF1 vs. AF2	AF1 vs. AF3	AF2 vs. AF3
1	RI	87.96±3.14	89.75±3.49	100.0±0.00	92.88±3.10
2	DT	82.55±3.97	90.19±3.84	92.47±2.73	89.31±3.69
3	CV	78.29±3.13	78.69±4.35	100.0±0.00	89.94±4.06
4	WD	75.51±3.47	82.00±4.11	95.76±2.73	84.06±3.86
5	PI	75.15±3.69	92.13±3.43	96.59±2.48	74.44±4.28
6	CI	74.07±3.59	84.50±4.31	94.71±3.36	77.69±4.45
7	EN	65.57±4.07	78.63±4.36	87.76±4.06	72.75±4.68
8	AP	64.77±3.48	80.03±4.62	89.94±4.16	71.21±4.92
9	NO	63.39±3.50	77.69±4.54	84.47±3.94	67.02±4.83
10	DF	48.45±3.98	64.25±4.05	68.65±4.27	58.12±4.97
11	BW	46.24±3.74	60.31±4.78	66.06±4.21	61.13±4.34

RI: regularity index; DT: distance to a template; CV: coefficient of variation of atrial period; WD: depolarization width; PI: principal component analysis index; CI: cluster analysis index; EN: Shannon entropy; AP: mean atrial period; NO: number of baseline points; DF: dominant frequency; BW: signal bandwidth.

TABLE II
FEATURES SELECTED BY THE BB ALGORITHM FOR DIFFERENT VALUES OF m

Number of Selected Features (m)	Features Selected										
1	RI										
2	RI	NO									
3	RI	NO	CV								
4	RI	NO	CV	DT							
5	RI	NO	CV	DT	PI						
6	RI	NO	CV	DT	PI	BW					
7	RI	NO	CV	DT	PI	BW	DF				
8	RI	NO	CV	DT	PI	BW	DF	CI			
9	RI	NO	CV	DT	PI	BW	DF	CI	AP		
10	RI	NO	CV	DT	PI	BW	DF	CI	AP	WD	
11	RI	NO	CV	DT	PI	BW	DF	CI	AP	WD	EN

Abbreviations as in Table I.

VI. RESULTS AND DISCUSSION

To assess the effectiveness of the proposed system, we carried out different experiments focusing the attention on: 1) the discrimination capability of each extracted feature; 2) the accuracy of the system at varying the number of features selected; and 3) the behavior of the system at changing the number of training patterns. In all experiments, we characterized the classification performances in terms of percentage overall accuracy (OA%). The learning of the SVMs was performed by the sequential minimal optimization (SMO) algorithm [33].

A. *Discrimination Capability of Each Single Feature*

To obtain a robust evaluation of the discrimination capability of each single feature, we carried out 1000 sets of experiments. In each experiment, 50 out of the 100 labeled patterns were randomly selected as training set, and the remaining 50 patterns defined the test set. The classification performances of each of the 11 extracted features were assessed individually for each experiment, and then averaged over the 1000 experiments. The model selection of the classifier (i.e., identification of the most suitable free parameters C and σ) was carried out for each experiment according to a grid search strategy.

In Table I, the mean and dispersion values of OA% are reported for the 3-class problem and the specific 2-class problems (i.e., AF1 vs. AF2, AF1 vs. AF3, and AF2 vs. AF3). The features are ranked on the basis of the measured accuracy for the 3-class problems. RI outperformed all the other features, providing very good individual results. The remaining features can be grouped in four different sets according to their performances: 1) DT and CV proved quite good discrimination capabilities (OA% $>$78); 2) WD, CI, and PI had average performances, (OA% $\sim$ 75); 3) EN, AP, and NO did not provide satisfactory results (OA% $<$70); and 4) DF and BW exhibited poor performances (OA% $<$50). From the analysis of the results obtained for the 2-class problems we note that, with few exceptions, the behavior of the single feature is similar to the one exhibited with the 3-class problem. Moreover, as expected, the subproblem AF1 versus AF3 proved to be less critical than the others.

These results have a data mining value with respect to the problem of classifying AF signals. In fact, the experiments assess the intrinsic capability of each parameter to describe the considered problem, and thus, from a physiological viewpoint, help relating the salient characteristics of the atrial signal to the organization of AF evaluated in accordance with the clinical practice. In particular, we observed that the most discriminating parameters (i.e., RI and DT) are those quantifying the variability in the morphology of the atrial activations. This finding is expected as the reference Wells' classification is mainly based on observing the variability over time of the signals' morphology. The good performance provided by measuring the dispersion of the atrial intervals (by CV computation) and the duration of the atrial waves (by WD computation) can be explained again on the basis of the Wells' definition of AF organization. Indeed, more disorganized AF is manifested by wave fragmentation, which is visually observed by the expert cardiology and automatically quantified by CV and WD, as fragmentation results in increased activation times variability and activation waves duration. On the contrary, parameters based on signal quantization (EN, NO) and frequency analysis (DF, BW), though indicated as good descriptors of the complexity of AF [14], [15], [34], resulted less related to Wells' classification, as they exhibited the worse individual performance in classifying the AF types.

B. *Accuracy of the System Versus the Number of Selected Features*

Feature selection, performed through the BB algorithm, allowed us to analyze the performances of the proposed system by varying the number of selected features. To this end, we used the same training and test sets generated for the experiments of Section VI-A and adopted the same grid search strategy for determining the best values for the SVM parameters. Table II reports the subsets of features selected by the BB algorithm for different values of m. The distribution of the features selected for $m = 1$ (RI), for $m = 2$ (RI and NO), and for $m = 3$ (RI, NO, and CV) are depicted in Fig. 3, where the better discriminative

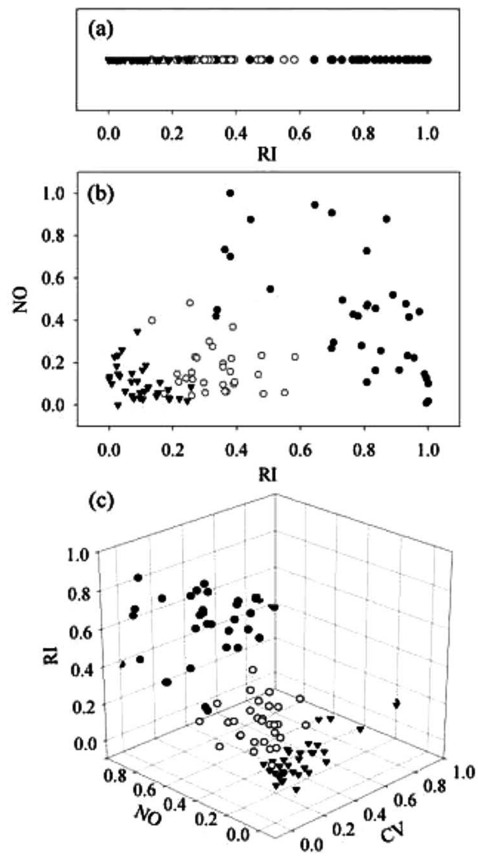

Fig. 3. Distribution of the features in a feature space of dimension (a) $m = 1$, (b) $m = 2$, and (c) $m = 3$. The optimal m features selected by the BB algorithm were represented: regularity index (RI), number of baseline points (NO), and atrial period coefficient of variation (CV).

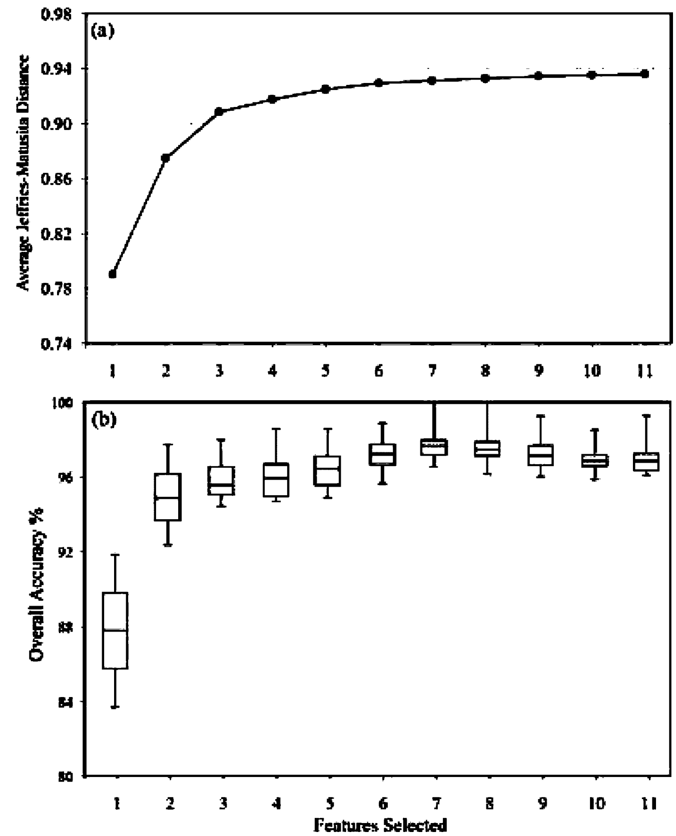

Fig. 4. Performances of the system versus number of selected features. (a) Average JM distance computed on the subsets of features selected by the BB algorithm. (b) Percentage overall accuracy (statistics over 1000 realizations) obtained with the features selected by the BB algorithm with 50% of labeled patterns considered in the learning of the classifier.

capability by employing an increasing number of features can be easily appreciated looking at the separation of symbols corresponding to the three classes.

When, as in our case, an optimal search algorithm is adopted, the subset of features selected for a generic $m = i$ is independent from the one selected for $m = i + 1$. This means that a feature belonging to the former subset does not necessarily belong to the latter. However, we obtained that, as soon as a feature is inserted into the eligible subset for a given value of m, it is always selected also for greater values of m. This aspect is very interesting as it allows us to perform a new ranking of the features based on the value of m corresponding to the selected subset with the lowest cardinality they belong to. Such analysis provides more detailed information about the usefulness of the features when considered together. As an example, even if NO proved poorly effective in the individual classification, it gave an important contribution to class separation if associated with RI. In fact, these two indexes provide complementary information, as they automatically quantify the two major aspects that are accounted for during manual Wells' classification, i.e., the changes in morphology of the activation waves and the perturbations of the isoelectric line of the atrial signals, respectively. Also BW and DF proved to be relatively effective when employed together with other features, although they exhibited the worst discrimination capabilities when considered singularly. Once again, this result can be explained considering that they provide frequency-domain information which is mostly uncorrelated with that coming from morphological analysis (by RI) and signal quantization (by NO). On the contrary, even if CI and WD proved individually rather effective, their information contribution seem redundant as they were inserted in the eligible subset only for high values of m. Also EN has a similar behavior and does not seem to be useful at all, as it is considered only when $m = 11$.

The JM distance as a function of the number of selected features is reported in Fig. 4(a). As m increases, the JM distance increases and saturates around 0.94, which corresponds to good separability among the information classes (theoretical complete separability is given by the upper bound $\sqrt{2}$). Even though the behavior of JM distance monotonically increases with m, its improvement is weak for a large number of considered features (e.g., it increased from 0.932 to 0.936 while m ranged from 7 to 11). Hence, one may expect to obtain nearly optimal classification results by considering only a subset of features. This observation is confirmed by looking at the average OA% obtained versus the number of selected features [Fig. 4(b)]. The results indicate that the classification performances improved as the number of selected features increased up to seven, while the addition of further features did not improve accuracy. This behavior is explained by the poor generalization ability of the

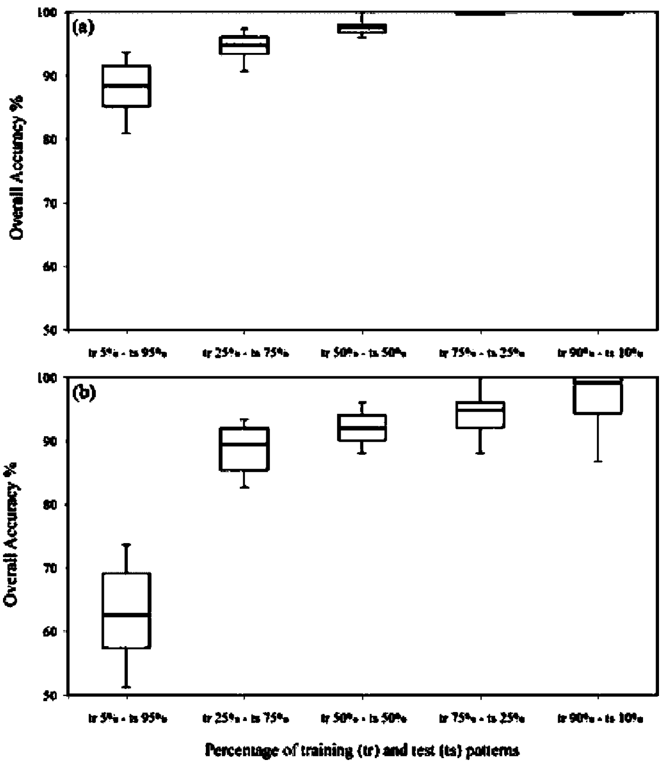

Fig. 5. Percentage overall accuracy (statistics over 1000 realizations) versus different percentages of training and test patterns exhibited by: (a) the SVM classifier, (b) the ML classifier (the seven features selected by the BB algorithm were employed).

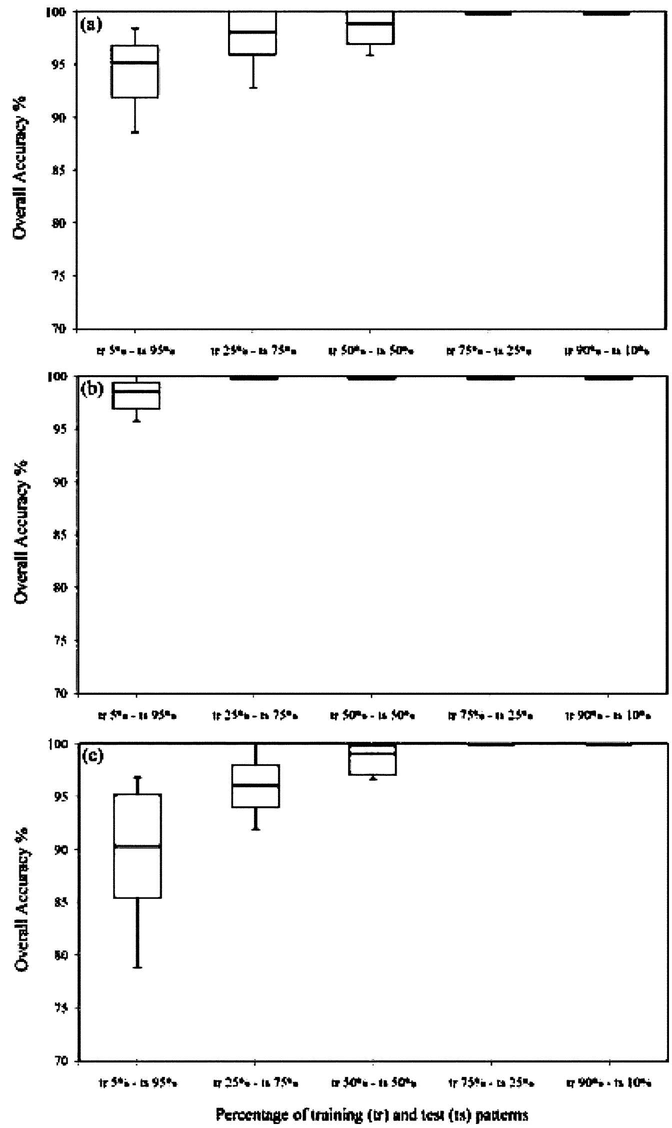

Fig. 6. Percentage overall accuracy (statistics over 1000 realizations) exhibited by the SVM classifier versus different percentages of training and test patterns: (a) AF1 versus AF2, (b) AF1 versus AF3, (c) AF2 versus AF3 (the seven features selected by the BB algorithm were employed).

classifier that, given the small number of training samples available, overfitted the training data when more than seven features were considered (Hughes phenomenon [35]). Overall, the results of the classification analysis are highly satisfactory, since with the optimal subset of features (i.e., RI, NO, CV, DT, PI, BW, and DF) the median accuracy is equal to 97.65%. These results document the high performance of SVM classifiers combined with the feature selection performed by the BB algorithm. The classification accuracies are sharply higher than those reported by the classification schemes previously proposed for the automatic discrimination among different AF types [11], [12], and, as we will show in the next subsection, improved that of the traditional maximum likelihood classifier.

C. Robustness of the System Versus the Number of Available Training Samples

To assess the reliability of the proposed system in critical, but commonly encountered, operative conditions in which only a few labeled patterns are available, in the third set of experiments we tested the robustness of the system versus the size of the training set. In this test, we randomly chose 5, 25, 50, 75, and 90 samples for SVM training, and exploited the remaining test samples according to a grid search method. According to the results reported in the previous section, we considered the optimal subset of features identified by the BB algorithm for $m = 7$.

Fig. 5(a) reports the overall accuracy obtained by increasing the number of available training patterns from 5 to 90 and expressed as median and dispersion over 1000 different realizations of the training set. It is worth noting that the consistency of the reported values should be influenced by the decreasing number of available test patterns. However, the corresponding increase of the number of training patterns should counteract this effect, as the classifier is expected to work better with a larger training set. In fact, the accuracy on the test set improved by increasing the number of training samples, reaching 100% (i.e., the proposed SVM classifier never made any mistake over the ten realizations) when 75 training patterns were employed. In comparison with the technique described in [12], where the same subdivision between training and test sets was exploited reporting an accuracy of 85.5%, a sharp increase in the performance can be observed. To further prove the effectiveness of the proposed SVM classifier, we compared its results with those obtained for the same set of experiments by a Gaussian maximum likelihood (ML) classifier. By comparing Fig. 5(a) and (b), we note that SVM outperforms ML, particularly when few training samples are used. Even in such critical conditions, the SVM system exhibited high stability (e.g., median OA% was 88.3% with only five training samples), thus proving particularly attractive for solving practical application problems. As shown in Fig. 6, this good behavior is observed also for the individual 2-class problems. Even though limit cases on the number of training samples should be carefully evaluated, the development of systems capable to model the classification problem efficiently (with good generalization ability) also with small training sets is recommended for practical applications, such as in our problem where manual classification of AF types can be limited to short data sets as it is subjective and time-consuming.

VII. Conclusion

In this paper, we considered the problem of characterizing the degree of organization of human AF from the analysis of intracardiac electrograms, and proposed an advanced system for the automatic classification of AF types in accordance with traditional clinical criteria [5]. To properly define the automatic classification system, we first studied the individual behavior of the indices commonly adopted in the literature for the automatic quantification of AF organization. This analysis pointed out the complexity of the problem, suggesting implementation of a multivariate approach. Multivariate analysis was carried out by a feature-selection technique identifying the subset of features that better discriminated the investigated AF types. We found that classification accuracy was optimized using a subset of seven features, including indices quantifying morphological variations of the atrial activations and indices detecting perturbations of the isoelectric line of the atrial signals. The selected features were given as input to an automatic classifier based on the SVM technique, which merges important properties such as high generalization capability, easy architecture definition, and learning associated with a convex cost function, with high classification rates.

The main advantages of using the proposed system are: 1) the applicability to short-time windows (4 s) that fits well the needs of clinical applications; 2) the very high accuracy provided by the automatic classification system that allows a precise identification of AF types; and 3) the robustness to small numbers of training samples that encourages the implementation without the need of long training times. These peculiarities suggest the proposed system as a very promising tool for the automatic evaluation of the organization of AF in the clinical practice. Indeed, it fosters the implementation of an automatic, real-time, approach to the classification of AF electrograms that, when exploited as an alternative to the subjective and time-consuming manual analysis, should be far more effective in supporting the catheter-based ablation treatment of paroxysmal AF.

References

[1] P. A. Wolf, R. D. Abbott, and W. B. Kannel, "Atrial fibrillation as an independent risk factor for stroke: The Framingham study," *Stroke*, vol. 22, pp. 983–988, 1991.

[2] F. Gaita, R. Riccardi, and R. Gallotti, "Surgical approaches to atrial fibrillation," *Card. Electrophysiol. Rev.*, vol. 6, pp. 401–405, 2002.

[3] K. Nademanee, J. McKenzie, E. Kosar, M. Schwab, B. Sunsaneewitayakul, T. Vasavakul, C. Khunnawat, and T. Ngarmukos, "A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate," *J. Amer. Coll. Cardiol.*, vol. 43, pp. 2044–2053, 2004.

[4] F. Ravelli, L. Faes, L. Sandrini, F. Gaita, R. Antolini, M. Scaglione, and G. Nollo, "Wave similarity mapping shows the spatiotemporal distribution of fibrillatory wave complexity in the human right atrium during paroxysmal and chronic atrial fibrillation," *J. Cardiovasc. Electrophysiol.*, vol. 16, pp. 1071–1076, 2005.

[5] J. L. Wells, Jr., R. B. Karp, N. T. Kouchoukos, W. A. MacLean, T. N. James, and A. L. Waldo, "Characterization of atrial fibrillation in man: Studies following open heart surgery," *Pacing Clin. Electrophysiol.*, vol. 1, pp. 426–438, 1978.

[6] K. T. Konings, J. L. Smeets, O. C. Penn, H. J. Wellens, and M. A. Allessie, "Configuration of unipolar atrial electrograms during electrically induced atrial fibrillation in humans," *Circulation*, vol. 95, pp. 1231–1241, 1997.

[7] K. T. Konings, C. J. Kirchhof, J. R. Smeets, H. J. Wellens, O. C. Penn, and M. A. Allessie, "High-density mapping of electrically induced atrial fibrillation in humans," *Circulation*, vol. 89, pp. 1665–1680, 1994.

[8] F. Gaita, L. Calo, R. Riccardi, L. Garberoglio, M. Scaglione, G. Licciardello, L. Coda, P. Di Donna, M. Bocchiardo, D. Caponi, R. Antolini, F. Orzan, and G. P. Trevi, "Different patterns of atrial activation in idiopathic atrial fibrillation: Simultaneous multisite atrial mapping in patients with paroxysmal and chronic atrial fibrillation," *J. Amer. Coll. Cardiol.*, vol. 37, pp. 534–541, 2001.

[9] B. Zrenner, G. Ndrepepa, M. R. Karch, M. A. Schneider, J. Schreieck, A. Schomig, and C. Schmitt, "Electrophysiologic characteristics of paroxysmal and chronic atrial fibrillation in human right atrium," *J. Amer. Coll. Cardiol.*, vol. 38, pp. 1143–1149, 2001.

[10] H. Li, J. Hare, K. Mughal, D. Krum, M. Biehl, S. Deshpande, A. Dhala, Z. Blanck, J. Sra, M. Jazayeri, and M. Akhtar, "Distribution of atrial electrogram types during atrial fibrillation: Effect of rapid atrial pacing and intercaval junction ablation," *J. Amer. Coll. Cardiol.*, vol. 27, pp. 1713–1721, 1996.

[11] V. Barbaro, P. Bartolini, G. Calcagnini, S. Morelli, A. Michelucci, and G. Gensini, "Automated classification of human atrial fibrillation from intraatrial electrograms," *Pacing Clin. Electrophysiol.*, vol. 23, pp. 192–202, 2000.

[12] L. Faes, G. Nollo, R. Antolini, F. Gaita, and F. Ravelli, "A method for quantifying atrial fibrillation organization based on wave morphology similarity," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 12, pp. 1504–1513, Dec. 2002.

[13] A. Capucci, M. Biffi, G. Boriani, F. Ravelli, G. Nollo, P. Sabbatani, C. Orsi, and B. Magnani, "Dynamic electrophysiological behavior of human atria during paroxysmal atrial fibrillation," *Circulation*, vol. 92, pp. 1193–1202, 1995.

[14] A. C. Skanes, R. Mandapati, O. Berenfeld, J. M. Davidenko, and J. Jalife, "Spatiotemporal periodicity during atrial fibrillation in the isolated sheep heart," *Circulation*, vol. 98, pp. 1236–1248, 1998.

[15] V. Barbaro, P. Bartolini, G. Calcagnini, F. Censi, S. Morelli, and A. Michelucci, "Mapping the organization of atrial fibrillation with basket catheters. Part I: Validation of a real-time algorithm," *Pacing Clin. Electrophysiol.*, vol. 24, pp. 1082–1088, 2001.

[16] P. M. Narendra and K. Fukunaga, "A branch and bound algorithm for feature subset selection," *IEEE Trans. Comput.*, vol. C-31, no. 9, pp. 917–922, Sep. 1977.

[17] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[18] B. Schölkopf, C. Burges, and A. J. Smola, *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA: MIT Press, 1999.

[19] L. Bruzzone and F. Melgani, "Classification of hyperspectral remote-sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[20] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[21] I. Guyon, J. Weston, S. Barnhill, and V. N. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, pp. 389–422, 2002.

[22] S. Shkurovich, A. V. Sahakian, and S. Swiryn, "Detection of atrial activity from high-voltage leads of implantable ventricular defibrillators using a cancellation technique," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 2, pp. 229–234, Feb. 1998.

[23] L. Faes and F. Ravelli, "A morphology-based approach to the evaluation of atrial fibrillation organization," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 4, pp. 59–67, Jul.–Aug. 2007.

[24] L. Faes, G. Nollo, M. Kirchner, E. Olivetti, F. Gaita, R. Riccardi, and R. Antolini, "Principal component analysis and cluster analysis for

measuring the local organisation of human atrial fibrillation," *Med. Biol. Eng. Comput.*, vol. 39, pp. 656–663, 2001.
[25] P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*. New York: McGraw-Hill, 1978.
[26] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension to multiclass cases of the Jeffreys–Matusita distance," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 6, pp. 1318–1321, Nov. 1995.
[27] L. Bruzzone and S. B. Serpico, "A technique for feature selection in multiclass cases," *Int. J. Remote Sens.*, vol. 21, pp. 549–563, 2000.
[28] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.
[29] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.
[30] V. N. Vapnik, *Estimation of Dependences Based on Empirical Data*. New York: Springer-Verlag, 1982.
[31] T. M. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," *IEEE Trans. Electron. Comput.*, vol. EC-14, no. 3, pp. 326–334, Jun. 1965.
[32] J. Mercer, *Functions of Positive and Negative Type and Their Connection with the Theory of Integral Equations*. London: Philos. Trans. Roy. Soc., 1909.
[33] J. Platt, "Fast training of support vector machines using sequential minimal optimization," in *Advances in Kernel Methods: Support Vector Learning*. B. Schölkopf, C. Burges, and A. Smola. Cambridge, MA: MIT Press, 1998, pp. 185–208.
[34] S. Lazar, S. Dixit, F. E. Marchlinski, D. J. Callans, and E. P. Gerstenfeld, "Presence of left-to-right atrial frequency gradient in paroxysmal but not persistent atrial fibrillation in humans," *Circulation*, vol. 110, pp. 3181–3186, 2004.
[35] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.

Giandomenico Nollo (M'07) was born in 1960. He received the Degree in physics from the University of Trento, Trento, Italy, and the Ph.D. degree in physiology from the University of Torino, Torino, Italy.

He was initially a Research Fellow, and from 1989, was a Researcher in the Division of Medical Biophysics, Institute for Scientific and Technologies Research (ITC-irst).

Since 2000, he has been with the Biophysics and Biosignals Laboratory, Department of Physics, University of Trento, where he teaches biomedical instrumentation and monitoring of biosystems for the Laurea in Physics course of the Faculty of Science. His current research interests include monitoring of human biosignals and linear and nonlinear modeling of cardiovascular data.

Mattia Marconcini (S'06) received the Laurea (B.S.) and Laurea Specialistica (M.S.) (*summa cum laude*) degrees in telecommunication engineering, and the Ph.D. degree in communication and information technologies in 2002, 2004, and 2008, respectively, all from the University of Trento, Trento, Italy.

He is currently with the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. He conducts research on transfer learning and domain adaptation classification and image segmentation problems within the framework of several national and international projects. His current research interests include the area of machine learning, pattern recognition, and remote sensing.

Dr. Marconcini was the recipient of several awards including being a finalist at the Student Prize Paper Competition of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain.

Luca Faes (M'07) was born in Calceranica, Trento, Italy, in 1973. He received the Laurea (M.S.) degree (*summa cum laude*) in electronic engineering from the University of Padova, Padova, Italy, in 1998, and the Ph.D. degree in electronic devices from the University of Trento, Trento, in 2003.

He was a Research Fellow in the Medical Biophysics Division, ITC-irst, Trento, until 2000, where he was engaged in research on system identification and modeling. He is currently with the Biophysics and Biosignals Laboratory, Department of Physics, University of Trento. His current research interests include time series analysis, digital signal processing, and system modeling applied to mechanisms of cardiovascular control and atrial fibrillation.

Francesca Bovolo (S'05–M'07) received the Laurea (B.S.) and Laurea Specialistica (M.S.) (*summa cum laude*) degrees in telecommunication engineering, and the Ph.D. degree in communication and information technologies in 2001, 2003, and 2006, respectively, all from the University of Trento, Trento, Italy.

She is currently with the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. She conducts research on multispectral and synthetic aperture radar images, and very high-resolution images within the framework of several national and international projects. Her current research interests include the area of remote sensing image processing.

Dr. Bovolo has served on the Scientific Committee of the SPIE International Conference on Signal and Image Processing for Remote Sensing since 2006. In 2007, she served on the Scientific Committee of the IEEE Fourth International Workshop on the Analysis of Multi-Temporal Remote Sensing Images (MultiTemp 2007), Leuven, Belgium, July 2007, and of the IEEE International Geoscience and Remote Sensing Symposium 2007 (IGARSS'07), Barcelona, Spain, July 2007. She is a Referee for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, *International Journal of Remote Sensing*, *Photogrammetric Engineering and Remote Sensing Journal, Remote Sensing of Environment Journal*, and *Pattern Recognition Journal*. She was the recipient of numerous awards including the first place in the Student Prize Paper Competition of the 2006 IEEE International Geoscience and Remote Sensing Symposium, Denver, CO.

Flavia Ravelli was born in 1960. She received the Laurea (M.S.) degree (*summa cum laude*) in physics from the University of Trento, Trento, Italy, in 1986, and the Ph.D. in physiology from the University of Torino, Torino, Italy, in 1991.

Between 1992 and 1994, she was a Postodoctoral Researcher in the Department of Anatomy and Physiology, University of Torino, and thereafter, was a Researcher in the Division of Medical Biophysics, Institute for Scientific and Technological Research (ITC-irst), Trento. Since 2000, she has been with the Biophysics and Biosignals Laboratory, Department of Physics, University of Trento. Her current research interests include cardiac electrophysiology and the mechanisms of atrial arrhythmias, especially mechanoelectrical feedback mechanisms and atrial fibrillation.

Lorenzo Bruzzone (S'95–M'98–SM'03) received the Laurea (M.S.) degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

From 1998 to 2000, he was a Postdoctoral Researcher at the University of Genoa. In 2000, he joined the University of Trento, Trento, Italy, where he is currently a Full Professor of Telecommunications and teaches remote sensing, pattern recognition, and electrical communications, and is also the Head of the Remote Sensing Laboratory in the Department of Information and Communication Technology. He conducts and supervises research on analysis of multitemporal data, feature extraction and selection, classification, regression and estimation, data fusion, and machine learning within the framework of several national and international projects. He is an Evaluator of project proposals for many different governments (including the European Commission) and scientific organizations. He is a Referee for many international journals and has served on the Scientific Committees of several international conferences. He is a member of the Managing Committee of the Italian Inter-University Consortium on Telecommunications and a member of the Scientific Committee of the India–Italy Center for Advanced Research. His current research interests include the area of remote sensing image processing and recognition. He is the author or coauthor of 60 scientific publications in referred international journals, more than 120 papers in conference proceedings, and 7 book chapters.

Prof. Bruzzone is a member of the International Association for Pattern Recognition and of the Italian Association for Remote Sensing (AIT). He was the General Chair and Co-Chair of the First and Second IEEE International Workshop on the Analysis of Multi-Temporal Remote-Sensing Images (MultiTemp), and is currently a member of the Permanent Steering Committee of this series of workshops. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. He was recipient of numerous awards including the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium, Seattle, WA, and the Recognition of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Best Reviewers in 1999. He was the Guest Editor of a Special Issue of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING on the subject of the analysis of multitemporal remote-sensing images in November 2003. From 2004 to 2006, he served as an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and is currently an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.