Optimization of ECG Classification by Means of Feature Selection

Tanis Mar*, *Student Member, IEEE*, Sebastian Zaunseder, Juan Pablo Martínez, Mariano Llamedo, and Rüdiger Poll

***Abstract*—This study tackles the ECG classification problem by means of a methodology, which is able to enhance classification performance while simultaneously reducing the computational resources, making it specially adequate for its application in the improvement of ambulatory settings. For this purpose, the sequential forward floating search (SFFS) algorithm is applied with a new criterion function index based on linear discriminants. This criterion has been devised specifically to be a quality indicator in ECG arrhythmia classification. Based on this measure, a comprehensive feature set is analyzed with the SFFS algorithm, and the most suitable subset returned is additionally evaluated with a multilayer perceptron (MLP) to assess the robustness of the model. Aiming at obtaining meaningful estimates of the real-world performance and facilitating comparison with similar studies, the present contribution follows the Association for the Advancement of Medical Instrumentation standard EC57:1998 and the same interpatient division scheme used in several previous studies. Results show that by applying the proposed methods, the performance obtained in similar studies under the same constraints can be exceeded, while keeping the requirements suitable for ambulatory monitoring.**

***Index Terms*—Association for the Advancement of Medical Instrumentation (AAMI), classification, ECG, feature selection (FS), linear discriminant analysis (LDA), multilayer perceptron (MLP), sequential forward floating search (SFFS).**

I. Introduction

Automated classification provides inestimable aid for long-term electrocardiography, which is a commonplace in patient monitoring, both in bedside and in ambulatory settings. Indeed, a large number of approaches using a variety of techniques have been proposed for this task, easing the diagnosis of arrhythmic changes as well as further inspection, e.g., heart rate variability or heart rate turbulence analysis. However, due to the huge amount of data and/or the need for online classification present in these situations, additional requirements arise concerning the complexity of classification algorithms. Especially during online ambulatory monitoring, the computational resources are very limited in order to fulfill the requirement of low energy consumption to enhance running times [1], [2]. Therefore, suited methods are required to enable high performance classification even in these unfavorable environments.

Simple classifiers, such as linear discriminants (LD) [3]–[7] or *K*-nearest neighbor classifiers [8], [9], have been used succesfully up to our days. Other researchers have made use of more complex classifiers, based on techniques such as fractal analysis [10], chaotic modeling [11], bispectral coherence analysis [12], or artificial neural networks (ANN). The latter technique deserves to be mentioned aside due to the large number of studies that make use of it. Multilayer perceptrons (MLPs) are the most popular family of ANN within ECG classification [13]–[18], but other ANN paradigms, such as fuzzy neural networks [14], [19]–[25], radial basis networks [26], [27], support vector machines [28], [29], and self-organizing maps [13], [25], [30], [31], have been also widely applied.

Likewise, a large number of possibilities have been proposed about which features to use to describe the ECG. Apart from the RR interval, which is used in most studies, almost every single published paper proposes a new set of features to be used, or a new combination of the existing ones. Among others, morphological features extracted directly from the ECG like amplitudes and peak widths [9], [32], features based on different transforms, e.g., wavelet transform (WT) [15], [33], [34], principal component analysis [26], [35] and Hermite functions [23], [29], [31], as well as statistical features, e.g., variances [14], [36], have been proposed.

Due to the large number of available features, some authors have already availed themselves of feature selection (FS) methods to reduce the dimensionality of the classification problem [6], [7], [15], [23], [29], [37], [38]. These methods involve a process wherein a number of subsets of the available features are evaluated, and the best one is selected for application on the learning algorithm. The best subset contains the least number of dimensions that most contribute to the application's performance; the remaining, unimportant dimensions, are discarded. However, none of the aforementioned studies has thoroughly

Manuscript received July 27, 2010; revised November 13, 2010; accepted January 12, 2011. Date of publication February 10, 2011; date of current version July 20, 2011. This work was supported in part by theFraunhofer IPMS in collaboration with the Institut für Biomedizinische Technik, at the Dresden University of Technology, Dresden, Germany, and in part by the MCINN (Spain) under Project TEC2010-21703-C03-02, and in part by the CIBER-BBN, an initiative of the Instituto de Salud Carlos III (Spain). *Asterisk indicates corresponding author.*

*T. Mar is with the Wireless Microsystems Department, Fraunhofer IPMS, Dresden, PLZ 01109 Germany, in collaboration with the Dresden University of Technology, Dresden, PLZ 01069 Germany (e-mail: tanismar@hotmail.com).

S. Zaunseder is with the Wireless Microsystems Department, Fraunhofer IPMS, Dresden, PLZ 01109 Germany, in collaboration with the Dresden University of Technology, Dresden, PLZ 01069 Germany (e-mail: sebastian.zaunseder@ipms.fraunhofer.de).

J. P. Martínez is with the Communications Technology Group, Aragon Institute for Engineering Research (I3A), University of Zaragoza, Zaragoza 50018, Aragón, Spain, and also with the CIBER de Bioengeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza 50018, Spain (e-mail: jpmart@unizar.es).

M. Llamedo is with Aragon Institute for Engineering Research, University of Zaragoza, Zaragoza, Aragon, Spain (e-mail: llamedom@unizar.es).

R. Poll is with the Institut für Biomedizinische Technik, Dresden University of Technology, Dresden, PLZ 01069 Germany (e-mail: ruediger.poll@tu-dresden.de).

Digital Object Identifier 10.1109/TBME.2011.2113395

tackled the issue within the ECG classification field by using a comprehensive feature set and a complex FS procedure.

Finding the right feature combinations is, indeed, a hard task. The computational requirements of exhaustive search methods (those which test all possible subsets) increase exponentially with the number of features in the original set. This effect, called the *curse of dimensionality* [39], makes this kind of methods impracticable for sets with over a dozen or so features. Suboptimal methods have, thus, to be used. From the many available, the sequential forward floating search (SFFS) algorithm, proposed by Pudil [40], has been found to outperform both sequential and fixed parameter FS algorithms [41], [42], finding in most occasions solutions very close to the optimal one thanks to the automatic control of the search method.

The rest of this paper is organized as follows: Section II introduces the chosen database together with the use that it was given throughout in this study. Also, it presents results of previously published studies on the topic, with a brief analysis of the performance measures used for these results. In Section III, the feature sets used in this study are described, and the method used for the FS process, as well as the applied criterion, is explained. Section IV deals with the classifier models applied. In section V, results are exposed, leading to the discussion in Section VI. Finally, a conclusion is drawn in Section VII.

II. Developing and Comparing Classifiers

Even considering that most approaches described in the literature use the same source for ECG recordings, there are several factors that render the comparison of these studies' results almost impossible. Moreover, the obtention of a realistic estimation about the real-world performance of the algorithms must also be regarded. In what follows, these problems are addressed and good practices to cope with them, which consequently have been also used in this study, are outlined.

A. Data Preparation

In this study, all the ECG data used have been obtained from the MIT-BIH Arrhythmia Database (MITDB) [43]. Originally, over 109 000 beats that the database contains are individually labeled as belonging to one of 15 possible beat types. However, in order to foster common proceedings, the Association for the Advancement of Medical Instrumentation (AAMI) proposed a standard for the evaluation of ECG classifiers [44], which recommends to group all present morphologies in six classes according to their physiological origin (see Table I). From these six classes, the standard recommends to ignore records containing paced beats when evaluating classifiers, and some authors have proposed to ignore unknown beats too, for being too poorly represented and of no help for further classification purposes [6], [7], [28], [29].

Applied to the MITDB, this standard leaves 44 records which should be divided using an interpatient division scheme such as the one proposed in [5] if a realistic estimation of the real-world performance is desired. Thus, following this scheme, we devoted 22 records (DS1) of the MITDB for development and training, while the other 22 records (DS2) are used only for final performance evaluation so they have no influence on the selection of the final classifier model (see Table II).

TABLE I
AAMI Grouping Scheme

Heartbeat Group Name	Morphologies Included	% of the Total	Number of Patterns
Normal (N)	N, e, j, L, R	82,78	90631
Ventricular Ectopic (V)	V, E	6,60	7236
Supraventricular Ectopic (S)	A, a, J, S	2,54	2781
Fusion (F)	F	0,73	803
Paced (P)	/, f	7,32	8010
Unknown	U	0,03	33
Total		100	109494

TABLE II
MITDB Division Scheme for Interpatient Classification

Inter-patient datasets	Number of beats per AAMI class N	S	V	F	Total
DS1	45662	943	3778	413	50796
DS2	44061	1830	3208	388	49487
Total (DS1 + DS2)	89723	2773	6986	801	100283

Records 102, 104, 107 and 217, which contain paced beats, are not included in DS1 nor in DS2.

DS1 includes records 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 and 230.

DS2 includes records 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233 and 234.

To conform DS1 and DS2, information from both ECG leads present in the MITDB was considered throughout in this study. Prior to feature acquisition, the same preprocessing as in [5] was applied: first, two median and a low-pass filter were used in order to remove noise and baseline wander. Afterward, QRS fiducial points were read from the MITDB annotation files. Further, heartbeat segmentation was achieved by executing *ecgpuwave*[1] separately on each ECG channel.

B. Performance Measures

The confusion matrix provides a complete description of any classification results. However, results are usually displayed using indices or factors that describe specific aspects of the classification. In the ECG classification field, those studies which follow the AAMI standard and interpatient division scheme have preferred indices such as the multiway accuracy [5], [7], [29], the *j* index [6], or the unweighted mean of sensitivities [28], [29] to summarize their results on a single value and thus facilitate the selection of the best performing model. A mathematical description of these indices is displayed in Table III.

However, due to the large imbalance found between the number of beats in each class, multiway accuracy was found not to be the most representative index for the overall quality of the classification. The number of true positive N beats, which indeed says little about the detection of most important arrhythmias (V and S beats, according to the AAMI standard), dominates the final value of the accuracy index. In fact, if all beats were labeled as normal, multiway accuracy would still retain a high value of over 89%. The *j* index, sum of the sensitivities and positive predictivities of V and S classes, is much more appropriate

[1] *ecgpuwave*: found at http://www.physionet.org/physiotools/softwareindex.shtml.

TABLE III
DESCRIPTION OF PREFERRED PERFORMANCE MEASURES IN HEARTBEAT CLASSIFICATION

Confusion Matrix

Reference Label \ Algorithm Label	**n**	**s**	**v**	**f**	$\sum$
N	X_{NN}	X_{NS}	X_{NV}	X_{NF}	O_N
S	X_{SN}	X_{SS}	X_{SV}	X_{SF}	O_S
V	X_{VN}	X_{VS}	X_{VV}	X_{VF}	O_V
F	X_{FN}	X_{FS}	X_{FV}	X_{FF}	O_F
$\sum$	A_N	A_S	A_V	A_F	X_{Tot}

Grouping Symbols

C is the number of classes.
be $I, J \in \{N, S, V, F\}$
X_{IJ} is the number of examples of class I classified as J.
thus, $\forall I \neq J$, X_{IJ} are wrong classified examples.
and, $\forall I$, X_{II} are correctly classified examples.
$O_I = \sum_{\forall J} X_{IJ}$, Total number of examples originally belonging to class I.
$A_J = \sum_{\forall I} X_{IJ}$, Total number of examples labeled by algorithm as class J.
$X_{Tot} = \sum_{\forall I \forall J} X_{IJ} = \sum_{\forall I} O_I = \sum_{\forall J} A_J$

Relevant Factors

True Positives of I (TP_I) $= X_{II}$
True Negatives of I (TN_I) $= X_{Tot} - O_I - A_I + X_{II}$
False Positive of I (FP_I) $= A_I - X_{II}$
False Negative of I (FN_I) $= O_I - X_{II}$

Performance Measures

Accuracy of I (Acc_I) = (TP_I + TN_I)/(TP_I + TN_I + FP_I + FN_I)
Sensitivity of I (Sens_I) = TP_I/O_I
Positive Predictive Value of I (PPV_I) =TP_I/A_I
Multiway Accuracy $= \sum_{\forall I} \mathrm{TP}_I / X_{Tot}$
***j* index** =Sens_V + Sens_S + PPV_V + PPV_S
Unweighted Mean Sensitivity $= \sum_{\forall I} \mathrm{Sens}_I / C$

to assess the effectiveness in discriminating ECG arrhythmias. This value, however, ignores how well F and N beats have been classified. The unweighted mean of sensitivities is computed giving the same importance to each of the classes. This not only accounts for the class size imbalance, but also makes the relative influence of F beats much bigger than for any other class, being its identification not as critical as that of S or V beat.

In order to avoid the aforementioned drawbacks, in this study we proposed an alternative performance measure, which is described in Section III-B.

C. Results of Recommendation Conform Studies

Until present time, most studies apply training and evaluation schemes that are not conform to the methodology described earlier, impeding the objective comparison of the achieved results. Fortunately, since the publication of the AAMI standard, the number of works in which the recommended practices are followed has been increasing steadily. Moreover, a small number of authors have also employed an interpatient division scheme in their studies in the last few years. The results obtained by this last group of authors, which are, thus, directly comparable, can be found in Table IV.

III. FEATURES

A. Feature Sets

The larger the number of features in a feature set, the most likely it is that for any given task an optimal subset exists among them. Therefore, aiming at having greater chances of finding the best subset of features for classification, a very comprehensive feature set was desired. However, if all features found in the literature were introduced, the number will be indeed too large, rendering a complete FS to be computationally unfeasible, so a constraint of some type was necessary to keep the number within reasonable limits. Having in mind the aim of this study to be most useful for ambulatory classification, the constraint was to use only causal features, which do not have to wait until the end of the recording to be computed. Hence, only those features that used information available at the moment of analysis were used, admitting at most information from the next beat to be considered for classification.

The resulting set, named C, contains a total of 71 features, divided (as shown in Table V) into the following categories.

1) Temporal features, which already proved in different studies to be the most relevant [29]. This category includes heart rate features, which were the only features computed just once for both channels, and features obtained from the segmentation information yielded by *ecgpuwave*.
2) Morphological features, also previously assessed as being of great relevance, made the bulk of the feature set. Direct samples from the ECG signal and computed measurements such as area, power, or extrema were included.
3) Statistical features completed the feature set, including different order moment-based indexes and histogram variance.

Unlike in temporal features, which were acquired from time-domain signals exclusively, the WT of the ECG signal was used to obtain some of the features belonging to the morphological and statistical categories. WT features were based on the same heartbeat intervals as the features obtained from the time domain, but using the scales 2, 3, 4, and 5 of the WT ECG signal. The wavelet used was the quadratic spline wavelet, as it allows a very efficient implementation as a filter bank, and its suitability for ECG processing has already been proved by other authors [2], [6], [7], [46].

Additionally, to enable comparative evaluation of the proposed methods, an already proposed feature set was built up. De Chazal *et al.*'s realization [5] was chosen as it was the first one to propose interpatient dataset distribution scheme, and hence a referent for all posterior studies using this scheme. This feature set consists of RR intervals, heartbeat intervals, and morphology features, obtained from both ECG channels after the preprocessing and segmentation procedures described before. However, in order to comply with the same constraint as feature set C, the original set was slightly modified to contain only causal features. The resulting dataset, containing 26 features, was named H.

Finally, all features both in C and in H were individually normalized, by computing the necessary scaling to make the features from DS1 signals be mean 0 and variance 1, and normalizing the corresponding features from DS2 with the obtained scaling factors.

TABLE IV
RESULTS OF AAMI CONFORM, INTERPATIENT TRAINED CLASSIFIERS, EVALUATED WITH DS2

Classifier		Normal			Supraventricular			Ventricular			Fusion		Multiway	Unweighted	*j*	κ
Paradigm	**Acc**	**Sens**	**PPV**	**Acc**	**Sens**	**PPV**	**Acc**	**Sens**	**PPV**	**Acc**	**Sens**	**PPV**	**Accuracy**	**Mean. Sens**	**index**	**index**
LD [5]	87.8	87.0	99.2	94.6	75.9	38.5	97.4	77.7	81.9	92.4	89.3	8.6	86.2	83.2	2.764	0.532
LD* [45]	94.5	94.4	99.4	95.9	87.7	47.0	99.4	94.3	96.2	98.4	74.0	29.1	94.0	87.7	3.234	0.754
LDA [6]	81.9	80.1	99.4	86.4	86.1	19.6	95.2	70.6	60.9	96.0	81.7	14.3	79.8	79.7	2.372	0.421
SVM [28]	-	86.2	-	-	82.6	-	-	80.9	-	-	54.9	-	85.6	76.2	-	-
SVM [29]	-	75.1	-	-	89.3	-	-	86.9	-	-	80.7	-	-	83.0	-	-
LDA** [7]	79.7	77.6	99.5	95.1	76.5	41.3	98.2	82.9	88.0	83.1	95.4	4.2	78.0	83.1	2.887	0.412

Values computed from the confusion matrices given in the respective publications. Where confusion matrix were not given, some values could not be computed (marked with '-').
* Patient adapting: Requires expert intervention.
** Feature set optimized for classification of N, S and V' classes, where V' class included V and F beats.

TABLE V
FEATURES IN C, DISTRIBUTED BY CATEGORIES

Features	Time Domain	WT Domain*
Temporal	Previous RR, current RR, average RR, average RR of the last 10 beats, QRS duration, T-Wave duration, P-Wave flag.	
Morphological	Downsampled (10 samples) QRS, downsampled (9 samples) T-Wave, QRS area, QRS power, QRS max, QRS min, QRS Max-Min ratio, peak width at 70% Max, peak slope, beat area, beat power, beat max, beat min, beat Max-Min ratio.	Max(3,4,5), Min(3,4,5), difference between Max and Min (3,4,5), distance (in samples) between Max and Min (3,4), power(2,3,4,5), power ratio(3-2,4-3,5-4).
Statistical	QRS variance, QRS skewness, QRS kurtosis, QRS histogram (20 slots) variance, beat mean, beat variance, beat skewness, beat kurtosis, beat histogram (20 slots) variance.	Mean (3,4), standard deviation(3,4), skewness (3,4).

* Numbers between parenthesis represent the scales on which the feature was obtained.

B. FS Procedure

The SFFS procedure, mathematically described in [40], can be explained as follows.

Let X be the original set, of size $|X| = n$ features, and $J(X_k)$ the evaluation measure to be maximized, where $J()$ represents the criterion function to be used (defined as $J : X_k \subseteq X \to \mathbb{R}$), and X_k the subset of size $k < n$ to which it is applied. Let X_k be, on any given point, the subset that has higher $J(X_k)$ from all the subsets of size k evaluated up to that point. Then, on any k except for $k = n$, a forward step is accomplished by evaluating $J(X_k + x)$ for all possible x, x being any of the features not already included in X_k, and keeping as the best subset of size $k + 1$ features the subset that maximizes $J(X_k + x)$, named as X_{k+1}. Afterward, backtracking is executed by evaluating those subsets X'_k of size k that are obtained by removing any of the $k + 1$ features from subset X_{k+1} one at a time. If for any of these subsets $J(X'_k) > J(X_k)$, then X'_k is kept as the new best subset of size k, substituting X_k. Backtracking steps are carried out as long as the results improve, reducing each step the size of the evaluated subsets by one feature. Whenever the results stop improving through backtracking, a forward step is carried out again, and the whole process is repeated again until a better subset at any possible level is found.

As a result, the SFFS procedure returns n subsets; each of which yields the highest criterion function value for its size. These subsets are, henceforth, named *SFFS subsets*.

C. Optimization Criterion Function

Unless intrinsically embedded in the system (like weights in some ANN), the FS process can follow two schemes: filter and wrapper [47]. The criterion function $J()$ used in filter schemes to evaluate the candidate subsets is independent of the classifier model, being typically simplified probabilistic measures. Wrapper schemes, on the contrary, use some kind of performance measure based on the results of testing the candidate subsets with the classifier itself. Although wrapper schemes tend to be much more computationally expensive than filters, they take into account the structural characteristics imposed by the classifier [40], and thus can be easily tailored to the specific problem to solve. Therefore, a wrapper scheme based on the linear discriminant classifier (LDC) was applied in this study (further described in Section IV-A).

Whatever be the method used to obtain the confusion matrix, an evaluation by means of a performance measure is necessary, so that subsets' performances can be compared. As described in Section II-B, none of the available performance measures fully represents the quality of the classification in the ECG analysis context. Therefore, we propose a new performance measure that tackles specifically the problem of providing in a single value complete information about how *good* an ECG classification has been. To this end, this new index was chosen to be a combination of two values: the *j* index, which specifically evaluates the discrimination of the most important ECG arrhythmias (S and V beats), and the *Kappa* (κ) index, which globally evaluates the confusion matrix [48].

This index, despite having been proposed as an evaluation coefficient several decades ago, and its potential convenience had never been applied before in the heartbeat classification. From its definition,

$$\kappa = \frac{\sum_{\forall I} TP_I - \sum_{\forall I} \bar{D}_I}{X_{T}\mathrm{ot} - \sum_{\forall I} \bar{D}_I} \qquad (1)$$

where

$$\bar{D}_I = \text{Ponderated Detections} = (O_I A_I) / X_{T\mathrm{ot}} \qquad (2)$$

it can be seen that it evaluates the global quality of the classification much better than other aforementioned performance measures: like the multiway accuracy, it represents also a complete evaluation of the confusion matrix (in a single value and weighting each beat equally), but it is much less influenced by imbalance.

The resulting combined index, which we named $j\kappa$ index $I_{j\kappa}$, takes into account the misclassification and the imbalance present between all the considered classes, thanks to the included κ index, and at the same time emphasizes the discrimination of the most important arrhythmias (S and V), thanks to the *j* index (I_j)

$$I_{j\kappa} = w_1\kappa + w_2 I_j. \tag{3}$$

As *j* takes values in the 0–4 range and κ in the 0–1 range, w_1 was set to 1/2 and w_2 to 1/8, so that both factors influence equally the overall result. Consequently, $I_{j\kappa}$ takes values between 0 and 1, where 1 indicates perfect classification.

IV. Classifier Models

A. LDC

As mentioned earlier, a serious hurdle for wrapper methods, which apply a direct classifier performance measure as a criterion function, is that many classifiers require a considerable amount of time to be trained and evaluated. This makes a complete FS procedure, where maybe thousands of different subsets have to be tested upon the classifier, unfeasible in a reasonable time period.

Nevertheless, the LDC has not only been already applied successfully in heartbeat classification [4]–[7], but also its simplicity enables the training and evaluation process to be completed very quickly, rendering the whole FS procedure feasible in an acceptable period of time, and the classifier suitable for eventual implementation in ambulatory low-power devices. Thus, LDC has been applied in this study as the classifier upon which the confusion matrix $I_{j\kappa}$ was obtained to guide the SFFS.

As in [5], a slight modification was introduced, weighting the likelihood function so that all classes influenced equally the classifier's train process, in spite of the difference in the number of examples between them. Also, our tests showed that the best results were achieved when the prior probability was set equal for all the considered classes.

B. MLP

The MLP belongs to the class of supervised learning networks, in which the discriminative power is gained through a preliminary learning phase, where labeled examples are presented to the network. The most common training strategies, also used in this study, is the *backpropagation* (BP) algorithm [49]. It works by computing the error between the returned and the known, desired output, employing it to adjust the MLP weights. Although the training process requires a rather long time, the implementation and execution of a trained MLP are very simple, making this paradigm very much suited for classification on ambulatory settings. On the other hand, its characteristics make this paradigm very inadequate to guide the FS process. The random initialization makes MLPs' results not constant, which renders the FS procedure unreliable if only one MLP is evaluated for each tested subset. The unreliability could be overcome by training many MLPs for each tested subset, and performing statistical analysis to obtain a result that would lead to the next step in the FS process. Unfortunately, due to the many subsets tested by the SFFS procedure, plus the relatively long time that training each MLP requires, the time and resources that a reliable MLP–SFFS procedure would take are beyond our capabilities.

TABLE VI
DS1 Division Scheme for MLP Evaluation

MLP evaluation datasets	Number of beats per AAMI class N	S	V	F	Total
Eval1 DS1	23379	442	1936	40	25797
Eval2 DS1	22283	501	1842	373	24999
Total (DS1)	45662	943	3778	413	50796

Eval1 DS1 comprises data from the recordings 109, 114, 118, 119, 124, 201, 203, 205, 215, 220 and 223.
Eval2 DS1 comprises data from the recordings 101, 106, 108, 112, 115, 116, 122, 207, 208, 209 and 230.

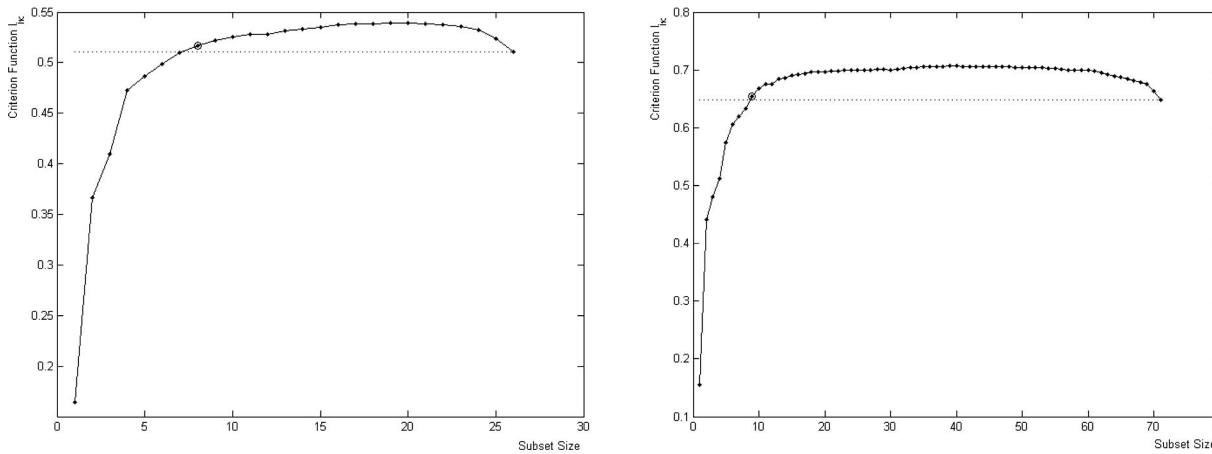
Therefore, in this study, the MLP paradigm was only applied to classify ECG arrhythmias with the best suited *SFFS subsets* both from C and from H datasets. The values of the different parameters governing the MLP were determined by applying twofold cross-validation on DS1, training with onefold the MLP parametrized with the desired combination, and evaluating with the remaining one, and vice versa, averaging the results. Again, this subdivision was performed interpatiently in such a way that all heartbeat classes were similarly represented in each of the folds, as shown in Table VI. MLPs with a single hidden layer of 25 neurons were used, and trained with a learn rate of 0.25 and a momentum of 0.03 to avoid getting stuck into local minima. The number of training cycles was chosen to be the one for which the mean results from the twofold evaluation began to get worse, i.e., when symptoms of overlearning appeared.

C. Classifier Combination

As mentioned earlier, information from both ECG leads was considered throughout the whole study. Except for the heart rate ones, all features were obtained separately for each channel, and the two resulting feature sets applied independently to perform classification. The posterior probabilities obtained after classification with each feature set were then combined using the Bayesian product integration scheme [5] and finally, each heartbeat was labeled with the class with higher posterior probability after the combination.

V. Results

The described LD-based SFFS procedure was executed both on H and on C feature sets. Interpatient division scheme was applied in order to avoid positively biased results, conducting the whole procedure using signals just from DS1, and reserving the data in DS2 for posterior evaluation. The SFFS algorithm run on H dataset tested around 1200 possible feature combinations, while when analyzing the C dataset, over 39 500 combinations were evaluated. Both quantities are much larger than those reported in previous FS attempts for ECG classification, but still



(a) (b)

Fig. 1. Performance achieved on DS1 data by LDC for each of the subsets in (a) *SFFS subsets*$_H$ and (b) *SFFS subsets*$_C$. The dashed line represents the performance achieved with the complete feature set in each case. The circles point out the smaller *SFFS subsets* with which these performances are exceeded: H_{SFFS} and C_{SFFS}, respectively.

small if compared to the total number of possible combinations, 2^{26} and 2^{71}, respectively.

As mentioned earlier, the SFFS algorithm does not yield a single result, but rather a collection of them for each of the feature sets where it is applied, *SFFS subsets*$_H$ and *SFFS subsets*$_C$ in this study. In order to assess the suitability of this feature reduction procedure for our data, a single LDC was retrained and evaluated with DS1 for each of the *SFFS subsets*, separately on H and on C. The results, displayed in Fig. 1, exhibit a common drawback of applying FS procedure on comprehensive sets: in many occasions, the subset with the highest criterion value has still a very large number of features. Therefore, in this study, aiming at reducing the classifier complexity as much as possible, but without making its performance worse, we selected as the most suited for our purposes the classifier with the smaller number of features that achieved at least the performance obtained with the original feature set on DS1.

The smaller subset accomplishing this criterion contains eight features in *SFFS subsets*$_H$, and nine in *SFFS subsets*$_C$. These subsets, which we named H_{SFFS} and C_{SFFS}, respectively, comprise the following features.

1) H_{SFFS}: Previous RR, current RR, RR average, downsampled QRS (samples 2, 5, and 8), QRS duration, downsampled T-wave (sample 9).
2) C_{SFFS}: Previous RR, current RR, RR average, beat min, beat max, QRS max–min ratio, peak slope, max–min difference on WT scale 3.

After evaluating the performance of the SFFS procedure with the matched classifier (LD) on DS1, the original feature subsets (H and C) and the most suited ones (H_{SFFS} and C_{SFFS}) were tested on DS2 to carry out the final evaluation of the classifier model. Additionally, these subsets were also tested on DS2 with the MLP classifier, in order to analyze the generalizing capability of the FS procedure in the case where the criterion function and the classifier paradigm do not match. At the same time, this analysis also tackles the suitability of the MLP for heartbeat classification, in direct comparison to the LD classifier. In

TABLE VII
CONFUSION MATRICES FOR C_{SFFS} ON (a) LDA CLASSIFICATION PARADIGM AND (b) MLP CLASSIFICATION PARADIGM

(a)	*LDA*	Algorithm			
		n	**s**	**v**	**f**
Reference	**N**	37384	2726	691	3260
	S	60	1517	237	16
	V	45	225	2782	156
	F	137	1	50	200

(b)	*MLP*	Algorithm			
		n	**s**	**v**	**f**
Reference	**N**	39497	2778	771	1015
	S	122	1523	93	92
	V	104	235	2783	86
	F	125	6	20	237

Table VII, complete classification description is displayed in the form of the confusion matrices for the results obtained by applying the C_{SFFS} feature set with either classifier paradigm. These matrices provide insight into the beat-by-beat performance and ease future comparison attempts by other authors. For the rest of studied feature sets, results for both LDA and MLP classifications are given through the considered performance measures in Table VIII.

VI. DISCUSSION

A. ECG Feature Optimization

The results of the FS procedure displayed in Fig. 1 show that in both evaluated feature sets, being able to find the most adequate combinations of features, allow us to improve the performance even when classification is carried out with a fraction of the original number of features. The original performance is already exceeded with 8 out of 26 features in the H feature set, and with 9 out of 71 in the C feature set. The capability of the

TABLE VIII
CLASSIFIER PERFORMANCES ON DS2 OBTAINED FOR THE MOST RELEVANT STUDIED CLASSIFIER MODELS

	Feature Set	Normal Acc	Normal Sens	Normal PPV	Supraventricular Acc	Supraventricular Sens	Supraventricular PPV	Ventricular Acc	Ventricular Sens	Ventricular PPV	Fusion Acc	Fusion Sens	Fusion PPV	MW Acc.	Unw. M. Sens.	j index	κ index	$j\kappa$ index
LDA	H	85.7	85.1	98.7	91.4	58.1	23.3	94.6	71.1	56.6	94.3	61.6	8.2	83.0	69.0	2.092	0.447	0.485
	H_{SFFS}	87.4	86.4	99.4	92.6	70.8	29.2	94.0	75.6	53.0	95.8	63.9	11.4	85.0	74.2	2.287	0.500	0.536
	C	85.4	85.7	97.6	93.4	76.6	35.0	95.4	79.0	61.9	94.1	22.1	3.2	84.4	65.9	2.525	0.472	0.552
	C_{SFFS}	86.0	84.8	99.3	93.4	82.9	33.9	97.2	86.7	74.0	92.7	51.6	5.5	84.6	76.5	2.776	0.511	0.602
MLP	H	80.3	78.5	99.2	93.0	78.5	32.0	90.3	83.9	38.6	94.0	74.0	9.1	78.8	78.8	2.332	0.411	0.497
	H_{SFFS}	86.5	85.6	99.0	94.1	80.4	36.7	96.8	83.1	71.9	92.6	49.2	5.2	85.0	74.6	2.722	0.510	0.595
	C	79.3	77.3	99.3	93.2	76.3	32.1	94.2	90.1	53.0	89.2	53.9	3.9	77.9	74.4	2.515	0.405	0.517
	C_{SFFS}	90.0	89.6	99.1	93.3	83.2	33.5	97.3	86.8	75.9	97.3	61.1	16.6	89.0	80.2	2.794	0.599	0.649

TABLE IX
RELEVANT INDICES FOR AAMI STANDARD AND INTERPATIENT DIVISION CONFORM STUDIES, INCLUDING PRESENT STUDY'S ONES

Classifier Paradigm	Multiway Accuracy	Unweighted Mean. Sens	j index	κ index	$j\kappa$ index
LDA (C_{SFFS})	84.6	76.5	2.776	0.511	0.602
MLP (C_{SFFS})	89.0	80.2	2.794	0.599	0.649
LD [5]	86.2	83.2	2.764	0.532	0.612
LD* [45]	94.0	87.7	3.234	0.754	0.781
LDA [6]	79.8	79.7	2.372	0.421	0.507
SVM [28]	85.6	76.2	-	-	-
SVM [29]	-	83.0	-	-	-
LDA** [7]	78.0	83.1	2.887	0.412	0.567

*Patient adapting: Requires expert intervention.

**Feature set optimized for classification of N, S and V' classes, where V' class included V and F beats.

method to successfully narrow down large feature sets is, thus, demonstrated.

It is also worth noticing that although the performance of the reduced feature sets (H_{SFFS} and C_{SFFS}) is just over the one achieved with the complete feature sets (H and C) on the evaluation with DS1, the former sets exhibit much higher performance in the final test with DS2, as shown in Table VIII. This is in accordance with the well-known fact that models with a smaller number of features tend to generalize much better than those with many features, as they are less conditioned by the distribution of the training data. Moreover, the fact that the performance achieved with the reduced feature sets also exceeds the one obtained with the complete ones when using the MLP to classify shows that the improvement obtained through FS does not only hold for the classifier used as criterion, but for other classifier paradigms too.

In addition to the analysis of the FS process itself, it is also interesting to identify the most relevant features. Inspecting H_{SFFS} and C_{SFFS} subsets, we can observe that in both cases the *previous RR*, *current RR*, and *RR average* features were present. This indicates the uttermost importance that heart rate features have on ECG classification. Looking at the results from H separately, and even considering that the isolated samples make no sense as individual features in a complete classifier model, the fact that four out of the five morphology features in H_{SFFS} belong to the QRS, and just one to the T-wave, reveals the greater importance for ECG classification that the QRS morphology has over the T-wave morphology.

Regarding the results for the C feature set, which contained a much larger number of features, including many statistical ones, it is remarkable to observe that the six nonheart rate features that complete C_{SFFS} are all morphological features too, and, even more noteworthy, that five of them represent extrema. This fact suggests that extrema value posses the highest discriminative power among all morphological features. Also worth noticing is the fact that just one of the features is a WT value, pointing out that although features obtained through this transformation allow a more robust detection of heartbeat segments, they add little information to the one provided by the untransformed temporal ECG features, at least for heartbeat classification purposes.

These results reinforce, upon a more comprehensive analysis, the outcomes of previous FS experiments conducted on ECG [5], [29], where temporal and morphological features were found out to be the most adequate ones to perform ECG classification. Moreover, they further detail the characteristics of the most relevant features, beyond their category.

B. Classification Performance

Table VIII displays a detailed description of the performances achieved by the most relevant classifier models studied. As already stated in the previous section, the comparison between the results obtained with the complete feature sets and the ones obtained with their respective reduced feature sets confirms that applying the SFFS procedure with a suitable criterion function leads to an improvement in the final performances. Although some performance descriptors are not very conclusive (considering the LD classifier, the SFFS procedure enabled an increase of the *unweighted mean sensitivity*, while the *multiway accuracy* practically retained its value, but for the MLP classifiers it happens almost the opposite), the κ and j indices, and hence the $I_{j\kappa}$ index achieved improve with the feature subsets returned from the FS procedure, for both original feature sets and both classifier paradigms.

In spite of the large number of studies in which the MLP classifier paradigm has been applied for ECG classification, none among them could be found in which the results were evaluated in conformance with the AAMI standard and interpatient dataset distribution. Yet, results show that, when applying reduced feature sets, the MLP can clearly outperform LD in the task of heartbeat classification. Comparing the results of both paradigms when working with C_{SFFS}, an improvement in the range of 4% can be observed in the *multiway accuracy* and in the *unweighted mean sensitibity*, and some points are gained in the other indices. Nevertheless, it should be noted that these numbers are just orientative of the possible improvement, as, due to their random initialization, successive evaluations of the MLP with the same feature set could yield different results.

C. Comparison With Previous Studies

Comparing the achieved performance results with those of previous studies provides further insight on the suitability of the proposed techniques. As mentioned, this comparison can only be objectively done among those studies following the same constraints. Thus, the performances obtained in this study on both classifier paradigms with the C_{SFFS} feature set have been only compared with the results of the other AAMI conform studies that followed the interpatient division scheme (see Table IX).

First of all, by comparing the results published in [5] with the ones obtained for the H dataset using LDC (see Tables IV and VIII), it can be observed that the former are clearly better, although both measure the same features applied on the same classifier. The difference in performance is likely to be due to the modification carried out to adapt the features in [5] to the causality criteria described in Section III-B. That is probably the reason why, even after applying the aforementioned optimization techniques, the results obtained with LD do not represent a clear improvement with respect to previous studies.

However, the ones achieved with the MLP outperform all previous nonadapting proposed methods. These results evince that the nonlinear classification capabilities of this type of ANN are extremely suitable to perform heartbeat classification, which is intrinsically nonlinear too. Moreover, they also show, in the context of this study, a greater generalization capability of ANNs when compared to algorithmic methods such as LDA, suggesting that they may be a more appropriate tool for ECG heartbeat classification.

VII. Conclusion

In this study, we have studied how the application of a suitable FS procedure can lead to the improvement of classifier's performances and simultaneously reduce their complexity, which can be of great help to improve online ECG monitoring, especially in ambulatory settings. To this end, a new performance measure index was introduced, which deals specifically with two major issues in heartbeat classification measurements: class imbalance and relative importance of the possible arrhythmias. Using this criterion function, the algorithm was executed upon two sets of 26 and 71 causal features, respectively. The first set contained features used in a previous study; it served to evaluate the capability of the method to improve previously proposed classifier systems. The second had a twofold purpose: first, the capability of the method to narrow down comprehensive datasets was assessed; second, by including only features suitable for online monitoring in the examined set, further insight into which are the most suited features for this kind of setting was accomplished. In what follows, the returned subsets were used to carry out ECG classification using two different classifier paradigms. The achieved performance results on both models prove the suitability of the $I_{j\kappa}$ index driven SFFS algorithm to improve performance while reducing classifier complexity, and reveal the capability of the MLP to outperform linear classifiers in the heartbeat classification field.

References

[1] X. F. Teng, Y. T. Zhang, C. C. Y. Poon, and P. Bonato, "Wearable medical systems for p-health," *IEEE Rev. Biomed. Eng.*, vol. 1, pp. 62–74, 2008.

[2] S. N. S. Zaunseder and W.-J. Fischer, "Prolonged wearable ECG monitoring—A wavelet based approach," in *Proc. IEEE Sens. Conf.*, 2007, pp. 1197–1200.

[3] W. Chang, K. Lin, and R. Lee, "Pattern recognition technique to QRS complex classification," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Images 21st Cent.*, Seattle, WA, 1989, pp. 32–33.

[4] L. Senhadji, G. Carrault, J. Bellanger, and G. Passariello, "Comparing wavelet transforms for recognizing cardiac patterns," *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 167–173, Mar./Apr. 1995.

[5] P. de Chazal, M. O'Dwyer, and R. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[6] M. Llamedo and J. P. Martinez, "An ECG classification model based on multilead wavelet transform features," in *Proc. Comput. Cardiol.*, 2007, vol. 34, pp. 105–108.

[7] M. Llamedo and J. P. Martinez, "Heartbeat classifier using feature selection with database generalization criteria," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 616–625, Mar. 2011.

[8] J. Rodriguez, A. Goñi, and A. Illarramendi, "Real-time classification of ECGs on a PDA," *IEEE Trans. Inf. Technol. Biomed.*, vol. 9, no. 1, pp. 23–34, Mar. 2005.

[9] I. Jekova, G. Bortolan, and I. Christov, "Assessment and comparison of different methods for heartbeat classification," *Med. Eng. Phys.*, vol. 30, no. 2, pp. 248–257, Mar. 2008.

[10] K. Lai and K. Chan, "Real-time classification of electrocardiogram based on fractal and correlation analyses," in *Proc. 20th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Hong Kong, 1998, pp. 119–122.

[11] M. I. Owis, A. H. Abou-Zied, A.-B. M. Youssef, and Y. M. Kadah, "Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 7, pp. 733–736, Jul. 2002.

[12] L. Khadra, A. S. Al-Fahoum, and S. Binajjaj, "A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1840–1845, Nov. 2005.

[13] Y. H. Hu, S. Palreddy, and W. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, Sep. 1997.

[14] S. Osowski and T. H. Linh, "ECG beat recognition using fuzzy hybrid neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1265–1271, Nov. 2001.

[15] Z. Dokur and T. Olmez, "ECG beat classification by a novel hybrid neural network," *Comput. Methods Progr. Biomed.*, vol. 66, no. 2–3, pp. 167–181, Sep. 2001.

[16] H. G. Hosseini, D. Luo, and K. J. Reynolds, "The comparison of different feed forward neural network architectures for ecg signal diagnosis," *Med. Eng. Phys.*, vol. 28, no. 4, pp. 372–378, May. 2006.

[17] J. Pardey, "On the use of artificial neural networks in a commercial holter algorithm," in *Proc. Comput. Cardiol.*, Sep. 17–20, 2006, pp. 485–488.

[18] E. D. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of ECG signals using Lyapunov exponents," *Comput. Methods Programs Biomed.*, vol. 93, no. 3, pp. 313–321, Mar. 2009.

[19] F. M. Ham and S. Han, "Classification of cardiac arrhythmias using fuzzy ARTMAP," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 4, pp. 425–430, Apr. 1996.

[20] P. Bozzola, G. Bortolan, C. Combi, F. Pinciroli, and C. BroHet, "A hybrid neuro-fuzzy system for ECG classification of myocardial infarction," in *Proc. Comput. Cardiol*, 1996, pp. 241–244.

[21] S. Barro, M. Fernndez-Delgado, J. A. Vila-Sobrino, C. V. Regueiro, and E. Snchez, "Classifying multichannel ECG patterns with an adaptive neural network," *IEEE Eng. Med. Biol. Mag.: Q. Mag. Eng. Med. Biol. Soc.*, vol. 17, no. 1, pp. 45–55, Jan./Feb. 1998.

[22] Y. Wang, Y. S. Zhu, N. V. Thakor, and Y. H. Xu, "A short-time multifractal approach for arrhythmia detection based on fuzzy neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 9, pp. 989–995, Sep. 2001.

[23] S. Osowski and L. T. Hoai, "Analysis of features for efficient ECG signal classification using neuro-fuzzy network," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 25–29, 2004, vol. 3, pp. 2443–2448.

[24] L.-Y. Shyu, Y.-H. Wu, and W. Hu, "Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1269–1273, Jul. 2004.

[25] I. Niwas, S. S. Kumari, and V. Sadasivam, "Artificial neural network based automatic cardiac abnormalities classification," in *Proc. 6th Int. Conf. Comput. Intell. Multimedia Appl.*, Aug. 16–18, 2005, pp. 41–46.

[26] N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, and M. Strintzis, "ECG pattern recognition and classification using non-linear transformations and neural networks: A review," *Int. J. Med. Inf.*, vol. 52, no. 1–3, pp. 191–208, 1998.

[27] A. S. al Fahoum and I. Howitt, "Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias," *Med. Biol. Eng. Comput.*, vol. 37, no. 5, pp. 566–573, Sep. 1999.

[28] K. Park, B. Cho, D. Lee, S. Song, J. Lee, Y. Chee, I. Kim, and S. Kim, "Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function," in *Proc. Comput. Cardiol.*, Sep., 2008, pp. 229–232.

[29] G. Lannoy, "Feature relevance assessment in automatic inter-patient heartbeat classification," in *Proc. BIO SIGNALS 2010—Int. Conf. Bio-Inspir. Syst. Signal Process.*, pp. 13–20.

[30] S. Palreddy, Y. H. Hu, V. Mani, and W. Tompkins, "A multiple-classifier architecture for ECG beat classification," in *Proc. 7th IEEE Workshop Neural Netw. Signal Process.*, 1997, pp. 172–181.

[31] M. Lagerholm, "Clustering ECG complexes using hermite functions and self-organizing maps," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 838–848, Jul. 2000.

[32] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Med. Eng. Phys.*, vol. 28, no. 9, pp. 876–887, Nov. 2006.

[33] G. K. Prasad and J. S. Sahambi, "Classification of ECG arrhythmias using multi-resolution analysis and neural networks," in *Proc. TENCON 2003—Conf. Convergent Technol. Asia-Pacific Region*, Oct., vol. 1, pp. 227–231.

[34] O. T. Inan, L. Giovangrandi, and G. T. A. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2507–2515, Dec. 2006.

[35] P. de Chazal and B. Celler, "Selecting a neural network structure for ECG diagnosis," in *Proc. 20th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Hong Kong, 1998, pp. 1422–1425.

[36] Y.-H. Chen and Sung-Nien Yu, "Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network," *Pattern Recognit. Lett.*, vol. 28, pp. 1142–1150, 2007.

[37] P. de Chazal and R. Reilly, "A comparison of the use of different wavelet coefficients for the classification of the electrocardiogram," in *Proc. 15th Int. Conf. Pattern Recognit.Proc.*, vol. 2, Barcelona, Spain, 2000, pp. 255–258.

[38] N. Acur, "Classification of ECG beats by using a fast least square support vector machines with a dynamic programming feature selection algorithm," *Neural Comput. Appl.*, vol. 14, pp. 299–309, 2005.

[39] R. E. Bellman, *Dynamic Programming*. New York: Dover, 2003.

[40] P. Pudil, F. Ferri, J. Novovicova, and J. Kittler, "Floating search methods for feature selection with nonmonotonic criterion functions," in *Proc. 12th IAPR Int. Pattern Recognit. Vol. 2—Conf. B: Comput. Vis. Image Process. Conf.*, Oct. 9–13, 1994, vol. 2, pp. 279–283.

[41] 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.

[42] P. Somol, P. Pudil, J. Novovičová, and P. Paclík, "Adaptive floating search methods in feature selection," *Pattern Recogn. Lett.*, vol. 20, no. 11–13, pp. 1157–1163, 1999.

[43] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, Jun. 2000.

[44] *Testing and Reporting Performance Results of Cardiac Rhythm and ST-Segment Measurement Algorithms.* Association for the Advancement of Medical Instrumentation (AAMI), Arlington, VA, EC57: 1998.

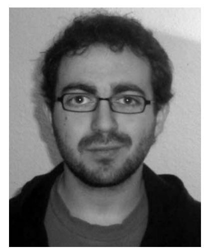
[45] P. de Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2535–2543, Dec. 2006.

[46] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, Jan. 1995.

[47] L. Molina, L. Belanche, and A. Nebot, "Feature selection algorithms: A survey and experimental evaluation," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, 2002, Dec. 9–12, pp. 306–313.

[48] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, pp. 37–46, 1960.

[49] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*. Cambridge, MA: MIT Press, 1986, pp. 318–362.



Tanis Mar (S'05) was born in Soria, Spain, in 1984. He received the M.S. degree in telecommunication engineering from the University of Zaragoza, Zaragoza, Spain, in 2009. He is currently working toward the M.S. degree in computational neuroscience at the Bernstein Center, Berlin, Germany.

From 2006 to 2007, he spent a year studying at the Dresden University of Technology, Dresden, where he was a Research Assistant from 2009 to 2010. His current interests include biomedical signal processing, in particular in its application in cardiovascular treatment and neuroscience development.

Sebastian Zaunseder was born in Nuremberg in 1981. He received the Diploma degree in electrical engineering from the Dresden University of Technology, Dresden, Germany, in 2007. Between 2004 and 2005, he spent one year studying at the University of Zaragoza, Zaragoza, Spain as an exchange student.

Since 2007, he has been a Research Associate at the Fraunhofer Institute for Photonic Microsystems, Dresden, Germany. His research interests include biomedical signal processing. Particularly, the applied combination of signal processing methods to expand the usability of ambulatory long-term recordings from variable origin is in the scope of his work.

Juan Pablo Martínez was born in Zaragoza, Aragon, Spain, in 1976. He received the M.S. degree in telecommunication engineering, and the Ph.D. degree in biomedical engineering from the University of Zaragoza (UZ), Zaragoza, Spain, in 1999 and 2005, respectively.

Since 2000, he has been an Assistant Professor at the Aragon Institute of Engineering Research, UZ, where he has been an Associate Professor since 2007. He is also with the Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza. His current research interests include biomedical signal processing, with main interest in signals of cardiovascular origin.

Mariano Llamedo was born in Buenos Aires, Argentina, in 1979. He received the M.Sc. degree in electronic engineering from the National Technological University of Buenos Aires (UTN-FRBA), Buenos Aires, Argentina, in 2005.

From 2005 to 2008, he was with the Department of Electronic Engineering, UTN-FRBA, as an Assistant Professor and Research Fellow. Since 2008, he has been involved as a Researcher with the Aragon Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain. His research interest include biomedical signal processing, with main interest in signals of cardiovascular origin.

Rüdiger Poll was born in Köslin, Germany, in 1944. He received the M.D. degree from the Medical Academy Dresden, Dresden, Germany, and the Dipl.Ing. degree from the Dresden University of Technology, Dresden, in 1976 and 1981, respectively.

Before retiring in 2010, Dr. Poll was a Physician in the Department of Cardiology at the Medical Academy Dresden, and a Professor of Biomedical Engineering with the Faculty of Electrical and Computer Engineering at the Dresden University of Technology, where he was the Head of the Institute of Biomedical Engineering, from 2002 to 2010.