

Support Vector Black-Box Interpretation in Ventricular Arrhythmia Discrimination

The Advantages of the Support Vector Method Make It Attractive for Clinical Applications

In our companion article [1], a new discrimination algorithm, based on the analysis of ventricular electrogram (EGM) onset, was proposed in order to discriminate between supraventricular and ventricular tachycardias (SVTs and VTs) in implantable cardioverter defibrillators (ICDs). Due to the absence of a detailed statistical model for the ventricular activation, this algorithm was based on a support vector method (SVM) learning machine [2], plus bootstrap resampling [3] to avoid overfitting. This SVM classifier was trained with available arrhythmia episodes, so that it can be viewed as containing a statistical model for the differential diagnosis. However, the black-box model character of any learning machine presents two main problems in a clinical environment:

- the nature of the problem remains in an obscure, inaccessible, mathematical formula;
- the cardiologist in charge of programming the criterion in the device has no knowledge of the underlying mechanism of this algorithm.

A solution is the extraction of the statistical information enclosed in the black-box model. As neural networks encrypt the model into a complex, nonlinear, mathematical formula, they are not easy to interpret at all. But in contrast to backpropagation networks, the SVM could be more appropriate for this purpose, given that the support vectors represent the critical samples for the classification task.

In this article we propose two SVM-oriented analyses and their use in building two new differential diagnosis algorithms based on the ventricular EGM

onset criterion. The following approaches are suggested:

- 1) a geometrical analysis of the input feature space and its relationship to the critical samples (i.e., the support vectors); and
- 2) a study of the relevance of the activation time state.

As was demonstrated in the companion article, an incremental learning procedure should be used for each algorithmic implementation in order to reduce the inter-patient variability as new information about the patient (i.e., new arrhythmia episodes) becomes available. Note that the records in Base C (training control group) and Base D (independent test group) have been described in the companion article [1].

Geometrical Analysis

In the preceding article [1], discrimination between SVTs and VTs from Base C episodes (38 SVTs and 68 VTs from 26 patients) taking into account the ventricular EGM onset criterion was achieved. For this purpose, the EGMs during tachycardia (and its preceding SR) were low-pass filtered (50 Hz), cycle segmented, synchronized to the R-wave, and averaged, producing a tachycardia template plus an SR template. The samples contained in the 80 ms preceding the R-wave in the SR and in the tachycardia templates were used as a single input feature vector for each episode. A radial basis function (RBF) kernel SVM was trained, fixing the free parameters (kernel width σ and margin-losses trade-off C) with the bootstrap resampling method in order to avoid the overfitting to the training set. The resulting nonlinear SVM classifier had 35 support vectors (106 total feature vectors), 22

José L. Rojo-Álvarez¹, Ángel Arenal-Maiz²,
Antonio Artés-Rodríguez¹

¹Escuela Politécnica Superior,
Universidad Carlos III, Madrid

²Department of Cardiology,
Hospital GU Gregorio Marañón, Madrid

corresponding to saturated coefficients (eight SVTs and 14 VTs), and 13 corresponding to nonsaturated coefficients (five SVTs and eight VTs). Figure 1 depicts the SVT and VT time samples for the episodes in the whole set, together with the preceding SR time samples; this figure also represents the support vectors for the three kinds of rhythms. The morphological similarity among these critical feature vectors can be seen there. We propose the geometrical comparison of the whole set input space to the obtained support vectors, as the later represents the critical samples for the classification.

Methods

For a given set of column vectors

$$\mathbf{V} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \quad (1)$$

we will denote its covariance matrix as:

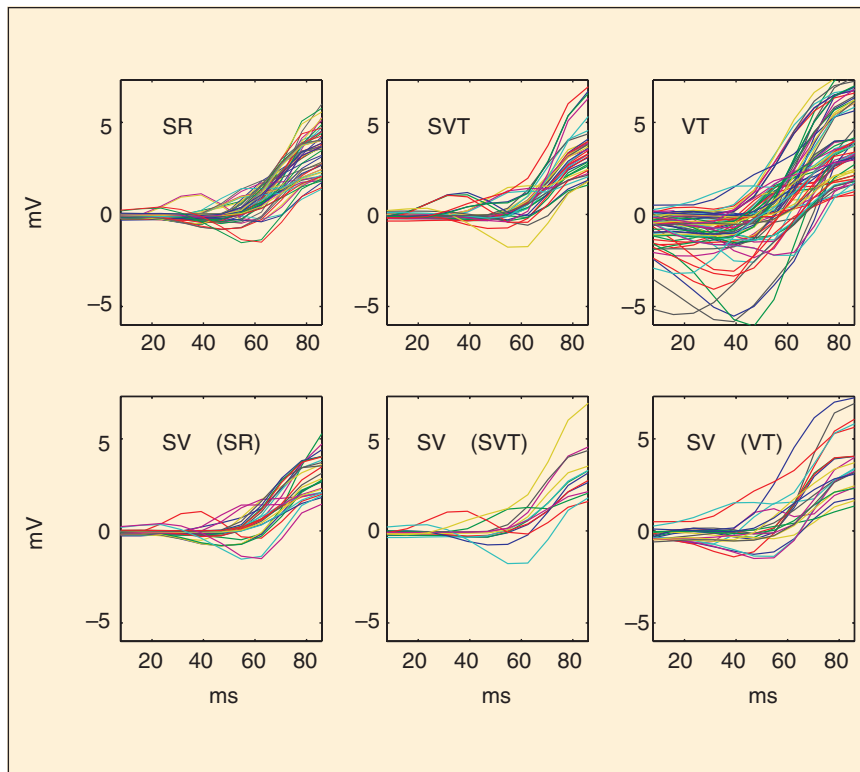
$$\Sigma_{\mathbf{V}} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (2)$$

where $\bar{\mathbf{x}}$ is the sample average of \mathbf{V} . It can be shown [4] that this matrix can be factorized (principal component analysis, PCA), thus

$$\Sigma_{\mathbf{V}} = \mathbf{U} \cdot \Lambda \cdot \mathbf{U}^T \quad (3)$$

with Λ a diagonal matrix of eigenvalues $\{\lambda_1, \dots, \lambda_N\}$, and the columns of \mathbf{U} being an orthonormal base of \mathbf{R}^d (eigenvectors) when the matrix is full rank. The PCA for the SR, SVT, and VT covariance matrices was performed in both the entire set and the support vector set. This led to $d = 11$ eigenvectors plus their corresponding eigenvalues in each case. In order to evaluate the scatter degree in every subset, a measure of the distance among the eigenvectors, weighted by their eigenvalues, was obtained. For example, the distance between two of the eigenvectors of the SR subset is:

$$d_k^{SR} = \|\lambda_i^{SR} \mathbf{v}_i^{SR} - \lambda_j^{SR} \mathbf{v}_j^{SR}\|_1. \quad (4)$$



1. Representation of the SR, SVT, and VT onsets (80 ms preceding the R-wave) in the episodes of Base C. Top: whole set waveforms. Bottom: support vectors.

Table 1. Averaged Distances Between Pairs of Eigenvectors.

	SR	SVT	VT
Whole set (100 x d)	27 ± 36	26 ± 33	86 ± 98
Support Vector Set (100 x d)	20 ± 29	35 ± 58	65 ± 105

This distance is averaged over all possible pairs (i, j) of different eigenvectors in each subset. The same scatter was obtained for the remaining five subsets.

Results

Table 1 assembles the mean ± standard deviation for the averaged subset distances. The scatter in the whole set is similar between SRs and SVTs, whereas it is markedly higher in VTs. The scatter in the support vector set is higher in SVTs with respect to SR, while the VTs scatter is still much greater than in the rest of cases. The relative scatter increasing in support vector SVTs with respect to the support vector SR is due to the fact that critical SVT samples are those that are more different from the SR vectors. Even in these circumstances, VT vectors are more sparse than SVT vectors. So, a reasonable approach is to center the searching of SVTs using their distance to the SR, excluding as VTs those cases that are away from it.

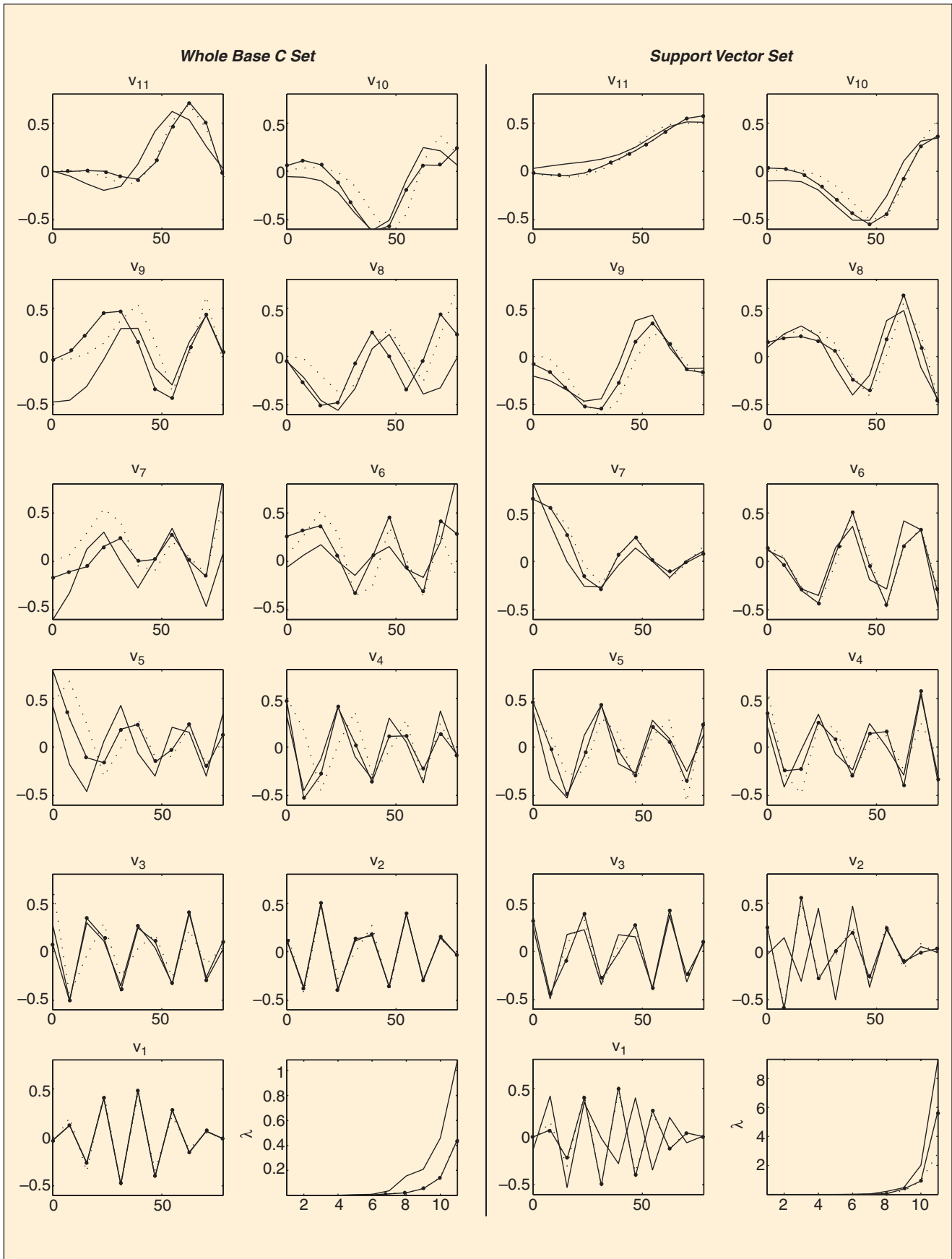
Figure 2 shows the eigenvectors and their eigenvalues. Differences in the whole set appear especially in the most significant (greater eigenvalues) vectors (\mathbf{v}_5 to \mathbf{v}_{11}), whereas in the support vector set they appear mainly in the least significant (minor eigenvalues) vectors ($\mathbf{v}_1, \mathbf{v}_2$). The most significant one, \mathbf{v}_1 , differs greatly in each set; however, it points in both cases to important differences between atrial rhythms (SR and SVT) and ventricular rhythms in the starting 40 ms. In the support vector set, \mathbf{v}_1 and \mathbf{v}_2 seem to point at differences in two complementary regions: early and late. Hence, critical differences are to appear in two time intervals: early and late activation.

Conclusions

We propose to group the tachycardia according to their distance from the SR. Also, it is convenient to cluster the SVT vectors, excluding as VT vectors those ones with features being far from SR in any direction. Two different time sections, early and late activation, should be considered, taking into account the different voltage amplitudes expected in these zones.

Figure 3 represents this scheme. The distance from the tachycardia EGM to the SR EGM can be calculated, in a first step, using the difference between the records in the 80 ms previous to the R-wave:

$$f(t) = EGM^{SR}(t) - EGM^T(t). \quad (5)$$



2. Eigenvectors for SR (dotted), SVT (continuous dotted) and VT (continuous) for the covariance matrices of SR, SVT, and VT vectors in the whole Base C set (left) and in the Support Vector set (right), together with their corresponding eigenvalues.

The proposed statistical parameters stem from the distinction between early and late activation; that is:

$$V_1 = \int_{t=-80\text{ms}}^{t=t_c} f(t) dt;$$

$$V_2 = \int_{t=t_c}^{t=0\text{ms}} f(t) dt \quad (6)$$

$$\mathbf{v} = [V_1, V_2]. \quad (7)$$

For clustering the SVT vectors, let us consider $\{\mathbf{v}_i\}_{i \in \text{SVT}}$ as the set of available SVT vectors. The mean and the covariance matrix are defined as:

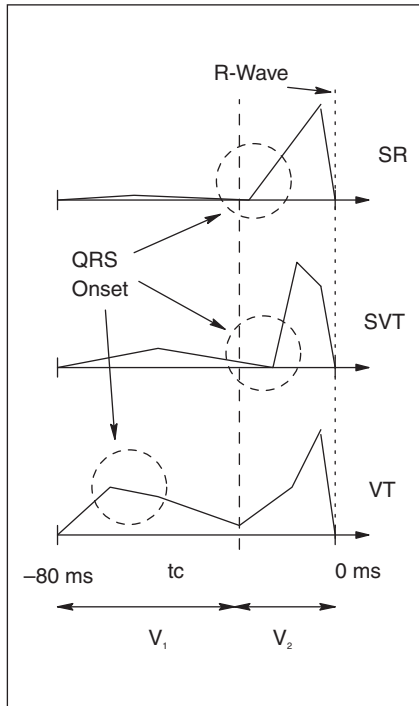
$$\mu_{\text{SVT}} = \frac{1}{N_{\text{SVT}}} \sum_{i \in \text{SVT}} \mathbf{v}_i \quad (8)$$

$$\Sigma_{\text{SVT}} = \frac{1}{N_{\text{SVT}}} \sum_{i \in \text{SVT}} (\mathbf{v}_i - \mu_{\text{SVT}})^T (\mathbf{v}_i - \mu_{\text{SVT}}) \quad (9)$$

where N_{SVT} is the number of SVT episodes. We propose the following transformation for each observed vector:

$$\mathbf{v}' = [V'_1, V'_2] = \Sigma_{\text{SVT}}^{-1} (\mathbf{v} - \mu_{\text{SVT}}). \quad (10)$$

This transformation will make the SVT vectors centered and uniformly distributed around the origin. The VT vectors will be far from the origin in any direction, and discrimination in this space can be



3. Scheme of the ventricular EGM onset according to early (V1) and late (V2) activation.

performed easily with a bounding sphere of radius R_o .

About the First Derivative and the Splitting Time

At this point two open issues remain. First, the changes were no longer evaluated employing the first derivative in the companion article because of the use of a learning method, though the hypothesis of ventricular EGM onset included the changes as a part of the discrimination criterion [1]; in this new scheme approach, no longer based on a black-box model, the use of the first derivative should be again tested. Second, the determination of a convenient t_c splitting time has not yet been stated.

We consider both subjects simultaneously. To take into account the first derivative, we will use another definition of the EGM function:

$$f(t) = \left| \frac{\partial \text{EGM}^{\text{SR}}(t)}{\partial t} \right| - \left| \frac{\partial \text{EGM}^{\text{T}}(t)}{\partial t} \right| \quad (11)$$

where the absolute value has been included in order to study only the intensity of the changes, rather than their morphology.

In the Base C episodes, both the onset waveform function in Eq. (5) and the derivative function in Eq. (11) were studied. The value of t_c was changed at all the possible values between -80 and 0 ms. For each value of t_c , feature vectors given by Eqs. (6) and (7) were obtained, and an RBF SVM classifier was built. Optimal RBF width was found with bootstrap resampling of the actual risk [1]. For a 128-Hz sampled signal, there were 11 possible values for t_c , and the error rate was determined for each of them in both schemes.

Figure 4(a) shows the results for both schemes. The onset EGM waveform works better for $-70 < t_c < -40$ ms, while the optimum range for the derivative-based scheme has an optimal range of $-40 < t_c < -15$ ms. The minimum error rate is reached by the derivative-based scheme.

Under our present working scheme, the first derivative should be considered, as it enhances the high-frequency components in the onset. This can be seen as a cancellation of the slow components of the T-wave, which interfere with the QRS onset, leading to a clearer definition of the onset potentials [see Fig. 4(c)]. The range

of appropriate values for t_c can be explained from the representation of the medians of SVT and VT vectors in Fig. 4(b); along this interval, the statistics represent separately and clearly the early onset activation (where the VT activity tends to be greater) from the late-onset activation (where the SVT energies trend to be greater).

The first derivative should be used for the proposed scheme. It also has been shown that it is possible to determine the optimum t_c instant by bootstrap resampling. For Base C, the splitting time obtained is $t_c = -30$ ms.

First Algorithmic Implementation

The procedure for the first algorithmic implementation is as follows:

- 1) For each available episode:
 - a) Obtain the templates for SR and T as given by Eq. (11).
 - b) Obtain the features given by Eq. (6) and Eq. (7).
 - 2) Estimate the SVT mean and covariance matrix with Eq. (8) and Eq. (9).
 - 3) Obtain the normalized statistics as given by Eq. (10).
 - 4) Determine a threshold R_o for the SVT region: for example, R_o excluding 100% of the individual observed VT episodes.
 - 5) Classify a new episode according to:

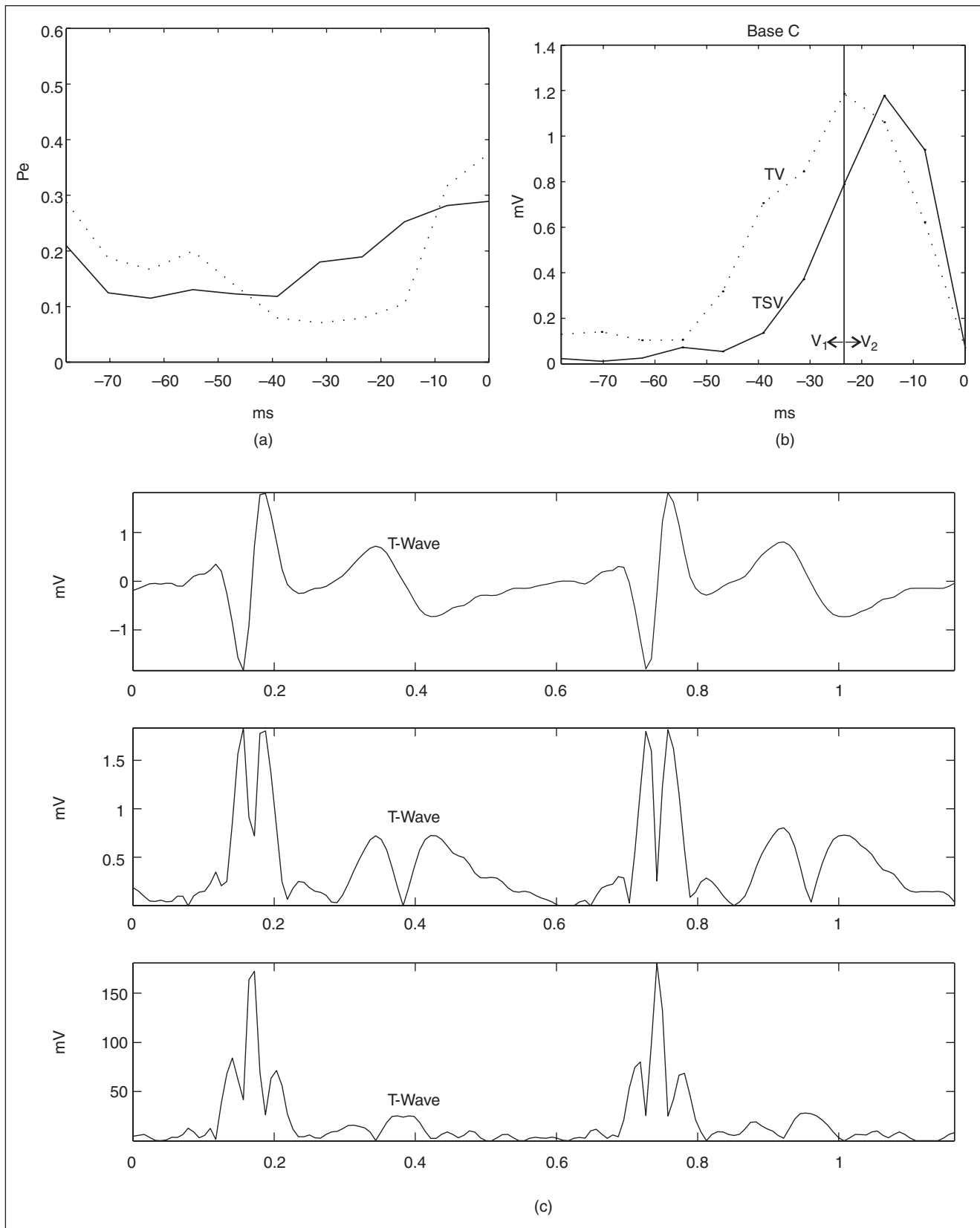
$$\left\{ \begin{array}{l} \|\mathbf{v}'_i\|_2 < R_o \rightarrow \text{SVT} \\ \|\mathbf{v}'_i\|_2 > R_o \rightarrow \text{VT} \end{array} \right.$$

where $\|\mathbf{v}'_i\|_2$ denotes the L_2 norm of the feature vector.

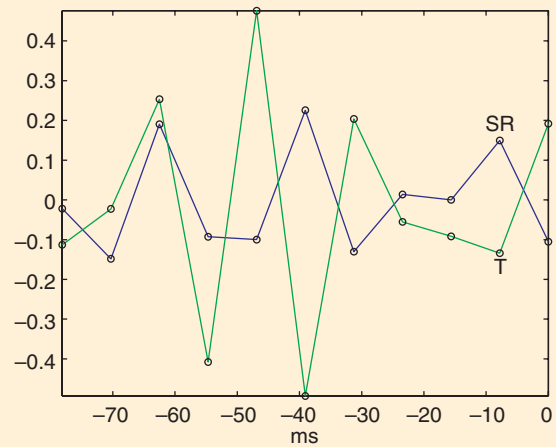
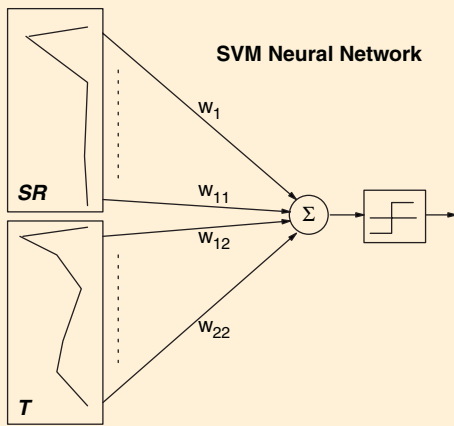
Relative Importance of the Activation Time

The relative importance of each time instant can be analyzed with the aid of the SVM. A nonlinear machine would lead to a better classification function, but it will be hardly useful for interpretation purposes. On the other side, a linear machine will give clearer information about the relevance of each time sample according to the corresponding weight. This will be the starting point to the next algorithmic implementation.

Two linear methods were used: the classical Fisher linear analysis [5] and the SVM with linear kernel. Figure 5(a) represents the input feature space for each episode, consisting on the beat-averaged, derived, and rectified EGM onset (80 ms

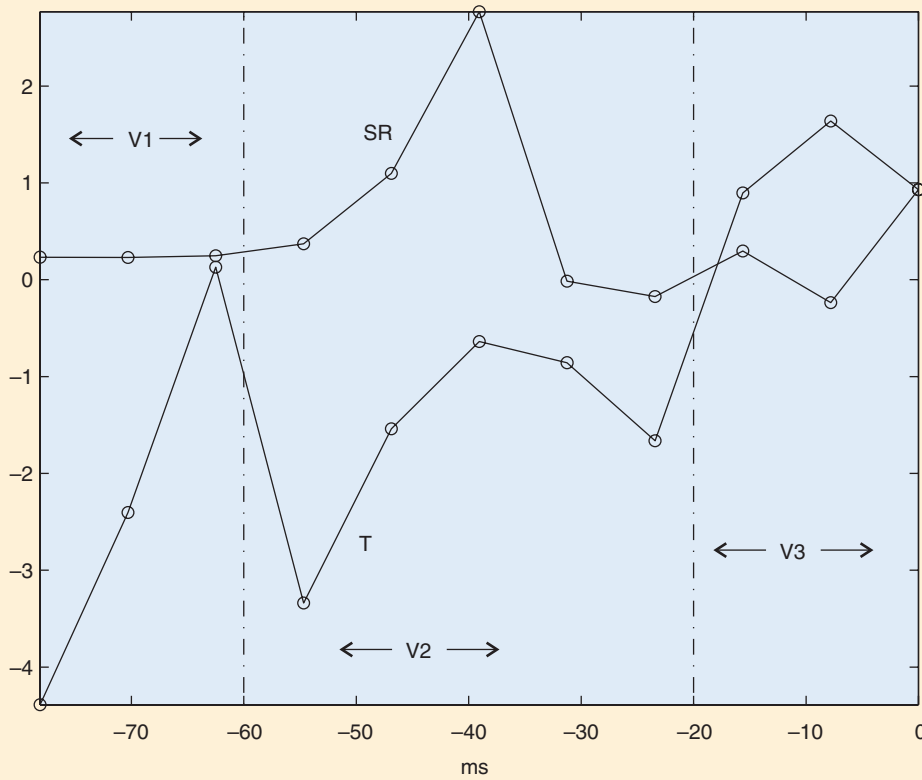


4. (a) Error rate for the waveform-based (continuous line) and the derivative-based (dotted line) scheme as a function of the splitting time t_s . (b) Medians of the SVT (continuous line) and the VT (dotted line) derived, rectified ventricular EGM onsets; during the early activation, VT is greater than an SVT, whereas during the late activation changes in the SVT are greater. (c) Effect of the first derivative on the T-wave. Top: EGM during SR. Middle: rectified EGM. Bottom: derived, rectified EGM. Note that the ending T-wave is lesser in this case.



(a)

(b)



(c)

5. Relative importance of the onset samples. (a) Linear classification scheme. (b) Weights obtained from the Fisher classifier. (c) Weights obtained from the SVM classifier: This points to the existence of three different states in the initial ventricular activation—early, transient, and late.

previous to the R-wave) for both the tachycardia and the preceding SR.

Figure 5(b) and (c) depicts the weights for the Fisher and for the SVM classifiers, respectively. The Fisher analysis does not exhibit any clear behavior, whereas SVM allows us to consider three different regions, namely early, transient, and late activation. From a statistical analysis viewpoint, differences in the results between both classifiers are due to the fact that SVM makes the decision based on the critical samples (and hence tracing a more accurate boundary) while the Fisher classifier takes into account all of the population (and it results in a rougher boundary).

Three different time regions can be considered in the analysis. The following statistics can be obtained for each episode, according to Eq. (11) [see Fig. 5(c)]:

$$V_1 = \int_{t=-80\text{ms}}^{t=-60\text{ms}} f(t)dt \quad (12)$$

$$V_2 = \int_{t=-60\text{ms}}^{t=-20\text{ms}} f(t)dt \quad (13)$$

$$V_3 = \int_{t=-20\text{ms}}^{t=t_0\text{ms}} f(t). \quad (14)$$

The normalization with respect to the available SVT vectors is the same as described in the preceding section.

A Priori Knowledge: Bootstrap Regularization

Some numerical considerations should be taken into account before proceeding. The normalization with respect to the SVT episodes is based on the estimation of the SVT covariance matrix using the feature vectors and its subsequent inversion. This is an ill-posed problem, as far as small errors in the estimation of the covariance matrix can easily lead to great errors when estimating its inverse, and a regularization process should be included in order to reduce the numerical instability. We propose the use of bootstrap resampling to provide a criterion to get the optimum regularization parameter γ , and thus work with a regularized version of the covariance matrix:

$$\Sigma_r = \Sigma + \gamma \mathbf{I}. \quad (15)$$

The proposed criterion is to determine the value of γ minimizing the classification error in the observed episodes.

A range for $\gamma \in (10^{-4}, 5)$, in a logarithmic scale, was explored. For each value we generated 200 bootstrap resamples of the statistics in Eqs. (12), (13), and (14). To avoid rank deficiencies in the

covariance matrix, only nonrepeated vectors in each resample were considered. The mean of each resample was extracted, and the vectors were multiplied by the resample covariance matrix regularized as given by Eq. (15). The minimum error reached in each resample was determined as a function of R_o .

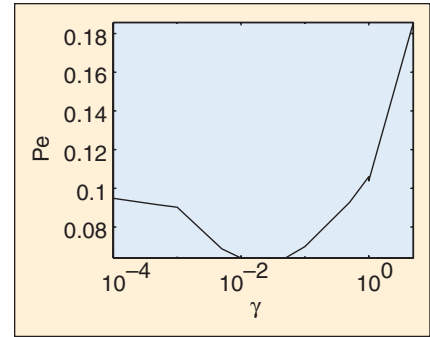
Figure 6 shows the replicated averaged error as a function of the regularization parameter. For high values of γ , the regularized covariance matrix becomes a biased estimation of the true matrix, leading to a high error rate. For small values of γ , there is not enough regularization and the error increases as well. A range of values in the order of magnitude of 10^{-2} appears to be suitable in terms of error rate.

Bootstrap regularization in the estimation of the covariance matrix leads to a better class separation and better numerical stability.

Second Algorithmic Implementation

The procedure for the second algorithmic implementation is as follows:

- 1) For each available episode:
 - a) Obtain the templates for SR and T as given by Eq. (11).



6. Bootstrap predicted error rate as a function of the regularization parameter.

- b) Obtain the statistics V_1, V_2, V_3 , as given by Eqs. (12), (13), and (14); the feature vector is:

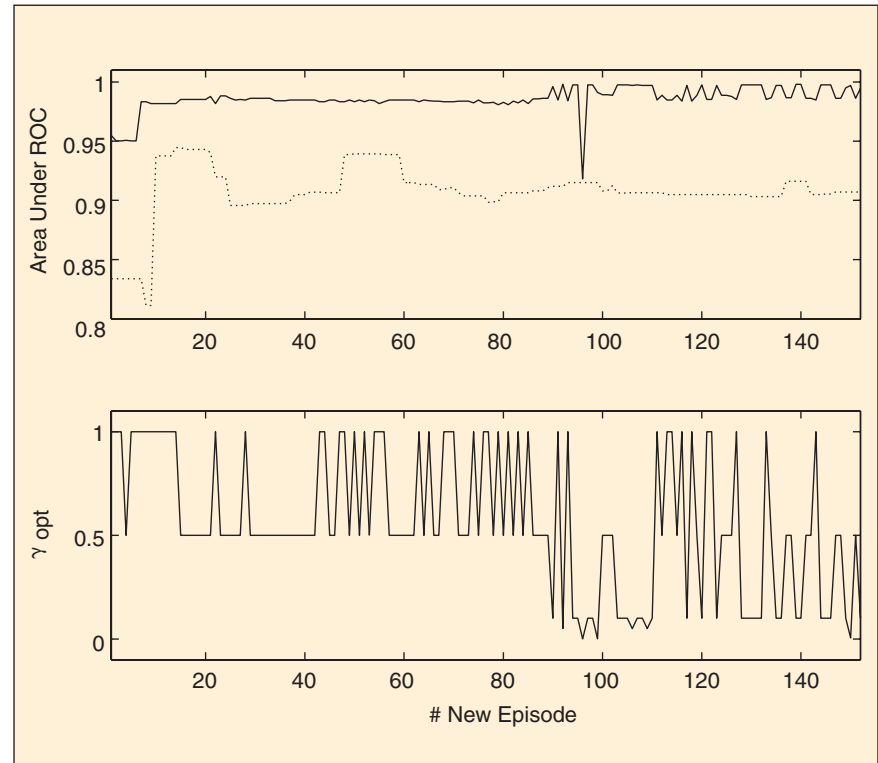
$$\mathbf{v} = [V_1, V_2, V_3].$$

- 2) Estimate the SVT mean and covariance matrix as given by Eqs. (8) and (9).

- 3) Regularize the estimated covariance matrix according to Eq. (15).

- 4) Obtain the normalized statistics as given by Eq. (10).

- 5) Determine a threshold R_o for the SVT region: for example, the R_o excluding 100% of the individual observed VT episodes.



7. Incremental learning for the second algorithmic implementation. Top: evolution of the area under the curve for the regularized (continuous line) and nonregularized (dotted line) procedure. Bottom: regularization parameter found by the bootstrap resampling.

6) Classify a new episode according to:

$$\begin{cases} \|\mathbf{v}'_i\|_2 < R_o \rightarrow SVT \\ \|\mathbf{v}'_i\|_2 > R_o \rightarrow VT. \end{cases}$$

Incremental Learning

As stated in the companion paper [1], an important question in any arrhythmia discrimination algorithm is to progressively tune it to the available observations of the particular patient. We named this process incremental learning (IL). For the above proposed algorithmic implementations, the statistics that must be brought up to date when new episodes become available are the SVT vector average and its covariance matrix. The latter is the most critical, because it has to be inverted for vector normalization. We propose the determination of the regularization parameter γ according to the minimum error on the available samples at each step using the bootstrap resampling as an error estimation procedure.

For the second algorithmic implementation, a patient from Base D was selected, with a total of 257 episodes (77 SVTs, 180 VTs). These episodes were divided into three subsets:

1) An initial set containing 2% of episodes, representing those available from the patient electrophysiologic study and the ICD implantation procedure.

2) A test subset (68% of episodes) whose contents will be progressively included (one episode at each step).

3) An independent validation subset (30% of episodes) for estimation of the area under the curve.

For these subsets, the SVT average vector and the covariance matrix were estimated and the regularized matrix was obtained with 1000 bootstrap resamples at each step. The range of values of the regularization parameter was:

$$\gamma = [0 \quad 10^{-5} \quad 10^{-4} \quad 10^{-3} \quad 5 \cdot 10^{-3} \quad 10^{-2} \quad 5 \cdot 10^{-2} \quad 10^{-1} \quad 0.5 \quad 1] \quad (16)$$

The area under the curve was estimated in the validation subset for both the regularized and the nonregularized covariance matrix.

Figure 7 shows the evolution of the areas under the curve, together with the optimal regularization parameter. The regularized algorithm shows high separation capabilities between SVTs and VTs starting from the earliest stages, and they increase as new samples are known. On the other hand, the absence of regularization leads to an inconsistent statistic with decreased separation capabilities, which do not improve with the addition of new samples. Note that the value of the regularization parameter is minor, on average, at the first steps.

Including an IL procedure in the proposed algorithmic implementations leads to numerical stability and, hence, increased classification performance. The bootstrap resampling offers a convenient way of regularization for this purpose.

Comparison with Previous Methods

Finally, an extensive comparison with a number of arrhythmia discrimination criteria is presented. The following algorithms were tested:

- The Rate Criterion, comparing the averaged cardiac rates of tachycardia episodes.
- The Correlation Waveform Analysis (CWA) [6], comparing a 160-ms, SR template to the tachycardia beats using the correlation coefficient:

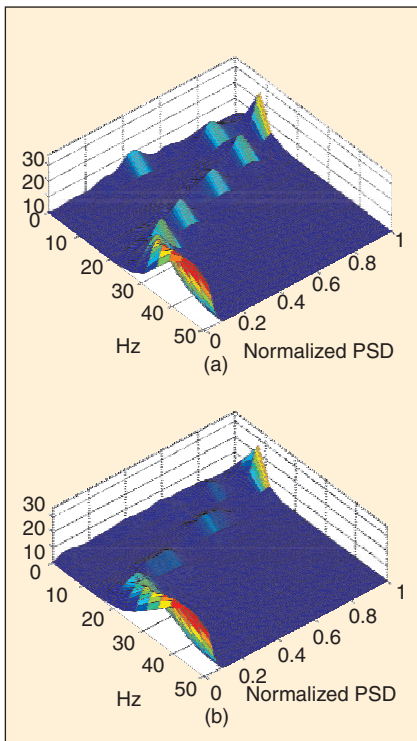
$$\rho = \frac{\sum_{i=1}^N (t_i - \bar{t})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^N (t_i - \bar{t})^2 \sum_{i=1}^N (s_i - \bar{s})^2}} \quad (17)$$

where t_i, s_i are the i th samples of the SR template and the tachycardia beat, respectively; \bar{t}, \bar{s} are the template and the tachycardia beat averages; and N is the number of samples. The episodes were compared using their beat-coefficients median.

- The direct training of an MLP with the SR and the tachycardia templates used for the CWA. To avoid overfitting, the actual error was bootstrap-estimated at each descent step, stopping the backpropagation when the error no longer decreased.
- The spectral analysis of the QRS complex [7], which was based on the hypothesis of a prominent high-frequency content in SVT beats [see Fig. (8)]. The implementation used an MLP trained with five spectrum samples of the QRS complex, and the training was stopped as described in the previous paragraph.
- The QRS Width Criterion [8], which consists of measuring the duration of a single beat, classifying it (related to the SR beat width) as narrow (supraventricular) or wide (ventricular origin).
- The Prediction Error Criterion [9], which proposes the AR modeling of the SR EGM and the observation of the prediction error in the tachycardia EGM: a reduced error means that the tachycardia is correctly modeled by the SR parameters, whereas an increase in the error is supposed to stem from a ventricular (hence with different morphology) tachycardia.
- Support Vector Learning with incremental margin (SVMm) [1].
- Support Vector Learning with incremental cost (SVMc) [1].
- First Algorithmic Implementation (V2d).
- Second Algorithmic Implementation (V3d).

For all of these methods, the area under the curve was obtained for the control test (Base C), validation test (Base D), and individual validation test (single patient from Base D), which are shown in Table 2.

The discriminatory power yielded by the rate criterion in the control set might be due to the nature of the SVT episodes (treadmill test, leading to slow SVTs). In the single patient, the rate is also an adequate discriminant between SVT and VT, but it fails in the spontaneous episodes in



8. Spectral analysis of the QRS complexes of (a) the SVT and (b) the VT episodes of Base C (estimated with Parzen's windows). Note that the SVT is a bandpass process and the VT is a low-pass process.

Table 2. Areas Under the Curve for the Tested Arrhythmia Discrimination Criteria.

	Rate	CWA	MLP	Spectral	Width	AR	SVMm	SVMc	V2d	V3d
Base C	0.98	0.94	0.97	0.97	0.86	0.62	0.99	0.68	0.95	0.96
Base D	0.85	0.86	0.97	0.86	-	-	0.92	0.88	0.93	0.98
Individual	0.95	0.99	0.99	0.89	-	-	0.92	0.97	0.98	0.98

general, pointing either to a “lucky” choice of the single patient or to a high dependence on the patient of the decision threshold. The same situation is found for the CWA: in fact, for both algorithms, only the results in the independent group are coherent with the previously published work. The spectral analysis shows a high discrimination in the training set, but it fails in the validation and in the single patient set, which makes it a nongeneralizing scheme. Both the QRS width and the AR criteria were only evaluated in the control set, as their areas were low even in this group. The MLP classification led to a high discriminatory power in all of the sets. The IL-based schemes showed a high separability in the patient data; also, the second implementation exhibited a remarkably high area in all the sets, so that it is wholly comparable to the nonlinear MLP.

We can conclude that the MLP classifier and the second algorithmic implementation are the best implementations. The MLP is a global, nonlinear classifier, which allows it to trace general boundaries that the local RBF (and hence the SVM) cannot describe; however, it is still a clinically noninterpretable formulation. On the other hand, the second algorithmic implementation is a simple but effective and transparent criterion.

Conclusions

The SVM leads to an interpretable black-box model, which is an interesting situation in clinical applications. It has been successfully used for the design of an SVT/VT discrimination criterion, which can compete with a global, nonlinear learning machine and with other currently proposed algorithms. One of the drawbacks of the SVM is the uselessness of the sigmoidal kernel, which is an interesting field of research. Finally, further studies (exclusively focused on SVTs with bundle branch block) should corroborate these findings.

Acknowledgments

Thanks to M. Ortiz for her kind help with the database management and to F.J. Escribano for the useful comments and review of the manuscript.

José Luis Rojo-Álvarez received the bachelor's degree in telecommunication engineering from the Universidad de Vigo, Vigo, Spain, in 1996, and the Ph.D. in telecommunication from the Universidad Politécnica de Madrid, Madrid, Spain, in 2000. He is an assistant professor at the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Spain. His research interests include statistical learning theory, digital signal processing, and complex system modelling, focusing on EKG and intracardiac EGM signal processing, arrhythmia-genesis mechanisms, robust analysis of HRV, echocardiographic imaging, and hemodynamic function evaluation.

Ángel Arenal-Maíz received the Doctor en Medicina degree from the Universidad Complutense de Madrid in 1980, and he spent his residence term in the Hospital General Universitario Gregorio Marañón (HGU GM), Madrid, from 1981 to 1986. He received a cardiac electrophysiology fellowship at the Montreal Cardiology Institute in 1986-1987 and then joined the Arrhythmia Unit of HGU GM. During 1991 and 1992 he was working at the Basic Electrophysiology Laboratory with Dr. Stanley Natel. Dr. Arenal is now in charge of the Arrhythmia Unit of the HGU GM, Madrid, Spain. His research interests are focused on the automatic discrimination of cardiac arrhythmia in implantable devices, on the study of arrhythmia mechanisms, and on clinical cardiac electrophysiology.

Antonio Artés-Rodríguez was born in Alhama de Almería, Spain, in 1963. He received the Ingeniero de Telecomunicación and Doctor Ingeniero de Telecomunicación degrees, both from the Universidad Politécnica de Madrid, Spain, in 1988 and 1992, respectively. He is now

an associate professor at the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Spain. His research interests include detection, estimation, and statistical learning methods and their application to signal processing, communication, and biomedicine.

Address for Correspondence: José Luis Rojo Álvarez, Depto Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, 28911-Leganés, Madrid, Spain. Phone: +34 1 624 87 69. Fax: +34 1 624 87 49. E-mail: jlrojo@tsc.uc3m.es.

References

- [1] J.L. Rojo-Álvarez, A. Arenal-Maíz, and A. Artés-Rodríguez, “Discriminating between supraventricular and ventricular tachycardias from EGM onset analysis,” *IEEE Eng. Med. Biol. Mag.*, vol. 21, pp. 16-26, Jan./Feb. 2002.
- [2] C.J.C. Burges, “A tutorial on support vector machines for pattern recognition,” *Data Mining and Knowl. Discovery*, vol. 2, no. 2, vol. 1-32, 1998.
- [3] B. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1998.
- [4] A. Papoulis, *Probability Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1991.
- [5] A.D. Gordon, *Classification Methods for the Exploratory Analysis of Multivariate Data*. New York: Wiley, 1977.
- [6] S.A. Stevenson, J.M. Jenkins, and L.A. DiCarlo, “Analysis of the intraventricular electrogram for differentiation of distinct monomorphic ventricular arrhythmias,” *PACE*, vol. 20, pp. 2730-2738, 1997.
- [7] K. Minami, H. Nakajima, and T. Toyoshima, “Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network,” *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 179-185, 1999.
- [8] T. Klingenheben, C. Sticherling, M. Skupin, and S.H. Hohnloser, “Intracardiac QRS electrogram width. An arrhythmia detection feature for implantable cardioverter defibrillator. Exercise induced variation as a base for device programming,” *PACE*, vol. 8, no. 2, pp. 16-17, 1998.
- [9] C.J. Finelli, J.M. Jenkins, and L.A. DiCarlo, “Detection and identification of cardiac arrhythmias using an adaptive, linear-predictive filter,” in *Proc. IEEE Computers in Cardiology*. London, U.K., 1993, pp. 177-180.