A method for context-based adaptive QRS clustering in real-time

Daniel Castro*, Paulo Félix, and Jesús Presedo

Abstract—**Continuous follow-up of heart condition through long-term electrocardiogram monitoring is an invaluable tool for diagnosing some cardiac arrhythmias. In such context, providing tools for fast locating alterations of normal conduction patterns is mandatory and still remains an open issue. This work presents a real-time method for adaptive clustering QRS complexes from multilead ECG signals that provides the set of QRS morphologies that appear during an ECG recording. The method processes the QRS complexes sequentially, grouping them into a dynamic set of clusters based on the information content of the temporal context. The clusters are represented by templates which evolve over time and adapt to the QRS morphology changes. Rules to create, merge and remove clusters are defined along with techniques for noise detection in order to avoid their proliferation. To cope with beat misalignment, Derivative Dynamic Time Warping is used. The proposed method has been validated against the MIT-BIH Arrhythmia Database and the AHA ECG Database showing a global purity of 98.56% and 99.56%, respectively. Results show that our proposal not only provides better results than previous offline solutions but also fulfills real-time requirements.**

Index Terms—**Adaptive clustering, Electrocardiogram (ECG), Dominant Points, Dynamic Time Warping, QRS clustering.**

I. INTRODUCTION

NOWADAYS the surface electrocardiogram (ECG) is recognized as an invaluable tool for monitoring heart condition, since its analysis provides decisive information that can reveal critical deviations from normal cardiac behavior. Recent developments in mobile sensors and mobile computing have enabled new scenarios for continuous ECG monitoring as an inexpensive tool for the early detection of some cardiac events [1], especially in those cases where symptoms appear intermittently.

As the monitoring period increases, the interpretation task becomes more time consuming and decision-support tools are needed to help cardiologists to reduce the time spent on it. If a continuous follow-up is required, these tools become imperative. Their main aim is to provide the cardiologists with a summary of all the acquired signals, enhanced with a fast locating of those anomalies detected.

Cardiac arrhythmias are the most relevant among the ECG findings. There are two main sources of arrhythmias: an automatism disorder, that is, a set of alterations in the beat activation point due to changes in its location or activation

frequency; or a conduction disorder, that is, an abnormal propagation of the beat wavefront through the cardiac tissue. They both have an effect on the ECG, affecting the beat morphology and/or beat rhythm. In order to support their identification, a method for separating the beats by their activation point and conduction pattern should be provided.

Beat classification arises as the task of assigning each beat in an ECG a label identifying its physiological nature. Machine learning techniques have been applied to this task by estimating the underlying mechanisms that produce the data of a training set. The main drawback of this approach is its strong dependence on the pattern diversity present in the training set. Thus, inter-patient differences show that it cannot be assumed that a classifier trained on data of a large set of patients will yield valid results on a new patient [2]–[4], and intra-patient differences show that this cannot be assumed even for the same patient throughout time. In addition, class labels only provide gross information about the origin of the beats in the cardiac tissue, loosing all the information about their conduction pathways. This approach does not distinguish the multiple morphological families present in a given class, as occurs in multifocal arrhythmias.

In contrast, beat clustering aims at dividing the ECG recording in a set of beat clusters, each one of them preserving some similarity properties. Previous proposals have focused on an offline approach, from a priori maximum number of clusters [5]–[8] and they imply processing the ECG signal once the acquisition has been completed. This approach has given good noise robustness, but as a side effect a single morphology is usually replicated in several clusters and rare beat morphologies can be missed. It also omits the dynamic aspect of ECG and, in particular, ignores the temporal evolution of morphologies. Furthermore, the detection of critical events can be deferred too long to provide timely attention. For all these reasons, a dynamic online approach must be considered.

In this paper, we present a real-time method for adaptive beat clustering, with a potential application not only as a previous step for classification [9], but also as a summary about those beat morphologies present in a certain period, their temporal evolution and variability, or even to detect the presence of alternating morphologies. The proposed method emulates the experts behavior in exploiting the temporal context for assigning each new beat to the most appropriate cluster. To this end, clusters are continuously adapting to the temporal evolution of beat morphologies, and they can be dynamically created, merged or modified, resulting in a variable number of clusters.

Beat clustering requires extracting from the ECG a set of representative measurements for every beat. Bibliography
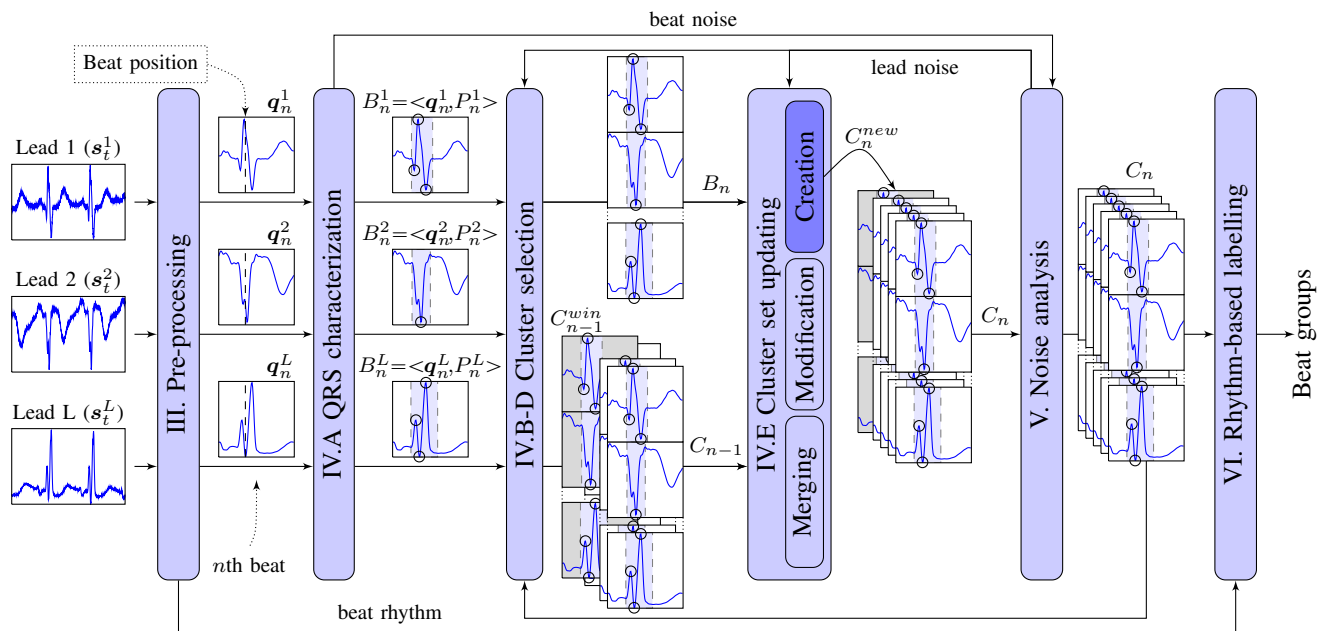
Fig. 1. Flow-chart of the method proposed for QRS clustering. The different stages involved in beat processing are shown and the creation of a new cluster is exemplified. Each block refers to the corresponding section.

shows a variety of proposals for beat representation that can be grouped into four categories: *morphological features* where the signal amplitude is directly used [2], [4], [7], [9]; *segmentation features* like area, amplitude or interval duration from beat waves [2], [4], [7], [10] or amplitude and angle values from vectorcardiogram [3]; *statistical features* derived using high order cumulants [4], [11], [12] and finally, *transformed space features*, defined in an alternative space using different transforms like Karhunen–Loewe transform [13], Hermite basis functions [4], [5], [12], discrete Fourier transform [9], [14] or wavelet transform [3], [14]. All the previous works complete their feature sets with rhythm information to complement their description capabilities.

The present paper proposes a new approach to represent a beat by reducing the QRS complex to a set of relevant points and support regions. This representation has some nice properties for beat clustering: it is stable against the usual variability and the presence of noise in the ECG, and it explicitly represents the temporal location of some QRS features.

The proposed method processes a real-time multilead ECG signal through a set of data-driven stages as shown in Fig. 1. In order to obtain comparable results, signals from two standard ECG signal databases are used as data source (section II). The pre-processing stage comprises real-time beat detection and baseline filtering (section III). Then, a fixed-length signal segment is selected for extracting and characterizing the QRS complex (subsection IV.A). QRS complexes are compared to the current set of clusters following a context-based criteria to obtain the best matching cluster (subsections IV.B–D). Afterwards, the current cluster set is updated in one of three ways: creating a new cluster, modifying the most similar one or merging two or more clusters (subsection IV.E). The next stage performs a noise analysis for each lead in order to detect noisy

intervals and avoid the processing of noisy beats or discard the clusters created from them (section V). Finally, the beats are classified by their rhythm type and a set of groups with common morphology and rhythm is obtained (section VI).

The ECG databases have been processed and their beat class labels have been used to validate the purity of the final cluster and group sets (section VII). These results are discussed in section VIII along with the conclusions of the work.

II. ECG SIGNAL DATABASES

The ECG databases recommended by the ANSI/AAMI EC57 [15] standard for reporting the performance of arrhythmia detectors were used for validation purposes: the MIT-BIH Arrhythmia Database and the AHA ECG database.

The MIT-BIH Arrhythmia Database [16] [17] can be referred to as the golden standard for beat clustering and classification tasks and it is the reference database for almost all the literature in this field. This database is composed of 48 recordings of ambulatory ECG, obtained from 47 different patients which comprise a very complete set of examples of common and rare arrhythmias. Each record has a duration of 30 min, and includes two channels with the same leads in almost all of them: a modified-lead II (MLII) in the first one and lead V1 in the second one. MLII was replaced by lead V5 in three records and V1 was replaced by MLII, V2, V4 or V5 on eight records. The signals were digitized at $f_s = 360$Hz and bandpass filtered with cutoff frequencies at 0.1 and 100Hz. All beats present in the database were annotated by at least two expert cardiologists, and assigned a class label using a 16 label set.

The AHA ECG database was compiled by the American Heart Association and it is composed of 155 recordings of ambulatory ECG digitized at $f_s = 250$Hz containing the most relevant types of ventricular arrhythmias. Each record is three

hours long with two channels but only the last 30 minutes have been manually annotated by experts.

III. PRE-PROCESSING

The major drawback of processing long-term ECG signals is the presence of a high level of noise with multiple manifestations —baseline wandering, power line interference or electromyographic activity— so an initial filtering stage needs to be performed. The efforts have been focused on filtering the baseline wandering, as this is the most relevant source affecting the reliability of the clustering algorithm because of the distortion it can cause on the QRS morphology.

A. Baseline filtering

The baseline filtering is performed through the estimation of the baseline wandering and its posterior elimination from the original ECG. To achieve this goal, each signal is processed by two sequentially connected median filters of 200ms and 600ms, respectively, as described in [2]. A global delay of 400ms is added by this process, independently of the sampling rate.

B. Beat detection

In order to carry out an evaluation of the clustering method separately from beat detector and provide a fair comparison framework for future algorithms, we used the beat position provided by the ECG databases as fiducial points.

A beat detector is required for real scenarios, affecting the quality of results. However, it virtually does not increment the global computational complexity as QRS detection represents only a small fraction of it. The global delay would not be affected either, since the beat detection can be performed concurrently with baseline filtering with a shorter delay.

IV. CLUSTERING

In this paper the following notation is used. Bold face variables (e.g., x) to represent vectors and sequences, lower case alphabets with subscripts to represent their components (e.g., x_i) or superscripts when a temporal index is used as a subscript (e.g., x_n^i), upper case alphabets (e.g., X) to represent sets, and calligraphic letters to represent functions (e.g., $\mathcal{F}$).

The aim of the clustering method is to group in real-time those beats which share the same activation area and propagation pattern in the cardiac tissue. The propagation of the electrical impulse through the heart is reflected in the ECG as a sequence of waves corresponding to the activity in the atria and ventricles. P wave and QRS complex are of particular interest since they show the atria and ventricular depolarization, respectively, thereby providing a fingerprint of the conduction pathway. Since the noise level of ambulatory signals makes the detection of the P wave a very difficult task, we focus on the QRS morphology to characterize the beat. As a consequence, beats with different atrial activation points or even with nodal activation points can share the same QRS morphology, and will be assigned to the same cluster. In the absence of a P wave analysis, rhythm information can be

useful in some cases to detect this situation, as explained in section VI.

The proposed method is intended to provide a real-time resume of the ECG signal as a dynamic set of clusters. Whenever a new beat is detected, the set of clusters is updated. Let C denote a time series $C = \{C_1, C_2, ..., C_n, ...\}$ where $C_n = \{C_n^c \mid 1 \leqslant c \leqslant N_n\}$ is the set of clusters that represents those QRS morphologies which appeared from the beginning of the recording up to the beat n, C_n^c is the set of beats assigned to cluster c and N_n the number of clusters, both at beat n.

The clustering follows a data-driven flow triggered by the beat detector as shown in Fig. 1. Each new beat is compared to the set of clusters C_n using concordance and dissimilarity measures in order to find the best matching and giving preference to those in its temporal context. A detailed explanation of this process follows in the remaining subsections: first, a beat characterization and a template-based representation of the clusters; next, the alignment technique and the measures used to compare beats and clusters; afterwards, the criteria for cluster selection and, finally, the process of updating the set of clusters.

A. QRS characterization

We adopt a strategy inspired by dominant point detection [18] to characterize the QRS morphology through its constituent waves. Let $S = \{s_t \mid s_t \in \mathbb{R}^L \wedge t \in \mathbb{N}\}$ denote a time series which represents the multilead ECG signal, where L is the number of leads and $s_t = (s_t^1, ..., s_t^L)$ is the vector of samples at time t for all leads. When the nth beat is detected with fiducial mark at time t, a fixed-length subsequence of $w = w^- + w^+$ samples (w^- before and $w^+ - 1$ after t) is selected from each lead l to represent its QRS complex (see Fig. 2):

$$\boldsymbol{q}_n^l = \{q_1^l, \ldots, q_j^l, \ldots, q_w^l\} \tag{1}$$

where $q_j^l = s_{t-w^--1+j}^l$. We set $w^- = \lceil 0.1 \times f_s \rceil$ and $w^+ = \lceil 0.2 \times f_s \rceil$, which are wide enough to capture the longest QRS of abnormal beats [19] (V, F and f beat types, typically).

Most of this section describes operations over one lead, so for the sake of notation simplicity the l superscript will be obviated unless multiple leads are involved.

We define the **curvature** $\mathcal{K}$ at $q_j \in \boldsymbol{q}_n$ as:

$$\mathcal{K}(q_j, \boldsymbol{q}_n) = \max_{i \in I_j^-, k \in I_j^+} \cos \widehat{q_i q_j q_k}, \tag{2}$$

where $2 \leqslant j \leqslant w-1$. The terms I_j^- and I_j^+ denote the intervals used for calculating the curvature, and they are given by:

$$I_j^- = \{i \mid (j-\theta) \leqslant i < j \ \wedge$$
$$\forall a \in (i,j), (\max_{b \in (a,j)} \Delta q_{j,b} - \Delta q_{j,a}) < \rho_{min}\} \tag{3}$$

$$I_j^+ = \{k \mid j < k \leqslant (j+\theta) \ \wedge$$
$$\forall a \in (j,k), (\max_{b \in (j,a)} \Delta q_{j,b} - \Delta q_{j,a}) < \rho_{min}\} \tag{4}$$

where $\Delta q_{j,x} = |q_j - q_x|$. The term θ is the maximum physiologically meaningful width of a QRS wave between its peak location and its left or right end, so that it is the upper limit
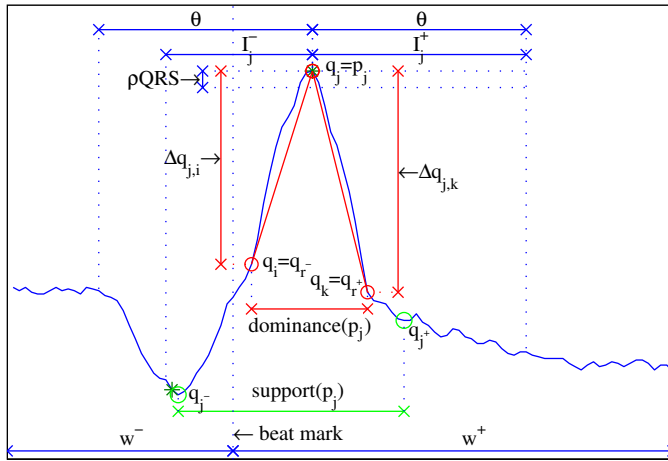
Fig. 2. Example of a subsequence q_n^l for lead $l = 2$, of a beat at sample 4102, $n = 13$, from record 108 of the MIT-BIH database. The relevant points detected by the algorithm are marked with *. The parameters involved in the relevant point detection are also shown.

of I_j^- and I_j^+. The term ρ_{min} is the minimum height for a signal deflection to be considered physiologically relevant and, therefore, to be excluded from the calculation of curvature.

We define the **dominance region** of a sample q_j as:

$$dominance(q_j) = [r^-, r^+], \qquad (5)$$

where $r^- = \min(\arg\max_{i \in I_j^-} \cos \widehat{q_i q_j q_k})$ for any fixed $k \in I_j^+$ and $r^+ = \max(\arg\max_{k \in I_j^+} \cos \widehat{q_i q_j q_k})$ for any fixed $i \in I_j^-$.

We define the set of **dominant points** of q_n as:

$$D_n = \{p_j \,|\, p_j = q_j \wedge j = \arg\max_{a \in dominance(q_j)} \mathcal{K}(q_a, \boldsymbol{q}_n) \wedge \Delta q_j > \rho_{min}\}, \qquad (6)$$

where $\Delta q_j = \min(\Delta q_{j,r^-}, \Delta q_{j,r^+})$.

We define the set of **relevant points** of q_n as:

$$R_n = \{p_j \,|\, p_j \in D_n \wedge \Delta q_j > \rho_{QRS}\} \qquad (7)$$

where ρ_{QRS} is the minimum height for a signal deflection to be considered a relevant QRS wave, with $\rho_{QRS} > \rho_{min}$. If $R_n = \varnothing$, then it is redefined as: $R_n = \{\arg\max_{p_j \in D_n} \Delta q_j\}$.

The limits of $dominance(q_j)$ can be located at any point in the edge of a wave. Since we are interested in capturing its full extent, the **support region** of a relevant point $p_j \in R_n$ is defined as:

$$support(p_j) = [j^-, j^+] \qquad (8)$$

where $j^- \leqslant r^-$ and $j^+ \geqslant r^+$ are the sample numbers nearest to r^- and r^+ where slope sign changes:

$$j^- = \min_{i \in I_j^-}(\{i \,|\, i \leqslant r^- \wedge \forall a \in [i, r^-], \Delta q_{j,a} > \Delta q_{j,a+1}\}) \qquad (9)$$

$$j^+ = \max_{k \in I_j^+}(\{k \,|\, k \geqslant r^+ \wedge \forall a \in [r^+, k], \Delta q_{j,a} < \Delta q_{j,a+1}\}). \qquad (10)$$

In consequence, adjacent support regions can now overlap.

A relevant point p_j is said to be in a concave wave if $q_j > q_{j-}$ and $q_j > q_{j+}$. Otherwise, it is said to be in a convex wave. The wave height is defined as $\Delta p_j = \min(\Delta q_{j,j^-}, \Delta q_{j,j^+})$.

Finally, the nth beat is represented by the QRS signal segment and the set of its relevant points and support regions $P_n = \{(p_j, support(p_j)) \,|\, p_j \in R_n\}$ and denoted by:

$$B_n = <\boldsymbol{q}_n, P_n> . \qquad (11)$$

There is not a consensus in the literature about the limits for width and height of QRS complex or individual QRS waves. The AAMI standard [20] recommends a minimum amplitude of $50\mu V$ and duration of 10ms for a QRS wave to be detected, and $150\mu V$ for peak-to-peak QRS amplitude, with a minimum duration of $70ms$. On the other hand, the AHA [21] and CSE [19] report lower amplitude for QRS waves (down to $20\mu V$ and 10ms), based on measures over averaged beats with increased Signal-to-Noise ratio. These limits were not established for physiological reasons, but for signal noise level or instrumentation limitations. Nothing is stated about maximum QRS width beyond a reference to case-based duration values (e.g. the CSE study [19] reports a maximum QRS width of 210ms). In our case, due to the Signal-to-Noise ratio present in the ambulatory signals, the value of ρ_{QRS} is set to $150\mu V$ in order to avoid the detection of small waves caused by noise and the value of θ is set to 100ms to accept QRS waves with a maximum width of 200ms. The value of ρ_{min} is set to $50\mu V$ following the AAMI standard [20] and will be useful to detect noise contaminated QRS complexes.

In order to perform the comparison between a new beat B_n and the current set of clusters C_{n-1}, each cluster C_{n-1}^c is represented by a template:

$$T_{n-1}^c = <\boldsymbol{q}_{n-1}^c, P_{n-1}^c>, \qquad (12)$$

where $\boldsymbol{q}_{n-1}^c = \{q_1^c, \ldots, q_w^c\}$ is derived from the QRS of the beats assigned to C_{n-1}^c as will be explained in section IV-E.

B. QRS temporal alignment

In order to compare a beat B_n with a cluster template T_{n-1}^c, a temporal alignment of $\boldsymbol{q}_n$ and $\boldsymbol{q}_{n-1}^c$ is performed using Dynamic Time Warping (DTW) [22]. The aim of this alignment process is twofold. First, to correct any temporal misalignment due to a misplaced fiducial mark. And second, to reduce the contribution of the height and width variability of the QRS waves to the dissimilarity measure.

DTW was previously used for this purpose in [6], providing a relation $\boldsymbol{m} = (m_1, ..., m_K)$ between $\boldsymbol{q}_n$ and $\boldsymbol{q}_{n-1}^c$ called warping path, with $m_k = (x_k, y_k) \in [1, w] \times [1, w], k \in [1, K]$ and $K \geqslant w$. Each m_k represents the alignment of the index x_k of $\boldsymbol{q}_n$ with the index y_k of $\boldsymbol{q}_{n-1}^c$ under three conditions: first, $m_1 = (1, 1)$ and $m_K = (w, w)$; second, $x_i \leqslant x_k$ and $y_i \leqslant y_k \,\forall i < k$; and third, $m_{k+1} - m_k \in \{(1, 1), (1, 0), (0, 1)\}$. These conditions preserve the time-ordering of points and prevent some point being missed in the alignment.

Let $\mathcal{G}$ denote the cost function associated to a warping path defined by $\mathcal{G}(\boldsymbol{m}) = \sum_{k=1}^{K} \mathcal{G}_l(x_k, y_k)$, where $\mathcal{G}_l(x, y) = |q_x - q_y^c|$ is the local cost function associated with each element of $\boldsymbol{m}$. The optimal warping path is the one that minimizes $\mathcal{G}$.

DTW aligns the original signal samples, so any component k from $\boldsymbol{m}$ where $m_{k+1} - m_k \in \{(1, 0), (0, 1)\}$ introduces a

(a) Non-aligned QRS complexes   (b) Non-aligned derivatives

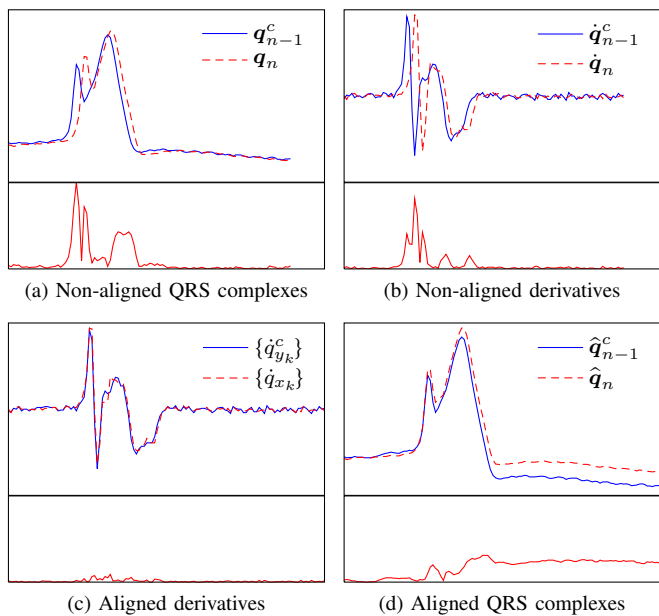(c) Aligned derivatives   (d) Aligned QRS complexes

Fig. 3. Example of the DDTW alignment process. For each subfigure, the upper solid and dashed lines represent the beat and template data, respectively; bottom solid lines represent their absolute difference. (a) Shows q_n and q_{n-1}^c subsequences; (b) $\dot{q}_n$ and $\dot{q}_{n-1}^c$ derivative approximations; (c) optimally-aligned derivatives of length $K \geqslant w-1$; (d) subsequences $\hat{q}_n$ and $\hat{q}_{n-1}^c$ of length $K+1$ obtained from the optimally-aligned derivatives. In (b) and (c), the difference is equivalent to the local cost function $\mathcal{G}_l(x,y) = |\dot{q}_x - \dot{q}_y^c|$. Notice the temporal increment $K-w$ in (c) and (d) as a result of the process. The absolute difference may also be increased in (d) outside the support regions, but this is irrelevant.

horizontal segment into the aligned signals. Since this can lead to unacceptable distortion of the QRS, we adopt the Derivative Dynamic Time Warping (DDTW) [23] which uses the estimation of the derivative instead of the signal itself. The derivative is approximated by the first difference: $\dot{q}_n = (\dot{q}_1, ..., \dot{q}_{w-1})$ and $\dot{q}_{n-1}^c = (\dot{q}_1^c, ..., \dot{q}_{w-1}^c)$ with $\dot{q}_x = q_{x+1} - q_x$ and $\dot{q}_y^c = q_{y+1}^c - q_y^c$.

We imposed some additional conditions on the selected warping path so as to restrict the alterations of aligned signals. A global restriction is set in the search process by defining a warping window δ (named Sakoe-Chiba band [22]) to limit the temporal distance between aligned samples so $|x_k - y_k| < \delta$. We set $\delta = 5$, which corresponds to a distance of 14ms with $f_s = 360$Hz, and is long enough to deal with small misalignments of beat marks. A local restriction is also used to limit the number of times the same sample can be aligned, setting a slope constraint: $m_{k+a} - m_k \notin \{(0,a),(a,0)\}, \forall a > \lambda$. We set $\lambda = 2$ to limit the variation in the amplitude of the aligned signals. Both conditions together allows the DDTW to cancel out wave differences up to three times in amplitude and up to δ samples in width.

Finally, after the optimal warping path m is found, the aligned signals $\hat{q}_n$ and $\hat{q}_{n-1}^c$ are obtained with coordinates $\hat{q}_{k+1} = \hat{q}_k + \dot{q}_{x_k}$ and $\hat{q}_{k+1}^c = \hat{q}_k^c + \dot{q}_{y_k}^c$ for $k \in [1,K]$, where $\hat{q}_1 = q_1$ and $\hat{q}_1^c = q_1^c$. Fig. 3 shows the result of the alignment of the current QRS complex and a template.

C. Template matching

In order to assign a beat B_n to an existent cluster C_{n-1}^c, we design a similarity calculation that only considers the

difference between signals over the support region of each relevant point, thus limiting the comparison to the constituent waves of the QRS. Given a pair $(p_j, [j^-, j^+]) \in P_n$, we are interested in verifying whether T_{n-1}^c contains a similar wave in the same location of the QRS. In order to do so, the interval $[\breve{j}^-, \breve{j}^+]$ of q_{n-1}^c aligned with the interval $[j^-, j^+]$ of q_n must be obtained (see Fig. 4). To this end, the warping path m is used to map j and $[j^-, j^+]$ from q_n into $\hat{q}_n$ obtaining the equivalent index $\hat{j}$ for the relevant point, and $[\hat{j}^-, \hat{j}^+]$ for its support region where: $\hat{j} = \max\{k \mid x_k = j\}$, $\hat{j}^- = \min\{k \mid x_k = j^-\}$ and $\hat{j}^+ = \max\{k \mid x_k = j^+\}$, being $(x_k, y_k) \in m$. Afterwards, since $\hat{q}_{n-1}^c$ and $\hat{q}_n$ are already aligned by the application of DDTW, the same interval $[\hat{j}^-, \hat{j}^+]$ from $\hat{q}_{n-1}^c$ is selected and mapped into q_{n-1}^c using the warping path. The interval $[\breve{j}^-, \breve{j}^+]$ is given by $\breve{j}^- = y_{\hat{j}^-}$ and $\breve{j}^+ = y_{\hat{j}^+}$.

Once the aligned intervals are obtained we proceed to evaluate the concordance of both vectors. The segment q_{n-1}^c is said to **concord with** q_n at p_j and denoted by $q_{n-1}^c \approx_{p_j} q_n$ if q_{n-1}^c contains a deflection in $[\breve{j}^-, \breve{j}^+]$ with height $\Delta p_j^c > \rho_{min}$ likely to be considered a significant waveform, where $\Delta p_j^c = \min(|q_{peak}^c - q_-^c|, |q_{peak}^c - q_+^c|)$ being $peak = \arg\min_{i \in [\breve{j}^-, \breve{j}^+]} q_i^c$, $q_-^c = \max_{i \in [\breve{j}^-, peak]} q_i^c$ and $q_+^c = \max_{i \in [peak, \breve{j}^+]} q_i^c$ for p_j in a convex wave (see Fig. 4a). Otherwise, min and max functions are interchanged.

We define the **concordance ratio** of q_{n-1}^c with respect to q_n at p_j, denoted by $\mathcal{C}_{p_j}(q_{n-1}^c, q_n)$, as:

$$\mathcal{C}_{p_j}(q_{n-1}^c, q_n) = \frac{\min(\Delta p_j, \Delta p_j^c)}{\max(\Delta p_j, \Delta p_j^c)} \quad (13)$$

if $q_{n-1}^c \approx_{p_j} q_n$. Otherwise, $\mathcal{C}_{p_j}(q_{n-1}^c, q_n) = 0$.

We define the **local dissimilarity** of q_{n-1}^c with respect to q_n at p_j, and denoted by $\mathcal{D}_{p_j}(q_{n-1}^c, q_n)$, as a weighted relative area difference between $\hat{q}_{n-1}^c$ and $\hat{q}_n$ at the interval $[\hat{j}^-, \hat{j}^+]$:

$$\mathcal{D}_{p_j}(q_{n-1}^c, q_n) = \left(\frac{(\Delta A_j^-)^2}{A_j^-} + \frac{(\Delta A_j^+)^2}{A_j^+} \right) \times \frac{1}{A_j^- + A_j^+} \quad (14)$$

The terms ΔA_j^- and ΔA_j^+ represent the areas under $a = |\hat{q}_n - \hat{q}_{n-1}^c|$ over the intervals $[\hat{j}^-, \hat{j}]$ and $[\hat{j}, \hat{j}^+]$, respectively (see Fig. 4b). Computing the area at each side of $\hat{j}$ independently allows us to minimize the effect of vertical misalignment or amplitude variation on each interval by subtracting their own median. Using the trapezoidal rule:

$$\Delta A_j^- = \sum_{k \in (\hat{j}^-, \hat{j})} a_k + \frac{1}{2}(a_{\hat{j}^-} + a_{\hat{j}}) - (\hat{j} - \hat{j}^-)M^- \quad (15)$$

$$\Delta A_j^+ = \sum_{k \in (\hat{j}, \hat{j}^+)} a_k + \frac{1}{2}(a_{\hat{j}} + a_{\hat{j}^+}) - (\hat{j}^+ - \hat{j})M^+ \quad (16)$$

where $M^- = \text{median}_{k \in [\hat{j}^-, \hat{j}]} a_k$ and $M^+ = \text{median}_{k \in [\hat{j}, \hat{j}^+]} a_k$.

The terms A_j^- and A_j^+ represent the areas below $[\hat{j}^-, \hat{j}]$ and $[\hat{j}, \hat{j}^+]$ intervals of $\hat{q}_n$, respectively (see Fig. 4c). They
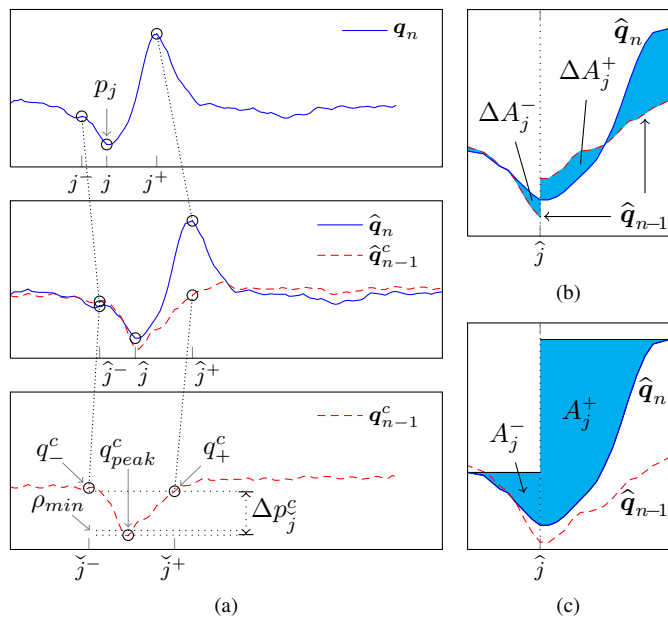
Fig. 4. Parameters involved in calculating the concordance and local dissimilarity of a cluster with respect to a beat at a relevant point p_j. Solid and dashed lines represent beat and template data, respectively. (a) Shows parameters involved in concordance checking; the support region $[j^-, j^+]$ and the derived intervals $[\hat{j}^-, \hat{j}^+]$ and $[\check{j}^-, \check{j}^+]$ are drawn; (b) signal segments within the interval $[\hat{j}^-, \hat{j}^+]$ with a vertical alignment performed in the subintervals $[\hat{j}^-, \hat{j}]$ and $[\hat{j}, \hat{j}^+]$ independently; ΔA_j^- and ΔA_j^+ areas are shaded; (c) the same signal segments shown in (b) with A_j^- and A_j^+ areas shaded.

are estimated using the trapezoidal rule:

$$A_j^- = \left| \sum_{k \in (\hat{j}^-, \hat{j})} \widehat{q}_k + \frac{1}{2}(\widehat{q}_{\hat{j}^-} + \widehat{q}_{\hat{j}}) - (\hat{j} - \hat{j}^-)\widehat{q}^- \right| \quad (17)$$

$$A_j^+ = \left| \sum_{k \in (\hat{j}, \hat{j}^+)} \widehat{q}_k + \frac{1}{2}(\widehat{q}_{\hat{j}} + \widehat{q}_{\hat{j}^+}) - (\hat{j}^+ - \hat{j})\widehat{q}^+ \right| \quad (18)$$

where $\widehat{q}^- = \max_{k \in [\hat{j}^-, \hat{j}]} q_k$ and $\widehat{q}^+ = \max_{k \in [\hat{j}, \hat{j}^+]} q_k$ for p_j in a convex wave. Otherwise, $\max$ is replaced by $\min$.

We define the **piecewise similarity** of q_{n-1}^c with respect to q_n, denoted $\mathcal{PS}(q_{n-1}^c, q_n)$, as the sum of two bounded contributions from the set of concordant and non-concordant relevant points:

$$\mathcal{PS}(q_{n-1}^c, q_n) = \sum_{p_j \in R_n} \mathcal{C}_{p_j}(q_{n-1}^c, q_n) \cdot \text{sig}(\mathcal{D}_{p_j}(q_{n-1}^c, q_n))$$
$$- \max_{p_j \mid q_{n-1}^c \not\approx_{p_j} q_n} \mathcal{D}_{p_j}(q_{n-1}^c, q_n) \quad (19)$$

where the sigmoid function $\text{sig}(x) = 1 - \alpha x / \sqrt{1 + (\alpha x)^2}$ keeps the contribution of each point in the interval $[0, 1]$. The parameter α determines the decrease rate of the contribution as the local dissimilarity increases and, as a consequence, the weight of a high dissimilarity value in the final piecewise dissimilarity. We limit the contributions of those points out of a range of admissible local dissimilarity. In order to set such range we consider the effect of amplitude and temporal variability of QRS waves. A temporal misalignment of one signal sample between the QRS waves of $\widehat{q}_n$ and $\widehat{q}_{n-1}^c$ can

lead to local dissimilarities of around 20%. Then we set an interval of $[0, 25\%]$ with a slightly greater upper bound and $\alpha = 4$ so the maximum contribution out of this interval is 0.30.

The previous measure is asymmetric since it depends on the relevant points and areas of one signal, so we define the **similarity** between B_n and T_{n-1}^c as:

$$\mathcal{S}(q_n, q_{n-1}^c) = \mathcal{PS}(q_{n-1}^c, q_n) + \mathcal{PS}(q_n, q_{n-1}^c) \quad (20)$$

thus obtaining a value which captures the concordance, similarity and morphological complexity of both signal segments.

This measure allows us to select the most similar template within a set, but we need a reference scale to evaluate the degree of matching. To this end, we define the **normalized piecewise similarity** as:

$$\overline{\mathcal{PS}}(q_{n-1}^c, q_n) = \mathcal{PS}(q_{n-1}^c, q_n)/|R_n| \quad (21)$$

and the **normalized similarity** as:

$$\bar{\mathcal{S}}(q_n, q_{n-1}^c) = \mathcal{S}(q_n, q_{n-1}^c)/(|R_n| + |R_{n-1}^c|), \quad (22)$$

where R_{n-1}^c is the set of relevant points of q_{n-1}^c.

D. Cluster selection

The occurrence of different QRS morphologies in a segment of ECG signal is usually limited by a reduced set of activation points and conduction pathways. Thereby we expect that the majority of QRS complexes in an ECG recording share their morphology with some of the QRS complexes present in a short previous temporal interval. For that reason, the search for the cluster $C_{n-1}^{win} \in C_{n-1}$ that best matches a beat B_n is first performed in the set of clusters present in its temporal context $C_{n-1}^{ctx} \subset C_{n-1}$ (see Fig. 5). We define the temporal context as the set of τ beats previous to B_n, $\tau\text{-}ctx^-(B_n) = \{B_{n-i} \mid 1 \leqslant i \leqslant \tau\}$ and $C_{n-1}^{ctx} = \{C_{n-1}^c \in C_{n-1} \mid \exists B_i \in \tau\text{-}ctx^-(B_n) \wedge B_i \in C_{n-1}^c\}$. The context length is set to $\tau = 15$ beats, the number of beats displayed in the typical 10s ECG strip used by cardiologists for a heart rate of 80 beats/min. This context is long enough to include every QRS morphology present in multifocal arrhythmias. Throughout this section, the l superscript is used to denote the lead.

The similarity measure is used to identify the most similar template for each lead as $sim^l = \arg\max_c \mathcal{S}(q_n^l, q_{n-1}^{c,l})$ and then the best matching cluster is obtained by majority vote as $sim = \text{mode}\{sim^l \mid l \in [1, L]\}$. If multiple clusters are selected, a second vote is performed to obtain a single one using the normalized similarity.

Beat B_n is assigned to C_{n-1}^{sim} if the condition $\bar{\mathcal{S}}(q_n^l, q_{n-1}^{sim,l}) > \gamma$ is fulfilled in all leads. Then $C_{n-1}^{win} = C_{n-1}^{sim}$. We set a value of $\gamma = 0.30$ which corresponds to the maximum contribution of a point with local dissimilarity outside the admissible interval.

When B_n and C_{n-1}^{sim} are not similar enough, a new comparison is performed within the subset $C_{n-1} - C_{n-1}^{ctx}$, obtaining the most similar cluster C_{n-1}^{sim*}. If B_n and C_{n-1}^{sim*} are similar enough, the beat is assigned and $C_{n-1}^{win} = C_{n-1}^{sim*}$. Otherwise, the beat is not assigned to any existing cluster and its most similar cluster C_{n-1}^{win} is selected between C_{n-1}^{sim} and C_{n-1}^{sim*} using the same voting criteria previously seen.
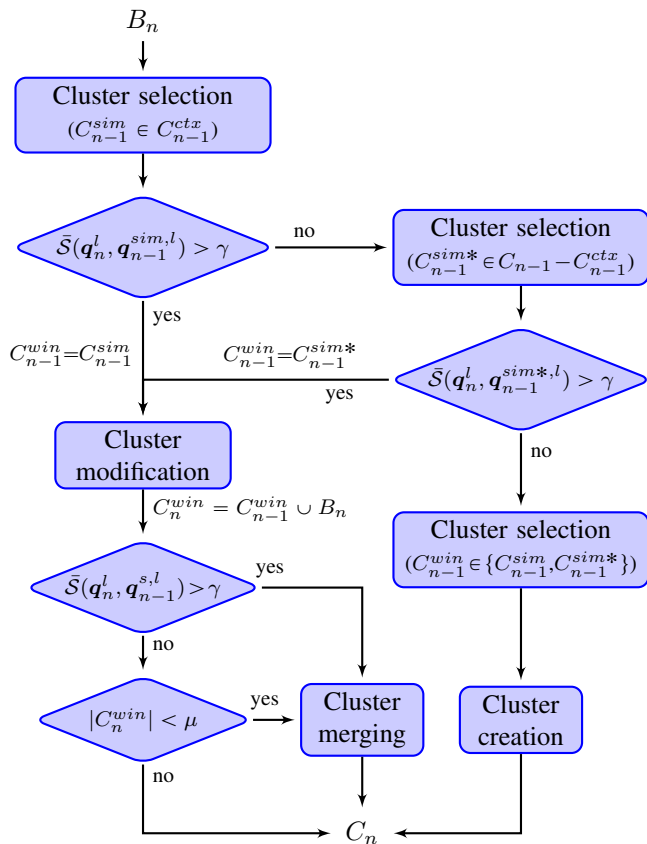
Fig. 5. Flow-chart of the cluster selection and cluster set updating processes.

E. Cluster set updating

In order to adaptively respond to the changing behavior of the ECG, clusters must be dynamically created, modified or merged whenever a new beat is detected:

1) Cluster creation: If B_n is not assigned to C_{n-1}^{win}, a new cluster C_n^{new} is created and its template is initialized for each lead using the beat representation: $T_n^{new,l} = B_n^l$. Then the cluster set is updated as $C_n = C_{n-1} \cup \{C_n^{new}\}$.

2) Cluster modification: If B_n is assigned to C_{n-1}^{win}, then the cluster is updated to $C_n^{win} = C_{n-1}^{win} \cup \{B_n\}$ and the template $T_{n-1}^{win,l} = <\boldsymbol{q}_{n-1}^{win,l}, P_{n-1}^{win,l}>$ is modified to $T_n^{win,l}$ for each lead. To this end, $\boldsymbol{q}_n^{win,l}$ is calculated from $\boldsymbol{q}_n^l$ and $\boldsymbol{q}_{n-1}^{win,l}$:

$$\dot{q}_{y,n}^{win,l} = (1-\beta)\dot{q}_{y,n-1}^{win,l} + \beta \operatorname{mean}(\{\dot{q}_{x,n}^l \,|\, (x,y) \in \boldsymbol{m}\}) \quad (23)$$

$$q_{y,n}^{win,l} = q_{y-1,n}^{win,l} + \dot{q}_{y,n}^{win,l} \quad (24)$$

where $q_{1,n}^{win,l} = q_{1,n-1}^{win,l}$. The term β is the coefficient of the exponential update. Setting a value for β implies a trade-off between plasticity and stability of the cluster template. We set $\beta = 1/8$ so the last 16 beats assigned to the cluster provide 90% of the contributions to the current template. This allows the template to be adapted to the evolution of the QRS morphology. Afterwards, the set of relevant points and support regions $P_n^{win,l}$ is obtained from $\boldsymbol{q}_n^{win,l}$ and assigned to the template: $T_n^{win,l} = <\boldsymbol{q}_n^{win,l}, P_n^{win,l}>$.

3) Cluster merging: During the clustering process, different clusters can evolve to represent the same QRS morphology, so they should be merged. This situation is common when

QRS complexes suffer from transient distortions in their morphology due to intrinsic variability, which makes them fall below the similarity threshold for their proper clusters. In this case, a new cluster is created which is subjected to an initial transient period during which the template for each lead can evolve getting closer to its most similar cluster.

In order to detect this situation, we define a relation *closest* which links each cluster with its most similar one among previous clusters. The relation is set for each new cluster as $C_n^{win} = closest(C_n^{new})$. As templates evolve with new assigned QRS complexes, the relation could have to be updated.

When B_n is assigned to C_{n-1}^{win}, the cluster is updated to C_n^{win} and its most similar cluster may change. The *closest* relation is updated when multiple clusters in the same set than C_{n-1}^{win}, be it either C_{n-1}^{ctx} or $C_{n-1} - C_{n-1}^{ctx}$, fulfill the assigment condition $\bar{S}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l}) > \gamma$ in all leads (see subsection IV-D). Since C_n has been updated, getting more similar to B_n, the best matching cluster C_{n-1}^s for B_n is selected within the remaining subset of clusters, where $s = \arg \max_{C_{n-1}^c \neq C_{n-1}^{win}} (\sum_{l=1}^L \mathcal{S}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l})/L)$ and the oldest of both clusters, C_{n-1}^{win} or C_{n-1}^s, is set as the new most similar cluster for the other. Then the clusters C_n^{win} and C_n^s are checked for possibly merging, since their templates have evolved to be similar enough to the last beat (see Fig. 5). The condition for merging those clusters will be analogous to the condition for beat assignment: $\bar{S}(\boldsymbol{q}_n^{s,l}, \boldsymbol{q}_n^{win,l}) > \gamma'$, where $\gamma' = 0.40$, which corresponds to the maximum contribution of a point with local dissimilarity over 20%. The threshold value is increased with respect to γ since both signal segments are promediated templates with increased Signal-to-Noise ratio.

Afterwards, the special case of clusters within its transient period is considered. We establish the duration of this transitory state in terms of the number of assigned beats. Hence, if B_n is assigned to C_{n-1}^{win} and $|C_n^{win}| < \mu$, the cluster is checked for merging with its closest one $C_n^s = closest(C_n^{win})$ (see Fig. 5). We set μ as the minimum number of beats assigned to the cluster to confirm it represents an independent morphology and we consider $\mu = 10$ enough for this purpose.

When two clusters C_n^{win} and C_n^c are merged, the cluster set, cluster template and *closest* relation are updated accordingly (we suppose that $C_n^c = closest(C_n^{win})$):

1) C_n^c is updated to $C_n^c = C_n^c \cup C_n^{win}$.
2) $\boldsymbol{q}_n^{c,l}$ is calculated by merging $\boldsymbol{q}_{n-1}^{c,l}$ and $\boldsymbol{q}_n^{win,l}$:

$$\dot{q}_{y,n}^{c,l} = (1-\beta)\dot{q}_{y,n}^{c,l} + \beta \sum(\{\dot{q}_{x,n}^{win,l} \,|\, (x,y) \in \boldsymbol{m}\}) \quad (25)$$

$$q_{y,n}^{c,l} = q_{y-1,n}^{c,l} + \dot{q}_{y,n}^{c,l} \quad (26)$$

where $q_{1,n}^{c,l}$ remains unmodified.
3) The template is modified to $T_n^{c,l} = <\boldsymbol{q}_n^{c,l}, P_n^{c,l}>$, where $P_n^{c,l}$ is the set of relevant points and support regions obtained of $\boldsymbol{q}_n^{c,l}$.
4) The *closest* relation is updated by removing the (C_n^c, C_n^{win}) pair and replacing C_n^{win} by C_n^c in all the pairs where the former appears.

After the cluster C_n^{win} gets merged, all the modified pairs of the *closest* relation are checked for merging. Afterwards, the C_n^c cluster is also checked for merging with its closest one.

V. Noise-cluster proliferation control

The dynamic creation of clusters entails the problem of identifying those QRS complexes contaminated by noise that could cause the proliferation of clusters. When noise appears in an ECG lead, the QRS can show different changes. In some cases, ECG segments with low Signal-to-Noise ratio present a high number of waves, and these can be detected by an abnormal number of dominant points. In other cases, domain knowledge is needed to discern if changes respond to a new QRS morphology or to a noisy version of a previous one. Some alterations are well known and described in literature, but others can be challenging even for an expert cardiologist, who compares every beat with those present in its temporal context in order to identify if it is related to a repetitive morphology change, a noisy interval or an isolated noisy beat. We follow a twofold, beat-based and context-based, approach to detect noisy QRS complexes. Noise can be present in one or more leads, and the influence of each lead in the clustering of a given beat will be analyzed in the following.

A state variable $lead\_noise_n^l$ is defined to denote the existence of a *noisy interval* in lead l containing the nth beat. Such noisy interval begins when the first noisy beat is detected in that lead just after a sequence of previous noise-free beats. The beats contained within this interval can be either noisy or noise-free in l and such condition will be represented as an attribute of the beat denoted by $beat\_noise_n^l$. During the noisy interval, the characterization of any new beat B_n can represent not only QRS waves but also noise artifacts, so the cluster selection and assignment rules (subsection IV-D) are modified in lead l: the dominant points of the beat are ignored, replacing $\mathcal{S}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l})$ by $\mathcal{PS}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l})$ and $\bar{\mathcal{S}}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l})$ by $\overline{\mathcal{PS}}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{c,l})$. Therefore, the condition for beat assignment is: $\overline{\mathcal{PS}}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{win,l}) > \gamma$.

The ending of a noisy interval in lead l is set just before $\kappa = 3$ contiguous beats are considered noise-free in that lead. A beat B_n is considered *noise-free* if $beat\_noise_n^l = false$ and it is assigned either to its winner cluster with $\bar{\mathcal{S}}(\boldsymbol{q}_n^l, \boldsymbol{q}_{n-1}^{win,l}) > \gamma$ or to a new cluster.

The *beat-based noise detection* is triggered when a new beat B_n is detected. Then the number of waves in the QRS is estimated as $|D_n^l|$ for each lead l. If $|D_n^l| > \eta$, B_n is considered noisy in lead l and $beat\_noise_n^l$ is set to $true$. The term η is the maximum number of waves that could be present in a noise–free QRS. We set $\eta = 6$ to admit complexes with Q, R and S waves, mixed with R', S' or spikes as occurs in paced, fusion or bundle branch block beats. Additionally, if $|R_n^l| > \eta$, the QRS complex is considered distorted with relevant waves caused by noise. If this happens in some but not all leads, such leads are ignored for cluster assignment and updating, but if it happens in all leads, the beat is considered failed and it is assigned to its most similar cluster.

The *context-based noise detection* is activated when a new cluster is created for B_n. Then two possible explanations are explored. A first explanation represents a hypothesis of noise, and every noisy beat will be assigned to its most similar cluster. A second explanation simply represents a change in morphology, so creation of new clusters is allowed. The evalu-

ation of these hypothesis is performed over a temporal context of τ beats defined as $\tau-ctx^+(B_n) = \{B_{n+i} \,|\, 0 \leqslant i < \tau\}$ which is long enough to check the evolution of cluster diversity.

A hypothesis of noise, denoted by $hyp\_noise_{n+i}^l$, is set over each lead for every beat $B_{n+i} \in \tau-ctx^+(B_n)$. The hypothesis is initialized using the current noise state for the first beat, $hyp\_noise_n^l = lead\_noise_{n-1}^l$, and the value of the previous beat for each new one, $hyp\_noise_{n+i}^l = hyp\_noise_{n+i-1}^l$. This value can be further modified for a beat B_{n+i} under three circumstances:

- If B_{n+i} is considered noisy in lead l by the beat-based noise detection, $beat\_noise_{n+i}^l = true$, then the hypothesis is set to $hyp\_noise_{n+i}^l = true$.
- If a new cluster is created for B_{n+i}, we define L_{n+i}^{noise} as the set of leads with $\bar{\mathcal{S}}(\boldsymbol{q}_{n+i}^l, \boldsymbol{q}_{n+i-1}^{win,l}) \leqslant \gamma$, which are responsible for its creation and candidates to be set as noisy. For these leads, we check the assignment under noisy interval conditions, $\overline{\mathcal{PS}}(\boldsymbol{q}_{n+i}^l, \boldsymbol{q}_{n+i-1}^{win,l}) > \gamma$. If the assignment condition is now fulfilled in all leads, the signal still resemble a previous QRS morphology and we set $hyp\_noise_{n+i}^l = true, \forall l \in L_{n+i}^{noise}$.
- If B_{n+i} is the last of κth consecutive noise-free beat, we set $hyp\_noise_{n+i-j}^l = false, \forall j \in [0, \kappa-1]$.

Once the τ beats have been processed, the detection of noisy leads is addressed. Let $C_{n+\tau-1}^{ctx-new}$ denote the set of new clusters created in $\tau-ctx^+(B_n)$. If $|C_{n+\tau-1}^{ctx-new}| > \tau/3$ we consider that an abnormal cluster proliferation exists since the appearance of more than 5 new QRS morphologies within an interval of 15 beats is an extremely rare condition. Hence any lead l that is a common cause of creation, that is, $l \in L_{n+i}^{noise}$, for all i such that $B_{n+i} \in C_{n+i}^c \wedge C_{n+i}^c \in C_{n+\tau-1}^{ctx-new}$, is discarded as noisy, setting $hyp\_noise_{n+i}^l = true, \forall i \in [0, \tau-1]$.

Finally, the hypothesis is reviewed for every beat $B_{n+i} \in \tau-ctx^+(B_n)$ which caused the creation of a new cluster so as to decide about its noisy condition. If $hyp\_noise_{n+i}^l = true, \forall l \in L_{n+i}^{noise}$ then we set $beat\_noise_{n+i}^l = true$, the created cluster is deleted and its beats assigned to its closest cluster. Afterwards, $lead\_noise_{n+i}^l$ is updated for the whole $\tau-ctx^+(B_n)$ using the $beat\_noise_{n+i}^l$ data to set the begining and ending of noise intervals for each lead. The resulting $lead\_noise_{n+\tau-1}^l$ is the new noise state used in the analysis of the subsequent beats.

VI. Rhythm analysis

The absence of an analysis of the P wave leads to the inability to discriminate premature, normal or ectopic beats which share a common QRS morphology. All previous clustering proposals include rhythm information within their beat characterization and claim its separation capabilities for this kind of arrythmias. In order to make our results comparable, we include a rhythm processing stage that allows us to separate those beats into different groups.

A. RR-Interval characterization

The beat to beat interval (RR) between normal beats, commonly known as NN interval, is the result of a non-stationary stochastic process regulated by the sympathetic and
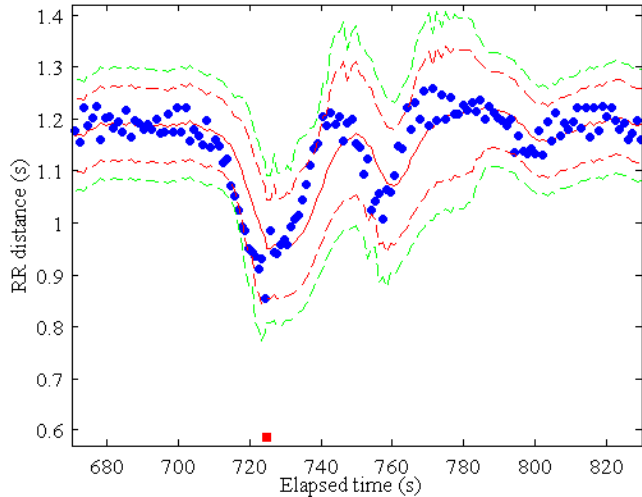
Fig. 6. An excerpt of the temporal evolution of RR distance for record 117 of MIT-BIH database is shown. Solid points and squares represent normal and premature beats respectively. The solid line represents $\widehat{NN}_n$ and the dotted lines enclose the $\pm 2\hat{\sigma}_n$ and $\pm 3\hat{\sigma}_n$ intervals.

parasympathetic nervous systems. This implies that the RR value for beat B_n, denoted by RR_n, should be put in context using the rhythm of the surrounding beats to analyze its normality. To this end, we model the NN series as a stochastic process with marginal distribution $\mathcal{N}(\overline{NN}_n, \sigma_n^2)$ at the nth beat. The mean is estimated as $\widehat{NN}_n = \theta RR_{n-1} + (1-\theta)\widehat{NN}_{n-1}$, with $\theta = 0.2$ for RR_{n-1} values labeled as normal (as explained in next section). Otherwise, $\widehat{NN}_n = \widehat{NN}_{n-1}$. The standard deviation is estimated as $\hat{\sigma}_n^2 = \sum_i NN_i - \widehat{NN}_i$ using the last τ beats with normal rhythm. Fig. 6 shows the evolution of the RR in an excerpt with a premature beat from the record 117 of the MIT-BIH database to illustrate this point.

The first complete context, $\tau - ctx^-(B_{\tau+1})$, is used to initialize $\widehat{NN}_n$ and $\hat{\sigma}_n$. Let TC denote the set of the first τ values of RR and let K_i denote any subset of RR values from consecutive beats. A value $RR_j \in TC$ is said to be normal if $\exists K_i$ such that $RR_j \in K_i \wedge \sigma_{K_i} < 0.1$, where σ_{K_i} is the normalized standard deviation of the K_i set. Let $TC_N = \{RR_j \mid RR_j \text{ is normal}\}$. If $TC_N = \varnothing$, then $TC_N = \{RR_j \mid RR_j \in [\overline{TC} - 2\sigma_{TC}, \overline{TC} + 2\sigma_{TC}]\}$, where $\overline{TC}$ and σ_{TC} are the mean and standard deviation of TC. If $TC_N = \varnothing$ then the context in TC is repeatedly moved forward one beat until $TC_N \neq \varnothing$ for a context $\tau - ctx^-(B_{\tau+1+i})$. Finally we set $\widehat{NN}_n = \overline{TC}_N$ and $\hat{\sigma}_n = \sigma_{TC_N}$, $\forall n \in [1, \tau + 1 + i]$.

B. Rhythm labeling

The aim of rhythm processing is the discrimination of those abnormal RR values associated with arrhythmic beats from the normal ones. To this end, the rhythm of the B_n beat is characterized through a vector $\boldsymbol{rr}_n = (RR_n, RR_n^-, RR_n^+, \widehat{NN}_n, \hat{\sigma}_n)$ where RR_n^- and RR_n^+ are the RR values for the previous and next beat, respectively. Then, the model described in the previous section is used to establish a range of validity for the RR_n value which allows us to detect any alteration in the normal rhythm.

ΔRR_n	Rhythm label for the previous beat						
	GP	**P**	**N_**	**N**	**N_+**	**C**	**D**
$\in (3\hat{\sigma}_n, \infty]$	D^6 C	$D^{2,6}$ C	D	$D^{1\|10}$ N_+	$D^{2\|11}$ N_+	D	D
$\in (2\hat{\sigma}_n, 3\hat{\sigma}_n]$	C	C	N_+	N_+	N_+	$D^{5,6}$ N_+	$D^{6\|9}$ N_+
$\in [-2\hat{\sigma}_n, 2\hat{\sigma}_n]$	N	N	N	N	N	N	N
$\in [-3\hat{\sigma}_n, -2\hat{\sigma}_n)$	GP^8 N_-	N_-^1 GP	N_-	$P^{3,4,6}$ $GP^{3,8}$ N_-	N_-	P^6 GP^8 N_-	P^4 N_-
$\in [-\infty, -3\hat{\sigma}_n)$	GP	GP	$P^{3,4,7}$ $GP^{3,12}$ N_-	$P^{3,4,7}$ GP^3 N_-	P^4 GP	$P^{4,7}$ GP	P^4 GP

1. $RR_n > RR_n^- + 4\hat{\sigma}_n$
2. $RR_n > RR_n^- + 3\hat{\sigma}_n$
3. $RR_n < RR_n^- - 3\hat{\sigma}_n$
4. $RR_n < RR_n^+ - 3\hat{\sigma}_n$
5. $RR_n^- > \widehat{NN}_n + 3\hat{\sigma}_n$
6. $RR_n^+ > \widehat{NN}_n + 3\hat{\sigma}_n$
7. $RR_n^+ > \widehat{NN}_n - 3\hat{\sigma}_n$
8. $RR_n^+ < \widehat{NN}_n - 3\hat{\sigma}_n$
9. $RR_n^+ > \widehat{NN}_n - 2\hat{\sigma}_n$
10. $RR_n^+ > RR_n^- + 4\hat{\sigma}_n$
11. $RR_n^+ > RR_n^- + 3\hat{\sigma}_n$
12. $RR_n^+ < RR_n^- - 3\hat{\sigma}_n$

We use seven rhythm labels to discern between four beat rhythm types: normal, with (C) or without (N, N^-, N^+) compensatory pause, premature (P), group of prematures (GP) and delayed (D). The explicit domain knowledge contained in Table I models, for each rhythm type, the relation of an RR value with the normal rhythm from its temporal context. It also reflects the dependence of the rhythm type for an RR value on the rhythm type of its adjacent beats. This model allows us to assign a rhythm label to the beat B_n based on the $\boldsymbol{rr}_n$ values and the rhythm label of the previous beat.

VII. RESULTS

We have applied our clustering method to all the records of the MIT-BIH database. The parameters and threshold values of this method have been neither trained nor adjusted to fit this database. These values have been justified by physiological reasons, or by the expertise or intuition of experienced cardiologists. The method shows low sensitivity to small changes in parameter values; the results either improve or worsen slightly. Although better results could be obtained by a fine tunning of these parameters for each specific database, this is not the aim of this work but proving the validity of the present approach for continuous ECG monitoring.

For each record, a set of clusters is obtained reflecting the QRS morphologies present in them. Afterwards, the rhythm labels are used to split each cluster into groups of beats with the same rhythm type. In order to compare our results with the proposal in [5] under equivalent conditions, we adopted a fixed number of clusters (25) as the maximum number of groups. Thus, if that limit is exceeded for a record, a merging process is applied to obtain a reduced set of groups with the most prevalent rhythm and morphologies. If necessary, we

TABLE II

NUMBER OF CLUSTERS PER RECORD IN MIT-BIH DB. N COLUMN INDICATES THE NUMBER OF CLUSTERS CREATED USING QRS MORPHOLOGY; N_{RR}, THE NUMBER OF GROUPS AFTER USING RHYTHM LABELS; AND N_{RR}^c, THE NUMBER OF GROUPS AFTER THE MERGING PROCESS.

Rec.	N	N_{RR}	N_{RR}^c	Rec.	N	N_{RR}	N_{RR}^c
100	4	7	7	201	15	32	10
101	4	6	6	202	9	18	18
102	10	13	13	203	33	87	25
103	10	12	12	205	14	20	20
104	16	25	25	207	61	96	25
105	10	16	16	208	28	63	22
106	27	49	18	209	10	26	9
107	11	21	21	210	27	65	13
108	22	35	7	212	5	8	8
109	13	18	18	213	17	29	11
111	8	10	10	214	21	36	11
112	4	7	7	215	16	32	8
113	5	9	9	217	28	50	19
114	8	13	13	219	14	24	24
115	11	11	11	220	2	5	5
116	10	15	15	221	14	24	24
117	4	5	5	222	8	19	19
118	3	8	8	223	23	43	17
119	6	13	13	228	14	25	25
121	5	7	7	230	3	5	5
122	1	1	1	231	5	6	6
123	3	5	5	232	4	9	9
124	14	18	18	233	24	48	15
200	20	46	16	234	5	7	7

keep merging the groups with the lowest number of assigned beats until the maximum is reached. Table II shows the results before and after the merging process.

Each group is labeled with the majority class label of the beats assigned to this group from the database. An assigned beat is considered as correctly grouped if the class label in the database match the label of the group. A confusion matrix is obtained for each record comparing both labels for every beat. These matrices are summed to obtain the global confusion matrix for the whole validation set shown in Table III. Table IV shows the results using the AAMI class labels obtained from MIT-BIH labels as described in [15]. Table V shows the results on AHA database.

A. Real-time considerations

The proposed method processes an ECG recording with a bounded time delay comprising the intrinsic latency and computation time for the different stages. Only the baseline filtering and rhythm labeling stages present an intrinsic latency due to non-causality: 400ms and one beat, respectively. Since they both are executed concurrently, their contribution to the delay is given by the maximum of both. The computation time can be evaluated through the time complexity which is constant in all stages but two: QRS alignment and template matching. In both cases, the time complexity is constant for the best case, corresponding to beats assigned to a cluster in the context; this happens in 95.35% of the total number of beats in MIT-BIH database. The time complexity is linear ($\mathcal{O}(|C_n|)$) for the worst case, which corresponds to beats assigned to a new or an out-of-context cluster, representing the remaining 4.65% of the total number of beats. Given the

high degree of parallelism in both stages and the computational cost of processing a single cluster, it can be guaranteed that a beat is clustered before the next one arrives even with a set of hundreds of clusters. In order to support this claim, the MIT-BIH Arrhythmia database was processed with a non-optimized, non-parallelized MATLAB implementation of the method using a single core of an Intel Q9550 CPU. The following computation times for a single beat were obtained: QRS characterization (maximum, mean): 4ms, 3ms; cluster set selection: 323ms, 16ms; cluster set updating: 110ms, 5ms; noise analysis: 9ms, <1ms; and rhythm-based labelling: 10ms, 5ms. Globally, the whole method summed a maximum of 358ms and mean of 32ms. Additionally, reducing groups for validation required a maximum of 190ms.

B. Clustering performance measures

Purity is usually used to measure the goodness of a clustering method. Nevertheless, in a multiclass problem like this one, after characterizing the clusters, the values of sensitivity (Se), positive predictivity (+P), and false positive rate (FPR) should also be provided for each class, in order to obtain a multidimensional measure of the quality of the results from the perspective of each class involved. A global purity of 98.56% is obtained for MIT-BIH Arrhythmia database (98.84% with AAMI class labels) and a 99.56% for the AHA ECG database. The other values are shown in the last row of Tables III–V.

VIII. DISCUSSION AND CONCLUSION

We propose a new clustering method to dynamically separate QRS morphologies as they appear in a multichannel ECG signal, representing them with a dynamic number of clusters. This objective has not been previously addressed in the literature and only some partial solutions can be found, all of them restricted to offline processing [5]–[7] and some of them even limited to single channel signals and to a fixed subset of beat classes [6], [7].

The performance of the QRS clustering technique without using rhythm data shows a global purity of 97.15% and 99.43% for MIT-BIH and AHA databases, respectively. This confirms the validity of our approach, since no method, as far as we know, neither offline nor online, achieve this performance without using *RR* derived information. As expected, the main source of error is the group of supraventricular classes A, N, J, j, and e, that can only be separated using P wave and rhythm information. This is the reason for the difference between both results since AHA database does not contain this type of beats.

The validation results after using rhythm labels show a high sensitivity and positive predictivity for almost all classes while the purity increases. We observe in accordance with [5], that the largest number of errors in Table III are caused by beats with similar morphology. Fusion (F) of ventricular (V) and normal (N) beats are included with N or V clusters and viceversa. The same occurs with N, paced (/) and fusion of N and / beats (f). Finally, beats with supraventricular or nodal activation points (A, N, J, j, and e) with similar QRS are wrongly clustered when the rhythm information does not

TABLE III
CONFUSION MATRIX RESULTING FROM CLUSTERING MIT-BIH ARRHYTHMIA DATABASE. THE FIRST ROW CORRESPONDS TO THE ANNOTATION LABELS OF THE DATABASE, AND THE FIRST COLUMN, TO THE DOMINANT ANNOTATION LABEL IN THE CLUSTERS. Se, $+P$ AND FPR DENOTE THE SENSITIVITY, POSITIVE PREDICTIVITY AND FALSE POSITIVE RATE FOR EACH BEAT CLASS RESPECTIVELY.

	N	L	R	a	V	F	J	A	S	E	j	/	e	f	Q	!
N	74618	0	3	15	167	108	11	196	2	1	60	1	15	38	9	0
L	0	8059	0	0	1	2	0	0	0	0	0	0	0	0	2	0
R	32	0	7245	0	1	3	29	8	0	6	6	0	0	0	0	4
a	7	0	0	120	1	0	1	2	0	0	5	0	0	0	0	0
V	131	1	0	11	6848	83	0	20	0	0	0	0	0	1	4	6
F	26	0	0	2	56	606	0	2	0	0	0	0	0	0	0	0
J	1	0	0	0	0	0	42	0	0	0	0	0	0	0	0	0
A	44	4	3	2	25	0	0	2317	0	0	0	0	1	0	0	59
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	2	0	10	0	0	0	0	93	0	0	0	0	0	0
j	145	0	0	0	0	0	0	0	0	0	158	0	0	0	0	0
/	9	0	0	0	1	0	0	0	0	0	0	7010	0	127	0	0
e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f	9	0	0	0	5	0	0	0	0	0	0	13	0	816	15	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
!	0	8	0	0	12	0	0	0	0	6	0	0	0	0	0	403
Se	99.46	99.84	99.89	80.00	96.09	75.56	50.60	91.04	0.00	87.74	69.00	99.80	0.00	83.10	9.09	85.38
+P	99.17	99.94	98.79	88.24	96.38	87.57	97.67	94.38	-	88.57	52.15	98.08	-	95.10	100.00	93.94
FPR	1.79	0.00	0.09	0.01	0.25	0.08	0.00	0.13	0.00	0.01	0.13	0.13	0.00	0.04	0.00	0.02

TABLE IV
CONFUSION MATRIX FOR MIT-BIH DB USING AAMI CLASS LABELS.

	N	S	V	F	Q
N	89957	342	176	113	50
S	204	2648	26	0	0
V	134	31	6951	83	5
F	26	4	56	606	0
Q	18	0	6	0	7984
Se	99.58	87.54	96.34	75.56	99.32
+P	99.25	92.01	96.49	87.57	99.70
FPR	3.57	0.22	0.25	0.08	0.02

TABLE V
CONFUSION MATRIX FOR AHA ECG DATABASE. CLASSES FROM THE MIT-BIH LABEL SET WITHOUT ASSIGNED BEATS ARE OMITTED.

	N	V	F	E	/	Q
N	318739	472	138	7	16	150
V	229	32090	217	0	9	66
F	33	115	885	0	16	1
E	0	0	0	5	0	0
/	5	23	26	0	3128	1
Q	11	32	0	0	0	354
Se	99.91	98.04	69.91	41.67	98.71	61.89
+P	99.75	98.40	84.29	100.00	98.27	89.17
FPR	2.07	0.16	0.04	0.00	0.02	0.01

result determinant for their discrimination. This kind of errors represent 66% of the total.

Besides the clustering performance, the analysis of the number of clusters (N) generated for each record shows that it remains reduced: 34 records (71%) have 15 or less clusters; 12 records (25%) have between 16 and 30 clusters, both included and only two records (4%) has more than 30 clusters. The high number of clusters for record 207 is caused by the presence of an episode of ventricular flutter with QRS complexes replaced by irregular waves. The clustering results would be improved if an specific detection method were used for these kind of arrhythmias with absent QRS complex.

The N_{RR} column of Table II shows a general increase in the number of groups in all records with a mean of 24 groups per record. This increment reflects the presence of different rhythm labels in the beats assigned to the clusters although they do not always belong to different beat classes. Records with irregular rhythm like those with auricular fibrillation, or sudden rhythm change will render the RR information useless to discriminate premature or delayed beats. In such cases, the RR does not provide information about the beat activation point and the discrimination cannot be performed without an analysis of the P wave.

Since no other method for online clustering has been reported, we compare our clustering performance with existing offline proposals. Only the work of Lagerholm et al. [5] provides comparable results since others [6], [7] are designed to deal with a concrete subset of beat types and perform the evaluation over a single channel (usually the one with lower noise). Compared to [5] our method provides a slightly better purity (98.56% vs 98.49%). The sensitivity is improved in our work for 11 out of the 16 beat classes and slightly worsened for 3. Special mention should be made for classes a, A, R, j and E where the improvement is remarkable. Let us remember that [5] rely on a SOM with 25 clusters to represent the different beat classes. This approach has two main drawbacks, first the

clusters get saturated with dominant morphologies present in the learning stage, while rare morphologies are ignored as well as new morphologies that appear afterwards. Second, the generated clusters are redundant in those records with a low number of morphologies. In contrast, our method dynamically adapts the number of clusters to the number of morphologies detected.

The results of our proposal confirms the relevance of the temporal context for beat clustering. It allows us to switch from an offline to an online analysis achieving the same or even better results, and to address the temporal evolution of a beat morphology that otherwise would be projected into multiple clusters. The results of the experiments on ECG standard databases also show the adequacy of the present method for real-time ECG monitoring.

Our proposal provides the cardiologists with the information about the morphological diversity within a desired time frame and its temporal evolution. This information allows them to promptly detect the different conduction patterns and evaluate its relevance. It also can be useful for arrhythmia detection and classification which can be later addressed either automatically by classification algorithms or manually by the cardiologists.

In conclusion, we have presented an adaptive, multichannel context-based method for clustering beat morphologies in real-time that has been validated over the whole MIT-BIH Arrhythmia and AHA ECG databases with performance results that outperform its offline counterparts in the field.

REFERENCES

[1] R. Sutton, "Remote monitoring as a key innovation in the management of cardiac patients including those with implantable electronic devices," *Europace*, vol. 15, no. suppl 1, pp. i3–i5, 2013.

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

[3] M. Llamedo and J. P. Martínez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 616–625, 2011.

[4] G. de Lannoy, D. François, J. Delbeke, and M. Verleysen, "Weighted conditional random fields for supervised interpatient heartbeat classification," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 241–247, 2012.

[5] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ECG complexes using Hermite functions and self-organizing maps," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 838–848, 2000.

[6] D. Cuesta-Frau, J. C. Pérez-Cortes, and G. A. García, "Clustering of electrocardiograph signals in computer-aided holter analysis," *Comput. Meth. Prog. Bio.*, vol. 72, no. 3, pp. 179–196, 2003.

[7] D. Cuesta-Frau, M. O. Biagetti, R. A. Quinteiro, P. Micó-Tormos, and M. Aboy, "Unsupervised classification of ventricular extrasystoles using bounded clustering algorithms and morphology matching," *Med. Biol. Eng. Comput.*, vol. 45, no. 3, pp. 229–239, 2007.

[8] J. L. Rodríguez-Sotelo, D. Cuesta-Frau, and G. Castellanos-Domínguez, "Unsupervised classification of atrial heartbeats using a prematurity index and wave morphology features," *Med. Biol. Eng. Comput.*, vol. 47, no. 7, pp. 731–741, 2009.

[9] V. Krasteva and I. Jekova, "QRS template matching for recognition of ventricular ectopic beats," *Ann. Biomed. Eng.*, Sep 2007.

[10] 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, 2006.

[11] 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, 2001.

[12] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 582–589, 2004.

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

[14] Z. Dokur and T. Olmez, "Ecg beat classification by a novel hybrid neural network," *Comput. Meth. Prog. Bio.*, vol. 66, no. 2-3, pp. 167–181, 2001.

[15] *Testing and Reporting Performance Results of Cardiac Rhythm and ST-segment Measurement Algorithms*.   ANSI/AAMI Standard EC57:1998/(R)2008.

[16] A. L. Goldberger, L. A. N. 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, 2000.

[17] G. Moody and R. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001.

[18] W.-Y. Wu, "An adaptive method for detecting dominant points," *Pattern Recogn.*, vol. 36, no. 10, pp. 2231–2237, 2003.

[19] The CSE WORKING PARTY, "Recommendations for measurement standards in quantitative electrocardiography," *Eur. Heart. J.*, vol. 6, no. 10, pp. 815–825, 1985.

[20] *Cardiac Monitors, Heart Rate Meters, and Alarms*.   ANSI/AAMI Standard EC13:2002/(R)2007.

[21] P. Kligfield, L. S. Gettes, J. J. Bailey, R. Childers, B. J. Deal, E. W. Hancock, G. van Herpen, J. A. Kors, P. Macfarlane, D. M. Mirvis, O. Pahlm, P. Rautaharju, and G. S. Wagner, "Recommendations for the standardization and interpretation of the electrocardiogram: Part i." *Circulation*, vol. 115, no. 10, pp. 1306–1324, 2007.

[22] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, no. 1, pp. 43–49, 1978.

[23] E. J. Keogh and M. J. Pazzani, "Derivative dynamic time warping," in *1st SIAM Int. Conf. on Data Mining, Chicago, IL, USA*, 2001.

Daniel Castro received the B.Sc and M.Sc. degrees in Physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1998 and 1999, respectively. He is currently a researcher and Ph.D. candidate at the Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), University of Santiago de Compostela. His research interests include signal processing and its application to biomedical signals.

Paulo Félix received the B.Sc and Ph.D. degrees in Physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1993 and 1999, respectively. He is currently a researcher at the Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), University of Santiago de Compostela. His research interests include temporal reasoning, machine learning and signal processing.

Jesús Presedo received the B.Sc and Ph.D. degrees in Physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1989 and 1994, respectively. He is currently a researcher at the Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), University of Santiago de Compostela. His research interests are biomedical digital signal processing, heart rate variability, nonlinear dynamics, soft computing and the development of ubiquitous healthcare systems.