Computing and data processing

Ventricular beat classifier using fractal number clustering

H. Bakardjian
Institute of Biomedical Engineering, Medical Academy, 1 G. Sofiski, Sofia 1431, Bulgaria

Abstract—A two-stage ventricular beat 'associative' classification procedure is described. The first stage separates typical beats from extrasystoles on the basis of area and polarity rules. At the second stage, the extrasystoles are classified in self-organised cluster formations of adjacent shape parameter values. This approach avoids the use of threshold values for discrimination between ectopic beats of different shapes, which could be critical in borderline cases. A pattern shape conventionally called a 'fractal number', in combination with a polarity attribute, was found to be a good criterion for waveform evaluation. An additional advantage of this pattern classification method is its good computational efficiency, which affords the opportunity to implement it in real-time systems.

Keywords—Classification, ECG, Extrasystole, Fractal number, QRS


1 Introduction

A fully automatic ventricular beat classifier could support an arrhythmia discriminator in Holter-type or bedside patient monitoring.

Most of the efficient classifiers evaluate pattern shapes through template matching with cross-correlation techniques (Lanza et al., 1990; Forster and Handwerker, 1990; Salerno et al., 1987; Lin and Chang, 1989; Lin et al., 1988) or multiparameter statistical classification (Paparella et al., 1987; Gelsema et al., 1988; Merri et al., 1989). Other classification approaches are the distance quantification in some feature space (Wolberg and Mangansarian 1990; Sornmo et al., 1981) and the application of expert knowledge and fuzzy labelling of events (Barro et al., 1990).

These algorithms are currently accepted and put into practice, but they label the examined events according to some relatively synthetic class borders with statistical or threshold criteria, thus a priori not being able to react flexibly enough in all cases. A satisfactory classification could also be attained by these methods, but at the cost of computational complexity and processing-time expansion. The attempts to reduce the error factor formally lead inevitably to classification limitations [e.g. single feature statistics (Kohn, 1989)].

The purpose of this paper is to present another classification approach which tries to follow the natural phenomena more closely. After beat identification and characteristic features extraction, the algorithm implemented performs the most important section: investigation of the grouping of events (beat patterns) into naturally self-organised clusters.

2 Method

2.1 Adaptive beat identification

Accurate ventricular beat identification is a basic preliminary task, prior to any further processing. It is valid for almost any kind of automatic ECG analysis. The following beat classification procedure sets additional requirements for beat identification.

2.1.1 Signal preprocessing. To amplify the QRS characteristic features common in all leads and to suppress slower waves, as the first part of a two-stage preprocessing, a fast procedure to compute the pseudospatial velocity is applied, which combines the event information from $L$ leads:

$$Y(i) = \sum_{j=1}^{L} \left[ X_j(i + 1) - X_j(i - 1) \right]$$

where $X_j(i)$ is the amplitude of data sample $i$ in lead $j$, and $Y(i)$ is the current pseudospatial velocity value.

To improve the beat detection prerequisites, the Menard derivative (Friesen et al., 1990) is employed on the computed data $Y$:

$$Z(i) = -2Y(i - 2) - Y(i - 1) + Y(i + 1) + 2Y(i + 2)$$

where $Z$ is the preprocessed data used for beat identification decisions only (Fig. 1).

2.1.2 Iterative decision rule. The decision rule should detect all beat events occurring, independently of their shape or peak amplitude, to respond to the further classification requirements. An adaptive iterative procedure is implemented on the $Z$ data to extend the advantages of the preprocessing. Each identified beat modifies a weighted mean peak amplitude estimator and a weighted mean $R - R$ interval estimator, both defining the adaptive search features of this detection technique.
The peak amplitude estimator (PAE) accounts for the past several ventricular beats and the very last detected peak amplitude value (PAV) on the detection threshold. The new weighted mean PAE is updated by a forgetting factor of 1/33, and the actual detection threshold (DT) value pursued on the preprocessed data is set initially at 33 per cent of the adapted PAE. This unusually low detection threshold is possible due to the specific preprocessing and is indispensable because a series of ectopic beats could set the current PAE too high. For each beat $b$, the following relationships are valid:

$$PAE(b) = \frac{32PAE(b-1) + PAV(b)}{33}$$

and

$$DT(b) = \frac{33}{100} PAE(b)$$

so that

$$DT(b) = \frac{32PAE(b-1) + PAV(b)}{100}$$

A forward search is performed using an adaptive search length, defined in a similar way by a variable, called the weighted mean $R-R$ interval estimator (RIE). Each new interbeat interval $R-R(b)$ updates this estimator by a forgetting factor of 1/8:

$$RIE(b) = \frac{7RIE(b-1) + R-R(b)}{8}$$

and the actual search length for the next beat is set at 200 per cent of the current RIE. As usual, a refractory blanking period is provided as well.

If, in a first pass, a normal or low-amplitude beat cannot be detected within the current search length, an iterative procedure performs new back searches, each with a 25 per cent decrease in the current detection threshold (PAE remains unchanged), until the next beat is identified. Finally, the successful detection of beat $b$ after $r$ recursions has established the following local threshold value:

$$LDT(b) = 2^{-2r}DT(b)$$

The Z peak position for each beat is found by searching for the maximum of Z data after the detection point. These locations are beat fiducial marks, labelling the identified ventricular beats.

Finally, the two adaptive weighted mean estimators can already be updated for the next beat search, taking into account the last Z peak amplitude and $R-R$ interval.

2.2 Ventricular beat onset and offset

Ventricular beat onsets and offsets are searched in the ECG data first derivative (simple slope)

$$dx_j(i) = x_j(i) - x_j(i - 1)$$

for each lead $j$ individually. A new beat location marker is placed on the first X data point, with actual slope $dx_j$ exceeding a preset but variable slope threshold after the previously found fiducial mark. We proceed further using an iterative search to the left and right of the beat marker. The first scan in each direction proceeds within a maximum search length, until a region with slope lower than an initial value of the slope threshold is located, and its inner end is found (Fig. 1b). The maximum search length in both directions is set equal to the current $R-R$ interval estimator. If, in any direction, this procedure is unsuccessful, the slope threshold is incremented, and an iterative search is repeatedly undertaken, until the onset and offset points of the ventricular beat for the investigated lead $j$ are found and labelled.

2.3 Ventricular beat characterisation parameters

Attempting to describe ventricular beat patterns, we tested the following input parameters: ventricular beat signed and absolute area, duration, $R$-amplitude, 'sign', fractal number of raw data and fractal number of the first derivative. Some of these parameters were further selected according to their efficiency in the corresponding classification stages.

2.3.1 Beat 'sign'. This feature is rather conditional, because different parameter values reflect different types of polarity variation only for the purpose of comparison. The parameter beat 'sign' is rather artificial, but it should be