DIFFERENCE-BASED PARAMETER SET FOR LOCAL HEARTBEAT CLASSIFICATION: RANKING OF THE PARAMETERS

Irena Ilieva Jekova, Ivaylo Ivanov Christov, Lyudmila Pavlova Todorova

Centre of Biomedical Engineering "Prof. Ivan Daskalov" - Bulgarian Academy of Sciences
Acad.G.Bonchev str. Bl.105, 1113, Sofia, Bulgaria, e-mail: irena@clbme.bas.bg

The initial step in the diagnosis of cardiac dysfunctions is the detection and classification of different types of heartbeats in the electrocardiogram (ECG). The aim of the present work was to assess the ability of a difference-based parameter set, containing information for an amount of morphological heartbeat features and for the RR intervals, to provide an adequate local heartbeat classification without requiring manual annotation of patient specific heartbeats. The study was performed on all ECG recordings from the internationally recognised MIT-BIH arrhythmia database. The results showed that the proposed difference-based parameter set could be a suitable tool for classification of ventricular ectopic beats.

Keywords: heartbeat classification, morphological parameters, RR intervals

1. INTRODUCTION

The initial step in the diagnosis of cardiac dysfunctions is the detection and classification of different types of heartbeats in the electrocardiogram (ECG). Some arrhythmias appear infrequently, and in order to capture them the clinicians use Holter devices. The use of specific algorithms for automatic analysis of ECG recordings may facilitate the analysis of the very long Holter ECG recordings. A particular aspect of the developed algorithms for automatic beat classification is the adopted learning strategy. It is studied, paying attention to the organization of the classifiers' training set, and considering two main approaches: *local* learning set and *global* learning set [1,2,3,4]. In the first case the learning set is customized to the tested patient, while in the latter it is built from a large ECG database. Traditionally the local learning set requires a cardiologist to annotate a set of heartbeats of the patient which are used afterwards as a basis for the classification. It is obvious that this is associated with time-consuming manual editing of the patient's ECG recording. On the other hand, the capacity of the global learning set to classify new records without additional training is balanced by a lower accuracy, since the morphology of the QRS complexes (of one and the same type) differ not only from patient to patient, but also from lead to lead of a same individual. Besides the learning strategy, an important point is the parameter set, which is selected to characterize the heartbeats. Some of the most popular ECG descriptors are based on assessment of the QRS complex morphology [1,3,4]. Other authors use time-frequency based parameters [4] or features calculated by QRS template matching procedures, based on different transforms, e.g. Karhunen-Loève transform [5], Hermite functions [6,7].

The aim of the present work was to assess the ability of a difference-based parameter set, containing information for an amount of morphological heartbeat

features and for RR intervals, to provide an adequate local heartbeat classification without requiring manual annotation of patient specific heartbeats.

2. METHOD

2.1 ECG Signals

The study involved all 48 ECG recordings from the MIT-BIH arrhythmia database. Each recording has a duration of 30 min and includes two leads - the modified limb lead II and one of the modified leads V1, V2, V4 or V5 [8]. The sampling frequency is 360 Hz and the resolution is 200 samples per mV. The heartbeats were recognized by the fiducial points in the database and the original database annotations were accepted. The study was focused on the analysis of the five largest heartbeat classes in the MIT-BIH arrhythmia database (almost 95 % of all beats), which feature with particular behavior of the QRS morphology: (i) Normal beats (N) - about 75000 cases; (ii) Premature Ventricular Contractions (V) - about 7130; (iii) Left Bundle Branch Blocks (L) - about 8070; (iv) Right Bundle Branch Blocks (R) - about 7260 and (v) Paced beats (P) - about 7020. In our study, we used the full-length MIT-BIH files, without selection based on the quality of the signal.

Before analysis we applied the following preprocessing procedures:

- a notch filter for elimination of the power-line interference, implemented by moving averaging of samples in one period of the interference;
- a low-pass filter for suppression of the tremor noise, realized by moving averaging of samples in 30 ms time-interval, thus having a first zero at about 35 Hz;
- a high-pass recursive filter for drift suppression with cut-off frequency of 2 Hz.

2.2 Morphological parameters

A method for calculation of a large collection of morphological descriptors was applied to all QRS complexes annotated as N, V, L, R or P in the MIT-BIH arrhythmia database. An isoelectric, baseline segment is searched for by starting from the QRS fiducial point back on the time axis up to 120 ms. The segment is found if eight successive differences between adjacent samples are less than a preset value and the difference between the end samples of the segment is lower than the same value. The earliest sample of this segment is defined as an isoelectric point.The QRS pattern recognition technique identified the onset and the offset of the QRS complex by simultaneous analysis of the two ECG leads. Then a set of morphological descriptors representing information of the amplitude, area, specific interval durations and measurements of the QRS vector in the vectorcardiographic (VCG) plane were calculated as follows:

- Eleven descriptors were extracted by individual assessment of each lead:

- ***Pp*** - Maximal amplitude of the positive peak;
- ***Pn*** - Maximal amplitude of the negative peak;
- ***ArP*** - Area of the positive samples in the identified pattern;
- ***ArN*** - Area of the negative samples in the identified pattern;

- ***Ar*** - Area of the QRS complex - the sum of the absolute values of the ECG samples in the identified pattern; (***Ar***=***ArP***+***ArN***);
- ***Av*** - Sum of the absolute values of the velocities in the pattern interval;
- ***No*** - Number of samples crossing a threshold of 70% of the highest peak amplitude;
- ***Ima*** - Time-interval from the QRS complex onset to the maximal positive peak;
- ***Imi*** - Time-interval from the QRS complex onset to the maximal negative peak;
- ***S1*** - QRS slope velocity calculated for the time-interval between the QRS complex onset and the first peak;
- ***S2*** - QRS slope velocity calculated for the time-interval between the first peak and the second peak.

• One descriptor representing the time-interval between the onset and the offset of the ventricular contraction was derived by the simultaneous leads analysis – ***Width***.

• Three descriptors were calculated from the single-plane VCG formed by the two leads:

- ***VCGamp*** - Maximal amplitude of the VCG vector;
- ***VCGsin*** - Sine component of the angle of the maximal amplitude vector;
- ***VCGcos*** - Cosine component of the angle of the maximal amplitude vector.

2.3 RR intervals

For each heartbeat the preceding RR interval, was measured - ***RRBefore***. We selected to analyze ***RRBefore***, since it normally remains stable for the N, L, R and P, and is shortened for the V beats.

2.4 Difference-based parameter set

We formed a difference-based parameter set by calculating the mean values of the above described parameters and the differences between these mean values and the respective parameter values for each heartbeat, following equation (1):

$$DPi = 100 * \frac{Pi - mean(Pi)}{mean(Pi)} \quad , [\%] \tag{1}$$

where ***Pi*** is the parameter with index ***i***;
mean(Pi) is the mean value of the ***Pi*** parameter;
DPi is the calculated difference parameter with index ***i***.

3. RESULTS

The descriptors included in the difference-based parameter set were calculated with MATLAB 7.0 and afterwards were processed with STATISTICA 6.0. The statistical distributions of all morphological difference descriptors were studied and 4 of them (***DPp***, ***DPn***, ***DAr***, ***DWidth***) present significant divergences between N, L, R, P beats on the one hand and V beats on the other. Aiming to strengthen these distinctions they were combined in a summary parameter, calculated as follows:

$$DTotal = \frac{(DPp + DPn + DAr + DWidth)}{4} \tag{2}$$

The categorized histograms of ***DPp***, ***DPn***, ***DAr***, ***DWidth***, ***DTotal*** and ***DRRBefore*** for the 5 heartbeat classes are represented in Fig. 1 and Fig. 2.

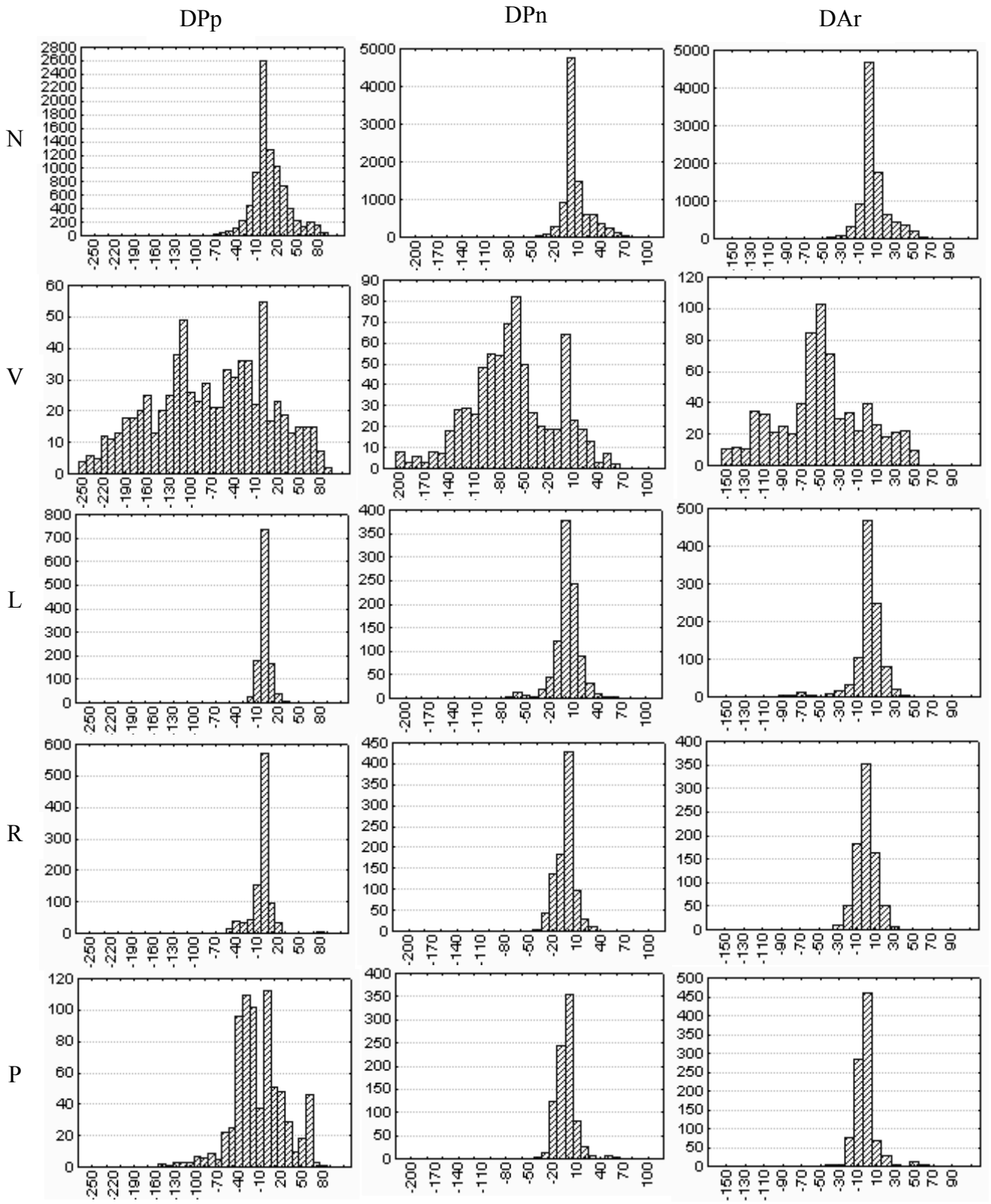

Fig. 1. Categorized histograms of DPp, DPn and DAr obtained for the 5 heartbeat types – N, V, L, R and P. X-axis – parameters' values (measured in percents) Y-axis – number of observations

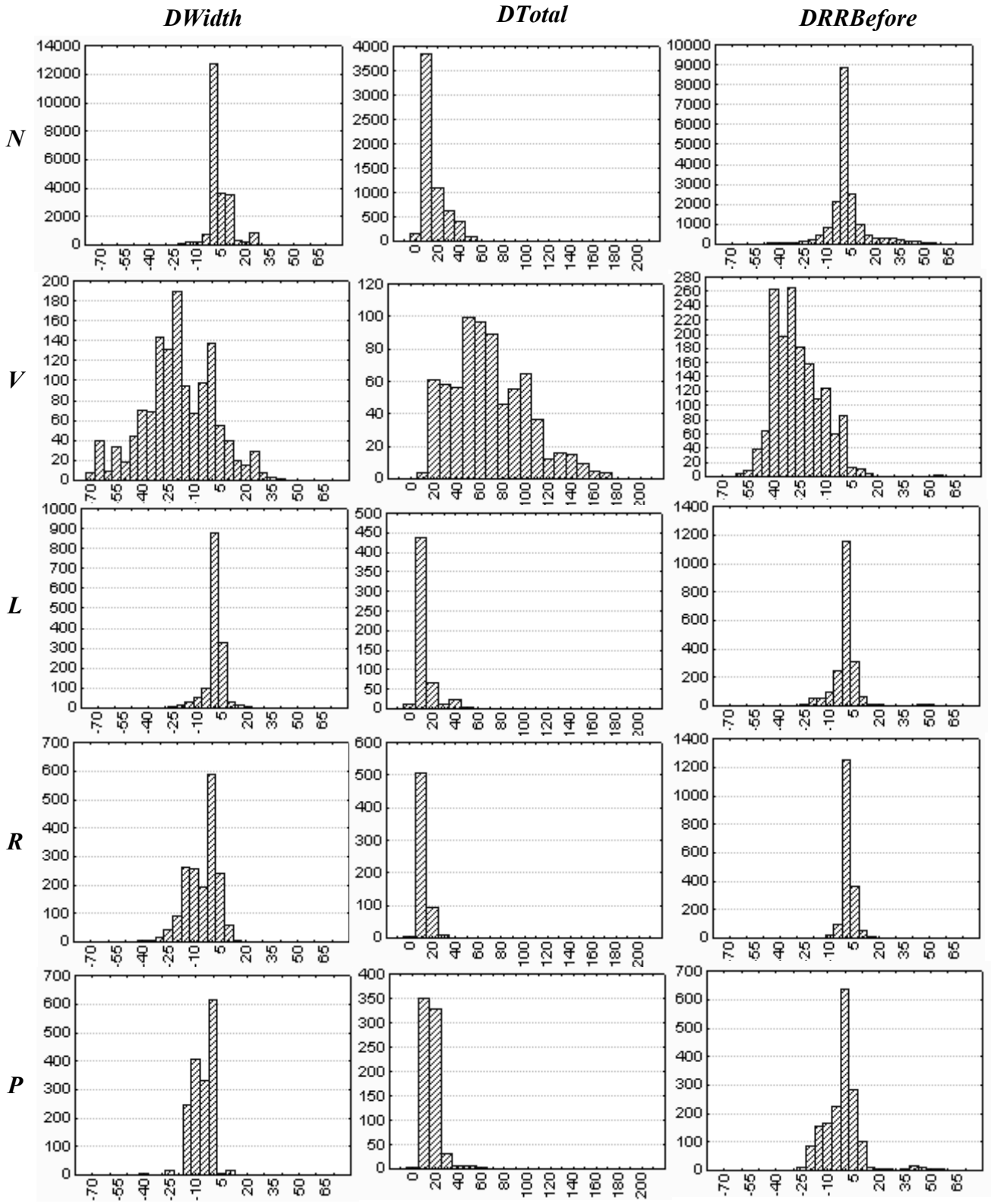

Fig. 2. Categorized histograms of DWidth, DTotal and DRRBefore obtained for the 5 heartbeat types – N, V, L, R and P.
X-axis – parameters' values (measured in percents)
Y-axis – number of observations

4. Discussion and Conclusion

The presented categorized histograms show significant differences in the distributions of ***DPp***, ***DPn***, ***DAr***, ***DWidth***, ***DTotal*** and ***DRRBefore*** for V beats, compared to N, L, R and P beats. Therefore, these parameters offer a suitable tool for classification of ventricular ectopic beats. The slight overlapping between the morphological difference parameters observed for V beats and for all other beats can be explained with the presence of deviations in the waveforms of N, L, R and P beats, as well as with the presence of V beats which resemble the morphology of the predominant patient's QRS complexes (in respect to this parameter set). As it could be expected, the values of ***DRRBefore*** are negative for V beats, due to the shortened RR interval preceding them. The partial coincidence between the distributions of ***DRRBefore*** for V beats and for all other beats could be justified by the existence of interpolated V beats (the so called 'sandwiched' V beats) which lead to shortened ***RRBefore*** for the following N, L, R or P beat.

Observing the categorized histograms, it is obvious that the discrimination between N, L, R and P beats could not be done by means of this difference-based parameter set, since they usually form the patient's sustained rhythm and significant differences are present neither in the morphological descriptors, nor in the RR intervals.

The general advantage of the proposed parameter set is its patient specific orientation, which eliminates the necessity of manual time-consuming annotation of certain amount of heartbeats. It is absolutely suitable for implementation in Holter devices, where the analysis of the long ECG recording is performed offline. For real-time operating ECG monitors a learning phase would be required in order to calculate initial mean values of the estimated parameters.

5. References

[1] Chazal P., M. O'Dwyer, R. Reilly, *Automatic classification of heartbeats using ECG morphology and heartbeat interval features*, IEEE Transactions on Biomedical Engineering, Vol. 51, pp. 1196–1206, 2004.

[2] Hu Y., S. Palreddy, W. Tompkins, *A patient-adaptable ECG beat classifier using a mixture of experts approach*, IEEE Transactions on Biomedical Engineering, Vol. 44, pp. 891-900, 1997.

[3] Christov I., I. Jekova, G. Bortolan, *Premature ventricular contraction classification by the K^{th} nearest neighbours rule*, Physiological Measurement, Vol. 26, pp. 123-130, 2005.

[4] Christov I., G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, K. Egiazarian, *Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification*, Medical Engineering & Physics, Vol.28, pp.876-887, 2006.

[5] Moody G., R. Mark, *QRS morphology representation and noise estimation using the Karhunen-Loève transform*, Computers in Cardiology, Vol. 16, pp. 269-272, 1989.

[6] Laguna P., R. Jane, S. Olmos, N. Thakor, H. Rix, P. Caminal, *Adaptive estimation of QRS complex wave features of ECG signal by the Hermite model*, Medical & Biological Engineering & Computing, Vol. 34, pp. 58-68, 1996.

[7] Lagerholm M., G. Peterson, G. Braccini, L. Edenbrandt, L. Sörnmo, *Clustering ECG complex using Hermite Functions and self-organizing maps*, IEEE Transaction on Biomedical Engineering, Vol. 47, pp. 838-848, 2000.

[8] MIT-BIH Arrhythmia Database http://www.physionet.org/physiobank/database/mitdb/