CLUSTERING OF HEARTBEATS FOR AUTOMATED ECG HOLTER ANALYSIS

Václav Chudáček*, Martin Hanuliak**, Lenka Lhotská*

* Czech Technical University in Prague/Gerstner Laboratory, Technická 2, Prague, Czech Republic
** Medical Technologies Prague, Czech Republic

chudacv@fel.cvut.cz

Abstract: Holter ECG monitoring is used for monitoring patients with possible diseases of heart such as arrhythmias or silent ischemia. Key task that is necessary for real-life application is processing of huge amounts of data – clearly task for computer system. In this article we propose simple thus fast method for automatic heartbeats clustering based on features that would have to be computed anyway for diagnosis. Computing should not take longer than five minutes on P4 2Ghz with reasonable results.

Introduction

Holter [1] ECG monitoring is nowadays common method used for long term monitoring of patients with possible diseases of heart such as arrhythmias or silent ischemia. There are some differences to standard 12-electrode ECG measuring that disallow us usage of standard approaches known from standard 12-lead ECG. First difference is in time span - we are able with holter to measure 24-hours recordings, but 48-hour recording are also used. Such a long recording brings enormous amount of data – circa in magnitude of hundreds of megabytes. Second difference is electrode placement – there are many of measuring systems ranging from 2-electrode to 12-lead [2]. Electrode placement positions are defined but unlikely as in standard 12-electrode system actual placements of electrodes differ. Usually electrodes are placed to suit best to the patient. Such inconsistent approach brings many problems in signal analysis and evaluation. Third problem lies in measurement itself – patient is not lying in doctors office – but he is spending his day as usual – moving, running, sleeping – in short doing things that require large elasticity on the side of algorithm structure.

Although there are many methods published for ECG holter beat clustering [5, 6] drawbacks of these methods, although very often with high percentage of correctly classified beats are at least two.

Firstly they are either not robust enough to cope with noisy signal and/or tested on standard signal patterns. Or secondly computational methods needed for acquiring results are very time consuming. If we are to offer our clustering method for use in everyday medical practice we need to have all the computations including preprocessing and filtering done in up to five minutes.

Approaches to clustering of heart beats in holter system can be divided into two groups. First one is clustering based on direct shape description [4], with some varieties using some mathematical simplification of the curves. Disadvantage of such an approach seems to be high computation demands and fact that we are usually computing something not usable for further use for example for diagnostic purposes. Second approach is using parameterization of the curve usually by important ECG-wave description [8,9,11]. Disadvantage can be to difficult computation for parameters – such as analysis of ECG signal by wavelet analysis.

Possible further computation – getting from features to clusters are very wide [3,10] but again they do differ very much in computation time.

Measurement and Data

The aim of this study was to find out interesting points suitable for clustering and try to evaluate them.

As to simulate real life condition we start computation after 24-hour measurement is done and we are loading the signal into the computer. We used 3-lead configuration (V1 V3 V5) and we measured 15 24-hours recording and had on each recording at least 30min annotated by experts.

Preprocessing

First thing after filtering 50Hz, we detect noisy parts of the signal, paced beats – parts of the signal we will not cluster. Then using adaptive filtering we find possible R peaks, beginnings and ends of QRS-complex and end of T-wave. Directly measured parameters from analysis are in Table 1. Then we create median class from first ten to twenty consecutive beats. We count intervals of QRS and T wave duration, maximal amplitudes of Q, R, S, T waves and also areas under the Q, S and T wave. We are not taking into account P-wave because of its difficult detection on holter recordings.

ISSN: 1727-1983 © 2005 IFMBE

Table 1: Directly measured parameters

QRSon	Onset of QRS complex
Rpeak	Maximum R-deflection
QRSoff	End of QRS complex
Toff	End of QRS complex

Table 2: Computed parameters

ampQ	Maximum Q-deflection
ampS	Maximum S-deflection
STsegment	segment of 80 ms beginning at QRSoff
Ton	Onset of T-wave

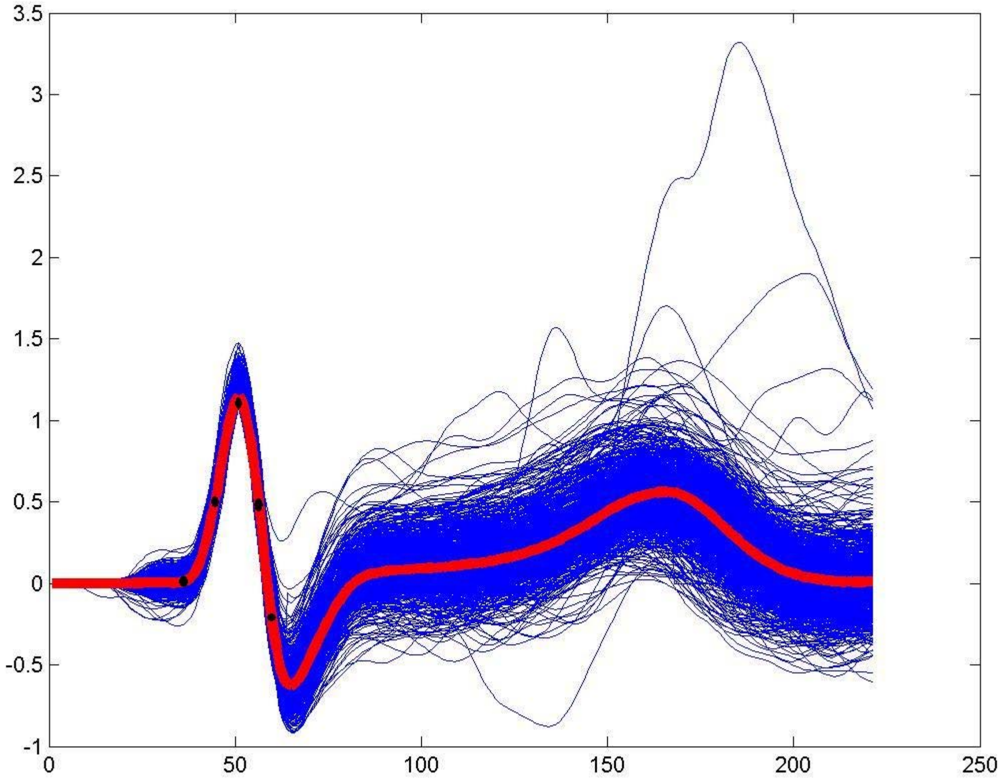

Figure 1: Cluster after QRS-complex clustering with black points for direct shape differentiation

Feature Extraction

For feature computation we used only points computed from basic beat analysis as shown in Table 3. Although these can not possibly cover all spectrums of beat shapes it seems that in usual cases these are sufficient.

Table 3: Points computed during basic beat analysis

widthQRS	Width of QRS complex
ampR	Absolute amplitude of R peak
ampS	Absolute amplitude of S peak
ratRQ	R:Q ratio

ISSN: 1727-1983 © 2005 IFMBE

ratRS	R:S ratio
polT	polarity of T wave
ampT	Absolute amplitude of T wave
ratRT	R:T ratio
areaT	Area under T wave
areaST	Area under ST segment
slopeST	Slope of ST segment

As ST-segment was taken part of the beat beginning at the end of QRS-complex with 80ms duration. T-on parameter was set according to equation (1).

$$Ton = QRSoff + RR/16 \qquad (1)$$

Additionally to these we computed delta-wave indicator and also P wave polarity although we did not use P-wave parameter for final computation.

Having in mind the facts mentioned above we tried to develop fast and robust clustering method that will divide heartbeats into clusters. Further computations can be made on individual clusters with relation to expected diagnosis.

As we stated at the beginning we also implemented direct shape comparison method – method that is the most

precise but has a drawback in extremely large amount of computations needed especially with 500Hz signal as in our case. We tried to ease this computational burden by selecting just five points on QRS-complex and we used this straight shape comparison only in limited cases

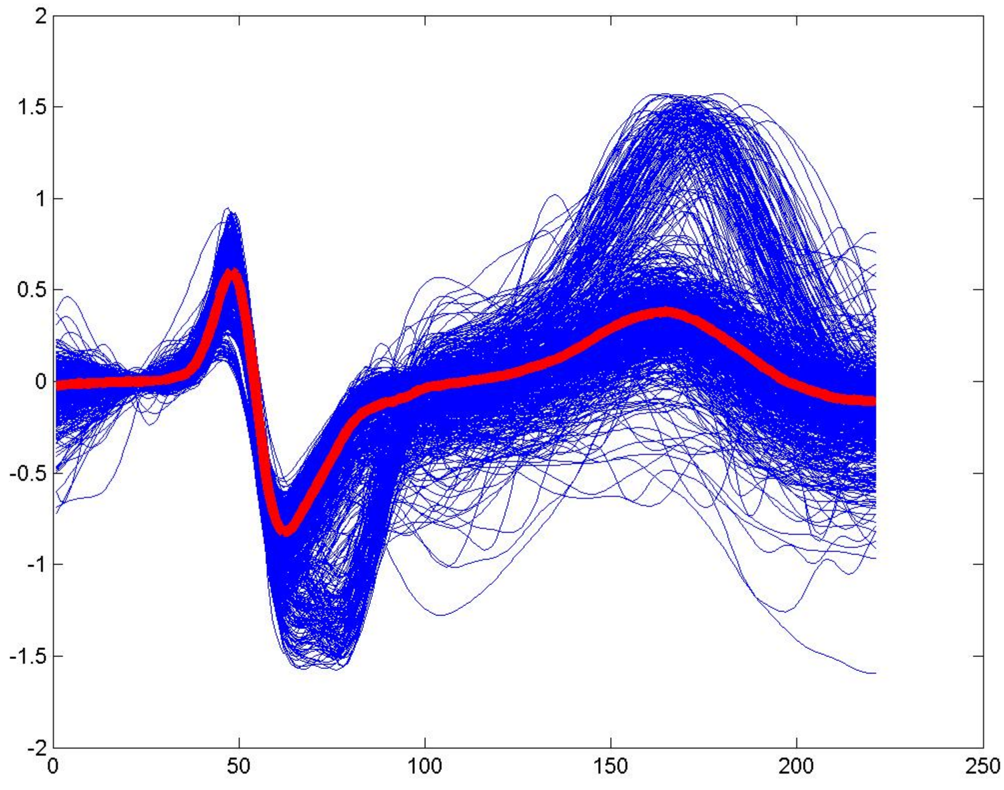

Figure 2: Example of cluster from middle of clustering process

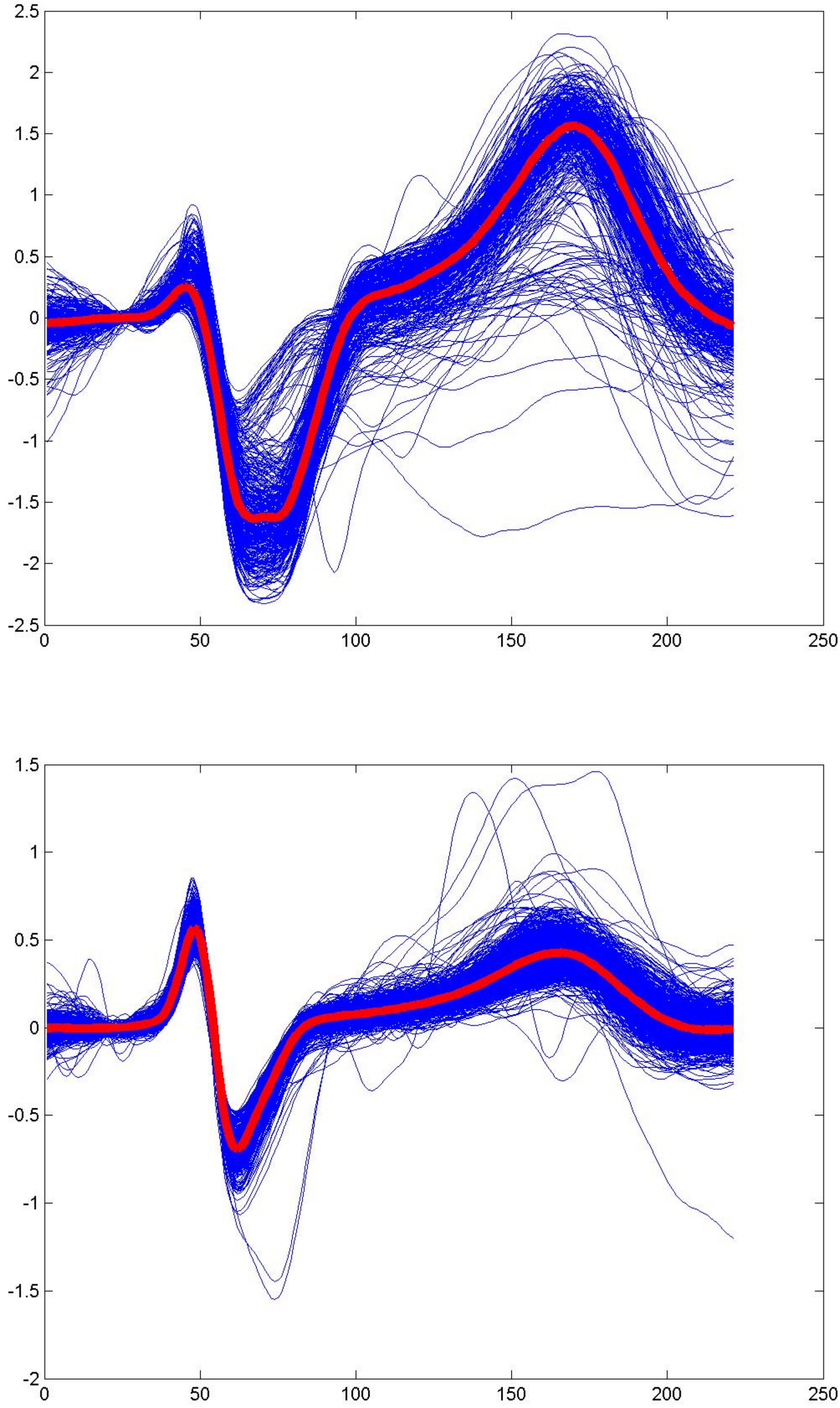

Figure 3: Two sub clusters from cluster on Figure 2

ISSN: 1727-1983 © 2005 IFMBE

Clustering

After having features computed the actual clustering is very easy to perform. We make first median class as described above. Then we go beat after beat – we measure parameters compute features and than give the beat into cluster according to his features. In the next step we take next beat use the decision tree, that is set according to the first ten beats median. We either add investigated beat into the cluster or create new cluster. After the whole signal is processed we make median of each cluster. Median and also the signals for comparison are simplified down to five equidistant points representing QRS complex and five equichronic points from T wave. Then we compare median with the rest of the signal within the cluster and for the signals that differ too much from the median – it means usually 10-20% deviation from median – we create new cluster – where only non-accepted beats from this cluster can go.
After doing so with whole signal we compute medians in each class and compare it with the class extreme – if it is larger than we use direct shape comparison technique mentioned above.

Results

For testing of the clustering algorithm we have used 15 full 24 hours long recordings. On each recording we had about 30 minutes of annotated signal usually with several ventricular beats and other abnormalities. As a correctly clustered beat was taken such that majority of a cluster belonged to one group – results are in the Table 4.

Table 4:

Type of beat	Correctly classified	Incorrectly classified	Success ratio
N	44792	2396	94.6%
V	4618	327	95.4%
S	791	83	89.5%

Conclusion

Our aim in this stage of work was not to classify computed cluster so we evaluated just if there are within a cluster beats classified into more then one group. This had happened only in the very small clusters(like up to 10 beats in a cluster) and can be explained by error of beat detection – detected beat was either noise or was in noise embedded. We have developed algorithm so far in Matlab, but anyway its computational times are at about 10minutes for one 24 hour signal – and so we can say that we should be able to get the computational times of the whole process below 5minutes in C++.

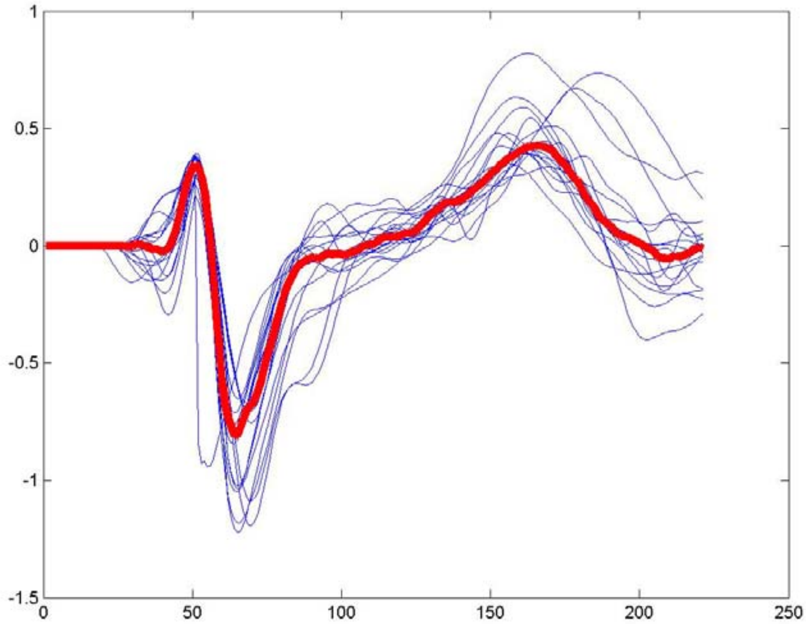

Figure 4: Example of one of the final clusters

Acknowledgement

This work has been supported by the research program "Information Society" under Grant No. 1ET201210527 "Knowledge-based support of diagnostics and prediction in cardiology ".

References

[1] HOLTER, N.J. (1961): 'New methods for heart studies', Science **134** pp.1214.

[2] WILSON, G.F.(2002): 'A Comparison of Three Cardiac Ambulatory Recorders Using Flight Data', The international journal of Aviation Psychology, *12*(1), pp. 111–119

[3] DASKALOV I.K., DOTSINSKY I.A., CHRISTOV I.I. (1998): 'Developments in ECG acquisition, preprocessing, rameter measurement, and recording', IEEE Eng. Med. Biol. 17, pp. 50-58

[4] SORNMO, L., BORJESSON, P.O., NYGARDS, M., PAHLM, O. (1981): 'A method for evaluation of QRS shape features using a mathematical model for the ECG', IEEE Trans. Biomed. Eng. 28, pp. 713-717.

[5] WOLF, T., NAGY, A.B. (1981): 'Analysis of ECG waveforms by clustering methods' Adv Cardiol. **28,** pp. 105-108

[6] DE CHAZAL, P., O'DWYER, O., REILLY, R.B. (2004): 'Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features', IEEE Transaction on Biomedical Engineering, **51**, No. 7

[7] CUESTA-FRAU, D., PEREZ-CORTES, J.C., ANDREU-GARCIA, G. (2003): 'Clustering of electrocardiograph signals in computer-aided Holter analysis', Computer methods and programs in Biomedicine 72, pp. 179-196

[8] SAXENA, S.C., SHARMA, A., CHAUDHARY, S.C. (1997): 'Data compression and feature extraction of the ECG signals' International Journal of Systems Science **23(5)**, pp. 483-498,

[9] LAGERHOLM, M., PETERSON, C., BRACCINI, G., EDENBRANDT, L., SORNMO, L. (2000): 'Clustering ECG complexes using hermite functions and self-organizing maps' IEEE Trans. Biomed. Eng., vol. 47, pp. 838–848

[10] HU, Y.H., PALREDDY, S., TOMPKINS, W.J. (1997): 'A patient-adaptable ECG beat classifier using a mixture of experts approach' IEEE Trans.Biomed. Eng., vol. 44, pp. 891–900

[11] CHRISTOV, I., BORTOLAN, G. (2004): 'Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks' Physiol. Meas. **25,** pp. 1281–1290