Feature Selection in Body Surface Potential Mapping

Vaclav Chudáček, Michal Huptych and Lenka Lhotská, *Member, IEEE*

***Abstract* — Body surface potential mapping is an extension of the conventional electrocardiography. Body surface potential mapping, measured from multiple places on the thorax, is a graphical representation of cardiac activity that enables us to acquire more data for more complex analysis of heart cycle. In this paper new features are presented and the methodology of obtaining them is described. In the last part of the paper usability of the features is demonstrated on dataset consisting of pregnant, non-pregnant and after-delivery women.**

I. INTRODUCTION

STANDARD 12-lead ECG is still the only universally accepted practical method used for diagnostics of heart diseases. However, it is not optimal and has its limitations as many research and clinical studies had shown (e.g. [1], [2]). Body surface potential mapping (BSPM) is an extension of the conventional electrocardiography that provides refined non-invasive characterisation of cardiac activity. The BSPM systems may use 30-300 electrodes. Increased spatial sampling on the body surface provides more in-depth information on the cardiac generated potentials, and thus in many cases exhibits better diagnostic value [3]. However, the BSPM has its drawbacks as well, e.g. more complicated measurement and electrode positioning. The most frequently used body surface maps are isopotential maps - giving a distribution of the potential at a specific moment, and isointegral maps - providing a distribution of the sum of potentials over a specified time interval. After the years the BSPM was used on experimental field the number of data seems to be large enough to enable us new knowledge about the way how the heart works.

Therefore the aim of this paper focuses on feature extraction from the body surface potential maps that should allow us by using the means of artificial intelligence to find possibly useful knowledge, so far not routinely analysed, for description of interesting events on the human heart.

Manuscript received June 30, 2006. This work has been supported by the research program "Information Society" under grant No. 1ET201210527 "Knowledge-based support of diagnostics and prediction in cardiology".

V. Chudáček is a PhD. student of Czech Technical University in Prague, Czech Republic, dept. of Cybernetics, Gerstner laboratory. (Contact e-mail: chudacv@fel.cvut.cz).

M. Huptych is a PhD. student of the Czech Technical University in Prague, Czech Republic, dept. of Cybernetics, Gerstner laboratory.

L. Lhotská is an assoc. professor of the Czech Technical University in Prague, Czech Republic, dept. of Cybernetics, Gerstner laboratory.

II. DATA ACQUISITION AND PREPROCESSING

A. Data acquisition

Cardiag 112.2 system (Czech-made BSPM device) with 80 electrodes in 16 x 5 equidistant matrix was used for measurement - see Figure 1. The system allows recording of standard ECG, vector-cardiograph and BSPM.

B. Data preprocessing

After all 95 signals consisting of 12 standard ECG leads, 3 orthogonal leads (VCG) and 80 mapping leads, signals are filtered with 50Hz notch filter to remove power-line frequency, and with set of adaptive filters for removing of breathing and MEG artefacts.

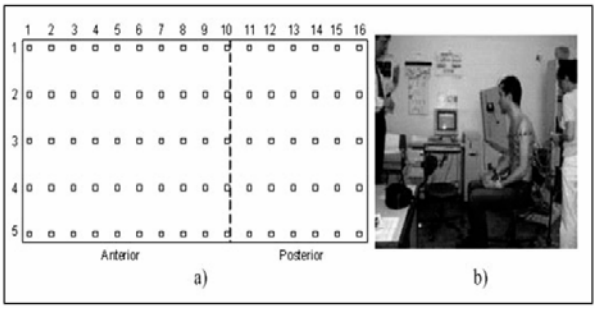
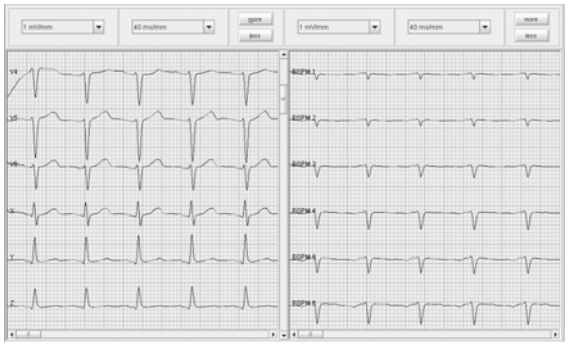


Figure 1: Electrode grid of the Cardiag 112.2 system

After the filtering of the signal, simple analysis was performed to find the erroneous signal – such as those with bad contact or signals too small for further consideration. Such signals were replaced with linear combination of it's four neighbors.

C. Signal analysis

Further step in the preprocessing phase was signal analysis – example of the program window of signal analysis see on Figure 2. We used wavelet-based signal detector that uses signal decomposition that uses discrete wavelet transformation to detect P, Q, R, S and T-waves onsets and offsets. It is important to mention, that the analysis was performed on all signals separately – thus set of 95 independent results was acquired for each parameter of the analysis searched.



Map creation

Figure 2: Window of signal analysis

The last step in the preprocessing phase was the creation of the maps – from which most of the features are then acquired. *Isopotential maps* were computed using bicubic transformation as a map in one time instant. The *isointegral maps* of QRS and QRST complexes were computed as an sum of isopotential in the given time interval.

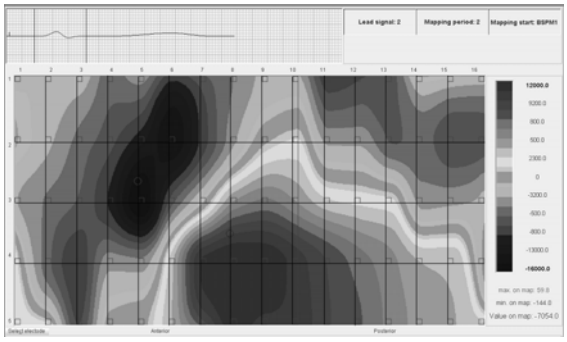


Figure 3:Isopotential map

III. FEATURE EXTRACTION

Feature extraction and feature selection are very important steps in the process of knowledge discovery and in classification systems. Proper selection of features may significantly influence the success rate of classification. Use of irrelevant and weakly relevant features can decrease the accuracy. The features can be selected either automatically or manually. Automatic selection can be viewed as a state space search where each state represents a single combination of features. The goal of the search is to find the state with the highest value of the evaluation function that characterises the success rate of classification with the corresponding features. It is obvious that such an evaluation function is only an estimation of the success rate of the classification, because the training set is limited. The transition operator is feature adding or deleting. The average accuracy of cross-validation is usually used as an evaluation function. Manual selection, on the other hand, is a more or less intuitive process based on experience.

One of the most important aspects of the ECG classification systems is reliable analysis of ECG records, which enables significant values to be identified on the measured signal. This analysis is a necessary condition for correct classification.

We have analysed a number of parameters that can be computed from the measured signals. Based on our previous experience with ECG signal pre-processing and classification [4], [5], and a number of experiments we have identified several parameters that may contain the significant information.

Features that were selected for further consideration are described in the following sections followed by the way of their computation.

A. *ECG features*

Basic features from standard ECG consist of:

- important intervals (QRS, QTc, PR) - QRS is interval between Q-wave onset and S-wave offset, QTc is corrected QT interval using formula by Fridericia [10]:

$$QT_c = \frac{QT}{RR^{\frac{1}{3}}} \quad (1)$$

- amplitudes of waves (P, R, T) where in case of R-wave the biggest amplitude is taken, in case of P and T – waves maxima of positive and negative parts of wave were taken
- intervals and amplitudes are computed for each lead of the system separately so 95 separate measurements are computed

B. *VCG features*

- two description parameters of the QRS-loop – QRSw (width) and QRSa (angle)
- two description parameters of the T-wave loop – Tw (width) and Ta (angle)

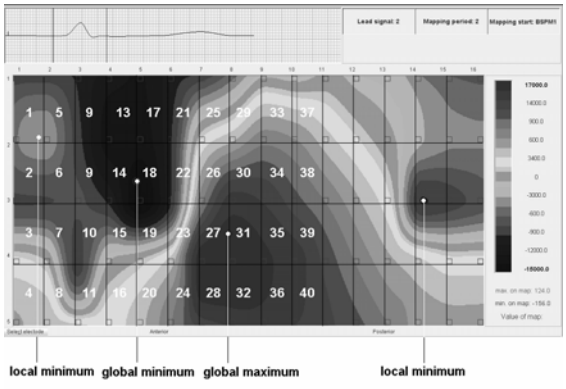


Figure 4: Isointegral map with 40 equally spaced squares

C. BSPM global features

For computation of local and global maxima of the isointegral map, it was divided into 40 equal squares (see Figure 4), that were used further for description of the map features.

- Position of global maximum and minimum of isointegral maps (QRS and QRST)
- Position of local maxima and minima of isointegral map, for example of such a case see Figure 4
- Zero-line of the map. Position of zero line of the isointegral map has been correlated to various changes on the heart. To be able to compare with other BSPM lead systems we represent zero-line by the consecutive array of squares the zero-line passes

Figure 5: Representation of the zero-line

D. BSPM local features

- Distribution of positive, negative and neutral pixels in each of 40 leads from anterior part of thorax (see Figure 6).

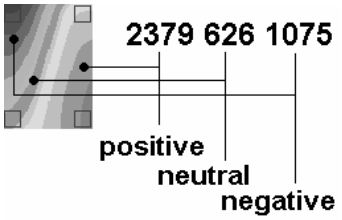


Figure 6: Example of detailed description of one square from the Figure 4.

IV. RESULTS

In the previous chapter extraction of features was described and in this section results using features from BSPM are presented.

Data for the test consisted of 88 women - in pregnancy (44) after delivery (26) and control group (18). We simulated the same experiment done in [8] using the parameters shown in the Table 1. We used the above described features and we can say that although results suffer from dependency on small data set, and problems with medical data in general, such as non-heterogeneity of biological systems and noise induction during measurement period, some interesting results have been obtained. An example of a generated decision tree using Weka software tool [6] is shown on the Figure 7.

We have found following interesting (important) differentiating features:

Description of vectorcardiographic T-loop Tw and Ta quite unambiguously differentiate the reference group from the pregnant and after-delivery groups.

Position of spatial maximum of isopotential map differentiates the group of pregnant women from the reference group. Position of spatial maximum of pregnant women has been usually found caudally (low in the direction of y axis) than in the reference group Tmax-Toff parameter includes information about the angle of the heart axis, meaning heterogeneity of repolarization.

The groups of pregnant and after-delivery women are merging partially. The group of pregnant women that is stronger has "pulled" nearly half of the after-delivery group. Our analysis is thus in correspondence with medical observations indicating that changes of electrical field of the heart caused by changed geometry of thorax are covered by other more significant changes caused by pregnancy and remaining for certain period after delivery; for example, changes of inner conditions of the body (hormonal changes, retaining fluids) cause changes of electric conductivity of tissues.

TABLE 1.SELECTED PARAMETERS OF DATA MEAN VALUES, SD – STANDARD DEVIATION, QT – LENGTH OF QT INTERVAL, QTC – LENGTH OF QT RELATED TO RR INTERVAL, TA – AMPLITUDE OF T-LOOP, TW – ANGLE (WIDTH) OF T-LOOP

	Healthy control	Pregnant	After delivery
QT[ms]	342	322	325
SD	13	12.8	11.78
QTc[ms]	382	419	413
SD	35.5	39	33
Ta	597	413	426
SD	141	151	133
Tw	8.1	11	12.5
SD	4	8.5	8.5

However, we have not confirmed unambiguously several expected characteristics, namely extremes of isointegral maps, position of the extremes of isointegral maps. We have analysed the original data with respect to these results and have come to the following conclusion. Measurement of both extremes is burdened by:

- objective error of the measurement – drift of signal zero has higher influence on map minimum because most of pronounced T waves are in positive leads;
- error given by current skin conductivity of the measured patient, this error contributes more significantly to increase of maximum;
- error caused by incorrect contact

V. CONCLUSION

At present body surface potential mapping is the most complete visualisation of the heart activity mapped on body surface. However its problem is the interpretation of the majority of measured data, with the exception of measurement of isopotential maps (measurement of differences of map values from healthy etalon). This study has tried to discover new relations in measured evaluated data that might contribute to deeper understanding of electrophysiological heart activity influenced by physiological changes of thorax geometry (heart position in thorax).

Decision tree:

```
Ta <= 547.6988:
:...TmaxToff > 37:
:   :...pozice_maxima_z_mapy03x <= 37: 1 (3)
:   :   pozice_maxima_z_mapy03x > 37: 2 (3)
:   TmaxToff <= 37:
:       :...pozice_maxima_z_mapy08x <= 59:
:           :...T_max_poz82 <= 0.78626: 2 (32.6/11.5)
:           :   T_max_poz82 > 0.78626: 3 (15.4/2.9)
:       pozice_maxima_z_mapy08x > 59:
:           :...T_max_hod01 <= 378.8304: 1 (2)
:               T_max_hod01 > 378.8304: 2 (8)
Ta > 547.6988:
:...Tw > 14.5955: 3 (2)
    Tw <= 14.5955:
        :...T_max_hod46 > 148.9341: 2 (5)
            T_max_hod46 <= 148.9341:
                :...T_max_poz75 <= 0.7622: 2 (2.1)
                    T_max_poz75 > 0.7622: 1 (14.9/1.9)
```

Size of decision tree : 10.
Accuracy: 79,5%

```
(a) (b) (c) <-classified as
---- ---- ----
18          (a): class 1
   2 42     (b): class 2
  16 10     (c): class 3
```

Rules:

```
Rule 1: (14/1, lift 4.3)
Ta > 547.6988
Tw <= 14.5955
T_max_hod46 <= 148.9341
T_max_poz75 > 0.7622
```

Figure 6: Decision tree used for classification of dataset consisting of pregnant women

We have designed a new tool for visualization of body surface potential maps. We have developed several modules for further features extraction, and we have tried to acquire information hidden in body surface potential maps. Concluding from the above mentioned results we can state that heart changes due to pregnancy are covered by larger changes of conductivity of thorax caused by changes in the body. Further investigation that will provide more discriminative set of features is required for revealing relations between heart position and BSPM.

ACKNOWLEDGMENT

Authors would like to thank professor Kittnar and dr. Mlcek from 1st Medical School of Charles University for the help with the data acquisition.

REFERENCES

[1] Kubota I, Ikeda K, Ohyama T, Yamaki M, Kawashima S, Igarashi A, Tsuiki K, Yasui S. Body surface distributions of ST segment changes after exercise in effort angina pectoris without myocardial infarction. American Heart Journal, 110: 949-955, 1985.

[2] Yanowitz F, Vincent M, Lux RL, Merchant M, Green LS, Abildskov JA. Application of body surface mapping to exercise testing: S-T80 isoarea maps in patients with coronary artery disease. American Journal of Cardiology, 50: 1109-1113, 1982.

[3] Kornreich F, Appropriate electrode placement in evaluating varied cardiac pathology. In Liebman J (ed) Electrocardiology '96. From the cell to the body surface. publ World Scientific 1997: 83-92.

[4] Lhotská, L., Fejtová, M., Macek, J., Novák, D.: Feature Extraction From Biological Signals: A Case Study. In: IEEE 4th International Conference on Intelligent Systems Design and Application. Piscataway: IEEE, 2004, vol. I,II, pp. 139-144

[5] Macek, J., Lhotská, L., Peri, D.: Evaluation of ECG: Comparison of Decision Tree and Fuzzy Rules Induction. In: Cybernetics and Systems 2004. Vienna: Austrian Society for Cybernetics Studies, 2004, pp. 713-718

[6] See5 demo version 1.19. [online] <http://www.rulequest.com/see5-info.html>

[7] WEKA(Waikato Enviroment for Knowledge Analysis), University of Waikatom Dep. Of Computer science, New Zealand. [online] <http://www.cs.waikato.ac.nz/~ml>

[8] Lechmanová, Kittnar, O., Mlcek, M., Slavíček J, Dohnalová A, Havránek, J., Kolařík J., Pařízek A., QT Dispersion and T-Loop Morphology in Late Pregnancy and After Delivery in Physiol. Res. 51: pp. 121-129, 2002

[9] Witten, I. and Frank, E. Data mining: practical machine learning tools and techniques with Java implementations, San Francisco: Morgan Kaufmann, 2000.

[10] Fridericia, L.S. (1920) Die Systolendauer im Elektrokardiogram bei normalen Menschen und bei Herzkranken. Acta Medica Scandinavica, **53,** 469-486