

ECG CLASSIFICATION USING CD-HMM FOR EDUCATIONAL PURPOSES

Ing. Dina KIČMEROVÁ, Doctoral Degree Programme (2)
Department of Biomedical Engineering, FEEC, BUT
E-mail: xkicme00@stud.feec.vutbr.cz

Supervised by: Prof. Ivo Provazník

ABSTRACT

The automatic detection and classification of cardiac arrhythmias is important for diagnosis of cardiac abnormalities. Modern education is not possible without the use of specialized program tools and visual demonstrations of new methods. This work is focused on classification of normal sinus rhythm and premature ventricular contractions of the human heart. The proposed method uses wavelets for the feature selection and extraction (searching for a local maximum in the contour envelope successfully detects R-peaks) and Continuous Hidden Markov models for the classification. The ECG data is taken from standard MIT-BIH arrhythmia database. The method was applied in a custom-made educational tool used in a Introduction to Medical Informatics master course.

1 INTRODUCTION

Information about development of new algorithms and methods are very important in educational system. Subjects should follow recent research results and techniques.

The medical domain is one of the areas in which automatic processing systems are most frequently applied. This is quite natural because modern medicine generates huge amounts of data, but at the same time there is often a lack of data understanding. New methods can help in dealing with this problem, they can simplify and usually speed up the processing of large volumes of data. New algorithms work in time-frequency domain and combine some advantageous characteristics known from classical methods – mainly they allow frequency analysis with time information about analyzed features.

The automated detection and classification of cardiac arrhythmias is important for diagnosis of cardiac abnormalities. In practice, unsupervised methods may be required to detect arrhythmias in changing environment. Hidden Markov models are often used for such tasks in speech recognition but also in ECG processing. To obtain improved results, HMM can be combined with wavelet transform [3].

The work is focused to the detection of premature ventricular contractions (PVC) among normal sinus rhythm (NSR). Premature ventricular contractions - these early depolarizations begin in the ventricle instead of the usual place, the sinus node. They are very

common, and are sometimes perceived as a palpitation. They often occur without the patient being aware of it at all. PVC's occur in bigeminy, trigeminy, quadrigeminy, ventricular tachycardia, ventricular fibrillation, etc.

An increased frequency of PVC's in patients with heart disease is statistically predictive of ventricular fibrillation and sudden death. In patients with some types of heart disease, PVC's or ventricular tachycardia do indicate an increased risk of serious arrhythmias. Therefore this work is focused on their detection.

2 METHODS

2.1 CONTINUOUS WAVELET TRANSFORM

In the proposed method, input data were transformed by continuous wavelet transform (CWT):

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left(\frac{t-b}{a} \right) f(t) dt \quad (1)$$

where a is a scale and b is a time shift. Time-frequency spectrum enables to measure time-frequency changes in spectral components. Interpretation of a time-frequency resolution by CWT is as follows: CWT represents time-frequency decomposition realized by correlation of signal $f(t)$ with basic functions derived from the mother wave $\psi(t)$. Haar function was used as the mother wavelet ψ for its simplicity:

$$\psi_H(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2} \\ -1 & \frac{1}{2} \leq t < 1 \\ 0 & \text{others} , \end{cases} \quad (2)$$

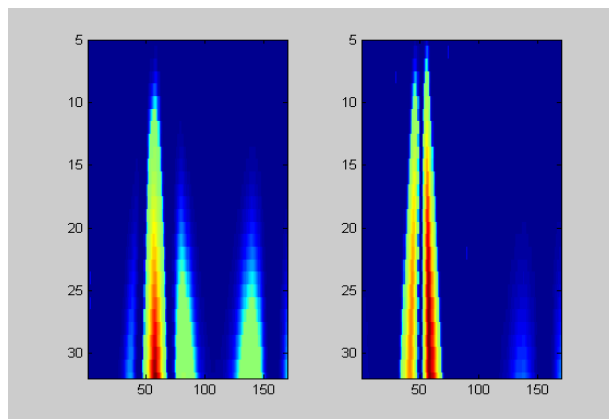


Fig. 1: Continuous wavelet transform of the raw time data series (left panel – premature ventricular contraction, right panel – normal sinus rhythm).

Three problems are considered as fundamental in hidden Markov modelling applications:

- 1) Estimation of hidden Markov model parameters from a set of representative training data (parameters include state transition probabilities, output probabilities).
- 2) Efficient calculation of $P(O|\lambda)$ - the probability that a given observation sequence was generated by a particular hidden Markov model λ .
- 3) Determination of X^* , the most likely underlying state sequence corresponding to a given observation sequence O such that $P(O|X^*,\lambda)$ is maximum.

The importance of solving the first problem is obvious; model parameters must first be estimated before the models can be used for classification purposes.

Baum-Welch algorithm was used as a training method to find hidden Markov model parameters A, B with the maximum likelihood of generating the given symbol sequence in the observation vector.

The probability of state occupation must be calculated. This is done efficiently using the so-called Forward-Backward algorithm.

The solution to the second problem is often used as the basis for a classification system. By computing $P(O|\lambda)$ for each of i possible models and choosing the most likely model, classification can be inferred.

An alternative classification approach uses the solution of the third problem to find the single best state sequence which maximizes $P(O|X^*,\lambda)$. Classification can then be inferred by choosing the model with the most likely best state sequence, which requires less computation than determining the most likely model.

The logarithmic Viterbi algorithm was used for the recognizing. It determines the most probable route to the next state, and remembers how to get there. This is done by considering all products of transition probabilities with the maximal probabilities already derived for the preceding step. The largest such is remembered, together with what provoked it.

Scaling the computation of Viterbi algorithm to avoid underflow is non-trivial. However, by simply computing of the logarithm it is possible to avoid any numerical problems.

3 RESULTS

The proposed method was used in an educational graphical program in Matlab environment. The program allows selection of a number of states and initial transition probabilities of the HMM model. After the reading of signals, students can visually follow whole classification process. They can also see results of the wavelet transform of the signal and the training process. They can compare results of the classification using transformed and raw ECG data.

The ECG data were taken from the standard MIT-BIH arrhythmia database automatically as a sequence of 250 samples (100 samples before R-waves and 150 after the R-waves). Graphical interface of the realized program is depicted on Fig.3.

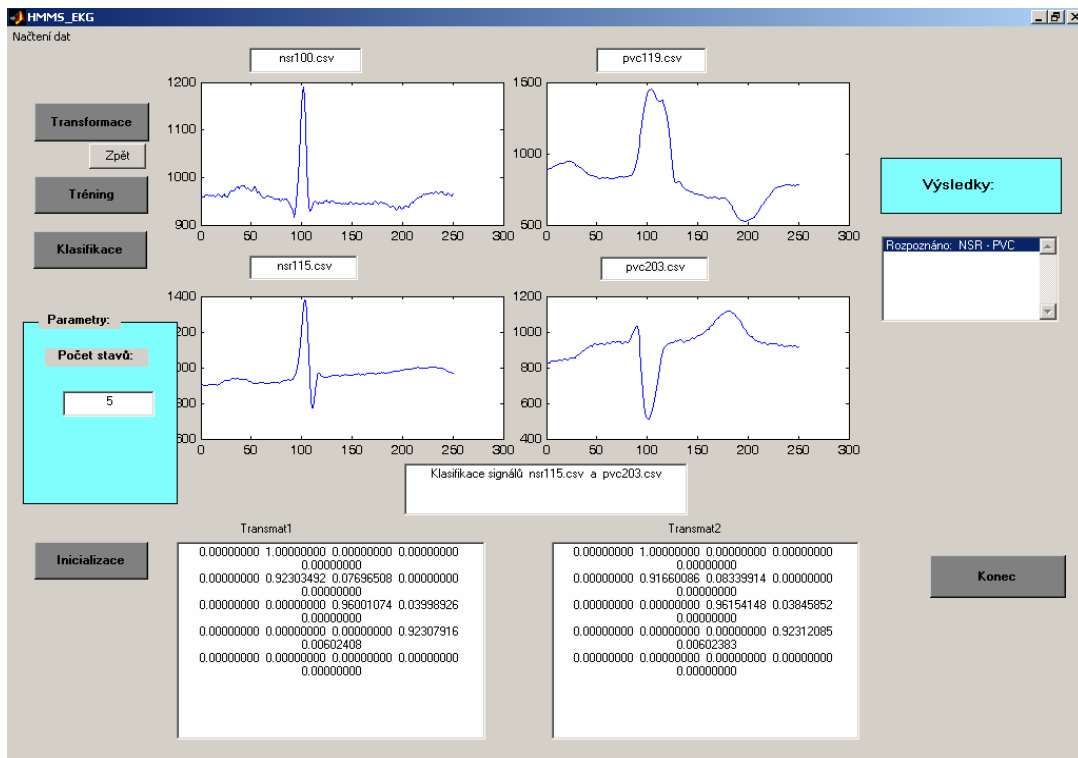


Fig. 3: Custom-made program tool interface

4 DISCUSSION AND CONCLUSION

An educational tool – a program in Matlab environment is proposed. Students can visually follow all classification process. An algorithm employs unsupervised way of the classification. This work is focused on classification of normal sinus rhythm and premature ventricular contractions. There is demonstrated that wavelet methods can be used to generate an encoding of the ECG which is tuned to the unique spectral characteristics of the ECG waveform features. Left-to-right CD-HMM was used.

ACKNOWLEDGEMENT

The research was supported by the grants 102/04/0472 and 305/04/1385 from GACR and the Research Programme of Brno University of Technology MSM 0021630513.

REFERENCES

- [1] Kičmerová, D.: Detection and classification of ECG signals in time-frequency domain. Diploma Thesis. Brno University of Technology, 2004.
- [2] Rabiner, L R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2):257–286, 1989.
- [3] Bardoňová, J., Provozník, I., Nováková, M., Nováková, Z.: Hidden Markov model in wavelet analysis of myocardial ischemia in rabbit. In Computers in Cardiology, p. 419–422, Boston, 2000.