Comparison of Linear and Non-linear Features for Intrapartum Cardiotocography Evaluation – Clinical Usability vs. Contribution to Classification

Václav Chudáček[1], Jiří Spilka[1], Michal Huptych[1], George Georgoulas[2], Petr Janků[3], Michal Koucký[4], Chrysostomos Stylios[5], Lenka Lhotská[1]
[1] dept. of Cybernetics CTU in Prague,
[2] dept. of Obstetrics and Gynaecology FN Brno,
[3] dept. of Obstetrics and Gynaecology 1.LF CUNI in Prague
[4] Dept of Computer Applications in Finance and Management, TEl of Ionian Islands, Lefkas
[5] Technological Institute of Epirus, Artas, Greece
chudacv@fel.cvut.cz

1 Introduction

Instrumental evaluation of the fetal well-being during delivery is more than hundred years old. Auscultation - sensing of the fetal heart rate (fHR) using a fetal stethoscope - introduced by Pinard in 1876 - was replaced in 1960's by electronic fetal monitoring(EFM) with cardiotocography (CTG - recording of fetal heart rate and force/pressure of contractions) as the most important representant.

Although introduction of the EFM was accompanied by large expectation, since it offered continuous fetal surveillance, meta-analysis of large multicentric studies reported did not prove any significant improvements in the delivery outcomes. Some studies additionally disapproved any evidence of advantages of continuous monitoring compared to intermittent one. Moreover EFM became the main suspect for increased rate of cesarean sections.

In order to improve interpretation and thus lower the number of asphyxiated neonates CTG guidelines were introduced [1]. Even though the guidelines are available for more than twenty years poor interpretation of CTG still persists with inter-observer as well as intra-observer assessment variations [2].

First attempts of automatic CTG analysis [3] followed FIGO guidelines that describe morphological changes in CTG. Those morphological features became fundamental for almost all methods that attempt to classify fetal status. The extraction of morphological features was improved by Bernades [4] and resulted in development of SisPorto - automatic system for CTG analysis.

Linear and nonlinear methods used for fHR analysis were mostly derived from adults HRV research. This field was thoroughly investigated and a general agreement on HRV analysis exists [5]. The statistical description of CTG tracings was employed in work of Magenes [6] and then in following study of Goncalves [7]. Another approach of fHR analysis is to examine frequency content by spectral analysis. This analysis was performed by many research groups and recent paper [8] gives a short overview of works which analyzed fHR spectrum. The fHR was also analyzed by various wavelets with different properties [9]. The different estimations of fractal dimension were reviewed by [10]. The most successful nonlinear methods for fHR analysis so far are approximate entropy (ApEn) and sample entropy (SampEn) [7].

2 Data

The fHR signals used in this work were measured externally using Doppler ultrasound or internally by scalp electrode. Our data set consists of 476 delivery recordings. The data were obtained at the Obstetricians ward of General Teaching Hospital in Prague on STAN S21 devices. The STAN system allows acquisition of the fetal heart rate via scalp electrode as well as the noninvasive Doppler ultrasound acquisition as depicted on Figure 1. In 60% (280) of all cases the measurement of fHR was done by ultrasound, in the rest electrode was attached to the scalp of the fetus. Signals were then annotated by five experts with at least five years of praxis as obstetricians; the process is described in chapter 5.

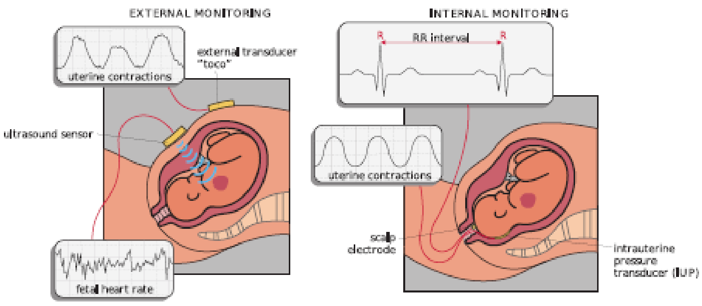

Fig. 1. Recording of the fetal heart rate and uterine activity [9]

Recordings obtained with the scalp electrode have usually fewer missed values and are in general less noisy.

All the recordings had to be checked for patient anamnesis and only one fold pregnancies delivered during 38^{th} – 42^{nd} week of pregnancy were chosen for the final database.

Additionally umbilical artery pH was obtained as and objective evaluation of hypoxia, where pH value of less then 7.15 was considered pathological. The neonatal acidemia is defined as pH below 7.05 - these values were suggested in the work of Sundstrom [11].

Nevertheless there exists other works suggesting other values 7.10 or 7.15[10]. Considering these facts and on recommendation by obstetricians at the CUNI we used border pH of 7.15.

Pathological recordings are very hard to get in most fields of clinical medicine and obstetrician ward is no exception. Therefore in our case we have decided to classify the fetuses as normal (i.e. without sustained hypoxia) if having pH above or equal to 7.15 and abnormal otherwise.

3 Methods

To be able to obtain and evaluate features from the fHR signal following consecutive steps were necessary.

Data pre-processing

Artifacts removal; interpolation; choice of appropriate segment; and detrending of the signal were the preprocessing steps undertaken. Detailed description of the preprocessing phase is out of scope of this article therefore we will mention only segment selection in greater detail since it might play important role in the evaluation process of clinical usefulness of our approach. The selected segment is always twenty minutes long, selected as close to delivery as possible, excluding signals with large amount of noise, usually during the active pushing.

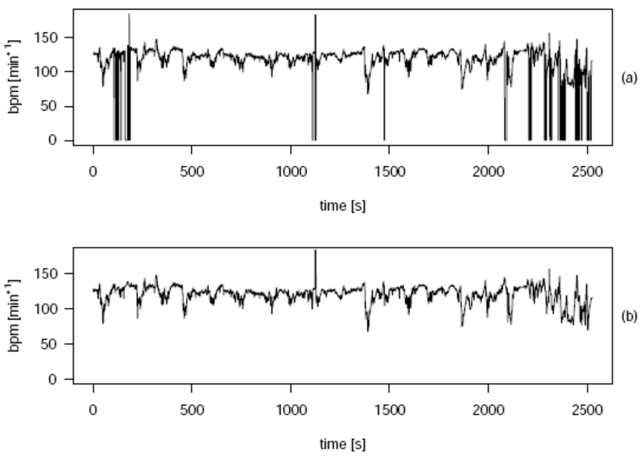

Fig. 2. Example of raw fHR signal (a) and signal after preprocessing (b).

Feature extraction

From the aforementioned selected segment linear and nonlinear features were computed. Many of these features are known in the field of HRV processing, but completely unknown in obstetrics and their benefits has yet to be proven.

The linear features computed were: Description of the fHR baseline using mean, median and SD measures, SDNN, RMSSD, mean of RR interval, and NN50 from the time domain.

The non-linear features computed were: Fractal dimension of attractor, fractal dimension of waveform, entropy, and complexity. The particular methods used to compute the features were: correlation method for estimation of attractor dimension; Higuchi's, variance, and box counting method for estimation of waveform fractal dimension; approximate and sample method for estimation of entropy and also the Lempel Ziv Complexity.

Feature evaluation

Two fold evaluation of the features was employed. Firstly all features were investigated for their informational gain (separately as well as in combinations) and (semi-)automatic feature selection was performed which resulted into several sets of features ranging from the fully-automatic selection set to sets based on clinical guidelines or easy-to-interpret features.

Afterwards features were used for data classification with objective and subjective classification using algorithms such as decision trees, SVM, k-nn or neural networks.

4 Results

When computing the results we have utilized 10-fold cross-validation using following classifiers: Naive Bayes, Support Vector Machine (SVM), and C4.5 decision tree, all implemented in WEKA. Short description and further references for these methods can be found in [12].

For the SVM, the polynomial kernel and penalty parameters $C = 1$ were used. The classification results are presented in Table 1. From all performance measures, the specificity is of major importance since a classifier with higher specificity causes lower number of false alarms that leads to lower rate of unnecessary intervention. Regarding the specificity, the SVM performed best. However, statistical tests revealed that difference between individual classifiers is statistically insignificant on $p<0.01$ confidence level.

We achieved classification results of 78% of sensitivity and 70% specificity using nonlinear features only. These results are well comparable to inter-observer variability [9].

Table 1. Performance of the best feature sets using the best classifier

All in [%]	NaiveBayes	SVM	C4.5 Tree
Accuracy	73	72	65
Sensitivity	84	78	74
Specificity	64	70	57
AUC	79	74	69

5 Discussion and conclusions

For the first we attempted to compare various types of features on one fairly large dataset the same one for all the experiments. We have found several promising new features namely baseline standard deviation, sample entropy and Higuchi's fractal dimension, that might be useful for everyday use in the obstetrician wards and that could perform better then the features so far used according to the FIGO guidelines. Nevertheless the final say will have the manufacturers of the devices, that will need to implement them – activity recently underway.

Acknowledgement

This work was supported by the research program No. MSM 6840770012 “Transdisciplinary Research in the Field of Biomedical Engineering II” of the CTU in Prague, sponsored by the Ministry of Education, Youth and Sports of the Czech Republic.

6 References

[1] FIGO. Guidelines for the use of fetal monitoring. International Journal of Gynecology & Obstetrics, 25:159–167, 1986.

[2] I. Amer-Wåhlin, I. Ingemarsson, K. Marsal, and A. Herbst. Fetal heart rate patterns and ECG ST segment changes preceding acidaemia BJOG, 112(2):160–165, Feb 2005.

[3] B. Guijarro-Berdinas, A. Alonso-Betanzos, and O. Fontenla-Romero. Intelligent analysis and pattern recognition in ctg signals. Artif. Intell., 136:1–27, 2002.

[4] G. Georgoulas, C. D. Stylios, and P. P. Groumpos. Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machines. International Journal on Artificial Intelligence Tools, 15:411–432, 2005

[5] O. Palomaki, T. Luukkaala, R. Luoto, and R. Tuimala. Intrapartum cardiotocography – the dilemma of interpretational variation. J Perinat Med, 34(4):298–302, 2006.

[6] R. Mantel, H. P. van Geijn, F. J. Caron, J. M. Swartjes, E. E. van Woerden, and H. W. Jongsma. Computer analysis of antepartum fetal heart rate: 2. Detection of accelerations and decelerations. Int J Biomed Comput, 25(4):273–286, May 1990a.

[7] J. Bernardes, C. Moura, J. P. de Sa, and L. P. Leite. The Porto system for automated cardiotocographic signal analysis. J Perinat Med, 19(1-2):61–65, 1991.

[8] P. Hopkins, N. Outram, N. Löfgren, E. C. Ifeachor, and K. G. Rosén. A comparative study of fetal hrv analysis techniques. Conf Proc IEEE Eng Med Biol Soc, 1: 1784, 2006.

[9] B. Guijarro-Berdinas, A. Alonso-Betanzos, and O. Fontenla-Romero. Intelligent analysis and pattern recognition in ctg signals using a tightly coupled hybrid system. Artif. Intell., 136:1–27, 2002.

[10] A. M. Fraser and H. L. Swinney. Independent coordinates for strange attractors from mutual information. Physical Review A, 33(2):1134–1140, 1986.

[11] L. Cao. Practical method for determining the minimum embedding dimension of a scalar time series. Physica D, 110:43–50, 1997.

[12] Ian H. Witten and Eibe Frank (2005) "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.