Evaluation of Feature Subsets for Classification of Cardiotocographic Recordings

V Chudacek[1], J Spilka[1], B Rubackova[1], M Koucky[2],
G Georgoulas[3], L Lhotska[1], C Stylios[4]

[1]Czech Technical University Prague, Czech Repub
[2]dept of Obstetrics and Gynaecology 1LF CUNI, Prague, Czech Repub
[3]Georgia Institute of Technology, Atlanta, USA
[4]Technological Institute of Epirus, Artas, Greece

Abstract

Electronic fetal monitoring - continuous recording of the cardiotocogram (CTG) - consisting of fetal heart rate (fHR) and tocographic signals, is a method used for intrapartal evaluation of the well being of the fetus. In this paper we evaluate several subsets of features for outcome classification of the pregnancy based on the CTG recording of the last 20 minutes preceding actual delivery. The features subsets are created based on PCA, information gain and Group of Adaptive Models Evolution (GAME) neural network feature selection algorithm. Our data set consisted of 104 intrapartum recordings including 18 pregnancies with acidemia reflected in umbilical artery pH<7.20. The results show that the best subset consisting of mix of time-domain and non-linear features performs consistently over the whole data set with sensitivity and specificity over 70%, which is well comparable to interobserver variations.

1. Introduction

Being born is one of the most critical events in our live. After an intrauterine growth and development a baby is going to establish itself as an independent individual. To handle the labour stress, a fetus has to be equipped with a defense mechanism. A good understanding of the way individual fetus reacts to the stress of labour helps us to intervene in appropriate manner when a fetal defense has been activated, but before risk of long-term sequel increases.

To be more specific, during labour fetus can repetitively suffer by oxygen insufficiency and as a consequence the metabolic acidosis can be developed. Severe hypoxic injury can leads to neurodevelopment disability and cerebral palsy or even death. Electronic fetal monitoring (EFM) can be used to diagnose the hypoxia and thus prevent these adverse outcomes.

EFM has become a generally accepted method for the fetal surveillance during pregnancy and labour and offers important information about the fetus behavior. EFM involves the cardiotocogram (CTG), which consists of two signals, i.e. the intrauterine pressure (IUP) and the instantaneous activity of fetal heart (fHR) or as recently introduced fetal electrocardiogram (fECG). The introduction of CTG monitoring to clinical practice significantly reduced the incidence of birth asphyxia; however, on the other hand it has contributed to the rise of cesarean section [1]. This has been a consequence of poor interpretation of fetal behavior and CTG signals. The guidelines for CTG were introduced [2, 3] in order to improve the interpretations and thus lower the number of born children with acidosis and decrease the rate of cesarean sections as well.

Even though the guidelines are long time available, the poor interpretation of CTG still persists [4, 5]. The variation is not only inter-observer but also intra-observer [6, 7]. The introduction of fECG ST analysis, provided by STAN, to clinical practice improved the outcome of labour [8, 9]. A recently published paper [1] concluded that the main reason for the poor outcomes lies in the generally poor standard of CTG interpretation and the contribution of human factor as well. Therefore, more intense training on CTG interpretation should be performed [4, 5] or more cost-effective solution by developing an expert system [1] should be used.

2. Data used

Data used for this particular study were collected on STAN S21 system that allows acquisition of the fetal heart rate via scalp electrode – method used in about one fifth (21) of our recordings. Those recordings had more accurate signals comparing to the rest of the fHR

obtained using ultrasound electrode used with the same device.

The dataset consisted of 104 intrapartal recording was selected containing 18 recording with umbilical artery pH < 7.20 – considered pathological. The rest of signals had umbilical artery pH > 7.20 and thus were considered normal. All the deliveries were in between 38th and 41st gestational weeks.

All the recordings had been 20 minutes long with the starting point not further then 25 minutes before the actual delivery – disconnecting of the electrodes.

3. Data preprocessing

The fHR signals were initially preprocessed - where preprocessing involved partial removal of artifacts and, more importantly, correction of the bpm levels.

This type of correction is presented in Figure 1 where fHR is shown in raw and corrected state and also the intrauterine pressures signal is depicted there as well.

The fHR contains a lot of artifacts caused by mother and fetal movement or displacement of the transducer. In general, the amount of missing data ranges between 20% and 40% of all the data. The preprocessing part is therefore of a major importance and significantly affects consequent analysis of the fHR. We employed the preprocessing process used by Bernardes et al. [10, 12]. That is, whenever there is difference between successive beats higher than 25 bpm the linear interpolation is applied. The interpolation is between the first of two samples and first sample of a new stable segment, where five successive beats with difference lower than 10 beats between them are consider as a stable segment. The results are shown in Figure 2. Let's note that artifacts occur mostly at the end of labour because it is very stressful period.

The pre-processing of the CTG data is very important step and in our case consisted of several steps. Firstly gaps in the signal containing only zeros had to be dealt with. Gaps shorter then 3 minutes were linearly interpolated. The longer then 3 minutes gaps were considered as segment division (final segment did not contain signal with longer then 3min gap). In the second step artifacts were removed. All peaks differing more than 25% from the median of preceding 3 beats were replaced by that median. Then resampling of the irregularly sampled data (CTG as well as TOCO signal) to the 4 Hz sampling frequency took place using Hermite transformation. As the last pre-processing step we have filtered the data using moving average filtering. As a result of the pre-processing we have obtained one 20 minute long segment as close to the delivery as possible from each recording.

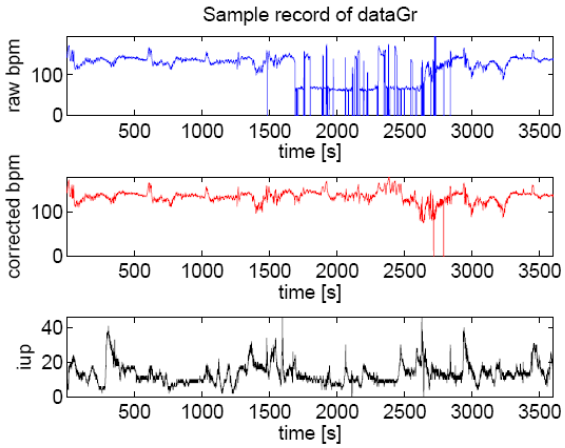

Figure 1: Correction of the bpm levels

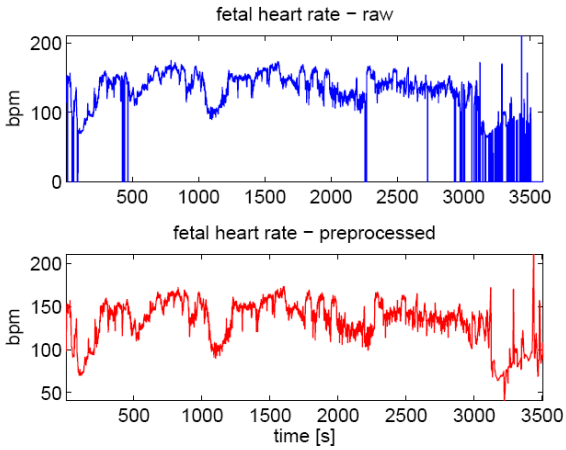

Figure 2: Artifacts removal. The raw signal with artifacts (upper) and signal after preprocessing stage (lower).

4. Feature extraction and selection

From the preprocessed 20 minute long segments 41 features had been extracted. Those features can be divided into several categories.

First are the morphological features introduced by the FIGO guidelines that describe the shape and changes of the baseline such as baseline mean and median, number of accelerations and decelerations present in the recording, long-term and short term variability, interval index etc. More detailed description of the features can be found in [2, 13].

The second category of features were those used routinely in adult HRV evaluation such as NN50, RMSSD, as well as the frequency features describing the amounts of energy in different energy bins [14].

The last category of features we have computed is the set of nonlinear features. Nonlinear features had been proven to reveal clinical relevant information of fHR that is not apparent in time and frequency domains [15]. This category included sample and approximate entropy, Poincare plot descriptors; and fractal dimensions characteristics [16, 17].

For the selection of the feature subsets with more generalized descriptive capabilities we have used three different methods – PCA and information gain methods as provided by WEKA [18] software and GAME feature selection algorithm [19]. The features included at least in two of the three subsets were used to create the fourth subset of features – the metaselection subset.

5. Results

The three used selection methods provided three slightly different subsets of features that can be seen in the Table 1. The metaselection subset was acquired from the subsets whenever the 2 out of 3 rule was satisfied.

Before the classification stage we had to balance the feature sets. We have used the synthetic minority oversampling technique as described in [20].

Table 1: Features selected according to the method used. The detailed description of the features can be found in the references mentioned in the section 2. II in the abbreviations stands for interval index, Fd – fractal dimension, DFA – detrended fluctuation analysis index, LTV – long term variability, LF_HF ratio of the energies in Low and High frequency bins, ApEn approximate entropy

Feature selection method	Features selected
PCA	baselineSD, deltaTotal, meanLTV, medianLTV, fdBoxCountDs, ApEn, meanII, baselineMean, medianII
InfoGain	baselineSD, deltaTotal, LF_HFMF, medianLTV, meanSTV, meanII, fdBoxCountDs, FdHiguchiDs, baselineMean, stdLTV
GAME-NN selection	baselineSD, meanII, FdHiguchiDl, baseline_Mean, ApEn, DFA_1, FdBoxCountDs, deltaTotal, LF_HF
MetaSelection	baselineSD, deltaTotal, meanII, medianLTV, fdBoxCountDs, ApEn baselineMean,

For the classification - evaluation of the feature sets we have used the GAME neural network and the results are presented in the Table 2.

Table 2: Results of the evaluation of the feature subsets

Feat. subset	Sensitivity [%]	Specificity [%]
PCA	69,0	71,0
InfoGain	65,5	64,5
GAME-NN selection	72,0	65,0
MetaSelection	75,0	72,5

6. Discussion and conclusions

We have evaluated performance of several subsets from the 41 computed features. The best results achieved were achieved by the set handpicked based on the automatic selection method results. Results of more than 70% sensitivity and specificity are comparable with the interobserver variability. Nevertheless there are some issues namely the oversampling of the pathological cases as well as the exact start of the recording that had to be focused in the future in order to confirm the general nature of the features selected.

It should be helpful to include other additional information about the state of the patient such as age and possible risk factors as well as utilization of the information provided by the STAN system such as QRST ratio or width of the QRS complex.

Acknowledgements

This research project has been supported by the research program MSM 6840770012 "Transdisciplinary Research in the Area of Biomedical Engineering II" of the CTU in Prague, sponsored by the Ministry of Education, Youth and Sports of the Czech Republic

References

[1] Steer PJ. Has electronic fetal heart rate monitoring made a difference. Seminars in Fetal and Neonatal Medicine 2008; 13:2–7.

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

[3] NIH. Electronic fetal heart rate monitoring: research guidelines for interpretation. National Institute of Child Health and Human Development Research Planning Workshop. Am J Obstet Gynecol Dec 1997;177(6):1385–1390.

[4] Doria V, Papageorghiou AT, Gustafsson A, Ugwumadu A, Farrer K, Arulkumaran S. Review of the first 1502 cases of ECG-ST waveform analysis during labour in a teaching hospital. BJOG Oct 2007;114(10):1202–1207.

[5] Westerhuis M, Kwee A, van Ginkel A, Drogtrop A, Gyselaers W, Vissera G. Limitations of ST analysis in

clinical practice: three cases of intrapartum metabolic acidosis. BJOG an International Journal of Obstetrics and Gynaecology 2007;114:1194–1201.
[6] Bernardes J, Costa-Pereira A, de Campos DA, van Geijn HP, Pereira-Leite L. Evaluation of interobserver agreement of cardiotocograms. Int J Gynaecol Obstet Apr 1997; 57(1):33–37.
[7] Palomaki O, Luukkaala T, Luoto R, Tuimala R. Intrapartum cardiotocography – the dilemma of interpretational variation. J Perinat Med 2006;34(4)
[8] Amer-Wahlin I, Hellsten C, Noren H, Hagberg H, Herbst A, Kjellmer I, Lilja H, Lindoff C, Mansson M, Martensson L, Olofsson P, Sundstrom A, Marsal K. Cardiotocography only versus cardiotocography plus st analysis of fetal electrocardiogram for intrapartum fetal monitoring: a swedish randomised controlled trial. Lancet Aug 2001; 358(9281):534–538.
[9] Noren H, Amer-Wahlin I, Hagberg H, Herbst A, Kjellmer I, Marsal K, Olofsson P, Rosen KG. Fetal electrocardiography in labor and neonatal outcome: data from the swedish randomized controlled trial on intrapartum fetal monitoring. Am J Obstet Gynecol Jan 2003;188(1):183–192.
[10] Bernardes J, Moura C, de Sa JP, Leite LP. The porto systém for automated cardiotocographic signal analysis. J Perinat Med 1991;19(1-2):61–65.
[11] de Campos DA, Bernardes J, Garrido A, de Sa JM, Pereira- Leite L. Sisporto 2.0: a program for automated analysis of cardiotocograms. J Matern Fetal Med 2000;9(5):311–318.
[12] Goncalves H, Rocha AP, Ayres-de Campos D, Bernardes J. Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery. Med Bio Eng Comput 2006;44:847–855.
[13] Georgoulas G, Gavrilis D, Tsoulos IG, Stylios CD, Bernardes J, Groumpos PP Novel approach for fetal heart rate classification introducing grammatical evolution. Biomedical Signal Processing and Control 2 (2007) 69-79
[14] Georgoulas G, Stylios CD, Groumpos PP. Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using svm. IEEE Trans Biomed Eng 2006;53(5)
[15] Van Leeuwen P, Bettermann H. The status of nonlinear dynamics in the analysis of heart rate variability. Herzschrittmachertherapie und Elektrophysiologie 2000;11 (3):127–130.
[16] Pincus S. Approximate entropy (ApEn) as a complexity measure. Chaos 1995;5 (1):110–117.
[17] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol Jun 2000;278(6):H2039–H2049.
[18] Witten I.H, Frank E., "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.
[19] Kordík P.: Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution, PhD thesis, CTU Prague, September 2006
[20] N.V. Chawla, L. Hall,W. Kegelmeyer, SMOTE: Synthetic Minority Over- Sampling Technique, J. Artif. Intell. Res. 16 (2002) 3241–3357.

Address for correspondence

Vaclav Chudacek
Department of Cybernetics, Czech Technical University in Prague
Technicka 6, 166 37 Praha 6, Czech Republic
chudacv@fel.cvut.cz