INVESTIGATION AND COMPARISON OF DIFFERENT SCALE DEPENDENT FEATURES FOR FETAL HEART RATE CLASSIFICATION

George Georgoulas, Chrysostomos Stylios[1], and Peter Groumpos

Laboratory for Automation & Robotics, University of Patras, 26500, Patras, Greece
Tel. +30 2610997293 Fax +302160997309 Email: georgoul@ee.upatras.gr

[1]Dept. of Communications, Informatics and Management, TEI of Epirus,
47100 Kostakioi, Artas, Epirus, Greece, Fax +302681075839 stylios@teleinfom.teiep.gr

Abstract: This research work compares the classification results of Fetal Heart Rate signal using three different feature sets. The Discrete Wavelet Transform is employed to extract three different sets consisted of scale and time-scale dependent features from the Fetal Heart Rate signal. The three sets of features are classified using the method of Support Vector Machines (SVM) with RBF kernels. The experimental data set are 40 intrapartum recordings and the extracted three different sets of features are entered to SVM to classify the FHR. The classification results for the three data sets are compared with respect to their ability to characterize fetal condition. The best classification performance achieved was 90%.

Keywords: Discrete Wavelet Transform, Support Vector Machines, Fetal Heart Rate.

1. INTRODUCTION

Electronic Fetal Monitoring (EFM), usually named cardiotocography, has been widely used for antepartum and intrapartum fetal surveillance. EMF refers to the continuous recording and monitoring of Fetal Heart Rate (FHR) and Uterine Activity (UA), also known as cardiotocogram (CTG), which is depicted in Fig.1. In daily obstetric practice, obstetricians largely rely on information from the FHR. During the final period of labour and especially during the stressful delivery process, the risk of developing fetal hypoxia is increased. Monitoring of FHR is extensively used as an indirect screening test on fetal acid base balance (Geijn, 1996).

In every day practice, obstetricians monitor and interpret FHR. However, extensive studies on FHR reliability have shown surprisingly poor inter-observer and intra-observer agreement in tracing interpretation (Bernardes, *et al*., 1997). This inconsistency in interpretation and the increase of false positive diagnosis have raised the question of whether a reliable and reproducible interpretation of the FHR patterns can be developed. The answer to this question may be the deployment of new methodological tools, considering new indices more responsive to normal and pathological fetal conditions.

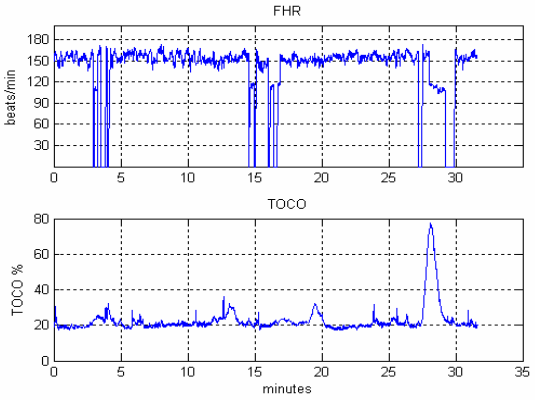


Fig. 1. Typical CTG, with the FHR in the upper part and the UA in the lower part

Therefore, some algorithmic approaches for the interpretation of the FHR that have been recently proposed are greatly helped by the technological advances in computers along with new signal processing methods (Arduini, *et al.,* 1993; Berdinas, *et al*., 2002; Bernardes, *et al*., 1991; Magenes, *et al*., 2000; Cazares, *et al*., 2001; Chung, *et al*., 1995; Dawes, *et al*., 1995; Jezewski, and Wrobel, 1993; Krause, 1990; Maeda, *et al*., 1990; Mantel, *et al.*, 1990a; Mantel, *et al.*, 1990b; Salamelekis, *et al.*, 2002; Skinner, *et al.*, 1999; Taylor, *et al.*, 2000). The employment of mathematical and algorithmic

approaches has led to a reduction of inter and intra-observer variability but there is still room for improvement regarding the identification of emergent situations which induce unacceptable stress on the fetus.

In this research work we concentrate on the final minutes of labour, seeking for appropriate indices that can trigger an alert if the fetus is on the verge of severe compromise (metabolic acidosis that may lead to cerebral palsy or even death).

We examine a novel hybrid method to discriminate fetuses suspicious of developing acidemia, based on features extracted mathematically from the FHR signal. The proposed method consists of two main stages; the first one deals with the extraction of a set of features, which are based on the coefficients produced by the application of the Discrete Wavelet Transform on the FHR signal. The second stage manipulates those coefficients through a Support Vector Machine (SVM) classifier to make a decision whether the particular FHR trace is to be characterized as "normal" or "suspicious".

In the past few years, researchers from the field of applied mathematics and signal processing have developed powerful wavelet methods for the multiscale representation of signals. This kind of representation allows the decomposition of a signal into a number of scales, each scale representing a particular "coarseness" of the signal under study (Mallat, 1998). Moreover, the localized nature of the wavelet transform makes it suitable to deal with non-stationary signals. As a result wavelet transform is particularly useful for medical signal processing and has been used for a number of biomedical applications (Unser, and Aldrubi 1996). Wavelet analysis has been used with quite a success for the analysis of the interbeat intervals of adults (Thuner *et al.*, 1998; Ivanov *et al.*, 1996). It has also been used for the analysis of FHR during the second stage of labour (Salamalekis *et al.*, 2002).

Support Vector Machines have been recently developed in the framework of statistical learning theory (Vapnik, 1995) and have proved highly successful in a number of classification studies (Burges, 1998, Veropoulos *et al*., 1999). Their experimental success in practical and difficult classification problems due to their ability to generalize well for unseen data, even when the training set is quite small, were the main reasons for selecting them for this particular problem

This paper is structured as follows, section 2 presents a background introduction concerning the basic theory of DWT and the SVMs. Section 3 presents the proposed classification procedure and how it was used in the experimental data set. Section 4 compares and discusses the classification results, and in section 5 some conclusions and future directions are drawn.

2. MATHEMATICAL BACKGROUND

Fetal Heart Rate signal is a biological signal that reflects the time varying influence of the fetus' autonomic nervous system and its components, the sympathetic and parasympathetic branch (Parer, 1997). As most physiological signals, FHR is non-stationary, where the main non-stationarities are the baseline and the acceleration/deceleration events (Jejewski *et al,* 2003). As a result, traditional Fourier analysis is not suitable for this particular type of signal, unless it is restricted to sufficiently short segments.

2.1 Wavelet Transform

A quite novel signal processing method is the wavelet analysis. The way wavelet analyses localizes signal's information in the time-frequency (time-scale would be me a more appropriate term), makes it especially suitable for the analysis of non-stationary signals as an alternative to the classical short-time Fourier transform.

The wavelet transform is a decomposition of the original signal onto a set of basis functions called wavelets. Those basis functions are obtained from a single prototype wavelet, which is referred to as the "mother wavelet" $\psi(t)$, by dilations and contractions (scalings), as well as shifts:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{1}$$

In the case of discrete wavelet transform, the dilation and translation parameters a, b are restricted only to discrete values leading to the following expression:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}}\psi\left(\frac{t-nb_0a_0^m}{a_0^m}\right) \tag{2}$$

For practical purposes the simplest and most efficient discretization comes by choosing $a_0 = 2$ and $b_0 = 1$ (dyadic grid arrangement)

$$\psi_{m,n}(t) = \frac{1}{\sqrt{2^m}}\psi\left(\frac{t-n2^m}{2^m}\right) = 2^{-m/2}\psi\left(2^{-m}t-n\right) \tag{3}$$

For this particular discreterization and with careful choice of the wavelet we can have a non-redundant decomposition of a signal on an orthonormal wavelet base. What it is even more interesting is that this kind of decomposition can be implemented using a cascade of FIR filters. For a discrete signal $x[i]$, i=0,...,M–1, the wavelet coefficients are given by:

$$T_{m,n} = 2^{-m/2}\sum_{i=0}^{M-1} x[i]\psi\left(2^{-m}i-n\right) \tag{4}$$

For an orthonormal wavelet base, the information stored in a wavelet coefficient $T_{m,n}$ is not repeated elsewhere. As it is obvious, each wavelet coefficient encompasses information not only concerning the scale but also the time (window) that "produced" this information.

Different mother wavelets give rise to different classes of wavelets, and hence the distribution of the wavelet coefficients of the decomposed signal can be quite different. However, according to (Thuner *et al.*, 1998) in their work involving the analysis of heartbeat intervals, the results obtained were similar experimenting with different types of mother wavelets. In this work we experimented with a variety of mother wavelets in order to find the best one concerning each on of the specific implementations.

2.2 Support Vector Machines

SVMs are universal feed-forward networks pioneered by Vapnik (Vapnik, 1995). Support Vector Machines are used for pattern classification and nonlinear regression. In pattern classification –and particularly in the case of a binary or dichotomization problem i.e. a problem with only 2 classes- the problem can be stated as follows:

Having a training set $S = \left\{\left(\mathbf{x}_i, y_i\right)\right\}_{i=1}^{l}$, where each point $\mathbf{x}_i$ is a p-dimensional vector, the input pattern for the *i*-th example, and $y_i \in \{-1,1\}$ is a label that specifies to which one of the 2 classes the point $\mathbf{x}_i$ belongs to, find a discriminating function that maps $\mathbf{x}_i$ to y_i. SVM classifiers try to solve this pattern classification problem by:

1) Nonlinear mapping of an input vector into a high-dimensional feature space $\boldsymbol{\varphi} : \Re^p \to \Re^m, (m > p)$
2) Construction of an "optimal" hyperplane for separating the features in the high-dimensional feature space (Haykin, 1999).

This "optimal" hyperplane:

$$f\left(\mathbf{x}\right) = \left\langle \mathbf{w} \cdot \boldsymbol{\varphi}\left(\mathbf{x}_i\right)\right\rangle + b \tag{5}$$

can be constructed by solving the following quadratic optimization problem:

$$\text{Minimize } \frac{1}{2}\mathbf{w}^T \cdot \mathbf{w} + C\sum_{i=1}^{l}\xi_i \tag{6}$$

$$\text{Subject to } y_i\left(\mathbf{w}\left\langle \cdot \boldsymbol{\varphi}\left(\mathbf{x}_i\right)\right\rangle + b\right) \geq 1 - \xi_i \tag{7}$$
$$\xi_i \geq 0\,,\ i = 1, 2, ..., l$$

The dual problem, which is in fact the one to be solved, is:

$$\text{Maximize} \tag{8}$$
$$\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} a_i a_j y_i y_j \left\langle \boldsymbol{\varphi}\left(\mathbf{x}_i\right) \cdot \boldsymbol{\varphi}\left(\mathbf{x}_j\right)\right\rangle$$

$$\text{Subject to } \sum_{i=1}^{N} y_i \alpha_i = 0 \tag{9}$$
$$\text{where } 0 \leq \alpha_i \leq C \ \ , \ i = 1, 2, ..., l$$

where parameter C is determined by the user. Larger *C* corresponds to assigning a higher penalty to errors (Burges, 1998). The discriminating function is finally given by:

$$f\left(\mathbf{x}\right) = sign\left(\sum_{i=1}^{l} y_i \alpha_i \left\langle \boldsymbol{\varphi}\left(\mathbf{x}_i\right) \cdot \boldsymbol{\varphi}\left(\mathbf{x}\right)\right\rangle + b\right) \tag{10}$$

The points for which a_i >0 are called Support Vectors. They are the most difficult patterns to classify and usually are a small portion of the training set. If the nonlinear mapping function is chosen properly, the inner product in the feature space can be written in the following form:

$$\left\langle \boldsymbol{\varphi}\left(\mathbf{x}_i\right) \cdot \boldsymbol{\varphi}\left(\mathbf{x}_j\right)\right\rangle = K(\mathbf{x}_i, \mathbf{x}_j) \tag{11}$$

where K is called the inner-product kernel. By using an appropriate symmetric positive semi-definite kernel, it is guaranteed that it corresponds to an inner-product in a Hilbert space. Therefore, one does not even have to know what the actual mapping is (Burges, 1998).

3. PROPOSED PROCEDURE AND EXPERIMENTS

The data set consisted of 40 FHR signals and it was divided into 2 subsets. Acidemia was defined for this study based on the umbilical artery pH<7.1. Therefore, in the first subset we included those signals that belonged to fetuses with umbilical artery blood pH less than 7.1 and in the second subset, those that belonged to fetuses with umbilical artery blood pH more than 7.2. All FHR records had been acquired during the final stage of the labour and, in fact, as close as possible to the delivery. This means that the data sets were time-biased free and a direct association could be made between the segment of the signal used and the fetal outcome. The recordings had durations ranging from 20 minutes to 1 hour.

3.1 Artifact removal-Data segmentation

FHR signal is a very noisy signal with a lot of spiky artifacts and even periods of missing data due to the movement of the baby and the stress induced during the labour. In order to eliminate this kind of "noise"

we implemented a noise detection and elimination algorithm (Bernardes, *et al.*, 1991).

In the data set, we focused on the final minutes of the recordings – those that are closer to delivery. Other recent experiments have shown that the final minutes before the delivery are those that affect more the value of the umbilical artery pH (Georgoulas, *et al.*, 2005). Therefore, we used only the 10 and 5 minutes of each recording closest to the delivery. It must be mentioned that in some of the recordings the final 1-2 minutes had to be excluded (before the artifact removal stage had taken place) because the FHR signal was totally obscured by noise.

3.2 Feature extraction

In this work, we examined and compared 3 different feature sets. All features are extracted using the coefficients of the DWT of the FHR signal. Therefore, we performed 3 different experiments. In the first one we used the entropy of the coefficients at each scale. In the second one we used isolated values of the wavelet coefficients. And in the third one we used a statistical measure of the coefficients contained in a window that spans only part of the available coefficients at each scale.

First feature set. For each FHR signal and for the corresponding time segment (10 and 5 minutes) we carried out discrete wavelet transform up to scale 6. For each scale, we calculated the corresponding entropy S of the (discrete) distribution p_i of normalized energies of wavelet coefficients (i.e. squared magnitudes) (Shannon entropy measure)

$$S(p) = -\sum_i p_i \log(p_i) \qquad (12)$$

Therefore, for each signal we extracted 6 values for each scale of decomposition. We applied this to both the 10 and 5-minute segments (the 5 minute segments correspond to the second half of the 10 minute segment).

Second feature set. For each FHR signal and for the corresponding time segment (10 and 5 minutes) we carried out discrete wavelet transform up to scale 5. For each scale we selected the wavelet coefficient with the maximum absolute value. This value, together with the "time" index *n,* were the features that we used to characterize each scale. Therefore, for each FHR signal we extracted 10 features. We did this both for 10 and 5-minute segments.

Third set. For each FHR signal and for the corresponding time segment (10 and 5 minutes) we carried out discrete wavelet transform up to scale 5. For each scale we used a sliding window and for the wavelet coefficients inside that window we calculated the standard deviation seeking for the location of the window that maximizes that quantity (Fig. 2). The location of the window for which we have the maximization of the calculated standard deviation, along with the time at which this happens, (the centre of the window) were the features that we used to characterize each scale. Therefore, for each signal we extracted 10 features. We did this both for 10 and 5-minute segments.

3.4 Classification

Depending on how the inner-product kernel is generated, different learning machines can be constructed with quite different non-linear decision surfaces. In this work we used only RBF learning machines, where the kernel function is:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x} - \mathbf{x}_i\|^2\right) \qquad (13)$$

and the width σ^2 is specified a priori by the user and is common for all the kernels.

As mentioned, the parameter C (Eq. 7 and eq. 9) is a user defined variable. We determined both parameters C and σ using an experimental training validation procedure, testing various configurations of the learning machines.

Due to the small number of labeled data, in order to test the performance of our classification scheme, we used multifold cross-validation (Haykin, 1999). We divided the 40 cases into 4 (non-overlapping) subsets, each one with 5 examples from the "normal" and 5 from the "suspicious-abnormal" group. The SVM classifier was trained on all subsets except for one, and the validation performance was assessed on the subset left out. We repeated this procedure 5 times, each time using a different subset for testing.

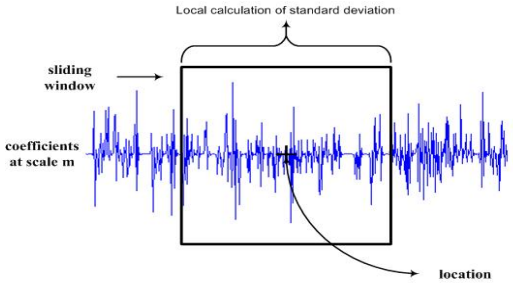


Fig. 2. Feature extraction using a sliding window for finding window for the greatest standard deviation

4. EXPERIMENTAL RESULTS

First feature set. Using the "scale-dependent" entropy and for a time duration of 10 minutes, we achieved a maximum classification rate of 82.5 % (for more than one mother wavelet). By reducing the time duration to 5 minutes we achieved a classification rate of 90 % (90% for the "normal" group and 90% for the "abnormal" group). This is

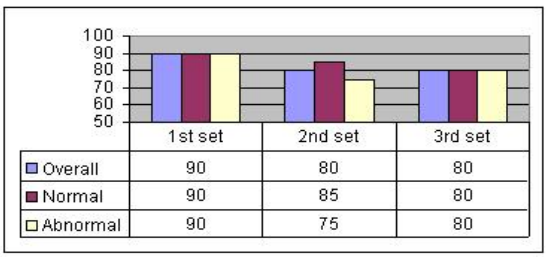
something that we expected since it is in accordance with some other results (Georgoulas *et al.,* 2005) where for the same data set we used as a global statistic the scale dependant standard deviation, and achieved the same classification performance. The only difference was that in the experiments involving the scale-dependent entropy we used biorthogonal wavelets and the time duration was 5 minutes, whereas in the case of the scale-dependent standard deviation, we achieved this rate for symmlets and the time duration had to be restricted to 3 minutes.

Second feature set. Using the time-scale features for time duration of 10 minutes, we also achieved a classification rate of 82.5 % (both for symmlets and coiflets (Daubechies, 1994)). However, when the time duration was reduced to 5 minutes, the maximum classification rate dropped to 80% (using coiflets).

Third feature set. The results of this third set were worse than the previous two. For time duration of 10 minutes the maximum classification rate achieved was 72.5, and for 5 minutes segments the classification performance was slightly improved, and equal to 77.5 overall classification rate.

	1st set	2nd set	3rd set
Overall	82.5	82.5	72.5
Normal	80	80	70
Abnormal	85	85	75

a. 10 minute segments



	1st set	2nd set	3rd set
Overall	90	80	80
Normal	90	85	80
Abnormal	90	75	80

b. 5 minute segments

Fig. 3. Best results for the 3 feature sets, for time duration equal to a) 10 minutes and b) 5 minutes.

5. CONCUSIONS

Among the 3 different feature sets, the first one that used of the wavelet coefficients in each scale performed very well, compared both to our previous findings (Georgoulas *et al.*, 2004a; Georgoulas *et al.*, 2004b; Georgoulas et al., 2005) and to other groups findings (Salamelekis et al., 2002). It seems that 90% is close to the upper limit of the classification performance concerning this particular data set. We suspect that there must be some outliers in this set. On the other hand, the other 2 feature sets failed to meet our expectations – the results were worse compared to those reported in the first set and compared to other work (Georgoulas, *et al.*, 2005). This is probably an indicator that these features are not the best for use in this particular task. The very restricted number of cases makes it quite difficult to use more features spanning more effectively the time-scale plane.

It is mentioned that the indices used to discriminate normal fetuses from those which may be at risk, is by no means a gold standard. It is known that babies with severe acidosis (pH 7.0 or less) will subsequently be normal in a percentage of 90%. However, this is considered immediate outcomes that one would prefer to avoid (Parer, 1997).

Using two distinct thresholds instead of one, which is the usual practice, we tried to make sure that the two different sets would actually include representative cases for normal in the one set and risky to develop acidemia in the other set. This approach naturally prompted the use of a "binary" classifier. In future work we are planning to try a regression approach employing cases with umbilical artery pH values also spanning the range (7.1, 7.2) that have not been used here. However these primarily results indicate that algorithmic procedures can be found to discriminate normal from acidemic outcome, something that was questionable in early 90s (Dawes, 1994).

In conclusion, we must mention that at the beginning of the experiments we expected the time-scale parameters to perform better than the scale dependent global measure of entropy. Thus, it is obvious that the time-scale feature selection needs to be refined and this is another research direction that we will focus on in future work.

ACKNOWLEDGMENTS

Authors thanks Prof. João Bernardes, Dep. Gynecology & Obstetrics, Porto Faculty of Medicine, Porto, Portugal for providing the FHR data collected within the Research Project POSI/CPS/40153/2001, from Fundação para a Ciência e Tecnologia, Portugal.

Funding for this research was provided by the program EPEAEK II: Archimedes -Research Support in TEI of the Ministry of National Education and Religious Affairs of Greece.

REFERENCES

Arduini, D., G. Rizzo, G. Piana, A. Bonalumi, P. Brambilla and C. Romanini (1993). Computerized Analysis of Fetal Heart Rate: Description of the System (2CTG). *J. Mater Fetal Invest.*, **3**, 159-163.

Berdinas, B. G., A. A. Betanzos and O. F. Romero (2002). Intelligent analysis and pattern recognition in CTG signals using a tightly coupled hybrid system. *Artificial Intelligence*, **136**, 1-27.

Bernardes, J., C. Moura, J.P.M. de Sa and L.P. Leite (1991). The Porto system for automated cardiotocographic signal analysis. *J. Perinat. Med*., **19**, 61-65.

Bernardes, J., A. C. Pereira D. A. de Campos, H. P. Van Geijn, L.P. Leite (1997). Evaluation of interobserver agreement of cardiotocograms. *Int J. Gynecol. Obst.,* **57**(1),33-37.

Bernardes, J., D. A. de Campos, A. C. Pereira, L.P. Leite and A. Garrido (1998). Objective computerized fetal heart rate analysis. *Int. J. Gynecol. Obst*., **62**, 141-147.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data mining and Knowledge Dicovery*, **2**, 121-167.

Chung, T. K. H., M. P. Mohajer, X. J. Yang, A. M. Z. Chang and D.S. Sahota (1995). The prediction of fetal acidosis at birth by computerized analysis of intrapartum cardiotocography. *Br. J. Obstet. Gynaecol.,* **102**, 454-460.

Daubechies I. (1994). Ten Lectures on wavelets. SIAM, Philadelphia, PA.

Dawes, G. S (1994). Computerized Fetal Heart Rate analysis. In A critical appraisal of fetal surveillance E.P. van Geijn and F.J.A. Copray editors, Elsevier Science. 311-314.

Dawes, G. S., M. Moulden, and C. W. Redman (1995). Computerized analysis of antepartum fetal heart rate. *Amer. J. Obstet. Gynecol.,* **173**(4), 1353-1354.

Geijn, H. P. V. (1996). Developments in CTG analysis. *Bailliers Clin. Obstet. Gynaecol.,* **10**(2), 185-209.

Georgoulas G., G. Nokas, C. Stylios and P. P. Groumpos (2004). Classification of Fetal Heart Rate during labour using Hidden Markov Models. *Proc. of IJCNN 2004 Intern. Joint Conference on Neural Networks & Fuzzy Systems,* Budapest 25-28 July 2004 (CD-ROM).

Georgoulas G., C. Stylios, J. Bernades and P. Groumpos (2004). Classification of CTG using Support Vector Machines. *Proc of 10th IFAC Symposium on Large Scale Systems: Theory & Applications, July 26-28, 2004, Osaka, Japan.*

Georgoulas G., C. Stylios, and P. Groumpos (2005). Classification of Fetal Heart Rate using scale dependent features and Support Vector Machines. *Proc of 16th IFAC World Congress, Prague.*

Haykin,S. (1999). *Neural Networks: A Comprehensive Foundation*. 2nd ed. Englewood Cliffs, NJ, Prentice Hall.

Ivanov P. C., M.G. Rosenblum, C.K. Peng, C.K. Peng, J Mietus, S. Havlin, H.E. Stanley, and A.L. Goldberger (1996). Scaling behaviour of heartbeat intervals obtained by wavelet-based time series analysis. *Nature,* **383**, 323-327.

Jezewski J., K. Horoba, A. Gacek, J. Wrobel A. Matonia and T. Kupka, (2003) Analysis of nonstationarities in fetal heart rate signal, Inconsistency measures of baselines using acceleration/deceleration patterns. *in Proc. 7th ISSPA,*34-38.

Krause, W. (1990). A computer aided monitoring system for supervision of labour. In: *Computers in perinatal medicine* (K. Maeda (Ed)), 103-111. Elsevier Science, Amsterdam, The Netherlands.

Maeda, K. (1990). Computerized analysis of cardiotocograms and fetal movements. *Bailliers Clin. Obstet. Gynaecol.,* **4**(4), 1797-813.

Magenes, G., M.G Signorini and D. Arduini (2000). Classification of cardiotocographic records by neural networks. *D.* Neural Networks. *Proc. of IEEE-INNS-ENNS Int. Joint Conference,* **3**, 637-641.

Mallat S (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.*, **11** (7), 674-793.

Mantel, R., H.P. van Geijn, F.J.M Caron, J. M. Swartjes, E. E. van Woerden and H. W. Jongsma (1990a). Computer analysis of antepartum fetal heart rate: Baseline determination. *Int. J. Biomed. Comput.,* **25**(2), 261-272.

Mantel, R., H.P. van Geijn, F.J.M Caron, J. M. Swartjes, E. E. van Woerden and H. W. Jongsma (1990b). Computer analysis of antepartum fetal heart rate: Detection of accelerations and decelerations. *Int. J. Biomed. Comput.*, **25**(2), 273-286.

Parer, J. T. (1997). *Handbook of fetal heart rate monitoring*. Philadelphia, Pennsylvania: W. B. Saunders Company.

Salamalekis, E., P. Thomopoulos, D. Giannaris, I. Salloum, G. Vasios, A. Prentza and D. Koutsouris (2002). Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks. *Br. J. Obstet. Gynaeco.,* **109**(10), 1137-1142.

Skinner, J. F., J. M. Garibaldi and E. C. Ifeachor, (1999). A Fuzzy System for Fetal Heart Rate Assessment. I*n Proc of the 6th Fuzzy Days Conference*, Dortmund, Germany, 20-29.

Taylor, G. M., G. J. Mires, E. W. Abel, S. Tsantis, T. Farrell, P. F. W. Chien and Y. Liu (2000). The development and validation of an algorithm for real time computerized fetal heart rate monitoring in labour. *Br. J. Obstet. Gynaecol.,* **107**, 1130-1137.

Thuner S., M. C Feurstein., and M.C Teich (1998) Multiresolution wavelet analysis of heatbeat intervals discriminates healthy patients from those with cardiac pathology. *Physical Review Letters*, **80**(7), 1544-1547.

Unser, M., and A. Aldrubi (1996). A Review of Wavelets in Biomedical Applications. *Proc. of IEEE*, **84**(4), 626-638.

Vapnik, V (1995). The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Veropoulos, K., Cristianini, C. Campbell (1999). The Application of Support Vector Machines to Medical Decision Support: A Case Study. Proceedings of the ECCAI Advanced Course in Artificial Intelligence.