

RESEARCH

Open Access

Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine

Niranjana Krupa^{1*}, Mohd Ali MA¹, Edmond Zahedi², Shuhaila Ahmed³, Fauziah M Hassan⁴

* Correspondence: niranjana.krupa@gmail.com

¹Department of Electrical Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi, Malaysia.

Full list of author information is available at the end of the article

Abstract

Background: Cardiotocography (CTG) is the most widely used tool for fetal surveillance. The visual analysis of fetal heart rate (FHR) traces largely depends on the expertise and experience of the clinician involved. Several approaches have been proposed for the effective interpretation of FHR. In this paper, a new approach for FHR feature extraction based on empirical mode decomposition (EMD) is proposed, which was used along with support vector machine (SVM) for the classification of FHR recordings as 'normal' or 'at risk'.

Methods: The FHR were recorded from 15 subjects at a sampling rate of 4 Hz and a dataset consisting of 90 randomly selected records of 20 minutes duration was formed from these. All records were labelled as 'normal' or 'at risk' by two experienced obstetricians. A training set was formed by 60 records, the remaining 30 left as the testing set. The standard deviations of the EMD components are input as features to a support vector machine (SVM) to classify FHR samples.

Results: For the training set, a five-fold cross validation test resulted in an accuracy of 86% whereas the overall geometric mean of sensitivity and specificity was 94.8%. The Kappa value for the training set was .923. Application of the proposed method to the testing set (30 records) resulted in a geometric mean of 81.5%. The Kappa value for the testing set was .684.

Conclusions: Based on the overall performance of the system it can be stated that the proposed methodology is a promising new approach for the feature extraction and classification of FHR signals.

Background

Cardiotocograph (CTG) is a graphical representation of fetal heart rate (FHR) and uterine activity (UA), also termed as electronic fetal monitoring, and has been an indispensable part of antepartum and intrapartum fetal surveillance [1] for four decades. A typical CTG is depicted in Figure 1. Heart rate (HR) contains reliable information about the synergic activity of the autonomic nervous system (ANS) that regulates the heart beat dynamics [2]. Parameters from the HR signal provide interesting hints about the generation of disease conditions and hence can be used to differentiate pathological states [3]. The analysis of FHR signal used in monitoring the fetal well being is a powerful tool in establishing the development of the nervous system of the fetus during the last period of pregnancy, starting from the 25th week of gestation [4]. However,

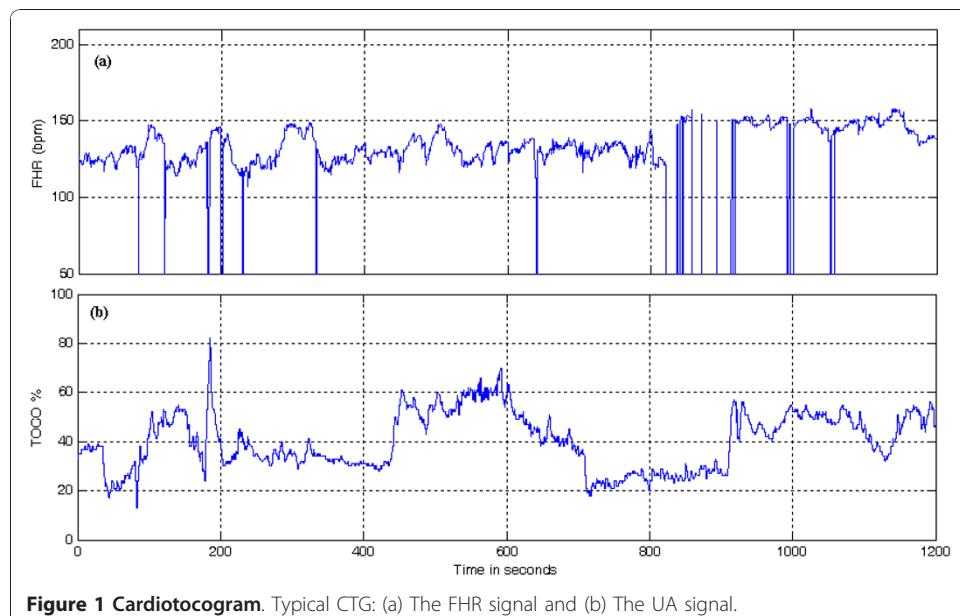


Figure 1 Cardiotocogram. Typical CTG: (a) The FHR signal and (b) The UA signal.

the current clinical practice of visual interpretation of CTG shows a high degree of inter-observer and intra-observer variability [5] due to its large dependency on the expertise and experience of the clinician(s) involved [6]. Advances in signal processing and pattern recognition techniques and skepticism over inconsistency in FHR interpretation paved the way for computerized methods [7]. The computerization of non-stress-test (NST) involves two principal problems namely, feature extraction and classification, and subsequent interpretation [6]. It is important to mention that the interpretation phase requires a contextual analysis of all the physiological, pathological and clinical aspects needed in assessing the well being of the fetus.

Diverse approaches have been investigated using conventional and artificial intelligence techniques for feature extraction [3,8-13] and also to come out with diagnostic systems. Computerized CTG analysis systems were developed [14-16] adhering to the guidelines issued by the International Federation of Obstetrics and Gynecology (FIGO) [17]. A non-linear back propagation artificial neural network (ANN) was introduced for the interpretation of NST records [18] and around the same time a software program to predict fetal acidosis at birth was also developed [19]. ANN was used to distinguish between the normal and pathological fetal conditions [20,21]. An expert system [22] capable of handling uncertainties in FHR interpretation was introduced, which was later transformed to a fuzzy system [23]. Yet again, an expert system called NST-EXPERT [24,25] was developed and was later named computer aided fetal evaluator (CAFÉ) system [6]. A method using wavelet analysis and self organizing map was developed to diagnose fetal hypoxia based on the scale-dependent features extracted from the FHR [26]. Another method based on wavelet transform and cumulative holder exponent was proposed [27] for real-time fetal monitoring during labor. In [1] scale-dependent features extracted from wavelet components of the FHR signal were categorized using a support vector machine (SVM). SVM classifier was employed yet again to predict the risk of metabolic acidosis in newborns [28], to detect fetal distress [29,30] and to discriminate healthy fetuses from the ones with hypoxia [31]. Recently,

a new approach based on FHR variability analysis involving Lempel Ziv complexity index and multiscale entropy was proposed for the early identification of intrauterine growth-restricted (IUGR) fetuses [32].

The above discussion shows that although some approaches have shown promising results, none has been widely accepted and there is still room for improvement [1,28] to reach the ultimate goal of a completely reliable method to assess fetal well being with minimal intervention from obstetricians. In an effort to achieve this goal we propose an innovative approach in this paper for FHR classification combining empirical mode decomposition (EMD) and SVM techniques.

EMD has been used in several biomedical applications (introduced for the study of ocean waves in 1998 [33]) such as artifact reduction in electrogastrogram and to extract the lower esophageal sphincter pressure in the gastro-esophageal reflux disease [34,35]. There has been a noticeable contribution from EMD based methods in processing electrocardiogram (ECG) signals [36,37]. It has also been used in the analysis of heart rate variability (HRV) [38] and high frequency FHR variability [39]. In addition to the above applications EMD has been employed for CTG signal enhancement [40].

SVM is a powerful supervised machine learning tool introduced recently in the framework of statistical learning theory [41]. It is used in a number of applications for both pattern classification and non-linear regression [42-44]. It has the ability to generalize well on unknown data without requiring the domain knowledge, and even when the sample size is small [45] which make it an attractive solution in difficult pattern recognition problems. Here, statistical features are extracted from FHR signals using EMD and later classified using the SVM classifier.

Methods

A brief explanation on the two important techniques (EMD and SVM) employed in this work is provided in the beginning of this section. Later a detailed description of the proposed methodology for the extraction of statistical features and classification of FHR signals as ‘normal’ and ‘at risk’ (Figure 2) is given. Finally, the four important stages in this work, namely, data acquisition, preprocessing, feature extraction and classification, are defined.

Empirical Mode Decomposition

EMD, proposed by Huang *et al* [33], is a method to decompose non-linear and non-stationary time series into several monotonic components termed as intrinsic mode functions (IMF) of different time scales. The most appealing nature of EMD is its dependency on the data-driven mechanism which does not require a priori known basis unlike Wavelet and Fourier transform.

The EMD method identifies all the local maxima and minima for a given input signal $x(t)$ which are connected by spline curves to form the upper and the lower envelopes,

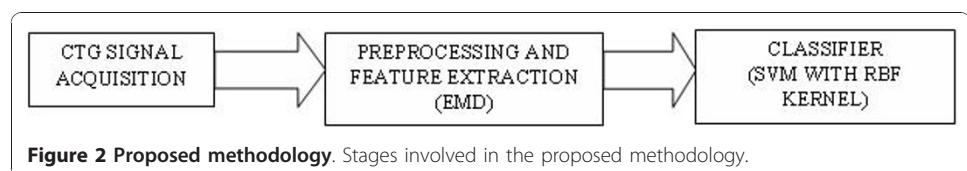


Figure 2 Proposed methodology. Stages involved in the proposed methodology.

$e_{\text{up}}(t)$ and $e_{\text{low}}(t)$, respectively. The mean of the two envelopes is calculated as $m(t) = [e_{\text{up}}(t) + e_{\text{low}}(t)]/2$ and is subtracted from the signal using $q(t) = x(t) - m(t)$.

An IMF $c_i(t)$ is obtained if $q(t)$ satisfies the two conditions of IMF, these are, the number of extrema and number of zero crossings is either equal or differs at most by one, and the envelopes defined by the local maxima and minima are symmetric with respect to zero mean. This procedure is called as the sifting process. Then $x(t)$ is replaced with the residual $r(t) = x(t) - q(t)$. If $q(t)$ is not an IMF, $x(t)$ is replaced with $q(t)$. The above process is repeated until the residual satisfies the stopping criterion called sum of difference, S_D shown in equation (1), which is generally set between 0.2 and 0.3.

$$S_D = \sum_{t=0}^T \frac{|q_{k-1}(t) - q_k(t)|^2}{q_{k-1}^2(t)} \quad (1)$$

At the end of this process the signal $x(t)$ would result in N IMFs and a residue signal as in equation (2).

$$x(t) = \sum_{n=1}^N c_n(t) + r_N(t) \quad (2)$$

Where n represents the order of IMFs, $n = 1$ to N and r_N denotes the final residue which can also be considered as an IMF, shown in equation (3).

$$x_d(t) = \sum_{i=1}^{N_f} c_i(t) \quad (3)$$

Here, N_f is the total number of IMFs including the residue. The signal $x(t)$ is decomposed such that the lower-order components represent fast oscillation modes and higher-order components represent slow oscillation modes. A detailed explanation of the method is provided in [40].

Support Vector Machine

SVM is a supervised learning tool that can be used for pattern classification. The main goal of SVM is to construct an optimal hyperplane as the decision surface in such a way that the margin of separation between the closest data points belonging to different classes is maximized. SVM is based on the principle of structural risk minimization method [46]. In a binary classification problem which is of interest in this work, each one of the set of points belongs to either one of the two classes.

Consider a training set $\{(x_i, d_i)\}_{i=1}^l$ where x_i is the input pattern for the i^{th} sample and $d_i \in \{-1, +1\}$ is the corresponding desired output and l is the number of observations. If the input patterns belonging to two different classes are linearly separable then there exists a hyperplane that maximizes the margin of separation. For the optimal hyperplane the Euclidian norm of the weight vector w is minimum and at the same time satisfies the constraints in equation (4).

$$d_i(w^T x_i + b) \geq 1 \quad \forall_i \quad (4)$$

This is a constrained quadratic optimization problem that may be solved using the method of Lagrange multipliers. Pattern classification problems in real life are not linearly separable. Here, SVMs depend on two mathematical operations: non-linear mapping of an input vector into a high-dimensional feature space and construction of an optimal hyperplane for separating the features.

Non-linear mapping is performed in accordance with the Cover's theorem [46] on the separability of patterns. Non-linearly separable patterns in the input space when transformed to a high dimensional feature space, they can be linearly separable with high probability. Therefore for each input pattern vector x_i in the m_0 dimensional input space we define a vector consisting of a set of real-valued functions $\{\psi_i(x) \mid i = 1, 2, \dots, m_1\}$, as shown by $\psi(x) = [\psi_1(x), \psi_2(x), \dots, \psi_{m_1}]^T$ that map the m_0 dimensional input points to m_1 dimensional new feature space. An optimal hyperplane in the feature space is found as in equation (5).

$$\begin{aligned} & \text{Minimizing } \frac{1}{2} w^T w + C \sum_{i=1}^n \zeta_i \\ & \text{subject to: } d_i(w^T \psi(x_i) + b) \geq 1 - \zeta_i \\ & \text{and } \zeta_i \geq 0 \quad \forall_i \end{aligned} \tag{5}$$

ζ_i are called slack variables, a set of non-negative scalar variables; they measure the deviation of a data point from the ideal condition of pattern separability. Parameter C is a user specified positive value that controls the trade-off between maximizing the margin and minimizing the error. $\psi(x_i)$ is the non-linear mapping of input patterns from input space to feature space. The optimal discriminating function is given by equation (6).

$$f(x) = \text{sign} \left[\sum_{i=1}^n d_i \alpha_i (\psi^T(x_i) \psi(x)) + b \right] \tag{6}$$

The coefficients α_i are derived from the maximization of dual Lagrangian as in equation (7).

$$\begin{aligned} Q(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j d_i d_j (\psi^T(x_i) \psi(x_j)) \\ & \text{subject to: } \sum_{i=1}^n \alpha_i d_i = 0 \\ & \text{and } 0 \leq \alpha_i \leq C \quad \forall_i \end{aligned} \tag{7}$$

The points for which $\alpha_i > 0$, are called support vectors. The term $\psi^T(x_i) \psi(x_j)$ represents the inner product of two vectors in the feature space. We may introduce the inner-product kernel denoted by $K(x_i, x_j)$, written as shown in equation (8).

$$K(x_i, x_j) = \psi^T(x_i) \psi(x_j) \tag{8}$$

A kernel function is a function in the input space and hence, we may use the inner-product kernel $K(x_i, x_j)$ to construct the optimal hyperplane without having to perform explicitly the non-linear mapping [42,46].

For classification problems dealing with medical data where the numbers of data in different classes are unbalanced, some researchers [1,28,47] have proposed the use of different penalty parameters in the SVM formulation as in equation (9).

Minimizing,

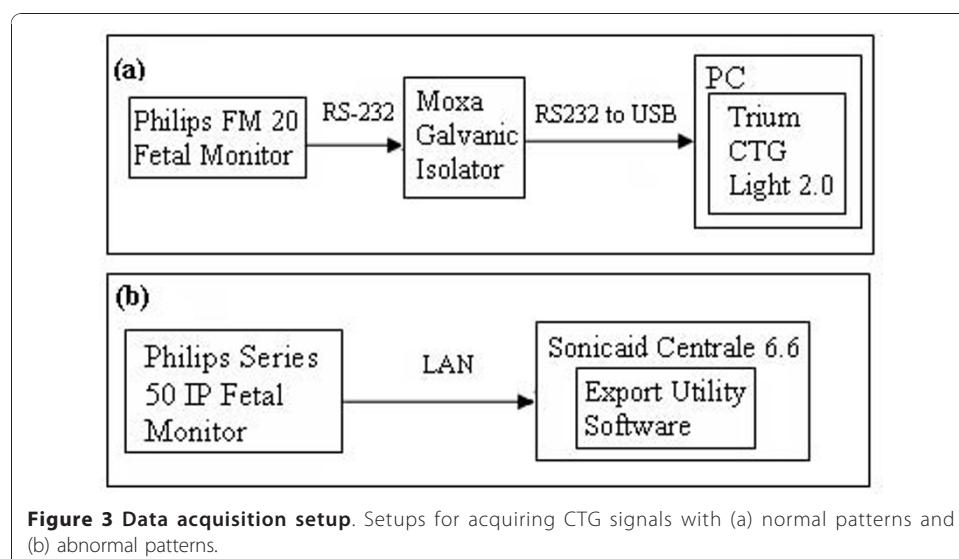
$$\frac{1}{2} w^T w + C^- \sum_{i:y_i=-1}^n \zeta_i + C^+ \sum_{i:y_i=+1}^n \zeta_i \quad (9)$$

subject to: $d_i(w^T \psi(x_i) + b) \geq 1 - \zeta_i$
and $\zeta_i \geq 0 \quad \forall i$

C^+ and C^- are the penalty parameters used to penalize more heavily the undesired type of error, and the errors related to the class with the smallest population [47].

Data Acquisition

CTG is a routine non-invasive fetal monitoring tool based on ultrasound Doppler combined with an external pressure transducer to record uterine activity. In this technique a transducer placed on the mother's abdomen transmits an ultrasound beam towards the fetal heart. The FHR is derived from the Doppler shifted echoes created by contractions of the fetal heart. An autocorrelation method is used to compare successive heart signals and test for similarity. CTG signals used in this research work were recorded at Universiti Kebangsaan Malaysia Medical Center (UKMMC). Data acquisition was carried out with the approval of UKMMC's ethical committee and after obtaining informed consent from all subjects. All 15 subjects were with singleton pregnancy and gestation age ranging from 34 weeks to 40 weeks. Since, the proposed study required both normal and abnormal patterns of FHR we had two different setups. The CTG signals with normal patterns were recorded from the day care clinic using an antepartum fetal monitor (Philips FM 20) and a software (Trium CTG Light 2.0 from Trium Analysis Online GmbH) as in Figure 3a. A galvanic isolator was used between the fetal monitor and the computer for safety purpose. CTG signals with abnormal patterns were obtained from the archived data on a Huntleigh server (Sonicaid™



Centrale, Labour Management System) at UKMMC, using the Export Utility Software as shown in Figure 3b. These data were recorded with the help of a fetal monitor (Philips Series 50IP). All the signals acquired using the above mentioned commercially available softwares had a sampling frequency of 4 Hz.

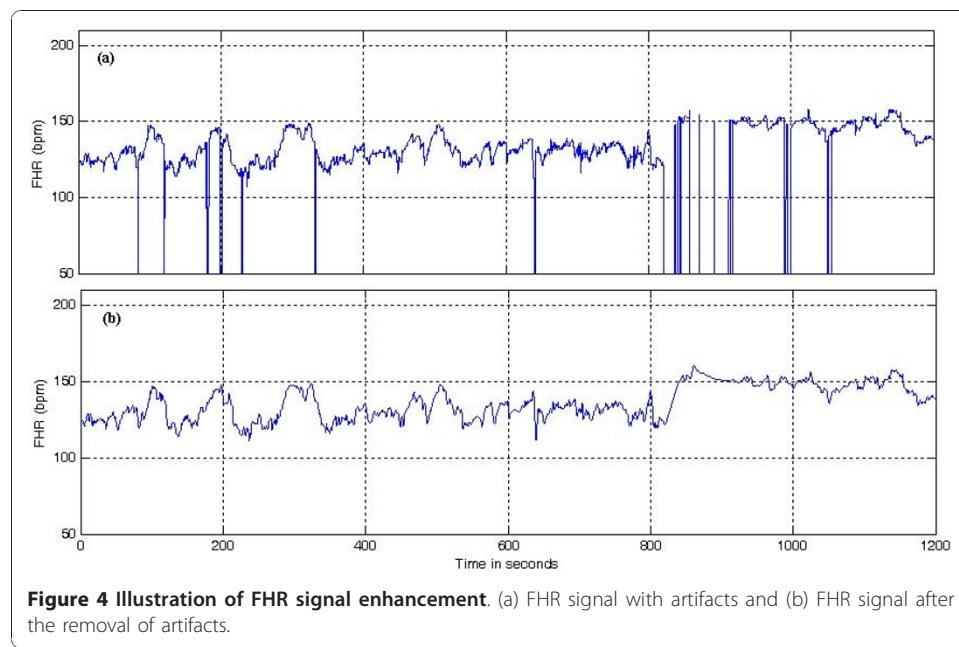
Pre-processing and Feature Extraction using EMD

Preprocessing

Recorded FHR signals may possess missing beats and spiky artifacts (Figure 4a) due to the displacement of the transducer because of maternal or fetal movement and stress induced after the onset of labour. These artifacts are generally present and difficult to eliminate from the source. Missing beats in FHR can be about 20% - 40% of the data, especially during the final stages of labour [28,40]. For this reason, the quality of the FHR signal is estimated based on the number of missing beats and a poor quality signal is not subjected to further analysis [40]. In this work, missing beats are removed using a recursive algorithm and high frequency noises are suppressed using a method based on EMD explained in [40]. Once missing data segments are eliminated, the FHR signal is decomposed using EMD into several monotonic components (IMFs) as explained earlier. It is reported that the lower order IMFs which may contain noise have zero mean [36,40,35]. Hence, a statistical t-test is used to determine the high frequency noise components which are then separated from the signal in the EMD domain (equation (10)).

$$\begin{aligned} H_0 : \text{mean}(c_{P_M S}(t)) &= 0 \\ H_1 : \text{mean}(c_{P_M S}(t)) &\neq 0 \end{aligned} \quad (10)$$

Where, $c_{P_M S}(t)$ is the M th-order partial sum of the IMFs.



This test is applied first on the lowest order IMF and later on the partial sum $c_{P_M S}(t)$ for $M = 1, 2, \dots$ as shown in equation (11), until a partial sum $c_{P_i S}(t)$ is obtained, where the mean significantly deviates from zero.

$$c_{P_M S}(t) = \sum_{i=1}^M c_i(t) \quad (11)$$

The above procedure indicates the number of IMFs that can be considered as noise, P_f (noise order). Since this technique might result in over smoothing of the FHR signal, the noise order is estimated using equation (12) [40].

$$P_f = \min(P_t, 3) \quad (12)$$

Then, the first P_f components are eliminated from the set of IMFs for a given FHR signal. A denoised FHR signal (Figure 4b) can be obtained by applying the partial construction method on the remaining components.

Feature Extraction

For each of the remaining components (Figure 5) a standard deviation is estimated (equation (13)).

$$SD(C_i) = \left[\frac{1}{N-1} \sum_{j=1}^N [C_{i,j} - M(C_i)]^2 \right]^{\frac{1}{2}} \quad (13)$$

for $i = P_f + 1$ to N

These standard deviations (noise order, P_f+1 onwards) are considered as statistical features of the FHR signal and are used as inputs to the SVM classifier. A similar approach has been employed in extracting statistical features from the wavelet coefficients, obtained from discrete wavelet transform (DWT) [1,48,49].

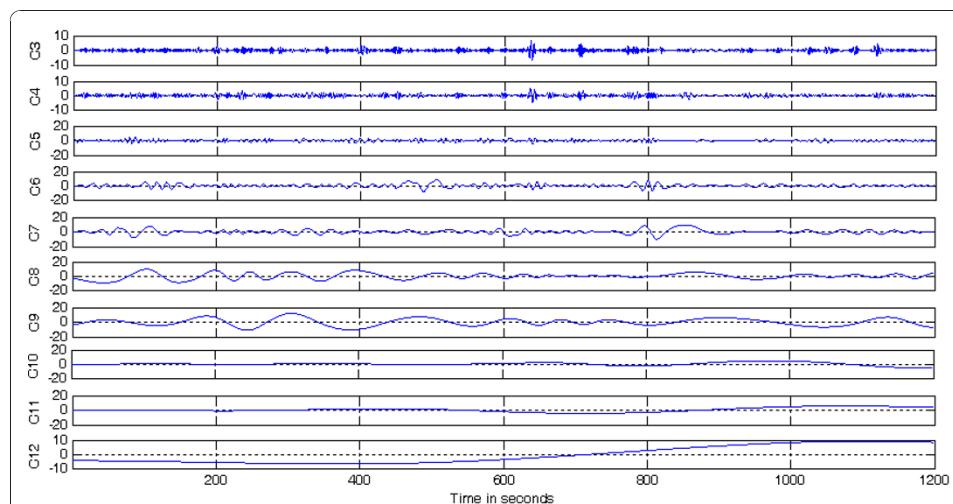


Figure 5 Decomposed components after signal enhancement. Remaining components (C3-C12) of a decomposed FHR signal used for statistical feature extraction, after the removal of noisy lower order IMFs (C1 and C2) based on P_f .

Classification using SVM

SVM Classifier

The classification stage follows the feature extraction stage. The main objective of this stage is to classify the FHR signals as normal or at risk represented as '+1' and '-1' respectively. The entire process involves a training set and a testing set of data instances. A training set consists of features extracted from the decomposed FHR signal, also called the attributes, and the labels '+1' (normal) or '-1' (at risk) which are the target outputs. SVM should produce a model to classify the data instances in the testing set which consists of only features. The classification depends on the inner-product kernel used that produces different learning machines and hence different decision boundaries [46]. In this work, based on the statistical features extracted from FHR signals the radial basis function (RBF) kernel was used (equation (14)) to obtain reliable results. Parameters C and γ are specified by the user.

$$k(x_i, x_j) = \gamma \exp\left(\|x_i - x_j\|^2\right) \quad (14)$$

The following procedure was employed during the classification stage:

(a) Scaling the training and testing data sets.

(b) General grid search method is considered an intractable problem and estimating accuracy for all possible combinations of C and γ is a time consuming process. Therefore, exponentially increasing values were considered initially to find a better possible set of values for C and γ that yielded better accuracy. Finally, C and γ values thus obtained are varied slightly to gain the best possible accuracy. The estimated set of values for C and γ are 4 and 2, respectively.

(c) Training SVM using the chosen C and γ values to achieve the best cross-validation accuracy (CSV) possible.

(d) Predicting the output of the testing set.

(e) Estimating the accuracy of classification.

Since unbalanced data are used in this work the ratio of C^+/C^- are set to the inverse of the corresponding cardinalities of the classes.

Description of the Data Set

A total of 129 FHR signals of 20 minutes duration were collected from pregnant women at gestation ages of 34 to 40 weeks. Missing beats were eliminated from all signals using the recursive algorithm. CTG traces were submitted to two experienced obstetricians (20 years in practice) for visual inspection who were asked to classify the traces as 'normal' (+1) or 'at risk' (-1). From the 129 records 29 showed disagreement between experts, therefore were eliminated. Out of the remaining 100 signals, 90 (30 normal and 60 at risk) signals were considered for the purpose of training and testing so that the normal and at risk signals ratio can be maintained at 1:2 during training and also in estimating the cross-validation accuracy. Hence, the proposed methodology was tested on 90 FHR recordings having mutually agreed interpretation from two experienced obstetricians.

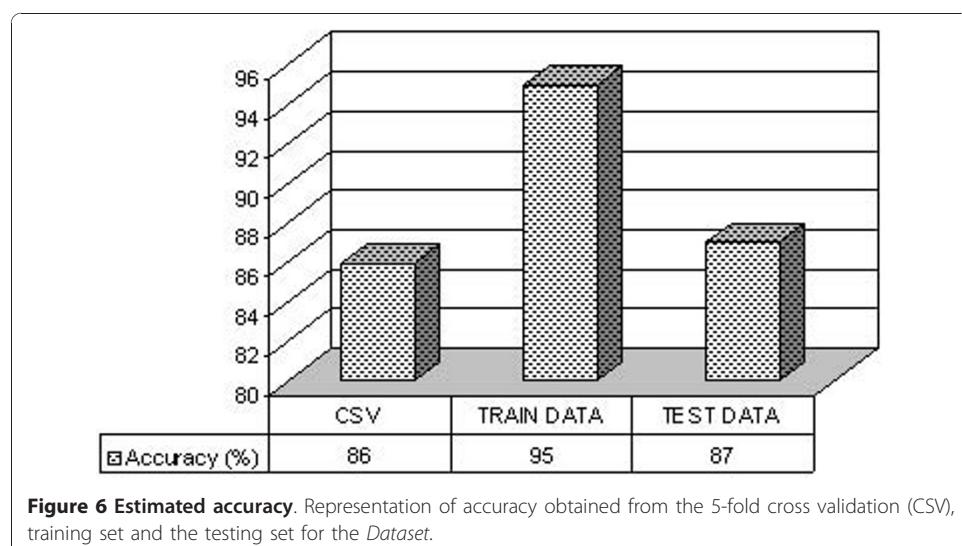
From the 90 FHR signals in the data set, *Dataset*, a training set and a testing set consisting of 60 and 30 data instances, respectively, were created. The training set had 20 'normal' and 40 'at risk' data instances and the testing set was composed of 10 'normal'

and 20 ‘at risk’ samples. The training set was further divided into five subsets, as explained later in the results section, in order to obtain more reliable results. Data instances in the *Dataset* consisted of a maximum of 10 statistical features (standard deviation) extracted from the remaining decomposed components (IMFs) of the FHR signal, after the separation of noisy components (preprocessing method). Since the number of remaining decomposed components of 90 FHR signals in the set varied from 7 to 10, a maximum of 10 input features or parameters were considered. Features in the *Dataset* extracted from 90 FHR signals were used as inputs to the SVM classifier.

Results

As a first step of validation a 5-fold cross validation method was employed using the FHR signals in the training set. This method is used for validating the systems, especially, when there is small number of data [50,1,49]. In this process, the training set was further divided into 5 non-overlapping subsets of 12 signals each. All the 5 sets had 4 instances that belonged to the ‘normal’ group and 8 from the ‘at risk’ group. The SVM classifier was trained with 4 subsets out of 5 and the 5th subset was used as the validation set and the procedure is repeated 5 times in order to find the 5-fold cross validation accuracy. The values for (C , γ) parameters are found using a systematic grid search method [1]. Because of the unbalanced data set used in this work the penalty parameters (C^+/C^-) ratio was maintained to be the inverse of the corresponding cardinalities of the classes (1/20/40). A 5-fold cross validation accuracy of 86% was obtained for the *Dataset* as shown in Figure 6.

In the next stage of evaluating the system performance, we used both the testing set and the training set and estimated the accuracy of prediction. The testing data set prediction accuracy was 87% and the training set accuracy of classification was 95%. Since the data used in the work is unbalanced the accuracy of prediction may not be the best measure [1,28].



In order to address this problem a method proposed [51] to estimate the geometric mean G_M (equation (15)) is used, whereas, S_T (sensitivity) and S_P (specificity) are estimated using equations (16) and (17), respectively.

$$G_M = \sqrt{S_T * S_P} \quad (15)$$

$$S_T = \frac{t_p}{(t_p + f_n)} \quad (16)$$

Where t_p and f_n stands for *true positives* and *false negatives*, respectively.

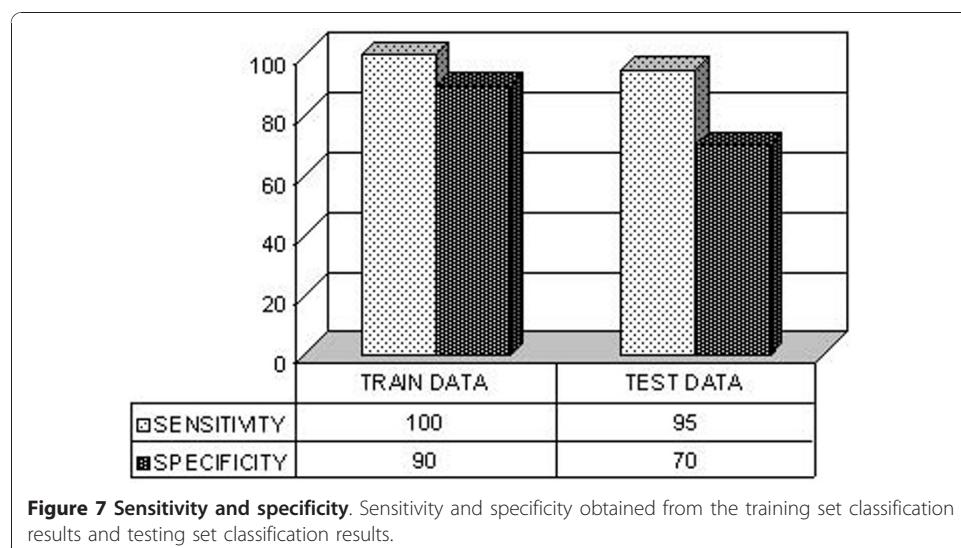
$$S_P = \frac{t_n}{(t_n + f_p)} \quad (17)$$

Where t_n and f_p stands for *true negatives* and *false positives*, respectively.

The S_T percentage obtained for the training data classification was 100% and S_P was 90%, and for the testing data set classification S_T was 95% and S_P was 70%, as shown in Figure 7. The t_p , t_n , f_p , and f_n values used for estimating S_T and S_P of the training and the testing data sets are taken from the confusion matrix shown in Figure 8a and Figure 8b, respectively. From S_T and S_P values the geometric mean G_M was estimated according to equation (15) [1,28], for both training and testing data: 94.87 (training set) and 81.55 (testing set).

In order to eliminate any agreement arrived by chance, the results of the SVM classifier and the mutually agreed visual interpretation results from the two experienced obstetricians were compared using inter-rater agreement Kappa statistics [50,52,6]. The kappa value k is estimated using equation (18).

$$k = \frac{p(a) - p(e)}{1 - p(e)} \quad (18)$$



		Human Expert			
		-1	1		
Algorithm Output	-1	40	2	42	
	1	0	18	18	
		40	20	60	
		Kappa		0.923	
		Standard error		0.0535	
		95% CI		0.818 to 1.028	
				Kappa	
				0.684	
		Standard error		0.147	
		95% CI		0.396 to 0.972	

Figure 8 Estimation of Kappa values. Presents confusion matrix and kappa value for the (a) training set, (b) testing set.

Here, $p(a)$ is the relative observed agreement and $p(e)$ is the hypothetical probability of agreement by chance. A 95% confidence interval (CI) is used in this work as shown in Figure 8. In Figure 8a and 8b the kappa value 0.923 with 95% CI was obtained for the training set and 0.684 for the testing set, respectively. The kappa values presented show a good agreement between the mutually agreed visual interpretation results from the two experienced obstetricians and the SVM classifier results, for both testing and training data.

Discussion

The proposed methodology showed that the statistical features extracted from the decomposed components using EMD yields good classification performance (based on kappa values). It is difficult to make direct comparison of the proposed method with others as the approaches are different. However, based on the methods employed for statistical/mathematical feature extraction and classification an indirect comparison of the proposed method with those developed by other researchers is provided. Thuner *et al* [53] employed wavelet transform based on the standard deviation corresponding to information contained in the coefficients 4th and 5th, and managed to achieve a complete separation between the class of healthy adults and adults with cardiac pathology.

Salamalekis *et al* [26] in their intrapartum study of diagnosing fetal hypoxia, from 10-minutes FHR patterns, used only the 2nd, 3rd and 4th wavelet coefficients and obtained a sensitivity of 97.9% and specificity of 83.3%. The use of fetal pulse oximetry may have contributed for the high performance value of their methodology. Georgoulas *et al* [48] considered standard deviations corresponding to all the six wavelet coefficients as input features and achieved an overall classification performance of 90% for 3 minutes window. They also [1] used features based on the entropy measure of the wavelet coefficients and obtained a maximum geometric mean of 83.67% with a sensitivity of 75% and a specificity of 93.33% on 5-minutes segment. In all of these works [1,48,26] FHR traces were associated to umbilical artery pH values and since there is no consensus (gold standard) regarding its threshold, different values were considered to discriminate normal fetuses from those at risk. All these studies very concerned with the prediction of metabolic acidosis during the intrapartum period. More recently, Kampouraki *et al* [49] extracted features using statistical methods and signal analysis methods (Wavelet Transform) but from the adult heart rate and achieved an accuracy of 100%.

In our study employing statistical features extracted based on EMD, no single feature could be identified that was capable of discriminating the 'normal' from 'at risk' classes, therefore all the statistical features were considered as inputs to the SVM classifier. With 20 minutes FHR signals, a sensitivity of 100% and specificity of 90% were achieved for the training set. Whereas, the testing set classification rates showed significant difference in the value of sensitivity (95%) and specificity (70%), even though the penalty parameters were set to handle the burden on the high false positive rate because of the imbalance nature of the data set used. SVM classifier has been successfully used in FHR feature classification [1,28,48,30,29,31], but most of the work in this field is based on extracting morphological features and providing the classification using SVM.

Conclusion

A new method of statistical feature extraction from FHR signals using EMD is proposed in this work. The features extracted from the decomposed components were further classified as 'normal' and 'at risk' by the SVM classifier.

Because of the lack of gold standards in evaluating the performance of intelligent systems [54] the proposed method was validated using the mutually agreed visual classification results from two experienced obstetricians. Validation of intelligent approaches based on the visual interpretation of a team of obstetricians has been used in evaluating several systems [6,54,23]. A high inter-observer and intra-observer variability in visual inspection is reported in the literature; however, it is important to note that it is largely dependent on the expertise and experience of the clinician involved [6]. Therefore, visual classification results from obstetricians with over 20 years experience were used in this study.

The inter-rater agreement kappa values obtained for the training set (0.923) and the testing set (0.684) showed good agreement of the proposed methodology with the mutually agreed visual classification results of two experts. This proved the viability of the method and its potential for further application.

The major limitations of this method are: the sifting process used in EMD is time consuming and the number of decomposed components varies with respect to the signal resulting in some empty spaces in the feature set. The preprocessing stage was helpful in removing the noisy components (IMFs). At this stage we can state that the results are quite promising. However, to be of clinical significance the proposed methodology requires extensive validation on a bigger data set.

As a future work, the proposed methodology would be applied to FHR signals of different durations and also would be extended for multi-class classification (here only two class labels were considered). Another possible extension is the classification of FHR by associating the traces to apgar scores. The effect of different sampling rates on features extracted using EMD and its effect on the SVM classifier also need further investigation.

Acknowledgements

The authors would like to thank Universiti Kebangsaan Malaysia (Grant UKM-AP-TKP-07-2009) for supporting this work. The authors would also like to express their gratitude to Prof. Dr. Muhamad Abdul Jamil M. Yassin for his support and guidance.

Author details

¹Department of Electrical Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi, Malaysia.. ²School of Electrical Engineering, Sharif University of Technology, Tehran, Iran. ³Department of Obstetrics and Gynecology, Universiti Kebangsaan Malaysia Medical Center, Cheras, Malaysia. ⁴Department of Obstetrics and Gynecology, Cyberjaya University College of Medical Sciences, Cyberjaya, Malaysia.

Authors' contributions

NKB was involved in the conception, design, development and analysis of the system and also drafted the manuscript. MAMA participated in the conception, design and analysis of the system, and helped in drafting the manuscript. ZE was involved in the conception of the study and provided critical review that helped in improving the manuscript. SA participated in conception, data acquisition and interpretation of FHR data. HFM was involved in the FHR data interpretation and analysis. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 10 November 2010 Accepted: 19 January 2011 Published: 19 January 2011

References

1. Georgoulas G, Stylios CD, Groumpos PP: Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machines. *Int. J. AI Tools* 2006, **15**:411-432.
2. Kamath MV, Fallen EL: Power spectral analysis of heart rate variability: a noninvasive signature of cardiac autonomic function. *Crit. Rev. Biomed. Eng* 1993, **21**(Suppl 3):245-311.
3. Signorini MG, Magenes G, Cerutti S, Arduini D: Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. *IEEE Trans. Biomed. Eng* 2003, **50**:365-374.
4. Lecanuet JP, Schaal B: Fetal sensory competencies. *Eur. J. Obstet. Gynecol. Reprod. Biol* 1996, **68**:1-23.
5. Bernardes J, Costa-Pereira A, Ayres-de-Campos D, Geijn HP, Pereira-Leite L: Evaluation of interobserver agreement of cardiotocograms. *Int. J. Gynaecol. Obstet* 1997, **57**:33-37.
6. Guijarro-Berdinas B, Alonso-Betanzos A, Fontenla-Romero O: Intelligent analysis and pattern recognition in cardiotocographic signals using a tightly coupled hybrid system. *Artif. Intell* 2002, **136**:1-27.
7. Georgoulas G, Gavrilis D, Tsoulous IG, Stylios CD, Bernardes J, Groumpos PP: Novel approach for fetal heart rate classification introducing grammatical evolution. *Biomedical Signal Processing and Control* 2007, **2**:69-79.
8. Mantel R, Geijn HP, Caron FJ, Swartjes JM, Woerden EE, Jongsma HW: Computer analysis of antepartum fetal heart rate: 1. Baseline determination. *Int. J. Biomed. Comput* 1990, **25**:261-272.
9. Mantel R, Geijn HP, Caron FJ, Swartjes JM, Woerden EE, Jongsma HW: Computer analysis of antepartum fetal heart rate: 2. Detection of accelerations and decelerations. *Int. J. Biomed. Comput* 1990, **25**:273-286.
10. Dawes GS, Moulden M, Redman CW: Computerized analysis of antepartum fetal heart rate. *Am. J. Obstet. Gynecol* 1995, **173**:1353-1354.
11. Taylor GM, Mires GJ, Abel EW, Tsantis S, Farrell T, Chien PFW, Liu Y: The development and validation of an algorithm for real time computerized fetal heart rate monitoring in labour. *Br. J. Obstet. Gynaecol* 2000, **107**:1130-1137.
12. Ulbricht C, Dorffner G, Lee A: Neural networks for recognizing patterns in cardiotocograms. *Artif. Intell. Med* 1998, **12**:271-284.
13. Fontenla-Romero O, Alonso-Betanzos A, Guijarro-Berdinas B: Adaptive Pattern Recognition in the Analysis of Cardiotocographic Records. *IEEE Trans. Biomed. Eng* 2001, **48**:1188-1195.
14. Bernardes J, Moura C, Sa JPM, Pereira-Leite L: The Porto system for automated cardiotocographic signal analysis. *J. Perinat. Med* 1991, **19**:61-65.
15. Bernardes J, Moura C, Sa JPM, Pereira-Leite L, Geijn HP: The Porto system. In *A Critical Appraisal of Fetal Surveillance*. Edited by: HP. van Geijn, FJA. Copray. Elsevier Science New York; 1994:315-324.
16. Ayres-de-Campos D, Bernardes J, Garrido A, de Sa JPM, Pereira-Leite L: SisPorto 2.0: a program for automated analysis of cardiotocograms. *J. Matern. Fetal Med* 2000, **9**:311-318.
17. Rooth G, Huch A, Huch R: Guidelines for the use of fetal monitoring. *Int. J. Gynecol. Obstet* 1987, **25**:159-167.
18. Kol S, Thaler I, Paz N, Shmueli O: Interpretation of non-stress tests by an artificial neural network. *Am. J. Obstet. Gynaecol* 1995, **172**:1372-1379.
19. Chung TKH, Mohajer MP, Yang XJ, Chang AMZ, Sahota DS: The prediction of fetal acidosis at birth by computerized analysis of intrapartum cardiotocography. *Br. J. Obstet. Gynaecol* 1995, **102**:454-460.
20. Magenes G, Signorini MG, Arduini D: Classification of cardiotocographic records by neural networks. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00)* volume 3: 24-27 July 2000 Como, Italy; 2000, 637-641.
21. Magenes G, Signorini MG, Sassi R, Arduini : Multiparametric analysis of fetal heart rate: comparison of neural and statistical classifiers. *IFMBE Proceedings Medicon 2001-IX Mediterranean Conference on Medical Engineering and Computing: Part I* Pula, Croatia; 2001, 360-363.
22. Ifeachor EC, Keith RDF, Westgate J, Greene KR: An expert system to assist in the management of labour. In *Proceedings World Congress on Expert Systems* volume 4: 16-19 December 1991; Orlando. Edited by: Liebowitz. Oxford: Pergamon Press; 1991:2615-2622.
23. Skinner JF, Garibaldi JM, Ifeachor EC: A fuzzy system for fetal heart rate assessment. *Proceedings 6th Fuzzy Days on Computational Intelligence, Theory and Applications: 25-28 May 1999* Dortmund, Germany; 1999, 20-29.
24. Alonso-Betanzos A, Moret-Bonillo V, Devoe LD, Searle JR, Banias B, Ramos E: Computerized antenatal assessment: The NSTEXPERT project. *Automedica* 1992, **14**:3-22.
25. Alonso-Betanzos A, Guijarro-Berdinas B, Moret-Bonillo V, Lopez-Gonzalez S: The NST-EXPERT project: the need to evolve. *Artif. Intell. Med* 1995, **7**:297-313.

26. Salamalekis E, Thomopoulos P, Giannaris D, Salloum I, Vasilios G, Prentza A, Koutsouris D: 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. Gynaecol* 2002, **109**:1137-1142.
27. Struzik ZR, Wijngaarden WJ: Cumulative effective Hölder exponent based indicator for real time fetal heart beat analysis during labour. *CWI Report INS-R0110* 2001, 1-13.
28. Georgoulas G, Chrysostomos Stylios D, Groumpas PP: Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using support vector machines. *IEEE Trans. Biomed. Eng* 2006, **53**:875-884.
29. Magenes G, Pedrinazzi L, Signorini MG: Identification of fetal suffrage antepartum through a multiparametric analysis and a support vector machine. *Proc. Of the 26th Annual Int. Conf of the IEEE EMBS: 1-5 September 2004* San Francisco, USA; 2004, 462-465.
30. Lunghi F, Magenes G, Pedrinazzi L, Signorini MG: Detection of fetal distress through a support vector machine based on fetal heart rate parameters. *Computers in Cardiology* 2005, **32**:247-250.
31. Warrick PA, Kearney RE, Precup D, Hamilton EF: System-Identification noise suppression for intra-partum cardiotocography to discriminate normal and hypoxic fetuses. *Computers in Cardiology* 2006, **33**:937-940.
32. Ferrario M, Signorini MG, Magenes G: Complexity analysis of the fetal heart rate variability: early identification of severe intrauterine growth-restricted fetuses. *Med. Biol. Eng. Comput* 2009, **47**:911-919.
33. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH: The empirical mode decomposition and hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. Lond. A volume 454: 8 March 1998*, 903-995, (1998).
34. Liang H, Lin Z, McCallum RW: Artifact reduction in electrogastrogram based on the empirical model decomposition method. *Med. Biol. Eng. Comput* 2000, **38**:35-41.
35. Liang H, Lin Q-H, Chen JDZ: Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease. *IEEE Transactions on Biomedical Engineering* 2005, **52**:1692-1701.
36. Blanco-Velasco M, Weng B, Barner KE: ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in Biology and Medicine* 2008, **38**:1-13.
37. Salisbury JI, Sun Y: Assessment of chaotic parameters in nonstationary electrocardiograms by use of empirical mode decomposition. *Ann. Biomed. Eng* 2004, **32**:1348-1354.
38. Echevarría JC, Crowe JA, Woolfson MS, Hayes-Gill BR: Application of empirical mode decomposition to heart rate variability analysis. *Med. Biol. Eng. Comput* 2001, **39**:471-479.
39. Ortiz MR, Bojorges ER, Aguilar SD, Echeverría JC, González-Camarena R, Carrasco S, Gaitan MJ, Martínez A: Analysis of high frequency fetal heart rate variability using empirical mode decomposition. *Computers in Cardiology* 2005, **32**:675-678.
40. Krupa BN, Mohd Ali MA, Zahedi E: Application of empirical mode decomposition for the enhancement of cardiotocograph signals. *Physiol. Meas* 2009, **30**:729-743.
41. Vapnik VN: *The Nature of Statistical Learning Theory* New York: Springer-Verlag; 1995.
42. Burges CJC: A tutorial on support vector machines for pattern recognition. *Data Mining Knowledge Discovery* 1998, **2**:121-167.
43. Muller KR, Mika S, Ratsch G, Tsuda K, Scholkopf B: An introduction to kernel-based learning algorithms. *IEEE Trans. Neural Netw* 2001, **2**:181-201.
44. Veropoulos K, Cristianini N, Campbell C: The application of support vector machines to medical decision support: a case study. *Advanced Course in Artificial Intelligence* 1999, 1-6, (ACAI'99).
45. Duin RPW: Classifiers in almost empty spaces. *International Conference on Pattern Recognition (ICPR'00) volume 2: 3-8 September 2000* Barcelona, Spain; 2000, 1-7, Proc. 15th.
46. Haykin S: *Neural Networks: A Comprehensive Foundation*. 2 edition. Prentice Hall: New Jersey; 1999.
47. Osuna E, Freund R, Girosi F: Support Vector Machines: Training and applications. *MIT A. I Lab., A. I. Memo AIM - 1602* 1997, 1-42.
48. Georgoulas G, Chrysostomos Stylios D, Groumpas PP: Classification of fetal heart rate using scale dependent features and support vector machines. *Proc of 16 thIFAC World Cong.: 4-8 July 2005 Prague*; 2005.
49. Kampouraki A, Manis G, Nikou C: Heartbeat time series classification with support vector machines. *IEEE Trans. on Information Technology in Biomedicine* 2009, **13**:512-518.
50. Alonso-Betanzos A, Mosquera-Rey E, Moret-Bonillo V, Baldonedo del Rio: Applying statistical, uncertainty-based and connectionist approaches to the prediction of fetal outcome: a comparative study. *Artificial Intelligence in Medicine*, Elsevier Science 1999, **17**:37-57.
51. Kubat M, Matwin S: Addressing the Curse of Imbalanced Training Sets: One-Sided Selection. *Proceedings of the 14th International Conference on Machine Learning: 8-12 July 1997* Nashville, Tennessee, USA; 1997.
52. Cohen J: A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960, **20**:37-46.
53. Thuner S, Feurstein MC, Teich MC: Multiresolution wavelet analysis of heartbeat intervals discriminates healthy patients from those with cardiac pathology. *Physical Review Letters* 1998, **80**:1544-1547.
54. Guijarro-Berdinas B, Alonso-Betanzos A: Empirical evaluation of a hybrid intelligent monitoring system using different measures of effectiveness. *Artif. Intell. In Medicine*, Elsevier 2002, **24**:71-96.

doi:10.1186/1475-925X-10-6

Cite this article as: Krupa et al.: Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. *BioMedical Engineering OnLine* 2011 **10**:6.