

# A Novel Scheme to Classify EHG Signal for Term and Pre-Term Pregnancy Analysis

Sindhiya Arora

Department of Instrumentation and Control  
Netaji Subhas Institute of Technology  
Delhi University, Delhi, India

Girisha Garg

Department of Instrumentation and Control  
Netaji Subhas Institute of Technology  
Delhi University, Delhi, India

## ABSTRACT

Early prediction of premature pregnancy reduces neonatal death and helps in adoption of treatment well suited for the pre-term pregnancy state. There are scads of work done in the area of term and pre-term pregnancy analysis like artificial intelligence, regressive models, and higher order statistical models. This paper proposes a four-level decomposition of Electrohysterography (EHG) signals using Discrete Wavelet Transform (DWT) based on pyramid algorithm to obtain the final feature vector matrix. Classification is done using Support Vector Machines (SVM) by dividing the data into test and training sets. It is validated on a well known benchmark database from Physionet Database. The proposed method can be used for real time implementation owing to low computational cost, high speed and its feasibility to be implemented on hardware. The encouraging experimental results show that the technique gives an accuracy of 97.8% and can be a promising tool for investigating the risk of preterm labor.

## General Terms

Labor time detection, non-invasive techniques, pregnancy analysis

## Keywords

Discrete wavelet transform, labor time detection, term and pre-term pregnancy, Support Vector Machines, EHG.

## 1. INTRODUCTION

In analysis of pregnancy for labor period detection the use of non – invasive techniques is highly encouraged. One such promising technique is the uterine Electrohysterography (EHG) signals. The EHG records correspond to the activity of the uterine muscles. The main events extracted from the uterine EHG are the contractions (CT). The electrical activity during pre-term labor (labor prior to 37 weeks of completed gestation) is significantly different from the activity of term labor. These differences prove to be helpful in identifying the nature (term or preterm) of the delivery. Owing to its simple and non- invasive nature this technique finds huge acceptance in hospitals. The premature delivery can be a threat to the child if not detected timely, as it may lead to birth of handicapped child with mental, neurological or behavioral abnormalities. But, among 80% cases it causes neonatal death. Thus, knowing the time of labor can help in adoption of treatment well suited for the preterm pregnancy state.

Due to the complications involved it is an extremely difficult task to predict the labor time hence, it is necessary to rely only on a technique having very high accuracy of separating the term pregnancy from the pre-term pregnancy.

## 2. RELATED WORK

### 2.1 Review of literature

A review of previous literature shows that immense work has been done in noninvasive techniques for determination of preterm delivery. Several model-based approaches [1], [5], [6], [7], [12], [13] have been proposed in this area. Jerzy et. al [1] classified the term and pre-term data using the Lagrangian Support Vector Machines (SVM). In [5] Marwa Chendeb et. al used wavelet transform and then classified using artificial neural networks and SVM. Various linear and non- linear processing techniques were used by G. Fele et. al [6]. In [12] classical techniques of data analysis, such as Principal Component Analysis (PCA) and Discriminant Analysis (DA) have been used. In [7], [13] Artificial Neural Networks (ANN) has been used on the EHG signals quantified by finding the means and standard deviations of the power spectrum. But, still this area lacks an effective practical method to access whether the uterine signals have entered the phase of uterus activity-burst that may indicate labor time.

### 2.2 Motivation

The following drawbacks in prior related works have led us to develop a method that mitigates them-

- Reducing the data dimensionality, like in PCA, may lead to loss of significant information [12].
- Classical statistical methods have the disadvantage of being probabilistic and computationally intensive .
- The low accuracies of prediction [1], [5].

In order to overcome these problems we present a novel methodology for automatic identification of pre-term/early pregnancy.

**Table 1: Techniques Used For Term / Pre – Term Labor Time Classification**

AUTHOR	YEAR	FEATURE EXTRACTION	CLASSIFICATION
Jerzy Sikora et. al	2011	–	Lagrangian Support Vector Machines
Marwa Chendeb et. al	2010	Wavelet Packet Transform	Neural Networks (NN) + Support Vector Machines (SVM)
Maartje Vinken et. al	2009	–	Artificial Neural Network
G. Fele et. al	2008		Linear(rms, peak, median, auto-correlation + non-linear

			(Lyapunov exponent, correlation dimension, entropy)
William Maner, Robert Garfield	2007	–	Artificial Neural Network
H. Lenman	1999	–	Principal Component Analysis and Discriminant Analysis

### 2.3 Organization

The rest of the paper is organized as follows: Section III contains the details of implementation. Section IV contains detailed description of the proposed system. Section V describes the methodology. Section VI shows the results. In Section VII, conclusion is presented.

## 3. DETAILS OF IMPLEMENTATION

### 3.1 Database

The uterine EHG signals available on the Physionet database have been utilized for the study. The Electrohysterogram records (uterine EHG records) included in the Term-Preterm Electrohysterogram Database (TPEHGDB) were obtained from 1997 to 2005 at the University Medical Centre Ljubljana, Department of Obstetrics and Gynecology. The records were obtained during regular check-ups either around the 22nd week of gestation or around the 32nd week of gestation.

### 3.2 Platform

The proposed system is implemented on an Intel I-3 processor. MATLAB has been used as the implementation platform during this research. MATLAB provides good functionality for research and experimentation purposes.

## 4. PROPOSED SYSTEM

The system used for automatic identification of the EHG signals is described using a block diagram shown in Figure 1.

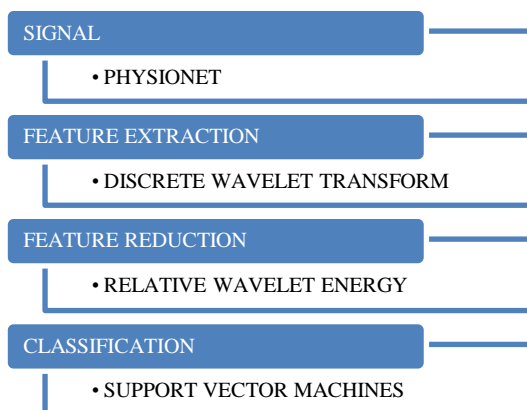


Fig 1: The Proposed System

### 4.1 Signal Acquisition

The uterine EHG signals of a total of 146 pregnant women were downloaded from Physionet. One fourth of the data records were separated into test set (36) and the rest (110) as training set. The training set was grouped further into four groups according to the time of recording (before or after the

26th week of gestation) and according to the total length of gestation (term delivery records— pregnancy duration  $\geq 37$  weeks and pre-term delivery records— pregnancy duration  $< 37$  weeks).

### 4.2 Feature Extraction

The dimension of the EHG signals is too high to be used directly for classification. Therefore it is important to apply feature extraction techniques in order to obtain informative features. Among all the feature extraction techniques used for signal processing, wavelet transform is found to be most successful. This is due to the reason that wavelet transform works in time and frequency domains. Therefore Discrete Wavelet Transform (DWT) is used to extract the features.

### 4.3 Feature Reduction

Further the features representing the burst of electrical activity are extracted by applying Relative Wavelet Energy (RWE) on the EHG signals. This feature reduction tools helps in further cutting down the feature vector matrix which is then sent for classification.

### 4.4 Classification

The classification stage is used to discriminate between different classes, term and pre- term pregnancy state. A classifier plays a very important role and thus it should accurately separate the data into the two groups. For this work Support Vector Machines (SVM) is used for classification purposes. The classifier is first trained using the training set and then tested using the test set.

## 5. METHODOLOGY

### 5.1 Wavelet transform

A crucial part of the EHG processing consists of transforming the information acquired from the signals into a small number of components which represent the uterine activity. Traditional Fourier transform methods (Fast Fourier transform and Short time Fourier Transform) have proved to be extremely insightful over the years as a feature extraction technique. However, there are various limitations while applying these techniques like they are not suitable to extract features localized simultaneously in time and frequency domain. Due to this reason they cannot be used to analyze transient signals especially when it is required to generate features for detection and discrimination for critical applications. Over the past several years, the methods based on Wavelet Transform (WT) (Stationary WT, Discrete WT, Wavelet Packet) have received a great deal of attention for extracting information from the EHG signals. This can be accounted to the Multi resolution analysis (MRA) which makes WT the most suitable candidate for analysis of frequency content of non stationary events which is a prerequisite for EHG signals.

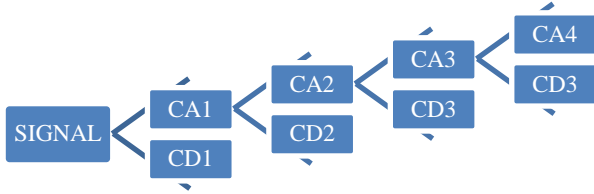
WT decomposes a signal into small waves with energy concentrated in time called wavelets. Wavelets are the scaled and shifted copies of the main pattern, so-called the mother wavelet. The mother wavelet function is defined by equation (1), where  $b$  is translation parameter and,  $a$  as scale parameter.

$$\Psi_{a,b} = \frac{1}{\sqrt{a}} \Psi \left( \frac{t-b}{a} \right) \quad \dots(1)$$

DWT analyses the signal using MRA by decomposing the signal into approximation and detailing information by employing two functions: scaling and wavelet function as shown in equation 1. The approximation coefficient is

subsequently divided into new approximation and detailed coefficients.

This process is shown in Fig. 2 which is carried out iteratively producing a set of approximation coefficients (CA) and detailed coefficients (CD) at four different levels of decomposition.



**Fig 2: Decomposition of a signal into approximation and detailed components**

For a given signal  $x(t)$  the DWT decomposition representing approximation and detailing coefficients can be represented by equation (2):

$$x(t) = \sum_{k=-\infty}^{k=+\infty} C_{N,k} \Phi(2^{-N}t - k) + \sum_{j=1}^N \sum_{k=-\infty}^{k=+\infty} d_{j,k} 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad \dots(2)$$

where  $C$  represents approximation coefficients at level  $N$ , while  $d_{j,k}$  ( $j = 1$  to  $N$ ) represents detailed coefficients or wavelet coefficients at level  $j$ .  $\psi(t)$  is the wavelet function, while  $\Phi(t)$  is a companion function, named as scaling function.

## 5.2 Relative Wavelet Energy

If the scaling functions and wavelets form an orthogonal basis, Parseval's theorem relates the energy of the signal  $x(t)$  to the energy in each of the components and their wavelet coefficients. The energy  $E_j$  of the detailed signal at each resolution level  $j$  is given by:

$$E_j = \sum_k |d_{j,k}|^2 \quad j = 1 \text{ to } N \quad \dots(3)$$

And the total energy for all the levels is given by:

$$E_{Total} = \sum_{j=1}^{N+1} E_j \quad \dots(4)$$

The wavelet energy can be used to extract only the useful information from the signal about the process under study. For this work the concept of relative energy has been used. RWE gives information about relative energy with associated frequency bands and can detect the degree of similarity between segments of a signal. RWE is defined by the ratio of detail energy at the specific decomposition level to the total energy. Thus the relative energy is given by:

$$RWE = \frac{E_j}{E_{Total}} \quad \dots(5)$$

RWE resolves the wavelet representation of the signal in one wavelet decomposition level corresponding to the representative signal frequency. Thus this method accurately detects and characterizes the specific phenomenon related to the different frequency bands of the EHG signal. RWE gains an advantage over DWT based feature extraction in terms of speed, computation efficiency and classification rate.

## 5.3 Support Vector Machines

The approach of SVM is to find an optimal hyper-plane in the multi-dimensional space of features, which would separate the classes being considered with the largest distance (margin) to the nearest training data points. The data (input vectors) that ensure this safety margin are called the support vectors. This optimal hyper plane is constructed in such a way that it maximizes the minimal distance between itself and the learning set.

$$g(y) = \text{sgn} \left( \sum_{i=1}^{L_s} d_i \alpha_i K(y_i, y) + b \right) \quad \dots(6)$$

Maximize

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l d_i d_j \alpha_i \alpha_j K(y_i, y_j) \quad \dots(7a)$$

Subject to

$$\sum_{i=1}^l \alpha_i d_i = 0 \quad 0 \leq \alpha_i \leq C \text{ for } i = 1, 2, \dots, l \quad \dots(7b)$$

$$K(u, v) = \exp \left( - \frac{\|u - v\|^2}{2\sigma^2} \right) \quad \dots(8)$$

## 6. RESULT

In this paper, DWT is exploited to decompose the EHG signal into (five) frequency bands. Then features based on RWE are extracted from DWT decomposed matrix. One of the most important issues while addressing DWT for signal analysis is choosing the appropriate wavelet and decomposition level. After performing several tests db2 wavelet and four decomposition levels were chosen for this study. In order to achieve good generalization capability, it is important to have the number of training data points several times larger than the number parameters being estimated. This limitation impeded the decomposition level to be increased beyond 6 and for lesser decomposition levels i.e. less than 4, there is insufficient detailing in terms of frequency bandwidth preventing the model from giving accurate results. The features extracted from DWT and RWE are used to classify the signals using SVM.

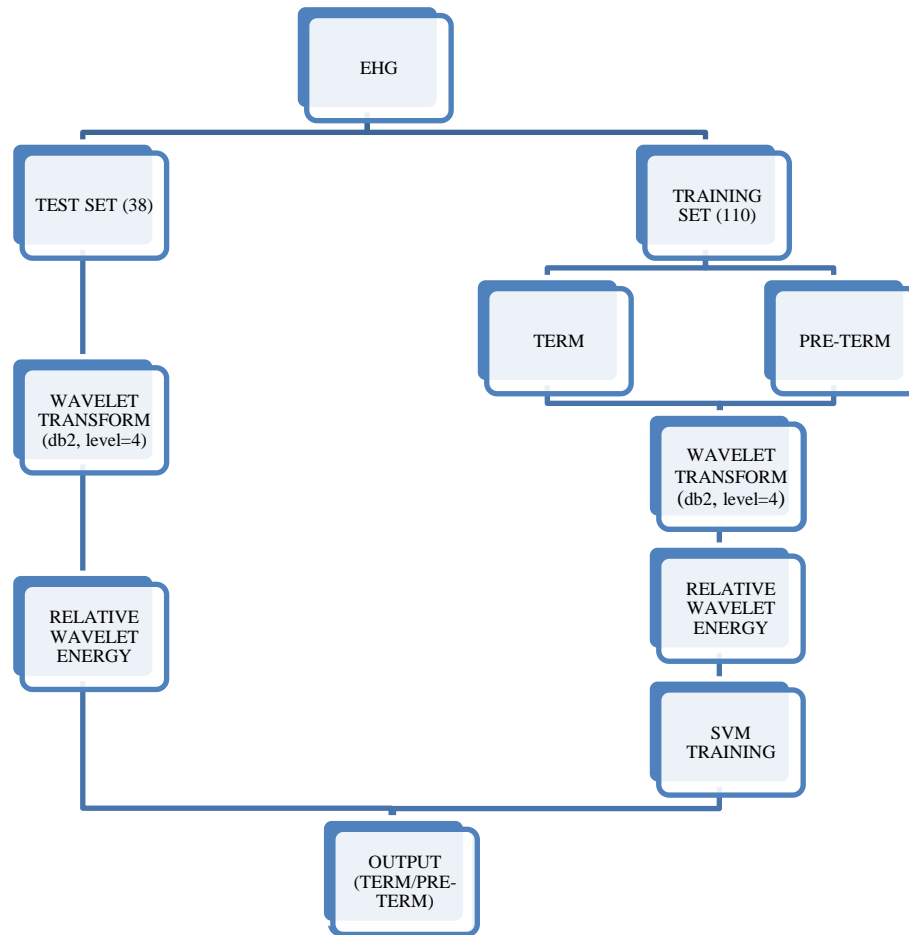


Fig 3: Steps involved in feature extraction and classification

The linear SVM classifier cannot be used in this case because of high and non linear dimensional space. Therefore kernel mapping is required to implement the SVM classifier. In this case the Gaussian kernel is used. The classifier is trained using the training set.

A total of 2336 features were obtained after feature extraction. Out of which, 608 features constitute the test matrix and the remaining 1728 features constitute the training matrix. The performance achieved for the training set is 100%. For the test set, 581 features were correctly identified, thus, the accuracy achieved is 95.6%. The average accuracy achieved is 97.8%. The comparison of average accuracies of this method (DWT + SVM) and the different methods [1], [5] is shown below.

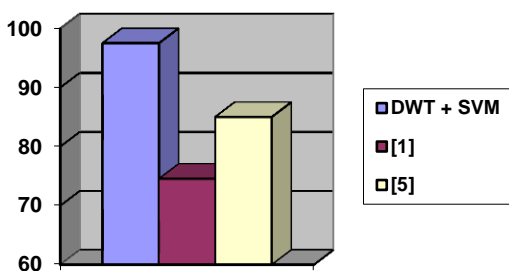


Fig 4: Comparison of average accuracies of various methods.

## 7. CONCLUSION

In this paper DWT based feature extraction is incorporated to classify the EHG signals using Support Vector Machines based classifier. The RWE of different EHG signals were obtained which provided a compact and accurate feature space to achieve better classification rate and reduced computational load. Due to its memory efficient capability this concept can generalize the EHG processing to much larger datasets. The SVM based classifier presented has a good performance in classifying the 2 stages of labor and can be easily implemented on hardware for real time response. The most redeeming feature of proposed system is its capability to achieve excellent computational efficiency without any loss of information. This sets apart the system from the other automatic classifiers falling in the same levels of accuracy without any requirement of topnotch PCs. Owing to the suitably high accuracy achieved using the synergy of SVM and RWE technique, the methodology presented in this paper can be utilized for designing automatic EHG classifiers which can be used by the gynecologist for detecting premature birth risk.

## 8. REFERENCES

- [1] Jerzy Sikora, Adam Matonia, Robert Czabanski, Krzysztof Horoba, Janusz Jezewski, Tomasz Kupka—“Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals”, 2011.
- [2] Girisha Garg, Shruti Suri, Vijander Singh, Rachit Garg, “Wavelet Energy based Neural Fuzzy Model for

- Automatic Motor Imagery Classification”, Vol. 28, No. 7, Aug 2011.kk
- [3] Girisha Garg, Vijander Singh, J.RP Gupta, A.P. Mittal, Sushil Chandra, “Model Computer Assisted Automatic Sleep Scoring System Using Relative Wavelet Energy Based Neuro Fuzzy”, Vol. 8, Issue 1, Jan 2011.
  - [4] Marwa Chendeb, Mohamad Khalil, David Hewson, Jacques Duchene, “Classification of non stationary signals using multiscale decomposition”, Dec 2009.
  - [5] G. Fele-Zorz, G. Kavs, Z. NovakAntolic, F. Jager, “A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups”, Apr 2008.
  - [6] William L. Maner and Robert E. Garfield, “Identification of Human Term and Preterm Labor using Artificial Neural Networks on Uterine Electromyography Data”, Vol. 35, Number 3, 2007.
  - [7] Gang Wang, Zhizhong Wang, Weiting Chen, Jun Zhuang, “Classification of surface EMG signals using optimal wavelet packet method based on Davies-Bouldin criterion”, Sept, 2006.
  - [8] M. B. I. Reaz, M. S. Hussain and F. Mohd-Yasin, “Techniques of EMG signal analysis: detection, processing, classification and Applications”, Jan, 2006.
  - [9] H. Eswaran, J. D. Wilson, P. Murphy, H. Preissl and C. L. Lowery, “Application of wavelet transform to uterine electromyographic signals recorded using abdominal surface electrodes”. Vol.11, No. 3, 2002
  - [10] The EHG dataset has been downloaded from this site <http://www.physionet.org/cgi-bin/atm/ATM>.
  - [11] Leman, H.; Marque, C.; Gondry, J.; Dept. of Biomed. Eng., Compiegne Univ., “Use of the electrohysterogram signal for charaterization of contractions during pregnancy” France, Vol- 46 Issue: 10, Oct. 1999.
  - [12] Vinken, Maartje P. G. C. MD; Rabotti, Chiara MSc; Mischi, Massimo PhD; Oei, S Guid MD, PhD – “Accuracy of Frequency Related Parameters of the Electrohysterogram for Predicting Preterm Delivery: A Review of the Literature”, Vol. 64, Issue- 8.
  - [13] Polikar R., “The Wavelet Tutorial”. Available: <http://users.rowan.edu/~polikar/wavelets/wttutorial.html>, 1999.