A Survey of Computerized Fetal Heart Rate Monitoring and Interpretation Techniques

Miao Liu, Yuzhong Shen, Mark Scerbo

Abstract — Fetal heart rate (FHR) reveals crucial information about fetal conditions and FHR monitoring is a standard clinical procedure widely used during pregnancy and labor. FHR tracing can be interpreted as reassuring, non-reassuring, and ominous patterns. Accurate interpretation of fetal state from FHR tracing can prompt necessary interventions before irreversible conditions occur and reduce the number of unnecessary cesarean sections. Although FHR tracing can be obtained automatically using electrocardiograms and cardiotocography, FHR monitoring and interpretations are conducted manually by nurses or obstetricians, resulting in substantial inter-observer and intra-observer disparities. There have been great needs to automate FHR monitoring and interpretation in order to produce accurate and consistent clinical judgements and practices. A number of methods have been developed to address various issues in computerized automatic FHR monitoring and interpretation, including FHR signal modeling and representation, feature extraction, and pattern classification. FHR signals can be considered as stochastic processes driven by the cardiovascular control system and both linear and nonlinear methods have been proposed to model FHR signals. A wide range of parameters, such as morphological features, temporal domain parameters, frequency domain parameters, and nonlinear parameters, have been defined to represent different FHR features. Finally, various classification methods have been utilized to categorize FHR signals into reassuring, non-reassuring, and ominous patterns. This paper provides a comprehensive review of various approaches in FHR signal modeling, feature extraction, and pattern classification and discusses several methods that are under investigation and can possibly improve FHR modeling and classification.

Index Terms — Fetal heart rate, modeling, feature extraction, spectrum analysis, pattern classification.

I. INTRODUCTION

Electronic fetal heart rate monitoring (EFM) plays important roles before and during labor [1]. It reveals crucial information about fetal conditions and provides close assessment of high-risk mothers. EFM provides a quantitative tool for evaluating the synergetic control activities performed by the nervous system that is critical in children’s later development. One important use of EFM is to prompt the obstetricians to perform necessary interventions before irreversible damages to the fetus take place. Since its introduction at Yale University in 1958 [2], EFM has been increasingly adopted in the monitoring of fetal condition and evaluation of the intrauterine environment on fetal welfare. In the 1970s, EFM became a common clinical practice thanks to the fact that uncontrolled trials demonstrated the use of EFM reduced intrapartum fetal death rate by a factor of more than three [3].

Even though EFM is now a standard procedure in labor and delivery, there are still debates concerning with its effectiveness and consistency [3]. EFM has increased the number of instrumental vaginal deliveries and cesarean sections performed. Interpretations of FHR patterns are qualitative, subject to inter-observer and intra-observer disparities due to varied clinical experiences and personal inconsistencies at different times. Manual FHR monitoring and interpretation bring both much physical and psychological stresses to the medical personnel, often resulting in erroneous observations. However, the arguments about the effectiveness of FHR monitoring are mainly concerned with the accurate interpretation of FHR tracings. Hence, there have been great needs and research efforts to develop computerized FHR monitoring and interpretation to produce accurate and consistent clinical judgments and practices [4].

A number of methods have been developed to address various issues in computerized automatic FHR monitoring and interpretation, including FHR signal modeling and representation, feature extraction, and pattern classification. As a first step, it is beneficial to build mathematical models that can gain insights into the underlying structures of FHR signals. Mathematically, FHR signals are one-dimensional random signals driven by the cardiovascular control system. Classical parametric signal model have been used for FHR modeling [5, 6], but they alone are too simple to describe all aspects of such a complex system. More advanced model have recently been proposed to unveil more information embedded in FHR signals [7]. To facilitate FHR feature extraction, additional analyses can be applied to various mathematical FHR models. Power spectrum analysis with Fourier transform (FT) is a classical method for this task, but FT can not provide temporal and spectral information simultaneously. Several advanced tools,
such as wavelet transform [8, 9] and Matching Pursuit [10], overcame this drawback and have been utilized for FHR analyses. In addition, various nonlinear features such as multi-scale entropy (MSE) [11] and fractal dimension (FD) [12] have been proposed to tackle the nonlinear characteristics of FHR signals. The final step in automatic FHR monitoring and interpretation is to classify FHR signals into reassuring, non-reassuring, and ominous patterns. Various recently developed classifiers have been utilized for this task. This paper provides a comprehensive review of various approaches in FHR signal modeling, feature extraction, and pattern classification and discusses several methods that are under investigation and can possibly improve FHR modeling and classification.

The remainder of this paper is organized as follows. Section II introduces typical FHR signal acquisition, modeling and representation methods. Section III reviews widely used linear and nonlinear parameters as well as their calculations. Section IV presents FHR patterns and their classification. In Section V we discuss several advanced issues that can possibly improve FHR modeling and classification. Section VI concludes this paper.

II. FETAL HEART RATE SIGNAL

A. Data Collection

Both electrocardiograms and cardiotocography are used to acquire FHR signals [13]. Electrocardiograms (ECG) can be obtained from a fetal scalp electrode or from the maternal abdomen, followed by peak detection to determine the $T_{ao}$-duration of fetal cardiac cycles. Cardiotocography (CTG) can be obtained using ultrasound-Doppler method through an abdominal probe. Fetal heartbeats are detected from the envelope of ultrasound wave reflected from the moving parts of fetal heart. CTG also records uterine activities through an external pressure transducer. Values of fetal heartbeats can be obtained using ultrasound-Doppler method through an abdominal probe. Fetal heart rate (HR) can be calculated using the equation: FHR [bpm] = 60 000/$T_{ao}$ [ms]. A typical fragment of CTG is illustrated in Fig. 1.

ECG allows accurate albeit time-consuming identification of fetal interbeat intervals and has better performance in terms of signal-to-noise ratio [15]. However, it bears two disadvantages: one is that there are cases of missing FHR extractions, and the other is that the maternal and fetal ECGs must be separated offline, a process that does not guarantee to produce perfect results. In contrast, CTG is able to obtain FHR under almost all conditions. The ultrasound method has become the standard approach since early 1970 because it is a noninvasive method and can be used during both pregnancy and labor.

B. Fetal Heart Signal Modeling and Representation

Since FHR signals are 1-D time series, traditional parametric model together with time-frequency analysis based on power spectral analysis [16] have been widely used to model and represent FHR signals. From a clinical point of view, FHR signals contain three components, namely, baseline, variability, and deceleration and acceleration. Most FHR research was focused on modeling FHR variability that indicates fetus’ ability to withstand the stress of delivery and is more difficult to detect than other FHR patterns such as acceleration and deceleration. In the early 1980s, W. Jarisch et al. [5] modeled FHR variability as a scalar output of a multivariate nonlinear stochastic process. They adopted the Kalman filter algorithm as part of an approximate maximum likelihood estimator of the process parameters. Their model is depicted in Fig. 2.

W. Jarisch’s model is relatively simple since the model’s order is predetermined and is only two. Maria G. Signorini, et al. [15] evaluated FHR characteristics by calculating FHR mean and variance based on power spectral density (PSD) estimation. They used an auto-regression (AR) model to model the FHR signal, and employed Akaike’s information criterion (AIC) [17] to determine the optimal model order. Finally, the PSD of the AR process, or the Fourier transform (FT) of auto-correlation function was calculated for further feature extraction and quantitative analysis.

Philip A. Warrick et al. [18] used discrete cosine transform (DCT) for spectrum analysis. They adaptively selected the resolution of analysis and used DCT to describe the spectrum at short-term and longer-term scales. To address situations where signals are not uniformly sampled, Cao et al. [19] proposed to use Lomb periodograms of FHR signals. Lomb periodogram analysis is the least squares fitting of the data to the sinusoid $A \cos(2\pi ft) + B \sin(2\pi ft)$ and it is suitable for analysis of unequally sampled signals.

Most of aforementioned studies were based on analyses
similar to the Fourier transform and auto-regression model. However these algorithms have limitations in the study of long-term variations of FHR as well as transient deviations, because Fourier transform only provides frequency information of the entire FHR signal and temporal information is lost in Fourier transform. On the other hand, wavelet analysis is a powerful multi-resolution tool that can provide localized information in both temporal and frequency domain and has been utilized in FHR signal processing [20]. Mathematically, a set of wavelet are formed by scaling and translating the mother wavelet. Wavelet transform provides both the scale and time dependent energy distribution.

Consider that classical Fourier transform may not represent signal that have stationary characteristics, and the wavelet transform may not represent signals whose Fourier transform has narrow frequency support, M. Akay et al [10] chose Matching Pursuit method [21] to examine FHR variability. The basic idea is a nonstationary signal can be expanded by waveforms whose time-frequency properties can be adapted to its local structures. The FHR signal can be decomposed into waveforms selected from a dictionary of time frequency wave atoms \( \psi \gamma(t) \), which are the dilations, translations, and modulations of a single window function \( \psi(t) \)

The linear expansion of \( f(t) \) over a set of vectors was selected from the dictionary (D) to match its local structures. This expansion can be done by successive approximations of \( f(t) \) with orthogonal projections on elements of D as follows

\[
f = \langle f, \psi \rangle \psi \gamma + Rf,
\]

where \( Rf \) represents the residual vector after approximating \( f \) in the direction of \( \psi \gamma \). The power of \( f(t) \) can be written as:

\[
\|f\|^2 = \|\langle f, u \gamma \rangle u \gamma\|^2 + \|Rf\|^2
\]

Thus the original signal is decomposed into a sum of dictionary wave atoms that are chosen to best match its residues.

With above FHR modeling and representation tools, the features of FHR signal could be easily analyzed using the methods discussed in Section III.

C. Artifact Removal and Denoising

Due to measurement error and circumference interference, signals are often contaminated by noise. To extract FHR features for accurate interpretation, the noise and artifact must be eliminated or attenuated. The most common method is smoothing of the time series based on a moving average.

Cao et al. suggested that noise due to fetal movement or displacement of the transducer, is usually “spiky” [19], and proposed a statistical filter based on outlier detection: a data point was regarded as “artifacts” if, within local window of width \( w \), it was either beyond the median value \( \pm b \) bpm, or beyond \( R \) folds of the interquartile range (IQR). The IQR was defined as the difference between \( p1 \) and \( p2 \) percentiles of the local window. If IQR = 0, a small value of \( b \) was selected, representing low variability. The optimal parameters of \( p1, p2, b, R, \) and \( w \) were determined by simulations.

Papadimitriou et al. [22] utilized wavelet transform for FHR signal denoising. They extracted wavelet coefficient module maximum (WTMM) which carried FHR variation—the most important information of FHR signals. The FHR signal variation and noisy components were distinguished by computing the Lipschitz regularity exponents [21]. If the Lipschitz regularity exponent is positive, the amplitude of the WTMM should increase with the scale, corresponding to signal variation. On the contrary, negative Lipschitz regularity exponents suggested that amplitudes decreased with the wavelet scale, indicating presence of noise. This method is especially useful for the removal of additive noise.

III. Fetal Heart Rate Features

The FHR modeling and representation methods discussed in the previous section are useful in unveiling the embedded information in FHR signals. By describing those informative components in forms of various features parameters, quantitative analysis of FHR patterns can be conducted later on. Generally, FHR features can be classified into three categories [6]. The first category is “morphological” features that have been mainly used by obstetricians for FHR interpretation in the past three decades. Two others types, temporal and frequency features, are also natural choices since FHR signals are essentially time series.

A. Morphological Features

Before reviewing the feature extraction methods, we first describe several commonly used terms in FHR monitoring. **Baseline:** the mean level of the fetal heart rate when it is stable, without acceleration or decelerations. The normal FHR range is between 120 and 160 bpm [1]. **FHR Variability:** beat-to-beat or short-term variability is the oscillation of the FHR around the baseline within the range of \( \pm 5 \) to \( \pm 10 \) bpm [1]. Long-term variability is a slower oscillation of heart rate and has a frequency of 3 to 10 cycles per minute and amplitude of 10 to 25 bpm [1]. **Acceleration:** a transient increase of heart rate of 15 bpm or more, lasting 15 seconds or more [1]. **Deceleration:** a transient decrease of fetal heart rate below the baseline level of 15 bpm or more, lasting 10 seconds or more [1].

From these definitions, it can be seen that the FHR baseline should be extracted first in order to perform any FHR analysis. Various methods have been proposed to extract the FHR baseline in the last three decades. In early 1980s, based on the principle that the baseline follows slow changes but not rapid changes in FHR, Dawes et al. [23] applied digital low-pass filtering to FHR signals using a moving window with a length of 3.75 seconds.

Dawes’s method is easy to implement. However, it suffered from four problems [24]: 1) poor baseline fitting at the beginning of the FHR recording; 2) baseline elevated in the case of acceleration; 3) elevation of baseline level caused by repeated large swings of FHR; and 4) baseline did not
adequately follow shifts in long term fluctuations. To address those problems, Mantel et al. [24] proposed an algorithm based on an iterative process of filtering and trimming, which was widely adopted for estimating baseline. It consists of the following steps: removal of components of the FHR signal associated with accelerations and decelerations; linear interpolation across the gaps, and low pass filtering (this procedure is repeated three times). The filtering process is based on a bidirectional low-pass filtering that ensures linear phase characteristic.

In recent years, several new methods have been proposed for baseline estimation. Instead of computing a moving average of 3.5 s or 2.5 s, L. Jimenez et al. [25] proposed a three-step procedure based on beat-to-beat derivative amplitude analysis: 1) Calculate the FHR average using a Hanning window with a length of 27; 2) Calculate the first derivative of FHR (dFHR), eliminate segments with dFHR exceeding the level of 1.0 bpm/s, average the remaining (μ), and select the pieces lasting more than 15 s and with amplitude between μ±10 bpm as possible baseline segments; 3) Interpolate the selected segments by cubic spline. The result is illustrated in Fig. 3. P. Warrick et al. [27] combined signal processing methods and neural networks to detect the FHR patterns of baseline, acceleration and deceleration.

In order to evaluate different baseline estimation algorithms, T. Kupka et al. [28] developed a statistical method for FHR modeling using a predetermined baseline component. Variations, accelerations, and decelerations are modeled separately using real FHR signals. The baseline extracted by an estimation algorithm is compared with that in the artificial FHR signal. The difference provides quantitative comparisons between different baseline extraction algorithms. Fig. 4 illustrates the artificial FHR signal generation process.

### B. Time-domain Parameters

For an FHR signal of length \( N \), the typical time-domain parameters are defined as follows [14]:

1. Mean value of FHR: \( \mu = (1/N) \sum_{i=1}^{N} f_{FHR}(i) \)
2. Standard deviation: \( \sigma = \sqrt{(1/(N-1)) \sum_{i=1}^{N} (f_{FHR}(i) - \mu)^2} \)
3. Delta:
   \[
   \delta = \frac{1}{m} \sum_{i=1}^{m} \max_{i \in [i-240, i+240]} \left( f_{FHR}(i) \right) - \min_{i \in [i-240, i+240]} \left( f_{FHR}(i) \right)
   \]
   where \( m \) is signal recording period in minutes.
4. Short-term variability: \( STV = \sum_{i=1}^{10} \left| f_{FHR}(i + 1) - f_{FHR}(i) \right| \)
   where \( f_{FHR}(i) \) is the value of the signal \( f_{FHR}(i) \) taken very 2.5 s (i.e., once very ten samples \( f_{FHR}(i) \) taken every 2.5 s and \( f_{FHR}(i+1) \) taken every 2.5 s plus 1 s)
5. Interval index: \( I = STV / (\sigma f_{FHR}(i)) \)
6. Long-term irregularity (LTI): defined as the inter quartile range \([1/4, 3/4] \) of the distribution with \( m(f) = \sqrt{f_{FHR}(i) - f_{FHR}(i+1)} \).
7. Total value of the Delta:
\[ \delta_{\text{total}} = \max_{m=1}^{\infty} \left( f_{\text{run}}(i) \right) - \min_{m=1}^{\infty} \left( f_{\text{run}}(i) \right) \]

C. Frequency-domain Parameters

Since a close relationship exists between the presence of spectral components and the activities of neural cardiovascular control system [15], PSD analysis and time frequency analysis tools can provide quantitative measurements of the activities performed by autonomic nervous system.

According to [29], the frequency range was partitioned into three bands: 1) very-low-frequency (VLF) band: 0-0.05 Hz; 2) low-frequency (LF) band: 0.05-0.15 Hz (referred to as Mayers’ waves); 3) high-frequency (HF) band: 0.15-0.5 Hz, which corresponds to fetal movements. The ratio of energy in the LF and HF bands is a standard measure used for adults and is presumed to convey the balanced behavior of the two branches of the autonomic nervous system.

Other frequency feature selection methods exist. For instance, [15] suggested that the frequency range can be divided into four bands: 1) VLF 0-0.03 Hz, related to long period nonlinear contribution; 2) LF: 0.03-0.05 Hz, mainly correlating with neural sympathetic activity; 3) Movement frequency (MF): 0.15-0.5 Hz, indicating fetal movements and maternal breathing; and 4) HF: 0.5-1 Hz, indicating fetal breathing. [15] also used the quantity LF/(HF+MF) to represent the autonomic balance between neural control mechanism from different origins.

D. Nonlinear Parameters

If observed over long periods of time, FHR signal are highly irregular and bear certain self-similarity structures, representing a nonlinear system behavior. Researchers have proposed the following parameters to model the FHR’s nonlinear nature.

(1) Approximate Entropy and Sample Entropy

Pincus [30] proposed a family of statistics, namely, approximate entropy (ApEn), to measure system complexity. A greater likelihood of remaining close (high regularity) produces smaller ApEn values, and vice-versa. Given a sequence of points, \( x(n) = x(1), x(2), \ldots x(N) \), the dimension of a vector (template, sub-sequence), and \( r \), a threshold for measuring similarity. The procedure of computing ApEn is shown below:

1) Assemble \( N-m+1 \) vectors, each of size \( m \), as follows:
\[ X^n(i) = [x(i), x(i+1), \ldots x(i+m-1)] \quad i = 1, \ldots N-m+1 \]

2) Compute the distance, \( d[X^n(i), X^n(j)] \) between each template and other templates (including itself) as
\[ d\left( X^n(i), X^n(j) \right) = \| X^n(i) - X^n(j) \| , 1 \leq p \leq \infty \]

3) For each template \( X^n(i) \), find the matched template number \( N^n(i) \) such that \( d[X^n(i), X^n(j)] \leq r \), and calculate the probability that any template \( X^n(j) \) matches \( X^n(i) \) as \( C^n(i) = N^n(i)/(N-m+1) \). Then take the natural logarithm of each \( C^n(i) \) and average them over \( i \), the theoretical value of the approximate entropy is defined as
\[ \Phi^n(r) = \frac{1}{\ln N} \sum_{i=1}^{N-m+1} \ln C^n(i) \]

4) By increasing the dimension of template to \( m+1 \) and repeating the previous steps to find \( \Phi^{m+1}(r) \), the theoretical value of the approximate entropy is defined as
\[ ApEn(m,r) = \lim_{N \to \infty} \ln \left[ \Phi^{m+1}(r) - \Phi^m(r) \right] \]

for a finite length of \( N \) points, the value of the ApEn is given by
\[ ApEn(m,r,N) = \Phi^{m+1}(r) - \Phi^m(r) \]

ApEn has three drawbacks [15]: 1) bias effects, due to including self-matching of each template to avoid \( \ln(0) \) in the calculation; 2) dependence on the record length; and 3) lack of consistency. That is, if ApEn of one sequence is higher than that of another, it does not remain higher for all conditions tested. To address these drawbacks, Richman and Moorman [31] proposed sample entropy (SampEn), an improved ApEn algorithm. In the computation of SampEn, 1) self-matches are not counted; 2) only the first \( N-m \) sub-sequence of length \( m \) are considered; 3) the conditional probabilities are not estimated in template fashion in step 3. Then
\[ \Phi^n(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} C^n(i) \]

and
\[ ApEn(m,r) = \lim_{N \to \infty} \ln \left[ \Phi^{m+1}(r) - \Phi^m(r) \right] \]

\[ ApEn(m,r,N) = \ln \left[ \Phi^{m+1}(r) - \Phi^m(r) \right] \]

(2) Multiscale Entropy

Both ApEn and SampEn are single index for analyzing the regularity of the time series and they cannot reveal the underlying dynamics of the generating system. To overcome this inefficiency, Costa [11], introduced the multiscale entropy (MSE) to represent the dynamics of the generating system. To compute MSE of a time series \( \{x_i\} \) of \( N \) points, construct consecutive coarse time series of different scale, \( \{x_i^{(r)}\} \), as
\[ x_i^{(r)} = \frac{1}{r} \sum_{j=[i-0.5r]}^{[i+0.5r]} x_j, 1 \leq j \leq N/r \]

which is a function of the factor \( r \), with length of \( N/r \). Then entropy of each sequence \( \{x_i^{(r)}\} \) is calculated and plotted.

Fig. 5. Multiscale sample entropy (MSE) results (mean ± S.D) (Original figure courtesy of Dr. H. Cao [19])

Differences in the Entropy at different time scales can help understand the time series in terms of regularity and structure
(i.e. short versus long range). Fig. 5 shows the MSE results of SampEn value as a function of scale for reassuring and non-reassuring patterns. We can see that nonreassuring FHR controlled by certain pathology, has less variability and thus have smaller SampEn than reassuring FHR in each scale.

(3) Fractal dimension (FD) and Self-Similarity Degree

Fractal geometry was first developed by Mandelbrot in 1970s. A Fractal is generally “a rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced-size copy of the whole [32].” It bears the property of “self-similarity” and was used to depict rugged, irregular phenomena by fractal dimension. A. Kikuchi et al. [12] employed FD as a parameter to study FHR variability during different gestational ages. They use Higuchi’s method [33] to calculate the FD for a given series .

They first built an equally decimated series

\[ F_i = \left\{ F(n), n = 1, 2, \ldots, N \right\} \]

where \( F_i \) is Gaussian Brackets, and both \( \tau \) and \( i \) are integers, and then calculated the length of the curve \( L_i \) as

\[ L_i = \left( \sum_{j=1}^{\tau} F(i + j \tau) - F(i + (j-1) \tau) \right) \left( N - i \right) / \left( i - j \right) \]

The curve length for the time interval of \( \tau \) was defined as the average value over \( \tau \) sets of \( L_i \), and specified by \( \langle L(\tau) \rangle \).

If \( \langle L(\tau) \rangle \approx \tau^{-d} \) within the range \( \tau_{\min} \leq \tau \leq \tau_{\max} \), the curve is self-affine with fractal dimension \( D \) in this range. Fig. 6 illustrates their experimental result with the curve length \( \log \langle L(\tau) \rangle \) was plotted as a function of \( \log \tau \). It shows a straight line with a sharp bent at a certain \( \tau \), which is called characteristic (critical) time scale [12]. The characteristic time scale can be found in most analysis and it suggests there are two self-affine fractal features of FHR. Table 1 shows the result of application of the introduced algorithm to FHR signals at different gestational ages.

IV. FHR PATTERN CLASSIFICATION

Clinically, FHR recording may be interpreted as reassuring, nonreassuring and ominous [1].

**Reassuring patterns:** FHR patterns associated with fetal well-being and positive outcomes, which include “mild variable decelerations (less than 30 seconds in duration with rapid return to baseline), early decelerations (concurrent “mirror image” decreased with contraction) and accelerations without other changes” [1].

**Nonreassuring patterns:** also known as “warning” patterns, suggesting decreasing fetal capacity to cope with the stress of labor, and including “fetal tachycardia” defined as a baseline heart rate greater than 160 bpm, “fetal bradycardia” defined as baseline heart rate less than 120 bpm, “saltatory variability” which proceeds by leaps rather than by gradual transitions, “variable decelerations associated with a nonreassuring pattern”, and “late decelerations with preserved beat-to-beat variability” [1].

**Ominous Patterns:** patterns suggesting possible fetal compromise, comprising “persistent late decelerations with loss of beat-to-beat variability”, “non-reassuring variable decelerations associated with loss of beat-to-beat variability”, “prolonged severe bradycardia”, “sinusoidal pattern”, and “confirmed loss of beat-to-beat variability not associated with fetal quiescence, medications or severe prematurity” [1]. The illustration of these FHR patterns can be found in [1].

From the research guidelines for interpretation of EFM proposed in 1997 [35] to recent FHR pattern notification guidelines and suggested Management algorithm [3], the
technologies for FHR have been improved clinically. However there are still difficult cases for interpreting FHR pattern manually. To this end, diverse classifiers with different efficiencies have been designed to indicate fetal conditions, based on the input feature parameters. The classifiers can be categorized as linear and nonlinear, and with learning and without learning.

G. Georgoulas et al. [36] used two hidden Markov models (HMM) to classify FHR tracing of hypoxic and normal newborns. One HMM for the “normal” case and one for the “hypoxic” case were estimated by using k-means clustering algorithm. Each model is described by a continuous density hidden Markov model (CDHMM) along with outputs modeled by Gaussian density. Finally, they used the multifold cross validation scheme to evaluate the performance of their method and obtained very encouraging results.

Afterwards, the same author [14] utilized support vector machines (SVMs) to predict the risk of metabolic acidosis fetus. They extracted temporal and frequency features and morphological features from a data set of intrapartum recordings. After reducing feature space dimension, the features were put into SVMs for detection and validation. SVMs are a set of related supervised learning methods used for classification and regression. They belong to a family of generalized linear classifiers. A special property of SVMs is that they simultaneously minimize the empirical classification error and maximize the geometric margin.

Cao et al. [19] labeled each tracing with an outcome of “reassuring” or “nonreassuring” by using logistic regression models (Multivaviant Regression model) [37] and Wald’s tests of significance. Values of ROC area, sensitivity, specificity, and positive and negative prediction accuracy rates were computed for significant measures in the regression models.

Among various classifiers, artificial neural networks (ANNs), which are capable of non-linear statistical data modeling and decision making, are the most prevalent tools employed by researchers. G. Magenes et al. [38] used two neural classifier- multilayer perceptron (MLP) trained with adaptive back propagation algorithm and self-organizing map (Kohonen maps) to discriminate among fetal behavioral states and classify fetal pathological states. O. F. Romero et al. [39] applied a principal component analysis (PCA) to make the system independent of the baseline and employed MLP to analyze CTG records. G. Vasilos et al. [9] utilized self-organizing-map ANNs for classification of FHR tracings based on wavelet-transform for detection of acidemia.

V. SEVERAL ADVANCED ISSUES

Until recently, only a few studies of systematic mathematical FHR analysis have been conducted [19]. The model of FHR variability is rather simple, characterized by the auto-regressive model driven by stationary random noise [5, 28]. However, the FHR signal is a nonstationary process in which the dynamics at different gestational ages and in different cases are different, such as existence and absence of deceleration. So the modeling of variability is worth further study [40]. Other nonlinear characteristics, such as nonnormality, asymmetric cycles, bimodality, and sensitivity to initial condition, were proposed in the literature [17] and some new models have been developed to extract these nonlinear characteristics in various applications. The autoregressive conditional heteroscedastic (ARCH) model [41] found its applications in finance and was used to describe different variability in different periods. The threshold autoregressive (TAR) model [42] is capable of describing signals in different area of state-space. Although some of these nonlinear characteristics were studied qualitatively in FHR analyses, the latest methods such as ARCH and TAR were not utilized. Moreover, some studies [43] showed heartbeats are a point process whose structure has not been characterized by previous models. Inverse Gaussian model was proposed to describe this process [7]. We suggest studying those models and their validation methods for further in-depth data mining of structures in FHR signal.

Besides building appropriate signal models, there are two other issues. One is there are “so many” feature parameters that can be extracted from FHR and some of them have redundant information. Including all features directly into the classifier would increase the computational burden and might have less contribution to reducing misclassification rate. Hence it is worth to study the correlation of the features and select suitable subsets. The other issue is how to choose an effective classifier. This requires us to compare the performance of different classifiers, and conduct detailed analysis of the pattern distribution in the feature space.

VI. CONCLUSION

FHR monitoring is a standard clinical procedure widely used during pregnancy and labor. In this paper, we provided a comprehensive review of the developments in various aspects of computerized FHR monitoring research, including FHR signal modeling and representation, FHR features (including morphological features, temporal and frequency-domain parameters, and nonlinear parameters), and FHR pattern classification. We also discussed several methods that are under investigation and can possibly improve FHR modeling and classification. Much work is still needed to achieve accurate and automatic FHR monitoring and interpretation. To accomplish this goal, latest advances in various fields, such as signal processing and classification methods, and information from other sources, such as the PH value, should be utilized.

ACKNOWLEDGMENT

The Author would like to thank M.D. A. Kikuchi, Dr. H. Cao, and Dr. Georgoulas for allowing us to use their figures and data, and M.D. Downs for providing us the full-length report of his published article.