Fetal Electrocardiogram Extraction: A Case-Study in Non-linear System Identification

Mustafa Fanaswala, Student Member, IEEE

Abstract—The medical community has expressed a strong need for developing non-invasive techniques of monitoring the fetal electrocardiogram (ECG) which is vital for prenatal diagnostics. The dynamic environment of the human body and associated interfering signals lends itself to the adoption of an adaptive technique for fetal ECG extraction. The objective of this paper is to survey and compare several methods of fetal ECG extraction. In particular, real and synthetic ECG signals are used to compare the performance of Volterra filters to standard techniques like the Least-Mean Square (LMS) and the Normalized Least-Mean Square (NLMS) algorithms. In addition, a more recent blind source separation technique of Independent Component Analysis (ICA) is also examined. The poor performance of linear techniques such as the LMS and NLMS underscore the non-linear nature of the problem which is efficiently solved using Volterra filters.

Index Terms—Fetal ECG Extraction, Volterra Filters, Polynomial Filters, Non-linear System Identification.

I. INTRODUCTION

The electrocardiogram (ECG) is a simple, non-invasive medical test that reflects the electrical activity of the heart. The rhythmic contraction of the heart is a result of the depolarization and re-polarization cardiac phenomena that generates the characteristic ECG waveform. As a result of its electrical nature, the ECG signal travels over the conducting medium of the human body and can be measured by an array of electrodes placed at various sites. But in the case of a developing fetus, the fetal ECG is not easily accessible. Invasive techniques of monitoring the fetal ECG involve penetration of tissue and membrane which are highly sensitive to infection given the delicate intra-uterine environment. As a result, there is an increasing demand for viable non-invasive techniques for measuring the fetal electrocardiogram. Monitoring the fetal ECG permits the determination of cardiac parameters and, in turn, the determination of cardiac pathologies such as cardiac arrhythmia, acidosis and uterine contraction information [1].

Although alternative techniques such as the echocardiogram (which measures the mechanical activity of the heart) and the Doppler ultrasound have been successfully used in determining cardiac parameters, the fetal ECG offers a more direct visual technique for the interpretation of cardiac activity. However, the fetal ECG is overwhelmed by the maternal electrocardiogram and other interfering signals such as base-line wandering, power-line interference, the electromyogram and random electronic noise. Furthermore, the extraction of the fetal ECG by conventional fixed-weight filters is complicated due to the spectral overlap of the interfering signals and possible non-stationarity of the system being identified. Such a situation suggests the possible use of adaptive filtering techniques which is the focus of this paper.

The paper is structured as follows. In Section II, the problem is formally defined and formulated as an adaptive noise cancellation scheme. Section III provides background on Volterra filters and the advocated algorithm. A discussion of the simulation results are presented in Section IV. Finally, some concluding remarks are made in Section V.

II. THEORY

The problem of fetal ECG extraction can be formulated broadly either as an adaptive noise cancellation scheme or a blind source separation problem. The more recent approach of blind-source separation assumes the presence of a set of independently generated source signals which are linearly mixed in an unknown environment to produce another set of observed signals.

\[ \mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \]

where,

\[ \mathbf{s}(t) = [s_1(t), s_2(t), ..., s_m(t)]^T \]

and,

\[ \mathbf{x}(t) = [x_1(t), x_2(t), ..., x_m(t)]^T \]

such that \( \mathbf{s}(t) \) is the vector of source signals, \( \mathbf{x}(t) \) is a vector of observations and \( \mathbf{A} \) is an unknown nonsingular mixing matrix. Such a formulation presents a unique solution under fairly general conditions except for an arbitrary scaling of each signal source and possible permutation of the indices [2]. It is thus possible to find a de-mixing \( \mathbf{W} \) matrix defined as;

\[ \mathbf{y}(t) = \mathbf{W}\mathbf{x}(t), \]

where \( \mathbf{y}(t) \) are estimates of the source signals. Thus the observed signals are used to estimate the generating independent source signals. The underlying principle involved in the solution to such a formulation is called independent component analysis. ICA imposes statistical independence on all the individual components of the output source signals.

On the other hand, adaptive noise cancellation uses a ‘primary’ input containing the corrupted signal and a ‘reference’ input containing noise correlated in some unknown way with the primary noise. The reference input is adaptively filtered and subtracted from the primary input to obtain the signal estimate. Adaptive filtering before subtraction allows
for the treatment of inputs that are deterministic or stochastic, stationary or time variable [3]. The process of subtracting noise from a signal could result in an increase in the power of the output noise. However, control of the filtering and subtraction by an appropriate adaptive process can be effectively used to accomplish significant noise reduction. This paper concentrates on the latter approach of adaptive noise cancellation.

**Fig. 1. Biological model of ECG signals**

The biological model used in the formulation of a solution is depicted in Figure 1. MECC\(_{\text{source}}\) is the signal source that generates the maternal ECG which is not accessible because it lies deep beneath tissue. Consequently, an electrode placed in the thoracic region captures MECC\(_{t}\) which is considered to be a faithful representation of MECC\(_{\text{source}}\). In other words, the effects of the MECC\(_{\text{source}}\) signal travelling to the thorax is assumed negligible. However, the effect of the MECC\(_{\text{source}}\) signal travelling over the abdomen cannot be ignored because the signal travels through several biological interfaces which may generate reflections and distort the signal characteristics. The tissue, bone and other bodily fluids that the signal travels through is collectively modeled as an unknown system which transforms MECC\(_{t}\) to MECC\(_{a}\). The fetal heart denoted by FECG\(_{\text{source}}\) also generates an independent fetal ECG which is captured over the abdomen as FECG\(_{a}\). This is assumed to be the same as FECG\(_{\text{source}}\) after ignoring the negligible effects of travelling to the abdominal surface. At the abdomen, MECC\(_{a}\) and FECG\(_{a}\) are assumed to additively interfere and this composite signal CECG is recorded. Thus, we are presented with two observed signals; the predominantly maternal ECG signal and a composite maternal plus fetal signal. The aim of the adaptive noise cancellation scheme is to identify the biological system which transforms the maternal ECG so that it can be subtracted from the composite signal to separate the fetal ECG signal.

**Fig. 2. Adaptive Filter Structure for FECG Extraction**

An adaptive filter can be used for the system identification task as shown in Figure 2. The maternal ECG signal is used as input to the adaptive filter to identify the transforming system \(h\_1\) by attempting to remove all correlation between the input and desired signals. The composite signal is used as the desired signal. If the filter is successful in its attempt to remove the correlation between the input and desired signal, then subtracting the filter output from the composite signal should ideally provide the fetal ECG signal which is uncorrelated with the maternal ECG because they are generated by independent sources. This fact is further corroborated by the principle of orthogonality which states that the error is statistically uncorrelated with the input when the filter reaches its optimum.

**III. Volterra Filters**

**A. Background**

The selection of a suitable algorithm to adapt the filter is of primary importance to obtain acceptable performance. The choice of a suitable algorithm is made on the basis of desired complexity, speed of convergence and tracking behavior in the case of non-stationary systems. This paper examines conventional LMS based filters applied to the fECG extraction problem to demonstrate the need for non-linear techniques such as Volterra filters. In specific, the approach outlined in [4] is advocated.

The celebrated LMS and NLMS are stochastic gradient-based algorithms which estimate the gradient by its instantaneous value. An implicit assumption is that the output of the system being modeled can be expressed as a linear combination of the inputs. It was shown in [4] that the biological system that transforms the maternal ECG is in fact a non-linear system. The use of linear techniques result in poor performance which is demonstrated in Section IV.

A Volterra filter approximates the non-linear input-output relationship by a polynomial of sufficient order. Consequently, Volterra filters are also called polynomial filters. The Weirstrass existence theorem states that any smooth and continuous function can be approximated to an arbitrary
degree of accuracy provided that a polynomial of sufficient order is used. The use of polynomials in approximating functions is a popular choice because of the desirable properties they possess of being differentiable and continuous almost everywhere.

It is known that a memoryless nonlinear system can often be described in terms of an appropriate series expansion like the Taylor series [5]. A nonlinear system with memory can be represented by an extension of such an extension known as the Volterra series expansion that relates the input and output signals of the system as:

\[ y(t) = h_0 + \sum_{n=1}^{\infty} \int h_n (\tau_1) x(t - \tau_1) d\tau_1 + \sum_{n=2}^{\infty} \int \int h_n (\tau_1, \tau_2) x(t - \tau_1) x(t - \tau_2) d\tau_1 d\tau_2 + \ldots \]

Such a nonlinear system represented by a Volterra series expansion is completely characterized by the multidimensional functions \( h_p (t_1, \ldots, t_p) \) called the Volterra kernels. A suitable approximation to the input-output relationship is obtained by truncating the series expansion up to 'p' terms, where 'p' is called the order or degree of the Volterra series expansion.

An important observation that can be made is that if all the filter coefficients \( h_p (\tau_1, \ldots, \tau_p) \) are set to be zero, except \( h_1 \), then the expression reduces to that of the familiar linear filter. Thus, a Volterra filter can at least approach the performance of a linear filter. An important property of the Volterra which permits its discussion under generalized optimal filter theory is its linearity with respect to the kernel coefficients. It is evident from the input-output expression that the filter coefficients appear linearly while the nonlinearity of the expansions can be attributed to the cross products of the delayed input values. As a consequence of this property, Volterra series can be conceptualized as extensions of linear system models. In fact, the Volterra series can be thought of as a two stage filter; a preliminary transformation stage where the input undergoes a polynomial expansion to a desired degree and a secondary stage comprising of the expanded input as the new input. The secondary stage with the new input can be thought of as a linear filter. Consequently, traditional LMS based techniques can be used to adapt the filter coefficients.

Drawing a parallel from transform-domain filters, the polynomial expansion can be considered as a transformation, albeit a non-linear one, which hopefully offers the same advantages as linear transforms to the frequency domain. However, on further thought, it is evident that the non-linear polynomial transformation does not offer the same advantage of faster convergence speed. The philosophy behind transform-domain adaptive filters is that the input is expected to be less correlated in the transformed domain than the original input. A perfect decorrelation is offered by the Karhunen-Lo‘eve transform which cannot be performed in real-time. Nevertheless, the band-partitioning property of several transforms like the Discrete Cosine Transform (DCT) offer sufficient decorrelation such that the autocorrelation matrix of the transformed input is very close to being diagonal in nature. This effectively reduces the eigenvalue spread hence leading to faster convergence as every mode converges more or less uniformly. On the other hand, a polynomial expansion brings in more correlation because it simply introduces higher-order terms of the same input. This leads to a subsequent increase in eigenvalue spread resulting in slower convergence.

It is also worthwhile to note the huge increase in computational complexity that the polynomial expansion affords. A filter of length 3 results in 20 independent coefficients. The disadvantages of slower convergence and high computational cost might discourage the use of Volterra filters but there exist several efficient algorithms that address the convergence and complexity issues. The choice of Volterra filters in this application was particularly made because of the increase in performance that it achieved. The minimum error that even the optimum linearly constrained Weiner filter can obtain is not sufficient for the task at hand and this fact is used to justify the use of Volterra filters.

B. Proposed Algorithm

For simulation purposes, the approach outlined below is adopted. The presented filter is frame-based such that the input is broken down into non-overlapping frames of length M. The algorithm presented below is the same as described in [4]. We define \( d(n) \) as the abdominal composite signal (CECG) and \( x(n) \) as the thoracic maternal ECG signal (MECG). The adaptive filter aims to extract the fetal ECG signal \( e(n) \) represented by \( \text{fECG}_f \). The inputs to the adaptive algorithm for the \( i \)-th frame are the sequence \( d_i (m) \) and the vector sequence \( \mathbf{x}_i (m) = [x_i (0) \ x_i (1) \ldots x_i (J) (0)] \). The sequences \( x_i (m) = x(iM + m) \) and \( d_i (m) = d(iM + m) \). The vector sequence \( \mathbf{x}_i (m) \) is composed of the sequence \( x_i (m) \) and its J time-derivatives. The time derivatives help in incorporating the dynamics of the signal \( x(n) \). In vector notation, the inputs to the polynomial network are the vector \( \mathbf{d}_i \) and the matrix \( \mathbf{X}_i \), where
\[ \mathbf{d}_i(m) = [d_i(0) \ d_i(1) \ \ldots \ d_i(N-1)] \]
\[
\begin{bmatrix}
x_i(0) & \hat{x}_j(0) & \ldots & x_i^{(j)}(0) \\
x_i(1) & \hat{x}_j(1) & \ldots & x_i^{(j)}(1) \\
\vdots & \vdots & \ddots & \vdots \\
x_i(N-1) & \hat{x}_j(N-1) & \ldots & x_i^{(j)}(N-1)
\end{bmatrix}
\]

The input matrix \( \mathbf{X}_i \) for each block is passed through a polynomial expansion \( \Gamma(\mathbf{X}_i) \) resulting in a matrix \( \mathbf{P}_i \) in which each row is the corresponding expanded row of \( \mathbf{X}_i \). The adaptation of the filter coefficients is achieved by minimization of the mean-squared error objective criterion given by:

\[ \mathbf{w}_i^* = \arg \min_{\mathbf{w}_i} \| \mathbf{d}_i - \mathbf{P}_i \mathbf{w}_i \|^2, \]

where \( \mathbf{w}_i^* \) is the optimum coefficient vector. The weight vector is obtained by direct solution of the normal equations given by:

\[ \mathbf{w}_i^* = (\mathbf{P}_i^T \mathbf{P}_i)^{-1} \mathbf{P}_i^T \mathbf{d}_i \]

The approach taken above can be viewed as a least squares approach using a pulse-shaped window rather than an exponentially shaped window. The optimum filter coefficients can be used to estimate the transformed version of the maternal ECG (MECG) which can then be subtracted from the composite abdominal signal (CECG) to generate the filter error. This provides an estimate of the fetal ECG signal (fECG).

\[ \hat{\mathbf{e}}_i = \mathbf{d}_i - \mathbf{P}_i \mathbf{w}_i \]

IV. RESULTS

For simulation purposes, real world ECG data as well as synthetically generated data was used.

A. Real ECG Data

The DalsY database [6] is a publicly available database that contains real ECG data recorded from a pregnant woman. The data consists of recordings from eight different electrodes placed at different sites on the mother's body. Three channels comprise of recordings from the thoracic region while the remaining five channels are abdominal recordings. Two channels representing a predominantly maternal thoracic ECG and a composite signal with clear fetal contributions were chosen for simulations.

The need for non-linear filtering is evident from the results obtained using the LMS and NLMS algorithms which are shown in Figures 3 and 4.

It is clear from the last subplot in both these figures that the LMS based algorithms are not capable of cancelling the maternal component that is present in the abdominal signal. As a result, the fetal signal still has significant maternal portions of comparable signal strength. The NLMS does perform slightly better than the LMS as is evident from the lower strength that the maternal signal has in the extracted fECG signal. Nevertheless, such a signal is worthless for all practical purposes.

The failure of these algorithms suggests that either the excess error \( J_{\text{excess}} \) is large or that the minimum error achievable \( J_{\text{min}} \) itself is large. We can test this hypothesis by solving the problem using the optimum Weiner filter which serves as a benchmark for the minimum error achievable. The results of using the Wiener filter are shown in Figure 5. Not surprisingly, the linearly constrained Wiener filter is also unable to completely remove the maternal component. This result confirms our suspicion that \( J_{\text{min}} \) itself is unreasonably large which compels us to re-examine the assumption that the desired output can be expressed linear combination of the input samples.
Having justified the use of non-linear techniques, the results of using the proposed frame-based Volterra filter is shown in Figure 6. A polynomial of order 3 and the first two derivatives of the input were used in setting up the polynomial filter which is capable of completely eliminating the maternal component. In fact, even in areas where there is significant overlap between the maternal and fetal components, the Volterra filter is able to successfully extract the fetal signal. However, it is unreasonable to expect successful extraction when there is complete overlap between the maternal and fetal case due to the nature of the problem formulation. A pathological case when this occurs is also shown in Figure 6 between samples 1650 and 1700. A small occluded fetal peak can be seen at the leading edge of the maternal signal in the composite signal. The polynomial filter recognizes this ambiguity and produces a peak, albeit not the correct fECG signal, at that point. This can be seen in the last subplot where the second PQRST waveform is incompletely extracted.

For the sake of comparison, independent component analysis (ICA) was also performed on the data set using all the eight channels present in the DaIsY database [6]. The fetal signal extracted using ICA is shown in Figure 7 where it can be seen that even completely occluded fetal signals (between samples 150 and 200) are properly extracted. The results obtained are visually more impressive but the disadvantage of using an ICA approach is the fact that it requires more than two channels to satisfactorily isolate each independent component. The polynomial filter, on the other hand, allows use of a simple dual lead system. The use of ICA is a consequence of setting up the problem as a blind source separation scheme rather than as noise cancellation.

B. Synthetic ECG Data
A dynamic model for generating synthetic ECG signals [7] was used to create artificial data. Results on the real ECG data are only comparable visually because the actual fECG signal is not available for objective comparison. As a result, it would be useful to create synthetic data to objectively compare the results of the different techniques used. The dynamic model mentioned before was used to generate different signals based on varying fetal-maternal signal-to-noise ratios. The fmSNR is defined as [8]:

$$fmSNR = 10 \log_{10} \left( \frac{\sum_{n} e(n)^2}{\sum_{n} x(n)^2} \right) = 10 \log_{10} \left( \frac{\sum_{n} fECG_{n}(n)^2}{\sum_{n} mECG_{n}(n)^2} \right)$$

A system diagram depicting the generation of the synthetic ECG data [8] is shown in Figure 8. The artificial maternal ECG is passed through a transformation that involves a time-delay to simulate signal reflections from biological interfaces and a non-linear transformation using the inverse tangent function.
Now that the actual fetal signal is available for comparison, the results of the polynomial filter can be expressed in a more tangible manner. Figure 9 depicts the correlation results in comparison with the NLMS algorithm.

\[ q\text{SNR} = 10\log_{10}\left(\frac{\sum_n \hat{e}(n)^2}{\sum_n (e(n) - \hat{e}(n))^2}\right) \]

\[ = 10\log_{10}\left(\frac{\sum_n \text{(Estimate of fetal signal)}^2}{\sum_n \text{(Difference between actual and estimated)}^2}\right) \]

It is obvious that the NLMS cannot achieve the performance or the close match to the actual system that the polynomial filter provides. Also, the performance decrease of the polynomial filter is more graceful and allows for useful correlation even up to -20dB fetal-maternal SNR. The quality SNR for the extracted signals are presented in Figure 10.

**V. Conclusion**

This paper presented a case-study of a particular application of non-linear system identification using polynomial filters. The problem of fetal electrocardiogram extraction presents itself as a non-linear system identification problem when set up in an adaptive noise cancellation framework. The specific frame-based approach advocated in this study is shown to be sufficiently powerful in providing a close match to the actual system as evidenced by both visual and objective comparison. In comparison to traditional LMS based algorithms, the polynomial filter clearly presents a huge jump in performance. However, this increase in performance comes with the associated costs of added complexity in terms of exponential increase in number of inputs within the transformed space and increase in eigenvalue spread of this input. As a result, iterative LMS based techniques are expected to converge much slower. The frame-based algorithm, however, directly solves the normal equations for each block thus bypassing the convergence issue with a further increase in complexity because of the matrix inverse involved. Additionally, the tracking performance of the filter may suffer as a consequence of the block based approach if the system changes its behavior with time. Yet, the fact remains that biological systems are organic in nature and are not known to change their behavior very fast. Current efforts at exploiting the redundancy in the expanded polynomial terms may accomplish significant reduction in complexity and increase the convergence speed of adaptive Volterra filters.

**References**


