2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Multi-modal Framework for Fetal Heart Rate Estimation: Fusion of Low-SNR ECG and Inertial Sensors*

Arash Shokouhmand, *Student Member, IEEE*, Clarel Antoine, *M.D.,* Bruce K. Young, *M.D.,* Negar Tavassolian, *Senior Member, IEEE*

***Abstract*— This study presents a novel multi-modal framework for fetal heart rate extraction, which incorporates wearable seismo-cardiography (SCG), gyro-cardiography (GCG), and electrocardiography (ECG) readings from ten pregnant women. Firstly, a signal refinement method based on empirical mode decomposition (EMD) is proposed to extract the desired signal components associated with fetal heart rate (FHR). Afterwards, two techniques are developed to fuse the information from different modalities. The first method, named early fusion, is intended to combine the refined signals of different modalities through intra-modality fusion, inter-modality fusion, and FHR estimation. The other fusion approach, i.e., late fusion, includes FHR estimation and inter-modality FHR fusion. FHR values are estimated and compared with readings from a simultaneously-recorded cardiotocography (CTG) sensor. It is demonstrated that the best performance belongs to the late-fusion approach with 87.00% of positive percent agreement (PPA), 6.30% of absolute percent error (APE), and 10.55 beats-per-minute (BPM) of root-mean-square-error (RMSE).**

***Clinical Relevance*— The proposed framework allows for the continuous monitoring of the health status of the fetus in expectant women. The approach is accurate and cost-effective due to the use of advanced signal processing techniques and low-cost wearable sensors, respectively.**

I. Introduction

Stillbirth, defined as the death of a fetus after 24 weeks of gestational age, is a critical public health problem [1]. Statistics indicate that nearly 2.5 million stillbirths occur globally every year, motivating the need for proactive fetal monitoring to reduce fetal mortality [1]. An important vital sign to monitor for the wellbeing of a fetus is the fetal heart rate (FHR), which should fall within the range of 120-160 beats per minute (BPM) from the 24th week of gestation [2]. Techniques for continuous monitoring of FHR could help expectant mothers be informed of the health state of the fetus, and undergo necessary intervention procedures as soon as FHR goes beyond the standard range, posing serious health risks to the fetus.

Ultrasound cardiotocography (CTG) enables the continuous monitoring of fetal heart sound through auscultation, and provides obstetricians with information regarding FHR and uterine contractions [3]. CTG requires pregnant individuals to visit the clinic intermittently. Despite being a non-invasive modality, CTG has not conclusively been proven to be a safe technology as it irradiates the fetus with ultrasound frequencies [4]. Furthermore, it is prone to missing information due to fetal movement during monitoring [5]. Other than CTG, FHR monitoring is carried out through fetal electrocardiography (fECG) by placing low-noise ECG electrodes on the abdomen of the mother [5]. Consequently, a mixture of maternal ECG (mECG) and fECG is achieved which is named abdominal ECG (aECG). Over the past decade, a huge amount of research, primarily based on blind source separation (BSS) [6], [7] and adaptive filtering [8], [9], has been dedicated to extracting fECG imposed by noise components with signal-to-noise ratio (SNR), occasionally as low as -20 dB [10].

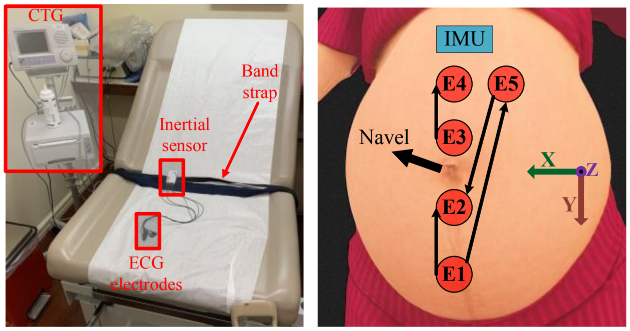
Recently, wearable seismo-cardiography (SCG) and gyro-cardiography (GCG) have been widely employed for wearable heart monitoring and disease detection. Our research group has reported on the application of SCG and GCG modalities on FHR estimation [11], where three inertial sensors were used to measure the abdominal movements caused by fetal heartbeats. Although the potentiality of FHR extraction was demonstrated, the reported PPA was 75.20% since the vibrations caused by fetal cardiac activity may not fully transfer to the abdomen, especially from the 28th to the 37th week of gestation due to the greasy layer called vernix, which dampens the vibrations produced by the heartbeat of fetus [12].

In this work, a multi-modal framework is developed where ECG and SCG/GCG modalities are employed in a fusion-based context for FHR extraction. It is assumed that data fusion among these modalities would enhance the appearance of fetal heartbeat components. The proposed framework employs *off-the-shelf* sensors, not necessarily designed for low-noise measurements. *To the best of our knowledge, this is the first study addressing the fusion between ECG and SCG/GCG modalities for fetal heart rate estimation.* The organization of the paper is as follows: In Section II, the experimental setup, the pre-processing, and the FHR extraction methodology are explained. Experimental results are discussed in Section III, and the paper is concluded in Section IV.

*This work was supported by the National Science Foundation (NSF) under award number 1855394.

A. Shokouhmand and N. Tavassolian are with the Department of Electrical and Computer Engineering at Stevens Institute of Technology, Hoboken, NJ 07030, USA. (e-mail: ashokouh@stevens.edu).

C. Antoine and B. K. Young are with the Department of Obstetrics and Gynecology, New York University Grossman School of Medicine, New York, NY 10016 USA.



(a) (b)

Fig. 1. (a) The experimental setup for data collection: an adjustable bed, the wearable sensors, CTG machine, and band straps. (b) ECG electrodes and IMU layout on the pregnant abdomen; X, Y, and Z show the axes for IMU recordings.

II. Methodology

In the following sub-sections, the experimental setup, signal processing, and sensor fusion techniques are explained in detail.

A. Experimental Setup and Measurement Protocols

Ten normal healthy pregnant women participated in this study and fetal heart rates were measured at the Department of Obstetrics and Gynecology at New York University Grossman School of Medicine (NYUGSM) after obtaining informed consent. The patients' experimental protocol was approved by the Institutional Review Board (IRB) of NYUGSM under study number i18-00564. The subjects' average (standard deviation) gestational age was 37.11 (2.56) weeks, with body mass index (BMI) falling within the range of 28.91 ± 4.46 Kg/m^2. Also, the average (standard deviation) age of the subjects was 32.8 (4.27) years old. Fig. 1 depicts the experimental setup.

The subjects were asked to lie down on an examination table, shown in Fig. 1 (a), for 5 minutes in a supine position. Then, an inertial measurement unit (IMU) from Shimmer Sensing was secured around the top part of the subject's abdomen using a band strap. The IMU sensor was used to record the rotational and linear vibrations of the abdomen. Furthermore, the electrodes of a four-lead ECG sensor were placed along the navel-to-chest direction as demonstrated in Fig. 1 (b). The layout of ECG electrodes (E1-E5) was designed such that they covered as many paths (purple arrows) on the abdomen as possible. According to Fig. 1 (b), the X, Y, and Z axes correspond to the directions along right-to-left, top-to-bottom, and dorso-ventral, respectively. All sensors recorded the data at a sampling rate of 512 Hz. In addition, an ultrasound CTG machine was employed to record the ground-truth FHR. After the measurement, the recorded data was transferred to a computer for further processing.

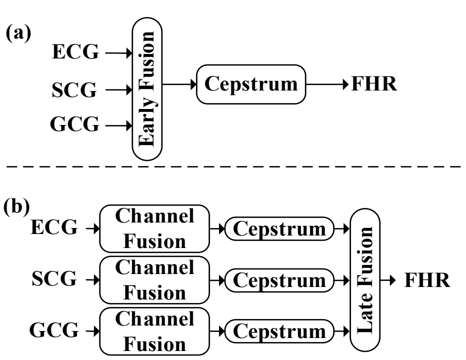


Fig. 2. Fusion-based FHR extraction flow graphs: (a) Early fusion, (b) Late fusion.

B. Signal Pre-processing and EMD Refinement

In order to remove the baseline wandering and motion artifacts from the recordings, ECG and SCG/GCG signals were filtered using a third-order zero-phase Butterworth band-pass filter with cut-off frequencies of 0.8-40 Hz and 0.8-25 Hz, respectively. The achieved signals are assumed to contain components associated with fECG, mECG, and noise, which should be separated from one another. On the other hand, mechanical activities, i.e., SCG/GCG, are delayed compared to electrical activities, i.e., ECG, in the heart, provoking the need for spectral analyses rather than time-domain methods. To this end, we designed an algorithm based on empirical mode decomposition (EMD), coined EMD refinement (EMDR), to keep those components that are highly correlated with fetal cardiac activities. EMDR starts with decomposing a signal into its intrinsic mode functions (IMFs), the top ones of which, carrying important information of the cardiac activity, are kept, whereas the remaining IMFs are discarded. In this work, the first six IMFs are kept as they comprise higher frequency components. Each IMF_i is scored by a factor of $\xi_i (i = 1,2,3,4, \dots)$, which is defined as follows:

$$\xi_i = \left(\sum_{k=\left\lfloor \frac{2N}{f_s} \right\rfloor}^{\left\lceil \frac{3N}{f_s} \right\rceil} |A_i[k]|^2\right) / \left(\sum_{k=1}^{N/2} |A_i[k]|^2\right), \qquad (1)$$

where $A_i[k]$ denotes the representation of i-th IMF in the spectral domain, N shows the signal length, and f_s implies the sampling frequency of the signal. The score for each IMF is determined based on the proportion of the information concentrated in a specific range to the total energy of the IMF. According to the Nyquist rate and assuming that the sample number $k = \frac{N}{2}$ in FFT domain corresponds to $f = \frac{f_s}{2}\ Hz$ in the frequency domain, $2N/f_s$ and $3N/f_s$ would imply 2 and 3 Hz in the frequency scale, respectively. This means that the IMF's are scored according to a frequency range, at which most of the energy of fetal FHR components, 120-160 BPM, exists. Thus, by scoring the IMF's, they can be ranked based on their information of fetal cardiac activity. Hence, considering the top scores as $\hat{\xi}_j$ ($\hat{\xi}_j > \hat{\xi}_{j+1}$) and their respective IMF's as $\widehat{IMF}_j$, the refined signal $\hat{x}$ would be reconstructed by using the top four IMF's out of six as follows:

$$\hat{x} = \hat{\xi}_1 \times \widehat{IMF}_1 + \hat{\xi}_2 \times \widehat{IMF}_2 + \hat{\xi}_3 \times \widehat{IMF}_3 + \hat{\xi}_4 \times \widehat{IMF}_4. \qquad (2)$$

Hence, EMDR could improve the signal components at a certain frequency range. EMDR is applied on every channel of ECG, SCG, and GCG, turning the signals more suitable for the multi-modal signal fusions illustrated in Fig. 2. As shown in this figure, two fusion techniques are designed: early-fusion (Fig. 2 (a)) and late-fusion (Fig. 2 (b)) as elucidated in the following sections.

D. Early-fusion FHR Estimation

In order to implement a meaningful fusion among the modalities, we employ the spectral contents of the signal since the nature of ECG, SCG, and GCG are different from one another in the time domain. Furthermore, having considered the signal refinement proposed above, the enhanced frequency components within the range of 2-3 Hz are deemed to contribute to a constructive effect in the spectral domain. The time-frequency (TF) representation of each channel is calculated by applying a Morse continuous wavelet transform (CWT) as defined below:

$$\Upsilon_{P,\upsilon}(\omega) = U(\omega) a_{P,\upsilon} \omega^{\frac{P^2}{\upsilon}} \times \exp\left(-\omega^{\upsilon}\right), \quad (3)$$

where $U(\omega)$, P^2, υ, and $a_{P,\upsilon}$ characterize the unit step function, time-bandwidth product, symmetry of the Morse wavelet, and normalizing constants, respectively. In this study, the values of υ and P are set to 4 and 120, respectively. Once the refined signals from each modality are transformed into their TF representation, the values of each TF are normalized to a zero-mean unit-variance distribution, i.e., $\sim\mathcal{N}(0,1)$. Then, the normalized TF's are averaged within their respective modalities, normalized again, and averaged across the three modalities. Indeed, the fusion operator generates a single TF characteristic that represents the combined information from the three modalities with an emphasis on the components within 2-3 Hz, which is relevant to fetal heart activities. An inverse continuous wavelet transform (ICWT) is then performed on the fused TF to convert it back to time domain. Subsequently, to obtain the FHR values, cepstrum is performed on the signal to calculate the lag at which the time-domain signal appears to have a higher energy. As the FHR value varies within the range of 110-180 BPM throughout the entire experiment, the desired peak is to be detected within the corresponding lag values, i.e., 333-545 ms. Hence, FHR is calculated as follows:

$$FHR\ (BPM) = \frac{60(second)}{lag\ (second)}. \quad (4)$$

E. Late-fusion FHR Estimation

The late-fusion flow graph is shown in Fig. 2 (b). According to this structure, each modality undergoes a channel fusion procedure as described above. As such, ECG, SCG, and GCG channels are fused within their respective modalities separately. Once the fusion at each modality is performed, three TF representations corresponding to ECG, SCG, and GCG are obtained. Unlike the early-fusion technique, the TF associated with each modality is converted back to the time domain with no data fusion at this stage. Then, cepstrum is applied on each time-domain signal, and FHR values are estimated from each modality separately. Thus, three estimates of FHR are acquired through SCG, GCG, and ECG channels. Data fusion is then conducted on the three FHR estimates by averaging the values. In fact, the three modalities are assumed to introduce some amount of error in the FHR estimate values. Yet, late-fusion is meant to reduce the total error by moderating the FHR discordance among the sensors. This will be further discussed in the experimental results.

TABLE I. RMSE OF ESTMATION BY MODALITY-ONLY AND FUSION-BASED METHODS

Subject	SCG	GCG	ECG	Early fusion	Late fusion
1	12.07	11.69	12.65	12.37	**9.52**
2	16.25	17.15	15.85	16.04	**14.70**
3	13.10	12.01	14.22	12.01	**11.30**
4	12.07	11.86	11.44	12.81	**9.22**
5	17.31	**12.34**	13.64	14.16	12.54
6	8.34	11.26	10.11	9.77	**7.34**
7	14.62	14.20	14.91	16.43	**12.65**
8	12.75	12.38	12.69	12.14	**10.00**
9	12.74	12.54	9.85	**9.42**	9.68
10	10.54	10.56	10.59	9.52	**8.57**
Mean ±std. (BPM)	12.98 ±2.47	12.60 ±1.76	12.60 ±1.96	12.47 ±2.38	**10.55 ±2.10**

III. Experimental Results and Discussion

In this section, the experimental results are presented and a comparison between the proposed techniques and those in literature is made.

A. Performance Evaluation

In order to evaluate the performance of the proposed methods, the signals are segmented into 10-second frames, from each of which FHR values are estimated. The estimate values are then compared with the ground-truth values, i.e., CTG readings, through three metrics called positive percent average (PPA), i.e., the percentage of time the proposed method generated a valid FHR within 10% of the readings from CTG, absolute percentage error (APE), and root-mean-square error (RMSE). Table I summarizes the RMSE results for fusion-based and single-modality scenarios for the ten subjects, where single-modalities include signal refinement and channel fusion. According to Table I, late-fusion suggests the best performance among all by 10.55±2.10 BPM of RMSE, whereas early-fusion is reported by 12.47±2.38 BPM, indicating superior accuracy to the single-modality scenarios such as SCG, GCG, and ECG with 12.98±2.47, 12.60±1.76, and 12.60±1.96 BPM, respectively, where GCG introduces the least error. The difference in performance between early and late fusions results from the potentially destructive effect caused by the complex values of TF while fusing the sensors in early-fusion. Furthermore, for subjects 5 and 9, the best RMSE's were achieved by the GCG modality and early-fusion, respectively, although late-fusion performs more robustly in other cases. Fig. 3 makes a comparison of the APE

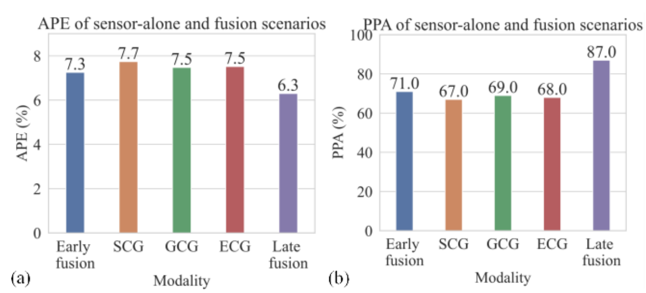


Fig. 3. Performance of the proposed method and the sensor-alone scenarios: (a) APE, (b) PPA.

TABLE II. COMPARISON WITH OTHER WORKS IN LITERATURE

Modality	PPA (%)	RMSE (BPM)	CI (low, high)	Reference
Low SNR ECG/SCG/GCG	**87.00±4.81**	**10.55±2.10**	**-1.9, -0.19**	**Proposed method**
SCG/GCG	75.20±11.81	11.40±2.17	N/A	[11]
fECG	83.40±15.40	4.80±2.00	N/A	[13
PCG/fECG	N/A	N/A	-8.84, 8.24	[14]

and PPA metrics. As illustrated in Fig. 3 (a) and (b), the late-fusion model outperforms early-fusion as well as single-modality settings by a margin of 1.00% and 17.00% in terms of APE and PPA, respectively. Unlike the RMSE values mentioned in Table I, PPA=71.00% and APE= 7.30% of early-fusion suggest better performances compared to the single-modality scenarios. The worst performance belongs to the SCG modality with 12.98 BPM, 7.70%, and 67.00% of RMSE, APE, and PPA, respectively. In summary, late-fusion offers a more robust and reliable estimation of FHR values.

B. Comparison with Other Methods

As the proposed framework leverages both abdominal ECG and SCG/GCG modalities, where CTG provides the ground-truth value, it is compared with the three most similar wearable-based works in the literature. In [11], SCG/GCG are employed to extract the FHR, where CTG is used as the reference. In [13], FHR extraction is performed using a small AN24 sensor from Monica Healthcare, which was used to record fECG. Furthermore, a sensor fusion of phono-cardiography (PCG) and fECG addresses FHR extraction in [14], with which the proposed method is compared. These works were selected for comparison since they leverage a partially-similar setup to ours, i.e., inertial sensors, abdominal ECG sensors, and a sensor fusion algorithm, respectively. Table II reports on the performance comparison between the proposed method and other works. As seen in Table II, the proposed late-fusion framework outperforms the other methods in the literature in terms of PPA by a large margin of 4.00%, whereas in terms of RMSE, it suggests weaker performance than the fECG-based method (10.55 vs. 4.80 BPM). This weaker performance stems from using commercial off-the-shelf ECG sensors in our setting. Furthermore, despite employing a single-IMU setting, the proposed method could estimate the FHR values more accurately than the triple-IMU-based framework in [11] (RMSE: 10.55 vs. 11.40 and PPA: 87.00% vs. 75.20%). The fourth column of Table II represents 95% confidence interval (CI) (lower bound, higher bound) for the difference between the reference and the estimates of FHR. This metric is only reported for the fusion algorithm in [14]. Comparing the proposed method and [14], it is concluded that most of the differences between the reference and the estimates of FHR in our method lie within a smaller range around zero, i.e., (-1.9, -0.19), in comparison to [14], i.e., (-8.84, 8.24), implying higher confidence by the proposed method.

IV. CONCLUSION

This study addresses the fusion of fECG and inertial sensors to monitor FHR for the first time. The proposed framework benefits from an EMD-based signal refinement, which enhances fetal cardiac activity components in both ECG and SCG/GCG readings. Furthermore, two sensor fusion methods are designed to improve the accuracy of FHR estimation. A PPA of 87.00% demonstrates the efficiency of our approach for FHR extraction. Future studies include source separation methods to cancel the maternal components from both ECG and inertial readings. Furthermore, channel/sub-signal scoring will be added to each modality to select the channels/sub-signals indicating higher SNR values within the range of 2-3 Hz. Moreover, other patterns for ECG placement will be investigated, where some of them might potentially contribute to high-SNR fECG components.

REFERENCES

[1] W. H. O. (WHO), "Stillbirth," 2018. [Online]. Available: https://www.who.int/health-topics/stillbirth.

[2] S. P. Von Steinburg *et al.*, "What is the 'normal' fetal heart rate?," *PeerJ*, vol. 1, p. e82, 2013.

[3] R. G. Kennedy, "Electronic fetal heart rate monitoring: retrospective reflections on a twentieth-century technology," *J. R. Soc. Med.*, vol. 91, no. 5, pp. 244–250, 1998.

[4] S. B. Barnett and D. Maulik, "Guidelines and recommendations for safe use of Doppler ultrasound in perinatal applications," *J. Matern. Fetal. Med.*, vol. 10, no. 2, pp. 75–84, 2001.

[5] J. Behar, F. Andreotti, S. Zaunseder, J. Oster, and G. D. Clifford, "A practical guide to non-invasive foetal electrocardiogram extraction and analysis," *Physiol. Meas.*, vol. 37, no. 5, p. R1, 2016.

[6] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 5, pp. 567–572, 2000.

[7] P. P. Kanjilal, S. Palit, and G. Saha, "Fetal ECG extraction from single-channel maternal ECG using singular value decomposition," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 1, pp. 51–59, 1997.

[8] R. Martinek *et al.*, "Non-invasive fetal monitoring: A maternal surface ECG electrode placement-based novel approach for optimization of adaptive filter control parameters using the LMS and RLS algorithms," *Sensors*, vol. 17, no. 5, p. 1154, 2017.

[9] E. R. Ferrara and B. Widraw, "Fetal electrocardiogram enhancement by time-sequenced adaptive filtering," *IEEE Trans. Biomed. Eng.*, no. 6, pp. 458–460, 1982.

[10] E. Fotiadou and R. Vullings, "Multi-Channel Fetal ECG Denoising With Deep Convolutional Neural Networks," *Front. Pediatr.*, vol. 8, 2020.

[11] C. Yang, C. Antoine, B. K. Young, and N. Tavassolian, "A Pilot Study on Fetal Heart Rate Extraction from Wearable Abdominal Inertial Sensors," *IEEE Sens. J.*, vol. 19, no. 22, pp. 10773–10781, 2019.

[12] E. M. Graatsma, B. C. Jacod, L. A. J. Van Egmond, E. J. H. Mulder, and G. H. A. Visser, "Fetal electrocardiography: feasibility of long-term fetal heart rate recordings," *BJOG An Int. J. Obstet. Gynaecol.*, vol. 116, no. 2, pp. 334–338, 2009.

[13] W. R. Cohen *et al.*, "Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes," *Acta Obstet. Gynecol. Scand.*, vol. 91, no. 11, pp. 1306–1313, 2012.

[14] M. Mhajna *et al.*, "Wireless, remote solution for home fetal and maternal heart rate monitoring," *Am. J. Obstet. Gynecol. MFM*, vol. 2, no. 2, p. 100101, 2020.