

使用信噪比正则化 LightGBM 的腹部源胎儿 心电信号提取方法

韩 亮^{1,2}, 蔡文涛¹, 蒲秀娟^{1,2}, 罗统军¹, 黄 谦¹

(1. 重庆大学微电子与通信工程学院, 重庆 400044; 2. 生物感知与智能信息处理重庆市重点实验室, 重庆 400044)

摘 要: 胎儿心电信号(Fetal ElectroCardioGram, FECG)能反映胎儿健康状况. 但是, 由于其信噪比相对较低, FECG 仍未能在临床上得到广泛应用. 如何有效提取高质量的 FECG 仍是一个巨大挑战. 为此, 本文提出一种使用信噪比正则化 LightGBM(Light Gradient Boosting Machine)模型的 FECG 提取方法. 针对原始母体腹壁混合信号, 首先使用传统滤波方法进行噪声抑制, 然后再使用快速独立成分分析(Fast Independent Component Analysis, FastICA)从中分离得到母体心电信号(Maternal ElectroCardioGram, MECG)估计和 FECG 估计, FECG 估计中残留的 MECG 成分是 MECG 的一种非线性变换. 改进传统 LightGBM 模型, 在目标函数中增加 FECG 的基于互相关系数的信噪比作为正则项, 构建信噪比正则化 LightGBM 模型, 并使用该模型拟合这一非线性变换. 将 MECG 估计经由所拟合的非线性变换得到 MECG 成分的最优估计, 并将其抑制, 提取得到高质量的 FECG. 采用真实腹部源心电信号数据集进行实验, 结果显示本文提出的方法的灵敏度为 99.9%, 阳性预测值为 99.1%, F_1 分数为 99.5%, 基于互相关系数与基于特征值分析的信噪比分别为 6.0 dB 和 6.1 dB. 实验结果表明, 本文提出的方法是有效的且具有更好的性能.

关键词: 胎儿心电信号; FastICA; 基于互相关系数的信噪比; 正则化; LightGBM

中图分类号: TN911.72; TH701

文献标识码: A

文章编号: 0372-2112(2022)08-1793-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20211472

Method on Abdominal Electrode-Sourced FECG Extraction Utilizing the Improved LightGBM with Signal-to-Noise Ratio Regularization

HAN Liang^{1,2}, CAI Wen-tao¹, PU Xiu-juan^{1,2}, LUO Tong-jun¹, HUANG Qian¹

(1. School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China;

2. Chongqing Key Laboratory of Bio-perception & Intelligent Information Processing, Chongqing 400044, China)

Abstract: The fetal electrocardiogram(FECG) reflects health status of fetus. However, the FECG has not been widely used in clinical practice due to its relatively low signal-to-noise ratio(SNR). How to effectively extract high-quality fetal ECG signal remains a great challenge. In this paper, FECG extraction method utilizing improved LightGBM(light gradient boosting machine) with SNR regularization is proposed. Firstly, the raw maternal abdominal mixed signals are denoised by conventional filtering method. Then, FastICA(fast independent component analysis) is used to separate the maternal electrocardiogram(MECG) estimation and FECG estimation. The residual MECG component in FECG estimation is non-linear transform of MECG and the non-linear transform is fitted by the improved LightGBM with SNR regularization, which is constructed by adding a regularization term, the SNR of extracted FECG, to the objective function of conventional LightGBM. The SNR is estimated using cross correlation. By MECG estimation undergoing the fitted non-linear transform, the residual MECG component in FECG estimation is obtained. At last, the high-quality FECG is extracted by suppressing the estimated MECG component. The real data are adopted to verify the proposed FECG extraction method. The sensitivity, positive predictive value and F_1 score of the proposed FECG extraction method are 99.9%, 99.1%, and 99.5%, respectively, and the SNR based on cross correlation and singular value decomposition are 6.0 dB and 6.1 dB, respectively. The experiment results indicate that the proposed method is effective and has better performance.

Key words: fetal electrocardiogram(FECG); FastICA; the SNR based on cross correlation; regularization; LightGBM

1 引言

胎儿电子监护能反映胎儿健康状况,便于各类胎儿疾病的早期诊断与及时治疗,从而降低围产儿的发病率^[1].目前在临床中使用最多的胎儿电子监护技术是胎心宫缩图^[2],但利用胎心宫缩图难以获取可靠的瞬时胎儿心率变异信息^[3].胎儿心电信号(Fetal ElectroCardioGram, FECC)能无创地对胎儿健康状况进行长期连续监护,并实时提供高精度的胎儿心率信息,具有广阔应用前景.

非侵入式 FECC 采集方法置电极于孕妇体表,能长期连续监护胎儿健康状况^[4].但采用该方法从母体腹壁获得的 FECC 受到以母体心电图(Maternal ElectroCardioGram, MECG)成分为主的各种干扰的污染^[5].FECC 和 MECG 成分在时域和频域都大部分重叠在一起,且 MECG 成分的时域幅度和频域能量均远大于 FECC,因此,如何从母体腹壁混合信号中有效抑制 MECG 成分是一个巨大挑战.

基于非侵入式采集方法的 FECC 信号提取方法分为组合源方法和腹部源方法^[6].组合源方法需置电极于母体胸部和腹部,主要包括自适应滤波^[7]、人工神经网络^[8]、支持向量机^[9]等.腹部源方法仅需置电极于母体腹部,便于临床应用,主要包括模版相减法^[10]、卡尔曼滤波^[11]、小波变换^[12]、盲源分离^[13]等.

应用于 FECC 提取的机器学习方法通常使用均方误差(Mean Squared Error, MSE)损失函数作为目标函数,将输出信号与目标信号差值作为 FECC.但是, MSE 损失函数仅表示输出信号与目标信号之间的均方误差最小,并不意味着 FECC 的质量最好. LightGBM^[14,15] 是一种基于梯度提升决策树的集成学习模型,常用于分类和回归等机器学习任务. Outram^[16] 将互相关分析用于估计 FECC 的信噪比,其信噪比越高,说明 FECC 质量越好.为此,本文提出一种使用信噪比正则化 LightGBM 的 FECC 提取方法.针对原始母体腹壁混合信号,先使用传统滤波方法进行噪声抑制;再使用 FastICA 从中分离得到 MECG 估计和 FECC 估计;构建信噪比正则化 LightGBM 模型并使用该模型估计残留于 FECC 估计中的 MECG 成分,将其抑制后最终获得高质量的 FECC.

2 腹部源 FECC 提取基本原理

2.1 FECC 采集

基于非侵入式采集方法的 FECC 提取方法分为组合源方法和腹部源方法,如图 1 所示.

其中,腹部源方法仅需采集母体腹壁混合信号,主要分为两类:一类是模版相减法,针对心电信号在时域的特点,利用母体腹壁混合信号的相关信息生成其

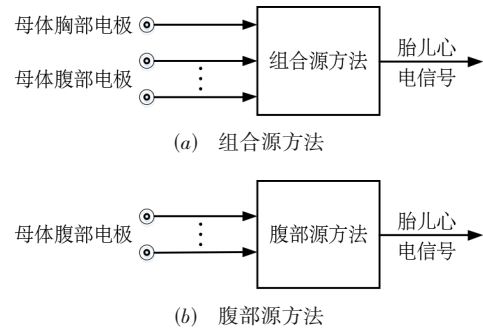


图 1 FECC 采集示意图

中包含的 MECG 成分模板,然后将其抑制;另一类是盲源分离法,该方法基于空间技术,通过使用相关源信号的空间分布信息提取 FECC.针对原始母体腹壁混合信号,运用这两大类腹部源方法能够提取 FECC,但其中仍残留有 MECG 成分等噪声.

2.2 FECC 提取原理

为从母体腹壁混合信号中提取高质量的 FECC,本文提出一种使用信噪比正则化 LightGBM 的 FECC 提取方法,如图 2 所示.

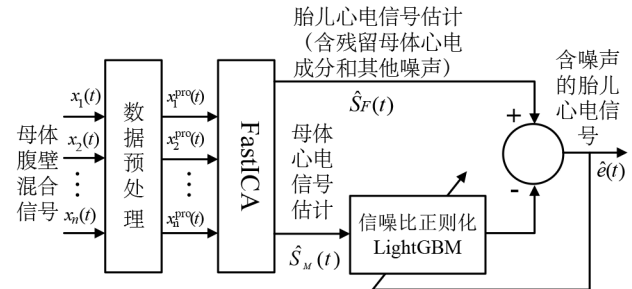


图 2 使用信噪比正则化 LightGBM 的 FECC 提取方法

置电极于孕妇腹壁,采集得到母体腹壁混合信号 $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$. $\mathbf{x}(t)$ 包含有 MECG 成分、FECC 成分和其他噪声.首先进行数据预处理,使用传统滤波方法^[17]抑制噪声,得到 $\mathbf{x}^{\text{pro}}(t) = (x_1^{\text{pro}}(t), x_2^{\text{pro}}(t), \dots, x_n^{\text{pro}}(t))^T$.采用线性瞬时混合模型对 $\mathbf{x}^{\text{pro}}(t)$ 建模^[13]:

$$\mathbf{x}^{\text{pro}}(t) = \mathbf{A} \cdot \mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

其中, $\mathbf{s}(t) = (s_1(t), s_2(t), \dots, s_k(t))^T$ 表示 k 个独立源信号, $\mathbf{n}(t)$ 表示残留的少量噪声, \mathbf{A} 表示混合矩阵.

然后,使用 FastICA 从 $\mathbf{x}^{\text{pro}}(t)$ 分离得到 MECG 估计 $\hat{S}_M(t)$ 和 FECC 估计 $\hat{S}_F(t)$. $\hat{S}_F(t)$ 包含胎儿心电信号 $d(t)$ 、残留的 MECG 成分 $m(t)$ 和噪声 $\eta(t)$ 三种成分:

$$\hat{S}_F(t) = d(t) + m(t) + \eta(t) \quad (2)$$

其中, $m(t)$ 是 $\hat{S}_F(t)$ 中最主要的干扰.众所周知, MECG 成分是 MECG 的一种非线性变换^[8]:

$$m(t) = f(S_M(t)) \quad (3)$$

其中, $f(\cdot)$ 表示非线性变换, $S_M(t)$ 为 MECG.

由式(3)可知,如果能得到 MECG 成分与 MECG 之间的非线性变换 $f(\cdot)$ 的最优估计 $\hat{f}(\cdot)$, 则可以利用 MECG 估计 $\hat{S}_M(t)$ 得到 FECG 估计 $\hat{S}_F(t)$ 中的 MECG 成分 $m(t)$ 的最优估计 $\hat{m}(t)$, 如式(4)所示:

$$\hat{m}(t) = \hat{f}(\hat{S}_M(t)) \quad (4)$$

进而将其抑制,得到含噪声的 FECG 的最优估计:

$$\hat{e}(t) = \hat{S}_F(t) - \hat{m}(t) = \hat{S}_F(t) - \hat{f}(\hat{S}_M(t)) \quad (5)$$

以 MECG 估计 $\hat{S}_M(t)$ 为输入信号,以 FECG 估计 $\hat{S}_F(t)$ 为目标信号,以 MSE 损失函数为目标函数,使用人工神经网络^[8]与支持向量机^[9]可以得到 MECG 成分与 MECG 之间的非线性变换 $f(\cdot)$ 的最优估计 $\hat{f}(\cdot)$, 此时目标信号 $\hat{S}_F(t)$ 与输出信号 $\hat{f}(\hat{S}_M(t))$ 之间的差值即为 FECG. 但是,上述 FECG 提取方法存在局限性,由式(5)可知, MSE 损失函数最小仅表示目标信号 $\hat{S}_F(t)$ 与输出信号 $\hat{f}(\hat{S}_M(t))$ 之间的均方误差最小,并不意味着提取得到的 FECG 的质量最好. Outram^[16]使用 FECG 的信噪比来评价 FECG 质量,信噪比越高表示 FECG 质量越好. 因此,本文构建信噪比正则化 LightGBM 模型估计 MECG 成分与 MECG 之间的非线性变换 $f(\cdot)$.

2.3 信噪比正则化 LightGBM 模型

LightGBM^[14]是对梯度提升决策树的高效实现框架,在不损失预测准确率的前提下具有更快的训练速度、更高的效率和更低的内存消耗. LightGBM 用于解决回归问题,通常采用 MSE 损失函数作为目标函数,其定义如下:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|y_i - f(\mathbf{x}_i; \theta)\|_2^2 \quad (6)$$

其中, $L(\theta)$ 表示损失函数, \mathbf{x}_i 表示输入向量, y_i 表示相应的目标值, $f(\mathbf{x}_i; \theta)$ 表示在输入向量 \mathbf{x}_i 和网络参数 θ 条件下的 LightGBM 网络的输出.

由 2.2 节的分析可知: MSE 损失函数不能直接表征提取得到的 FECG 的质量. 为此,本文在传统 LightGBM 模型的目标函数中增加 FECG 的信噪比作为正则项,构建信噪比正则化 LightGBM 模型,其目标函数优化为

$$\min_{\theta} J(\theta) = L(\theta) + \lambda R(\theta) \quad (7)$$

其中, $L(\theta)$ 为损失函数,本文选择 MSE 损失函数,其定义如式(6)所示; $R(\theta)$ 为正则项,表示 FECG 的信噪比,其定义如式(8),即

$$R(\theta) = \text{SNR}(y_i - f(\mathbf{x}_i; \theta)) \quad (8)$$

其中, $f(\mathbf{x}_i; \theta)$ 表示在输入向量 \mathbf{x}_i 和网络参数 θ 条件下

的 LightGBM 网络的输出, $(y_i - f(\mathbf{x}_i; \theta))$ 表示 FECG 估计. λ 为加权系数,用于控制损失项和正则项的权重.

根据式(7)和式(8),信噪比正则化 LightGBM 模型优化后的目标函数为

$$\min_{\theta} J(\theta) = L(\theta) + \lambda \times \text{SNR}(y_i - f(\mathbf{x}_i; \theta)) \quad (9)$$

由于胎儿心电信号信噪比越高表示 FECG 质量越好,所以 λ 取负值. 改进后的目标函数的最小化同时反映了 MECG 成分的最优估计和 FECG 的最佳质量.

2.4 FECG 信噪比

FECG 信噪比的计算方法如下^[18].

步骤 1 使用传统 QRS 检测算法^[19]确定 FECG 的 R 峰位置. 以 R 峰位置为基准将 FECG 分为等长的 L 段, 每段均含有一个完整的 QRS 波群. 构建一个 K 行 L 列的矩阵 $M_{K \times L}$, 其中, K 为每段采样点数.

步骤 2 基于互相关系数估计的信噪比如下:

$$\text{SNR}_{\text{cor}} = 10 \times \log \sqrt{\frac{\eta}{1-\eta}} \quad (10)$$

其中, $\eta = \frac{2}{L(L-1)} \sum_{i=0}^{L-2} \sum_{j=i+1}^{L-1} M(i)^T M(j)$.

步骤 3 基于特征值分析估计的信噪比如下:

$$\text{SNR}_{\text{svd}} = 10 \times \log \sqrt{\frac{\lambda_1^2}{\sum_{i=2}^L \lambda_i^2}} \quad (11)$$

其中, λ_i 是矩阵 $M^T M$ 的特征值, λ_1 是矩阵 $M^T M$ 的最大特征值.

SNR_{cor} 和 SNR_{svd} 越高,说明 FECG 质量越高. 在高信噪比条件下, SNR_{cor} 和 SNR_{svd} 都具有很高的估计精度,但是,在低信噪比条件下, SNR_{cor} 的估计精度要明显优于 SNR_{svd} , 这是由于 SNR_{cor} 估计方法不依赖信号成分的精确估计^[16]. 因此,本文选择 SNR_{cor} 作为信噪比正则化 LightGBM 模型的正则项.

3 FECG 提取方法

步骤 1 首先使用传统滤波方法^[17]对原始母体腹壁混合信号 $\mathbf{x}(t)$ 进行噪声抑制,得到输入信号 $\mathbf{x}^{\text{pro}}(t) = (x_1^{\text{pro}}(t), x_2^{\text{pro}}(t), \dots, x_k^{\text{pro}}(t))^T$.

步骤 2 使用 FastICA 从 $\mathbf{x}^{\text{pro}}(t)$ 中分离得到 MECG 估计 $\hat{S}_M(t)$ 和 FECG 估计 $\hat{S}_F(t)$.

步骤 3 构建信噪比正则化 LightGBM 模型. 考虑到 MECG 传导至腹壁电极时可能包含多条路径^[8], 以 MECG 估计 $\hat{S}_M(t)$ 及其 J 维时间导数作为模型的输入向量 \hat{S}_M , 以 FECG 估计 $\hat{S}_F(t)$ 作为模型的目标向量 \hat{S}_F , 具体如下:

$$\hat{S}_F = [\hat{S}_F(1) \hat{S}_F(2) \dots \hat{S}_F(k)]^T \quad (12)$$

$$\hat{\mathbf{S}}_M = [\hat{S}_M(1) \hat{S}_M(2) \cdots \hat{S}_M(k)]^T$$

$$= \begin{bmatrix} \hat{S}_M(1) & \hat{S}_M(1) & \cdots & \hat{S}_M^{(j)}(1) \\ \hat{S}_M(2) & \hat{S}_M(2) & \cdots & \hat{S}_M^{(j)}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{S}_M(k) & \hat{S}_M(k) & \cdots & \hat{S}_M^{(j)}(k) \end{bmatrix} \quad (13)$$

使用输入向量 $\hat{\mathbf{S}}_M$ 和目标向量 $\hat{\mathbf{S}}_F$, 训练 2.3 节提出的信噪比正则化 LightGBM 模型, 其信噪比正则项为基于互相关系数的信噪比 SNR_{cor} , 得到 MECG 成分与 MECG 之间的非线性变换 $f(\cdot)$ 的最优估计 $\hat{f}(\cdot)$.

步骤 4 再次将 MECG 估计 $\hat{S}_M(t)$ 及其 J 维时间导数构成的输入向量 $\hat{\mathbf{S}}_M$ 输入已训练好的信噪比正则化 LightGBM 模型, 得到 $\hat{S}_F(t)$ 中残留的 MECG 成分的最优估计 $\hat{m}(t)$, 即

$$\hat{m}(t) = \hat{f}(\hat{S}_M(t)) \quad (14)$$

步骤 5 根据式 (5), 从 FECG 估计 $\hat{S}_F(t)$ 中减去 MECG 成分的最优估计 $\hat{m}(t)$, 提取得到高质量的胎儿心电信号 $\hat{e}(t)$:

$$\hat{e}(t) = \hat{S}_F(t) - \hat{m}(t) = \hat{S}_F(t) - \hat{f}(\hat{S}_M(t)) \quad (15)$$

4 实验结果与分析

4.1 实验数据

本文使用 PhysioNet 2013 挑战赛^[20]SET-A 中的 7 组真实腹部源心电图数据进行实验, 每组数据中包含 4 路腹壁混合信号, 其采样频率为 1 000 Hz, 采集时间为 1 min. 在 A15 中截取一段长为 2 000 个采样点 (40~42 s) 的数据绘制于图 3 (纵坐标表示相对幅度, 横坐标表示采样点).

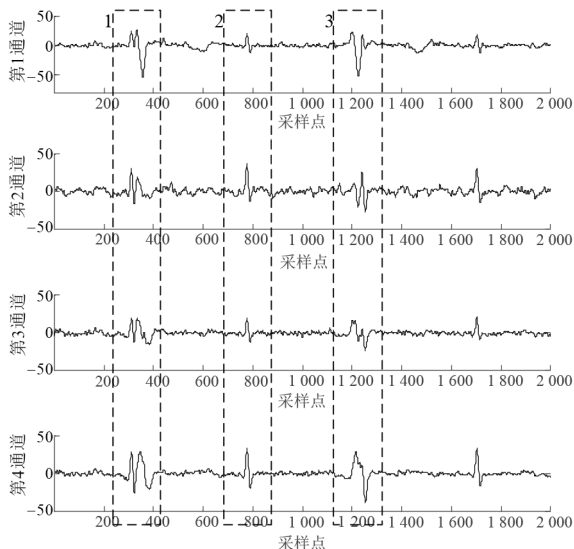


图3 A15(40~42 s)预处理后数据

在图 3 中, 虚线框 1~3 分别显示了 MECG 成分与 FECG 部分重叠、不重叠和完全重叠的情况.

4.2 性能评价方法

主观评价通过观察 FECG 波形是否理想和 MECG 成分残留情况来分析 FECG 提取方法的性能.

客观评价方法如下.

(1) 临床评价指标

使用传统 QRS 检测算法^[19]确定提取得到的 FECG 的 R 峰位置, 并依据 PhysioNet 2013 挑战赛^[20]SET-A 给定的 FECG 的 R 峰的标准位置, 计算灵敏度 (Se)、阳性预测值 (PPV) 和 F_1 分数^[21,22]:

$$\text{Se} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (16)$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (17)$$

$$F_1 = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \times 100\% \quad (18)$$

其中, TP (True Positive)、FP (False Positive) 和 FN (False Negative) 分别表示正确检测的、错误检测的和漏检的 FECG 的 R 峰位置个数. 灵敏度 (Se)、阳性预测值 (PPV) 和 F_1 分数的值越高, 表示性能越好.

(2) 信噪比指标

采用如 2.4 节所述 SNR_{cor} 和 SNR_{svd} 评价提取得到的 FECG 质量.

4.3 本文方法的实验结果与分析

本文所有实验采用的计算机 CPU 为 Intel Core i7-9700K, CPU 主频为 3.60 GHz, 内存为 32 GB, 操作系统为 Windows 10 专业版, 编程工具为 Python 和 Matlab R2019a, 机器学习库使用 Scikit-Learn 和 Lightgbm3.3.0.

采用第 3 节所述的使用信噪比正则化 LightGBM 的 FECG 提取方法, 对 4.1 节所述真实腹部源心电图数据进行实验. 首先对原始母体腹壁混合信号进行噪声抑制, 然后使用 FastICA 分离得到 MECG 估计 $\hat{S}_M(t)$ 和 FECG 估计 $\hat{S}_F(t)$, 参数设定为源数量 $n=4$. 选取 A15 第 40~42 s 的实验结果绘制于图 4.

由图 4 可知, 信号经 FastICA 分解后输出的独立成分中, 前两个独立成分为噪声, 第 3 个独立成分含有较好形态的 MECG, 选择其作为 MECG 估计 $\hat{S}_M(t)$, 第 4 个独立成分含有较为明显的 FECG, 选择其作为 FECG 估计 $\hat{S}_F(t)$. 进一步仔细观察图 4 中的第 4 路信号, 可以发现其中残留了明显的 MECG 成分和其他噪声, 需要对其进行进一步抑制.

使用信噪比正则化 LightGBM 模型进一步抑制 FECG 估计 $\hat{S}_F(t)$ 中残留的 MECG 成分. 使用网格搜索法进行参数优化后确定的模型参数见表 1. 选取 A15 第 40~42 s 的实验结果绘制于图 5.

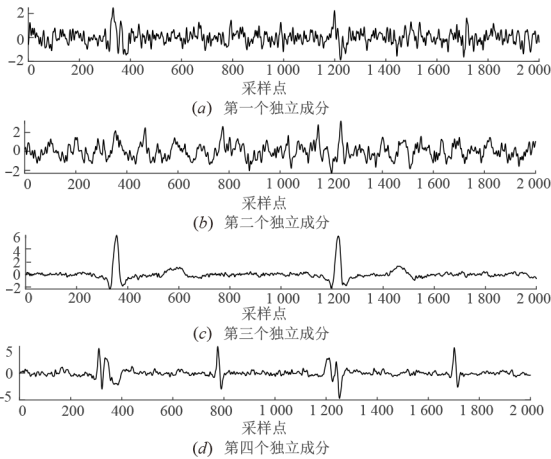


图4 A15(40~42 s)经FastICA分解得到的独立成分

表1 信噪比正则化 LightGBM 模型参数

参数	参数描述	参数值
J	时间导数	20
λ	目标函数加权系数	-1
max_depth	树的最大深度	7
num_leaves	树的最大叶子数量	31
learning_rate	学习率	0.1
feature_fraction	建树的特征选择比例	0.9

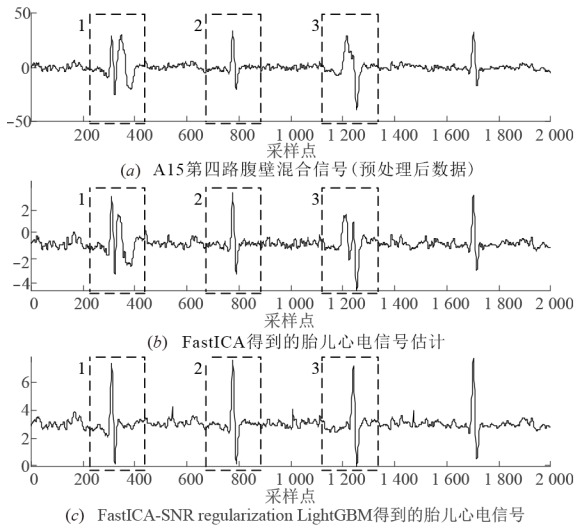


图5 A15 (40~42 s)经FastICA-信噪比正则化LightGBM提取的FECG

由图5可知,与FastICA分离得到的FECG估计 $\hat{S}_F(t)$ 相比,在MECC成分与FECG成分部分重叠、不重叠和完全重叠(虚线框1~3)情况下,经FastICA-信噪比正则化LightGBM均能提取得到的清晰完整的FECG,且其中残留的MECC成分明显减少,FECG质量更高。

采用4.2节所述客观评价方法对经FastICA和经FastICA-信噪比正则化LightGBM提取得到的FECG进行计算,结果分别见表2和表3。

表2 使用FastICA的FECG提取方法的性能评价

数据集	SNR _{cor} /dB	SNR _{svd} /dB	Se/%	PPV/%	F ₁ /%
A04	1.8	2.1	100	78.1	87.7
A08	3.2	3.4	99.2	69.2	81.6
A15	2.0	2.4	97.0	81.1	88.4
A22	0.9	1.2	96.8	96.0	96.4
A32	4.3	5.0	98.7	96.8	97.7
A43	5.4	5.7	100	98.1	99.1
A59	4.0	4.2	98.7	98.7	98.7
平均值	3.1	3.4	98.6	88.3	92.8

表3 使用FastICA-信噪比正则化LightGBM的FECG提取方法的性能评价

数据集	SNR _{cor} /dB	SNR _{svd} /dB	Se/%	PPV/%	F ₁ /%
A04	4.4	4.5	100	99.2	99.6
A08	5.8	6.0	100	98.5	99.2
A15	7.2	7.3	100	99.3	99.6
A22	5.3	5.4	100	99.2	99.6
A32	6.7	6.8	100	99.3	99.7
A43	7.1	7.3	100	99.4	99.7
A59	5.5	5.6	99.3	98.7	99.0
平均值	6.0	6.1	99.9	99.1	99.5

由表2和表3可知,使用FastICA直接提取的FECG质量较低,需进一步抑制其中残留的MECC成分和其他噪声.与使用FastICA的FECG提取方法相比,使用FastICA-信噪比正则化LightGBM提取得到的FECG的各项性能指标均有显著提升,其中SNR_{cor}平均提高了2.9 dB,SNR_{svd}平均提高了2.7 dB,PPV平均提高了10.8%,F₁分数平均提高了6.7%,Se在接近99%的情况下也平均提高了1.3%.这充分表明使用信噪比正则化LightGBM模型能有效抑制残留的MECC成分和其他噪声,提取高质量的FECG。

综上所述,主观和客观性能评价都充分表明本文提出的使用信噪比正则化LightGBM的FECG提取方法能够从母体腹壁混合信号中提取得到完整清晰的FECG,且其质量明显高于使用FastICA直接提取的FECG。

4.4 对比实验结果与分析

为进一步验证本文提出的使用信噪比正则化LightGBM的FECG提取方法,使用4.1节所述真实心电数据进行实验,首先对原始母体腹壁混合信号进行噪声抑制,然后分别进行如下对比实验。

实验1:分别使用TFBSS,TFBSS-LightGBM和TFBSS-信噪比正则化LightGBM提取FECG,选取A15第40~42 s的实验结果绘制于图6。

实验2:分别使用RobustICA,RobustICA-LightGBM和RobustICA-信噪比正则化LightGBM提取FECG,选取A15第40~42 s的实验结果绘制于图7。

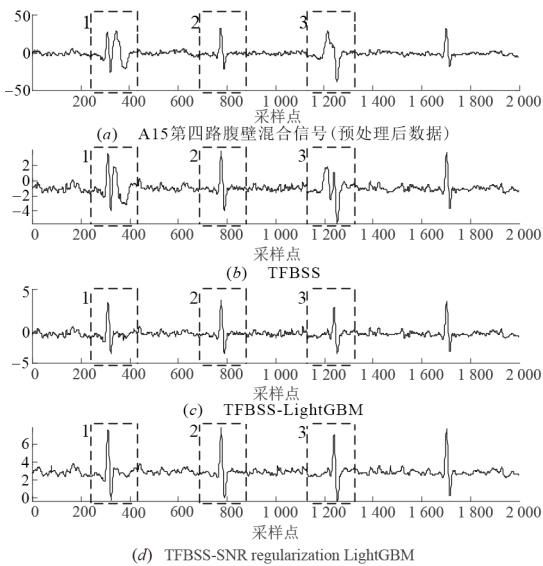


图6 A15(40~42 s)对比实验1

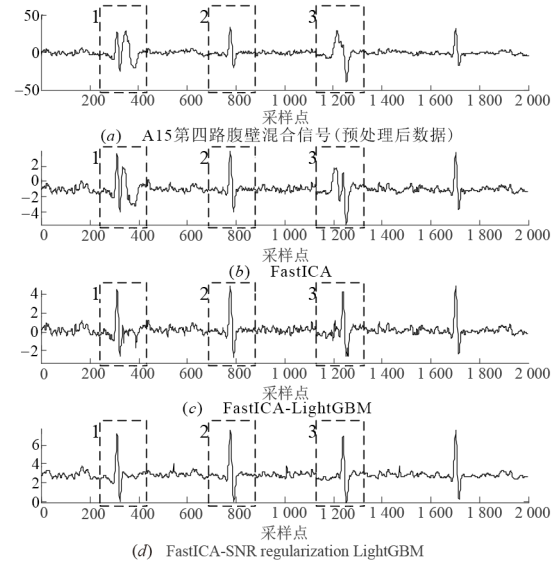


图8 A15(40~42 s)对比实验3

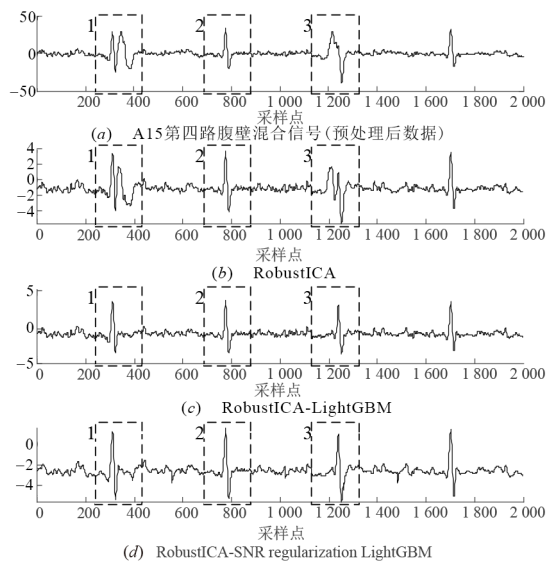


图7 A15(40~42 s)对比实验2

实验3: 分别使用 FastICA, FastICA-LightGBM 和 FastICA-信噪比正则化 LightGBM 提取 FECG, 选取 A15 第 40~42 s 的实验结果绘制于图 8.

其中, TFBS 参数设定为 $n=4$, $N_j=256$, $N_l=256$, $\text{tol}=1/256$, RobustICA 的参数设定为 $\text{prewhi}=\text{true}$, FastICA 的参数设定为 $n=4$; 信噪比正则化 LightGBM 的参数如表 1 所示; 除不含目标函数加权系数 λ 这个参数外, 传统 LightGBM 的其余参数与信噪比正则化 LightGBM 的参数一致.

由图 6~图 8 可知:

(1) 使用 TFBS, RobustICA 和 FastICA 直接提取 FECG, 在 MECG 成分和 FECG 成分部分重叠(虚线框 1) 的情况下, TFBS, RobustICA 和 FastICA 均未能有效抑

制 MECG 成分;

(2) 采用传统 LightGBM 抑制 MECG 成分后, 在 MECG 成分和 FECG 成分部分重叠、不重叠和完全重叠(虚线框 1~3) 情况下, TFBS-LightGBM, RobustICA-LightGBM 和 FastICA-LightGBM 均较为有效地抑制了 MECG 成分;

(3) 采用信噪比正则化 LightGBM 抑制 MECG 成分后, 在 MECG 成分和 FECG 成分部分重叠、不重叠和完全重叠(虚线框 1~3) 情况下, TFBS-信噪比正则化 LightGBM, RobustICA-信噪比正则化 LightGBM 和 FastICA-信噪比正则化 LightGBM 均有效抑制了 MECG 成分和部分噪声, 且其性能分别优于 TFBS-LightGBM, RobustICA-LightGBM 和 FastICA-LightGBM, 这表明信噪比正则化 LightGBM 模型比传统 LightGBM 模型更适用于 FECG 提取.

采用 4.2 节所述临床评价指标和信噪比指标计算方法分别对实验 1~3 提取得到的 FECG 进行计算, 对每种方法在 7 组数据集上实验结果的客观性能评价指标取平均值, 如表 4 所示.

由表 4 可知:

(1) 使用 TFBS, RobustICA 和 FastICA 直接提取的 FECG 的客观评价指标均不太高, 其质量相对较低;

(2) 分别使用 TFBS-LightGBM, RobustICA-LightGBM 和 FastICA-LightGBM 三种方法提取的 FECG, 相对于使用 TFBS, RobustICA 和 FastICA 直接提取的 FECG, 其客观评价指标有所提高, 这说明盲源分离方法结合 LightGBM 是有效的, 其中, FastICA-LightGBM 方法优于 TFBS-LightGBM 和 RobustICA-LightGBM 方法;

(3) 分别使用 TFBS-信噪比正则化 LightGBM, RobustICA-信噪比正则化 LightGBM 和 FastICA-信噪比正

表 4 FECC 提取方法性能对比(平均值)

方法	SNR _{cor} /dB	SNR _{svd} /dB	Se/%	PPV/%	F ₁ /%
TFBSS	2.0	2.3	94.5	77.8	84.2
RobustICA	2.1	2.8	93.1	81.0	86.3
FastICA	3.1	3.4	98.6	88.3	92.8
TFBSS-LightGBM	2.8	3.2	99.5	98.6	99.0
RobustICA-LightGBM	4.9	5.0	99.7	98.9	99.3
FastICA-LightGBM	5.1	5.3	99.7	98.9	99.3
TFBSS-SNR regularization LightGBM	4.5	4.7	99.5	97.1	98.2
RobustICA-SNR regularization LightGBM	5.9	6.0	99.9	99.0	99.5
FastICA-SNR regularization LightGBM	6.0	6.1	99.9	99.1	99.5

则化 LightGBM 三种方法提取的 FECC,其客观评价指标明显提高,这说明本论文提出的信噪比正则化 LightGBM 模型在性能上优于传统 LightGBM 模型,更适用于 FECC 提取,其中, FastICA-信噪比正则化 LightGBM 提取的 FECC 的客观评价指标最高,这说明 FastICA-信噪比正则化 LightGBM 方法的性能优于其他 FECC 提取方法。

综上所述,本文提出的使用信噪比正则化 LightGBM 的 FECC 提取方法无论在主观视觉评价还是在客观性能指标上都明显优于其他传统的 FECC 提取方法。本文提出的信噪比正则化 LightGBM 模型比传统 LightGBM 模型更适用于 FECC 提取,采用本文提出的信噪比正则化方法有效提高了 FECC 提取方法的性能。

5 总结

本论文提出一种使用信噪比正则化 LightGBM 的 FECC 提取方法。针对母体腹壁混合信号,首先使用传统滤波方法进行噪声抑制,然后使用 FastICA 分离得到 MECC 估计和 FECC 估计,再使用信噪比正则化 LightGBM 模型拟合 MECC 成分与 MECC 之间的非线性变换,抑制 FECC 估计中残留的 MECC 成分,得到高质量的 FECC。实验结果表明,本文提出的使用信噪比正则化 LightGBM 的 FECC 提取方法是有效的,且无论在主观视觉效果还是在临床评价指标和信噪比评价指标上,均优于基于盲源分离的 FECC 提取方法,也优于盲源分离结合传统 LightGBM 的 FECC 提取方法。本文提出的 FECC 提取方法能够从母体腹壁混合信号中直接提取 FECC,更便于临床应用。

除 LightGBM 模型外,还有大量的机器学习模型被广泛应用于生物医学信号处理领域,因此,研究如何将本文提出的信噪比正则化方法与其他机器学习模型相结合,提取得到更高质量的 FECC 是下一步研究方向。

参考文献

[1] MARTINEK R, KAHANKOVA R, JEZEWSKI J, et al. Comparative effectiveness of ICA and PCA in extraction of fetal ECG from abdominal signals: Toward non-invasive fe-

tal monitoring[J]. *Frontiers in Physiology*, 2018, 9: 648.

- [2] BEHAR J, ANDREOTTI F, ZAUNSEDER S, et al. A practical Guide to non-invasive foetal electrocardiogram extraction and analysis[J]. *Physiological Measurement*, 2016, 37(5): R1-R35.
- [3] JEZEWSKI J, WROBEL J, MATONIA A, et al. Is abdominal fetal electrocardiography an alternative to Doppler ultrasound for FHR variability evaluation? [J]. *Frontiers in Physiology*, 2017, 8: 305.
- [4] TAHA L Y, ABDEL-RAHEEM E. Fetal ECG extraction using input-mode and output-mode adaptive filters with blind source separation[J]. *Canadian Journal of Electrical and Computer Engineering*, 2020, 43(4): 295-304.
- [5] WARMERDAM G J J, VULLINGS R, SCHMITT L, et al. Hierarchical probabilistic framework for fetal R-peak detection, using ECG waveform and heart rate information [J]. *IEEE Transactions on Signal Processing*, 2018, 66(16): 4388-4397.
- [6] KAHANKOVA R, MARTINEK R, JAROS R, et al. A review of signal processing techniques for non-invasive fetal electrocardiography[J]. *IEEE Reviews in Biomedical Engineering*, 2020, 13:51-73.
- [7] KAHANKOVA R, MARTINEK R, BILIK P. Non-invasive fetal ECG extraction from maternal abdominal ECG using LMS and RLS adaptive algorithms[C]//*Proceedings of the Third International Afro-European Conference for Industrial Advancement(AECIA 2016)*. Marrakesh: Springer, 2016. 258-271.
- [8] ASSALEH K. Extraction of fetal electrocardiogram using adaptive neuro-fuzzy inference systems[J]. *IEEE Transactions on Biomedical Engineering*, 2007, 54(1): 59-68.
- [9] HAN L, PU XIU-JUAN, CHEN XIAO-JUN. Method of fetal electrocardiogram extraction based on ν -support vector regression[J]. *IET Signal Processing*, 2015, 9(5): 430-439.
- [10] TSUI S Y, LIU C S, LIN C W. Modified maternal ECG cancellation for portable fetal heart rate monitor[J]. *Biomedical Signal Processing and Control*, 2017, 32: 76-81.

- [11] 韩亮, 蔡文涛, 蒲秀娟, 龙灵. 结合FastICA与EKF的腹部源胎儿心电信号提取[J]. 仪器仪表学报, 2021, 42(7): 116-125.
HAN L, CAI W T, PU X J, LONG L. Abdominal electrode-sourced FECG extraction utilizing EKF combined with FastICA[J]. Chinese Journal of Scientific Instrument, 2021, 42(7):116-125. (in Chinese)
- [12] LIMA-HERRERA S L, ALVARADO-SERRANO C, HERNÁNDEZ-RODRÍGUEZ P R. Fetal ECG extraction based on adaptive filters and Wavelet Transform: Validation and application in fetal heart rate variability analysis [C]//2016 13th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City: IEEE, 2016: 1-6.
- [13] ANDREOTTI F, RIEDL M, HIMMELSBACH T, et al. Robust fetal ECG extraction and detection from abdominal leads[J]. Physiological Measurement, 2014, 35(8): 1551-1567.
- [14] KE G, MENG Q, FINLEY T, et al. Lightgbm: A highly efficient gradient boosting decision tree[C]//Advances in Neural Information Processing Systems 30(NIPS 2017). Long Beach: Curran Associates, Inc., 2017: 1-9.
- [15] 孙全明, 曲志坚, 任崇广. 基于粒子群优化和LightGBM的情景感知多式联运推荐[J]. 电子学报, 2021, 49(5): 894-903.
SUN Q M, QU Z J, REN C G. Context-aware multi-modal transportation recommendation based on particle swarm optimization and LightGBM[J]. Acta Electronica Sinica, 2021, 49(5): 894-903. (in Chinese)
- [16] Outram N J. Intelligent Pattern Analysis of the Foetal Electrocardiogram[D]. Plymouth: University of Plymouth, 1997.
- [17] VARANINI M, TARTARISCO G, BILLECI L, et al. A multi-step approach for non-invasive fetal ECG analysis [C]//Computing in Cardiology Conference. Zaragoza: IEEE, 2013: 281-284.
- [18] MARTINEZ M, CALPE J, SORIA E, et al. Methods to evaluate the performance of fetal electrocardiogram extraction algorithms[C]//Computers in Cardiology. Rotterdam: IEEE, 2001: 253-256.
- [19] BEHAR J, OSTER J, CLIFFORD G D. Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data[J]. Physiological Measurement, 2014, 35(8): 1569-1589.
- [20] GOLDBERGER A L, AMARAL L A, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals[J]. Circulation, 2000, 101(23): E215-E220.
- [21] JOHNSON A E W, BEHAR J, ANDREOTTI F, et al.

Multimodal heart beat detection using signal quality indices[J]. Physiological Measurement, 2015, 36(8): 1665-1677.

- [22] 张昱, 刘开峰, 张全新, 王艳歌, 高凯龙. 基于组合-卷积神经网络的中文新闻文本分类[J]. 电子学报, 2021, 49(6): 1059-1067.
ZHANG Y, LIU K F, ZHANG Q X, WANG K G, GAO K L. A combined-convolutional neural network for Chinese news text classification[J]. Acta Electronica Sinica, 2021, 49(6): 1059-1067. (in Chinese)

作者简介



韩亮 男, 1975年生于陕西西乡. 生物感知与智能信息处理重庆市重点实验室副主任, 重庆大学微电子与通信工程学院副教授. 主要研究方向为信号处理和图像处理.
E-mail: hanliangaa@cqu.edu.cn



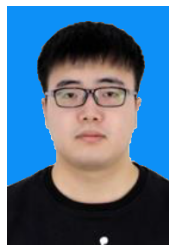
蔡文涛 男, 1994年生于四川都江堰. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为生物医学信号处理.
E-mail: caiwentao5@cqu.edu.cn



蒲秀娟 女, 1979年生于四川隆昌. 重庆大学微电子与通信工程学院讲师. 主要研究方向为生物医学信号处理.
E-mail: puxj@cqu.edu.cn



罗统军 男, 1998年生于广西北流. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为生物医学信号处理.
E-mail: luotongjun@cqu.edu.cn



黄谦 男, 1998年生于四川成都. 重庆大学微电子与通信工程学院硕士研究生. 主要研究方向为生物医学信号处理.
E-mail: Hqian@cqu.edu.cn