

Levenberg-Marquardt Algorithm-Based Gaussian Model Research for ECG

Jie Wu, Tingting Huang, Biqiang Zhang, Baojin Liu, Wei Song

Nanyang Institute of Technology, Nanyang Henan
Email: jacks_wu@163.com

Received: Oct. 7th, 2019; accepted: Oct. 22nd, 2019; published: Oct. 29th, 2019

Abstract

Aimed at the waveform characteristics of main components (P wave, QRS complex and T wave) to an electrocardiogram (ECG), this study proposes a novel method based on Levenberg-Marquardt algorithm (LMA) to build the Gaussian model for researching ECG signal, and then to character the features of each component. Firstly, the wavelet decomposition-based preprocessing method is employed to cancel baseline shift and background noise existing in ECG signal. Secondly, ECG signal is segmented into periodic signal based on the distance between tow sequence R peaks as well its ratio of 1:2. Finally, LMA-based Gaussian Mixture model is built to character each component waveform. The performance of the proposed method has been evaluated by using the 48 ECG records of MIT-BIH arrhythmia database, and the higher average of R-square value 0.9668, used to measure the goodness of fit, has been achieved.

Keywords

Electrocardiogram, Wavelet Transform, Gaussian Fitting, Levenberg-Marquardt Algorithm, Trust-Region Method

基于Levenberg-Marquardt算法的心电高斯模型研究

吴 杰, 黄婷婷, 张弼强, 刘保进, 宋 伟

南阳理工学院, 河南 南阳
Email: jacks_wu@163.com

收稿日期: 2019年10月7日; 录用日期: 2019年10月22日; 发布日期: 2019年10月29日

摘要

针对心电信号中各构成成分(P波、QRS波、T波)波形特征,本研究提出了一种基于伯格-马夸尔特算法创建心电高斯模型以实现心电波形特征研究进而对心电特征进行定性评价研究。首先,基于小波分解对心电信号进行预处理,以消除基线漂移以及背景噪声;其次,以心电信号各R峰间距为参考,按照1:2的分割比例对心电信号进行周期分割;最后,基于Levenberg-Marquardt (LM)算法创建心电高斯混合模型以刻画各成分波形特征。本算法通过对MIT-BIH心律失常数据库中48条心电数据进行分析,利用其拟合优度均值为0.9668对本算法可行性进行了验证。

关键词

心电信号, 小波变换, 高斯拟合, LM算法, 置信域算法

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来心血管疾病的患病率和死亡率均呈现持续增长的趋势,并且心血管疾病死亡人数连续多年位居疾病死亡之首。根据《中国心血管报告 2018》[1],2016年农村因心血管疾病死亡人数占全部因疾病死亡人数的45.50%(见图1(a)),城市因心血管疾病死亡人数占全部死亡人数的43.16%(见图1(b))。依据心血管疾病患病率持续上升的趋势,推测我国现有2.9亿患者,其中心脏病占有很大比重。我国心血管病防治工作面临严峻挑战。

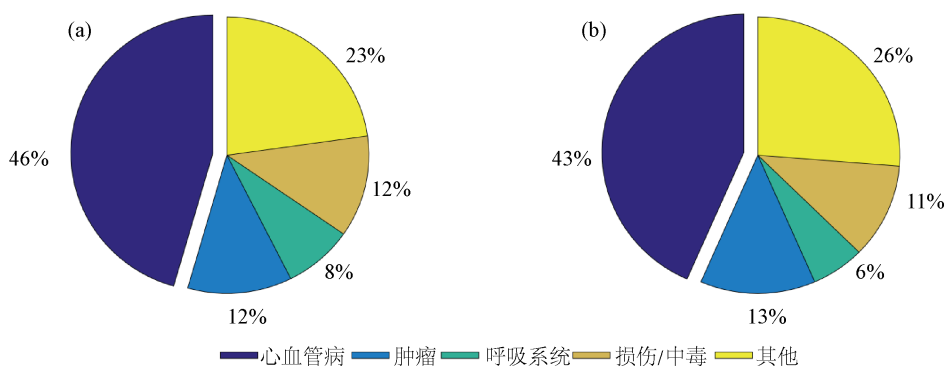


Figure 1. The main cause of death in rural and urban residents in China in 2016

图 1. 2016 年中国农村和城市居民主要疾病死因构成

传统心血管疾病治疗最大弊端是需持续长时间心电信号记录,从而导致绝大多数心血管疾病患者错过了最佳理疗期。随着计算机技术,数字信号处理手段等先进技术的不断发展,借助计算机和数字信号处理手段对心电图信号进行检测、处理和分析已逐步取代传统诊断方式。心电图(ECG)包含大量用于分析心脏功能的生理信号,是用来诊断心脏疾病的主要依据之一。表征心电信号的形态特征是诊断心脏性疾病、心率变异性分析、生物统计学和心电信号自动分类中最重要的步骤,这些系统的性能很大程度上依

赖于心电信号描述的准确度。而心电信号的特征提取、识别及分类已经被国内外研究人员引入了许多种算法。一般可用的特征表示方法包括形态学[2]、时间信息[3]、小波变换[4]、高阶统计量(HOS) [5]、Hermite基函数[6]、ECG波间隔[7]、选定的采样点[8]、QRS波段镜像高斯模型[9]等。

心电信号特征提取经历了从单个节拍特征提取到海量心电信号特征提取的发展过程。基于滤波技术和决策规则的方法计算效率高, 是任何心电分析的理想方法。一般来说, 预处理阶段采用各种信号处理技术来增强心电波的特征和抑制噪声, 但大多数方法都存在一定的缺点。如, 基于小波变换的预处理存在母波和尺度的选择问题, 基于基准点描述特征存在难度较大和精度较低的问题。对于真实的心电信号波形形态各异, 单一的决策规则很难描述各种病理信号。周群一将 QRS 复合波建立数学模型并应用在心拍的分类, 但该方法只描述了 QRS 复合波的形状, 却丢失了其他部分信息。

本文提出了一种基于简单的小波变换和 Levenberg-Marquardt 算法建立函数模型表述心电节拍信号形态的方法。本文结构组织如下。第一部分讲述基于小波分解对心电信号进行简单的预处理——去除基线漂移以及背景噪声, 第二部分介绍了以心电信号各 R 峰间距为参考, 按照 1:2 的分割比例对心电信号进行周期分割, 第三部分详细介绍了建立数学模型的两个步骤。最后给出了实验结果。使用 MIT-BIH 心律失常数据库对本文提出的方法进行了实验评估。

2. 方法

根据算法实现原理, 心电信号高斯模型拟合主要分为三部分如图 2。

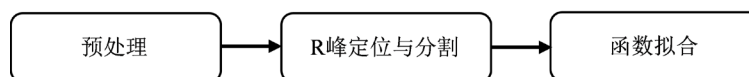


Figure 2. Block diagram of Gaussian fitting of ECG signals
图 2. 心电信号高斯拟合流程框图

2.1. 小波变换处理

在信号处理中, 傅里叶变换在处理非平稳信号及刻画信号在时域中的局部特性上具有局限性。由此衍生出了小波变换。小波分解采用有限时间分布且衰减的函数作为基, 使得小波变换具有了傅里叶变换没有的时频分析能力, 其原理如公式(1)所示。

$$C_{\varphi} = \int_{-\infty}^{+\infty} \frac{|\hat{\varphi}(w)|^2}{|w|} d\omega < +\infty \quad (1)$$

小波变化广泛应用在信号处理、数字图像处理。为了实现很好的处理效果, 在实际应用中, 学者在傅里叶变换基础上衍生出许多小波, 其中 dB 小波具有正交性、指数多项式消失距和连续紧支性等特性, 因此本研究采用 Daubechies 小波作为母小波进行小波变换。研究[10]表明心电信号基线频率段为 0.15~0.3 Hz, 鉴于此本研究根据心电信号特征选用 dB8 小波对心电信号处理, 心电信号原始波形如图 3 所示, 其去除基线漂移效果如图 4 所示。

2.2. 心电信号的切割

第二步中, 算法通过定义一组规则来识别和定位感兴趣的点。如 R 峰峰值位置、P 波起始点、T 波终点等。在 MIT-BIH 心电信号数据库中, 每组数据都经过两位心电图专家独立对信号进行标注, 标注内容包括信号质量、节律变化、节拍局部说明及每个心电节拍的类型。其中每个心电节拍的类型都准确地标注在该节拍的 R 波波峰位置, 方便了广大研究人员进行研究分析。本文就以注释位置为标准, 每个 RR

间隔的 2:1 划分每个心电节拍[11], 避免心律不齐信号切分的完整性, 保证了切割后的心电节拍中都包含有 P 波、QRS 复合波和 T 波等主要波形。对心电波形图(图 5)心拍分割处理的两个阶段及实验结果如图 6, 图 7 所示。该切割方法在心律失常的情况下依然可以按要求将心电节拍切割。

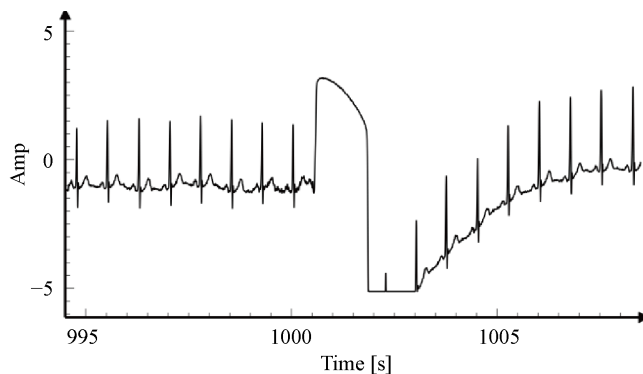


Figure 3. Original waveform of the ECG signal
图 3. 心电信号原始波形图

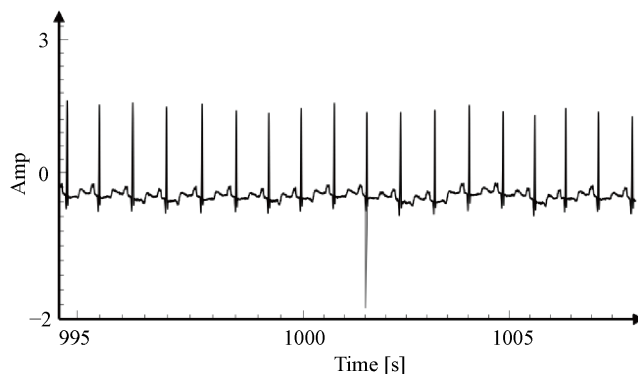


Figure 4. ECG signal removal baseline drift plot
图 4. 心电信号去除基线漂移图

2.3. 建立高斯模型

高斯函数可用式(2)表示, 式中 y_0 为峰高, μ 为峰的中心, σ 为高斯曲线的峰宽。高斯函数是非线性函数, 对公式(3)取对数将其线性化, 最终得式(4)线性函数表达式

$$y = y_0 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

$$\ln y = \ln y_0 - \frac{(x-\mu)^2}{2\sigma^2} \quad (3)$$

$$\ln y = \ln y_0 - \frac{\mu^2}{2\sigma^2} + \frac{2\mu x}{2\sigma^2} - \frac{x^2}{2\sigma^2} \quad (4)$$

将函数系数分别计为 c_1, c_2, c_3 , 则

$$c_1 = \ln y_0 - \frac{\mu^2}{2\sigma^2}, \quad c_2 = \frac{\mu}{\sigma^2}, \quad c_3 = -\frac{1}{2\sigma^2} \quad (5)$$

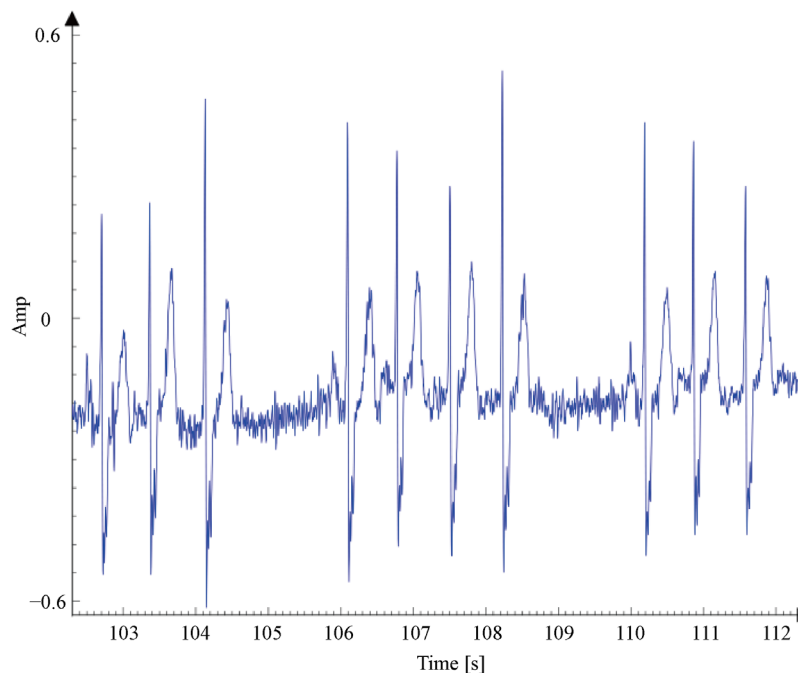


Figure 5. Original waveform of the ECG signal
图 5. 心电信号原波形图

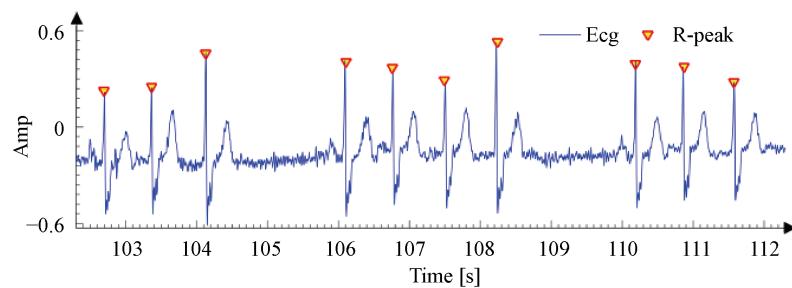


Figure 6. ECG signal R peak location map
图 6. 心电信号 R 峰定位图

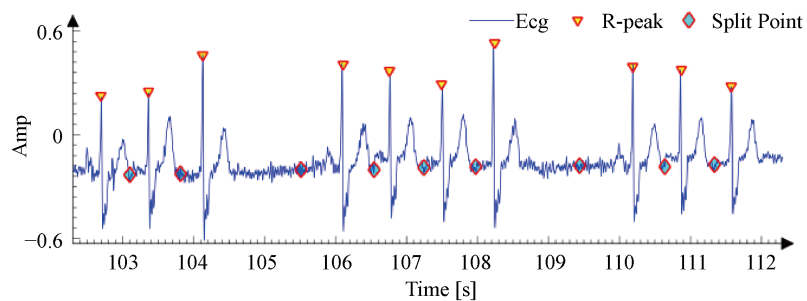


Figure 7. ECG signal segmentation
图 7. 心电信号分割图

公式(4)可以表示为 $\ln y = \begin{bmatrix} 1 & x & x^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$, 对于单高斯函数拟合, 可以计为 $Y = A_{1 \times 3} C_{3 \times 1}$

对于多高斯函数公式(6)所示

$$f(y, n) = \sum_{m=1}^n y_m \frac{1}{\sqrt{2\pi}\sigma_m} e^{-\frac{(x-\mu_m)^2}{2\sigma_m^2}} \quad (6)$$

其 y 为长度为 n , 对应的公式(4)可以表示为

$$\begin{bmatrix} \ln y_1 \\ \vdots \\ \ln y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \quad (7)$$

简记为 $Y_n = A_{n \times 3} C_{3 \times 1}$, 计算可得 $C = (A^T A)^{-1} A^T Y$ 。数组 C 中都可以由参数 y_0, μ, σ 计算出来。但由于多峰的复杂性, y_0, μ, σ 只是估计值因此要采用最小二乘[12]拟合求出最优的高斯函数参数。高斯函数拟合转化为了非线性最小二乘问题, 最小二乘问题可分为线性和非线性两类。线性最小二乘问题的解是封闭式的确定性的, 即 $x = (A^T A)^{-1} A^T Y$ 。对于非线性最小二乘问题通常采用迭代法求解, 在计算中, 有许多优化算法用于处理最小二乘问题, 使用较少的计算量达到很好的结果。比如梯度下降算法、高斯-牛顿算法、共轭梯度法、列为伯格-马夸尔特算法和信赖域算法等。心电信号的高斯拟合是非线性的最小二乘问题, 目标函数可以表示为

$$F(x) = \sum_{i=1}^m f_i^2(x) \quad (8)$$

其中 $x = (x_1, x_2, x_3, \dots, x_n)^T$, 通过优化使 $F(x)$ 最小化, 本实验选用了在曲线拟合中应用广泛的列为伯格-马夸尔特算法和信赖域算法进行实验并对比二者结果。

2.3.1. 列为伯格-马夸尔特算法

列为伯格-马夸尔特算法(简称 LMA 或 LM), 也称为阻尼最小二乘(DLS)。它是建立在高斯牛顿算法和梯度下降算法之间的优化算法, 在许多情况下, 即使初始值距离最终的最小值很远, 它也可以找到解决方案。LM 算法是解决非线性最小二乘问题的重要算法, 主要应用于曲线拟合问题。LM 算法流程如下: 把目标函数式(8)在 x^k 处一阶泰勒展开, 用 $\phi_i(x)$ 表示可

$$\begin{aligned} \phi_i(x) &= f_i(x^k) + \nabla f_i(x^k)^T (x - x^k) \\ &= \nabla f_i(x^k)^T x - \left[\nabla f_i(x^k)^T x^k - f_i(x^k) \right], i = 1, 2, 3, \dots, m \end{aligned} \quad (9)$$

则

$$\begin{aligned} \psi(x) &= \sum_{i=1}^m \phi_i^2(x) \\ &= \left(\begin{bmatrix} \nabla f_1(x^k)^T \\ \vdots \\ \nabla f_m(x^k)^T \end{bmatrix} x + \begin{bmatrix} \nabla f_1(x^k)^T x^k - f_1(x^k) \\ \vdots \\ \nabla f_m(x^k)^T x^k - f_m(x^k) \end{bmatrix} \right)^T \left(\begin{bmatrix} \nabla f_1(x^k)^T \\ \vdots \\ \nabla f_m(x^k)^T \end{bmatrix} x + \begin{bmatrix} \nabla f_1(x^k)^T x^k - f_1(x^k) \\ \vdots \\ \nabla f_m(x^k)^T x^k - f_m(x^k) \end{bmatrix} \right) \end{aligned} \quad (10)$$

用 $\psi(x)$ 近似表示 $F(x)$ 进每一步迭代。

$$\text{令 } A_k = \begin{bmatrix} \nabla f_1(x^k)^T \\ \vdots \\ \nabla f_m(x^k)^T \end{bmatrix}, B = \begin{bmatrix} \nabla f_1(x^k)^T x^k - f_1(x^k) \\ \vdots \\ \nabla f_m(x^k)^T x^k - f_m(x^k) \end{bmatrix}, f_k = \begin{bmatrix} f_1(x^k) \\ \vdots \\ f_m(x^k) \end{bmatrix}。 \text{式(10)可以用式(11)表示}$$

$$\psi(x) = (A_k x - B)^T (A_k x - B) \tag{11}$$

求导并令导数为零 $A_k^T A_k x = A_k^T (A_k x^k - f_k)$ ，将一个正定对角矩阵加入到 $A_k^T A_k$ 中，步进增量式变为下面形式 $x - x^k = -(A_k^T A_k + \lambda I)^{-1} A_k^T f_k$ 。其中 I 是一个单位矩阵， λ 是一个正实数，当 $\lambda = 0$ 时，LM 算法即为高斯牛顿算法。当 λ 很大时， $x - x^k \approx -A_k^T f_k$ ，LM 算法就变成了梯度下降算法。 λ 的值会在优化中不断变化，控制着 x 的前进方向，也控制着 x 每次变化的量。在优化前期类似梯度下降算法使 x 快速收敛到极值点附近；在优化后期类似高斯牛顿算法， x 变化速度逐渐变小，使 x 稳定收敛达到最优，LM 算法在每次迭代中不仅要和高斯牛顿算法一样更新 m 个函数 $f_i(x^k)$ 的函数值和一阶偏导数，还要更新 λ 。

2.3.2. 置信域算法

置信域算法(Trust-region methods) [13]，它于 1970 年由 Powell 提出。该方法从给定初始解出发，迭代计算，直至达到满意的结果为止。对于 $f(x)$ 是在 R^n 上的二阶连续可微函数，定义现在点的邻域 Ω_k

$$\Omega_k = \{x \in R^n \mid \|x - x_k\| \leq \Delta_k\} \tag{12}$$

在式(12)中， Δ_k 为置信域半径，在这个邻域中极小化目标函数的近似二次模型，得到近似极小点 S_k ，其中 $\|S_k\| \leq \Delta_k$ 。置信域方法的子模问题是

$$\begin{cases} \min q^{(k)}(s) = f(x_k) + g_k^T s + \frac{1}{2} s^T B_k s \\ \text{s.t. } \|s\| \leq \Delta_k \end{cases} \tag{13}$$

其中， $s = x - x_k$ ， $g_k = \nabla f(x_k)$ ， B_k 是一个对称矩阵。 Δ_k 根据 $q^{(k)}(s)$ 对目标函数 $f(x)$ 的拟合程度来调整置信域，对于 S_k ，定于以下两个参量

$$Ared_k = f(x_k) - f(x_k + s_k), \quad Pred_k = q^{(k)}(0) - q^{(k)}(s_k), \quad r_k = \frac{Ared_k}{Pred_k} \tag{14}$$

其中 $Ared_k$ 为目标函数下降量为实际下降量， $Pred_k$ 为函数模型的下降量记为预测下降量。 r_k 用来衡量目标函数与实际函数的一致性程度。计算步骤如下：

- 1) 给出初始值 x_0 ，置信域半径 $\bar{\Delta}$ ， $\Delta_0 \in (0, \bar{\Delta})$ ， $\varepsilon \geq 0$ ， $0 \leq \eta_1 \leq \eta_2 < 1$ ， $0 < \gamma_1 < 1$ ， $k = 0$
- 2) 如果 $\|g_k\| \leq \varepsilon$ ，则停止计算。
- 3) 求模型子问题，得到 S_k
- 4) 计算 $f(x_k + s_k)$ 和 r_k 其中

$$x_{k+1} = \begin{cases} x_k + s_k, & r_k \geq \eta_1 \\ x_k, & r_k < \eta_1 \end{cases} \tag{15}$$

- 5) 按照下列规则校正置信域半径，

$$\Delta_{k+1} \in [0, \gamma_1 \Delta_k], \quad r_k < \eta_1 \tag{16}$$

$$\Delta_{k+1} \in [\gamma_1 \Delta_k, \Delta_k], \quad r_k \in [\eta_1, \eta_2) \tag{17}$$

$$\Delta_{k+1} \in [\Delta_k, \min\{\gamma_2 \Delta_k, \bar{\Delta}\}], \quad r_k \geq \eta_2 \tag{18}$$

- 6) 得出 B_{k+1} ，修正 $q^{(k)}$ ， $k = k + 1$ ，循环第 2 步。

3. 实验与分析

3.1. 拟合评价标准

为了确定拟合模型对实际的心电数据是否具有良好的逼近性，采用了误差均方根(RMSE)、误差平方

和(SSE)、总平方和(SST)和确定系数(R-square)等拟合优度统计方法。计算方法如下:

$$SSE = \sum_k^n (f(x_k) - y_k)^2, SST = \sum_k^n \left(f(x_k) - \frac{1}{n} \sum_k^n f(x_k) \right)^2, R\text{-square} = 1 - \frac{SSE}{SST} \quad (19)$$

其中 SSE 统计参数计算的是拟合数据与原始数据对应点的误差的平方和, 其越接近于 0, 说明模型选择和拟合的结果越好, RMSE 和 SSE 效果一样, 这两个参数都表现在拟合结果和原始数据点与点层面的结果。R-square 是通过数据整体的变化来描述一个拟合结果的好坏。确定系数越接近 1, 说明拟合结果对原始数据的解释能力越强。在研究中使用 SSE 与 R-square 作为拟合结果的评价指标。

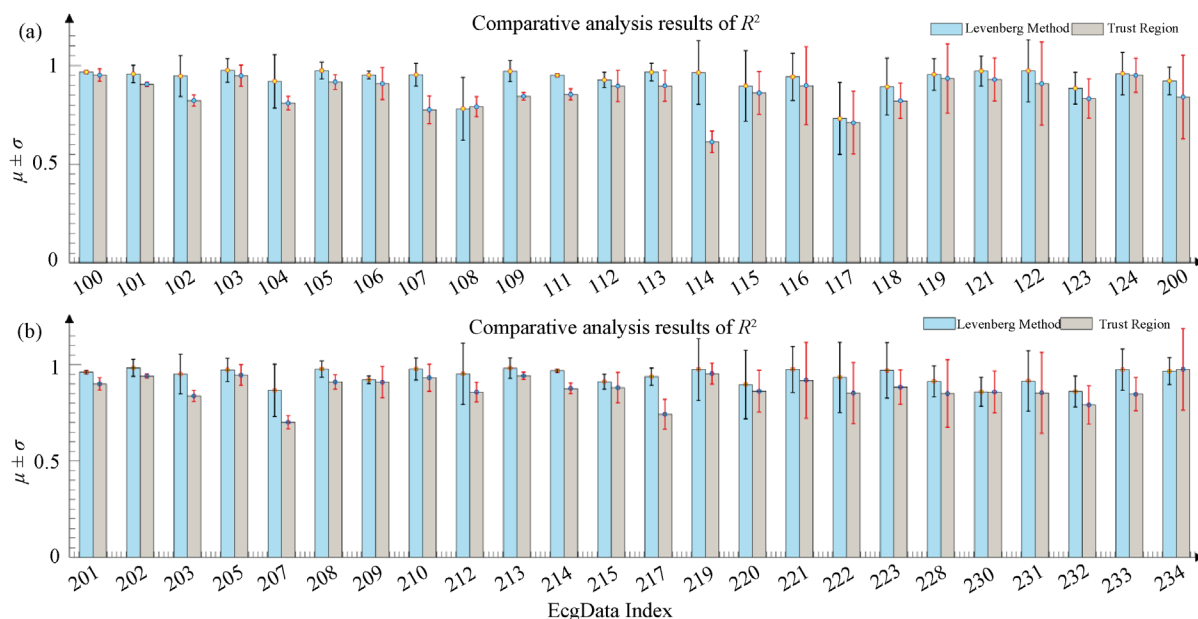


Figure 8. Goodness of fit between LM algorithm and TR algorithm

图 8. LM 算法与 TR 算法拟合优度

3.2. 实验结果

本研究采用 MIT-BIH 心律失常数据库对提出的心电信号拟合方法进行了评价。它包含 48 小时的两个通道心电记录采样在 360 赫兹。提出的算法在 2.5 GHz Intel(R) Core(TM) i5-5200U CPU 上使用 MATLAB 2018b 实现, 并使用数据库中第一通道采集的心电信号进行测试。

根据心电信形态, 采用不同的高斯个数为模型, 在本研究中, 对比了三到七个高斯函数模型对心电节拍拟合的效果。以 10,000 组心电节验证了多高斯拟合结果, 确定以 6 个高斯函数作为本研究的拟合模型。并采用置信域方法与列为伯格 - 马夸尔特算法对比了 SSE 与 R-square。两种方法拟合优度对比如图 8 示, LM 算法普遍优于 TR 算法。

参考文献

- [1] 胡盛寿, 等. “中国心血管病报告 2018”概要[J]. 中国循环杂志, 2019, 34(3): 209-220.
- [2] Dima, S.M., Panagiotou, C., Mazomenos, E.B., et al. (2013) On the Detection of Myocardial Scar Based on ECG/VCG Analysis. *IEEE Transactions on Biomedical Engineering*, **60**, 3399-3409. <https://doi.org/10.1109/TBME.2013.2279998>
- [3] Kutlu, Y. and Kuntalp, D. (2012) Feature Extraction for ECG Heartbeats Using Higher Order Statistics of WPD Coef-

- ficients. *Computer Methods and Programs in Biomedicine*, **105**, 257-267.
<https://doi.org/10.1016/j.cmpb.2011.10.002>
- [4] Yang, H., Kan, C., Liu, G. and Chen, Y. (2013) Spatiotemporal Differentiation of Myocardial Infarctions. *IEEE Transactions on Automation Science and Engineering*, **10**, 938-947. <https://doi.org/10.1109/TASE.2013.2263497>
- [5] El Bouny, L., Khalil, M. and Adib, A. (2018) ECG Signal Filtering Based on CEEMDAN with Hybrid Interval Thresholding and Higher Order Statistics to Select Relevant Modes. *Multimedia Tools and Applications*, **78**, 13067-13089. <https://doi.org/10.1007/s11042-018-6143-x>
- [6] De Lannoy, G., François, D., Delbeke, J. and Verleysen, M. (2012) Weighted Conditional Random Fields for Supervised Interpatient Heartbeat Classification. *IEEE Transactions on Biomedical Engineering*, **59**, 241-247. <https://doi.org/10.1109/TBME.2011.2171037>
- [7] 陈志博, 李智, 李健, 等. 基于 RR 间期和多特征值的房颤自动检测分类[J]. 生物医学工程学杂志, 2018, 35(4): 550-556.
- [8] Mondéjar-Guerra, V., Novo, J., Rouco, J., Penedo, M.G. and Ortega, M. (2019) Heartbeat Classification Fusing Temporal and Morphological Information of ECGs via Ensemble of Classifiers. *Biomedical Signal Processing and Control*, **47**, 41-48. <https://doi.org/10.1016/j.bspc.2018.08.007>
- [9] Zhou, Q.Y., Liu, X. and Duan, H.L. (2006) ECG Beat Classification Using Mirrored Gauss Model. 2005 *IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, 17-18 January 2006, 5587-5590. <https://doi.org/10.1109/IEMBS.2005.1615752>
- [10] 李文涛. 心电信号噪声处理的研究[J]. 数字技术与应用, 2016(8): 85-86.
- [11] Banerjee, S., Gupta, R. and Mitra, M. (2012) Delineation of ECG Characteristic Features Using Multiresolution Wavelet Analysis Method. *Measurement: Journal of the International Measurement Confederation*, **45**, 474-487. <https://doi.org/10.1016/j.measurement.2011.10.025>
- [12] 李颖, 林洪生. 基于相对误差的曲线最小二乘拟合[J]. 沈阳师范大学学报, 2012, 30(3): 338-342.
- [13] Wright, S.J., Goldberg, A.V., Johnson, D.S., *et al.* (2000) Trust-Region Methods. Elsevier, Amsterdam.