

# 基于双正交小波的 ECG 信号自适应压缩编码算法

邹宇华, 骆德汉, 庄家俊

(广东工业大学 信息工程学院, 广东 广州 510006)

**摘要:** 提出 1 种基于  $\text{Bior4.4}$  双正交小波的 ECG 信号压缩算法, 在频率域上对小波系数进行连续分割, 采用自适应的十进制至二进制数据格式转换机制来编码小波系数. 从 MIT-BIH 数据库中选择 8 种不同的 ECG 信号作为测试数据, 算法的平均均方根误差比 (PRD) 在  $0.05\% \sim 4.70\%$  之间, 平均压缩率 (CR) 在  $1.5:1 \sim 33:1$  之间. 实验结果表明本算法具有压缩率高、计算复杂度低、编码速度快的特点.

**关键词:** ECG 信号压缩; 小波变换; 自适应编码

中图分类号: TP274 R318

文献标识码: A

文章编号: 1007-7162(2009)01-0044-04

心电信号的压缩和编码主要是通过检测和消除心电信号中的冗余信息来实现的, 压缩的方法大体可划分为 3 类: (1) 直接数据压缩法<sup>[1-2]</sup>. 通过原始采样数据直接分析处理、消除冗余信息, 主要算法有 TRAZTEC、CORTESE、FAN、LADT 算法等. (2) 参数提取法<sup>[3-5]</sup>. 提取信号的一组特征参数, 一般要根据对信号的一些先验知识, 用这些参数对信号进行分类, 如语义分析法、神经网络法等, 该算法的难点是对信号进行恢复. (3) 变换压缩法<sup>[6-13]</sup>. 对信号进行某种时/频域分析, 以更好地检测和消除冗余数据, 主要方法有 FDT、DCT、KLT、HT 及 DWT 等.

在变换压缩法中, 小波变换 (DWT) 方法是近年来的研究热点, 它具有良好的时域和频域局部分析特性, 通过在信号的小波分解域对小波系数进行量化来实现信号压缩. 研究表明, 常用的 Daubechies 小波函数系并不具有最优信号压缩性能, 双正交小波函数系才是最佳小波函数, 其中最优综合压缩性能的小波函数是  $\text{bior4.4}$  和  $\text{bior2.8}$ <sup>[14]</sup>.

著名的嵌入式小波编码算法, 例如嵌入式零树小波 (EZW)<sup>[6]</sup>、多级树集合分裂 (SPHT)<sup>[7]</sup> 和分层零编码 (LZC)<sup>[8]</sup> 等, 都是一种位平面或分层位扫描编码方法. 这类方法能实现对压缩数据的渐进传输, 即首先传输信号的轮廓数据, 然后传输信号的精细数据. 这种位平面或分层位扫描编码方法可视为一种横向扫描技术, 其主要不足是每一级编码都会扫描全部小波系数, 这就增大了扫描路径的复杂度, 减慢了算法的运行速度.

针对这一问题, 本文提出一种能进行横向和纵向位扫描的分块优先编码方法, 具有压缩率高、计算

复杂度低、编码速度快的特点.

## 1 编码算法原理

首先利用  $\text{Bior4.4}$  小波对经过去均值处理后的心电信号数据进行小波变换, 生成相应的低频、高频子带小波系数. 每次扫描过程中, 对数据位首先进行横向扫描, 再进行纵向扫描. 这种编码方法不需要进行零树分析, 每个小波系数只扫描一次, 这就使编码执行非常简单. 对数据位的纵向扫描, 采用自适应的十进制至二进制数据格式转换机制来编码小波系数. 算法的具体过程如下.

### 1.1 数据预处理

对 ECG 信号进行离散小波变换前, 先对信号进行去均值处理, 使采样数据由无符号描述转换为以零或接近零的数值为中心的互补描述, 这样数据的全局动态范围变小, 经小波变换后得到的小波系数均方差也变小, 故可用较少的码字进行编码.

### 1.2 离散小波变换

将去均值后的 ECG 信号数据利用样条小波  $\text{Bior4.4}$  进行离散小波变换, 以获取最优的综合压缩性能<sup>[14]</sup>. 分解级数为 5 级. 对信号进行 5 级分解后, 所得的小波系数序列是按频率从低到高排列的, 可依次划分成一个低频系数集 (即第 5 分解级低频小波系数  $c_{A_5}$ ) 和 5 个高频系数集 (即第 5 至第 1 分解级的高频小波系数  $c_{D_5} \sim c_{D_1}$ ). 其中, 第 5 分解级低频部分小波系数的能量贡献率约占信号总能量的  $99.5\%$ <sup>[15]</sup>. 图 1 给出了 MIT-BIH 数据库 No. 117 信号经过均值去除处理后, 用  $\text{Bior4.4}$  双正交小波进行 5 级小波变换得到的系数幅值图.

收稿日期: 2008-10-16

作者简介: 邹宇华 (1984-), 男, 硕士研究生, 主要研究方向为测试计量技术与仪器.

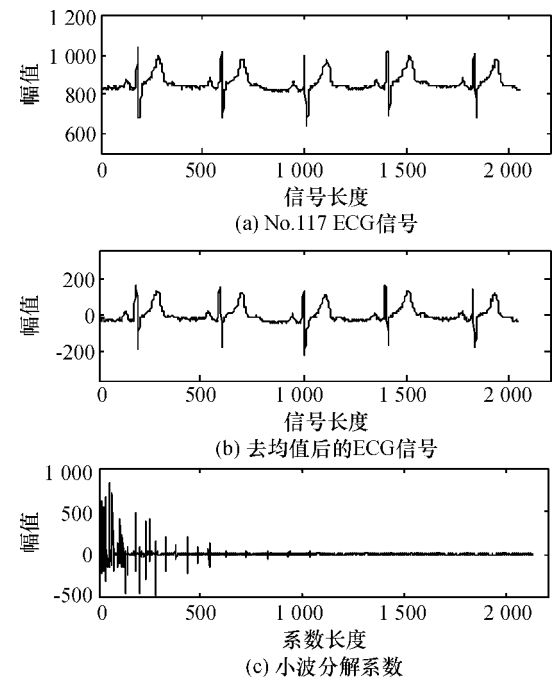


图 1 Record 117 心电信号的预处理与小波分解

1.3 自适应进制转换

根据小波变换的滤波器组实现原理<sup>[16]</sup>, 若信号长度为  $n$ , 滤波器的长度为  $2N$ , 则 1 次小波分解后得到的平均系数和细节系数的长度均为

$$L = \left\lceil \frac{n-1}{2} \right\rceil + N \tag{1}$$

本文所使用的测试数据采样长度为 2 048, 经 5 级小波分解后可得到 6 个小波系数集  $cA_5$ 、 $d_5 \sim d_1$ 。根据式 (1), 可知这 6 个系数集的长度依次为, 72, 72, 136, 263, 518, 1 028, 则小波系数序列的总长度为 2 089。在信号编码中需要将小波系数由十进制数转换为二进制码。对于不同的系数子集, 使用不同的二进制位长来表示系数值, 另外用 1 个符号位来表示系数的正负。每个系数子集的二进制位长按照式 (2) 计算:

$$n_s = \lceil \log_2(\max\{|c_s|\}) \rceil + 1. \tag{2}$$

其中  $n_s$  表示子集系数二进制表示码的位长,  $c_s$  ( $s = 1, 2, \dots, 6$ ) 依次表示子集  $cA_5$ 、 $d_5 \sim d_1$ 。

1.4 编码输出

本算法编码器的输出流由 4 个部分组成: (1) 全局信息  $H_1$ , 长度为 30 位, 包含信号的长度 (12 位)、信号的平均值 (11 位)、小波类型 (4 位)、小波变换级数 (3 位) 等信息; (2) 分集信息  $H_2$ , 长度为 4 位, 表示各分集系数编码所需的码位长度; (3) 符号位  $Sign$ , 长度为 1 位, 表示小波系数的正负符号信息; (4) 系数编码表  $Coef$  表长由分集信息给出的码位长度确定, 存放小波系数的二进制表示码。对于一组样本长度为  $N$  的 ECG 信号数据, 经小波变

换  $L$  级分解后, 可获得 1 个低频系数集和  $L$  个高频系数集, 全系数编码时, 输出流将包括 1 个全局信息、 $(L+1)$  个分集信息、 $N$  个符号位和  $N$  个系数编码表。

2 实验结果与性能分析

在实验中从 MIT-BIH arrhythmia 数据库里分别选取 Records 101、117、121、202、209、212、215、222 等 8 组 ECG 信号的头 2 048 位样本值作为测试数据, 所有数据的采样频率均为 360 Hz, 采样精度均为 11 位。使用 Matlab R2007b 作为测试平台。

2.1 性能指标

使用 2 个指标来衡量压缩编码算法的性能, 即压缩比 CR 和均方根误差比 PRD。具体计算公式为:

$$CR = \frac{N_{in}}{H_1 + (H_2 \times 6) + N_{sign} + N_{coef}}, \tag{3}$$

其中,  $N_{in}$  表示输入信号的数据位流总长度 ( $11 \times 2\,048$  位),  $H_1$  表示全局头信息的长度 (11 位),  $H_2$  表示分集头信息的长度 (4 位, 共有 6 个分集头信息),  $N_{sign}$  表示符号位的总数 (2 048 位),  $N_{coef}$  表示所有小波系数的编码位流长度。

$$PRD = \sqrt{\frac{\sum_{i=1}^N [x_{in}(i) - x_{rec}(i)]^2}{\sum_{i=1}^N x_{in}(i)^2}} \tag{4}$$

其中,  $N = 2\,048$ ,  $x_{in}(i)$ 、 $x_{rec}(i)$  分别表示原始 ECG 信号数据和重构的 ECG 信号数据。

2.2 实验分析

图 1 给出了 record 117 心电信号的预处理和小波变换结果。可以看到, 均值消除后信号幅值降低到  $-200 \sim 200$  范围内, 这样就能有效降低编码小波系数所需的码位长度。经过 5 级小波变换后, 根据各个系数集的系数最高幅值, 可得到  $cA_5$ 、 $d_5 \sim d_1$  的系数编码位长分别是 10, 9, 9, 8, 6, 5 (bits)。

编码器通过对小波系数的自适应编码获得优良的压缩性能, 用最短的码位长度来表示小波系数。对全部小波系数进行编码时, 可以在 PRD 小于 0.05% 的条件下达到 2:1 的压缩比。按照小波分解级从低到高的次序, 逐步忽略高频系数集  $d_1 \sim d_5$ , 即不对该系数集进行编码, 解码时对未被编码的子集系数作置零处理, 可以显著提高信号的压缩比, 不过均方根误差比也相应增大。从 MIT-BIH 数据库中选择 8 种不同的 ECG 信号作为测试数据, 算法的平均均方根误差比 (PRD) 在 0.05% ~ 4.7% 之间, 平均压缩率 (CR) 在 1.5:1 ~ 33:1 之间。图 2 给出了编码器对 8 组 ECG 信号的压缩性能比较。表 1 给出了其中 4 组 ECG 信号的编码压缩率和解码失真度数据。

表 1 编码器对 4 组 ECG 信号的压缩性能比较

| 性能指标         | ECG 信号 | Encode A II | Om it cD <sub>1</sub> | Om it cD <sub>1-2</sub> | Om it cD <sub>1-3</sub> | Om it cD <sub>1-4</sub> | Om it cD <sub>1-5</sub> |
|--------------|--------|-------------|-----------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 压缩率 (CR)     | 121    | 2 1855      | 3 1202                | 4 8699                  | 8 1007                  | 14 5061                 | 27 1749                 |
|              | 212    | 1 6681      | 2 3994                | 3 9118                  | 7 2090                  | 13 8634                 | 27 1749                 |
|              | 215    | 1 5528      | 2 4058                | 4 3190                  | 7 9185                  | 15 2113                 | 29 7596                 |
|              | 222    | 1 5684      | 2 4434                | 4 4416                  | 8 3406                  | 16 8497                 | 32 8876                 |
| 均方根误差比 (PRD) | 121    | 0 1104      | 0 1330                | 0 1710                  | 0 3719                  | 1 0484                  | 2 3701                  |
|              | 212    | 0 0871      | 0 1333                | 0 4595                  | 1 7330                  | 2 9986                  | 4 7131                  |
|              | 215    | 0 0570      | 0 1220                | 0 4162                  | 1 4926                  | 3 0920                  | 3 7068                  |
|              | 222    | 0 0831      | 0 1340                | 0 4903                  | 1 2600                  | 1 8013                  | 2 2506                  |

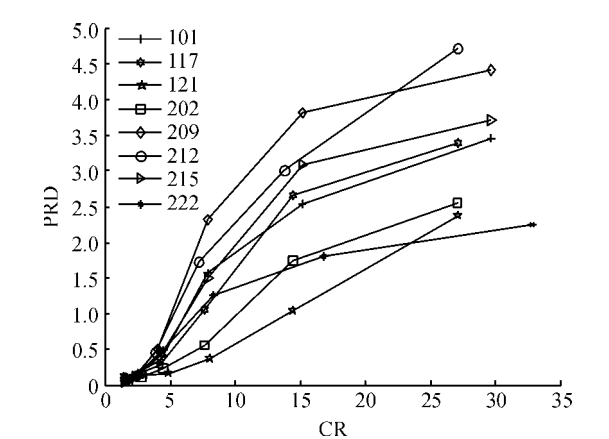


图 2 编码器对 8 组 ECG 信号的性能比较

本文所提出的算法在实现高压缩率的同时, 重构信号仍能准确还原原始信号的关键信息, 并且通过对高频小波分解系数的取舍, 可以在一定程度上消除或减弱原始信号中仪器引致的干扰噪声. 图 3 图 4 分别给出了 record 117 和 record 202 在不同压缩率下的重构信号, 可以看到在 CR = 8: 1 的条件下, 算法能够以接近或低于 1% 的失真率重构出比原始信号更平滑的波形, 并完整保留了用于临床诊断的关键信息.

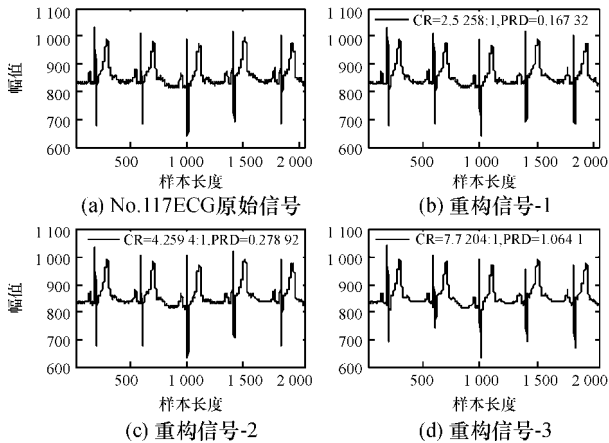


图 3 Record 117 在不同压缩率下的重构信号

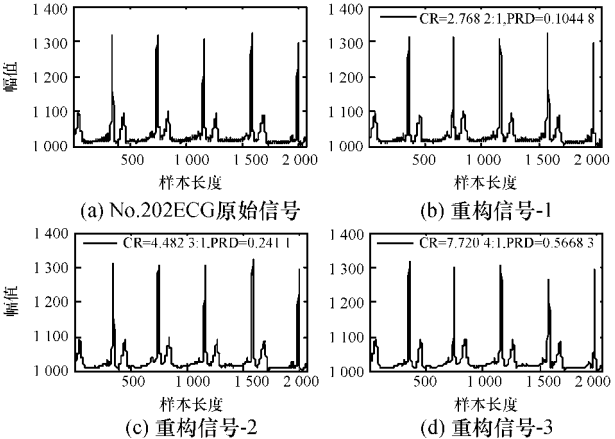


图 4 Record 202 在不同压缩率下的重构信号

本文编码方法与常用的零树编码方法相比, 输出码流只有 ‘0’、‘1’ 两种符号, 而且不需要进行零树分析, 这就使编码执行非常简单. 而且采取了分块优先编码的策略, 各系数子集可以并行编码, 从而可显著提高算法运行速度. 每个小波系数只扫描一次, 避免了重复编码, 压缩性能相对较优. 表 2 给出了本文算法与几种编码算法的性能比较数据.

表 2 本文算法与其它编码算法的压缩性能比较

| 算法                              | Records | PRD    | CR       |
|---------------------------------|---------|--------|----------|
| Proposed                        | 117     | 1 0641 | 7 723    |
|                                 | 117     | 2 6553 | 14 5061  |
|                                 | 119     | 3 4759 | 12 7349  |
| M ankandan et al <sup>[9]</sup> | 117     | 0 956  | 8 5: 1   |
|                                 | 117     | 2 089  | 18 7: 1  |
|                                 | 119     | 3 7166 | 21 8: 1  |
| Blanco et al <sup>[17]</sup>    | 117     | 1 1760 | 8 24: 1  |
|                                 | 117     | 2 5359 | 17 40: 1 |
|                                 | 119     | 5 0474 | 18 02: 1 |
| Benzid et al <sup>[18]</sup>    | 117     | 2 55   | 16 24: 1 |
|                                 | 119     | 5 1268 | 17 43: 1 |
| Lu et al <sup>[7]</sup>         | 117     | 1 18   | 8: 1     |
| H ilton <sup>[10]</sup>         | 117     | 2 6    | 8: 1     |
| Tai et al <sup>[19]</sup>       | 117     | 0 94   | 8: 1     |

### 3 结论

本文提出一种基于 Bior4-4 双正交小波的心电信号压缩算法,采取分块优先编码策略,实现自适应编码。经实验分析,该算法具有如下特点: (1) 算法实现简单,编码速度快; (2) 通过横向和纵向位扫描、带间编码获取高压缩率; (3) 实现嵌入式传输,即低频小波系数均嵌入到编码输出流前端。

### 参考文献:

- [1] Hamilton P S, Tompkins W J. Compression of the ambulatory ECG by average beat subtraction and residual differencing [J]. IEEE Trans Biomed Eng 1991, 38(3): 253-259.
- [2] 李刚, 刘蓉, 林凌. 一种基于复合编码的心电数据压缩算法 [J]. 中国生物医学工程学报, 2002, 21(6): 557-562.
- [3] 骆德汉, 许广桂, 邹宇华, 等. 基于多阶人工神经网络的 ECG 信号诊断模型研究 [J]. 仪器仪表学报, 2008, 29(1): 1-5.
- [4] Ghohm Hossein H, Luo Dehan, Reynolds K J. The comparison of different feed forward neural network architectures for ECG signal diagnosis [J]. Medical Engineering & Physics 2006, 28(4): 372-378.
- [5] 张浙亮, 吕维雪. 基于混合神经网络的 ECG 数据压缩 [J]. 浙江大学学报: 工学版, 2000, 34(1): 71-76.
- [6] Shapiro J M. Embedded image coding using zero trees of wavelet coefficients [J]. IEEE Trans Signal Process 1993, 41(12): 649-665.
- [7] Lu Z K, Kim D Y, Pearlman W A. Wavelet compression of ECG signals by set partitioning in hierarchical trees (SPIHT) algorithm [J]. IEEE Trans Biomed Eng 2000, 47(5): 849-856.
- [8] Lin W K, Bugress N. Lossless zero-tree coding for color image [C] // 32nd Asilomar Conference on Signals Systems and Computer CA, Monterey, 1998, 231-235.
- [9] Dohn A, Nguyen T Q, Tompkins W J. ECG compression using discrete symmetric wavelet transform [C] // The 17th IEEE International Conference on Medicine and Biology Canada Montreal QC, 1995.
- [10] Hilton M, Michael L. Wavelet and wavelet packet compression of electrocardiogram [J]. IEEE Trans Biomed Eng 1997, 44(5): 394-402.
- [11] Sabarimala M, Ankandan M, Dandapat S. Wavelet threshold based ECG compression using USZZQ and Huffman coding of DSM [J]. Biomedical Signal Processing and Control 2006(1): 261-270.
- [12] 王培康, 沈凤麟. 用 DCT 压缩 ECG 数据的新方法 [J]. 中国生物医学工程学报, 1995, 14(4): 332-338.
- [13] 郑凯梅, 余生晨. 基于小波变换的心电信号无损压缩算法 [J]. 山东生物医学工程, 2003, 22(2): 8-10.
- [14] 李梦龙, 康彬, 戚华溢, 等. 红外光谱小波压缩中最佳小波函数的选取 [J]. 化学通报, 2002, 65(9): 648-653.
- [15] Phoong S M, Kim C W, Vaidyanathan P P, et al. A new class of two channel biorthogonal filter banks and wavelet bases [J]. IEEE Trans Signal Process 1995, 43(12): 649-665.
- [16] Mallat S G. A theory for multiresolution signal decomposition: the wavelet representation [J]. IEEE Trans Pattern Anal Machine Intell 1989, 11(7): 674-693.
- [17] Blanco-Velasco M, Cruz-Roldan F, Godino-Lobente J I and Bamer K E. ECG compression with retrieved quality guaranteed [J]. IEEE Electron Lett 2004, 40(23): 1466-1467.
- [18] Benzid R, Marir F, Boussaad A, et al. Fixed percentage of wavelet coefficients to be zeroed for ECG compression [J]. IEEE Electron Lett 2003, 39(11): 830-831.
- [19] Tai S C, Sun C C, Yan W C. A 2-D ECG compression method based on Wavelet transform and modified SPIHT [J]. IEEE Trans Biomed Eng 2005, 52(6): 999-1008.

## Adaptive ECG Signal Compression Algorithm Based on Biorthogonal Wavelets

Zou Yu-hua, Luo De-han, Zhuang Jia-jun

(Faculty of Information and Engineering Guangdong University of Technology, Guangzhou 510006 China)

**Abstract** It proposes a novel ECG signal compression algorithm, based on biorthogonal wavelet Bior4-4 that can successively partition the wavelet coefficients in the space frequency domain and encode them using adaptive decimal to binary conversion. This algorithm is tested with 8 different records from MIT-BH arrhythmia database, an average percent root mean square difference as around 0.05% ~ 4.70% for an average compression ratio of 1.5:1 to 33:1 obtained. Experimental results show that the proposed algorithm is more efficient in compression and simpler in implementation and in computation.

**Key words** ECG signal compression, wavelet transformation, adaptive coding