

# 基于 PCA 特征和融合特征的 ECG 身份识别方法

杨向林, 严 洪, 任兆瑞, 宋晋忠, 姚宇华, 李延军

(中国航天员科研训练中心, 北京 100193)

**摘 要:** ECG 作为一种活体生物特征用于身份识别在国际上引起了广泛重视. 针对基于解析特征的 ECG 身份识别方法对特征点检测精度要求较高的缺点, 提出一种仅需 R 波峰值点检测的 ECG 身份识别方法, 该方法通过有针对性的设定相应阈值, 将 PCA 特征和小波融合特征方法相结合. 实验结果表明该方法优于 PCA 特征方法、波形特征方法和小波特征方法, 既减少了特征点检测的复杂性和特征点检测不准确带来的误差, 又可获得较高的识别率, 是一种实时、高效算法.

**关键词:** 主成分分析; 小波分解; 融合特征; 心电图; 身份识别

**中图分类号:** TP301.6 **文献标识码:** A **文章编号:** 1673-4785(2010)05-0458-06

## A method based on the PCA feature and fusion feature for ECG human identification

YANG Xiang-lin, YAN Hong, REN Zhao-rui, SONG Jin-zhong, YAO Yu-hua, LI Yan-jun

(China Astronaut Research and Training Center, Beijing 100193, China)

**Abstract:** As a new biometric for identification, ECG attracted widespread attention in the international community. The method was based on the analytic feature for ECG identification and required high precision for fiducial points detection. To overcome this disadvantage, a method, in which only the peak point of the R-wave detection is needed, was proposed. As for setting the relevant threshold, this method combined the PCA feature method and fusion feature method based on wavelet decomposition. The experiment demonstrates that the method proposed in this paper is better than the PCA feature method, waveform feature method, and wavelet feature method. This method, which not only reduces the complexity and error of the fiducial points detection but also achieves high accuracy, is a realistic and efficient algorithm.

**Keywords:** principal component analysis (PCA); wavelet decomposition; fusion feature; electrocardiogram (ECG); identification

1903 年荷兰医学家 Einthoven 发明心电图仪以来, 心电图 (electrocardiogram, ECG) 的发展已有 100 多年, 但应用范围仍局限于医学诊断和治疗. 2001 年瑞典学者 I. Biel 等拓宽了 ECG 的应用范围, 提出利用 ECG 作为生物特征进行身份识别, 该技术得到国际社会广泛关注, 相关学者对其进行了研究<sup>[1-9]</sup>. 其中李中建等从医学角度选用 6 000 例样本对 ECG 身份识别的可行性进行了系统研究, 通过不同环境季节 (温暖和寒冷环境)、心率、时间 (1 年以上)、药物 (心得安等影响心率的药物) 等分组实验, 研究表

明 ECG 满足生物特征识别的惟一性和稳定性要求, 并成功识别 6 000 例个体身份<sup>[2-3]</sup>. 2001 年 Irvine 等在读书、驾驶等 7 种状态下进行实验, 研究表明多生理状态下 (心率明显变化情况下) ECG 身份识别是可行的<sup>[4]</sup>.

ECG 是心肌细胞动作电位在体表的综合效应, 波形变化具有很强的规律性, 为一种准周期信号. 典型 ECG 每一周期由 P 波、QRS 波、T 波和 U 波组成. P 波反映心房肌的除极过程, QRS 波和 T 波分别反映了心室肌的除极和复极过程, U 波产生机理不明, 心房肌复极过程被 QRS 波掩盖而无法观测<sup>[5]</sup>.

ECG 身份识别技术假设 PQRST 波形在一定的时期内保持相对稳定, 对于正常人该假设成立, 即使在

焦虑、压力、运动等心率变化情况下QRS波形仍然保持稳定,即ECG满足生物身份识别特征稳定性的要求。同时受体型(例如肥胖)、年龄、体重、情绪、性别、心脏位置、大小、心脏几何形状、胸部构造、心脏生理特征等影响,个体ECG差异较大<sup>[5]</sup>,因此ECG满足生物身份识别特征惟一性的要求。即同一人ECG波形变化的内部差异小于不同人ECG波形间个体差异,因此ECG可作为一种生物身份识别特征。

ECG作为一种活体身份识别特征,有着较高的防伪性,既可单独用于身份识别,也可以和指纹等其他生物特征相结合开发多生物特征识别系统,它避免了指纹等其他身份识别技术被模仿或窃取的隐患。ECG为一维信号、数据量小、处理简单,减少了计算存储等资源消耗,而且ECG采集方便,甚至可以直接在2个指尖采集<sup>[5]</sup>。

国内外诸多研究人员对ECG身份识别进行了大量研究,提出了诸多算法。2001年L. Biel利用ECG信号的30个幅度和间期为特征,通过特征矩阵的相关运算减少冗余特征,最后选择10个间期和幅度特征,采用有监督的模式识别方法进行ECG身份识别<sup>[1]</sup>。2004年Palaniappan除了采用ECG信号时域波形的间期和幅度特征外,还采用了波形因数特征,利用多层感知器——后向传播神经网络和简化模糊ARTMAP(fuzzy adaptive resonance theory map)进行分类识别<sup>[6]</sup>。2006年Wang除了利用时域波形的幅度和间期特征外,还利用了主成分分析(principal component analysis, PCA)方法降维提取特征,利用 $K$ 阶邻接距进行身份识别<sup>[7]</sup>。以上基于解析特征的ECG身份识别方法在实验中取得了较好效果<sup>[5]</sup>,但对特征点定位精度要求较高,要求检测P波、QRS波和T波中的多个特征点。然而实际中很多人的ECG波形和理想ECG波形差异较大,例如无P波、双峰值T波和P波、T波倒置等情况,设计一种多个特征点定位的普遍算法一直是ECG信号处理中的难点和热点。本文提出了一种仅需R波峰值点检测的ECG身份识别方法,将基于PCA特征方法和基于小波分解融合特征的方法相结合,这样既保证了识别正确率,又可同时降低误拒率和误识率。

## 1 基于PCA特征和融合特征的身份识别方法

ECG身份识别属于模式识别范畴,ECG身份识别包括ECG预处理、特征提取、分类识别等过程。本

文方法流程如图1所示,先将输入ECG信号进行预处理,然后采用基于PCA的特征提取方法进行ECG身份识别粗分类,找出被试者的所有可能候选模板,减小模板库,然后采用基于小波分解的特征融合方法在减小的模板库中进行分类识别。



图1 ECG身份识别流程框图

Fig.1 Flow chart for the ECG identification

ECG身份识别系统的主要评价指标除了识别正确率之外,还有误拒率(false reject rate, FRR, 指拒绝接受应该接受的身份的概率)和误识率(false accept rate, FAR, 指接受不该接受的身份的概率)。这二者之间存在一定矛盾,FRR降低,则FAR升高,反之FRR升高,FAR降低,因此需要根据现实情况寻求某种折衷。本文为克服FRR和FAR之间的矛盾,在基于PCA特征的身份识别方法中设置较低的FRR,避免可能的候选者被拒绝,在基于小波分解的融合特征身份识别方法中设置较高的FRR,这样避免接受非法的候选者。将2种方法相结合,使FRR和FAR同时降低,提高ECG身份识别的正确率。

ECG信号的消噪和特征点检测采用参考文献[10]中的方法。其中特征点检测算法仅用参考文献[10]中的三样条小波检测R波峰值,避免了检测Q波、S波峰值点等问题。

### 1.1 基于PCA的ECG身份识别方法

PCA方法是一种考察多个数值变量间相关性的多元统计方法,它通过少数几个主成分来解释多变量的协方差结构,即将高维数据降为低维数据,保留原始变量信息的同时,去掉数据间的冗余信息,既减小了数据量,又突出了数据特征。图2为模板库中30个试验者ECG信号的PCA特征(横坐标为PCA特征向量序列,纵坐标为PCA特征向量值),30个试验者的PCA特征差异较大。同时ECG信号的PCA特征稳定,如图3所示,图中带方框的实线和带圆圈的虚线分别表示模板库中试验者A和另外4个试验者的PCA特征,虚线为试验者A不同时刻的PCA特征。从图3中可以看出,试验者不同时刻的PCA特征与模板库中试验者A的PCA特征基本重合,并且与模板库中其他PCA特征差异较大,因此ECG信号的PCA特征可用于身份识别<sup>[8-9]</sup>。

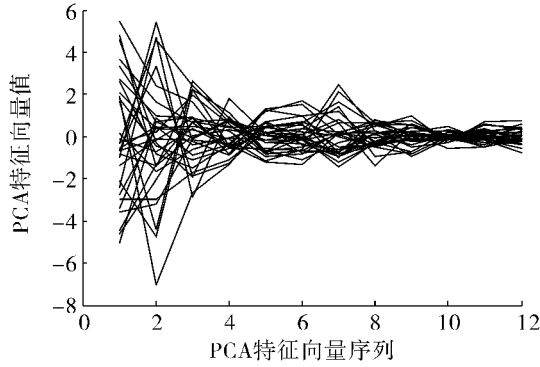


图2 模板库中30个试验者的PCA特征

Fig. 2 PCA features of 30 subjects in template database

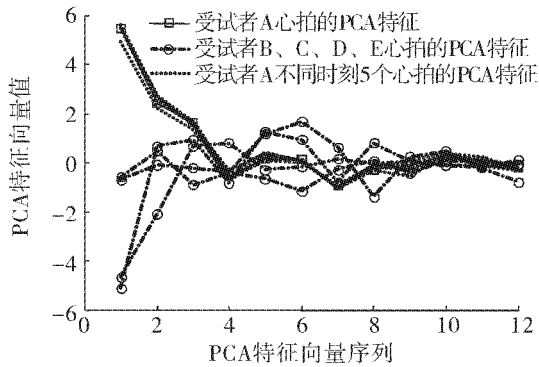


图3 ECG的PCA特征

Fig. 3 PCA features of ECG

算法过程如下所述:

1) 假设训练集  $Z = \{Z_i\}_{i=1}^C$  包含  $C$  类, 每一类  $Z_i = \{z_{ij}\}_{j=1}^{C_i}$  中包含  $C_i$  个心拍, 总共含有  $N = \sum_{i=1}^C C_i$  心拍,  $z_{ij}$  为一个心拍, 所有心拍都是以 R 波峰值点为基准点对齐, 心拍长度为  $N$ , 小于长度  $N$  的填零补齐.

2) 计算训练集的协方差矩阵:

$$S_{\text{cov}} = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} (z_{ij} - \bar{z})(z_{ij} - \bar{z})^T.$$

式中:  $S_{\text{cov}}$  为  $N \times N$  方阵,  $\bar{z} = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} z_{ij}$  是训练集样本心拍的平均心拍.

3) 计算协方差矩阵  $S_{\text{cov}}$  的特征值和特征向量, 并按从大到小顺序对特征值进行排序, 得到特征值序列  $\lambda_i$  和对应的特征向量序列  $\psi_i, i=1, 2, \dots, N$ .

4) 出于降维目的, 取  $M$  个主成分, 且  $M < N$ , 定义前  $M$  个主成分的累计贡献率为

$$\eta = \sum_{i=1}^M \lambda_i / \sum_{i=1}^N \lambda_i.$$

通过设定的累计贡献率选择主成分个数  $M$  的取值. 累计贡献率的大小反映了前  $M$  个主成分对原数据

所含信息量保留的多少, 通过实验,  $\eta$  取值为 85% 即可.

5) 令  $\psi = [\psi_1 \psi_2 \dots \psi_M]$  为变换矩阵, ECG 信号的 PCA 特征为

$$y_{ij} = \psi^T (z_{ij} - \bar{z}).$$

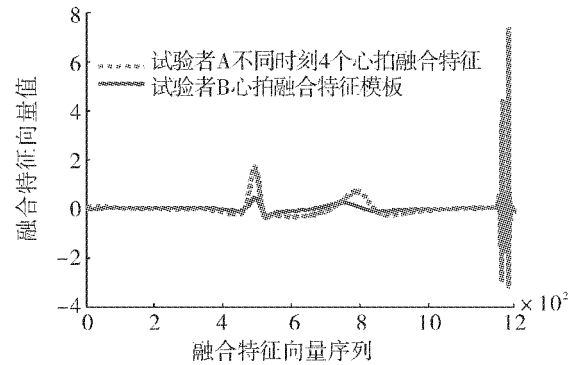
6) 构建 ECG 身份识别模板数据库为  $Y = [y_1 y_2 \dots y_c]$ .

7) 识别过程, 通过验证 ECG 构建心拍阵列  $X_i = \{x_{ij}\}_{j=1}^{C_i}$ , 利用  $t_{ij} = \psi^T (x_{ij} - \bar{z})$  从 ECG 心拍中提取 PCA 特征, 将提取的 PCA 特征和模板库中的 PCA 特征进行比较, 选择模板库中欧氏距离最小的模板为识别结果.

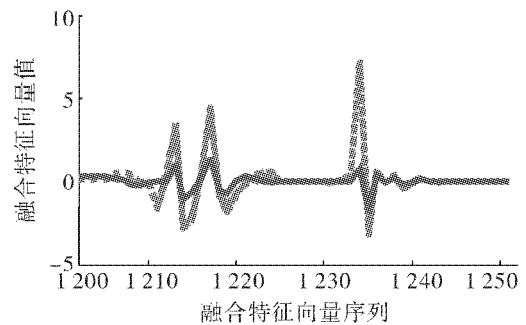
8) 将  $X_i$  中的所有心拍采用上述方法进行分类识别, 将  $C_i$  次分类识别结果的所有模板作为候选模板. 再采用基于融合特征的 ECG 识别方法在候选模板库中进行最终分类识别.

## 1.2 基于小波分解融合特征的 ECG 身份识别方法

不同试验者 ECG 的波形和小波特征差异较大, 并且特征稳定, 对波形特征和小波特征进行融合, 融合特征向量的差异显著增强, 如图 4 所示. 采用该特征向量进行身份识别抗噪声能力增强, 并且在心率变异性下仍保持较高的识别率<sup>[10]</sup>.



(a) ECG 的融合特征总体图



(b) (a)中局部放大图

图4 ECG 的融合特征图

Fig. 4 Fusion features of ECG

[illegible]

基于小波分解融合特征的 ECG 身份识别方法在缩小后的候选模板库中进行最终分类识别,通过对 30 名受试者 ECG 模板信号的融合特征向量进行相关系数阈值训练.融合特征向量的相关系数如图 5 所示,融合特征的相关系数阈值通过有监督学习设为 0.80.对 35 名受试者的另外 40 段数据在缩小后的候选模板库中进行身份识别验证,正确率为 97.5%.将本文方法和 Irvinea 所提的 PCA 特征方法、波形特征方法以及 Chan 所提的小波特征等方法进行比较<sup>[11-12]</sup>,如表 2 所示,可以看出该方法优于其他 3 种方法,识别率进一步提高,并且算法简单

快速,对噪声干扰影响较小,自适应性更强.

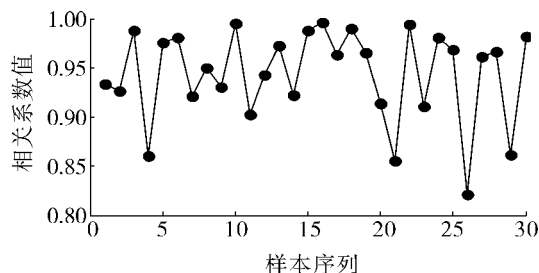


图 5 30 个模板基于小波分解的融合特征的相关系数

Fig.5 The correlation coefficients of fusion feature based wavelet decomposition for 30 templates

表 2 35 个试验者的 40 段 ECG 信号身份识别实验结果(  $n = 35$  )

Table 2 The experiment results of the identification for 40 sessions from 35 persons (  $n = 35$  )

身份识别方法	阈值	拒识率/%	误识率/%	识别率/%
PCA 特征		7.5	25.0	67.5
波形特征	0.78	5.0	12.5	82.5
小波特征	0.85	2.5	10.0	87.5
本文方法	0.80	0.0	2.5	97.5

### 3 结束语

本文提出了一种仅需检测 R 波峰值点的 ECG 身份识别方法,该方法首先采用基于 PCA 特征身份识别方法进行粗分类,选出模板库中所有可能的候选模板,然后针对少量的候选模板采用基于融合特征的 ECG 身份识别方法进一步识别分类.由于有针对性设定相应阈值,该方法既避免了多个特征点检测带来的繁琐和误差,又同时降低了 FAR 和 FRR.将本文方法与基于 PCA 特征、波形特征和小波特征的 ECG 识别方法相比较,该方法优于其他 3 种方法,准确性高,实时性好.

ECG 作为一种生物特征用于身份识别有着广阔的应用前景.本文算法取得了较好的实验效果,但只是在小样本量条件下进行的. ECG 身份识别技术的研究还处于起步阶段,还有诸多问题需要进一步研究.如:研究方便、有效的用于 ECG 身份识别的信号采集方法;研究 ECG 信号的变化(如心率变化、波幅变化)、电极位置偏差以及心脏突发等对 ECG 身份识别的影响及相应的有效 ECG 身份识别方法;另外普及 ECG 身份识别在实际中的应用,评价 ECG 身份识别算法的有效性,建立有效、科学的 ECG 身份识别标准数据库也是亟待解决且具有重大意义的研究工作.

### 参考文献:

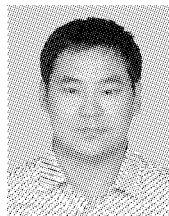
- [1] BIEL L, PETERSSON O, PHILIPSON L, et al. ECG analysis: a new approach in human identification[J]. IEEE Transactions on Instrumentation and Measurement, 2001, 50(3): 808-812.
- [2] 李中健,王庚勤,贾耀勤,等.应用心电图检查技术在活体个人识别中的研究[J].中国实用医刊,2008,35(4): 23-25.  
LI Zhongjian, WANG Gengqin, JIA Yaoqin, et al. Investigation on the possibility of electrocardiography in the identification of individuals[J]. Chinese Journal of Practical Medicine, 2008, 35(4): 23-25.
- [3] 李中健,井艳,张巧玲,等.活体个人心电图波形唯一性探讨[J].医药论坛杂志,2008,29(4): 45-46.  
LI Zhongjian, JING Yan, ZHANG Qiaoling, et al. The discussion of the uniqueness of ECG wave[J]. Medicament Forum Magazine, 2008, 29(4): 45-46.
- [4] IRVINE J M, WIEDERHOLD B K, GAVSHON L W, et al. Heart rate variability: a new biometric for human identification[C]//International Conference on Artificial Intelligence. Las Vegas, USA, 2001: 1106-1111.
- [5] 杨向林,严洪.基于心电图身份识别技术综述[J].仪器仪表学报,2009,29(增刊): 695-699.  
YANG Xianglin, YAN Hong. Survey on ECG based human identification technique[J]. Chinese Journal of Scientific Instrument, 2009, 29(Suppl): 695-699.

- [6] WANG Y, PLATANIOTIS K N, HATZINAKOS D. Integrating analytic and appearance attributes for human identification from ECG signal[C]//Proceedings of Biometrics Symposiums. Baltimore, USA, 2006: 1-6.
- [7] PALANIAPPAN R, KRISHNUN S M. Identifying individuals using ECG beats[C]//International Conference on Signal Processing and Communications (SPCOM). Bangalore, India, 2004: 569-572.
- [8] CASTELLS F, LAGUNA P, BOLLMANN A, et al. Principal component analysis in ECG signal processing: Article ID 74 580[J]. EURASIP Journal on Advances in Signal Processing, 2007(1): 98.
- [9] 张敏,李陶深,钟淑瑛. 基于 MATLAB 的主成分分析方法(PCA)的实现[J]. 广西大学学报:自然科学版, 2005, 30(S2): 74-77.  
ZHANG Min, LI Taoshen, ZHONG Shuying. Implementation of principal components analysis method based on Matlab[J]. Journal of Guangxi University: Natural Science Edition, 2005, 30(S2): 74-77.
- [10] 杨向林,严洪,李延军,等. 基于小波分解和数据融合方法的 ECG 身份识别[J]. 航天医学与医学工程, 2009, 22(4): 296-301.  
YANG Xianglin, YAN Hong, LI Yanjun, et al. Human identification based on ECG with wavelet decomposition and data fusion method[J]. Space Medicine & Medical Engineering, 2009, 22(4): 296-301.
- [11] IRVINE J M, ISRAEL S A, SCRUGGS W T, et al. Eigen

pulse: robust human identification from cardiovascular function[J]. Pattern Recognition, 2008, 41: 3427-3435.

- [12] CHAN A D C, HAMDY M M, BADRE A, et al. Wavelet distance measure for person identification using electrocardiograms[J]. IEEE Transactions on Instrumentation and Measurement, 2008, 57(2): 248-253.

#### 作者简介:



杨向林,男,1984年生,研究实习员,主要研究方向为生物医学信息处理和模式识别.负责承担中国航天医学工程预先研究项目1项,申请国家发明专利1项,发表学术论文13篇.



严洪,男,1962年生,研究员,主要研究方向为生物医学信号处理和图像处理.



任兆瑞,男,1978年生,助理研究员,主要研究方向为科研管理.

## 中文核心期刊——《计算机工程与应用》2011年征订启事

《计算机工程与应用》杂志是由信息产业部华北计算技术研究所主办的、面向中高级计算机专业工作者的学术刊物,系中国计算机学会会刊、中国电子学会一级会刊,计算机工程与应用学会学报,计算机类中文核心期刊,中国期刊方阵双效期刊,被列为“中国科学引文数据库来源期刊”、“中国学术期刊综合评价数据库来源期刊”,《中国期刊网》、《中国学术期刊(光盘版)》、英国 Inspec/SA、俄罗斯《文摘杂志》、美国剑桥科学文摘(CSA)及《乌利希期刊指南》、波兰《哥白尼索引》、《日本科学技术社中国文献数据库》收录期刊.

本刊是一本面向计算机全行业的综合性学术刊物,覆盖面宽、信息量大、报道及时是本刊的服务宗旨.主要栏目:博士论坛;理论研究;产品、研发、测试;网络、通信与安全;数据库、信号与信息处理;图形图像、模式识别;工程与应用.

多年来,本刊坚持走学术与实践相结合的道路,在内容上既注重理论的先进性又兼顾实用技术的广泛性,在促进学术交流的同时,推进了科技成果的转化.计算机科技工作者的良师益友,计算机工程技术人员的必备手册.

读者对象:计算机相关专业科研人员,工程项目决策、开发、设计及应用人员,大专院校师生.

订阅方式:本刊为旬刊,大16开,每期信息量60余万字,每月1日、11日、21日出版,邮局订阅代号:82-605,每期定价27.5元,全年36期总订价990元,全国各地邮局均可订阅.

通讯地址:北京市619信箱26分箱 邮编:100083

联系电话:010-51615542

咨询信箱:tjit@public2.bta.net.cn

主页:http://www.ceaj.org

户名:《计算机工程与应用》杂志社

账号:805903228608094001

开户银行:中国银行北京北极寺支行