

小波变换和支持向量机相融合的 ECG 身份识别算法

彭 辉^{1,2}

(1 浙江大学 计算机学院, 浙江 杭州 310018; 2 浙江警官职业学院, 浙江 杭州 310018)

摘 要: 心电图信号(ECC)受到设备、检测者心情等影响,含有噪声且特征具有高维性,为了提高身份识别的正确率,提前出一种小波变换和支持向量机相融合的 ECG 身份识别算法. 首先小波变换对采集的心电图信号进行预处理,消除各种噪声;然后采用相关分析对 ECG 特征进行降维处理,消除冗余特征;最后将特征向量输入到支持向量机进行训练,建立基于 ECG 的身份识别模型. 仿真结果表明,该方法提高了身份识别正确率和速度,有一种有效的身份识别算法.

关键词: 心电图信号;身份识别;小波变换;相关分析;支持向量机

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2013)03-0152-04

A Method Based on Wavelet Transform and SVM of ECG Human Identification

PENG Hui^{1,2}

(1 Department of Computer, Zhejiang University, Hangzhou 310018, China;

2 Zhejiang Police Vocational Academy, Hangzhou 310018, China)

Abstract: The electrocardiography signal includes noise and has high dimension due to the acquisition of equipment, testing the mood, in order to improve the identification accuracy, this paper proposes a new ECG signal identification method based on wavelet transform and support vector machine. Firstly, wavelet transform is used to eliminate the nose of ECG signal, and then the correlation analysis method is used to reduce the feature dimension; lastly, support vector machine is used to build ECG signal identification model. The simulation results show that the method improves the identification accuracy and speed, it is efficient identification algorithm.

Key words: ECG signal; identification; wavelet transform; correlation analysis; SVM

1 引言

传统身份识别方法有口令、证件、密码等,它们存在易复制、遗忘和易丢失等缺陷,法不能满足信息安全的要求^[1]. 生物特征的身份识别技术具有携带方便、不易伪造等优点,成为当前身份识别的主要研究方向. 生物特征有指纹、掌形、人脸、虹膜、和步态等,这些生物特征用于身份识别时,存在各自缺陷,如手指有伤痕时,拒识率高;视网膜识别技术伤害眼睛^[2-3]. 大量研究表明,心电图(electrocardiography, ECG)具有惟一性,易采集,不易复制,提供了一种新的身份识别技术,成为研究的热点^[4].

国内外学者对基于 ECG 的身份识别进行了相关研究,提出了一些基于 ECG 信号的身份识别方法^[5]. 然而 ECG 信号存在噪声信号,传统消噪方法无法准确消除噪声干扰. 在 ECG 的身份识别过程中,首先要提取多个特征点,然后采用一定方法消除特征间的冗余信息,最后采用分类器进行识别^[6]. 因此如何选择消除冗余信息和分类器设计也至关重要. 目前主成分分析提取特征法易破坏特征之间相关性^[7]. 神经网络等识别方法,在样本较大的时候才能保证获得较高的身份识别率,增加了识别成本,且基自身存在过拟合、网络结构复杂等缺陷^[8-9]. 支持向量机作为一种新的机器学习算法,要求小样本小,

解决了“维数灾”等难题,较好的克服了神经网络存在的不足,十分适合于用于 ECG 的身份识别^[10].

为了提高 ECG 的身份识别正确率,提出一种小波变换、相关性分析和支持向量机相结合的 ECG 身份识别方法,并通过具体心电图身份数据库进行仿真实验,验证其有效性和可行性.

2 ECG 的身份识别算法

2.1 ECG 信号预处理

ECG 的身份识别是指通过对心电图进行分析,以达到鉴别人的身份的目的. 在 ECG 信号中,由于工频和肌电对 ECG 信号采集的干扰,会产生高频噪声,工频干扰噪声的频谱为 50Hz,肌电干扰噪声的频谱为 5~2000Hz,其与 ECG 信号的频段重叠部分较大. 同时由于其它因素影响,如基线漂移等,产生一些低频噪声. 设 $f(t)$ 采集到的原始 ECG 信号, $e(t)$ 为噪声信息,那含有噪声的 ECG 信号可以描述为:

$$s(t) = f(t) + \sigma * e(t)$$
(1)

式中, σ 为噪声水平.

要获得较高的身份识别率,最为关键就是消除信号中的噪声,但是由于 ECG 信号不仅包括了低频噪声,而且包括了高频噪声,而传统除噪算法无法全面消除噪声.

2.1.1 消除 ECG 信号中高频噪声

(1) 采用小波变换对 ECG 信号进行多尺度分解,此处小波分解尺度为 5.

(2) 采用阈值法对多尺度小波系数进行滤波,消除 ECG 信号中的高频噪声.

(3) 用消除了高频噪声干扰的小波系数重构 ECG 信号,得到没有高频噪声的 ECG 信号.

消除 ECG 信号中高频噪声的具体流程如图 1 所示.

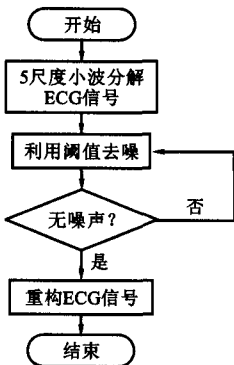


图 1 ECG 信号中的高频噪声消除流程

2.1.2 消除 ECG 信号中低频噪声

(1) 采用小波变换对 ECG 信号进行多尺度分

解,此处小波分解尺度为 8.

(2) 将低频噪声对应的小波系数置零进行滤波,消除 ECG 信号中的低频噪声.

(3) 用消除了低频噪声干扰的小波系数重构 ECG 信号,得到没有噪声的 ECG 信号.

消除 ECG 信号中高频噪声具体流程如图 2 所示.

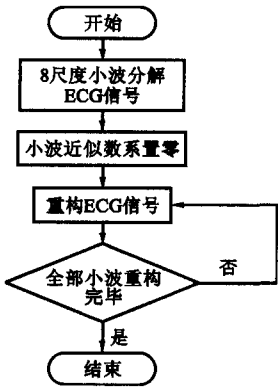


图 2 ECG 信号中的低频噪声消除流程

2.2 ECG 信号特征提取

ECG 信号包括 QRS 波、P 波和 T 波特征,其中, QRS 波特征由 Q 波起点、S 波终点、R 波的峰值点、Q 波和 S 波段峰值点组成; P 波和 T 波的特征点的包括 P 波、T 波峰值点及其起点、终点. 首先采用小波对 EDG 信号进行多尺度小波分解,然后对 QRS 波特征点进行提取,最后将 R 波峰值点的提取为关键,以 R 波峰值点为基准,提取 P 波和 T 波的特征点. 具体流程图 3 所示.

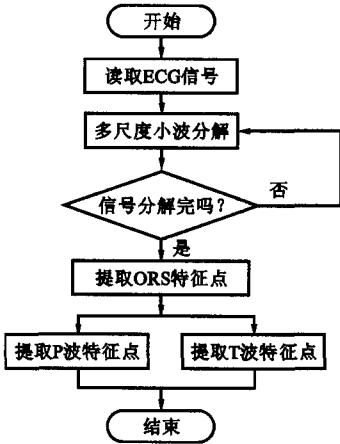


图 3 ECG 信号特征提取

2.3 降低 ECG 信号特征维数

经过特征提取算法获取的原始特征集,其包含 12 个特征,这些特征之间有一定的重复信息,如果全部输入到 ECG 信号分类器进行身份识别,会使识别效率降低,且会对识别正确率产生不利影响. 相关

性分析法可以选择最有利于身份识别的 ECG 信号特征,降低 ECG 信号特征维数.具体实现步骤如下:

(1) 首先对各 ECG 特征的身份识别适应度函数值,并结合支持向量机分类器,对单个特征的分类能力进行判别.

(2) 各 ECG 特征间的相关系统根据式(2)和式(3)计算得到.

$$s_{ij} = \frac{1}{k-1} \sum_{L=1}^k \left(x_i^L - \frac{1}{K} \sum_{L=1}^K x_i^L \right) \left(x_j^L - \frac{1}{K} \sum_{L=1}^K x_j^L \right) \quad (2)$$

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ij} \cdot S_{ji}}} \quad (3)$$

式中, K 为样本总数, x_i^L 表示特征集中第 L 个样本的第 i 个特征值, r_{ij} 表示第 i 个特征和第 j 个特征间的线性相关系数,如果 i 与 j 相等,那么 $r_{ij} = 1$ 为自相关系数.

(3) 设置门限值 δ ,本研究设置 $\delta = 0.95$.当原始 ECG 信号特征间的相关系数 δ 时,那么就根据步骤(1)的结果,选择最优的 ECG 信号特征,剔除其余特征,对原始 ECG 信号特征进行降维,将选择的 ECG 信号特征构成一个的特征子集.

2.4 ECG 的身份识别分类器设计

通过上述步骤后,就获得了比较重要的 ECG 信号特征,可以将这些特征输入到 ECG 信号分类器进行学习,建立基于 ECG 的身份分类模型.当前分类器相当的,其中支持向量机作为一种新型的机器学习算法,非线性分类能力相当强,需要较小的 ECG 信号样本,因此本研究选择支持向量机作为 ECG 的身份识别分类器.

然而支持向量机用于 ECG 建模过程中,如果选择最优参数(C 、 p 、 g)至关重要,目前大多数采用经验方法确定最优参数,这会对 ECG 信号的身份识别结果产生不利影响,为此,本研究采用全局寻优能力强的遗传算法对支持向量机参数优化,解决支持向量机在实际应用过程参数选择问题.

3 仿真实验

3.1 仿真环境

CPU 为 P43.0G,内存配置为 1G,显示分辨率 1024 × 768,操作系统是 Windows 2000,开发工具 visualC++6.0.

3.2 ECG 信号降噪

采用小波变换对图 4 中 ECG 信号进行降噪处理,得到的结果如图 5 所示.对降噪前后的 ECG 信

号处理比较可知,降噪后的 ECG 信号比较平滑,且基线明显消除,这说明采用小波变换可以很好达到去除 ECG 信号中的噪声信息目的.

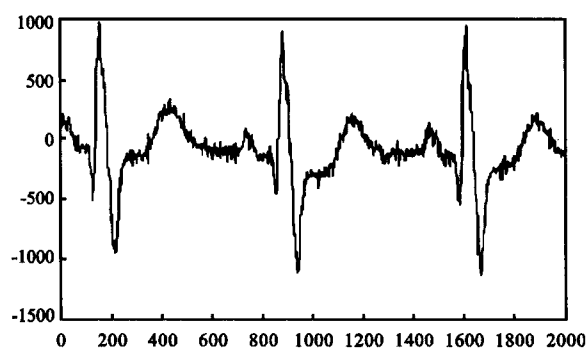


图 4 含噪声的 ECG 信号

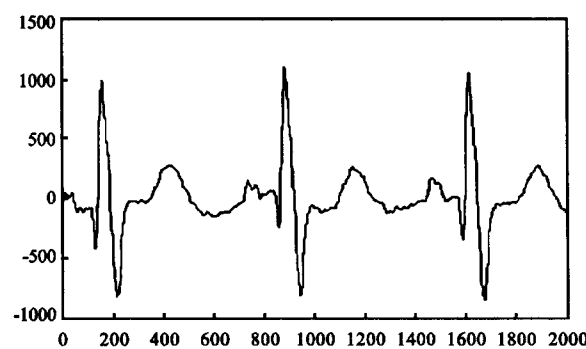


图 5 去噪后的 ECG 信号

3.3 ECG 信号特征和降维

首先提取 ECG 信号的 QRS 波、P 波和 T 波特征,然后采用相关分析对提取特征进行降维处理,得到的最优特征子集的特征数为 7 个.

3.4 结果与分析

对 ECG 信号样本进行降维后 7 个特征作为支持向量机的输入,实际 ECG 信号类别作为输出,进行训练,得到最优 ECG 信号识别模型的参数为 $C=100$, $g=1.25$, $p=0.144$,从而建立了 ECG 的身份识别模型.为使本算法的结果具有可比性,采用没有去噪的支持向量机模型(SVM1)、去噪但没有降维的支持向量机(SVM2)模型、去噪且降维处理,但没有网络搜索参数的支持向量机(SVM3)、去噪且降维处理的 BP 神经网络模型(BPNN).模型评价为识别正确率、误识率和拒识率.各模型的身份识别结果见表 1.

对表 1 中各模型的识别结果进行分析可知,在所有识别模型中,本研究提出的身份识别算法各项评价最优.可以得到如下结论:

(1) 相对于 SVM1,识别正确率提高了 6.5%,这说明采集原始 ECG 信号含有各种噪声信息,采用

小波变换可以有效的消除噪声信息对身份识别结果的不利影响,可以进一步提高了身份识别正确率.

表 1 不同模型的 ECG 信号身份识别性能比较

识别模型	正确率	误识率	拒识率
SVM1	82.5%	13.5%	4.0%
SVM2	97.5%	1.5%	1.0%
SVM3	98.0%	1.5%	0.5%
BPNN	80.5%	15.0%	4.5%
本研究模型	99.0%	1.0%	0.0%

(2) 相对于 SVM2,本研究身份识别模型的正确率提高了 1.5%,且拒识率为 0,本研究具有很好鲁棒性和可靠性,这比结果表明,提取的 ECG 信号特征之间具有一些重复信息,通过相关性分析可以消除一些无用或冗余特征,减少了计算复杂度,且可以提高了 ECG 信号的身份识别正确率.

(3) 相对于 SVM3,本研究的身份识别模型正确率也有所提高,这说明采用遗传算法比传统网格搜索算法可以找到更优的支持向量机参数,获得了性能更优的支持向量机,从而使识别性能有所提高,表明采用遗传算法对支持向量机的参数进行优化是更可行和有效的.

(4) 相对 BPNN,所以基于支持向量机的 ECG 信号身份识别模型的识别性能要好,这说明支持向量机很好的克服了神经网络的过拟合、泛化能力差等缺陷,十分适合于 ECG 信号的建模.

由于身份识别实时性要求比较高,因此采用仿真实验对模型的识别速度进行比较,得到的对比结果如图 6 所示.

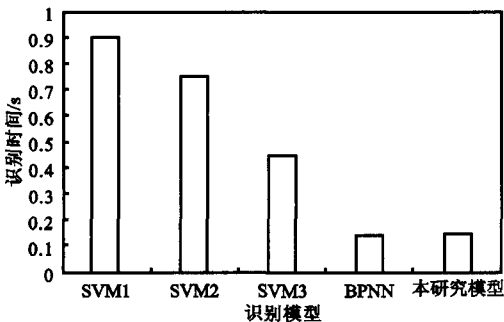


图 6 各模型的识别速度比较

从图 6 可知,本研究识别模型的识别速度最高,这表明经过消噪处理、降低特征维数以及分类器参数优化等操作,使基于 ECG 的身份识别模型获得较高的识别效率,拓宽了 ECG 信号在信息安全领域的应用范围.

4 结束语

ECG 信号具有唯一性、难以伪装等优点,是身份识别领域一个具有应用价格研究课题,本研究提出一种小波去噪、相关分析降维和支持向量机相融合的身份识别方法. 仿真结果表明,相对其它识别方法,本研究识别方法提高了身份的正确率,误识率最低,拒识率为 0,识别结果十分稳定. 然而基于 ECG 信号的身份识别还处理初级研究阶段,还有诸多问题有待进一步研究,如方便、有效的 ECG 信号采集方法、建立科学的 ECG 身份识别标准数据库等,这也是我们以后将要进行的研究重点.

参考文献:

[1] 田捷,杨鑫. 生物特征识别技术理论与应用[M]. 北京电子工业出版社, 2005.

[2] 陈添丁,郎燕峰. 单导程心电图的身份识别[J]. 计算机工程与设计, 2006,27(24):4770-4774.

[3] Biel L, Pettersson 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.

[4] Chiu C C,Chuang C M,Hsu C Y. A novel personal identity verification approach using a discrete wavelet transform of the ECG signal[C]// Proceedings of the 2008 International Conference on Multimedia and Ubiquitous Engineering Washington D C: IEEE,2008; 201-206.

[5] Agrafioti F, Hatzingakos D. ECG biometric analysis in cardiac irregularity conditions[J]. Signal, Image and video Processing, 2008,3(4):1863-1703.

[6] 邢国全. 自适应滤波器在 ECG 信号处噪中的应用[J]. 北京生物医学工程, 2005,24(3):209-210.

[7] 杨向林. 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.

[8] Rajoub B A. An efficient coding algorithm for the compression of ECG signals using the wavelet transform[J]. IEEE Trans Biomed Eng April, 2002,49(4):355-362.

[9] 师黎,朱民杰. 基于遗传算法优化 BP 神经网络在心电图身份识别中的应用[J]. 中国组织工程研究与临床康复,2010,14(43):8069-8072.

[10] 舒双宝,罗家融,徐从东,等. 一种基于支持向量机的人脸识别新方法[J]. 计算机仿真,2011,28(2):280-283.

作者简介:

彭辉 女,(1980—),硕士,讲师. 研究方向为计算机硬件开发、数据恢复与数字取证、图像处理.