

DOI:10.3969/j.issn.1673-4785.201204027

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20120803.1628.001.html>

脑电信号的小波变换和样本熵特征提取方法

张毅¹, 罗明伟¹, 罗元²

(1. 重庆邮电大学 自动化学院, 重庆 400065; 2. 重庆邮电大学 光电工程学院, 重庆 400065)

摘要:针对现有的采用单一的特征提取算法对运动想象脑电信号识别率不高的问题,提出一种结合小波变换和样本熵的特征提取方法.通过小波变换,把脑电信号进行3层分解,抽取对应于脑电 β 节律频带的小波系数的能量均值和能量均值差,并结合脑电信号的样本熵组成特征向量,使用支持向量机分类器对左右手运动想象脑电信号进行分类.结果表明,结合小波变换和样本熵的特征提取方法明显优于仅采用小波变换、样本熵以及其他传统的特征提取方法,得到的最高正确识别率为91.43%.

关键词:脑电信号;样本熵;小波变换;支持向量机;特征提取

中图分类号:TP18; R318 **文献标志码:**A **文章编号:**1673-4785(2012)04-0339-06

EEG feature extraction method based on wavelet transform and sample entropy

ZHANG Yi¹, LUO Mingwei¹, LUO Yuan²

(1. College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2. College of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Considering the issue of low recognition rate for electroencephalograph (EEG) signal of motor imagery by using current single feature extraction method, a feature extraction method based on wavelet transform and sample entropy is presented in this paper. The EEG signals are decomposed to three levels by wavelet transform and the average energy and its difference of wavelet coefficient corresponding to the β rhythm of EEG signals are computed. The feature vector is composed of the average energy, its difference of wavelet coefficient and sample entropy of EEG signals. Finally, the left-right hands motor imagery EEG signals are classified by a support vector machine classifier. The experimental results show that the feature extraction method combining wavelet transform and sample entropy is much better than the ways of only using wavelet transform, sample entropy, or others, and its highest recognition rate is 91.43%.

Keywords: electroencephalograph signal; sample entropy; wavelet transform; support vector machine; feature extraction

脑-机接口 (brain-computer interface, BCI) 是不依赖于大脑外周神经与肌肉系统,在人和计算机或外部设备之间建立起来的一种通信系统.它能够为肢体残疾但思维意识正常的患者提供一种新型的对外信息交流手段,并在残疾人康复、正常人辅助控制、娱乐等领域有着广泛的应用前景^[1-4].

特征提取是 BCI 研究的关键技术之一,常用的方法有 FFT^[5] (fast Fourier transform)、AR^[6-7] (auto-regressive)、AAR^[8] (adaptive auto-regressive)、小波变换^[9-11]、样本熵^[12] 等方法. FFT、AR 和 AAR 通过把幅度随时间变化的脑电信号变换为脑电功率随频率变化的谱图,从而提取出脑电信号的频域特征.然而这些方法都只适合分析平稳信号,对于非平稳的脑电信号,具有很大的局限性.小波变换是一种典型的时频分析法,具有多分辨率特性,在时域和频域均具有良好的分辨率,适用于非平稳信号处理.因此,

收稿日期:2012-04-28. 网络出版日期:2012-08-03.

基金项目:科技部国际合作项目(2010DFA12160);国家自然科学基金资助项目(51075420).

通信作者:罗明伟. E-mail:350874723@qq.com.

小波变换非常分析合适脑电信号. 然而, 大脑是一个典型的非线性系统, 脑电信号可以看作是它的输出, 小波变换却不能反映出脑电信号的非线性特征. 而样本熵是一种非线性分析法, 能通过度量脑电信号的复杂度来反映它的非线性特征, 但不能反映出脑电信号的时频特征. 为了尽量充分地提取出脑电信号的特征, 提高其正确的识别率, 本文提出了一种结合小波变换和样本熵的特征提取方法. 该方法通过结合小波变换提取出的时频特征和样本熵分析提取出的非线性特征, 得到了能反映出脑电信号时域、频域和非线性特征的特征向量. 利用支持向量机(support vector machine, SVM)对左、右手运动想象脑电信号进行分类, 得到的最高正确识别率为 91.43%, 明显高于仅采用小波变换、样本熵以及其他传统方法^[6-7]的正确识别率.

1 基于 Emotiv 的脑电采集系统

本文所使用的脑电采集仪是如图 1 所示的 Emotiv, 其采样频率为 128 Hz. 电极是按照国际 10-20 标准电极安放法安放, 其安放位置如图 2 所示.



图 1 Emotiv 脑电采集仪

Fig. 1 Emotiv EEG acquisition instrument

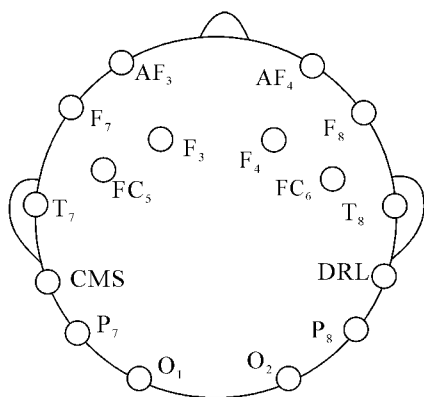


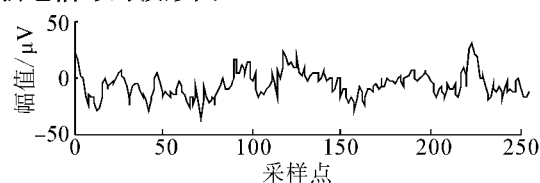
图 2 Emotiv 电极安放位置

Fig. 2 Electrode positions on Emotiv

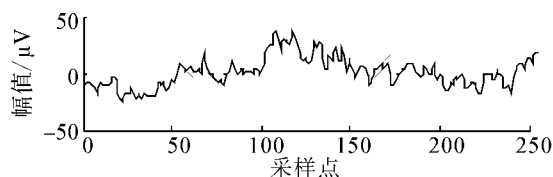
图中“CMS”和“DRL”为参考电极. 选取 4 个健

康受试者(年龄为 22 ~ 24 岁), 分别采集左、右手运动想象脑电信号. 以一个受试者进行左手运动想象为例, 对整个数据采集过程进行说明: 实验开始时, 让受试者静坐并保持放松状态, 当受试者听到口令时, 开始想象左手运动, 大约 1 min 后停止想象, 休息 2 s 后再次进行同样的操作, 总共进行 5 次实验. 按上述实验步骤, 分别采集 4 个受试者进行左、右手运动想象的脑电信号. 以持续 2 s 的脑电信号作为一组样本, 一个受试者进行一种运动想象可获得 150 组样本, 每组样本有 256 个数据.

图 3 是受试者进行左、右手运动想象的 FC₅ 通道脑电信号的波形图.



(a) 想象左手运动



(b) 想象右手运动

图 3 想象左、右手运动的 FC₅ 通道脑电信号的波形

Fig. 3 The EEG signals of FC₅ channel for imagery left-right hands movement

2 结合小波变换和样本熵的脑电信号特征提取

2.1 小波变换特征提取

人在想象单侧手运动时, 其对侧相应初级感觉运动皮层区的脑电 μ 节律 (8 ~ 12 Hz) 和 β 节律 (14 ~ 30 Hz) 节律幅值降低, 这种现象称为事件相关去同步(event-related desynchronization, ERD); 而同侧脑电 μ 节律和 β 节律幅度升高, 称为事件相关同步(event-related synchronization, ERS). 根据这一特征, 可使用 μ 节律和 β 节律来分析左、右手运动想象脑电信号. 而小波变换能把信号的整个频带划分为多个子频带, 因此可使用小波变换来分析左、右手运动想象脑电信号. 为了减少特征向量的维数, 本文仅分析 β 节律. 设 $x(n)$ 表示实验采集的 EEG 离散信号, 则 $x(n)$ 的离散小波变换定义为

$$C_{j,k} = 2^{-j/2} \sum_{n=-\infty}^{\infty} x(n) \bar{\psi}_{j,k}(2^{-j}n - k) = \langle x(n), \psi_{j,k} \rangle, j, k \in Z.$$

式中: $\psi(n)$ 为小波基函数, j, k 分别代表频率分辨率和时间平移量. 采用 Mallat 算法, 对信号进行有限层分解, 即

$$x(n) = A_L + \sum_{j=1}^L D_j. \quad (1)$$

式中: L 为分解层数, A_L 为低通逼近分量, D_j 为不同尺度下的细节分量. 设信号 $x(n)$ 的采样频率为 f_s , 则式(1)中的 $A_L, D_L, D_{L-1}, \dots, D_1$ 各分量所对应的子频带依次为 $[0, \frac{f_s}{2^{L+1}}], [\frac{f_s}{2^{L+1}}, \frac{f_s}{2^L}], [\frac{f_s}{2^L}, \frac{f_s}{2^{L-1}}], \dots, [\frac{f_s}{2^2}, \frac{f_s}{2}]$. 将信号进行小波分解时, 分解的层数将视具体信号的有用成分和采样频率而定. 本文分析的左、右手运动想象脑电信号的采样频率为 128 Hz, 信号的有用成分是 14 ~ 30 Hz 的 β 节律. 因此, 本文选用 db5 小波对脑电信号进行 3 层分解, 即 $x(n) = A_3 + D_3 + D_2 + D_1$, 则各分量对应的子频带见表 1.

表 1 小波分解的各层频带范围

Table 1 Frequency band range of each level of wavelet decomposition

分解信号	频带范围/Hz	分解的层数
D_1	32 ~ 64	1
D_2	16 ~ 32	2
D_3	8 ~ 16	3
A_3	0 ~ 8	3

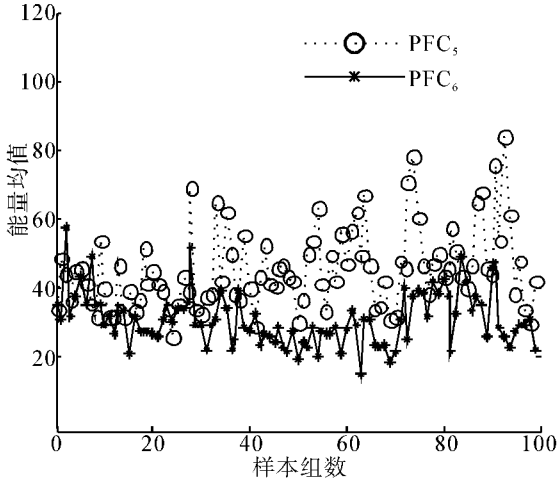
小波系数能表达信号在时域和频域的能量分布, 因此利用小波系数的能量能反映出脑电信号的时域和频域特征. 由表 1 可知 D_2 (16 ~ 32 Hz) 在脑电信号的 β 节律频带范围附近, 因此, 可提取对应于 D_2 频带的小波系数的能量均值作为特征量. 同时, 为了进一步突出想象单侧手运动引起的 FC_5 、 FC_6 通道脑电信号的幅值差异, 还提取了小波系数的能量均值差 PS 作为特征量, 即

$$PS = PFC_5 - PFC_6.$$

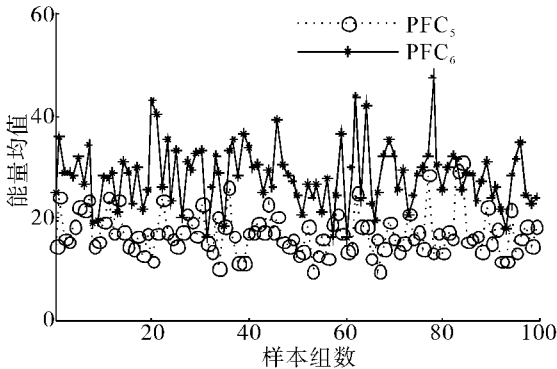
式中: PFC_5 为 FC_5 通道的能量均值, PFC_6 为 FC_6 通道的能量均值.

对左、右手运动想象任务, 各选取 100 组样本数据计算, 得到小波系数的能量均值和能量均值差分别如图 4 和图 5 所示. 从图 4 可以看出: 想象左手运动时, FC_5 通道的能量均值 PFC_5 大于 FC_6 通道的能量均值 PFC_6 ; 想象右手运动时, FC_5 通道的能量均值 PFC_5 小于 FC_6 通道的能量均值 PFC_6 . 因此, 脑电

信号经小波分解后, 其对应于 β 节律频带的小波系数的能量均能作为区分左、右手运动想象脑电信号的特征量. 从图 5 可知, 利用小波系数的能量值均值差提取得到的特征很好地反映出了左、右手运动想象脑电信号的差别.



(a) 想象左手运动



(b) 想象右手运动

图 4 小波系数的能量均值

Fig. 4 The average energy of wavelet coefficient

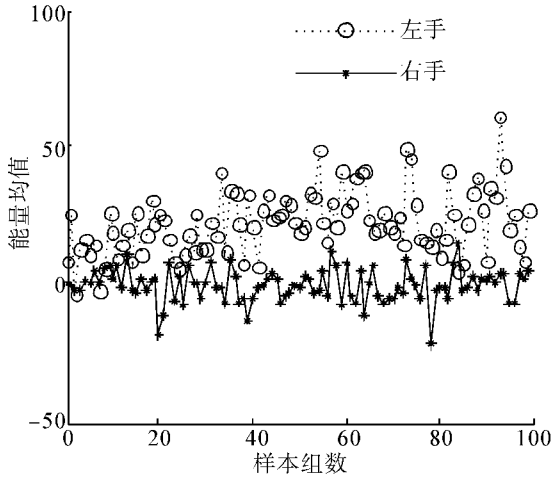


图 5 小波系数的能量均值差

Fig. 5 The average energy difference of wavelet coefficient

2.2 结合小波变换和样本熵的特征提取

小波变换能反映出脑电信号的时频特征,但不能反映出它的非线性特征. 而样本熵是一种非线性分析法,能通过度量脑电信号的复杂度来反映它的非线性特征,但却不能反映出它的时频特征. 针对这2种方法的不足,本文将结合小波变换和样本熵来提取脑电信号的特征,进而提高脑电信号的正确识别率.

样本熵是在近似熵的基础上提出的一种改进的复杂度测量方法^[13]. 它不但具备近似熵所有的优点,而且避免了近似熵中统计量不一致的问题.

设一维时间序列为 $\{X(i)\}, i = 1, 2, \dots, N$, 其样本熵可通过如下计算得到:

1) 将序列 $\{X(i)\}$ 按顺序组成 m 维矢量,即

$$Y_m(i) = [X(i) \ X(i+1) \ \dots \ X(i+m-1)], \\ i = 1, 2, \dots, N-m+1.$$

2) 对每一个 i 计算矢量 $Y_m(i)$ 与矢量 $Y_m(j)$ 之间的距离,即

$$d[Y_m(i), Y_m(j)] = \max_{k=0,1,\dots,m-1} |X(i+k) - X(j+k)|, \\ i, j = 1, 2, \dots, N-m+1; i \neq j.$$

3) 给定阈值 $r (r > 0)$, 对每一个 i 值统计 $d[Y_m(i), Y_m(j)]$ 小于 r 的数目,记为 $A_i^m(r)$,并记此数目与总的距离数目 $N-m$ 的比值为 $B_i^m(r)$,即

$$B_i^m(r) = A_i^m(r) / (N-m).$$

4) 求 $B_i^m(r)$ 对所有 i 值的平均值,记为 $B^m(r)$,即

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r).$$

5) 将序列 $\{X(i)\}$ 按顺序组成 $m+1$ 维矢量,再重复1)~4),得到 $B_i^{m+1}(r)$ 和 $B^{m+1}(r)$.

6) 序列 $\{X(i)\}$ 的样本熵为

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \{-\ln[B^{m+1}(r)/B^m(r)]\}.$$

在实际计算中,由于序列长度有限,因此最后得到序列长度为 N 时的样本熵的估计值

$$\text{SampEn}(m, r, N) = -\ln[B^{m+1}(r)/B^m(r)].$$

显然,样本熵的值与 m, r 和 N 有关. 通常取 $m = 2, r = 0.1 \sim 0.25 \text{ SD}$ (SD表示原始时间序列的标准差). 在本文中取 $m = 2, r = 0.2 \text{ SD}, N = 256$.

利用样本熵提取脑电信号的非线性特征的具体步骤为:首先对脑电信号加滑动时间窗,以1 s长度为滑动时间窗(即128点),计算脑电信号的样本熵,窗口每次移动一个采样点,并计算下1 s时间窗的脑电信号的样本熵,直到计算出最后1 s时间窗的脑电信号的样本熵为止,从而获得该样本数据中脑电信号样本熵的时间序列,然后将这一组样本熵

序列叠加平均,即获得一组样本数据的样本熵.

对左右手运动想象的 FC_5 通道脑电信号,各选取100组样本数据计算,得到 FC_5 通道脑电信号的样本熵如图6所示. 从图6知,想象左、右手运动的脑电信号的样本熵虽有部分重叠,但仍具有一定的区分度. 由此可见,样本熵在一定程度上能反映左、右手运动想象脑电信号的特征.

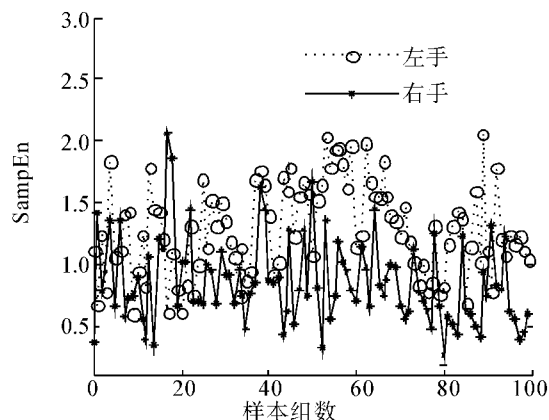


图6 FC_5 通道脑电信号的样本熵

Fig.6 The sample entropy of EEG signals of FC_5 channel

对采集的 FC_5, FC_6 通道的左、右手运动想象脑电信号,用样本熵特征提取方法可得到二维特征向量,用小波变换可得到三维特征向量,结合2种方法可得到五维特征向量,即为本文的特征提取向量.

2.3 SVM分类

SVM是一种基于统计学习理论的模式识别方法,能有效地避免传统分类方法所存在的过学习、维数灾难、局部极小等问题,并且在小样本条件下仍然具有良好的泛化能力. 本文将利用SVM对左、右手运动想象脑电信号进行分类. 首先选取SVM的核函数,然后利用训练样本计算出SVM的参数(如拉格朗日乘子 a 、最优超平面的法向量 w 和偏置值 b 等),并确定判别函数 $f(x)$,最后输入测试样本,根据 $\text{sgn}(f(x))$ 的值,输出类别. 如果 $\text{sgn}(f(x))$ 为1,则判别结果为想象左手运动的脑电信号;如果 $\text{sgn}(f(x))$ 为-1,则判别结果为想象右手运动的脑电信号.

3 实验结果分析

在采集的150组样本数据中,选取75组作为训练样本,剩下75组作为测试样本. 把用小波变换提取的三维特征向量、用样本熵提取的二维特征向量以及两者相结合的五维特征向量分别输入SVM分

类器中,得到不同受试者在不同的特征提取方法下的正确识别率见表2,其中SVM的核函数采用的是线性核函数。

表2 不同受试者在不同的特征提取方法下的正确识别率

Table 2 The right recognition rates of different subjects using different feature extraction methods %

特征提取方法	受试者1	受试者2	受试者3	受试者4
小波变换	85.33	83.33	88.00	80.67
样本熵	70.67	75.33	60.67	75.33
小波变换 + 样本熵	91.43	89.33	90.00	86.67

从表2可以看出,在对左、右手运动想象脑电信号的识别中,使用小波变换和样本熵相结合的特征提取算法得到的最高正确识别率为91.43%,仅使用小波变换和样本熵得到的最高正确识别率分别为88.00%和75.33%,而在文献[6-7]中,仅使用AR特征提取方法得到的最高正确识别率分别85.00%和75%。因此,使用小波变换和样本熵相结合的特征提取算法的正确识别率明显高于仅采用小波变换或样本熵的特征提取算法的正确识别率,同时也高于传统的使用AR特征提取方法的正确识别率。

4 结束语

本文提出了一种结合小波变换和样本熵的特征提取方法。分别用小波变换和样本熵对左、右手运动想象脑电信号进行了分析,并抽取出相应的特征组成特征向量,输入SVM分类器中,实现了左、右手运动想象脑电信号的分类。研究结果表明,结合小波变换和样本熵的特征提取算法对左右手运动想象脑电信号的正确识别率明显高于仅采用小波变换、样本熵以及其他传统特征提取算法的正确识别率,其最高正确识别率达到91.43%。因此,该方法在BCI中具有应用价值。

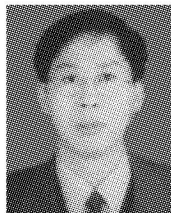
参考文献:

- [1] MUGLERR E, BENSCH M, HALDER S, et al. Control of an internet browser using the P300 event-related potential [J]. *Int J Bioelectromagn*, 2008, 10(1): 56-63.
- [2] 王斐, 张育中, 宁廷会, 等. 脑—机接口研究进展[J]. *智能系统学报*, 2011, 6(3): 189-199.
WANG Fei, ZHANG Yuzhong, NING Tinghui, et al. Research progress in a brain-computer interface [J]. *CAAI Transactions on Intelligent Systems*, 2011, 6(3): 189-199.
- [3] TANAKA K, MATSUNAGA K, WANG H. Electroencephalogram-based control of an electric wheelchair [J]. *IEEE Trans Robotics*, 2005, 21(4): 762-766.
- [4] REBSAMEN B, GUAN G T, ZHANG H H, et al. A brain controlled wheelchair to navigate in familiar environments [J]. *IEEE Trans Neural Syst Rehabil Eng*, 2010, 18(6): 590-598.
- [5] POLAT K, GÜNES S. Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals [J]. *Expert System with Applications*, 2008, 34(3): 2039-2048.
- [6] 徐宝国, 宋爱国. 单次运动想象脑电的特征提取和分类 [J]. *东南大学学报: 自然科学版*, 2007, 37(4): 629-633.
XU Baoguo, SONG Aiguo. Feature extraction and classification of single trial motor imagery EEG [J]. *Journal of South-east University: Natural Science Edition*, 2007, 37(4): 629-633.
- [7] 张毅, 杨柳, 李敏, 等. 基于AR和SVM的运动想象脑电信号识别 [J]. *华中科技大学学报: 自然科学版*, 2011, 39(suppl.2): 103-106.
ZHANG Yi, YANG Liu, LI Min, et al. Recognition of motor imagery EEG based on AR and SVM [J]. *Journal of Huazhong University of Science & Technology: Natural Science Edition*, 2011, 39(suppl.2): 103-106.
- [8] PFURSTCHER G, NEUPER C. Motor imagery and direct brain-computer communication [J]. *Proceedings of the IEEE*, 2001, 89(7): 1123-1134.
- [9] 李明爱, 王蕊, 郝冬梅. 想象左右手运动的脑电特征提取及分类研究 [J]. *中国生物医学工程学报*, 2009, 28(2): 166-170.
LI Ming'ai, WANG Rui, HAO Dongmei. Feature extraction and classification of EEG for imagery left-right hands movement [J]. *Chinese Journal for Biomedical Engineering*, 2009, 28(2): 166-170.
- [10] RAFIEE J, RAFIEE M A, PRAUSE N, et al. Wavelet basis functions in biomedical signal processing [J]. *Expert Systems with Applications*, 2011, 38(5): 6190-6201.
- [11] 徐宝国, 宋爱国, 费树岷. 在线脑机接口中脑电信号的特征提取与分类方法 [J]. *电子学报*, 2011, 39(5): 1025-1030.
XU Baoguo, SONG Aiguo, FEI Shumin. Feature extraction and classification of EEG in online brain-computer interface [J]. *Acta Electronica Sinica*, 2011, 39(5): 1025-1030.
- [12] 周鹏, 葛家怡, 曹红宝, 等. 基于样本熵的运动想象分类研究 [J]. *信息与控制*, 2008, 37(2): 191-196.
ZHOU Peng, GE Jiayi, CAO Hongbao, et al. Classification

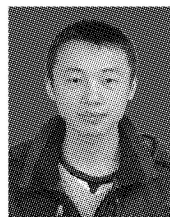
tion of motor imagery based on sample entropy[J]. Information and Control, 2008, 37(2): 191-196.

- [13] RICHMAN J S, MOORMAN J R. Physiological time-series analysis using approximate entropy and sample entropy [J]. American Journal of Physiology—Heart and Circulatory Physiology, 2000, 278(6): H2039-H2049.

作者简介:



张毅,男,1966年生,教授,博士生导师,博士,重庆市人工智能学会理事,重庆邮电大学智能系统及机器人研究所所长,英国Essex大学机器人研究中心访问学者.主要研究方向为机器人技术及应用、生物信号处理及应用、模式识别.主持完成省部级及其他科研项目10余项,获国家发明专利4项,发表学术论文60余篇,其中30余篇被SCI、EI、ISTP检索,出版专著1部、教材2部.



罗明伟,男,1987年生,硕士研究生.主要研究方向为脑电信号处理及应用,发表学术论文5篇.



罗元,女,1972年生,教授,博士.主要研究方向为数字信号处理及应用、机器视觉、数字图像处理及应用.主持或参与国家自然科学基金、科技部国际合作项目、教育部回国留学人员科研项目等10余项,获国家发明专利6项,发表学术论文50余篇.

2012年第2届计算机科学与网络技术国际会议

2nd International Conference on Computer Science and Network Technology(ICCSNT 2012)

2012第2届计算机科学与网络技术国际会议由东北师范大学主办,IEEE 哈尔滨分会以及辽宁师范大学、大连交通大学、哈尔滨师范大学、哈尔滨工业大学等学术团体协办,定于12月29-31日在长春召开.

第1届ICCSNT会议于2011年底在哈尔滨师范大学成功召开.会后,会议组织流程及论文集通过了IEEE的审核,所有录用的文章均已经进入IEEEExplore数据库且被EI核心检索.

本次会议旨在为学术、工业及商业界的专业人士在计算机科学、计算机应用、网络工程领域的最新进展的讨论提供一个广泛交流平台.

会议采用网上投稿方式,仅接收符合会议征稿范围的研究型全英文论文.

作者投稿的论文必须具有以下要素:充分的背景分析;明确新颖的理论基础;仿真或实验验证;论文需按照会议网站提供的模版进行排版;能够合理地使用公式、曲线图及表格直观清晰地展示研究成果,且结论论述充分;参考文献格式正确,数量不能少于10个;页数不超过5页,不少于3页半.

会议网站: <http://www.iccsnt.org/>.