

基于皮肤电信号与文本信息的双模态情感识别系统^①

张力行¹, 叶 宁^{1,2}, 黄海平^{1,2}, 王汝传^{1,2}

¹(南京邮电大学 计算机学院、软件学院、网络空间安全学院, 南京 210023)

²(南京邮电大学 江苏省无线传感网高技术重点实验室, 南京 210023)

通讯作者: 张力行, E-mail: 253053474@qq.com

摘 要: 人机交互离不开情感识别, 目前无论是单模态的情感识别还是多生理参数融合的情感识别都存在识别率低, 鲁棒性差的问题. 为了克服上述问题, 故提出一种基于两种不同类型信号的融合情感识别系统, 即生理参数皮肤电信号和文本信息融合的双模态情感识别系统. 首先通过采集与分析相应情感皮肤电信号特征参数和文本信息的情感关键词特征参数并对其进行优化, 分别设计人工神经网络算法和高斯混合模型算法作为单个模态的情感分类器, 最后利用改进的高斯混合模型对判决层进行加权融合. 实验结果表明, 该种融合系统比单模态和多生理参数融合的多模态情感识别精度都要高. 所以, 依据皮肤电信号和文本信息这两种不同类型的情感特征可以构建出识别率高, 鲁棒性好的情感识别系统.

关键词: 皮肤电信号; 文本识别; 特征提取; 情感识别; 加权融合

引用格式: 张力行, 叶宁, 黄海平, 王汝传. 基于皮肤电信号与文本信息的双模态情感识别系统. 计算机系统应用, 2018, 27(11): 103-108. <http://www.c-s-a.org.cn/1003-3254/6633.html>

Bimodal Emotion Recognition System Based on Skin Electrical Signals and Text Information

ZHANG Li-Xing¹, YE Ning^{1,2}, HUANG Hai-Ping^{1,2}, WANG Ru-Chuan^{1,2}

¹(School of Computer Science, Software and Cyberspace Security, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

²(Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Human-computer interaction is inseparable from emotion recognition. At present, there is a problem of low recognition rate and poor robustness both in single-modality emotion recognition and multi-physiological parameters fusion emotion recognition. Therefore, a fusion emotion recognition system based on two different types of signals is proposed, that is, a dual-modality emotion recognition system that integrates physiological parameters of skin electrical signals and text information. Firstly, by collecting and analyzing the characteristic parameters of the corresponding emotional skin electrical signals and the emotional keyword features of the textual information, the artificial neural network algorithm and the Gaussian mixture model algorithm are designed as a single mode emotion classifier, respectively. The Gaussian mixture model weights the decision layers. Experimental results show that this kind of fusion system has higher accuracy than multi-modality emotion recognition combined with single mode and multiple physiological parameters. Therefore, based on the two different types of emotional characteristics of skin electrical signals

① 基金项目: 国家自然科学基金 (61572260, 61373017, 61572261, 61672297); 江苏省“六大人才高峰”(2010DZXX026); 中国博士后科学基金 (2014M560440); 江苏省博士后科研基金 (1302055C); 江苏省重点项目研究与发展计划 (BE2015702, BE2017742); 江苏省优秀青年科学基金学者 (BK20160089)
Foundation item: National Natural Science Foundation of China (61572260, 61373017, 61572261, 61672297); “Six Talent Peaks” of Jiangsu Province (2010DZXX026); China Postdoctoral Science Foundation (2014M560440); Postdoctoral Research Foundation of Jiangsu Province (1302055C); Research and Development Plan of Key Projects of Jiangsu Province (BE2015702, BE2017742); Science Foundation for Outstanding Young Scholar of Jiangsu Province (BK20160089)

收稿时间: 2018-04-10; 修改时间: 2018-04-27; 采用时间: 2018-05-08; csa 在线出版时间: 2018-10-24

and text information, an emotion recognition system with high recognition rate and sound robustness can be constructed.

Key words: kin electrical; text recognition; signal feature extraction; emotion recognition; weighted fusion

情感计算不仅要让现在的计算机拥有智力能力,还要让它具备一定的情感交互能力,这样有利于人们与计算机有更好的交流.要实现情感计算首要问题是情感识别,情感识别直接影响着能否实现情感计算^[1].皮肤电信号作为生理参数只受自主神经系统和内分泌系统的分配,而且它相比其他生理参数更易于采集.但是它的缺陷也同样明显,如信号弱,抗干扰能力差等.文本信息作为人类最常用的交流手段之一,其本身蕴藏着丰富的信息资源,但是文本信息的情感识别也存在着很多缺陷,如主观性强等.所以这两种单模态的情感识别系统显然不能保证情感识别系统的准确性.但是将这两类不同类型信号的模态融合互补了单模态情感识别的缺陷,提高了分类器的识别准确度.它的优势在于,当一个通道的特征或识别过程受到影响或者缺失时,另一个通道能很好的进行补足,这样识别系统就会有更好的准确性和鲁棒性.

2004年, Kim等人将多生理参数进行融合,他们使用音频和视频作为实验素材采集被试的心电信号、脉搏信号、皮肤温度和皮肤电导信号,进行情感识别研究,识别率达到了78.4%^[2].2004年东京大学的Wang等^[3]使用生理信号的传感器和动画文本实现在线交流情感.2006年, Lee等人^[4]融合心电信号和皮肤电信号这两类生理参数识别人类情感,情感识别率达到80.2%.2017年陈鹏展等人^[5]通过融合语音信号与文本信息这两类信号也提高了识别率.

本文以皮肤电信号和文本信号这两类不同类型的信号为基础^[6],提出了它们各自的特征情感分析和融合算法^[7].并对各自的识别结果进行加权融合,构建双模态分类器.并对融合之后的识别率与单模态情况下的识别率和 Kim, Lee等人利用多生理参数融合识别情感进行了比较.实验结果表明不同类型的双模态情感识别系统具有更高的识别率.

1 特征提取

1.1 皮肤电信号的特征提取与数据处理

由于皮肤电信号比较微弱,易受到机器干扰,肌电干扰,电磁干扰等的影响,所以要对采集的皮肤电信号

去噪处理.本文采用小波变换进行去噪处理^[8].小波变化有其特有的优势,它在低频部分和高频部分对与频率分辨率和时间分辨率所展现出来的处理效果完全不同,前者具有较高的频率分辨率和较低的时间分辨率,而后者具有较高的时间分辨率和较低的频率分辨率.因此小波变换对于信号的自适应性非常强,尤其适用于生理信号^[9]的分析.信号 $f(t)$ 连续小波变换的定义如下:

$$WT_f(a, \tau) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(t) \psi^* \left(\frac{t-\tau}{a} \right) dt \quad (1)$$

其中, $WT_f(a, \tau)$ 为小波变换系数, a 为尺度因子, τ 为平移因子, $\psi(t)$ 为小波母函数, 它经过拉伸和移位之后得到小波基函数 $\psi_{a,\tau}(t)$.

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-\tau}{a} \right) \quad (2)$$

连续小波变换存在逆变换的条件是存在容许条件,容许条件公式:

$$C_\psi = \int_{\mathbb{R}} \frac{|\psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (3)$$

连续小波变换的逆变换公式为:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} WT_f(a, \tau) \frac{1}{\sqrt{a}} \psi \left(\frac{t-\tau}{a} \right) d\tau \quad (4)$$

但是在实际应用中,我们通常要将连续小波变换化为离散小波变换,主要是为了方便计算机进行分析、处理.通过对小波基函数 $\psi_{a,\tau}(t)$ 中的 a 进行幂级数离散化,对 τ 进行离散取值,即可得到:

$$\psi_{j,k}(t) = a_0^{-\frac{j}{2}} \psi(a_0^{-j} t - k\tau_0) \quad j = 0, 1, 2, \dots, k \in \mathbb{Z} \quad (5)$$

则离散小波变换为:

$$WT_f(j, k) = \int_{\mathbb{R}} f(t) a_0^{-\frac{j}{2}} \psi(a_0^{-j} t - k\tau_0) dt \quad (6)$$

对信号进行小波变换离散化实质上就是对信号分解,可以得到信号的低频分量 A 和信号的高频分量 D .其中 A 可以继续小波分解为低频信号和高频信号.本文采用了 `wdencomp` 函数实现信号去噪.去噪前后的信号如图1和图2所示.

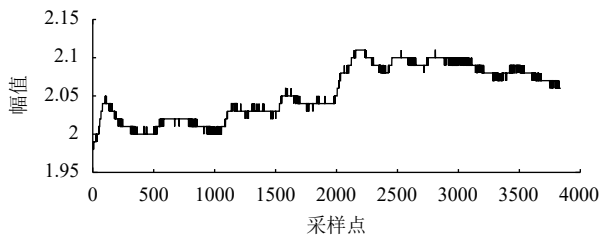


图1 原始皮肤电信号图

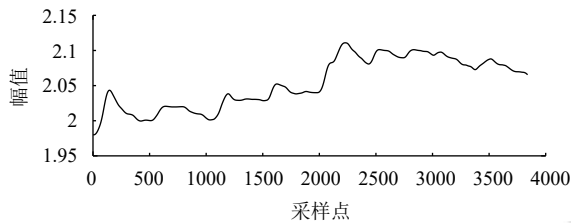


图2 去噪后皮肤电信号图

在对信号处理之后,删去无效数据,留下有效数据.参照德国 Augsburg 大学^[10]特征提取的方法,提取了信号的时域和频域中最能代表皮肤电信号变化的统计值作为情感识别研究的原始特征.在时域中,提取了皮肤电信号的:平均值,中位数,标准差,最大值,最小值,最小值比率,最大值比率,最大最小差值,一阶差分的均值,一阶差分的中位数,一阶差分的标准差,一阶差分的最大值,一阶差分的最小值,一阶差分的最小值比率,一阶差分的最大值比率,二阶差分的均值,二阶差分的中位数,二阶差分的标准差,二阶差分的最大值,二阶差分的最小值,二阶差分的最小值比率,二阶差分的最大值比率 22 个时域特征.信号的一阶差分检测局部极值点,二阶差分检测局部拐点.为了提取皮肤电信号的频域特征,先对皮肤电信号进行离散傅里叶变换,然后计算频率均值、中值、标准差、最大值、最小值、最大最小差值,得到 6 个频域特征.根据公式提取的 28 个统计特征的特征值的取值范围处在不同的数量级,为了方便后续处理,所以对他们进行归一化处理,使各特征值的取值范围在(0,1)之间.

在提取特征过程中涉及到的主要计算公式如下:

(1) 均值

$$u_x = \frac{1}{N} \sum_{i=1}^N X_n \quad (7)$$

(2) 标准差

$$\sigma_x = \left\{ \frac{1}{N-1} \sum_{n=1}^N (X_n - v_x)^2 \right\}^{\frac{1}{2}} \quad (8)$$

(3) 归一化

$$X'_i = \frac{X_i - X_{i\min}}{X_{i\max} - X_{i\min}} \quad (9)$$

(4) 一阶差分

$$1d_n = X_{n+1} - X_n \quad (10)$$

(5) 一阶差分的绝对值均值

$$\delta_x = \frac{1}{N-1} \sum_{n=1}^{N-1} (X_{n+1} - X_n) \quad (11)$$

(6) 二阶差分

$$2d_n = X_{n+1} - X_n \quad (12)$$

(7) 二阶差分绝对值

$$\gamma_x = \frac{1}{N-2} \sum_{n=1}^{N-2} (X_{n+2} - X_n) \quad (13)$$

(8) 最小值比率

$$\min Ratio = \frac{Min}{N} \quad (14)$$

(9) 最大值比率

$$\max Ratio = \frac{Max}{N} \quad (15)$$

在以上各式中 N 表示采集的信号总个数, X_n 表示第 n 次采集的信号, Min 表示采集信号的最小值, Max 表示采集信号的最大值.

1.2 文本信息的特征提取

文本信息的特征提取主要是对文本进行语法和语义的分析,通过对语句的拆分、去除冗余信息、去除停当词、分词、标注词性等提取出表达文本情感倾向性的情感词.采用大连理工大学信息检索研究室的情感词汇本体库^[11]来完成上述句子分词工作.即把一个完整的句子拆分成了若干个单独的情感词.采用布隆过滤器去除掉所有文本的停用词,以减少文本特征向量的维度和不必要的运算量.

特征提取采用信息增益 X ^[12],名词 Q 的 X 值定义为:

$$X(Q) = - \sum_{k=1}^n p(A_k) \log P(A_k) + P(Q) \sum_{i=1}^n p(A_k) \log P(A_k|Q) + P(\bar{Q}) \sum_{i=1}^n p(A_k|\bar{Q}) \log P(A_k|\bar{Q}) \quad (16)$$

式中, $p(A_k)$ 表示 A_k 类出现的概率, $P(Q)$, $P(\bar{Q})$ 表示名词 Q 出现在训练样本集和不出现样本训练集的概率. $P(A_k|Q)$, $P(A_k|\bar{Q})$ 表示在 $P(A_k|Q)$ 名词 Q 出现和不出现时出现 A_k 类的概率. X 值越高,对句子情感的判别帮助越大.

2 分类器模型创建

2.1 单通道皮肤电情感识别模型

单通道皮肤电情感识别模型创建思想是:对原始信号进行小波去噪处理之后留下有效数据进行特征选择,以最少的特征个数和最高的识别率来识别情感.再通过创建训练样本与测试样本,利用神经网络算法^[13]对情感进行分类,获得皮肤电情感识别结果.基于单通道的皮肤电信号模型分类器的识别框图如图3所示.

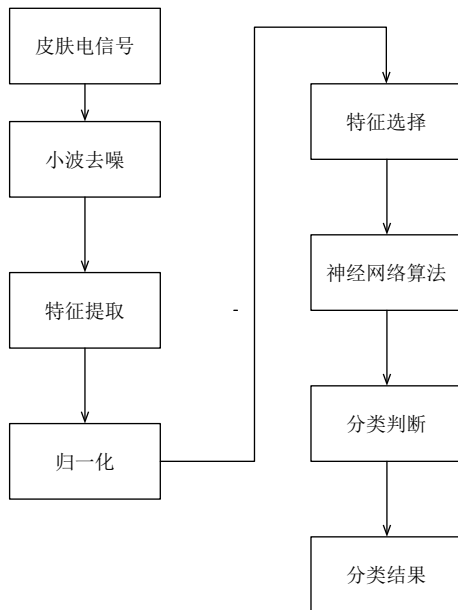


图3 皮肤电情感识别图

2.2 单通道文本情感识别模型

文本情感识别模型的主要思想是通过大连理工大学信息检索研究室提供的中文情感词库对句子中的情感关键词进行匹配判断.通过对文本内容进行预处理、特征提取和向量转化,然后通过高斯混合模型分类器进行情感状态的分类识别.基于单通道文本分类器识别框图如图4所示.

2.3 双模态融合识别模型

本文对两种单模态分类器分别采用了神经网络算法和高斯混合模型算法来进行4种情感识别.即高兴、悲伤、愤怒、平静.识别完成后再利用改进的高斯混合模型进行判决层融合^[14].高斯混合模型是M个成员密度的加权和,公式如下:

$$P(x_i|\lambda) = \sum_{i=1}^m a_i p_i(x_i; u_i, \Sigma_i) \quad (17)$$

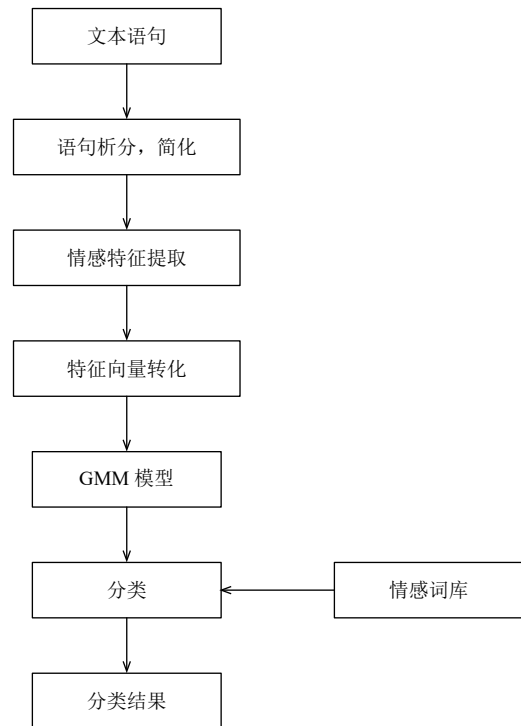


图4 文本情感识别图

式中, x_i 为第 i 个 D 维随机向量; $p_i(x_i; u_i, \Sigma_i)$ ($i = 1, 2, \dots, M$) 为成员密度, a_i 为第 i 个单高斯分布的权值, 且 $\sum_{i=1}^m a_i = 1$; 每个成员密度均为单高斯分布函数, 它是 D 维变量的关于均值矢量为 u_i , 协方差矩阵为 Σ_i 的函数, 即:

$$p_i(x_i; u_i, \Sigma_i) = \frac{1}{[2\pi^{\frac{D}{2}} |\Sigma_i|^{\frac{1}{2}}]} \exp \left\{ -\frac{1}{2(x_i - u_i)^T \Sigma_i^{-1} (x_i - u_i)} \right\} \quad (18)$$

完整的高斯混合分布密度公式如下:

$$\lambda_i = \{a_i, u_i, \Sigma_i\} \quad i = 1, \dots, m \quad (19)$$

其中, GMM 模型的参数估计采用 EM 算法迭代完成, 使得 GMM 表示的样本分布概率最佳. 由于单通道在工作时都会存在一定干扰, 当文本信息分类器受到噪声干扰, 那么他的性能就会下降; 皮肤电信号受到仪器采集的影响, 如抖动, 碰撞或者仪器本身受到的基线漂移干扰, 此时性能都会有所下降. 此时就要考虑在融合前各个子分类器的置信度, 所以本文采取自适应加权融合算法让两个通道信息进更新和融合. 各分类器加权系数^[15]根据其当前样本可靠性进行动态调整, 置信度高的分类器所占权重更高, 算法以自适应的方式

找到每个分类器的最优加权因子, 利用得到的加权因子实现双模态数据融合, 获得最终的分类结果. 两个子分类器给出的情感类的 GMM 似然度记为 $P(z|\lambda_k)$, 这里我们假设待测样本为 z , k 代表四种情感类别, 取值 1-4. 不同情感类别得出的最大似然度直接决该分类器的置信区间. 子分类器的融合权值公式如下:

$$I_n = \sum_{1 \leq i \leq j \leq 4} \frac{|\ln(P(y|\lambda_j)) \ln(P(y|\lambda_j))|}{\left| \sum_{k=1}^4 \ln(P_n(y|\lambda_k)) \right|} \quad (20)$$

子分类器的分类结果公式如下:

$$A_n = \max\{P_n(y|\lambda_k)\} \quad (21)$$

子分类器的判决置信度的高低与样本所处概率分布模型的非重叠区域有关, 似然值越分散, 判决置信度越高, 性能越好. 最后通过对计算出来的两个子分类器的判决进行加权融合, 加权融合公式如下:

$$Y = \sum_{n=1}^2 I_n A_n \quad (22)$$

式中, Y 为融合之后的最终结果, A_n 为子分类器结果.

双模态混合模型识别框图如图 5 所示.

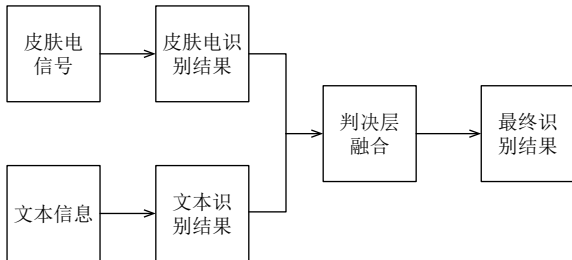


图 5 双模态情感识别图

3 实验结果分析

实验中采用高斯混合模型作为融合算法, 通过 4 种类别的训练样本集和测试样本集完成对两个单模态和一个多模态分类器实验, 4 种类别包括高兴、悲伤、愤怒、平静. 每个训练样本集包括每种情感的 150 条皮肤电样本与 150 条文本信息样本, 每个测试样本各含 50 条样本.

通过对三种模态, 即两个单模态和一个融合模态的实验之后, 图 6 显示了三个不同模型产生的实验结果. 实验表明, 融合之后的情感识别模型对每类情感识别精度明显提高. 平均识别率也由单模态的 70%, 80% 提高到双模态融合后的 90%.

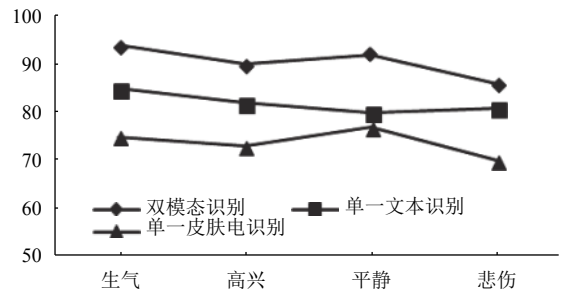


图 6 3 种方法识别率对比图

而对比于多生理参数融合, 2004 年, Kim 等人对多生理参数进行融合, 他们使用音频和视频作为实验素材采集被试的心电信号、脉搏信号、皮肤温度和皮肤电导信号, 进行情感识别研究, 识别率达到了 78.4%. 2006 年, Lee 等人融合心电信号和皮肤电信号这两类生理参数识别人类情感, 识别率达到了 80.2%.

实验结果表明, 两种不同类型信号模态融合的情感识别, 不仅可以提高各自单模态的情感识别率, 而且较之于多模态生理参数的融合情感识别率也大幅度提高.

表 1 显示采取双模态情感识别系统之后对四类情感的误判率只有 6%, 10%, 8%, 14%. 比较于单模态的情感识别率有了大大的提高.

分类	生气	高兴	平静	悲伤
生气	94	6	0	0
高兴	2	90	0	8
平静	0	2	92	6
悲伤	4	2	8	86

4 结论与展望

本文采用加权融合的方式将皮肤电信号识别情感与文本信息识别情感的结果进行融合, 与仅利用皮肤电信号识别情感和仅利用文本信息识别情感进行了比较, 结果表明无论是识别率还是识别精度和鲁棒性都得到了大幅提高. 本文为当前单纯的利用一种方式识别情感提供了一种新的思路, 实验也证明是确实可行的.

目前的情感识别研究要么采取的是单通道情感识别, 要么采取的是多生理信号的融合或同一类型不同特征的融合, 而本文将生理信号与文本信息两种不同类型的信号进行加权融合, 通过对这两种不同来源数据的分类结果再次融合, 实现了不同类型信号情感的双模态识别系统的研究. 通过对单模态皮肤电信号、文本信息的分类实验以及双模态融合后的情感识别实

验,实验结果表明,相比于一般的单模态和多生理信号融合的情感识别,该种模式系统识别率更高,鲁棒性更强。

参考文献

- 1 Vinciarelli A, Pantic M, Bourlard H. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 2009, 27(12): 1743–1759. [doi: [10.1016/j.imavis.2008.11.007](https://doi.org/10.1016/j.imavis.2008.11.007)]
- 2 Kim KH, Bang SW, Kim SR. Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biological Engineering and Computing*, 2004, 42(3): 419–427. [doi: [10.1007/BF02344719](https://doi.org/10.1007/BF02344719)]
- 3 Wang H, Prendinger H, Igarashi T. Communicating emotions in online chat using physiological sensors and animated text. *Conference on Human Factors in Computing Systems*. Vienna, Austria. 2004
- 4 Lee CK, Yoo SK, Park YJ, *et al.* Using neural network to recognize human emotions from heart rate variability and skin resistance. *Proceedings of 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. Shanghai, China. 2006. 5523–5525.
- 5 陈鹏展,张欣,徐芳萍.基于语音信号与文本信息的双模态情感识别. *华东交通大学学报*, 2017, 34(2): 100–104.
- 6 曹梦思.基于脑电信号的中文情感词的情感识别[硕士学位论文].北京:北京邮电大学,2012.
- 7 付丽琴,毛峡,陈立江.基于改进的排序式选举算法的语音情感融合识别. *计算机应用*, 2009, 29(2): 381–385.
- 8 Li CW, Zheng CX, Tai CF. Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on Biomedical Engineering*, 1995, 42(1): 21–28.
- 9 Alfaouri M, Daqrouq K. ECG signal denoising by wavelet transform thresholding. *American Journal of Applied Sciences*, 2008, 5(3): 276–281. [doi: [10.3844/ajassp.2008.276.281](https://doi.org/10.3844/ajassp.2008.276.281)]
- 10 Kim J, André E. Emotion recognition using physiological and speech signal in short-term observation. In: André E, Dybkjær L, Minker W, *et al.*, eds. *Perception and Interactive Technologies*. Berlin, Heidelberg. Springer. 2006. 53–64.
- 11 徐琳宏,林鸿飞,潘宇,等.情感词汇本体的构造. *情报学报*, 2008, 27(2): 180–185. [doi: [10.3969/j.issn.1000-0135.2008.02.004](https://doi.org/10.3969/j.issn.1000-0135.2008.02.004)]
- 12 申红,吕宝粮,内山将夫,等.文本分类的特征提取方法比较与改进. *计算机仿真*, 2006, 23(3): 222–224. [doi: [10.3969/j.issn.1006-9348.2006.03.061](https://doi.org/10.3969/j.issn.1006-9348.2006.03.061)]
- 13 Wang GJ, Wang ZL, Li LX. An affective model of artificial psychology [computational affection model]. *Proceedings of 2005 International Conference on Communications, Circuits and Systems*. Hong Kong, China. 2005. 1061–1064.
- 14 黄程韦,金赟,王青云,等.基于语音信号与心电信号的多模态情感识别. *东南大学学报(自然科学版)*, 2010, 40(5): 895–900. [doi: [10.3969/j.issn.1001-0505.2010.05.003](https://doi.org/10.3969/j.issn.1001-0505.2010.05.003)]
- 15 蔡莉莉.基于数据融合的语音情感分析与识别[硕士学位论文].南京:东南大学,2005.