

基于 Fisher 准则与 SVM 的分层语音情感识别^{*}

陈立江¹ 毛 峡¹ Mitsuru ISHIZUKA²

¹(北京航空航天大学 电子信息工程学院 北京 100191)

²(Department of Information and Communication Engineering, University of Tokyo, Japan)

摘 要 针对说话人无关的语音情感识别,提出一个分层语音情感识别模型,由粗到细识别悲伤、愤怒、惊奇、恐惧、喜悦和厌恶 6 种情感。每层采用 Fisher 比率从 288 个备选特征中选择适合该层分类的特征,同时将 Fisher 比率作为输入参数训练该层的支持向量机分类器。基于北京航空航天大学情感语音数据库和德国柏林情感语音数据库,设计 4 组对比实验,实验结果表明, Fisher 准则在两两分类特征选择上优于 PCA, SVM 在说话人无关的语音情感识别推广方面优于人工神经网络(ANN)。在两个数据库的基础上得到类似结果,说明文中分类模型具有一定的跨文化适应性。

关键词 语音情感识别,说话人无关, Fisher 准则,支持向量机
中图分类号 TP 391.42

Multi-Level Speech Emotion Recognition Based on Fisher Criterion and SVM

CHEN Li-Jiang¹, MAO Xia¹, Mitsuru ISHIZUKA²

¹(School of Electronic and Information Engineering, Beihang University, Beijing 100191)

²(Department of Information and Communication Engineering, University of Tokyo, Japan)

ABSTRACT

To solve the speaker independent emotion recognition problem, a multi-level speech emotion recognition system is proposed to classify 6 speech emotions, including sadness, anger, surprise, fear, happiness and disgust from coarse to fine. The key is that the emotions divided by each layer are closely related to the emotional features of speech. For each level, appropriate features are selected from 288 candidate features by Fisher ratio which is also regarded as input parameter for the training of support vector machine (SVM). Based on Beihang emotional speech database and Berlin emotional speech database, principal component analysis (PCA) for dimension reduction and Artificial Neural Network (ANN) for classification are adopted to design 4 comparative experiments, including Fisher+SVM, PCA+SVM, Fisher+ANN, PCA+ANN. The experimental results prove that Fisher rule is better than PCA for dimension reduction, and SVM is more expansible than ANN for speaker independent speech emotion recognition. Good cross-cultural adaptation can be inferred from the similar results of experiments based on two different databases.

^{*} 国家自然科学基金项目(No. 61103097 60873269)、中日国际科技合作项目(No. 2010DFA11990)资助

收稿日期: 2011-03-21; 修回日期: 2011-09-08

作者简介 陈立江,男,1984 年生,博士研究生,主要研究方向为语音信号处理、情感识别等。E-mail: clj@ee.buaa.edu.cn。毛峡,女,1952 年生,教授,博士生导师,主要研究方向为情感计算、模式识别、人机交互等。Mitsuru ISHIZUKA,男,1949 年生,教授,博士生导师,主要研究方向为人机交互、情感计算、仿生代理等。

Key Words Speech Emotion Recognition, Speaker Independent, Fisher Criterion, Support Vector Machine

1 引言

通过语音传递信息是人类最重要的能力之一, 语音中的情感信息可影响人们的交流状态. 如, 说话人用不同情感表达某语句时, 听者可能会有不同的反应, 所谓“听话听音”就是这个道理. 因此研究如何利用计算机识别语音中的情感有助于实现和谐的人机情感交互. Johnson 等^[1]的实验表明, 人耳能容易地识别高兴、悲伤、愤怒等情绪, 即使是说话人无关并且经过各种干扰的语音信号, 识别率也能超过 70%. 随着模式识别和情感计算理论不断发展, 利用计算机识别语音情感的研究受到众多学者的重视. 然而, 计算机的表现仍远逊于人类, 语音情感识别还未能走入人们的日常生活. 语音情感识别主要存在两大难点: 如何寻找有效的语音情感特征; 如何建立适当的情感识别模型.

语音信号是一维信息, 语音情感特征隐藏于时域和频域之中. 语音情感识别的研究表明基频、语调、重音、语速、停顿等低频韵律信息与情感唤醒度具有较强的相关性^[2-3], 同时高频信息能有效区分某些特定情感, 例如愤怒时急促的呼吸声、恐惧时的哆嗦声等. Mao 等^[4]和 Schuller 等^[5]结合不同时间尺度特征识别语音情感, 得到较好的识别结果. Mao 等^[6]采用参数滤波与分形维数提取语音情感特征并通过实验验证新特征与情感具有较好相关性. 刘佳等^[7]利用协方差描述子和黎曼流形进行噪声环境下的语音情感识别取得较好效果. 上述研究均采用固定特征, 但实际上某些特征可能只适用于某些情感的识别.

离散情感模型和维度情感模型是最具有影响力的两种情感模型. 离散情感模型最早由 Darwin 提出, 并且被 Tomkins^[8], Ekman^[9]和 Izard^[10]继承和发扬. 这些情感模型中归纳 2 种到 14 种基本情感或主要情感, 更加精细的情感可由基本情感混合而成. 最著名的离散情感模型是 Ekman 的 6 种基本情感(悲伤、愤怒、惊奇、恐惧、喜悦、厌恶)模型. 维度情感模型最早由 Wundt^[11]提出, 离散的情感类别被放置在两维或三维的情感空间中. 其中愉悦度-激活度-优势度(Pleasure Activation Dominance, PAD)情感空间得到多数研究者的认可. 然而, 并没有专门的研究指明各基本情感或情感维度与语音信号之间的联系.

本文借助 Fisher 准则探索不同情感特征与情感的语音表现之间的内在联系, 达到特征降维的目的, 同时与主成分分析(PCA)进行对比, 验证 Fisher 准则的有效性. 同时, 本文将离散情感模型和维度情感模型与情感语音特征紧密结合, 建立一个分层语音情感识别模型. 在该模型中每层采用一到两个分类器进行两两分类, 层数越低, 分类越粗糙; 层数越高, 分类越精细, 最终实现 6 种基本情感分类.

2 基于 Fisher 准则的特征选择

高维特征虽然包含丰富的信息, 但是往往存在大量的冗余, 影响识别效果, 同时高维数据的巨大计算量也影响到实际应用, 因此必须对众多备选特征进行降维处理.

Fisher 准则是一种传统的线性判别方法, 在模式识别领域得到广泛应用^[12-13], 其基本原理是寻找特征空间的某个投影子空间, 使得所有特征点在该子空间得到最好的分类.

设样本集 $X = \{x_1, x_2, \dots, x_n\}$ 中共有 n 个样本, 分 C 类, 每个样本 x_i 均为 L 维列向量, 定义类间离散度 S_b 和类内离散度 S_w :

$$S_b = \frac{1}{n} \sum_{j=1}^C n_j (m_j - m)(m_j - m)^T,$$

$$S_w = \frac{1}{n} \sum_{j=1}^C \sum_{x_i \in X_j} (x_i - m_j)(x_i - m_j)^T,$$

其中 x_j 、 n_j 和 m_j 分别为第 j 类样本的子集、总个数和均值点, m 为所有样本的均值点.

Fisher 比率:

$$F_r = \text{diag}(S_b / S_w),$$

其中, 符号“/”表示矩阵的对应元素分别相除, 函数 $\text{diag}()$ 取矩阵对角线上的元素, F_r 为 L 维列向量, 其中每个元素反映对应特征的二分类能力.

首先将语音信号进行分帧处理, 帧长为 256 个采样点, 帧移为 128 个采样点, 提取每帧信号的能量、过零率、能零积、基频、一至三阶共振峰、频谱质心、频谱截止频率、自相关密度、计盒维数及 5 个 Mel 子带能量, 再计算各特征对时间的一次导数和二次导数, 得到 48 维特征序列. 然后分别计算各特征序列的最大值、最小值、平均值、标准差、偏度和峰度, 最终得到 288 维备选特征.

本文所提取的特征之中, 能量、过零率、能零积、

基频、共振峰都属于传统的语音特征,其中能量和基频反应语音的低频韵律特性,过零率和共振峰反应高频信息,能零积则综合能量和过零率的信息.频谱质心与频谱截止频率反映语音信号的频率分布特性,均属于音质特征.首先对每帧数据进行快速 Fourier 变换:

$$F(k) = \left\| \sum_{n=1}^N x(n) \exp(-i2\pi k \frac{n}{N}) \right\|, \\ k = 1, 2, \dots, N,$$

其中 $F(k)$ 表示第 k 个频率点的幅值, $x(n)$ 表示输入语音信号, N 为语音信号帧长(256).

频谱质心 F_{cent} :

$$F_{\text{cent}} = \frac{\sum_{k=1}^{\frac{N}{2}} k \cdot F(k)}{\sum_{j=1}^{\frac{N}{2}} F(j)},$$

频谱截止频率 F_{cut} 是指某一频率点,满足

$$\frac{\sum_{k=1}^{F_{\text{cut}}} F(k)}{\sum_{j=1}^{\frac{N}{2}} F(j)} = 0.85.$$

自相关密度反映语音信号的短时频谱分布随时间的变化,计盒维数反映语音信号的非线性特性.文献[6]中讨论该两种特征与情感之间的相关性.首先通过五组全极点滤波器对语音信号进行不同程度的低频提升,然后计算各滤波结果的一阶自相关系数:

$$\rho_k = \sum_{n=1}^{N-1} Y_k(n) Y_k(n+1),$$

其中 $Y_k(n)$ 表示第 k 个全极点滤波器在第 n 点的输出结果.自相关密度:

$$C_d = \lg \left[\sum_{k=1}^4 (\rho_{k+1} - \rho_k)^2 \right].$$

用一个正方形的网状栅格将语音信号波形覆盖,网格的边长为 A , $N(A)$ 为包含语音信号波形的网格数,计盒维数:

$$D_b = \lim_{A \rightarrow 0} \left(\frac{\lg N(A)}{\lg \left(\frac{1}{A} \right)} \right).$$

Mel 频率符合人耳对声音频率的感知特性, Mel 子带能量:

$$E_{mk} = \sum_{n=1}^N [Y_{mk}^2(n)],$$

其中 $Y_{mk}(n)$ 表示语音信号通过第 k 个 Mel 子带滤波器在第 n 点的输出结果.

3 分层语音情感识别模型

无论是离散情感模型还是维度情感模型均存在局限性.在离散情感模型中,情感类别与语音情感特征不存在一一对应的关系,例如喜悦、愤怒和惊奇均会导致语音信号的基频和能量升高.在维度情感模型中,语音情感特征向量并非均匀分布在维度情感空间中,当情感激活度较低时,人们甚至选择沉默,自然无语音情感特征可言.根据以上分析,本研究拟对现有的情感模型做如下改进:新模型借鉴离散情感模型由基本情感到精细情感的逐步分类思想,并保留连续情感模型空间维度的概念,其特点是,划分情感空间的维度与语音情感特征紧密相关.图1是各离散情感类别在新的情感空间中的分布示意图.

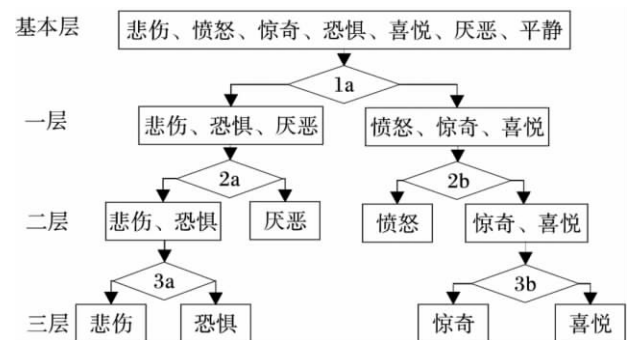


图1 分层语音情感空间模型

Fig. 1 Multi-level speech emotion recognition system

图1的情感模型有两个基本问题需要解决: 1) 分层空间的结构,即分为几层; 2) 采用哪些语音特征进行情感空间的划分.平静语音在分层情感空间中仅处在顶层基本空间,其余各层空间结构划分过程中不予考虑.首先对6种基本情感采用穷举法,列出所有可能的第一层空间划分情况,并针对每种情况计算所有备选特征的 Fisher 比率,根据排名前5个 Fisher 比率的平均值选择最优的第1层空间划分,同时将该5个特征用于第1层语音情感识别.以此类推,得到更高层的最优空间划分情况及用于每层语音情感识别的语音情感特征.经过筛选得到最优的分层语音情感空间模型具有4层结构:基本层包含所有离散情感类别;第1层包含1个分类器,将情感分为趋于远离(悲伤、恐惧、厌恶)和趋于靠近(愤怒、惊奇、喜悦)两大类;第2层和第3层各包含两个分类器将情感逐步细分.

由于 SVM 具有较好的推广性能,因而在说话人无关的语音情感识别研究中得到广泛应用^[14-15]。由于本文致力于说话人无关的分层语音情感识别研究,因此采用鉴别能力和推广特性较好的 SVM 作为分类器。

支持向量机的本质是采用核函数将特征非线性地映射到高维空间,并在高维空间中寻找最大间隔分类面。常见的核函数有多项式核函数、RBF 核函数和 Sigmoid 核函数。本文采用 Sigmoid 核函数:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{1 + \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|)}.$$

在线性不可分问题中,文献[16]引入松弛变量 ξ_i 和惩罚系数 C ,求解软边界二次规划问题。 ξ_i 可作为训练样本关于广义分离超平面的偏差。 $C > 0$ 是一个自定义的惩罚因子,它决定对错分样本的惩罚程度,用来控制样本偏差与机器泛化能力之间的平衡。 C 越大,惩罚就越大,对错分样本的约束程度也就越大。惩罚因子 C 的取值与具体应用有关:当情感特征的区分性能较好时,可以适当提高惩罚因子,以提高 SVM 的分类能力。当情感特征的区分性能较差时,必须适当减小惩罚因子,以保证分类器有较好的泛化性能。本文的分层语音情感识别模型中各分类器采用的惩罚因子取决于该层语音情感识别所采用的语音情感特征的 Fisher 比率:

$$C_F = \exp\left(K_F \frac{1}{N_f} \sum_{i=1}^{N_f} F_r(i)\right),$$

其中 F_r 表示各输入特征的 Fisher 比率。 K_F 由实验设定为 15。

人工神经网络(Artificial Neural Networks, ANN)可实现复杂的映射函数。Mao 等采用 HMM 与 ANN 相结合的方式提高语音情感的识别率^[17],并与 ANN 进行对比实验。本文使用的 ANN 采用多层感知器拓扑结构,节点输出函数采用 Sigmoid 函数:

$$S_{ig}(x) = \frac{2}{1 + e^{-2x}} - 1.$$

为了与 SVM 进行对比,各 ANN 分类器只进行两两分类,因此输出节点数为 1。输入节点数与输入特征维数相同,且输入特征与对应 SVM 的输入特征完全相同。

4 语音库及实验结果

4.1 数据库

优质的情感语料是语音情感研究的基础。由于汉语是音调语言,其发音方式与西方语言存在差别,

因此国际现有的一些西方语言情感语料库仅适合作为参考。本研究基于北京航空航天大学情感语音数据库和柏林情感语音数据库。北京航空航天大学情感语音数据库包含 15 名发音人员(7 男 8 女) 20 句适合表达各种情感点录音脚本,7 种情感(包括平静),每种情感重复 3 次,共 6 300 句情感语料。为保证数据库中情感语料的有效性,邀请不少于 20 名评测者对每句情感语料进行主观评测,根据评测结果为每句情感语料的可信度属性赋值。根据情感语料的可信度值从全部语料集中选择 2 套语料子集,每套子集中的情感语料至少被 $p\%$ 评测者正确识别($p = 50, 70$)。情感语料全集用 S 表示,则各情感语料集合关系为 $S \supseteq S_{50} \supseteq S_{70}$ 。柏林情感语音数据库包含 10 名发音人员(5 男 5 女) 10 句录音脚本,7 种情感,其中的情感语料在主观评测中被超过 60% 的评测者正确识别^[18]。

4.2 特征选择实验结果

本文分层语音情感空间模型的确立过程也是各层语音情感识别的特征选择过程。图 1 中各分类器均进行二分类。在北京航空航天大学情感语料库全集 S 和柏林情感语料库二者的并集上计算各备选特征的 Fisher 比率,表 1 为各分类器 Fisher 比率排名靠前的特征。

由表 1 数据可知,随着分类层数的增加, Fisher 比率逐渐减小。如 1a 分类器 Fisher 比率最大值达到 1.07,而 3b 分类器 Fisher 比率最大值不到 0.1,说明越精细的分类越困难。其中惊奇与喜悦(3b 分类器)最难区分,其次是悲伤与恐惧(3a 分类器)。另外,随着分类层数的增加,所需特征统计量的阶数呈现递增现象,如 1a、2a 和 3a 分类器所需统计量多为一阶和二阶,如最大值、最小值、标准差等,3a 和 3b 分类器出现高阶统计量如峰度,说明情感划分越精细,越需要特征分布的精细结构来进行情感识别。

4.3 分层语音情感识别实验结果

本文基于北京航空航天大学情感语音数据库的 S50、S70 数据集以及柏林情感语音数据库分别设计 4 组对比实验,包括 Fisher+SVM, PCA+SVM, Fisher+ANN, PCA+ANN,分别验证 Fisher 准则的特征选择效果和 SVM 对说话人无关语音情感识别的可推广性。实验中随机选择数据库中的 3 名女说话人和 3 名男说话人的语料作为训练数据,其余说话人的语料作为测试数据。对于混合情感分类器的训练采用数据库中与混合情感类别相对应的情感语音数据进行训练,例如 2a 分类器,采用数据库中的悲伤和恐惧语音作为 A 类,厌恶语音作为 B 类,训练 SVM

和 ANN 分类器. 训练每层分类器采用的特征种类和数量由实验确定. 对于按照 Fisher 比率排列的特征, 由序列前向搜索算法确定参加训练 SVM 或 ANN 的特征个数. 对于按照 PCA 由大到小排列的奇异值, 由序列前向搜索算法确定参加训练 SVM 或 ANN 的奇异值(主成份)个数. 训练各层分类器的特征数如表 2 所示.

表 1 各分类器特征 Fisher 比率
Table 1 Fisher rates of features for each classifier

分类器	特征序号	Fisher比率	统计量	导数	基本特征
1a	52	1.072420	最小值	零次导数	自相关密度
	1	0.961068	最大值	零次导数	能量
	161	0.909948	标准差	一次导数	能量
	177	0.854875	标准差	二次导数	能量
	145	0.833322	标准差	零次导数	能量
2a	100	0.472027	平均值	零次导数	自相关密度
	52	0.275581	最小值	零次导数	自相关密度
	4	0.228755	最大值	零次导数	自相关密度
	97	0.226631	平均值	零次导数	能量
	1	0.178080	最大值	零次导数	能量
2b	163	0.312398	标准差	一次导数	能零积
	179	0.292202	标准差	二次导数	能零积
	161	0.278100	标准差	一次导数	能量
	191	0.260381	标准差	二次导数	第四 Mel 子带能量
	67	0.258701	最小值	一次导数	能零积
3a	81	0.142356	最小值	二次导数	能量
	177	0.139190	标准差	二次导数	能量
	65	0.137351	最小值	一次导数	能量
	161	0.136651	标准差	一次导数	能量
	163	0.103730	标准差	一次导数	能零积
3b	168	0.089046	标准差	一次导数	频谱截止频率
	184	0.043731	标准差	二次导数	频谱截止频率
	54	0.041251	最小值	零次导数	基频
	251	0.040918	峰度	零次导数	第三共振峰
	280	0.040316	峰度	二次导数	频谱截止频率

表 2 训练各层分类器的特征数
Table 2 Number of features for each classifier

降维方法	1a	2a	2b	3a	3b
Fisher	4	5	3	3	4
PCA	5	6	4	5	3

识别结果如图 2 所示. 图中 3 种不同深度的柱形分别代表分层语音情感识别中第 1 层到第 3 层的平均识别率.

对比图 2 中的 4 组实验结果可知. 在两两分类多层语音情感识别中, 采用 Fisher 准则进行特征选

择优于采用 PCA 进行特征降维, 各层识别率均有提高. 其原因为该模型符合 Fisher 准则使得所有特征点在子空间得到最好分类的特点, 而 PCA 则在子空间对原始数据进行最小二乘意义下的最佳描述. 由于 ANN 的推广性能不如 SVM, 因此在说话人无关语音情感识别中, ANN 的识别结果明显低于 SVM 的识别结果. 最优组合为 Fisher+SVM.

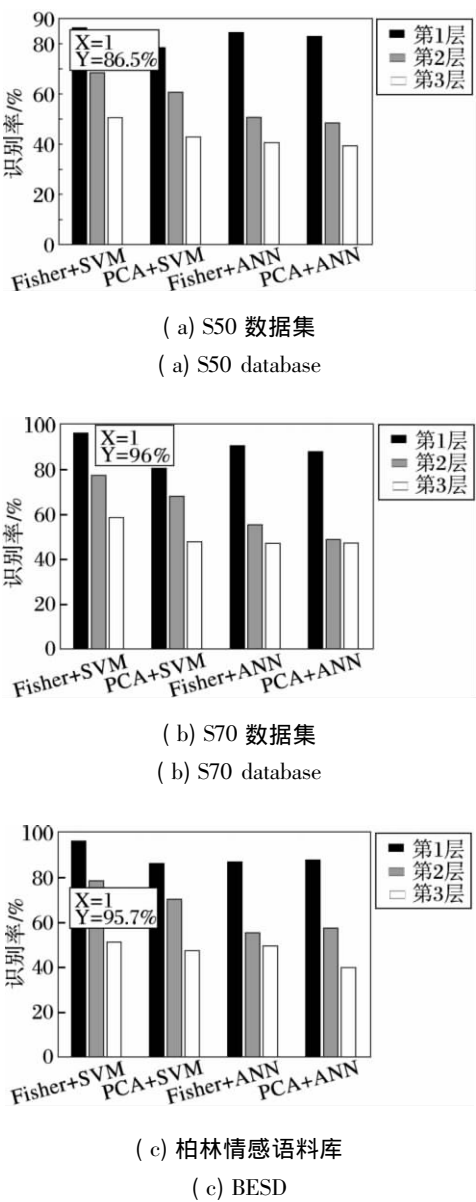


图 2 基于不同数据集语音情感识别结果

Fig. 2 Results of emotion recognition based on different databases

图 2(a) 与图 2(b) 的结果说明语料库的质量对识别结果有较大影响, 基于 S70 数据集的识别率比基于 S50 数据集的识别率提高约 10%. 图 2(c) 的结果说明, 本文提出的 Fisher+SVM 分层语音情感识别

模型同样适用于德语情感语音识别, 基于该数据库的第一、第二级识别结果与 S70 数据集相当, 而第三级识别结果与 S50 数据集相当. 分析其原因是由于该数据库情感语料的数量较少, 第三级精细情感识别模型训练不够充分, 尤其是厌恶、悲伤等情感识别率较低, 导致平均识别率不高.

5 结 束 语

为提高说话人无关的语音情感识别性能, 本文从如何寻找有效的语音情感特征以及如何建立适当的语音情感识别模型出发, 通过 Fisher 准则探索语音特征与语音情感表现的内在联系, 并通过建立基于 SVM 的多层语音情感识别模型提高语音情感识别性能. 基于汉语和德语两个数据库的实验得到类似结果, 说明本文的分类模型具有一定的跨文化适应性.

实验结果表明恐惧、厌恶等个别情感的识别率仍然有待提高, 后续研究可从以下两方面进行: 1) 结合特殊情感信息进行情感识别, 如惊奇时句尾的基频上扬, 恐惧时的哆嗦声等; 2) 采用情感维度空间模型结合模糊理论识别某种情感的存在概率.

参 考 文 献

- [1] Johnson W F, Emde R N, Scherer K R, *et al.* Recognition of Emotion from Vocal Cues. *Arch Gen Psychiatry*, 1986, 43(3): 280-283
- [2] Park C H, Sim K B. Emotion Recognition and Acoustic Analysis from Speech Signal // *Proc of the International Joint Conference on Neural Networks*. Portland, USA, 2003, IV: 2594-2598
- [3] Banziger T, Scherer K R. The Role of Intonation in Emotional Expressions. *Speech Communication*, 2005, 46(3/4): 252-267
- [4] Mao Xia, Chen Lijiang, Fu Liqin. Multi-Level Speech Emotion Recognition Based on HMM and ANN // *Proc of the WRI World Congress on Computer Science and Information Engineering*. Los Angeles, USA, 2009, VII: 225-229
- [5] Schuller B, Rigoll G, Lang M. Speech Emotion Recognition Combining Acoustic Features and Linguistic Information in a Hybrid Support Vector Machine-Belief Network Architecture // *Proc of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Montreal, Canada, 2004, I: 577-580
- [6] Mao Xia, Chen Lijiang. Speech Emotion Recognition Based on Parametric Filter and Fractal Dimension. *IEICE Trans on Information and Systems*, 2010, 93(8): 2324-2326
- [7] Liu Jia, Chen Chun, Ye Chengxi, *et al.* Speech Emotion Recognition Based on Covariance Descriptor and Riemannian Manifold. *Pattern Recognition and Artificial Intelligence*, 2009, 22(5): 673-677 (in Chinese)
(刘佳, 陈纯, 叶承羲, 等. 基于协方差描述子和黎曼流形的语音情感识别. *模式识别与人工智能*, 2009, 22(5): 673-677)
- [8] Tomkins S S. *Affect, Imagery, Consciousness*. New York, USA: Springer, 1962
- [9] Ekman P. An Argument for Basic Emotions. *Cognition and Emotion*, 1992, 6(3/4): 169-200
- [10] Izard C E. *Human Emotions*. New York, USA: Plenum Press, 1977
- [11] Wundt W M. *Grundzuge der Physiologischen Psychologie*. 5th Edition. Leipzig, Germany: Engelmann, 1874
- [12] Guo Yuefei, Shu Tingting, Yang Jingyu, *et al.* Feature Extraction Method Based on the Generalised Fisher Discriminant Criterion and Facial Recognition. *Pattern Analysis & Applications*, 2001, 4(1): 61-66
- [13] Wang S, Li Deyu, Wei Yingjie, *et al.* A Feature Selection Method Based on Fisher's Discriminant Ratio for Text Sentiment Classification // *Proc of the International Conference on Web Information Systems and Mining*. Shanghai, China, 2009: 88-97
- [14] Lin Yilin, Wei Gang. Speech Emotion Recognition Based on HMM and SVM // *Proc of the 4th International Conference on Machine Learning and Cybernetics*. Guangzhou, China, 2005, VIII: 4898-4901
- [15] Lu Tao. Speech Emotion Recognition Based on SVM. Master Dissertation. Yanshan, China: Yanshan University, 2007 (in Chinese)
(芦涛. 基于 SVM 的汉语语音情感识别研究. 硕士学位论文. 燕山: 燕山大学, 2007)
- [16] Cortes C, Vapnik V. Support-Vector Networks. *Machine Learning*, 1995, 20(3): 273-297
- [17] Mao Xia, Chen L J, Zhang Bing. Mandarin Speech Emotion Recognition Based on a Hybrid of HMM/ANN. *International Journal of Computers*, 2007, 4(1): 321-324
- [18] Burkhardt F, Paeschke A, Rolfes M, *et al.* A Database of German Emotional Speech // *Proc of Interspeech*. Lissabon, Portugal, 2005: 1517-1520