

# 基于组合核函数SVM的说话人识别方法

樊持杰<sup>1</sup>, 司巧梅<sup>1</sup>, 徐岩<sup>2</sup>, 张丹<sup>1</sup>, 蔡春华<sup>1</sup>, 于旭<sup>3</sup>

1. 牡丹江师范学院工学院, 牡丹江 157012
2. 牡丹江大学信息与电器工程学院, 牡丹江 157011
3. 青岛科技大学信息科学技术学院, 青岛 266061

**摘要** 鉴于应用支持向量机进行说话人识别过度依赖于选择核函数的问题, 提出一种基于组合核函数支持向量机(SVM)的说话人识别方法。对多项式核函数、径向基核函数进行线性加权, 构建既具有全局核函数优点又具有局部核函数优点的组合核函数, 并通过多重网格搜索调节权重系数使组合核函数适用于当前数据分布, 确定组合核函数SVM的最优参数, 实现对说话人的有效识别。对TIMIT数据集和含噪声数据集的仿真实验显示, 基于组合核函数SVM的说话人识别性能明显优于单一的多项式核函数、径向基核函数和线性核函数。

**关键词** 说话人识别; 支持向量机; 组合核函数; 多重网格搜索

**中图分类号** TN912.34

**文献标志码** A

**doi** 10.3981/j.issn.1000-7857.2015.01.016

## Speaker recognition method based on combination of kernel functions of SVM

FAN Chijie<sup>1</sup>, SI Qiaomei<sup>1</sup>, XU Yan<sup>2</sup>, ZHANG Dan<sup>1</sup>, CAI Chunhua<sup>1</sup>, YU Xu<sup>3</sup>

1. School of Engineering, Mudanjiang Normal University, Mudanjiang 157012, China
2. School of Information and Electrical Engineering, Mudanjiang University, Mudanjiang 157011, China
3. School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China

**Abstract** In speaker recognition systems, if the original data distribution is unknown, the choice of inappropriate kernel functions will result in poor support vector machine (SVM) learning performance. Thus a speaker recognition method based on a multi-grid search of parameters and a combination of kernel functions is proposed in this paper. First, the method constructs a hybrid kernel function by linearly weighted polynomial and RBF kernels. Then it proposes a multi-grid search method to adjust the weights, and thus the hybrid kernel function can adapt to the current data distribution. Finally, a SVM classifier is trained to obtain the classification results. Simulation experiments on TIMIT datasets and noisy datasets show that the recognition performance of SVM classifiers using a combination of kernel functions is better than that using linear kernels, polynomial kernels, and RBF kernels. Therefore, the proposed method can effectively improve the performance of speaker recognition systems.

**Keywords** speaker recognition; support vector machine; combination of kernel functions; multi-grid search

利用机器学习算法进行说话人识别<sup>[1-3]</sup>是近年来的热门研究领域。所谓说话人识别, 就是根据测试语音判断说话人是训练集中的哪一位。说话人识别有着广泛的用途, 例如

在司法领域, 可以根据现场留下的声音判断指证疑犯; 对于家居领域, 根据语音识别可以有效地取代密码门和防盗门, 使智能大厦和智能家居变得越来越普及。说话人识别系统

收稿日期: 2014-06-20; 修回日期: 2014-11-07

基金项目: 黑龙江省教育厅科学技术研究项目(12533074)

作者简介: 樊持杰, 副教授, 研究方向为数据挖掘与支持向量机, 电子信箱: fanchijie@163.com

引用格式: 樊持杰, 司巧梅, 徐岩, 等. 基于组合核函数SVM的说话人识别方法[J]. 科技导报, 2015, 33(1): 90-94.

可按其识别、训练时所用文本是否相同,分为文本相关(text-dependent)和文本无关(text-independent)两种类型。

目前说话人识别技术主要分为三大类:基于声学特征的长时平均、说话人相关特征建模和基于有监督的学习策略。基于声学特征的长时平均的说话人识别方法是最早出现的识别方法,其基本思想是:频谱表示或基音等声学特征经过长时平均后,滤除了语音变化对声学特征的影响,剩下的就与说话人相关的部分。对谱特征而言,长时平均代表了说话人的声道形状。Kinnunen等<sup>[4]</sup>指出该方法在一些比较难的与文本无关的说话人辨识系统中应用得比较成功,并分析了这种方法的不足在于丢失了太多的说话人相关信息,必须有较长(>20 s)的语音才能获得稳定的长时语音统计。

不同于基于声学特征长时平均的识别策略,说话人相关特征建模方法进行识别时,将测试语句中的语音声学特征和特定说话者的包含相同语音内容的模型相比较。该识别方法的优势在于能够较好地把握住说话人的差异。Sakoe等<sup>[5]</sup>借助于动态时间归整方法进行说话人识别,Togneri等<sup>[6]</sup>总结了基于隐马尔可夫模型进行说话人相关特征建模的方法。对于无限制文本,必须在训练或识别之前进行语音切分,由于显式切分计算复杂度较高,非显式的切分方法应用更为普遍。Rosenberg等<sup>[7]</sup>提出基于矢量量化方法,Higgins等<sup>[8]</sup>提出基于K最邻近原则方法,都可以归为基于非显式切分进行说话人识别的方法。

基于有监督的学习策略是一类最新的方法,其识别的过程是一个训练样本的学习过程。这就要求有充分的样本供分类器学习,目前用的较多的分类器包括神经网络、支持向量机和决策树等。该类方法对样本数量的要求比较高,充足的样本可以使得各种分类模型都取得较好的效果。当样本数目不是特别多时,确定分类模型最有效的参数就成为这类方法取得较好效果的关键。

Wang等<sup>[9]</sup>研究了说话人识别神经网络算法的性能。首次应用支持向量机(support vector machine, SVM)进行说话人识别的工作来自Gish等<sup>[2]</sup>,然而由于SVM的性能较大程度上依赖于核函数的选择,所以应用SVM进行说话人识别在识别性能上经常会有波动<sup>[10]</sup>,当选择的核函数比较适合于当前训练数据时其识别效果比较好,当选择的核函数不太适合于当前训练数据时其识别效果则往往比较差<sup>[11,12]</sup>。鉴于此问题,本研究将全局核函数与局部核函数相结合,提出一种基于组合核函数SVM的说话人识别方法。

## 1 组合核函数构建

任何一个满足Mercer条件的正定函数都可作为SVM的核函数,建立从原始样本空间到特征空间的一种隐式映射。常见的两种类型的核函数为多项式核函数和径向基核函数,其表达式分别为

$$K(\mathbf{x}, \mathbf{z}) = [(\mathbf{x} \cdot \mathbf{z}) + 1]^d \quad (1)$$

$$K(\mathbf{x}, \mathbf{z}) = \exp(-\|\mathbf{x} - \mathbf{z}\|^2 / \sigma^2) \quad (2)$$

SVM分类算法的核心在于选择核函数,不同的核函数对应的SVM性能有着很大的差别,核函数的选择一直是SVM研究的难点。多项式核函数和径向基核函数是最为常用的两种核函数,前者为全局核函数,后者为局部核函数,两者分别适用于不同的数据分布。为了使核函数既有全局核函数的优点又具有局部核函数的优点,将多项式核函数与径向基核函数进行线性加权组合,构建的组合核函数为

$$K(\mathbf{x}, \mathbf{z}) = u[(\mathbf{x} \cdot \mathbf{z}) + 1]^d + v \exp(-\|\mathbf{x} - \mathbf{z}\|^2 / \sigma^2) \quad (3)$$

式中, $u, v$ 为加权系数,且为正数; $d$ 为多项式核函数的幂指数; $\sigma$ 为径向基核函数的宽度参数。下面对式(3)所示组合核函数进行证明。

设给定一个有限点的数据集合 $\{x_1, x_2, \dots, x_l\}$ ,  $K_1, K_2$ 分别为多项式核函数和径向基函数相对于这个集合的Gram矩阵。对正数 $u, v$ 和任意向量 $\beta \in R^l$ ,其中 $R^l$ 为 $l$ 维欧氏空间,有

$$\beta^T (uK_1 + vK_2) \beta = u\beta^T K_1 \beta + v\beta^T K_2 \beta \geq 0 \quad (4)$$

因此证明式(3)为核函数。

构建的组合核函数是将全局核函数与局部核函数线性加权,当加权系数 $u=0$ 时组合核函数退化为径向基核函数,当加权系数 $v=0$ 时组合核函数退化为多项式核函数,通过适当调整 $u$ 和 $v$ ,使组合核函数适用于不同的数据分布。

## 2 组合核函数SVM及算法

### 2.1 组合核函数SVM模型

设训练样本集合为 $\{(x_i, y_i), i=1, 2, \dots, l\}$ , 则组合核函数SVM模型的优化问题可描述为

$$\min_a \frac{1}{2} \sum_{i,j=1}^l a_i y_i a_j y_j K(x_i, x_j) - \sum_{j=1}^l a_j \quad (5)$$

且目标参数需满足约束条件

$$\sum_{i=1}^l a_i y_i = 0 \quad a_i \in [0, C] \quad (6)$$

式中, $K(x_i, x_j) = u[(x_i \cdot x_j) + 1]^d + v \exp(-\|x_i - x_j\|^2 / \sigma^2)$ 表示组合核函数在训练样本 $x_i$ 和 $x_j$ 上的取值; $l$ 为训练样本个数; $a_i, a_j$ 为拉格朗日乘子; $\mathbf{a}$ 为拉格朗日乘子组成的优化向量, $\mathbf{a} = (a_1, a_2, \dots, a_l)$ ;  $y_i, y_j$ 为样本类别标签; $x_i, x_j$ 为训练样本特征向量; $C$ 为惩罚因子。

当求出优化问题的解 $a_i$ 之后,组合核函数SVM模型的决策超平面可表示为

$$f(\mathbf{x}) = \text{sgn} \left[ \sum_{i=1}^l a_i y_i K(\mathbf{x}, \mathbf{x}_i) + y_j - \sum_{i=1}^l y_i a_i K(\mathbf{x}, \mathbf{x}_i) \right] \quad (7)$$

### 2.2 最优参数确定

组合核函数SVM模型包含5个参数 $u, v, \sigma, d$ 和惩罚因子 $C$ ,需要确定最优参数以使模型达到最佳效果。

关于SVM最优参数选取,国际上尚无公认方法,目前常用的方法有基于网格搜索的参数寻优和基于各种优化算法的参数寻优。基于网格搜索的SVM最优参数确定方法的主要思路是,首先将参数 $C$ 和 $\sigma$ 在一定范围内取值,对于取定的 $C$ 和 $\sigma$ ,把训练集作为原始数据集利用 $K$ 折交叉验证法得到在

此组  $C$  和  $\sigma$  下的训练集验证分类准确率, 最终取使得训练集验证分类准确率最高的一组  $C$  和  $\sigma$  作为最优参数。由于网格寻优通常可并行操作, 基于网格搜索寻找最优参数所需时间不比基于优化算法寻优时间多, 是网格搜索的优点。

本研究采用多重网格搜索确定最优参数。对于组合核函数 SVM 模型的 5 个参数  $u, v, \sigma, d, C$  确定搜索空间, 借鉴 SVM 经典求解软件 LibSVM 的策略, 采用指数范围搜索, 将参数  $C$  和  $\sigma$  的变化范围设定为  $2^{-10}, 2^{-9}, \dots, 2^0$ ; 将  $d$  的变化范围设定为  $2^0, 2^1, 2^2, \dots, 2^8$ ; 将  $u, v$  设定为开区间  $(0, 1)$  内的正数, 搜索步长为 0.1。该方法的主要特点为

1) 如果存在多组参数值对应最高的分类准确率, 一次网格搜索法选取第一组参数值, 而多重网格搜索法选取参数  $C$  最小的一组参数值。因为  $C$  如果设置过大会导致学习产生过拟合, 所以从达到最高验证分类准确率的参数组合中选取  $C$  最小的一组参数值是最佳选择。

2) 一次网格搜索寻优, 尤其是对于具有指数范围的参数而言, 一次步长的调整导致搜索范围跨越过大, 很难搜索到全局最优解。而多重网格搜索, 是从第一次网格寻优确定的最优点开始, 进行一次或多次网格寻优, 此时搜索范围局限在第一次搜索确定的最优点附近, 使搜索范围更加精细, 增大了寻找到全局最优解的可能性。

### 2.3 算法设计

组合核函数 SVM 模型算法设计如下。

输入原始样本集合  $\{(\mathbf{x}_i, y_i), i = 1, 2, \dots, l\}$ , 其中  $y_i$  为样本类别标签,  $\mathbf{x}_i$  为训练样本的特征向量,  $l$  为训练样本数。

输出分类决策函数:

步骤 1, 将数据集分成 10 份, 轮流将其中 9 份作为训练数据, 1 份作为测试数据, 进行 10 次实验;

步骤 2, 对于每一次实验, 根据多重网格搜索寻优方法, 计算组合核函数 SVM 优化问题, 即

$$\begin{aligned} \min_a & \frac{1}{2} \sum_{i,j=1}^l a_i y_i a_j y_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^l a_j \\ \text{s.t.} & \sum_{i=1}^l y_i a_i = 0 \quad 0 \leq a_i \leq C \end{aligned} \quad (8)$$

可得优化问题的解为

$$\mathbf{a}^* = (a_1^*, a_2^*, \dots, a_l^*)^T \quad (9)$$

步骤 3, 计算位于开区间  $(0, C)$  的  $\mathbf{a}^*$  的任意一个分量  $a_j^*$ , 并据此计算

$$b^* = y_j - \sum_{i=1}^l y_i a_i^* K(\mathbf{x}_i, \mathbf{x}_j) \quad (10)$$

构建决策函数, 表达式为

$$f(\mathbf{x}) = \text{sgn}[g(\mathbf{x})] \quad (11)$$

式中,

$$g(\mathbf{x}) = \sum_{i=1}^l a_i^* y_i K(\mathbf{x}_i, \mathbf{x}) + b^* \quad (12)$$

步骤 4, 根据每一次实验的分类精度, 确定最优参数, 并利用最优参数进行 SVM 学习, 得到相应的决策函数。

算法流程如图 1 所示。

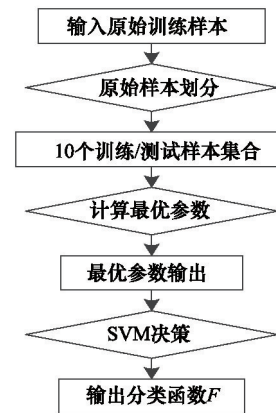


图 1 算法流程

Fig. 1 Flow chart of the algorithm

### 3 实验分析

实验平台为 Intel Core2 Duo CPU T6500, 2.10 GHz, 2.00 GB RAM, Windows 7 操作系统, 选择 Matlab 7.0 软件进行实验。

#### 3.1 客户流失数据集实验

##### 3.1.1 数据来源

实验的客户流失数据集为 Churn 数据集, 来自 UCI 标准机器学习数据库。该数据集将客户流失定义为无线手机客户不再使用电信公司的服务, 其中数据窗口为连续 3 个月, 将 3 个月期间的客户通话记录、账单及客户服务等数据转换为各个属性值作为模型输入, 模型输出被定义为第 4 个月的客户流失状态。将流失的客户定义为正类, 保留的客户作为负类。该电信公司客户流失数据集包含 2850 个负类样本, 483 个正类样本, 负类与正类之比达到 5.9, 是一类典型的不平衡数据集。本实验从正类样本中随机抽取 50 个作为训练样本, 25 个作为测试样本, 从负类中随机抽取 1000 个作为训练样本, 500 个作为测试样本。正负样本比例为 1:20。

##### 3.1.2 性能评价指标

由于实验是一个不平衡数据分类实验, 因此传统算法的分类准确率评价指标不能完全反映出分类器的性能, 为有效进行不平衡数据分类问题的分类器性能评价, 实验选取正确率  $P$ 、召回率  $R$  和  $F_1$  值<sup>[13]</sup>作为分类性能评价指标, 计算公式为

$$P = \frac{n_1}{n_2} \quad R = \frac{n_1}{n_3} \quad F_1 = \frac{2RP}{R+P} \quad (17)$$

式中,  $n_1$  为事实属于此类且分类正确的样本数;  $n_2$  为被判为此类的样本数;  $n_3$  为属于此类的总样本数。

可以看出, 只有算法的正确率和召回率都较高时, 算法的性能才更优越。  $F_1$  同时兼顾了分类算法的准确率和召回率, 可以看作是准确率和召回率的一种加权平均, 其最大值为 1, 最小值为 0。

##### 3.1.3 实验方法与结果分析

采用 C-SVC 算法作为分类算法, 利用 10 折交叉验证<sup>[14]</sup>确

定最适合分类算法的参数值,对组合核函数和径向基核函数、多项式核函数、线性核函数分别进行5次实验,取5次实验的平均值。各核函数在正类上的实验结果如图2所示,在负类上的实验结果如图3所示。从图2和图3可以看出,在客户流失数据集上,利用多重网格搜索确定最优参数的组合核函数在分类性能上明显优于其他核函数。充分说明组合核函数可以更好地对原始数据进行拟合,对于数据分布未知的分类问题往往能取得更好的效果。

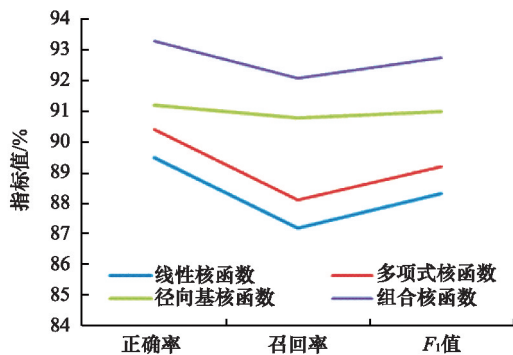


图2 不同核函数在正类上的分类性能

Fig. 2 Comparison of classification performance of different kernel functions on positive class

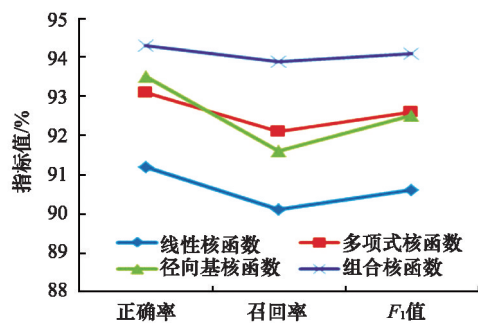


图3 不同核函数在负类上的分类性能

Fig. 3 Comparison of classification performance of different kernel functions on negative class

### 3.2 TIMIT数据集实验

#### 3.2.1 不同核函数实验

实验数据来源于 TIMIT 语料数据库<sup>[15]</sup>,该数据库共包含 630 个说话人和 2 个目录 (TRAIN 目录和 TEST 目录),每个目录下又分别包含 8 个文件夹 (DR1~DR8)。每个说话人 10 个句子,每个句子约 3 s。选取 TEST 目录中的 168 个说话人模型,每个说话人模型有 5 个训练句子 (2SA+3SI) 和测试句子 (5SX)。

实验用滤波器为

$$H(z) = 1 - 0.97z^{-1} \quad (18)$$

用该滤波器对语音进行预处理,以帧长 25 ms、帧移 10 ms 分帧,提取 12 维 MFCC 特征和 12 维一阶差分 MFCC 特征<sup>[16]</sup>。

选择 C-SVC 作为分类算法,分别对组合核函数和径向基核函数、线性核函数、多项式核函数在说话人识别数据集上的识别性能进行实验对比。实验涉及多个类别,因此选择成对分类策略<sup>[17]</sup>,即从多个类别中每次抽取 2 个类别,故对  $M$  类问题,需要做  $M(M-1)/2$  次分类。分别进行 10 次实验,取实验结果的平均值作为最终识别率。实验结果如表 1 所示,其中的参数值是经过 10 折交叉验证确定的最优参数。

表 1 TIMIT 数据集上不同核函数的识别精度

Table 1 Classification precision of different kernel functions on TIMIT dataset

核函数	最优参数	识别率/%
线性核函数	$C=512$	82.25
多项式核函数	$d=10, C=32$	87.10
径向基核函数	$\sigma=0.12, C=10$	91.25
组合核函数	$u=0.8, v=0.2, \sigma=0.1, d=6, C=8$	97.36

从表 1 可以看出,在核函数各自选用最优参数的情况下,对基于核函数 SVM 的说话人识别系统的识别性能而言,基于径向基核函数方法的识别性能优于多项式核函数和线性核函数,而基于组合核函数的识别性能明显优于径向基核函数。这主要是由于组合核函数通过调节线性加权系数能够较好地将局部核函数与全局核函数的优点综合,可更有效地拟合原始数据,而且利用多重网格搜索寻找最优参数,因此达到更高的识别精度。

#### 3.2.2 不同分类算法实验

选用 BP 神经网络算法、决策树 C4.5 算法与本研究算法的识别精度进行实验对比。训练数据和测试数据与上述核函数实验相同。

本研究算法依然选取上述核函数实验中组合核函数的最优参数,即  $u=0.8, v=0.2, \sigma=0.1, d=6, C=8$ 。对于 BP 算法,设定网络结构  $n=3$  层,网络权重(初始化) $w=0$ ,学习率  $h=0.1$ 。对于决策树 C4.5 算法,设定拆分方式为完全拆分  $S=2$ ,子节点包含最少子节点数  $M=10$ ,决策树的复杂度  $O=0.5$ 。实验结果如表 2 所示。从表 2 可以看出,相对于神经网络 BP 算法和决策树 C4.5 算法,本研究算法具有更好的说话人识别精度。这主要是由于本研究算法能够较好地拟合原始数据,一定程度上避免数据的过拟合,提高了识别性能,而 BP 算法和 C4.5 算法容易陷入局部极小值,导致过拟合,从而降低了识别性能。

表 2 TIMIT 数据集上不同算法的识别精度

Table 2 Classification precision of different classifiers on TIMIT dataset

分类算法	最优参数	识别率/%
BP 算法	$n=3, h=0.1, w=0$	85.5
C4.5 算法	$S=2, M=10, O=0.5$	89.2
本算法	$u=0.8, v=0.2, \sigma=0.1, d=6, C=8$	97.36

### 3.3 含噪声数据集实验

实验采用含噪声数据集,分别对组合核函数方法和传统的基于径向基核函数、线性核函数和多项式核函数方法的性能进行测试对比。

实验由80个说话人组成(其中男女各40人),每个说话人说出6个句子,其中2个句子是每个人都要说,另外4个句子则每个人不同。每个说话人随机选取3个句子组成训练集,剩下的3个句子组成测试集。采用的语音为纯净语音再附加各种信噪比的高斯白噪声,在3个月内分3次进行录制,训练语音时长为10 s,用时长为2 s的语音进行识别。采样率为11025 Hz,帧长为30 ms,帧移为15 ms,进行 $1-0.95z^{-1}$ 预加重,用汉明窗加窗。

选择C-SVC作为分类算法,信噪比取5、10、15、20 dB,分别对基于不同核函数方法的识别率进行实验测试。分别进行10次实验,取其平均值作为最终识别率。实验结果如表3所示,其中参数值是经过10折交叉验证确定的最优参数。

表3 含噪声数据集上不同核函数的识别精度

Table 3 Classification precision of different kernel functions on noisy dataset

核函数	最优参数	识别率/%			
		5 dB	10 dB	15 dB	20 dB
线性核函数	$C=256$	75.1	82.8	88.2	91.8
多项式核函数	$d=8, C=64$	77.6	85.1	89.6	95.2
径向基核函数	$\sigma=0.12, C=8$	78.2	86.3	91.2	96.5
组合核函数	$u=0.6, v=0.4, \sigma=0.1, d=6, C=10$	83.8	90.1	95.3	98.2

从表3可以看出,各种基于核函数方法的识别率随着信噪比值的增加而增大。但在信噪比值较小时,基于传统核函数方法的识别率较低,而基于组合核函数的方法则具有较好的抗噪声性能,能够有效提高系统的鲁棒性。

## 4 结论

基于组合核函数SVM的说话人识别方法,是将属于全局核函数的多项式核函数和属于局部核函数的径向基核函数进行线性加权组合,通过多重网格搜索可以有效地确定最优参数。由于设定有效参数值可使组合核函数适用于不同数据分布,因此在原始数据分布未知的情况下,本研究方法可有效提高说话人识别精度。TIMIT数据集和含噪声数据集实验表明,基于组合核函数SVM的识别性能明显优于线性核函数、多项式核函数及径向基核函数。未来研究方向是直接从原始数据中进行学习,更有针对性地构建适合于现有数据集的核函数。

### 参考文献(References)

[1] Reynolds D A, Rose R C. Robust text-independent speaker identification

using Gaussian mixture speaker models[J]. IEEE Transactions on Speech and Audio Processing, 1995, 3(1): 72-83.

[2] Gish H, Schmidt M. Text-independent speaker identification[J]. IEEE Signal Processing Magazine, 1994, 11(4): 18-32.

[3] 张亮. 说话人识别中语音增强算法的研究和系统实现[D]. 重庆: 重庆大学, 2009.

Zhang Liang. Speech enhancement algorithm research and system implementation for speaker recognition[D]. Chongqing: Chongqing University, 2009.

[4] Kinnunen T, Li H. An overview of text-independent speaker recognition: From features to supervectors[J]. Speech Communication, 2010, 52(1): 12-40.

[5] Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1978, 26(1): 43-49.

[6] Togneri R, Pullella D. An overview of speaker identification: Accuracy and robustness issues[J]. IEEE Circuits and Systems Magazine, 2011, 11(2): 23-61.

[7] Rosenberg A, Soong F. Evaluation of a vector quantization talker recognition system in text independent and text dependent modes[J]. Computer Speech and Language, 1987, 22(4): 143-157.

[8] Higgins A L, Bahler L G, Porter J E. Voice identification using nearest-neighbor distance measure[C]. IEEE International Conference on the Acoustics, Speech, and Signal Processing, Minneapolis, USA, April 27-30, 1993.

[9] Wang G W, Luo S X, He L, et al. Application BP neural network in the speaker recognition based on chaos particle swarm optimization algorithm[J]. Advanced Materials Research, 2013, 765: 2805-2808.

[10] 刘雪燕, 李明, 张亚芬. 基于PCA和多约简SVM的多级说话人辨识[J]. 计算机应用, 2008, 28(1): 127-130.

Liu Xueyan, Li Ming, Zhang Yafen. Hierarchical speaker identification based on PCA and multi-reduced SVM[J]. Computer Applications, 2008, 28(1): 127-130.

[11] You C H, Lee K A, Li H. GMM-SVM kernel with a Bhattacharyya-based distance for speaker recognition[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2010, 18(6): 1300-1312.

[12] Fisher W M, Zue V, Bernstein J, et al. An acoustic-phonetic data base [J]. Journal of the Acoustical Society of America, 1987, 81(Suppl 1): 92-93.

[13] Vapnik V. The nature of statistical learning theory[M]. Berlin: Springer Publishing Company, 2000.

[14] 兰均, 施化吉, 李星毅, 等. 基于特征词复合权重的关联网页分类[J]. 计算机科学, 2011, 38(3): 187-190.

Lan Jun, Shi Huaji, Li Xingyi, et al. Associative web document classification based on word mixed weight[J]. Computer Science, 2011, 38(3): 187-190.

[15] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]. 14th International Joint Conference on Artificial Intelligence, Adelaide, Australia, December 10-14, 1995.

[16] Nakagawa S, Wang L, Ohtsuka S. Speaker identification and verification by combining MFCC and phase information[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(4): 1085-1095.

[17] Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415-425.

(责任编辑 韩星明)