

联合 HMM-UBM 与 RVM 的声纹密码识别算法

胡志隆^{1a}, 文 畅^{1b}, 谢 凯^{1a}, 贺建彪²

(1. 长江大学 a. 电子信息学院; b. 计算机科学学院, 湖北 荆州 434023;

2. 中南大学 信息科学与工程学院, 长沙 410083)

摘 要: 针对声纹密码识别中声纹文本信息利用率低和噪音干扰的问题, 提出隐马尔科夫模型-通用背景模型(HMM-UBM)融合相关向量机(RVM)的声纹识别算法。利用 HMM-UBM 对语音信号进行时序建模, 使用 RVM 学习得到每位注册话者语音的分类信息。通过对待识别话者建立 HMM 模型, 并将 RVM 作为分类器进行判决决策得到分类结果。实验结果表明, 与 GMM-UBM 算法和 GMM-SVM 算法相比, 在无噪声环境下, 该算法错误接收概率降低 7%~9%, 识别正确率提高 4%~5%, 在低信噪比环境下, 其识别正确率提高 5%~12%。

关键词: 声纹识别; 隐马尔科夫模型; 通用背景模型; 相关向量机; 语音信号

中文引用格式: 胡志隆, 文 畅, 谢 凯, 等. 联合 HMM-UBM 与 RVM 的声纹密码识别算法[J]. 计算机工程, 2018, 44(11): 129-134.

英文引用格式: HU Zhilong, WEN Chang, XIE Kai, et al. Voiceprint password recognition algorithm fusing with HMM-UBM and RVM[J]. Computer Engineering, 2018, 44(11): 129-134.

Voiceprint Password Recognition Algorithm Fusing with HMM-UBM and RVM

HU Zhilong^{1a}, WEN Chang^{1b}, XIE Kai^{1a}, HE Jianbiao²

(1a. School of Electronic Information; 1b. School of Computer Science, Yangtze River University, Jingzhou, Hubei 434023, China;

2. College of Information Science and Engineering, Central South University, Changsha 410083, China)

[Abstract] Considering shortcomings of noise interference and low utilization in voiceprint recognition, a voiceprint recognition algorithm which is based on Hidden Markov Model-Universal Background Model(HMM-UBM) and Relevant Vector Machine(RVM) is proposed. The algorithm uses HMM-UBM to model speech signals in time sequence, and learn the classified information of each registered speaker by using RVM. It is established the HMM model for the speaker who needs to be identified, and the results of classification is obtained by using RVM as classifier. Experimental results show that compared with the GMM-UBM algorithm and GMM-SVM algorithm in a noiseless environment, the probability of error acceptance of proposed algorithm is reduced by 7%~9% and the accuracy of recognition is improved by 4%~5%. In low SNRR environment, the accuracy of recognition is improved by 5%~12%.

[Key words] voiceprint recognition; Hidden Markov Model(HMM); Universal Background Model(UBM); Relevance Vector Machine(RVM); speech signal

DOI: 10.19678/j.issn.1000-3428.0049075

0 概述

语音识别是一种生物识别技术, 获取方法简单且成本低廉。相比于人脸识别、虹膜识别等技术, 语音识别使用者的接受程度更高, 因此, 其被广泛应用于医疗、社保、金融及公共场所的安全认证等领域。而声纹密码识别作为文本相关的语音识别方法, 用文本的上下文关系和话者声道信息保护说话人的信息安全, 与文本无关的语音识别方法相比, 具有较高

的安全性^[1]。

目前, 语音识别系统常采用高斯混合模型-通用背景模型(Gaussian Mixture Model-Universal Background Model, GMM-UBM), 该算法应用话者的声道信息并通过全局背景模型解决训练样本少的问题, 在文本无关的语音识别中取得较好的识别效用。此外, 随着人工神经网络(Artificial Neural Network, ANN)、支持向量机(Super Vector Machine, SVM)等机器学习方法的不断成熟, 也出现类似 GMM-SVM 的融合

基金项目: 国家自然科学基金(61272147); 湖北省教育厅科学技术研究计划指导性项目(B2015446)。

作者简介: 胡志隆(1994—), 男, 硕士, 主研方向为密码技术、图形图像处理; 文 畅(通信作者), 讲师、硕士; 谢 凯, 教授、博士; 贺建彪, 副教授、博士。

收稿日期: 2017-10-25

修回日期: 2017-12-04

E-mail: wenchang2016paper@163.com

算法^[2-6]。然而,GMM 模型仅单一反映话者的声道信息而忽略语音文本的上下文关系,因此不适用于声纹密码识别。同时,ANN 是一个高度非线性的大型网络,需要大量的训练样本才能得到效果较好的模型,导致其难以应用于实际。

针对上述方法的不足,本文提出一种隐马尔科夫模型-通用背景模型(Hidden Markov Model-Universal Background Model,HMM-UBM)联合相关向量机(Relevance Vector Machine,RVM)的声纹密码识别算法,该算法采用 HMM-UBM 模型,利用语音的文本信息及其话者的声道信息进行时序建模。同时针对 HMM 模型分类决策能力较差的问题,本文融合相关向量机作为分类器,做最后的判决决策。

1 声纹密码识别算法流程

指定文本的声纹密码识别系统主要流程包括语音信号的预处理、特征参数提取、特征建模和相

似性度量等模块,其中特征参数的选取及相似性度量量的方法决定系统的识别效率^[7]。本文针对传统声纹密码识别系统中的相似性度量模块进行了改进,提出了基于 HMM-UBM 联合改进 RVM 的声纹密码识别算法,算法流程如图 1 所示。首先对注册语音、训练语音及待识别语音进行预处理,包括分帧、加窗、预加重等;其次采用梅尔倒谱系数(Mel-Frequency Cepstral Coefficients,MFCC)对处理过的语音提取特征参数;然后将其作为输入序列,训练得到 HMM-UBM 模型后计算每位注册说话人语音与训练语音的匹配得分,归一化后组合成一个特征向量,将每位注册说话人的特征向量提供给 RVM 训练,得到语音分类信息,即 RVM 分类器;最后对待识别语音采用同样的方法得到其对应的特征向量,输入训练好的 RVM 模型进行分类决策,并最终取得分类结果。

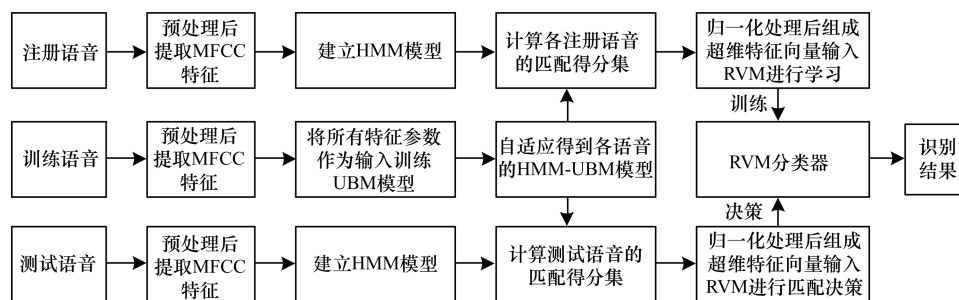


图 1 声纹密码识别算法流程

2 HMM-UBM 模型

说话人识别系统中常用的 GMM-UBM 模型虽然拟合了说话人的声音特性,但忽视了文本内容对识别效果的影响,不适用于特定文本的声纹密码识别^[8-10]。因此,本文采用 HMM 来对相关性进行建模,同时采用全局高斯混合模型作为 UBM 来表达说话人声道特征在训练样本中的分布概率。模型构建流程如图 2 所示。

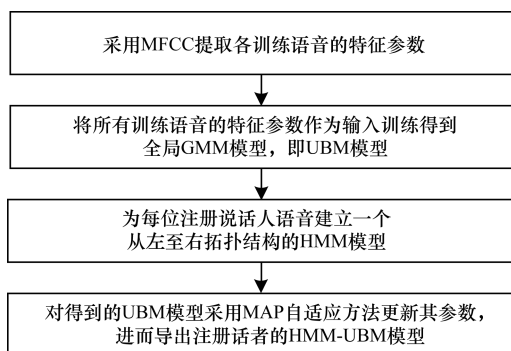


图 2 HMM-UBM 模型构建流程

与 GMM-UBM 模型相似,由于 UBM 的均值矢量利用率较高,因此自适应时仅更新该参数,更新过程如下:

1) 设注册话者的输入特征矢量为 $\{x_i | i = 1, 2, \dots, t\}$, 计算其在所有训练样本中的概率分布,若话者对应第 i 个训练样本,则其概率分布为:

$$P_r(i|x_t) = \frac{\omega_i p_i(x_t)}{\sum_{j=1}^M \omega_j p_j(x_t)} \quad (1)$$

其中, $p_i(x_t)$ 为第 i 个训练样本的密度函数, ω_i 为第 i 个训练样本的权重系数, $P_r(i|x_t)$ 为第 i 个训练样本的后验概率,表示在测试语音的特征矢量为 x_t 的条件下,测试语音对应第 i 个训练语音的概率。

2) 利用 $P_r(i|x_t)$ 和均值向量进行从分统计,统计式为:

$$E_i = \frac{1}{n_i} \sum_{t=1}^t P_r(i|x_t) x_t \quad (2)$$

3) 通过所有训练数据产生新的均值统计量,更新 UBM 第 i 个混合分量的均值矢量得到第 i 个分量

的 HMM-UBM 模型, 如式 3 所示。

$$\hat{\mu}_i = \alpha_i^m E_i(\mathbf{x}) + (1 - \alpha_i^m) \mu_i \quad (3)$$

3 相关向量机

RVM 是一种基于贝叶斯稀疏核的分类算法。与 SVM 相比, RVM 可以计算出样本输出的后验概率分布, 更适用于多分类问题, 并且其核函数不需要限定为对称正定核, 从而可以得到更加稀疏的解^[11-15]。因此, 本文选用其作为分类器, 进行最后的决策, 得到更加准确的识别结果。

RVM 是建立在贝叶斯理论上, 通过最大化参数概率分布的后验概率得到最优参数值的模型。设给定的训练样本集为 $\{\mathbf{x}_i, t_i | i = 1, 2, \dots, N\}$, $\mathbf{x} \in \mathbb{R}^d$ 为输入矢量序列, $t \in \{0, 1\}$ 代表类别号, 则目标 t_i 可以采用式(4)的回归模型来获得。

$$t_i = y(\mathbf{x}_i) + \varepsilon_n \quad (4)$$

其中, ε_n 是均值为 0、方差为 σ^2 的噪声, $y(\mathbf{x})$ 为 RVM 的分类模型, 其定义为:

$$y(\mathbf{x}) = \sum_{i=1}^N \omega_i K(\mathbf{x}, \mathbf{x}_i) + \omega_0 \quad (5)$$

其中, $K(\mathbf{x}, \mathbf{x}_i)$ 为核函数。本文采用高斯核函数 $G(\mathbf{x}_i, \mathbf{y}_j) = \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma_1^2})$, ω_i 为模型的权重, ω_0 为偏差。

设目标 $\{t | t = t_1, t_2, \dots, t_N\}$ 独立同分布, 则整个训练样本的似然函数可以表示为:

$$P(\mathbf{t} | \boldsymbol{\omega}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{t} - \boldsymbol{\varphi}\boldsymbol{\omega}\|^2\right\} \quad (6)$$

其中, $\mathbf{t} = (t_1, t_2, \dots, t_N)^T$, $\boldsymbol{\omega} = (\omega_0, \omega_1, \dots, \omega_N)^T$, $\boldsymbol{\varphi}$ 为 $N \times (N+1)$ 矩阵, 如式(7)所示。

$$\boldsymbol{\varphi}(\mathbf{x}) = \begin{bmatrix} 1 & K(\mathbf{x}_1, \mathbf{x}_1) & K(\mathbf{x}_1, \mathbf{x}_2) & \cdots & K(\mathbf{x}_1, \mathbf{x}_N) \\ 1 & K(\mathbf{x}_2, \mathbf{x}_1) & K(\mathbf{x}_2, \mathbf{x}_2) & \cdots & K(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K(\mathbf{x}_N, \mathbf{x}_1) & K(\mathbf{x}_N, \mathbf{x}_2) & \cdots & K(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \quad (7)$$

假设式(6)中的 $\boldsymbol{\omega}$ 和 σ^2 采用最大似然估计求解, 结果通常使权重参数 $\boldsymbol{\omega}$ 中大部分元素不为 0, 从而导致过拟合。为了避免过学习的问题, RVM 对每个权重参数加上先决条件: 使其几率是分布在 0 周围的正态分布, 如式(8)所示。

$$p(\boldsymbol{\omega} | \boldsymbol{\alpha}) = \prod_{u=0}^N (\omega_u | 0, \alpha_u^{-1}) \quad (8)$$

其中, $\boldsymbol{\alpha}$ 为 $N+1$ 维超向量。

根据贝叶斯公式直接求得参数 $\boldsymbol{\omega}$ 的后验分布:

$$p(\boldsymbol{\omega} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) = (2\pi)^{-(N+1)/2} |\boldsymbol{\Sigma}|^{-1/2} \times \exp\left\{-\frac{1}{2}(\boldsymbol{\omega} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\omega} - \boldsymbol{\mu})\right\} \quad (9)$$

其中, $\boldsymbol{\mu} = \sigma^{-2} \boldsymbol{\Sigma} \boldsymbol{\varphi}^T \mathbf{t}$, $\boldsymbol{\Sigma}$ 为协方差, $\boldsymbol{\Sigma} = (\sigma^{-2} \boldsymbol{\varphi}^T \boldsymbol{\varphi} + \mathbf{A})^{-1}$, $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$, $\sigma^2(\mathbf{x}) = (\boldsymbol{\beta})^{-1} + \boldsymbol{\varphi}(\mathbf{x}) \boldsymbol{\Sigma} \boldsymbol{\varphi}(\mathbf{x})$ 。

对于一个给定的样本 $\mathbf{x}$, 最终其输出的概率分布为:

$$p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2) = \int p(\mathbf{t} | \boldsymbol{\omega}, \sigma^2) p(\boldsymbol{\omega}, \boldsymbol{\alpha}) d\boldsymbol{\omega} = (2\pi)^{-N/2} \cdot |\sigma^2 \mathbf{I} + \boldsymbol{\varphi} \mathbf{A}^{-1} \boldsymbol{\varphi}^T|^{-1/2} \exp\left\{-\frac{1}{2} \mathbf{t}^T (\sigma^2 \mathbf{I} + \boldsymbol{\varphi} \mathbf{A}^{-1} \boldsymbol{\varphi}^T)^{-1} \mathbf{t}\right\} \quad (10)$$

式(10)中的未知量为超参数 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$, 则求解输出概率分布的问题转化为求解超参数 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$, 本文通过最大化法来求解参数, 如式(11)所示。

$$p(\mathbf{t} | \mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \int p(\mathbf{t} | \mathbf{x}, \boldsymbol{\beta}) p(\boldsymbol{\omega} | \boldsymbol{\alpha}) d\boldsymbol{\omega} \quad (11)$$

通过最大化式(11), 来更新 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的值:

$$\alpha_i^{\text{new}} = \frac{1 - \alpha_i \Sigma_{ii}}{m_i^2} \quad (12)$$

$$(\boldsymbol{\beta}^{\text{new}})^{-1} = \frac{\|\mathbf{t} - \boldsymbol{\varphi} \mathbf{m}\|^2}{N - \sum_i (1 - \alpha_i \Sigma_{ii})} \quad (13)$$

经过多次学习, 大部分超参数 α_i 会趋于无穷, 而对应的权重向量 $\omega_i = 0$, 少部分权重向量不为 0 的训练样本 $\mathbf{x}_i$ 即为相关向量, 通过相关向量得到其分类模型并作为最优分类超平面, 以对输入的测试样本进行识别。

4 HMM-UBM-RVM 声纹密码识别算法

HMM 算法具有较强的时序建模能力, 处理连续动态信号时表现优异。但 HMM 是基于先验知识的统计学习方法, 其分类决策能力较差。而 RVM 是基于贝叶斯稀疏核的回归分类算法, 具有较强的分类效果和泛化能力。本文提出 HMM-UBM-RVM 声纹识别算法, 将 2 种方法进行融合, 具有较强的时序建模能力和分类效果。

设训练样本中包含 n 位说话人, 每位说话人包含 10 条语音, 对于第 i 位说话人, 其特征序列分别为 $\{i\text{-voice0}, i\text{-voice1}, \dots, i\text{-voice9}\}$ 。其中 $i\text{-voice0}$ 为其注册声纹密码, 则该模型的实现流程如图 3 所示。

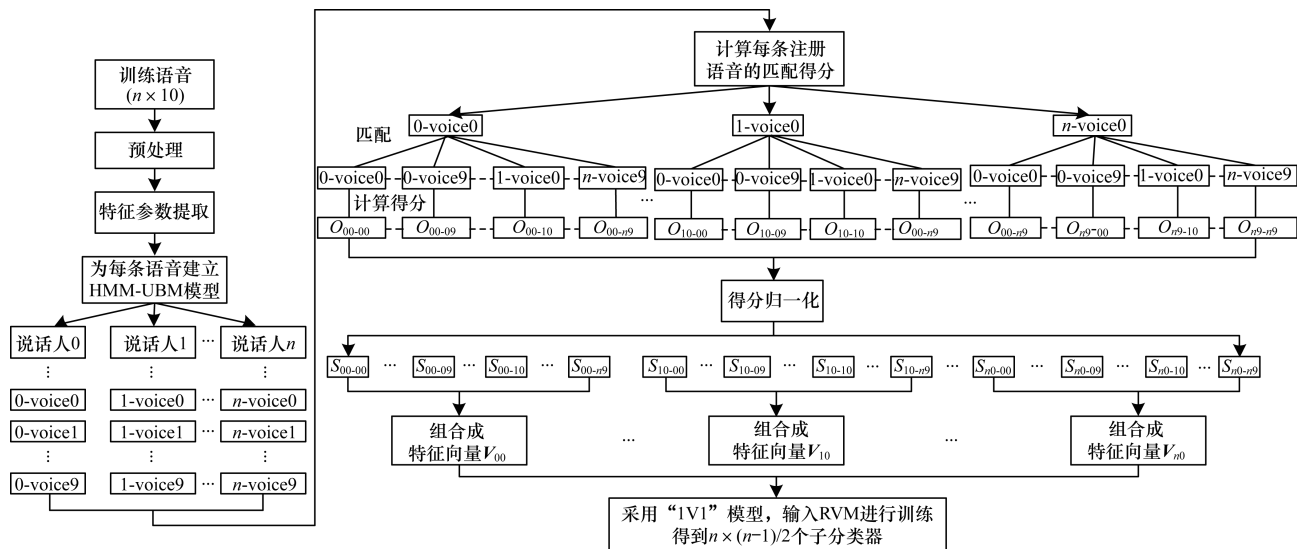


图 3 HMM-UBM 联合 RVM 实现流程

联合 HMM-UBM 与 RVM 算法实现流程如下:

- 1) 录制语音信号。对语音信号进行预处理,消除干扰信息后提取其 MFCC 特征参数。
- 2) 对每位说话人的语音建立 HMM-UBM 模型。
- 3) 用式(14)计算训练样本中每条注册语音对其他语音的匹配得分:

$$O_{t0-ik}(X_{t0}) = \log_p(X_{t0} | \lambda_{ik}) - \log_p(X_{t0} | \lambda_{UBM}) \quad (14)$$

其中, X_{t0} 为训练样本中第 t 位说话人注册语音的 HMM 特征序列, λ_{ik} 为训练样本中第 i 位说话人第 k 条语音的 HMM-UBM 特征序列, λ_{UBM} 为背景模型的特征序列。

- 4) 对匹配得分进行归一化:

$$Score_{t0-ik} = \frac{O_{t0-ik} - O_{\min}}{O_{\max} - O_{\min}} \quad (15)$$

5) 将每条注册语音归一化数据组成超维向量 V 。例如,对于第 t 位说话人,其注册语音的超维特征向量为 $V_{t0} = \{Score_{t0-00}, Score_{t0-01}, \dots, Score_{t0-n9}\}$ 。

6) 将得分矢量输入到 RVM 进行学习,直到 RVM 迭代次数到预设次数为止(本文取最大迭代次数为 300 次)。至此完成 RVM 分类器的训练阶段,得到 $n \times (n-1)/2$ 个二分类的子 RVM 分类器。

7) 对待识别的输入语音信号进行训练,得到 HMM 模型。用式(14)计算其得分,归一化后组成待识别语音的特征向量 V_{vp} , $V_{vp} = \{Score_{vp1}, Score_{vp2}, \dots, Score_{vpn}\}$ 。

8) 将待识别语音的特征向量 V_{vp} 输入到 RVM,然后对该向量进行非线性映射。为提高识别精度,本文采用“1V1”模型进行多分类识别,如图 4 所示。

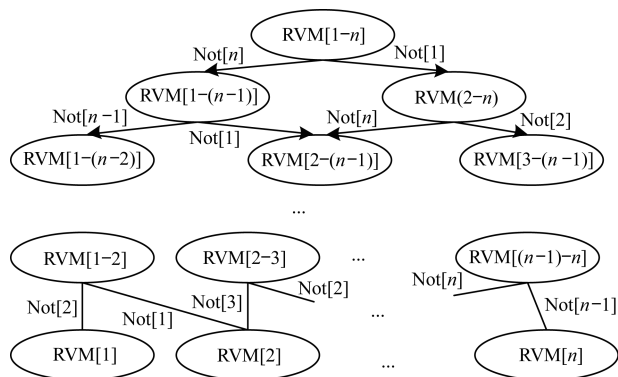


图 4 RVM“1V1”分类模型

5 实验结果与分析

采用本文设计的声纹密码识别系统作为测试平台,系统运行界面如图 5 所示。

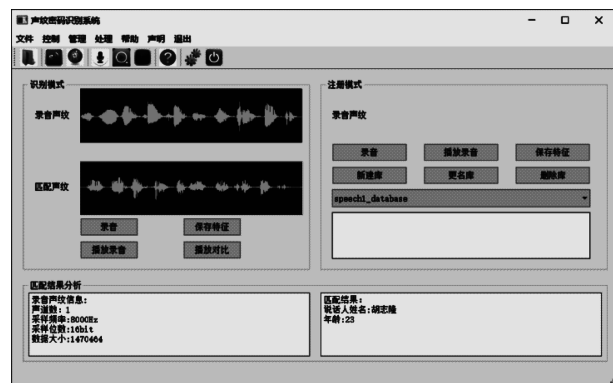


图 5 系统运行界面

5.1 数据库及评测方法

本文采用的数据库是 TIMIT 语音库,包含来自美国 8 个主要方言地区的 630 个人,每个人包含 10 段 3 s~6 s 的语音。其中,2 段为方言句子,每个人的方言句子内容相同,其余语音内容不同。

方言句 1 内容为“*She had your dark suit in greasy wash water all year*”,方言句 2 内容为“*Don't ask me to carry an oily rag like that*”。

本文采用错误接收概率 (FA) 和错误拒绝概率 (FR) 评判声纹密码识别系统的性能,其表达式如式 (16) 和式 (17) 所示。

$$F_{FA} = \frac{n_{NRVC}}{n_{NRVT}} \quad (16)$$

$$F_{FR} = \frac{n_{RVW}}{n_{RVT}} \quad (17)$$

其中, n_{NRVC} 表示测试语音为非注册语音时,识别为注册语音的概率, n_{NRVT} 表示采用非注册语音作为测试语音的实验次数, n_{RVW} 表示测试语音为注册语音时,识别错误的概率(当且仅当测试语音与识别出的注册语音内容和对应说话人均匹配时,才认为识别正确), n_{RVT} 表示采用注册语音作为测试语音的实验次数。

5.2 UBM 混合数对 HMM-UBM 模型的影响

通用背景模型的高斯混合数越大,说话人声道特征分布概率越精确,但计算复杂度相应增加。因此,本文针对这些参数做以下实验。

采用数据库中不含噪的 dr3 部分语音,以其中每个人的方言句 1 作为注册语料,以该语音库全部 6 300 条语音作为训练集训练 UBM 模型。实验中 HMM 的状态数及高斯混合度均取为 4,UBM 模型的高斯混合数分别取:8,16,32,64,128,256,512,1 024。

在测试时,随机选取 dr3 部分的 1 条语音作为测试语音进行识别,重复实验,取 FA、FR 均值。实验结果如图 6 所示。

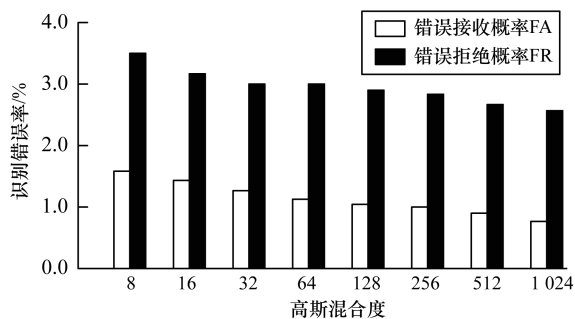


图 6 UBM 高斯混合度对识别率的影响结果

测试结果表明,随着通用背景模型的高斯混合度增高,错误接收概率及错误拒绝概率均有小幅度下降。当混合度超过 512 时,错误接收概率小于 1%,尽管错误拒绝率仍有 3% 左右,但错误拒绝对用户信息安全影响较小。因此,取 UBM 模型的高斯混合度为 512。

5.3 RVM 采用不同核函数对模型识别率的影响

采用数据库中不含噪的 dr3 部分语音,以其中每个人的方言句 1 作为注册语料,以该语音库全部

6 300 条语音作为训练集训练 UBM 模型。在实验中,RVM 分类器核函数分别选多项式核、sigmoid 核和高斯核,分别选取 dr3 语料中 15 位、30 位、45 位、60 位说话人的注册语音重复进行识别率测试。实验结果如表 1 所示。

表 1 RVM 采用不同核函数对模型识别率的影响

人数	识别率/%		
	多项式核	sigmoid 核	高斯核
15	95.33	97.00	100.00
30	92.17	95.83	99.17
45	90.11	93.22	98.89
60	87.00	90.08	97.92

从表 1 可以看出,随着测试人数的增加,高斯核的测试精度明显高于多项式核和 sigmoid 核。因此,RVM 均采用高斯核作为核函数进行分类决策。

5.4 系统识别率测试

本文所提出的声纹密码识别方法与常用于语音识别的 GMM-UBM、GMM-SVM 及 HMM-UBM 算法进行对比,实验采用 TIMIT 语音库中所有说话人的方言句 1 作为注册语音,为对应话者建立模型,采用全部 6 300 条语音作为训练集训练 UBM 模型,抽取每位话者的注册语音及随机两条其他语音进行识别,共测试 $630 \times 3 = 1\,890$ 次,取识别正确率、FA 及 FR 均值,实验结果如表 2 所示。

表 2 不同算法的识别率比较

算法	识别正确率	FA	FR
GMM-UBM 算法	90.42	10.63	9.08
GMM-SVM 算法	91.06	9.37	8.73
HMM-UBM 算法	91.16	5.79	5.40
HMM-UBM-RVM 算法	96.98	1.11	3.81

从表 2 可以看出,本文方法与 GMM-UBM 和 GMM-SVM 识别算法相比,大幅降低了识别算法的错误接收概率和错误拒绝概率。可以看出,针对文本相关的声纹密码识别,采用更具有时序建模能力的隐马尔科夫模型-通用背景模型,该模型可反映声纹的文本信息及话者的声道信息。相比之下,GMM 模型的 GMM-UBM 算法和 GMM-SVM 算法都忽视了文本信息,会将说话人的其他语音识别为其注册语音,从而导致错误接收概率较大,影响用户信息安全。同时,本文方法在 HMM-UBM 模型基础上采用 RVM 作为分类器,回避了该模型分类决策能力弱的问题。

5.5 系统抗噪性能测试

对测试语音进行高斯白噪声加噪处理,信噪比分别为 0 dB、5 dB、10 dB、15 dB、20 dB、25 dB、30 dB。系统抗噪性能测试结果如图 7 所示。

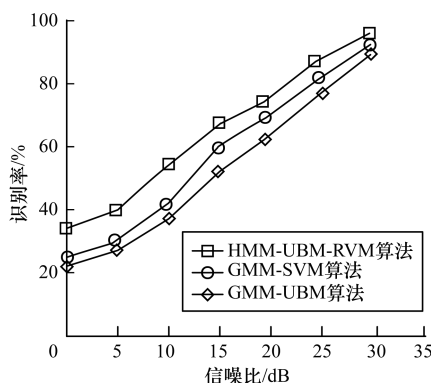


图7 系统抗噪性能测试结果

由图7可以看出,本文算法在各信噪比环境下,识别率均优于基于GMM-SVM和GMM-UBM识别算法。尤其在低信噪比环境下,其优越性更加明显。一方面是因为该方法采用HMM-UBM模型,兼顾语音的文本信息及话者的声道信息,更适用于文本相关的声纹密码识别;另一方面,由于HMM是基于先验知识的统计学习方法,并不具备良好的分类能力,将其与相关向量机相融合,发挥HMM的时序建模能力和RVM的分类决策能力,能够提高其抗噪性能识别精度。

5.6 应用测试

本文方法用于模拟门禁系统,实现声纹密码开锁。实验采用实验室中20位说话人(12男,8女),每位说话人50条语音,语音内容为8位0~9的随机数,取每位话者其中一条语音作为声纹密码,注册其模型,以全部1000条语音作为训练集,训练UBM模型。对每位说话人进行测试,测试语音采用说话人现场说出5条对应的注册语音及5条对应的非注册语音,记录正确开锁率(说话人与密码相匹配并开锁成功及说话人与密码不匹配并开锁失败2种情况视为正确开锁)。以基于GMM-UBM、GMM-SVM和HMM-UBM算法的模拟门禁系统作为对比,采用同样的方法进行测试,对比结果如表3所示。

表3 算法应用结果对比

算法	正确开锁率		平均 开锁率
	男性话者	女性话者	
GMM-UBM 算法	85.83	87.50	86.65
GMM-SVM 算法	89.17	88.75	88.96
HMM-UBM 算法	91.08	90.13	90.61
HMM-UBM-RVM 算法	96.33	93.75	95.04

6 结束语

声纹密码识别广泛应用于各种场所的安全认证,而常用语音识别系统大多采用传统的GMM-UBM模型及改进算法。GMM模型虽能较好反映说

话人的声道信息,但忽视语音内容对识别正确率的影响,不适用于固定文本的声纹密码识别。本文提出HMM-UBM联合RVM的声纹密码识别算法,利用隐马尔科夫模型的时序建模能力得到语音的文本信息,采用UBM模型解决训练样本不足的问题,并通过相关向量机对测试语音进行分类决策。实验结果表明,该算法在进行文本相关的说话人识别时,识别效果优于GMM-UBM算法和GMM-SVM算法,具有较好的应用价值。

参考文献

- [1] ALI Z, ALSULAIMAN M, ELAMVAZUTHI I, et al. Voice pathology detection based on the modified voice contour and SVM[J]. Biologically Inspired Cognitive Architectures, 2016, 15: 10-18.
- [2] ZHANG J, CHEN X M. A research of improved algorithm for GMM voiceprint recognition model[C]// Proceedings of Control and Decision Conference. Washington D. C., USA: IEEE Press, 2016: 5560-5564.
- [3] KINNUNEN T, LI H. An overview of text-independent speaker recognition: from features to super vectors[J]. Speech Communication, 2010, 52(1): 12-40.
- [4] 胡青, 刘本永. 基于卷积神经网络的说话人识别算法[J]. 计算机应用, 2016, 36(增刊): 79-81.
- [5] 陈黎, 徐东平. 基于SVM-GMM的开集说话人识别方法[J]. 计算机工程, 2011, 37(14): 172-174.
- [6] 展领, 景新幸. 基于VQ-MAP和SVM融合的说话人识别系统[J]. 计算机工程与应用, 2011, 47(13): 136-138.
- [7] 李宁, 徐守坤, 马正华, 等. 自适应语音识别算法仿真研究[J]. 计算机仿真, 2011, 28(8): 181-185.
- [8] 于娴, 贺松, 彭亚雄, 等. 基于GMM模型的声纹识别模式匹配研究[J]. 通信技术, 2015, 48(1): 97-101.
- [9] SARKAR A K, TAN Z H. Text dependent speaker verification using un-supervised HMM-UBM and temporal GMM-UBM[C]// Proceedings of Interspeech. San Francisco, USA: [s. n.], 2016.
- [10] CHEN Y, HONG Q. Voiceprint verification based on two-level decision HMM-UBM[C]// Proceedings of the 1st IEEE International Conference on Information Science and Engineering. Washington D. C., USA: IEEE Press, 2009: 3556-3559.
- [11] 卓著, 李辉. PCA变换下的GMM-SVM话者确认研究[J]. 小型微型计算机系统, 2015, 36(3): 637-640.
- [12] LOPEZ-OTERO P, DOCIO-FERNANDEZ L, GARCIA-MATEO C. A fisher-voice-SVM language identification system[M]. Berlin, Germany: Springer, 2012: 381-391.
- [13] 白静, 张雪英. Gaussian核SVM在抗噪语音识别中的应用[J]. 计算机工程与设计, 2009, 30(17): 4061-4063.
- [14] 杨敏, 赖惠成, 董九玲. 基于改进HMM-RVM混合模型的人脸识别方法研究[J]. 激光杂志, 2015, 36(11): 44-47.
- [15] 郑泽萍, 王万良, 郑建伟. 基于保局部核RVM的说话人识别方法[J]. 计算机工程, 2011, 37(14): 208-210.

编辑 赵辉