

短文

基于GMM-UBM和GLDS-SVM的英文发音错误检测方法

李宏言, 黄申, 王士进, 梁家恩, 徐波

1. 中国科学院自动化研究所数字内容技术研究中心 北京 100190

2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190

收稿日期 2009-3-19 修回日期 2009-10-21 网络版发布日期 接受日期

摘要

将语种和说话人识别的方法应用到英语发音错误检测系统, 提出一种基于广义线性区分序列支持向量机 (Generalized linear discriminant sequence based SVM, GLDS-SVM) 的发音错误检测方法. 主要创新点为: 1) 提出一种基于状态拼接的特征规整方案, 增强SVM对发音特征的建模能力; 2) 提出一种基于多模型融合的模型训练策略, 该策略可以更加充分地利用训练数据, 并在一定程度上解决了由于真实发音错误数据缺乏造成的正负样本不均衡的问题; 3) 将GLDS-SVM与基于通用背景模型GMM (Universal background models based GMM, GMM-UBM) 的方法进行融合, 以进一步提高发音检错性能. GLDS-SVM和GMM-UBM的融合系统在仿真测试集和真实测试集上的等错误率 (Equal error rate, EER) 分别达到9.92%和16.35%. 同时, GLDS-SVM在模型占用空间和运算速度方面均比传统径向基函数 (Radial basic function, RBF) 核方法具有明显优势.

关键词 [计算机辅助语言学习](#) [自动发音错误检测](#) [支持向量机特征规整](#) [多模型融合策略](#)

分类号

Automatic Mispronunciation Detection for English Learners by GMM-UBM and GLDS-SVM Methods

LI Hong-Yan, HUANG Shen, WANG Shi-Jin, LIANG Jia-En, XU Bo

1. Digital Content Technology Research Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

Abstract

The paper proposes an efficient generalized linear discriminant sequence based SVM (GLDS-SVM) based mispronunciation detection method. Firstly, in order to enhance the ability of describing pronunciation characteristics, we introduce an improved SVM feature normalization scheme based on state-concatenated operation. Then, we propose a novel multi-model strategy for model training to make full use of samples and solve the problem of data unbalance caused by lack of the actual mispronunciation corpus. Finally, we combine GLDS-SVM with universal background models based GMM (GMM-UBM) to further improve the performance. The fused system by these two methods achieves 9.92% and 16.35% in equal error rate (EER) for simulation set and real set, respectively. Meanwhile, GLDS-SVM processes a higher computation speed and smaller model size than traditional radial basic function (RBF) kernel.

Key words [Computer assisted language learning \(CALL\)](#) [automatic mispronunciation detection](#) [support vector machine \(SVM\) feature normalization](#) [multi-model fusion strategy](#)

DOI: 10.3724/SP.J.1004.2010.00332

通讯作者 李宏言 hyli@hitic.ia.ac.cn

作者个人主页 李宏言; 黄申; 王士进; 梁家恩; 徐波

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF \(2432KB\)](#)

▶ [\[HTML全文\] \(OKB\)](#)

▶ [参考文献 \[PDF\]](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

相关信息

▶ [本刊中 包含“计算机辅助语言学习” 的相关文章](#)

▶ 本文作者相关文章

· [李宏言](#)

· [黄申](#)

· [王士进](#)

· [梁家恩](#)

· [徐波](#)