自动化学报 2009 35 (9): 1193-1198 ISSN: 0254-4156 CN: 11-2109/TP

论文与报告

说话人识别中的因子分析以及空间拼接

郭武, 李轶杰, 戴礼荣, 王仁华

1. 中国科学技术大学多媒体计算与通信教育部-微软重点实验室 合肥 230027

收稿日期 2008-6-11 修回日期 2009-1-3 网络版发布日期 接受日期

摘要
联合因子分析可以有效拟合混合高斯模型中的说话人和信道差异, 在说话人识别中得到广泛应用. 一般情况下, 该算法在对说话人和信道两个载荷矩阵进行联合估计时, 说话人残差矩阵无法发挥作用, 信道载荷矩阵的因子数不能提高. 本文提出说话人载荷矩阵、说话人残差载荷矩阵采用串行的训练模式, 在信道载荷矩阵训练中采用矩阵拼接的方法, 能够有效提高识别率; 在NIST SRE 2008年核心测试数据库的五个部分分别达到等错误率3.3%, 5.1%, 5.0%, 5.3%和5.0%.

关键词 说话人识别 联合因子分析 本征音因子 说话人确认 期望最大化

分类号 TN912.34

Factor Analysis and Space Assembling in Speaker Recognition

GUO Wu, LI Yi-Jie, DAI Li-Rong, WANG Ren-Hua

1. Ministry of Education-Microsoft Key Laboratory of Multimedia Computing and Communication, University of Science and Technology of China, Hefei 230027

Abstract
Factor analysis is a model of the speaker and session variability in Gaussian mixture models and is widely used in text-independent speaker recognition. There exist two issues when the loading matrices of the eigenvoice and eigenchannel are estimated jointly. First, the speaker diagonal matrix (residual) will not take effect; second, the channel factors can not be very large. In this paper, the loading matrices of eigenvoice and the diagonal are calculated serially and different eigenchannel matrices are assembled to form a large channel loading matrix. The performance can be improved by the proposed algorithm. In the NIST speaker recognition evaluation (SRE) 2008 core test corpus, the equal error rates (EERs) of the five sub sessions were 3.3%, 5.1%, 5.0%, 5.3%, and 5.0%.

Key words Speaker recognition joint factor analysis eigenvoice speaker verification expectation maximization

DOI: 10.3724/SP.J.1004.2009.01193

通讯作者 郭武 guowu@mail.ustc.edu.cn

作者个人主页 郭武; 李轶杰; 戴礼荣; 王仁华