

doi:10.1631/FITEE.1500380

题目: 采用背景人声简化特征集的说话人识别直方图均衡化方法

概要: 本文提出了一种用于说话人识别技术的直方图均衡化方法。该方法采用了一套增补简化特征集,用以在训练数据和测试数据较短时改进说话人识别的效果。该增补特征集采用选择算法或聚类算法从背景人声中派生得到。当输入语音数据样本不足时,本文提出的方法可作为构建直方图的特征归一化方法使用。另外,该方法作为一种 *i-vector* 归一化方法,源于一种目前较为先进的基于 *i-vector* 的概率线性判别分析 (Probabilistic linear discriminant analysis, PLDA) 说话人识别系统。在输入语音和增补集中,用于直方图均衡化的样本值序号均按升序进行估计。新的序列号则按不同种类的序号之和进行排列。随后,该方法采用最新的序列号得出了测试语音样本的累积分布函数。本文将这一方法与倒谱均值归一化 (Cepstral mean normalization, CMN) 方法、倒谱均值和方差归一化 (Cepstral mean and variance normalization, MVN) 方法、直方图均衡化 (Histogram equalization, HEQ) 方法和欧洲电信标准协会模拟前端方法进行了比较。此外,在一具体算例中将该方法性能与采用模糊 *C-means* 和 *K-means* 算法的贪婪选择算法进行了比较。采用 YOHO 和 ETRI 数据库对特征空间进行评估。测试集采用 Opus VoIP 编码器进行了模拟。本文还采用了 2008 美国国家标准技术研究所说话人识别评测语料库对该 *i-vector* 系统进行了评测。试验结果表明,与传统特征归一化方法相比,当采用所提出的方法时,平均系统性能可得到有效提升。

关键词: 说话人识别; 直方图均衡化; *i-vector*