

短文与研究通讯

多分形谱簇研究及其在说话人识别中的应用

周宇欢, 张雄伟, 付强, 徐鑫, 王金明

解放军理工大学指挥自动化学院

摘要:

语音是一种复杂的非线性信号, 这使得基于线性系统理论发展起来的传统说话人识别技术性能难以进一步提高。本文提出了多分形谱簇分析方法, 用于分析语音信号的非线性特征, 并应用于短语音(2秒)说话人识别。通过对Cantor集的仿真实验, 发现不同标度区能反映出系统不同阶段的生长规律, 因此可用一组连续变化的多分形谱分层次地表征系统的分形特性, 即多分形谱簇分析方法。然后结合语信号的分形特点, 提出一种语音的多分形谱簇特征(Multifractal Spectrum Cluster Feature, MSCF)的提取方法。最后将几种非线性特征与短时谱特征结合用于说话人识别, 基于TIMIT数据库50人的实验表明, 非线性特征与短时谱特征互补性较强, 特别是MSCF与MFCC、LPC特征结合, 使得系统的误识率下降到0.8%。

关键词: 说话人识别 多分形谱簇 标度区 高斯混合模型

Research on Multifractal Spectrum Cluster and Its Application in Speaker Recognition

ZHOU Yu-Huan, ZHANG Xiong-Wei, FU Qiang, XU Xin, WANG Jin-Ming

Institute of Command Automation, PLAUST

Abstract:

Speech is a complicated nonlinear signal, so traditional speaker recognition technology based on the linear theory is difficult to be further improved. Hence, the multifractal spectrum cluster analytical method is proposed, and applied to the analysis of nonlinear characteristic of the speech signal in speaker recognition of short speech. Through extensive experiments for Cantor sets, it is found that sub-scaling ranges, which were neglected by traditional multifractal method, actually reflected the growth pattern in different growth stages. Therefore, in order to fully consider the fractal characteristics contained in different scaling range, the multifractal spectrum cluster analytical method is proposed to describe the multi-level fractal characteristics accurately and comprehensively. Then, according to the characteristic of the speech signal, an extraction method of speaker multifractal spectrum cluster feature (MSCF) is proposed, which could combine with short-term spectral feature in feature layer effectively. Finally, the combinations of several nonlinear features and short-term spectral feature are applied to speaker recognition. Experiment results based on the TIMIT show that nonlinear feature and short-term spectral feature are highly complementary, which make the error rate of speaker recognition system decrease obviously, especially the combination of MSCF, MFCC and LPC can reduce the error rate to 0.8% in short speech speaker recognition.

Keywords: speaker recognition multifractal spectrum cluster scaling range; Gaussian mixture model

收稿日期 2011-01-25 修回日期 2011-07-26 网络版发布日期 2011-12-25

DOI:

基金项目:

2009江苏省自然科学基金资助(BK2009059)

通讯作者:

作者简介:

作者Email: zhouyuhuan250@sina.com

参考文献:

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(888KB)
- ▶ [HTML全文]
- ▶ 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
- ▶ 引用本文
- ▶ Email Alert
- ▶ 文章反馈
- ▶ 浏览反馈信息

本文关键词相关文章

- ▶ 说话人识别
- ▶ 多分形谱簇
- ▶ 标度区
- ▶ 高斯混合模型

本文作者相关文章

- ▶ 周宇欢
- ▶ 张雄伟
- ▶ 付强
- ▶ 徐鑫
- ▶ 王金明

PubMed

- ▶ Article by Zhou, Y. H.
- ▶ Article by Zhang, X. W.
- ▶ Article by Fu, Q.
- ▶ Article by Xu, X.
- ▶ Article by Wang, J. M.

1. 薛峰, 俞一彪. 倒谱域特征分量置信度分析及说话人识别应用[J]. 信号处理, 2010,26(1): 127-131
2. 叶蕾, 郭海燕, 杨震. 基于压缩感知重构信号的说话人识别系统抗噪方法研究[J]. 信号处理, 2010,26(3): 321-326

文章评论

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text"/> 7742