数据库、信号与信息处理

汉语语音识别中基频特征的直接声学建模方法

黄 浩[1]，哈力旦[2]

1.新疆大学 信息科学与工程学院，乌鲁木齐 830046
2.新疆大学 电气工程学院，乌鲁木齐 830046

摘要　　提出了隐条件随机场对断续基音频率序列进行直接声学建模的方法，该方法针对汉语语音中基频值在清音段连续，浊音段断续的特点，根据隐条件随机场区别于隐马尔可夫模型的重要特性——无需对观察值采用统一的建模方式，直接对不连续基频值与连续谱特征观察值一起进行声学建模。大词汇语音库上的汉语带调音节分类实验表明，隐条件随机场下对断续基音频率序列的直接建模较使用清音段人工平滑基频特征的识别率有明显的提高，还给出了与不同区分性准则训练的隐马尔可夫声学模型的实验性能的比较。
关键词　　隐条件随机场　汉语语音识别　声学模型
分类号　TN912.34

Direct F0 incorporation for acoustic modeling in Mandarin speech recognition

HUANG Hao[1]，Halidan[2]

1.Department of Information Science and Engineering，Xinjiang University，Urumqi 830046，China
2.Department of Electrical Engineering，Xinjiang University，Urumqi 830046，China

Abstract

 Hidden Conditional Random Fields（HCRFs） based acoustic modeling is proposed by directly using discontinuous fundamental frequency（F0） sequences for Mandarin speech recognition.The method is based on the fact that F0 observations are continuous in voiced portion in Mandarin speech and missing in unvoiced portion，and HCRFs are more suitable for integrating such non-uniform features.Tonal syllable classification tasks are carried out on continuous speech database.Results show HCRFs trained on discontinuous F0 are significantly better than those trained on smooth F0 sequences from artificial interpolation.Comparisons with hidden Markov models under various training criteria are also given.

Key words　hidden conditional random fields　Mandarin speech recognition　acoustic modeling

通讯作者　黄 浩 hwanghao@gmail.com

扩展功能

本文信息
- Supporting info
- PDF(725KB)
- [HTML全文](0KB)
- 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
- 文章反馈
- 浏览反馈信息

相关信息
- 本刊中 包含"隐条件随机场"的相关文章
- 本文作者相关文章
- · 黄 浩
- · 哈力旦