

智能感知与识别处理

基于动态贝叶斯网络的音视频连续语音识别和音素切分

吕国云¹;蒋东梅²;赵荣椿²;蒋晓悦²;侯云舒²;孙阿利²;H. Sahli³;W. Verhelst³

西北工业大学, 西安傅立叶电子科技¹

西北工业大学²

VUB³

收稿日期 2007-1-24 修回日期 网络版发布日期 2007-7-6 接受日期

摘要 构造了两个单流单音素的动态贝叶斯网络 (DBN) 模型, 以实现基于音频和视频特征的连续语音识别, 并在描述词和对应音素具体关系的基础上, 实现对音素的时间切分。实验结果表明, 在基于音频特征的识别率方面: 在低信噪比 (0~15dB) 时, DBN模型的识别率比HMM模型平均高12.79%; 而纯净语音下, 基于DBN模型的音素时间切分结果和三音素HMM模型的切分结果很接近。对基于视频特征的语音识别, DBN模型的识别率比HMM识别率高2.47%。实验最后还分析了音视频数据音素时间切分的异步关系, 为基于多流DBN模型的音视频连续语音识别和确定音频和视频的异步关系奠定了基础。

关键词 [动态贝叶斯网络](#) [音视频](#) [语音识别](#) [音素切分](#)

分类号

DOI:

对应的英文版文章: [A7010482](#)

通讯作者:

吕国云 tech@fftchina.com, lvguoyun101@sohu.com

作者个人主页: 吕国云 蒋东梅 赵荣椿 蒋晓悦 侯云舒 孙阿利 H. Sahli W. Verhelst

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF](#) (907KB)

▶ [\[HTML全文\]](#) (0KB)

▶ [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](#)

相关信息

▶ [本刊中 包含“动态贝叶斯网络”的相关文章](#)

▶ 本文作者相关文章

· [吕国云](#)

· [蒋东梅](#)

· [赵荣椿](#)

· [蒋晓悦](#)

· [侯云舒](#)

· [孙阿利](#)

· [H Sahli](#)

· [W Verhelst](#)