研究简报

基于汉语视频三音素的可视语音合成

赵　晖, 唐朝京

国防科技大学电子科学与工程学院 长沙 410073

摘要
为了合成具有真实感的视频序列，该文提出一种基于汉语视频三音素的可视语音合成方法。根据汉语的发音规律和音素与视素的对应关系，该文提出"视频三音素"的概念。在此基础上，建立隐马尔可夫(HMM)训练与合成模型，在训练过程中使用了视频音频联合特征，并加入了动态特征。在合成过程中，连接视频三音素HMM模型形成句子HMM，并从中提取特征参数，合成可视语音。从主观和客观评估结果来看，合成视频的真实感强，满意度较高。

关键词　[可视语音合成](#)　[视频三音素](#)　[隐马尔可夫模型](#)　[联合特征](#)

分类号　[TP391.42](#)

Visual Speech Synthesis Algorithm Based on Chinese Visual Triphone

Zhao Hui, Tang Chao-jing

College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China

Abstract
In order to synthesize real video sequence, a visual speech synthesis algorithm based on Chinese visual triphone is proposed. According to Chinese pronunciation principle and the relationship between phoneme and viseme, conception of 'visual triphone' is presented. Hidden Markov Model(HMM) is established based on visual triphones. In the training stage, combined features including visual features and audio features are used. In the synthesis stage, sentence HMM is constructed by concatenating triphone HMMs, from which the feature parameters are extracted. From the result of subjective and objective evaluation, the synthesized video is real and satisfied.

Key words　[Visual speech synthesis](#)　[Visual triphone](#)　[Hidden Markov Model(HMM)](#)　[Combined features](#)

DOI :

通讯作者

作者个人主页　　赵　晖; 唐朝京

扩展功能

本文信息

▸ Supporting info
▸ PDF(341KB)
▸ [HTML全文](OKB)
▸ 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
▸ 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"可视语音合成"的相关文章

▸本文作者相关文章
· 赵　晖
· 唐朝京