



学科导航4.0暨统一检索解决方案研讨会

基于机器学习的语音驱动人脸动画方法

<http://www.fristlight.cn> 2007-06-25

[作者] 陈益强;高文;王兆其;姜大龙

[单位] 中国科学院计算技术研究所, 哈尔滨工业大学计算机科学与工程系

[摘要] 语音与唇动面部表情的同步是人脸动画的难点之一。综合利用聚类和机器学习的方法学习语音信号和唇动面部表情之间的同步关系, 并应用于基于MPEG-4标准的语音驱动人脸动画系统中。在大规模音视频同步数据库的基础上, 利用无监督聚类发现了能有效表征人脸运动的基本模式, 采用神经网络学习训练, 实现了从含韵律的语音特征到人脸运动基本模式的直接映射, 不仅回避了语音识别鲁棒性不高的缺陷, 同时学习的结果还可以直接驱动人脸网格。最后给出对语音驱动人脸动画系统定量和定性的两种分析评价方法。实验结果表明, 基于机器学习的语音驱动人脸动画不仅能有效地解决音视频同步的难题, 增强动画的真实感和逼真性, 同时基于MPEG-4的学习结果独立于人脸模型, 还可用来驱动各种不同的人脸模型, 包括真实视频、2D卡通人物以及3维虚拟人脸。

[关键词] 机器学习;人脸动画;语音驱动

语音与唇动面部表情的同步是人脸动画的难点之一。综合利用聚类和机器学习的方法学习语音信号和唇动面部表情之间的同步关系, 并应用于基于MPEG-4标准的语音驱动人脸动画系统中。在大规模音视频同步数据库的基础上, 利用无监督聚类发现了能有效表征人脸运动的基本模式, 采用神经网络学习训练, 实现了从含韵律的语音特征到人脸运动基本模式的直接映射, 不仅回避了语音识别鲁棒性不高的缺陷, 同时学习的结果还可以直接驱动人脸网格。最后给出对语音驱动人脸动画系统定量和定性的两种分析评价方法。实验结果表明, 基于机器学习的语音驱动人脸动画不仅能有效地解决音视频同步的难题, 增强动画的真实感和逼真性, 同时基于MPEG-4的学习结果独立于人脸模型, 还可用来驱动各种不同的人脸模型, 包括真实视频、2D卡通人物以及3维虚拟人脸。

[存档附件1](#)

[我要入编](#) | [本站介绍](#) | [网站地图](#) | [京ICP证030426号](#) | [公司介绍](#) | [联系方式](#) | [我要投稿](#)

北京雷速科技有限公司 Copyright © 2003-2008 Email: leisun@fristlight.cn

