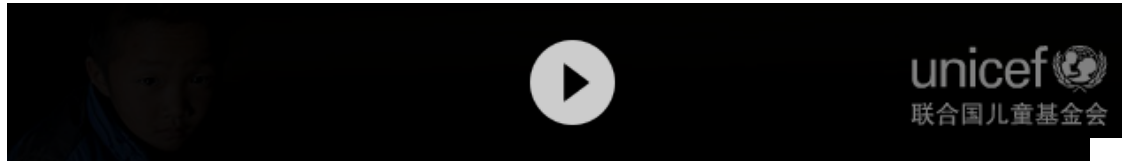


[新浪首页](#)[新闻](#)[体育](#)[财经](#)[娱乐](#)[科技](#)[博客](#)[图片](#)[专栏](#)[更多](#)[移动客户端](#)[新浪科技](#) > [业界](#) > [正文](#)[新闻](#)[快乐男声](#)

迪士尼也来研究人工智能 将AI用于动画制作

2017年08月18日 18:46 [雷锋网](#)

5



雷锋网 AI科技评论按：众所周知，卡耐基梅隆大学在计算机科学方面的研究名列前茅，而迪士尼有意将计算机科学技术引入动画制作。他们与卡耐基梅隆大学合作建立的实验室近日发表了一篇论文 A Deep Learning Approach for Generalized Speech Animation，利用深度学习的方法，来生成看起来自然的语音动画。这篇论文已被SIGGRAPH 2017收录。

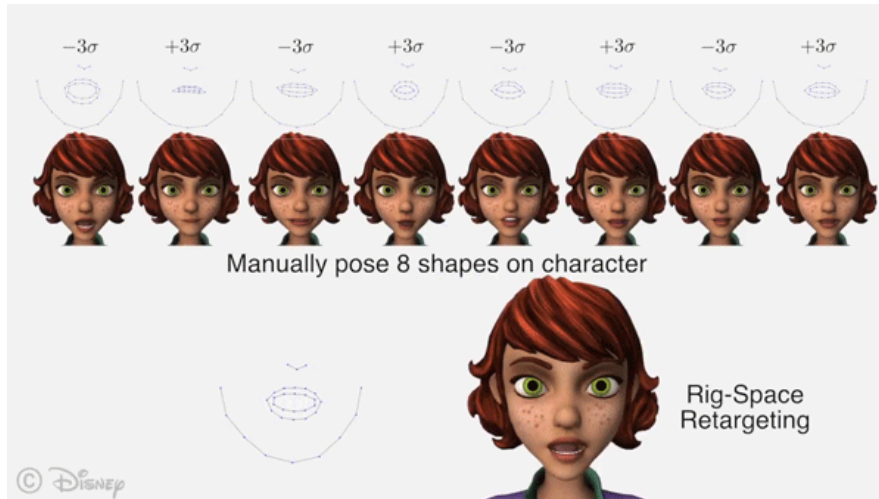
他们引入了一种简单而有效的深度学习方法，来自动生成看起来自然的，能够与输入语音同步的语音动画。这种方法使用滑动窗口预测器，可以学习到从音位标签输入序列到嘴型运动的任意非线性映射，能精准捕捉自然动作和可视化的协同发音效果。

这种方法有几个吸引人的特性：它能实时运行，只需要进行非常少的参数调节，能很好的泛化到新的输入语音序列，很容易编辑来创建风格化和情绪化的语音，并且与现有的动画重定向方法兼容。

[创事记](#)[+ 更多](#)[联发科迎来“最艰难时刻”](#)

或许谁都不希望看到这样一个景象，不久的将来搭载 Helio X30的某某... [详细]

迪士尼实验室表示，他们工作中的一个重点是开发出能高效生成语音动画，并将其轻松地整合到现有作品中的方法。他们的论文中详述了这种端到端的方法，其中包括机器学习的一些设计决策。在论文中，通过动画片段中不同的人物和声音，演示了泛化的语音动画结果，包括唱歌和外语输入。这种方法还可以根据用户的语音输入实时生成灵活的语音动画。



雷锋网 AI科技评论将论文部分内容编译如下：

前言

语音动画是生成逼真的角色动画中重要且耗时的一部分。从广义上讲，语音动画是一种这样的任务：改变图形(或机器人)模型的面部特征，使嘴唇的动作与发出的声音同步，形成一种在说话的感觉。作为人类，我们都是面部表情的专家，糟糕的语音动画可能会让人分心，不愉快，产生困惑。例如，当看到的嘴型和听到的声音不一致时，有时会让观众以为自己听到的是另一种声音（McGurk和MacDonald的论文，1976）。对于实际的角色动画来说，高保真语音动画至关重要。

目前在电影和 videogame 制作中使用的传统语音动画方法通常趋向于两个极端。一种做法是，高预算的产品通常会采用表演捕获技术或雇一个大型的专业动画制作团队，这样花费巨大，而且很难大规模复制。例

高端智能电饭煲

——微压快煮 香甜Q弹——

4L大容量 货到付款



高端智能电饭煲

——微压快煮 香甜Q弹——

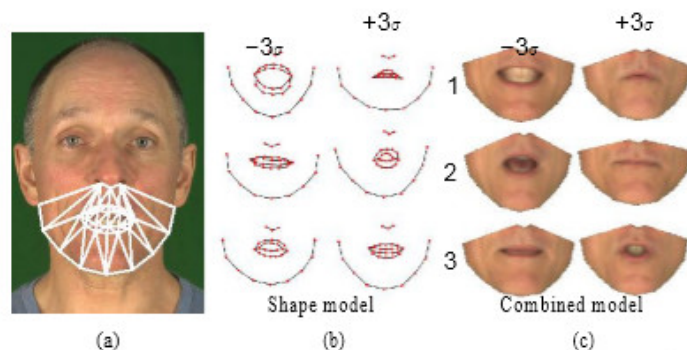
4L大容量 货到付款



如，目前没有什么好的生产方法，可以跨多种语言，划算且高效地生成高质量的语音动画。另一种做法是，对于成本低、内容多的产品，可能会使用简单的唇形库来快速生成质量相对较低的语音动画。

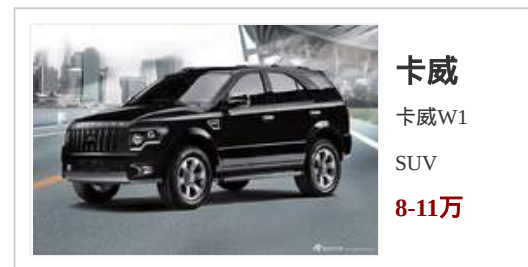
最近，人们对开发出自动生成语音动画的数据驱动方法越来越感兴趣，以找到将这两个极端折中的解决办法（De Martino等的论文，2006；Edwards等的论文，2016；Taylor等的论文，2012）。但是，以前的工作需要预先定义一组数量有限的唇形，还必须将这些唇形混合起来。简单的混合函数限制了可以建模的视觉语音动态的复杂度。所以我们另辟蹊径，计划利用现代机器学习方法，直接从数据中学习视觉语音的复杂动态。

我们提出了一种自动生成语音动画的深度学习方法，这种方法提供一种划算且高效的手段，能大规模地生成高保真的语音动画。例如，我们用100多个自由度，在电影特效制作级别的人脸模型上生成逼真的语音动画。我们工作中的一个重点是开发一种高效的语音动画方法，可以无缝地整合到现有的作品生产中。



我们的方法使用连续的深度学习滑动窗口预测器，这是受Kim等人在2015年发表的一篇论文的启发。滑动窗口的方法意味着预测器能够在持续讲话的输入语音描述和输出视频之间表示复杂的非线性回归，也自然包括语境和协同发音效果。我们的研究结果展现了在Kim等人之前的决策树方法上利用神经网络深度学习方法带来的改进。

使用重叠的滑动窗口更直接地将学习集中在捕捉局部范围的语境和协同发音的效果上，比起循环神经网络和LSTM（Hochreiter和Schmidhuber的论文，1997）等传统的序列学习方法，更适合预测语音动画。



科技前沿

+ 更多



大脑功能的副产品：我们为何会做梦？

使用机器学习的主要挑战之一是：要以一种对所需的最终目标有用的方式，恰当地定义学习任务(例如选择什么样的输入/输出和训练集)。我们的目标是让动画师能轻松地将高保真的语音动画合并到任何rig上，对任何说话者都适用，并且易于编辑和风格化。

我们将我们的机器学习任务定义为，从单个作为参照的说话者中，学会产生具有中性语音的高保真动画。通过聚焦作为参照的面部和中性的语音，我们可以低成本且高效地收集一个全面的数据集，这个数据集能充分地描述出语音动画的复杂特性。大的训练数据集使得我们能够使用现代机器学习方法，可靠地学习语音运动中细微的动态变化。

与之前程序化的生成语音动画的研究相比（De Martino等的论文，2006；Edwards 等的论文，2016；Taylo等的论文，2012），我们的方法能直接从数据中学会自然的协同发音效果。

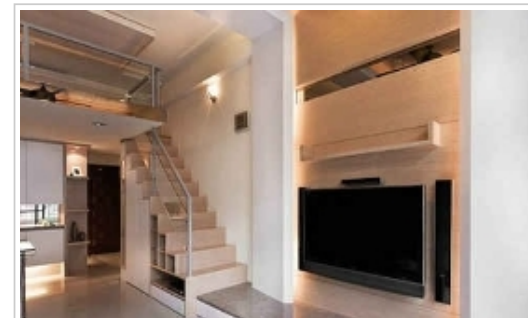
我们将输入定义为文本(音位标签)，意味着可以学习与说话者无关的从语境到语音动画的映射。

我们只需要现成的语音识别软件自动将任何说话者的语音转换成相应的音位描述。因此，我们的自动语音动画可以泛化到任何说话者，任何形式的语音，甚至是其他语言。

局限性和未来的研究

主要的实际局限是，我们的动画预测是依据AAM参数化法生成的参考面部来制作的。这使我们这种方法能泛化到任何内容，但是对特征进行重定位会引入潜在的错误源。当提出重定向模型的初始特征设置时，必须小心谨慎，以保持预测动画的逼真度。幸运的是，对每个角色，这个预计算步骤只需执行一次。展望未来，一个有意思的研究方向是使用真实的动画数据来开发针对自动语音动画的数据驱动重定位技术。

只从中性的语音中学习，我们可以得到一个具有鲁棒性的语音动画模型，它可以泛化到任何语音内容。目前，在动画中添加表情和情感还是艺术家的工作，在未来，一个有趣的方向是从许多具有情感的语境（生气、伤心等）生成的训练数据中，训练一个更大的神经网络，使预测的面部动作更接近于真实的情感。



小户型复式房



俄罗斯旅游



JD.COM 京东



¥ 418.00

2/12

一个主要的挑战是如何既划算，又高效地收集一个综合数据库用于训练。如果没有一个够全面的训练集，使用现代机器学习技术会存在困难，因为深度学习等方法通常是严重欠约束的。可能的方向是大规模地收集杂乱的数据(例如从公共视频存储库中收集)，或者开发能自适应地选择收集哪种视频的主动学习方法，以使总收集成本最小化。

进一步的泛化性可以从具有多种面部特征(男性、女性、圆脸、方脸、肥胖、消瘦等)的多个讲话者中训练一个语音动画模型，并在预测的时候选择与动画角色模型最匹配的特征。这种方法

可以根据人物的说话风格，泛化到不同脸型的不同面部表情。再一次说明，如何高效地收集综合训练集是一个很大的挑战。

雷锋网 AI科技评论编译

论文地址：http://www.yisongyue.com/publications/siggraph2017_speech.pdf

标签： 人工智能 动画 语音

苹果汇

+ 更多



创意还是败笔？揭秘你不知道的苹果产品

5 条评论 | 5 人参与



微笑的夏草冬虫
退出

请输入评论

分享到微博

发布

最新评论



小朱_zwb [广东惠州]

让AI制作动画哆啦A梦吃惊相当期待呢！

8月18日22:29 | 举报

赞 | 回复



宇宙1530 [山东淄博]

迪士尼也来研究人工智能 将AI用于动画制作

8月18日21:51 | 举报

赞 | 回复



耗子和兔子的小窝 [江苏南京]

【迪士尼也来研究人工智能 将AI用于动画制作】卡耐基梅隆大学在计算机科学方面的研究名列前茅，而迪士尼有意将计算机科学技术引入动画制作。近日，双方合作建立的实验室发表了一篇论文，旨在利用深度学习的方法，来生成看起来自然的语音动画：

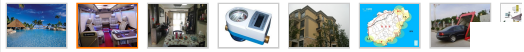
8月18日21:16 | 举报

赞 | 回复

还有2条评论，更多精彩评论>>



柯斯达改装



【券后¥289】李宁气垫跑步鞋



新浪科技公众号

“掌”握科技鲜闻 (微信搜索Techsina或扫描左侧二维码关注)

这锅真的太值了！

德国老牌 顺丰包邮



2折



吉利熊猫

5万以下



荣耀畅玩5C 标准...

999



柔白紧致美体乳

160

新浪众测·免费送

更多好物在等你



明基爱眼显示器EW2775ZH

价格：¥1399 数量：10台

立即申请



台电T10平板电脑

价格：¥1399 数量：10台

立即申请



推荐阅读

迪士尼用AI捕捉观众观影情绪 脑洞大开？

迪士尼用AI捕捉观众观影情绪 脑洞大开？[详细]

7月27日 22:11 迪士尼用AI捕捉观众观影情绪 脑洞大开？

评论 | 分享

外媒：百度DuerOS领跑中国语音交互市场

中新网7月31日电 在2017百度AI开发者大会宣布全面开放赋能后，对话式人工智能系统DuerOS不仅相继落地了美的智能语音盒、中信国安广视的G1智能高清机顶盒和极米无屏电视...[详细]

7月31日 15:43 百度 人工智能 语音

评论 | 分享

搜狗地图智能副驾发布 开拓人工智能落地新方向

中新网7月18日电 当下人工智能无疑是最受追捧的概念之一，几乎所有的顶尖互联网企业都在布局人工智能，但目前来看能够真正落地的应用却乏善可陈。[详细]

7月18日 15:48 搜狗 人工智能 语音 评论 | 分享

哈佛用NBA比赛数据生成报道，评测各模型效果 | 数据集+论文+代码

哈佛用NBA比赛数据生成报道，评测各模型效果 | 数据集+论文+代码[详细]

7月28日 13:30 哈佛用NBA比赛数据生成报道，评测各模型效果 | 数据集+论文+代码 评论 | 分享

谷歌新论文：让机器人依靠视觉识别抓取特定物体

谷歌新论文：让机器人依靠视觉识别抓取特定物体[详细]

7月10日 13:30 谷歌新论文：让机器人依靠视觉识别抓取特定物体 评论 | 分享



¥68.00



¥49.00



¥49.80

客厅灯

冲锋衣

床垫

沙发垫

男内裤

淘广告

 <p>韩式婚纱摄影</p>	<p>40岁女人征婚</p> <hr/> <p>找女生</p> <hr/> <p>团购婚纱照</p>	 <p>俄罗斯八日游</p>	<p>学习动画制作</p> <hr/> <p>智能传感器</p> <hr/> <p>韩式婚纱照</p>
--	--	---	---

 <p>婚纱摄影前十名</p>	 <p>丽江婚纱摄影</p>
 <p>创意婚纱照</p>	 <p>学习动画制作</p>

18/29岁潮男请进！
真的很帅 GO >



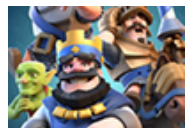
应用中心

新浪公益

新浪游戏

新浪视频

新浪科技



▶ 带领皇室家族赢得胜利
靠反应力还不如背赛道
动脑筋喂小怪物吃饼干



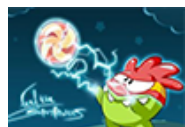
▶ 轻松延长手机续航时间
练就梦寐以求的好身材
帮助保持健康生活方式



▶ 随心所欲打造完美效果
银河系最全新番齐放送
复杂世界里一个就够了



▶ 进入机器梦境寻找真相
疯狂逃离恐怖的压碎机
冰锋暗影古色庙殿狂奔



▶ 斗巫师夺回被偷的糖果
屯兵练兵出征攻占堡垒
伪装胖猪混敌营救伙伴



▶ 国内最专业的手机地图
帮你一起保护你的颈椎
随时随地轻松订火车票

[新浪简介](#) | [About Sina](#) | [广告服务](#) | [联系我们](#) | [招聘信息](#) | [网站律师](#) | [SINA English](#) | [通行证注册](#) | [产品答疑](#)

Copyright © 1996-2017 SINA Corporation, All Rights Reserved

新浪公司 版权所有