

# 文本如何转换成动画? 迪士尼通过AI自动生成动画

■ 映 寒

迪士尼研究所和罗格斯大学的科学家共同发表了关于AI文本生成动画模型的论文。这种算法只要在输入的文本中描述某些活动即可,不需要注释数据和进行大量训练就能产生动画。研究人员进一步提出了端到端模型,这种模型可以创建一个粗略的故事版和电影剧本的视频,用来描绘电影剧本中的文字。此外,这个系统还可用于生成训练端到端神经系统的训练数据。

## 仅训练了996个电影剧本

将文本转换为动画并不是一项简单的任务,大多数将文本转换为视频的工具不能处理复杂句子,因为输入的句子和输出的动画都没有固定的结构。为了克服这种问题,研究团队构建了一个包含多个组件模块的神经网络。

这个网络由几个部分组成:一个可以自动将文本与剧本场景描述隔离开的脚本解析模块、一个自然语言处理模块(使用一套语言规则简化复杂句子,并从简化句子中提取信息,转化为预定义的动作表示),以及一个将所

述表示转换为动画序列的生成模型。

研究人员说,简化方法使得提取关键脚本信息变得更加容易,并且为此,他们的系统自主地确定给定的代码片段是否包含特定的句法结构,然后将其拆分并组合成更简单的句子,递归处理它,直到无法进一步简化。

接下来,系统将“协调”句法关系相同、功能相同的句子。最后,词汇简化器将简化后句子中的动作,与预定义库中的52个动画匹配(通过同义词词典扩展到92个)。

然后,在一个名为Cardinal的管道中将动作输入,并在一个流行的视频游戏引擎Unreal中创建预可视化。利用预定义的动画库、预加载的对象以及可用于创建角色的模型,最终,这个系统可以生成一个3D动画视频。

为了训练这个系统,研究人员从IMSDb、SimplyScripts和ScriptORama5等可自由获取资源的电影剧本数据库中,搜集了超过1000个剧本,从中选取了996个,编写了场景描述语料库。这个语料库由525708个描述组成,包含1402864个句子,其中920817个(超过40%)句子中至少有一个动词。

## 帮助作家提高编写效率

迪士尼一直十分关注AI领域的进展,多年来一直尝试将AI技术融入自己的各项产业当中,也曾多次与大学或其他研究机构合作,开发了一系列训练模型。

2018年1月,迪士尼的研究人员发布了一项可以使动画AR角色与实际的物体互动的黑科技。这项研究做到了AR体验与现实的互动,给动画带来了更多乐趣。

2018年9月,迪士尼AI研究中心还尝试了让机器人完成像超级英雄一样的特技动作。迪士尼的特技实验包括训练机器人的神经网络来控制机器人,以此完成上天入地、爬行、划船等动作。

这次,迪士尼又将文本转换视频的技术进一步加强,使模型自动生成动画。

“从自然语言文本自动生成动画可以应用到许多领域中,如电影脚本编写、教学视频和公共安全等,通过实现更快的迭代、原型设计和概念验证,对剧本编写特别有价值。我们开发了一个能够处理复杂句子的文本到动画系统,目的不是要取代作家,而是要使

他们的工作更有效率,并减少乏味的工作。”研究团队介绍道。

## 68%测试者认为动画合理

在一项定性测试中,22名参与者以5分制标准,来评估系统生成的20个动画。例如,如果所显示的视频对文本来说是合理的动画,则视频中描绘了多少文本信息,以及视频中有多少信息存在于文本中。68%的参与者认为系统通过输入剧本生成了“合理”的动画。

研究人员认为,除了系统本身的局限性之外,生成动画的“不合理”成分也与文本中关于行动的模糊性有关。他们承认这个系统并不完美,它的动作和对对象列表并不是详尽无遗的。有时候,词汇简化不能将动词(如“watch”)映射到相似的动画(“look”)中,或者只能为原句中有很多主语的动词创建几个简化的句子。

内部评价和外部评价显示了该系统性能的合理性。研究人员计划在今后的工作中,重点关注如何能使系统充分地利用文本中描述的话语信息,来解决文本中关于行动模糊性的问题。



日前,日本东京政府在东京一处商业设施内展开机器人实用测试活动。本次测试包括机器人引导、检测来客体温、远距离临场等内容。

新华社发  
(东京都政府供图)

## 为传统工厂 装上“AI大脑”

在西服智能工厂里,“人工智能+5G技术”可以实现对拆线、板型检验过程中10多道工序精细化检测识别,并自动计算出产品合格率和员工工时达标率,强化生产过程的可视化、透明化、可预测、自适应能力……这是中国联合网络通信股份有限公司与雅戈尔联手打造的全球最大的“全连接5G+西服智能工厂”。

不仅是雅戈尔,“中国联通工业AI(人工智能)解决方案”已成功落地国内多个工厂。在上海某汽车内饰材料生产车间内,针对面料质检的外观及品控要求,中国联通AI质检解决了复杂的瑕疵成像采集问题,并通过模型迭代成长机制,让系统不断提高瑕疵检测准确率。这一系统能实现对多款服饰面料、十几种瑕疵的检测,缺陷检出率超92%,大大降低了对专业质检人员的依赖。

基于中国联通“AI大脑”,依托在人脸识别、人体识别、物体识别、环境识别、语音识别、自然语言处理等方向自主研发的30余项科技成果,结合智能监控、智能机械臂、智能对话等设备与技术形成的垂直功能,“中国联通工业AI解决方案”有力支撑了工业质检、工业合规等多个场景。

作为中国联通创新能力建设的一部分,联通大数据积极探索运营商大数据技术及应用,不断夯实大数据、AI、区块链三大数字化转型基础,基于多年实践形成自主研发产品体系及服务运营体系,对内全面赋能的同时,面向政务、金融、文旅、公共安全、工业制造等行业打造了一批数据智能应用及服务,赋能组织数字化变革,目前已服务3000家以上企业客户。(本报综合)

## 中国联通发布AI视频彩铃城市名片

■ 秦 天

10月18日,由工业和信息化部、江西省人民政府主办,以“VR让世界更精彩——育新机 开新局”为主题的2020世界VR产业大会云峰会在江西南昌拉开序幕,中国联通视频彩铃携多项技术成果亮相展会,共同见证5G智慧未来。

作为中国联通的5G战略产品之一,5G视频彩铃业务不断创新,颠覆传统音频彩铃单一的业务形式,将短视频引入彩铃设置阶段,并带动用户发挥创意DIY专属视频彩铃,视频彩

铃的AI新形态将为用户打造通话交互新体验。大会现场,中国联通视频彩铃发布了视频彩铃AI城市名片,吸引了众多参观者驻足。

观众与南昌绿地中心双子塔、八一南昌起义纪念馆、南昌之星摩天轮、滕王阁、奥林匹克体育中心、婺源、三清山等江西地标的合拍,实时生成视频,进而一键设置成为自己的视频彩铃,个性化定制的家乡“城市名片”让更多人领略家乡风采,了解家乡文

化。据了解,该产品已成为当前各个城市的重要创新“宣传窗口”。

同时,在联通在线公司展区现场,中国联通视频彩铃还特别定制了江西全省11个地市共计12个城市名片,展示江西形象,诠释其经济底蕴,引领城市数字经济发展新潮流。中国联通视频彩铃打造的AI城市名片既可以满足用户更深层次个性化需求,同时为每一座独具特色的城市贴上标签,用5G赋能城市数字化发展。