

人工智能为 3D 影视行业注入新动力

赵天奇

北京聚力维度科技有限公司,北京 100089

摘要 介绍了 3D 视觉产生的原理和人类对 3D 视觉的探索历程,分析了 3D 内容的制作方式。伴随计算机视觉和人工智能技术发展,提出并研发了一种新的 3D 转制技术——人工智能 2D 转 3D,该技术可根据 2D 图像自动生成 3D 内容,在无损品质下,将 3D 转制效率提升了 1000 余倍。并指出基于这项技术,可开发 3D 直播、3D 浏览器等新应用,也有望彻底解决 3D 内容匮乏问题。

关键词 3D 影视;深度估计;人工智能 2D 转 3D;3D 内容制作

自 2009 年《Avatar》上映之后,中国观众对 3D 电影不再陌生,3D 电影在整个影视行业也占据重要位置。根据国家广电总局公布的数据,2011 年中国大陆电影票房总收入达 131.1 亿元,上映的 3D 电影超过 30 部,其中 14 部票房超过 1 亿元,总票房超过 50 亿元,约占全年票房的 40%^[1]。而 2016 年中国大陆 457 亿元全年电影票房中,3D 电影的票房超过 294 亿元,占比超过 64%^[2]。

3D 电影的商业表现令人惊叹,但 3D 影视的制作技术却原地踏步。传统的 3D 实拍、2D 转 3D 技术需要消耗大量的时间与人力,制作成本高昂,这使 3D 影视作品至今无法像普通影视内容一样类型丰富并广为普及。本文提出的人工智能 2D 转 3D 技术,不仅将现有 3D 内容制作效率提升 1000 余倍,而且能够精确、稳定地转制所有类型的影视内容,有望彻底解决 3D 内容匮乏问题,为影视行业注入新动力。

1 3D 影视的发展及优势

1.1 3D 影视的原理及发展历程

人的视觉具有立体感,其产生原因之一是视差。

人左右瞳孔间距约为 65 mm,这使得左右眼看到的世界略有差异。将一个物体放在距鼻子约 20 cm 处,分别闭上左眼或右眼进行观察,会发现该物体与鼻子的相对位置关系在左眼和右眼的视野中有所不同,左眼看到的物体相对偏右,右眼看到的则相对偏左,这就是视差,它能使人脑产生强烈的立体感(图 1)。单眼通过调焦行为也能产生一定立体感,这源于眼睛的生理基

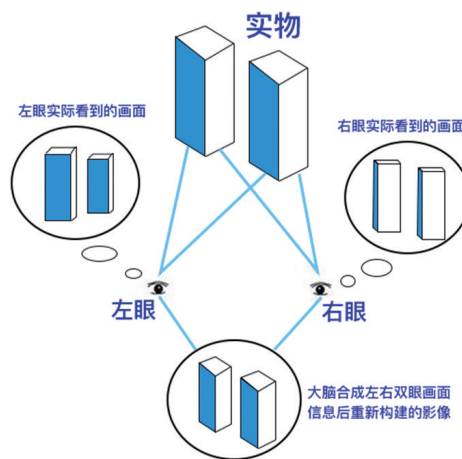


图 1 视差原理示意

Fig. 1 Schematic diagram of parallax principle

收稿日期:2018-04-24;修回日期:2018-05-03

作者简介:赵天奇,聚力维度创始人兼首席执行官,研究方向为计算机视觉和人工智能,电子信箱:zhaotianqi@12dms.com

引用格式:赵天奇. 人工智能为 3D 影视行业注入新动力[J]. 科技导报, 2018, 36(9): 66-72; doi: 10.3981/j.issn.1000-7857.2018.09.009

础。当物体距离较远时,水晶体变为扁平;物体较近时,水晶体变为凸出。因此,随物体的远近而引起毛状肌的伸缩作用,从而产生焦距的变化使人们可以区分远近不同的景物。

此外,生活经验也可以使人产生立体感。若两个物体具有相同的体积,人的经验判断会得出远处的物体较小,近处的物体较大;如果一个物体(A)遮盖了另一个物体(B)的一部分,通常A物体距离人更近,并且,距离越近的物体色彩越鲜明,距离越远色彩变灰。

3D影视则是利用左右视差原理,使人对画面产生立体的视觉效果。

16世纪,人们已经开始用不同的颜色分别为左右眼绘制有一定规律差异的图像,通过滤光镜观察,使人产生立体视觉效果。17世纪末到18世纪初出现了“立体镜”,可为左右眼睛分别提供独立的视觉通道。19世纪科学家曾尝试构造一种不借助辅助装置就能观察到立体画面的方法,但以失败告终。直到1838年英国科学家发明一种名为“反光式立体镜”(reflective stereoscope)的装置,人们才得以观看到立体的画面。直到1849年,Davis发明第一台可用于摄影领域的立体观片装置,名为“透镜式立体镜”(lens stereoscope)(图2)。

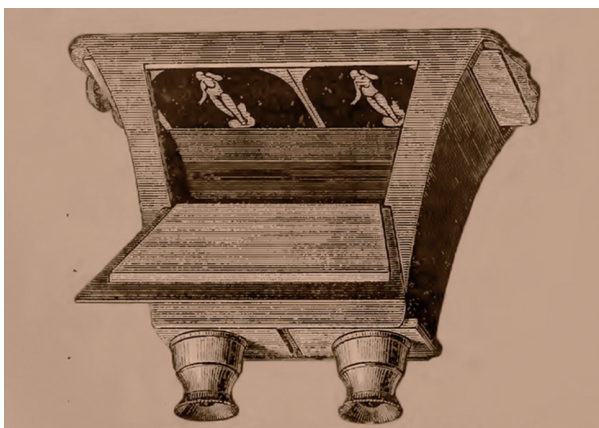


图2 1849年Davis Brewster的立体镜设计图

Fig. 2 Stereoscope design of Davis Brewster in 1849

19世纪末电影发明后,科学家尝试用电影来表现运动的立体视觉图像,3D观看设备由此起步。1922年9月27日,Harry和摄像师Robert制作的世界第一部3D电影《The Power of Love》上映,1952年,美国好莱坞也推出3D立体电影《Bwana Devil》。

3D影视内容,在20世纪50年代开启了短暂黄金时代。早期的3D电影都是以展示立体效果为主,常以指

向观众的枪、扔向观众的物体为噱头,让观众爆出一身冷汗。主流电影人多对这种形式的电影抱有排斥态度,20世纪90年代3D电影渐渐退出大众视野,直到2009年,电影《Avatar》上映,其全片采用3D视效,场面恢宏、逼真,才正式引爆3D商业电影市场。

1.2 3D对影视表达的优势

Neil Postman说“媒介即隐喻”。媒介形式不仅是一种载体,还会影响其所传播的内容,这在3D这种媒介形式上表现尤为明显。

1) 可增加现场感。

现场感是3D赋予电影最具突破性的魅力。例如,在电影《Avatar》中,James用其研究长达数10年的3D技术,将观众带入一手创造的潘多拉星球:Na'vi人奔跑在高空中的藤蔓上;男主驾驭他的飞龙与女主角翱翔在潘多拉的天地之间;水母状的种子自由飞舞;漫天的火光等3D技术,让观众与主角一起身临其境游历“潘多拉”。

2) 烘托故事氛围。

3D不同于文字表述,它可以用震撼的视效烘托故事气氛。《Titanic》于1997年上映2D版,15年之后,James推出转3D版《Titanic》,再次轰动全球。谈到为何转3D,这位导演说:“3D效果把细腻的情感交流变得更加生动,尤其是一些感人的时刻,在3D的渲染下,感受会更加强烈,视觉感官才能提升到另一个新境界。”^[3]

在文学创作中,烘托是一种从侧面渲染来衬托主要写作对象的表现技法。而在影视作品中,3D则通过其真实,让观众进入画面里的世界,与角色感同身受,从而更好地理解作品。

3) 影响叙事结构。

3D能影响观众对影片的叙事解读。《一代宗师》也曾先后推出2D和3D版本。导演王家卫对这两个版本的解读是:“2D版就像中国的水墨画,大写意,讲的是整个布局、讲的是一个时代,而3D是工笔人物,着重关注的是这个时代中,几个人的变化和命运。2D版的武林是遥远的,3D是亲切的回归。”也有媒体给出2D版是诗,而3D版是小说的评价。

关于3D技术的使用,艺术家们仍在探索。一方面让观众更加沉迷于电影所讲述的故事,拍手称赞;另一方面,也因为某些不当尝试,导致观众头昏眼花。

从媒介形式上看,任何媒介技术都有其两面性。语言:直接、简练但不易保存;文字:容易存储,却不够

生动。3D当然也不例外,找到充分发挥其核心价值的途径才是关键。

2 3D影视内容制作现状

《Avatar》引燃观众对3D艺术魅力的追捧后,为了将3D影视一举推向繁荣,电影公司、电视台乃至3D电视生产商纷纷通过不同合作模式进行制作尝试,但在创新制作技术上,却显乏力。

2.1 目前3D制作的商业模式

由于传统3D内容制作技术耗费成本高、需求人力多、周期长,无法满足电视台长时间、少循环的播放需求,所以目前的3D内容形式仍以电影为主,3D制作商业模式也主要依赖资金雄厚的电影制片商独立制作。

1) 电视厂商合作传媒公司。

电视生产厂商为了销售3D电视设备,纷纷向影视传媒公司伸出橄榄枝,开始了硬件设备方联手内容方的制作模式。如三星宣布与梦工厂、Technicolor建立战略合作伙伴关系,推出三星独家的3D节目;索尼宣布与探索传播、IMAX公司在美国合资建立首个专门的3D电视网;松下宣布与美国最大的卫星电视服务商DirecTV Group建立合作。

2) 传媒巨鳄独立制作。

也有一些大的传媒公司,依靠其自身雄厚的财力与人力基础,独立展开3D内容制作。如传媒巨鳄迪斯尼在3D内容制作上选择了独立尝试,在其旗下频道ESPN开设3D体育频道,并进行为期1年试播,试播期间每天节目不少于2小时。

3) 电视台合作。

在中国,电视台通过多家合作方式制作3D内容。例如,2012年1月1日,中国的3D试验频道成立,就由中央电视台、北京电视台、天津电视台、江苏电视台、上海电视台、深圳电视台6家单位开办,首播4.5小时,重播2次,每天播出时间共13.5小时。

4) 综合产业基地。

此外,发展出一种包含影视策划、剧本筹备、拍摄剪辑、技术支持、后期制作、传播交易在内的整合3D影视产业链。如北京、天津、深圳等地相继建立以3D影视制作为核心的产业园或基地。但业界多方联动并未能使3D影视内容短缺状况得到根本改观。消费者可接触到的3D内容仍十分有限,观看3D内容相比普通2D节

目,仍是昂贵的体验。

2.2 目前3D内容制作技术

导致3D内容匮乏的根本原因是制作技术。3D内容制作技术主要有2种:3D双机实拍和2D后期转3D。

1) 3D实拍技术。

3D实机拍摄是用2台摄影机精确同步输出左、右眼各自对应的影像,再通过3D视频剪辑后,合成真实震撼的立体效果的技术。电影《Avatar》的多数画面,就采用这种方式制作而成。采用这种方式制作而成的3D电影,立体效果非常震撼。但是,一部实机拍摄的3D电影也需要耗费巨大的资金、人力和时间成本。

首先,拍摄设备昂贵。一个立体摄像机的架子、带数码的跟踪设备超过600万欧元,其厚重的机身外加厚重的摇臂,也不便移动。

其次,并不是所有镜头都适合双机实拍。当双目摄像机距离拍摄物体过近,会造成视差过大,超出人眼习惯,使观众产生头晕、头疼等严重不适感;而距拍摄物体过远,可能会使透视效果严重变形,空间距离看起来过近或过远等。

在后期上,3D剪辑除了要完成2D剪辑的所有步骤,还要完成景深、汇聚面、视差的校正和调整,耗时甚多且无成熟流程。

2) 2D后期转3D。

另一种3D影视制作方式是2D后期转3D,利用的仍是2D视频素材。传统2D转3D技术是将原本只有XY方向2个维度的2D图像,通过计算机和人工处理而得到画面内各被摄物体的纵深信息(即Z方向),并借助此信息“构建”出一对具有水平立体视差的左右眼图像。主要包括以下流程(图3)。



图3 2D转3D技术主要制作流程

Fig. 3 The main production process of 2D to 3D Technology

(1) 图像分割。通常是用计算机软件加人工绘制的方式将画面中每个物体单独分割开(图4)。分割的精细程度取决于要求的转换质量及投入的时间和费用。

(2) 创建深度图。每个被分开的面都需要被分配一个深度信息,这些各自独立的深度信息组合成为画面的深度图(图5)。这是一个不断重复的过程,需要根据立体效果不断调整。



图4 红色区域是对人物进行分割

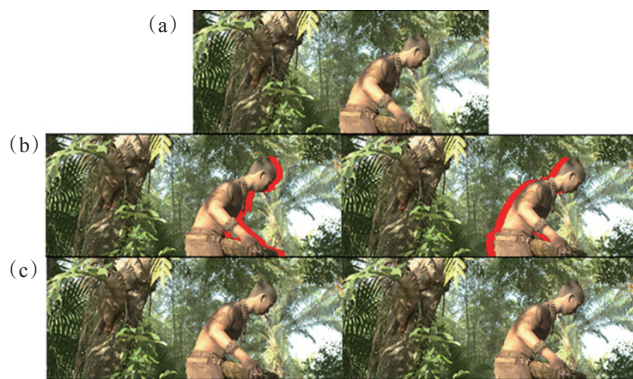
Fig. 4 The red area is the segmentation of the characters



图5 不同灰度代表不同的立体深度

Fig. 5 Different grayscale represents different stereoscopic depth

(3) 背景补图。在2D转3D过程中,原始的2D素材有时被当作左右立体图中的一个画面,通过位移生成另外一个画面;有时也会被当作介于左右画面之间的中间画面,相较于前一种做法,这样做的好处是由于立体产生的位移平均分配给了左右两个画面,提高了画面精度(图6)。图像经过位移后,在物体边缘会形成部分空隙,需要根据背景信息进行重新填补。



(a) 原图;(b) 红色部分为左右产生的平移区域;

(c) 红色空白部分补图完成后的画面

图6 2D转3D时背景补图

Fig. 6 Background completion of 2D to 3D technology

(4) 渲染合成。检查立体效果,如果立体画面看起来太“平”,就要对画面中的主要物体进行深度微调、三维建模及重新贴图渲染等操作。然后将字幕、立体图像序列和音频文件整合成完整立体影片。有数据显

示,2D转3D的成本可以控制在300万~500万人民币之间。而多数镜头采用3D实拍的《Avatar》耗资近5亿美元。虽然两者效果不能直接比较,但实机拍摄和后期转制3D资源耗费的差距从中可见一斑。

3 人工智能2D转3D技术

传统转3D技术相较于双机实拍,虽然将3D影视内容制作的时间和资金成本大幅降低,但3D内容仍然没能走向普及。艺恩网发布的一份调研报告,2015年PGC的2D产量达约15000集,在各平台不去重的情况下,优酷土豆节目覆盖率达到95%^[4]。而通过国内知名影评论坛豆瓣电影统计,2015年国内公映的3D电影仅69部,3D电视剧、3D综艺、3D体育、3D新闻、3D科教等节目则几近于无。

内容匮乏的主要原因在于现有制作水平的落后、制作周期长、成本消耗高。目前,国内每分钟2D画面的3D制作成本最低2万。以此推算,曾一度火爆网络的《太子妃升职记》,全剧37集,每集20分钟,转制成3D版本的费用就需要1480万,堪比其实际制作成本。并且,需要100人规模的团队,大概花费2年时间才能完成。

针对转3D技术在推动3D内容普及方面遭遇的困境,北京聚力维度科幻成真实验室历时3年研发出一种基于人工智能的3D内容自动转制技术。

3.1 基本原理

2016年,人工智能走进大众视野。谷歌人工智能程序AlphaGo以4:1的战绩,打败人类围棋世界冠军李世石,轰动世界。近年来,人工智能技术已飞速融入各个领域。在金融、安防、自动驾驶、医疗等领域,都可见人工智能踪影,并且占据十分重要的位置。

人工智能2D转3D技术是人工智能技术在影视制作领域里的一项颠覆和创新,主要基于深度神经网络其构成如图7所示。

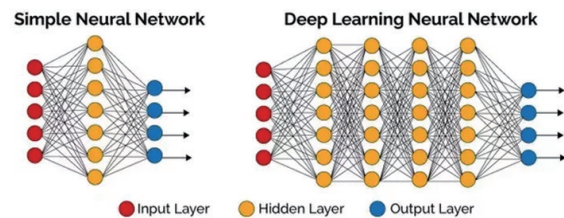


图7 左侧为简单神经网络,右侧为深度神经网络

Fig. 7 The left side is a simple neural network and the right one is a deep neural network

深度神经网络是近几年出现的新的的人工智能技术,它继承了浅层神经网络对于单个神经元的仿生,同时又增加了神经网络的层数,分为输入层、输出层和中间的隐藏层。

人工智能2D转3D技术就是把特征提取的工作交给神经网络来做,输入层可以是最原始的数据,不需要人工预先处理,而隐藏层可以实现对原始数据进行特征提取,最后在输出层实现效果。目前神经网络有多种方式,卷积神经网络、递归神经网络等,各自适合于不同领域,其中卷积神经网络模仿了生物视觉神经皮层的结构,适合于视觉领域。但这种方式需要大量的数据进行训练,如果数据量不够就无法训练成功。

3.2 人工智能2D转3D系统

人工智能2D转3D系统可以进行高效率、高精度2D转3D生产工作(图8)。该系统针对当前电影后期制作主要靠人工,极少自动化的现状,以及专门人工智能技术无法解决内容丰富多变的电影问题,给出多种新技术方法。

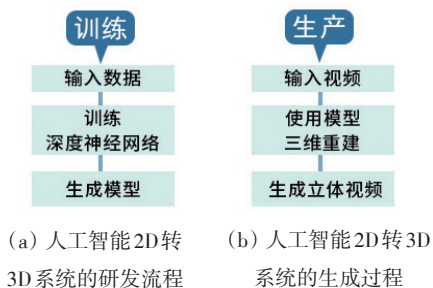


图8 人工智能2D转3D系统的研发流程与该系统生成3D过程

Fig. 8 Development flow of artificial intelligence 2D to 3D system, and to use the system to generate 3D process

在理论上,采用一种新的逐层认知深度神经网络模型。通过对神经网络各层的逐层学习与训练实现认知的立体化、健全化、通用化,从而满足电影内容复杂多变的智能自动处理问题。在技术上,针对影视2D转3D找到一种新的视频特征。这种特征只与边界属性有关,不因对象是什么而改变,可以满足边界属性不变的多样对象电影制作的自动化。在应用上,把所有的自动化算法都跟资深人工生产团队的制作流程相结合,所以能够在多样化的生产中解决实际问题,提升整体处理效率。

3.3 转制效果

从实现效果看,人工智能系统将院线电影的3D转

制效率提升1000余倍。使用传统转3D技术转一部90分钟的电影,需要100人团队,花费大约3个月时间才能完成。而在人工智能转3D技术下,同等上映品质,仅需要1人,花费1周时间即可完成。除了有1000余倍的高效率之外,人工智能转3D技术还具有精确、稳定、全能的优势。

1) 精确性。在传统转3D技术中,画面生成的立体效果精确度取决于分割层次。而分割情况,又与提供的转制时间、资金花费、质量要求及立体制作人员的技术水平密切相关。通过精选数据集和优化算法,科幻成真实验室让人工智能2D转3D系统能自动将画面里的每一个物体、元素都精确分割、绘制,最大限度还原画面的真实立体感。图9中,(a)原图用传统手工转3D方式三维重建,人眼很难看到右上角的半张人脸。而从灰度图(b)可见人工智能系统精确绘制出了右上角的人脸灰度。为了检验人工智能系统,对原图进行调色观察,才看到右上角的人脸(图9(c))。



图9 人工智能系统高于人眼识别能力
Fig. 9 Artificial intelligence system is higher than human eye recognition ability

2) 稳定性。在转制3D的过程中,除了单帧画面的立体分割,还要关注连续帧的立体效果,才能做出在影院观看的活灵活现的3D影视作品。传统转3D方式,由于工序复杂、流程较慢,很多电影制作方为了赶档期会

将整部电影的3D转制工作分包给不同的3D制作公司。由于不同公司使用的3D转制软件不同,不同立体设计人员对立体感大小的把控也不同,导致播放连续画面时,会出现同一人物、物体在同一场景、同一位置,时而给观众观感较近,时而较远的抖动感。这一点,在人工智能转3D系统中,完全不用考虑。超高的效率,使得系统可以按照编入的统一设计标准,完成整部作品的3D转制工作。

3) 全能性。人工智能转3D系统在影视画面中的识别率高达99%。无论是都市剧里的现代街道,还是古装剧里的飞檐塔楼都能使用该系统完成3D转制工作。为实现3D内容的量产,中国将努力并积极地与互联网视频平台合作建立3D媒资库,将成百上千的2D影视资源用人工智能系统转制成3D资源,丰富人们的3D观看体验。

4 展望

精神分析学理论解释了电影“似梦”特性,通过“移情”与“想象”,让观众在观影中获得审美满足感,而3D电影的“立体”“虚拟现实”梦境在一定程度上营造着深度“立体幻觉”与真实的触觉感官刺激^[5]。可以遇见的未来是,在3D双机实拍、传统2D转3D技术尤其是人工智能2D转3D技术的共同作用下,人们的娱乐文化生活将会发生巨大改变。

4.1 催生新应用

新技术带来的不仅是丰富的3D内容,还将为3D设备提供多样化的应用,包括3D直播和3D浏览器等。

1) 3D直播。基于人工智能2D转3D系统,科幻成真实验室已经研发出的3D直播应用。这项应用,就是要将电视频道正在播放的普通电视节目实时转换成3D节目。转换后的3D节目与原播放内容延迟不超5 s。新闻、体育、军事、音乐等,任何类型、任何频道都能够实现实时转换。另外,这项技术也能应用到网络视频平台,并将其变成3D网络平台。

不久的将来,不仅在家中可以看到3D电影,而且可以通过电视或网络观看实时转换的3D新闻、3D电视剧、3D体育赛事等。

2) 3D浏览器。3D浏览器是科幻成真实验室基于人工智能2D转3D系统研发的另一项3D应用。目前,

人们只要打开电脑上网,大多都会用到浏览器。因此,3D浏览器将为人们带来新的体验。例如,浏览3D淘宝网,犹如穿越,将人们融入商品所展示的环境中,切身感受拥有这件商品后的美妙情境。因此,3D浏览器网购所带来的极有可能是巨大的消费热潮。又如,计划假日行程,用3D浏览器搜索游览地的图片,就能将风景真实般地展现在人们眼前。此外,用3D浏览器观看3D动漫也是一种享受,立体化后的漫画,会让人产生一种错觉,随着画格的翻动,仿佛已经不是再看漫画,而是体验一部3D大电影。

4.2 3D观影由被动变为主动

目前,3D的精彩放映内容主要集中在影院。消费者只是可以选择去看或者不去看,但在播映哪些内容上几乎没有选择权,是一种被动式的观影体验。在人工智能2D转3D制作技术下,影视内容的质量和产量都会有大幅提升,越来越多的精彩影视节目不仅包括电影,其他如电视剧、体育比赛、记录片、综艺节目等都有望促成3D版,为用户提供多样选择。同时,丰富的内容也会提升3D电视、3D投影机、3D平板电脑等3D硬件设备的实用性,进而带动硬件设备的发展。

4.3 重新塑造3D观影感受

观影作为一种观众与电影作品之间的互动行为,集合着观众对于作品的视觉感知、剧情理解与观影想象等元素,3D电影因其空间纬度变迁带来全新观影与审美体验,升级着“幻觉”与“梦境”效应^[6]。目前,由于3D内容的生产成本较高,消费者3D体验可选择性较差,票价在某种程度上,会影响消费者对3D内容的评价。传统3D内容制作方根据给定的制作周期、资金等制定转制完成的3D效果标准。当转制效果有瑕疵时,较高的票价会使观众有一定落差。

随着技术进步,可观看、选择的内容增多及体验花费的降低,消费者会有更多机会与更强烈的意愿去体验3D,站在一个更客观、公正的立场上去评判3D影视内容。3D是显示技术上的重要突破,也是人类传播方式上的重要改变,并期待人工智能2D转3D技术能为影视行业乃至人类传播方式变革添加浓墨重彩的一笔。

参考文献(References)

[1] 喻德术. 洋3D给力 2011完成130亿[N]. 法制晚报, 2012-01-

- 03(17).
Yu Deshu. Foreign 3D is awesome 2011 complete 13 billion [N]. Legal Evening Newspaper, 2012-01-03(17).
- [2] 聂宽冕, 冯赛琪. 3D电影这下有些尴尬了[N]. 北京日报, 2018-01-27(10).
Nie Kuanmian, Feng Saiqi. 3D movies are a bit embarrassing [N]. Beijing daily, 2018-01-27(10).
- [3] 朱梁. 工艺繁复“令人抓狂”电影大片怎样“2D转3D”. 北京日报[N]. 2012-05-11(6).
Zhu Liang. The complex process is "crazy" How does movie blockbuster "2D turn 3D"[N]. Beijing daily, 2013-05-11(6).
- [4] 2016年PGC产业生态研究报告[EB/OL]. (2017-02-20)[2018-03-20]. <http://www.useit.com.cn/thread-14631-1-1.html>.
2016 PGC Industrial ecology research report[EB/OL]. (2017-02-20)[2018-03-20]. <http://www.useit.com.cn/thread-14631-1-1.html>.
- [5] 付茜茜. 国产3D电影与观众认同:一次观影调查的思考[J]. 今传媒, 2016(6): 67-70.
Fu Qianqian. Domestic 3D movies and audience identity: A reflection on the survey of[J]. Today's Massmedia, 2016(6): 67-70.

Artificial intelligence is a new driving force for film and television industry

ZHAO Tianqi

Beijing Dream World Technology Co., Ltd., Beijing 100089, China

Abstract The principle of 3D-vision generation and the exploration process for human 3D-vision are introduced, and the way of making 3D content is analyzed. With the development of computer vision and artificial intelligence technology, a new transformation technology from 2D to 3D has been proposed and developed. The technology can automatically generate 3D content according to 2D images and the efficiency of 3D transformation has increased by more than 1,000 times without loss of quality. It is pointed out that based on this technology, new applications such as 3D live broadcast and 3D browser can be developed, and it is also expected that the problem of 3D content shortage is completely solved.

Keywords 3D stereoscopic film; depth estimation; artificial intelligence 2D to 3D; 3D content production ●



(责任编辑 卫夏雯)