

# 从数字媒体技术应用于专业职业教育视角 浅析虚拟人制作、应用和教学

● 江 玥



**[摘要]** 近年来,虚拟人逐步成为职业教育数字媒体技术应用专业新的培养方向。基于上述教学需求,从职业教育视角出发,对虚拟人相关教学(制作与应用)思路的探讨有着重要的实际意义。本文在论述虚拟人定义、分类、技术、应用和制作的基础上,建议可遵循先易后难、先掌握(应用)后研发(底层技术)、先普及后择优的思路来开展虚拟人教学实践。同时,基于职业教育的初衷和未来虚拟人发展的广阔前景,建议可采用工学一体化课程教学方式开展相关教学。

**[关键词]** 数字媒体技术应用;职业教育;虚拟人;教学思路

## Q 虚拟人定义

虚拟人(Virtual Human)是指具有数字化外形的虚拟人物,依赖显示设备存在,并拥有人的相貌(像人)、人的举止(能说话、能举手)以及人的思想(与人交流)。由于数字技术的运用,一定程度上虚拟人又被称作数字人。虚拟人其实质依然是一种人机对话方式。为此,虚拟人需具备“听、说、读、看、思”功能。其中,“听”“读”“看”是在人机交互过程中,对来自人的输入数据的识别;“思”是在人机交互过程中,为实现人的指令而对来自人的输入数据的处理;“说”是在人机交互过程中,对来自人的输入数据进行处理后所给出的反馈。具备虚拟人形象且具备“听、说、读、看、思”综合能力的虚拟人可定义为广义的虚拟人。相对于广义虚拟人而言,运用虚拟人像合成技术所生成的具体虚拟人物形象可定义为狭义的虚拟人,这种虚拟人形象充当了人机交互界面的角色。现阶段,虚拟人一般应用于承担一些重复性高的工作,从而达到降低成本、提高效率

的目的;长远来看,“人机耦合”是人工智能领域的主流趋势,人类与基于 AI(Artificial Intelligence)技术的虚拟人相得益彰将能创造出更多的劳动价值;而未来,无人工干预,以 AI 技术为支撑的广义虚拟人是虚拟人的主要发展方向。

## Q 虚拟人技术

为实现虚拟人所包含的“听、说、读、看、思”功能,基于深度学习(多层 ANN,即人工神经网络,Artificial Neural Network)的人工智能技术为核心,将 ASR(自动语音识别技术, Automatic Speech Recognition)、NLP(自然语言处理, Natural Language Processing)、TTS(从文本到语音, Text To Speech)、计算机视觉、视频通信、生物识别(人脸识别、表情识别)、深度神经网络渲染等各种 AI 技术进行综合利用,以生成具备感知、分析、认知、表达、行为等能力的智能虚拟人,并通过数据输入、数据处理、数据输出模块,实现人机交互,如表 1 所示。

表 1 虚拟人功能、技术与模块

功能	技术	模块	广义虚拟人	狭义虚拟人
虚拟人像	虚拟人像合成技术(深度学习、建模)	人机交互界面	具备	具备
听	ASR(自动语音识别技术, Automatic Speech Recognition)	数据输入	具备	
看	计算机视觉、视频通信、生物识别(人脸识别、表情识别)	数据输入	具备	
读	NLP(自然语言处理, Natural Language Processing)	数据输入	具备	
思	智能数据库(数据分析、数据生产)	数据处理	具备	
说	TTS(从文本到语音, Text To Speech)	数据输出	具备	

需要强调的是基于人工智能的深度学习技术的发展，对于虚拟人发展所起的显著推动作用(譬如，深度学习技术在图像识别、语音识别、自然语言处理方面的应用与突破)。深度学习技术正在不断演化并展现出充分的潜力。随着硬件技术的进步，未来具备更大规模和更快速度的深度学习模型将在虚拟人发展与应用领域继续扮演重要的角色。

### Q 虚拟人分类

虚拟人可以按照其形象风格、所处空间、是否交互、底层技术进行分类。第一，按照虚拟人形象风格分类，可分为Q版虚拟人、二次元虚拟人、仿真虚拟人。第二，按照虚拟人所处空间维度分类，可分为2D虚拟人、3D虚拟人、全息虚拟人。第三，按照是否与人交互分类，可分为交互型虚拟人和非交互型虚拟人。第四，按照广义虚拟人制作底层技术思路分类，可分为半虚拟虚拟人(以真人为后台支撑，即有人干预，实现“听、说、读、看、思”)和全虚拟虚拟人(以智能数据库为后台支撑，即无人干预，实现“听、说、读、看、思”)。第五，按照狭义虚拟人制作(虚拟人像合成技术)底层技术思路分类，可分为基于深度学习技术的虚拟人和基于建模技术的虚拟人。如表2所示。

表2 虚拟人分类标准与种类

虚拟人分类标准	虚拟人种类
形象风格	Q版虚拟人、二次元虚拟人、仿真虚拟人
所处空间维度	2D虚拟人、3D虚拟人、全息虚拟人
是否与人交互	交互型虚拟人、非交互型虚拟人
广义虚拟人制作底层技术思路	半虚拟虚拟人(有人干预)、全虚拟虚拟人(无人干预)
狭义虚拟人制作底层技术思路	基于深度学习技术的虚拟人和基于建模技术的虚拟人

以上不同类型的虚拟人，在技术成熟度、制作成本、制作周期、应用范围方面均呈现出较大区别。简言之，半虚拟虚拟人技术成熟度高、制作成本低、制作周期短、应用范围广；全虚拟虚拟人仍处于研发阶段，制作成本高、制作周期长，尚未被广泛应用。

### Q 虚拟人应用

现阶段，虚拟人已应用于金融服务、直播电商、新闻报道等服务行业。首先，以光大银行数字人为例，虚拟人已应用于金融服务业。光大银行推出3D形象数字人工“小璇”和2D数字人应用于现实业务中：前者以金融服务专家及生活助理形象示人，适用于不同金融场景及生活场景；后者在实际业务场景中，通过客户端与客户智能交互问答完成“随心贷”业务，在高效满足核实客户贷款意愿，实现网贷业务“面谈面签”监管要求的同时，将传统外呼方式改变为

客户与数字人进行人机交互的方式，进而能让客户体验到更加真实、更有温度的金融服务。

其次，以淘宝直播间虚拟形象“张小Z”为例，虚拟人已应用于直播电商行业。其直播间虚拟形象“张小Z”不是单纯的算法，而是由动作捕捉而来，背后有真人动作，还有人配音(即在底层技术层面属于有人化操作类型)，比较接近真人。

最后，以虚拟主播“N小黑”“N小白”为例，虚拟人已应用于新闻报道领域。每日经济新闻推出的虚拟主播“N小黑”“N小白”，均依靠“深度神经网络渲染”技术，通过对预先录制的2D视频素材进行分析和学习，再实现对语言、动作、表情、音色的“还原”，能做到虚拟人与真人的1:1真实复制。如表3所示

表3 虚拟人应用行业与类型

应用行业	虚拟人	虚拟人类型
金融服务	光大银行“小璇”	3D建模半虚拟
金融服务	光大银行2D数字人形象	2D半虚拟
电商直播	淘宝直播“张小Z”	3D建模半虚拟
新闻报道	“N小黑”“N小白”	2D深度学习半虚拟

可见，虚拟人具备广泛的市场潜力和发展前景。数据显示，2021年我国虚拟人带动的产业市场规模为人民币1074.9亿元；其中，核心市场规模为人民币62.2亿元；预计2025年，上述产业市场规模和核心市场规模将分别达到人民币6402.7亿元和人民币480.6亿元，呈现强劲增长态势。此外，鉴于半虚拟和全虚拟两种广义虚拟人底层技术思路在成熟度、制作成本、制作周期等方面的差异，现阶段，2D半虚拟主播的应用范围相对较广。

### Q 虚拟人制作

广义虚拟人制作，其底层技术思路存在半虚拟(有人干预)和全虚拟(无人干预)的差异；无论采取哪一种思路，都涉及基于大量AI技术应用所开发的功能模块。由于篇幅有限，本文将着重讨论狭义虚拟人的制作。

制作狭义虚拟人(虚拟人像合成)的底层核心技术包括：第一，深度学习。基于深度学习模型架构，采用神经渲染技术(譬如：生成对抗网络GAN, Generative Adversarial Nets)，通过对大量数据的分析和学习，让机器理解并实现对人与物在不同空间维度中形态与特征的模拟，包括表情、声音、动作等。前文所提及的每日经济新闻所推出的虚拟主播“N小黑”“N小白”，均是基于这种思路采用“深度神经网络渲染”技术生成的。由于这种技术其逻辑是通过录制视频图像进行分析和学习后进行合成，所以，最终生成的2D虚拟主播的呈现效果不能超越视频素材本身。但综合而言，目前该技术日臻成熟，优势明显。

以3D建模为例,可通过设备扫描、软件建模以及设备扫描加软件建模三种方式来完成。设备扫描方式指使用3D扫描设备扫描真实人体以获得分型结果,从而最终搭建出“无限接近”真人的三维人体模型;软件建模方式指使用建模软件(3Dmax、Maya、Unity3D等)进行三维建模,创作出虚拟形象;设备扫描加软件建模方式,是指使用建模软件来优化和调整通过使用扫描设备对真人进行扫描得来的真人模型。然而,由于制作周期长、费用高,这种3D虚拟人制作技术尚未普及。

第二,基于以上两种底层技术思路,一些专业虚拟人制作软件平台已投入实际应用中。这些专业虚拟人制作软件平台的使用与发展,较大地促进了虚拟人在相关行业的应用。以软件Adobe Character Animator为例,用户可以根据其需求创建专属的2D动画虚拟人角色。基于人工智能技术,运用该软件平台,可以通过摄像头和麦克风等设备实现自主实时面部追踪绑定、语音识别和动作捕捉,从而创造2D虚拟人角色的动画。这款应用程序速度非常快,并与其他Adobe应用程序集成在一起,实现了无缝衔接的动画制作工作流程。

## Q 虚拟人教学

近年来,虚拟人逐步成为职业教育数字媒体技术应用专业新的培养方向。虚拟人制作和应用可包含于全媒体运营师等相关工种的工作职能范围内。通过相关教学,使学生能对虚拟人的设计、制作,以及其在全媒体等相关领域的应用有较深入的了解,使学生具备以虚拟人为人机交互载体,将文字、声音、影像、动画、网页等信息内容策划和加工后向目标受众进行精准分发、有效传播和特定营销等能力,帮助学生能在毕业后尽快适应全媒体运营师等相关技术岗位,并较好地完成全媒体运营工作任务。

基于上述教学需求,从职业教育的视角出发,对虚拟人制作与应用的相关教学思路的探讨有着重要的实际意义。具体而言,可以从现实需求、难易程度、学生基础、未来岗位等维度来进行相关探讨。首先,从相关产业对虚拟人的现实需求维度出发,基于本文第四部分和第五部分所作的有关论述,出于对技术成熟度、制作成本、制作周期等因素的考量,半虚拟2D虚拟人为目前金融服务、直播电商、新闻

报道等虚拟人主要应用场景的首选。其次,从与虚拟人制作和应用相匹配的技术要求维度出发,半虚拟2D虚拟人较其他类型虚拟人的制作思路而言实操性更强,更易于项目落地实施。再次,从学生学习基础维度出发,半虚拟2D虚拟人制作流程和技能较其他类型虚拟人制作流程和技能而言,更易于为学生所理解和掌握。学生可在掌握半虚拟2D虚拟人制作和应用的基础上,尝试学习制作和应用其他类型虚拟人。最后,从未来相关工作岗位的维度出发,熟练掌握半虚拟2D虚拟人的制作和应用,可以帮助学生在金融服务、直播电商、新闻报道等行业从事相关岗位的工作。在此基础上,继续学习并熟练掌握其他类型虚拟人的制作和应用,有利于相关专业学生在上述行业内的可持续发展与提升。

基于以上认识,本文建议在虚拟人制作与应用教学中可遵循先易后难、先掌握(实际应用)后研发(底层技术)、先普及后择优的思路来进行教学实践。同时,基于职业教育的初衷和未来虚拟人应用的广泛潜力,建议可采用工学一体化课程的教学方式来开展相关教学。具体而言,应遵循工学一体化课程教学的指导原则,在“高仿真”或“全真”环境下,将理论教学与实践操作相结合,将具有代表性的工作项目(由企业真实工作演化得来)作为教学导向,将典型工作任务(对应行业先进技能)作为教学载体,将项目教学、四阶段教学、角色扮演、翻转课堂等方法作为教学方法,将学生作为教学中心,让学生在完成工作任务的过程中学习如何工作,从而掌握在未来实际工作岗位上必须具备的专业技能和职业素养。

## 参考文献

- [1]裴亚民,张彬.光大银行“数字人”建设实践[J].中国金融电脑,2022(03):35-39.
- [2]覃凯.人工智能背景下AI虚拟主播直播带货创新应用研究[J].商场现代,2022(05):14-16.
- [3]刘学东,万高迈.浅析虚拟主播在财经媒体中的应用和未来发展——以每日经济新闻AI电视为例[J].采写编,2022(10):9-11.

## 作者简介:

江玥(1982-),女,汉族,山东济南人,硕士,讲师,深圳技师学院,研究方向:职业教育、学生发展与教育、数字媒体传播。