

创意基因：AIGC 的本质与进化



章节导语

在这个数字智能迅猛发展的时代,AIGC(AI generated content,人工智能生成内容)正以革命性的力量重塑着创意设计的版图。从简单的文本生成到复杂的视觉创作,从单一的音频合成到沉浸式的多媒体体验,AIGC 不仅在改变内容创作的方式,更在重新定义创意的本质。本章将带学生探索 AIGC 的核心原理,追溯其发展轨迹,洞察其在创意领域的多元应用,并思考它为设计行业带来的机遇与挑战。

本章将深入剖析 AIGC 的技术内核,特别是新的扩散模型(diffusion model)如何突破传统生成技术的局限,实现更高质量的创意内容生成。通过对比分析各类 AIGC 平台的特点和优势,学生将了解如何选择适合自己需求的创意工具。同时,本章也将探讨 AIGC 与传统艺术创作的深度对话,思考技术与艺术融合的新可能。

作为未来的设计师,理解并掌握 AIGC 技术将提高学生的竞争力。本章的学习不仅会帮助学生把握 AIGC 的技术脉络,更重要的是启发学生思考如何将这些新型工具融入创意实践,在人机协作中开创属于自己的设计新天地。让我们一起揭开 AIGC 的神秘面纱,探索人工智能与创意设计碰撞的无限可能吧!



学习目标

知识目标:

- (1) 深入理解 AIGC 的核心原理,特别是扩散模型等前沿技术的工作机制。
- (2) 系统掌握 AIGC 在文本、图像、音频、视频等领域的发展动态。
- (3) 准确认识主流 AIGC 平台的特点及其在设计领域的应用价值。

能力目标:

- (1) 培养运用 AIGC 工具进行创意设计的实践能力。
- (2) 提升在人机协作环境下的创新思维和问题解决能力。
- (3) 发展对新技术的学习适应能力和批判性思维能力。

素养目标:

- (1) 树立正确的技术伦理观,理解 AIGC 在创作中的辅助定位。
- (2) 培养持续学习的意识,保持对新技术发展的敏感度。
- (3) 建立开放包容的创新意识,勇于探索技术与艺术的融合可能。



1.1 解码 AI 创意：生成式人工智能的内核

在这个数字化浪潮汹涌的时代,AIGC 正在悄然改变我们对创意生产的认知。本节将深入探讨 AIGC 的核心原理、技术基础和应用前景,帮助读者全面理解这一革命性技术如何重塑创意产业。我们将从 AIGC 的定义出发,详细介绍其底层技术,特别是最新的扩散模型,并通过生动的例子展示 AIGC 在各个领域的应用。同时,我们也将探讨 AIGC 带来的机遇与挑战,以及它如何改变创作者的角色和工作方式。

1.1.1 AIGC：创意生产的新纪元

AIGC 全称“人工智能生成内容”,代表了一种全新的创意生产方式。它不同于传统的人工智能技术,不再只是感知、理解和决策,而是将重心放在了“生成”上——利用智能算法创造出文字、图像、声音、视频等多种形式的创意内容。AIGC 的技术基础来源于深度学习中的各种生成模型。在 AIGC 发展的早期,生成对抗网络(GAN)和变分自编码器(VAE)是两个主要的算法。这些算法就像是勤奋的学生,通过大量的“学习”和“练习”,逐渐掌握了内容的抽象特征和生成规律。在实际应用中,我们只需给出一些指示(如文字描述、关键词、参考图片等),这些算法就能按照指示生成相应的内容。这种方式不仅提高了创作效率,还为创意表达开辟了新天地。

1.1.2 扩散模型：AIGC 的“新宠”

近年来,一种叫作扩散模型的新技术在 AIGC 领域大放异彩,成为当前最先进的生成模型之一。扩散模型的工作原理很有趣,它模拟了一个双向过程:一个逐步添加“杂质”的正向过程和一个逐步“净化”的反向过程。在正向过程中,模型慢慢向原始数据添加“杂质”,直到数据变得面目全非。这就像是把一幅清晰的画作慢慢涂抹,最后变成一团模糊的色彩。而在反向过程中,模型学习如何从这团模糊中一点点还原出有意义的内容。这就像是画家从一团模糊的色彩中,一笔一笔地勾勒出清晰的轮廓,最终呈现出一幅完整的画作。扩散模型的优势在于它能创造出更高质量、更丰富多样的内容,同时训练过程更加稳定。这种优势使得扩散模型在各种 AIGC 应用中表现出色。例如,大家可能听说过的 DALL-E 3 和 Stable Diffusion 等图像生成工具,它们就采用了扩散模型的原理,能够根据文字描述生成图像。

1.1.3 AIGC 的工作原理与应用示例

让我们用一个生动的例子来理解扩散模型是如何工作的。假设我们想要一张“一只可爱的柯基犬在海滩上奔跑”的图片(图 1-1)。扩散模型首先会创造出一片“混沌”,就像是一张涂满随机颜色的画布。然后,它会根据我们的描述,慢慢地从这片混沌中勾勒出图像。它会考虑“柯基”“可爱”“海滩”“奔跑”等关键词,确保生成的图像包含这些元素。最终,一幅生动的画面会呈现在我们眼前:沙滩上,一只短腿但活力十足的柯基犬正在海边欢快地奔跑,它的表情显得十分可爱。AIGC 的应用范围非常广泛,不仅限于图像创作,



在音乐领域也大有可为。例如,当我们哼出一段简单的旋律,AIGC 就能以此为基础,通过类似的过程,创作出一段全新的音乐。这种方式创作的音乐往往连贯和谐,仿佛是由专业音乐人精心创作的作品。



图 1-1 一只可爱的柯基犬在海滩上奔跑

1.1.4 AIGC 的优势与潜力

AIGC 的魅力在于,它打破了内容创作的人力限制,使得大规模、个性化的内容生产成为可能。想象一下,在传统创作模式下,画一幅画或写一首歌可能需要很长时间。而借助 AIGC 技术,我们可以在很短的时间内生成大量的创意内容。这不仅提高了创作效率,还为个性化内容的大规模生产提供了可能。同时,AIGC 创作的开放性和不确定性,也为创意带来了无限可能。每一次生成都可能有意想不到的结果,这种“惊喜”往往能激发创作者的灵感,引导他们探索新的创意方向。在这种技术的帮助下,每个人都有机会成为内容的创作者和创新者,不再受限于传统技能的掌握程度。

1.1.5 AIGC 与人类创作者的关系

然而,我们也要认识到,AIGC 并不能取代人类创作者,而是成为创作者的得力助手。AIGC 可以帮助创作者快速实现初步构想,生成大量备选方案,但最终的艺术判断和创意决策仍需要人类的智慧。在 AIGC 时代,创作者的角色正在从“亲力亲为”转变为“指挥调度”,这要求我们具备更高层次的创造思维和审美能力。作为未来的设计师和创意工作者,深入理解 AIGC 的原理和应用,将帮助我们在这个新时代中占据先机,创造出更多令人惊叹的作品。让我们一起拥抱这项新技术,探索创意的无限可能!

1.2 技术进化论：AIGC 的里程碑时刻

AIGC 的发展历程是人工智能技术不断迭代与内容创作领域持续融合的过程。通过回顾 AIGC 的发展脉络,我们可以识别出几个重要的里程碑时刻,这些时刻不仅标志着技



术的突破,也预示着内容创作范式的转变。

1.2.1 AIGC 的早期萌芽

AIGC 的故事可以追溯到 20 世纪 50 年代。1950 年,被誉为人工智能之父的艾伦·图灵在其论文《计算机器与智能》中提出著名的图灵测试,为人工智能的发展指明了方向。尽管这个时期的技术还无法实现复杂的内容生成,但为后续的发展奠定了重要的理论基础(图 1-2)。

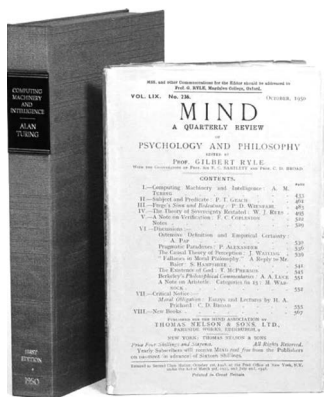


图 1-2 《计算机器与智能》

1.2.2 深度学习与 GAN 的突破

真正的突破要到 21 世纪初。2012 年,Krizhevsky 等人提出的 AlexNet 在 ImageNet 大规模视觉识别挑战赛中取得突破性进展,标志着深度学习时代的到来,深度学习为 AIGC 的发展奠定了坚实的技术基础。2014 年 10 月,Ian Goodfellow 等人首次提出生成对抗网络(GAN)的概念,这一突破性进展为 AIGC,特别是图像生成领域带来了革命性的变化。

1.2.3 扩散模型和大语言模型的兴起

2015—2020 年,扩散模型逐渐兴起。2015 年,Sohl-Dickstein 等人提出扩散模型的早期版本。2020 年 6 月,Jonathan Ho 等人提出去噪扩散概率模型(DDPM),大幅提高了扩散模型的实用性和生成质量。这一时期,大语言模型也迎来了飞速发展。2017 年 6 月,Vaswani 等人提出的 Transformer 模型为大规模语言模型的发展铺平了道路。随后,OpenAI 相继发布了 GPT 和 GPT-3 模型,展示了大规模语言模型的巨大潜力,掀起了 AIGC 在文本生成领域的一场革命。

1.2.4 多模态生成的突破

2021—2023 年,AIGC 迎来了多模态生成的突破。2021 年 1 月,OpenAI 发布 DALL-E,首次展示了大规模文本到图像生成的能力。2022 年,Midjourney、Stable



Diffusion 和 DALL-E 2 相继问世,标志着 AIGC 正式进入大众视野。同年 10 月,Google 发布 AudioLM,展示了 AI 在长形式音频和音乐生成方面的能力。2023 年 2 月,Runway 发布 Gen-1 视频生成工具,AIGC 正式进入多模态时代。一个月后,OpenAI 发布 GPT-4,展现出更强大的多模态理解和生成能力(图 1-3)。

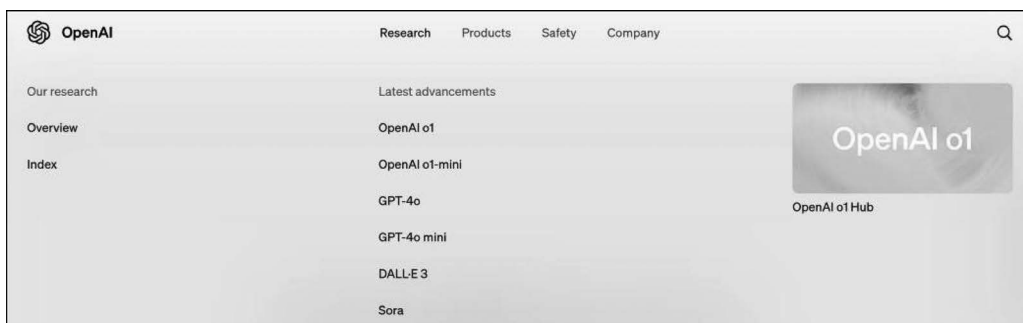


图 1-3 OpenAI 官网

1.2.5 AIGC 的深化与普及

从 2024 年至今,AIGC 继续呈现快速发展态势。多模态融合进一步深化,不同模态之间的融合更加紧密和自然,实现了文本、图像、音频和视频的无缝融合。针对不同行业 and 用户需求的专业 AIGC 工具不断涌现,在建筑设计、医疗诊断、金融分析等专业领域的应用不断深入。与此同时,相关部门开始制定 AIGC 作品的版权保护指南,为 AIGC 的健康发展提供法律保障。AIGC 相关课程也开始进入高校课堂,旨在提高学生的 AIGC 素养(表 1-1)。

表 1-1 AIGC 发展关键里程碑

年 份	事 件	意 义
1950	艾伦·图灵提出图灵测试	为人工智能的发展指明方向,奠定理论基础
2012	AlexNet 在 ImageNet 挑战赛中取得突破	标志深度学习时代开始,为 AIGC 发展奠定技术基础
2014	Ian Goodfellow 等人提出 GAN	为 AIGC 特别是图像生成领域带来革命性变化
2015	Sohl-Dickstein 等人提出扩散模型早期版本	扩散模型开始兴起,为后续 AIGC 发展埋下伏笔
2017	Vaswani 等人提出 Transformer 模型	为大规模语言模型的发展铺平道路
2018	OpenAI 发布 GPT 模型	展示了大规模语言模型的潜力
2020	OpenAI 发布 GPT-3	掀起 AIGC 在文本生成领域的革命
2020	Jonathan Ho 等人提出 DDPM	大幅提高了扩散模型的实用性和生成质量
2021	OpenAI 发布 DALL-E	首次展示了大规模文本到图像生成能力
2022	Midjourney、Stable Diffusion、DALL-E 2 相继发布	AIGC 正式进入大众视野
2022	Google 发布 AudioLM	展示 AI 在长形式音频和音乐生成方面的能力



续表

年 份	事 件	意 义
2023	Runway 发布 Gen-1 视频生成工具	AIGC 进入多模态时代
2023	OpenAI 发布 GPT-4	展现更强大的多模态理解和生成能力
2024	OpenAI o1 模型发布、DeepSeek V3 发布	推理能力突破性提升,开启 AI 深度认知新阶段;V3 以更低成本实现 GPT-4 级性能,推动 AI 技术平民化
2025	DeepSeek R1 发布、阿里巴巴 Qwen 2.5-Max 发布、产业深化及监管完善	R1 在专业领域创造性价比新高,中国 AI 跻身全球第一梯队;AI 技术从通用向专业化转变,行业生态更趋成熟

这些进展不仅体现了 AIGC 技术的快速发展,也展示了我国在这一领域的创新能力和实践应用。作为未来的设计师和创意工作者,我们要积极拥抱这些新技术,同时保持对传统文化和人文价值的敏感,在技术与艺术的交融中创造出更多优秀作品。AIGC 的发展历程告诉我们,技术创新永无止境,而每一次突破都可能带来创作方式的革命性变化。

1.3 创意产业的数字化身：AIGC 的多元应用

AIGC 正以令人惊叹的速度和深度渗透到创意产业的各个领域,成为推动数字创意经济发展的新引擎。从文学创作到新闻报道,从视觉设计到游戏开发,从音乐制作到影视制作,AIGC 正在这些领域掀起一场颠覆性的“内容生产革命”。这场革命不仅改变了内容的创作方式,也正在重塑整个创意产业的生态。

1.3.1 AIGC 的核心优势与创作革新

AIGC 技术的核心优势在于其强大的生成能力和高效的创作流程。通过深度学习算法,AIGC 工具能够理解和模仿人类创作的模式,生成符合特定风格、主题或要求的内容。这不仅大幅提高了内容生产的效率,还为创作者提供了无限的灵感来源和创意可能性。想象一下,一个作家可以在几秒内生成数十个故事大纲,一个设计师可以瞬间创造出数百种 logo 方案(图 1-4),这就是 AIGC 带来的革命性变化。

在文学创作领域,AIGC 工具正在成为作家的得力助手。这些工具可以根据作者设定的人物、情节、背景等要素,自动生成段落内容,辅助作家进行写作。更高级的 AI 写作助手甚至能够提供情节建议、人物塑造和文风模仿等功能,极大地拓展了作家的创作空间。例如,一位科幻作家可能会使用 AIGC 工具生成未来世界的详细描述,或者探索不同的情节发展的可能性(图 1-5)。这不仅加快了写作速度,也为作品注入了更多的创意元素。

1.3.2 AIGC 在新闻和内容生产中的应用

新闻和内容生产领域也正在经历 AIGC 带来的深刻变革。大语言模型已经能够根据关键信息生成基本的新闻稿件,甚至可以模仿特定的写作风格。这使得记者可以将更多



图 1-4 Midjourney 生成式工具进行 logo 设计

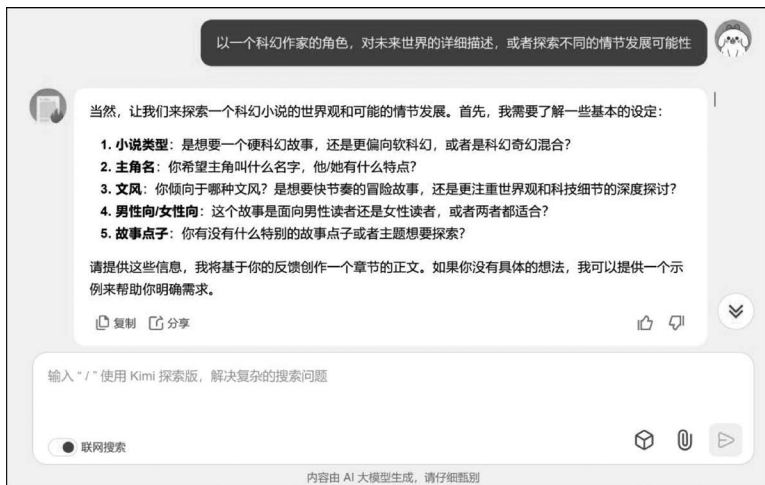


图 1-5 Kimi 生成式 AI 工具进行文本创作

精力投入深度报道和调查性新闻中,而将日常新闻的撰写交给 AI 助手。同时,AIGC 在内容摘要、多语言翻译和个性化内容推荐等方面的应用,正在重塑整个媒体行业的内容生产和分发流程。想象一下,一篇新闻报道可以瞬间被翻译成数十种语言,并根据不同读者的兴趣进行个性化呈现,这就是 AIGC 为媒体行业带来的变革。

1.3.3 AIGC 在视觉设计和游戏开发中的创新应用

在视觉设计领域,AI 图像生成工具正在成为设计师的“创意助手”。这些工具能够根据文本描述生成高质量的图像,为设计师提供丰富的创意素材和灵感来源。设计师可以使用这些工具进行快速的概念验证和原型设计,大幅缩短了设计周期。此外,AI 还在平面设计、UI/UX 设计、品牌标识设计等具体应用场景中发挥着越来越重要的作用(图 1-6)。例如,一个品牌可以使用 AIGC 工具快速生成数百个 logo 方案,然后由设计师进行筛选和优化,这大幅提高了设计的效率和创意多样性。



图 1-6 即时 AI 生成式工具进行 UI 设计

游戏开发是 AIGC 应用最为广泛和深入的领域之一。从游戏场景生成到 NPC(非玩家角色)对话系统,从角色动画到背景音乐,AIGC 几乎触及了游戏开发的每个环节。这不仅大幅提高了游戏开发的效率,还为创造更加丰富、动态和个性化的游戏体验提供了可能。想象一个开放世界游戏,其中的每个 NPC 都有独特的性格和背景故事,每个场景都是动态生成的,这样的游戏体验将会无比丰富和真实。

1.3.4 AIGC 在音乐和影视制作中的革新

在音乐领域,AI 作曲工具正在改变传统的音乐创作模式。这些工具能够学习特定风格的音乐,并创作出类似风格的乐曲。一些流行音乐人已经开始尝试将 AIGC 融入音乐创作过程,为专辑注入新鲜灵感。此外,AI 还在音乐制作、混音、编曲等技术层面发挥着越来越重要的作用。例如,一个音乐制作人可能会使用 AI 工具生成数十种不同的编曲版本,然后从中选择最佳的一个进行进一步优化。

影视制作领域也正在经历 AIGC 带来的变革。从剧本创作到视觉特效,从配音配乐到后期制作,AI 正在重塑整个影视制作流程。特别是在虚拟人物创建、场景生成和视频编辑等方面,AIGC 显示出巨大的潜力。想象一部电影的每个场景都可以根据导演的描述自动生成,每个群众演员都是 AI 创造的虚拟角色,这将大幅降低制作成本,同时提高创作的自由度。

1.3.5 AIGC 的多模态整合与未来展望

AIGC 技术进步的另一个重要方向是多模态整合。通过整合文本、图像、音频等多种模态的输入,AIGC 系统能够生成更加复杂、富有创意的内容。这为创意产业带来了全新的可能性,例如仅凭一段文字描述就能自动生成完整的短视频或音乐短片(MV)。这种多模态整合不仅提高了创作效率,也开启了全新的创意表达方式。



然而,需要指出的是,AIGC 在各领域的应用还处于发展阶段。尽管在某些特定任务上已经能够产出高质量内容,但在整体质量、创意性和一致性上仍有提升空间。AIGC 要真正赋能创意产业,还需要在算法、数据、应用场景等方面持续突破。同时,我们需要警惕 AIGC 可能带来的一些问题,如版权纠纷、内容真实性验证、人工智能伦理等。

AIGC 的发展也对创意产业从业者提出了新的要求。如何有效利用 AIGC 工具提升创作效率?如何在保持人类创意主导地位的同时充分发挥 AI 的优势?如何在快速变化的技术环境中保持竞争力?这些都是从业者们需要深入思考的问题。未来的创意工作者可能需要成为“人机协作”的专家,既要精通传统的创意技巧,又要熟悉各种 AI 工具的使用(表 1-2)。

表 1-2 AIGC 在各创意领域的多元应用

应用领域	代表性 AIGC 工具/技术	主要功能	应用案例
文学创作	GPT-4、Claude、通义千问	内容生成,情节构建,人物塑造	辅助小说创作,自动生成新闻稿
视觉设计	Midjourney、Stable Diffusion、即梦、秒画	图像生成,风格转换,图像编辑	概念设计,品牌视觉识别,UI/UX 设计
游戏开发	NVIDIA GauGAN3、网易伏羲、有灵 AI	场景生成,NPC 对话系统,角色动画	自动生成游戏关卡,动态 NPC 行为
音乐创作	Suno、Udio、乐盒 AI	作曲,编曲,音频生成	AI 辅助作曲,自动生成背景音乐
影视制作	Runway、Gen-3、Stable video diffusion、svd、pika、可灵 AI、清影 AI、通义万相	视频生成,特效制作,虚拟人物创建	AI 短视频制作,虚拟主播
多模态内容	GPT-4o、智谱清言等	跨模态内容理解与生成	图文互动创作,智能内容分析

1.4 AIGC 平台大比拼：寻找最佳创意伙伴

2024 年,AIGC 技术的飞速发展推动全球 AIGC 平台市场进入一个新的竞争阶段。从文本、图像、音频到视频生成,各个领域都呈现出激烈的创新角逐态势,展现了人工智能在创意领域的巨大潜力和广阔应用前景。这场技术革命不仅改变了内容创作的方式,也正在重塑整个创意产业的生态系统。

1.4.1 文本生成领域的竞争格局

在文本生成领域,与先进的国际模型相比,中国本土模型的实力差距正在迅速缩小,呈现出百家争鸣的局面。以 OpenAI 的 GPT 系列和 Anthropic 的 Claude 为代表的模型在多语言处理和复杂任务解决方面仍保持领先,特别是在科研写作和跨学科分析上表现突出。这些模型展现了强大的语言理解和生成能力,能够处理各种复杂的语言任务,从创意写作到技术文档,从市场分析到学术论文,无所不包。



与此同时,中国的大语言模型如 DeepSeek(深度求索)公司的 DeepSeek 大模型、科大讯飞的讯飞星火、智谱的 GLM 系列和百度的文心大模型等产品在中文处理能力上已经取得了显著进步(图 1-7~图 1-10)。这些本土模型在理解中国文化语境、处理中文特有表达方式等方面显示出独特优势,反映了 AI 技术在适应不同文化和语言环境方面的进展。例如,在古诗词创作、成语运用、文言文翻译等中国特色文学形式上,本土模型展现出了明显的优势。



图 1-7 DeepSeek(深度求索)公司的 DeepSeek 大模型



图 1-8 科大讯飞的讯飞星火

这种竞争格局不仅推动了技术的快速迭代,也为用户提供了更多样化的选择。对于创意工作者来说,这意味着他们可以根据具体需求选择最适合的工具,无论是需要处理国际化内容还是深耕本土市场。例如,一个需要创作跨文化营销文案的广告公司可能会选