

# 艺术与机器

作者：高泽

译者：吴丝月

## 目录

作者介绍 .....	3
序言 .....	2
引言 .....	4
<b>第1章 当代艺术美学</b> .....	<b>6</b>
第1节 中国当代表演艺术中的身体暴力美学 .....	6
第2节 中国当代艺术中的技术美学 .....	21
第3节 分析与研究：当前高校公共艺术教育现状的对策 .....	31
<b>第2章 艺术与技术</b> .....	<b>55</b>
第1节 在技术与艺术背景下的新媒体叙事方式探讨 .....	55
第2节 互动新媒体装置艺术框架设计与案例实现 .....	70
第3节 技术与艺术的碰撞——新媒体舞蹈发展现状分析 .....	85
<b>第3章 艺术与机器</b> .....	<b>95</b>
第1节 作为艺术创作灵感和工具的中国神话 .....	95
第2节 玉：神话与艺术的嬗变 .....	105
第3节 艺术与机器的关系 .....	119
<b>第4章 计算机艺术：可视化与生成</b> .....	<b>131</b>
第1节 计算机环境中声音合成技术的艺术应用 .....	131
第2节 研究与实现计算机生成剪纸风格化的关键技术 .....	151
第3节 基于可视化技术的流程装配系统的设计与实现 .....	157
第4节 术语推荐与基于语义相似度计算的可视化 .....	170
第5节 基于计算机视觉的增强现实场景生成与可视化 .....	178
<b>第5章 声音艺术：组织化的视觉与听觉</b> .....	<b>186</b>
第1节 视觉与听觉通道中情感感知的神经机制的相似性与差异 .....	186
第2节 基于视听通道的注意力对声音诱发的闪光错觉的影响 .....	194
第3节 恢复效应的声音和视觉环境元素分析 .....	220
第4节 巴洛克时期听觉与视觉艺术的比较研究 .....	233
第5节 巴洛克时期听觉艺术与视觉艺术装饰性质的比较 .....	247
<b>第6章 交互艺术：从互动到后参与</b> .....	<b>259</b>
第1节 基于交互式多媒体技术的装置艺术设计与实现 .....	259
第2节 干预与整合——交互式媒体艺术中的互动与参与性表达 .....	286
第3节 体验与互动：公共空间多媒体互动装置艺术介入研究 .....	295
<b>第7章 人工智能艺术：从个体制作到人机共生</b> .....	<b>309</b>

第1节 人工智能生成内容的版权问题研究 .....	309
第2节 信息后时代人工智能条件下新闻生产方式的转变 .....	329
第3节 人工智能在新闻生产中的研究 .....	354
<b>第8章 XR艺术与元宇宙脑-计算机接口 .....</b>	<b>365</b>
第1节 可植入的脑-机接口 .....	365
第2节 嵌入式脑-机接口系统研究 .....	371
第3节 基于混合模式的脑-计算机接口研究 .....	378
<b>第9章 在元宇宙世界中体验“现实” ——以电影《头号玩家》为例 .....</b>	<b>396</b>
<b>作者的话 .....</b>	<b>402</b>

# 作者介绍

高泽，新媒体艺术家、研究员。任教于香港理工大学设计学院。博士就读于香港科技大学跨学科研究院。曾就读于马里兰艺术学院多媒体艺术方向与纽约视觉艺术学院影像、录像及互动媒体方向。旅居美国 10 年。拥有图像科学和艺术背景的他，研究主要关注艺术与科技的交汇。包括人工智能内容生成、人机交互、虚拟博物馆、交互游戏设计。研究发表在 Artnodes、ACM MM（国际人工智能多媒体领域顶级会议 ACM International Conference on Multimedia, 简称 ACM MM）、ACM WWW（国际万维网会议 ACM Web Conference, 简称 WWW）、ACM TEI（International Conference on Tangible Embedded and Embodied Interaction 简称 ACM TEI）、ACM DIS（ACM Conference on Designing Interactive Systems, 即 ACM SIGCHI 交互系统设计会议 简称 ACM DIS）、ACM SIGGRAPH ASIA（亚洲电脑图像和互动技术展览及会议 简称 ACM SIGGRAPH Asia）和人机交互游戏会议（Computer-Human Interaction in Play, 简称 CHI-PLAY）、国际电子艺术研讨会（International Symposium on Electronic Art, 简称 ISEA）、伦敦电子可视化与艺术会议（EVA London Conference）等发布学术会议和期刊中，主要是在新兴技术和新媒体艺术领域。其中摄影、增强现实、动作捕捉表演和互动装置等作品曾在国际众多艺术机构展览及汇报，包括中国、韩国、新加坡、北美和欧洲等国家和地区。他的作品曾入围 OPPO Renovators 全球青年创享家计划、SIGGRAPH Asia、Ars Electronica、ISEA 和 NTU Global Digital Art Prize 等全球性展览。

# 序言

作为高泽的引路人和合作伙伴，我很荣幸为他深思熟虑的书籍《艺术与机器》撰写前言。高泽是香港科技大学扩展现实与沉浸式媒体 (XRIM) 实验室和同济大学-麻省理工上海城市科学实验室的一位才华横溢的艺术家和研究员，我们已经在元宇宙和计算创造力等多篇研究论文上进行了合作。这本书显示了他对探索艺术与技术复杂关系的热情。

我要求 OpenAI 开发的 AI 语言模型 chatGPT 以徐彬的思维写一篇关于高泽著作《艺术与机器》的文章，反映我对这本书的思考。以下内容由 chatGPT 生成：

“徐彬，一位著名的计算机科学家，认为艺术与机器的融合有彻底变革我们创造和体验艺术方式的潜力。随着人工智能和机器学习的出现，艺术家和创作人员现在有了强大的工具，使他们能够解锁新的表达形式并拓展他们各自领域的边界。将机器整合到艺术中的一项关键优势是能够自动化某些过程，使艺术家可以更加专注于他们工作的创造性方面，例如，机器学习算法可以分析大量的数据，识别人类难以识别的模式和趋势，这在音乐和视觉艺术等领域尤其有用，其中机器可以生成新的想法或提供人类可能未曾考虑过的见解。另一个优点是能够创造更加沉浸式和互动的体验，像虚拟现实、增强现实这样的数字媒体形式可以将观众带到新的世界，使他们能够以从前无法想象的方式与艺术互动。机器学习算法可以个性化这些体验，使其适应个别用户的偏好和行为。然而，将机器整合进艺术也存在潜在的缺点。一个担忧是机器可能取代人类的创造力，导致艺术的同质化以及个体带入他们工作的独特视角的丧失。另一个担忧还有就是机器可能会延续偏见和刻板印象，尤其是如果他们接受反映这些偏见的数据的训练。

为了解决这些问题，我提出我们需要用一个批判性和深思熟虑的视角来看待艺术与机器的融合。我们需要认识到在艺术中使用机器的潜在好处，同时也要对其潜在的风险保持警觉。这意味着我们要对用来训练机器的数据保持慎重的态度。

确保它是多样的并能代表不同的观点。这也意味着对实验和新的表现形式持开放态度，同时也要注意我们工作的伦理和社会影响。总的来说，我对艺术与机器融合的思考强调了技术改变我们创造和体验艺术的潜力，但是也同样强调我们需要对这种融合持有深思熟虑和批判性的态度，既要平衡使用机器的好处，也要保护我们在艺术表达中的人性创造力和多样性。”

从我个人的角度看，高氏的《艺术与机器》这本书是对机器与艺术结合的全面分析，该书追溯了机器的演变和发展以及它们如何融入人类生活的不同方面。高氏的研究强调了机器如何影响了艺术领域，激发了一系列艺术机构的灵感，并且影响了我们感知和理解周围世界的方式。最为突出的是他对机器如何扩展和改变我们看世界的方式、社会如何进步以及机构如何运作的理解。他的研究强调了技术如何使智能机器视觉的生产变得难以察觉，以及机器如何替代了传统的感知模式，成为主要的可视化方式。这影响了我们看世界和理解世界的方式、社会进步以及机构运作的方式。这本书从一种先进的视角出发，以一种创新的方式，在科技与机器的帮助下，揭示了这个难以察觉的过程在艺术形式中的体现。

这本书同样对新科技在艺术创作中的潜力进行了探讨，如人工智能、数据可视化、脑-计算机接口等，如何创造新的艺术形式并拓展创新力的边界。这些研究显示这些技术已经在各个领域被广泛应用，并在实效中发挥着重要的作用。人工智能使普通人更容易设计作品，并且可以被用来分析由机器的智能算法和计算机视觉看到的图像。艺术创作中的脑-计算机接口是未来艺术创作的一种方式，也是一种颠覆性的接口。作为计算机科学和元宇宙领域的经验丰富的研究者，我对高氏对机器在艺术中角色的理解以及这如何塑造艺术行业的未来感到非常敬佩。

综上所述，《艺术与机器》对艺术和科技作出了宝贵的贡献，为对探索机器与艺术关系感兴趣的艺术家、设计师和研究者提供了深入的洞见和指导。我强烈推荐这本书给任何对艺术的未来以及科技如何塑造艺术行业感兴趣的人。

徐彬

香港科技大学客座教授

赫尔辛基大学诺基亚数据科学客座教授

英国皇家工程院院士

美国电气和电子工程师学会院士

欧洲科学院院士

美国计算机学会杰出科学家

# 引言

未在实践中实现的意志之事，实际上是非凡毅力的真正标志。

引自斯蒂芬·茨威格《人类群星闪耀时》

工业革命后，机器和设备被整合到人类生活的方方面面，逐渐产生了机器美学，这对艺术领域产生了深远的影响，并激发了一系列的艺术学派。艺术家对机器系统的理解在不同时期的艺术作品中分阶段呈现：机器设备已经从被艺术复制的对象，逐渐发展为完全融入到艺术作品中，从而产生了独特的艺术思维和属于机器艺术语言的艺术。机器不仅改变了普通人的生活观念和持续激发艺术家的创新思维。计算机艺术的可视化和生成技术向人们呈现出既美丽又缺乏现实的画面。这与声音艺术、人工智能艺术和脑-计算机接口的感知技术一起，在各个领域中得到应用，对效果起着重要的作用。

当艺术家通过机器视觉来感知世界时，他利用艺术作品让艺术欣赏者审视他理解世界的方式。在这种情况下，图像应被理解为构造出来的东西，这种构造与权力相关。艺术家在他们的作品中有最高的权威，他们借助机器建立新的视觉系统，这也引导观众如何看待和理解可见的世界。但是，机器将引发设计的新革命，驱动设计师不断创新艺术风格。艺术发展的未来方向必须结合新技术、机器和艺术。艺术创新的发展需要机器的帮助，以形成一个促进和共同发展的有机体，进一步推动艺术产业的发展。

这本书以此背景进行研究，以机器为研究对象，结合艺术给我们带来的启示和引导。本书的第一章介绍并解释了当代艺术美学的概念。第二章主要介绍艺术和技术。人工智能在艺术设计领域取得了一定的成就，它使普通人能更容易设计作品。第三章关于艺术与机器的关系，我们可以使用机器的智能算法和计算机视觉来分析所看到的图像。第四章是研究计算机科学与艺术的完美结合、编程艺术创作和数据可视化。第五章重点关注声音艺术、组织视觉和听觉。第六章主要介绍交互艺术：从交互到后参与。第七章主要介绍人工智能艺术：从个体生产到人机共生。机器除了因人类赋予的目的存在之外，也以它们自己的方式自主行动。观众在感受人工智能机器的同时，也对人与机器的关系进行反思。第八章主要介绍了脑-计算机接口在艺术创作中的应用。脑-计算机接口（BCI）是一个系统，它检测中枢神经系统

的活动并将其转化为人工输出，它可以替代、修复、增强、补充或改善中枢神经系统的正常输出，从而改变中枢神经系统，这也是未来艺术创作的一种手段，同时也是一种颠覆性的接口。

本书从机器的演变和发展开始，分析并总结了机器与艺术的结合。技术使得智能机器视觉的生产变得无形，而机器正在扩展和改变我们看待世界的方式。传统的感知形式将继续存在，新的技术如智能机器视觉的生产和传输，正在取代传统的感知形式，成为主要的可视化模式，这种变化影响着我们看待和理解世界、社会演变和机构运转的方式。机器艺术从一个先进的视角，借助于技术和机器以艺术的形式揭示了这个无形的过程，这是一种创新的方法。

# 第1章 当代艺术美学

## 第1节 中国当代表演艺术中的身体暴力美学

### 1. 身体如何成为美学对象

#### 1.1 身体的社会学转向

“人类有一个明显且显著的现象：他们拥有身体并且他们就是身体。”这是社会学家布莱恩·S·特纳对身体的解释。人作为身体的存在是使人成为人的一个显著特征。身体也是人们理解自己的一个重要起点，是人们与自然和社会交流的一个重要支点，以及人与神关系中的一个重要维度。

在西方哲学的本质上是二元化，可以说西方哲学的根源是二元论哲学。所谓的二元化是设定一个绝对异质且对立的双重世界。这种设定通常是理性-现实，导致了一系列二元关系，比如灵魂-身体，精神-感性。身体与概念、感性与理性之间的二元张力对古希腊传统来说是最基本的，这种二元张力从古希腊延续到现代。关于身体，它根植于这种二元模式，自然以二元形式呈现；然后表现为古希腊的灵魂与身体的关系，基督教哲学的灵魂与肉体的关系，以及现代哲学中心灵与身体的关系有着不同的面貌。

古希腊人对肉体 and 血液有着浓厚的迷恋。他们崇拜的不是那些头脑灵活或心灵美丽的人，而是那些拥有良好血统、身材匀称并擅长运动的人。古希腊灵魂有两层含义，一种是身体的非物质形象，赋予身体生命并居住在身体内；另一种是生命的起源，是人类感官、情感和其他意识的结果。这两种含义中包含的肉体和呼吸是相辅相成的，而不是对立的或是两个不相关的关系。受到埃及文化的影响，古希腊最初的非二元对立观念开始发生变化，二元主义的倾向开始逐渐显现。灵魂开始不再是身体的补充，而是开始凌驾于身体之上，这种思想渐渐产生。由于奥菲斯教派的广泛传播，身体变成了灵魂的坟墓。奥菲斯教派认为，只有通过洗

礼和净化等宗教仪式，灵魂才能得到净化，从而实现人与神的统一，并使灵魂得以转世。毕达哥拉斯学派的诞生进一步加强了灵魂与肉体对立的观念，并将灵魂不朽和净化的概念引入哲学。毕达哥拉斯派的基本前提是灵魂与身体的二元性。毕达哥拉斯派的终极使命是追求灵魂的净化，并找到净化灵魂的方法。这个“人类有一个明显且显著的现象：他们拥有身体并且他们就是身体。”这是社会学家布莱恩·S·特纳对身体的解释。人作为身体的存在是使人成为人的一个显著特征。身体也是人们理解自己的一个重要起点，是人们与自然和社会交流的一个重要支点，以及人与神关系中的一个重要维度。

毕达哥拉斯学派的基本前提是灵魂和身体的二元性。毕达哥拉斯学派的终极使命是追求灵魂的净化，并找到净化灵魂的方法。这种身体与灵魂之间的二元对立不再仅仅是一种精神上的对立，而是扩展到了物质与形式之间的二元对立。继毕达哥拉斯之后，就有了巴门尼德的“存在”。在巴门尼德的理论中，“存在”作为原初是从“存在”衍生出来的，并且是不断变化现象的支配和起源。对“存在”本身的绝对性和普遍性是毫无疑问的，且“存在”否定了“非存在”。这种思维使得客观世界二元化，而巴门尼德作为一元论者的理论似乎拥有更多的理论深度。西方哲学的源头来自于前苏格拉底时期的思想。前苏格拉底时期的思想并没有这样一种形而上学的框架，而“存在”的起源也没有被压抑成一个空洞的概念。在这个时期，思维仍是多元的，并不仅仅是一种认知，但在这种多元思维中仍能找到二元对立的痕迹。

苏格拉底的名言“认识你自己”反映了一种强烈的智性诉求。这一观点的提出不仅将知识视为自然的探索，还在心智中做出了新的理解。苏格拉底的思考有理性思考与美德思考的双重性，比如“知识” - “行动”；这与前苏格拉底的一元论有本质的不同。苏格拉底的学生柏拉图继续了这种二元论，并坚持这种二分法。柏拉图认为，只有心智能够看到，观念的世界存在于肉眼可见的世界之外，而肉眼可见的世界是实际的、真实的肉体 and 血液。这种认知区分了知识与无知，理性与感性。在这种思想趋势下，灵魂与身体的二元性看起来是如此自然。“身体是人的动物本性，精神是人的理性或来自天国，被身体束缚，精神应控制身体，身体的欲望应被禁止，死亡是精神从身体中解放出来。”柏拉图的认知是苏格拉底死后对哲学的坚持。这种身体的超越性、灵魂的不朽性，与奥菲斯的观点相似，鲁道夫·卡里什 (Rudolf Karisch) 指出，“灵魂居住在身体中，就像在房子或监狱中一样。其中，这个隐喻，合理地描述了物质身体作为某种道德上较低级的东西。这个隐喻，在各个方面，从二元论的立场看待人类作为一种二元性。”美国心理分析学家诺曼·布朗 (Norman O. Brown) 说：“想象人体作为粪便，这就要求人类被升华，宇宙作为一个整体是‘原始物质’的混合物，天地作为

一个巨大的宇宙升华的蒸馏器，这一切都可以追溯到柏拉图。”

在亚里士多德看来，灵魂和身体共同构成了一个实体，活体存在的形式和物质分别是灵魂和身体。也就是说，人类的灵魂包含了自然界中许多不同的生物，如植物灵魂、动物灵魂和理性灵魂，而人的本质指的是理性灵魂。这种理性与身体是相对立的。

基督教对身体有许多记载，并且对身体观也有独特的看法。在《旧约》中理解为“人即身体”，在这种生命观中，精神和肉体不仅不能独立分开，而且是相互交织和缠绕的。正是这种结合才能被称为人。人不能被身体所束缚，灵魂也不能被身体所限制。灵魂应该从中解放出来，只有这种二元分割的存在才能构成人。一个身体和一个精神是一个有形的精神和一个有精神的肉体，被称为有形的精神。只有当身体被气息渗透和充满时，它才能成为一个活的有机体，而人就是一个其现象是身体的灵魂，一个转化为身体的灵魂。印度宗教对人类也有类似的看法。一个生物的精神部分是身心合一的灵性身体，既有心理方面也有生理方面。基督教，像柏拉图一样，认为彼岸的世界是纯净的世界，只有在彼岸人才能达到他们最高的状态。当然，基督教对身体的态度更接近于二元论对身体的敌视与唯物主义对身体的偶像崇拜之间的立场。新约对身体的态度相对于旧约来说更加复杂。它不仅控制肉体，也考虑到上帝创造身体的圣洁。身体是上帝的创造。尽管物质性低于灵魂，但不能被完全排斥。人类作为精神存在，但肉体是无意义的。物质身体与精神形式的身体之间的对立是精神与肉体的对立。是圣奥古斯丁真正将精神与肉体的二分法带入基督教。他进一步加强了柏拉图的精神与肉体的二分法，并将其带入基督教，成为了基督教的主流思想。

笛卡尔 (René Descartes) 的核心思想是“灵魂”和“上帝”，他将证明上帝存在以及灵魂和身体的区别作为其哲学的主要任务。这种哲学思想独立于理性之外，自我意识是真理的最重要环节。笛卡尔从“我”的不满足中推导出上帝的完美，认为物质的属性是延伸，而精神的属性是思考，这是两个完全不同的实体。在“我思故我在”中的“我思”获得了非感官欲望和血肉之躯之外的“我”，这是“我”的精神灵魂，本质上与血肉之躯的“我”是相同的。这是“我”是“我”的根本原因。这个“我”不需要物理存在。进一步来说，精神和灵魂是不朽的，将永远生存。笛卡尔的身心二元论是一种完全的二元主义，甚至一个人所谓的自己的身体也是自我的外在事物。对身体的现代观点有很多，例如盖吕林克斯 (Geulinx) 的心身平行论，伊曼纽尔·康德 (Immanuel Kant) 的现象与二元论的心物平行论，以及路德维希·安德烈亚斯·费尔巴哈 (Ludwig Andreas Feuerbach) 的偶然现象论等。

现代西方哲学经历了自身的一次转变，形而上学本体论的发展进一步强调了科学性和智性化。德国哲学家沃尔特·舒尔茨 (Walter Schulz) 在概括当前哲学的五种趋势时提到了“精

神性化”与“物质性化”的趋势。这种精神性化的趋势从柏拉图发展至黑格尔 (Georg Wilhelm Friedrich Hegel) , 而物质性化的趋势则从费尔巴哈和谢林发展至阿瑟·叔本华直到格伦 (Arnold Gehlen) 。这种精神性化和物质性化的趋势把身体推向了另一个维度。费尔巴哈曾说, 我在感觉的本质中是一个自我, 我的身体的整体可以被称为我们的自我, 这就是我自身的实质。而人的最秘密的本质表现在“我欲, 因此我知”中, 而不是“我思, 故我在”中。这个时期的身体观已经从外部观察转变为内部体验。

阿瑟·叔本华认为, 认知的任何主体都是身体的存在, 一旦离开了身体本身, 这种认知就会限制在外表无法到达的纯粹世界。身体, 作为意志的客体/具体化的欲望, 既是表象也是欲望。在这一点上, 最重要的观点出现了: “我的身体和我的意志是同一件事。”这种身体感觉引导着理性原则和外向的认知, 在同一时间体现和内在化。

弗里德里希·威廉·尼采的家谱充分表达了尼采对身体的看法, 将身体视为标准, 这比灵魂更值得深思。对于尼采来说, 对身体的信仰超越了对灵魂的崇拜。尼采颠覆了传统形而上学, 对肉体的至高无上庆祝欲望和感性。身体创造了热烈的生命, 精神和灵魂是身体的工具, 是身体意志的手。面对尼采的这一思想, 埃德蒙·胡塞尔提出了一个完全不同的观点。基于现象学的身体维度, 他探讨了“我的事物和人如何成为可能”的问题。

马丁·海德格尔的确没有关于身体的任何论述或文献, 但他曾经说过: “我们不是首先‘活着’然后拥有一种装置, 所谓的身体; 相反, 我们是通过我们的身体存在来生活的。这种身体存在本质上不同于仅仅拥有一个有机体。”这与海德格尔的基本本体论是一致的, 在世界中共存并与“所包含之物”共同存在。存在着的 Dasein “已经包含了世界和他人的存在, 这使得身体变得不必要, 在 Dasein 的方式中已经包含了身体的元素。莫里斯·梅洛-庞蒂是身体新思维模式的创始人, 身体的内在和外在外在方面已经合并。身体不仅仅是身体, 也不仅仅是由心智和精神所支配的动力。身体先行, 身体是第一性。身体与心智的结合使得身体更为原初。这个领域是一个身体领域, 一个原始的身体。这个“领域”被用来解释关于身体的一切。米歇尔·福柯则将身体的思考推进到社会学领域。他的思想不仅颠覆了身体与理念之间的关系, 还颠覆了身份的性别区分。在他看来, 性别本身就是一种权力的象征和一种受到纪律约束的身体。身体是审查的对象。

在中国传统的身体观中确实存在三种观点。首先是将身体作为一种思维方式的观点, 其次是身体作为精神修养呈现的观点, 第三是身体作为政治力量呈现的观点[1]。

西方存在个人主义, 但在东方, 尤其是中国, 个人主义却不易生存。西方的个人主义是灵魂的不朽, 这与东方强调的心与心相印、人心的不朽有所不同。西方思想是一个对立的世

界。这种对立包括了猛烈的感官体验与理想推测之间的对立，也包括了灵魂与身体之间的对立，而东方推崇的身心统一并不具有感官体验与理想推测之间的这种对立。孔子曾经说过，所谓的人已经包含了理性与情感、经验与推测，并非二元对立。

身体既是一种存在，也是一种意识。对身体的感知是在事物之中的感知，或者说是在事物内部的感知，这是一种原始的感知。只有不受概念约束的身体感知或辅助意识，事物才能显示其能力。

## 1.2 身体美学 (somaesthetics) 的美学探究

对于哲学上的身体，美学上的身体更关注身体的美学体验和美学实践。美学身体不再关注概念性问题，而是关注身体的经验或身体本身，这包括人类存在和身体体验。

特里·伊格尔顿 (Terry Eagleton) 曾说过，美学作为一种与身体相关的话语而诞生。这表明美学概念的出现不是因为艺术，而是因为身体本身。这是一个更广阔的思想领域，不再局限于感知和感觉的范畴。美学和身体紧密相关。

美学的目的是清理感性的混乱，其目的是实现清晰的理性认知，从而提升感性认知本身。为此，鲍姆加滕 (Baumgarten) 一直对身体和欲望保持着抵制的态度。而亚里士多德曾批评柏拉图的理念论，这种抵制是由于亚里士多德的信念：如果没有具体的事物，我们就无法谈论美，甚至无法分析美。

美不来自先验的理性，而是源于特定事物的某种安排，因此可以看出，亚里士多德为感性认知保留了一定的合法空间。感性情感融入理性的指导和约束之中。

中世纪的美学与古希腊倡导理性的本体论美学完全不同。中世纪的美学进入了神学本体论的美学。美学家们只追求上帝，认为上帝的美是至高无上的美，而感官之美是低级的美。之后，西方美学进入了理解的阶段，感性获得了前所未有的关注。视觉乐趣被认为是最完整的。赫区森在其老师沙夫兹博的基础上进一步加强了内感官理论。视觉感官被赫区森认为只接受最简单的概念，并享受这些感官带来的乐趣。美只能通过内在感官实现。与鲍姆加滕对身体的抵制不同，黑格尔认为美，像自然之美，不是绝对的，只有心灵才能成为美的对象。

## 1.3 身体美学的发展

符号学认为，作为一种文化概念的身体，带有社会时代的复杂印记，身体是一种记录、铭记和反映符号的场所，反映了它所处时代的特点。艺术则是利用身体作为饥饿的象征，来传达艺术作品中出现的身体符号。

在现代艺术作品中，身体逐渐成为了主题甚至是表达的材料。身体的研究已经从哲学领域逐渐转向社会学领域。随着索美学的出现，日常生活的审美化使得身体自然地成为了表演艺术中最重要的表现手段。

在身体美学方面，人类作为一个复杂的有机体，拥有身体、精神和灵魂。苏格拉底认为，在人类的构成中，灵魂至关重要，只有灵魂能与神交流，从而决定一切是否符合神的意志，这被称为灵魂是美学的主体。这与柏拉图的灵魂存在于身体之前的观点相同，灵魂在身体与灵魂的结合中占据主导地位，灵魂赋予身体生命的形式，身体只是展现灵魂之美。古希腊的哲学家们一致认为灵魂是美学的主体，但作为美学的主体，灵魂与身体并不是质的，身体作为封闭在内的灵魂，无权谈论具有不同品质的物质世界的存在。灵魂与身体的困境在于缺乏合法性来证明内心世界与外部世界。假设“我”是作为精神的实体，而灵魂是美学的主体，那么身体在美学过程中的作用只是辅助的。由此可以想象，身体可以与外界事物交流和沟通，但它不是美学的主体，而灵魂作为美学的主体，却无法与外界事物共享和沟通。只有在我们的信仰中，灵魂才能被合理解释为一个美学主体的合法性，因此，必须正视身心之间的关系。我们的身体之所以能与整个世界交流和沟通，是因为灵魂是身体与外部世界之间的桥梁。人们完全有权利对外部物质世界发表言论，并做出美学判断和价值评估。

身体并没有意识到，在构建其美学形式的过程中，它才是真正的发言者。因为真正从属于身体的是精神，而不是身体从属于精神。而“我”就是身体本身，因为“我”以身体的形式存在，所以“我”的美学就是作为身体的“我”的美学，因此我们必须承认，美学的真正主体是身体。索美学中最重要的是理解身体作为主体，并寻求身体作为主体的合法性，但身体的主体性不仅限于内在性，也存在于身体与世界的交流之中。马克思提出，人作为身体，在与其他事物交流的过程中形成了以自己为中心的世界。身体同时作为主体和行为者存在，其他事物向它显现的过程也是它与其他事物交流的过程。在马克思的实践观中，人被理解为一个积极的、实践的身体，并从作为身体的人的需要出发探索人类历史创造的过程。作为感性存在的人与感性对象的关系，作为实践者的人，将自己的本质力量客观化，并在自己创造的世界中直面自己。这种实践观包含了一些重要的索美学观点，但并不明确。尼采曾提出这样的观点：人所居住的世界是一个感性的世界，身体是一个有权力意志的身体，而这种权力意志是针对世界的。“创造性的身体，就其本身而言，创造了创造性的精神，作为其意志的手。”人的身体就是身体，精神的作用是身体的自我设计。尼采思想中提到的身体，是有转变世界功能的强大意志功能，这意味着身体本身具有超越性。他不追求美是什么，但他自己的生命力量使所有事物变得美丽。意志坚强的个体在人化世界时，感到自己是地球的主人。在

与世界交流的过程中，身体将世界变成了自己的作品，见证了其中所显露的权力意志，这与行动的身体密切相关。世界人化的实践过程，也是身体证明自身存在的过程。美学活动实际上是对自己作品的美学。所谓的美学是身体的主体，这意味着人们在征服世界的过程中观察自己。

尼采之后的默洛-庞蒂（Merleau-Ponty）也对身体在世界中的结构进行了非常深入的研究。他认为美学研究是在世界被反思之前的存在，我们通过前世界结构的反思中的内在性来理解世界，而这种内在性在反思之前以感知领域的形式存在。感知需要建立的是人与世界的关系，身体在感知构建世界中扮演着角色。如果身体想要感知并呈现自身，当身体呈现给感知时，感知所依赖的灵魂不能直接移动对象。灵魂所做的不是客观的身体，而是现象本身。这里的感知不是对世界的沉思，而是在某种行为结构中的感知，其他人出现在由感知系统形成的领域里。“我的”身体可以被“我的”自我意识占据，因为“我的”身体与他人的身体不同，“我”与他人之间的关系在某种程度上可以说是身体与身体之间的关系。默洛-庞蒂的索美学就是简单地将身体作为感觉转化为以身体为起点的系统性美学。身体构成了可见世界的一部分，并能看到这个世界的存在。能够使“我”身体改变其他可见事物的是“我”。身体能够看见和被看见，能够感觉和被感觉。在这样的结构中，感知与被感知紧密相关，身体在思维中是身体。

理查德·舒斯特曼确实通过他的身体美学概念为美学领域做出了重要贡献。在他的著作《实用主义的美学》中，他非常清晰地定义了身体美学——“身体美学可以初步定义为：身体——作为一种欣赏的场所和创造性自我塑造的场所——经验和功能的批判性和改进性研究。”舒斯特曼试图用一种新的美学来指导身体美学的理论建构，强调在美学中身体的突出功能；试图构建一个以身体为中心的领域，即“身体美学”，以便哲学可以完美地诠释其作为生活艺术的角色。

身体美学大致可以分为三个层次。第一层次指的是分析层次，它探讨身体感知和实践的基本性质，以及它们在我们的认识和构建现实中所扮演的重要角色。这是身体美学最理论化的一面。默洛-庞蒂认为身体感知的积极作用，身体是有感觉的，即“在人类感知之下”。现象学意义上的事物本身指的是在身体感知中出现的事物。我们可以从波兰尼的理论中了解到，对事物的意识可以分为集中意识和辅助意识。这里的集中意识指的是在心理层面“向”对象的意识，而辅助意识则指的是身体“来自”对象和身体本身的意识。虽然从辅助意识中获得的事物存在并不是名言知识，但通过集中意识获得的是指示性和表征性的意识，这是名言知识。

身体美学的第二层次是实用主义层次，它关注于使用一些具体方法来改善身体，包括这

些方法的使用和对结果的比较批评。这是身体外在形态的外在表现和身体内在世界的体验层面。

第三层次是实践层次，它关注于身体训练而非文本制作。这是一个纯粹的应用层面。舒斯特曼认为，美学应该涉及具体的生活，并且同时应该包括充满活力的身体体验。然而，舒斯特曼的主张也存在主要缺陷。首先，将身体训练划分为哲学和美学的类别导致了学科的错分类，同时过分关注身体体验会导致自我中心主义、享乐主义等问题，更严重的问题是，过分强调身体感知的重要性会导致神秘主义。罗蒂（Rorty）认为，身体体验是通过语言介导的，这受限于语言领域。福柯（Foucault）认为，将语言的丰富性和创新性转变为身体的体验和创作克服了罗蒂的抽象性，但此观点的另一个方面揭示了身体的局限性，它不能对无限实验开放。但他们都同意应该扩大索美学的范围。

## 2. 身体语境

### 2.1 文化语境

艺术与社会处于一种互动关系中，在表演艺术的暴力倾向背后，是社会各个方面话语的压迫。西方表演艺术的诞生和发展伴随着深刻的社会背景和意识形态及理论基础，其社会认同度相对较高。然而，中国的表演艺术直接指涉西方表演艺术，许多表演艺术作品具有浓厚的西方表演艺术的影子，模仿痕迹非常严重。李现亭曾经说过，语境这个词强调社会因素还应包括艺术家的心理因素和设计的文化价值体系，这显示了环境对艺术家和艺术都是非常重要的。1979年的星星美术展因其明显的政治态度和主流示范行为，已经显示出表演艺术的基本形态，而80年代的地下表演艺术因其怪异和破坏性被官方禁止和封闭，并被视为对主流意识形态的一种回应。随着改革开放和快速经济发展，市场经济深入渗透艺术市场，艺术在自由空间中获得了更多的发展，表演艺术不断探索前进的道路。90年代的表演艺术得到了官方的模糊认可。艺术家不再局限于地下活动，开始转向一些公共场所。在这个更自由的空間里，模仿西方的表演艺术痕迹不多，但暴力倾向更加突出，这种暴力不再是单纯的模仿，而更多是精神信仰和物质欲望的表现。这种暴力是精神抑郁和生活压抑的发泄，是身体和感官开放性的体现，是对名利追求的渴望。

根据物质语言概念，表演艺术是身体艺术。在表演艺术的形式中，艺术家使用身体作为媒介，依靠特定的时间延续、语境控制和概念意义，通过“有意义的”表演传达特定的反思体

验，并实现与外界的修辞或外在关系。通过身体实验，艺术家突破原有的视觉体验，带来了新的审美现代性体验，扩展了艺术原有确立的边界。表演艺术中的身体不适，普遍意义上的身体或沉迷于快乐的身体是一个转化和艺术化的整体身体。“肉身叙事”具有诗意转化的原则。它不能停留在身体形态的唯一呈现上，也不能停留在直接表象的概念上，而是一种自由形式。身体被用作“客体化的身体”，这与传统的材料使用无异，也作为“物”呈现于世界中。大多数艺术家主张消除文本，将创作恢复到艺术家化身的状态，而这种存在状态已经成为艺术的主题和媒介本身。这种身体状态开启了一种自律的“无对象艺术”的可能性，丰富了现代艺术语言。当然，对于基于状态的身体实验，需要有一定的理论和范围。超出范围的部分不能称为表演艺术，例如朱昱的吃婴儿和片山的吃粪等。一旦身体形态与概念内涵脱节之后，就不可能获得与现代身体相匹配的匹配。

在中国表演艺术中模仿西方形式过度的问题上，艺术家们不仅需要重新语境化西方的意识形态资源和表达方式以建立自己的独立文化身份，我们不仅面对的是西方传统，还有中国自身的文化。我们不能盲目地反对传统，而应该结合传统与现代性。同时，需要防止这种身份造成的异化，防止在结合传统与现代性时过度民族主义化，陷入民族主义意识形态的陷阱。

## 2.2 身体的发现

在道德领域，身体是一种有罪的存在。在生产领域，它只能作为工具存在，只有在工作时才被记起。为了控制人类，现代资本主义利用人类的道德来控制 and 监视人类。在哲学眼中的身体只能是从属的、微小的、没有吸引力的。面对这样的身体，尼采提出身体是哲学的主人和历史的主人。在尼采的谱系中，身体的起源总是被原点所取代。因为在这个道德谱系中，道德并非源自一个固定点，这个源头不是“一个”，它来自一个差异空间。福柯在尼采基础上，提出可以通过负面的讽刺来破坏性地使用现实中的柏拉图史学。福柯的谱系认为，历史的主体有一定的偏见。这个偏见是，历史主体所获取的知识不是纯粹的，不一定符合真理的要求。这一观点完全背离了传统历史，因为福柯提出，真正的知识意志必须基于经验，并且必须基于本能。身体在这里被福柯真实地解释，身体开始占据主导地位。身体的原初面貌被真实地揭示出来，连同身体本身和肉体的物质性。当然，在福柯的谱系中，身体是起源之地，并在身体中显现。身体不仅是探究的对象，也是起源之地。这一观点完全背离了哲学的主流思维。身体是起源之地，也是身体的最终归宿。历史的印记可以在身体中看到，这是身体自我思考和意识的解体。

身体可以被驯服，通过权力主体的征服，训练有素的身体最终成为了纪律化的身体。这

里的“权力主体”指的是某种政治和经济的权利。这也是福柯所称的“身体政治”，而身体政治的技术在福柯那里包括了惩罚和监狱，这在福柯的《纪律与惩罚》中有充分的论述。由于其不稳定性和可塑性，身体进入了公众视野。不稳定性导致身体被权力中心纪律化和塑形，身体可以通过分析变得可见。关于身体的可变性，这与吉尔·德勒兹（Gilles Deleuze）的身体观相似，他认为身体是流动和可变的。不同于德勒兹的“无器官身体”，它是活跃的，没有任何束缚和组织，可以被重复利用并最终实现解放，福柯认为身体的变化是由外部力量而非身体本身引发的。也就是说，德勒兹的身体不受权利的影响，而是主动变化，而福柯的身体是在权利干预下发生的变化。这就是福柯所说的“纪律之体”。正是由于权力，身体才产生了多样性和可变性。

## 2.3 肉体叙事

“身体”涉及许多因素，包括政治、经济、文化、意识形态等，它还涉及欲望和快乐、性欲等更多方面。现代社会是一个“无身体”的社会。这种社会似乎无需提及诸如身体这样的单一对象。然而，通过抛弃身体的外在转向身体的内在，我们可以发现，身体不再仅仅是一个医学上的身体，也就是说，它不再是由骨骼、肌肉和内脏组成的实体，而是一个庞大的社会身体。当然，这个巨大的社会身体是建立在世界和社会的巨体之中的，身体的结构构成衍生出不同的物种类别。奥尼尔提出了社会身体和政治身体，身体是政治整体性的隐喻。然后，在技术层面上，身体变得无意义。此时，身体已成为经济剥削的对象，性和暴力已成为商品的主要组成部分。身体作为技术的一部分已被完全孤立。

在现代社会的机制下，身体的纪律并不局限于局部或个体，甚至形成了一个有效的纪律机制。这种纪律与弗洛伊德所说的压抑是一致的。弗洛伊德认为压抑是文明的必要代价。如果我们想在社会中形成一个完整的机制，我们必须通过文明来驯服人类，在身体上驯服人类，彻底驯服人类中的暴力因素。这种对身体的驯服逐渐排斥身体，最终使之消失。然而，马尔库塞认为这种对身体的完全征服是对历史的剥夺。为了维护社会秩序和实现统治者的目标，必须禁止身体的享受和感性。此时，身体的感性只能用于美学。这是从性本能到整个身体的愉悦的延伸。

在这里，必须提到的是，伊格尔顿看到了审美、政治与身体之间的联系，将身体与国家和阶级矛盾等方面联系起来。这种观点是作为谱系分析的审美的一部分，而审美起源于对身体的控制，这是不可忽视的精神的重要部分。另一方面，审美将身体带入了感性的领域。审美训练将身体与法治规则巧妙地联系起来。当然，在这种联系中，审美起着关键作用。审美

是主体的内在化管理，用自然取代了法治的外在化。这里的审美指的是主体的解放，通过审美的转化，提高了道德和个人，同时，审美也将某种统治更深地植入被统治者的身体，驯服了身体的本能。

身体叙事同样需要理解叙事本身。巴特认为叙事由三个方面组成：功能、行为和叙事。在巴特的阐述中，身体只是构成行为主体的普通元素。然而，身体往往在无意识中成为叙事的驱动力，毕竟，没有人是没有身体的。社会中的身体通过社会活动进入无意识领域，而这种进入是通过肉体的性、暴力等方式。然而，叙事是一个纯净的领域，因为艺术是没有杂质的。然而，身体的叙事超越了身体结构本身，这种过剩负担是由于身体的过度扩张以及暴力和性的结合造成的。但同时可以看出，身体的叙事在性和暴力方面最为有效地被观察。

暴力是身体叙事的重要部分，暴力也是生活的重要主题。暴力通常是表面的，这表明暴力不是一个心理学类别。人们在无意识中的攻击能力会形成巨大的压力，这种压力虽然看不见，但在巨大的压力过后，它会逐渐释放，并最终回归平静。在弗洛伊德的世界里，这种攻击是本能发现的实现，是潜伏在无意识中的死亡本能。当然，社会的文明过程有效地遏制了无端的暴力。当然，人们的攻击只能在小范围内进行，这个范围也属于文明社会纪律的安全空间。因此，文学叙事中的暴力情节应运而生，在这里剑光剑影的宣泄是暴力本能。当文学暴力的发展引发了电影电视暴力时，这种暴力的叙事在视觉商品中作为影响力符号被越来越广泛地使用。在影响力符号中隐藏着看与被看的问题。人们仰望着他们的身体，同时也被观察着。现代社会的身体已经超出了想象的范畴。身体蕴含着革命性的能量，但身体仍然无法完全承担所有的社会关系。

### **3. 身体暴力**

#### **3.1 暴力的历史根源**

暴力是人类历史上反复思考和分析的一个主题，它的根源深深地植入我们存在的结构中。现代汉语词典将暴力定义为强制力量；具体来说，是国家的强制力。牛津词典将其描述为“为了违反、破坏或滥用而施加的身体力量”。

恶性侵犯是人类独有的一种侵犯形式。这种侵犯具有破坏性，这种破坏性也是人类所独有的，因为无论是侵犯还是破坏都不来自动物的本能。暴力是这种违反和破坏的产物，它从人类存在的最初就产生了。最终，暴力被视为一种在人类世界中延续的社会现象。实际暴力

的源头、动机和目的甚至是一种完全相反的状态。艺术作品中的暴力根本不是普通意义上的社会暴力。这种暴力只属于艺术的范畴，因此是一种非暴力现象。在文化中的暴力，在良性的人类侵犯中，文化暴力的动机与弗洛姆的话相符合：“个体，对种族的生存必要，并且对生存具有适应性”。弗洛姆还指出，暴力杀戮不是人类的普遍本能，而是文明的邪恶结果。随着文明的发展，人类开始相互杀戮，而根本的解决办法是将尊严还给人类，将中心地位还给人类，使人类不再是物质文明、政治和经济等的奴隶。人类暴力的基础通常被认为是对人类残忍欲望的满足。归根结底，暴力来自人类共同创造并为之骄傲的文明。

现实中的暴力每天都在我们周围发生。例如，由于医患纠纷的频繁发生，医务工作者有时会受到伤害甚至被杀害。还有各种虐待猫狗的事件，甚至是周围的小人因为一些小事而发生的争执。今天的中国经济发展迅速，但随着经济的快速发展，也带来了更多的矛盾，这是暴力事件持续发生的重要原因。在如今这个和平的时代，我们是否低估了暴力带来的一系列问题。和平时期的暴力完全是个人心理因素和社会环境影响的结果。与真实的暴力相比，艺术中的暴力似乎要温和得多，但其影响不容小觑。

暴力情结的概念是在弗洛伊德的情结理论下提出的。弗洛伊德将情节视为在无意识活动中，有意识地压抑的欲望，是来自创伤性经历的潜意识本能冲动。艺术中的情结包括观念、情感和意图，而在作品中表现出来的暴力情结是一种倾向于暴力的艺术语言。这种情结受到社会环境、个人隐私等多方面的影响。从美学的角度来看艺术与暴力的关系，一般公众认为美是令人愉悦的，与暴力无关；人们想当然地认为暴力不会在艺术中反映出来。然而，在艺术发展的每一个时期，暴力都是不可忽视的因素；它直接或间接地反映在艺术作品中，这样的作品出现产生了“暴力美学”这一术语。

## **3.2 暴力美学的探索**

### **3.2.1 社会暴力**

现代化的进程伴随着社会暴力的滋生。这种社会暴力具有双重属性，即自然属性和社会属性。社会暴力通常分为三类：攻击、暴力和暴力犯罪。社会暴力的自然属性，广义上指的是当个体的生存利益受到威胁时发生的侵略行为。这种暴力通过个体逐渐发展到社会领域。其社会属性是，特定环境衍生的道德问题最终导致社会暴力的发生。环境对暴力有巨大影响。不同的人对社会暴力有不同的看法。暴力可以损害个体自由和社会秩序，但它也是社会进步

的必要代价。弗里德里希·恩格斯曾说：“至于暴力，在历史上还有另一个角色，那就是革命；至于暴力，正如马克思所说，它是任何一个旧社会在孕育一个新社会时的助产士；暴力是为社会运动开辟道路并破坏僵化和死寂的政治局面的武器。”

在中国，社会暴力问题有着悠久的历史，比如由土地问题引起的暴力。这种社会暴力的出现通常伴随着王朝的更替。这一现象表明，社会通过创造自己的制度或机制来支持自己的组织，并通过这些组织构建社会秩序，但这种社会秩序需要依靠暴力来维持，并通过限制、控制和合理使用暴力来维持。这正是米歇尔·福柯在《纪律与惩罚》中提出的，监狱这一高度集中的方式使社会获得权力。

暴力美学在人类诞生之初就已存在，当人类开始创造艺术时，这种暴力之美也随之出现。当然，这种美通常用来炫耀个人的力量及该力量带来的胜利。这种暴力之美是人类对力量原始的赞美。这种力量能够带给人类生存的空间。《论语》记录了周朝的“大舞”。这种对历史记录的艺术性使用一直持续至今。“庄子”曾记载了赵王文舞剑的场景，他们昼夜搏斗，超过 100 人在此过程中被杀伤，人们对此从未感到厌倦。这些都是中国古代对暴力之美的记录。在西方国家，对暴力之美也有所推崇。古希腊的青铜器上雕刻着战争等对抗场景。还有陶器上的特洛伊战争描绘。由此可见，西方崇拜力量之美。

### 3.2.2 电影和电视中的暴力

在 20 世纪 70 年代，嘉禾电影公司与李小龙先生、成龙先生合作制作了一系列功夫电影，如李小龙的《猛龙过江》、《斗士》等，以及成龙主演的《飞马》。这些电影在国内外上映后，获得了高度评价。在这些赞誉中，“暴力美学”这个词第一次出现，人们所认可的“暴力美学”也是电影和电视中的“暴力美学”——“以仪式化的方式创作电影和电视中的暴力动作场景。”最重要的特征是，电影和电视作品中的暴力场景通过后现代的创作方法被溶解进没有具体意义的情节中；这些情节最终被象征化，使其与电影和电视作品相关联。美学元素，包括视觉和听觉方面。这类电影加剧了打斗的动作，强化了打斗场景的表象形式，有意识地将宣传暴力的形式提升到另一个层次，并故意削弱电影的社会和道德功能。将国内电影中的“暴力美学”提升到成熟阶段的是吴宇森在 1980 年代执导的“英雄系列”，这些电影完美地将电影中的暴力风格化。这一系列的电影为世界展示了一种新的暴力审视，这种审视可以被视为东方美学的视角，这种审视与之前电影和电视作品中的纯粹暴力斗争相分离。这个新视角让观众看到电影不再是一部“暴力电影”，观众可以从电影暴力的渲染中感受到“有意义的形

式”。在外国电影中也有很多暴力美学的表现。例如，昆汀·塔伦蒂诺在他的导演电影中对“暴力风格化”的探索也掀起了一波“暴力美学”。由罗伯特·罗德里格斯和弗兰克·米勒共同执导的《罪恶之城》三部曲将“暴力美学”推向极致。三部影片的主人公都在试图为一个死去或垂死的女友报仇。正义的唯一方式是以暴制暴，侵犯女性的肇事者都有官方社会背景。这种社会地位的悬殊使得主人公们最终不得不用鲜血换取正义和尊严。

在“暴力美学”的影响下，近年来国内电影和电视作品对暴力的处理已经提升到了另一个层次，形成了一种诗意和舞蹈性的趋势。例如《夜宴》和《十面埋伏》等热门电影，这种趋势使得人们观赏的不仅是肉体的搏斗，还有独特的舞蹈表演，而暴力场景在视觉上呈现的美学进一步削弱了暴力的残酷性。

出乎意料的是，表演艺术已经成为当代文化语境中的一个流行词汇。这个意外并不是偶然的，它是现代主义现代艺术中“表演艺术”发展的结果。

作为表演艺术的主要创作对象，身体通过对身体的发现和使用揭示了表演艺术的文化内涵。在表演艺术中的行为是完成整个艺术实施过程的主要动力，而行为需要身体这个载体来支撑完成。这是表演艺术与传统画架艺术的基本区别。身体承载行为，行为创造作品。表演艺术的出现和发展改变了当代艺术的形式。

在当代艺术中，身体不再仅仅是身体，而是一个政治的身体、经济的身体、文化的身体、医疗的身体以及世界的身体。这个多层次的身体为表演艺术带来了更丰富的层次。然而，这个丰富的层次也给表演艺术带来了更多的批评，因为它不再是单一的身体直接冲击社会的承受能力以及政府的规章制度。使用身体作为艺术的主要材料，身体是表演艺术最主要和最直接的表达场所，身体叙事和作品结合成一个整体。

当身体成为表演艺术表达的如此重要场所时，伴随人类发展的暴力在此刻也同样被凸显。无论是对物体的暴力还是对自己的暴力，它都已经成为了表演艺术的一种表达方式。但并不是所有的暴力行为都可以归于表演艺术，我们还应该区分社会暴力、道德暴力、游戏暴力以及其他形式的暴力与表演艺术中的暴力。表演艺术中的暴力倾向是国际性的，达米恩·赫斯特向公众展示了浸泡在福尔马林中的尸体。他那颗镶满真钻的真实骷髅头《为了上帝的爱》也大受欢迎。而穆娜·哈顿想向公众展示她的心脏和胃部。玛丽娜·阿布拉莫维奇向观众提供了72种工具，让观众对她做任何他们想做的事，她永远不会反击。等等，但这些活动大多数是理性的，而不是纯粹的暴力。

许多国内艺术家认为西方是一个非常开放的社会，中国也应该如此开放，所以不管是什么样的行为，它都属于艺术，但西方并不是大家想象中的那样。比中国的人权限制要严格得

多，这种限制不仅仅是外部力量施加的，也是自身施加的。例如，达米恩·赫斯特曾有过对他人暴力的想法是经过协商的，但最终没有成功实施，因为这种暴力在西方不能被称为艺术，而是简单的暴力事件。因此，应该区分不同类别的暴力。

表演艺术中反映的伦理问题不仅仅是对认知的一种抵抗，更是道德话语解构过程中的一个重要环节。在表演艺术中，知识、道德和美学是相互支持的组成部分，这是当然的。当然，“体感美学”的思考和构建并非无可争议，但最重要的仍然是身体本身，以及体感美学和美学体验本身。

不可否认，表演艺术的发展是中国当代艺术的一个重要组成部分，其发展将会更加多元化，并将呈现出更多样的艺术形式。

## 第2节 中国当代艺术中的技术美学

### 1. 引言

#### 1.1 背景及意义

##### 1.1.1 背景

“什么是艺术”的问题已经不再有一个普遍的定义。从现代艺术到现在的艺术风格演变历史已经失去了在艺术研究中的中心地位。仅仅从强制性形式演进的角度来讨论艺术，已不再是艺术史的讨论主题，艺术也不再是一个排除风格而遵循时间线索的成就。就美学而言，仅仅讨论艺术在道德和行为上的美已不复存在。世界在发展和变化，整个艺术世界也在不断地发展和变化。多元的艺术概念在同一时间并存，当代艺术更为复杂。人们可以在多个维度上观看艺术和审视美学。此外，当今的观众也可以从技术美学的审美视角中解放出对作品概念风格的惯性观看。

技术越是熟练，越能解决目的性与规律性之间的对抗。有目的地使用技术的规律性有利于自由形式的形成和美的境界的实现。“技术美学”被认为是在创造艺术的过程中由技术创造出的美。简单地说，就是在将科研成果转化为生产实践的过程中所呈现的美学价值，新技术将促进艺术生产的快速发展。

本文主要围绕“技术美学”，通过对技术、工艺和艺术的分析，逐步探讨中国当代艺术中的技术美学。

##### 1.1.2 研究意义

本节在艺术的生产关系中分析技术、工艺与艺术，并逐渐探讨什么是技术美学以及中国当代艺术中的技术美学，旨在为当代艺术作品提供一个衡量尺度和智性观看的视角。

## 1.2 研究的目的是与研究内容

### 1.2.1 研究目的

在当代艺术中，艺术的复杂性愈加突出，艺术的界限极为宽广。普通观众甚至无法辨别同样的现成物品，哪个是艺术作品而哪个不是。如今，艺术的观看不再仅仅依靠模仿对象的手写美学。通过分析技术美学的美学价值，为当代艺术的评价尺度和观看方法提供一个可能的切入点具有重要意义。

### 1.2.2 研究内容

结合当代艺术中艺术生产者的实际案例，分析技术和艺术之间的关系。通过技术、工艺和艺术的分析，阐明技术的作用，有目的地使用技术的规律性特征，以及获得形式自由和美的境界的目标。

将逐步探讨什么是美学和技术美学，然后分析中国当代艺术中的技术美学，以便为中国当代艺术作品的衡量尺度和智性观看提供一个可能的切入点。

结合个人作品分析技术美学在个人艺术创作中遇到的问题，以及对作品创作的影响。并在最后一章总结其在当代艺术中的意义和局限性。

## 1.3 国内外研究现状

“技术美学”这一概念直到 20 世纪 50 年代末才正式提出。与其他美学概念相比，它出现得相对较晚。在“技术美学”在世界各地蓬勃发展的时候，中国的一些人还在处于“技术美学”辩论的状态。现在，技术美学已经被绝大多数美学家所认可，但是在我撰写论文期间检索到的理论研究成果并不是很多，它们偶尔被涉及，并且与当代艺术联系不大。缺乏一种系统的分析来讨论它在当代艺术中的意义。

## 1.4 创新点

在总结前人经验和讨论的基础上，本节从艺术生产关系出发，分析技术和工艺，阐述其方式，明确技术的作用，并提炼出基于技术的美学价值，即技术美学。通过对中国当代技术

美学的分析，在多彩而复杂的当代艺术中，我们寻求评价和观看的可能切入点。

## 2. 技术与艺术

### 2.1 紧密联系

技术与艺术被认为是孪生兄弟，这是东西方人对技术和艺术达成的共识。艺术离不开技术；技术从未离开艺术片刻。有些人认为技术和艺术是一枚硬币的两面，不可分割；有些人认为掌心与掌背就是技术与艺术，彼此融为一体。在我看来，技术与艺术绝不是对立的，它们只是探索真理的不同方式。艺术创作离不开技术，技术也不能从艺术中孤立出来。

技术与艺术互为一体，艺术与技术具有彼此的概念。这是 16 世纪和 17 世纪的事情，也就是欧洲文艺复兴时期。19 世纪各自独立一直被认为是一种倒退，与文艺复兴人文主义的思想倾向相反。因此，我们不应当在艺术与技术之间做出严格的划分。

### 2.2 技术

“技术美学”一词最初由捷克斯洛伐克艺术家帕特尔·图奇在 20 世纪 50 年代末首次使用，并很快流行开来。在深入分析中国当代技术美学之前，必须认识到技术美学是在将科学研究应用于生产实践的过程中呈现的美学价值，不可避免地首先要分析技术与艺术生产之间的关系。

社会的发展是由技术革命推动的。艺术革命者在艺术生产关系中使用新技术创作时，将促进新的艺术发展。旧的生产关系将不可避免地被新的生产技术打破，新的艺术形式最终将相继出现，显示出多元化的发展。诸如录像艺术和摄影艺术等新艺术形态将不可避免地转变传统的生产和感知方式。

人们使用技术来解决问题，新技术的应用能够产生新的事物。医学是利用技术获得健康，战术是利用技术取得胜利，魔术是利用技术获得新奇，这在各行各业都是如此。精湛的技术是为了解决发生的问题，生产出满足使用者和接受者的结果，所有的操作部署都是针对对象的。因此，使用技术就是为了获得作品的完美与极致。鞋匠制鞋，农民耕地，厨师烹饪，各司其职，发挥所长，互补所短，各有其专业使命，并将精力集中在一个主题上。每个人都使用自己的技术来完美解决领域内的问题，以达到生产者 and 接受者的完美要求。这样，技术已

经成为解决问题的更好方式,技术的应用和创新也为旧的艺术生产方式和感知方式的转变提供了其存在的合法性。

制造商为何不遗余力地追求产品的卓越?这应该是对完美目标的追求,对自创产品的热爱。在生产之前,无论是谁都有一个预定的目的,农民希望获得丰收,厨师追求美味。对所有行业来说,制作活动的过程并非目的。相比之下,获得产品比生产过程更重要。制作出优质产品是生产者欲望的对象。

艺术创作欲望的对象是达到生产者和接受者所要求的完美,就像医学中的健康和战术中的胜利一样。艺术生产的实践活动是生产者使用技术获取完美作品的欲望,而接受者在作品中找到符合自己审美的感知。技能(技术)的使用是有意义的,因为它帮助实现专业目的。在艺术生产的实践中,技术是必需的,因为它帮助实现艺术作品的完美目的。所有行业都是这样的。

德国建筑博物馆曾经以“今日设计展览,工业还是艺术?”作为展览标题,这是艺术对现代工业技术的反思。技术是实现艺术生产实践的重要手段,但作为人与外界调试过程的一部分,艺术本身无法自解其义。美学恰恰是解释艺术的科学。技术在艺术生产中的作用非常重要,但并非所有的生产技术都是有价值的。有一种美学价值是与生俱来的;此外,当技术介入艺术实践时,艺术本身是不可区分的,当一个不可或缺时,我们也认为它有价值。在不同人的眼中,这两个价值的价值是无法等量齐观的。

当沃尔特·本雅明(Walter Benjamin)进行艺术生产的研究时,他清楚地指出,艺术在内容和形式上有问题,同时也在技术上存在问题。他认为,技术不仅仅是艺术生产的手段,而且在社会关系中也有明显的落后性或进步性。在一定程度上,新的技术因素对艺术的发展有巨大的推动作用,这改变了艺术的形式和功能,导致艺术的多元化趋势。在艺术生产的实践中,技术因素是生产力,代表了某一阶段的艺术发展水平。基于他的研究结果,本雅明得出了技术决定论的结论。同时,本雅明将这个时代称为“机械复制时代”。随着这个时代的到来,现代艺术作为新技术应用的必然结果取代了传统艺术。技术创新不仅带来了大量的艺术作品,而且也导致大量艺术作品流向市场,造成了繁荣的艺术市场,并且艺术作品越来越成为日常生活的一部分。

## 2.3 工匠精神——以画架画中工匠精神的拥护者冷军

### 为例

在讨论技术时，我们不得不谈论工匠精神。工匠精神是指在工作中追求完美和精益求精的精神理念，它被视为工匠精神。工匠精神被认为是严谨细致、坚持专注的意志品质，对技术因素的最高监视。

然而，在艺术生产的关系中，有些艺术制作者羞于将工艺（技术）作为创造手段来关注。他们认为工艺（技术）是低下和不精炼的手段。当人们首次提到工匠精神时，他们想到的是“工匠”，这个词已经被无限贬低，认为它是重复的机械操作和外来工人的工作。实际上，这是对工匠精神和工匠这个词的误读，工匠的精神和工匠本身是不能等同的。

从事某一行业并具有丰富经验的人被称为工匠。由于各种原因，工匠只能反复从事特定的操作，而在实践中获得的丰富经验无法提升到理论层面。重复的操作方法只能用于生产过程中，无法获得更完美的作品。作品自然缺乏自己的主观想法和精神性。技术成为工匠生产艺术作品的障碍，也导致艺术界不高看工匠的作品。自古以来，工匠精神就有贬义色彩。在绘画艺术中，也有“工艺性”一说，指的是那些没有灵气、定式化、僵化且缺乏变化的作品。无论是画家还是评论家都一致认为，绘画应该避免平庸和工艺性，唯一平庸和工匠的就是线条画。现在那些没有才能和精神性的艺术家很容易被称为“工匠”，而那些定式化和僵化的作品也被人们贬值。

与工艺不同，工匠精神对艺术生产实践是不可或缺的。如果没有工匠精神，艺术创作可能会草率，技术不会更新和长期使用，落后的技术将阻碍艺术生产的发展，妨碍艺术的进步。坚持工匠精神将促进技术创新，进而推动艺术的发展。

熟练的技术有利于更完美地实现艺术创作的目的并获得艺术价值。艺术生产不能以某一方式概括。整个艺术世界处于持续发展之中。工匠和工匠精神也随之变化。拥有工匠精神的工匠可以成为艺术家。

盲目专注于技术似乎成为了一些艺术制作人的魔障，他们无法摆脱这种自娱自乐的状态。表面上看是这样，但实际上他们是工匠精神的拥护者，他们是新时代的“工匠”，正如钱学森所说：“科学技术与文学艺术之间的互动规律应该被我们有意识地研究。”艺术批评应该从多个维度评价艺术家，不能因一点就下结论。画架画家冷军就是这样一个坚持者。他只把自己当做一名工匠，试图用精湛的技术创造作品，人为地阻止社会的浮躁。

冷军的作品曾经被批评为“超限的绘画”。他认为这样的作品超越了材料本身的限制，但极大地提升了材料的表现力。这是对取得成就的认可。不同于西方出现的超现实主义绘画，冷军的作品不再仅仅是用绘画方法创造照片质感，展现当代社会现实的冷漠一面。而是以追求完美的态度，技术上达到完美状态，用作品代替了一种不同的生命力——最真实甚至“超现实”的图像效果，给观众留下深刻印象并传染给观众。

作为一个杰出的艺术家，冷军对自己身份的态度非常清晰而不可思议。他认为自己只是一个工匠，一个画家，一个玩手的人，不是思想家，不是社会学家，不是公共知识分子，或者简单地用时间打磨自己的作品，目的是让观众在作品前停留一段时间。以工匠的精神，我使用传统人类文化遗产中有价值和经典的东西人为地阻止了社会的浮躁。冷军坚持的正是亚里士多德认为的“善”，他所做的一切都是为了它，技术成为他实现“善”的手段。坚持工匠精神，最终的成就是像冷军这样的艺术家被记录在中国艺术史上，创作了诸如《五角星》和《小罗肖像》等优秀作品[3]。

## 2.4 技术承载道

作者认为，技术不仅能够承载思想，还能解决这些思想需要解决的问题，允许通过技术手段表达思想——换句话说，就是“技术携带道”。

引用古老的中国文献《易经》中的话：“形而上者谓之道，形而下者谓之器。”这里，“道”指的是自然界中的根本原则或法则，“器”指的是工具或设备。“技”存在于道和器之间，既是连接它们的桥梁，也是器得以获得道的媒介。电影制作人詹姆斯·卡梅隆在讨论他的工作方法时强调，他采用3D技术的目的是为了吸引沉迷于电脑、电视的观众回到电影院，而不仅仅是为了展示技术有多先进。有效的技术可以更容易地实现作品的完美形态；同时，当作品达到最完美的状态时，好的技术会悄无声息地消除技术的痕迹，只展现艺术的价值。卡梅隆的这种技术使人们“忘记技术”的原则，与清代中国著名画家石涛所说的“无法有法，此之谓至法”相呼应。技术是艺术创作的一个组成部分，也可以承载艺术的道。两者相互作用，相互依赖。因此，在创意表达领域中，技术与艺术本身一样重要，我认为它应当被视为评估艺术作品的一个有效标准。

## 3. 技术与美学

### 3.1 什么是美学

Zangwill 认为：“艺术是美学价值的创造和构建。”那么美学是什么呢？德国美学家 Geiger 认为：“美学是一门关于价值的科学，是一门关于美学价值的形式和法则的科学。”实际上，这个说法并没有被绝大多数人所认可，也没有一个公认的美学定义。然而，接下来必须对美学给出定义。从美学出现到现在，美学被无数美学家定义过。总结起来，基本上有五种类型：

美学是研究美的科学。

美学是艺术哲学。

以美学经验为中心，将美与艺术的科学视为美学。

美学词汇的语言分析科学。

认为美学是研究美学价值的科学。

在中文语境中，“科学”这个词汇非常重大。只有形成一个完整的知识体系，才能被称为“科学”。对于那些刚接触美学的艺术研究者来说，当他们看到“美学是什么科学”时，有些人可能会感到气馁。实际上，这里提到的科学不是我们在中文中所说的自然科学，而是学科的意思。这里需要特别说明一下。

第一个定义，认为美学是研究美的科学。这在中文里的字面意思非常清晰。但在西方语言中，会出现意义上的问题。在西方语言中，感性 (sensibility) 是研究美的一门学科。这与中文强调对象的物质性质相反，西方更强调主体性。《大英百科全书》应该不会同意上述说法，并在美学的定义中增加了“品味” (taste)：美学是关于美和品味的哲学研究。这样，在中文表达中，它变成了美与趣味的并列情况，不符合中文表达的含义。另外，在中文里，认为美学是研究美的科学，而在西方语言中，美是指美的本质，这与中文说的所有美都是美的说法不同。然而，西方现代美学曾经将美的本质视为一个错误的命题。从这个角度看，东西方对这第一个定义就没有共识，更不用说所有美学家之间了。

定义 2：艺术哲学被视为美学。简单来说，就是研究艺术之美为什么是美的。技术、社会等的美并不是艺术之美，其本质是美。研究艺术的美，艺术的本质就是西方所说的艺术哲学。这个定义在西方被绝大多数人所认可，从巴图到黑格尔，都得到了尊重，但这并不符合

中国的逻辑。首先，从语言表达逻辑来看，艺术包含在美学之中，艺术比美学小。此外，从中国美学经验的角度来看，排除艺术之外的其他美学问题，如社会美和技术美，是不合理的，这也不符合中国文化传统中的艺术思维。

定义三：以美学经验为中心，认为研究美与艺术的科学是美学。这样的总结应该说是把能考虑到的所有因素都包含在这个定义中，看起来似乎全面，但失去了它的逻辑性。什么经验是定义的中心，那么美学经验就是美学的基础，美从何来，美来自于美学经验。那么，不包含艺术的美之袋怎么办？如果艺术被包含在内，那么在概念中就无需讨论对美与艺术的研究；如果艺术不能被包含，那就等于是把艺术推到外面，也就无法谈论美学。这个总结由美学家李泽厚给出，也被大多数国内美学家所认可，但普遍认可的并不具有真理性，因为真理可能掌握在少数人手中。

定义四：认为美学词汇的语言分析科学是美学。这是分析美学关于美学的定义。仅仅使用语言分析的方法将一个感性的学科当作逻辑来对待，并排除了传统美学中涉及的美学领域的研究，在美学界应该没有共识。

定义五：认为研究美学价值的科学是美学。如何判断一个事物是否具有美学价值是这个定义的难题？如果绝大多数人同意它的美，就判定它具有美学价值。但当只有一个人说它是美的时候，它就没有美学价值吗？将价值论应用到美学的定义上在一定程度上是符合美学的复杂性的，并且有利于对一个事物进行美学评论。

这五个定义仍是最受支持的，每个定义需要来回长时间地解释，而且它们都还不够令人满意和有说服力，更不用说用一个定义来解释美学了。美学是什么？百家争鸣，百花齐放，我认为给它们一个积极和共通的定义并不容易。借用西方物理学界定义物理的方式也是一个不错的选择。美学是美学家的活动。

## 3.2 技术美学

在 20 世纪 50 年代末，捷克斯洛伐克艺术家帕特尔·图奇首次提出了“技术美学”的术语，并自此在一些国家流行起来。

技术之美是技术美学的核心，而技术是技术之美的核心，艺术最初的含义是指技术满足艺术制作者和艺术接受者的完美要求。艺术制作者可以利用完美的技术服务于艺术接受者，艺术接受者也可以找到与自己的审美相匹配的感知，对双方都非常适用。技术是实现美的境界的手段，技术之美是技术应用中产生的审美价值，技术之美是技术美学的核心价值。

在我国，美学家李泽厚首先提出了“技术之美”的范畴，并将技术之美、自然之美、社

会之美和艺术之美列为美学的四大类别。既然有技术之美，那么自然就有技术美学。学者屠屠写了一本名为《技术美学》的书。艺术之美是人创造的，这种美被视为“艺术美学”；那么，由技术创造的艺术之美应该被视为“技术美学”。技术美学是技术与美学的融合。技术是人们解决问题的方式，换句话说，是将研究成果应用到实践中的科学；美学是研究美的最高形式。通俗地说，技术美学是科学研究成果在转化生产实践过程中的审美价值。

### 3.2.1 奥拉维尔·埃利亚松

在当下阶段，没有人知道艺术作品究竟是什么，对艺术作品的限制似乎已经不存在了，清洁工甚至难以判断是否应该清理掉艺术馆中的一堆灰尘。试图从技术美学的视角来解释当代艺术中的作品，是一种新的可能性。

奥拉维尔·埃利亚松运用各种技术工具控制如光、雾、水和空气温度这些善变的自然物质，创造出触发各种自然现象的环境，营造出幻觉和沉浸式体验。当我第一次看到埃利亚松的作品时，我认为它令人惊叹，但还是有些不明确。基于技术美学的视角，尝试思考技术的目的和美学的价值，并不难发现艺术家想要利用人们的好奇心和个人体验来迫使观众思考这些自然现象如何构成我们所知的世界。以及人类如何转变这一自然，我们可以逐渐尝试评估艺术家创作艺术的目的，以及作品本身的影响力。

1993年创作的作品《美》(Beauty)，被认为揭示了埃利亚松对人造与自然的表达。从多孔软管喷出的水雾可以被多个聚光灯照亮，在墙上投射出彩虹光晕，而软管和照明设备就像彩虹一样清晰可见。

2008年，埃利亚松在美国纽约东河完成了他的巨型装置作品《纽约市瀑布》(The New York City Waterfalls)。艺术家使用水循环和虹吸原理技术，将东河的河水从与自由女神像同样高度的金属支架上喷出，创造出人造瀑布。

2010年的作品《你的盲旅客》(Your Blind Passenger)是丹麦阿肯现代艺术博物馆委托创作的。艺术家制作了一个90米长的通道，通道内部只有1.5米的能见度，并且周围环绕着雾气。当参观者失去视觉，他们只能使用其他感官来导航和指引自己。

2003年完成于英国伦敦泰特美术馆涡轮厅的《天气项目》是埃利亚松最著名的作品。埃利亚松直接在雾中创造了日落的画面：雾是通过使用雾机将糖水制成带有甜味的雾化气体而形成的；而日落则是由数百个橙黄色单频光拼凑成一个巨大的发光球体。同时，艺术家在大厅的天花板上覆盖了银色箔，底部是镜子，形成了“太阳”和观众的清晰镜像，使得在“太

阳”下行走或躺卧的观众似乎处于梦境之中。像其他作品一样，埃利亚松揭示了作品背后的机制，并使用雾机的轰鸣声和人造太阳的供电系统在明亮的地方来削弱模拟自然过程的浪漫主义和作品的情感方面的宏伟。

### **3.2.2 埃利亚松的创意意图被认为是对让·鲍德里亚的拟像 (Simulacrum) 理论的模仿**

模拟 (Simulation) 理论不同，拟像手段仅仅是艺术家用来表达“反模拟”的概念性手段。埃利亚松直接复现自然风景，意图让真实的事物更加真实。目的是为了让观众关注到他所思考的活动，而不只是他所创造的虚拟影像[4]。

他观点的表达隐藏在一个事实之中，那就是他通过自己特别设置的“裂缝”来展现自己的思想，这有点像艺术手法中的魔术，允许观众在好奇心的驱动下发现并思考其中的神秘。这样，技术成为了他“魔法”的关键，成为幻觉和拟像之间的桥梁。

# 第3节 分析与研究：当前高校公共艺术 教育现状的对策

为了分析我国高等院校公共艺术教育的现状,我们首先必须弄清楚什么是公共艺术教育。公共艺术教育的定义因人而异,仁者见仁,智者见智。大多数学者认可的、更有说服力的定义是:公共艺术教育指的是在广大群众中广泛开展的一种审美教育形式,即通过教师的艺术知识教学,利用音乐、摄影、舞蹈、绘画、书法、雕塑、文学、戏剧、电影、表演艺术等艺术手段,对学生进行情感教育的总称。公共艺术教育是一种集娱乐、欣赏和基础知识于一体的综合性教育。目前,我国高等院校大力发展公共艺术教育的目的,在于优化学生的知识结构,提高学生的审美能力和艺术鉴赏水平,培养学生的情操,通过艺术教育的渗透,塑造学生高尚的心灵。此外,在高校做好公共艺术教育工作,也有助于培养学生的创新思维,培养学生的感知力,使当代大学生更顺利地培养成为全面发展的高素质人才。高校的公共艺术教育是一种在大学生中广泛进行的审美教育形式。它对大学生进行艺术方面的培训和培养,使大学生能够在课堂上掌握人类艺术审美的体验,并充分感受艺术之美的感染。为了达到身心全面和谐发展,它以一种微妙的方式具有良好的审美素质。

近年来,我国普通高等院校的公共艺术教育工作相对顺利地展开,总结起来,主要表现在以下几个重要特征:

## (1) 普通高校高度重视公共艺术教育

随着素质教育理念的深入人心,普通高等院校越来越重视大学生的公共艺术教育发展,并在人力、物力和财力上投入巨大。各大高校认真执行教育部关于公共艺术教育的规章制度,如《学校艺术教育工作条例》、《全国普通高等院校公共艺术课程指导计划》等。绝大多数普通高校都开设了公共艺术课程,并大力加强公共艺术教育课程建设,深化公共艺术教学改革,进一步提高了普通高校艺术教育的水平。同时,各大高校开设的公共艺术课程数量也显著增加,公共艺术教学逐渐规范化,教学质量得到了极大提高。许多高校还根据自身实际教学发展的需要,增设了一些具有学校特色和地方特色的艺术教育课程。例如,电影评论、摄影艺术、油画欣赏、交响乐欣赏等,这些课程大多是各大高校根据学生与教师的实际情况开展的公共艺术教育课程。

## (2) 将校园文化活动发展成为艺术教育的重要载体。各大高校积极开展各种丰富多彩

的校园文化活动,逐步将这些活动发展成为普通高校公共艺术教育的重要载体,无论是举办大学生艺术节还是各种大学生歌唱、绘画等形式的比赛,这些丰富多彩的校园文化活动有效地提升了大学生的艺术素质,使得所有大学生都能参与到艺术活动中,为我国普通高等教育机构营造了良好的艺术氛围。

(3) 公共艺术教师队伍得到发展,师资结构趋向专业化和合理化。虽然在后续章节还会分析普通高校公共艺术教育存在的问题,其中之一就是教师队伍的薄弱,但不可否认的是,近年来,仍然有一批受过专业训练、教育程度较高的音乐、艺术教师加入到高校公共艺术教育的教师队伍中。许多艺术教育教师也在不断地学习、观摩和实践,以此不断提升自己的艺术修养。

(4) 普通高校公共艺术教育的投入持续增加,硬件和软件设施大幅度改善。高等院校在公共艺术教育上的投资不断增加,无论是教学设施还是教育资源,都有了大幅度的提升,这为学生提供了更好的学习条件,也为艺术教育的深入开展奠定了坚实的基础。

在许多高校的搬迁过程中,规划和建设了一批学生艺术活动中心和多媒体艺术教育室。这充分显示了普通高等院校对公共艺术教育的重视,学生艺术社团的迅速且良好发展,以及丰富的艺术活动,这些都有助于为校园文化创造良好的艺术环境,营造健康、高雅的校园艺术氛围。这些艺术设施的大量增加也反映了我国普通高等院校公共艺术教育的快速发展。

通过上述分析,我们不难发现,艺术教育是普通高等院校公共教育的重要内容,也是普通高等院校提升素质教育的重要手段和措施。可以说,艺术教育是素质教育的重要组成部分,是青少年学生健康成长不可或缺的因素,也是当代大学生实现全面发展、提升个性的重要手段。通过研究普通高等院校公共艺术教育的现状,我们可以找出当前公共艺术教育中存在的问题和不足,进一步提高公共艺术教育的水平,并为普通高等院校创新公共艺术教育寻找新的方法。

## **1.我国普通高等院校公共艺术教育的现状**

### **1.1 我国普通高等院校公共艺术教育发展的现状**

近年来,艺术教育问题受到了广泛关注。几乎有一个一致的观点:没有艺术教育,就无法谈论素质教育。没有艺术教育,人们丰富的想象力从哪里来?如何培养人们敏锐的感受力?如何塑造生活的境界?技术产品,也在通过文学艺术的学习,努力培养他们对美的热爱、追求和信仰。那么,艺术教育的好处是什么?毫无疑问,艺术教育是素质教育不可分割的一

部分，但艺术教育难道仅是学习乐谱、听几首著名歌曲、欣赏几幅画作、观看几出著名戏剧吗？被引导学习艺术的同学们熟悉艺术史，他们往往感到困惑。特别是当代西方艺术史模糊了艺术与非艺术的界限，颠倒了美与丑的关系，拒绝在艺术中传递真和善，使艺术变成了一种“快餐式”的精神消费……等等。“究竟什么是真正的艺术？”“哪种艺术具有现代性？”这些问题不仅困扰了一般的艺术鉴赏者，也困扰了艺术学生、艺术工作者，甚至是当代艺术家。在真正的现代意义上，什么是艺术？艺术的现代性是否由现代主义和后现代主义代表？当代艺术观点的混乱是前所未有的。一些艺术家将真善从艺术中剔除，使艺术成为纯粹形式美的对象或形式游戏，以至于虽有美，却没有美的意识传达；一些艺术家站在挑战所有人类传统价值观的立场上，肆意将艺术变成了丑陋的表演；一些艺术家将自己变成了前卫，将艺术创作变成了适合宣泄个人情绪的产品，艺术失去了沟通人类情感和心灵的功能。这种情况的出现严重干扰了艺术教育的顺利进行，甚至基于这个例子的艺术教育违背了艺术教育的宗旨。因此，先于艺术教育的问题是艺术观念的教育问题。在上文中，我们已初步阐释了高校公共艺术教育的意义，并理解到在高校发展公共艺术教育在我国进一步推动高等教育中发挥了重要作用。然而，尽管我国普通高等教育机构的公共艺术教育发展迅速，但仍然存在许多问题。我国普通高等教育机构中的公共艺术教育的现状是什么，可以总结如下？

### **1.1.1 尽管我国普通高等院校的公共艺术教育发展迅速，但公共艺术教育的观念相对滞后**

近年来，经济的持续发展也为文化交流创造了更多的机会，艺术的发展与传播在中国获得了前所未有的机遇。然而，由于传统教育体系的影响，在我国的普通高等院校中，艺术教育被看作是其他技能教育的教学附属物，无论是在教学的整体安排上还是在教育和教学的观念上，都没有扩展其自身特有的发展空间，也没有获得最佳的发展环境。尽管许多高校已逐渐开始重视公共艺术教育并开展相应的课程，但从学科建设来看，普通高等院校的公共艺术教育仅是在高就业压力下对专业人才培养的一种“调剂教育”。它没有满足全面推进素质教育的要求，甚至还有很长的路要走。我国目前的教育机构在提高公共艺术教育的公众意识方面，受传统教育的影响表现出对公共艺术教育的错误观念。此外，我国教育部门也成立了相应的管理部门，对普通高等院校的公共艺术教育进行特别指导、监督和总体管理。组织定期培训和相对规律的检查，以测试培训成果，使我国的各大高校在开展公共艺术教育时有法可依，从而促进了我国高校公共艺术教育的顺利发展。目前，虽然许多高校管理者已经意识到

公共艺术教育在素质教育中的重要作用,但他们仍需要进一步提高对素质和教育功能的理解。可以说,普通高等院校的公共艺术教育应遵循党的教育方针和国家对培养高素质人才的要求,坚定把握先进文化方向,把学校的艺术教育作为培养高素质人才和传播先进文化的重要手段,贯彻到学校教育的各个方面。

### **1.1.2 我国普通高等院校的公共艺术教育趋向专业化**

目前,我国的普通高等院校基本上已经开始设立公共艺术专业,这使得高校的公共艺术教育趋向专业化。然而,仍有许多高校对公共艺术教育的概念和内涵缺乏科学和完整的理解。他们认为公共艺术教育就是艺术技能教育,即“音乐教育 + 美术教育 = 艺术教育”的模式。这种认识是公共艺术教育专业发展的巨大障碍。我们必须克服我国当前普通高等教育机构对艺术教育概念和内涵的狭隘理解,否则将导致公共艺术教育实践中的盲目性和非理性。当前,我国高校的公共艺术教育重点在于营造良好的校园艺术教育氛围。在校园文化建设中,公共艺术教育是最活跃的部分,公共艺术教育已逐渐成长为校园文化的重要支撑点,直接影响校园文化和艺术的水平与品味。通过普通高校公共艺术教育的发展,将有助于我国高校营造一个健康、高雅、积极的教育氛围,并允许当代学生有更多机会在国内外体验各类艺术形式。艺术的基本知识也提高了大学生的审美欣赏能力。

### **1.1.3 我国普通高等院校的公共艺术教育教学计划越来越完善**

公共艺术教育在我国高等教育中已经开展了很长时间,其教学方法和教学计划也日益改进。高校公共艺术教育的学科建设以全面、科学的方向设计和发展。校园公共艺术教育的氛围也逐渐从冷漠转变为浓厚。事实上,公共艺术教育的范围非常广泛,并不仅仅依赖于传统艺术教育概念中相对单一的音乐、文学鉴赏等艺术课程。传统的公共艺术教育课程已无法完成成为大学生提供终身艺术教育的使命。因此,我国普通高等院校的公共艺术教育不断加强艺术教育课程体系的建设。在这一过程中,必须构建以人为本、规范化和科学化的公共艺术课程教学体系,这是发展公共艺术教育的关键。因此,我国普通高等院校的公共艺术教育应该建立艺术教育课程体系并创建优秀课程,使学生能够真正接受系统的艺术教育。同时,还需要重视普通高校艺术教育教材的建设,逐渐建立大学艺术教育教材的评审体系,并积极推荐一批质量高、特色明显且受学生欢迎的艺术教育教材。

通过上述分析,我们不难发现,我国普通高等院校公共艺术教育的当前发展现状是发展速度快、规模大,但质量仍需进一步提升。也就是说,为了真正充分发挥艺术教育的重要作用,我们仍然有许多工作要做。只有认识到普通高等院校公共艺术教育的现状,发现存在的问题,并运用科学合理的方法解决这些问题,才能使普通高校的大学生在不久的将来在公共艺术课程中获得更多收益,更多地感受到人文关怀和艺术的影响。艺术教育的成果肯定会在提高他们的整体素质上得到凸显。

## 1.2 我国普通高等院校公共艺术教育存在的问题

通过对我国高校公共艺术教育现状的研究,我们总结了普通高等院校公共艺术教育的几个特点,并且在公共艺术教育快速发展的过程中,也发现了一些不可避免的问题,主要集中在以下几个方面:

公共艺术教育尚未获得教育管理者更多的关注。与其他学科教育相比,公共艺术教育在中国高等院校的地位并没有在素质教育中得到足够的重视。与当前素质教育的发展相比较,我国普通高等院校公共艺术教育的发展水平仍然不够。因为我国高校公共艺术教育的发展历史曲折而漫长,其发展过程经历了停滞-发展-退潮-再发展四个阶段。结果是,在实施素质教育的过程中,我国高等教育的管理者对公共艺术教育的认识不足。当前,我国高校的公共艺术教育机构不够完善,运行体系不畅,同时缺乏政策支持和制度保障。高校公共艺术教育在职责划分、发展定位和相应政策的实施等方面缺乏必要的制度保障,这也是公共艺术教育发展的根本障碍。

教师的弱势制约了公共艺术教育的快速发展。我国普通高等院校公共艺术教育教师队伍的综合素质和专业素质需要进一步提升。这既是公共艺术教育课程改革的需要,也是教育者和受教育者自身的需求。从教育者的角度来看,我国普通高校从事公共艺术教育的教师大多是艺术院校毕业的。这些教师在校期间受到专业艺术的限制,其教育和培训模式相对固定和单一,不可避免地存在着强调艺术技能而忽视知识的教育缺陷,高等艺术教育的知识范围狭窄,人文知识浅薄。对于受教育者来说,公共艺术教育的对象是来自不同专业的大学生,具有不同的知识背景和需求,他们需要教师在公共艺术教育课程中提供更广泛的指导,这进一步要求我们的高等艺术教师需要具备广泛的知识来满足学生的需求,提高教师个人素质势在必行。可以说,公共艺术教师在教育中扮演着引导学生提升艺术修养的角色。此外,当前我校的公共艺术教育尚未形成一套统一标准的教材,这也导致教师使用教材更加随意,公共艺术教育难以形成一定的知识体系和教育规模。在这个过程中存在很多随机性,还有专业艺术

院校压缩使用教材和根据专业教学设置课程的情况。这表明从事公共艺术教育的教师缺乏对公共艺术教育自身教学规律的研究，无法满足高校公共艺术教育的发展需求。

在高校中对公共艺术教育的认识仍存在偏差。无论是受教育的大学生、学生家长，还是社会的其他各个领域，都在一定程度上对高校公共艺术教育存在一些误解，这也限制了高校公共艺术教育的发展。就大学生而言，许多人认为，在大学教育中公共艺术教育是可有可无的，甚至将公共艺术教育视为一种娱乐和欣赏活动，认为它不能与专业教育相提并论。当代大学生对公共艺术教育的理解本质上忽视了公共艺术教育在提高自身综合素质方面的作用，这导致他们参与的热情不高。大部分大学生偏好艺术欣赏，但对专注于技能训练和公共艺术教育理论研究的课程兴趣较少。公共艺术教育难以涉及音乐和艺术之外的其他教学内容，这使得公共艺术教师难以与学生的思想和感情进行沟通。

高校公共艺术教育效果的评估机制尚不够完善。目前，我国高等院校的公共艺术教育还处于探索阶段，在组织管理、制度保障、师资队伍、课程及教学设施等方面都有待提高。其中，高校公共艺术教育效果的评估机制是其中较为显著的问题和缺陷之一。在深化高校公共艺术教学改革的过程中，只有不断加强管理措施和完善评估机制，才能为高校公共艺术教育的发展提供足够的保障。由此可见，我国的普通高等教育机构应当逐步构建和完善以艺术课堂教学为核心，辅之以艺术素质教育讲座和艺术社团活动的教育体系领导思想，并进一步改进校园公共艺术教育评估机制，以改变高校公共艺术教育的现状，真正实施素质教育。

高校公共艺术教育的管理机构不够完善，专业化的机构建设也不够完善。经过调查研究可以发现，在我国普通高等院校中，目前负责公共艺术教育的管理机构非常少，而且专门从事高校公共艺术教育管理的人才也相对短缺。只有少数高校建立了公共艺术教育中心。有些学校缺少领导学校公共艺术教育工作的机构。为了弥补这一缺陷，一些高校不得不依赖学生工作组织、共青团委员会、学生工作部门或者学校的艺术系来负责全校的艺术教育工作。从本质上讲，这导致了普通高校公共艺术教育的缺失。

普通高校的艺术教学和课程设置与教育部对公共艺术课程的要求之间存在很大的差距。具体表现在课程设置不完整、教学内容单一、课时不足、选修学生不够热情等问题。只有少数高校为公共艺术教育设置了必修课程，绝大多数高校仍将艺术课程作为选修课程设置。同时，许多高校受到教师力量不足、教学设施不完善等因素的制约，公共艺术教育的教学课程种类少，根本无法满足素质教育的要求。即便与中小学相比，高等艺术院校在艺术教育方面也存在一定的差距。不仅没有统一规范的教材，而且在专业艺术教育的教学内容、方法和评价标准上也缺乏规范，这导致了对专业艺术院校使用的教材和按照专业教学设置课程的情况

被压缩。

### **1.2.1 我国普通高校的公共艺术教育机制不够完善**

从管理者的角度来看，我国普通高校公共艺术教育机制的缺陷是显而易见的。尽管国家多年来一直倡导素质教育，但并没有真正落到实处，也没有真正改变应试教育主导地位的现状。对素质教育的理解不足导致在国家教育体系中缺乏专注于公共艺术教育的教育机制。分数和考试仍然是学生、教师和家长关注的焦点。艺术类考生在高考中的低文化分数线甚至将艺术教育降低为进入高等教育机构的跳板。公共艺术教育在塑造学生完美人格和提高整个民族素质方面独特且不可替代的作用也被完全忽视了。我国普通高校公共艺术教育发展的政策和规章相对滞后，组织结构不够完善。但对普通高校的管理者来说，缺乏对普通高校公共艺术教育的统一规划和领导，高校公共艺术教育的管理不到位。因此，我国普通高校的公共艺术教育管理体系尚未建立和完善，公共艺术教育没有集中的管理部门，仍处于多层管理或管理无序的状态。这也导致了一些高校公共艺术教育的断断续续和随意性，以及用活动代替课程、用讲座代替课程的情形。

### **1.2.2 我国普通高校公共艺术教育缺乏政策和财政支持**

为了共同维护良好的文化环境，教育、文化、宣传出版等部门必须“共管”，有效加强对青少年各类影视节目和文化出版物的管理，严禁污染人们思想的“精神垃圾”流入学校和社会。严厉打击文化市场上的各种犯罪行为，净化社会文化环境，并确立音乐和艺术教育在教育中的地位。根据《中国教育改革和发展纲要》的要求，应在教育法规和制度中确立音乐和艺术教育的地位。各级教育部门和学校应将音乐和艺术教育纳入重要议程和教育发展计划中，并使音乐和艺术教育逐步法制化、规范化。加强艺术教育课程建设。要促进艺术教育的可持续发展，必须抓住艺术教育课程建设的核心，全面提高普通高校艺术教育的管理水平和课程教学水平，使普通高校艺术教育进入规范化、制度化的轨道。构建以人为本、标准化和科学化的艺术课程教学体系是发展艺术教育的关键。为此，需要建立艺术教育课程体系，创设精品课程，使学生能够真正接受系统的艺术教育。高水平的教师是高校艺术教育优异成果的保证。同时，艺术教育必须有高水平的教学条件和设备，建设良好的多媒体视听教室和学生社团活动，购置各种乐器及必要的绘画、书法、舞蹈道具。要提高师资和教学设备的水平，

需要高校管理者对艺术教育有正确的认识，对艺术教育的投入不仅要有保障，而且要逐年增加。艺术教育的资金投入应被视为人才培养的重要前提。

### **1.2.3 我国普通高校公共艺术教育师资保障的缺失**

近年来，我国各高校的公共艺术教育取得了长足进步，公共艺术教育的管理体制逐渐进入了既定轨道。高等院校也相继开设了一系列公共艺术教育课程，这是我国高校公共艺术教育持续发展的不争事实。然而，在高校实施公共艺术教育的过程中，如何建设一支高水平的教学团队始终是一个尚未完全解决的问题。现阶段我国高校的公共艺术教育教师是否满足艺术教育工作的要求，的确值得共同探讨。面对公共艺术教育的现实，我们应当充分看到，在普通高校公共艺术教育不断发展壮大的过程中，各类人才齐聚此领域。虽然为教师队伍注入了新活力，但同时也造成了公共艺术教育教师队伍素质参差不齐的局面。在普通高校，一些公共艺术教育教师显示出在教学中的一系列文化修养不足、文化底蕴薄弱、科研能力弱、职业精神和教学能力差、教育观念和理念滞后等问题。显然，这样的教师不能完成新世纪培养公共艺术人才的任务。这些教师的存在，在一定程度上制约了我国普通高校公共艺术教育的发展，也极大影响了公共艺术教育的质量。针对这个问题，我们应当通过现象看本质。部分原因在于，现在从事公共艺术教育工作的教师在其学生时代接受的是相对单一的学校培训模式。其次，这些公共艺术教师并非专业的艺术教育者，他们很多是边学边教，甚至有些教师是从一些相关领域兼职担任课程，因此在教授公共艺术教育相关课程时显示出知识狭隘和人文知识的肤浅，艺术素质差，艺术理论基础薄弱等问题。此外，公共艺术教育领域的少数教师能够树立终身教育的理念，并将其付诸于公共艺术教育活动中，教师自身也未能不断丰富提升。在当今推进素质教育深入发展的今天，高校的公共艺术教育迫切需要一个能干的教师团队来完成公共艺术教育改革的历史任务。可以说，只有真正建立起一支高素质的公共艺术教师队伍，才能确保艺术教育工作能够取得实效[6]。

### **1.2.4 我国普通高校公共艺术教育缺乏统一的教材**

在我国普通高等院校中，公共艺术教育的教材编制存在许多问题，其中最突出的问题是缺乏统一编纂的公共艺术教育教材。在编制我国普通高等院校公共艺术教育教材的过程中，教材缺乏整体设计，编纂工作不够有针对性，往往倾向于片面。此外，在编纂过程中还缺少

相应的管理系统,对编纂目标的确定也存在单一性。在公共艺术教育材料编纂中,各自为政,教材的选择和编纂各不相同,这增加了教学难度,极大降低了我国普通高等院校公共艺术教育的教学效果。由于缺乏统一编纂,我国普通高等院校公共艺术教育教材的编纂不能反映团队合作精神 and 集体智慧,导致缺乏对教材使用的全面衡量和评价,使得编纂教材的专家无法评估教材,并找出其优点和不足,这使得教材的编纂过程越来越简单、粗糙,缺乏系统性。目前,我国高校的公共艺术教材不能真实反映学生使用自编教材在知识、能力、情感等方面的培养效果,也不能反映大学生学习公共艺术课程的成长和进步,以及公共艺术教育教师的技能、发展和学校公共艺术教育的文化特色。

## **2. 高校公共艺术教育的意义与功能**

### **2.1 普通高等院校公共艺术教育的意义**

我们在前文已经提到了艺术教育的定义。实际上,艺术教育,尤其是公共艺术教育,有狭义和广义之分。广义上讲,作为高校美学教育教学的核心,公共艺术教育是每一位大学生都应当接受的教育。高校美学教育教学的主要功能是培养大学生的审美素质、文化修养和其他品质,旨在培养出全面发展的高素质人才。由此可见,艺术教育对于当代大学生素质的构建和培养具有重大的意义 [7]。

#### **2.1.1 公共艺术教育有助于提高大学生的思想道德水平**

想要促进大学生受到多种因素的影响,其中思想道德素质具有决定性的意义。在我国普通高等院校实施公共艺术教育,可以使当代大学生潜移默化地接受思想道德素质教育,从而树立正确的审美观,形成科学进步的人生观、价值观和道德观,进而塑造一个健全的人格和健康的心理,最终实现大学生全面和谐发展的目标。由此可见,普通高校的公共艺术教育能有效提升大学生思想、道德、科学文化以及身心健康等整体素质的发展和提高。高校公共艺术教育通过培养大学生的审美意识,进一步提高了大学生的思想道德水平,使得教师能够更好地、更直观地让学生在课堂上体验到真、善、美的存在。普通高校思想道德教育的核心任务,是在普通高校的课堂上对大学生进行系统的道德影响,培养适当的伦理行为。普通高校的公共艺术教育反映了伦理与教育的统一。这种统一通常在大学生不感受外部压力的情况下渗透到他们的内心,使得大学生总是在不知不觉中接受艺术作品的熏陶。这样,高校的公共艺术

教师能够将道德教育融入到艺术作品中，使得大学生更易接受，在作品教学中获得愉悦、净化情感、美化心灵。

### **2.1.2 普通高等院校的公共艺术教育能够拓宽大学生的文化视野并丰富他们的文化修养**

所谓艺术，往往通过典型的艺术形象反映一个时代的生活或者一个时代人们的精神面貌。作为我国新时代的大学生，普通高校的公共艺术教育正好能够弥补当代大学生因知识有限或经验不足而对每个时期的生活或文化不够了解的问题。通过欣赏不同时期的艺术作品，大学生能认识到不同时期、不同国家和地区，乃至不同民族的各种具体而生动的不同生活场景。普通高校的公共艺术教育可以使大学生了解不同时期、不同国家和地区，甚至不同民族的生活场景，以及那个时代生活中的各种人物形象，理解这些人物的性格特点和思想特征。通过公共艺术课程的学习，当代大学生拓宽了生活视野，认识到了社会发展的现实，并了解了人类发展的历史和真理。可见，所有真实的艺术作品都在不同程度和不同性质上具有认知和启发作用，对于丰富大学生的文化视野发挥着举足轻重和重要的作用。

### **2.1.3 普通高等院校的公共艺术教育可以培养和训练大学生的创造能力，提高科学文化知识的理解和记忆能力，促进他们智力的发展**

普通高校的公共艺术教育能使大学生不断提升道德和文化修养，不断完善审美能力，同时也能够使他们不断提高创造能力。可见，普通高校的公共艺术教育增加了大学生对科学文化知识的理解和记忆，促进了大学生智力的发展。在当前的普通高等教育中，我们的教育工作者应该充分发挥普通高校公共艺术教育的优势。学校应开设更多与艺术相关的教育课程，为大学生提供更多的艺术实践活动机会。只有在这些实践活动中，大学生才能充分发挥他们的想象力和创造力，使他们的各方面素质在这些实践活动中得到最大程度的发展和利用。

### **2.1.4 公共艺术教育同样可以有效促进大学生的心理健康教育并提高大学生的个性修养**

在当今社会竞争加剧、社会关系更加复杂、学习和生活压力日益增大的背景下，大学生

的心理健康正经受严峻考验。无论是就业还是升学,许多大学生都存在不同程度的心理障碍和心理疾病。公共艺术教育在提高人的心理素质方面起着重要作用,也是有效解决大学生心理问题的良药。善用公共艺术教育能够缓解大学生的紧张状态,为生活增添愉悦的态度,在公共艺术教育课程中有效地解决心理问题。通过以上讨论不难发现,我国普通高等院校的公共艺术教育是不可或缺的。通过在普通高校开展公共艺术教育,可以培养和提高大学生的整体素质,这也是必要的手段和重要途径。我国高等院校教育中,普通高校的公共艺术教育具有不可替代的独立地位和不容忽视的重大意义。

## 2.2 普通高等院校公共艺术教育的作用

下面,我们将从以下几个方面阐释普通高等院校开展公共艺术教育在促进学生全面发展方面的重要作用:

普通高等教育机构开展公共艺术教育,可以促进大学生思维能力的发展。作为培养大学生想象力的最佳途径之一,高校的公共艺术教育能够有效促进大学生创造性思维能力的发展。在艺术教育的审美过程中,在教师的引导下,大学生通过不断的探索和创造形成多维度和多层次的思维模式。在思维过程中,他们的发散性思维可以得到有效的培养和发展促进。

普通高等教育机构开展公共艺术教育,能够有效提升大学生的身心素质。在普通高校实施公共艺术教育,可以使大学生获得更多的展示和交流机会。通过这些学习和交流的机会,大学生可以进一步丰富内心世界的情感表达,这有助于增强他们的人际交往技巧并扩展自己的人际交际范围。此外,我国普通高等院校公共艺术教育的发展,也能够有效提高大学生的审美建设,培养他们的审美情趣。

普通高等院校开展公共艺术教育,有助于提升整个校园的文化素养。在普通高等院校中,通过举办丰富多彩、艺术性强、健康雅致的文化活动,我们可以有效提高大学生在课堂上学习艺术知识的速度。雅致的艺术活动在普通高校的公共艺术教学与学习中形成了良性循环,使得高校的学术氛围和娱乐氛围在校园文化中有机和谐地结合起来,使普通高等院校的公共艺术教育更好地发挥了启迪大学生思维、净化大学生心灵以及疑惑解惑的作用。

### 2.2.1 艺术教育是全面提高大学生素质的有效途径

对于普通高等院校而言,公共艺术与素质教育是和谐统一的。在普通高等院校中实施公共艺术教育,是素质教育的极其重要的手段和工具,也是培养全面发展的综合性人才的

最有效途径之一。普通高等院校的公共艺术教育能够提高当代大学生的联想能力、想象能力、创造性想象力、抽象思维能力等，并且公共艺术教育能够培养大学生的思维，塑造大学生的性格，调节大学生的情感，增强大学生的智力发展。它在各个方面都有着不可替代的作用。

### **(1) 大学公共艺术教育有助于大学生确立正确的世界观和人生观**

正如我们所知，这个世界上的一切都是错综复杂和多样化的，这很容易对大学生的世界观和人生观产生一定的影响。艺术作品的创作和欣赏能在大学生世界观和人生观的形成中发挥积极而正面的作用。这主要体现在优秀的艺术作品蕴含着艺术家对生活 and 世界的真挚情感。这种情感持续地感染和影响受教育的大学生，激发他们对真实生活的热爱和对美好事物的珍惜。新世纪的大学生通常具有丰富的感情、鲜明的个性、对新事物的敏感和较强的可塑性。高校公共艺术教育的发展，可以让他们通过对艺术作品的审美欣赏获得精神上的愉悦和享受，精神世界的丰富也有助于改善大学生的个性。众所周知，大学阶段是学生的世界观、人生观和审美观逐一成熟的发展阶段。在这个阶段，大学公共艺术教师应通过教育和引导，使大学生理解世界和历史发展的规律。社会历史是人民创造的，这是不可改变的历史法则，大学生也将在这种教育中接受社会是从低级向高级发展的。之所以在艺术美德方面受到教育，是为了发现美、认识美、欣赏美，这是全人类的共同理想。只有这样，大学生才能在公共艺术教育的过程中树立一个充满真理、善良和美丽的世界观和人生观。

### **(2) 大学公共艺术教育可以激发大学生的爱国热情和集体主义精神**

对于整个社会而言，当代大学生是未来发展的强大动力和骨干力量。这种力量必须具备正确的人生方向。因此，大学生必须具有健全的人格和高尚的情操，才能在推动社会发展的过程中追求真、善、美的品质，并发现一个真、善、美的生活，这样才有力量使整个社会和国家的的发展朝着繁荣、活力和生机的方向发展。普通高等院校为大学生提供公共艺术教育，可以不断塑造大学生的完美人格，这不仅是当代社会发展的需要，也是我们民族和国家发展的需要。在新时代，高等院校继续重视精神文明建设，精神文明建设的一个重要目的是树立大学生的民族自尊心和民族自信心。为学校思想教育打下坚实的基础。我们的高等院校必须跟上社会改革的快速稳健步伐，为我国社会经济的腾飞在思想领域打下良好的基础。因此，普通高等院校对大学生的公共艺术教育是激发大学生爱国思想、培养大学生集体主义精神和奠定大学教育民族精神的最佳途径之一。

### **(3) 大学公共艺术教育的发展对于当代大学生增强对科学和文化的热爱非常有益**

在普通高校开展公共艺术教育有助于当代大学生增强对科学和文化的热爱,激发他们对知识的渴望和对真理的追求。公共艺术教学的发展可以有效调整大学生的思维方式,促进他们的观察能力、理解力和想象力的提升。对于大学生来说,公共艺术教育所蕴含的力量可以激发他们对更深奥知识的热爱、对未被理解真理的追求以及对飞速发展科学的探索。大学生利用在大学教育中学到的专业知识,不断地探索、开发和创造,开辟新的途径以获得更多的认知。在当今科学发展中,艺术文化与自然科学的关系已转变为相互渗透和整合。这种高度的精致性和高度的整合性展现了以整合为主流的社会发展特征。可以看出,在当前的高等教育中,我们应培养具有创新思维的复合型人才,结合文科和理科知识,这是我国高等教育国际教育改革创新趋势的标志。长期以来,我国的高等教育普遍存在着重视专业知识教育而轻视综合素质教育的隐患,而艺术素质教育是最缺失的一部分。在我们未来的高等教育中,必须防止大学生只学习高度专业化和狭隘的专业知识而缺乏创业和创新能力的悲剧。可以说,高校的公共艺术教育肩负着提高我国高等教育效果的重要使命,对于大学生精神境界的升华、想象力和创造力的发展、思维方法的拓展都有着不可估量的影响。

### **(4) 大学公共艺术教育可以培养大学生的正确审美观**

大学公共艺术教育能够帮助学生形成正确的审美观念,使他们学会欣赏美、创造美的能力,加强对他们素质的培育,使大学生在道德、智力、体质、美育等多方面实现全面而平衡的发展。同时,培养具有全面发展的人才也是我们高等院校教师的责任和义务。公共艺术教育可以培养大学生正确的审美观念,并培育欣赏美和创造美的能力。在当前,艺术不再是普通大众难以企及的东西,而是呈现出普及化的趋势。

这同样是世界现代史上一个值得注意的现象。可以说,艺术已经变成了人们文化生活中难以触及的存在。遗失的内容。因此,为了提高大学生的文化和艺术素养,培养具有全面发展的综合型人才,大学的公共艺术教育应当得到普及和推广。在我国社会发展的现阶段,培养具有全面发展的人才已经成为我国教育的首要任务,而普通高等学校的公共艺术是我国素质教育全面发展的重要组成部分。艺术的美好为人所共知。这种美可以使大学生在接受教育的过程中感到愉悦和流连忘返。在我国高校公共艺术教育领域,传统国家艺术作品和外国艺术品的利用仍然远远不够。因此,如何拓宽大学生的审美视野,提高他们在普通高校公共艺术教育中的审美品位,已成为高校公共艺术教育中亟待解决的最紧迫问题。公共艺术和审美

观念的形成可以说是血肉相连的联系。作为艺术的重要组成部分，公共艺术为学生带来最直接的美，是真、善、美的有机结合。因此，大学生的公共艺术教育可以使他们沉醉在美的世界中，并且在净化大学生的心灵和提高他们的审美能力方面发挥重要作用。

### 2.2.2 艺术教育是提升校园整体文化品位的有效途径

公共艺术教育是普通高等学校实施美育的主要内容和方式。公共艺术教育具有其他教育方法所不可替代的独特作用，是我国普通高校教育不可缺少的重要组成部分。公共艺术教育是加强社会主义精神文明建设，潜移默化地提高大学生道德水准，培养大学生高尚情操，推动大学生智力和身心健康发展的有利手段。公共艺术教育的发展已经成为教育改革的重要组成部分，它已成为学校素质教育的突破口之一。中外古今的教育实践告诉我们，艺术性思维是将三维思维转化为平面表现或表达，这本身就是一种创造。大学公共艺术教育是培养学生对美追求的最有效途径之一。如果一个人从小就受到美的教育，使他们感知并高度赏识所有美好的事物，那么他们的高尚人格就会逐渐完善，世界也将因此充满爱与美。因此，艺术教育是提升整个校园整体文化品位最有效的方式。大学时期是大学生思想观念和性格形成的关键阶段。公共艺术教育的实施，对大学生思想观念、价值取向和行为方式的形成和改善具有潜移默化的影响。近年来，我国高校不断扩大招生规模，大学生思想活动的独立性、选择性、变异性和差异性特点越来越明显。

艺术文化是人类文明发展的积累。它是卓越、经典、健康和进步精神文化的结晶。因此，在高校中进行公共艺术教育有助于使当代大学生在经济发展的浪潮中具备更深的思想、更厚的文学底蕴和更专注的审美力。这种全面的提升也对大学生完成高等教育后步入社会的发展起到了重要的保障作用。

高校的普遍公共艺术教育是提升校园整体文化品味的有效方式，主要反映在以下几个方面：

从政策角度看，高校公共艺术教育符合我国当前高等教育中的以人为本思想，改变了高校的现有教育观念，使大学生成为校园文化的实践者、开发者和受益者。校园文化对大学生的理念、价值取向和行为产生了微妙而深远的影响。自1994年以来，我已通过法令使公共艺术教育成为基础教育的核心科目。与此同时，相关方面还制定了《国家艺术教育标准》，显示出高等艺术教育在我国基础教育中获得了独立和平等的地位。在我国，艺术教育和高雅艺术的普及尚未真正获得像在国外那样重要的地位，但教育当局和高校在制定和完善校园文

化的政策与措施时应以人为本，制定适合当地和学校发展的政策。艺术教育已进一步纳入规范化教学轨道，并得到大力支持，普及高雅艺术活动。

从技术角度来看，在我国普通高校的公共艺术教育方面，需要积极建立相关机制，增强学校管理层和教职工的公共艺术培养意识。校园文化建设是一个系统而繁琐的工程。在不断推进的过程中，普通高校的公共艺术教育是最重要的组成部分之一。从2005年到今天，我国包括教育部、文化部、财政部等多个部门非常重视普通高校的公共艺术教育，并联合举办了多项公共艺术教育普及活动。其教育效果非常显著。政府对公共艺术教育的重视不仅赢得了师生的赞誉，也得到了良好的社会反响。然而，在欢呼公共艺术教育的发展时，我们还必须清楚地认识到，真正帮助学生理解和接受公共艺术教育不可能一蹴而就。关键是要为高校的大学生提供长期的教育。引导广大大学生培养艺术氛围。可见，我国各大高校在公共艺术教育机制建设上仍需加大力度，大力建立长期机制，确保学生审美素质培养的连续性。

从教育的角度来看，普通高校公共艺术教育的发展在加强大学生艺术品位的引导上扮演着重要角色。此外，它还可以整体提升我国普通高校的艺术品位。在高校中推广高雅艺术将进一步促进当代普通高校的学术繁荣和文化发展。普通高校应从政策和体制的角度激励师资力量开展公共艺术教育，并有效引导学生开展“向着真善美和提升”的审美活动。普通高校还应大力支持并加强各类大学生艺术团体的建立和繁荣的引导，给予学生展示个性、风采和增强自信的机会。

从载体的角度来看，普通高校是公共艺术教育扩展和传播的有效艺术传播载体。无论是视频、网络、报纸还是广播，这些大众传媒都可以成为公共艺术教育和传播的载体。普通高校应承担更多艺术传承的责任和使命，用新的方式宣传和推动普通高校公共艺术教育的发展，使更多的学生能够接触和了解普通高校的公共艺术教育。通过在高校中开展各种学生活动，广大公共艺术教师可以让大学生以身临其境的方式感受艺术的震撼，从而在学校的公共艺术课程中产生极大的兴趣。普通高校公共艺术教育的发展经历了多次传播和积累，现在已经积极融入时代元素，对于高校的发展，已经成功地从“进入校园”转变为“繁荣校园”。

### **3. 解决普通高等院校公共艺术教育问题的有效对策**

#### **3.1 从学生兴趣出发，把握学生主体特征**

众所周知，兴趣是学生最好的老师，也是学生学习艺术课程的基本动力之一。如果学生对公共艺术课程不感兴趣，只是被动地跟随老师学习，将直接影响公共艺术教育的效果。因

此，高校的公共艺术教师应充分发挥艺术的独特魅力，运用多种教学方法来激发学生对公共艺术学习的兴趣。激发学生对公共艺术教育兴趣主要体现在明确目标上，让大学生清楚地了解学习公共艺术课程是多么巨大的动力。同时，公共艺术教师应在了解学生在艺术课程的专业水平和文化课程的成就基础上，为学生的公共艺术课程发展水平制定清晰、严谨和适宜的学习目标与学习计划，让大学生在公共艺术课程的学习中有章可循，对不同的学生实施因材施教、循序渐进的教学方法。只有这样，我们才能根据不同学生的特点制定出能够激发学生学习兴趣的教学方法，让大学生在原有艺术的基础上轻松完成他们的短期学习目标和长期学习目标，从而在原有艺术的基础上取得进步。为了更好地激发学生的学习兴趣，普通高校的艺术系可以定期组织学生开展“文化周”、“艺术节”等活动，组织学生自编自导自演，进行校园广泛展示成果。有了这样一个表演舞台，学生对艺术学习的理解会更加深入，也能激励更多的学生对艺术学习产生兴趣，从而达到把握学生主体特征进行高校艺术教育的目标。在学生的艺术教育和学习过程中，教师应积极创造条件，使每个学生都有机会在艺术课堂上表达和锻炼，让学生有成功的体验，从而激发他们对艺术学习的兴趣。教师可以组织学生举办展览，让学生欣赏艺术作品，与他人分享创作和体验，从而认识到自己的位置和差距，调整艺术学习的目标和计划。此外，根据当前学生的心理特征，公共艺术教师应努力为学生创造一个公平竞争的氛围，让学生的优势在各种比赛中得到发挥，这不仅实现了自我价值，也激发了他们学习艺术的热情。

### **3.2 加强师资队伍建设和，充分发挥教师主观能动性，**

#### **做好普通高校公共艺术教育工作**

普通高校的公共艺术教育工作离不开一支高素质教师队伍的支撑。只有当大多数教师树立正确的教育理念，充分发挥自身的主观能动性，他们才能坚持以人为本，推动素质教育，将教育改革与发展的战略主题实施到普通高校的公共艺术教育中，从而实现党的教育方针政策，满足当代社会的要求。我国早期现代著名教育家蔡元培先生认为艺术教育具有“陶冶情操”的作用，能够以美启真，以美引善。通过蔡元培先生的理论，不难发现，在当今以人为本的思想指导下，普通高校的公共艺术教育不仅是教育自身规律的需要，也是我国社会发展的必然要求。从2002年9月1日起，我国教育部计划在全国范围内实施《学校艺术教育工作规定》，其中明确了学校艺术教育的目标。面向现代化、面向世界、面向未来，实行面向全体学生、提供分类指导、因地制宜、注重实践的政策。

可以看出,通过普通高校的公共艺术教育,大多数公共艺术教师应该让大学生充分了解我国现有的优秀民族艺术和文化传统。在教学过程中,教师应积极向学生介绍国内外的优秀艺术成果,并借鉴这一过程,努力提高学生们的艺术鉴赏水平。通过参加艺术鉴赏培训和高校间的交流活动,提高了普通高校艺术教师的素质和水平,进一步发挥了他们的热情。首先,根据高校艺术教师的兴趣爱好,应持续组织高校艺术教师外出学习、参加函授课程等方式学艺,以便高校艺术教师能进一步发展自己的艺术特长。从各个方面出发,对高校艺术教师的成就给予奖励和激励,使普通高校的公共艺术教师能提高在艺术教学中的热情和创造性。一方面,从教师的成就需要出发,实施成功激励。我们相信,引导教师创造成就,支持教师取得成就,并为普通高校的艺术教学提供设备支持和财政支持,以满足普通高校艺术教师的成就感,完成对艺术教师最大的激励。例如,创造条件让艺术教师的经验论文在会议上交流,发表在报刊上,并在教师中推广,给艺术教师加入各级教育学会和研究会的机会,并优先考虑专业职称评定和论文奖励的机会。赋予他们,从而营造一个“只要有行动,你就有地位,只有行动才能有地位”的积极氛围,让学有所成、做有所为的艺术教师感受到被关怀和激励。另一方面,为了满足教师发挥自身优势和创造自我特色的需要,学校投入了大量的财物资源,为他们创造了更好的条件。

### **3.3 在普通高校做好公共艺术教育工作必须贯彻科学发展观**

在高校的艺术教育中,“人”无疑坚定地占据着主导地位。无论是教育的发起者还是接受者,人都是核心单元。因此,如何在艺术教育中做到以人为本,高中艺术教师喜欢歌曲和合理分配教学资源,高校艺术教育必须在科学有效地进行艺术“渗透”方面下足文章,以便进一步形成学校、教师与学生之间积极互动的良好局面,在此情形下,实现人性化教学已成为当务之急。

首先,高校艺术教育的参与应该是广泛的。作为高校艺术教育的直接受益者,大学生的广泛参与毋庸置疑,但目前应该把学生从课堂中“解放”出来,让当代大学生的艺术教育观念从纯粹为了学分而上艺术课的被动行为,转变为通过在校接受艺术教育培养情操和愉悦心灵的主动行为;另外,应积极寻求“大教育”的局面,通过动员各种社会资源,使高校艺术教育不仅局限于校园范围内,也应实现请进来和走出去的局面,使参与者的广度能扩展到广泛的艺术圈乃至整个社会。另一方面,艺术类别的广度应基于提高和加强学生的综合能力和

知识覆盖面，以艺术为主要知识结构，涵盖音乐、美术、戏剧、舞蹈、书法等不同的艺术类别。强调艺术与文化意义上的基本艺术概念，并从广阔的艺术视角和视野思考新问题，使学生们能在基础学科的认知过程中广泛接触到各种类型艺术的经典之作，理解并掌握各种艺术形式的语言表达方式，为自己的专业修养寻找参照和对比。我国普通高校传统的公共艺术教育模式主要是基于课堂灌输经验。大多数教师在课堂上给学生传授理论知识，课堂氛围非常严格，学生难以接受。

随着多元化教育形式的不断发展和公共艺术教育的不断深化，高校教学资源的整合越来越受到重视。正是因为如此，普通高校在培养学生人文精神、道德修养、思维智力等方面的特殊功能越发突出。众所周知，我们伟大的中华民族拥有悠久而辉煌的历史文化遗产，它们是不可替代的丰富和艺术宝藏。目前，我国的文化艺术正不断朝着多元化竞争的方向发展，稳步前进，这要求普通高校公共艺术教育中传统文化与现代文化之间的关系不断整合与合理化。鉴于当前普通高校大学生身心发展的新特点，普通高校的公共艺术教育不仅要关注现代艺术作品的欣赏，而且要不断比较过去艺术作品的欣赏，使当代大学生能理解到每个艺术类别的发展起源、历史背景和社会功能。通过这样的分类学习，学生们在接受现代文化的同时，不忘光辉的历史和文化，充分发挥多媒体的教学功能，调动学生的视觉、听觉、声音等不同感官，激发学生的学习兴趣，在课堂上实现不仅可以与世界文化艺术面对面交流和沟通，从而全面提高学生的综合艺术审美能力；很多当代大学生在孩童时期就学习了乐器演奏、舞蹈、书法和绘画等艺术技能，但对艺术的本质理解和掌握尚待提高，所以让他们从浅入深，从感性到理性，达到更高的审美境界。同时，我们必须使他们成为具有创新思维、创造能力和健康审美的新一代。

### **3.4 为了在普通高校中做好公共艺术教育工作，必须在体制上进行改革**

面对现代科技的冲击，我国普通高校公共艺术教育存在诸多问题，我们必须在体制上进行创新。为了改变我国普通高校现存的落后现象，我们不仅要学习和借鉴国际先进的经验，还要根据中国高校的特点，构建适应现代科技发展的公共艺术教育体系，以确保公共艺术教育的顺利实施。

### 3.4.1 明确普通高等学校公共艺术教育的教育目标和价值标准

公共艺术教育与专业艺术教育之间存在显著差异。普通高校的公共艺术教育并不是用专业艺术教育的标准来培养和评价学生，而是通过普通高校开展公共艺术教育课程，有效培养学生的欣赏能力，提高大学生的艺术判断力，并增强大学生的艺术素养。与普通高校的公共艺术教育相比，专业艺术学院通常过于注重专业技能培养，这对公共艺术教育来说显然过于艰难。对于非艺术专业学生的公共艺术技能的培养，不能使用其他学科的固定标准来衡量普通高校非艺术专业学生的艺术素质。我们只能说，普通高校公共艺术素质教育的目标是通过艺术塑造大学生的灵魂，并全面提高大学生的素质。普通高校中的公共艺术教育是一种具有特殊性的教育，它应遵循普通高校公共教育的规律以及公立高校艺术教育的规律。普通高校的公共艺术教育在整个高等教育中扮演着独特的角色。它集中了音乐、艺术、书法、影视、舞蹈等艺术手段和艺术教育内容，不断培养大学生感受美、欣赏美和创造美的能力。我国高校的公共艺术教育应旨在促进大学生的全面发展，提高综合艺术知识水平。只有实现这一目标，我们才能满足现代社会发展的需求，使普通高校的公共艺术教育与时俱进，满足我国素质教育的要求[8]。

### 3.4.2 在高等院校中构建合理的公共艺术教育课程体系

目前，我国普通高等院校的公共艺术教育正处于一个新的发展时期。在进一步推广这种教育的过程中，加快构建一个创新的、符合时代新要求的课程体系尤为重要。在高等院校中构建一个合理的公共艺术教育课程体系是非常有必要的，因为科学课程体系建设的核心目的在于提高人民素质，并培养学生的创新精神和实践能力。在这一理念的指导下，普通高校的公共艺术教育应该通过具体的课程教学目标、课程结构、课程标准、教学过程、教材的开发与管理，以及相关课程改革的组织和实施来进行渗透和落实。课程目标是学校教学工作的核心，也是实现教育标准和教育目标的保障。普通高校公共艺术教育课程目标的确立应以审美教育为核心，关注学生感受和欣赏艺术之美的能力培养，使他们形成健康向上的审美情趣和正确的审美观念，并通过艺术审美教育拓宽知识视野，培养创新精神和实践能力，最终塑造高尚完美的人。当前，我国高校公共艺术教育的课程设计仍存在问题。为普通高校公共艺术教育构建一个合理的课程体系，有利于为实现公共艺术教育标准和教育目的提供坚实保障。这个课程体系的构建主要体现在课程设置、课程结构、教学方法以及教材建设。鉴于当

前我国公共艺术教育的教育现状，课程体系仍需要大幅改进。课程体系的建立和完善不可能一蹴而就，必须是一个从普及到提高，从低水平到高水平的过程，为不同层次的学生提供相应的课程，并在课程内容的选择上精益求精。

### 3.4.3 在普通高校中创新公共艺术教育的科学教学方法

首先，从学生的兴趣出发，注重艺术欣赏能力的培养和提高。一方面，高校的公共艺术教育教师应向学生介绍相关的公共艺术教育材料，帮助他们了解作品的作者、其个性、生活经历、时代背景、宗教知识等。另一方面，教师在教学过程中应持续激发学生对这些经典公共艺术作品运用想象力，加深个人理解，并培养学生欣赏艺术作品的 ability。

其次，高校的公共艺术教育应与应试教育不同。应该放弃评估和标准答案。艺术教育的目的应该是发现和欣赏美，从而充分激发大学生的个体审美热情和潜能。普通高校的公共艺术教育实际上是一种欣赏教育，实质是一个理解事物和把握创作者意图的过程。公共艺术课程中的这种欣赏教育最终不仅会丰富学生的审美情趣，拓宽他们的文化视野，还会使当代大学生深入到艺术的本质，使他们知道如何用自己的眼睛在生活中感知艺术之美。

第三，在教学形式上，普通高校的公共艺术教育课程应与大学生的学习特点充分结合，解决大学生当前关注的审美问题。应安排不同形式的教学，避免公共艺术教育中课程选择缺乏和形式僵化的问题。

第四，积极开展“第二课堂”等活动，增强公共艺术教育课程的实践性。公共艺术教育不应停留在教师授课和学生听课的层面，而应鼓励学生亲自参与其中。任何形式的艺术教育都有其自己的语言表达。只有掌握了这种语言表达，我们才能真正理解艺术作品，理解艺术的内涵，并提高欣赏水平。因此，只有通过积极开展“第二课堂”等实践活动，组织大学生更多地动手和动脑，亲身体验公共艺术教育过程中的情感活动，然后将艺术之美的欣赏与自己的实际经验相结合，这样的艺术教育效果才是最好的。

第五，通过艺术欣赏教学，使学生掌握艺术欣赏的基本方法。只有当学生掌握了更多科学的方法，他们的审美能力才能得到提高。对比是艺术欣赏的一种重要学习方法，特别是作品的艺术风格和技巧，比如不同绘画和风格的作品，往往难以用言语清楚地描述，但如果把这些作品放在一起进行比较，它们的艺术特征就可以轻易区分。在艺术欣赏中，不要孤立地看待一件作品，而是要将其与其他作品联系起来，从多个方面和层次进行比较，以此来激发和指导学生，并调动学生积极学习。

### **3.5 为了在普通高校中做好公共艺术教育，必须提高编写艺术教材的能力**

目前，我国普通高校公共艺术教育教材编写存在许多问题，其中最突出的表现是缺乏统一编写的高校公共艺术教育教材。在我国普通高校编写公共艺术教育教材的过程中，缺乏教材的整体设计，公共艺术教育教材的编写不够有针对性，往往偏向单一。此外，在编写过程中还缺乏相应的管理体系，确定编写目标时也存在单一性。在公共艺术教育编写中，各自为政，教材的选择和编排差异大，这增加了教学难度，大大降低了我国普通高校公共艺术教育的教学效果。教材内容方面，公共艺术教材应与小学及中学生和艺术专业学生使用的教材不同，属于普通大学生的通用艺术教材。如何准确把握知识的广度、内容的深度和技能训练的选择等问题是很难的。过分关注艺术学科知识内容，缺乏跨文化、跨学科内容和现代媒体技术，存在明显的知识倾向，缺乏对学生需求和兴趣的考虑。

### **3.6 为了在普通高校做好艺术教育工作，必须建立和完善科学的评价体系**

对于普通高校的艺术教育而言，建立和完善科学的评价体系是普通高校艺术教育的有力保障。普通高等教育机构艺术教学评价标准的合理性将直接影响教学评价结果的准确性。因此，只有建立并不断完善评价体系，才能合理地实现艺术教学评价的预期目的。在一般意义上的教学中，评价方案的指标体系包括普通高校艺术教学的指导思想、艺术教学的教师队伍、艺术教学的教学条件、教学建设与改革、教学管理、学风、教学效果以及特色项目等。在艺术学院的艺术教育设计教育评价体系的评价标准创新实践是基于教育部对普通高校教学水平的评价体系。因此，我们必须建立一个符合中国艺术教育现实、能够鼓励艺术院校形成办学特色的评价体系。对于不同类型、层次和规模的普通高校，需要根据自身艺术教学的发展建立适当的教学评价体系，单靠统一的评价指标是不够的。艺术没有固定的评价标准，欣赏水平也是各不相同。如果机械地使用相同的评价标准，很容易扼杀学生学习艺术的兴趣和才能。相反，这并未达到艺术院校评价艺术教学的目的。

## 4. 我国普通高校公共艺术教育的发展

### 4.1 我国普通高校公共艺术教育的发展潜力

近年来,随着国家在高校全面实施素质教育、深化教育改革,素质教育一直是高等教育关注的重点。面对激烈的竞争,各高校越来越意识到我国普通高校公共艺术教育的巨大发展潜力。当前社会和个人对大学生个人素质的要求越来越高,而公共艺术课程能有效提升大学生各方面的能力和素质。在素质教育中,审美教育一直是“德智体美”四育中比较薄弱的环节,但随着人们对审美教育理解的逐步深入,艺术教育已成为我国教育政策的重要组成部分,得到了前所未有的关注。从1999年到2002年,我国的公共艺术教育一直在不断发展,这可以从各种文件的发布中看出。我国教育部先后发布了《学校美术教育工作规定》、《国家学校美术发展纲要(2001-2010)》、《全国普通高校公共艺术课程教学指导计划》等文件。这些文件再次强调了公共艺术教育的发展,必须合理化公共艺术教育课程的教学结构。公共艺术教育与教学的实践、实现公共艺术教育教育思想的基本途径、公共艺术教育的课程设置都表明,公共艺术教育在未来拥有无限的发展潜力。我们知道,公共艺术教育课程并非由教师在课堂上简单指导学生欣赏,而是根据学生不同的专业方向和发展特点激发学生的潜能。设立艺术公共必修课和艺术公共选修课是为了面对未来越来越多的普通专业大学生。高校的公共艺术教育课程必须从各个角度开设,以寻找更丰富的发展路径。普通高校是我国高等教育的脊梁。在普通高校实施公共艺术教育,是我国发展素质教育的一个重要途径。普通高校艺术教育的问题关系到学生个性和整体审美素质的提升,关系到国家精神文明建设问题,关系到和谐社会的健康发展。可见,在未来很长时间内,普通高校公共艺术教育的发展仍将占据大家实现,吸引更多教育学者的关注。高校公共艺术教育的发展是无限的,它将在越来越多的方面对当代大学生产生各种影响[10]。

### 4.2 我国普通高校公共艺术教育的未来展望

从20世纪末到21世纪初,学科综合已经在基础教育中的各种艺术课程改革与发展中成为一个热门议题。这一点在高校的公共艺术教育中得到反映,以高校的一般艺术教育研究为核心,结合音乐、舞蹈、美术、戏剧等姐妹学科,将各种公共艺术课程整合到一起,进行全面的学习。这种类型的公共艺术教育将得到普遍接受。全面研究国际公共艺术教育的发展趋势,不难发现我国高校公共艺术教育的基本现状尽管存在一定问题,但在问题中我们也看

到了它的巨大发展潜力。我们有理由相信，普通高校公共艺术教育课程的专业化和精细化将导致形成一个各艺术类别的重要联盟，这是公共艺术教育未来的必然发展方向。如果高校的公共艺术教育能够及时转变自身，真正占据高校素质教育的一席之地。因此，高校的公共艺术教育工作者应该承担更多的责任和义务，使高校公共艺术教育成为高校素质教育的领导者。

尽管我国普通高校的公共艺术教育已积累了相当丰富的经验，艺术教育理论也日益丰富，但大多数理论研究仍然是基于相对宏观的视角，并依赖于现有的研究文献。在已有的研究成果中，对高校公共艺术教育实际实践的研究与各方面存在问题之间仍有较大的差距。因此，本节的作者主要结合我国普通高校公共艺术教育的概念与现状，运用问卷调查、访谈和文献研究等多种研究方法，并充分结合我国高校多年的教学经验。通过撰写论文过程中的讨论与总结，对教学实践经验进行对比分析，综合分析了大多数国内高校公共艺术与教学的相关信息与材料，并结合高校公共艺术课程特别是课程目标、课程内容、课程设置、课程评价等方面进行了调查研究，规范、梳理并系统化地结合横向比较与调查分析方法，对公共艺术教育的目标、内容、实施、意义与功能进行了研究，进一步明确了现存问题，并提出了针对性的对策和建议。

## 参考文献

- [1] 程静. 当代艺术审美判断标准的几点思考——读《当代艺术的争议》有感. 神州, 2019(9):1.
- [2] 程诺. 市场导向环境下当代艺术美学的生存与发展. 石家庄学院学报, 2021, 23(1):5.
- [3] 包佩佩, 方伟. 当代艺术美学下的模糊美感. 艺术与设计: 理论版, 2019(5):2.
- [4] 倪朝然. 当代艺术美学及中国本土艺术的现代发展历史与现状分析. 艺术品鉴, 2020(2).
- [5] 王杰, 孟凡军. 美学人类学: 构建当代美学与艺术批评的新体系. 社会科学家, 2020(5):9.
- [6] 王帆. 中国当代艺术中的可爱美学研究. 河南大学, 2019.
- [7] 段云东. 从“艺术的终结”到“领域的扩张”——美国“十月派”艺术写作与美学的新进展. 西南大学学报: 社会科学版, 2019, 045(002):157-171.
- [8] 高凌予. 艺术美学与生活美学的反思. 视界, 2020.
- [9] 翟楚霞. 异域新声——从跨文化视角看意大利当代艺术展. 美术, 2018(7):3.
- [10] 段云东. 从“艺术的终结”到“领域的扩张”: 美国“十月派”的艺术写作与美学的新进展. 西南大学学报: 社会科学版, 2019, 45(2):16.

## 第 2 章 艺术与技术

### 第 1 节 在技术与艺术背景下的新媒体叙事方式探讨

#### 1. 技术与艺术

##### 1.1 技术与艺术的关系

凯文·凯利 (Kevin Kelly) 曾说：“技术和艺术创造了新工具，新工具引领新艺术，新艺术再次孕育新工具，这是一个无限循环”，正如技术与艺术之间的关系，从大约 15000 年前法国拉斯科洞穴的壁画到今天的 3D 打印艺术、虚拟现实、AR 技术、生物基因等，艺术创作中艺术媒介的转变对图像创作和叙事方式产生了巨大的影响和变化。

技术的进步、人们适应能力的提升，以及由此产生的概念和习惯，使得古代那些美丽的艺术经历了深刻的变革。在所有艺术中，有一种现象无法再像以前那样观察和对待物质，因为这些物质也受到现代科学和实践的影响。在过去的 20 年里，无论是物质还是空间和时间，都已经不是过去的模样。我们期待艺术的伟大创新能够改变整个艺术技术，从而对艺术创作产生影响，并最终以迷人的方式改变我们的艺术观念。这是保罗·瓦勒里 (Paul Valery) 在《无所不在的征服》中所表达的观点。正是技术推动了工具的转变并产生了新的艺术形式，比如数字艺术、装置艺术、视频艺术和互动艺术。新的世界观察方式和新艺术促进了新工具的产生，例如在科幻电影中出现的机器人。随着科技的发展，现在你只需要学会使用 Arduino 编程软件和一些传感器，机器人爱好者也能制作简单的机器人；一种新的艺术流派，觉醒主义，激励科学家创造更加“智能”的计算机，艺术家运用他们的想象力将两幅经典杰作“混搭”创造出新的艺术效果，受到 Google 赞助的创意团队开发了一项新技术，允许计算机“混搭”两幅图片并创造自己的作品[1]。

总结来说，随着数字时代的到来，技术与艺术之间的关系变得越来越密切。技术与艺术不可分割，艺术也离不开技术。没有科学的发展，就没有科学与艺术的提升。科学技术的发

展为艺术家提供了更多的媒介和创作可能性，同时，艺术也为科学技术提供了想象力和艺术思维。在技术的推动下，21世纪的艺术已经覆盖了数字艺术、生物艺术、虚拟现实等。借助技术的帮助，艺术家转变了艺术创作的媒介，为人们带来了微观世界和超现实主义等体验。这些作品关注现代技术和新媒体的使用，被定义为“新媒体艺术”。

## 1.2 艺术创作中的概念与媒介

### 1.2.1 概念与媒介

媒介作为艺术创作工具之一是非常重要的，但概念更为重要，概念指导媒介的使用。马歇尔·麦克卢汉（Marshall McLuhan）等人认为“衣服是人类皮肤的延伸，车轮是脚的延伸，相机和望远镜是眼睛的延伸，技术是身体和概念的延伸。”那么，艺术就是概念延伸中的一个分支。艺术与技术之间的关系一直在讨论中。艺术家创作的概念是艺术家情感记忆的反馈。艺术家在其作品中展现自己的个人情感、信仰、认知和哲学背景，以探索人类身份。通过艺术重新理解世界、自然和自我是一种方式。艺术在很大程度上激发了人们的思考，指的是引导人们的创造性思维。因此，作品可以分为两类。一类是只有艺术家个人情感和审美认知，这是艺术家的“共情”。在与情感融合的过程中，南齐画家谢赫在其著作《古画品录》中提倡艺术创作应该反映艺术家自己的思想和感受。传统艺术作品具有这样的特点，艺术家表达情感，观众可以从作品中感受到艺术家带来的“美”的体验，或者感受到艺术家情感带来的情感。

### 1.2.2 媒介与记忆

艺术家与观众之间的互动，艺术家通过技术传达思想，这种互动往往源于观众在体验艺术家作品时的共鸣，新技术和手段可以提供更多这样的便利，这是从艺术家个人的情感记忆升级到群体情感记忆的过程。莱克认为，“保留印象是记忆的功能，回忆则会使其瓦解。本质是记忆是保守的，而回忆是解构的”，并且“意识只能在有记忆痕迹的地方出现”。越来越多的艺术家在他们的创作中将‘记忆与经验’看作是一种重要的艺术表达。

珍妮特·卡迪夫（Janet Cardiff）的40声部赞美诗，展示在大都会博物馆的修道院分馆，其作品的核心是一个十二世纪的小教堂，其中40个扬声器在修道院中以环形排列，每个扬

声器都发出声音。这件作品通过创新的媒介使用，以声音艺术的形式，激发了观众对于记忆和空间的深层次体验，使得个人的感受在群体中得到了共鸣和放大，展现了媒介与记忆在艺术创作中的重要性。

托马斯·塔利斯于 1579 年创作的宏大合唱《除主外，无人可信》的一段，卡迪夫选取了西方音乐中最为壮丽的杰作之一，（图 2 珍妮特·卡迪夫 4. 赞美诗的一部分）将其分解成普通的独立声部，在你站在单个扬声器前时，你将听到一个单独的声音，但当你站在整个作品的中心，听到 40 个声部合并在一起时，这种宏伟的效果，两者形成鲜明对比，许多人因此而流泪。这项作品具有强烈的集体记忆情感。如果不是基督教信徒，应该很难产生共鸣。集体情感记忆驱动着所有来参观现场的观众。技术创新允许艺术家在表达作品时使用扬声器、传感器等技术手段，带来了传统油画画架无法提供的体验。

通过技术创新改变观看方式的一项重要技术是 VR。VR 带来的视觉体验是前所未有的。加上立体环绕声，它为观众带来了全新的视觉体验和艺术体验。虽然它还不够成熟，但我们努力的年轻艺术家试图将这种体验融入到艺术作品中。在一次 TED 演讲中，艺术家利用 VR 技术，将战后的一个小女孩带到你的眼前。通过 VR 眼镜，你会感觉自己站在她的对面。这种视觉冲击比报纸、电视和计算机带来的阅读类型或视频类型的冲击更为直观。通过想象一个战后带着悲伤和荒凉站在你面前的小女孩，比观看视频带来的冷漠更容易给人留下深刻印象。技术为我们的艺术创作带来了更多的可能性，艺术也为技术提供了更多不受约束的空间。技术正在将其变为现实。就像法国 19 世纪作家福楼拜所说：“艺术越来越科学化，科学越来越艺术化，两者在山脚分开，总有一天他们将在山顶相会。”

### 1.2.3 新媒体的重要性

自工业革命以来，科学技术突飞猛进，艺术创作的媒介也随着时代的更新而发生了变化。因此，技术对媒介来说至关重要。摄影在 1939 年诞生，相机的出现带来了摄影艺术，当摄影艺术出现时，人们说绘画死了，因为在此之前，艺术创作是油画、版画、壁画和雕塑的时代，无论是古典主义、巴洛克还是印象主义。绘画并未死去，从印象主义到现在，包括立体主义和超现实主义，绘画恰恰展示了那些被摄影所回避的物理和心理元素。随着科技的发展，在第一台相机出现后，艺术家解放了双手，能在最短的时间内带来艺术作品。而且，这类作品具有强烈的复制性，这在同期其他艺术表现方式中是无与伦比和超越的。艺术家们也开始了新的尝试，特别是在相机得到改进、便于携带和私人化之后，摄影成为艺术家表达艺术的

重要方式。相机的出现改变了艺术家观察世界和叙事的方式，改变了创作媒介，使得新的艺术视角出现在观众面前。艺术家们大胆使用新媒体，站在前沿。他们的嗅觉极为敏锐。录像机和电视机的出现为视频艺术奠定了基础。向我们展示了他们的疯狂和激情。媒介本身具有时代的特征。Fluxus 运动的一员：韩裔美国艺术家南·准·派克，是视频艺术的大师。尽管南·准·派克学习的是古典音乐，但在遇到约翰·凯奇后，受到约翰·凯奇的实验音乐影响，他将视频和电视机变成了装置作品。南·准·派克在他的作品“电视佛陀”中也加入了禅的色彩，表面上看“电视佛陀”指的是电视机在人们家中所占的位置，取代了家庭祭坛中佛像原有的位置。南·准·派克曾对记者说，他是一个社会思想家，所以我们将电视媒介变成高艺术。知识的爆炸和融合，视频技术和电视的出现为南·准·派克提供了物质基础。他对禅宗和噪音音乐的研究，确立了具有自己时代特征的思想。意识与物质的结合创造了这位视频艺术家。

中国的第一部视频艺术作品由被称为“视频艺术之父”的张培力在 1988 年创作。事实上，很难说这项作品在精确意义上是“视频艺术”，因为它使用视频作为工具来记录和再现一场行为艺术，这部 180 分钟的视频作品现在看起来异常枯燥乏味，因为它是在摄像机前不断地打碎、粘合、再次打碎一块玻璃的过程。但正是这样一个实验过程，选择录像机作为创作媒介的方法，为中国的视频艺术打开了大门。这不仅是媒介的创新，也是中国艺术家意识的创新。张培力在后期放弃了绘画，转而在视频、摄影和装置进行艺术创作。他的作品侧重于展示人与世界关系的悖论。他以一种孤立而中立的立场，自由而深刻地思考人与世界的关系。创造了一个庞大、综合且强有力的艺术关系，促进了中国新媒体艺术的发展。在科技背景下诞生的媒介具有独特的魅力。它是一种超越所有已知类型的艺术形式。这种新的影像形式或叙事方法是以前从未感知过的。这种魅力将吸引观众的视觉、听觉和心理记忆。记忆是大脑与物质的结合。作为一种物质，媒介本身的特殊性将导致不同的思考，因此艺术家们会尝试使用新媒介来传达他们的观点。

随着科学技术的影响，艺术的边界已经极大地扩展，包含了生物学过程、遗传工程、人工智能等技术，创造出新的表现形式。艺术与科学技术的交汇孕育了新媒体艺术形式，比如微生物艺术、遗传艺术，以及整合了先进技术如人工模拟装置和磁悬浮的作品。现在的艺术家们正在探索使用生物媒介作为他们创作表达的画布和工具。

## 1.3 新媒体的一些叙事方法

### 1.3.1 “重复性”叙事

在艺术作品中，“重复”一直作为一种艺术表现技巧出现。反复的重复本身就是一种力量的状态，因此许多艺术作品通过重复一个图像或概念来强化作品的力量，强调主体的重复性。多斯特效应本身就是重复的一种表现形式，意味着图片的某个部分与整个图片相同，并通过递归的视觉形式产生无限循环，使得图像在时间和空间中产生奇妙的变化。通过重复，可以使主题更加清晰和统一，视觉冲击力更强，可以增强视觉艺术效果，给观众留下深刻印象。实际上，在古代陶瓷艺术作品中就会有重复的图案。符号的重复带来审美的兴趣；版画艺术本身就是重复的创造。安迪·沃霍尔重复的罐头、重复的梦露玛丽莲反映了那个时代工业革命带来的影响，艺术作品也是通过流水线作业生产的，放弃了艺术技巧和原创性。他的所有作品都是复制品，他想用复制品取代原创。草间弥生的重复点点出现在她的绘画中，将自己的精神状态外化。无论是绘画还是装置，它们都充满了童真和令人眩晕的效果。艺术可以是无聊的吗？或者无聊可以是艺术吗？无聊，一个形容词，是人作为主体时的一种状态。在心理学中，无聊是当关注的对象不符合个人价值观时的心理体验。阿瑟·叔本华曾说：“生活是一大堆欲望。如果欲望得不到满足，就会受苦。如果得到满足，就会无聊。生活在痛苦和无聊之间摆动。”生活中充满了无聊，它也出现在许多文学作品中，比如大卫·塞林格的文学作品《麦田里的守望者》的主人公：霍尔顿·考尔菲尔德，一个闲散的中产阶级，没有什么是他习惯看到的，而正是这样一个人，却赢得了无数读者的喜爱[2]。

### 1.3.2 “无聊”中的艺术

“无聊”是现代人的专利。人们的无聊感一天天增加。无聊从边缘现象蔓延到现代社会的每个角落。无聊是一种独特的心理特征。原因可以归纳为两个主要方面。一是自我调节情绪的能力，二是外部环境的刺激，所以当观众观看艺术作品时，我们可以总结出两种现象：首先，观众不理解艺术家想要表达什么，不符合他们预期的价值观。会有一种厌倦的心态，当你感觉到作品不知道它在说什么时，你会感到无聊。其次，艺术作品本身想要表达无聊的想法，但它透露出另一个观点，这种无聊变得非常有趣。许多艺术作品都是这种情况。例如，

一些画作只有简单的色块。当你没有与颜色产生心理共鸣时，你对作品的认知可能会导致无聊。然而，本文是讨论技术和艺术背景。接下来，新媒体艺术创作对艺术家艺术创作的艺术影响和方法，比利时出生、早期在墨西哥工作的建筑师艺术家弗朗西斯·艾丽丝 (Francis Alys) 开始于 1991 年作为艺术家参加展览。弗朗西斯·埃利斯的艺术作品是在户外创作和完成的，他对日常生活有非常细致的距离观察。使用绘画和表演、摄影、纪录片和视频，创作的作品具有强烈的个人风格，同时反映了他对当前社会经济环境的真实记录和反思。弗朗西斯·艾丽丝关于他创作的观点是：“我花很多时间在城市里走来走去，许多作品的概念在走动时突然出现。作为一名艺术家，我的艺术相当于移动。我的作品是记录和成功的指南。语言的创造和城市的创造齐头并进，我的每一个创作都是我正在创造的另一个故事的片段，是我试图绘制的城市地图。”他的许多作品都与走路有关，在城市的街道或贫民窟的土路上等。关于走路，加布里埃尔·奥罗兹科 (Gabriel Orozco) 和莎拉·斯泽 (Sarah Sze) 都是在走路时受到启发创作艺术的。加布里埃尔·奥罗兹科的作品“Penske”在纽约市捡拾被丢弃的垃圾并重新组装。拍照的方式记录了这一刻纽约场地的气氛。加布里埃尔·奥罗兹科的作品同样无聊且荒诞，看似简单的装置或摄影，总是透露出他处理物体的变化无常的方式，偶尔创造出随机特性或完全相反的方向。萨拉·斯泽也以使用牙签、海绵、灯泡、塑料瓶等日常物品的现场特定装置而闻名，并且她为高线公园(High Line Park)闲逛的人设计了一种现代鸟类栖息地 - 仿生木头覆盖的鸟舍。萨拉·斯泽对材料和环境的敏感性把材料重新带回到艺术家自己的价值观，并专注于形式之美。这是她与其他艺术家的一个重大区别。艺术家对步行的热爱总是一个接一个。日本声音艺术家铃木俊南的“声音的踏板”即是以这种形式进行创作。

尽管弗朗西斯·艾丽丝的作品是基于行走的，但在技术的指导下，弗朗西斯仍然采用视频作为展示他作品的方式，比如他的代表作《实践的悖论》。1997 年，弗朗西斯的作品以视频形式呈现在观众视野中。视频展示了弗朗西斯弯腰努力在墨西哥城推动一个巨大的冰块前行。后来，冰块慢慢地向前移动，它融化了，变得小了一点，变成了冰块，然后是鹅卵石，最终变成了一滩水。弗朗西斯也从一开始的艰难推动，变成了在城市小巷中穿梭的游戏。整个作品花费了艺术家大量的时间，艺术家没有取得任何实际结果，这只是对他行为的记录。艺术家这种徒劳的努力似乎很无聊，但也赢得了认可。

另一件作品《排练》发生在一条陡峭的土路上。一辆小红色甲壳虫在鼓乐队的伴奏下一次又一次地爬上坡，然后再滚下来。这部作品长达 29 分钟又 34 秒，画面中有车辆经过、狗吠声和路人穿梭。最终，汽车没有越过山坡，而是选择了后退，退出观众的视野。这种无聊的荒谬让观众感到有趣又悲伤。在艺术作品中，事件的无聊不一定是无聊的事情，就像硬币

的两面。弗朗西斯·艾丽丝的作品《无聊》就是同性质的，但它让我们感受到艺术家艺术创作的独特性，以及他对社会生活环境和当前政治形势的反思。实际上，艺术家在创作艺术作品时，没有任何责任或义务为观众带来任何感官或心理上的冲击和刺激。自杜尚以来，艺术更多地关注观念的表达，而不是盲目追求美或美的观念。现实主义，古代中国文人画非常重视艺术意境的传达，尤其是元代艺术家的艺术作品。艺术作品总是激发人们的思考。我们以“无聊”为创作的起点，一个想象的空间，那么我们就可以在艺术创作中享受这种无聊状态，用自我意识改变这种感知，也是可能的吗？正如桑塔格所说：“如果我们感到无聊，我们应该问自己是否正确处理了我们的关切。”因此，我们可以得出结论，以“无聊”作为叙事形式的艺术往往并不无聊。

## 2. 新媒体材料

### 2.1 无生命介质

#### 2.1.1 介质生成的体验

体验的前沿 - VR 技术改变视觉体验

随着当今技术的发展，虚拟现实（VR）和 AR 技术已经涉及到我们生活的所有领域，尤其是在虚拟现实游戏和虚拟现实影视领域，像是一些 VR 游戏比如 Stean 的游戏 Tilt Brush1，你可以在 3D 环境中进行绘画，另一个 AR 应用是 HoloStudio Holoens，通过这个软件，你可以利用手势在三维空间中构建 3D 模型。同时，使用 3D 打印技术，将创建的全息图打印出来。VR 技术与电影的结合，这项技术的进步，颠覆了传统的电影技术，以前的构图、场景移动、剪辑等方法都不成立，需要重新考虑 360 度镜头下的新镜头语言，相机会随着故事的发展而移动。观众仿佛置身于电影之中。通过在空间中戴上的 VR 眼镜，建立起空间的景深感，从而创造出全新的空间纬度和时间维度。这种新视角带来的场域本质是不存在的，但观众极大地沉浸在这个虚拟世界中。为什么原本没有表达自己存在立场的观众会对故事产生兴趣？在王拓、刘彦涛和周键的工作空间演变中，艺术家通过将互动技术、动态捕捉技术、实时视频生成技术等多媒体综合方法加入到作品中，完成了视频作品。观众可以与虚拟世界中的元素互动。虚拟互动，通过这个作品来思考身体与世界之间、真实与虚拟的存在之间的关系。在虚拟视觉环境中，视觉和听觉增强了虚拟世界中的真实感，比如约翰·克纳特鲁普和

马兹·丹布斯波的艺术电影《狗屋》和《The Doghouse》，在戴上有着“家庭宴会”这样的故事情节的眼镜和耳机的瞬间，每个人在与真实物体结合时，都从自己的角度开始了全心全意、全意识的体验。在 VR 电影中，观众诞生于虚拟环境中，但无法改变或对虚拟世界产生任何影响。当你佩戴 VR 技术，作为一种工具、手段或艺术呈现方式时，它自然成为艺术家的艺术创作和表达手段新宠并不令人惊讶。VR 技术不断在改进和发展。虽然 VR 技术仍有一定的缺陷，长时间佩戴 VR 眼镜会让体验者感到头晕，而且因为它是虚拟现实，不会带来真实的触感。但我们可以预测的是，未来的观看方式将会是一个具有时间维度的三维虚拟现实。真实场景与虚拟场景之间的无缝连接将会是未来 VR 叙事方法。

### 2.1.2 控制与技术的破坏

技术与艺术的发展使艺术家更加注重在使用技术后艺术中观众的体验。在 2016 年北京媒体艺术双年展“技术与伦理”中，有一个参与性体验作品“人性”。“与机器共舞”，这是由 Bill Vorn 和 Louis-Philippe Demers (Philip Demers 和 Bill Vorn) 合作制作的机器人表演，作品安装在观众的身上。当与人类结合时，只有当机器人部件与人类一起表演时才是一个完整的作品。观众可以自由移动或将自己的身体交给机器控制，从而创造出人与机器之间的互动。在表演中，机器不会在人体中诱发疼痛，但会唤起身体受惩罚的体验。通过这种方式，艺术家质疑了人类存在的现状，艺术家 Gary-Hill 也参与其中。“人机之间的舞蹈”这项工作看起来像是一场游戏，但它带给我们更加敏感和深刻的思考。我们发明了机器。在技术的驱动下，生活方式发生了变化。技术背后，是人类被机器统治还是人统治机器。凯文·凯利在《不可避免》中也提出，当我们的的大脑被广泛开放时，这个社会是否被机器统治，人类中的精英只是被机器统治的人才。我们没有被机器统治吗？100 年前，人类没有手机和电脑。通信只能通过信件和面对面的交流来完成。要了解最新消息，读报纸。现在我们已经成为手机的“奴隶”，早上醒来的第一件事就是刷微信朋友圈或者看最早的新闻，看看昨晚是否有什么大事发生。你不必早上去股票大厅，你可以直接从手机上检查，中午吃午饭也可以从手机应用程序上订餐。将会有专人送餐。（或许将来，机器人将负责送餐和快递），即使是在地铁上的空闲时间，大多数人也在低头看手机读书、看电视节目或者玩游戏。

技术的控制与破坏是一个复杂的话题，它涉及到了人类与机器之间的界限、相互作用和未来的关系。技术的发展无疑增强了我们生活的方方面面，从通信到娱乐，从工作到个人生活，技术无处不在。然而，这种普遍性也引发了担忧：我们是否过于依赖技术？我们是否正

在失去对技术的控制，以至于我们自己的行为 and 思想也受到了机器的影响？

在这种背景下，艺术家们利用他们的作品来探索这些问题，挑战我们对技术的接受程度，以及我们愿意允许技术进入生活的深度。例如，“与机器共舞”这样的作品，通过对身体和机器之间的物理连接进行实验，提出了关于自主性、控制和共生的问题。这些作品并不是要提出答案，而是要激发更深入的思考和对话。

随着技术的不断进步，我们作为社会成员的角色也在不断地被重新定义。凯文·凯利（Kevin Kelly）在《不可避免》一书中提出的观点，强调了技术发展的不可避免性，以及它如何塑造我们的思维和行为模式。他的观点提出了一个关键的问题：我们是否能够维持对技术的控制，或者我们正逐渐成为它的从属？

这种担忧反映在我们日常生活中对技术的依赖上。从早晨醒来检查手机，到使用各种应用程序简化日常任务，技术已经成为我们生活中不可或缺的一部分。然而，这种便捷同时也带来了潜在的风险：隐私泄露、注意力分散、人际关系的质量下降，以及对现实生活的逐渐疏远。

艺术家通过他们的作品来探讨这些现象，为公众提供了一个反思和质疑现代技术生活方式的空间。他们的作品促使我们思考，同时也提醒我们，尽管技术可能会带来许多便利和好处，但我们作为人类，需要保持对它的警觉和批判性思维，以确保我们不会失去对自己生活的控制。

总结来说，技术的控制与破坏是一个不断发展的议题，需要我们不断地评估和重新考虑我们与机器的关系。艺术在这一过程中扮演着重要的角色，不仅作为一种表达和反思的媒介，也作为激发公众意识和促进社会对话的工具。

我们的生活总是与手机密切相关，许多年轻人说很难想象没有手机的一天。每件事都有两面性。当我们感谢科技带给我们便利和快捷的生活时，我们是否看到了科技背后对人类存在方式的巨大改变？当然，科技不仅带来了便利，也带来了在战后，核武器可以在瞬间消灭成千上万的人，而且还有长时间难以消除的核辐射，就像俄罗斯的切尔诺贝利和日本的福岛核电站爆炸一样，把整个区域变成了无人区。换句话说，就是受限区域，切尔诺贝利事件近20年后，一个真实夫妇在优酷上的自媒体节目“旅行”通过各种杂乱的程序再次进入。我们看到的是一些人曾居住的工业区。一座荒城的场景，天空是蓝色的，树木被杂草丛生，路边有诱人的红苹果，居民楼里是人们匆忙逃离时留下的日常用品，一切看起来都非常陌生，靠近核电站泄漏中心区域的辐射值仍然很高，你需要穿着防辐射服装才能接近，但约有15千人生活在辐射区。我们无法停止技术的进步，因为技术进步是必要的。“逆风行走”与艺

术家 Bontaro Dokuyama 合作，在艺术作品“福岛行走”中，也就是 Bontaro Dokuyama 的家人住在日本核电站泄露的福岛排除区外。福岛排除区边缘居民希望在未来生活在低辐射中的愿望反映在艺术家创作的不同风格的头盔上。就像切尔诺贝利一样，居住在福岛的人们紧急疏散，但辐射对人类的破坏并没有停止。受辐射区域内的所有食物和水源都无法食用，所有商店都关闭了。排除区外的人们只能依靠从其他地方发送的未受污染的食物生存。耳机的材料来自快递员使用的包装盒。包装盒上也会有感谢的语言：“总有一天我们会送你福岛的美味桃子”，这是福岛人民的心态。他们乐观地认为，即使在这种不可逆转的辐射影响下，核污染的放射性物质仍可以通过呼吸、皮肤伤口和消化道被吸收到体内，并诱发类似白血病的慢性辐射病。

核污染的影响不是几年内就能轻易消除的。化学上的解释依赖于污染中元素的半衰期。对于像铯、锶和钚这样的放射性元素造成的环境污染，它们的半衰期分别是 30,000 年、29,000 年和 24,000 年。在成千上万的年份里，这些化学物质会衰减到原来浓度的一半，因此它们会造成土地的长期污染，这是一种难以消除的长期污染。我们无法停止技术的进步。核电站不可想象的巨大发电能力也可以惠及人类的生产和生活。然而，作为艺术家，我们可以通过体验让观众感受到技术背后的影响，这可以使我们以不同的方式思考。面对技术与生存的对抗[3]。

## 2.2 生命媒介

### 2.2.1 生物艺术—生命

自从 1996 年现代工程学创造了第一只人工克隆羊“多莉”以来，“克隆”生物技术为进一步发展奠定了基础，引起了广泛关注并激起了社会担忧。克隆技术的最大优势在于它几乎可以复制被克隆本体上的所有特征，通过胚胎细胞可以进行克隆的技术。人们对克隆可能带来的问题、以及克隆技术可能误用造成的灾难感到恐惧。

同时，使用生物材料作为媒介的艺术也在发展并进入公众视野。生物艺术是用来指代艺术家使用活的有机物质创作的艺术，这些有机物也被科学家使用：细菌、细胞系、分子、植物、体液和组织，甚至活的动物。对生物艺术的分类没有严格的定义，正如任何艺术媒介一样，关于某种材料和艺术实践是否可以被归类为生物艺术一直存在争议。例如，当艺术家使用生物材料时，这种材料必须是生命的天才即是生物艺术。

2009年7月，现代与当代艺术史系讲师弗朗西斯·斯特雷西在《自然评论》杂志的“科学与社会”专栏发表了《生物艺术：美学背后的推理》一文。文中描述了西方生物艺术的起源，在1936年，爱德华·斯泰钦在纽约现代艺术博物馆展出了翠蓝色翠雀花的作品，这些花朵奇异而美丽。因为这些花都是化学实验的产物，艺术家根据达尔文的生物学观点“自然选择”，挑选了合适且美观的花朵在展览馆展出。在此之前的1933年，英国科学家亚历山大·弗莱明爵士（Alexander Fleming）的“细菌绘画”作品实际上已经模糊了艺术家和科学家之间的界限。爱德华·斯泰钦和亚历山大·弗莱明爵士发起的生物艺术研究的首次热潮之后，二战很快爆发。二战结束后，欧美国家相继进入经济复苏时期，科技迅速发展。在接下来的几十年里，随着计算机技术的飞速发展，人工智能在英国和在2009年7月，现代与当代艺术史部门的讲师弗朗西斯·斯特雷西在《自然评论》杂志的“科学与社会”栏目发表了《生物艺术：美学背后的推理》一文。文中描述了西方生物艺术的起源。1936年，爱德华·斯泰钦在纽约的MoMA画廊展出了翠蓝色翠雀花作品，这些花朵奇异且美丽。因为这些花都是化学实验的产物，艺术家按照达尔文的“自然选择”生物学观点，选择了适宜且美丽的花朵在展览厅展出。在那之前的1933年，英国科学家亚历山大·弗莱明爵士（Alexander Fleming）的“细菌绘画”作品实际上已经模糊了艺术家与科学家之间的界限。爱德华·斯泰钦和亚历山大·弗莱明爵士发起的生物艺术研究的首次高潮之后，二战迅速爆发。二战结束后，欧美各国陆续进入经济复苏时期，技术迅猛发展。在随后的几十年里，随着计算机技术的飞速进步，人工智能在英国和美国得到了广泛研究，艺术家们也尝试使用这一新技术进行创作。然而，生物艺术在世纪初出现，但在这漫长的时期里鲜有涉及。直到20世纪80年代末和90年代初，1990年，由美国国会资助的“人类基因组计划”正式启动，生物科学被引入了艺术家的艺术视野。

### 2.2.2 艺术家对生物媒介的理念

信息技术革命对未来人类思维方式和存在方式产生了巨大的影响和改变，我们每个人都有着深刻的体验。然而，生物技术对艺术发展有何影响和作用呢？生物技术一直被置于实验室中，它们的论文发表并赢得了各种世界级的奖项，但是先进的生物技术很少出现在公众面前。当生物技术成为艺术表达的媒介时，艺术背后的思考也带给我们不同的视觉盛宴。自从转基因技术被人类使用以来，它就一直饱受质疑。在科学机构的帮助下，艺术家爱德华多·卡茨（Eduardo Kac）成功培育出了2000年出生的绿色“荧光兔阿尔巴”（fluorescent rabbit Alba）。

艺术家爱德华多·卡茨使用 GFP（绿色荧光蛋白）技术将水母的 DNA 注入到母兔的胚胎中，阿尔巴出生后在紫外线下会显示出绿色荧光。自那以后，基因工程引起了一致关注，但是与卡茨一起开发阿尔巴的科学机构拒绝养这只兔子，因此由生物基因创造的无法归类的生物体的伦理问题遭到了公众的质疑，阿尔巴的诞生。

它并不是为了满足人类对新物种的好奇心，第二是正视转基因动物作为一个生命体的存在。爱德华多·卡茨 (Eduardo Kac) 的另一个作品将他的 DNA 转移到了一个红色的花茎中。植物就是植物，动物就是动物，该作品题为“植物动物” (Pinmal)，卡茨对这项工作说：

“地球上的生命经历了 40 亿年的演化，植物特性和动物特性在很早的时候就已经分离。这项工作不仅仅是挑战了 40 亿年的历史。植物的进化和人类的进化，更重要的是，它增加了进化历史的史诗，在这个史诗中未来物种之间的界限将会被跨越。”他的作品没有生物链金字塔或层次结构的排序，而是构建了生命的水平光谱。在这样的水平范围内，所有生物都是构建生命共同体的概念。巧合的是，法国艺术家马里昂·拉瓦尔-詹泰 (Marion Laval-Jeantet) 的行为艺术作品“愿马与我共生”，将马的血浆注入他的身体，使他变得更像一个“人马杂交体”，常识是普通人类不可能接受除了自己血型之外的血液，更不用说动物的血液，会发生排斥反应，严重情况会导致人的生命。为了这项工作，马里昂在过去的几个月里接受了几次马的免疫蛋白和其他物质的注射，以增强免疫系统对马血细胞的耐受性。最终，她成功完成了工作，没有任何不良反应。在表演期间，她引导着为她提供血浆的马，穿着特制的类似马蹄的高跷。表演结束后，艺术家本人表示，注射马血后，她感觉自己和以前不一样了。马里昂的尝试是大胆和新颖的。这种人马共生的可能性呈现给观众，触发了人与动物共生的思考。艺术家改变了生命的密码，这是持久和不可动摇的。

### 2.2.3 生物媒介艺术的伦理认知

奥隆·卡茨 (Oron Catts) 的《无害皮夹克》(2008) 生物艺术在使用材料方面与过去的艺术有很大不同。生物艺术与生命科学实验室密切相关，艺术家所使用的技术也非常先进，为了创作和完成一个作品，生物艺术家需要掌握一些生物技术知识，如 DNA 技术。当然，科学家的帮助和共同制作也是必要的。许多生物媒介艺术家的作品都是在科学家的帮助下实现的。生物媒介艺术与技术之间的关系非常紧密。它触及到社会科学：如医疗（例如试管婴儿）、食品安全和生物伦理学。

2008 年，生物艺术家奥隆·卡茨使用生物材质创作的备受赞誉的作品《无害皮夹克》

在纽约现代艺术博物馆展出。选择“皮夹克”是因为组织培养与艺术项目（Tissue Culture & Art Project, 1996 年与伊奥纳特·祖尔（Ionat Zurr）合著）对于奥伦·卡茨的项目。这个项目首先是对人类用来覆盖身体的工具——“服装”的反思。

这项作品是通过将人类和小鼠的干细胞放置在一个实验用玻璃生物反应器中，并喂养小牛胚胎血清来培养的。它是一种依赖营养生存的有机生命。在展览期间，因为它看起来像“皮肤”，这个类似“夹克”的生物细胞疯狂地增殖，袖子开始脱落，生物反应器也变得堵塞。最终，博物馆不得不停止营养供应，这使得生命现象停止了。这个由细胞生成的“皮夹克”是为了保暖，它反映了人类与生态系统之间的关系。然而，艺术家也明确表示，他们的计划不是要推出一个新的创新替代产品，而是要质疑人类和其他生命的开发利用。这一结果引发了广泛的公众争议：生物的使用在哪里是伦理的？生物基因技术的发展使我们违背了达尔文物种进化论，我们相信人类是某种猿的一个分支，之后我们慢慢直立行走。那么，以艺术名义带来的人工生命作品将有何不同呢？科学研究不是一蹴而就的，科学研究的结果未必在未来有任何用处，但是由于科学技术的某些行为或特点，人们已经变得对技术本身着迷。有一些固有的差异和鸿沟是难以弥合的，但是由于人类技术的发展，这些界限正变得越来越模糊。技术正在改变我们看待周围世界的方式，以及我们对待自然和生命的方式。

#### 2.2.4 中国生物艺术的发展现状

近年来，随着科学技术的发展，“生物”相关的词汇在我们的视野中出现得越来越多，特别是在生物技术、生物工程、生物伦理学、生物制药以及外国艺术的生物领域。生物艺术已逐渐被认可与发展。国外有许多生物艺术实验室，为使用生物作为媒介的艺术家们提供一个专业高效的创作环境。在中国，专业的生物艺术家寥寥无几，但许多多媒体和跨媒体艺术家已经开始利用生物材料进行艺术创作。例如，80 后多媒体艺术家陆扬，其作品《No.1 僵尸青蛙水下芭蕾（NO.1 Zombie Frog Underwater Ballet）》，这件作品是关于通过电膝跳反应控制死后青蛙心智的作品，只有半身的青蛙在水下随着无线电波的节奏跳舞。然而，死后的青蛙仍能移动，反映了意识的控制和对身体的补偿。准确来说，这只是一件基于生物背景的作品，就像她的其他作品，新媒体作品《子宫战士（Womb Warrior）》、《癌症宝宝（Cancer Baby）》和关于宗教的《茶狂金刚核（Cha Furious King Kong Core）》。所有这些都贯穿着“半人半机”的生物概念。

当然，还有其他艺术家也致力于生物艺术领域，比如微生物艺术家陈幼桐和生物艺术家

李山。

陈幼桐使用微生物作为艺术创作的媒介。他的艺术作品《植物》是一件关于“死亡与繁殖”的作品，是通过使用微生物、土壤和死去的植物制成的。这件作品通过将分解纤维素的微生物移植到死去的植物种子中，恢复了种子的绿色生机。虽然死亡与繁殖一直是艺术表达的永恒主题，但陈幼桐用全新的媒介进行表达，显得更具当代感。

2007年，李山与艺术评论家张平杰完成了具有植物特点的艺术作品《南瓜》。李山的艺术项目《南瓜计划》是一个可以在当前生物艺术阶段实施的主题，因为《南瓜计划》中的南瓜不是自然生长的，它是艺术家被迷住的人工生态生长的南瓜，一个经过人工处理的生物符号。

在未来时态中，使用生物学作为媒介的艺术家依赖于实验室、新的发现以及生物实验室的视角；这也是生物艺术与其他艺术门类的区别所在。他们最大的贡献是，他们的作品在很大程度上总是指向我们不知道的某些事物或领域。未知的空间非常广阔，这决定了生物媒介艺术的独特先锋性，它将超越和突破许多人的认知，挑战伦理，因此它也具有相当大的不确定性和复杂性，生物艺术家面临着许多问题。在未来，艺术将发展出更多的可能性出现在观众的视野中。使用生物学作为媒介的艺术将触及公众身心需求的双重挑战，但艺术家的艺术创作仍应遵循以人为本的人文精神，避免生态破坏。随着技术的快速发展，艺术家和公众对艺术的认识也将越来越广泛。生物艺术使用基因、细胞或活的动物作为创作媒介。

#### 2.2.5 在自己的艺术创作中使用新媒介

自当代艺术发展以来，随着技术的发展，艺术创作的概念和媒介经历了巨大的变化。艺术家们关注社会与政治、生活与自然，以及哲学与技术。艺术家们让公众对生活环境有了更多的了解，我们对自己的感知和解释也更加丰富。我希望通过我的作品理解这些概念和媒介。

《它看起来很美》创作的起源：我们生活在当今社会，而当前中国食品安全是人们生活中讨论最多的问题。各种新的食品术语如“毒姜”和“苏丹红鸭蛋”出现在我们的视野中。快速的经济扩张和发展带来了快节奏的生活，但也引发了食品安全问题。作为自然的晶体，无论是植物还是动物细胞，在显微镜下都极其美丽，通过细胞产生的生物艺术可以看到美丽的细胞作品，比如可以为设计治疗方案提供逆转人类视网膜退化和失明线索的斑马鱼视网膜或用于制作纺织品和绳索的血管植物等。然而，为了最大化利润，一些人的价值观被扭曲，原本的动植物被加工成对人类有害的食物。不仅是食品，人类对美的追求从未停止。人造美

人——面部和身体塑形也成为我们讨论的问题。然而，一些人对美的追求变得病态，但总有一些特立独行的艺术家做相反的事情。另一方面，法国行为艺术家奥尔兰（ORLAN），她反对传统的利用整容手术使自己看起来怪异的概念，通过四次手术操作，在她的脸上、额头和太阳穴制造了一对肾上腺腺瘤。突然的结节，同时通过视频向世界传达她的“痛苦与欢乐”的手术过程，这项工作奥尔兰想要创造一个与社会和传统冲突的社会形象，以促进女性权利。通过作品《看起来很美》考虑我们当前的生活环境和审美价值的偏差[5]。

具体实施和过程：“看起来很美”分为两部分完成，第一部分由 $7*7=49$ 个仿生蛋（通常称为假蛋）完成，无壳的假蛋黄是由藻酸钠加水制成，并加入柠檬黄溶液，倒入模具中然后加入氯化钙使其凝固。蛋白也是如此，但不添加色素。成分还包括明矾和明胶。柠檬黄会影响人的神经系统，特别是儿童吃多了会有注意力缺陷多动障碍（ADHD），明矾中的铅也会影响智力发展，导致儿童智力迟钝和老年人痴呆。 $7*7=49$ 天，这是佛教传给中国的传统丧葬仪式，是一个拯救亡灵的大法会。仿生蛋（无壳）按仪式放置在白色瓷盘中。不断地干燥和收缩。未受精的蛋本意是作为供人们食用以提供营养的细胞。尽管仿生蛋在外观上相似，其本质上却有质的不同。第二部分由我家庭成员的头部照片完成。这是由最流行的、病态扩大的眼睛、超高的鼻梁和超尖的下巴组成的“蛇”。“精致的脸”给了我一种怪异和丑陋的精神感受。两部分相互衬托和补充，以此反映我们生活的当下时代，我们所处的社会，以及人与人之间的思考和人们的审美价值观。

技术使人类在生产和生活中得到满足。我们今天的艺术面临着一个新的环境。艺术家们思考如何在新环境中转换不同的语言来创作艺术作品。在我们关注审美和思想表达之后，科学和技术的应用也成为了艺术家改变媒介的方式和方法。在依赖了某些媒介的特殊性之后，他们思考生活环境、社会问题和政治生态，并且也有一小部分艺术作品。可能只是纯粹地想表达无聊这一主题。随着今天艺术的发展，艺术家们创作艺术作品不再单一追求美。审美、概念和材料已经成为艺术创作的三大核心问题。时间和空间、情感和记忆是艺术家艺术创作的主线。观众的参与、互动和体验将成为艺术作品的风向标。在艺术作品中，如重复、无聊等视角是艺术作品个性特写的调味品。由科学和技术产生的新媒体艺术将始终激发公众的伦理思考，并且艺术的未来发展有许多可能性。新媒体艺术活跃在技术和艺术交叉的各个领域。新的特性、新的恐惧、新的机遇和新的认知以迅速的速度占据我们的认知和生活。因此，利用技术服务于艺术是新媒体艺术家的责任。未来的艺术建立在跨媒介和跨学科的艺术表达方式之上。

## 第2节 互动新媒体装置艺术框架设计与 案例实现

### 1. 互动新媒体艺术及互动方法的发展

#### 1.1 互动新媒体艺术简介

长期以来，绘画和雕塑是所有艺术的代表。现在，电影、视频、声音和计算机都成为艺术家手中的工具。如今，人们在参观现代艺术展览时能够看到多种艺术形式，而互动新媒体艺术是其中重要的一部分。

互动新媒体艺术起源于19世纪早期法国艺术家马塞尔·杜尚的现成品艺术。二战后，它逐渐发展成为一种独立的艺术形式，在西方现代艺术中占据了重要位置。通过对现成产品的组合、修改和挪用，互动新媒体艺术给参与者提供线索和灵感，并激发人们思考其背后的深层内涵。互动新媒体艺术不是画廊中的展品，而是一个让观众和听众扩展想象力并反思作

品背后意义的空间。

杜尚的现成品艺术被视为互动新媒体艺术的起源。像“自行车轮”、“喷泉”和“雪铲”这样的作品颠覆了人们对艺术的传统观念，引发了对艺术本质的反思，并给人们带来了全新独特的体验。杜尚认为他的作品是有趣且富有玩味的。这是互动新媒体艺术与传统艺术最大的不同，它允许参与者从作品中获得乐趣和感受，而不仅仅是关注美学。

互动新媒体艺术有一些显著的特点：它需要一个可以与观众互动的独立空间，并且会根据不同的空间而变化；与观众的交流是互动新媒体艺术的重要部分；它对使用其他艺术手段是开放的，比如绘画、雕塑、音乐、电影、摄影、计算机、网络等，并不受艺术类别的限制；常常让观众体验到在现实生活中很少感受到或无法感受到的事物，并激发参与者的思考；随着时间和地点的变化，互动新媒体艺术作品也可能发生变化。

经过数十年的发展，互动新媒体艺术包含了越来越多的艺术作品，形式各异、材料不同、内容和效果各异，例如激光光艺术、视频艺术、电影艺术、计算机艺术和互联网艺术，以至于将这些作品归入同一类别似乎变得困难。但它们确实有一个共同的特点，那就是强调观众与艺术作品之间对时间和空间的新理解，而“互动新媒体艺术”是最能描述这些作品关键特征的名称。这也是互动新媒体艺术近年来越来越受欢迎的原因。

## 1.2 互动新媒体艺术互动方式的发展

互动新媒体艺术的互动模式随着科学技术的发展不断演进。在其起源之初，互动新媒体艺术更多的是利用增强的视觉效果来创造具有心理暗示的环境，让参与者融入其中并体验作品的整体。杜尚的《喷泉》和《L.H.O.O.Q.》是这方面的代表作，通过强烈的对比激发参与者的思考。1958年克莱因的《虚空》将观众置于一个空旷的纯白空间中，而阿曼次年的《满足》则用垃圾填满了同一个房间，使得观众只能从房间的窗口观看。这两部作品的回响和对比使人们反思对自然环境的破坏。这些早期作品提供的大部分参与方式允许观众看、听和感受作品，而不会让参与者改变它。少数作品提供通过机械方式进行简单互动的方式，比如杜尚的《自行车轮》：观众可以通过转动轮子来改变作品的状态。

光线和反射的神奇效果也是艺术家们热衷使用的一种方法。福利亚蒂曾经利用观众无法同时凝视两组光源材料的特性，创作艺术作品，将光影投射在60厘米间隔的细棒上，并悬挂在展览空间中的细棒上，导致眼睛的来回移动在棒之间的空间中产生真实图像的错觉。萨马拉在1966年的“镜子之屋”创造了一个无限的虚幻三维空间。参与者的动作通过镜子反馈，形成特殊的互动效果。让参与者成为作品的一部分。

20 世纪 60 年代是互动新媒体艺术从机械时代进化到电子时代的意义深远的时期。探索运动、光动和早期虚拟艺术运动的艺术家们广泛使用技术设备来创造各种运动和光效果。电机、电子、热、液压和磁力运动出现在这些艺术家的作品中，有时与光效果结合，制作弧光灯、舞台聚光灯、各种幻灯机，填充着钠或汞，无论是白色还是彩色的灯泡，直的或弯曲的霓虹灯或荧光灯。在 1968 年纽约现代艺术博物馆的展览上，让·迪普伊的“心跳尘埃”展示了一些独特的技术应用，这使他赢得了由 EAT（艺术与技术实验）组织的比赛。

在这一时期，激光和全息技术也被应用于互动新媒体艺术，作为展示和与观众互动的手段。美国艺术家洛威尔·克罗斯与大卫·图多尔以及科学家兼雕塑家卡森·杰弗里斯合作，创作了混合媒体作品“Video-Laser I, II”，将激光视觉展示与音乐和戏剧表演结合起来，被视为互动新媒体艺术的一个明显贡献。同时，电视、视频、投影等也开始出现在互动新媒体艺术作品中，而交互式有线电视的使用甚至预示了 90 年代互联网的出现。在 1978 年马勒的视频作品《前意识》中，观众坐在装置前，他们的头部会替换视频中的原始图像并成为图像的一部分，给参与者留下了深刻的印象。

艺术家对计算机技术的应用起源于 1952 年。早期的艺术家主要使用计算机绘制二维或三维图形图像，即巧妙地利用计算机屏幕编程。其中一个经典的艺术家是川口喜弘 (Yoichiro Kawaguchi)。在他的作品《漂浮》(Float)中，浮动和移动的海洋生物是直接通过计算机图形绘制的。在这项作品中，一组金属球体被用来形成曲面，其质地表达了生活在亚热带岛屿上的微生物的皮肤颜色。随着艺术家对计算机的理解加深，电子互动事件被引入到艺术作品中。莉莉安·莱恩 (Liliane Lijn) 的著名作品《梦想结构》(Dream Structures)是由微型计算机制成的雕塑，能够将光反应为声音。由铁质物件、光学玻璃和合成纤维组成的作品随着声音的变化而美丽而神秘地变化。

20 世纪 80 年代以后，随着计算机技术的不断发展，人机交互技术变得日益丰富。从鼠标的发明到 WIMP（窗口、图标、菜单、指针）界面，从单机到互联网，从 CRT（阴极射线管）显示器到虚拟现实技术，这些人机交互技术已经成为计算机技术的重要组成部分，并因此使得人机交互成为了一个独立的学科。艺术家们对新技术的热情一直保持着，从皮奥特·科瓦尔斯基 1983 年在巴黎 Electronica 展览上的“互动场”到莫奔 2005 年在上海创意中心的“冻结感觉 - 情感市场”。人机交互技术在互动新媒体艺术作品中的应用展现出更先进、更广泛和更深入的特点。

## 1.3 人机交互技术

### 1.3.1 人机交互技术概述

自 1946 年第一台数字计算机出现以来，计算机技术取得了惊人的发展，并已被广泛应用于社会生活的各个领域。人机交互技术随着趋势迅速发展。人机交互是研究人们如何与计算机通信以及它们如何相互影响的技术。计算机交互界面的优劣对计算机的使用有着重要的影响。交互性差、操作繁琐的计算机可能导致用户错误增加、效率降低和产生疲劳。这使得如何操控和管理计算机这一强大工具成为一个重要的研究课题。因此，它在 20 世纪 70 年代开始被美国、日本等发达国家重视，并在 1980 年代取得了巨大发展。

在计算机发展的早期，人们使用各种开关、指示灯和打孔纸进行人机交互。在输入时，使用设备将记录在打孔纸上的程序和数据传输到计算机，并通过指示灯输出运行过程和结果，调试只能用开关来完成。这些笨重的计算机大多用于科学研究和军事部门，使用不便，且完全没有艺术性。这一时期的交互被称为非交互式批处理模式。

随着计算机处理能力的提高和软件技术的发展，作业控制语言和交互式命令语言相继出现，以及键盘、单一显示器、光笔等工具。用户可以使用命令语言、功能键和文本菜单与计算机进行交互。在这一时期，人机交互进入了文本字符交互时代。

1963 年，由麻省理工学院开发的 Sketchpad 绘图系统和加州 RI 国际开发的鼠标为图形用户界面的实现奠定了坚实的基础，并创造了发展空间。1984 年，苹果公司在 MacIntosh 计算机上推出了首个多窗口系统，提供了对话框、滚动框、下拉菜单等更多功能。后来，微软公司开发的 Windows 系统和麻省理工学院的 X-Window 系统允许用户通过观察、指向和点击来操作计算机。这些作品将人机交互推入了基于二维 WIMP（窗口、图标、菜单、指针）的图形用户界面时代。

自 20 世纪 80 年代以来，将文本、图形图像、声音和视频集成在一起的多媒体技术迅速发展，增加了从计算机到人的输出带宽。1980 年 TCP/IP 协议的成功开发和 1990 年代 WEB 技术及相应浏览器的出现，促进了互联网的发展和应用，帮助人们从浩瀚的数字海洋中获取所需信息。由 Flash 等多媒体技术实现的 Web 交互已成为流行的交互方式。我们正处于多媒体和超媒体交互的时代。

近年来，视频识别、眼动追踪、动作识别、三维输入、语音识别、面部表情识别和手写

识别等技术已经出现并日趋成熟, 这些技术将继续增加从人类到计算机的输入带宽。虚拟现实技术和新感测技术等各种技术的发展, 使人们有可能在沉浸式虚拟环境中与计算机通信。这些技术的出现和发展将逐渐改变我们与计算机交流的方式。人机交互的发展方向是集成人类各种感官通道的多通道交互, 以完成交互。

### 1.3.2 人机交互中重要技术的当前状态

目前, 人机交互的新技术层出不穷, 并不断成熟。这些技术推动了人机交互的发展, 并决定了人机交互的未来方向。关于这一重要技术及相关文献的研究已经出现, 诸如: 董世海、王恒的《人机交互》、罗世坚、朱尚尚、孙守谦的《人机界面设计》、董世海的《人机交互的进展与挑战》、王宏兵、曲玉忠、徐冬梅、王瑶的《人机交互的几个关键技术》、方志锋、吴晓波、马维娟的《人机交互技术研究的新进展》以及宫洁敏、王宪清的《人机交互技术的进展与发展趋势》等。

计算机视觉: 科学研究表明, 人类 80% 的信息是通过视觉获得的。使计算机具有类似人类的视觉功能, 成为许多科学家的梦想, 而这个梦想正在逐步变为现实。当前的计算机视觉主要是使用摄像机来替代人眼, 用计算机来替代人脑。计算机分析摄像机捕获的图像序列, 并解释场景中的事件, 从而实现在二维或三维空间中的自然人机交互。

最早也是至今仍有影响力的计算机视觉框架理论是由 David Marr 在 20 世纪 70 年代末提出的。与人工智能的其他分支一样, 计算机视觉与认知心理学和生物计算密切相关。这些基础学科的研究进展可以为计算机视觉提供新的思路或处理方法。也许完全实现模拟人类视觉是不可能的, 但部分实现无疑有利于改善人类的生活条件。目前, 计算机视觉的商业系统已经出现, 并已成功应用于医疗图像处理、工业装配线生产、自动产品检测、智能交通、现场监控、数字图书馆、智能搜索引擎、数字娱乐、数字电影等领域。市场现已达到每年数十亿美元, 并且仍在迅速增长。同时, 计算机视觉也是眼动跟踪、手势识别和表情识别等其他交互技术的基础, 并在人机交互的发展中发挥着重要作用。

语音交互: 随着计算机便携性的趋势, 键盘和显示屏已成为计算机微型化的大障碍。与计算机的语音交互可以让人们摒弃键盘输入, 并通过显示器接收信息。语音交互主要由以下三个部分组成: 语音识别、语言理解和语音合成。

语音识别被认为是 2000 年至 2010 年信息技术领域十个重要技术发展技术之一, 并逐渐成为人机交互的关键技术。当前主流的语音识别技术基于统计模式识别的基本理论。近年来,

中文语音识别研究发展迅速，研究水平基本与国外同步，逐渐从实验室走向实用化。

自然语言理解是计算机科学和人工智能领域的一个重要方向，它研究能够实现人与计算机使用自然语言有效沟通的各种理论和方法。进入 1980 年代后，自然语言理解的应用研究得到广泛开展。现在一些领域的商业化、产业化以及典型实用系统已经实现，例如智能短消息服务、智能搜索引擎、智能聊天机器人等。

语音合成技术，也称为文本到语音技术（Text to Speech），简称 TTS，主要解决将文本信息转换成可听见的声音信息的问题，使机器“像人一样说话”，这是语音交互的一个重要部分。目前，语音合成技术正在进入一个相对成熟的阶段，许多关键技术问题已经基本得到解决，并已在许多领域得到广泛应用。例如：电话账单查询、语音信息咨询、信息查询等。

手写识别：键盘是向计算机输入信息最常见的方式。与传统的手写方式相比，尤其是中文输入，键盘必须通过输入法将字母转换成文字来完成输入。手写识别应运而生。手写识别意味着计算机通过收集写作时的信号，记录下笔迹的上下、每个像素在空间上的位置以及每一笔之间的时间关系。这些收集到的信号被处理以便根据特定规则提取特征，并最终实现识别，将其以常见格式存储在计算机的电子文档中。手写技术可以分为两类：在线识别和离线识别。

虚拟现实：虚拟现实（简称 VR），也称为精神技术，是在 20 世纪末发展起来的。虚拟现实技术的历史可以追溯到 1960 年代。1965 年，美国学者 Ivan Sutherland 在其博士论文中提出了虚拟现实的概念，这是最早的虚拟现实研究。由虚拟现实技术代表的新型人机交互技术旨在探索人与机器之间自然和谐的关系，使人机界面从视觉感知发展到包括视觉、听觉、触觉、力感、嗅觉和运动感觉。与通常意义上的多媒体相比，这项技术将人与计算机之间的信息交互渠道从二维（声音和图像）拓展到多维（声音、图像和其他人类功能感官），并且显示的图像从平面变换到立体。因此，可以说它为人机界面的发展开辟了一个新领域。目前应用的虚拟现实系统可以分为四类：桌面虚拟现实系统、沉浸式虚拟现实系统、增强虚拟现实系统（允许参与者看到真实环境中的物体）和分布式虚拟现实系统[8]。

由于虚拟现实技术的卓越互动特性，它已经对医学、通讯、教育、军事、航天和工业生产等领域产生了显著的影响。例如虚拟战场、医学模拟、虚拟旅游景点等。

虚拟现实技术的发展仍处于初级阶段。仍然有一些技术难题需要克服，例如形成和交互的实时性。这些问题也导致了当前虚拟现实的研究方向：通过计算机图形学建立实时三维视觉效果，研究虚拟人机交互接口，以及研究虚拟现实的应用。随着这些研究的深入，虚拟现实技术将得到进一步完善和实用化。

多通道人机交互：多通道界面的概念出现在 1970 年代初，当时有人提出了“会话式计算机”的概念，即可以以人们日常生活中相互交流的方式与人沟通的计算机，包括声音、手势、表情和视线等。这也是多通道人机交互研究的目标。自 1980 年代后期以来，多模态用户界面 (Multimodal User Interface) 已成为人机交互技术研究的一个新领域。

多通道人机交互技术的主要特点包括：使用多种感官和应用通道、在三维空间中直接操作、允许不精确的交互、双向交互以及隐式交互。多通道交互的优势在于：可以利用多种设备和交互方法的互补性；强调声音/音频通道，弥补了现有界面在这方面的不足；交互可以更加灵活。同时，它也有自己的特殊问题需要研究，例如多通道整合、人体工程学问题、描述方法、软件结构、界面评估等。

由于计算机发展的“小型化”和“便携性”，以及人们对自然和谐交互方式的需求，多通道交互肯定会成为未来人机交互发展的一个重要方向。

## 1.4 互动新媒体艺术中人机交互技术应用的研究

### 1.4.1 互动新媒体艺术中的人机交互特性

在动态新媒体艺术中的互动是最重要的关键环节之一。互动的质量直接决定了一个互动新媒体艺术是否成功。良好的互动可以吸引观众，引导观众与互动新媒体艺术作品进行互动，并触发观众的联想或反思，这对基于互动的互动新媒体艺术具有重大意义。在互动新媒体艺术作品中使用互动技术，应既考虑作品的艺术表达，也要考虑技术的可行性。对于艺术装置来说，如果无法表达作品的主题，不管技术多新颖，都是没有意义的。同样的道理，不管艺术想法多么伟大，如果技术上不可能实现，那也只能等待未来。一个成功的互动新媒体艺术作品是技术与艺术的完美融合。关于互动新媒体艺术中人机交互的特性和总结，请参考 Jennifer Preece、Yvonne Rogers 和 Helen Sharp 的《交互设计 - 超越人机交互》。



图 2-1: 数字舞蹈视频装置 DUO (图片来源: <https://www.alioop.com/blogue>)

在互动新媒体艺术中应用人机交互具有五个特点: 强烈的互动启发性、简单易用的交互方式、符合作品主题和展览环境、可靠的交互体验和不显眼的交互。

强烈的交互启发性: 当观众参观艺术展时, 在每件作品前的停留时间非常有限。在短暂的时间内, 吸引并引导观众成功与互动新媒体艺术作品进行交互非常重要。因为观众不能花费五分钟甚至十分钟的时间去研究和学习如何与作品互动。在人机交互中有一个可学性的要求, 即用户可以立即开始并且不需要太多努力就能胜任任务, 例如 Nelson 在 1980 年提出的“10 分钟规则”。对于展示作品中的人机交互, 对易学性的要求更为严格, 希望观众能够凭直觉或几秒钟的学习就能操作。在这种情况下, 作品的交互方式必须非常具有启发性。启示意味着给出提示, 比如门把手暗示“转动”、“推或拉”, 按钮暗示“按”。因此, 互动新媒体艺术作品中的交互大多使用具有强烈暗示性的物理对象来引导观众参与, 使得作品的交互具有更好的启发性。

互动简单易用: 观众与艺术作品的互动不同于学习如何开启电脑, 更不同于花费 20 分钟学习如何创建一个网页。互动是为了让观众体验作品并理解其背后的含义。因此, 装置作品不能要求参与者执行复杂的操作或等待操作结果, 否则参与者会因为不耐烦而放弃互动。互动新媒体艺术的交互必须简单、直接、快速且易于使用。具体来说, 包括两个原则: 功能的视觉简洁性和及时的反馈。功能的视觉简洁性意味着作品的整体交互功能对观众是可见的或可预测的, 并且可以通过一两个操作完成。让观众容易知道接下来做什么? 如何去做? 并成功完成交互操作。快速及时的反馈意味着艺术品能够快速返回与观众活动相关的信息, 如

视频、声音、触感等。符合作品主题和环境：交互是整个互动新媒体艺术作品的重要部分，其作用是引导观众参与并帮助参与者体验和感受作品。为了确保作品的艺术表达和吸引力，人机交互的应用必须与作品的主题和展览环境相符。作为艺术作品的一部分，交互的物理设备、交互方式、交互过程和反馈信息（声音、视频等）应该给观众美好的体验，充满美感，且符合作品的整体风格。互动新媒体艺术作品能否吸引观众参与，交互的吸引力非常重要。为了提高作品中人机交互的吸引力，应该做到：触发情感，给观众带来惊喜、乐趣或愉悦的感觉；抓住兴趣，激发观众的创造力，让他们感到迷人并从中获得成就感。最后但同样重要的是，作品中的人机交互能够启发人们。在体验了交互过程之后，参与者将受到启发并理解整个作品的意义。

可靠的交互：互动新媒体艺术作品通常在室内或户外公共场所展出，人机交互大多使用较为精密的设备。因此，交互的可靠性也是需要关注的问题，包括交互功能的可靠性和安全性。交互必须能够在复杂的公共环境中持续、稳定和安全地工作，否则会导致整个作品的失败，甚至对参与者造成伤害。

隐形交互：在人机交互的未来发展中，隐形交互是一个重要的发展方向。对于互动新媒体艺术来说，同样重视隐形交互。交互只是作品的一部分，观众的焦点应该在互动新媒体艺术作品本身上。观众的年龄、专业背景和水平各不相同，不可能每个参与者都精通电脑。一个好的交互应该让参与者专注于体验作品，而不是被如何交互所分心。

#### 1.4.2 人机交互在互动新媒体艺术中的应用

传感器和软件技术的发展以及计算机处理能力的提升，使得早期的开关触发式人机交互方法在互动设备中仍有良好的应用。1986年，Jon Goldman 的作品“在兰普雷的嘴巴里”（IN A LAMPREY'S MOUTH）在康涅狄格州哈特福德的 Real ART WAYS 展出时，就使用了红外传感器。当观众走过作品下方时，传感器会触发计算机播放一段古典音乐。红色的充气装置和愉快的音乐为参与者提供了一种新奇而美妙的体验。

台湾艺术家陈见泰和 Maclau 的作品“风铃”受到了泰国和尼泊尔寺庙的启发，使用竹子制作带有传感器的乐器。当观众的指尖拨动作品时，就会播放出美妙的风铃声。在日本媒体艺术节获奖作品展中，Kinamori 和 Juna mirror 的作品“幻影影”（Phantom Kage）也展示了当观众用传感装置触摸圆锥体时，计算机改变 CG 投影，创造出一个真实与虚拟共存的奇幻 CG 世界。

胡介鸣在 2004 年上海双年展上的多媒体装置作品“向上” (Up, Up) 中, 在堆叠的电视画面上, 一个红色的人影不断向上攀爬。当遇到外界声音时, 他会做出停下、或者跌落、或者加速的互动反应, 这些反应促使人们反思生活的无常。奥尔纳·波图加利 (Orna Portugal) 等人在 2004 年创作的作品“心跳” (Heartbeats) 则是基于心跳传感器的。当参与者的手触摸桌上的传感器时, 他们的心跳将控制一个虚拟生命的诞生、成长和消失。这些作品使用感应设备来感知外部环境, 将信号传输到计算机, 通过设计好的软件进行处理, 然后生成相应的动作, 以实现与观众互动的目标。这种互动设计方法通常能够取得相当好的艺术效果。

这些作品的创新之处在于它们如何将科技和艺术融合, 以及它们对于参与性体验的重视。胡介鸣的“向上”运用了声音交互, 将观众的声音转化为影响作品中人物动作的输入, 而“心跳”则将生物反馈作为交互的核心, 以此探索身体、技术和艺术之间的关系。这些作品都充分体现了互动新媒体艺术的核心特点, 即在观众参与的基础上创造富有深度的艺术体验, 并借此引发观众对于生活、技术和艺术本身的思考。

命令式交互和传统图形化交互 (例如 WIMP: 窗口、图标、菜单、指针) 是人机交互技术历史上的重要里程碑。然而, 由于它们对于输入设备如键盘和鼠标的依赖, 难以将其完全整合进互动新媒体艺术中。在互动新媒体艺术中使用键盘和鼠标作为交互手段时, 很难维持艺术作品的艺术吸引力, 使之不与办公用品产生联系。

网络交互, 网络让互动新媒体艺术家的创作手段前所未有地灵活。互联网可以成为展示平台: 1999 年, 爱德华多·卡茨 (Eduardo Kac) 的网络互动新媒体艺术作品《创世纪》使用了网络作为其作品的展示舞台, 参与者可以通过网络操作真实的装置。互联网上传的信息也被用来合成装置: 台湾互动新媒体艺术家陈见泰和 Maclau 的《互联网动作》将网络流量转化为另一种形式的展示, 使网络成为了一种无形的生命。当然, 互联网是一个信息的海洋, 艺术家可以搜索他们需要的信息来创造自己的互动新媒体艺术作品。

2005 年 5 月 1 日, 著名的法国多媒体艺术家莫本 (MoBen) 在上海展示了互动新媒体艺术作品《全球情感地图》。该作品通过莫本原创的软件, 在互联网上同时获取世界不同地区的信息, 根据情感关键词的点击率比例, 将其绘制成一个三维的地球仪。

人机交互技术正朝着多媒体和多渠道的方向发展。三维输入设备、手势识别技术、注视追踪技术、表情识别技术、语音识别技术、虚拟现实技术等相继出现。但并不是应用最新科技就能创造出最佳艺术作品。精密的输入设备需要用来实现准确的手势识别、运动追踪、注视追踪和面部表情识别。在互动新媒体艺术作品中使用如此昂贵和复杂的设备并不现实, 毕竟观众想要看到的是艺术作品, 而不是戴上 3D 眼镜和数据手套参与一个实验。语音识别技

术现在发展相当快,其在身份确认上的应用已经达到了99%的准确率,但艺术家想要的是能够给观众留下深刻印象的艺术作品,而不是能理解每一句话的机器人。那些不一定高度精确,但可以融入到艺术作品中的新技术,将会受到艺术家们的青睐。例如2002年,日本艺术家児玉幸子和竹野美奈子的作品《与声音共舞的磁浮波》,容器中的磁性流体根据参与者声音的变化不断地进行各种形态的变形。投影在屏幕上的磁性流体有着不同的形状和美丽的变化。在这项作品中,观众的语言并没有被准确识别,但它实现了惊人的艺术效果。

计算机视觉技术因其对观众干扰小、设备要求简单和非接触式交互的优势,受到互动新媒体艺术家的高度赞扬。目前,基于计算机视觉的互动新媒体艺术作品非常多,例如著名的“文字雨”作品。这个作品使用摄像机收集实时图像,计算机分析参与者的边界线,然后控制文字落下和停止的位置。许多互动新媒体艺术家利用这种互动方式实现了文字雨作品的不同版本。在2005年上海举行的“电子花园”国际新媒体艺术展中,韩国艺术家安东尼·洪的互动作品“电子鱼F12”通过分析参与者的影子位置,实时控制虚拟游泳鱼和观众的互动。另外,还有许多采用高科技的互动新媒体艺术作品,甚至可能促进新型人机交互方式的产生。例如,麻省理工学院2003年的作品“Audiopad”结合了电磁感应器追踪系统和投影仪,为我们构建了一个奇妙的虚拟表演平台。参与者可以通过移动不同的塑料块来操作虚拟界面,并控制音乐播放。可以说这是一个成功地将虚拟现实技术与艺术及音乐结合的作品。

克里斯塔·索默勒(Christa Sommerer)和洛朗·米尼奥(Laurent Mignonneau)在“电子花园”展览上的作品“互动植物”允许观众与真实植物互动,然后控制虚拟植物的生长。台湾互动新媒体艺术家林淑敏的视频作品“内功”允许参与者使用脑波与作品中的虚拟鱼和莲花互动。

## 2. 交互设计与实施

交互模块是艺术作品架构中的重要组成部分。本文中的人机交互是由物体触发的。通过玉触发器激活视频内容的播放;这一章将详细讨论本作品的图像识别技术,并提出与本作品相符合的人机交互技术的相关应用,从而实现剪纸相关视频的播放。

### 2.1 作品交互中实现兔石识别与视频反馈

在互动新媒体艺术作品“兔”的人机交互应用中,目的是使作品更好地与观众互动,研

究计算机技术与艺术的结合，并探究在互动新媒体艺术中应用人机交互的正确方法。在参考国内外互动新媒体艺术作品中的交互方式基础上，我们在作品“兔”中使用真实物体作为交互手段，将真实和虚拟图像进行叠加反馈。使得观众能以日常方式操作作品，并能获得在现实生活中看不见的新奇体验。

### 2.1.1 数字图像处理与识别原理

本作设计中使用的图像处理与识别技术主要参考以下文献，如：陈传波、金宪基《数字图像处理》、陈树海、傅鲁祥《实用数字图像处理》、何东健《数字处理图像》、余成波《数字图像处理与 MATLAB 实现》以及 Rafael C. Gonzalez、Richard E. Woods《数字图像处理》第二版。

其中，图像处理与识别技术主要集中在灰度图像与校正、去噪与平滑、二值化、数学形态学、边缘检测和轮廓提取等方面。

### 2.1.2 作品交互模块的算法过程

该作品基于计算机视觉交互技术，其开发环境为 Microsoft Visual C++ 6.0，主要使用的功能库为 OpenCV 5 beta 版本。OpenCV 是由英特尔公司开发的功能库，主要由 C 语言函数和少量 C++ 类组成，用于图像处理和计算机视觉。由于 OpenCV 提供了许多优化的底层图像处理和界面处理函数，开发者可以直接实现简单高效的界面操作和图像处理，无需调用复杂且效率不高的 Windows API 函数。另一方面，OpenCV 是开源的，开发者可以根据自己的需要重新编译，这样非常灵活。

## 2.2 《兔》中的人机交互设计

人机交互设计是《兔》作品设计的重要部分。为了成功设计互动艺术装置，以下将详细分析和设计具体的交互方法。

### 2.2.1 交互要求与环境分析

互动新媒体艺术作品的成功依赖于观众参与的互动效果。因此，在设计本作品的初期阶段，应特别注意人机交互部分的需求。尽可能广泛地调查作品未来的展示对象和环境，并同

时注意调查人机交互涉及的硬件和软件环境，以提高交互的可行性、稳定性和易用性。

由于《兔》将在开放式展览中展出，我们访问了上海多伦现代艺术博物馆、证大现代艺术博物馆、上海科技馆，对展厅中的互动作品进行了研究，并做出了以下判断。

展示时间：10 至 12 小时。

展览环境：室内，有工作人员。

目标观众：普通观众，具有一定的知识水平和文化素养。

交互要求：简单、可靠、新颖、可操作性强。

涉及交互的设备：电脑、传感器、投影仪等。

涉及交互的软件：自行开发。

### 2.2.2 观众特性分析

为了确保作品交互的适当性和清晰性，有必要了解观众的技能 and 经验，预测观众对不同交互设计的反应，并分析观众的特性。《兔》的观众是广大公众，其中很大一部分是儿童。他们的电脑技能相对较弱，因此交互操作不能复杂，并且需要明确的提示。由于作品基于传统中国文化，观众对“兔子”应该很熟悉，而剪纸兔子视频的出现将给观众带来新奇体验。

### 2.2.3 交互模型的建立

人机交互模型是描述人机交互系统中交互机制的结构化概念模型。目前已经提出了许多模型，如用户模型、交互模型、人机界面模型、评估模型等。这些模型从不同的角度描述了交互过程中人和机器的特性及其交互过程。在人机交互领域的模型研究中，较早提出并且有影响力的模型是诺曼的执行-评估循环模型。在这个模型中，诺曼将人机交互过程划分为四个阶段：意图阶段、选择阶段、行动实施阶段和评估阶段。

由于交互式新媒体艺术作品的交互是以观众为中心的，并在展览的特殊条件下进行，因此我们从用户认知的角度来建模。

### 2.2.4 交互任务设计

在作品中，用户拿起玉石并将其放置在柜子上后，触发计算机识别玉石的轮廓、边缘、面积和特征等参数，然后分别触发对应的图片和视频。通过这个过程，参与者可以通过选择

不同的玉石来触发不同的视频和图片，观看后可以理解这个作品背后的意义。

作品设计为 1.5（长度）× 0.8（宽度）× 1.1（高度）的复古风格箱型木质桌子，所有电脑、相机、投影仪等设备都隐藏在桌子内部，以满足隐式交互的要求。桌面的前部（扇柄部分）是一块磨砂玻璃窗口，允许参与者将手放在上面进行交互。玻璃窗口一方面可以暗示参与者将玉石放在其上，另一方面也有利于相机通过玻璃捕捉玉石的轮廓。随着参与者将玉石放在窗口上并对其产生影响，投影显示出图片和视频信息。观众通过影响交互平台周围的光线与作品交互，并最终完成剪纸兔视频的展示，获得成就感。整个作品的交互过程简单方便。观众可以通过直觉或简单学习进行交互，最终完成剪纸“兔”相应信息的展示。

## 2.3 作品中的人机交互应用

### 2.3.1 作品交互中物理对象的应用

自然人机交互是指使用人们的日常技能进行交互，而这些交互技能无需特别培训或仅需少量培训就可以获得。人们生活在物理世界中，在漫长的生活经验中积累了丰富的日常技能。如果人们能在与计算机互动时使用这些技能，将大大减少人们的认知负担和所需的培训量。

正是因为人们长时间生活在物理世界中，人们对物理世界中许多事物的操作方法和功能有一种直觉，一种不言自明的感觉。例如，当我们看到一个把手时，就知道这个把手是可以抓握的。如果门上的把手是球形的，我们知道它应该被旋转；如果把手是弓形的，两端固定在门上，那么我们知道把手是不能旋转的。人们看到大面积的光滑表面时，会想到在上面书写。可以看出，人们对物理世界的直觉非常丰富，而这种直觉可以为作品交互设计带来极大的帮助。适当利用人们在物理世界中的经验可以帮助用户很好地理解作品的交互。

为了使观众与作品之间的交互简单方便，我们在设计作品交互时利用了人们的日常生活经验，并使用人们非常熟悉的物体作为交互的触发手段。在作品的交互设计中，包括了玻璃窗和空白剪纸画框。相机将图像投射在玻璃窗上，当观众看到时，他们会想到通过在玻璃窗上放置物体来改变图像，并会用手尝试测试。

### 2.3.2 作品交互中计算机视觉的应用

基于摄像头的计算机视觉交互技术使人们能够突破现有的人机交互方法，如键盘、鼠标

和游戏手柄。使用数字摄像机捕捉用户的动作并区分用户的操作意图，使得人与机器之间的交互变得更加直接、自然和亲密，从而缩小了人与机器之间的距离。交互的必需性。计算机视觉技术对观众的干扰较小，设备要求简单，并且可以实现非接触式交互。它已经在许多交互式新媒体艺术作品中得到应用，并受到交互式新媒体艺术家的高度赞誉。正是因为这些优点，我们在作品的交互中采用了计算机视觉技术。

为了使观众与作品之间的交互尽可能简单，我们使用摄像头追踪磨砂玻璃上物体的轮廓变化，并通过投影仪提供实时反馈，以便观众能够理解交互的结果。另一方面，为了确保作品的艺术表现力和吸引力，并让观众关注作品本身，我们将摄像头和其他计算机设备隐藏在交互平台中，避免分散观众的注意力。

### 2.3.3 作品交互中增强现实的应用

增强现实系统使用了一些与虚拟现实技术相同的硬件技术，但两者之间最关键的区别是，虚拟现实用软件模拟的虚拟世界取代了真实世界；而增强现实则是在真实世界的背景中添加补充的虚拟信息。例如，增强现实维修系统使用先进的增强现实技术，在真实场景的图像上实时叠加维护所需的辅助信息。这系统可以帮助维修人员更好地完成维护任务。讲师可以实时将增强的真实对象投影到大屏幕上，并通过真实对象的信息和叠加信息与学生进行交流。在教育领域，魔法书是一个经典例子。魔法书看起来与普通书籍无异，也有图片和文字。但当你通过相机在监视器上看书页时，可以看到虚拟对象，这非常有趣。

为了增加作品中交互的吸引力，触发观众的情感，使他们感到参与并获得现实生活中无法获得的新奇体验。我们在作品交互设计中采用了简单的增强现实技术，叠加了真实视频和虚拟剪纸兔图形，并将反馈给观众。当参与者将手放在窗户上时，投影会实时显示红色轮廓的变化，直到虚拟剪纸兔图形出现。观众通过影响周围的光线与作品互动，并最终完成剪纸兔相应信息的显示。

本节研究了在倡导科学与艺术融合的背景下，人机交互技术在交互式新媒体艺术中的应用。自 20 世纪以来，随着科学技术的快速发展，人们比以往任何时候都更感受到科学与艺术融合的重要性。特别是计算机技术已被广泛接受并在艺术领域中得到了广泛应用。由于我国长时间以来艺术和科学的分离，技术和艺术领域的人才无法很好地交流，导致我国的交互式新媒体艺术更多地是模仿而缺乏创新。如何有机地结合交互技术与艺术，已成为一个重要话题。希望通过对国内外在交互式新媒体艺术中人机交互技术应用的研究，能够帮助国内交

互式新媒体艺术家和人机交互技术研究者相互理解彼此的领域,并合作设计出完美融合艺术与技术的作品。为我国技术与艺术融合提供一些参考和借鉴。

## **第3节 技术与艺术的碰撞—新媒体舞蹈 发展现状分析**

在当今快速发展的科技时代,人们的生活方式不断变化。人际交流已经从书信演变到微信,出行从看地图问路到卫星导航,获取信息从报纸杂志变为终端推送……我们周围的一切似乎都在变得自动化和智能化,曾经在科幻电影中期待的和大量想象中的场景已经变成现实。自然地,在人们的物质需求不断增长的同时,他们的文化需求也在增长。舞蹈作为一种舞台

艺术和视听艺术，显然无法以其传统的表现形式满足观众的审美需求。技术与艺术的碰撞，“新媒体舞蹈”应运而生[9]。

## 1. 新媒体舞蹈简介

### 1.1 新媒体舞蹈的概念与要素

新媒体舞蹈特征在于基于信息、技术与艺术语言的互动。它强调概念性、当代性、艺术性和思想性，并通过互动和参与来完成。依托信息技术，它使用文本、声音、图像、动态图像乃至表演作为载体。这是一种实时的、互动的、体验式的艺术。其作品的本质呈现了一种非物质的数据形态。信息文化的艺术形式是科学与艺术、理性与感性、现实与虚拟、大众与个性的融合。

所谓的“新媒体舞蹈”与“多媒体”概念不同，因为它随着科学技术的进步而不断发展和更新，其内容和形式也将继续丰富。它是在进展中，而不是已完成的。长期以来，理论界总是将技术和艺术视为两个界限，而不是一个统一体。然而，作为“20世纪以来舞蹈与媒体技术互动艺术实践的结晶，新媒体舞蹈是剧场舞蹈艺术通过媒体技术来解释其现代-后现代概念的基本方式；换句话说，它是舞蹈艺术借助电影电视和互联网等新媒体技术进行艺术创新的结果。”新媒体舞蹈和其他新媒体艺术类别充满活力。它们继续挑战传统艺术理论，并以创新作品激发观众的关注。

### 1.2 新媒体舞蹈艺术的发展

经过长期发展，新媒体舞蹈艺术大致可以分为两个发展方向：一是利用电视技术和网络技术，产生了专门为屏幕存在的实验性舞蹈短片，在中国，这种艺术形态被总结为“电视舞蹈”，其实质是利用新媒体作为载体来传播传统舞蹈的创新手段。这个方向的新媒体舞蹈发展已经非常成熟，它主要依靠电视和互联网等新媒体进行传播，并通过视频制作和视频编辑的技术手段，实现与剧场舞蹈不同的视觉效果。舞蹈动作的三个要素“时间、空间和力量”在屏幕上被打破并重新组织，舞蹈背景可以通过视频制作自由改变，它渐渐形成了表演空间的“屏幕化”、表演方法的“片段化”和语言结构的“蒙太奇”特点。

在舞台艺术、舞台技术和舞蹈的结合下，运用各种新媒体技术的舞蹈表演艺术形式已逐渐成为“新媒体舞蹈”概念的主流。它以互动表演、装置表演等艺术形式为代表，是传统剧场舞蹈的直接发展。本文也专注于这一发展方向。

## 2. 新媒体技术在舞蹈表演线性发展阶段的应用

### 2.1 二十世纪上半叶

摄影技术相对较早出现，开始发展是在 1870 年代和 1880 年代。进入二十世纪初，这项技术应用于剧场舞蹈主要体现在舞蹈的记录上。

原始的投影技术，其应用可以追溯到我国传统文化“皮影戏”，在 20 世纪初，主要以“幻灯投影机”的形式出现在舞台上。1924 年，德国剧场导演皮斯卡托在其剧场作品《红色时代讽刺》和《绝望》中使用了历史材料图片，创造了图像与舞台表演结合的先例。因此，传统的以简单布景为特点的舞台艺术开始变得丰富多彩，“舞台的三面墙”逐渐开始被打破。

### 2.2 二十世纪下半叶

1984 年，在美国进行了将计算机编程与拉班标记相结合的实验，并为苹果笔记本电脑专门设计了“拉班标记软件 4.0”。这款软件的出现标志着舞蹈发展进入数字时代的开始。它不仅提高了舞蹈的记录质量，还将人类运动特征整合到数字系统中。然而，随着视频录制技术的发展，这款软件的舞蹈记录功能逐渐被削弱。1989 年，第一个编舞软件“生命结构”诞生。坎宁安是第一个使用这款软件进行编舞的人。在 1990 年代初，像“计算机路径系统”和“邻居”这样的第一批舞蹈使用了这款软件。编排的作品被公开表演，这也是新媒体舞蹈历史上最早的尝试之一。同时，坎宁安还进一步发展了他的“偶然编舞”，开始使用计算机的无限可能性来取代原来抽签的方法。技术在舞蹈编排中发挥作用，创作了如“海鸟”、“邻居”和“输入”等作品。

当前的动作捕捉技术是在摄像机的辅助下进行的。摄像机采集运动图像并将它们输入到计算机中进行实时渲染和其他数字处理，然后根据需要以不同形式输出生成的图像。在实际应用中，为了提高动作捕捉技术的准确性，通常会使用 3-5 台摄像机同时进行采样。

数字投影技术，在计算机的发明和大量集成电路的出现之后，投影进入了数字时代。1989 年，第一台 LCD 投影仪诞生，这快速引发了 LCD 投影的热潮，LCD 投影技术开始进入舞台。

LED 技术始于 1970 年，绿色光管在 1970 年代末期被开发出来。在 1980 年代中期，第一代高亮度 LED 问世。这个时候，只有红、黄、绿三种颜色的灯光。1990 年代初，日本开发出了蓝光。这时，LED 灯具变得成熟，在 1990 年代末，LED 显示屏逐渐被应用到舞台上。

它结合了光电技术、微电子技术与计算机技术和视频技术。如今，LED 技术已经完善并广泛应用于各种剧场和晚会中，用相对成像较差的幕布投影替换。与此同时，剧场中的环境照明也逐渐更新为 LED 光源。

## 2.3 二十一世纪初

实时渲染技术起源于 1960 年代。最初出现的是线框模式渲染，着色语言于 1985 年出现，体积渲染技术则在 1988 年出现，而渲染软件 RenderMan 在 1990 年被开发出来，标志着非实时渲染技术的出现。进入 21 世纪，实时渲染技术开始应用于舞台，并开始向原子级 3D 实时渲染技术的方向发展。

全息投影技术起源于 1960 年代激光的出现，以及 1970 年代末彩虹全息图像显示的诞生。在 1980 年代之后，激光全息技术迅速发展。直到 21 世纪，全息技术逐渐发展并开始进入舞台表演领域，但这种类型的全息并不是真正意义上的“全息”，它仍需要使用透明薄膜和其他材料作为介质来形成图像。

现在已经可以实时追踪面部，并将其投影到面部上，实现快速变妆的效果，这也意味着可以随时为舞者“换装”。

可穿戴设备是基于 LED 技术或无线网络技术来实现特定视觉效果或人机交互的道具。如今，它的技术含量似乎并不高，但通过创意表达，它也可以形成独特的舞蹈形式。

根据新媒体技术和舞蹈艺术在表演中的作用和各自比例，新媒体舞蹈大致可以分为三个类别。首先，舞者仍是表演的主要形式，新媒体技术纯粹是辅助性的，用以提高舞蹈艺术的水平。两者都是不可或缺的。如果它们独立存在，将变得无意义。第三，科技是主要的表现形式，舞者起辅助作用或根本不需要真实的舞者。

## 3. 新媒体技术与舞蹈的结合

### 3.1 新媒体技术协助舞蹈艺术的表现形式

新媒体技术在舞蹈艺术中纯粹的辅助角色主要体现在舞台空间的扩展和背景氛围的创造上。在传统剧场舞蹈中，舞台布景相对单一且固定。在节目或场景切换时，传统舞台布景笨重且成本高昂的缺陷一览无余。观众在观看舞者表演时也需要自行创造更多的想象空间。然而，随着投影技术和 LED 技术等成像技术的应用，各种新型电子设备开始进入舞台，大型 LED 屏幕现已成为各种演出和派对的标准配置。在某些演出中，LED 屏幕甚至可以替代

所有传统布景。在舞蹈表演中，LED 屏幕可以呈现与舞蹈相匹配的图像。无论是静态的还是动态的，真实的或虚拟的，它都能直接给观众一个更清晰的环境信息传达和更直观的视觉体验。它打破了传统剧场舞蹈“三面墙”的空间限制。根据视频图像的“移动”特性，它可以借助各种先进的投影和成像技术随意创造各种真实或虚拟的空间。这可以直接将剧场中的观众带入编舞者创造的任何空间，将传统舞蹈的单一感官体验升华为三维的视觉冲击。

例如，2012 年春节联欢晚会上的舞蹈作品《鸟之恋》就为普通人提供了体验新媒体舞蹈艺术之美的机会。该舞蹈发生在深蓝和浓密森林的一角。美丽的凤尾竹隐藏在月光下，使得孔雀的美丽融入到处幽静的环境中。两只美丽的孔雀在静谧的森林中缠绵地跳舞，周围静静环绕着蝴蝶，这是人们想象中的仙境。随着两只孔雀之间的“爱情”发展，舞台背后出现了孔雀的屏幕，舞者的动作与之无缝匹配。

《鸟之恋》可说是多年来春节联欢晚会舞蹈中最佳的舞台效果之一。它是杨丽萍舞蹈作品《孔雀》的一个片段。它仍然延续了杨丽萍“孔雀舞”的风格特点，但与传统的“麻雀精神”在服装、音乐和舞台风格上都经历了巨大的变化。除了服装和音乐的进步，其华丽的 3D 效果舞台背景完全得益于 2012 年 CCTV 春节联欢晚会演播厅使用尖端技术重构的舞台结构。这支舞蹈的视频制作团队“蜀虎影像”为《鸟之恋》创造了国内舞蹈作品中前所未有的大量编舞视频。春节联欢晚会的舞台上各个角度都布置了 LED 显示屏，视频展示呈现出压倒性的趋势，LED 显示屏的任何角度都需要大量的模型、特效、后期制作、动画、渲染等工作最终呈现出高清的舞蹈画面。正是通过大量的数据制作和画面拼接，以及使用极其精确的同步控制和播放技术，《鸟之恋》的舞台美效果不仅给现场带来了 360 度全方位的视觉冲击，而且还确保了现场直播的效果。

在新媒体技术扩展舞台空间和创造背景氛围的功能方面，有一种更前卫的表现形式，那就是实景演出。它是将舞蹈艺术、装置艺术与数字艺术相结合的综合性表达。例如，张艺谋的印象系列作品“印象刘三姐”、“印象西湖”、“印象丽江”，这三部作品将演员置于真实场景的广阔空间中，强调浓厚的地方文化特色。在“印象西湖”中，大量的新媒体装置被用来通过灯光、雨幕和演员的完美结合，将表演推向高潮。许仙与白素贞的传奇故事在如梦似幻的场景中生动地表达了出来。

### 3.2 新媒体技术与舞蹈艺术互动融合的表演形式

“互动性”是新媒体舞蹈艺术的一个重要特点。一些作品是由新媒体技术与舞蹈艺术的互动融合产生的。这类作品也可以被称为“多媒体舞蹈”。没有了物理舞蹈，图像就会变得

无意义，没有了视频图像，舞蹈结构就无法建立。两者缺一不可，相互支撑。它的出现是“舞蹈与媒体技术互动”在真正意义上的艺术实践的结晶。起初，它只是编舞家试验新技术背景下舞蹈发展可能性的一个实验性舞蹈方法。然而，由于其独特的视觉体验，近年来已被广泛应用。它被用于综合剧场的音乐会和表演中。

舞者的身体动作被计算机感知，由于不同的舞蹈动作产生相应的光影效果。在实时互动表演中，数字图像呈现也因舞蹈本身的随机性而变得不可重复。最初，在剧场中的投影仅被用作传统布景的替代品，并且只扮演了协助舞蹈表演的角色。后来，艺术家开始使用预制的图像进行互动舞蹈表演。然而，这需要舞者和视频图像提前排练，以达到互动的效果，这并不是即时的新媒体舞蹈表演。在不断的探索中，编舞家分析舞台成像的媒介，如玻璃、水、雾甚至空气；思考如何让人机对话更自然，并给予及时的反馈；更先进的技术手段和数字艺术形式，如动画技术、动作捕捉技术、运动追踪系统、高清摄影和播放设备、3D 投影技术、全息投影、无线设备、传感器、视频与舞者互动系统、互动安装与舞蹈的融合等如何应用于舞台表演。新媒体技术与舞蹈艺术的结合已经开始向即时互动的方向发展。在不同技术应用的作用下，新媒体舞蹈艺术变得丰富多彩。编舞家、舞台设计师、计算机程序员甚至工程师都参与到这种互动游戏中。

在这种互动游戏中，法国表演团队 Adrien M / Claire B 可以被视为先驱者，而他们在 2014 年在互联网上广泛传播的一部新媒体舞蹈作品“Pixel”是他们的成功案例之一。这部作品由法国艺术家 Adrien Mondot 和 Chaire Bardainne 共同创作，它使用 3D 投影技术创造出一个充满想象力的虚拟空间。在这个空间中，11 位拥有不同舞蹈风格的演员与数字图像互动融合，在虚拟与现实之间徜徉，充满了想象力。荧光雨滴击中舞者的白色雨伞，虚拟的光环围绕着舞者。起伏的光之地毯随着舞者的摇摆形成一圈光图案。嘻哈舞与动态光环相匹配，创造出一则诗意的小故事，充满创造力。

Adrien M / Claire B 设计的这些看似复杂的舞蹈是使用他们自己的视频创作工具 eMotion（意为电子动画，是一个实时互动系统）创造的，它集成了图像和虚拟视觉。在舞台上呈现的像素图像都是经过算法计算生成的，最终形成一个梦幻般的舞台世界，它与现实世界的信息相结合并相互碰撞。“每一个舞台场景都是精确的构图，我们使用 eMotion 来实现这一点，并用 QIab 将声音融入到构图中，” Adrien 说。

Adrien Mondot 和 Chaire Bardainne 自 2004 年以来一直在探索如何将映射和投影技术与舞蹈相结合。那时他们创作了“Cinematique”。这部新媒体舞蹈剧在 2011 年的“中法文化之春”艺术节中巡回了我的国家。

在我的国家，互动新媒体舞蹈作品相对而言是比较成功的，例如 2008 年北京奥运会开幕式的“画卷”环节。中心舞台由 44,000 个 LED 模块组成，展现了一幅巨大画卷的视觉效果。在“画纸”上，舞者变成了画笔，自由摇摆，勾画出一幅巨大的水墨画。演员和 LED 画卷融为一体，相互诠释。在这个虚拟空间中呈现的视觉效果脱离了传统舞蹈的时空限制，生动地表达了中国悠久的文化遗产。

#### 4. 我国新媒体舞蹈的发展状况

在电视舞蹈方面，中国舞者白志群自 1990 年代以来开始全面探索和实践，并创作了以“扇舞丹青”为代表的电视舞蹈作品。如今，随着电影电视技术的成熟，电视舞蹈在我国已进入一个成熟阶段。

在新媒体剧场舞蹈方面，我国的发展相对滞后，原创作品数量较少。尽管中国出现了越来越多的“中国创造”，但大部分新媒体舞蹈作品还是“中国制造”或直接从国外成功案例引进。如上所述，2004 年由法国艺术家创作的“智能流光”在我国 2011 年的“中法文化之春”活动中演出时，仍被赞叹为富有创造性。而 2007 年创作的《斯特拉文斯基的 3D 春之祭》（在我国演出时命名为“异空间春之祭”）在 2013 年我国国家艺术团表演季作为剧目演出，各种媒体也以“中国首个 3D 多媒体交响音乐会”的标题进行了报道。另一个例子是 2015 年中央电视台春节联欢晚会节目“辉煌”，所使用的技术甚至已在各种国内时装秀中广泛应用，仍被主持人尊称为“高科技”。由此可见，中国新媒体舞蹈艺术的发展仍处于引进的初级阶段，与国外最高发展水平存在 5 到 10 年的差距。

##### 我国新媒体舞蹈发展存在的问题

首先，技术手段不足。大部分新媒体技术都是从国外引进的，如今广泛使用的计算机技术和 LED 技术等。因此，我国在探索将某项技术应用到舞台的可行性方面已经处于落后的起跑线上。

其次，舞蹈的发展在我国仍局限于基于传统舞台的编舞，缺乏个性化，对新媒体舞蹈的理解还处于“新鲜事物”的阶段。尽管北京舞蹈学院新媒体编舞、北京理工大学表演与仿真技术实验室、上海戏剧学院虚拟实验室等专业和机构也有所涉猎，但大多数参与者是艺术人才，而科技人才的连接不够，一些创意思法因各种因素而无法转化为舞台作品。

同时，经济因素也是限制新媒体舞蹈发展的一个方面。虽然我国过去 40 年的改革开放经济上取得了巨大变化，但文化发展并未与之相适应。由于新媒体舞蹈的先锋性质，任何新的创意表达都需要在舞台上进行实验性探索，并且需要团队操作，而不是传统舞蹈作品足以

应对的一个编舞者模式。高成本和多方面智力支持的需求成为了限制我国国情下新媒体舞蹈发展的一个重要因素。

## 5. 解决我国新媒体舞蹈发展中存在问题的解决方案

举办各种新媒体舞蹈节和舞蹈比赛以加强宣传力度。可以在中国舞蹈“荷花奖”和“电视舞蹈大赛”等国家舞蹈奖项中设立新媒体舞蹈比赛项目，以提升公众的认知度和认可度。也可以设立国际新媒体舞蹈展演和比赛，以加强与其他国家新媒体艺术家的交流和学习。

在高等院校舞蹈专业的教学中设立新媒体舞蹈课程，以培养新型人才。因为新媒体舞蹈艺术的创造性需要创新思维和想象力的支持，有必要加强高校之间以及不同专业之间的学术交流，并将各种交叉学科或相关学科的内容联系起来进行教学，使课堂教学内容尽可能丰富。鼓励学生尝试新媒体舞蹈，并为他们提供必要的技术和设备支持。同时，这也有利于激发舞蹈专业学生对这一专业学习的兴趣和热情。

提供一定的政策支持。首先，新媒体艺术是一种不断发展的艺术形式，需要大量的投入。因此，各级政府需要建立一个相对完善的投资机制。其次，鼓励各级艺术团体创作和演出新媒体艺术，如新媒体舞蹈，并通过剧院演出的票房收入来确保其发展，形成良性循环。第三，支持建立和发展一批新媒体艺术实验室和艺术团体，并广泛接纳舞蹈、舞台艺术、计算机技术等领域的人才加入。

### 5.1 新媒体舞蹈编舞技术的思考

在新媒体舞蹈中，创意表达取代了传统舞蹈作品中舞蹈结构的重要性，成为最重要的方面。传统舞蹈作品注重“意义”和“情感”，其成功与否依赖于舞蹈结构的质量，尤其是情节舞（叙事舞）。而新媒体舞蹈作品则注重于“新”，因为它们更注重对观众的强烈视觉冲击，使得创意表达成为评判其成功与否的主要标准。

提出了新的舞蹈词汇。传统舞蹈的“语言”是从劳动、祭祀、生殖崇拜等发展而来的，也可以说是源自自然。在新媒体的语境下，舞蹈动作的表现力被极大地扩展。例如，舞蹈中的跳跃、转圈、翻转等技术动作对舞者的能力有一定要求。由于人体的局限性，即便是优秀的舞者也只能完成更困难的动作，但不能无限制地发展。而媒体为此带来了足够的空间。从这个意义上说，新媒体舞蹈的出现可以被描述为数千年舞蹈发展中的一次重大变革。

在传统舞蹈中，舞者是绝对的主体，但在新媒体舞蹈中，舞者的主体地位被虚拟角色所取代。新媒体舞蹈是图像语言与身体语言的结合，它给观众带来的感受是两者融合的结果。

这也要求保证舞蹈动作的流畅性和图像互动的和谐性。

团队创作取代了独立编舞的创作。在传统舞蹈作品中，编舞者和舞者不一定有关连。编舞者先创作作品然后由演员进行表演是常见的情况。在新媒体舞蹈作品中，这种情况不会发生。新媒体舞蹈的独特性和不可重复性特征，以及它对电子技术、视频制作、舞台机械等的要求，使得其编舞必须得到团队的支持，并且只有在各行各业的人才合作下，才能创作出优秀的新媒体舞蹈作品。

## 5.2 新媒体舞蹈的发展趋势

新媒体舞蹈的发展在很大程度上受到科技发展的制约。在科技发展的基础上，作者认为新媒体的发展趋势将在以下几个方面得以体现：人机互动特征将更加明显。随着实时人机互动技术和实时投影技术的逐步完善，新媒体舞蹈将渐渐成为剧院舞蹈的主流形式。随着量子计算机等新型计算技术的发展，舞蹈表演中的人机互动将变得更加复杂和迅速。此外，随着增强现实（AR）、虚拟现实（VR）、混合现实（MR）等新媒体技术的进一步发展，新媒体舞蹈可能会融入更多的互动和沉浸式体验，为观众提供一种全新的观赏维度。观众不再仅是被动的接受者，而是可以通过与舞蹈表演的实时互动成为参与者，从而深化他们的体验。跨学科融合的趋势将进一步加深，新媒体舞蹈的发展需要艺术家、科学家、技术专家以及教育者之间的密切合作[10]。

舞台效果更加立体和精细。在视觉方面，成熟的LED显示技术正发展为3D显示，灵活的LED屏幕也已经出现。随着3D全息投影技术的进步，舞台影像正向无介质化方向发展，科幻电影中的虚拟成像效果正在逐步变为现实。在听觉方面，数字音乐的发展使得舞蹈音乐更加多样化，建筑声学 and 环绕声的发展也会使舞蹈音乐具有更强的环绕效果。在触觉方面，剧场中已经有使用风、雨、雾等元素的实验性作品，并将在未来更加多样化。在嗅觉方面，已经出现了气味打印机，未来可能会进入新媒体舞蹈领域。

表演空间将走进公共空间。随着视频技术的完善，舞蹈表演不再局限于剧场内，它可以通过光影的视觉效果在任何地方隔离出表演空间。广场、道路、市中心，甚至屋顶或水面的任何地点都有可能成为新媒体舞蹈的表演空间。

任何人都可以成为编舞者。随着编舞软件和网络技术的发展，新媒体舞蹈艺术的普及是一个不可避免的趋势。在某些平台的帮助下，即使是业余爱好者也可以创作出虚拟舞蹈作品，并通过网络进行展示和传播。

## 参考文献

- [1] 乔彦菊. 艺术与技术融合的个体呈现与人才聚合分析——评《新媒体装置：公共艺术中的科技创新》. 科技管理研究, 2020, 40(8):1.
- [2] 刘琼. 艺术与技术的变奏：从机械美学到生物艺术. 文艺理论研究, 2020(4):11.
- [3] 张列, 张强, 崔金银, 等. 展览中的艺术与技术. 建筑技术, 2020(5):6.
- [4] 范振坤. 场景理论框架下艺术与技术学科发展的新维度. 艺术百家, 2019(3):7.
- [5] 戴欣也. 转专业学生职业认同度的反思：以河北传媒学院艺术与技术专业为例. 文艺

青年, 2021.

[6] 冷君琦. 《数字潮流: 电子时代的艺术》: “艺术与技术关系”的回顾与反思. 当代艺术家, 2020(2):3.

[7] 戴燕良. 民间艺术与技术融合视角下的文化创意产业发展研究. 中国国家博览, 2019(8):2.

[8] 刘志毅, 冯建明. 艺术与技术融合下的新媒体设计分析. 山西广播电视大学学报, 2019(4):4.

[9] 刘伟. 现代艺术设计与民间艺术内涵的融合. 艺术科技, 2019, 32(02):182.

[10] 顾甜甜, 刘世旭, 李家恒. 马克思主义与当代艺术活动实践关系的简要分析// 2019年7月建筑技术与管理学术交流会议论文集. 2019.

## 第3章 艺术与机器

### 第1节 作为艺术创作灵感和工具的 中国神话

自古以来, 神话与人类社会的发展紧密相关。早在人类社会的原始时代, 劳动人民就以艺术的方式加工对自然力量的崇拜与恐惧, 以及对无法解释的自然现

象的理解，创造了最早的神话。神话中拥有超自然力量和明显个性的神祇，表达了古人对残酷自然环境的抗争以及对更好生活的无尽渴望。每个地区因其独特的发展和进步，形成了完全不同的神话体系。丰富的神话传说也成为了后代艺术创作的取之不尽的源泉，形成了独特的艺术创造形式乃至思维方式。本文试图结合古今中外的神话艺术创作，以及我自己在创作中的经历和感悟，简要讨论对此的理解和体会。

## 1. 古代中国神话与相关艺术创造

在古代，九州大地剧烈变化，托举浩瀚天空的四根天柱崩断。天无法遮盖世界，地不能容纳万物。不熄的夜火肆虐，凶猛的洪水席卷大地；野兽吞食人类，恶鸟掠走老人。这个世界上的生灵遭遇毁灭，景象骇人听闻。

在昆仑山上，有一位中国的女神，女娲，她感受到了世界的哀嚎。她熔化五彩石来补天之缺，断龟之肢以代替天柱，支撑天地，并杀死狂暴的龙来拯救中原的人民。她焚烧芦苇化为灰烬，堆积起来以治理洪水。

至此，天空得以修补，四极已立。洪水退尽，中原得以安宁。野兽死亡，人民安康。而创造和教化人类、守护人间的大地之母，终究耗尽了她的能量，陷入了漫长的沉睡。

这是中国最为人熟知的女娲补天神话。然而，与古希腊古罗马所代表的西方结构严谨的神话体系不同，大多数古代中国的神话传说来源于笔记、散文和诗歌，包括《诗经》、《楚辞》、《淮南子》和《山海经》。《左传》、《国语》、《庄子》、《韩非子》、《墨子》、《三五历纪》等等。因此，版本复杂，见解不一。大部分相关的艺术作品都是以还原神话内容、再现神话故事场景为主题，主要形式为壁画、插图和石刻像。

### 1.1 混沌初开——古代中国神话的出现

在远古时代，人类社会的主要矛盾还是人与自然的对立。生产力和生产关系的简单性导致人类对于日夜更替、日月更变、洪水和旱灾等自然现象感到恐惧和困惑。由于迫切希望理解自然，人们以自身为基准，想象世间万物，并无意识地对它们进行艺术加工。在这种思维之下，人类改造自然和与之斗争的过程，变成

了早期的神话和传说。

在原始公社制度下，集体劳动和生活是当时人类的常态。因此，与自然作斗争和改造自然成为了古代神话的主要内容。同时，从集体中涌现出来的能力强、技艺出众的人被所有成员崇敬，被赋予各种传说，成为了早期的“神”。

目前，中国最早记录神话的文献是《诗经》。例如，在《大雅·生民》中记载：“钜人生焉，维姜嫄生焉。维姜之烝，民之初生，生民如何？降生稷矣，实维后稷。”大致意思是姜嫄踏天帝大人的大拇印而生下了后稷。后稷亦被后世称为农业之神，在世间受到祭祀。

先秦古籍《山海经》是一部集最多古代神话且内容复杂的奇书。同时，它也是研究神话学的意义重大的作品。《山海经》分为《山经》和《海经》两大部分。其中，《山经》分为五部：《南山经》、《西山经》、《北山经》、《东山经》和《中山经》；《海经》分为《东海经》、《南海经》、《西海经》和《北海经》；另外还有《大荒经》。《山经》主要记载古代历史、禽兽、宗教等“天地之物”；《海内经》和《海外经》分别记载国内外地区的奇特地貌，而《大荒经》描述了熟悉的黄帝、女娲、禹等诸神的神话材料。

《山海经》的内容丰富，涵盖范围极为广泛。后人认为它除了在神话学上具有不可替代的价值外，还为地理、医药、民俗、宗教等学科提供了珍贵的研究资料。但在传承的过程中，并非所有学者都持正面意见。司马迁在《史记》中曾说：“至于山海经之属怪，余不敢道也，至于禹贡”，而鲁迅先生也在《中国小说史略》中称其为古代巫书。因此，今天的《山海经》也是后人修补加工过的版本。

《山海经》对后代文学艺术产生了巨大的影响。最早的《山海经》图文记录在传播过程中已经失传。目前学术界公认的较为重要的古代绘图版本有清代吴任臣的《山海经广注》及日本的《异鸟兽图》。吴任臣的《山海经广注》出版于1667年，此版刻本共有图144幅，其中71幅大多取自中国明代胡文焕的《山海经图》，画风继承了六朝张僧繇和宋代舒芜。然而，《山海经广传》的教材流传广泛且杂，包括乾隆地图版（图3-1）、近文堂图等。《异鸟兽图》成书于江户时期（中国明清时期），初步研究，《全像山海经》“导论”书中有76幅图，其中66幅也见于胡文焕的《山海经图》。在着色、神名、风格和构图上有许多相似之处，但也加入了一些日式风格。

除了《山海经》的各种插图和图册之外，古代还有大量的岩画、壁画和铸造雕像，同样展现了对神或神话的描绘。这些艺术作品不仅反映了古人对自然和神秘事物的认知，也体现了他们的想象力和创造力，为我们今天理解古代文化和信仰提供了直观的窗口。

例如，在连云港市将军崖的岩画中，第三组东侧有四个与古代传说中的神有关的人面，面部特征由短线和点表示，其中两个头上雕刻有类似羽毛的装饰，面孔之间混杂着点和符号。将军崖岩画也是中国发现的最古老的岩画。著名考古学家苏秉琦先生称之为我国最早的天书。在阴山的岩画中，有古代巫师向天献祭和祈祷以娱乐神灵。在随县雷古墩1号墓中的棺槨上，也有持双戟守护的神的雕像。



图 3-1: 《山海经》注释 (图片来源: [https://www.sohu.com/a/531937307\\_121124714](https://www.sohu.com/a/531937307_121124714))

不仅岩画和壁画，商周青铜器上的动物纹样也在沟通人与神的世界中扮演了重要角色。现存出土青铜器上的动物纹样多为饕餮、夔、龙、囚、凤、肥遗等。这些想象中的动物形象都是中国古代神话中的经典神兽。

可以看出，从原始时代到古代，人们的神话思维非常发达。在这个时期，神话和巫术仪式往往是相伴随的。神话内容也与生产和生活密切相关。神话类型主要有传承神话、发明神话、战争神话等。相关的艺术作品都出现在当时人们用于祭祀或记录的形式中，它们没有自己独特的艺术表达或语言，并不仅仅为了艺术作品本身而出现。

## 1.2 时代的变迁——中国神话的演变

古代神话世代口头传下来，很少有专门的古籍对其进行整理和总结，加之时代变迁不可避免地导致了缺失和变异，许多神话有着不同的版本和说法，内容本身也发生了微妙的变化。

中国古代神话大致经历了从精神性到神性再到人性的阶段。在朝代的更迭中，通过后代的加工和转化，发展成为了横向的多元化。

首先，由于对故事完整性和因果关系合理化的追求，人们在传播过程中对其进行加工结合，重新组织原初的神话。比如本文开头提到的《淮南子·览冥训》中记载的女娲补天的故事，就是将各种女娲的传说整合的结果。《山海经》对女娲的描写笔触极少，《淮南子》也未提及女娲补天的前提，但在后世广泛流传的各种版本中，是因为水神共工与火神祝融斗争，共工败后愤而触碰不周山，引发天下大变。后来女娲忍受不了人民的苦难而上天补天。

其次，为了宣扬自己的理论，后来的思想家们有意识地艺术性地加工神话，使之更便于寄托其思想并增加其意义。春秋战国时期的儒家致力于“神”的人性化和奇异故事的合理化，这一点在史书中有记载。比如，在古代，西王母的形象还是兽、神、人的混合体，更接近于怪物的形象。作为刑罚之神，西王母居住在被弱水包围的玉山（也有说是昆仑山），外有烈火山。到了西周时，《穆天子传》记载了周穆王与西王母相处的场景，西王母开始有了人性，开始变得世俗化。到汉代《汉武内传》和《汉武故事》中，西王母的形象已完全人性化。到了明代《西游记》中，因道教的影响，西王母的形象也是从西王母传说演化而来。至此，西王母已被完全人性化。

最后，古代神话随着人类社会的发展而不断演进，神话开始变成宗教，这也是整个人类历史中的普遍现象。在封建社会的统治下，宗教力量强大，神话的演变也受到了极大的影响。西汉时期，中国神话从以英雄、神、祖先和图腾为核心的古代神话体系变为以五帝体系为核心的神话体系。东汉时期，佛教传入中国，出现了佛教神话。魏晋南北朝时期，佛教为了巩固其主导地位和社会发展，得到了前所未有的发展。

在神话的演变过程中，产生的艺术作品，虽然仍以附加形式出现，但开始呈现多样化。无论是形象还是目的，都更为清晰。例如，在艺术表现上，我们可以看到壁画、雕塑和文学作品中的神话人物和故事随着时间的流逝而发生变化，反映了社会的价值观和宗教信仰。

在这样的发展中，神话不再仅仅是古老的传说，而是成为了一种能够表达哲学思想、历史观念、民族精神和文化认同的媒介。神话中的神祇和英雄形象逐渐与当下的文化价值观相融合，被赋予了新的意义和象征。

例如, 道教的兴起使得原本的民间信仰和神话故事被重新解读和包含进道教的宇宙观中。道教中的神仙体系, 如八仙的故事, 就是从各种民间传说中提炼并融入了道教哲学的典型例子。而在文学作品中, 如《西游记》对神话进行了再创造, 将佛教、道教以及儒家思想融合其中, 创造了一个复杂而丰富的神话宇宙, 这反映了明朝时期的宗教和文化多元性。

从这些演变中, 我们可以看出, 随着社会的发展, 中国神话不断地适应新的文化环境, 体现了人们对于宇宙、社会和人生的认识与追问。这些神话的转变和发展, 不仅丰富了中国的文化传统, 也为我们理解古代人民如何在不断变化的世界中寻找意义和指导提供了窗口。

敦煌壁画中的“降魔变相图”收藏于法国国家图书馆, 它详细描绘了佛教神话中舍利弗与罗睺罗的战斗场景。从这幅画卷中, 我们不难得出结论, 当时的神话艺术作品已经打破了原有的传统束缚, 开始追求更加丰富的表现技巧。而且, 的确为后来的《西游记》、《封神演义》等神魔小说创造了先例。



图 3-2: “降魔变相图卷”

(图片来源: [www.360doc.com/content/18/0203/17/19519242\\_727460250.shtml](http://www.360doc.com/content/18/0203/17/19519242_727460250.shtml))

古中国神话体系无疑为后世的艺术和文学创作营造了一个庞大的世界。在传播和演变的过程中, 以古代神话为主题的艺术作品也形成了它们独特的艺术风格。后人学者认为, 中国的神话是中国文化的源头, 它极大地促进了中国艺术和文化创作的繁荣, 并且对民族精神的形成产生了深远的影响。

## 2. 当代杰出神话题材的艺术创作

### 2.1 出现与毁灭 - 中国当代艺术中的神话再创造

20 世纪初, 中国的封建社会走到了尽头。社会动荡, 内外侵扰导致的政治经济打击, 同时间文化和艺术也经历了强烈的冲击。直到 1980 年代改革开放之后, 艺术界才真正稳定下来。然而, 中国人似乎已经完全接受了西方现代和当代艺术的表达方式, 而中国本土文化似

乎在衰退。当代艺术如何与本土文化建立内在联系，成为艺术家们面临的首要问题。

在这个时期，国内艺术创作中出现了一个不容忽视的现象。许多艺术家创作了大量明显不同于主流社会意识形态的作品，这些作品基于神话、民间传说以及其他神秘与幻想的内容。黄永录在 1999 年威尼斯双年展上展出的《人与九兽》无疑具有重大意义。首先映入眼帘的是一个老式指南针，上面立着一个铜制的人形。他的手指向展厅。当观众站在与他相同的位置时，他将在展厅的墙壁上看到一个词“ANCIA”——意为古老。绕他稍微走动一下，这个词就变成了“RANCIA”——自由翻译是陈旧。而当观众继续走向展厅的门时，接着他就能看到“FRANCIA”这个词的全貌——自由翻译是法国。然而，即使他能清晰地读到这些词，前面还有树干和九只长爪怪兽挡住了道路。



图 3-3：黄永录，《蛇杖》

(<https://m.027art.com/art/gzh17/8886185.html>)

安装艺术中怪兽的原型来自《山海经》，但其含义远远超出了对传统中国神话的解读。由于他当时代表法国参加展览，这件装置艺术更像是他从一个相对模糊的视角，对中西文化以及历史上中国殖民主义问题的个人见解。

在《20 世纪中国艺术史》中，陆鹏评论黄永录是“用东方打西方，用西方打东方”。在某种程度上，这实际上是相当准确的。他后来的作品《蛇杖》就是这种观点的一个确认。这件作品的灵感来源于《出埃及记》中的一个故事，这个故事被犹太教和基督教都使用——摩

西带领人们逃离灾难，陷入绝望之中，上帝显现了神迹。黄永砅动员了不同的宗教符号，让它们在冲突中共存，达到了矛盾与和平之间的奇异平衡。



图 3-4: 蓬莱山 (<https://www.027art.com/art/gzh14/8896905.html>)

在当代社会的复杂背景环境下，以神话为题材的艺术作品放弃了过去二元对立的艺术思维模式。虽然黄永砅作品的原型来自于神话或民间记忆，但通过艺术家对自己作品的“重新编码”，原有结构被改变，形成了一种新的叙事方法。这种方法更适合当代视觉艺术，更适合表达作品自身的意义，更加紧张并且更能打动观众。

另一位当代艺术家蔡国强在日本地球艺术节的大型动态装置作品《蓬莱山》同样基于中国的神话传说。（图 4）准确地说，应该是取自秦汉时期的神话观念。战国时期，六国不断战争，战火纷飞。百姓长时间没有稳定的居所，因此梦想着超脱凡间的仙境，在那里不仅可以逃避战争的伤害，还能养生长寿。受神话思维的启发，追求一个适合自己需要的仙境。秦汉统一后，统治者在自己欲望的驱使和对长生不老的极端追求下，命令大小官员寻找蓬莱仙境，因此从上到下掀起了对神的狂热信仰。巫师变成了炼金术士。为了自己的名利，不惜迎合统治阶级，炒作传闻中的蓬莱、阿房、满洲里等海上之州，统称为东方蓬莱仙境。当然，其中

有些人对当时的文化做出了很大的贡献。这与蔡国强的创作意图不谋而合。他希望通过美丽的神话之地“蓬莱”，反映东亚国家间对岛屿所有权的持续争端，并希望以太平洋战争 70 周年为契机，呼吁东亚内部和平。火药爆破画《岛屿》是在当地稻草艺术大师的指导下由儿童完成的，完成了近 1000 件由稻草制成的作品，最终围绕岛屿形状的装置形成龙的形状。同时，蔡国强的作品也围绕死亡和战争展开。由于艺术表达的特殊性，它可以被称为一场无声的战斗。自从火药被广泛用于武器制造和实战以来，艺术创作的意义变成了一个避免血腥和残酷的材料。那些艺术是对血腥场面的追求，可以通过火药爆炸的真实感受在艺术中得以反映，并通过艺术唤起人类对战争残酷性的直观感受。蔡国强的作品投入了巨大的精力以唤起人们对战争的恐惧。在这里也需要提到他的艺术手法。火药最初是人类征服和改造自然发展的产物，但现在其在战争中的应用发挥着不可替代的作用。在这个意义上，它也与他想要表达的理念形成了非常强烈和惊人的对比。就真正想要参与全球政治而言，能够获取一些信息无疑是重要的，但更重要的是，不忽视我们现在所处的时代背景和全球形势。纷扰不断的冲突和战争，不论大小，艺术家赋予了自己在战争中思考的责任，提取思想，通过不同的艺术表达描述其目的，并呼吁世界珍惜来之不易的和平。

蔡国强的艺术实践表明，他使用的材料和主题——火药和神话传说——虽然看似矛盾，但实际上是用来探讨和平与冲突、生与死、传统与现代之间复杂而微妙的关系。他的作品通过对传统文化的重新解读和象征性材料的使用，提出了对当代社会问题的深刻反思，特别是在处理记忆、历史和文化遗产方面。通过这样的艺术创作，蔡国强不仅在视觉上提供了震撼人心的体验，而且在情感和智识上挑战了观众，促使人们重新考虑和平与共存的可能性。

## 2.2 万象更新 - "神话思维" 的出现

狭义上的神话思维是指古人关于他们的精神世界的认知和思维模式，这个世界在神话中被描绘。例如，当道教在秦汉时期盛行时，那时的人们相信不朽之道，并且相信修炼各种仙术可以延年益寿。在那个时代，许多至今仍被传颂的著名道家人物就此诞生，普通人也形成了一种独特的思维模式。

在现代科学世界，随着先进技术和理性思维的影响，似乎不可能产生新的神话思维。然而，我们可以从上述事实看出，越来越多的艺术作品在寻求追随传统文化的脚步时，选择神话作为切入点。围绕它所产生的艺术作品以及艺术家工作的行为和思维模式，也形成了一种不容小觑的独特新力量。

许多艺术家选择从中国神话或碎片化的民间记忆出发，重构他们自己独特的创意思想。他

们可能会从古老的神话传说和当地的民间习俗出发，结合古代形象和所有可用的视觉材料，寻找与自身目的交汇点，并重写他们的艺术语言，使其作品具有新的意义。例如，邱志杰、蔡国强、黄永钻等；将神话和宗教融入自己的作品中，同时表达思想，他更注重叙事过程本身，具有浓厚的文学色彩，如孙逊等。

这些艺术作品已经超越了20世纪80年代初中国当代艺术的初期阶段，它们不是“挪用主义”的产物。因此，我个人认为，随之而来的艺术思维模式是全新的，与任何时期的神话思维都不同的新一代思维。这种思维模式的出现不仅意味着由它衍生的艺术作品与神话主题密切相关，而且还显示出中国当代艺术正在逐渐摆脱西方艺术创作体系对我们的影响。这也是一种脱离以前某个时期生产的僵化统一的艺术作品的方式，在不断的改进和转变中，艺术家自己的风格已经融入其中。因此，这种新鲜而有力的新力量是艺术创作领域中不可或缺的重要力量。

### **3. 神话体系与以古希腊和古罗马为例的艺术作品**

在古希腊时期，已经形成了严密的神话体系，主要表现形式是文学作品，例如《荷马史诗》、《神谱》等。内容涵盖了非常广泛的范围，主要基于战争和创世纪的英雄史诗。同时，在当时的陶器和其他日常用品上出现了各种几何图形，如由极细的笔触描绘的各种文学或三角形和圆形图案。内容包括人物、鸟类、野兽等，后来演变成神话人物或口耳相传的事物。

与中国神话不同，在大多数时期的中国神明没有喜怒哀乐，无欲无求，具有不朽和至高无上的特性。古希腊的神则更接近人类。就像人类一样，他们有喜悦、悲伤、快乐和自私的欲望。他们与人类唯一的不同是神拥有无限的生命。除此之外，他们与人类无异。这样的神似乎更贴近当时古希腊人的生活和内心的精神需求。

在这一时期，举世闻名的古希腊雕塑也应运而生。雕塑成为古希腊不可磨灭的成就之一，同时它也影响了远东后来的艺术作品。古希腊雕塑最成功的特点在于艺术家们形成了自己严格的比例和视角关系。他们无差别地、细致地观察和研究他们所描绘的对象，并以神为主题制作了大量生动严谨的雕塑。

这一时期的古希腊男性雕像有一种单一的形式，大多数都是站立的，一条腿在前，另一条腿在后，双肩上有头发，双手垂于身旁并紧握，身体裸露。当地的宗教观认为，为了展示神与人的不同，并反映他们完美的姿态，最好不要有衣物遮盖。无论他们的地位、年龄大小，古希腊的雕像总是面带一种平静、不变的含蓄微笑。

与古希腊不同，古罗马直到公元前三世纪之后才模仿古希腊的神话体系，并编撰了自己的

神话传说——一套关于自身诞生和发展的丰富而有逻辑性的传说体系。与希腊神话最明显的不同是，在罗马神话中，人是主体并发挥主要作用，神只是偶尔介入。后来，随着罗马的军事扩张和被占领国家的影响，罗马的神话体系也发生了变化和扩张。罗马神话的最具代表性作品是罗马的神庙。如万神殿和朱庇特大神庙。

神话的再创造已经成为当前艺术环境中许多不可忽视的主题之一，而丰富的中国古代传统文化是取之不尽的源泉。在西方体系的影响下，中国当代艺术也在寻找自己的起源。越来越多的艺术家，在表达自己情感的同时，从传统出发，运用自己独特的艺术表达方式，利用民间文化和记忆来转化和替代传统资源，这成为中国当代艺术的一个重要力量。

## 第2节 玉：神话与艺术的嬗变

在神话与艺术相撞的过程中形成的人工结晶是玉器。著名的英国人类学家爱德华·伯纳特·泰勒在他的著作《原始文化》中对文化进行了详细的讨论。他提到，在广泛的民族学意义上，文化或文明指的是作为“社会人”的存在所具备的知识、信仰、艺术、习俗、道德和

法律以及其他任何才能和习惯的总和。也就是说，自人类诞生以来，文化便随之出现，而文化是物质财富与精神财富的总和。它包括了物质文化、制度文化和精神文化。

神话是古代祖先及后来人民集体智慧的结晶，是在历史循环中创造的精神财富，也是无数代人口口相传下来的创造物。作为人类最古老的文化遗产，它是世界文学的原始来源。它具有高度的哲学和艺术价值，自古以来一直是创作的源泉，它是不朽的。它孕育了文人和民间艺术家的创作，并影响了人类的精神和思想几千年。它对后代文化和文学的发展产生了深远的影响。

世界上所有民族都有自己丰富而衰退的神话，与更为系统和成熟的古希腊神话、古罗马神话、古埃及神话等相比，古代中国的神话显得有些枯萎。沈延宾先生在他的文章《中国神话研究》中曾说：“关于中国神话的书籍一直非常散乱，没有被汇编成书的作品。”因此，整理一本关于中国神话的书籍是极其困难的。神话记录在系列著作的零散片段中，如《山海经》、《诗经》和《淮南子》，在口头时代，神话的意义似乎停留在神话文本的层面。也就是说，所谓的文章和著作是神话，并非如此。这些《山海经》的神话文本不等同于神话，它们并不以这种单一形式流传，当代学者也重新解读并发现了神话在多种形式中的新价值和意义，跨学科研究为神话的生存开辟了新天地。在他的文章《伟大的传统：神话与艺术史的人类学转向》中，叶舒宪总结了神话与艺术视野的融合现象，并提出了当代研究中几个较为流行的领域，例如：1. 从史前时期到古代的玉器图像研究；2. 青铜器图像研究；3. 萨满艺术研究；4. 汉代肖像研究等。充分证明了从古至今，在神话永恒的特性下，古人为当代研究提供了直观的视觉符号材料，从而确认了神话研究的无限潜力。

## 1. 神话思维与美学意识的出现

### 1.1 神话

对于神话，当代人并不会感到陌生，它出现在现实生活的各个角落、文学作品、影视艺术、二次元文化等等。今天中文中的“神话”一词是日语的翻译，并且它的根源在古希腊词“muth”或“mythos”，意为“故事、传说、叙述”等。但要给它一个确切的定义，即什么确切是“神话之言”，有点措手不及。在研究过程中，许多专家学者对它的定义有不同的看法，不同的学者有不同的观点。著名的日本神话学家大林太良曾这样描述这一现象：“可以毫不夸张地说，研究这个问题的学者有多少，就有多少神话定义。”

一般来说，神话最初出现在原始社会，是原始祖先以自己的方式想象和幻想，按照自己

的心理需求，解释世界起源、人类起源、未知自然和周围神秘力量的传奇故事。简单地说，它是关于神的故事。作为一种特殊的精神产品，神话是人类思维能力和语言技能发展到一定阶段的产物，它反映了这一时期的社会物质生产条件和相应的精神风貌，也是人类劳动实践和征服自然的产物。

## 1.2 神话思维

神话思维是用来创造神话的想象力，它蕴含于神话之中，是一种关于物质世界和精神世界的现实性的把握和认识的方式。这种方法与在文明时代建立后基于概念上成熟的逻辑思维不同。它是在从无知时代向文明时代过渡的思维发展过程中形成的特定思维方式。它体现了原始祖先对周围自然和生活的一种神性理解，是人类认识自然和控制自然的手段，也是在原始社会萌芽中出现的对超人自然的解释。它标志着人类心智的演化模式，从低级的原始思维进化到高级的逻辑思维。

神话思维的特点主要体现在两个方面：一方面，意味着在心理活动的过程中思维或想象始终停留在具体的表面层次。这可以理解为原始祖先在认识周围环境时，对具体事物的感知形成的直观印象，在心中再现时从未经过加工。另一方面，思维和想象总是“不断切割，合理化和混沌”的情感，而作为主体的人始终附着于客体。这可以理解为在原始祖先的认知过程中，当客观事物没有呈现在人面前时，人的心智无法形成一个概括而清晰的情感形象。

由于原始人类的生产力水平相对较低，抽象思考的能力也相对较低。虽然也有数量的积累，但没有从数量变化到质量变化的转变，不能在积累大量表面现象的基础上进行有效的改进，没有从感性认识到理性认识形成逻辑思维，更不用说从感性认知形成理性认识的升华了。

## 1.3 艺术思维

正如上文提到的，原始人类的思维模式是神话思维，那么它与艺术思维有何种联系，我们不可以用科学认知来考量神话思维。作为人的主观幻想，神话思维处于混沌、荒谬和缺乏理性逻辑的领域之中。它是整个人类社会进程的童年时代，具有最原始和最低级的思维方式。然而，与神话思维相比，不难发现，神话思维包含了某些艺术形象思维的元素，反映了艺术思维的特征。

首先，我们必须理解艺术思维的含义，那就是创造性的想象力，作为形象思维的基本元素，它是以概念重现具体表象，并伴随着一系列处理转换、分解和组合来创造新的艺术形象。

艺术思维的主要特征主要包括两个层面：其一是艺术思维的创造想象过程是直观的。其

二是表象的概括过程，艺术思维能够归纳和总结社会生活。因此，艺术思维创造的形象在一定程度上能够反映社会生活的现象，从而反映社会生活的本质意义。

艺术思维与神话思维的关联在于两者都通过象征和隐喻来传达深层次的意义。神话思维通过讲述神秘的故事来探索人类、自然和宇宙的起源，而艺术思维则运用象征、隐喻和创造力来表达个人和集体的情感、想法和社会批判。在神话故事中，人类的恐惧、愿望和道德观念常以寓言的形式体现，这与艺术作品中传达复杂情感和社会评论的方式有着本质上的相似性。

艺术思维对于形象的捕捉和创造不仅限于视觉艺术，它还包括音乐、文学、舞蹈和戏剧等领域，在这些领域中，形象和隐喻通常用于传达超越文字和语言的情感和理念。因此，虽然神话思维和艺术思维在形式上可能不同，但它们都致力于通过形象和故事传递人类经验的深层次意义。

## 1.4 神话思维与艺术思维的关系

在这里，我们不探讨神话与艺术的发生问题，无论是神话不属于艺术，还是艺术起源于神话。我们的关注点应该是导致早期人类艺术形成的因素，以及为什么历史遗留下来的艺术作品的形象特征与功能特征相应地呈现。这归结于原始思维模式背景下艺术的出现和发展。

在创造新的艺术形象方式方面，神话思维的想象与艺术思维的想象是一致的，两者都是对表象进行处理、分解以及重塑和组合。值得注意的是，二者在本质上实际上是有区别的：首先，思维方式不同，神话思维属于感性思维，它是感性逻辑的类比和混沌的聚合；艺术思维是感性思维上升为理性思维的产物，它受到接受逻辑的指导 and 约束；其次，创造艺术形象的目的不同，神话思维有明确的功利目的，并且在艺术创作中是无意识产生的；艺术思维主要是为了审美目的，艺术创作是有意识形成的。

马克思曾经说过：“希腊神话不但是希腊艺术的仓库，而且还是它的土壤”，这也适用于理解中国古代玉器发展历史的认识。一方面，中国古代神话为中国古代玉器的创作提供了原始材料和艺术的精神意义，另一方面，它利用无意识的幻想方式进行创造性想象，对秦代以后艺术家的艺术思维发展产生了巨大影响。

神话思维与艺术思维之间的联系表明，艺术的发源和发展是与人类的思维方式、宗教信仰、社会结构和文化传统密不可分的。早期的艺术形象往往根植于神话故事和宗教仪式中，这些故事和仪式不仅为艺术家提供了丰富的视觉和主题素材，也为艺术创作提供了社会功能和文化意义。因此，虽然神话思维和艺术思维在目的和方法上存在差异，但它们都共享着形

象思维的特质，并在人类文化和艺术的发展中发挥着不可或缺的作用。

## 2. 神话与艺术的载体之玉器

### 2.1 玉的区分

东汉许慎在《说文解字》（卷一）中记载：“玉，石之美者，有五德”，这是最早见于许慎“释词”。中华书局1963年版第10页。对“玉”的定义多次提出，指出既然玉可以被描述为“石之美”，这就意味着玉与石的关系。可以看出，玉是从石的自然属性中衍生出来的，这表明玉在材料上比一般用于生产工具的石料更加美丽。在长期寻找石制生产工具的过程中，古人发现了一种特殊的石头，它比普通石头更坚韧、更美观，被用来制作工具或装饰品。从视觉上看，玉的颜色丰富多彩，有些甚至透明明亮；从听觉上，敲击时的声音清脆悦耳；在这三种美感的影响下，原始先祖们用“玉”这个词给它赋予了新的身份，这是玉和石分化的起点。

### 2.2 玉的起源

玉石被称为神石和巫石。首先，我们来看古人是如何解释“玉”的。在甲骨文中，“玉”字写作“仿”，像是用绳子穿起来的几块玉石的图像。“巫”在甲骨文中是“事”的意思，而在《说文解字》中记载：“巫，亦，以玉为神，从玉”，根据清代段玉裁的注释：“巫可以用玉来做神事，所以其字从玉。”古代的巫师被称为“巫”，在《说文解字》中也有记载：“巫，祝也，女能隐形，舞于神者也。男能压神，在男与巫。”从这里可以看出，作为象形字，“巫”是两个玉字重叠的结果，它像是巫女挥舞袖子跳舞的样子。因此，这是“巫”和“玉”之间起源相关的解释，现在被学术界认为是巫与玉神早期的象形字。

人必须处理无形的、看不见的和无法言说的事物，这就是巫术、艺术和宗教，是一种形而上的活动。由于后者的需求，最早的精神生产者 and 心智劳动者就此产生，即那些集原始哲学家、神学家和艺术家于一身的巫师、术士、女巫和巫女。他们的任务是通过各种艺术性的巫术活动或巫术性的艺术活动来同情、捕捉并掌握事物的神秘力量和神秘属性，从而为人类的利益控制它们。

由此可见，作为一种独特的存在，巫医在我国原始社会中存在。作为原始宗教巫术的主人，他或她也是巫术实践的参与者。他或她具有特殊的宗教哲学功能和技能，在与神对话之后，他或她将神的意志传达给人民。他或她是古人中的特殊人物，也可以被认为是原始宗教

权威下的领袖。到目前为止，我们对他们的了解仍然处于初级阶段，我们只能看到女巫的坟墓，但我们还没有找到他们的居住地和巫术活动的遗迹，现在我们只能通过大量的文献和考古挖掘来推测女巫的社会地位和历史贡献。

玉石之所以能够与普通石头区分开来，被定义为“神之石”，不仅因为它自身的物质属性，更重要的是，它还具有其他的物质和精神属性，如文化属性、社会属性等，并且使用符号作为载体，形成文化观念的表现。基于拟灵论，比如自然崇拜的观念。人们将玉视为一种精神上的神或神器，其雕刻的形式自然与神性有关。在原始时期，首次出现被定义为“神色”的玉，这赋予了它不同的意义。除了“以玉为兵”，在劳动实践中使用玉作为生产手段外，它更重要的一点是“物之神”。玉不仅作为神渗透进原始祖先的内心，也作为神物的概念存在，并且还是巫师与神沟通的重要媒介。创造与神沟通的渠道是制作玉，而神即是玉。

## 2.3 玉

### 2.3.1 玉

玉从自然物提升为器物后，它承载了丰富而多重的文化内涵，其本身的概念也有了质的飞跃。在漫长的石器制作过程中，玉的固有质地和色泽给人们带来了美好的感受，并且常常被制作成装饰品来装饰自身。在史前遗址早期的大多数玉器是生产工具和装饰品，这证明了人们最早关注的内容偏向于玉的实用价值和审美意义，实用性和装饰性是原始人对玉的首要需求。这些装饰品本质上是“从胚胎形态过渡到成熟形态的史前艺术作品”。

早期玉文化的主要特点是玉具有原始的审美价值，使用玉作为装饰品来装饰自己和生活，这是原始祖先赋予玉的原始社会功能。在新石器时期，作为巫师手中的祭祀器物，玉与巫术紧密相连，玉被赋予了神秘的力量，在原始宗教中作为“玉器”发挥着重要作用。正是因为原始审美和神性的双重作用，玉的功能发展和审美特征展现出继承和变化。

### 2.3.2 玉学理论基础的哲学取向

上述提到的玉学理论基础的哲学内容包括神学和美学。

## 1) 神话是神学的基础

神话是在讲述神的故事，而神学体系的形成前提必然需要对神的信仰，所以神话是神学发生的必要条件。在对神的信仰下形成的神话为巫术的发生提供了基本条件，而巫术通过巫术开辟了神话的新领域，巫术的成熟是神学确立的必要过程。

赋予玉的神色，通过不同的表现形式和借助载体的帮助，从古至今被传承下来，并且不同阶层和不同身份群体对玉的功能有着象征性的理解，如祈祷、祝福、获得某些力量等。

## 2) 艺术与美学

美学是研究人与世界之间美的关系的学科，美学的对象是审美活动。艺术之所以成为艺术，除了人的生产条件之外，更重要的是它具有审美价值，即审美特征 [6]。

艺术价值。人们通过拥有艺术价值来表现他们的身份、地位、财富和权力。不同历史时期和不同身份的群体对这种艺术价值表现出了渴望，目的是进一步满足个体的审美需求和精神追求。

首先，祭祀。原始社会进行巫术活动，巫师有与神交流的能力，用于交流的重要神器便是玉。玉作为祭祀器皿，拥有极其重要的社会地位。原始居民认为玉美丽、坚韧又沉默，相信玉具有理解神意的能力，因此玉得到了作为“玉器”的新身份。国家出现之后，“玉器”的身份演变为“王玉”，经常被用作皇帝祭祀活动中的“礼器”。《周礼》记载了一个玉制，包括“璧”、“琮”、“瑗”、“璋”、“黄”和“璜”六种器物，它们被用来祭祀天地四方。可见，在古代的礼仪活动中，玉占据了重要位置，且用于礼仪活动的玉器种类繁多。

其次，权力、等级和财富的象征。从迄今为止发掘的原始社会墓葬中，可以推断出墓主的身份高低在所埋葬的玉器类型和数量上有所体现。国家形成之后，玉的实用礼仪和器具被严格规定，并被编入典籍作为文字记载。这样做的原因是朝代的统治阶级为了维持礼仪制度，避免混乱和阶级层次的越界。

第三，审美装饰。玉器因其材质的特殊性被原始祖先用于装饰和美化。在新石器时代，已有使用玉器进行装饰的习俗。原始祖先将玉制作成各式各样的饰品，主要用于佩戴。从古代一直到明清时代留下的玉器艺术品显示，人们的审美能力不断提升，工艺艺术的发展也得到了积极推动。

第四，玉葬。夏鼎在其文章《汉代玉器的传统续变与汉代玉器》中定义了玉葬，主要是为了保存尸体的目的而制作的玉器。值得一提的是，汉代的玉衣在中国古代玉器史上以其完

整的形状和精湛的制作工艺而独树一帜。学术界普遍认同的玉葬功能是认为玉可以保护尸体，助魂飞升，成仙。

第五，吉玉。玉自古以来就被奉为神物，可以向神明祈求吉祥，得到神灵的庇佑，以及避邪挡煞，因此被视为吉祥之玉。

### 3. 神巫世界中的古代玉器

新石器时代开始出现大量的玉器，随着自豪势头的快速发展，新石器时代中期影响最大，相互之间可比拟，最典型的是辽东地区的红山文化玉和长江流域的良渚文化玉，共同代表了中国石器时代玉文化发展的最高成就。

#### 3.1 红山文化玉

##### 3.1.1 概述

在中国东北的辽河流域发掘的红山文化，距今约 5500-5000 年，出土了大量的玉器。作为新石器时代的重要文化遗址，玉在这一时期得到了发展和繁荣，以动物形玉器最为鲜明。20 世纪 30 年代，红山文化被发现。1954 年，它被命名为“红山后遗址”。红山文化遗址位于辽宁、内蒙古和河北的交界处。

红山文化的玉器材质主要是透闪石，颜色有黄绿色、黄白色、深绿色和浅蓝色。出土的玉器工具主要包括：玉璧、玉斧、玉凿、玉锥、玉棒等；主要装饰品类有：玉环、玉镯、玉珠、玉管、玉臂饰、玉饰件、玉菱形饰物、绿松石护身符等；主要人物形象类有：渔人、人面饰物等；动物形象玉器主要包括：玉龙、玉猪龙、玉鸟、玉龟、玉鱼、玉蚕等；特殊类别有：勾云形器、玉璧、环形器、三孔形器、玉璜、玉环等。其中，动物形象玉器是红山文化玉器中艺术成就最高的，其文化特点也是最典型的。

在红山文化的玉器中，有玉龙、玉鸟和云形玉，可以看作是物的神的“玉器”。

##### 3.1.2 红山玉龙

红山玉龙的形象特征

龙的形象在中国传统文化中占有重要的民族色彩，中国人用“龙的传人”来形容自己的民族，

封建时期的皇帝用“真龙天子”来自称，以显示其崇高至极的王权。新石器时代的玉器开始出现龙的形象。在《说文解字注》中记载：“龙能大能小，能长能短，遇春风则天，遇秋风则渊。”《尔雅·释兽》中对龙的解释是：“角如鹿、头如蛇、目如兔、项如蛇、腹如幻、鳞如鱼、爪如鹰、掌如虎、耳如牛。”红山玉龙证明了红山文化的玉器崇拜色彩已达到一个成熟阶段。红山文化玉龙主要有两种形态，即三星达拉“C”形龙和玉猪龙（玉兽形龙）。这两者的造型特点相同但又有所不同，都是以猪的头部特征作为龙的头部，以蛇的身体作为龙的身体。红山玉龙的形象应当是在图腾崇拜现象成熟后，以猪和蛇两种动物作为崇拜对象的基础上，综合了两者的形象特征创造出来的，这也更符合原始社会的生产和生活环境。

红山玉龙的整体呈蜷曲状态，尾端不相连，龙头以猪的特征为主，兽头丰满，眼睛突出，鼻孔为圆形，鼻间有多重褶皱，吻部前伸，口中有暴露的獠牙，背部有穿孔。

在原始社会中，猪具有一定的实用功能，随着养猪能力的出现，猪具有强大的繁殖能力，也是原始祖先的食物来源，在这个时期，猪象征着丰收和食物的丰富。

在他的《红山玉龙多面解析》一文中，肖冰先生提到，一旦猪进入水中，它们会变成“江豚”或“河猪”，它们对风暴或风浪非常敏感，常被用作吸引雨水的“依靠”，并被当作能引发暴雨或洪水的神器。

基于上述猪的重要社会功能，原始祖先崇拜猪作为图腾，甚至用它作为部落的名字。

《山海经》和《中山经》记载：

至于大马之山，十中有九山，一百八十四里。其十六位神均为人首猪身。

大致来说，从苦山起，经过扈山和山至大七山，共有十座山，一千一百八十四里，其中有十六位神，皆为猪面[图4]。在他的《蒙古艺术研究》中，阿穆尔·巴图说：红山文化的原始人最崇拜猪龙为神，甚至他们的至高无上的皇帝都有“猪人”的特征。由此可见，原始部落驯养家猪并将猪崇拜为部落图腾并非不可能。

红山文化的玉龙身体部分采用了蛇的躯干形象特征，蜷曲成圆形，表面光滑，身形细长，无角、无鳞、无鳍、无爪、无足。

古代时期，植被茂密、丛林蔓生，毒蛇和猛兽频繁出没，古代祖先与之共存，并深受其影响，同时它们也提供了生存和生活的物资。正是因为对蛇的恐惧，古人将蛇变成了崇拜的对象，被视为土地的象征，得到供奉和崇拜。

《山海经》和《海外北经》记载：

钟山之神，名曰烛阴，视为昼，钻为夜，吹为冬，呼为夏，不饮不食，呼吸为风。身长

数千里，处于无冬之地。是物，人面蛇身，赤色，居于钟山之下。

在《山海经》和《北大荒经》中，也有记载：

在西北海之外，赤水之北，有张尾山。那里有神，人面而蛇身，赤色，目视为直，钻为昏，视为明，不饮不食不寝，风雨为食，为九烛阴，即烛龙。

文中所述的烛龙或烛阴是一位人面蛇身的神，神力极为广大，身体呈赤色，睁眼则为白昼，闭眼则为黑夜，不需食物和饮水，不需睡眠，吹出的是寒冷的冬气，呼出的是炎热的夏风，以风雨为食。

在这一时期，农业生产有了一定程度的发展，原始的生产活动形式多样化，捕鱼和狩猎生产形式开始出现。然而，这一时期的生产力水平仍然不高，生产活动始终受到自然的制约，原始祖先高度依赖于自然条件，如阳光和雨水；在他们眼中，水是不可预测的，原始祖先的命运和生存权被紧紧控制在其中。因此，水中的各种生物，包括蛇，都会让原始祖先感到敬畏，并认为这种生物也是神秘的，与神有关。将其作为图腾崇拜，必将带来风调雨顺，并起到驱邪避灾的作用。因此，红山玉龙被用作原始图腾崇拜对象。

龙象征着带来雨云。在《山海经》和《北大荒经》中，有这样的记载：

轩辕与蚩尤交战，黄帝在蓟州之野，应龙袭击蚩尤。应龙蓄水。蚩尤求风师雨师，使风雨得势。黄帝是天之女，雨停了，于是杀了尤。黎阳无法恢复，居所不雨。

《山海经·海内东经》记载：

雷神有雷令，身为龙体人首，腹部隆起，在无西。

作为农业经济的强大保障，水灾和干旱经常限制作物的收成。原始祖先改造自然的能力低下，因此他们只能通过祈雨作为确保农业收成的最有利活动，通常采取巫术形式进行。这也显示了用龙祈雨的真实性。

龙象征着权力和地位。古代书籍记载，古代帝王是龙的化身。伏羲是创造之神，是中华民族的人文始祖。他统一了中国的各个部落，成为中华民族的第一位王，并且在神话中，伏羲是一位人面蛇身的神。在《山海经·海内东经》中记载：

大人出自雷泽，华胥生伏羲。

上述记载是说，一个名叫华胥的女子去一个名为雷泽的地方游玩，看到了一个巨大的脚印，踏上去后怀孕，并生下了伏羲，伏羲有着人头蛇身的形态。此外，据说伏羲创造了中华民族的图腾龙，这只龙有着鳄鱼的头、鹿的角、须鲸的须、蟒蛇的身体、红鲤的鳞片、巨蜥的腿、雀鹰的爪和大白鲨的尾巴。从此，龙的后代就出现了。

龙象征着生命的繁衍。神话中的女娲是伏羲的妹妹，同为创造之神，但也是人面蛇身的

神，据传她在黄河边模仿自己的形象创造了人类，创建了人类社会和婚姻制度，炼就五彩石以补天，断龟足以支撑天柱，被誉为地之母。在《山海经·西山经》中记载：

有十位神，被称为女娲的肠子，他们变成了神，在栗广之野；横穿道路。这原本意味着有十位神和人，他们是变成神的女娲的肠子，在广阔田野中，像肠子一样，成为了道路的形状。传说中，神并不完全死去，他们在死后，身体的各个部分化身为世界上的万物，以另一种形式继续存在。这里以女娲为例，不仅女娲自己是生命的创造者，还将自己转化为其他生命体，从而永生。

### 3.1.3 赤山玉鸟

#### 1) 赤山玉鸟的图像特征

在他的文章《红山文化玉鸟的图像意义与艺术风格》中，李北雷提到，红山文化中发现了“祭祀葬坑”、“祭坛遗址”和“女神像”。祭坛中心“出土了玉鸟或鸟类的形象。根据墓葬规模和随葬品的推测，葬于祭祀葬坑中的人物具有极高的社会地位，拥有统治权和宗教地位，很可能是当时拥有政治和战争之神等各种权力的巫师。因此，对当时情况的分析表明，赤山玉鸟应该是当时原始信仰图腾元素的崇拜对象，在祭祀中被视为“玉器”。许多古文献记录了古代对鸟的崇拜。

以《山海经》为例，《山海经》和《中山经》记载：

其神是人面鸟身。以冀山为首，从玄山到蛮渠山，共有九座山，长 1670 里，其山之神是人面鸟身。

《山海经》和《海外北经》记载：

在渔江之北，有神人面鸟身，两条青蛇作耳，两条青蛇踏足。

在北方，有一位名为禺强的神，是风水之神，其形态是人面鸟身，耳朵上挂着两条青蛇，脚下踩着两条青蛇；

《山海经》和《海外东经》中记载：

东方句芒，鸟身人面，驾两条龙。

东方木神句芒，鸟身人面，座下是两条龙。

在红山文化遗址中，有许多玉制的鸟形器物，其中大部分是扁平的、呈飞翔状，鸟头的形状非常大且圆润，尾部大多是方形的，鸟的背部是直的且背部有孔，雕刻的线条极其简单，

大多是浅刻的。

## 2) 赤山玉鸟的功能象征

在原始社会的图腾崇拜中，鸟类被作为图腾崇拜的一大部分原因是它们具有飞翔的能力，因此赤山玉鸟也被用作“玉器”，以服务于神灵。原始的祖先对周围的自然环境充满敬畏，尤其是当他们看到超出自己认知能力的神奇现象时，相信这些现象具有特殊的神力，可以成为神灵的心灵和吉祥的象征。《山海经》和《南山经》记载：

有鸟，其状如鸡，五彩并文，名为凤皇，德之首文，翼之顺文，背之义文，胸之信文，腹之忠文，是为鸟也，食自然，歌而舞，见则天下太平。

在《山海经》和《西山经》中记载了一种奇特的鸟类，它的形状像鸡，身覆五彩斑斓的颜色，被称为凤皇。它头上的文叫做“德”，翅膀上的文叫做“顺”，背上的文叫做“义”，胸前的文叫做“仁”，腹部的文叫做“信”。

《山海经》和《西山经》中的记载还表明：

有鸟，其形状像鸟又有人面，蝎子身和狗尾，其名自称，见之则大旱。存在一种鸟，有着人的面孔和鸟的形态，身如猕猴，尾如狗，它的名字来源于它的叫声，出现它则预示着将有干旱。

在《山海经》和《西山经》中记载了一种奇特的鸟类，它的形状像鸡，身覆五彩斑斓的颜色，被称为凤皇。它头上的文叫做“德”，翅膀上的文叫做“顺”，背上的文叫做“义”，胸前的文叫做“仁”，腹部的文叫做“信”。

《山海经》和《西山经》中的记载还表明：

有鸟，其形状像鸟又有人面，蝎子身和狗尾，其名自称，见之则大旱。存在一种鸟，有着人的面孔和鸟的形态，身如猕猴，尾如狗，它的名字来源于它的叫声，出现它则预示着将有干旱。

《山海经》及《北山经》记载：

此处鸟类众多，形状如鼠而有鸟翼，其声如羊，可用作兵士。

这里的大多数鸟类是玉鸟，如蝙蝠，它们看起来像老鼠并且有翅膀，它们的叫声像羊，可以用来防御军事灾难。

在《中间·易经·燕山篇》中记载：

有一种鸟，形状像喜鹊，蓝色的身体，白色的嘴，白色的眼睛，白色的尾巴，名叫青庚，能够抵御疫情，它的叫声似乎在呼唤自己的名字。

鸟类拥有这种特殊的飞行能力，而在古代社会中，鸟被赋予了各种功能，比如传播和平、驱除妖魔、从危险中拯救人们、传递信息等。

在原始社会中，先祖们在举行祭天仪式时，常常使用某种媒介与“天帝”沟通，而鸟类因其自由飞翔的能力，也成为巫术活动中人与神之间重要的“连结”。因此，赤山玉鸟最重要的角色是在祭祀活动中充当与神沟通的媒介，更准确地说，赤山玉鸟是神的使者，用来传达神意和散播吉祥。

## 4. 良渚文化玉器

### 4.1 概述

良渚文化位于长江下游地区，距今约 5300 至 4200 年，主要分布在太湖流域，北至江苏常州地区，南至钱塘江地区。1959 年，以良渚遗址命名正式得名。

良渚文化的玉器在材料的选择上具有原生地特性，绝大多数为具有微纤维结构的透闪石。良渚文化玉器的颜色主要包括青色、浅黄色、橙色、深绿色、赭色、乳白色等。

良渚文化的墓葬中，随葬品以玉器为主体，种类繁多。出土的玉制工具主要包括：玉斧、玉铲、玉凿、玉纺轮等；装饰品类有：玉璜、玉镯、玉管、玉珠、玉坠、玉龟、玉鸟、玉蝉、玉鱼、玉柱状器物、玉锥形器物、玉牌饰、玉带钩等；礼仪类主要有：玉璧、玉琮、玉冠形器、玉叉形器、组合项链等。

在良渚文化的玉器中出现了大量具有原始宗教属性的玉器，诸如玉璧、玉琮等，已被学术界广泛认可。现在其功能通常被视为一种神学工具，用于祭祀天地，象征财富和地位。这是良渚玉文化的一个突出特点。

### 4.2 良渚玉琮

#### 4.2.1 良渚玉琮的形象特征

首先，在《周礼·考工记·玉人》中记载：“大琮十分有二寸出，射四寸，厚寸。”《白虎通·文质》说：“圆中齿身玄外曰琮。”当郑玄注解《周礼》时，他说：“琮，八方如地。”根据东汉许慎的《说文解字》，清代乾隆始于文字。

一般来说，良渚玉琮在形状上较大，结构严谨且平衡，出土数量最多。有不同大小的规格和形状，形态上可以分为“内圆外圆”的圆柱和“内圆外方”的方柱两类。至于为什么会有这样圆形与方形相结合的复合形态，主要原因是原始祖先们认为天圆地方。中国文明中根深蒂固的“天圆地方”观念促使玉琮的外观相对出现，并长久以来被用作祭祀天地四方的特殊祭器。

良渚玉琮几乎都有装饰纹样，其中最引人注目的是“神徽纹”或“神人兽面纹”。装饰主要位于玉琮的外壁，雕刻有上下层叠的神人及兽面复合装饰，装饰的主体是神像，其面部呈倒梯形，圆眼睛周围重圈，两侧眼角有小三角形，鼻子宽大，鼻翼以弧线勾勒，嘴巴宽大，内部雕有两排平整的牙齿。头顶戴有高耸宽大的冠冕。神人之下，雕刻有兽面，圆眼重圈，鼻子宽大，嘴巴宽大，两侧有尖牙。整体纹饰雕刻精良，手工精细，反映了高超的工艺水平。

对于这种精美的装饰纹样的特征，学术界有不同的看法，作者赞同黄剑剑先生和姚世琦先生提出的“神人驭神兽说”。他们认为，“神徽纹”应该划分为两部分，上半部分是戴羽冠的神人，下半部分是独立的神兽，神人与神兽既独立又相互联系，呈现出神人控制神兽的形象。

神人驾驭神兽的形象在古代书籍中有所记载，它在《山海经》和《海外南经》中被记录如下：

在南方有一位名为祝融的火神，他有兽的身体和人的面孔，驾驭着两条龙。

在《海经》和《海外西经》中，记载着西方有一位名为西方收获的金神，左耳挂着一条蛇，骑着两条龙。

在西方，有一位叫做收获的金神，左耳挂着蛇，驾着两条龙。

或许这种装饰纹样象征着神灵征服神兽的形象。它清晰地表现出神与人位于神兽之上，代表了征服者与被征服者之间的关系，而神兽可能是神与人的狩猎对象，二者之间是敌对关系。神灵狩猎神兽的形象也在古书中有所记录，在《山海经》和《东方大荒经》中有记载：

在东海有流斜山，流入海中长达七千里。山上有一种兽，形状像牛，身有斑纹而无角，单足，当它进出水时会引起风雨，其光辉如日月，声音如雷，名为槌。黄帝取得了它，取其皮为鼓，拔了雷兽的骨，声闻五百里，以震慑世界。

东海有一只凶猛的兽叫做槌牛，被黄帝所获，因此他用它的皮制成了鼓，然后用雷神的骨头制成的鼓槌击之，声音传五百里，震撼世界。

#### 4.2.2 良渚玉琮的功能象征

良渚玉琮的功能之一是作为祭祀之物。学术界对玉琮的功能有一个相对统一的理解，即它是一种“玉器”，是巫师用来祭祀神明、沟通天地的工具，而刻画在琮上的“神徽”或其他装饰纹样则是巫师在进行祭祀时必须附着和使用的力量。玉琮被认为是一种用于天地祭祀的礼器，这主要基于《周礼》的记载，其中提到“以黄琮奠地”，并且它体现了“天圆地方”学说中的“外方内圆”形象，受到了天圆地方观念的影响。

第二是葬礼功能。良渚人将玉琮置于墓葬之中，应是用来保护死者，并具有镇墓、避邪、驱逐恶灵的效果。通常而言，拥有玉琮的主人皆为氏族贵族，因此玉琮同时也是死者权力、地位和财富的象征。在《周礼》中记载：“秉璧以收尸。”意味着使用玉璧等玉器来作为收殓遗体的一部分仪式，从而反映了玉在古代葬俗中的重要地位。

## 第 3 节 艺术与机器的关系

### 1. 从“艺术机器人”开始

在 21 世纪，人类已经进入了一个以数字技术、人工智能技术和生物基因技术为特征的技术时代。技术与艺术的互动融合已成为一种不可抗拒的趋势。技术已成为当代艺术创作中

不可避免且重要的因素，它对艺术形式产生了深远的影响，并带来了一系列重大的变革。这种影响不仅涉及创作技巧、创作材料和媒介，也涉及艺术观念。可以说，这些变化具有划时代的意义，与艺术史上那些伟大的发展时期相媲美。

最近在法国的国家大皇宫画廊开幕的“艺术家与机器人”展览让我们一窥未来。该展览汇集了机器人技术、计算机生成艺术和算法艺术。与那些仅仅炫耀技术手段的展览不同，为大皇宫精选的科学技术艺术作品具有很高的艺术价值，代表了那些不仅使用最新技术进行创作，而且关注技术对人类社会影响的艺术家们积极讨论的作品。这些顶尖技术艺术家使用的高科技工具为我们提供了一个重新评估技术对艺术家创作影响的新机会。

显然，策展人已经仔细安排了参观展览的顺序。三个展览区中展品的布置呈现出以“人为本”的逻辑顺序，并完美诠释了展览标题中“机器人”的含义。展览的第一部分包含了几个人拟人化的机器人。它们有身体和四肢，用手臂进行绘画或雕刻。然而，这些可动的“四肢”仍然是僵硬和滑稽的，并不能与人类甚至动物运动的协调性相比。第二个展区的“机器人”不再有可见的拟人形态，它们存在于芯片中，但它们能够创造艺术作品的形状和动态，并具有与公众互动的能力。第三部分是整个展览中最令人不安的。这一部分让我们看到，人工智能拥有超越人类身体限制的惊人知识和能力。图 3-5：埃德蒙德·库查乌，米歇尔·布雷特，《蒲公英》，数字图像，2006 年。

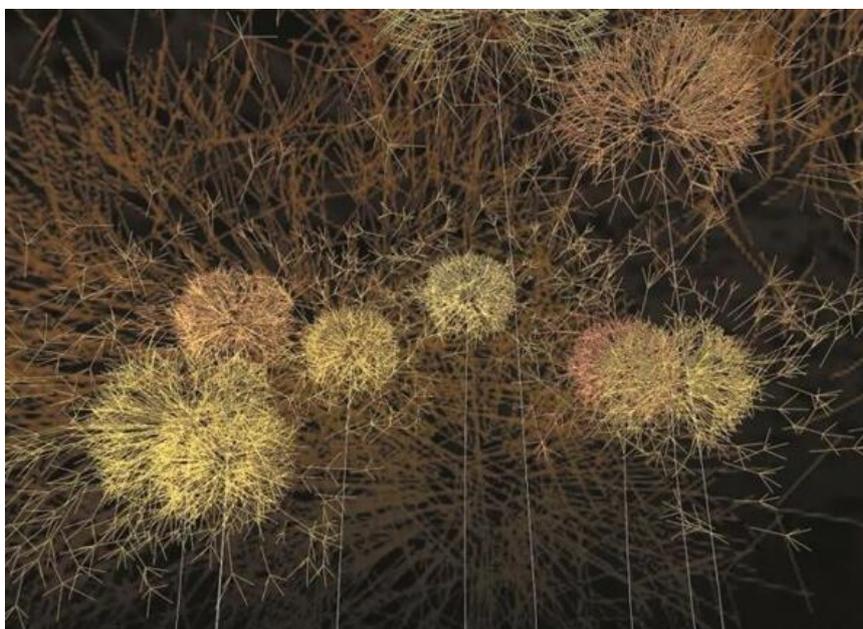


图 3-5：大皇宫国家画廊（图片来源：[https://www.sohu.com/a/234037527\\_534797](https://www.sohu.com/a/234037527_534797)）

艺术家们利用机器人来创作传统艺术所不能及的艺术品，这在人工智能出现以前的现代艺术史上，并非是新奇的事。早在 1956 年，匈牙利裔法国艺术家尼古拉斯·舍费尔（Nicolas

Schöffer) 即造出了能够创作艺术的“CYSP 1”型机器人, 而瑞士先锋艺术家让·丁格力 (Jean Tinguely) 同样在 1959 年就创作了《参与-自动机器》(M é ta-matic) , 让机器人在画布上作画, 并在同年的首届巴黎双年展上引起热议。2016 年, 村上隆的作品《机器人罗汉》也是人们非常容易联想到的同类作品。可以说, 艺术家们探索“机器人”的历史, 已经超过了半个世纪。图 3-6: 尼古拉斯·舍弗尔的 CYSP1 装置, 1956 年。



图 3-6: 马多克斯艺术

(图片来源: [collection.sina.com.cn/exhibit/zlxx/2018-07-13/doc-ihfhfwmu4271642.shtml](http://collection.sina.com.cn/exhibit/zlxx/2018-07-13/doc-ihfhfwmu4271642.shtml))

“艺术家与机器人”展览聚焦于日益先进的机器人技术给艺术创作领域带来的新体验。在一个更智能的社会中, 艺术家们探索了包括人工智能在内的新技术, 这无疑改变了艺术的创作、展示、传播和保存方式。40 位艺术家的作品创造了一个沉浸式的数字互动艺术体验, 这不仅扩展了我们身体的感官体验, 也颠覆了我们对空间和时间的概念。此外, 越来越复杂的软件和代码赋予了机器更多自发的创造, 包括无限的表现形式, 以及与观众的各种互动活动。

“艺术家与机器人”展览系统地解释了互动概念的诞生和发展。图 3-7: 机器人艺术, 莱昂内尔·穆拉, 2017 年。

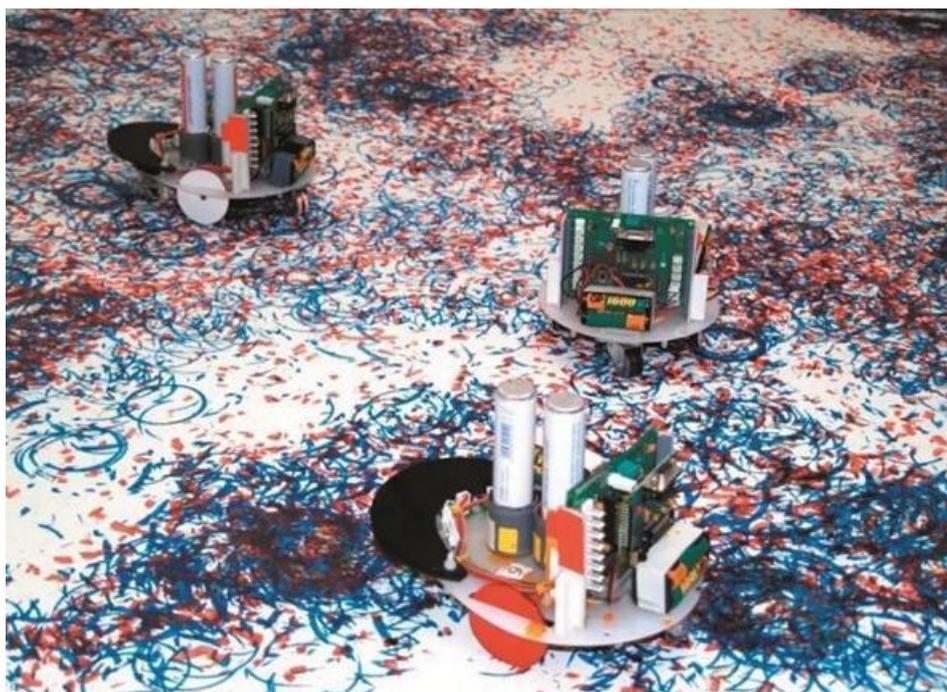


图 3-7: 艺术家 (图片来源:

[https://www.researchgate.net/figure/230807-ink-on-canvas-130-x-180-cm-C-Leonel-Moura\\_fig3\\_283620229](https://www.researchgate.net/figure/230807-ink-on-canvas-130-x-180-cm-C-Leonel-Moura_fig3_283620229))

可以看出, 21 世纪的艺术将会更加“技术化”, 正如今天的艺术家总是在使用新材料和新技术一样, 但是对于艺术家来说, 科技的发展提供的不仅仅是多样的材料和技术创作手段, 还有观念的变化。同时, 这些作品给观众带来的感受绝不仅限于艺术的形式或审美意义。它们可能会让观众担忧, 尽管人工智能可以帮助人类在许多领域取得前所未有的突破, 但它也可能超越人类, 反过来控制甚至奴役人类。事实上, 自 1951 年以来, 数学家图灵就怀疑了计算机独立思考的可能性。如今, 备受争议的超人类主义先驱雷·库兹韦尔甚至预测了在不久的将来出现绝对形式的人工智能。图 3-8: 彼得·科格勒, 《维度》, 计算机生成视频装置, 2018 年。



图 3-8: 计算机生成的视频装置 (图片来源:

<https://cloud.tencent.com/developer/news/304843>)

在技术主导的 21 世纪, 技术与艺术的互动融合是一种不可抗拒的趋势。技术对艺术形态产生了深远的影响, 涵盖了艺术生产系统的所有方面, 包括艺术生产的主体、生产方法、产品形态和评价体系。在讨论技术与艺术之间的关系时, 我们不能只看到技术发展带来的艺术媒介的突破, 也不能简单地将科学对艺术的影响视作艺术的推动。我们应该思考这些变化是否改变了我们对艺术本质的原始理解, 艺术在面对技术时如何保持其自身的身份和价值。思考艺术的身份和价值实际上隐含了时代的疑问, 即对技术可能带来的未来不确定性的担忧。这些疑问不仅是艺术界内部的讨论, 也触及到公众对于科技进步可能对人类文化、社会结构乃至个人生活产生冲击的深层担忧。艺术作为人类情感和思维的表达方式既是技术变革的产物, 同时也是对这些变革的反思和评论。当艺术家们利用数字媒体、人工智能、虚拟现实等工具创作时, 他们不仅仅是在使用新工具, 更是在探索和挑战艺术的边界。艺术品的价值可能不再单单依附于它的物理形态, 而是它的创作过程、观众的参与度以及艺术品与观众之间建立联系的能力。然而, 技术的介入也带来了关于作者权、艺术创作归属和原创性的问题。例如, 当一个由人工智能创作的艺术品出现时, 我们应如何评价其艺术性? 它的价值是否与人类艺

术家的作品等同？在艺术作品的创作和评价过程中，传统的美学标准是否仍然适用？总的来说，技术与艺术的结合提出了许多问题，也为探索这些问题提供了新的途径。艺术在这一过程中不仅保持了其独特的身份和价值，而且加深了我们对人类自身的理解——无论是作为创作者还是作为技术的使用者。对艺术身份和价值的思考，归根结底是对人性、社会和生存本质的更深层次反思。

## 2. 从邱志杰的《宇宙全图》开始

尤伦斯当代艺术中心将于2019年1月19日至5月5日举办邱志杰个展“邱志杰：笼罩全图”。这次展览是尤伦斯当代艺术中心建筑翻新完成后的首个展览。此次展览是对邱志杰自2010年以来创作的地图系列作品最全面、系统的整理和呈现。展览涵盖了艺术家在各种媒介上的地图作品，并重点展示了“世界地图项目”中最具代表性的部分，包括一套基于文化研究的24幅纸本水墨地图作品“万物系列”（2015-2017）；过去几年艺术家创作的3幅策展地图也将在展览中展出。展览还包括桌面地图装置作品“世界末日的思考”（2016），三个展示艺术家思考和创作过程的视频作品，以及艺术家与京东人工智能研究院共同开发的互动软件装置“京东AI世代地图”。此外，展览还包括艺术家为尤伦斯当代艺术中心翻新后的新空间创作的水墨壁画“艺术生态地图”，这是艺术家迄今为止创作的最大的场地特定地图作品，展览结束后将继续作为尤伦斯项目的一部分，并见证这个重要机构在当代艺术生态系统中的独特地位。

邱志杰具有多元的身份和创作媒介。从他早期跨越多个领域的艺术实践，到后来的策展和教育实践，他作为艺术家、理论家、教育者和策展人的一系列身份在他的地图作品中得到体现，为最终在形式创作中的融合奠定了基础。展览标题“世界地图”，源自中世纪拉丁语对世界地图的称呼：大约2010年，邱志杰开始绘制最初用于描述作品之间关系的“地图”，逐渐发展成为研究、写作、幻想和行动相结合的“世界地图项目”，并在随后的一百多幅地图作品系列中，由山水笔墨所构建的坐标系统凝练地将概念、个体、对象、事物和事件编织在一起，为观众提供了相互理解它们的可能性。作为借鉴他人之长的制图者，地图在过去十年邱志杰的艺术实践中承载了多重功能：它不仅是自我意识、作品框架、展览计划、思想交流、政治摄影，也承担了历史和研究概要的功能。如图3-9所示：



图 3-9: (图片来源: <https://www.sohu.com/picture/289764628>)

艺术家在 2016 年新加坡双年展上展出的《地图之图》聚焦于艺术家对制图史的研究。作品从中世纪开始,标记着重要地理事件如大地理发现和郑和下西洋等人类探索的历史,以及不同国家地图绘制方法的演变历程,在这件作品中得到呼应,向观众揭示了制图学和全球政治的发展之间的密切联系——正如艺术家所言,制图者握有以其选择的方式呈现世界的权力。在“万物系列”(2015-2017)中,这是“世界地图项目”系列中的代表作,邱志杰展示了一位自然主义者的兴趣和视野,从身体、命运、动物、植物、疾病、医学、情感、记忆、故事、神话、工具、职业等主题出发,对社会学、政治学、历史学、人类学等多学科的许多事物进行了发散性的文化排序和研究。系列中的 24 张地图可以被视为强调自我建构性的独立作品,彼此之间又相互呼应;它们的构图从左至右端对端相连,形成了一张将客观事实与个人经验相融合的知识谱系。《记忆之图》(2017)系列关注记忆生成的机制以及这一概念的隐藏政治:记忆作为一种情感产品既高度个人化,又可被以各种方式使用,修改和重塑的特性与国家、政治、历史等命题紧密相关。在地图中,每一个具体概念的内涵和外延,各种概念之间合逻辑或非逻辑、显性或隐性的联系,在艺术家由山川组成的水墨拓扑中逐渐揭示出来。图 3-10:



图 3-10: (图片来源: [https://www.sohu.com/a/289764628\\_626855](https://www.sohu.com/a/289764628_626855))

在“世界尽头的思想”桌面地图装置中，“群岛”系列的拓片运用了众多的“主义”作为线索，非线性地分类了人类思想历史上的不同流派和重要议题。在巨幅壁画“艺术生态地图”中，邱志杰勾勒了自1989年以来逐步形成的艺术产业生态系统。在这幅地图中，艺术教育体系、官方艺术体系和私立机构群被划分为三个区域。地理元素的命名和出人意料的空间关系显示了艺术家独特的诗意和幽默感。在中国当代艺术发展史中，这些轶事颇具讽刺意味。如图3-11所示：



图 3-11: (图片来源:

[m.cfbond.com/zclb/detail/20190119/1000200000019271547892535868898294\\_1.html](http://m.cfbond.com/zclb/detail/20190119/1000200000019271547892535868898294_1.html))

由于地图具有思维导图的特质，邱志杰经常使用它们作为策划各种展览的蓝图，这些展示策展观念的地图也包含了艺术家对其所处艺术生态系统的反思。《为什么表演》是艺术家为同名 2016 年展览创作的策展地图。在梳理表演艺术的历史、类型和理论的基础上，它审视了当下时代“表演”的政治和社会本质。《无尽之图》（2018）最初是为第 67 届威尼斯双年展的中国馆制作的；这一次，艺术家将原始的草图地图扩展为水墨地图，在展览中首次呈现。艺术家将地图主体的轮廓转化为中国的太极和两仪图。象征着中国传统中“无常”的寓言和古人为应对“无常”而发展出的一系列生命哲学，被放置在地图左半部的“变易元”中。在右半部的“传承元”极中，代表继承、延续和永存的概念和形象出现。这与邱志杰对展览主题的思考相一致：正是在不断的变化和传承中，文化和艺术才能持续繁荣发展。在艺术家为展览“后 1989 艺术与中国：世界剧场”（2017-2018）创作的地图中，代表中国政治思想史和过去 40 年全球化进程的两条山脉构成了社会政治背景的框架。它回应了中国当代艺术的曲折历史河流；许多接近现实且相当自传性的细节，被邱志杰 30 年的教学、研究以及对艺术史的个人经验和见解高度凝练。如图 3-12 所示：



图 3-12: (图片来源: <https://www.cafa.com.cn/cn/news/details/2110675>)

展览还包括视频和互动软件装置,展示了邱志杰在地图创作中的思维过程。“墙上的项目”将邱志杰之前的壁画作品结合成视频,“思维导图动画”记录了思维导图软件组织和生成信息的过程,“京东人工智能生成地图(现场表演版视频)”展示了艺术家如何使用人工智能软件进行创作。此外,在与京东人工智能研究院的何晓东博士合作开发的人工智能软件装置中,艺术家继续拓展自己地图创作的界限。邱志杰地图系列中的高频词汇和包含数万常用词的京东词库被导入该软件装置。观众通过语音或打字输入任何有效词汇后,软件将自动关联生成邱志杰地图景观和图形元素的组合,以艺术家创作风格的地图。作为地图系列的技术延伸,AI地图反映了艺术家长期以来跨媒介创作的倾向。通过文本语义层面的无限扩展,地图已进化为一个具有无限可能性的自动增长系统。如图 3-13 所示:



图 3-13: (图片来源: <https://www.cafa.com.cn/cn/news/details/2110675>)

## 参考文献

- [1] 刘志毅, 冯建明. 艺术与技术融合下的新媒体设计. 山西广播电视大学学报, 2019(4):4.
- [2] 封雪. 艺术与科技结合: 借助多媒体营造艺术教学研究氛围. 软件: 电子版, 2018.
- [3] 王超奇. 基于杭州地区特色的公共艺术水景创作与图像技术应用研究. 浙江工业大学, 2018.
- [4] 刘伟. 现代艺术设计与民间艺术的结合. 艺术与技术, 2019, 32(02):182.
- [5] 叶笛. 苏州科技大学媒体与视觉艺术学院精选画作. 编辑之友, 2018.
- [6] 叶天赏. 运用“适应性”与“跨界”教学培养学生解决问题的素养: 以“技术”与“艺术”为例. 台湾教育, 2018(709):93-101.
- [7] 张千. 无常与永恒: 艺术与技术关系的反思. 艺术观察, 2022(4):3.
- [8] 张婷. 庆祝绿色冬奥会——访清华大学艺术与技术创新基地副主任余仓碧. 国家电网, 2021.
- [9] 江山, 张丽洁. 新文科建设背景下艺术与科技的跨界融合: 基于全国声音与音乐技术大会的思考. 人民音乐, 2021.
- [10] 魏继贇. 城市住宅室内空间老化设计中的艺术与技术. 2021.

# 第4章 计算机艺术：可视化与生成

## 第1节 计算机环境中声音合成技术的艺术应用

### 1. 声音艺术表达的革命性扩展

从声音合成技术的历史发展角度来看,对艺术类别的研究势必会集中于最初对人类生理结构进行粗糙借鉴和模仿的声音合成技术,声音艺术表达形式的长期探索,以及市场的及时选择。这一总体思想直接或间接地展示了人声合成技术已成为电子音乐中一个非常热门的历史现象。

#### 1.1 声音合成通过技术扩展了人声

声音合成技术是一项非常早期的发明,与我们在音乐制作中常见的大多数合成器相比,我们会惊讶地发现,人声合成技术最初并非用于制作音乐,而是用于通信研究,作为减少电话信号传输过程中带宽的一种方式。随后,偶然间它逐渐被应用到音乐领域,并为公众所熟知。声音合成技术主要分为两类:第一类是语音合成技术,主要用于计算机和通信领域。另一类用于创作和处理制作中的声乐和音乐方面,称为声乐合成技术。

在声音合成技术中,“技术”首先属于技术类别,指的是处理原材料或半成品的技术和方法,而劳动生产活动通常指的是在生产实践过程中形成的知识技能和操作方法。在文学和艺术领域,它是从某种特定技术视角对作品在特定音乐环境中的解读。然而,在技术的深层结构中,声音合成技术包含了两种不同形式的表达:一种是纯粹的技术性创造性思维,另一种是具有审美体验的艺术创新[1]。一般来说,如果我们只关注技术评估方面的创造,那么就用某一领域的技术探索取代了整体艺术探索,并形成了使用特定技术标准来衡量声音合成作品质量的做法。在艺术创作过程中,如果技术手段和审美体验巧妙地链接起来,通过特定的声音形式感知和表达主体的审美品味,并期望缩小观众的“心理距离”作为艺术创作的概念,那么它就是以审美为主体的艺术创作。声音合成的“技术”层面不仅指导了技术领域中的具体追求,而且在这一特定语境中促进了声音合成在音乐中的持续形成和发展。

推动声音合成技术在音乐和电影中持久应用的一个重要原因是,合成的声音为艺术家提

供了更多的艺术表达空间，且合成的声音可以更好地与自然人声融合。当声音合成技术首次应用于流行音乐时，就像任何新事物一样，被“处理”的电子人声引起了相当大的争议。人们常常认为，天然的人类声音已经被赋予了无限的表达力，为什么还要使用合成的人声，以及这种电子人声到底属于什么类型的电子声音。对新事物的这种困惑主要表现在对合成技术的不理解和排斥恐惧上。因为大多数人对人声的音色和音域都非常熟悉，声乐一直是艺术家们的一种受欢迎的表演艺术。相反，声音合成技术在流行音乐中的应用，由于其更广泛的声音范围，为许多习惯于传统人声的听众和声乐歌手带来了很大的不适应。因此，他们认为通过合成的人声来表达情感和音乐对象往往无法触动人心，所以在概念上，未来派前卫作曲家希望人们从另一种思维方式来审视声音合成的电子音乐，这种令人耳目一新的概念变革将逐渐适应观众和传统艺术家。

在音乐中使用声音合成技术并没有所谓的“正确方式”，但在手段方面仍然存在不同的分类方法。每个创作者在自己的知识体系的影响下都有自己的偏好，虽然没有固定的界限划分，但可以简单分为以下几个部分：

### **1.1.1 模仿人声**

当人声合成技术最初被发明时，其目的是模仿人类的语音，使用简单而粗糙的设备来创建一个人工喉部，尝试模仿人声发出的声音。这种声音试图以更令人满意的方式模拟人声的各种发声特性，而且合成技术可以更方便地控制或突出某种音质，后来这种模仿的声音合成技术广泛应用于卡通和科幻电影中。在电影制作中，模仿出的声音往往会听起来像是“真实”的声音，而附加的声效则突出了要表现的声音的特点，使声音更符合角色特性。例如，在电影《小黄人》中小黄人的声音制作，就是通过改变速度和调整音调的方法，改变正常的配音，从而创造出小黄人独特的语言系统。

### **1.1.2 氛围暗示**

这种合成不需要模仿真实的人类声音，它只需要人声的特征。就像由 MIDI 产生的弦乐组感觉一样，它给予不太挑剔的观众强烈的氛围感：这是由许多小提琴演奏的“弦乐组”。在大型晚宴的电影场景中，如果在声音编辑中叠加多个人声轨道，语言之间会有相位偏移，喧闹的人声将自动产生一种环境声音，使得任何语言的信息都无法听清，从而创造出人声的

现场氛围。

### 1.1.3 合成音色

电子音乐旨在通过声音混合体与真实人声的结合，将表演者从传统的限制中解放出来，因此这种风格的音乐的声学特性是前所未有的。例如，合成的声音可以被视为一种带有“未加工”人声的乐器，在捕捉到原始人声的声音特征的基础上，在合成器上创造出复杂独特的音色，突破了传统歌唱中人声音域和音符变化范围的限制，满足创作者的个性化创作需求。这种合成声音在现代流行音乐和科幻电影中广泛使用。

合成人声是一个创造性的过程，简单地认为只要将声音输入机器，就能自然生成，这是一个错误的假设。声音合成技术涉及声学、语音学、模拟技术和数字技术等，掌握和理解技术和理论知识将更好地激发我们在声音合成中的创作。人声的原材料可以是单一的或者是杂乱的片段，创建人声的方法有很多种。近来，声音合成不仅仅是计算机生成的“超现实”声音，尽管这通常是最为人所熟知的角色之一，但对于声音设计师和音乐制作人来说，如果人声经过精心设计，可以将单调直接的声音色彩绘制成丰富多彩的人声画面，将创作者大脑中的灵感变为现实。高品质的人声采样和声音合成技术无限放大了人声的可能性，并将在未来的发展中被更多的非技术爱好者所使用。声音合成技术通常只是一种可以产生不同形式人声的技术，是创作者使用的工具，但只有在艺术家和创作者的创作下，我们才能避免让这项技术变得平庸，优化人们的音乐审美体验，并为人们的视听活动带来不同层次的感官体验。

## 1.2 简要历史回顾：人声合成技术

人声合成技术几个世纪以来一直是人类的梦想，要理解它们是如何发展成现在的形式，必须对历史进行一定的回顾。在本章中，从机械时代的语音合成到计算机时代的电子音乐，简要概述了人声合成发生、发展和形成过程中的重要里程碑和技术。

### 1.2.1 起源：19 世纪末至 20 世纪初

#### (1) 话筒 (Talkbox)

话筒的起源通常被认为是在 20 世纪早期，最初是与吉他一起使用的，称为“哇效果”(Wow Effects)。话筒的原理是使用具有带通滤波器特性的踏板，通过连续上下移动来改变

吉他声音信号的频率，交替增减高低音调，类似于 EQ 效果上的扫频，从而产生所谓的“哇”声音。使用它，音乐家能够产生类似于元音、辅音或简单短语的词语，但这还不是传统意义上的声码器（一种将原始语音信号与处理过的电子语音信号混合的乐器），它本质上是一个效果器，允许音乐家通过改变嘴型来修改乐器的声音，并且音乐家通过改变口型来控制 and 调整声音，将用于塑造声音的频率和声音信号加载到乐器中，通常是吉他或键盘。

话筒通常有一个效果踏板，放置在地板上，通过一个塑料气管与扬声器相连，这个连接器连接了放大器扬声器输出和普通乐器输出，踏板上的脚踏开关指示声音是面向话筒还是普通扬声器，管子的另一端放置在麦克风旁边，伸到表演者的嘴边或足够近以便能够复制声音。

当话筒被激活时，放大器中的声音通过话筒的扬声器和塞进表演者口中的管子来复制。表演者可以改变嘴巴的形状和舌头的位置，产生声音并通过话筒的扬声器重放，就像效果本身在说话一样。嘴巴的形状改变了声音的空间场，但它也影响了声带发声时产生的谐波，这种“形状”声音从表演者的口中输出并进入麦克风，形成一种乐器/语音混合物。如果我们正常发声，音高是由声带控制的，但通过话筒，音高变化可以是任意乐器，话筒取代了发声道，该乐器大多与吉他或键盘有关。电吉他中丰富的谐波是通过唇动声音产生的，有效地允许吉他“说话”。

音频使用数字音频分析工具 Logic 的内置效果 Channel EQ 来数字化分析话筒的声音效果变化。在图中可以清楚地看到，处理过的吉他声音的水平比吉他原声轨道大 5-10dB，这表明话筒的主动扬声器会放大其合成音频和原始音频的总水平。锯齿形状的频率变化表明存在以噪声为导向的音色，该音色的频率集中在人声上。

话筒是一种经典的声音效果器，以其独有的声音表现力和辨识度而著称，它在 60 年代和 70 年代的摇滚和朋克音乐中随处可见。例如，Pitt Drake 的专辑《Forever》（1964 年）中就有经过话筒处理的歌曲，同样地，Smith Spaceship 在 1975 年发布的专辑《阁楼中的欢乐》中的歌曲《Sweet Emotion》，话筒在歌曲开始时创造了和声效果，以及 Eagles 乐队在 1979 年专辑《The Long Run》中的歌曲《Those Shoes》也使用了话筒效果。

## (2) 声箱器 (Sonovox)

声箱器是由工程师吉尔伯特·怀特 (Gilbert White) 在 1939 年 1 月发明的一种设备，声箱器的发现非常非凡。有一天，怀特在刮胡子时，当他在喉咙连接处刮胡茬时，他注意到可以随着剃须刀片在脖子周围来回移动而产生一种“嗡嗡”的声音。出于好奇，怀特尝试用嘴巴和舌头的运动来形成单词，并通过这种奇怪的替代方式，他能够发出奇特的声音。

声箱器的原理和喉部发声的原理是相同的，与插入嘴中的连接管不同，它使用一个小型扬声器附着在表演者的喉部。根据仿生学的原理，表演者通过模拟说话，借助舌头、嘴唇和其他语言器官的活动以及改变嘴巴的形状来形成语言，声箱器取代了喉部声带振动，并通过小扬声器放大产生的基础声音来产生一个可识别的声音。最初，声箱器被用作医疗设备，用于恢复经过喉切除手术的患者们的语言能力，作为一种人工喉部的设备。然而，最初的声箱器在发声时音调变化很小，产生的声音类似于机器人的讲话。如今，人工喉部已经可以做得足够小，通过牙科手术的帮助，可以根据患者的口语动作来改变音调，以产生更自然的声音。

声箱器曾被用作电影的噱头，成为电子音乐效果的一部分。声箱器首次出现在 1940 年的电影《You'll Find Out》中，当时主演凯·凯瑟（Kay Kyser）和他的乐队在表演中，而其主唱卢高西（Lugosi）使用声箱器在一次降神会中模拟了一个死者的声音。声箱器后来还被用于电影《小飞象》（Dumbo, 1941 年）、《不情愿的龙的声音》（The Voice of the Reluctant Dragon, 1941 年）和《三个妻子的故事》（A Letter to Three Wives, 1949 年）。

### (3) 语音合成机

最初的语音合成技术是科学家沃尔夫冈·冯·肯佩伦（Wolfgang von Kempelen）的“声学机械语音机”，这是一种能够产生单一声音和一些声音组合的声音设备，需要手动操作声音。

该机器的主要部分是一个压力室，它被构造成一个粗糙的模拟人类发声系统。风箱被比喻为人的肺，振动的簧片作为声带，皮革管道作为声道，通过改变皮革管道的形状，可以产生不同的元音。辅音则使用手指来模拟和控制四个独立通道的收缩，而对于爆破音等声音，根据声道的模式模拟舌头和嘴唇的使用。他的研究表明，人类的发声道，即声带和嘴唇之间的腔体是声学的主要部分，并且喉部是语言产生的中心。但这台机器也有负面影响，因此它只能模仿类似人类的声音，这没有特别的意义。

在随后的年份中，科学家惠斯通（Wheatstone）基于冯·肯佩伦的说话机进行了进一步的改进，并能够产生元音和大多数辅音。这些声音的组合，甚至产生完整单词的可能性，在一定程度上是声音合成技术实现的最简单的声音合成。惠斯通的改进增强了机器的灵活性和发声的自然度，使得机器能够更接近于真实的人声。这些早期的发声设备为后来的电子语音合成和计算机语音合成技术奠定了基础。

## 1.2.2 发展与成熟：1940 年代-1980 年代

### (1) 声码器/语音合成器

首个完整的电子合成器是由斯图尔特 (Stewart) 制作的，这是一个带有蜂鸣器作为激励源和两个共振电路来模拟声道的声音共振的设备。这个合成器能够产生带有两个最低共振峰的单个元音，但不能产生辅音和辅音发声。然而，第一个被认为是语音合成器的是霍默·达德利 (Homer Dudley) 在 1939 年纽约世界博览会上展示的 VODER，它是后来在 20 世纪 30 年代中期由贝尔实验室开发出的声码器 (语音编码器) 的前身，也是人类首次用电子手段合成成人声。

原始的 VODER 是一种用于将分析后的语音转换为缓慢变化的声学参数，然后在驱动合成器中重建近似原始语音的声音信号的设备。VODER 使用手动操纵杆选择声音和噪声源，脚踏板控制基频，源信号是十个频段的带通滤波器，其输出级别用手指控制，使其具有更高的可理解性和可读性。

该设备的 10 个按钮是一个共振控制箱，用于控制带通滤波器的增益，并大致确定声音信号的频谱内容。所有的滤波器都由噪声源或蜂鸣振荡器激发，手腕杆用于切换激发源。周期性激发用于产生单调的音调，白噪声激发用于产生音高，脚踏控制周期性激发产生音调，最后将声音合成并输出，放大后发送到扬声器。

Voder 最初是作为一种机器设计的，用于研究人声通过铜线电话的传输，以及二战期间的加密语音传输。然而，在访问了达德利大学之后，波恩大学语音学系主任维尔纳·迈耶尔·埃普勒 (Werner Meyer-Eppler) 于 1948 年认识到了 Voder 和电子音乐之间的联系，并发表了一篇关于使用电子手段产生音乐概念、电子合成语言的论文，并在德国成立了一个电子音乐工作室，成为了德国“电子音乐 (Elektronische Musik)”的奠基人。

### (2) 慕格电子音响合成器

1971 年，声音合成器 (vocoder，即“声码器”) 开始进入流行音乐领域。作曲家温迪·卡洛斯 (Wendy Carlos) 和罗伯特·摩格 (Robert Moog) 共同构建了第一个专门为音乐制作设计的音乐合成器，那就是由键盘、电子电路、振荡器、滤波器等组成的 Moog 合成器系统。这个设备被命名为“频谱编解码器”，受到 Voder 启发，在音乐合成器中加入了音乐编码器。在电影《发条橙》(A Clockwork Orange) 的配乐中，大量使用了声音合成技术，而且声音

合成器的“歌唱”被用来创造一系列前卫的电子音乐作品。在经典恐怖电影《闪灵》（The Shining）的音乐中，使用托马斯的诗《怒火的日子》（Days of Rage）作为歌词结构，通过合成器和人声的重组，营造出了宗教、讽刺和死亡的特定意象。

德国伟大的电子音乐乐队 Kraftwerk 之所以成为不会被遗忘的电子先锋，是因为他们在《Autobahn》（高速公路）专辑中风格的突变。正是在那时，他们得到了一台小型的 Moog 合成器，从早期充满沉重而漫长的电子段落和生活噪音的录音，转变到首次在音乐中加入歌词，首次全面使用电子合成器，这对整个流行音乐界也具有深远的意义。合成器丰富的和声效果、循环的操控旋律和清脆的鼓声，将定义许多未来电子音乐大师的未来方向。Kraftwerk 乐队随后将电子音乐与机器人结合，在高度工业化的风格中创造了前所未有的视听效果。他们凭借音乐创新性和机器人的音乐表现力成为了这个乐队的标志。代表作《The Model》（模特儿）、《We Are the Robots》（我们是机器人）和《Man Machine》（人机）都是通过 Moog 合成器的效果处理制作而成的。

### 1.2.3 20 世纪末期，进入了电脑音乐时代

当计算机不再是只有大型机构和大学才能使用的复杂处理设备，而成为知识经济时代重要的工具和载体时，现代计算机时代就开始了。与大多数机械时代设备相比，机械设备通常只能完成特定的任务，其变化性非常有限，而计算机是通用设备，可以根据不同的程序完成任何功能。今天，无论是家用、商用还是工业用，计算机已经在许多领域得到了广泛应用。

确实，计算机在音乐制作上的潜力超出了传统合成器。1957 年，贝尔实验室 (Bell Labs) 成为第一个使用计算机生成声音制作音乐的机构。通过计算机生成的声音，作曲家们可以进行特殊处理来完成音乐制作，这一过程不受传统乐器和演奏者的限制，开启了声学效果的新可能性。

法国作曲家蒂萨·索纳米的 1992 年电子音乐作品《发生了什么》也使用了声音处理方法，这些方法基于作家萨默斯的文本，分解人声来讲述一个恐怖故事。在音乐作品中，故事是通过不清晰的声音来叙述的，而且在无线电波或其他影响信号的作用下，声音的音调会突然增加，使得声音信号变得分离和断裂，整个作品充满了超现实主义。

#### 1.2.4 中国语音合成技术的发展现状

随着中国信息技术的不断进步，人机交互变得越来越频繁，这也使得中文语音合成技术成为一个热门的研究方向。随着 13 亿中国人广泛使用中文进行交流，中文语音合成技术的应用前景非常广阔，它带来的市场实用性和经济效益将是不可估量的。显然，中国日益实用的人声合成技术已经使得中文合成技术从学术研究转变为最有前景的技术之一。

在 20 世纪中叶，英语语音的文本到语音 (TTS) 技术首次被发明。在 1980 年代，中国也相继开始对汉语语音合成进行研究，虽然起步相对谦卑，但到目前为止，基本上已经达到了与国际研究水平同步的地位。汉语的合成技术经历了最初的共振峰合成和线性预测编码 (LPC) 合成，现在采用的是国际公认的 PSOLA (音高同步叠加线性预测) 算法来合成汉语。

在国家 863 计划、国家自然科学基金等相关项目和研究项目的支持下，中国的语音合成技术取得了显著进展。例如，中国科学院声学研究所的 KX-PSOLA 系统、联想集团 1993 年推出的“联想佳音”、中国科学技术大学在 1995 年推出的 KDTALK 系统等。此外，中国科学技术大学后来推出的 KD-2000 中文转换技术，虽然语音的自然度需要提高，特殊语言的发音存在“机器人调”的问题，但在提高语音可懂度和拼接技术的处理上，中文合成的完整性和清晰度已得到了极大的改进。

中国的语调具有声调性，一个汉字的发音有一个音节，这个音节由元音、辅音和声调三大要素构成。在汉语的发音过程中，汉字会受到相邻的字或词的影响，声调会发生变化，句子中包含丰富的韵律特征，导致需要输出的数据量巨大。为了使中文语音合成技术不仅发音清晰，而且听起来更自然，需要语音学、声学、计算机语音研究等多领域的专家不懈努力，以推动中文合成技术的发展和持续改进。

## 2. 声音艺术表达的技术扩展

尽管任何类型的语音合成技术都是历史发展的必然产物，但就声码器而言，不同阶段的声码器也具有不同的时代特征。当音乐声码器作为电子音乐的一个流派产品产生并流行起来时，每个阶段都呈现出多种复杂的表现特征。在这一章中，作者总结了音乐声码器的演变过程，以归纳音乐声码器创作的风格。

### 2.1 人声重构：声码器

声码器 (Vocoder) 的发展始于 20 世纪 30 年代中期，由霍默·达德利 (Homer Dudley)

发明，它是一种用于分析和合成实验性人工语言的研究设备，目的是为了监测语音的结构以提高电话线路的语音传输能力。

原始的语音信息通常在 300~3400Hz 之间，因此通信系统中电话的最初设计带宽大约是 3000Hz。然而，当我们分析语音的均衡器 (EQ) 时，我们发现提高语音易懂度和突出说话者个人特征的频率往往超出了这 3000Hz 的范围。这就是为什么当我们在电话中交谈时，感觉对方的声音与现实中的声音相比有较大差异，甚至有时候难以听清“s”和“f”音的原因。

声码器的结构主要由两部分构成：分析部分和合成部分。它的原理是先将声音信号分隔成不同的频率带，然后对每个频率带的信息进行分析。分析器根据声音样本的频谱特征，通过一系列窄带滤波器进行测量；合成部分则会分别处理每个频带，并最终生成一个动态的声音信号。

### 2.1.1 声码器结构

在 20 世纪 50 年代，声码器开始被艺术家用来处理音乐中的人声信息。声码器的基本结构是模拟人类声音在声道中振动的物理特性。人类的声道由鼻腔、口腔、喉部组成的一个充满腔室的装置，声带在腔室中的振动产生语音，声音具有多种谐波成分，通过声道的过滤方式来进行处理，某些频率会被增强，某些频率会被减弱，最终产生复杂的音色。

声码器有三个主要组成部分：带通滤波器、包络跟踪器和放大器。

#### (1) 带通滤波器

滤波器可以调整语音信号的谐波含量，带通滤波器的作用是将声音信号分割成不同的频率带，每一个频率带负责分析对应的声音频率。

带通滤波器是一种过滤器，它允许选定范围内的频率通过，并阻止该范围之外的频率。例如，如果带通滤波器控制 600 到 1500Hz 之间的范围，那么 200Hz 和 4000Hz 的频率就不会通过。

声码器通常有两组带通滤波器，根据分组的分类，将声音信号分配到不同的频率范围，每个频带都有指定的中心频率和带宽，采用减法合成的方式，将每个频带限制在一个非常狭窄的频率范围内，然后发送到它自己的包络跟踪器。同时，载体加载的声音信号也通过另一组带通滤波器按照与语音信息一致的频带分类分割成不同频率。

## (2) 包络跟踪器

包络跟踪器是一种感测声音信号随时间变化的响度变化形状的装置。包络跟踪器决定了分割的频率能多快地影响处理后的声音信号。如果速度快且间隔短，则每个频带的振幅将容易被检测到；然而，如果包络跟踪器的响应时间长，就容易造成频带的振幅控制不易出现间隙。

包络跟踪器响应声音信号的响度，包络描述了响度的变化，我们可以看到，语音信号的包络曲线是一条随时间变化而信号音量升高和降低的线条。

## (3) 放大器

放大器，顾名思义，是一个电路元件，用于放大声音信号并控制放大量。在模拟合成器时代，任何声音信号的输出音量都需要通过 VCA (Voltage Controlled Amplifier, 电压控制放大器) 来控制，这是控制放大器音量的增益。在数字时代，声音信号是通过数字设备控制的，但仍然习惯称之为 VCA。此时，连接到包络跟踪器的语音信号被发送到放大器，将声波形和载波信号的合成输出加以控制，打上最终声音信号的烙印。

## 2.1.2 不同时期的声码器

### (1) 慕格模块化合成器

慕格模块化合成器实际上是由一系列模块组合而成的，每个模块执行特定的功能，如音高调制或控制各种滤波器以校正声音。这些模块的组合为用户提供了自动化操作来修改声波形和减法合成，简化了漫长而乏味的声音制作过程，摩格模块化合成器的出现彻底改变了音乐界。温迪·卡洛斯 (Wendy Carlos) 在一次采访中曾说：“上世纪 60 年代典型的电子音乐设置是由一群身穿白大褂的技术人员在一个杂乱无章的实验室工作台上完成的，而音乐的最重要组成部分，旋律、节奏、管弦乐编排、和声与复调，都很难实现，而且大多数技术人员并非严格训练有素的音乐家。所以，当我们听到这些‘土气的’声音效果时，我们如何让公众看待电子音乐？但是罗伯特·摩格的合成器改变了这一切，它将大量的设备与一个触键键盘结合起来。随后，温迪·卡洛斯和罗伯特·摩格进一步完善了合成器，将声码器添加到合成器中，这样声音或电子属性的作曲家或钢琴家就能创造出真正的电子音乐。

## (2) 软件声码器模型

在计算机上实现软件声码器，为音乐家提供了更深层次的可编程性。与模拟时代的大多数机械设备相比，计算机的通用软件具有独特的优势：机械设备在操作过程中只能完成一定范围的可变功能，并具有一定的局限性。使用计算机上的操作软件和插件，将能够根据开发者编写的程序执行任何功能。最近，随着个人计算机时代的到来，计算机不再仅由大公司和大学使用，标准化的计算机概念被整合到计算机音乐的各个分支中，计算机的处理能力越来越强。随着软件功能的不断更新，全数字的声码器软件应运而生，并变得更加灵活。软件声码器的主要流派包括 Native Instruments 的 Vokator、Prosoniq 的 Orange Vocoder、Apple 的插件 EVOC 20 PolySynth 等。

### 分析参数

与大多数声码器相似，语音编码始于声音信号的分析。EVOC 20 PolySynth 的分析和合成部分包含一个滤波器组，其带宽与合成组的带宽相匹配，并生成其均衡器曲线，该曲线可以从 75 到 8000Hz 调整。滤波器的最低和最高带宽是可配置的，带内的频率将被保留，而带外的频率将被切除。

此外，人类的声音通过声带的振动区分清音和浊音，而声码器合成的声音信号必须具有清晰和浊音。因为合成引擎无法单独识别清音和浊音信号，EVOC 20 PolySynth 包括一个 U/V 检测。在检测和分析信号的单元中，检测诸如“S”、“K”、“P”等清音的声音信号，将声音信号的相关部分替换为“噪声”信号，如果遇到“a”、“oh”、“e”等元音，则不会影响声音信号的变化，直接作为声音信号输出。人类的发音往往是有声和清音的结合，像“vocation”和“base”这样的英语发音就是浊音和“噪声”的混合。由于被噪声信号替换的静音信号带宽较宽，分析滤波器组可以检查多个频带是否产生频谱变化。每个滤波器都有自己的包络跟踪器，随着输入的声音和音频段被分析，跟踪频率变化的量。通过这种方式，每个频带的包络跟踪器可以实时生成动态控制信号。

### 合成信号

分析后的声音信号被送到一个合成滤波器，在那里 EVOC 20 PolySynth 滤波器通过视觉控制声音频率。显示框使用水平线作为分界线，蓝色部分用于分析部分，绿色部分用于合成部分，标志下的蓝色水平条用于确定滤波器允许的最高和最低频率，长度的变化代表滤波器

分析和合成所需的频率范围。

在处理滤波器的合成过程中，需要注意滤波器的频率设置。例如，如果分析信号缺少高频，那么最终合成的输出声音信号也将缺少高频。根据人耳的听觉习惯，高频的缺失会导致低频的加重，这将显得模糊且听不清楚。为了增强人声的清晰度，有效提升高频，并使所有频带的水平均衡化，以避免失真效果。

### 多声部合成器

EVOC 20 PolySynth 的独特多声部合成器可以将优美的人声合成为合唱声音。大多数人在发声时只能发出单一的声部，而 EVOC 20 PolySynth 总共有“poly”（多声部）、“mono”（单声部）、“legato”（连音）三种模式，不同模式的切换和组合可以使原始声音信号翻倍并叠加，改变人声的和声构成，产生更有“丰富感”的声音。

原声部分和加倍信号同时输出时，会增加声音信号的输出水平，容易产生声音的失谐。在合理压缩信号的过程中，使得水平变化越小越好，这样原始声音与合成的双倍信号的融合才会更好。

## 2.2 声音艺术的电子转向：语音合成

语言是人与人之间主要的交流方式，而语音合成——即自动生成语音波形，几个世纪以来一直是人类的梦想。使用计算机进行语音合成的探索可以追溯到 20 世纪 50 年代商用计算机的初期。事实上，这一遥远的里程碑在早期的澳大利亚科学与工业研究自动化计算机委员会 (CSIRAC) 项目中就已经讨论过。在整个 20 世纪 50 年代和 60 年代，计算机行业被昂贵的大型计算机所主导，这些计算机只能由学术机构和商业企业负担得起。

语音合成中的文本到语音 (Text-to-Speech, TTS) 合成方法包括两个主要阶段。第一个是文本分析，它将输入文本转录为语言或其他语言表示形式；第二个是语音波形的生成，在这个阶段，语音输出是通过语音和韵律信息生成的。这两个阶段通常被称为高层次和低层次的集成。

生成合成语音最简单的方法是预录制一段长的自然语音样本，比如单个词或句子。这种级联方法提供了高质量的语音质量，但有限的词汇量只适用于一些通告和信息系统。然而，显然我们无法为数据库创建所有的文本和通用名称。因此，对于无限制的语音合成 (TTS)，我们只能使用短的语音信号，如音节、音素等。

另一种常用的语音合成方法是共振峰合成，这是一种声源-滤波器模型。这种方法经常

被称为末端类比，因为它的模型仅由声源和共振峰频率组成，而不包括声道的其他物理属性（Flanagan 1972）。激励信号可以是单调的基频（F0）或清晰的噪声。这两种类型的混合激发也可以产生有声辅音和气音。

共振峰合成器利用滤波器参数来合成语音，其构造类似于模拟人体发声器官，因此能够自然地仿制人类的声音。共振峰合成方法的优点是，它合成的语音质量高于文本到语音（TTS）合成，并且具有很强的韵律调节能力。基频特性和共振峰是分离的，这意味着可以自由地修改基频和其他特征来产生不同的音色。然而，由于参数设置的复杂性，利用共振峰合成技术合成的语音自然度还有提升的空间。

共振峰合成语音可以很好地控制声音的基频，但要达到语音声音的自然度较为困难。第三种类型的拼接语音合成基于播放不同长度的预录制自然声音样本，这种方法正变得日益流行。其中最常见的是 FD-PSOLA 方法和 TD-PSOLA 方法。

FD-PSOLA（频域脉冲同步叠加）方法最初由法国电信（CNET）开发，基本计算分为三个步骤：首先，在分析阶段，原始语音信号被分割成分离的、重叠的短时分析信号；其次，在合成阶段重组和叠加这些片段；最后，在合成步骤操作中基频的基础上通过改变间隔来实现，并且持续时间的改变取决于语音是重复还是保存。原则上，基频的校正也意味着持续时间的变化。

TD-PSOLA（时间域脉冲同步叠加）方法是将自然语音变为可变长度的单元。这些单元可以是词、音节、音素，以及基周期、基本周期，并基于 TD-PSOLA（时间域基频同步叠加）算法。如果恒定音节之间的距离小于两个关键带宽，则通过修改重叠部分，并且原始波形与合成波形之间的映射来确定所需的短时合成信号序列。如果插值超过了两个带宽，则必须改变额外的中间原型的基频，并修改合成信号的幅度，从而在听觉上感知到信号的延迟变化。TD-PSOLA 方法的优点是计算量和存储量要求低于其他方法。

过去几十年间，语音合成技术得到了很好的发展，并且越来越多地融入到新的应用领域中。对于大多数应用来说，合成语音的可懂度和清晰度已经达到了可接受的水平。然而，在韵律、文本预处理和发音方面，仍有许多工作要做，以实现更自然的语音合成。

### **3. 人类语言的细节和发音规则**

人们想要理解人类语音合成技术，我们必须从人类语音合成的物理特性开始，也就是从人类的声音出发，本章从研究人类语言的细节入手，找出人类语音合成的发音规律，并结合人们对电子音乐的认知过程或心理过程的探索，来探讨语言和声音合成之间的关系。

## 3.1 人类发声的精确捕捉：声带与声学之间的关系

任何声音的产生都是由物体的振动引起的，物体在外力的影响下发生振动并产生声音，一旦振动停止声音也会停止。例如，当我们敲打鼓皮时，鼓皮会发出声音，我们可以用肉眼检测到鼓皮的振动，但当我们用手握住鼓皮时，鼓皮会慢慢停止振动，声音也会逐渐消失。同样地，人类的声音是由声带的振动产生的，当从肺部呼出的空气通过声道时，声带就会振动。当我们正常说话时，我们把手放在喉咙的位置，我们可以明显地感觉到振动。

所谓的声带在构造上相当复杂。声带位于喉部中央，由软骨和不同的肌肉组织构成，连接上部的咽腔和下部的通气管，是一个重要的发音器官。声带是两条坚实且有弹性的韧带，其前后端固定在软骨上，两个声带之间的空隙被称为声门，它由周围的软骨和肌肉调节。当我们在自然状态下正常呼吸时，声门是处于开放状态；而在发声时，声带闭合，肌肉收缩，软骨开始连续移动，从肺部呼出的气流撞击声门，声带在伯努利效应的影响下闭合，气压形成周期性的变化，使得声带放松或紧张，从而产生声波。配合吸气和呼气，通过控制声带的长度和弹性，使得声音在高度和强度上发生变化。

人类发声器官的生理结构主要分为三个要素：呼吸系统、声带和发声道，从声学的角度来看：动力源、振动器和共鸣器。呼吸是发音的能量源泉，声带在能量的激励下振动，形成在两个声带间隙上共鸣的气柱，使声音在共鸣器的作用下发出美妙动听的声响。

### 3.1.1 呼吸系统

提供发声所需的呼吸能量。人类的呼吸系统包括与鼻、口、喉、肺和隔膜相关的肌肉。肺部语言的动力站，由肺部生成的气流循环是人们发声的驱动力。连接肺部和喉部的是运输气流的气管，气流通过气管和支气管到达喉部，在声带之间的调节作用下，发出不同类型的语音。横膈膜是呼吸器官中更为重要的呼吸肌肉，肺部本身并不伸缩，而是通过在吸气和呼气的周期性运动中扩张和收缩胸腔容量来完成的。吸气时，横膈膜下降，胸腔容量持续扩大，使空气快速进入肺部，便于吸气。相反，呼气时，横膈膜逐渐放松，上升以支撑整个胸腔，使胸腔体积变小，帮助空气从肺部排出。

### 3.1.2 共鸣系统

改变声音的特性和音色。人体的共鸣器官由口腔、鼻腔、胸腔等构成。鼻腔、口腔和喉腔统称为上部共鸣体，胸腔称为下部共鸣体。胸腔这一共鸣体，通常用于低音区域。当我们以较低的音调说话时，可以强烈感觉到胸腔的振动，这是由于支气管压迫的空气在胸部形成了一个支点，产生了胸腔共鸣。相反，当我们提高音调甚至使用假声说话时，随着音调的升高，声音的振动越来越少，上部共鸣体的振动逐渐增强。

人类的发声器官与乐器不同，它由一系列可以改变自身形态和张力的共鸣器组成，因此声音具有一定的特性并获得复杂的音色。口腔是语言发音的主要共鸣腔，发音器官包括：唇部、上下牙齿、上颚、舌头等。舌头是口中最为活跃的共鸣腔，与其他器官的协调配合是产生不同元音辅音的必要条件。鼻腔在共鸣中也扮演一定的角色，与口腔相通，并且是通过头腔的唯一途径。当我们发声时，软腭抬起，呼出的气流以集中的方式流出，从而产生口腔共鸣。当我们触动软腭向下至舌面时，软腭中部振动，集中气流在鼻腔内流动并流出，从而产生鼻腔共鸣。头腔共鸣是在口腔共鸣形成的基础上，再加上一定的鼻腔共鸣，声音从鼻腔进入头腔，使得声音拥有丰富的泛音，从而产生头腔共鸣。

### 3.1.3 发音系统

塑造不同的发音部位以产生可辨认的语言。语言的产生依赖于不同的元音和辅音，主要使用的是口腔，它不仅是一个共鸣器官，也是一个发音器官。口中不同部位的协作，如上下唇、上下牙齿、软腭的配合，以及大小形态和弹性的不断变化，导致了多样化的语言形成。只有合理利用发音部位并正确区分咬字和吐字所需的相关部位，才能把握好字母发音的重要因素，使发音清晰准确。

人类男性和女性音色的差异主要是由于共振峰（formants）的差异，这类似于乐器音色的不同，而人类共振峰的频率依赖于声道的形状。女性的声道通常比男性的短，共振峰的频率也相应不同。因为较短的共鸣管比较长的共鸣管共振频率更高，就像乐器中的短笛和长笛，短笛通常比长笛高一个八度。这表明即使发出相同的元音字母，女性的声道也会产生比男性更高的共振峰频率。此外，身体自身的咽部和口腔的长度和大小也会影响共振峰区域的面积。通常女性比男性小，不同的声带振动，音色也不一致，对于声音的分类，这也是判断声音特点的基础。

## 3.2 准确掌握发声机制：咬字与吐字的节奏

声音合成和语言技术包含许多特殊的概念和术语,理解不同的声音合成及其分析方法的工作原理对于语音生成、发音音韵学以及其他一些相关研究至关重要。

在正常的语言交流过程中,人们将单词的发音视为一个整体,而不是将其分割成几个部分。例如,单词 "speak" 应该发音为 "speak" 而不是 "s-p-e-a-k"。英语强调音节的结构特点,在中文发音中我们注意到音节的完整性,这样做既使得词语清晰易懂,又有助于语调的变化。

人声合成技术在语音发音方面的挑战是巨大的。每个国家的语言发音都不相同,语言的发音有一定的地域性和民族性限制。人们通常在学习外语时,如何准确发音会感到无助,如果使用本国语言去模仿外语的话,不准确的发音除了会引起混淆外,还容易导致错误。通过国际音标的发音和基础语音学知识,以及在发音过程中英文和汉语的语言发音规则,本节为声音合成技术中的语音发音提供了良好的基础。

### 3.2.1 英语语音特点

随着语音学研究的发展,语音学家一直在寻找不同国家语言发音的分类方法,希望找到不同语言之间的基本语音共性,从而促进人们对外语发音的认识和理解。1988年,国际语音学会的国际音标部门开发了一套“国际音标音素表”,这至今仍是人们学习外语时最常用的方法。国际音标的重点是任何单一的元音和辅音都可以分解为语音元素,这使得熟悉国际音标的人容易掌握发音的方式。

英语一共有 26 个字母,包括 5 个元音字母 (A, E, I, O, U) 和 21 个辅音字母。后面一句不全 (English based on the country)

国际音标的分类总结了 26 个字母的发音,并将其划分为 48 个音素。

英语在音系系统上更为复杂,因为英语单词的发音常常因位置变化或与不同单词的组合而有所不同。

举例: low[lou ]—coward]' kau gd] good[gu d]-noon

[no:n]

### (1) 元音字母

英文中的元音字母有: a, e, i, o, u, 作为音节的核心, 被分为单元音和双元音。音节可以分为开音节和闭音节。

### (2) 辅音字母

英语中辅音发音的分类通常有三种方法。第一种是使用声带是否振动作为分类标准之一。发音时, 如果在喉咙处感觉到轻微的振动, 则发出的是浊辅音; 如果振动不明显, 则为清辅音。第二种是根据肺部气流穿过声门时的方式作为发音的分类方法:

半元音: [w]

第三种是根据发音时气流在不同腔室中的共鸣作为发音分类的方法。

## 3.2.2 汉语语音特点

### (1) 音节和音素

如果根据汉语的特点来分析语音的发音, 每个汉字的发音在语言系统中是一个自然的发音单元, 通过各种形式的声调组合, 形成了一个基本的声韵平仄形式, 我们称这种自然、有韵律的声音为音节。音素是构成音节的最小单位或最小语音段。

### (2) 元音、韵母、声调

在语音学中, 汉字通常被分为“元音”和“韵母”两种类型, 单词的发音是元音或韵母在音节中的发音。

前者部分称为声母, 统称为元音, 后者部分称为韵母, 统称为韵母。同时, 汉语是一种有声调的语言, 根据一定的发音规则相互协调和组合, 产生不同的音高变化, 这就是所谓的升降调。

元音的发音方式

发音是指发音器官的活动部分与固定部分接触产生气流阻碍和突破阻碍的方法。一般发音有三个阶段:

阻碍: 发音部位创建阻碍以阻止气流通过, 这需要阻碍部位的位置精确。保持阻碍: 持续保持阻碍阶段。

解除阻碍：解除阻碍

元音的发音在这三个阶段中相对较短，这与器官的呼吸控制息息相关。阻力足以成为发音受阻部位接受的阻力，阻力是在不断储存空气以移除障碍时气流强度的变化，从而元音的清晰度被直接影响。

元音的发音

元音的发音都是由于声带的振动产生的，并且不受发音器官的阻碍，因此声音可以变得非常响亮，并且可以无限延长。因此，在声乐表演中，歌唱中的声音和旋律都是由韵母来发挥的。

元音的发音是利用口腔共鸣腔的形状变化来实现的。在发音过程中，声带不断振动，声波进入不同腔室的共鸣体，元音则借助不同的部位改变焦点。

单元音：

由单个元音构成的元音称为单元音。单元音包括 a、o、e、ê、i、u、ü、i（前）、i（后）、er。以上单元音的发音要点总结如下：

复合元音：

由两个或三个元音组成的元音称为复合元音。复合元音的发音不是几个元音简单相加，而是在发音时根据嘴唇和舌头的形状逐渐变化到另一个元音的元音。在变化过程中，根据口腔开合的大小，它被划分为前响复合韵母和后响复合韵母。

声调的发音方式

汉语是一种声调语言，我们称一个音节或一个词在实际发音中的音高形式为声调。普通话有四个平仄声调，称为“阴平、阳平、上声、去声”。

在语言训练中，我们不仅要注意元音和韵母的协调，还要注意到自我组合过程中产生的变化。这就要求我们根据汉语的特点和结构组成进行综合训练，展现出汉语独特的韵味。

### 3.3 声音情感的终极追求：语言与音乐的融合

语言和音乐与人类的日常生活和文化交流密切相关，尽管音乐和语言广泛而普遍，但这两种形式的多样性和复杂性是人类特有的。因此，语言和音乐不仅可以激发人们的强烈情感特征，而且甚至能够相互影响情感变化和认知感受。语言是人类声音的一种复杂而有意义的序列，它在语言声音系统研究中被划分为两个主要领域：语音学和音韵学。语音学是研究人类发音的学科，包括语音的声学结构、语音变化的规律以及听觉和声音等感知机制的研究。音韵学是研究语音的声音模式的学科，包括音素的组合，比如音节如何构成高级词汇和句子，

语音如何随着语境的变化在听者心中产生心理变化等。

为了理解这两种重叠研究方法之间的差异,我们尝试考虑在说话时重音的加强和语调的变化。在正常的语言表达中,即使句子的特定头部和短语没有被强调,人们也能感知到音高运动是如何调整变化的。如果是研究音调或重音的长度和强度等声学参数,那么它属于语音学。然而,如果研究的是不同音高的位置和结构,那么它就隶属于音韵学。

### 3.3.1 语言中的音高对比

语言的复杂语调元素主要是由于音高之间的对比,而重音通常通过音高运动的变化表现出来,如音域的扩展和音阶的增加。语言的特点在于声带振动的基频(F0),振动次数越多,频率越高,音调相对越高,反之亦然。声带的厚度和紧张程度也对音高有相应的影响。声带厚且长,紧张度低,音调低;声带薄且短,紧张度强,音调高。

### 3.3.2 语言中的节奏变化

尽管诗歌节奏的研究有着悠久的历史,但在语音学领域中,对音乐中人声节奏的研究相对较少。语言在音乐节奏中的长度受到多种因素的影响,例如不同语音元素的持续时间受到发声的限制,不同旋律所需的音节速度差异导致了歌词的“随意”和“清晰”。

### 3.3.3 语言中的旋律变化

语言和音乐的旋律在结构和认知处理上具有相似之处。旋律包含大量的音调序列,而语言在正常的口语交流中,包含了传达情感、句法和重音的信息。旋律之所以能被记住,是因为它使听众产生许多心理模式,这些模式被转化成有意义的关系网络。

音乐和语言之间的差异在于,音乐中的旋律通常具有相对稳定的音高特征,而语言旋律则不然。音乐歌唱中的旋律往往围绕中心音进行演化和衍生,而语调中的旋律通常根据语言学的需要来组织,比起说话来,更具有审美性。

### 3.3.4 音色

音色,这是一个含糊的术语,通常指的是声音的独特音质。与音高和价值相比,音色的

鉴别更加个性化和多面性。不同的人，声带振动器产生复杂的。【最后一句不完整】

声音的数量和声波是有区别的。在复音中，有许多声波，其中频率最低且振幅最大的被称为基音，其余频率高于基音或者是其基频整数倍的声波，被称为泛音。

### 3.3.5 空间性

斯托克豪森认为，除了声音的高低、长短、强弱和音色等四个特性之外，还有另一个特性，那就是声音的空间位置。他的代表作《青年之歌》是第一个将序列技术应用于空间声音设计的音乐作品。它于 1953 年首演，周围是五组扬声器。

## 第2节 研究与实现计算机生成剪纸风格化的关键技术

剪纸是中国最古老的传统艺术形式之一,因为剪纸所使用的工具和材料在日常生活中非常普遍,应用领域也非常广泛,所以无论是北方地区还是江南、中原或边疆地区都能看到不同种类的剪纸,它是我国广大人民群众喜闻乐见的一种民间艺术。剪纸艺术在刻画人物形象时具有自己独特的特点,这是由于生产工具“剪”和生产材料“纸”的生产条件所限制,它只能通过镂空、鲜明的色彩对比以及物体的连接来刻画不同剪纸的形象,由于这些限制,剪纸艺术不能展示宏观的大场面和非常复杂的层次。剪纸艺术家需要丰富的想象力和高度的概括能力,只有这样,才能反映现实生活中最动人的形象,它们常常使用夸张、变形的技巧和精美的线条来描绘具有装饰性、教育性和讽刺性的艺术形象。

然而,在早些年,剪纸的现状并不乐观,生产和传播“剪纸”人物和场景不可避免地存在限制。一是手工剪纸效率低下,二是制作剪纸的成本相当高,三是对剪纸的可重用性有很大限制。相比之下,如果传统民间艺术如剪纸由计算机生成,可以克服上述问题,它不仅能真实再现剪纸技艺,而且易于生产、可重复使用,更重要的是,现在是一个信息爆炸、互联网爆炸的时代,将剪纸作品保存在计算机中,通过现代网络技术使剪纸艺术传遍全世界,让更多的人重新了解剪纸,使中国的民间传统艺术再次得到发扬。在这个系统中,我们使用非真实感渲染(NPR)技术让计算机生成剪纸风格化的人物、动物和场景。

在计算机图形学领域,首先出现的是写实绘图技术,随后出现了非写实绘图技术。非写实绘图技术是在写实绘图技术的基础平台上发展起来的。实际上,写实绘画早于计算机图形学的诞生。在1620年代,西欧艺术家就能够使用写实绘图工具画出非常逼真的画作。使用写实感绘制的图像具有非常逼真的特点,这几乎与人们通常所看到的真实世界一致,符合每个人的视觉习惯,易于为大家所接受。凭借上述优势,写实绘画很早就深入人心,这对于多年来计算机图形学围绕写实绘图技术为中心的发展提供了很好的解释。通过上述介绍,不难看出写实绘图技术已经成为应用范围广泛的成熟技术。

然而,由于写实感绘制的图像反映现实如此精确,艺术家在画作中加入艺术感来表达已变得困难。正是在这个时候,非写实绘画应运而生,为印象派艺术家所利用。这给了艺术家极大的自主权来传达他们想要表达的情绪,这使得通过非写实绘画技术绘制的事物具有强烈

的艺术感。当然，通过非写实绘画，画家也可以将当时的内在情感加入到他的作品中，并通过不同的色彩、不同的光线和不同的笔触来传达艺术家想要展现的情绪。

中国浙江大学 CAD&CG 国家重点实验室在非真实感研究工作中主要涉及卡通动画生成、水彩风格图形绘制、基于图形的中国画风格图形绘制等。在水彩图形绘制中，他们提出了一个三维计算机水彩画笔模型，包括笔刷单元选择、颜料模拟、颜料扩散和整体控制 4 个部分，模型结构简单，可以直接应用在传统三维表面上，绘制的结果具有手工绘制水彩画的效果。在绘画风格的图形描述中，利用可编程图形硬件设计了一个实时墨效果绘制系统，用于三维山体模型，在该系统中，首先建立一组对应不同层次的纹理图像，图像中的笔触纹理不按手工绘画统一分布，构成了明暗变化的纹理序列。系统运行时，首先检测山体模型的位置确定它所在的层次，然后根据山体模型的光照值混合相应的纹理，山体底部混合了北京的云雾图，最后获得具有中国画风格的图画。

首先是艺术领域。艺术家可以通过非写实绘画技术表达他们的艺术构思，达到“善人看到善，智者见智”的效果，使得所绘制的图像产生多种不同的艺术风格。这主要体现在以下几种绘画类型中：钢笔画、铅笔画、水彩画、水粉画、水墨画、印象主义画和中国画。

第二是教育学领域。图像可以传递更多信息，其传输方式简单，且相对于文字具有很大的优势。这使得儿童更容易接受信息，因为他们的认知规律倾向于偏好简单、色彩鲜明的图片，所以需要非写实绘图技术的协助。鉴于这一优势，许多研究人员已经在教育领域加入了非真实感绘制 (NPR) 的研究。

另一个重要应用领域是各行各业的工业场合，这也是近期的热点研究。例如，在 CAD (计算机辅助设计) 领域，当你想要显示一个零件的细节时，无论是零件的尺寸、大小还是间距。

## 1. 非写实绘图技术的分类与研究方向

非真实感绘图技术可以根据不同的方式进行多种分类，其中基于笔触的渲染 (Stroke Based Rendering) 和基于图像的渲染 (Image Based Rendering) 是根据实现方式来分类的。不需要用户干预的绘图和需要用户干预的绘图是根据是否需要用户干预来划分的。2/2.5 维图像绘图和 3D 图像绘图是按照维度来分类的。基于模型的非真实感绘图、基于图像的非真实感绘图、混合型、交互式 2D、3D 绘画系统也是 NPR 的分类。像钢笔画、铅笔画、水彩等都属于基于模型的 NPR。基于图像的 NPR 不仅处理单一图像，也处理视频流。交互式 2D 和 3D 绘画系统允许用户创造具有非常不同风格的场景。

NPR 图像绘制的主要研究方向包括:

半调技术: 主要是图像马赛克、有序抖动、扩散误差等。

钢笔画模拟: 主要是钢笔绘画、点画法、铅笔画的模拟。

绘画模拟: 主要用于水粉画、水彩画、水墨画的模拟。

## 2. 网格模型

我们在现实生活中看到的物体是具有三维感的三维模型, 但如何在计算机上表示三维物体呢? 通常使用 3D 模型, 这主要用于表示现实中物体的三维多边形。这意味着, 只要是自然界中存在的物体, 都可以通过三维模型来表示。3D 模型通常使用专业的软件工具生成, 如 3D 建模工具, 例如 3DMax、MAYA、PhotoShop 和 SoftImage 等 3D 建模软件。生成三维模型的另一种方式是通过三维扫描仪或三维数字化仪。它目前是现实世界中物体进行 3D 建模的主要工具。它的工作原理是使用三维扫描仪将现实世界中的立体色彩信息转换为计算机可以直接识别和处理的数字信号。另一种 3D 模型生成方法是基于图像的建模和渲染

(Image-based Modeling and Rendering, 简称 IBMR)。

计算机中三维模型的构成有两个主要方面, 一方面是网格。网格是由许多点云组织起来, 以形成三维模型的结构。另一个方面是纹理。它既可以指通常所说的物体表面的不平整纹理, 也可以指那些让物体表面显得平滑的彩色贴图。

总结来说, 网格模型可以说是真实物体的一种表现, 它通过许多空间三角形片来表示三维物体。网格模型分割指的是: 根据一定的拓扑结构, 将那些封闭的网格多面体, 根据其表面几何特征、拓扑特征, 分解为一定数量的子网格, 每个子网格具有简单的形状意义, 并且相互连接。然后, 本文基于这种网格模型, 改进了网格模型的分割算法。网格模型的分割算法有很多种, 其中按照网格模型的规则性分为规则网格分割、半规则网格分割和任意机制网格分割算法。

## 3. 剪纸风格化图案生成算法

图案是剪纸中非常重要的一部分, 它决定了剪纸的基本特征。然而, 在现实生活中看到的剪纸结构复杂且图案众多, 但通过观察, 图案的组成还是有许多相似的特征, 例如, 大多数剪纸图像的轮廓都是非常基本的图案结构; 在剪纸中, 动物头部的面部特征, 如鼻子、眉毛、眼睛、嘴巴等, 几乎都是用特殊的图案来绘制的。

## 3.1 剪纸图案分类

根据剪纸图案的结构和特点，我们将其分为以下三种不同的类别，即：基础结构图案、特殊图案和装饰图案。

### 基础结构图案

基础结构图案主要用于绘制对象的大致物理特征，通常包括头部、躯干、四肢和尾巴。它们主要用于绘制剪纸动物的身体轮廓，图案的基本组成部分主要包括直线、圆弧、贝塞尔曲线、B样条曲线等，通过这些可以简单地绘制出剪纸动物的大体轮廓。

### 特殊图案

这里的特征图案可以细分为两个子类别：特性图案和特定图案。其中，特性图案主要指那些能反映剪纸对象特征的图案。例如，马、公鸡等动物的尾巴。

在将这些特征图案添加到相应的剪纸图像中时，可以实现画龙点睛的效果，让剪纸图像更加生动，更贴近实际手工剪纸的效果。当然，这些特征图案并不是固定不变的，它们可以通过控制点来调整图案的大小、线条的粗细等，以适应不同的剪纸图像需求。另一个小的分类是特定图案。这种特定图案主要指的是动物面部的图案，比如眉毛、眼睛、鼻子和嘴巴。

### 装饰图案

从装饰图案的分类来看，它反映了中国剪纸表现形式的风格化，因为自然中的场景是复杂多变的，但通过装饰图案的概括，它变得有条不紊且具体化。这种经过提炼和加工的简洁明了的图案，对于剪纸艺术来说更具有艺术性和装饰性，这也是我们剪纸的民族特色。通过对许多剪纸图像的分析、整理和综合，可以得出在装饰中最常用的图案主要是锯齿图案、花卉图案和新月图案这三种。

## 3.2 装饰性剪纸图案生成算法

### 新月图案生成算法

新月图案在剪纸中是一种更为常见的装饰图案,许多复杂的剪纸图案都是基于新月图案进行变换和组合得到的。在此系统中,通过绘制贝塞尔曲线作为基本框架来生成新月图案。

### 花卉图案生成算法

剪纸中的花卉图案也可以称为梅花图案,它与新月图案一样,也是剪纸中非常常用的装饰图案,其生成算法同样非常简单:将类似泪滴的图案限制在一个圆形之内。在此系统中,用户可以根据不同的需求设置相应的花瓣数量,以便生成的花瓣效果不仅结构简单,而且装饰效果非常美观。

### 锯齿图案生成算法

我们可以将锯齿图案想象为类似木梳的结构,即一些短线分布在一条曲线的一侧。我称那条曲线为主线,因为它是产生那些短线的基础。首先,确定主线的两个端点 X 和 Y,只要在曲线的起始端点 X 和结束端点 Y 之间确定一个中间点,就可以大致确定主线的曲线,这里也是用 B 样条曲线来绘制的,然后使用样条曲线插值这些点的中间,以获得锯齿图案的主线。在图案的分类中提到了基础结构图案,在这里也会使用锯齿图案,通常这个地方使用的锯齿图案的主线长度相对较长,曲线也非常大,并且变化的程度相对较陡峭,这需要许多控制点来把握主线的大致形状,例如主线的长度,以及弯曲的程度。

许多短线排列在主线的一侧,每个短线的形状都根据锯齿主线的形状和曲率来确定。我们可以称这些小短线为小锯齿,小锯齿的整体结构沿主线的方向排列,小锯齿的法线垂直于主线,并且其长度可以视为小锯齿的高度  $h$ ,它们之间的间隔也将按照后续的算法来确定。

## 3.3 特定剪纸图案生成算法

特定剪纸图案是我们从复杂剪纸图案中区分出来的一类图案。正如其名,特定剪纸图案用于那些特定的且不会改变的图案,例如眼睛、鼻子和嘴巴的图案,一旦制作出来后,它们

就可以被应用于不同的动物图像上, 仅需根据动物图像的大小进行尺寸调整, 特定图案的形状则保持不变。这些特定图案主要被用在动物形象的面部结构上。本节主要介绍特定剪纸眼睛的两种形状和相应的生成算法。

第一种哺乳动物的眼睛生成算法很简单。它是由两个相对称的新月形条纹生成的。在这个系统中, 作者没有为这个算法专门生成一个眼睛图案, 那是因为这个图案的生成非常简单, 只需要由两个对称的新月形图案生成, 不需要为其特别设计, 用户可以绘制一个新月形图案, 然后通过复制可以生成另一个新月形图案, 再调整相应的两个新月形图案的位置, 就可以得到你想要的特定图案——眼睛图案。

上述内容主要介绍了中国传统艺术中剪纸的主要特征。解释了剪纸图案对剪纸艺术的重要性。随后, 作者对剪纸图案进行了三种分类: 基础结构图案、特定图案和装饰图案。通过将剪纸图案分类, 我们可以将不同的图案放入不同的图案类别中, 这样我们可以找到图案并在以后修改和再次使用它。上文介绍了剪纸图案分类中的一种——装饰图案, 主要介绍了三种常见的装饰图案生成算法: 新月形图案生成算法、花卉图案生成算法和锯齿形图案生成算法。并在系统中分别生成了这三种装饰图案, 用户可以通过控制参数和图案的调整点来控制装饰图案的形状, 方便用户编辑。上文描述了特定图案中生成眼睛图案的两种算法, 并详细描述了第二种眼睛的生成。

# 第3节 基于可视化技术的流程装配系统的设计与实现

## 1. 研究背景与意义

由智能制造引领的“工业 4.0”时代已经到来，航空发动机在智能制造领域是一个非常重要的领域。航空发动机制造能力是一个国家综合实力的象征，最关键的是能够衡量一个国家综合实力的研发和制造水平，并且在整个行业中具有极高的经济、军事和政治价值。航空发动机是西方发达国家发展的优先行业，也是外部封锁的核心技术[5]。

本节研究的内容基于实习公司的产品 UltraCAM（复杂自由曲面零件的数控编程软件）。该产品旨在开发航空发动机关键零部件整体叶盘工艺的软件系统，目的是为用户提供如整体叶片、叶片、叶轮、风扇等高质量的制造服务，并为航空发动机等用户提供一套多加工策略、易学易用的编程工具，同时为用户提供完整的整体叶片平台制造工艺解决方案。通过软件平台，可以导入用户给定的模型文件，经过 CAD 系统、工艺规划系统、CAM 系统以及后处理系统等，并最终输出机床可识别的 CNC 程序机床代码，然后可以放到机床上进行加工。目前，系统已完成整体叶盘的多个工艺，如：叶片粗加工、叶片精加工、前后缘加工、过渡圆角精加工、流道粗加工、流道精加工、流道粗加工（周向侧铣）、内插铣削、叶尖精加工等多种叶片工艺，但仍存在许多问题：

(1) 在 UltraCAM 系统生成的数控程序机床代码放到机床上加工时，缺乏在可视化中进行干涉、过切和欠切检查，这导致铣削工具更容易磨损，也极大地影响了整体叶盘的质量和开发周期。由于缺乏对工艺参数的反馈，特别是当发生干涉、过切和欠切时，没有给出相应的参数修改方案，导致工艺参数设置只能通过工艺人员的经验来进行，而对于经验较少的普通用户来说，使用难度大大增加。

(2) 无法组装多个工艺。多个工艺没有作为一个整体来考虑，导致在机床加工时，操作者完成一个工艺的加工后，需要旋转机床一个角度，然后再处理下一个工艺。这样做非常耗时且费力，不利于从机床中解放机床操作员。此外，手动操作的疏忽直接导致整体叶盘加工

失败，给企业和公司带来巨大的经济损失。

(3) 工艺计算中缺少用户停止和调试。现有的工艺计算缺少用户停止和调试功能，导致用户在计算过程中不能停止计算，只能徒劳地等待直到工艺计算结束；缺乏调试功能导致工艺人员在计算失败时无法调试工艺参数，特别是工具参数。

工艺组装系统具有非常重要的研究意义，具体如下：

(1) 减少机床手动操作时间。在实际的机床加工中，机床操作员在加工一个叶片后需要转动机床一个角度，这非常耗时且费力。通过工艺组装系统，操作员无需进行这些操作，从而将操作员从繁琐工作中解放出来，节省操作时间，提高机床加工效率，为企业节省大量成本。

(2) 降低软件使用门槛。由于工艺设置不再仅仅基于工艺人员的工作经验，它为普通用户在设置工艺时提供了推荐解决方案。这样，即使是经验较少的用户也可以更容易地使用软件，进行工艺规划和编程，从而提高了软件的可用性和工艺的准确性。

(3) 提高交互能力。因此，设计和实施工艺组装系统是目前非常有益的事情，它可以为工业界和学术界的创新和产品创新提供强有力的支持。该项目从工艺计算开始，到多个工艺的组装，再到后期的可视化，最终满足用户需求。因此，对这一主题的研究具有广泛的应用前景和重要意义。

## 2. 国内外相关研究现状

### 2.1 可视化

可视化（Visualization）是指运用计算机图形学和图像处理技术，将数据转换为屏幕上显示的图形或图像，并进行交互式处理的理论、方法和技术。即将传统意义上无法看见的事物或现象呈现给专业人员，使他们能够观察到模拟和计算的整个过程。其目的是依靠人们强大的视觉观察能力，促进对观察数据的更深入理解，并培养对潜在过程的洞察力。

可视化技术包括科学计算、工程计算、图像处理、计算机图形学、人机界面等多个学科和技术。其过程是将数据信息转换为图像信息，在保持数据信息完整性的同时，将这些信息转换为人类视觉系统可以接受和理解的形式。这个过程基本上可以分为三个子过程：数据处理、可视化映射和渲染。

## 2.2 视觉仿真

3D 视觉仿真是一种集成了计算机图形技术、多媒体技术和数据库技术的新技术。它最早在美国得到发展。在曼哈顿计划中，随着科学家们为军事需求开发先进技术，仿真技术成为预测气候变化和武器生产测试的关键技术。因此，自 20 世纪 80 年代以来，仿真领域普遍重视建模技术，并将人工智能技术、图形技术和知识工程技术引入仿真建模过程。随着计算机技术和仿真技术的发展，一方面，仿真系统的结构复杂且规模不断扩大，正向分布式仿真发展；另一方面，仿真系统在复杂系统的设计分析和决策管理中发挥的作用日益重要，导致对仿真信心的要求显著提高。因此，仿真系统功能的智能化、集成化、仿真软件设计和自动化是仿真软件系统发展的主要趋势。计算机图形仿真技术的使用不仅可以验证加工程序的有效性，还可以减少试样的成本，节省机床的加工和调试时间，缩短新零件的开发周期。因此，CNC 机床仿真技术具有重要意义。加工过程仿真仍处于起步阶段，目前存在以下问题：

仿真的加工形式较少，研究范围狭窄，加工理论复杂，并且不同的加工方法和工具形状的加工模型有很大的不同。同时，大多数当前的仿真系统进行的是几何仿真，即工具位置轨迹的干涉校准、工件和工具等，也被称为数控验证。物理仿真与实际切割过程之间存在较大差距，主要考虑的是理想的切割状态。

仿真的局限性制约了其系统的发展。当前仿真系统发展的趋势是应用 C++ 语言和面向对象的方法。为了使图像更加立体逼真，使用 OpenGL 进行可视化模块的开发。OpenGL (Open Graphics Library) 定义了一个跨编程语言、跨平台编程接口的规范，是一个专业的图形图像程序接口。OpenGL 是基于 SGI 的 3D 图形库的通用共享的开放式 3D 图形标准，它是一个功能强大且易于调用的底层图形库，独立于操作系统和窗口系统，由大量强大的图形函数组成，并整合了所有图形变换、照明、像素操作、表面建模和光栅选择等计算机图形算法。

在使用 OpenGL 进行 3D 图形渲染时的基本步骤包括：

**构建场景模型：**根据基本图形单元建立场景模型，并对构建的模型进行数学描述。

**放置场景模型：**将场景模型放置在三维空间中的适当位置，并设置视点。

## 2.3 根据应用需求计算模型中所有对象的颜色、光照条件、纹理贴图方法

### 虚拟装配

虚拟装配是指使用数字化和虚拟现实等技术来完成产品设计和装配工作。通过在计算机上的仿真，可以直观地识别产品的潜在设计问题，这些问题在生产前容易得到解决，从而提高效率并降低成本。外国对虚拟装配技术的研究始于上世纪 90 年代，一直持续到现在。虚拟装配技术的发展大致可以分为三个阶段：提出虚拟装配理论、构建虚拟装配系统和深入研究虚拟装配。

#### 2.3.1 外国对虚拟装配技术研究的现状

美国目前处于虚拟装配技术的前沿。华盛顿州立大学和美国国家标准与技术研究院 (NIST) 共同开发了 VADE 系统。它结合了 CAD 虚拟环境，可以更直观地显示模型数据，能够对模型进行拆卸和组装等操作，以形成一个完整的组装模型，并能验证结果。他们已经开发出第二个系统并且非常成熟。该系统可以应用于产品验证和装配培训等各个方面，具有良好的应用前景[6]。

德国在虚拟装配领域的研究也相对较早开始，而且弗劳恩霍夫虚拟现实实验室开发的虚拟装配规划原型系统在研发后不久就在慕尼黑电脑展上获得了奖项。该系统建立了虚拟环境模型和虚拟人体模型，并且还使用了一些输入和输出设备，允许操作者与系统进行交互，并控制虚拟假人和三维模型的运动，从而模拟现实中的整个装配过程。

美国宇航局的哈勃望远镜项目同样如此。美国宇航局应用虚拟现实技术对空间站的操作进行虚拟模拟。为了在轨对哈勃空间望远镜进行维护，美国开发了哈勃虚拟维护装配系统，该系统可以在地面上模拟望远镜在太空中的实际操作，使宇航员能够在地面上清晰地了解实际操作任务和维护环境，并能在虚拟环境中完成维护，达到培训的目的。该系统现在已经训练了多名宇航员，并完成了四次哈勃在轨维护任务，对其进行了修复和升级。

### 2.3.2 国内虚拟装配技术的研究现状

清华大学的 CIMS 研究中心在装配设计、规划和建模等多个方面进行了深入研究。他们在 Pro/Engineer 商业 CAD 平台上开发了一个虚拟装配系统。该系统实现了拆卸模拟、生成工艺文件和装配过程规范等功能，并且已经被应用在产品开发中。

北京航空航天大学在虚拟现实技术领域的研究在中国也是较早的，同时也是中国非常权威的单位之一。他们对虚拟原型和虚拟装配基础技术进行了研究，并且在虚拟现实视觉方面取得了良好的理论和成果。

西北工业大学的教授分析了现有装配方法的不足，并提出了基于协调关系的装配路径规划方案。像邵毅这样的教授使用 CATIA 的二次开发接口实现装配路径规划算法，开发了 DMSS 虚拟装配系统，并通过实例验证。结果表明，这种方法可以有效解决大多数组件的装配路径解决问题，并具有良好的实际应用价值。

上海交通大学也研究了虚拟装配中的约束表达、识别与确认、以及导航问题。在这项研究中，提出了装配约束的表达方法，并描述了约束与自由度之间的关系，解决了多重约束下的自由度问题。然后，在完成约束识别和约束确定的基础上，进行了组件姿态的导航。这一方法已经在许多工业环境中得到了验证。

### 2.3.3 过程装配

装配是每个产品形成过程中的关键活动。特别是在复杂结构产品的装配中，这个过程是按层次和单元进行的。在每个装配单元中，为了保证产品装配要求和装配质量，组合过程具有更严格的工艺技术要求，同时也有装配过程顺序位置的要求。

当前，过程装配面临着以下问题：

#### 1) 低效率和高成本

对设计人员提出了较高的要求：需要阅读图纸、行业标准和工艺规范等，必须具备丰富的生产实践经验，工艺准备的效率难以提高，因此无法缩短产品开发周期和降低产品及企业成本。

#### 2) 产品装配信息不准确

在传统复杂产品装配过程规划中，图纸是装配信息和几何信息的主要来源。由于图纸是由产品设计师绘制的，信息的掌握主要依赖于工艺设计师对图纸的理解，因此在信息理解上存在

模糊性和不完整性。

### 3) 难以保证装配规划的质量

传统的装配过程规划主要由工艺人员完成，由于个人知识和经验的局限，在相同的生产条件和生产环境下，可能会准备不同的装配过程，因此装配过程规划具有很多经验性、主观性。这使得难以保证装配过程规划的质量。

### 4) 无法满足多变和快速的市场需求

在当今市场竞争中，产品是核心。现有的装配过程规划已不足以满足快速多变的市场需求。计算机辅助技术和信息技术的发展进步，现有的装配方法严重限制了企业的发展。

### 5) 无法为操作人员和相关部门提供统一的概念

随着市场需求的日益多样化和市场竞争的日益激烈，生产制造的速度应当加快。现有的装配过程规划方法已无法满足产品市场快速变化的需求，以及制造系统效率高、灵活性高、信息集成度高的要求，这对现有的装配过程规划提出了新的要求。

## 2.3.4 可视化装配

在计算机上模拟组件的动态装配过程和现实中的装配过程被称为装配可视化 (Assembly Visualization)。也就是说，利用仿真技术和计算机图形学在计算机上模仿产品的实际装配过程。装配仿真以一种直观的方式展示、验证和改进产品装配，从而缩短产品开发周期并降低成本。装配仿真可以分为三个层次：

演示产品装配过程，验证装配能力，并为过程装配规划提供一个可视化的人机交互环境。

进行装配分析的模拟。利用虚拟环境技术进行动态模拟，形成模拟的设计可行性分析 (DFA) 评估环境。

基于虚拟现实 (VR) 技术的装配设计。利用虚拟现实技术的交互特性，进行虚拟原型的抓取、放置等操作，形成了一个装配设计环境。

虽然三个层次的装配仿真对产品开发具有重要意义，但它们仍处于研究和探索阶段。因为虚拟技术尚不成熟且设备昂贵，第二和第三层次的研究在中国不易进行更深入的研究和大规模推广。但相比之下，第一层次的研究更具实用性且发展迅速。

在装配仿真方面，国外已经有装配软件系统，例如波音公司应用虚拟现实仿真来训练和指导操作员进行装配；国内的研究主要在产品的设计阶段，模拟装配过程，进行运动碰撞干涉检查、装配顺序等，并在设计阶段获取装配信息。然后通过这些信息，再次进行设计，以

解决产品装配中的各种问题，即为装配而设计（Design for Assembly, DFA）。

工艺装配系统主要涉及三个部分，即工艺的计算、工艺的装配和工艺的视觉模拟，本文的主要研究内容如下：

对 UltraCAM 系统工艺计算模块进行了改进，以构建工艺装配系统 VPAS 工艺计算。这包括工艺工具点计算的过程，通过表面偏移、表面分级、表面交点、计算工具轴向向量等步骤来获得最终的工具点，在此基础上，通过多线程技术和串行计算优化，同时增加了用户停止计算和调试计算结果的功能[7]。

通过建立工艺装配树来装配工艺，设置旋转面等，实现工具在多个工艺之间的安全进给和撤出。

设计并实现加工过程的视觉仿真。通过使用 Qt 创建界面，利用 Open CASCADE 进行工具路径的过程视觉仿真，在仿真中同时进行工具干涉检查、刮削检查、欠切检查等，并在出现干涉点、过切点和欠切点时为用户提供相应的解决方案。

### 3. 相关技术简介

本文的主要工作是设计并实现工艺装配系统 VPAS，涉及工艺计算模块、装配过程模块和过程可视化模块。本章首先介绍整体叶盘的相关知识和定义，然后介绍工艺装配系统中使用的工具和 IGES 文件与 STEP 文件，接着介绍 VPAS 的界面开发语言 Qt，以及计算和可视化模块所需的开源几何内核库 Open CASCADE，最后介绍相关算法和 3D 建模方法。

#### 3.1 整体叶盘的定义

由于传统叶轮存在的缺点：采用焊接加工，即将复杂加工分解为多个叶片加工，然后利用焊接技术将所有叶片和叶轮盘整合在一起。因此，德国、英国等国家共同开发了一款名为 EJ200 的大型航空发动机，该发动机采用了一体化叶片盘（Blisk）结构。随后，美国开发了一种名为 F119-PW-100 的发动机，其风扇和转子采用了一体化叶盘结构。后来，民用发动机也采用了一体化叶盘结构。

一体化叶片是航空发动机和燃气轮机的核心结构部件，将传统的分离式转子叶片和叶轮盘整合为一个整体，简化了结构，提高了飞机发动机和燃气轮机的性能与可靠性。整体叶盘通常使用镍基高温合金和钛合金作为材料。

整体叶盘的核心部分是叶片。叶片加工的工艺流程通常包括以下几个步骤：

- 1) 将用户提供的叶片模型文件（STEP 或 IGES 文件）导入 UltraCAM 系统；

- 2) 对导入文件中的叶片进行重构, 以获得光滑的叶片表面;
- 3) 添加叶片所需的加工工艺, 并进行计算;
- 4) 对成功计算的工艺进行后处理操作, 获取机床加工的机器代码;
- 5) 使用机器代码在机床上加工每个叶片, 以获得最终的整体叶盘产品。

### 3.1.1 整体叶盘的分类

整体叶盘有几种分类方法。如果按照建模方法来分, 整体叶盘可以分为开放式整体叶盘、半开放式叶盘和封闭式整体叶盘以及整体叶盘模型; 如果按照在发动机上的位置来分, 可以分为压气机的整体叶片盘和风扇的整体叶片; 如果按照加工方法来分, 可以分为一体式整体叶盘和焊接式整体叶盘。

与传统的叶盘结构相比, 整体叶盘具有结构复杂、直径大、刚性弱、叶片薄、流道狭窄和受力变化性大的特点。

### 3.1.2 整体叶盘涉及的相关定义

为了便于后续在工艺计算中的描述, 需要对整体叶盘进行定义规范。UltraCAM 系统中的所有模型都源自标准的 IGES 文件或 STEP 文件。由于这些模型文件各不相同, UltraCAM 对它们重新进行了定义, 我们以开放式叶盘为例进行说明。

开放式整体叶盘模型由三个部分组成: 叶片、外环(shroud)和轮毂(hub)。其中, 带型曲面(belt profile)和轮毂曲面(hub profile)都是绕 x 轴旋转的旋转表面, 而叶片型曲面(vane profile)是一个整体的自由形曲面。在实际加工过程中, 带型表面并不存在, 也就是说, 只有轮毂表面和叶片型表面。在 UltraCAM 系统中, 带型表面只是一个辅助表面, 用以确保叶片型曲面存在于带型曲面和轮毂曲面之间。开放式整体叶盘模型从下到上依次是轮毂、叶片和带型。

UltraCAM 系统使用的坐标系统是笛卡尔坐标系统, 或称为右手坐标系统。其中 x 轴与整体叶盘的轴线共线, 它们也是整体叶盘的旋转轴。对于开放式叶盘, z 轴方向指向主叶片(相对于邻近的叶片, 主叶片指的是 UltraCAM 软件中要加工的叶片), 并且从轮毂面到带型面, 必须通过主叶片。坐标原点 O 位置是在 x 轴上任意选取的。

## 3.2 工具介绍

工具在叶片加工的计算和模拟可视化中扮演着关键角色,因此有必要了解不同的机加工工具。在 UltraCAM 系统中目前有六种工具类型,分别是球头刀具(ball head cutters)、圆鼻端铣刀(round nose end milling cutters)、平底刀具(flat bottom cutters)、端铣刀(end milling cutters)、LOLL 刀具以及 BARR 刀具。

其中,球头刀具和平底刀具是使用最为频繁的刀具。球头刀具主要由一个圆柱体和一个半球体组成,而平底刀具则由一个圆柱体和一个圆台体组成。

## 3.3 三维通用文件介绍

随着工业制造的发展,国内外出现了大型三维 CAD 系统软件,其中主要代表有 UG 等。然而,由于产品初始数字化发展的限制,每个系统都有自己的派系,有自己的产品信息模型和自己的图形系统。

等一下。要做到这一点,将不同 CAD 系统中的产品导入到 UltraCAM 系统中必须是相同的通用文件格式。常用的通用标准文件有 IGES 文件和 STEP 文件。

### 3.3.1 IGES 文件

IGES 文件是由美国国家标准局和工业界在 1975 年开发并实施的。IGES 文件接口不仅能够处理三维线框元素、等距偏移曲线、表皮及表皮边界,而且还使设计师能够在两个完全不同的系统之间可靠地直接交换数据。

IGES 文件是 ASCII 格式的,包括五个区域:起始部分(Start Section)、全局部分(Global Section)、目录部分(Directory Section)、参数部分(Parameter Section)和终止部分(Terminate Section)。

起始区域(Start Section)。这个区域包含两个主要功能:记录图纸文件的来源以及生成该文件的系统的名称。

全局区域(Global Section)。这个区域描述了 IGES 文件使用的文件名分隔符、参数分隔符和记录分隔符等信息。

目录区域(Directory Section)。提供整个 IGES 文件的索引,需要包含每个实体的属性等信息。实体索引记录的顺序是任意的。

参数区域 (Parameter Section) 。包含与每个实体相关联的参数数据。

终止区域 (Terminate Section) 。IGES 文件的最后一行, 每个区域都包含一个区域标识符和区域的总数。

### 3.3.2 STEP 文件

STEP 产品数据交换标准指的是由国际标准化组织 (ISO) 制定的系列标准 ISO 10303 “产品数据的表达与交换”。这个标准的主要目的是为了解决制造业中计算机环境下的设计和制造 (CAD/CAM) 中的数据交换和企业数据共享问题。

STEP 文件的架构分为三层: 物理层、逻辑层和应用层。

应用层: 这是属于特定应用的顶层, 即包括应用协议和与应用协议相对应的抽象测试集的应用层;

逻辑层: 中间层是逻辑层, 逻辑层是从实际应用中抽象出来的完整产品模型, 但与具体实现无关, 该层包括各种集成资源;

物理层: 底层是物理层, 它包括多种实施方法, 赋予文件在计算机上的具体实现形式。

## 3.4 QT 开发框架

### 3.4.1 Qt 简介

Qt 是在 1991 年由 Qt 公司 (当时为 Trolltech, 后被 Nokia 收购, 现为 Digia 的一部分) 开发的一个跨平台 C++ 图形用户界面应用程序开发框架。Qt 既可以用于图形用户界面 GUI 程序的开发, 也可以用于非 GUI 程序的开发, 如控制台工具和服务器。它能够全面支持 C++ 应用程序, 高效地进行开发, 实现数据显示与业务逻辑处理的分离, 并且具有良好的软件可移植性。Qt 使用面向对象的框架, 其中包括特殊的代码生成扩展, 即宏, 还包括一个元对象编译器 (moc) 。Qt 极其容易扩展, 是真正的组件编程。

Qt 可以很好地支持网络编程、数据库编程、多线程编程, 并为读者提供了强大的在线帮助文档以供查询。此外, Qt 还具备国际化的考量, 通过使用 Unicode 编码、使用 Linguist 进行相关翻译等方式, 实现应用程序的翻译目的。你也可以使用 OpenGL 进行 3D 图形可视化。

### 3.4.2 Qt 开发的优势

如果您熟悉 Windows 平台的软件开发,尤其是图形用户界面方面的开发,您必然对 MFC、VCL、ATL、Openwin 等不陌生。Qt 在功能上基本相当于这些框架,而且 Qt 还有几个显著的优势:出色的跨平台特性、丰富的 API、面向对象设计、对图形渲染的支持、对 XML 文件的支持以及大量的开发文档。具体来说,有以下几个优点:

Qt 具有良好的跨平台特性:你可以将相同的 Qt 代码直接放在不同的编译平台下进行编译并直接运行,能够构建不同类型的应用程序。Qt 支持不同的操作系统平台,包括:Windows 98、Windows 2000、Windows XP、Windows Vista 和 Windows 7、Windows NT、Unix/X11-Linux、Sun Solaris、SGI IRIX、IBM AIX 等。

面向对象:Qt 在 C++中使用了良好的封装机制,使得自身非常模块化、可重用,并且便于用户开发各种应用程序。除此之外,Qt 与其他应用程序框架不同的地方在于 Qt 提供了信号/槽机制。

安全类型 (SIGNAL/SLOT) 替代了传统的回调函数,这使得各种元素之间的协同工作变得非常简单。

丰富的 API:Qt 包含 250 多个 C++类,一方面提供基于模板的集合、文件、I/O 设备、目录管理、日期/时间类,另一方面还能处理正则表达式。

通过支持 OpenGL,实现了 2D 和 3D 可视化的渲染。除此之外,还有大量的开发帮助文档,支持 XML 等。

### 3.4.3 Qt 类库

Qt 发布了构建高级跨平台应用程序所需的所有类库,因为 Qt 的 C++类库模块提供了丰富的代码块集合,用于构建各种应用程序。因此,Qt 的类库总共包含以下类别:

1) 强大的图形用户界面 (Advanced GUI):Qt 提供了开发嵌入式和 PC 系统应用程序所需的功能。Qt 图形用户界面的主要特点如下:

具有完整多样的表单控件,例如:表格、对话框、按钮、标签、视图等。

在实际编程中,Qt 可以智能地进行缩放;

具备完全自定义用户界面的能力;

2) 多线程功能 (Multithreading):众所周知,多线程编程是一种资源密集型操作,在

同时操作多个 UI 界面时，不会冻结用户界面，然后充分利用系统资源。Qt 的固有多线程特性可以充分降低并行编程的难度，其同步功能也使得开发者更容易地使用多核架构。

3) 基于 OpenGL 的图形：由于 OpenGL 用于开发可在多个平台上执行的应用程序，Qt 能够进行跨平台移植。虽然 OpenGL 完美支持 3D 图形，但存在一个问题：OpenGL 无法创建应用程序的用户界面。现在，Qt+OpenGL 的组合完美地解决了这个问题。Qt 使得在应用程序中使用 3D 图形变得容易，并且还可以利用系统资源来最大化性能。

4) 对象间通信：在开发图形用户界面 UI 时，频繁的系统崩溃的根本原因是不同组件之间如何进行通信。Qt 对此的解决方案是使用信号/槽机制。

(Signal/Slot) 机制。例如，当用户点击某个动作时，将会发出一个信号，然后，相应的槽函数会对这个信号作出响应，系统将执行相应的动作。

5) 支持网络连接：Qt 提供了相对简单的跨平台网络连接，同时支持多种网络通信协议。您还可以访问各种格式的 HTML、媒体文件和 XML 的数据。

6) 支持 XML：前文谈到支持 XML，这是因为 Qt 包括了一个类似于 SQL 的 XQuery 查询语言，用于实现 XML 文档的查询。XQuery 执行少量代码即可实现相对高性能的 XML 查询，并且可以将读取在 XML 中的数据转换到应用程序中。

7) 支持数据库：其强大的包容性能完美地将应用程序与数据库集成。Qt 支持主要的数据库驱动程序，如 SQLite、ODBC、DB2、MySQL、Oracle、Sybase 等，并且能够以各种方式呈现它读取的数据。

### 3.5 Open CASCADE 技术简介

Open CASCADE 是由法国 Matra Datavision 公司开发的一个开源几何内核库，它是一个面向对象的 C++ 类库，专为需要快速开发程序的特定领域设计。OCCT 可用于开发 3D 和 2D 通用几何计算机辅助设计 CAD 系统、计算机辅助制造 CAM 系统以及计算机辅助工程分析 CAA 系统。强大的 OCCT 库能够加快程序开发的速度。

Open CASCADE 由六个模块组成，分别是建模数据模块 (Modeling Data Module)、基础类模块 (Basic Class Module)、可视化模块 (Visualization Module)、建模算法模块 (Modeling Algorithm Module)、应用框架模块 (Application Framework Module) 和数据交换模块 (Data Exchange Module)。建模数据模块也被称为数据结构模块 (Data Structures Module)，因为它主要提供 3D 和 2D 几何的数据结构。一个模块可以由一个或几个工具箱 (toolbox) 组成，一个工具箱在结构上等同于一个共享库。多个类构成一个包 (package)，一个或多个包构

成一个工具箱。由于分工不同，每个模块负责不同的功能。

在主流的几何系统中，数据结构是由形状驱动的。但在 Open CASCADE 主导的系统中，情况完全不同，其原则是由引用键的使用驱动的。Open CASCADE 数据框架的原则是：引用键以标签的形式实现，所有数据，包括几何数据、形状数据等，都通过标签的树状结构来保存。所有属性都具有相同的等级。

具体来说，数据保存在一个称为 TDF\_Label 的数据结构中。TDF\_Label 代表一个地址或表示方式，可以存储的数据是属性 Attribute，它包括整数、实数、字符串等。对于每个 Label，同一类的 Attribute 只能有一个。也就是说，如果有多个整数，则需要一个 IntegerArray 属性，而不是单个的整数属性。

上述文章主要描述了在过程装配系统 VPAS 中的相关技术、工具、开发环境和算法。首先介绍了整体叶片的知识，然后介绍了 VPAS 中使用的各种工具和输入文件，介绍了 Qt 开发环境中的模块和 Open CASCADE 的特点，最后介绍了曲面交叉算法和 3D 建模方法，为后续的可视化方面奠定了基础。

# 第 4 节 术语推荐与基于语义相似度计算的可视化

## 1. 研究背景与意义

随着互联网的快速发展和大数据时代的到来，世界正逐渐向智能化和多样化方向发展。互联网的普及和信息的共享使世界变成了一个整体。新兴事物不断涌现，数据量持续增加，在“互联网+”的背景下，传统的互联网知识组织和知识分享方法越来越难以满足人们的需求，如何使用这些数据的组织，已成为一个研究热点。语义网、智能搜索、自然语言处理、人机对话、机器翻译、智能机器人等的快速发展，都预示着互联网正在经历一场革命——语义时代。与文本处理相关的技术在人机交流的“语义时代”扮演着重要的角色。

根据不完全统计，当前互联网中，预计 2015 年大约 80% 的网络数据以非结构化文本的形式存在，且有 44.6% 的网络用户表示网络信息的主要问题是“信息重复太多”。因此，非结构化数据的语义相似度计算方法对于优化搜索引擎和处理网页语义至关重要。研究术语语义相似度计算方法在发展语义时代中扮演着重要角色。过去，大多数计算方法都是基于结构化数据的形式，通过结构化特征来测量语义相似度，这在实践中往往有很多不足。基于非结构化数据文本计算术语的语义相似度随着语料库规模的增加也增加了很多计算难度。特别是面对数据的指数级增长，如何有效处理和探索新术语及术语之间的关系，已成为术语学者的一个难点。因此，提高相似度计算方法的性能可以减轻术语学者的工作压力，有效探索术语之间的关系，并为自然语言处理多任务打好基础；它还可以提高机器翻译的准确度，提升人机交互的程度，并有助于自动问答系统、自动文摘和语义搜索引擎的发展。

信息数据的快速增长也导致了新术语的出现，术语的提取和确定正面临着巨大的困难。在不同的国家或语言中，同一术语的表达是不同的，这就需要各国的术语专家进行推荐和确定。对于 ISO 文件中的术语，更需要国家专家进行确定后才能编写。为了解决各国专家交流困难的问题，我们开发了术语推荐系统，为来自不同国家的术语学家提供一个交流和确认提取术语的沟通平台。已经通过术语推荐系统确认的术语可以编写到 ISO 文件中。但是随着网络的发展，即使一些新术语出现，但在语义上也是同义词或近义词，对于这类术语，我们需要在早期进行语义判断，排除相似或近似的术语，以使 ISO 术语文件更加完整和科学。因此，

本文提出了一种改进的混合语义相似度算法,对需要推荐的术语进行语义判断,排除同义词或近义词,并通过术语推荐系统推荐术语。

知识图谱是一种新型的学科领域,是科学、信息等微学科多学科发展成果的新发展,能够科学、直观地挖掘科学知识及其相互联系,促进科学知识共享以及科学知识和技术的深入研究和探索。对于术语学者来说,他们更倾向于将相似术语进行可视化,这能直观反映术语之间的隐含关系以及与相似意义的术语之间的关联,也为绘制知识图谱奠定了基础。因此,对术语的语义相似词进行可视化研究在信息挖掘和知识图谱绘制等应用中具有重要意义。针对这些问题,本文提出了一种基于词向量的语义相似度计算方法,用于基于大规模非结构化数据搜索和过滤术语相似性,然后将其进行可视化显示。

## 2. 国内外研究现状

术语语义相似度计算方法在多个研究领域都有广泛的应用,例如:网络语义学、知识工程、机器翻译、人机交互等领域,而术语标准的制定与这些领域的发展直接相关。最早的术语语义相似度计算方法是 Manber 提出的 Sif 方法,其思想是通过字形匹配来纯粹地计算相似度; Garcia-Molina 等人提出了基于向量空间模型的 SCAM 原型,其通过计算向量空间的距离来衡量术语的语义相似度;香港理工大学的 Si 和 Leong 利用文档中关键词计数的方法来衡量术语的相似度,并提出了 CHECK 原型。随着研究的进展,结构化数据概念之间的语义相似度成为了研究的焦点。结构化数据表示可以是本体、XML、框架、WordNet 和 HowNet 等,基于本体结构的语义相似度计算方法主要有以下成果。

基于本体距离的语义相似度计算方法:基本思想是通过计算本体中两个概念词路径的长度来表示它们之间的语义距离。典型的代表性算法有最短路径算法 (Shorted Path algorithm)、加权链接算法 (Weighted Links algorithm)、Wu 和 Palmer 算法、Leacock 和 Chodorow 算法等。尽管这类算法具有最小的计算复杂度并且易于实现,但其主要缺点是,使用这种方法的前提是假设本体分类系统中的每条边在计算过程中都处于同等重要的位置,这显然是不现实的。对于复杂本体结构的语义相似度计算, J.M. Kim 等人提出了 CP/CV 概念传播方法。

基于本体内容的语义相似度计算方法:这种方法的基本思想是通过父节点的信息来衡量两个节点的语义相似度,父节点包含的信息越多,其下的两个子节点就越相似;反之,相似度越小。这种方法通常考虑的是概念在本体中的共享信息量,比如它们共同的属性、关系或者在本体中的共同父类。通过比较这些共享信息的丰度,可以对概念之间的语义相似度进行定量的评估。

基于本体属性的语义相似度计算方法：不同的事物具有不同的属性和特征，具体的事物有更多的属性，抽象的事物则属性和特征较少。因此，可以通过计算两个概念节点之间的公共属性数量来衡量节点间的语义相似度，即共有属性越多，节点的语义相似度越大。反之，共有属性越少，语义相似度也越小。这种方法认为，共有属性的数量能够反映两个概念在语义上的接近程度。

混合语义相似度计算方法：它是一种综合性方法，全面考虑了上述三种方法，包括位置距离、信息内容、概念词属性等多种因素的关系。代表性的算法包括：Li 方法以及 Marco 等人提出的 SSA 算法。这种方法旨在通过结合不同的计算维度来提高语义相似度计算的准确性和鲁棒性，从而更全面地捕捉概念之间的语义关联。

近年来，中国在计算语义相似度方法的研究也取得了一定成果。邓志宏等人在第 18 届全国数据库学术会议上将本体技术引入到 Web 检索模型中，同时扩展其功能语义，以解决 Web 信息处理中的语义缺失问题。万洁等人开发了一种使用本体知识技术原型的信息检索系统。该系统基于本体的语义相似度计算，语义上扩展用户的查询意图，使得检索结果更加精确，减少了冗余。基于本体的语义表达能力和相似度计算，王存刚等人提出了 MIRS 模型，深入挖掘用户查询请求的语义信息，使其隐含的信息得到表达。王兰成等人使用语义相似度计算方法，将检索请求与相似节点概念词语义上扩展，完成检索文档的过滤，使最终的搜索结果更加符合用户的查询请求。曹天等人基于 VSM 和词共有特征的术语语义相似度算法，正确率和平均召回率都有了很大的提高。基于属性理论，潘倩红等人在术语的属性上建立了坐标系统，并通过属性坐标系统表示术语向量，通过计算向量之间的匹配距离来衡量术语的语义相似值。

### **3. 计算相关语义相似度的方法与原理介绍**

数据结构对象可以分为结构化和非结构化两种，而当前大多数的语义相似度计算方法都是基于结构化文本或文件。这些计算方法大多基于这些结构化数据；随着互联网的发展，数据呈指数级增长，非结构化数据的语义相似度计算方法不断更新，逐渐成为一个研究热点。例如，使用机器学习的语义相似度算法。本章主要描述了文献[8]中涉及的计算方法和原理。

#### **3.1 基于本体距离的语义相似度计算方法**

基于本体距离的语义相似度计算方法的基本思想是计算本体中两个节点概念词之间的最短路径长度。路径越长，节点之间的语义距离越大，节点之间的相似度越低；相反，路径

越短，其节点之间的语义距离越小，相似度越大。不同的方法在改进点上各有不同，但基本理念都是集中于计算其最近距离。

### 3.2 基于本体内容的语义相似度计算方法

基于本体内容的语义相似度计算的思想是通过两个节点概念词共享的概念信息量来衡量它们的相似度。如果两个节点概念词有更多的共同信息内容，这两个概念节点词的语义相似度就更高。相反，则更小。在本体中，每个子节点都是其父节点的具体化，子节点的信息比父节点更详细、全面，父节点代表抽象信息，但子节点包含在父概念节点中。因此，通过计算两个子节点最近共同父节点所包含的信息来表达两个子节点的语义相似度。

### 3.3 基于本体属性的语义相似度计算方法

基于本体属性的语义相似度计算方法的基本概念是根据两个节点共有属性的数量来衡量相似度。不同的事物可以根据它们自身的属性特征进行区分。这种方法仅考虑其自身的属性集合，并不考虑节点在本体中的位置信息以及它所包含的内容信息，因此它能充分利用本体的属性集合进行操作。

### 3.4 混合语义相似度计算方法

混合语义相似度计算方法的基本思想是综合考虑节点之间的距离、节点的内容信息、节点的属性以及其他因素，全面计算语义相似度。这种方法能够全面反映节点词的语义相似性。

### 3.5 Word2vec 简介

Google 在 2013 年开源的 Word2vec 工具能将单词转换成实数向量，在自然语言处理(NLP)领域获得了极大的青睐。通常认为 Word2vec 是一种深度学习模型，它可以通过对大规模语料库的学习建立词向量模型，并通过词向量模型来计算非结构化语料库中单词的语义相似度，以发现同义词或相似词。

词向量表示目前包括两种，一种是 One-hot 表示法，这种方法通过对词汇进行顺序编号，每个词汇是一个非常长的向量，向量的维度是词典的大小，只有对应编号的位置是 1，其余都是 0。例如：“手机”被表示为 $[0, 0, \dots, 1, \dots, 0]$ ，“电脑”被表示为 $[0, 1, \dots, 0, \dots, 0]$ 。但这种方法有两个缺点：维度灾难和词汇鸿沟。如果文档的词汇量太大或者语料库很大，那么构

成词向量的维度将难以估计；这种表示不反映单词之间的词汇边界和单词之间的联系。另一种词向量表示是分布式表示法。它将每个单词映射到一个 K 维的实数向量中，通过计算单词之间的距离来衡量语义相似度。如果可以表示为[0.792, -0.177, -0.107, -0.109, 0.542...]等。

### 3.6 Skip-gram 原理

Skip-gram 语言模型是 Word2vec 中的一个重要模型。与 CBOW 模型不同，CBOW 是通过上下文预测目标词的过程，而 Skip-gram 模型恰恰相反，它是通过目标词来预测其上下文的过程。

### 3.7 Prefuse 简介

Prefuse 是一个可扩展的软件框架，它基于 Java 开发，可以很好地开发信息可视化交互程序，也可以用来独立构建程序。Prefuse 支持由表格、图形、树组成的数据结构，现有组件可以帮助完成布局、颜色、形状等功能；它包含了许多与用户交互的功能，能够实现动态动画，改变可视化的变形效果，以及位置大小和语义的变化。更新数据可以动态显示，支持将数据库查询结果集与 prefuse 内部信息数据的关系映射，并使用良好的 API 进行定制。

## 4. 改进与应用混合语义相似度计算方法

### 4.1 方法论概述

改进的混合语义相似度计算方法基于模糊优化的理念，通过科学计算导出权重，消除了基于专家经验判定的主观误差。该方法使用分词技术对输入信息和文本进行分割，然后使用语义扩展搜索内容进行语义扩充，最后利用改进的混合语义相似度计算方法查询和计算资源库。

### 4.2 相关技术介绍

分词技术是文本挖掘和语义分析的关键技术，也是搜索引擎的必要技术。对于用户输入的搜索信息，必须对输入的信息进行分词处理才能达成目的。英文分词相对简单，因为英文单词之间有空格分隔，但中文之间没有空格这样的标点分隔，中文的最小单位是词，而一个词有时可以代表一个意义，所以中文分词相对困难。例如，在“我喜欢学习计算机”中，“喜欢”、“学习”、“计算机”可以构成一个有独立意义的句子。通过计算机，我们必须对其

进行分析和划分，总结分割后的词汇的语义信息，然后进行组织和整合，得到具体和合理的语义，这需要一定的分词技术来完成。中文分词技术实际上是基于一种合理有效的方法对句子中的词汇进行拆分，并在词与词之间添加标记，例如“我非常喜欢学习计算机”，合理的划分是：“我 / 非常 / 喜欢 / 学习 / 计算机”。分词技术目前广泛应用于文本分类、人工智能、自动标注、知识工程等领域。同时，也有许多主流的成熟分词软件，例如：基于 Python 语言的结巴分词；NLPIR 中文分词软件；基于 HTTP 协议的开源中文分词 HTTPCWS；SCWS 简中分词系统；Phpanalysis 分词系统；基于 Java 的开源中文分词组件 MMSEG4J；盘古分词等。

### 4.3 混合语义相似度计算方法的改进

模糊优化理论和方法起源于 20 世纪 70 年代，模糊优化的焦点是将不确定的模糊优化信息转化为确定的优化信息。对于给定的不确定信息，可以获得一个确定的边界信息，然后得到一个确定的结果。本文主要借鉴模糊优化排序的思想，科学确定混合语义相似度计算方法的权重，避免专家的主观错误，使计算结果更加可靠和准确。

### 4.4 基于混合语义相似度计算方法的术语推荐系统

#### 4.4.1 系统开发背景

在现实生活中，数据量呈指数型增长，新术语不断产生，现有术语的定义也需要不断更新，因此国际 ISO 组织的术语标准文件需要不断更新和补充。术语标准文件的统一制定对推动术语学的发展具有重大作用，也能促进世界一体化的协调发展，同时还能推动自然语言处理、人工智能、机器翻译等学科的快速进步。然而，庞大的数据量给术语学家带来了巨大的困难，特别是在存在语言障碍、文化障碍和在术语统一标准制定上有很大分歧的不同国家之间。鉴于此，开发了 this 术语推荐系统，以提供一个供各国术语专家交流的平台。在这个平台上，每个术语学家都可以提出术语推荐和定义更新，并且还可以综合其他专家提出的术语意见并表达自己的观点。同时，系统还具有翻译功能，使来自不同国家的术语学家能够交流术语。

然而，随着数据量的增加，语义上新出现的术语可能与原有术语相似，如果术语学家推荐的术语在术语文件中有同义词或相似词，这些术语就可以被淘汰和减少。因此，在推荐的

早期阶段,有必要计算术语的语义相似度,并搜索标准术语文件中是否存在具有同义或相似意义的术语。在术语语义相似度计算方法方面,系统采用了一种改进的混合语义相似度计算方法,所有提交到术语推荐系统的术语都需要事先使用它来进行语义判断,然后再进行审查和推荐。

#### 4.4.2 系统开发环境

术语推荐系统是在 Windows 7、500G 硬盘、4G 内存、2G 独显的硬件环境下开发的,开发语言基于 PHP + Mysql,使用 Tiki 组件进行构建和管理。

#### 4.4.3 系统功能浏览

该系统主要针对专家群体开发,因此用户群体仅限于注册用户和非注册用户,注册用户通常是专家或超级专家等,不同用户拥有不同的权限和功能。系统仅推荐使用改进的混合语义相似度计算方法来处理术语,对于 ISO 术语标准文件中已有相似术语的情况,将不会进行推荐。

#### 4.4.4 系统实现

由于系统开发的目的是为国际 ISO 术语专家提供一个交流平台,因此整个页面采用英文显示。本文中的术语语义相似度算法旨在辅助系统工作,所以系统的开发过程不是重点,文本主要通过页面进行实际展示。

## 5. 大规模非结构化术语的语义相似度计算与可视化

智能化和信息化时代带来了数据的指数级增长,不同领域中的新术语不断涌现。术语更新工作需要跟上时代的步伐,特别是 ISO 国际术语标准文件的制定和更新需要不断进步,但面对海量数据,也给术语学家带来了巨大的困难。术语之间的关系难以挖掘,特别是当数据量庞大时,使用有效方法进行术语关系挖掘更是困难重重。因此,术语关系的可视化逐渐成为一个研究热点。另一个问题是,如何从大量的非结构化数据中找到术语的相似性?过去,大多数术语相似度计算方法使用结构化数据文本来计算术语的语义相似度。在本文中,我们通过使用词向量来计算术语的语义相似度,并搜索术语的相似性。这种方法主要利用

Word2vec 进行建模训练，将语料库中的单词转换为词向量，并通过调用词向量来计算术语的语义相似度，从而消除了结构化数据的一些缺点，并实现了大规模非结构化数据的语义相似度计算。

通过计算词向量中术语的语义相似度，得到术语的相似词，使用 Prefuse 组件中的 GraphTreeView 模型来可视化这些相似词。可视化显示可以方便地找到一个术语衍生出的其他语义相似词，这样术语学家可以通过可视化显示观察术语之间的关系，为术语学家提供参考价值，同时也为后期绘制知识图谱奠定了基础。

# 第5节 基于计算机视觉的增强现实场 景生成与可视化

## 1. 研究背景和意义

中国拥有五千年深厚的历史文化，如指南针等“四大发明”、陶瓷、青铜器、长城等在人类生活的许多领域都取得了杰出的成就，对人类文明的发展作出了重要贡献。然而，随着时间的推移，这些珍贵的文化遗迹、古建筑等长期未能得到强有力的保护，受到人为或自然条件的破坏，每年都以一定的速度从地球上消失，它们是中国历史的直接见证，是世界文明史上不可或缺的成员，其保护已成为当今亟待解决的紧迫问题。在当今社会发展的科技进步中，增强现实技术在图形技术、显示技术、交互技术、传感技术等方面应运而生，并已广泛应用于教育、娱乐、医疗、古迹修复等领域，在这样的背景下，利用增强现实技术对祖先留下的宝贵财富进行数字化，不仅可以减少现实条件下对它们的损害，而且还可以长期保存这些珍贵的遗产。

增强现实指的是利用计算机图形学、多媒体、人机交互、立体显示等技术，使得在现实场景中不存在的信息与现实场景中的真实信息共同存在于同一个场景之中。具体来说，增强现实是在现存的真实场景的基础上，提供虚拟信息来增强场景，将虚拟信息叠加在真实场景的适当位置，给用户带来全新的视觉体验，体验者完全感觉到虚拟信息是真实场景信息的一部分，并使体验者在场景中有一定的沉浸感。

增强现实场景的生成和可视化是增强现实技术中一个极具挑战性的课题，它在工业、军事、娱乐、医疗、教学、古建筑修复等领域的应用不容小觑。圆明园的数字再现和2010年上海世界博览会上震惊中外的青花瓷碗的数字复制，是该技术在国内领域的重大应用。增强现实场景是将计算机生成的虚拟对象与真实场景融合的场景。可视化（Visualization）是指利用计算机图形学和图像处理技术，将计算得到的数据转化为通过显示设备展示的图形或图像，并进行其他操作的过程。因此，对于一个增强现实系统而言，其场景的生成和可视化是不可或缺且相辅相成的两个部分。

增强现实场景的一个关键特征是虚拟对象和真实场景的精确对齐，这是增强现实系统中

的关键技术，即注册定位技术的核心目的。在方法上，注册定位的基本方法可以分为两类：基于硬件的跟踪和基于视觉算法分析。由硬件设备实现的跟踪注册操作相对较快，但成本较高，安装配置复杂，有时精度也不一定满足要求。相比之下，基于计算机视觉的注册定位技术仅需要一两个普通的数字相机，具有成本低、适应性强的特点。计算机视觉方法指的是计算机对场景图像的识别及其转换为数字信息的过程。

了解图形几何形态，以获取图像场景中的结构信息，这些信息可以被处理、存储和理解。综合考虑基于硬件的跟踪方法和基于视觉算法分析方法的优缺点，为了降低系统构建成本，本项目采用箱型场景模型代替古建筑内部环境模型，使用计算机视觉的三维注册方法构建增强现实系统。通过图像采集与预处理子系统收集并预处理真实的三维场景，通过相机标定子系统获得三维注册信息，通过三维注册子系统添加虚拟对象，增强真实场景信息，生成增强现实场景，最后通过三维显示子系统实时显示恢复后的三维场景，为最终实现古建筑室内三维场景展示奠定技术研究基础。

## 2. 国内外增强现实技术的研究现状

增强现实（AR）技术起源于1960年代，当时哈佛大学的教授伊凡·萨瑟兰发明了透视图头戴显示器（STHMD），用于展示计算机生成的3D图形。然而，在接下来的20年左右时间里，由于受限于当时的硬件设备条件和图形研究水平，AR技术并没有取得显著的进展，甚至连AR的概念都不是非常清晰。

20世纪末，AR系统发展相对较快，各种增强现实系统层出不穷，应用领域多样，主要包括：

1986年，Furness开发了VCASS系统，该系统利用头盔式显示器在飞行员的视野中显示操作信息，如距离和设计目标。

1986年，北卡罗来纳大学教堂山分校开发了STHMD系统，用于生物化学和建筑可视化。

1993年，美国哥伦比亚大学的Feiner教授等人设计了一个基于知识的增强现实系统。该系统可以将相关的技术指导信息叠加到激光打印机上，维修人员可以通过将信息叠加在自己的视线中成功完成维修工作。

20世纪90年代初，波音公司的汤姆·考德尔（Tom Caudell）和他的同事在他们设计的辅助布线系统中首次创造了“增强现实”（Augment Reality, AR）这一术语。在这个系统中，由简单线条和文本提示信息组成的布线路径实时显示在佩戴立体头盔显示器（S-HMD）的工作人员的视野中，这可以帮助从事该领域工作的人员逐步完成打包或拆卸过程，从而减少

出错的机会。后来，出现了许多种增强现实应用系统，主要集中在医疗、制造与维护、机器人行动路径规划、娱乐和军事等方面，但由于系统自身的原因，这些研究的系统并没有发挥其应有的使用价值，也未被投入生产实践。

随着研究的深入，国际上出现了如 IWAR、ISAR、ISMAR、ISWC、VRAIS、ICAT、IEEE VR 等重要的国际会议，为研究者提供了充分交流的空间。由于这是一个跨学科的研究领域，不同学科之间的交流也在增加，吸引了越来越多的研究者参与合作[9]。

国内对 AR 的研究起步相对较晚，增强现实系统的研究主要集中在高校和科研单位，目前关于增强现实系统的研究主要包括浙江大学 CAD&CG 国家重点实验室、北京理工大学、国防科技大学、中国科学院自动化研究所、华中科技大学、上海大学、同济大学、西安交通大学、西安电子科技大学等。

增强现实具有虚拟和真实结合、三维注册、实时交互的特点，因此被广泛应用于各个领域，主要应用于以下几个方面：

在医疗应用中，AR 系统可以作为手术辅助工具，麻省理工学院的人工智能实验室使用 AR 技术将核磁共振成像 (MRI) 或 CT 扫描的三维图像叠加在病人的身体部位上，通过头盔显示器 (HMDs) 等显示设备为医生提供“透视之眼”的功能。AR 系统还可以通过图像或文本信息为医生提供精确的手术导航，医生可以根据提示逐步完成手术操作，这可以大大提高手术的成功率。

在军事领域，AR 系统通过视觉图像或文本信息为士兵提供全方位的战场增强信息，包括地理位置信息、目标、友军邻位以及支援等，这可以极大地提高士兵的战斗效能。

在考古领域，IntraCom 在欧盟的支持下，开发了一个利用 AR 技术对古希腊遗迹进行现场恢复的电子导游系统 (ARCHEOGUIDE)。该系统由计算机服务器、无线网络和个人数据助理 (PDA) 客户端组成，旨在让游客能够在现场看到古迹恢复后的效果。

注册定位与显示是增强现实系统中两项最关键的技术。目前，国际上使用的注册方法主要包括基于跟踪设备的注册定位、基于计算机视觉的注册定位以及基于视觉与跟踪设备相结合的混合注册定位。

基于跟踪设备的注册定位是通过一些硬件传感器来获取摄像机的方位参数，目前常用的传感器包括：磁场传感器、惯性传感器、超声波传感器、光学传感器和机械传感器。例如，磁场传感器，磁场感应跟踪系统使用线圈作为信号发生器，根据磁发射信号与磁感应信号之间的耦合关系来确定被测物体的空间方向，通过测量接收器的磁通量可以确定接收器与信号源之间的相对距离。机械传感器通过机械关节的物理连接来测量运动摄像机的位置和方向。

结果准确直观，但设备复杂，不适合典型应用。

在基于计算机视觉的注册方法中，目前主要有两种类型的方法：基于标识点的注册和基于自然特征的注册。基于标识点的注册方法主要是在场景中放置某种标记，通过提取标记的信息来进行注册。基于自然特征的注册方法的基本原理是通过提取一些容易提取且具有仿射不变性（如 Sift, Surf, Fern 等）的特征，并在二维图像中找到这些特征与三维世界中的帧坐标的对应关系来进行注册。与硬件跟踪相比，计算机视觉的注册方法具有使用范围广、场景成本低的特点，但由于其大量的计算，某种程度上导致了跟踪的延迟，并且会出现漂移。

虽然在过去的几十年中增强现实技术得到了很大的发展，但仍有许多问题限制了增强现实技术的推广。首先是定位问题，特别是 AR 系统的定位已成为决定系统成功与否的关键。系统的显示问题也不容忽视，大多数现有的 AR 显示设备都有亮度和对比度低、分辨率低和视场窄等缺点。同时，AR 系统的高设备成本也在一定程度上限制了 AR 技术的发展和普及。因此，基于特征的计算机视觉注册方法和显示技术的研究仍然是这一领域的研究热点。

### **3. 增强现实系统的关键技术**

#### **3.1 相机内部参数标定**

在相机姿态估计过程中，需要预先知道相机的内部参数，包括相机的焦距、相机主点的坐标等，而相机内部参数的确定也称为相机标定。针对相机标定已经出现了许多方法，但这里主要描述以下两种方法。

(1) Kruppa 方程法: Kruppa 方程法是通过获取相机拍摄得到的图像来进行相机内部参数标定的方法，并且无需预设标定块来了解场景中的几何模型。通过相机获取不同视点的图片之间的映射关系，并计算两幅图像之间的极点几何关系，以获得 DIAC 的约束方程。每个 Kruppa 方程只对 DIAC 的 6 个未知参数提供两个独立约束，因此给定三幅图像以及每对视图之间的基础矩阵，假设相机在运动过程中内部参照保持不变，就可以确定相机的内部参照矩阵。

(2) MATLAB 工具箱法: MATLAB 工具箱中的相机标定方法是由 Klaus Strobl 等人开发的，它简化并改进了张正友等人的标定方法，也是本文采用的相机内部参考矩阵的标定方法。该方法使用预设的棋盘格标定板进行标定。MATLAB 工具箱有两种图像设置的形式，当图片小时可以使用标准界面，当图片相对较大时，可以使用内存高效方式，逐一下载图片进行角点检测。

## 3.2 跟踪与定位技术

相机的跟踪与定位指的是在虚拟与现实场景融合的过程中,对相机的运动姿态进行实时检测,改变虚拟相机的位置以便在场景的特定位置添加虚拟对象信息。目前,常用的跟踪与定位技术主要有基于硬件和视觉的两种跟踪定位方法。

基于硬件的跟踪与定位技术主要利用硬件设备的特征,通过分析接收到的信号或硬件设备本身的物理连接来完成移动过程中的位置和方向估计。

基于计算机视觉的注册方法主要是基于相机拍摄的图像,使用图像之间的映射关系,分析图像信息来估计相机移动,最终实现虚拟信息与真实场景的融合。基于计算机视觉的注册方法不需要特别昂贵的硬件设备作为注册的基本条件,只需使用相机拍摄的图片作为识别对象,与基于硬件的方法相比有效降低了注册成本。基于计算机视觉的注册分为基于标记点的注册和基于自然特征点的注册。

## 3.3 显示技术

立体显示技术是增强现实系统中一个非常关键的技术。增强现实系统可以使用多种显示设备来显示叠加的混合场景,目前常用的显示设备有传递式头盔显示器、投影设备、移动设备 PDAPC 显示等。头盔显示器有视频透视头盔显示器 (Video see-through, 简称 VST) 和光学透视头盔显示器 (Optical see-through, 简称 OST) 两种, VST 可以使用图像处理方法来校正灰度和颜色, OST 结构简单且便于携带。为了使 OST 能够产生正确的遮挡效果, Bimber 和 Frohlicht 提出了一种投影照明技术。其中,代表性的投影显示方法包括直接投影显示模式、分光器准直显示模式、后投影准直显示模式等。这些显示方法占地面积大,主要用于大中型飞行模拟器。

为了降低系统显示的成本,该项目使用 OPENGL 3D 图形软件通过 PC 的普通显示设备来呈现融合的虚拟现实场景。

## 4. 混合特征提取与匹配

标志板现在被用来覆盖那些由于时间的流逝而自然特征不明显的区域,或者由于光照等外部条件的影响而无法提取自然特征的区域。由于标志板放置在原有自然特征的部分,此时标志板的位置也就是场景中自然特征的位置,标志点的坐标也就是三维空间中自然特征的坐标。因此,在特征提取过程中,图像中标志点的像素值替代了未放置标志板时无法提取的自

然特征点的像素值,在图像中未放置标志板的地方,仍然提取原有三维环境空间的自然特征,最后将自然特征和标志特征混合到同一空间中[10]。

## 4.1 混合特征提取算法

获取的原始图像不能被计算机直接处理,必须先转换成由像素块组成的二维矩阵形式的数字图像,计算机才能对其进行灰度、二值化等处理。真实场景中的自然特征包括点特征、线特征、轮廓特征和颜色等。由于具有较强的抗遮挡能力,点特征已经成为增强现实跟踪和定位算法中更常用的特征。

## 4.2 混合特征匹配算法

当距离小于某个阈值时,认为两个视图中的关键点是匹配的,以此找到两个视图中对应点的二维像素坐标,为增强现实系统使用计算机视觉方法估计相机运动打下基础,进而计算相机的旋转矩阵和平移向量,找出相机的投影矩阵,并还原真实场景中的三维点坐标。

# 5. 相机姿态估计的基础知识

从第二章的系统框架来看,当第三章的特征提取完成后,所实现的是根据特征提取结果来估计相机的位置。相机姿态估计是指计算相机的旋转、平移以及相机运动的方向,涉及相机的针孔模型、射影几何学、基础矩阵等数学知识,因此,在分析相机姿态估计算法之前,有必要先介绍这些基础理论。

## 6. 两种相机姿态估计算法

相机运动状态的估计是注册过程的一个非常关键的部分。目前,对于计算相机基础矩阵,改进的八点算法在相机视野中出现大量平面场景的情况下会出现退化。为了扩展算法的使用范围,提高算法的稳定性,提出使用五点算法替代改进的八点算法,以实现相机姿态估计。这不仅在使用过程中增强了算法的自由度,还提高了注册的准确性和稳定性。

### 6.1 八点算法

在介绍改进的八点算法之前,需要了解空间中对应的两幅图像之间存在一定的几何关系,这种关系可以通过基础矩阵来唯一确定。根据极线几何学的知识,可以通过两幅图像中相对应的匹配点信息来确定这个基础矩阵。

## 6.2 改进的八点算法

由于八点算法对噪声等因素敏感, Hartley 等人提出了改进的八点算法。在获得关键点的二维匹配点之后, 首先对关键点的坐标进行变换, 将所有匹配点的质心平移到图像的中心。

### 6.3 五点算法

根据八点算法, 所谓的五点算法是指用图像中的 5 个对应点计算基础矩阵的方法, 区别在于使用八点算法计算基础矩阵时, 不需要预先知道相机的内部参数, 但是五点算法在计算基础矩阵时需要预先知道相机的内部参数信息。相机内部参数的准确性是估计相机姿态的关键, 因此在这个实验中, 使用了 MATLAB 工具箱, 用相机拍摄了 25 张棋盘格照片进行校准, 最终获得了相机的内部参数矩阵  $K$ 。

## 参考文献

- [1] 曲思遥.(2020). 国内动画电影产业信息的视觉设计探索. 湖北工业大学.
- [2] 赵萌.(2019). 基于混合密度网络的音乐编舞算法研究. 北京邮电大学.
- [3] 孙佳丽.(2019). 基于模型驱动的电视剧本创作辅助工具软件的研究与开发. 武汉理工大学.
- [4] 王鑫.(2019). 基于 Processing 的计算机生成艺术应用研究. 艺术与设计:理论版,2019(1), 2.
- [5] 孙迪, 詹锐, 张家湾等.(2018). 基于大量图像的真实故事可视化生成. 计算机辅助设计与图形学报,30(8), 11.
- [6] 刘逸清, 廖志林.(2021). 山水画的精神渊源与气化美学的生成. 中国美学研究,2021(2), 19.
- [7] 陈伟.(2020). 信息可视化与视觉媒体艺术设计在材料生产设计中的应用——《材料科学基础》回顾. 机械设计,2020(3), 1.
- [8] 刘淑红.(2018). 新媒体艺术下信息可视化设计的反思与创新. 艺术与设计:理论版,2018(1), 3.
- [9] 邱晨宇.(2021). 在虚拟学习场景中构建 3D 空间与交互设计的研究.
- [10] 刘佳瑶, 黄航琼, 安洁.(2020). 视觉页面生成方法、装置、计算机设备及存储介质:CN112015501A[P].

# 第5章 声音艺术：组织化的视觉与听觉

## 第1节 视觉与听觉通道中情感感知的神经机制的相似性与差异

### 1. 相关综述

#### 1.1 情绪认知的神经机制研究

情绪是快速变化的心理和生理现象,反映了人类有机体适应不断变化的环境的适应模式,包括人对客观事物的态度和体验以及行为反应。在现实生活中,人们在完成各种活动时不可避免地会受到不同情绪的影响。例如,当我们与他人交谈时,我们可以通过他们的面部表情和声音感知对方当前表达的情绪。这些来自视觉和听觉通道的信息通过不同的感官器官转换成电信号,然后通过大脑的渐进式认知,形成了感知、判断、记忆和行动的指令[1]。

詹姆斯是美国心理学的创始人,他在1884年首次提出人的情绪体验源自于个体对外周生理反应的感知反馈,这主要指的是外周神经系统支配下的内脏器官和腺体的活动,比如心率和血压等自主神经功能的变化,以及骨骼肌的运动和变化。詹姆斯的观点强调,身体对刺激的外周生理反应是情绪体验产生的必要前提,而负责情绪体验的只有感觉皮层和运动皮层,并不涉及大脑的任何特殊中心,这可以说是一种情绪的“外周决定论”。自那以后,一些研究者试图找到与不同类型的基本情绪相对应的外周生理反应模式,他们发现当人类经历不同情绪时,确实存在如心率(HR)、血压(BP)和手指温度(FT)等外周指标上的一些差异。近年来,使用心率变异性(HRV)等指标的研究也表明,不同的基本情绪有不同的特定自主活动模式。

不同于情绪的“外周决定论”,也有一些研究者更加关注情绪活动的中枢神经机制。例如,美国情绪生理学家坎农(Cannon)认为,自主神经活动触发的外周生理反应可能只是情绪产生的非特异性表现,真正决定情绪性质的是大脑皮层中心。这一观点也得到了某些研究证据的支持,尤其是近期的神经功能成像研究:包括眼眶额皮层(OFC)、腹内侧前额皮

层 (vmPFC)、杏仁体、脑岛、扣带皮层、下丘脑和感觉皮层等多个皮层可能是处理情绪信息、产生情绪感受和调节情绪表现的关键脑结构,并提出不同性质的情绪认知可能需要不同中枢神经系统回路的参与。一些研究者已经整理出积极情绪与消极情绪之间的生理和大脑反应差异,并认为处理不同情绪的大脑机制之间存在相同的部分和不同的部分。

情绪认知处理的大脑半球优势是情绪研究中的一个热门话题,但目前尚无共识性的结论。在研究情绪信息诱发的认知处理的半侧神经机制方面,主要有右半球偏侧理论和情绪维度理论。一些研究者整理了多项脑电波 (EEG) 和功能性磁共振成像 (fMRI) 的研究,发现正常人在情绪处理过程中确实存在大脑半球偏差现象,并提出了右半球偏侧理论,认为右大脑半球是处理所有情绪信息的优势半球。在许多脑损伤患者的研究结果也支持情绪处理的右半球偏差观点。例如,研究者发现左脑损伤患者在重新识别面部表情的任务中表现显著优于右脑损伤患者。还发现,在 37 名脑损伤受试者识别六种基本面部表情的实验中,尽管所有参与者都能正常识别愉悦表情,但在识别负面情绪,特别是恐惧和悲伤情绪时受到影响,因此研究者认为情绪认知障碍与右侧顶叶皮层和右侧中前下皮层的损伤有关。对语音的情绪类型判断的研究以及在两耳分别听情感和非情感词对的研究也显示了情绪词汇处理的左耳优势(即右半球优势)。但在持续进行的情绪认知研究中,越来越多的发现不支持右半球偏斜理论。一些研究者认为积极情绪和消极情绪是由不同的情绪处理系统完成的,大脑左半球在处理积极情绪信息时占优势,而在处理消极情绪时右大脑半球占优势,提出了情绪加工的情绪价值理论。情绪价值理论认为,情绪信息处理的半球偏差可能取决于信息传达的情绪价值。一些大脑成像研究也证实了积极和消极情绪之间大脑处理机制的分离,支持情绪处理过程存在特殊的大脑半球专业化观点,左半球更多地参与积极情绪的处理,右半球更多地参与消极情绪的处理。然而,一些研究者综合分析了连续十年内 65 项不同大脑成像研究的结果,他们得出的结论不支持情绪维度理论,而是认为情绪活动的半球差异复杂且具有很大的区域特异性。可以看出,在研究情绪的不对称性,即偏侧化过程中,右半球偏侧理论和情绪维度理论都无法完全解释情绪偏侧处理的全部结果。来自认知神经科学的最新证据对这两种理论提出了质疑,趋向于认为情绪是自动和控制处理的结合。当前理论认为,情绪处理是一个自动化的过程,主要激活大脑的左半球。处理的后期阶段需要消耗注意力资源,主要激活大脑的右半球。

总结来说,从研究方法的角度看,情绪认知研究主要集中在行为研究方法与认知神经科学和电生理学的结合上。在以往关于情绪认知的研究中,大多数都是关于情绪识别的周围和中央神经机制,以及识别不同情绪所需的神经回路的讨论,并且对于情绪处理的大脑半球优势没有共识性的结论。以下是对视觉和听觉通道的情绪认知机制研究现状的回顾。

## 1.2 视觉与听觉通道的情绪感知研究

在日常生活中，人们可能会从不同的感官通道接收情绪信息，面部和声音是非常常见的情绪信息传递媒介。来自面部和声音的情绪信息通常是复杂且多变的，相似的表情或声音可能表达完全不同的情绪，在这个过程中，无法感知和理解这些非言语的情绪信号往往会导致交流障碍。对大脑皮层的研究使研究人员认识到，来自不同感官通道的信息的融合和整合是认知的重要组成部分。然而，在传统的大脑模型中，认为大脑有分开的脑区和网络来专门处理来自面部和声音的信息，以及一个更高级的多感官脑区负责整合来自不同感官通道的信息。多感官整合被延迟到“单感官”处理之后发生。这种观点来源于早期的动物神经解剖学研究，这些研究表明，躯体感觉、听觉和视觉皮层之间几乎没有关联。在这个原则被广泛认可的情况下，许多研究人员探索了大脑对于面部和声音情绪信息的处理机制。

面部是用于研究视觉通道情绪的一种非常特殊的刺激材料，它具有独特的属性，自身拥有丰富的表情，并且与情绪有着非常密切的关联。作为常见的情绪信息来源，面部包含了大量的情绪信息，情绪面部感知能力是情绪智力的一个非常重要的组成部分，是否能准确识别他人的情绪面部以及做出恰当的反应，这部分反映了个体社会能力的发展，与个体社会互动、交流的成功与否相关。因此，使用情绪面部作为视觉通道的刺激材料来研究情绪感知是非常常见的。

面部处理的神经基础已经被一定程度上研究了，研究表明双侧颞下回（脑内侧颞叶的一部分，常被称为面部融合区 FFA）和右侧前额区的额下回与面部处理有关。还有其他功能性磁共振成像（fMRI）研究表明，颞下回的激活与面部处理过程有关。也有观点认为，面部处理涉及到颞下回、右侧海马回以及颞叶前皮层。此外，对于面部处理中颞上沟（superior temporal sulcus, STS）的重要作用也得到了确认。利用颞上沟区域作为面部选择区的研究证实了颞上沟的激活强度和模式能够区分参与者在视觉工作记忆任务中维持的信息是否为面部。

基本表情理论认为，由于面部肌肉结构的决定，存在六种基本表情——愤怒、恐惧、惊讶、愉悦、厌恶和悲伤——这些表情在人类中普遍存在。在这一理论的影响下，上述六种基本面部表情的研究在视觉通道的情绪感知研究领域占据了主导地位。许多研究者认为，不同情绪面部的处理模式和倾向也存在差异，例如，实验发现，当参与者在任务中观看悲伤面部表情时，悲伤强度的增加与左侧杏仁核和右侧颞叶活动的增加有关；愤怒面部表情的强度增加与眶额皮层和前扣带皮层活动的增加有关，这一实验不支持杏仁核是参与愤怒表情处理的

重要脑区的假设。也有研究显示，杏仁核的激活与愉悦和恐惧情绪面部的处理有关。同时，也有研究确认颞上沟在情绪面部识别中扮演着重要角色。不同情绪面部的处理模式和倾向也存在差异，许多研究考察了对积极和消极情绪面部的注意偏向，认为消极面部比积极面部更能吸引注意力[4]。

在情绪面部的研究中，情绪面部识别和情绪面部感知都是关键的研究方向。很多关于面部识别的研究已经产生了有意义的成果，例如情绪面部识别中的性别差异、年龄差异、跨文化研究以及情绪面部识别的 ERP 成分研究，以及对特殊群体情绪面部识别的研究。许多关于视觉情绪感知的研究集中在探究情绪感知与注意力及意识之间的关系。长期以来，许多研究者相信，情绪感知，尤其是负面情绪的感知，不仅可以在没有注意力参与的情况下发生，甚至可以在没有意识的情况下发生。这一点在心理学界引发了许多争论，近年来，使用 fMRI 和 ERP 技术的相关研究成果层出不穷。虽然这些利用最新技术的研究可以很好地记录与情绪活动相关的大脑生理活动，但这些心理和生理指标最终揭示的结果之间存在显著的差异。

情绪声音也是研究情绪处理机制的常见刺激，因此许多研究者探索了情绪声音感知的神经机制。除了初级听觉皮层（A1），人类听觉的神经机制已经被广泛研究，研究人员发现语言处理区可能位于双侧额上回。研究人员在事件相关功能性磁共振成像实验中扫描了健康的成人志愿者，提供了额上回对人类发声反应特异性的证据。同时，研究者还发现双侧内侧和下颞叶脑区与语言理解有关，而声学特征处理更多发生在背侧颞叶区域。左侧后上颞皮层（Wernicke 区）对听觉语言理解特别敏感。听觉注意力与布罗卡区 44（BA44）有关。在此基础上，情绪声音的研究也进入了人们的视野。

多项研究已经确认了颞叶和额叶皮层在听觉情绪感知中的作用。研究表明，当参与者听到情绪声音刺激时，双侧的反应区域总是出现在额颞回的后部初级听觉皮层，概率纤维追踪显示这些“情绪语音区”（VAVs）与同侧的内侧膝部直接结构连接，这是早期听觉皮层的主要输入源。研究还显示，听情绪声音的参与者总是激活外侧颞叶（上颞回和/或中颞回）。与中性声音刺激相比，当呈现愤怒声音刺激时，额上回的激活显著增强。双侧额上回的激活反应与快乐和愤怒语调的情绪强度的线性关系表明，这一区域的血流动力学响应的增加是由积极和消极情绪引起的，不依赖于声学特性。额叶皮层对情绪声音刺激的反应也进行了研究，研究表明，在情绪声音刺激下，受试者激活了右侧前额叶皮层。在听觉情绪识别任务中，参与者激活了小脑和额叶区域，这表明这些区域参与了情绪的处理。研究者致力于研究情绪声音与语义处理机制之间的关系，并发现双侧眶额叶的反应与受试者听到的听觉刺激的情绪信息有关。

情绪声音认知中前额叶和颞叶皮层的半球倾向性也被研究。许多研究的结论支持情绪声音处理的右手优势，研究者使用功能性磁共振成像（fMRI）来研究处理不同音素的情绪声音和声音刺激的神经机制，要求十名男性根据所表达的情绪（愤怒、愉悦、悲伤、中性）或单词的音素特征（特别是首辅音）来区分单词，并且对前额叶和听觉皮层的分析显示，在情绪处理过程中，相比于语言处理，受试者右脑半球的活动性增加。在情绪声音的评估中，右侧后颞中回和双侧额叶皮层显示激活，在情绪语义判断中，左侧颞叶前部和内侧额叶皮层显示激活。还得出结论，在听觉刺激下，与解决情绪表现相关的激活主要发生在受试者大脑的右下顶叶、双侧额叶和其他皮层，且在双侧背外侧额皮层的左右半球不对称性中展现，激活不依赖于刺激的声学结构或情绪类型。另一方面，一些研究对情绪声音的右半球偏斜提出了质疑。为了研究处理情绪声音感知的大脑网络，研究者设计了识别五种不同基本情绪（快乐、悲伤、愤怒、恐惧和厌恶）的声音的实验，并进行了功能性磁共振成像（fMRI）研究。结果显示，在语音感知过程中，一些独特的大脑网络支持语言和语调信息的处理。大脑处理不同情绪类型的神经网络有所不同（或部分交叉），其中前额叶外侧部分的激活是半球特异性的，即左半球的激活与快乐情绪有关，而右半球的激活与恐惧情绪有关。

在认知领域中，对视觉通道和听觉通道的情绪感知的研究主要集中在结合行为研究方法、认知神经科学和电生理学等多种手段，而基本情绪面孔和基本情绪声音则是视觉和听觉通道情绪感知研究中的主要刺激材料。在以往对视觉通道和听觉通道情绪感知的研究中，大部分都集中在研究视觉通道和听觉通道情绪识别的中枢神经系统机制。然而，视觉通道和听觉通道的情绪感知机制的相似之处和差异尚未被完全揭示出来。目前还没有明确的答案来解释视觉通道的情绪感知是否需要意识和注意力的调节，以及视觉通道和听觉通道情绪感知的大脑偏侧化问题也没有得到清晰的阐述。

## 2. 多感官皮层研究和情绪感知

近年来关于人类及其他灵长类动物多感官处理的研究迅速积累，直接对传统的分开研究视觉和听觉机制的研究方法提出了挑战。越来越多的研究者认同，处理来自不同感官通道信息的神经系统并不是彼此独立的，而是相互整合的。多感官整合在感知、记忆等方面发挥着重要作用。多感官中心也已经成为近年来国际研究的热点话题。关于多感官大脑区域的经典研究表明，有一些大脑区域与多个感官区域相连，它们对具有多感官输入的单个神经元的反应，以及在这些区域有损伤的患者的行为缺陷研究，都可以证明这些皮层区域能够处理多种感觉。

灵长类动物的经典多感官皮层包括额上沟和额叶皮层。许多神经影像学研究表明在额上沟区域存在多感官处理的可能性。例如，对猴子额上沟的多感官区域（STP）的神经记录显示，这个大脑区域的某些神经元可以同时对视觉和听觉刺激作出反应。研究表明，额上沟的多感官区域位于听觉和视觉相关皮层的界面上，并且它们接收来自各自基础感官区域的感官信息。听觉皮层中的多感官过程可能通过与额上沟的互动发生。直接研究前额叶皮层中的多感官处理的研究相对较少。在视觉和听觉任务上，猴子的前额叶神经对听觉和视觉刺激都有反应，前额叶皮层包含对视觉、听觉和体感输入都有反应的神经元。也有实验证明疼痛、热感、触觉、视觉、听觉等都被投射到前额叶皮层的同一部分，并在不同程度上汇聚到同一个神经元上。

然而，一些解剖学研究表明，多感官整合的神经机制也可能存在于初级皮层中。越来越多的研究发现，之前认为只负责分析和处理单一感觉刺激的感觉皮层，也可以受到其他感觉刺激的影响，并直接参与多感觉信息的整合。这些结果也有助于探究视觉刺激（如面部表情）和听觉刺激（如非语言情感声音）是否有共同的情感处理通路。

在自然环境中，大多数情感信息是通过多个感官通道来刺激的，而以往的研究促使我们仔细探索情感处理是否也具有类似的多感官大脑区域。仅仅专注于特定的感官过程和某个感官大脑区域，并不能为后续认知大脑中情感感知的神经机制提供足够完整的证据。

在情感的视听整合研究中，也取得了相当的成果，这些成果支持了面部和声音中情感信息整合过程的存在，并为研究视觉和听觉通道中情感直觉的多感官皮层研究提供了有意义的研究证据。研究表明，当面部表情和情感声音所表达的情感信息一致时，参与者处理情感刺激的效率会得到提高，与单一通道刺激和双通道刺激所表达的情绪类型的不一致相比。此外，通过一个通道获得的情感信息可以改变另一个通道中情感信息的处理，例如，如果伴随着恐惧（与中性相对）声音，面部表情更有可能被认为是恐惧的。还有研究表明，当视听信息同时呈现时，视觉信息可能首先被处理，并影响随后的视听信息处理关系。即使在明确的指示下参与者忽略在另一个感官通道中同时传达的信息，并且不受注意力资源的约束，如果不同通道表达的信息不一致，参与者的反应也会产生偏差。

视听整合是发生在较低级别的单一感觉大脑区域还是在较高级别的多感觉皮层，目前的研究还不够清晰。一些研究者通过PET研究发现，在视听刺激（情感面孔和声音）下，左侧颞中回(MTG)的活动显著增强，与单独呈现任何一种刺激（情感面孔或声音）相比。颞叶在视听情感整合中的作用也在一系列fMRI实验中得到证明。一些研究表明，面部和声音情感的整合发生在多感官区域，如杏仁核、眼眶额皮层和后上颞沟。然而，这些情绪的视听整

合大脑区域是位于面部选择区还是语言选择区，或者是独立于前两者之外的，目前还没有得到清晰和完整的研究。也许，面部和声音的情感信息整合可以在情感处理的早期阶段发生，通过面部皮层和声音优先网络之间的互联。近期的识别模型表明，面部和声音的整合可能发生在非高级感知区域内，这可能是因为在颞回内侧的梭状回和颞回中的回语言选择区域的结构直接连接和功能连接。一些研究者通过功能连接性研究得出结论，当参与者识别熟悉的人时，梭状回面部区和颞上沟是连接的，并且评估人的熟悉度不一定涉及多模态高级大脑皮层，而可能是直接在听觉和视觉区域之间共享的。有趣的是，研究在猕猴身上显示，听觉皮层区域有相当比例的语言选择性神经元对视觉和听觉刺激都有反应，并且对视听刺激的反应比单独的声音或面部刺激更强烈。这些研究也为情感感知的多感官皮层研究提供了相当大的帮助。

同样，除了上文已经提到的额上沟和额叶皮层之外，在传统上被认为是处理单一感觉的大脑区域中，视听情感感知的多感官皮层也可能发生。自1970年代以来，对灵长类动物和人类颞皮层的研究表明，颞皮层不仅在听觉信息处理中起着重要作用，而且还具有复杂的视觉、运动感知信息处理以及各种信息存储的记忆功能。一些研究者探索了后上颞皮层对面部-语言多感官情感刺激的影响。在这项研究中，后上颞沟和前颞回分别区分了中性和情感视听整合刺激。对参与者恐惧和中性刺激的大脑激活分析确认了超颞回在多模态情感刺激整合中的作用，特别是在恐惧语言处理的过程中。这一发现支持了这样的假设：传统上被认为是听觉皮层的极颞回区域，也可能是一个多感官处理的大脑区域。作为传统意义上的视觉皮层，条纹体外侧视皮层可能不仅仅处理视觉信息，而且是多感官的处理区域，并在感觉整合、运动和感觉-运动整合中发挥重要作用。声音和面部刺激似乎在传统上被认为只处理其中一种模态刺激的大脑区域中发生了交互作用，跨通道的感觉信息绑定不仅发生在大脑的高级多感官区域，也许还发生在更早的初级感觉大脑区域中。

总结来说，虽然在传统研究中，研究者们一直专注于单一模式的感觉感知和神经生物学，并且认为感觉信息在低级别的单感觉皮层中被感知性地处理，并提交给更高级的多功能脑区进行高级处理。但现实世界中的行为和感知是由来自多重感觉源的信息整合驱动的，越来越多的证据表明，多感觉整合的神经基础扩展到了早期感觉处理皮层。前辈们对多感觉皮层的研究已经取得了一定的成果，如上颞沟和前额叶等大脑区域已显示出它们在处理视觉和听觉通道中的情感信息方面的参与，因此这两个大脑区域被认为是经典的多感觉大脑区域。

视听整合的研究为多感官皮层的研究提供了一些帮助，但是跨视觉和听觉通道的情感感知多感官皮层的神经机制尚未明确研究，尤其是处理情感信息的多感官皮层的哪些部分，是否存在于传统意义上更高级的多感官脑区，或者可能出现在处理单一感觉的传统低级脑区。

在感知多感官脑区中不同情绪的激活差异也尚未得到很好的研究。

通过对中国有关视觉、听觉以及视听整合情绪认知研究的学习和总结，我们对视觉与听觉通道的情感感知机制有了更深入的了解。以往的研究为本研究提供了坚实的理论基础：面部和非言语声音的情感感知的神经基础，以及跨通道情感信息感知的多感官皮层的存在。基于这一理论，本研究将根据自己的实验目的设计相关实验，并从认知神经层面探索人类在视觉与听觉通道中情感感知的神经机制的相似性和差异[5]。

针对视觉通道和听觉通道情感感知的神经机制的研究，主要考虑以下几个方面：首先，在社会活动中，来自面部和声音的情感往往是复杂和多变的，相似的表情或声音可能代表完全不同的情绪，在社交过程中，无法感知和理解这些非言语的情感信号往往会导致沟通障碍。其次，以往的研究常常认为大脑拥有独立的脑区和网络来专门处理来自面部和声音的信息，但面部和带有情感信息声音的情感感知的神经机制的相似性和差异性并没有得到很好的研究。第三，有限的视听整合研究探索了一些视觉-听觉整合的神经基础，而从视觉和听觉通道处理情感信息的多感官皮层并没有被明确研究过，这个脑区是来自高级多感官皮层，它接收来自低级单感官皮层的信息进行统一处理，还是可能出现在处理单一感觉的传统意义的低感官皮层中，这仍然存在争议并且尚未被明确证实。第四，如果这个多感官皮层存在，它是否能够感知不同类型情绪的不同感知尚未被明确研究。因此，研究视觉和听觉通道中情感感知的神经机制是必要的。

# 第2节 基于视听通道的注意力对声音诱发的闪光错觉的影响

## 1. 综述

### 1.1 多感官整合及其神经机制

当我们在喧闹的人群中寻找朋友时，如果那个人在挥手并大喊，我们很容易就能锁定目标。为了帮助我们更快地完成搜寻任务，来自不同感官通道的信息（视觉：挥手；听觉：尖叫）不仅进行了相互作用，而且还汇聚成了一个整体的、有意义的表现。这些感官系统之间的交互和汇聚被称为多感官整合 (Lewkowicz & Ghazanfar, 2009; Talsma et al., 2010)。多感官整合有两种主要的行为输出形式。第一种是多感官错觉效应，它展示了不同感官信息整合的效果，例如腹语效应 (Hairston, Wallace, Vaughan, Stein, Norris, & Schirillo, 2003)、McGurk 效应 (McGurk & MacDonald, 1976)、定格效应 (Vroomen & de Gelder, 2000) 和双重闪光错觉 (Shams et al., 2000) 等。第二种是多感官行为改善效应，如冗余信号效应 (RSE)，在这种效应中，不同感官系统同时呈现的刺激比单独呈现同一刺激时的反应更快、更准确。

自 19 世纪末以来，多感官整合领域引起了广泛的兴趣。在近几十年中，许多研究利用先进的神经影像学和电生理技术来探讨多感官整合发生的时间点和大脑区域。多感官整合能够发生在一些皮层和皮层下的大脑区域。上丘 (Superior Colliculus, SC) 是中脑的一部分，含有大量的多感官神经元，这些神经元在整合来自本体感觉、视觉和听觉通道的信息中扮演着重要角色 (Fairhall & Macaluso, 2009; Meredith & Stein, 1996; Wallace, Meredith, & Stein, 1998)。上颞沟 (Superior Temporal Sulcus, STS) 作为一个联合皮质区域，在对象识别层面调节多感官过程的增益 (Werner & Noppeney, 2010b)，特别是来自不同通道的生物相关刺激，包括语言 (Senkowski, Saint-Amour, Gruber, & Foxe, 2008)、面部/声音 (Ghazanfar, Maier, Hoffman, & Logothetis, 2005) 以及真实物体 (Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Werner & Noppeney, 2010a)。顶叶的后部区域，如上顶叶和顶内沟，能够通过预测性的运动控制来调节多感官促进效果 (Molholm, Sehatpour, Mehta, Shpaner, Gomezramirez, & Ortigue, 2006; Werner & Noppeney, 2010a)。后顶部区域和前运动皮层在空间的指导和控制作用上也很重要，对不同感官通道的神经信号整合也很重要 (Bremmer et al., 2001; Driver & Noesselt, 2008)。此外，前额叶皮层的神经元被认为参与了语义层面的跨通道连接 (Fuster, Bodner, &

Kroger, 2000)。例如，腹侧前额叶皮层调节语义分类的多感官促进效果 (Sugihara, Diltz, & Romanski, 2006; Werner & Noppeney, 2010a)。此外，感官整合还可以影响较低皮层水平的活动，例如，初级视觉皮层 (Martuzzi et al., 2007; Romei, Murray, Merabet, & Thut, 2007)、初级听觉皮层 (Calvert et al., 1997) 和初级体感皮层 (Cappe & Barone, 2005; Zhou & Fuster, 2000)。这些假设的单一通道感觉区域也被认为参与了多感官过程 (Ghazanfar & Schroeder, 2006)。

多感官整合也被认为在不同大脑区域之间存在解剖学上的联系。一方面，与感觉相关的皮层下结构及其与相应的皮层区域的连接在多感官处理中起作用。这种连接包括内侧膝状核与初级听觉皮层之间的连接，以及外侧膝状核与初级视觉皮层之间的连接 (Noesselt, Tyll, Boehler, Budinger, Heinze, & Driver, 2010)。上丘区域中的多感官整合也显示出由皮层输入的调节 (Jiang et al., 2001)。另一方面，皮层之间的连接可以调节多感官整合。例如，同步的听觉刺激可以通过增加低层视觉皮层和听觉皮层的连接，改善视觉层面的激活以及视觉感知敏感度 (Lewis & Noppeney, 2010; Romei, Murray, Cappe, & Thut, 2009)。

综合来看，多感官整合能够在多个神经水平上发生 (即皮层下区域、联合皮层以及最底层皮层区域)，这表明多感官整合可以被许多因素所调节 (Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1986a, 1986b; Stein & Meredith, 1993)。早期研究表明，多感官刺激的强度、时间一致性和空间一致性是多感官整合的决定因素。虽然多感官整合被认为是一个典型的自动化处理过程，但它也受到诸如注意力 (Talsma & Woldorff, 2005) 这样的自上而下的因素的影响。

## 1.2 感官支配效应

在自然环境中，我们对事物的感知通常来自多个感觉通道，多感觉通道的信息导致了感觉通道之间的相互作用。这种多感觉通道之间的相互作用可以表现为通道之间的信息整合和通道之间的信息竞争 (Chen & Zhou, 2013; Driver & Noesselt, 2008; Koelewijn et al., 2010; Spence, 2011; Talsma et al., 2010)。感觉通道支配效应是指在处理特定属性 (例如时间属性) 时，某个感觉通道 (如听觉通道) 占主导地位，影响另一个感觉通道 (如视觉通道) 对这一属性的处理。

在正常情况下，来自同一对象的多感官通道信息会相互融合，即存在一致性 (对应性，参见 Spence, 2011)。例如，从近到远的汽车声并不会听起来像火车声。多感官通道信息之间的这种一致性和冗余性使得识别对象更加容易。然而，当环境非常嘈杂，存在许多感官输入源，且到达感受器的速度非常快时，每个感官通道的输入信息可能会发生冲突，导致感知

错误。多通道信息处理的主流理论是“最优线索整合”理论，该理论认为大脑以加权平均的方式将多感官通道的空间或时间维度信息整合成统一的感知表征，并且每个感官通道根据其对应感觉刺激空间（或时间）信息的可靠性来确定其在多通道信息整合中的权重（Alais & Burr, 2004; Hartcher-Ó'Brien & Alais, 2011; Witten & Knudsen, 2005）。

在视听整合的过程中，的确存在一种优先的引导作用，即视觉在空间维度上主导听觉处理，而听觉在时间维度上主导视觉处理（Talsma et al., 2010）。视觉通道支配效应是指在视觉通道与其他感官通道发生冲突时，视觉通道占据主导地位。最典型的视觉通道支配效应包括腹语效应、Colavita 效应和橡胶手幻觉。其中，腹语效应，也被称为腹语现象，指的是对听觉刺激的空间感知位置偏向于同时呈现的视觉刺激（Alais & Burr, 2004; Bertelson & Radeau, 1981; Choe, Welch, Gilford, & Juola, 1975; Howard & Templeton, 1966; Pick, Warren, & Hay, 1969）。例如，在观看腹语表演时，我们通常感觉声音来自腹语师旁边的傀儡，好像是傀儡在说话而不是腹语师；当我们看电视或电影时，声音似乎来自演员的嘴巴，而不是真正的音源扬声器。相反的效应，即使存在，即视觉刺激的空间感知位置移至同时呈现的听觉刺激位置，其效应量也非常小，这表明视觉位置不会因为与其位置不同的听觉信号而改变。尽管在大多数双通道互动情况下，视觉通道是占主导地位的。但在时间维度上，听觉通道通常占主导地位。听觉通道支配的最典型现象包括声音诱发的闪光幻觉、听觉驱动效应和时间腹语效应。其中，声音诱发的闪光幻觉是指在 100 毫秒内连续或同时呈现的不等数量的听觉声音伴随着视觉闪光时，被试感知到的视觉闪光数量等于听觉声音的数量，这一现象说明听觉信息在处理视觉信息时具有优先权，即听觉信息可以支配视觉信息（Shams et al., 2000, 2002）。当一个视觉闪光伴随两个听觉声音时，单个视觉闪光被错误地感知为两个视觉闪光，也被称为裂变错觉。当两个闪光伴随一个听觉声音时，两个视觉闪光被错误地感知为一个视觉闪光，也被称为融合错觉。研究表明，融合错觉的效果比裂变错觉的效果要弱（Shams et al., 2000; Wozny, Beierholm, & Shams, 2008）。

### 1.2.1 感官支配效应理论

#### (1) 指向性注意解释

Posner, Nissen 和 Klein（1976 年）通过指向性注意的方式来解释视觉的主导性。指向性注意理论认为，与其他感官通道（例如听觉）的刺激不同，视觉刺激并不会自动捕获注意力。因此，参与者必须主动将他们的注意力指向视觉刺激。这种主动的注意力定向需要消耗认知

资源，导致留给其他感官通道刺激的注意资源减少。Colavita (1974) 也提出了注意力在视觉主导性中发挥作用的假设，该假设后来受到 Egeth 和 Sager (1977) 以及 Sinnott, Spence 和 Soto-Faraco (2007) 的影响和支持。这些研究者发现，调节注意力资源可以影响视觉主导效应，但不能消除它。例如，通过增加听觉刺激的比例，可以使注意力更多地指向听觉通道，但视觉主导效应只是被削弱而没有被消除(Egeth & Sager, 1977; Sinnott, Spence, & Soto-Faraco, 2007)。

## (2) 模态适宜性解释

通道适宜性理论由 Welch 和 Warren (1980) 提出。该理论认为，每个通道中的刺激只有在该通道适宜的维度中才能被最佳处理 (Ward, McDonald, & Lin, 2000)。对于视觉，最适宜的维度是空间维度，而对于听觉，最适宜的维度是时间维度 (Bertelson & Aschersleben, 2003; Shams 等人, 2000)。视觉用于处理空间信息，听觉用于处理时间信息。视觉主导效应的原因是因为视觉对于空间定位比听觉更有用，这是由于视觉输入在视网膜上是以空间关系映射的。这种空间映射在高级视觉处理区域得以保持和使用。因此，当视觉和听觉信息发生冲突时，这种冲突以偏向视觉通道的形式得到解决。此外，当听觉刺激的空间位置与当前任务无关时，听觉刺激不会被完全处理，因此干扰性较小。

大多数关于感官通道冲突的研究要求参与者识别刺激的某些空间属性（例如，位置、倾斜程度或形状）。由于视觉是最适合评估所需特征的感官通道，因此它具有优先权，同时其他通道中的信息会被调整以消除冲突。相反，当执行时间处理时，听觉系统占据优先地位，视觉将被调整以适应听觉输入。相应地，最近的研究表明，跨通道效应的方向并不是静态的，而是随着每个通道传递的信息的可靠性而变化。

变化 (Witten & Knudsen, 2005)。例如，当视觉输入是理想的，听觉信息对于定位视听刺激并不有帮助；但当视觉输入不理想时，听觉信息对于定位就变得有帮助了。在视觉-本体感觉互动中也发现了类似的结果 (Ernst, Banks, & Bühlhoff, 2000)。因此，人类大脑会根据特定环境优化地整合信息：当通常占据主导地位的通道中的信息不足时，次要通道的效应开始变得显著。

### 1.2.2 不同认知层次上的感官优势效应

跨通道冲突可能发生在早期的感知和语义处理中，也可能发生在反应选择和反应执行中

(Colavita & Weisberg, 1981)。不同感官通道之间的互相优势不对称性取决于处理的层次。在某些运动和时间相关任务中，例如同步节奏性手指敲击，听觉干扰项对视觉信号同步敲击任务的干扰比视觉干扰项对听觉信号同步敲击的干扰要大 (Mayer, Franco, Canive, & Harrington, 2009)，即听觉优势。这表明在时间维度的处理中，听觉优势可能发生在反应层面。

一项功能性磁共振成像(fMRI)研究表明，感官通路优势的具体方向并不是由当前行为任务的任务需求决定的，而是由不同的认知处理水平决定的，例如早期和晚期 (Chen & Zhou, 2013)。这项研究向参与者呈现了双通道视听刺激，要求参与者在忽略另一通道的刺激的同时，要么注意视觉，要么注意听觉。通过控制视听通道的三个水平（一致性、前反应水平不一致性和反应水平不一致性）之间的关系，将跨通道冲突分为前反应水平和反应水平。结果显示，在早期反应之前的认知处理水平上，视觉干扰项对听觉处理造成的干扰量显著大于听觉干扰项对视觉处理造成的干扰量，表明在早期处理水平上视觉占主导。相反，在晚期反应水平上，听觉干扰项对视觉处理造成的干扰量显著大于视觉干扰项对听觉处理造成的干扰量，表明在晚期反应水平上听觉占主导。在神经水平上，发现早期的视觉优势效应和晚期的听觉优势效应由不同的神经网络负责。前反应水平阶段的视觉优势效应主要激活默认模式网络；而反应水平阶段的听觉优势效应则激活了一些执行功能的网络。

## 1.3 声音引发的闪光错觉现象

### 1.3.1 定义

Shams 等人 (2000, 2002) 首次提出听觉信息可以主导视觉信息的处理。该研究使用了两个间隔为 57 毫秒的听觉刺激，并伴随一个与第一个听觉刺激间隔为 23 毫秒的视觉刺激，然后让参与者判断所感知的视觉闪光的数量。研究发现，大多数参与者感知到两次视觉闪光。因此，实验结果表明，当一个视觉闪光伴随着两个听觉声时，单个视觉闪光会被错觉为两个视觉闪光，这种现象通常被称为裂变错觉 (Shams et al., 2000, 2002)。Andersen 等人 (2004) 采用了 Shams 的研究范式，并设置了多个视觉闪光刺激伴随一个听觉刺激，允许参与者报告视觉闪光的数量。研究发现，除了 Shams 发现的裂变错觉，还发现当两个视觉闪光伴随一个听觉声时，两个视觉闪光被误感为一个视觉闪光，这种现象通常被称为融合错觉 (Andersen, Tiippana, & Sams, 2004)。因此，声音诱发的闪光错觉是指在 100 毫秒内呈现的视觉闪光刺

激与听觉声刺激的数量不等，且感知的视觉闪光数量与听觉声的数量相等[7]。

### 1.3.2 原因

#### (1) 主观判断偏差

Shams 等人 (2002 年) 提出了一个伴随多个听觉声刺激的视觉闪光刺激，要求参与者对视觉闪光刺激的数量做出判断。目的是探索听觉信息是否可以主导视觉信息。研究发现，声音引发的闪光错觉基于他们自己的主观视觉感知，而不是由于任务难度或知觉偏见导致的 (Shams 等人, 2002 年)。随后, Andersen 等人 (2004 年) 采用了 Shams 等人的范式, 添加了不同强度的声音刺激, 并要求参与者判断视觉闪光的数量。结果显示, 参与者感知的融合错觉由于声音强度的变化而改变, 也证明了当外部刺激输入改变时, 错觉现象随着受试者对刺激的主观判断而变化 (Andersen 等人, 2004 年)。在神经层面上, Bolognini 等人 (2011 年) 使用经颅直流电刺激 (tDCS) 作用于参与者的颞叶、枕叶和顶叶, 检查了参与者在判断视觉闪光数量时大脑皮层的活动。研究表明, 在裂变错觉中, 当使用阳极 tDCS 时, 颞叶皮层的兴奋性增强, 听觉处理改善, 枕叶皮层的兴奋性减弱, 减少了视觉处理, 从而增加了受试者判断错觉现象的比例; 当使用阴极 tDCS 时, 颞叶皮层的兴奋性减弱, 枕叶皮层的兴奋性增强, 即视觉处理改善, 听觉处理减少, 因此参与者判断错觉现象的比例减少。然而, 融合错觉由于 tDCS 的输入而没有改变, 可能是因为在激活初级视觉皮层时融合错觉已经达到了最大值。因此, 通过 tDCS 揭示了枕叶皮层和颞叶皮层之间的因果关系, 并显示出 tDCS 可以影响与多感官视听整合相关的视觉体验 (即, tDCS 作用于参与者的的大脑)。可以改变参与者的常规视觉体验) (Bolognini, Rossetti, Casati, Mancini, & Vallar, 2011)。这些研究表明, 由于参与者对刺激输入不确定, 他们在报告视觉闪光数量时依赖自我主观感知。

#### (2) 视觉敏感度

McCormick (2008 年) 在 Shams 等人 (2002 年) 研究的基础上操纵了对比度和位置变量, 以探索参与者是否由于视觉感知敏感度的变化和判断标准的变化而感知到错觉闪光。该研究首先测量了参与者的阈值, 然后要求参与者完成位置检测任务。其中, 一半的高对比度白色或黑色闪光在所有位置 (左和右) 同时出现, 而在另一半的测试中, 一个低对比度的白色或黑色闪光出现在其中一个位置, 要求参与者判断是否出现了低对比度的闪光。同时, 一半的尝试将呈现一次声音 (只在低对比度的闪光同时发生时), 而另一半将呈现两次声音 (伴

随着两次闪光)。因此,这种方式可以计算在错觉和非错觉测试中敏感性的变化和判断标准的转移。结果表明,在错觉测试中,参与者的视觉感知敏感度和判断标准降低。研究表明,错觉闪光现象的发生是由于参与者的视觉敏感度降低和他们的判断标准降低。Kumpik 等人 (2014 年) 通过降低视觉刺激的空间可靠性来探索视觉敏感度对声音诱发的闪光错觉的影响。使用信号检测理论,发现声音诱发的闪光错觉在视野的周边区域更为常见,参与者感知到的裂变错觉数量大于融合错觉数量,这表明声音诱发的闪光错觉由受试者对视野中的闪光灯光敏感度决定 (Kumpik, Roberts, King, & Bizley, 2014)。

### (3) 神经机制

Shams (2001 年) 通过视觉诱发电位 (VEPs) 探索了声音诱发的闪光错觉对视觉感知的影响,这种跨通道整合现象是否存在于通道特异性 (特定模态) 的视觉路径中。该研究主要比较了有声音呈现和无声音呈现时的 VEPs,发现这些跨通道效应发生在枕叶皮层层面上,更重要的是,在定性分析中,错觉闪光的 VEPs 与物理闪光刺激的 VEPs 非常相似。结果表明,视觉皮层中产生的两种感知现象具有相似的大脑机制 (Shams, Kamitani, Thompson, & Shimojo, 2001)。此外,Shams (2002 年) 使用  $\gamma$  波段 (gamma band) (>30Hz) 振荡响应进一步探索了声音诱发的闪光错觉现象和非错觉现象的神经基础。通过记录事件相关电位,发现当错觉闪光现象发生时, $\gamma$  波段的振荡明显高于没有错觉的尝试次数,并且音视交互只在错觉测试中发生。因此,这些结果清楚地表明声音可以调节视觉皮层区域的处理,并且调节的强度可以影响最终感知的输出。Watkins 等人 (2006 年) 使用功能性磁共振成像 (fMRI) 通过视觉映射来研究声音诱发的裂变错觉是否可以发生在早期视觉皮层区域。结果表明,声音改变了参与者的感知,反映了受试者的主观感知而不是物理刺激,并且发现视觉双闪错觉在 V1 皮层有更高的激活水平 (Watkins, Shams, Tanaka, Haynes, & Rees, 2006)。相反,他们对融合错觉的研究发现,错觉视觉闪光在 V1 皮层的激活水平较低 (Watkins, Shams, Josephs, & Rees, 2007)。此外,经过全脑分析后,发现激活的脑区还包括颞上沟和上丘 (两者都有较高的激活水平)。Jiang 等人 (2007 年) 使用 fMRI 技术基于视觉感知的增益和损失来探索声音诱发的闪光错觉的神经机制。也就是说,当视觉闪光与听觉刺激同时呈现时,听觉刺激会影响视觉闪光的感知:当听觉刺激的数量大于闪光的数量时,参与者感知到错觉闪光,即感知增益;当听觉刺激的数量小于闪光的数量时,参与者感知到的闪光数量减少,即感知损失。结果显示,声音诱发的闪光错觉现象的变化是双向的,既可以产生感知增益,也可以产生感知损失。感知增益中的错觉感知激活了左侧额上回、左侧前额叶皮层和右侧小脑,这

反映了与工作记忆的整合处理相关的神经网络。而感知损失则与内侧枕部皮层和丘脑的激活有关，这些脑区与早期视觉处理相关。

Watkins 等人 (2006 年) 的研究使用功能性磁共振成像 (fMRI) 技术，同时呈现一个闪光和两个声音，或在不同的视野位置 (外围视野和中央凹区域) 呈现两个闪光，并改变了视听刺激之间的时间间隔，以便让参与者判断感知到的闪光数量。随后，研究人员使用血氧水平依赖性 (BOLD) 信号来表示双闪光错觉在初级视觉皮层上的反应。研究探究了双闪光错觉是否影响初级视觉皮层的活动。研究发现，双闪光错觉效应可以显著影响人类视觉皮层的神经活动，而这种效应对视听刺激的刺激间隔 (ISI) 非常敏感，表明视听刺激间隔对于跨通道神经互动至关重要，且视听刺激导致视觉和听觉感觉系统之间的双闪光错觉效应的跨通道神经互动。fMRI 的结果还表明，与中央视野相比，双闪光错觉效应在外围视野中更强，这与行为实验的结果一致，表明行为表现与 BOLD 信号密切相关，反映了大脑神经活动，并进一步暗示外围视觉可能涉及超越视网膜定位的高级脑处理和视觉感知。以上的 fMRI 结果表明了不同感觉系统中有关跨通道的时空信息的神经互动。

Mishra 等人 (2007 年) 利用事件相关电位 (ERP) 技术研究声音诱发的闪光错觉在神经生成层面上的表现。该研究主要使用了不同数量的视觉刺激 (0、1 和 2) 和不同数量的听觉刺激 (0、1 和 2) 的组合，在其中一种视觉刺激和两种听觉刺激被呈现在两种不同的情境中 (听-看-听和听-听-看)，因此有 10 种不同类型的视听刺激组合。参与者的任务是判断闪光的数量。研究发现，双闪光错觉 (即分裂错觉) 在第二个听觉声音呈现后 30-60 毫秒产生了视觉皮层活动幅度的早期调制 (Mishra 等人, 2007 年)。随后，Mishra (2008 年) 使用事件相关电位技术 (ERPs) 研究了融合错觉，并发现融合错觉在第二次闪光出现后 80-112 毫秒产生了视觉皮层活动范围的早期调制。

研究表明，声音诱发的闪光错觉现象主要在人脑的初级视觉皮层的神经层面上发生。然而，目前对声音诱发的闪光错觉的神经机制的研究主要是基于单一感觉皮层，例如初级视觉皮层和初级听觉皮层。随着研究的深入，越来越多的研究发现，声音诱发的闪光错觉不仅发生在单一感觉皮层，还与多个脑区互动，并且声音诱发的闪光错觉的神经机制得到了更加清晰的研究和阐述[8]。

#### **(4) 个体差异**

声音诱发的闪光错觉中听觉优势效应是一种典型的个体间有差异的现象。Mishra 等人 (2007 年) 的研究发现，这种错觉现象在个体之间的差异很大，发生闪光错觉的比例在个

体间的范围从 3%到 86%不等。Haas, Kanai, Jalkanen, 和 Rees (2012 年) 认为, 人脑结构的差异导致了个体对声音诱发的闪光错觉的敏感度差异。该研究遵循了 Shams (2000 年、2002 年) 的研究范式, 但是在视野的上下位置呈现了视觉刺激, 并且在参与者判断闪光数量时, 通过 fMRI 记录了其大脑活动。结果显示, 当初级视觉皮层 (BA17 &18) 的局部灰质体积较低时, 参与者更容易感知错觉, 且灰质体积与受试者感知的闪光错觉数量成反比, 无论视觉刺激是呈现在上方还是下方。这表明, 个体大脑中初级视觉皮层的灰质体积与声音诱发的闪光错觉有关, 即闪光错觉与个体大脑的结构差异相关 (Haas, Kanai, Jalkanen, & Rees, 2012 年)。

Cecere, Rees, and Romei 等人在 2015 年的研究中发现, 声音诱发的闪光错觉的视听刺激时间窗大约为 100 毫秒, 而  $\alpha$  波段的振荡活动周期也大约为 100 毫秒, 因此他们认为  $\alpha$  波段的振荡周期可以调节视听事件整合的时间窗。该研究使用脑电图 (EEG) 记录参与者在完成诱发声音闪光错觉任务时的数据。研究发现, 个体  $\alpha$  频率峰值 (IAF) 与错觉时间窗大小之间存在正相关关系。然后要求参与者在执行相同任务时, 对其枕叶皮层施加经颅交流电刺激 (tACS), 以调节振荡活动 ( $IAF \pm 2\text{Hz}$ ), 在 IAF 或非峰值  $\alpha$  频率下进行。研究比较了 IAF、IAF-2Hz 和 IAF+2Hz 的 tACS, 发现错觉的时间窗随  $\alpha$  频率的振荡活动而变化。研究表明,  $\alpha$  振荡可以调节闪光错觉的整合时间窗。因此, Cecere 等人认为, 闪光错觉是由于在快速视觉刺激处理的关键时间窗内, 连续的声音诱导视觉皮层的兴奋性发生突变引起的现象 (Cecere, Rees, & Romei, 2015)。

现有研究显示, 在声音诱导的闪光错觉中存在个体差异, 而这些个体差异与个体的初级视觉皮层灰质体积和  $\alpha$  波段的振荡周期有关。然而, 目前对声音诱导闪光错觉现象的个体差异的研究还相对较少, 需要进一步的研究和讨论。

#### **(5) 注意力对声音诱发闪光错觉的影响**

Mishra 等人 (2007) 利用事件相关电位 (ERPs) 进行了声音诱发闪光错觉的神经生成研究。他们发现, 双闪光错觉 (即裂变错觉) 在第二个听觉声音呈现后的 30-60 毫秒产生了视觉皮层活动范围的早期调制 (Mishra 等人, 2007)。因此, Mishra (2010) 基于他们 2007 年的研究得出结论, 声音诱发的光学错觉不仅仅是跨通道整合的结果, 而且应该有值得注意的影响。该研究使用 ERP 技术增加了刺激的位置 (视野高位和低位) 以及是否注意 (注意和不注意) 到 Shams 范式, 即参与者在判断闪光次数时被要求注意视野高位或低位的闪光刺激。研究表明, 在高视野和低视野刺激下, PD120/PD110 组分 (在低视野中) 位于腹侧枕颞

部条状视觉皮层，而随后的 PD180 和 ND250/ND240 组分（在低视力中）位于上颞回，而上述的枕颞附属物提到的皮层和上颞回被认为是多感官认知通道信息交互的大脑区域（图 2-3）。这些 ERP 组分与 Mishra 等人（2007）在先前研究中发现的声音诱发双闪光错觉的神经基础组分有很高的相似性，并且在这项研究中增加的注意力使这些组分得到了增强。这些结果表明，注意力的空间分配对跨通道交互与声音诱发闪光错觉现象之间的关系有着强烈的影响，说明双闪光错觉并不都是自动整合处理的结果，并且顶下认知过程参与了视觉和听觉信号的错误整合（Mishra 等人，2010）。

不同感官产生的信息被整合以形成感知。同时，注意力的认知过程也可以改变参与者的感知。经颅磁刺激 (TMS) 已被发现能够打断与选择性注意力相关的角回和上边缘回的联系，那么，如果注意力参与声音诱发的闪光错觉现象，角回的破坏将会影响受试者感知的闪光错觉数量的变化。因此，Kamke、Vieth、Cottrell 和 Mattingley（2012）使用经颅磁刺激研究了与注意力相关的皮层神经网络在声音诱发视觉闪光错觉现象中的作用。该研究主要使用了不同组合的视听刺激（一个视觉刺激伴随两个听觉刺激，一个视觉刺激，两个听觉刺激）和不同的视听刺激时间间隔（70 毫秒和 160 毫秒），并且视听刺激同时呈现，要求参与者判断闪光的次数。结果显示，当 TMS 作用于右角回（AG）时，报告闪光错觉的参与者比例显著下降。因此，他们认为破坏的角回增强了参与者在多感官事件中的真实感知，有效地减少了无关联的听觉刺激对视觉事件处理的影响，降低了听觉和视觉刺激错误整合的可能性。然而，当呈现单一视觉刺激或者当视听刺激有较长时间间隔时，TMS 并不改变参与者感知。总之，角回涉及到在声音诱发的闪光错觉中调节视觉和听觉刺激整合的过程，但该研究仅限于裂变错觉现象，融合错觉现象是否也会有类似表现还需要进一步研究（Kamke, Vieth, Cottrell, & Mattingley, 2012）。

## 1.4 内源性注意力与外源性注意力

注意力在选择相关通道、空间位置和与任务有关的对象方面扮演着重要角色，它有助于区分哪些是相关的，并排除那些无关的信息。内源性和外源性注意力机制都参与了这一注意力过程。内源性注意力，也被称为自主或目标驱动的注意力，涉及更有目的和集中的目标处理。比如，当有人告诉你你的朋友在红色的桌子那里时，你会被引导去注意红色的桌子。相反，外源性注意力，也被称为非自愿或刺激驱动的注意力，可以由外界显著的感官事件以反射性的方式诱发（Hopfinger & West, 2006）。比如，你的朋友穿着鲜艳的衣服在人群中显得格外突出。

在视觉系统的研究中，内源性和外源性注意力常被认为是两种不同的注意系统，它们具有不同的行为效应和独特的神经基础（Chica, Bartolomeo, & Lupi nez, 2013; Mysore & Knudsen, 2013）。与内源性注意力不同，外源性注意力不需要认知资源，且不太容易受到干扰（Chica & Lupi nez, 2009）。由外源性注意力触发的效应比由内源性注意力触发的效应更快且更短暂（Jonides & Irwin, 1981; Shepherd & M ller, 1989）。神经成像研究揭示了这两种机制大部分由前额-顶部网络调控。然而，研究也表明内源性注意力与背侧注意系统有关。

外源性注意力与腹侧注意网络相关联。

腹侧注意网络位于右侧，包括右侧颞顶联合（TPJ）、右腹侧额叶皮层（VFC）、部分中额回（MFG）和下额回（IFG）。腹侧注意网络参与非自愿（刺激驱动）的定向，它将注意力引向显著事件（Chica et al., 2013）。背侧注意系统是双侧的，包括上顶叶的前眼领域（FEF）、顶内沟和前额叶皮层。背侧注意网络参与自主（自上而下）的定向，其活动在提示呈现后增加，这可以指示何时、何地以及哪些对象需要注意力定向（Corbetta & Shulman, 2002）。Chica 等人（2013）提出假设，背侧额顶网络参与内源性注意力的定向和外源性注意力的定向，而腹侧额顶网络部分参与与任务相关事件的重定向（Chica et al., 2013）。ERP 研究表明，内源性和外源性注意力调节刺激处理的不同阶段。具体来说，内源性注意力影响 N1（Hopfinger & West, 2006）和 P300（Chica & Lupi nez, 2009），而外源性注意力仅调节 P1 组分（Chica & Lupi nez, 2009; Hopfinger & West, 2006）。然而，以前的研究表明，P1 的幅度可以被内源性注意力调整，而 N1 的幅度则由外源性注意力调节。

内源性和外源性注意力之间的关系也在其他模型中得到了检验。研究显示，内源性和外源性注意力机制构成了两种不同的注意系统，但它们共享有限的系统容量。在这个有限容量的系统中，内源性和外源性注意力并不是独立的，相反，它们会争夺对注意力的控制（Godijn & Theeuwes, 2002; Yantis, 1998）。在内源性和外源性注意力机制之间的竞争中获胜的一方将控制注意力，并决定何时何地需要注意力定向。无论内源性和外源性注意力是两个不同的系统，还是单一注意系统的两种模式，大多数研究至少表明这两种机制在调节刺激处理方面是不同的。这意味着内源性和外源性注意力可能以不同的方式调节多感官的整体性。

## 1.5 多感官整合与内源性注意的相互作用

注意力能够从大量的感观信息中选取适当的刺激，帮助大脑将来自不同感官通道的有价值信息整合为认知的整体（Giard & Peronnet, 1999）。相反，由于其显著性增加，一个整合的多感官刺激在复杂环境中能更有效地捕获注意力。最近，关于多感官整合与注意力相互作用

用效应的研究以惊人的方式增长。一些研究已经表明多感官整合可以独立于注意力，而其他研究则发现注意力可以调节多感官整合（Bertelson, Pavani, Ladavas, Vroomen, & de Gelder, 2000; Bertelson, Vroomen, de Gelder, & Driver, 2000; Spence & Driver, 2000; Vroomen, Bertelson, & de Gelder, 2001; Vroomen, Driver, & de Gelder, 2001）。然而，目前尚不清楚在什么情况下以及通过哪些机制多感官整合与注意力相互作用（Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Alsius, Navarra, & Soto-Faraco, 2007; Harrar, Spence, & Makin, 2014; Talsma et al., 2007; Talsma & Woldorff, 2005）。

为了解释多感官整合与注意力之间的关系，研究人员提出了许多假设。例如，Talsma 等人（2010）认为多感官整合对注意力有刺激驱动的影响，但自上而下的注意力定向也影响多感官处理（Talsma 等人，2010）。De Meo, Murray, Clarke 和 Matusz（2015）在他们的综述中提出，早期的多感官整合独立于自上而下的注意力控制（De Meo, Murray, Clarke, & Matusz, 2015）。多感官整合与注意力之间的交互作用也被认为取决于整合发生的处理级别（Koelewijn 等人，2010）。这些研究集中在自上而下的注意力控制（在这种情况下是内源性注意）与多感官整合的交互作用。然而，内源性和外源性注意力可能以独特的方式调整多感官处理，并且这两种机制在调节多感官处理方面可能有所不同。

### 1.5.1 基于空间的内源性注意对多感官整合的影响

注意力可以通过指导或暗示性的视觉中央提示，被定位到特定的空间位置，比如注视点的左边或右边，这通常被称为集中的空间注意（对一个位置的注意，见图 2-5 和表 2-1）。注意力也可以锚定到多个空间位置，例如同时锚定到左边和右边，这通常被称为分散的空间注意（对多个位置的注意，见图 2-5 和表 2-1）。先前的研究表明，内源性注意的选择可以促进对于过度注意（预期）空间位置的单通道信号（视觉或听觉-A）的反应，与未被注意（无意）位置相比（Coull & Nobre, 1998; Li, Chen, Han, Chui, & Wu, 2012; Tang et al., 2013）。这种类似的注意效应（注意 vs. 不注意）也在多感官通道的刺激中发现，例如同时呈现的视觉和听觉刺激（视听-AV）（表 2-1）。例如，未被注意的视听刺激引起的 P1 分量的振幅，比未被注意的视听刺激引起的振幅更大（Talsma et al., 2007）。当不同通道的空间一致的刺激同时呈现时，在集中（Li, Wu, & Touge, 2010; Wu, Li, Bai, & Touge, 2009）和分散（Li, Yang, Sun, & Wu, 2015）的情况下，多感官整合在空间注意条件下有显著提高。多感官整合的提高通常伴随着大脑对多感官刺激的反应，这些反应与大脑对单一通道刺激组分结合后的反应是

分离的 (例如, AV vs A+V) 。这种非线性反应是多感官整合的特征 (De Meo et al., 2015; Giard & Peronnet, 1999) 。

基于空间的选择性注意被认为以非线性方式调节多感官整合。具体来说, 注意力位置刺激的事件相关电位 (ERP) 反应在中央区域刺激呈现后 280 毫秒增强 (Li, Wu, & Touge, 2010), 在前额中央区域刺激呈现后 100 毫秒的反应也比非注意力位置的刺激更强 (Talsma & Woldorff, 2005) , 甚至在大约 40 毫秒的早期  $\gamma$  波段振荡中也能观察到这种现象 (Senkowski et al., 2005) 。选择性空间注意还能调节多感官整合的更高层次, 例如视觉刺激流中的唇部动作与听觉刺激流中的语言之间的交互 (Fairhall & Macaluso, 2009) 。在一项功能性磁共振成像 (fMRI) 研究中, 当同时呈现的语言和唇部动作匹配时, 内源性注意增强了皮质和皮层下层的激活, 包括额上沟、条纹视皮层、额外条纹视皮层和上丘。

此外, 在许多情况下, 空间注意通常分布在不同的位置 (如左右) , 而不仅仅是一个位置。在分散的空间注意条件下, 当同时呈现的听觉刺激出现在一个空间一致的位置时, 视觉目标的感知敏感度可以通过视听互动得到增强 (Frassinetti, Bolognini, & Lãdavas, 2002; Gao, Li, Yang, Yang, Tang, & Wu, 2014) 。此外, 通过在呈现后 100 毫秒刺激腹侧枕颞和上颞叶区域的神经活动, 视听互动可以由空间一致性来调节 (Teder-Sãlejárvi, Russo, McDonald, & Hillyard, 2005) 。总结来说, 大量研究发现内源性空间注意可以增强多感官整合。

### 1.5.2 基于通道的内源性注意对多感官整合的影响

在多感官刺激流中, 可以通过使用指令或意图来对特定的通道进行内源性选择性注意的分配。例如, 在嘈杂的环境中阅读书籍时, 需要集中注意力于与任务相关的通道, 即视觉刺激中阅读书籍和页面的动机, 同时忽略与任务无关的通道, 如听觉通道中的噪音。请注意, 特定通道可以加速低级皮层区域的信息处理, 这也被称为“优先进入效应” (Vibell et al., 2007) 。

当参与者注意视觉 (Wu et al., 2009) 、听觉 (Li et al., 2010) 或视听刺激流 (Giard & Peronnet 1999) 时, 对视听刺激的行为响应和事件相关电位 (ERP) 反应都得到了增强。然而, 集中注意力 (专注于一个通道) 和分散注意力 (注意多个通道) 的条件以不同的方式调节多感官处理 (见表 2-1: 专注于一个感官模态 vs. 注意多个感官模态) 。在特定通道的基于通道的选择性注意条件下, 多感官整合对行为表现的影响减少甚至完全消除 (Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008) 。当听觉和视觉刺激都被定位时, 多感官整合反应和准确性是最

优的 (Barutchu, Freestone, Innes-Brown, Crewther, & Crewther, 2013)。此外, 当注意力分散在通道之间而不是专注于一个特定通道时, 感官门控可以被调节 (Anderson & Rees, 2011; Talsma et al., 2007), 例如多感官整合发生得更早, 比如在刺激呈现后的 100 毫秒内 (P50 组分) (Giard & Peronnet, 1999; Talsma et al., 2007)。当注意力分散到多感官通道时, 早期多感官 ERP 处理 (Magnée, De Gelder, Van Engeland, & Kemmer, 2011; Talsma et al., 2007) 以及与上颞皮层区域多感官相关的 fMRI 反应都得到了增强 (Degerman et al., 2007)。

尽管在专注于一个通道和分散注意力于多个通道的条件下多感官神经处理都得到了增强, 但多感官神经处理的增强效应也与零效应 (Talsma et al., 2007, 注意力听觉刺激任务) 或多感官行为的负面效应相关 (Degerman et al., 2007)。确切地说, Talsman 等人 (2007) 和 Degerman et al. 等人 (2007) 发现了多感官神经加工的增强。然而, 在前者的研究中, 当参与者注意听觉通道时, 视听刺激的行为反应与单独听觉刺激的行为反应之间没有发现显著差异。在后者的研究中, 当参与者注意视觉或听觉通道时, 他们对视听刺激的行为反应比注意多种感官通道时更为准确。

此外, 一项事件相关电位 (ERP) 研究显示, 与集中通道选择注意相比, 在分散通道选择注意的情况下, 多感官行为的增加与神经处理的减少相关 (Mishra & Gazzaley, 2012)。这些不一致的结果表明, 集中通道注意和分散通道注意调节的多感官处理是不同的。然而, 确定行为表现与潜在多感官神经活动之间联系的方向是困难的。实验参数和任务的差异可能会导致不同的神经变化和行为结果。尽管反映脑活动的神经响应比行为数据更敏感, 但对于由于实验差异导致神经响应变化的所有解释都应该基于行为表现 (Cappe, Murray, Barone, & Rouiller, 2010)。

内源性注意力通过基于空间位置或通道的选择, 在多个阶段影响多感官行为表现。监控两个位置或通道的成本高于监控特定的通道或位置, 这与当注意力分配到位置或通道之间时, 在额顶区域或颞上皮层的激活更为强烈有关 (Degerman et al., 2007; Santangelo et al., 2010)。此外, 在两个位置监控多个通道的行为消耗小于在一个位置监控多个通道, 这一差异与左右前楔叶的活动增加有关 (Santangelo et al., 2010)。结果表明, 这两种类型的选择性注意力不是独立的; 对单一位置的多个位置的注意或集中与对多个通道的注意或对单一通道的注意相交叉。一项 ERP 研究表明, 相对于空间位置, 吸引注意力到感觉通道可能更容易增加信息处理的速度 (Vibell et al., 2007)。此外, 多感官整合与内源性自上而下的注意力控制之间的交互受到多感官刺激的语义一致性的影响, 尽管多感官刺激的语义一致性并不直接调节早期的多感官处理 (Fort, Delpuech, Pernier, 2007)。

Yuval-Greenberg 和 Deouell (2007) 的研究显示, 改变一致性尝试的比例会导致注意力控制的转移, 以及一个会间接调节多感官整合的输出。

近期研究已经在注意力内部做出了区分, 并发现内源性注意力可以通过通道注意的形式影响多感官整合过程, 也就是说, 对通道的注意力会削弱参与者未关注通道中的信息处理, 并增强被关注通道中的信息处理 (Talsma, 2015)。同时也发现, 同时注意两个通道的成本高于关注特定通道, 这与在通道间分配注意力时额顶区或颞上皮层的活动增强有关

(Degerman et al., 2007; Santangelo et al., 2010)。我们最近的一项研究还发现, 与关注一个通道相比, 同时关注两个通道可以影响声音诱发的闪光错觉现象, 这表明分散通道的注意力可以影响声音诱发的闪光错觉, 但仅限于分裂错觉。通道依赖的注意力调节感觉优势效应的一个可能解释是优先进入原则, 即两个刺激同时呈现时, 首先被注意到的参与者倾向于认为它们先出现。优先进入效应的发生可能是由于参与者对视觉/听觉通道的内源性注意, 也可能是由于视觉/听觉刺激本身外源地捕获了参与者的注意力 (Posner et al., 1976; Turatto, Benso, Galfano, & Umilt, 2002)。因此, 当参与者以内源性方式将注意力指向视觉通道时, 参与者倾向于认为在两通道测试中视觉刺激先于听觉刺激; 当参与者以内源性方式将注意力指向听觉通道时, 参与者倾向于认为在两通道试验中听觉刺激先于视觉刺激。许多研究使用不同的方法来指向内源性注意力的一个通道, 其中最常见的是改变不同通道试验中的比例。例如, Sinnett 和 Spence (2007) 在他们的实验中将单通道视觉目标的比例增加到 60%, 同时将单通道听觉目标的比例降低到 20% (其余 20% 的试验是两通道尝试), 以调查内源性通道注意对 Colavita 视觉优势效应的影响, 实验结果发现 Colavita 视觉优势效应变大; 当单通道视觉目标的比例为 20%, 单通道听觉目标的比例为 60% 时, Colavita 视觉优势被减小。因此, 这也表明这种操纵可以成功地内源地将测试的注意力转移到更频繁发生的通道上 (Sinnett & Spence, 2007)。

## 1.6 多感官整合中适应效应对注意过程的影响

影响感知系统神经活动强度的一个重要机制是, 由刺激重复触发的自下而上的适应效应。当完全相同的刺激被重复时, 后出现的刺激会受到先前刺激的影响, 导致其诱发的神经激活减少, 这种现象被称为神经活动的重复抑制 (RS), 也就是适应效应 (Ringo, 1996; Sobotka & Ringo, 1994)。适应效应是一种稳定的现象, 在多个感觉通道中普遍存在, 持续时间从毫秒 (Sobotka & Ringo, 1996) 到分钟 (Henson, Shallice & Dolan, 2000), 甚至天 (Van Turennout, Ellmore, & Martin, 2000), 不同大脑区域中的适应效应持续时间不同 (Barron, Garvert &

Behrens, 2016)。

### 1.6.1 视觉通道中的适应效应

在视觉通道中，当同一视觉刺激重复出现时，参与者对该刺激的响应速度更快，错误率也更低 (Ferrand & Grainger, 1992; James, Humphrey, Gati, Menon, & Goodale, 2000; Schacter & Buckner, 1998; Tulving & Schacter, 1990)，且该刺激引起的视觉皮层神经活动的幅度在刺激重复后显著降低 (Grill-Spector, Kushnir, Edelman, Itzhak, & Malach, 1998; Malach et al., 1995)。例如，大脑皮层 MT 区域 (V5) 的神经活动由于视觉刺激在单一方向上的重复而显著减少 (Tootell et al., 1995)；初级视觉皮层的 V1 区域由于同一方向的光栅重复出现而产生适应效应，V1 对方向的适应效应随着光栅在正交方向上的方向逐渐变化而逐渐减小 (Tootell et al., 1998)；外脑皮质和下颞叶皮层 (IT) 由于相同物体的重复呈现而产生适应效应 (Buckner, Koutstaal, Schacter, Wagner, & Rosen, 1998)。

适应效应的大小受到诸多因素的影响，比如适应刺激的数量、适应刺激与检测刺激之间的关系 (Henson, Rylands, Ross, Vuilleumeir, & Rugg, 2004; Henson, Shallice, & Dolan, 2000; Sayres & Grill-Spector, 2006)。例如，Grill-Spector 和 Malach (1999) 在一项关于视觉适应的功能性磁共振成像 (fMRI) 研究中发现，视觉系统侧枕叶皮层 (LOC) 的信号强度随着刺激重复的次数增加而减少，并在适应刺激重复 6-8 次后减少到最小。在视觉适应的研究中，研究者通过比较在相同适应刺激情况下不同检测刺激对适应效应的影响，探究了视觉刺激的认知处理过程的神经机制 (Grill-Spector et al., 1998; Kourtzi & Grill-Spector, 2005; Winston, Henson, Fine-Goulden, & Dolan, 2004)。例如，Grill-Spector 通过一次只改变探测刺激的一个属性，而保持其他属性与适应刺激一致，研究了视觉刺激的大小、位置、视角和亮度对适应效应的影响。研究发现，侧枕叶皮层后部神经元的适应效应会因视觉大小和位置的变化而减弱，而侧枕叶皮层前部神经元的适应效应没有改变。此外，在面部处理的适应效应研究中，发现当重复呈现具有相同身份和不同表情的面孔时，颞下回后部和颞上沟后部的 BOLD 信号会减弱。当重复呈现具有相同表情和不同身份的面孔时，颞上沟中部的 BOLD 信号会减弱。这表明大脑中处理面部身份信息 and 表情信息的区域存在神经解剖学上的分离，负责身份信息编码的区域位于颞下回和颞上沟的后部，而负责表情信息编码的区域位于颞上沟的中部 (Winston et al., 2004)。

除了脑成像研究之外，神经电生理学研究也发现了视觉适应效应。这些研究主要关注视

觉适应效应中事件相关电位 (ERPs) 的时间进程和信号强度 (功率) 。随着时间的推移, 由视觉适应效应引起的 ERP 变化通常在刺激呈现后发生。

大约在 200 毫秒时 (Doniger, Silipo, Rabinowicz, Snodgrass, & Javitt, 2001; Gruber & Müller, 2005) 。在没有干扰刺激的情况下, 视觉适应效应最初产生在 160-190 毫秒; 而在有干扰刺激的情况下, 视觉适应效应被延迟, 直到刺激呈现后 200 毫秒才出现 (Doniger et al., 2001; Gruber & Müller, 2005; Schendan & Kutas, 2003) 。在信号强度方面, 重复刺激会削弱视觉皮层在伽马频段 (高于 40 赫兹) 的神经活动 (Gruber & Müller, 2005) 。最近对猴子进行的颅内电极研究进一步显示, 当视觉刺激重复时, 视觉皮层中神经元的放电率显著降低。同时, 在伽马频段 (40-90 赫兹) 中, 一级视觉皮层 V1 和高级视觉皮层 V4 之间的神经活动同步性 (一致性) 显著增加。在神经元放电率降低的情况下, 视觉皮层中神经活动增加的同步性维持了视觉信号的处理效率[9]。

### 1.6.2 听觉通道中的适应效应

在听觉通道中, Belin 和 Zatorre (2003) 率先使用 fMRI 技术研究由不同听觉刺激引起的听觉适应效应。实验分为两种条件: (1) adapt-Speaker, 即由同一个声音发出的 12 个不同音节的听觉; (2) Adapt-Syllable, 即由 12 个不同声音发出的相同音节的听觉。结果显示, 在声音适应条件下重复的声音引发的神经活动在右侧颞上沟前部显著减少, 与音节适应条件相比, 这表明右侧颞上沟前部对于个体声音的特征化是特异的。除了 fMRI 研究之外, 听觉通道的神经电生理学研究也显示, 听觉刺激的开始或结束会引发听觉 N1 成分, 这是一个具有大约 100 毫秒孵化期的稳定听觉诱发电位 (Näätänen & Picton, 1987) 。研究表明, 声音刺激之间的时间间隔越长, 诱发的 N1 成分的幅度越大 (Hari, Kaila, Katila, Tuomisto, & Varpula, 1982; Okamoto, Ross, Kakigi, Kubo, & Pantev, 2004; Rosburg et al., 2010) 。

当声音刺激重复时, 第一个声音刺激引发的 N1 成分的振幅最大, 随后由后续声音刺激引发的 N1 成分的振幅显著减少 (Budd, Barry, Gordon, Rennie, & Michie, 1998; Fruhstorfer, 1971; Fruhstorfer, Soveri, & Järvilehto, 1970; Okamoto & Kakigi, 2014; Ritter, Vaughan, & Costa, 1968) 。

最近的一个听觉通道 EEG 研究, 使用纯音作为听觉刺激, 通过操纵适应刺激的呈现长度和适应刺激的数量, 研究了适应刺激的时间特性对人类听觉诱发电位的影响 (Lanting, Briley, Sumner, & Krumbholz, 2013) 。结果显示, 听觉适应主要受到适应刺激持续时间的影响, 持续时间越长, 适应效应的量越大。同时, 适应刺激呈现的次数越多, 适应效应减弱得

越快。此外，听觉 P2 成分比 N1 成分更容易受到适应刺激的影响 (Lanting et al., 2013)。

### 1.6.3 多感官整合过程中的适应效应

在多感官整合过程中，神经活动在刺激之前不仅影响单一通道的刺激处理，还影响感官通道的优势效应的结果。最近的一项研究利用多感官整合中的视觉优势效应，一种 Colavita 效应的变种，来检验这一现象。其研究范式与经典范式的区别在于：（1）明确告知参与者实验中存在一定数量的双通道测试；（2）在两通道测试中，如果参与者同时听到和看到，他们被要求同时按下两个反应键。实验结果显示，在双通道测试中，参与者首先按下视觉键然后按下听觉键的测试比例显著大于先按下听觉键再按下视觉键的测试比例，表明视觉优势效应的存在。此外，以这种方式进行的平均测试数量，包括先按下视觉键再按下听觉键和先按下听觉键再按下视觉键的数量，都大于 40，这允许进行足够的测试来检查 Colavita 视觉优势效应的神经机制，并为调查在刺激呈现之前神经活动的变化对跨通道优势方向的影响的神经机制提供了可能性。进一步的研究分析了刺激前神经活动对感官优势方向的影响。结果显示，前额叶神经活动的增强和默认网络神经活动的减弱可以预测当前双通道测试的视觉优势效应，例如左侧额下回 (inferior frontal gyrus, IFG)，与当前测试相比，当视觉优势低 (Trial N) 和视觉优势高 (Trial N) 时，前一次试验 (Trial N-1) 的 BOLD 反应显著更高；默认网络 (如后扣带回) 比当前试验视觉优势低 (Trial N) 时的负激活更强，并且在前一次试验 (Trial N-1) 视觉优势高 (Trial N) 时也是如此。

此外，研究还发现，在刺激呈现之前听觉皮层中的神经活动增加可以预测听觉优势的大小。例如，左侧颞上回 (superior temporal gyrus, STG) 对前一次试验 (Trial N-1) 的 BOLD 反应明显高于当前听觉优势低的试验 (Trial N)。相比之下，当前试验在之前试验中视觉优势高 (Trial N) 和视觉优势低 (Trial N) 时，并没有 BOLD 信号强度的显著差异 (图 2-6C 右下)。先前的研究表明，前额叶神经活动的升高和默认网络神经活动的减少增强了额下回与初级和次级视觉皮层以及背侧视觉通路之间的神经连接 (Miller & Cohen, 2001; O'Reilly, Herd, & Pauli, 2010)。结合以上结果，一方面，额下回 (IFG) 的自上而下调节促进双侧枕叶回 (MOG) 的视觉处理，以编码高质量的感知特征；另一方面，背侧视觉通路的神经活动得到增强，这提高了视觉表征进入感觉运动特征化的效率，并最终增强了视觉优势效应。听觉优势效应是由于刺激前听觉皮层中的神经活动增加，从而增强了听觉皮层与感觉运动区域之间的神经连接。因此，多感官通道竞争的结果取决于感觉皮层与额感觉运动皮层和默认

网络之间的动态相互作用。

## 2. 提出的问题

随着多感官整合研究的深入，越来越多的研究者不仅研究多感觉整合现象本身，还开始关注注意力在多感官整合中的作用。一些研究认为多感官整合现象是自动化的，顶层注意力不影响多感官整合过程 (Vroomen, Bertelson, & de Gelder, 2001)；也有观点认为在双通道刺激同时得到注意的相对于非注意条件下会发生多感官整合 (Talsma & Woldorff, 2005)。因此，研究者们认为注意力可以影响多感官整合。然而，大多数早期提出多感官整合不受顶层控制影响的研究侧重于空间注意 (Vroomen et al., 2001)，他们认为影响多感官整合的研究只关注了注意与非注意条件的区别，忽略了注意内部的区分。Mishra 等人的研究 (2007) 发现声音诱发的闪光幻觉在第二个听觉声呈现后的 30-60 毫秒产生了视觉皮层活动范围的早期调制 (Mishra et al., 2007)。Mishra 等人的研究 (2010) 进一步探讨了刺激位置 (视野的高低) 与注意 (注意与不注意) 之间的互动对声音诱发的闪光幻觉的影响。结果显示，PD120/110 成分 (在低视野) 位于腹侧枕颞部条状视觉皮层，而随后的 PD180 和 ND250/240 成分 (在低视野) 位于上颞回 (枕皮层和上颞回被认为是多感官通道信息互动的脑区)。这些结果说明声音诱发的闪光幻觉并非都是自动整合处理的结果，而是顶层注意力参与了视听信号的错误整合 (Mishra et al., 2010)。此外，Kamke 等人 (2012) 利用经颅磁刺激 (TMS) 研究了注意相关皮层神经网络在声音诱发视觉闪光幻觉现象中的作用。结果显示，当 TMS 作用于右角回 (AG) 时，报告闪光幻觉的参与者比例显著减少。因此，他们认为破坏的角回增强了参与者在多感官事件中的真实感知，有效减少了无关听觉刺激对视觉事件处理的影响，减少了听觉和视觉刺激错误整合的可能性。然而，当呈现单一视觉刺激或视听觉刺激时间间隔长时，TMS 并不改变参与者感知。因此，角回涉及调节声音诱发的闪光幻觉中视觉和听觉刺激的整合，但该研究仅限于分裂幻觉现象 (Kamke et al., 2012)。

此外，最近的研究表明，注意力除了可以指向空间位置，还可以指向某个感官通道 (Talsma, 2015)。而且，Wilschut, Theeuwes, 和 Olivers (2011) 也认同在认知处理层面，注意力系统包含了定向 (空间转移) 和选择 (通道选择是其中之一) 等组成部分。在行为表现上，对通道的注意会削弱参与者在未被注意的通道中的信息处理能力，而在被注意的通道下的信息处理则得到增强。此外，根据偏向竞争模型的假设，选择性注意力会增强对选定信息的感官神经反应，并抑制与之无关的反应 (Mishra & Gazzaley, 2012)。因此，当注意力集中在一个感官通道上时，该通道内信息的神经反应得到增强，而被忽略通道内信息的神经反应

则被抑制。在这项研究中，运用了三项研究来研究基于通道的注意力对声音诱发的闪光幻觉的影响（见图 3-1），这些研究聚焦于在调节通道试验比例时产生的内源性注意力，预刺激注意通道的适应性，以及集中和分散的通道注意力，其目的是为了系统地调查通道注意对声音诱发闪光幻觉中的听觉优势现象的影响。为了能够为注意力影响的多感官整合过程提供来自通道注意的经验支持。

## 2.1 基于通道的内源性注意对声音诱发闪光幻觉的影响

自从 Shams 等人 (2000) 发现了声音诱发闪光幻觉现象以来，许多研究者开始研究影响声音诱发闪光幻觉的因素，研究者们认为自上而下的注意力可以影响声音诱发的闪光幻觉。Mishra 等人 (2007) 的研究发现，声音诱发的闪光幻觉在第二个听觉声呈现后的 30-60 毫秒产生了视觉皮层活动范围的早期调制 (Mishra et al., 2007)。Mishra 等人 (2010) 的研究进一步探讨了刺激位置 (视野的高低) 与注意 (注意与不注意) 之间的互动对声音诱发的闪光幻觉的影响。结果显示，PD120/110 成分 (在低视野) 位于腹侧枕颞部条状视觉皮层，而随后的 PD180 和 ND250/240 成分 (在低视野) 位于上颞回 (被认为是多感官通道信息互动的脑区)。这些结果说明声音诱发的闪光幻觉并非都是自动整合处理的结果，而是顶层注意力参与了视听信号的错误整合 (Mishra et al., 2010)。此外，Kamke 等人 (2012) 利用经颅磁刺激 (TMS) 研究了注意力相关的皮层神经网络在声音诱发视觉闪光幻觉现象中的作用。该研究主要使用了不同组合的视听刺激 (一个视觉刺激伴随两个听觉刺激，一个视觉刺激，两个听觉刺激) 以及不同视听刺激间隔 (70 毫秒和 160 毫秒)。结果显示，当 TMS 作用于右角回 (AG) 时，报告闪光幻觉的参与者比例显著减少。因此，他们认为破坏的角回增强了参与者在多感官事件中的真实感知，有效减少了无关听觉刺激对视觉事件处理的影响，减少了听觉和视觉刺激错误整合的可能性。然而，当呈现单一视觉刺激或视听觉刺激时间间隔长时，TMS 并不改变受试者感知。因此，角回涉及调节声音诱发的闪光幻觉中视觉和听觉刺激的整合，但该研究仅限于分裂幻觉现象 (Kamke et al., 2012)。

在最近的研究中，对注意力的内部差异化分析后，发现内源性注意力可以以通道注意力的形式影响多感官整合过程，在多感官整合中，对通道的注意会削弱未被注意通道中信息的处理，而被注意通道下的信息处理则被增强 (Spence, 2011; Talsma, 2015)。还发现，注意两个通道的成本高于只注意一个特定通道，这与在通道间的额顶区域或颞区时给予注意力相

关的更强烈激活有关 (Santangelo et al., 2010)。最近的研究还发现, 相较于只注意一个通道, 同时注意两个通道可以影响声音诱发闪光幻觉的现象, 表明分散注意力至多个通道可以影响声音诱发的闪光幻觉, 但只限于分裂幻觉。通道注意调节感官优势的一个可能解释是优先进入原则, 即两个刺激同时呈现时, 参与者首先注意到的刺激, 往往会被认为是先出现的。优先进入效应可能是由于参与者对视觉/听觉通道的内源性注意, 或者是由于视觉/听觉刺激本身外源地抓住参与者的注意力 (Posner et al., 1976; Turatto, 2002)。因此, 当参与者以内源性方式将注意力指向视觉通道时, 在双通道试验中, 参与者倾向于认为视觉刺激先于听觉刺激出现; 当参与者以内源性方式将注意力指向听觉通道时, 在双通道试验中, 参与者倾向于认为听觉刺激先于视觉刺激出现。许多研究使用了不同的方法来指向内源性注意力到其中一个通道, 最常见的使用方式是改变不同通道中的试验比例。例如, Sinnett 和 Spence (2007) 在一个实验中将单通道视觉目标的比例增加到 60%, 同时将单通道听觉目标的比例降低到 20% (剩余的 20% 的试验是双通道尝试), 来调查内源性通道注意力对 Colavita 视觉优势效应的影响, 实验结果发现 Colavita 视觉优势效应变大; 当单通道视觉对象的比例是 20%, 单通道听觉对象的比例是 60% 时, Colavita 视觉优势效应变小。因此, 这也显示出这样的操纵可以成功地将受试者的注意力内源地转移到发生频率更高的通道上。

在先前的研究基础上, 通过将注意力内源地分配给视觉通道或听觉通道, 检验内源性通道注意力是否会影响声音诱导的闪光幻觉的听觉优势效应, 即通过改变不同通道测试的比例, 将受试者的注意力内源地转移到多个通道的发生频率上, 检验内源性注意力是否可以影响声音诱导的闪光幻觉的听觉优势效应。研究采用了经典的声音诱导的闪光幻觉范式

(Shams 等人, 2002 年), 并分为 3 个实验, 实验 1 是基线实验, 即视觉通道测试、听觉通道测试和视听双通道测试的比例是 1:1:1; 实验 2 是对视觉通道的内源性注意力, 即视觉通道测试、听觉通道测试和视听双通道测试的比例是 3:1:1; 实验 3 是对听觉通道的内源性注意力, 即视觉通道测试、听觉通道测试和视听双通道测试的比例是 1:3:1。我们假设内源性注意力对视觉/听觉通道可以影响感官优势效应的方向, 从而改变声音诱导的闪光幻觉的听觉优势效应, 具体表现为当注意力内源地指向视觉通道时, 闪光幻觉效应减少; 当注意力内源地指向听觉通道时, 闪光诱导效应增加。

## 2.2 注意通道适应对声音诱导闪光幻觉的影响

自从 Shams 等人在 2002 年发现了声音诱导闪光幻觉现象之后, 许多研究者开始研究影响声音诱导闪光幻觉现象的原因和影响因素。研究者们发现, 声音诱导闪光幻觉现象的主要

原因之一是视觉敏感度。在 McCormick (2008 年) 的研究中, 使用信号检测理论来计算参与者在声音诱导闪光幻觉中的视觉感知敏感度。研究发现, 参与者在两声音条件下 (即, 当参与者感知到闪光幻觉时) 的视觉敏感度显著低于伴随一声音的一次闪光。在 Kumpik (2014 年) 的研究中, 也发现当参与者感知到幻觉性闪光时, 他们的视觉感知敏感度明显低于非幻觉条件。因此, 显示了声音诱导闪光幻觉现象是由于参与者的视觉感知敏感度降低, 使得参与者感知到幻觉现象。此外, 研究发现, 自上而下的注意力可以影响声音诱导的闪光幻觉。Mishra 等人 (2007 年) 的研究发现, 声音在第二个听觉声音呈现在 30-60 毫秒时诱导了闪光的幻觉, 这是对视觉皮层活动范围的早期调节效应 (Mishra 等人, 2007 年)。Mishra 等人 (2010 年) 的研究进一步发现, 声音诱导的闪光幻觉并非全部是由于自动集成处理的结果, 自上而下的注意力参与了视觉和听觉信号的错误集成 (Mishra 等人, 2010 年)。我们最近的一项研究也发现, 同时关注两个通道相比只关注一个通道, 可以影响声音诱导的闪光幻觉现象, 表明分散于通道的注意力可以影响声音诱导的闪光幻觉, 但仅限于分裂幻觉。

影响感觉系统神经活动强度的一个重要机制是由自下而上的刺激重复触发的适应效应, 即当完全相同的刺激重复出现时, 后续刺激会受到最初呈现的刺激的影响, 导致其诱发的神经激活度下降, 这种神经活动的重复抑制 (RS) 被称为适应效应 (Grill-Spector & Malach, 2001; Ringo, 1996; Sobotka & Ringo, 1994)。适应效应是一种稳定的现象, 在多个感官通道中普遍存在, 持续时间从毫秒 (Sobotka & Ringo, 1996) 到几分钟 (Henson, Shallice, & Dolan, 2000) 甚至几天 (Van Turennout, Ellmore, & Martin, 2000), 不同脑区的适应效应持续时间不同

(Barron, Garvert, & Behrens, 2016)。刺激前的神经活动不仅影响单一通道的刺激处理 (Lanting, Briley, Sumner, & Krumbholz, 2013), 而且还影响感觉通道的优势效应的结果。研究 (2013 年) 检验了刺激前神经活动对多感官整合中 Colavita 视觉优势的影响, 结果表明呈现的刺激预示了前额叶神经活动的增加和默认网络神经活动的减弱能够预测当前两通道试验的视觉优势效应。此外, 还发现刺激出现前听觉皮层的神经活动升高可以预测听觉优势的程度。

由于声音诱导闪光幻觉现象, 这可能是因为参与者的视觉敏感度降低, 再加上基于通道的注意力可以影响声音诱导的闪光幻觉现象。在先前研究的基础上, 通过操纵注意通道的适应现象, 考察了刺激出现前感官系统内部神经活动变化对感官通道优势方向的影响, 即基于通道的适应是否能够影响声音诱导的闪光幻觉现象。研究分为两个实验, 将适应范式插入到经典的声音诱导闪光幻觉范式之前, 实验 4 是基于视觉通道的适应对声音诱导闪光幻觉的影响, 实验 5 是基于听觉通道的适应对声音诱导闪光幻觉的影响。我们假设注意通道的适应可以影响感官优势效应的方向, 从而可能导致声音诱导的闪光幻觉, 即听觉优势效应的变化。

## 2.3 集中和分散通道注意力对声音诱导闪光幻觉的影响

随着多感官整合研究的深入,越来越多的研究者不仅研究多感官整合现象本身,也开始关注注意力在多感官整合中的作用。有研究表明,多感官整合是一个自动化的过程,不受来自注意力的自上而下控制的影响(Vroomen, Bertelson, & de Gelder, 2001);也有研究提出,多感官整合仅在双通道刺激被注意到的情况下发生,与未被注意到的条件相比(Talsma & Woldorff, 2005)。因此,人们认为注意力在多感官整合中起着重要的作用。然而,早期提出多感官整合不受自上而下控制影响的研究大多是关于空间注意力的(Vroomen et al., 2001);他们认为,关于注意力能够影响多感官整合的研究只关注了影响多感官整合的注意力和非注意力条件之间的差异,而忽略了注意力内部的区分。例如,近期的研究表明,除了空间位置,注意力还可以指向感官通道(Talsma, 2015)。此外,Wilschut, Theeuwes 和 Olivers (2011)也认为,在认知处理水平上,注意力系统包括定向(空间转移)和选择(通道选择是其中之一)等组成部分。在行为表现上,对空间位置的注意力可以增强参与者对该位置信息的感知;而对通道的注意力会削弱未被注意的通道中的信息处理,并增强受注意的通道中的信息处理。因此,我们假设多感官整合不受空间注意力影响的结论不适用于基于通道的注意力。根据偏向竞争模型的假设,选择性注意力增强了对选定信息的感官神经反应,并抑制了无关的反应(Mishra & Gazzaley, 2012)。因此,当注意力集中在一个感官通道上时,该通道内信息的神经反应得到增强,而被忽略通道内信息的神经反应被抑制,从而跨通道刺激的整合可能不会产生。

在过去,研究声音诱导闪光幻觉时要求参与者只需关注视觉通道的刺激来判断视觉闪光的数量,而来自听觉通道的声音刺激只是被动地呈现给参与者,甚至明确告诉参与者他们不需要关注声音刺激。基于以往类似的研究,研究3通过将注意力指向不同的感官通道,调查了不同通道的注意力集中和分散对多感官整合中声音诱导闪光幻觉现象的影响。该研究采用了经典的声音诱导闪光幻觉范式(Shams et al., 2002),一方面,使用对视觉单一通道的注意力作为基线任务,另一方面,要求参与者关注视觉通道以确定视觉闪光的数量,并要求参与者关注听觉通道来判断声音刺激的数量。例如,在实验6的分散条件下,参与者需要首先报告视觉闪光的数量。

在实验7中,除了实验6中要求的报告视觉闪光的数量之外,还增加了任务难度:参与

者还需要报告听觉声音的数量是否与视觉闪光的数量相匹配。这样的设计目的不仅是为了调查在一定的注意力资源条件下，主动关注听觉通道的声音刺激是否会分散注意力资源，从而导致声音诱导闪光幻觉效应的增强或减弱，并在此基础上，进一步探究不同任务难度对声音诱导闪光幻觉中听觉优势效应的影响。为了调查集中和分散注意力对声音诱导闪光幻觉现象的认知神经机制，使用 fMRI 技术探讨了集中和分散注意力对声音诱导闪光幻觉现象的认知神经机制。我们假设在主动关注听觉通道的条件下，声音诱导的闪光幻觉会受到注意力资源分配程度的影响。并且作为一个稳定的听觉优势现象，声音诱导的闪光幻觉不受任务难度的影响。

### **3. 内源性通道注意力对声音诱导闪光幻觉中裂变错觉的影响**

以前的研究表明，内源性注意力可以以通道注意力的形式影响多感官整合过程 (Talsma, 2015)，即，对通道的注意将削弱对未被注意通道中信息的处理，并且未被注意的通道中信息的处理得到加强。因此，当参与者以内源性方式将注意力指向视觉通道时，参与者倾向于认为在双通道试验中视觉刺激先于听觉刺激；当参与者以内源性方式将注意力指向听觉通道时，参与者倾向于认为在双通道试验中听觉刺激先于视觉刺激。Sinnott 和 Spence (2007) 将单通道视觉对象的比例提高到 60%，单通道听觉对象的比例降低到 20%，以研究内源性通道注意对 Colavita 视觉优势的影响，并发现 Colavita 视觉优势效应变大；当单通道视觉对象的比例是 20%，单通道听觉对象的比例是 60% 时，Colavita 视觉优势效应减小。在我们的研究中，通过改变...【不完整】

在通道试验中，将比例内源地转移到出现频率更高的通道上，并且是否内源性注意视觉通道或听觉通道能影响声音诱导闪光幻觉的听觉优势效应。我们的结果发现，当注意力内源地指向视觉通道时，裂变错觉的数量显著减少；当注意力内源地指向听觉通道时，裂变错觉的数量增加。因此，我们可以论证，通过操纵通道试验的比例，我们也可以成功地将测试的注意力转移到出现频率更高的通道上。

通道注意调节感官优势的一个可能解释是优先进入的原则，即两个刺激同时呈现，首先被注意到的参与者倾向于认为它们先出现。优先进入效应可能是由于参与者对视觉/听觉通道的内源性注意，或者可能是由于视觉/听觉刺激本身外源地抓住了参与者的注意力 (Posner 等, 1976 年; Turatto 等, 2002 年)。因此，当注意力被指向某个感官通道时，对

这个通道中信息的神经反应被增强,而对被忽略的通道中信息的神经反应被抑制。声音诱导闪光幻觉现象是多感官整合中的一种听觉优势现象,涉及视觉通道和听觉通道,在本研究中,当注意力内源地指向视觉通道时,视觉通道中信息的神经反应被增强,听觉通道中信息的神经反应被削弱,行为结果表现为参与者判断视觉闪光数量的正确率得到提升,然后裂变错觉效应减少;当注意力内源地指向听觉通道时,听觉通道中信息的神经反应被增强,视觉通道中信息的神经反应被削弱,行为结果是参与者判断的视觉闪光数量更容易受到听觉声音刺激的干扰,这使得被试反应的正确率降低,进而表现为裂变错觉效应增加。

#### **4. 通道适应对声音诱导闪光幻觉中裂变错觉的影响**

以往的研究显示,在视觉通路中,当相同的视觉刺激重复出现时,参与者对该刺激的反应速度更快,错误率也更低 (Ferrand et al., 1992; James et al., 2000; Schacter et al., 1998; Tulving et al., 1990),而且这种刺激引起的视觉皮层神经活动的幅度在刺激重复后显著减少

(Grill-Spector et al., 1998; Malach et al., 1995)。本研究的实验 4 结果显示,视觉闪光适应的刺激数量会影响裂变错觉,即对于 F1B2 情况,视觉闪光适应的数量相同时,准确率是适应状态的两倍,这与以前研究的结果一致,表明当相同的视觉刺激重复出现时,参与者对该刺激的反应更快,错误率更低。然而,融合错觉不受视觉闪光适应数量的影响,即对于 F2B1 条件,一次视觉闪光适应与两次适应之间没有显著差异。这一结果表明,本实验中使用的经典声音诱导闪光幻觉范式 (Shams et al., 2000) 是有效的,视觉适应条件只能在一定程度上影响声音诱导闪光幻觉现象,不能消除这种现象。我们推测,裂变错觉可能受到适应过程的影响,这是因为适应刺激与检测刺激之间的关系以及适应过程对视觉通道敏感性的变化。具体来说,首先,先前的研究显示适应效果受到适应刺激与检测刺激之间关系的影响

(Grill-Spector et al., 1999; Kourtzi et al., 2005; Winston et al., 2004),适应刺激与检测刺激的物理属性一致时,参与者的正确率高,错觉的量小。其次,视觉闪光适应两次比适应一次占用更多的注意力资源,这削弱了视觉通道的敏感性,所以在裂变错觉中,存在适应两次的错觉量大于适应一次的现象。融合错觉不受视觉闪光适应过程的影响,这可能是由于融合错觉的不稳定性所致。

## 5. 集中与分散的通道注意力对声音诱导闪光幻觉中的裂变错觉效应的影响

以往的研究显示，除了对空间位置的关注，注意力还可以指向一个感官通道 (Talsma, 2015)。这表现为当注意力集中于一个通道时，参与者不关注的通道中的信息处理会减弱，而被注意的通道中的信息处理则会增强。因此，当注意力集中在一个感官通道上时，对该通道内信息的神经反应会被增强，而对被忽略通道内的信息的神经反应则会被抑制。在主动关注听觉通道的声音刺激的条件下，表明对裂变错觉的效果减少，这与先前的研究结果一致，表明与集中单一的视觉通道相比，当注意力分散到视觉和听觉两个通道时，参与者会影响声音诱导闪光幻觉中的裂变错觉的大小，但它不会影响融合错觉以增强或减弱其效果。该研究结果显示，无论是在基于视觉的单一通道注意条件下还是在视听双通道注意条件下，被试更容易产生裂变错觉而不是融合错觉，而且裂变错觉更容易受到注意力资源分配程度的影响。

# 第3节 恢复效应的声音和视觉环境元素分析

## 1. 背景调查

由于现代化进程的加速和经济的快速发展,人们在快节奏的生活中常常感到疲倦和压力。同时,研究显示疲劳和心理压力之间有很大的相关性。

疲劳和压力对人们身体健康的影响分为生理和心理两个方面。一些外部因素,如噪音、繁重的工作负担、单调等,容易引起身体疲劳和压力,同时也会导致人们产生一些不适感,例如疲劳、嗜睡、不适、情绪不稳定等,进而降低人们的判断力、记忆力、注意力和工作效率;长期的疲劳和压力会导致免疫功能下降,甚至引发精神疾病,也可能导致不必要事故,从而带来更严重的社会危害。

同时,我们的日常工作和学习环境充满了各种声音,如交通噪音、对话、鸟鸣等。从环境保护的角度来看:任何妨碍人们正常生活、休息、学习和工作,以及干扰人们想要听到的声音,使人不适,产生厌恶感的声音都是噪音。对人来说,噪音无论在生理上还是心理上都有严重的危害。研究表明,人们长期暴露在过多的噪音环境中会引起一定的心理问题,如增加人们的烦恼感,降低人们的认知能力、理解能力等,甚至会引起严重的病理变化,如听力损伤等。而且噪音通常还会加剧人们的疲劳和压力。只有在具有适宜的声学舒适度的环境中,人类的疲劳和压力才能得到缓解,即具有恢复效应。因此,如何提高环境的声学舒适度成为了一个关键问题。

我们通常采取被动的方式来应对声音,常用的被动处理方法是减少噪音危害,即通过吸音和隔音来达到降噪的目的,通过降低室内外环境的声压级来维持安静的环境。如今,人们对声环境的舒适度提出了更高的要求。当前,日常生活中噪音源的类型逐渐增多,应对噪声变得越来越困难,仅仅维持低声压级以实现降噪,使得被动噪声控制技术难以应对声环境恶化的趋势。同时,研究发现,单纯依靠降低声压级的手段,并不一定能提高环境的声学舒适度。

当声压级降低到一定数值时,对场所的声音舒适度评价将不再仅依赖于声压级,而往往会受到其他因素的影响,如声源的类型、个人特征等,这些因素在评价网络中将扮演重要的

角色。

由于环境的声学舒适度与许多因素相关，我们不能仅从声压级的角度被动地考虑，而应尝试采用多角度的主动处理方法。

“声音景观”（soundscape）的概念在这方面做出了巨大贡献，其核心思想是将声音作为一种积极因素来对待，并借鉴景观处理的方式来处理声音。“声音景观”是“声学”和“景观”两个概念的综合体，更加注重声音、景观与人们主观感受之间的关系。

这改变了传统噪声研究的思路，不再简单地将声压级作为噪声控制的目标，而是关注各声源之间的平衡以及声音、环境与人之间的和谐关系，并更加注重人们对声音的主观听觉感受。

研究表明，环境能缓解人们的疲劳和压力，即具有一定的恢复效果，特别是当人们处于他们喜欢的环境中时。至于声音是否能缓解人们的疲劳和压力，即是否具有恢复效果，目前还没有系统的研究。需要系统地深入探索不同声音类型及其不同呈现方式的恢复效果。

研究已经证实了视觉和听觉之间的相互作用，这两种感官可以相互影响并加强彼此的效果。也就是说，一个舒适的视觉环境可以减少某些声音引起的主观烦恼感，从而改善声音环境。视觉和听觉互动对恢复效果的影响尚待探索。

目前，人们在日常生活中面临的疲劳和压力主要来自工作方面，由于激烈的竞争和繁重的工作任务，办公室员工的疲劳和压力尤为显著。现代的办公空间普遍采取开放式办公的做法，因此本研究以开放式办公空间为例，探索具有恢复效果的声音和视觉元素。目的是能够以匹配建筑空间设计的方式，在开放式办公空间中植入声音和视觉元素，帮助人们缓解疲劳，减少压力，并发挥相应的恢复作用。

总的来说，关于声音的“恢复效应”的系统研究尚未进行。在办公空间的设计中，通常只考虑到视觉形象的减压和缓解疲劳作用，并没有系统地关注声音以及视觉和听觉互动对人们生理和心理的恢复效应。

基于这一研究背景，我们对不同类型的声音进行详细研究，探索它们在缓解和恢复人类疲劳与压力方面的作用，找出它们物理特性的共性，并从视听互动的角度考察声音在实际办公空间中的作用，为办公空间的设计提供合理化建议，希望能够提出有效的声音环境设计方法，从而改善人们对办公空间的整体感受。

## 2. 国内外理论研究综述

### 2.1 国内外恢复性环境相关研究动态与趋势

恢复性环境的研究起源于国外，早在 1990 年代，美国密歇根大学的两位学者塔尔伯特 (Talbot) 和卡普兰 (Kaplan) 首次提出了“恢复性环境”的概念。他们对现场的被试进行了为期两周的观察，主要观察对象是被试的心理状态，并且经过研究发现，现场环境可以对人的精神疲劳起到一定的恢复效果。

1991 年，芬兰坦佩雷大学的心理学教授卡莱维·科尔佩拉 (Kalevi Korpela) 研究了个人偏好的环境种类的恢复效应。他选择了 100 多名青少年作为参与者，在实验期间，他们被要求写下他们在日常生活中最喜欢的环境和地点，结果显示，受试者最喜欢的是家和自然环境，因为他们在家和自然环境中感到非常放松，所以疲劳和注意力能在一定程度上得到恢复。研究表明，个人偏好的环境属于恢复性环境，并且具有恢复的作用。

上述研究发现，户外环境和个人喜欢的环境具有一定程度的恢复作用，这表明并非所有环境都能被称为“恢复性环境”，具有“恢复”作用的地点和环境需要具备必要的特征。卡普兰 (Kaplan) 总结了一个“恢复性环境”必须具有的四个因素：离开 (Being Away)、延伸性 (Extent)、魅力 (Fascination)、兼容性 (Compatibility)。离开意味着离开日常生活的地方，进入一个新奇、不同的环境；延伸性，意味着人们居住的新环境具有一定的广度，其容纳能力足够大，个体可以在其中自由地、不受拘束地活动，不必担心走出界限；魅力意味着环境必须足够有吸引力和吸引人，以确保个人能够长时间保持注意力；兼容性，人们所处的环境与当时的心情应当是一致的、协调的、和谐的。只有在这样的环境中，人们的注意力才会一定程度上得到恢复，疲劳也会相应得到缓解。

1996 年，卡莱维·科尔佩拉 (Kalevi Korpela) 和特里·哈蒂格 (Terry Hartig) 选择了校园内七十多名学生作为实验参与者，对恢复性环境的因素和特征进行了相关研究。研究显示，个人偏好的环境由于不同因素，其恢复特性也有所不同，相应的恢复强度顺序为：“延伸性”、“兼容性”、“远离”和最后的“魅力”。

2001 年，卡莱维·M·科尔佩拉 (Kalevi M Korpela) 和特里·哈蒂格 (Terry Hartig) 等人探讨了“个体”、“偏好”和“恢复性”的关系，并进一步展示了人的恢复性体验、自我调节以及对偏好环境的适应之间的相互关系。研究表明，偏好的环境与不受喜欢的环境在恢复特性上有显著的差异，而高度偏好的自然环境具有典型的恢复效果；与“迷人”和“有趣”

的特性相比，偏好的环境和体验更多地与自我调节相关的特征相关。

2005年，意大利学者丽塔·贝托（Rita Berto）设计了一套主观评估实验，以验证恢复性环境能否帮助人们从精神疲劳中恢复。首先为参与者创建一定的注意力疲劳效应，然后呈现某些视觉刺激，包括恢复性环境、非恢复性环境和几何图形，实验结果显示除了恢复性环境组外，其余两组的注意力得到了改善，从而验证了假设；她还撰写了关于卡普兰的文章，“自然的恢复益处：朝向一种综合性框架”。在此中，注意力恢复理论得到了测试。

2008年，为了探究在其偏好地点拥有恢复性体验的决定性因素，芬兰的科尔佩拉·卡勒维（Korpela Kalevi）等人对两个城市的受访者进行了调查，并发现了以下重要因素：连续或频繁地接触到喜爱的地方，以及个人在生活中关于自然体验的经历或背景，如对自然的热爱，或童年生活经验。还有一些社会生活和压力中的情境因素。同时，研究发现，在偏好的环境中，恢复的强度和体验时间的长度是相关的。

在中国，“恢复性环境”的研究还处于初级阶段，主要是基于国外的研究成果，并在此基础上进行拓展。

2010年，北京师范大学发展心理学研究所的苏谦和辛自强对复合环境研究的理论、研究方法和进展给予了详细的介绍。

2010年，西南科技大学的研究人员石明达、闫文静、石冰冰等人向50名参与者展示了自然环境和城市环境的照片，并让受试者根据自己的环境感知以恢复性量表的形式评估其恢复性程度。相对于城市环境，人们对自然环境的偏好更高，恢复效果也更强。同时，对卡普兰夫妇的注意力恢复理论中的恢复性环境的四个因素进行了深入研究，并指出了提高城市环境恢复效果对这四个因素的响应方法。

2010年，谭家伦、唐杏芬等人从一个新的角度对“恢复性环境”进行了研究，主要是为了调查旅游环境中的压力恢复力。他们讨论了乡村游客的生活压力、感知休闲调节策略与健康之间的关系，研究表明休闲调节策略可以缓解生活压力对个人健康的不利影响。

2014年，南京大学的郭永瑞、张杰等学者也开始研究旅游环境的恢复作用，主要通过社会调查分析和测试旅游者的恢复性环境感知方法，验证了恢复性环境感知量表的一致性、兼容性、新颖性、逃避性、迷人性这五个维度，结果显示它具有良好的信度和效度。同时，旅游者对复合环境的心理感知过程是逐步的，如果旅游环境整洁有序、和谐、一致且新颖，旅游者更容易在环境中被迷住，产生远离生活、逃避现实的感觉，进而实现身心健康以及疲劳压力的恢复效果。

2012年，中国民族大学心理学博士池莉萍进行了一个关于日常典型环境恢复力顺序的

研究，以 200 多名初中生为研究对象，以特殊性、偏好性和地方依恋作为“感知恢复量表”的校准效度，研究显示效度指标合理，可以用于实证研究。同时，比较了四种环境的恢复力差异，结果显示恢复力的顺序是家庭、自然环境、学校和文化环境。

同年，池莉萍具体考察了注意力与恢复之间的关系，研究了青少年依恋环境的情绪启动和注意力恢复功能，并选取 90 多名初中生作为参与者，通过主观实验探讨了青少年依恋地点对情绪和注意力恢复的差异。研究发现，如果参与者对某个环境产生了依恋，他们的注意力在一定程度上得到了恢复。最终实验证明了依恋环境有助于注意力的恢复。

实际上，在 2011 年，北京林业大学的陈聪、赖映辉和吴建平也讨论了注意力的恢复，并研究了不同环境中有意注意力和反思注意力的恢复情况。研究表明，自然环境在有意注意力的恢复属性上比其他人造环境发挥更高的作用。

## 2.2 国内外声音景观的动态与趋势

早在 1929 年，芬兰地理学家格拉诺就提出了“声音景观”的概念。加拿大音乐教育家莫雷·沙弗 (Murray Schafer) 在 20 世纪 60 年代末首次提出了“声景观”这一概念。与传统的声学概念不同，他致力于寻找一种新的方式，使环境声学成为一种积极的研究，从而在声学中加入了景观的元素，将“声学”和“景观”整合起来，并更加关注声音、景观与主观人类情感之间的关系。

其他一些外国学者也对声音景观进行了一些研究，主要是为了对其进行评估，例如评估城市声音环境对健康的影响以及对景观欣赏价值和体验价值的影响。随后，需要使用环境心理学的统计分析方法来量化相应的评估参数，然后得出最终的研究结论。

英国谢菲尔德大学建筑学院的教授姜健对城市公共开放空间中的声音景观概念进行了详细研究，并提出了声音景观的三个要素，主要包括声音、听众和环境。同时，在“声音景观”研究范围内，“声音”的研究不仅包括人们愿意听到的声音，也包括他们不想听到的声音。

这涉及到声音的分类，实际上，在巴西亚马逊对声音景观的研究中，声音被划分为三个类别，即自然声音、技术声音和与人类相关的声音，声音景观中不同类型的声音会产生不同的效果。同时，发现声音给人的印象会根据声音的呈现方式而有所不同，比如频率、起止变化、顺序等。例如，研究发现声音持续时间越短，人对声音的感觉就越尖锐。

对于听众来说，在很大程度上，由于其即时状态，即其在环境和空间中的目的和行为，将会影响他们对声音景观的评价。不同的人，在带有不同目的和行为的声景观中，他们对

声音景观的舒适度感知是不同的，因此他们也会相应地给出不同的声音景观满意度评价。

环境空间的形状和封闭程度、空间界面的颜色和材质以及其中的景观元素，如绿化，都会影响声音的舒适度，并相应地影响对声音景观的主观评价。

在进行了有关声音景观的相关讨论之后，姜健建立了一个声音景观的预测、描述、评估和设计系统。

从国内角度来看，中国关于“声音景观”的研究并不多，在近几年有一些研究进展，并取得了一些成果。

浙江大学建筑学系的葛剑教授对城市公园的声音景观及其设计方法进行了相关讨论。她以日本的佐贺森林公园为例，将声音景观设定的概念作为一个新的切入点，融入到公园景观的研究和设计中，并最终提出了公园声音景观的配置要素，以及设计程序和处理方法。

华南理工大学的吴朔萱教授和袁晓梅教授对具有中国文化特色的声音景观进行了研究，并从中国古典园林文化的角度探讨了古典中国园林的声音景观。研究发现，与西方物理声学的声音景观不同，古典中国园林的声音景观基于自然声和人文，更加直观、感性和诗意，涉及到诗歌、文学、绘画等多种传统艺术形式，因此在声音景观的设计要素方面具有更多的传统中国文化的人文特性。

## 2.3 国内外开放式办公空间研究趋势与动态

20世纪50年代初，著名的德国学者 Quickborner Teams 首次提出了开放式办公的概念。随后，开放式办公模式在各个国家迅速发展，中国的开放式办公模式也越来越普遍，由于建筑结构形式的变化，框架结构和剪力墙结构的普遍应用为开放式办公空间的实现提供了技术支持，目前新的办公空间基本上都是开放式办公空间。

其空间形态的特点在于打破了传统每个工作位置之间的墙壁界限，只用简单的隔板或家具进行分隔，视线透明，整体整洁有序。

开放式办公空间形态自由，对不同办公类型具有良好的适应性，平面和空间形式灵活多样，一般包括蜂窝型、群组型、俱乐部型等。其中，蜂窝型是最典型的办公空间形式，其特点是整个办公区作为一个均匀的单元布局，单元分布均匀，具有强烈的规律性，单元间有走廊连接，人员密度很大，空间隐私性较低。

开放式办公空间之所以逐渐取代传统的封闭办公室，并且迅速发展、被广泛使用，正是因为它具有以下优点：

易于管理，有利于管理者指导和监督员工，了解员工的工作状态；同时，它也便于管理

人员与员工之间以及员工之间的联系，有利于增强办公人员之间的互动和交流。

节省空间，提高室内空间的综合利用率，可以容纳更多的人员和设备，进而降低建筑成本。

实现设备和能源共享，可以提高设备使用效率，并在供电、照明、供暖等方面节约能源消耗。

在空间形态上，开放式办公空间宽敞，扩大了室内外景观范围，视野更加开阔，增加了办公人员与空间环境之间的沟通与互动，也活跃了办公环境的氛围。

尽管如此，我们也不能忽视开放式办公空间中存在的一些问题：开放式办公空间在视听感知和其他方面打破了传统封闭办公室的隔离，但同时也破坏了空间的隐私性、领域感和安全感，如语言隐私，这对用户的心理产生负面影响；同时，由于空间的统一性，也会导致空间的单调性。

同时，开放式办公空间的声学环境问题也越来越突出，没有封闭隔断的自由透明空间，再加上人员密度高，无疑会造成噪声干扰。噪声干扰会让员工感到烦躁，导致对工作环境舒适度的评价下降，严重影响人的注意力，降低工作效率和工作热情。

由于噪声污染和人员干扰，再加上工作压力大，员工很容易产生不良的心理感受，即紧张、抑郁、疲劳等，甚至会引起头晕、眩晕、疲乏等身体不适。

这些问题可能会对办公室工作者在心理和生理上都产生负面影响。

有一项调查显示，北美超过 70% 的开放式办公室的人员每天都在工作，而在开放式办公环境中度过的时间占到 50% [三月]。由此可见，对办公室员工来说，大部分白天时间都是在办公室度过的，这严重影响了办公室工作者的身心健康。特别是在近几十年中，出现了“病态建筑综合症”的概念，它指的是人们长时间生活和工作在办公楼中引起的相应病理反应，如头痛、嗜睡、头晕、恶心等。当然，造成这种情况的原因有很多，噪声干扰是一个非常重要的因素，如果不加以适当控制，的确会对办公人员产生负面影响。

美国弗吉尼亚理工大学和州立大学的 Tonya L. Smith-Jackson 以及北卡罗来纳州立大学的 Katherine W. Klein 对开放式办公环境中无关语言干扰对家庭作业表现和心理负担的影响进行了研究，研究表明无关语言干扰会导致家庭作业表现的准确率下降，并造成心理负担的增加。

可以看出，不适宜的办公环境不仅危害员工的身心健康，而且还影响员工的工作效率，因此创造一个舒适的办公环境变得越来越重要。

众所周知，在安静的环境中人们更容易保持稳定的注意力，这要求将背景噪声控制在合

理的水平；然而，在安静环境中人类语言理解度的提高并不利于维护开放式办公室的隐私，所以保持办公环境的舒适性需要多种因素的协调。目前，在开放式办公空间，传统做法是综合运用吸音、阻隔和掩蔽技术来减少噪声污染，同时适当降低语言的易懂度，并提高办公空间的基础声学舒适性。

确实，虽然通过一些手段可以满足基本的声学舒适性，但单调的开放式办公空间仍然容易使人感到疲劳和压抑。随着社会经济的快速发展，人们更加渴望一个更加人性化的办公空间，希望能够享受现代办公的便利性，更能体会到人性化和尊重感。

实际上，早在二战后，“园林式办公室”的概念就在 20 世纪 50 年代诞生了。这种新型办公空间打破了传统办公环境的僵硬和冷漠，布局更自由，颜色材料更加多样和协调，最重要的是在工作环境中加入了景观元素，即工作区周围的空间和茂密的植被。通过增加人与自然的互动，不仅丰富了视觉内涵，而且增强了员工之间的情感交流，有利于改善更和谐的人际关系。

近年来，开放式办公空间的研究焦点也转移到了与心理舒适度相关的研究上，主要是检验不同类型的开放式办公对人们心理状态的影响。

瑞典 Gävle 大学的学者 Helena Jahncke、Staffan Hygge 等人进行了办公空间的恢复性研究。他们讨论了开放式办公环境中的认知表现和恢复情况，让参与者在模拟的开放式办公环境中，在两种噪音下工作，然后提供了四种恢复性条件（水声和图像、水声、静音、办公室噪音），研究表明水声和图像的恢复效果是最好的，其次是水声，办公室噪音的恢复效果是最差的。这项研究为开放式办公的恢复性方面提供了一定程度的可能性，并为开放式办公的设计提供了新的思路。

瑞典斯德哥尔摩大学的学者 Aram Seddigh 和 Erik Berntson 进行了有关办公空间与注意力之间关系的研究。这项研究探讨了注意力需求和办公室类型对健康和工作表现的影响，尤其是与分心、认知压力、精神疲劳和人格解体有关的指标。在人类效率和总体健康方面，研究探讨了工作中的注意力需求与六种办公室类型之间的交互作用。研究显示，在单间办公室中，人们的注意力分心和认知压力相对较好，而高注意力需求的工作如果在开放式办公环境中进行，将会导致员工分心和认知压力的增加。

同时，Aram Seddigh 和卡洛琳学院的 Cecilia Stenfors 对开放式办公空间与人们认知家庭作业表现之间的关系进行了深入的调查研究，他们主要观察了正常工作条件下以及在小型、中型和大型开放式办公环境中的记忆得分。研究表明，人们在小型开放式办公环境中的认知家庭作业表现要比在大型开放式办公环境中更好。

马来西亚学者 Shafaghat, A., Keyvanfar, A.等人从整体空间设计和舒适度的角度对可持续开放式办公空间设计进行了研究,以增强员工的舒适度和对工作效率的满意度。他们对一些开放式办公空间的设计元素进行了权重分析,例如灵活空间、听觉干扰等,并希望未来能够形成开放式办公的评价指标。

## **3. 声音偏好问卷调查**

### **3.1 问卷的目的和步骤**

#### **3.1.1 问卷调查的目的**

在日常生活中,人们会听到各种各样的声音,有些声音人们可以明显注意到,而有些则可能不会意识到;有的声音令人愉悦,有的声音则难以接受,也就是说,人们对不同的声音元素有不同的关注度和偏好。如果了解人们在生活中更加重视哪些声音,喜欢哪些声音,以及不喜欢哪些声音,不仅可以获得人们对不同声音的偏好信息,还可以为后续的研究内容和实验设置做准备,在正式实验之前更有针对性地收集声音信号作为正式实验的实验声音。因此,有必要找出不同人对不同声音的关注度和偏好程度。

为此,作为实验的初步调查,进行了一项关于声音偏好的问卷调查。

#### **3.1.2 问卷调查的步骤**

1) 设计问卷: 问卷中的问题符合简洁形式和有针对性内容的原则, 问卷包含了“关怀的声音”、“喜悦的声音”、“不喜欢的声音”、“想要听到的声音”等几个关键问题。

2) 分发问卷: 由于该调查采取问卷与访谈相结合的形式, 除了现场分发问卷和在线分发问卷之外, 还进行了现场访谈, 在与受访者交流的过程中进行现场记录。同时, 确保发放的问卷数量达到一定数额, 以便能回收足够数量的有效问卷。

3) 问卷结果分析: 对收集到的问卷进行分类, 剔除无效问卷, 整理、统计和分析有效问卷, 并使用数据分析软件如 Excel、SPSS 等来进行数据分析。

## 3.2 调查对象

因为偏好程度属于心理学范畴，不同年龄、不同性别甚至不同职业背景的人对声音的偏好程度有所不同，并且应该表现出明显的年龄和性别特征。为此，本次调查对象分为4个年龄组：0-20岁，21-40岁，41-60岁，以及61岁以上。为了更详细地了解人们的观点，本研究采用了问卷调查和深入访谈的方法，同时确保调查对象中男女比例大致平衡。

## 3.3 问卷内容

问卷包含7个问题，涵盖两种类型的问题：开放式问题和基于量表的问题。

开放式问题包含5个主题，其目的是了解人们在日常生活中关注和在意哪些声音，喜欢哪些声音，不喜欢哪些声音，以及哪些声音会增加体重感和缓解疲劳。需要受访者给出主观性的答案，对于不便现场作答的受访者则通过访谈方式进行记录。

标题：“你在日常生活中经常听到哪些声音？”目的是调查人们在日常生活中关注的声音，并明确人们对不同声音的关注程度；

标题：“在日常生活中，你听到的哪些声音是你喜欢/不喜欢的？”目的是调查人们对声音的偏好，确定人们喜欢/不喜欢什么样的声音，以及认为什么样的声音听起来舒适/不舒适；

标题：“当你感到疲劳时，什么声音可以帮助缓解/加剧疲劳？”目的是调查人们想要增加/想要移除的声音类型，并识别人们想要增加到他们环境中以帮助缓解疲劳的声音/移除什么声音以避免造成疲劳。

量表部分包含2个问题，分别是“当你感到疲劳时，听到某些声音能帮助缓解/加剧疲劳吗？”主要目的是调查在人们心中声音是否具有缓解或加重疲劳的效果，以及这种效果的程度。量表使用里克特五点量表，每个问题的答案选项为5个，分别是：完全没有，似乎有一点能力，比较能，相当能，特别能。要求受访者做出选择来回答。

开放式问题和量表类型问题以有序的方式交错组合。

有关问卷的详细信息，请参阅文章末尾的附录部分。

## 3.4 问卷调查概述

受访者将被划分为4个年龄组：0-20岁，21-40岁，41-60岁，61岁及以上，将为每

个年龄段分发问卷，以此确保受访者基本上涵盖所有年龄层，从而可以从整体和年龄段角度观察声音偏好的特征。

问卷在天津水上公园、天津文化中心、天津南江小学进行了分发，并通过互联网向开放办公室的员工进行了分发。

本次共发放了 540 份问卷，实际回收了 440 份，其中无效问卷 184 份，有效问卷 256 份，有效率为 58.18%。

### **3.4.1 有效问卷中受访者的年龄描述**

本次调查将被划分为 4 个年龄段：0-20 岁，21-40 岁，41-60 岁，61 岁及以上。从表 2-2 可以看出，受访者基本上涵盖了每个年龄段，0-20 岁的有 118 人，占总数的 46.09%；21-40 岁的有 73 人，占总数的 28.52%；41-60 岁的有 34 人，占总数的 13.28%；61 岁以上的有 31 人，占总数的 12.11%。

### **3.4.2 有效问卷中受访者的性别描述**

调查中共有男性 138 人，占受访者总数的 53.9%；女性有 118 人，占受访者总数的 46.1%，受访者的性别比例相对平衡。

## **4. 声音偏好分析的主观评价实验**

### **4.1 实验音调的获取**

在“人们日常生活中关注的声音”声音偏好问卷中，挑选了单一声音源（包括喜好与不喜好的声音）进行现场录音，而录制的限制性声音是从英国图书馆声音档案的项目数据库以及 BBC 声音图书馆中获取的。

录音地点选择：天津大学北洋广场、天津意式风情区、天津滨江道商业街等。录音时需要确保良好的录音条件，即天气条件较好、背景噪音较低、风速较小、人员较少以及避开较安静的场合，以避免噪声对录制声音的干扰。录音时间为五至十分钟，并且在后期制作中满足声音编辑的需要。录音设备包括高灵敏度单声道麦克风、EDIROL R-4Pro 高清录音器。

调查中共录制了 62 种声音，收集的声音经过分类筛选，并使用专业音频编辑软件 Adobe

Audition CS6 进行统一处理，将声音截取并拼接至相同长度的 20 秒。

由于本实验的目的主要是研究受试者对声音类型的偏好，不涉及声压级的因素，并且应排除声压级因素的干扰，因此需要将信号音的声压级调整至一致，以便作为实验声音。在预实验中，受试者普遍认为 55dB 较为舒适，所以在半消声室使用耳机播放，利用 Norsonic140 精密声音分析仪进行调制测量，使得 20 秒声音信号的等效连续声级 LAeq 为 55dB，即 LAeq,20s = 55dB。保存为实验音。经处理后的有效实验音为 38 段。

## 4.2 参与者的选择

选定的参与者均为天津大学的学生，所有人的年龄大约为 25 岁，他们的专业包括建筑学、规划学、土木工程、化学工程、材料学等。总共 30 名参与者，其中男生 15 名，女生 15 名；男女比例为 1:1。

为了确保参与者选择的公平性，防止无关因素对实验产生不良影响，所选的 30 名参与者在最近的工作学习生活状态、身心健康状况及工作态度上基本保持相同或相似。实验前，还需询问参与者最近的工作学习生活状态，如有过度劳累或对声音有反感等异常情况，则不允许参加实验。

## 4.3 实验场地和实验仪器

实验场地选在天津大学建筑学院的半消声室内，实验中将耳机与笔记本电脑连接，使用耳机播放声音信号。耳机的型号是：AKG K702 头戴式监控耳机，其频率响应范围为 10-39800Hz。

## 4.4 声音偏好分析的主观评价实验

### 4.4.1 实验流程

本次评价实验的目的是探究人们对不同类型声音的偏好程度，需要参与者对他们听到的声音进行七级主观评价，以评定声音的“好听”程度。

实验场地为天津大学建筑学院的半消声室，实验声音通过耳机播放，每个实验声音为 20 秒，实验总时长约为 13 分钟。对于每个实验声音，参与者需要在 20 秒内评价他们所听到的声音的偏好程度，评价采用七级语言量表，即从“喜欢”到“不喜欢”、“中性”、“有

些”、“相当”和“特别”七个级别。每个等级都给出一个分数，将“喜欢”作为正分，“不喜欢”作为负分，用于实验后的统计分析。

#### 4.4.2 实验结果的主观评价

实验结束后，统计参与者对声音偏好评价的得分，计算得出平均值，并根据偏好程度的平均分数对 38 个声音信号进行排序编号。

从建筑声音钻到喷泉的声音，被评价为“不快乐”，而从雷声到流水声，则被评价为“愉快”；其中，建筑声音钻是偏好度最低的声音，其偏好度被划分为-2.21，而流水声是最受欢迎的声音，其偏好度为 2.43。

本节中实验室的主观评价实验，以及上述声音偏好问卷调查的目的都是为了获得声音偏好的程度，两者之间存在一定的区别：

1) 问卷调查是由受访者通过想象和回忆来填写的，会受到现场许多因素的影响，如调查环境和受访者的情绪；而在半消声室进行的声音偏好分析的主观评价实验，则是在实验室环境下考察参与者的听觉状态，排除了其他许多因素的干扰；

2) 在问卷调查中，受访者回忆的声音通常是复合声源，并且同时存在背景噪声；而在实验室中参与者听到的实验声音是单一声源的声音信号；

3) 问卷调查是在现场进行的，能够更好地反映现场的听觉状态，相对来说更为真实；而在半消声室进行的主观评价实验中，由于隔绝了背景噪声，可能会存在听觉失真。

为了了解之前的研究结果与实验室实验是否一致，参照主观评价实验中对喜好与不喜欢的打分方法，通过以下设置对问卷结果进行打分：你喜欢的声音类型，在问卷中的得票数计为正分；你不喜欢的声音类型，在问卷中的得票数计为负分。

得分的主观评价实验和问卷的统计得分的相关性分析可以看出，两者之间存在正相关关系，皮尔逊相关系数为 0.688，显著性系数为 0.000，这是非常显著的，由此可以看出人们实际听到的声音与记忆中的声音是非常一致的。

因此，可以继续开展实验室主观评价实验，以研究声音的偏好性和恢复性。

本节分析了人们在日常生活中关心的声音的声音偏好的主观评价，探索了参与者对声音的偏好，获得了声音类型的偏好程度，并分析了问卷的统计得分和问卷的统计得分之间的相关性，两者之间的相关性非常显著，结果表明人们实际听到的声音与记忆中的声音是一致的。同时，本章为后续实验中声音的筛选准备了声音材料。

# 第4节 巴洛克时期听觉与视觉艺术的 比较研究

## 1. 课题的起源与意义

“巴洛克”作为一种特殊的艺术风格,在欧洲文化艺术史的发展中发挥了决定性的作用。巴洛克风格与欧洲古典主义风格相对立又共存,它更注重艺术的方法、形式和技巧,对艺术本身的回归和自由发展起到了重要作用,这不仅激发了浪漫主义的出现,也对20世纪现代主义的发展具有重要的启示意义。然而,如此重要的风格,自从其诞生之初,就被主流学术界贬低和嘲笑,巴洛克被认为是古怪的、夸张的和不规则的,直到19世纪80年代才逐渐得到肯定。20世纪初,以德国为首的艺术评论家对巴洛克进行了大量研究。由于历史原因,当中国学术界能够基本上对待巴洛克或在过去忽视它时,我们发现对它的了解非常少。虽然自1990年代以来一些研究成果逐渐发表,但巴洛克的研究可以说才刚刚开始,特别是对巴洛克时期各种艺术学科之间关系的研究尚未开始。本文试图通过比较巴洛克时期的听觉和视觉艺术得出结论,艺术风格是在特定的人文背景下受艺术潮流影响而形成的,不同艺术领域之间的关系不容忽视,它们之间既有共同的本质和特点。在这方面,巴洛克艺术尤为明显。能够正确把握巴洛克时期艺术之间的特殊关系,对我们更好地研究巴洛克音乐和其他艺术科目具有重大意义,它在我们未来的音乐教育工作中也起着重要的作用。

## 2. 研究对象与文献综述

艺术的分类方法有很多种,常用的主要有三种类型:第一,艺术形态的存在可以分为空间艺术、时间艺术以及空间和时间的艺术;第二,艺术形态的感知方式可以分为视觉艺术、听觉艺术、视听艺术以及形象艺术;第三,根据艺术形态的创作可以分为造型艺术、表演艺术、语言艺术以及综合艺术。本文主要研究以音乐为基础的巴洛克时期的听觉艺术和以绘画、建筑、雕塑为代表的视觉艺术作为研究对象,并比较两者之间的相似性或差异性。

对巴洛克艺术的研究出现在许多著作中。例如, Kenneth Clarke 的《艺术与文明:欧洲艺术与文化史》(易莹翻译), 雅克·德比克的《西方艺术史》(徐庆平翻译), 戈姆布里希的《艺术发展史》, 温妮·海德·米奈的《巴洛克与洛可可》, 蔡良玉、梁茂春的《世界艺术史》, 李绍林的《欧洲艺术史》等。这些作品大多将巴洛克视为一个具体的历史时期,

并对每个艺术学科的风格特征或最具代表性的艺术家和艺术作品进行了表面分析。在讨论各种艺术之间关系的著作中有王宏剑主编的《艺术概论》，但只有少数几句话，讨论较少。

在本部分的期刊论文中，与本主题相关的研究主要包括吴葳的硕士论文《巴洛克时期音乐与建筑共性的比较研究》（河南大学，2009年），黄馨和张磊的《从巴赫到库哈斯：建筑与音乐同步性研究》（《山西建筑》，2004年），梁宣城的《巴洛克艺术中的鲁宾斯与巴赫》（《文艺研究》，2010年）等。可以看出，这些论文主要是对巴洛克时期音乐与个别艺术之间关系的探索，唯一与作者论文标题相似的是高胜寒的《巴洛克视觉艺术与音乐艺术的形式比较》（《安徽文学》，2008年），但遗憾的是这篇论文篇幅较小，讨论不够透彻和深入。撇开“巴洛克”时期的定义，关于音乐与其他姐妹艺术关系的研究有100多篇论文，包括听觉艺术与视觉艺术的融合，视听艺术的辩证，音乐与建筑关系的探讨，以及音乐与绘画的共性和个性研究。

### **3. 巴洛克时期听觉与视觉艺术概述**

#### **3.1 巴洛克时期概述**

##### **3.1.1 巴洛克的起源**

关于“巴洛克”一词的起源，大致有两种理论。一种认为，“Barok”这个词来源于葡萄牙语“barroco”，指的是不规则珍珠的撞击。另一种理论则由法国思想家和作家卢梭提出，他认为“巴洛克”这个词来自意大利语单词“baroco”，这是一个形容词，用来指逻辑中三段论（演绎法）中牵强附会的论证。这两个词，“barroco”在葡萄牙语中和“baroco”在意大利语中的根源都是相同的意思，它们表示“复杂结构”或“装饰”。

如果我们试图从根源的含义上解释它，不难发现与之直接相关的词汇正好是17世纪初在罗马教皇统治下诞生的一种建筑形式，我们将其视为巴洛克艺术的原型，一种新艺术风格的发展趋势，但这种艺术风格在当时并未真正定型。

据说，“巴洛克”作为一个术语，首次被一个名叫布洛斯的意大利人在16世纪末用于建筑艺术，之后可以看到在意大利、西班牙、法国等国家，这个术语经常用于珍珠产品的相关文本中。也有说法，直到17世纪末，巴洛克最初应用于艺术评论，通常指的是各种非传统、怪异、偏离常规的事物，主要被建筑评论家用来评论当时的建筑。到了18世纪，它被

用作贬义词，通常指违反自然规律和古典艺术标准的情况。直到 19 世纪中期，它更多的是一个贬义而非艺术风格的名称。直到 1888 年，德国艺术家 H. Wölfflin 发表了《文艺复兴运动与巴洛克》一书，系统讨论了巴洛克风格，并且这种系统的整理与概括的研究充分肯定了巴洛克文化打破传统艺术的积极意义，从而确定了巴洛克作为一种艺术风格的概念。

巴洛克作为一个时期的概念可以在现有资料的记录中找到，且在艺术史的编纂中有所体现。除了雅各布·布克哈德 (Jacob Burckhardt) 在 1885 年的《意大利艺术鉴赏指南》(Guide to the Appreciation of Italian Art) 之外，还有上述的沃尔夫林 (Heinrich Wölfflin) 的《文艺复兴运动与巴洛克》(Renaissance and Baroque)，在这部作品中，他将“巴洛克风格”描述为历史现象，但在他的《艺术史的基本概念》(Basic Concepts of Art History) 一书中，他也尝试将巴洛克理解为一个跨历史的场景和一组判断特征的组合。雷蒙德·努安西奥 (Raymond Nacientes) 在 1895 年将 17 世纪的文明特征概括为巴洛克主义 (Baroqueism)，从而，学术界逐渐形成了将 17 世纪视为巴洛克时代的概念，甚至有巴洛克音乐、巴洛克文学、巴洛克哲学等参照物。进入 20 世纪，西方学者对巴洛克进行了更深入的研究，赋予了它不同的含义，但将其理解为一种艺术风格仍然是主流。

### 3.1.2 社会、历史和文化背景

自 17 世纪初以来，资本主义经济的发展、自然科学的进步、现代哲学的兴起和现代西方美学的崛起构成了巴洛克时期的主要社会历史文化背景。在这样的社会情境下，艺术的创造性发展进入了一个真正完全摆脱神学束缚的阶段，不再仅仅是宗教服务的工具，艺术家也不再是手工艺人的同义词，而是成为了知识学者的一员。

17 世纪的欧洲处于不断的动荡之中，政治领域出现了资产阶级与封建贵族两大势力的对抗，宗教领域有天主教和新教之间的斗争，科学领域涌现出许多新的发现，哲学领域也有激烈的辩论，以及世纪末的情绪波动，这些都汇集到了 17 世纪社会的大熔炉中。美国艺术史学家威廉·弗莱明 (William Fleming) 说：“巴洛克时期是一个不可抗拒的现代力量与固定不变的传统事物相对抗的历史时期。在一系列神学争议、哲学讨论、科学辩论、社会剧变、政治领域的动荡和国家间的战争之后。”

结果既是巴洛克的诞生，也是现代历史的开端。

(1) 其中，“一系列神学争议”被弗莱明列为主要原因。自中世纪以来，宗教信仰在欧洲引起国家间战争和国内动乱的情况并不少见。16 世纪和 17 世纪的欧洲是宗教改革的重

要时期，新教的兴起和天主教的反击在整个欧洲带来了无休止的宗教战争和社会动荡。1562年至1598年间，法国爆发了持续30多年的“胡格诺战争”，这是天主教徒和加尔文主义者之间的冲突；荷兰的独立战争是加尔文主义者在国内发起的战争，希望摆脱天主教西班牙人的殖民统治；英国的清教徒运动是代表资产阶级的新教徒与保皇党之间的斗争；1618年至1648年的“三十年战争”是当时最广泛和最具影响力的战争，也是加尔文派与主教官派之间的斗争，它涉及到欧洲许多国家和地区，给人们带来了毁灭性的灾难和深远的影响。

(2) 社会的动荡带来了思想的混乱，人们的审美意识也随之发生了变化，不再停滞于文艺复兴时期那种优雅、宁静、细腻和圆润的美，而是追求一种带有强烈动荡感的巴洛克风格。

巴洛克艺术形式在不同国家的发展时间是不同的，原因主要是由于每个国家的地域因素和历史文化因素的差异。巴洛克艺术，起源于17世纪初的意大利教皇罗马，并不是在那个时候真正确立为一种艺术风格和形式，而只是一个发展中的艺术趋势。随后，它很快传播到欧洲其他国家和地区，以及受到圣座影响的一些拉丁美洲国家。在文化、宗教和政治环境与意大利相似的国家，巴洛克艺术形式迅速传播，而在一些地区巴洛克艺术则遭到排斥。总的来说，意大利的态度一直主导着这一时期包括音乐、建筑、绘画、雕塑等艺术形式的发展。例如，德国的作曲家依赖于这种类型。像亨德尔、吕利、鲁本斯、伦勃朗等艺术家的作品……

人文主义思想的传播和新教相对宽松的意识形态环境，促进了现代西方科学的兴起和文学艺术的繁荣。巴洛克时期在艺术、文学、科学和哲学史上诞生了许多伟大的大师，包括在文艺史上的英国的约翰·多恩(John Donne)和弥尔顿(Milton)；西班牙的塞万提斯(Cervantes)；在法国，高乃依(Corneille)、拉辛(Racine)和莫里哀(Molière)，拉辛以其细腻的人物分析为例，其核心是反封建思想；荷兰画家鲁本斯(Rubens)和伦勃朗(Rembrandt)，代表荷兰画派的主要代表鲁本斯的创作是一种结合了荷兰传统艺术的宏伟、华丽风格，充满了浪漫激情的笔触，伦勃朗的作品延续了宗教主题的世俗化，人物的矛盾显得更加深刻；西班牙的维拉斯凯斯(Velasquez)和穆里略(Murillo)；意大利雕塑家贝尔尼尼(Bernini)和建筑师博洛米尼(Borromini)等人。在哲学和科学历史上包括培根(Bacon)、笛卡尔(Descartes)、莱布尼茨(Leibniz)、伽利略(Galileo)、开普勒(Kepler)、牛顿(Newton)等许多同样重要的人物，他们为现代思想奠定了基础。从那时起，欧洲人民的精神生活可以用一句简短而准确的话来概括：“他们一直以17世纪天才们提供的智力资本为生活的基础。”

### 3.2 巴洛克时期的听觉艺术 - 音乐概述

听觉艺术不同于视觉艺术，主要指的是通过有组织的声音结构来表达的艺术形式。音乐

是听觉艺术中最重要的代表性艺术。

西方音乐史上的巴洛克时期通常按时间划分为从1600年第一部歌剧的诞生到1750年巴赫逝世的一个半世纪。这种时间的划分是相对的，因为巴洛克音乐的风格特征在文艺复兴晚期就已经出现，并延续到早期古典主义时期，首先是为了便于学习和研究，其次，这一时期的音乐确实具有内在的风格统一性。如同巴洛克时期的其他艺术学科一样，巴洛克音乐是一种宏伟、激情和生动的艺术；它强调情感的表达和充满戏剧性的对比，非常注重细节上的装饰；它也具有大胆的艺术理念和技巧，并有尽可能多地综合诸多艺术的倾向。

巴洛克音乐风格的形成最重要的驱动力是文学和艺术潮流的影响。巴洛克时期科学与文化的迅速进步，以及实验科学的发展，促进了乐器制造业的繁荣，尤其是弦乐小提琴的制造。在意大利，那时有三位小提琴制造业的巨人，而在这一时期小提琴的制造开始关注材料的选择。古希腊的神学思想开始解体，人文学科的发展使新思想影响了音乐的发展，音乐家们抛弃旧观念以发现声音的起因和本质，这一点在巴赫应用十二平均律的法则和拉莫《和声》一书的出版中得到了凸显，后者被称为音乐史上第一部和声学理论。

西方殖民主义带来的资本主义的发展和中产阶级的兴起，以及宫廷中相对宽松的音乐体系，对巴洛克音乐的发展起了积极的推动作用。一方面，中产阶级积累的财富为音乐的发展带来了巨大的可能性，主要体现在商业歌剧院的出现和新型铜版印刷技术的使用上。市民阶层的赞助也变得更加普遍，他们通过向音乐家学习音乐、购买印刷乐谱和歌剧门票，支持音乐的发展，音乐的世俗化增长成为了不可避免的趋势。另一方面，宫廷成为了重要音乐文化的中心，如法国的路易十四宫廷和意大利的美第奇宫廷。

巴洛克时期是复调音乐发展的鼎盛时期，也是向主调音乐时期的过渡，根据风格特征主要可以分为早期、中期、晚期。早期（大约1600-1640年）是巴洛克音乐风格形成的时期，音乐家们力图摆脱文艺复兴时期的传统音乐创作方法，努力追求和探索新的音乐风格；中期（大约1640-1690年）是巴洛克音乐风格定型的时期，作曲家们开始在音乐创作中采用共通的音乐语言，典型的巴洛克音乐体裁和形式最终确立；晚期（大约1690-1750年）是巴洛克音乐的顶峰时期，相比前两个阶段，这个时期的音乐创作主要是完善期，由音乐大师巴赫和亨德尔代表，将巴洛克音乐的发展推向了辉煌的顶峰，同时也为未来音乐的持续发展奠定了基础。

巴洛克音乐的风格特征主要体现在五个方面：

1. 通奏低音：通奏低音是巴洛克音乐的标志性特征，与文艺复兴时期的复调编织截然不同。它由旋律和和声伴奏组成，也就是作曲家在作曲时仅写出旋律和低音，低音旁边用数

字标记和弦中的音符位置，然后由演奏者根据这些标记即兴演奏和声。通奏低音通常由大提琴、低音提琴或者大管演奏，而键盘乐器和鲁特琴则负责填充和声。通奏低音的产生导致和声学的诞生，和声体系的形成又导致了大小调式的产生，从而结束了教会调式的时代。

2. 两种惯例：意大利作曲家蒙特威尔第在 1605 年《田园集》第五卷的序言中提出了“两种惯例”的观点，他认为文艺复兴时期的复调传统，也就是“古风”，可以称为“第一惯例”，而他自己采用的新的作曲技术可以称为“第二惯例”。他强调在使用“第一惯例”时，“和声（指复调音乐）是歌词的主人”，即形式先于内容；在“第二惯例”中则相反，“歌词应该成为和声的主人”，即内容先于形式。

3. 协奏曲风格：“协奏曲”最初指的是独奏者与合唱团之间的对比，后来被应用到器乐作品中，个别乐器与一组乐器、一组乐器与乐队之间时而独立演奏，时而合作演奏，形成特殊的对比效果。这种风格是巴洛克时期管弦乐作品的共同特征。

4. 即兴演奏：巴洛克时期音乐的独特特点反映在其即兴演奏风格上，主要体现在和弦的填充处理、装饰性处理、主题变化以及终止的使用上。

5. 情感因素：贯穿巴洛克时期，音乐家们希望在音乐中注入更多的情感，大家都关注音乐如何能够触动听众情感的审美问题。他们认为音乐的主要目的是唤醒人们的情感。一方面，音乐家开始从音乐的各种素材中探索和实验，寻找音乐表达各种情感和情绪的方式和手段，因此这一时期的音乐首次出现了指示作品速度和情感的标记，例如：慢板大于快板，轻音，强音等。另一方面，这一时期发展了一种称为“情感理论”的理论，它系统地解释了音乐在表达情感方面的作用，18 世纪许多音乐家和学者对此进行了讨论，如蒙特威尔第、卡尔·马塔松等。然而，这一时期的感情与后来浪漫主义音乐家表达的情感不同，它们不是个体情感，而是一种类型化的、群体性的情感，是一种普遍意义上的情感因素，整体意义上的情感形式。

### 3.3 巴洛克时期的视觉伪经

#### 绘画、建筑和雕塑概述

视觉艺术是使用某些材料来塑造可供人们观赏的直观艺术形象的塑造艺术，包括电影电视、绘画、雕塑、建筑、实用装饰艺术和手工艺。本文主要以绘画、建筑和雕塑为例。

巴洛克作为一种艺术现象，在天主教会极力反对宗教改革的时期达到鼎盛，此时在罗马，即圣座所在地。由意大利著名文艺复兴时期建筑师和建筑理论家维尼奥拉 (Vignola) 设计的罗马耶稣会教堂被称为第一部巴洛克作品，它代表了循道主义向巴洛克风格的过度转变。技

法主义是 16 世纪末的一种艺术风格，主要追求奇特和不寻常的效果，比如以扭曲和不协调的方式表现空间，以及以夸张拉长的比例表现人物。"巴洛克艺术被赋予了推动集体情感的义务，通过令人眩晕的意义来满足信徒，这样的艺术具有高度的戏剧性，是一种与文艺复兴时期的清晰性和理性相对立的无比强烈的冲动。视觉艺术热爱曲线和斜线，剧烈扭转，做壮观的游戏，展示一切能引起人们惊叹和钦佩的东西，各种特点共同作用以获得整体效果，各种物质被使用、刺激和结合在一起，与技术壮举共同发挥基础作用。"(1)在绘画方面，宏伟、充满动感、精巧的透视变化、戏剧性的构图、起伏波动、反映无限空间，并与理想的光线形成对比，使画作产生统一的协调感；在建筑方面，主要表现在外观的自由度、对动态的追求、对丰富装饰和雕刻的偏爱、浓烈的色彩，常见的是交错的曲面和椭圆形空间；在雕塑方面，艺术家表现真实世界的技术趋向成熟和完美，展示气势成为最重要的特征，雕塑的人物不再被刻画成静态或休息的姿态，而是永远处于无尽的运动之中。

巴洛克时期的视觉艺术风格特征总体上主要体现在以下四个方面：

与文艺复兴时期的理性宁静和谐相比，巴洛克时期的视觉艺术强调非理性的无限幻觉和错觉，努力打破和谐与平静，在绘画和雕塑中充满了紧张的喜剧气氛，而在建筑上则体现在丰富多变的结构上，充满激情和强烈的动感。内容方面也是令人不安的。

空间关系的处理，巴洛克时期的视觉艺术非常重视空间关系的处理，展现艺术形象的空间立体感，绘画依靠光影造型，追求空间深度，建筑和雕塑强调层次和深度，试图打破平面限制。

综合艺术处理，巴洛克艺术也强调全面的艺术处理。这种综合一般包括建筑、雕塑和绘画的总和，绘画与雕塑的总和，形态与环境的总和，以及艺术本身各个部分的总和。

强烈的宗教色彩。因为巴洛克艺术得到了宗教的赞助，它不可避免地充满了强烈的宗教内涵。

## **4. 巴洛克时期听觉艺术与视觉艺术情感性的比较**

### **4.1 情感源泉及其发展**

情感是艺术中非常重要的审美特质。所谓情感，指的是人们的心理状态，如快乐、愤怒、悲伤和幸福，这些都反映了人们对 K 世界中对象和现象的主观态度。换句话说，情感是人类对客观现实反映的一种特别形式，是人们对客观事物是否符合自己需要和目的的心理反应。当客观事物满足人的需求时，人们会对它们产生满意感、亲近感和快乐感；反之，他们会产

生不满、厌恶和愤恨。从广义上讲，文学艺术是情感的艺术，没有情感就没有艺术。伟大的俄国批判现实主义大师列夫·托尔斯泰在《艺术论》中曾写道：“在个人的心理中唤起一段时间所经历的感受，并激发这些感受，就需要通过动作、线条、色彩和文字所表达的形象来传达这样的情感，使其他人也能体验到同样的情感，这就是一种艺术活动。”无论是悲伤或愤怒，还是舒适和快乐的情感，都可以通过文学创作来宣泄。因此，强烈的情感和真实的情感能够激发强烈的创作冲动和创作欲望，而孤寂和伪造的感情是无法达到创造状态的。

艺术情感的源头有着悠久的历史。早在古希腊时期，亚里士多德便认为艺术起源于人类模仿的天性，因此所有艺术都是模仿，但不同艺术门类之间存在差异，“即使用于模仿的媒介不同，取材的对象不同，采用的方法也不同。”对于音乐艺术来说，其模仿的对象是人类情感。亚里士多德将音乐与人的情感生活相联系，认为音调反映了愤怒与温和、勇敢与节制以及所有相对立的品质和其他性情。他主张“当我们聆听节奏和音调时，我们的情绪也会随之改变。”在和谐的和声与节奏中，似乎存在着一种与人类心智的和谐或亲缘关系，因此有些哲学家说心灵是和谐的，而另一些人则说心灵拥有和谐。亚里士多德之所以将音乐视为“最接近现实的模仿”，主要是因为他认为艺术所模仿的现实主要是指“行动中的人”，音乐可以模仿和反映人的情感，这是一种以人为本的艺术，因此这种艺术与现实的关系尤为紧密。亚里士多德的这一观点不仅代表了古希腊美学中的人文精神，也代表了贯穿于西方音乐美学两千多年历史的音乐情感理论的开端。古罗马哲学家、作家西塞罗在其《论神性》中说：“德谟克利特不认可没有热情的人能成为伟大的诗人。”这一观点也证明在古希腊德谟克利特时代，情感已经存在于文学艺术创作之中。

在欧洲中世纪时期，基督教神学主导了人们的精神生活。神学家认为，表达人的本性和人性对于对上帝的信仰不利，是一种罪恶，所以他们反对用艺术来表达人类情感，主张将艺术的感官色彩减至最低，使之成为服务于宗派观念的工具。中世纪基督教神学的奠基人圣奥古斯丁宣称，上帝是所有美的本质、美的源泉以及衡量一切美的标准。因为对上帝的绝对信仰和顺从，圣奥古斯丁认为所有的艺术都是不可取的，因为它通过个别事物的美来表达和作用于人的感官，而没有上帝的人的感官记忆中的美，与上帝的美相比是遥不可及的。从某种意义上讲，圣奥古斯丁也承认了音乐艺术的情感作用，他认为音乐艺术的魅力能够超越语言和逻辑思维的限制。他说：“那些狂喜的人不是在唱歌词，他的欢乐歌词没有文字，它沉浸在喜悦之中。”

而且他在歌唱的心中充分表达了自己的情感，尽管在那时他并不理解歌词的意思。然而，出于对上帝的信仰和保护，圣奥古斯丁将音乐艺术的吸引力视为潜在的危险，认为情感的表

达会使人疯狂，抑制理性，阻碍对上帝的信仰，因此情感的表达应该被抑制。

文艺复兴时期，人文主义者掀起了一场人性解放的运动，代表了新兴资产阶级的先进立场。人文主义者积极倡导根据世俗和现世的价值标准观察世界，用人的意志和情感决定人的行为，情感的作用得到了更大的关注。在这一时期，情感音乐美学找到了成长的理想土壤。艺术家不再将音乐视为宗教的工具，而是以人文主义态度看待音乐，认为音乐是一种表达内在情感和展现人性的艺术。

到了巴洛克时期，艺术有了一个共同的美学特征，即强调激情状态并有意打破固有的平衡与协调，以强化这种表达的力量。无论是在哪一种艺术形式中，巴洛克艺术家都在追求“震撼效果”，他们非常重视艺术作品的情感，为了使观众的心产生共鸣，他们努力揭示他们心中的情感活动。巴洛克时期最重要的音乐美学特征是由意大利作曲家蒙特威尔第提出的情感理论，他自称为“激动风格的创造者”。

18世纪是西方现代史上一个重大变革的时期。笛卡尔被认为是“身心协同论”的创始人，他认为人类有六种原始情感：惊奇、爱、厌恶、欲望、喜悦、悲伤，所有其他情感都源自这六种原始情感。当人们欣赏艺术作品时，他们的心中会感到一种智性的喜悦，这与对其他事物的爱恨不同，属于一种精神上的情感。法国启蒙思想家狄德罗认为：“按情感和兴趣来描述，这是诗人的才能；情感越强烈，剧本的兴趣就越浓；没有情感的品质，任何笔调的写作都不可能触动人心。”这一观点也可以被认为是在强调情感在创作和欣赏中的作用。狄德罗在他的《论绘画》中还说：“作者首先必须触动我，震惊我，撕碎我的心，吓到我，使我颤抖，流泪，生气，然后如果你还有力气讨好我的眼睛。”心碎、恐惧等都是情感活动，因此在狄德罗看来，绘画中的情感表达应该放在对形式美的追求之上。在视觉艺术的创作中，狄德罗试图摒弃绘画中华丽细腻的创作手法，反对顺应贵族社会品味的新古典主义绘画，热衷于描述和赞扬第三等级的新生活主题风格。

进入19世纪之后，情感美学的发展进入了繁荣时期。俄罗斯进步思想家、文学理论家别林斯基说：“情感是事物本性最重要的驱动力之一；没有情感，就没有诗人，也就没有诗。”伟大的俄国作家列夫·托尔斯泰对情感问题进行了更多的阐述，他认为艺术感染力的深度取决于以下三个条件：1. 传达的感情有多独特；2. 这种感情传达得有多清晰；3. 艺术家的真诚程度。由此可以看出，在整个文学艺术思想发展史中，创作中情感的动机作用和欣赏中的感染作用始终受到重视，并且越来越被强调。在音乐中的情感美学观念在这方面是最成熟和系统化的。最重要的代表人物是黑格尔，他强调音乐的内容是情感的表达，认为只有情感才是音乐应该专有的领域。在这个领域内，音乐可以扩展来表达所有不同特殊情感，所有不同

阴影的灵魂中的喜悦、快乐、幽默、轻佻、任性和振奋；各种程度的焦虑、烦恼、悲伤、痛苦和哀伤，以及像崇敬、崇拜和爱情这样的情感，都属于音乐的特殊领域。汉斯利克与黑格尔的观点不同，他在《论音乐之美》一书中突出表现了他对情感的美学观点，他指出：“音乐不能表达情感，音乐的内容是音乐运动的形式。”音乐只能表达音乐的力度，并用音乐的运动来模仿情感的运动。表达一个定义明确的情感或激情并不是音乐艺术的功能。音乐不描述任何情感，既不是确定的情感，也不是不确定的情感。只有通过演奏，情感才直接在音乐中显露出来。听音乐时的情感是通过演奏或歌唱而被唤起和激发的。尽管汉斯利克反对音乐中情感的表达，但他肯定了音乐中情感的存在。

## 4.2 巴洛克时期音乐和视觉艺术中的情感表达

在巴洛克时期，对激情状态的强调以及为了增强这种表达的强度而故意打破固有的平衡与和谐，成为了那个时代所有艺术学科共有的美学特征。

巴洛克音乐中情感美学的概述，行卫凯在其著作《情感艺术的美学课程》中指出：“这一时期的音乐夸张而奇特，情感的表达过分强化，音乐宏伟而缺乏规律性，打破了平衡与稳定，创造出一种似乎不协调但实际上富有节奏的特殊力量，从而唤起人的情感；在高度赞扬理性价值的同时，也充分肯定了情感的价值，并将理性与感性互为补充，这就是‘理性主义情感美学’。”他还认为“理性主义情感美学”的原因，是由于当时人文主义的延续和具有表达功能的修辞学的影响。

巴洛克时期最重要的音乐美学特征是意大利作曲家蒙特威尔第提出的情感理论。在他的歌剧创作中，蒙特威尔第创造性地发展了许多自文艺复兴以来作曲家们通常使用的音乐表现技巧。对巴洛克作曲家来说，特别是在声乐音乐中，这种表现是一种向听众指明他们想要表达的特殊感情的方式，如喜悦、愤怒、悲伤、仇恨或爱情。为此，他们应用了一系列被称为“感性”的音乐概念。例如，在亨德尔的合唱《哈利路亚》中，“万王之王，万主之主”的歌词以同一音调歌唱，以强调上帝的独一无二。在合唱《我全如迷失的羔羊》中，“迷失”一词被编写给男高音和女高音同时演唱，伴随着他们自己的旋律，从其他人或“迷失”的旋律中脱颖而出。旋律、节奏、重音、重复以及其他音乐概念都被用来加强歌词中所包含的情感。人们认为，上升的旋律、音符的上升表达了希望和渴望，随着旋律的持续上升，而下降的旋律、音符的下降音高表达了悲伤或痛苦（1）。但蒙特威尔第与其他作曲家的区别在于，他不再坚持逐字表达歌词的起伏，而是努力用音乐来表达人的内在活动，并突出情感色彩的渲染。在他的《田园集》第八卷中，他展示了他追求的美学理想是利用音乐构建一个“情感

类比”。为了实现这一理想，他创造了许多当时的新表现手段，如管弦乐的颤音和拨奏。此外，他首次使用主导动机的写作方法，并确立了咏叹调三部曲，大胆使用如七度音等不和谐音程，音乐表达情感的力度被大大加强，表达的领域和幅度也被大大拓宽。

在巴洛克时期的音乐中，情感因素也体现在了情感标记的出现上，例如“慢板” (adagio)、 “快板” (allegro)、 “轻音” (piano)、 “强音” (forte)等。这些速度和情感标记的使用丰富了表达多种情绪和感情的方式和手段。

情感性在视觉艺术中也十分显著，波洛米尼设计的圣伊沃教堂和瓜里尼设计的圣殇教堂就采用了尖塔设计，并且基本上以对空间和光线的陶醉，以及对无限的陶醉为特点。这种设计使人的视觉也随之螺旋式向上，产生了一种飘浮和上升的感觉，也反映了对上帝无限视野的渴望，对未来世界的向往，以及对个人渺小感的体现。

鲁本斯是典型巴洛克艺术风格的大师级画家，他的作品更是情感的宣泄和浪漫表达。华丽的色彩和奢侈的表现使鲁本斯成为巴洛克艺术的杰出代表。热情成为了他几乎所有作品的主调，这在《乡村节日》和《亨利四世出征战场》中表现得尤为突出。在《村庄节日》中，我们可以看到这样一个场景，“画中的舞蹈看起来像是一个无休止的狂野舞蹈，其他的次要线条也倾向于上述的中心线条，使得整个动作变得如此强烈，以至于令人目眩。”丰满的女性身体和肌肉发达的男性是上乘之作。

现实主义的比例在鲁本斯的作品中常常被夸大，例如《机械人美女》、《披着毛皮的海伦》等作品。另一个例子是他的作品《十字架上》，这是一个生动的震撼情感表达。《被举起的十字架》是鲁本斯于1609年创作的，这是文艺复兴时期许多艺术家都创作过的宗教题材，观者在其中体验到宁静与超然、神圣与永恒。鲁本斯的《十字架》描绘了一个被活生生的肉体摧残的暴行，充满了挣扎和痛苦，让我们体验到一种紧张和不安的感觉。在表达内容上，文艺复兴时期的艺术家强调人性中的理性，而巴洛克则表达了人性中的情感，理性让人冷静自持，情感让人沉溺和宣泄。在构图上，巨大的十字架是画作的主题，对角线穿过画面，围绕其移动的一群人都倾斜和扭曲。在透视上，由于在画面设计中使用了倾斜的结构线条，画面具有强烈的前进感。鲁本斯努力将每一个人物连接成一个巨大的整体，将耶稣的身体作为画面的中心，其他人物围绕倾斜的十字架形成一个巨大的S形，这些身体似乎被某种看不见的力量拉动，凸起的肌肉让人感觉体内有一种无法抑制的力量将肌肉向外挤压，展现了一种巨大的动力和情感状态。

#### 4.3 巴洛克时期听觉艺术与视觉艺术的情感性比较及

## 差异

通过对巴洛克艺术情感表达的综合案例分析,不难看出巴洛克时期的听觉艺术和视觉艺术主要集中在情感表达中理性与非理性的比较上。

理性指的是概念、判断、推理、综合等思维方式和思考活动,这不仅是逻辑推理的能力,也包括对事物的协调性和差异性的敏感程度。它是避免艺术家想象力缺乏的重要环节,是艺术家进行艺术活动的指挥棒。理性属于理智范畴内的抽象思维,它与事物的本质、内容的联系有关。

巴洛克时期正是哲学思想史上理性主义思潮占主导地位的时期,笛卡尔、开普勒、莱布尼茨等一些哲学家和思想家用理性的态度批判性地验证并重构世界。理性主义者主张以判断和推理的抽象思维活动作为衡量一切客观事物的标准,认为世界上的所有客观事物都必须以理性的尺度来衡量,只有被理性所确认的事物才是真正的客观存在,只有经过理性检验的知识才能被称为真理。

在巴洛克音乐理论的研究和讨论中,笛卡尔和马兰·马尔桑试图对和声的数学基础进行理性化的探索,虽然并未成功,但为后人积累了大量的理论基础。一百多年后,法国作曲家拉莫解决了音乐美学的问题,并著有《和声与声学》一书在流传。新和声理论的出现不仅为音乐创作实践提供了丰富的理论基础,同时也反映了在音乐和艺术领域倡导科学理性精神的潮流。

余润阳先生在《音乐美学课》中提到,“巴洛克时期的音乐美学继承了自文艺复兴以来产生的一些基本概念,在某种意义上,巴洛克音乐美学是文艺复兴音乐美学的进一步延伸。”不同之处在于,巴洛克时期的音乐理论家和作曲家在研究和阐述一些具体问题时趋于更加理性。从上述例子可以看出,巴洛克时期的音乐家在他们的审美观念中确实有理性主义的倾向,他们相信音乐和人类情感之间有一种对应的象征关系,每一种情感都可以通过特定的音乐形式来表达。在这种信念的指导下,他们寻求澄清和理性化音乐与情感之间的关系。这一概念在蒙特威尔第、A.斯卡拉蒂、巴赫、亨德尔、拉莫等音乐家的作品中可以清晰地看到,早期意大利歌剧中的风格化和体裁化的表现手段,以及德国音乐家研究的音乐修辞理论,也同样适用于这种美学观念。

蒙特威尔第是巴洛克时期音乐美学情感理论的典型代表,在他的音乐理论和实践中,随处可见“他创造性地发展了音乐表达方式,不再拘泥于歌词的表面意义,而是追求表达内心深处的情感。”他使用了前所未有的新音乐表现技巧,如乐队的颤音和拨弦、主导动机、不

协和音程等。他认为人心深处有三种情感：愤怒、温柔和谦卑，而人的高音区、中音区和低音区三个声部区域，用音乐语言以激动、柔和和适度的方式来表达这三种主要情感。这种兴奋的情感表达是他所谓的创作风格——“激动风格”。

作为巴洛克时期最具代表性和成就的音乐大师，巴赫的作品最充分、最完美地体现了巴洛克理性主义对情感的审美理念。他的音乐完美融合了严谨的逻辑思维和深邃的情感表达，庄严中蕴含着精致、严格而自由，高度理性的形式结构中包含着丰富的情感内涵，使得人们内心深处那不可见、无意识的情感体验被准确地表达于有意识的音乐形式的安排之中。以《马太受难曲》为例，这部作品最佳地体现了理性主义的情感审美。《马太受难曲》基于圣经对耶稣受难的记载。在这部作品中，高度理性的构思与情感内容极端完美地结合，每一项技术手段的安排都服务于特定的表达，情感真挚深沉，表达了作曲家对当时社会实际生活的真实感受。

汉德尔是巴赫同时代的音乐巨人，他在巴洛克理性音乐创作中也具有典型的重要性。据说汉德尔脾气暴躁，情绪波动很大，这种性格在他的音乐创作中得到了体现，从他的作品中我们可以清晰地感受到那宏大起伏的情感世界。但这种情感似乎被某种理性所控制，他的音乐结构极其清晰，旋律线条突出，合唱部分特别宏伟，并且充满各种对比和曲线结构的严谨和细致，不管情感的起伏有多大，不管情绪多么激动，他总是追求一种清晰、和谐和理性的音乐风格，并在内心深处毫无保留地展示真实感受。这种巧妙的创作方法基于他卓越的创作能力。他可以自由地表达内在的情感，不受任何因素的限制。

巴洛克时期盛行的音乐修辞同样非常理性化，它研究的对象不仅包括音乐的数学基础，还包括其情感和直觉方面，包含了一种称为“情感编码”的美学概念，这是由文艺复兴时期兴起的情感美学和笛卡尔代表的理性主义趋势结合而成的。马特松（Mattheson），巴洛克时期晚期音乐美学情感理论的著名代表，曾经说过：“如果音乐家想要打动别人，他必须知道如何仅仅用声音及其组合来纯粹地表达一些心里话，不借助文字，而且要表达得让听众能够听出来。”就好像是一场真实的演讲，能够充分理解和领会它的意图、思想、意义，以及每一段落和句子的语调。可见，在巴洛克时期，理性主义的概念贯穿了音乐家的整个创作过程，他们相信，只有通过借用和应用理性化的修辞原则，音乐情感的准确性和广泛性才能得到真正和真实的表达。

与巴洛克时期的音乐相比，在绘画、建筑和雕塑中，情感表达也同样非常突出。但是，视觉艺术的艺术们通常采用一种被称为非理性主义的创作概念：浓重的装饰性因素、明暗对比的结构元素、奇异的人物造型，以及大师作品选择中那种似乎来自另一个世界的生命力。

这种非理性的情感，与文艺复兴时期理性的静谧和谐形成了对比，同时也与音乐中的理性主义情感形成对照。

在17世纪的意大利，随着歌剧的诞生、剧院的建立以及市民街头的活跃生活，为了迎合这种氛围，艺术家们创作了与社会和谐相契的戏剧性创作风格。例如，巴洛克画家们使用对角线、弧线等方式来构图，以实现逼真的艺术画面效果，通常使用明暗对比和透视法，一些艺术家还强调光影和色彩的作用，使得画面极具动态感和戏剧性，给人一种无限空间的错觉。

巴洛克建筑师们将建筑物转化为令人眼花缭乱的混合体，其中包括扭曲的石柱、彩色大理石、弯曲的圆顶天篷、复杂的光线运用、大量的雕塑和绘画等，突出了在追求气势基础上的夸张、浪漫、激情和非理性特点，给人一种惊奇的视觉感受。尽管许多作品是宗教作品或基于宗教主题，但它们常常展现出渴望自由的世俗思想。巴洛克教堂是非理性情感表达的典型例子，大多数巴洛克教堂在主题清晰、形式丰富多彩和夸张方面与玉台集合一样，并且建筑师们通过透视和光线“造假”，形成一种壮丽的气氛。博洛米尼圣卡洛教堂依附于从剧场和歌剧中散发出的信息，教堂的正面就像是一个夸张、热情的舞台背景。帕尔米尼创造的空间被称为“通过雕塑表达的戏剧事件的舞台”，他并不像博洛米尼那样使空间本身具有动态特性，在科尔纳罗祭坛的设计中，他将剧场的光线追逐固化为由华丽的赛璐珞石和精致雕刻组成的背景，隐藏的窗口开口处的光线使这一场景充满活力，主要人物雕塑的形态极为瞬间，似乎在讲述一个故事。

巴洛克时期的艺术作品中，情感性非常突出，主要是因为这一时期的共同美学特征——情感美学的繁荣。任何艺术形式都专注于激情的状态，艺术家们故意打破文艺复兴延续的平衡与和谐，以加强这种表达。在艺术创作过程中，情感与理性始终紧密相连，交织在一起，相互依存，相互影响。首先，任何艺术作品的情感都包含着一定的理性，理性通过情感得以表达。其次，艺术家通过观察客观事物来触发情感，然后通过艺术作品理性地表达这种情感。第三，从自然人类运动的规律来看，任何实践活动都是从感性认识上升到理性认识的过程，不同的是艺术创作最终将理性认识还原到感性认识，而这种感性形象，无论是自然表达还是自然再现，或多或少都掺杂了艺术家的情感、观念和思想。第四，在情感方面，艺术中的理性表现为一种控制力。就像巴洛克时期的音乐家们，即使是最强烈的情感也能在可调度的范围内控制，形成理性主义的情感。在视觉艺术家眼中，为了达成他们的创作目的，他们必须尽一切可能表达情感，甚至超出理性的范围，艺术家在创作自己的作品时陷入纯粹难以理解的推理、思考，或被特有的精湛技艺所现实的压抑，并不奇怪，情感也无法得到充分的表达。

# 第5节 巴洛克时期听觉艺术与视觉艺术

## 装饰性质的比较

### 1. 装饰艺术的含义

通常来说，装饰艺术是造型艺术中的一个分支，包括三个方面：首先，它指的是生活实用产品的装饰加工过程以及供欣赏的饰品的处理，如染织技术、陶瓷技术、家具工艺、象牙雕刻、玉雕、装饰画等；其次，它指的是为建筑物、器物、服饰等装饰的图案或图样、纹饰、花纹等；第三，它指的是不作为独立艺术形式出现的属性，但存在于某一主体艺术之下的不同程度，如隶属于音乐，成为装饰音乐；隶属于建筑，成为装饰建筑；隶属于绘画或雕塑，成为装饰画和装饰雕塑。装饰艺术非常重视形式之美，在此基础上形成了一种具有丰富多样性和统一性、强烈对比和和谐的艺术特征。换言之，自然界中各种事物的形象通过变形和夸张的加工，使其变成具有形式美的图案、纹样、色彩和形态，从而产生一种具有装饰效果的艺术，这就是我们所说的装饰艺术，即装饰性艺术。

装饰性可以在任何艺术形式中表现，如音乐、舞蹈、戏剧、文学、绘画、雕塑、建筑等。从整个艺术发展史来看，最典型的装饰艺术是巴洛克时期产生的洛可可艺术。“洛可可”这个词的含义有更多的解释，在这里主要参考《大不列颠百科全书》和《大美的百科全书》中的例子。《大不列颠百科全书》对洛可可艺术的解释如下：这是一种艺术风格，起源于18世纪初的巴黎，特点是放松、优雅、高贵、典雅，并且大量使用曲线和自然形态作为装饰，形状是“C”形和“S”形，旋涡线的形状，通常是不对称的而不是对称的，颜色明亮而柔和，象牙白和金色是其流行色彩。

洛可可艺术的解释是：“18世纪欧洲艺术的一种表现手法，源于巴洛克风格，为装饰而生，以不对称的图案特质取胜，如曲线和反曲线，以及自然岩石、贝壳、花朵、藤蔓和叶子的形态。”从这两部权威参考书籍中，可以看出洛可可艺术风格最注重的是其装饰性，因此有些人直接称洛可可艺术为装饰艺术。

洛可可艺术，作为法国王公贵族从追求权力到追求奢侈与华美转变的典型反映，它通过旋转的线条和华丽、精致的装饰风格来表现。在洛可可艺术中，画作的主题多是贵族阶级的享乐生活和男女爱情。安东尼·瓦托是洛可可绘画的创始人，代表作为《胖船西摩斯岛》。

受当时宫廷审美观念的影响，与洛可可艺术相对应的，是一种优雅、精致、巧妙、有趣的同时又华丽的宫廷装饰风格的植物风格。他的倡导者是巴洛克时期古钢琴音乐最重要的作曲家——库普兰。

无论是洛可可艺术还是华丽风格，一个是巴洛克艺术发展到一定时期的产物，或几个巴洛克艺术风格的分支，装饰性特征始终是它们的共同点。

## 1.1 巴洛克时期在音乐和视觉艺术中装饰性特征的表现

巴洛克时期是装饰艺术发展中的一个繁荣时期，而“巴洛克”这个词本身就具有过于雕琢的装饰意味。综合性被认为是巴洛克艺术最典型的风格特征之一，而综合性的装饰性质独具魅力。现代主义建筑家称巴洛克艺术仅仅是画面艺术，“以复杂的形式。剪影堆叠的装饰来追求与建筑本质无关的艺术兴趣。”通过分析各种巴洛克艺术作品，不难看出巴洛克艺术确实注重装饰，而这种装饰是基于综合性的。巴洛克艺术在建筑上无处不强调艺术形式的综合性，重视空间绘画的综合，同时吸收文学、戏剧、音乐等其他艺术因素的元素作为想象和创作的来源。

空间与形态的连续性是并行不悖的，但这并不仅仅是通过柱子、重复的断檐和山花等装饰来完成的。就像由波洛米尼设计的圣卡洛教堂一样，它将空间作为建筑的组成元素，以及空间的连续性、独立性和变化作为设计原则，以构成动态空间的状态。瓜拉尼的设计空间概念源自希腊十字、拉丁十字、八角形、圆形或圆顶元素的序列，通过特殊的结合将它们融合成一个连贯、开放的整体。在他们丰富的绘画技巧基础上，巴洛克画家还致力于构图、笔法和韵律的工作，并且使用阴影和透视法使画面活跃起来。音乐艺术家在他们的音乐创作中强调情感和戏剧性的对比，非常注重细节上的装饰，使用大胆的艺术理念和技巧，并尽可能地融合各种艺术的倾向。

与文艺复兴艺术的朴素和简约风格相比，巴洛克艺术显得格外华丽和壮观。这种风格的变化主要是由于教会观念的改变，文艺复兴之后的天主教会和宗教贵族将艺术视为快乐的源泉，在地上过着享乐主义的生活，他们花费巨资以满足对快乐的渴望。因此，一系列如绘画、建筑和雕塑的巴洛克艺术都是非常奢华和壮观的。为了达到他们所谓的豪华风格，艺术家们在形象设计和材料选择上不遗余力。精心的雕刻和装饰设计从未缺席，且经常使用五彩大理石、宝石、青铜和黄金等贵重材料在当时的建筑和雕塑中。例如，意大利建筑师贝尼尼在罗

马圣彼得大教堂的代表作品，其“宝座天蓬”的设计充分反映了巴洛克艺术的奢华与壮丽。整个天蓬堆叠着装饰品，包括每个平面的接合处和角落，以及螺旋缠绕的柱子、金色调和繁复的装饰反映了整个天蓬的贵族气质。代表着弗兰德斯学派的鲁本斯，也是巴洛克装饰风格的有力倡导者。这些画家服务于信奉主教之死的教会组织和宫廷贵族。

巴洛克时期的绘画艺术紧密地将审美欣赏与宗教情感联系在一起，创作出独特的画作。对于热爱宫殿式装饰教堂并定期举行音乐仪式以增强宗教信仰的天主教徒来说，他们相信艺术本身可以表达上帝的善意，能够显露和加强信仰。基于此，鲁本斯为了增加信徒们的信心，努力将他的信仰与晚期巴洛克时期新的、更具装饰性的艺术风格——当时流行的洛可可艺术相匹配。洛可可艺术首先在法国出现，然后迅速从法国传播到欧洲各国。

受到这种艺术风格的影响，音乐中出现了一种壮丽、优雅、精致、灵巧且极具装饰性的炫目风格。它的倡导者是巴洛克时期古钢琴音乐最重要的代表作曲家，库普兰。他的作品中经常使用一种典型的“断奏风格”，尤其是一些装饰音的记谱，他认为这反映了作品的灵魂和本质。库普兰的《密斯蒙妮卡》是装饰音乐的有力表现。装饰音的音乐表达也体现在对装饰音的系统化和统一化上，这成为了巴洛克时期音乐作品的重要部分，后来的作曲家在创作中基本上都使用了这些符号。

在德国，像汉德尔和巴赫这样的音乐家也在一定程度上受到了影响。例如，汉德尔的神剧《弥赛亚》运用了一些丰富装饰的复调写作，多种合唱风格，块状和弦风格的质地和由格式模仿的对位质地，偶尔以齐唱的形式出现，并与复杂的织体结合使用和弦质地。巴赫的勃兰登堡协奏曲中的一些多变的片段也体现了这种风格。意大利作曲家D. 斯卡拉蒂留下的大约600首键盘奏鸣曲精致且装饰丰富，具有鲜明的洛可可风格。斯卡拉蒂的每首奏鸣曲类似于练习曲，要求演奏者使用某种技术和音乐解释解决问题，包括双手交叉弹奏、快速重复、跨越整个键盘的琶音模式，以及需要使用拇指弹奏的乐段。其中最具代表性的是巴赫。

巴赫曾经聆听过布克斯特胡德（Buxtehude）的演奏，布克斯特胡德的幻想曲风格影响了巴赫后来的作曲。这种幻想曲风格不受旋律主题的限制，可以充分展现作曲家的才华和隐藏的和声。巴赫采纳了这种风格，创作了七首托卡塔（toccatas），每一首都有独特的个性和设计，允许演奏者展现华丽的触键技巧。巴赫的管弦乐作品中最为人所知的是勃兰登堡协奏曲第一号中间的Gigue舞曲，以其特别装饰性的节奏和高音提琴声部而著称。巴赫的作品中有许多装饰音符。有些还发展成了非常复杂的装饰性乐句。

在巴赫的音乐中，装饰音符与其他音符处于同等地位，所有装饰音都在拍子上，重音通常比装饰音的第一个音符更为突出。例如，在他的《B小调管弦乐组曲第2号》的第二乐章

中，长笛和第一小提琴以齐奏形式演绎出华丽的旋律主题，中间夹杂着装饰性的辅助音，这个主题重复了三次。他的《英国组曲》和《法国组曲》也使用了丰富而华美的装饰音符。巴赫的D小调双小提琴协奏曲非常经典，但他创造了新鲜有趣的音乐。在这部作品中，双小提琴交替出现，重点展示了主乐器独特的演奏技巧和华丽的装饰音，反映了巴洛克音乐的装饰性风格。

## 1.2 巴洛克时期听觉与视觉艺术中的装饰性比较及差异

在他的《美学》第三卷中，黑格尔评论到，希腊化时期的艺术“刻意雕琢精美以赢得观众的喜爱，人们在从伟大艺术的简朴与自然向姿态的艺术转变中又向前迈进了一步。”同样的评价似乎也适用于对巴洛克艺术风格的评价，这表明在一个时代中有一个艺术的时代。但丁曾经说过：

当事物在表面上看起来并不装饰，但内在实际上是装饰过的，那就是一种好的装饰方式。这里他肯定了一种简朴的艺术精神。巴洛克艺术的装饰性与雕饰时期艺术家们不同，后者热衷于表达装饰性，并追求宏伟、精致和炫目的效果，这是巴洛克听觉艺术和视觉艺术在装饰用途上的共同特点之一。但从装饰艺术使用的结果来看，中间存在许多微妙的差异。

从音乐的角度来看，装饰风格的使用晚于其他艺术门类。在巴洛克时期，受当时绘画和多种艺术风格的影响，音乐家开始寻求新的方式来表达情感和音乐语言。音乐家们赋予了音乐更强的生命力和热情，强调情感的表达和充满戏剧性的对比，更加注重对音乐的装饰。巴洛克音乐的装饰性在巴洛克器乐作品中得到了充分体现。通奏低音的使用提升了器乐音乐中高音部分的重要性，这是一个主要的和声结构，它与文艺复兴时期的复调结构区别开来。高音部分，也就是旋律部分，必须被突出，这就要求创造更加华丽和装饰性的高声部来与其他声部区分开。有时这种华丽和装饰性的高声部也要求演奏者即兴创作，即兴作品使高声部更加华丽、丰富且更具特色。这种新的音乐质地的使用也是因为装饰性的高声部符合当时人们的审美趋势。从某种意义上说，音乐中装饰特征的使用带来了积极的结果。

转向以绘画、建筑和雕塑为基础的视觉艺术。在这类艺术中使用装饰性，并非巴洛克时期的一个特征，而是产生独特巴洛克艺术的装饰特征的一个典型特点。这个时期的艺术之所以被称为巴洛克艺术，在很大程度上是由于这一时期艺术的装饰性质。怪诞、复杂、扭曲等形容词，都与过度装饰联系在一起，难怪许多历史研究者用贬义词来评价巴洛克艺术。仔细

想来，这些贬义术语似乎更适合描述巴洛克时期的建筑、绘画和雕塑。特别是在巴洛克时期的中后期，这一现象更为突出。

## 2. 巴洛克时期听觉艺术与视觉艺术的宗教性质比较

### 2.1 巴洛克时期的宗教发展

14 世纪和 16 世纪是欧洲历史上一个重要的转折点 —— 一个“文艺复兴”的时期。这场强调人类解放的运动影响了一部分支持并赞助复兴运动的教皇。在人文主义的影响下，基督教组织掀起了一场改革运动，这导致了观念的变化，这与之前正确解释基督教信仰不同，天主教会的权威受到了冲击。在这种情况下，天主教会面临的主要任务是“改变现状，消除弊端，整顿内部以巩固其地位。”反宗教改革的序幕正式开始。从 1545 年开始，教皇召集了多次的特伦托会议。会议内容主要反映了反宗教改革的目的，例如宣布教皇为最高权威，神圣且不可侵犯，天主教为正统，其他新教徒为异端。为了囚禁和控制人们的思想，企图打压所有新观念和概念的传播，包括艺术在内的所有宣传手段都被用来巩固教会的权威和地位。天主教与新教之间的战争一直延续到巴洛克时期，更大的影响包括法国的胡格诺战争、荷兰的独立战争、英格兰的清教徒运动，以及 1618-1648 年的三十年战争。当时，最广泛和最有影响的战争是欧洲三十年战争。

三十年战争是由 1618 年在捷克共和国针对哈布斯堡家族的起义触发的，它包括由教皇和波兰支持的哈布斯堡集团，以及由法国、丹麦、瑞典、荷兰和德国新教联盟组成的反哈布斯堡联盟，得到了英国和俄罗斯的支持。战争持续了四个阶段，即捷克、丹麦、瑞典和全欧洲的混战，并以反哈布斯堡集团的胜利而结束。神圣罗马帝国只是名义上存在，罗马天主教会的权威和地位遭受了沉重的打击。尽管自 17 世纪中叶以来欧洲不断发生战争，但大多数都是世俗战争，宗教派别之间的和平共存逐渐成为了一种情况。

在这一时期，为了反对改革运动，天主教会创立了影响力极大的耶稣会。耶稣会，也被称为耶稣公司，是最有力量的大型修会之一，致力于维护天主教会的权威。作为一种新型的修会组织，耶稣会具有严格的纪律，活动地点不受限制，没有统一的服装规定，还保持了念诵六祷文的宗教传统。在天主教会的指导下，耶稣会是一个多任务组织，旨在重新夺回教会对艺术的控制权。从 16 世纪末到 18 世纪，在耶稣会稳步发展的时期，进行了大规模的工程建设，在所有崇拜天主教的地区，他们主持建造了许多宏伟的教堂建筑和新学校，开设修道院，创立了各种兄弟会和善会，发起了避静、朝圣、宗教节日和其他活动。耶稣会在成员中培养

了专业知识，并在现代文化和艺术领域培养了大量名人。

教堂的建设是耶稣会保护自身权威所采取的有效措施之一，它为教会提供了集会、组织和宣传机构，也是展示天主教会规模和权威的有力手段。因此，教会大力建造教堂，耶稣会教堂和圣依纳爵教堂是新型教堂的典范，其精美、丰富、巨大和豪华的艺术特点是巴洛克艺术的典型风格。有一段时间，这种“耶稣会风格”的艺术被用来描述巴洛克艺术。

实际上，巴洛克艺术并不仅仅是耶稣会风格的体现。耶稣会在各个方面的成功很大程度上依赖于罗马教廷强大家族的支持和赞助，特别是在财政资源方面。财务问题经常导致赞助者和受雇艺术家之间的争议，他们通过对财务基金的控制来握有所有的主动权，包括建筑师、画家、风格等的选择。根据他们自己的愿望，他们将教堂顶部的平顶改为拱顶，没有考虑更好的声学效果，保留了装饰主教座位的权利，并设立了家族的主祭坛等。在教堂装饰的问题上，无论是耶稣会还是艺术家都没有绝对的发言权，特别是耶稣会。在这种情况下，耶稣会、赞助者和艺术家之间的冲突加剧了。资金问题一直困扰着耶稣会，直到1660年代，当约翰内斯·保卢斯·奥利瓦成为耶稣会会长，这改变了经济状况。在奥利瓦的领导下，耶稣会积累了足够的资金来雇佣当时最有声望的艺术家，这些艺术家能够根据当时的观念和艺术追求来装饰耶稣会教堂。

## 2.2 巴洛克时期音乐和视觉艺术中的宗教特征表现

在漫长的人类历史进程中，各种艺术的基本主题大多具有宗教性质或涉及宗教性质。长时间以来，宗教艺术作为主导艺术，引领了文学艺术的发展。艺术与宗教之间存在着密切的内在联系，艺术是宗教宣传和说服的工具，宗教感则通过艺术得以表达。“只有当宗教受到那些激发所有伟大艺术的同类事物的启发时，它才显得虔诚而有活力。”一般来说，一个伟大的宗教时代能够创造出一个伟大的艺术时代。

在16世纪和17世纪，欧洲的封建势力逐渐衰落，资产阶级力量越来越强大，但两者之间的斗争仍然处于僵持阶段，彼此不相让步，各有得失，有些国家建立了过渡性的君主制，其他国家则建立了新的资产阶级共和国。在宗教上，天主教和新教之间的斗争也同样激烈且势均力敌。这样动荡的时代在某种程度上有利于艺术的发展，因为无论哪种情况，艺术都被视为双方如此重要的宣传武器，他们相信艺术的优势是胜利的力量和象征。因为是天主教的基地，意大利的罗马成为了巴洛克艺术的发祥地和中心。社会和科学的发展动摇了教会的权威，教会的教义原则上脱离了科学的经验证据，宗教比以往任何时候都更依赖于说服手段，而且在天主教的中心罗马，巴洛克教堂得到了大力建设。作为一种说服和宣传的手段，教堂

的形式变得尤为重要。巴洛克建筑非常奢华和壮丽，外表自由奔放，线条曲折，色彩复杂而华丽，给人一种动态变化的感觉，再加上与之相匹配的透视深度、远处的壁画和夸张、栩栩如生的雕像结合成一种宏伟、多彩、神秘的氛围和充满激情的艺术结构。借助油画和其他艺术手段，描绘了赎罪的故事，使天堂和上帝的形象不断出现在信徒的脑海中，使信徒心理上肯定了教会的权威。这种艺术风格被教会所利用，他们打算通过建造或扩建教堂来加强和扩大天主教会的权力，他们不仅支持而且鼓励，投入了大量的资金用于其发展，而艺术家们也乐此不疲，全力以赴，充分展现了他们的才能，留下了大量带有宗教色彩的辉煌作品。

人们普遍认为，第一件具有巴洛克风格的艺术作品是最有代表性的建筑——位于意大利罗马的耶稣会教堂，它采用了一些典型的巴洛克建筑风格。例如，成对的壁柱、重叠的山花、起伏的檐口、巨大的漩涡等等。宗教建筑作品主要是教堂，但其他作品还包括由贝尼尼建造的罗马圣玛丽亚德拉大教堂、维多利亚的科尔·埃尔纳罗小堂，以及由米开朗基罗·布奥纳罗蒂建造的佛罗伦萨圣洛伦佐教堂的新祭衣室，伦敦的圣保罗大教堂由克里斯托弗·雷恩设计，以及由弗朗西斯科·博罗米尼建造的罗马圣伊沃教堂等等。这一系列耶稣会教堂的内部设计逐渐形成了一种模式，广泛应用于17世纪其他建筑物的建设。

巴洛克时期的雕塑同样具有浓厚的宗教性质，最具震撼力的杰作是贝尼尼的祭坛雕塑《圣德蕾莎的祭坛》，由罗马圣玛丽亚教堂的科纳罗小堂组成，这标志着贝尼尼雕塑艺术的巅峰之作。雕塑描绘了圣德蕾莎在她的幻觉中与上帝相遇的瞬间。德蕾莎苍白的面容无力地仰倒，双眼微闭，唇瓣半开，仿佛在呻吟，宽大的长袍向下垂落，整个身体被飘浮的云朵抬升，显得飘忽不定，在她面前站着一个淘气可爱的小天使，手持着带火的金色箭，射向熟睡的圣徒的心脏，这支箭仿佛已经穿透了圣德蕾莎的心脏，让她感到既痛苦又甜蜜。这座雕塑充满了矛盾，肉体正在经历无与伦比的痛苦，而心中似乎有一股幸福的暖流注入，重新点燃了圣德蕾莎对美好爱情生活渴望已久的沉寂而冰冷的女孩心。为了增强戏剧效果，贝尼尼在人物上方增加了由几束镀金钢制成的神圣光芒，金色的光线照耀在白色大理石雕刻的天使和圣德蕾莎身上，使观众产生了幻觉，增加了对上帝的敬畏，并且渲染了雕塑的神秘氛围，激发了观者更大的宗教热情。

巴洛克时期的绘画和艺术作品同样带有强烈的宗教色彩，例如画家卡拉瓦乔的《基督的安葬》，鲁本斯的《抬上十字架》、《抬下十字架》、《东方人拜访圣婴》以及《圣母升天》，这些都是基督教的主题。伦勃朗的绘画如同反映圣经的思想史，既蕴含了旧约的神秘与宏伟，也表达了新约的兄弟情谊精神。在伦勃朗的作品中，历史与神话再次被提升，目的在于基于信仰之上揭示宗教的真实情感。

在音乐体裁风格方面，巴洛克时期产生了较大的宗教声乐体裁，如清唱剧、康塔塔、钦定曲等，还有一些使用协奏曲风格的宗教作品。就作曲家而言，巴洛克时期的音乐家都参与了宗教体裁的作品，或为宗教组织服务，或从中受益。

清唱剧是一种类似歌剧的声乐题材作品，由旁白人叙述，不设舞台动作和服装布景，强调合唱的作用，大部分内容基于圣经主题，因此不同于歌剧，清唱剧的原文本指的是毗邻教堂的祈祷室（祈祷堂或小礼拜堂，“oratory”），因为这种类型的音乐作品在此地演出而得名。卡瓦列里的《灵魂与身体的剧》，于1600年上演，被认为是第一部清唱剧。巴洛克时期最重要的清唱剧作曲家是贾科莫·卡里西米（Giacomo Carissimi, 1605-1674）。到了巴洛克晚期，汉德尔在这一音乐体裁的创作上达到了新的高峰。

康塔塔通常由一系列宣叙调、咏叹调、二重唱和合唱组成。内容可以使用宗教主题也可以使用世俗主题，根据主题的选择可以分为两类：世俗康塔塔和宗教康塔塔。宗教康塔塔用于路德教会的礼拜仪式中，巴赫在这一体裁的创作上达到了完美的顶峰。阿莱桑德罗·斯卡拉蒂是康塔塔最重要和最多产的作曲家，作品超过600部。

钦定曲（Passion）是一种基于新约圣经中耶稣受难记的清唱剧风格作品。它起源于意大利，后在德国新教中得到发展。在5世纪，钦定曲的雏形在罗马天主教的仪式中出现。14世纪下半叶，由“苏格钦定曲”所组成的单声部圣歌成为钦定曲音乐发展的第一个阶段。到了15世纪初，多声部技术的使用使钦定曲进入了第二阶段的发展，即“苦难回应曲”。17世纪中期之后，钦定曲音乐发展成为第三阶段，即“钦定曲清唱剧”，借鉴了清唱剧和歌剧的创作技巧，使用赞美诗的调子，并增加了对圣经之外自由演绎的歌词。钦定曲最重要的作曲家是舒茨（Schütz），而巴赫则在巴洛克时期将钦定曲提升到了最高水平。

17世纪是宗教音乐发展的又一个高峰时期，在这一时期中，宗教音乐因“第二习惯”的影响而发生了巨大变化，导致了一种协奏曲风格的宗教音乐出现。长期以来作曲家们遵循并坚持的帕莱斯特里纳式的写作风格受到了动摇，尽管他的复调风格被视为教会中唯一可接受的音乐语言。蒙特威尔第首当其冲，他在1610年出版的《圣母弥撒》和《晚祷》就是这种老与新风格的典型代表。

通过音乐创作表达对宗教的信仰，巴赫是最典型的代表。巴赫是一位虔诚的基督徒，他的音乐就像他的信仰一样，是虔诚、崇高、纯洁和神圣的。他的大部分音乐作品表达了对上帝的信仰和对世界和谐意义的体现。在巴赫看来，世间万物都是上帝的恩赐，音乐也是由上帝赐予的圣物，音乐是与上帝对话的方式，是通往天堂的阶梯，因此音乐的全部意义在于赞美和感谢上帝，以及通过这种特别的方式感受上帝的存在。巴赫一生都生活在德国，为教堂

和宫廷工作。晚年时住在莱比锡，他被委以莱比锡市托马斯学校以及包括圣托马斯教堂、雅各布教堂和圣布拉吉斯教堂在内的四座教堂的工作。在这一时期，巴赫为教堂创作了大量宗教作品，包括《马太受难曲》、《约翰受难曲》、《圣诞神曲》等。

复活节剧（Easter Play）、以及两百多部教堂康塔塔和无数的管风琴作品，反映了巴赫惊人的作曲能力和他对宗教虔诚的信仰。

## 2.3 巴洛克时期宗教性在听觉与视觉艺术中的对比及

### 差异

在社会历史发展的过程中，宗教与艺术在一定程度上存在着相互关联和相互渗透的关系。宗教一直使用艺术作为信仰表达的重要形式，一方面为艺术提供创作源泉，另一方面又限制和规范艺术创作，其主要目的是使艺术成为宣传和说服的工具。总体上，宗教主导的时代也是艺术辉煌的时代。然而，由于教会在内容或形式上的限制，艺术的个性发展受到阻碍，显得过于规范化，从而影响了正常的审美趣味和艺术观念，使得受宗教影响的艺术失去了一般艺术的审美特征，形成了只满足教会需要的“公共风格”。从另一个角度来看，宗教艺术有助于增强宗教的影响力，在一定条件下可以为宗教地位的巩固和发展提供重要保障。

从宗教艺术作品的发展来看，巴洛克艺术仍然是宗教艺术的主导状态，但宗教作品逐渐世俗化。这种世俗化在听觉艺术和视觉艺术中有不同的体现，绘画、建筑、雕塑等视觉艺术的世俗化主要是通过宗教作品的具体形象来展示世俗内容，而音乐则主要通过宗教体裁的转变来反映其世俗性。这两者之间的内在联系受到两个重要原因的影响：社会因素和艺术家自身观念的转变。

巴洛克时期同时也是宗教盛行的时期。17世纪上半叶，天主教与新教之间的三十年战争（1618-1648）对艺术产生了深远的影响。代表中产阶级利益的新教和代表旧封建王朝的天主教都利用新的巴洛克艺术风格来服务于各自的目的。换句话说，这种新风格既与罗马教会的反宗教改革有关，同时也是新教改革活力的一部分。三十年战争之后，新教确立，随着17世纪下半叶政治经济的逐渐稳定，文化和艺术得到了前所未有的发展。西方的殖民扩张和资本主义的发展促进了中产阶级的崛起，宫廷的艺术生活也变得更加繁荣。欧洲各国的宫廷成为了艺术发展的中心，尽管教会仍然支持艺术的发展，但其重要性显著下降。各种艺术的发展成为了宫廷不可或缺的优雅元素。这是宗教艺术作品世俗化的社会原因。

对于天主教而言，自马丁·路德的宗教改革以后，天主教会受到了巨大冲击，为了恢复

教会的权势，他们进行了自我转型，比如创建了耶稣会这样的传教士团体，成立了奥拉托利会，进行了圣餐仪式的革新等活动。17世纪初，教会取得了一定的胜利，大量的宗教建筑或祭坛画等作品需要完成，大批艺术家聚集到罗马，为艺术发展注入了新鲜血液。为了满足宗教的需求，这一时期的大部分艺术作品都带有浓厚的宗教色彩，无论是音乐、建筑、绘画还是雕塑，有的作品是对旧有宗教体裁的新诠释，有的是对新主题的新解读和挖掘，有的是直接使用宗教内容，也有的是运用宗教形式。艺术家们在完成宗教任务的同时，也承接了一些非宗教性的工作，宗教作品的世俗性质初现端倪，宗教作品在某种意义上并不一定全部服务于教会，有的作品只是使用了宗教主题，然后逐渐与宗教分离。这是宗教艺术作品世俗化的另一个原因：艺术家创作观念的变化，这种变化是基于符合社会潮流的发展。

鲁本斯是巴洛克时期多产而杰出的画家，作品涵盖宗教和世俗题材。他自幼接受了天主教的洗礼，后来成为了宫廷画家，宗教是鲁本斯画作中的一个重要主题。他为安特卫普大教堂创作的两幅祭坛画《上十字架》和《下十字架》，确立了他作为宗教画家的地位，其他宗教画作还包括《三美神》、《复活》、《爱的乐园》和《狮子狩猎》。鲁本斯是宗教画家，也是伟大的人文主义画家，他通常在关于宗教神话的作品中描绘世俗人物。他的大多数作品反映了他对生活的热爱和对更好生活的渴望。他喜欢通过描绘健康、丰满而充满活力的形象来表达自己的审美理想和兴趣。在他许多宗教画作中都可以看到世俗生活中的人物。例如，在他的宗教画《三美神》中，左侧的女神是以他的妻子海伦娜为模特的。在《爱的乐园》作品中，也可以看到他的妻子和他自己的影子。还有许多类似的作品，在这里就不一一列举了。鲁本斯的宗教主题画作继承并发扬了意大利传统，倾向于散发出宏伟的比例和宫殿式的主题，形成了既是天主教又是世俗的风格。他的宗教作品的宗教性质仅限于题材，所有宗教中推崇的神秘主义或禁欲主义的概念已被抛弃，在画中涉及的如圣母玛利亚、基督和使徒们的宗教形象，都被现实生活中的真实血肉所取代，所有古代神祇都在王公贵族的宫殿中以真实的名字和面貌复活。表面上虽然笼罩着教会的一层轻纱，但实际上所展现的风俗、习惯、思想和情感都是世俗的。

卡拉瓦乔的绘画作品总是描绘了维纳斯和圣母，他强烈主张将自然主义的理念应用到绘画中。在他的作品《圣马太的灵感》中，圣母和英雄的形象来源于社会地位较低的男人、女人和孩子们，耀眼的光线从画面右侧斜射而来，在墙面的背景上反射出强烈的对角线，并且光线的变化增强了画面的立体感和空间感，给人一种超强的自然真实感。

雕塑是一种比其他艺术形式更易于表现的对象。巴洛克时期的雕刻家们能够创造真正的三维作品，放置在真实的环境中供人观赏，并使用特殊的创作技巧来唤起更大的宗教热情。

雕塑不仅能够叙述故事，相较于绘画，它还能更直观地展现人物的情节，因此巴洛克雕塑非常受欢迎。巴洛克雕塑家贝尔尼尼为罗马圣玛丽亚德拉维多利亚教堂的一个小教堂创作的《圣德肋撒的幻觉》展示了那些宗教狂热者，在被刺伤后虽饱受巨大痛苦，但仍坚持对上帝的忠诚，并表现出极端享受的状态。这个充满幻觉的宗教传说，在贝尔尼尼巧妙的创作方法下，给这个传说增添了世俗的色彩。

在巴洛克时期，与视觉艺术相比，宗教音乐的世俗化主要体现在宗教音乐类型的世俗转变上。首先是歌剧的世俗化。神剧是一种与歌剧区别开来的宗教音乐剧种，主要是因为其内容主要是圣经故事。实际上，神剧有两种类型：一种是使用拉丁语，以圣经为主题的，属于真正意义上的宗教音乐，称为“拉丁神剧”；另一种是用意大利语演唱，不局限于教堂，向公众表演，内容自由不仅限于圣经主题，因此被称为“世俗神剧”。17世纪下半叶，神剧开始从罗马传播到其他意大利城市如威尼斯和佛罗伦萨，以及德国和英国的许多地方。世俗化趋势越来越明显，以至于神剧甚至可以在公共歌剧院上演，女歌手可以参与演唱，而咏叹调的使用被给予了更多的强调。到了17世纪末和18世纪初，拉丁神剧逐渐消失，而世俗神剧继续发展。

其次是康塔塔世俗化。康塔塔与神剧相似，内容既可以是宗教的，称为宗教康塔塔；也可以采用世俗主题，称为世俗康塔塔。最初，康塔塔主要以宗教主题为基础。在德国，用于路德教会议式的宗教康塔塔得到了巨大的发展，这个体裁在巴赫的创作中达到了完美状态，并且在巴赫的创作中完成了最自然的转变。巴赫的早期作品主要创作了宗教康塔塔，其中最具有特色的使用了赞美诗曲调，比如第4首《基督躺在死亡的锁链中》和第80首《我们的上帝是坚固的堡垒》。在巴赫晚期，作为莱比锡大学音乐俱乐部的指挥，这为世俗音乐开辟了道路，康塔塔转向了世俗创作，如这一时期著名的《农民康塔塔》和《咖啡康塔塔》。

值得一提的是，从宗教艺术作品创作目的的角度来看，巴洛克艺术的目的不再仅仅为教会服务，而是呈现出多样化的目的。不过，像巴洛克时期建立的教堂，在音乐中创作的诗歌和弥撒曲等宗教音乐，它们最初的目的是为教会建立，所有的设计，包括结构、形式和创作技巧，都是以特定的方式安排和布局的，无论花费多少时间，其作用不会改变，具有特定的宗教目的。然而，它们的蓬勃发展得到了教皇和贵族的极大支持，因此在这种情况下，巴洛克艺术经历了一定程度的道德缺失，并且为了维护教皇和贵族的权威，他们不惜代价建造奢华的艺术作品来衬托神的威严，不顾下层阶级的苦难，继续在思想上麻痹和统治下层阶级，对他们进行榨取和剥削。这样的艺术不是真正的艺术，它只是统治的工具。

## 参考文献

- [1] 郭大飞. 听觉的建构：空间设计中声音的语法翻译研究. 南京艺术学院, 2014.
- [2] 罗珊. 听觉与视觉的完美结合：动画在音乐教学中的应用探讨. 中国音乐教育, 2001(11): 2.
- [3] 康小云. 钢琴声音与艺术表现关系探究. 乐器, 2003(2): 2.
- [4] 郝可. 姚大钧 | 寻求声音的本体. 东方艺术, 2013(11): 1.
- [5] 韩洁. 从视觉到听觉：声音标志艺术探索. 南京艺术学院学报：音乐与表演版, 2013(4): 6.
- [6] 李欣宇. 用声音“看”世界：声音语言表达的“可视性”研究. 科技传播, 2021, 13(3): 3.
- [7] 赖泽毅. 简论装置艺术中声音可视化的理解. 2020.
- [8] 全明霞. 电视专题片音乐美学与叙事表达分析. 新闻爱好者, 2020.
- [9] 张汉超. 电影电视声音中听觉思维的建立和培养. 新电影, 2018(3): 4.
- [10] 丁江. 主持人主导的“视觉语言”与“听觉语言”. 中国广播电视学报, 2021(12): 5.

# 第6章 交互艺术：从互动到后参与

## 第1节 基于交互式多媒体技术的装置 艺术设计与实现

在当前全球化、日益精密的科学技术以及各种文化的融合的时代，装置艺术在双年展和国际数字互动艺术交流展等大型展览中占据越来越多的席位。这一现状表明，装置艺术在当代艺术中扮演着越来越重要的角色。此外，随着艺术世界中感知倾向理论的发展，越来越多的艺术家倾向于关注观众的体验感受，更加注重观众在他们的作品中的感受、想象力和沉浸感，而交互式装置艺术首当其冲。在这种结合了艺术与技术的交互装置中，交互式多媒体技术发挥了重要作用，丰富和发展了装置艺术的展示效果，艺术与技术的结合也更加紧密。本文以此为起点，论证装置艺术呈现出新的变体：交互式装置艺术时代的兴起之始。目前，将各种交互技术融合到装置艺术作品中有许多方式，本文的主线是使用交互式多媒体相关技术来阐述原创交互视频装置的设计与实现过程，重点是使用 FLASH 软件实现图像互动的实验过程，并分享作品的设计过程和经历。本章将解释本文的背景、意义、研究现状、研究方法和组织结构，以便观众对整篇文章的结构有一个清晰明了的了解。

### 1. 研究背景与意义

#### 1.1 研究背景

自 2010 年以来，多媒体技术研究进入了一个稳定期，期间各种针对交互式多媒体应用技术的目标研究变得流行起来。除此之外，与艺术相关的技术研究也开始活跃。交互式多媒体技术的应用研究主要集中在视觉展示上，近年来，其研究重点逐渐向其他方向倾斜，这表现在教学方法的变革、展示效果的转变和互动应用的创新上。另外，近年来，艺术不再是少数艺术家的自我娱乐，而是与公众的亲合力越来越强，特别是新媒体的发展，让更多人有机会观看新媒体艺术展览。新媒体艺术的魅力之一是，许多作品融入了交互艺术，无论是交互

式游戏、交互式电影，还是交互式广告，“互动”这个词汇的使用越来越频繁，在这样的环境下，交互式多媒体技术也不出所料地成为装置艺术的一部分，丰富了装置艺术的视觉可观性，为装置艺术的空间拓展和设计执行手段的丰富性提供了广阔的空间。这为装置艺术家的创作和装置作品的互动方式提供了广阔的创造空间。

通过相关文献阅读，我了解到交互式多媒体在装置艺术中的应用是一种新型的课题，加之当今的设计越来越多元化和多学科交叉发展，多媒体技术，尤其是交互式多媒体技术，在一定程度上促进了装置艺术的成长。与传统装置艺术相比，交互式多媒体技术的应用对于装置艺术的效果充满了惊喜。因此，交互式装置艺术频繁被提及，其互动功能自创始以来就依赖于计算机，而综合的展示效果使交互式装置富有魅力。优势在于其丰富生动的表现力和可操作的互动性，这显示了多媒体技术对装置在视觉、触觉和听觉效果的渲染和衬托具有重大意义，它使观众能够沉浸在作品中，在与作品的互动中完成体验，产生无尽的联想和余味。在这样的背景下，本节的焦点是基于计算机科学的交互式多媒体技术在装置艺术中的应用，通过开题最终采纳了在装置艺术中设计和实现交互式多媒体技术的方法。

此外，本节是基于上海市优秀青年教师奖励基金项目“艺术设计领域交互式多媒体技术应用研究”结论的深化研究，为本节的后续工作奠定了坚实的理论基础和技术支持。研究成果将具体化为交互式装置。作品的创作首先考虑解决方案的可行性，由于交互式多媒体技术的多样性和技术难度，经过长时间的思考，倾向于使用自己的专业知识，交互式视频装置是首选的类型。理由如下：首先，多媒体技术的发展及其在艺术领域的应用已开始占据主导地位，改变了传统媒体对人们生活的影响，特别是交互式多媒体技术，很大程度上满足了艺术家，尤其是新媒体艺术家对作品创新的需求，也满足了现代人作为观众寻求创新和刺激的心理；其次，在艺术领域的装置艺术创作中，艺术与技术的关系日益密不可分，交互式多媒体技术的介入被赋予了生动独特的活力。创意、观念和设计理念也在新技术的变革下逐渐升华到更高的层次。

在上述背景的基础上，本文设计并制作了交互式视频装置作品《自行之路》，这是一件个人原创作品，基于计算机科学的交互式多媒体技术进行设计与实现。具体包括亲身实践经验和设计过程、交互实施以及最终编辑结果[1]。

## 1.2 研究意义

目前，在国内艺术界，交互式多媒体技术仍然局限于小范围的“话语”讨论，在最好的情况下也只是现代艺术类型的外在形式。在形式上，几乎都是对国外装置艺术的模仿、借鉴、

修改和应用现成技术。许多国内的雕塑家和画家被重新包装,获得了新的头衔:装置艺术家,这被戏称为中国式的“暴发户艺术家”,他们缺乏对交互式多媒体技术等新技术手段的理解和投入,更关心的是如何完成一个作品,或者像堆砌作品一样堆放东西。因此,装置艺术本质上需要新的学问和新技术来支撑。这些技术既不能直接拿来,也不能简单地从中国传统中挖掘,而需要引入一些新的技术表达手段。然而,国内的“现实主义”爱好者们骨子里对“技术”的存在极其蔑视,这种蔑视阻碍了中国当代艺术特别是装置和新媒体艺术的发展。这个项目恰恰是试图寻找一个平衡点:将“技术简化和模板化”作为研究的起点和突破口,使更多的艺术家不再“望技术而却步”,不再担心不懂新技术,担心如何创新作品,也不再担心技术人员不理解自己的设计理念,让艺术家能够因为掌握了技术而使装置艺术更加生动有趣,通过“互动”更加贴近公众和生活。这是本课题的理论意义。

此外,计算机科学中的交互式多媒体技术是基于计算机操作的,但与平面软件或非线性编辑操作不同,它要求作者们能够使用基本的软件语言进行编程,例如Flash软件使用的Action Script、Director软件使用的lingo,以及微控制器编程所需的Pic basic或C+语言。然而,艺术专业的学生毕竟不是理工科专业,学习的基础不同,达到的目的和结果也有所不同。本文的研究希望对交互式多媒体技术在装置艺术中的应用做出新的尝试,试图将基于计算机科学的交互式多媒体技术应用到原创的交互式视频装置《非自行之路》中,专注于使用Flash软件等Action Script编程来完成交互。同时,寻找一条将复杂的交互式多媒体技术与装置艺术结合的路径。这项研究的结果将为没有电子科学基础的艺术师们在理论和实践中掌握核心技术铺平道路,这也是本研究的实践意义和价值。

## 2. 研究现状

### 2.1 国内外研究现状

在中国,交互式装置(也称为互动装置)作为一种新兴艺术形式,它伴随着艺术观念的更新和科学技术的巨大进步,无论是在表现形式还是与观众的互动上都非常具有创新性。此外,新媒体影像艺术近年来在中国也已经起步并开始流行,一些由清华大学、上海大学、武汉大学代表的高等院校已经相继开始在新媒体视频装置艺术领域进行研究。2003年,上海大学主办了《九国数字艺术展》,2004年,北京清华大学与德国艺术与媒体中心(ZKM)和荷兰艺术展览协会(V2)合办了北京首届新媒体艺术展,此后也举办了多次展览,其中交互式视频装置占了相当比例的作品。

此外,许多国内艺术家自费出国学习和发展,现在他们回到国内艺术界带来了新的活力,国内的交互式装置艺术在这样的氛围中也开始迅速发展,首先是以专业课程的形式进入了大学,在一些大学逐渐开展了这门课程的教学,使得以交互式装置为载体的新媒介也开始逐渐被艺术家们接受和理解;虽然现在这类展览在北京、上海和深圳更为频繁,但本文认为,在不久的将来,交互式装置艺术的展览在中国将会更加普遍,这必将带动整个国内艺术氛围的提升。

国外关于交互式装置艺术最著名的研究机构是美国麻省理工学院的媒体实验室(MIT Media Laboratory),其前瞻性和创造性已经得到了国际学术界的认可。2010年3月9日,媒体实验室向全世界媒体开放,波士顿环球报的记者罗伯特·坎贝尔(Robert Campbell)评论说,该实验室是全世界最美丽的30座建筑之一,跨学科的研究小组将在这里进行超过400个项目的研究和开发。后来,MIT推出了一系列关于新媒体艺术的公开课程,通过互联网这一媒介免费向全世界开放,这表明艺术作品的更新发展与新技术的发展紧密相连。一般研究交互式装置艺术的人多来自学术界,一些著名的艺术学校和研究中心包括德国卡斯鲁尔的ZKM(科学与媒体艺术中心),日本东京的ICC(互动艺术中心),科隆媒体艺术学院,澳大利亚新南威尔士大学美术学院,以及奥地利林茨的AEC(多媒体艺术中心)以及前面提到的美国MIT等,这些都是与艺术学校紧密相连的非常知名的艺术和研究中心。这些中心致力于新媒体和新艺术表现形式的研究,并自然而然地辐射到交互式多媒体技术和交互式装置艺术。

当前国内外“交互式装置艺术”研究现状,然后从研究的高度来看,目前国内对“交互式装置艺术”的研究更多是围绕交互式装置艺术的概念、互动模式、交互式装置表现形式等理论研究的方向,此外还包括交互式装置艺术介入展示设计、产品设计、公共艺术、博物馆和新媒体艺术应用研究。这些学者或研究生也关注于交互式装置艺术的概念、交互式装置艺术作品的观众分析、创作与互动方法的分享,以及概念和形式的理论描述,而技术与艺术的真正结合非常少见,特别是在装置艺术中使用交互式多媒体技术的例子更是少见且很少被讨论。

总的来说,交互式多媒体技术在交互式装置艺术研究的应用,是伴随着近年来新媒体艺术的发展以及交互技术的进步而发展起来的。大多数技术仅限于艺术家在创作过程中的小范围讨论,艺术家本身更倾向于研究作品的艺术形式和设计概念,不太会将某个特定的交互式多媒体技术在装置艺术中的尝试写入学术论文或是专业书籍中。因此,相关的论文资料不仅相对较少,甚至几乎是不存在的,也没有成熟和系统的指导,这种研究资料的缺乏给本文的

研究带来了一定的难度。然而，现实总是隐藏着一个机会，这样的研究情况也为本文提供了机会，有着创新和发挥的空间。

## 2.2 交互式多媒体技术在交互式装置艺术中的应用现状

对于“交互式多媒体”这个术语，人们并不陌生，它是指利用计算机交互集成技术和数字通信网络技术来处理各种媒介——文本、图形、图像、视频和声音，使得多种信息建立逻辑联系，整合成一个交互式的系统；这个交互式系统被广泛地称为“交互式多媒体”。在“交互式多媒体”这个词后面加上“技术”二字，作品的“门槛”变得更高，许多艺术家宁愿绕道而行，对技术色变。这种情况清晰地概括了交互式多媒体技术在装置艺术中的应用情况，而“技术”似乎已经成为了一个难以逾越的鸿沟，艺术与技术之间的一座断桥。

简化技术和降低执行难度是让“断桥”变得通畅的方法之一。在新媒体艺术兴起之际，交互式多媒体技术介入艺术领域，开始改变人类感知、反思以及与世界互动的方式，同时也改变了我们与周围世界的沟通方式。许多技术人员通过参与或与艺术家合作完成交互式艺术作品，如何以更简单的方式实现相同的互动效果，成为了大家关注的焦点。加上全球范围内新媒体艺术中心或研究机构的建立，加速了艺术与技术的融合，这在装置艺术中尤为突出，即装置艺术，特别是后现代装置艺术，已不再是达达主义的聚集，也不再是杜尚那种典型的现成品艺术作品“泉”（小便斗）的代表，它已经完全变成了一个反叛的青少年，更多地考虑与观众的互动，有趣的互动是装置艺术的标志性逗号，因此提出了交互式装置艺术的概念。交互式装置与传统装置最显著的不同在于交互式多媒体技术的介入，其创新的互动效果标志着装置艺术进入了一个新的阶段。

新阶段的标志性事件是2010年上海世博会，各展馆中的多媒体展示效果充分体现了交互式多媒体技术带来的新感觉和新体验。这种新的视听和互动感受已成为展示趋势，并且越来越本土化，更加普遍地用于各种展览馆，如：博物馆、科技馆、企业展览厅等，它们被称为“多媒体展厅”。另外，从近年来双年展的趋势来看，交互式装置艺术占据的比重越来越大，以交互式多媒体技术为技术核心的应用，从技术层面、认知层面和审美层面极大满足了艺术家和观众对互动的需求。在这种环境下，基于计算机科学的交互式多媒体技术正被越来越多的艺术家所关注，并试图应用在交互式装置艺术作品中。然而，当前的情况是：由于艺术家的成本考量，尽管交互式多媒体技术越来越发达，但很少有艺术家首次尝试螃蟹，将这

项技术应用于创作作品，即使有作品，也仅是艺术家改变电路板，利用红外感应装置设计互动效果，展示和互动效果一般。因此，交互式多媒体技术在装置艺术中的应用还不够普及，在众多原因中，成本和技术壁垒是作品生存难度的关键。在中国期刊全文数据库的范围内，尽管对装置艺术的讨论很多，但很少或根本没有提及交互式多媒体技术在装置艺术中的使用。这是中国交互式多媒体技术应用于交互式装置艺术研究的一个无奈现状。

### **3. 交互式装置艺术与交互式多媒体技术**

交互式多媒体技术已经给人们的视听生活带来了变革，这种结合了文字与图像、拥有出色声音与影像的交互式多媒体技术渗透到我们生活的各个方面，应用越来越广泛。特别是在交互式装置艺术领域，其应用极大地增强了人们对抽象事物的理解和对某一过程的模拟与感受，并将装置艺术的表现力引入了一个新的境界。因此，本章的上半部分主要阐述艺术设计领域中装置艺术的概念和范围，然后从交互的角度引出交互式装置艺术的概念，即在装置艺术的基础上扩展交互式装置艺术。此外，还总结并结束了国内外交互式装置艺术的发展，并研究了交互式装置艺术的互动模式、界面和沉浸感，为后续作品的交互设计和互动模式选择铺平了道路。本章的下半部分以“交互”为起点，重点介绍交互式多媒体的概念和主要技术；描述了使用 FLASH 软件实现交互技术的方法；详细描述了交互式多媒体技术的应用及其独特的应用优势。通过这样的方式，从交互技术到多媒体技术，可以直观地理解在交互式装置艺术中使用的主要交互式多媒体。

#### **3.1 交互式装置艺术概述**

本文研究了交互式多媒体技术在装置艺术中的应用，因此，本文专注于对装置艺术，特别是作为数字艺术领域新媒体分支的交互式装置艺术的大量文献阅读和研究，近年来交互式装置艺术日益成为新媒体艺术家的新宠。

##### **3.1.1 装置艺术**

装置艺术是当代艺术的主角，这一点是不可动摇的，这从画家、雕塑家、甚至医生转向装置艺术家的趋势中可见一斑，在近年来的过去的双年展和新媒体艺术展览中，装置艺术也承受了重要地位的冲击。有人说，“展览在当代变成了一种装置”，这表明精心阐述概念或理念的艺术日益转变为装置艺术：它变成了一个整体的空间，它也变成了对物与物之间关系

的诠释，并最终成为艺术作品的基础。因此，当我们谈到装置艺术时，不得不提到装置艺术的著名先驱作品：法国艺术家马塞尔·杜尚在19世纪早期的《泉》是装置艺术的开创性作品，杜尚持有题写“R·Mutt 1917”侧边的那件作品被送往博物馆展览，当时的观众对这件作品感到疑惑，这件作品展示了杜尚对艺术的质疑，什么是艺术？艺术与生活有何不同？《泉》向人们表明，艺术或非艺术仅取决于人们如何选择，人们将它放在什么位置，它就会成就它变成的性质，杜尚的思想为当代艺术打开了自由的大门，二战后逐渐发展成为独立的艺术形式，并在西方现代艺术中占据了重要的位置。

因此，装置艺术不是使用笔和纸作为工具的传统艺术，而是已经发展成为一种观念艺术，这与绘画、雕塑和壁画等传统艺术有所区别。这种差异反映在当今的艺术被划分为两大类：架上艺术和非架上艺术，而装置艺术及其变体交互式装置艺术属于非架上艺术。因此，装置艺术的概念可以概括如下：在特定的时间和空间环境中，艺术家以艺术性和有效性选择、利用、转化和结合人类日常生活中的物质和文化实体，使之能够演绎出一种展示个体或群体丰富精神和文化意义的新艺术形式。

### 3.1.2 交互式装置艺术

交互式装置艺术（INTERACTIVE INSTALLATION ART），其词源来自英文，字面上被解释为“互动的，或者是交互式设备”，或者可以理解为“一个具有直接且持续的双向电子或通信系统用于人机对话”的硬件设备。它是一种跨界艺术，它属于新媒体艺术领域中的数字艺术，作为新媒体的一个新分支。本文将交互式装置艺术的形象比作一次性游戏艺术，当然，这是有偏差的，对于交互式装置的描述有很多种，可以大致总结如下：

- 1) 它是艺术家借助某种媒介来表达观点和反思的艺术结晶。
- 2) 它是艺术家生活经验、感知和情感的延伸。
- 3) 它也可能是自我贬低和不知所措的表达。
- 4) 它是一个不需过度解读的隐喻。
- 5) 是一种形而上学的神秘暗示。
- 6) 它是设计方面的禅意启示。
- 7) 它是艺术家向公众或少数人传达的情感的宣泄和释放。

除了上述关于装置艺术与传统艺术之间的差异描述之外，Richard Schiffer 在文章《Cychan 的物理学：触摸的政治》中写道：当手通过其触觉能力象征性地描述眼睛通过直接观察所经

历的空间时，我们称之为“绘画”。绘画保留了两种物理邂逅或冲突的细节，即正在被绘制的内容和绘制的主体，也就是艺术家的主题和造型对象，但交互式艺术模糊了这一概念，并且与传统绘画相比，有第三种“物质性”——公众的参与。观众的干预和参与是交互式艺术不可或缺的一部分。因此，交互式艺术“既不是纯粹的艺术，也不是纯粹的游戏，艺术与游戏的双重性是它的天生特征。”这种互动实现了从作者一方的创造性思维到接收方的双方之间的主动思考的转变。”。移植交互式艺术的概念，再加上多媒体技术的应用，交互式装置艺术应运而生，成为装置艺术的另一个维度。

因此，基于以上观点，交互式装置艺术的概念可以描述为新媒体类别下的一种新兴艺术表现手段。它是一种基于某特定计算机硬件和软件平台，通过计算机硬件实现人机互动的艺术形式，并且本质上是基于硬件装置媒介的交互艺术，能够使观众参与、沟通乃至融入作品，成为作品不可分割的一部分。交互式装置艺术伴随着创作者想要传达的不同信息，所呈现的艺术效果也各不相同。人们通过参与作品和与作品的互动成为作品的一部分，艺术反过来与公众更为亲密。互动使得装置艺术家所表达的创作动机和思想概念更深植人心，也可以说交互式装置艺术是艺术家情感的物质化表达，互动的效果是衡量交互式装置作品的核心。

## 3.2 交互式装置艺术的发展

### 3.2.1 国外的交互式装置艺术发展

交互式装置艺术是在装置艺术的基础上发展起来的，最早的交互式装置可以追溯到1946年美国宾夕法尼亚大学开发的世界上第一台电子计算机的诞生，自那时起，它的生命历程就与计算机平台有着千丝万缕的联系。在此后近70年的时间里，计算机变得越来越普及，技术与艺术的结合也变得越来越频繁，这两个事实成为了交互式装置艺术最终成就的坚实基础。特别是后期以交互多媒体为代表的新媒体的出现，参与到装置艺术中并与之融合，交互式装置艺术因此而产生。

在交互式装置艺术最为兴盛的时期，艺术家丁克·哈根（Dink Hagen）是首先提出“交互媒体”概念的著名代表。但在六十年代末，当时的作品尚未摆脱单调乏味，主要是视频和声音媒介。韩国艺术家南俊（Nam June Paik）也被称为视频艺术之父，因为他在这一时期创作了大量的作品，其中一个作品是在展览馆内设置一个电视机，并使用一个巨大的磁铁干扰电视成像。在晚年，他专注于互动性的研究，尽管互动模式倾向于简单，但他的作品也长期

成为艺术展览中交互艺术作品的代表。1968年，计算机图形艺术首次在伦敦的计算机艺术巡回展览中亮相，计算机艺术作为一个独立的艺术形式，在艺术设计领域开辟了新世界。同时，电子媒介逐渐介入，不同形式的艺术表现催生了不同类型的媒介——视频艺术、数字艺术等，新媒体艺术的概念开始成型。到了1970年代中期，多媒体的概念出现了，再加上其在信息传递和表达上的多样性、互动性和整合性，成为其巨大的优势，并开始被世界接受。

自那以后，图像、光、声等媒体被用于装置艺术，但相对来说，技术只是一种简单机械的操作，例如电子控制的钢琴、放置在玻璃容器中的一组通电的发光二极管，或者一个编号的灯可能会成为交互式装置艺术。然而，这些并不是真正意义上的交互式装置，观众的参与很少在整个展览空间中被考虑，并融入作品之中。直到1970年代中期，才出现了以摄像头和传感器为主要技术的装置作品，收集观众的各种动作和行为，通过计算机完成图像合成，再通过投影机在大屏幕上显示，这种作品的随机性和互动方式成为了之后交互式装置艺术的模板和核心互动模式。

在1980年代，多种多样的交互式装置愈发耀眼，随着多媒体技术的提升，交互方式变得越来越多样化。进入90年代，个人电脑的普及，互联网技术越来越多地沉浸并迷住人们，这种网络媒体的出现使得交互式装置艺术有了新的研究和尝试，自此，艺术的主要领域不断更新扩大，趋向于观念艺术，并突出人文关怀，互动性趋向于趣味化和成熟化。

自新世纪开始的近十年间，各国纷纷设立了研究中心，世界上最重要的机构包括：美国麻省理工学院（MIT）的媒体实验室、日本东京的ICC媒体艺术中心、德国的ZKM科技媒体艺术中心、奥地利林茨的电子艺术节以及荷兰的V2媒体艺术中心。他们的研究推动了交互技术的进一步改善。特别是眼动追踪器、追踪球、操纵杆、笔板、红外传感器等的应用，结合多媒体技术，加上投影、分屏仪器、无缝巨幕拼接等技术的提升，使得交互式装置作品的表现力不断扩展。在当今的主要双年展中，交互式装置的比重越来越高，已成为一种趋势发展方向，赢得了观众的关注和好评。

### 3.2.2 中国交互式装置艺术的发展

#### (1) 中国的装置艺术

在交互式装置艺术的发展谱系中，不可避免地要提到装置艺术这一大的艺术类别。装置艺术是英文“Installation Art”一词的直译，最初意指“安装”，后来这个术语主要在剧场领域使用，一般指舞台布景和可进行拼贴、排列、移动和拆卸的部分。在中国，装置艺术经历

了从最初不被理解、关注和接受的边缘艺术,到如今逐渐进入当代艺术世界主流领域的过程,而在 20 世纪 80 年代,装置艺术从国外传入中国。

中国装置艺术的兴起源自 1985 年 11 月 18 日在中国美术馆举办的“劳森伯格艺术展”,这对中国年轻艺术家产生了重要影响。然而,当许多年轻艺术家创作装置艺术时,他们还处于一种不甚了解的阶段,对装置艺术的概念、内涵和源头理解不足。改革开放后,年轻艺术家们通过非官方渠道获得了一些有限的信息,但对装置艺术的理解仍然非常陌生。他们的目的是利用装置艺术来打破文化大革命以来僵化的艺术体系,这也许可以解释为什么四川美术学院的两名学生创作的作品被贴上了“伤痕艺术”的标签,就像在国家艺术展览中获奖一样。正如评论家范迪安所言:“以新的造型模式展现艺术家的现代意识,甚至允许出现含糊的所谓‘反艺术’内容,以此质疑现有的艺术标准和原则。”至此,装置艺术一词被引入当代艺术,描述那些与传统艺术形式截然不同的作品,并毫无悬念地超越了传统的架上绘画,进入了主流艺术的激流。

随之,越来越多的中国艺术家开始关注装置艺术,一个里程碑事件是 1989 年在北京中国美术馆举办的“中国现代艺术展”,在这里许多艺术家默契地采用了装置艺术的形式,特别是在材料和特殊的表现感上。80 年代后半期不仅是先锋艺术的高潮期,也是年轻人出国的高潮期。这两个运动的叠加结果是,前者几乎被后者掏空。到了 90 年代初,至少有几千名年轻艺术家出国。中国先锋艺术运动中最活跃的艺术家和评论家大多数离开了中国,转而在巴黎、纽约、东京等城市继续他们的工作。因此,他们对西方现代艺术有了更现实和具体的理解。这一事实告诉我们,90 年代初,中国现代艺术的低谷期,许多作为现代艺术推动者的先锋报纸和期刊相继停刊,许多年轻艺术家和评论家也开始了一段“自我流放”的海外旅程,吸收外国当代艺术的养分。

在 90 年代中期,位于首都师范大学的美术馆 1994 年举办了文化部批准的“中韩日国际现代艺术展”,而 1995 年之后,举办了“北京-柏林艺术交流展”,“位移:四人装置艺术展”以及朱金石、宋冬、尹秀珍等人的个人装置艺术展,这些展览成为了那一时期北京重要的先锋艺术展览空间。在北京王府井翰墨艺术中心,徐冰以两只猪为主体制作了装置作品“文化动物”,赵半狄的“月光”,焦英奇的种植与编织作品,都给人留下了深刻的印象。此外,讨论会画廊和复中美术馆也经常有一些小型的装置艺术展览。90 年代中期,上海、广州等地的装置活动也开始活跃起来。

进入 21 世纪,随着观众艺术水平的普遍提高,中国的装置艺术越来越符合国际化的趋势。中国与海外艺术家之间的交流频繁,装置艺术展览在概念、质量、规模和制作标准方面

都有显著提升。这个时期，如蔡国强、徐冰等居住海外的中国艺术家也回到中国，回到他们的家乡，他们的代表作引人注目；这些各种展览反映了过去 10 年中国艺术家的最高水平。

回顾中国装置艺术的发展，许多艺术家如云烟般涌现，随后又淹没在当代艺术的洪流中，而那些始终坚持艺术信仰、不断以各种方式探索装置艺术语言和当代社会精神生活表达的艺术家是极其宝贵的。今天，纵观中国装置艺术的发展，尽管它短暂而辉煌。在当代艺术发展的背景下，装置艺术频繁出现在各种展览中，特别是像双年展这样的大型展览中，它强烈地引起了观众的共鸣和热情追求。中国装置艺术正逐渐以更成熟的方向发展和延续。

## (2) 中国的互动装置艺术

在中国，互动装置艺术因为依赖计算机技术和网络技术，而比外国落后了几年。中国的计算机图形艺术始于 1980 年代，随后的十年里，新媒体艺术开始涌现。1988 年，张培力创作了中国第一件视频艺术作品《30 X 30》，这是中国艺术作品中第一次使用视频作为媒介，他也被誉为中国视频艺术之父。在 1990 年代，国内的装置艺术和行为艺术不断发展，新媒体艺术也开始崛起；到了 90 年代中期，当中国艺术家的作品开始受到西方的重视时，中国本土的装置艺术也开始得到官方的认可，新媒体开始以装置艺术为载体进行设计和创作，但那时在中国，互动装置艺术仍处于起步阶段，并且这一时期主要是装置艺术、身体艺术和行为艺术的主导。在 90 年代中后期，中国美术学院的画廊举办了被称为中国第一次视频艺术展览的“现象与图像”，互动装置艺术也开始在中国出现，这一时期的新媒介是简单使用视频、声音和光媒介，而那时的互动装置艺术的互动模式大多是模仿外国作品。

随着学成归国的前沿艺术家们的回归，许多来自海外的先进技术和理念被引入，导致中国互动装置艺术的发展深受外国艺术的影响。自此，自 21 世纪初以来，中国的互动装置艺术发展迅速，越来越多的年轻艺术家开始尝试让作品不再沉默，而是与观众互动，并在反映某一主题和某种思想的前提下，做出有趣的互动表达，以获得更多关注。因此，中国的互动装置艺术从一开始的复制和模仿，到现在拥有了独立思想和表达，独立的话语权，独立的文化遗产和独特的艺术风格，并开始受到国际艺术评论家和公众的广泛关注。

## 3.3 互动装置艺术的互动模式

迪特·丹尼尔斯曾经提出问题：互动性是一种意识形态还是一项技术？在今天对互动性的讨论中，一方面是关于通过媒体构建的人际关系的社会意识形态，另一方面是关于人机通信的技术可行性。因此，装置艺术的互动性无非分为两大类，即：沉浸式和界面式。随着虚

虚拟现实技术和智能环境技术的发展，“沉浸式互动艺术”很可能成为未来装置艺术发展的主要方向。

### 3.3.1 界面式

界面式互动新媒体作品目前是国内外展览中最常见的艺术形式。其人机互动模式是：观众可以通过语言、触摸、手势或操控摇杆、遥控器等身体动作与作品进行互动，通过观众面前的屏幕或者“互动墙”生成反馈结果。

### 3.3.2 沉浸

沉浸式互动装置主要用于大型展览之中，如博物馆、世界博览会、双年展等环境，相对而言，展品的空间宽敞，适合多人与作品互动，用户体验感强，沉浸感十足，并且不受界面限制，个人自由度相对较大。

例如，加拿大卡尔加里艺术家凯特琳德·布朗设计的装置作品《云》，是由6000个明亮或烧毁的灯泡组成的互动装置，其互动设计是允许观者尝试拉动那些灯泡绳，他们可以随意拉动。你也可以选择关闭组成云朵的灯泡之光。这样，观众也成为了这个作品的一部分。这种互动行为创造了一个巨大的闪烁灯光云，在其中观众沉醉，享受它本身带给整个装置的变化，而作者利用沉浸式互动装置让人们从不同的角度看待和处理过剩物资（灯泡）。

## 3.4 互动多媒体技术

### 3.4.1 互动多媒体的概念

“媒介是人类感官的延伸”，互动多媒体作为一种新的媒介，给人们的生活带来了质的变化。通过文献阅读，本文发现近年来，媒体与人的沟通方式不断在变化，从单一媒介到多媒体，再从多媒体到互动多媒体，已转向颠覆性的变革。百度百科曾介绍：单一媒介指的是依赖单一方式进行通信。多媒体，顾名思义，是单一媒体的组合，它集成了文本、图形、图像、声音等各种通信方式，更为多样化。互动多媒体在此基础上更进一步，具有更强的实时互动性，是集成了“视频、声音、文本、数据和流媒体”的互动式通信。

Vaughn 将“多媒体”定义为文本、图形艺术、声音、动画和计算机传输视频的结合。

台湾当代艺术家吴定武在文章《计算机多媒体在视觉艺术中的应用》中进一步指出，多媒体是利用计算机、视频和通信等信息技术，对文本、图表、静态图像、动态图像、计算机动画、背景音乐、旁白和数字音效等材料进行数字化编辑、存储和转换，以全面展现或突出概念。因此，互动多媒体的概念可以归纳为：使用计算机互动集成技术和数字通信网络技术处理各种媒介表示的文本、图形、图像、视频和声音，使各种信息建立逻辑联系，整合成一个互动系统；这个互动系统被广泛称为“互动多媒体”。

### 3.4.2 互动技术

提到互动多媒体技术，与之相关联的互动技术及其密切相关。互动技术，也被称为交互技术，根据百度百科所述：“互动技术”这一概念最初由智能互动新媒体国际有限公司提出：“一种通过先进的视频运动捕捉系统和成熟的三维游戏引擎，将传统空间转变为新颖互动空间的技术。”在互动空间中，人们可以通过身体动作（如身体摆动、行走、手势等）控制相关的音频和视频以及游戏元素。互动技术打破了传统的静态空间模式，将观众作为空间的一部分进行整合。在搜索“互动技术”时，百度百科首先对“交互”这个词进行了定义：指的是自然界和社会中智能体、数据、技术知识传输和交换活动，从信息论的角度看，是将各种信息材料集中到一定区域的活动，是有形的文本信息载体。互动技术是使用某些手段来实现互动目的。可见，在中国对互动技术的研究相对年轻，其确切定义还不甚明确。

目前，在商业领域，根据使用方式，互动技术通常被划分为三种类型：墙面互动技术、地面互动技术和桌面互动技术。虽然成本相对较高，但应用越来越火热，最初多见于卡拉OK、餐厅等娱乐场所，而现在其应用也在高校、大型展览、商业广告、游戏等其他领域逐渐增多。互动技术在艺术领域的应用研究也已经开始，例如，在2008年的上海双年展上，上海交通大学软件学院的数字艺术实验室展示了“虚拟鱼池”，这是一种使用虚拟现实技术的地面互动方式。维也纳的音乐墙，一般费用约为30万元人民币，属于墙面互动技术。

### 3.4.3 互动多媒体技术

理解互动多媒体技术首先要理解多媒体技术。多媒体技术是利用计算机对文本、图形、图片、动画、声音和视频等多种媒体信息进行数字化处理、分析、传输以及互动应用的技术。互动多媒体技术可以帮助艺术家们理解这个概念的定义及其应用的实际意义，并且挑选出对

艺术创作有帮助的互动多媒体技术类别，使用这些技术可以帮助他们取得更好的作品效果。互动多媒体技术，包括触摸互动技术、手势互动技术、笔交互技术和语音激活互动技术，这些都是目前已知应用较多的互动技术，但因为其较高的门槛，虽然一直受到装置艺术家的关注，但使用并不广泛。

该互动方案是在 Adobe Flash CS5.5 软件平台下构建的，并且基于 Flash ActionScript 3.0 的实现，使用 Flash 的动画界面和后台语言来实现互动解决方案，主要面对两个问题：电影剪辑的随机组合以及控制电影剪辑的播放。结合以上两个功能可以实现预期的效果。使用这项技术实现互动分为三个阶段：1) 实现电影剪辑的随机组合。2) 实现单按钮的视频随机播放。3) 实现多按钮的视频随机播放控制。这种方法的优点是：不需要强大的技术团队配合，成本低廉，且只需会使用 Flash ActionScript 3.0 编程的技术人员就能实现界面互动，并且可以在短时间内完成工作。

除了使用 Flash 软件实现互动方案之外，手势互动也是最佳方式之一，但是由于后面提到的工作中具体技术应用受到技术研发的限制，并没有应用手势互动技术。作为一种自然和直观的互动通道，手势互动在人机互动过程中扮演着重要角色。基于 Kinect 的手势互动技术结合了手势定位、手势检测与识别，能够自动提取操作者的深度信息，并提供动态捕捉、图像识别、麦克风输入、语音识别、社区互动等功能，无需使用任何控制器，就可以依靠摄像头捕捉操作者在三维空间中的动作趋势和关节运动轨迹，可用于各种类型的显示和交互系统。与界面互动相比，手势互动技术在提升观众的沉浸式体验方面具有优越性。

此外，视频编辑技术也是本项目实践中的主要多媒体技术之一，它是交互式装置艺术中常用的技术。许多交互式装置艺术作品的设计都是基于图像的。这项技术使用电影电视后期制作中常用的软件来编辑和调色所拍摄的素材。它通常包括素材采集、剪辑、后期调色等过程。在互动多媒体展示中，图像往往是主角，因此编辑技术往往伴随着多媒体作品的诞生而存在，并与之并肩发展。

#### **3.4.4 互动多媒体技术的应用领域**

作为数字通信的颠覆者，互动多媒体技术改变了与人沟通的传统方式。目前，最常见的应用是在大型展览活动中，如科技馆、规划馆、博物馆等地，尤其是科技馆用来讲述海洋、宇宙诞生等科普知识的场馆，对许多儿童来说，实现了形式与神的解释和展示，清晰明了。韩立伟在文章《多媒体技术在博物馆展览中的应用》中提到：“作为辅助展示的有效手段，

多媒体技术满足了信息时代和多媒体时代人们对博物馆的新要求,有助于发挥博物馆在宣传、教育、科研及其社会功能方面的作用,并对文化传承和推广具有广阔的前景。”正确处理和应用这项技术,将对文化和文化事业的发展大有裨益。

其次,互动多媒体技术在艺术设计领域的应用,在当今的大型展示艺术、雕塑、纤维艺术、视频、音乐以及其他新媒体,特别是装置艺术中频繁应用,并且其展示效果和互动效果稍优于传统艺术。本质上,互动多媒体艺术并不缺乏吸引力,通过改变人们之前与传统艺术交流的形式,它能从观众的个人感受中获得更广阔和深刻的信息,进而触发观众的思考,带来更多艺术家的思索。

此外,互动多媒体技术在日常生活娱乐中的应用更为常见,许多超市促销、酒店广告展示、酒吧多媒体音乐墙、卡拉 OK 包厢设计乃至大型购物中心、广场等都开始尝试使用多媒体技术进行展示、推广和宣传。除了服务行业之外,互动多媒体技术也被用于婚庆庆典、戏剧、舞剧、传统戏剧中,它们开始尝试使用多媒体技术来设计舞台布景,丰富的模拟场景仿佛触手可及,更好地呼应剧情需求,视觉上更为逼真,使得观众在观看剧目的同时感受到的沉浸感和真实感更为强烈。最后,互动多媒体技术在学校教学中的应用,不同于传统的黑板单一媒介教学,互动多媒体在教学中的应用和普及是教学方法创新的一章意义重大的篇章,也是对传统教学方法的补充和延伸。

### **3.4.5 互动多媒体技术应用的优势**

首先,互动多媒体技术对观众具有很强的吸引力,在 2010 年世界博览会的各个馆展中,多媒体技术的应用使得每个馆展都变得光彩夺目,特别是美国馆、加拿大馆、沙特馆以及国内的中石油企业馆等极大地吸引了公众,为观众带来了新奇和神秘的展示效果。其次,它对观众具有很强的可观性:强烈的视觉冲击力、丰富多变的展示效果极大地提高了观众对作品的可观性。例如,许多虚拟现实的互动装置引领观众进入一个虚拟的梦幻世界,随着互动的进行,观众沉浸在场景中,成为美妙体验中互动装置的一部分。最后,它对社会具有强大的经济效益。任何一个互动多媒体项目的开发与设计都具有巨大的经济效益,仅水晶石公司的世博展示项目就为公司带来了巨大的利益。因此,互动多媒体技术与装置的结合,这种实时互动让观众既能看到宣传内容,也增加了其知名度,娱乐了大众,可以说是一举三得,是社会经济效益的创造者之一。

## **4. 互动装置艺术作品系统的总体结构**

### **4.1 总体结构**

它包含四个主要内容，分别是：计算机多媒体处理子系统、信息检测与采集子系统、传输子系统和媒体呈现子系统。由此结构衍生出的互动视频装置艺术的展示效果清晰地说明了这一结构的合理性。

#### **4.1.1 计算机多媒体处理子系统**

计算机多媒体处理子系统作为互动作品的基础，相当于电脑上的 CPU，就像人脑一样。信息接收并直接传输至装置，通过互动程序发出指令，控制图像的生成和播放。信息检测与采集子系统的信息通过传输子系统直接在媒体呈现子系统中成像，并判断观众的互动反馈，通过计算机程序、计算机控制和技术处理，将程序语言“翻译”成视觉语言，这是整个互动装置艺术作品的核心。

#### **4.1.2 信息检测与采集子系统**

信息检测与采集子系统就像人的眼睛或指尖，主要用于检测人体传感信息，常用的工具有红外传感器等，能够快速侦测到某个场域内人的存在，除此之外，温度传感器可以通过人体温度感知“人”进入工作空间，然后向计算机多媒体处理子系统发出指令，它是感知外部信息的重要触角。

#### **4.1.3 传输子系统**

传输子系统就像人的敏感神经末梢，是连接人类感官与大脑的线路，是多媒体处理子系统和信息检测与采集子系统之间的桥梁。其畅通无阻是确保互动装置交互性的前提。它分为无线和有线两种传输介质。

#### 4.1.4 媒体呈现子系统

媒体呈现子系统就像人的面孔，它吸引公众注意力的所在，是作品外观的体现。以互动视频装置为例，精致美观的视频图像是吸引观众的重要部分，除了图像之外，通过声音的渲染和光影的变化，为观众实现 360 度的感官刺激，使观众产生参与互动的兴趣。此外，投影媒介的变化也直接影响到作品的呈现效果，比如近年来流行的球形巨幕，世博会加拿大馆的多面体六面屏等，视觉上呈现出强烈的视觉冲击力和新鲜感。

### 4.2 作品的创作过程

在创作过程中，经过初步调研、文献研究和讨论之后，互动视频装置作品《路非我独行》的创作过程得以产生。

### 4.3 用户分析

#### 4.3.1 用户的范围和类型

由于互动视频装置着重于交互性，交互是以观众为中心的。只有了解潜在的观众是谁，才能对用户进行深入研究。因此，对于互动视频装置作品的潜在观众，主要对中国大陆的潜在观众群体进行以下分类分析：

(1) 从性别相似性来分，可以将观众分为男性和女性。男性观众更重视视觉冲击力，而女性观众则更关注观看后的感受和启发。

(2) 从年龄段来分，可以简单划分为 18-22 岁的大学生；22-25 岁的公司白领；25-30 岁的青年人中期观众；30 岁以上的青年后期观众；40-50 岁的中年观众分为几个类别。他们的共同特点是热爱艺术或从事艺术设计工作、艺术教育等职业，并且在艺术展览活动中更为活跃。

(3) 从文化层次来分，可以简单划分为高等教育及以上。

除了上述的性别结构、年龄结构、文化结构之外，还有决定作品本身潜在观众群特性的生活差异结构和职业结构的不同。在这些因素中，职业和文化结构起决定性作用。经过分析，其观众主要针对的是学习设计或从事设计的人员、大学艺术教师，以及热爱当代艺术的考古爱好者，因此，作品在媒体呈现子系统的设计与选择上应具有一定的美学性，界面设计力求

精美，画面剪辑精致，并且色彩分级接近电影质感，以吸引这些专业观众。

#### 4.3.2 用户体验与需求分析

用户体验分析，简称 UE 分析，是艺术作品创作流程图中“性能评估”部分。在装置作品中，把用户称作观众更为合适。用户体验，顾名思义，是对观众在体验互动视频装置时的感受和互动体验进行的分析。这一节中互动视频装置作品的互动设计属于界面交互类型，所以可以从三个方面分析观众的纯主观感受：用户需求、装置作品的互动可操作性和互动反馈，也可以包括用户情感体验、界面偏好、生理和心理喜好及互动方式的分析。

此装置面向的是具有一定美学倾向和对中国当代艺术关注的普通公众。这群人热爱艺术，对各种艺术展览和前卫艺术展充满热情。因此，此类用户的需求是新颖多变的，像是视觉刺激或怀旧倾向，并且愿意体验和接受新事物。基于以上用户特性，图像装置界面设计的视觉感必须具有一定的美学和视觉冲击力，以满足用户的视觉习惯、情感和体验需求，吸引这类用户的注意。

#### 4.3.3 用户认知模式分析

本项目发现，在互动图像中，特别是针对基于界面的互动图像设计，用户的认知方式更容易被动态视频而非静态界面所吸引。经过问卷调查和现场模拟测试后，判断用户对装置作品的互动操作便利性非常敏感，这直接决定了您的作品能否在其中。

为了在大空间和场馆中吸引足够多的观众，因此，在作品的设计过程中，采用了三幕并置窗口的形式开场视频，以主观镜头播放用户在车内通过车窗看到的移动景象。一旦用户被吸引并愿意参与互动，进入现场，作为信息检测与收集子系统的红外传感器将发出指令给多媒体处理子系统，控制媒体展示子系统停止视频播放，进入互动界面。这样的设计易于操作，只要用户进行移动和点击鼠标，互动就展开，接下来的视频开始播放。因此，如果想要互动效果更好，视频装置作品的互动可操作性尤为重要，观众应该非常容易投入，无需费力思考如何进行互动、如何做手势、移动步骤或点击鼠标。

#### 4.3.4 用户交互行为及预期效果分析

将作品的娱乐属性作为吸引观众的重要手段无疑是有效的。因此，在研究互动视频装置

的过程中，发现互动行为决定了互动视频装置的吸引力指数。当然，手势互动、触摸屏等手段更容易引起观众的互动行为和兴趣，但考虑到这种设计和技术支持的成本，更倾向于选择界面互动方法，这种方法相对传统和经典，只需通过增强作品本身的内涵和由此引发的反思程度来增强互动效果，预期效果会更加乐观。此外，在研究过程中，模拟了互动效果，并对用户参与互动行为的反馈进行了实验，包括在预展期间记录和收集用户对视频装置作品的反馈，这极大地帮助了研究改善用户体验，并成为重要的依据之一。在具体测试过程中，随机对20位顶尖的参与者进行了访谈，他们普遍的反应是通过参与互动和观看作品而被触感和反思。

本节首先从装置艺术的概念引入互动装置的概念，并对国内外互动装置艺术的发展进行了详细介绍；分析了互动装置艺术的交互模式，并分别列举了交互模式中的界面和沉浸感。此外，讨论了互动装置艺术作品系统的整体结构和各子系统之间的连接，最后解释了作品的创作过程和用户分析。

## 5. 作品《路不由己》的设计与实现

在早期对互动多媒体技术的理解和对互动装置艺术的研究之后，本节将利用互动多媒体技术来完成作品及其互动部分，该作品命名为《路不由己》，试图结合图像、计算机和投影技术，实现作品与观众之间的互动。

### 5.1 设计理念

本节利用汽车不断前行的窗口作为生活的隐喻，采用公路电影的方式作为视频装置的开篇内容，分为三个屏幕，代表从司机视角的三个窗口。在观众进入之前，它将持续播放，直到观众进入现场，与作品互动，并切换到一个静态界面。界面上有7个静止屏幕，代表“我”的生活和几个代表性的生活快照。与观众互动的想法是，观众任意选择一个快照作为视频回放的“起点”，以此开始切换到这个生活画面，其余六段随后随机播放，从而让观众看到因为他（她）的介入，“我”的生活经历的顺序发生了变化，以此告诉观众：你的一次无意碰撞可能改变了“我”的生活。通过这样的互动，引出了一个反思：善待你周围的人，你的一次无心行为可能导致她的生活轨迹发生变化，从而导致她的生活不由自主地发生转变。

互动影像装置的具体概念：静态界面是一个任意可选的界面，观众随意选择一个起点开始播放影像的实质性章节。画面以“我”的生活的7个片段作为视觉呈现的核心，在一个10平方米的场域内，在暗示客观时间持续旅行的三幕窗口移动镜头的大屏幕上，通过观众

随时的打断，进入时间选择界面。在界面中由时间空间构建的无序的7个时间片段中，每一个“不速之客”既是参观者，也是闯入“我”生活的主角，他们的出现和1-7起始时间片段的选择，观众自由选择一个时间回放的起点，随机改变“我”的故事、惯性和方向。

每个被选择的时间片段可以被理解为一个时间点，在这个时间点上，一个时代结束而另一个时代开始，惯性继续存在。使用这个镜头来做一个隐喻：在一个时间片段的结束时，“我已经死去，我没有意识到这一点，而是另一个人在为我活着。这个影像应该是在通往未来的路上最宝贵的礼物，献给我的过去，和那些无法描述的生活与感受。期待我回顾我的人生前半段，并预测我的人生后半段的可能性，这件作品是“我”与世界之间的一次庄严、公开的对话。

### 5.1.1 设计理念

#### (1) 回忆与沉思的美

美丽散布在世界的每个角落，以各种形式存在，没有高贵与低贱、善与恶的区分，我们取所需，并按照我们的欲望赋予它意义。有些美丽具有温度，让人向往光明；有些美丽使人无力，却也获得平静；有些美丽激荡人心；还有些美丽因为慰藉而永不熄灭。这件作品不试图讲述一个深奥的寓言，而是希望挑选几段过去的记忆和对过去生活及未来的几分期待，并以影像还原的形式表达本文对不同时间空间的沉思。

本质上，这件作品并不试图探讨任何深奥的主题，而只是期待在与作品的互动中微笑，同时提出一个问题：我们一生努力探索的某些事物，难道仅仅是我们的想象，空手而来，费时费脑之后仍空手而归，那么何不将其视为“黄粱一梦”，简单地一笑置之，云淡风轻呢？如同善待春夏秋冬一样，善待生活中的酸甜苦辣、苦乐和悲喜，生活中的一切都只是路过，路过之后就放下，这就是获得幸福生活的源泉。

#### (2) 时间与场域

时间的概念：我们常常思考拥有一刻、一指弹或一瞬间需要多长时间。依照《僧一行》的说法，1念为1念，20念为1顷，20顷为1弹指，20弹指为1罗预，20罗预为1须臾，1昼夜有30须臾。转换结果：须臾=48分钟，弹指=7.2秒，顷=0.36秒，罗预=1念=0.018秒。须臾>弹指>顷>念=1思。

借用上述数据，本文在一个刻意规划的时间和场域中放置了7个关于自身的冥想片段，

试图表达一种看似神秘的前世、看似无意识的今生，所有这些不过是一个午后冥想的瞬间，仅仅是一瞬，是整个人类时间线上的一刻，以一个指弹或一念之间，这种奇妙而不容置疑的领悟在互动视频装置的形式中得以再现和视觉化。

在视频装置中，时间的解读不可避免地引出另一个词：“场域”。这个术语在艺术界的使用可以追溯到心理学的格式塔学派的马克斯·韦特海默（Max Wertheimer, 1880-1943）受当时物理学法则的影响，认为就像物理世界的“场域”一样，在人类的直观世界中也应存在一个非常类似的“场域”。相对于人类的视觉世界，可以称之为“视觉场域”，与人类的生活、学习等情境相关的是“感知场域”。在作品《路，它不由自主》中，空间是视频装置作品展示所必需的、甚至不可或缺或不可缺失的元素，此时的空间成为观众与作品交流的媒介和平台，如图像、界面的视觉图像和音乐等融合在一起，形成一个沟通的气场，特别是互动过程，使得参与者与其场域，以及本文中勾勒的真实与虚拟交错的空间完成“能量”和“气”的转移，感受本文中的生活片段。

### 5.1.2 听觉元素

为了突出时间交错的感觉，将使用爵士乐、老上海风格以及无歌词的纯音乐作为配乐，整个影像没有对白，只有文字叙述，混合音乐讲述作者对某一思想的冥想。

### 5.1.3 文本剧本创作

视频装置就像是切割了一段电影片段给你看，这种盲目的反思充分诠释了图像对艺术的教训：所有图像和所有经验都必须被拖入同一空间，使它们成为公共的，在一个共同的尺度前相互比较。在《电影史》中，戈达尔（Godard）将伊丽莎白·泰勒（Elizabeth Taylor）的微笑搭配在乔治·史蒂文斯（George Stevens）拍摄的集中营镜头上。艺术谴责的责任最终会被见证的责任所掩盖。剪贴和搭配让不同的事物共存。这一概念直接影响了本文的创意构思，所创作的互动视频装置是基于碎片化镜头的随机拼接和播放。因此，剧本设定由七组镜头组成，这七组镜头都有意识流风格的文本，没有特定的顺序，每组镜头的副本如下：

1) 那场雨，把她淋回了1930年，她看见了自己的前世，她撑起一把伞，优雅地走过雨巷，雨滴穿透了所有的风情。她想要画出一个前世，她没见过前世，也不可能见到。但知道自己曾美丽，她画出了前世走过的楼梯，传递信息的笔触，镜前易装的日子，以及因爱她而

熄灭的诸多心火。

2) 这个女婴的诞生是对奶奶最大的回报。在婴儿的手上挂上那个古老的小银锁，道路也随之加快了速度（镜头花卉穿过）。新生儿，通透无色，但世界是巧妙的，让她绽放，嬉戏于草间花丛，尽管有千万条奇异的路等她来选择。

3) 学会发呆地看一朵花，让时间流逝，不惜一切代价。这就是她想要感知的世界，在路上行走，想要触摸周围的一切，仿佛触摸能让她感觉到这些东西也像人一样拥有生命。感知树下的阴影，感知对世界的一瞬思考和困惑。感觉到自己被落花环绕，她不忍踏出步伐踩碎一朵花，相信自己一生中永远不会随意伤害生命、善良和美丽，所以生命的步伐有了节奏。她忍不住捡起了自己的童年。

4) 想尽一切办法去翻越那红色的绳索，不想把失败翻落在自己手中，但不想如此，未来的路况怎能预设，做一个自己最喜欢的样子，机会只有一次。

5) 生活的步伐不均，无缘无故地流淌通过生命。那些已经走过的路径，就像一片落在地上的香灰，被描述为静止的，想要拾起，却又无法捡起。她每天看见自己，但却看不见变化。悄无声息地，她从少女快进到了职业女性。

6) 曾经，她童年的愿望在树枝间破碎，而今天，她的思绪也被囚禁在树隙中，谁给了钥匙，却打不开胸口上的锁。她的时间从一个教室迁移到另一个教室，不知不觉间，她的路途变得苍白。

7) 半日读哲学，读诗，甚至美学，中国传统学问和佛学，她依然如同雨天的衣物，无处安放。回望过去的路，漫长且曲折，想要抹去那些忧伤，印迹却是如此之深，宛如在梦中迷雾森林里徘徊，索性，放下书本，上路，睡到下一世。

## 5.2 互动设计方案

(1) 利用红外感应和微控制器技术来控制计算机实现互动，控制 7 段视频的随机播放。

(2) 利用计算机屏幕的投影仪交互，Flash 编程实现 7 个视频的随机播放。

(3) 使用 Kinect 的手势交互完成互动，即结合手势定位、手势检测与识别以及触摸屏技术的界面互动。

考虑到制作成本，最终选择了第二个方案：界面互动。为了有效且可行地整合程序、音乐和图像，这个作品的互动设计选择了 Flash 软件作为主要程序编写工具，其强大的 Action Script 2.0 能够控制 7 个视频的播放。这一创作主要以互动装置的形式呈现，在投影屏幕前设有一个红外感应区，当观众进入该区域，计算机控制屏幕上快速移动的窗景视频将会减慢，

如果观众参与活动，将被程序控制进入静态界面。

## 5.3 制作流程

### 5.3.1 确定作品的计划和结构

正如上文所引用的：“所创建的互动视频装置作品的整体结构包含四个主要内容，即：计算机多媒体处理子系统、信息检测与采集子系统、传输子系统和媒体呈现子系统。其中，计算机多媒体处理子系统的核心技术是在 Flash 软件中编写 Action Script 2.0 实现程序，由计算机实现数据控制；信息检测与采集子系统通过界面互动中的鼠标控制来实现；传输子系统起到承上启下的作用，连接前两个子系统，并通过数据串口将计算结果连接到投影仪，即媒体呈现子系统；媒体呈现子系统通过投影仪投射到大屏幕上，屏幕以三幕并列的形式呈现，以突出媒体呈现子系统和图像素材的特点，展现其鲜明的一面。”

### 5.3.2 确定分镜头脚本

分镜头脚本是图像装置创作中不可或缺的准备工作的，它的作用是指导摄影师进行拍摄，指导编辑进行后期制作，以及帮助主创人员了解脚本的蓝图。分镜头脚本详细描述了每一个镜头的内容、角度、运动、镜头时间以及配合的声音和特效等，确保创作过程中每个环节都能准确地实现创作初衷。

### 5.3.3 拍摄

摄影设备包括佳能 400D 单反相机、三脚架、拍摄用柔光灯、反光板、遮光罩、手持稳定器等。具体拍摄场景分为四组：

1) 在雨天的新天地房屋和公寓里，拍摄第 7 组关于过去生活中穿旗袍的老上海女人的镜头，以及第 6 组关于年长的“我”的镜头。

2) 雨天的上海建桥学院，拍摄第 3 组和第 4 组关于“我”成长和教学生涯的镜头。

3) 阳光明媚的上海建桥学院，拍摄第 1 组和第 2 组关于“我”婴儿和童年时期的镜头，成人镜头和工作镜头。

4) 阳光明媚的上海建桥学院，拍摄第 5 组镜头。

### 5.3.4 视频编辑和后期制作过程

#### (1) 使用软件 Sony Vegas

在PC上较为流行的编辑软件有 Adobe Premiere 和 Sony Vegas。Premiere 支持的插件最多，且与 AE (Adobe After Effects)、3ds Max、Maya 这些视频效果和 3D 软件的兼容性好，但对计算机资源的占用非常高，需要较好的计算机配置才能流畅使用，适合对后期效果和高级合成要求较高的视频。而 Vegas 占用的计算机资源相对较少，低配置的电脑也能够流畅预览，低配置电脑导出视频的速度也是 Premiere 的近 10 倍，支持的视频格式特别多，基本上不会遇到需要转换格式才能编辑的情况，这次视频的效果和合成要求非常小，因此选择 Vegas 作为编辑软件。

#### (2) 整理素材并放置于统一文件夹中

对于场景较多、跨越时代较宽的视频项目，最好根据场景和时代将素材存放在不同的文件夹中，这样可以减少编辑寻找素材的时间，降低出错率。文件夹命名不能使用中文，文件的整个路径也应避免使用中文。使用英文和数字命名会相对稳定，软件运行不易崩溃。

#### (3) 导入素材

根据现场标识板上的镜头编号，将现场记录员记录的所有成功的镜头导入到剪辑软件的素材库中。

#### (4) 初剪

初剪的目的是筛选出所有无瑕疵的镜头，剪除每个镜头开拍后和停拍前不能使用的部分，按照事情发展的逻辑顺序和时间顺序理顺故事，并确定没有逻辑或时间上的剧情问题，以便及时重新拍摄。那些拍摄多次但可以使用的镜头，也要并行存放在单独的轨道中，方便剪辑时选择。

#### (5) 音效处理

在苹果机上最专业的音频软件是 Pro Tools。该软件算法复杂，音频、MIDI、视频都能得到良好支持。由于其不同的算法，在单纯的音频方面，其播放和录音的音质比我们现在在

PC 上流行的各种音频软件要好得多。在 PC 上的音频软件中，Adobe Audition（Adobe 收购 Cool Edit 后的新名称）应用最广泛，而 Nuendo 是大多数专业人士的选择。

（Cubase 和 Nuendo 来自同一家公司，他们有相同的界面，但 Nuendo 在与视频结合的音乐制作方面更加强大。）由于我使用的是 PC，我选择了 Nuendo 作为我的音效录制软件。根据需要，将视频需要添加的音效添加到影视剪辑软件的素材库中。如果找不到合适的音效，需要自己录制音效，使用 Nuendo 软件进行录制。录制时，应该将初剪的视频导入到音频软件中，然后创建一个新的轨道，这样便于录制音效，并且方便影片的声音和画面对齐。

## **(6) 精剪**

根据想要达到的效果对故事叙述的顺序进行非线性编辑，比如使用插叙、倒叙、平行双线索等手法。适当使用视觉蒙太奇技巧，使效果更丰富。剪掉初剪中冗长的部分，这些部分不影响剧情。视频的初剪节奏感相对较弱，应根据镜头感和初剪的音效来做节奏和氛围感。

## **(7) 调色**

目前已知的顶级专业调色软件有两款，DaVinci Resolve 和 Nucoda Film Master。本工作中使用的软件是 Nucoda Film Master。DaVinci 目前是最广泛尝试的顶级调色软件，但硬件安装的最低配置相对较高，目前 1 万元以下的 PC 台式机和 2 万元以下的 MAC 台式机不允许安装。Film Master 的原始硬件比 DaVinci 更昂贵也更专业，但最低配置限制较小，六七千元的 PC 台式机可以安装使用，但调色后，如果需要编辑，1080P 的电影每帧是 8 兆字节，每秒 200 兆字节，没有双通道固态硬盘，软件中无法实时播放。你需要导出视频才能流畅播放。

Nucoda Film Master 的调色优势包括：支持数字中间片处理的各个方面，无限层数的色彩增强，可以在剪辑片段、序列中应用风格和效果，准确分割图像或场景的区域，实现特效参数的动态分级和动画，每一层可以添加、减去和结合任何形状的 Alpha 通道，跟踪和操作独立变换中的形状或基础样本。因此选择使用 Nucoda Film Master 进行调色。根据每个剧情的需要，调整每个镜头的色调，并注意控制一组镜头之间统一的色调关系。在调色过程中，应密切关注整体画面的色彩平衡和视觉连贯性，确保最终成片的视觉效果既符合剧情氛围，又保持整体风格的一致性。

### 5.3.5 精剪

1) 首先调整画面构图，原始影片是使用车载吸盘拍摄的，因此不便调整角度，所以构图有些偏上。使用 Vegas 软件对画面进行放大，以调整视觉中心点。

#### 2) 剪辑素材

通过停车再继续拍摄的方法，让女主角每次穿过花坛时换一次衣服，象征着成长了几年。使用的辅助设备是车载吸盘，车辆在每个经过的花坛前停下，等待演员换装后，然后同步启动。因此，需要无缝衔接车辆减速和重新加速的动作。就像画面一样，你应该仔细调整转换曲线，并反复调整，直到感觉到车辆中间有一个暂停的感觉。

#### 3) 添加音频

使用 Vegas 将片段导出为 wmv 格式的小样本，导入到 Nuendo 中，观察脚步声并添加行走的声音，并调整使声音和画面完美对齐。添加所需的背景音乐和音效。

#### 4) 导出视频

从 Nuendo 中导入完成的音轨回到 Vegas，使用 Vegas 导出未压缩的 AVI 格式超高清视频。

#### 5) 调色过程

调色需要解决的问题包括：主体不突出、与背景无法分离、景深无法打开、对比度过低、细节不清晰、色彩层次不够丰富以及边缘过于柔和。

##### 1) 初级调色，恢复白平衡。

由于白平衡已经非常准确，只需要根据人物肤色微妙调整，所以不过多介绍。

##### 2) 一级调色，呈现出需要的大体色调。

因为当时道具不够精致，在阴天环境下，光影相对较弱，导致裙子的颜色与背景过于接近。并且穿裙子应该是夏天的氛围，实际上这是春天拍摄的，所以树叶的颜色应该是深绿色，灰暗绿色，鲜亮的翡翠绿，原始片效果如图 4-11。所以首先将整个画面重调区域的绿色推向蓝色系统，即从黄绿色和中绿色向蓝色系统转变，使树叶符合夏天的色调。中灰色的地面大体上不受影响且不变色。

##### 3) 次级调色，将局部乱序的颜色调整回正常值。

从色谱上不难看出，裙子的绿色不够，因为绿色并没有明显多于蓝色和红色，所以不会出现过于鲜亮的绿色。勾勒出裙子区域的形状，并设置内外羽化的范围。在动态遮罩范围内调整重色、中色和亮色至绿色。使裙子的绿色整体接近原片。在色谱和具体参数中也容易看

出，裙子已经转为绿色。

4) 三级调色，控制前、中、后场景的景深和对比度。

这个镜头的前后场景比较虚化，不需要过多调整。然而，中间场景不够锐利，与后场的区分度不够突出。因此，在形状动态遮罩区域增加裙子的色阶，使中间和后场景清晰分离。

对比度过大，颜色显得过于鲜亮，因此在形状动态遮罩区域将裙子的饱和度降低至 70%，这样裙子在细节上既得到了锐化，也获得了适当的明亮色彩，与前后场景显著区分开来，同时又不会过于跳脱，与整个画面的色调不协调。

本节重点讨论交互式装置艺术的概念和设计理念，并对设计构思做出详细规划。详细解释了整个作品的设计过程。主要完成了 Flash 编程的交互设计，并组织和分析了拍摄与剪辑的技术，最后讨论了交互式装置作品的交互效果评估，实现了交互视频装置作品《道非自》的制作，为其他艺术家进行类似创作提供了技术参考。

## 第2节 干预与整合 - 交互式媒体艺术 中的互动与参与性表达

### 1. 干预 - 交互式媒体艺术中公众参与的互动表达

#### 1.1 理解媒介与相关概念

在《旧唐书·张兴政传》中，首次提到“媒体”一词：“观古今之用人，必因媒介。”这里的“媒介”指的是使得双方产生关系的人或物。在《文中子·为相》中说：“见誉而喜，媒介于欲也。”“媒体”一词历来是指位于两者之间的中介物或装置。

在新媒体艺术之父马歇尔·麦克卢汉的著作《理解媒介：人的延伸》中，对媒介的解释如下：“在一般大众的观念中，媒介仅仅是承载信息和知识内容的容器，一种表现方式，本身是简单和表面的，而传递媒介内容是它的功能，在现实中，媒介相对于所传达的信息和内容有着强烈的反应性。”媒介的形式就是媒介，而媒介的内容、信息和知识只不过是媒介的另一种形态，媒介作为一种形式是活跃的、动态的，对信息和知识有着显著的影响，并决定了信息的清晰度和结构。在艺术界，“媒介”通常体现在艺术家和技术人员的合作中，使用不同的艺术表达方法，比如计算机技术结合高科技、网络艺术、视频装置艺术等，与公众共同创作。

交互式媒体艺术不同于传统艺术的单向传播，它采用了一种新的参与互动的表达形式，具有强烈的虚拟性和概念性。通过新型媒体构建非现实的幻象，对当前问题进行假设或虚构，并通过电视、视频、互联网以及自然硬件安装等手段积极投入到流行文化中。艺术家通过交互式媒体艺术整合了他们的理性和感性思考，并使用科学技术和现代媒介技术作为技术支持的艺术形式。新媒体艺术是创造、欣赏、携带和评价艺术表现方法的手段，以计算机和互联网为技术支撑，引入新的艺术形态，并因此在艺术的审美感受、体验和思考的许多方面产生了深刻的变化。

20世纪初，由于印刷等新媒介的出现和无线电、计算机及电子通信等科学技术的发展，传统艺术得到了扩展。其中，立体主义和未来主义的出现标志着传统艺术不再是艺术家创作的唯一方式，它打破了延续数千年的西方艺术创作规则，创造了革命性的观察视角和表现世

界的新技术，不再将近远的透视作为创作标准，而是破坏和解构画面上所有的物体形象，并主观地进行拼接和组合。[不完整，重新输入]

巴勃罗·毕加索的《一位女士的肖像》是一件运用新材料完成的拼贴画作，这种作品将艺术表达从画布延伸到更多元的平面之上。未来主义试图在二维空间中追求二度甚至三度空间，并且揭示出肉眼无法看见的解构和四度空间（时间）。从原始社会发展的角度来看，由于技术水平的限制，未来主义和立体主义竭尽全力在平面上追求动感，它们的出现和迅速成长在很大程度上促进了艺术从二维向三维的发展，打破了时间和空间的固有墙壁，艺术表现方式从纸张到画布的转变，作品之美从静态向动态的追求，推动了当今互动媒体艺术的出现。

## 1.2 交互媒体艺术的互动特性和公众参与体验的特点

1917年，杜尚将从商店购买的一个“小便器”命名为“泉”，并将其展出于博物馆，这开了“现成品艺术”的第一枪；1962年，在裴南俊的《向约翰·凯奇致敬》中，他在音乐表演期间打碎鸡蛋、向观众撒大豆，打破了观众的期望。这些传统的艺术表达方式让人感觉新颖，但也让人一时间感到难以理解和困惑。20世纪50年代，欧洲出现了实验性艺术活动，充满了大量以工业和建筑材料为主体的艺术；60年代，便携式摄影和录像机械设备慢慢出现，先进的艺术家开始结合这项技术，这是媒体艺术的起点。到了80年代，录像艺术在世界各地的展览中已经占据了一席之地，并逐渐开始成为与货架艺术和装置艺术并驾齐驱的主要艺术媒介，在这个时候，艺术家完成一个作品不再是一个人独自闭门造车，而是与每一个领域的技术人员合作，传统意义上的观众地位也在改变，不再仅仅是“接受者”，而可以主动触摸、感受和与作品互动，从被动的观看者成为作品的发起者甚至作品的一部分。所有这些都使得媒体艺术的互动特性越来越明显，互动艺术作品也变得越来越多样化，从最初的触摸式高压气瓶、摄像机追踪人类视觉，到今天借助尖端科学技术创造的互动艺术展品，这些都增强了展示技术，例如无人机。

交互媒体艺术无论在内容还是形式上，都与传统艺术形式拉开了距离，形成了鲜明的对比。创作者和观众不再是彼此独立的，而是在观众的参与下共同完成整个作品。在观众与作品之间的互动中，由于交互媒体艺术的双向性质，观众从被动的“旁观者”转变为“参与者”和“互动体验者”，从被动变为主动。

## 1.3 交互媒体艺术在移动终端上的互动特性体现

随着社会和科技的发展，手机作为我们日常生活中的通信工具，极大地改变了人们在日

常生活中的虚拟世界与现实世界的交流方式。现在，手机除了基本的通讯功能外，还有更多的智能应用，比如发送和接收电子邮件、在线购物、视频通话、手机支付、手机远程控制等功能。手机已经成为便携式桌面电脑的“迷你版”，同时也为艺术家提供了新的艺术创作平台。例如，2015年在拉斯维加斯举办的 CES（国际消费电子展）上的“看不见的展览馆”，只需在平板或手机上安装此程序的 App，然后把设备对准作品，设备上就会立即出现三维的建筑、花草树木，而展台上的作品本身仍只是一个长方形或正方形的物体；当你把设备对准《蒙娜丽莎》，她会回以微笑，并且画中的花朵会悄悄绽放，把这幅著名画作变成了一个动画。这种新的感应器的使用，结合了装置艺术和现代科技，使得展览作品能够在时间和空间上延伸。

移动互动媒体艺术也在我们周围发生，它不仅实现了与外部环境的相互影响，而且在艺术家与不同领域技术人员的合作下，作为对作品进行实时远程控制和交互的媒介。例如在2014年10月，一个年轻艺术家团队合作的实时互动作品“跳跃”，这个作品利用了当下流行的“弹幕”元素，观众在现场听音乐会的同时，可以使用手机 APP 参与到作品中。观众在手机上输入的文本信息，会实时滚动显示在舞台背景的屏幕上。利用移动媒体，每个观众都可以通过自己的手机成为移动互动媒体艺术项目的参与者和体验者。随着科技的发展，手机等电子移动终端将为艺术家提供更多的创作方式和可能性。

#### 1.4 交互媒体艺术的互动性与游戏艺术的关系

在人们的成长经历中，游戏总是代表着乐趣和无忧无虑。在现代生活中，适度的游戏还可以用来减压和舒缓情绪。游戏的存在不仅仅是为了让人们感到“有趣”，还可以在互动过程中体验到不同的乐趣和内心的情感波动。现在，网络游戏和移动游戏的普及表明，游戏的互动性是游戏的一个显著特征，同时也是在线视频游戏吸引人的地方，因为玩家在游戏中可以与来自世界各地的不同人互动。游戏作为一种互动媒介，将玩家与现实世界紧密联系在一起。游戏的互动性和社交性是这类游戏不可或缺的部分。游戏软件和计算机硬件构成了一个庞大的互动媒体平台，玩家在游戏操作影响着游戏世界的走向。在游戏中，玩家可以与不同的人接触，现实中可能永远不会有交集的两种类型的人在这里可能会相遇。这种互动性使每个人都变得平等，观众体验到一种暂时的、只存在于虚拟世界的新型沟通和交流方式。在线视频游戏的魅力始终在于它能够让玩家之间无障碍地连接、相互交流。这种前所未有的人与人之间的关系不受身份、地位、教育、背景等现实因素的限制，彼此之间的互动经验超越了游戏的边界，也超越了人类沟通的边界。

与传统视频游戏相比，网络游戏具有高度的互动特性，实现了“人-机-人”互动通信模式。这与传统单机版游戏的僵硬和单调不同，网络游戏更加注重玩家在游戏过程中的沟通和合作。玩家可以根据游戏设置随意改变自己在游戏中的体型、身份等，弥补现实中的不足，完成现实世界中无法实现的愿望。在游戏中与其他玩家“PK”或攻击地下城 Boss，可以宣泄玩家在现实中承载的压力和不满。《你画我猜》是一款流行的休闲互动游戏，可以在台式电脑上或在手机或平板电脑上进行。

来自世界各地不同身份和背景的人们进入同一个游戏房间，操作触屏设备进行绘画，玩家可以在空闲时间通过平台上的交流来绘画和分享个人审美需求，随时进行游戏和互动。《你画我猜》的社交属性是其独特属性之一，在游戏中，人与人之间的关系模式发生了变化，不再是单一的游戏对抗，游戏的走向是基于不同玩家的操作而变化的。玩家既可以像艺术家一样在这个平台上绘画，让其他玩家猜测你的画作，也可以在游戏过程中与其他玩家交流，并相互呼唤出对方的想法和个人游戏体验。

与像“你画我猜”这样的休闲游戏不同，视频游戏的互动性质是另一种表达形式。视频游戏在绘画、雕塑、建筑、音乐、文学、舞蹈、戏剧和电影这八大艺术形式之后，被认为是“第九艺术”。高德说：“每一种艺术的最高使命是通过幻觉实现对更高现实的幻想。游戏在某种程度上与小说、电影和喜剧一样，本质上是一种‘虚拟现实’，这些艺术作品都是创作者在大脑中的想象、幻想和创意的产物，这些狂野的想象是创作者对现实生活的感性理解，某种程度上反映了其精神的吸引力，揭示了创作者内心世界的真相。”

冯梦波是一位当代艺术家，他在作品中反映了自己内心虚拟现实的真实性，有些评论家这样说：冯梦波是最有能力在游戏中进行艺术创作的艺术家，同时也是最能在创作中玩游戏的人，他兼具二者，并有效地将两者融合在一起，作品富有艺术性。冯梦波自 20 世纪 90 年代中期就开始了他的“玩耍”，他的早期作品将经典的政治图像如《红灯记》和《红色娘子军》处理成游戏角色，在其中，角色扔出的武器是可乐罐，带有娱乐性的政治精神。在韩雅轩 1994 年的展览《游戏：长征》中，有两张描绘小红军和春丽在游戏决斗的图像。春丽作为卡普空公司制作的视频游戏《街头霸王》的女主角，拥有男性无法比拟的战斗能力。

1997 年，冯梦波制作了一部互动媒体作品《智取威虎山》，这是著名游戏《毁灭战士》和样板戏《智取威虎山》的混搭。横眉冷对的英雄人物杨子荣成为游戏的主角，观众可以选择自己的虚拟角色扮演英雄或反派。2002 年，在德国卡塞尔的文献展上展出了互联网项目 Q4U（图 2-10，图 2-11），在该项目中冯梦波用一幅卡通化的自我形象替换了一个名为“Quake”的游戏角色，一手持摄像机，另一手拿着血色步枪。“在迷宫般的城堡里，与敌对

近战的形象完全相同，震耳欲聋的爆炸声和激烈的 TECHNO 音乐，耀眼的火光和血雾让整个大厅明亮如白昼，剧情和故事不多，只有枪和子弹，伤亡的枕头……通过让玩家不是用常规的键盘、鼠标或游戏操纵杆，而是通过在跳舞板上灵巧的步伐参与游戏，艺术家不仅改变了游戏的互动模式，也激发了参与者从自身角度深入思考现实社会中的暴力问题和人性的阴暗面。”

## 1.5 互动媒体艺术中的虚拟现实体验

虚拟现实在英文中被翻译为“Virtual Reality”，也被称为精神领域和幻象，它本身的含义是用来表达一些现实事物的属性，这些属性是虚幻的。虚拟现实是通过现代科学计算机技术来生成一个三维的、立体的、可感知的非真实世界，并为观众提供视觉、听觉和触觉等感官模拟，使其能够产生一种沉浸感。

吴博反这样解释虚拟现实：“人们用手指控制计算机，这等同于使用数字来编造一个虚拟世界。现在，这个‘手指’就像是造物主对编造的虚拟世界的手，成为了‘创造’的手指。

‘0 和 1’变成了搭建基础的‘原材料’，其中数字本身在现实生活中的本质意义有所不同，它们可以创造出一种虚幻的、人工的‘现实’，一个栩栩如生的‘现实’。”虚拟现实不是纯粹的技术，而是一种艺术表达，允许用户使用技术来表达“不可能完成”的事物。用户在虚构世界中的沉浸式体验与现实世界无异，虚拟现实体验是实现“不可能”的幻想的一种方式，比如新媒体艺术家克里斯塔·索默雷（Christa Sommerer）在 1997 年创造了“生命空间”（Life Spacies），随后与劳伦特·米尼欧（Laurent Mignonneau）在 1999 年创造了《生命空间 II》（Life Spacies II），这系列作品为观众提供了生动的互动体验，通过计算机控制的人机界面，观众可以通过触摸屏输入一段文本或绘制一个图形，输入的信息将被转换成特定的语言符号，通过“形态转换器”将观众输入的文本转换为具有实体、互动和自主行为的有机体，观众还可以通过触摸不断影响它们的行为变化和生存状况，并与它们进行互动。

从这个例子中可以看出，在互动媒体艺术作品中，观众扮演着“参与者”和“消费者”的角色，而艺术家则是“生产者”。一个互动媒体艺术作品中最重要的一环是“参与者”的整个介入过程，参与过程中的每一个参与者因其各自的审美需求、审美品味不同，使得参与作品时有不同的参与体验和感受，也因此会写下不同的艺术评论，产生不同的化学反应，这是互动媒体艺术与传统艺术最明显的不同之处，也是其自身魅力的源泉。虚拟现实技术的发展便利了“生产者”和“消费者”之间的实时互动，艺术家杰弗里·肖（Jeffery Shaw）在德国卡尔斯鲁厄的媒体艺术中心在 1988 年至 1991 年间设置的互动装置“可读城市”（The

Legible City) 结合了虚拟现实技术, 观众需要像平常一样骑在屏幕前的自行车上, 当车轮旋转时, 自行车会持续将不同的速率和转弯数据传送给传感器, 电子屏幕上显示的变化三维文字被编辑成城市中的建筑物, 观众骑行通过它, 仿佛穿行在真实的城市街道中。虚拟技术的研发和进步允许艺术家将自己脑海中的幻想变为现实, 并邀请观众参与到他的幻想世界中, 得益于精湛的数字技术, 人们可以在虚拟场景中旅行, 时间和空间的界限在虚拟现实面前轻易被打破。

## **2.融合 —— “生产者”与“消费者”之间的角色转换**

### **2.1 “旁观者”到“参与者”——角色反转与双重身份**

角色替换思维是互动媒体艺术的典型设计思维, 它与传统艺术在本质上有不同。在传统艺术中, 观众是单方面接收艺术家通过作品传递的信息, 他们是“旁观者”的身份, 不需要反馈和反应。而互动媒体艺术的目的是将观众从“旁观者”的位置转变为“参与者”的位置, 观众在接收互动媒体艺术作品的信息的同时, 也参与到创作过程中, 从而改变了作品的表现方式和互动体验的不同。

在过去的传统艺术中, 艺术作品通常由艺术家一方完成, 艺术家是创作的主体, 观众则是被动地接受作品、欣赏作品, 观众的体验和想法不能影响、改变作品本身, 作品始终是外在的、固有的, 虽然“一千个读者就有一千个哈姆雷特”, 不同的观众因为不同的审美需求、不同的欣赏角度, 所以每个人对一个作品都有不同的理解和感知, 但这些只存在于观众内心世界中, 观众始终只是一个“旁观者”, 看着艺术家的创作作品, 心中充满自我的审美呼声无法实时互动和影响作品。在互动媒体艺术作品中, 观众从被动变主动, 从“旁观者”转变为“参与者”, 角色的转变是互动媒体艺术作品互动特性最直接的体现。

在互动媒体艺术作品中, 观众不仅仅是从“旁观者”转变为“参与者”的位置转换, 而且还拥有了“创作者”的双重身份。在这里, 艺术家和观众之间的互动不仅是可能的, 而且是必要的。在作品中, 观众占据了创作主体的位置, 主动参与互动, 影响作品的变化。此时, 它形成了对作者权威的转移, 使得公众直接进入艺术作品的体验中, 观众成为了作品不可或

缺的一部分，允许艺术家和观众进行自然对话和沟通，并共同完成作品的创作。2010年世界博览会德国馆的一个互动装置作品“力量之源”就是一个例子。这个装置本身由40万个发光二极管组成的金属互动球体制成，通过观众的呼喊来控制金属球的运动，观众的呼喊越大声，金属球的摆动就越剧烈，在现场的左右两边，哪边的观众呼喊高，金属球就会摆向哪边，同时摆动中，金属球表面会显现出和谐城市的图像，声音的变化直接影响球体的颜色和表面城市图案的变化。官方对这个装置作品的意义做出了解释：“希望通过这个互动环节，表达‘目标一致性’的原则，因为城市的和谐需要每个人共同参与和干预。”

“力量之源”这个互动装置利用技术手段将欲传达的概念赋予作品，使观众得以参与进来。在与互动装置进行沟通的同时，互动装置作品也就得以完成，同时观众也体验到了作品本身想要传达的概念，并结合个人的审美需求体验到不同的愉悦以及作品引发的思考。充分展示了观众从“观众”到“参与者”以及“创作者”双重身份转变的另一个互动装置叫做“城市互动”（图3-3）。在这个作品中，年轻的奥地利艺术家使用互动媒体技术将一些“不可能”的艺术转变为互动媒体艺术作品，通过计算机技术将两个城市连接起来。当一个城市的互动装置上的观众在木板上来回移动，产生互动时，这时木板连接的计算机会将数据传输到另一个城市的同样装置上，这时另一个城市的装置会发生跳动。这个作品只能通过观众的参与，观众与装置之间的互动来完成。观众的身份从传统的“旁观者”完全转变为作品的“参与者”，观众不再是被动的“观看者”，而是参与到作品的创造和互动中的“创作者”。

## 2.2 “生产者”与“消费者”相互交织和依赖

互动媒体艺术的重要性在于现场的存在，艺术家将他们的艺术概念与科学技术结合起来，并与现场观众一起进行创造性旅程，共同完成互动艺术作品。互动媒体艺术具有很强的公众参与性，对参与者没有限制和要求。在互动媒体艺术作品中，观众既是艺术作品的“生产者”也是“消费者”。在观看作品时，观众是接受艺术家展示作品的“消费者”；而在参与作品互动时，观众则变成了作品的“生产者”，在作品的“生产”过程中与艺术家进行交流对话，同时产生情感共鸣。

《云》是加拿大卡尔加里艺术家 Caitlind Brown 设计的一个原比例互动灯光装置。艺术家让参与者站在装置旁边，可以选择取下或安装灯泡，这种互动行为创造了一个巨大的、可触摸的、闪烁的灯光云。该装置由 Caitlind Brown 使用超过 6,000 个可以点亮和熄灭的灯泡，以及钢材和金属拉绳制作而成。这个设计重新想象了废弃材料，从不同的艺术角度处理过剩的材料。

在传统的艺术展览与欣赏活动中，艺术家与观众是相对独立的个体，观众对作品产生的所有感受与审美情趣只在自己心中，双方因作品引起的内在共鸣缺乏沟通，这种审美活动只能通过作品的描述或作品的中介传达给观众，观众只能选择“接受”或“不接受”，存在局限性。但在互动媒体艺术中，这种单向的“孤独欣赏”感已不复存在，观众可以自由发挥想象，进而参与到作品中。在这里，当观众与作品以及艺术家产生“情感共鸣”时，艺术家与观众之间的关系变得越来越亲密，双方在共同成为艺术作品的“生产者”的同时，也是作品的“消费者”，相互依赖和不可分割，任何一方的缺席或不存在都会导致作品完成度的缺失。在互动媒体艺术中，由于身份的转变，观众与艺术家之间不可分割且相互依存的关系，可以获得不同的审美感受，这是个人体验与审美需求与作品本身最本质的融合，而整合的产物是以个体为主体创造力的艺术审美呼吁。这种互动创作的主体性在互动媒体艺术的表达中变得越来越明显。

## 2.3 干预与整合：互动媒体艺术应当“以人为本”

互动媒体艺术的魅力始终在于其“以人为本”的参与性互动特征，它打破了传统艺术形态和观众的单向沟通方式，在与观众互动时，有力地传达了作品本身所代表的实践意义的艺术理念，艺术家与观众共同成为创作的主体，而作品的对象也是“人”。互动媒体艺术在一定程度上依赖于科学技术的艺术，这在今天似乎是互动媒体艺术的一个限制，它自身的发展必须是“以人为本”的，首先要满足大多数观众的需求，科技的使用不应让互动媒体艺术变成普及艺术的高墙，而应该因为科技的使用让艺术家的理念更加清晰简单地表达给观众，不再晦涩，不再“高高在上”，让科技服务于互动媒体艺术，成为“为90%的人创造的艺术”。

从未来主义和立体主义的出现，到录像艺术和装置艺术的逐渐成熟，再到复制媒介、多媒体到网络科技的迅猛发展，这一切不断推动互动媒体艺术的发展，使得观众的角色从“旁观者”转变为“参与者”。这一角色的替换，实现了传统艺术作品对观众的单向信息传递向公众在作品创造中的积极参与的转变。互动媒体艺术通过邀请观众参与作品的互动，主动影响观众的思想 and 行为，并努力消除观众与观众、观众与艺术家、观众与艺术作品之间的距离。通过独特的方式替换观众和艺术家的角色，并产生一个新的身份——“生产者”和“消费者”。这一段落不仅阐释了互动媒体艺术的参与性互动是如何表达的，还说明了互动媒体艺术是如何连接、沟通并转变作品与观众之间的关系的。通过这种转变，互动媒体艺术不仅扩展了艺术的边界，而且增强了艺术的社会性和参与性，这使得艺术不再是高高在上的象牙塔，而是变成了公众可以触及并影响的平台。互动媒体艺术的这种开放性和包容性，使得每一个

参与者都有机会通过自己的行为 and 选择来影响艺术作品的最终形态, 从而在艺术的创作和体验过程中实现了双向沟通和共同创造。

互动媒体艺术始终建立在科学和技术的基础上, 它利用媒介传达创作者的真实观念和情感。互动媒体艺术从“介入”人们的日常生活, 到积极影响观众的思想, 有效地将观众与互动媒体艺术的关系“整合”在一起, 不断探索如何更好地利用科技服务于人民, “以人为本”, 通过互动媒体艺术的手段塑造并改变世界, 使艺术与时俱进, 借助科技让更多人接受和欣赏, 更好地发展“中国化”的互动媒体艺术。

# 第3节 体验与互动：公共空间多媒体互动装置艺术介入研究

## 1. 多媒体互动装置

### 1.1 多媒体互动装置的概念

#### 1.1.1 对多媒体互动装置的基本理解

装置艺术作为一种艺术形式,其存在具有一定的历史,关于它的概念在本文中不再赘述。简单地说,装置艺术是场所、材料和情感综合展示的艺术。

在本文中讨论的多媒体互动装置艺术中,“装置”主要指的是与多媒体相关的信息输入、输出的硬件和软件,以及一些辅助的综合材料,比如操纵杆、各类传感器、扬声器、屏幕、投影仪等。在这其中,多媒体技术是核心和艺术表现的来源,它的创作理念和目标来源于多媒体技术展现的新表达方式以及它带来的新互动体验,而综合材料只是作为辅助表现手段。与多媒体互动装置不同,传统的装置艺术主要使用一些物理材料来展示其物质本身的特点。

#### 1.1.2 多媒体互动装置中的核心技术

要理解多媒体互动装置的概念,我们首先需要了解一些与多媒体互动装置艺术相关的术语,以及它们与互动媒体装置艺术的关系。例如,经常可以看到的新媒体、数字媒体等概念。新媒体是一个相对概念,在当前世界对新媒体的定义并不统一,我们可以暂时将由现代技术推动的新的信息载体定义为新媒体。数字媒体技术包含在多媒体艺术中。多媒体指的是单一媒介的复合形式,而多媒体则更为多样化。从单一媒介到多媒体,再从多媒体到互动多媒体,媒介与人沟通的方式不断变化。本文涉及的互动装置艺术中的多媒体技术主要是指数字化媒介技术。

原始的多媒体互动技术仅仅是原定的操作和输出,并不是真正意义上的互动。多媒体技术的应用范围限于广播、电视、广告、影视制作等领域。随着现代科学技术的发展,公共艺

术的表达也有了更广阔的发展空间,科技与公共艺术结合的多媒体互动装置艺术在艺术表达上丰富多彩,扩展了公共艺术的体验模式和表现空间。例如,将传统的静态视觉体验如浮雕、雕塑等转变为动态的图像、光影变化的视觉体验,并且还可以添加声音。同时,它也创新了公共艺术的互动形式,使得传统的心理上不可见的体验和互动变得显性化和具有娱乐性,使得作品更具吸引力和趣味性。例如,从传统的观赏性或可触摸的雕塑和浮雕变为红外感应和声音感应,从而使得装置作品随着公众参与的变化而相应变化。多媒体技术的介入最终使得公共艺术更容易被展示出来,并被更多的人看到和触摸。

### 1.1.3 多媒体互动装置中的互动

互动这一概念源自英文单词“INTERACTIVE”,有时也被称为交互。本文中提到的互动主要是指在公共空间中,由多媒体互动设备产生的互动。目前,随着互动艺术作品的快速发展,艺术设计中的“互动”一词有了相对成熟的定义。在《互动设计基础实验》一书中定义互动为:“互动的概念是我们通过设计互动作品来支持人们的日常生活和工作,尤其是通过创造用户体验,增强和拓展人们的生活方式、沟通和互动。”“互动设计决定了通信对用户的价值以及他们经历的感觉的质量。”作品的互动设计更多的是一种感知,而不仅仅是一个物理的对象,比一个具象的实体。互动设计处理的是无形的过程,并且适应个人的需求和偏好,这是一种全新的设计类型。互动也经常被用在人们的日常生活中,例如人与人之间的互动。

本节重点讨论多媒体互动装置与公共空间中公众之间的互动。本文将多媒体互动设备中的体验和互动分为显性和隐性两种类型。公共空间中多媒体互动装置的隐性互动要求作品触发公众的思考和心理交流,需要公众反复体验和参与其中,并从中获得自己的思考,但其互动形式不会改变作品的外在形态。而显性形式不仅会与公众产生心理互动,还会使公众通过积极参与和与装置艺术作品的互动,改变作品的形态。研究多媒体互动装置中的互动有助于为多媒体互动装置的设计概念提供思路。

## 1.2 多媒体互动装置艺术的特点

### 1.2.1 互动性

多媒体互动装置最显著的特点是它在特定的地点展出，并在观众的互动中呈现。互动本身具有许多特点，如不确定性、衍生性、多元性、偶然性，在多媒体互动装置中，需要观众的参与、体验和互动，不同观众的互动过程不尽相同，会产生各种心理感受。在观众互动过程中，艺术会产生无限的衍生物，比如观众与其他参与者或某些多媒体材料之间的互动，材料的连锁反应，衍生和发展新的体验元素；多媒体互动装置也因其科技属性具有多种交互方式。体验和互动的经验是不可控的，而技术带来的多样化互动方法进一步增加了多样化的展示结果。多媒体互动装置的每一次互动体验也是由并发性产生的，一个随机的观众被装置的某个方面吸引，进而产生了与多媒体互动装置的体验和互动。

### 1.2.2 体验性

与传统公共艺术传达给观众的孤独感相比，多媒体互动装置的互动性为观众带来了活力和亲和力，使得观众有足够的空间去感受自己。在传统的公共接触艺术过程中，观众必须忽视展示空间并专注于作品，相反，多媒体互动装置艺术包含了空间，空间和公众参与也是作品的一部分，这为观众提供了对其所处环境的整体观察、介入和反思的机会。多媒体技术更能让观众感受到身临其境的体验。在世博会期间，中国馆“东方足迹”展区的动态“清明上河图”就是一个例子，该作品基于北宋画家张择端的《清明上河图》，使用12台投影仪同时作业，将原本静态的画面动态化，连续播放。巨型投影屏幕长128米，高6.5米，几乎占据了展区的整个圆形墙面，除了投影部分外其余地方光线昏暗。这使得观众在观看时能够完全沉浸其中。

### 1.2.3 高科技

多媒体互动装置是技术与艺术的跨界融合。在信息技术发展的背景下，科技与艺术的结合使得艺术不再神圣不可侵犯，公众能够更容易地进入公共艺术领域参与体验和互动。多媒体互动装置艺术已逐渐成为公共艺术前沿的一个分支。科技的发展拓宽了公共艺术的范畴，

计算机技术与数字媒体技术使得当前的科技成为一种新的审美可能性，它带来的数字体验正在共同的空间中重新演绎了一种新的审美。例如，在旧金山市中心的街道上，就有结合多媒体互动的装置艺术作品，其中一个作品是当两个人分别坐在长凳的两端时，通过握手连接长凳两端的扶手，长凳中设置的电路将会连接起来，美妙的音乐将通过扬声器传播。这是通过人体触摸长凳上的金属片，利用人体可以导电的原理，结合装置中的科技设计完成的。设计不仅让观众体验到科技的乐趣，也使人與人之间更加亲近，并蕴含了城市文化的内涵。

#### **1.2.4 实时性**

与人们对传统艺术作品的反思相比，多媒体互动装置具有实时特性，人们可以通过接触装置或通过装置的传感器，快速地得到及时的信息反馈。观众能够获得在普通生活中难以获得的心理愉悦。从另一个角度看，多媒体互动装置的展示过程和结果是通过时间的累积完成的。它打破了时间的界限，使作品具有了延时性。多媒体互动装置通过在较长的时间内拉伸其维度来展示其完整性。观众也能够在整个延时期随时获得实时反馈。艺术家 Daniel Rozin 创造了一种与人实时互动的多媒体互动装置，就像一个机械化的镜子，一个挂在墙上的等边三角形，配有可以旋转的黄色线条，通过角度产生图像。装置配备了隐藏摄像机，当人在 8 英寸的距离范围内时，装置将随着人的动作产生不同的图案，人的身形将以艺术剪影的方式显示出来。

#### **1.2.5 选择与控制**

选择与控制权是指观众在参与多媒体互动装置交互过程中的选择权，自行选择控制对象，这也是创作者和多媒体技术以及计算机系统给予的选择结果。观众拥有控制作品体验结果的权力，这来源于观众选择的结果和心理控制感受。与传统公共艺术的被动感受和体验相比，选择与控制权可以给观众带来极大的心理满足感。这种权力强烈地加深了他们对艺术作品的记忆，并提高了他们对公共艺术的好感，进一步缩小了公众与公共艺术之间的距离，深入到公共生活中。观众的选择不仅影响了他们自己的体验，而且可能影响到后续观众的体验，因为多媒体互动装置常常会记录这些互动，并将其作为未来互动的一部分。这种参与和控制的机制，使得每个观众都成为了艺术创作的一部分，这不仅改变了艺术品的含义，也改变了观众与艺术品之间的关系。

### 1.2.6 灵活性

多媒体互动装置首先可以将观众置于三维空间环境中，它具有开放多样的创作条件，通常来说，它只提供短期展览，并不具备收藏功能，也不适合长期展出。与其他公共艺术作品相比，多媒体互动装置具有较高的灵活性，它的许多方面都可以临时改变，创作者可以在展览期间更改其组合方式，也可以根据时间和场馆因素进行重新组合，甚至可以根据观众的本质进行调整。

### 1.2.7 娱乐性

娱乐性也是多媒体互动装置的一个主要特征和新颖之处。所提到的多媒体互动装置中的交互，不仅仅是一般传统定义下的交互，它包含了计算机计算和数字媒体技术体验的融合。互动装置中的娱乐性质源自于视频游戏的参与性质。亚里士多德认为，游戏可以作为工作的调剂品，因为它能给灵魂以休息，然而它似乎包含了生活的所有乐趣。这种快乐是人类缺乏的更为高雅的事物的终极追求。技能、力量和韧性的竞争在历史上每一种文化中都扮演着重要角色，比如当前的奥林匹克运动会，无论是参与仪式还是纯粹为了乐趣和节日。而这种竞争是游戏无区别追求的。“游戏的形式给社会生活以超越和生物本能的形式，这强调了游戏的价值。正是通过游戏，人类社会表达了他们对生活和世界的解释，在这方面，如果说游戏已经演变为文化，不如说在文化的最早阶段，就有游戏的特征，文化在游戏的氛围和形式中前进。”<sup>3</sup> 像游戏一样，多媒体互动装置作品需要观众的介入来完成，观众可以通过互动装置艺术品中的娱乐获得身心的放松和锻炼，以获得最终的心理体验。多媒体互动装置的娱乐性也可以让公众积极参与，从互动过程中获得体验和互动的乐趣，亲近艺术，并接受文化内涵，而不是在传统艺术环境中被动填鸭。例如，商场中的投影足球游戏吸引人们积极参与互动。

### 1.2.8 空间性

组成传统艺术的材料由特定的物质定义：画布、石材或胶片。构成多媒体互动装置的主要“材料”是空间和多媒体技术。这并不意味着装置艺术是无形的。由于其空间和技术的本质，成为多媒体装置的材料通常的定义是它存在于公共空间中。“装置将空气中、中性的、

公共的空间转变为个人的作品,并且它邀请观众体验这些公共空间作为作品的整体空间。”<sup>4</sup> 这意味着艺术装置实际上是一种空间展示。在2014年活力悉尼光与音乐节上,艺术家们在广场地面上布满了装有强力敏感反应器的光箱。观众可以在这个空间内随意奔跑和踩踏,从一个光箱跳到另一个光箱,随着人站在其上,光的颜色将会有不同的变化。这个多媒体互动装置通过照明的视觉体验和公众的互动,创造了一个充满欢乐的空间和场地氛围。

## 2. 多媒体互动装置艺术的设计理念和应用意义

### 2.1 多媒体互动装置的互动设计理念

多媒体互动装置拥有三种最基本的体验方式,视觉体验、听觉体验和触觉体验,它们可以通过光影、形态和声音等多种形式来表达。从技术角度来看,光影具有明暗和色彩图案或光的变化,形态具有变化和重组,声音具有音乐和立体声,声音强度的变化等等。基于这些体验方法,依靠现代多媒体技术的表现技巧,可以创作出具有感染力的多媒体互动装置作品。

在多媒体互动装置的创作和构思中,需要考虑所使用体验模式的载体,是突出视觉光效,还是特定使用声音表现,或者强调形态或综合使用,作品应以能够感染公众情感为原则,装置本身应该有其自身的一套技术理论体系,更需要有一个带有体验性和互动性的互动设计理念,通过互动技术作为媒介,设计表现形式和体验来获得观众的互动和响应,实现人与多媒体互动装置及环境之间的互动。应该考虑到公共艺术的艺术功能的延伸。这种“艺术—人—新艺术”的双向互动形式的创新性和不确定性极大地激发了公众的参与和热情。艺术家和观众的身份已经模糊并转移,公众已从单纯的欣赏者变为艺术创造者。“<sup>5</sup> 群众是动力和环境的主体,而在公众参与中的互动和体验也是需要考虑的关键点。”

相比之下,传统装置艺术通常使用熟悉的物质材料来创作,而多媒体互动装置则是将传统材料与多媒体媒介相结合。它们都在公共空间中具有装饰性和参与性空间。在传统装置艺术中,人们通常是被动接受装置艺术。多媒体互动装置艺术的吸引力来源于人们的心理习惯和好奇心。多媒体互动装置主要利用现实主义、尺度、颜色、传统材料和多媒体设备,将公共艺术变成围绕人们的互动艺术形式,从而具有亲和力,互动装置的主题也是平易近人的。它不仅可以传递参与的乐趣感,还可以通过多媒体设备传达更多的信息。

多媒体互动装置的设计需要清楚地认识到目标吸引的受众及其影响范围。在特定的主题和环境中,通常更容易吸引儿童和年轻人,并使他们无私地参与到富有游戏性的公共艺术中去。卡尔·约翰·埃克罗斯在公园中的多媒体互动装置,一端设有麦克风,另一端是听筒,

顶部有按钮，通过这些按钮可以听到不同的声音效果，声音与互动装置相互作用。这个多媒体互动装置的观众主要是公园里玩耍的儿童。

## 2.2 多媒体互动装置中的互动设计

### 2.2.1 公共空间中多媒体互动设备的隐性互动

“隐性互动”基于对隐性知识的延伸。基于身体的隐性知识是主体在认知过程中生成和应用的隐性知识。这种隐性知识很大程度上植根于人类身体功能的使用，或者工具的使用。例如，当艺术家创作时，这些艺术技巧中存在着大量的隐藏元素。这种类型的隐性知识是具有垄断性的，隐性知识被限制进入公共交流领域或被共享。这种隐性知识的显性化需要建立信任机制，可以通过内部思考和外部交流获得。此外，它通常还需要反复的动手实践和体验。公共空间中多媒体互动装置的隐性互动要求作品触发公众的思考和心理交流，要求公众反复体验和参与并从中获得思考。

多媒体技术通过视频、声音等装置在公共空间中的参与，结合了技术与各种常规材料，承载了一些信息元素，在信息传递或装饰环境中发挥作用，这种互动装置的艺术形式具有隐性的特点，即公众与作品产生共鸣，但互动形式并不改变作品的外在形态。

例如，在上海购物商场的情人节装置艺术中，装置材料通过LED灯和亚克力板制作，并且形状是一个倒置的酒杯，人们可以停下来观看或走进放大的酒杯内部，感受酒杯中的光影世界，而酒杯的内部空间放置了一个视频装置，以宣传其赞助的红酒品牌主题，起到商业宣传的作用。在外观上，通过独特的塑形艺术设计可以增强购物中心的文化氛围，让人们能体验和参与其中，并通过观众主动按动开关来了解装置想要传达的信息，观众在欣赏艺术作品的同时也提升了审美情趣。这种涉及多媒体技术的互动装置艺术作品不仅具有商业与审美性，还具有娱乐性和互动性，并通过放置视频装置准确传达了它想要表达的信息。在顺应公众心理兴趣的同时，它强调了公众的主体性。人们可以选择观看视频也可以不观看，观看与否则不影响他们在装置中的体验。人们的参与和互动在这种情境下，装置的外观和表达没有变化和受到影响，但公众可以通过感受互动装置的氛围接近体验空间，反映了在公共空间中无形体验和互动性的应用。

### 2.2.2 公共空间中多媒体互动设备的显性互动

公共空间中多媒体互动设备的显性互动，通常包括图像和视频的视觉感受、使用音频的听觉感受、通过传感器实现的触觉感受等，或者综合运用多种方式进行表达。显性形式不仅会与公众产生心理互动，而且实现公众主动参与并与作品的安装互动以改变作品的形态。这使得互动装置艺术作品打破了时间的界限，成为一种动态的艺术作品，其展示在观众的动态参与和互动中完成。装置成为一个过程，而不是一个静态的艺术作品。装置作品不是它自己，而是它与观众的互动共同构成了作品的完整性。

Travesias De Luz 的“漂浮灯光”装置是一种基于光和色彩的互动装置，将光的概念转变为城市空间中所有年龄段的公众都可以参与的游戏。该装置使用了一个 10\*3cm 的低分辨率屏幕，由 100 个圆形彩色灯泡组成，每个灯泡中央都配有一个转换开关，访客可以触摸它们，可以点亮或熄灭它们，并且还可以在设备上留下创意性的信息和图案。

## 2.3 多媒体互动装置中的体验设计

传统的艺术展览向每位观众传达一种孤独感，让他们单独面对和沉思艺术作品。这种传统的艺术作品展览使得观众不得不忽略整体展览空间，包括他们自己在这个空间中的位置。相比之下，由于多媒体互动装置拥有空间的完整性，多媒体互动装置作品能够建立起一个观众群体。多媒体互动装置的观众绝对不是孤影单行，而是一个观众群体，像这样的艺术空间能够被大量观众观看和互动。这些参与的群众是多媒体互动装置的一部分，并通过参与完成展览的完整性。这也是结合空间设计并具有体验性和互动性的多媒体互动装置艺术与传统艺术展览之间的本质区别。它甚至可以创造出一个不需要任何共同过去的大众文化群体，一种新型的无条件群体，不同于宗教和政治、部落文化群体，展现出现代化多媒体互动装置艺术的巨大潜力。

这种互动装置与传统展览和公众的现场观看体验不同，因为后者相比之下过于有针对性，通常情况下观众无法充分反思自己，他们无法对自己所处的环境和空间进行反思。另一方面，装置艺术在一定程度上帮助这些群体观察和思考自己，给予他们自我展示的空间。换句话说，它给予他们足够的空间和环境，感受他们所处的空间和群体以及自身。

纽约 2015 年在时代广场的互动艺术装置“心跳 (Heart Beat)”来自布鲁克林设计工作室 Stereotank，是由环保材料制成，并在心形雕塑的顶部设置了不同音质的鼓面，使得每个鼓都有自己独特的声音。观众可以使用鼓槌敲打它，当鼓声响起时，设备中的呼吸灯会闪烁，

旁边的管道也可以直观地看到变化并听到声音。设计师鼓励游客四处移动,并希望每个人都敲打不同的音乐节奏,感受在声音振动和光影变化中独特的“爱之协奏”和“心跳”的感觉。在这件公共艺术作品中,公众不仅是独自欣赏艺术作品,而是一起参与其中,并通过艺术作品创造新的创作。

## 2.4 多媒体互动装置的设计用途与意义

### 2.4.1 设计目的

多媒体互动装置的设计目的包括以下几点:

首先,它丰富了艺术语言的表现性,增加了信息的传递,创作者通过多媒体技术这一媒介来表达作品。这使得观众更容易获取信息并接受画面。

其次,为观众提供全新的视觉体验。创作者将新的数字技术与材料相结合应用于装置作品中,使得观众能够获得难以从传统装置艺术中获得的技术与艺术交叉融合的丰富体验。

第三,扩展内容表达的空间,增强作品的空间维度。多媒体技术以其独特的功能扩展了艺术作品的空间维度,使空间成为材料的一部分,并全方位感染观众的感官情感。

第四,提高艺术创作效率。多媒体技术的使用和计算机软件操作的速度减少了艺术创作的时间。与需要打磨加工等待成型的传统艺术创作如雕塑相比,多媒体技术对图像和空间的塑形更为便捷和直观。

第五,增加公众的参与热情。通过引入视频、光影、声音等数字媒体技术,极大地满足了公众的游戏心理,让人们在光影交错的时空之间流连忘返,打破时间的界限,使人成为多媒体互动装置的一部分。

### 2.4.2 设计意义

在一定程度上,艺术作品反映了人们的生活方式和生活环境,并标志着人们社会生活现代化的过程,因此艺术的科技发展对人们生活的现代化有着积极的意义和推动作用。多媒体互动装置中艺术的技术化是这一点的良好证明。

多媒体互动装置使接触艺术成为一种生活方式,并且这些装置被放置在公众可以到达的地方,这对当代艺术家和公众都非常有益,因为它能在更大范围内传播艺术与文化。公众通

过参与和体验多媒体互动装置，将会喜欢上这种艺术形式，进而理解公共艺术，参与公共艺术，提升公众审美，传播艺术文化，形成良性循环。多媒体互动装置通过增强公共艺术的艺术吸引力，更能够承载和大力传播公共艺术的时代社会价值和责任。多媒体互动设备超越了传统艺术作品单一信息传递的局限。

当代城市的技术发展和消费形态减少了人与人之间社交接触和个体对话的机会，这导致了我们的视觉主导了其他感官，将公众转变为城市中纯粹的观众。当公众面对一个地方时，味觉、视觉、嗅觉、听觉和触觉的指令、交流和处理需要达到平衡。涉及多媒体技术的互动装置艺术作品使公众能够参与其中，去体验和互动，激发公众的全感官体验，这样具有多样化体验方法的艺术作品平衡了公众的各种感官。在多媒体互动装置艺术中，众所周知的元素不是直接挪用，而是在空间布局和材料方面进行了进一步的处理和整合。

这种独特的艺术形式在增强艺术作品的亲和力的同时，也增添了趣味性，与观众的互动激发了他们的兴趣和创造力。这使得多媒体互动装置不仅用于传播文化和乐趣，还用于进一步反思装置艺术作品的效果。将之前的单一的单向审美提升到一个新的水平。

互动装置艺术让我们看到了科学与艺术感性和理性的结合，多媒体介入的互动装置不仅打破了传统，为互动装置提供了全新的互动方式，而且打破了时间的界限，丰富了互动的形式和空间，扩展了公共艺术的范围，其延伸性和实时互动性，突破了时间的限制，使公众能够在公共艺术中感受到体验和互动的可能。

### **3. 多媒体互动装置艺术体验性和互动性的实现**

#### **3.1 多媒体技术的使用**

多媒体技术是多媒体互动装置艺术核心与艺术表达的源泉，其创作理念和目的源于多媒体技术能够展示新的表达方式和带来新的互动体验，综合性的材料只是辅助性的表现手段。利用多媒体技术进行艺术创作有利于把握社会时尚话题和文化热点，快速有效地进行艺术表达和传播，其艺术符号非常容易与流行文化融合，形成具有先锋性和前卫性质的独特艺术风格。多媒体技术具有信息载体数字化、实时性、虚拟性、多样性和互动性的特点。

多媒体互动设备中较为常见的技术包括与多媒体相关的信息输入、输出硬件软件和一些辅助性的综合材料，例如操纵杆、力敏传感器、红外传感器、光敏传感器和声敏传感器、音频输入输出设备、屏幕、投影仪等。多媒体互动设备的整体架构通常包括信息检测和采集设备、计算机处理设备、传输系统以及信息输出设备。以作者的毕业设计作品实践为例，信息

检测和采集设备是红外感应投影仪和自行车踏板，当检测到有人踩踏踏板的信息时，信息从动能转化为电能经过传输系统传输到投影仪，投影仪接收信号输出带有图像的照明信息。投影仪的红外感应系统感知到有人踩踏在投影屏幕上时，信息被传输到计算机处理系统，然后投影仪会根据信息旋转方向，将屏幕投射到附近没有人踩踏的空白区域。

## 3.2 多媒体互动装置的体验

本节将从技术视角研究多媒体互动装置的创作方法和体验方式。多媒体技术涉及互动设备常用的创作工具，如图像、声音、灯光、网络等。多媒体互动设备具有综合性的交叉特点，这要求有多媒体技术支持，通常包括操纵杆、声音输入设备、传感器、屏幕等。完成此类多媒体互动装置作品，要求创作者既要有使用多媒体技术的能力，也要具备艺术创造力和美学素养，因此这类艺术作品对创作者提出了更高的要求，需要跨学科和综合能力。在创作期间，使用多媒体技术应该展现出创作者想要表达的艺术体验，并且需要合理结合软硬件来实现预期效果。

### 3.4.1 视觉体验

视觉体验是互动装置的主要体验形式，自传统艺术表现方式以来，视觉体验一直是表达的焦点。例如，传统雕塑、浮雕和传统装置艺术等，它们主要是通过视觉语言来体验艺术作品，面对公共艺术作品的展示时，观众通常会被强烈吸引。具有视觉特征的公共艺术作品能够在第一时间吸引观众的注意。多媒体技术对互动装置的介入带来的视觉语言是多样化和变化多端的。多媒体为互动装置带来了选择性，以及信息传输的多样化，如通过LED屏幕或光的变化来表达图片和图形。例如，在购物中心或机场常能看到的LED游戏屏幕，观众除了可以欣赏到装置的外观，还可以随时通过感应参与游戏，以放松身心。

### 3.4.2 听觉体验

听觉体验是仅次于视觉体验的表现方式之一，其中音乐和音效是听觉体验模式的重要组成部分，在公共空间营造现场氛围中起着作用。传统的公共艺术通常难以结合听觉和视觉体验因素来体验，表达方式相对单一，而涉及多媒体技术的互动装置使得视觉和听觉元素的融合成为可能。声音元素的融入可以进一步激发观众的共鸣。例如，在2011年上海音乐学院

国际电子音乐周的“触声-触动声音”互动装置展览中，作品的声音立方体就是基于观众共振和循环的声音创造了一个音视空间。这个发光的立方体的每一面都反映了空间的视觉反射，这种反射不断地受到观众声音行为影响的效果。

### 3.4.3 触觉体验

触觉体验方法是体验互动装置最具特色和直接的方式，也是它与其他传统公共艺术作品的区别。可以触摸的公共艺术装置使艺术作品不再是在大厅中禁止触摸的孤独展品。相反，它是可以与公众互动的艺术，这使得公共艺术作品更加贴近观众。例如，纽约著名的时代广场上的互动心形装置。它的名字叫做“BIGNYC”，高十英尺，发光的“心”由400根由奥地利灯具公司大力支持的透明LED照明照亮的亚克力塑料管组成，可以在时代广场折射光线。当人们触摸这个心形传感器装置时，它的亮度会增加，跳动得更快，因为人类能量通过多媒体技术转化为光的视觉体验。这个公共装置主要通过触摸敏感的装置体现装置艺术的主题，当人们触摸装置的开关时，触摸引起的装置中的视觉变化向人们传达正能量。

## 3.3 多媒体互动装置中体验性与互动性的良性实现

首先，需要一个合适的地点。一个好的多媒体互动装置作品，如果放置在不适当的位置，将无法突出其特点和美感，也无法表达其意义。例如，如果互动媒体设备中有投影设备，应根据环境调整投影的亮度。一个主题是促进人们的低碳环保，使用公共交通工具的多媒体互动设备原本放置在公共交通使用的地方，例如地铁、公交车站等公共空间，它失去了原本的意义。相反，应该将其放置在车库或私家车主高频经过的地方，如城市街道、公园等公共场所，以便更好地向其观众展示。在促进环保方面发挥作用。梅赛德斯-奔驰 Smart 公司在国外街道上的路口交通灯处设置了“跳舞交通灯”，并在交通灯附近设置了一个动态捕捉室，路人可以走进去跳舞，动态捕捉室中的传感器读取舞者的动作，这些动作将反映在交通灯上，这不仅使等红灯的时间不再枯燥，而且确保行人可以安全地等待交叉路口的交通灯，防止危险的交通事故发生率。

其次，精确的信息传递。与只提供审美体验的传统公共艺术不同，多媒体互动艺术装置在审美过程中可以同时承担信息传播的功能。设计有艺术性且富有趣味性的作品，在传递信息的同时，可以微妙地影响和提高公众的审美水平。

第三，它需要科技的支持。涉及高科技的多媒体艺术互动装置不仅需要有的安装创意，

还需要依赖科技的支持，先进的技术设备能够实现多媒体互动装置的独特之处。技术参与的多媒体互动装置艺术作品是信息采集的输入设备，包括红外传感器、压敏传感器和计算机处理系统，这些都是通过计算机程序进行编程的。构成还包括与媒体体验系统相结合，用来反馈互动信息，用来生成图像、声音等，通常由音频、视频和其他播放设备组成。科技的融合是多媒体互动装置表达独特互动方式的基础媒介。【这一段有问题，重新录入】

第四，加强公众的主体性，把握公众的心理兴趣，多媒体互动装置需与观众的要求达成一致，准确把握观众的心理需求。它还应根据时代变化进行更新，否则将导致观众的抵制，导致互动过程失败和公众参与度的流失。因此，多媒体互动装置应充分发挥公共艺术的社会热情和文化价值，提高公共文化生活质量。加强公众的主体性。创作者应在创作构成中遵循易用性，使公众能够轻松便捷地参与，同时为公众创造有意义的体验。加拿大艺术家 Caitlind Brown 的“云”是一个由灯泡制成的云，它是一个互动灯光装置，在该装置中，公众可以取下或安装灯泡，或通过拉绳来熄灭和点亮灯泡。这个多媒体互动装置由钢材、金属拉绳和超过 6000 个亮的和烧毁的灯泡组成。该设计重新想象了废物，艺术性地重新处理了过剩材料，它不仅具有娱乐性，还推广了环保概念，其形状和视觉体验非常吸引人且富有创意。

在娱乐与审美需求之间取得平衡，多媒体互动装置为参与的公众带来感官愉悦。然而，一件优秀的公共艺术作品除了增强其自身的娱乐性外，还应增强其艺术传承和文化情感交流。审美需求不应该因娱乐而被削弱，一件好的多媒体互动装置艺术作品应该在娱乐与审美之间取得平衡。在日本东京银座的索尼大厦索尼广场，有一个水晶音乐雕塑，音乐装置的设计团队希望通过这个多媒体互动装置来表达互动体验的概念。整个装置结合了周围的环境，其照明随着美妙的音乐伴奏而变化，在装置前有一个捐款箱，摄像头连接到传感器，当有人在捐款箱中投入硬币时，灯光会转变成另一种闪烁的图案，相应地对人们的互动做出反应。该装置不仅让人们体验到互动的乐趣，也成为了银座的一道美丽风景。

## 参考文献

- [1] 孙欣, 尹晓峰. 空间与互动: 从全景艺术到互动艺术的叙事转化. 美术研究, 2018(4):3.
- [2] 李治路. 从互动到概念的新媒体艺术研究. 新闻与传播, 2022(1):2.
- [3] 李立峰. 论初中英语课堂教学中的互动教学艺术. 学术界, 2020(32):2.
- [4] 韩淑君. 公共空间的互动艺术创作. 中国字文化, 2019(10):2.
- [5] 王万锋. 师生互动激发小学语文课堂智慧之花绽放. 学术期刊, 2022(7).
- [6] 钱维宇, 张毅. 互动体验趋势下的城市景观设计研究. 工业工程设计, 2020, 2(1):5.
- [7] 丁子甲. 虚拟公共空间: 在线互动媒体. 福建质量管理, 2018, 000(008):24-25.
- [8] 穆一文. 从幕后到台前: 声乐表演艺术与大众文化的互动艺术. 青少年时代, 2019, 000(011):71.
- [9] 李昌益. 集体阅读教学中的互动艺术. 知识百科, 2019(2X):1.
- [10] 杨春萌. 教学理解艺术的初步研究. 教育进展, 2022, 12(6):8.

# 第7章 人工智能艺术：从个体制作 到人机共生

## 第1节 人工智能生成内容的版权问题研究

### 1. 人工智能生成内容概述

#### 1.1 人工智能生成内容的概念

人类有着悠久的使用机器帮助创作作品的历史。机器拥有高效的信息处理能力和精准的行为执行能力，这使得作品的传播形式极为丰富。很多学者对机器生成内容的定义提出了自己的见解。例如，《计算机艺术的法律问题》一书的作者乔纳森·C·杰克逊在书中提到，机器生成的内容应该是“计算机艺术”而不是“计算机作品”。人们使用技术让计算机运行特定程序创造内容，这些内容可以在显示屏上展示，或者通过某种复制形式固定并展示出来，作者认为这种“计算机艺术”的根源仍然是“人类的思维”。许多学者在对这类生成内容的命名和应用上还是持保守态度。他们大多数不使用“计算机作品”这一说法，而是称之为“计算机生成内容”、“计算机生成作品”等其他名称。

#### 1.2 人工智能生成内容的特征

##### 1.2.1 低成本、持续且高效的输出

随着芯片和服务器等硬件设施的不断发展，机器的读写能力迅速提高。在其发展对比中，2002年世界上计算速度最快的计算机仅刚刚达到2016年单个NVIDIA显卡的读写速度。特别是当人工智能通过云技术使用时，它不仅在质上改变了人工智能产生结果的速度和准确度，而且使得人工智能在花费比人脑更少的时间内做出更精确的判断。人类需要思考和利用的是，AI技术可以产生的大量、多线程的持续输出。

人工智能在新闻报道领域的最早应用是 Automated Insights 公司开发的为美联社等不同用户服务的 Wordsmith 产品。仅在 2013 年，Wordsmith 就生成了 3 亿种不同类型的内容，它每秒钟可以生成 10 篇文章，而且生成的内容还涵盖了 100 种不同类型的文章。预计在第二年，为数千个移动应用生成的内容量也可能是前一年的三倍。

### 1.2.2 人工智能生成内容涉及多个领域

人工智能生成的内容包括文本、音频、图片、视频以及互动内容，覆盖了新闻、诗歌、剧本、图片、视频、歌曲、电影甚至立体模型等不同类型的內容。其根本原因在于，人工智能生成的内容是通过代码编程的一系列算法的结果，包括无监督学习、深度强化学习、迁移学习、生成对抗网络等算法，使用自然语言模型进行文本处理和文本生成，利用语音合成进行音频生成，或使用计算机视觉技术进行图像识别、风格转换或图像生成。因此，只要对相关算法和模型进行调整，人工智能技术就可以应用于不同的生成内容领域。

例如，在 2017 年的双 11 期间，阿里巴巴使用人工智能鲁班制作了 4 亿张促销海报；在 2017 年腾讯 Media+峰会上，腾讯写作人工智能“梦想编剧”在嘉宾讲话结束后的一秒钟内写出并发布了新闻稿；Adobe 的 Project VoCo 项目能够在学习了某人的声音之后模拟该人的声音，并且修改相应的文本内容，说话内容也会随之改变；Nvidia 的 AI 可以在学习了脸部特征的基础上生成不存在的人脸照片；2017 年 5 月，展鲁文化出版了由微软小冰人工智能创作的现代诗集《阳光失去了玻璃窗》；上海博森数据开发的古诗词生成机器人可以随机作诗，或者根据指定的词生成藏头诗；纽约大学的研究人员使用人工智能学习了多种经典剧本之后，然后编写了一个名为《Sunspring》的科幻剧本，并将其制作成了一部 9 分钟的电影；华盛顿邮报开发的人工智能 Okun 能够根据用户的偏好为不同的客户定制广告；索尼 CSL 实验室开发的人工智能具有自动作曲、互动作曲、和声、变奏、渲染等功能；维多利亚大学开发的人工智能可以通过深度学习神经网络将一个无表情的脸部照片变成一个笑脸照片，等等。

### 1.2.3 人工智能生成的内容与人类创作的内容无法被普通人区分

人工智能生成的内容或人类创作的作品主要用于音乐、绘画、小说等领域，人工智能生成的内容与人类生成的内容之间没有本质的差异。人工智能输出的本质是代码，程序员或程序设计师通过程序代码编程，将人类常用的语法和语义表达编入其中，因此，对于查看内容

的第三人来说，它们的表现结果是相同的。

除了相同的形式表达之外，人工智能生成的内容与人类创作的内容在内容上也趋于一致，其中一些作品的相似度之高使普通观众难以区分。例如，在音乐领域的人工智能创作中，巴黎索尼计算机科学实验室的 Gaetan Hadjeres 和 Francois Pachet 开发了一个基于音乐家约翰·塞巴斯蒂安·巴赫合唱曲目的人工智能作曲系统，他们使用 352 首巴赫歌曲的音符将其转换成数据，创作出了 2503 部作品，然后邀请了 1600 名音乐家和音乐专业人士进行欣赏，其中超过一半的人认为这些作品是巴赫本人创作的。

## 1.3 可以由 AI 生成的内容类型

### 1.3.1 第一类人工智能生成的内容

这种类型的人工智能生成行为将向人类发出指令，并使用计算机实现它们的想法，就像人类在现实中使用纸和铅笔进行创作一样。如果创造力源自人的智力活动，个人或团体即使没有计算机的存在，也可以通过其他方式实现其想法。计算机在人类创造过程中只扮演了工具的角色。例如，使用文字编辑软件 Word 进行文学作品创作，使用 Photoshop 进行照片编辑，使用 AutoCAD 进行建筑作品创作等。

以新闻写作为例，自 2014 年人类开始使用机器人进行写作并投入新闻生产以来，机器人新闻已经产生了巨大的经济和社会价值，迄今为止，世界上许多国家的新闻产业已经开始实际使用人工智能进行新闻创作，如美联社使用的写作机器人“Wordsmith”，洛杉矶时报的机器人作家“Quakebot”，腾讯财经使用的“梦想编剧”，以及新华社使用的“快笔小新”。对于需要收集多方面数据且工作量较大的新闻写作，人工智能使用大数据快速整合和收集信息，其写作效率和产出是人类难以匹敌的。

例如，由 Automated Insights 开发的 Wordsmith 本质上是一个语言生成引擎，它能够实现从基础数据到知识、洞察力等更高层次内容的跳跃。人工智能研究包括这两种重要技术，即数据库知识发现和自然语言处理，并将两种技术相结合。

生成内容的步骤主要包括：1) 获取数据：基于报导对象的各种形式的数据和材料，以及客户提供的各种数据，涵盖文本、图片、图表等各种形式的数字内容。2) 分析数据：解读数据的内在联系和相关背景，通过数据分析过程勾勒出数据之间的有机关系。3) 提炼观点：解释目标客户的数据和数据所反映的趋势，并将此类信息放在更宏大的背景中，可以细分为行业、社会乃至国家级别进行认识，比较两者的输出结果，从而提出一些可行的执行建

议。4) 结构和格式: WordSmith 使用其自然语言生成功能来叙述从前面的分析和提炼中获得的想法, 并根据需要生成各种形式的文本。5) 发布: WordSmith 将生成的文章发送到客户指定的平台。

### 1.3.2 第二类人工智能生成的内容

除了上述作为工具的人工智能之外, 神经网络和机器学习技术也得到了发展和利用, 人工智能技术实现了飞跃性发展。人工智能在创造力方面越来越接近人类, 其创造活动不是基于人类的具体指令, 而是在预设程序的抽象框架下进行, 其中一些生成的产品满足授予知识产权的条件。

这类人工智能具备机器学习能力, 它可以通过主动学习解决一些没有预定解决路径的问题。今天我们处于大数据时代, 大数据技术使得人工智能机器学习如虎添翼, 可以通过海量的信息进行学习和判断, 在发展过程中逐步纠正偏差错误, 判断越来越成熟。此外, 应该注意的是, 在内容生成中, 人工智能能够基于其获得的信息创建类似人类思维的内容, 这也意味着它在一定程度上逐渐摆脱了人类的预设计。近年来, 人工智能的“深度学习”取得了突破性进展, 它也模糊了程序预设和独立思考之间的界限, 这是通过模拟人脑神经网络的结构, 通过深度神经网络, 使得人工智能在没有预设稳定规律的情况下, 通过大量数据分析的相关程度, 然后用来模拟人脑来分析数据。

在可预见的未来, 随着大数据、神经网络、深度学习和算法的持续发展, 第二类人工智能生成的内容将越来越多地出现在人类的视野和日常生活中。

## 1.4 人工智能的发展阶段

自人类开始使用计算机作为提高工作效率的工具以来, 我们打开了通往新世界的大门, 对计算机及其后续的人工智能的持续深入研究开发出了许多成果, 根据人工智能系统的“智能”程度, 作者们将其分为几个阶段。按照人工智能的智能程度, 分为三个阶段:

第一阶段, 准智能阶段。这一阶段始于 20 世纪 80 年代, 伴随着相机和摄像机的出现在创作作品的这一阶段, 机器以工具的形式在人类创作过程中发挥辅助作用, 这并不助长作品的创造性, 而只是实现人类创意的载体或工具。因此, 这一阶段并不是我们现在所说的人工智能, 而只能称为准智能。

第二阶段, 人工智能阶段。这一阶段的特点更为明显的是使用各种软件, 比如一些可以自动创作的软件, 应用场景在音乐、艺术和新闻等领域。作者认为, 尽管一些人工智能通过

学习和判断生成新的内容，但这一阶段人工智能接近人类智能的能力并不是自发生成的，而是在编程时预先植入的路径。在这个阶段，设计师使用机器学习和模拟大脑神经网络技术，可以根据需要收集数据并学习不同的操作逻辑，这一阶段的人工智能可能会根据人类的抽象设定生成不同的内容和作品，但它仍然没有脱离程序员设定的算法。例如，报纸和期刊使用的人工智能可以使用给定的数据结构、报告主题、常见模板等算法生成新闻内容。然而，这类人工智能仍然依赖于人类给定的模型来操作，一旦涉及到创造性思维、批判性思维、社会理解等，准人工智能阶段的机器就会显得捉襟见肘。

第三阶段，超人工智能阶段。人工智能发展到这个阶段，成为能够模拟生物神经系统的智能系统。这项技术仍然处于技术的前沿，这种人工智能的运作模式是测试人类对不同音乐和音色的感知，以创造出满足人们期望的音乐作品。基于对人脑的研究，研究人员发现，通过让听众听一个和弦，听众对和弦的期待越大，引导和弦的共鸣就越大。

应用到人工智能领域，通过人们对音乐满足或不满足的期待，它代表的产品是 MUASCT，可以根据这个原则为给定的旋律生成相应的期待。日本研究者已经成功地借助人工智能解读了人类的思维和想象，并能将大脑的运作机制应用到人工智能的研究中，使现有的人工智能进入到一个新的阶段。

目前，在人类生活中能够设计的人工智能仍然处于第二阶段，即算法智能阶段，人工智能仍然需要依赖人类的操作和指令来生成相应内容。在这个阶段，还需要大量的人为因素存在，人工智能的“智能”表现仍然在既定的模式和逻辑中，通过数据建模过程中的连续训练，安排和组合现有的数据或赋予它近似人类的判断能力。

## **2. 人工智能生成内容的著作权问题**

### **2.1 人工智能是否能成为作者**

#### **2.1.1 自然人作者的视角**

在现行著作权法的范围内，获取著作权的途径有两种原始方式，一种是通过自然人进行某种智力劳动创作而获得，另一种是法律虚构，即在特定的作品如职务作品和委托作品中，单位、法人或其他组织基于著作权法的规定或合同协议获得著作权，成为著作权的主体。

如果机器能成为著作权的主体，那么作为人工智能载体的机器、软件和算法本身是不能行使它们所拥有的权利的，如果无法行使它们所拥有的权利，那么授予它们权利的意义便无

从谈起。著作权法调整的是人类社会内部的社会关系，而不是人与机器、人与动物之间的关系。

在人工智能与自然人之间的差异方面，人类的创造是包含人类思想、情感和表达的智力活动，只能通过自然人的血肉之躯来实现。从历史唯物主义的角度来看，人类的意识是社会性的，人之所以能够思考，是因为他们的思考能力是社会长期劳动和时间的产物，这一属性是计算机永远无法模拟的。人类大脑是由数亿个神经元组成的神经网络，每个神经元的复杂性远远超过计算机的运算，人脑能够根据不同的环境变化做出相应的调整 and 变化，而人工智能则是基于人类设定的模式、算法、具体数据来穷尽所有可能性，这一过程是机械的、无意识的。因此，尽管机器人在效率和准确性方面可以远远超过人类，但过程仍然不是有意识的，不能拥有与人类相同的思考能力，也不能进行创造性的智力活动。

从工具主义的角度来看，目前人工智能仍在人类的控制之下运作，即它是人类劳动的对象，人类是这一过程的主体。权利主体和权利对象是相互对应的，其法律地位不可互换。主体与对象的关系是确定的，权利对象只能是法律支配的对象。在著作权法下亦是如此，作品确定和权利归属的要求离不开以人为主体的核心要素。可以看出，权利主体和权利对象之间的不同划分是不能相互转换的，现行法律体系否认主体在任何情况下成为主体的可能性，同时也不存在支持对象或物归对象拥有法律权利的可能性。在现行的著作权体系下，严格来说，人工智能本身只能局限于对象，当然也不能转化为权利的主体。

### 2.1.2 法人制作者的视角

在现有著作权法体系下，除了自然人可以成为作者并享受权利之外，单位、法人或其他组织也可成为作者。在单位、法人或其他组织提供了主要的物质和技术条件，使得员工能够按照特定目的创作时，虽然作品最终是由自然人完成的，但成果主要被单位、法人或其他组织的资源条件所使用。为了更好地优化资源配置，协调作者对其作品的独占权与公众对他人作品的独占权，实现公共利益与个人利益之间的平衡，立法者将作品的作者身份规定为单位、法人或其他组织。

从这种法律虚构的视角来看，现行法律体系下的法律已经突破了作者必须是自然人的限制。可以看出，尽管目前第一类人工智能生成作品的归属还没有法律障碍，人类仍然可以是作品的作者。但是随着科技的进步，在不远的将来，人类将面临第二类人工智能生成内容所带来的大量法律问题。届时，这些机器可能创作出符合原创性标准的作品，我们如何定义作

品的作者将成为必须面对的问题。

立法者可能已经为我们未来的法律发展提供了一些思路。单位作品和人工智能生成器之间有一定的联系和一定的区别。相似之处在于，单位、法人或其他组织和机器都不是自然人，没有人类所具有的创造性思维。不同之处在于，单位、法人或其他组织为自然人提供了一定的物质和资源帮助，使得自然人可以按照代表单位特定目的的意志进行创作，而人工智能则利用计算机模拟人脑神经网络进行深度学习，可以独立完成作品的创作。

历史上，法人获得独立法律人格是因为一些组织需要永久存在，并且在法律上有资格处置它们所拥有的财产或偿还其附带的债务。从激励理论的角度来看，知识产权制度设立的最终目的是为了促进全人类文明的进步。从这个角度来看，如果人工智能被赋予法律人格，也是对创新创造的一种激励，有利于科学技术、文学、艺术等领域的发展。目前，欧盟已经尝试通过法律手段将人工智能定位为主体，并由欧盟法务委员会向欧盟提出了建议，其内容是提议将人工智能定位为“电子人”，应当作为人工智能机器人进行注册，并且还可以赋予其特定的权利和义务，这个“电子人”也应当有资格缴纳税款，欧盟这一尝试很可能对传统的民事法律体系产生影响。

在当前的民事法律体系中，民法的两个基本制度包括民事主体和民事客体。从法理学的角度来看，如何从主客体二分法中分解问题，是我们需要深入思考的。在现行法律体系中，对主体和客体有严格的区分。客体没有人所具有的意志和意图表示，应当被称为客体，即权利的客体。虽然在现阶段，人工智能被视作“人”并赋予其一定的主体资格，但在民法理论中找到相关依据颇为困难。在民法中，我们规定有两类人，一类是自然人，另一类是作为投资者提供物质信息的法人，它们都享有民事权利并承担相应的义务。机器人既不是具有思维和意识的自然人，也不是提供作品中投资的法人，它们的法律虚拟仍需要在法理学中进行研究和讨论。然而，随着人工智能技术的快速发展，主体和客体之间的界限变得越来越模糊。借鉴人类使用法律模拟的经验，这种方法突破了作品作者必须是自然人的界限，并且法律模拟方法并没有导致自然人利益的流失，反而极大地推动了文化领域和人类社会的发展。因此，为人工智能专门设计的独立法律体系也可能遵循这一模式，这可以在上述欧盟将机器人分类为“电子人”的计划中看出。

从洛克的劳动理论视角来看，这种创作行为本质上是一个劳动过程。通过创作行为的劳动，将想法具体化为作品，对劳动所产生的独有权利的认可，符合洛克劳动理论所提倡的财产获得条件。从哲学层面思考人工智能的劳动，类比于法人作者制度，公司雇员让人工作，但实际的财产所有权确实属于公司，在这个过程中公司并未进行劳动。如果人工智能作为作

者，那么人工智能进行的创作就像作者一样进行了劳动，但财产所有权需要另外规定。

因此，曹原认为，人工智能可能是作品的事实上的作者，但成为法律意义上的作者很大程度上将是一个价值判断和政策选择的问题。因为，根据第二种人工智能产生内容的情形，人工智能将脱离人类控制生成可以被视为作品的内容。

然而，从版权法对版权保护的意义上来看，版权保护的最终目的是鼓励作者创作更多内容，并赋予作者特定的权利，从而平衡个人利益和社会利益。但从这个角度来看，我们将面临一个问题：人工智能不需要以这种方式得到奖励，它也能进行大量的创作，这样的权利保护对计算机没有特别的激励意义，真正需要被保护和激励的是开发这些人工智能的自然人或单位组织。然而，实际上，开发人工智能的自然人或单位组织并不是直接从事作品创作的主体，因此我们可以考虑人工智能的实际作者和目前法律意义上的作者，自然人或组织仍然是法律上的作者，但对于机器人可以称之为事实上的作者。

从这个角度看，法律是一种不断调整和适应社会发展的社会规范。从法人作者打破了作品必须由自然人创作的背景来看，现代法律中的“人”的概念已经发生了变化，现代法律中的自然人早已与传统的自然人分离，并且增加了法人的存在。因此，在未来的法律发展中，我们可以结合法人作者的概念，为人工智能产品的作者进行制度创新。

## 2.2 人工智能生成内容是否可以成为作品

根据中国的《著作权法实施条例》第二条，所指的作品必须具备三个条件：（1）文学、艺术和科学领域中思想或情感的表达；（2）原创性；（3）能以某种形式被复制。至于人工智能生成的内容能否成为作品，根据人工智能生成的内容是否可以被人类预设，可以将作者分为第一类人工智能和第二类人工智能。第一类人工智能的本质还是同质化创作，设计者可以预测其输出结果，输入结果相同，输出的结果必然相同，它更像是人的手臂延伸的工具。但第二类人工智能则不同，虽然设计者可以设计和制造人工智能，但人工智能本身具有机器学习和模拟大脑神经网络的能力，可以通过没有预设算法或规则的自主学习，模拟大脑分析、构思、安排、选择数据组合生成内容，这样的人工智能使得程序算法与思维的界限进一步模糊。因此，以下是从作品的构成要素和原创性的角度，对第一类和第二类人工智能生成内容进行的分析和研究。

### 2.2.1 第一类人工智能生成的内容能否成为作品

对于第一类人工智能生成的内容，以及基于人类预设规则的计算机衍生作品，此时人工智能仅存在作为提高效率的人类手臂延伸的工具，并且许多国家已经开始研究关于计算机衍生作品的版权保护问题。

人工智能内容生成的最早问题出现在 1950 年代，当时计算机技术刚刚起步，美国的法律学者们开始研究和讨论机器人创建系统。一些美国计算机专家可以利用计算机开发的软件每小时创造 4000 首歌曲，但当他们试图为这些歌曲申请版权时，美国版权局以以前从未为机器创作的作品注册过版权为由，拒绝注册他们的作品。在 1970 年代，美国国会建立了“新技术时代作品使用审查委员会”，目的是审查新技术如何影响版权，并在长时间的调查研究后发布了相关报告，认为计算机程序仅作为被动创作的工具存在，并没有直接参与创作行为。1993 年，日本版权审查委员会第九小组的“计算和生成内容部门”发表了报告，认为计算机在人类创作过程中仍是道具，不能成为创作的主体。

因此，作为第一类人工智能生成内容，它是按照既定的模型、算法进行的，无论哪种人工智能产生的内容都是相同的，不能体现独创性，而且既定的模型、算法属于思想的范畴，所以第一类人工智能产生的内容不属于著作权法中的作品。

### 2.2.2 第二类人工智能生成的内容能否成为作品

对于第一类人工智能生成的内容，学者们之间并没有太多争议，但我们需要更多关注第二类人工智能生成的内容。因为从科技发展的模式来看，计算机不可避免地不仅仅会局限于辅助工具，而且很可能发展成为参与创造过程的主体，即当人工智能生成的内容发展为第二类生成内容时，现有的版权法体系必须进行重新调整。但如果仅仅是第一类生成的内容，它仍然属于现有版权法保护的范畴。

对于第二类人工智能生成的内容，这类内容的一个特点是，从外部很难区分人工智能生成的内容和人类创造的作品。版权法中关于思想表达的二分法指的是版权法不保护如思想或情感等抽象的思维方式，而只保护人类推理和联想的观点表达。人类思维的表达与第二类生成内容的表达也有所不同，人类思维的表达是基于大脑的思维联想、观点的整理，而第二类生成内容是基于大数据分析、算法和语言的整合，第二类生成的内容是基于“计算机学习”和模拟“神经网络”按照一定逻辑整合数据。

除此之外，第二类人工智能的一个特点是人类无法预测其最终输出。第一类人工智能的本质仍然是同质化的创造，设计者可以预测其结果的输出，输入的结果是相同的，它们的输出结果必须是一样的。第二类人工智能则不同，尽管设计者可以设计和制造人工智能，但人工智能本身具备机器学习和模拟大脑神经网络的能力，可以在没有预设算法或规则设置的情况下主动学习，模拟大脑分析、组合、整理、选择数据组合来生成内容，这样的人工智能使得程序算法和思维的界限进一步模糊。

从原创性的角度来分析第二类生成内容的版权性，作品的一个重要特征是它必须具备原创性，其内在含义是作品应当是独立完成的，并且具有最低程度的创造性。相对于最低创造性步骤和独立完成这两个要素，最低创造性步骤在判断原创性时比独立完成更为重要，而独立完成对于确定作品所有权更为关键。在司法实践中，最低创造性更多是一种主观判断，其标准最终由人类决定。偏向于人格权的学者认为，作者创作的作品是作者思想的体现，作者的想法和创意以某种载体或形式表达出来。代表性学者康德认为，作品是作者个性的具体表现形式，所谓的版权即是作者人格权的外延或表现。作品应当反映作者的个性，且表达方式应当是可感知的，这也是立法者保护版权的合法性基础。

作为一个民法国家，作品不仅使作者具有财产权，还代表了作者的人格权，其表现不是普通的财产或作者的人格反应。黑格尔是人格权理论的代表之一，他认为作品是作者意志意识的外在表达，作者利用其创作的作品来展示作者的精神和个性，其作品也可以展示他的个性印记，因此他应当享有作品的版权，为了在社会利益和个人利益之间取得平衡，立法者给予作者的个人权和财产权以一定的保护。然而，由人工智能生成的内容是根据某种逻辑算法在人类语言中整合数据的结果，这一过程很难说是作者人格的体现。

作者认为，相比之下，作品是否必须反映作者人格才能成为作品是有争议的。因为，法人的作者如何主张单位、法人及其他组织的人格得到反映呢？法人作者是一个单位法人提供物质资源，使自然人根据具体目的进行创作，完成的作品是否反映法人的人格是难以界定的。例如，德国版权法就没有对法人作者进行法律上的模拟，因为作品无法反映作者的人格和精神，直接创作只能是具有创造能力的自然人，即“创作者原则”，所以不存在作品是否反映作者人格的问题。

此外，在一些司法实践中，对作品原创性的判定会关注新作品与现有作品在表达上是否有所不同。如果只遵循这个标准，人类创作与人工智能创作之间的界限已变得愈发模糊，例如，巴黎索尼计算机科学实验室的 Gaetan Hadjeres 和 Francois Pachet 开发的人工智能作曲系统针对音乐家约翰·塞巴斯蒂安·巴赫的合唱曲目录。该系统将 352 首巴赫歌曲转换成 2,503

条数据，然后邀请了 1,600 名音乐家和音乐专业人士进行鉴赏，其中超过一半的人认为这些作品是巴赫自己的作品。如果比较人类智能与人工智能，目前最先进的人工智能机器测试显示，人工智能的智力已达到四岁儿童的水平。有鉴于此，一些学者认为，作者身份制度的主观判断标准逐渐使版权制度关于原创性理解的客观标准发生了转移。由于版权的原旨是关注创造性本身，那么版权法不应仅保护人类产生的创造性，也就是说，如果人工智能能够满足其判断标准，被认为具有一定的创造性，那么人工智能生成的内容也应得到版权法的保护。

观念和创造需要通过某种载体或形式来表达。其代表学者康德认为，作品是作者人格的具体体现，所谓的版权，即是作者人格权的外在化或延伸。作品应反映作者的个性，表达应当是可感知的，这也是立法者保护版权的合法性所在。

作为一个民法系国家，作品不仅令作者拥有财产权，还代表了作者的人格权，其表现不是普通财产或作者人格的反应。黑格尔是人格权理论的代表人物之一，他认为作品是作者意志意识的外在表达，作者通过其创作的作品来展示其精神和个性，并且他的作品还可以展示他的个性印记，因此他应当享有作品的版权。为了在社会利益和个人利益之间取得平衡，立法者为作者的个人权和财产权提供了一定程度的保护。然而，由人工智能生成的内容是根据某种逻辑算法在人类语言中整合数据的结果，这一过程很难说是作者人格的体现。

然而，作者认为相比之下，作品是否必须反映作者人格才能成为作品是有争议的。因为法人作者如何主张其作品反映了单位、法人及其他组织的人格？法人作者是指一个法人单位提供物质资源，使得自然人根据特定目的进行创作，完成的作品是否反映法人的人格则难以界定。例如，德国版权法就没有对法人作者进行法律上的模拟，因为作品无法反映作者的人格和精神，直接创造只能是具有创造能力的自然人，即“创作者原则”，因此不存在作品是否反映作者人格的问题。

在某些司法实践中，对作品原创性的判断确实会关注新作品与现有作品在表达上是否存在差异。如果单纯遵循这一标准，人类创作与人工智能创作之间的界限确实越来越模糊，例如巴黎索尼计算机科学实验室的 Gaetan Hadjeres 和 Francois Pachet 开发的人工智能作曲系统，该系统针对音乐家约翰·塞巴斯蒂安·巴赫的合唱曲目录进行了开发，将 352 首巴赫歌曲转换成了 2503 条数据，并邀请了 1600 名音乐家和音乐专业人士进行欣赏，其中超过一半的人认为这些作品是巴赫自己的作品。如果从人类智能与人工智能的对比来看，当前最先进的人工智能机器测试显示，人工智能的智力已经达到了四岁儿童的水平。有鉴于此，一些学者认为，作者身份制度的主观判断标准逐渐转变了版权制度对原创性理解的客观标准。既然版权的含义是聚焦于创造性本身，那么版权法不应只保护人类产生的创造性，也就是说，如果人

工智能能够满足其判断标准，应当被认为具有一定程度的创造性，那么人工智能生成的内容也应能够受到版权法的保护。

## 2.3 对人工智能生成内容保护的意义

### 2.3.1 激发人工智能开发者的积极性

经济功能被认为是知识产权保护最重要的目的之一，在本文讨论的版权法领域内，通过授予作者或传播者一定的权利来实现经济功能，最终促进整个文化市场的繁荣和发展。激励论认为，应当给予作者或传播者某些专有权利来获得经济回报，最终激励他们产生更多内容丰富的文化。尽管作者创作的最初目的不一定是为了获得经济回报，但如果不保护其经济权利，将导致文化市场由于违法成本低而出现大量抄袭现象，最终可能导致劣币驱逐良币的现象。如果作者或其背后的投资者所付出的努力得不到相应的回报，愿意投入智力创新的人会越来越少。因此，一方面给予作者或传播者一定的经济利益可以得到保护，另一方面可以鼓励作者创作更多作品，最终为市场增添更多作品。

与此同时，人工智能生成内容的低成本和高效率特点将在市场上大量增加此类内容，人类作品的生成速度和效率远不及人工智能，人类作者创作的原始动力和热情可能会受挫，作为创作知识源泉的作者正面临着枯竭的危险，市场上的作品数量将变得极其庞大。

然而，随着科学技术的进步和发展，人工智能创作是根据特定算法对大数据进行整合和计算，而不是创造性的精神劳动。对于第一种类型的人工智能生成内容，其生成的内容仍然受到人类的控制，所以根据激励理论，最终是激励者的操作而不是人工智能本身，在目前一些国家的法律和实践已经注意到了这个问题，例如英国等普通法国家从法律实用主义的角度出发，人工智能生成的内容属于机器的操作者，所以如果是第二种类型的人工智能生成内容，工作者的智能可能独立于人类控制生成新的内容，那么激励理论就不再适用于激励操作者的行为，未来的法律可能更倾向于激励人工智能的程序员。

### 2.3.2 平衡市场对人工智能生成内容的利益分配

除了激励人们创作之外，版权法还可以平衡人工智能生成内容各方面的利益。法律根据创意成果的创作和传播方式来分配利益，这涉及到许多主体，包括传播者、创作者和公众之

间的利益关系，以及众多传播者和衍生作者之间的利益平衡。作为知识产权基本理论之一的利益平衡论认为，版权所有者的利益、公众的利益以及整个社会的利益之间应当保持平衡。

版权法之所以受到冲击的根本原因，在于现有的版权法难以平衡新的利益关系。如果能够妥善处理作者对创作的热情与人工智能之间的平衡，版权法将符合社会 and 不同利益团体的要求。知识产权的确立是为了保护公共利益，如果利益关系没有得到平衡，过度保护或者保护不足权利持有者追求利益的权利，都是违背其原初意图的。

人工智能生成的内容作为一种新的内容来源，对现有的版权系统产生了一定的影响。原本由人类花费大量精力创造的内容，现在因为人工智能的低成本、高效率的大量输出而变得泛滥，人类精心安排的利益分配制度框架将变得不平衡。如果人工智能生成的内容不受版权保护，市场上将出现大量与人类创作的作品实质上没有太大区别且未受版权保护的作品，由于权利未受保护，这些作品将被置于公共领域。这可能的后果是，对于那些可能使用作品的人来说，市场上公共领域的人工智能生成内容大量，使用这些内容不需要支付与版权相关的成本，相关市场的版权授权和转让交易将大幅减少。如果版权价值只能依靠作者的声誉或高度原创的作品来产生，那么人类创作的社会经济价值将逐渐趋向于零。人类创作作品的经济动力不断减弱，作品对人类的市場吸引力下降，这将导致市场失去流动性，最终人类将被排除在市场之外。目前，人工智能的工作原理始终是利用数据整合创作作品，如果来自人类的作品减少，那么人工智能可分析的数据也会减少，导致的结果是，在人工智能将人类排除在市场之外之后，新的创新作品没有及时补充，人工智能只是使用现有的作品进行循环整合，作品市场将变得庞大而荒凉。

此外，与作者创作的作品中版权的直接销售相比，作品带来的传播价值所创造的经济价值——也就是邻接权——往往更大。传播者和投资者在作品的传播中扮演了重要角色，使得原有的版权市场不断扩大。因此，基于上述条件形成的版权市场，如音乐和新闻产业，已经受到冲击，使得人类在发展和进步上的空间逐渐缩小。同时，允许人工智能生成的内容在未授予版权的情况下进入公共领域，也将导致人类在文学和艺术领域过度依赖人工智能进行创作。

### 2.3.3 避免滥用内容的问题发生

侵占，在字典中的解释为傲慢地声称拥有。标题的内容与民法中的抢先权类似，但是抢先权内容的前提条件是所获得的财产应该是动产，而使用标题所指的是版权法中的内容作为

无形财产的知识产权，因此侵占的内容与抢先权之间还是存在差异的。

未能保护由人工智能生成的第二类内容也可能导致滥用内容的问题。那些对人工智能生成内容没有任何权利的人，因为人工智能生成的内容没有权利归属，可能会声称自己是权利持有者的机会行为，即如果人工智能生成的内容流入公共领域，并且其内容形式与人类创造的内容无法区分，这可能会导致某些人声称对人工智能生成的内容拥有权利。利用邻接权等从事作品的二次利用，即传播等，由此产生的内容的经济利益问题将会产生更大的影响。例如，如果由人工智能生成的第二类内容小说被改编成电影和电视作品，其作品本身未带有任何权利就流入公共领域，而从该作品改编的电影和电视作品也没有权利基础。最终的结果是，一些本不应受到保护的主体反而获得了保护，而一些不存在的权利持有者声称拥有权利并实际享有这一权利。

这可能会产生以下后果：首先，版权法规定了权利的期限，如果没有权利持有者，那么其规定的期限就无从谈起，也无法进行改编；其次，一旦相关行业察觉到滥用内容的问题存在，就无法确认内容是否由人工智能生成，因此不可避免地会对行业的许可业务产生影响；第三，由于难以区分人工智能作品和人类作品的差异，如果人类智慧生成的内容未受到保护，这就是一种傲慢的问题，所以只要不被谴责，就难以追究责任。

根据上述分析，我们可以将人工智能生成内容的作者区分为事实上的作者与法律意义上的作者，自然人或组织作为法律作者，人工智能作为事实上的作者。对于第二类人工智能，在一定程度上，我们可以运用法律虚构来模拟它为作者，这为我们解决人工智能不能成为作者的问题提供了解决方案。

此外，对于人工智能生成内容的版权能力应该通过区分不同类型和子情境来讨论：第一种类型的人工智能生成内容不能反映出缺乏原创性，因为它是按照已建立的模型、算法进行的，由此人工智能产生的内容是相同的，而已建立的模型和算法属于思想范畴，因此在第一阶段产生的内容不属于版权法中所指的作品，不能成为版权法意义上的作品；第二种类型的人工智能生成内容，因为它生成的内容不是预设的，并且介入作品的创造过程生成与人类表达形式相同的内容，从原创性的客观标准来看，它是构成原创性的条件，所以有可能产生具有原创性的作品。

此外，从激励理论、平衡理论以及避免内容侵占的角度来看，保护第二类人工智能产生的内容具有极大的意义。从目前社会对人工智能发展的热情来看，一方面保护由人智生成的内容并给予作者或传播者一定的经济利益可以得到保障，另一方面也可以激励作者创作更多作品，最终丰富和促进文化市场的繁荣发展。从利益平衡的视角保护人工智能生成的内容，

可以平衡传播者、创作者和大众的利益，以及众多传播者与衍生作者之间的利益平衡。另外，从滥用内容的问题来看，如果人工智能生成的内容得到保护，可以避免生成的内容流入公共领域，阻止投机者对孤儿作品的投机行为。

### **3. 国际组织和一些发达国家对人工智能生成内容版权的保护及其启示**

#### **3.1 国际组织和一些发达国家对人工智能生成内容版权的保护**

国际组织和世界上一些国家已经开始对人工智能生成内容进行一系列的讨论和研究。例如，自 20 世纪 70 年代以来，英国政府对计算机生成内容进行了立法研究，到 2016 年日本政府成立了专门机构进行人工智能生成内容的立法研究。这些研究的目的是为了保护与人工智能相关的利益相关者，并防止在法律争议中无法依靠。尽管目前，在现有的人工智能立法讨论中，绝大多数国家仍然局限于机器的概念，并且对于人工智能概念进入法律条文持谨慎态度。机器在人类创造过程中以工具的形式发挥辅助作用，这并不有助于作品的创意，而是人类创意实现的载体或工具，而人工智能则是模拟研究人类智力的研究、开发技术科学，它涉及包括理论、方法、技术等非常丰富的内容。然而，从人类智力发展的阶段来看，作为工具的机器是人工智能发展的基础阶段，各国对机器生成内容的研究作为人工智能的基础技术对人工智能的立法研究具有重要意义。因此，以下是尝试整理不同国家和组织中机器和人工智能在法律规范中的定性和现状，以及总结其经验，以便更好地推动中国人工智能生成内容版权法律研究的进行。

##### **3.1.1 国际组织对人工智能生成内容的保护**

1979 年，联合国教科文组织和世界知识产权组织组织了一次关于如何使用计算机技术在版权范围内生成内容的讨论会议，代表们对此进行了探索。关于机器生成内容的讨论并不顺利，这个问题的讨论从 1980 年持续到 1982 年，在不同政府专家组成的委员会分析和论证之后，该组织发布了《关于解决由计算机系统产生的内容及其创作引发的版权问题的建议》。

就目前机器发展的现状而言，没有必要对版权法进行原则性改变，现行的版权法足以涵盖计算机生成内容的保护范围。具体来说：1) 计算机仅作为创作作品的辅助工具使用，其用于该目的应被视为技术应用；2) 符合国际条约和各国法律规范是计算机生成内容受版权保护的基本要求；3) 关于计算机生成内容的创作者，创作主体应是给予计算机原创性元素的人，如果软件设计者对生成内容有直接影响，可以在作品中标注为合著者；4) 涉及的计算机生成内容应不足以按照版权法的标准被称为作者生成的内容。

联合国教科文组织和世界知识产权组织在 1991 年组织了一次会议，讨论修订有关保护文学和艺术作品的伯尔尼公约条款。计算机生成内容在组委会的初版讨论中被提及，但后来被排除在外。涉及计算机生成内容的讨论包含以下内容：1) 所涉及的计算机生成内容应符合版权法的标准，且人的贡献程度不足以称之为作者生成的内容；2) 必须操作创作作品的人应为版权所有者；3) 法律不应认可作品的人格权；4) 其生成的内容的版权保护期限应为 50 年。

世界知识产权组织 (WIPO) 在 2017 年 10 月发布的一份出版物中讨论了对人工智能生成内容的保护，并提出了两种处理人工智能生成内容的方法。其中一种观点认为不应保护人工智能，因为从版权法的角度来看，计算机不是作者，因此不应授予版权保护。这一观点得到了欧洲联盟法院的支持，认为作品必须能够反映作者的智力创作并体现作者的个性，从而表明只有自然人才能成为作品的作者。另一种观点是将生成内容的权利授予程序员，如印度、爱尔兰和新西兰等国目前的做法。

### **3.1.2 英国对人工智能生成内容的保护**

英国是第一个在版权问题上积极反映计算机技术影响的国家，在不同国家的知识产权法律中，人工智能生成内容的法律定位和所有权划分已经得到明确规定。1973 年，英国就此问题组织了一个特别委员会研究版权法的修订，并在 1977 年总结了一份工作报告，该报告指出计算机产生的内容应被视为作为工具使用而生成的内容，内容的版权所有者仍然是一个或多个操作内容生成机器或向其发出指令的自然人，成为其内容的作者。

然而，在 1981 年，委员会对计算机生成内容的作者应该是负责计算机处理的法律责任人这一事实进行了一些调整。1988 年英国《版权法》规定，如果一个文学、戏剧、音乐或艺术作品是由计算机产生的，那么作者应该是负责创作该作品所需安排的人；计算机作品的版权保护期限为 50 年；计算机生成的内容没有人格权；计算机生成的内容是在没有任何人

类参与的情况下产生的。此外，计算机生成的作品被定义为“没有人类作者的计算机生成”。这项规定背后的思想是，通过承认创建能够生成作品的计算机所做的工作，所有人类作者都是作品作者的例外。

### 3.1.3 南非对人工智能生成内容的保护

深入研究人工智能主要在英国和美国等普通法国家进行。南非作为非洲的一个国家，也在人工智能内容方面有一定的法律实践。其 1978 年版权法规定，作者应该是实际操作生成作品的人，即使是计算机生成的内容，也仍然追溯到自然人的作者。在南非的一个案例中，2006 年 Haupt 诉 Brewers Marketing Intelligence (Pty) Ltd. 及其他人案件中，计算机生成内容中的版权问题成为讨论的一个焦点。其最高法院的最终决定显示，尽管计算机生成的作品和计算机辅助的作品之间的差异没有权威性的确定，但就案件本身而言，人为生成内容的版权问题是案件走向的关键，比如作品的定位、作者的确定和权利的归属等。授予权利的原因是因为创作作品的人需要被激励，但如果计算机生成的内容被授予计算机权利，最终还是会追溯到自然人，因此版权问题的讨论仍应围绕人类进行。

### 3.1.4 澳大利亚对人工智能生成内容的保护

澳大利亚版权修订委员会在 1993 年发布了一份关于计算机软件保护的报告。除了正常讨论计算机软件保护的问题外，报告还讨论了计算机生成内容的问题，报告的态度与澳大利亚版权法一致，认为版权所有者必须是自然人。从这个角度出发，澳大利亚基本上遵循英国法律模式，限制计算机生成作品中知识产权的所有者为操作计算机软件的人，并要求操作者对创造行为承担法律责任。

然而，在 2002 年的 Desktop Marketing Systems Pty Ltd v. Telstra Corporation Ltd 案件中，原告主张其是由计算机软件自动生成的电话簿的知识产权所有者。对此，本案的审判法官 Keane 指出，对于文学作品，其版权的基础在于人类的创造。如果一个产品仅仅是在人类操作下由程序锁生成的，那么它不能称为一个作品，自然也就不具有版权。法官 Perram 持有这样的观点：如果人类对计算机程序的控制可以被视为对创作过程的调整，那么确实可以成为作者。然而，如果计算机程序的操作者对最终产品的形式没有控制权，那么所谓的创作过程不能被视为操作者对产品生产的足够独立的智力贡献。基于上述讨论，本案的原告不被视

为作者，而电话目录在版权法中不能被视为一个作品。

因此，在澳大利亚的知识产权法律体系中，尽管在政策层面对 AI 生成器持有成熟的态度，为 AI 操作者成为生成产品的作者提供了制度上的开放性；然而，在实践层面，人工智能操作者被确认为作者的标准较高，即操作者对生成物质形态必须具有一定程度的控制权，才能满足认定标准。目前，在澳大利亚的司法实践中，纯粹由人工智能产生的产品不能被认定为可以享有版权的作品。

### 3.1.5 美国对人工智能生成内容的保护

在 20 世纪 70 年代，美国国会成立了版权新技术利用委员会 (CONTU)，并于 1978 年发布了一份报告。报告指出，计算机、摄像机和其他机器应该被用作缺乏自主性的工具，而不是因为机器人能够生成内容并将计算机作为作者，就改变以人为中心的版权保护标准，正如一些学者所认为的。此外，就目前的科技发展而言，计算机生成的内容仍然受到人类操作的控制，计算机对内容的产出没有影响，因此在版权法中计算机不能成为作者。

但在 1986 年，一份发表在《电子信息时代的知识产权》的报告指出，美国技术评估局认为 CONTU 的主张有争议。因为它认为计算机生成的内容不能仅作为工具使用，计算机仅与人类互动是值得怀疑的，因为否认计算机的自主性将否认它在创造过程中没有价值。我们应该讨论，在这种情况下，计算机在人类操控下生成内容，计算机是否可以成为人类作者的共同作者并受到法律保护，我们不能仅仅因为其工具性就避免讨论这个问题。如何保护计算机生成的内容以及如何平衡程序员和计算机软件用户的利益，需要认真研究和讨论。

目前，上述关于计算机生成内容的版权问题在美国国会的相关立法中还没有得到实践。然而，在司法领域，美国关于计算机相关版权的实践大约分为以下几个阶段：在 19 世纪，美国联邦最高法院在 Trade Mark Cases 案件中的判决中指出，版权是对智力劳动权的保护制度，作者创造的内容应当是原创的，内容要体现作者的精神意志，作品是由人产出，并应该是确定作者身份的前提条件；(2) 在 20 世纪，随着相机等摄影工具开始进入人们的生活，美国法律学者讨论它们所产生的作品是否具有版权，例如在 *Jewelers Circular Publishing Co. v. Keystone Publishing Co* 案件中，法院认为版权作品不必反映作者的个性或者人格，相应的人格理论是民法版权法的一个概念，其目的也是为了更好地将所有照片纳入版权法的保护范围，但同时，这一阶段削弱了作者身份的必要性要求；在 20 世纪末，在 *Williams Electronics Inc. v. Artic International, Inc.* 案中，法院认为通过操作视频游戏生成的视频可以产生版权，游戏

程序的创造者不是游戏程序的使用者。设计者在游戏中构想了所有可能展现的情景和图像及其原因，而用户只是展示了游戏的某个特定过程，因此，通过操作游戏生成的演示视频的版权应属于程序员。

从上述案例分析来看，美国法院在理解版权和计算机生成内容方面的判决不断变化。特别是在人类为计算机准备了足够的数据材料之后，计算机生成内容权利的归属更多的是集中在程序的程序员身上，而不是计算机的用户。

## 3.2 国际组织和一些发达国家对人工智能生成内容版权保护的告知

总结来看，目前全球大多数国家尚未详细讨论人工智能生成内容的版权问题，而已经认识并试图从法律和政策角度应对这一问题的大多数国家是普通法国家。其中，以英国、南非和新西兰为代表的一些英联邦国家，承认人工智能生成内容的可著作权性，并将此类作品的作者视为人工智能的操作者；虽然澳大利亚在法律中没有明确规定，但对一些人工智能生成内容的可著作权性有一定程度的政策认可。

此外，在普通法国家的司法实践中存在不同的认知标准，如英国和南非基本接受了纯粹由人工智能生成内容的著作权归属问题，而澳大利亚对受版权保护作品的创作主体资格标准显得更为严格，对人工智能操作者的要求不仅限于普通的操作行为，还要求他们对创作作品的物质形式有一定程度的控制，纯粹的人工智能生成内容在澳大利亚法律下是不具有著作权的。尽管美国长期以来一直在讨论这个问题，并相信只要符合版权法的标准，这类作品通常可以被授予版权，但在司法实践中对人工智能创作作品的认定和归属仍然缺乏一致的做法，程序员和操作者目前都可能成为人工智能创作作品的作者。

日本以外的民法国家很少尝试从法律和政策角度应对人工智能内容生成的问题。由世界知识产权组织（WIPO）代表的国际组织并没有试图以公约的形式统一由人工智能内容创作引发的版权问题。从1982年的建议和1991年的初步讨论来看，WIPO对计算机生成内容概念的理解也在变化，显而易见的是，计算机不再被视为辅助创作的技术手段，而是开始被考虑为创作的“主体”。从1991年WIPO最终撤回计算机生成内容版权问题讨论的趋势来看，至少在目前，国际组织已经将人工智能生成内容的版权问题委托给各国的国家法律，并选择采取不干预的态度。

因此，目前关于机器人和人工智能的法律讨论仍处于探索阶段，其根本原因是人工智能

技术的发展有一定的局限性，仍然处于人类的控制之下。因此，可以理解其他国家立法者对人工智能进入法律领域持谨慎态度，但另一方面，越来越多的国家和组织开始认真考虑人工智能的立法可能性，当前的法律体系并不妨碍我们对人工智能进行研究和讨论，而可以借鉴其他国家的先进经验，使我们的国家在从机器的法律态度到人工智能的法律研究方面更快发展。

## 第2节 信息后时代人工智能条件下新闻生产方式的转变

### 1. 相关概念的定义

#### 1.1 什么是后信息时代？

美国思想家尼古拉斯·卡尔在他的著作《浅薄》中提出，信息时代互联网的普及和快速发展已经将人类带入海德格尔所说的“从深刻到浅薄”模式，这种基于信息的思维方式因其片段化、简短和混乱的特征，导致了忙碌生存和记忆外包现象的出现。媒介工具变化带来的人类思维方式的变化，包括机器在内，最终导致人们质疑“是人主宰机器，还是机器主宰人？”然而，就在我们担忧互联网技术的日新月异和快速发展可能引起人类习惯的改变和复杂的伦理冲突时，时代的车轮不会拒绝技术的胜利。1946年，世界上第一台由美军定制的电子计算机ENIAC的问世，标志着人类从工业时代进入信息时代。在接下来的半个世纪里，正如摩尔定律所预测的那样，计算机处理器和内存的性能分别增加了两千万倍和十亿倍。互联网以不可阻挡的优势进入了人类生活的各个方面，信息爆炸式增长，各种新媒体的发展和应用使公众能够通过网络获取和传输无限信息。我们已经进入了信息后时代。

根据麻省理工学院媒体实验室的尼古拉斯·尼格罗庞特，信息后时代是一个“个性化时代”，媒体提供的信息不再是广泛撒网的盲目大众传播，追求的不是大而全，使所有人获得同质化服务，而是使传播的受众从大众过渡到越来越小的群体，信息变得极其个性化，直至仅针对个体。机器对人的了解程度可与人与人之间的默契相媲美，以便信息传播可以针对不同的人群和不同的个体。同时，我们关注个体的不同需求，提供个性化和多样化的定制信息服务。这意味着信息后时代的信息传播基于对个人习惯、状态和心理需求的深入了解。这种精准识别和信息传递需要依靠最先进的智能技术来实现。我们已从信息时代进入到信息后时代，包括今日头条和网易新闻客户端等算法引导的新闻分发渠道推送状态就是验证这一点，从观众的点击、页面停留时间、评论状态分析公众的新闻偏好，以实现有针对性的新闻信息定制推送。尽管今天的推送算法仍处于粗放阶段，但随着技术的发展和完善，机器推送将继续变得精准和细致。

## 1.2 人工智能的定义

根据麻省理工学院的帕特里克·亨利·温斯顿，人工智能就是研究计算机怎样才能做到以前只有人类才能做的智能工作。约翰·麦卡锡认为，人工智能是关于使机器行为看起来好像它们是有智能的。这与“中国房间”实验的逻辑一致，该实验认为不可能实现高度发达的人工智能为基本立场，即使人工智能看起来非常智能，但机器本身并不理解它在做什么，因此不能被判断为具有智能。这被认为是对强人工智能的遗漏。《经济学家》商业理论主编吴辰提出，“AI是点亮未来科技创新的力量”，并且提炼和分析大数据的技术和过程即是人工智能。东京大学副教授松尾豊认为人工智能是软件和程序的一部分，但与其他程序相比，它被设计用来根据实际情况自动改变行为，并做出适当和细致的动作。加州大学伯克利分校AI系统中心创始人斯图尔特·拉塞尔和计算机科学家彼得·诺维格将人工智能的定义主要集中在20世纪50年代和60年代建立的方法上，这些方法允许计算机以人产生智能的方式开发机器智能。硅谷风险投资家吴军在拉塞尔和诺维格的概念基础上提出，人工智能分为1.0和2.0时代，狭义的人工智能指的是传统的人工智能方法，广义的人工智能指的是允许计算机通过图灵测试的方法，包括数据驱动方法和超级计算方法。

在1950年，人工智能之父图灵在《心智》杂志上发表了一篇题为《计算机与智能》的文章，提出了一种测试机器智能的方法，如果一个人无法判断与他背后通信的是机器人还是人类，那么这台机器就具有人工智能。图灵测试的判断方法至今仍被用作衡量机器是否具有强人工智能的标准之一。20世纪50年代的计算机科学家们认为，如果计算机能够做到包括语音识别、机器翻译、自动摘要或文本撰写、战胜人类国际象棋冠军以及自动回答问题等事情，那么这些计算机可以被视为具有人工智能。半个世纪后，今天的计算机已经做到了这些事情，但今天的科学界并不认为计算机拥有人工智能。这表明随着科技的不断发展，对人工智能的定义和判断也在随着时代的进步而不断推进。人工智能的研究领域也随着时代的发展而不断扩展。尽管研究人员对于人工智能的确切定义并不完全一致，但他们都认同AI是使用计算机技术精炼和分析大数据，最终服务于人类的最佳方式。作者认为人工智能(Artificial Intelligence)是一个包含了许多技术和系统的总称，包括强和弱两个判断维度，强人工智能强调机器能够像人一样思考，并做出相应的行为反馈，能够自由通过图灵测试。而弱人工智能则指的是没有自主意识的机器智能。技术的最终方向是，是否要实现突破到强人工智能，或者是停留在弱人工智能的阶段，剥离机器能力与意识的关系，作为人类生存的优秀辅助替代工具，使人类生活智能化和技术化，这是历史演进和发展的唯一旅程。在21世纪，随着

后信息时代的到来，各种算法和处理器的巨大改进，大数据的蓬勃发展，技术的进步使得人工智能再次成为议程。正如摩尔定律所预测的那样，计算机数据处理能力不断提高，算法也在不断进步。海量的信息流使得每个网民都深陷其中，任何人都可以成为信息的接收者、发布者和反馈者，不再受时间和空间的限制。谷歌作为最早抓住互联网发展早期机遇的公司之一，搭上了技术更新迭代的快车，其搜索引擎的发展收集了全网几乎所有的数据而不是少量数据，其领地不断增加和扩大。这使他们能够在特定的数据分析中，实现样本与整体的等同可能性。2005年，谷歌利用全球互联网上的数以万亿计的语料库，开发了机器翻译系统，并在NIST主持的评估和交流中，以巨大优势击败了全世界的所有机器翻译团队，成为该领域的先锋。因此，2005年被认为是大数据的元年。搜索引擎的出现和谷歌翻译质量的突破，显著提高了机器的自然语言处理能力，这是人工智能发展的必要技术基础。

人工智能虽然表面上只是计算机科学研究的一个分支，但实际上它包括了统计学、哲学、神经学、语言学、心理学等自然科学、社会科学和技术科学三方面的交叉学科，人工智能的发展相应地影响了许多领域。当现有领域遇到大数据和机器智能时，势必会产生变化和新的形态。对于新闻传播领域来说，研究传统媒体新闻生产框架如何不可避免地受到冲击和重塑，有必要对新闻生产方法的现状有一个清晰的认识。

### 1.3 新闻生产方法的定义

新闻生产指的是将新的和变化的事实加工处理成为新闻作品的过程。新闻生产方法是用来完成这一完整过程的具体方法。传统意义上的新闻生产必须经过严格的过程顺序和严格的控制。传统新闻生产依附于传统新闻机构，这些作为权威和官方新闻出版者的机构对新闻和信息的流通拥有绝对的控制权。当以报纸、期刊、广播和电视为主导的传统媒体从事大规模信息生产和传播时，它们牢牢占据了生产、处理、制作新闻信息的港口，这种对新闻信息的绝对控制使它们拥有独特的话语权。在传统的自上而下的媒体环境中，新闻框架的议程设置过程通常由权威媒体主导，它作为政府宣传的前线号角，线性地向观众输入信息内容。由于信息传播的自上而下特性，专业记者的手工生产内容模式产生了大量丰富的新闻内容，渠道的单一性使新闻信息到达率高，传播效率和传播效果在一定程度上可以量化。传统媒体在舆论引导和控制方面拥有绝对的优势。

随着技术工具的创新，它带来了新的生产传播工具的变化。互联网的技术创新带来的实时速度、信息内容的海量性、检索的便捷性、交流过程的互动性、通信的全球范围以及通信形式的多媒体性，打破了传统媒体的垄断局面。新媒体时代媒体融合加强和互联网的快速

发展,使得新闻来源越来越复杂,网络信息输出渠道不再单一,呈现出多元化发展趋势。新媒体的生产和编辑信息具有全天候和全覆盖的特点,信息源和信息总量更为复杂,每个人都可以成为新闻发布者,参与上传、制作和与新闻的实时互动,在这个时候新闻生产朝着专业记者生产模式和用户生产的二合一状态发展。新闻信息和公众舆论变化迅速,热点频发,在某些情况下,技术工具的发展分散并降低了信息到达率。由报纸、杂志、广播和电视主导的传统媒体三巨头,以门户网站和新闻网站为主的在线媒体,以及以微信、微博等为主的社交媒体,共同构成了一个新的新闻生产领域。在互联网时代,依托科技的力量,网民可以通过互联网表达自己的观点,形成强大的协同效应。舆论领域日益增强的力量与技术的发展密切相关,这提供了一个平等参与的平台,弥补了传统新闻生产过程中缺乏反馈和互动的不足,满足了公民政治参与交流的不同层次的需求,对政治问题和社会问题的关注与表达反映了公众参与民主化进程的愿景。当时间的车轮进入后信息时代,又有了机器人新闻,新闻生产的主体再次扩大。

## 2. 新闻生产形式的变化

2016年3月,由谷歌DeepMind团队开发的人工智能程序“AlphaGo”以4:1的总比分战胜了世界围棋冠军、韩国九段棋手李世石,引起了整个网络的关注,并掀起了一波有关人机大战话题的公众舆论高潮。人工智能或对人类生存的威胁成为时下热议的话题。AlphaGo通过学习大量的人类棋谱,通过不同神经网络“大脑”的协作来提升棋艺,这与谷歌图片搜索引擎识别图片的结构机制相符,是一种渐进的、基于梯度的策略改进方法。与AlphaGo人工智能程序类似的,还有基于语音识别的服务,如iPhone的语音助手Siri和微软小冰,包括搜索引擎、机器翻译等在内的自然语言理解技术,数据挖掘与知识发现能力技术,以及基于视觉的计算机视觉与图像处理服务。这场人机围棋大战是对人类的及时警醒,我们已经生活在被弱人工智能环绕的环境中。中国工程院院士李德毅认为,这场人机大战的意义在于充分说明了数据已经成为产业的核心资源,信息产业的算法已经从一个以数据+处理器为核心的计算架构,过渡到了一个以算法+数据+处理器为核心的架构,这将开启不确定性人工智能的时代。

换句话说,在处理器技术和算法达到一定水平的情况下,任何人拥有的数据越大,处理数据的方式越先进,它就拥有越高的价值。AlphaGo的庞大数据库和超级推演速度,凸显了AI围棋选手神经网络算法的聪明和熟练。然而,机器学习的超快速度持续地颠覆着现有的学习架构模式。2017年1月,DeepMind在德国慕尼黑的DLD会议宣布推出新版本的AlphaGo

Zero, 同年 10 月, AlphaGo Zero 采用全新的学习方法, 在仅仅三天的训练后, 它以 100:0 的比分击败了之前与李世石对战的上一代 AlphaGo。新一代的 AlphaGo 完全抛弃了人类的棋谱, 通过深度学习方法进行了数百万局的自我对弈, 打出了真正的人工智能围棋。这一战役探索了深度学习技术的极限, 表明一个完全自我学习的系统是完全可能的, 并且能够出色地完成数据分析和选择任务。一旦这些人工智能技术应用到围棋之外的其他地方和领域, 它们可以最大化整合优势, 实现最优策略。

人工智能算法和技术的进步使得新闻的制作在资源收集、信息整合、内容表达和信息分发等方面与传统新闻生产方式大相径庭。

## 2.1 资源收集：以全集数据作为样本进行收集

在《通信革命》一书中, 费德里克·威廉姆斯提到, 在过去两个世纪的媒体发展中, 通信技术的最重要发展是早期技术成就的结合, 而不是突然出现的新发明。在通信的技术革命中, 连接结构、记忆方式和内容复制发生了根本性的变化。早期的远程有线技术如摄影、磁带录音和印刷使得语音、文本和图像能够直接通过不同的渠道传输到遥远的地方。现代电子计算机和电子高速网络是数字信号普遍性的直接体现。然而, 它们获取信息的能力极其有限, 没有有效、便捷和廉价的方式来保留数据和信息。进入新世纪后, 谷歌推出了“眼镜”移动设备, 如果将摄像头对准目标物体拍照, 设备可以自动检索并识别目标物体的相关信息。换句话说, 在一个集中化且高效更新的巨大云更新库中, 信息获取的成本大大降低, 效率也大大提高。尽管设备的成本很高, 但它也给我们提供了一个思路: 移动数据可以被挖掘和利用。数据主义认为, 宇宙是由数据流构成的, 任何现象或实体的价值在于它对数据处理的贡献。

在过去, 记者基于能够被人直接识别和处理的语音、文本、图像、视频等数据来收集和接触资料。然而, 随着语音识别技术、视觉识别技术等人工智能技术的兴起, 一旦人工智能被应用到传感器生成的数据上, 就会增添新的数据收集方式和渠道。全集数据作为样本数据量已经不再是一个空洞的概念, 学会并善用它极为重要。如果用传统的手动收集方式进行问卷调查, 收集所需的研究数据, 不仅工作量巨大、耗时, 而且无法获取全集样本数据。如果问卷内容设置中存在某种倾向性, 它将影响到最终的调查结果。但是, 机器学习可以帮助新闻从业者从庞大的数据仓库和语料库中得到他们想要的全集数据, 监督学习可以迅速发现已知输入信息与已知输出信息之间的联系。今天, 当你将传感器与数据挖掘和深度学习结合时, 你会发现许多新的新闻报道角度和内容。比如, 手机中的 GPS 直接收集每个人的移动数据, 截至 2017 年 6 月, 中国移动、中国联通和中国电信的移动用户总数分别是 8.67 亿、

2.69 亿和 2.3 亿。如果每人平均每天通讯一次，每天将产生大约 13 亿条位置数据。这些数据记录了大量人们的日常出行和旅游历史轨迹，在记者看来，这虽然是一种无法直接识别的非传统信息资源，但在机器看来，它却是个人生活习惯的完整而详细的记录。通过基于位置和社交网站数据分析这些全集数据，可以为人们推送个性化的交通信息服务，并根据观众的点击等反馈数据，进一步了解城市族群的基本习惯。这是一个积极的互动和修正过程。

通过微软的 T-Drive 系统和 T-Finder 系统，通过收集和分析出租车司机的全集数据轨迹，实现了智能驾驶路线推荐。基站连接数据提供的数据库非常庞大，一旦转换为与用户密切相关的人类活动轨迹数据文本，它就能模仿城市社会生活的日常印象，并对城市的交通状况有一个全局的了解。及时预防和治疗城市群的交通问题和疾病传播。在过去的两年中，非常流行的可穿戴设备和手机自带的运动 APP 都是用户运动的记录，如果公众对这些应用程序开放更多的授权，软件就可以记录更详细的人类日常活动。通过分析多个不同的数据集，可以确定用户在每个时间段的生活状态。例如，通过收集某个年龄段用户的全集的时间段、持续时间和睡眠时间，可以分析该年龄段的生活习惯的规律性和健康状况，并进行健康风险监测。公民自主参与数据记录，最终使用全集数据形成新闻报道内容的这个过程，是一种新的生产参与模式。许多传感器新闻使用在某一时间段内的全集报告作为样本来报道研究。2014 年 8 月，美国公民新闻网站率先尝试，通过对 NASA 卫星感应系统收集的 2 万平方英里沿海侦察图像和数据的统计和分析，团队的报道团队推导出路易斯安那州海岸的历史和现实环境画面，以视觉新闻的形式揭示了当地的严重海岸萎缩和土壤侵蚀。

当信息以语音、文本、图像和视频的形式呈现时，它是可以被直接识别和收集的数据。语音识别是人机交互的入口。语音识别与自然语言理解的结合，将为人们与周围设备高效交互创造更便捷有效的语音界面。在人工智能技术的快速发展下，信息传播的连接结构已经得到了质的提升。早在 2014 年底，百度首席科学家吴恩达及其团队发布了第一代端到端深层语音识别系统 Deep Speech，即便在嘈杂的环境中也能拥有 94% 的语音识别准确率。在 2018 年两会报告中，人民日报新媒体中心首次引入了人工智能语音识别系统。这意味着语音识别系统已经进入新闻领域，在重大会议的视频直播中，该系统可以使音频与中英文文字翻译实时无缝连接，这极大地提高了新闻发布的速度和传播的效率。语音识别系统在第十三届全国人民代表大会第一次会议、两会新闻发布会、部长通道、代表通道、委员通道等的现场视频直播中被采用，有效提高了文本记录的准确率和发布速度，赢得了用户的广泛认可。语音识别系统是人工智能技术发展的第一个门槛。谷歌的 AlphaGo 最初是从进入语音识别领域开始的，使用深度学习方法将语音转换为文本，使机器能够翻译超过 90 种语言，只有在其成功

后，它才逐渐推进其在视觉识别系统和其他领域的研究。在人工智能领域，图像、文字、声音等的整合、分析、描述和处理是计算机模式信息处理的基本内容。依托互联网网络的海量数据和强大算法，可以实现包括声音、文本、视频等数据在内的全集的综合分析和提取。

In becoming a reality. 当社会逐步走向数据化的过程时，制度设计的最终结果可能会与设计系统的基本出发点发生矛盾。在真实环境中可能隐藏着一些未被发现的不平等问题，比如“贫富固化”的争议问题，使用完整数据集作为样本量可以揭示事实和趋势，有助于把握宏观趋势以及从时间和空间层面收集信息，更大范围地了解阶级固化何时开始，以及受到哪些社会力量和社会结构的影响。从全数据集中收集的数据量适合用于调查性新闻、环境新闻、公民参与新闻和无人机新闻报道。当样本数据量足够大时，数据可以帮助记者发现由于信息有限或其他原因难以进行的新闻问题，简化手工收集的效率，同时做到科学和客观。

## 2.2 信息整合：人工智能技术助力实现整合优化

目前，新闻媒体涉及并且对人工智能技术产生较大影响的四项技术包括：机器学习、自然语言生成技术、语音识别系统技术、视觉识别系统技术。机器学习是人工智能的核心内容之一，主要使用归纳与综合方法来获取所需要的数据信息。机器学习可以帮助记者基于已知输入，访问云数据平台获取大规模相关信息，并从中提取结论。信息论创始人香农认为，信息是用来消除随机不确定性的东西。在信道传播过程中，总会有噪声和干扰。互信息指的是两个随机变量之间的相关过程，这一过程最小化了信息中存在的不确定性。使用“互信息”来衡量数据之间的相关性，可以提高新闻信息的准确性。通过机器学习，人们能够提高解决未知问题的效率和准确性，排除信道中的无用干扰，实现对信息相关性的判断，并帮助尽快输出搜索控制知识和启发式知识。

新闻工作的核心任务是筛选正确的信息并纠正内容。机器算法可以提供尽可能多的信息，获取新闻的相关性和多维度信息，以便横向和纵向地进行新闻研究。机器学习同时不断地验证随机变量，并使用交叉验证来分析和比较数据集，以消除不确定性并实现交叉熵接近零，从而获得最接近新闻真相的数据材料。在实践中，如果能量化两件事情之间的相关性并对大数据文本进行统计，就会发现一些表面上看似无关的信息实际上高度相关。使用机器算法实现重复验证的过程，不仅可以避免“黑天鹅效应”并在理论上获得最优的整合内容，还可以使新闻更加精准，成为真实历史材料的流动。国际调查记者联盟组织超过 400 名记者参与分析高达 2.6TB 数据量的“巴拿马文件行业阴影报告”，在人工智能技术的帮助下，记者们对各种电子邮件、文件及其他原始数据的深入分析变得快速而清晰。

自然语言生成技术的设计目的是使计算机能够像人类一样表达和写作,最终生成符合语言规范的高质量文本内容。文本的自动生成通常会经过文本规划阶段和表层形态阶段。机器人撰写新闻是自然语言技术发展到现在一定阶段的产物。在新闻写作中直接应用机器人除了,新闻记者编辑还可以使用自然语言生成技术来自动产生文本,以此来协助新闻分析,更快地整合有效信息,使新闻工作变得简单方便。

语音识别系统能够实时地将语音转换成文本,同时也能将文本转换为语音。2018年“两会”上语音识别系统的引入便是一个例证。这项技术极大地方便了新闻视频的编辑和新闻声音的录入。例如,它可以自动为新闻栏目匹配字幕翻译,无论是外语频道的外语识别还是本地频道方言的识别,都能快速实现字幕录入,降低生产成本,减轻编辑的负担。在现场视频直播、新闻会议等方面,它节省了时间和精力,使得文本信息和声音信息能够同步整合。未来的配音工作可以使用机器来自动化完成。

视觉系统识别技术是让机器代替人眼来测量和判断输入的目标。视觉识别技术中的图像识别技术使得机器能够根据图像中包含的颜色、物体、纹理、背景乃至空间构成关系等元素的特征来获取有效信息。谷歌的计算机视觉专家 Tobias Weyand 和 James Philbin,以及像 Ilya Kostrikov 这样的工程师开发并训练了图像识别机器 PlaNet,它能够通过分析任何照片上的像素来判断照片拍摄地点。斯坦福大学的研究人员开发的算法模型“gaydar”可以通过照片来判断人们的性取向。这意味着面对一个不熟悉的图像,新闻编辑无法用肉眼识别从照片中获得的结构化和非结构化信息,而机器算法能够判断它并将其转化为可识别的数据信息。这不仅有助于为新闻原始数据自动适配合适的图像材料,还足以帮助记者编辑发散性思维,找到人眼无法判断的隐藏信息,寻找与新闻主题相关的内容。算法还可以对模糊照片进行高清场景还原。语音识别系统和视觉识别技术的结合将使自动为新闻视频和电视节目找到匹配的声效和图像成为可能,这为视觉新闻制作提供了极大的便利。

在整合新闻的过程中,理解权威新闻媒体对某些新闻禁用词和谨慎使用词的规定是至关重要的。以新华社为代表的权威新闻媒体通常会根据社会环境和网络文化的持续发展来设定新闻用词的边界和禁忌。例如,在新华社发布的《新华社新闻信息报道中禁用词和慎用词(2016年7月修订)》中提到,提到“一带一路”战略时不使用,而是使用“一带一路”倡议,不使用“践行‘八荣八耻’”这一表述,而应使用“践行社会主义荣辱观”等等。如果算法及时输入新闻禁用词和谨慎用词标签,并且不断跟进权威媒体的更新,实时跟进信息的更新,加入到数据库中,那么机器算法就可以协助新闻从业者审核新闻内容,以确保不会出现原则性错误。通过这种方式,算法不仅可以自动化地检查和筛选文本内容,还可以为编

辑提供一种保障机制，确保新闻报道符合现行的政策和语言使用规范。这对于提高新闻发布的准确性和遵守媒体伦理具有重要意义。

## 2.3 内容表达：机器人新闻写作的崛起

美国学者尼尔·博斯曼提出，技术变革对生态是非加法性的。新媒体的出现并不是增加或减少某些东西，而是改变一切。直到机器算法的崛起之前，新闻信息的生产是专业从业者或由主动性推动的受众根据特定模式和框架构建社会的活动。然而，在2015年9月，腾讯财经的自动化新闻写作机器人“梦想编写者”与新华社的写作机器人“快笔小新”联合打破了这一层次。新闻生产所需的专业壁垒被打开，算法内容生成被纳入新闻生产的方式之一。写作机器人的存在，就是人工智能进入新闻领域后，在内容表达层面出现的新生态。

在2017年4月24日，新华社的特派机器人记者佳佳采访了美国《连线》杂志创始人凯文·凯利。这是世界上首次有机器人作为记者对人类进行采访和交流。与国外相比，中国在智能机器人进入新闻领域的渗透还比较短暂，但发展速度非常快。机器人新闻写作是算法生成新闻内容新时代的集中体现。在人工智能技术的推动下，传统媒体的生产纬度和经度已经被打破。在媒体领域表现较好的机器人新闻写作是由新华社带领的，诞生了大量的机器人记者编辑。传统媒体包括新华社的“快笔小新”，浙江日报报业集团的媒体传播平台中的媒体写作机器人，第一财经的“DT草稿王”，南方都市报的“小南”，广州日报的“阿通”，钱江晚报和微软小冰的“小冰”机器人记者，腾讯开发的在线媒体“梦想编写者”机器人，百度的“度秘评论”，今日头条的“小明”机器人，中国地震台网的写作机器人等。如果人工智能进入新闻行业，语音领域最重要的技术是语音识别系统。

然后，文字的领域是自然语言处理系统。在机器人新闻写作中使用的，且仍在进步的技术是自然语言处理技术。新闻稿是遵循一定语言规范和词序结构的文本综合体，机器人新闻写作，如同人类写作一样，需要经过收集数据 - 分析数据 - 规划文本 - 陈述实施的过程。对机器人来说这并不难。每一步都对应着一项技术。当机器人读取大量人类编辑模板中的新闻稿时，它可以利用自然语言技术、关键词生成技术、摘要提取技术、文本信息综合技术、重复文本技术等，将相应的新闻相关内容嵌入到特定模板中。北京大学计算机科学系。

万晓君，科技研究所的研究人员认为，“新闻是自然语言处理最标准化的文本。”事实证明，在今天的体育赛事分析、娱乐快讯、新闻综述、财经报道等领域，机器人完成新闻稿的质量又快又好。以2016年5月29日的新闻项“四川绵阳安县发生4.3级地震”为例，官方通知中仅6秒的编辑时间显示机器人写作迅速、准确且强悍。早在2014年3月17日洛杉

级发生 4.4 级地震时，人工智能机器人仅用了三分钟就编写出了成功的新闻稿。目前，由机器人记者代表的人工智能可以写一些风格化的信息，但更高级别的智能交互尚未实现。机器人可以从海量数据中快速分析出所需的部分信息，根据固定模板进行填充和发布，完全自动化新闻的制作。摒弃了过去必须有类参与痕迹的环节，这意味着这时的生产主体从人变为了机器。

在 2017 年的两会上，作为一款集成了语音识别、图像识别和情感识别技术的现象级产品，新华社不仅以类似记者的形式在两会现场进行采访，而且还在 30 多个媒体、专业杂志和海外媒体平台广泛传播了两会的相关报道。2017 年 11 月 23 日，新华社 iSi 机器人赢得了“王选新闻科技”一等奖。iSi 新闻机器人赢得了中国媒体行业唯一的科技奖项，这表明机器人进入新闻领域后，新技术的优势已经得到了业界的一致肯定，而且随着技术的创新，人工智能将在未来的新闻产业中发挥更大、更重要的作用。

不用说，新闻机器人作为可以自动收集数据和编写内容的计算机程序，在速度和数量上具有无法比拟的人类优势。在财经新闻、体育事件等规律性新闻报道中，它们可以迅速捕捉最新内容进行发布，提高“信息”新闻的数量和质量。四川九寨沟发生 7.0 级地震时，中国地震台网的写作机器人在短短 25 秒内就报道了这条新闻。今日头条的机器人“张小明”能在 2 秒内完成稿件，并上传至媒体进行发布，在里约奥运会期间 24 小时监控事件状态，平均每天撰写 30 篇赛事报道。强大的算法计算能力使得机器人能够在大量复杂的数据摘要中游刃有余。综合性、具体性、效率性和准确性。这样记者就不再为编写快速新闻感到疲惫，不必花费大量时间进行数据收集和验证，而是可以集中精力思考和撰写更深入的新闻报道，降低劳动成本。

记者的采访、编辑和写作是对社会环境进行重构的动态和建设性过程，这些过程中包含了不可或缺的伦理理想、道德情操和同理心。然而，当今的机器人仅仅重复简单且大量的数据处理操作，并不具备强人工智能的能力，因此无法掌握需要深度思考、社会观察和构建的新闻类型。因此，短期内，写作机器人不会导致记者和编辑大规模失业。在写作机器人席卷新闻行业的浪潮下，新闻编辑应该与机器智能相互补充和协调，使得人类编辑能够从大量信息中抽身，保持客观和清醒，从而发现有新闻价值的话题，制作出更具可读性的新闻。

## 2.4 内容分发：算法推送机制带来的个性化定制体验

算法推送机制是技术对新闻行业影响最大的领域之一。在传统媒体时代，主要依靠报纸的流通被动接收信息，互联网仍然延续了新闻编辑主动选择并推送新闻的操作方式。进入移

动互联网时代后，社交媒体分发渠道开始抢夺新闻产品流量，微博和微信凭借其庞大的用户基础和用户粘性，能够控制新闻信息的流向。还有一些新闻聚合媒体依靠先进的算法捕捉整个网络信息平台的内容，并将其推送给读者。据埃利·帕里泽尔（Eli Pariser）在《过滤泡泡：互联网没有告诉你的事》一书中所述，用户接收到的信息被机器算法所操纵。我们可能不会收到我们真正需要的新闻信息。而且我们是愿意全盘交出控制权的。

对于许多移动社交平台，包括腾讯QQ和新浪微博，用户每一次点击、签到、分享、发布图片和心情都是个体不断建立用户属性模型的过程。当数据量达到一定水平时，算法将根据个人场景维度和社会关系维度等多维数据判断用户的兴趣和爱好，社交网络在你之前就会知道你需要什么、想看什么。这对个性化新闻推送来说意义重大。在人工智能技术的驱动下，越来越多的终端和设备连接到网络，实现人机之间的自然对话和智能交互，新的数据平台不断涌现。

新闻聚合媒体完全依靠算法来推送新闻信息，用户的每一次点击都能为算法提供数据。算法可以根据用户的历史搜索、订阅、页面停留时间、阅读开启频率、点赞分享、评论等数据进行分析，从而模拟用户画像的流动。如果用户资源信息在多个资源数据源之间共享，例如一个新闻聚合平台同时了解您的购物历史和旅行历史，平台就可以结合个人兴趣爱好来实现更精准的点对点推送和本地化推送。随着时间的积累，媒体平台对用户数据和信息的把握越来越清晰和深刻，平台媒体可以获取点击率和巨大的流量，抢占商业利益。对于用户而言，算法推送机制的不断创新一方面可以让他们在获取想了解的信息方面得到最全面的答案，节省寻找类似新闻信息的成本和精力，这是一个动态的流动过程，比如今天你可能想了解房地产政策的新内容，点击和转发后，第二天算法仍会向你推荐相关信息，及时全面地满足用户的信息需求；另一方面，它可能导致用户沉迷于算法编织的熟悉世界，并形成对算法的依赖[3]。

然而，算法并不会自主筛选推送内容是否健康和积极。当传统新闻编辑被程序员编写的算法取代时，可能导致质量低下、错误频出的新闻内容在互联网上流行。2016年，Facebook解雇了所有人类编辑，让机器算法接管了原本由专业记者负责的热门话题的选择和排名工作。结果，在这项决策生效后的三天内，算法就推送了一则关于梅金·凯利是民主党总统候选人希拉里·克林顿的秘密支持者的假新闻。尽管后来澄清说这条假新闻背后有一个新的审核团队，但它仍然是在缺乏专业记者的大错误背景下发布的。对于有经验的媒体专业人士来说，判断假新闻并将其排除在人们可接受的内容之外是常有的事。但对于算法来说，它们无法仅通过文章的点击量和新闻来源来判断是否为假新闻。尽管对于大多数人来说，算法推送机制

带来的个性化新闻定制体验是令人兴奋的。但从长远来看，用户只能看到他们想看的，屏蔽掉他们不想看的一切，这将使人们变得更加极端和固执，加剧刻板印象。以今日头条和FACEBOOK为首的聚合媒体已将工具性和利益追求推向极致。

在2017年9月11日，《人民日报》发表了一篇关于“不要让算法困在‘信息茧房’”的讨论文章，提出算法会为人们创造一个信息茧房，并且算法不能决定一切。2017年10月5日，《人民日报》又发表了一篇批评今日头条、一点资讯和UC头条等代表的智能新闻客户端的文章，人民日报的评论文章对新闻聚合媒体敲响了警钟。文章“不能让算法决定内容”提出算法应该回归服务内容的角色，通过“人工推荐+智能筛选”的组合方式推送新闻和相关信息，而不仅仅是看点击和转发的数量，这加剧了低俗标题党的泛滥。《警惕算法创新向反面发展》强调，即使是最好的渠道也应该有守门人，盲目沉浸在算法创造的只有一个观点的世界中，这不仅是偏执的，而且可能会造成侵权。

算法并非是为了抓住边缘而存在的。获取新闻信息的目的是为了消除不确定性，理解和认识我们周围的世界。盲目依赖算法推送机制带来的新闻信息，可能会导致信息过量或内容冗余，算法机制带来的“信息茧房”加剧了自我封闭，创造了一个不真实的“模仿环境”，使我们离真相越来越远。一个好的新闻出口应该提供正确的价值导向和多样化的信息，而不是以盈利为导向的方法。然而，另一方面，算法推送机制反映了以受众为中心的理念，旨在为观众提供最佳的用户体验。无论技术创新如何发展，公众始终需要保持独立思考和判断的能力，并提高媒介素养。善用算法，真正做到我们所需要的。

### **3. 新闻产业结构的转变**

#### **3.1 新闻生产方式的变化**

##### **3.1.1 线性、有序的新闻生产方法的解构**

在过去，新闻生产过程有一套统一的固定标准。从线索收集的早期阶段——采访——写作——编辑——报道组织和策划都是由专业媒体和专业记者共同完成的，新闻生产过程也是各种权力博弈、新闻背景和新闻场的决定过程，这些都决定了新闻生产的最终面貌。随着互联网的繁荣和快速发展，互联网用户数量和网站总数持续上升，截至2016年12月，中国互联网用户规模达到7.31亿，手机互联网用户规模达到6.95亿，互联网渗透率达到了53.2%，

同时，中国“.CN”域名总数达到2061万个，占中国域名总数的48.7%。这种“零门槛”的进入条件值得注意的是，在世界范围内，包括中国在内的无处不在的互联设备、机器和系统之间的非结构化数据量的爆炸性增长。手机和物联网的低消耗数据存储技术不断成熟，这不仅允许新闻编辑直接从大数据网络中检索他们需要的任何数据，而且还可以利用机器算法的高效率完成数据的选择。新闻生产设备的变化直接导致了观众获取信息习惯的变化。在今天信息爆炸的世界里，每个人都可能成为潜在的传播者和潜在的观众，以及各种信息设备的市场化，包括手机、平板电脑、电脑等，满足了信息获取设备的支持，所以新闻信息生产者的身份不再局限于某个单位和某个媒体，而是属于每一个个体。每个人都可以成为事件的目击者或第一人称，所以生产者身份的多样性允许他们随时随地拿起手机记录当下。没有必要等待专业媒体搬运重装备匆忙赶往现场。

同时，机器算法的不断完善使得网络的智能推送能够以更精确地优化信息的到达率，这种便捷且迅速的人工智能互动体验随着个人兴趣和关注的差异而持续变化。最终实现了个性化的信息服务，满足不同人群的个人信息需求。换句话说，传统媒体发布的新闻大而全，媒体向你展示你所看到的。现在是你想看什么，媒体就给你什么。在新技术下，公众不仅有选择权，还有生产和传播的权利。新闻生产的主体和新闻生产的对象不仅仅是新闻人，还包括普通人甚至机器人，特别是以今日头条上的自媒体渠道和网络上的社交媒体为代表，它们对新闻的处理实现了各主体的动态互动。过去，人们可能只关注传统媒体发布的最热门新闻消息，但数据算法改变了一切。有序的线性传播方式已经不再适用。无论是今日的世界新闻还是历史上的今天，只要与个人的实际需求相关，它就是观众的关注点。这种跨越时空的价值实现已经成为人们新闻的滋养。原有新闻传播内容的范围已得到扩展。技术的发展影响了人们的阅读方式，新闻的控制、评论和设置也影响了人们对内容的接收和接受。从2015年腾讯发布由机器人“梦想编写者”编写的第一条财经新闻“8月CPI同比上涨2.0%，创12个月新高”，到2015年11月7日，新华社正式启动机器人写作项目“快笔小新”，并且快笔小新被用于经济报道和财经报道以及事件报道中，这表明人工智能已经进入了需要深度学习的信息生成和传播领域，媒体组织重视技术发展并积极将最新技术引入专业领域，以最大化效率和效益。这不仅是媒体组织面对日益复杂的生活环境的积极响应，也是技术进步趋势下的必然选择。“落后就要挨打”，只有跟上时代潮流，顺势而为，我们才能立于不败之地。2016年初，在湖北省的两会上名为“云”的机器人记者进行了全媒体传播。人工智能的交互和探索从未停止。作为传统新闻形式，机器人在编写新闻时不仅快速，而且客观准确。如果新闻发布不再仅仅通过记者的第一手采访，那么包括新闻Feed在内的一些新闻形式的写

作方式的改变，将不可避免地强迫媒体组织进行内部人员设置的转变。传统的、线性的、有序的新闻生产方式不可避免地会解体。传统媒体的新闻生产过程已经发生变化，它不再是唯一的新闻传播渠道，但是传统媒体的权威性需要得到维护，在这种情况下，应坚持传统媒体的严谨、认真、可靠和负责任的特点。在重新处理新闻材料时，我们关注观众反馈并保持应有的专业性[4]。

### 3.1.2 对话式和社会化新闻生产模式的出现

随着传统的、线性的、有序的新闻生产模式的解散，分散的、对话式的和社会化的新闻生产模式已经出现，新闻生产的主体和对象已经多样化。

自从2016年百度“度秘”机器人评论员推出后，它便同步实时直播了奥运篮球赛事，旨在创造个性化的观赛体验。在比赛过程中，度秘不仅会使用文本、图片和语音的结合为用户解读比赛过程，而且还会将全网实时精选的粉丝评论推送到用户端，点对点的独家服务完全满足用户对大型体育赛事的个性化需求。这种实时直播的对话式新闻生产方式颠覆了过去的生产模式，是对单点场景下新闻的私人定制。这种基于用户反馈的新闻互动生产是新技术的集中体现。换句话说，基于度秘，未来的现场体育赛事可以实现私人定制评论。无论是受众范围狭窄的小众比赛，还是存在语言障碍的国际比赛，度秘所代表的新应用都可以使用。

实现实时翻译加实时评论，并且伴随有表情和语调词汇的使用，这种对话式的方式提供你想要获取的新闻信息，提供了极大的便利，从某种程度上溶解了信息传播的壁垒，真正满足了不同层次公众的需求。

通过研究今日头条“小明机器人”的发布新闻发现，“小明机器人”主要关注体育赛事和财经新闻。根据北京大学教授万小军的观点，机器写作涉及两种方式，一种是原创稿件，另一种是稿件的二次创作。原创稿件主要使用现有的结构化数据，直接生成新的稿件，比如“小明机器人”的赛事实时评论新闻写作。原创稿件利用自然语言生成技术，从结构化数据出发表达和生成自然语言陈述。对于体育赛事，只需对接到组委会的数据库，借助数据进行一些数据分析、文档规划和表述实施，就可以自动生成简短的赛事信息；对于长篇赛事报道也是同样的做法，机器算法只需要收集现场文本，然后利用机器学习算法整理出精彩描述和关键词，最终生成赛事报道。除了原创新闻外，另一种是通过二次创作生成的稿件。使用自动摘要技术，从现有的文本新闻信息中提取所需材料，或者基于同一事件的多条新闻报道自动生成长篇事件摘要。这是通过新闻采集 - 段落划分 - 段落排序 - 段落选择和合并的过

程来实现的。

以2018年2月26日“小明机器人”编写的一场体育赛事报道为例，有九篇足球报道。从标题判断，“西甲第25轮塞维利亚2:5床单军团没有面子”，“法甲第27轮巴黎圣日耳曼完虐马赛胜利”以及“意甲第26轮罗马0:2遗憾失利”这三篇文章中的“没有面子”、“完虐”、“抢劫”和“遗憾”等词语，已经不再是简单的机械性比分展示，而是有了个性化媒体报道属性的评价。从报道形式和具体内容来看，机器人在事件报道上与人类的写作相比并无太大差别，都使用图文结合的方式来阐述赛事过程，基本上可以替代人类编辑在体育短消息新闻写作上的工作。

在“西甲第25轮塞维利亚2:5床单军团无面子”的新闻中，对阵的是马德里竞技。最终，马德里竞技客场以5:2战胜塞维利亚，格里兹曼、科克、迭戈·科斯塔为客队贡献进球，而萨拉维亚为主队挽回了一些颜面。在这篇报道的第一段中，清晰地指出了“何时、何地、发生了什么、谁参与了、为什么”的五个要素。这意味着机器体育新闻写作在引言部分做得到位，符合新闻报道的基本要求，是一篇完整的新闻报道。然而，通读全文后，还是在细节上存在问题，在上述文章中，详细描述了双方进球的过程。0-1、0-2、0-3、0-4至1-5进球段落有详细的描述，但最终比分是2-5，在塞维利亚最后一个进球的部分并没有描述。机器人写道1-5部分已结束。“1-5! 很快，裁判吹响了比赛结束的哨声。”这里存在数据错误和文本描述错误，是“缺失了新闻信息”。而且这个详细错误到3月份也没有被更正。这种常规信息的缺失在传统媒体中很少发生。传统媒体对稿件的错误有零容忍度。所有在报纸上发表的稿件都会经过严格细致的人工审查。然而，网络媒体发布的新闻内容常常会有打字错误或句子不通顺，机器人新闻写作缺乏严格细致的检查，难以避免出错。在某些情况下，互联网时代的在线稿件对质量的要求并没有传统媒体时代那么谨慎。然而，不可否认的是，无论是体育赛事还是财经信息，都依赖大量数据，面对高度结构化的人群，有着强烈的垂直需求。但传统新闻编辑面对每周大量的体育比赛，难以做到面面俱到。机器人的出现弥补了这一短板，无论是热门比赛还是小众赛事，它提供相同的新闻信息，这满足了一些群体的个性化需求。新闻信息从面向大众到面向个人的转变已经实现。必须知道，机器写作之所以能在体育新闻和财经新闻领域大放异彩，是因为这两个领域往往拥有详尽的高质量信息数据库和图表材料，机器写作在清洗文本材料和文本数据方面面临的困难较小。但在其他领域，中文信息的质量更加混乱，清洗困难，编写模板代码的工作量巨大。这两者都使得机器新闻写作在其他领域难以施展。

在社交媒体的时代，确实每个人都可以成为数据的携带者和传播者，任何人都能通过手

机制作内容。新闻生产不再是媒体人的专利，无论是用户提供新闻线索、提供新闻背景材料，还是直接参与新闻的制作和传播，都是独立新闻生产的体现方式。社交新闻生产的出现极大地压缩了新闻制作的时间，并且我们无法知道何时会触发新一轮的舆论热点。

传统的新闻守门人通常由专业人士处理，编辑对新闻来源可靠性的判断至关重要。机器辅助判断假新闻并非不可能，至少可以帮助人们更有效地完成守门工作。《华盛顿邮报》在2012年开发的写作机器人“真相告诉者”能够记录新闻报道过程中的文本、声音和图像，并快速将其具体内容与防伪数据库进行比对，一旦发现假新闻就立即发出警报。卡内基梅隆大学的客座教授迪安·波默劳（Dean Pomerleau）构想建立一个庞大的假新闻数据库，通过不断收集新数据来训练算法识别假新闻，随着假新闻数据库的扩大，识别假新闻的能力也将提高。从源头开始，需要使用云计算数据，实时记录虚假来源并不断更新，作为辅助判断的重要参考。新闻内容中提到的具体单位和机构名称是否准确和权威，都是机器算法可以完成的事情。输入到训练样本中的数据越多，其准确性就会越高，机器人的新闻检查行为将通过关键词过滤实现，大大提高了机器人协助人类识别假新闻的可能性。因此，当今的新闻生产需要人机合作才能达到最佳解决方案。

## 3.2 双向流动的资本与技术

### 3.2.1 资本主导资源流动

从2015年5月到2017年12月，国家共发布了12项与人工智能相关的政策和计划。2015年5月的《中国制造2025》提出加快智能制造发展，聚焦智能装备和产品的发展，并推动智能生产过程。2016年3月，人工智能被纳入《国民经济和社会发展第十三个五年规划纲要》。2016年4月，提出了机器人产业发展规划，强调到2020年自主品牌工业机器人的产出发展和技术指标。2016年5月提出的《互联网+》人工智能三年行动实施计划，在前一个计划的基础上进行了细化。

到2018年，国家将聚焦人工智能层面需要注意的关键点，即建立人工智能产业体系，应用和技术部分领先。2016年7月，重点放在基于大数据分析的人工智能发展上。2017年3月，“人工智能”首次被写入国家政府工作报告。2017年7月，国务院印发《新一代人工智能发展规划》的通知中确立了三步走战略，人工智能首次上升为国家战略层面。强调到2025年，部分技术和应用应达到世界领先水平，核心产业规模超过4000亿元，相关产业规

模超过 5 万亿元；到 2030 年，人工智能的理论、技术和应用将达到世界领先水平，核心产业规模超过 1 万亿元，带动相关产业规模超过 10 万亿元。

从规划中可以看出，国家政策和规划的推动是对人工智能发展的一个渐进式、循序渐进的布局。在厚度中有精细，精细中又有深度的要求。这表明，人工智能技术的蓬勃发展已经成为中国构建智能社会和智慧城市的基础，也是中国跻身创新型国家、走在世界高科技发展前沿的重要措施。2017 年 12 月，《关于促进新一代人工智能产业发展的政策》相继发布，包括广东、浙江、安徽、贵州、江西、山东、陕西、吉林、黑龙江、上海、北京、重庆等多个省市都推出了发展人工智能产业的相关政策。从资金到政策支持，提出了大量的红利，足以看出国家和地方政府对人工智能产业化的重视程度。在整体环境的资本支持和红利驱动下，高科技人才的流动和涌入势必将加速，这有利于形成人工智能企业集群和产业发展高地，推动地区科技发展和技术落地。

除了政府持续出台政策发挥政府的引导作用之外，以百度、阿里、腾讯为首的 BAT 大型科技巨头也早已进入该领域实现产业布局。早在 2013 年，百度就广泛招募人工智能人才，并基于人工智能不断完善其现有核心业务，包括搜索和信息流。阿里巴巴推出了机器人项目 ET，并在不同场景环境下演化出 ET 大脑，比如能够通过交叉分析温度、风力和气压等指数实现环境预警的 ET 环境大脑，这在推动环境新闻领域的发展中起到了重要作用。腾讯设立了 AI 实验室，专注于计算机视觉、语音识别、自然语言处理和机器学习等领域。与此同时，腾讯约有一千名技术人员从事人工智能研究，与百度的规模相当。科技巨头在人工智能领域的强力布局加剧了竞争，并同时提高了技术发展和进步的速度。政府政策 + 企业产业布局 + 高质量人才的涌入引导了相关资源的流动。腾讯的用户基础庞大，其 QQ 月活跃用户达到 8.5 亿，微信月活跃用户达到 9.63 亿，这意味着它拥有源源不断的海量数据资源，一方面，庞大的数据库有利于新闻生产的精确数据采集，实现样本的全面收集，另一方面，庞大的用户基础更有利于算法准确描绘用户的个人兴趣习惯实现更好的垂直服务。资本和政策使得数据能够在最大程度上被收集和集中，高质量人才的培养和引入能够发挥最大的协同效应以促进技术和算法的更新，而场景落地是最终的要求。就像使用传感器进行新闻采集，使用机器人进行消息写作，使用算法推荐进行智能推送。这也与 2017 年 12 月发布的新一代人工智能产业发展三年行动计划的希望不谋而合。技能的产业化和集成应用是重点，需要推动人工智能与实体经济的深度融合。

### 3.2.2 技术改变市场行为

在 2017 年，借助人工智能技术，百度收集了大量用户数据并分析用户搜索模式，向用户发送目标推荐内容，实现了第二季度净利润 44.14 亿元人民币，同比增长 82.9%。利润的大幅增长，反过来又促使百度加强使用人工智能技术优化搜索引擎和智能推荐。在通过市场机制将技术创新成果转化为资本的过程中，最终实现技术创新成果的市场化，不仅有助于提高企业的创新能力和竞争力，使企业能够持续进行新的研发，实现创新的可持续性，还能振兴市场，为整个市场行为带来变革。在写作机器人出现之前，传统媒体无法想象，不仅新闻写作的工作会在一定程度上被替代，而且他们还将与机器合作共同编写稿件。但事实不言自明。2015 年，中国地震台网发展出写作机器人后，直接连接到各个平台，无需记者和编辑自动编写并完成地震信息的发布，然后通过新闻客户端、微博、微信等渠道为公众提供最新的地震报告。传统和新媒体都引入了机器人写作，如南都的小南、腾讯的梦想家写手、新华社的快笔小新、《华盛顿邮报》的真相告诉者、《纽约时报》的开花、《卫报》的 open001 等，这是技术发展给市场带来的最直观变化[5]。

科学技术研究院与新闻媒体进行深度合作与研发，为媒体提供可以快速写作的机器人。在新技术介入后，媒体获得了直观的效益，自然会增加对新技术的资本投入和技术引进，为传统生产模式注入新的动力和新的方式。正如主要媒体使用机器人写作一样，快速实现人工智能产品货币化的公司将率先获得增长机会，并且当成果被应用时，他们可以在获得商业利益和经济利益的同时，再次吸引资本流入，实现正向循环，加速技术发展。北京大学计算机科学与技术研究院与南方都市报合作推出了“小南”写作机器人，与今日头条合作推出了“小明 bot”，与广州日报合作推出了“阿通”机器人。作为高校的科研机构，通过人工智能技术的研究，获得资本，然后通过资本投入和产品输出实现行动者与利益相关者的一致性。技术、人才、资本、市场和环境在高科技产业发展中不可或缺。媒体写作机器人的应用是技术资本化的重要结果和体现。大数据+算法的分发模式使得技术与市场融合后长尾信息的传播价值达到顶峰。例如，对于春节新闻信息，通过对手机地理位置数据和阅读数据的挖掘，算法可以分析潜在用户并有针对性地推送地区性春节火车信息。南方都市报的小南写作机器人就对春节火车票进行了专业的写作和推送。除了技术资本化的实现，人工智能技术对大数据的清洗和处理能力也让我们看到了新闻采集过程中的新动力，在后信息时代，无论是互联网和社交媒体生成的数据，政府部门发布的数据，还是物联网传感器捕获的数据，或者个人和组织保留和传播的数据等，都可能成为新闻源，数据的开放性和共享特性一方面促进了技

术的更新和信息的互动。另一方面，隐私泄露的持续发生也迫使用户在互联网上自觉自我审查，屏蔽某些敏感语言和涉及私人信息的部分。技术的发展吸引了资本注入，同时也带来了市场行为 and 用户理念的变化。

## **4. 人工智能技术对新闻生产的影响是有益的**

### **4.1 更科学、客观、快速地获取多维度新闻信息，提高了新闻的准确性**

信息的本质在于它必须致力于消除读者的不确定性。尤其是在这个信息过载的时空中，没有任何新闻能保持长期的新鲜感。信息过载最常见的现象是谣言和传闻的源源不断，而在来得及辨别真伪之前，相关的新闻信息就迅速掠过水面，不留痕迹。在重大事件中官方权威声音的缺乏，不可避免会导致观众对信息的不确定性增加。这一矛盾导致我们偏离了信息的本质。因此，包括文本、图片、视频和直播在内的生产形式的灵活组合和应用，不仅使新闻更具可读性，消除了不确定性，而且还对应了媒体融合的意义轨迹。

当观众独立使用搜索引擎寻找答案时，任何输入的文本、图片或行为或情感的表达都可以成为机器量化的工具，数据化使一切都可以被总结和解构，从而找到最符合个人理解信息本质的方法和方式。因此，通过图像、声音、文本和数据的整合，机器算法捕捉到的具有相关性和因果性的数据可以将被忽视的分散认知结合成一个相对完整的画面。先进的人工智能技术可以帮助编辑和记者将繁重的内容转化为互动的视觉信息产品，新闻生成过程变得更快更准确，能够降低劳动力成本并提高市场效率。可见，人工智能与大数据技术能够促进新闻生产的全媒体化，整个生产系统更加开放、个性化和多样化。

互联网技术的迅猛发展带来了信息量的大幅增加，这在很大程度上满足了人们对信息的需求。面对海量的信息，用户很难迅速获得他们想要的那部分信息，因此信息到达率和使用率并没有得到提高。虽然直接搜索新闻信息的做法是基于个人兴趣，但有时候观众并不容易通过搜索关键词获取他们真正想要的信息。此时，机器算法的出现向人们发送智能推送，帮助人们选择他们想要看的和应该看的内容。最早采用个性化推荐技术的是 20 世纪 90 年代的亚马逊，它使用协同过滤算法将个性化推荐推向前台。无论是基于内容的过滤还是协同过滤，都是基于对观众偏好的挖掘，FIT 中国网络和《经济学人》在这方面做得更好。FIT 中文直接使用超链接推送相关、及时、个性化的新闻，根据用户阅读时间的长短，并且其算法推荐

的新闻内容不仅有正面评价，还有截然相反的观点评价，并根据用户的反应不断调整市场策略。这为解决信息冗余和同质化问题提供了思路。算法提供了个性化的新闻信息服务，但并不排斥提供相反的观点。这可以使得点对点的个性化新闻在科学和客观的路线上满足观众的个性化定制服务。无论观众是否同意不同的观点，但至少让他们知道不同的意见是存在的。人们对信息的渴望永远不是单一和固化的[6]。

此外，通过准确预测观众的偏好，可以实现更精准的广告投放，以最小的投资获得最大的输出效果，从而获得市场利益。例如，在阅读技术新闻时，广告商可以根据对观众的点对点数据分析，投放与技术产品相关的广告。这种方法极大地增加了潜在用户转化为目标用户的可能性。

## 4.2 新闻策划中的“意外”参数可以增强用户对新闻媒体的粘性

好奇心理学是人类活动最基本的三大需求之一，它是人类的一种天生特征。人类有对新奇事物的需求，根据个人对新奇度的需求水平进行新闻策划，因此，如果在算法参数中设置“意外”参数，从数据分析中推导出个人对内容的厌倦值，量化用户的情绪表达，同时利用不规则的情绪测试问卷，掌握用户的情绪趋势图，避免接触过多同质化信息带来的厌倦感，这可以帮助跳出信息茧房。这样用户始终保持对新闻信息的新鲜感，同时也增强了用户对新闻网站平台或新闻客户端的粘性。例如，利用迈克尔·科顿的 KAI 量表，测量公众对新事物的接受程度。

通过与不同数据源平台分享数据以了解观众的日常生活，智能设备的普及和传感技术的发展使人们的每一举动都记录在云数据中，寻找两件事之间的相关性可以帮助找出观众是否有新奇需求。例如，如果一个人在外就餐时频繁光顾不同的餐厅，可以推断此人有对新奇事物的需求。如果一个读者在社交平台或新闻聚合平台上频繁在不同的频道和栏目之间切换，阅读新闻信息，也可以推测他有新奇需求。然后对于有新奇需求的人，可以推送他们以前未见过的新闻类型和新闻栏目。通过持续的测试来实现最佳效果。技术的发展影响了人们的阅读方式，新闻的控制、评论和设置也影响了人们对内容的接收和接受。通过精准地了解 and 预测用户的行为和需求，新闻媒体可以提供更加个性化和吸引人的内容，这不仅可以增加用户对新闻产品的兴趣和参与度，还可以促进用户之间的互动和讨论，形成更强的社区粘性。

避免基本归因错误，从宏观角度把握抽象社会力量和社会结构在新闻事件中的作用。在

评估他人行为时，即使人们拥有足够的证据支持，他们通常倾向于寻找个人原因，最终总是指向特定的个人，高估了个人或内在因素的影响力。相比于寻找环境或外部原因，抽象的社会力量和社会结构的角色往往被忽视。记者需要能够通过表面迅速把握实质，发现更有意义的事实。这是机器单独无法完成的，但这是专业的媒体专业人士在机器的帮助下可以做到的。

《纽约时报写作指南》已经规定，应避免在职位称呼前使用性别区分，比如“女性”经理和“男性”保姆，以免形成一种全因某一性别所导致的现象，这不仅有职业基础上的性别歧视之嫌，也不利于维护新闻的客观性。这种主导解释不仅是记者和编辑的错误，也使社会形成了对某一性别群体的错误刻板印象。在社交媒体上有关女性驾驶员驾车事故的新闻络绎不绝，甚至“女司机”这一词汇已成为固定术语，放大了性别在道路交通事故中的作用，并忽视了其他客观因素。当听到有关驾车时发生的奇异新闻时，人们首先会反应可能是女司机。其余因素则被忽略了。新闻媒体在选择新闻要素时存在一定程度的故意强调和排除。通过机器算法，新技术可以用来使用穷尽法对报告的道路交通事故进行统计和分析，从而整合实际情况，并从宏观角度把握社会环境甚至媒体本身报道中的问题和倾向。只有避免基本归因错误，并放弃为了找到噱头而舍弃新闻的客观公正元素，我们才能从表面看到实质，避免主观假设，发现新闻的真相。

### **4.3 避免陷入峰终定律，让观众在面对新闻宣传和公关轰炸时能保持客观和清醒的判断**

2002年诺贝尔经济学奖得主丹尼尔·卡尼曼邀请受试者参加冷水实验，得出了人的身体中存在着经验自我和叙事自我两种自我。经验自我指的是人们在每个瞬间的意识，而叙事自我则通常只记得事件的高潮和最终结果。在叙事自我中，对整个经历的记忆和价值等同于事件的高潮和结束时的感受之和，然后确定平均值。人类是自主实体，经常操纵叙事自我所经历的故事和记忆。叙事自我不关心时间的长度和细节，只关注事情的高潮和结果。

例如，如果媒体想要掩盖真相或让社会事件的热度迅速降温，他们通常会选择删除帖子和评论来达到舆论控制的目的，或者用无关信息来淹没关键信息。对于新闻工作者和宣传者来说，如果观众是由经验自我的体验所引导，那么宣传控制的效果将大大降低。因为观众会记住每一个时刻的感受。但是如果观众受到峰终定律的支配，他们会有意识地回顾整个故事，很快忘记这里的不愉快细节，并迅速为下一个娱乐狂欢撤离。

批评的结果是显著的。人工智能依赖于庞大的数据集，基于云计算所拥有的巨大数据库

的机器算法可以在任何时候调动关于事件的任何细节。然后，观众在做出判断时，可以选择在任何时候调动所有相关信息。如果公众恰好是新闻事件的参与者，并留下了言语的痕迹，这一次的意识痕迹将会更加准确。对于观众来说，避免陷入峰终定律，避免让叙事自我为自己做决定，有助于摆脱媒体精心制造的信息陷阱，从而避免形成对其周围世界的虚幻认知。人的记忆是有限的，让人们在某些时刻远离以叙事自我为准则，转而以经验自我为基础，借助算法数据实现细节和感受的回溯，以便在面对新闻事件时拥有更加专业的判断和影响力，不再轻易被公关和宣传手段所左右。

## **5. 不良影响**

### **5.1 在算法上形成依赖，导致信息时代的知识悖论**

在信息过剩的时代，人们输入的信息越多，获取有效知识就越困难。先进技术不仅带来了信息传播速度的增加，还促进了真实信息与误导信息的传播。机器算法对内容的偏好和深度定制迎合了个人思维和兴趣，使得人们对兴趣之外的内容毫无收获，从而陷入信息茧房的循环。这种主动选择和被动创造相互构成的“后真相”世界，使得观众离真相越来越远。与基于新闻事实的个人思考相比，人们往往根据主观假设重塑甚至扭曲新闻事实，客观事实对公众观点的影响小于人们自己的主观感受和个人信念。再加上算法的推荐过滤机制，人们看到同质化信息和相同类型信息的概率大大提高。

例如，媒体框架议程对“扶老被讹”现象进行了一系列长期报道。自2006年的彭宇案以来，几乎每一起“扶起倒地老人”都成为热点新闻案例，“扶老被讹”似乎已经成为新闻中的常态，而“老人”标签被污名化。深圳大学法学院的应飞虎教授从2004年到2015年10月收集了媒体报道的149起“扶老”案例，其中32起被确认为真正的碰瓷，84起是老人设局的。最终有80%的真相被澄清。在此基础上，新闻聚合平台推送来自多方的相关内容，两个阶段的叠加之下，会导致人们对道德衰退感到恐慌，认为老人真的变坏了，认为这种不良现象和法院的不公处理已经成为常态，形成了对周围世界宏观环境的错误感知。形成了关于老年群体的刻板印象。这种刻板印象还加剧了人们对法治的不信任和安全感的丧失。如果不及时有效地回应，这种模仿环境的刻板印象和负面效应将是巨大和长久的。

### **5.2 语音克隆技术使伪造新闻来源变得更加容易，也**

## 使得判断新闻真伪变得更加困难

卡内基梅隆大学语言技术研究所开发的 Festvox 和巴黎公司 Candyvoice 开发的语音克隆应用具有相同的功能，两者都可以通过输入个人声音到软件中进行复制和重构。之后，你想让软件说的任何话都将基本上与你自己的声音相同。然后 Adobe 推出了一款产品，允许用户提供 20 分钟他人的录音来复制任何人的声音。加拿大蒙特利尔的技术公司 Lyrebird 开发的语音克隆应用将时间缩短到只需 1 分钟的录音，就能完美复制他人的声音，并合成你想听的任何音频内容。在 2018 年 3 月，百度再次刷新了语音克隆技术领域的记录，借助人工智能下的神经网络技术，只需一秒钟就可以复制任何人的声音。而这些语音克隆技术在欺骗语音识别设备方面有 95% 的成功率。

情感是人类共有的，语音克隆能够复制声音的情感特征。如果将语音克隆技术应用于新闻领域，个性化新闻的定制将再次添砖加瓦，这样的声音定制服务可以使人们选择任何他们喜欢的声音来阅读新闻或播放内容，并且未来可能会有专业的点对点私人语音播放频道。个性化定制服务最大程度地满足观众需求。同时，记者在新闻采访中经常使用语音录音机来保存当前的语音采访内容，录音内容也是新闻写作的可靠原材料之一。但是，当技术泛滥时，采访的录音和受访者提供的音频证据是否还具有完整的可信度？这可能导致更多的新闻伪造和制作假新闻的成本显著降低。它使新闻谣言的滋生和传播变得更加容易和方便。揭穿谣言的成本和难度大大提高。

### 5.3 技术的误用侵犯了观众的隐私权

当人工智能捕获新闻生产所需的源数据时，它基于云计算技术，从全网的数据库平台捕获新闻素材。只要用户连接到互联网，就会一直生成行为数据，无论是聊天历史、拍照、分享评论、上下班出行等，都会实时生成大量数据。这些记录作为云数据的一部分，传输到云端和各种传感器上。一方面，人工智能利用自己巧妙的深度学习算法、强大的计算能力和庞大的数据库，对云数据进行极其快速的计算和挖掘分析，从而实现对个体演员在个性偏好等特征上的画像，以实现新闻信息的精准推送和广告的精准营销；另一方面，当用户享受人工智能技术带来的便利时，个人的综合数据资源也频繁地被庞大的机器算法公司和组织调用和外包。而我们对隐私调用的内容和范围一无所知。

根据腾讯社会研究中心和 DCCI 互联网数据中心联合发布的《网络隐私安全与在线欺诈研究分析报告（2017 年第一季度）》，96.6% 的安卓应用程序会获取用户隐私权限。在后信

息时代，个人变得透明。人们的各种信息数据已成为可以打包买卖的资源。在缺乏个人信息保护法的情况下，公众的个人信息数据的收集和管理缺乏使用标准和法律保障，应避免技术滥用侵犯观众的隐私权。

## 5.4 导致员工的结构性失业

自 2014 年 7 月以来，美联社已经让机器人记者每季度撰写 3000 篇新闻报道。目前，大型媒体中的大多数机器人记者用于财经和体育新闻。对于传统的体育记者和财经记者来说，这些只需要通过数据处理和解释的新闻信息包含了巨大的工作量。而人工智能机器人可以完美处理繁琐的数据分析操作，比人类记者更快、更好、更便宜、更准确。通过收集互联网上的巨大信息数据，可以最大限度地获取公众舆论，并且也可以为专业人士的舆论分析带来便利。在当今的数据新闻学中，这种方法尤为常见。

如果只是简单描述信息和数据，不加评论，那么在这种最基础的新闻写作编辑领域，机器人记者就拥有了一席之地。《纽约时报》使用 BLOOSOM 分析用户对新闻推送的偏好，《卫报》使用 OPEN001 直接完成从内容选择到排版编辑再到整合生成的一系列操作。《华盛顿邮报》将人工智能技术应用于新闻验证，整个从新闻编辑到验证再到发布的过程可以由人工智能应用程序完成。国内的《钱江晚报》聘请微软小冰机器人作为特约记者，完成“浙江 24 小时”栏目的新闻稿任务。这表明，如果新闻从业者从事的是简单的新闻工作，不寻求改变、提高业务能力和自我价值，那么他们可以被机器智能所取代。过去，新闻业是以领域的核心为基础的，如今正在发生变化，而且在未来其影响范围将会更大。在后信息时代的背景下，我们不再缺乏媒体，不再缺乏媒体人，不再缺乏设备，我们最缺乏的是内容，是与时俱进的媒体新思维。

## 5.5 数据集中在科技寡头手中，使得新闻成为资本集

### 团的附庸

事实上，目前从事人工智能相关内容研究的都是网络或资本巨头，它们在互联网时代持有大量数据，不断合并企业和技术，控制巨大的流量入口，并实现极权主义与技术的高度整合。如果新闻媒体需要调用数据资源进行新闻研究，它需要向数据所有者寻求数据支持。由百度代表的中国互联网巨头 BAT 引领了中国 AI 的发展，并聚集了大量关于互联网 AI、企业 AI、感知 AI 和自主 AI 的数据，这客观上产生了技术寡头垄断和数字鸿沟，加剧了数据

垄断的情况。

在媒体融合的背景下，网络资源虽然分散且丰富，但只有政府和财团寡头才会掌握顶层设计，从分发、整合、媒体到平台运营都受到各方的严格规范，资本的高度集中很容易使新闻媒体成为金融资本和垄断集团的附庸。虽然依靠科技力量使网民获得了平等参与的平台，满足了民众不同层次的政治参与性沟通，但面对自上而下的权力中心网络审查和政治干预，以及科技寡头的垄断，未来的传播环境和舆论生态是否乐观，仍有待观察。资源技术的高度集中很容易使新闻成为资本集团附属品的产物。

# 第3节 人工智能在新闻生产中的研究

## 1. 人工智能新闻生产概述

### 1.1 新闻生产模式的演变

新闻生产的演化过程主要根据生产主体的变化来划分。最初参与新闻生产的主体是专业记者，他们拥有话语权，这种模式下的新闻生产被称为“专业生产”。（1）随着技术的发展，互联网的出现为人们提供了自由表达的平台，每个人都变成了记者，每个人都有了一个话筒的时代来临了，新闻生产的主体不再是专业记者，接收新闻的观众有时变成了新闻生产者，具有双重身份，“用户生产”出现了。随着技术的进一步发展，计算机拥有了人类智能，可以像人类一样通过相关算法模板生产新闻，所以人工智能新闻生产的算法生成模式和专业生产、用户生产一样。这些是由于不同的生产主体而经历的三种新闻生产模式，在一定程度上，它们依赖于媒体技术的发展。

### 1.2 专业记者手工生产内容的模式

在计算机和互联网平台出现之前，专业新闻机构和新闻从业者垄断了新闻生产的权利，并掌握了大部分新闻资源和传播渠道。因此，在传统的新闻生产过程中，从新闻的采集到新闻信息的编辑、生产和播报，都是在新闻组织内部进行的。在新闻信息和产品传播给观众之前，新闻生产的过程对观众是封闭的，他们无法看到新闻生产的过程，更不用说参与新闻生产的过程，这使得新闻生产成为了专业记者的垄断性专有模式，我们称这种模式为专业记者生产内容的手工模式。

在手工模式下，新闻生产的创作主体是专注于特定话题的专业记者，他们向观众提供他们认为有价值 and 意义的新闻。这种模式是线性的、单向的沟通方式，无法进行相应的反馈，对应的是标志和印刷媒体的时代。

在人类出现初期的很长一段时间内，人类只能通过表情和动作等原始古老的方式传递信息。由于原始人信息交换方式的简单性，信息交换的复杂性和信息传播的范围都受到了严重限制，一些简单的信息只能面对面交换。随着生产方式的改变，原始人群的沟通面不断扩大，为了实现更好的信息交换和传播，口语的出现是不可避免的，它能够表达具体的事物。但同样地，口头传播也有保存不便、容易扭曲的缺点，一个能够将信息从人脑的记忆系统中分离

出来并能够保存下来的符号系统应运而生——文字。文字能够永久保存信息，并穿越地域限制进行沟通。同时，造纸术和印刷技术的发明使得文字的传播更加广泛，人类也因此进入了印刷媒体的时代。

造纸术允许相关信息被永久保留，而印刷则使信息能够大量生产，正式地，这两项技术的发明和特点导致了报纸的出现和纸介传播时代的开始。无论是在口头交流时代还是纸介传播时代，我们可以明确的一点是，传播的生产者都是人类，传播的信息也都是由人类产生的。纸介传播的一个特点是点对面的传递，新闻信息的生产者与新闻信息的接收者之间通常是一对多的关系，在通信过程中，信息生产者作为信息发送者的“点”，以纸质媒体报纸为中介，向这个“面”的观众群进行单向信息传输。

### 1.3 专业生产与用户生产共存的 Web 2.0 模式

在 Web 1.0 时代，捕获、发布和组织信息是各种互联网门户网站的核心角色，通过为用户提供新闻聚合平台来浏览信息，信息趋势从互联网传递给用户。进入 Web 2.0 时代，信息的传输不再仅仅是从互联网到用户，这时用户也可以利用 Web 平台来生产内容，并传递给其他人，以此来区分传统网站员工领导的内容生产模式，被定义为 Web 2.0。（1）以 2009 年微博元年为标志，Web 2.0 时代正式到来。与 Web 1.0 相比，Web 2.0 是一个方向性和结构性的变化，Web 2.0 时代的互联网将每个网民视为一个信息节点。人们可以理解，Web 2.0 的应用使得万维网从一个网站的集合转变为为用户提供网络应用的服务平台，通过信息共享使得数据资源更加丰富。

随着电子信息技术和计算机技术的发明及广泛应用，特别是博客和微博的出现，标志着 Web 2.0 时代的来临。技术的发展使得“人人都有麦克风”的全民记者时代到来，新闻生产的权利不再局限于专业新闻生产机构和新闻从业者，接收信息的公众具有双重身份，有时他们也进行新闻生产，即用户生产。纵观新闻传播的历史，可以发现每一次技术的发展都会触发新闻生产方式的变化。从印刷媒体时代到电子媒体，印刷技术创造了报纸，无线电技术创造了广播和电视，互联网技术使新媒体崛起。新闻生产模式也由过去的专业记者生产，变为现在用户也参与新闻生产。坚实的技术基础使得在这个世界上信息的传播越来越方便，曾经的地球村的猜想已经变为现实。

观察这两种新闻生产模式，尽管新闻生产的主体已经发生了变化，但在生产过程中除了人类智能之外，并没有其他智能介入。随着技术的不断发展，参与新闻生产的软件已被赋予人工智能，它可以完全独立于人类进行生产和创作。

## 1.4 算法生成内容与专业生产和用户生产齐头并进

在大数据时代,使用人工智能算法从无限的数据仓库中提取知识已经成为社会生存的基础。人类生活在社会系统复杂性呈指数级增长的时代,在一门被一些术语称为“社会物理学”的新科学的黎明时期。当我们通过移动设备进行各种活动时,它将不断自动分析所有领域。

传统记者通过密集且有时危险、昂贵的实践调查新闻事件,试图揭示新的事实和社会趋势,并且利用他们的叙述天赋、经验、价值观、创造力和直觉为观众转化这些事实。随着“社会物理学”新领域的出现,叙事——讲故事的艺术,也正在成为一种科学尝试,使用大量的语言学 and 自然语言研究的知识领域的人工智能算法。AI 算法可以在短时间内将事实转换成可读的故事,这就是人工智能新闻生产的诞生。AI 驱动的新闻生产基于两大技术支柱:一是自动从巨大的数据库中使用新的“社会物理学”概念提取新知识的计算机软件,二是自动将这些知识转换成可读内容的算法。

技术的发展使新闻生产的主体再次发生变化,之前无论是专业记者还是用户生产的新闻,其生产者都是有形的人类。而计算机技术、算法编程等使新闻生产的主体机械化、模拟化。随着技术的持续成熟,通过算法生成内容的机器人记者已经与专业记者和生产新闻内容的用户站在了同一阵线。

## 2. 人工智能新闻生产的概念与技术

### 2.1 人工智能新闻生产的概念

人工智能新闻生产的具体形式是新闻生产主体的仿效和机械化,其生产主体是依赖于计算机互联网平台的智能软件,因为它能像记者一样生产新闻,所以也被称为新闻写作机器人。因此,人工智能新闻生产表现为新闻写作机器人通过互联网平台收集大量信息和数据,通过大数据技术专业处理数据并获取有价值信息,然后利用自然语言处理技术,根据相关算法和模板,将获取的有价值信息转化为人类可以理解的语言,最终形成一篇完成的新闻报道。

### 2.2 人工智能新闻生产的技术

新闻媒体技术的发展必然带来新闻生产方式的改革与创新,在未来,大数据、云计算、人工智能等新技术将为人类开辟新的新闻生产方式。

### 2.2.1 大数据的基础

近年来，由于社会信息化和网络化的持续发展，大数据受到了人们的高度关注。“大数据”这一术语越来越多地被提及，并渗透到日常生活的许多领域，在许多行业中发挥着重要作用。

大数据是一个广泛的概念，尽管信息领域中有很多关于“大数据”的定义，但这些定义都不变地强调了一个“大”的词。然而，大数据不仅仅用来描述数据量的大小，其真正内涵应该包括处理数据的速度和能力。大数据发展的终极目标是人们能够从各种类型的海量数据中快速获取高价值信息，而没有价值或尚未发现其价值的大数据在某种意义上是资源的浪费。

目前，更统一的理解来自于互联网数据中心（IDC）的大数据定义，它包含了四个基本特征：体量（Volume）、多样性（Variety）、速度（Velocity）和价值（Value），即所谓的4V特征。

大数据的4V特征使得大数据与传统的数据概念有所不同。海量数据只是强调数据的数量非常大，而大数据与海量数据还是有区别的，大数据并不只是用来描述数据量大，最重要的是强调处理数据的技术，也就是说，要能够使用技术在复杂的大量数据中，用较少的时间对所需数据进行爬取和专业分析，并获取有价值的信息。

### 2.2.2 自然语言处理

自然语言处理，英文缩写为NLP（Natural Language Processing），是人工智能的一个子领域，可以简单理解为使机器能够理解中文、英文等人类语言的技术。例如，如果一台机器既懂中文又懂英文，它就可以充当两者之间的翻译。再比如，如果电视能理解观众的语言，那么观众就可以不用按钮，直接用说话来控制电视，选择他们喜欢的节目。自然语言处理的范围主要包括文本阅读、语音合成、语音识别、自动分词、词性标注、句法分析、自然语言生成、文本分类、信息检索、信息抽取、文本校对、问答回答系统、机器翻译、自动摘要和文本蕴涵等。

自然语言处理技术的探索可以追溯到1960年代。这项技术是计算机出现之后才存在的新技术，随着计算机的持续发展，相关人员在这方面的探索也取得了一系列进展。在过去的20年里，随着互联网技术的发展，自然语言处理技术也得到了极大的发展，有效地推动了网络核心能力的提升。比如，互联网的信息检索能力。起初，当人们在搜索引擎中输入关键

词时，他们得到的结果往往是大量与需求无关的信息。后来，这种现象变得越来越少，搜索到的信息基本上是与搜索者需要最贴切的信息。这种搜索质量的持续提升是自然语言处理技术不断改进的结果。

自然语言处理技术使得机器智能成为现实，帮助机器像人类一样理解和生成自然语言。这让机器人足够智能，再加上机械技术赋予的行动能力，使得机器人能够根据人的意愿做所有工作，包括替代人类冲锋陷阵。随着大数据的使用以及云计算提供的高性能计算能力，深度学习技术在语音和图像识别领域取得了显著进展，机器人的智能将继续提高并逐渐接近人类。类似机器人的机器人几乎可以做所有人类的工作，甚至理论上，它们可以完全将人类从劳动中解放出来。

### **3. 人工智能新闻生产的过程和特点**

#### **3.1 人工智能新闻生产的过程**

在人工智能新闻生产的过程中，肉眼无法看见的机器人坐在电脑前像人一样撰写新闻稿，这是一个依赖于计算机和互联网平台来收集数据，并使用智能算法和语言处理技术来处理有效数据，从而自动生成新闻稿的智能过程。

新闻生产机器人作为媒体领域的新事物，它与传统新闻生产者最大的不同在于新闻生产方法的差异。在构建新闻作品的过程中，传统新闻生产者之间形成了专业标准、流程规范和组织文化。在传统的新闻工厂，专业记者和编辑像工厂流水线上的工人一样，以日常周期推出新闻产品。在传统新闻生产中，对新闻产品的生产也有具体的操作指导方针。导语、标题和消息都有相应的写作规范，甚至新闻格式、叙事方法等都有相应规则可循。编辑部也有自己的操作规范，包括编辑会议、选题规划会议，甚至各种编辑规定和评估系统来约束编辑部的运作。然而，在人工智能新闻的生产中，由于大数据和自然语言处理技术的支持，整个生产过程可以在无需人类参与的情况下进行，而且可以将获取的有效新闻数据自动转换为人类能够理解的新闻文章。人工智能新闻生产过程主要是数据收集、数据处理、新闻写作和海关发行，具有高效率、强大的数据处理能力和及时响应紧急情况的优点。虽然新闻生产流程的名称看似相似，但仔细观察会发现人工智能新闻生产过程与传统新闻生产过程有很大的不同。

### 3.1.1 数据采集

在传统新闻生产中，新闻采访是新闻生产的开端，也是为新闻写作、新闻编辑及其他后续工作收集相关材料，为生产打下基础。在传统的新闻生产过程中，新闻采访要求记者亲自赴新闻事件发生地，贴近生活、贴近事件当事人，同时以高度的责任感和历史感去观察、倾听和思考。可以说，新闻采访在传统新闻生产过程中扮演着举足轻重的角色。搜集并分析新闻事实材料是记者为报道新闻而进行的一种专业活动，这是所有新闻工作的基础和前提，也是每个记者都应掌握的基本技能。

在人工智能新闻生产过程的第一步是数据采集，这不需要人类记者在现实生活中亲自观察、倾听和思考，一切变得智能化和自动化。在“互联网+”大数据时代，人类信息已被数字化传输到互联网平台。新闻写作机器人只需要连接到互联网平台，所有人类活动将被持续和动态地监控，只要发出命令，新闻写作机器人就可以自动收集数据库中的所有数据，其数量和时间跨度是无限的，所需时间也极短。

通过比较，我们可以发现人工智能技术在新闻生产中的应用减弱了人类记者的角色，并增强了智能生产在新闻生产行业中的主导地位[8]。

### 3.1.2 数据处理

人工智能新闻生产过程中的第二个步骤是数据处理，这对应于传统新闻生产过程中的新闻编辑，即新闻记者进行新闻采访后，必须立即筛选和编辑所获得的第一手信息，以提取有价值的信息。在传统新闻生产中，这一过程主要依赖于人类记者的专业能力，他们能够根据新闻的价值和特性，从新闻采访中获得的信息中选取能吸引观众的有用信息。

在人工智能新闻生产过程中，要从大量数据中提取有用信息，就需要使用大数据、云计算等技术。如前所述，海量数据只是强调数据量非常大，并且大数据与海量数据之间还存在区别，大数据不仅可以用来描述大量的数据，最重要的是强调处理数据的技术，即能够使用技术在复杂的大数据中，用更少的时间爬取和专业分析所需的数据，并获取有价值的信息。

在人工智能新闻生产中，通过敏感词汇表和垃圾词汇表的结合，过滤掉行业中存在的垃圾信息，同时，通过句法分析，有效捕捉相关有用信息，这样的双层保险方法确保了爬取信息的有效性。首先，系统通过定义结构化对数据库中的数据进行分类；然后根据用户所需的标准收集数据；最后，系统准确过滤类似内容和用户反馈来收集有效数据。在这个过程中，

不需要消耗任何人力，只需要一个命令，就可以直接抓取有效数据。

对于新闻生产来说，技术不仅仅是为了收集大量数据，更重要的是分析数据以获得有意义的信息，这是技术的最终目的。在这个过程中，数据处理的机械化和智能化节省了大量的人力和物力资源，且效率是人类所无法比拟的。

### 3.1.3 撰写新闻

在传统新闻生产过程中，记者的新闻写作是将采访信息转化为成果的重要步骤，也是检验记者写作水平的重要环节，同时也是记者与公众沟通的关键一步。记者通过新闻作品向观众传播他们想要表达的信息，尽管这是基于客观材料，但写作的主体是人，所以以这种方式传播的信息不可避免会带有人的主观性。

人工智能新闻生产基于大数据和自然语言处理技术，整个过程可以在没有人类记者的情况下进行，能够选取有效的数据，首先根据数据选择一个好的报道角度，然后自动选择适当的词语，最后根据新闻报道模板生成新闻。这一步骤在 AI 新闻生产过程中更注重脱离劳动和自动化生成新闻产品。

### 3.1.4 发布环节

在传统的新闻生产过程中，新闻产品在面向公众之前必须经过守门员的审核，传统新闻记者或专业组织是新闻产品编辑发布之前的守门者，守门者决定哪些信息被使用以及哪些信息可以传播给观众。

在人工智能新闻的生产中，前期的生产过程几乎没有人的参与，但在最终的编辑发布环节，作为守门者的人类记者是不可或缺的。因为虽然新闻机器人拥有在一定程度上学习人类智能的能力，但它们在道德伦理方面的判断能力并不与人类相同，因此守门者仍然是新闻机器人人工智能生产过程中不可或缺的环节。这也是传统记者未来发展的方向，即新闻信息的收集者逐渐聚焦于新闻生产的守门者。

以新华社的智能写作机器人“快笔小新”为例，它根据每个业务部门的需求定制发布模板，自动抓取数据并生成稿件，但最终稿件只有经过各业务部门的编辑和审核后才能发布。

## 3.2 人工智能新闻生产的特点

### 3.2.1 智能化生产技术

在传统新闻生产中，虽然已经有各种各样的技术用于新闻生产，甚至包括互联网技术，但在这些生产过程中并没有类似人类智能的出现，只能说这些技术依赖于人的操作来完成新闻作品的生产。可以说，在整个过程中，新闻生产主要是人工的，技术工具在这里只是被操作，完全没有智能性可言。

在人工智能新闻生产的过程中，因为它依赖的生产技术是大数据和自然语言处理技术，这两种技术能够实现无需人类参与的数据智能捕捉和语言文字的智能生成，使得新闻写作机器人具有类似于人的智能。这种智能化不仅体现在数据的自动化处理和信息的快速提取上，还体现在语言的生成和文章的构建上，它能够模拟人类的语言表达和写作风格，自动产出符合新闻标准的稿件。

这些智能生产技术能使信息来源变得更加广泛。由于人眼能看到的有限，很多时候人类记者在报道新闻时会忽视事物深处隐藏的众多因素，从而丢失了大量信息源。智能技术能够通过设备与人类相连，直接获取人体发出的信号，并将其转化为数据进行分析。另一方面，在互联网时代，网络上的数据也是极其庞大的，智能技术能够自动侦测网络数据，一旦数据发生波动，它能迅速捕获数据，并通过自然语言处理技术，把数据转换成人类可以阅读的新闻报道。

例如，智能技术可以实时监控地震数据、压力数据、温度数据、空气质量等人眼无法直接观测的数据。虽然之前的技术也能检测到这些数据，但现在的智能技术不仅实现了监控，还能根据这些数据进行分析、判断，甚至能像人类一样自动通知人类其结果。

### 3.2.2 模仿生产主体

在传统新闻生产过程中，新闻生产的主体，无论是职业记者还是用户，都是具体的人类。在人工智能新闻生产过程中，你看不到新闻写作机器人坐在电脑前敲击键盘进行新闻写作，因为新闻写作是一个智能技术，也就是说，在人工智能新闻生产过程中，计算机可以自动产出新闻内容，并实现对新闻生产主体的模仿。

在新闻行业中使用智能技术变得越来越成熟，人工智能可以自动捕获数据，使用模板生成内容，并开始应用于各种报道领域，模仿性的新闻写作机器人逐渐替代人类记者成为新的新闻生产主体和新闻写作者。新闻生产进入了一个新的发展阶段。

在人工智能新闻生产过程中，模仿性新闻写作机器人进行数据收集、数据处理以及新闻撰写，这些都是人类无法看到的过程。人类能看到的只是在很短的时间内，新闻写作机器人会展示已经写好的新闻文章，而人类需要做的就是进行检查和审阅。这一生产过程重新定义了传统记者的工作内容。

另一方面，不难发现，在整个人工智能新闻生产过程中，除了最后一步新闻发布需要人类记者的控制以外，前面的步骤都由智能机器人执行，这基本实现了生产主体的“机械化”。由于生产主体的机械化，它可以全天候收集数据、分析数据和生成新闻，新闻生产可以在任何时间和任何地点完全实现，这在一定程度上节省了人力，并极大地提高了新闻的时效性。

### 3.2.3 生产过程高效率

传统记者在报道金融、体育、灾害等新闻时，常常需要处理大量的新闻数据，而新闻写作机器人可以在几秒钟内撰写文章并产出新闻报道，这使得传统记者对这种速度感到惊叹。由于技术特性，新闻写作机器人已经在金融财经、体育赛事、突发灾害等模式化报道中得到了广泛的应用。

人工智能新闻生产基本上可以实现新闻同步，实时监控情况的动态发展，自动捕获相关数据，分析得到有效结果，而人工新闻则要慢上半拍，甚至依赖机器完成数据处理。在当前社会，速度无疑是新闻媒体的竞争点，因此人工智能新闻生产的效率成为了各大新闻媒体的有力竞争武器。随着技术的不断发展，相关技术更加成熟，人工智能新闻生产不仅能够实现速度，还能达到质量保证。

在竞争激烈的媒体行业中，取胜的关键在于时效性。人工智能新闻生产由于其新闻生产主体的模仿和机械化，新闻写作机器人能够全天候检测全球互联网网络数据，并在第一时间收集相关数据，然后根据预先设定的算法模板瞬间产出新闻产品。人工智能新闻生产的广泛应用在一定程度上提高了新闻生产的效率。

自2013年以来，美联社已经与自动化洞察（Automated Insights），一家机器人研究公司合作，开发了一种用于财经新闻报纸新闻生产和报导的新技术。这种创新技术可以转换任何不熟悉的数据格式，将其翻译成分析语言，通过算法总结内容的前因后果，并分析数据背

后的特性和趋势，从而自动生成相关的信息图表、文章、报告等。如何推送文章，甚至在网页上放置文章的位置，如何更改点击率低的文章标题等问题，都可以通过分析和总结网络用户的网页中存在的规律来解决，这可以大大提高生产效率并实现生产过程的高效率[10]。

## 参考文献

- [1]孙欣,尹晓峰.空间与互动:从全景艺术到互动艺术的叙事演变.艺术研究,2018(4):3.
- [2]李治路.从互动到概念的新媒体艺术研究.新闻与传播,2022(1):2.
- [3]李立峰.谈初中英语课堂教学中的互动教学艺术.学周刊,2020(32):2.
- [4]韩淑君.公共空间中的互动艺术创作.汉字文化,2019(10):2.
- [5]王万峰.师生互动激发小学语文课堂智慧之花绽放.学术期刊,2022(7).
- [6]钱薇钰,张毅.互动体验趋势下城市景观设计研究.工业工程设计,2020,2(1):5.
- [7]丁子甲.虚拟公共空间——互联网互动媒体.福建质量管理,2018,000(008):24-25.
- [8]穆一文.从幕后到台前:声音表演艺术与大众文化的互动.青春岁月,2019,000(011):71.
- [9]李禅怡.小组阅读教学中的互动艺术.百科知识,2019(2X):1.
- [10]杨春萌.教学理解艺术初探.教育进展,2022,12(6):8.

# 第 8 章 XR 艺术与元宇宙脑-计算机 接口

## 第 1 节 可植入的脑-机接口

### 1. 获取胞外神经元放电数据的方法

神经元放电数据的获取是脑-计算机接口技术研究的基础。在此，我们将神经元放电数据分为两类：真实神经元放电数据和模拟信号数据。真实神经元放电数据：通过神经元放电信号采集系统获得，这些数据包含了反映大脑工作机制的信息，是脑-计算机接口设备的待处理信号。模拟信号数据：基于对真实神经元放电信号特性的研究，建立模型，并通过数学模拟得到已知背景信息的神经元放电信号模拟数据。模拟信号数据主要用于脑-计算机接口技术中神经元放电信号处理算法的测试。本章将描述如何获取这两种类型的信号。

#### 外细胞神经元放电信号简介

神经元放电（动作电位或尖峰）信号可以通过使用微电极外部记录技术来获得。

从生理学知识中我们了解到，神经元在静息状态下，由于细胞膜两侧钾、钠和氯离子浓度差异驱动下的离子跨膜运动以达到平衡时，在细胞膜两侧形成的电位差被称为静息膜电位。当神经元处于静息状态，典型的静息电位值为-65 毫伏（mV）。当神经元受到阈值的刺激时，膜电位的速率、膜电位极性的返回和恢复，这波膜电位的变化被称为动作电位。典型的动作电位由 1 个单位（1U）的峰值电位和后继电位组成。峰值电位包括快速去极化的正向 FB 电位变化和快速复极化的负电位变化。后继电位是指紧随其后的慢电位变化部分。动作电位的产生引起神经元局部兴奋性的一系列变化。大致在动作电位的超发射阶段，无论使用多大的刺激电流，在那个位置都不能诱导动作电位，这个阶段被称为绝对不应期。在随后的短暂时间内，可以使用远超阈值的电流刺激在这个位置激发动作电位，而且动作电位的幅度会更小，这被称为相对不应期。

神经元中发生的动作电位，在动作电位传递过程中，其幅度和时间进程保持不变。不同神经元发出的动作电位的形状和大小会根据神经元自身与微电极之间的传导路径的属性而有所变化。在微电极尖端附近有众多神经元，它们发出的动作电位被微电极所收集。此外，微电极记录还会受到外部环境噪声、采集电路本身的噪声、电极漂移等的干扰，因此最终收集到的胞外神经元放电信号是所有这些因素叠加。

在接下来的文章中，将基于这些真实神经元放电信号的特性建立一个数学模型，以生成已知的背景信息（例如，动作电位释放时间、平均释放频率、动作电位形状等），并将使用神经元放电模拟数据来测试算法的性能。

## **2. 神经元放电信号采集系统**

神经元放电信号属于在强噪声背景下的低频弱信号，这对信号采集系统的设计提出了高要求。本节将在充分研究神经元放电信号特性的基础上，详细阐述信号采集系统的设计和实现。

### **2.1 神经元放电信号采集系统的性能要求分析**

神经元放电信号极其微弱，仅有微伏级别，并且极易受到外部环境噪声、采集电路噪声、极化电压等的干扰。外部环境噪声包括 50Hz 的工频干扰、高频电磁干扰等。工频干扰通常以共模信号的形式存在，其幅度从几伏到数十伏不等。采集电路本身的噪声主要是器件噪声。例如，传感器自身的噪声决定了传感器的最终分辨率；生物电信号放大器输入端的短路噪声限制了能够检测到的生物电信号的最小幅度。极化电压是由于电极的极化造成的，即电极与电解液溶液之间的相互作用形成双层，当有电流流动时，电极与电解液界面电位变化，极化电压通常有数十毫伏。在采集和处理原始信号的过程中，这些噪声会叠加在生物信号上，如果没有采取合理的处理方法，生物信号将被完全掩盖。

为了有效地采集神经元放电信号并从噪声中分离出来，神经元放电信号采集系统需要满足以下性能要求：

高输入阻抗：以确保微弱的生物电信号不会因输入端的负载而降低。

高信噪比（SNR）：放大信号的同时尽可能减少噪声的放大，以便更清晰地检测到神经元信号。

高共模抑制比 (CMRR) : 减少共模噪声 (如工频干扰) 对信号的影响。

低漂移和低噪声电路设计: 最小化由电路自身引起的噪声和漂移, 保证信号的稳定性和可靠性。

带宽选择: 神经元信号的频率一般较低, 因此采集系统的带宽应适当, 既要防止高频噪声, 也要避免信号失真。

适应性和灵活性: 系统应能适应不同的实验环境和需求, 包括不同类型的电极和生物信号的多样性。

数据处理能力: 具备有效的信号滤波、去噪、放大和数字化处理能力, 以便后续的信号分析和解码。

为应对这些挑战, 神经元放电信号采集系统的设计通常采用高精度的放大器、精确的滤波器和先进的数据处理技术。同时, 系统设计应考虑抑制或校正电极极化带来的影响, 以及实时监测和补偿环境噪声和设备自身噪声。通过这些措施, 可以确保有效地采集并保留神经元放电信号, 为脑-计算机接口技术和神经科学研究提供可靠的数据。

因此, 为了检测微弱的神经元放电信号, 必须对其进行放大、滤波等处理, 以提高信号的信噪比。用于生物放大的放大器不仅要有高放大系数、非常低的噪声水平, 还要有高共模抑制比和高输入阻抗。为了达到降低噪声的目的, 还需要采用合理的屏蔽措施、合理的电路板布局和可靠的接地。由于我们选择的实验对象是鸽子, 其背负能力只有几十克, 而且鸽子的脑空间很小。这就要求脑电采集设备必须小巧、轻便且易于佩戴。

## 2.2 神经元放电信号采集系统的硬件设计与实现

生物电信号采集系统的一般流程。由微电极收集的微弱神经元放电信号, 先后经过放大、过滤, 并通过模拟-数字转换存储, 最终利用软件分析所采得的信号数据。在我们的设计中, 神经元放电信号采集系统由模拟前端和后负载系统组成。模拟终端完成神经元放电信号的采集、放大和过滤。它直接放置在鸽子的头部。后负载系统实现从模拟前端发送信号的模拟-数字转换, 并将数字化的信号数据存储于 TF 卡中。

### 2.2.1 模拟前端电路

众所周知, 整个神经元放电信号采集系统电路的噪声水平是由其第一级放大电路决定的, 因此在模拟前端电路的第一级, 我们选择了仪表放大器 AD8221。AD8221 是专为生物医学仪器设计的仪表放大器, 具有低噪声、高输入阻抗、高放大倍数、小封装等优点。其主要参数包括: 最大输入电压噪声为  $8\text{nv}@1\text{kHz}$ , 最大输入偏置电流为  $0.4\text{nA}$ , 最大输入温度漂移为  $0.3\mu\text{v}/^\circ\text{C}$ , 输入阻抗为  $100\text{G}$ 。运算放大器使用的是低噪声 TLV2264, 电路中使用的电阻和电容的精度为 1%。

仪表放大器作为神经元放电记录装置的第一级, 如图(a)所示。当  $R_1$  是  $1\text{K}$  时, 放大系数设置为 500, 微电极收集的微弱 EEG 信号分别送到仪表放大器的  $\text{IN}+$  和  $\text{IN}-$  端子。为了消除如基线漂移这样的低频噪声,  $\text{OP}1$ 、 $\text{C}2$  和  $\text{R}2$  形成了一个高通滤波器, 其截止频率设置为  $0.7\text{Hz}$ 。第二级是放大和滤波电路。放大系数的设定原则是放大输入信号的幅度, 使其能满足后续主芯片模拟-数字转换单元的动态输入电压范围。 $\text{R}5$  和  $\text{C}3$  形成一个低通滤波器, 截止频率设置为  $3\text{kHz}$ 。因此, 在模拟端, EEG 信号被放大到足够大的程度, 以减少环境噪声在信号传输期间对信号质量的影响。由于神经元放电信号有正负之分, 所以需要有一个电平升高电路。电平升高是通过电压跟随器实现的, 电压由供电电压上的  $\text{R}7$  和  $\text{R}8$  分压确定。这里我们取  $\text{R}7$  为  $400\text{k}$ ,  $\text{R}8$  为  $100\text{k}$ , 所以电平升高电路的输出电压值是供电电压的五分之一。为了提高这个电压的稳定性, 选择了一个与电阻  $\text{R}8$  并联使用的电容  $\text{C}5$ 。模拟前端直接放置在鸽子的头部。

### 2.2.2 背部负载系统

背部 I 系统由 *iboshan* 主控核心片、TF 卡和电池组成。主控核心是 ATmega80 供电的锂聚合物电池。模拟前端和背部负载系统分别独立供电。模拟前端使用两个  $180\text{mAh}$  的电池串联供电, 背部负载系统使用一个  $650\text{mAh}$  的电池供电, 三个电池的总重量不超过 25 克, 这可以确保神经放电信号采集系统能够连续工作 12 小时以上。背部负载系统将从模拟前端发送过来的信号转换为模拟信号, 并将数字化的信号数据存储在 TF 卡中。主控芯片的模拟-数字转换模块选择内部  $2.56\text{V}$  参考电压, 采样频率设置为  $6\text{kHz}$ 。使用 EEG 采集设备收集的真实神经元放电信

号数据测试新算法，或者比较两种算法的性能是非常不明智的，因为真实数据中的动作电位发放时机、优势神经元数量、动作电位形状、噪声水平等参数是未知的，因此无法给出算法的客观评价。解决方案是建立信号的数学模型，用已知背景信息模拟仿真数据，然后使用它们来测试算法。

模拟数据是根据相关文献中描述的方法生成的。在这种方法中，细胞外神经元放电被看作是背景噪声、单个单位放电和多单位放电的线性叠加。区域 1、2 和 3 分别代表背景噪声、多单位放电和单个单位放电信号区域，这些区域是根据区域内神经元与记录微电极尖端（中心）的距离来划分的。在图中， $a$  代表信号幅度， $d$  代表神经元与中心的距离， $s$  指的是某个神经元的放电信号。神经元放电信号的幅度与其与微电极尖端（中心）的距离成反比。在三个区域中构建神经元放电模拟信号是基于动作电位数据库的。这个数据库包含了从真实神经元放电信号中提取的数十到数百个动作电位的波形数据。

#### (1) 背景噪声模拟

背景噪声（区域 1）用于模拟较远距离（超过微电极尖端 150 微米）的动作电位放电。首先，从动作电位数据库中随机抽取一定数量（通常是数十个）的动作电位。对于每一个动作电位，生成一个满足均匀分布的时间序列间隔，并且背景噪声是这些神经元放电序列的线性叠加。每个序列的权重与发出这个动作电位的神经元与微电极尖端（中心）之间的距离成反比。通常，在生成的背景噪声上叠加一个白噪声信号，以模拟外部环境噪声、电路自噪声、极化电压等干扰。最后，对背景噪声进行归一化处理，使其幅度小于动作电位检测阈值的 0.5 倍（例如，阈值可以设置为背景噪声均值加上噪声标准差的几倍）。

#### (2) 多单位放电模拟

多单位放电（区域 2）用于模拟更近距离（微电极尖端 50 微米至 150 微米）的动作电位放电。从数据库中选择  $N$  个动作电位，并且为每个动作电位生成满足泊松分布的时间序列，动作电位的平均发射频率是  $(20/N)$  Hz。多单位放电信号是这些放电序列的线性叠加。通过这种方式，最终多单位放电信号的动作电位平均发射频率是 20Hz，并且最后将多单位放电模拟信号的幅度进行归一化处理，使其幅度等于动作电位检测阈值的 0.5-1.5 倍。

### (3) 单个单位放电模拟

单个单位放电用来模拟距离记录微电极尖端最近（50 微米以内）的优势神经元素的放电。单个单位放电模拟信号的生成方法与多单位放电模拟信号的生成方法相似，但参数不同。单个单位放电信号的幅度设置为动作电位检测阈值的 1.5-4 倍。每个神经元放电序列的平均释放频率为 5Hz，并且有一个 2ms 的不应期。

## 第 2 节 嵌入式脑-机接口系统研究

### 1. 嵌入式脑-机接口系统信号选择

大多数脑活动都能引起脑电图 (EEG) 信号或血流的变化。这些脑活动可以通过监测生理信号来记录。根据信号来源的分类, 脑-机接口可以分为两大类, 一类是基于电信号的, 另一类是基于影像的, 这两种方法都有各自的优点和缺点。

#### 1.1 神经成像方法选择

##### 1.1.1 脑电图 (EEG)

脑电图 (EEG) 通过头皮记录脑内神经元的一些放电活动, 并且通过在头皮上放置多个电极, 有可能记录一段时间内的脑活动。EEG 被认为是最常用的脑电记录方法, 它可以检测到毫秒级别的脑活动, 不仅具有高时间分辨率, 而且具有操作简单、安全性高、成本低廉、无不良副作用等优点。

##### 1.1.2 磁脑图 (MEG)

磁脑图 (MEG) 利用功能性神经成像技术来识别和分析大脑的磁场, 并且在临床上被广泛使用。它最早由 David Cohen 于 1968 年发现, 并开始屏蔽室内使用由导电铜管制成的探测器进行检查, 以减少背景噪声。最近, 开发了更为敏感的传感器, 如 SQUID 超导量子干涉装置, 以捕捉更准确的 MEG 信号。MEG 特别有效, 特别是在癫痫和脑瘤患者中, 它可以帮助识别癫痫或肿瘤或其他块状病变患者的正常大脑功能区域。MEG 主要使用电磁波工作, 因此它可以与 EEG 联合使用, 为 EEG 提供补充信息。

##### 1.1.3 功能性磁共振成像 (fMRI)

功能性磁共振成像是一种非侵入性技术, 用于在大脑活动期间测量血氧水平

的变化。功能性磁共振成像具有很高的空间分辨率, 适合于识别大脑的活跃区域。然而, 功能性磁共振成像的时间分辨率非常差, 范围在 1 秒到 2 秒之间。fMRI 对头部运动很敏感, 在测试过程中需要保持身体静止, 否则容易产生误差。

#### **1.1.4 皮层脑电图 (ECoG)**

皮层脑电图通过侵入性电极测量大脑活动, 即需要打开头骨, 直接将电极放置在大脑上进行测量。由于距离神经元较近, ECoG 具有高时间和空间分辨率, 获得的信号优于 EEG, 并且信噪比也非常高。然而, 它需要暴露脑区域, 所有测量只能在手术室中进行, 这也是一个高风险的过程。

#### **1.1.5 正电子发射断层扫描 (PET) 和单光子发射计算机断层扫描 (SPECT 或 SPET)**

正电子发射断层扫描 (PET) 用于观察体内的代谢过程, 与 SPECT 类似。进行 PET 检查时, 会向体内注射一个放射性源, 而在 PET 检查对象体内, 由于注射的放射性核素会发射出一对伽玛射线。利用这些伽玛射线, 可以构造出非常清晰的图像, 但 PET 设备和可注射的放射性源昂贵且操作成本高。

#### **1.1.6 光学成像 (功能性近红外 (fNIR))**

光学成像 (功能性近红外 (fNIR)) 是一种新技术, 应用近红外光谱学原理和神经血管耦合原理, 实现对大脑功能的成像。近年来, fNIR 技术已被应用于脑损伤功能监测、脑-机接口、自然环境下的高级认知、发展心理学、异常心理学等领域的研究, 并取得了一定的成果。与目前成熟的大脑功能成像技术如功能性磁共振成像 (fMRI) 和正电子发射断层扫描 (PET) 等相比, fNIR 技术具有更经济、更便携、无噪音、非侵入性、连续实时、对实验过程中受试者的动作不敏感等特点, 也可以与 fMRI、EEG、TMS 等技术结合使用, 对于未来的临床医学、认知心理学等领域的研究具有广阔的应用前景。

脑-机接口系统在本节中设计的是一个便携式实时在线控制系统, 对时间分辨率有很高的要求, 而对空间分辨率的要求相对较低。直观地比较上述方法表明,

EEG 的时间分辨率很高，信号采集设备的成本低，设备体积小，加上设计成本和安全性考虑，最终认为 EEG 最适合作为本文的信号方案。因此，本文使用 EEG 信号作为设计的脑-机接口系统的输入信号。

## 1.2 EEG 信号选择

根据 EEG 产生的方式，EEG 脑-机接口系统常用的信号有自发 EEG 信号和诱发 EEG 信号。

### 1.2.1 自发 EEG

自发 EEG 信号是大脑本身发出的信号，不需要外部刺激，也不占用感觉通路。自发 EEG 是最符合脑-机接口原始定义的信号，其应用前景非常广阔，是未来脑-机接口研究的关键方向。然而，自发 EEG 信号波动较大，非常容易受到受试者情绪和外部刺激的干扰，信号特征往往不明显。目前，自发 EEG 较为流行的应用是基于运动想象的脑-机接口 (MI-BCI) 的运动感觉节律，这类脑-机接口系统在使用前需要长时间训练，且系统稳定性不高，信号识别率低。

### 1.2.2 诱发 EEG

诱发 EEG 需要特定的外部刺激来诱发产生规律、稳定且具有明显特征的 EEG 信号。参与者通常不需要进行针对性训练，系统稳定性和识别率更高，更具有实用性。基于实用性、稳定性和高识别率的需求，诱发 EEG 被用作系统的输入信号。

目前，脑-机接口系统中常用的诱发 EEG 信号包括 P300、稳态视觉诱发电位 (SSVEP) 和运动启动视觉诱发电位 (mVEP)。P300 信号是在受试者暴露于罕见或令人惊讶的刺激后近 300 毫秒出现的内源性 EEG 信号。该信号通常由“奇异球”范式刺激产生，要求用户参与一个随机刺激序列，其中一个刺激的频率低于其他刺激。当这一罕见峰值与受试者相关联时，它会触发 P300 EEG 信号。P300 不需要任何训练，但 P300 也有一个缺陷，P300 需要重复刺激，当参与者熟悉这种刺激时，生成的 P300 信号质量会下降，在长期实验中，同一个参与者由于熟

悉刺激范式，常会减少对 BCI 的控制力。

稳态视觉诱发电位 (SSVEP) 是当受试者感知周期性视觉刺激时产生的脑信号，研究表明，当受试者接收闪烁图像、调制声音乃至某些振动时，会产生周期性的 EEG 信号。与此相对的，SSVEP 特指由周期性视觉刺激引发的 EEG 信号。图像闪烁刺激在视觉皮层产生 SSVEP，其频率与闪烁图像的频率相同（通常在 6 至 30 赫兹之间）。不同的目标闪烁序列对于不同的刺激应当是正交的或彼此接近的，以确保目标的可靠识别。SSVEP 基本上不需要受试者进行训练，但也有一个缺陷，SSVEP 刺激范式是周期性的闪烁刺激，在长时间的实验后受试者很容易感到疲劳，严重时，还可能使受试者产生癫痫，因此 SSVEP 只适合短期使用。

运动启动视觉诱发电位 (mVEP) 是使用运动刺激来在 EEG 中引发响应的一种方法。运动感知是视觉处理系统 (VPS) 的主要任务之一。尽管连续运动并不显示出明显的特征，但已经证明对象的突然运动能够引起皮层活动的突出响应。对视觉刺激的突然运动的响应在背侧通路周围产生神经活动。在大多数研究中，mVEP 在 EEG 上产生清晰的信号并且有强烈的反应，并且提供了参与者内部变异性低相关性的优势。mVEP 能够在对比度比率低至 2% 以及亮度值  $<20\text{cd/m}^2$  的情况下引起低对比度反应。与 SSVEP、P300 刺激相反，使用与运动相关的而不是闪烁的刺激来引发 mVEP，这些刺激的视觉疲劳低。BCI 设计者也对 mVEP 的更为优雅的刺激特性感兴趣，它比基于 SSVEP 的 BCI 使用更少的屏幕空间来实现更清晰的游戏界面。由于大脑天生就具有捕捉运动的能力，使用 mVEP 不需要长时间的训练即可成功使用。基于此，系统采用 mVEP 作为系统输入信号。

通常，mVEP 信号可以通过多种运动启动刺激来诱发，但刺激方式不同，诱发的 mVEP 信号也不同。常用的刺激范式通常包括水平运动、向内收缩、向外扩散和中心旋转四种类型，这四种方式刺激产生的 mVEP 信号的差异主要体现在 N2 点的振幅和空间分布上。

水平运动刺激产生的 mVEP 在参与者之间的差异最小，所获得的 EEG 信号的振幅大，稳定性高。因此，本节设计的嵌入式脑-机接口系统采用水平运动作为刺激诱导方法。

mVEP 分布通常包含三个主要的特征成分：P1、N2 和 P2。P1 成分主要在刺激后约 100 毫秒出现，P1 的特点是振幅小，参与者在进行 EEG 实验时通常不表

现出明显的性能。N2 成分通常在刺激后 60-200 毫秒之间出现。P2 成分通常在刺激后约 240 毫秒出现，它主要在空间上分布在枕叶区域-颞叶和顶叶区域，在通常情况下，P2 和 N2 更为明显，参与者之间的差异小，并且经常被用作特征信号。

## 2. 嵌入式脑-机接口系统架构

本节设计的脑-机接口系统架构包含四个部分：EEG 信号采集模块、特征提取模块、模式识别模块和视觉刺激反馈模块。其中，信号采集模块包括一个 EEG 放大器和蓝牙模块，负责收集用户的 EEG 信号，放大和滤波采集到的信号并通过蓝牙传输到嵌入式系统。视觉刺激和反馈模块有两个主要功能，一是为参与者提供启动运动的视觉刺激，通过该刺激诱导 mVEP 信号，二是通过边界中红色矩形框的移动通知参与者指令是否被正确识别。特征提取模块负责提取携带 mVEP 信号特征的 EEG 信号，然后将其传递给模式识别模块，由它进行分类，以便根据结果信号解码出受试者的意图，最终将解码结果传递给反馈模块。此外，系统识别的指令也可用来控制外部设备，如玩跑酷游戏，控制假肢等。

### 2.1 EEG 信号采集模块

本节设计的脑-机接口系统中的信号采集模块主要包括 EEG 采集电极、EEG 放大器和蓝牙数据接收器。该系统的电极由氯化银制成的干电极构成，并通过自行焊接制作。与传统的湿电极相比，干电极在 EEG 信号采集时无需涂抹导电膏，在采集信号后也无需清洁皮肤，且长期使用时需要重复涂抹导电膏，在使用上更加麻烦。干电极直接与皮肤接触即可采集信号，使得系统使用更加方便，因此系统采用干电极进行 EEG 信号的采集。

EEG 放大器使用自行研发的双通道 EEG 放大器。EEG 放大器由锂电池供电，设计为两个通道，为了提高共模抑制比，放大器特别设计了右腿驱动电路，图片中的白色耳夹作为右腿驱动。放大器共使用 3 个电极，其中一个参考电极，另外两个是数据采集电极。EEG 放大器电路框图，所采集的 EEG 信号首先由 8422 放大器放大，接着进行直流校正，然后通过二阶巴特沃斯低通滤波器对 EEG 信号进行滤波，滤波完成后，数据送至单端旋转差分驱动器产生一个能被 AD 识别

的信号, 该信号经过 AD 转换后发送至主控芯片。数据随后通过蓝牙发送至嵌入式设备。

为了使系统操作更加方便, 系统取消了传统的直接使用数据线实现 EEG 放大器与 PC 之间通信的方式, 改用蓝牙通信进行数据传输。EEG 放大器内部集成了蓝牙功能, 在发板方面, 系统设计有可插拔的蓝牙模块。

## 2.2 嵌入式硬件平台

嵌入式硬件平台采用了友善之臂在线购买的 NanoPC-T2 开发板。NanoPC-T2 使用了三星 Cortex-A9 架构的 S5P4418 处理器, S5P4418 处理器共有四个核心, 动态操作频率 400M-1.4GHZ, 配备了 1GB DDR3 RAM 和 8GB eMMC, 还集成了 WiFi 模块, 并增加了千兆以太网和电源管理功能。为了满足工业客户的需求, nanoPC-T2 配备了标准散热片, 解决了 4418 芯片长期以来的发热问题。NanoPC-T2 几乎是同等尺寸板卡中功能最丰富的微型嵌入式主板, 并配备了更丰富的视频输出接口, 支持 HDMI 1080P 完美输出, 不仅支持各种流行的显示设备, 还能完美适配友善之臂生产的各种电容和电阻屏。

在选择 NanoPC-T2 开发板的同时, 本文还购买了友善之臂的 7 英寸电容屏。开发板和屏幕的物理图。本节在程序开发中使用了串口调试, 并且 NanoPC-T2 开发板没有配备标准串口, 为了方便调试, 本文购买了一个串口适配板 TTL 到 RS232 串口适配板。

为了提高本系统的便携性, 开发板由锂电池供电。锂电池规格为 5V 10000mAh, 能够提供最高 2A 的输出电流, 而开发板的输入电压要求为 5V, 输入电流需求为 2A, 这完全满足了整个系统的供电需要。

## 3. 嵌入式脑-计算机接口系统分类算法

在脑-计算机接口系统中, 收集的脑电信号需要进行分类。常见的分类器包括: 线性判别分析 (Linear Discriminant Analysis, LDA)、主成分分析 (Principal Component Analysis, PCA)、神经网络分类器、支持向量机 (Support Vector Machine, SVM) 等。这些分类器中, 神经网络分类器和支持向量机都需要大量的数据进行训练, 不适用于本系统。LDA 分类器计算量小, 计算速度快, 与本文设计的脑-

计算机接口系统不同，市场上大多数脑-计算机接口都使用 PC 平台，PC 平台的处理器性能强大，计算速度快。而本文中脑-计算机接口系统使用的是嵌入式处理器，其计算能力与 PC 平台存在一定的差距，因此这里选择计算需求相对较小的 LDA 分类器。此外，LDA 分类器在脑-计算机接口的二分类任务中表现良好。在过去的国际脑-计算机接口比赛中，LDA 分类器的稳定性和分类准确性已经被多次验证。

本节首先比较了脑-计算机接口系统中常用的输入信号，最终选择运动启动视觉信号作为系统的输入信号，并详细介绍了运动启动视觉信号。介绍了系统的总体框架，并最后简要介绍了本节中使用的 EEG 放大器、嵌入式硬件平台和分类算法。

# 第 3 节 基于混合模式的脑-计算机接口

## 研究

### 1. 基于 EEG 信号的 BCI 系统

#### 脑电信号概述

当大脑活跃时，大脑皮层细胞产生的生物电会随着空间和时间发生变化，通过放置在头皮表面的电极可以检测到每个点上的电位差随时间的变化。EEG 信号是大脑思维活动过程的直接映射，它可以全面反映大脑的功能状态和大脑思考的电活动，而大脑思考状态的不同会改变大脑皮层不同区域和频段上的 EEG 信号。对 EEG 信号的生理学和医学研究表明，根据 EEG 产生的机制，它们可以分为自发 EEG 和诱发 EEG[2]。

自发 EEG 是指在没有外部环境刺激的情况下，人可以自主调节大脑皮层产生电活动的节律性变化，其电位随时间改变。自发 EEG 具有较强的随机性和非平稳性，噪声大，信号特征不明显，并且容易受到用户自身心理状态和外部环境的影响；诱发 EEG 是指当人在一定的外部刺激（如光、声）作用一段时间后，在大脑的特定区域产生稳定的电位响应，其特点是在外部刺激下不需要被试长期训练就可以获得较为明显的特征信号，而且在刺激信号和诱发电位之间存在清醒、放松、心理紧张、情绪激动、婴儿配置频率、深度麻醉、深度睡眠、缺氧及有机性脑损伤、儿童配置频率、抑郁、患有精神疾病时的变化。

在基于 EEG 信号的 BCI 系统中，根据 EEG 信号产生的不同机制，分为诱发型和自发型 BCI。诱发型 BCI 在外部刺激设备的帮助下作用于用户的某一感觉器官（视觉、听觉等），可以在大脑皮层的特定区域检测到一种特殊的规律性诱发 EEG 信号，并使用该 EEG 信号实现控制命令。因为诱发电位具有强烈的规律性和稳定功能，因此受到各行各业研究者的青睐。然而，诱发型 BCI 需要用户某一感觉通路的配合，也需要使用某种外部刺激设备，导致系统应用受限。自发型 BCI 以自发 EEG 信号为输入，系统不需要用户的感觉器官配合，也不需要使用刺激装置，系统设计简洁实用，应用空间大，是目前 BCI 研究内容的重点，但这类系统要求用户长时间训练，且由于自发 EEG 的独特特性，这类系统的整体识

别率较低。常见的诱发型 BCI 系统有视觉诱发电位(VEP)和 P300, 自发型 BCI 系统则有运动想象 ERD/ERS 和思维任务。

## 2. 基于 SSVEP 的 BCI 系统

当人类视觉神经系统接收到颜色、图案或闪光等视觉刺激时, 从视觉通路产生并可在大脑头皮上测量到的特定电位是视觉诱发电位 (VEP)。不同类型的视觉刺激产生不同的诱发电位, 利用视觉刺激与诱发电位之间的相关性来确定被选中的目标并实现特定控制。

根据不同的刺激频率, VEP 被分为瞬态 VEP 和 SSVEP。当前大多数基于 VEP 的 BCI 系统是基于 SSVEP 的。当多个高频视觉刺激作用于视觉神经系统时, 当下一个刺激出现时, 前一个刺激引发的响应尚未消失, 多个刺激产生的 VEP 电位被多次叠加, 形成含有稳定频率的类似正弦波的 SSVEP。SSVEP 是多个具有稳定频率的视觉刺激彼此叠加的结果, 其周期性显著。SSVEP 电位与刺激频率之间存在对应关系: 与刺激频率相关的频率分量出现在由固定频率闪烁刺激在大脑皮层枕部视觉区产生的 SSVEP 电位的谱中, 这里的相关性指的是相等或整数倍的关系。因此, 在基于 SSVEP 的 BCI 系统中, 可以通过检测 EEG 信号中的频率特征根据刺激频率与诱发电位之间的锁定关系来识别相应的控制命令[3]。

基于 SSVEP 的 BCI 系统已经研究了很长时间。国内的清华大学高尚凯、程明、王一军等人对 SSVEP 进行了深入研究, 利用 SSVEP 信号开发了一个快速的电话拨号系统, 系统中的每个数字方块以不同的频率 (6-9Hz) 闪烁, 通过提取和分类 SSVEP 频率特征, 可以识别用户的视线注视区域, 从而确定命令, 实现手机拨号。

目前基于 SSVEP 的 BCI 系统研究相对成熟, 普遍可用, 系统结构也相对简单和成熟。SSVEP 信号相对稳定, 通常情况下, 参与者无需训练就能获得良好的系统性能, 然而, 该系统需要使用外部刺激设备, 并且要求受试者的视觉感觉通路的协作, 长时间的闪烁刺激会导致参与者感到视觉疲劳, 甚至可能诱发癫痫。

## 3. 基于 P300 的 BCI 系统

事件相关电位 (ERP) 指的是当受试者经历特定的外部刺激并处理这些事件

时，大脑诱导的电位变化。P300 是 ERP 的一种类型，在当前的 BCI 研究中经常作为 P300 来使用。P300 由 Sutton 于 1965 年发现，指的是在大脑受到新颖事件刺激后约 300 毫秒发生的一个前向峰值。刺激事件对受试者来说不常见或不可预测，但与受试者相关。P300 电位是由目标刺激和非目标刺激的结合产生的，目标刺激和非目标刺激的概率不同，大致比例为 2:8。目标刺激指产生事件相关电位的刺激，其他刺激为非目标刺激。P300 的幅度取决于目标刺激的相关程度和发生概率。因此，在基于 P300 的 BCI 系统中，可以通过检测 P300 波形与事件发生时间的对应关系来确定相应的指令指示。

在基于 P300 的 BCI 系统中，由 Farewell 等人设计的虚拟打字系统是最经典的，刺激界面由一个 6\*6 的字符矩阵组成，随机闪烁包含目标字符的一行或一列，闪烁诱发 P300 电位，如果生成 P300 电位时闪烁的行和列相交于目标字符所在的点，那么该交点上的字符就是目标字符。后来，许多基于通信的 BCI 系统继续使用这一范式，并成功应用于严重运动障碍患者。其中最著名的应用之一是美国 Wadsworth 研究中心利用基于 P300 的 BCI 技术为 ALS（肌萎缩侧索硬化症）患者设计了一个字符打印系统，患者可以使用这个系统向他人发送电子邮件。

## **4. 运动想象与思维任务混合模式 BCI 系统范式**

可行性分析与运动想象与思维任务混合模式的构建

### **4.1 可行性分析**

基于运动想象的 BCI 系统是基于受试者在不同肢体运动想象时的 ERD/ERS 生理现象，通过脑电信号中特定频段的能量强度和分布区域的变化来判断受试者想象的内容，进而输出相应的特定控制信号，以实现对外部设备的控制功能。然而，这种方法存在缺陷，因为与运动想象相关的大脑区域在解剖上彼此靠近，当进行简单的运动想象时，不同运动想象任务之间的可区分性较差，且基于运动想象的脑电信号是自发性的脑电，易受到用户的思维活动、心理情绪以及环境的影响，这导致不同受试者，甚至同一受试者在不同时间，ERD/ERS 的生理现象不明显。由于基于运动想象的 BCI 系统是基于 ERD/ERS 生理现象的空间分布特征，这直接导致运动想象范式的应用受到限制[4]。

提高基于运动成像 ERD/ERS 的脑-计算机接口系统性能的方向通常包括以下几个方面：改进系统实验范式以提升系统性能\*\*；通过多种脑-计算机接口模式的融合设计新的混合范式；改善现有的信号处理方法\*\*，如特征提取和分类方法，或提出新的处理方法。

运动成像 ERD/ERS 是自发性脑电，具有很强的自主性，并且不需要外部刺激。目前结合运动成像形成混合脑-计算机接口的模型有 P300 和 SSVEP。基于 P300 和 SSVEP 的脑-计算机接口需要引入外部刺激，并且还需要受试者的外部感觉系统的参与，容易造成眼睛疲劳和不适。基于上述考虑，为了避免给受试者造成不适，同时也为了提高基于运动想象的脑-计算机接口系统的性能，提高不同运动想象任务的区分度，使 ERD/ERS 现象更加明显，本节尝试将一些特殊的思维任务引入基于运动想象的脑-计算机接口系统，并将两种模式整合，构成一个复杂的混合脑-计算机接口系统。

本文提出的混合模式与现有的常见混合模式不同。目前对于混合脑-计算机接口的研究是同时产生两种具有不同机制特征的脑电信号作为系统的输入信号，不同模式的脑电信号对应不同的控制命令，在同一系统内使用两种不同模式的脑电信号来实现控制。而本节提出的基于运动想象和特殊思维任务融合的混合脑-计算机接口系统，与现有的混合脑-计算机接口系统不同，它只有一个输入信号，通过这种脑电信号实现控制，而这种脑电信号是由运动想象和特殊思维任务的融合产生的，即将一些特殊思维任务作为肢体运动想象任务，这样的融合模式使得脑电信号同时包含了两种模式的特征，相对于现有的混合模式，既提高了系统性能，也简化了实验过程。基于运动想象的脑-计算机接口系统通常是通过受试者想象自己的左手或右手做简单肢体运动想象来获取脑电信号，例如：想象左右手食指运动，想象握拳和松拳运动等。本文提出的复杂混合脑-计算机接口在获取脑电信号时进行了改进，将一些特殊思维任务整合到受试者的左右手运动想象中，受试者根据某些特殊思维任务执行左右手想象运动，如空间思维任务和语言思维任务。这种改进既是两种模式的混合，也是对实验设计任务的改进。

将这两种模式融合在一起的原因是：两种脑-计算机接口模式都是自发性的脑-计算机接口，用于控制的脑电信号仅通过想象产生，不依赖于任何外部刺激，也不需要受试者感觉通道参与，不会给受试者带来更多不适，两者都需要受试者

经过一段时间的训练，系统设计简单，属于脑-计算机接口的真正意义。基于运动想象的脑-计算机接口系统主要是通过受试者自发想象某些肢体运动，如左手、右手、脚、舌头等部位的移动；基于思维任务的脑-计算机接口系统是通过受试者自发想象某些特殊思维任务，如保持静止、在空间中旋转、写字和起草、心算等任务。两者都不需要受试者实际的身体运动参与，而是通过自发想象某些动作来实现。

这两种模式产生的电生理活动存在一些相似之处，大脑皮质反应区域有共性，并且在最终产生的脑电信号的特征上也有某些相似性[5]。

## 4.2 混合模式的构建

基于与运动想象和思维任务相关的上述生理基础，本文提出了一个新的复杂混合脑-计算机接口模式，该模式基于运动想象与特殊思维任务的融合。具体描述如下：想象肢体执行某些特殊思维任务，将想象左手运动与空间思维任务结合起来，即使用想象左手旋转立方体来使其执行空间三维旋转，而不是简单的想象左手运动；将想象右手运动与言语思维任务结合起来，即用想象右手默写和默背古诗来替代简单的想象右手运动。通过将想象左手运动与空间工作结合，大脑右半球（特别是在 C4 通道处）8-12Hz 频带的能量减少能力被叠加，同时左半球（特别是在 C3 通道处）8-12Hz 频带的能量增强能力被叠加，从而增加了大脑左半球和右半球之间的能量差异，使得产生的 ERD/ERS 现象更为明显；通过将想象右手运动与言语思维结合，大脑左半球的 8-12Hz 频带的能量减少能力被叠加，右半球的 8-12Hz 频带的能量增强能力被叠加，使得在这一频带内大脑右半球和左半球的能量差异增大，从而使得产生的 ERD/ERS 现象更为明显，以此来增加在执行不同想象活动时的区分度，这直接影响最终的分类准确率。

## 5. 脑-计算机接口 (BCI) 可以在混合现实环境中感知

### 虚拟对象

脑-计算机接口 (Brain-Computer Interface, BCI) 是 1976 年由加利福尼亚大学洛杉矶分校的 Jacques J. Vidal 提出的。简单来说，这项技术是建立大脑与外部

机器之间的接口和互连路径，以实现信息的直接交换。

具体而言，人体的所有体验和感觉都可以归因于神经元的电信号传导，而脑-计算机接口通过某些手段感知和检测大脑中神经元的电活动，并将它们转换成外部信号，控制机器。或者相反，将外部事件转换成电脉冲信号，传回大脑，干预影响神经元的电活动，使受试者能够获得相应的感觉和体验。

一个完整的脑-计算机接口过程包括四个步骤：脑电信号采集、信息解码与处理、重新编码和反馈。在第一步的信息采集阶段，根据不同的采集方法，脑-计算机接口技术可以分为侵入式方法和非侵入式方法（如图 8-1 所示）。



图 8-1: 脑-计算机接口（图片来源：<https://freed.news/game/35345.html>）

非侵入式是指在人类或动物大脑外部佩戴脑-计算机接口设备，通过收集头皮脑电图（EEG）来获取大脑信息，这可以用于简单的信号判断和反馈，或用于 VR 和 AR 游戏应用中的基本手势控制；侵入式的脑-计算机接口设备可以直接植入到人类或动物大脑的灰质或颅腔中，以获得相对高频率和准确的神经信号。

此外，它不仅能通过读取 EEG 信号来控制外部设备，还可以通过精确的电刺激在大脑中产生特定的感觉。尽管侵入式脑-计算机接口的实现比非侵入式脑-

计算机接口更加困难，但更高的频率和高精度信号传输确立了它作为脑-计算机接口最终实现的必要性。

脑-计算机接口技术从通过意识控制机械、信息读取和行为预测，到今年 Neuralink 发布的视频中展示的带有脑植入芯片的猴子通过思维玩视频游戏，这项技术不断地给人类带来惊喜。如今，它又有了新的突破：约翰霍普金斯大学的研究人员通过在混合现实环境中集成神经刺激，开发出了“感知”虚拟物体的能力。

该研究由美国国防高级研究计划局 (DARPA) 资助，是神经增强操作 (NEO) 项目的一部分，旨在研究神经多路复用，比如大脑能在多大程度上通过脑-计算机接口完成其感官和肌肉的感知，以及操作者是否能够在使用大脑控制其他通道或维度的同时，使用手和感官与计算机进行交互。

通过神经刺激感知虚拟物体的研究集中在神经多路复用的一个极端版本上，不同的信号发送到大脑的同一区域，并且神经刺激将现实混合到同一部分，以观察大脑的感觉运动区域如何成功地运作。

就研究结果而言，视觉和听觉的增强在研究中是可行的，但触觉还有改进空间，并且在嗅觉、味觉或本体感觉方面尚未实现。未来，研究人员希望通过直接的神经输入改变这一切。

## **6. 即将到来的虚拟现实终极体验**

### **6.1 虚拟变成现实**

在今天的活动中，VR 的研究呈爆炸性增长，专家们正在寻找创新的方法，使虚拟体验更加沉浸和逼真。现在，VR 技术已经超越了视觉和听觉，我们已经开发出允许用户触摸虚拟物体、感受风和温度变化，甚至在 VR 中品尝食物的技术。

然而，尽管取得了这些进展，人们仍然不会将虚拟环境误认为是真实世界。技术还不够先进，如果我们仍然必须依赖传统的头显和其他可穿戴设备，VR 永远无法实现其真正的承诺。

《穿越大西洋的隧道》，这个电视虚拟艺术装置，建立了蒙特利尔和巴黎之

间的连接，这两个城市实际上相隔数千英里。

隧道使得两地的数百人得以相遇。在每一端，一个两米直径的管道使我们联想到线性穿越我们的地球，就好像它是在地下挖掘的，在蒙特利尔当代艺术博物馆的中央一侧喊出声来，在巴黎蓬皮杜中心的底层另一侧。

例如，隧道在大西洋下运行，位于两点之间的路线并不是海底的模拟，而是一个象征性物质的块体，在这里地质层让位给图标层。它们是在两种文化的历史中摄取的图片层，每个人在挖掘时都可以揭示出来。集体探索揭示了稀有或熟悉图片的碎片，这些图片是唤醒参与者集体记忆的机会。帮助我们闲逛和与人交谈，这些遗迹将每个人的挖掘路线转变为一次独特的体验，变成一个由声音和图片组成的个人组合，在一个通过他们的动作构建的三维空间中。在挖掘时，参观者可以与大西洋彼岸的伙伴交谈。他们声音的回响锚定在空间中，它们使每个人都能找到与对方相遇的方向。在双方的挖掘者面对面会面之前，需要六天时间来建造和铺设象征性空间。

在物理限制之外，空间随即成为时间的函数。在那里，速度不是加快会面的最佳方式，而是指定每个人在信息中位置的一种方式。隧道的架构由每位访客在移动时创造，决定了图片的编辑，就在他们构建新空间的同时。

由新挖掘的隧道改变和塑造的图像唤起了风景的实质，随着每位探险者/访客决策的结果而重新定义自己。它们在时间上的顺序和在空间上的组合既不仅仅是预定的元素，也不是随机化的元素。通过他们处理的事物，以及通过挑选的图片，这些都来源于每位访客自己的挖掘方式。如果我们无法掌握我们将要发现什么，我们发现的东西取决于我们做事的方式。如果我们让自己享受挖掘高速时那诱人的即时感觉和欣快的能力，我们不会遇到与我们在仔细和好奇地探索发现的元素时看到的相同的图标遗迹。每个人对文档中某些细节的兴趣解释了之后会出现的主题和语义发展。于是，写作过程不再关注一个确定建立的声音和图片的积累，而是关注由于访客的探索行为而创造它们出现条件的过程。

隧道下的大西洋的组合机会和决心定义了结果架构，使得世界探索类似于我们当前的生活经验。"Gadevu"，为大西洋隧道开发的基本版本的代理，已经成为我们可以用于复杂数据库的动态和直观探索的 Z-A Profiler。结合自发的行动和对话，马丁·马塔隆 (Martin Matalon) 组成的音乐随着事件的发展而改变，并围

绕个人路线组织，图片也是如此。

电视虚拟活动 - 即在交互式象征空间中远程连接人们 - 由四个虚拟摄像机拍摄。他们获取的内容会自动混合和编辑，同时考虑到每个参与者的言论。他们可以在反拍事件中发现，他们自己的实时图片在他们刚刚挖掘的空间中漂浮。在隧道的两边相遇之前，他们将无法看到对方。到目前为止，交换基本上是由声音组成的，然后变得可视化。当会面实现时，其他人最终可以走同样的路或创建新的路，就好像他们在集体探索共享记忆。

## 6.2 现实 2.0

神经现实指的是由直接与人脑相关的技术驱动的现实。传统的虚拟现实依赖于用户对外部刺激的身体反应（例如，挥动控制器在屏幕上挥动虚拟剑刃），而神经现实系统则通过脑-机接口（BCI）直接与用户的大脑接口。

值得注意的是，这项技术不是遥不可及的科幻未来，它是非常真实的。

知识回顾：脑-机接口是连接我们的大脑到机器的手段，它们可以是侵入性的（需要某种植入物）或非侵入性的（依赖于电极或其他外部技术来检测和引导脑信号）。专家预测，BCI的进步将开启人类进化的新纪元，因为该设备有潜力彻底改变我们治疗疾病、学习和交流的方式等等。简而言之，它们将革命性地改变我们看待和与周围世界互动的方式。

事实上，一些公司已经在新兴的神经现实系统领域进行创新。

EyeMynd 由物理学家 Dan Cook 于 2013 年创立，旨在创建一个 VR 系统，用户可以通过他们的思维导航虚拟世界，无需打破沉浸感的控制器。

"无论是游戏还是其他什么，在虚拟现实，你不想一直担心你的手在做什么，"Cook 去年 11 月在一次采访中说。纯粹的脑波控制会更好。它将提供更满意的体验，并能够实现更高水平的沉浸感。你可以忘掉你的身体，只专注于你面前发生的事情。"

Cook 将这种体验比作做梦，他说："在梦中，你可以在不移动脚的情况下奔跑。做梦和想象会产生我们可以读取的大脑信号。有了这个，你不需要眼睛就可以跑。看，不用耳朵听，不用手和脚，我们可以绕过所有这些。"

EyeMynd 的系统属于非侵入性类别, 这意味着它不需要用户植入任何类型的设备。相反, 他们只需佩戴一个包含 EEG 传感器的头戴设备, 来追踪他们的脑波。

EyeMynd 并不是唯一一家探索使用外部技术来检测脑电波, 从而使虚拟现实体验更加流畅的公司。总部位于波士顿的初创公司 Neurable、生物信息学公司 EMOTIV 和社交网络巨头 Facebook 都在研发非侵入式设备, 这些设备允许用户用他们的思维来导航虚拟世界。

然而, 正如音频技术初创公司 OSSIC 的技术总监 Joy Lyons 在 2016 年洛杉矶虚拟现实夏季博览会 (VRLA) 上所说的, 无论多么先进, 创造新现实的理想硬件并不是一个外部头盔, 而是“大脑中的芯片”。

### 6.3 心中的世界

今年早些时候, 埃隆·马斯克创立了 Neuralink 公司, 目标是开发尖端技术, 通过一系列植入的电极将人类大脑与数字世界连接起来。在马斯克宣布后不久, Braintree 创始人 Bryan Johnson 宣布了一个类似的决定, 投资 1 亿美元以释放人脑的力量, 希望能编程我们的神经代码。Johnson 的公司 Kernel 正在努力开发世界上第一个神经假肢。

马斯克预测, 我们最终将能够创造出与现实无法区分的计算机模拟, 并且, 如果这些大脑接口实现了, 它们可以作为我们体验这些模拟的平台, 允许我们不仅仅是看到一个逼真的数字世界, 而且能触摸它, 真正感受它。

在一份关于 Neuralink 启动的详细报告中, Tim Urban 详细介绍了这项技术可能对我们理解现实的影响。我们不需要外部硬件, 比如耳机、手套或头戴设备来欺骗大脑相信我们面前的虚拟环境是真实的, 我们可以编程和触发大脑的同一部分, 让大脑认为体验就是现实。

Urban 说: “你当然不需要一个屏幕, 因为你可以让你的视觉皮层展示一个虚拟屏幕。或者用你所有的感官走进一个 VR 电影。你几乎可以免费体验任何事情。”

当你在这个新现实中咬下一片“披萨”时, 当你尝试真正的披萨时被激活的大脑部分会被触发, 并提供一种仿佛你在吃披萨的感觉; 当你身处虚拟大西洋海

岸时，同样的部分会被触发并提供真实世界中呼吸海风的感觉。

确定这是否在技术上真的可行并不容易，并且克服与技术无关的障碍将带来额外的挑战（例如开发人脑及其所有神经元的全面图谱）。选择性脑部手术是一个非常具有争议的话题，过去的实验并没有产生如此有希望的结果。Neuralink 和其他志同道合的公司将需要数年的研究才能准备好人工植入物。即便如此，他们还需要克服监管和伦理障碍。

尽管大脑-机器接口 (BCI) 研究正在迅速发展，所以虽然一个能够有效地将新世界直接植入大脑的电极系统可能看起来像是科幻小说中的未来，但它实际上并非遥不可及。毕竟，二十年前我们都认为 Rift 今天提供的虚拟现实体验是难以想象的，但现在你只需 399 美元（限时优惠，它将从原来的 600 美元永久降至 500 美元）就可以把它带回家。

正如 Cook 所说，我们可能认为今天人类通过思想导航虚拟世界还很遥远，但实际上并没有那么远，“在十年内，这将变得显而易见。”

## 7. XR 艺术研究案例

第一场演讲是关于法国艺术家莫里斯·贝纳永 (Maurice Benayoun) 讨论 NFT、区块链以及他四十多年来的革命性作品。

我坐在莫里斯·贝纳永所谓的“拷问椅”上，一个头戴设备紧绑在我的前额，感应并与我的脑电波互动。在我对面，墙上钉着的屏幕上，是一个不断变异的、类似液体的物体。想象一下你在严冬中向空中倒沸水，它瞬间变成了冰，它看起来就像那样。



图 8-2: 贝纳永的艺术作品

说这一幕看起来就像直接来自好莱坞科幻电影并不夸张。只不过，这些电极并没有利用我的记忆来提取我的大脑中的任何秘密或间谍信息——实际上，这个三维（3D）的、有点像生殖器大小的物体，迭代了我的大脑波信号所定义的“友谊”概念。

“当你的大脑专注于‘友谊’这个词时，头戴设备通过脑电图（EEG）实时捕捉你的大脑活动，并将其转化为屏幕上不断变化的形状，这看起来像是在动的大脑汁液，”贝纳永在告诉我尝试的这个互动艺术装置时说道。

“Value of Values”（VoV）是艺术家贝纳永的脑力结晶，这是一个基于区块链的艺术项目，源于对抽象人类价值观的文化背景和感知经济价值的好奇，这些价值观包括“爱”、“自由”、“权力”和“正义”等，贝纳永的团队为了研究目的确定了42个价值观。它始于2016年，最初仅仅是一个互动艺术装置，但自那时以来已经演变成更多。

在贝纳永早期的作品中，潜藏着一种命中注定的预言，即技术会说服我们重新发明自己，很明显他总是领先于其他人一步。以“Art After Museum”为例：早在1993年，贝纳永就已经通过一个当代艺术收藏品进入了虚拟世界，该收藏品横跨数字世界和现实世界，拓展了当时还大体上未知的虚拟现实（VR）技术。

在分散于多个地方的实体立方画廊空间内，艺术作品以计算机数据库的形

式构建，并被投影在墙壁上，艺术家们必须与工程师合作，充分发挥场地的潜能。贝纳尤恩说：“展出的艺术作品是生动的——它们不断地进化、转换、消失，并与外界及观众进行对话。”

时间快进到 2022 年，我们现在可以在家中的舒适环境里，通过虚拟观赏平台参观世界上最伟大的博物馆，并在元宇宙中浏览各种富有极大想象力的艺术空间，这些空间触及了收藏家、画廊乃至整个艺术界。如果《艺术后博物馆》对艺术产业的未来走向提供了一个不安的预测，那是因为像贝纳尤恩这样的人敢于走未被探索的道路。

或者，考虑一下贝纳尤恩在 1994 年创作的虚拟现实视频游戏《上帝是平的吗？》。在这个类似修道院小屋的 VR 环境中，玩家追寻上帝的形象：面对一个砖墙构成的世界，当他们向前移动时前面的墙壁变得空心，如果决定改变方向，他们就会打开一个以前从未存在的新走廊，揭示了艺术史上不同文化中的上帝形象。

这个游戏给我传达了一种内省的感觉。贝纳尤恩说：“这就像生活，每个人都定义自己的路线。它也可以是对世界的隐喻，因为我们以为自己在其中移动的这个平坦且无限的空间，事实证明是一个有限且球形的空间。它将观众置于宇宙的中心。”

我不禁打量了一下我们正在进行这次谈话的大房间。就在门边有两个机器人手臂：它们由伺服电机控制，伴随着沉浸式的电影音乐，它们在一场棋局中进行头对头的较量（“它们试图弄明白人类冲突的逻辑”），一位工作室助理很友好地向我展示了这台机器。旁边是一个大型的骨架隧道装置，有一个圆形的屏幕开口，通道逐渐变细，看似从地面上消失。

那是贝纳尤恩 1995 年引人注目的虚拟现实创作《大西洋隧道》的另一种呈现。“迈克尔·森世理（现在是伦斯勒理工学院艺术系的教授）偶然发现了《上帝是平的吗？》并联系了我，希望在法国和加拿大之间做些事情。我提出了一个跨洲虚拟现实项目，它在巴黎的蓬皮杜中心和蒙特利尔当代艺术博物馆之间的访客建立了联系，以参加国际电子艺术协会会议，”他回忆道。

这一次，虚拟性成为了核心元素。来自隧道两端的参观者使用操纵杆“挖掘”由图像、绘画、声音和地图组成的文化材料，每向前一步，他们就更接近于另

一家博物馆的对应参观者。在巴黎和蒙特利尔通过万维网进行首次虚拟会面之前，他们挖掘了五天多的时间。自 1995 年首次亮相以来，该项目已经扩展到新德里、首尔、圣何塞、香港和南京等地。

由法国驻地的罗马尼亚艺术家康斯坦丁·布朗库西创作的一系列青铜雕塑，预定出现在先锋派的布鲁默画廊，当它们在 1926 年抵达纽约港时，受到了密切的审查。尽管这个系列被命名为《太空中的鸟》，布朗库西并没有专注于鸟的物理特性，而是其运动：一座倾斜的椭圆平面形雕塑描绘了这个生物拉长的身体、头和喙。

"虽然艺术品不需缴纳关税，但让海关官员相信那件高长、细长的抛光青铜作品是艺术品，实在有些牵强，因此他们像对待制造的金属物品一样对其征收了关税，"贝纳尤恩说。

随后的审判完全改变了艺术作品的法律定义。贝纳尤恩补充说："这就是艺术家们如何移动界线的。"艺术本身是一种媒介，某物可以具有无限的主观艺术表达，这一命题自古就有，但这确实让我思考了 NFT 反对者对于什么应该和不应该被视为艺术的想法。

"区块链改变了市场对数字艺术的认知，人们开始将数字艺术视为艺术。在 90 年代，除了在专门的画廊或节日里，几乎不可能展示数字艺术。运营和维护一个‘数字画廊’的成本简直太高了，"贝纳尤恩解释说。

他在称赞区块链技术具有解放性之前持谨慎态度："关于 NFT 令人困惑的地方在于，它本身并不是艺术作品或物件。你实际上是在买卖艺术作品的所有权，并在区块链上保存数据。在区块链出现之前，主要是画廊——所谓的守门人——决定什么是重要的，人们应该购买什么。

"画廊大约收取 50% 的佣金，艺术家可以期待他们的作品能够公之于众。现在区块链移除了中间人，因为人们可以自行组织，就像 Uber 司机可以直接与顾客交易服务一样。"

这是否意味着艺术家能够筹集到大量资金呢？"如果我想出售一个 NFT 画作，我需要在区块链上铸造它，这需要花钱，然后在 OpenSea 上展示它，这也需要钱。有很多初期投资，所以围绕 NFT 艺术的所有热闹实际上并不像看起来那么解放，"他补充说。"除非你非常有名，否则你不可能从中赚到很多钱。它所做的是为艺

术家打开市场，赋予人们成为创造者的权力。”

然后，自从 NFT 进入主流媒体和公众心理以来，FOMO（错过的恐惧）一词成了每个人的口头禅。从我过去与艺术家和画廊主的交谈中，大家普遍对数字艺术的普及和它将导致的市场饱和度感到不安。当艺术家仅仅为了金钱而创作时，艺术还保留着什么伦理和文化价值呢？

贝纳尤恩提到了美国艺术家杰夫·昆斯：“他以其气球动物雕塑而闻名于世，这些作品唤起了珍贵的童年记忆。但这是艺术的全部意义吗？”事实证明，昆斯今年早些时候推出了他的第一个 NFT 项目《月相》，并计划将雕塑送入太空。

最近，一些最令人兴奋和创新的虚拟现实（VR）体验结合了 VR 和现场演员。《暴风雨下》体验就是最好的例子之一，因为它提供了独特的互动和参与体验。它还为演员如何在大流行期间成功地远程工作提供了很好的模板，同时为观众提供了无与伦比的体验。

将 VR 与现场真人演员结合起来提供了我们迄今为止看到的一些最好的体验。

例如《爱丽丝，虚拟现实游戏》和《真实的虚拟恐怖》。这两种体验都在 Pippi 中心展出，并由 DV 制作。与《暴风雨下》不同，这两种体验都是在观众面前使用现场演员，而不是在远程地点。

所有小组成员都同意，由于制作的复杂性和雄心壮志，这是他们体验过的最好的 VR 体验。

《无天花板》是一份关于沉浸式艺术和娱乐行业的在线出版物和播客，它列出了 150 多种在远程进行的互动剧场体验，而没有在我们的 VR 中。这让我们很好地看到娱乐行业是如何适应大流行的，也可以为传统表演公司寻找数字化运营的灵感提供很好的资源。



图 8-3: DV 集团的《真实的虚拟恐怖》

虽然随着技术的发展，维护 VR 和 AR 互动艺术作品可能会遇到困难，但这并不适用于像雕塑和建筑这样的静态 XR 艺术。这些静态作品比互动 VR 艺术体验更适合长期收藏。因此，对于那些希望日后转售 VR 艺术的人来说，收集静态 3D 加密艺术可能是一个更好的长期投资。

许多画廊对 XR 技术感到不适，并且如果他们打算展示 VR 或 AR 艺术，将需要帮助。销售、购买和收集 XR 艺术的概念尚未得到常规化。随着越来越多的艺术家对这一概念感到舒适并开始销售他们的 XR 艺术，这种局面将变得不那么

尴尬。

加密艺术变得非常流行，速度非常快，关于它将如何整合或改变 XR 艺术的销售和收藏方式，仍然有许多问题。虽然它与图像和视频很契合，但它缺乏完全支持 XR 艺术的方面。

在加密艺术收藏家中出现了一种新趋势，他们购买 XR 艺术并在线创建自己的数字画廊。这与传统的收藏实践非常不同，看看传统收藏家如何对这种新的收藏方法做出反应将会非常有趣。



图 8-4: 由塞缪尔·阿森诺-布拉萨德创作的 VR 雕塑和建筑

## 参考文献

- [1] 尾崎太一, 王凯. 大脑解码与脑-电脑接口. 《科学世界》, 2022, (4): 10.
- [2] 肖锋. 脑-电脑接口与身体革命. 《探索与争鸣》, 2021, (9): 10.
- [3] 林新源, 张绍敏, 陈光第等. 一种基于眼动追踪和脑-电脑接口技术的人机交互方法: CN111158471A. 2020.
- [4] 钟鹤. 脑-电脑接口, 现实版的《黑客帝国》. 《风和电流代》, 2020, (11): 2.
- [5] 蒋慎, 李铮. 信息互动的“湿艺术”: 脑-电脑接口的方法差异化, 目的分层和战略机遇. 《今日传媒》, 2020, 28(5): 4.

# 第9章 在元宇宙世界中体验“现实”

## ——以电影《头号玩家》为例

近年来，虚拟现实（VR）不再是一个新词，许多领域尤其是游戏和电影行业都进行了相应的研究。电影《头号玩家》便是将两者结合的产物。它想象了一个现实与虚拟交织的未来社会，人们的生活发生了翻天覆地的变化。这个“真实”的VR世界是好是坏？

### 1. 未来的VR世界

《头号玩家》是一部科幻冒险电影，改编自同名小说，由著名导演斯皮尔伯格搬上银幕。该片通过熟悉的屏幕形象唤起了80后和90后的记忆。在向经典致敬的同时，导演也没有忘记为我们描绘一个无限可能的未来世界——2045年高度发展的VR技术，虚拟世界已成为人们的另一个家，这个新家被称为“绿洲”。在这里，你可以通过改变外观、使用昵称等方式成为任何人，并且可以在不离家的情况下体验户外活动。VR设备在我们生活中无处不在，人们大部分时间都花在虚拟世界中。

在“绿洲”中，人们可以轻松地将《回到未来》中的“德洛里安飞行汽车”放入口袋，独立创造出好莱坞动画中的“钢铁巨人”，甚至变身为日本动画中的“高达”体验飞行和战斗，这些熟悉的事物吸引了越来越多的玩家进入VR世界。在电影中，“绿洲”的创始人哈利迪留下了3把钥匙在VR世界，告诉世界谁能依次找到它们，就可以获得终极复活节彩蛋——控制“绿洲”的权力，全世界的VR爱好者都为之疯狂，争先恐后地找方法通关三个游戏。每天，无数玩家在VR世界中竞争，希望获得线索并成为第一个得到钥匙的人，但到目前为止，还没有人成功通过第一关，因此在“绿洲”中突破关卡已经成为未来人们日常活动的一部分。

## 2. 元宇宙中的虚拟化身

虚拟化身在元宇宙虚拟环境中不仅仅是一种独特的存在方式，更广义的定义还包括在线论坛、即时通讯软件和社交媒体中的平面化身，甚至可以追溯到早期基于文本的多人在线游戏中的文本用户名。

1992年，卢卡斯影业制作了一款名为Habitat的游戏软件。这也是历史上第一款图形化在线游戏，其形式非常接近今天的在线游戏。在游戏历史上首次使用了“Avatar”（虚拟化身）来代表玩家的虚拟形象。

如今，越来越多的游戏允许玩家使用游戏提供的模板系统进行适当的个性化。例如，《侠盗猎车手》允许玩家个性化游戏中主角的服装、装备和发型；其他游戏，如元宇宙项目PUGG，甚至允许用户上传自己的平面照片，然后使用游戏内部的3D模型系统转换成三维模型。

这个过程被玩家生动地称为“扫脸”。在沉浸式网络的时代，虚拟化身的展示和角色也经历了革命性的变化。

弗兰克·比奥卡（Frank Biocca）认为，历史上从未有过任何媒体像虚拟现实那样专门致力于扩展人类的感官感知。虚拟现实的终极目标是人类感知、认知甚至智慧的扩展。这种扩展主要依靠化身作为中介来完成。

一些学者对虚拟化身社会效果也持有负面评论，并认为虚拟现实虽然是涉及最多感官和最多使用身体的媒介，但本质上是一种“脱身”的技术。安妮·巴尔萨莫（Anne Balsamo）认为，虚拟现实通过实际身体的物质压制实现了其概念上的拒绝。因此，关于虚拟现实“去身体化”的真正担忧是，它切断了用户与大部分物理世界的物理联系。

围绕元宇宙的一个神话是，它提供了“非肉身自我”的终极可能性。

自古典哲学时代以来，人类存在的生物性和可逝性被认为是人类自我提升的最大障碍。

柏拉图提出了“厌恶身体”（Somatophobia）的概念，并认为意识单独存在而不带身体是一件好事，因为“用身体探索任何事物，灵魂显然要被愚弄。”

对前现代人来说，身体是一个不可逾越的障碍，人们一生都在与疾病和死亡挣扎。

笛卡尔直接将人类存在分解为两个完全相反的部分：“思考实体”（res cogitans）和“广阔实体”（res extensa），这一区分后来在本体论中规定了“自然”和“广阔实体”之间的“精神”差异。从那时起的 300 多年里，西方哲学一直在试图“驯服”身体。

17 世纪之后，由于文艺复兴时期对身体的赞美，长期被压抑的身体逐渐摆脱了神学的束缚，笛卡尔的传统逐渐走向了尽头。

尼采坚决否认了意识的主导地位和灵魂的假设，创造了身体哲学的传统，并认为一切都必须从身体开始，以身体为准则。

莫里斯·梅洛-庞蒂（Merleau-Ponty）认为：“我不是站在我的身体前面，而是在其中，或者说我就是身体。”笛卡尔主义所倡导的意识优先被弗洛伊德的精神分析理论所挑战，后者创造了一个打破笛卡尔式我思的心理过程模型。

对于弗洛伊德来说，思想产生的过程是主体的直接反思所无法触及的。在他的自传中，罗兰·巴特（Roland Barthes）不遗余力地列举了自己的习惯，并认为起源于身体的差异是个性的标记，人与人之间的根本差异是“我的身体与你的身体不同”。

他从尼采哲学的角度解释了身体与意识的关系，即任何人的差异不再来自于思想、意识或精神，而是根本来自于身体。

社会心理学家乔治·赫伯特·米德（George Herbert Mead）认为，游戏是人类构建自我（自我的成因）过程中的一个重要机制。在他 1934 年的著作《心灵、自我与社会》中，他指出社会行为本身是一个互动过程，人类通过共享的符号系统进行交流，游戏也是通过符号进行意义交换。

通过在游戏中的角色扮演，儿童可以在进入真实社会之前完成练习。米德认为，人类通过社会互动中的他人认识和理解自己，并最终完成自我的建立，而游戏的过程为“镜中的我”的发现提供了最佳场所。

自“第二人生”等第一代元宇宙的兴起以来，在线身份的呈现实现了新的飞跃：用户化身变成三维虚拟化身并进入虚拟环境，在实时与其中的人、物和环境互动。

围绕虚拟化身，许多新的次级市场已经出现。例如，在“第二人生”（Second Life）中，曾有许多虚拟公司提供个性化的服装和外观形象给用户，一些公司专门

为虚拟化身推出了社交平台，促进用户在虚拟环境中进行真实的社交互动，围绕共同或相似的身份形成真正的认同感，进而形成独特的社区和文化。

从这个角度来看，虚拟世界中的所有互动和文本都构成了参与者的表达。化身可以被视为参与者的另一个自我。

### 3. 真实的虚拟与现实生活

在现阶段，我们的 VR 技术能够带给人们的体验仅限于观看和通过头戴设备进行简单的互动，它还不能达到电影中真实互动体验的水平。在电影中的未来，人们将穿戴一整套 VR 装备，以获得真实的触觉体验并充分享受 VR 带来的乐趣。用户不仅能看和听，还能在虚拟世界中“触摸”物体。在《头号玩家》中的 X1 套装是一种可以带来全身触觉体验的设备。当你在舞会上拥抱某人时，你将感到你的手臂紧紧地环绕在你的腰间。当你在战场上被敌人击中时，你将在伤口处感到疼痛。

通过编剧和导演的奇思妙想，我们得以窥见一个现实与虚拟交织的未来世界。当用户在“绿洲”中的角色死亡时，它将变成虚拟世界中的金币。人们可以使用这些金币来购买高级装备，而“绿洲”中的稀有装备可以卖成真实货币。在现实生活中订购的物品将由无人机送货上门，实现虚拟与现实之间的无缝连接。通过这种方式，未来的 VR 世界似乎已经变成了一个真实的社会。

虽然电影中“绿洲”创造的梦想非常美丽，但现实生活中的残酷事实却无法被它遮盖。城市中的贫富差距高得离谱，到处是垃圾山，普通市民的生活环境恶劣。上层阶级继续通过高科技剥削普通人。他们瞄准了“绿洲”的控制权，想要通过 VR 赚取巨额利润。人们在 VR 世界中的真实个人信息不应当向他人透露。一旦被犯罪分子获得，个别用户甚至会在现实生活中受到伤害。有些人沉迷于 VR 世界带来的幻想，不断用金钱累积他们在“绿洲”中的虚拟形象，已经放弃了多年来改变世界的想法……”

### 4. “活”在真实中

回顾过去，让我们看看当前的“绿洲”是什么？快速发展的互联网可以说是

一个无形的“绿洲”。通过互联网，我们可以拥有视听共感和交互式沟通。虽然它不像电影中的 VR 世界那样真实，但仍有许多人沉迷于虚幻的网络世界，拒绝与外界沟通，浪费他们美好的青春。

VR 不是一个梦幻机器，它带来的现实感是为了让我们在体验中放松和享受生活，而不是忘记和逃避真实世界，斯皮尔伯格也说过：“VR 不能取代真实生活”。VR 技术不仅仅是一种令人上瘾的娱乐方式，它实际上是引导人们探索世界和促进社会发展的手段。弗洛伊德说过，梦想之所以产生，是因为现实中许多事情无法实现，所以人们倚赖梦想，而 VR 技术可以让梦想在现实中成真。除了在游戏行业的应用外，VR 技术还可以模拟现实中的飞行员训练，减少实际训练所消耗的人力和财力资源。尽管许多人已经开始探索 VR 技术在不同领域的应用，公众似乎更关注其在游戏、电影、购物等娱乐方式上的新发展。这是一个我们需要反思的真实问题。VR 的原始目的是帮助人类改变世界，变得更好，而不是让人们仅仅享受它带来的快乐。

实际上，VR 带给我们的“现实”只是一种幻觉，而真正的生活是由我们自己的经历创造的。VR 技术的发展必然会带来一种不同的生活方式。无论是现实还是虚拟，人们都必须“活”在真实之中！

## 参考文献

- [1] 李强. 平行世界就在你眼前：观后《头号玩家》感. 机场, 2018.
- [2] 陈洁, 陈朝阳, 朱文奇. 《头号玩家》中的 VR 虚拟技术. 影视文学, 2019(9):3.

# 作者的话

人类创造了艺术和技术作为工具，使他们的生活更加美好。区别于这个地球上其他一切以及人造设备的是，人类有能力通过艺术记录、实践和传达他们的精神情感，而技术在改善我们的生活的同时，也改变了我们的生活方式和行为。技术与艺术是不同的，但它们密切相关，我们共有的人性是连接它们的纽带。探索技术与艺术之间的联系，实际上就是探究我们是谁。

无论是艺术实践的旅程还是生活的旅程，它既是终点也是起点，既是结束也是开始。

作者想要感谢所有亲爱的朋友和导师们，他们对我的书稿提出了宝贵意见和帮助。特别是特里斯坦·布劳德教授、徐彬教授、曲华民教授、张康教授和佩德罗·桑德教授。这本书的成就是我们共同努力的结果，也是我在香港三年里所思考的。时间永远不够，会面永无止境。正如中国著名作家顾城所说：

现在草在结种。

它的叶子被风摇晃。

我们站立着沉默，这是极其可爱的。

高泽

2023年3月6日撰写于中国香港清水湾大埔仔