

文章编号:1003-9104(2025)03-0029-07

# 媒介文艺史视野下人工智能 声音艺术的新变\*

王樱子

(1.杭州师范大学 人文学院,浙江 杭州 311121;2.杭州师范大学 文艺批评研究院,浙江 杭州 311121)

**摘要:**长期以来,媒介文艺史对声音的关注几乎等同于对音乐的关注,语音、噪音等其他类型的声音元素则长期被忽视。随着人工智能技术的发展,媒介文艺史中的声音元素呈现出多元化趋势,不同类型的声音被重新纳入文艺史叙述之中。这一变化推动了听觉主体性的建构,使声音不再只是视觉的附庸。从口传时代的具身唱奏,到书写印刷时代的谱写沉默,从机械复制时代的唱片留声,到数智时代的“造音”工程,声音艺术以迭代的方式不断演变。各类前沿的人工智能声音艺术实践,例如人工智能朗读带来听觉快感,人工智能 ASMR 艺术引发通感与联觉体验,均重塑了人们对声音艺术审美效应的理解。

**关键词:**人工智能;声音艺术;听觉快感;声音联觉

**中图分类号:**J02

**文献标识码:**A

在媒介文艺发展史中,除了音乐,戏剧中的台词、诗歌中的吟诵以及影视作品中的声音效果等,也包含非常丰富的声音元素。受限于早期听觉技术的发展,声音艺术在实践和理论层面长期处于发展缓慢、体系混乱的状态。人工智能技术的出现,为声音艺术提供了更加客观、全面与多元的实践空间,同时也促使其理论研究呈现出显著的跨学科特征。

媒介文艺史中的声音问题,不仅涵盖传统音乐学所积累的成熟研究成果,也包括人工智能艺术中所使用的多种声音材料,以及这些材料在艺术中所处的位置与所发挥的功能。与口传时代、书写印刷时代、机

械复制时代相比,人工智能时代声音艺术的生成方式发生了根本性的变化,声音艺术的审美效应也展现出前所未有的听觉快感与通感联觉特征。

## 一、从“乐音”到“声音”： 人工智能声音艺术听觉材料的泛化

在媒介文艺史中,音乐一直是接受最为广泛、研究最为深入的声音艺术形式。然而,借助当代先进的声音技术重新“聆听”文艺史,可以发现音乐只是声音表现的一种形式,而语音、噪音等非音乐性声音的艺术价值日益凸显。事实上,在诸多艺术门类中,音

\* 基金项目:本论文为 2023 年度国家社科基金青年项目“中国新媒介文艺中的声音问题研究”(项目编号:23CZW063);2023 年度上海市教委科研创新计划“新媒体艺术理论基本问题研究”(项目编号:2023SKZD15)阶段性成果之一。

作者简介:王樱子(1993—),女,汉,安徽安庆人,文学博士,杭州师范大学人文学院、文艺批评研究院讲师。研究方向:新媒介文艺,声音研究与听觉文化。

乐以外的声音元素广泛存在,如文学中的讲述声、戏剧中的台词、影视中的声音特效(如变形与合成)等,均构成了丰富的听觉层次。即便在音乐这一经典的声音艺术中,其构成也远不止乐音(即具有固定音高的声音),还包括诸如演唱者的换气声、琴弦的摩擦声、录音设备的失真效果及现场杂音等非乐音成分。这些以往常被忽视的声音细节,借助人工智能等新技术被重新发现与认知,逐渐成为声音艺术中具有独特感知价值的重要听觉材料。

在传统语境中,“声音研究”常与“音乐研究”密切相关,甚至被视为后者的延伸或补充。20世纪以来,随着声音技术的发展,诸如未来主义音乐观与“具体音乐”(musique concrète)等艺术实践,逐渐打破传统音乐对声音材料秩序性与旋律结构的依赖,推动了声音艺术的理念革新。未来主义(Futurism)发端于20世纪初的意大利,涵盖政治、经济与科技等多个领域,强调近代科技与工业交通对人类物质生活的深刻影响,并主张精神世界亦应随之转型。在音乐领域,未来主义强调技术对艺术感知方式的重塑,认为所有可被声音技术探测与记录的听觉材料,皆可成为音乐创作的有效组成部分。然而,受限于当时的声音技术条件,未来主义音乐观在很大程度上停留于理论倡导阶段,难以有效转化为可持续的音乐创作实践。尽管未来主义主张拓展听觉材料的范畴,但其创作仍普遍保留对声音“悦耳性”的基本要求。因而,未来主义音乐家所选用的声音艺术元素,往往仍具有类似乐音的音色特征与协调性。这使得其对传统音乐观的挑战更多表现为形式上的突破,而非彻底的范式转换。

受未来主义音乐观的影响,“具体音乐”流派从根本上突破了“音乐等同于乐音”的传统观念。该流派的创始人皮埃尔·沙费(Pierre Schaeffer)是一位在音乐创作、声音实验和理论研究方面均有重要贡献的作曲家与音乐学者。自1948年起,沙费开始系统探索传统乐音之外的声音材料,并在1948至1958年间创立“具体音乐”并推动其发展。自1953年起,他进一步转向“实验音乐”的创作,持续拓展声音艺术的边界。在沙费的影响下,各类听觉材料被广泛引入音乐创作实践,使之在作品中与传统乐音并置使用,试图打破乐音中心主义,确立非乐音声音的主体地位。这一倾向在美国作曲家、先锋派艺术家约翰·凯奇(John Cage)的一系列声音艺术实验中得到了集中体现。20世纪50年代,约翰·凯奇在美国黑山学院及

德国达姆施塔特现代音乐暑期培训班等场所,开展了一系列实验音乐教学和艺术活动。他通过演讲与组织声音艺术事件剧,广泛引入工业生产声音、生活噪音,甚至无声等艺术元素,推动了偶然音乐、无声音乐及声音事件剧等艺术形式的发展。“声音艺术并不只是给多媒体影像配音,它是一种对传统音乐的批判。这包括对传统音乐和传统乐器的分解和改造,或者音乐概念的抽象化”<sup>[1]126</sup>,如非乐音材料音乐和非乐器音乐。在约翰·凯奇的影响下,后续者在声音艺术实验中高度重视多样听觉材料的运用。这些创作突破传统音乐创作的范式,将生活中的各种声音纳入艺术范畴,打破听众对音乐作品的固定预期,促进了音乐理论界对听觉材料创新的研究与批判性反思。

在跨媒介的前卫声音艺术中,听觉材料呈现出更加多元化的发展趋势。20世纪60年代,由乔治·麦西纳斯(George Maciunas)发起,乔治·布莱希特(George Brecht)、白南准(Nam June Paik)、拉蒙特·桑顿·扬(La Monte Thornton Young)等艺术家共同参与的激浪派(Fluxus),融合工业技术制品、偶发行为和观众参与等元素,通过跨媒介的形式创作了多样化的声音装置。这一运动使声音艺术的实践与研究超越了传统的音乐范畴。在声音技术快速发展的背景下,乐音以外的多种听觉材料开始广泛应用于音乐创作。以约翰·凯奇和激浪派为代表的艺术家,将音乐现场的偶发声视为艺术事件,率先倡导听众对语音与噪音的聆听,甚至提出“聆听声音即是聆听生活本身”的理念。学界将这一时期的声音艺术创作主要成果归纳为五类,“一是延续音乐的传统,以音乐创作为基本范式,结合技术手段,发展出实验音乐的类型;二是结合绘画、雕塑、装置、行为艺术和观念艺术的形式来创作声音艺术;三是从文学的角度探究文本与发声的关系;四是利用广播、电影、电视、录像等媒介技术,探索声音在广播和视听语言中的创意表现;五是源自环境生态、非乐音材料污染方面的考量,强调声音生态的观念,创作指涉环保、公共艺术、声景规划等议题”<sup>[2]99</sup>。

进入人工智能时代,原本分散的声音艺术实践逐渐汇聚,形成了当代声音艺术的历史脉络,发展趋势愈加清晰。以人工智能“文生音乐”(Text-to-Music)为例,日常生活中的声音以及各类非音乐性质的声音艺术材料,在这一创作形式中获得了前所未有的重视。乐音与非乐音材料在人工智能声音艺术中,从传

统的对立逐步转向协同共存。“文生音乐”是指借助人工智能技术,通过文本提示词生成音乐内容的创作方式。擅长结合人工智能与音乐要素进行实践的提示词工程师深刻认识到,声音艺术不仅限于乐音的规律排列,也可具有随机性和无序性。传统观念通常认为,音乐是由音符按照一定规律排列而成的乐音作品。然而,在文生音乐的听觉素材库中,风声、雨声、对话声、环境声及各类生活噪音等非乐音频繁出现,同时人声也可以被自由编辑。这表明,在人工智能时代,听觉材料的多样化已成为“文生音乐”等新兴音乐创作模式的核心特征,而非以往仅存于先锋实验中的偶发现象。人工智能“文生音乐”的创作者将日常生活声音引入艺术作品,挑战了传统音乐艺术的界限,将生活经验转化为审美对象。其核心意图在于颠覆传统艺术观念,彰显一种全新的艺术态度。

借助人工智能技术,可利用的声音艺术材料变得更加丰富。在以“文生音乐”为代表的艺术作品中,声音可细分为乐音、语音和生活噪音等类别,各类声音特性使曲式自动生成,音乐创作得以多元化发展。人工智能声音艺术不仅广泛吸收物理学和声学的技术成果,也融合音乐与人声的艺术特质,实现了技术与艺术的深度结合。

## 二、从“唱奏”到“生成”： 人工智能声音艺术发声机制的演进

声音在媒介文艺史中不仅作为审美的听觉材料存在,更是一种具有大众传播属性的发声媒介。声音的振动特性使其在生产、记录与传播上存在技术难题,因此每一次媒介与技术的进步,皆成为声音艺术的重要革命节点。口耳相传时代、书写印刷时代、机械复制时代、数字智能时代的声音艺术创作,均与声音技术的发展密切相关,呈现出同步演进的态势。

口传时代的声音艺术主要由自然声响(如风声、雨声)和人类具身性“唱奏”——身体介入的声音表现形式——构成。“自然声音指的是所有非中介的声音,它们发生在声音艺术发明之前,或者它们现在没有任何形式的传输或记录。”<sup>[3]</sup>声音的发声依托人体的口舌等器官,因而声音艺术本质上具有强烈的具身性。声音艺术的生产与传播深受身体条件的限制,其听觉记忆多为感知性和印象性的体验。本雅明(Walter Benjamin)在1936年专门论及声音艺术早期传播的特性,指出由于声音艺术依赖人体,其所留下

的听觉痕迹即“经验”(Erfahrung)。“讲故事的人取材于自己亲历或道听途说的经验,然后把这种经验转化为听故事人的经验。”<sup>[4]</sup>“口耳相传”的发声与传声模式极大限制了声音信息的传递效率,导致信息在传播过程中大量缺失与扭曲。由于信息不完整,听觉审美主要局限于“悦耳”的乐音。进入书写印刷时代,声音通过文字和曲谱等视觉符号得以稳定保存与传播。然而,本雅明认为,这一阶段反而造成了口传经验的断裂。他说:“口口相传的经验是所有讲故事者都从中汲取灵思的源泉。那些把故事书写下来的人当中,只有佼佼者才能使书写版本贴近众多无名讲故事人的口语。”<sup>[4]</sup>在口耳相传的过程中,言说者与倾听者共享同一场域,声音交流具有双向动态性。讲述者会根据听众的反应不断调整言说,声音经验因此呈现流动性。

进入书写印刷时代,故事由口头传递转为书面表达,声音的动态特性被削弱,信息传递变得静态且单向,听觉经验的流动因此中断。“小说诞生于离群索居的个人”,小说家闭门独处,“写小说意味着在人生的呈现中把不可言说和交流之事推向极致”。<sup>[4]</sup>然而,声音景观学者雷蒙德·谢弗(R. Murray Schafer)认为,前工业时代的声音景观因声音类型单一而显得纯净。在自然环境中,气候和生物共同构成了丰富的“声景”。在录制和分析这一时期的音响环境时,谢弗及其团队发现,前工业时代的声景拥有宽广的音频,其声学视野(Acoustic Horizon)辽阔,聆听者甚至可以从远处辨识家园的方位。“前工业革命的另一个特点——高保真音响,使‘声学地平线’可能延伸许多英里。因此,来自听者自己社区的声音可以在不可忽视的距离内听到,加强空间感和位置感,并保持与家的关系。”<sup>[5]</sup>这一时期的声音传播依托人类身体(口舌、耳朵),具有本雅明所说的“原真性”:“第一重是此时此地的原真性,讲述故事本身即是一个完整的动态事件,亦是文学作为艺术作品的审美体现;第二层特性则是同时性,在口耳相传的时代,故事的讲述和聆听是同时发生的,讲故事的人与听故事的人实时互动。”<sup>[6]</sup>

在书写印刷时代,声音艺术经历了一段短暂的“沉默”——完全依赖视觉的记录方式(文字与曲谱)将有声艺术转化为无声的抽象符号,口舌与现场表现随之消失,声音的时间性与动态性被固定在纸张上。然而,这种凝固并非对口传时代声音艺术传播的彻底

颠覆或替代,而是一种兼容并蓄的“迭代”过程。声音的时间性与动态性导致其传播速度缓慢、质量受限,难以适应当时的传播技术水平,视觉符号形式不能满足书写印刷时代的信息传递需求。

机械复制时代的到来得益于工业革命的推动,机械复制极大地促进了声音艺术的发展。人们对声音发生原理的认识不断加深,更先进、精确且多样的声音生产、保存、传播与改造技术及其媒介因此诞生。声音的时间被凝固,而空间得以拓展。以留声机和广播为开端,媒介声音(Mediated Sound)通过录音、电话和广播等渠道传播,实现了声音的机械化保存和基本还原式复制,显著突破了传统聆听在时空维度上的限制。“声音关心公共声学,而这种声学是为了成千上万的人能够在他们的家庭环境中收听而制作的。生产商以非常谨慎的方式设计产品的特性,这可以称为声音艺术的声学结构。”<sup>[3]21</sup>口传时代的演唱与作曲家的谱写,随着留声机的普及而逐渐被划分为多个隐形的“声学空间”。现代社会中,人们通过双层玻璃或背景音乐将噪音隔离在外,致使声音被压制或相互混杂,有效信息退化为无效的噪音。个体与环境因此被噪音形成的“声墙”(Sound Wall)所隔绝,声音对象本身呈现出低保真(Low Fidelity)特征。耳机与随身听(Walkman)的发明催生了完全个人化的聆听需求,耳机成为可随个体移动而变化的微型声学空间。为了追求极致音质,耳机设计既注重内部线圈的精确振动以实现高保真播放,又不断强化降噪功能。封闭式头戴耳机通过物理隔离,甚至采用主动降噪技术,发出反向声波抵消外部噪音。在这一阶段,声音艺术的保存、传播与接受方式获得质的飞跃,但声音创作依然依赖人工声音原型(如朗诵、演唱、演奏),通过录音保存或以曲谱形式记录下来。

进入数字时代后,真实的动态音波被转化为完全虚拟的数字信号,传统的主体演唱与演奏过程被省略。声音被拆分为独立的艺术组件,储存在人工智能的造音库中,等待用户调用。在电子合成器和数字音乐工作站出现前,声音艺术的创作依赖人声或乐器的直接演奏。人工智能技术则通过将声音可视化以及信号编码等手段组合各类声音元素,直接生成艺术作品,实现了人与声音的间接互动。比如在一首音乐作品的创作过程中,即使使用播放录制音色的电子琴,创作者仍需通过“弹奏”实现声音的艺术创作;但是在数字时代,“造音”完全省略演唱与演奏过程,无需

练习,直接以“生成”的方式实现从视觉指令到听觉信号的转换,即“文生音乐”模式。

这种生成模式深刻影响了媒介文艺史中研究者对声音的定位与理解。传统音乐学中的声音理论研究多聚焦于主观聆听体验,即探讨声音的“悦耳”与“刺耳”等主客观感知差异。而数字媒介对声音的物质性和传播性的变革,更清晰地揭示了声音研究的理论优势。这一阶段的声音研究表明,听觉文化仅为声音研究的一条路径,而对声音问题的构成作用有限。因此,噪音研究作为声音领域的重要分支,自然隶属于声音研究而非单纯的听觉文化研究范畴。“在这个‘噪音’已经成为一种声音类型的时代,它作为另一种音乐的地位,以及它在整个历史中的变化值得思考。噪音不仅仅是技术本身,它提供的可能性引发了从符号到信号再到系统表示的进步。”<sup>[7]173-178</sup>人工智能生成的艺术作品展现出显著的媒介融合特征,反映出艺术创作的自动化趋势,这一趋势与西方后工业社会的发展密切相关。二战后,伴随着工业生产模式和信息技术的转型,自动化理论被提出来。在艺术领域,一些先锋派艺术家将人机关系的物理学及控制论中的概率自动化理论引入创作,融合“随机性”与“自动化”因素,推动了艺术实践的发展。

依托数字时代声音艺术技术的发展,数字“造音”在方法上实现了重大突破。声音信息的生成不再依赖人体声源,而是可直接由数字信号转换为声音信号。在接受层面,“悦耳”不再是唯一标准,“听觉中心主义”被打破,人与机器的交互构建了平行且虚拟的数字声音艺术空间。人工智能时代的声音艺术呈现出前所未有的多元共鸣形态。以声音为核心素材的人工智能有声读物和“文生音乐”软件,突破人声诵读、演唱及乐器演奏的传统具身限制,对音色进行最大化且富有想象力的扩展,声音与文字、图像等多种艺术元素的融合激发了新的审美效应。人工智能声音艺术作品不仅在创作方式上呈现多样化,其接受效果也呈现共生生态性,观众通过与装置的互动,借助触觉激发声音,并通过声音的可视化将之转换为视觉形式,实现了多感官在艺术中的深度交融。

### 三、听觉快感与通感联觉:

#### 人工智能声音艺术审美效应的突破

人工智能有声读物、“文生音乐”等声音艺术作品不仅展现了当代声音艺术的新面貌,也推动了理论

界对声音审美效应的反思。长期以来,在视觉主导的社会,声音作为听觉对象常被视为视觉的补充,表现出明显的“感官第二性”。然而,人工智能技术在声音审美效应的建构中发挥了关键作用,打破声音的附庸地位,使声音理论经历了从依附视觉、与视觉对立,到主体性逐步确立的转变。人工智能时代的声音艺术不仅带来了独特的听觉快感,还实现与其他感官的多元共鸣与通感融合,形成了奇妙的审美化学反应。

以人工智能有声读物为代表的声音艺术作品主要由纯声音材料构成,较少涉及视觉与听觉的多感官融合,其听觉快感与媒介属性紧密相关。将其与同处于人工智能语境下的网络文学进行比较可见,作为视觉阅读材料的网络文学文本常以邵燕君提出的“爽感”作为主要美学效果。人工智能有声读物的去风格化听觉快感在某些方面与“爽感”存在共通之处,同时也表现出差异。两者的共性主要体现在创作主题的去精英化和目的上的欲望分层。“在网络文学的作者和读者看来,‘意义’不是外在灌输的,而是读者内在欲望的需求,但却是在‘基本需求’被满足基础上的高层需求,二者密不可分。”<sup>[8]</sup><sup>19</sup>听觉快感与视觉爽感的差异,源于两者所依托的媒介形式不同。中国有声读物市场被认为是全球第二大市场。人工智能有声读物依托语音克隆(Voice Cloning)、自动化声音追踪(Automatic Soundtracking)等技术,实现“文转声”(Text-to-Speech, TTS)作品的自动生成,取代了传统的人工诵读,使“讲故事的人”转变为“声音程序”。传统声音艺术作品的演奏与录制通常耗时较长,而人工智能“文转声”模式通过算法生成,大幅节约了时间成本。更为重要的是,传统人声诵读常带有明显的个体色彩,而人工智能声音在发音吐字上极为精准,节奏与语速均匀,情感表达平稳。因此,人工智能有声读物在听觉体验上更具情绪抚慰效果,而非强烈的情绪激发。在实际应用中,人工智能有声读物因其温和平静的声音特质,被用户称为“助眠神器”,体现了其独特的去风格化美学。这一美学特征表明,人工智能环境中的受众并非总追求声音技术无限接近人声。音频恐怖谷理论(Audio Uncanny Valley)指出,人工智能生成的声音若过于接近人声,可能引发听众的不适感。“模拟的声音越接近自然声音,人耳就越能捕捉到其中的细微差异,并将其标记为奇怪或不可思议。”<sup>[9]</sup><sup>8</sup>

除了单一声音媒介形式,人工智能艺术中更常见

的是听觉与视觉相协同的跨媒介形式。因此,需在研究听觉快感基础上,进一步探讨人工智能声音艺术中的多感官融合及声音联觉(Synesthesia)现象。“联觉”类似于中国古代的“通感”,指的是“一种罕见的体验,其中刺激的一个属性会唤起与第一个无关的第二个体验。例如,在词汇—味觉联觉中,单词唤起了味觉的体验。其至少有 60 种已知的联觉变体,包括颜色联觉、味道联觉、触摸联觉和声音联觉等”<sup>[10]</sup><sup>1</sup>。作为融合视觉、听觉、触觉等多感官体验的媒介,并且以非乐音元素为主,人工智能 ASMR 艺术成为近年来代表性的声音艺术形式之一。ASMR,即自发性知觉经络反应(Autonomous Sensory Meridian Response),指人体通过视觉、听觉、触觉、嗅觉等感官刺激,在头皮、颅内、背部或其他部位产生的一种愉悦且独特的感觉体验,俗称“耳音”或“颅内高潮”。“ASMR 是一种感觉状态,其中特定的视觉和听觉刺激持续触发头皮和脖子上的刺痛感,有时扩散到背部和四肢。这些触发性刺激在本质上通常是社交性的、亲密的(例如他听到耳语或看着某人梳头),并常常引起一种平静而积极的情绪状态。”<sup>[11]</sup><sup>361</sup>纵观人工智能“文生音乐”、人工智能有声读物等多种声音艺术作品,传统音乐学中对秩序感的追求并未因对非乐音元素的重视而失效。非乐音元素并非抵抗乐音,也非通过对立构建自身主体性,而是追求多元化的发声形式。人工智能 ASMR 艺术作品主要包含耳语声、咀嚼声、柔软材料揉擦声等听觉愉悦的非乐音元素,极少涉及音乐或语音,体现了声音艺术对非乐音素材的特殊运用。由于乐音具有固定音高,语音具备明确的符号意义,两者的指向性较强,容易使听众集中注意力,情绪趋于外向和紧张。相比之下,日常非乐音元素使听众无需理性地分析或重组声音材料,仅需以纯粹、松散的方式聆听就可随机放松、获得幸福感。在人工智能 ASMR 艺术作品中,音乐模拟的日常声音环境通常表现为主播轻声哼唱以安抚听众入睡,采用低音量、模糊音高的哼唱,而非响亮准确的吟唱。这种非乐音元素信息模糊,更趋向于背景化,远离传统音乐形式。大多数情况下,听众更期待听到低沉的嗡嗡声、风声以及抚摸清脆材料时产生的柔和非乐音,而非具有明确意义指向的语言声音。听众的放松感主要源自非乐音元素,而非语言所带来的认知紧张。一些学者形象地将这种体验比喻为对听众耳朵的抚摸,这种体验可创造声音中的亲密感,拉近主播与听

众的距离,增强安全感。此外,非乐音元素通过模糊不同声音间的界限,形成多种感官交织的“声音联觉”现象。实验研究已证实,声音联觉不仅与视觉密切相关,也涉及运动等感官体验。“受试者能感知到的声音通常是简单的非语言声音(例如哗哗声、敲击声或呼呼声),它们在时间上与视觉闪光或连续的视觉运动相关联。静止场景中的眼球运动(视网膜运动)通常不会引起声音。”<sup>[12]</sup><sup>650</sup>人工智能技术使声音联觉现象进一步融入艺术作品中,这不仅体现了跨媒介生成模式的革新,也促使跨媒介审美效应成为重要议题。从技术层面看,人工智能不仅实现了声音的可视化(Visualization),还推动了其他感官信息的数据声音化(Data Sonification),使多感官体验在艺术创作中得以深度融合。“声音化是指使用非语音音频来传达信息。具体地说,声音化是为了促进沟通或解释,将数据关系转换为声学信号中的感知关系。”<sup>[13]</sup><sup>4</sup>这种可视化与声音化的艺术效果是“产生某种感觉”,更是“生成明确的信息”。人工智能通过信号转换,实现声音与视觉等信息形式的互通,使观众能够以信息的形式对作品进行二次感知。

人工智能声音艺术作品常借鉴经典的声音景观概念。正如现实中的声音景观融合多种感官信息,人工智能声音艺术通过装置结合视觉、触觉、嗅觉等多感官元素,在艺术空间内构建虚实交织的微缩声音景观。借助智能声音化技术,任何数据皆可转化为可听信息,使感官边界模糊。此类作品常与人机交互技术结合,通过视觉、嗅觉、触觉等媒介激发声音效果,实现多感官通感体验。图像与声音被平等视为可供创作编织的素材。“艺术家将数据视为可雕刻的东西,通过规定参数和音阶,以及在软件中过滤、处理和调整声音,使数据成为艺术创作的表现材料。”<sup>[14]</sup><sup>155</sup>在人工智能声音艺术作品中,声音以“信号”形式存在,算法精密地将之编排成艺术品。在过去的视觉霸权时代,声音信号通常被转化为可视化数据;而在人工智能时代,这些可视化数据又被算法转回为可听的艺术作品。这种双向转换在某种程度上体现了对视觉霸权的有效抵抗。

总而言之,人工智能技术对非乐音元素进行大规模采样、合成、模拟和变形,并将其重组为全新的声音艺术作品。其核心理念是将音乐概念拓展为涵盖乐音、噪音乃至无声在内的所有声响,突破传统发声方式和演奏技巧的限制,使声音生产在本质上转为程序

编写。人工智能声音艺术的发展使语音、噪音等元素成为理论研究的核心,声音艺术不再等同于音乐艺术,而获得了更为丰富的内涵。与音乐所强调的“悦耳性”不同,语音更为强调声音信息的清晰性,其对意象的表现更为生动;噪音则由于不具有固定音高和均匀节奏,而更适合彰显艺术作品的先锋色彩。纵观声音艺术的发展历程,尽管许多艺术家、理论家缺乏深厚的乐理基础或精湛的创作技巧,但他们专注于自我表达,忠实于声音素材的真实性,为声音艺术的多元化创作提供了丰富的思想资源和广阔的实践空间。

(责任编辑:徐智本)

#### 参考文献:

- [1]董东.论数字化生成中的音乐与声音[J].黄河之声,2019(20).
- [2]芦影.声音艺术的美学观念与形式语境[J].文化遗产,2018(1).
- [3]Lars Nyre. Sound Media: From Live Journalism to Music Recording[M]. London: Routledge, 2008.
- [4][德]汉娜·阿伦特编,张旭东、王斑译.启迪:本雅明文选[M].北京:生活·读书·新知三联书店,2008.
- [5]Kendall Wrightson. An Introduction to Acoustic Ecology[J]. Soundscape: The journal of acoustic ecology, 2000(1).
- [6]王樱子.从“讲故事的人”到“生成故事的人”[N].文汇报,2025-02-28(9).
- [7]Doug van Nort. Noise/Music and Representation Systems[J]. Organised Sound, 2006(11).
- [8]邵燕君.从乌托邦到异托邦——网络文学“爽文学观”对精英文学观的“他者化”[J].中国现代文学研究丛刊,2016(8).
- [9]Avdeeff M. Artificial Intelligence & Popular Music: SKYGGE, Flow Machines, and the Audio Uncanny Valley[J]. Arts, 2019(8).
- [10]Banissy M J, Jonas C, Cohen Kadosh R. Synesthesia: An Introduction[J]. Frontiers in psychology, 2014(5).
- [11]Stephen D. Smith, Beverley Katherine Fredborg, Jennifer Kornelsen. An Examination of the Default Mode Network in Individuals with ASMR[J]. Social Neuroscience, 2017(12).
- [12]Melissa Saenz, Christof Koch. The Sound of Change: Visually-induced Auditory Synesthesia[J]. Current Biology, 2008(15).
- [13]Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. Sonification Report:

Status of the Field and Research Agenda [ C ]. International Conference on Auditory Display, 2010.

[ 14 ] Parkinson, Adam & Tanaka, Atau. Making Data Sing: Em-

bodied Approaches to Sonification [ C ]. Conference: International Symposium on Computer Music Modeling and Retrieval, 2013.

## New Changes in Artificial Intelligent Sound Art from the Perspective of Media Literature and Art History

WANG Ying-zi

( 1. School of Humanities, Hangzhou Normal University, Hangzhou, Zhejiang 311121;

2. Institute of Literary and Art criticism, Hangzhou Normal University, Hangzhou, Zhejiang 311121 )

**Abstract:** For a long time, the attention to sound in the history of media liberal art has been almost equal to the attention to music, and other types of sound elements, such as voice and noise, have been neglected for a long time. With the development of AI technology, sound elements in media liberal art history have shown a diversified trend, and different types of sound have been reintroduced into liberal art historical narratives. This change has contributed to the construction of auditory subjectivity, so that sound is no longer just an appendage of vision. Sound art has evolved in an iterative manner, from the embodied singing of the oral tradition era, to the silence of the writing and printing era, from the sound recording of the mechanical reproduction era, to the “sound-making” project of the digital intelligence era. Various cutting-edge AI sound art practices, such as the auditory pleasure brought about by AI reading aloud, and the experience of communication and association triggered by AI “ASMR” art, have reshaped people’s understanding of the aesthetic effect of sound art.

**Key Words:** Artificial Intelligence; Sound Art; Auditory Pleasure; Sound Synesthesia

( 上接第 12 页 )

[ 8 ] Jean-François Lyotard. L’inhumain: Causeries sur le temps [ M ]. Éditions Galilée, 1988.

[ 9 ] Hubert Damisch. Traité du trait/Tractatus tractus [ M ]. Paris: Réunion des musées nationaux, 1995.

[ 10 ] 人工智能画作拍出三百万天价, AI 抢饭碗抢到艺术家头上? [ EB/OL ]. [https://baike.baidu.com/tashuo/browse/](https://baike.baidu.com/tashuo/browse/content?id=0cbc6fa2b775cc38f30e04c0)

content?id=0cbc6fa2b775cc38f30e04c0, 2021-03-02.

[ 11 ] Roland Barthes. La chambre claire: Note sur la photographie [ M ]. Paris: éditions du Seuil, 1980.

[ 12 ] [ 德 ] 瓦尔特·本雅明著, 王涌译. 艺术社会学三论 [ M ]. 南京: 南京大学出版社, 2017.

[ 13 ] [ 法 ] 乔治·迪迪-于贝尔曼著, 吴泓缈译. 看见与被看 [ M ]. 长沙: 湖南美术出版社, 2015.

## Can AI Painting Replace Human Painting: A Discussion Starting from “Details” in Paintings

WU Yu-yu

( Department of Chinese Language and Literature, East China Normal University, Shanghai 200241 )

**Abstract:** “Detail” is the key to distinguishing between human painting and AI-generated painting. Through a close reading and historical review of art history, it is evident that the details in human painting encapsulate secrets of the era, history, and individual experience, as well as the energy of innovative thinking. These details generate irreplaceable “Punctum” and “Aura”, triggering emotional resonance and a sense of historical reverence. The analysis and deconstruction of the AI painting principles reveal its technical limitations in perceiving, integrating, and experiencing complex details, as well as its potential in constructing topological structures capable of rich generative creativity beyond details. The future of painting lies not in opposition or substitution, but in a pluralistic vision of differentiated co-existence.

**Key Words:** Artificial Intelligence; Painting; Detail; Algorithm