

人工智能时代AI配音对配音行业的影响分析

胡邦睿

(深圳大学,广东深圳,518060)

版权说明: 本文是根据知识共享署名 - 非商业性使用 4.0 国际许可协议进行发布的开放获取文章。允许以任何方式分享与复制, 只需要注明原作者和文章来源, 并禁止将其用于商业目的。

摘要: 当前, AI 技术为人类服务的现象在人类社会的方方面面得到体现, 而当下的人工智能语音配音, 对传统的配音行业产生了很大的影响。AI 配音以方便、快捷、成本低廉等优势, 在配音行业里都得到了使用者的青睐。本文从播音与主持艺术专业的视角, 采用文献资料法与对比分析法, 讲述 AI 配音在配音行业中兴盛的现象、对传统配音行业的影响以及未来发展的趋向。本文认为配音员应结合自身独特优势, 把 AI 配音作为辅助工具, 合理利用 AI 配音技术的便利对自己创作的配音作品进行优化。本文旨在为配音行业的配音员们如何应对 AI 配音兴起的挑战, 构建人与 AI 合作发展的模式提供理论参考, 希望能为作为配音行业持续地注入活力。

关键词: AI 配音; 配音行业; 对比分析; 优势结合

DOI: <https://doi.org/10.62177/fapad.v2i2.1230>

一、引言

人工智能 (AI) 虽被大众视作新兴领域, 但其理论早在上世纪成型。1950 年图灵在《计算机器与智能》中提出“机器能够思考吗?”的核心命题, 1956 年麦卡锡正式提出“人工智能”概念并给出定义 AI 依托计算机软硬件迭代才从理论构想实现规模化落地。^[1]当前国产 AI 相较 ChatGPT 等海外模型中文适配性更优, 豆包、DeepSeek 是本土化优化的典型代表。AI 语音合成技术的突破让 AI 配音具备规模化商用基础, 其低成本高效率的优势已深度渗透配音行业, 也引发了“人类配音能否被替代”的行业争议。当前 AI 配音行业发展稳中向好, 技术上可实现精细化语音合成, 应用端已获市场与官方传媒认可, ElevenLabs 等工具用户规模庞大; 但针对“人机协同”创作模式、传统配音艺术传承的相关探讨仍显不足。

二、核心概念阐释

作者简介: 胡邦睿 (2003—), 男, 深圳大学本科, 韩国汉阳大学硕博在读, 研究方向为播音主持、传播学、艺术产业, E-mail:3262088139@qq.com。

基金项目: 无。

（一）人工智能

当今是人工智能 AI 进入百姓寻常家的时代，但人们对于人工智能的概念本身仍相对模糊。人工智能（Artificial Intelligence，简称 AI）在当前时代是对社会各方面、各行业产生巨大影响的新兴事物。比如日常使用的智能语音助手、复杂的自动驾驶系统、医疗影像诊断的辅助工具和金融市场风险评估的智能模型等等。

人工智能可以被定义为旨在“让计算机系统具备像人类一般智能行为”的科技。核心目标是让机器能模拟、延伸甚至扩展人类智能的能力，并且能像人类一样思考、学习和解决问题。人工智能科技，其中包括以下几个子领域：机器学习、深度学习、自然语言处理、计算机视觉、图形生成等等。

（二）人工智能配音

人工智能 AI 配音是人工智能技术在媒体中的应用体现，也是人工智能技术与播音主持艺术相结合的重点研究方向。AI 配音，是指通过声音合成软件，将文字转变成语音，再根据具体的场景，选择合适的文字或指定文字进行配音。AI 配音主要应用于新闻播报中，能够实时进行音频合成、音频编辑及语音朗读。^[2] 以下部分具体解释说明 AI 配音的原理：AI 配音技术分为“收集数据并整理、语音特征提取、语音模型训练和语音合成并优化”四个部分。首先，AI 会收集并记忆大量语音数据。为了能够进行数据和数据之间的匹配和深度学习。第二点是 AI 对不同的语音特征的提取。在完成第一步的数据准备后，AI 就会进入声音特征提取的阶段。AI 会对收集到的语音数据进行分析，提取出能代表这些语音特点的关键信息。比如，声音的音高、声音的音色和语速等等。这些特征是 AI 配音需要提取的信息，通过提取这些特征信息，AI 能更好地理解人类语音发声的特点。第三点是模型训练，前两个步骤之后，AI 数据库里保存了处理好的数据和提取的语音特征，接下来会开始训练 AI 模型。训练时，AI 将之前准备的语音数据和对应的文字信息输入至模型中，模型会开始学习语音和文字之间的联系。最后一点是合成与优化，用户输入一段文字，AI 的训练模型就能根据之前训练痕迹生成相对应的语音，然后经过“损失函数”反复优化。

（三）配音产业

配音是一门多元化的语言创作艺术，在不同的时代背景下有不同的艺术创作形式与审美特征。^[3] 配音要做到赋予人物、旁白等角色鲜活生命力的艺术创作，具有相当的专业性。

因此也是一门专业实践。配音领域涵盖了影视、动画、游戏、广告和有声读物等多个领域。先从配音的历史说起，广义上来说，传统人工配音是一个相当有历史的行业。在中国古时候，有“京中有善口技者”的口技表演，在古典时代的古希腊有声音演员，在欧洲中世纪出现的吟游诗人等等，这些声音演绎者都有出现模仿自然、动物和剧情场景人物声音的表演，从广义上来看，这些可算作配音艺术的早期雏形。到了近现代，有声电影《爵士歌王》诞生，这标志着文化和娱乐领域新时代的到来，从此开始有了早期的配音产业。但当时的技术和设备非常有限，只能将声音进行简单录制，在人工匹配画面后进行同步播放。此后，影视行业发展蒸蒸日上，录音技术和设备获得巨大进步，配音产业也逐渐成型。

配音产业的结构大概可分为上游：产出部分，是负责配音作品的生成的人或机构，比如配音演员和配音工作室。中游：是负责推广和发布配音、含有配音的作品的渠道的公司和机构，比如游戏开发商和影视制作公司。下游是零零散散的发布平台，以及各个平台上零零散散的发布者，还有配音作品的受众。简而言之，上中下游的顺序，就是配音作品从产生到进入公众视野，再到通过演绎进入观众们的脑子里的顺序

现代的配音产业，一方面，观众对影视作品的要求越来越高，优质的影视作品肯定需要优质的配音演绎，如此一来。观众对影视配音的需求会更高，甚至推动不少专业配音的培训机构和院校的产生和发展。而另一方面，互联网和电脑技术在不断进步，使得配音行业不仅局限于影视剧配音，在人们的娱乐需求强烈的当下，网络广播剧、短视频、有声读物和游戏接踵而至，它们的配音质量需求也随受众审美提高而提高，以上就是配音产业的发展和现状。

三、人工智能时代 AI 配音的特征及问题分析

在当前配音领域，AI 配音以其具有高效率、高度稳定性和低成本的优势，必将成为一项重要的辅助工具。AI 配音技术展现出了改变行业规则的巨大潜力，但配音不仅是一项技术，更是一种艺术创作。人类配音演员的声音情感表现和对文字的艺术处理优势，目前仍是 AI 难以达到的高度。

（一）AI 配音具有高效率的特征

AI 配音具有极高的创作效率，得益于当今 AI 技术的发达。AI 的信息处理能力和内容生产能力远超人类数十倍甚至上百倍，因此，AI 配音具有“工业级别”的内容生产能力。AI 配音具有高效率的特征有如下三大表现：“批量生成”、“快速迭代与修改”和“自动化音画同步”。

1. 批量生成

用户只需输入文案或素材，多数 AI 配音软件能根据文案背景和浅层情感表达要求，在很短的时间内完成配音任务。肖笛在他的文章里告诉读者：人工智能配音以其强大的稳定性和技术性，能大幅提高成品制作效率，快速生成大量的配音素材。2023 年 10 月 18 日，中文在线集团介绍该公司于 2021 年推出 AI 主播，一天可以录制 500 万字。两相对比之下，人工智能配音一天就能完成传统配音 8 个月的工作量，生产效率惊人。^[4]

2. 快速迭代与修改

AI 在写作的时候可以通过快速修改文本的参数进行调整来达到录音要求，而不需要重新录制。但是如果在传统录音途径中，遇到录错的内容就需要进行重新录制，甚至不能保证重新录制的一遍也保证需求。

3. 自动化音画同步

在视频制作方面，AI 能自动分析视频画面和节奏是否与自己配音的节奏契合，从而调整时长与停顿，减少人工校对时间。甚至很多应用的 AI 可以做到短视频创作者上传画面之后，AI 可以自动生成匹配时长和节奏的配音。

可以看出，人工智能极大提高了视频、音频生产的效率，短视频、营销视频、广告片等各类视频的制作只要不求精细打磨和高质量，都能做到快速生产。

（二）AI 配音具有高度稳定性的特征

AI 配音有人类无法比肩的稳定性，其稳定性突出表现在“无生理限制”、“很难受到外部因素影响”、“错误率极低”等优势上。

1. 无生理限制

AI 在新闻和配音领域被称为“铁嗓子”，因为它完全没有人类配音演员的生理限制。一般来说，就本人自己而言，完全专注配音一小时，嗓子必定会受损，声音质量也会下降。就算是专业的配音员连续工作 2 小时后，普遍也会出现声带疲劳、音色沙哑等问题，但是 AI 完全没有类似困扰，它可以 24 小时不间断生成语音，这种稳定性应用在应急广播和公共服务等领域对人类社会的运行和生产有很大帮助。假如一个地区发生了地震或遭遇突发事件，需要密集发布语音通告，并且根据互联网数据

分析后实时播报当前情况以提醒民众，AI语音播报是很好的选择，AI能确保每条信息都有完全一致的播报质量。并且从时间维度上看，AI能做到一套声库无限次的调用，而人类配音员的音色会随健康状况和年龄变化而自然产生变化，在常规的、不追求绝对质量和艺术水准的配音行业部分需求中，这种稳定性广受用户喜爱。

2. 很难受外部因素影响

最简单的例子——不通过机械控制手段，AI的播报不会被“打断”。可以说明AI具有极好稳定性的原因是对比人类配音员，AI拥有强大的“抗干扰能力”，除去部分物理上的破坏和信号干扰因素，AI配音无法受到外界干扰。举例说明：人类配音员容易受环境噪音、情绪波动等干扰因素影响，哪怕是很多专业的播音员和配音员，比如节目开始录制前，心理承受能力不太好的配音员接收到不好的消息时，例如对象提分手，亲人逝世等消息甚至一些小事而导致情绪产生巨大波动，必然会影响配音演员的状态和作品质量，而AI配音完全不会受到这些因素的影响。

3. AI语音生成的错误率低

AI拥有强大的动态纠错能力：AI配音系统有自然语言处理系统，能够自动识别并纠正文本中的歧义发音，比如在影视字幕配音中，不断更新数据库的AI能够精准识别“他背起了机枪”与“他背叛了队长”中“背”字的多音字的区别，但是人类需要凭借经验判断，在一些需要快速备稿等准备不足的情况下容易因疏忽而导致错误，AI配音很难出现这样的情况。人类配音演员的犯错误具有发生的随机性，而AI配音会随着错误不断被更新到数据库的神经网络中而不断减少，这种稳定性与前文的高效的优点结合是相当完美的搭配。AI配音很少会因“产生错误”而被用户鄙弃。

（三）AI配音具有低成本的特征

成本低廉也是AI配音的巨大优势，毕竟商业社会中的用户们都相当顾虑成本这一因素。AI配音的低成本主要体现在“时间成本低”和“雇用成本低”两方面。

1. 时间成本低

时间也是人生最重要的货币，节省了时间相当于节省一切。无论是在前期的语音生成还是在后期制作这一环，AI在时间成本节省上有着巨大的优势，AI配音借助上文的地错误率的特征，可以不间断生成音频。开始后，AI会对文本进行非常快速的分析，按照用户需求并进行二十四小时不间断输出音频。这种配音方式无须专业的人员进行配音，在较短时间内

生成用户所需的配音素材，在提升配音整体效率的同时极大地降低了配音工作的成本。能够极大节省时间成本。^[5] 凭借稳定高效的特征，当AI配音已经结束工作时，人类配音员可能还正在做熟悉文稿的准备工作，节省的大量时间可以进行其它生产步骤，进一步提高全流程的生产效率。

2. 雇用成本低

再高级的AI配音的价格也不会比聘请配音演员贵，用户们肯定会依据需求来选择最优性价比的选项。如此一来，尤其是在短视频平台上，更多“小制作、大内容”的影视作品呈现在观众眼前。^[6] 举例说明：专业配音员一小时的报价通常在300-5000元之间（根据水平和名气差异估价），这个范围只是预估，在现实中差值甚至更大。而AI配音的成本可降低到聘请配音演员价格的百分之一甚至不到，如果对后期处理不高的要求的用户甚至可以使用免费版本的AI配音。做个假设：制作100小时的有声课程，人工配音需支付约15万元（按配音员300元/小时计），而AI配音仅需不到一万的价格，上文提到，如果创作者没有很高的拟真的质量要求，可以用最普通的AI，使用最基础的配音功能，做到零成本完成创作。

（四）AI 配音存在的问题及分析

总结以上 AI 配音的多项优势，可以得出“AI 配音在很多方面几乎完败传统人工配音”的结论。但是，这并不表示“人类一败涂地”。凡事具有两面性，AI 技术虽然有高效和低成本等巨大优势，但是在模拟人类情绪表现和艺术性处理以及文字细节的处理上还仍存在与人工配音对比的巨大差距。

1.AI 配音模拟人类的情绪表现仍不尽人意

尽管 AI 配音的拟真程度相当高，可以在深度学习和建立模型后不断进行自我完善，但是 AI 技术还不足以强大到如科幻电影中“智械危机”一般拥有自主思考和创作的能力，所以 AI 配音在情绪表达方面多数情况下仍无法与配音演员媲美。由于现阶段的技术限制，无论 AI 有多拟真，在表达的时候还是会缺乏符合人类正常的情绪波动。而配音演员是人类，他们所配音的角色本身也是人类，或至少拥有人类的语言行为。但人类拥有极其复杂的情绪和行为动机，可以做到毫无规律，完全自主产生情绪，这是当今最强大的 AI 也难以模仿的，因为当前 AI 配音本质上是对已拥有的信息收集并处理，在深度学习后建立思维模型。虽然 AI 现阶段已有不少可以模拟复杂情感的配音技术，但是在配音艺术的核心领域——角色塑造方面，AI 还是无法实现比肩人类配音演员的真情实感。^[7]尤其是演绎有血有肉的角色时，AI 配音对角色声音仍是仅停留在表层特征的模仿，包括一些可能会产生“开心”、“难过”的情绪语调变化，但仍然难以通过正确表达人物情绪特征以构建立体的角色人格。就以我来说体验感最好的“ElevenLabs”举例，这款 AI 的中文配音目前尚不成熟，但英文配音的技术很先进，情感语速都处理的相当不错。经过笔者尝试，虽然它能模拟人类说话时的基本语调变化，但在处理复杂情感时，与达到专业配音员的水平仍相差甚远，更不用说超越。诸如 ElevenLabs 一般的 AI 可以做到精准复刻《流浪地球》里的 MOSS 那种冰冷的机械音，但是处理复杂背景下人类产生的情绪和内在语义对它们来说难度仍过于超纲。比如，笔者将丘吉尔在二战时期一段激情澎湃的演讲“我们将战斗到底”输入至 ElevenLabs 的 AI，并做出描述“1940 年 5 月，德军在西欧战场势如破竹，英法联军被迫从敦刻尔克撤退。丘吉尔在此背景下发表演讲，旨在鼓舞英国军民抵抗到底的决心，请用英式英语，且要用带有一丝悲壮但充满激情、坚决的情感播讲”。结果不尽人意，AI 生成了一段听起来毫无感染力且语速均匀的音频，只有一点有规律的微弱情绪起伏。

2.AI 配音的艺术性表现不佳

配音的艺术性其实也是情绪的演绎，本段强调的“情绪”表现，和前文提到的基本的情绪处理不同，艺术性的表现需要做到更细致入微。而 AI 配音在这样的艺术性方面表现不佳，

本文通过一部作品中人与 AI 的表现来说明 AI 配音的艺术性表现为何仍有欠缺。笔者对游戏配音颇感兴趣也颇有研究。在波兰大型 3A 大作《巫师 3》当中，为主角“利维亚的杰洛特”配音的演员名为道格·科克尔（Doug Cockle），这部游戏花了近十年时间制作和打磨，科克尔通过数年的角色塑造，将杰洛特的沧桑、冷静与坚毅却又时而矛盾的性格融入声线。例如，在“血腥男爵”的任务线中，科克尔巧妙控制声带，让声调产生微妙的停顿和颤抖，从而传递出主角与男爵对白时产生的惊讶和愤怒面对家庭悲剧时的复杂情绪，而在一款 AI 模拟角色的软件“猫箱”中，也存在着角色杰洛特，以及中文配音员的声音模拟 AI，让“AI 杰洛特”说了几句剧情台词之后，尽管也有和主角贴切的低沉和沧桑的感觉，但几乎完全没把剧情中角色的声音表现演绎恰当。因为 AI 配音只能通过储存的技艺和算法模型模拟出情感标签（如悲伤、喜悦等），它们是无法拥有人类那样的生命体验和情感的，因此也无法产生良好的艺术表达效果。

3.AI 配音在文字细节处理上仍有不足

其实 AI 配音也难以完全做到“准确配音”，甚至有些时候会犯很多低级错误，人类在出错后可能会及时发现改正，避免后续出错，而 AI 在播报时，如果有多音字识别错误或重音和停连的处理有误，AI 会一直保持这样的错误直至进行程序更改或打断纠正。虽然人们普遍认为 AI 最强大的是其稳定性和低错误率，但其实有过使用 AI 配音的经历就会发现，AI 有时在处理口语中不常见的多音字时往往会出错，也很难每次精准地处理词句的轻重格式。尽管技术人员会在 AI 的神经网络知识库中频繁更新人类社交时出现的语流音变、多音字词等经常改变的要素，但 AI 处理时仍会出现错误。因为中文有很多的新兴产生的多音字和同形异义词等复杂字词，而这些多音字又会产生频繁的官方读音改变，这些复杂的难题成为 AI 配音一道难以逾越的鸿沟。更有甚者是一些深层的文化意象，比如巫山云雨、孔乙己的长衫，面对这些“意中意”，AI 几乎无法做到准确处理，导致生成的配音和文本的深刻的内涵产生割裂感，用户和受众在感受到这种割裂感时会有强烈反感情绪。而配音演员经过培训和训练之后可以很快掌握如何应对每一种不同情感和人文环境下的配音，这是 AI 目前无法做到的。

四、AI 配音对配音行业的影响分析

人工智能 AI 配音技术的出现本质上是为配音行业更好地发展并为配音演员们提供便利，但由于其具有高效低成本且稳定等优势，使得人工智能 AI 配音对配音行业和人工配音产生了冲击，对传统的配音演员们的配音模式产生了较大影响和挑战，但机遇和挑战永远共存。一项新兴技术冲击当下行业，导致就业市场的变化甚至是行业经济模式的变化时，也必然蕴藏着新的机遇。

（一）AI 配音对配音行业造成冲击是必然现象

AI 配音为配音行业带来的对于行业内人来说更大的就业竞争压力是必然的。因为 AI 配音有着极高的效率和极低的成本，必然会导致配音行业中简单的工作受到冲击，比如部分简单的有声读物旁白、广告中简单常规的配音等。制作方为了节省时间和成本，通常会选择物美价廉的 AI 配音。举例说明：一家小型的有声读物制作公司，在过去人工智能配音还未出现或不发达的时候，可能需聘请多名配音演员完成不同书籍的录制，或是给少数甚至单个配音演员高薪酬以完成作品录制。而现在使用 AI 配音，只需一人操作软件就能完成大量内容的配音工作，最多只需一个具有熟练使用 AI 功能进行创作的配音演员做修改润色，对音频做后期处理，相比前两种方案可以节省数十甚至上百倍的成本。这方面的工作转移给 AI 配音完成，直接导致了传统配音演员失去此类工作机会，进而导致其它有人工配音需求的岗位的就业竞争压力增大。

（二）AI 配音将逐渐改变大众对配音作品的审美标准

在工作应用上，AI 配音已经渗透到社会的各行各业，在娱乐上，大量短视频和游戏等娱乐应用也频繁出现 AI 配音。这样的广泛使用会使 AI 配音在潜移默化中、一定程度地降低人们对配音创作的标准，会导致很多“失真”的现象的产生，因为受众听习惯了“非真”的音频。前文提到，AI 配音主要是基于大量已收集的数据进行学习和模仿而创作的，缺乏人类配音演员独特的创造力和具有个性化的表达，如果频繁地使用 AI 配音可能会导致配音风格逐渐趋同，缺乏多样性，艺术性降低等严峻问题。长此以往，如果 AI 配音的使用在大众视野中扮演主要角色，受众会逐渐适应这种语言表达风格，并逐渐认为这是正常的配音表现，大众审美会因为 AI 配音的渗透而逐渐降低。

（三）AI 配音会潜移默化地影响传统配音从业者的创作标准

虽然人类配音员拥有自身音色和情绪的独特性，且能在专业领域和发挥创作的自主性，但在很多时候会在潜移默化中受到 AI 影响，因为 AI 配音采样于专业配音演员的音频录制，因此在字音字词

的基础发声上非常标准，有些时候也能够依据模型对文本做出相应基本的情绪处理。心理学上，人类有跟风的倾向，而配音演员也有生活和娱乐，也会接触到各种 AI 配音的视频和音频，因此配音演员在长时间接触 AI 配音的情况下，会一定程度上受到影响，从而导致在创作时会在潜意识里跟风。简而言之，配音演员广泛接触到 AI 配音会逐渐导致配音演员本身的配音艺术创作受到影响。根据前文内容，在人工配音的影视作品中，人类配音演员能够根据角色不同的性格和背景等要素赋予角色独特的声音特点，但 AI 配音在模仿此类风格创作时，很难做到精准和独特，可能会出现多个不同角色的配音风格相似和情感表现失真的情况，从而影响作品的艺术质量。在配音市场出现此类失真和风格趋同的作品越多，配音行业的从业者——配音演员们需要付出改变大众审美和纠正打磨 AI 配音的努力就要越大，要克服自己在专业领域内潜意识中受到 AI 配音的影响就越大。

（四）AI 配音的广泛应用可能导致传统配音行业技能体系重构

结合前文观点，笔者认为随着人工智能 AI 配音的广泛使用。一定程度上会产生“行业技能体系重构”的危机。简而言之，现在的 AI 配音功能先进，能够高效保证质量地完成常规配音任务。但 AI 配音也正在改变配音行业的能力评价标准。传统的配音演员要用数十年打磨的“基本功”，如呼吸的控制、音色改变的技巧和方言配音等专业内容的学习及其它各种技能的掌握，将会面临贬值的风险。结合笔者自身在播音与主持艺术专业的学习经历：在大二至大三过程中有大量需要配音录制完成的作业。为了提升效率、节省时间成本，笔者和很多同学想法一致，认为 AI 配音相当便利而频繁使用，作为一名播音主持艺术专业的学生尚且会产生这样的想法，更不用说以流量和金钱至上为信条的非专业人士，他们对“情声气结合”、“语流音变”之类的专业技巧不感兴趣，反之更关注如何使用 AI 能更高效快捷完成作品。这种思想存在于许多需求方之中，如果受到大众认可，会进一步导致配音行业的专业技能权重降低，需求方和制作者将会更关注交付速度和配音制作成本，而非传统意义上作品的“表演深度”。如果任由情况往该方向发展，配音行业的评价体系将会从“艺术专业能力主导”转变为“技术能力主导”。举例说明：同样是初出社会寻找工作的新人，如果是一位仅掌握配音技能，就算配音技能优秀的配音从业者，竞争力可能也会弱于精通 AI 参数调试和语音修复技术的专业人员，在游戏行业里的 NPC 配音、广告旁白等配音的选择上，绝大多数公司会选择语音合成和修复技术性更强的人员。笔者从配音行业的艺术和专业的保护角度出发，认为虽没有设想中一般严峻，但如果不加以干预，将会产生技能断层，很有可能会导致未来配音艺术的技能传承面临危机，因为当配音员缺乏练习基本功和配音技巧，且花费更多时间研究怎么用好用 AI 配音技术时，必然会忽视自我提升，最终导致难以掌握配音技巧，无法传承给后辈。

（五）AI 配音促进传统配音向“人机协作”新模式方向发展

AI 配音已然得到广泛被使用，其影响力不可小觑，但 AI 配音毕竟是人类开发的便利工具，并不是完全是一个要刻意打乱当下行业秩序的东西。如果 AI 配音想要在配音领域持续发挥光和热，且能够做到与传统配音员的配音方式和睦共处，就需要现在的配音从业者使用 AI 配音发展“人机协同”的新模式。人机协作的模式可以理解为，创作音频时，可以先采用 AI 生成基础配音，再由专业配音员进行情感和艺术方面的润色的方法，或类似的 AI 配音结合人工精修处理的模式，此类“AI 打底 + 人工精修”的新模式可以让配音音频的制作效率获得成倍提升，而且加入人工干预调整细节能确保情感表达和艺术性处理仍不丢失。可以说目前最好的模式即 AI 配音和人类协作的“人机协作”模式。

^[8] 其实早已有实践者践行“人机协同”的模式，比如在 AI 精品有声剧《凡人修仙传》中就有配音演员与 AI 配音合作打造作品的案例：2024 年配音演员赵乾景推出的 AI 有声剧《凡人修仙传》、游戏《赛

博朋克 2077》续作用 AI 还原已故配音演员声线的创作，均采用了 AI 配音与人类配音协同的模式。上述的两个例子都是人工智能 AI 配音和传统配音员配音结合的案例，是能够促进配音行业向好发展的新尝试。

五、结语

尽管目前 AI 配音仍有着多方面的技术不成熟，但 AI 技术仍是时下热点，许多公司的研发机构都在力求 AI 技术有更好的发展，能为人类提供更好的服务。人类社会需要给予 AI 时间以发展，等待这门技术不断革新，市场需求会不断促进 AI 配音技术的发展。同时，AI 的学习能力强，错误促使它们进步，AI 会依据神经网络模型在一定程度内自我总结，自我完善，也会详细地将存在的问题反馈给技术开发人员以促使自身进步。现如今的 AI 配音发展的迅猛趋势已然无法阻挡，但人机协作的新模式是当前配音行业走向繁荣的最佳途径。即使人工智能 AI 配音能相当还原地模仿人类的声线进行配音演绎，真正的艺术仍是闪耀发光的，爱美之心，人皆有之，人们爱听有艺术性的配音作品，而并非失真且僵化的声音表现。配音这门艺术的创作模式今后也许不再只是传统的、完全由配音艺术家们独立创作的模式，而是让配音员们获得从“只会配音”到“既会配音，也会用 AI 智能进行修音”的提升。^[9] 配音员们要利用好 AI 配音的优势进行更完美的配音艺术创造，才是配音艺术应该出现的美好结局。人工智能 AI 配音是技术驱动的必然趋势，但无论其有多便利和高效，也无法完全替代人类本身在艺术创作中的价值。配音行业需要通过人机协作以实现高效率与艺术创造的平衡，同时也要密切关注技术伦理和社会接受程度，以及配音行业的传承也是每一位从业者不可忽视的方面。这场配音行业变革的启示或许是：技术能复制声波的振动频率，但无法替代那些在话筒前颤抖的双手、调整呼吸时的片刻沉默，以及声带振动时传递的演绎者的灵魂。人声之所以动人，从来不只是因为精准的发声，而是那份无法被量化的生命质感——那本就应该独一无二的声音演绎。

利益冲突

作者声明，在发表本文方面不存在任何利益冲突。

参考文献

- [1] 戚凯. 人工智能新时代，繁华背后的挑战不可小觑 [J]. 世界知识, 2024(11): 14-17.
- [2] 刘慧君. AI 配音的优势与劣势探究——以《创新中国》中解说词配音的功能为例 [J]. 新闻研究导刊, 2023(07): 7-9.
- [3] 李林翰. 配音艺术的类别及发展现状 [J]. 四川文理学院学报, 2024(03): 126-130.
- [4] 肖迪. 人工智能配音对配音行业的影响分析 [J]. 声屏世界, 2024(20): 36-38.
- [5] 杜成菲. 人工智能技术在微课配音中的应用 [J]. 电声技术, 2024(2): 71-73.
- [6] 常亮. 人工智能技术在影视配音艺术中的应用研究 [J]. 河南经济报, 2024(10): 1-2.
- [7] 韩志浦. 智能语音技术下播音主持的探索与思考——以微软 AI 配音为例 [J]. 文学艺术周刊 2023(16): 80-82.
- [8] Sreenivasulu Reddy Mogal i, Olivia Ng; Jia Xin Tan, Thu Htet San, Kian Bee Ng. Voice - over anatomy lectures created by AI - voice cloning technology: A descriptive article [J]. [8] Anatomical Sciences Education. Volume 17, Issue 9. 2024. PP 1686-1693.
- [9] Yuyun Zhang; Xiangzhe Cui; Yiming Wang. Innovative Exploration of the Teaching Mode of Dubbing Major in the Era of Artificial Intelligence. [J] Journal of Educational Research and Policies. Volume 6, Issue 7. 2024. PP 100-104.