

多模态数据融合驱动下傩戏人物动态造型的 AI 设计机制

李武建 符柯

贵州民族大学 贵州贵阳 550025

摘要: 傩戏作为中国古老的民间祭祀戏剧, 具有独特的视觉造型和深厚的文化内涵。本文基于多模态数据融合技术, 构建了傩戏人物动态造型的 AI 设计机制, 系统整合图像、文本、音频与动作等多种数据类型, 通过语义抽取、风格生成与动作建模实现对传统人物造型的动态再现。研究提出一套包含用户反馈优化与多维评估的机制, 以提升生成结果的文化表达力与动态表现力。实践表明, 该机制在视觉还原与语义契合方面均取得良好效果, 为非遗数字化提供了新路径。

关键词: 傩戏; 多模态数据融合; 人工智能设计; 动态造型; 文化数字化

傩戏作为中国古老的民间祭祀戏剧, 承载着深厚的民族精神与地域文化象征, 其人物造型不仅体现了神灵信仰、驱邪镇煞等文化意义, 更通过夸张的面具、仪式化的动作与独特的服饰风格构成了一种高度程式化的视觉系统。随着时代变迁和数字媒介的普及, 传统傩戏的艺术表现正面临着传播断层、审美淡化和年轻群体认知缺失等挑战, 传统造型语言难以有效延续与拓展。与此同时, 人工智能技术, 特别是多模态数据融合方法的发展, 为传统文化的数字再造提供了新的可能。通过整合文本、图像、音频、动作等多源异质信息, AI 不仅能够理解傩戏人物的造型逻辑, 还能在此基础上进行风格生成与动态演绎, 从而拓展传统造型的表现边界。本文正是在这一背景下展开研究, 探讨如何通过 AI 驱动机制对傩戏人物进行动态造型设计, 旨在为传统非遗文化注入智能化表达的新路径, 并构建文化理解与技术演化之间的桥梁。

1 多模态数据在傩戏造型设计中的集成路径

傩戏作为一种具有浓厚仪式性的民间戏剧形式, 其人物造型在表达宗教观念与社会情绪方面具有极强的象征性。不同于一般戏曲角色的写实性, 傩戏人物常通过面具、服饰、身体动作与音乐节奏的组合来完成对“神”与“鬼”的艺术表现, 具有高度抽象与程式化的特征。然而, 这种以形象化语言承载文化隐喻的表达方式, 在数字化语境中常常面临误读、简化甚至遗失的问题。因此, 在进行 AI 驱动的傩戏人物造型生成之前, 必须构建能够反映其文化复杂性的多模态数据基础, 并厘清不同模态间的语义映射与功能定位。

在数据类型上, 傩戏人物的造型信息并非单一图像即

可还原, 而需从多个维度协同分析: 视觉图像主要指传统面具、服装、化妆与动态动作的视频资料; 文本文义则包括角色设定、剧本叙述、祭祀背景等描述性信息; 声音模态涵盖了角色唱词、锣鼓节奏、喊声、环境音等音频资料; 而动作数据可借助姿态识别系统(如 MediaPipe 或 OpenPose)对表演者动作进行三维捕捉, 提取动态特征。为了构建有效的训练样本集, 本研究在贵州沿河、湖南新化等保留傩戏传统的区域开展实地采集, 结合地方志、民间艺人口述与舞台演出录像, 形成图像 3000 余张、文本材料 50 余万字、语音数据约 200 分钟的初步样本库。

在多模态数据的融合方式上, 研究采用了联合表示(Joint Representation)的思路, 借助 Transformer 结构将不同模态的信息嵌入统一的语义空间。例如, 以“钟馗”为原型的角色, 其文化意涵为“镇煞驱邪、正气浩然”, 通过图像可获取其怒目圆睁的面部特征、披发长须的头饰造型; 文本则提供了“天师下凡”“威震鬼魅”等描述语义; 声音中则体现为厚重低沉的台词与急促有力的鼓点。这些信息在融合过程中, 通过构建多模态注意力机制将其关联, 最终形成一个具有“神性驱动”标签的高维语义向量, 可作为 AI 模型进行动态造型生成的关键参数。

此外, 多模态技术在提升傩戏造型生成准确性的同时, 也有助于突破传统静态图像生成的限制, 使生成结果具备动态化、语境化特征。例如在生成“跳神者”时, 通过识别“跳、顿、转、俯”等动作序列, 并与音乐节奏同步生成连续动态帧, 赋予人物造型以时间逻辑和仪式氛围。再如, 在处理“土地神”这类庄重、静态角色时, 模型则自动抑制过多动作,

增强面部肃穆表情的细节描绘,体现文化语义与视觉造型之间的内在一致性。

2 AI 驱动下的傩戏人物动态造型设计流程

在多模态数据基础搭建完成之后, AI 驱动的傩戏人物动态造型设计流程便成为技术落地的核心环节。由于傩戏造型不仅涉及静态外观风格的再现,更关乎动作、神态与文化语义的综合表达,因此必须构建一个以“语义识别—风格生成—动态模拟—反馈优化”为核心的设计机制。

2.1 设计机制框架构建

AI 驱动的傩戏人物动态造型并非简单的图像生成过程,而是需要整合语言、图像、动作等多种模态信息进行结构化建模与动态表达。本研究提出的 AI 设计机制共包括五个主要模块:数据预处理、语义抽取、风格迁移、动态造型生成与反馈优化。具体而言,数据预处理环节主要负责标准化输入图像、动作轨迹及文本描述,确保其可被神经网络高效识别与处理;语义抽取阶段通过自然语言处理技术(如 BERT)识别剧本中的人物特征关键词及文化意涵;风格迁移部分依托 StyleGAN 实现对傩戏面具、服饰风格的仿真重建;动态造型生成则采用基于时序控制的动作预测网络(如 MoGlow 或 PoseRNN),将角色动作动态与生成图像匹配;最后,系统还引入了用户反馈驱动的微调机制,用于提升生成结果的人物真实感与文化表现力。

该设计流程的创新之处在于打破了图像风格生成与动作控制的割裂状态,而是以“语义驱动”为纽带,将文化意象作为控制变量嵌入整个造型流程,使得 AI 输出的人物既具备形态上的拟真,又能体现精神与功能上的契合。

2.2 关键模型与技术路线

为实现傩戏人物复杂造型与动态的生成,本研究选用了多种 AI 模型进行协同搭建。在文本语义抽取方面,采用了双向 Transformer 结构(如 BERT 或 T5)进行中文文本的语义编码,重点识别出角色性格标签(如“庄严”、“狰狞”、“威武”)以及动作提示词(如“登台”、“驱鬼”、“旋转”);在图像生成方面,基于 StyleGAN2 架构进行傩戏面具与服饰风格的图像生成,同时加入了风格特征约束模块,以便于控制面具夸张度、色彩饱和度与服饰对称性等关键元素。

动态动作生成模块采用的是结合图像生成器的 PoseNet 和动作生成网络 MoGlow。以“跳神者”为例,系统首先根据输入文本中识别出的动作关键词“跳跃、扭身、前冲”生

成初步动作轨迹,然后在生成面具与服饰图像的过程中,同步控制人物四肢位置与服饰飘动状态,确保整体视觉的一致性与动态流畅性。为了避免动作突兀或失真,模型中还引入了动作正则化算法,通过调节动作幅度与速度变化曲线,使生成结果贴近真实傩戏表演状态。

此外,在图像渲染阶段,为实现面具光影细节与服饰褶皱的真实还原,还使用了神经渲染技术(Neural Radiance Fields, NeRF)对图像进行细节增强处理。借此,在不增加大量计算成本的前提下,显著提升了最终输出图像的表现力。

2.3 案例实施:三类典型傩戏人物动态造型生成

为验证 AI 设计机制的有效性,本文选取了三类在傩戏中具有代表性的角色:“牛头马面”、“土地神”与“跳神者”进行实验建模。这三类角色分别代表了威吓驱邪、庄严稳重与动作灵活三种典型傩戏表现类型,涵盖了从静态造型到动态动作的多种文化语义。

“牛头马面”造型具有极强的象征性,其头饰巨大、眼神凶狠、服饰对称且配色强烈。系统通过输入关键词“勾魂、凶猛、威吓”引导图像生成网络强化面部张力与颜色对比度,同时结合锣鼓声轨节奏控制动作生成模块,产生缓慢、威严的行走与转头动作,以强调其镇魂属性。

“土地神”则在视觉上更为稳定与厚重,其特点是面具小、五官平和、动作缓慢、语音沉稳。模型在训练过程中,抑制了大幅动作生成,通过动作网络添加低频抬手、转身等细节,同时提升面部表情中眉眼的静态表现,使其造型更贴近真实戏台表现。

“跳神者”是动作最为复杂的角色,常以连续快速的舞蹈动作配合鼓点完成驱神仪式。模型通过采集多段舞蹈视频,并结合文本提示“连跳、旋转、下蹲”,生成动态骨架序列,之后与面具图像叠加合成高帧率动态造型。实验中发现,跳神者动作生成质量的提升,主要得益于将动作模态与声音节奏进行同步建模,这一策略也为其他动态文化造型生成提供了借鉴。

3 傩戏造型 AI 生成的效果评估、优化机制及未来展望

3.1 动态造型生成效果的评估维度

傩戏人物的 AI 造型生成不同于一般静态图像处理,其结果不仅要求在外观上贴近传统造型,更要求在动作动态与文化语义表达上具备高度一致性。为此,本研究构建了以造型还原度、动态连贯性与文化语义准确率为核心的三维评估

体系。在造型层面,利用结构相似度指数(SSIM)和感知损失函数对生成面具与服饰图像进行相似度分析,确保关键特征如“怒目”“披发”等得以准确还原;在动态层面,通过光流检测与轨迹平滑度算法,检验生成动作的连贯性与节奏合理性,特别是在“跳神者”等高频动作角色中显现出明显效果;而在文化语义方面,结合文本—图像匹配机制与专家人工评分方式,评估生成结果是否准确体现诸如“威严”“驱邪”“守护”等传统文化内涵。综合多项实验表明,模型整体生成质量处于可接受范围,平均文化语义匹配评分达到4.2分(满分5分),基本符合傩戏艺术再现的标准需求。

3.2 用户反馈驱动的优化机制

为了提升模型适应性并满足用户多样化审美需求,本研究构建了基于用户反馈的闭环优化机制,使AI生成系统从单向输出走向交互进化。用户可通过平台对生成的傩戏造型在视觉形象、动作节奏、情绪表达等维度进行打分或文字反馈,系统则依据这些信息实时调整风格生成器或动作控制模块的参数配置。例如,当多位用户反馈“跳神者”动作表现不够生动时,系统会自动提升训练集中高频跳跃样本的权重,优化动作幅度和连续性;若面具造型被认为“神韵不足”,则在StyleGAN的风格迁移层中增强面部表情张力的生成倾向。通过这种反馈—调整—再输出的快速迭代机制,模型不仅不断提高生成精度,也逐渐形成对不同傩戏角色与风格偏好的适应能力,从而实现用户参与下的文化智能生成共创,进一步提升系统的实用性与文化表达力。

3.3 存在问题与改进方向

尽管现有AI造型系统在动态生成与文化还原方面取得初步成果,但仍存在一些技术与语义层面的不足,亟待改进。一方面,动作生成模块对复杂舞步和肢体组合动作的细节捕捉仍显不足,特别是传统傩戏中强调节奏变化与仪式感的部分,现有网络难以精确还原其复杂节奏关系;另一方面,AI对文化语义的理解仍存在“浅表化”倾向,诸如“土地神”“山神”等角色在生成过程中可能因训练样本不足或语义泛化而产生形象偏差,降低文化准确性。此外,在风格表现上也存在过拟合风险,部分模型训练结果中人物表情趋于模板化,缺乏风格变化的张力。为解决上述问题,未来可通过构建多层次傩戏文化知识图谱、引入更精细的动作捕捉数据以及优

化训练样本结构,提高模型在文化理解力、动作表达能力和风格生成多样性等方面的综合表现,推动傩戏造型AI生成朝向更真实、更灵活、更具人文深度的方向发展。

4 结论

本文围绕多模态数据融合驱动下的傩戏人物动态造型AI设计机制展开研究,构建了从数据采集、语义抽取、风格生成到动态输出与反馈优化的完整技术流程。通过引入图像、文本、音频与动作等多源信息,实现了对傩戏造型在视觉、动作与文化语义上的高度还原,有效提升了传统非遗形象的数字化表达能力。同时,用户反馈机制的引入为模型持续优化提供了路径,推动了人机共创在文化遗产领域的应用。未来,仍需在动作细节建模与文化语义理解上进一步提升,以实现更具表现力与情感温度的动态傩戏人物生成,为AI参与传统艺术再造开辟更广阔的空间。

参考文献:

- [1] 石宗燕,李晓盈.傩戏文化的传承与实践途径浅析[J].丝网印刷,2024,(12):92-94.DOI:10.20084/j.cnki.1002-4867.2024.12.026.
- [2] 应佳鑫.非物质文化遗产的活态传承——以池州傩面具的数字化传播为例[J].艺术市场,2025,(01):91-93.
- [3] 赵紫薇,杨先艺.数字媒介语境下武安傩戏文化基因模型的构建与应用[J].创意与设计,2025,(01):65-73.
- [4] 李志远.傩戏数字传播对乡村文化建设的启示[J].兰州文理学院学报(社会科学版),2025,41(01):2-8.
- [5] 汪先敏.基于新媒体背景的傩戏APP界面交互设计研究[J].印刷与数字媒体技术研究,2024,(S1):1-6+17.
- [6] 聂国红.梅山傩戏的文化特征和艺术价值[J].戏剧之家,2025,(02):54-56.
- [7] 王忠,舒怡鸣.叙事视角下的地域特色非遗文化再设计路径——以辰州傩戏为例[J].美术教育研究,2024,(23):38-40.
- [8] 桂雪儿.池州傩戏的传播困境及破局策略研究[J].戏剧之家,2024,(35):18-20.

基金项目: AI技术赋能非遗屯堡傩戏人物角色的数字化设计与设计策略研究,项目编号: 25GZGXRWJD0124。