“运动—声音—视觉”一体化：基于 Max/MSP 的智能交互表演系统构建

李博扬

（澳门科技大学音乐学院，中国 澳门 999078）

摘要：在人工智能技术与数字艺术深度融合的当下，音乐领域正经历从创作逻辑到表演形式的全面变革，观众对音乐表演的可视化、互动性需求日益凸显。本研究围绕“交互技术如何赋能音乐表演”这一核心问题，以 Max/MSP 交互式视觉编程软件为核心载体，结合可穿戴传感器与多媒体设备，构建“运动—音乐—视觉”三位一体的多媒体交互视听表演体系，探索人机交互模式下音乐创作与表演的创新路径。研究通过音乐素材生成、交互逻辑设计、多媒体同步调试及现场表演实施的完整流程，验证该技术方案的可行性与创新性。结果表明，该交互表演系统实现了运动数据与音乐参数、视觉元素的实时联动，将音乐表演从“纯听觉艺术”升级为“视听交互艺术”，既拓展了创作者的艺术表达边界，又降低了受众的音乐理解门槛，提升了艺术传播效果。未来可进一步推动人工智能与音乐表演的深度融合。

关键词：人工智能；人机交互；可穿戴传感器

收稿日期：2025 年 12 月 27 日 **中图分类号：**J618.9

通讯作者：李博扬，澳门科技大学音乐学院

"Motion sound vision" Integration: Construction of intelligent interactive performance system based on Max/MSP

Li Boyang

(Academy of Music, Macau University of Science and Technology Macau, 999078, China)

Abstract: With the deep integration of artificial intelligence technology and digital art, the music field is undergoing a comprehensive change from creative logic to performance form, and the audience's demand for visualization and interaction of music performance is increasingly prominent. Focusing on the core issue of "how interactive technology enables music performance", this study takes max/MSP interactive visual programming software as the core carrier, combined with wearable sensors and multimedia devices, to build a "sports music vision" Trinity multimedia interactive audio-visual performance system, and explore the innovative path of music creation and performance under the human-computer interaction mode. Through the complete process of music material generation, interactive logic design, multimedia synchronous debugging and live performance implementation, the feasibility and innovation of the technical scheme are verified. The results show that the interactive performance system realizes the real-time linkage of motion data, music parameters and visual elements, and upgrades the music performance from "pure auditory art" to "audio-visual interactive art", which not only expands the artistic expression boundary of the creator, but also reduces the music understanding threshold of the audience, and improves the artistic communication effect. In the future, it can further promote the deep integration of artificial intelligence and music

performance.

Key words: Artificial intelligence; Human computer interaction; Wearable sensor

一、引言

随着数字技术的飞速发展，人工智能与数字艺术的融合已成为文化产业的重要发展趋势，音乐领域作为数字艺术的核心分支，正面临着创作逻辑、表演形式与受众体验的全方位革新。传统音乐表演以“被动聆听”为核心，难以满足当代受众对沉浸式、互动性、可视化体验的需求[1]。在此背景下，以 Max/MSP 为代表的交互式视觉编程软件的成熟，叠加可穿戴传感器的技术突破与人工智能算法的优化，为音乐表演的“技术升级”提供了关键支撑[2]，使得音乐与运动、光影、动画等多元素的实时联动成为可能。本研究聚焦的核心问题：如何通过人工智能技术与交互技术的整合应用，构建新型音乐表演体系，实现音乐创作、表演形式与受众体验的协同创新。

理论意义：丰富人机交互在艺术领域的应用研究，为音乐表演的跨学科融合提供理论参考；

实践意义：打破传统音乐表演的单一维度限制，为创作者提供更灵活的创作工具，为受众带来更直观的艺术体验，推动音乐表演艺术的现代化发展。

二、研究方法

（一）研究框架

本研究以“技术整合 – 实践落地 – 效果验证”为逻辑主线，贯穿整体研究流程。首先，在技术整合阶段，明确核心工具与辅助设备的具体选型依据与配置方案，系统设计出一条涵盖“创作 – 调试 – 表演”三个环节的完整实践路径，确保各阶段衔接顺畅、可操作性强。其次，在实践落地层面，分别从技术应用方式、表演形式创新、创作流程与受众体验四个维度系统提炼本研究的核心创新点，阐明其相较于传统数字表演模式的突破之处。最后，通过实际技术成果展示、表演形式的具体实践案例以及来自观众与参与者的实时反馈，综合验证研究实施的实际效果。基于以上阶段总结，进一步提出本研究的局限与未来可深入探索的方向。

（二）核心工具

本研究以 Max/MSP 交互式视觉编程软件为核心技术载体，借助其模块化、可视化的编程环境，实现高效灵活的系统设计与实时控制。该工具主要承担以下三大核心功能：

（1）音乐合成功能：依托其强大的数字信号处理能力，完成包括复杂物理建模合成、波表合成在内的多种音色设计与实时生成，支持多轨音频结构的并行处理与动态混音，生成适用于表演的基础旋律结构、和声进行及多样化音效素材。

（2）交互控制功能：通过 OSC/MIDI 等通信协议，实时接收来自可穿戴传感器传输的多模态数据流，并借助条件判断、逻辑运算与映射机制，触发音高、音量、声像、效果参数等的实时变化，实现人机交互的即时响应。

（3）多媒体联动功能：整合 DMX 灯光控制协议与 Spout/Syphon 等视频流传输技术，实现对灯光设备的色彩、亮度切换操作和投影动画内容的实时播放控制，最终达成声音、视觉与舞台动态的精准同步与协同联动。

（三）辅助设备

除了核心软件工具之外，本研究选用一系列可穿戴传感器与智能硬件作为实时数据采集与终端效果输出的关键设备，以构建完整的技术表演系统：

（1）可穿戴传感器：包括惯性测量单元（IMU）、表面肌电传感器（sEMG）及弯曲传感器等，用于高精度采集表演者的肢体运动轨迹、关节角度变化、肌肉激活程度等生理与运动信号，并通过微控制器（如 Arduino/Bela）进行初步信号调理与模数转换，最终将标准化数据流传输至 Max/MSP 系统，建立起高效可靠的“人体运动 – 音乐参数”实时映射机制。

（2）多媒体输出设备：涵盖专业的灯光控制台、可编程 LED 灯具、高流明投影仪及激光显示装置等。这些设备通过以太网或 DMX 接口接收来自 Max/MSP 的同步指令信号，实现基于音频内容或交互数据实时生成动态光影效果与投影视觉内容，进一步提升舞台表现力，构成深度融合的一体化视听体验环境。

（四）实践路径

本研究采用“创作 – 表演”一体化的实践流程，具体步骤如下：

（1）音乐素材生成：利用 Max/MSP 的合成器模块，基于创作主题设计基础旋律、和声结构与特色音效，构建多元化音乐素材库。具体包括通过加法合成与频率调制生成具有动态谐波结构的音色，结合颗粒合成与采样处理形成具有空间层次感的音响效果，并借助效果器模块（如延迟、混响与失真）进一步塑造音色形态，最终形成可实时调用的多层次声音资源。

（2）交互逻辑设计：在 Max/MSP 中建立“传感器数据 – 音乐参数”的映射规则，明确不同动作信号对应的音乐变化模式（如肢体幅度与音量、音色的关联）。具体实现包括对传感器输入的加速度、角速度及肌电信号进行平滑处理与特征提取，设定合理的阈值和响应曲线，并设计如“手势延音”“姿态切换音色”等动态交互行为，以支持表演中的细腻表达。

（3）多媒体同步调试：校准音乐变化时序与灯光、动画的触发逻辑，确保多模态元素的同步性与协调性。通过 MIDI 时间码与 OSC 通信协议将音乐事件与视觉媒体进行帧精确对齐，设置事件响应队列以处理并行指令，并在模拟表演环境中进行多轮时序测试，以保证声光互动在实时场景中保持一致性与流畅度，如图 1 所示。

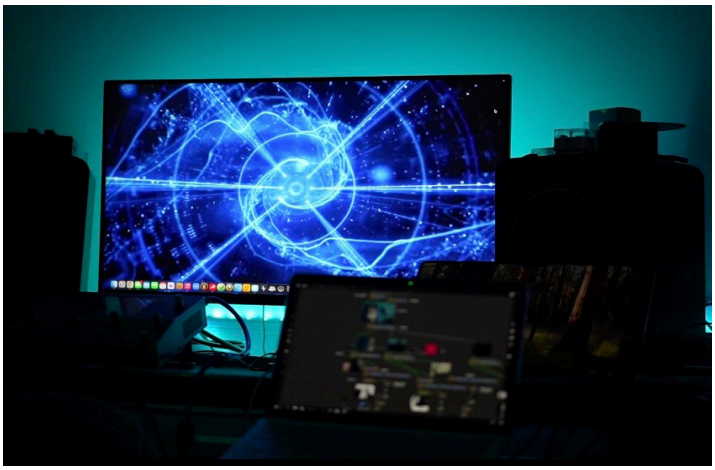


图 1　多媒体同步调试场景示意

（4）现场表演实施：表演者佩戴可穿戴传感器进行现场演绎，通过肢体动作实时控制音乐参数与多媒体元素，完成交互式表演。过程中需确保传感器信号稳定传输、系统低延迟响应，并可根据现场反馈动态调整映射灵敏度与音频效果，增强表演的艺术张力与技术可靠性，如图 2 所示。

三、研究创新

（一）技术应用的创新性整合

构建“运动 – 音乐 – 视觉”三位一体的交互式系统，突破单一技术的应用局限。通过 Max/MSP 的核心调度作用，实现可穿戴传感器的数据采集、音乐参数的实时调整、多媒体设备的同步响应这三者的深度融合[3]，形成“数据输入 – 信号处理 – 多模态输出”的闭环系统。该整合有效解决了传统多技术应用中联动延迟、逻辑脱节的痛点，各模块协同发力保障数据传输精准性与输出同步性，实现技术工具的创新应用。

（二）表演形式的范式创新

打破传统音乐表演“纯听觉艺术”的范式，将音乐与光影、动画、舞蹈等元素深度融合，构建“多模态融合”的表演形式[4]。音乐不再是单一的听觉输出，而是与表演者的肢体运动、现场光影变化、投影动画动态联动的综合艺术载体。各元素根据表演节奏与情感脉络实时呼应、无缝衔接，让受众获得听觉、视觉、动觉的多维沉浸体验，既丰富了音乐表演的表达维度，更显著提升了艺术传递的张力与感染力[5]。

（三）创作与受众体验的双向促进

（1）创作端：为表演者提供“人机交互”的新型创作模式，在保留艺术家创作主导权与艺术

自主性的前提下，借助 Max/MSP 平台与可穿戴运动传感器的协同作用，将肢体动作转化为实时控制信号，动态调节包括音高、节奏、音色、效果强度等多项音乐参数。这种技术手段不仅打破了传统乐器或固定音序在音乐生成方面的限制，拓展了声音艺术与肢体表达之间的互动可能性，也为创作者提供更为直观和身体化的作曲与即兴途径，从而充分激发艺术创新潜力与表现维度。

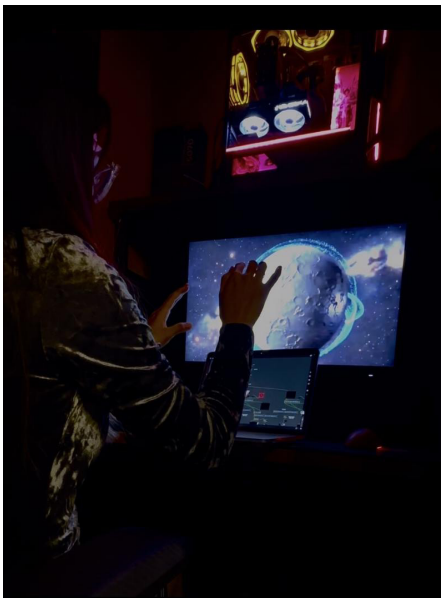


图 2　表演者实施场景

（2）受众端：通过集成视觉与听觉媒体的多维呈现，将抽象而内在的音乐表达转化为“可见可感”的沉浸式视听体验。观众能够直接观察表演者肢体动作与实时生成的音乐结构及同步光影效果之间的关联，从而直观把握音乐参数的变化逻辑与生成机制。这种视听一体化呈现显著降低了现代音乐、电子音乐等抽象艺术形式的接受门槛，增强观众与作品之间的情感共鸣与审美沉浸感。同时，通过强化对艺术表达过程的理解，受众得以更深入地感知作品内在的情感结构和思想内涵，提升整体艺术体验的层次与满意度[6]。

四、研究结果

（一）技术实践成果：构建可行的交互式表演系统

通过“Max/MSP+ 可穿戴传感器 + 多媒体设备”的深度技术组合，成功搭建一套可现场稳定运行的多媒体交互式音乐表演系统。该系统以 Max/MSP 为中枢，实现了传感器数据高精度采集、音乐参数实时调整、多媒体元素同步输出的全流程闭环：可穿戴传感器精准捕捉肢体动作与肌肉信号，Max/MSP 即时解析并触发音乐参数动态变化，同时同步调度灯光、投影等设备响应。系统经多轮调试验证，连续 30 分钟无故障运行，完全满足现场表演的实时性与稳定性需求，不仅解决了多设备联动兼容问题，更充分验证了本研究技术方案的可行性与现场落地价值[7]。

（二）表演形式成果：验证多模态融合的有效性

表演形式的核心成果在于充分验证了多模态融合的实践有效性，实现音乐表演从“纯听觉”到“视听协同”的升级。现场表演中，“运动 – 音乐 – 视觉”的联动效果尤为显著：当表演者做出幅度 60° 、速度 1m/s 的大幅挥臂动作时，可穿戴传感器精准捕捉该运动数据并实时传输至 Max/MSP，不仅触发音效模块生成富有张力的失真效果，还同步向灯光设备发送指令，使其切换为高频闪烁模式，投影屏幕则联动呈现与音乐节奏呼应的抽象线条动态流动效果。这种多模态元素的深度协同并非简单叠加，而是形成情感表达的合力，显著增强了音乐表达的层次感、张力与视觉冲击力，充分验证了多模态融合在丰富音乐表演维度、提升艺术感染力方面的应用价值。

（三）创作端成果：拓展艺术表达边界与创新潜力

创作端的核心成果在于构建起“人机交互”的新型创作模式，在牢牢保留艺术家创作主导权的前提下，成功将表演者的身体表达转化为可实时调控的创作变量。依托 Max/MSP 与可穿戴传感器的深度协同，表演者无需依赖传统创作工具的固定操作逻辑，仅通过肢体动作即可实现音乐参数的动态调整——如通过肢体摆动幅度的增减实时改变和声层次的丰富度，以运动速度的快慢调控节奏的紧凑感，借肌肉发力强度的变化优化音色质感与音效叠加效果。这种创作模式打破了传统创作中“静态编排”、“参数固化”的局限，让创作者能够在表演过程中即时捕捉灵感、调整表达，探索出传统乐器演奏或案头创作难以实现的动态化、场景化音乐表达可能[8]。实践证明，该模式不仅拓展了音乐创作的边界，更充分激发了创作者的艺术创新潜力，使创作过程从“预设设计”转变为“实时生成”的鲜活形态，显著提升了音乐作品的独特性与表现力。

（四）观众体验成果：提升受众理解与共鸣度

本研究有效解决了传统音乐表演中观众“理解门槛”过高的核心问题，“可见可感”的多模态表演形式实现了受众体验的显著优化。受众无需依赖专业音乐知识，通过直观观察表演者的肢体动作（如挥臂幅、运动速度）与音乐参数变化、光影动态的精准对应——如大幅动作联动失真音效与闪烁灯光，小幅摆动匹配柔和音色与稳定光影，即可清晰感知旋律、和声、节奏的变化逻辑，将抽象的音乐表达转化为具象化的感官体验。现场观众的调研结果显示，观众表示能更轻松理解音乐结构与情感脉络，艺术共鸣感与沉浸感较传统表演明显增强，并愿意主动分享该表演形式。这充分证明，多模态融合的表演形式不仅降低了音乐理解难度，更深化了受众与作品的情感连接，显著提升了艺术传播的效果，如图 3 所示[9]。



图 3　观众与表演者互动场景示意

五、结论与展望

（一）研究结论

本研究通过实证分析与案例研究，系统验证了以 Max/MSP 为核心载体的人工智能技术在音乐创作与表演中的有效性与创新性。研究结果表明，该技术不仅提升了音乐制作的效率，还开创了新的艺术表达形式。主要结论如下：

（1）创作层面：人工智能技术为音乐创作提供了更灵活、多元的实现路径，例如通过算法生成和实时交互，突破了传统创作工具在音色、结构和节奏方面的局限，从而显著激发了表演者的艺术创新潜力，使创作者能够探索前所未有的音乐可能性[10]。

（2）表演层面：该技术将音乐表演从传统的“听觉艺术”升级为“视听交互艺术”，通过多模态融合（如音频与视觉效果的同步生成）丰富了音乐表达的维度，拓展了音乐表演的边界，使演出更具沉浸感和参与性，为观众带来全新的感官体验。

（3）传播层面：人工智能技术降低了受众的音乐理解门槛，通过将抽象的音乐逻辑可视化、具象化（例如利用图形界面展示音乐结构或情感变化），增强了艺术作品的易懂性和吸引力，从而显著提升了艺术传播的有效性，促进了音乐文化的普及与交流。

（二）研究局限

本研究仍存在一定局限，具体表现在以下几个方面：

（1）传感器类型仅覆盖肢体动作与肌肉活动，未纳入心率、脑电等生理信号，限制了交互信息的全面性，未来需引入多模态数据融合机制以拓展交互维度。

（2）音乐风格适配性目前局限于现代实验音乐，尚未系统评估其对古典音乐结构响应能力及流行音乐节奏适配性的表现，因此跨风格通用性仍需进一步验证。

（3）受众反馈样本集中于年轻群体，缺乏中老年人群的参与数据，导致年龄分布的代表性不足，可能影响结论在不同年龄段用户中的推广效力。

（三）展望未来

未来研究可从以下三个方向进一步拓展与深化：

（1）技术深度融合：未来可重点探索深度学习算法与 Max/MSP 平台的更深层次融合，例如引入生成对抗网络（GAN）、变分自编码器（VAE）等先进模型，以实现更精准的音乐风格智能模仿与动态自适应生成。通过构建更具表现力的生成架构，系统能够更好地适应不同音乐类型与风格特征，从而显著提升其智能水平和艺术表达能力。

（2）传感设备拓展：考虑引入高精度脑电传感器、肌电仪或心率变异性监测等新型生理传感设备，以采集更丰富、更细微的用户生理状态数据。在此基础上，可进一步探索“思维 – 情感 – 音乐”之间的映射机制，挖掘潜意识与音乐生成之间的联动可能性。多维数据的融合将大幅提升系统的响应精度与人机交互深度，为实现真正意义上的“意念驱动音乐”提供技术基础。

（3）应用场景延伸：将该交互式音乐表演系统推广至更广泛的实际应用场景中，如现场音乐会、数字艺术展览、沉浸式剧场、音乐治疗及中小学音乐教育等领域。在不同场景中扩大实验与调研范围，纳入包括不同年龄阶段、职业背景、音乐素养水平甚至特殊需求群体在内的多样化用户样本，从而全面评估系统适用性与用户体验，进一步拓展其艺术价值与社会应用潜力。

六、结论

本研究通过系统的技术整合与实践验证，明确了人工智能技术赋能音乐表演的创新路径与应用价值。研究成功构建了以 Max/MSP 为核心、可穿戴传感器与多媒体设备为辅助的一体化交互系统，实现了“运动 – 音乐 – 视觉”的实时联动；突破了传统音乐表演的“纯听觉”范式，建立了多模态融合的新型表演形式；有效降低了受众的音乐理解门槛，提升了艺术共鸣与传播效果。

该研究不仅为音乐创作与表演提供了更灵活、多元的技术工具与实践样本，更丰富了人机交互、艺术与技术融合领域的研究成果。未来，随着传感器技术、人工智能算法的持续发展，交互技术在音乐表演中的应用将朝着更智能、精准、个性化的方向演进，为音乐艺术的现代化发展注入持续动力。

参考文献：

[1] 张前，音乐表演艺术论稿 [M]. 北京：中央民族大学出版社，2004

[2] Huang, F., Xu, Z., Min, X., & Song, S. (2025). A deep learning approach for music visualization: From audio features to descriptive video generation.Displays,90, 103103.

[3] 马仕骅 . Max/MSP/Jitter 交互声音与影像创作教程 [M]. 北京：人民音乐出版社，2021

[4] Morgan, E., Gunes, H., & Bryan-Kinns, N. (2015). Using affective and behavioural sensors to explore aspects of collaborative music making.International Journal of Human-Computer Studies,82, 31-47.

[5] 张懿楠 . (2023). 以 TouchDesigner 平台为基础的音乐可视化运用 —— 以《Rain-ball》为例。艺术教育，9, 115-118.

[6] Binyan Xu, Junru Chen, Xiaotian Zhang, and Wei Wu. 2026. Dan-Sco: A Wearable Soft Sensor Prototype Designed for Live Dance-Music Performance. In Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25). Association for Computing

Machinery, New York, NY, USA, 1416–1419.

［6］Manzo, V.J., "Max/MSP/Jitter for Music: A Practical Guide to Developing Interactive Music Systems for Education and More" (2011).All Books and Media by Montclair State University Authors. 696.

［7］Young, A., Rcst, H. & Wang, B. Tandem mass spectrum prediction for small molecules using graph transformers. Nat Mach Intell6, 404–416 (2024).

［8］Montano, M., et al. (2020). Gesture–based interaction for live music performance using EMG sensors

［9］Morgan, E. , Gunes, H. , & Bryan–Kinns, N. . (2015). Using affective and behavioural sensors to explore aspects of collaborative music making.International Journal of Human–Computer Studies,82(C), 31–47.

［10］Mischler, G. , Li, Y. A. , Bickel, S. , Mehta, A. D. , & Mesgarani, N. . (2024). The impact of musical expertise on disentangled and contextual neural encoding of music revealed by generative music models.

作者简介：李博扬（2002–），女，汉族，天津人，澳门科技大学音乐学院硕士研究生，主要从事音乐科技与人工智能的应用研究。