从三层混合逻辑到生成产线：AI 视觉生成的工程实现与产业迁移

YunFei M. Wang
（王·Micheal·雲飛）
Independent Researcher

摘要

本文为上篇《结构约束先行的视觉生成：AI 绘画与视频生成的三层混合逻辑》的延续。上篇提出了三层混合生成逻辑的理论框架——三维矢量结构约束全局元素、二维投影逻辑约束单帧空间关系与因果连贯性、位图逻辑填充单帧局部细节。本文聚焦该框架在工程与产业层面的展开：生产流程重构、算力调用逻辑、跨领域借鉴与产业迁移路径。本文分析三层架构如何将生成过程从依赖灵感的单件手工转变为依赖规范的流水作业，讨论以创作者本地终端算力为核心的分布式算力支持产业生态，论证该逻辑对 AI 音乐生成与 AI 配音生成的借鉴意义，并探讨传统 3D 设计软件企业的经验优势与核心项目组的关键作用。本文属于方法论研究，不包含实证数据。

关键词

AI 生成；生产流程重构；算力调用；跨领域借鉴；产业迁移

1 引言

上篇提出了三层混合生成逻辑的理论框架。本文为下篇，聚焦该框架在工程与产业层面的展开。核心问题包括：三层架构如何重构生产流程？算力调用逻辑如何按层设计？该逻辑对 AI 音乐与 AI 配音有何借鉴意义？产业迁移路径如何展开？对创作者提出了什么要求？

本文属于方法论研究，不包含实证数据。三层混合生成逻辑为概念性方案，具体实现需相关领域研究者进一步推进。

1.1 结构安排

本文按以下序列展开：

- 第 2 章：生产流程重构与成本降低，分析前期工作的 AI 辅助、后期工作的全面 AI 化、人类角色转变、生产成本降低，以及确认锁定与延续机制。

- 第 3 章：算力调用逻辑与并发优化，分析三层算力特征差异、分层调度架构、异构算

力池配置、动态算力分配策略、未来创作工作的算力构成，以及算力池与潜存场的类比。

- 第 4 章：跨领域借鉴——AI 音乐与 AI 配音。
- 第 5 章：产业迁移路径与核心项目组的关键作用。
- 第 6 章：对创作者的要求与创作门槛。
- 第 7 章：应用前景与验证方向。
- 第 8 章：与状态管理、感知隔离的协作及游戏/具身智能价值。
- 第 9 章：结论。

2 生产流程重构与成本降低

2.1 前期工作的 AI 辅助

三层混合生成逻辑的前期工作——三维矢量结构的构建、全局元素设定集的制定、力量体系设定函数集（指明确全局元素的关联接口、操控接口、响应规则的函数集合）的推导——均可由 AI 辅助完成。

三维矢量结构的自动生成。创作者定义全局元素的描述性约束，AI 根据这些描述生成三维矢量结构。

全局元素设定集构建辅助。AI 根据创作者的意图，辅助构建全局元素的设定集——角色特征、道具属性、场景布局、线索性细节特征。

力量体系设定函数集 AI 辅助推导。对于涉及魔法、法术、异能等超凡力量的设定，AI 可根据创作者定义的规则框架，辅助推导力量体系的函数集。

2.2 后期工作的全面 AI 化

三层混合生成逻辑的后期工作——二维投影逻辑的执行、位图细节的填充、帧间因果连贯性的校验——可全面交由 AI 完成。

二维投影逻辑的自动执行。AI 根据三维矢量结构，自动计算镜头参数、光照参数、阴影参数、物理模拟参数。

位图细节的自动填充。AI 在位图生成层填充纹理、色彩、微观细节。

帧间因果连贯性的自动校验。AI 根据因果连贯性约束，自动校验帧与帧之间的上下文

因果关系与全局环境泛化响应。

2.3 人类角色转变：从操作者到设定者

在三层混合生成逻辑中，人类的角色从"全流程操作者"转变为"设定者"。人类只需定义"生成什么"与"遵循怎样的设定集"，后续工作可交由 AI 完成。灵感将作为创作原驱力与视角窗口，作用于创作初期的作品内容框架和大体脉络，而非约束具体内容情节细节。

2.4 生产成本的大幅降低

成本类型	传统流程	三层混合生成逻辑
人力成本	全流程操作	设定构建
时间成本	反复重试	一次生成
修正成本	事后补救	事前约束

AI 生成工具的进步不是简单地降低创作门槛，而是将门槛从技术操作转移到设定构建；净效果因创作者类型而异。本方案将结构一致性的保证从生成模型的学习任务转移至架构协议的约束执行，使第三层生成模型只需在约束空间内优化局部细节。这大幅降低了生成模型对全局结构一致性的学习负担，减少了对海量端到端配对数据的依赖，并降低了无效试错带来的算力消耗。需要指出的是，第三层生成模型仍需训练，第一层与第二层的工程集成存在一次性投入；本方案降低的是"结构一致性训练成本"，而非"全部训练与工程成本"。

2.5 确认锁定与延续

用户首次确认生成结果后，系统将第一层的结构参数、瑕疵参数五元组与结构关系函数固化锁定，作为后续生成的硬约束与初始条件。后续生成以锁定参数包为第一层输出，仅在用户显式要求时局部解锁修改，修改生成新版本并保留旧版本为分支。

锁定参数按变化的驱动因素写入分层状态，而非按属性名称归类（详见文献[13]第 4 章）。驱动因素是角色固有属性的，写入本底键（身份级）——基础身高、基本比例、核心识别特征，在没有遭遇改变它们之前保持稳定。驱动因素是角色遭遇的，写入经验键（场景级关系），可随遭遇演进——角色因成长、受伤、变形导致的身高变化，作为事件记录；若变化不可逆，升格为本底键中的新身份级数据。驱动因素是场景即时状态的，写入活跃键（当轮表现）——角色此刻蹲着、坐着、弯腰，视觉高度不同，但这是姿态变化，不是属性变化。同一属性可以在不同驱动因素下归入不同层

级：基础身高在本底键，遭遇导致的身高变化在经验键并可升格，场景姿态在活跃键。归类的依据是"变化由什么驱动"，不是"属性叫什么名字"。

这使首次确认的结果成为后续生成的稳定基础，而非每轮重新解析与检索。

3 算力调用逻辑与并发优化

3.1 三层算力特征差异

三层混合生成逻辑的算力需求并非均匀分布。三个层级在计算类型、并行性、缓存潜力上存在根本差异。

维度	第一层	第二层	第三层
计算类型	几何处理、约束求解	投影几何、光照、物理模拟	扩散/Transformer 推理
并行性	中	中	高
缓存潜力	高	中	低
可复用性	跨帧、跨镜头、跨场景、跨用户	同镜头内帧间部分复用	帧间基本不可复用

第一层是前期成本高、持续成本低；第二层是每镜头中等、持续；第三层是每帧高密度、持续。算力调用逻辑必须按层设计。

3.2 分层调度架构

用户请求→任务分解器→三层任务队列→算力调度器→异构算力池。

第一层：CPU 集群，缓存优先，异步构建。第二层：中端 GPU，按镜头调度，预调度平滑峰值。第三层：高端 GPU/TPU，按帧调度，批处理合并。

3.3 异构算力池配置

算力类型	适用层级	原因
CPU 集群	第一层	几何运算适合 CPU，全局结构构建频率低

中端 GPU	第二层	中等并行度，中端 GPU 足够
高端 GPU/TPU	第三层	高并行度，需要大显存与高算力

三层架构天然适配异构算力池，按层分配算力，大幅降低总体成本。

3.4 动态算力分配策略

策略	内容	适用场景
优先级调度	高优先级用户优先分配算力	付费用户、紧急任务
批处理合并	多个用户的帧任务合并批处理	大量并发用户
预调度	根据任务预测提前分配算力	可预测的工作负载
弹性伸缩	根据实时负载动态增减算力	负载波动大的场景
缓存优先	优先复用已缓存的全局结构	重复角色/场景的项目

3.5 未来创作工作的算力构成

未来创作工作的算力构成将以创作者本地终端算力为核心，形成分布式算力支持产业生态，确保创作项目的著作权与数据安全。本地终端算力为核心，意味着创作者的设定集、全局结构、核心资产保存在本地，不依赖云端。分布式算力生态提供弹性扩展，用于第三层逐帧渲染等算力密集环节。

3.6 算力池与潜存场的类比

算力池可类比为潜存场——所有可能生成的公共基底，持续动态，待调用。项目实现的设定集函数可类比为现实涌现时携带的框架全局内禀属性——生成时携带的约束条件，决定生成什么、如何生成。

这一类比不能帮助我们准确描述潜存场本身，但可以帮助我们理解算力池与设定集函数之间的结构关系。算力池不是潜存场，设定集函数不是框架内禀属性——它们在存在方式、可变性、本体论地位上都不同。但通过结构相似性，这一类比帮助我们理解：为什么不同项目在相同算力池下产生不同结果——因为不同的设定集函数携带不同的约束条件；为什么算力配置需要按项目变化——因为设定集函数不同，算力需求

不同。

4 跨领域借鉴——AI 音乐与 AI 配音

4.1 AI 音乐生成的三层架构

当前 AI 音乐生成在局部旋律生成上表现优异，但存在与 AI 绘画类似的结构性缺陷：曲式混乱、动机发展不连贯、和声逻辑矛盾、节奏漂移。

三层混合逻辑	音乐领域对应内容
第一层：全局音乐结构	调性中心、曲式框架、主题动机、和声骨架、节奏网格
第二层：单段音乐关系与因果连贯性	配器逻辑、声部进行、动态变化、段落过渡、情感弧线
第三层：局部演奏细节	颤音、力度微调、音色变化、装饰音、呼吸感

音乐领域中的因果连贯性对应关系同样遵循这三个层次（详见文献[2]）。经典力学连锁对应和声进行的功能逻辑；非局域关联对应远距离段落之间的动机呼应与调性回归；非直接接触式操控对应多声部或多音轨在编曲中通常单轨独立编辑、但混响后各轨之间产生不依赖空间传播的相互影响——不是简单的线性叠加，而是共振式的耦合。这一现象在机制上对应"不经过空间中介的直接作用"。

4.2 AI 配音生成的三层架构

当前 AI 配音生成在单句语音合成上表现优异，但存在与 AI 绘画类似的结构性缺陷：角色声音不一致、情感弧线断裂、语境响应不足、呼吸与节奏不自然。

三层混合逻辑	配音领域对应内容
第一层：全局角色声音结构	音色特征、音高范围、语速基线、口音特征、发声习惯
第二层：单段语音关系与因果连贯性	情感表达、语气变化、对话响应、空间混响、呼吸节奏

第三层：局部发声细节	微停顿、气声、唇齿音、音高微波动、音量微调

配音领域中的因果连贯性对应关系同样遵循这三个层次（详见文献[2]）。经典力学连锁对应对话的因果逻辑；非局域关联对应远距离对话之间的情感呼应；非直接接触式操控对应情感共鸣或共情——配音通过语气、语调、语速、轻重音、顿挫间隔、音阶起伏、断句等操作，使同一语句内容在不同场景中准确表达出符合当下语境的语义与情感，在说者与听者之间达成情感共鸣或共情。这一共鸣基于双方各自类似经验：说者的声音是键层面的操作，听者的经验结构是值层面的存在，共鸣发生在经验结构的耦合处，不经过空间中的因果链传递。当前各类配音模型虽然也在尝试作情感表现方面的突破，但实际效果大多呈现要么过于夸张，要么表现力不足的现象，观众们也常因此而感到“出戏”。这很有可能和模型对角色情感层次的处理过于扁平化，缺乏复杂的结构化层次感有关。而本方案在配音领域的迁移，或许正是解决这一问题的有效途径。

本方案的配音三层架构在游戏中直接对应分层记忆结构（详见文献[13]）：第一层音色身份锚受本底键约束，第二层情感弧线受经验键与当前分支的坍缩链位置调制（详见文献[13]），第三层细节在约束内实时生成。外观三层同理。

4.3 跨领域借鉴的共同逻辑

领域	第一层	第二层	第三层
AI 绘画/视频	角色、道具、场景的三维矢量结构	镜头、光照、物理、因果连贯性	纹理、色彩、微观细节
AI 音乐	调性、曲式、动机、和声骨架	配器、声部、动态、段落过渡、情感弧线	颤音、力度、音色变化
AI 配音	音色、音高、语速、口音、发声习惯	情感、语气、对话响应、空间混响、因果连贯性	微停顿、气声、唇齿音、音高微波动

共同逻辑：先构建全局结构约束，再建立单段关系与因果连贯性，最后在约束框架内填充局部细节。人类定义"生成什么"与"遵循怎样的设定集"，AI 负责执行。

这一共同逻辑表明，三层混合生成逻辑不是特定领域的技巧，而是一套通用的生成架构原理。

5 产业迁移路径与核心项目组的关键作用

5.1 传统 3D 设计软件企业的经验优势

三层混合生成逻辑的工程实现，传统 3D 设计软件企业具有显著的经验优势。

优势维度	内容	对三层混合生成逻辑的价值
建模工具链	数十年积累的几何建模、拓扑编辑、约束求解工具	直接对应第一层三维矢量结构构建
场景设置系统	成熟的摄像机、灯光、材质、物理模拟系统	直接对应第二层二维投影逻辑约束
渲染引擎	光线追踪、光栅化、实时渲染的完整技术栈	直接对应第三层位图细节填充
程序化生成	节点系统、脚本系统、表达式系统	对应因果连贯性约束与非局域关联的规则定义
工作流程理解	对"建模→场景设置→渲染"分层工作流程的深刻理解	三层混合生成逻辑的架构基础
行业标准	文件格式、数据交换、协作流程的行业标准	三层混合生成逻辑的工程化基础

5.2 纯 AI 公司的结构性局限

局限表现	对三层混合生成逻辑的影响
缺乏 3D 工作流程经验	从 2D 生成直接跨越到视频生成，跳过建模与场景设置，可能缺乏对分层约束必要性的深度理解
缺乏专业工具链	无建模、场景设置、渲染的完整工具链，难以实现三层架构的工程整合
缺乏专业用户反馈	用户以普通消费者为主，难以获得专业级的结构一致性需求反馈

端到端思维惯性	追求"输入提示词→输出视频"的端到端方案，与三层混合生成逻辑的分层约束理念相悖

需要指出的是，上述局限不是能力问题，而是经验分布问题。纯 AI 公司也可能拥有 3D 背景的研究人员，或通过合作获得 3D 工作流程经验。局限的实质是：在架构设计时容易忽略分层约束的必要性，因为端到端范式在纯 AI 公司的经验分布中更为熟悉。

5.3 核心项目组的关键能力

企业经验优势是条件，不是决定因素。关键在于实际负责软件设计与编程的核心项目组。

能力	内容	对三层混合生成逻辑的价值
架构设计	三层架构的工程分解与模块划分	将理论框架转化为可实现的系统架构
算法实现	矢量约束求解、投影几何计算、位图生成推理	各层级核心算法的工程实现
系统集成	三层之间的数据流、控制流、约束传播	确保三层协同运作
性能优化	算力调度、内存管理、并行计算	支撑大规模并发用户
跨学科理解	3D 图形学、AI 生成、认知科学	理解三层架构的理论依据
工程迭代	快速原型、测试、反馈、改进	从概念到产品的持续迭代

5.4 产业迁移路径

路径	内容	特点
3D 软件企业主导	将 AI 能力集成到现有工具链中	经验优势最充分，转型成本较高
AI 企业与 3D 企业合作	通过 API 或插件方式集成	资源互补，协调成本较高

新型平台整合	从零构建三层混合生成平台	门槛最高，可能产生最完整方案

无论哪条路径，核心项目组的能力都是决定成败的关键，而传统 3D 设计软件企业的经验优势则是不可替代的条件性资源。

6 对创作者的要求与创作门槛

6.1 设定构建的认知门槛

设定构建的创作门槛，要求创作者本人构建对现实世界更清晰的认知，以及掌握对幻想世界展开合理想象的能力。

维度	内容	与三层架构的对应
对现实世界的清晰认知	理解物理规律、因果逻辑、社会结构、认知机制	第一层与第二层的设定基础
对幻想世界的合理想象	在理解现实规则的基础上，以自洽的方式改造规则	力量体系设定函数集、非局域关联、非直接接触式操控的规则设计

幻想世界的"幻想"不在于打破规则，而在于遵循不同的规则体系。要合理地想象一个幻想世界，首先必须清晰地理解现实世界的规则——否则无法知道哪些规则被改变了、改变后会产生什么后果。

简单来说，创作者应当是博学的，但不必事事精通。博学指向对现实世界多学科的基本认知；不必精通指向 AI 辅助下的专业分工——创作者不需要成为物理学家、程序员、3D 建模师，因为 AI 可以承担各层级的专业执行。创作者需要的是判断力与整合力，而非单一领域的深度专精。

6.2 力量体系设定函数集的构建需求

三层混合生成逻辑要求创作者在生成前为全局构建完善严谨的力量体系设定函数集——明确全局元素的关联接口、操控接口、响应规则。非局域关联与非直接接触式操控需要精确的规则定义。

6.3 剧情结构的合理性要求

因果连贯性需要剧情逻辑的支撑。创作者需要确保倾向因果连贯性与全局环境泛化响应的合理性。

6.4 通用创意设计指导工具的重要性

三层混合生成逻辑的广泛应用，需要依赖通用的创意设计指导工具——即帮助创作者构建合理设定集的方法论工具。这类工具在未来创作工作中具有重要性与必要性，并可能出现在高等院校相关专业的教科书内容中。

6.5 创作门槛提升对不同创作者的影响差异

创作门槛的提升，对不同创作者的影响存在显著差异。

对于成熟作者而言，三层混合生成逻辑带来的变化可能是微调。成熟作者通常已经具备构建自洽设定集的能力。三层混合生成逻辑将他们的这些能力显性化、结构化，从隐性经验转变为显性规范。

对于创作新手而言，创作的前期工作可能会变得比以往更复杂，工作量有所提升。新手可能缺乏构建自洽设定集的能力，也缺乏跨学科常识认知水平。因此，新手可能需要小组式协作完成创作——有人负责角色设定、有人负责场景布局、有人负责力量体系设计、有人负责剧情结构。这种小组式协作在当前的创作环境中也有类似的存在形式——影视编剧团队、游戏策划团队、动画制作团队。

6.6 后续工作方向

本方案的后续工作应倾向于在 AI 辅助下，让新手创作者更容易跨过初期适应阶段和成熟作品创作门槛，让专业化的流程更易于普通人理解与运用。将专业工作变得傻瓜式操作——这是本方案后续工作方向的精确概括。

7 应用前景与验证方向

7.1 应用场景

三层混合生成逻辑可应用于图像生成、视频生成、3D 生成、角色设计等领域。

7.2 验证方向

验证方向包括结构一致性指标（畸形率、边界清晰度）、时间一致性指标（帧间稳定性、角色一致性）、物理合理性指标（光影一致性、因果连贯性）、结构自洽性感知度指标。

结构自洽性感知度：在给定作品风格框架的前提下，衡量观众是否感知到与框架不相容的结构矛盾。判据分两层。其一，风格框架内的自洽——写实风格中光影矛盾、物理违规、因果断裂属于结构矛盾；风格化作品中，风格化的变形、夸张、简化不属于

结构矛盾，但违反该风格自身内部规则（如设定好的透视体系被破坏、已建立的光照逻辑被违背）属于结构矛盾。其二，风格框架的承诺兑现——作品在开头建立的风格承诺（写实、半写实、风格化、抽象）是否在后续生成中被持续兑现；风格转换若是有意为之且内部自洽，不算矛盾；风格漂移若非有意，则属于结构矛盾。

这一指标不衡量观众能否识别 AI 生成——观众因风格化而识别出"这不是照片"是正常的，不是框架的失败。它衡量的是：作品在其自身风格承诺内是否自洽。风格化作品可以让观众明知是生成的却仍觉得"自然"，写实作品可以让观众明知是生成的却因结构矛盾而感到"不自然"——区别不在"是否被识别"，而在"结构是否自洽"。

7.3 与现有技术的兼容性

三层混合生成逻辑可作为现有扩散模型的附加约束层，作为现有视频生成模型的结构引导，不替代现有技术，而是增强其结构稳定性。

8 与状态管理、感知隔离的协作及游戏/具身智能价值

8.1 与长程状态管理的协作

如第 4.2 节所述，配音三层与分层记忆的对应关系是：音色身份锚定于本底键，情感弧线受经验键与坍缩链位置调制，语音细节在约束内实时生成（详见文献[13]第 4 章、第 8 章）。

这一对应在工程实现中包含三点含义。其一，音色参数跨分支共享，而非逐分支独立维护——身份锚只需在本底键中定义一次。其二，情感表达随分支演化而不同，但始终从同一身份锚出发——分支之间差异的是情感弧线，不是音色本身。其三，语音细节按需生成，而非预烘焙全部分支可能性——细节在约束内实时填充，不占用预存储空间。

反过来说，没有状态约束，实时生成会失去身份锚；没有生成，坍缩选择只有文本，没有可感知的呈现。

8.2 与机器感知系统的协作

本方案的生成结果不是直接给其他 NPC 读取的，而是注册为引擎真值，经机器感知系统的感知查询接口裁剪后，才进入其他 NPC 的 L2 编码（认知对象数据库的 L2 感知编码层）。这确保生成侧与感知侧之间隔着一道信息隔离闸门，NPC 不能通过生成参数偷看其他角色的内部状态。

8.3 游戏与具身智能价值

在游戏中，本方案使 NPC 的外观与语音能够随分支、情感、伤势、环境实时变化，而不失去身份锚。NPC 从"预烘焙资产 + 脚本对话"变为"被锁定在自己感官世界里的观察者"——外观的身份锚受本底键约束，语音的情感弧线受经验键与坍缩链位置调制，局部细节在约束内实时生成。这使分支叙事中的角色呈现能够随命运分叉而合理变化，同时保持"同一个角色"的可辨识性。

在具身智能中，三层生成架构可迁移为分层动作生成：第一层是机器人本体的运动学与动力学约束（身份锚），第二层是任务级动作序列与因果连贯性，第三层是底层关节轨迹与力控细节。生成结果经本体传感器感知后进入状态更新，形成"生成—执行—感知—再生成"的闭环。这一迁移保持了与游戏场景相同的三层分工——结构约束、关系约束、细节填充——只是将角色外观替换为机器人本体，将情感弧线替换为任务因果。

9 结论

本文为上篇的延续，聚焦三层混合生成逻辑在工程与产业层面的展开。

生产流程重构方面，前期工作可由 AI 辅助完成，后期工作可全面交由 AI 完成，人类角色从操作者转变为设定者，生产成本大幅降低。算力调用逻辑方面，三层架构天然适配异构算力池，按层分配算力，大幅降低总体成本。跨领域借鉴方面，三层逻辑适用于 AI 音乐与 AI 配音，表明其为一套通用的生成架构原理。产业迁移方面，传统 3D 设计软件企业具有经验优势，但核心项目组的能力是决定成败的关键。

对创作者的要求方面，设定构建的认知门槛要求创作者构建对现实世界更清晰的认知与对幻想世界展开合理想象的能力。创作者应当是博学的，但不必事事精通。后续工作应倾向于在 AI 辅助下，让新手创作者更容易跨过初期适应阶段和成熟作品创作门槛，将专业工作变得傻瓜式操作。

参考文献

[1] YunFei M. Wang 结构约束先行的视觉生成：AI 绘画与视频生成的三层混合逻辑[PP/OL]. AiraXiv, 2026. eprint: 2611.0019
[2] YunFei M. Wang 因果律的形式化分析与异质框架比较——四维分解、耦合运算与回归趋势判别[预印本]. PhilPapers, 2026. 链接：https://philpapers.org/rec/WANAFA-4
[3] YunFei M. Wang 论经典 AI 训练方式认知局限的产生原因与破局思路——迈向基于多尺度机器感知与认知对象数据库的下一代人工智能[PP/OL]. AiraXiv, 2026. eprint: 2611.0009
[4] YunFei M. Wang 数学有效性的哲学解释：基于计算模拟视角的"键-值-信息-相"四层认知模式[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANPOT-8
[5] YunFei M. Wang 时间、多重基轴与平行现实：计算模拟中的时间分支与界面选择

[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANAPE
[6] YunFei M. Wang 关系实体：值层面纯关系结构的形而上学分析[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANRSA-2
[7] YunFei M. Wang 键-值-信息-相框架的回应能力：一项基于框架内部资源的批判性推演[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANTRO-23
[8] YunFei M. Wang AI 视觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0010
[9] YunFei M. Wang AI 听觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0011
[10] YunFei M. Wang AI 嗅觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0012
[11] Puffin-World: Unified 3D World Model, 2026.
[12] YunFei M. Wang AI 味觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0013
[13] YunFei M. Wang 从长上下文到分层状态：面向长程智能体的设计框架[PP/OL]. AiraXiv, 2026. eprint: 2611.0018
[14] YunFei M. Wang AI 触觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0014
[15] YunFei M. Wang 论 AI 通感（机器联觉）功能的可行性构建方法——基于五感感知体系与认知对象数据库的跨模态联想与验证机制[PP/OL]. AiraXiv, 2026. eprint: 2611.0015
[16] YunFei M. Wang 论 AI 认知对象数据库的设计与构建——基于五感感知映射的多模态认知对象数据体系与基础设施[PP/OL]. AiraXiv, 2026. eprint: 2611.0016
[17] Wolfram S. A new kind of science[M]. Champaign: Wolfram Media, 2002.

AIGC 使用声明

本文作者提出了论文的核心观点、理论框架、方案设计思路与全部逻辑原理。在文本撰写过程中，作者使用了 DeepSeek AI 模型辅助进行文献观点的归纳、部分文本的起草与修辞润色。作者对 AI 生成的所有内容进行了全面的审阅、批判性修改与最终定稿，并对论文的全部学术内容、观点、数据引用及结论负最终责任。