

重构影视工业：AI Agent 短剧全自动化生产流水线

2026 工业级落地框架 / 从单点工具盲目尝试到全流程自动化调度



降维打击：为什么传统的“作坊式”短剧生产必将被淘汰？

传统模式

约 3-5万 RMB

单集成本

7-15天 / 部

制作周期

导演/编剧/摄像/灯光/后期 (10+人)

人员配置

强依赖演员发挥与现场调控

剧本到成片转化率

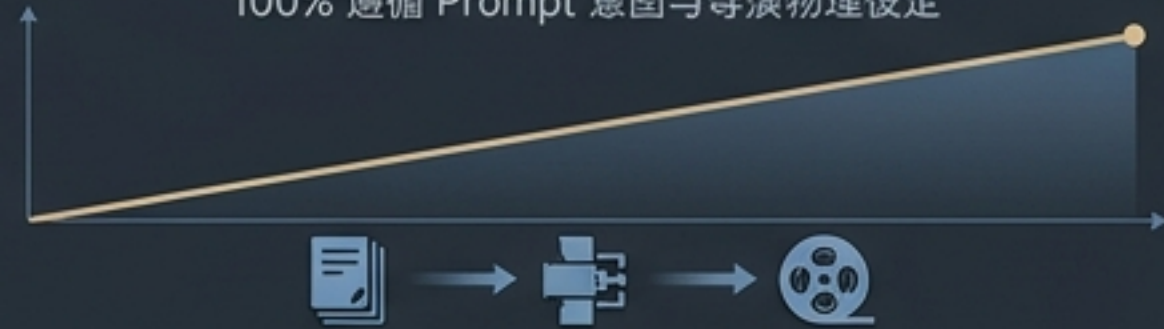
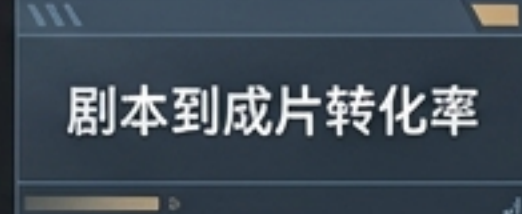
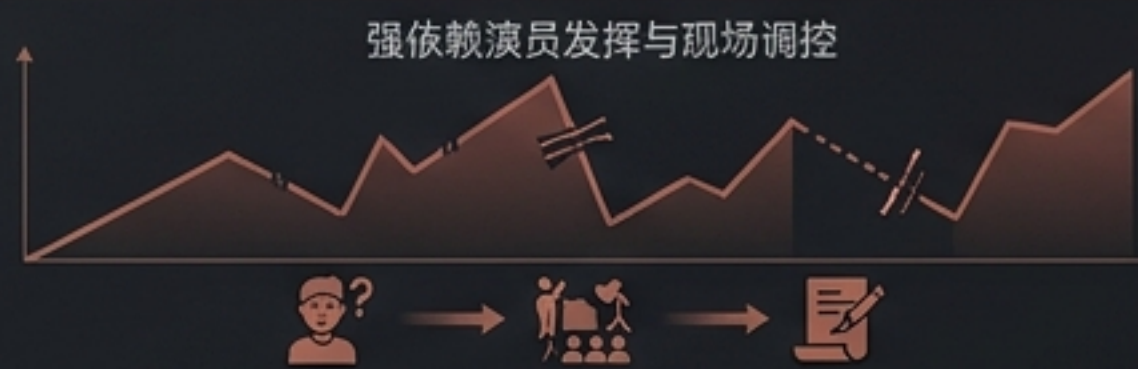
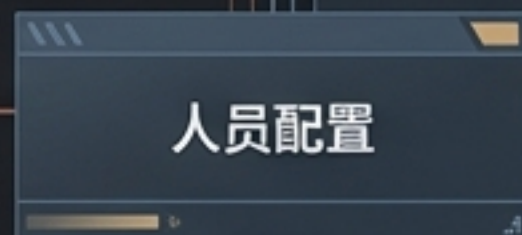
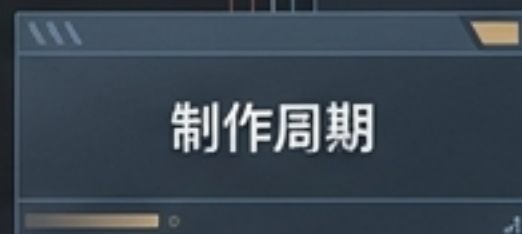
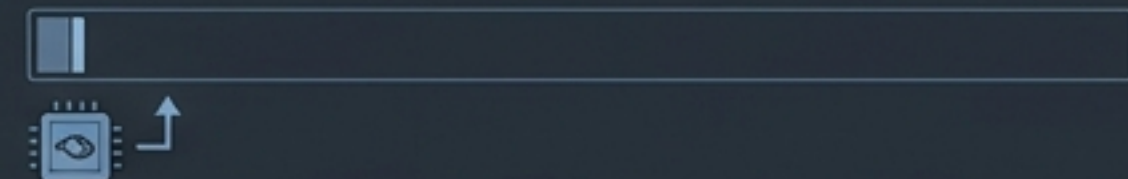
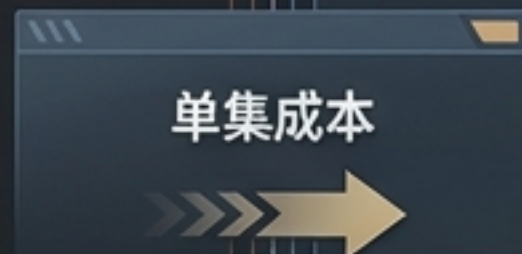
Agent 全自动流水线

趋近算力成本 < 500 RMB

4-6小时 / 部 (全自动渲染)

1名 AI 架构师 (超级个体)

100% 遵循 Prompt 意图与导演物理设定



你的下一代虚拟制片厂：4步自动化流水线

[车间 1 - 大脑]
剧本创意与角色一致性中心

[车间 2 - 骨肉]
视觉与动态生成引擎

[车间 3 - 灵魂]
声音对齐与后期封装

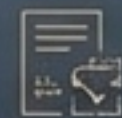
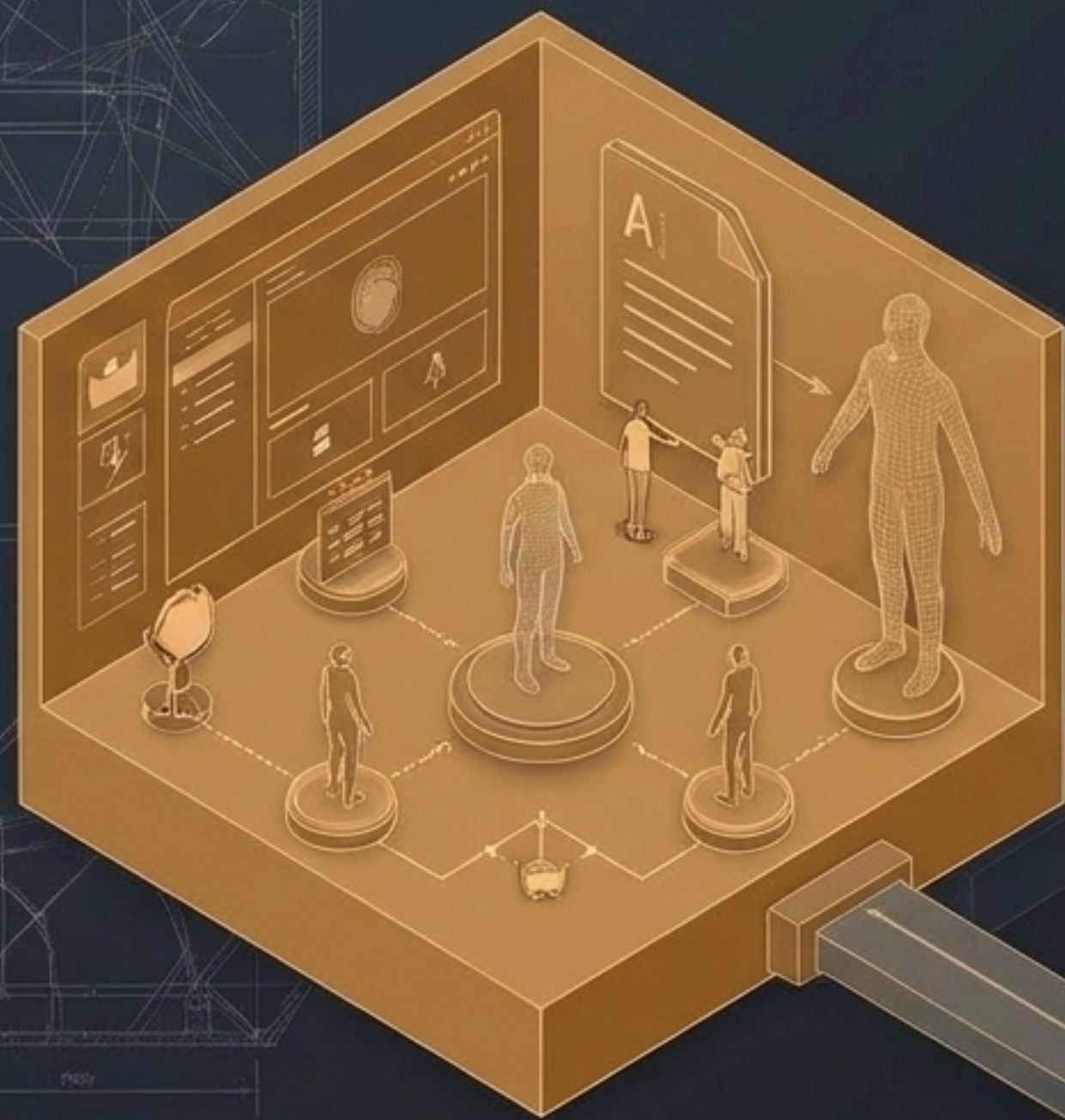
[车间 4 - 中控]
Agent 全自动化调度系统

[车间 2 - 骨肉]
视觉与动态生成引擎

[车间 4 - 中控]
Agent 全自动化调度系统



Part 1: 剧本与角色引擎 —— 构筑工业化的基石



剧本结构自动化



角色一致性控制

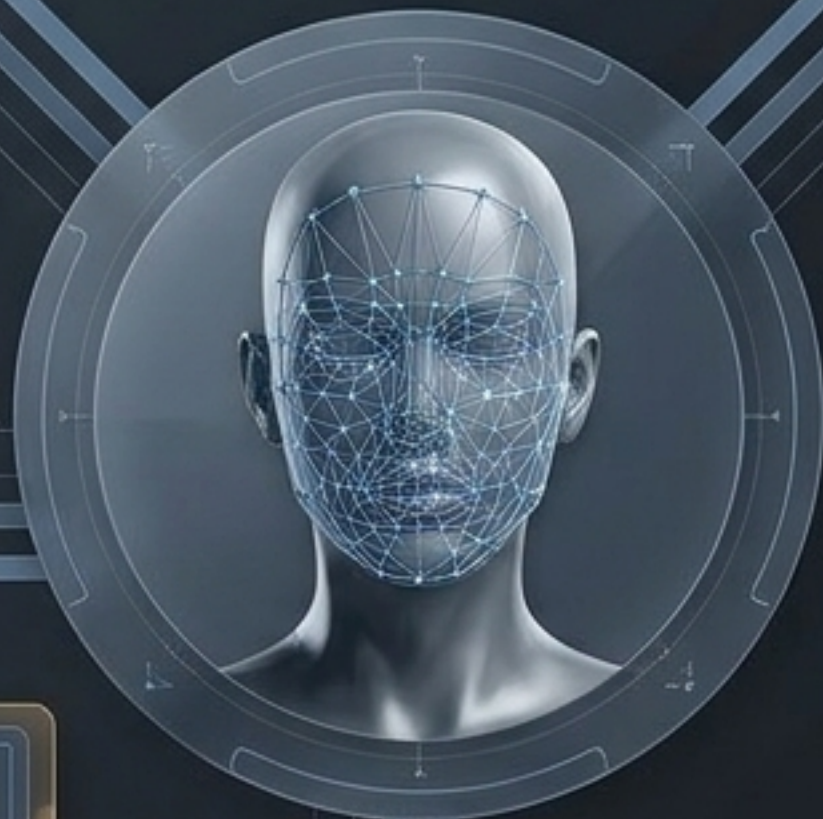


资产参考包裂变

角色不再崩坏：基于 Seed 与 Reference ID 的拓扑锁定



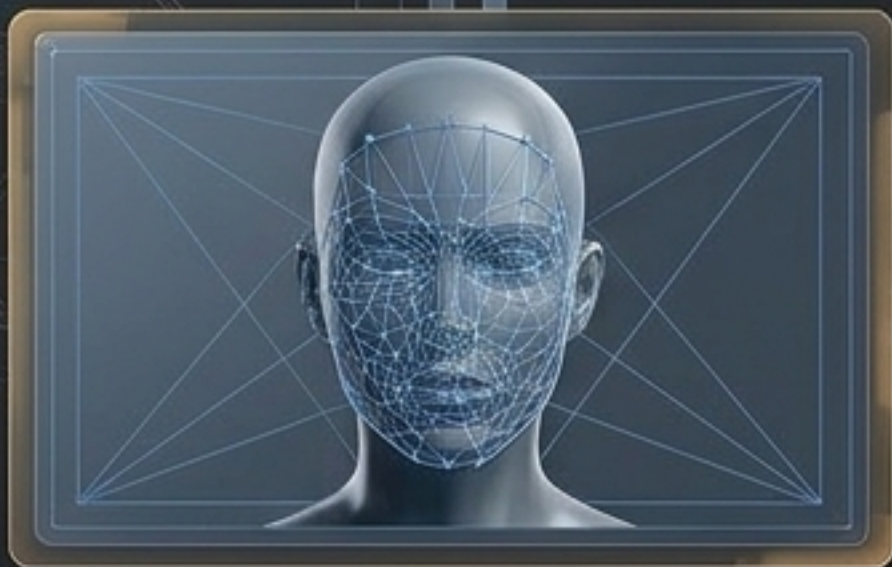
Seed



Reference ID



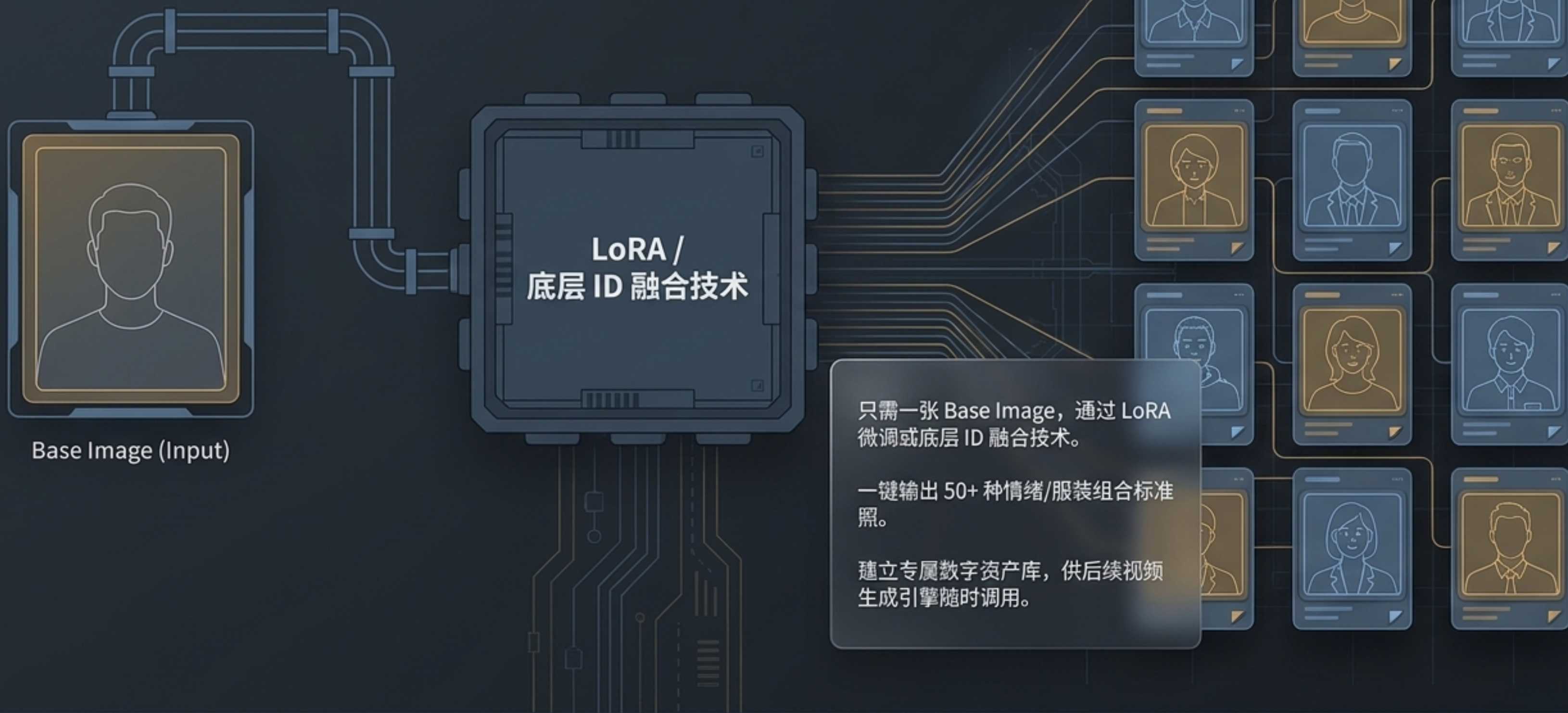
Seed



核心锚点：锁定底层骨骼与面部特征网格（Facial Mesh Topology）

- **Seed 继承：**跨场景维持全局物理特征不变量
- **Reference 权重控制：**在动态幅度过大时，强制回调面部特征（Weight = 0.85-0.95）

零成本构建数字演员库：角色资产裂变池



Part 2: AI 视频生成引擎 —— 突破物理常识与动态极限



Kling 3.0 vs Seedance 2.0

镜头语言控制

自动化渲染列队

2026 视频大模型双雄：Kling 3.0 与 Seedance 2.0 能力矩阵

Kling 3.0

物理规律遵循度：极高（适合都市/现实）

角色面部稳定性：工业级可用（95%+）

动态幅度：中（适合微表情/对话）

镜头语言控制力：强支持多轴运镜

生成耗时：较长

Seedance 2.0

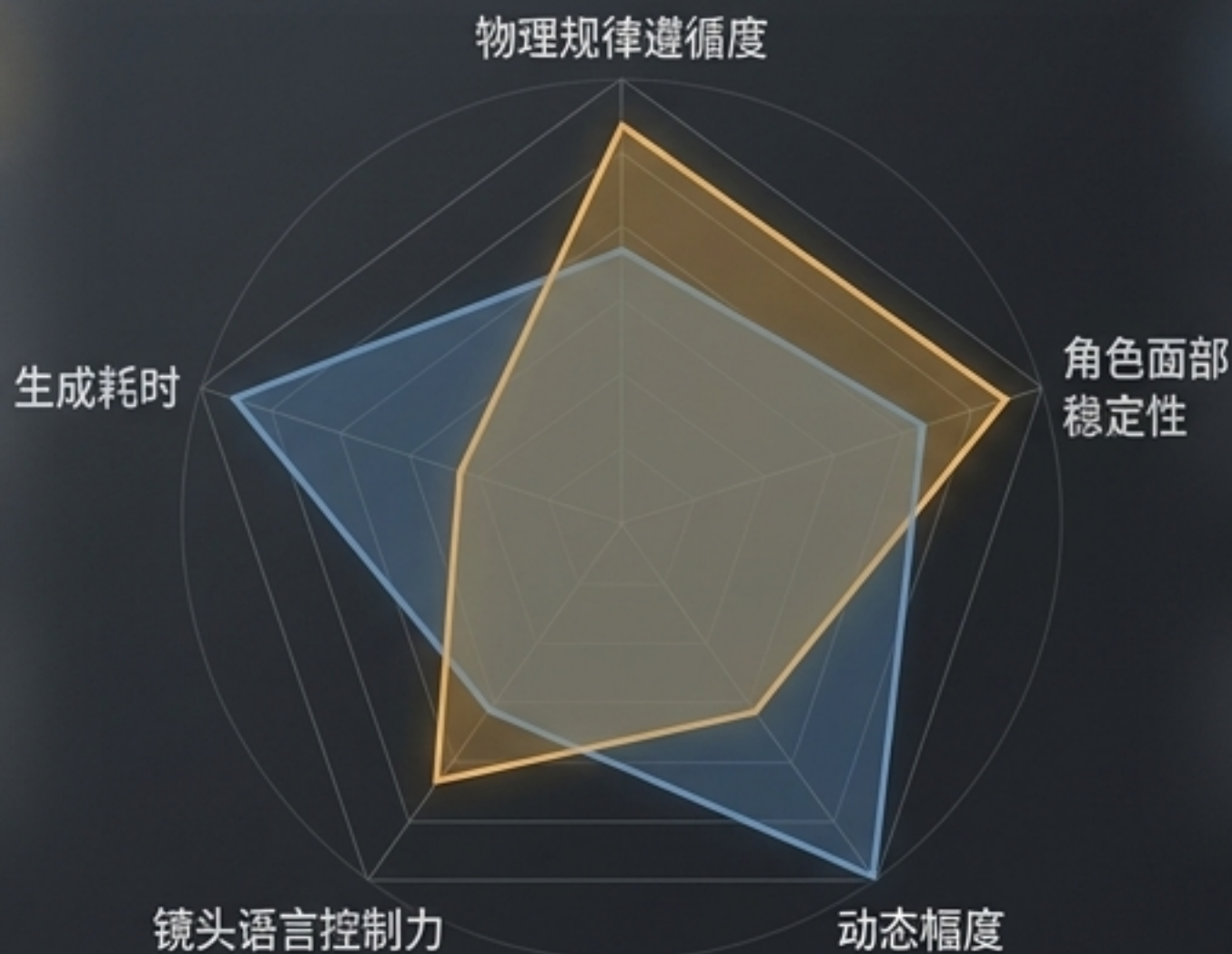
物理规律遵循度：中高

角色面部稳定性：工业级可用（95%+）

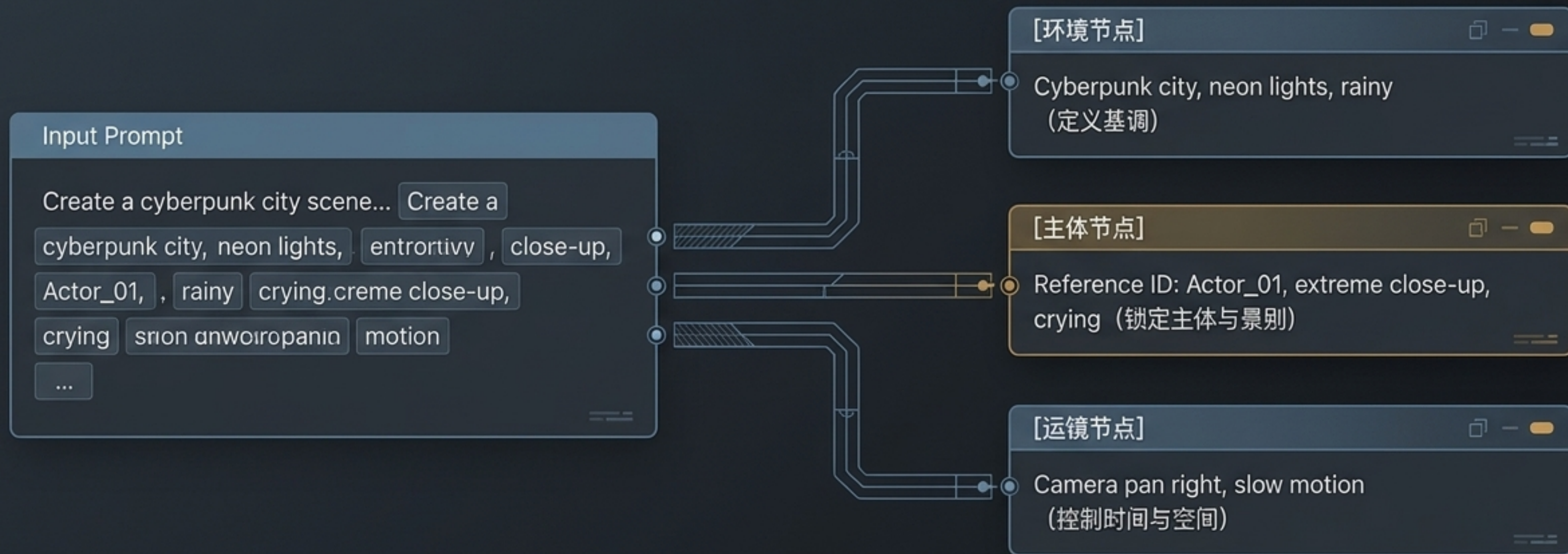
动态幅度：极高（适合玄幻/打斗）

镜头语言控制力：依赖后处理

生成耗时：极速



拒绝黑盒抽卡：将提示词转化为精确的参数控制流



通过结构化语法，将出片率从 20% 提升至 85% 以上。

Part 3: 音频封装与后期自动化 —— 赋予数字血肉以灵魂



多模态资产在时间轴上的毫秒级对齐

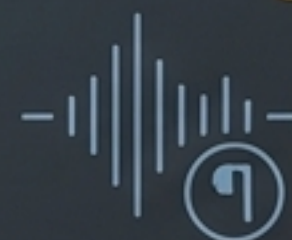


取代人工剪辑, Agent 直接生成 XML/EDL 格式文件, 实现多轨自动拼合。

LivePortrait 底层机制：基于声学特征的口型微表情驱动



1. 提取 TTS 语音中的音素 (Phonemes) 与振幅。



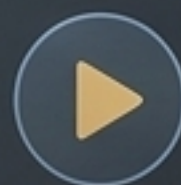
1.

2. 将音频特征映射为面部肌肉的微细牵引（如眉头微皱、嘴角抽动）。



2.

3. 实现 100% 毫无违和感的音画同轨。



3.

剪映 API 深度接入：从渲染直接到最终成片

```
> jianying_api.initialize(project_id="12345")  
  
> video_asset = load_asset("render_output_v4.mp4")  
> timeline.add_clip(video_asset)  
  
> srt_task = api.generate_subtitles(language="zh-CN", dynamic=True)  
> timeline.apply_subtitles(srt_task.result())  
  
> transitions = api.analyze_rhythm_and_apply_transitions(style=["Wipe", "Fade"])  
  
> color_grading = api.apply_global_luts(preset="Cinematic_Warm")  
  
> final_render = api.export_video(resolution="4K", fps=60, format="mp4")  
  
> print("Final Render Complete. Output path: " + final_render.path)
```

接收合成视频

API 自动识别并添加动态字幕 (SRT)

根据分镜节奏自动插入转场 (Wipe/Fade)

自动执行全局色彩校正 (LUTs)

一键输出 4K 60FPS 成片

Part 4: 系统级觉醒 — AI Agent 全流程 workflows

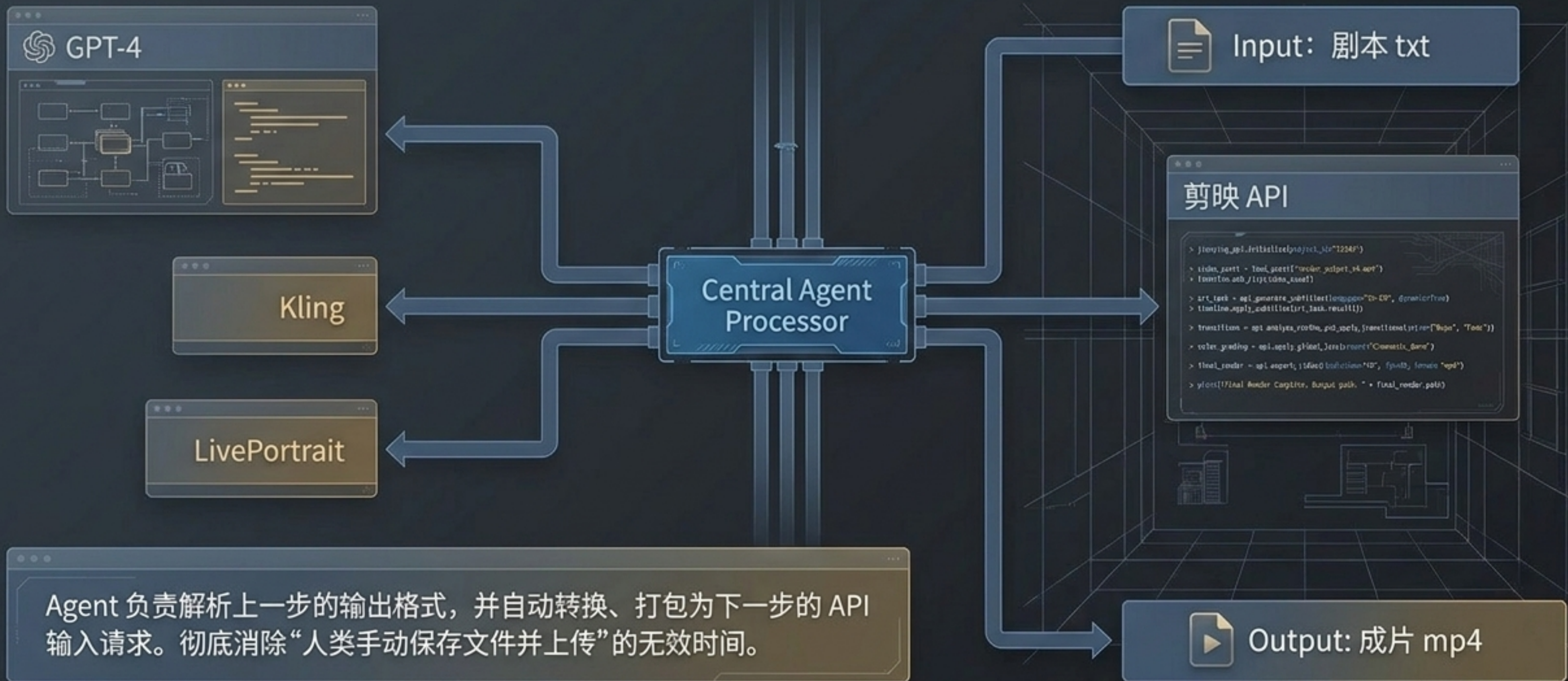
“

你不再需要学习如何使用一百个AI工具；你只需要学会如何指挥一个掌握了这一百个工具的超级智能体团队。

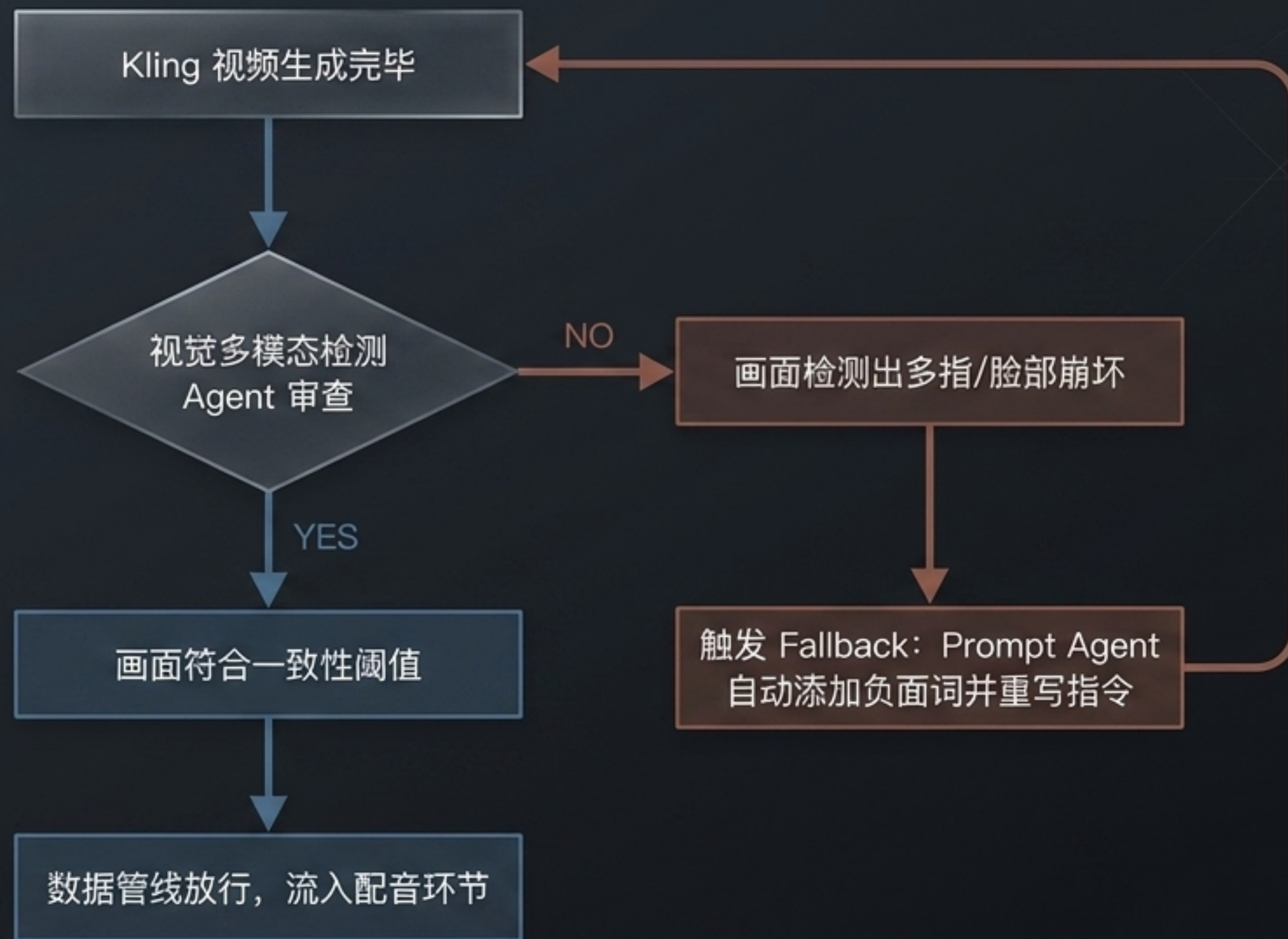
”



中控调度网络：让工具之间学会“互相交谈”



工业级稳定性：Agent 视觉质检与自愈机制



工业化实战：某头部短剧 MCN 自动化改造数据看板

人效比提升



+ 4000%

(从 1部/周 到 10部/天)

有效算力转化率



92%

(得益于自愈检测剔除废片)

全流程人力干预耗时



< 15 分钟

(仅限立项审批与最终审片)

影视工业的下一个十年，属于会写代码的导演



技术的终点，是让人类重归“造梦者”的本职。
把像素的搬运交给 Agent，把想象力的边界留给自己。
—— 你的超级制片厂，已准备就绪。