M5 · 第 14 讲

图生视频与镜头控制

给模型一个明确的运动范围

阅读约 30 分钟

练习约 30 分钟

产出：3 个稳定镜头

本节目录

开场：图片很稳，一动就崩

本课目标

一、首帧决定上限

二、运动描述四要素

三、运镜词表

四、时长与节奏

五、三类崩坏与对策

六、生成流程

七、30 分钟练习

八、检查清单

九、延伸阅读

开场：图片很稳，一动就崩

你很满意的一张图，丢进视频模型，人物开始融化、背景开始漂移、镜头乱转。你加了一句“电影感”，情况更糟。

图生视频的关键不是描述得多美，而是**给模型一个明确的运动范围**。

本课目标

- 用首帧 + 运动描述控制镜头；
- 选择正确的景别与运镜；
- 处理常见崩坏：脸变、手变、背景漂；
- 建立“先短后长”的生成节奏。

一、首帧决定上限

图生视频的第一帧，几乎决定了：

- 人物长什么样；
- 光线与色调；
- 构图与景别；
- 画面里能发生什么。

所以先做一张“可动”的首帧：主体完整、姿态自然、留有运动空间。特写和大动态不要放在同一镜。

二、运动描述四要素

> 主体动作 + 镜头运动 + 速度/节奏 + 环境变化

示例：

> 女孩缓慢抬头看向镜头，镜头轻微推近，雨滴从屋檐落下，节奏由慢到快，5 秒。

不要写“她想起往事”这类心理描述，模型无法执行。

三、运镜词表

运镜	效果	适合
固定	稳，突出动作	人物表演、产品展示
推近	聚焦情绪	特写、转折
拉远	交代环境	结尾、开场
横移	展示空间	场景、陈列
跟随	增强代入	行走、奔跑
环绕	立体感	产品、角色

一条镜头只用一个主要运镜。叠加两个，模型会打架。

四、时长与节奏

- 单镜 3–5 秒最稳；
- 超过 8 秒，崩坏概率明显上升；
- 动作复杂就拆成两镜；
- 慢动作适合情绪，不适合叙事推进。

“先短后长”：先用 3 秒验证动作，再延长或拼接。

五、三类崩坏与对策

崩坏	原因	对策
脸变	运动幅度大、分辨率低	减小动作、加人脸参考
手变	手部特写 + 运动	避开特写、藏手
背景漂	镜头运动过大	固定镜头、简化背景
闪烁	帧间不一致	降低运动幅度、换模型

崩坏	原因	对策
变形	首帧本身结构差	重做首帧，不要硬生成

六、生成流程

1. 定首帧（图生图或文生图）；
2. 写一条 3 秒运动描述；
3. 生成 2–3 个版本，选最稳的；
4. 稳定后再延长或做同场景其他镜头；
5. 全部镜头统一色调后再剪辑。

顺序反了，就会在崩坏的素材上反复修补。

七、30 分钟练习

1. 选一张自己的图作首帧（5 分钟）；
2. 写“固定镜头 + 一个动作”的 3 秒描述（5 分钟）；
3. 生成 3 个版本，记录崩坏点（10 分钟）；
4. 改成“推近 + 同一动作”，再生成 1 个版本（10 分钟）。

八、检查清单

- [ ] 首帧主体完整、姿态自然；
- [ ] 一条描述只有一个主要运镜；
- [ ] 单镜不超过 5 秒；
- [ ] 人物镜头有参考图；
- [ ] 崩坏时先改首帧或运动幅度。

九、延伸阅读

- 知识库 1.5 AI 视频

- 知识库 1.4 AI 绘画
- 知识库 3.3 前沿：空间计算与 VR

> 下一讲：配音、字幕与剪辑交付。你会把零散镜头变成一条能发出去的片子。
