M2 · 第 6 讲

图像与视频提示词

从一句话到分镜表

阅读约 35 分钟

练习约 30 分钟

产出：6 格分镜

本节目录

开场：单张好看，放到一起就不像一个系列

本课目标

一、图像八要素

二、风格词怎么选

三、一致性：三种可控手段

四、视频提示词：镜头语言

五、分镜表：先想清楚再生成

六、图生视频的工作流

七、30 分钟练习

八、五个常见坑

九、检查清单

十、延伸阅读

开场：单张好看，放到一起就不像一个系列

你生成了一张很满意的图，第二天想再做一张同风格的，结果人物变了、色调变了、构图也变了。做视频更明显：每个镜头单看都不错，拼起来却像五部不同的片子。

图像和视频的提示词，难点不在“写得美”，而在**可控与一致**。这一讲给两套公式和一张分镜表。

本课目标

- 用八要素公式写图像提示词；
- 用参考图、固定描述、种子保持一致性；
- 用分镜表把想法拆成可生成的镜头；
- 完成一组 6 格分镜。

一、图像八要素

> 主体 + 动作/状态 + 环境 + 风格 + 构图/镜头 + 光线 + 色彩 + 画质

示例：

> 一位穿深色风衣的年轻女性站在雨后的旧城街道，回头看向镜头，35mm 胶片质感，中景，侧逆光，冷灰与暖黄对比，浅景深，高清细节。

写法要点：

- 一个主体，最多两个动作；
- 风格词只选一个主风格，不要“既写实又赛博又水彩”；
- 光线决定质感，比堆形容词有效；
- 画质词放最后，别用它替代具体描述。

二、风格词怎么选

风格	关键词	适合
写实摄影	纪实、自然光、35mm、浅景深	产品、人物、场景
胶片	颗粒、过期胶片、暖调、漏光	情绪、故事感
商业插画	扁平、渐变、干净的形	海报、封面
3D 渲染	盲盒质感、柔光、塑料材质	角色、IP
极简图形	大面积留白、几何、双色	信息图、品牌

选一个主风格 + 一个辅助词就够。风格越多，画面越糊。

三、一致性：三种可控手段

1. 参考图

把满意的那张作为参考，要求“保持人物、服装、色调，改变场景与动作”。这是最直接的方法。

2. 固定描述块

把不变的部分写成一段固定文字，例如“短发、圆脸、左眉有疤、深蓝夹克”，每次原样粘贴，只改动作和环境。

3. 固定种子/参数

工具支持时，固定随机种子，再小步调整提示词。一次只改一个变量，才能知道是哪句起了作用。

四、视频提示词：镜头语言

视频不是“会动的图”，它多了四件事：

要素	写法	示例
景别	远景/全景/中景/近景/特写	中近景
运镜	固定/推/拉/摇/跟/环绕	缓慢推近
动作	一个明确动作	她抬头，雨滴落下
时长/节奏	3s/5s，快慢	5 秒，前慢后快

一个可用的镜头句：

> 中近景，镜头缓慢推近，女性抬头看向镜头，雨滴从屋檐落下，冷色调，5 秒，电影感。

五、分镜表：先想清楚再生成

镜号	画面	景别	运镜	时长	台词/字幕
1	雨夜街道，人物背影	全景	固定	3s	字幕：那天雨很大
2	她回头	中景	推近	3s	—
3	手中旧照片特写	特写	固定	2s	—

先填表，再逐镜生成。表格能暴露三个问题：镜头之间有没有逻辑、时长够不够、需不需要台词。

六、图生视频的工作流

1. 用图像模型做出每个镜头的首帧；
2. 选一张最稳的作为风格基准；
3. 逐镜生成 3–5 秒，不要一次做 30 秒；
4. 统一色调与节奏后再剪辑；
5. 配音、字幕、音乐最后加。

顺序不要反。先解决画面一致性，再谈叙事。

七、30 分钟练习

1. 选一个 15 秒的小故事（5 分钟）；
2. 填 6 格分镜表（10 分钟）；
3. 为第 1 格写图像提示词并生成（5 分钟）；
4. 为同一格写视频提示词，比较差异（5 分钟）；
5. 记录哪一句改动最影响结果（5 分钟）。

八、五个常见坑

1. **画面里塞文字**：生成文字容易错，字幕后期加；
2. **手部特写**：仍是弱项，用构图避开或后期修；
3. **一镜到底讲太多**：拆成 3–5 秒的小镜头；
4. **风格词堆叠**：互相冲突，画面变脏；
5. **只改提示词不换参考图**：一致性问题多半靠参考图解决。

九、检查清单

- [ ] 提示词包含主体、环境、风格、光线；
- [ ] 一组图有固定描述块或参考图；
- [ ] 分镜表写清景别、运镜、时长；
- [ ] 单镜头不超过 5 秒；
- [ ] 字幕与文字后期添加。

十、延伸阅读

- 知识库 1.4 AI 绘画
- 知识库 1.5 AI 视频
- 知识库 1.9 ComfyUI 工作流

> 下一讲：认识智能体。你会分清“聊天机器人”和“能干活的 Agent”，并判断哪些流程值得交给它。