

AI 短剧技术平台

3.1 脚本与文案创作平台

虽然 AI 短剧的最终呈现形式是视频，但其创作过程涉及大量文案工作，包括故事大纲、剧本对白、分镜头描述、角色设定及宣传文案等。因此，高效的文案生成能力对于提升 AI 短剧的整体创作效率至关重要。当前行业内常用的文案创作平台包括 DeepSeek、腾讯元宝、Kimi、智谱清言、通义千问、豆包、文心一言、秘塔 AI 等。

3.1.1 DeepSeek

DeepSeek 是由深度求索公司推出的大语言模型，专注于专业创作场景下的高效文本生成与智能知识处理。该平台深度融合先进的自然语言理解与生成技术，拥有优秀的长文本创作能力，能辅助完成商业报告、学术论文、营销文案及小说创作等多种文本生成任务。

其知识推理引擎支持精准的信息整合与逻辑分析，能够实现多轮交互对话、文本优化、内容摘要及结构化写作，为作家、企业用户和内容创作者提供高效可靠的智能写作解决方案，DeepSeek 平台页面如图3-1所示。



图 3-1

3.1.2 腾讯元宝

腾讯元宝是腾讯公司推出的大语言模型，聚焦于专业创作场景下的高效文本生成与智能知识计算。其知识增强引擎支持精准的信息检索与逻辑推理，能够进行多轮深度对话、文本润色、数据分析乃至代码生成，为创作者、企业和开发者提供一站式智能创作支持。

该平台的特色在于其搜索结果主要来源于微信公众号，而公众号内容普遍被认为是当前网络文本中质量较高的来源之一，腾讯元宝平台页面如图3-2所示。



图 3-2

3.1.3 智谱清言

智谱清言是由智谱 AI 推出的大语言模型，专注于高效创作场景下的智能文本生成与知识推理。该平台深度融合自然语言处理与知识图谱技术，具备强大的长文本理解与创作能力，可辅助完成商业分析、学术写作、品

牌文案、网络文学等多种专业文本创作任务。

其知识增强架构支持精准的信息提取与逻辑推演，能够实现多轮交互对话、内容优化、数据解读及结构化写作，为作家、企业用户和内容从业者提供高效精准的智能创作支持，智谱清言平台页面如图3-3所示。



图 3-3

3.1.4 通义千问

通义千问是由阿里巴巴集团推出的大语言模型，聚焦于智能文本生成与多模态内容创作。该平台深度融合自然语言处理与多模态交互技术，具备卓越的长文本理解与生成能力，可高效辅助创作者完成商业企划、技术方案、文学创作、新媒体内容等多种专业写作任务。

通义千问是一个庞大的模型系列，其最新 MAX 版本曾登顶全球 AI 模型排行榜，是创作者不容错过的重要平台。本书后续章节还将介绍通义千问的视频 AI 模型，通义千问平台页面如图3-4所示。



图 3-4

3.1.5 豆包

豆包是由字节跳动推出的多模态大模型，专注于创意内容生产。该平台深度融合自然语言处理与创意激发技术，具备强大的风格化写作能力，可高效辅助创作者完成短视频脚本、社交媒体文案、广告创意、轻小说创作等多种类型的互联网内容生产任务。

该平台除文本处理能力强外，因其与即梦采用相同的底层大模型，还具备出色的图像处理与修饰能力，基本可实现“用语言修改图像”。在短剧创作中，关键帧画面的修改可优先考虑使用此平台，豆包平台页面如图3-5所示。



图 3-5

3.1.6 文心一言

文心一言是由百度公司推出的大语言模型，聚焦于智能创作场景的高效文本生成与知识增强计算。该平台深度融合自然语言处理与多模态交互技术，具备卓越的长文本理解与生成能力，可精准辅助创作者完成分镜头脚本、学术论文、营销文案、网络文学等多种专业文本创作任务。

该平台的特色在于依托百度多年积累的网络文本库及百度文库资源，能够输出较为精准、实用的结果，为作家、企业用户和内容创作者提供全方位的智能写作支持，文心一言平台页面如图3-6所示。



图3-6

3.2 AI图像创作平台

如本书第2章所述，在AI短剧的创作过程中需要生成大量关键帧画面，因此创作者必须了解不同AI图像创作平台的特点。以下将分别介绍Midjourney、可灵、即梦、Liblib AI、通义万相、Stable Diffusion、ComfyUI、Lovart AI、造点等平台。各平台的具体使用方法详见本书附赠的飞书学习资料。

3.2.1 Midjourney

Midjourney是基于AI技术的图像生成工具，创作者通过输入文本描述即可快速生成高分辨率、风格多样的艺术作品，涵盖写实、卡通、动漫、抽象等多种类型。

平台支持以图生图功能，允许上传参考图片并结合文字指令调整生成结果，还能通过混合多张图片创作融合性作品。Midjourney是世界上首款实现大规模商业化的图像生成模型，在场景丰富度、创意性和视觉效果方面始终保持领先地位，不过在生成涉及中国文化符号的图像时，其表现尚有提升空间。

Midjourney绘画平台页面如图3-7所示。若访问不便，可使用其国内孪生平台“悠船”，其页面如图3-8所示。如需系统学习Midjourney，可参阅笔者在清华大学出版社出版的相关图书，如图3-9所示。

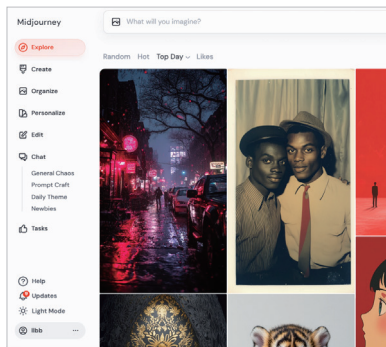


图3-7

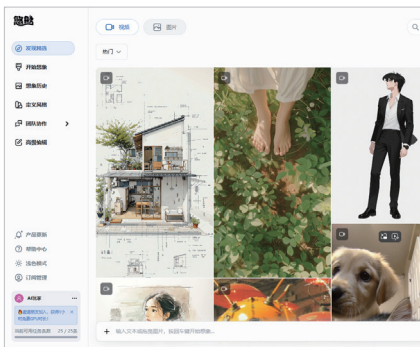


图3-8



图3-9

3.2.2 可灵

可灵平台提供文生图、图生图功能，支持多样化模板和中文语义精准生成。该平台创意表现力强，文字效果生成能力突出，可灵的生图页面如图3-10所示。

3.2.3 即梦

与可灵平台类似，即梦除支持生成视频外，还支持文字生成图片、图片生成图片等多种图像创作功能。创作者通过输入中文描述词即可快速生成超现实场景、人物肖像等多样化风格的视觉作品。平台提供智能画布功能，包含扩图、局部重绘、消除笔、高清放大等编辑工具，支持多图融合与细节调整，生成速度快且提供多版本选择，是许多创作者的常用平台之一，其生图页面如图3-11所示。

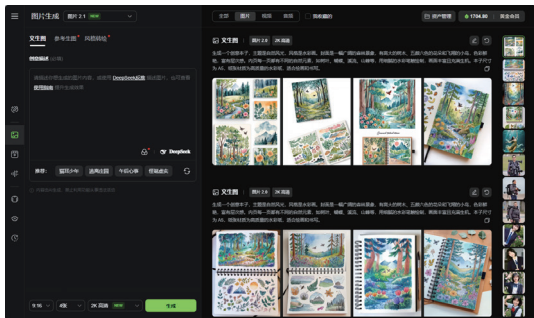


图3-10

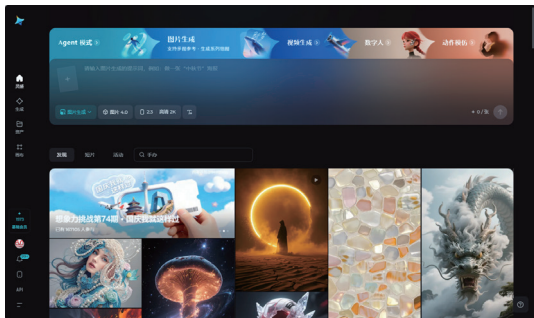


图3-11

3.2.4 Liblib AI

Liblib AI 是基于 Stable Diffusion 生态构建的综合性 AI 创作平台与大型创作者社区，现已升级至 2.0 版本。平台集模型分享、在线生成、模型训练和互动交流于一体，致力于打造一站式 AI 专业创作工作室，深受中文创作者欢迎。

平台托管了全球最大的图片风格开源模型库，覆盖插画、摄影、电商、海报、IP 等多种视觉风格，并提供极简生成器，支持图像与视频在同一页面生成。Liblib AI 2.0 整合了多类顶尖模型，包括 Qwen Image、F.1、Kontext、Nano Banana、Seedream4、MidjourneyV7 等图像模型，以及 Kling、Hailuo、Vidu、WAN 等视频模型，并内置 500 余种独家专业视觉特效，支持影视级品质输出。

平台兼容开源与闭源模型，保留 WebUI 和 ComfyUI 操作页面，支持创作者自由组合工作流，提供批量化处理功能，如扩图、高清修复、换脸、产品精修等，显著提升创作效率。该平台特别适合对绘图模型和工作流有一定了解的创作者，其生图页面如图3-12所示。



图3-12

3.2.5 通义万相

通义万相是阿里云推出的 AI 绘画平台，支持文生图、图生图功能。创作者可通过文字描述或上传图片快速生成多样化艺术作品，并自由选择水彩、油画、中国画等特色艺术风格。平台还提供了相似图生成功能。作为

阿里云通义大模型系列成员，该平台整合了文本、图像等多模态处理能力，适用于艺术设计、广告营销等多元场景，其生图页面如图3-13所示。



图3-13

3.2.6 Lovart AI

Lovart AI 是集成了多种前沿 AI 绘画模型的创意生成平台，作为 Liblib 的海外平台，通常率先整合国内无法使用的国外优秀模型，如 Midjourney、Nano Banana 等。

该平台的功能不仅限于图像生成，还接入了多款主流视频生成模型，例如 Sora 2、Kling 2.5 Turbo 以及 Wan 2.5 等。创作者完成图像创作后，可直接在平台内调用这些视频模型，将图片一键转化为视频内容，其平台创作页面如图3-14所示。

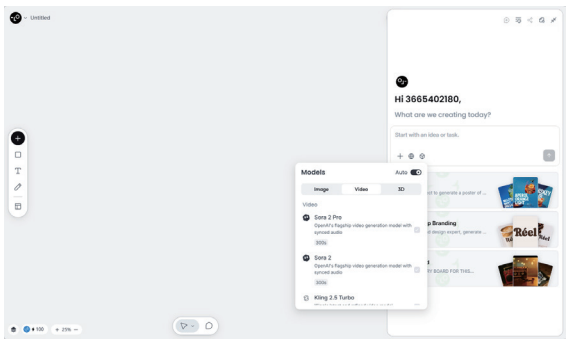


图3-14

3.2.7 造点

造点是夸克应用推出的集成化 AI 创作解决方案，深度融合 Midjourney 模型与自研技术，为中文创作者提供低门槛、高效率的 AI 绘画与视频生成服务。平台采用积分消费机制，成本约为海外平台的 50%，创作者可通过夸克应用或造点平台官方入口进入一体化创作环境。

造点能够高效生成视觉素材与分镜画面，显著提升剧情图像化与视频化的制作效率和创意落地能力，其生图页面如图3-15所示。

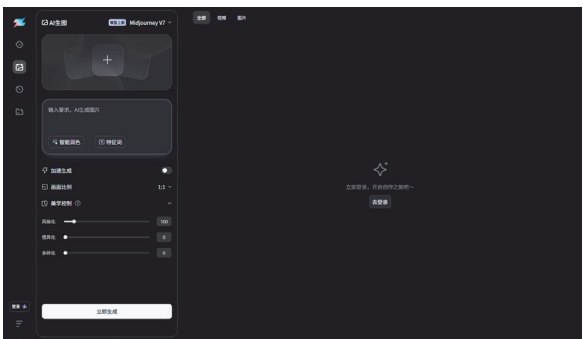


图3-15

3.2.8 Stable Diffusion

与前述商业化 AI 平台不同，Stable Diffusion 是一款开源的 AI 绘画工具，创作者可将其部署在本地设备中运行，无须额外购买算力。

除免费使用外，Stable Diffusion 功能强大，能够根据文字描述快速生成高质量图像。其核心原理基于“潜在扩散模型”，先将文本提示转换为机器可理解的语义向量，再从随机噪声图像开始，逐步去噪并融入文本信息，最终生成符合描述要求的逼真画面。

Stable Diffusion 支持多种创作模式，除基础的文生图功能外，还能实现图生图、局部重绘、图像修复等。创作者可通过调整采样步数、引导系数等参数精细控制输出效果。

作为开源项目，Stable Diffusion 拥有丰富的插件生态和模型库，可轻松生成写实、动漫、水彩等不同风格的作品，操作页面直观友好，便于新手快速上手。如需系统学习，可参阅笔者在清华大学出版社出版的相关图书，如图3-16所示。Stable Diffusion 的操作页面如图3-17所示。



图3-16

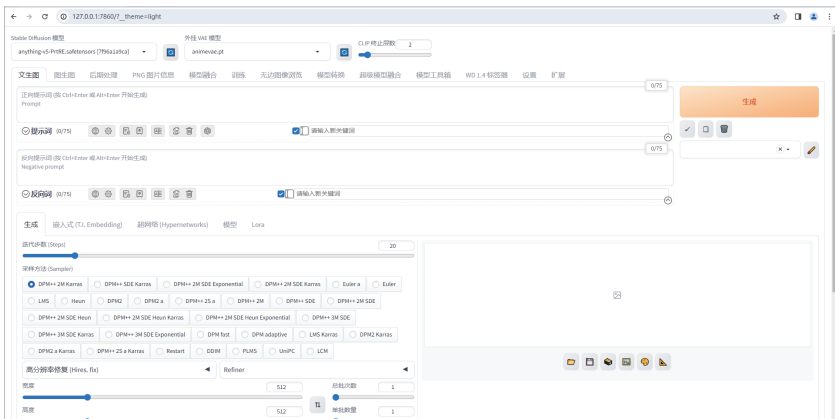


图3-17

3.2.9 ComfyUI

ComfyUI 与 Stable Diffusion 同为开源软件，支持本地化部署，其操作页面如图3-18所示。ComfyUI 采用节点工作流设计理念，将复杂的稳定扩散算法拆解为多个独立操作节点。这种设计不仅增强了工作流的灵活性和可定制性，还为创作者提供了丰富的定制选项和工作流复用能力。

如需系统学习 ComfyUI，可参阅笔者在清华大学出版社出版的相关图书，如图3-19所示。

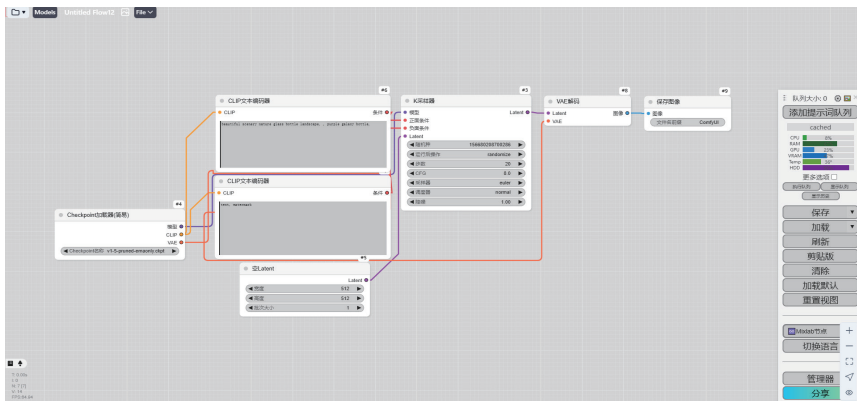


图3-18



图3-19

在实际应用中，ComfyUI 展现了强大的生成能力与高效性能。对初学者而言，其集成的众多插件节点和复杂操作流程可能带来一定的学习难度，但大多数创作者只需理解基本原理并掌握现有工作流的使用方法即可，

无须从头搭建复杂工作流。

随着功能持续扩展，ComfyUI 已从最初的图像生成工具逐步演进为支持视频制作、3D 建模、音频处理等多模态创作的综合型工作流环境。

总体而言，ComfyUI 是一个学习门槛较高但创作潜力巨大的 AI 平台，也是专业短剧创作者必须掌握的重要工具之一。

3.3 AI 视频创作平台

AI 视频创作平台是制作短剧的核心工具，其重要性不言而喻。当前市场上的可选平台较多，主要功能差异不大，但在部分特色功能上仍有所区别。创作者在选择平台时，既要考查平台功能，也要考虑使用成本，毕竟制作 AI 短剧往往需要反复尝试不同的提示词。以下将分别介绍可灵、即梦、Hailuo AI、通义万相、Vidu、RunningHUB 等平台。各平台的详细特点与使用方法，可参阅本书附赠的飞书学习资料。

3.3.1 可灵

可灵是由快手公司推出的 AI 视频生成平台，具备强大的文生视频和图生视频转换功能，能够生成 1080P 高清视频、模拟真实物理场景并支持长视频创作，为创作者提供高效的内容创作工具，其平台页面如图 3-20 所示。

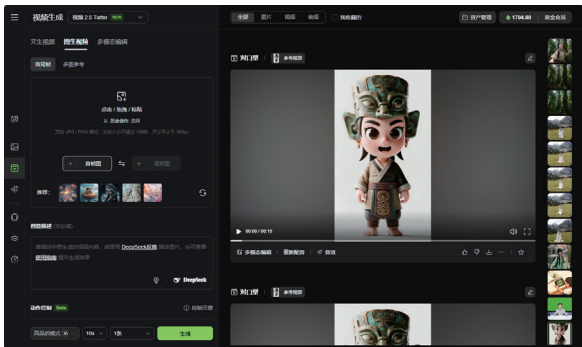


图 3-20

1. 基础功能

- 文生图功能：输入文字描述，即可快速生成对应图片。
- 图生图功能：上传参考图片并配合文字描述，生成符合要求的图片。
- 文生视频功能：根据文本或图片内容生成视频，支持 AI 自动续写后续视频内容。
- 图生视频功能：通过上传图片生成相应视频内容，特别适合基于静态图像创作动态视觉效果。
- 镜头控制功能：支持横移、摇镜头等镜头运动控制，增强视频的动态表现力。

2. 特色功能

- 首尾帧控制：自定义视频开头和结尾画面，提升视频的专业性与个性化程度。
- 运动笔刷：为图片中的元素指定运动轨迹，实现图片转视频时的定向运动效果（注：非所有版本均支持此功能）。
- 多模态视频编辑：在已生成视频基础上，通过新增图片或文字提示对视频元素进行增删、替换，支持精细

化创意调整。

- 视频音效：同步生成高质量、具有空间感的音效与背景音乐，增强沉浸式观看体验。
- 对口型功能：根据对话内容自动匹配人物口型，提升角色表现力与真实感。
- 高品质模式：在文生视频和图生视频中提供更高画面质量与动态一致性，显著提升生成效果。

3. 收费标准

平台每月赠送 166 个灵感值，用于生成图片和视频。灵感值用尽后，可购买灵感值（1 元兑换 10 个灵感值，100 个起购）或开通会员。黄金会员月费 66 元，包含 660 个灵感值及专属会员权益。

3.3.2 即梦

即梦是剪映旗下产品，基于先进的自然语言处理与图像生成技术，用户通过语言描述或上传图片即可快速生成高质量图像与视频内容，其平台页面如图3-21所示。

1. 基础功能

- 图片生成：将文字描述转换为对应图片。
- 智能画布：提供交互式画布页面，支持抠图、图像重组及基于提示词的图像重绘。
- 视频生成：将文字描述转换为视频，也支持以图片为基础生成视频内容。

2. 特色功能

- 多模态数字人：上传角色图片与音频，生成最长 30 秒的数字人视频。
- 动作模仿：基于静态图片与动作参考视频（或官方模板），生成角色复刻动作的视频，大幅降低虚拟偶像、节目分身的制作成本。
- 智能多帧：上传 2~10 张关键帧图片，AI 自动补全中间过渡画面，生成最长 54 秒的长镜头视频。支持设置帧间运镜时长（1~6 秒），并通过提示词控制运镜逻辑，适用于广告、产品展示等场景，在时长、自然度与成本方面优于同类产品。
- 主体参考：上传参考图片并使用 @ 符号指定主体，确保生成视频中主体外观、细节与姿态在场景切换、动作变化中保持一致，有效避免面部扭曲、服装突变等问题。

即梦平台另一大特色是采用最新的 Seedance 1.0 视频模型。该模型在 Artificial Anlys 权威评测中表现优异，在“文本转视频”和“图像转视频”两项任务中均位列首位，综合表现超越 Google Veo 3、快手可灵 2.0 及 OpenAI Sora 等主流模型，如图3-22所示。

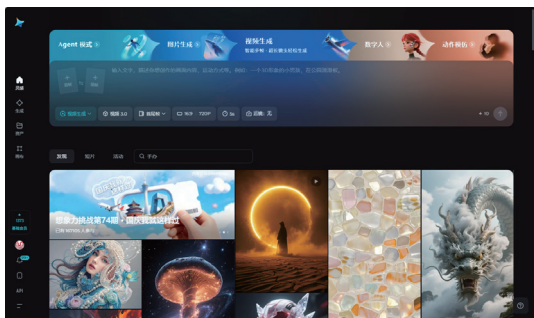


图3-21

Artificial Analysis Video Arena Leaderboard				
Text to Video		Image to Video		
Creator	Model	Arena ELO	95% CI	# Appearances
ByteDance Seed	Seedance 1.0	1301	-10/+9	10,869
Google	Veo 3 Preview	1250	-8/+10	12,310
Google	Veo 2	1131	-6/+7	14,868
Kuaishou	Kling 2.0	1114	-7/+8	13,199
Kuaishou	Kling 1.5 (Pro)	1053	-5/+5	25,220
OpenAI	Sora	1052	-5/+5	26,240
MiniMax	T2V-01	1039	-4/+4	47,460

图3-22

该模型通过超过 3000 个生成片段的实测试验，在镜头连贯性、细节保持与多镜头逻辑等方面实现技术突破，其核心优势如下。

- **多角度场景一致性：**一次性生成同一场景的多个连贯镜头，提升空间一致性与镜头语言理解能力。
- **复杂提示精准执行：**同步解析主体动作、摄像机运动与情绪节奏等复合指令，实现高精度控制。
- **高动态物理仿真：**在高速运动场景中保持背景自然运动与物理连贯性，避免空间扭曲或动作断裂。
- **多风格生成能力：**支持从照片级真实感到 3D 动画、手绘风格的多样化创作。

3. 收费标准

平台依据创作者活跃度每日赠送 60~100 个幸运积分，用于生成图片与视频。积分用尽后，可直接购买（1 元兑换 10 个积分，500 个起购）或开通会员。基础会员月费 79 元，包含 1080 个积分及专属会员权益。

3.3.3 Hailuo AI平台

Hailuo AI 是由 MiniMax（上海稀宇科技有限公司）自主研发的先进智能助手平台，融合前沿 AI 技术，为用户提供丰富高效的功能体验，其平台页面如图3-23 所示。

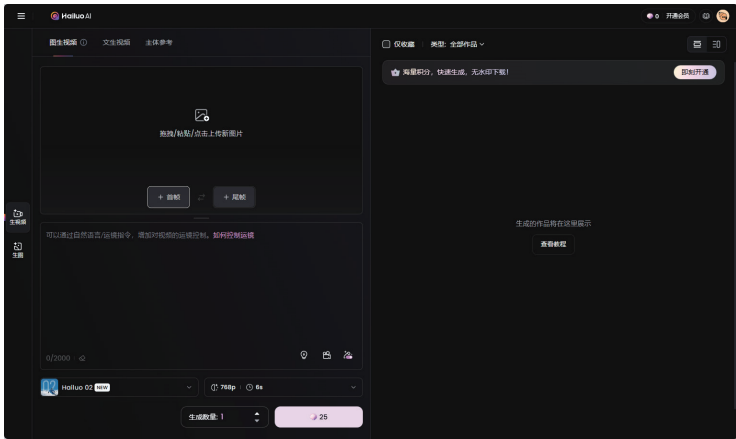


图3-23

1. 基础功能

- **长文本处理：**能够快速提炼长文本关键信息并生成内容摘要，帮助用户高效获取核心要点。该功能在处理大量信息时优势明显，可有效提升决策效率。
- **图片生成：**支持通过文本描述或上传参考图片生成高清图像，提供写实、3D 艺术等多种风格选项。免费用户单次可生成 5 张图片。
- **视频生成：**基于文本描述或上传图片直接生成视频内容，具备较快的生成速度。

2. 特色功能

- **主体参考：**上传包含人物或物体的图片后，AI 将精准识别并锁定主体的视觉特征（包括面部细节、服装纹理等）。在后续视频生成过程中，无论场景与动作如何变化，主体均能保持高度一致性与连贯性。此功能有效解决了 AI 视频生成中常见的角色形象“崩坏”问题，特别适合需要保持角色形象一致性的叙事内容创作。
- **首尾帧技术：**允许创作者指定视频的起始画面与结束画面，由 AI 自动计算并补全中间所有过渡帧，生成流畅的转场动画。该技术为实现复杂的镜头语言提供了更多可能性。

3. 收费标准

平台为新注册用户赠送一定数量的“贝壳”，用于生成图片和视频。“贝壳”用尽后，需开通会员获取“贝壳”。基础会员月费 68 元，每月可获得 1000 个“贝壳”，并享受专属会员权益。

3.3.4 Vidu AI

Vidu AI 是由北京生数科技有限公司研发的视频生成模型，核心团队源自清华大学 AI 研究院，具备视频生成、多镜头处理与个性化创作等核心能力，其平台页面如图3-24 所示。

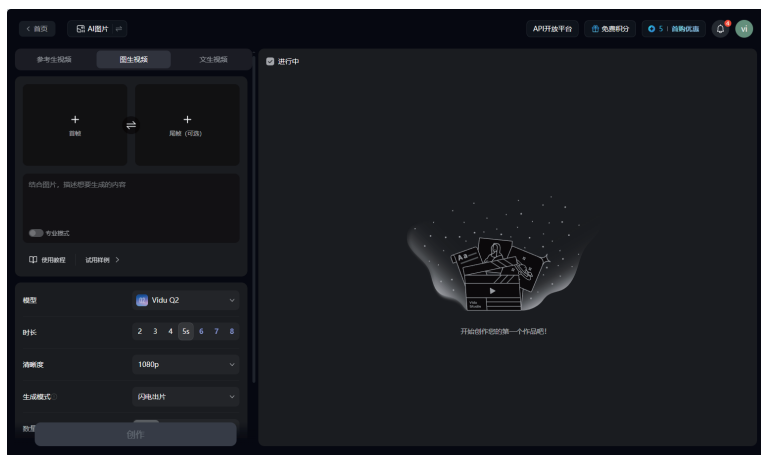


图3-24

1. 基础功能

- 文本转视频：输入文字描述即可生成相应视频内容。
- 图片转视频：上传静态图片后可转换为动态视频，实现图像动画化。
- 高清视频生成：支持生成最长 16 秒、分辨率达 1080P 的高清视频，并具有较快的推理生成速度。

2. 特色功能

- 主体参照：用户上传任意主体的单张图片后，Vidu AI 可锁定该主体形象，在切换不同场景时保持主体一致性，实现连贯的视频内容生成。
- 多主体一致性：新版本支持多图参考功能，可同时处理最多 7 个不同对象（包括人物、场景、道具等），实现多主体间的形象与位置协调，还可基于参考人物与站位图片直接生成分镜视频。
- 动漫风格生成：在动漫风格视频生成中表现优异，能有效维持角色形象在不同镜头下的稳定性，避免画面跳变与风格不一致的问题。
- Vidu Q2 模型：该模型的生成质量显著提升，尤其在角色表情、复杂动作和镜头控制方面表现突出。新增“微表情精细表演”“复杂动态模拟”“精准运镜”三大核心功能，支持生成具有电影级表现力的视频。用户可自由选择视频时长（2~8 秒），并在“闪电大片模式”与“电影大片模式”中根据需求平衡视频生成速度与画面质感。
- AI 音效：支持基于文本生成音效，可精准控制音效时长与出现时间点，允许多段音效叠加使用，有效降低音效制作门槛。

3. 收费标准

每日登录平台可获得 20 个积分，用于生成图片和视频。积分用尽后需开通会员获取积分。基础会员月费 69 元，每月获得 800 个积分（约可生成 200 段视频），并享受会员专属权益。

3.3.5 RunningHUB

RunningHUB 是海马云科技公司推出的全球首个基于开源生态的图形音视频 AIGC 应用共创平台。该平台以云端 ComfyUI 为核心，用户无须配置复杂的本地环境，打开网页即可进行 AI 应用开发与内容创作，其平台页面如图3-25 所示。



图3-25

1. 基础功能

- **快捷创作功能：**提供低门槛创作入口，内置图文生成、视频剪辑等预设模板。用户无须完成复杂流程，只需选择模板并输入文字描述、风格偏好等简单参数，即可一键生成基础作品，满足快速创作的需求。
- **AI 应用模块：**整合多场景 AI 工具，涵盖图生图、文生图、文生视频、音频生成等基础功能。每个应用均提供可视化操作页面，支持调整分辨率、风格强度等参数，适合新手用户直接使用，无须专业技术背景。
- **工作流管理：**作为核心功能之一，支持用户直接调用平台内置的标准化工作流（如“照片写实生成”“短视频特效制作”），也可通过拖曳节点自主搭建个性化工作流。支持工作流的保存、分享与复用，有效简化重复的创作流程。
- **模型库资源：**建立分类清晰的模型库，按功能划分为图像生成、视频处理、音频合成等类别，收录 Stable Diffusion 全系列、腾讯混元等主流模型，并定期更新热门资源。用户可免费或按需调用模型，无须本地下载部署，直接在云端运行。

2. 特色功能

- **超全节点插件体系：**作为行业最完整的 ComfyUI 开源节点平台，已集成 7000 余个即插即用节点，覆盖图像处理、生成算法等全品类功能，节点数量与稳定性领先同类平台。节点支持“搭积木式”组合，无须编程即可实现“换脸”“动漫生成”“运镜控制”等复杂效果，避免重复开发。
- **工作流应用化快速分发：**支持将自定义工作流一键封装为 AI 小程序或网页应用，直接分发至抖音、小红书等社交平台及私域场景。用户可快速将创意转化为可复用工具，无须额外开发即可实现作品落地，打通“创作→封装→分发”的完整转化路径。
- **极速更新的顶尖模型适配：**对行业前沿模型实现“当日级”接入，腾讯混元图生视频模型、通义万相 2.1、

FramePack 长视频技术等均在发布后 12 小时内完成平台适配。平台预装 40000 余个热门模型，涵盖图像、视频、音频全模态，用户无须手动部署调试，开箱即可调用高阶创作能力。

- **移动端便捷访问：**支持通过微信小程序访问，用户无须安装专业软件或依赖高配置硬件，可随时在手机端进行操作，实现真正的移动化、轻量化创作体验。

3. 收费标准

每日登录平台可获得 100 个 RH 币，用于生成图片和视频。RH 币用尽后，可直接购买（10 元兑换 2500 个 RH 币）或开通会员。轻享版会员月费 33 元，每月获得 8000 个 RH 币，并享受会员专属权益。

3.3.6 拍我AI

拍我 AI 是海外热门 AI 视频平台 PixVerse 的国内正版产品，由爱诗科技发布。该平台采用 Diffusion+Transformer (DiT) 基础架构，在时空建模中引入自研的时空注意力机制，突破了传统时空分离与全序列架构的限制，显著提升了模型对空间与时间的感知能力，在处理复杂场景时表现尤为出色，其平台页面如图3-26 所示。

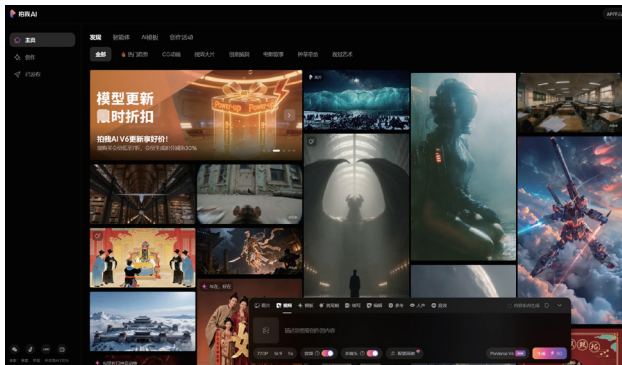


图3-26

1. 基础功能

拍我 AI 支持通过文本提示（文生视频）或上传图片（图生视频）快速生成高质量的视频内容。用户输入描述文字或上传单张图片，即可生成 5~8 秒的单视频片段，或者生成最多 5 段、总时长不超过 40 秒的连续视频。

2. 特色功能

- **多片段一致性生成：**能够实现多片段间主体、风格与场景的一致性控制。用户可一键生成 1~5 段情节连续的短视频，系统自动保持角色形象、画面色调与背景元素的整体连贯。
- **MagicBrush 运动控制：**支持通过手绘箭头精确控制视频元素的运动轨迹与速度。例如，在赛车场景中绘制长箭头可提升车辆移动速度，调整箭头方向可驱动宇航员完成挥手等动作，还可反向移动背景以增强动态效果。每个画笔都支持 6 个方向控制，实现了传统视频生成工具难以达到的精细化运动编辑。
- **角色一致性生成：**针对动漫游戏类用户，拍我 AI 推出 Character2Video 技术。上传人物正面图片后，系统可在 20 秒内完成角色建模，后续生成的视频将严格保持该角色的面部特征与服饰细节。

3. 收费标准

新用户注册即赠 120 个积分，每日登录可再获 60 个积分。积分用于生成图片与视频，用尽后需开通会员获取额度。基础会员月费 79 元，每月获得 1200 个积分，并享受会员专属权益。

3.3.7 绘想

绘想是由百度公司推出的 AI 创作平台，基于自研 MuseSteamer 视频生成大模型，专注于图像与短视频一体化创作。平台支持生成 1080P 高清内容，深度适配中文创作者需求，为个人用户与企业提供低门槛创意工具，其平台页面如图3-27 所示。



图3-27

1. 基础功能

- **短视频生成**：支持文生视频与图生视频两种模式。用户上传单张图片并搭配文字指令，即可生成最长 10 秒的高清视频，人物表情与镜头运动自然流畅。
- **创意特效**：内置毛绒世界、捏捏乐等丰富特效库，支持将图片素材一键转换为特效视频，有效提升视频的视觉表现力。

2. 特色功能

- **音视频一体化生成**：在生成视频时同步创作背景音效及人声台词，无须后期单独配音。中文语音细节还原度高，实现“画面、音效、台词”协同产出，降低多工具切换成本。
- **多模型版本选择**：提供 Turbo 版、Lite 版、Pro 版等多种模型选项。普通创作者可使用 Turbo/Lite 版快速生成基础视频，专业机构则可借助 Pro 版的高级参数（如动态细节、光影层次）打造影视级作品。

3. 收费标准

平台使用“想象力值”作为生成视频的消耗单位。用户可单独购买想象力值，10 元兑换 100 个想象力值，起购量为 100 个想象力值。也可订阅会员服务，月费 46 元，每月获得 462 个想象力值及其他会员权益。

3.3.8 Higgsfield AI

Higgsfield AI 是一款专注于特效与镜头运动的 AI 视频生成平台，其突出特色在于提供预设的电影级运镜模板，有效解决了传统文本提示难以准确描述复杂镜头运动的问题，使创作者能够轻松实现以往需专业摄影团队才能完成的动态视觉效果。

与依赖文本生成视频的传统 AI 工具不同，Higgsfield AI 允许用户仅上传一张静态图片，即可从超过 50 种运镜模板中选择或组合使用（如 360 度环绕、子弹时间、Snorricam 身体固定镜头等），模拟推轨、快速变焦、空中俯拍等专业拍摄效果。

平台通过智能算法确保画面主体与场景在大幅度运镜中保持一致性，避免出现画面“崩坏”问题，显著提升了生成视频的流畅度与专业质感。

用户也可以结合文本提示词进一步优化生成效果,但平台的核心优势仍在于其镜头模板库的丰富性与精准控制能力,特别适用于创意广告、电影短片、音乐视频等对动态视觉表现要求较高的领域,其平台页面如图3-28所示。

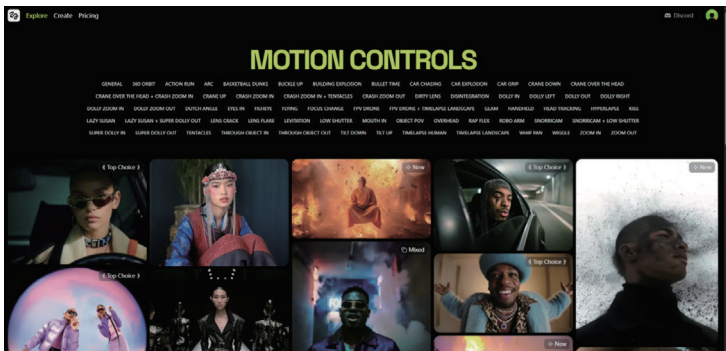


图3-28

3.3.9 LibTV

LibTV 是 LiblibAI 推出的首款专业 AI 视频创作产品，也是目前唯一一个同时为人类创作者和 Agent 设计的视频创作系统。其核心特色在于提供了一个“无限画布”操作界面，允许创作者在一个连续的空间内搭建剧本、分镜、镜头和剪辑流程，像排练电影一样组织整个创作项目。

在功能特性方面，该平台支持五大基础节点，新建、文本、图片、视频、音频，用户可以在画布上自由组合这些节点来构建复杂的工作流。此外，LibTV 还提供了丰富的图像和视频工具，如增强、扩图、多角度生成、打光、重绘、擦除、抠图等，以及视频高清、解析、剪辑和合成等功能。这些工具不仅能够满足专业创作者的多样化需求，还能通过自动化工作流大幅提升创作效率。

LibTV 的另一个显著优势在于其对多机位九宫格、剧情推演四宫格、25 宫格连贯分镜等实用工具的支持，以及电影级光影矫正和角色三视图生成等高级功能。这些特性使其特别适用于创意广告、电影短片、音乐视频等对视觉表现要求较高的领域，其平台页面如图 3-29 所示。

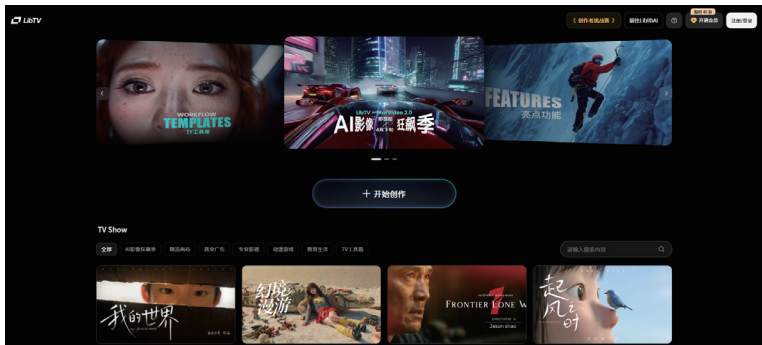


图3-29

3.3.10 Veo 3.1

Veo 3.1 是由谷歌 DeepMind 与 Google Labs 于 2025 年 10 月联合推出的新一代 AI 视频生成模型。其核心突破在于实现了导演级叙事控制能力，在延续前代音画同步生成的基础上，新增多图像场景控制、首尾帧过渡

编辑与物体级插删功能，赋予用户对视频内容、节奏及音频的精准创作控制，其模型操作页面如图3-29所示。

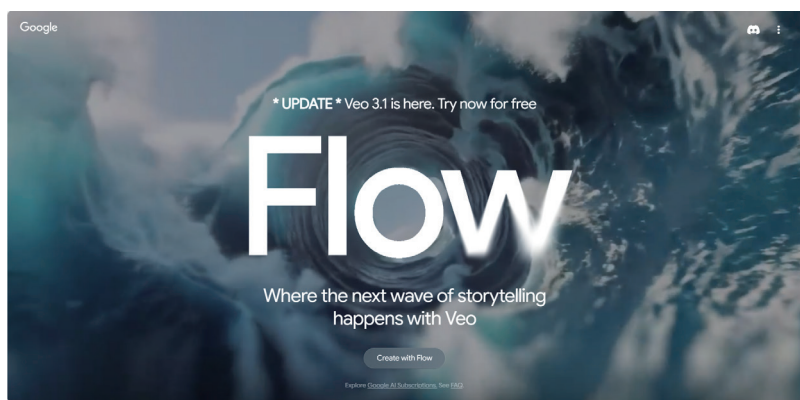


图3-30

该模型在物理模拟（如光影效果与物体运动）和分辨率方面显著超越前代，可生成长达 60 秒的 1080P 高清视频，并支持多角色复杂场景的连贯事件流生成。创作者可通过 Google Flow 工具或 Gemini 应用访问该模型，目前生成视频时长暂限制在 8 秒左右，且需配合英文提示词使用。

Veo 3.1 的推出标志着生成式 AI 正式迈入“视听一体”的新阶段，目前已应用于电影预演、广告制作等创意产业，未来有望拓展多语言支持功能。

近期全网流行的“刀切万物”解压视频主要由谷歌 Veo 3 模型生成。其核心技术在于“视觉—听觉同步渲染”，能够根据文本提示词自动生成高保真画面与匹配的清脆音效，实现沉浸式 ASMR 体验。“刀切万物”系列视频如图3-30所示。



图3-31

3.3.11 通义万相

通义万相是一款支持免费使用的 AI 视频生成工具，提供文生视频与图生视频两种创作方式。用户可通过

输入文字提示词或上传图片，自动生成具有影视级质感的视频内容。平台支持古风、科幻、动画等多种艺术风格，特别优化了对中式元素的表现能力，同时兼容多语言提示词与可变分辨率生成。

该工具不仅完全免费，还采用 Apache 2.0 开源协议开放使用。其 14B 与 1.3B 参数规格的模型权重及推理代码已在 GitHub、Hugging Face、魔搭社区等平台全面公开，用户可自由部署于各类本地或云端环境，其平台页面如图3-31所示。

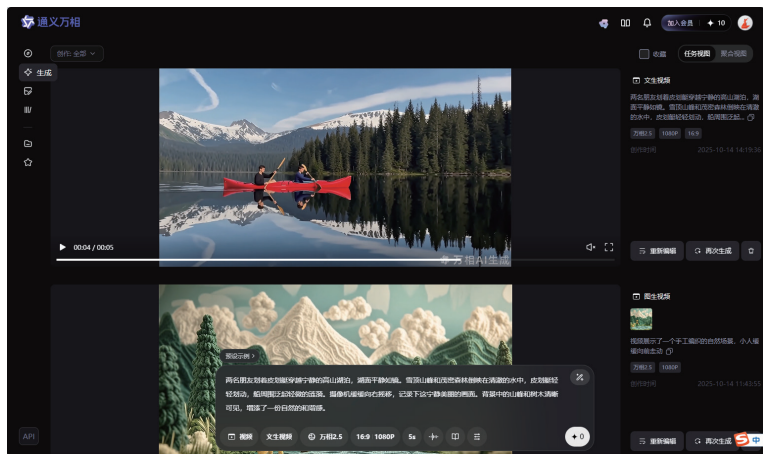


图3-32

1. 基础功能

- **图片生成**：支持通过文字描述或上传参考图像生成目标画面，提供水墨、油画、3D 卡通、二次元等多种艺术风格选项。
- **视频生成**：基于文本描述或输入视频，AI 可自动生成动作连贯的视觉内容，并支持添加运镜特效。

2. 特色功能

- **原生音画同步生成**：Wan2.5-preview 版本能够根据文本、图像或音频，同步生成包含人声、环境音效与背景音乐的完整视频，实现“音画同频”的一体化创作，避免后期拼接带来的不协调。
- **视频编辑与内容修改**：支持对已生成视频进行精细化编辑，包括调整局部色彩、删除指定对象、替换视频主角，以及基于手绘草图重新生成动作等，具备较高的创作可控性。
- **视频音效**：开启该功能后可自动为视频匹配环境音效或适配的背景音乐，增强内容沉浸感。
- **灵感模式**：通过增强画面丰富度与创意表现，可以突破提示词的表达限制，生成更具想象力的内容，适用于探索性创作场景。

3.3.12 通义万相Wan2.2-S2V

Wan2.2-S2V 与 Wan2.2-Animate 是阿里通义万相视频生成模型在 2.2 版本中的两个重要分支。以下将重点介绍 Wan2.2-S2V 模型的基本特性，如需查看完整技术文档，请参阅本书附赠的飞书在线学习资料。

通义万相 Wan2.2-S2V 是一款专注于音频驱动数字人视频生成的 AI 模型。该模型通过输入单张人像图片和一段音频，即可直接生成口型、表情、肢体动作与镜头运动高度同步的电影级视频内容，特别擅长维持长镜头中的人物一致性与叙事连贯性。

该模型支持多语言音频输入，能够自动匹配对应的口型动作与情感变化，可生成对话、歌唱、解说等多

种表演形式。同时，模型内置电影运镜语法，可自动实现推进、拉远、手持抖动等专业镜头效果。其采用的 FramePack 技术机制有效保障了长视频中人物身份、服装造型及场景环境的稳定性，并支持多人物同屏互动与复杂情绪表达的精准控制。

Wan2.2-S2V 显著降低了虚拟人视频制作的技术门槛，通过提供高度拟真的表演效果与专业镜头语言，使单人创作者也能快速制作具备电影质感的数字人短剧，极大提升了创作效率与内容表现力。该模型目前已在魔搭社区开放在线使用，其模型操作页面如图3-32所示。



图3-33

3.3.13 通义万相Wan2.2-Animate

通义万相 Wan2.2-Animate 是一款专注于动作生成与角色替换的 AI 模型。该模型能够基于一张静态角色图片和一段参考视频，生成动作与表情高度同步的电影级动态内容，在角色一致性保持和光影自然融合方面表现尤为出色。

该模型创新采用统一架构设计，同步支持“动作模仿”（Move 模式）与“角色扮演”（Mix 模式）两种生成方式。在动作模仿模式下，模型可将参考视频中的肢体动作和面部表情精准迁移至静态图片角色，实现形象动态化；在角色扮演模式下，则能在完整保留原视频动作、表情、背景及镜头运动的基础上，将视频中的原始角色无缝替换为目标图片角色。

Wan2.2-Animate 支持对真人、动漫形象及动物图片进行驱动处理，可广泛应用于创意短视频制作、舞蹈模板生成、动画制作、剧情改编、影视后期等多元场景。该模型已全面开源，开发者可通过 GitHub、Hugging Face、魔搭社区等平台获取完整模型权重与源代码，也可直接访问魔搭社区在线体验，其模型操作页面如图3-33所示。



图3-34

3.4 音乐创作平台

3.4.1 Mureka

Mureka 是由昆仑万维推出的 AI 音乐创作平台，依托自研 Skymusic 2.0 大模型与 Mureka O1/V6 推理模型，具备强大的词曲生成、旋律扩展及风格定制能力，可创作专业级完整歌曲并支持多场景商用，为音乐创作者与内容生产者提供高效的创作与变现工具，其平台页面如图3-34 所示。

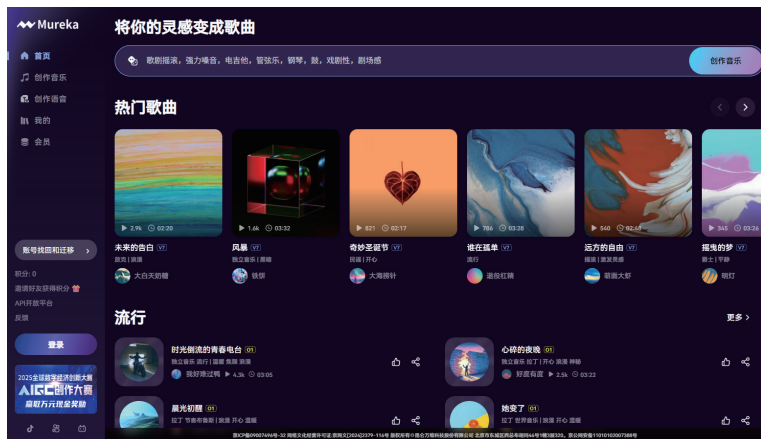


图3-35

1. 基础功能

- **文生曲功能**：输入场景描述（如“夏日海边的电子音乐”），即可全自动生成包含完整编曲与人声的歌曲，也支持纯音乐创作。
- **词生曲功能**：输入歌名与歌词（可区分主歌 / 副歌结构），一键完成旋律创作、编曲及演唱，生成结果支持直接试听与下载。
- **旋律扩展功能**：上传哼唱片段或旋律动机，智能扩展为完整音乐段落，有效解决创作中断问题。
- **风格匹配功能**：上传参考曲目或输入风格标签（如流行、古典、电子等），生成具有相似风格与情感表达的音乐作品。
- **基础编辑功能**：支持歌曲裁剪、局部修改、段落延长，并可分离人声与伴奏音轨，便于进行二次混音。

2. 特色功能

- **音色克隆功能**：上传 10 秒人声样本（清唱为佳），即可克隆生成专属音色，用于创作个性化演唱作品，平台也提供官方音色库及明星声线模仿选项。
- **双模型创作体系**：Mureka O1 模型提供专业级人声与电影级混音效果，Mureka V6 模型支持 10 种语言创作，显著提升歌词准确度与编曲质量。
- **版权交易生态**：生成作品自动获得区块链存证证书，可上架平台商店进行销售，支持单曲购买、版权授权及流量分成等多种变现方式。
- **音视频联动功能**：音乐生成后一键制作贴合情绪的带字幕视频，直接适配短视频平台发布需求，有效节省后期制作时间。

- 多语言创作：支持中、英、日、法等 10 种语言的歌词创作，可制作国风、拉丁节奏等跨文化风格音乐作品。

3. 收费标准

新用户登录可获得 20 个积分生成额度，有效期为一个月。积分用尽后，可开通会员继续使用。

- 基础会员：月费 88 元，每月可生成 180 首歌曲。
- 尊享会员：月费 168 元，每月可生成 440 首歌曲。

会员权益包括歌曲商用授权、上传音频作为参考歌曲、自定义歌手人声音色等专属功能。

3.4.2 天谱乐

天谱乐是由趣丸科技旗下唱鸭团队自主研发的全球首个多模态 AI 音乐生成平台。该平台基于专属多模态音乐生成大模型，支持文本、图片、视频、音频等多种输入方式生成音乐，可创作长达 3.5 分钟的专业级音乐作品，为个人创作者、专业音乐人及企业用户提供高效的音乐创作解决方案，其平台页面如图3-35 所示。

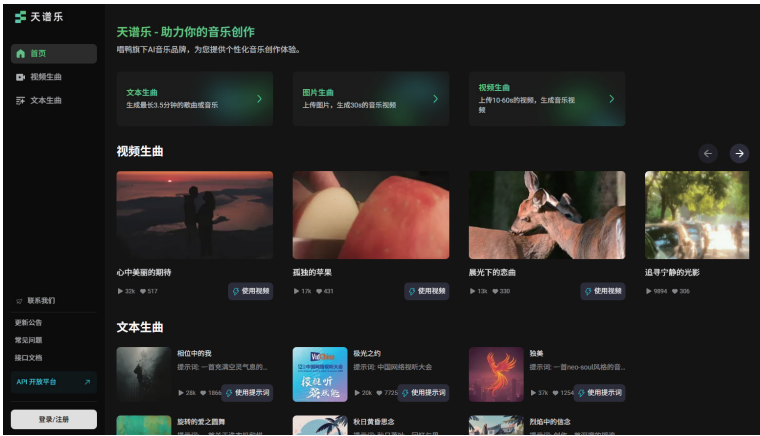


图3-36

1. 基础功能

- 文本生曲功能：输入关键词、歌词或场景描述（如“关于雨中漫步的歌”），系统通过长序列音乐语义建模生成契合的纯音乐或带人声歌曲，生成过程仅需 15~20 秒。
- 图片生曲功能：上传任意图片后，AI 通过影像识别算法解析画面内容、色彩与风格基调，自动生成包含人声唱词的完整歌曲，实现视觉与听觉的精准匹配。
- 视频生曲功能：支持上传 10~60 秒的 MP4、MOV 等格式视频，AI 逐帧识别画面明暗、情绪变化及场景特征，生成高度适配的配乐，可直接用于视频后期制作。
- 基础编辑功能：提供专家模式供专业创作者使用，支持对生成音乐的旋律、和弦、人声等元素进行精细调整，满足高阶创作需求。

2. 特色功能

- 多模态输入融合：核心优势在于支持文本、图片、视频、音频四种输入方式生成音乐，实现多源信息与音乐创作的深度融合，适配多样化创作场景。
- API 开放能力：提供企业级 API，可无缝集成至 Pro Tools、Unity 等专业软件，助力影视、游戏等行业实现批量配乐与音效生成。

3. 收费标准

- **免费权益：**用户每日登录可获得 200 个创作点，每生成一首音乐消耗 10 个创作点。
- **充值说明：**目前平台充值系统处于迭代升级阶段，暂不支持创作点购买或充值操作，用户仅可通过每日登录获取免费的创作点。

3.4.3 海绵音乐

海绵音乐是由字节跳动推出的免费人工智能音乐创作与分享平台，依托先进的 AI 技术，专注于中文歌曲创作优化，具备高效的灵感转化与个性化生成能力，可快速产出符合国内用户听觉偏好的音乐作品，为内容创作者与音乐爱好者提供便捷的创作工具，其平台页面如图3-36 所示。

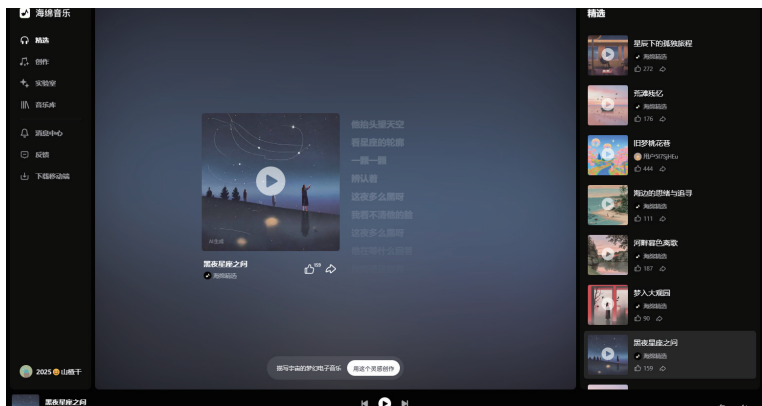


图3-37

1. 基础功能

- **灵感创作功能：**输入至少 5 个字的主题灵感（如“春夏秋冬，时间慢慢”）或上传图片，AI 可全自动生成包含歌词、旋律与演唱的完整歌曲，默认一次性生成 3 首音乐作品，单首时长约 1 分钟。
- **自定义写词功能：**支持输入歌名与歌词（可搭配主歌 / 副歌结构），提供曲风、心情、音色等参数选项，AI 完成旋律创作与演唱后，支持调整音色、修改歌词等操作。
- **作品管理功能：**生成的音乐可通过链接分享，支持自定义上传作品封面，平台设有精选榜单展示不同风格的优质 AI 生成歌曲。

2. 特色功能

- **中文优化能力：**针对中文歌曲优化吐字清晰度与演唱流畅性，减少电音感，曲风更贴合国内音乐审美偏好（如民谣、国风、流行等）。
- **多模态输入支持：**创新采用“文字 + 图片”双灵感输入模式，可基于视觉内容转化生成匹配情绪的音乐作品，实现跨模态创作联动。
- **轻量化创作体验：**页面简洁、操作门槛低，30 秒内即可完成歌曲生成，支持对已生成作品进行二次修改，适配非专业创作者的快速创作需求。
- **AI 改词功能（实验室测试阶段）：**支持基于原曲进行歌词改编，保持旋律不变，满足用户在熟悉旋律基础上创作新内容的需求。
- **AI 仿写功能（实验室测试阶段）：**可参考指定原曲的风格、节奏与情感基调，创作旋律、编曲风格相似的

歌曲，帮助用户快速产出同风格作品。

3. 收费标准

海绵音乐目前面向所有用户免费开放，无登录积分或使用额度限制，生成歌曲暂不收取任何费用。

3.4.4 音潮

音潮是由自由量级推出的人工智能音乐创作平台，依托自研“音潮音乐”大模型（采用行业领先的AR+NAR 架构），具备多模态创作与高精度风格控制能力，可生成具有工业级制作水准的完整歌曲，为专业音乐人及内容创作者提供高效创作工具，其平台页面如图3-37 所示。

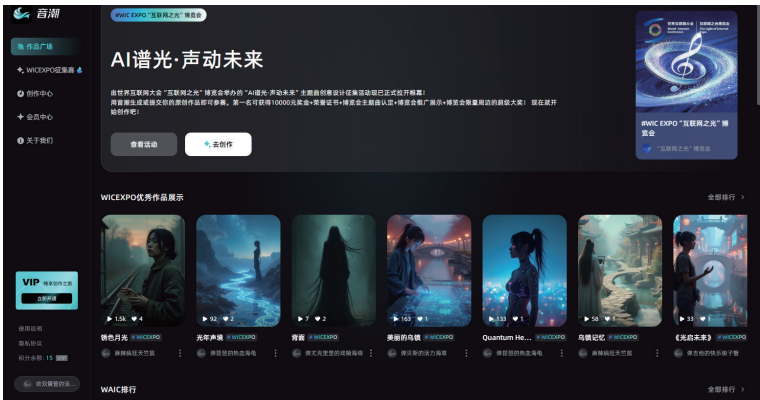


图3-38

1. 基础功能

- 极简写歌功能：输入灵感描述（如“爱情中国风，中文 + 笛子”）即可生成歌曲，无须专业知识，实现低门槛快速创作。
- 图片写歌功能：上传照片后，平台自动解析画面意境，生成贴合视觉氛围的歌曲。
- 自定义创作功能：支持输入歌名与歌词，选择流派（流行 / 中国风等）、情绪、乐器等参数；提供调速、段落修改、音色替换等编辑选项。

2. 特色功能

- 多模态创作体系：支持通过文本描述或照片上传生成带人声的完整歌曲，创作入口灵活多样。
- 专业级品质输出：生成的歌曲结构完整，编曲与和声符合专业规范，满足日常创作与分享需求。
- 版权激励生态：作品版权收益归创作者所有，可通过收益分成与 IP 运营实现商业变现。
- 全平台适配能力：支持 Windows、macOS、iOS、Android 等操作系统，用户可在不同设备上进行创作与管理。

3. 收费标准

平台创作需消耗相应积分，用户可通过开通会员获取更多积分及专属权益。

- 周会员：22 元，含 225 个积分（可写歌 45 次），享有赠分、商业授权、VIP 标识等权益。
- 月会员：60 元，含 665 个积分（可写歌 133 次），享有对应会员权益。
- 年会员：578 元，含 7200 个积分（可写歌 1440 次），享有对应会员权益。

Tunee 是由趣丸科技推出的国内首个对话式音乐创作智能体，基于自研 TemPolor V4（天谱乐 V4）多模态音乐生成大模型构建。该模型采用“需求理解+音乐生成”两段式架构，具备智能交互与专业级创作能力，可生成结构完整的多语种原创音乐及配套 MV，为专业音乐人及普通创作者提供低门槛创作支持，其平台页面如图3-38所示。

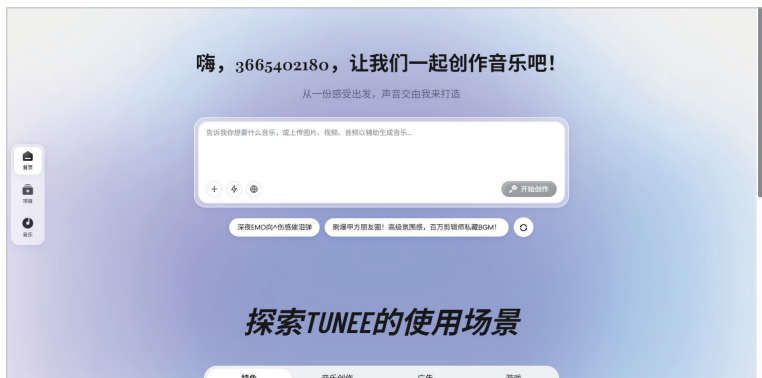


图3-39

1. 基础功能

- **对话式创作功能：**输入自然语言描述（如“缅怀青春的摇滚歌曲”）即可启动创作流程，AI 会主动追问细节并提供风格选项，无须使用专业音乐术语。
- **多模态触发功能：**支持上传音频、图片或视频文件，平台将解析内容意境与情绪特征，生成贴合输入素材氛围的音乐作品。
- **快速生成功能：**开启“快速模式”可跳过对话环节，直接生成符合基础需求的音乐内容，适合高效创作场景。
- **灵活编辑功能：**支持修改歌词、替换人声、融合乐器元素，提供调速、音轨分离（人声 / 鼓 / 贝斯等）及母带处理等后期制作功能。

2. 特色功能

- **对话式智能体架构：**具备强大的上下文记忆能力，能够记录创作者的风格偏好，通过多轮对话持续优化作品，实现“越用越懂需求”的伙伴式创作体验。
- **全链路内容生成：**集成一键 AI MV 制作功能，从分镜脚本、视觉风格设计到成片生成一站式完成，支持横屏 / 竖屏比例切换。
- **多语种创作能力：**支持中、日、英等 11 种语言的音乐生成，满足全球创作者需求，后续将针对不同语言进行差异化优化。
- **精准需求拆解：**通过联网搜索捕捉实时流行元素（如“抖音感”），将模糊创作需求转化为具体风格方案，提升创作精准度。

3. 收费标准

平台使用“创作点”作为计费单位，用户可通过免费额度及会员订阅获取。

- **Free:** 免费使用，每日刷新 150 个创作点（约 15 首歌曲），无商业授权，同时支持 5 个任务并行生成。

- **Basic**: 月费 18 美元, 每月 1800 个创作点 + 每日 150 个刷新点 (约 250 首歌曲), 含商业授权, 支持 5 个任务并行。
- **Plus**: 月费 38 美元, 每月 4000 个创作点 + 每日 150 个刷新点 (约 500 首歌曲), 含商业授权, 支持 5 个任务并行。
- **Pro**: 月费 89 美元, 每月 9200 个创作点 + 每日 150 个刷新点 (约 1300 首歌曲), 含商业授权, 支持 10 个任务并行。

3.4.6 Suno

Suno 是由海外公司推出的人工智能音乐创作平台, 依托自研的双核心模型架构——歌唱演绎模型 Bark 与器乐创作模型 Chirp, 具备秒级生成广播级歌曲的能力。该平台面向专业音乐人、短剧团队及短视频创作者等用户群体, 提供从灵感到成品的全流程创作支持, 其平台页面如图3-39所示。近期网络上广泛传播的《西游记》角色演唱视频即基于此平台制作, 视频截图如图3-40所示。

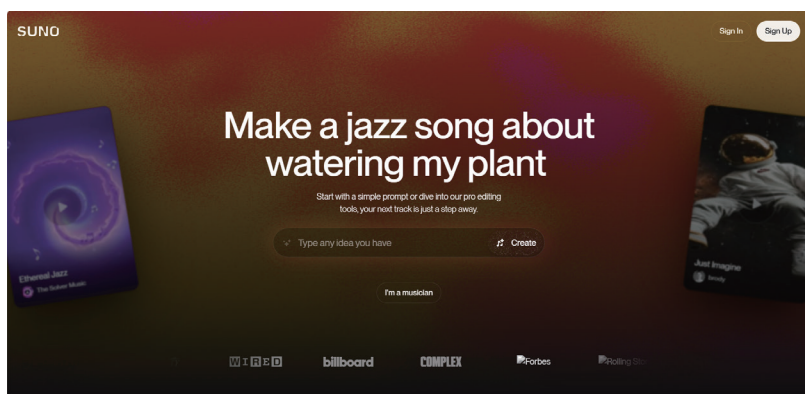


图3-40



图3-41

1. 基础功能

- **极速文本写歌功能**: 输入风格描述 (如“20 世纪 80 年代合成, 流行, 流畅爵士人声”) 或关键词 (如“太空、探索、孤独”), 系统可在 25 秒内生成带人声的完整歌曲, 提示助手可自动补全“128BPM Drop 节奏”等专业术语, 零乐理基础用户也可轻松创作。
- **人声定制功能**: 支持上传伴奏或清唱片段, 通过梅尔频谱对齐技术生成契合旋律的专业人声, 实现“即兴哼唱变专业演唱”的效果, 还可替换现有歌曲的人声部分。
- **自定义创作功能**: 支持输入歌名与歌词 (可由 AI 自动生成或手动录入), 提供流行、摇滚、古风等 12 种细分风格选择, 通过元标签定义歌曲结构; 同时支持多轨混音、EQ 调节及母带处理等专业编辑操作。

2. 特色功能

- **跨模态创作体系**: 融合文本描述、音频导入 (吉他曲 / 清唱)、Spotify 歌单分析等多种创作入口, 用户输入“机械运转回声 + 水滴金属声”等描述即可生成对应的环境音素材包。
- **专业级品质输出**: V4.6 版本支持生成最长 8 分钟音频, 采用动态压缩算法消除音频闪烁, 可导出 48kHz/24bit 无损 WAV 格式文件, 符合流媒体发行标准, 编曲和声效果媲美专业制作水准。
- **版权分级保障**: 付费创作者拥有作品完整版权, 可用于 YouTube 分成、Spotify 发行 (平台自动注册

ISRC 编码)；免费用户需标注 Made With Suno 信息，且仅限非商业用途使用。

- **全场景适配能力：**支持 iOS/Android 移动端实时协作（同步工程文件、远程调整参数），地理标签功能可自动匹配对应地区的民族乐器音色库，网页端兼容多操作系统。
- **社区灵感生态：**平台主页展示全球创作者作品，支持内容分享与交流；Inspire 功能可分析经典歌曲元素生成新曲，创作效率较传统采样方式提升 300%。

3. 收费标准

- **Basic Plan：**免费使用，每月 50 个积分（约生成 10 首音乐），无商业授权，支持 2 首音乐同时生成。
- **Pro Plan：**月费 10 美元，每月 2500 个积分（约生成 500 首音乐），含商业授权，享受快速生成通道，支持 10 首音乐同时生成。
- **Premier Plan：**月费 30 美元，每月 10000 个积分（约生成 2000 首音乐），含商业授权，享受快速生成通道，支持 10 首音乐同时生成。

3.5 后期合成软件

3.5.1 剪映

剪映是由字节跳动推出的视频剪辑与创作软件，依托自研智能算法与多媒体处理引擎，具备高效的视频剪辑、特效添加、音频处理及多轨道编辑能力，支持专业级视频制作并适配多平台内容分发。对于常规及中等复杂度的后期制作任务，该软件可提供完整的解决方案，其操作页面如图3-41所示。

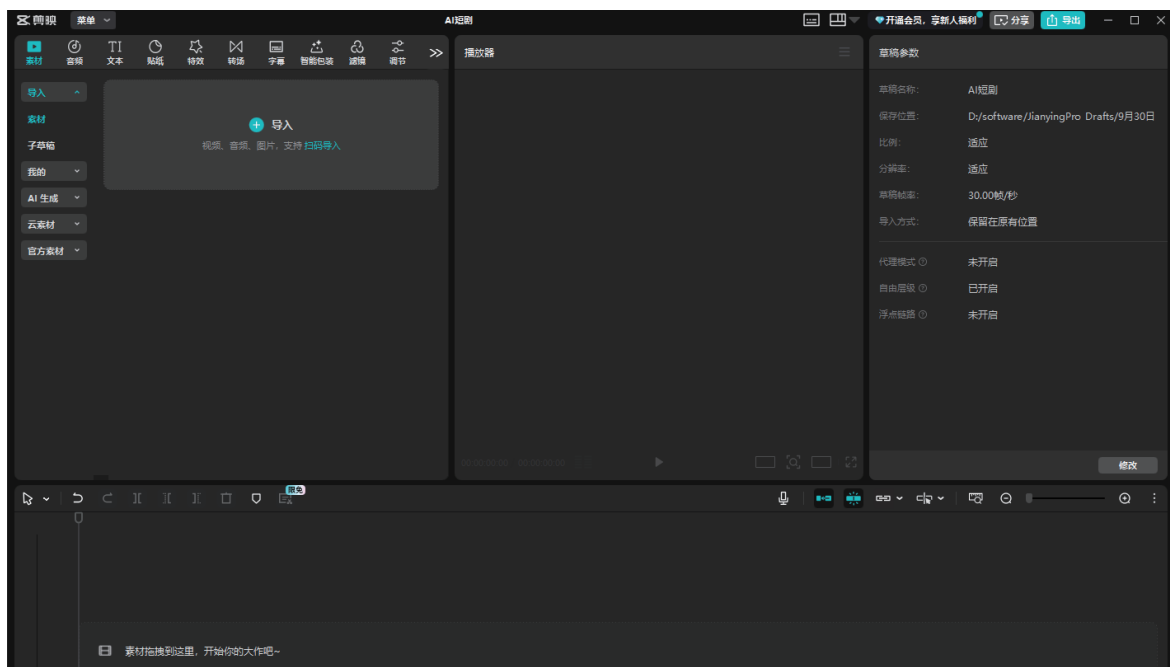


图3-42

1. 基础功能

- **剪辑功能**：支持对视频素材进行切割、删除、合并、排序等基础编辑操作。
- **音频功能**：提供背景音乐与音效添加，支持音频剪辑、音量调节及变声处理。
- **文本功能**：可为视频添加字幕、标题及贴纸等文本元素，提供多种字体样式与动画效果。
- **滤镜与特效**：内置丰富的视频滤镜、转场特效与画面特效，可一键应用以提升视觉表现。
- **变速功能**：支持常规变速与曲线变速，便于制作快动作或慢动作效果。
- **美颜美体**：提供智能美颜、塑形调整及美妆效果，可对人像画面进行优化处理。

2. 特色AI功能

- **AI 智能抠像**：基于先进 AI 技术，可一键精准分离视频或图片中的人物、物体等主体与背景，极大简化背景更换与创意合成流程。
- **图文成片**：输入文字内容或文章链接，AI 自动解析文本语义，智能匹配对应视频素材、背景音乐并生成字幕，快速将文字转化为结构完整的短视频。
- **AI 一键成片**：用户选择照片或视频片段后，AI 自动分析内容，智能匹配模板、转场效果与音乐卡点，一键生成节奏感强、质量高的短视频作品。
- **AI 数字人**：提供多种逼真虚拟数字人形象，输入文案即可驱动其进行讲解播报，显著降低口播类、知识讲解类视频的制作门槛。
- **AI 文案提词**：基于大模型技术，可根据用户输入的关键词或需求，自动生成拍摄所需的文案、脚本或口播稿，提供创作灵感支持。
- **AI 音乐生成**：输入视频内容或风格描述，AI 可自动生成或推荐氛围匹配的原创背景音乐，有效解决音乐版权与内容适配问题。
- **AI 降噪**：智能识别并分离视频或音频中的人声与环境噪声，有效抑制风声、电流声等杂音，显著提升音频纯净度。
- **AI 智能补帧**：通过 AI 算法生成中间帧，将低帧率视频转换为平滑流畅的高帧率画面，有效减少运动模糊现象。
- **即梦集成**：与专业 AI 图像生成工具——即梦深度联动，用户可在剪辑页面内直接使用 AI 生图、图生图、图像扩展等功能，为视频制作提供丰富素材来源。

3. 收费标准

剪映提供免费与付费两种使用模式。免费版本可使用基础剪辑功能及部分特效模板。付费 SVIP 服务月费约为 79 元，支持移动端与 PC 端权益通用，包含 AI 语音字幕、核心 AI 功能、智能抠像、商用素材库、4K 无水印导出及全平台云同步等全部高级权益。

3.5.2 Adobe Premiere Pro

Adobe Premiere Pro 是由 Adobe 公司推出的专业视频编辑软件。该平台基于强大的 Pro Video 引擎，具备深度集成的工作流体系，支持多轨道剪辑、高级色彩管理、音频混音及动态图形制作等核心功能，可处理从 4K 到 VR 的全格式视频制作需求。其符合好莱坞影视工业标准，为专业剪辑师、影视制作团队及创意工作者提供高效可靠的创作平台与行业协作支持，其操作页面如图3-42所示。

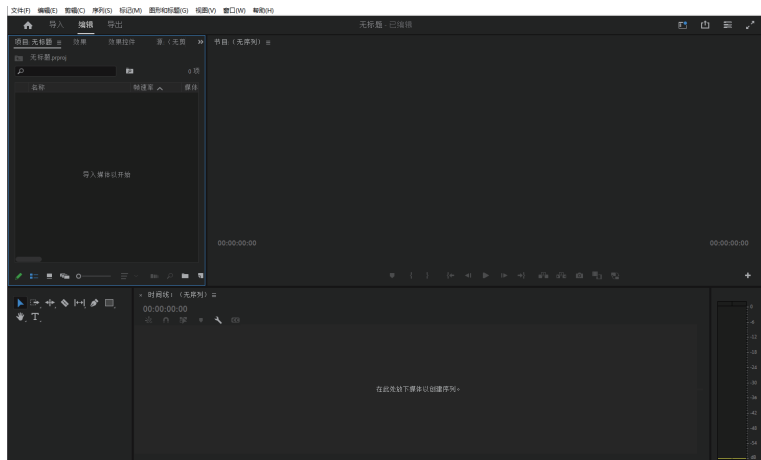


图3-43

3.6 一站式视频创作平台

一站式视频创作平台是指整合了从文本生成、图像设计到视频编辑全流程的云端或软件工具。其核心目标是通过统一的工作流消除传统多工具切换带来的操作壁垒，实现高效、低门槛的视频内容生产。

创作者输入剧本文本后，平台可自动将其拆解为分镜描述并生成对应的视觉画面。在编辑环节，用户可直接在云端完成剪辑、特效添加、内容修改等操作，平台还支持多角色协同审阅。这种全链路整合不仅显著提升了创作效率，还通过 AI 技术有效解决了角色一致性、跨镜头连贯性等传统制作难题，使从专业团队到个人创作者的多元化需求能在单一平台内实现闭环。

需要说明的是，由于此类平台高度自动化，其在创意表达精准度与流程可控性方面与传统制作流程相比存在一定局限。因此，它特别适用于快速生成样片，或者对制作时效要求较高、对画面精度要求相对宽松的项目场景。

以下将具体介绍几款代表性的一站式视频创作平台。

3.6.1 白日梦

白日梦是一款集成多项先进技术的智能视频制作工具，能够将输入文案转化为完整的视频内容。用户只需提供不超过 2000 字的文案，平台即可自动生成包含角色、场景、分镜和配音的成片，通过一站式流程大幅节省创作时间，其平台页面如图3-43 所示。

1. 基础功能

- **文生视频**：输入故事标题与正文等内容，平台自动生成包含角色、场景、分镜和配音的完整视频。
- **动态画面**：支持将静态图片转为动态视频，用户可在生成前选择旁白配音与背景音乐，并在预览页面调整视频封面、语速等参数。
- **AI 角色生成**：根据文案内容自动识别所有角色，并为每个角色设置性别、声音和形象特征。



图3-44

- 人物 / 场景一致性：保持视频中角色形象的连贯统一，提升叙事代入感与观看体验，同时提供多种视频风格适配不同故事主题与受众需求。

2. 特色功能

- 智能体模式：用户输入简单想法或关键词，AI 即可自动完成故事构思、脚本撰写与视频生成，显著降低创作门槛。
- AI 形象定制：通过“新建 AI 形象”功能，基于文字描述或上传照片定制专属数字人角色，支持重复调用，满足个性化制作需求。
- 丰富的素材库与模板：平台提供精选视频模板与多元素材资源，方便用户快速启动创作并获取灵感。

3. 收费标准

新用户注册可获得免费梦币，用于生成图片与视频。梦币用尽后需订阅会员获取额外额度。白银会员的月费为 29 元，赠送 5000 个梦币（约可生成 1000 张图片或 75 分钟视频），并享受其他会员专属服务。

3.6.2 MOKI

MOKI 是美图公司推出的 AI 视频短片创作平台，专注于为视频创作者提供动画短片、网文短剧、故事绘本及音乐 MV 等内容的智能生成与辅助制作服务。

该平台基于美图自研的奇想智能大模型（MiracleVision），集成智能剪辑、自动配乐、音效添加与字幕生成等核心功能，有效简化视频制作流程，提升创作效率，其平台页面如图3-44 所示。

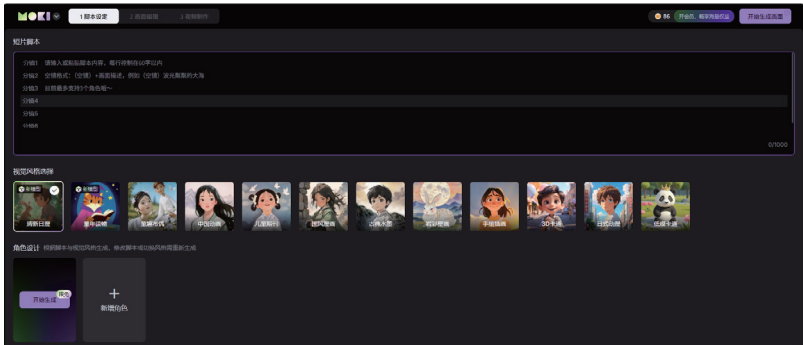


图3-45

1. 基础功能

- **AI 短片创作：**支持生成动画短片、网文短剧、故事绘本、音乐视频等多种类型内容，满足多领域创作需求。可基于简单故事创意快速构建视频脚本框架，提供情节设计、对话内容及场景转换建议，加速创意落地。
- **AI 智能剪辑：**自动识别视频关键帧并进行智能剪辑，通过图像与视频分析技术精准定位关键画面，避免人工筛选的烦琐操作。同时可根据视频节奏与情感变化自动调整剪辑节奏。
- **AI 配乐：**根据视频主题与情感氛围自动匹配背景音乐，无论是喜剧动画的欢快节奏、历史短剧的庄重氛围，还是音乐视频的浪漫情感，都能从版权音乐库中精准选取适配曲目。

2. 特色功能

- **AI 动画短片生成：**作为平台核心功能，可将用户的文字描述或创意想法直接转化为高质量动画视频。创作者无须掌握专业动画制作技能，即可快速生成风格多样的动画短片。
- **AI 创意广告生成：**提供专门的广告制作工具，用户上传商品图片后，可一键生成包含 360°环绕展示、毛绒化、膨胀飞起等流行特效的商品广告视频，极大简化电商视频制作流程。

3. 收费标准

平台使用积分制进行视频生成。用户可单独购买积分（10 元兑换 100 个积分，起购量为 100 个积分），或者通过订阅会员获取积分。基础会员的月费为 90 元，每月获得 1200 个积分（约可生成 200 段视频），并享受其他会员专属权益。

3.6.3 SkyReels

SkyReels 是由昆仑万维推出的全球首个集成视频大模型与 3D 大模型的 AI 短剧创作平台。该平台通过全流程自动化工具显著降低创作门槛，实现“一人一剧”的高效制作模式，其平台页面如图3-45 所示。

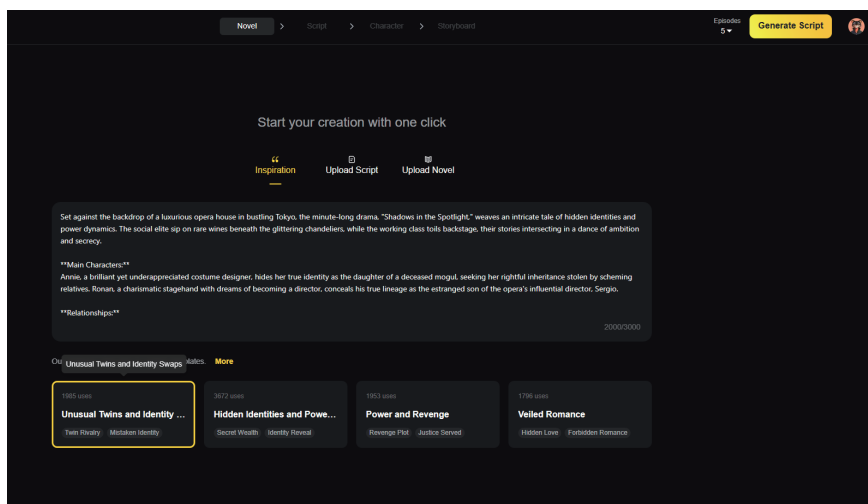


图3-46

1. 基础功能

- **剧本生成：**基于自研剧本大模型 SkyScript，支持输入概念或故事创意后一键生成完整剧本，包含情节冲突、爽点设计及情感表达等要素。同时提供剧本优化功能，可对现有剧本进行润色与结构化调整。

- **角色定制**：输入角色描述后，AI 自动生成对应的 3D 形象及配音，支持动漫、卡通、真人等多种艺术风格。
- **分镜生成**：通过分镜大模型 StoryboardGen，自动将剧本拆解为分镜脚本，并保持角色与场景的视觉一致性。
- **视频合成与编辑**：结合 WorldEngine 技术（AI 3D 引擎与视频生成模型融合），支持生成 1080P 60 帧高清视频，并提供图层融合、光照调整等后期处理功能。
- **音效配乐**：自动为生成视频添加适配的背景音乐与音效。

2. 收费标准

平台采用积分制进行视频生成，用户每月支付 7.9 美元可获得 1000 个积分。

3.7 同步生成配音视频的平台

3.7.1 纳米

纳米是由 360 集团推出的人工智能搜索与创作平台。在 AI 视频领域，该平台首创“多智能体蜂群”技术架构，基于 16 家主流大模型的协同调度，实现从脚本生成到成片输出的全流程自动化创作，支持长短视频、数字人口播等多种视频类型的制作，其平台页面如图 3-46 所示。

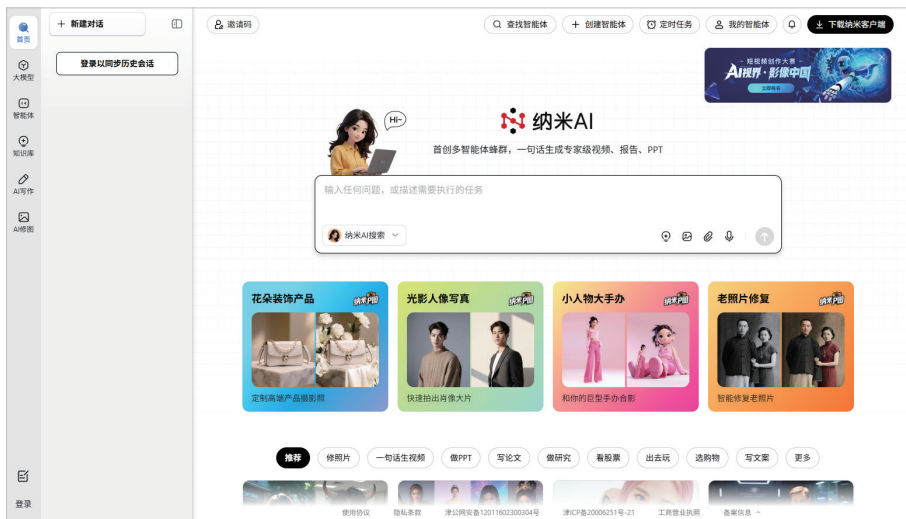


图3-47

1. 基础功能

- **文生视频功能**：输入包含主题、风格等信息的文字指令，AI 可自动生成 2 分钟以上的视频内容，支持微电影、时事新闻解读等多种类型，最长可生成 10 分钟视频内容。
- **图生视频功能**：上传单张图片后，AI 可一键生成带有分镜脚本与配乐的音乐短视频，实现静态图像的动态化呈现。

- **数字人视频生成功能**：上传个人形象及声音素材，可生成适用于电商带货、自媒体口播等场景的单人视频，支持多场景切换及服饰、动作的自定义设置。
- **基础编辑功能**：在视频生成过程中，支持时长调整、过渡效果设置、背景音乐选择等参数的自定义调整，满足基础的创意调整需求。

2. 特色功能

- **多智能体协同创作**：自动调用创意策划、智能绘图、配音、剪辑、后期等多类智能体分工协作，实现从意图分析、信息采集、脚本生成、图像创作到配音配乐等 11 个步骤的全自动化执行。
- **多角色口型对齐**：生成多角色视频时，能够精准匹配人物对话与口型动作，同时提供多音色配音选择，有效提升角色的表现力与真实感。在“记者采访《西游记》角色”等复杂场景中，能够保持整体风格的一致性。
- **全程透明可编辑**：视频生成过程全流程可视化，创作者可随时介入修改脚本、调整画面风格或替换背景音乐，避免“黑盒式”创作导致的反复生成问题。

3. 收费标准

平台使用“纳米”作为视频生成的计费单位。用户可单独购买纳米（10 元兑换 100 个纳米，起购量为 100 个纳米），或者通过订阅会员获取。基础会员的月费为 59 元，每月获得 1180 个纳米及其他会员权益。

3.7.2 Seko

Seko 是由商汤科技推出的 AI 视频智能体，作为行业首个采用“创编一体”模式的短片创作平台，该平台集成全行业主流多模态生成大模型，通过对话式交互实现从创意构思到成片输出的全流程自动化创作，其平台页面如图3-47 所示。



图3-48

1. 基础功能

- **文生视频功能**：输入包含主题与风格描述的文字灵感，AI 可自主完成剧本创作、分镜拆分、角色与画面生成，直接输出完整短片，支持短剧、创意短片等多种内容类型。

- 分镜生成与编辑功能：能够对复杂剧本进行精准拆分，解析详细分镜与角色台词，支持分镜复制、提示词上下文记忆及图像修改操作，确保剧情发展的连贯性。
- 配音配乐功能：提供 60 种自然音色选择，内置多样化背景音乐库，可根据视频风格智能匹配或手动挑选音效素材，无须跨平台获取资源。

2. 特色功能

- “创编一体”全流程自动化：基于智能体重构创作流程，整合创意策划、分镜生成、画面创作、配音剪辑等环节，用户仅需通过对话交互即可完成成片制作，无须跨平台操作。
- 多模型智能匹配：集成商汤日日新、即梦 4.0、可灵 2.1 等主流生成模型，可根据创作需求自动匹配最优技术方案，用户也可自主选择特定模型，无须额外部署环境。
- 超长分镜一致性保障：搭载高性能一致性图像生成模型，能够一次性生成角色、风格、光线保持高度统一的连续分镜，有效解决 AI 创作中的随机性问题。
- 专业级可视化编辑：提供直观的成片编辑器，视频生成过程全流程可视化，用户可随时修改脚本、调整画面风格、替换配乐，避免因反复生成造成的效率损失。

3. 收费标准

平台采用会员积分制，积分不支持单独购买，需通过订阅会员获取。标准会员的月费为 39 元，每月获得 500 个积分及其他会员权益。

3.7.3 橙星梦工厂

橙星梦工厂是由武汉风行在线技术有限公司推出的一站式 AIGC 创发一体平台，以“创作 + 分发”全链路服务为核心，依托前沿 AI 技术赋能视频短片、动态漫画、图片海报、原创音乐、小说剧本等多元化文娱内容创作，其平台页面如图3-48 所示。

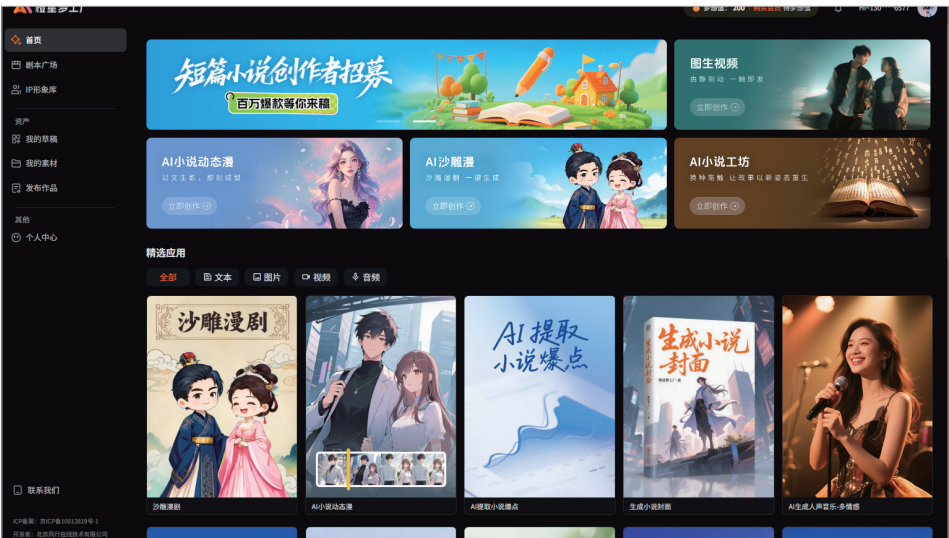


图3-49

1. 基础功能

- 文生视频功能：输入包含主题、风格、剧情走向等描述的文字创意，AI 可自动完成剧本生成、分镜设计与

画面创作，直接输出完整的短视频，特别适配漫剧、剧情短片等内容类型。

- **动态漫制作功能：**支持将静态图文内容转化为动态漫画视频，自动添加角色动作与场景过渡效果，无须专业动画制作技能即可实现内容动态化呈现。
- **剧本生成功能：**基于文本提示词智能生成小说剧本与短剧脚本，可根据创作需求调整剧情节奏与角色设定，为视频制作提供完整的内容基础。
- **配音配乐功能：**内置专业配音模块与音乐素材库，可根据视频风格智能匹配适配的配音音色与背景音乐，无须跨平台获取音频素材。

2. 特色功能

- **“创发一体”全链路服务：**整合 AI 创作与分发变现环节，创作内容可直接对接风行智能文娱分发渠道，实现从创意到收益的快速转化，无须额外对接第三方分发平台。
- **多场景内容适配：**除基础视频创作外，还可同步生成图片海报、原创音乐等配套内容，满足短视频平台、自媒体账号等多场景的内容输出需求。

3. 收费标准

平台使用“梦想值”作为内容生成的计费单位。用户可单独购买梦想值（10 元兑换 100 个梦想值，起购量为 100 个梦想值），或者通过订阅会员获取。基础会员的月费为 58 元，每月获得 660 个梦想值及其他会员权益。