

视频人工智能 与应用场景创建

Video++ MAKE VIDEO SMART

Video++ COO 董慧智



Video++ COO

董慧智

Video⁺⁺
www.videojj.com

Video++是国内最大的视频识别和多维信息流技术领导者，2015年硅谷圣经红鲑鱼评选的亚洲100强。

清华本科北大硕士，曾担任海知智能COO，获得北美创业大赛二等奖。创立影视投资公司嘉田文化，投资制作户外综艺节目《两天一夜》，在四川卫视和东方卫视播出，并获得广电总局评选的省级卫视，最具创新影响力栏目十强。



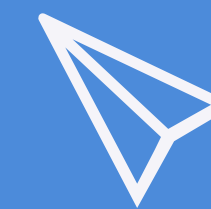
视频内容结构化



视频应用场景创建



开放平台生态



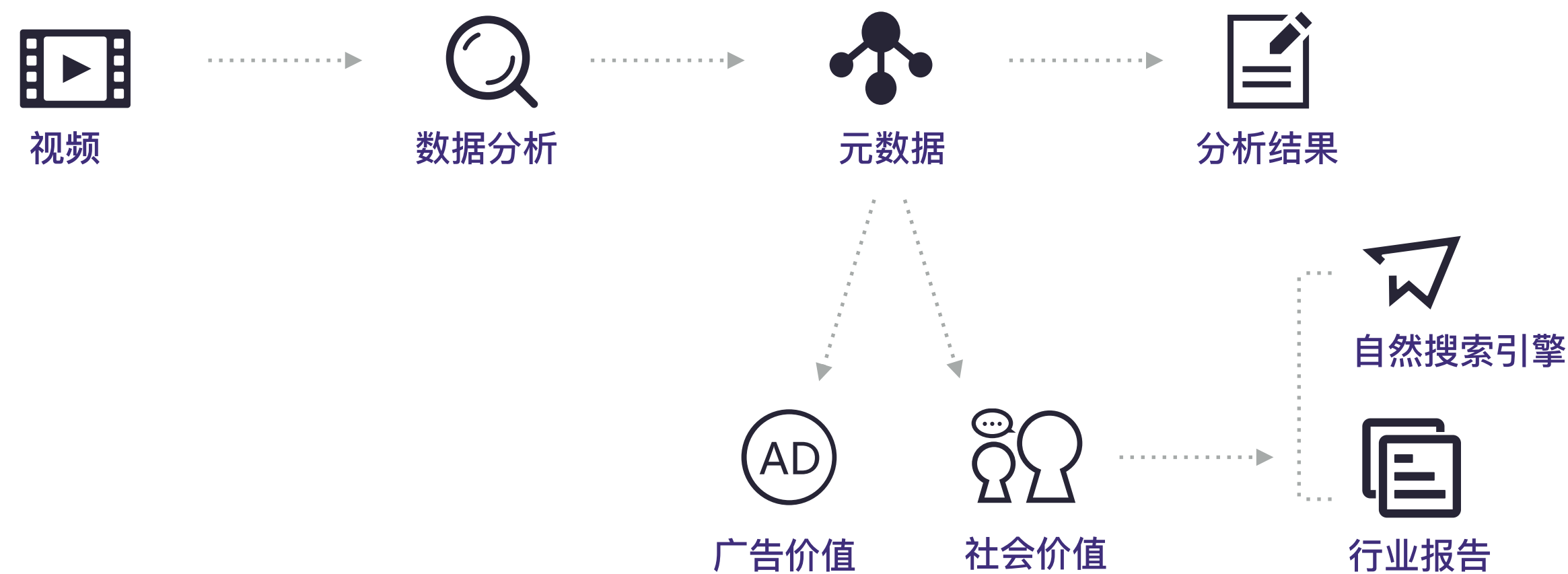
聊聊未来

专注于消费级视频的 人工智能技术

消费级视频——以观看为主要目的的视频，区别于安防等其他视频。

消费级视频的价值链：

- 1、明星或名人广告/电商价值
- 2、内容物品的植入广告/电商价值
- 3、视频内容相关的信息价值
- 4、深入视频内部的大数据价值



专注于消费级视频的 人工智能技术



视频识别与图像识别的不同:

--追踪 --时序分析;



消费级视频与安防等视频的识别要求不同:

--对误检率要求高 --对漏检率要求低 --系统级产品需求
--端对端的场景创建能力 --商业级样本库



消费级视频的关键扫描目标:

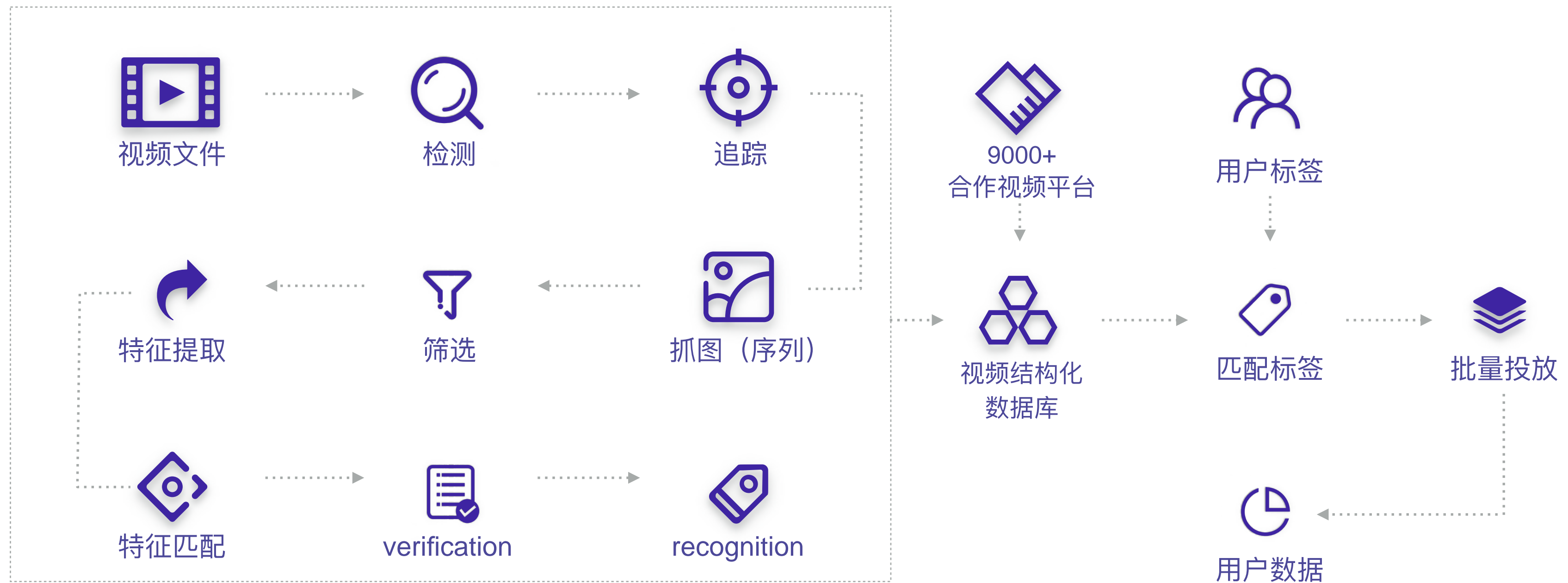
1、人物: 明星/名人/网红 2、物品: 汽车, 3C, 服装 3、场景: 海边, 酒店, 会议室
4、品牌: 可口可乐, 康师傅 5、语音: 自动识别字幕 6、语义: 场景故事 (story line)



消费级视频识别扫描的输出数据:

1、对象 2、片段起止时间 3、空间坐标 (轨迹流) 4、置信度

端对端视频扫描&全自动广告投放引擎 ——灵眸系统



视频对象追踪

- KCF跟踪算法** ● Kernelized Correlation Filters: 以跟踪对象为正样本, 环境为负样本训练的分类器
特点: 速度快, 通过傅立叶变化到频域进行计算, 避免了矩阵求逆的计算
- 多Scale** ● 多物体Scale, 最多支持7 Scale 计算
- 镜头切换检测** ● 独创镜头切换检测技术, 对光照、拉伸、渐隐变换的鲁棒性较高
第一步: 前后两帧进行直方图匹配检测, 分数低判定为切换, 分数高进入第二步;
第二步: KCF跟踪算法学习率调整, 直方图匹配分数高则调高学习率, 反之亦然
- ASMS单Scale** ● ASMS是mean shift的最新改进算法, 主要特点是速度极快
创新Scals算法, 处理形变



明星识别

通过VideoAi人脸识别人工智能算法，自动识别视频中出现的明星，平台与广告上可随意通过1.7万的库中挑选目标投放明星。用户可随时点击明星即可获得明星信息和代言商品

明星 | 名人 | 政要

17,851

1线

898

2,3线

6,378

海外

10,575

准确率/mAP

95.5%

识别特点

1 增加镜头分割作为预处理模块

2 根据消费级视频的特点，即单帧质量较差且受各类干扰因素较多，我们的识别框架摒弃了提取关键帧作为识别依据的做法，而是采用检测跟踪来提取整条人脸轨迹来进行分析的策略，设计思路是尽量获得更多帧的人脸图像信息来弥补单帧质量的不足

3 在跟踪过程中定期用检测的结果来修正跟踪结果，平衡检测跟踪两者在准确度和开销之间各自的优劣





商品识别-发掘视频宝藏 海量商品 随“心”淘

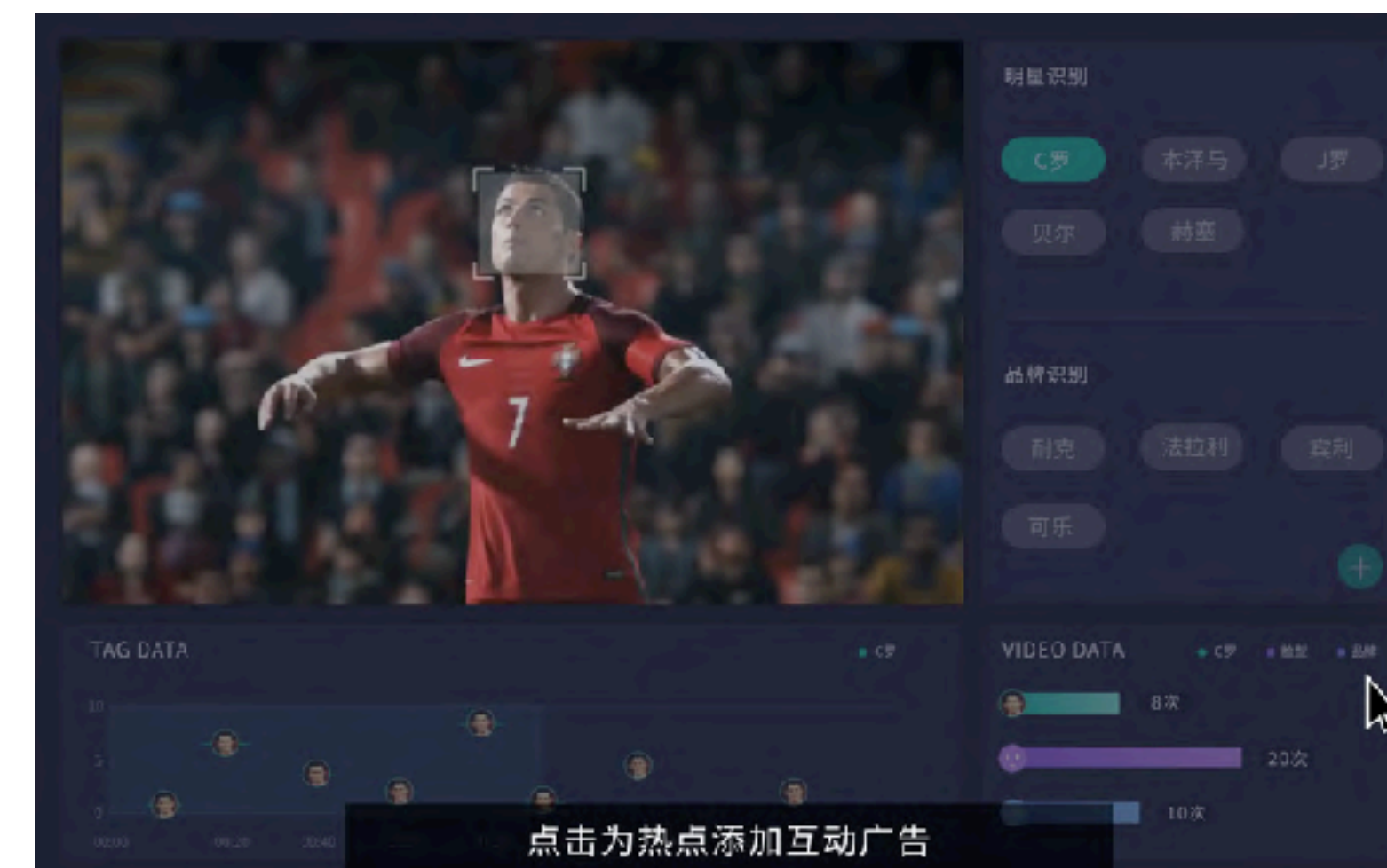
通过人工智能图像识别技术以及海量的物品库，可对视频中任何商品进行识别，看到喜欢的商品点击即可显示对应的电商链接，一键找出同款以及相似款

平台：将平台的用户流量变现，通过图像识别的商品匹配，将视频中的商品关联电商入口，转化为收益

用户：无中断，无干扰，视频中喜欢的商品一键加购

使用流程：

1. 上传 & 扫描视频商品内容
2. 通过人工智能技术识别商品
3. 关联商品的同款和相似款电商链接





物体识别

主要正对大众消费级的视频中的物体进行识别，相比较传统图片级别识别，视频里物体检测具有背景复杂，场景多变，物体遮挡严重等问题。目前已经训练模型，衣食住行共38大类，常见物体20类，另外针对广告价值较大的轿车以及手机重点训练

汽车准确率

90.1805%

手机准确率

90.205%

数码3C平均准确率/mAP

89.205%



品牌识别

驾驭内容之上将品牌信息最大化，针对Logo自动投放品牌广告，最大化品牌的曝光率，有效利用品牌的每次曝光

准确率/mAP

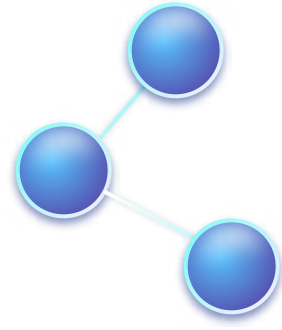
97%

品牌总量

229

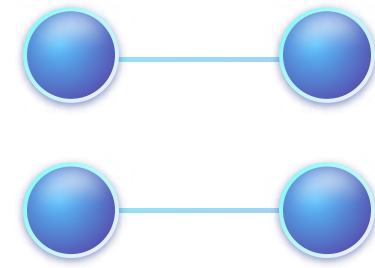


产品架构特点



算法部署插件化管理

目前人工实验室各个识别算法(明星识别, 物体识别, 品牌识别, 场景识别等)已经进入产品化阶段, 需要把实验室的研究推进到工程化开发, 我们通过docker将每个算法封装部署, 使得每个算法都能单独更新快速迭代。同时我们也能通过这种方式, 更加方便扩展, 随时增加算法类型 比如我们以后的商品识别, 语音识别等, 无论外部接入还是 原创算法 都能很好扩展



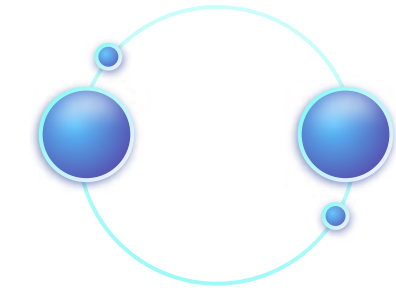
分析算法调度及与前端应用解耦

在实际生产环境中, 算法分析与应用是两套完全独立的系统。由于视频分析计算资源占用非常多, 因此必须跟前端应用解耦。目前的架构中有独立的调度系统, 使用任务队列的方式使得算法部分独立于前端应用。所有分析任务从前端发出后都不是直接调用算法, 而是通过调度系统从队列中去消费, 启动相应的算法分析, 最后由调度系统反馈到应用层。



VideoAi分析与VideoOS视频互动平台的深度结合

灵眸视频智能分析平台能够输出视频的结构化数据, 这些数据本身价值巨大, 如何实现这些价值需要与os的视频互动深度结合。灵眸与os 的视频数据可以天然互通, 无论是内部api方式或者数据库层面都能很好的支持。目前已经打通video os中的视频(即芒果, 乐视等视频平台中的视频)与灵眸分析视频的关联, 使得分析的视频数据能够直接应用到视频互动资源上, 让投放更加精准和多样化, 同时效率也更高。



样本管理系统和运营反馈机制

优质的训练样本对于提高识别精度非常重要, 因此我们设立有专门的数据采集团队, 负责训练样本采集 标注等工作, 以保证样本质量, 同时提供样本的管理系统, 方便录入, 筛选和管理。除了样本的管理, 灵眸还设立有运营反馈机制。运营人员会根据系统在运营中产生的实际识别结果进行抽样审核, 及时反馈到算法团队, 以便调整优化算法。

分析类型

可视化分析



明星/人脸



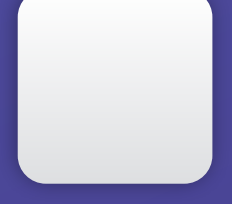
品牌



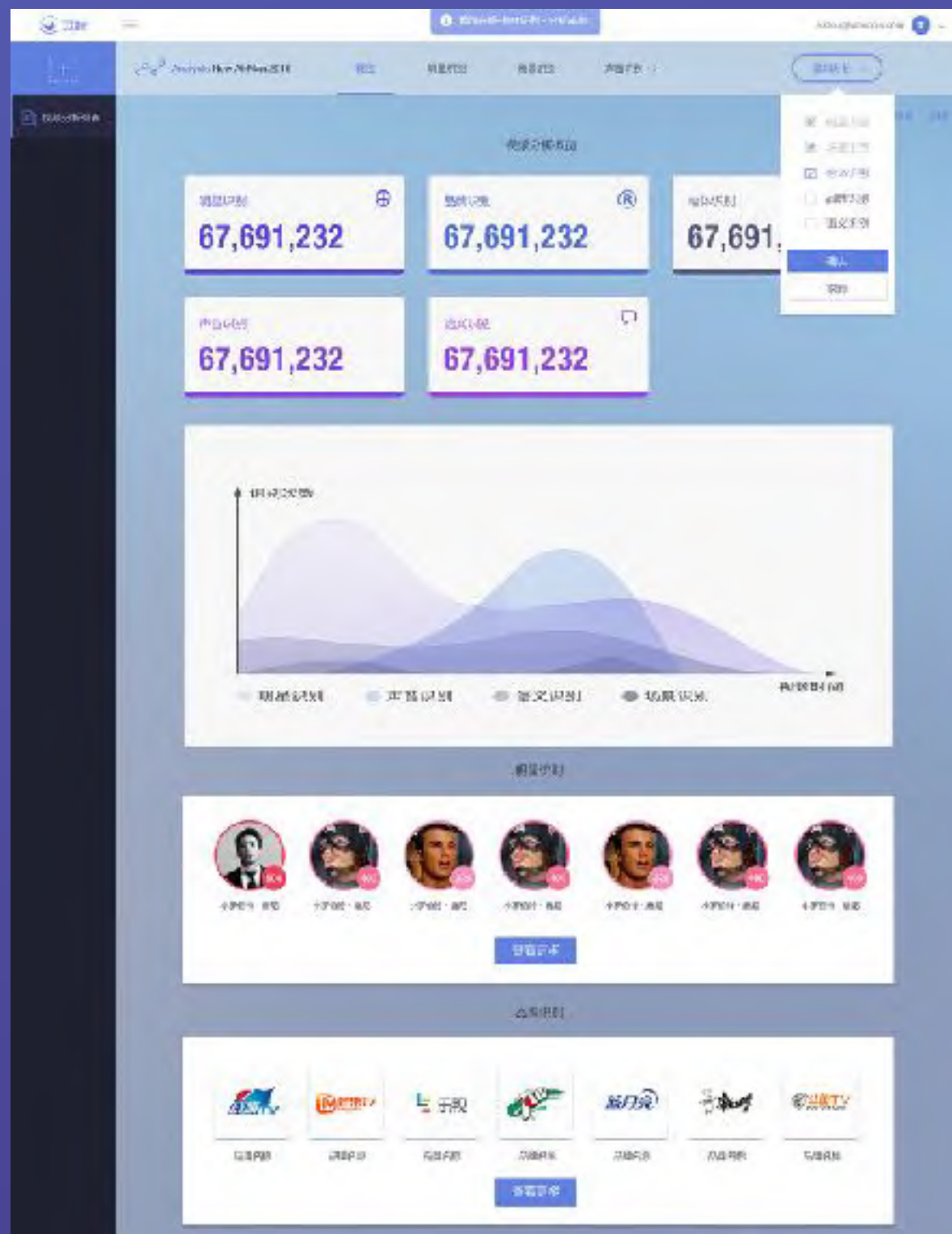
场景



物体



语音



出现次数



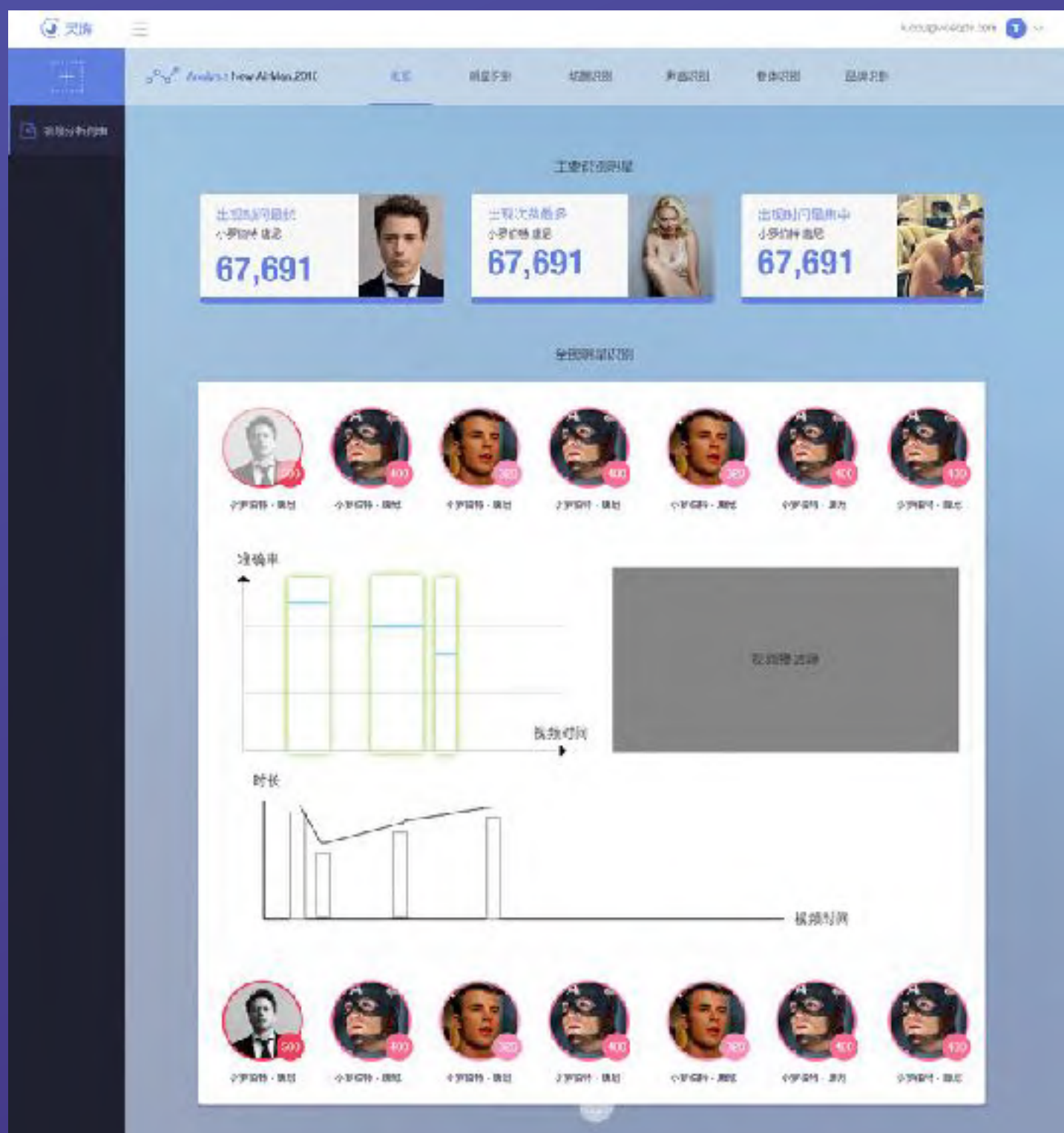
出现时间
地点



持续时长



显示位置



A blurred photograph of a man in a red shirt being embraced from behind by a woman in a black and white striped shirt. A white mouse cursor is positioned over the man's shoulder. The image is overlaid with a dark grid pattern.

扫描视频, 自动识别出现的明星, 品牌, 场景



视频内容结构化



视频应用场景创建



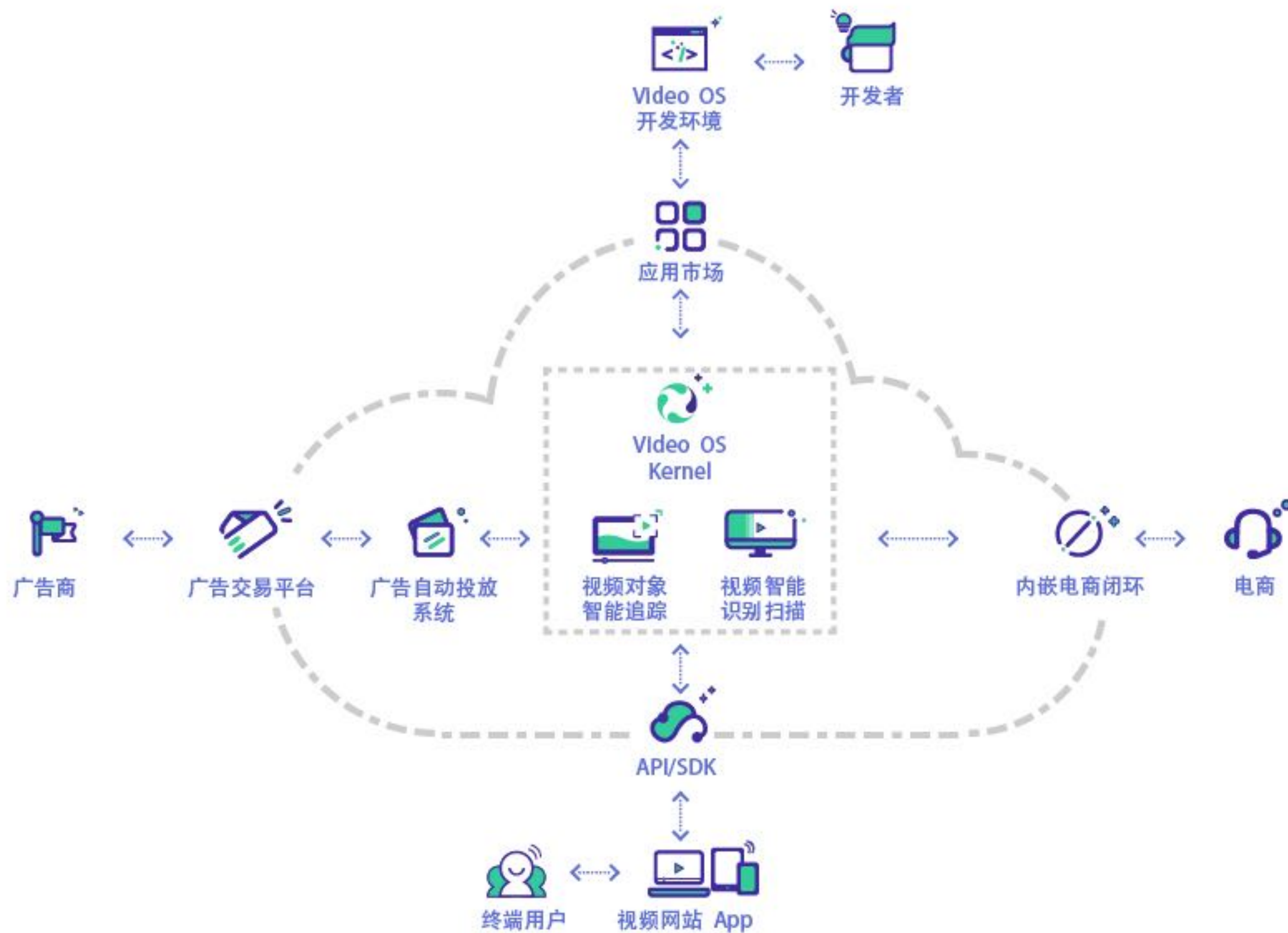
开放平台生态



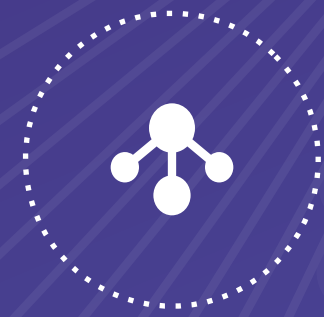
聊聊未来

Video OS 视频场景创建

让视频连接一切



视频应用场景构建 三要素



结构化视频——“模数转换”

可识别

可索引

可连接



体验美学创建情感链接

为用户提供价值

创造超越期待的体验

建立与用户的情感链接



大数据精准匹配用户情感需求

传统广告是“猜你所想”，只有用户
标签一个维度；

视频内容结构化提供了场景维度与
用户维度匹配；

视频电商场景

边看边买是伪需求——违反视频场景



• 快速收藏 一键加购

视频广告场景

隐性互动广告



还有幸福
可借

视频广告场景

显性互动广告——声波动效



视频信息场景

旅游

科莫多岛 Komodo



视频信息场景

教育



A video player interface showing a man in a yellow t-shirt with a graphic and the text "书呆子才喜欢科学!" and "and science is for nerds!". The video player includes a progress bar, a play/pause button, a volume icon, a full screen icon, and the text "高清".

视频信息场景

时尚



视频信息场景

CARTON



视频信息场景

美食



视频电商场景

边看边买是伪需求——违反视频场景



视频O2O场景

预告片——格瓦拉订票



视频O2O场景

直播订餐



互动游戏场景

直播投票

PANDA Girls
model名称
1号 韩敏英

king_A皇请韩敏英吃了 6 个佛跳墙，快来抢福袋吧！

由你来选！

倒计时
144

福 ×12

00:07:47

The image shows a live broadcast scene featuring a model in a pink top and black shorts. The scene is overlaid with various interactive elements: a 'PANDA Girls' logo with the model's name '1号 韩敏英', a chat message from 'king_A皇' about '佛跳墙' and '福袋', a '由你来选!' (Choose for yourself!) banner, a '倒计时 144' (Countdown 144) timer, a '福 ×12' (Fortune x12) reward indicator, and a '00:07:47' timer in the bottom right corner. The background includes a white office chair, a grey sofa, and a clothing rack.

互动游戏场景

卡牌游戏



互动游戏场景

直播爆灯

女神经常来 周二珂携三位美女陪你唱
榴莲娱乐 娱乐联萌 上次直播 6月18日 19:28

分享 下载APP

24,612 关注

直播公告

游戏群: 553800526 舞蹈群: 131216561
萌宠群: 511531483 微博: 榴莲娱乐Spk
@百眼 @cisco思科: 色情
我不是我你是谁: 666
小习叫我支援你: 这尼玛 好像有点吊
whamxn: 再点
九九香: 二珂好平啊
@百眼 sy_skye: 二珂快走吧
@百眼 男神偶巴巴: 66666
RaAAA: 刚刚进来这什么情况
夏哥哥来啦: 真会玩
coolhandsiq:
北北的游戏: 等二珂再点
姑娘你的短发真美: 我珂没胸会输吗
我不是我你是谁: 666
最爱gakki: 你妈哦,,色情游戏的模式
佛山_权仔: 我看到了下面
@百眼 静听北琴、: 小胸弟 讲道理
我不是女神么么哒: 老司机带带我老司机带带我

🗨️ 😊 🎁 🗑️ 📄



视频内容结构化



视频应用场景创建

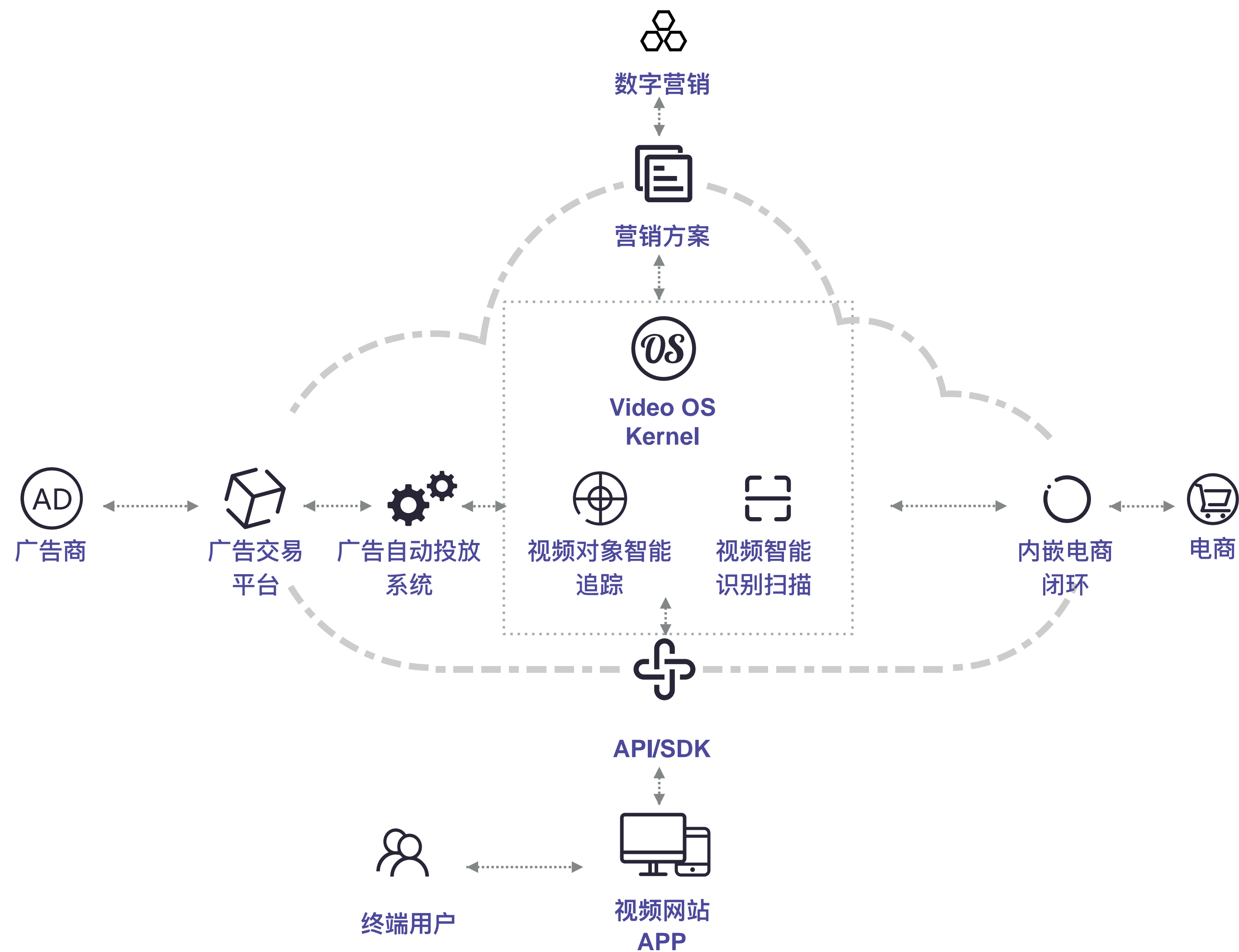


开放平台生态



聊聊未来

广告/电商/数字营销生态



视频云/直播云生态





视频内容结构化



视频应用场景创建



开放平台生态

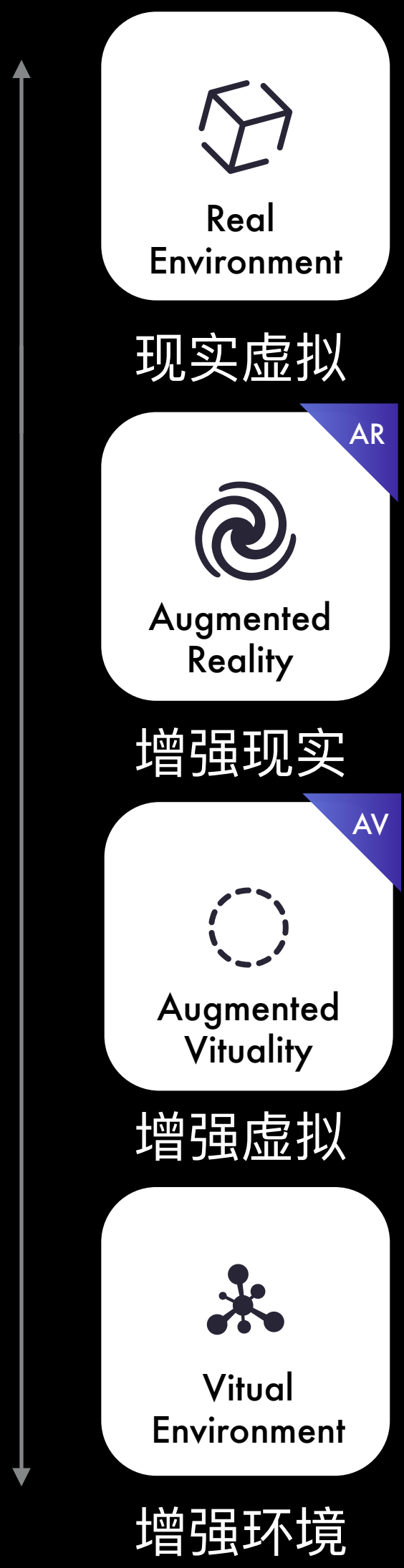


聊聊未来

VR-AR-MR



混合现实



现实虚拟

增强现实

增强虚拟

增强环境



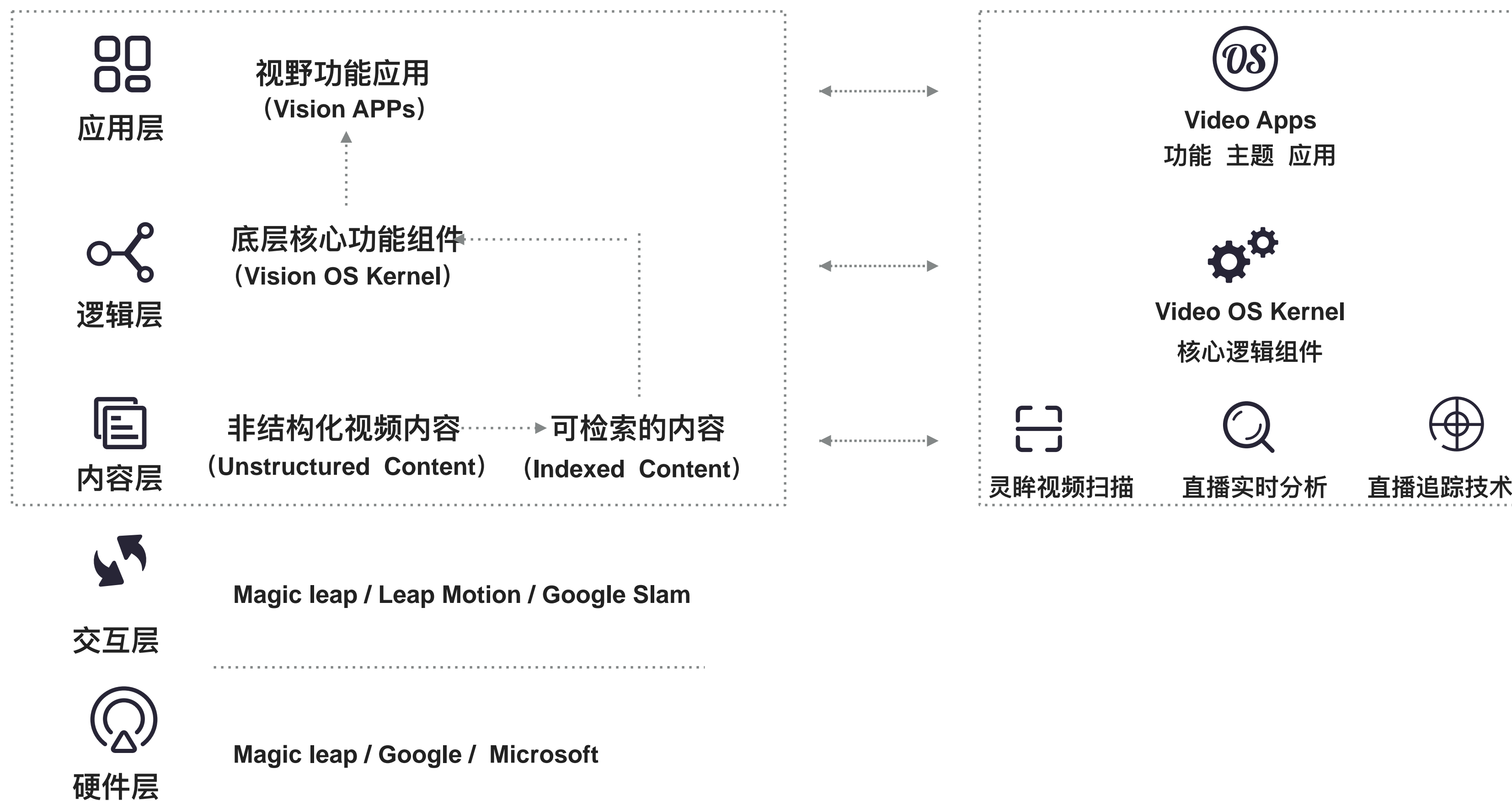
MR 下一代计算机

下一代计算机平台

虚拟光场投射 · 视觉
骨传导 · 听觉
SLAM · 空间建模
细微手势识别 · 交互



从 Video OS 到 Vision OS



THANK YOU

Video⁺⁺ MAKE VIDEO SMART