

# AI 在爱奇艺视频广告中的探索

原创 刘祁跃 DataFunTalk 2019-09-18



分享嘉宾：刘祁跃 爱奇艺科学家

编辑整理：李婷婷

内容来源：爱奇艺技术沙龙

出品社区：DataFun

注：欢迎转载，转载请留言



**导读：**本次分享的主题为 AI 在爱奇艺视频广告中的探索。AI 可以对视频内容、广告素材进行理解和加工，并应用于视频广告，尤其在广告的点位挖掘和生成、点位分析、素材创作等方面已产生价值，进而有益于广告库存提升和收入增加。本次分享的主要内容为：

- 背景介绍
- 场景：生成/推荐点位
- 效果：辅助广告素材创作

## 背景介绍

1. 广告的目的

# 广告的目的

在一定周期，让尽可能多的人产生消费

长周期：品牌广告  
短周期：效果广告

向合适的人  
传达消费价值

消费行为  
点击、下载、安装、激活、购买

爱奇艺技术沙龙

广告是在一定周期内，让尽可能多的人产生消费的一种方式。从周期来看，这个周期可长可短，长周期广告更多是品牌广告，目的是让客户记住该品牌，当产生消费需求的时候，使用该品牌的服务；短周期的则是效果广告，希望客户看到该广告时，当下就会做出一些消费的行为，比如点击、购买、下载、安装等等。在中间传递层，从广告的价值分析，向合适的人传达消费的价值，也就是让有需求的客户看到广告，并且让他感受到当前广告推荐的商品服务可以更好地满足他的消费需求。

2. 如何传达消费的价值

# 如何传达消费价值

匹配需求

展示价值

用户行为（搜索、浏览、点击等）  
统计特征（性别、年龄、地区等）  
场景（点位）+ 效果（广告素材体现）

爱奇艺技术沙龙

广告系统是一个非常复杂的系统，包括很多非常复杂的投放策略，各种各样的策略算法。它包含基本的两大功能：匹配需求和展示价值。

其中，我们如何匹配需求，比如从用户行为，人群特征等方面，了解用户喜欢做什么，从而投放匹配的广告，满足怎么比较好的传达广告的价值。这一阶段总的来讲就是通过匹配需求，把用户想看的广告实时推送到用户面前，带来广告主精准投放，并且能实现很好的展现广告的价值。

那么如何实现“展示价值”呢，我认为可能由这两个方面组成：场景和效果。场景就是我们经常听到的营销上的词语“场景化营销”，比如今天上海台风，雨伞被风吹翻折，在这个场景下，如果出现一个可以抵抗10级台风的雨伞广告，用户会很有动力买这把伞。

## AI（视频理解方面）的主要工作

生成/推荐点位

场景

辅助创作素材

效果

爱奇艺技术沙龙

从视频理解相关的 AI 范畴来讲，第一是如何找到这些适合场景化营销的场景，即广告点位，第二是摸索怎么去帮助广告展现出一个更好的效果。

### ■ 场景：生成/推荐点位

#### 1. 场景化示意

## 点位——场景化示意



创可贴



前情提要



Video In

爱奇艺技术沙龙

- **创可贴**：关键是找到一个适合出广告的场景，并且这个广告跟场景带有一定的延续性、相关性。现在已经可以规模化地为影视剧生成广告点位。

例子：用户吃饭买单时，出现支付宝广告：买单就用支付宝。

- **前情提要**：自动从上一集筛选并拼接多个小片段，能够对上一集内容做摘要，称为前情提要，在前情提要上可以贴广告。它是无中生有去制造出来的，可以根据广告主的要求自动生成不同的前情提要，作为广告点位。

例子：找到男女主高甜的场景，出美颜饮料广告，喝这种饮料，女人可以变得更美丽，让用户产生这样的关联。

- **video in**：将广告 in video，它是一种后期植入，更灵活的广告投放方式。一方面当一个剧拍得很好，后期上线后会不断引入更多客户，我们需要这种在后期可以上线的植入广告。另一方面前期植入需要客户、商务和片方导演、演员等频繁沟通协调，相对而言，后期植入的灵活性较大，也不影响前期拍摄。

例子：运动员在乒乓球场馆，他们身后有一个饮料瓶。这是后期植入的一个广告。

## 2. 视频广告如何实现场景化

# 点位——视频广告如何实现场景化

商业价值点=有消费需求的点位

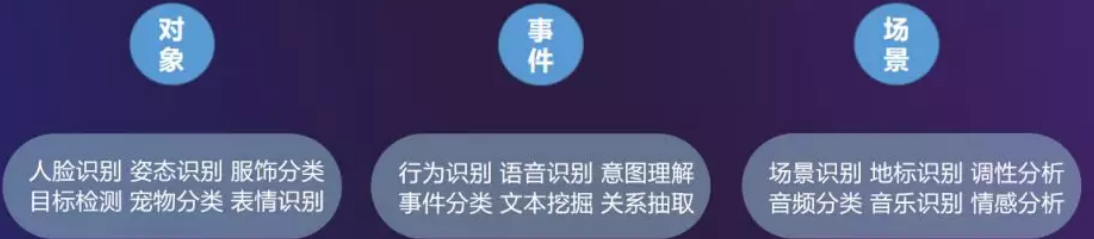
聚餐 ---- 想喝饮料	地铁 ---- 听点歌曲
海滩 ---- 希望防晒	亲吻 ---- 来束玫瑰

爱奇艺技术沙龙

刚刚的例子都显示出在一个具体的场景化中，一个商品具有什么样的效果，使得客户有身临其境的感觉，同时配合不同的说辞，很明显的感觉到广告的效果。所以核心的工作就是寻找和生成场景化的点位，这些点位同时是具有商业价值的点位，通俗来讲也就是具有消费需求的点位。从视频分析，我们会给视频打上非常多的标签。有的具有商业价值，有的具有创作价值。商业价值的点位提供给商业部门，实现变现；创作价值的点位提供给内容创作者，进行二次创作。商业价值点位不一定是精彩的点位，但它一定是能跟我们的消费场景有一定的关系。比如聚餐想喝饮料，地铁上听歌，海滩上防晒，亲吻时联想玫瑰花等等这些可能的消费场景。

### 3. 视频理解

## 点位——视频理解



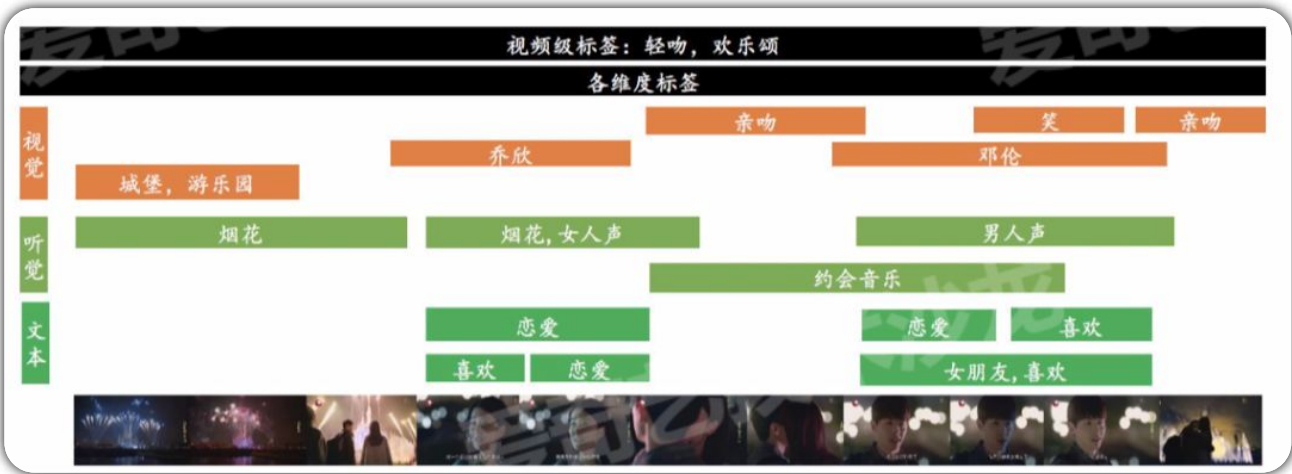
爱奇艺技术沙龙



为了找到消费场景，需要进行视频分析，也就是让机器去理解视频的内容。抽象来说，一个视频内容就是一个三元组，一个对象在一个什么场景，发生了什么事情。

- **关于对象**，这个对象可以是人，也可以是一个宠物，甚至是一个物体，对象的识别，我们可能去识别它的身份，穿着的装束，姿态等，加深我们对对象的理解；
- **关于事件**，最基础的事件可能就是我们的行为，更抽象一点的事件，比如婚礼，战争等更高层的语义。在事件的识别中，既有视觉上直接的分类识别，也会利用到语音的信息，音频的信息，台词 OCR 识别产生的文本内容等；
- **场景的识别**，即识别一种场景类型，比如这是海滩，会议室等，也有具体地标的识别，比如南锣鼓巷。包括它的调性，比如奢侈品，希望投放在一个高大上的环境比如西餐厅，同时也包括背景音乐等等，通过这些大量的底层的算法，我们可以形成对视频的密集标签，这些标签从各个维度去理解视频。我们可以创造出很多具有商业价值的标签，通过类似创可贴这样的广告方式进行售卖，但是客户要求会越来越高，比如更关注男女主之间的互动，或者客户需要更抽象一些的概念时，底层的标签不能满足。这是因为底层标签是孤立的，未产生关系。

4. 视频图谱



怎么实现对视频更进一步的理解，以下面的例子为主，我们可以从视觉、听觉、文本打各式各样的标签，也可以从部分到整体可以打上各种标签；有了这些底层标签之后，我们需要能达到满足2个需求：

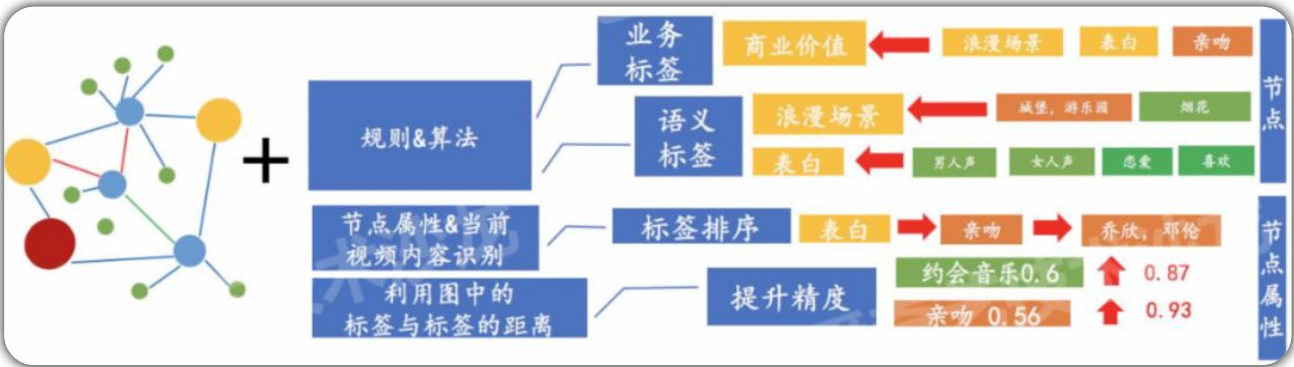
- ① 怎么去表达抽象的概念，比如浪漫；
- ② 如何判断哪些是重要的标签，哪些是非重要的标签，以及标签的精度。

当我们生成这些独立的标签之后，我们应该利用标签之间的相关性，提高标签的精度。

当各个识别算法给出这样的结论：

从场景识别，识别出宫廷；从物体检测，识别出手机；从人物装扮中，识别出古装。

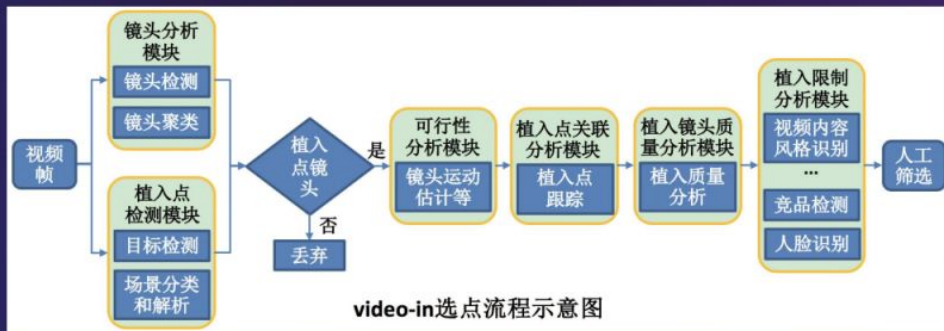
很明显物体检测出手机是有问题的，因为这些标签之间会很多的关系，比如同义词，上下级、相关性、互斥等，通过对标签之间的关系分析，我们可以提高标签识别的精度，对标签排序，找到一些更多高层标签，特别是业务更需要的标签，比如具有商业价值和创作价值的标签。



我们还会有一些外部的知识，当我们每一个算法去分析视频的时候，我们的算法大部分是机器学习的方式，尤其是深度学习，我们搜集了大量的带标注的数据，算法能实现的是这些标注数据里面体现出的知识，如果这个算法的背后模型的网络结构设计的足够好，训练的方法足够好，它就越有可能去逼近这些标注的数据，虽然我们也会做一些非监督类的算法，但这些模型学到的知识都是来自于训练的数据，这些训练数据都是一个领域范围内的数据，通过这些数据，我们会学习到哪些视频片段有什么样的行为，但是我们不知道发生这种行为的原因。比如我们识别出人物在跑步，流汗，我们很难得出人物想要喝水，想要补充体力等这些知识。这些知识不在我们训练数据里面，存在于外部数据，所以我们需要学习内、外部知识，对视频有更深入的理解，在这些深入理解的基础上，我们可以支持广告主更复杂的要求，它有很多的节点以及节点间关系，从而我们称它为视频图谱。

## 5. VideoIn 选点流程

## 点位——VideoIn选点流程



爱奇艺技术沙龙

我们要做一个饮料的后期植入，我们需要什么样的点位？简单总结一下：

- **品牌方要求：**空间足够大，不要被遮挡，镜头不要一直运动，不要运动模糊，具有场景化的需求，和明星同框等；
- **制片方要求：**解决与艺人的同框问题，不能跟竞品广告同时出现；
- **技术方要求：**技术层面是否可以植入等等。

### 6. 前情提要贴生产逻辑

## 点位——前情提要贴生产逻辑

### 精彩度分析

感官  
情节

### 转场点识别

场景/镜头  
对话

### 无效片段过滤

回忆  
主角缺失

### 剪辑逻辑

多样性  
顺序

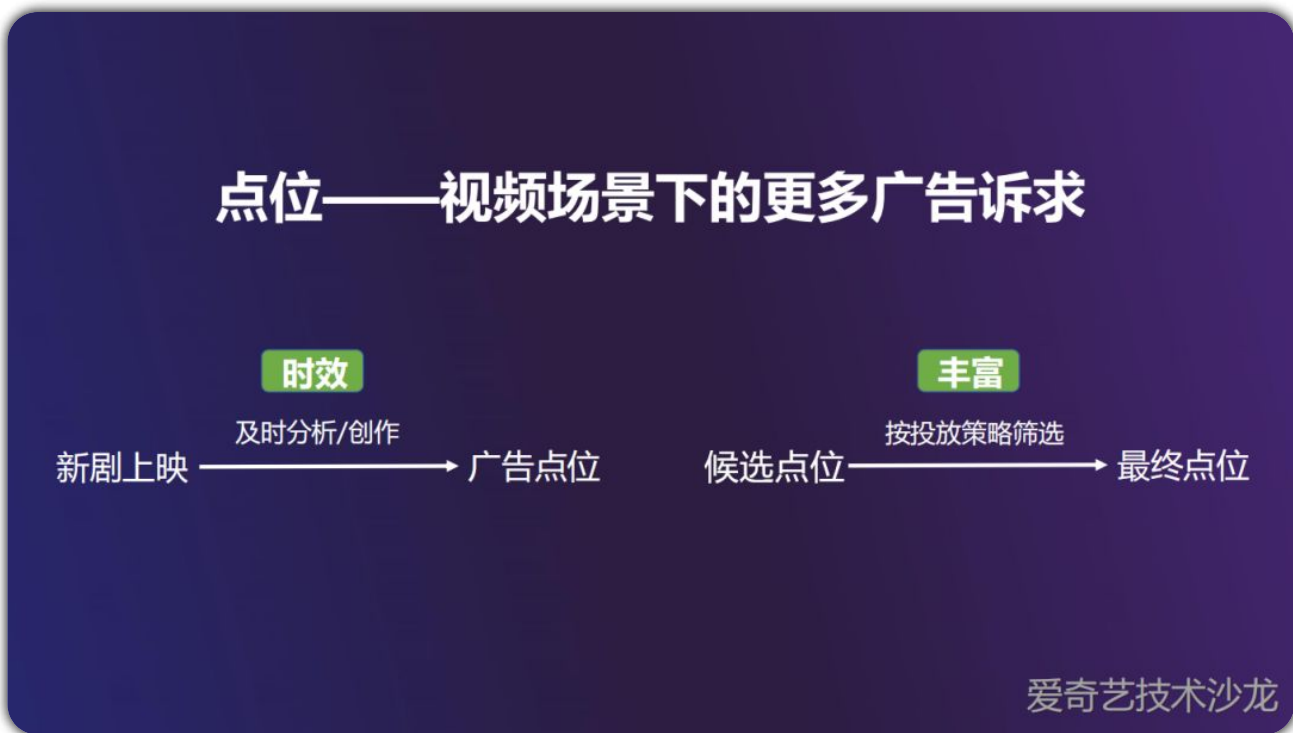
爱奇艺技术沙龙



对每一集剪辑出一个几十秒的片段集锦，它是一个兼顾内容价值和商业价值的一款产品。首先需要体现内容价值：

- ① **前情提要要精彩**，精彩是一个抽象的概念，可以拆为感官层面和情节层面；同时需要具有代表性，涵盖剧情关键的结点；
- ② **转场点识别**，主要分为场景/镜头和对话两方面，不能出现镜头/场景的突然切换，不能出现对话还未结束就被切割等；
- ③ **无效片段过滤**，比如一些回忆的镜头，一些主角缺失的镜头等；
- ④ **剪辑逻辑**，使得剪辑的视频非常丰富，也就是多样性，同时按照剧情发展的时间顺序。

## 7. 视频场景下的更多广告诉求



### ○ 时效性：

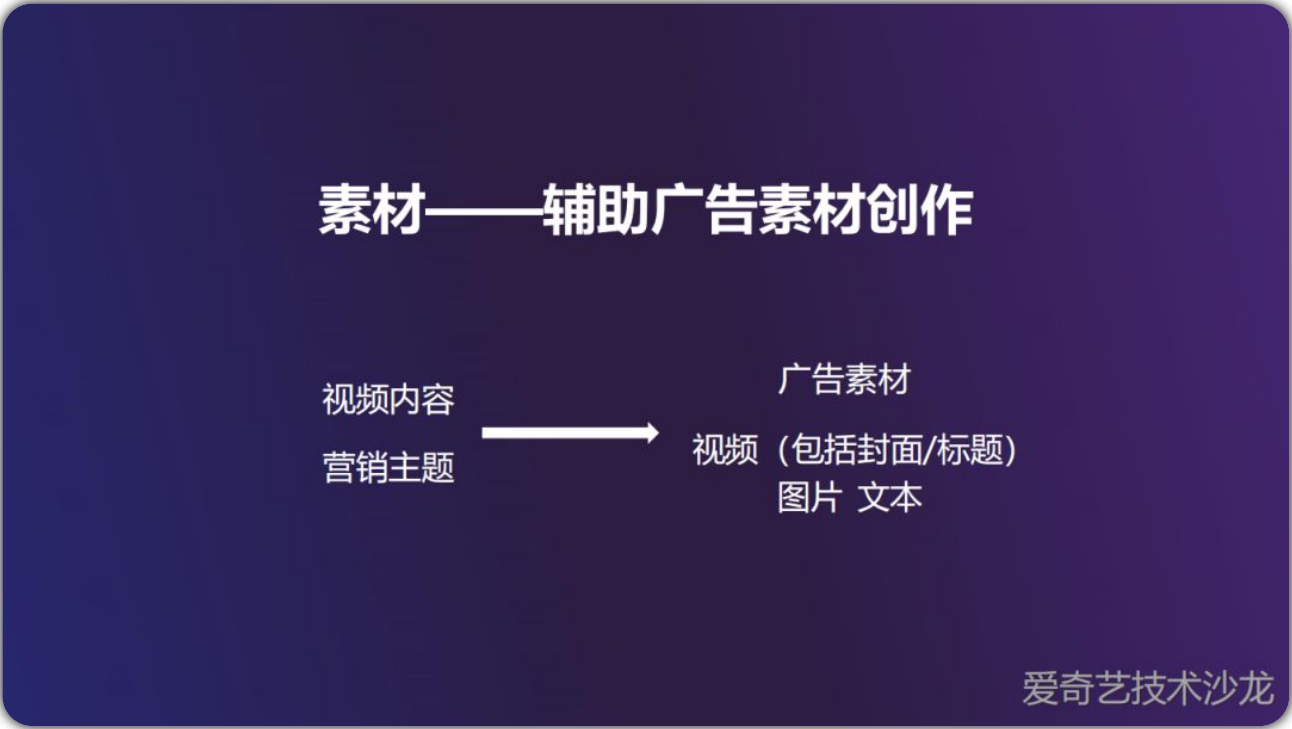
新剧刚刚上映，是否可以马上产生创可贴或者前情提要的点位，选择 AI 来做，可以克服人去做的成本问题，可行性的问题（时效要求），不可控的因素等，可以交付一个更具有鲁棒性的结果。

### ○ 丰富性：

让人去做一件事情时，想要在有限时间内发现有价值的点位，会筛选出最大可能的点位，数量有限。但一个视频可自动产生的广告点位是非常多的，我们在为广告主服务的时候，希望尽可能的满足客户需求，更关注广告点位的丰富性，这种丰富性可以提高广

告投放的灵活性。

效果：辅助广告素材创作



创作方面的探索，怎么通过图片生成视频，怎么选封面图，怎么生成标题等方面的探索。

1. 场景化下的效果展示



可以在广告方面也做类似的探索，比如扫地机器人，如果出现老人，广告的术语为：有扫地机器人，避免老人弯腰；如果出现一个比较懒的宅男，广告的术语为：解放家务活等，广告内容

会根据不同场景有不同的表现。

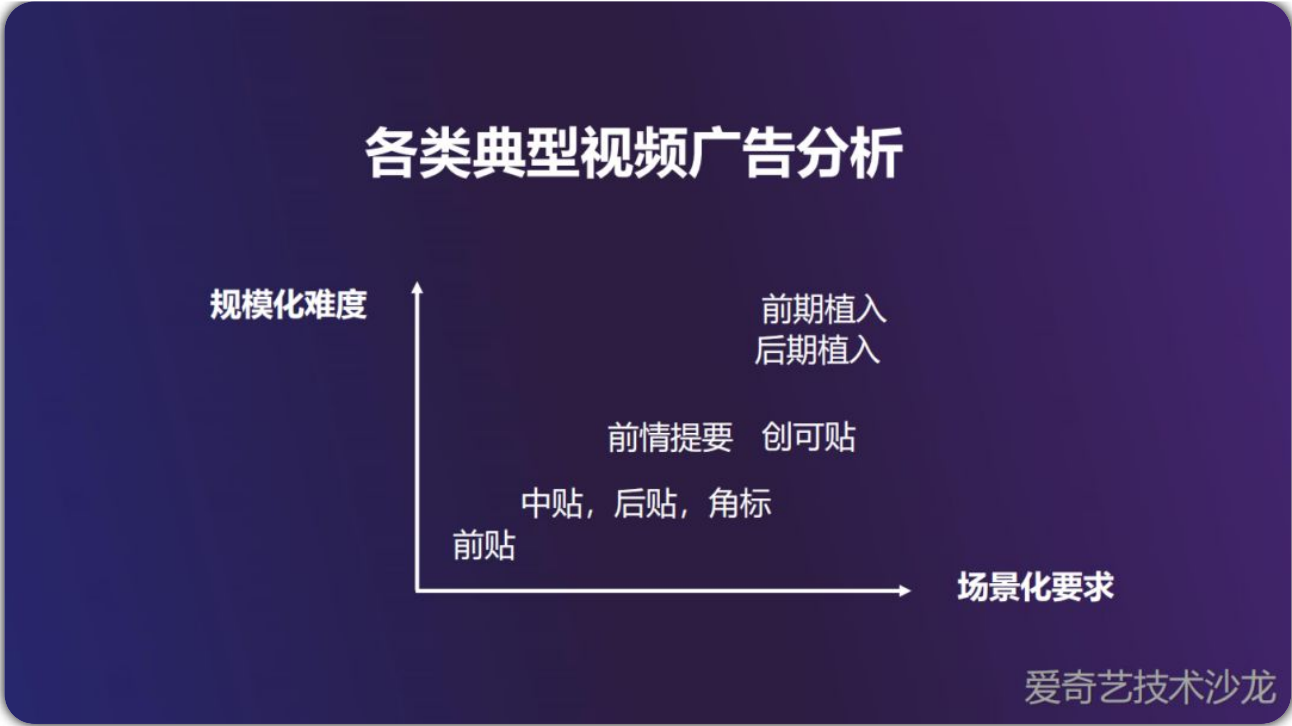
2. 视频广告的主要形式



视频广告的主要形式：

- 贴片类的：前贴、中贴和后贴；
- 浮层类的：创可贴、角标和前情提要贴
- 植入类的：前期植入、后期植入

3. 各类典型视频广告分析



我们会考虑在资源一定的条件下，评价各个广告形式实施的难度，规模化的程度。在实施难度方面：我们可以看到前贴对场景化、视频内容相关性要求最少；但植入对视频内容的要求非常多，需要大量的人工或机器做深入的分析；在规模化难度方面，植入很难规模化，但前贴是很容易规模化的。

4. AI 辅助视频场景营销



在视频场景下的营销之后会是一个非常广泛的场景，之前我们看到的很多广告，它不一定有视频的场景，但是我们走在大街上看到广告牌其实也是一个视频化的场景。只是说我们生活中看到的广告没有剧情感，但是当我们看一个故事，看一个游戏的时候，它有剧情的，所以它能提

供更多具有场景化的点位，所以视频平台，广告公司，内容创作机构都会一起来参与，那 AI 在里面，尤其是视频分析这一块 AI 的能力，主要会提供内容分析和素材的一个创作，在这样一个框架下，我们在视频场景做越来越多的探索。

### 嘉宾介绍

刘祁跃，爱奇艺科学家。从事视频分析相关算法和业务落地，其中将 AI 应用于广告是重要工作方向。

——END——

### 文章推荐：

[微博广告策略工程架构体系演进](#)

[Hulu：视频广告系统中的算法实践](#)

### DataFun：

专注于大数据、人工智能领域的知识分享平台。



一个「在看」，一段时光！👉