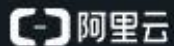




2017云栖大会·成都峰会  
THE COMPUTING CONFERENCE



# 人工智能时代下的视觉合成

主讲人：阿里云算法专家 禾朔

# 目录

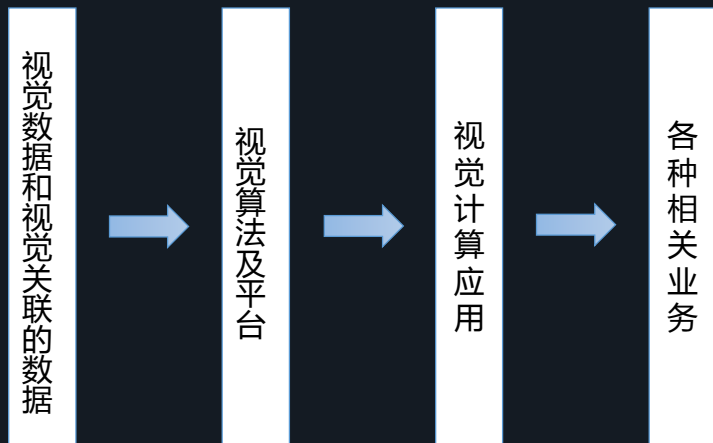
## “阿里云眼”总概

### 智能设计介绍

- 横幅布局设计

### 视频植入介绍

# “阿里云眼”视觉计算服务



## 阿里云眼业务方向



图像识别



图像搜索



视频分析



计算平台



合成



广告



设计



挖掘

- 阿里云上的视觉信息离线和实时智能分析处理中心，承载云上广泛、深入的视频图像分析、识别、搜索、生成和挖掘服务。
- 准、全、透的云上视觉已经成功应用于城市大脑、医疗健康、工业视觉、版权保护、搜索广告等等领域。

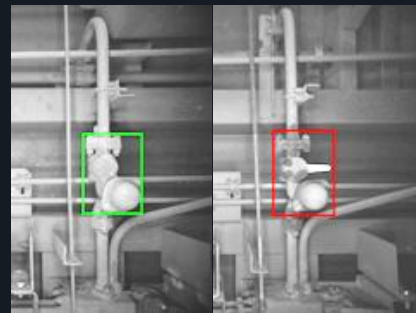
# “阿里云眼”效果展示



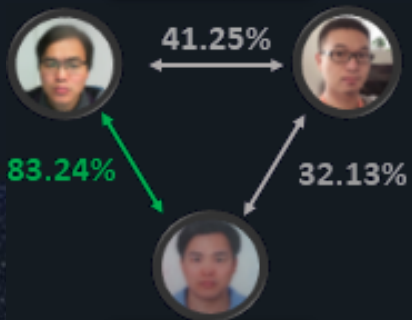
图像打标



智能设计



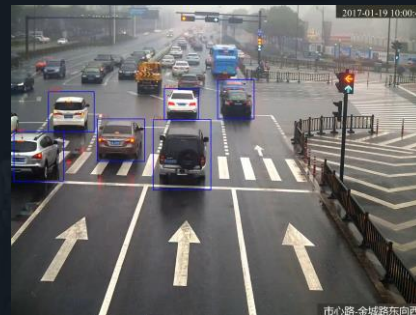
工业诊断



人脸识别



视频植入



交通监控



# “阿里云眼”效果展示



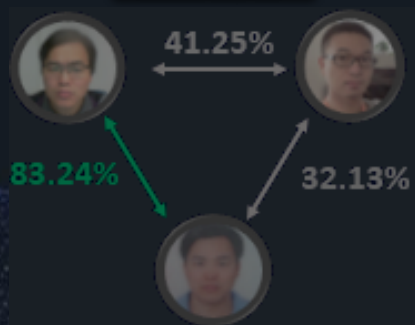
图像打标



智能设计



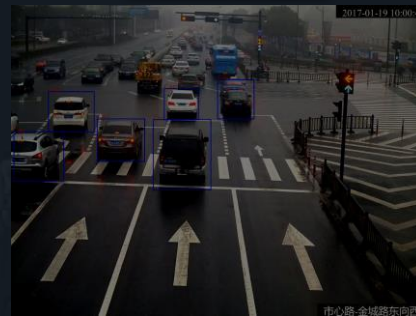
工业诊断



人脸识别



视频植入



交通监控

# 目录

“阿里云眼” 总概

智能设计介绍

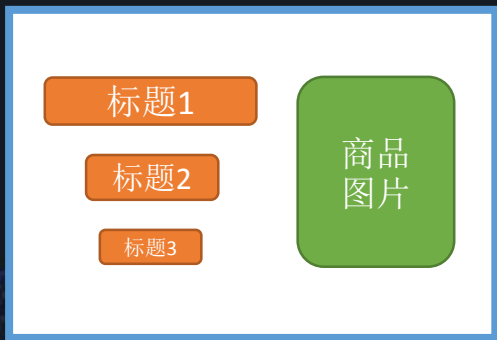
- 横幅布局设计

视频植入介绍



# 智能设计介绍

- 自动化可控视觉内容生成
- 图像或视频内容的布局、配色、背景搭配、字体合成、风格推荐、智能交互



春夏换杯  
纹理字体  
龍爪體



# 智能设计技术大图

业务层

平面设计

网店装修

出版印刷

游戏

影视

应用层

横幅设计

搭配推荐

页面设计

素材增强

设计审核

出版物排版

贴图设计

动画设计

功能层

横幅布局

背景配色

字体合成

评价与推荐

纹理合成

智能抠图

智能裁剪

风格转移

基础层

设计样本库

设计标注库

设计素材库

深度学习

机器学习

评测系统

服务平台

平台建设

服务平台

研发平台

计算平台

数据平台

监控平台





900 x 600



智能排版

下载设计

全部尺寸 ▾



900 x 600

# 时装面料盛会

辅料潮爆新

FASHION FABRICS



立即前往

智能排版

下载设计

全部尺寸



900 x 600



全部尺寸

换个风格

试试其他

下载设计



900 x 600



辅料潮爆新  
时装面料盛会  
FASHION FABRICS

立即前往



全部尺寸

布局风格



稳重大方



活泼动感



简约现代

色彩风格



浓墨重彩



清新优雅



尚酷黑白

试试其他

下载设计



900 x 600



全部尺寸

布局风格



稳重大方



活泼动感



简约现代

色彩风格



浓墨重彩



清新优雅



尚酷黑白

试试其他

下载设计



# 目录

“阿里云眼” 总概

智能设计介绍

- 横幅布局设计

视频植入介绍



# 横幅布局设计

- 设计合理的元素排列使得横幅具有特定的设计风格
- 元素包括“文字”“图片”“Logo”“背景”等
- 排列方式用每个元素的位置(x, y)和尺寸(w, h)表示



THEIR BEST FRIEND

THE BEST CHOICE

Dolls & stuffed toys

Different colors & shapes

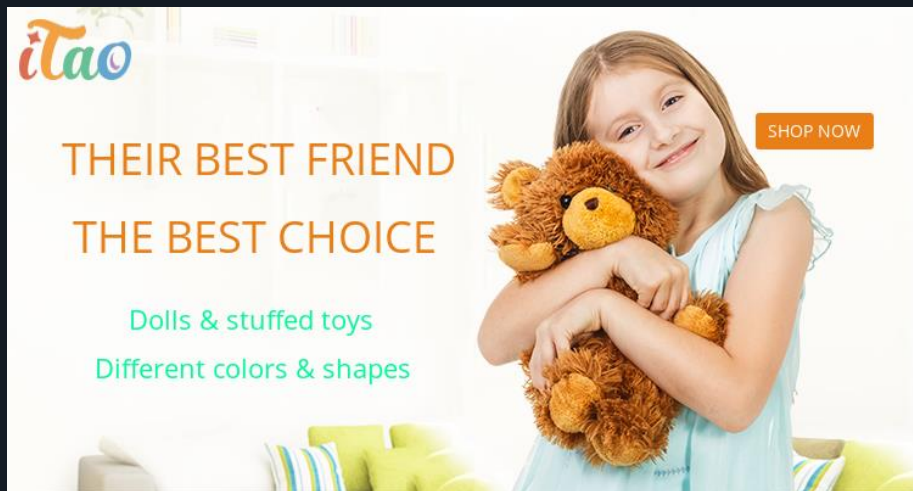


SHOP NOW



# 横幅布局设计

- 自动：最小化人工参与
  - 自动适应各种尺寸
  - 自动学习多种风格
  - 自动适应元素数量
- 可控：风格可选可预测
- 美观：符合审美习惯



THEIR BEST FRIEND

THE BEST CHOICE

Dolls & stuffed toys

Different colors & shapes

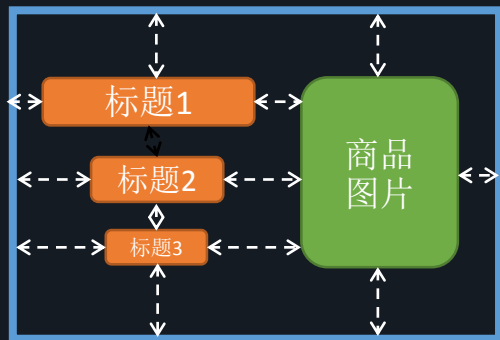


iTao

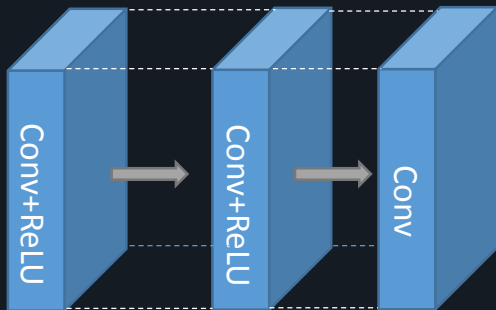
SHOP NOW



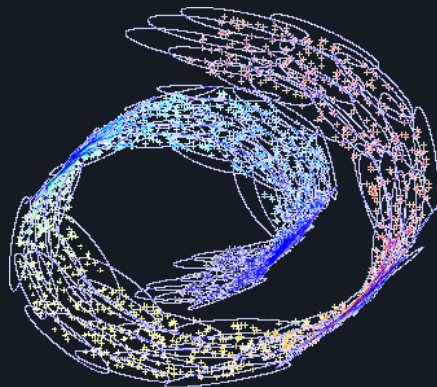
# 研发路线



- 设计模板
  - 可控性强
  - 计算成本低
  - 人工成本高
  - 通用性差
  - 扩展性差

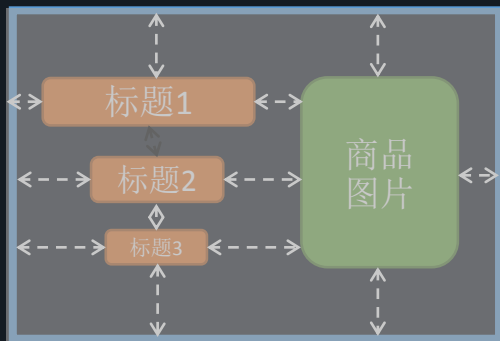


- 深度学习
  - 可控性弱
  - 计算成本高
  - 人工成本中
  - 通用性中
  - 扩展性差

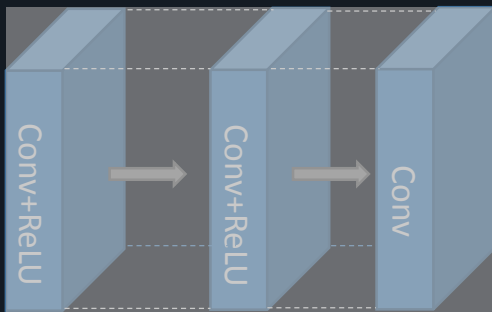


- 流型建模
  - 可控性强
  - 计算成本中
  - 人工成本低
  - 通用性强
  - 扩展性强

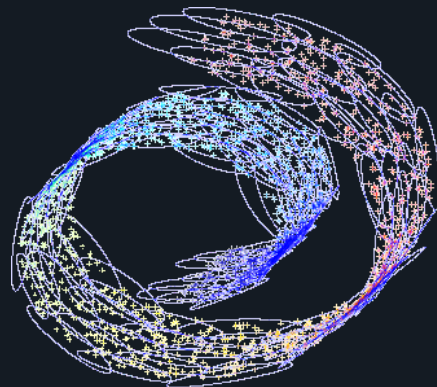
# 研发路线



- 设计模板
  - 可控性强
  - 灵活性差
  - 计算成本低
  - 人工成本高
  - 复用性差



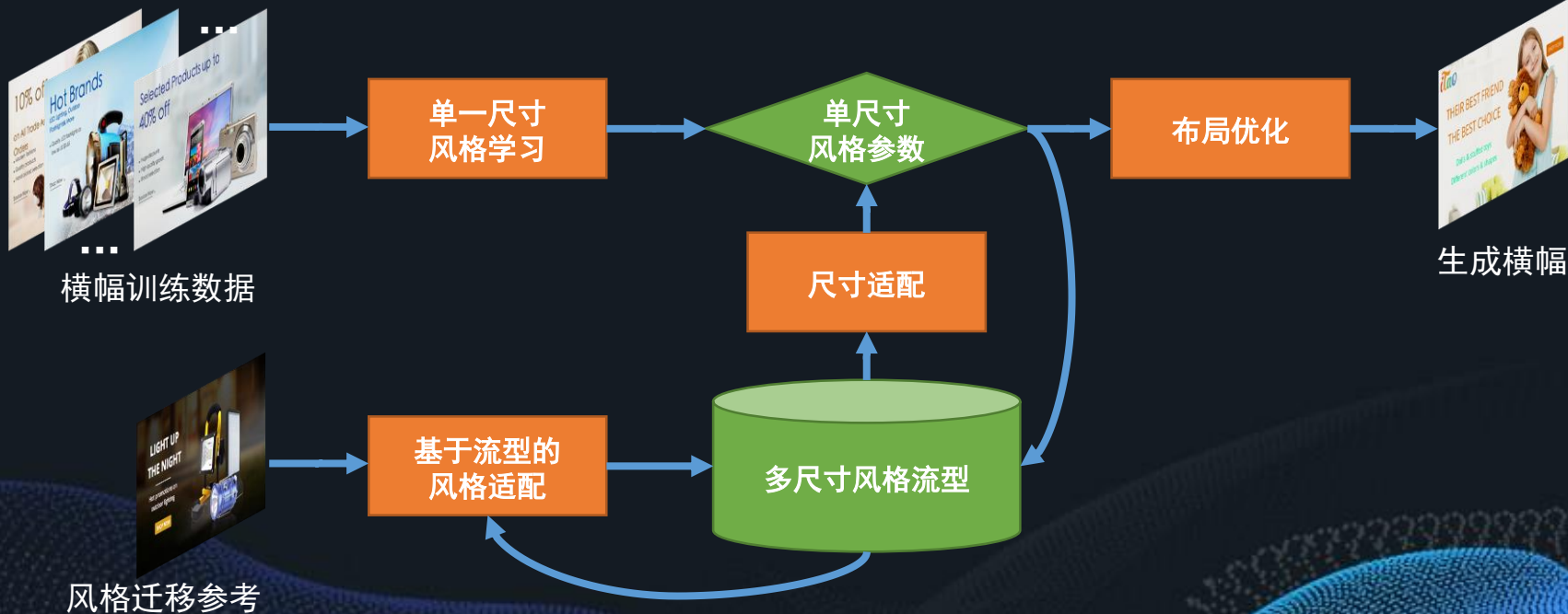
- 深度学习
  - 可控性弱
  - 灵活性强
  - 计算成本高
  - 人工成本低
  - 复用性差



- 流型建模
  - 可控性强
  - 灵活性强
  - 计算成本中
  - 人工成本低
  - 复用性强

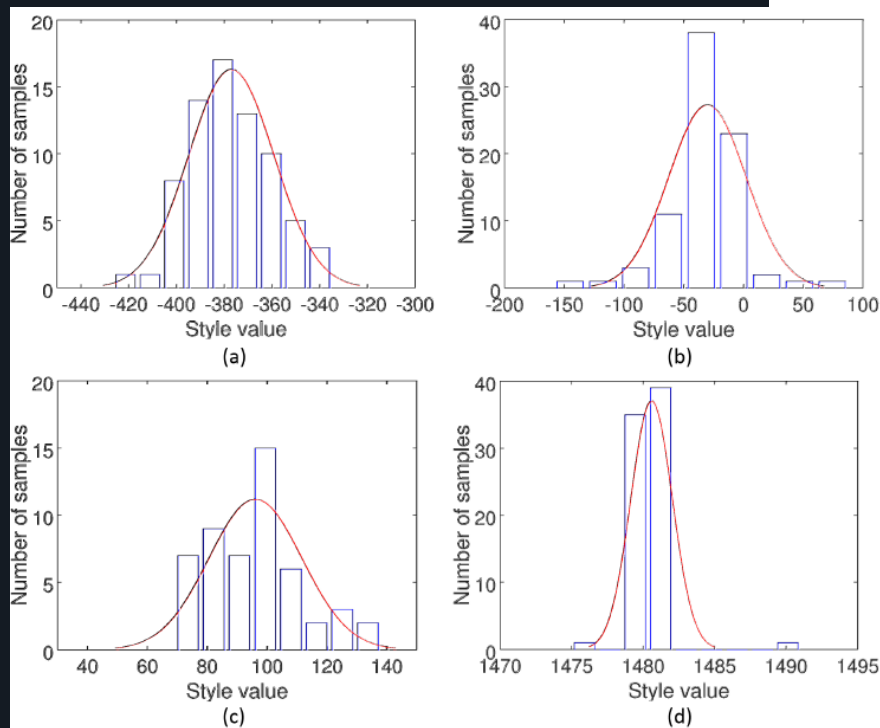


# 算法流程



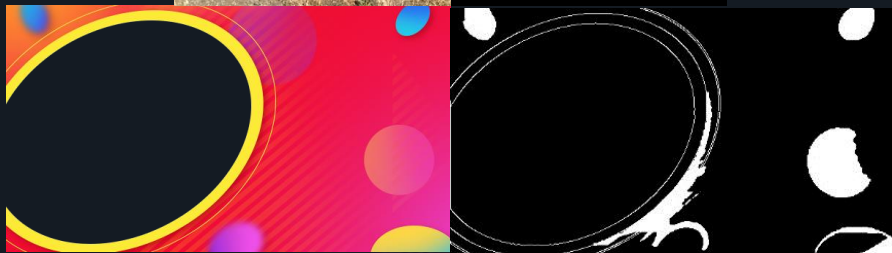
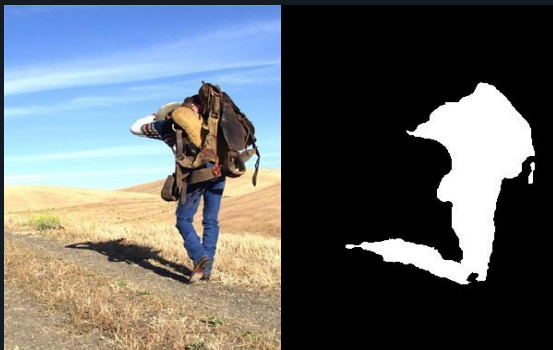
# 单一尺寸风格学习与布局优化

- 简洁的风格参数定义:  $\Theta$ 
  - 页边空白、相对位置、元素尺寸
- 基于高斯分布的风格模型能量函数
 
$$E(X, \Theta) = \sum_{k=1}^n Z_k^2 = \sum_{k=1}^n \left| \frac{f_k(X) - m_k}{\sigma_k} \right|^2$$
- 基于EM方法风格参数学习与布局优化
  - 风格学习: 已知横幅  $X$ , 求  $\Theta(\sigma, m)$
  - 布局优化: 已知  $\Theta(\sigma, m)$ , 求横幅  $X$



# 显著性检测

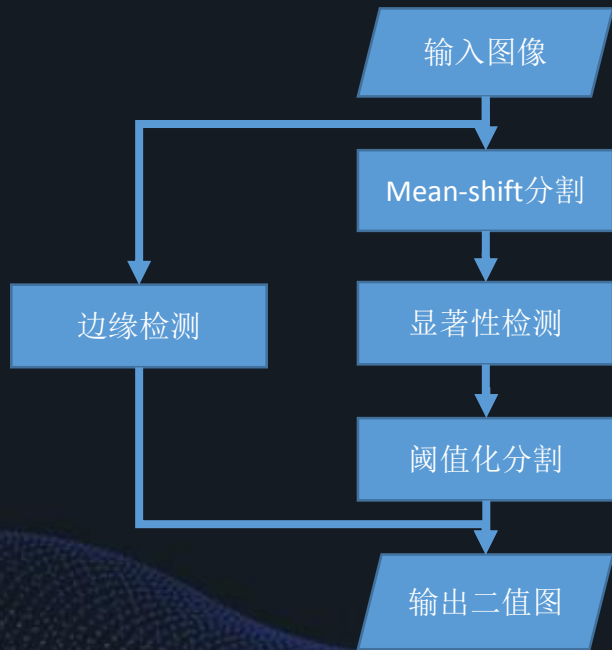
- 自动标记图像中“重要”区域
  - 自然图像 <-> 合成图像
  - 块状区域 <-> 线状区域
  - 重要区域跨越图像边界



输入图像

输出显著性mask

# 显著性检测



原始图



Mean-shift分割



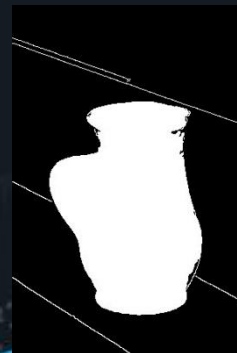
Canny边缘检测



显著性图



显著性二值图



结果图

# 尺寸适配

- 屏幕与窗口大小任意变化
  - 不同尺寸风格必须一致
  - 设计随尺寸变化不连续
- 设计多个尺寸代价不菲
- 简单拉伸压缩效果不佳





# 尺寸适配

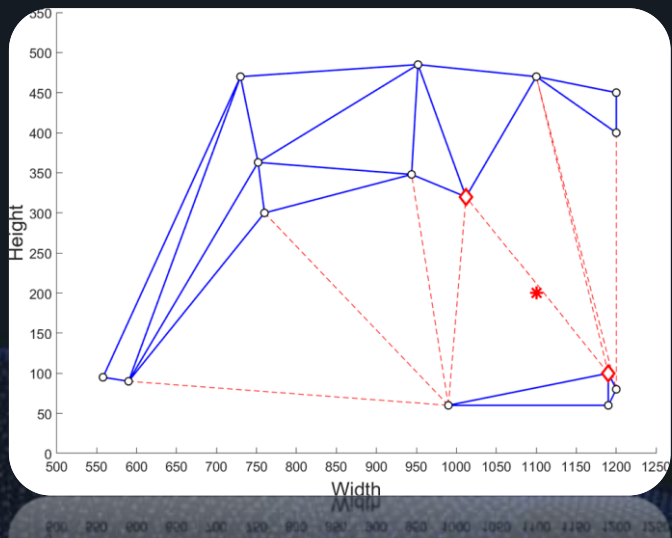
- 设计随尺寸变化不连续
- 直接插值效果难料



?

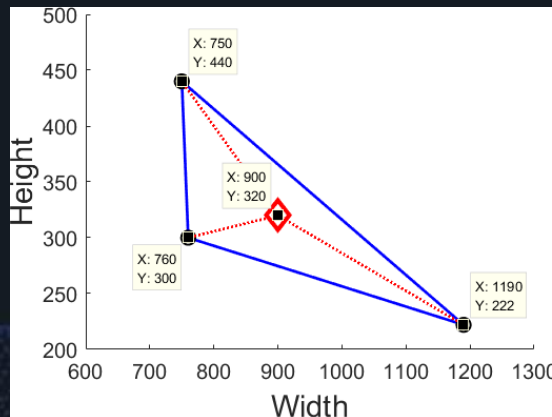
# 尺寸适配

- 基于流型分析的风格插值
- 考虑尺寸变化引起的设计突变



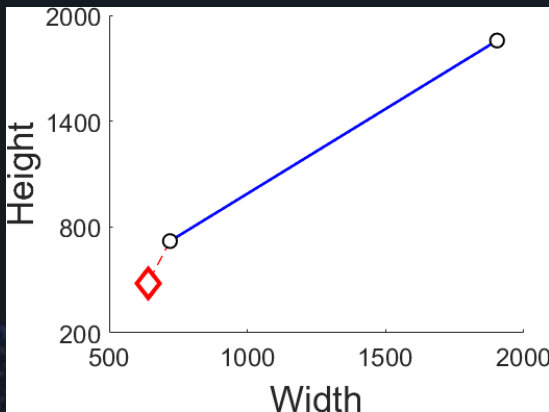
# 尺寸适配

- 多种插值情况全覆盖
- 三角形插值



# 尺寸适配

- 多种插值情况全覆盖
- 点插值：“等比例”缩放



$$S(\Theta', D') = \beta S(\Theta, D)$$

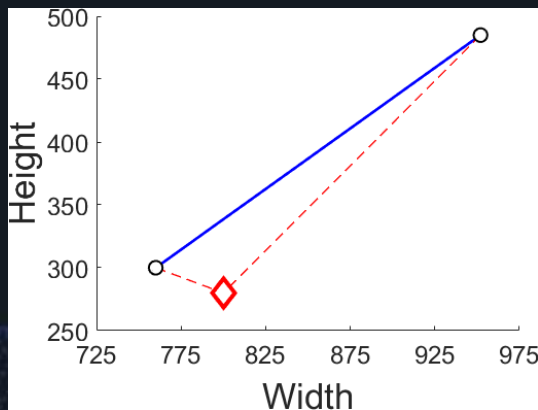
$$C(\Theta', D')/D' = C(\Theta, D)/D$$





# 尺寸适配

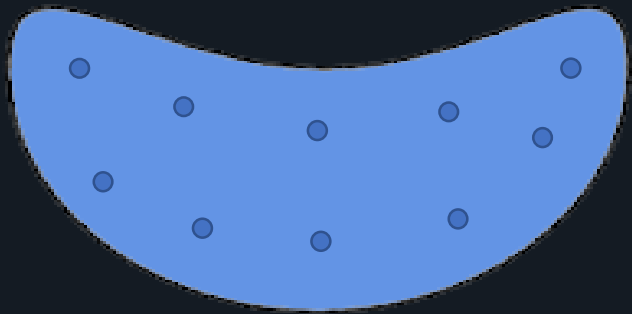
- 多种插值情况全覆盖
- 线插值：点插值加权组合





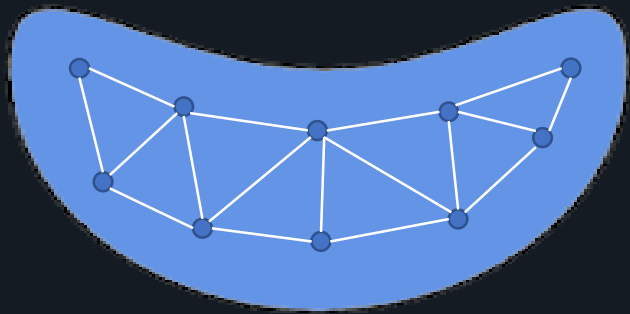
# 风格适配

- 基于流型分析的风格迁移



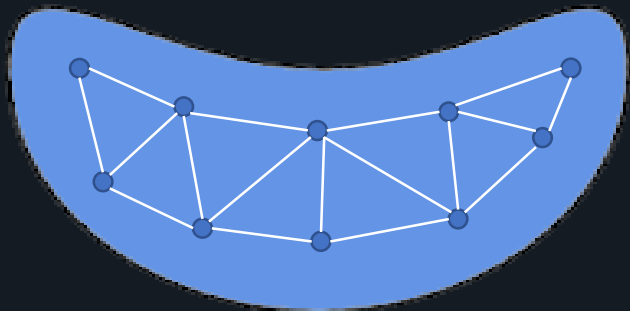
# 风格适配

- 基于流型分析的风格迁移

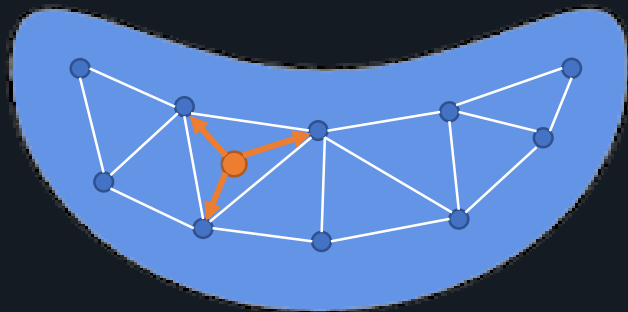


# 风格适配

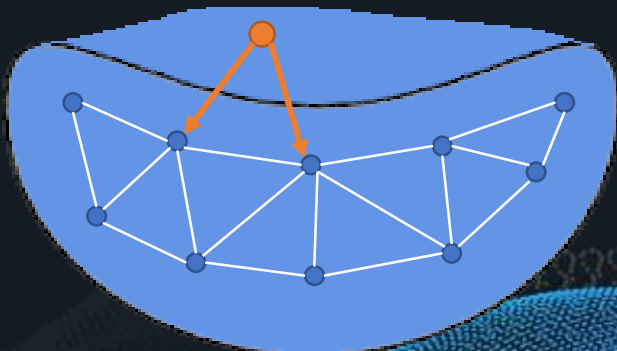
- 基于流型分析的风格迁移



内插式迁移

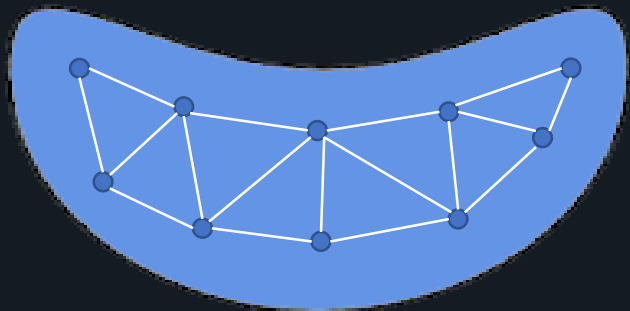


外插式迁移

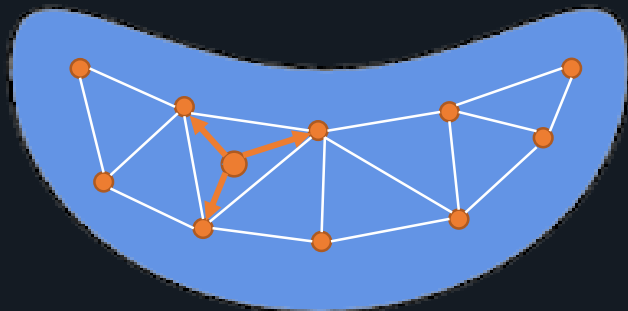


# 风格适配

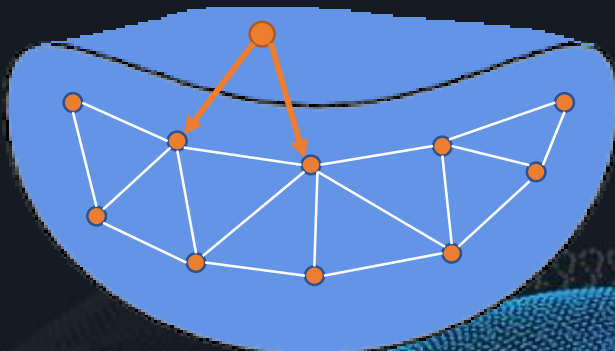
- 基于流型分析的风格迁移



内插式迁移

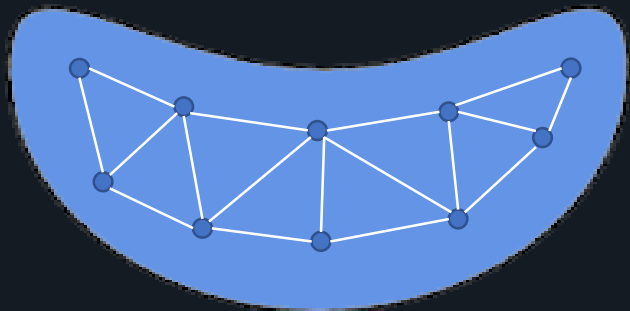


外插式迁移

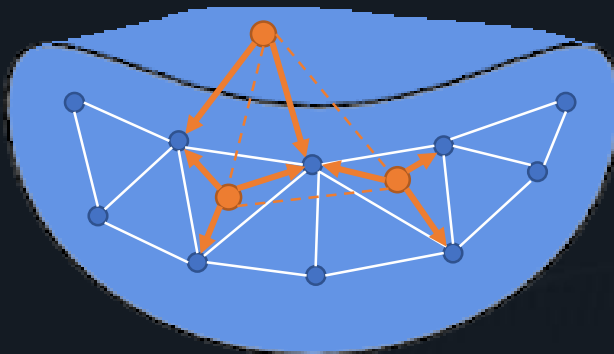


# 风格适配

- 基于流型分析的风格迁移



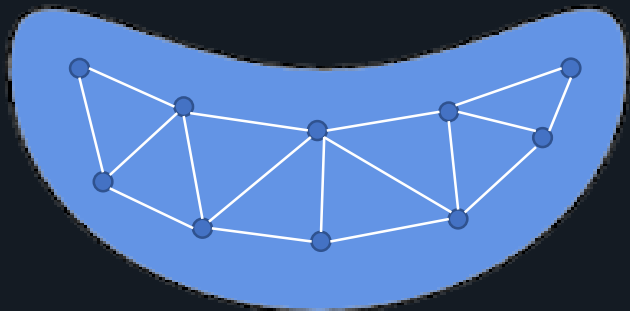
集合迁移



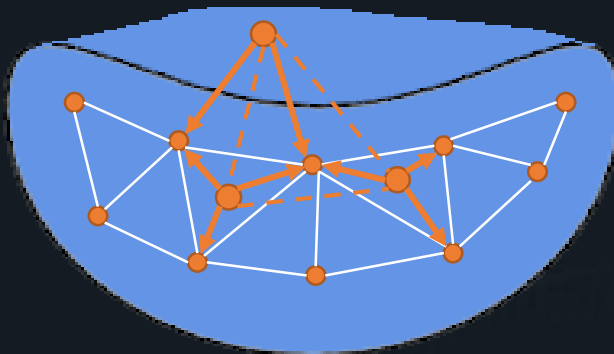


# 风格适配

- 基于流型分析的风格迁移

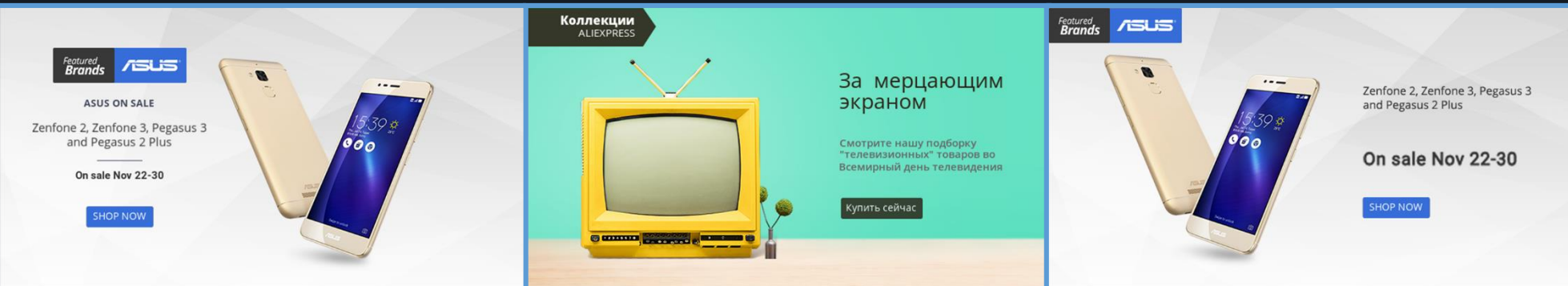


集合迁移



# 风格适配

- 基于流型分析的风格迁移



## 下一步工作

- 布局风格推荐
  - 依据布局草图
  - 依据商品特点
  - 依据业务类型
  - 依据人群偏好



01 稳重大方：平衡感强、庄重大方

02 活泼动感：错落张力、自由灵动

03 简约现代：左右分离、清新利落



# 目录

“阿里云眼” 总概

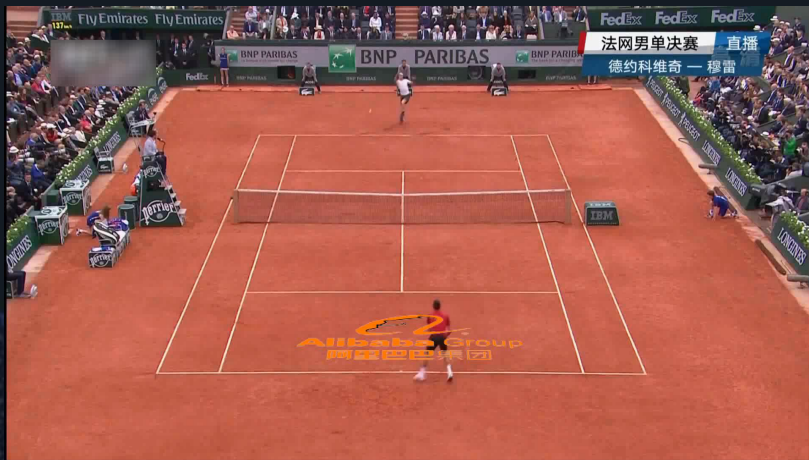
智能设计介绍

- 横幅布局设计

视频植入介绍

# 视频植入介绍

- 在视频中嵌入具有现场感的内容
- 嵌入内容包括：Logo、海报、商品、动画

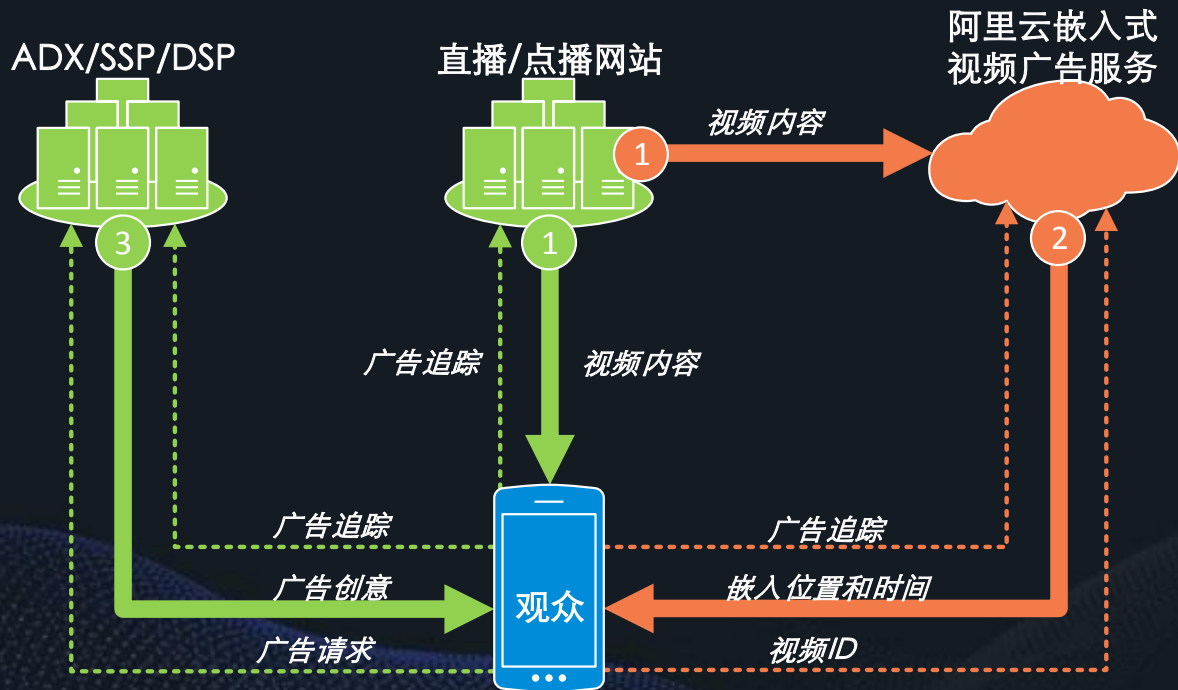




# 视频植入技术大图



# 智能视频广告植入系统



## 阿里云提供

- 视频广告嵌入服务
- 客户端技术示范
- 兼容主流DSP

# 三类植入技术



- 互动场景
  - 主播视频
  - 电视节目
  - 通过人机交互指定植入位置



- 特定场景
  - 体育比赛
  - 地标视频
  - 通过场地检测与跟踪获取植入位置



- 一般场景
  - 一般生活视频
  - 一般影视节目
  - 检测电视、墙壁、橱窗、门面、地面等位置作为植入点

低

难度

高

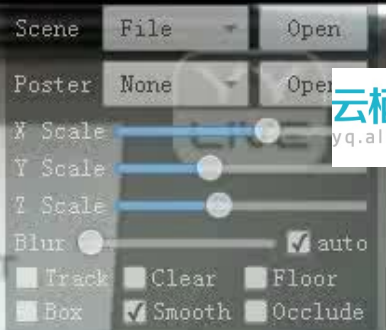
FPS:9

# 主播交互植入





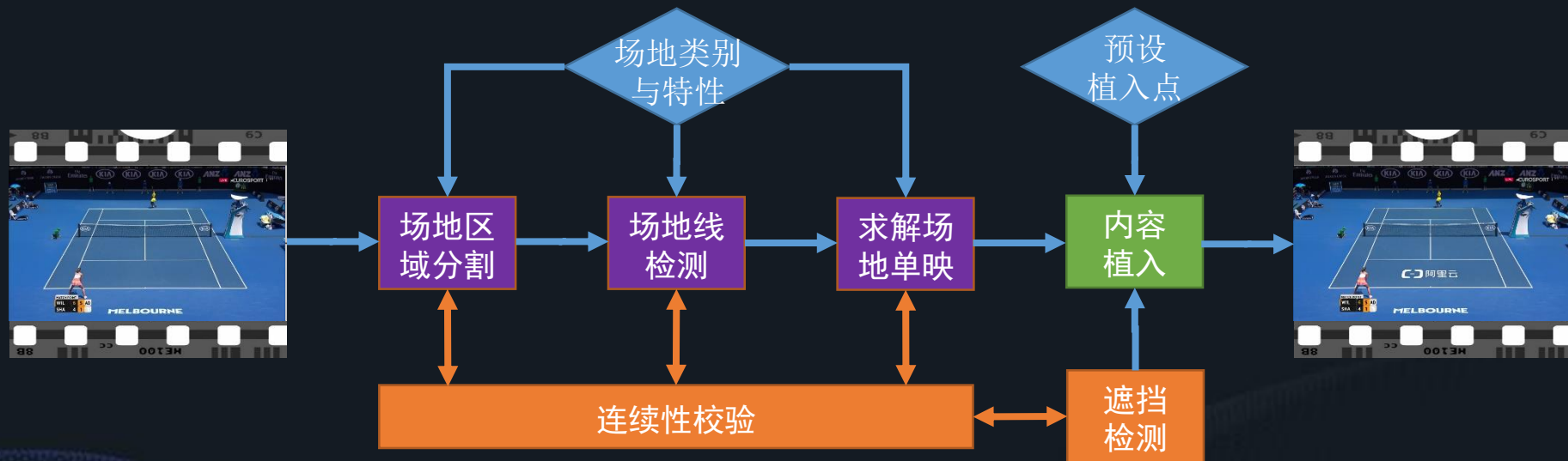
# 自动遮挡检测



云栖社区  
yq.aliyun.com



# 场地检测与跟踪



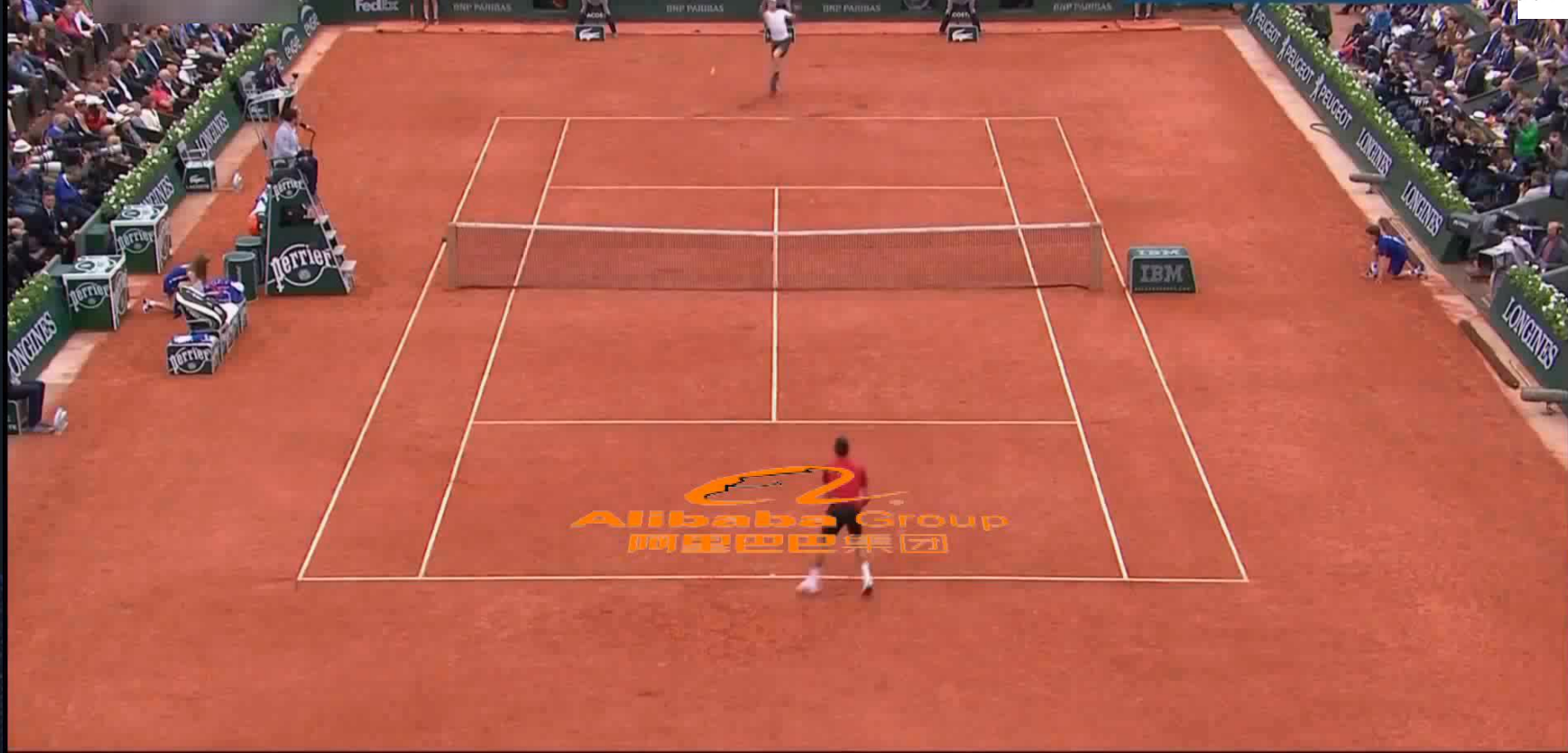


IBM 137m Fly Emirates Fly Emirates

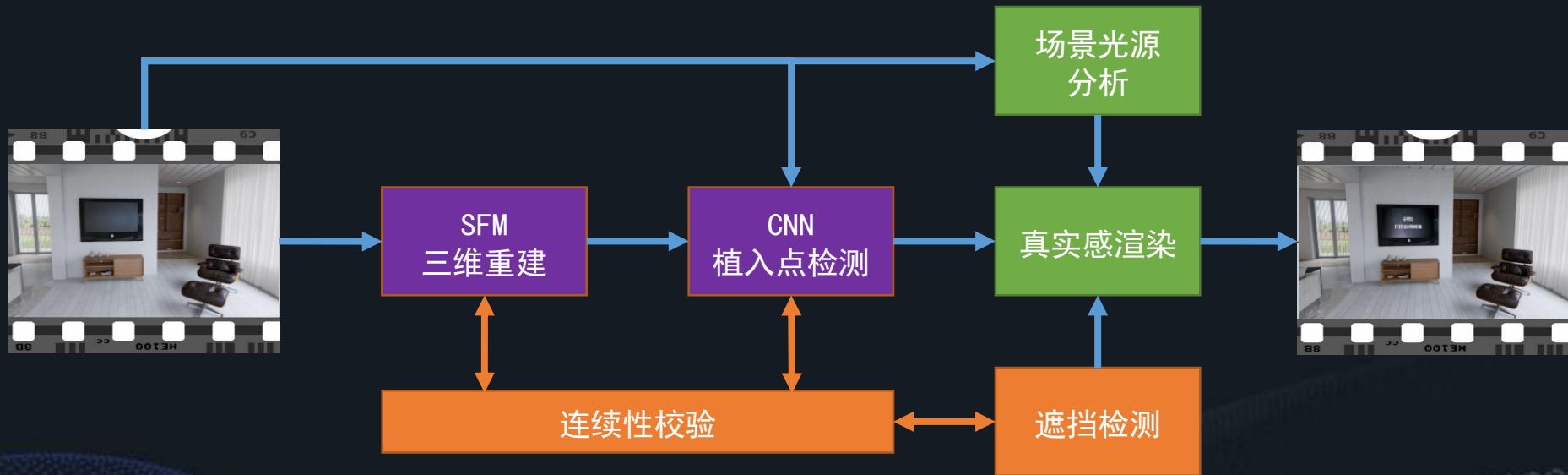
FedEx FedEx FedEx

法网男单决赛 直播  
德约科维奇 — 穆雷

云栖社区  
yq.aliyun.com



# 一般场景植入位置检测





# 一般场景植入



# 视频植入与增强现实的异同

## 相同点

在背景视频中添  
加更多视觉内容

采用相似的三维  
重建与渲染技术

## 视频植入

- 目标是云端服务
- 侧重植入效果
- 植入点未知
- 不依赖特殊硬件

## 不同点

## 增强现实

- 目标是终端应用
- 侧重响应速度
- 植入点通常已知
- 依赖深度传感器  
及立体显示设备



# 下一步工作

- 算法持续改进
  - 提高场地检测、遮挡检测、三维重建等核心算法的稳定性
  - 开发基于深度学习的植入点推荐算法
- 建设视频植入云平台
  - 利用阿里云存储（OSS）、云计算（ODPS）
  - 联合广告投放平台（DSP/SSP/ADX）开发完整的视频广告服务解决方案

# 飞天·智能

## APSARA INTELLIGENCE

2017云栖大会·成都峰会

5月23日 成都世纪城天堂洲际大酒店