



Alibaba A.I. Labs
阿里巴巴人工智能实验室

AR开发者平台体系及教育领域的探索

阿里巴巴高级技术专家 古鉴（坎特）

目录

- 1 AR 市场及阿里AR开放平台介绍
- 2 AR教育产品
- 3 研发中技术难点
- 4 未来的研发方向

个人介绍

11年AR技术的研究和创业的经验

5年AR的学术研究经验

- 新西兰HIT Lab NZ 实验室
- 新加坡国立大学

6年AR工业领域经验

- 梦想人软件科技有限公司, CTO, 联合创始人
- 暴风魔镜, AR及交互技术负责人-主要负责交互产品
- 阿里巴巴人工智能实验室, AR开放平台负责人



AR市场的分析



AR 硬件演进方案

- 手机AR 是现在AR的主流平台
- 可以看到定制AR硬件会是未来一段时间的主流
- AR 眼镜会是今后的最终形态.

Mobile AR



AR Customized Equipment



Smart Glasses



AR的价值

● AR的价值

- 视觉产生的震撼
- 自然交互方式产生便捷
- 虚拟和现实实时融合产生高附加价值

● AR的问题

- 震撼的效果要求3D的制作成本比较高
- AR算法有一定的技术门槛
- 定制化多，产生的价值无法被大量复制。
- AR只是附加值，很难成为主要的价值提供者

● AR平台能产生的价值

- 降低AR的技术门槛：提供AR SDK
- 降低AR的制作成本： 标准化3D制作流程
- 更便捷的生成AR内容： 内容编辑器

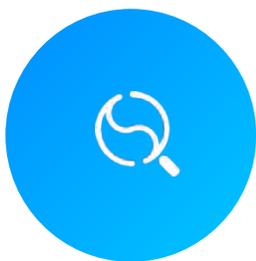


Ali Genie AR 开发者平台

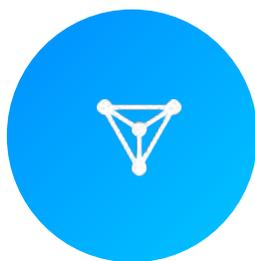
- 2017年10月，我们发布了Ali Genie AR 开发者平台



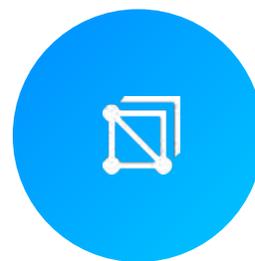
Ali Genie AR 开发者平台



2D识别追踪



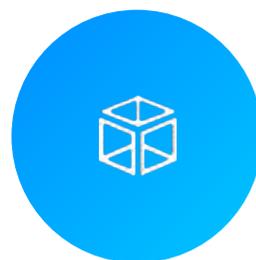
3D识别追踪



内容制作平台



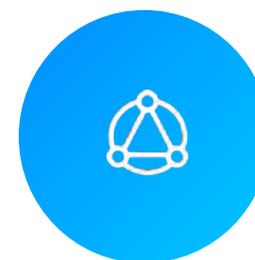
标记物训练工具



Aviex渲染引擎



稳定的云服务

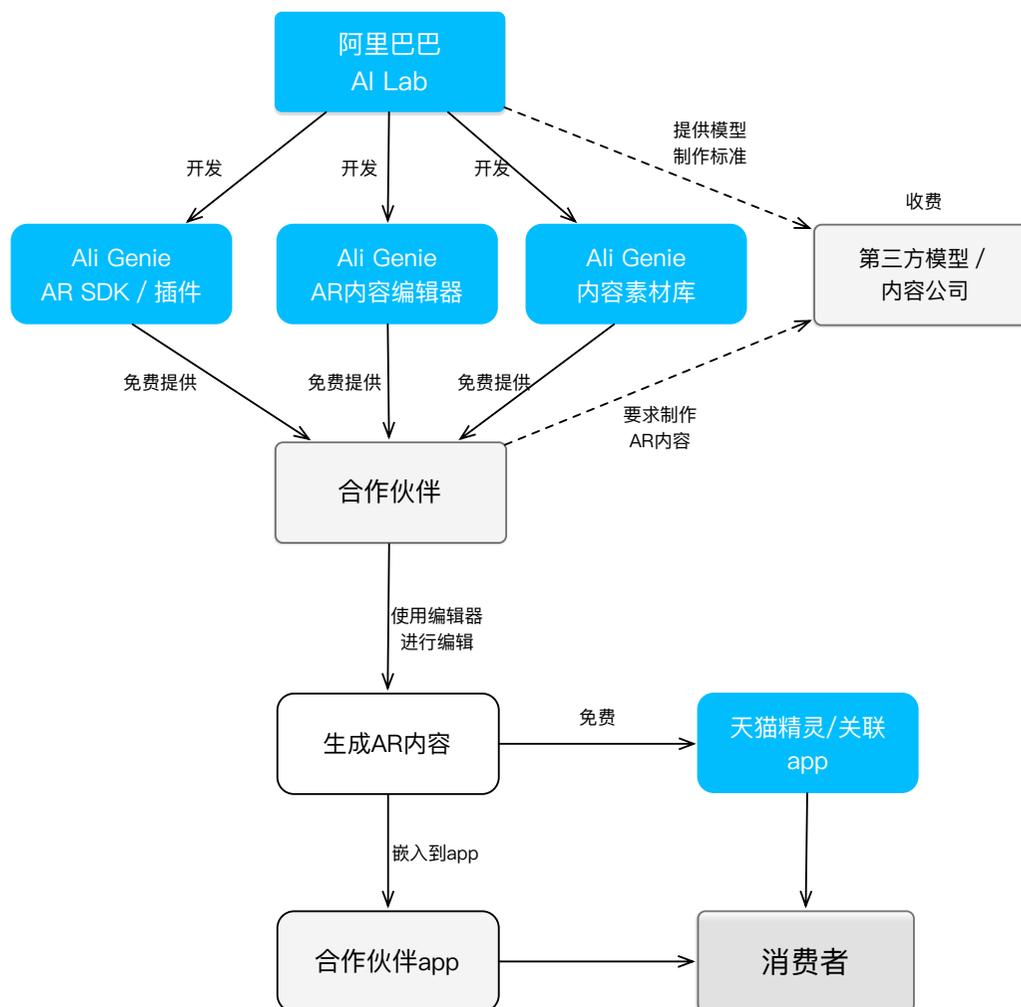


高品质3D素材库



GPU加速

内容合作模式

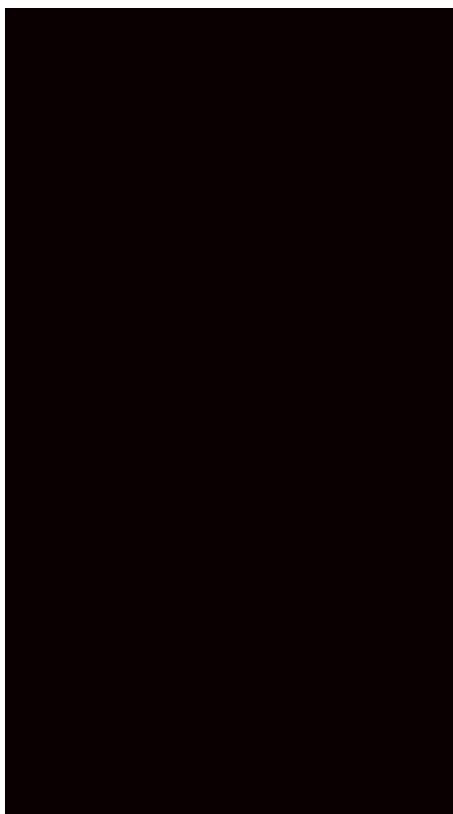


开放平台案例-阿里健康



开放平台案例

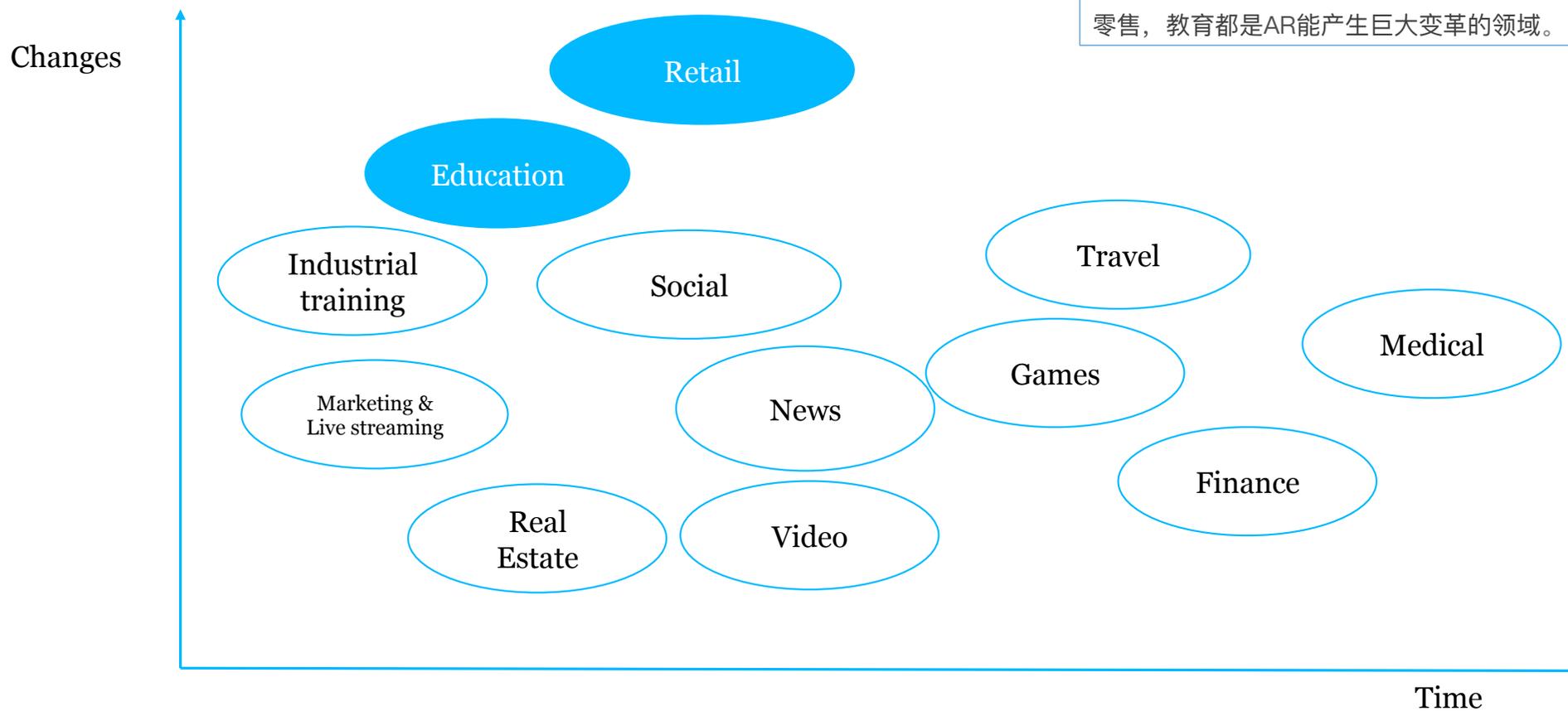
星巴克



阿里汽车

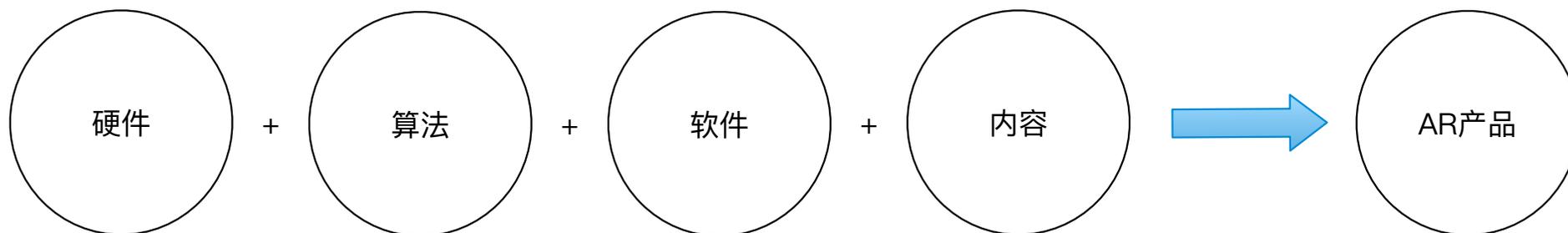


AR对各个领域的影响



思考AR开放平台如何产生更多的价值

- 选择AR适合的领域：大规模使用，能够快速制作，同时能够快速复制行业
- 深入了解某一领域的需求。
- 要以**硬件/实体物品**为**AR**载体，同时让**算法+软件+内容**对用户产生**高附加价值**。
- **零售和教育**是我们看到的AR能产生高附加价值的领域



选择儿童教育作为AR产品切入点

儿童教育市场中：早教市场前景广阔，估算目前规模1000亿元左右。目前购买天猫精灵的顾客中，大约 60% 左右已育人群，儿童用户是天猫精灵重要用户。

AR的强烈的和现实的结合性，比如说和实物的结合（书，积木等），导致它天生会和教育场景结合密切。

AR在教育领域的应用

Project name

AR 放大镜



AR 地球仪



AR 学习课件



AR 涂色



识花



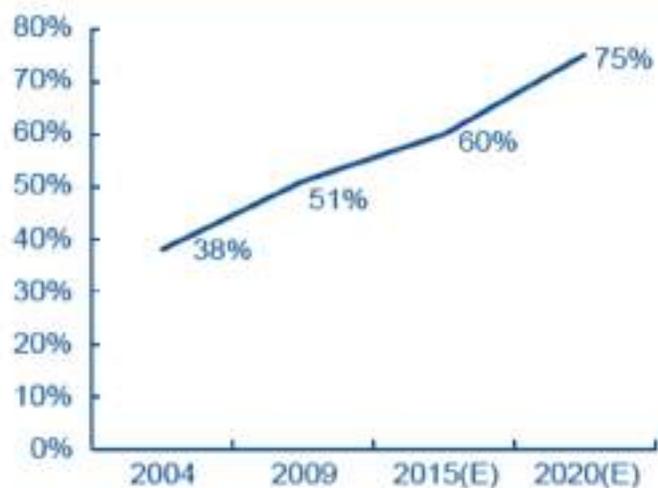
天猫书城



对教育的投入

教育对于儿童成长和未来终身发展有着重要的意义，教育支出主要集中于有孩子的家庭，有孩子的家庭教育支出显著高于没有孩子的家庭。

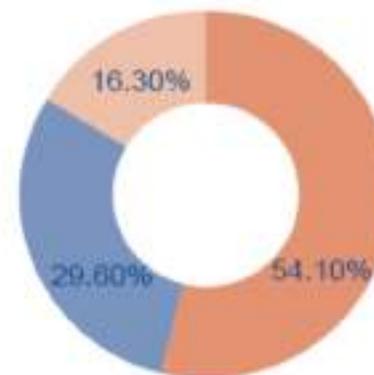
图 32: 2004-2020E 中国学前三年教育普及率



资料来源: 德勤, 中信证券研究部

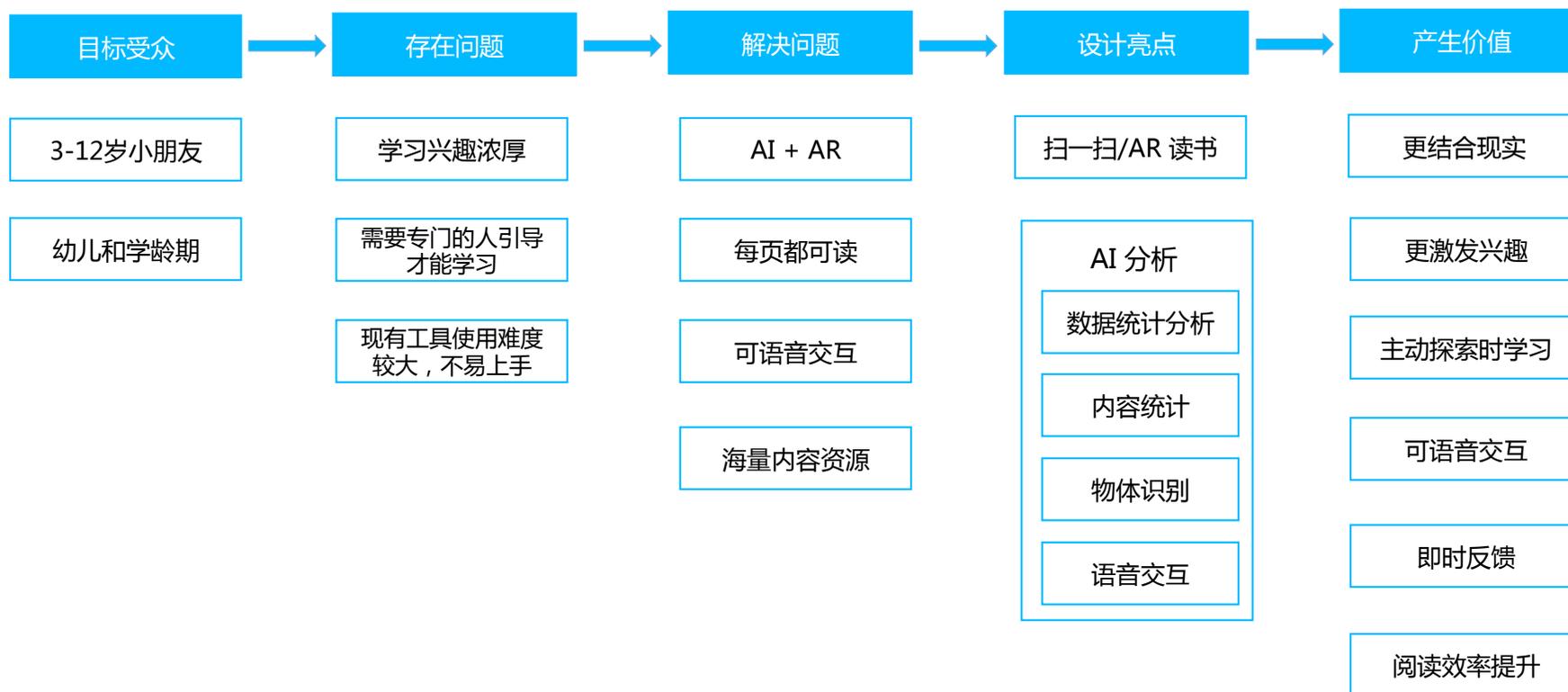
图 33: 中国家庭早教月花费分布

■ 小于1000元 ■ 1000-2000元 ■ 2000元以上



资料来源: 《2012 搜狐教育行业白皮书》, 中信证券研究部

设计方向 (AR+AI)



教育产品的探索

国图字帖

和国家图书馆合作的字帖，教小朋友学习写字。



天猫图书

和天猫图书合作，通过动态恐龙的方式展现图书

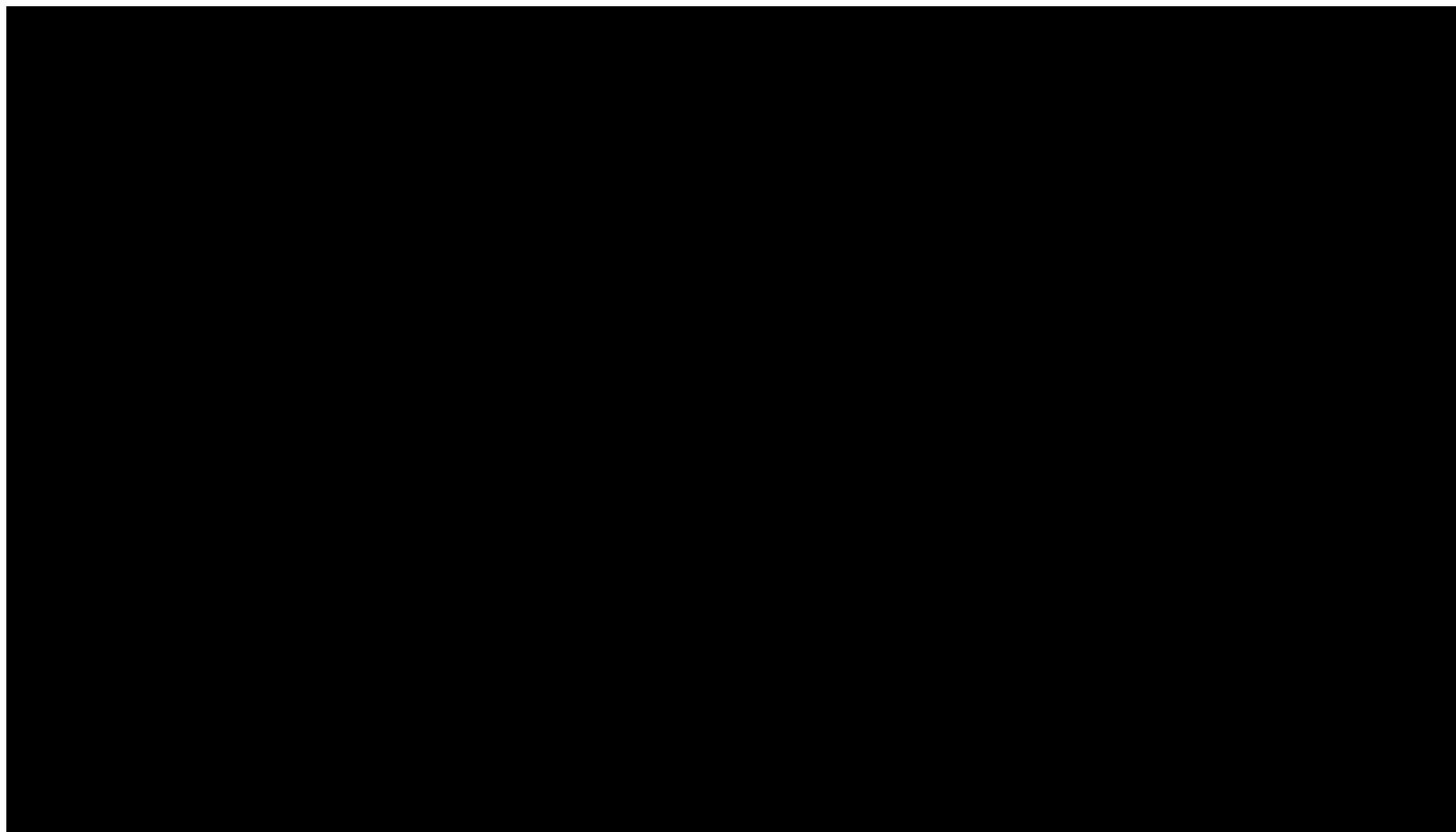


教育产品的探索

- 图像识别
- 物体检测
- 拟态引擎
- 多模态交互
- 情景感知



天猫精灵火眼+XHolder



用户反馈

很有魔性，小孩很喜欢玩

小孩跟着天猫精灵火眼读书超过一个小时

睡前喜欢听故事，不用我在读了

奶奶爷爷发音不标准，这个可以更好教他们英文

父母不在的时候也可以玩识字卡片



用户反馈



技术上的探讨

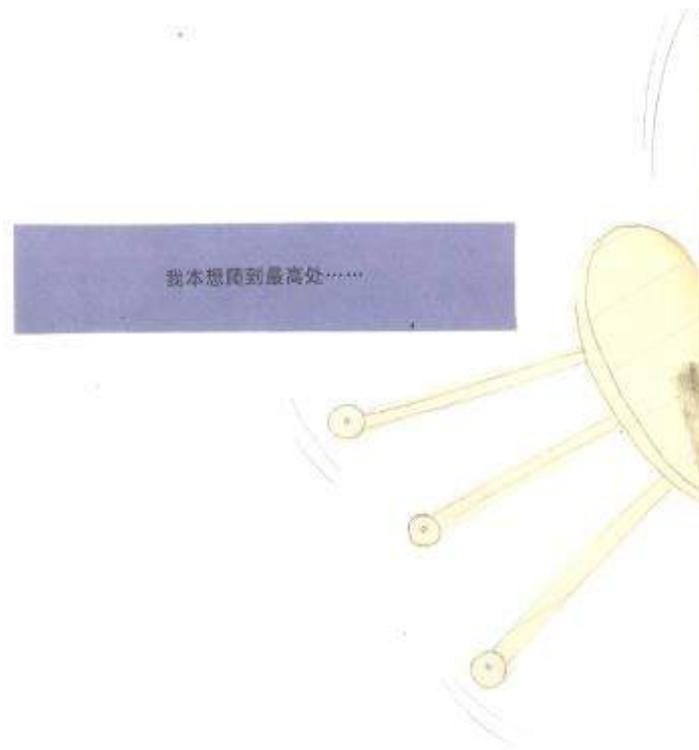
- 天猫精灵火眼产品的架构
- 声音和视觉结果的同步
- 识别追踪算法逻辑图
- 识别中遇到的一些难点
- 算法上一些新的探索
- 内容编辑器
- 3D制作标准

整体设计架构图

通过相机、渲染、算法的分离，并行处理，多重缓冲区，实现实时

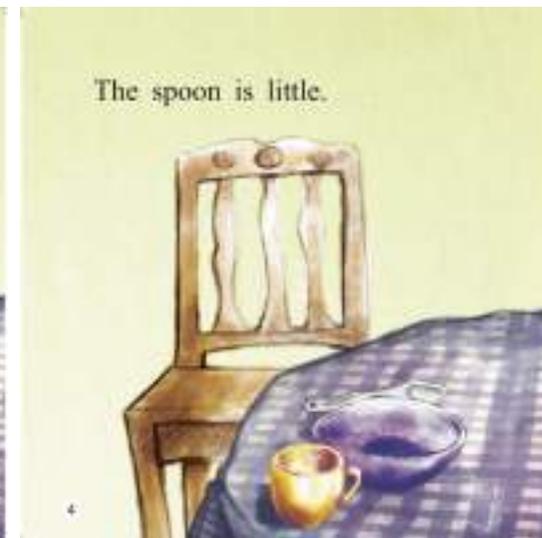
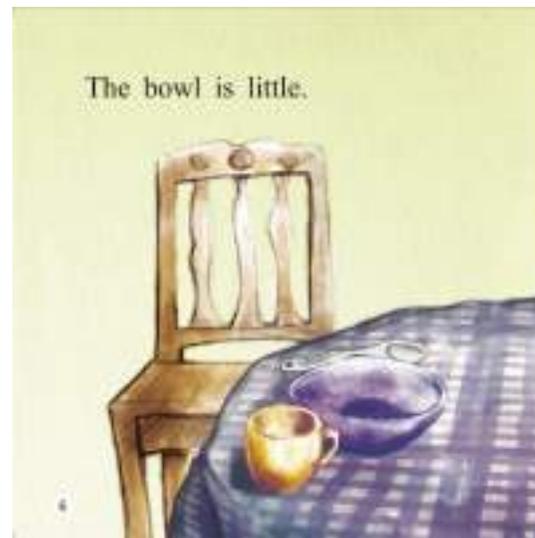


识别的问题



基于深度学习的OCR技术用于文本图片检索

- 同时利用visual feature + textual feature 提高图片识别的准确率
- 基于深度学习的OCR技术进行图片的文本特征提取，主要分为四步：
 - 第一步：文本检测。
 - 第二步：文本字符识别。
 - 第三步：文本特征提取。
 - 第四步：文字统计



后台编辑器



AR开发者平台3D模型输出规范

Model of the standard



模型参数

- 面数:2.5W-3W三角面 (Triangle)
- 要求：模型结构清晰，模型精致
- 模型尺寸缩放到30cm的立方体之内



贴图规范

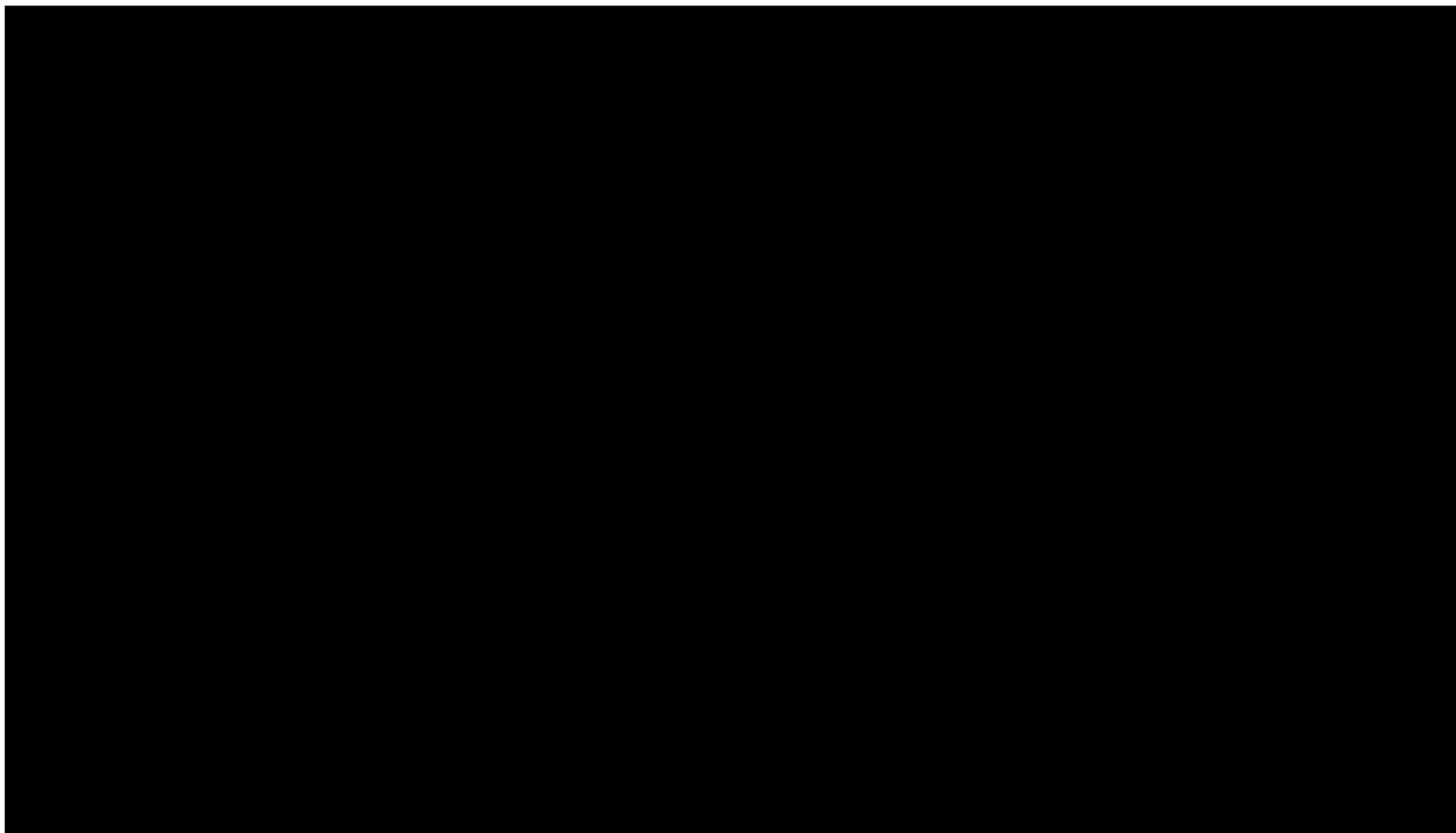
- 纹理贴图 1024*1024 PNG
- 法线贴图 1024*1024 PNG
- 材质贴图 1024*1024 PNG
 - ◆ R：金属
 - ◆ G：皮革
 - ◆ B：棉布等亚光材质



动画规范

- 世界中心点创建Dummy
- 命名Dummy为：RootBone
- 以RootBone为根节点的骨骼树
- 骨骼树下面不能LINK任何模型
- 动画必须要通过SKIN绑定实现
- 骨骼总数不能超过68根
- 每个顶点最多受4根骨骼权重影响

未来的展望



视频来自 HIT Lab NZ

未来的计划

- AR开发者平台更多的和不同领域进行深度合作。
- 教育平台会结合更多的内容。更多的教育3D课件
- 兼容不同的引擎和3D模型，动画，及交互是我们的目标。
- 语音+视觉+内容可以输出到不同的儿童教育硬件产品上。



Alibaba A.I. Labs
阿里巴巴人工智能实验室

Thanks