

# 从2D到3D，AR发展中的关键技术

侯晓辉 / 亮风台

# 目录

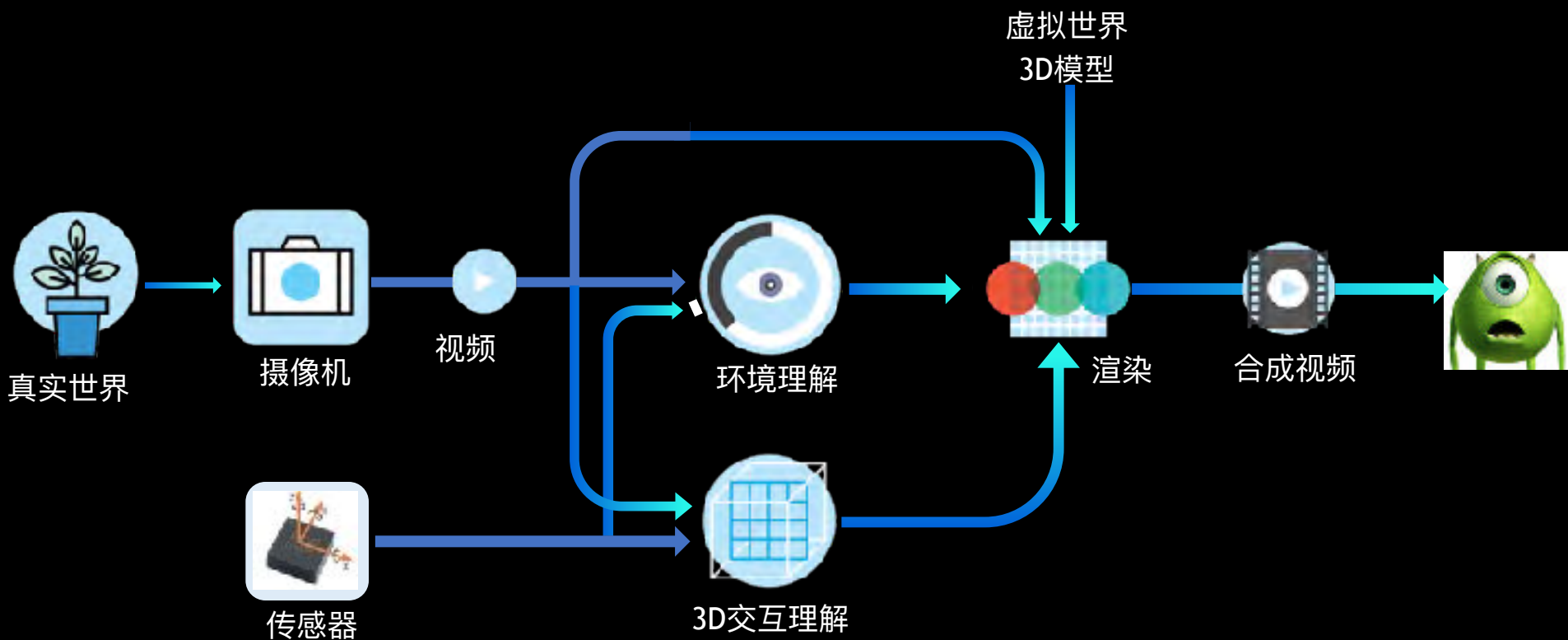
- AR简介
- AR原理
- AR现状
- AR 3D中的关键技术

# AR简介

优酷



# AR原理—多模态增强现实流程



# AR现状



**Pokémon GO**  
任天堂+Niantic  
下载量**逾亿**  
LBS+IMU+Camera  
3D场景定位, 重内容



北京

雅典

**AR奥运火炬传递**  
手机QQ, 亮风台SDK  
参与用户**逾亿**  
Camera  
2D跟踪识别, 高精度

# AR现状 (二)



# 从2D到3D

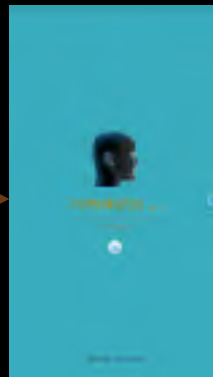


平面识别跟踪

2D

3D

3D



◀ 非平面  
物体识别

▼ 3D环境实时感知



# 非平面物体识别跟踪



跨视角识别跟踪  
处理非刚性形变

应用领域拓展

打破平面的限制  
规避高能耗计算

高效的深度学习算法  
手机端实时运行

旅游、展览、  
教育、娱乐...



# 几何理解： 三维场景感知

## Simultaneous Localization And Mapping (SLAM, 即时定位与建图)

- 机器人： 在一个未知的环境中实时的构建和更新地图，同时对物体自身的进行定位和跟踪。
- 常用技术： particle filter, Kalman filter, etc.
- 计算机视觉： 基于视觉传感器或以视觉为主的定位和三维重建技术
- 常用技术： sequential Structure from Motion (SfM), Bundle Adjustment
- 增强现实： 视觉传感器为主， visual SLAM (vSLAM)



# SLAM是什么

- 在一个陌生的环境里。。。

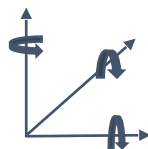
## 我在哪



我：相机(传感器)



哪：制图 (mapping)



在：定位  
(localization)

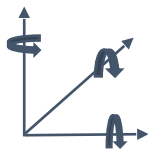
# SLAM是什么

- 在一个陌生的环境里。。。

我在哪



我：相机(传感器)



在：定位  
(localization)



哪：制图 (mapping)



# SLAM是什么

- 在一个陌生的环境里。。。

我在哪

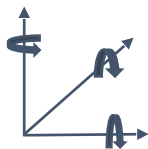


我：相机(传感器)

哪：制图 (mapping)

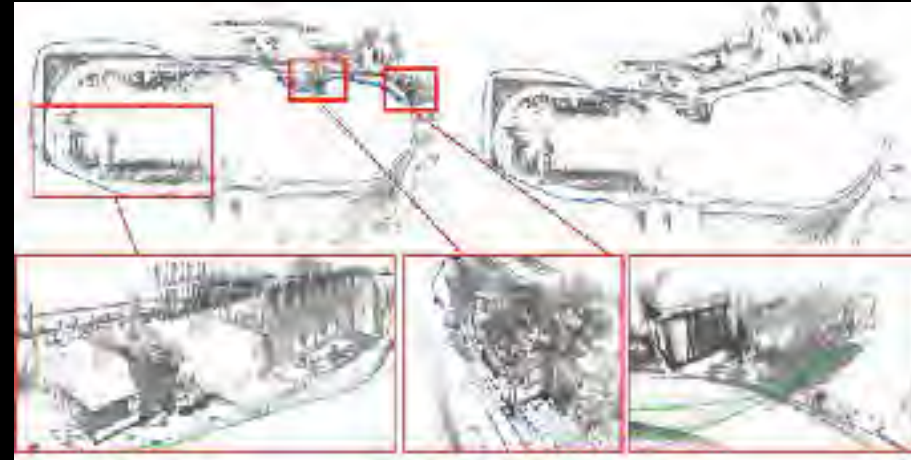
simultaneous

在：定位  
(localization)

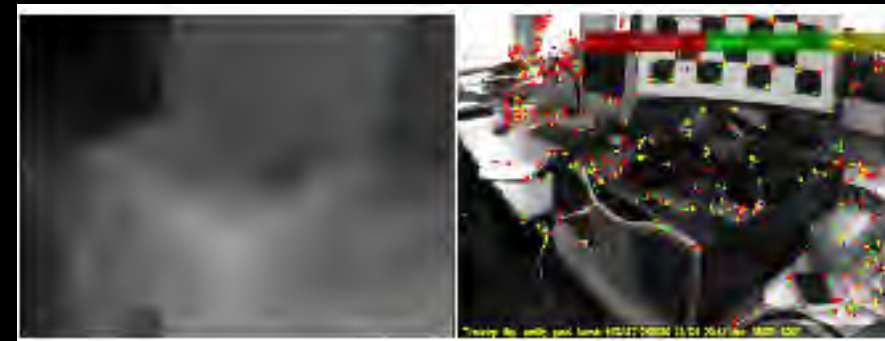


# Visual SLAM

- 由于计算资源的限制，AR中主要使用的是Visual SLAM
- 从三维场景建模的密度可以分为
  - 稀疏SLAM：PTAM，ORB-SLAM，等等
  - 半稠密：LSD-SLAM，等等
  - 稠密：DTAM，等等
- 目前AR中使用最多的是稀疏SLAM，尤其是PTAM的变种



LSD-SLAM, Engel, Schöps & Cremers, ECCV'14



DTAM (left) and PTAM (right), Newcombe, Lovegrove & Davison, ICCV'11

# 增强现实中的SLAM

- 低端移动设备，弱传感器，视觉传感器为主
- 实时要求，计算能力弱，模型大小限制
- 视频质量问题：抖动，模糊



# 算法优化

- 估计运算规模
  - 减小运算量：特征点，3D点，关键帧
  - 减小CPU占用率
  - 减小内存消耗
- 配置参数
- 好的实现

# SLAM算法规模估计

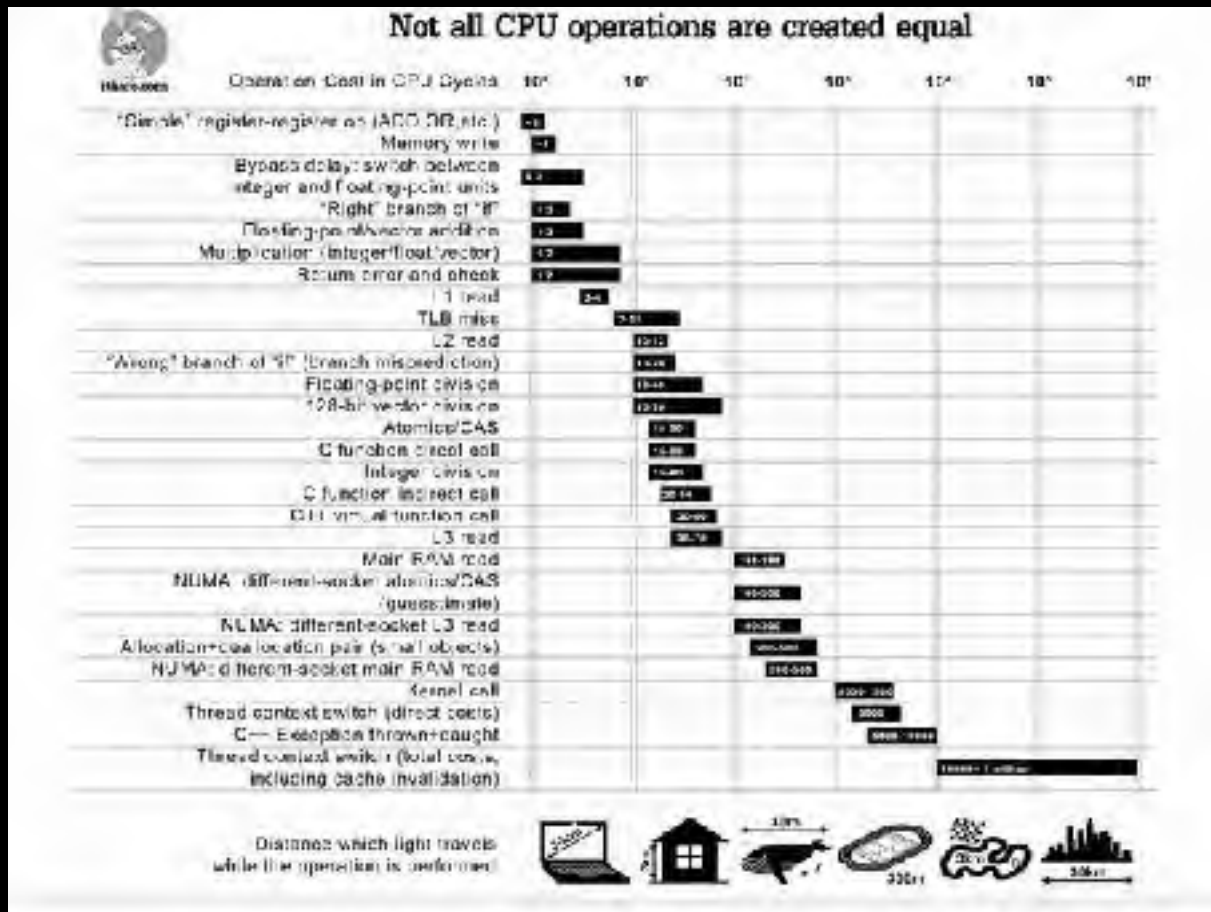
- 输入图像：  $640*480=307,200$
- 提点： 6000
- 均匀化： 600
- 匀速模型估计： 10
- 位姿优化： 20



# SLAM算法规模估计 (二)

- Local Bundle Adjustment
  - 帧数
  - 3D点数

# 工程优化



# 工程优化 (二)

- 优化线程模型
  - TBB
  - OpenMP
  - Lockless programming

# 工程优化 (三)

- 使用NEON指令
  - 降低功耗
  - 提升性能
  - 128位SIMD (Single Instruction, Multiple Data)

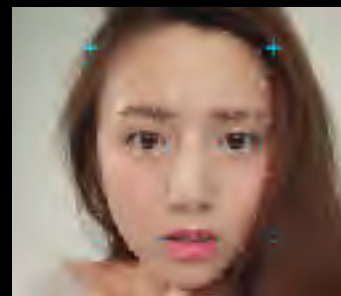
# 工程优化（四）

- 使用异构加速
  - OpenCL (Open Computing Language)
  - 第一个面向异构系统通用目的并行编程的开放式、免费标准，也是一个统一的编程环境
  - CPU, GPU, DSP
  - OpenCL 2.0
  - Tune OpenCL kernel 函数

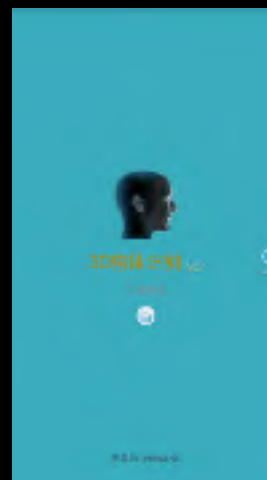
# 三维重建

- SfM
- MVS (Multi View Stereo)
- Meshing
  
- 需要稠密点云
- 目前移动端无法达到实时

# 从几何到语义

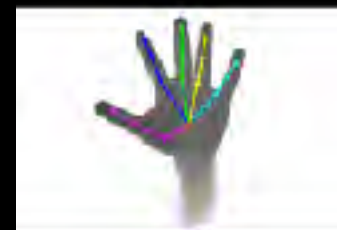


◀ 人脸  
检测定位



◀ 非平面  
物体识别

▼ 手势识别理解

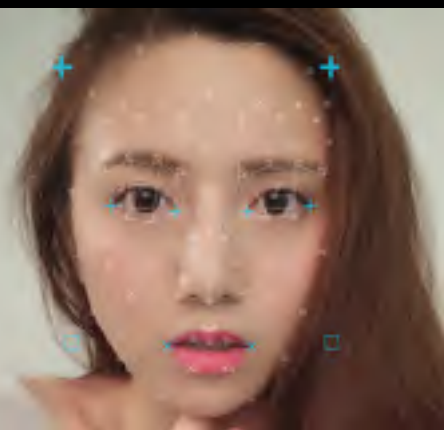


平面识别跟踪

语义

几何

语义



# meitu

## M-Face

为亚洲人研发的人脸识别技术



171个  
定位点

0.2秒  
快速识别



# 总结

- 增强现实兴起是软件、硬件和应用三方面的共同促进的
- 核心技术主要体现在智能交互理解，高效虚实融合渲染，以及高效的内容制作三个方面
- 智能理解中：跟踪和配准技术为当前核心热点，对场景的实时理解跟踪是增强现实当前的关键技术
- 对语义的理解是AR中未来的一个重点

谢谢!