



视觉定位与感知 (Visual Localization & Perception)

2018年秋季电子系大三选修

邹丹平 副教授 导航与感知所
电子信息与电气工程学院



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



自我介绍

- 邹丹平
 - 信息技术与电气研究院 (感知与导航研究所)
 - 微电子楼309
 - Email:dpzou@sjtu.edu.cn
- 主要研究方向：
 - 图像处理和三维重构方法
 - 计算机视觉定位技术
 - 无人机视觉导航



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

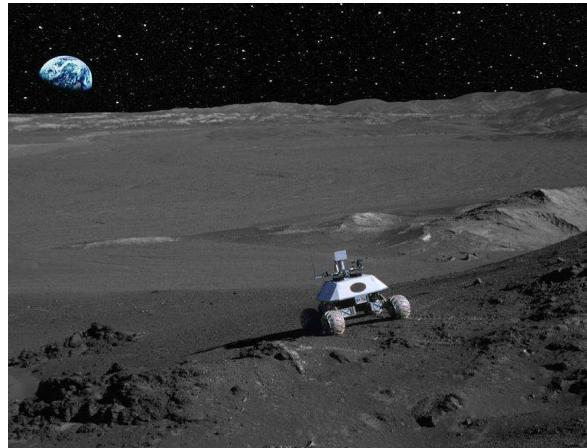
什么是“视觉定位与感知”



自主移动机器人



- **自主移动**-无人系统/机器人接收到目标地点指令后的**自主导航能力**。
- 所谓“**自主**”一般指无人工干涉情况下，仅依靠
 - 自身携带的传感器
 - 自身携带的计算资源
- 实现上述目标。





导航的基本要素



- 定位 (Localization) – 获得精确位置与姿态
 - 室外 - GPS/Beidou, GLONSS, Galileo (四大卫星定位系统)
 - 室内 – 动作捕获系统 (Vicon, Optitrack)
- 地图构建 (Mapping) – 获得周围环境的几何模型
 - 测绘 – 事先定点测量、拼接
- 路径规划 (Planning) – 在获得定位结果和环境几何模型后，通过最优搜索方法，找出行进路径
- 定位和地图是实现导航的基本前提



自主导航的核心问题：

- 在陌生环境下，在无外界辅助定位系统，也没有实现测绘好的地图，如何实现自我定位与地图构建？
- 仅仅通过自身携带的传感器：



视频摄像头 激光雷达



深度摄像头 超声测距传感器

+



受限计算资源（小型无人系统）



视觉定位与感知



- 视觉 - 以摄像头作为主要传感器，运行计算机视觉算法
- 定位 - 获取传感器自身的位置、姿态
- 感知 - 环境感知：场景地图构建、三维重建、场景理解与分割



通过视觉信息处理，解决自主定位和环境感知的问题



视觉定位与感知 – 主要技术点



- 定位:

- 光流 (optical flow)
- 位姿估算 (Pose estimation)
- 基于图像匹配的视觉定位

- 环境感知:

- 三角化 (Triangulation)
- 双目深度匹配 (Binocular stereo)
- 基于图像的多视角3D重构 (Multi-view stereo)



- 视觉SLAM



应用场景

- 基于视觉定位与环境感知方法已在无人机/移动机器人成熟应用，并开始自动驾驶中启动



双目摄像头



背部鱼眼摄像头

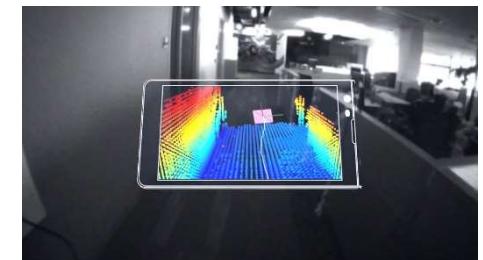


双目摄像头



应用场景

- 增强现实(AR)产品 (微软hololens， 谷歌Project Tango手机)





应用场景

- 苹果和谷歌在各自手机操作系统(iOS, Android)加入视觉定位功能

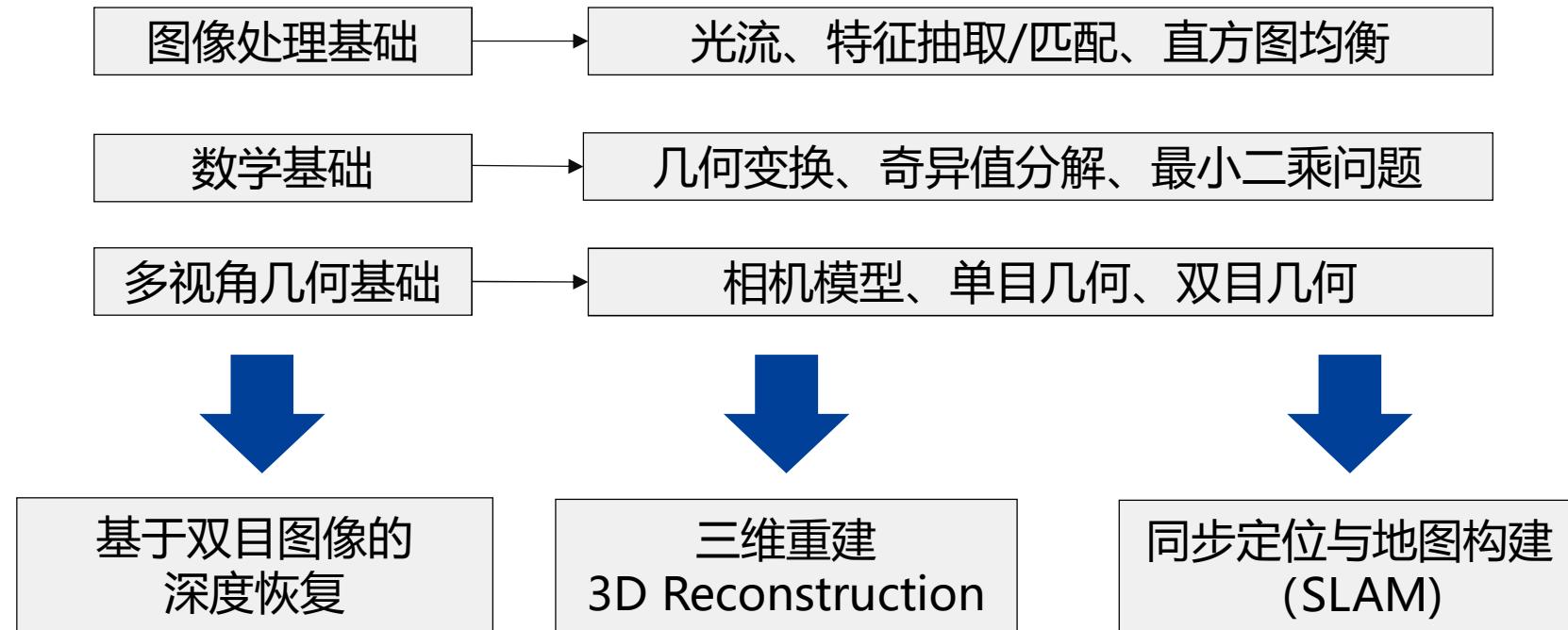


为了抢占移动端增强现实市场





课程内容安排

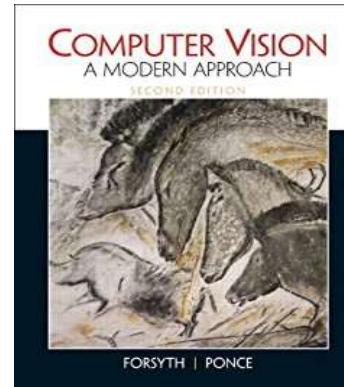
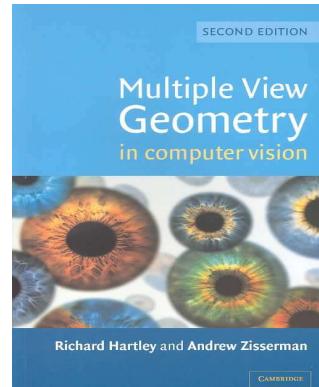




参考书目



- Richard Hartley & Andrew Zisserman, *Multiple View Geometry in Computer Vision*
- David Forsyth, Jean Ponce, *Computer Vision – A Modern Approach*



- 其他：
 - Richard Szeliski, *Computer Vision – Algorithms and Applications*
 - 高翔博士：《视觉SLAM十四讲》



课程成绩

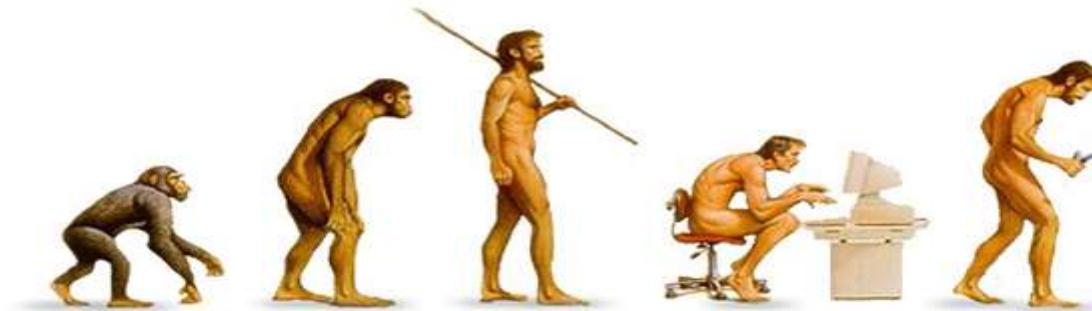


- 平时作业: 60% (5次)
- 期末大作业: 40%

- 助教: 胡瑜、唐玮璇
- 办公室: 微电子楼309



视觉定位与感知技术快速导览





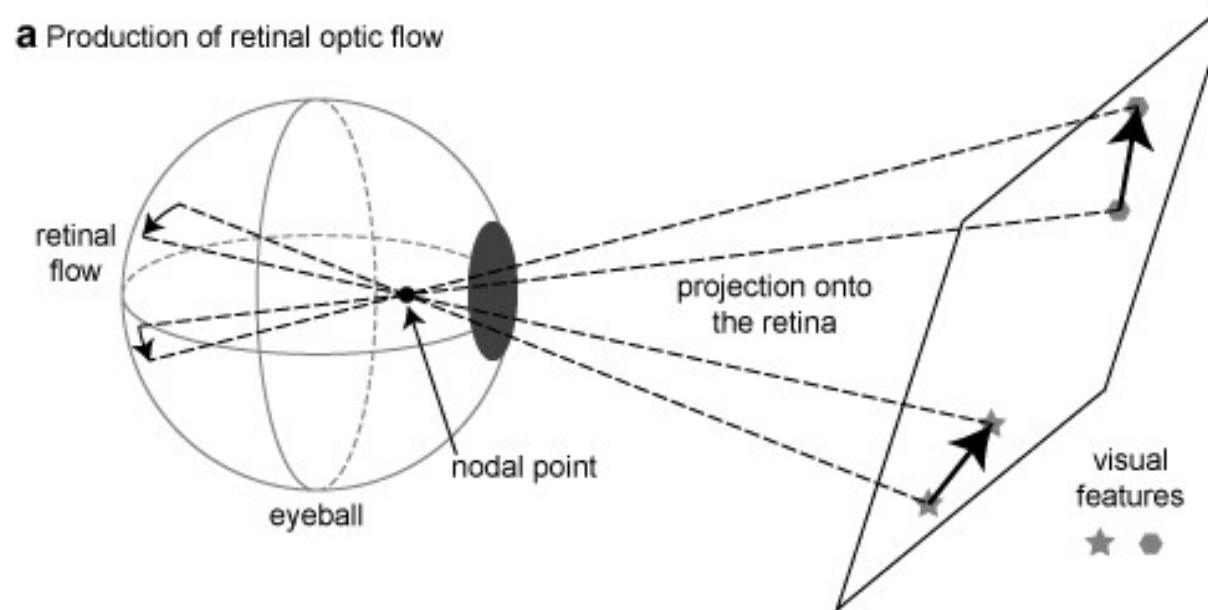
上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

光流 (Optical Flow)



光流

- 光流是观察者感知到的一种场景相对运动图案



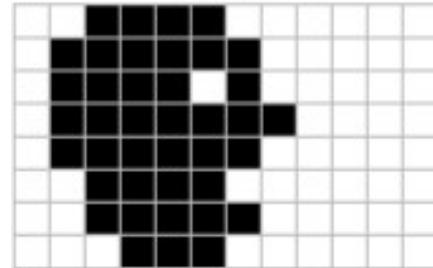


光流

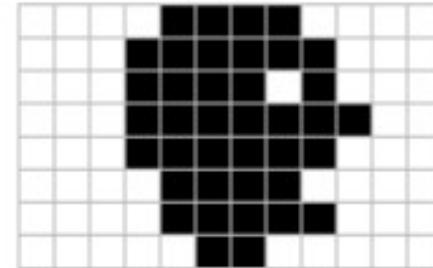
- 物体移动所产生的光流示意图

b Detection of optic flow in the image plane

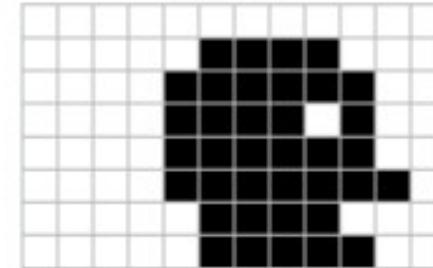
frame 1



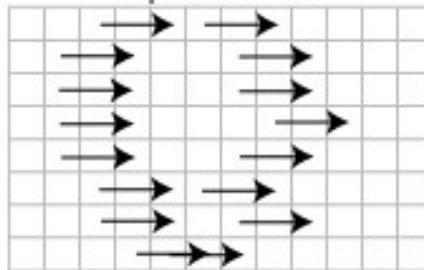
frame 2



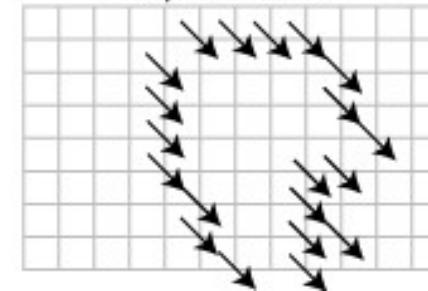
frame 3



optic flow 1-2



optic flow 2-3





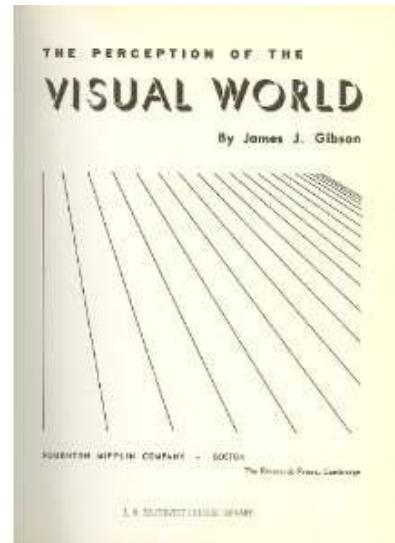
光流



- Optical flow is firstly introduced by James J. Gibson in 1940s



James J. Gibson
(1904-1979)



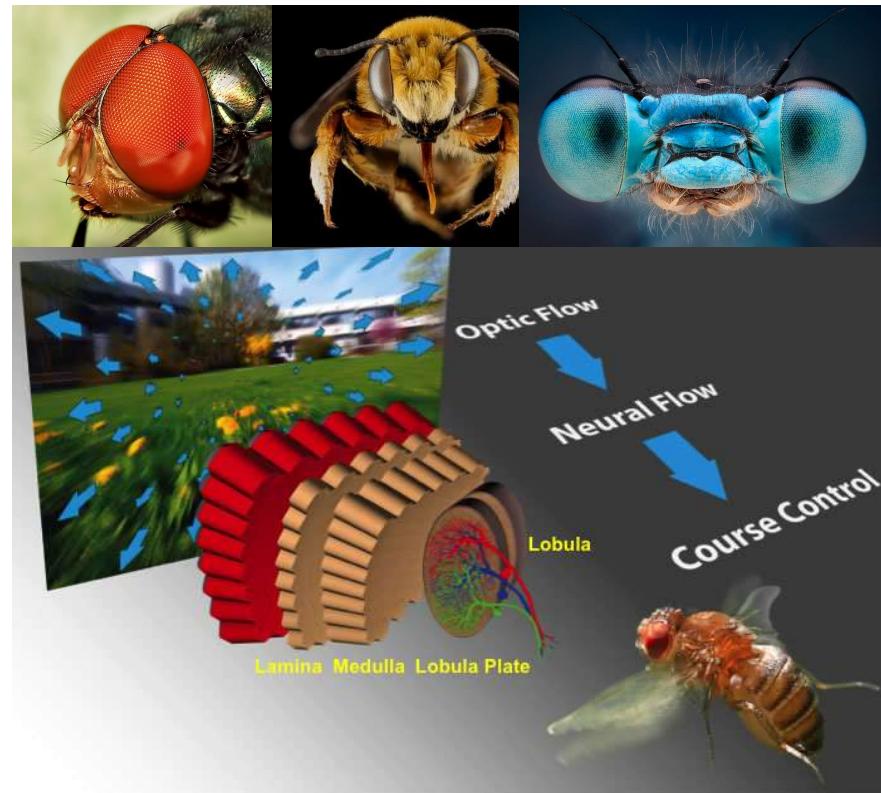
The Perception of the
Visual World

视觉世界的感知



光流

- 经过大量的生理和心理研究，光流信息对
 - 感知自我运动 (Ego-motion)
 - 感知距离 (Distance)
 - 感知形状 (Shape)
 - 感知目标移动 (Object movement)
 - 动物本身的移动控制 (Control of Locomotion)
- 起到至关重要的作用。





光流的计算方法和应用



- 光流计算的两种经典方法：

Lucas-Kanade 算法（稀疏法）

Horn-Schunk 算法（密集法）



稀疏光流



密集光流（每像素）





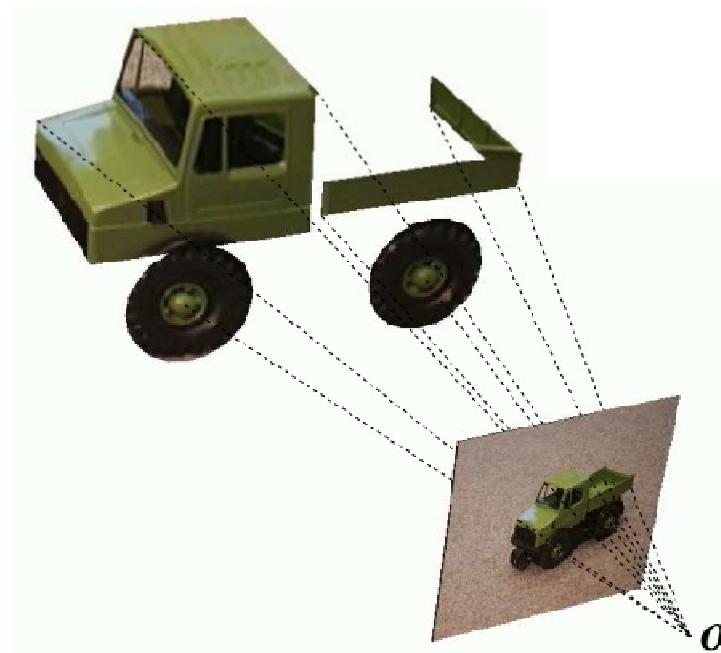
上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

位姿估算 (Pose Estimation)



位姿估算

- 位姿估算
 - 已知物体的几何模型和它在图像上的投影
 - 计算它与物体之间的相对位姿与姿态

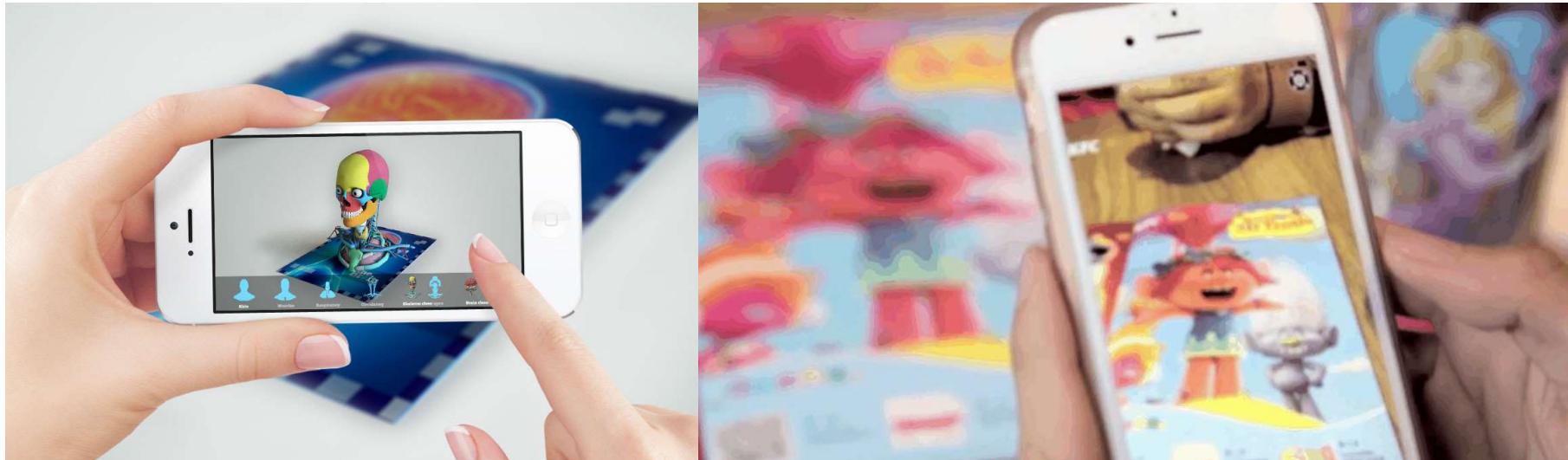




位姿估算



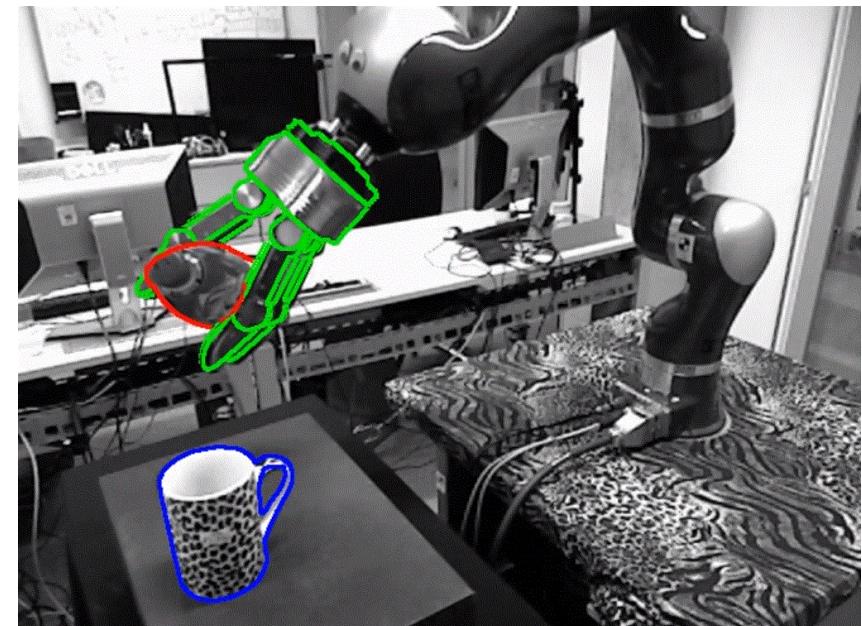
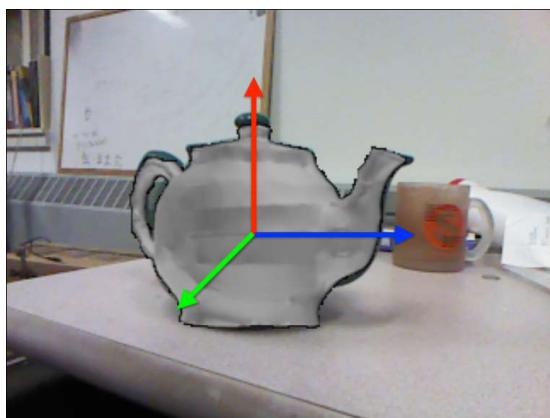
- 应用 - 虚拟现实





位姿估算

- 应用 - 机器人自主抓取





上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

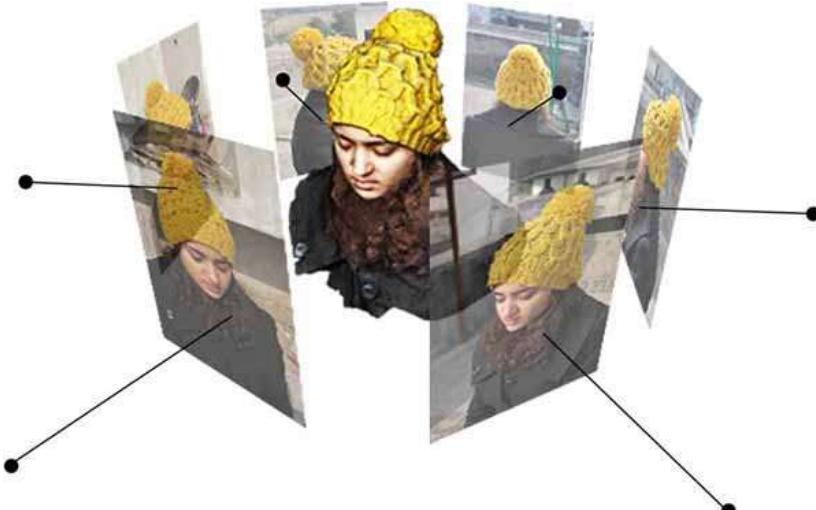


三维重构 (3D Reconstruction)



三维重构

- 从多个视角的图片中恢复场景和物体的三维结构





SLAM

Simultaneous (同步)
Localization (定位)
And (与)
Mapping (地图构建)



SLAM - 同步定位与地图构建技术

- 在未知环境下，定位和地图构建是一个相互耦合的问题：
 - 如果没有地图，无法建立参考坐标系，也就无法定位
 - 如果没有位置信息，地图信息无法进行拼接，形成完整的地图
- 它实际上是一个“鸡与蛋”的问题
该问题被称为SLAM问题 (1980年代)：



Simultaneous
Localization
And
Mapping



SLAM - 同步定位与地图构建技术

- 可用于SLAM的传感器
 - 测距+测向类



2D laser rangefinder



3D LiDAR



RGB-D camera

- 纯测向类



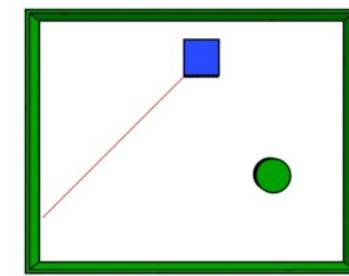
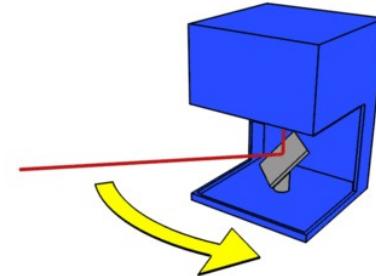
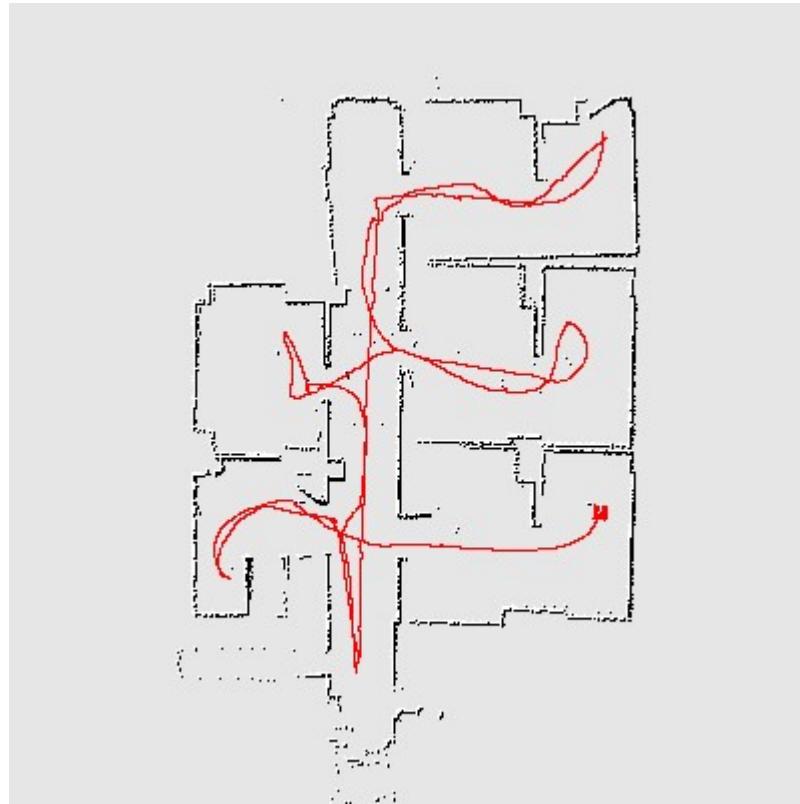
Video cameras





SLAM - 同步定位与地图构建技术

- 激光SLAM : 2D - LiDAR



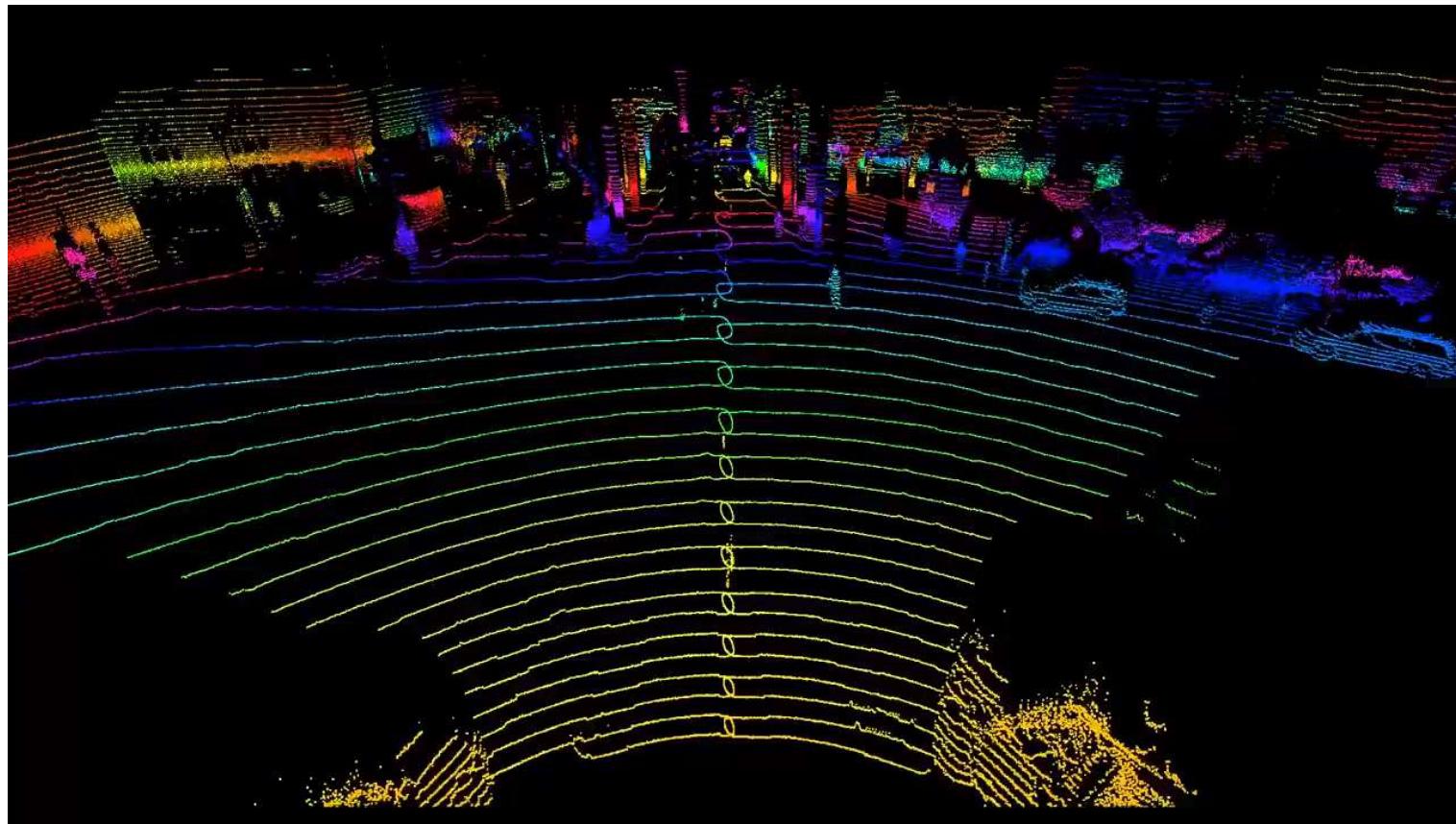
•

•



SLAM - 同步定位与地图构建技术

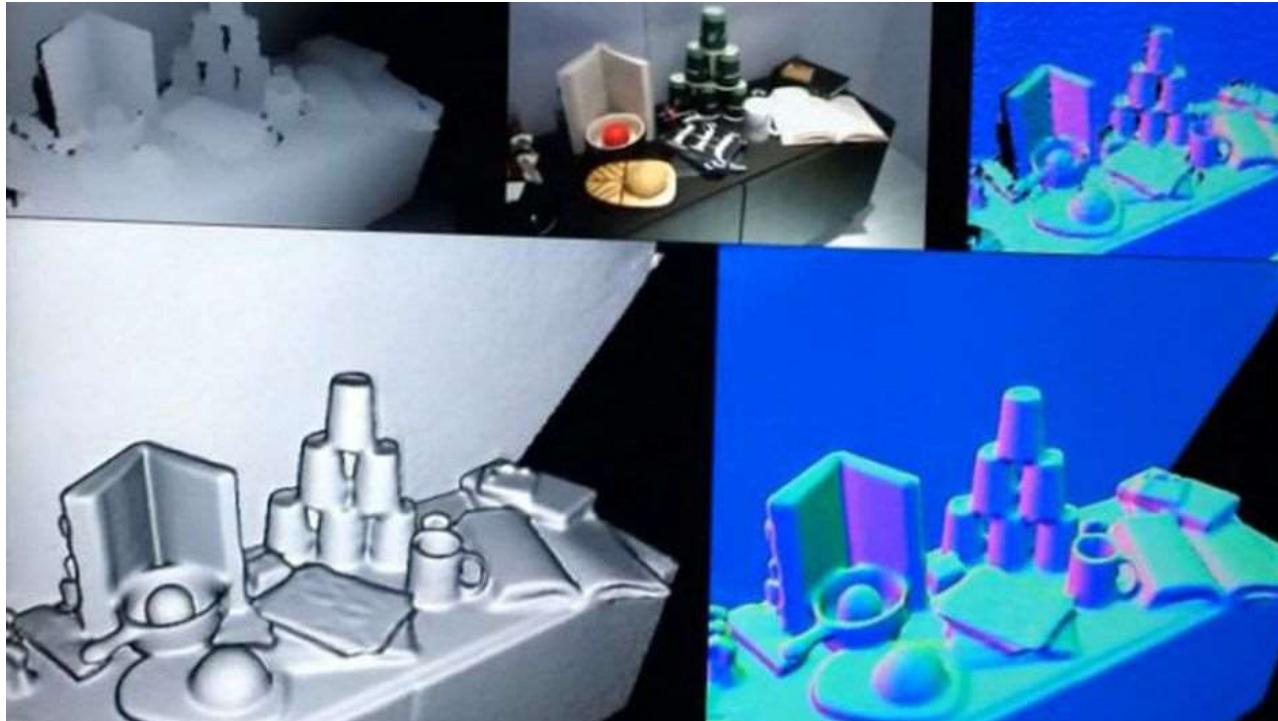
- 激光SLAM : 3D – LiDAR (自动驾驶)





SLAM - 同步定位与地图构建技术

- RGBD深度相机





SLAM - 同步定位与地图构建技术

- RGBD SLAM

SIGGRAPH Talks 2011

KinectFusion:

Real-Time Dynamic 3D Surface
Reconstruction and Interaction

Shahram Izadi 1, Richard Newcombe 2, David Kim 1,3, Otmar Hilliges 1,
David Molyneaux 1,4, Pushmeet Kohli 1, Jamie Shotton 1,
Steve Hodges 1, Dustin Freeman 5, Andrew Davison 2, Andrew Fitzgibbon 1

1 Microsoft Research Cambridge 2 Imperial College London

3 Newcastle University

4 Lancaster University

5 University of Toronto



SLAM - 同步定位与地图构建技术

- 视频摄像头 – 视觉SLAM
- 体积小，成本低，功耗低
- 无处不在：
 - 手机
 - 无人机
 - 车载

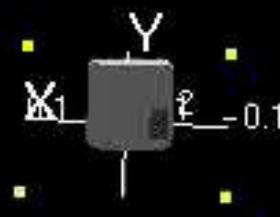
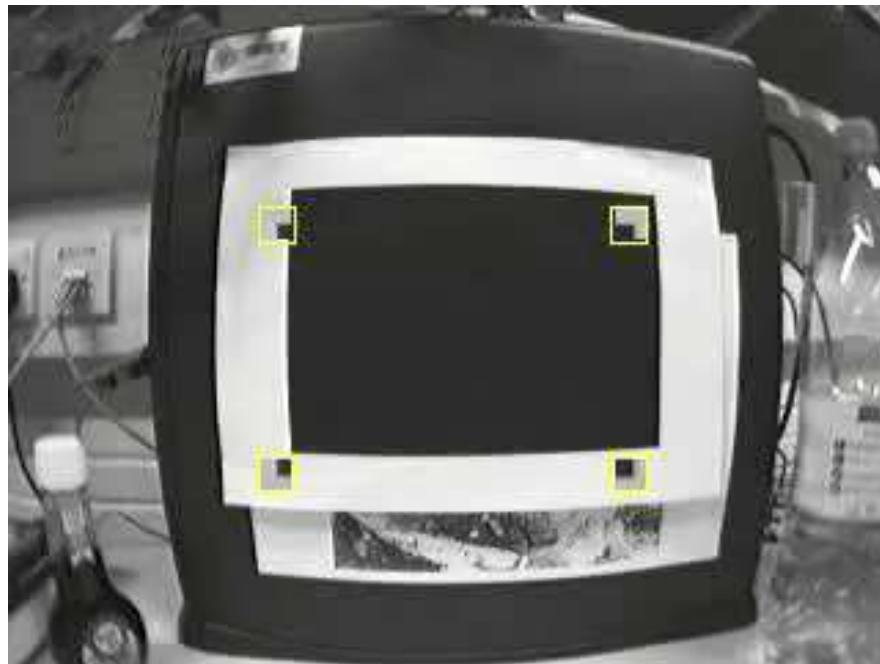


Video cameras



视觉SLAM代表工作

MonoSLAM (Andrew Davison, 2003)





视觉SLAM代表工作

- ORB-SLAM



Universidad
Zaragoza

1542



Instituto Universitario de Investigación
en Ingeniería de Aragón
Universidad Zaragoza

ORB-SLAM2: an Open-Source SLAM System
for Monocular, Stereo and RGB-D Cameras

Raúl Mur-Artal and Juan D. Tardós

raulmur@unizar.es

tardos@unizar.es



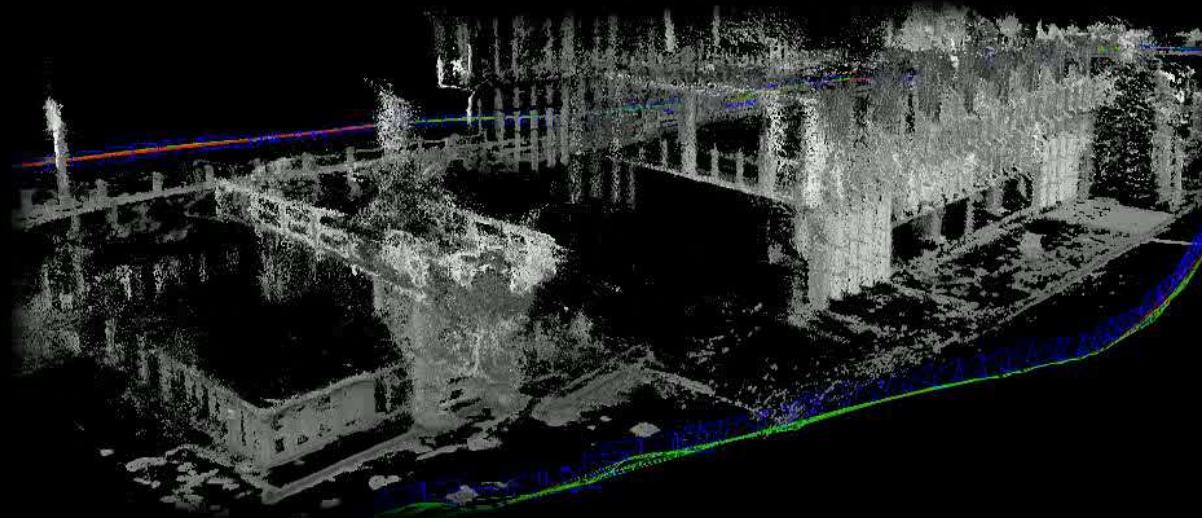
视觉SLAM代表工作



- LSD-SLAM

LSD-SLAM: Large-Scale Direct Monocular SLAM

Jakob Engel, Thomas Schöps, Daniel Cremers
ECCV 2014, Zurich



Computer Vision Group
Department of Computer Science
Technical University of Munich





视觉SLAM代表工作

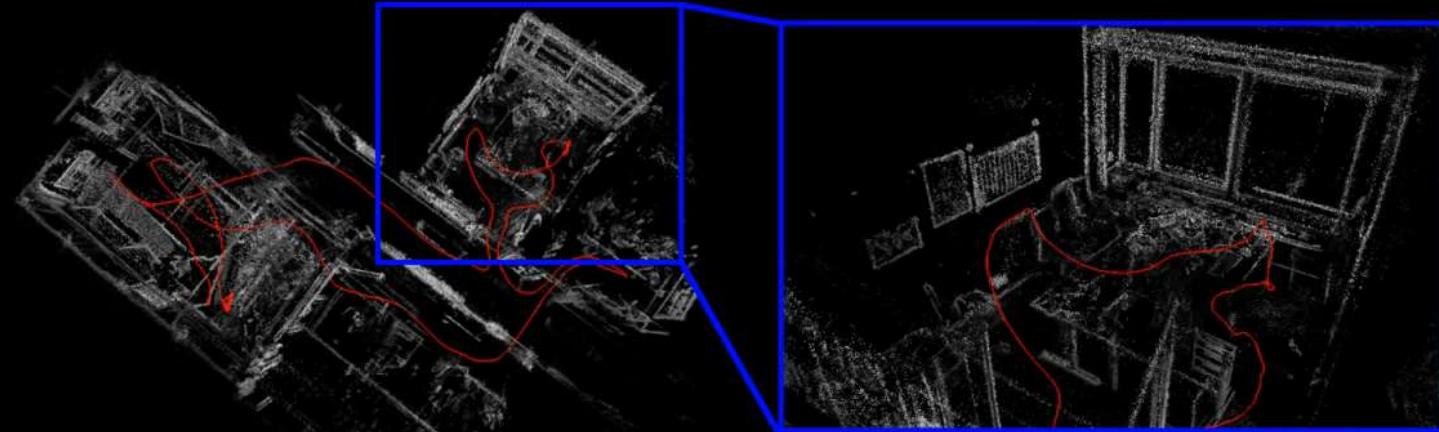


- DSO

Direct Sparse Odometry

Jakob Engel^{1,2}, Vladlen Koltun², Daniel Cremers¹

July 2016



¹Computer Vision Group
Technical University Munich

²Intel Labs



小结-SLAM



- SLAM是实现位置环境下自主导航的关键技术
- 同时解决定位与地图构建问题
- 可用于SLAM的传感器包含激光雷达(2D/3D)、RGB-D深度相机、视频摄像头
- 其中视觉SLAM在价格、体积、重量、以及普适性方面有较大应用优势
- 经过20多年发展，视觉SLAM已经成熟，有不少开源解决方案可供使用



小结-视觉SLAM



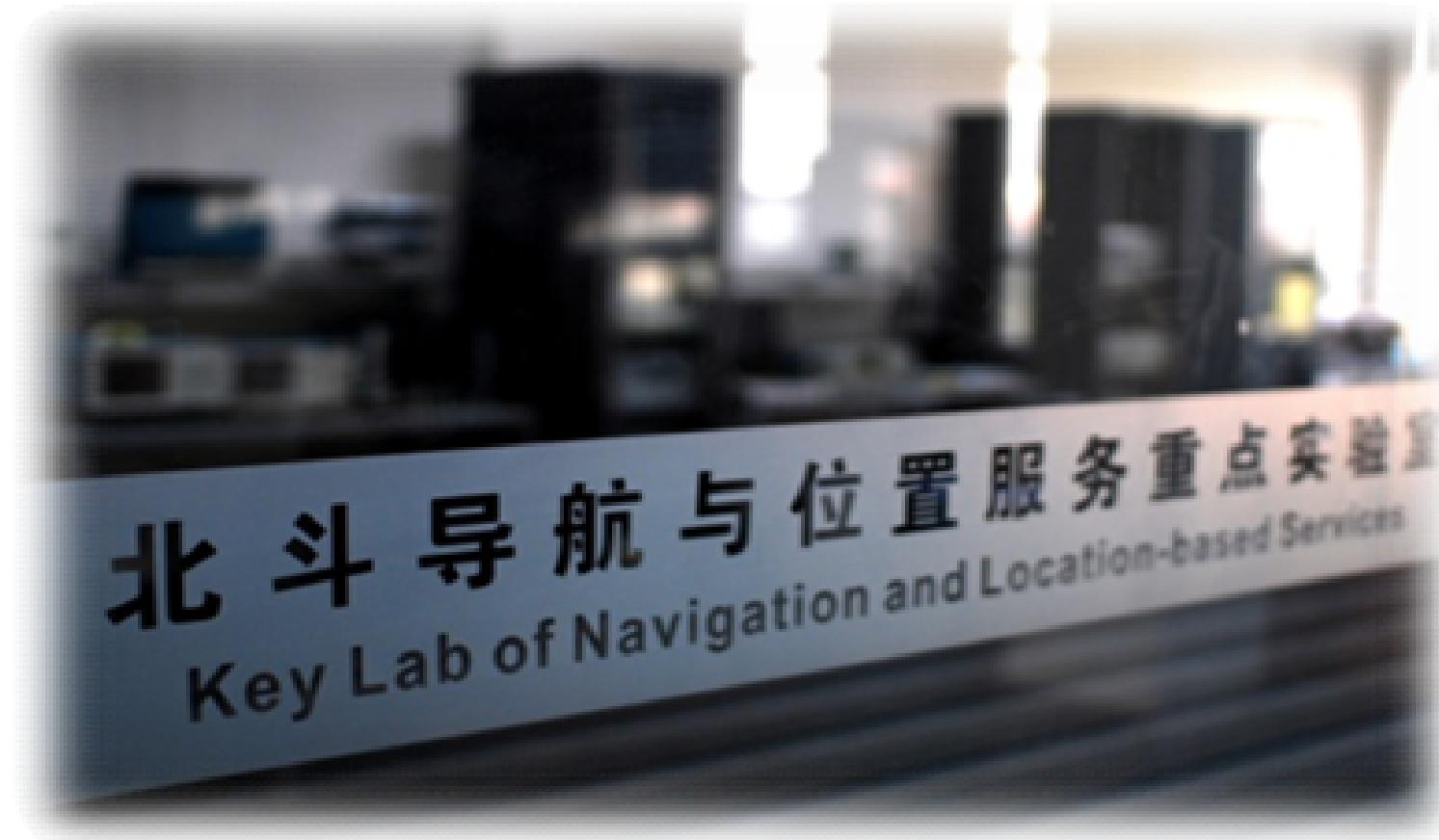
- 利用单个或多个摄像头作为测量输入，同时计算摄像头自身的6自由度位置姿态信息，和周围环境的三维点云信息。
- 随着计算机视觉技术的飞速发展，视觉SLAM算法框架基本成熟。
- 已经有成熟的工业级产品应用（Google Project Tango, Microsoft Holo Lens）
- 但对各细分应用领域还需要结合应用背景进行深度定制。

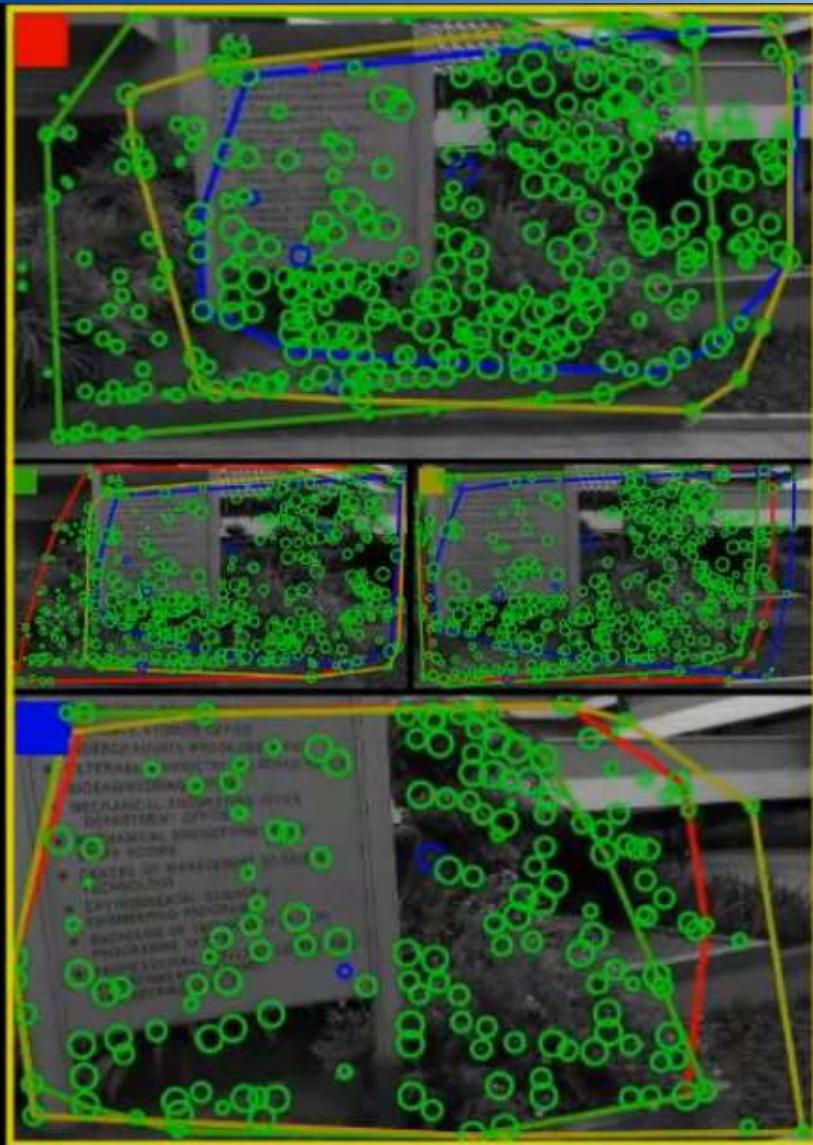


上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

导航与感知研究所

- 视觉导航与无人系统研究组



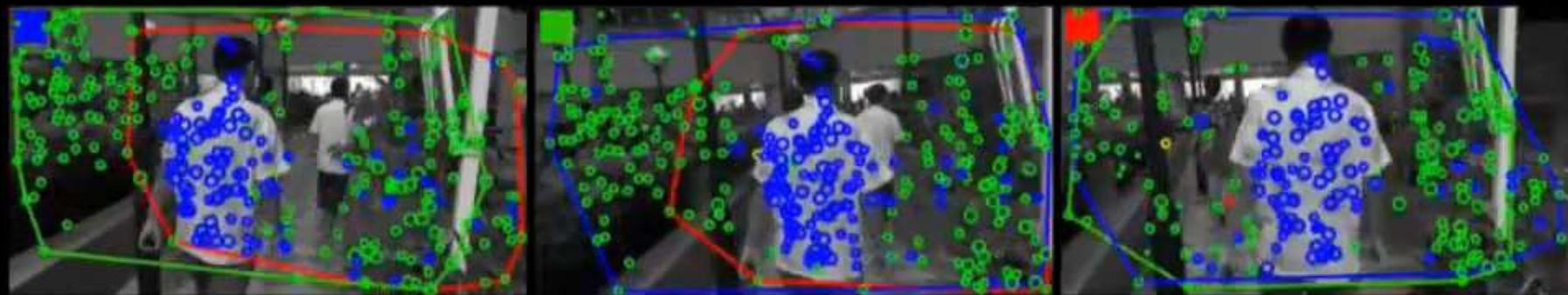


CoSLAM - courtyard example

2 x speed

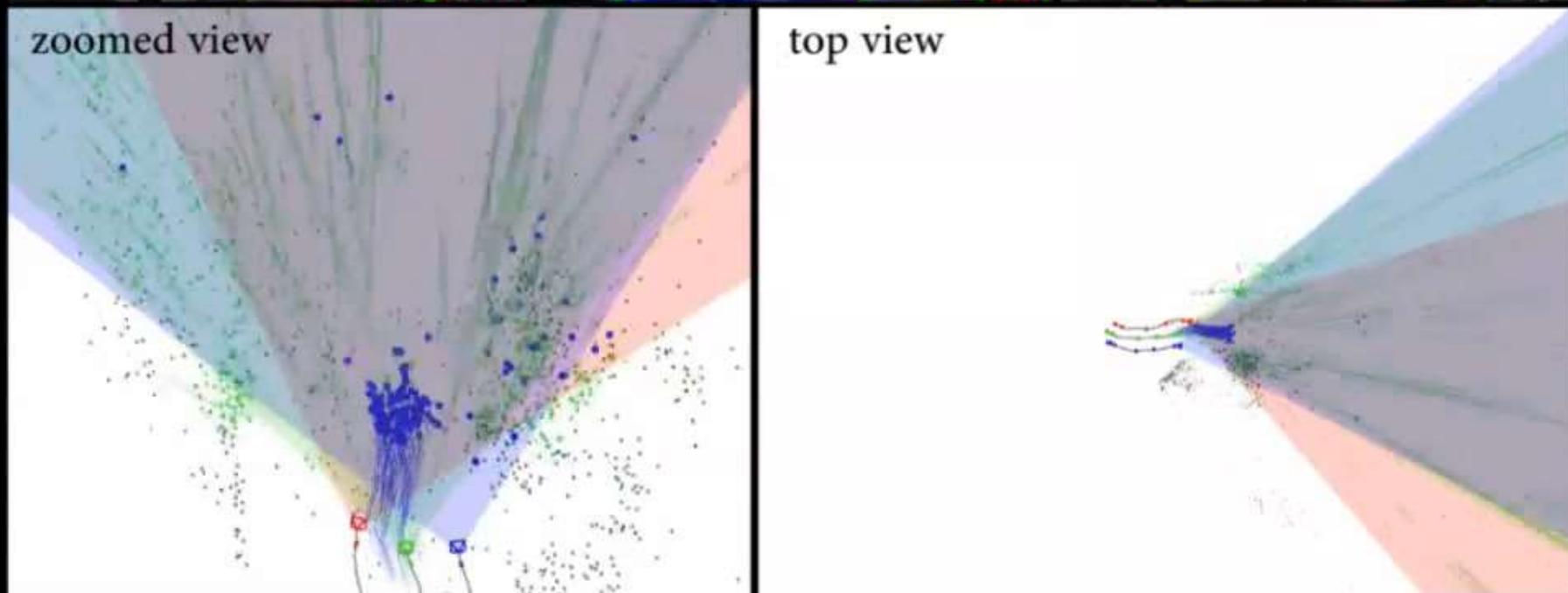
top view

zoomed view



zoomed view

top view



CoSLAM - dynamic environments (outdoor)

2x speed





StructSLAM:

Visual SLAM with Building Structure Lines



Hui Zhong Zhou, Danping Zou et al.

Shanghai Key Laboratory of Navigation and Location Based Services
Shanghai Jiao Tong University
April, 2014



Real system – Multi-Robot vSLAM



Orbiting a Moving Target with Multi-Robot Collaborative Visual SLAM

Jacob M. Perron*, Rui Huang*, Jack Thomas, Lingkang Zhang,
Ping Tan, Richard T. Vaughan
(* are first authors of equal contribution)

Autonomy Lab

GrUVi Lab



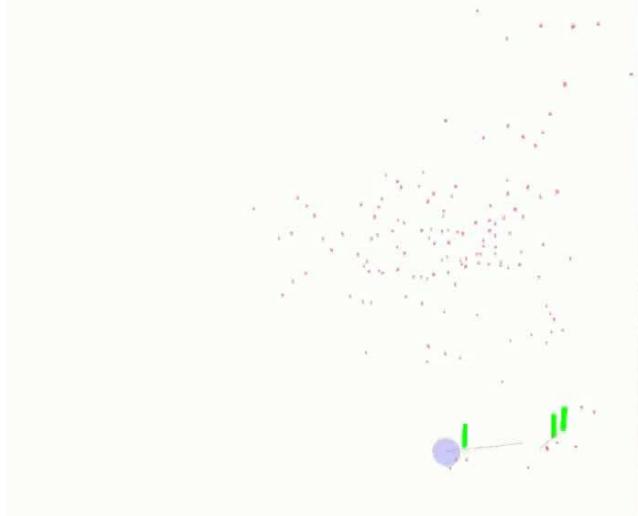
Simon Fraser University



NSERC Canadian Field Robotics Network



Onboard Visual SLAM /UAV





上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

Active Image Modeling





Active Image Modeling

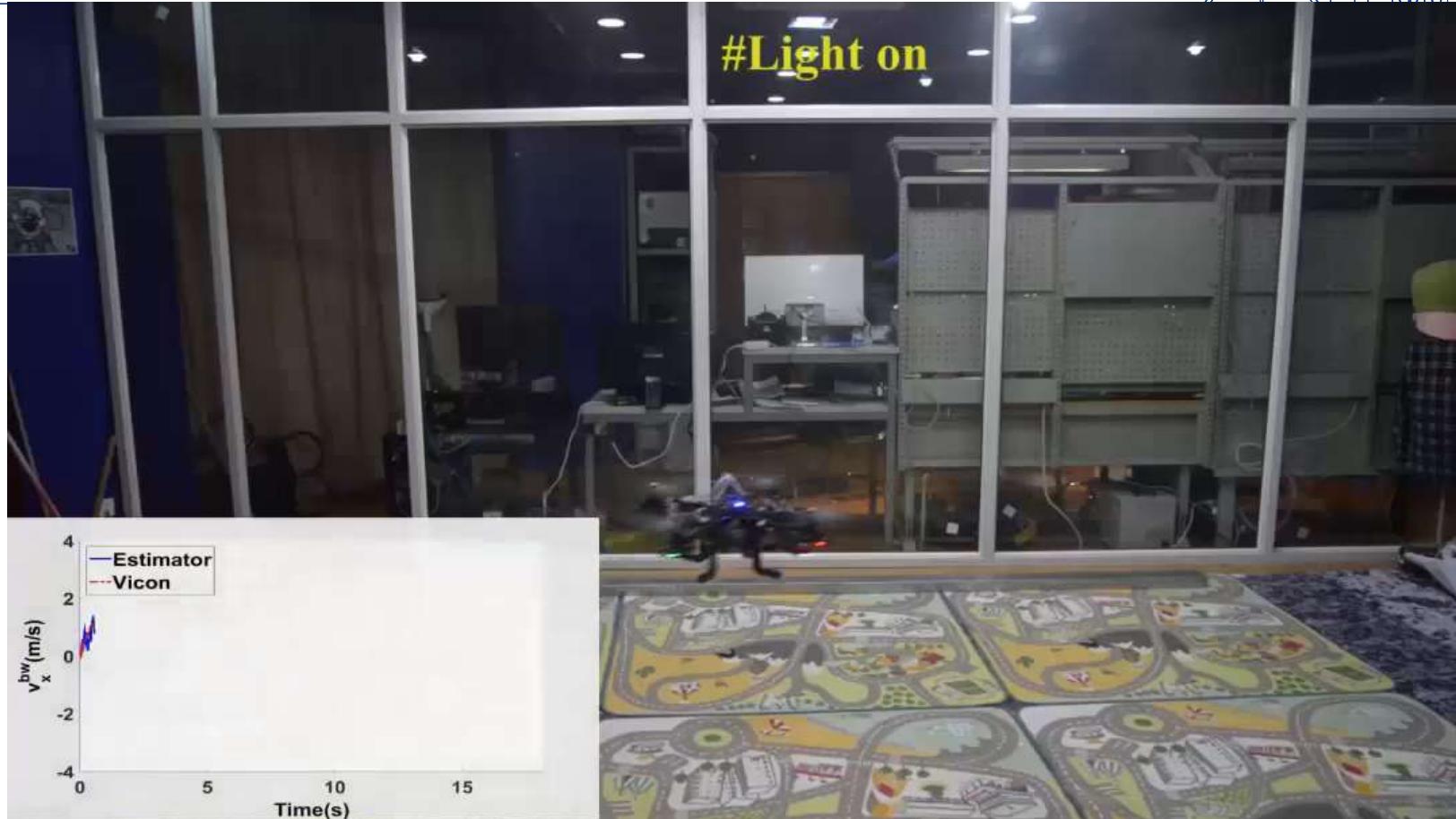


Outdoor Experiments

Eagle



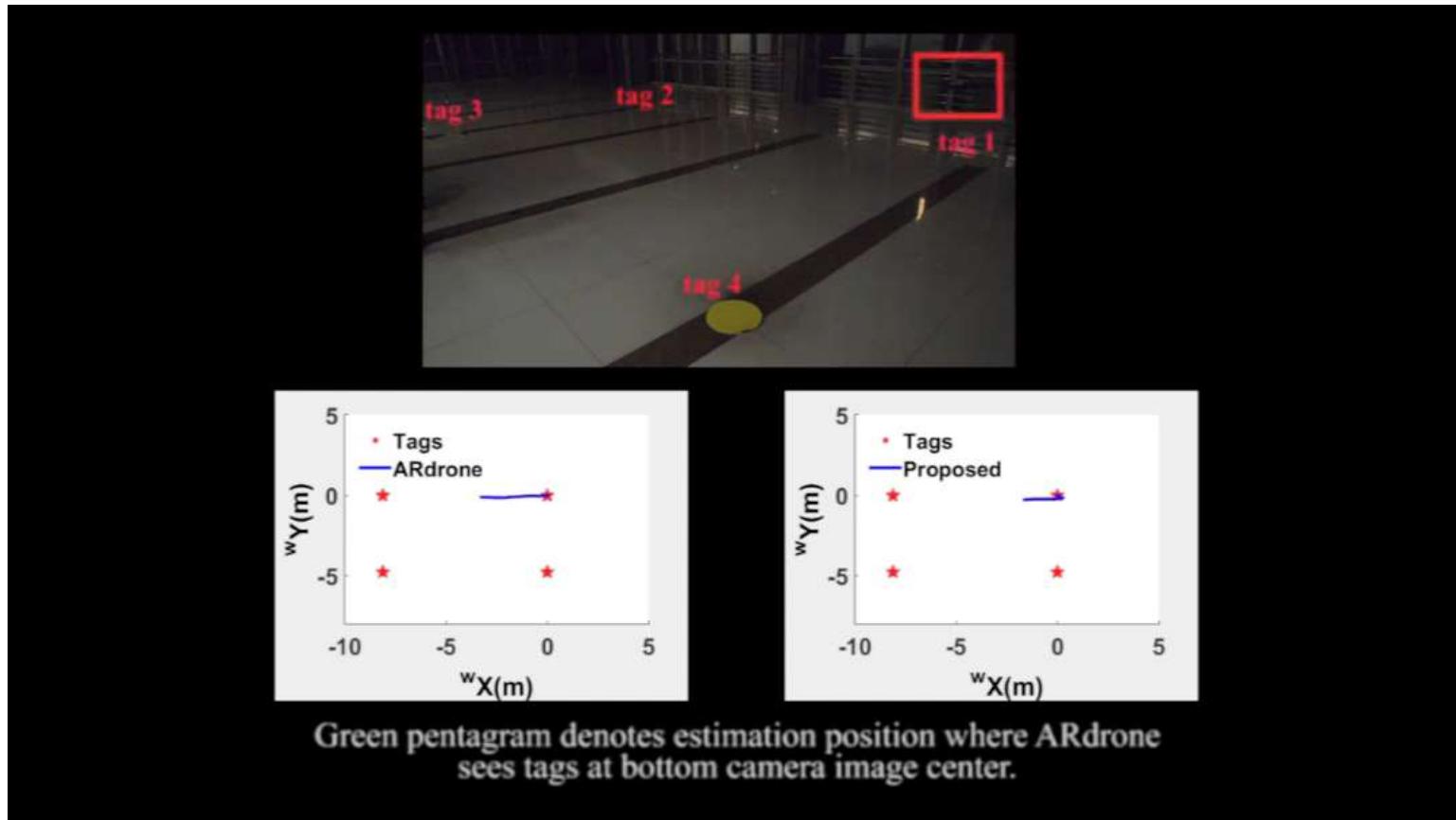
Dark scenes



R. Wang, D. Zou etc, A Robust Aerodynamics-Aided State estimator for Multi-rotor UAVs, IROS, 2017



Low texture + dark scenes



R. Wang, D. Zou etc, An Aerodynamics Model-Aided State estimator for Multi-rotor UAVs, IROS, 2017



<http://drone.sjtu.edu.cn/contest/>



谢谢!

