



# 机器人视觉定位导航

吴毅红，中国科学院大学，2020春季



# 视觉传感器

- 针孔相机
- 广角相机：鱼眼、反射折射等
- 深度相机：双目；结构光；TOF

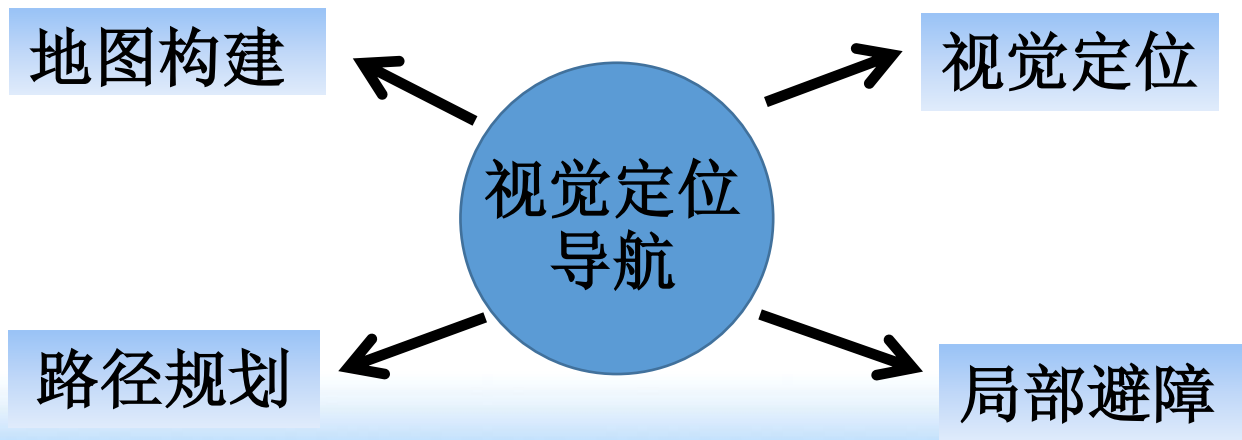
**2017年以来，3D摄像头异常火爆**

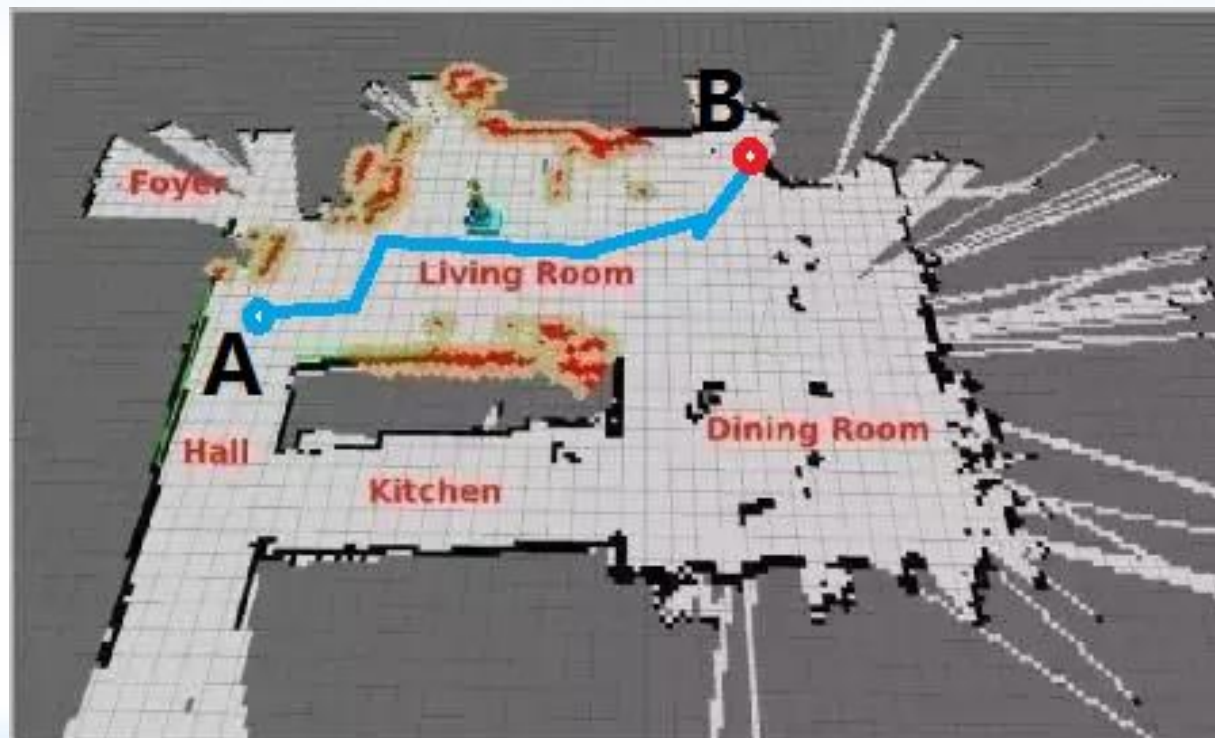
Kinect, 2009年6月，2012年发售，微软；  
2017年10月，停产  
2018年5月，宣布复活

- 实际应用场景中的机器人导航：  
激光、磁导航
- 视觉导航：学术，产业转化阶段
- 未来视觉导航将占主导地位：
  - 视觉定位导航与人或动物的智能更加接近
  - 视觉系统已被广泛应用20 多年
  - 视觉传感器成本低廉且普遍存在

# 视觉定位导航主要任务

- 智能机器人，自主移动。
- 自主导航是机器人的核心技术，指机器人通过传感器感知环境信息和自身位置，实现有障碍的环境中面向目标的自主运动
- 传感器：相机





# Learning To Explore Using Active Neural SLAM

Devendra Singh Chaplot, Dhiraj Gandhi, Saurabh Gupta,  
Abhinav Gupta, Ruslan Salakhutdinov  
Carnegie Mellon University, Facebook AI Research, UIUC

International Conference on Learning Representations, 2020

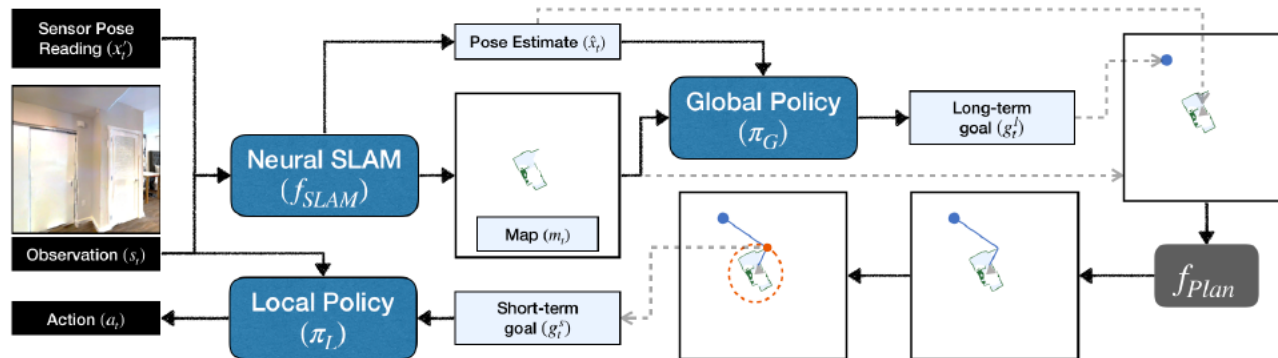
深度学习三大巨头之二的 Yoshua Bengio 和 Yann LeCun 牵头创办

2013年开始

<https://github.com/devendrachaplot/Neural-SLAM>

- A learned Neural SLAM module
- A global policy
- A local policy

To be interfaced via the map and an analytical path planner



**Figure 1: Overview of our approach.** The Neural SLAM module predicts a map and agent pose estimate from incoming RGB observations and sensor readings. This map and pose are used by a Global policy to output a long-term goal, which is converted to a short-term goal using an analytic path planner. A Local Policy is trained to navigate to this short-term goal.

Pyrobot API and ROS to implement the control commands and to get sensor readings

Not end-to-end

A Local Policy is trained to navigate



$$\epsilon_{sen} = p'_1 - p_1 = (x' - x^*, y' - y^*, o' - o^*)$$

$$u_{Forward} = (0.25, 0, 0), u_{Right} : (0, 0, -10 * \pi/180) \text{ and } u_{Left} : (0, 0, 10 * \pi/180).$$

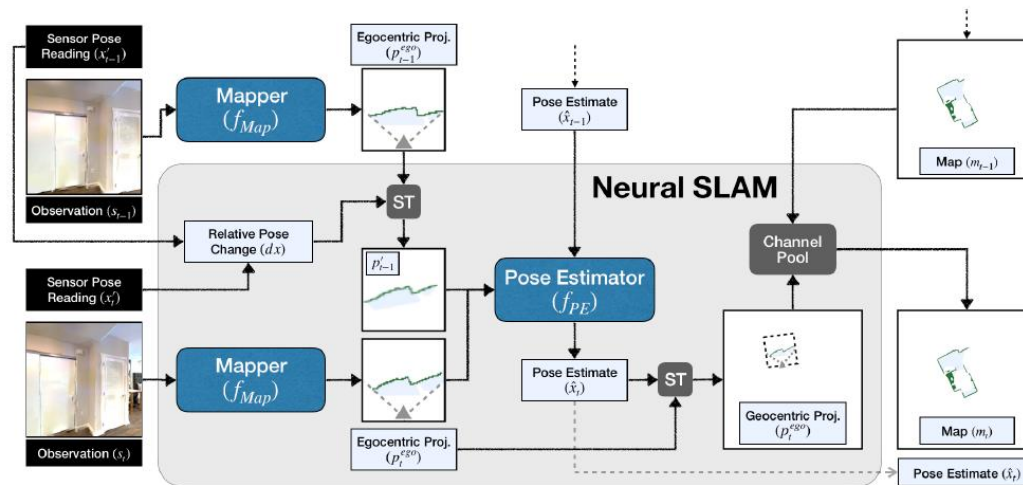
## 1. Map

The spatial map,  $m_t$ , is a  $2 \times M \times M$  matrix

The probability of an obstacle at the corresponding location.

The second channel denotes the probability of that location being explored.

## 2. SLAM



**Figure 2: Architecture of the Neural SLAM module:** The Neural SLAM module ( $f_{SLAM}$ ) takes in the current RGB observation,  $s_t$ , the current and last sensor reading of the agent pose  $x'_{t-1:t}$ , last agent pose estimate,  $\hat{x}_{t-1}$  and the map at the previous time step  $m_{t-1}$  and outputs an updated map,  $m_t$  and the current agent pose estimate,  $\hat{x}_t$ . 'ST' denotes spatial transformation.

$$m_t, \hat{x}_t = f_{SLAM}(s_t, x'_{t-1:t}, \hat{x}_{t-1}, m_{t-1} | \theta_S, b_{t-1})$$

where  $\theta_S$  denote the trainable parameters and  $b_{t-1}$  denotes internal representations of the Neural SLAM module. The Neural SLAM module can be broken down into two parts, a Mapper ( $f_{Map}$ ) and a Pose Estimator Unit ( $f_{PE}$ ). The Mapper outputs an egocentric top-down 2D spatial map,  $p_t^{ego} \in [0, 1]^{2 \times V \times V}$  (where  $V$  is the vision range), predicting the obstacles and the explored area in the current observation:  $p_t^{ego} = f_{Map}(s_t | \theta_M)$ , where  $\theta_M$  are the parameters of the Mapper. It consists of Resnet18 convolutional layers to produce an embedding of the observation. This embedding is passed through two fully-connected layers followed by 3 deconvolutional layers to get the first-person top-down 2D spatial map prediction.

### 3. Global policy and planner

Global Policy uses a convolutional neural network to predict a long-term goal.

### 4. Local policy

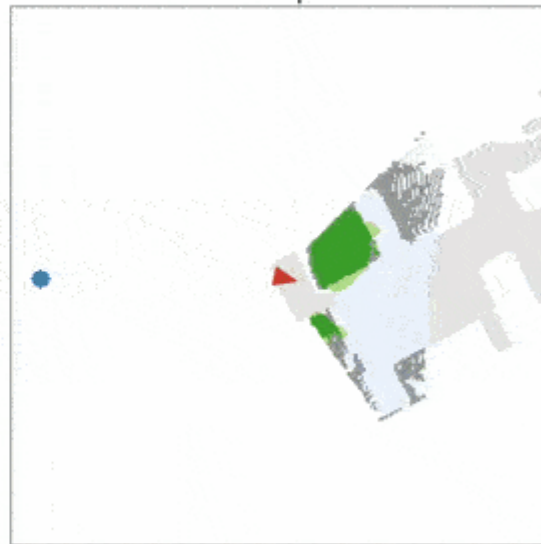
Local Policy is a recurrent neural network consisting of a pretrained ResNet18:

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. CVPR 2016.

Observation



Predicted Map and Pose



- 绪论
- 机器人视觉定位的理论和方法
- 机器人视觉同步定位与地图构建方法
- 机器人视觉建模的理论和方法
- 机器人导航路径规划
- 机器人视觉局部避障
- 总结与考试

## □ 机器人视觉定位的理论和方法

- Youji Feng, Lixin Fan, and Yihong Wu. Fast Localization in Large Scale Environments Using Supervised Indexing of Binary Features. IEEE Transactions on Image Processing, Vol. 25, No. 1, pp. 343-358, 2016.
- L. Liu, H. Li, and Y. Dai. Efficient Global 2D-3D Matching for Camera Localization in a Large-Scale 3D Map. ICCV 2017.
- Eric Brachmann, Carsten Rother. Learning Less is More - 6D Camera Localization via 3D Surface Regression. CVPR 2018. (Deep learning, <https://github.com/vislearn/LessMore>)

## □ 机器人视觉同步定位与地图构建方法

- H. Strasdat, J. M. M. Montiel, and A. J. Davison. Real-time monocular SLAM: Why filter?. 2010 IEEE International Conference on Robotics and Automation, 2010, pp. 2657–2664.
- R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE T-RO, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

## □ 机器人视觉建模的理论和方法

- R.A. Newcombe, S. Izadi, O. Hilliges et al. KinectFusion: Real-time dense surface mapping and tracking. ISMAR, 2011, pp. 127-136.
- M. Nießner, M. Zollhöfer, S. Izadi, et al. Real-time 3D reconstruction at scale using voxel hashing. ACM Transactions on Graphics, 2013, Vol. 32, No.6, pp. 169.

## □ 机器人导航路径规划

- Steven M. LaValle, James J. Kuffner, Jr. Randomized Kinodynamic Planning. The International Journal of Robotics Research, 20(5): 378-400, 2001. (RRT)
- Sertac Karaman, Emilio Frazzoli. Sampling-based algorithms for optimal motion planning. <http://arxiv.org/abs/1005.0416>. (RRT\*)

## □ 机器人视觉局部避障

- Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller. Playing Atari with Deep Reinforcement Learning. <https://arxiv.org/abs/1312.5602>.
- Yufan Chen, Michael Everett, Miao Liu, J.P. How. Socially aware motion planning with deep reinforcement learning. IROS 2017, Best student paper.

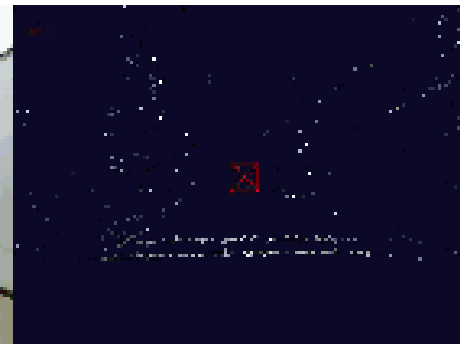
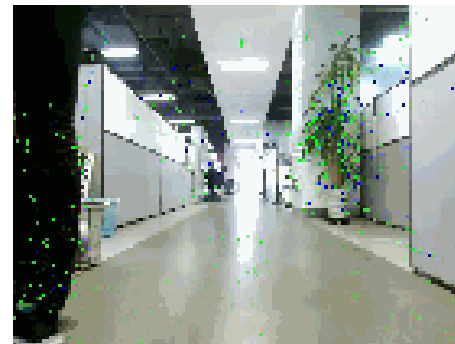
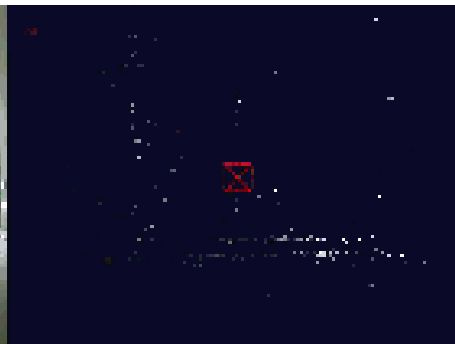
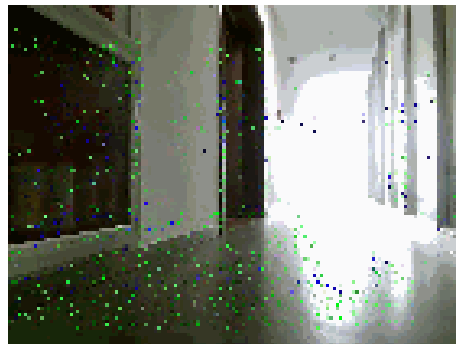
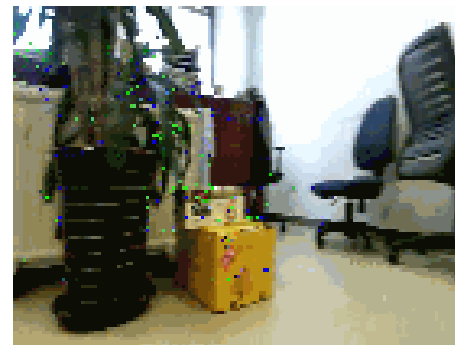
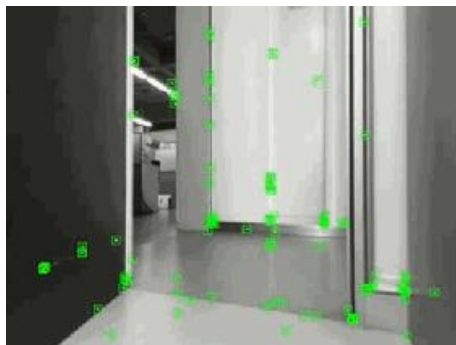


# 系统展示

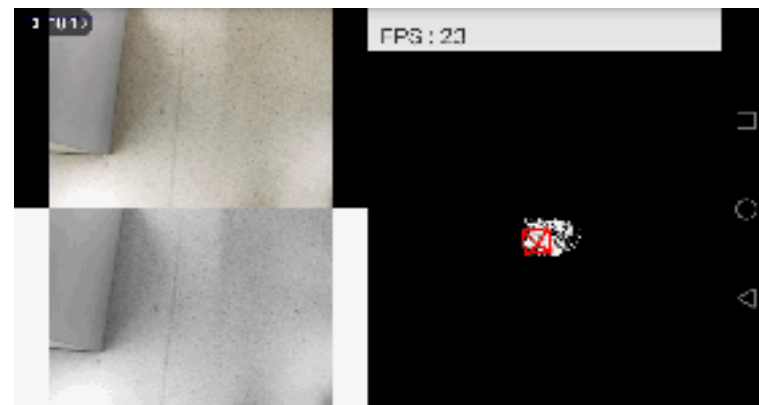
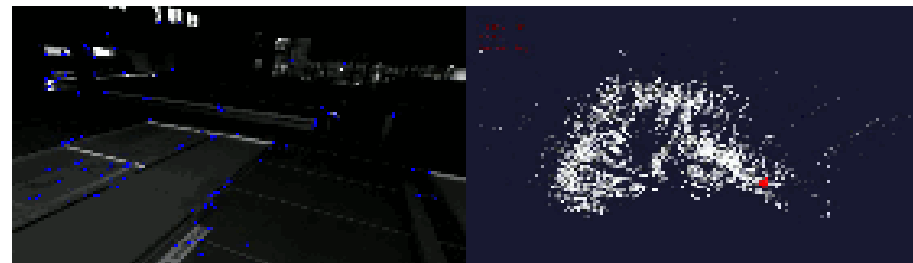
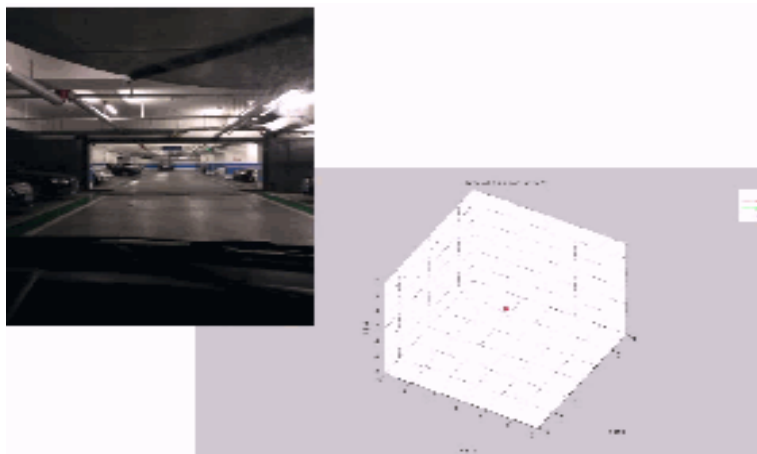
# Mapping



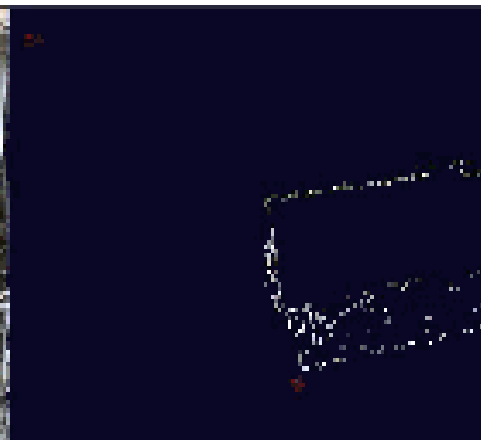
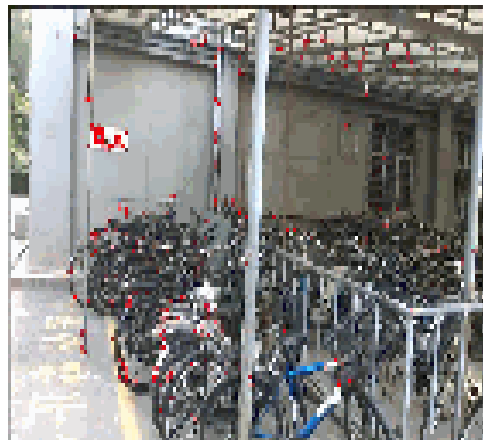
# SLAM in Challenge Environments



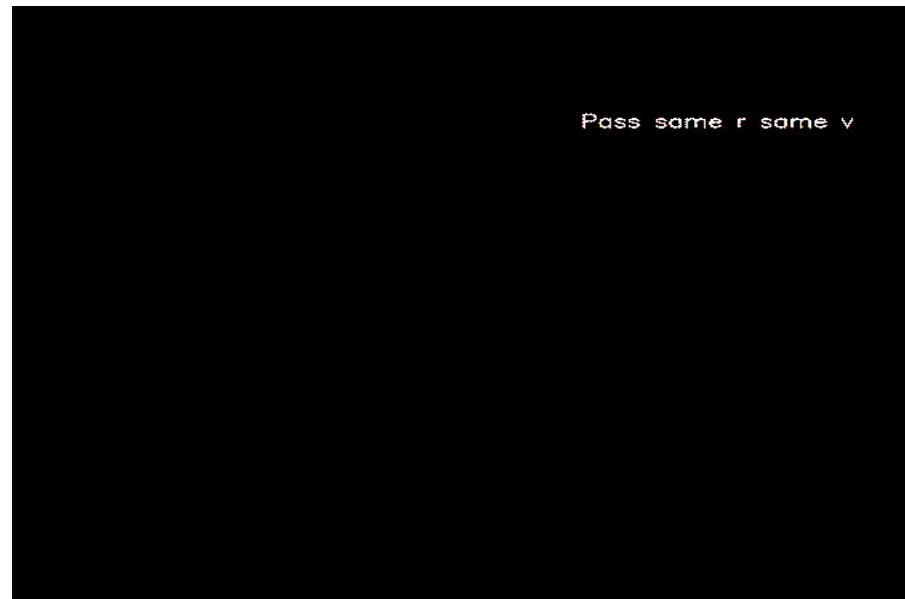
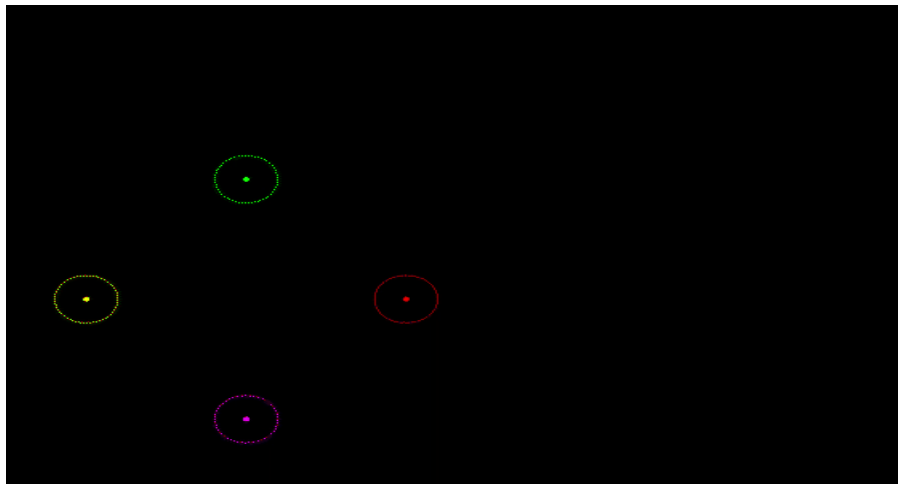
# SLAM in Challenge Environments

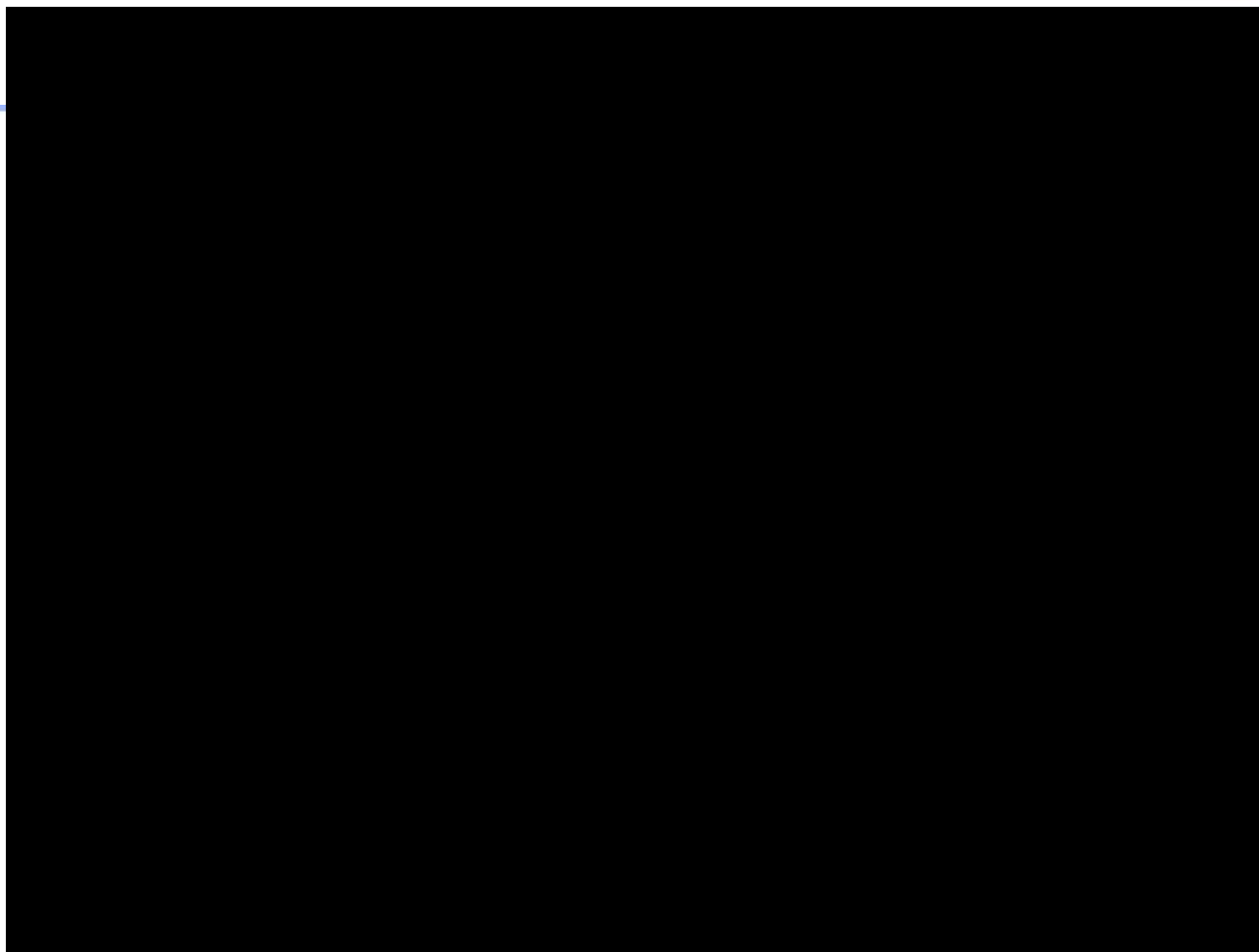


# Real Time Visual Localization in Large Scale Environments



# Obstacle Avoiding in Dynamic Environments



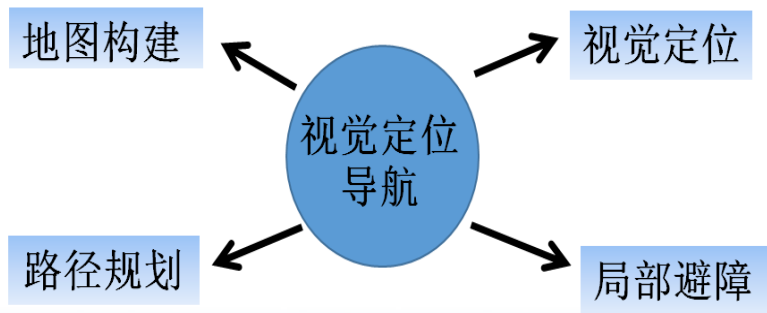


# 课后作业



# 课后思考题

- 你对无人驾驶有何看法？
- 了解无人驾驶的各种级别
- 了解无人驾驶的各种传感器
- 5G可以怎样推动无人驾驶？



**L0--人工驾驶** 由驾驶人全权操控车辆的前进后退，加速减速以及改变方向，在行驶过程中可以得到一些辅助系统的辅助，比如雷达盲点监测等，但是油门、刹车、方向盘都必须是驾驶员控制。

**L1--辅助驾驶** 行车系统在必要的时候或者在驾驶员允许的情况下就会去干涉车辆的加速减速，比如定速巡航系统，在驾驶员设定速度之后，就可以不应去操控油门，车辆就可以自动维持恒定速度行驶；或者紧急制动系统可以实时检测前方道路情况，一旦遇到紧急情况，系统便会自动紧急刹车。这个级别的主要操控者还是驾驶员，辅助操控者是行车系统，并且行车系统只是在某些特定的情况下才会起到辅助驾驶的作用，更多对于周边环境的监测还是人的意识。

**L2--半自动驾驶** 也称为部分自动化，这个级别的主要操控者是行车系统，辅助驾驶者是驾驶员，但如果一旦行车电脑故障或者死机，驾驶员需要立刻接替行车电脑来操控车辆。此时的行车电脑可以监测简单的周围环境，并可以自主操控车辆前进后退、加速减速以及左右转弯，但驾驶员需要时刻监测周围环境，在复杂路面时，驾驶员必须随时介入干涉。

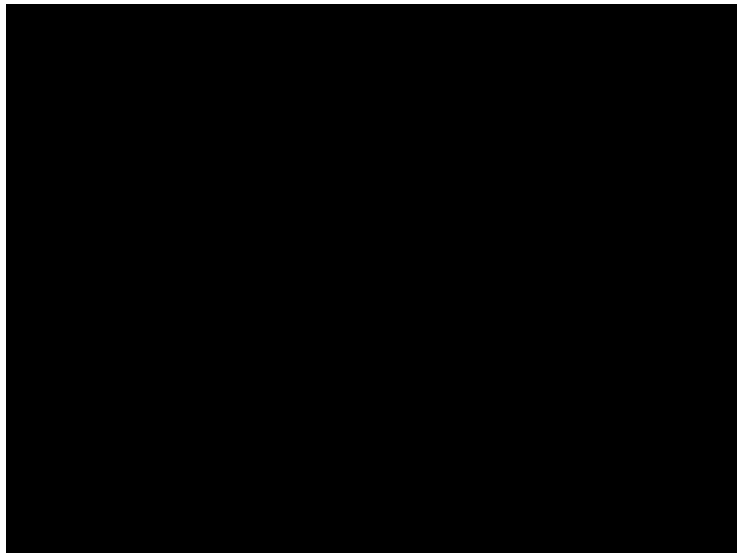
**L3--高度自动驾驶** 也称为有条件自动化，由无人驾驶系统自主完成对油门、刹车以及方向盘的控制，对于周边道路环境的监测也完全是由无人驾驶系统来完成，对于一些复杂一点的路面情况，系统会向驾驶员提出请求，驾驶员给出适当的应答，无人驾驶系统来完成相应的操作。驾驶员不需要过多的关注周围路面环境，而是给予无人驾驶系统以支援。

**L4--超高度自动驾驶** 也称为高度自动化，这个级别的自动驾驶汽车可以自主地完成对车辆的操控，对周边复杂环境地监测和判断，具有更高层次的人的意识，可以自主应对绝大部分的路面情况。

**L5--全自动驾驶** 驾驶员只需要输入目的地，无人驾驶系统自动规划路线，包括加油在内的各种操作都是车辆自主完成，可以应对百分之99.9999的复杂路面情况。打个比方，就相当于《变形金刚》中的车辆，基本具备人类的基本意识。总结 目前市场上的车辆基本上是在L1的级别，只具备一些基本的辅助驾驶功能。但是并不是L2、L3级别的车辆就不存在，如今各大车企的无人驾驶技术基本上都可以实现L2、L3的级别，不过这些车距离量产还有很远的距离，这些车要么在研究人员的实验室里，要么在试验测试道路上。

# 课后思考题

- 观看波士顿机器人SpotMini进行视觉定位导航的demo



- 问题:

- 叙述场景特色
- 叙述传感器
- 叙述其中视觉定位导航中的各模块任务

## About Spot Mini

---

SpotMini is a small four-legged robot that comfortably fits in an office or home. It weighs 25 kg (30 kg if you include the arm). SpotMini is all-electric and can go for about 90 minutes on a charge, depending on what it is doing. SpotMini is the quietest robot we have built.

SpotMini inherits all of the mobility of its bigger brother, [Spot](#), while adding the ability to pick up and handle objects using its 5 degree-of-freedom arm and beefed up perception sensors. The sensor suite includes stereo cameras, depth cameras, an IMU, and position/force sensors in the limbs. These sensors help with navigation and mobile manipulation.

## 波士顿动力创始人：SpotMini要成为机器人版的安卓平台

波士顿动力创始人Marc Raibert在CEBIT上的演讲题为“SpotMini产品平台：机器人的Android”，暗示机器人将成为像安卓一样无处不在的机器人开源平台。他设想，这些机器人被改造后可以在很多领域工作，比如建筑、安保、最后一英里的运送，甚至还可以在家里帮忙。

# 课后作业

## 了解四足机器人

- 综述

<https://mp.weixin.qq.com/s/kVbjZdw6t9wymCd84oXbpA>

- 国内外四足机器人团队

<https://mp.weixin.qq.com/s/oTELx9nJLDEnqAvdTquhjA>



# 课后思考题

了解BAT在机器人领域的布局，以及它们各自的特色

消费级机器人公司优必选有哪些消费级机器人？

你对人形机器人的前景如何看待？

# 了解BAT在机器人领域的布局，以及它们各自的特色



2018年3月，测试已获官方授权，百度 Apollo 自动驾驶汽车在亦庄周边的开放道路上进行公开路测



阿里：  
菜鸟小G；菜鸟小G2；  
菜鸟Gplus

朋友圈一篇名为《腾讯没有梦想》的文章引起了热议，文章开头的一句“腾讯正在丧失产品能力和创业精神，变成一家投资公司.....”

小布；小Q；  
2018年3月，成立了机器人实验室Robotics X研发机器人技术，  
张正友：三维视觉、表情识别



Lund University



京东无人送货车

谷歌5.5亿美元投资京东，将在全球合作开发零售解决方案

# CVPR 2018 自动驾驶展

- **Waymo:** alphabet公司旗下的无人车企业 Waymo, 这家公司首次将其无人车开到了 CVPR 的展台, 同时也是会议的赞助商. 传感器包括激光雷达、毫米波雷达以及摄像头等等.
- **通用 Cruise:** 通用旗下的自动驾驶公司, Cruise Automation 激光雷达与摄像头的组合安装在车顶.
- **福特 Argo AI:** 论文主要是关于基于视频的目标物识别、地图与定位技术、图像数据处理以及基于深度学习技术的视觉 SLAM.
- **Uber、Lyft Level5、AutoX、Ouster、Mighty.ai、DeepMap 以及 Aptiv&nuTonomy**等
- 百度, 阿里, 腾讯, 华为, 滴滴, 京东等.

总数接近 30 家, 占据了参展企业的四分之一。

# 课后作业

- 了解SpaceX的龙飞船卫星发射新闻
- SpaceX公司火箭精准着陆回收，用到的有可能是什么导航定位技术？
- 畅想SpaceX的“星链”计划对未来定位导航的影响

# 谢 谢

yhwu@nlpr.ia.ac.cn  
<http://vision.ia.ac.cn/>

