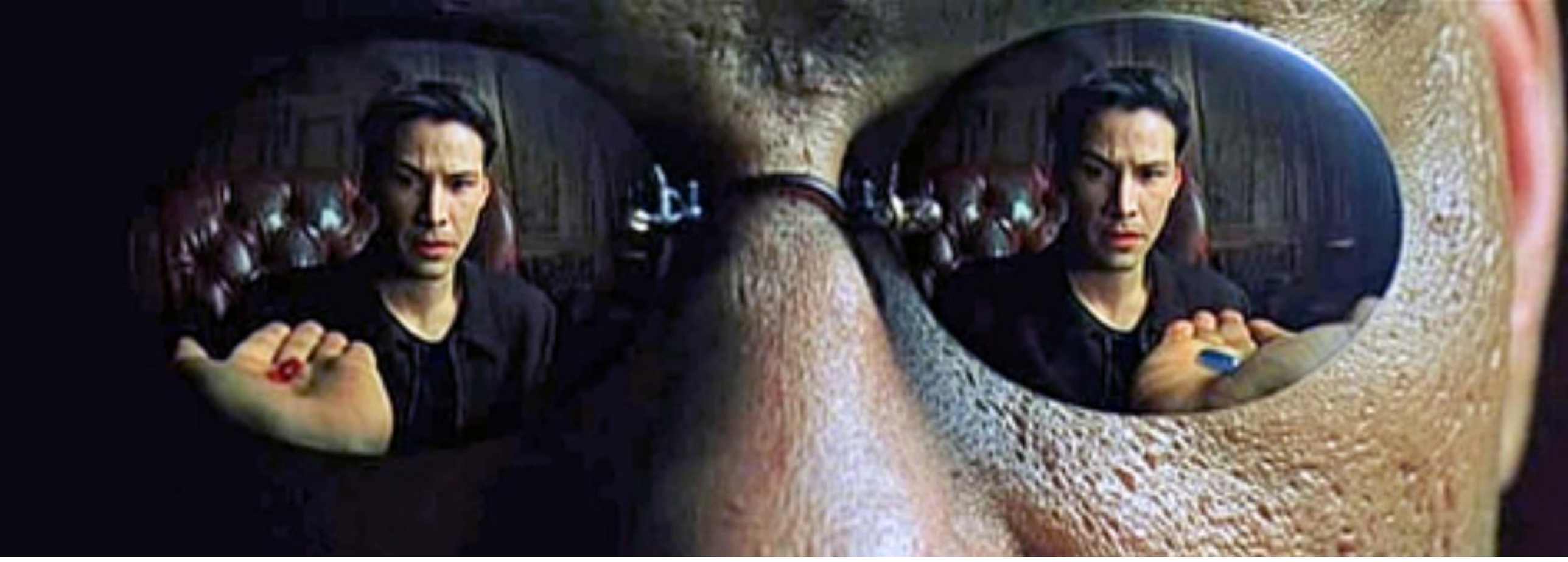


16.485: VNAV - Visual Navigation for Autonomous Vehicles

服务于自动驾驶的视觉导航技术

Luca Carlone

MIT2023年课程资料

Lecture 1 **(2024年整理)**

2025.11.3

VNAV Staff / 课程教学团队

MIT课程团队

—Instructor—

Luca Carlone

Rajat Talak

Guest Lecturers

Jared Strader

Yulun Tian

中国海洋大学引入该课程，授课团队：

■老师：范浩

■助教：安康，冯海林，杨浩迪，范继腾，杨雅麟，刘晨笑，黄傲宇

Lecture Outline / 大纲

- **课程基础（Logistics)**
  - VNAV 课程目的
  - 课程内容大纲
  - 作业 & 分数
  - 课程要求

- **Visual Navigation for Autonomous Vehicles （VNAV)**
  - 机器人革命 / The robot revolution.
  - 感知的重要性 / Why perception?
  - More on the VNAV

课程目的 / Goals of VNAV (1/3)

- **了解机器人的最新 视觉导航感知技术**

  (2D computer vision, localization, mapping,

  object pose estimation, semantic understanding,…)


[iRobot]


[Amazon]


[Forbes]

自动驾驶汽车，无人机，家用和服务机器人，虚拟现实、增强现实！

(Self-driving cars, drones, domestic and service robots,
virtual and augmented reality)

2025.11.3

MIT AEROASTRO

课程目的 / Goals of VNAV (2/3)

- 主要理论
  - 学习/开发/练习机器人研究所需的理论工具（**几何、优化**等)
  - 了解最先进的机器人感知算法（控制、轨迹优化、智能学习方面的选定主题)
  - 了解**机器人感知**中的未解决问题

- **Note 1**: If you don't like writing (some) **math**, you may not like this class
- **Note 2**: this course is much more theoretical
- **Note 3**: this is not an ML course (but we use ML in some labs/lecture)
- **Overarching goal**: prepare you (or refine your skills) to perform state-of-the-art research in robotics (not necessarily in robot perception)

课程目的 / Goals of VNAV (3/3)

- 实践（ practice ）
  - 学习/实践 **ROS** (**R**obot **O**perating **S**ystem)
  - 在逼真的模拟器中对最先进的感知、控制、机器学习实现进行**严格测试**（**Rigorous testing**）

Rigorous testing 严格测试：一种系统的、全面的测试方法，旨在确保产品或系统的质量和可靠性。

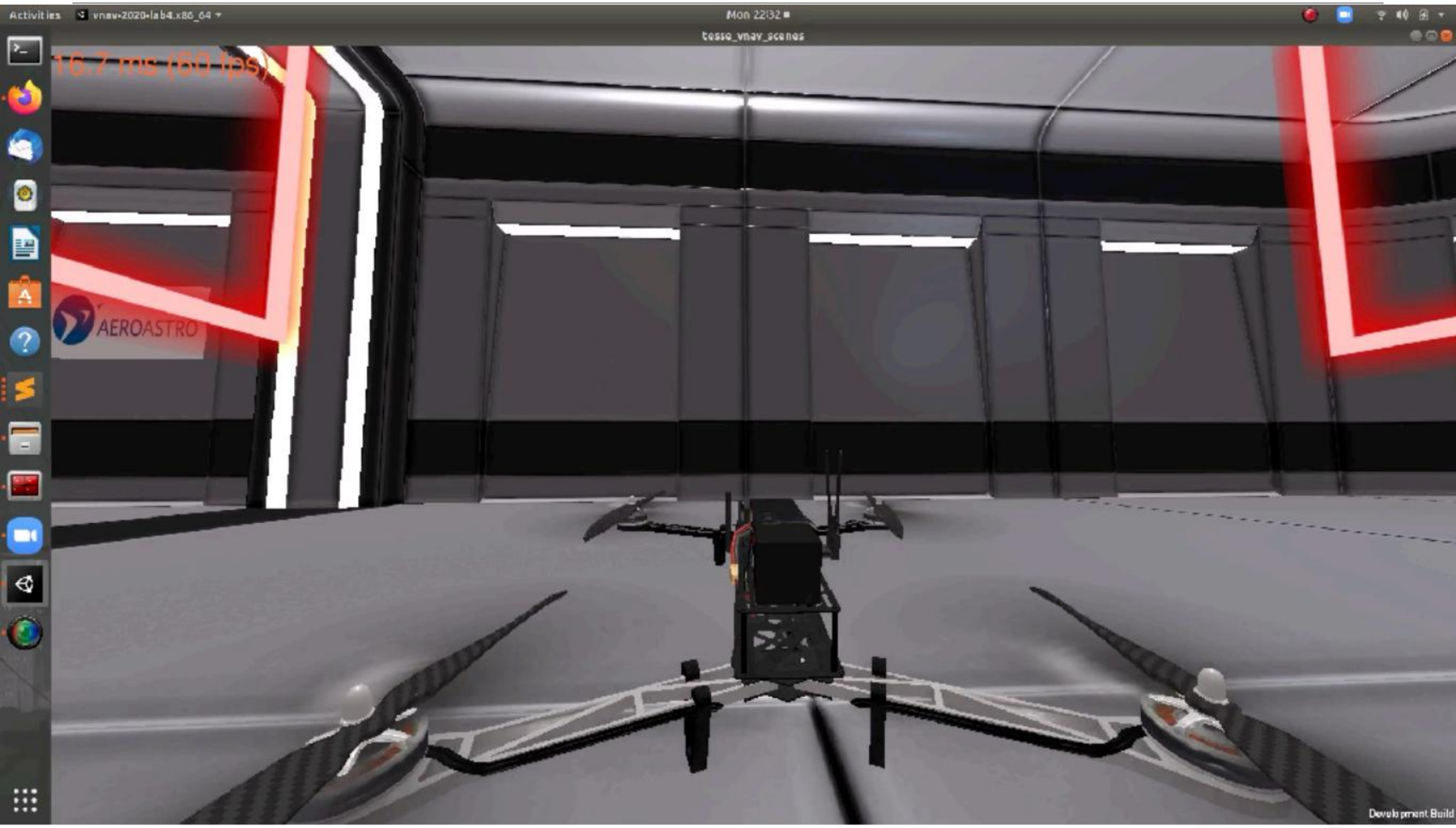
  - 在最终项目中使用**物理机器人**（下学期）
  - 了解**最先进方法**在实践中的**局限性**

- **Note**: If you do not like **coding (in C++)**, you may not like this class
  -> **C++ refresher**: https://youtu.be/F_vIB3yjxaM
- **Overarching goal**: prepare you (or refine your skills) to perform state-of-the-art research in robotics (not necessarily in robot perception)

Practice on Real Robots: Intel Aero Drone

MIT AEROASTRO

Practice in Realistic Unity-based Simulator



MIT AEROASTRO

课程内容大纲

- 理论课:
  - 1:30-3:20pm
    -- 信息南楼B103
- 实验课:
  - 3:30-5:30pm
    -- 信息南楼C108
  - Code, assignments, OH
- 老师办公室:
  - 信息南楼C409
  - 助教:信息南楼A304

Week	Dates	Lecture topic	Lab
1	Sep 6, 8	Introduction	Lab 1: Linux, C++, Git
2	Sep 11, 13, 15	3D Geometry	Lab 2: ROS
3	Sep 18, 20	Geometric Control	Lab 3: 3D trajectory following
4	Sep 25, 27, 29	Trajectory Optimization	Lab 4: 3D trajectory optimization
5	Oct 2, 4, 6	2D Computer Vision	Lab 5: feature detection
6	Oct 11, 13	2-view Geometry and Minimal Solvers	Lab 6: object localization
7	Oct 16, 18, 20	Non-minimal Solvers and Visual Odometry	Lab 7: GTSAM
8	Oct 23, 25, 27	Place Recognition	Lab 8: ML for robotics
9	Oct 30, Nov 1, 3	SLAM and Visual-Inertial Navigation	Lab 9: SLAM
10	Nov 6, 8	Advanced Topics	Final project
11	Nov 13, 15, 17	Advanced Topics	Final project
12	Nov 20, 22	Advanced Topics	Final project
13	Nov 23-26	Thanksgiving Break	
14	Nov 27, 29, Dec 1, 4 ,6, 11, 13	Guest Lectures and Students Presentations	Final project

Structure of VNAV

2025.11.3

课程内容大纲

- 理论课:
  - 1:30-3:20pm
    -- 信息南楼B103
- 实验课:
  - 3:30-5:30pm
    -- 信息南楼C108
  - Code, assignments, OH
- 老师办公室:
  - 信息南楼C409
  - 助教:信息南楼A304

Week	Dates	Lecture topic	Lab
1	Sep 6, 8	Introduction	Lab 1: Linux, C++, Git
2	Sep 11, 13, 15	3D Geometry	Lab 2: ROS
3	Sep 18, 20	Geometric Control	Lab 3: 3D trajectory following
4	Sep 25, 27, 29	Trajectory Optimization	Lab 4: 3D trajectory optimization
5	Oct 2, 4, 6	2D Computer Vision	Lab 5: feature detection
6	Oct 11, 13	2-view Geometry and Minimal Solvers	Lab 6: object localization
7	Oct 16, 18, 20	Non-minimal Solvers and Visual Odometry	Lab 7: GTSAM
8	Oct 23, 25, 27	Place Recognition	Lab 8: ML for robotics
9	Oct 30, Nov 1, 3	SLAM and Visual-Inertial Navigation	Lab 9: SLAM
10	Nov 6, 8	Advanced Topics	Final project
11	Nov 13, 15, 17	Advanced Topics	Final project
12	Nov 20, 22	Advanced Topics	Final project
13	Nov 23-26	Thanksgiving Break	
14	Nov 27, 29, Dec 1, 4 ,6, 11, 13	Guest Lectures and Students Presentations	Final project

Structure of VNAV

2025.11.3

课程内容大纲

- 理论课:
  - 1:30-3:20pm
    -- 信息南楼B103
- 实验课:
  - 3:30-5:30pm
    -- 信息南楼C108
  - Code, assignments, OH
- 老师办公室:
  - 信息南楼C409
  - 助教:信息南楼A304

Week	Dates	Lecture topic	Lab
1	Sep 6, 8	Introduction	Lab 1: Linux, C++, Git
2	Sep 11, 13, 15	3D Geometry	Lab 2: ROS
3	Sep 18, 20	Geometric Control	Lab 3: 3D trajectory following
4	Sep 25, 27, 29	Trajectory Optimization	Lab 4: 3D trajectory optimization
5	Oct 2, 4, 6	2D Computer Vision	Lab 5: feature detection
6	Oct 11, 13	2-view Geometry and Minimal Solvers	Lab 6: object localization
7	Oct 16, 18, 20	Non-minimal Solvers and Visual Odometry	Lab 7: GTSAM
8	Oct 23, 25, 27	Place Recognition	Lab 8: ML for robotics
9	Oct 30, Nov 1, 3	SLAM and Visual-Inertial Navigation	Lab 9: SLAM
10	Nov 6, 8	Advanced Topics	Final project
11	Nov 13, 15, 17	Advanced Topics	Final project
12	Nov 20, 22	Advanced Topics	Final project
13	Nov 23-26	Thanksgiving Break	
14	Nov 27, 29, Dec 1, 4 ,6, 11, 13	Guest Lectures and Students Presentations	Final project

Structure of VNAV

2025.11.3

课程内容大纲

- 理论课:
  - 1:30-3:20pm
    -- 信息南楼B103
- 实验课:
  - 3:30-5:30pm
    -- 信息南楼C108
  - Code, assignments, OH
- 老师办公室:
  - 信息南楼C409
  - 助教:信息南楼A304

Week	Dates	Lecture topic	Lab
1	Sep 6, 8	Introduction	Lab 1: Linux, C++, Git
2	Sep 11, 13, 15	3D Geometry	Lab 2: ROS
3	Sep 18, 20	Geometric Control	Lab 3: 3D trajectory following
4	Sep 25, 27, 29	Trajectory Optimization	Lab 4: 3D trajectory optimization
5	Oct 2, 4, 6	2D Computer Vision	Lab 5: feature detection
6	Oct 11, 13	2-view Geometry and Minimal Solvers	Lab 6: object localization
7	Oct 16, 18, 20	Non-minimal Solvers and Visual Odometry	Lab 7: GTSAM
8	Oct 23, 25, 27	Place Recognition	Lab 8: ML for robotics
9	Oct 30, Nov 1, 3	SLAM and Visual-Inertial Navigation	Lab 9: SLAM
10	Nov 6, 8	Advanced Topics	Final project
11	Nov 13, 15, 17	Advanced Topics	Final project
12	Nov 20, 22	Advanced Topics	Final project
13	Nov 23-26	Thanksgiving Break	
14	Nov 27, 29, Dec 1, 4 ,6, 11, 13	Guest Lectures and Students Presentations	Final project

Structure of VNAV

2025.11.3

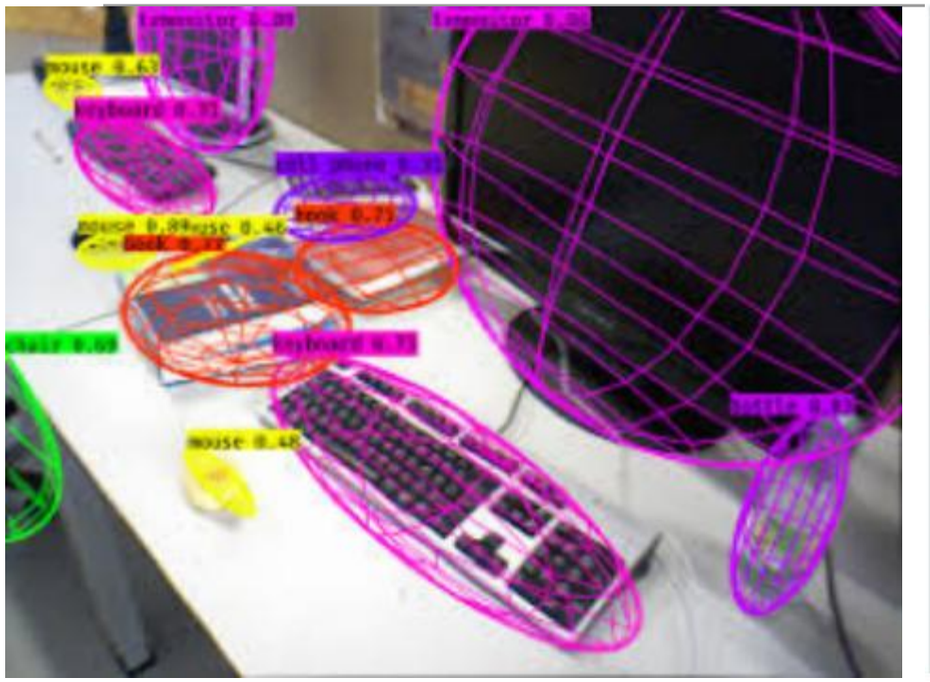
课程内容大纲

- Lectures:
  - M/W/F at 1-2pm EDT (35-225)
  - Mostly on the board
- Labs:
  - W at 3-5pm EDT (typically in 33-116)
  - Code, assignments, OH
- Office Hours:
  - OH: TBA on Piazza
- **Final project:**
  - **Push the boundary of the state of the art**

Week	Dates	Lecture topic	Lab
1	Sep 6, 8	Introduction	Lab 1: Linux, C++, Git
2	Sep 11, 13, 15	3D Geometry	Lab 2: ROS
3	Sep 18, 20	Geometric Control	Lab 3: 3D trajectory following
4	Sep 25, 27, 29	Trajectory Optimization	Lab 4: 3D trajectory optimization
5	Oct 2, 4, 6	2D Computer Vision	Lab 5: feature detection
6	Oct 11, 13	2-view Geometry and Minimal Solvers	Lab 6: object localization
7	Oct 16, 18, 20	Non-minimal Solvers and Visual Odometry	Lab 7: GTSAM
8	Oct 23, 25, 27	Place Recognition	Lab 8: ML for robotics
9	Oct 30, Nov 1, 3	SLAM and Visual-Inertial Navigation	Lab 9: SLAM
10	Nov 6, 8	Advanced Topics	Final project
11	Nov 13, 15, 17	Advanced Topics	Final project
12	Nov 20, 22	Advanced Topics	Final project
13	Nov 23-26	Thanksgiving Break	
14	Nov 27, 29, Dec 1, 4 ,6, 11, 13	Guest Lectures and Students Presentations	Final project

Final Projects (Samples)

课程大作业—举例



Object-based SLAM

Mapping MIT tunnels

Next we formulate the constraints also as quadratic functions of $\tilde{z}$. For $\boldsymbol{R} \in SO(3)$, we have the following three categories of constraints:

- Orthonormal rows: $\boldsymbol{R}\boldsymbol{R}^T = \boldsymbol{I}$, which induces 6 scalar constraints:

$$q_{\tilde{\boldsymbol{P}}_{ij}^r} = \tilde{\boldsymbol{r}}^T \underbrace{\begin{bmatrix} -\delta_{ij} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_3 \otimes ({}_3^+\boldsymbol{e}_{ij}) \end{bmatrix}}_{\tilde{\boldsymbol{P}}_{ij}^r} \tilde{\boldsymbol{r}} = 0$$

$$\forall 1 \le i \le 3;\ i \le j \le 3.$$

- Orthonormal columns: $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}$, which induces 6 scalar constraints:

$$q_{\tilde{\boldsymbol{P}}_{ij}^c} = \tilde{\boldsymbol{r}}^T \underbrace{\begin{bmatrix} -\delta_{ij} & \boldsymbol{0} \\ \boldsymbol{0} & ({}_3^+\boldsymbol{e}_{ij} \otimes \boldsymbol{I}_3) \end{bmatrix}}_{\tilde{\boldsymbol{P}}_{ij}^c} \tilde{\boldsymbol{r}} = 0$$

$$\forall 1 \le i \le 3;\ i \le j \le 3.$$

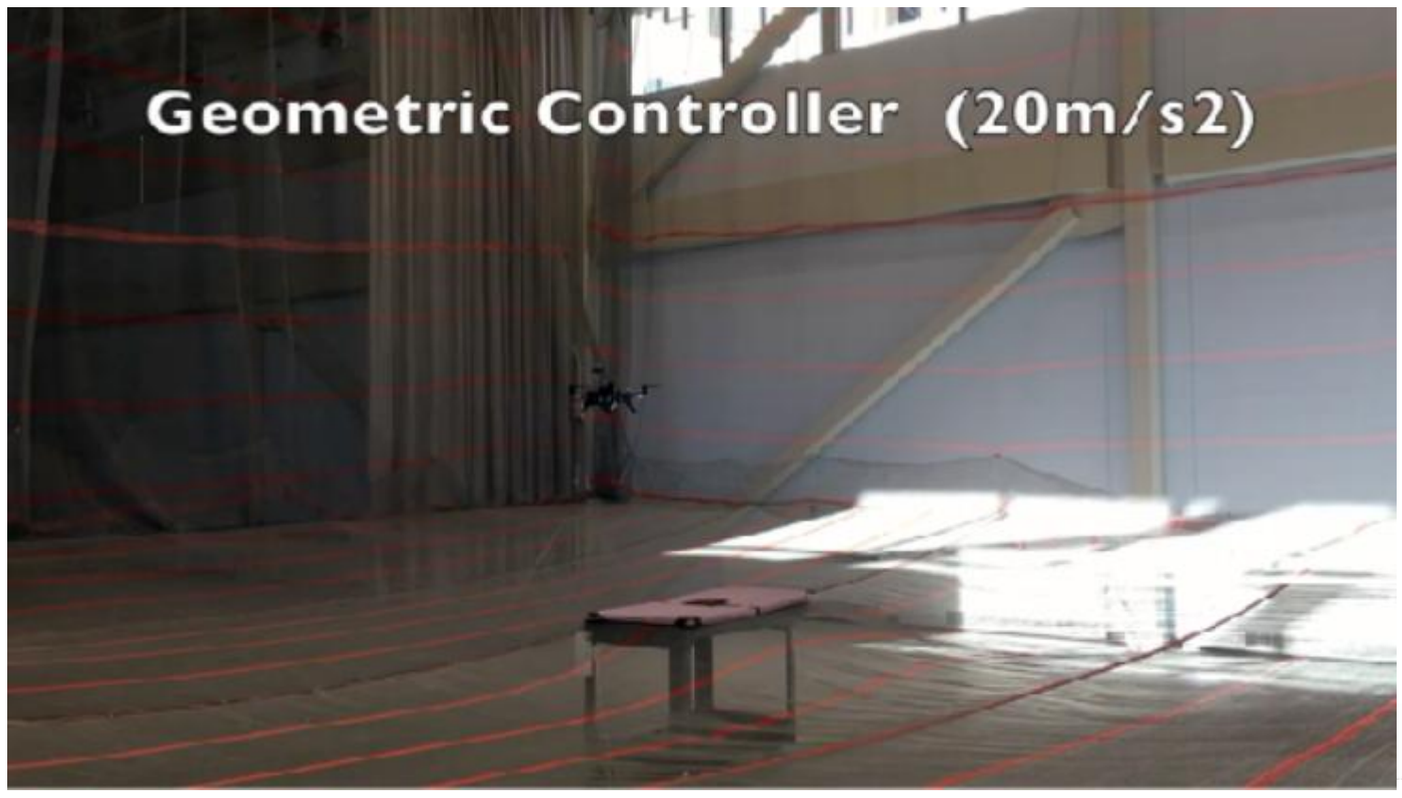
- Right-handedness: $(\boldsymbol{R}\boldsymbol{e}_i) \times (\boldsymbol{R}\boldsymbol{e}_j) = (\boldsymbol{R}\boldsymbol{e}_k)$, which induces 9 scalar constraints:

$$q_{\tilde{\boldsymbol{P}}_{ijk\alpha}^d} = \tilde{\boldsymbol{r}}^T \underbrace{\begin{bmatrix} 0 & \frac{1}{2}(\boldsymbol{e}_k \otimes \boldsymbol{e}_\alpha)^T \\ \frac{1}{2}(\boldsymbol{e}_k \otimes \boldsymbol{e}_\alpha) & ({}_3^-\boldsymbol{e}_{ij}) \otimes [\boldsymbol{e}_\alpha]_\times) \end{bmatrix}}_{\tilde{\boldsymbol{P}}_{ijk\alpha}^d} \tilde{\boldsymbol{r}} = 0$$

$$\forall (i,j,k) \in \{(1,2,3),(2,3,1),(3,1,2)\};\ 1 \le \alpha \le 3$$

The work in [6] has shown that lifting the constraints helps make further SDP relaxation tight. Because our decision

Robust 2-view



outcomes:

- ICRA’19-20, IROS’19, RSS’19, ICCV’19 papers
- “The VNAV class has been a deciding factor for the completion of my thesis.”

2025.11.3

作业 & 分数

- Assignment schedule
  - Lab exercises announced on **Monday**
  - Lab exercises due the following **Monday**
  - Lab handouts include:
    - Setup **instructions**
    - 3-5 **questions**
      - **Theoretical and**
      - **Experimental**
  - Work in teams of 2-3 students but questions may be individual

Week	assignment	Individual/ Team	Announced	Due
1	Lab 1: Linux, C++, Git	I	W Sept 06	W Sept 13
2	Lab 2: ROS	I	W Sept 13	W Sept 20
3	Lab 3: 3D trajectory following	T/I	W Sept 20	W Sep 27
4	Lab 4: 3D trajectory optimization	T/I	W Sep 27	W Oct 4
5	Lab 5: Feature detection	T/I	W Oct 4	W Oct 11
6	Lab 6: Object localization	T/I	W Oct 11	W Oct 18
7	Lab 7: GTSAM	T/I	W Oct 18	W Oct 25
8	Lab 8: Place recognition	T/I	W Oct 25	W Nov 1
9	Lab 9: SLAM	T/I	W Nov 1	W Nov 8
10-14	Final project (report, demo, presentation)	T	W Nov 8	M Dec 4, W Dec 6, M Dec 11, W Dec 13

Assignments & Grading

作业 & 分数

- **课程说明：**机器人集成小组项目1分为夏季和秋季两个部分，秋季主题为“视觉SLAM”，参考MIT课程，该部分占课程总分45%。
- **秋季课程包括8次理论课程，7次实验课程**；7次实验课程完成2023基于ROS1的LAB1-LAB6+LAB9（除LAB7、8）；实验完成时间为2周，晚于2周提交会适当扣分，晚交最高得分为满分的80%；
- **总分构成：**平时出勤分为10分，每个LAB1-LAB3，每个10分，LAB4-LAB6和LAB9为15分，每次实验都会给出评分依据，且在实验课上单独检查后给出分数（仅交实验报告不给分）；
- **实验检查的额外标准（实验分最终分=实验完成度*下述完成分比例）：**实验完成程度（作业展示+实验检查结果）30%’ +实验过程分解（各步骤思路）30% ‘+个人思考（总结分析）35%’ +格式5% ‘
- **实验报告应包含个人思考总结**，整体重复比例小于85%，高于85%默认为抄袭。最后，课程过程中遇到问题，咱们再灵活调整，解决问题。

Assignments & Grading

ICRA what?

ICRA: International Conference on Robotics and Automation

ROBOTICS SCIENCE AND SYSTEMS

RSS

IEEE/RSJ International Conference on Intelligent Robots and Systems

IROS

2025.11.3

Submission deadlines	
ICRA 2025	Sept. 15, 2024
RSS 2025	Feb. 1, 2025 (?)
IROS 2025	Mar. 1, 2025 (?)
ICRA 2026	Sep. 15, 2025 (?)

作业 & 分数 Assignments & Grading

- Syllabus: https://tinyurl.com/VNAV2023-syllabus
- All assignments are announced on Canvas: https://tinyurl.com/VNAV2023-canvas (lab handouts: https://vnav.mit.edu/)

- **[90%]** Lab exercises and code / **LAB系列实验**
- **[5%]** Participation and TA/LA evaluation
- **[5%]** Team members' assessment

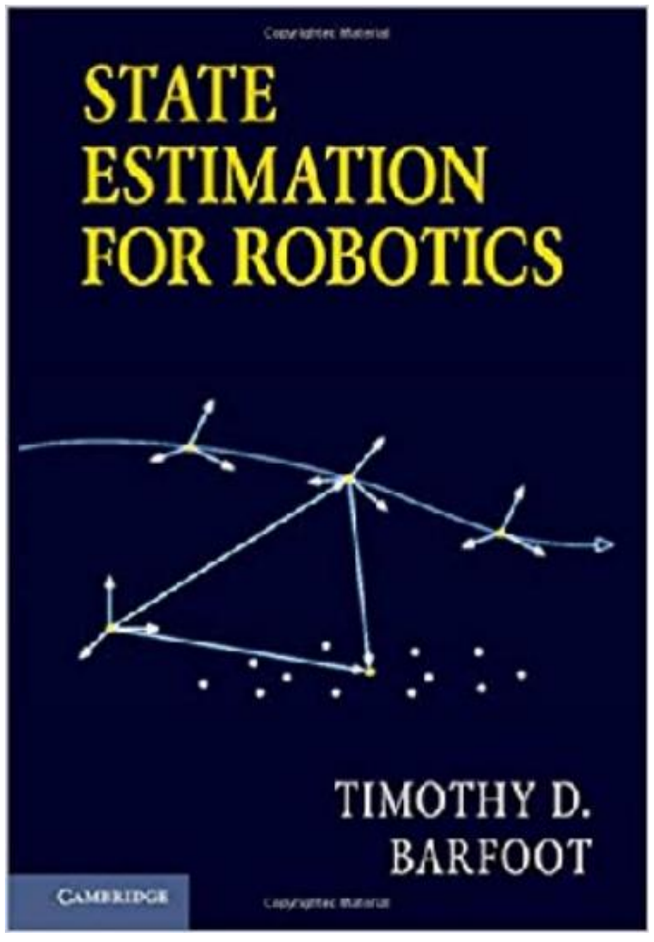
- **晚交作业扣分公式:**

  grade_late_submission = grade * max(1 - 0.15 * days_late, 0)

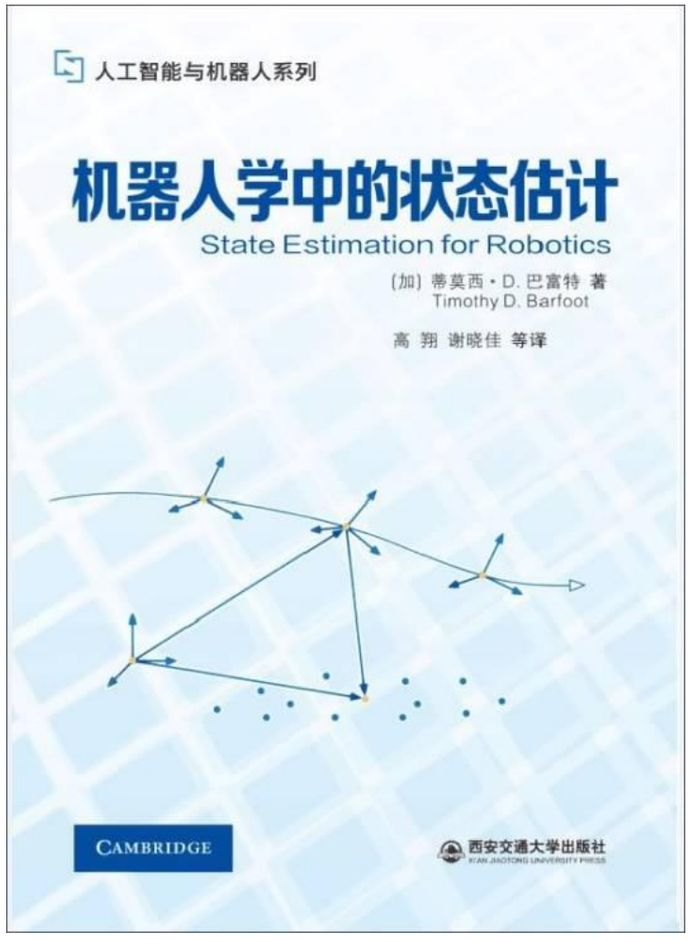
参考书籍

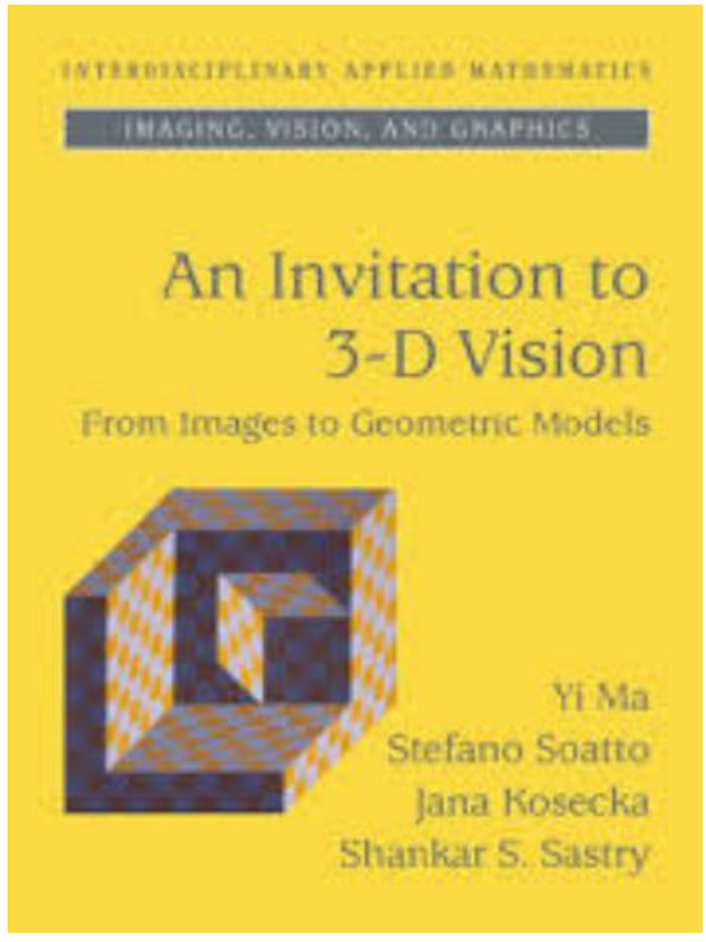
- Both textbooks are recommended, but not required
- Both available online

Textbooks



中文翻译版本





- Specific pointers to chapters in these books and other resources (papers) will be provided in each lecture

课程要求 / 先修课程（MIT）

- Requirements satisfied by VNAV:
  - 12 units: 3 - 2 - 7
  - Field Evaluation Subject in Course 16
- Prerequisites:
  - Programming and algorithms (16.35 or similar)
    程序与算法
    - Familiarity with coding and **C++**
  - Optimal estimation and control (16.32 or similar)
    优化估计与控制
    - PID, Kalman Filtering, …
  - Linear algebra (18.06 or similar)
    线性代数
- Good to have:
  - Independent experience (UROPs, competitions, etc.)
  - Some background in optimization

[online questionnaire distributed at the end of this lecture]

Lecture Outline / 大纲

- **课程基础（Logistics)**
  - VNAV 课程目的
  - 课程内容大纲
  - 作业 & 分数
  - 课程要求

- **Visual Navigation for Autonomous Vehicles （VNAV)**
  - 机器人革命 / The robot revolution.
  - 感知的重要性 / Why perception?
  - More on the VNAV

机器人革命：自动驾驶

- 【Telsa】特斯拉自动驾驶宣传片_哔哩哔哩_bilibili
- 小米自动驾驶技术_哔哩哔哩_bilibili

2025.11.3

机器人革命: 无人机Drones

https://www.youtube.com/watch?v=gsfkGlSajHQ

Skydio Autonomous Drones for Public Safety, Defense, and Energy | Skydio

地下城市挑战赛

The DARPA **Subt**erranean Challenge

拿下200万美元奖金，DARPA地下挑战赛冠军机器人长啥样？轮式、腿式、飞行，多形态齐上阵！ - 知乎 (zhihu.com)

【DARPA】地下城市挑战赛（Subterranean Challenge Urban ）速览_哔哩哔哩_bilibili

CERBERUS: Deployment at the DARPA Subterranean Challenge Urban Circuit_哔哩哔哩_bilibili

MIT AEROASTRO

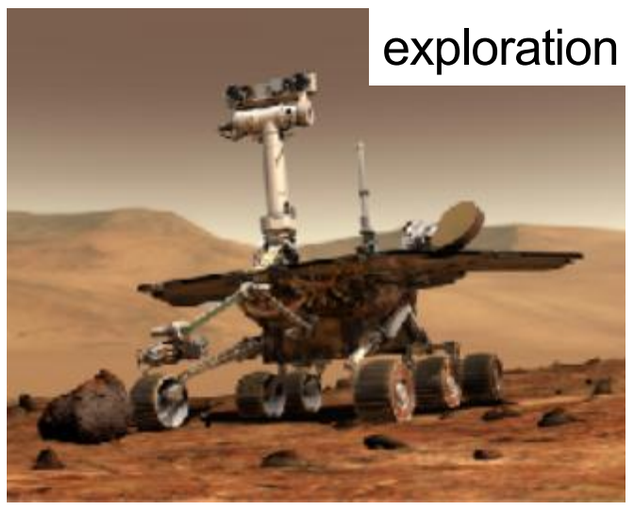
机器人革命 / The Robot Revolution

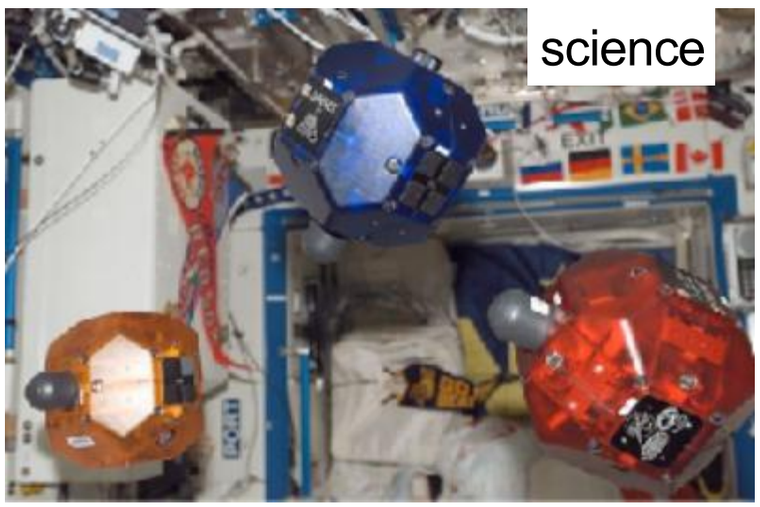
ground

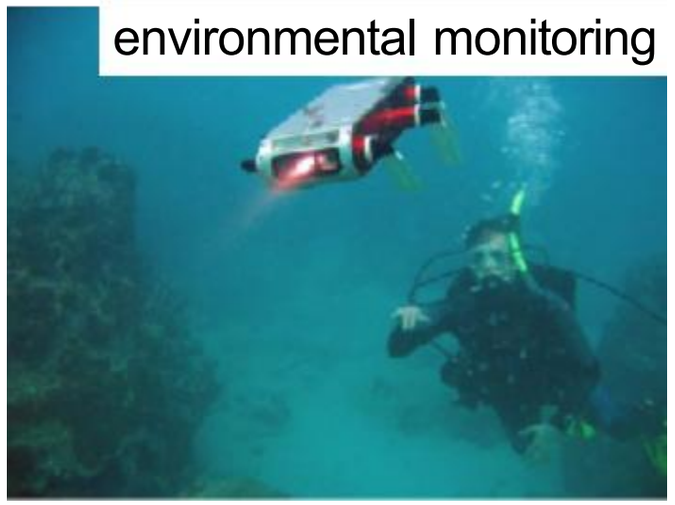
air

and more ...

space







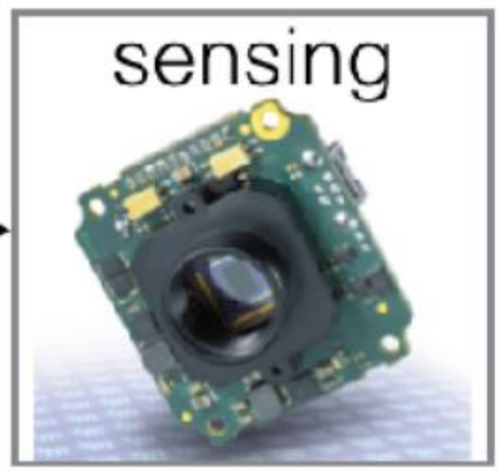
2025.11.3 **reasons for adoption**: faster, better, safer, cheaper, access

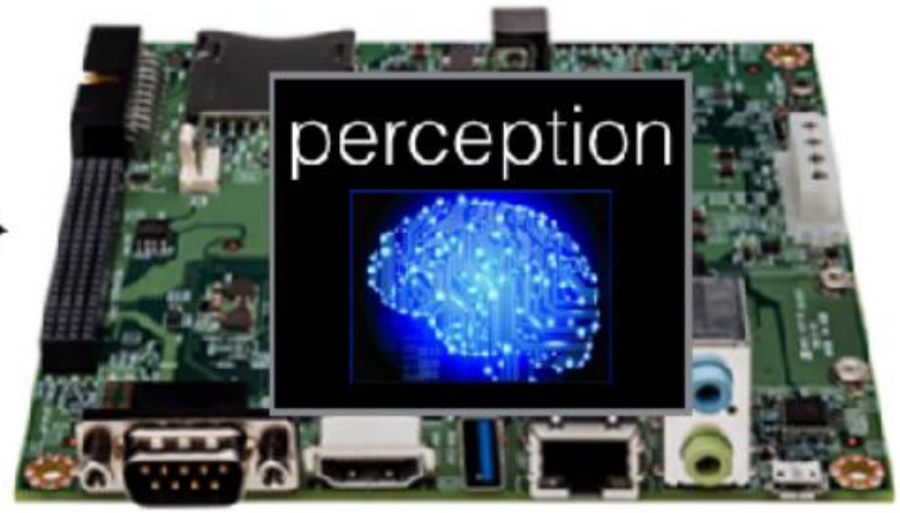
机器人感知

Robot Perception

real world

world model / representation

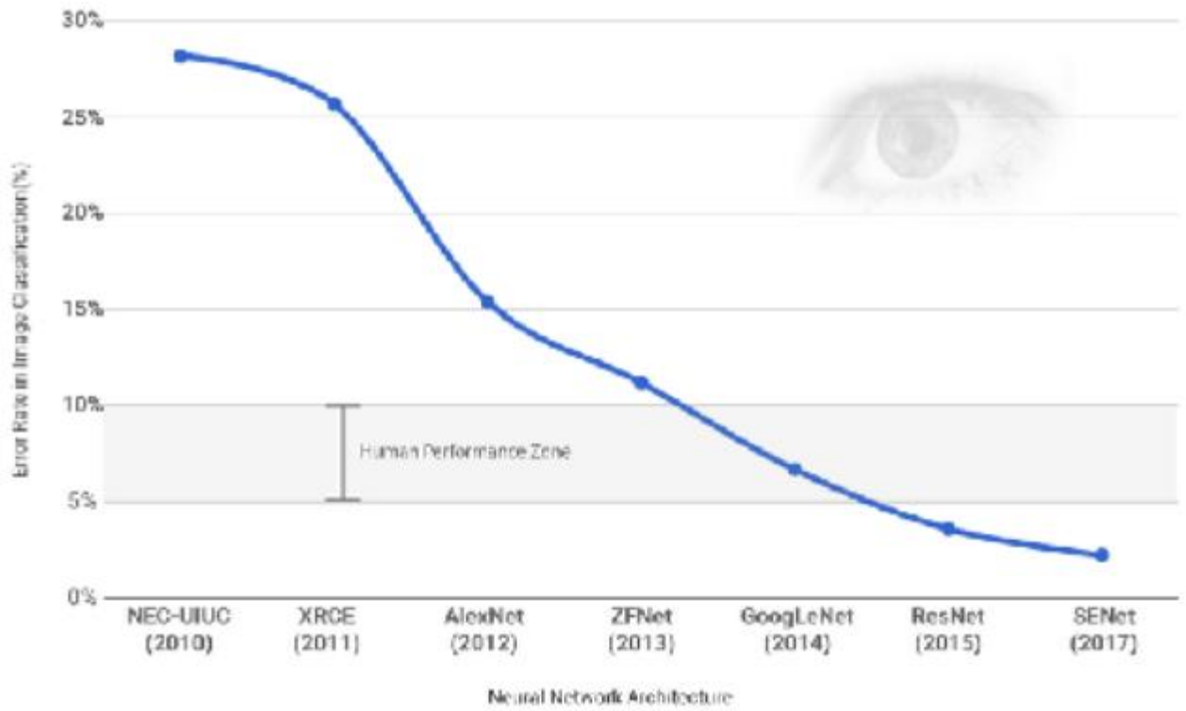




感知的重要性?

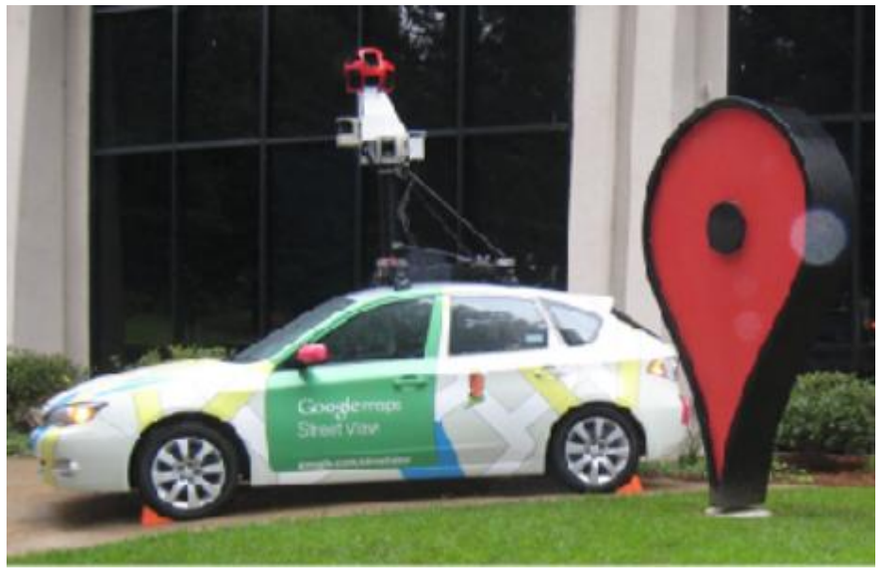
Why Perception?

Things are starting to work!



Google Street View

2025.11.3



Oculus Rift Goggles

感知的重要性！

- 骗过车机！纯视觉的特斯拉，眼神真的不太好！我自己是不敢用AP了_哔哩哔哩_bilibili

Perception success.. and its failures

Carefully designed graffiti and art stickers cause a neural network to misclassify stop signs as speed limit 45 signs or yield signs.

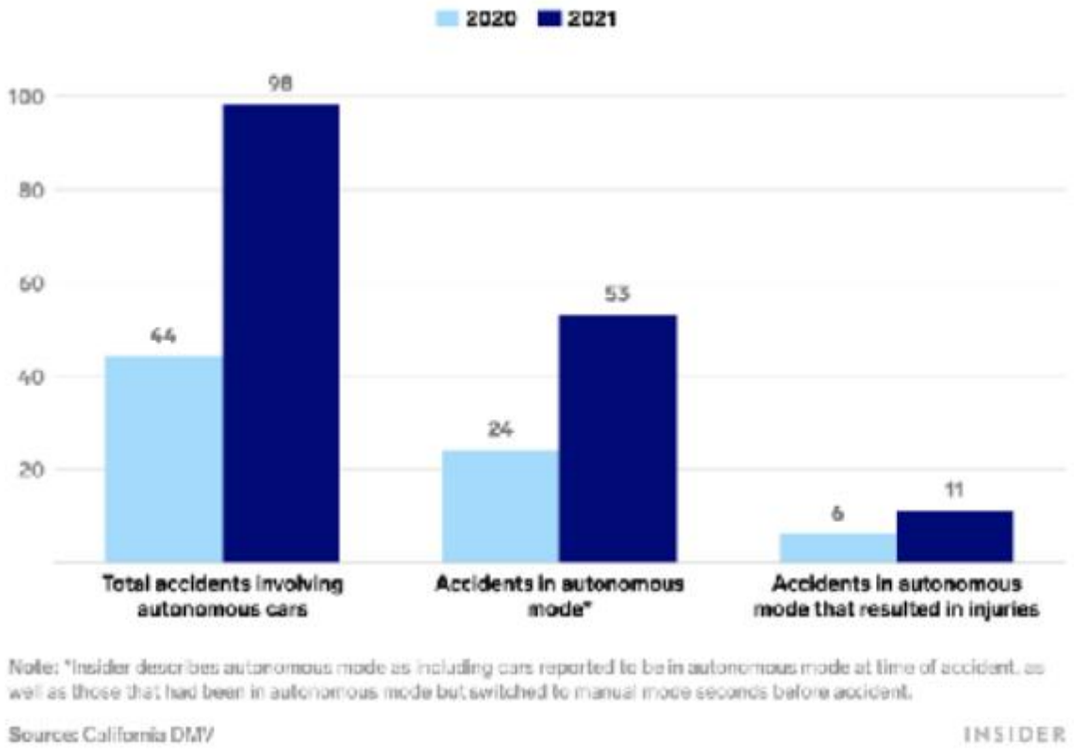
“Google employs a small army of human operators to manually check and correct the maps”
[Wired]

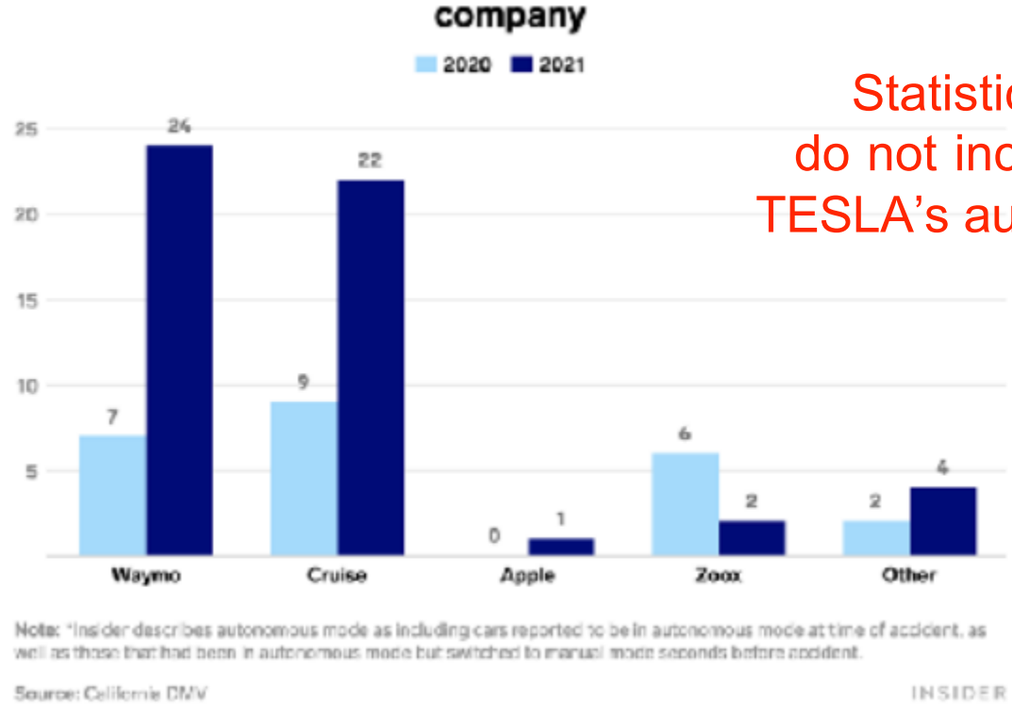
未来的路还很长。。。

Robust perception is crucial in safety critical applications

Waymo and Cruise self-driving cars took over San Francisco streets at record levels in 2021 — so did collisions with other cars, scooters, and bikes

自动驾驶碰撞





Statistics
do not include
TESLA's autopilot

2025.11.3

[https://www.businessinsider.com/self-driving-car-accidents-waymo-cruise-tesla-zoox-san-francisco-2022-1]

2 robotaxi crashes in San Francisco put focus on autonomous vehicle safety

A Cruise driverless taxi collided with a San Francisco Fire Department engine.

KPIX

aAugust18,2023/9.05p

"The fire engine was operating in Code 3 emergency mode, with lights and sirens. It's really a reminder to everyone you are required to yield whether it's a vehicle driven by a human operator or an autonomous vehicle," said SFPD spokesperson Sgt. Kathryn Winters.

It was unclear why the self-driving car did not yield. Cruise said in a statement it's investigating to better understand the problem.

About two hours later, another Cruise driverless car was struck by Dodge Charger in the Mission District.

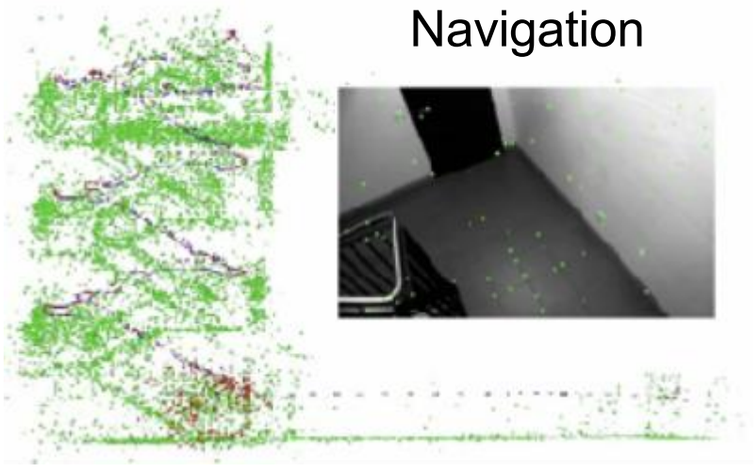
Why Perception?

Sensing Perception Autonomy and Robot Kinetics

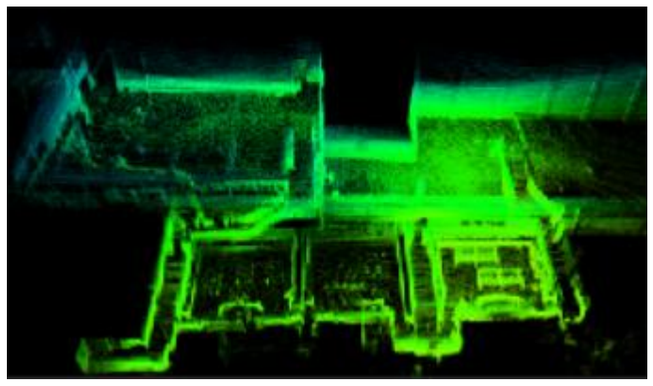
感知自主性和机器人动力学

Robust Perception, Localization and Mapping

Visual-inertial Navigation



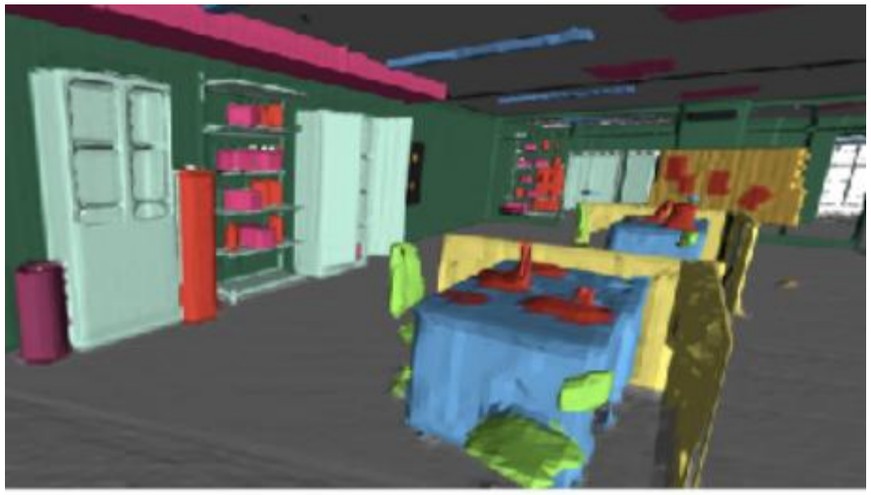
Lidar-based SLAM

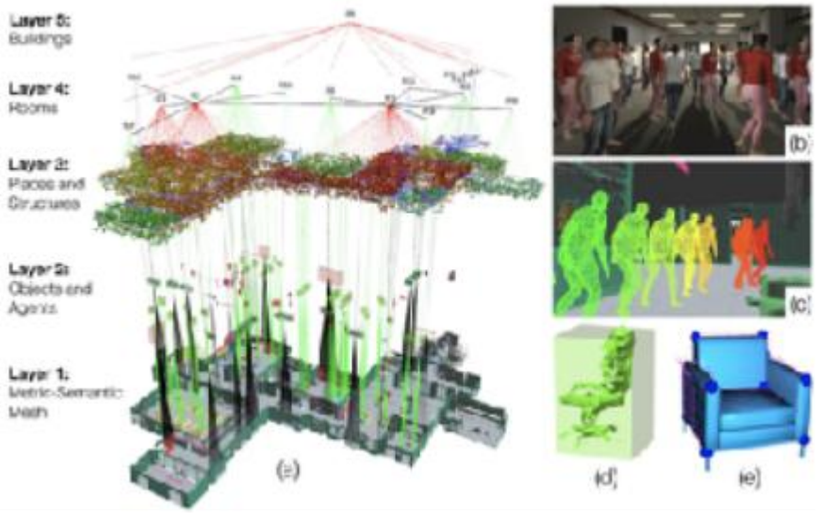


Certifiable Algorithms

High-level Scene Understanding (Spatial AI)

Kimera: Metrics-semantic SLAM

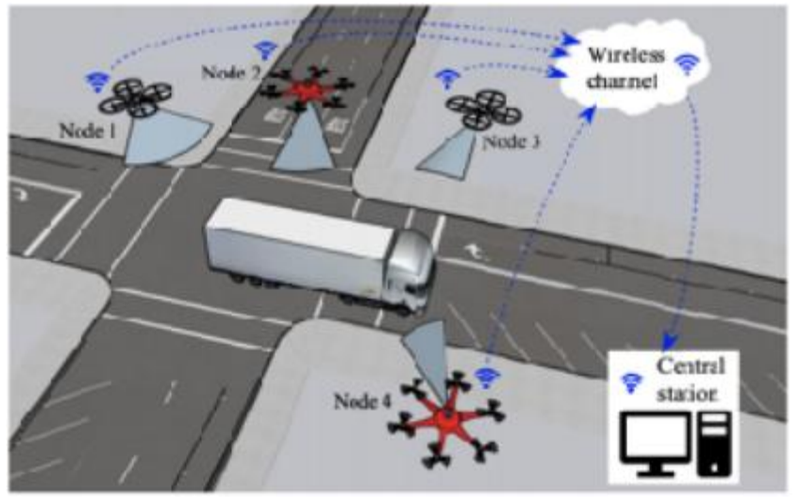




3D Scene Graphs

Co-design

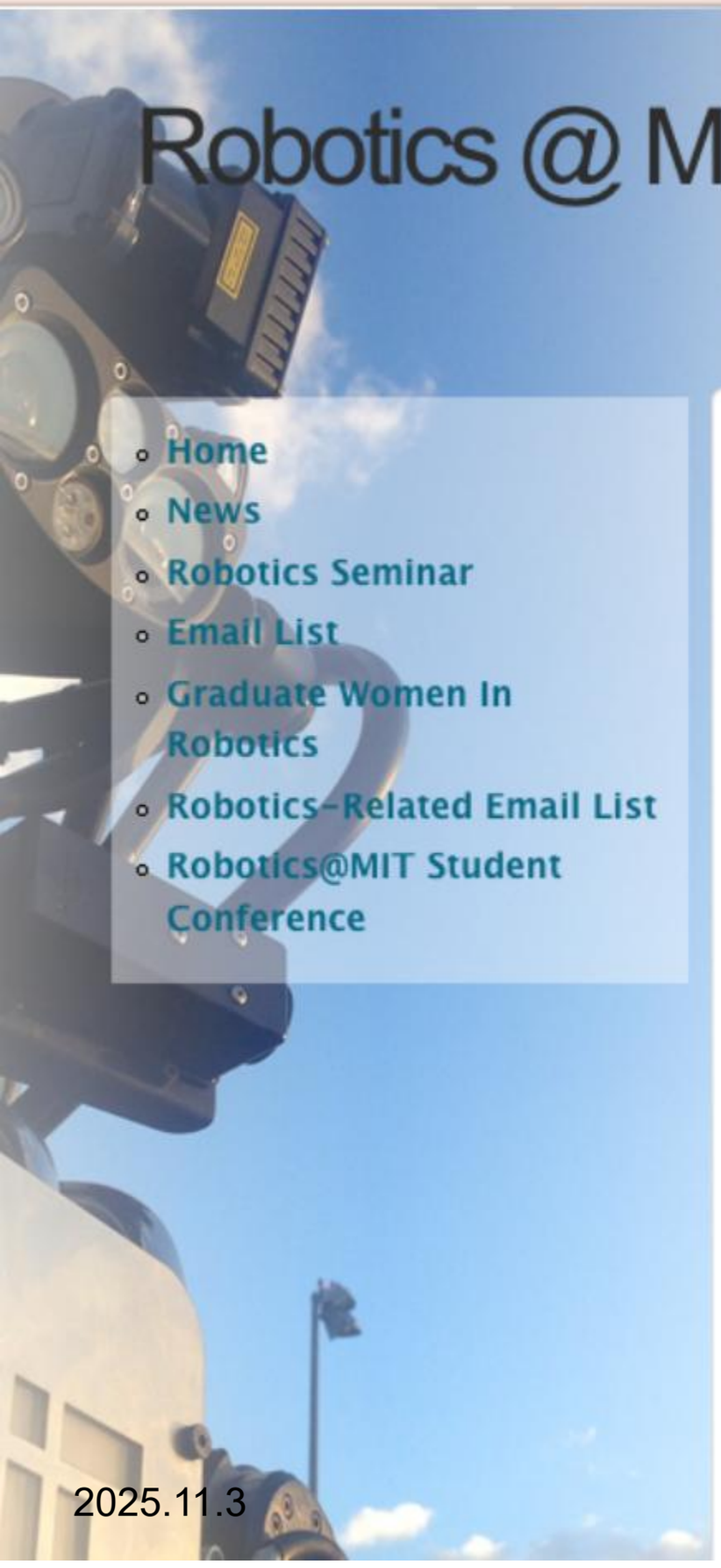
- Computation-communication co-design
- Control and sensing co-design



Soft Drones and Soft Aerial Manipulation

https://robotics.mit.edu

Robotics Groups at MIT (listed by Principal Investigator) | Robotics @ MIT



Robotics Groups at MIT (listed by Principal Investigator)

Perceptual Science Group	Ted Adelson
Improbable AI	Pulkit Agrawal
D'Arbeloff Lab: Robotics	Harry Asada
Lab website coming soon!	Navid Azizan
Center for Marine Robotics (at Woods Hole)	James Bellingham
Personal Robotics Group	Cynthia Breazeal
The SPARK (Sensing, Perception, Autonomy, and Robot Kinetics) Lab	Luca Carlone
Soft and Micro Robotics Laboratory	Kevin Chen
Control Networks Group	Domitilla Del Vecchio
Realm: REliable Autonomous systems Lab	Chuchu Fan

2025.11.3

Experimental course

Lab 1 at 3pm today in C108

Thank you for the attention!

2025.11.3

MIT AEROASTRO