

# **16.485: VNAV** - Visual Navigation for Autonomous Vehicles

Luca Carlone

Lecture 1





#### -Instructor-



Luca Carlone



Rajat Talak

#### Guest Lecturers-



Jared Strader



Yulun Tian

#### Teaching Assistants.



Jingnan Shi



Nathan Hughes



Dominic Maggio

Yulun Tian

### Lecture Outline

#### Logistics

- Goals of VNAV
- Structure
- Assignments & Grading
- Requirements & Pre-reqs

#### Visual Navigation for Autonomous Vehicles

- The robot revolution
- Why perception?
- More on the VNAV Staff



### Goals of VNAV (1/3)

Learn the state of the art in robot perception (2D computer vision, localization, mapping, object pose estimation, semantic understanding,...) for visual navigation



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

### Goals of VNAV (2/3)

### • Mens / Theory:



- Learn/develop/exercise theoretical tools necessary for robotics research (geometry, optimization, ...)
- Learn about state-of-the-art **algorithms** for robot perception (+ selected topics in control, trajectory optimization, learning)
- Get an overview of **open problems** in robot perception
- Note 1: If you don't like writing (some) math, you may not like this class
- Note 2: this course is much more theoretical than RSS (6.4200/16.405)
- 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 stateof-the-art research in robotics (not necessarily in robot perception)

# Goals of VNAV (3/3)

- Manus / Practice:
  - Learn/practice ROS (Robot Operating System)



- Possibility to use **physical robots** for final projects
- Learn limitations of state-of-the-art implementations

- Note: If you do not like coding (in C++), you may not like this class
   -> C++ refresher: <u>https://youtu.be/F\_vIB3yjxaM</u>
- Overarching goal: prepare you (or refine your skills) to perform state-ofthe-art research in robotics (not necessarily in robot perception)



### Practice on Real Robots: Intel Aero Drone





© Intel Corporation. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/



### Practice in Realistic Unity-based Simulator





#### • 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

#### • 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

#### • 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

#### • 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

#### • 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)



**Object-based SLAM** 

#### Mapping MIT tunnels

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

Orthonormal rows: RR<sup>T</sup> = I, which induces 6 scalar constraints:

$$q_{\tilde{P}_{ij}^r} = \tilde{r}^T \underbrace{\begin{bmatrix} -\delta_{ij} & \mathbf{0} \\ \mathbf{0} & I_3 \otimes (\frac{+}{3} e_{ij}) \end{bmatrix}}_{\tilde{P}_{ij}^r} \tilde{r} = 0$$
  
 $\forall 1 \le i \le 3; \ i \le j \le 3.$ 

Orthonormal columns: R<sup>T</sup> R = I, which induces 6 scalar constraints:

$$\begin{split} q_{\tilde{P}_{ij}^e} &= \tilde{r}^T \underbrace{ \begin{bmatrix} -\delta_{ij} & \mathbf{0} \\ \mathbf{0} & (_3^+ \boldsymbol{e}_{ij} \otimes \boldsymbol{I}_3) \end{bmatrix}}_{\tilde{P}_{ij}^e} \tilde{r} = 0 \\ \forall 1 \leq i \leq 3; \ i \leq j \leq 3. \end{split}$$

 Right-handedness: (Re<sub>i</sub>) × (Re<sub>j</sub>) = (Re<sub>k</sub>), which induces 9 scalar constraints:

$$q_{\tilde{P}_{ijk\alpha}^d} = \tilde{r}^T \underbrace{\begin{bmatrix} 0 & \frac{1}{2} (\boldsymbol{e}_k \otimes \boldsymbol{e}_\alpha)^T \\ \frac{1}{2} (\boldsymbol{e}_k \otimes \boldsymbol{e}_\alpha) & (\frac{1}{3} \boldsymbol{e}_{ij}) \otimes [\boldsymbol{e}_\alpha]_{\times} ) \\ \tilde{P}_{ijk\alpha}^d}_{\tilde{P}_{ijk\alpha}^d} \tilde{r} = 0$$

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

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."

Drone racing

# Assignments & Grading

- Assignment schedule
  - Lab exercises announced on Wednesdays
  - Lab exercises due the following Wednesday
  - 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	<b>T/</b>	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)	т	W Nov B	M Dec 4, W Dec 6, M Dec 11, W Dec 13



# Assignments & Grading

- no final exam but final project
  - Final demo, showcasing the project outcome
  - Team presentation, including a video abstract of the project
  - Technical report (formatted using ICRA guidelines)

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/l	W Sept 20	W Sep 27
4	Lab 4: 3D trajectory optimization	Τ/Ι	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)	т	W Nov B	M Dec 4, W Dec 6, M Dec 11, W Dec 13



### ICRA what?

ICRA: International Conference on Robotics and Automation





IEEE/RSJ International Conference on Intelligent Robots and Systems

IROS

#### Submission deadlines

ICRA 2024	Sept. 15, 2023
RSS 2024	Feb. 1, 2023 (?)
IROS 2024	Mar. 1, 2023 (?)
ICRA 2025	Sep. 15, 2024 (?)

### Assignments & Grading

- Syllabus: <a href="https://tinyurl.com/VNAV2023-syllabus">https://tinyurl.com/VNAV2023-syllabus</a>
- All assignments are announced on Canvas: <u>https://tinyurl.com/VNAV2023-canvas</u> (lab handouts: <u>https://vnav.mit.edu/</u>)
- [65%] Lab exercises and code
- [25%] Final project: demo, presentation, technical report
- [ 5%] Participation and TA/LA evaluation
- [ 5%] Team members' assessment
- Penalty on late submissions: grade\_late\_submission = grade \* max(1 - 0.15 \* days\_late, 0)
- Q&A with TAs and Instructors on Piazza: <u>https://tinyurl.com/VNAV2023-piazza</u>



### Textbooks

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



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



### Requirements & Pre-reqs

- Requirements satisfied by VNAV:
  - 12 units: 3 2 7
  - Field Evaluation Subject in Course 16
- Prerequisites:
  - Programming and algorithms (16.35 or similar)
    - Familiarity with coding and C++
  - Optimal estimation and control (16.32 or similar)
    - PID, Kalman Filtering, ...
  - Linear algebra (18.06 or similar)
- Good to have:
  - Independent experience (UROPs, competitions, etc.)
  - Some background in optimization
  - [online questionnaire distributed at the end of this lecture]



### Lecture Outline

#### Logistics

- Goals of VNAV
- Structure
- Assignments & Grading
- Requirements & Pre-reqs

#### Visual Navigation for Autonomous Vehicles

- The robot revolution
- Why perception?
- More on the VNAV Staff



### The Robot Revolution: Self-driving Cars



#### Oct, 2016

https://www.tesla.com/videos/full-self-driving-hardware-all-tesla-cars

O Naser Dadgostar. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/

https://www.cnn.com/2023/01/17/tech/ tesla-self-driving-video-staged/ index.html#:~:text=A%202016%20vide o%20that%20Tesla,testimony%20by%2 0a%20senior%20engineer.

### The Robot Revolution: Drones

### Introducing Skydio R1





© Skydio. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/

### The DARPA **Subt**erranean Challenge





### The Robot Revolution





medical applications



environmental monitoring



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

### **Robot Perception**

#### real world



#### world model / representation



© source unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/fag-fair-use/







### Why Perception?



2.1. COCO Detection Challenge



3. Places Challenges





Scene Parsing

Instance Segmentation



### Things are starting to work!



@ Google. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/

Google Street View





Oculus Rift Goggles

### Why Perception?

Perception success.. and its failures





Images: Evtimov et al © IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.ecu/help/faq-fair-use/

Carnouflage 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]

# A long road ahead

### 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





Note: "Insider describes autonomous mode as including cars reported to be in autonomous mode at time of accident, as well as those that had been in autonomous mode but switched to manual mode seconds before accident.

Source: California DMV



Note: "Insider describes autonomous mode as including cars reported to be in autonomous mode at time of accident, as well as these that had been in autonomous mode but switched to manual mode seconds before accident.

Car accidents in autonomous mode\* in California, by

Source: California DMV

INSIDER

29

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

INSIDER

#### Autonomous car collisions in California

### 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

August 18, 2023 / 9:05 PM

"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.

https://www.cbsnews.com/sanfrancisco/news/robotaxi-crashes-san-francisco-focus-autonomous-vehicle-safety/

### Why Perception?





### Why Perception?



[computer animation, courtesy of Lexus]





### Sensing Perception Autonomy and **R**obot **K**inetics

#### **Robust Perception**, **Localization and Mapping**







**Certifiable Algorithms** 

#### **High-level Scene Understanding (Spatial AI)**

Kimera: Metrics-semantic SLAM







### **Co-design**

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





Soft Drones and Soft Aerial Manipulation

**3D Scene Graphs** 



Jingnan Shi jnshi@mit.edu

- AeroAstro MS, 3rd year PhD

- Undergrad at Harvey Mudd College

- Likes to cook

- Have a shiba inu



#### Current Research

#### **Research Question:**

Can we implement object perception that is general, efficient and robust?



### **Dominic Maggio**

3rd year grad student

- M.S. Course 16, 2023
- B.S. (16 & 6-2), 2021



VNAV final project (2021)

 Visual Inertial Navigation at night using Infrared flashlights mounted to a drone

**Outside School** 

- Hometown: Olive Branch, Louisiana
- Building Rockets, RC Planes, Boats

#### Research

- High-altitude visual localization
- Using neural maps for visual localization











### Nathan Hughes

Research Goal: Actionable representations designed for robot interaction



2nd Year PhD Student (AeroAstro)



**Undergraduate Education:** WPI '16 (RBE / CS) **Research Focus:** 3D Scene Graphs

- Creation from visual-inertial data (Hydra)
- Richer scene representations
   R(3) R(4)



Figure: 3D Scene Graph Created by Hydra



Figure: Dory (my dog)



#### Yulun Tian

- Who am I: Postdoc at MIT AeroAstro working with Prof. Jon How
- **Research interests**: Perception, Optimization, Multi-Robot Systems







Optimization on manifolds



Spectral graph theory





Multi-Robot Distributed SLAM

### Robotics @ MIT

- Home
- News
- Robotics Seminar
- Email List
- Graduate Women In Robotics
- Robotics-Related Email List
- Robotics@MIT Student Conference



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

### Admission questionnaire

Admission questionnaire at following link:

http://tinyurl.com/VNAV2023-survey

#### Please fill in and submit by 5:00pm today

Lab 1 at 3pm today in 33-116

Thank you for the attention!

