Robotics and Autonomous Vehicles

John J. Leonard

Samuel C. Collins Professor of Mechanical and Ocean Engineering

Computer Science and Artificial Intelligence Laboratory (CSAIL)
Massachusetts Institute of Technology

PROFESSIONAL EDUCATION

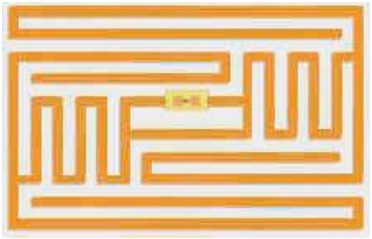
CSAIL
MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

OUTLINE

- Robots and the Internet of Things
- Sensing and data processing
- Simultaneous localization and mapping
- Self-driving vehicles
- Future research challenges

ROBOTS AND THE INTERNET OF THINGS

ROBOTS AND THE INTERNET OF THINGS

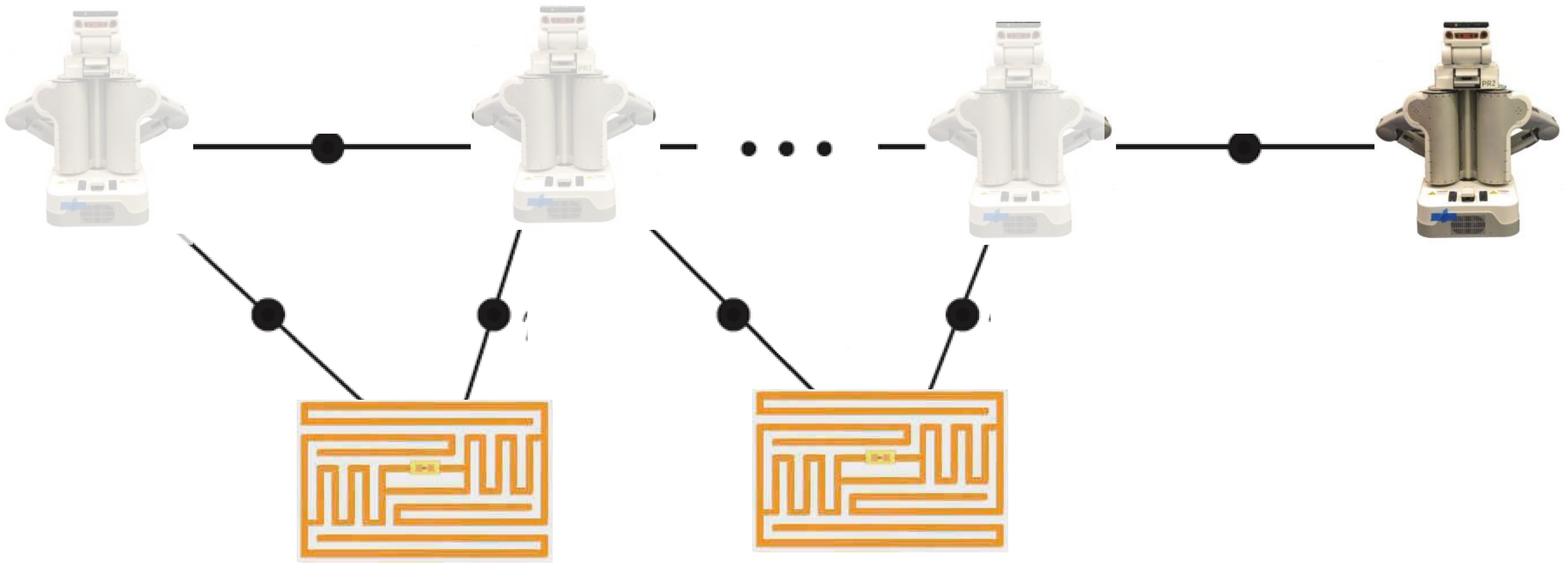


RFID

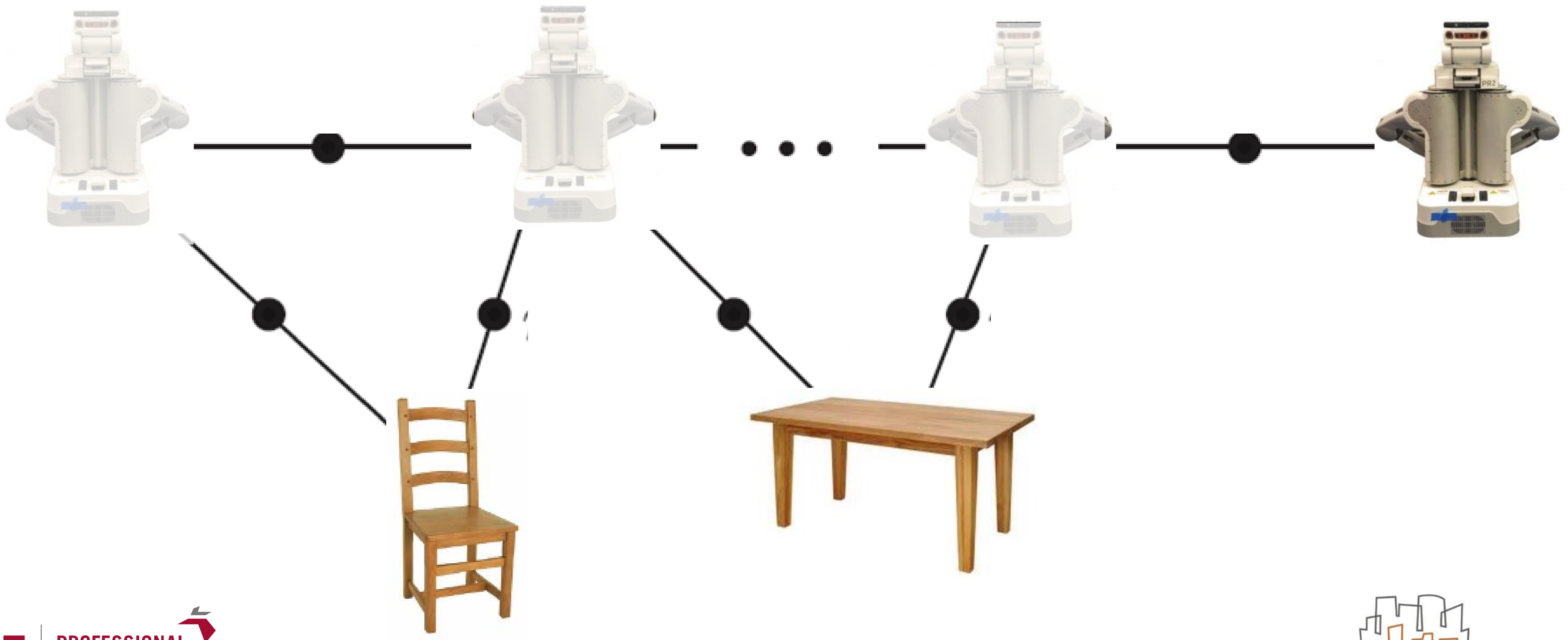
Objects

- Vision: What would it take to realize the equivalent of Google for the physical world?
- Ability to build and maintain an indexed database for entities in the world
- Rapidly answer queries for “what and who is where?” for the physical world
- Maintain this information as a function of time and place, as the world changes

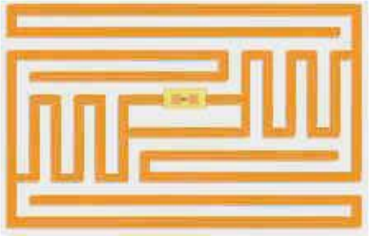
ROBOTS AND THE INTERNET OF THINGS



ROBOTS AND THE INTERNET OF THINGS



ROBOTS AND THE INTERNET OF THINGS



RFID

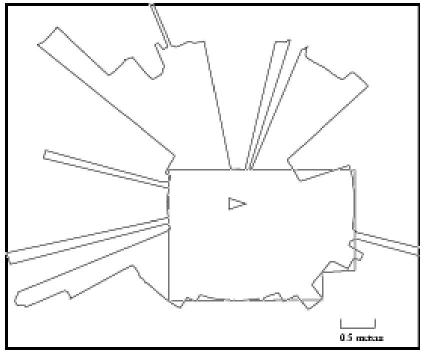
Objects

Questions:

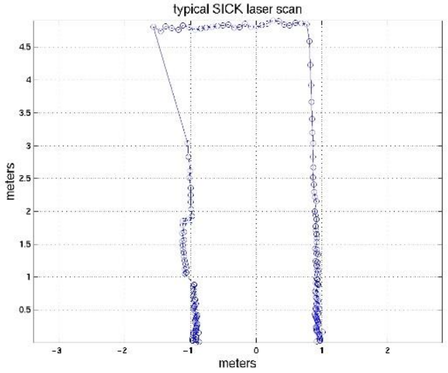
- What are the sensors available?
- What is the right map representation?
- How do we perform inference?
- How do we solve for data association?
- How do we build robust autonomous systems?
- How do we achieve long-term autonomy?

SENSING AND DATA PROCESSING

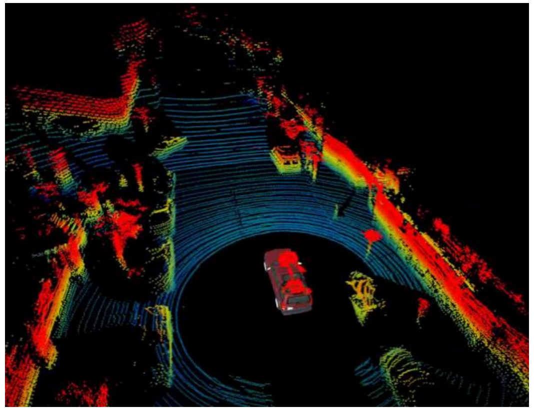
Sonar



2-D Lidar

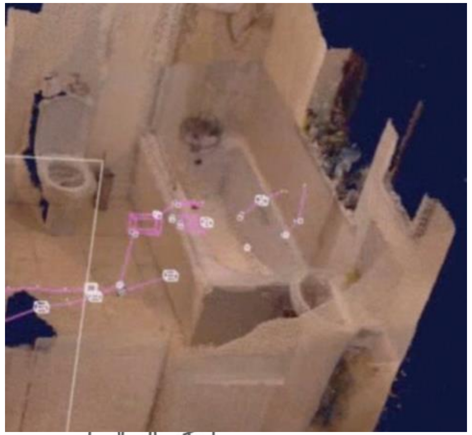


3-D Lidar



Vision

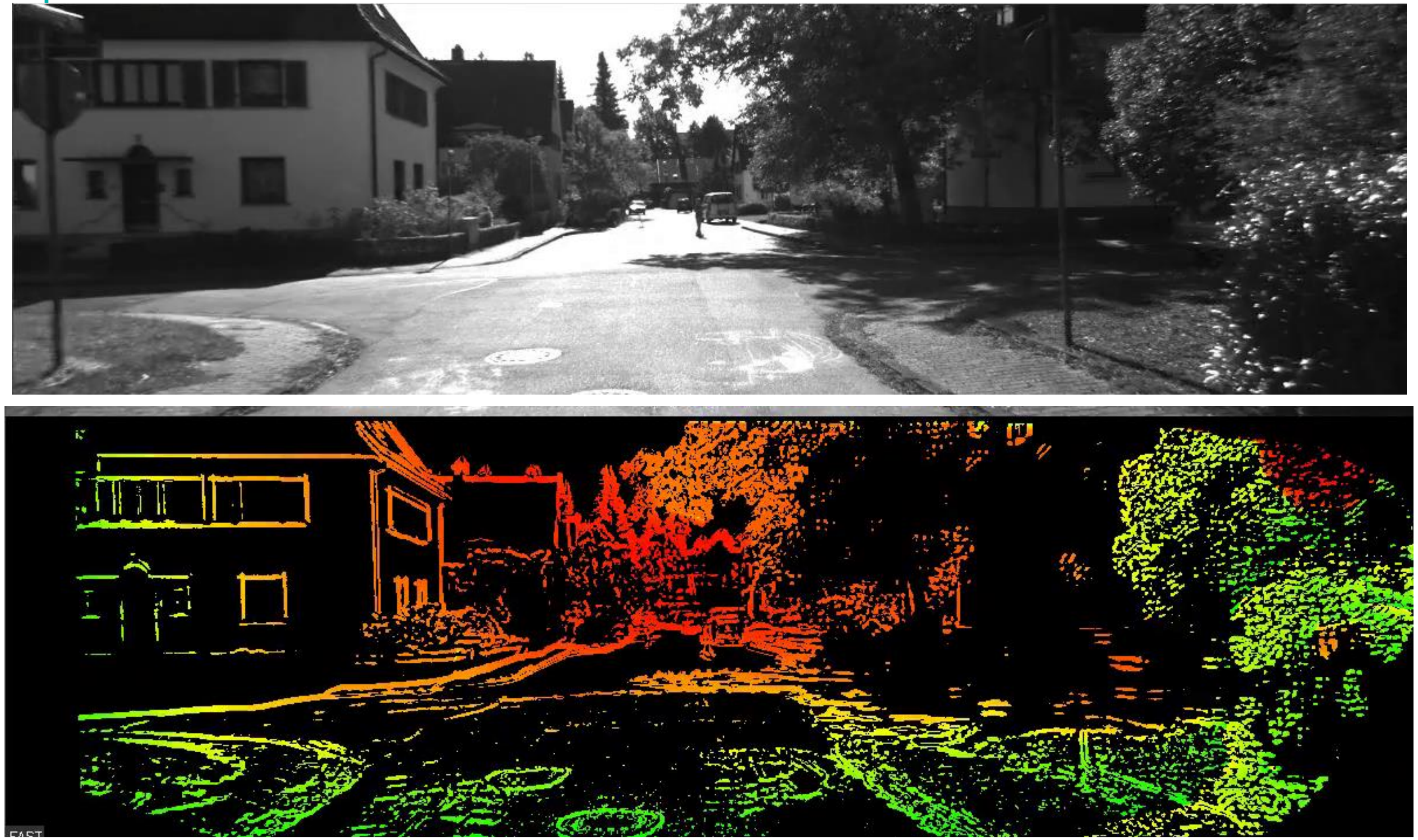
RGB-D



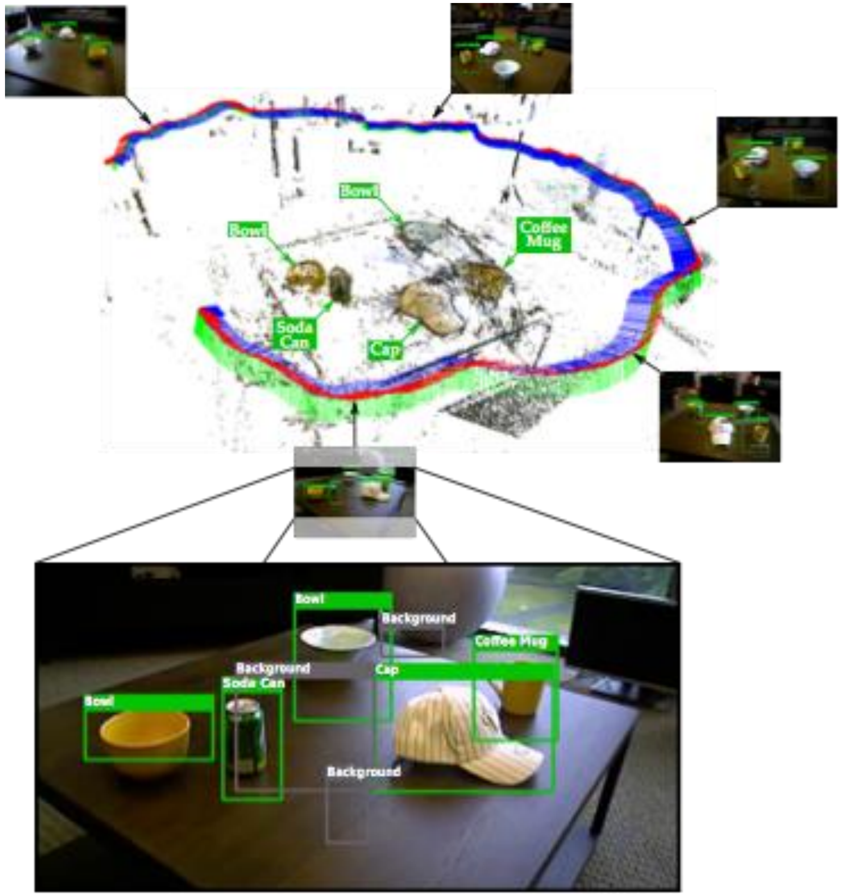
PROFESSIONAL EDUCATION

CSAIL
MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

COMPUTER VISION



Dense Stereo, Sudeep Pillai (MIT)

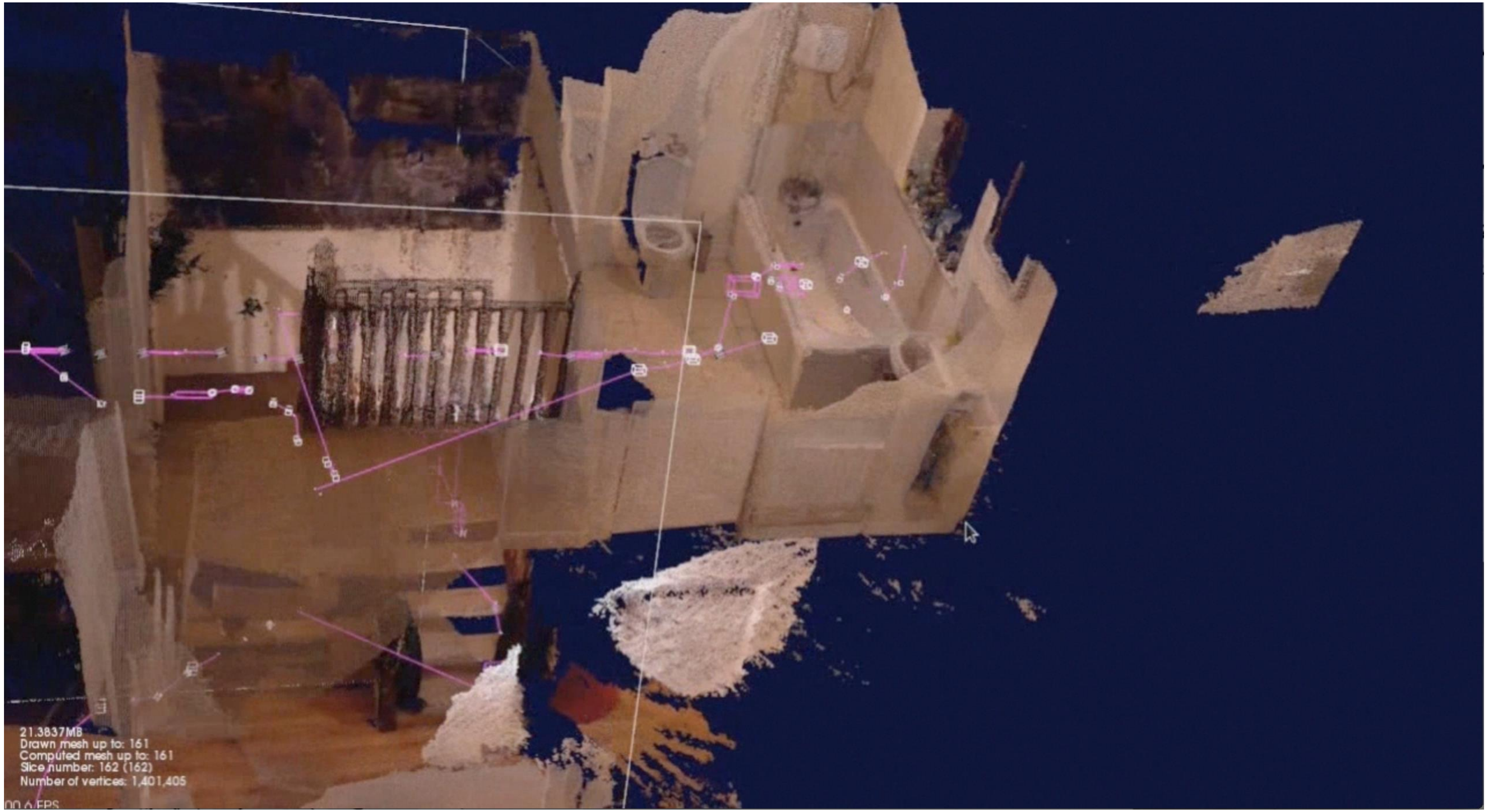


Location-aware object recognition, Sudeep Pillai (MIT)

DENSE RGB-D PROCESSING USING GPUS

Kinect

GPU



Whelan, National University of Ireland, Maynooth

VISUAL-INERTIAL NAVIGATION

Goal:

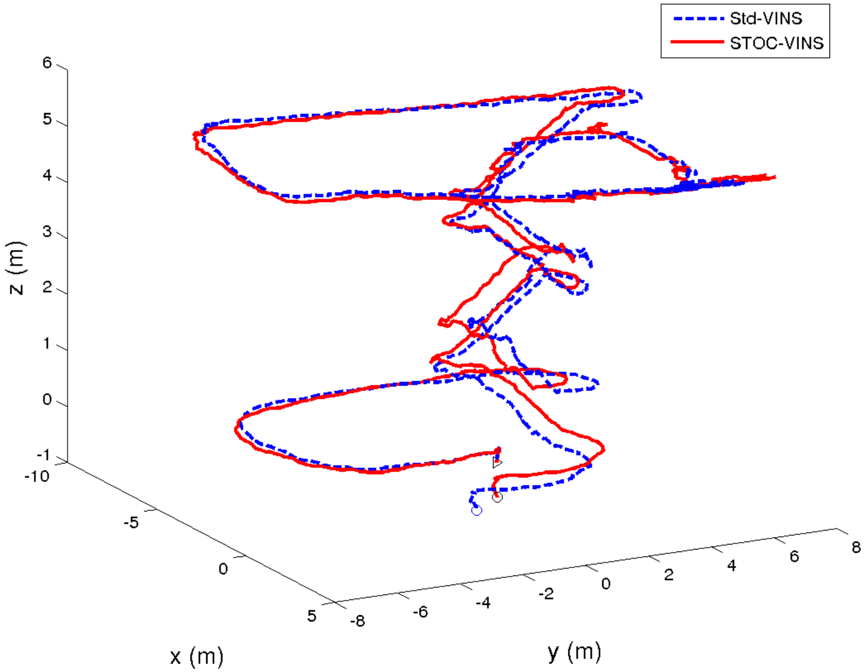
- combine dead-reckoning from inertial sensors (accelerometers, gyros) with visual information to estimate the motion of a robot or mobile device
- Typically performed with a Kalman filter
- Example: Google Tango

Guquan Huang, MIT

VISUAL-INERTIAL NAVIGATION

Goal:

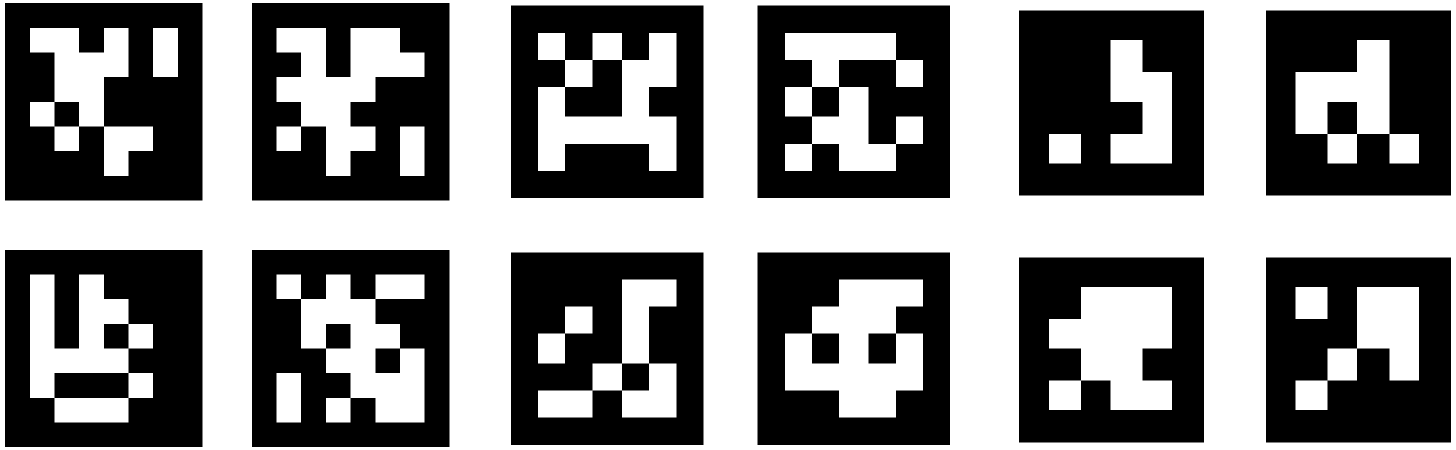
- combine dead-reckoning from inertial sensors (accelerometers, gyros) with visual information to estimate the motion of a robot or mobile device
- Typically performed with a Kalman filter
- Example: Google Tango



Guquan Huang, MIT

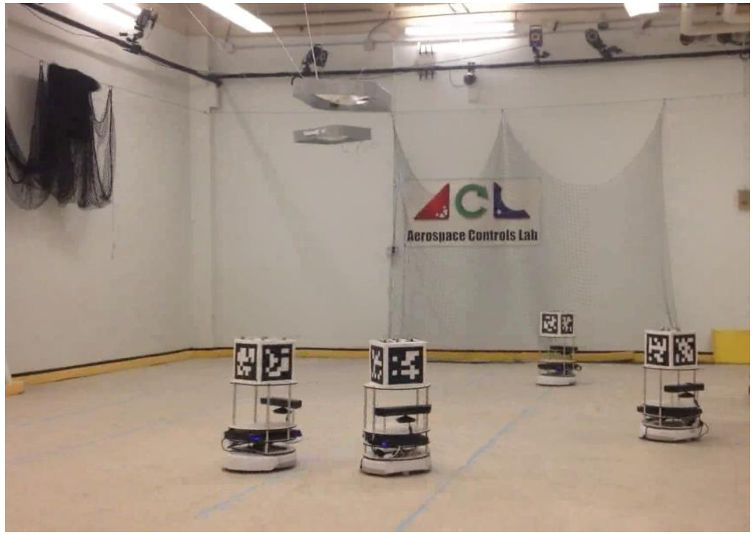
APRILTAGS

- Visual fidicuials developed at the University of Michigan by Edwin Olson
- Open source
- Enable testing of location-based sensing algorithms without solving for recogntion



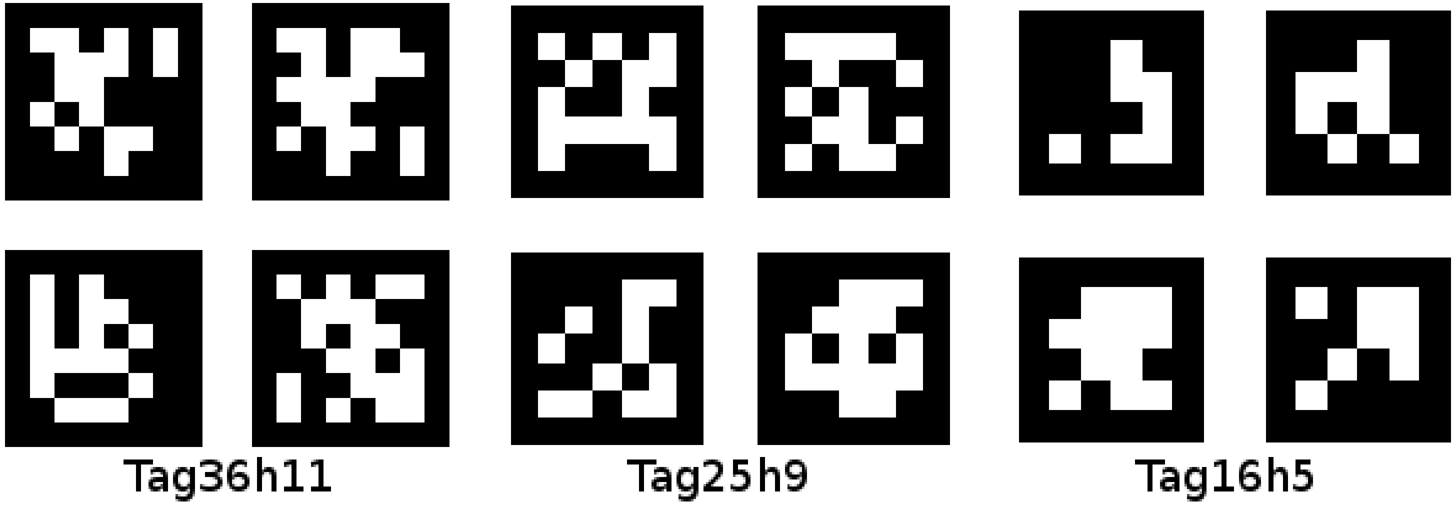
Tag36h11 Tag25h9 Tag16h5

https://april.eecs.umich.edu/wiki/index.php/AprilTags



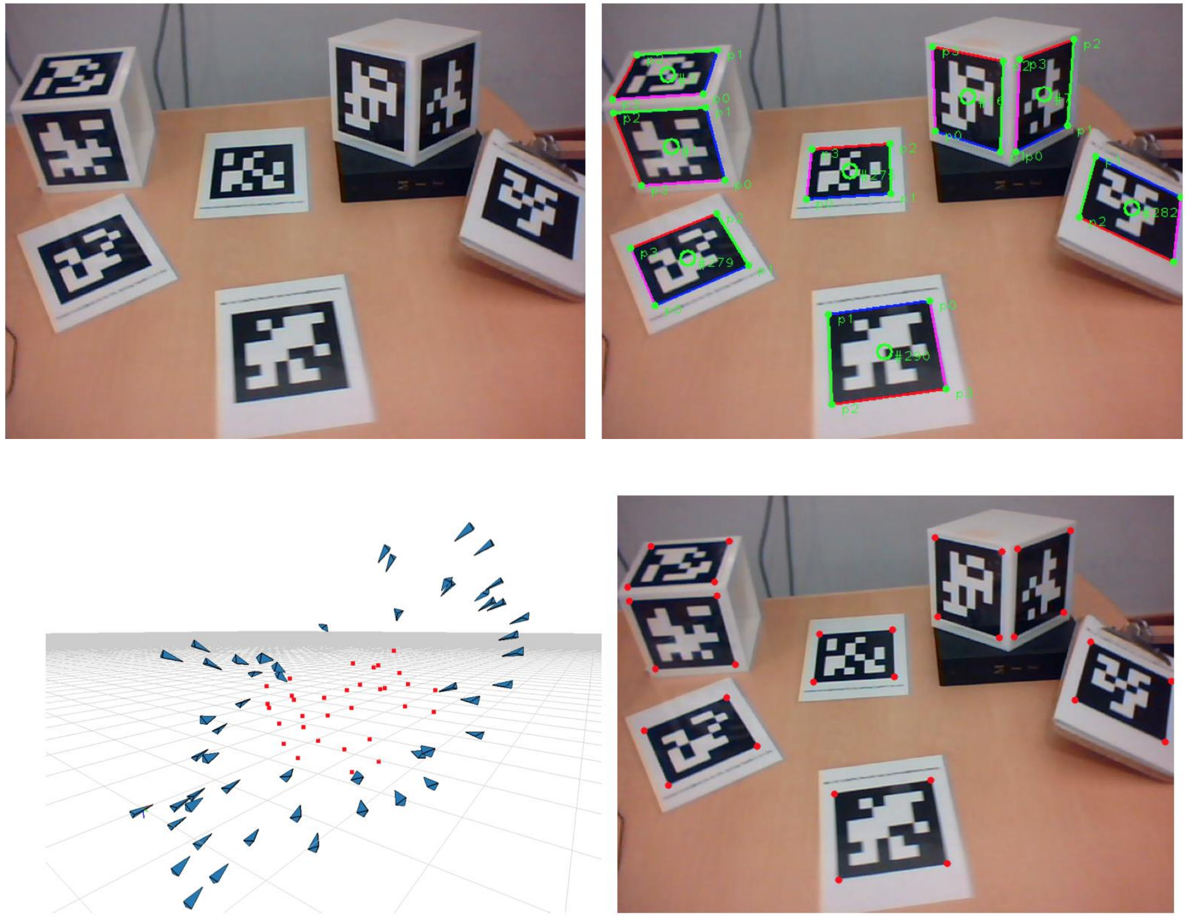
APRILTAGS

- Visual fidiciuals developed at the University of Michigan by Edwin Olson
- Open source
- Enable testing of location-based sensing algorithms without solving for recogntion



APRILTAGS

- Visual fidicuials developed at the University of Michigan by Edwin Olson
- Open source
- Enable testing of location-based sensing algorithms without solving for recogntion



Use of AprilTags to test an algorithm for robust mapping (Rosen, MIT)

SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM)

Goal: build a global map using local sensor measurements, while concurrently estimating the vehicle trajectory

Victoria Park Data Set
(Jose Guivant, ACFR)

Dead-reckoning

Landmarks

Measurements

Trajectory estimated by SLAM (Michael Kaess, Georgia Tech)

VICTORIA PARK LIDAR DATASET

Raw Data, Vehicle-relative

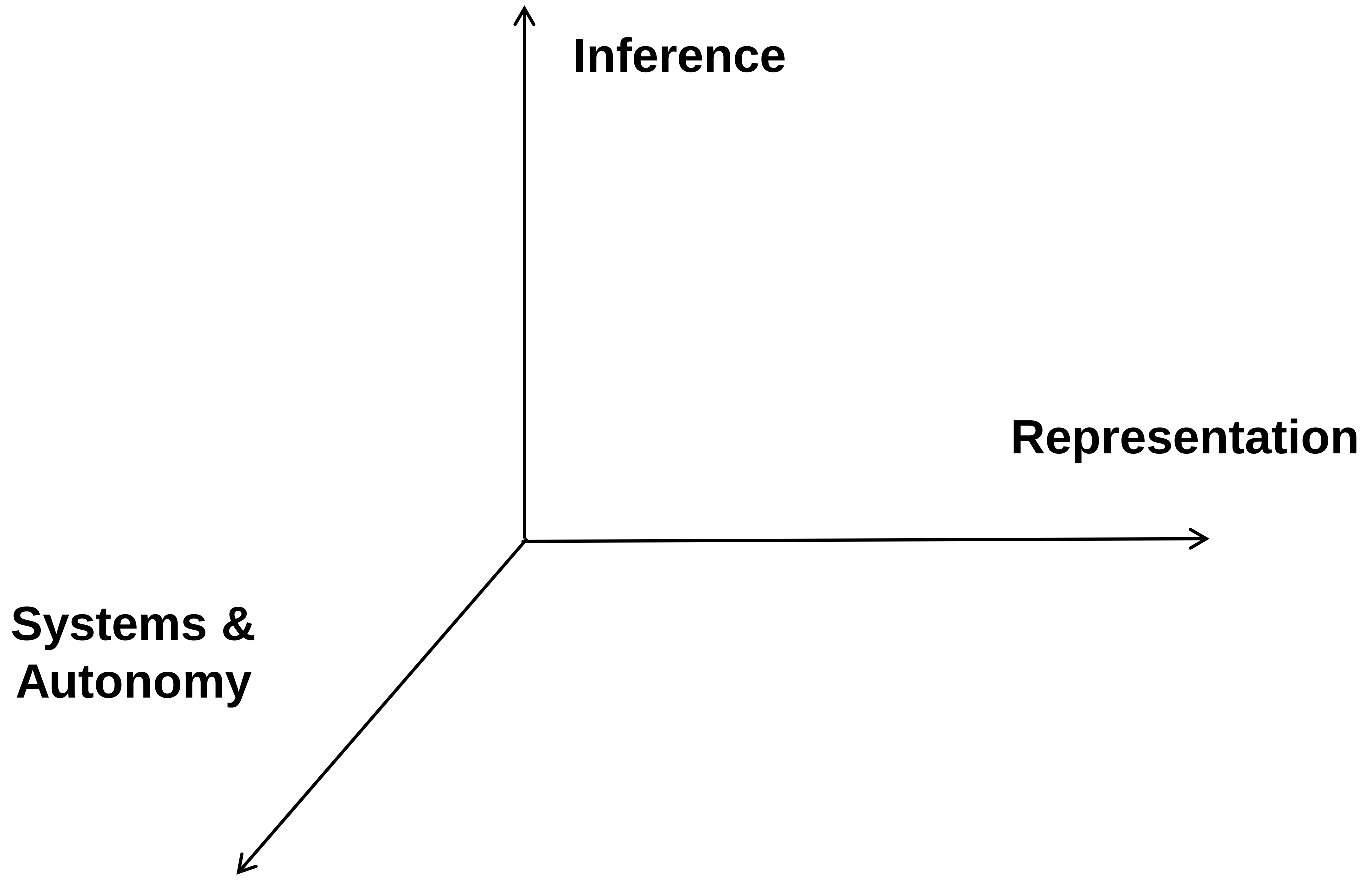
(Open source dataset courtesy Dr. Jose Guivant)

VICTORIA PARK LIDAR DATASET

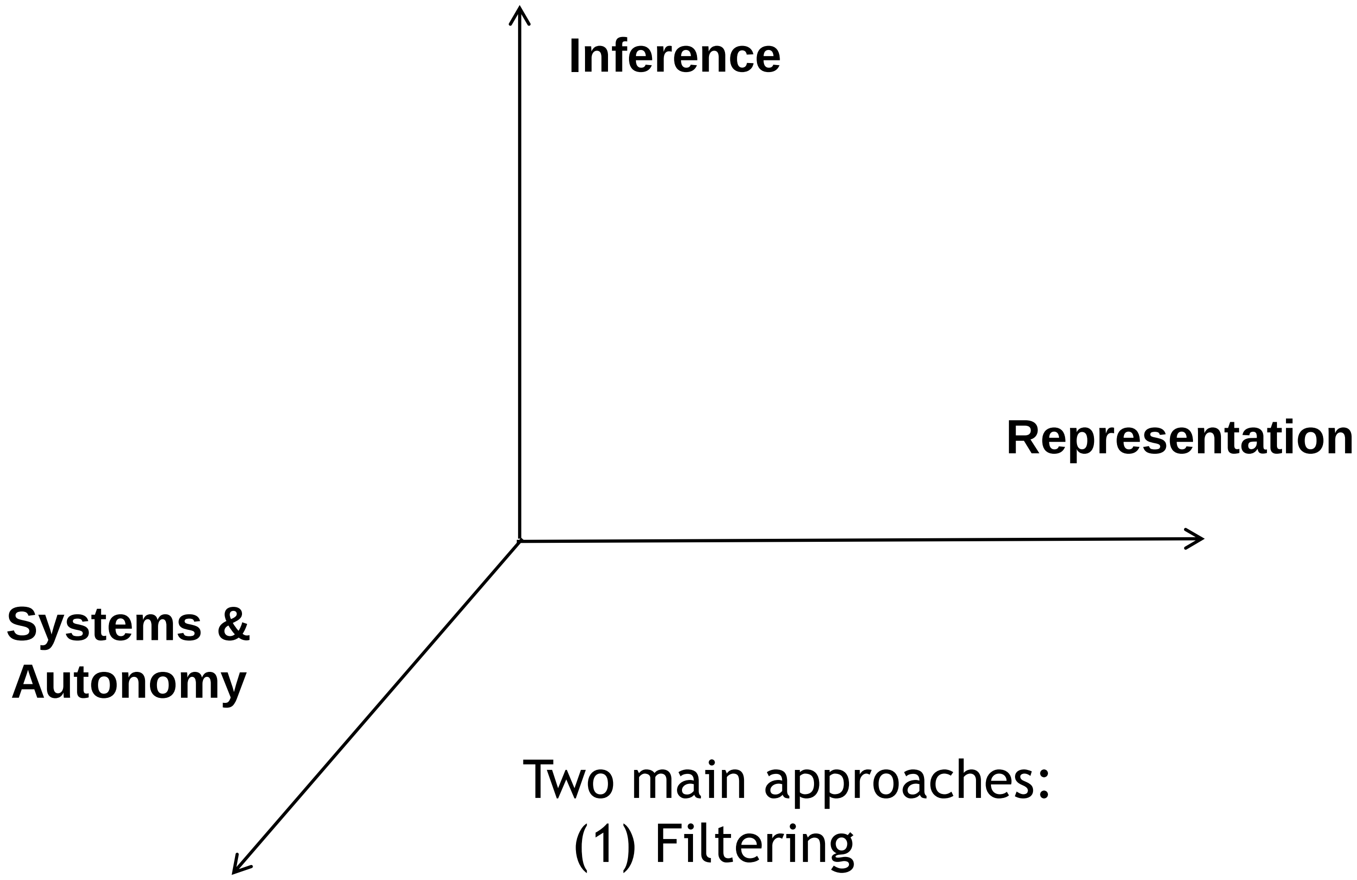
(Open source dataset courtesy Dr. Jose Guivant)

SLAM Processing
(Dr. Michael Kaess, GTech/MIT)

WHY IS SLAM DIFFICULT?



WHY IS SLAM DIFFICULT?

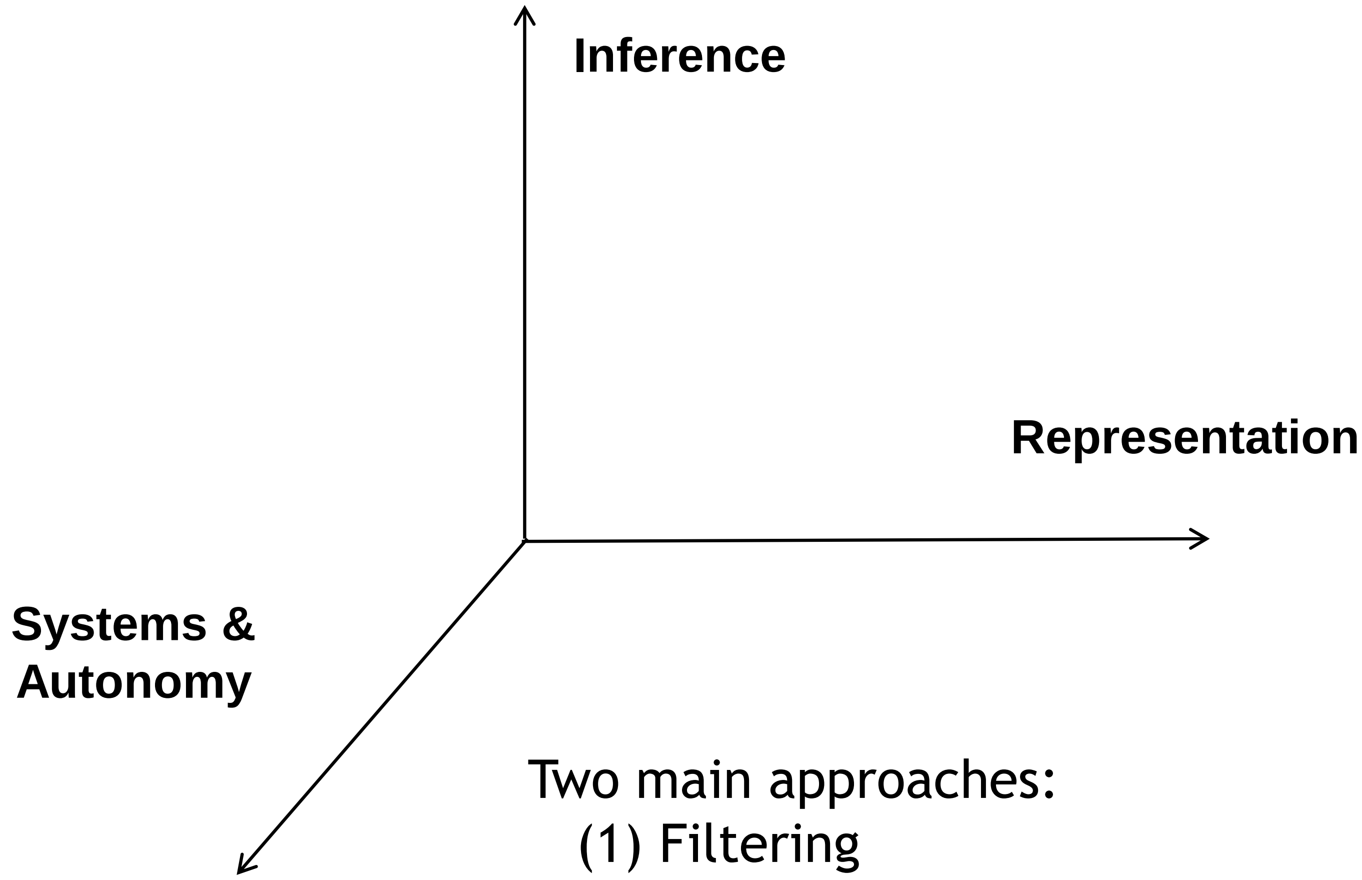


Two main approaches:

(1) Filtering
(2) Pose Graph Optimization

WHY IS SLAM DIFFICULT?

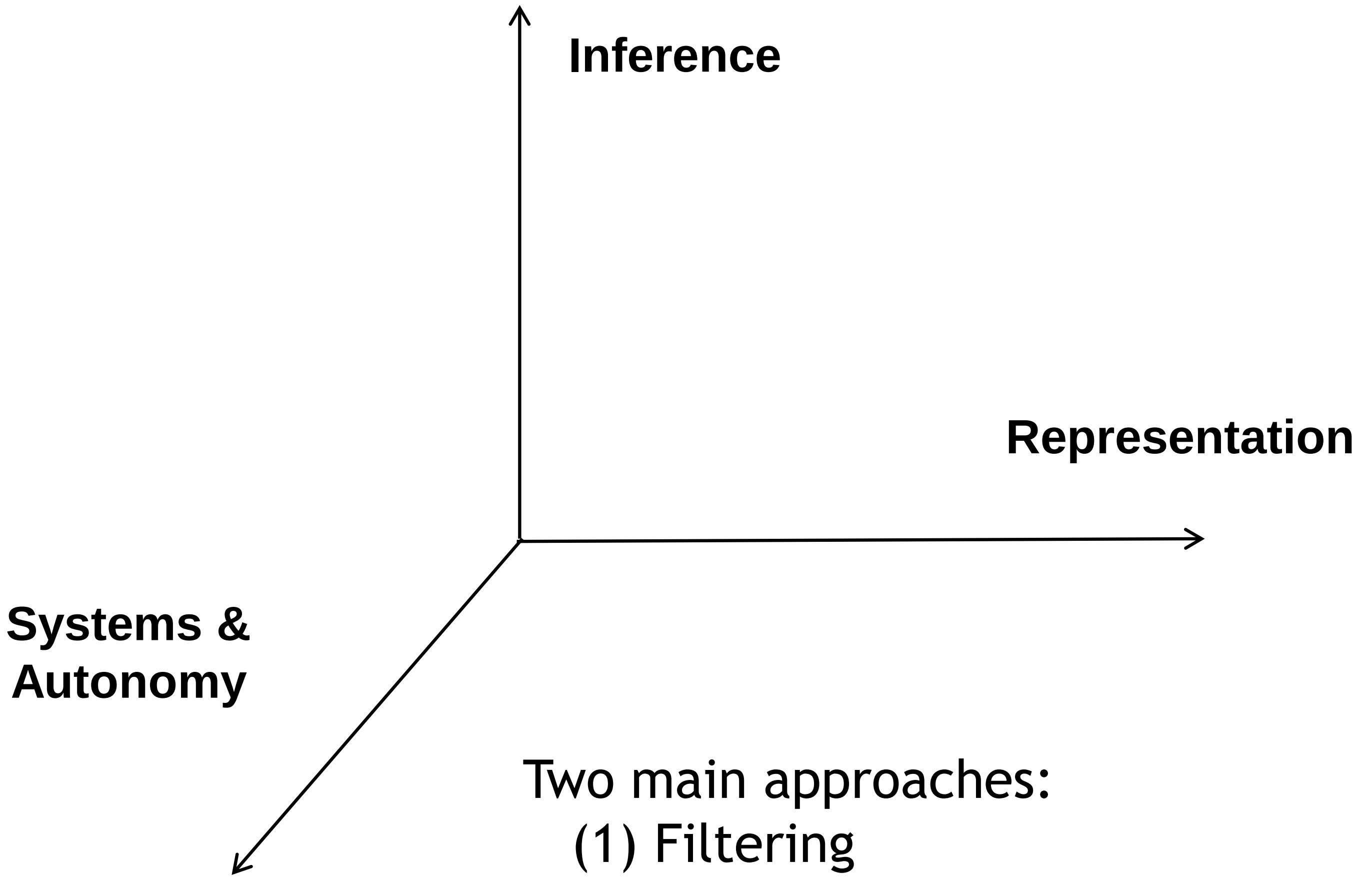
SLAM with sonar



Two main approaches:
(1) Filtering
(2) Pose Graph Optimization

WHY IS SLAM DIFFICULT?

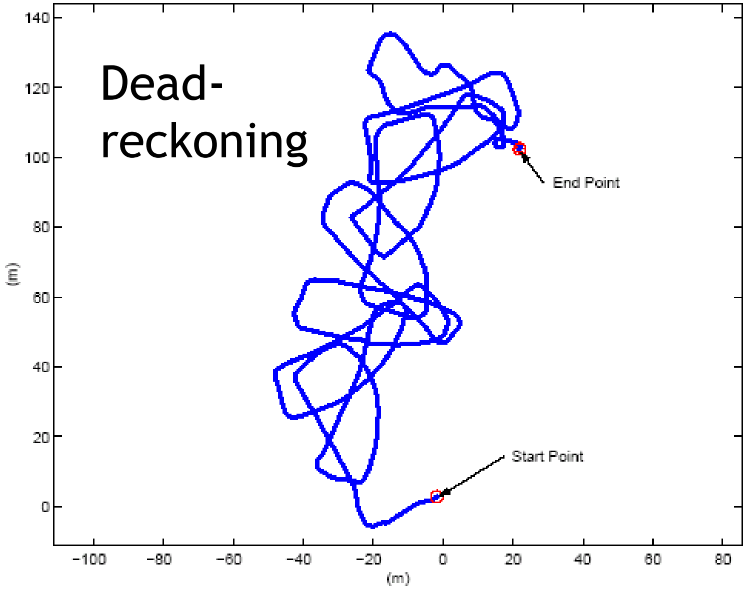
SLAM with sonar

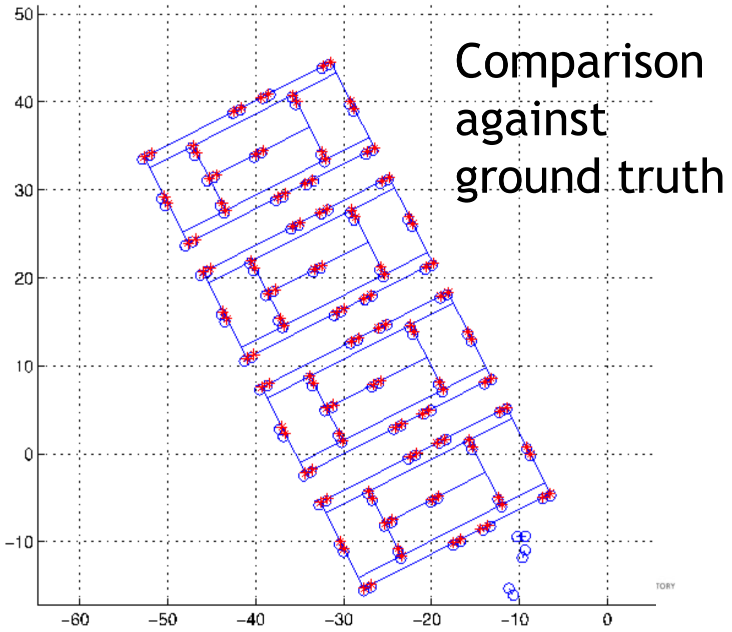


SLAM with vision

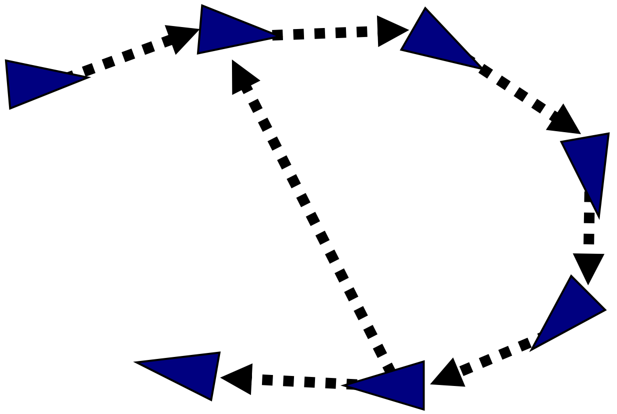
Two main approaches:
(1) Filtering
(2) Pose Graph Optimization

FILTERING EXAMPLE: 2-D SLAM WITH LIDAR DATA





POSE GRAPH OPTIMIZATION EXAMPLE (WITH LOOP CLOSING)



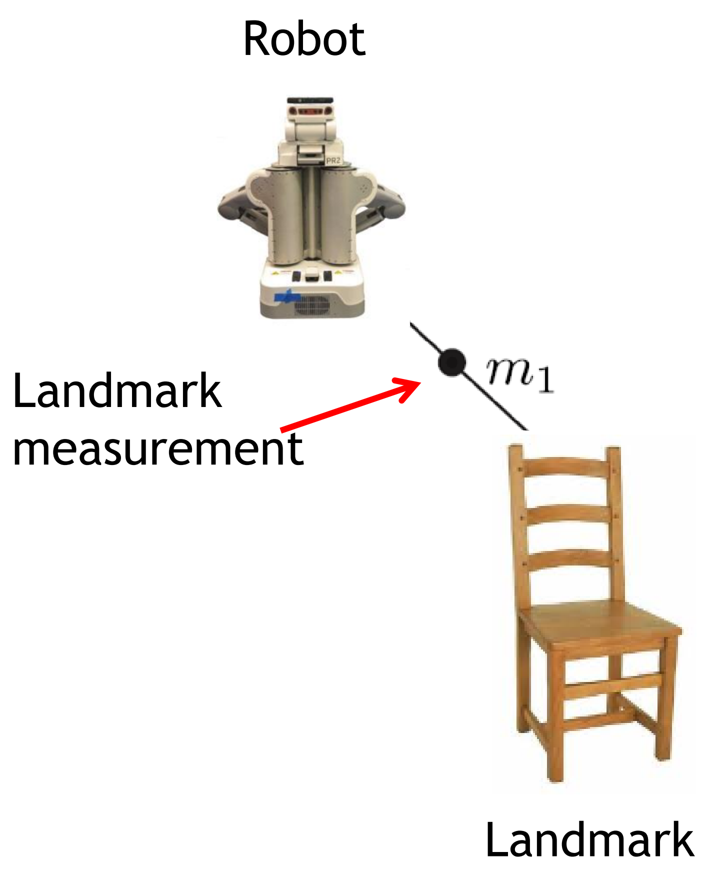
Pose Graph Representation:

- Sequence of vehicle poses with constraints between them
- Constraints are derived from measurements
- Non-sequential constraints are called loop closures

Solve as a least-squares optimization problem

MIT Killian Court Data Set (Olson, Leonard and Teller, 2006)

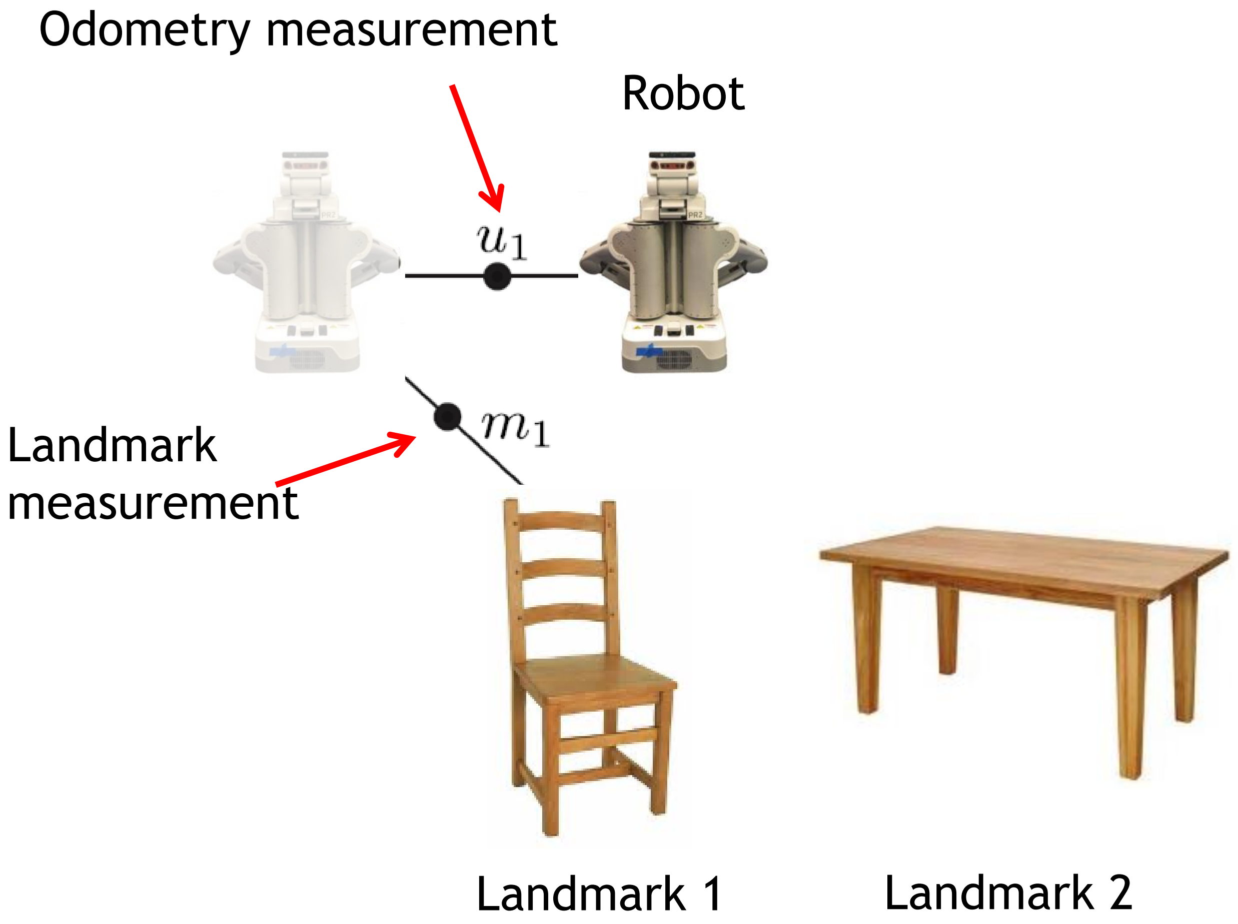
THE POSE GRAPH SLAM PROBLEM (T=0)



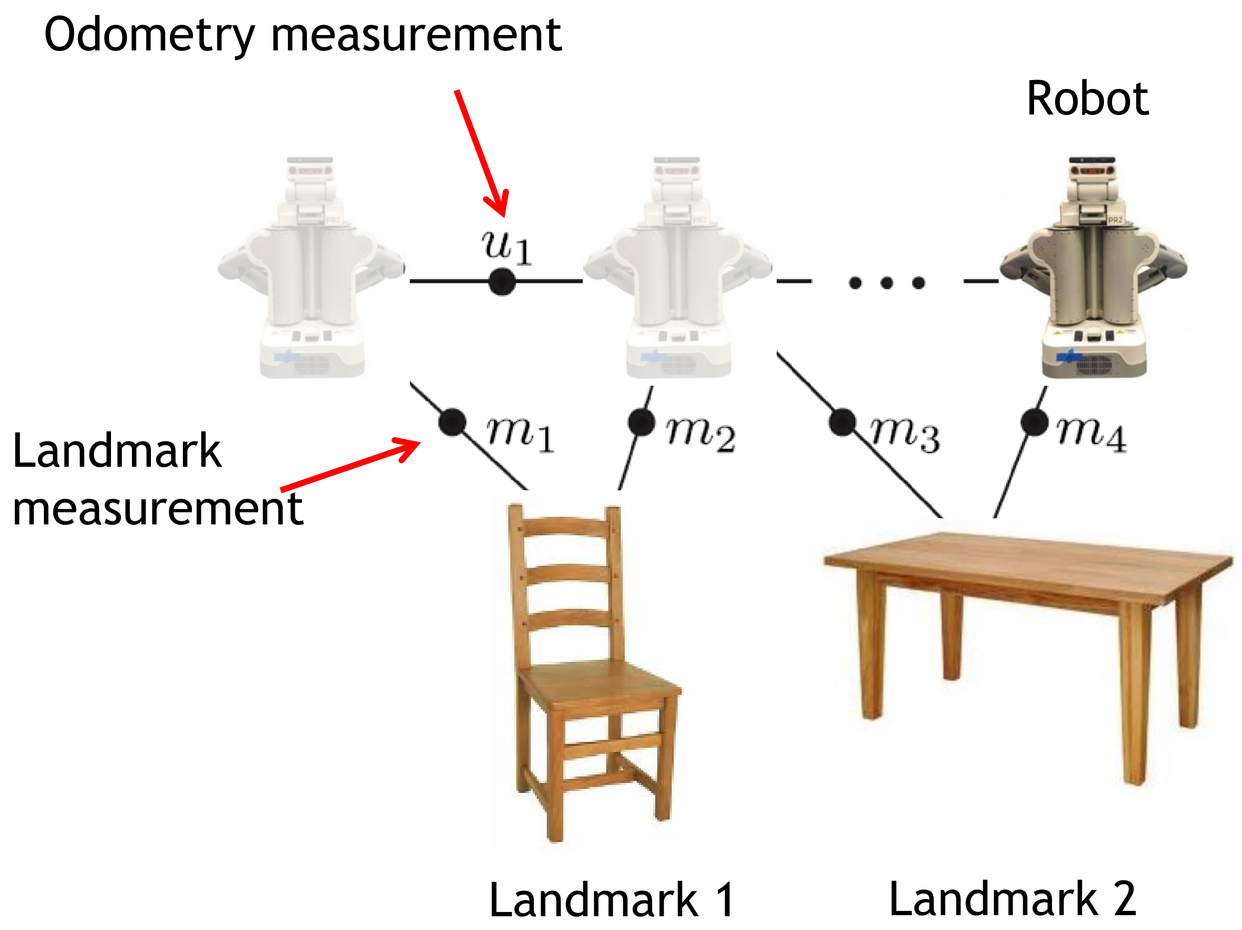
Onboard sensors:

- Wheel odometry
- Inertial measurement unit (gyro, accelerometer)
- Sonar
- Laser range finder
- Camera
- RGB-D sensors

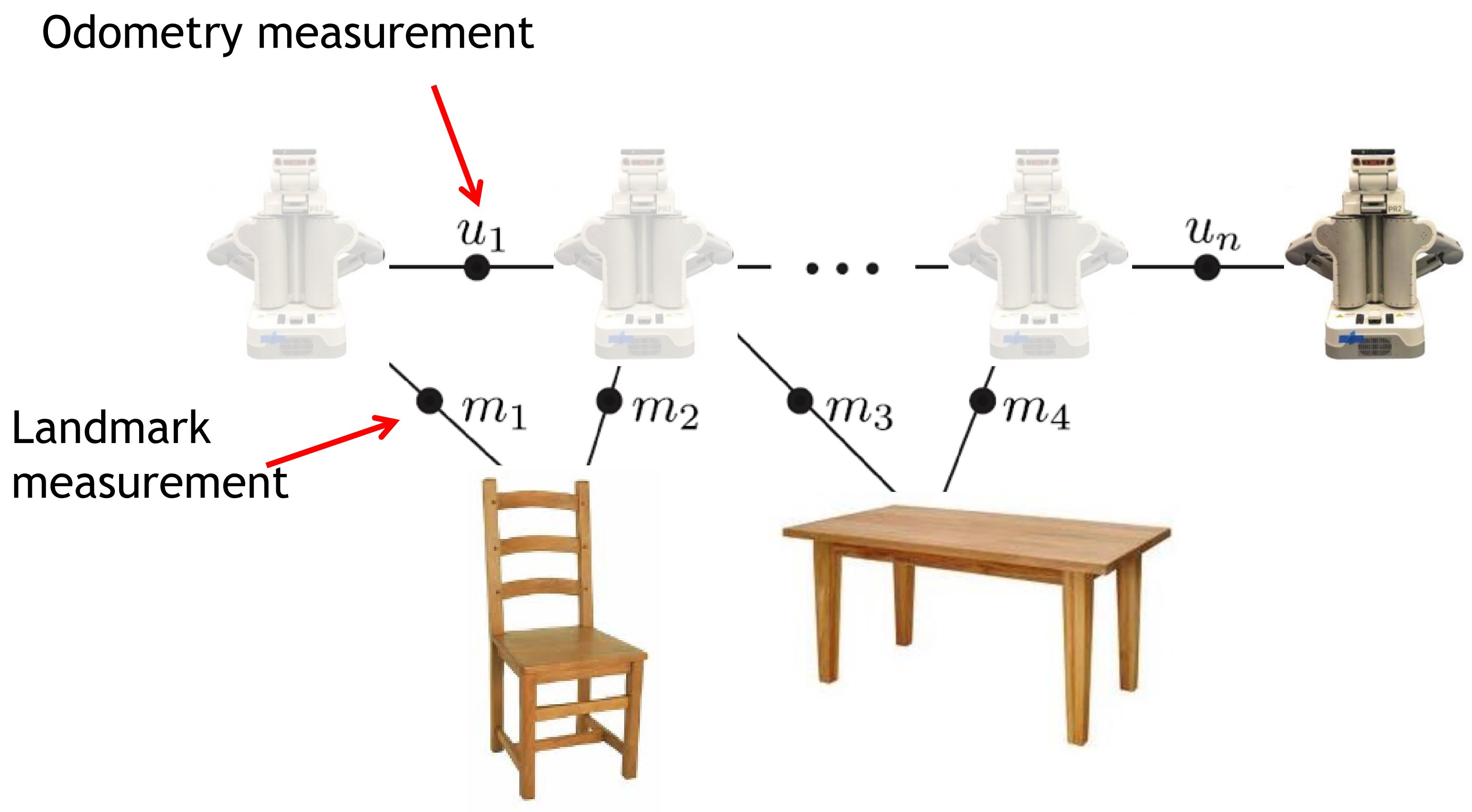
THE POSE GRAPH SLAM PROBLEM (T=1)



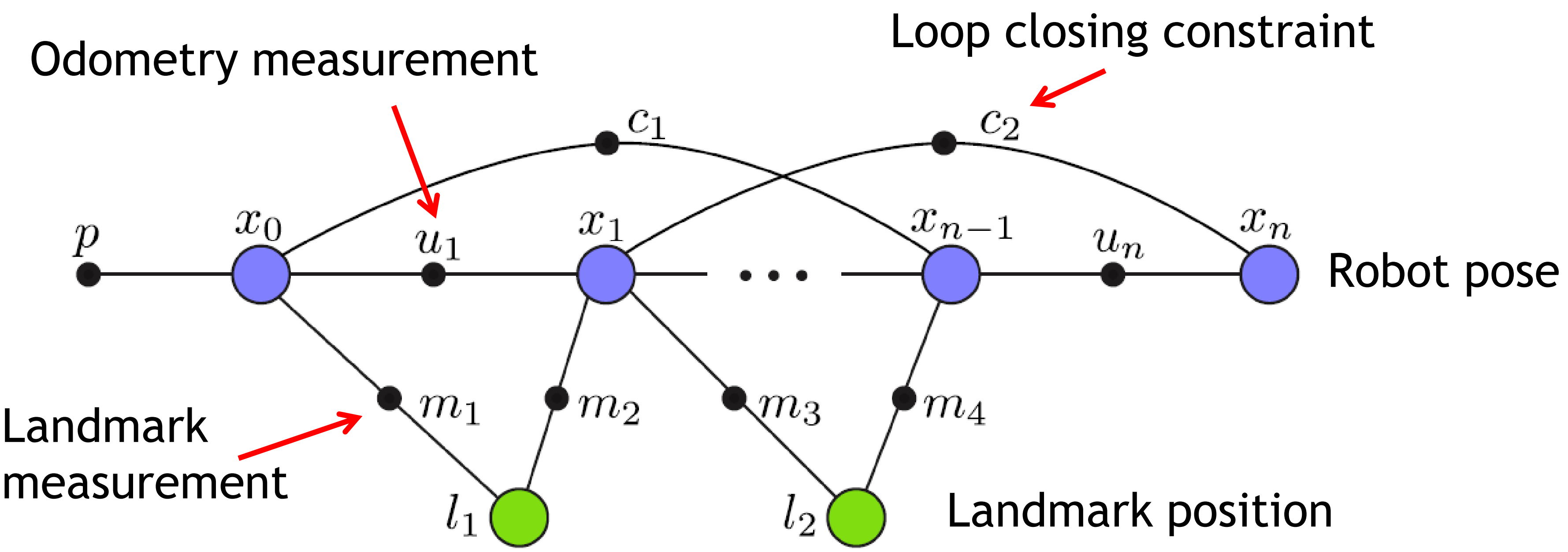
THE POSE GRAPH SLAM PROBLEM (T=N-1)



THE POSE GRAPH SLAM PROBLEM (T=N)

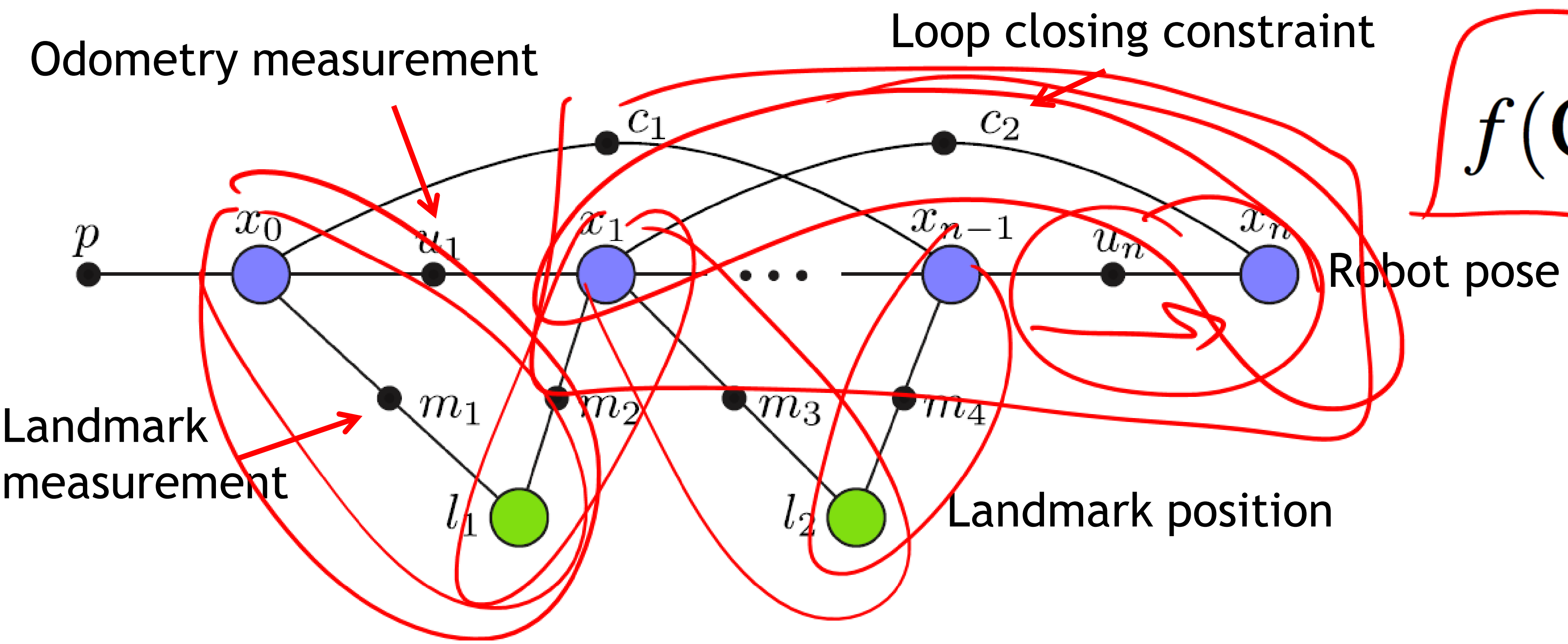


FACTOR GRAPH REPRESENTATION



Bipartite graph with **variable nodes** and **factor nodes**

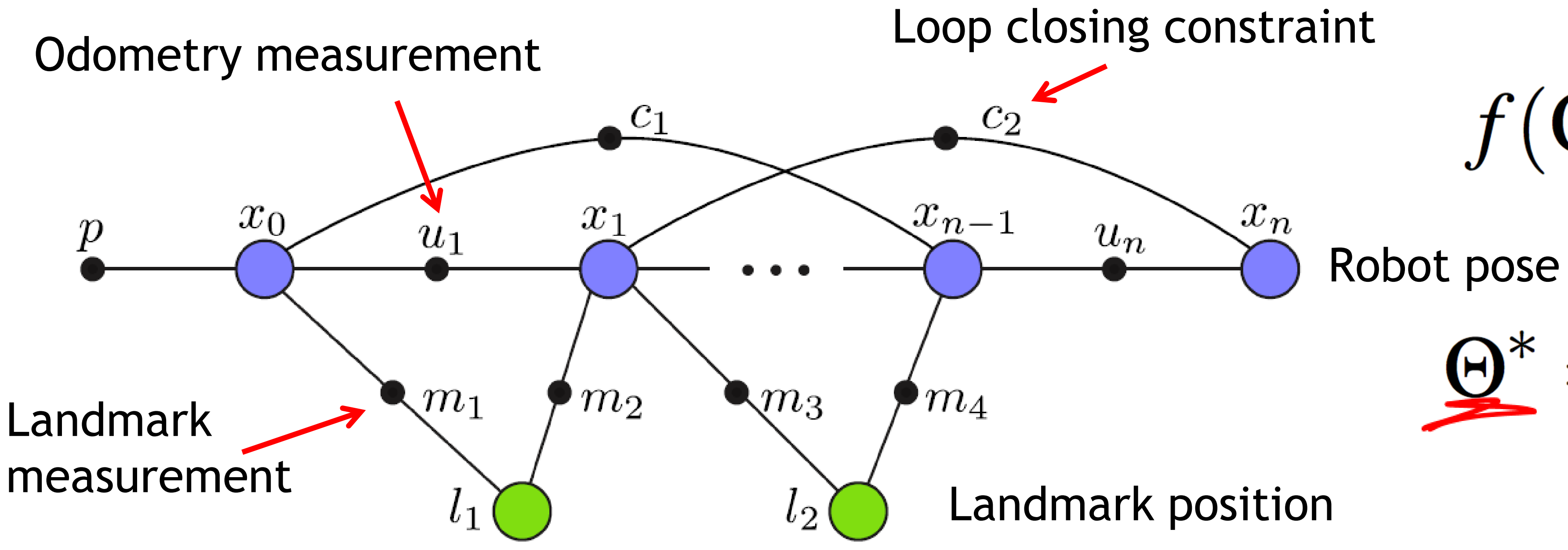
FACTOR GRAPH REPRESENTATION



$$f(\Theta) = \prod_i f_i(\Theta_i)$$

Bipartite graph with **variable nodes** and **factor nodes**

FACTOR GRAPH REPRESENTATION

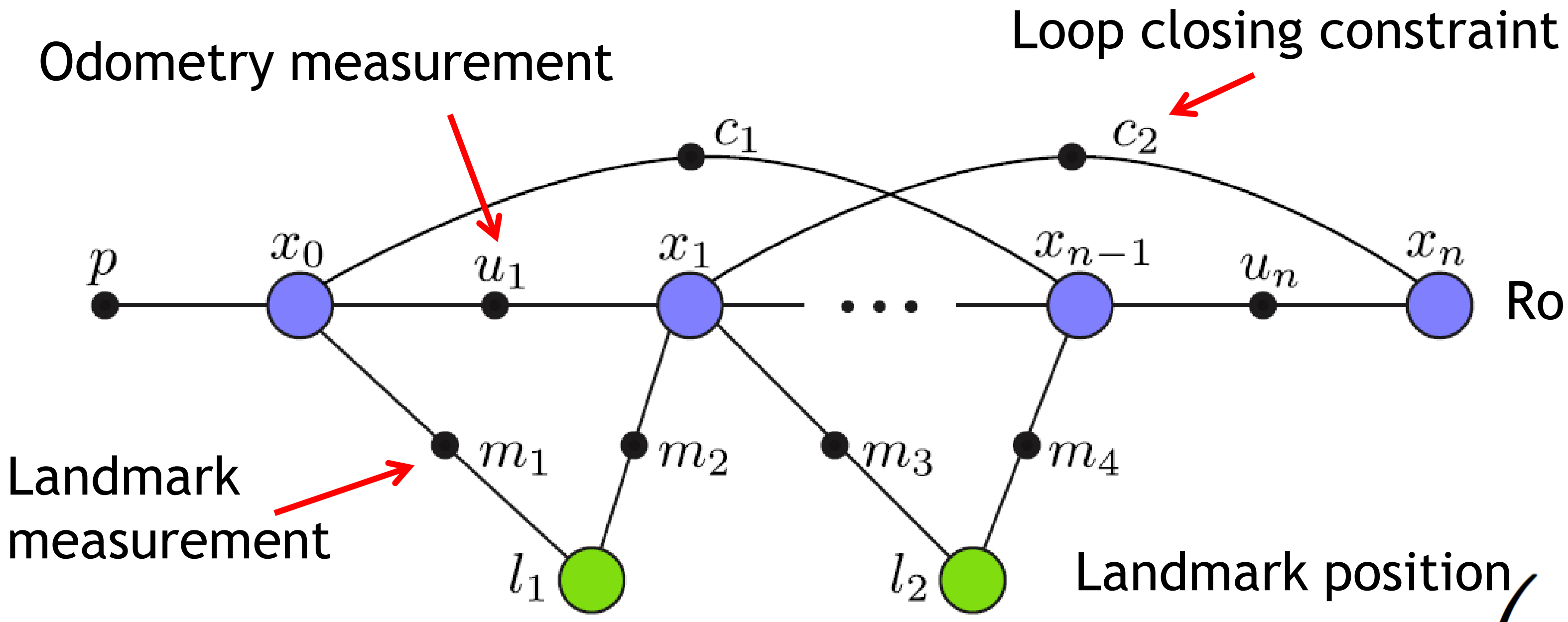


$$f(\Theta) = \prod_i f_i(\Theta_i)$$

$$\Theta^* = \arg\max_{\Theta} f(\Theta)$$

Bipartite graph with **variable nodes** and **factor nodes**

FACTOR GRAPH REPRESENTATION

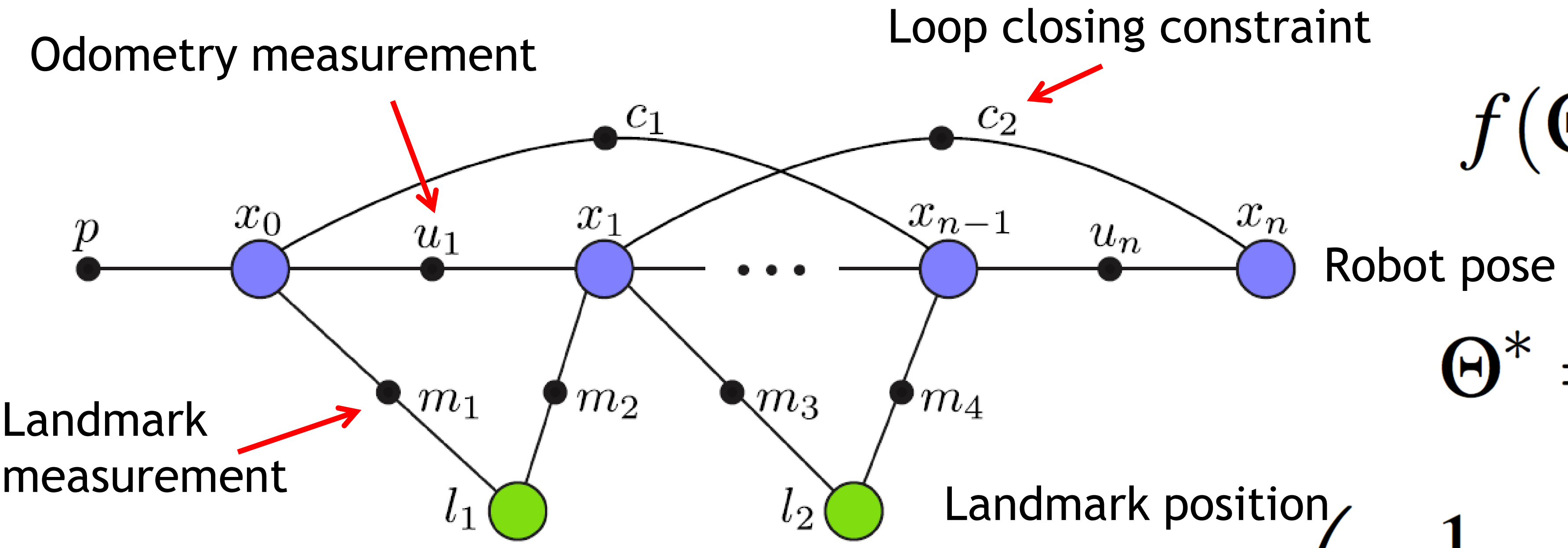


$$f(\Theta) = \prod_i f_i(\Theta_i)$$

$$\Theta^* = \arg\max_{\Theta} f(\Theta)$$

$$f_i(\Theta_i) = \mathcal{N}(h_i(\Theta_i); z_i, \Sigma_i) \propto \exp\left(-\frac{1}{2}\|h_i(\Theta_i) - z_i\|^2_{\Sigma_i}\right)$$

FACTOR GRAPH REPRESENTATION



$$f(\Theta) = \prod_i f_i(\Theta_i)$$

$$\Theta^* = \arg\max_\Theta f(\Theta)$$

$$f_i(\Theta_i) = \mathcal{N}(h_i(\Theta_i); z_i, \Sigma_i) \propto \exp\left(-\frac{1}{2}\|h_i(\Theta_i) - z_i\|^2_{\Sigma_i}\right)$$

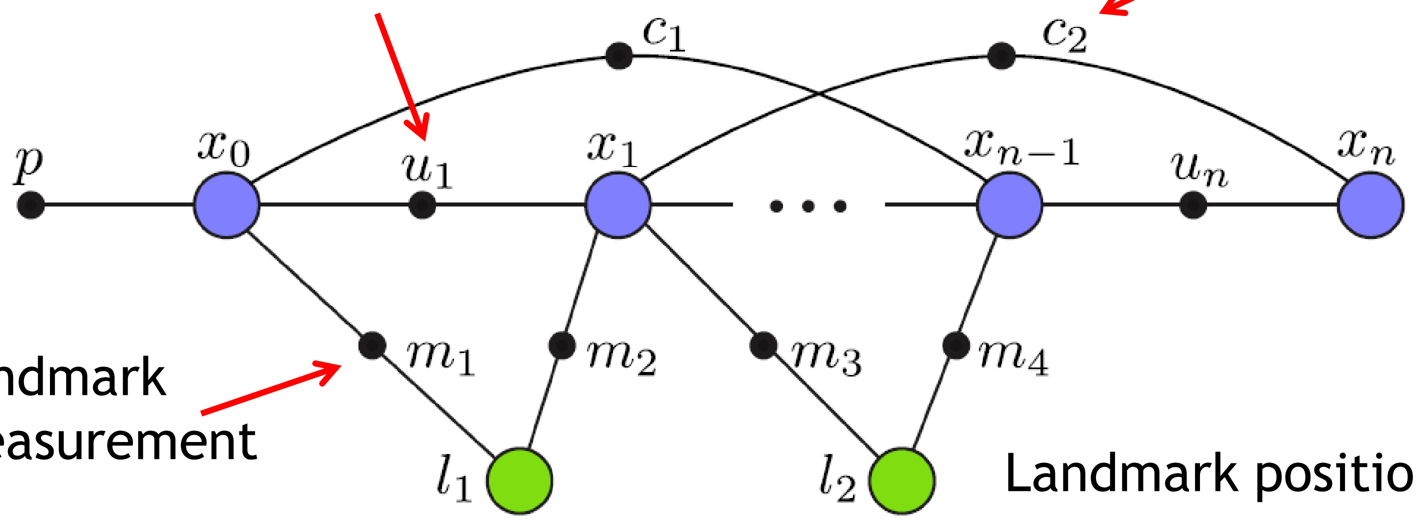
$$\arg\min_\Theta(-\log f(\Theta)) = \boxed{\arg\min_\Theta \frac{1}{2}\sum_i \|h_i(\Theta_i) - z_i\|^2_{\Sigma_i}}$$

FACTOR GRAPH REPRESENTATION

Implement with GTSAM, G2O, or iSAM:
See: https://openslam.org/g2o.html
https://collab.cc.gatech.edu/borg/gtsam
http://people.csail.mit.edu/kaess/isam/

Odometry measurement

Loop closing constraint



Robot pose

Landmark measurement

Landmark position

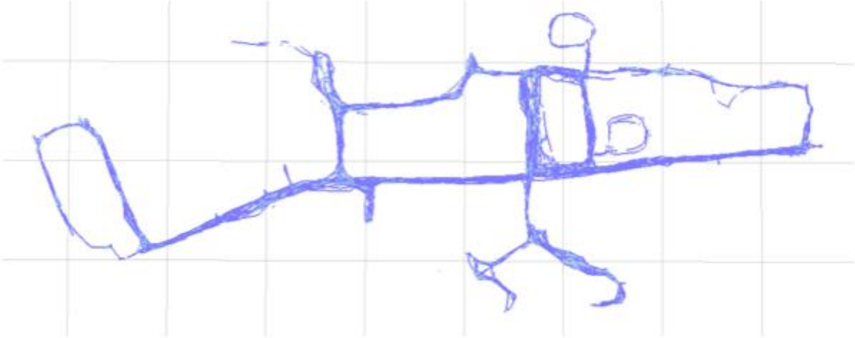
$$f(\Theta) = \prod_i f_i(\Theta_i)$$

$$\Theta^* = \arg\max_{\Theta} f(\Theta)$$

Enforcing sparsity in the underlying representation is critical for maintaining efficenciey and consistency of the SLAM solution

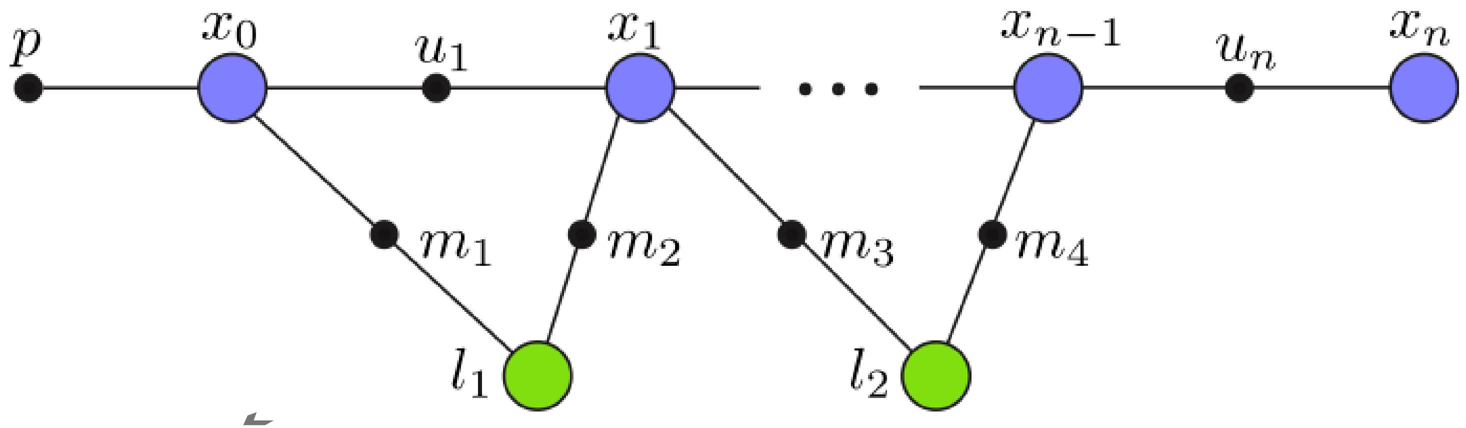
$$\arg\min_{\Theta} \left(-\log f(\Theta)\right) = \arg\min_{\Theta} \frac{1}{2} \sum_i \|h_i(\Theta_i) - z_i\|^2_{\Sigma_i}$$

EXAMPLE: 3D VISUAL SLAM FOR A B21 ROBOT

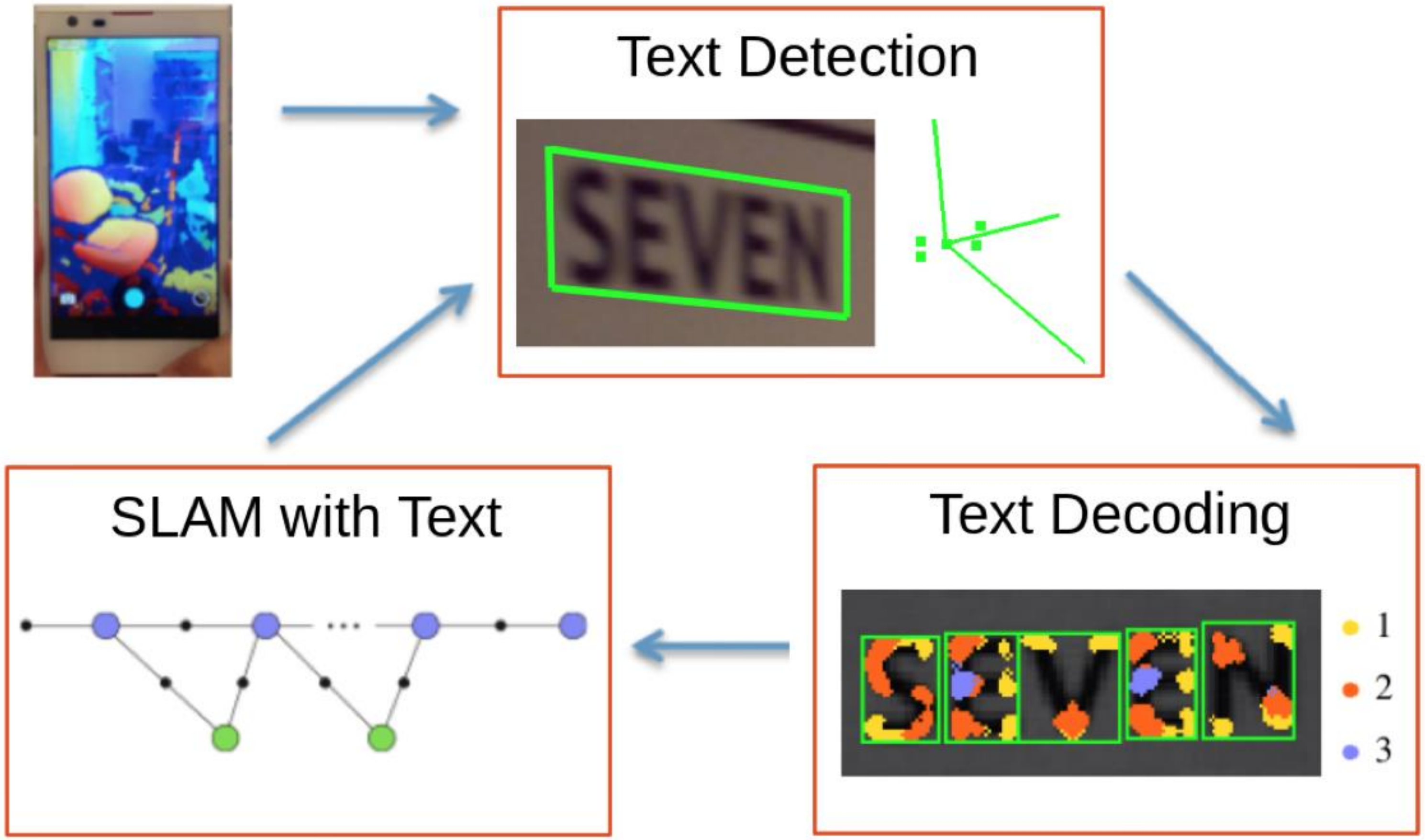


EXAMPLE: POSE GRAPH SLAM WITH TEXT LANDMARKS

- incremental Smoothing And Mapping (iSAM)
- Hidden states to be estimated are:
  - camera poses $x_{0..n}$
  - high-level landmark locations l_1, l_2.
  - These states are connected through prior p, odometry $u_{1:n}$ and landmark measurements $m_{1..}$



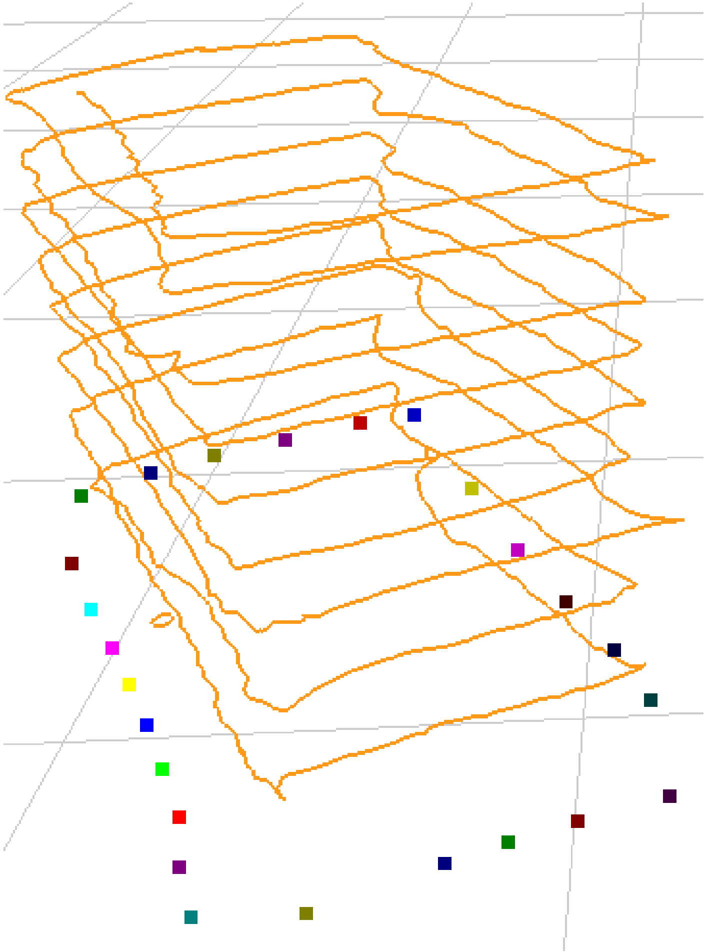
EXAMPLE: POSE GRAPH SLAM WITH TEXT LANDMARKS



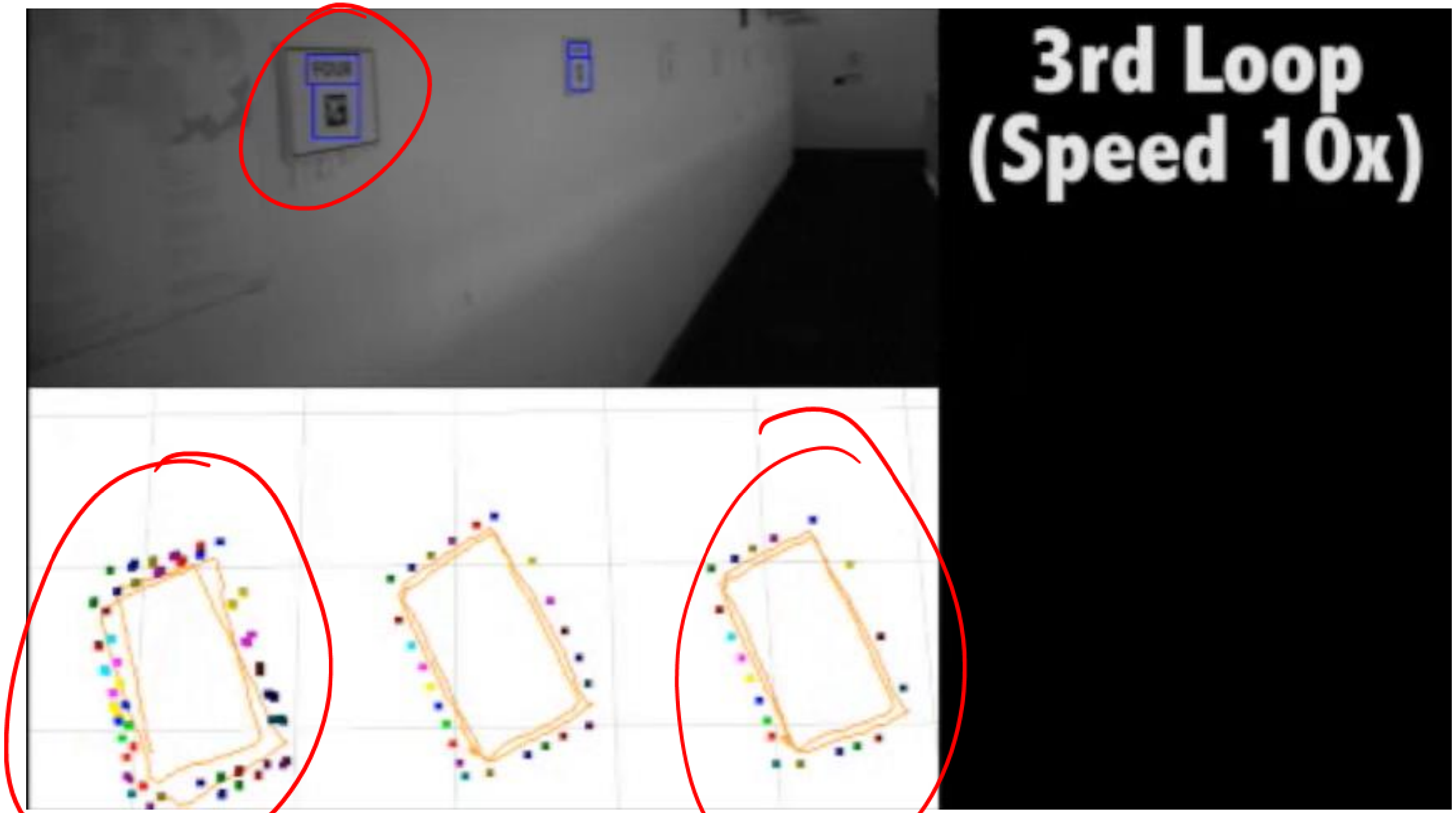
FOUR SCORE" dataset

- Gettysburg Address by Abraham Lincoln
- FOUR SCORE AND SEVEN YEARS AGO OUR FATHERS BROUGHT FORTH on THIS CONTINENT a NEW NATION, CONCEIVED in LIBERTY, and DEDICATED to THE PROPOSITION THAT ALL MEN ARE CREATED EQUAL.

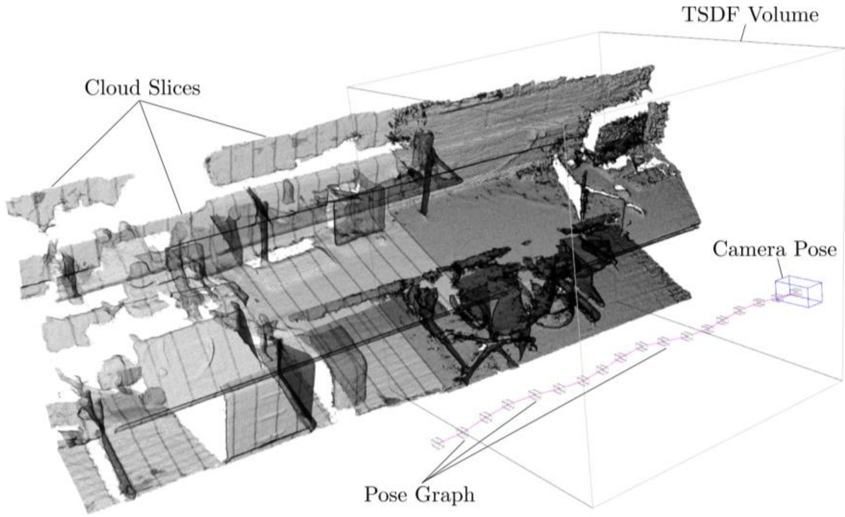
EXAMPLE: POSE GRAPH SLAM WITH TEXT LANDMARKS



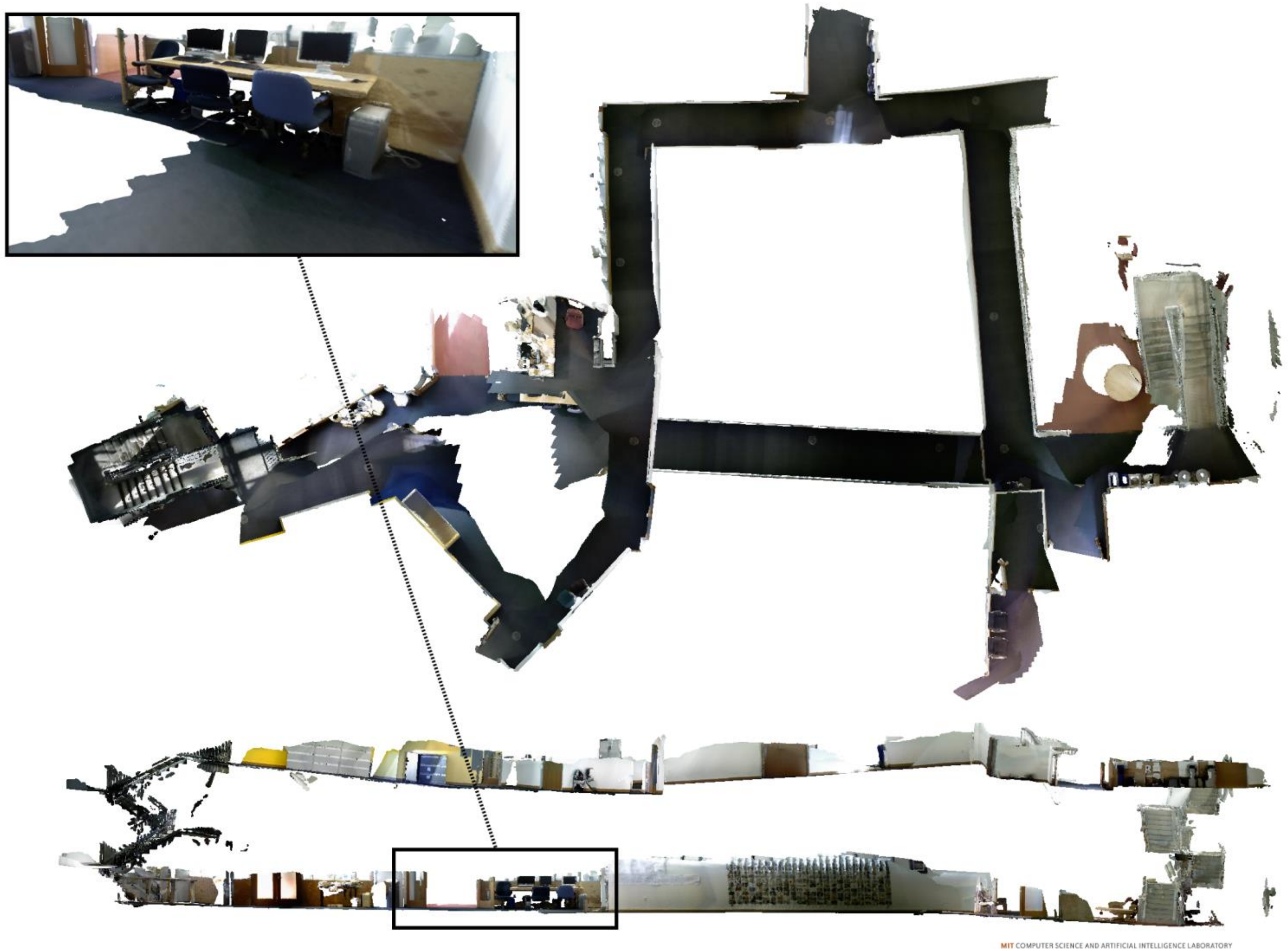
EXAMPLE: POSE GRAPH SLAM WITH TEXT LANDMARKS



EXAMPLE: DENSE MAPPING USING SPATIALLY EXTENDED KINECTFUSION (TOM WHELAN, NUIM)



et o



EXAMPLE: HUMANOID ROBOT LOCOMOTION (FALLON, MIT)

Kuindersma et al., Optimization-based locomotion planning, estimation, and control design for the Atlas humanoid robot, 2015

APPLICATION: SELF-DRIVING VEHICLES

APPLICATION: SELF-DRIVING VEHICLES

Incredible excitement in the field due to the Google Self-driving car project and related efforts by Uber, Tesla, and the major car companies

APPLICATION: SELF-DRIVING VEHICLES

Questions:

- Technological
- Economic
- Employment
- Ethical
- Legal
- Security
- Energy and the environment

APPLICATION: SELF-DRIVING VEHICLES

Location information is critical:

Where am I?

Where am I going?

How should I get there?

What is around me?

What will happen next?

What should I do next?

POTENTIAL BENEFITS OF SELF-DRIVING VEHICLES

Safety

- Over 5 Million vehicle crashes per year in the US
- 93% of accidents have human error as a primary factor
- Over 30,000 fatalities/year in the US from traffic accidents

Increased Road Network Efficiency

Recovery of Time Lost due to Commuting

Radically New Models for Personal Mobility and the Distribution of Goods and Services

HOW DO SELF-DRIVING VEHICLES WORK?

MIT DARPA Urban Challenge Vehicle (2007)

Blade cluster

- 10 blades each with two 2.33GHz dual-core processors ➔ 40 cores

A **lot** of sensors

- Applanix IMU/GPS
- 12 SICK Lidars
- Velodyne (~64 Lidars)
- 15 radars
- 5 cameras

6 kW generator

MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

HOW DO SELF-DRIVING VEHICLES WORK?

Blade cluster

- 10 blades each with two 2.33GHz dual-core processors ➔ 40 cores

A **lot** of sensors

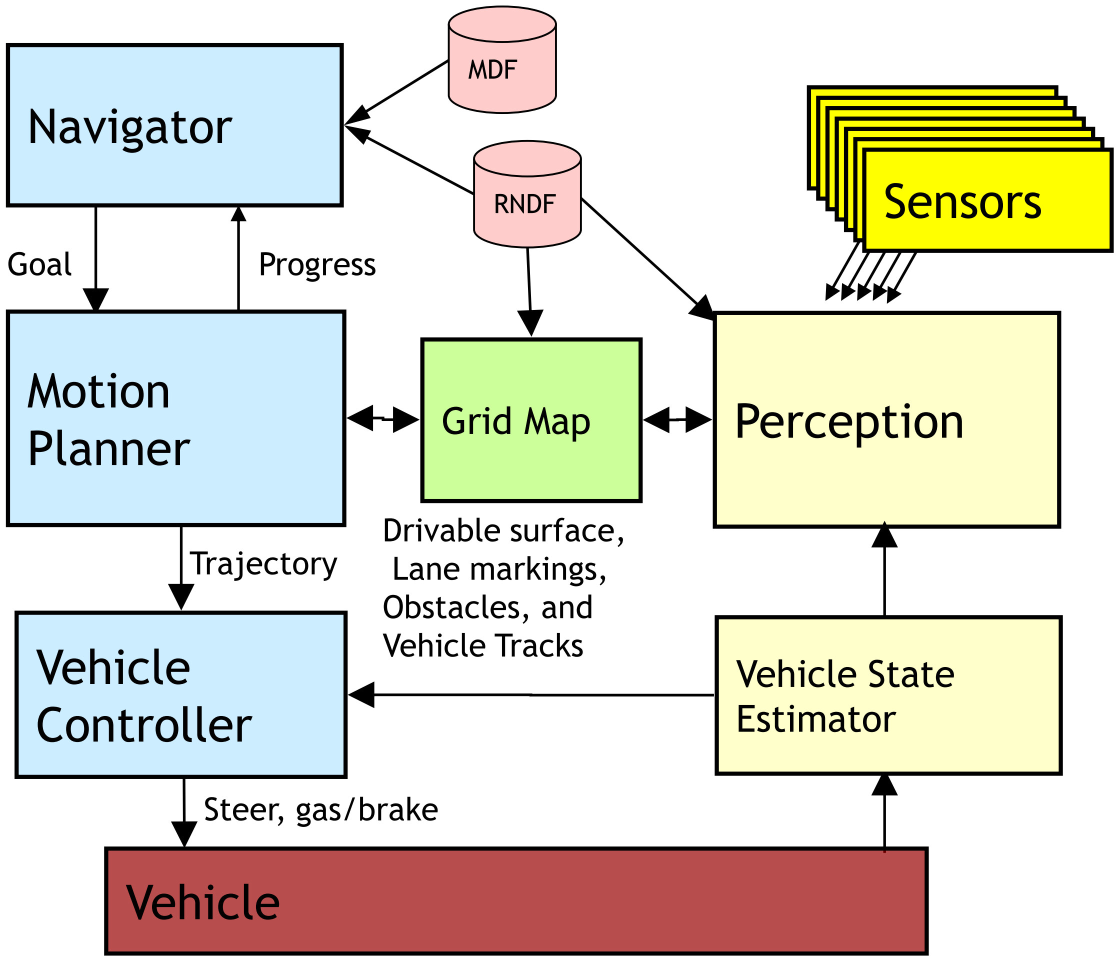
- Applanix IMU/GPS
- 12 SICK Lidars
- Velodyne (~64 Lidars)
- 15 radars
- 5 cameras

6 kW generator

MIT DARPA Urban Challenge Vehicle (2007)

MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

HOW DO SELF-DRIVING VEHICLES WORK?



HISTORY OF THE GOOGLE CAR PROJECT

November, 2007 – Carnegie Mellon University wins the DARPA Urban Challenge; Stanford finishes 2^{nd}

Jan, 2009 to Oct, 2010
- Key members of the CMU and Stanford DUC Teams hired to work at Google-X under the leadership of Sebastian Thrun
- Autonomous execution of ten 100-mile routes
- 140,000 miles autonomous driving at time of J. Markoff NYTimes article (October, 2010)

Oct, 2010 to Mar, 2012 – Prius phase
- Thrun TED Talk; Steve Mahan trip to Taco B ell video

Apr, 2012 to approximately Sep, 2013 – Lexus phase
- "Google Chaffeur" highway driving (e.g. 90% of journey)
- "Attention problem" identified with Level 3

Oct, 2013 (approx) to April, 2014 – Shift to City driving

May, 2014 to present – "Level 4" 25 mph vehicles

TECHNOLOGICAL CHALLENGES

Four key challenges for self-driving vehicles:

- Maintaining Maps
- Adverse Weather
- Interacting with People
- Robust Computer Vision (towards P_D=1.0, P_{FA} = 0.0)?

CHALLENGE 1: MAINTAINING MAPS

Nov 08th, 2013

Nov 12th, 2013

Driving across the Harvard Bridge, from Boston to Cambridge (4 days apart)

CHALLENGE 1: MAINTAINING MAPS

Nov 08th, 2013 Nov 12th, 2013

Driving across the Harvard Bridge, from Boston to Cambridge (4 days apart)

CHALLENGE 1: MAINTAINING MAPS

- Current approaches to self-driving vehicles rely on accurate a priori maps of the road surface for precisely estimating the vehicles's position

CHALLENGE 2: ADVERSE WEATHER

- Rain, fog, and snow can limit the range of sensors and produce “ghost” returns
- Snow obscures road surface markings, making location estimation difficult

CHALLENGE 3: INTERACTING WITH PEOPLE

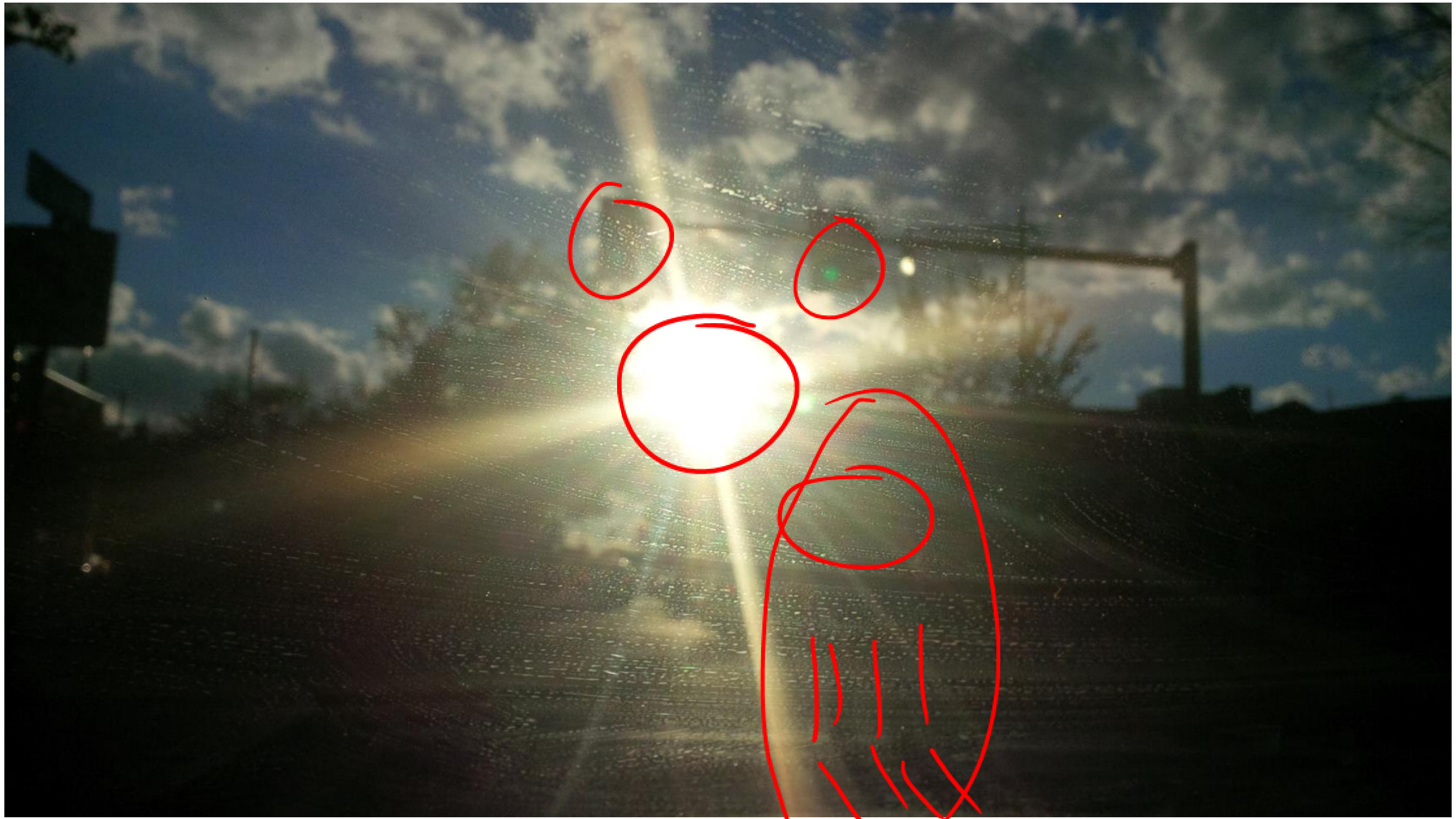
- Interacting with people is a major challenge for robotic perception
- Example 1:
  One police officer waves me through a red light; then another stops me at a green light

CHALLENGE 3: INTERACTING WITH PEOPLE

- Interacting with people is a major challenge for robotic perception
- Example 2: Making a left-turn at a busy junction with no traffic light; requires waving to another driver

CHALLENGE 4: ROBUST COMPUTER VISION

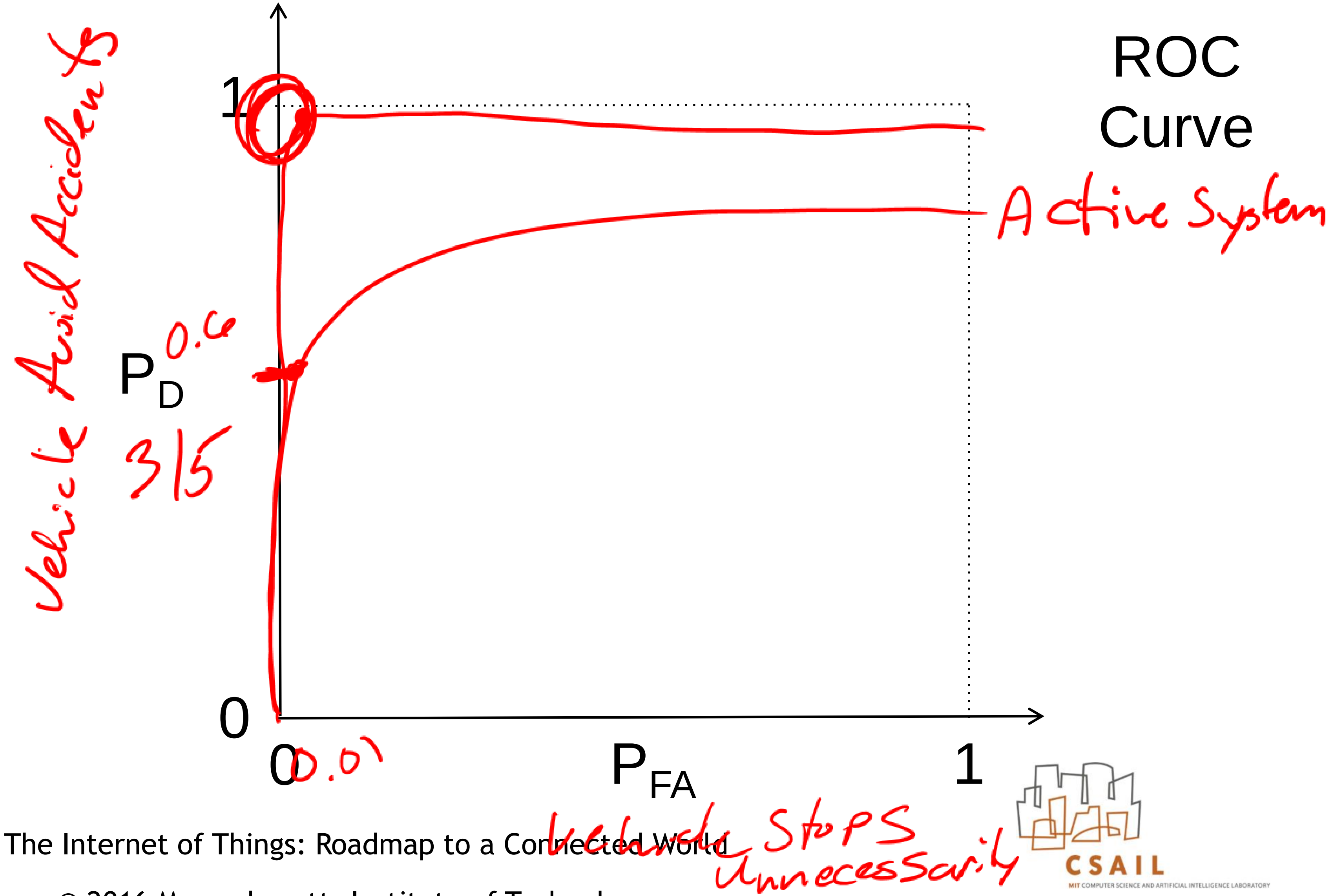
- Detector performance is typically shown using precision-recall (PR) or receiver operating characteristic (ROC) curves
- No sensor is perfect!
- P_D = Probability of Detection
- P_{FA} = Probability of False Alarm



What do you see in this picture?

CHALLENGE 4: ROBUST COMPUTER VISION

- Detector performance is typically shown using precision-recall (PR) or receiver operating characteristic (ROC) curves
- No sensor is perfect!
- P_D = Probability of Detection
- P_{FA} = Probability of False Alarm



NHTSA LEVELS OF AUTONOMY

No-Automation (Level 0): Driver is in complete and sole control of brake, steering, throttle, and motive power at all times.

Function-specific Automation (Level 1): Automation of one or more specific control functions

Combined Function Automation (Level 2): Automation of at least two primary control functions that operate in unison to relieve the driver of those functions.

NHTSA LEVELS OF AUTONOMY (CONT'D)

Limited Self-Driving Automation (Level 3): Enables the driver to cede full control of all safety-critical functions under certain conditions. The driver is expected to be available for occasional control, but with sufficiently comfortable transition time

Full Self-Driving Automation (Level 4): The vehicle performs all safety-critical driving functions and monitor roadway conditions for an entire trip. The driver is not expected to be available for control at any time. This includes both occupied and unoccupied vehicles.

THE BIG QUESTIONS GOING FORWARD

For Level 3 (partial autonomy):

Can humans be trusted to take control when necessary?

For Level 4 (full autonomy):

Can sufficiently robust perception be achieved for safe and reliable operation when human operator intervention is not possible?
(P_D approaching 1.0, P_{FA} approaching 0.0)

CONCLUSION

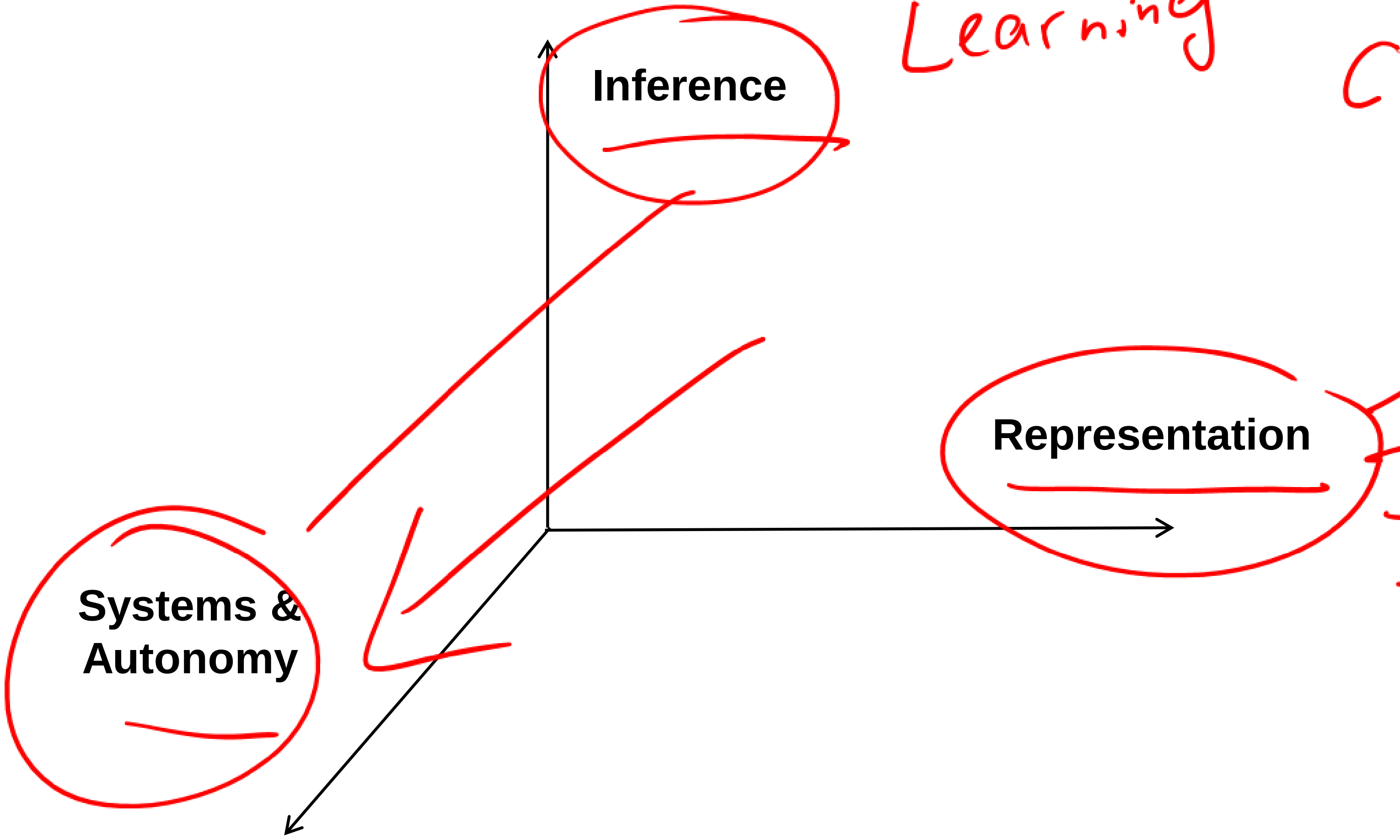
Its an exciting time to be working in mobile sensing and SLAM Engineers that can build real-time 3D perception, navigation and motion planning systems are in high demand:

- Virtual Reality
- Mobile Devices
- Self-Driving Vehicles
- Drones

Additionally, the connection to the Internet of Things offers a much broader set of opportunities spanning many industries

Knowing "what is where in the world?" is vitally important

WHY IS SLAM DIFFICULT?



Learning

$\lim_{t \to \infty}$

Cloud

objects
Semantics
Text
Dense 3D

The Internet of Things: Roadmap to a Connected World

THANK YOU!

John J. Leonard

Samuel C. Collins Professor of Mechanical and Ocean Engineering

Computer Science and Artificial Intelligence Laboratory (CSAIL)
Massachusetts Institute of Technology

END OF MODULE - Technologies