Mapping, Localization, and Self Driving Vehicles

John Leonard
Samuel C. Collins Professor of Mechanical and Ocean Engineering
Massachusetts Institute of Technology
August 20, 2015

Robotics Afternoon, The Center for Brains, Minds & Machines
MBL
Outline

• Technical Challenges for Self-Driving Cars

• A Historical Perspective on Robot Mapping and Localization

• Object-based Mapping
Education:
• University of Pennsylvania, BSEE (1987)
• University of Oxford, DPhil (1991)

History of MIT Positions:
• MIT Sea Grant AUV Lab (1991-1996)
• Dept. of Ocean Engineering (1996-2004)
• Dept of Mechanical Engineering 2005-present
• Artificial Intelligence Laboratory (2002-2004) and CSAIL (2005-present)

Current Position:
• Associate Department Head for Research, MIT MechE

Research Interests:
• Mapping and Localization for Autonomous Vehicles; Marine Robotics
Autonomous Underwater Vehicles

- Must compute a navigation solution in real time to achieve mission objectives and ensure safe operation
- Acoustic communications are hampered by severe bandwidth constraints
Fallon, et al., Relocating Underwater Features Autonomously Using Sonar-Based SLAM

© IEEE. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.
Source: Fallon, Maurice F., John Folkesson, Hunter McClelland, and John J. Leonard.
Outline

• Technical Challenges for Self-Driving Cars

• A Historical Perspective on Robot Mapping and Localization

• Object-based Mapping
“Tesla’s Musk minimized the challenges necessary to achieve a future where self-driving cars will become commonplace. ‘I view it as a solved problem’, said Musk, who compared autonomous cars with elevators that used to require operators, but are now self-service.”

Tesla CEO Elon Musk and Nvidia CEO Jen-Hsun Huang declare self-driving cars “solved”

by Bradley Berman    MARCH 18, 2015, 2:41 PM EDT

© Fortune. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.
I think Elon Musk is wrong ...
MIT DARPA Urban Challenge Team (2006-2007)

Leonard et al., JFR 2008; Karaman and Frazzoli, IJRR 2011; Huang et al., AR 2009
MIT Land Rover LR3 (Talos)

- 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

Figure removed due to copyright restrictions.
2007 Urban Challenge Results

Initially 89
Site Visit 53
Invited to NQE 35
Qualified 11
Finished 6

CMU 1st place
Stanford 2nd place
Virginia Tech 3rd place
MIT 4th place
2007 DARPA Urban Challenge – Collision between MIT and Cornell
2007 DARPA Urban Challenge – Collision between MIT and Cornell

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

2007 DARPA Urban Challenge – Collision between MIT and Cornell

Perception-based Navigation (PhD Thesis of Albert Huang, supervised by Prof. Seth Teller)

Playback speed: 3.3x

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


“Multi-Sensor Lane Finding in Urban Road Networks”, Albert Huang, David Moore, Matthew Antone, Edwin Olson, Seth Teller, RSS 2008
The Google Car is an amazing research project that might one day transform mobility

The technology of the Google Car, however, has been over-hyped and is poorly misunderstood

This has led many people to say that self-driving is a “solved” problem

“Just because it works for Google”, doesn’t mean it will work for everyone else
Just because it works for Google (using Lidar and precision a priori maps) doesn’t mean it will work for everyone else (using vision)

- Localization using *a priori* maps vs. GPS and vision
- Level 4 (100% autonomy; no human controls) vs. Level 3 (99% autonomy; human must be ready to take control)
- Mountain View CA vs. other locations (e.g. Boston)
Difficult Situations for Self-Driving Vehicles (in Boston)

Left turn across traffic

Changes to road surface

Traffic cops, crossing guards, police/fire

Winter weather
Police Officers Directing Traffic

Brookline, MA – November, 2013
Changes in Road Surface Appearance

Driving from
Boston to
Cambridge

Nov 08\textsuperscript{th}, 2013

Nov 12\textsuperscript{th}, 2013
Unsolved Challenges: Adverse Weather
Google: Lidar Localization with an a priori map

[Image]

© 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/.

https://plus.google.com/+GoogleSelfDrivingCars
What do you see in this Picture?
The Big Questions Going Forward

Technical Challenges:

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

• Technical Challenges for Self-Driving Cars

• A Historical Perspective on Robot Mapping and Localization

• Object-based Mapping
The Nobel Assembly at Karolinska Institutet has today decided to award The 2014 Nobel Prize in Physiology or Medicine with one half to John O’Keefe and the other half jointly to May-Britt Moser and Edvard I. Moser for their discoveries of cells that constitute a positioning system in the brain.
Simultaneous Localization and Mapping (SLAM)

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

Johannsson et al, ICRA 2013
Localization with an *a priori* map (Polaroid Sonar)
Why is SLAM Difficult?

Inference

Systems & Autonomy

Representation
Simultaneous Localization and Mapping (SLAM)

Goal: Generate globally consistent map from noisy local sensor data
Concurrently estimate the vehicle trajectory
Probabilistic Formulation of SLAM (Assume Data Association is Known)

Bayesian Belief Network:

Known measurements, want variables \to probabilistic inference problem
Q: What is the most important thing I learned up thru 2012?
A: Maintaining \textit{Sparsity} in the underlying representation is critical

\begin{align*}
\text{argmax}_\theta & \prod_i p_i(\theta) \\
\text{argmin}_\theta & \sum_i \| h_i(\theta) \|_2^2 \\
\text{argmin}_\theta & \| A\theta - b \|_2^2
\end{align*}

(Johannsson et al, ICRA 2013)

Pose Graph Optimization Algorithms:
[Lu&Milios 97, Konolige 04, Folkesson 04, Eustice 05, Frese 06, Olson 06, Dellaert 06, Grisetti et al. 10]

Question: What is the most important thing that I learned about SLAM since 2012?
Answer: Building and Maintaining Dense 3D Representations is possible

© Thomas Whelan, Michael Kaess, Mourice Fallon, Hordur Johannsson, John Leonard and John McDonald. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Whelan et al. RSS 2012 RGB-D Workshop (Sydney, Australia)
Kintinuous (Whelan et al. ‘12, ’13, ‘14)

- Extension of KinectFusion (Newcombe, et al. ISMAR ’11)
- Treat volumetric model as a cyclical buffer.
  - As region leaves the range of the buffer, extract the corresponding surface data.
  - As region enters the range of the buffer, initialise and track the new data.
- Connect with Pose Graph SLAM techniques to achieve loop closure

Kintinuous (Whelan et al. ‘12, ’13, ‘14)

Kintinuous Processing Pipeline (“Cloud Slices” connected to pose graph SLAM optimization)

Figure removed due to copyright restrictions.

“Deformation-based Loop Closure for Large Scale Dense RGB-D SLAM” by T. Whelan, M. Kaess, J. Leonard and J. McDonald, IROS 2013
Real-time Dense Loop Closure using Mesh Deformation

Deformation-based Loop Closure for Large Scale Dense RGB-D SLAM

Thomas Whelan, John McDonald
Department of Computer Science, NUI Maynooth

Michael Kaess, John J. Leonard,
Computer Science and Artificial Intelligence Laboratory (CSAIL),
Massachusetts Institute of Technology (MIT)

Whelan et al., IROS 2013 and IJRR 2014
Figure removed due to copyright restrictions.
Outline

• Technical Challenges for Self-Driving Cars

• A Historical Perspective on Robot Mapping and Localization

• Object-based Mapping
Vision for Future Research in Mobile Sensing

• David Marr:

  *Vision is the process of discovering from images what is present in the world and where it is.*

• We need an *object-based* understanding of the environment that facilitates life-long learning

• Let’s build rich representations that leverage knowledge of location to better understand about objects, and concurrently uses information about objects to better understand location

  – Sudeep Pillai: Monocular SLAM Supported Object Recognition (presented at RSS 2015)
Pillai and Leonard, RSS 2015: SLAM-Supported Object Recognition

MONOCULAR SLAM SUPPORTED
OBJECT RECOGNITION

Sudeep Pillai & John J. Leonard
Computer Science and Artificial Intelligence Lab
Massachusetts Institute of Technology
The hippocampus develops stable object-location mappings.

“what” stream

perirhinal cortex

parahippocampal cortex

“where” stream

lateral entorhinal cortex

medial entorhinal cortex

hippocampus

Conclusion and Future Research Challenges

Goals:

• My dream is to achieve *persistent autonomy* and *lifelong map learning* in highly dynamic environments

• Can we robustly integrate mapping and localization with real-time planning and control?

Open Questions:

• Robustness – we would love to have guarantees of performance, but we do not have them for most approaches

• Representation – how can we integrate many different types?

• We need dynamic scene understanding and robust vision (recent work in computer vision is very exciting, but current precision-recall curves indicate we have a long way to go)
Some Questions for Neuroscience in Relation to Spatial Memory and Navigation

• Do biological representations support multiple location hypotheses?
• Is there evidence for an “experience map” in the brain?
• Does “pose graph optimization” occur?
  – On-line during path execution?
  – Off-line after path execution?
• What really are the grid-cells doing?
  – Path integration only? Or path correction as well?
  – How is the correction performed?
• Could grid-cells serve as an “indexing mechanism” to facilitate what functions as a “search database”, providing a mechanism to store pointers to “what?” vs. “where?” information?
Resource: Brains, Minds and Machines Summer Course
Tomaso Poggio and Gabriel Kreiman

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.