Learning to Navigate

Vladlen Koltun
Intel Labs
Vision, like other sensory functions, has its evolutionary rationale rooted in improved motor control. Although organisms can of course see when motionless or paralyzed, the visual system of the brain has the organization, computational profile, and architecture it has in order to facilitate the organism's thriving at the four Fs: feeding, fleeing, fighting, and reproduction.

- Churchland, Ramachandran, and Sejnowski

A critique of pure vision
Navigation via SLAM?

- SLAM addresses perception, not action
- SLAM does not tell us how to explore, navigate, accomplish goals
  - The planner will take care of it…
  - But planners assume that the map is accurate and sufficiently complete
  - When this assumption breaks, the whole system can fail
Intel® RealSense™ Tracking Module T260 Features:

Wide Field of View Stereo Fisheye Pair:
- 170 deg circular FOV on 800 pix (dia) monochrome sensor

MEMS IMU
- Bosch BMI055 @100-400Hz

Vision Processor
- Myriad 2

Complete 6DoF Solution
- Entire SLAM algorithm runs onboard; poses streamed out over USB

Low Power
- ~1 watt
Navigation via SLAM?

- High “dynamic range” (e.g., driving to work, flying to ECCV)
- Poor localization (e.g., perceptual aliasing, walking through a field)
- Environments that change over time
- Reflexive perception-action shortcuts (e.g., duck behind corner to avoid danger)
- Prior knowledge (navigate to a grocery store in a new city, navigate to a restroom in a restaurant)
Classic sensorimotor control pipeline

- Perception system constructs a map of the environment
- Planner generates waypoints
- Continuous controller (e.g., proportional-derivative) actuates the system towards the waypoints
Prize for place cells

Discoverers of brain’s navigation system get physiology Nobel.

BY ALISON ABBOTT & EWEN CALLAWAY

Brain cells that make up the biological equivalent of a satellite-navigation system have garnered three scientists the 2014 Nobel Prize in Physiology or Medicine. The discovery of the cells sheds light on one of neuroscience’s great mysteries — how we know where we are in space.

John O’Keefe of University College London won half of the prize for his discovery in 1971 of ‘place’ cells in the hippocampus, a part of the brain associated with memory. Edvard and May-Britt Moser, who are married and jointly run a lab at the Kavli Institute for Systems Neuroscience in Trondheim, Norway, share the other half for their 2005 discovery of ‘grid’ cells in an adjacent brain structure, the entorhinal cortex. Along with other navigation cells, grid and place cells allow animals to keep track of their position. Both cell types were discovered

John O’Keefe, co-winner of a 2014 Nobel prize.
THE PSYCHOLOGICAL REVIEW

COGNITIVE MAPS IN RATS AND MEN

BY EDWARD C. TOLMAN

University of California

I shall devote the body of this paper to a description of experiments with rats. But I shall also attempt in a few words at the close to indicate the significance of these findings on rats for the clinical behavior of men. Most of the rat investigations, which I shall report, were carried out in the Berkeley laboratory. But I shall also include, occasionally, accounts of the behavior of non-Berkeley rats who obviously have misspent their lives in out-of-State laboratories. Furthermore, in reporting our Berkeley experiments I shall have to omit a very great many. The ones I shall talk about were carrying the food box and eats. This is repeated (again in the typical experiment) one trial every 24 hours and the animal tends to make fewer and fewer errors (that is, blind-alley entrances) and to take less and less time between start and goal-box until finally he is entering no blinds at all and running in a very few seconds from start to goal. The results are usually presented in the form of average curves of blind-entrances, or of seconds from start to finish, for groups of rats.

All students agree as to the facts. They disagree, however, on theory and explanation.
Do not ask whether they have a cognitive map, but how they find their way about

N.J. Mackintosh

University of Cambridge, UK

The publication of "The hippocampus as a cognitive map" (O'Keefe & Nadel, 1978) has had a remarkable impact, stimulating a huge amount of both behavioural and neurobiological research on spatial learning and memory, involving both laboratory and field studies, and employing a variety of novel techniques. The reviews of this general area of research provided by the previous contributors to this special issue attest to the progress that has been made since 1978. No one would now doubt that the hippocampus is implicated in the use of configurations of landmarks to locate a goal - in both mammals and birds, although the precise nature of that implication remains a matter of much speculation and debate. On balance, however, the behavioural evidence does not seem to have supported O'Keefe and Nadel's original hypothesis that true spatial or locale learning is quite distinct from simple associative learning and depends on the establishment of a cognitive map of the environment.
Semi-parametric Topological Memory for Navigation

Nikolay Savinov*  Alexey Dosovitskiy*  Vladlen Koltun
ETH Zürich  Intel Labs  Intel Labs

Abstract

We introduce a new memory architecture for navigation in previously unseen environments, inspired by landmark-based navigation in animals. The proposed semi-parametric topological memory (SPTM) consists of a (non-parametric) graph with nodes corresponding to locations in the environment and a (parametric) deep network capable of retrieving nodes from the graph based on observations. The graph stores no metric information, only connectivity of locations corresponding to the nodes. We use SPTM as a planning module in a navigation system. Given only 5 minutes of footage of a previously unseen maze, an SPTM-based navigation agent can build a topological map of the environment and use it to confidently navigate towards goals. The average success rate of the SPTM agent in goal-directed navigation across test environments is higher than the best-performing baseline by a factor of three.
Task: memory-based navigation

Walkthrough sequence (~5 min)

Current observation

Goal observation

Savinov et al., ICLR 2018
Topological memory for navigation

Semi-parametric topological memory

- Retrieval network
- Memory graph

(a) Localization
(b) Planning
(c) Waypoint selection

Current observation
Goal observation

Waypoint observation

Savinov et al., ICLR 2018
After being shown a 5-minute traversal video of a *previously unseen* maze...
On Evaluation of Embodied Navigation Agents

Peter Anderson  
Australian National University

Angel Chang  
Princeton University

Devendra Singh Chaplot  
Carnegie Mellon University

Alexey Dosovitskiy  
Intel Labs

Saurabh Gupta  
UC Berkeley

Vladlen Koltun  
Intel Labs

Jana Kosecka  
George Mason University

Jitendra Malik  
UC Berkeley and Facebook

Roozbeh Mottaghi  
Allen Institute for AI

Manolis Savva  
Princeton University

Amir R. Zamir  
Stanford and UC Berkeley

Abstract

Skillful mobile operation in three-dimensional environments is a primary topic of study in Artificial Intelligence. The past two years have seen a surge of creative work on navigation. This creative output has produced a plethora of sometimes incompatible task definitions and evaluation protocols. To coordinate ongoing and future research in this area, we have convened a working group to study empirical methodology in navigation research. The present document summarizes the consensus recommendations of this working group. We discuss different problem statements and the role of generalization, present evaluation measures, and provide standard scenarios that can be used for benchmarking.

Simultaneous Localization and Mapping (SLAM) [6], which focuses on building a map of the environment and localizing the agent within the map. Traditional SLAM systems largely focus on constructing metric maps using geometric techniques. Navigation itself (i.e., what to do, where to go) is rarely considered. Representations constructed by SLAM systems are often not suitable for traditional motion planning methods and are prone to error when the environment changes over time.

In contrast, biological systems appear to use more flexible representations, can robustly navigate without precise localization or a metric map, and can bring substantial prior knowledge to bear when planning in previously unseen environments.
Summary

- What matters is action
- SLAM is solving a problem that is both unnecessarily hard and incomplete
- Biological motivation suggests alternative approaches
- Semi-parametric topological memory (SPTM)
- Working group on empirical methodology in navigation research
Thank you