Autonomous Robotics

Simultaneous Localization and Mapping (GraphSLAM and Outlook)

Josiah Hanna
University of Wisconsin — Madison

Announcements

- Homework #3 has been released.
- Reading assignment for next week (Kinematics) due Monday at noon.
- Midterm in 3 weeks.

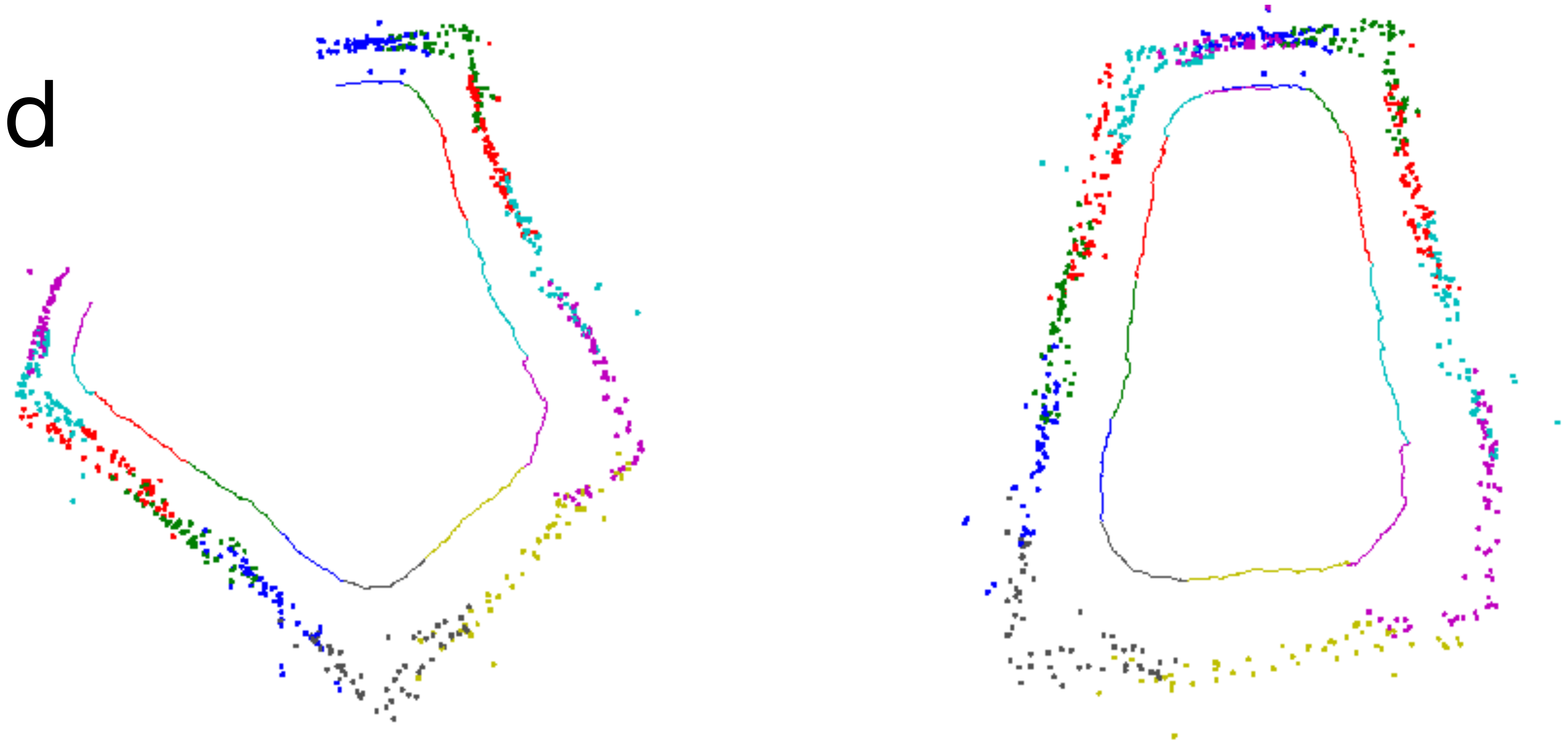
Learning Outcomes

After today's lecture, you will be able to:

- Identify open questions in SLAM research and applications.
- Describe approaches to visual and semantic SLAM.
- Discuss opportunities and challenges with integrating learning into state estimation and SLAM.

Robustness

- Hardware failure
- Robustness to incorrect loop closure
- Dynamic maps
- Integration of the front-end and back-end



Scalability

- Need to forget?
- Handling new types of sensors?
  - E.g., event cameras
- Integrating semantic information
- Handling resource constraints

Introduction to Visual SLAM

Visual SLAM: perform SLAM with a camera

Camera types:

- Monocular: cannot determine size of objects in image.
- Stereo: two cameras, enables sizing.
- RGB-D: also provides depth information.

Visual SLAM Front-end

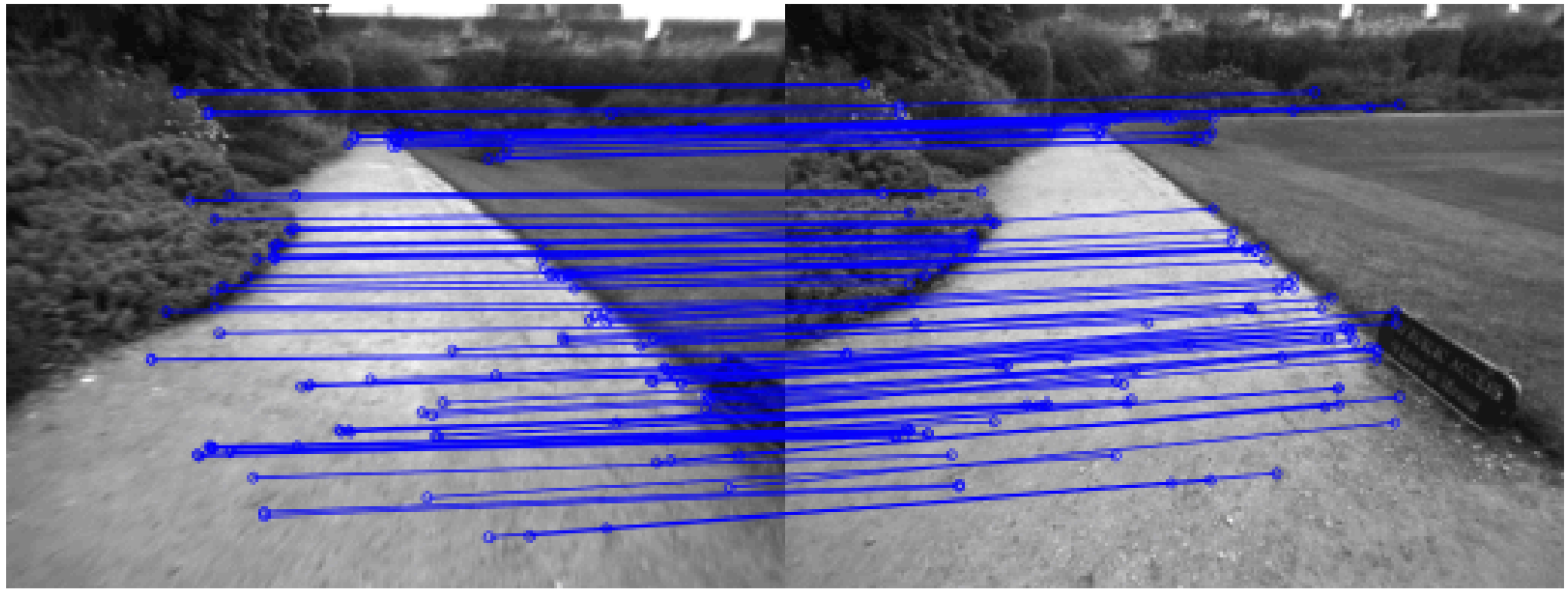
Compute distinct features and track them across images.

  Challenge: need to have distinct visual features (e.g., SIFT, SURF, ORB)

Alternatively, track pixels across frames.

  More sensitive to lighting changes, more robust to lack of features in environment.

Loop-closure: maintain a database of previous visual scenes and match against it.



ORB-SLAM: a Versatile and Accurate Monoculat SLAM System

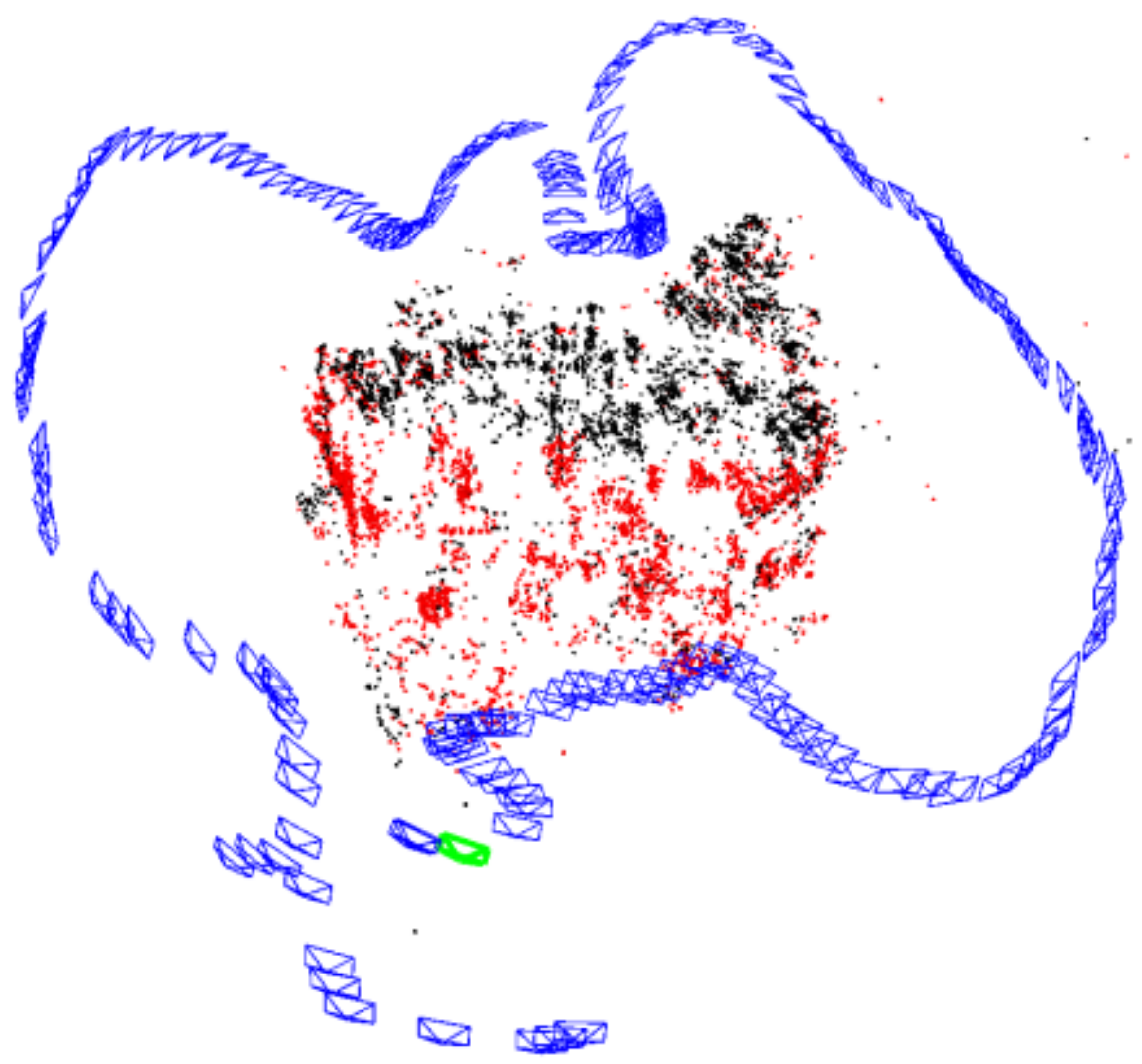
Visual SLAM: Back-end

Bundle adjustment: similar to GraphSLAM, used in computer vision to optimize camera poses and locations of features.

- Run an optimization to find most consistent location of features and sequences of camera poses.

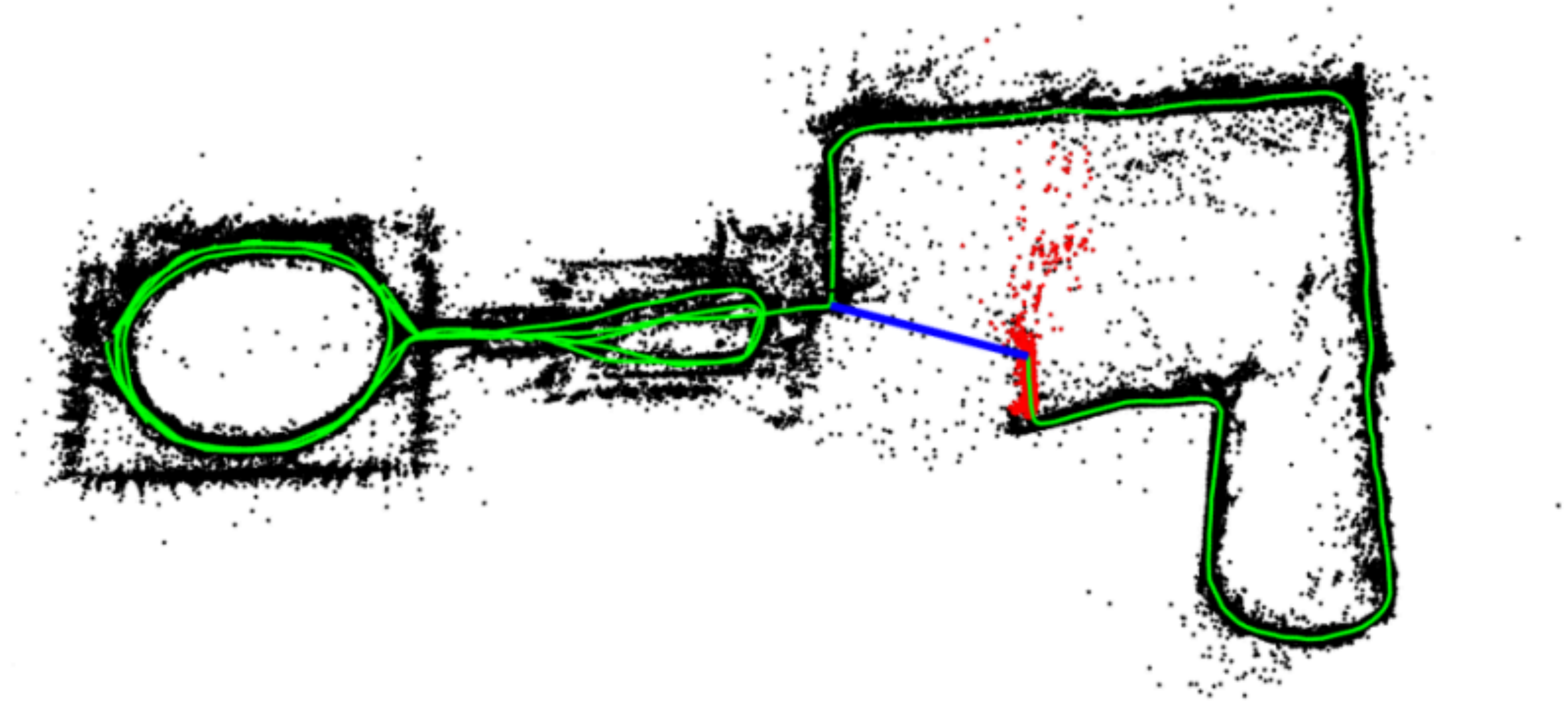
Cannot keep all observations in memory.

- Instead, use bundle adjustment to produce a local map.
- Merge local map with global map using GraphSLAM.



ORB-SLAM: a Versatile and Accurate Monoculat SLAM System

ORB-SLAM Example



Before loop closure

After loop closure

ORB-SLAM: a Versatile and Accurate Monoculat SLAM System

Visual SLAM: Challenges

Dynamic obstacles and occlusion

Map is just a collection of low-level image features (point cloud)

Semantic SLAM

So far we have discussed metric SLAM.

Not discussed in detail: topological SLAM

To make robots useful, they need semantic information about their environment.

Semantic SLAM: label a metric (or topological) map with semantic information.

Example: a landmark is in the kitchen.

How might this help metric SLAM?

Semantic Computer Vision

Semantic segmentation: classify all pixels in an image.

Object detection: Put bounding boxes around specific objects.

Instance segmentation: classify every pixel of distinct objects.

Map representations

Landmark maps: label each landmark with semantic label.

Occupancy grid maps: label each grid cell with semantic label.

Object-oriented: cluster map points together and provide one label to all of them (e.g., chair at x,y,z)

Use of semantic information

Semantic information can aid metric SLAM.

- Use semantic information to remove dynamic objects (e.g., person detection)
- Addressing scale ambiguity

Limitations of Semantic Slam

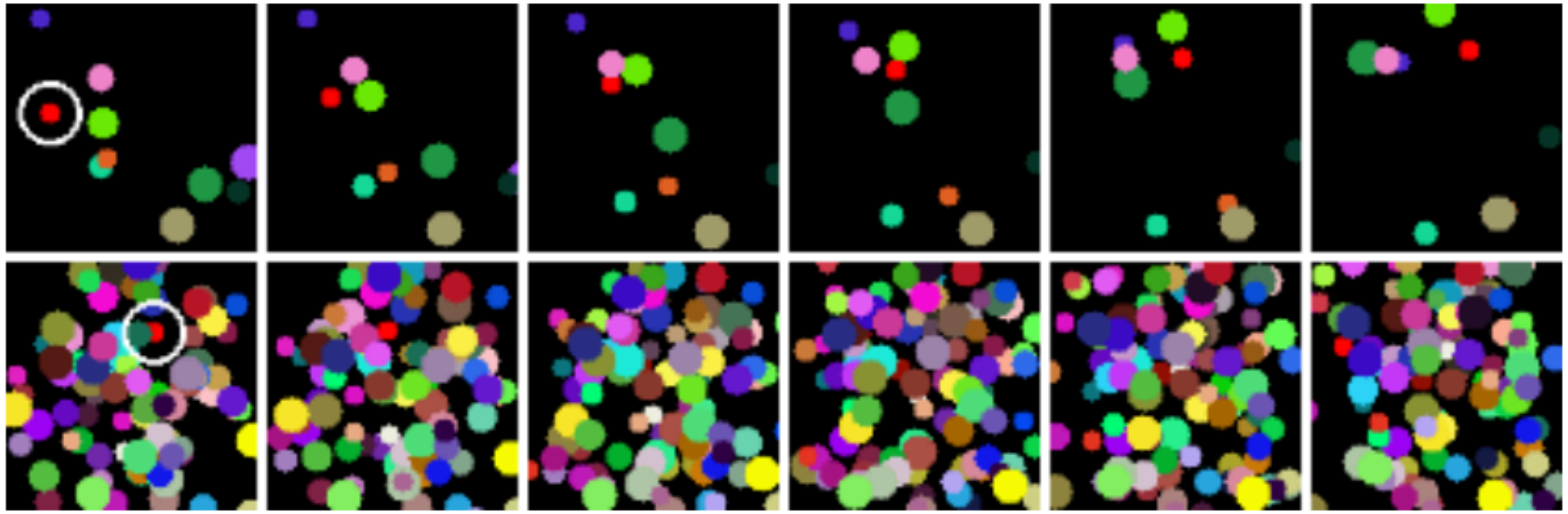
More computation needed to run modern neural networks

Learning State Estimation

So far we have assumed given observation and motion models.

Alternative idea: robot learns these through experience.

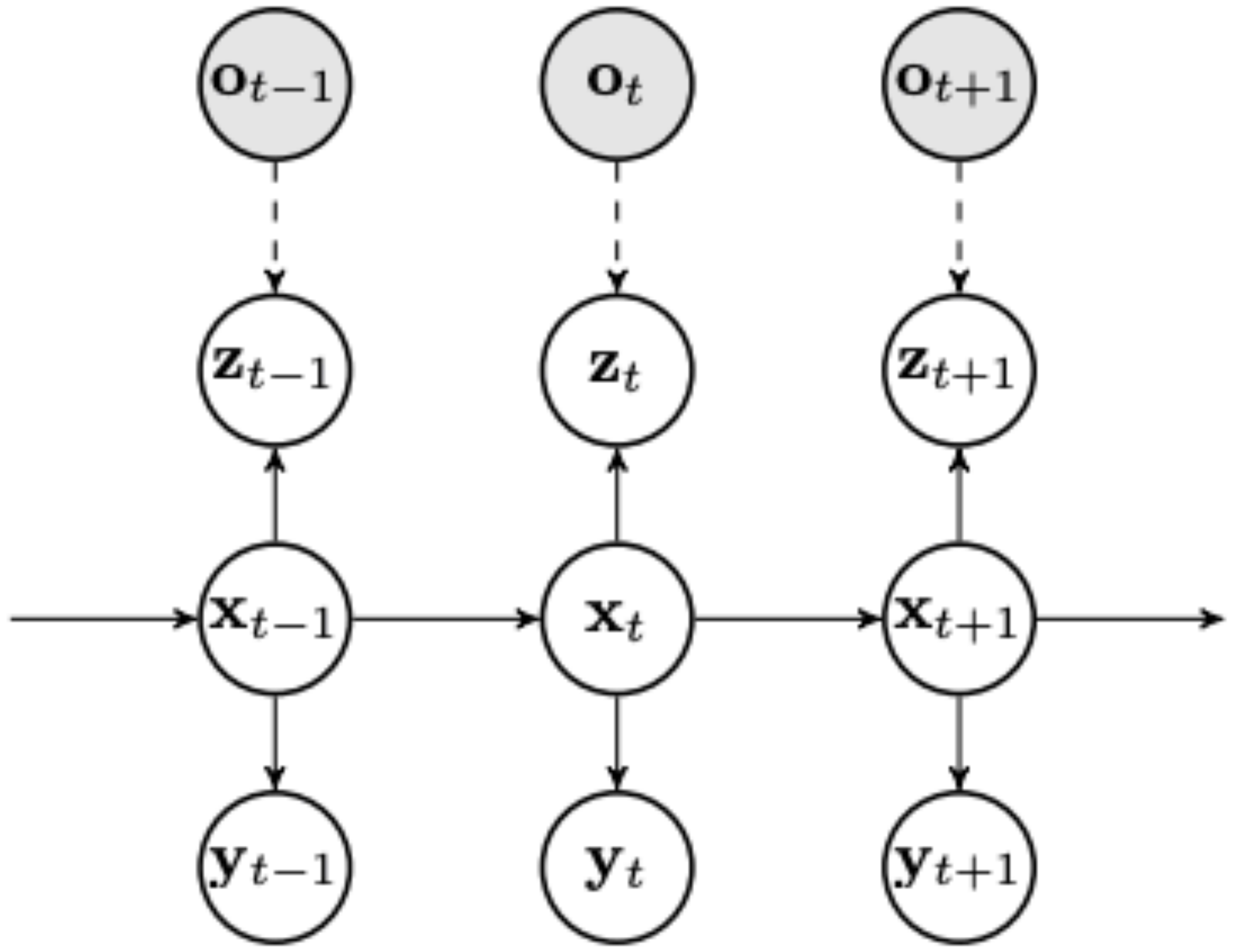
Potentially useful for handling high-dimensional, complex observations

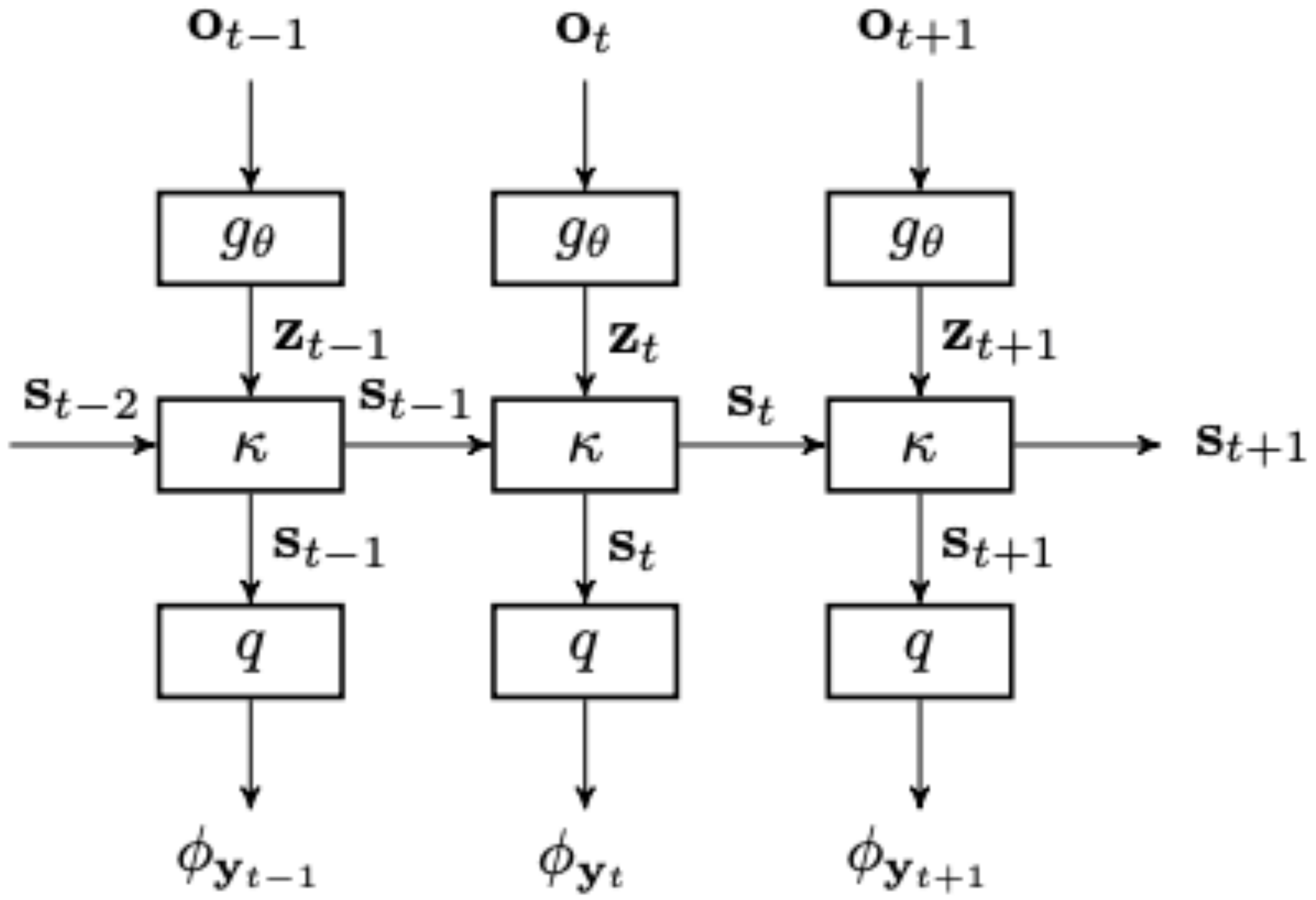


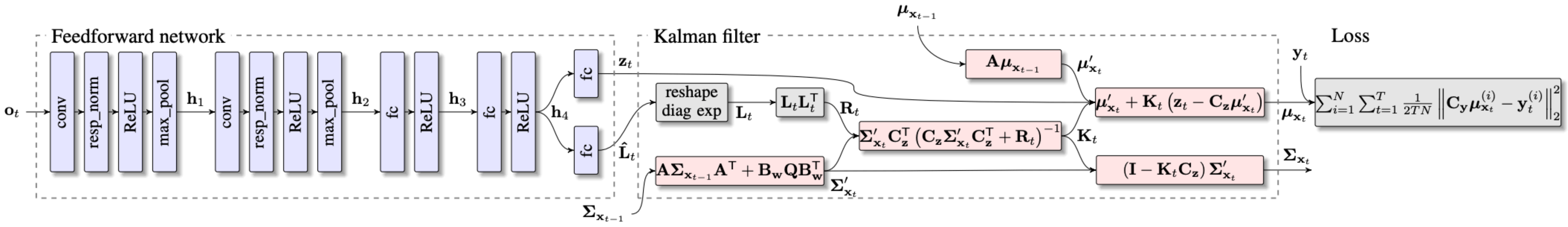
Tracking with distractions

Backprop KF: Learning Discriminative Deterministic State Estimators

Integrating deep learning into Kalman Filters







Backprop KF: Learning Discriminative Deterministic State Estimators

Integrating deep learning into Kalman Filters

State Estimation Model	**# Parameters**	**RMS test error** $\pm\sigma$
feedforward model	7394	0.2322 ± 0.1316
piecewise KF	7397	0.1160 ± 0.0330
LSTM model (64 units)	33506	0.1407 ± 0.1154
LSTM model (128 units)	92450	0.1423 ± 0.1352
BKF (ours)	**7493**	$\mathbf{0.0537} \pm \mathbf{0.1235}$

Kalman filter + deep models improves upon solely using deep learning tracking.

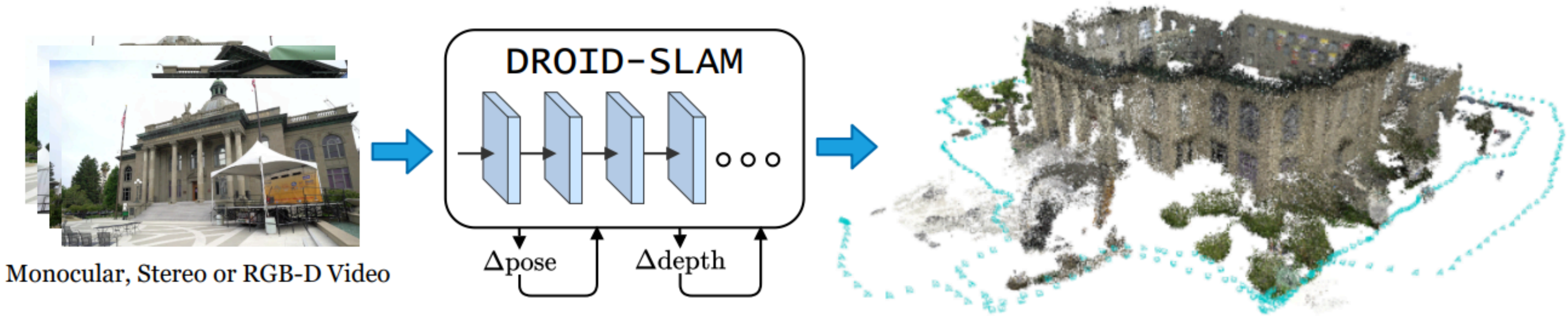
Backprop KF: Learning Discriminative Deterministic State Estimators

End-to-end SLAM

Similar ideas developed for SLAM, e.g., DROID-SLAM.

Key part of DROID-SLAM is a learned update operator (learned bundle adjustment)

Key limitation: require datasets with ground-truth to train the system.



Backprop KF: Learning Discriminative Deterministic State Estimators

Summary

- Discussed challenges in SLAM
- Introduced basics of visual and semantic SLAM
- Described ways to integrate learning approaches into SLAM and state estimation.

Action Items

- Kinematics reading for next week; send a reading response by 12 pm on Monday.
- SLAM assignment has been released.