Certifiable Perception for Robots and Autonomous Vehicles:

From Robust Algorithms to Robust Systems

Luca Carlone
Assistant Professor
Massachusetts Institute of Technology

@lucacarlone1

SPARK Lab

MIT AEROASTRO

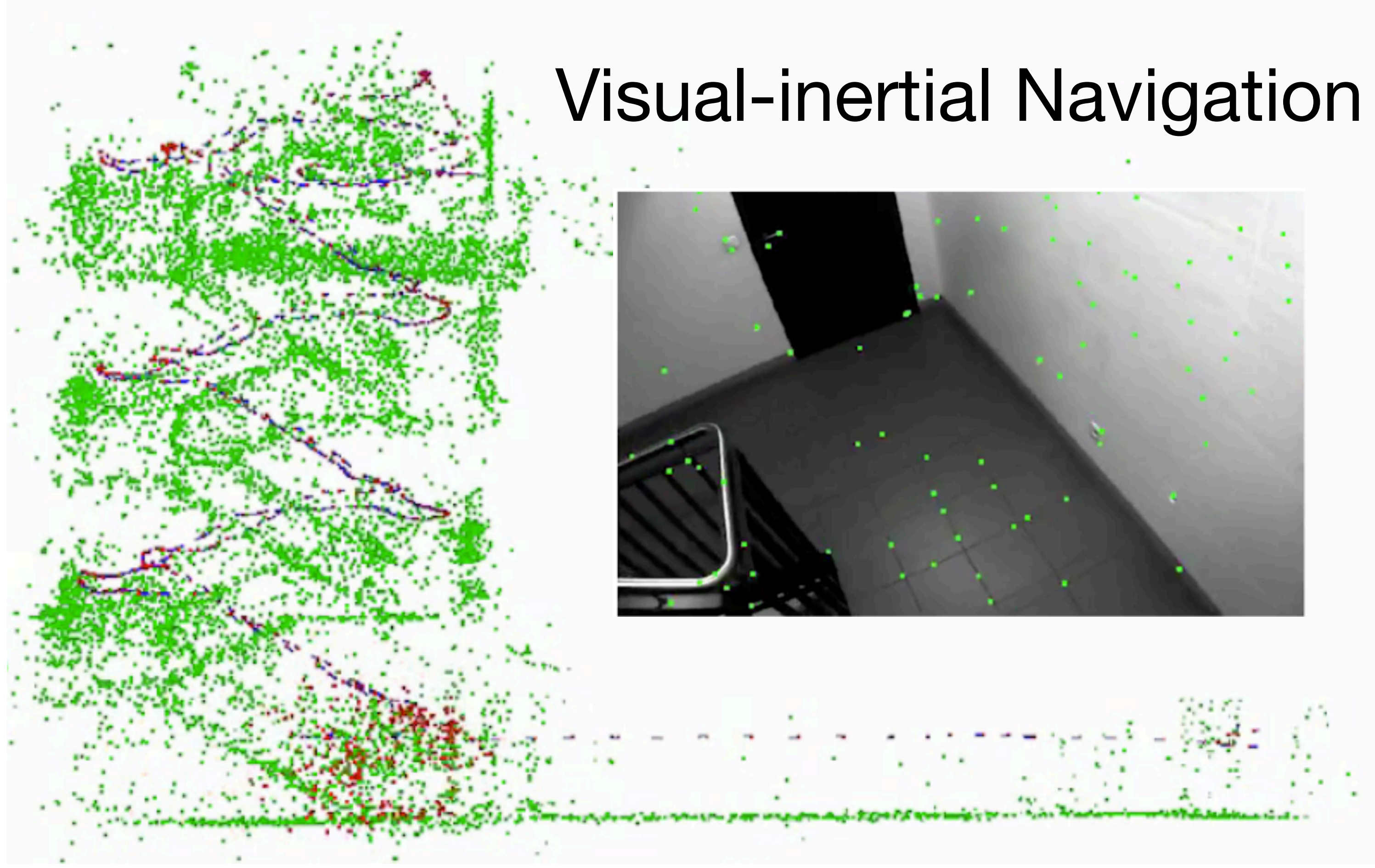
IPAM - Workshop II: Safe Operation of Connected and Autonomous Vehicle Fleets

LIDS

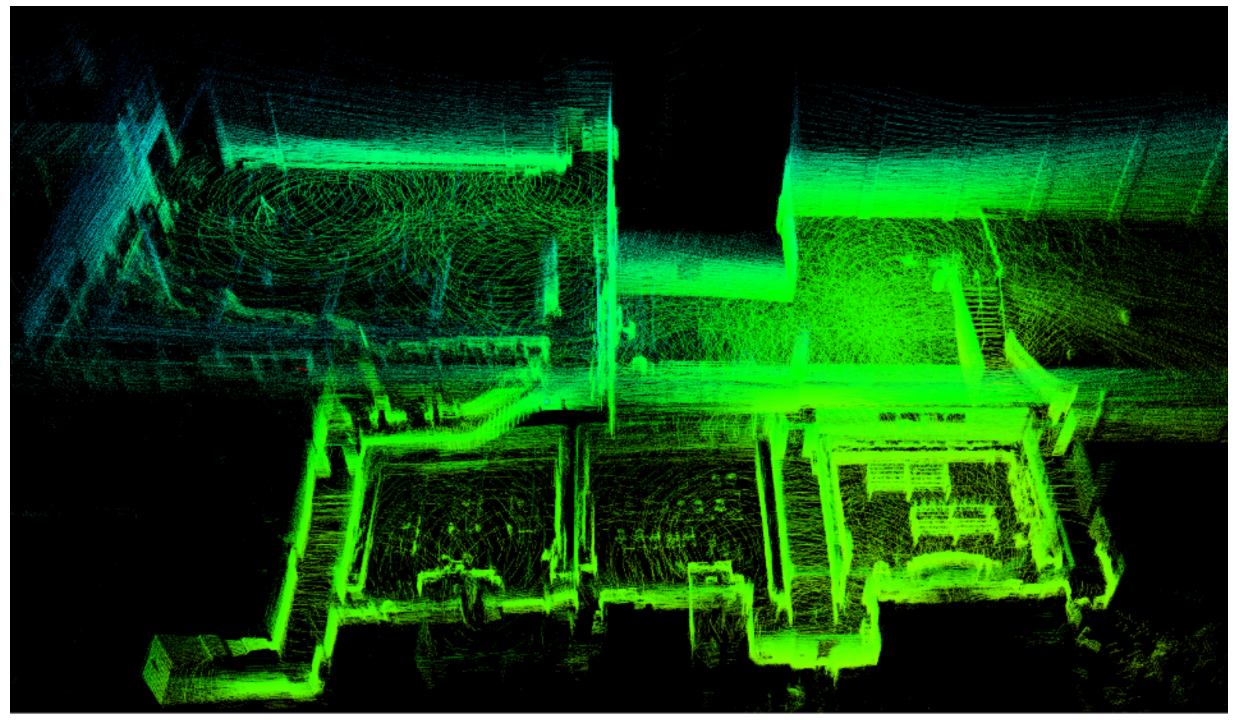
SPARK Lab — Sensing Perception Autonomy and Robot Kinetics

Robust Perception, Localization and Mapping

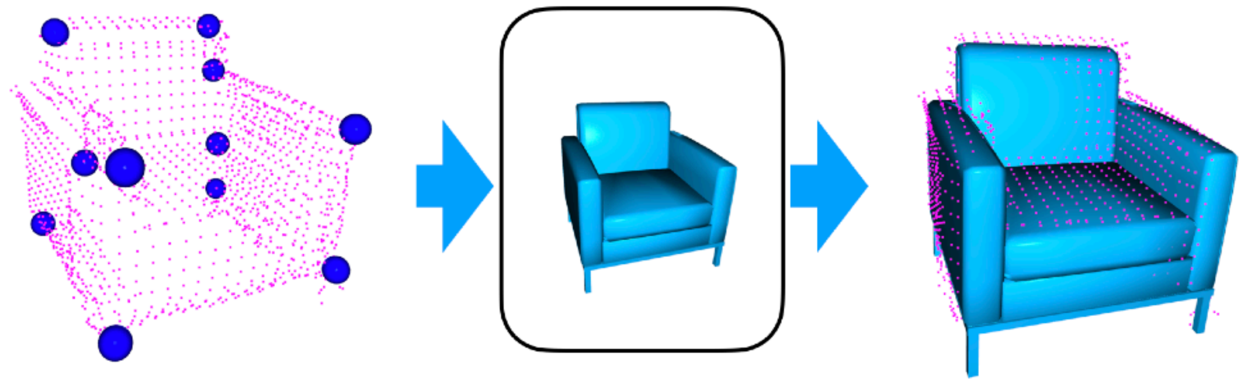
Visual-inertial Navigation



Lidar-based SLAM

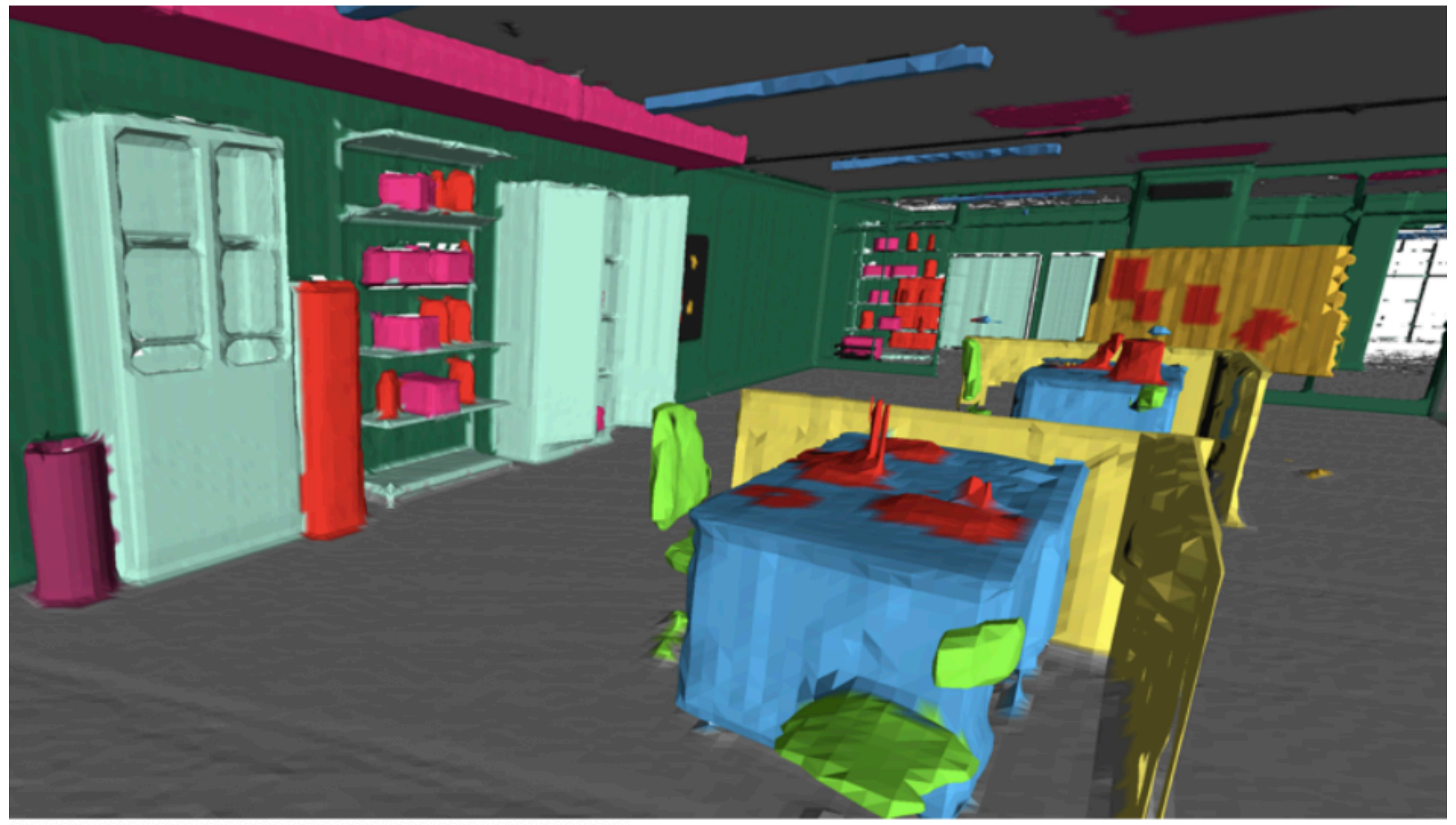


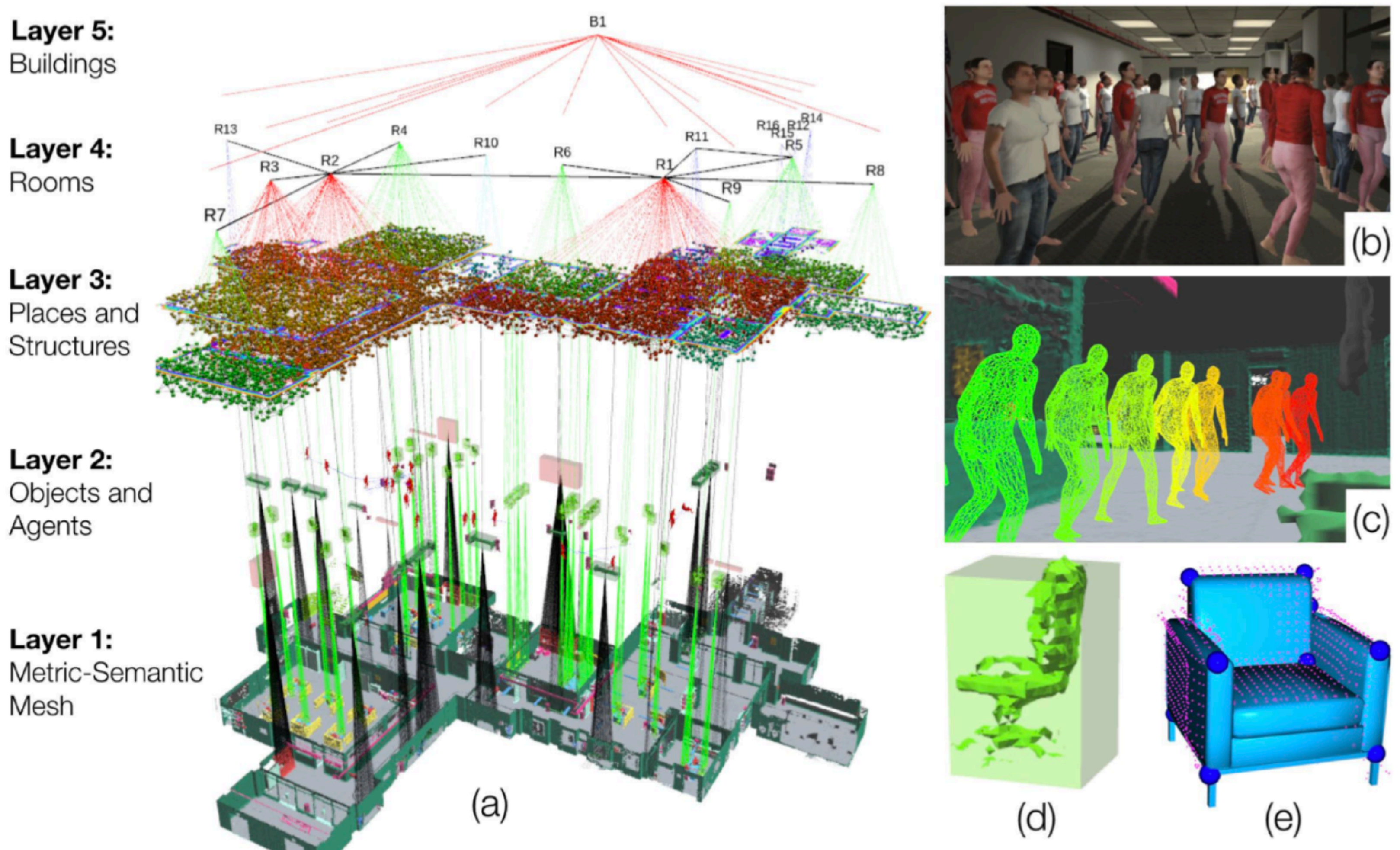
Certifiable Algorithms



High-level Scene Understanding (Spatial AI)

Kimera: Metrics-semantic SLAM

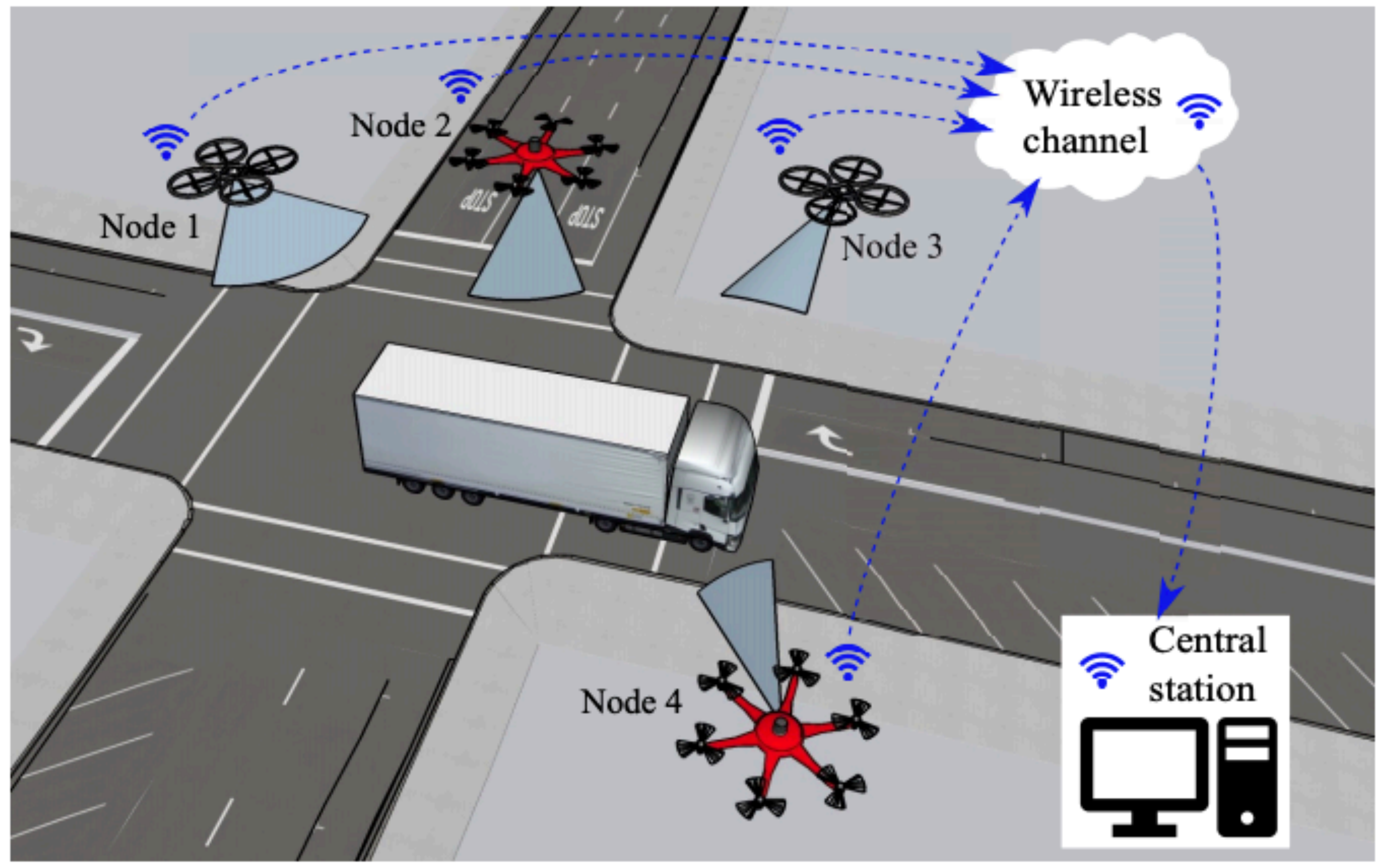


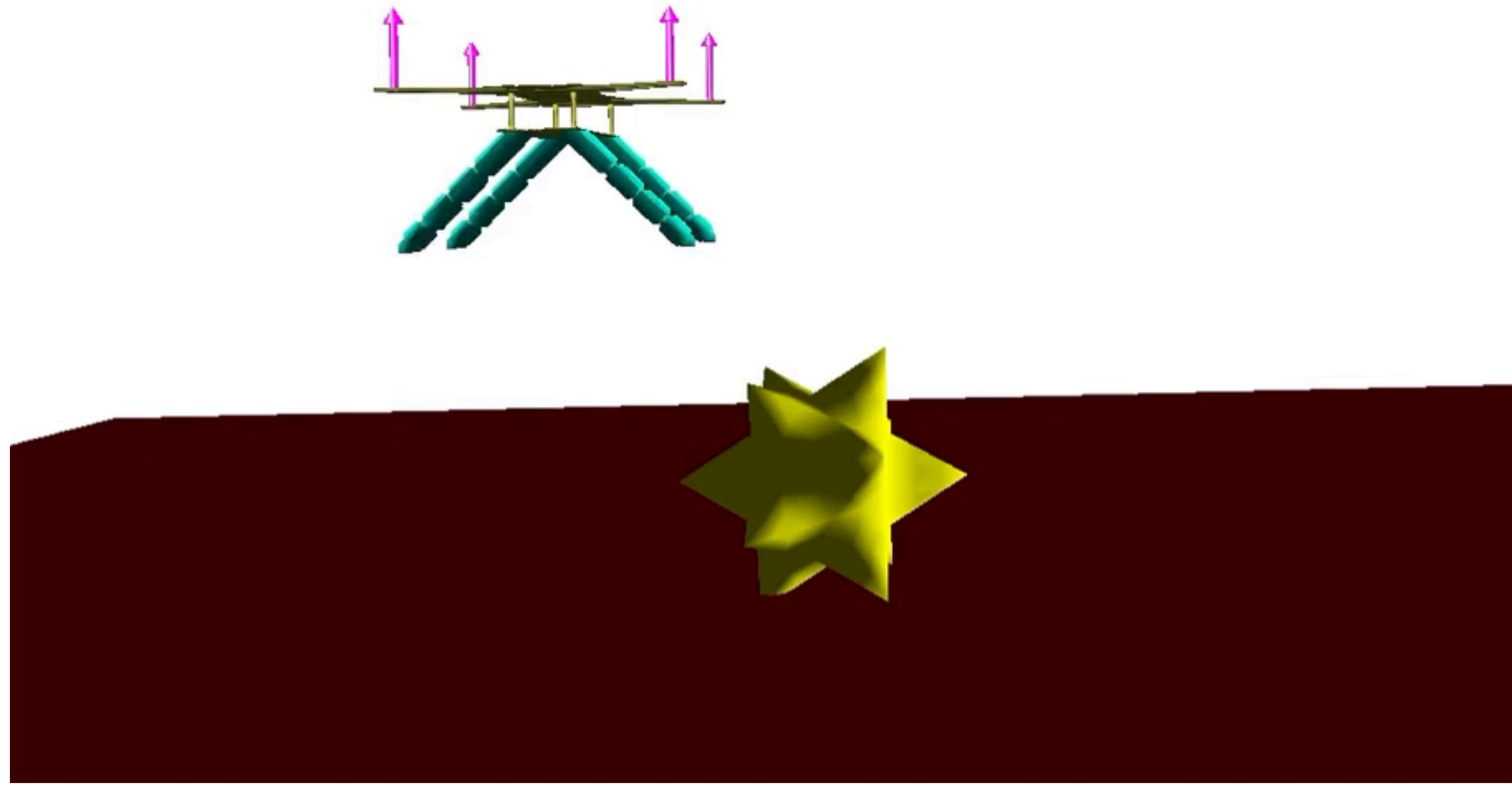


3D Dynamic Scene Graphs

Co-design

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

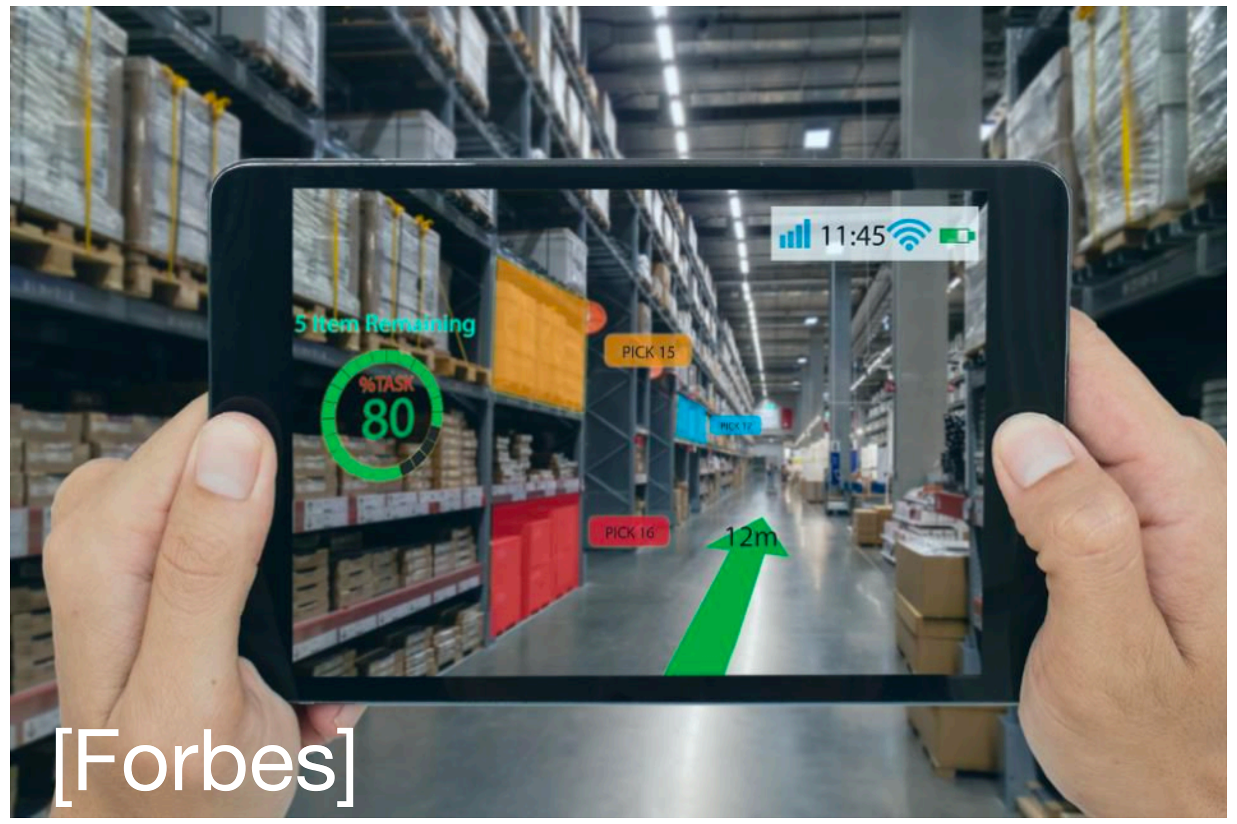




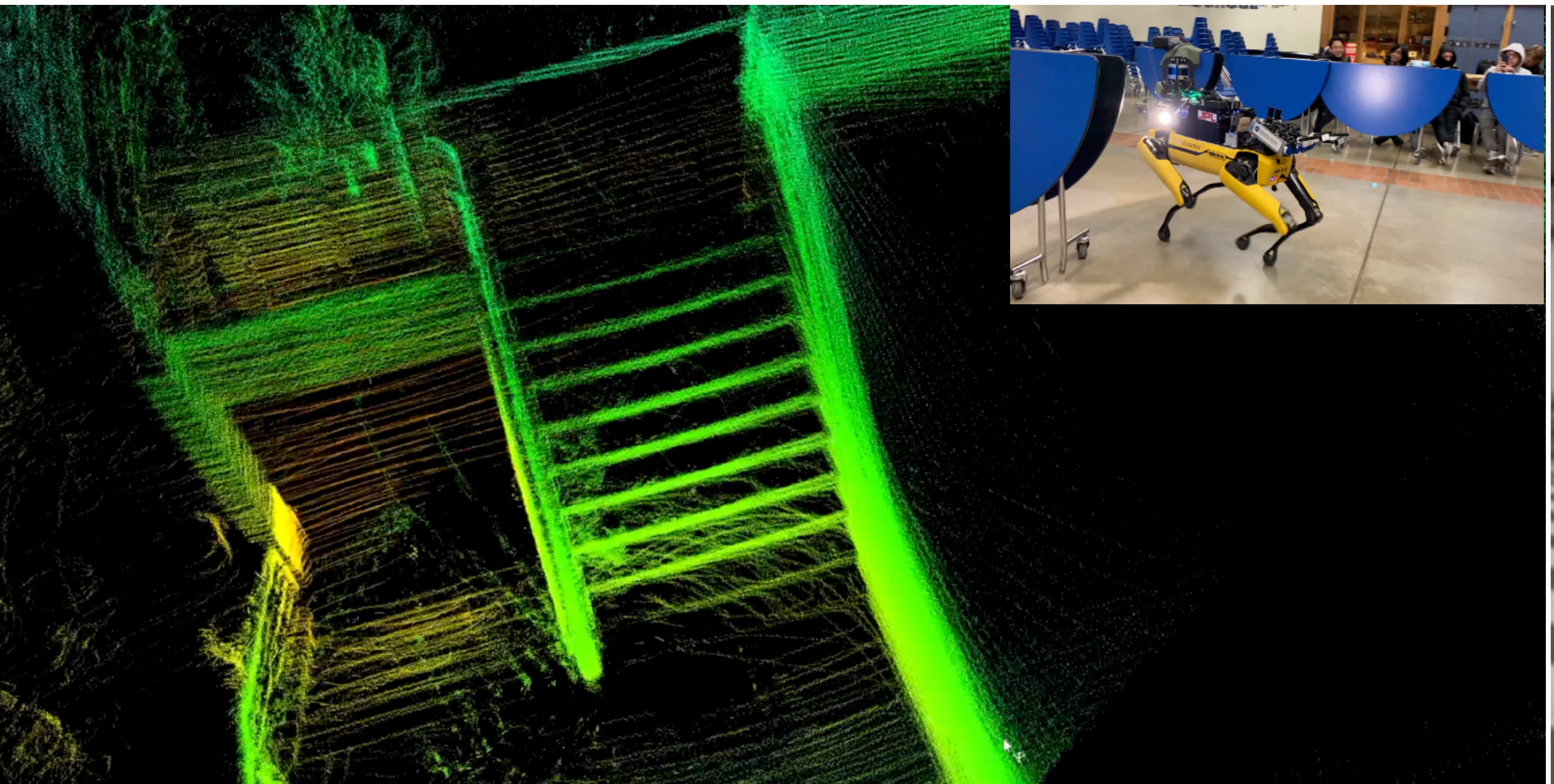
Soft Drones and Soft Aerial Manipulation

Spatial perception

- Localization and mapping
- Object detection / pose estimation
- Object tracking
- Semantic understanding
- ...

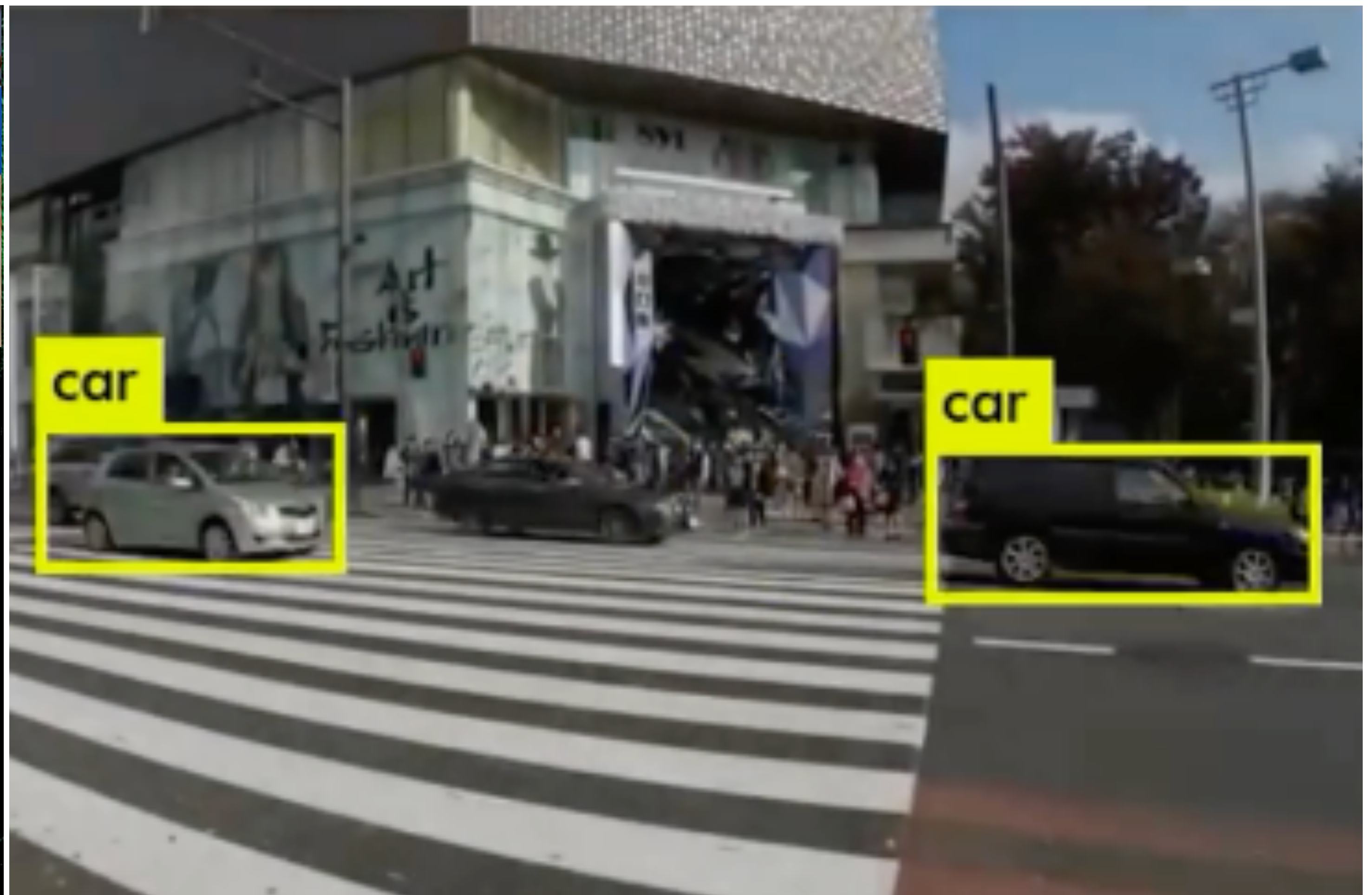


Opening the doors of perception



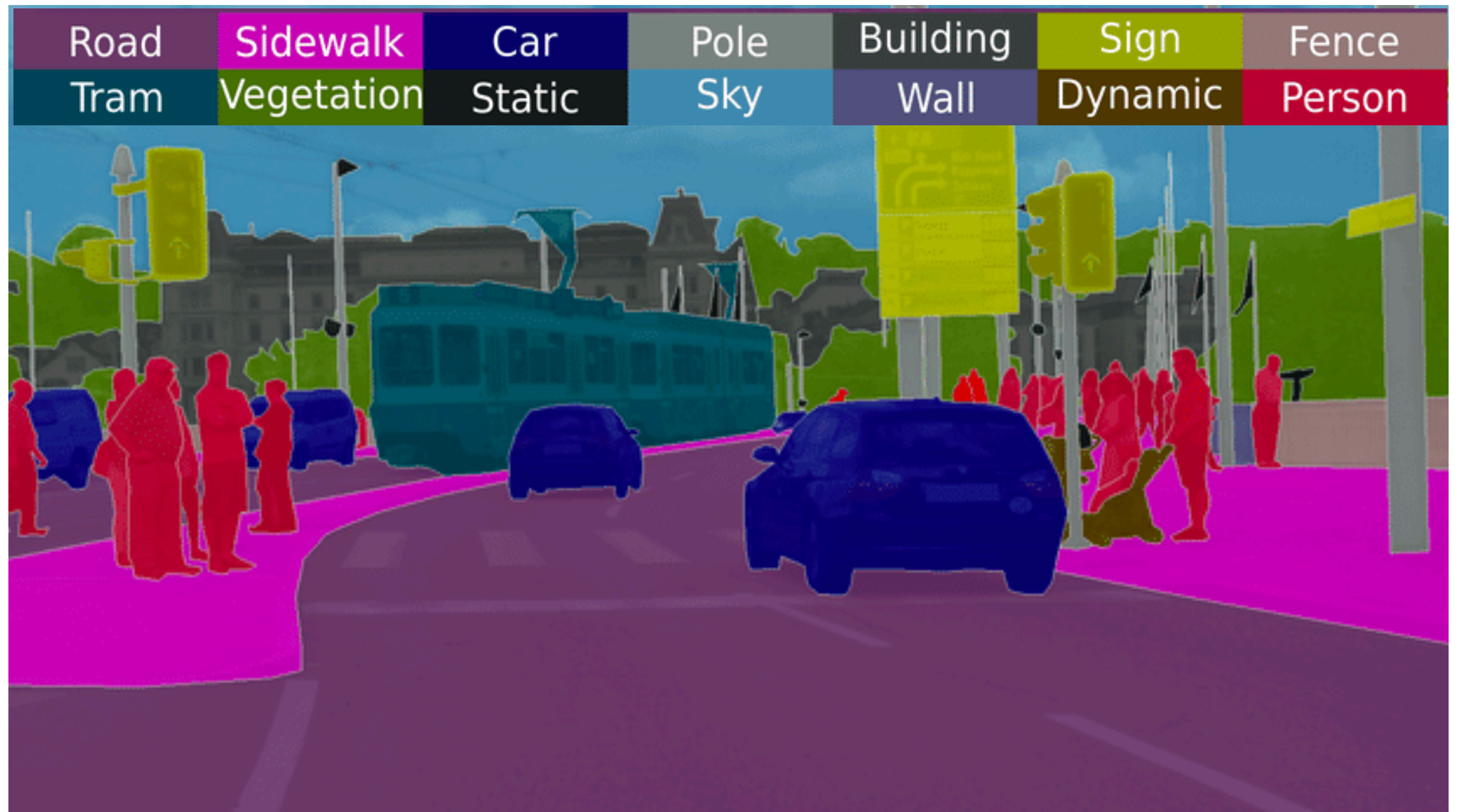
Localization and Mapping

[DARPA Subterranean Challenge]



Object Detection

[Yolo]



2D Semantic Segmentation

[Cityscape]

SLAMCORE

Aptiv

Skydio

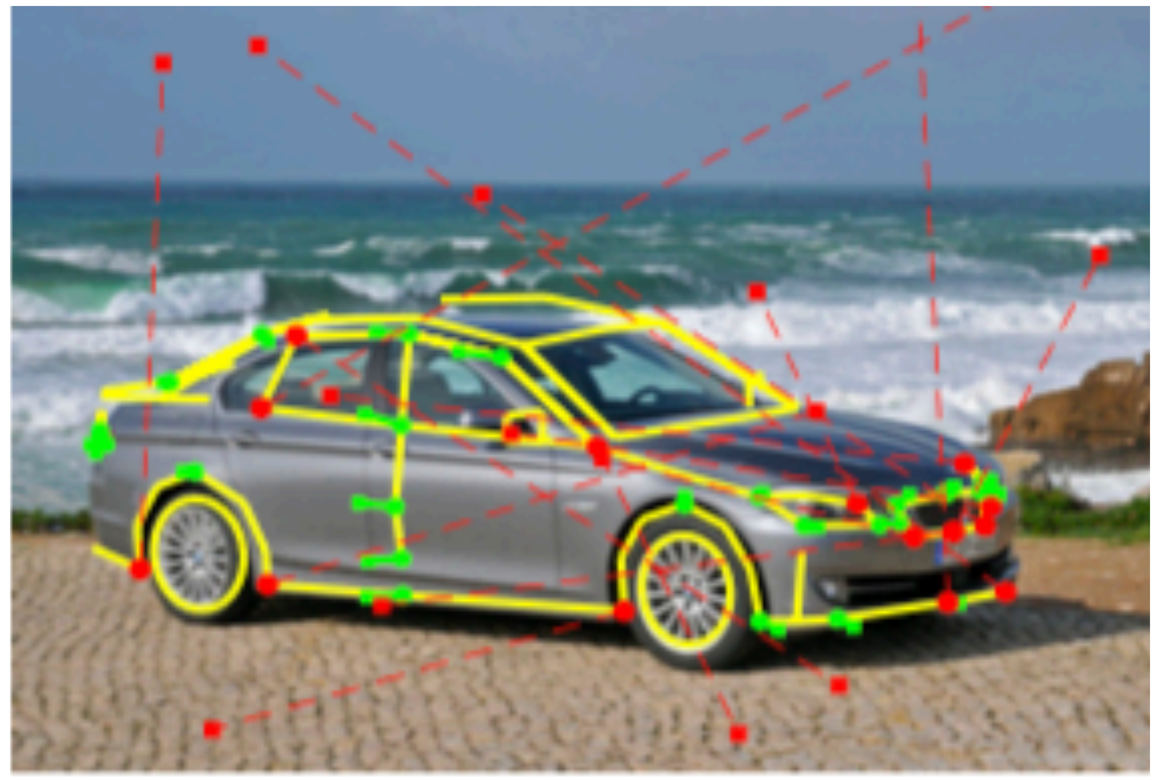
Oculus VR

With great power comes great responsibility

Images: Evtimov et al

Camouflage graffiti and art stickers cause a neural network to misclassify stop signs as speed limit 45 signs or yield signs.

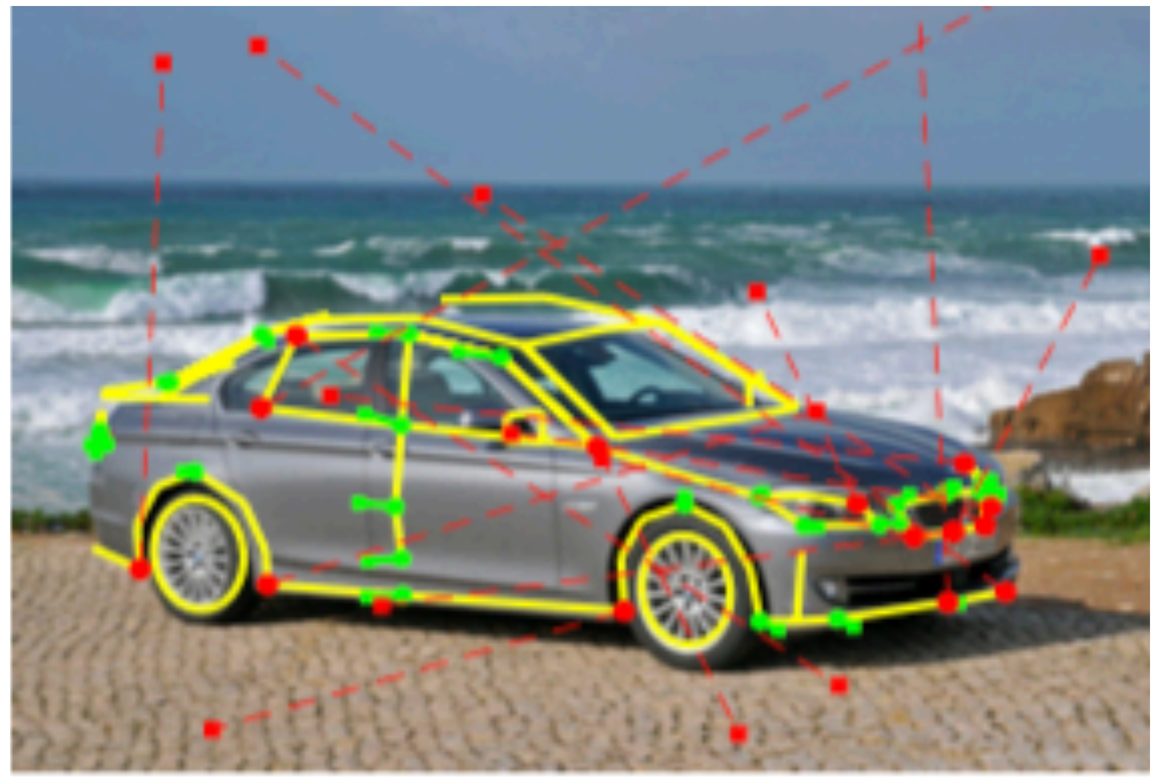
Key takeaways of this talk



In order to get low failure rates and performance guarantees we need to rethink current perception algorithms

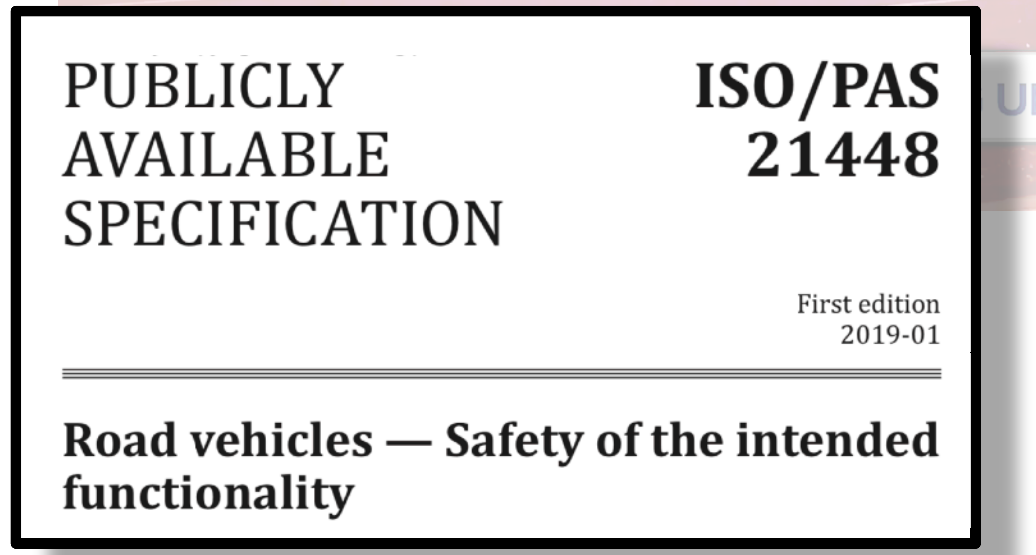
We need a theory of robust spatial perception: how to connect robust algorithms into a robust system?

Outline



Certifiable Perception Algorithms

Towards System-level Guarantees & Real-time High-level Understanding

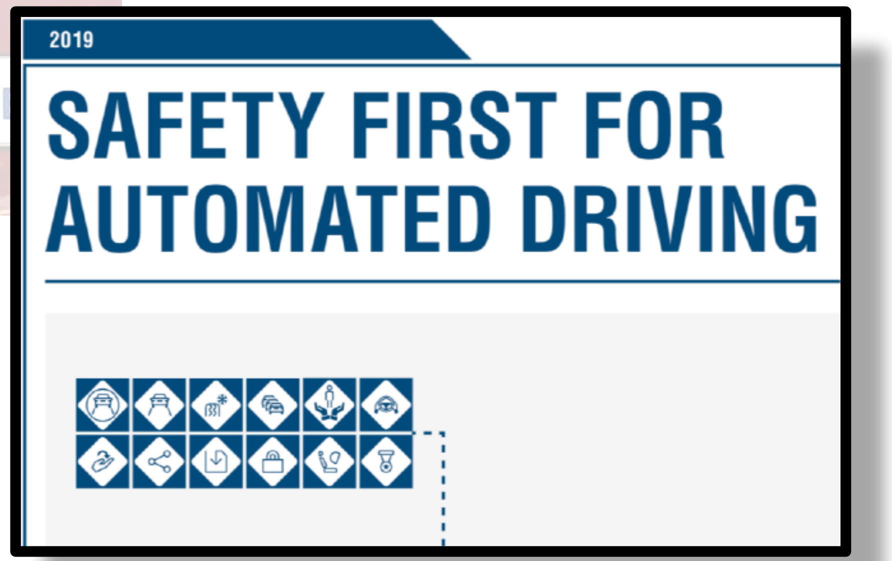


PUBLICLY AVAILABLE SPECIFICATION

ISO/PAS 21448

First edition
2019-01

Road vehicles — Safety of the intended functionality



2019

SAFETY FIRST FOR AUTOMATED DRIVING

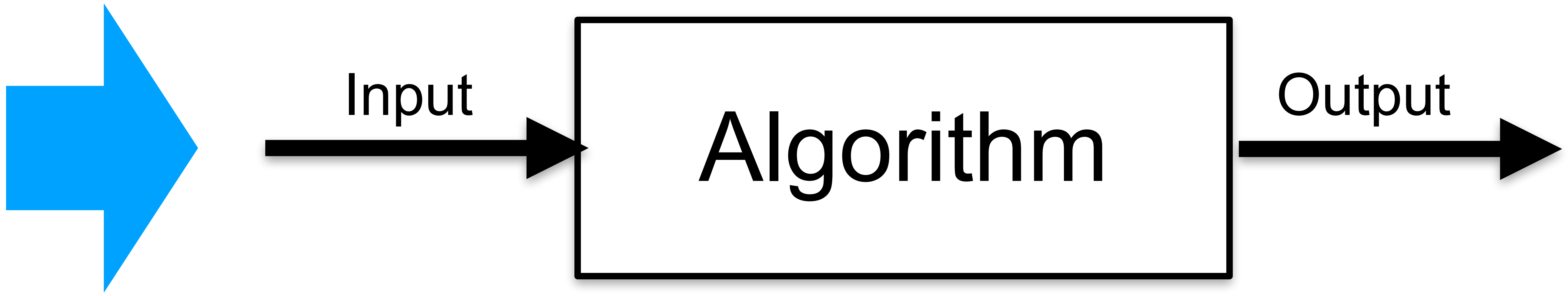


Image-based object localization

RGB images: [Gu&Kanade, CVPR’06][Lin, ECCV’14][Zhou, CVPR’15][Pavlakos, ICRA’17][Yang, CVPR’20]
Point clouds: [PointNetLK, CVPR’19][DCP, ICCV’19][SmoothNet, CVPR’19][TEASER, RSS’19, TRO’20]

(Generality)

Perception Front-end
Perception Back-end
Object localization in images
Feature detection
Model fitting/ Estimation
Object localization in point clouds
Feature detection
3D registration
SLAM (visual-inertial navigation, Structure from Motion)
Camera images
Inertial Measurements
Sensor Front-end
Factor Graph Optimization

Image-based object localization: perception issues

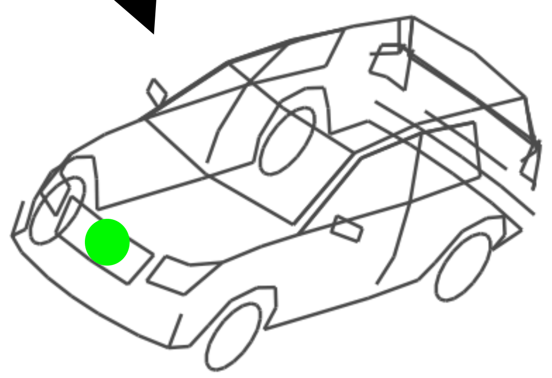
***ISSUE 1**: front-end (hand-crafted or deep-learned) can produce many mis-detections (not uncommon to have >90% outliers)*

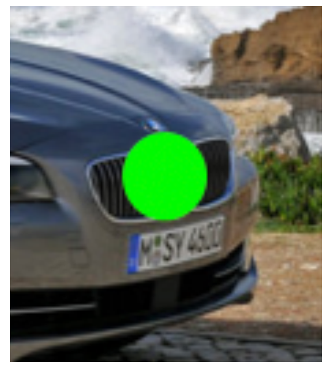
***ISSUE 2**: back-end may fail if there are many outliers*

RGB images: [Gu&Kanade, CVPR'06][Lin, ECCV'14][Zhou, CVPR'15][Pavlakos, ICRA'17][Xinke, RSS'19][Yang, CVPR'20]
Point clouds: [PointNetLK, CVPR'19][DCP, ICCV'19][SmoothNet, CVPR'19][TEASER, RSS'19, TRO'20]

Why does the back-end fail?

$$\min_{\boldsymbol{x} \in \mathbb{X}} \sum_{i \in \mathcal{M}} \| r(\boldsymbol{x}, \boldsymbol{y}_i) \|^2$$





Set of measurements

Residual

Estimate

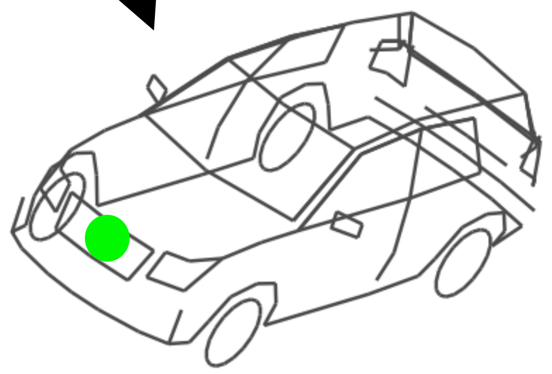
Measurement “i”

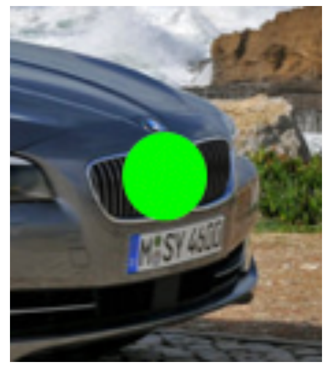
[Dellaert and Kaess, Factor Graphs for Robot Perception, FnT 2017]
[Barfoot. State Estimation for Robotics. Cambridge University Press 2017]

Why does the back-end fail?

Loss function

$$\min_{\boldsymbol{x}\in\mathbb{X}} \sum_{i\in\mathcal{M}} \rho(r(\boldsymbol{x},\boldsymbol{y}_i))$$





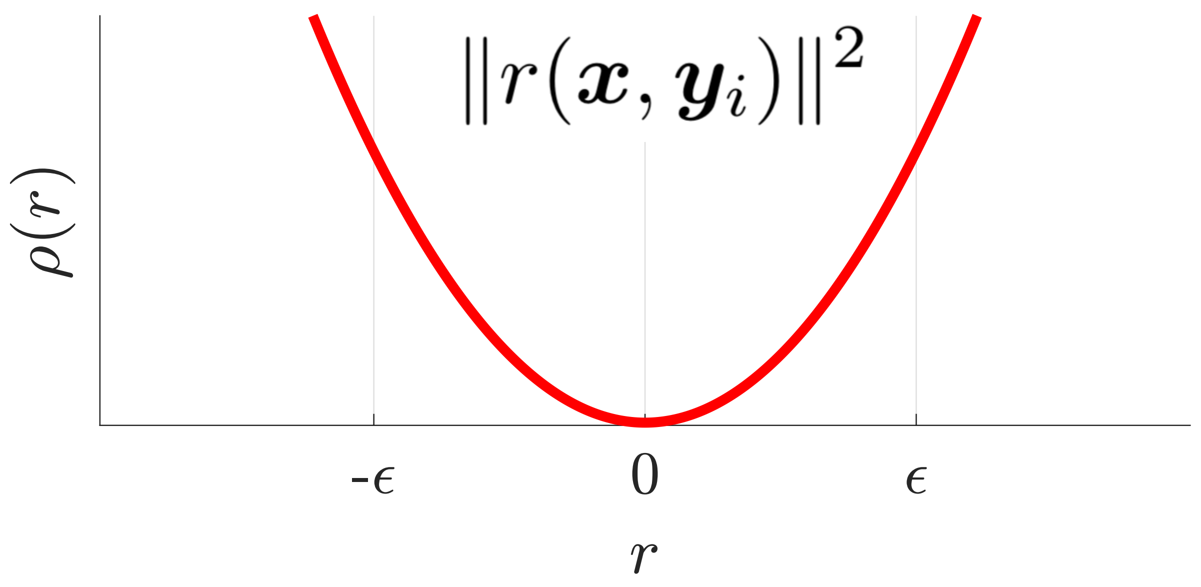
Set of measurements — Residual — Estimate — Measurement “i”

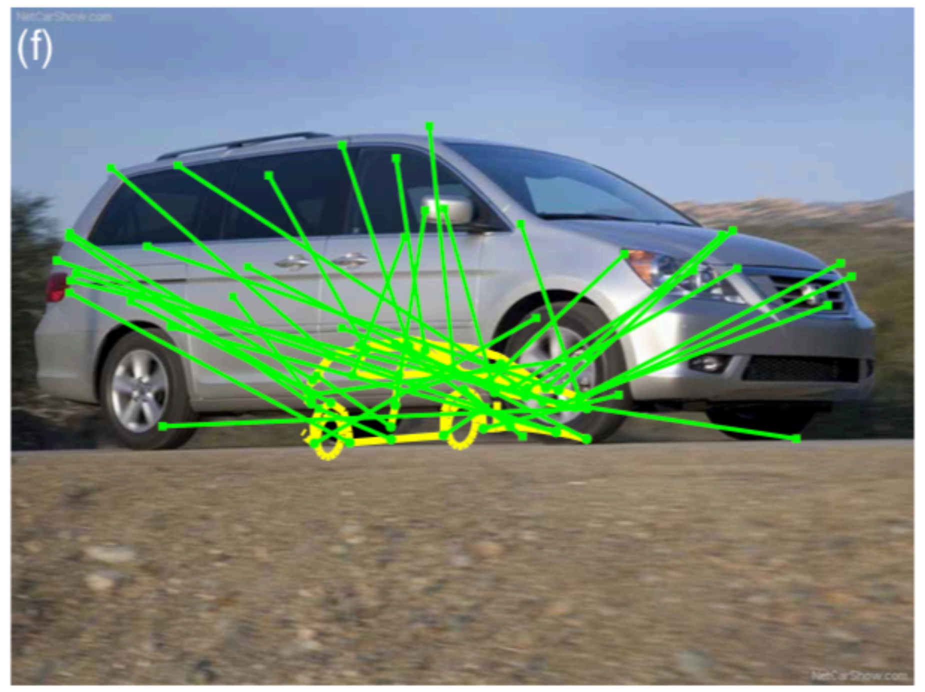
[Dellaert and Kaess, Factor Graphs for Robot Perception, FnT 2017]
[Barfoot. State Estimation for Robotics. Cambridge University Press 2017]

Why does the back-end fail?

$$\min_{\boldsymbol{x}\in\mathbb{X}} \sum_{i\in\mathcal{M}} \rho(r(\boldsymbol{x},\boldsymbol{y}_i))$$

Nonlinear Least Squares

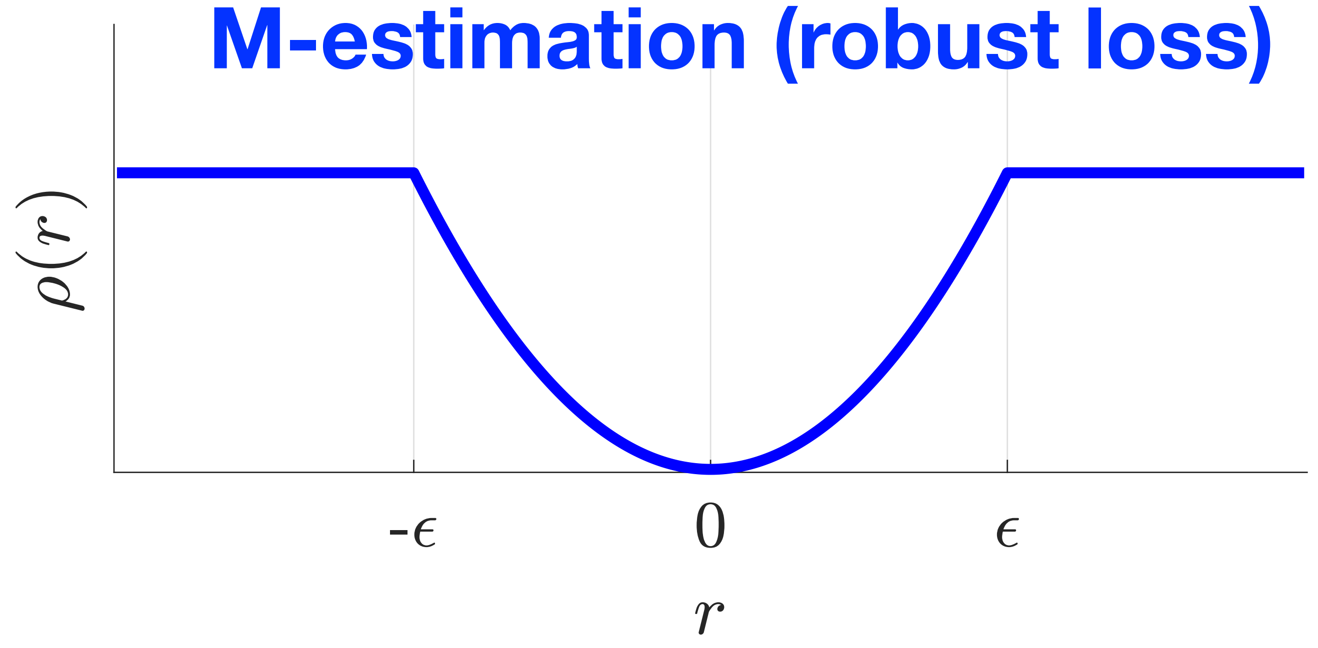


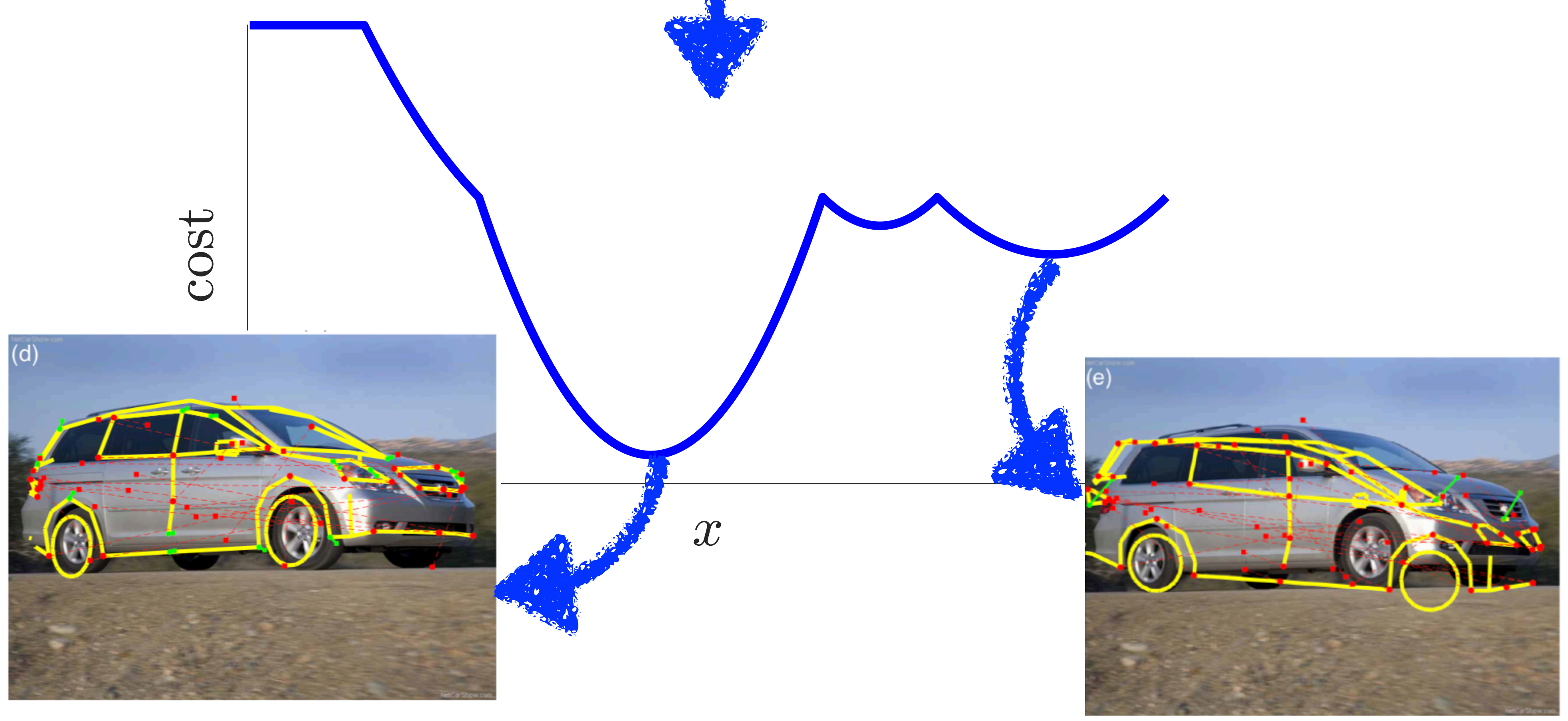


not robust to outliers

[Agarwal et al., ICRA'13] [Tavish and Barfoot, CRV'15]
[Chin and Suter, LCV'17] [Izatt et al., ISRR'17]
[Rosen et al., IJRR'18] [Doherty et al, ICRA'19]
[Chebrolu et al., Arxiv'20]

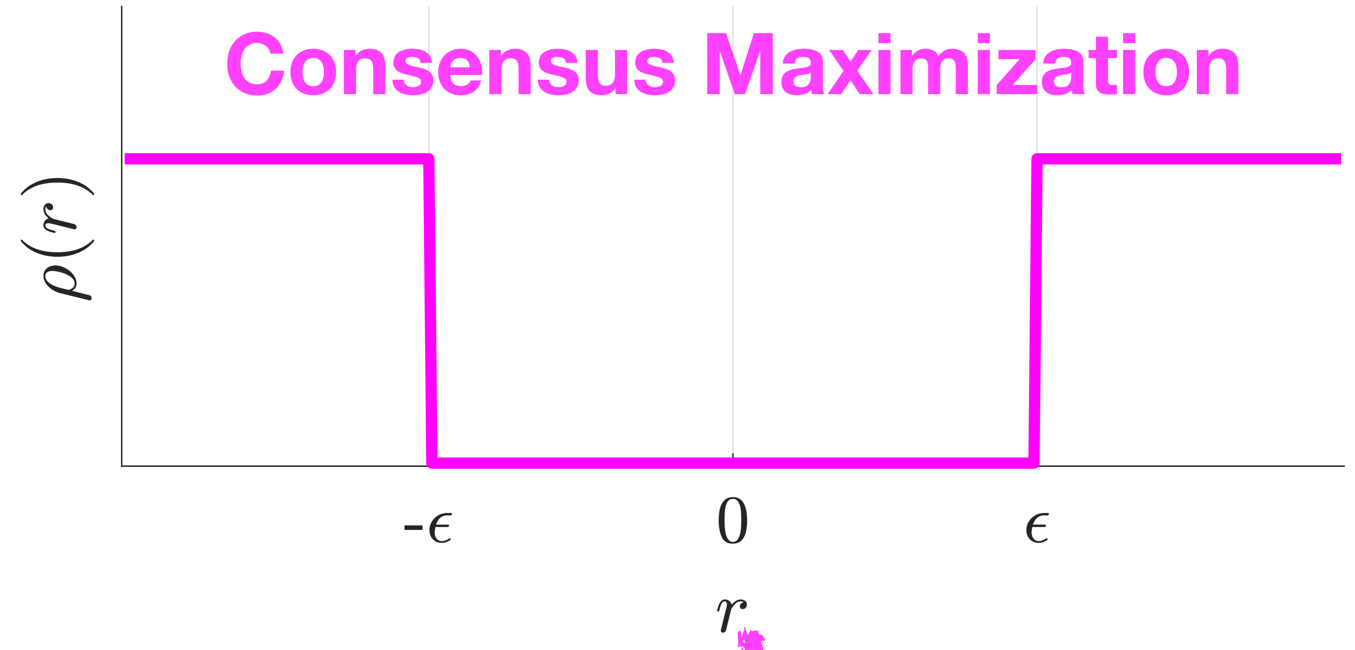
M-estimation (robust loss)

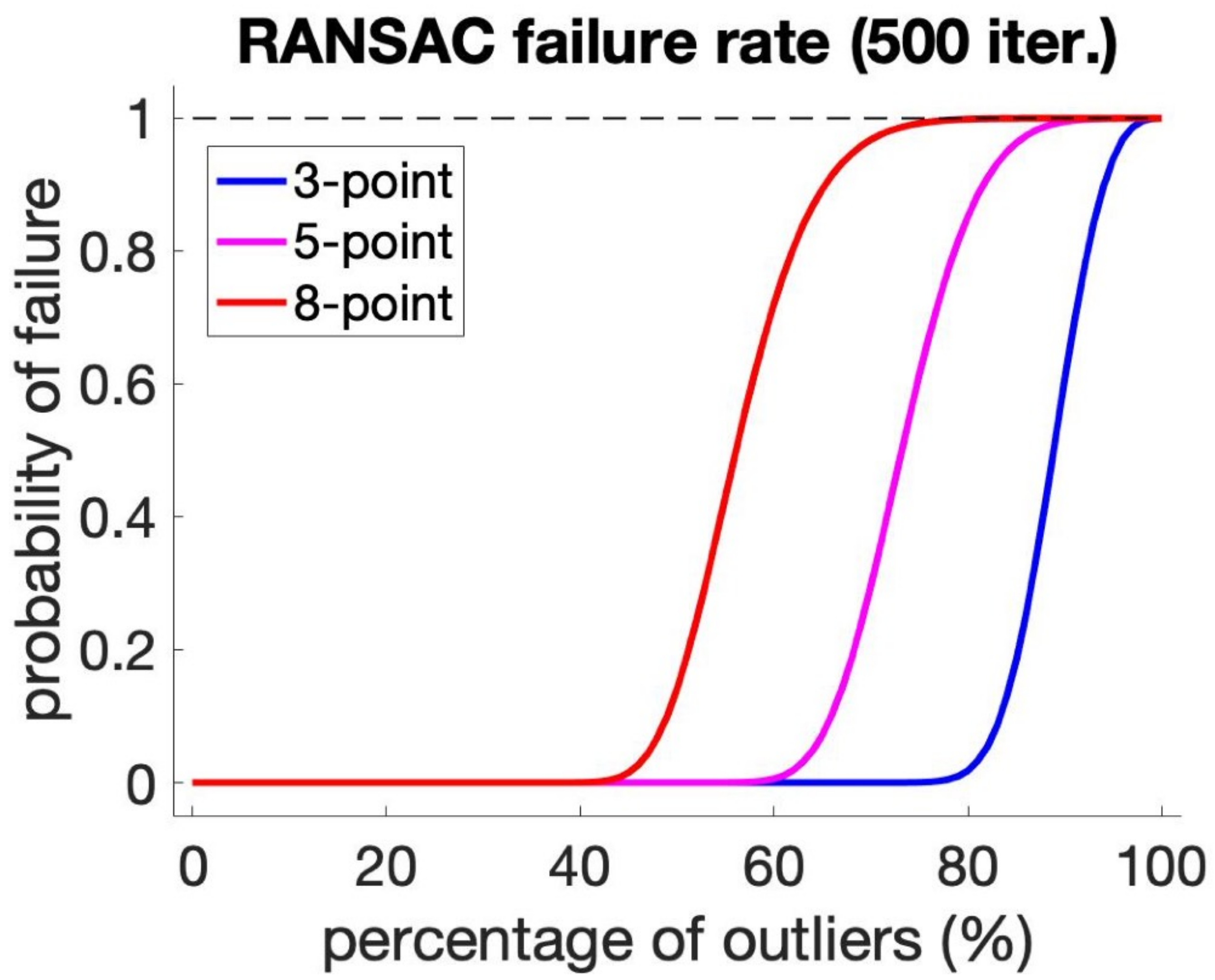




Local solvers:
- need initial guess
- brittle due to local convergence
- gradients?
- fail without notice

Consensus Maximization





RANSAC:
- fails with many outliers
- does not scale to "large" problems
- non-deterministic
- fails without notice

A new perspective: Certifiable Algorithms

Input
(measurements y_i)

Algorithm

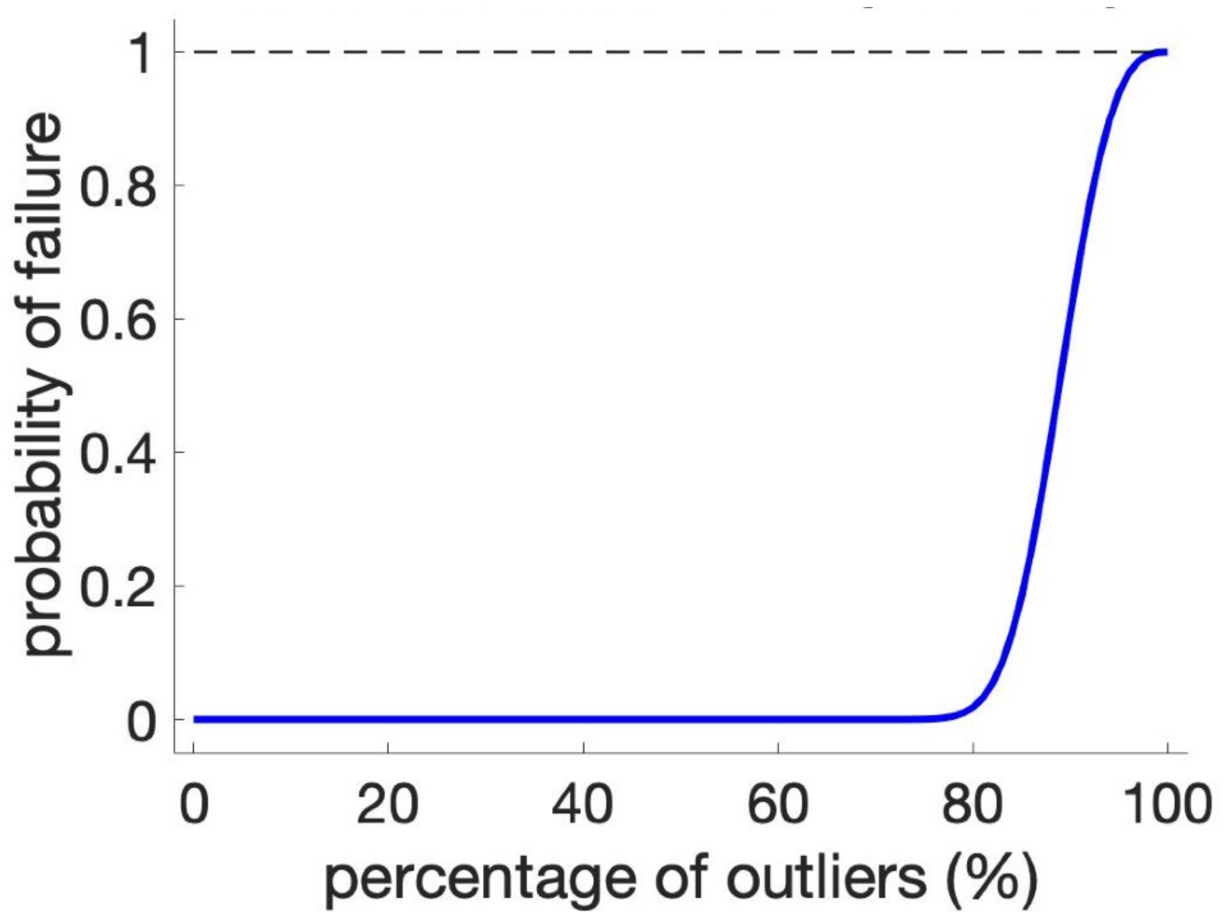
$$\min_{\boldsymbol{x} \in \mathbb{X}} \sum_{i \in \mathcal{M}} \rho(r(\boldsymbol{x}, \boldsymbol{y}_i))$$

Output
(estimate)

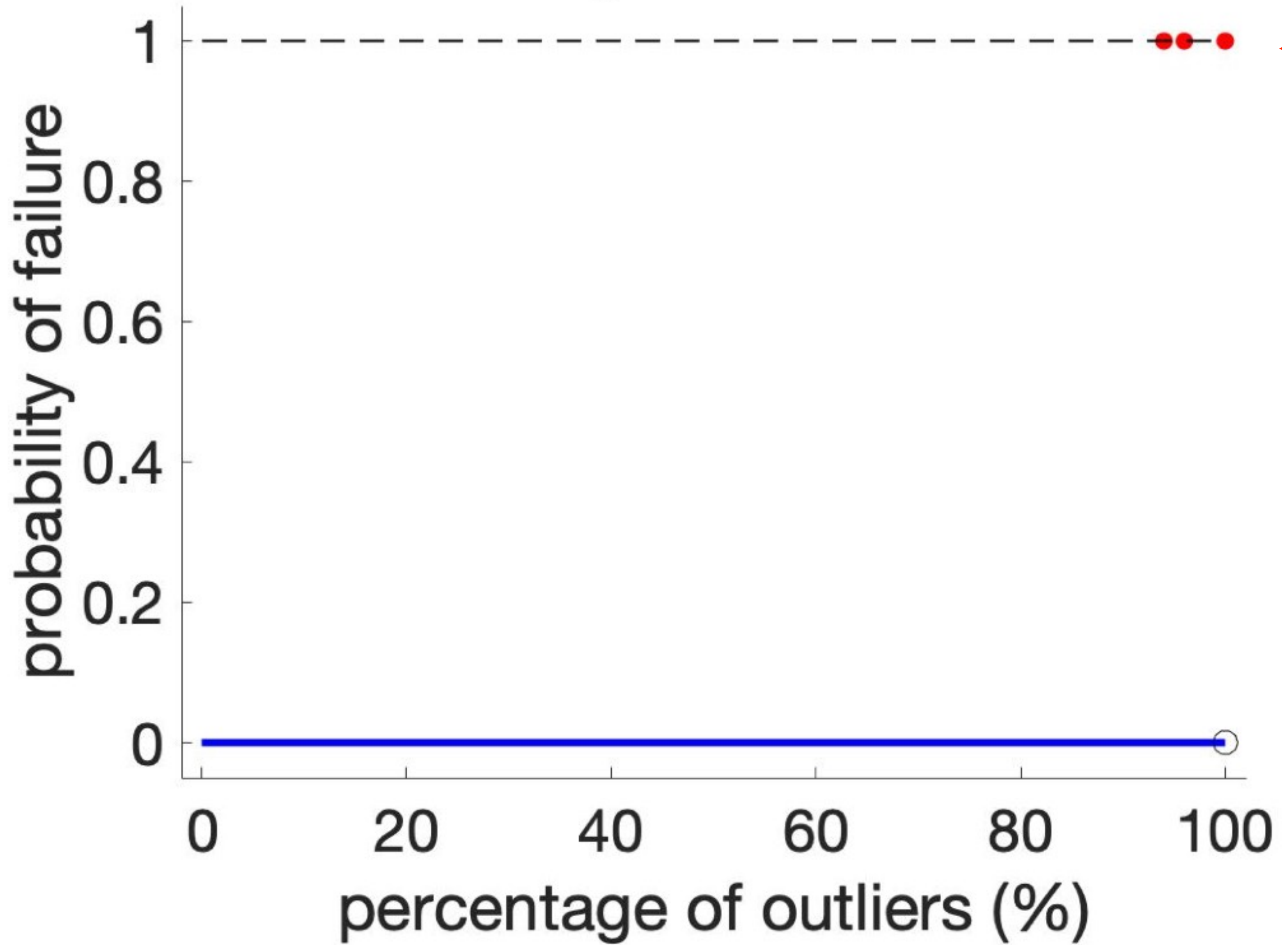
Certifiably robust

Certifiable algorithms: fast (i.e., polynomial-time) algorithms that solve outlier rejection to optimality in virtually all problem instances or detect failure in worst-case problems

RANSAC



Certifiable algorithms



Good news: worst-case instances are not commonly found in practice

Flatten the curve!

One ring to rule them all

Black-Rangarajan duality (IJCV 1996)

Lasserre hierarchy of relaxations (2001)

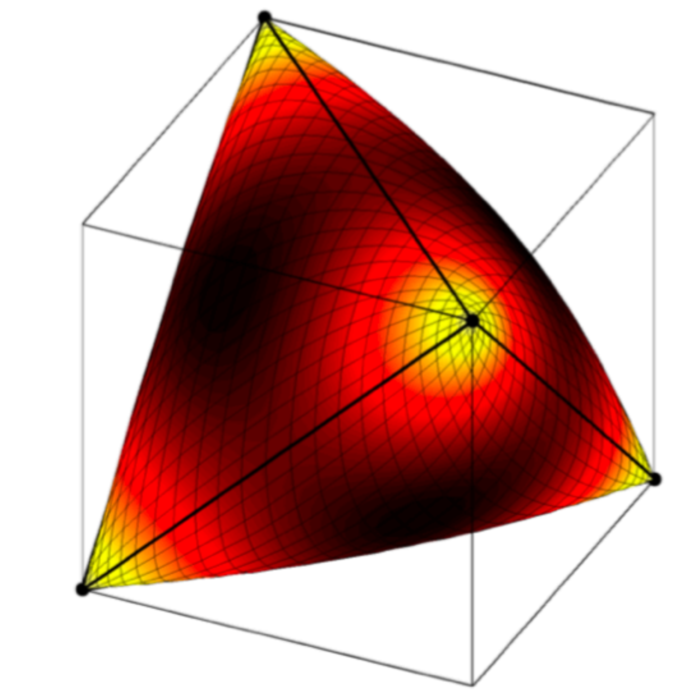
Robust Estimation (non-convex, hard)

$$\min_{\boldsymbol{x} \in \mathbb{X}} \sum_{i \in \mathcal{M}} \rho(r(\boldsymbol{x}, \boldsymbol{y}_i))$$

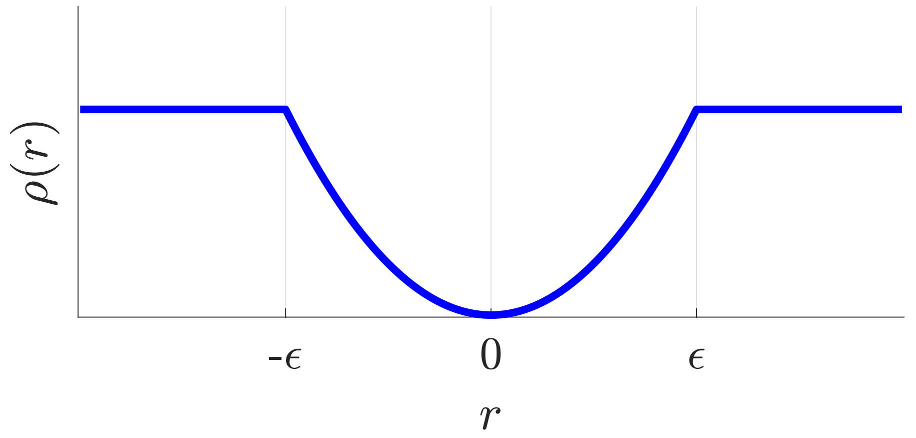
Optimization over the ring of polynomials (non-convex, hard)

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & f(\boldsymbol{x}) \\ s.t. \quad & h_i(\boldsymbol{x}) = 0, i = 1, \ldots, m, \\ & g_k(\boldsymbol{x}) \geq 0, k = 1, \ldots, l, \end{aligned}$$

Semidefinite relaxation (convex)



Optimality certificate



+ clever choice of robust loss function (truncated least squares)

+ leverage problem geometry

+ basis reduction

Theory enables:

Robust solvers:
- ✓ no initial guess
- ✓ globally optimal
- ✓ deterministic
- ✓ tolerate extreme outliers
- ✓ detect failures

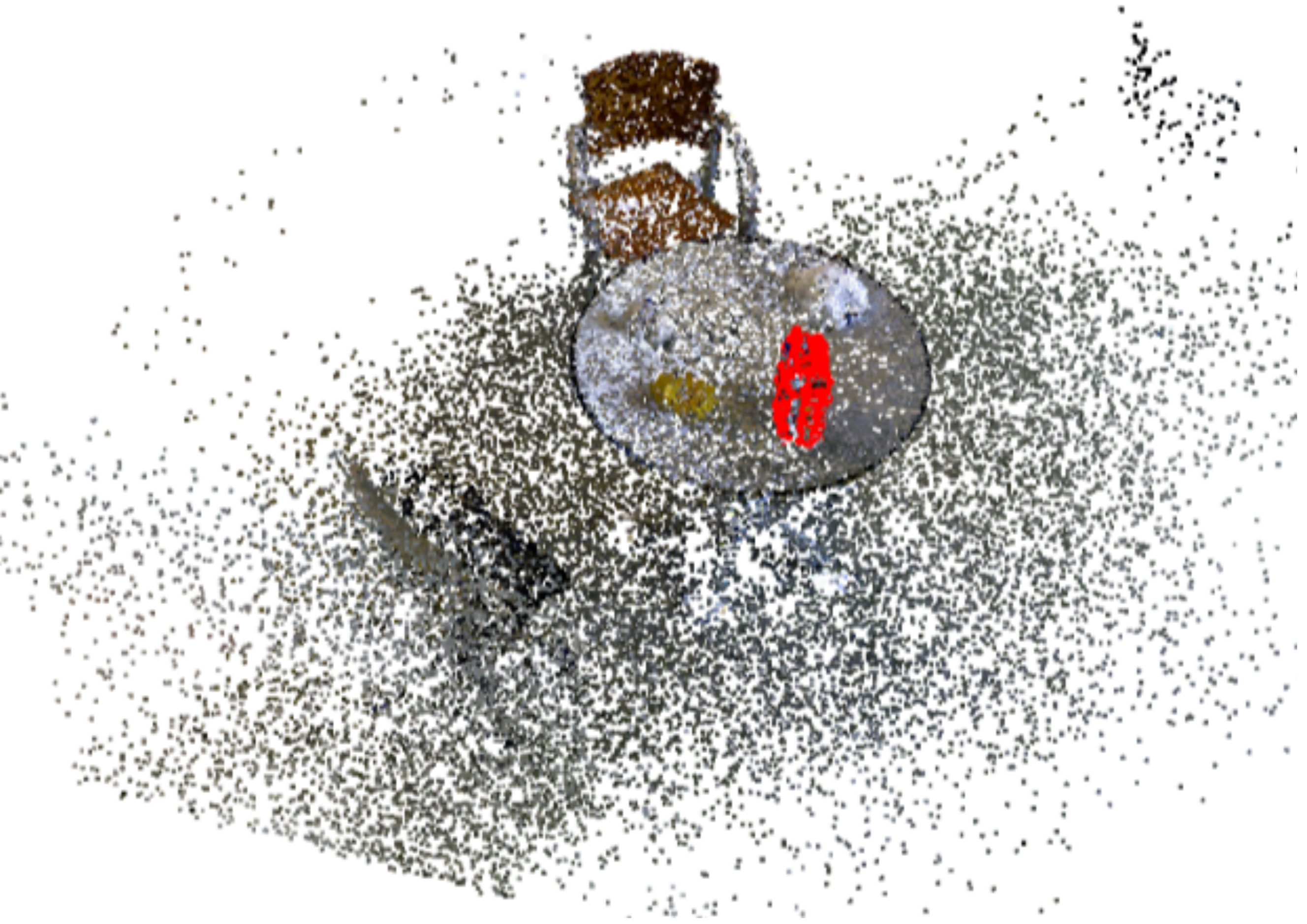
Fast certification techniques:
- ✓ check quality of given estimate (from RANSAC, local solvers, GNC)

Estimation contracts: establish when global optimum recovers correct pose

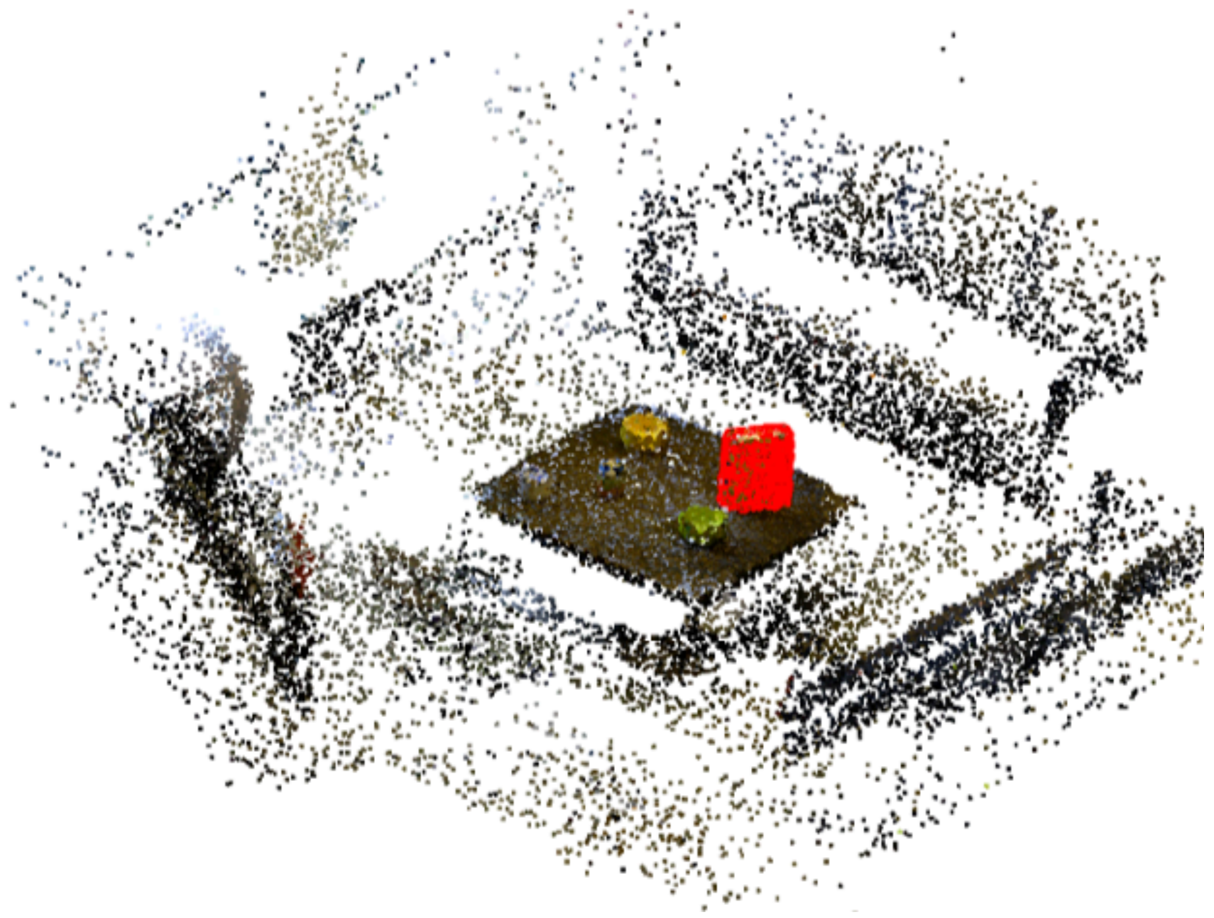
TEASER++: Certifiable object localization in point clouds

- RGB-D point cloud dataset
- ~500 FPFH correspondences

96.87% outliers

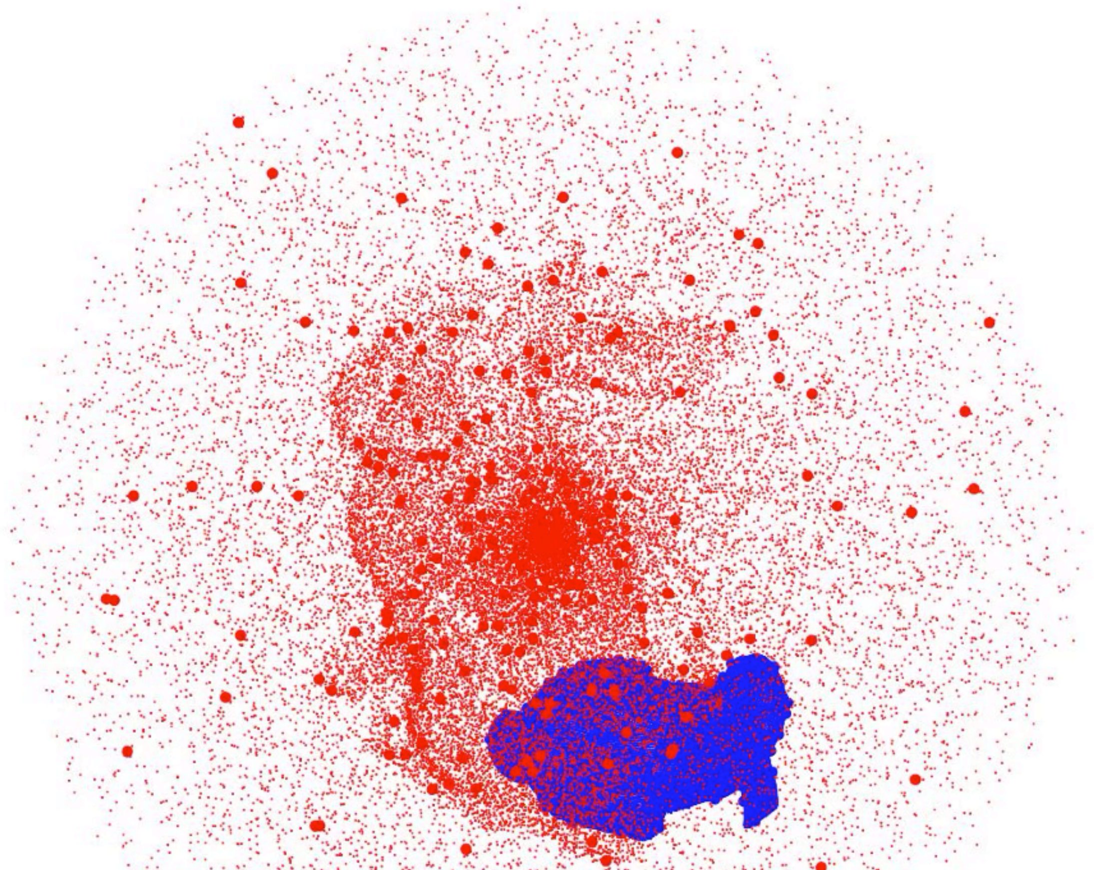


97.37% outliers

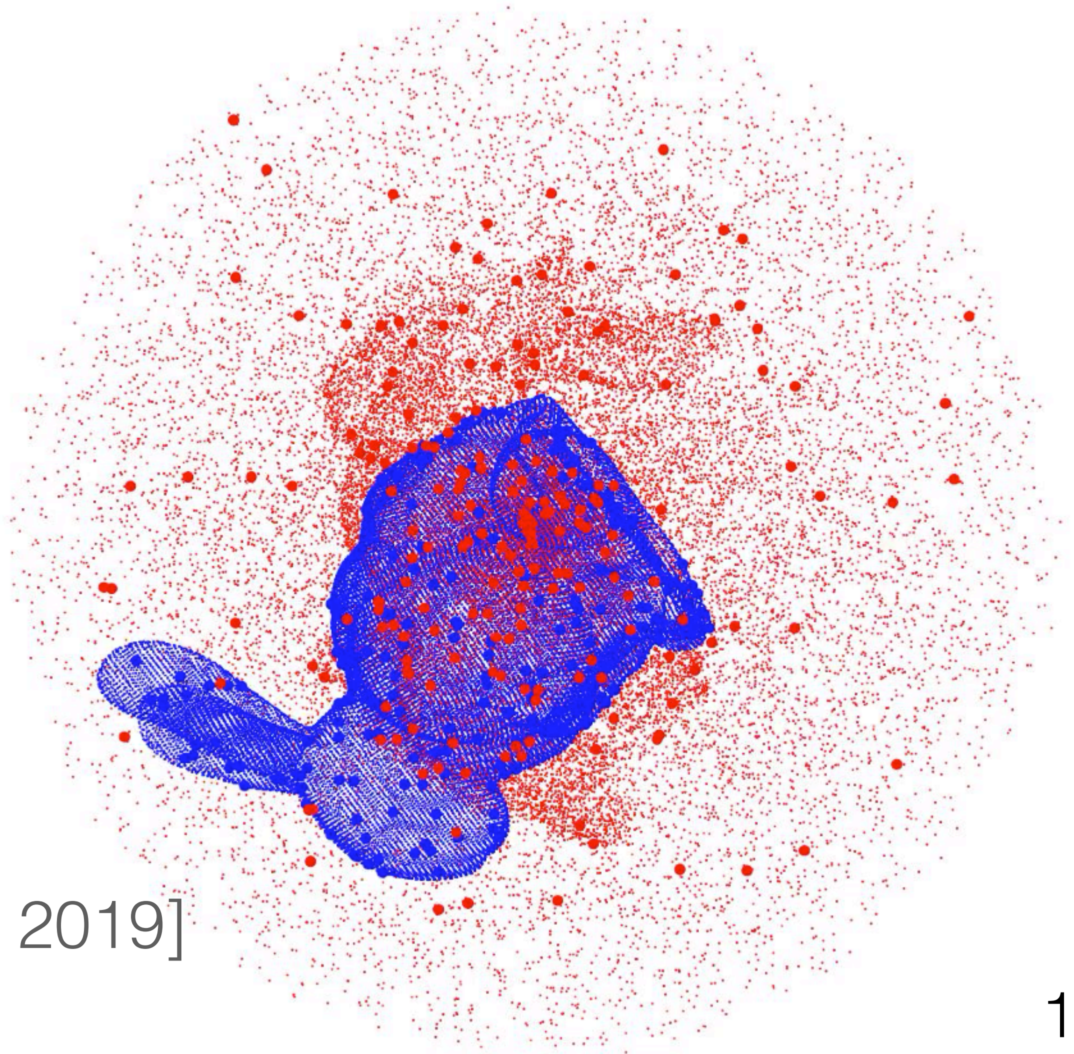


95.44% outliers

Teaser++ (proposed)



RANSAC



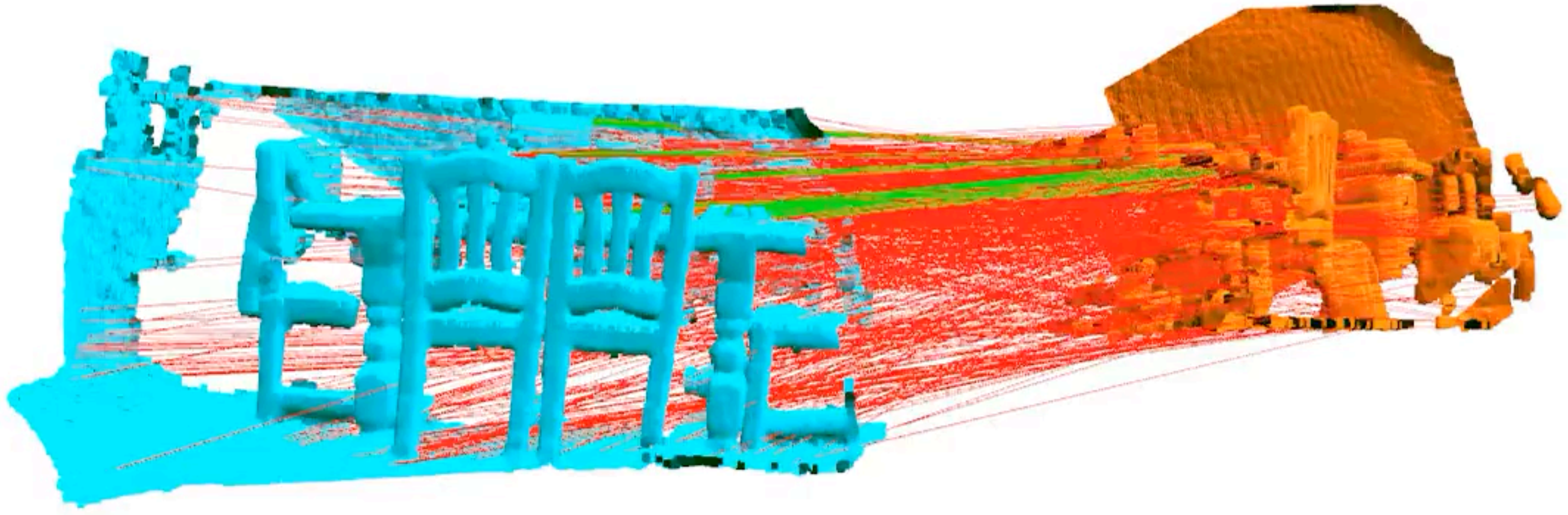
[Yang and Carlone. A Polynomial-time Solution for Robust Registration with Extreme Outlier Rates. RSS 2019]
[Yang, Shi, Carlone. TEASER: Fast and Certifiable Point Cloud Registration. TRO 2020]

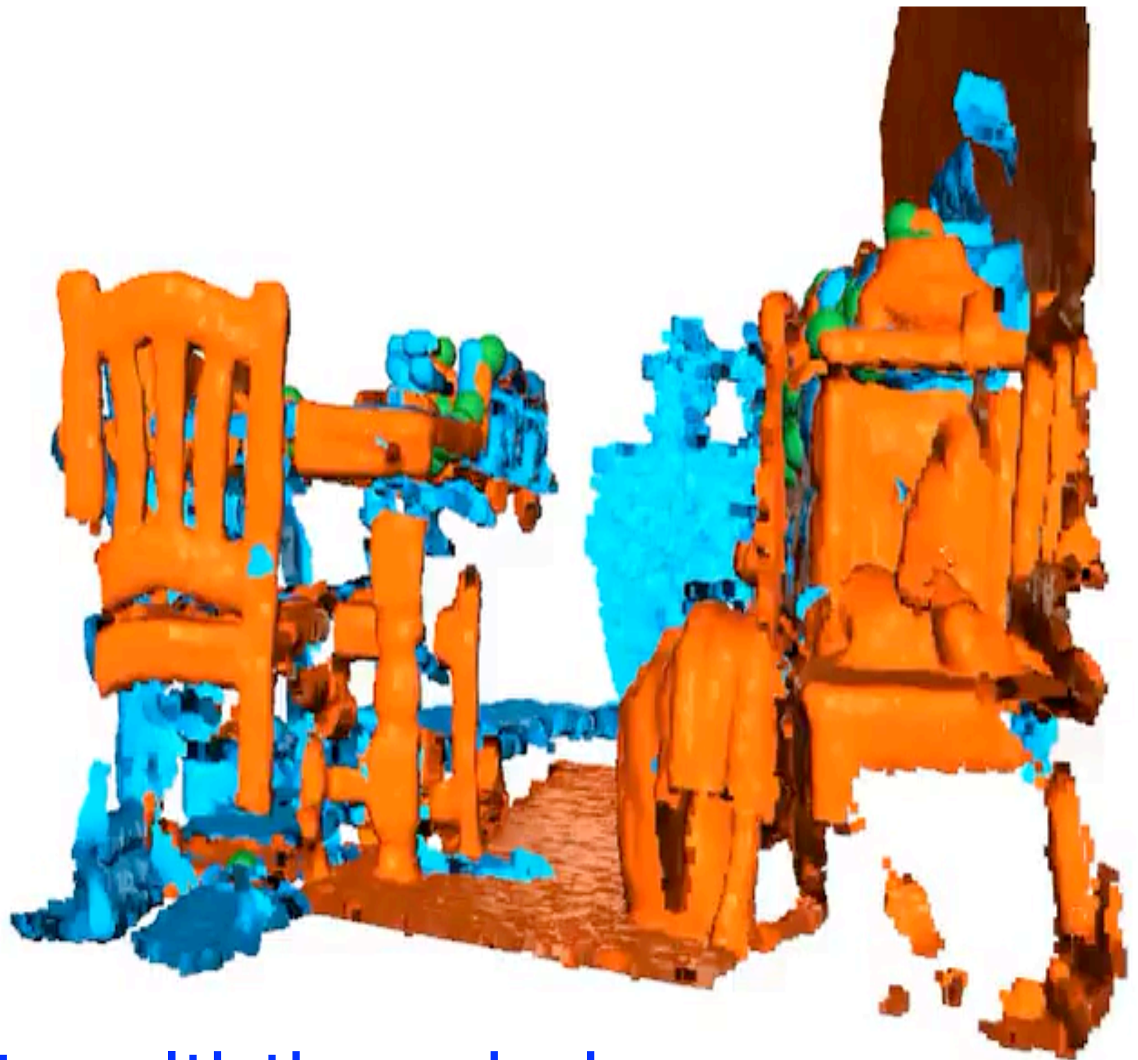
TEASER++: Certifiable object localization in point clouds

- 3DMatch dataset, ~1000 deep-learned correspondences (using 3DSmoothNet)

Success rate

Scenes	Kitchen (%)	Home 1 (%)	Home 2 (%)	Hotel 1 (%)	Hotel 2 (%)	Hotel 3 (%)	Study Room (%)	MIT Lab (%)
RANSAC-1K	90.9	91.0	73.1	88.1	80.8	87.0	79.1	81.8
RANSAC-10K	96.4	92.3	73.1	92.0	84.6	90.7	82.2	81.8
TEASER++	97.7	92.3	82.7	96.9	88.5	94.4	88.7	84.4
TEASER++ (CERT)	99.2	97.5	90.0	98.8	94.9	97.7	94.8	93.9





Green lines: **inlier correspondences**

Red lines: **outlier correspondences**

Fast multi-threaded open-source code: https://github.com/MIT-SPARK/TEASER-plusplus

[Yang and Carlone. A Polynomial-time Solution for Robust Registration with Extreme Outlier Rates. RSS 2019]
[Yang, Shi, Carlone. TEASER: Fast and Certifiable Point Cloud Registration. TRO 2020]

Estimation contract

x° **Input** (measurements y_i) →

TEASER++

$$\min_{\boldsymbol{x} \in \mathbb{X}} \sum_{i \in \mathcal{M}} \rho(r(\boldsymbol{x}, \boldsymbol{y}_i))$$

→ **Output** (estimate) $x^\star$

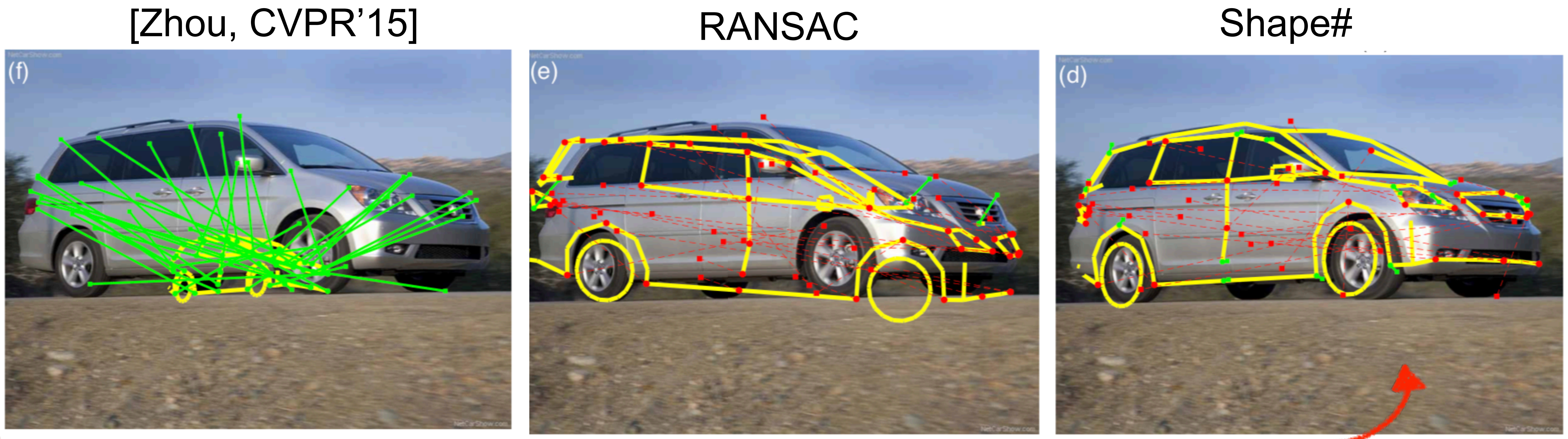
Theorem (Exact Recovery): if the measurements (i) contain at least 3 noiseless and non-collinear inliers, (ii) the outliers are not adversarial, and (iii) the certificate of optimality holds, TEASER++ recovers the true object pose.

Theorem (Exact Recovery in **Adversarial** Setting): if the number of noiseless and non-collinear inliers is larger than the number of outliers ($N_{in} > N_{out} + 3$) and the certificate of optimality holds, TEASER++ recovers the true object pose.

Shape#: Certifiable object localization in images

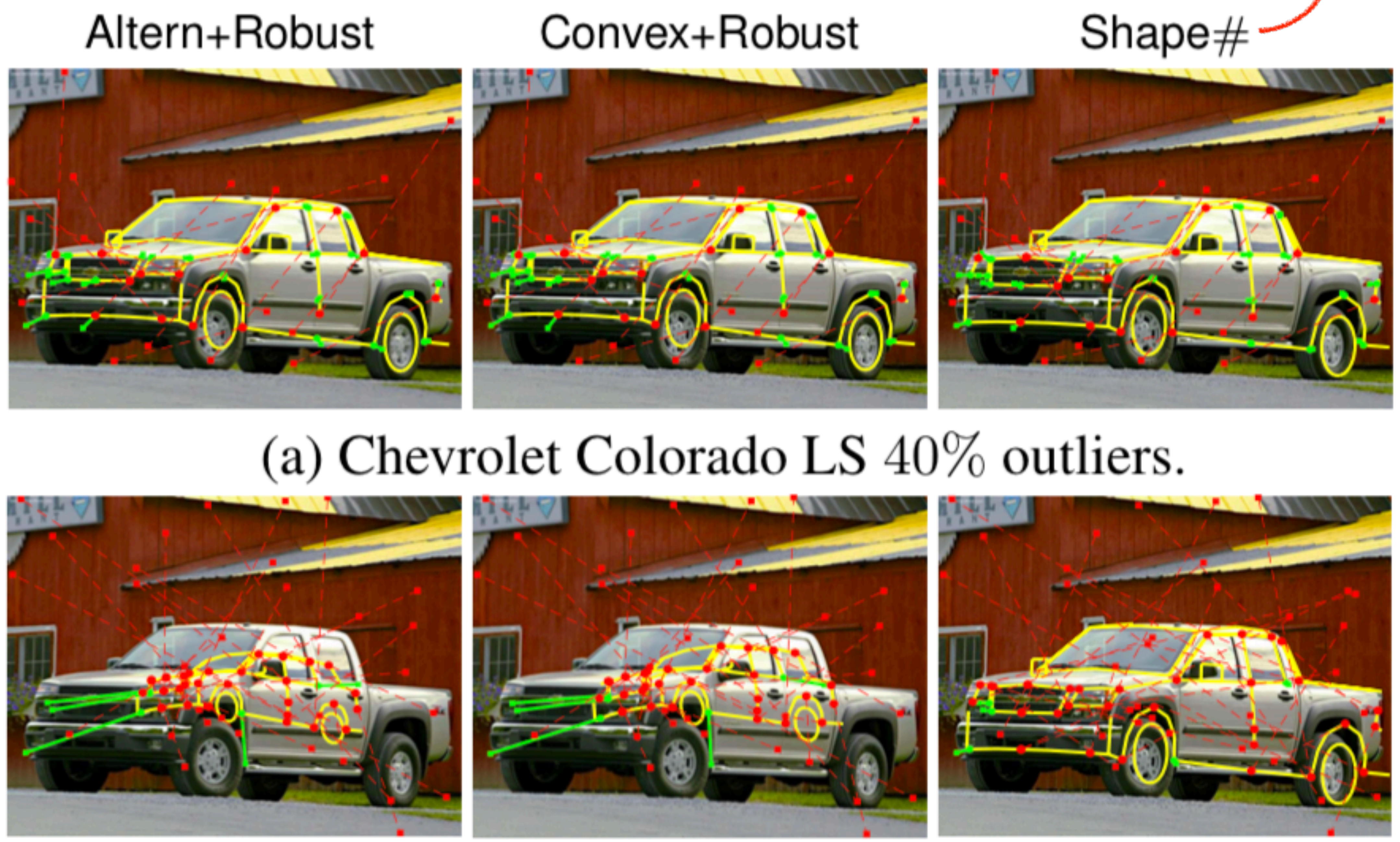
- FG3DCar dataset
- 300 car images with corresponding CAD models

70% outliers



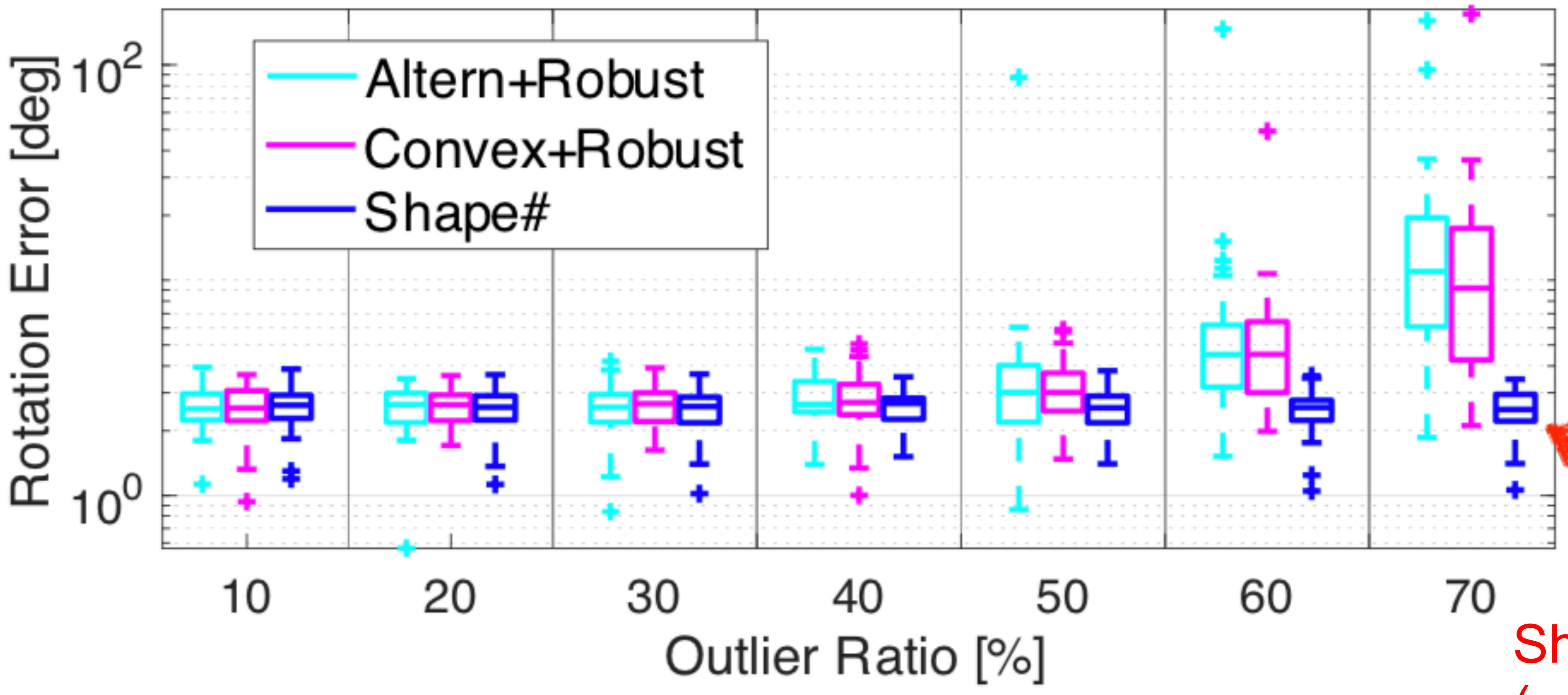
Proposed

Proposed



(a) Chevrolet Colorado LS 40% outliers.

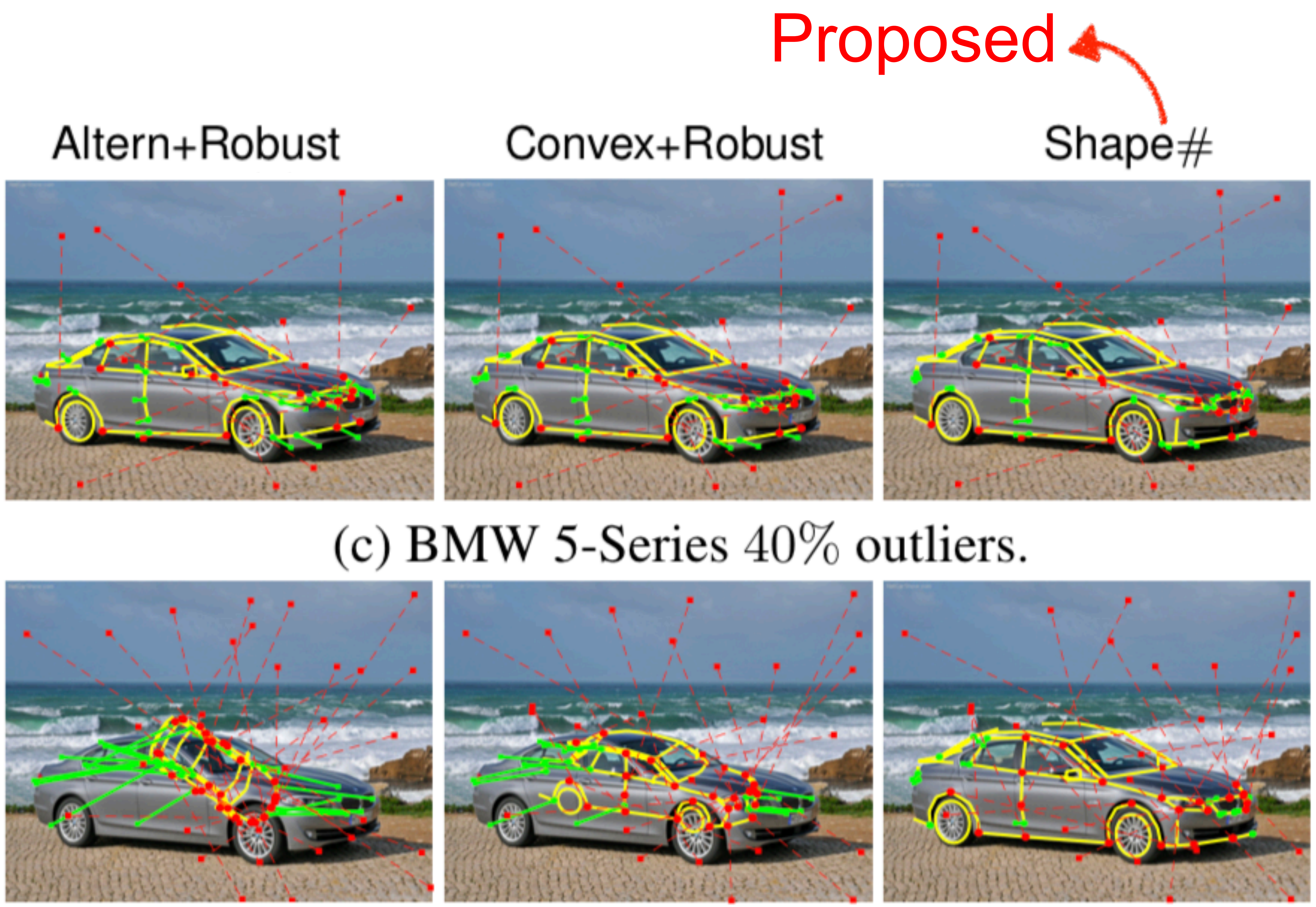
(b) Chevrolet Colorado LS 70% outliers.



Shape# (proposed) is best approach

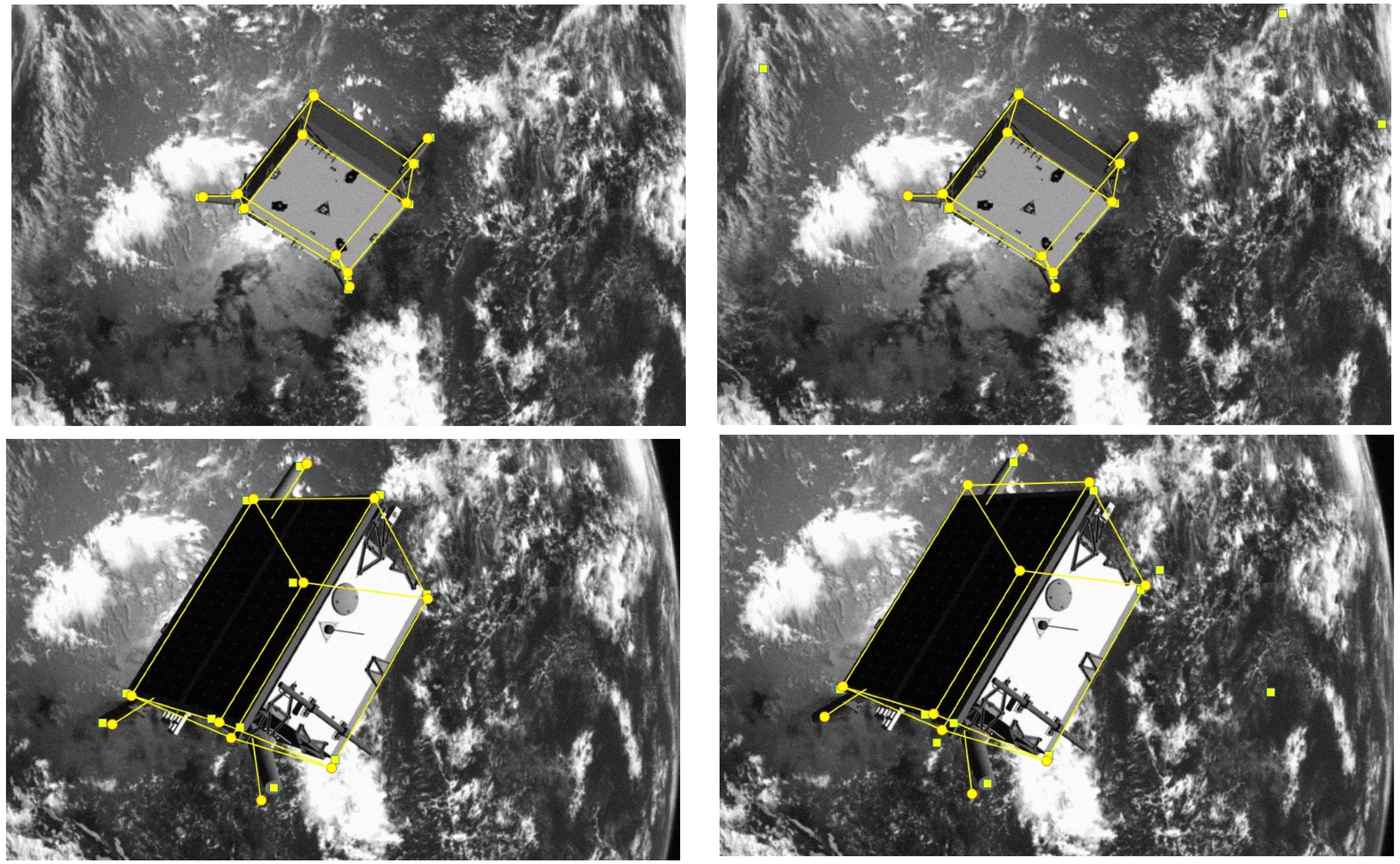
[Yang and Carlone. In Perfect Shape: Certifiably Optimal 3D Shape Reconstruction from 2D Landmarks. CVPR 2020]

Shape#: experimental results



(c) BMW 5-Series 40% outliers.

(d) BMW 5-Series 70% outliers.

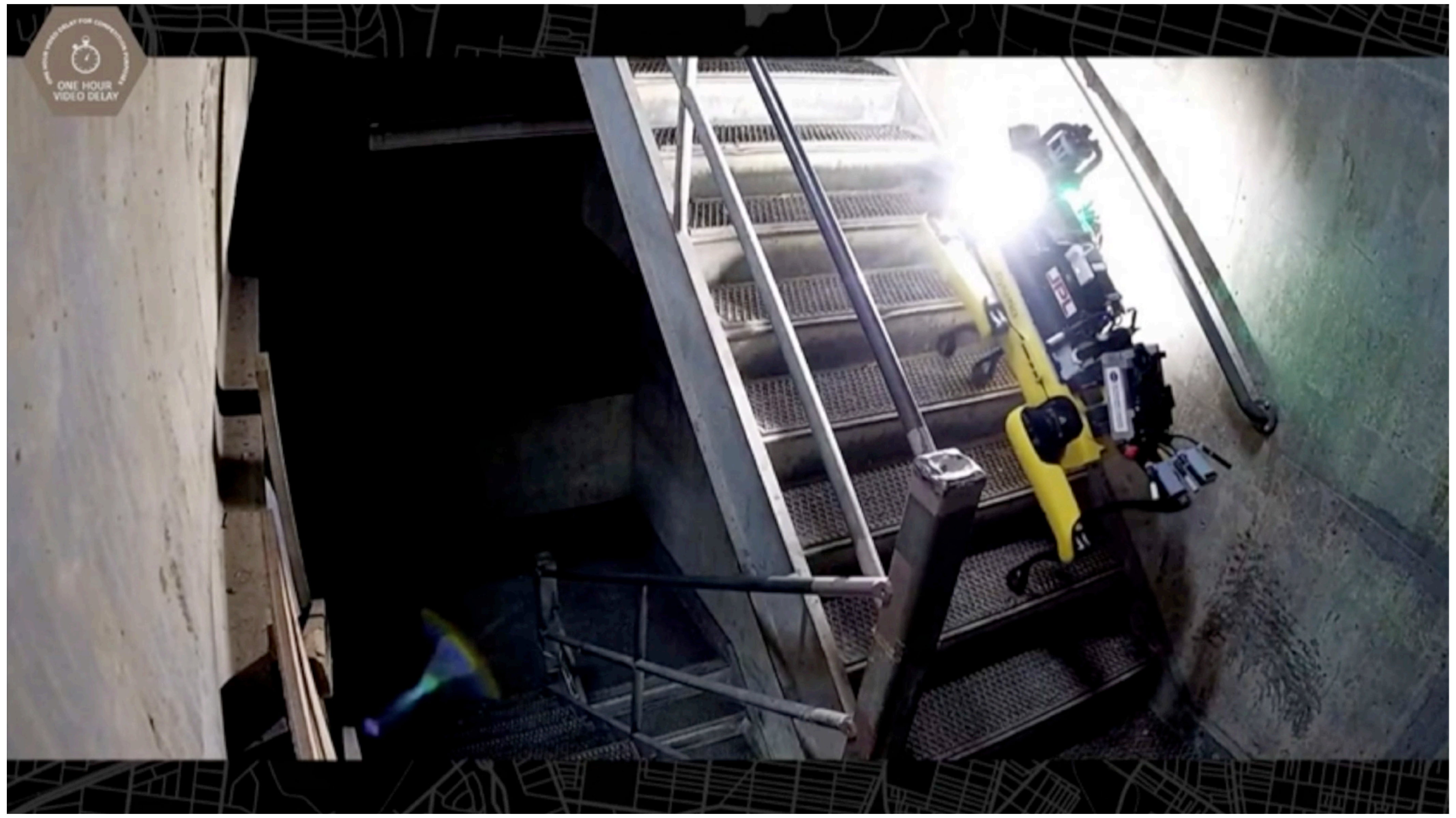


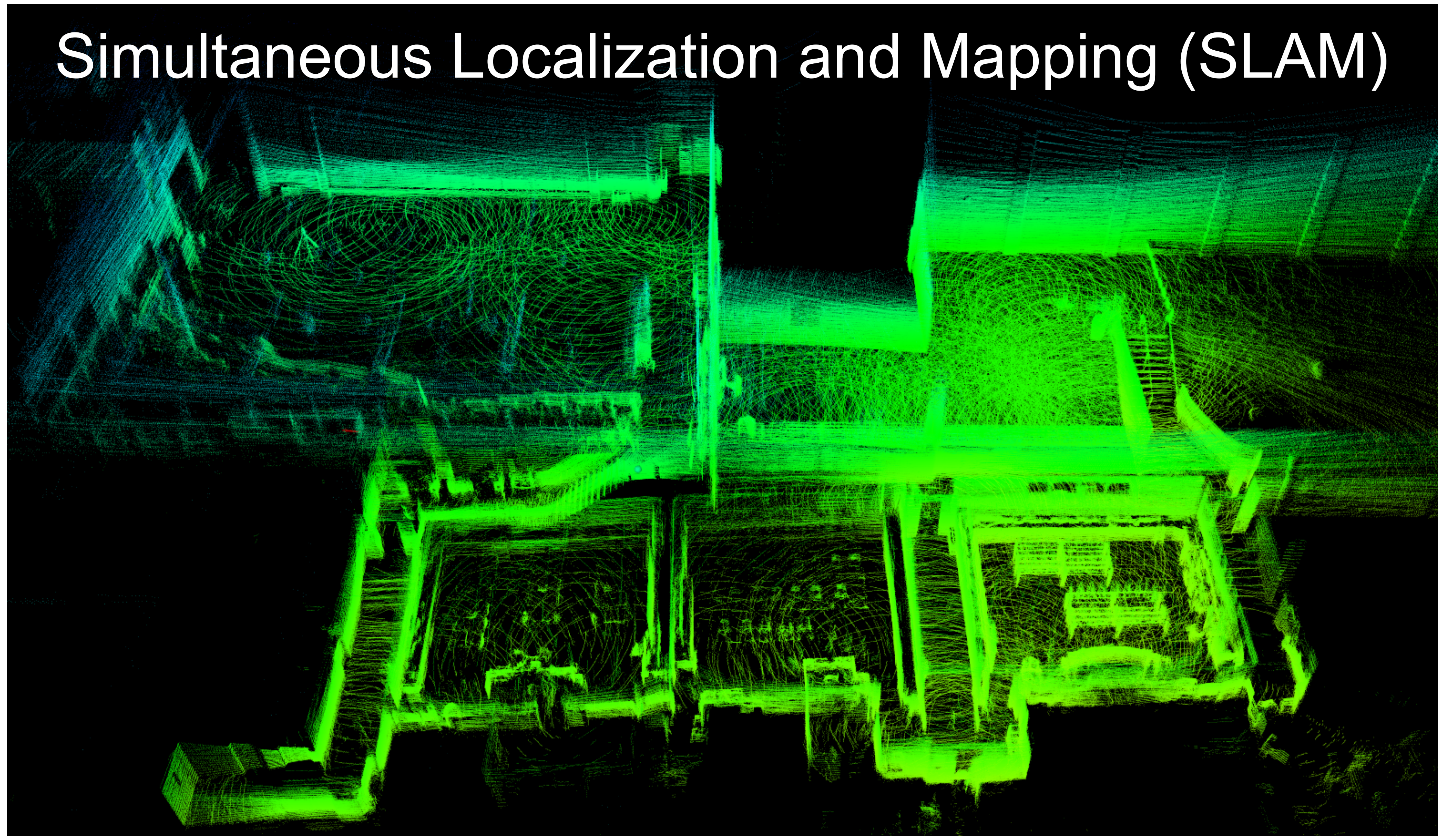
- Certifiable spacecraft pose estimation
- SPEED benchmarking dataset

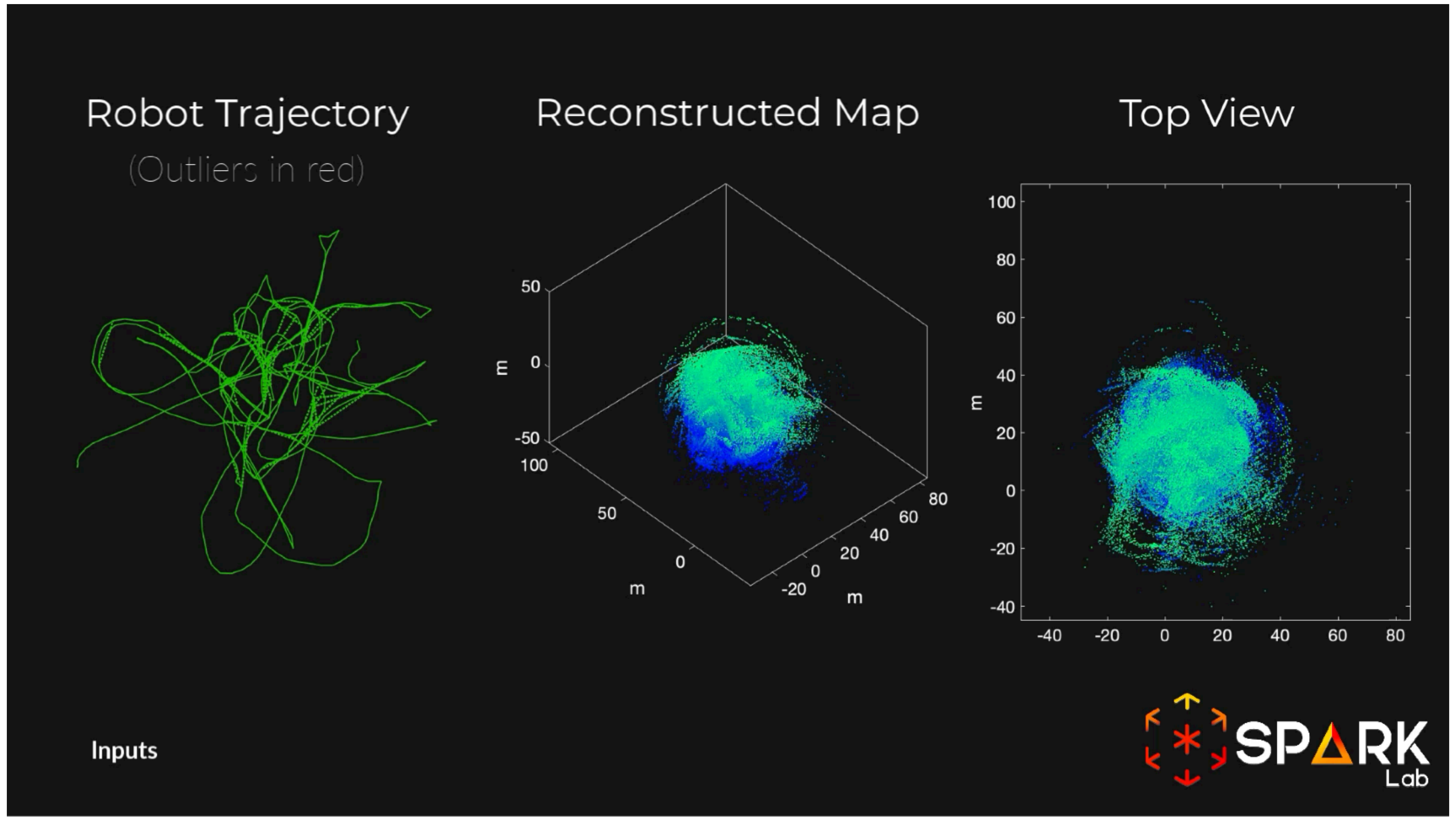
[Yang and Carlone. In Perfect Shape: Certifiably Optimal 3D Shape Reconstruction from 2D Landmarks. CVPR 2020]
[Yang and Carlone, One Ring to Rule Them All: Certifiably Robust Geometric Perception with Outliers, NeurIPS'20.]

Towards certifiable localization and mapping

DARPA Subterranean Challenge - in collaboration with JPL & Caltech



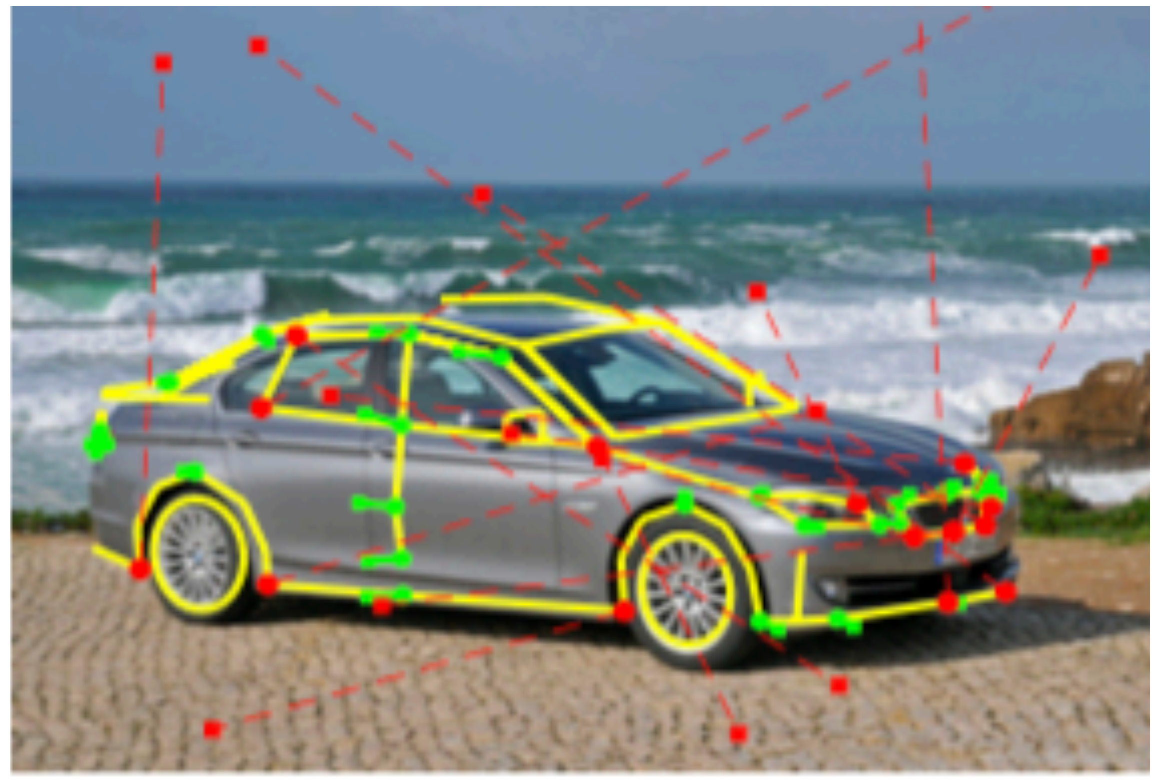




- Graduated Non-Convexity as a scalable (but not yet certifiable) tool for convexification

Yang, Antonante, Tzoumas, Carlone. Graduated non-convexity for robust spatial perception: from non-minimal solvers to global outlier rejection. RAL 2020. (best paper in robot vision at ICRA 2020)

Outline



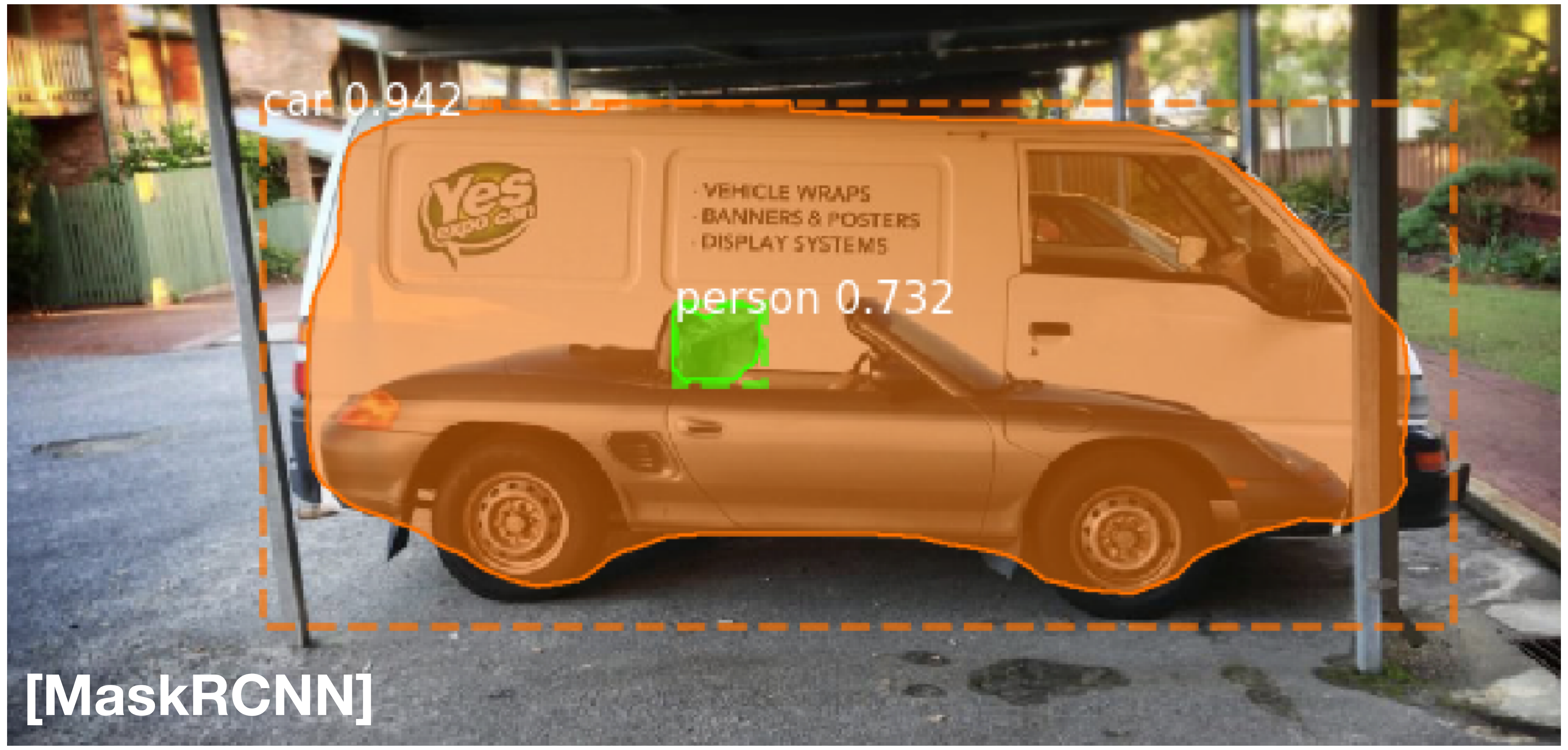
Certifiable Perception Algorithms

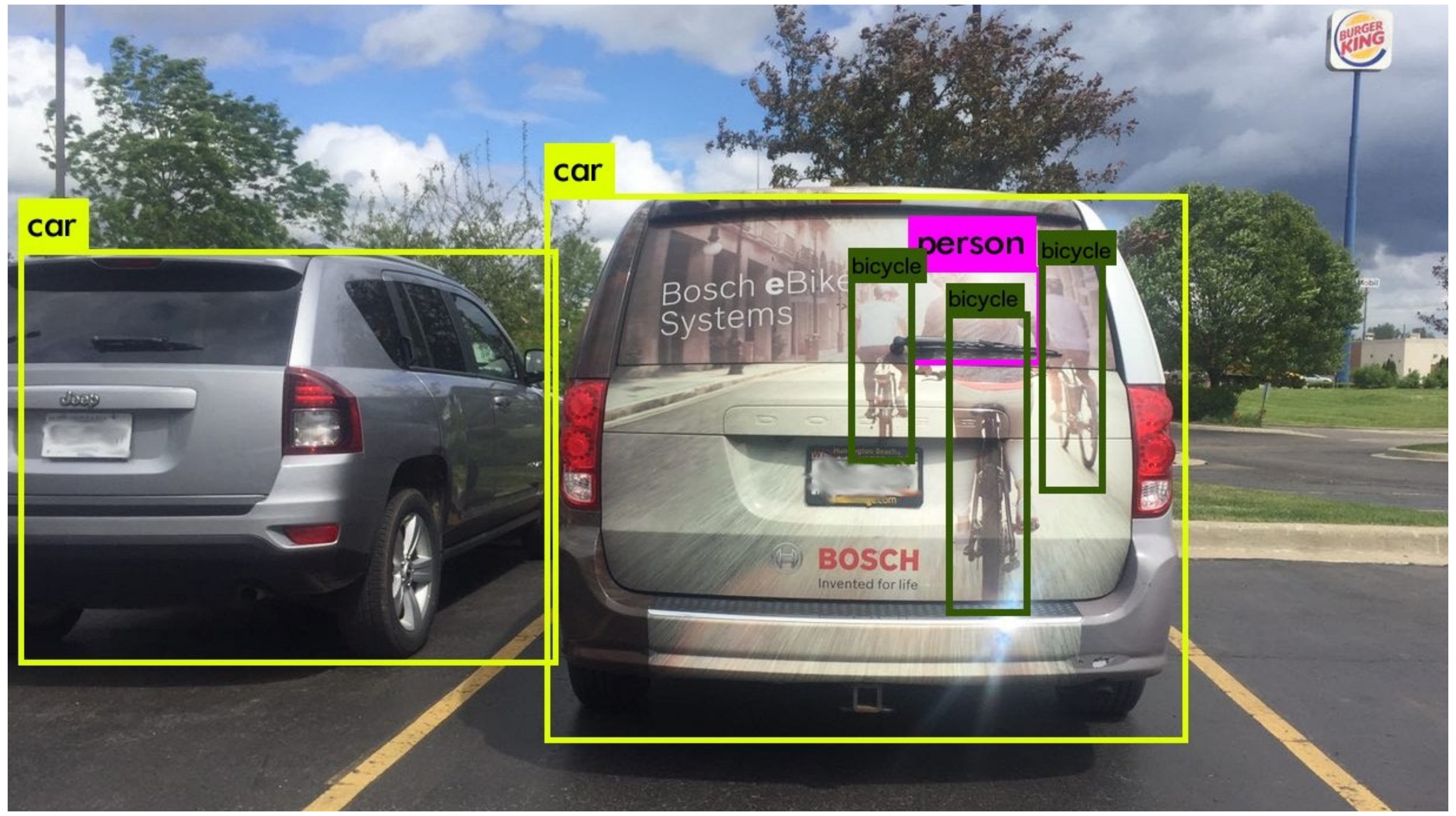
Towards System-level Guarantees & Real-time High-level Understanding

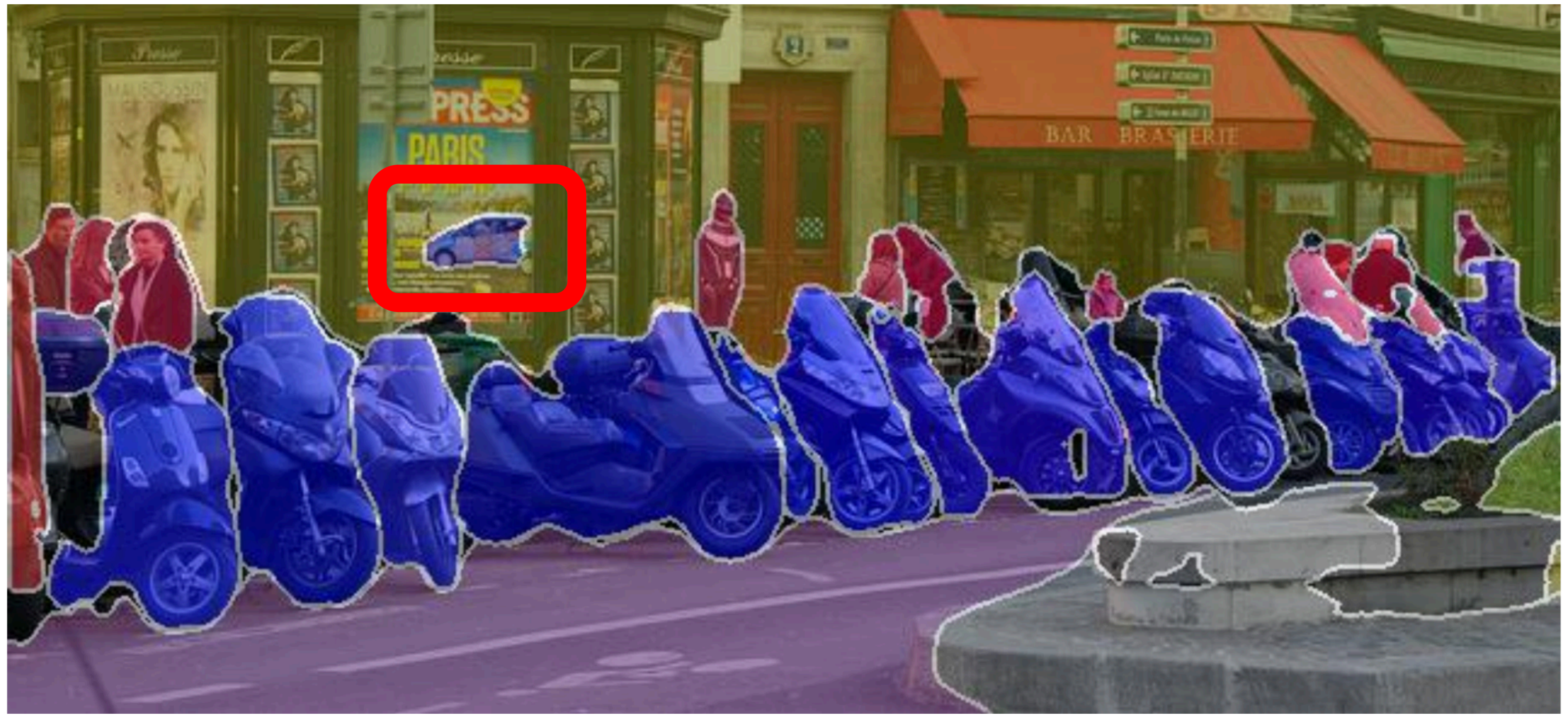
Robust perception requires high-level 3D understanding

Robustness of human perception relies on simultaneously understanding geometry, semantics, physics, and relations in 3D

2D understanding is doomed to fail





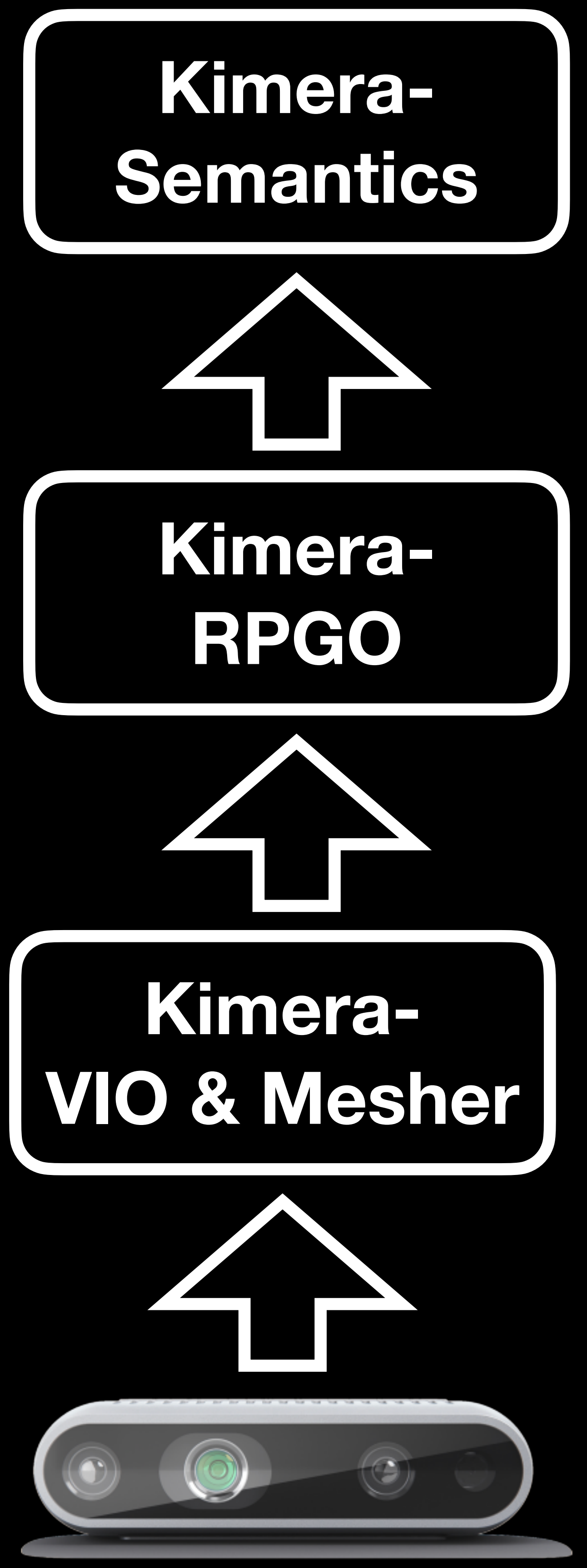


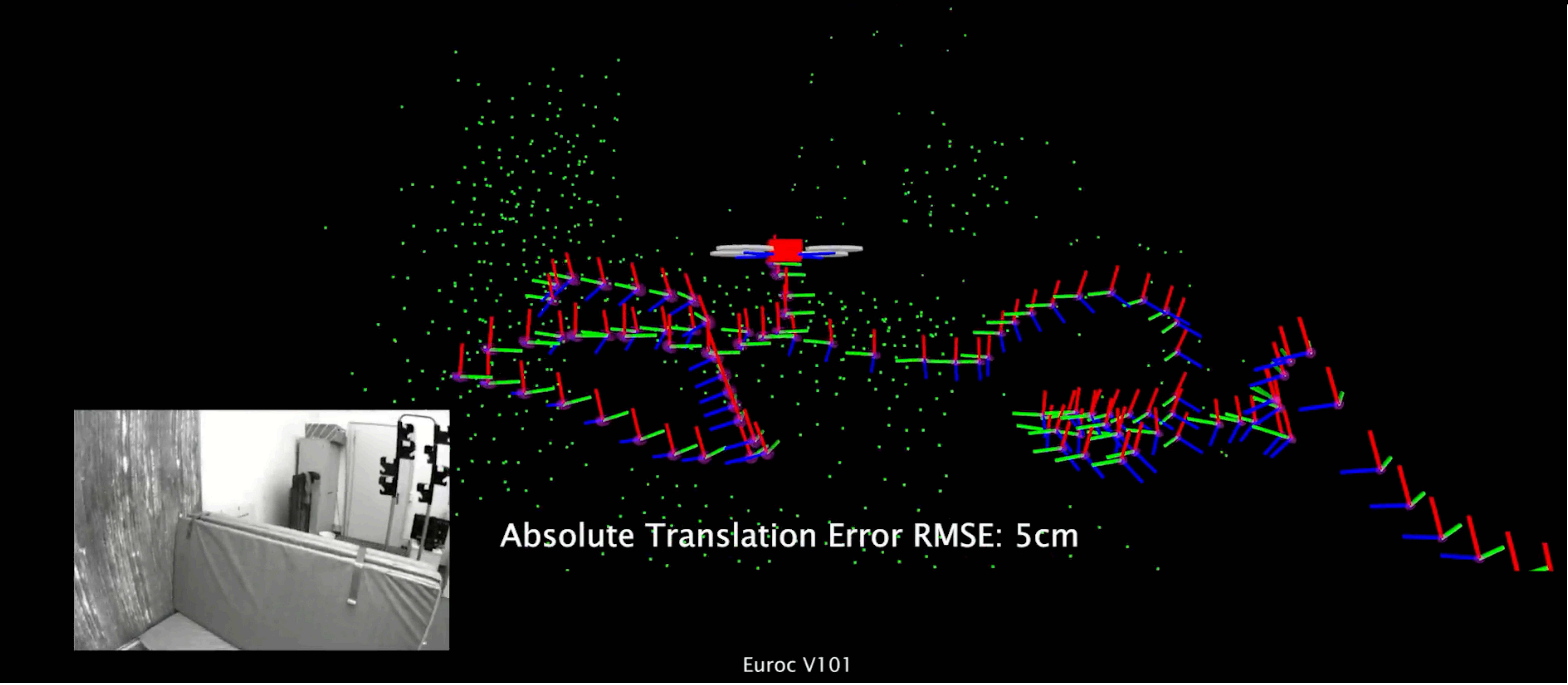
Robust perception requires: understanding
geometry, semantics, physics, and relations in 3D

Kimera: real-time 3D metric-semantic understanding

[Rosinol, Abate, Chang, Carlone. Kimera: an open-source library for real-time metric-semantic localization and mapping. ICRA 2020]

Kimera: real-time 3D metric-semantic understanding



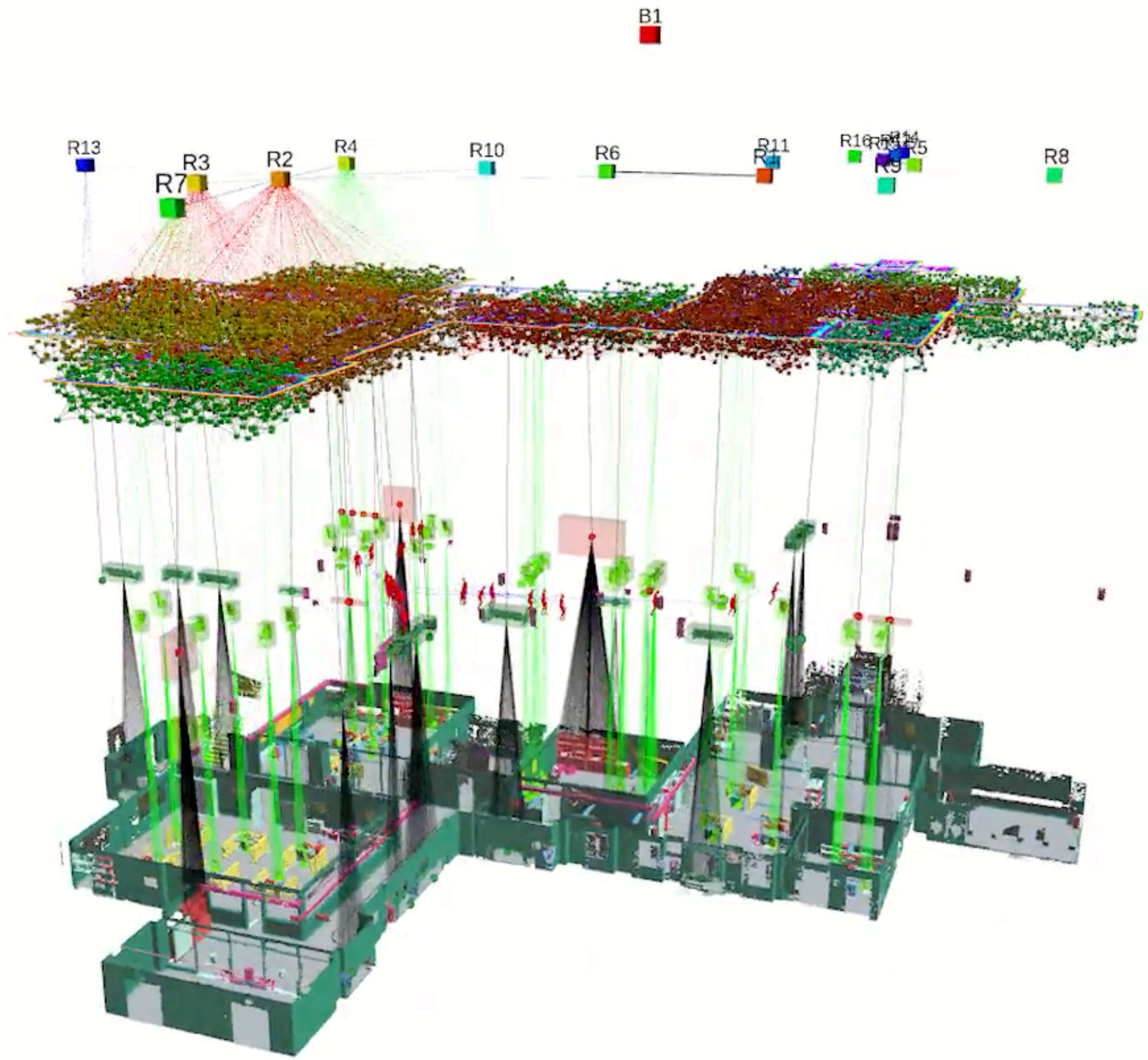


- **Insights**: IMU preintegration, GTSAM, Pairwise Consistency Maximization, VoxBlox, SemanticFusion

Fast multi-threaded open-source code: https://github.com/MIT-SPARK/Kimera

[Rosinol, Abate, Chang, Carlone. Kimera: an open-source library for real-time metric-semantic localization and mapping. ICRA 2020]

3D Dynamic Scene Graphs



Directed graph, where:

- **nodes** are *spatial concepts*
- **edges** represent spatio-temporal relations between concepts (e.g., inclusion at time t)

OPPORTUNITIES

- *reason over context and relations among objects to conclude on plausibility*
- *bridge task planning and motion planning*
- *human-robot interaction*

[Rosinol, Gupta, Abate, Shi, Carlone. 3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans. RSS 2020]

- **Goal**: 3D Dynamic Scene Graph from visual-inertial data in a heavily populated environment
- From SLAM algorithms to a **Spatial Perception Engine**, that infers geometry, semantics, a hierarchy of high-level spatial concepts and their relations

Conclusion

- Getting low failure rates and performance guarantees in spatial perception requires rethinking current algorithms
    ➡ **Certifiable algorithms** as a practical approach to get robust performance
- We need a theory of how to connect robust algorithms into a robust system
    ➡ **3D high-level understanding** is key to true robustness (enabled by Kimera and 3D Dynamic Scene Graphs)

- **Open questions**:
    - Certifiable algorithms for other problems
        - theory, implementations, learning
    - 3D Dynamic Scene Graphs
        - Tightly-coupled metric-semantic
        - Learning, physics, causality
    - Theory of robust perception systems

One Ring to Rule Them All: Certifiably Robust Geometric Perception with Outliers

Heng Yang and Luca Carlone
Laboratory for Information and Decision Systems (LIDS)
Massachusetts Institute of Technology
{hankyang,lcarlone@mit.edu}

NeurIPS'20

Monitoring and Diagnosability of Perception Systems

Pasquale Antonante David I. Spivak Luca Carlone
Massachusetts Institute of Technology
Cambridge, MA 02139
{antonap, dspivak, lcarlone}@mit.edu

Arxiv'20

References

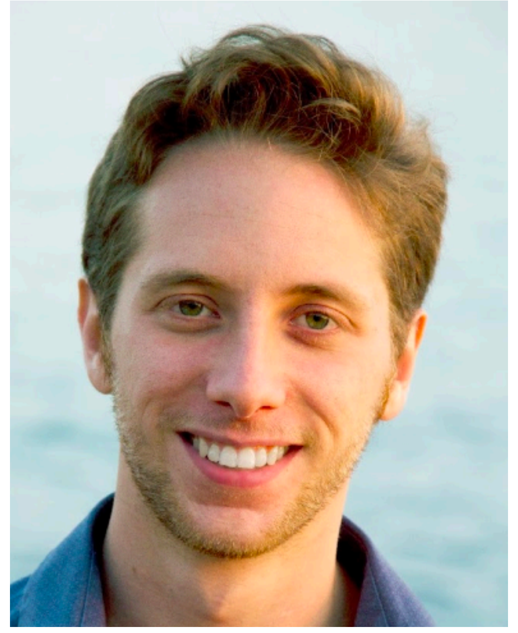
Certifiably robust perception algorithms and systems

- Yang and Carlone, One Ring to Rule Them All: Certifiably Robust Geometric Perception with Outliers, NeurIPS'20.
- Antonante, Spivak, Carlone, Monitoring and Diagnosability of Perception Systems, Arxiv'20.
- Yang, Shi, Carlone. TEASER: Fast and Certifiable Point Cloud Registration. T-RO 2020.
- Yang, Antonante, Tzoumas, Carlone. Graduated non-convexity for robust spatial perception: from non-minimal solvers to global outlier rejection. RAL 2020. (best paper in robot vision at ICRA 2020)
- Yang and Carlone. In Perfect Shape: Certifiably Optimal 3D Shape Reconstruction from 2D Landmarks. CVPR 2020.
- Yang and Carlone. A Polynomial-time Solution for Robust Registration with Extreme Outlier Rates. RSS 2019.
- Yang and Carlone. A quaternion-based certifiably optimal solution to the Wahba problem with outliers. ICCV 2019.
- Tzoumas, Antonante, Carlone. Outlier-robust spatial perception: Hardness, general-purpose algorithms, and guarantees. IROS 2019.
- Lajoie, Hu, Beltrame, Carlone, Modeling Perceptual Aliasing in SLAM via DCGM, RAL'19.

High-level understanding: 3D Dynamic Scene Graphs and Kimera

- Rosinol, Gupta, Abate, Shi, Carlone. 3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans. RSS 2020.
- Rosinol, Abate, Chang, Carlone. Kimera: an open-source library for real-time metric-semantic localization and mapping. ICRA 2020.

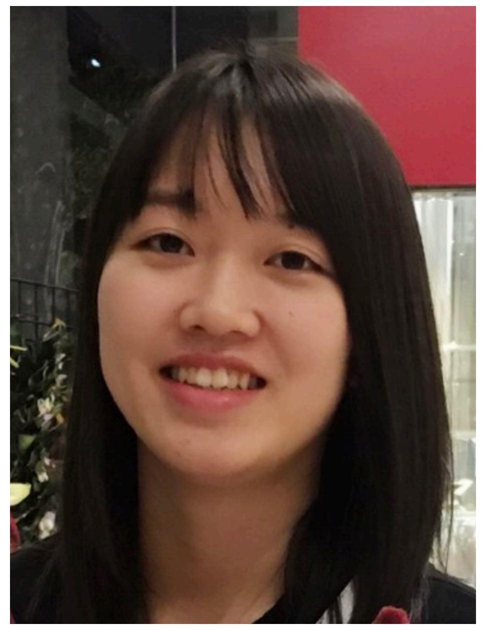
Heng Yang



Pasquale Antonante

Jingnan Shi

Vasileios Tzoumas



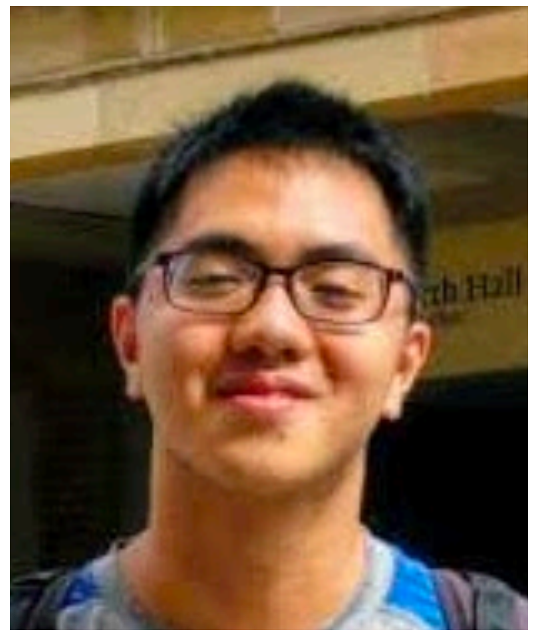
Siyi Hu

Luca Ballotta

Antoni Rosinol

Marcus Abate

Arjun Gupta



Yun Chang

Joshua Fishman

Andrew Violette

Thank you!

MIT AEROASTRO

ARL

MathWorks

DARPA

MIT LINCOLN LABORATORY

amazon robotics

Office of Naval Research ONR Science & Technology

Ford

Google