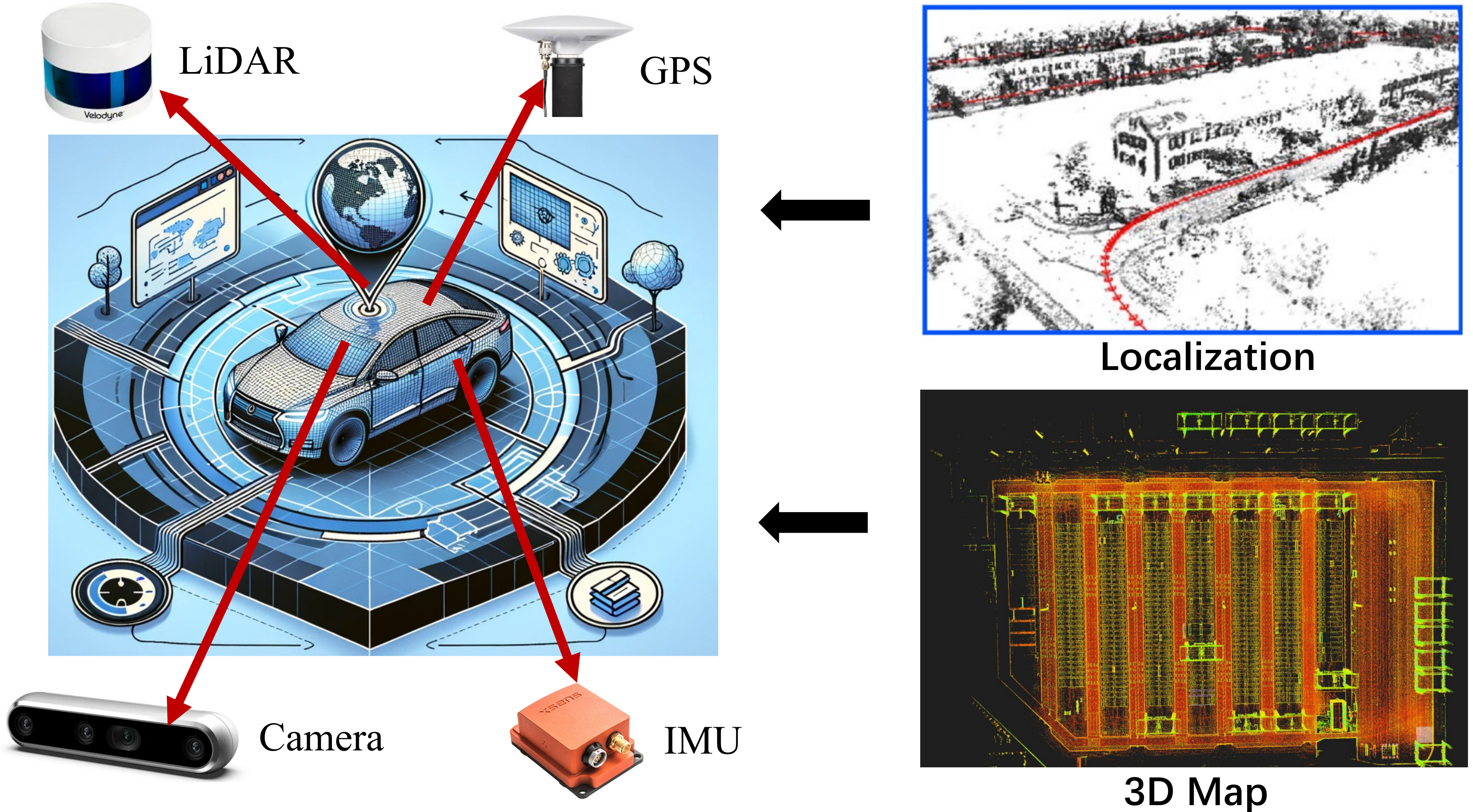
VILAM: Infrastructure-assisted 3D Visual Localization and Mapping for Autonomous Driving

Jiahe Cui[1,2], Shuyao Shi[2], Yuze He[2], Jianwei Niu*[1], Guoliang Xing*[2], and Zhenchao Ouyang[1]

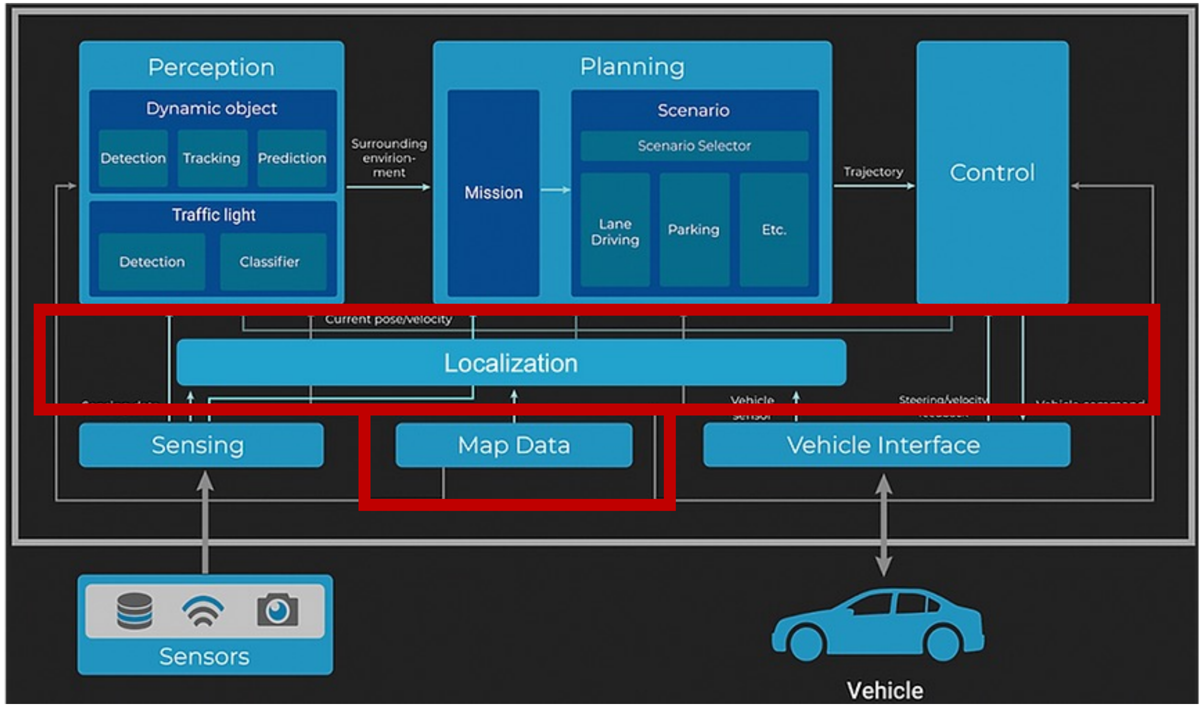
[1]Beihang University,

[2]The Chinese University of Hong Kong

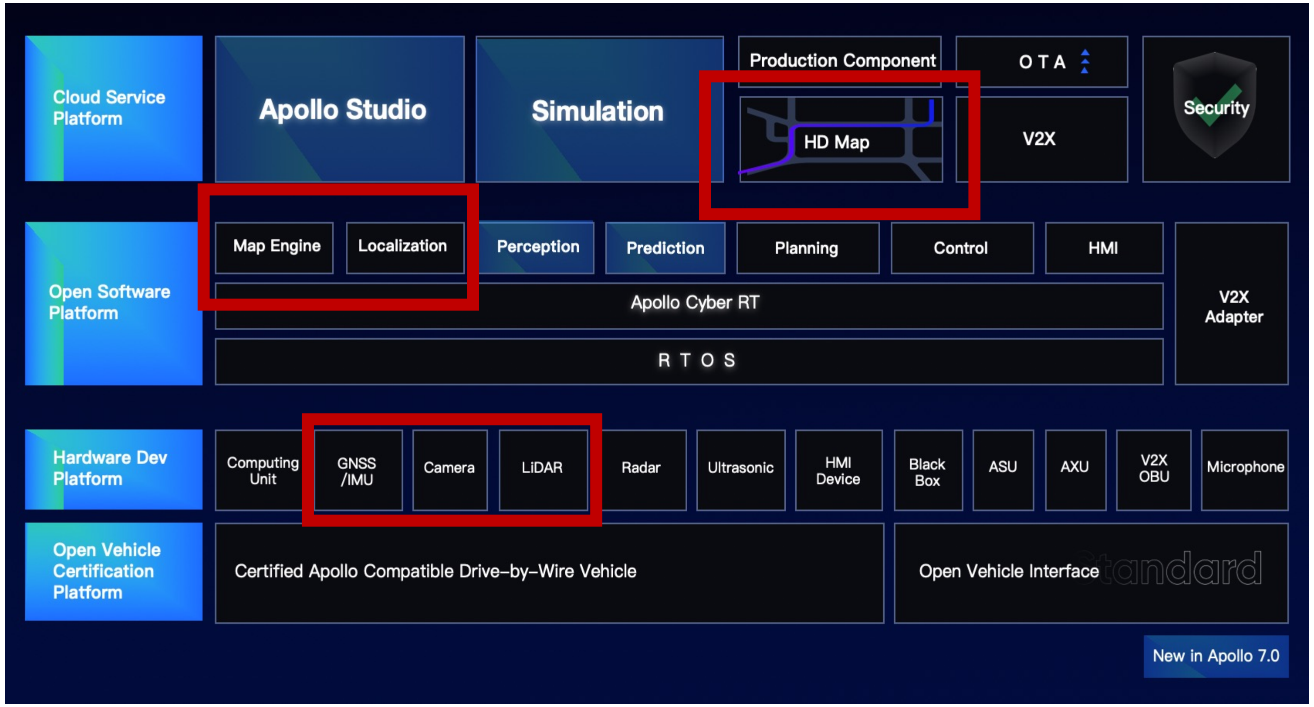
Localization and Mapping are Critical for Autonomous Driving



Localization and Mapping are Critical for Autonomous Driving



Autoware Autonomous Driving Stack



Apollo Autonomous Driving Stack

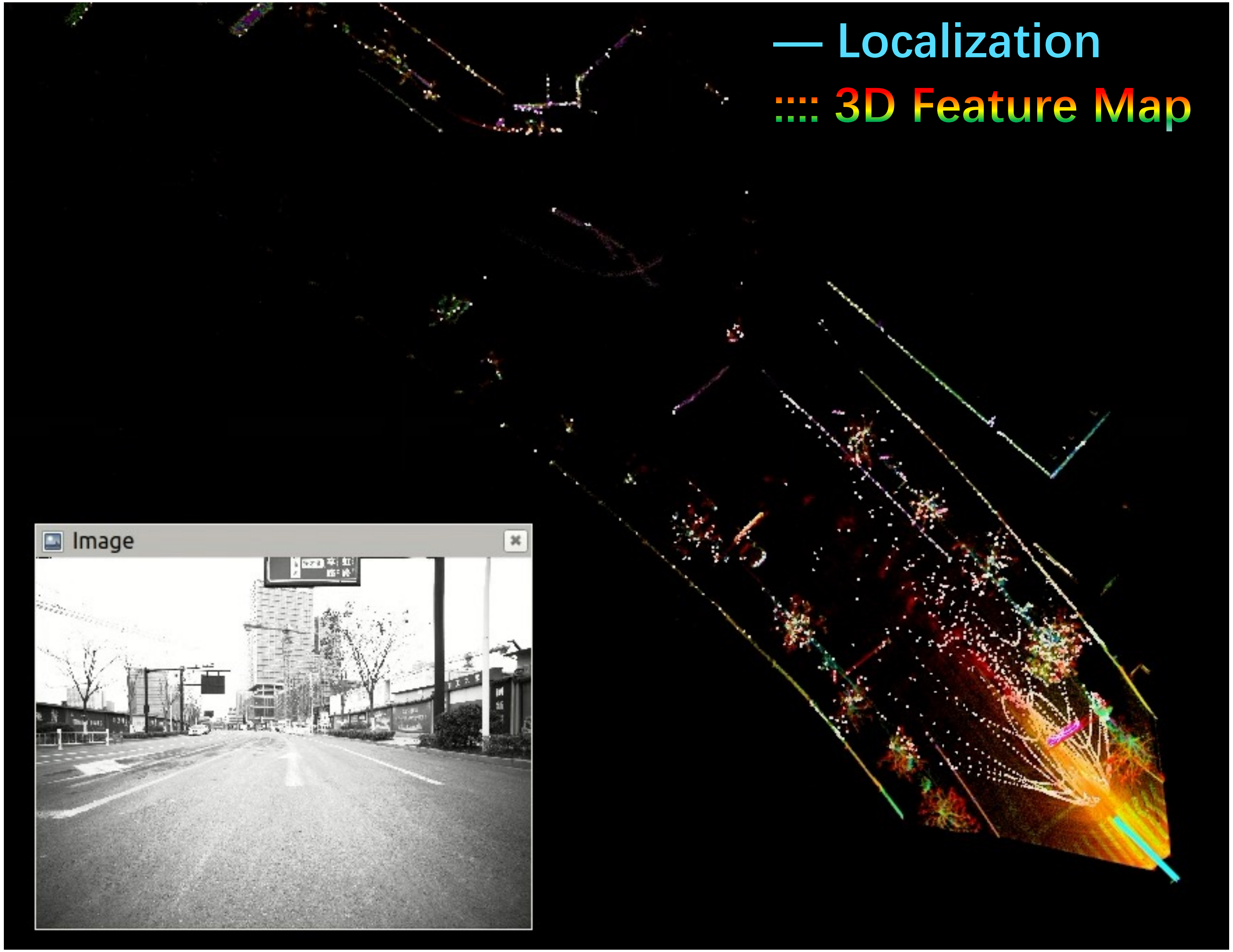
- Consumer-grade GPS: **5~10m**
- Google Map: **10~20m**
- Autonomous Driving's Requirement: **<< 1m**

Visual Simultaneously Localization and Mapping (Visual SLAM)

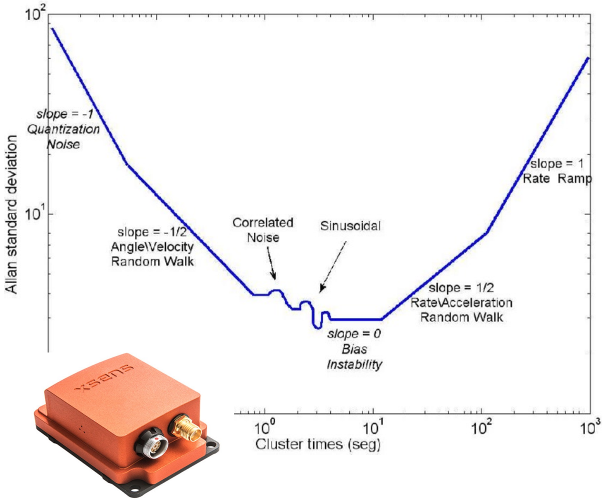
Camera

IMU

Low-cost sensors



Visual SLAM – Cumulative Drift



Inherent sensor distortion of camera / IMU

Feature extraction error

Feature mismatching

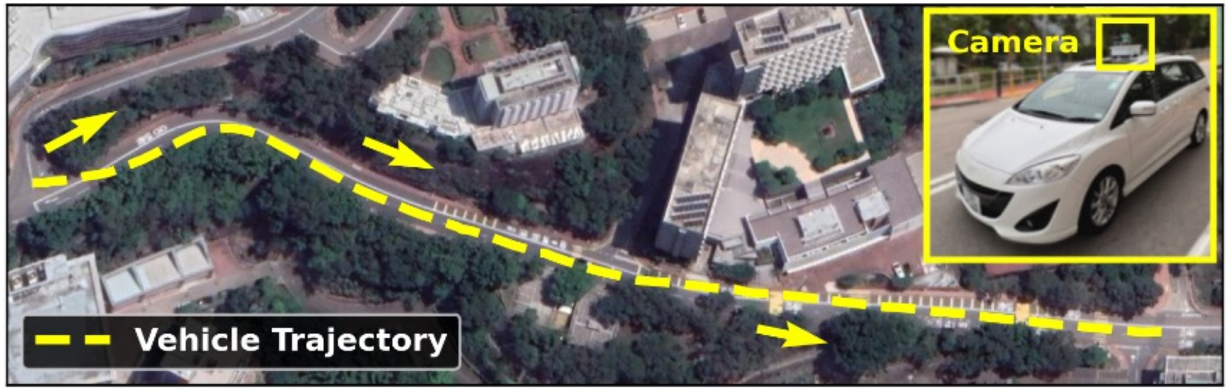
⬇ ⬇

Frame-by-frame accumulation

⬇

Cumulative drift !

Visual SLAM – Cumulative Drift



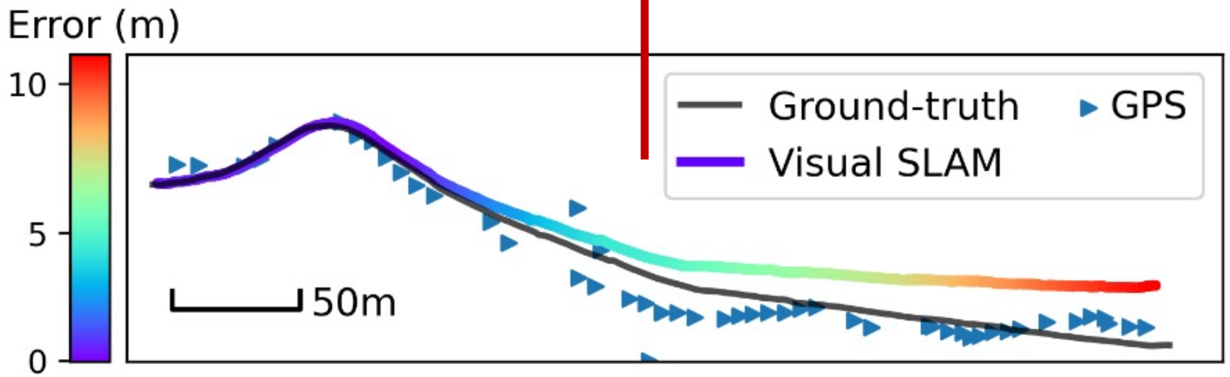
(a) The real-world test scene and the test vehicle.

- Inherent sensor noises of camera / IMU
- Feature extraction error
- Matching error
- …

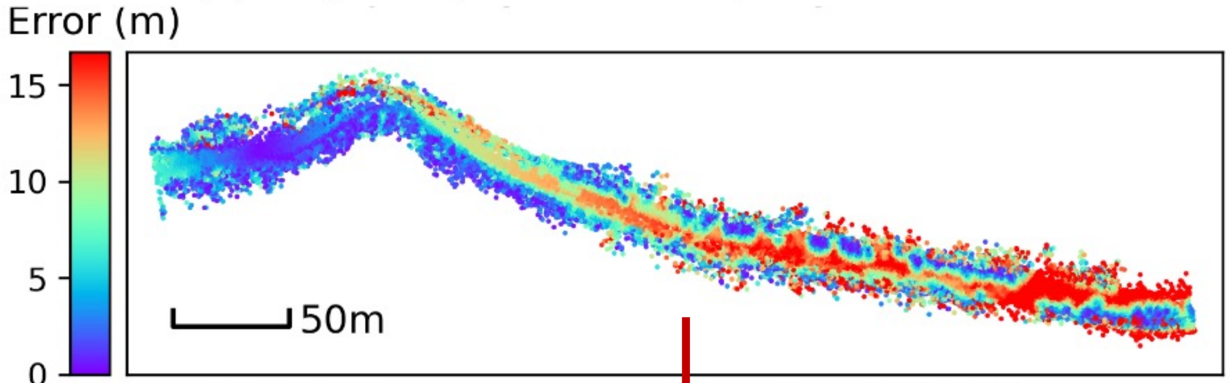
⬇

Cumulative drift !

Accumulated Drift of Localization（>10m)



(b) Trajectory estimated by visual SLAM.

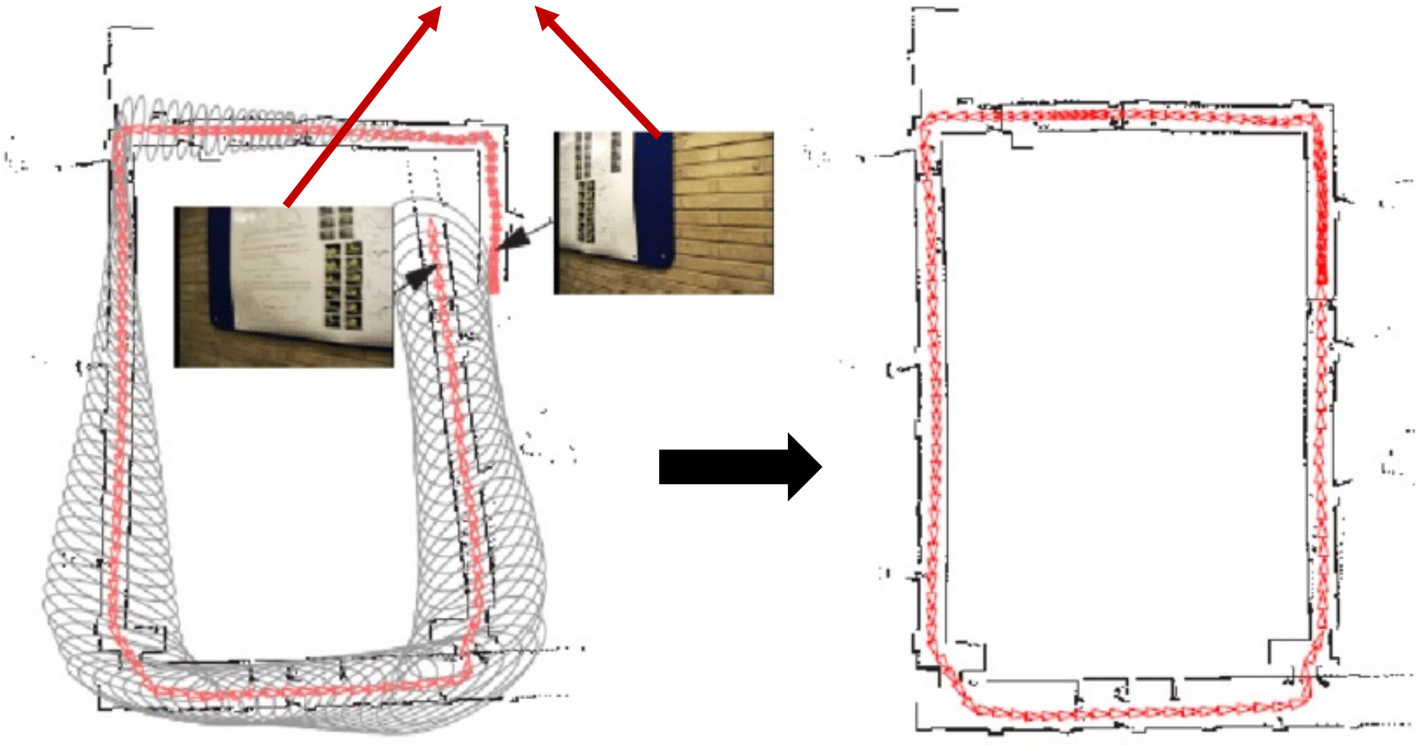


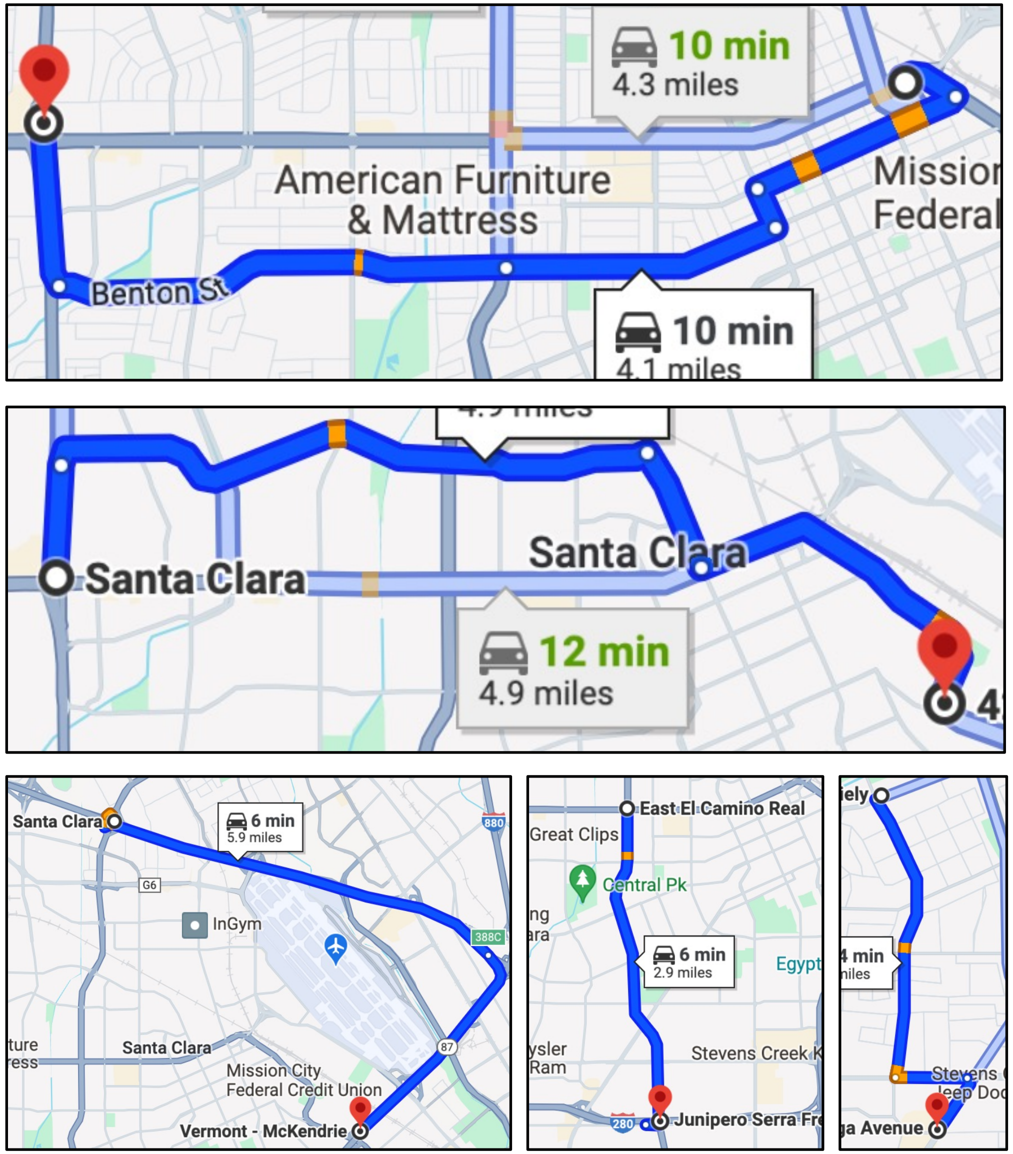
(c) 3D Map constructed by visual SLAM.

Accumulated Drift of Map（>15m)

Eliminating Cumulative Drift

Visited Place?

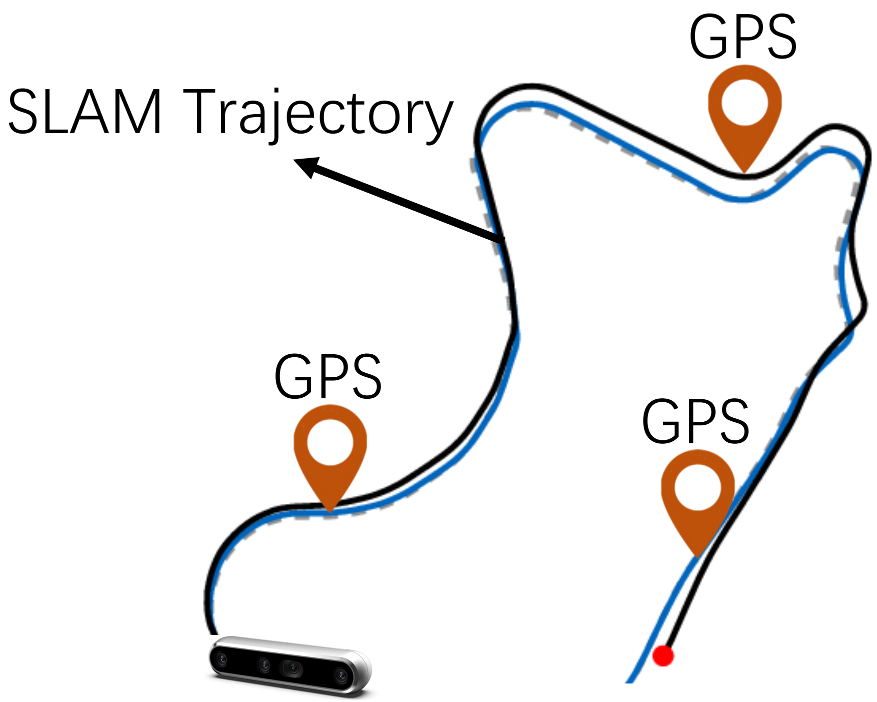


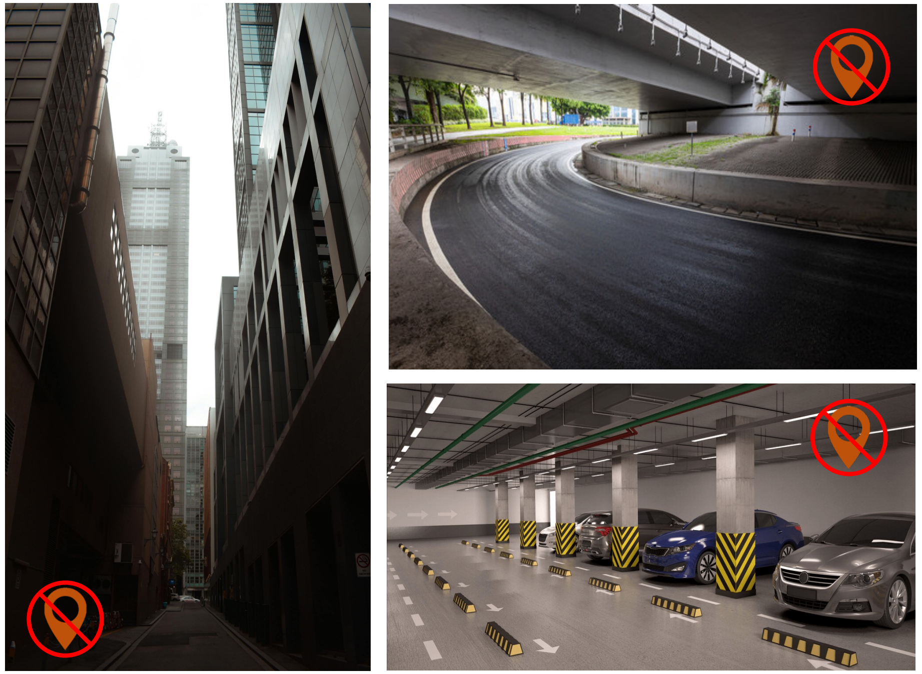


- Most existing studies employ **Loop-Closing** to correct **Cumulative Drift**

Autonomous Driving: No Closed Loops!

Eliminating Cumulative Drift





- Employ **GPS positions** to correct **Cumulative Drift**

GPS is not available in many environments

Consumer-grade GPS
Low accuracy (5~10m)

↔

High-end RTK-INS-GPS
Expensive ($4,000)

How to effectively eliminate the cumulative drift of SLAM?

1. Prior measurement of the environment

(Loop closing)

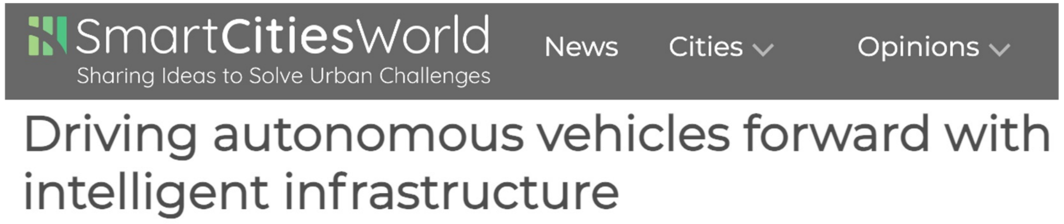
2. External localization

(GPS positions)

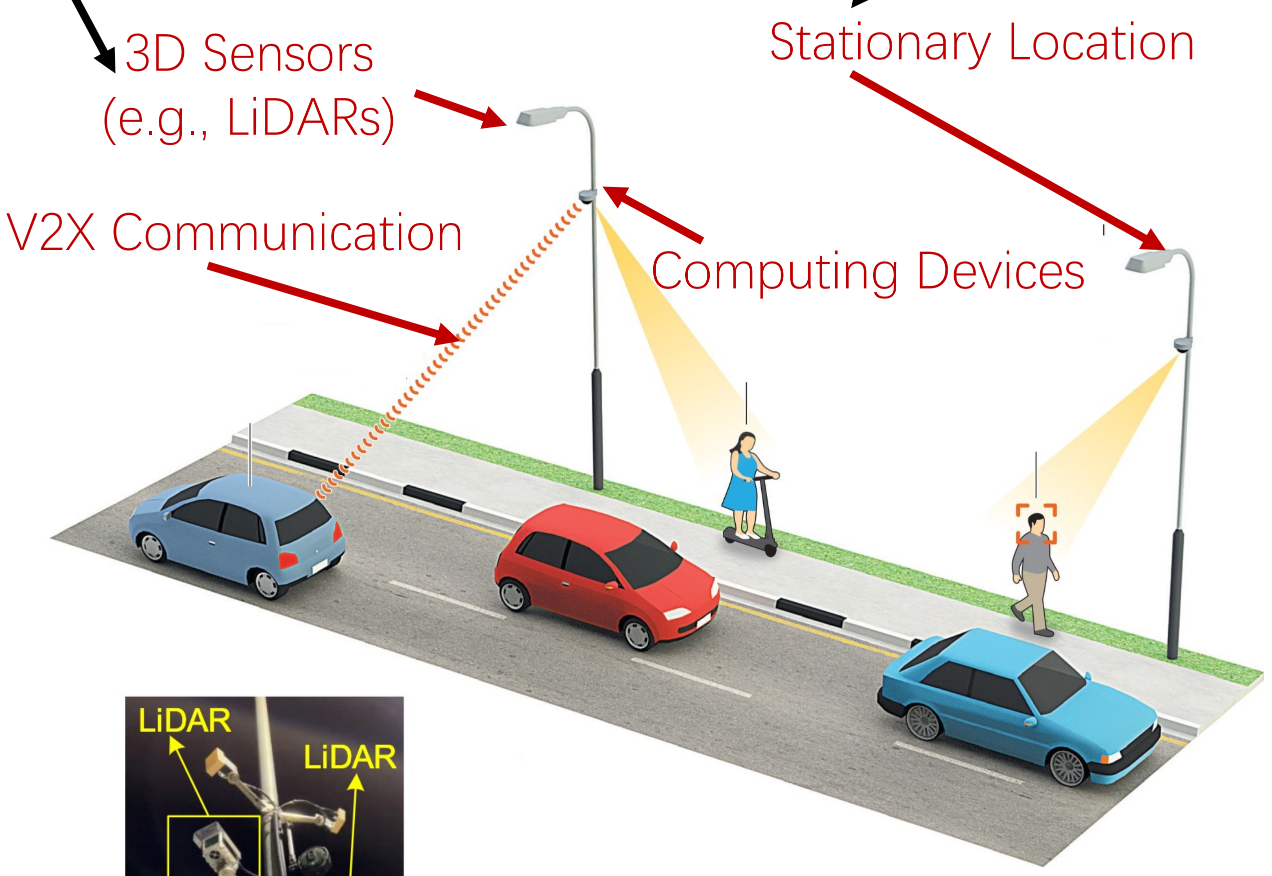
Smart Roadside Infrastructure

1. Prior measurement

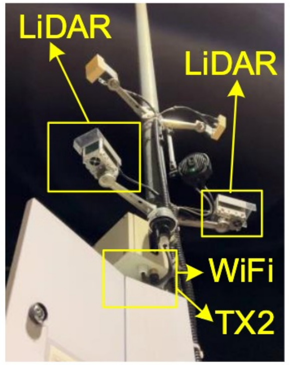
2. External localization



. . .

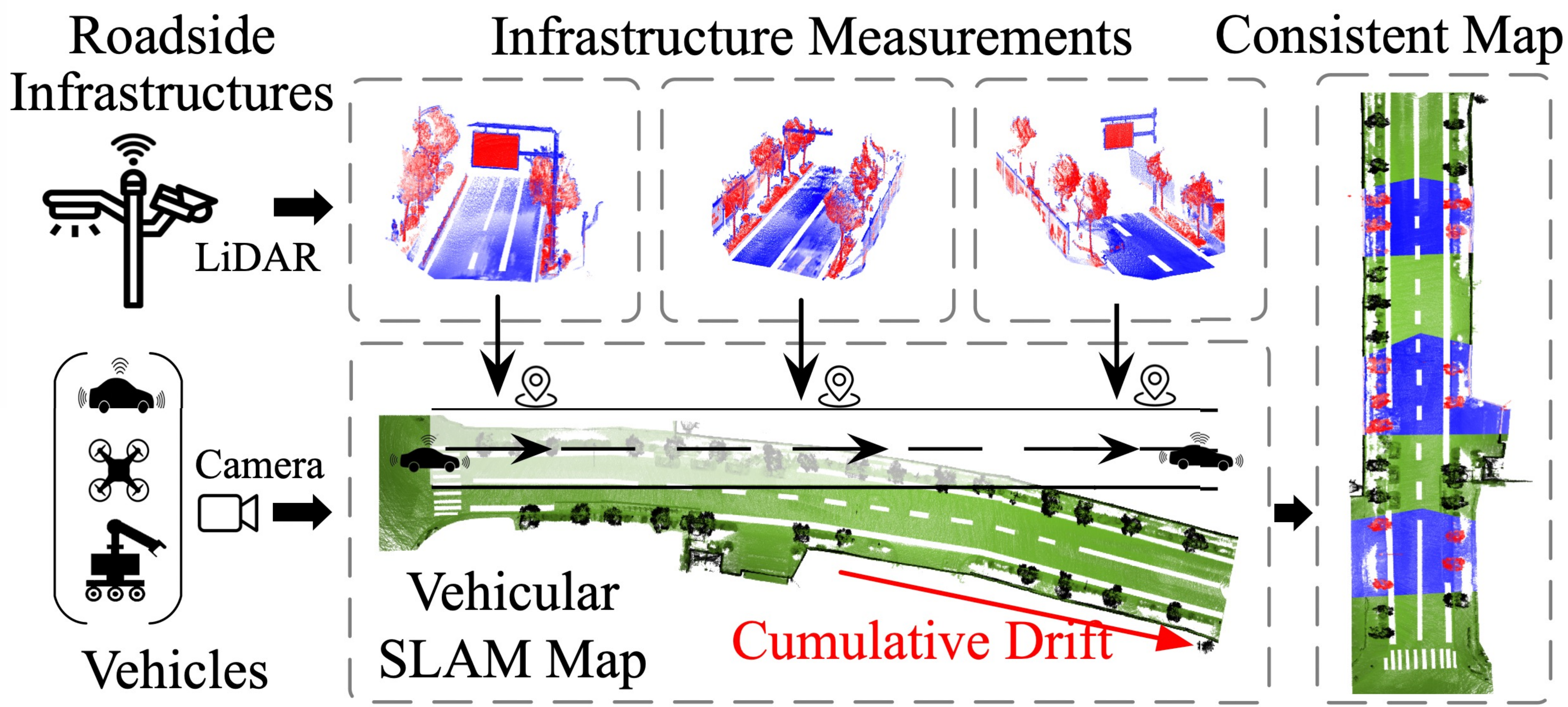


Source: The Straits Times



Smart roadside infrastructure

VILAM: Infrastructure-assisted 3D Visual Localization and Mapping for Autonomous Driving

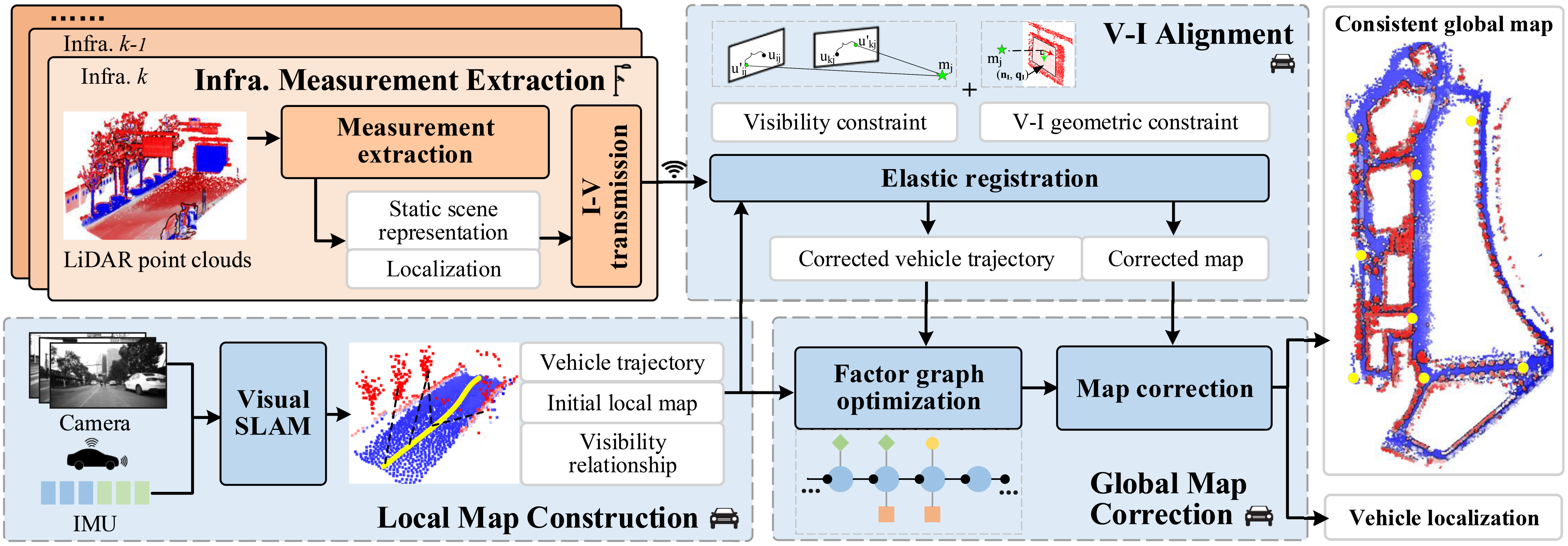


Key idea: Utilizing the accurate scene measurement from distributed infrastructures as global reference to correct the cumulative drift of vehicle SLAM

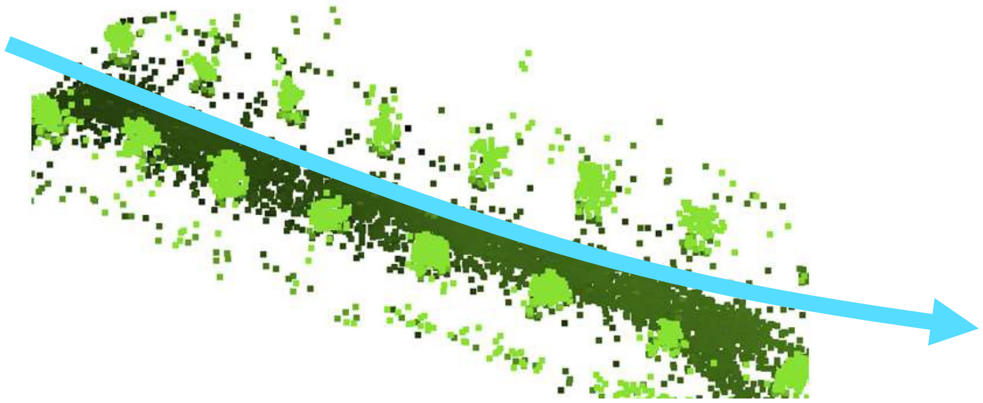
VILAM: Infrastructure-assisted 3D Visual Localization and Mapping for Autonomous Driving

Pipeline

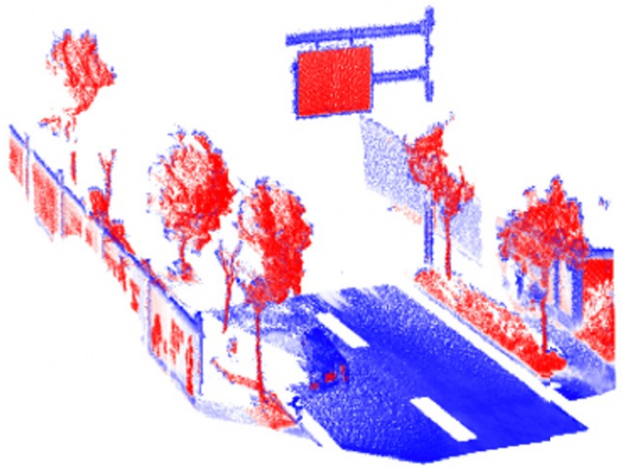
Infrastructure-side operation Vehicle-side operation



Vehicle-Infrastructure Alignment



Drifted Vehicle Trajectory
Drifted Vehicle-side SLAM Map



Infrastructure Measurement

↓ ↓

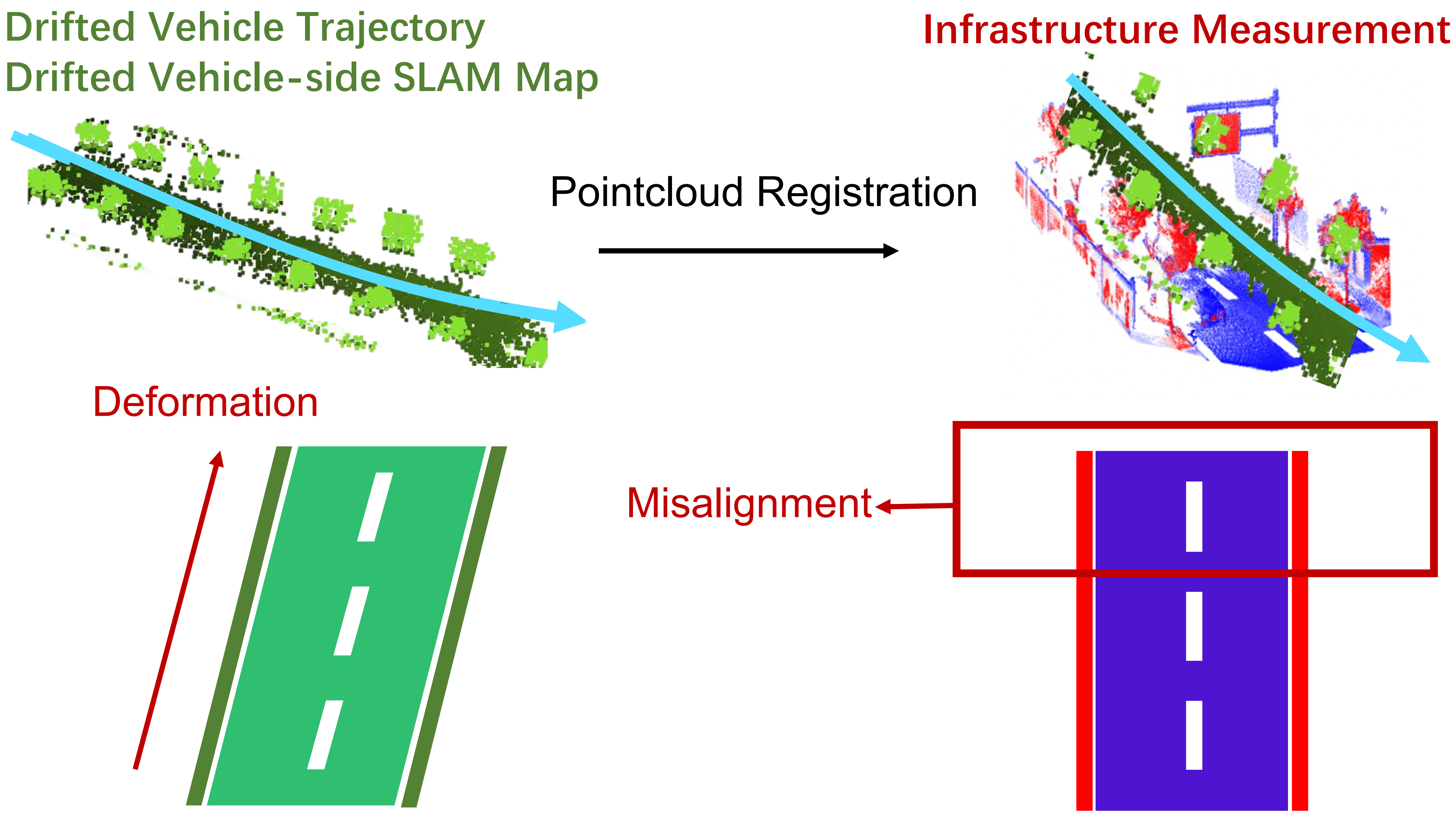
?

↓ ↓

Corrected Vehicle Trajectory

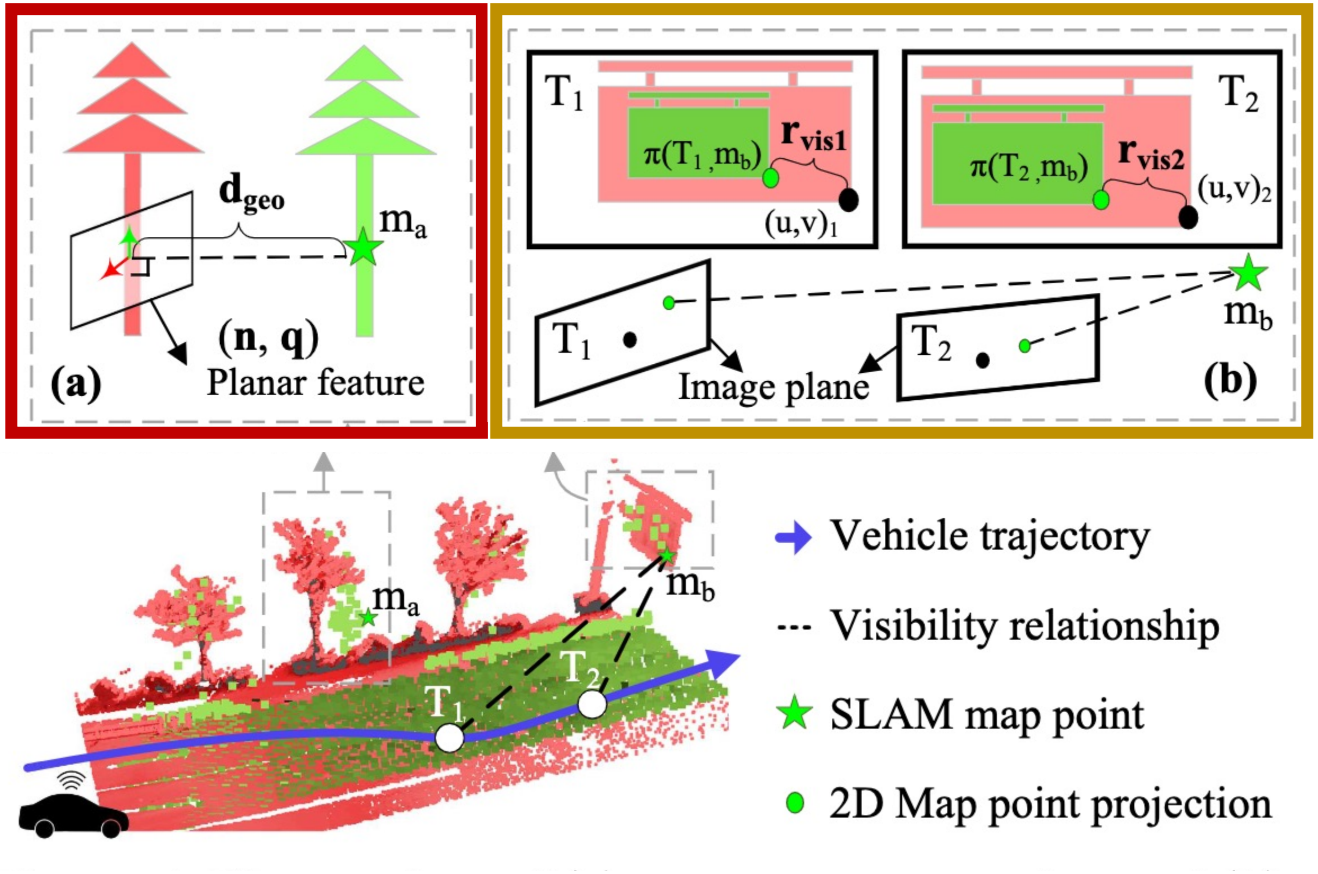
Corrected Vehicle-side Map

Challenge: Deformation of Vehicle-side Map



VILAM's Solution: Elastic Registration

■Elastic Registration



Infrastructure measurement   Vehicle-side local map

(a) Geometric Constraint:

$$\mathbf{d}_{\text{geo}}(\mathbf{m}_j) = (\mathbf{m}_j - \mathbf{q}_j)^T \cdot \mathbf{n}_j$$

(b) Visibility Constraint:

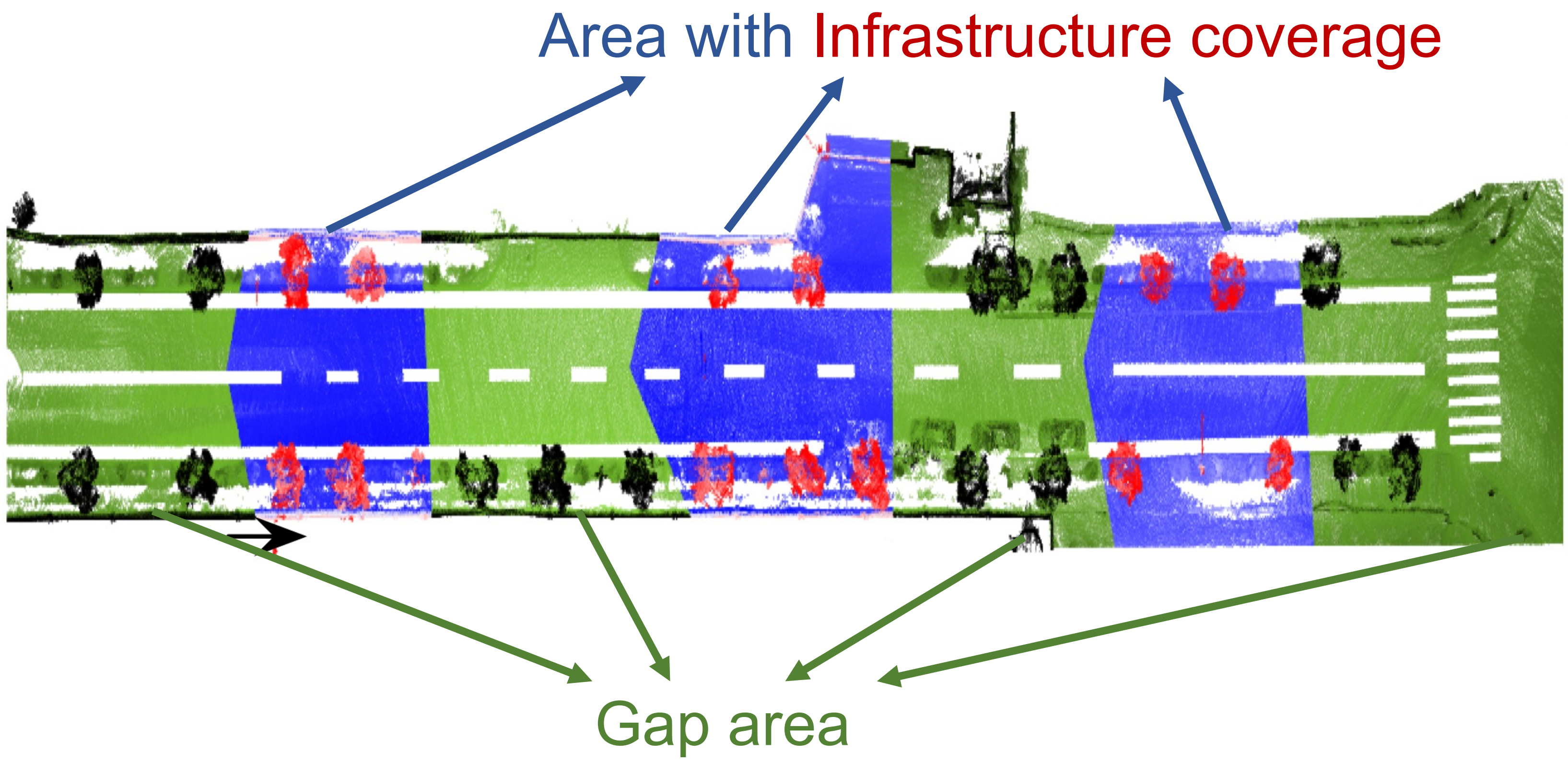
$$\mathbf{r}_{\text{vis}}(\mathbf{m}_j, \tilde{\mathcal{T}}) = \sum_i \left\| \pi\left((\mathbf{T}_i)^{-1}\mathbf{m}_j\right) - (u, v)^T \right\|$$

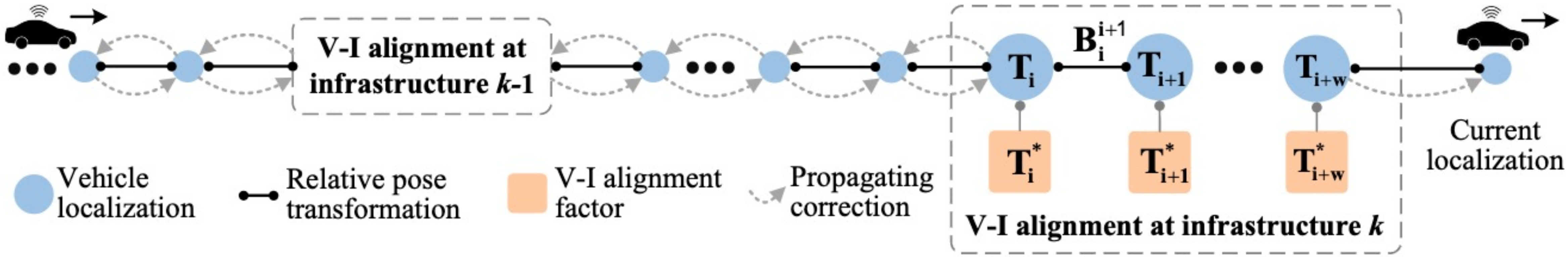
Cost Function (Geo and Vis)

$$\tilde{\mathcal{T}}^*, \tilde{\mathcal{M}}^* = \arg\min_{\tilde{\mathcal{T}}, \tilde{\mathcal{M}}} \sum_j \left\| \begin{bmatrix} \mathbf{d}_{\text{geo}} \\ \mathbf{r}_{\text{vis}} \end{bmatrix} \right\|_2^2$$

More to Considerate: Global Map Correction

Real-world Deployment:

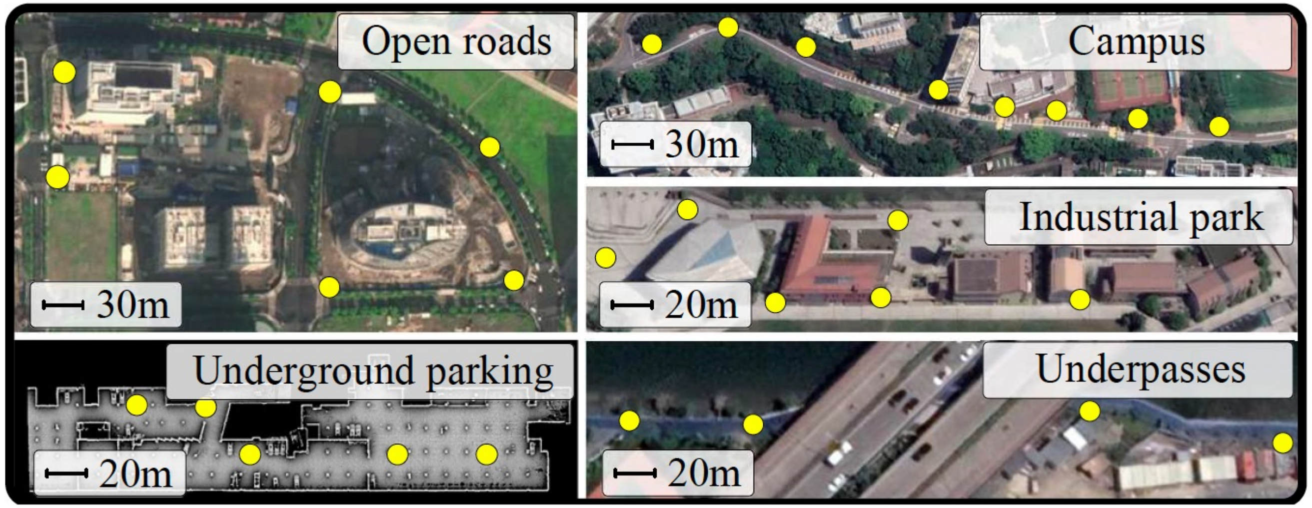




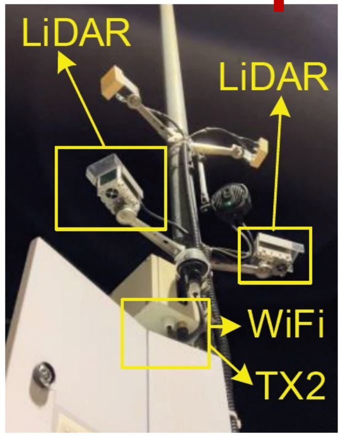
Testbed and Datasets

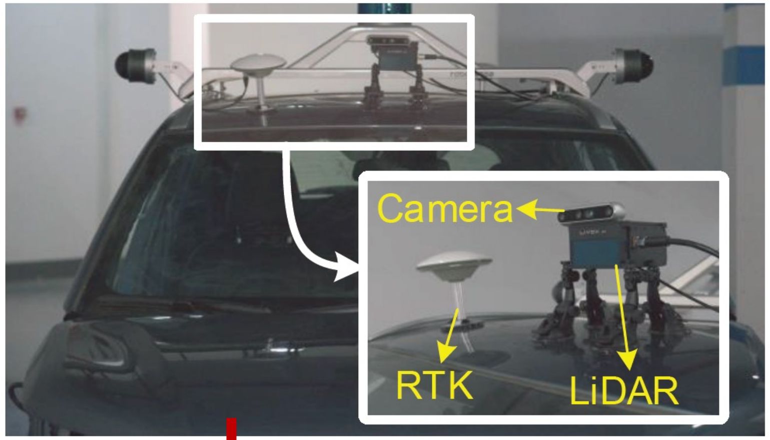
■ Test Scenarios

Scenario	Traj. length	Images	GPS	Infra. nodes
Open Road	6.6 km	45.1 k	~90%	13
Campus	1.3 km	15.4 k	~80%	8
Industrial park	5.9 km	40.2 k	~60%	9
Underpasses	0.4 km	3.6 k	~60%	4
Underground parking	3.4 km	33.0 k	0%	10



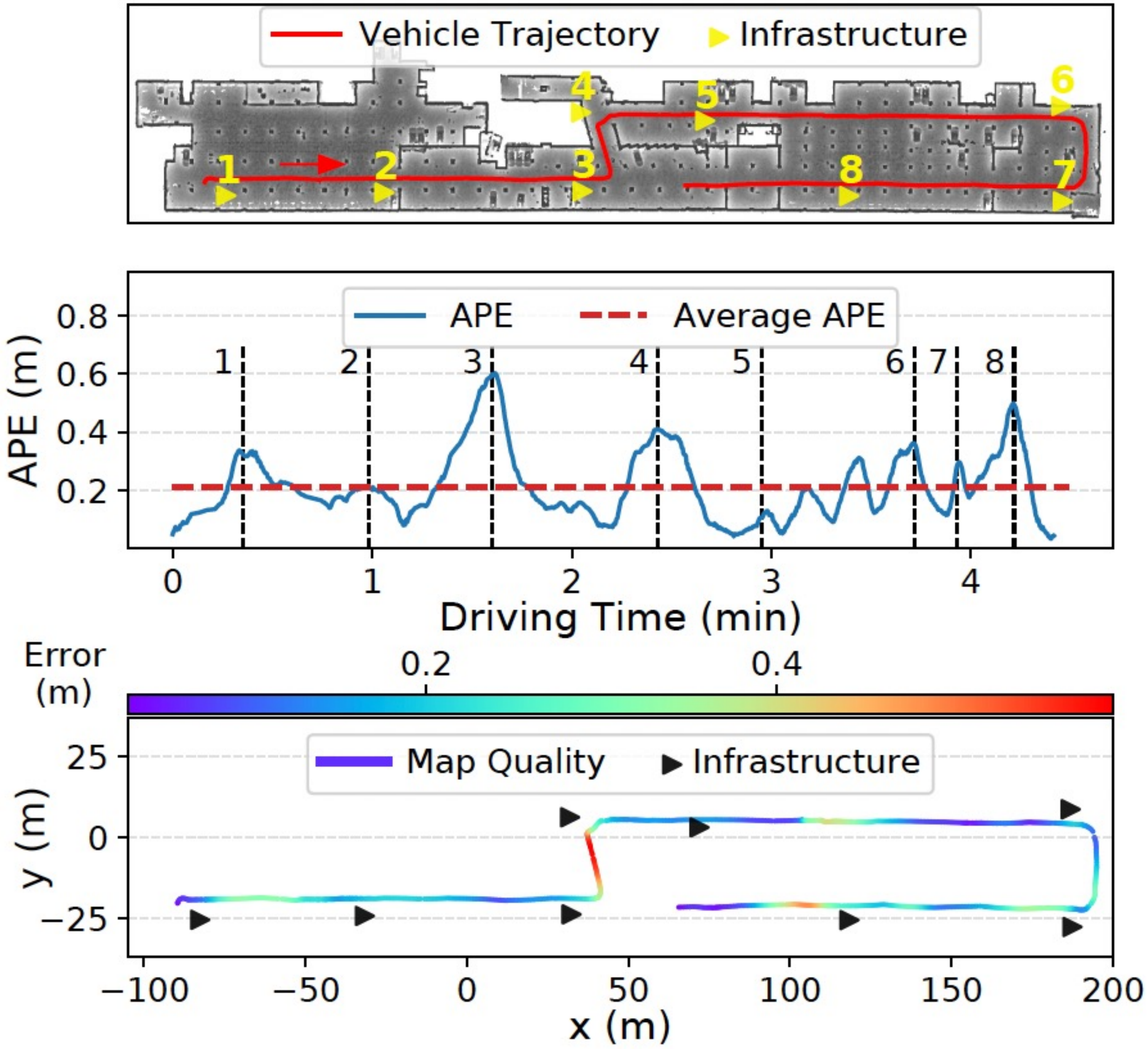
■ Roadside Infrastructure





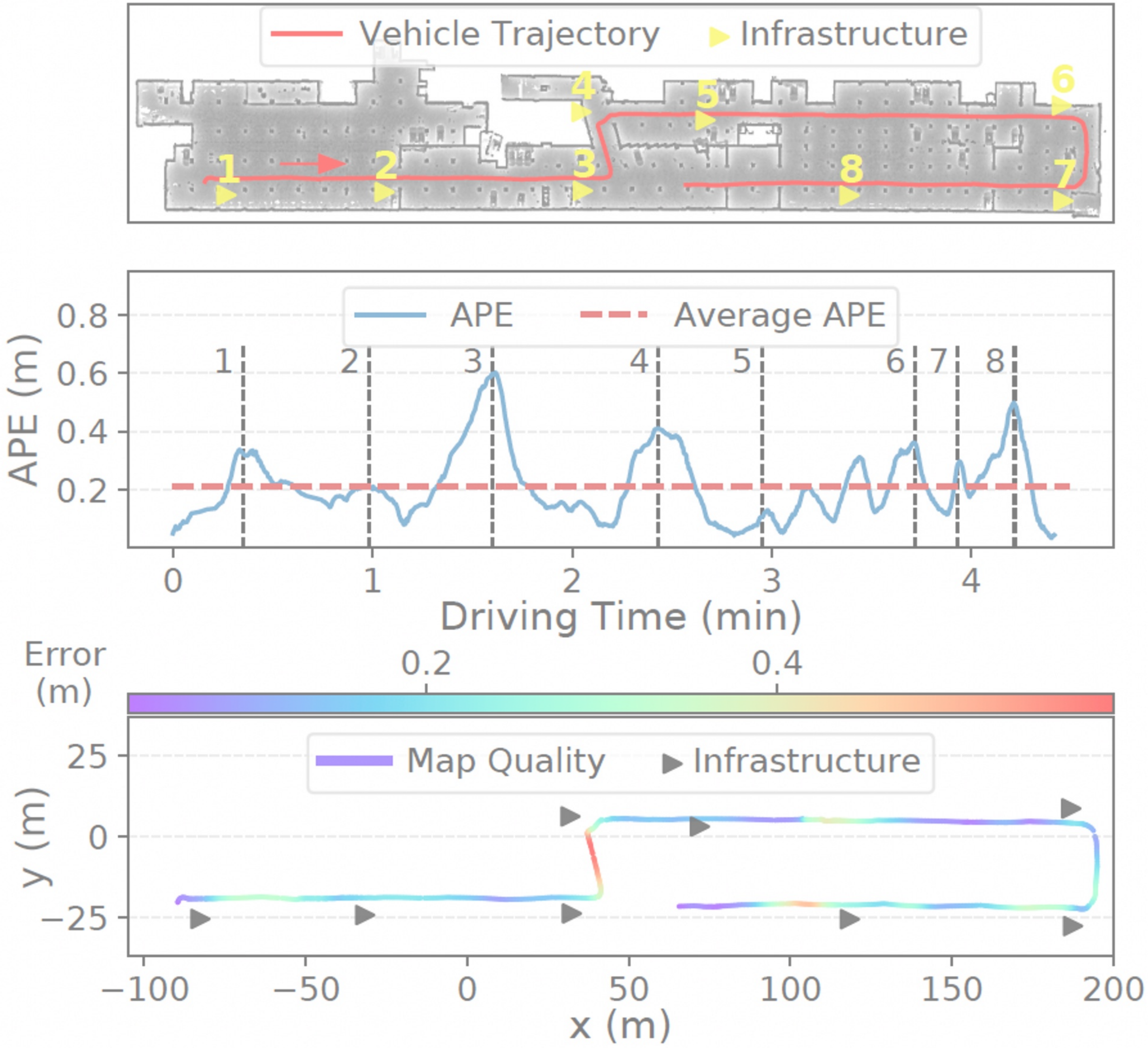
■ Test Vehicle

System Performance Evaluation

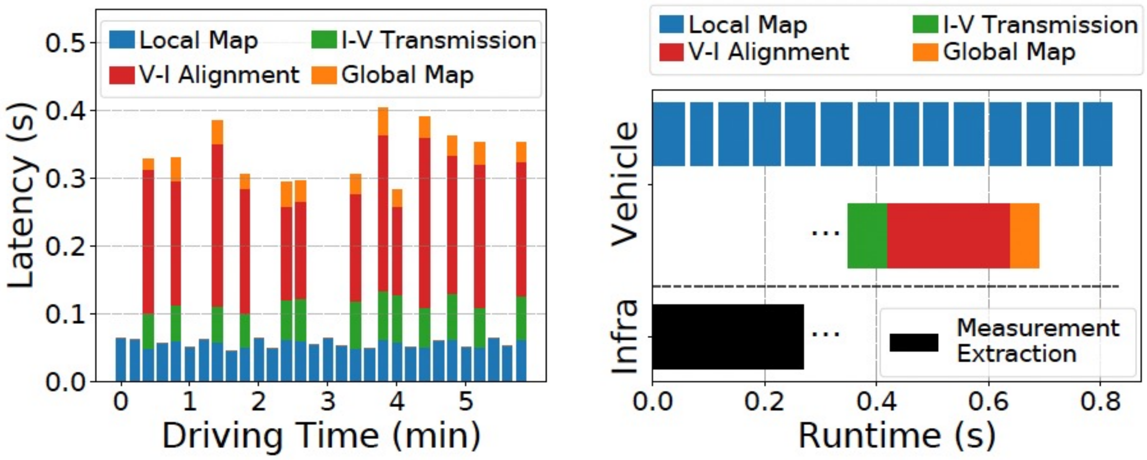


- End-to-end Performance

System Performance Evaluation

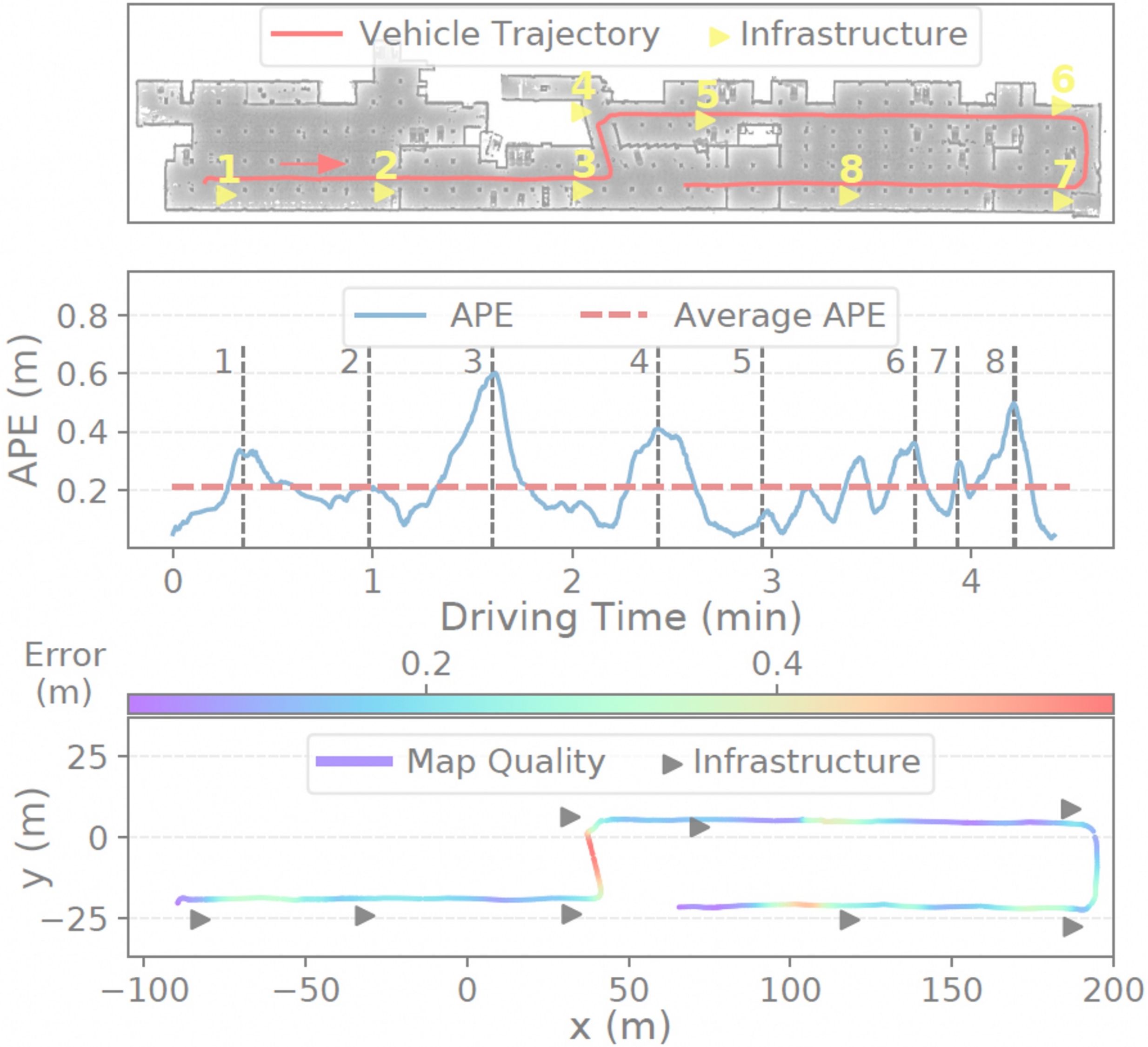


■ End-to-end Performance

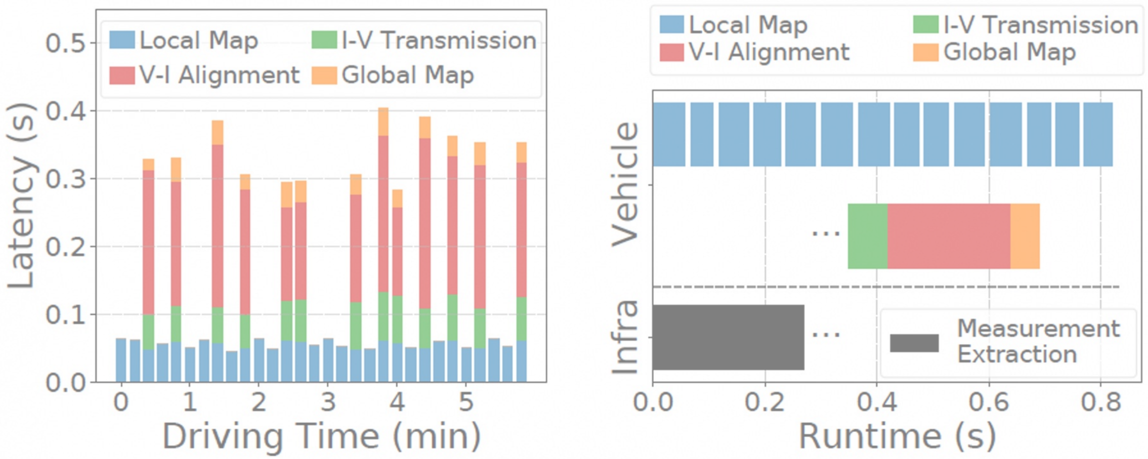


■ Latency of VILAM

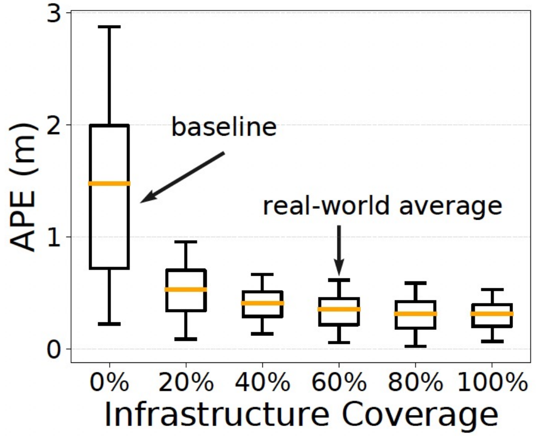
System Performance Evaluation



■ End-to-end Performance

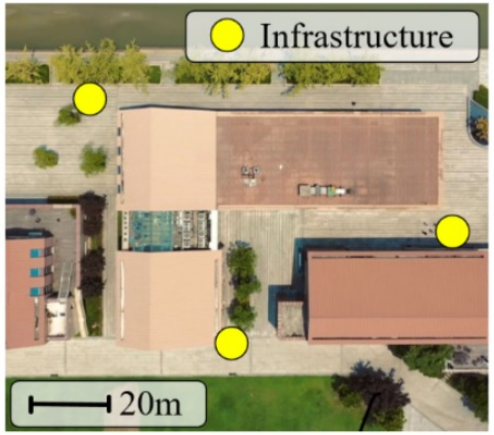


■ Latency of VILAM



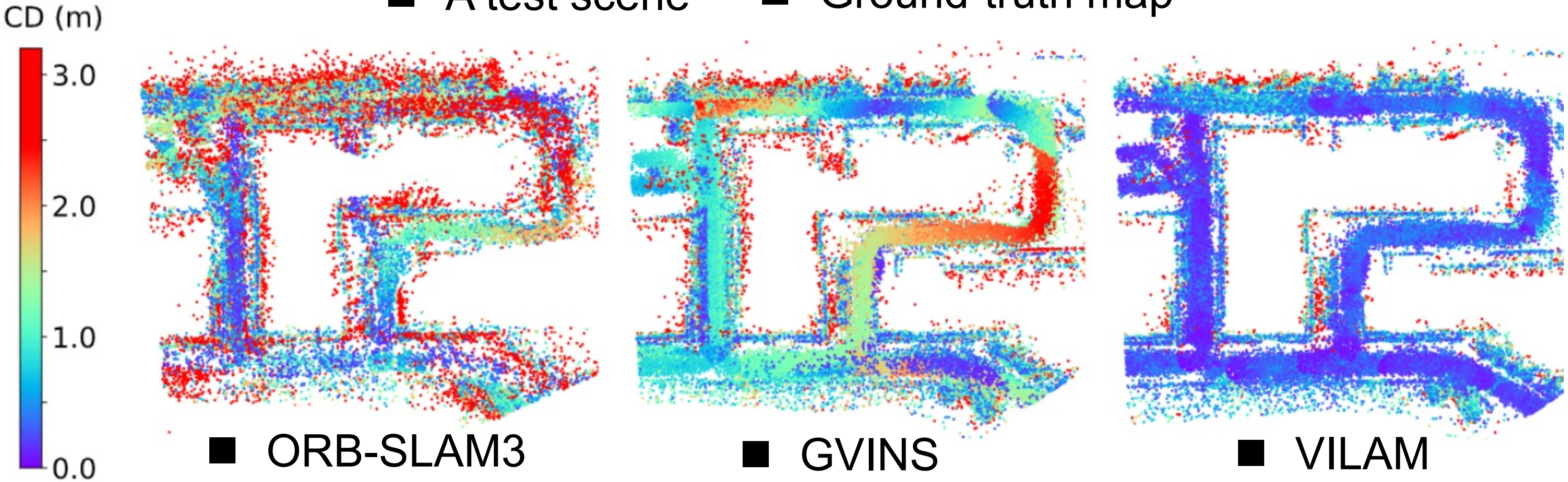
■ Performance of VILAM under different infrastructure density

System Performance Evaluation



■ A test scene

■ Ground-truth map



■ ORB-SLAM3

ATE: 3.58m

■ GVINS

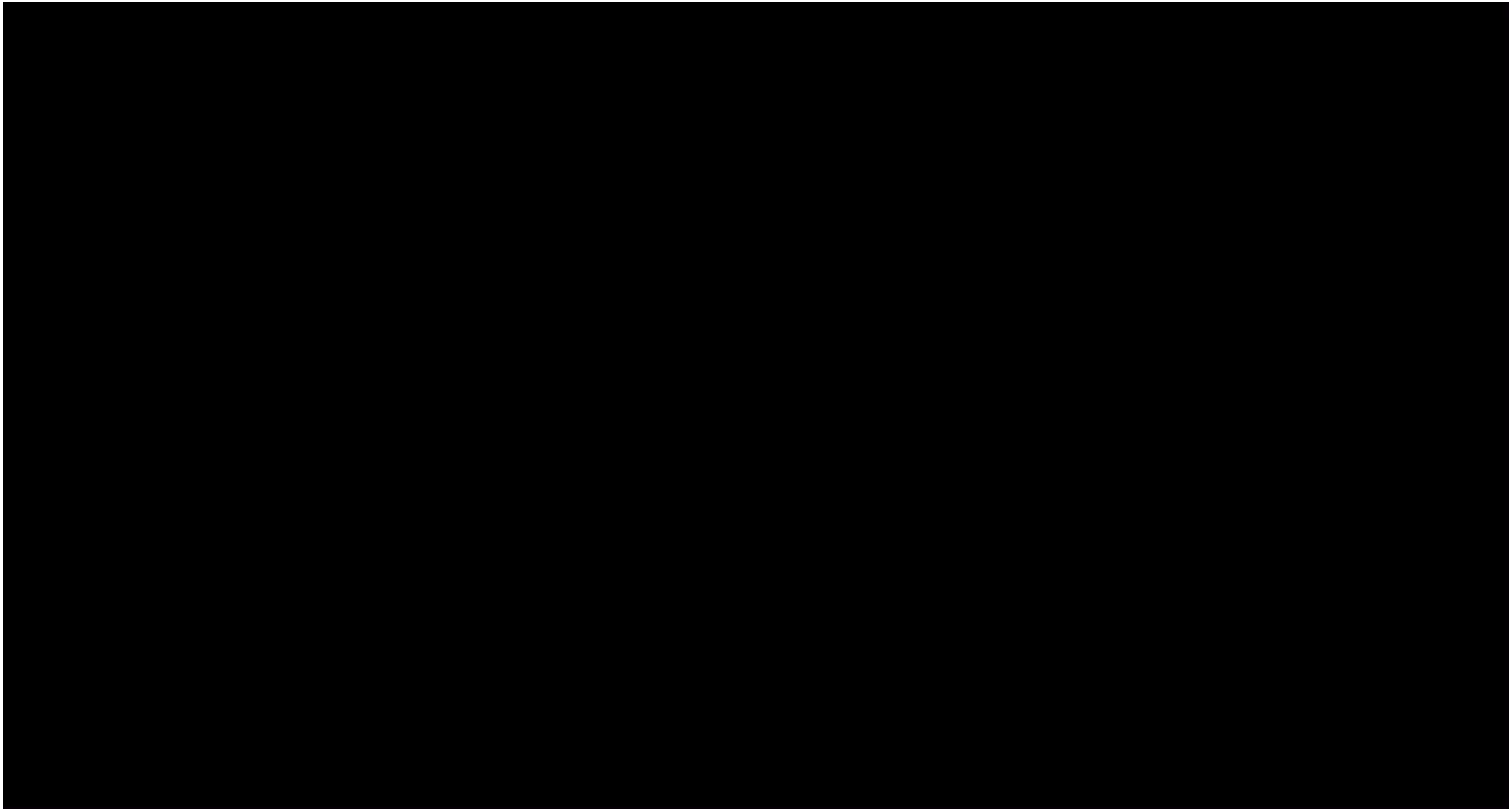
ATE: 3.19m

■ VILAM

ATE: 0.34m

VILAM can eliminate **85%** of the cumulative errors in state-of-the-art SLAM methods.

End-to-end Implementation Demo



Link to this video demo: https://youtu.be/lTlqDNipDVE

Conclusion

- ❖ Distributed infrastructure-assisted localization and mapping
  - Elastic registration to address the deformation of SLAM map
  - Lightweight factor graph-based approach to build a consistent map

- ❖ Extensive evaluation
  - Decimeter-level (~0.3m) localization
  - Globally consistent (~0.7m accuracy) map

Thanks for Listening!

- *Check our paper:* VILAM: Infrastructure-assisted 3D Visual Localization and Mapping for Autonomous Driving

- *Authors:* Jiahe Cui, Shuyao Shi, Yuze He, Jianwei Niu*, Guoliang Xing*, and Zhenchao Ouyang

- *Website:* http://aiot.ie.cuhk.edu.hk
  https://github.com/HViktorTsoi