Exploiting Sparse Semantic HD Maps for Self-Driving Vehicle Localization

Wei-Chiu Ma*, Ignacio Tartavull*, Ioan Andrei Bârsan*, Shenlong Wang*
Min Bai, Gellért Máttyus, Namdar Homayounfar, Shrinidhi Kowshika Lakshmikanth, Andrei Pokrovsky, Raquel Urtasun

* Denotes Equal Contribution

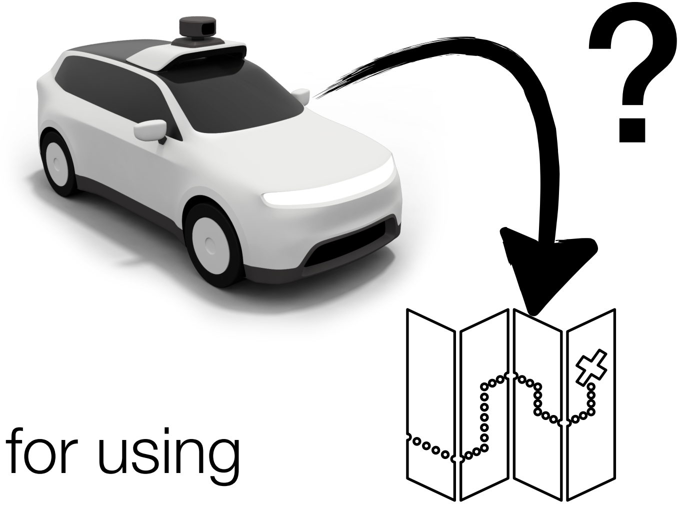
November 6, 2019

Uber

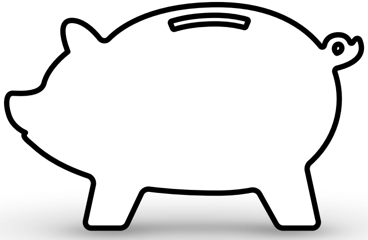
iROS MACAU 2019

Problem & Motivation

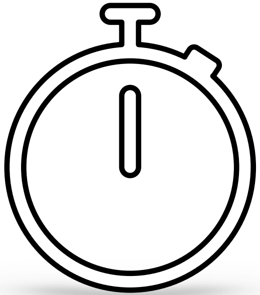
- Self-driving vehicles are complex robotic systems
- **Maps** can improve safety and performance of perception, motion forecasting and planning
- Precise **ego-localization** is required for using maps



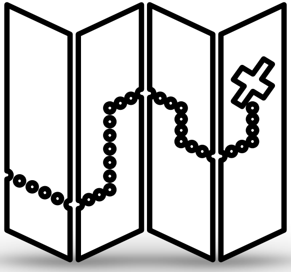
Localization Desiderata



Low **Cost** for
Map Building &
Storage

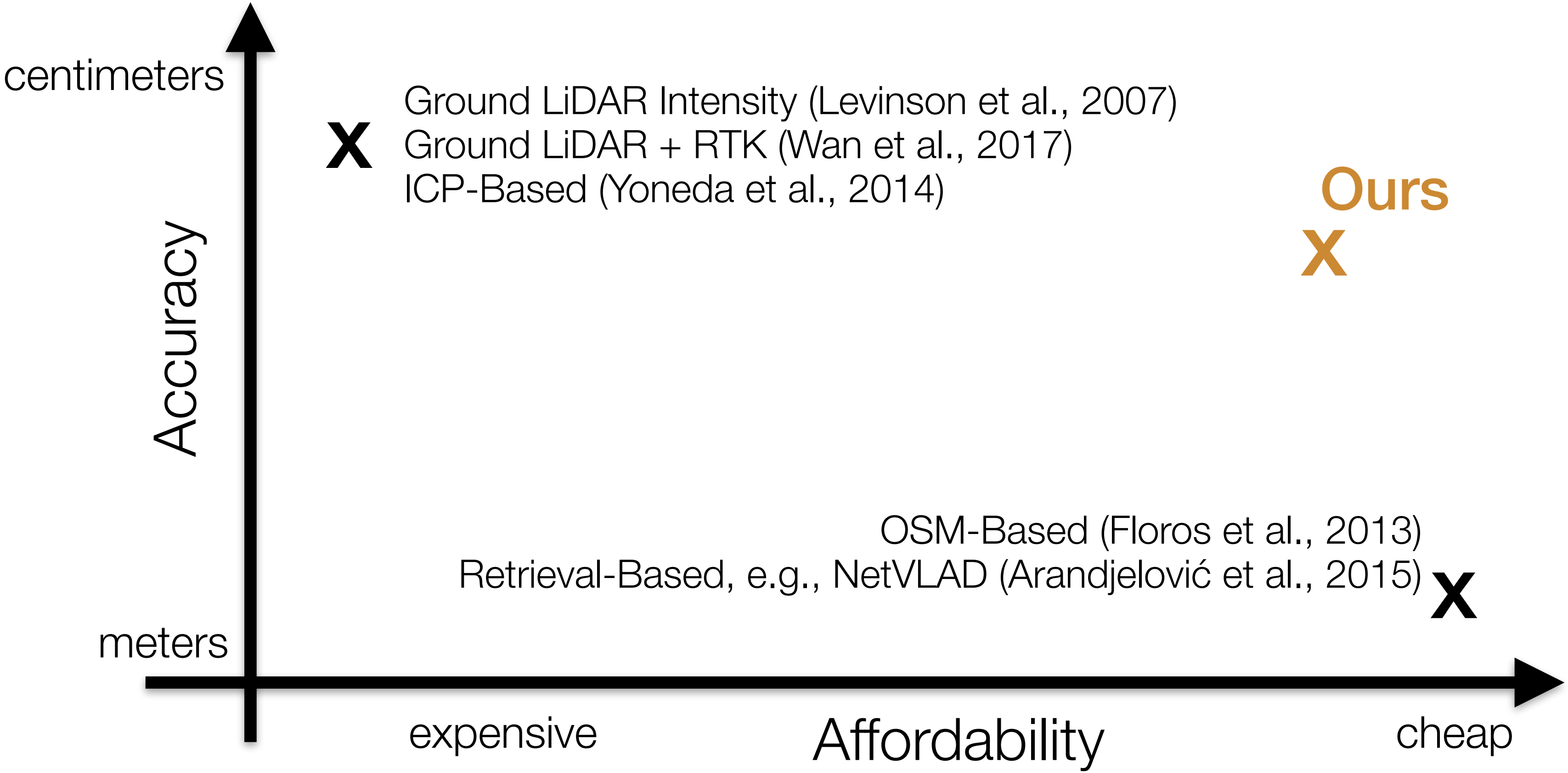


Real-Time
Inference



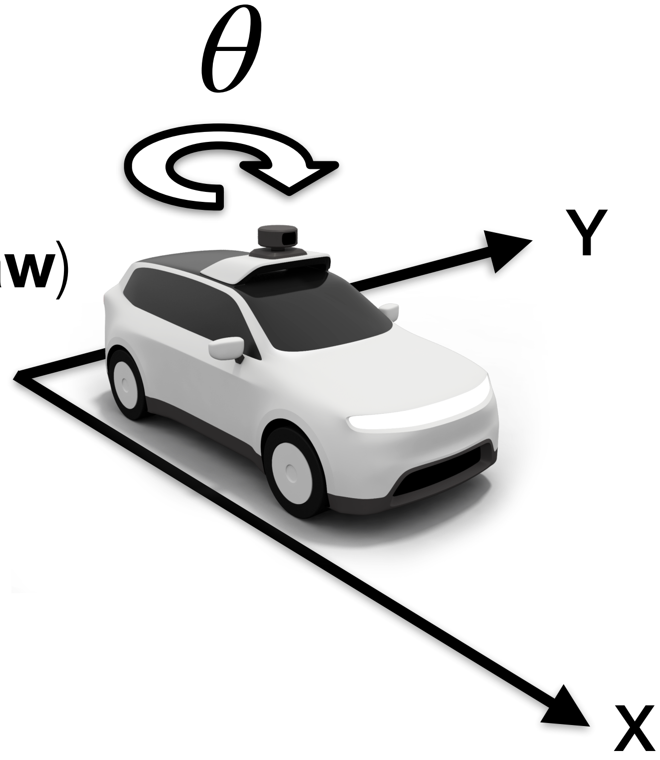
High **Accuracy**
(Centimeter-level)

Related Work



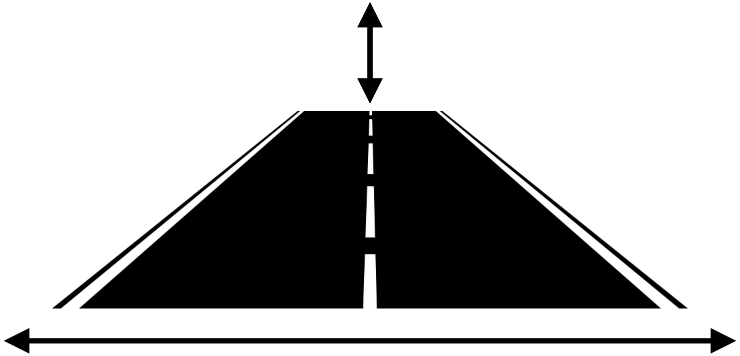
Problem Statement

- Online localization w.r.t. map
- Sub-meter accuracy
- Vehicle on ground: state = (**x, y, yaw**)



Proposed Method

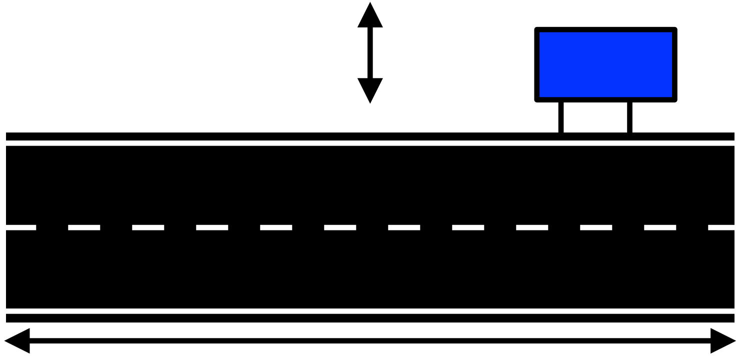
Perceived Lanes



Lateral Information

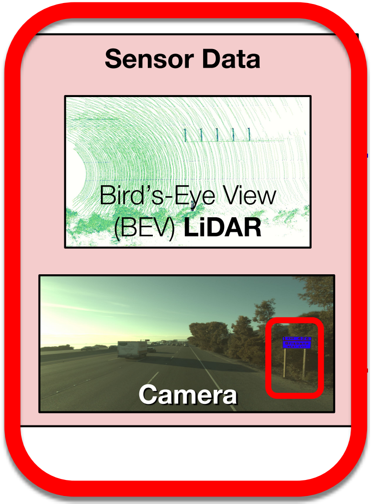
+

Perceived Signs



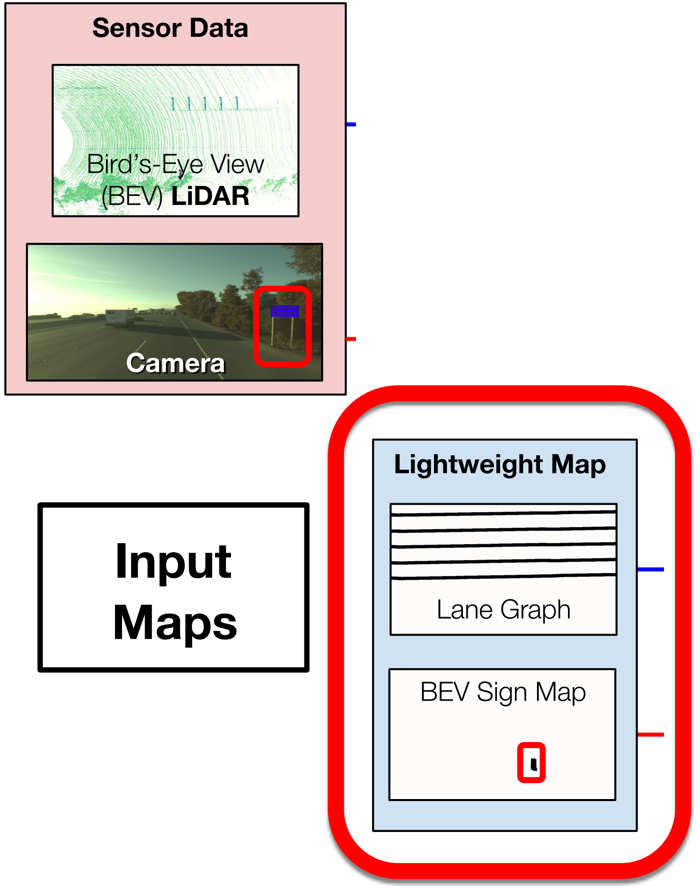
Longitudinal Information

Method Overview

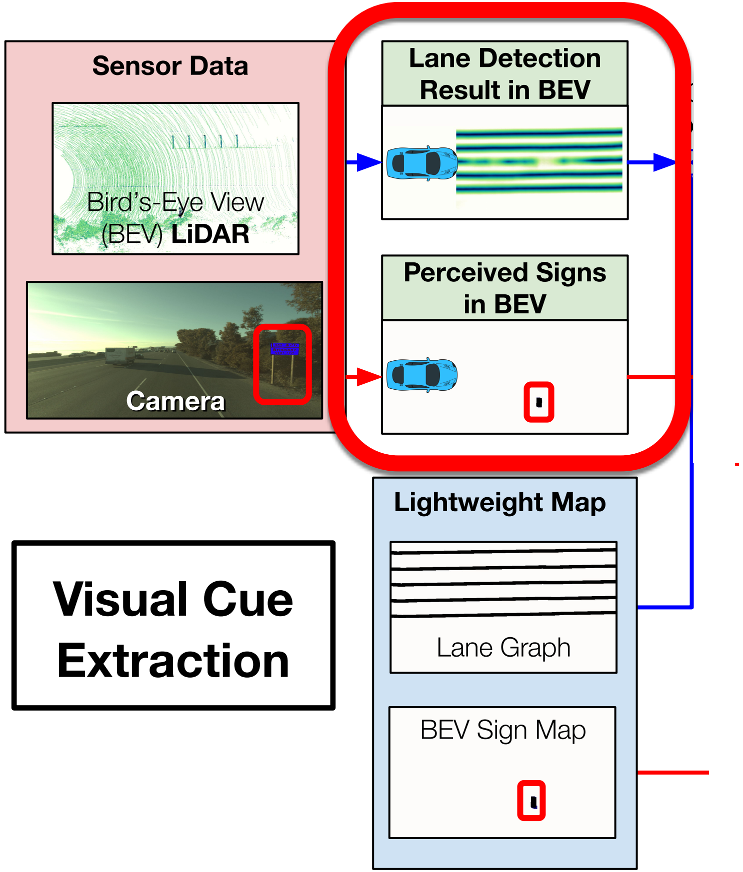


Input Sensors

Method Overview

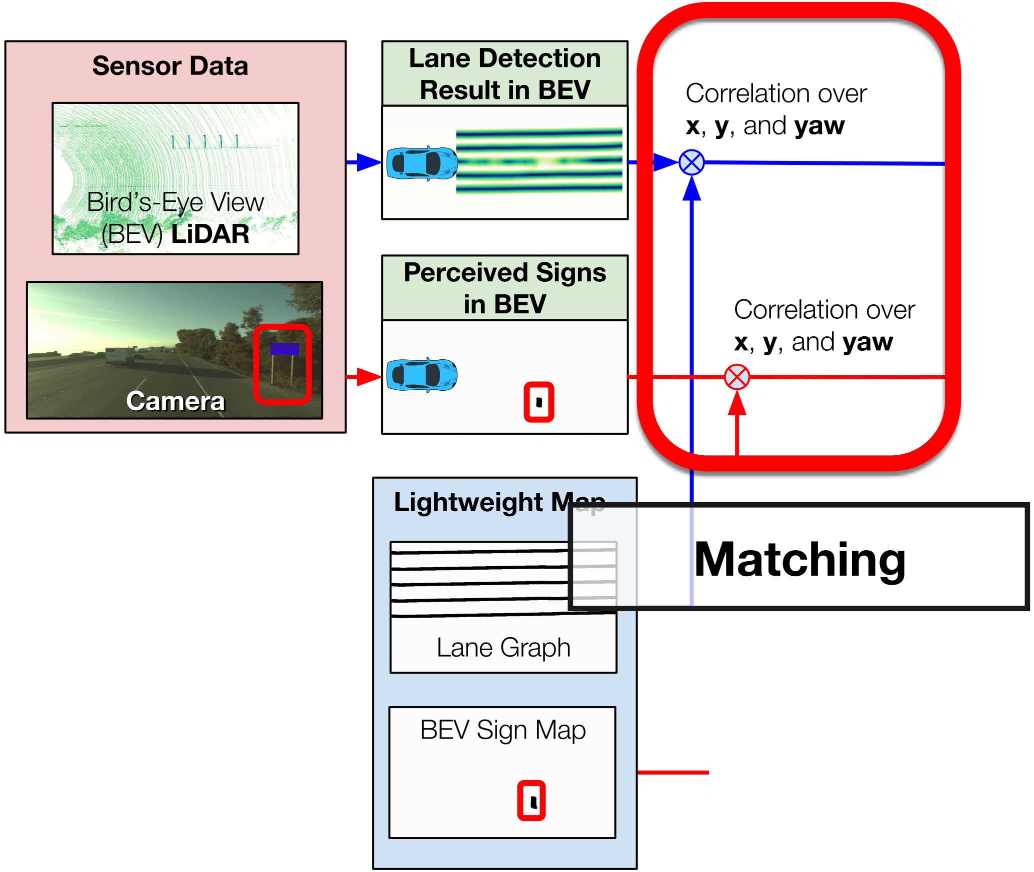


Method Overview

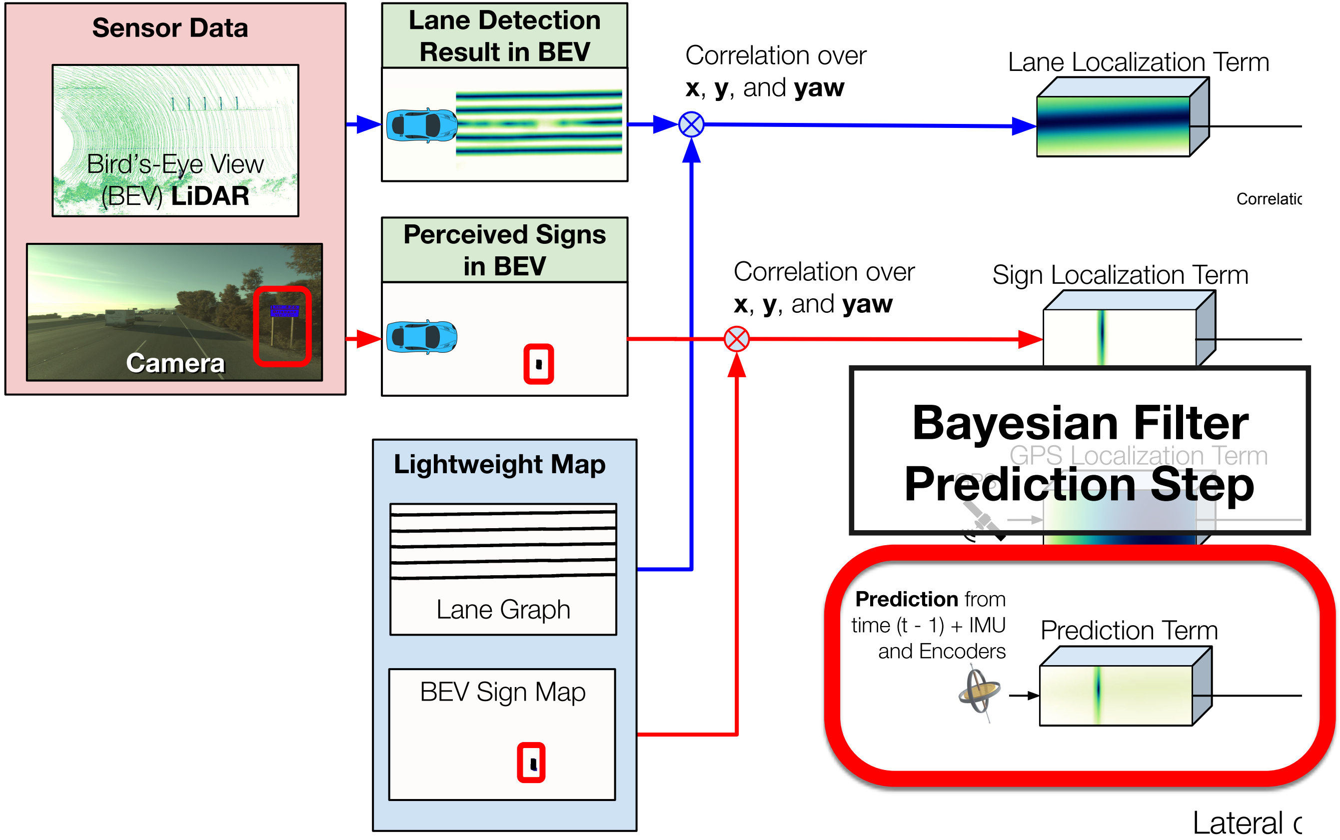


Visual Cue Extraction

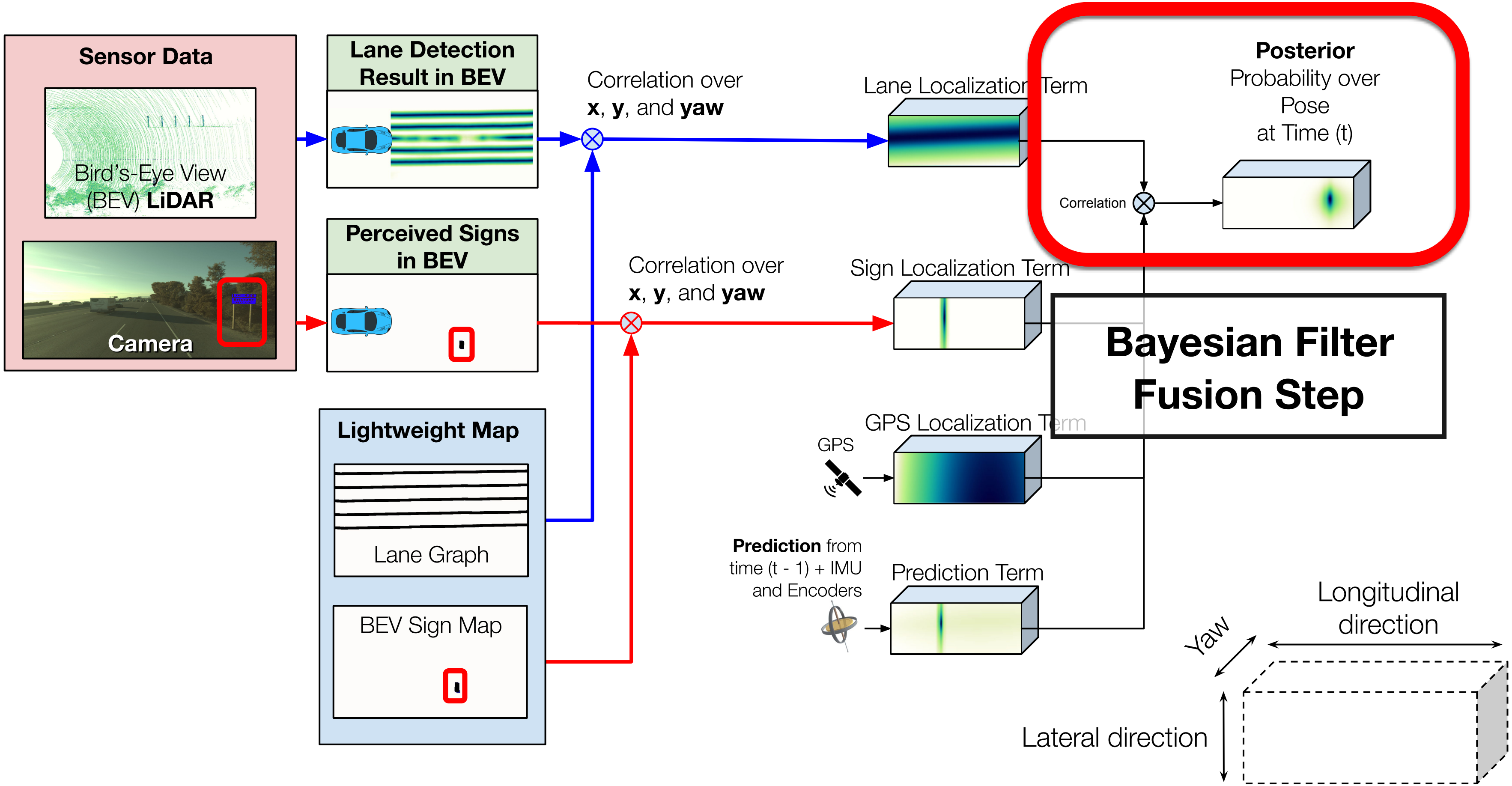
Method Overview



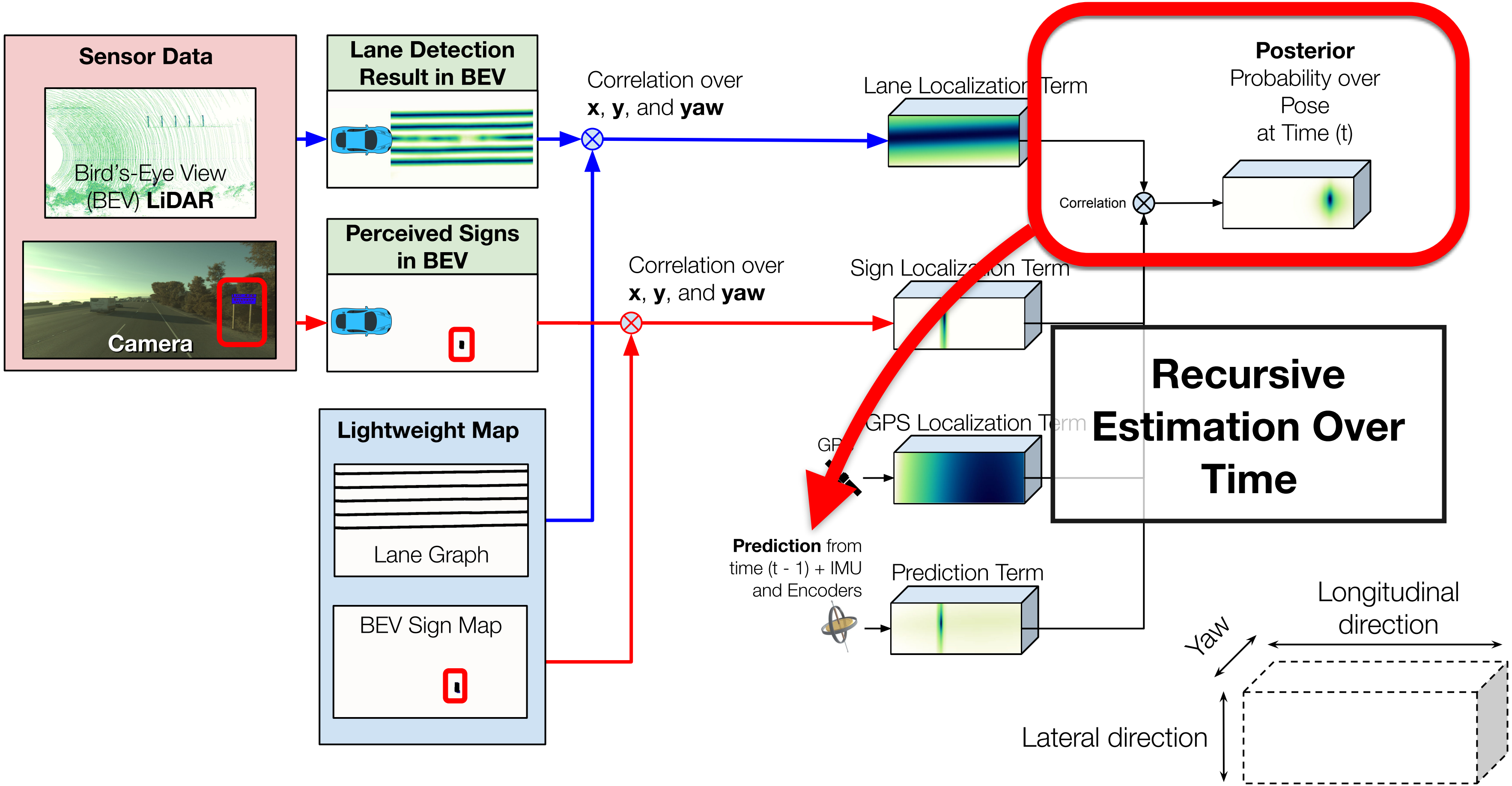
Method Overview



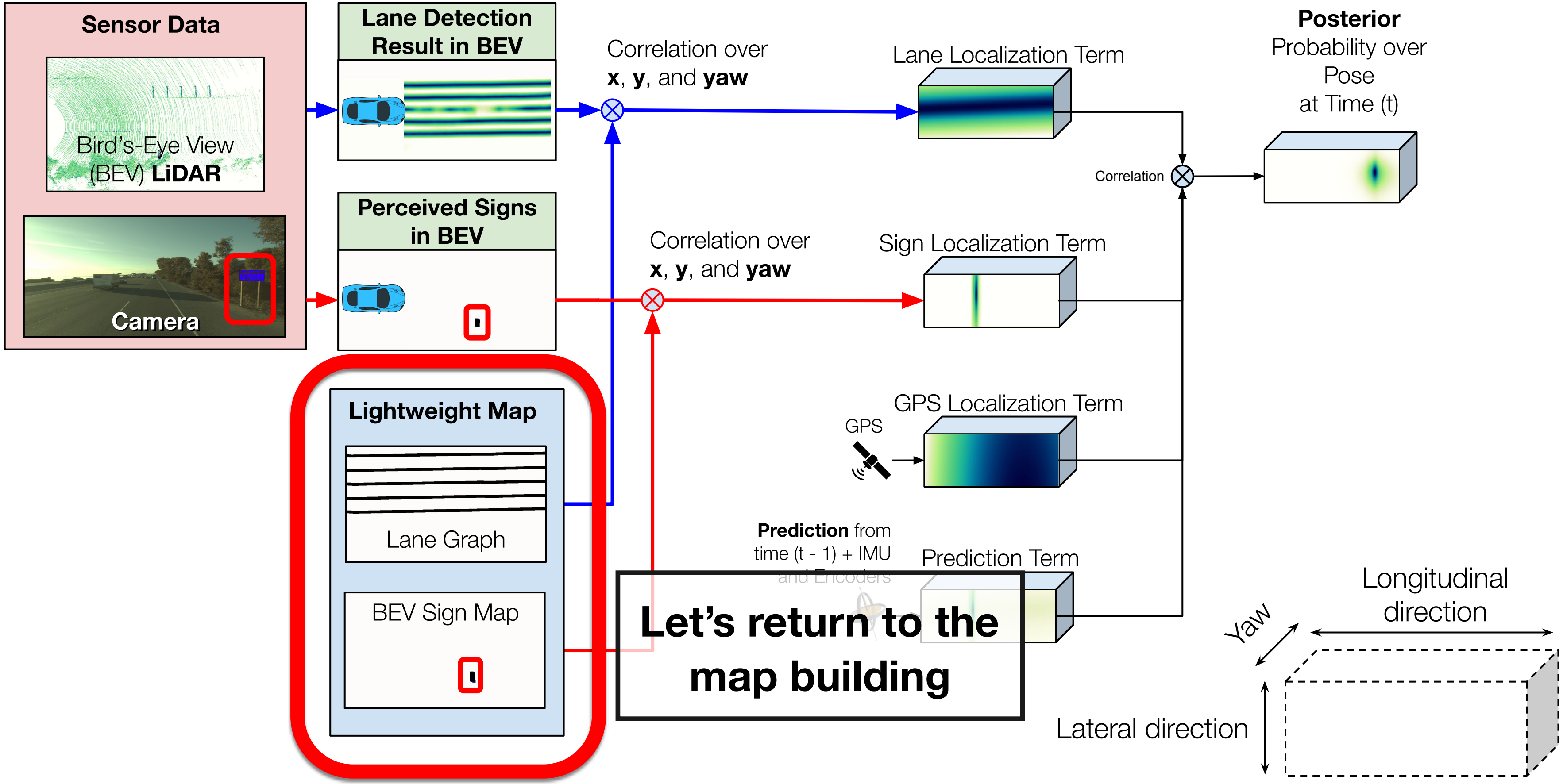
Method Overview



Method Overview

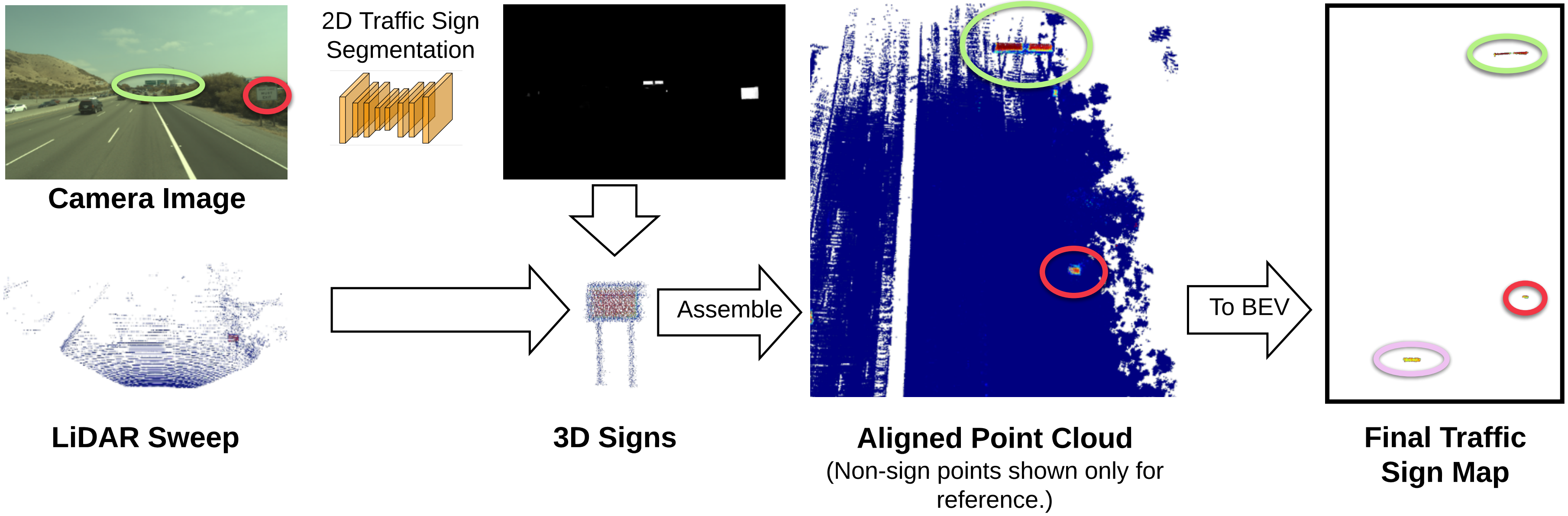


Method Overview

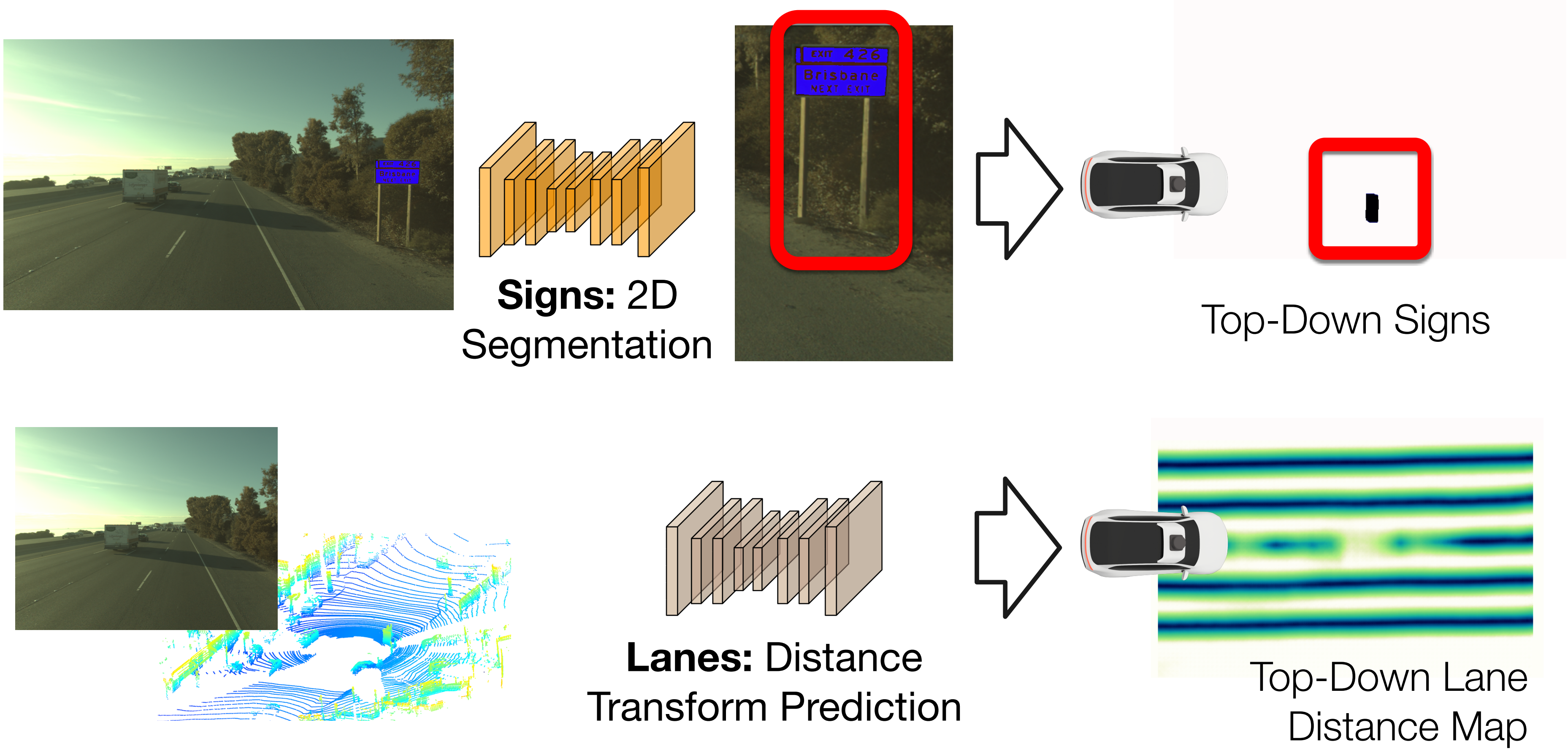


Let's return to the map building

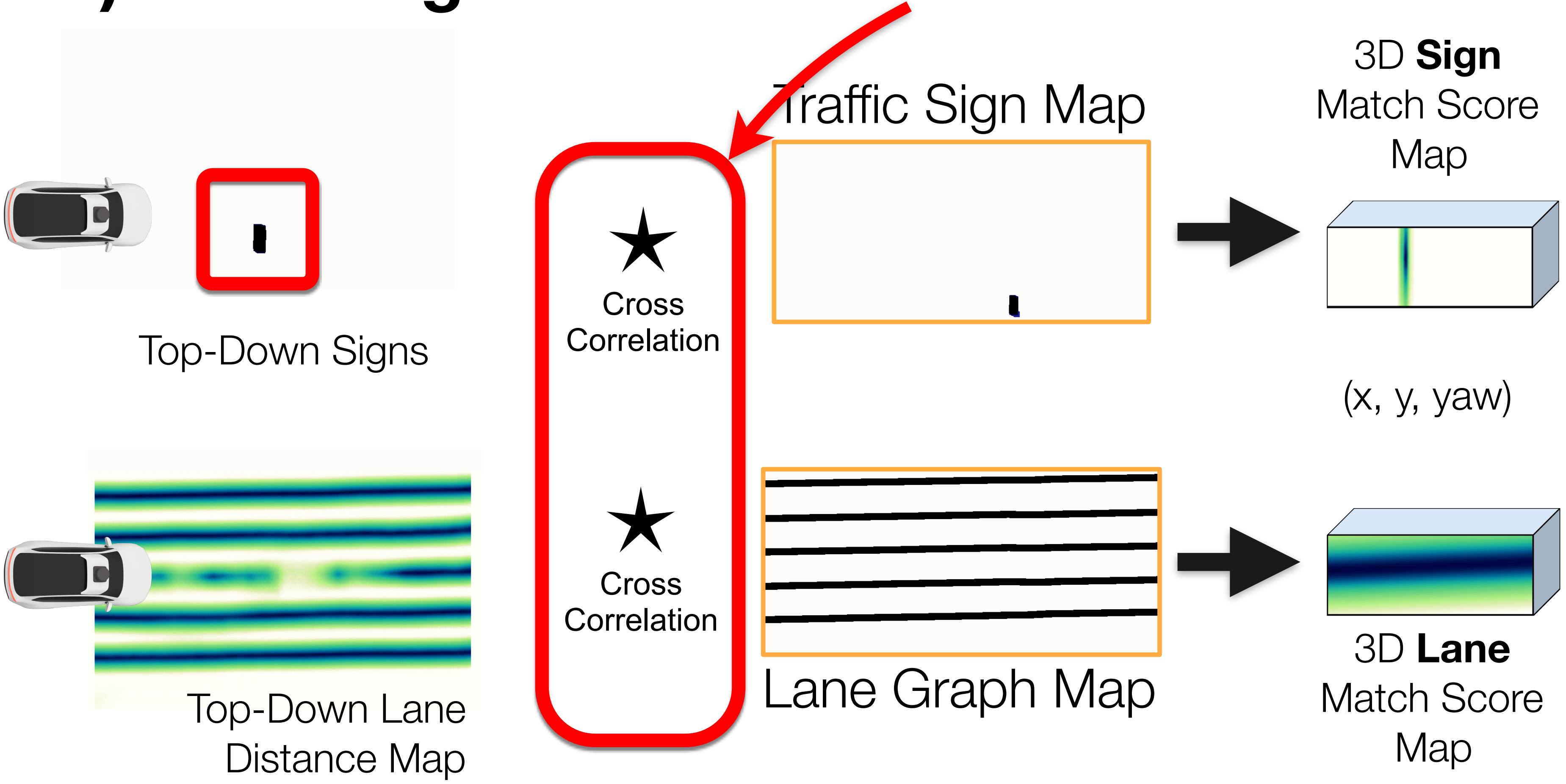
Offline: Sign Map Building Process



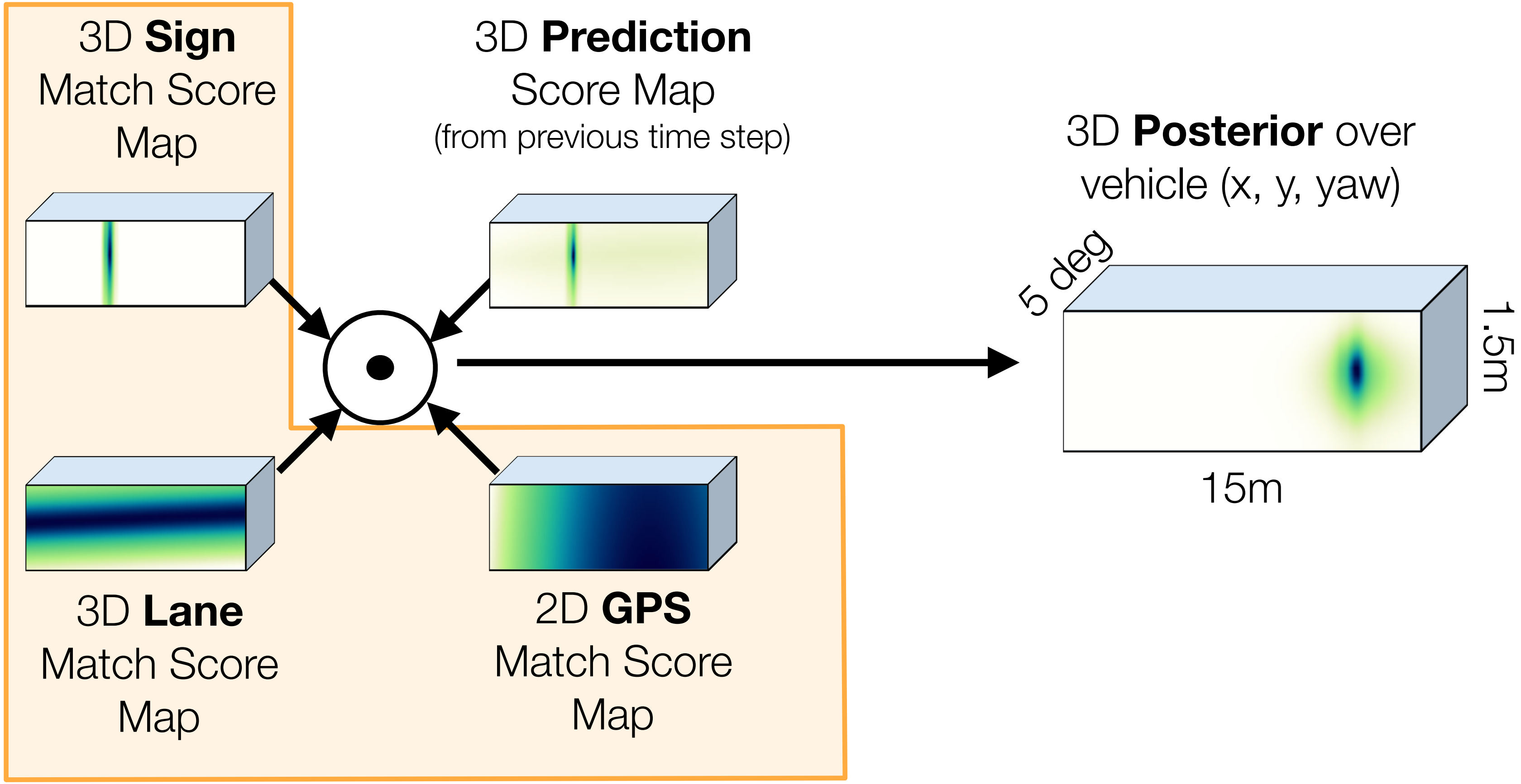
1) Visual Cue Extraction



2) Matching


Fast in Fourier Domain
Top-Down Signs
Cross Correlation
Cross Correlation
Traffic Sign Map
3D **Sign** Match Score Map
(x, y, yaw)
Lane Graph Map
3D **Lane** Match Score Map
Top-Down Lane Distance Map

3) Pose Filter



Dataset

- 312km of driving on multiple US highways
- Challenges:
  - High speed
  - Repetitive structures

Metrics

- Localization accuracy
  - Euclidean distance w.r.t. true pose
- Worst-case behavior critical
  - Report 50th, 95th and 99th percentiles

Experimental Results: Performance

Methods	Longitudinal Error (m)			Lateral Error (m)		
	Median	95%	99%	Median	95%	99%
Dynamics	24.85	128.21	310.50	114.46	779.33	784.22
GPS	1.16	5.78	6.76	1.25	8.56	9.44
GPS + Dynamics	1.59	6.89	13.62	2.34	11.02	42.34
Ours	**1.12**	**3.55**	**5.92**	**0.05**	**0.18**	**0.23**

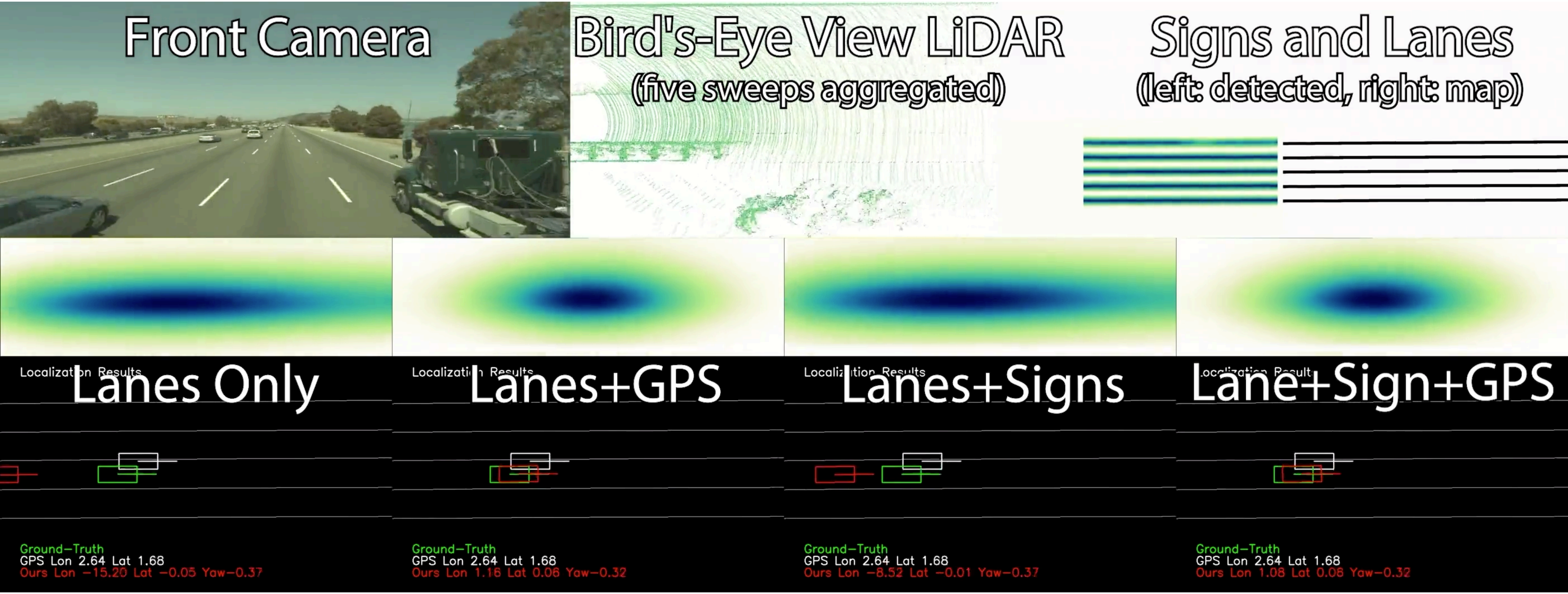
Experimental Results: Ablation Study

			Longitudinal Error (m)			Lateral Error (m)		
Lane	GPS	Sign	Median	95%	99%	Median	95%	99%
✓			13.45	37.86	51.59	0.20	1.08	1.59
✓		✓	6.23	31.98	51.70	0.10	0.85	1.41
✓	✓		1.53	5.95	6.27	0.06	0.24	0.43
✓	✓	✓	**1.12**	**3.55**	**5.92**	**0.05**	**0.18**	**0.23**

Experimental Results: Storage

Map Type	Storage (MB/km^2)	Approximate USA Road Network Storage (TB)
Full point clouds	1,447.00	1,138.47
Ground intensity	177.00	139.26
Ours (Signs + Lane Graph)	0.55	0.43

Qualitative Results



Discussion and Future Work

- Complementary semantic cues can enable accurate map-based localization on highways using a fraction of the storage required for traditional HD maps
- Reliable localization in the correct lane on >300km
- 3—4 orders of magnitude less storage than appearance based maps
- Future work:
    - Integrate with compressed appearance maps
    - Re-localization module

Thank you!

FAQ

- Unpainted roads?
    - Road boundaries are still a strong cue!
- Lack of road signs & off-road?
    - Can be mitigated with (compressed*****) appearance maps
- Longitudinal error?
    - Safety is much more related to lateral accuracy in the highway scenarios we evaluated.

*****) Wei et al., Learning to Localize through Compressed Binary Maps, CVPR ‘19

FAQ

- What if the maps are out of date?
  - Change detection + mapless driving.
  - No over-reliance on any one sensor or the maps.
- If you want sparse maps, why no visual SLAM/ORB-SLAM, etc.
  - Accuracy still not high enough in the lateral dimension.
- Why not LOAM?
  - We are planning to investigate more advanced LiDAR SLAM methods.