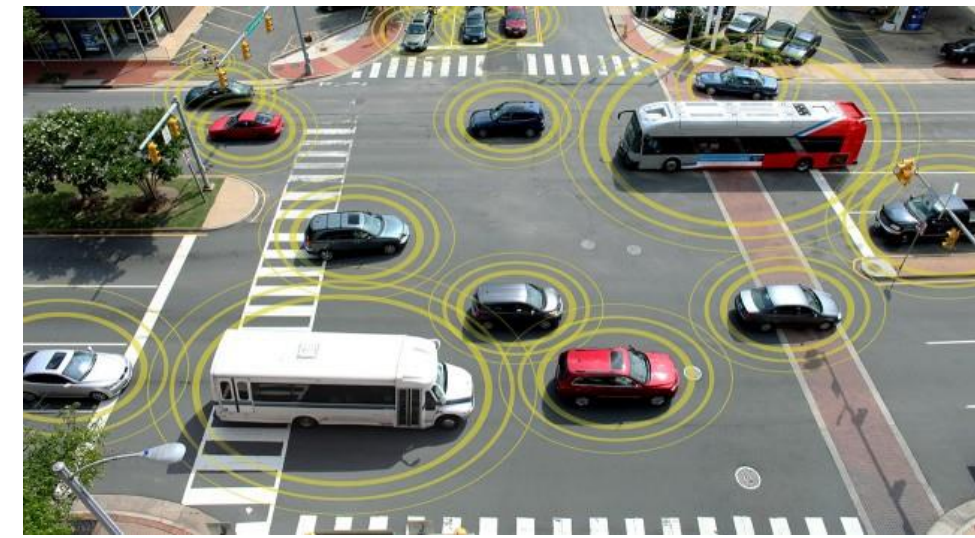


## Motivation & Goal



City of the Future

**Goal:** Fast and accurate 4D sensing of vehicles from multiple cameras.

**Application:**

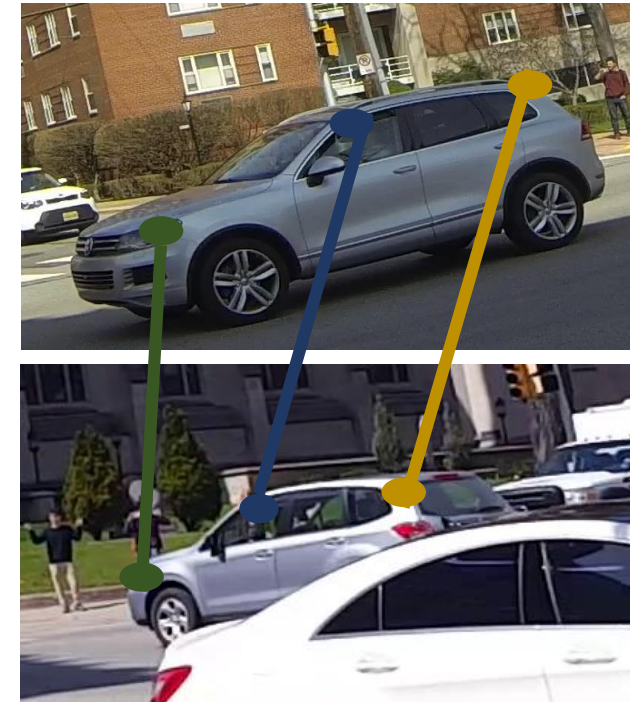
1. Traffic behavior understanding
2. Autonomous driving
3. Pollution analysis

## Challenge



Noisy and incomplete keypoint detection (*structured point*) prohibits accurate multiview car association and reconstruction by ray triangulation.

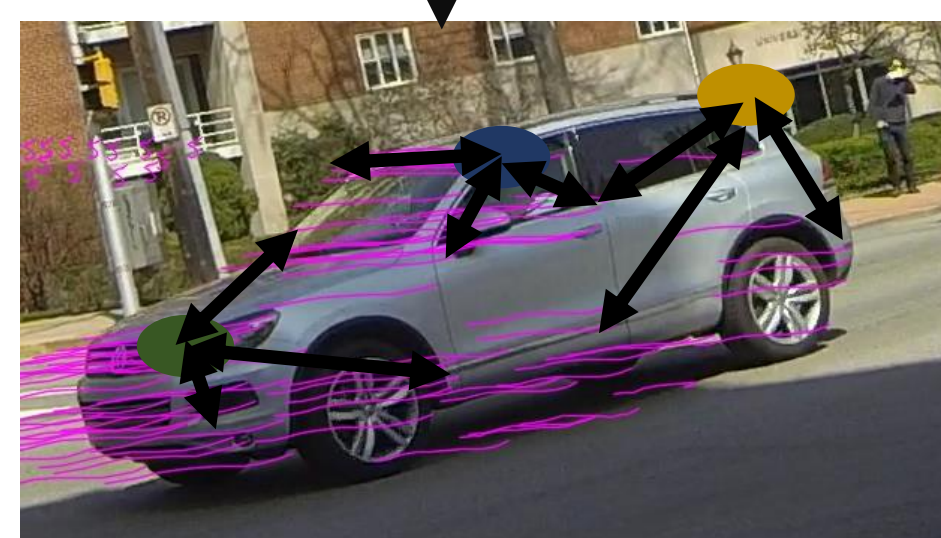
## Insight



**Structured point:** accurate matching but imprecise tracking

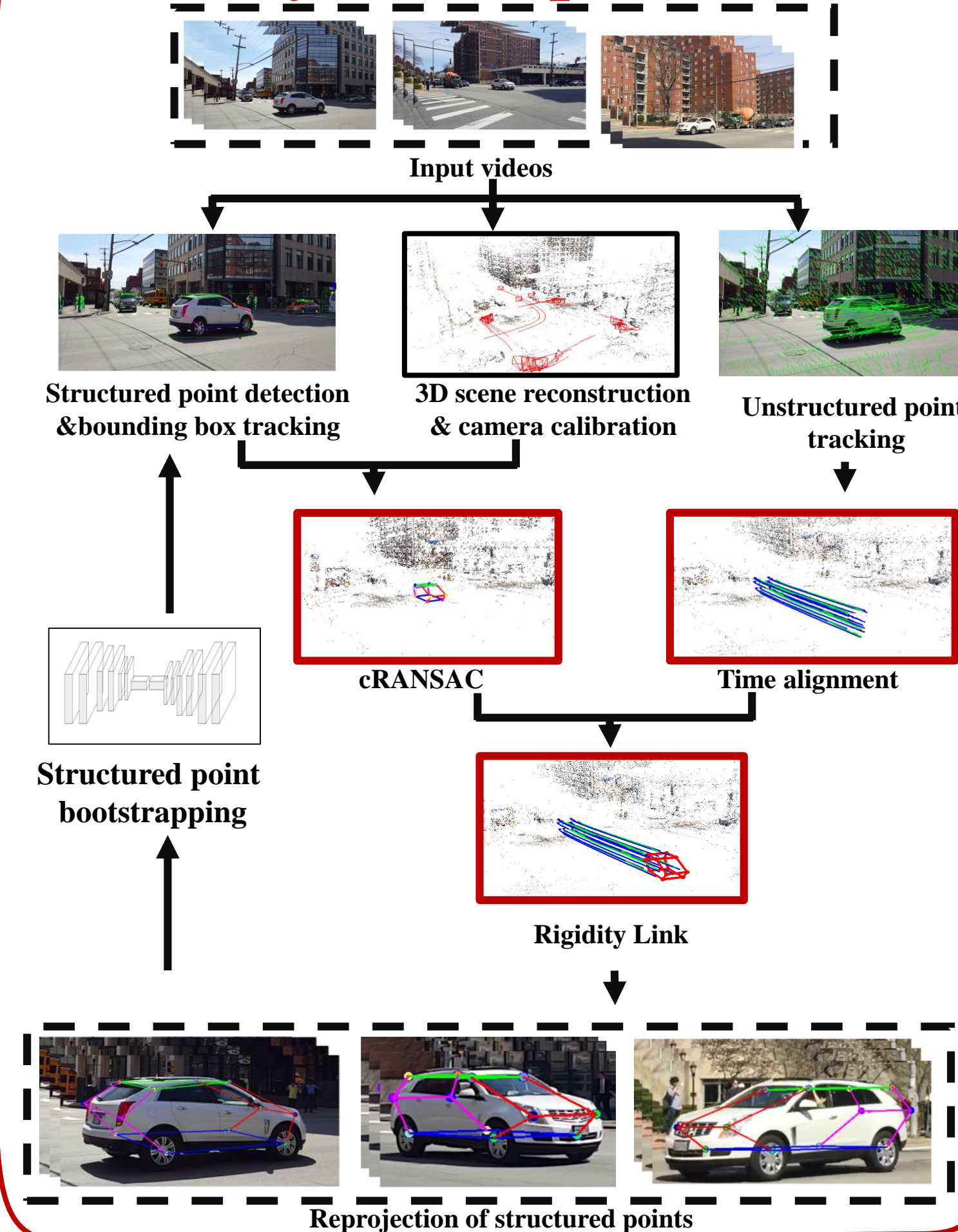


**Unstructured point:** precise tracking but inaccurate matching



1. The distances between the structured points and unstructured points are constant over time for *rigid deformation*
2. *No cross-view matching* of the unstructured points are needed

## System Pipeline



## Formulation

**cRANSAC:** Whole car multiview association

$$E_c = \text{Image reprojection cost} + \text{Symmetry cost} + \text{Length prior}$$

**Time alignment:** Motion coherency

$$E_S = \sum_o \sum_t (SE_o(t) - SE_o(t-1))^2$$

The car motion is computed from the unstructured points

**Rigidity Link:** Length preservation

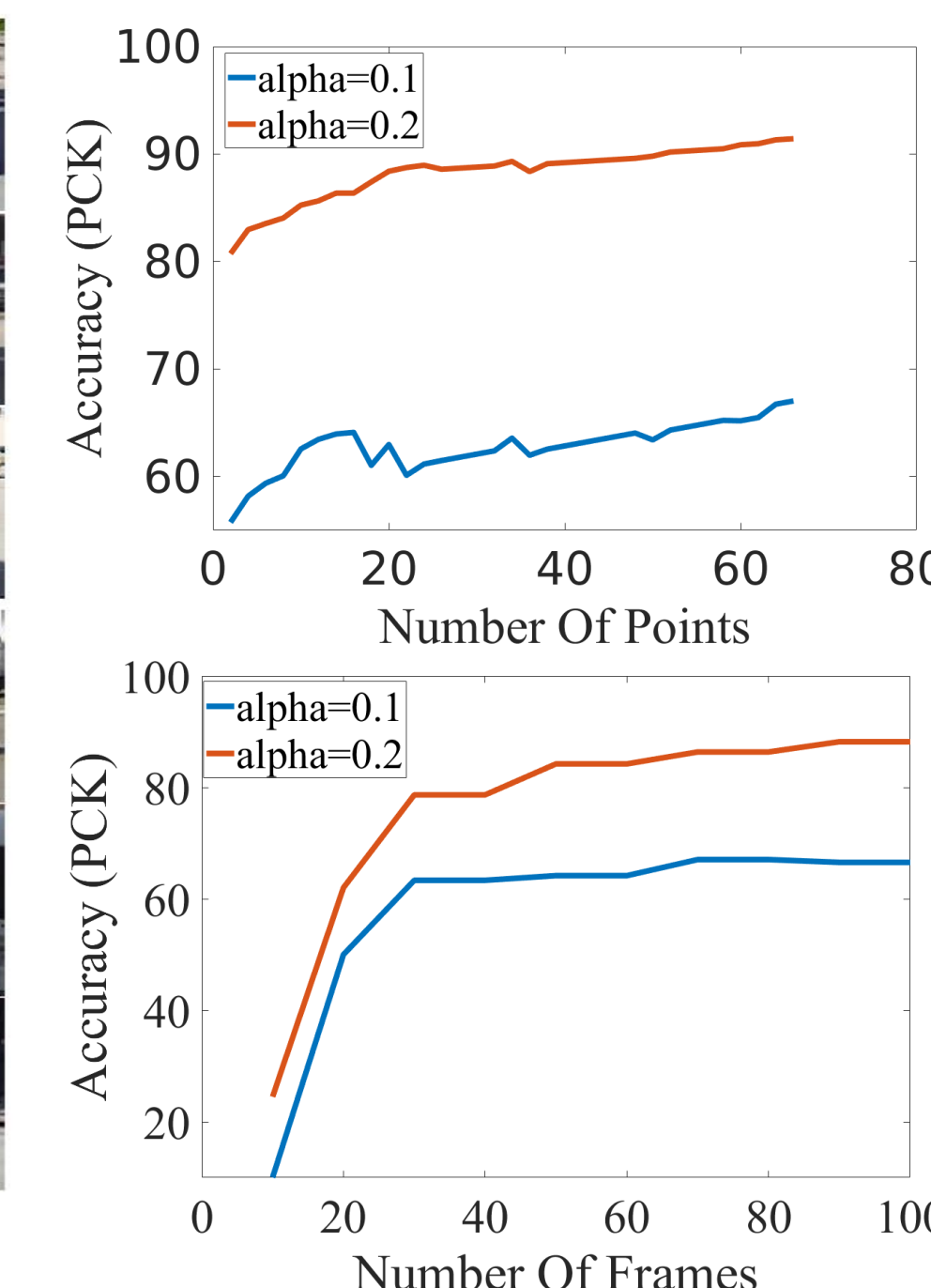
$$\min_s \sum_o \sum_f \sum_i \sum_j (\|R_o(f)S_i(f) + T_o(f) - U_j^c(f)\| - \|S_i(0) - U_j^c(0)\|)^2$$

Distance after deformation      Distance at the beginning

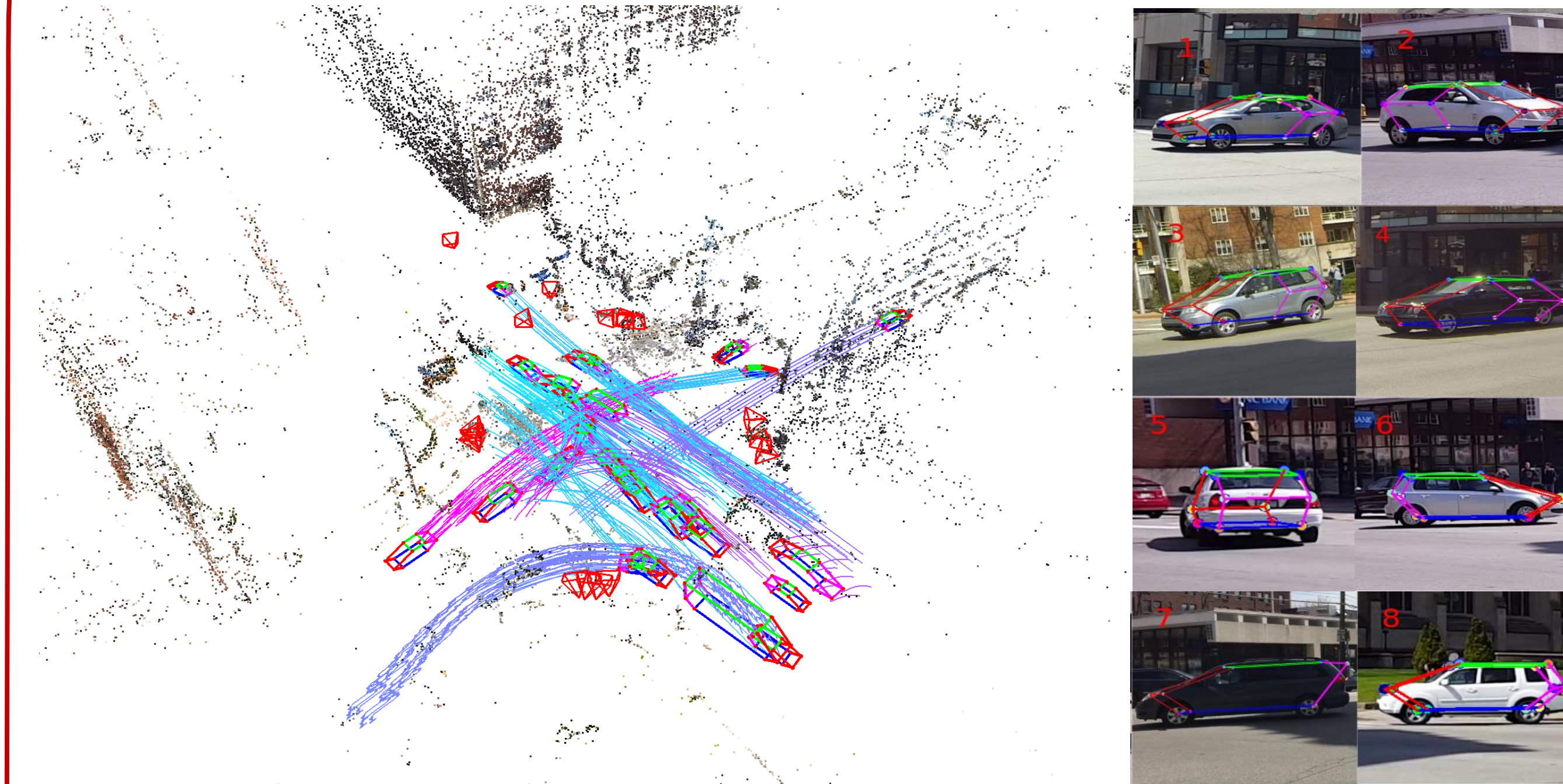
## Evaluation of 2D Structured Points



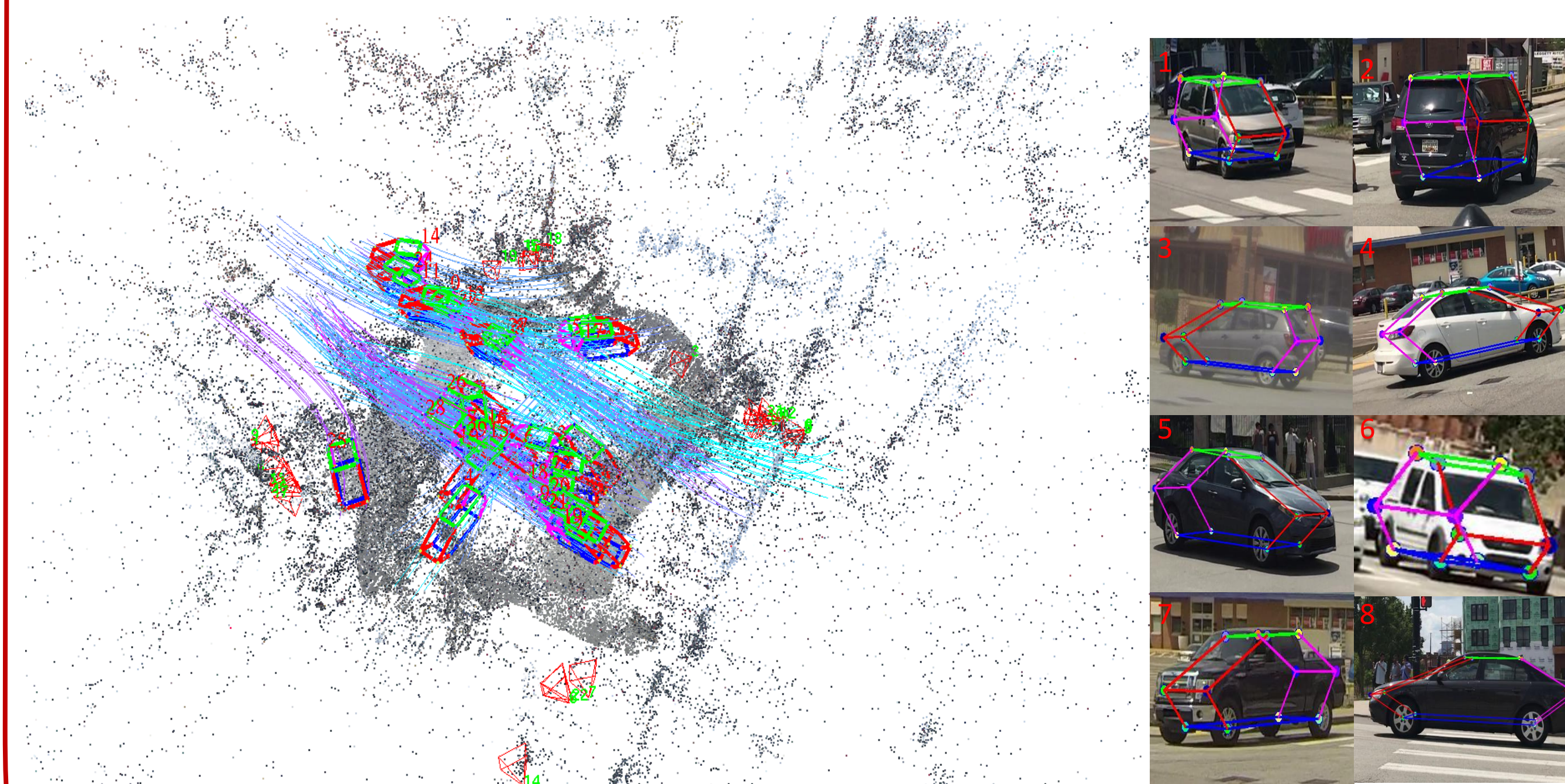
Input      Ground truth      Pretrained      Bootstrapped      Reprojection



## 4D Reconstruction



Fifth Ave. & Craig St. : reconstruction of 32/45 cars from 9 cameras @60fps for 3 minutes



Butler St. & 40<sup>th</sup> St.: reconstruction of 29/33 cars from 12 cameras @60fps for 3 minutes

## Acknowledgement

This work is supported by a Heinz Foundation grant, an NSF award CNS-1446601, an ONR grant N00014-15-1-2358, a CMU University Transportation Center T-SET grant, and a Qualcomm Innovation PhD fellowship.