Motion Estimation for Self-Driving Cars With a Generalized Camera

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Gim Hee Lee (glee@student.ethz.ch)
ETH Zürich, Switzerland

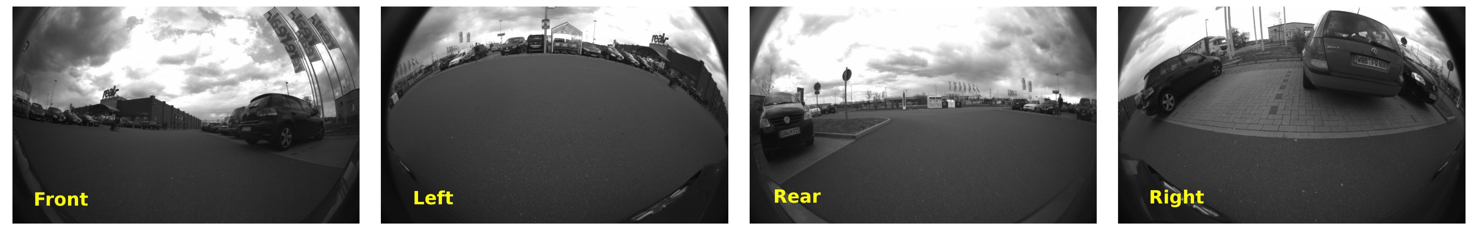
Friedrich Fraundorfer (friedrich.fraundorfer@tum.de)
Technische Universität München, Germany

Marc Pollefeys (marc.pollefeys@inf.ethz.ch)
ETH Zürich, Switzerland

Introduction

- **Objective**: Ego-motion estimation for a self-driving car equipped with a multi-camera system, i.e. a generalized camera.
- **Grobi** - Our self-driving car with a generalized camera made up of 4 fish-eye cameras with minimal/non- overlapping field-of-views.

- Sample images from the generalized camera.



- **Main contributions**:
  - New formulation of the generalized epipolar essential matrix from combining the generalized epipolar constraint (GEC) and Ackermann motion model.
  - Analytical 2-point minimal solution for ego-motion estimation with metric scale.
  - Investigation of and practical solution to the degenerate case of straight motion.

Generalized Camera Model

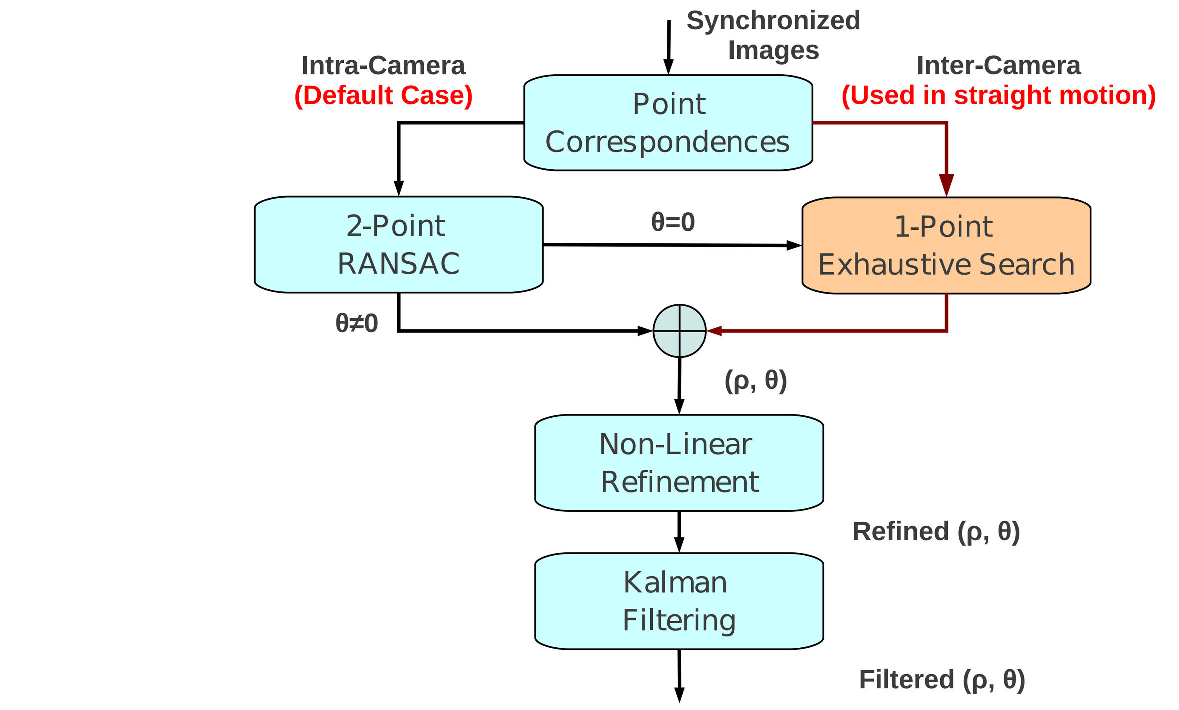
- Generalized epipolar constraint defined in [1].

$$l'^T \underbrace{\begin{bmatrix} E & R \\ R & 0 \end{bmatrix}}_{E_{GC}} l = 0 \qquad (1)$$

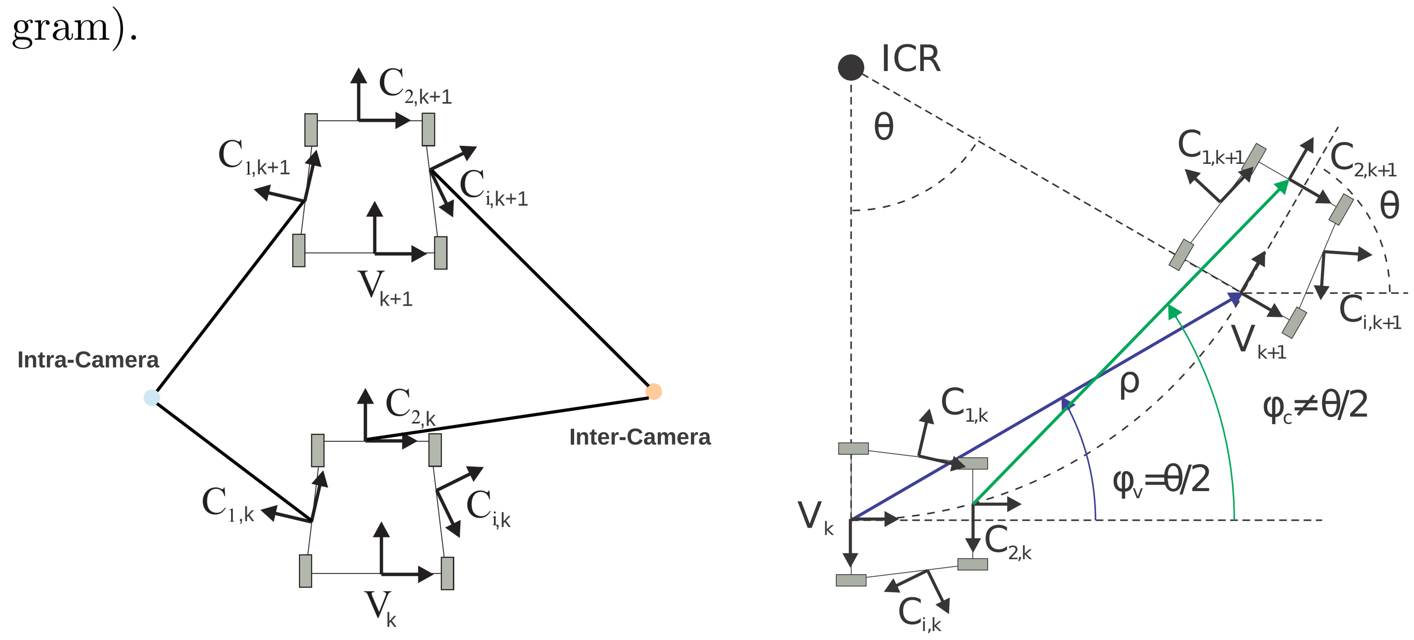
where E_{GC} is the generalized essential matrix. $l \leftrightarrow l'$ are point correspondences represented as 6-Vector Plücker lines. $E = \lfloor t \rfloor_{\times} R$ is the conventional essential matrix. R and t are the rotation and translation of the relative motion.

Motion Estimation

- System overview for motion estimation with the generalized camera on a car.



- Inter-camera and intra-camera point correspondences (left diagram).
- Relation between generalized camera and Ackermann motion (right diagram).



- Ackermann motion is parameterized by the relative yaw angle θ and scale ρ of the relative translation.

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad t = \rho \begin{bmatrix} \cos\varphi_v \\ \sin\varphi_v \\ 0 \end{bmatrix} \qquad (2)$$

- Putting Equation 2 into E_{GC} from Equation 1, we get the generalized essential matrix with Ackermann motion.

$$E_{GC} = \begin{bmatrix} 0 & 0 & \rho\sin\frac{\theta}{2} & \cos\theta & -\sin\theta & 0 \\ 0 & 0 & -\rho\cos\frac{\theta}{2} & \sin\theta & \cos\theta & 0 \\ \rho\sin\frac{\theta}{2} & \rho\cos\frac{\theta}{2} & 0 & 0 & 0 & 1 \\ \cos\theta & -\sin\theta & 0 & 0 & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad (3)$$

2-Point Minimal Solution

- Putting Equation 3 into Equation 1, we get the generalized epipolar constraint with Ackermann motion.

$$a\cos\theta + b\sin\theta + c\rho\cos\frac{\theta}{2} + d\rho\sin\frac{\theta}{2} + e = 0 \qquad (4)$$

a, b, c, d and e are coefficients computed from the Plücker line.

- 2-point correspondences are needed to solve for the 2 unknowns θ and ρ from Equation 4.
- This leads to the minimal solution of 3-degree polynomial with $\gamma = \sin^2\frac{\theta}{2}$ where the yaw angle θ can be solved in closed-form.

$$A\gamma^3 + B\gamma^2 + C\gamma + D = 0 \qquad (5)$$

A, B, C and D are the coefficients computed from the Plücker line correspondence $l \leftrightarrow l'$.

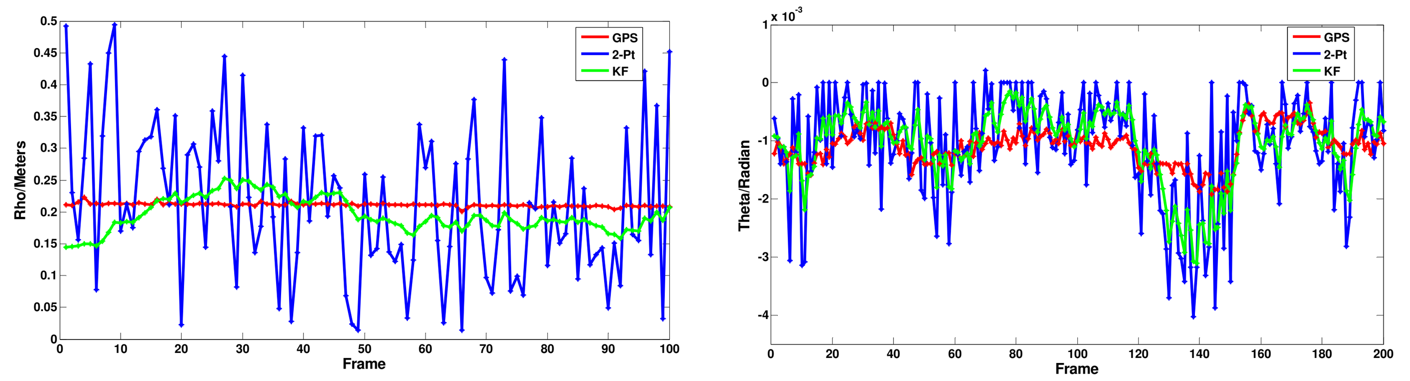
- The metric scale ρ can be solved by back-substitution.

Degenerated Case and Its Solution

- More intra-camera than inter-camera correspondences ⇒ intra-camera correspondences as the default case.
- Metric scale ρ cannot be uniquely computed when the car goes straight, i.e. $\theta = 0$ (when using intra-camera tracks only).
- **Solution**: Retrieve the metric scale with 1 additional inter-camera correspondence.

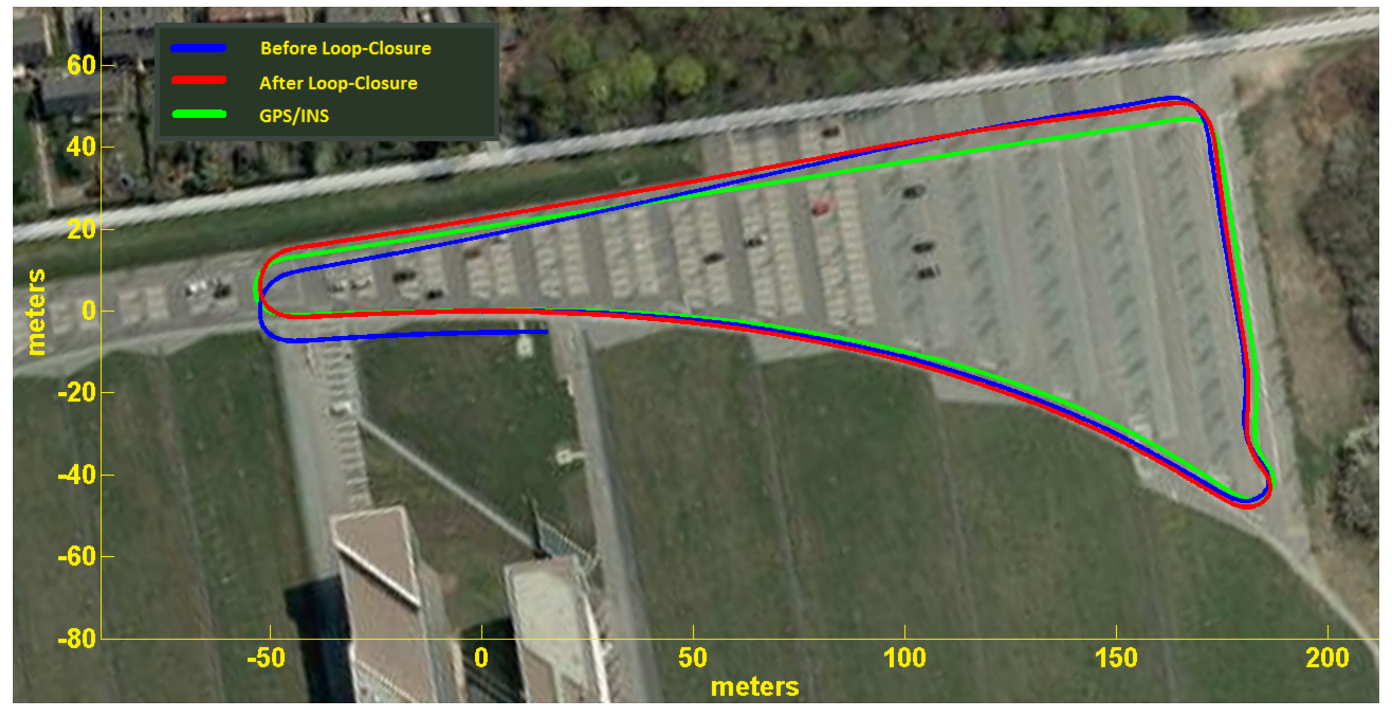
Kalman Filtering

- Two independent 1D Kalman filters with constant velocity prior to smooth out noisy estimates.
- Example of scales ρ (left) and yaw angles θ (right) between consecutive frames after Kalman filtering.

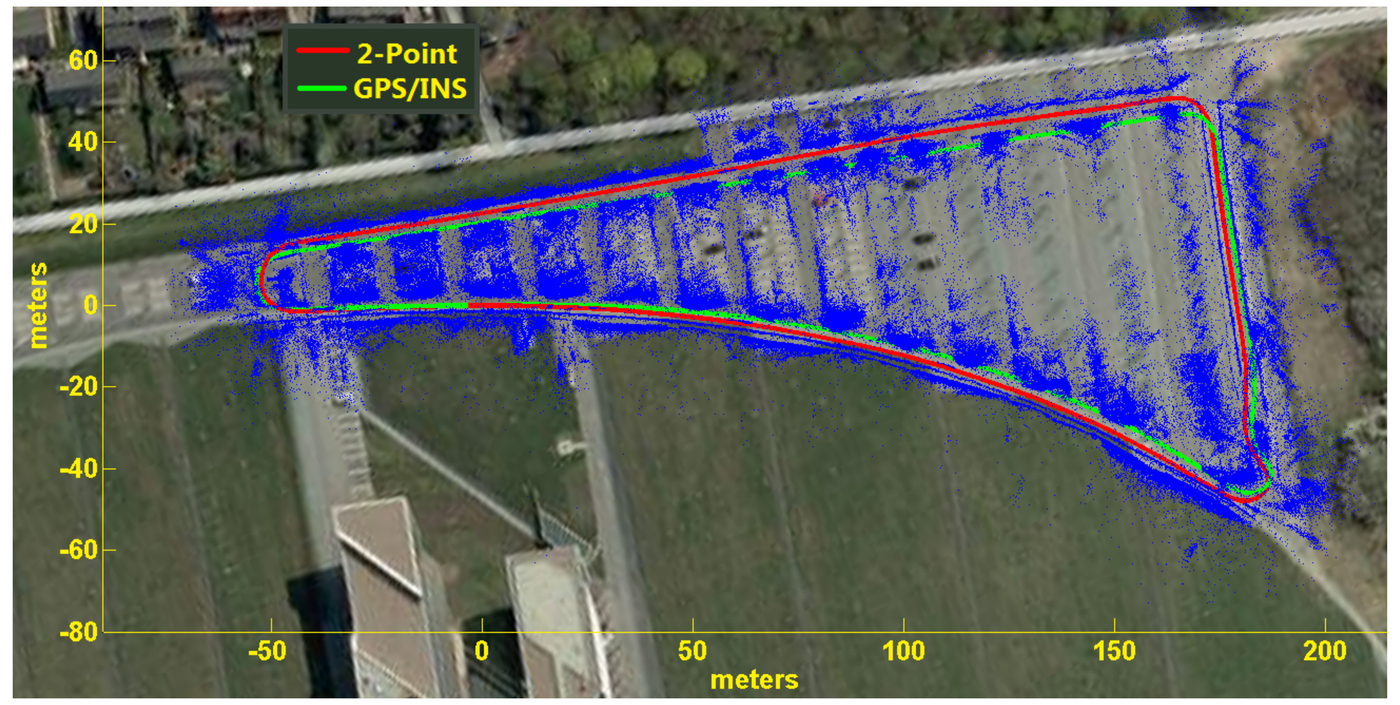


Results

- A total of 4 × 2500 images are used in the test of our algorithm.
- Additional 1 inter-camera correspondence used to compute scale for 79.9% when the car is moving straight.
- Trajectories before and after pose-graph loop-closure compared with GPS/INS ground truth.



- Top view of trajectory and 3D map points after pose-graph loop-closure and full bundle adjustment compared with GPS/INS ground truth.



References / Acknowledgement

[1] R. Pless, Using many cameras as one, CVPR 2003.

This work is supported in part by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant #269916 (v-charge) and 4DVideo ERC Starting Grant Nr. 210806.