Relative Pose Estimation for a Multi-Camera System with Known Vertical Direction

Gim Hee Lee (gimhee.lee@merl.com)
Mitsubishi Electric Research Labs, USA

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

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

IEEE 2014 Conference on Computer Vision and Pattern Recognition

Introduction

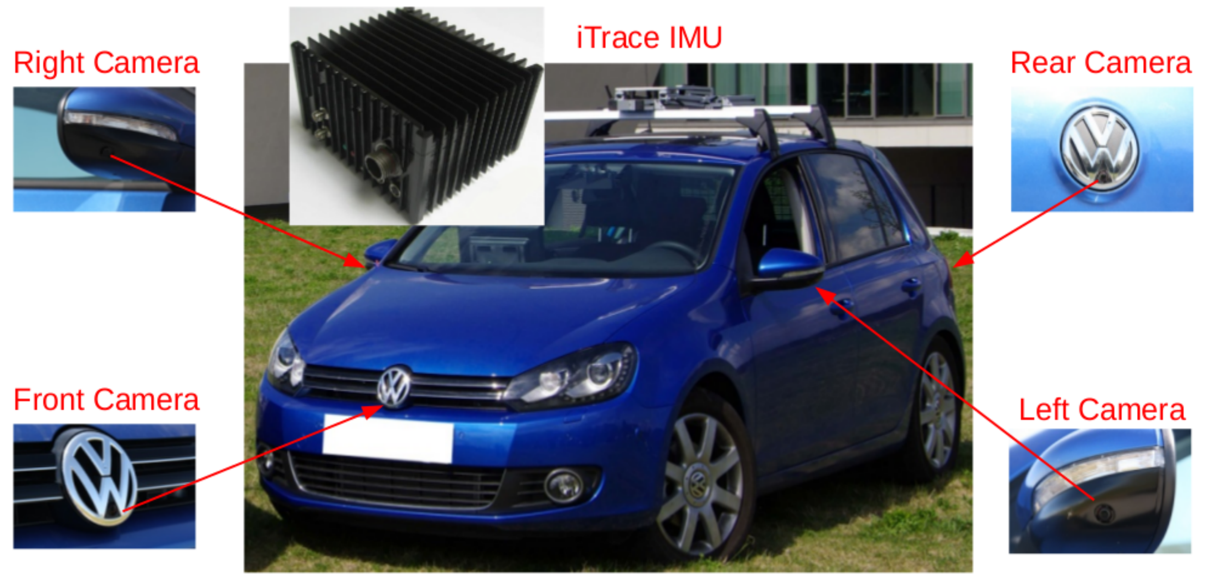
- **Objective:**

  Given:

  1. A calibrated multi-camera system, i.e. known intrinsics and extrinsics.
  2. The vertical direction (roll and pitch angles from the IMU) of two multi-camera system frames.
  3. Image point correspondences between the two multi-camera system frames.

  Find: The relative pose between the two multi-camera system frames.

- **Grobi** - Our self-driving car equppied with a multi-camera system and iTrace IMU.



Generalized Camera Model

- Relative pose R and t of a multi-camera system can be solved from the **Generalized Epipolar Constraint** defined in [1]:

$$l'^T_{ij} \underbrace{\begin{bmatrix} E & R \\ R & 0 \end{bmatrix}}_{E_{GC}} l_{ij} = 0 \qquad (1)$$

  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.

Existing Works and Their Limitations

- Linear algorithm [1] requires 17-point correspondences and minimal 6-point problem [2] gives 64 solutions, which are computationally expensive for RANSAC.
- Practical large-scale results were shown from our previous works that used the Ackermann (2-point) [3] and planarity (3-point) [4] motion models to reduce computational complexity.
- However, [3] and [4] cannot estimate the full 6-dof motion when the planarity constraint is violated.

Minimal 4-Point Algorithm

1. **Apply Roll and Pitch Angles**

   - Transform the Plücker line correspondences $l \leftrightarrow l'$ with the absolute roll ($R_r \leftrightarrow R'_r$) and pitch ($R_p \leftrightarrow R'_p$) angles from the correspondence frames.

$$\underbrace{\left(\begin{bmatrix} R'_pR'_r & 0 \\ 0 & R'_pR'_r \end{bmatrix} l'\right)^T}_{\hat{l}'} \underbrace{\begin{bmatrix} \lfloor \hat{t} \rfloor_{\mathbf{x}} \hat{R}_y & \hat{R}_y \\ \hat{R}_y & 0 \end{bmatrix}}_{\hat{E}_{GC}} \underbrace{\left(\begin{bmatrix} R_pR_r & 0 \\ 0 & R_pR_r \end{bmatrix} l\right)}_{\hat{l}} = 0 \qquad (2)$$

   - $\hat{l} \leftrightarrow \hat{l}'$ and $\hat{E}_{GC}$ are the Plücker line correspondence and generalized essential matrix after the transformation.
   - $\hat{E}_{GC}$ is made up of only a yaw $\hat{R}_y$ angle and translation vector $\hat{t}$ of the transformed relative motion.

2. **Minimal Problem**

   - The task now is to solve for $\hat{R}_y$ and $\hat{t}$, which we write as:

$$\hat{t} = \begin{bmatrix} \hat{t}_x \\ \hat{t}_y \\ \hat{t}_z \end{bmatrix}, \quad \hat{R}_y = \frac{1}{1+q^2} \begin{bmatrix} 1-q^2 & -2q & 0 \\ 2q & 1-q^2 & 0 \\ 0 & 0 & 1+q^2 \end{bmatrix} \qquad (3)$$

   where $q = \tan\frac{\hat{\theta}}{2}$ and $\hat{\theta}$ is the yaw angle that make up $\hat{R}_y$.

   - Putting Eq. (3) into Eq. (2) and with 4-point correspondences, we get a system of four polynomial eqautions of the form:

$$a_1\hat{t}_x q^2 + a_2\hat{t}_x q + a_3\hat{t}_x + a_4\hat{t}_y q^2 + a_5\hat{t}_y q +$$
$$a_6\hat{t}_y + a_7\hat{t}_z q^2 + a_8\hat{t}_z q + a_9\hat{t}_z + a_{10}q^2 + a_{11}q + a_{12} = 0 \qquad (4)$$

   where a_1 to a_12 are coefficients formed with $\hat{l} \leftrightarrow \hat{l}'$.

   - Apply the "Hidden Variable Resultant" method to get an 8 degree univariate polynomial

$$Aq^8 + Bq^7 + Cq^6 + Dq^5 + Eq^4 + Fq^3 + Gq^2 + Hq + I = 0 \qquad (5)$$

   where A to I are known coefficients formed from the coefficients of the system of polynomials.

   - Eq. (5) gives up to 8 real solutions for q, and $\hat{t}$ can be obtained from back substitution.

3. **Recover Relative Pose**

   - Finally, the relative pose R and t can be recovered from undoing the absolute roll and pitch transformation.

$$R = R_r'^T R_p'^T \hat{R}_y R_p R_r, \qquad t = R_r'^T R_p'^T \hat{t} \qquad (6)$$

Linear 8-Point Algorithm

- $\hat{E}_{GC}$ can be written as Eq. (7) which consists of 9 unique entries.

$$\hat{E}_{GC} = \begin{bmatrix} -e_1 & -e_2 & e_3 & r_1 & -r_2 & 0 \\ e_2 & -e_1 & -e_4 & r_2 & r_1 & 0 \\ e_5 & e_6 & 0 & 0 & 0 & r_3 \\ r_1 & -r_2 & 0 & 0 & 0 & 0 \\ r_2 & r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & r_3 & 0 & 0 & 0 \end{bmatrix} \qquad (7)$$

- Replacing $\hat{E}_{GC}$ from Eq. (2) with Eq. (7), we get

$$\mathcal{A}\mathcal{E} = 0 \qquad (8)$$

  where $\mathcal{A}$ is a 8×9 matrix formed with the eight Plücker line correspondences and $\mathcal{E} = [e_1,\ e_2,\ e_3,\ e_4,\ e_5,\ e_6,\ r_1,\ r_2,\ r_3]^T$, with $r_3 = 1$.

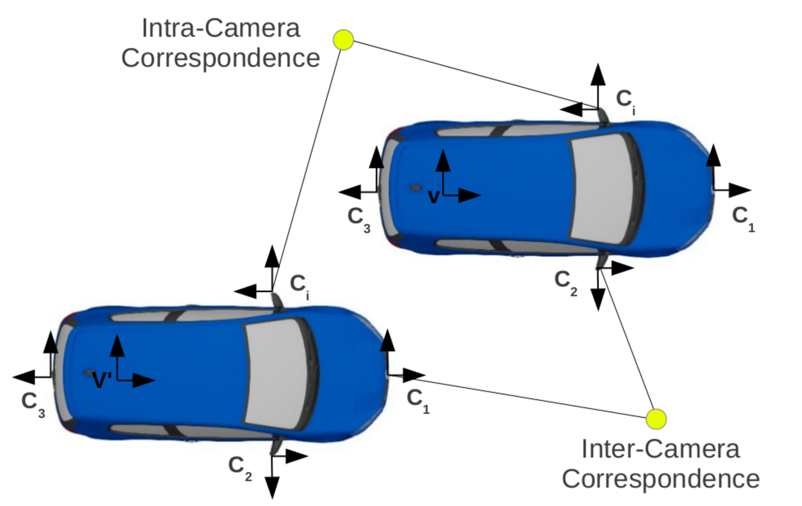
- **Degenerate Case**: Rank of $\mathcal{A}$ drops to 7 when there is no relative roll and pitch $\Rightarrow$ solution is no longer unqiue but given by a family of solutions ($\lambda \hat{E}_{GC}\ \lambda \hat{R}_y + \mu I$).
- We circumvent the ambiguity in the rotation matrix $\hat{R}_y$ by enforcing the constraint $\|e\| = 1$ on the elements of the essential matrix instead of $\|\mathcal{E}\| = 1$ while minimizing $\mathcal{A}\mathcal{E}$ using the SVD method.
- This is equivalent to solving

$$(\mathcal{A}_r\mathcal{A}_r^+ - I)\mathcal{A}_e e = 0 \qquad (9)$$

  where $\mathcal{A}_e$ and $\mathcal{A}_r$ are made up of the first six and last three columns of $\mathcal{A}$ respectively, and $\mathcal{A}_r^+$ is the pseudo-inverse of $\mathcal{A}_r$.

- An unique solution for e can be obtained by applying the SVD method on the matrix $(\mathcal{A}_r\mathcal{A}_r^+ - I)\mathcal{A}_e$.

Degenerate Case and Its Solution

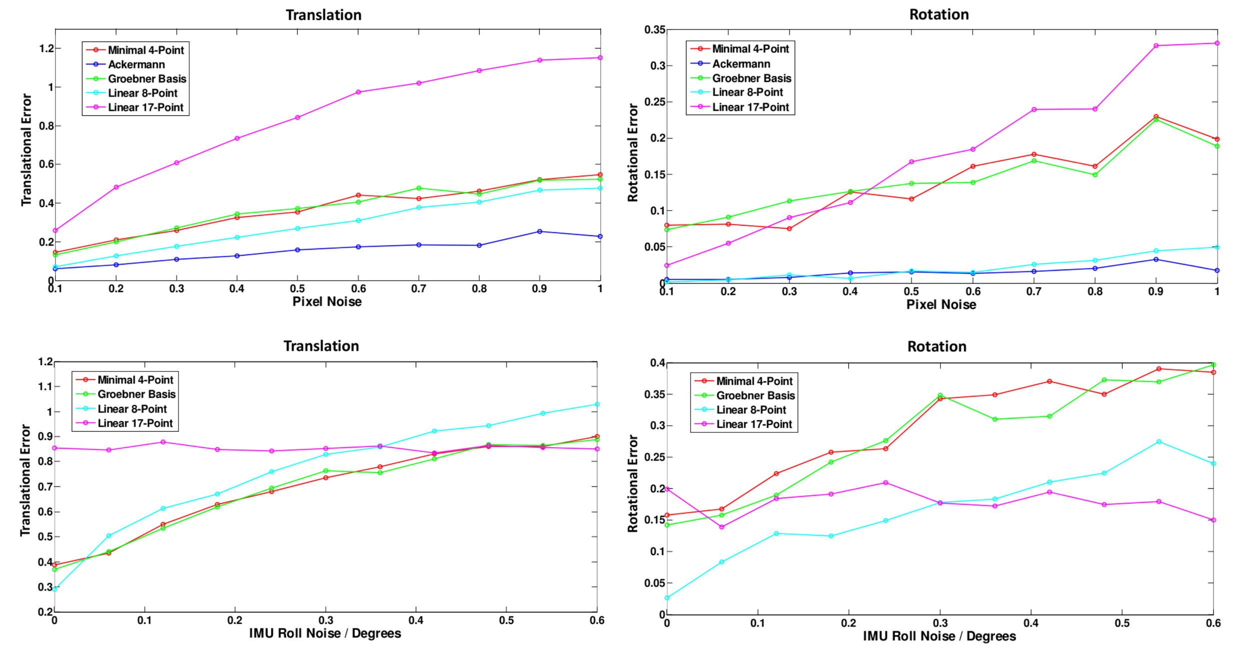


- Scale cannot be estimated from the 4-point and 8-point algorithms when the multi-camera system undergoes pure translation while having only intra-camera correspondences.
- We use an additional inter-camera correspondence to circumvent the scale problem in the degenerate case.

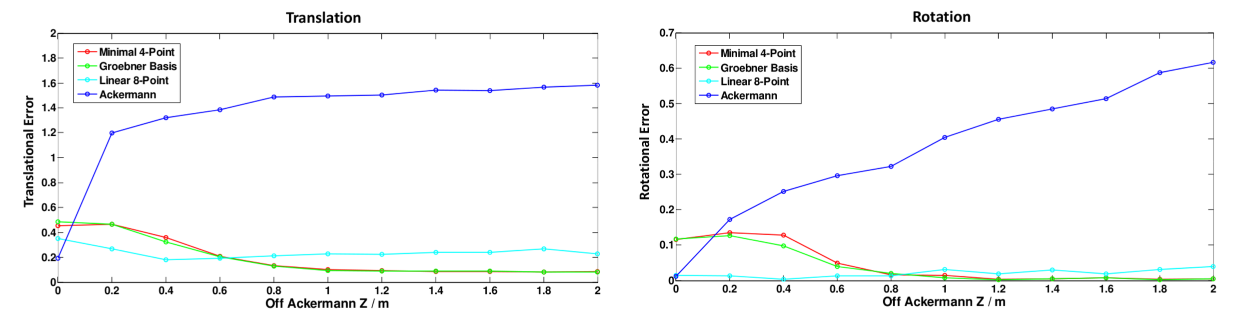
Results

- **Simulations**

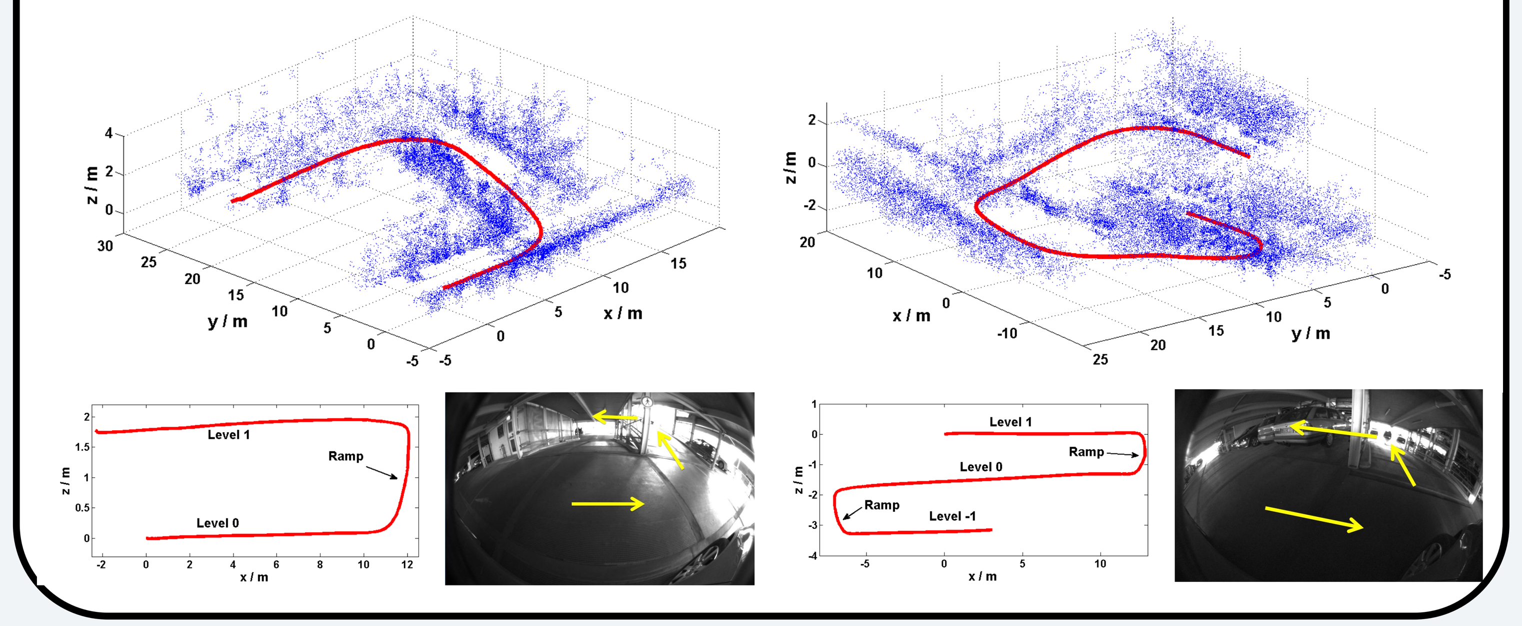
  1. Comparison with other methods under pixel and IMU noise.



  2. Comparison with [2] under Ackermann violation in the z-axis.



- **Real-World Datasets**



Acknowledgement and References

The first author was at ETH Zürich when this work was done. It was funded by the under the grant #269916 (v-charge) and 4DVideo ERC Starting Grant Nr. 210806.

1. R. Pless, "Using many cameras as one", CVPR 2003 .
2. H. Stewénius *et. al*, "Solutions to minimal generalized relative pose problems", OMNIVIS 2005.
3. G.H. Lee *et. al*, "Motion estimation for a self-driving car with a generalized camera", CVPR 2013.
4. G.H. Lee *et. al*, "Structureless pose-graph loop-closure with a multi-camera system on a self-driving car", IROS 2013.