Plane-ORB-SLAM

Sun Qinxuan

May 6, 2019

OUTLINE

Robust Optical-Flow Based Self-Motion Estimation

- V. Grabe, H. H. Bülthoff and P. Robuffo Giordano, "Robust optical-flow based self-motion estimation for a quadrotor UAV," *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, 2012, pp. 2153-2159.

Continuous Homography Constraint

- $\dot{\boldsymbol{X}} = \hat{\boldsymbol{\omega}}\boldsymbol{X} + \boldsymbol{v}$
- Consider a set of features located on a common plane of equation $\boldsymbol{N}^T\boldsymbol{X} = d$, where $\boldsymbol{N} \in \mathbb{S}^2$, $d \in \mathbb{R}$.

$$\dot{\boldsymbol{X}} = \hat{\boldsymbol{\omega}}\boldsymbol{X} + \boldsymbol{v}\frac{1}{d}\boldsymbol{N}^T\boldsymbol{X} = \left(\hat{\boldsymbol{\omega}} + \boldsymbol{v}\frac{1}{d}\boldsymbol{N}^T\right)\boldsymbol{X} = \boldsymbol{H}\boldsymbol{X} \tag{1}$$

- The matrix $\boldsymbol{H} \in \mathbb{R}^{3\times 3}$ is *continuous homography matrix*.
- $\boldsymbol{H}$ encodes the camera linear/angular velocity $(\boldsymbol{v}, \boldsymbol{\omega})$ and the plane structure $(\boldsymbol{N}, d)$.

Continuous Homography Constraint

- $\lambda \boldsymbol{x} = \boldsymbol{X}$
- $\dot{\boldsymbol{X}} = \dot{\lambda}\boldsymbol{x} + \lambda\dot{\boldsymbol{x}} = \dot{\lambda}\boldsymbol{x} + \lambda\boldsymbol{u}$

$$\boldsymbol{u} = \boldsymbol{H}\boldsymbol{x} - \frac{\dot{\lambda}}{\lambda}\boldsymbol{x} \tag{2}$$

- Continuous homography constraint

$$\hat{\boldsymbol{x}}\boldsymbol{H}\boldsymbol{x} = \hat{\boldsymbol{x}}\boldsymbol{u} \tag{3}$$

Extended 4-Point Algorithm

- Continuous homography constraint

$$\hat{\boldsymbol{x}}\boldsymbol{H}\boldsymbol{x} = \hat{\boldsymbol{x}}\boldsymbol{u} \tag{4}$$

- Stack the elements of $\boldsymbol{H}$ into the vector $\boldsymbol{H}^S = [H_{11}, H_{21}, \cdots, H_{33}] \in \mathbb{R}^9$ and rewrite (4) as

$$\boldsymbol{a}^T\boldsymbol{H}^S = \hat{\boldsymbol{x}}\boldsymbol{u} \tag{5}$$

- Stack all the $\boldsymbol{a}_i$ from n tracked features into $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]^T \in \mathbb{R}^{3n\times 9}$ and stack all $\hat{\boldsymbol{x}}_i\boldsymbol{u}_i$ into $\boldsymbol{B} = [\hat{\boldsymbol{x}}_1\boldsymbol{u}_1, \cdots, \hat{\boldsymbol{x}}_n\boldsymbol{u}_n]^T \in \mathbb{R}^{3n}$.

$$\boldsymbol{A}\boldsymbol{H}^S = \boldsymbol{B} \tag{6}$$

Extended 4-Point Algorithm

- The angular velocity is obtained directly from the IMU $\boldsymbol{\omega} = \boldsymbol{\omega}_{IMU}$.
- Subtract the rotational components from the perceived flow using the interaction matrix which relates $\boldsymbol{u}$ to $(\boldsymbol{v}, \boldsymbol{\omega})$.

$$\begin{bmatrix} u'_x \\ u'_y \end{bmatrix} = \begin{bmatrix} u_x \\ u_y \end{bmatrix} - \begin{bmatrix} -x_x x_y & 1 + x_x^2 & -x_y \\ -(1 + x_y)^2 & x_x x_y & x_x \end{bmatrix} \boldsymbol{\omega} \tag{7}$$

- Thus $\boldsymbol{H}$ reduces to

$$\boldsymbol{H} = \frac{1}{d} \boldsymbol{v} \boldsymbol{N}^T \tag{8}$$

- $\boldsymbol{N}$ spans $\boldsymbol{H}^T$ and $\|\boldsymbol{N}\| = 1$ and so $\boldsymbol{N}$ can be obtained by an SVD of $\boldsymbol{H}$.
- $\boldsymbol{v}/d = \boldsymbol{H}\boldsymbol{N}$.

Planarity Measures

- To test whether a certain group of observed features belongs to a common plane.
- Consider two different quantitative measures.
- First one,
  - $\boldsymbol{A}\boldsymbol{H}^S = \boldsymbol{B}$
  - $\text{rank}(\boldsymbol{A}) = \text{rank}([\boldsymbol{A}\ \boldsymbol{B}]) = 8$
  - $\sigma_i, i = 1, \cdots, 10$ are the singular values of $[\boldsymbol{A}\ \boldsymbol{B}] \in \mathbb{R}^{3n\times 10}$ in decreasing order.
  - The value σ_9 can be exploited as a measure of how well a certain set of features/optical flow meets the planarity constraint.
- Second one,
  - $\boldsymbol{H}^S = \boldsymbol{A}^\dagger \boldsymbol{B}$ is the LS solution.
  - Consider the "reprojection" vector

$$\boldsymbol{R} = \boldsymbol{B} - \boldsymbol{A}\boldsymbol{H}^S = \boldsymbol{B} - \boldsymbol{A}\boldsymbol{A}^\dagger \boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^\dagger)\boldsymbol{B} \tag{9}$$

  - another measure $r = \|\boldsymbol{R}\|/n$

Map Initialization in ORB-SLAM

- Parallel computation of two models: a homography $\boldsymbol{H}_{cr}$ and a fundamental matrix $\boldsymbol{F}_{cr}$

$$\boldsymbol{x}_c = \boldsymbol{H}_{cr}\boldsymbol{x}_r \quad \boldsymbol{x}_c^T \boldsymbol{F}_{cr}\boldsymbol{x}_r = 0 \tag{10}$$

- The homography is computed with the normalized DLT algorithm.
- Eight motion hypotheses are retrieved from the homography. [1]

[1] O. D. Faugeras and F. Lustman, “Motion and structure from motion in a piecewise planar environment,” Int. J. Pattern Recog. Artif. Intell., vol. 2, no. 03, pp. 485508, 1988.

Homography Constraint Consistent with ORB-SLAM

- Homography constraint

$$\boldsymbol{x}_c = \boldsymbol{H}_{cr}\boldsymbol{x}_r \tag{11}$$

with

$$\boldsymbol{H}_{cr} = d_r\boldsymbol{R}_{cr} + \boldsymbol{t}_{cr}\boldsymbol{n}_r^T \tag{12}$$

encodes the information of camera poses and the plane $\boldsymbol{n}_r^T\boldsymbol{X}_r = d_r$.

- DLT algorithm [2]

$$\hat{\boldsymbol{x}}_{ci}\boldsymbol{H}_{cr}\boldsymbol{x}_{ri} = 0 \tag{13}$$

rewrite (13) as

$$\boldsymbol{A}_i\boldsymbol{H}^S = 0 \tag{14}$$

Stacking $\boldsymbol{A} = [\boldsymbol{A}_1^T, \cdots, \boldsymbol{A}_n^T]^T \in \mathbb{R}^{3n\times 9}$ yields the similar equation (compared with the continuous homography constraint)

$$\begin{aligned} \min \quad & \|\boldsymbol{A}\boldsymbol{H}^S\| \\ \text{s.t.} \quad & \|\boldsymbol{H}^S\| = 1 \end{aligned} \tag{15}$$

The solution is the unit eigenvector of $\boldsymbol{A}^T\boldsymbol{A}$ with the least eigenvalue.

[2] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision". Cambridge University Press, 2003, second Edition.

Homography Constraint Consistent with ORB-SLAM

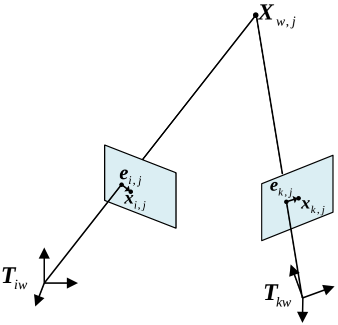
- Similarly to Ref.[[3]] the planarity measure can be defined:
- First one,
  - 9-th singular value σ_9 of $\boldsymbol{A}$.
- Second one,
  - $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{H}^S = \boldsymbol{A}\boldsymbol{v}_9 = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{v}_9 = \sigma_9\boldsymbol{u}_9$
  - $r = \|\boldsymbol{R}\|/n = \sigma_9/n$

[3] V. Grabe, H. H. Bülthoff and P. Robuffo Giordano, "Robust optical-flow based self-motion estimation for a quadrotor UAV," *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, 2012, pp. 2153-2159.

Homography Constraint in BA

- BA optimization in ORB-SLAM
  - The error term for the observation of a map point j in a keyframe i is

$$e_{i,j} = x_{i,j} - \pi_i(T_{iw}, X_{w,j}) \qquad (16)$$

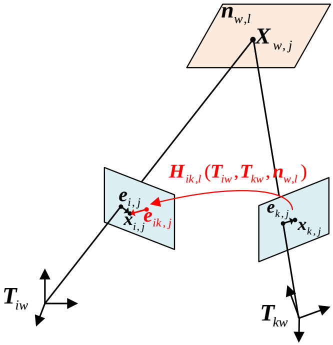


Homography Constraint in BA

- Homography constraint in BA
  - Given that the map point j is on a plane indexed by l,
  - Add an error term for the homography transform between two frames i and k w.r.t. the plane l (associated with the map point j) in the map.

$$\boldsymbol{e}_{ik,j} = \boldsymbol{x}_{i,j} - \boldsymbol{H}_{ik,l}(\boldsymbol{T}_{iw}, \boldsymbol{T}_{kw}, \boldsymbol{n}_{w,l}) \cdot \boldsymbol{x}_{k,j} \quad (17)$$

  - unknowns: $\boldsymbol{T}_{iw} \in \mathbb{SE}(3)$, $\boldsymbol{X}_{w,j} \in \mathbb{R}^3$, $\boldsymbol{n}_{w,l} \in \mathbb{S}^2$.



Homography Constraint in BA

- Optimization of $\boldsymbol{n}_{w,l} \in \mathbb{S}^2$.
  - The local update is defined in the tangent space of $\boldsymbol{n}_{w,l}$ at $\mathbb{S}^2$.
  - Local update $\boldsymbol{\zeta} \in \mathbb{R}^2$
  - updated variable $\boldsymbol{n}'_{w,l} = \mathrm{N}\left(\boldsymbol{J}_\zeta(\boldsymbol{n}_{w,l})\boldsymbol{\zeta} + \boldsymbol{n}_{w,l}\right)$. [4]
  - $\boldsymbol{J}_\zeta(\boldsymbol{n}_{w,l}) = \mathrm{null}(\boldsymbol{n}_{w,l}^T)$
  - $\mathrm{N}(\cdot)$ normalized a vector to unit length.
  - Jacobian matrix of the error w.r.t. $\boldsymbol{\zeta}$

$$\frac{\partial \boldsymbol{e}_{ik,j}}{\boldsymbol{\zeta}} = \boldsymbol{t}_{iw}\boldsymbol{x}_{k,j}^T(\boldsymbol{I} - \boldsymbol{n}_{w,l}\boldsymbol{n}_{w,l}^T)\boldsymbol{J}_\zeta(\boldsymbol{n}_{w,l}) \tag{18}$$

[4] W. Förstner. "Minimal Representations for Testing and Estimation in Projective Spaces." Asian Conference on Computer Vision, 2010.