1 State Parametrization

IMU State at Time k (16×1)

$$\mathbf{X}_{I_k} = \begin{bmatrix} \mathbf{q}_{I_k G} \\ \mathbf{p}_k^G \\ \mathbf{v}_k^G \\ \mathbf{b}_{g,k} \\ \mathbf{b}_{a,k} \end{bmatrix} \begin{matrix} \leftarrow \text{Global to IMU rotation unit quaterion} \\ \leftarrow \text{IMU position in global frame} \\ \leftarrow \text{IMU velocity in global frame} \\ \leftarrow \text{Gyro bias} \\ \leftarrow \text{Accelerometer bias} \end{matrix} \tag{1}$$

IMU Error State at Time k (15×1)

$$\widetilde{\mathbf{X}}_{I_k} = \begin{bmatrix} \widetilde{\boldsymbol{\theta}}_k^G \\ \widetilde{\mathbf{p}}_k^G \\ \widetilde{\mathbf{v}}_k^G \\ \widetilde{\mathbf{b}}_{g,k} \\ \widetilde{\mathbf{b}}_{a,k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\delta\theta}_I \\ \mathbf{p}_k^G - \hat{\mathbf{p}}_k^G \\ \mathbf{v}_k^G - \hat{\mathbf{v}}_k^G \\ \mathbf{b}_{g,k} - \hat{\mathbf{b}}_{g,k} \\ \mathbf{b}_{a,k} - \hat{\mathbf{b}}_{a,k} \end{bmatrix} \tag{2}$$

where

$$\delta\mathbf{q} = (\hat{\mathbf{q}}_{IG})^{-1} \otimes \mathbf{q}_{IG} \simeq \begin{bmatrix} \frac{1}{2}\boldsymbol{\delta\theta} \\ 1 \end{bmatrix} \tag{3}$$

MSCKF State at Time k $((16 + 7N) \times 1)$

$$\hat{\mathbf{X}}_k = \begin{bmatrix} \hat{\mathbf{X}}_{I_k} \\ \hat{\mathbf{q}}_{C_1 G} \\ \hat{\mathbf{p}}_{C_1}^G \\ \vdots \\ \hat{\mathbf{q}}_{C_N G} \\ \hat{\mathbf{p}}_{C_N}^G \end{bmatrix} \begin{matrix} \leftarrow \text{IMU State estimate} \\ \leftarrow \text{Camera pose 1 orientation estimate} \\ \leftarrow \text{Camera pose 1 position estimate} \\ \vdots \\ \leftarrow \text{Camera pose N orientation estimate} \\ \leftarrow \text{Camera pose N position estimate} \end{matrix} \tag{4}$$

MSCKF Error State $((15 + 6N) \times 1)$

$$\widetilde{\mathbf{X}}_k = \begin{bmatrix} \widetilde{\mathbf{X}}_{I_k} \\ \boldsymbol{\delta\theta}_{C_1} \\ \widetilde{\mathbf{p}}_{C_1}^G \\ \vdots \\ \boldsymbol{\delta\theta}_{C_N} \\ \widetilde{\mathbf{p}}_{C_N}^G \end{bmatrix} \tag{5}$$

2 IMU State Estimate Propagation

Integrate these with RK5 from t_{k-1} to $t_{k-1}+T$ where T is the sampling period of the IMU:

$$\dot{\hat{\mathbf{q}}}_{IG} = \frac{1}{2}\boldsymbol{\Omega}\left(\boldsymbol{\omega}_m - \hat{\mathbf{b}}_a - \hat{\mathbf{C}}_{IG}\boldsymbol{\omega}_G\right)\hat{\mathbf{q}}_{IG} \tag{6}$$

$$\dot{\hat{\mathbf{b}}}_g = \mathbf{0}_{3\times 1} \tag{7}$$

$$\dot{\hat{\mathbf{v}}}_I^G = \hat{\mathbf{C}}_{IG}^T\left(\mathbf{a}_m - \hat{\mathbf{b}}_a\right) - 2\left(\boldsymbol{\omega}_G^\times\right)\hat{\mathbf{v}}_I^G - \left(\boldsymbol{\omega}_G^\times\right)^2\hat{\mathbf{p}}_I^G + \mathbf{g}^G \tag{8}$$

$$\dot{\hat{\mathbf{b}}}_a = \mathbf{0}_{3\times 1} \tag{9}$$

$$\dot{\hat{\mathbf{p}}}_I^G = \hat{\mathbf{v}}_I^G \tag{10}$$

$$\dot{\widetilde{\mathbf{X}}}_I = \mathbf{F}\widetilde{\mathbf{X}}_I + \mathbf{G}\mathbf{n}_I \tag{11}$$

where

$$\boldsymbol{\Omega}(\boldsymbol{\omega}) = \begin{bmatrix} -\boldsymbol{\omega}^\times & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix} \text{ is } 4\times 4, \tag{12}$$

$\hat{\mathbf{C}}_{IG} := \mathbf{C}(\hat{\mathbf{q}}_{IG})$ is the rotation matrix form of the quaternion $\hat{\mathbf{q}}_{IG}$,

$$\mathbf{F} = \begin{bmatrix} -\hat{\boldsymbol{\omega}}^\times & -\mathbf{1}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ -\hat{\mathbf{C}}_{IG}^T\hat{\mathbf{a}}^\times & \mathbf{0}_{3\times3} & -2\boldsymbol{\omega}_G^\times & -\hat{\mathbf{C}}_{IG}^T & -\left(\boldsymbol{\omega}_G^\times\right)^2 \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{1}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix} \text{ is } 15\times 15, \tag{13}$$

$$\mathbf{G} = \begin{bmatrix} -\mathbf{1}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{1}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -\hat{\mathbf{C}}_{IG}^T & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{1}_3 \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix} \text{ is } 15\times 12, \tag{14}$$

and

$$\mathbf{n}_I = \begin{bmatrix} \mathbf{n}_g \\ \mathbf{n}_{wg} \\ \mathbf{n}_a \\ \mathbf{n}_{wa} \end{bmatrix} \begin{matrix} \leftarrow \text{Gyro noise} \\ \leftarrow \text{Gyro-bias rate-of-change noise} \\ \leftarrow \text{Accelerometer noise} \\ \leftarrow \text{Accelerometer-bias rate-of-change noise} \end{matrix} \tag{15}$$

is the system noise whose covariance matrix $\mathbf{Q}_I$ is computed offline during calibration. Note, $\boldsymbol{\omega}_G$ is the angular velocity of the spinning Earth in our global frame. The magnitude (i.e. rotational speed) is equal to $7.292\times 10^5\,\frac{\text{rad}}{s}$.

For the covariance,

$$\hat{\mathbf{P}}_k^- = \begin{bmatrix} \hat{\mathbf{P}}_{II,k}^- & \mathbf{\Phi}\left(t_{k-1}+T, t_{k-1}\right)\hat{\mathbf{P}}_{IC,k-1} \\ \hat{\mathbf{P}}_{IC,k-1}^T \mathbf{\Phi}\left(t_{k-1}+T, t_{k-1}\right)^T & \hat{\mathbf{P}}_{CC,k-1} \end{bmatrix} \text{ is } (15+6N)\times(15+6N) \tag{16}$$

where:

- $\hat{\mathbf{P}}_{CC,k-1}$ is the $6N \times 6N$ covariance matrix of the camera pose estimates (see section on State Augmentation),
- $\hat{\mathbf{P}}_{IC,k-1}$ is the correlation between the errors in the IMU state and the camera pose estimates,
- the state transition matrix $\mathbf{\Phi}\left(t_{k-1}+T, t_{k-1}\right)$ is computed by integrating (with RK5)

$$\dot{\mathbf{\Phi}}\left(t_{k-1}+\tau, t_{k-1}\right) = \mathbf{F}\mathbf{\Phi}\left(t_{k-1}+\tau, t_{k-1}\right), \qquad \tau \in [0, T] \tag{17}$$

  with initial condition $\mathbf{\Phi}\left(t_{k-1}, t_{k-1}\right) = \mathbf{1}_{15}$, and
- $\hat{\mathbf{P}}_{II,k}^-$ is obtained by integrating (with RK5)

$$\dot{\hat{\mathbf{P}}}_{II}\left(t_{k-1}, t_{k-1}+\tau\right) = \mathbf{F}\hat{\mathbf{P}}_{II}\left(t_{k-1}, t_{k-1}+\tau\right) + \hat{\mathbf{P}}_{II}\left(t_{k-1}, t_{k-1}+\tau\right)\mathbf{F}^T + \mathbf{G}\mathbf{Q}_I\mathbf{G}^T, \qquad \tau \in [0, T] \tag{18}$$

  with initial condition $\mathbf{P}_{II}\left(t_{k-1}, t_{k-1}\right) = \hat{\mathbf{P}}_{II,k-1}$.

3 State Augmentation

3.1 Camera Poses

For the $(N+1)^{\text{th}}$ image, the camera pose is estimated as

$$\hat{\mathbf{q}}_{CG} = \mathbf{q}_{CI} \otimes \hat{\mathbf{q}}_{IG} \tag{19}$$

$$\hat{\mathbf{p}}_C^G = \hat{\mathbf{p}}_I^G + \hat{\mathbf{C}}_{IG}^T \, \mathbf{p}_C^I \tag{20}$$

and the EKF covariance matrix is augmented according to

$$\hat{\mathbf{P}}_{k-1} \leftarrow \begin{bmatrix} \mathbf{1}_{15+6N} \\ \mathbf{J} \end{bmatrix} \hat{\mathbf{P}}_{k-1} \begin{bmatrix} \mathbf{1}_{15+6N} \\ \mathbf{J} \end{bmatrix}^T \tag{21}$$

where the Jacobian

$$\mathbf{J} = \begin{bmatrix} \hat{\mathbf{C}}_{CI} & \mathbf{0}_{3\times 9} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 6N} \\ \left(\hat{\mathbf{C}}_{IG}^T \, \mathbf{p}_C^I\right)^\times & \mathbf{0}_{3\times 9} & \mathbf{1}_3 & \mathbf{0}_{3\times 6N} \end{bmatrix} \text{ is } 6 \times (15+6N) \tag{22}$$

4 Correction Equations

Residuals (pretend everything has a k subscript...)

$$\mathbf{r}_n = \mathbf{T}_H \widetilde{\mathbf{X}} + \mathbf{n}_n \tag{23}$$

where $\mathbf{T}_H$ (an upper-triangular matrix) is obtained from the QR decomposition of $\mathbf{H}_o$ ($\mathbf{H}_\mathbf{x}$ in the original tech report, but I think that's an error...)

$$\mathbf{H}_o = \begin{bmatrix} \mathbf{H}_o^{(1)} \\ \vdots \\ \mathbf{H}_o^{(L)} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{T}_H \\ \mathbf{0} \end{bmatrix}. \tag{24}$$

Each $\mathbf{H}_o^{(j)}$ is the projection of $\mathbf{H}_\mathbf{x}^{(j)}$ onto the left nullspace of $\mathbf{H}_f^{(j)}$

T_H and $\mathbf{r}_n$ can be computed in $O(r^2 d)$ time using Givens rotations without having to form $\mathbf{Q}_1$ explicitly. [FIGURE OUT HOW TO DO THIS]

Each $\mathbf{H}_\mathbf{x}^{(j)}$ is a stack of Jacobians $\mathbf{H}_{\mathbf{x}_i}^{(j)}$ of the i^{th} measurement of feature j with respect to the state (only the entries corresponding to pose i are non-zero):

$$\mathbf{H}_{\mathbf{x}_i}^{(j)} = \begin{bmatrix} \mathbf{0}_{2\times15} & \mathbf{0}_{2\times6} & \dots & \mathbf{J}_i^{(j)} \left(\hat{\mathbf{X}}_{f_j}^{C_i} \right)^\times & -\mathbf{J}_i^{(j)} \hat{\mathbf{C}}_{C_i G} & \dots & \mathbf{0}_{2\times6} \end{bmatrix} \text{ is } 2 \times (15 + 6N) \tag{25}$$

Each $\mathbf{H}_f^{(j)}$ is a stack of Jacobians $\mathbf{H}_{f_i}^{(j)}$ of the i^{th} measurement of feature j with respect to the feature position:

$$\mathbf{H}_{f_i}^{(j)} = \mathbf{J}_i^{(j)} \hat{\mathbf{C}}_{C_i G} \text{ is } 2 \times 3 \tag{26}$$

In both equations,

$$\mathbf{J}_i^{(j)} = \frac{1}{\hat{Z}_j^{C_i}} \begin{bmatrix} 1 & 0 & -\frac{\hat{X}_j^{C_i}}{\hat{Z}_j^{C_i}} \\ 0 & 1 & -\frac{\hat{Y}_j^{C_i}}{\hat{Z}_j^{C_i}} \end{bmatrix} \tag{27}$$

where

$$\begin{bmatrix} \hat{X}_j^{C_i} \\ \hat{Y}_j^{C_i} \\ \hat{Z}_j^{C_i} \end{bmatrix} = \hat{\mathbf{C}}_{C_i G} \left(\hat{\mathbf{p}}_{f_j}^G - \hat{\mathbf{p}}_{C_i}^G \right) \tag{28}$$

are the 3D coordinates of feature j in the frame of image i.

$\mathbf{n}_n$ is a noise vector with covariance matrix

$$\mathbf{R}_n = \sigma_{\text{im}}^2 \mathbf{1}_r \tag{29}$$

where r is the number of columns in $\mathbf{Q}_1$.

Kalman Gain

$$\mathbf{K}_k = \hat{\mathbf{P}}_k^- \mathbf{T}_{H,k}^T \left(\mathbf{T}_{H,k} \hat{\mathbf{P}}_k^- \mathbf{T}_{H,k}^T + \mathbf{R}_{n,k} \right)^{-1} \tag{30}$$

State Vector Correction

$$\Delta \mathbf{X}_k = \mathbf{K}_k \mathbf{r}_{n,k} \tag{31}$$

State Covariance Correction

$$\hat{\mathbf{P}}_k = \left(\mathbf{1}_{15+6N} - \mathbf{K}_k \mathbf{T}_{H,k} \right) \hat{\mathbf{P}}_k^- \left(\mathbf{1}_{15+6N} - \mathbf{K}_k \mathbf{T}_{H,k} \right)^T + \mathbf{K}_k \mathbf{R}_{n,k} \mathbf{K}_k^T \tag{32}$$