ECE276A: Sensing & Estimation in Robotics
Lecture 13: Visual-Inertial SLAM

Instructor:

Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:

Mo Shan: moshan@eng.ucsd.edu

Arash Asgharivaskasi: aasghari@eng.ucsd.edu

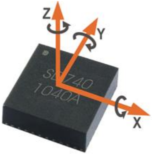
UC San Diego

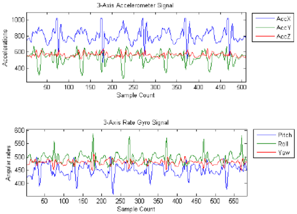
JACOBS SCHOOL OF ENGINEERING

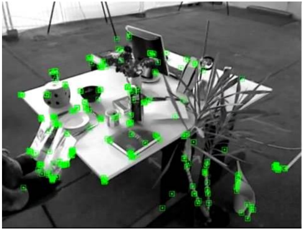
Electrical and Computer Engineering

Visual-Inertial Localization and Mapping

- **Input**:
  - IMU: linear acceleration $\mathbf{a}_t \in \mathbb{R}^3$ and rotational velocity $\boldsymbol{\omega}_t \in \mathbb{R}^3$
  - Camera: features $\mathbf{z}_{t,i} \in \mathbb{R}^4$ (left and right image pixels) for $i = 1, \ldots, N_t$







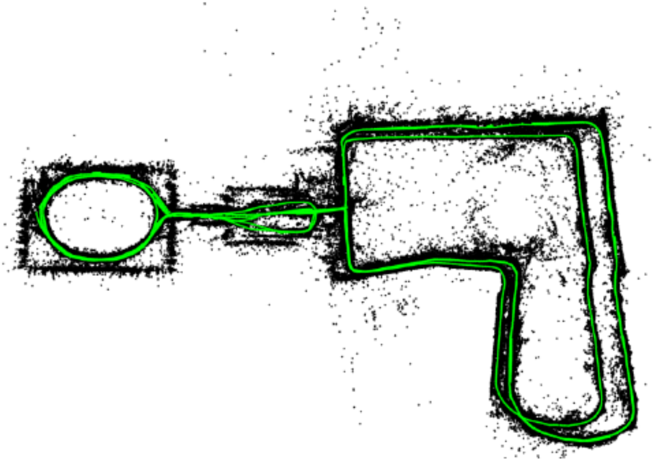
- **Assumption**: The transformation ${}_O T_I \in SE(3)$ from the IMU to the camera optical frame (extrinsic parameters) and the stereo camera calibration matrix M (intrinsic parameters) are known.

$$M := \begin{bmatrix} fs_u & 0 & c_u & 0 \\ 0 & fs_v & c_v & 0 \\ fs_u & 0 & c_u & -fs_u b \\ 0 & fs_v & c_v & 0 \end{bmatrix}$$

$$\begin{aligned} f &= \text{focal length } [m] \\ s_u, s_v &= \text{pixel scaling } [pixels/m] \\ c_u, c_v &= \text{principal point } [pixels] \\ b &= \text{stereo baseline } [m] \end{aligned}$$

Visual-Inertial Localization and Mapping

- **Output**:
  - World-frame IMU pose ${}_W T_I \in SE(3)$ over time (green)
  - World-frame coordinates $\mathbf{m}_j \in \mathbb{R}^3$ of the $j = 1, \ldots, M$ point landmarks (black) that generated the visual features $\mathbf{z}_{t,i} \in \mathbb{R}^4$



Visual Mapping

- Consider the mapping-only problem first
- **Assumption**: the IMU pose $T_t := {}_W T_{I,t} \in SE(3)$ is known
- **Objective**: given the observations $\mathbf{z}_t := \begin{bmatrix} \mathbf{z}_{t,1}^\top & \cdots & \mathbf{z}_{t,N_t}^\top \end{bmatrix}^\top \in \mathbb{R}^{4N_t}$ for $t = 0, \ldots, T$, estimate the coordinates $\mathbf{m} := \begin{bmatrix} \mathbf{m}_1^\top & \cdots & \mathbf{m}_M^\top \end{bmatrix}^\top \in \mathbb{R}^{3M}$ of the landmarks that generated them
- **Assumption**: the data association $\Delta_t : \{1, \ldots, M\} \to \{1, \ldots, N_t\}$ stipulating that landmark j corresponds to observation $\mathbf{z}_{t,i} \in \mathbb{R}^4$ with $i = \Delta_t(j)$ at time t is known or provided by an external algorithm
- **Assumption**: the landmarks $\mathbf{m}_i$ are static, i.e., it is not necessary to consider a motion model or a prediction step

Visual Mapping via the EKF

- **Observation Model**: with measurement noise $\mathbf{v}_{t,i} \sim \mathcal{N}(0, V)$

$$\mathbf{z}_{t,i} = h(T_t, \mathbf{m}_j) + \mathbf{v}_{t,i} := M\pi\left({}_O T_I T_t^{-1} \underline{\mathbf{m}}_j\right) + \mathbf{v}_{t,i}$$

- Homogeneous coordinates: $\underline{\mathbf{m}}_j := \begin{bmatrix} \mathbf{m}_j \\ 1 \end{bmatrix}$

- Projection function and its derivative:

$$\pi(\mathbf{q}) := \frac{1}{q_3}\mathbf{q} \in \mathbb{R}^4 \qquad \frac{d\pi}{d\mathbf{q}}(\mathbf{q}) = \frac{1}{q_3}\begin{bmatrix} 1 & 0 & -\frac{q_1}{q_3} & 0 \\ 0 & 1 & -\frac{q_2}{q_3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{q_4}{q_3} & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

- All observations, stacked as a $4N_t$ vector, at time t with notation abuse:

$$\mathbf{z}_t = M\pi\left({}_O T_I T_t^{-1} \underline{\mathbf{m}}\right) + \mathbf{v}_t \quad \mathbf{v}_t \sim \mathcal{N}\left(\mathbf{0}, I \otimes V\right) \quad I \otimes V := \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix}$$

Visual Mapping via the EKF

- **Prior**: $\mathbf{m} \mid \mathbf{z}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)$ with $\boldsymbol{\mu}_t \in \mathbb{R}^{3M}$ and $\Sigma_t \in \mathbb{R}^{3M \times 3M}$
- **EKF Update**: given a new observation $\mathbf{z}_{t+1} \in \mathbb{R}^{4N_{t+1}}$:

$$K_{t+1} = \Sigma_t H_{t+1}^\top \left(H_{t+1} \Sigma_t H_{t+1}^\top + I \otimes V \right)^{-1}$$

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + K_{t+1}\left(\mathbf{z}_{t+1} - \underbrace{M\pi\left({}_O T_I T_{t+1}^{-1} \underline{\boldsymbol{\mu}}_t\right)}_{\tilde{\mathbf{z}}_{t+1}}\right)$$

$$\Sigma_{t+1} = (I - K_{t+1} H_{t+1})\Sigma_t$$

- $\tilde{\mathbf{z}}_{t+1} \in \mathbb{R}^{4N_{t+1}}$ is the predicted observation based on the landmark position estimates $\boldsymbol{\mu}_t$ at time t
- We need the observation model Jacobian $H_{t+1} \in \mathbb{R}^{4N_t \times 3M}$ evaluated at $\boldsymbol{\mu}_t$ with block elements $H_{t+1,i,j} \in \mathbb{R}^{4\times 3}$:

$$H_{t+1,i,j} := \begin{cases} \frac{\partial}{\partial \mathbf{m}_j} h(T_{t+1}, \mathbf{m}_j)\Big|_{\mathbf{m}_j = \boldsymbol{\mu}_{t,j}}, & \text{if } \Delta_t(j) = i, \\ \mathbf{0}, & \text{otherwise.} \end{cases}$$

Stereo Camera Jacobian

- Consider a perturbation $\delta\boldsymbol{\mu}_{t,j} \in \mathbb{R}^3$ for the position of landmark j:

$$\mathbf{m}_j = \boldsymbol{\mu}_{t,j} + \delta\boldsymbol{\mu}_{t,j}$$

- Projection Matrix: $P = \begin{bmatrix} I & 0 \end{bmatrix} \in \mathbb{R}^{3\times 4}$ such that $\mathbf{m}_j = P\underline{\mathbf{m}}_j$

- The first-order Taylor series approximation to observation i at time t using the perturbation $\delta\boldsymbol{\mu}_{t,j}$ is:

$$\begin{aligned}
\mathbf{z}_{t+1,i} &= M\pi\left({}_O T_I \, T_{t+1}^{-1}\underline{(\boldsymbol{\mu}_{t,j} + \delta\boldsymbol{\mu}_{t,j})}\right) + \mathbf{v}_{t+1,i} \\
&= M\pi\left({}_O T_I \, T_{t+1}^{-1}\left(\underline{\boldsymbol{\mu}}_{t,j} + P^\top \delta\boldsymbol{\mu}_{t,j}\right)\right) + \mathbf{v}_{t+1,i} \\
&\approx \underbrace{M\pi\left({}_O T_I \, T_{t+1}^{-1}\underline{\boldsymbol{\mu}}_{t,j}\right)}_{\tilde{\mathbf{z}}_{t+1,i}} + \underbrace{M\frac{d\pi}{d\mathbf{q}}\left({}_O T_I \, T_{t+1}^{-1}\underline{\boldsymbol{\mu}}_{t,j}\right) {}_O T_I \, T_{t+1}^{-1} P^\top}_{H_{t+1,i,j}} \delta\boldsymbol{\mu}_{t,j} + \mathbf{v}_{t+1,i}
\end{aligned}$$

Visual Mapping via the EKF (Summary)

- Prior: $\boldsymbol{\mu}_t \in \mathbb{R}^{3M}$ and $\Sigma_t \in \mathbb{R}^{3M \times 3M}$
- Known: calibration matrix M, extrinsics ${}_O T_I \in SE(3)$, IMU pose $T_{t+1} \in SE(3)$, new observation $\mathbf{z}_{t+1} \in \mathbb{R}^{4N_{t+1}}$
- Predicted observations based on $\boldsymbol{\mu}_t$ and known correspondences Δ_{t+1}:

$$\tilde{\mathbf{z}}_{t+1,i} := M\pi\left({}_O T_I T_{t+1}^{-1} \underline{\boldsymbol{\mu}}_{t,j}\right) \in \mathbb{R}^4 \qquad \text{for } i = 1, \ldots, N_{t+1}$$

- Jacobian of $\tilde{\mathbf{z}}_{t+1,i}$ with respect to $\mathbf{m}_j$ evaluated at $\boldsymbol{\mu}_{t,j}$:

$$H_{t+1,i,j} = \begin{cases} M \frac{d\pi}{d\mathbf{q}}\left({}_O T_I T_{t+1}^{-1} \underline{\boldsymbol{\mu}}_{t,j}\right) {}_O T_I T_{t+1}^{-1} P^\top & \text{if } \Delta_t(j) = i, \\ \mathbf{0}, \in \mathbb{R}^{4\times 3} & \text{otherwise} \end{cases}$$

- EKF update:

$$K_{t+1} = \Sigma_t H_{t+1}^\top \left(H_{t+1}\Sigma_t H_{t+1}^\top + I \otimes V\right)^{-1}$$
$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + K_{t+1}\left(\mathbf{z}_{t+1} - \tilde{\mathbf{z}}_{t+1}\right)$$
$$\Sigma_{t+1} = (I - K_{t+1}H_{t+1})\Sigma_t$$

$$I \otimes V := \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix}$$

Lie Group Probability and Statistics

- The elements of matrix Lie groups do not satisfy some basic operations that we normally take for granted
- We need a different way to define random variables because matrix Lie groups are not closed under the usual addition operation:

$$\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\epsilon} \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \Sigma)$$

- **Idea**: define random variables over the Lie algebra, exploiting its vector space characteristics:

	perturbation	distribution
$SO(3)$	$R = \exp(\hat{\epsilon})\boldsymbol{\mu}$	$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$
$\mathfrak{so}(3)$	$\boldsymbol{\theta} \approx \log(\boldsymbol{\mu})^\vee + J_L^{-1}(\log(\boldsymbol{\mu})^\vee)\boldsymbol{\epsilon}$	$R = \exp(\hat{\boldsymbol{\theta}})$
$SE(3)$	$T = \exp(\hat{\epsilon})\boldsymbol{\mu}$	$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$
$\mathfrak{se}(3)$	$\boldsymbol{\xi} \approx \log(\boldsymbol{\mu})^\vee + \mathcal{J}_L^{-1}(\log(\boldsymbol{\mu})^\vee)\boldsymbol{\epsilon}$	$T = \exp(\hat{\boldsymbol{\xi}})$

Lie Group Probability and Statistics

- $SO(3)$ **and** $SE(3)$ **Random Variables**:

$$R = \exp(\hat{\boldsymbol{\epsilon}})\boldsymbol{\mu} \qquad T = \exp(\hat{\boldsymbol{\epsilon}})\boldsymbol{\mu} \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$$

  where $\boldsymbol{\mu}$ is a 'large' noise-free nominal rotation/pose and $\boldsymbol{\epsilon}$ is a 'small' noisy component in $\mathbb{R}^3$ or $\mathbb{R}^6$

- Note that $\boldsymbol{\epsilon} = \log\left(R\boldsymbol{\mu}^\top\right)^\vee$ and $\boldsymbol{\epsilon} = \log\left(T\boldsymbol{\mu}^{-1}\right)^\vee$

- Assuming $\boldsymbol{\epsilon}$ has most of its mass on $\|\boldsymbol{\epsilon}\| < \pi$, the pdf of R can be obtained using **Change of Density** with $dR = |det(J_L(\log(\boldsymbol{\mu})^\vee))|d\boldsymbol{\epsilon}$:

$$p(R) = \frac{1}{\sqrt{(2\pi)^3 \det(\Sigma)}} \exp\left(-\frac{1}{2}\left(\log\left(R\boldsymbol{\mu}^\top\right)^\vee\right)^\top \Sigma^{-1} \log\left(R\boldsymbol{\mu}^\top\right)^\vee\right) \frac{1}{|det(J_L(\log(\boldsymbol{\mu})^\vee))|}$$

- The choice of $\boldsymbol{\mu}$ and Σ as the mean and variance of R are justified:

$$\int \log\left(R\boldsymbol{\mu}^\top\right)^\vee p(R)dR = 0$$

$$\int \log\left(R\boldsymbol{\mu}^\top\right)^\vee \left(\log\left(R\boldsymbol{\mu}^\top\right)^\vee\right)^\top p(R)dR = \mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top] = \Sigma$$

Example: Rotation of a Random Rotation Variable

- Let $Q \in SO(3)$ and $\boldsymbol{\theta} \in \mathbb{R}^3$. Then:

$$Q \exp(\hat{\boldsymbol{\theta}}) Q^\top = \exp\left(Q \hat{\boldsymbol{\theta}} Q^\top \right) = \exp\left((Q\boldsymbol{\theta})^\wedge \right)$$

- Let $R \in SO(3)$ be a random rotation with mean $\boldsymbol{\mu} \in SO(3)$ and covariance $\Sigma \in \mathbb{R}^{3\times 3}$.

- The random variable $Y = QR \in SO(3)$ satisfies:

$$\begin{aligned} Y &= QR = Q\exp(\hat{\boldsymbol{\epsilon}})\boldsymbol{\mu} = \exp\left((Q\boldsymbol{\epsilon})^\wedge\right) Q\boldsymbol{\mu} \\ \mathbb{E}[Y] &= Q\boldsymbol{\mu} \\ \mathbf{Var}[Y] &= \mathbf{Var}[Q\boldsymbol{\epsilon}] = Q\Sigma Q^\top \end{aligned}$$

Visual-Inertial Odometry

- Now, consider the localization-only problem
- We will simplify the prediction step by using kinematic rather than dynamic equations
- **Assumption**: linear velocity $\mathbf{v}_t \in \mathbb{R}^3$ instead of linear acceleration $\mathbf{a}_t \in \mathbb{R}^3$ measurements are available
- **Assumption**: known world-frame landmark coordinates $\mathbf{m} \in \mathbb{R}^{3M}$
- **Assumption**: the data association $\Delta_t : \{1, \ldots, M\} \to \{1, \ldots, N_t\}$ stipulating that landmark j corresponds to observation $\mathbf{z}_{t,i} \in \mathbb{R}^4$ with $i = \Delta_t(j)$ at time t is known or provided by an external algorithm
- **Objective**: given IMU measurements $\mathbf{u}_{0:T}$ with $\mathbf{u}_t := [\mathbf{v}_t^\top, \boldsymbol{\omega}_t^\top]^\top \in \mathbb{R}^6$ and feature observations $\mathbf{z}_{0:T}$, estimate the pose $T_t := {}_W T_{I,t} \in SE(3)$ of the IMU over time

Pose Kinematics with Perturbation

- **Motion Model** for the continuous-time IMU pose $T(t)$ with noise $\mathbf{w}(t)$:

$$\dot{T} = T\left(\hat{\mathbf{u}} + \hat{\mathbf{w}}\right) \qquad\qquad \mathbf{u}(t) := \begin{bmatrix} \mathbf{v}(t) \\ \boldsymbol{\omega}(t) \end{bmatrix} \in \mathbb{R}^6$$

- To consider a Gaussian distribution over T, express it as a nominal pose $\boldsymbol{\mu} \in SE(3)$ with small perturbation $\hat{\delta\boldsymbol{\mu}} \in \mathfrak{se}(3)$:

$$T = \boldsymbol{\mu} \exp(\hat{\delta\boldsymbol{\mu}}) \approx \boldsymbol{\mu}\left(I + \hat{\delta\boldsymbol{\mu}}\right)$$

- Substitute the nominal + perturbed pose in the kinematic equations:

$$\dot{\boldsymbol{\mu}}\left(I + \hat{\delta\boldsymbol{\mu}}\right) + \boldsymbol{\mu}\left(\dot{\hat{\delta\boldsymbol{\mu}}}\right) = \boldsymbol{\mu}\left(I + \hat{\delta\boldsymbol{\mu}}\right)\left(\hat{\mathbf{u}} + \hat{\mathbf{w}}\right)$$

$$\dot{\boldsymbol{\mu}} + \dot{\boldsymbol{\mu}}\hat{\delta\boldsymbol{\mu}} + \boldsymbol{\mu}\left(\dot{\hat{\delta\boldsymbol{\mu}}}\right) = \boldsymbol{\mu}\hat{\mathbf{u}} + \boldsymbol{\mu}\hat{\mathbf{w}} + \boldsymbol{\mu}\hat{\delta\boldsymbol{\mu}}\hat{\mathbf{u}} + \cancelto{0}{\boldsymbol{\mu}\hat{\delta\boldsymbol{\mu}}\hat{\mathbf{w}}}$$

$$\dot{\boldsymbol{\mu}} = \boldsymbol{\mu}\hat{\mathbf{u}} \qquad \boldsymbol{\mu}\hat{\mathbf{u}}\hat{\delta\boldsymbol{\mu}} + \boldsymbol{\mu}\left(\dot{\hat{\delta\boldsymbol{\mu}}}\right) = \boldsymbol{\mu}\hat{\mathbf{w}} + \boldsymbol{\mu}\hat{\delta\boldsymbol{\mu}}\hat{\mathbf{u}}$$

$$\dot{\boldsymbol{\mu}} = \boldsymbol{\mu}\hat{\mathbf{u}} \qquad \dot{\hat{\delta\boldsymbol{\mu}}} = \hat{\delta\boldsymbol{\mu}}\hat{\mathbf{u}} - \hat{\mathbf{u}}\hat{\delta\boldsymbol{\mu}} + \hat{\mathbf{w}} = \left(-\overset{\curlywedge}{\mathbf{u}}\delta\boldsymbol{\mu}\right)^{\wedge} + \hat{\mathbf{w}}$$

Pose Kinematics with Perturbation

- Using $T = \boldsymbol{\mu} \exp(\delta\hat{\boldsymbol{\mu}}) \approx \boldsymbol{\mu}\left(I + \delta\hat{\boldsymbol{\mu}}\right)$, the pose kinematics $\dot{T} = T\left(\hat{\mathbf{u}} + \hat{\mathbf{w}}\right)$ can be split into nominal and perturbation kinematics:

$$\begin{aligned} \text{nominal}: \quad & \dot{\boldsymbol{\mu}} = \boldsymbol{\mu}\hat{\mathbf{u}} \\ \text{perturbation}: \quad & \delta\dot{\boldsymbol{\mu}} = -\overset{\curlywedge}{\mathbf{u}}\delta\boldsymbol{\mu} + \mathbf{w} \end{aligned} \qquad \overset{\curlywedge}{\mathbf{u}} := \begin{bmatrix} \hat{\boldsymbol{\omega}} & \hat{\mathbf{v}} \\ 0 & \hat{\boldsymbol{\omega}} \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

- In discrete-time with discretization τ, the above becomes:

$$\begin{aligned} \text{nominal}: \quad & \boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t \exp\left(\tau\hat{\mathbf{u}}_t\right) \\ \text{perturbation}: \quad & \delta\boldsymbol{\mu}_{t+1} = \exp\left(-\tau\overset{\curlywedge}{\mathbf{u}}_t\right)\delta\boldsymbol{\mu}_t + \mathbf{w}_t \end{aligned}$$

- This is useful to separate the effect of the noise $\mathbf{w}_t$ from the motion of the deterministic part of T_t. See Barfoot Ch. 7.2 for details.

EKF Prediction Step

- **Prior**: $T_t | \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$ with $\boldsymbol{\mu}_{t|t} \in SE(3)$ and $\Sigma_{t|t} \in \mathbb{R}^{6\times 6}$
- This means that $T_t = \boldsymbol{\mu}_{t|t} \exp(\widehat{\delta\boldsymbol{\mu}}_{t|t})$ with $\delta\boldsymbol{\mu}_{t|t} \sim \mathcal{N}(0, \Sigma_{t|t})$
- $\Sigma_{t|t}$ is 6×6 because only the 6 degrees of freedom of T_t are changing
- **Motion Model**: nominal kinematics of $\boldsymbol{\mu}_{t|t}$ and perturbation kinematics of $\delta\boldsymbol{\mu}_{t|t}$ with time discretization τ:

$$\boldsymbol{\mu}_{t+1|t} = \boldsymbol{\mu}_{t|t} \exp\left(\tau \hat{\mathbf{u}}_t\right)$$
$$\delta\boldsymbol{\mu}_{t+1|t} = \exp\left(-\tau \overset{\curlywedge}{\mathbf{u}}_t\right) \delta\boldsymbol{\mu}_{t|t} + \mathbf{w}_t$$

- **EKF Prediction Step** with $\mathbf{w}_t \sim \mathcal{N}(0, W)$:

$$\boldsymbol{\mu}_{t+1|t} = \boldsymbol{\mu}_{t|t} \exp\left(\tau \hat{\mathbf{u}}_t\right)$$
$$\Sigma_{t+1|t} = \mathbb{E}[\delta\boldsymbol{\mu}_{t+1|t}\delta\boldsymbol{\mu}_{t+1|t}^\top] = \exp\left(-\tau \overset{\curlywedge}{\mathbf{u}}_t\right) \Sigma_{t|t} \exp\left(-\tau \overset{\curlywedge}{\mathbf{u}}_t\right)^\top + W$$

where

$$\mathbf{u}_t := \begin{bmatrix} \mathbf{v}_t \\ \boldsymbol{\omega}_t \end{bmatrix} \in \mathbb{R}^6 \quad \hat{\mathbf{u}}_t := \begin{bmatrix} \hat{\boldsymbol{\omega}}_t & \mathbf{v}_t \\ \mathbf{0}^\top & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4} \quad \overset{\curlywedge}{\mathbf{u}}_t := \begin{bmatrix} \hat{\boldsymbol{\omega}}_t & \hat{\mathbf{v}}_t \\ 0 & \hat{\boldsymbol{\omega}}_t \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

EKF Update Step

- **Prior**: $T_{t+1} | z_{0:t}, u_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$ with $\boldsymbol{\mu}_{t+1|t} \in SE(3)$ and $\Sigma_{t+1|t} \in \mathbb{R}^{6\times 6}$
- **Observation Model**: with measurement noise $\mathbf{v}_t \sim \mathcal{N}(0, V)$

$$\mathbf{z}_{t+1,i} = h(T_{t+1}, \mathbf{m}_j) + \mathbf{v}_{t+1,i} := M\pi\left({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j\right) + \mathbf{v}_{t+1,i}$$

- The observation model is the same as in the visual mapping problem but this time the variable of interest is the IMU pose $T_{t+1} \in SE(3)$ instead of the landmark positions $\mathbf{m} \in \mathbb{R}^{3M}$
- We need the observation model Jacobian $H_{t+1} \in \mathbb{R}^{4N_{t+1}\times 6}$ with respect to the IMU pose T_{t+1}, evaluated at $\boldsymbol{\mu}_{t+1|t}$

EKF Update Step

- Let the elements of $H_{t+1} \in \mathbb{R}^{4N_{t+1}\times 6}$ corresponding to different observations i be $H_{t+1,i} \in \mathbb{R}^{4\times 6}$
- The first-order Taylor series approximation of observation i at time $t+1$ using an IMU pose perturbation $\delta\boldsymbol{\mu}$ is:

$$\begin{aligned}
\mathbf{z}_{t+1,i} &= M\pi\left({}_OT_I\left(\boldsymbol{\mu}_{t+1|t}\exp\left(\delta\hat{\boldsymbol{\mu}}\right)\right)^{-1}\underline{\mathbf{m}}_j\right) + \mathbf{v}_{t+1,i}\\
&\approx M\pi\left({}_OT_I\left(I - \delta\hat{\boldsymbol{\mu}}\right)\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right) + \mathbf{v}_{t+1,i}\\
&= M\pi\left({}_OT_I\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j - {}_OT_I\left(\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right)^{\odot}\delta\boldsymbol{\mu}\right) + \mathbf{v}_{t+1,i}\\
&\approx \underbrace{M\pi\left({}_OT_I\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right)}_{\tilde{\mathbf{z}}_{t+1,i}} \underbrace{-M\frac{d\pi}{d\mathbf{q}}\left({}_OT_I\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right){}_OT_I\left(\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right)^{\odot}}_{H_{t+1,i}}\delta\boldsymbol{\mu} + \mathbf{v}_{t+1,i}
\end{aligned}$$

where for homogeneous coordinates $\underline{\mathbf{s}} \in \mathbb{R}^4$ and $\hat{\boldsymbol{\xi}} \in \mathfrak{se}(3)$:

$$\hat{\boldsymbol{\xi}}\underline{\mathbf{s}} = \underline{\mathbf{s}}^{\odot}\boldsymbol{\xi} \qquad \begin{bmatrix}\mathbf{s}\\1\end{bmatrix}^{\odot} := \begin{bmatrix}I & -\hat{\mathbf{s}}\\0 & 0\end{bmatrix} \in \mathbb{R}^{4\times 6}$$

EKF Update Step

- **Prior**: $\boldsymbol{\mu}_{t+1|t} \in SE(3)$ and $\Sigma_{t+1|t} \in \mathbb{R}^{6\times 6}$
- Known: calibration matrix M, extrinsics ${}_O T_I \in SE(3)$, landmark positions $\mathbf{m} \in \mathbb{R}^{3M}$, new observations $\mathbf{z}_{t+1} \in \mathbb{R}^{4N_{t+1}}$
- Predicted observation based on $\boldsymbol{\mu}_{t+1|t}$ and known correspondences Δ_t:

$$\tilde{\mathbf{z}}_{t+1,i} := M\pi\left({}_O T_I \boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right) \qquad \text{for } i = 1, \ldots, N_{t+1}$$

- Jacobian of $\tilde{\mathbf{z}}_{t+1,i}$ with respect to T_{t+1} evaluated at $\boldsymbol{\mu}_{t+1|t}$:

$$H_{t+1,i} = -M\frac{d\pi}{d\mathbf{q}}\left({}_O T_I \boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right) {}_O T_I \left(\boldsymbol{\mu}_{t+1|t}^{-1}\underline{\mathbf{m}}_j\right)^{\odot} \in \mathbb{R}^{4\times 6}$$

- Perform the EKF update:

$$K_{t+1} = \Sigma_{t+1|t} H_{t+1}^\top \left(H_{t+1}\Sigma_{t+1|t}H_{t+1}^\top + I \otimes V\right)^{-1}$$

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} \exp\left(\left(K_{t+1}(\mathbf{z}_{t+1} - \tilde{\mathbf{z}}_{t+1})\right)^{\wedge}\right)$$

$$\Sigma_{t+1|t+1} = (I - K_{t+1}H_{t+1})\Sigma_{t+1|t}$$

$$H_{t+1} = \begin{bmatrix} H_{t+1,1} \\ \vdots \\ H_{t+1,N_{t+1}} \end{bmatrix}$$