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

Instructor:

Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:

Qiaojun Feng: qif007@eng.ucsd.edu

Arash Asgharivaskasi: aasghari@eng.ucsd.edu

Thai Duong: tduong@eng.ucsd.edu

Yiran Xu: y5xu@eng.ucsd.edu

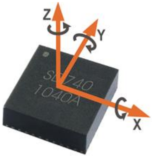
UC San Diego

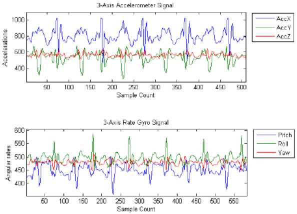
JACOBS SCHOOL OF ENGINEERING

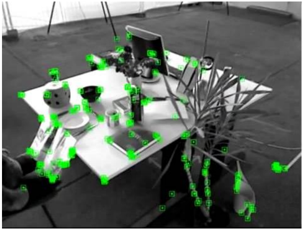
Electrical and Computer Engineering

Visual-Inertial Localization and Mapping

- **Input**:
  - IMU data: linear accleration $\mathbf{a}_t \in \mathbb{R}^3$ and rotational velocity $\boldsymbol{\omega}_t \in \mathbb{R}^3$
  - Camera data: visual features $\mathbf{z}_t \in \mathbb{R}^{4 \times N_t}$ (left and right image pixels)







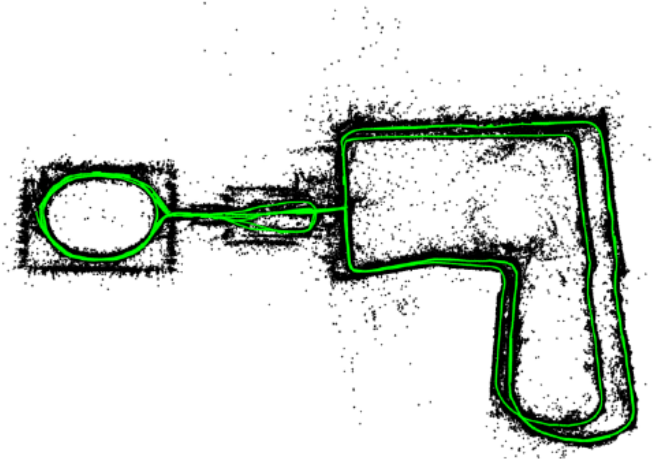
- **Assumption**: The transformation ${}_O T_I \in SE(3)$ from the IMU to the camera optical frame (extrinsic parameters) and the stereo camera calibration matrix M (intrinsic parameters) are known.

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

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

Visual-Inertial Localization and Mapping

- **Output**:
  - IMU pose ${}_W T_I \in SE(3)$ in the world frame over time (green)
  - World-frame coordinates of the point landmarks $\mathbf{m} \in \mathbb{R}^{3\times M}$ that generated the visual features $\mathbf{z}_t$ (black)



Visual Mapping

- Consider the mapping-only problem first
- **Assumption**: the inverse IMU pose $U_t := {}_W T_{I,t}^{-1} \in SE(3)$ is known
- **Objective**: given the visual feature observations $\mathbf{z}_{0:T}$, estimate the homogeneous coordinates $\underline{\mathbf{m}} \in \mathbb{R}^{4 \times M}$ in the world frame of the landmarks that generated the visual observations
- **Homogeneous coordinates**: $\underline{\mathbf{m}} := \begin{bmatrix} \mathbf{m} \\ 1 \end{bmatrix}$
- **Assumption**: the data association $\pi_t : \{1, \ldots, M\} \to \{1, \ldots, N_t\}$ stipulating which landmarks were observed at each time t is known or provided by an external algorithm
- **Assumption**: the landmarks are static, i.e., it is not necessary to consider a motion model or a prediction step

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 $\Sigma_t \in \mathbb{R}^{3M \times 3M}$
- **Observation Model**: with measurement noise $\mathbf{v}_{t,i} \sim \mathcal{N}(0, V)$

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

- 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:

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

Visual Mapping via the EKF

- **EKF Update**:

$$K_t = \Sigma_t H_t^\top \left(H_t \Sigma_t H_t^\top + I \otimes V \right)^{-1}$$

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

$$\Sigma_{t+1} = (I - K_t H_t)\Sigma_t$$

- $\tilde{\mathbf{z}}_t$ is the predicted observation based on the landmark position estimates $\boldsymbol{\mu}_t$ at time t
- We need the observation model Jacobian $H_t \in \mathbb{R}^{4N_t \times 3M}$ evaluated at $\boldsymbol{\mu}_t$
- Let the elements of $H_t \in \mathbb{R}^{4N_t \times 3M}$ corresponding to different observations i and different landmarks j be $H_{t,i,j} \in \mathbb{R}^{4 \times 3}$

Stereo Camera Jacobian

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

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

- **Projection Matrix**: $P = \begin{bmatrix} I & 0 \end{bmatrix}$

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

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

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 ${}_O T_I \in SE(3)$, inverse IMU pose $U_t \in SE(3)$, projection matrix P, new observation $\mathbf{z}_t \in \mathbb{R}^{4 \times N_t}$
- Predicted observations based on $\boldsymbol{\mu}_t$ and known correspondences π_t:

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

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

$$H_{t,i,j} = \begin{cases} M \frac{d\pi}{d\mathbf{q}}\left({}_O T_I U_t \underline{\boldsymbol{\mu}}_{t,j}\right) {}_O T_I U_t P^\top & \text{if observation } i \text{ corresponds to} \\ & \text{landmark } j \text{ at time } t \\ \mathbf{0} \in \mathbb{R}^{4\times 3} & \text{otherwise} \end{cases}$$

- Perform the EKF update:

$$K_t = \Sigma_t H_t^\top \left(H_t \Sigma_t H_t^\top + I \otimes V\right)^{-1}$$

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + K_t\left(\mathbf{z}_t - \tilde{\mathbf{z}}_t\right)$$

$$\Sigma_{t+1} = (I - K_t H_t)\Sigma_t$$

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

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

Lie Group Probability and Statistics

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

$$R = \exp(\hat{\boldsymbol{\epsilon}})\boldsymbol{\mu} \qquad T = \exp(\hat{\boldsymbol{\epsilon}})\boldsymbol{\mu}$$

where $\boldsymbol{\mu}$ is a 'large' noise-free nominal rotation/pose and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$ 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 $\|\epsilon\| < \pi$, the pdf of R can be obtained using **Change of Density** with $dR = |det(J_L(\boldsymbol{\epsilon}))|d\boldsymbol{\epsilon}$:

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

- 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**: the world-frame landmark coordinates $\mathbf{m} \in \mathbb{R}^{3\times M}$ are known
- **Assumption**: the data association $\pi_t : \{1, \ldots, M\} \to \{1, \ldots, N_t\}$ stipulating which landmarks were observed at each time t is known or provided by an external algorithm
- **Objective**: given the IMU measurements $\mathbf{u}_{0:T}$ with $\mathbf{u}_t := [\mathbf{v}_t^\top, \boldsymbol{\omega}_t^\top]^\top$ and the visual feature observations $\mathbf{z}_{0:T}$, estimate the inverse IMU pose $U_t := {}_W T_{I,t}^{-1} \in SE(3)$ over time

Visual-Inertial Odometry via the EKF

- **Prior**: $U_t | \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$ with $\boldsymbol{\mu}_{t|t} \in SE(3)$ and $\Sigma_{t|t} \in \mathbb{R}^{6\times 6}$
- The covariance is 6×6 because only the six degrees of freedom of $U_t \in SE(3)$ are changing
- **Motion Model**: with time discretization τ and noise $\mathbf{w}_t \sim \mathcal{N}(0, W)$

$$U_{t+1} = \exp\left(-\tau\left((\mathbf{u}_t + \mathbf{w}_t)\right)^\wedge\right) U_t \qquad \mathbf{u}_t := \begin{bmatrix} \mathbf{v}_t \\ \boldsymbol{\omega}_t \end{bmatrix} \in \mathbb{R}^6$$

- Note that $\mathbf{u}_t + \mathbf{w}_t$ is negative above since U_t is the inverse IMU pose
- Let the IMU pose in continuous time be ${}_W T_I(t) = T(t) = U^{-1}(t)$:

$$\dot{T} = T\hat{\mathbf{u}} \qquad TU = I \qquad \dot{T}U + T\dot{U} = 0$$
$$\dot{U} = -U\dot{T}U = -U\left(T\hat{\mathbf{u}}\right)U = -\hat{\mathbf{u}}U$$
$$U_{t+1} = \exp(-\tau\hat{\mathbf{u}}_t)U_t$$

Pose Kinematics with Perturbation

- Consider what happens with the pose kinematics

$$\dot{T} = -\left(\hat{\mathbf{u}} + \hat{\mathbf{w}}\right) T$$

if the pose is expressed as a nominal pose $\boldsymbol{\mu} \in SE(3)$ and small perturbation $\widehat{\delta\boldsymbol{\mu}} \in \mathfrak{se}(3)$:

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

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

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

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

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

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

Pose Kinematics with Perturbation

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

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

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

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

- 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

- Using the perturbation idea from the previous slide, converted to discrete time, we can re-write the motion model in terms of nominal kinematics of the mean of T_t and zero-mean perturbation kinematics:

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

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

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

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

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

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**: $U_{t+1}|z_{0:t}, u_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$ 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 $\mathbf{v}_t \sim \mathcal{N}(0, V)$

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

- The observation model is the same as in the visual mapping problem but this time the variable of interest is the inverse IMU pose $U_{t+1} \in SE(3)$ instead of the landmark positions $\mathbf{m} \in \mathbb{R}^{3\times M}$

- We need the observation model Jacobian $H_{t+1|t} \in \mathbb{R}^{4N_t \times 6}$ with respect to the inverse IMU pose U_t, evaluated at $\boldsymbol{\mu}_{t+1|t}$

EKF Update Step

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

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

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 $\mathbf{m} \in \mathbb{R}^{3\times M}$, new observation $\mathbf{z}_{t+1} \in \mathbb{R}^{4\times N_t}$
- Predicted observation based on $\boldsymbol{\mu}_{t+1|t}$ and known correspondences π_t:

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

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

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

- Perform the EKF update:

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

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

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

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