ECE276A: Sensing & Estimation in Robotics
Lecture 12: Visual-Inertial SLAM

Nikolay Atanasov

natanasov@ucsd.edu

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Outline

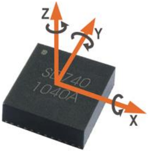
Visual-Inertial SLAM

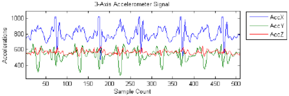
Visual Mapping

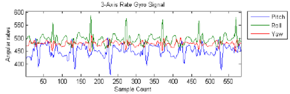
Visual-Inertial Odometry

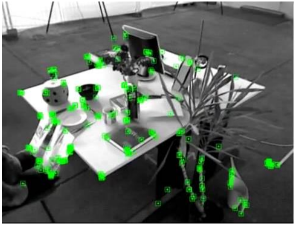
Visual-Inertial Simultaneous Localization and Mapping

- **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 $i = 1, \ldots, N_t$









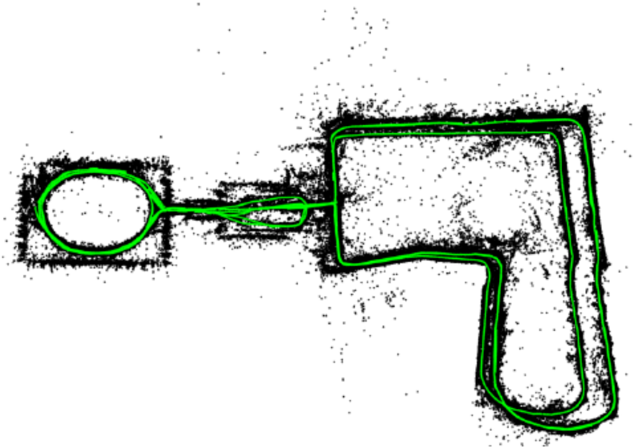
- **Assumption**: The transformation ${}_O T_I \in SE(3)$ from the IMU to the camera optical frame (extrinsic parameters) and the stereo camera calibration matrix K_s (intrinsic parameters) are known.

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

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

Visual-Inertial Simultaneous 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 $j = 1, \dots, M$ point landmarks (black) that generated the visual features $\mathbf{z}_{t,i} \in \mathbb{R}^4$



Extended Kalman Filter

Prior: $$\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$$

Motion model:
$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, W)$$
$$F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}), \quad Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$$

Observation model:
$$\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t), \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, V)$$
$$H_t := \frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0}), \quad R_t := \frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0})$$

Prediction:
$$\boldsymbol{\mu}_{t+1|t} = f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$$
$$\Sigma_{t+1|t} = F_t \Sigma_{t|t} F_t^\top + Q_t W Q_t^\top$$

Update:
$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(z_{t+1} - h(\boldsymbol{\mu}_{t+1|t}, 0))$$
$$\Sigma_{t+1|t+1} = (I - K_{t+1|t} H_{t+1}) \Sigma_{t+1|t}$$

Kalman gain: $$K_{t+1|t} := \Sigma_{t+1|t} H_{t+1}^\top \left(H_{t+1} \Sigma_{t+1|t} H_{t+1}^\top + R_{t+1} V R_{t+1}^\top \right)^{-1}$$

Outline

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 $\mathbf{m}$ are static, i.e., it is not necessary to consider a motion model or a prediction step for $\mathbf{m}$

Visual Mapping via the EKF

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

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

- 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:

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

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 step**: given a new observation $\mathbf{z}_{t+1} \in \mathbb{R}^{4N_{t+1}}$:

$$
\begin{aligned}
K_{t+1} &= \Sigma_t H_{t+1}^\top \left(H_{t+1} \Sigma_t H_{t+1}^\top + I \otimes V \right)^{-1} \\
\boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + K_{t+1} \Bigg(\mathbf{z}_{t+1} - \underbrace{K_s \pi \left({}_O T_I T_{t+1}^{-1} \underline{\boldsymbol{\mu}}_t \right)}_{\tilde{\mathbf{z}}_{t+1}} \Bigg) \\
\Sigma_{t+1} &= (I - K_{t+1} H_{t+1}) \Sigma_t
\end{aligned}
$$

- $\tilde{\mathbf{z}}_{t+1} \in \mathbb{R}^{4N_{t+1}}$ 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 (by Chain Rule)

- Observation model: $h(T_{t+1}, \mathbf{m}_j) = K_s \pi \left({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j \right)$
- How do we obtain $\frac{\partial}{\partial \mathbf{m}_j} h(T_{t+1}, \mathbf{m}_j) \Big|_{\mathbf{m}_j = \boldsymbol{\mu}_{t,j}}$?
- Let $P = \begin{bmatrix} I & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 4}$ and apply the chain rule:

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{m}_j} h(T_{t+1}, \mathbf{m}_j) &= K_s \frac{\partial \pi}{\partial \mathbf{q}} ({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j) \frac{\partial}{\partial \mathbf{m}_j} \left({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j \right) \\
&= K_s \frac{\partial \pi}{\partial \mathbf{q}} \left({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j \right) {}_O T_I T_{t+1}^{-1} \frac{\partial \underline{\mathbf{m}}_j}{\partial \mathbf{m}_j} \\
&= K_s \frac{\partial \pi}{\partial \mathbf{q}} \left({}_O T_I T_{t+1}^{-1} \underline{\mathbf{m}}_j \right) {}_O T_I T_{t+1}^{-1} P^\top
\end{aligned}$$

Stereo Camera Jacobian (by Perturbation)

- The Jacobian of a function $f(\mathbf{x})$ can also be obtained using first-order Taylor series with perturbation $\delta\mathbf{x}$:

$$f(\mathbf{x}+\delta\mathbf{x}) \approx f(\mathbf{x}) + \left[\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x})\right]\delta\mathbf{x}$$

- The Jacobian of $f(\mathbf{x})$ is the part that is linear in $\delta\mathbf{x}$ in the first-order Taylor series expansion

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

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

- First-order Taylor series approximation of the observation model:

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

Visual Mapping via the EKF (Summary)

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

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

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

$$H_{t+1,i,j} = \begin{cases} K_s \frac{d\pi}{d\mathbf{q}} \left({}_OT_I T_{t+1}^{-1} \underline{\boldsymbol{\mu}}_{t,j} \right) {}_OT_I T_{t+1}^{-1} P^\top, & \text{if } \Delta_t(j) = i, \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

- EKF update:

$$\begin{aligned} K_{t+1} &= \Sigma_t H_{t+1}^\top \left(H_{t+1} \Sigma_t H_{t+1}^\top + I \otimes V \right)^{-1} \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + K_{t+1} \left(\mathbf{z}_{t+1} - \tilde{\mathbf{z}}_{t+1} \right) \\ \Sigma_{t+1} &= (I - K_{t+1} H_{t+1}) \Sigma_t \end{aligned} \qquad I \otimes V := \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix}$$

Outline

Visual-Inertial SLAM

Visual Mapping

Visual-Inertial Odometry

Visual-Inertial Odometry

- Now, consider the localization-only problem
- We will simplify the prediction step by using kinematic rather than dynamic equations of motion for the IMU pose
- **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 IMU poses $T_t := {}_W T_{I,t} \in SE(3)$

How to Deal with an $SE(3)$ State in the EKF?

- Goal: estimate $T_t \in SE(3)$ using an extended Kalman filter
- $$SE(3) := \left\{ T = \begin{bmatrix} R & \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4} \;\middle|\; R \in SO(3), \mathbf{p} \in \mathbb{R}^3 \right\}$$
- Since T_t is not a vector, we face multiple questions:
  - How do we specify a "Gaussian" distribution over T_t?
  - What is the motion model for T_t?
  - How do we find derivatives with respect to T_t?

How Do We Specify a Gaussian Distribution in $SE(3)$?

- In $\mathbb{R}^6$, we can define a Gaussian distribution of a vector $\mathbf{x}$ as follows:

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

  where $\boldsymbol{\mu} \in \mathbb{R}^6$ is the deterministic mean and $\boldsymbol{\epsilon} \in \mathbb{R}^6$ is a zero-mean Gaussian random vector

- In $SE(3)$, we can define a Gaussian distribution of a pose matrix T using a perturbation ϵ on the Lie algebra:

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

  where $\boldsymbol{\mu} \in SE(3)$ is the deterministic mean and $\boldsymbol{\epsilon} \in \mathbb{R}^6$ is a zero-mean Gaussian random vector corresponding to the 6 degrees of freedom of T

- Example:
  - Let $T \in SE(3)$ be a random pose with mean $\boldsymbol{\mu} \in SE(3)$ and covariance $\Sigma \in \mathbb{R}^{6\times 6}$
  - For $Q \in SE(3)$, the random variable $Y = QT = Q\boldsymbol{\mu} \exp(\hat{\boldsymbol{\epsilon}})$ has mean $Q\boldsymbol{\mu} \in SE(3)$ and covariance $\Sigma \in \mathbb{R}^{6\times 6}$

What Is the Motion Model for a Pose Matrix T?

- Continuous-time kinematics of pose $T(t) \in SE(3)$ under generalized velocity $\boldsymbol{\zeta}(t) = \begin{bmatrix} \mathbf{v}(t) \\ \boldsymbol{\omega}(t) \end{bmatrix} \in \mathbb{R}^6$, expressed in body-frame coordinates:

$$\dot{T}(t) = T(t)\hat{\boldsymbol{\zeta}}(t)$$

- Discrete-time pose kinematics with **constant** $\boldsymbol{\zeta}(t)$ for $t \in [t_k, t_{k+1})$:

$$T_{k+1} = T_k \exp(\tau_k \hat{\boldsymbol{\zeta}}_k)$$

where $T_k = T(t_k)$, $\tau_k = t_{k+1} - t_k$, $\boldsymbol{\zeta}_k = \boldsymbol{\zeta}(t_k)$

How Do We Find Derivatives With Respect to a Pose T?

- In $\mathbb{R}^6$, the derivative of a function $f(\mathbf{x})$ can be obtained using first-order Taylor series with perturbation $\delta\mathbf{x} \in \mathbb{R}^6$:

$$f(\mathbf{x} + \delta\mathbf{x}) \approx f(\mathbf{x}) + \left[\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x})\right] \delta\mathbf{x}$$

- In $\mathbb{R}^6$, the derivative is $\left.\frac{\partial}{\partial \delta\mathbf{x}} f(\mathbf{x} + \delta\mathbf{x})\right|_{\delta\mathbf{x}=0}$

- In $SE(3)$, the derivative of a function $f(T)$ can be obtained using first-order Taylor series with perturbation $\delta\boldsymbol{\psi} \in \mathbb{R}^6$:

$$f(T \exp(\delta\hat{\boldsymbol{\psi}})) \approx f(T) + \left[\frac{\partial f}{\partial T}(T)\right] \delta\boldsymbol{\psi}$$

- In $SE(3)$, the derivative is $\left.\frac{\partial}{\partial \delta\boldsymbol{\psi}} f(T \exp(\delta\hat{\boldsymbol{\psi}}))\right|_{\delta\boldsymbol{\psi}=0}$

Visual-Inertial Odometry

- Now, consider the localization-only problem
- We will simplify the prediction step by using kinematic rather than dynamic equations of motion for the IMU pose
- **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 IMU poses $T_t := {}_W T_{I,t} \in SE(3)$

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 $\widehat{\delta\boldsymbol{\mu}} \in \mathfrak{se}(3)$:

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

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

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

$$\dot{\boldsymbol{\mu}} + \dot{\boldsymbol{\mu}}\widehat{\delta\boldsymbol{\mu}} + \boldsymbol{\mu}\widehat{\dot{\delta\boldsymbol{\mu}}} = \boldsymbol{\mu}\hat{\mathbf{u}} + \boldsymbol{\mu}\hat{\mathbf{w}} + \boldsymbol{\mu}\widehat{\delta\boldsymbol{\mu}}\hat{\mathbf{u}} + \cancelto{0}{\boldsymbol{\mu}\widehat{\delta\boldsymbol{\mu}}\hat{\mathbf{w}}}$$

$$\dot{\boldsymbol{\mu}} = \boldsymbol{\mu}\hat{\mathbf{u}} \qquad \boldsymbol{\mu}\hat{\mathbf{u}}\widehat{\delta\boldsymbol{\mu}} + \boldsymbol{\mu}\widehat{\dot{\delta\boldsymbol{\mu}}} = \boldsymbol{\mu}\hat{\mathbf{w}} + \boldsymbol{\mu}\widehat{\delta\boldsymbol{\mu}}\hat{\mathbf{u}}$$

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

Pose Kinematics with Perturbation

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

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

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

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

- 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 $T_t = \boldsymbol{\mu}_{t|t} \exp(\delta\hat{\boldsymbol{\mu}}_{t|t})$ with $\delta\boldsymbol{\mu}_{t|t} \sim \mathcal{N}(0, \Sigma_{t|t})$
- **Motion model**: nominal kinematics of $\boldsymbol{\mu}_{t|t}$ and perturbation kinematics of $\delta\boldsymbol{\mu}_{t|t}$ with time discretization τ_t:

$$\begin{aligned}
\boldsymbol{\mu}_{t+1|t} &= \boldsymbol{\mu}_{t|t} \exp\left(\tau_t \hat{\mathbf{u}}_t\right) \\
\delta\boldsymbol{\mu}_{t+1|t} &= \exp\left(-\tau_t \overset{\curlywedge}{\mathbf{u}}_t\right) \delta\boldsymbol{\mu}_{t|t} + \sqrt{\tau_t}\mathbf{w}_t
\end{aligned}$$

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

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

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

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

- 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 the IMU pose mean $\boldsymbol{\mu}_{t+1|t}$

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:

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

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**: Gaussian with mean $\boldsymbol{\mu}_{t+1|t} \in SE(3)$ and covariance $\Sigma_{t+1|t} \in \mathbb{R}^{6\times 6}$
- Known: stereo calibration matrix K_s, 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:

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

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

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

- EKF update step:

$$\begin{aligned} K_{t+1} &= \Sigma_{t+1|t} H_{t+1}^\top \left(H_{t+1} \Sigma_{t+1|t} H_{t+1}^\top + I \otimes V \right)^{-1} \\ \boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} \exp\left((K_{t+1}(\mathbf{z}_{t+1} - \tilde{\mathbf{z}}_{t+1}))^\wedge \right) \\ \Sigma_{t+1|t+1} &= (I - K_{t+1} H_{t+1}) \Sigma_{t+1|t} \end{aligned} \qquad H_{t+1} = \begin{bmatrix} H_{t+1,1} \\ \vdots \\ H_{t+1,N_{t+1}} \end{bmatrix}$$