Simultaneous Localization and Mapping using Extended Kalman Filter(EKF)

Hsiao-Chen Huang
Department of Electrical and Computer Engineering
University of California San Diego
hsh030@eng.ucsd.edu

***Abstract*—This project aim to solve simultaneous localization and mapping problem by using Extended Kalman Filter(EKF). Our goal is to use the EKF equations with a prediction step based on SE(3) kinematics from IMU data and an update step based on the stereo camera observation model to perform localization and mapping. In this work, we basically followed the Bayes filter probabilistic with Gaussian prior assumption for estimating the state of dynamic system. The difference between Kalman Filter ans EKF is that our motion model and observation model are not linear but affected by Gaussian noise. The results are shown in the final section.**

I. Introduction

One of the approaching in autonomous driving is building the high-definition (HD) map to help recognizing the environment while navigation. For a self-driving fleet to work efficiently, a highly-reliable infrastructure needs to exist for not just building the maps, but also updating the maps. Therefore, localization in real-time is also an important part because we must know where we are while mapping the environment information to our global map.

In this project, we are provided with synchronized measurements from an IMU and a stereo camera as well as the intrinsic camera calibration and the extrinsic calibration between the two sensors. We can separate this problem into two steps: 1) Localization: We implement the EKF prediction step to estimate the pose $T_t \in SE(3)$ of the IMU over time t using the motion model. 2) Mapping: We assume the predicted IMU trajectory from (1) above is correct and focus on estimating the landmark positions. Implement an EKF update step after every visual observation z_t with the unknown landmark positions $m \in \mathbb{R}^{3M}$ as a state, and keep track of the mean and covariance of m. The landmarks are static points in the world frame. Moreover, since the sensor does not move sufficiently along the z-axis, we assume that the z coordinates for all landmarks are 0 and focus only on estimating their xy coordinates.

II. Problem Formulation

A. IMU Localization via EKF Prediction

Here, we consider the localization-only problem using motion model to predict the pose. Given the IMU measurements $\{u_t\}_{t=0}^T$ with $u_t := [\mathbf{v}_t^T, \omega_t^T]^T$, where $\mathbf{v}_t \in \mathbb{R}^3$ is the linear velocity and $\omega_t \in \mathbb{R}^3$ is the rotation velocity in time t, we need to estimate the inverse IMU pose $T_t \in SE(3)$ overtime. Assume the inverse IMU pose is in Gaussian distribution with mean $\mu_{\mathbf{t|t}} \in SE(3)$ and covariance $\Sigma_{t|t} \in \mathbb{R}^{6\times 6}$. The covariance is 6×6 because only the six degrees of freedom of $\mu_{t|t}$ are changing via a perturbation $u_t \in \mathbb{R}^6$. We will update the mean and covariance of IMU pose via our motion model with time discretization τ and noise $\mathrm{w}_t \sim N(0, W)$:

$$\mu_{t+1|t} = \exp(-\tau\hat{u}_t)\mu_{t|t} \tag{1}$$
$$\Sigma_{t+1|t} = \exp(-\tau\tilde{u}_t)\Sigma_{t|t}\exp(-\tau\tilde{u}_t)^T + \tau^2 W \tag{2}$$

where we define

$$\hat{u}_t := \begin{bmatrix} \hat{\omega}_t & \mathbf{v}_t \\ 0 & 0 \end{bmatrix}, \tilde{u}_t := \begin{bmatrix} \hat{\omega}_t & \hat{\mathbf{v}}_t \\ 0 & \hat{\omega}_t \end{bmatrix}$$

$$\hat{\omega}_t := \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

$$\hat{\mathbf{v}}_t := \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$$

B. Landmark Mapping via EKF Update

Consider the mapping-only problem, assume the inverse IMU pose $T_t \in SE(3)$ over time is known. Given the visual feature observations $\{\mathbf{z}_t\}_{t=0}^T$ with $\mathbf{z}_t \in \mathbb{R}^{4\times N_t}$, where $z_{t,i} \in \mathbb{R}^4$ is pixel coordinates of thre observable landmark i correspondence between the left and the right camera frames, we try to estimate the homogeneous coordinates $\mathbf{m} \in \mathbb{R}^{4\times M}$ in the world frame of the landmarks. M is the number of landmarks, and N_t is the number of the landmarks that were observable at time t.

Perform EKF update step after every visual observation z_t in order to keep track of the mean and covariance of m, where the mean is $\mu_t \in \mathbb{R}^{4\times M}$ and covariance is $\Sigma_t \in \mathbb{R}^{3M\times 3M}$. First, we compute the predicted observation based on μ_t:

$$\hat{z}_{t,i} := M\pi({}_oT_iT_t\mu_{t,j}) \in \mathbb{R}^4, \qquad \text{i}= 1, ..., N_t$$

where M is the calibration matrix, and ${}_oT_i \in SE(3)$ is the extrinsics. Then, compute the Jacobian of $\hat{z}_{t,i}$ with respect to $\mathbf{m}_j$ evaluated at $\mu_{\mathbf{t},\mathbf{j}}$

$$H_{i,j,t} = \begin{cases} M\dfrac{d\pi}{d\mathbf{q}}({}_oT_iT_t\mu_{t,j}){}_oT_iT_tD, \text{i correspond to j} \\ . \\ \mathbf{0} \in \mathbb{R}^{4\times 3}, \text{otherwise} \end{cases}$$

We can update the mean and covariance of $\mathbf{m}$ by

$$K_t = \Sigma_t H_t^T(H_t\Sigma_t H_t^T + I \otimes V)^{-1} \tag{3}$$

$$\mu_{t+1} = \mu_t + DK_t(\mathbf{z}_t - \hat{\mathbf{z}}_t) \tag{4}$$

$$\Sigma_{t+1} = (I - K_tH_t)\Sigma_t \tag{5}$$

where D is the dilation matrix, and

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

III. Technical Approaches

A. IMU Localization via EKF Prediction

In the prediction part, we initialize our pose as

$$\mu_{t|t} = \begin{bmatrix} \mathbf{R} & \rho \\ 0 & 0 \end{bmatrix} \in SE(3)$$

where $\mathbf{R} = \mathbf{I}$ and $\rho = [0,0,0]^T$. Our noise w_t define with zero mean and $\mathbf{W} = \mathbf{I}$. We use the IMU data: linear velocity $\mathbf{v}_t$ and rotation velocity ω_t in each time stamp to update our pose using Eq.1.

B. Landmark Mapping via EKF Update

We initialize our landmark positions via the inverse function of observation model

$$\mathbf{z}_{t,i} := M\pi({}_oT_iT_t\mathbf{m}_j) + \mathrm{v}_t$$

where v_t is the measurement noise $\mathrm{v}_t \sim N(0, \mathbf{V})$. We use our observation measurements obtained in stereo camera, and the predicted inverse IMU pose T_t obtained in the prediction step over time, to compute the landmark positions $\mathbf{m}$. We perform the update step using Eq.4-5 to update the mean $\mu_t \in \mathbb{R}^{4\times M}$ and covariance $\Sigma_t \in \mathbb{R}^{3M\times 3M}$ of the landmark positions.

For the computation in Eq.4, we reshape $\mathbf{z}_t$ and $\hat{\mathbf{z}}_t$ to a vector with dimension $4N_t$ which concatenate every features into a vector. After multiplying with K_t, we change the homogeneous coordiantes μ_t via this perturbation for each feature. Moreover, since the sensor does not move sufficiently along the z-axis, the estimation for the z coordinate of the landmarks will not be very good. We assume that the z coordinates for all landmarks are 0, which means we set the third row of μ_t are all 0, and focus only on estimating their xy coordinates. After updating in every time stamps, we can obtain the updated landmark positions.

IV. Results

We first implement IMU localization via EKF prediction with our motion model. The results are shown in Fig1-3.

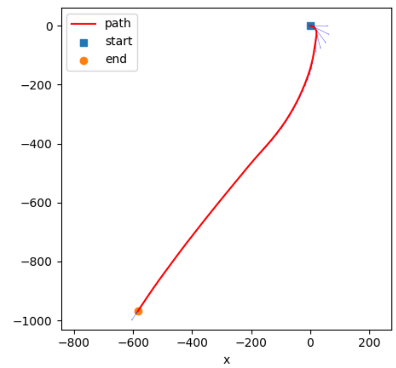

Fig. 1. The motion update of dataset42

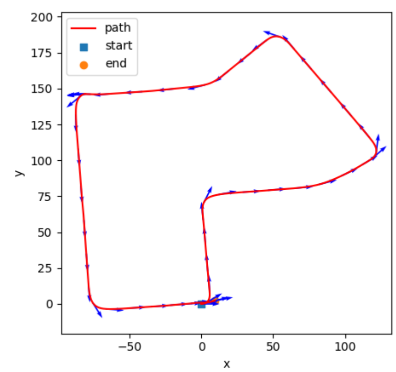

Fig. 2. The motion update of dataset27

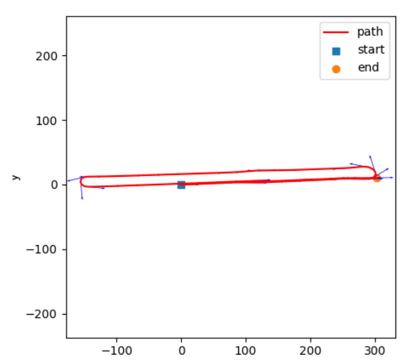

Fig. 3. The motion update of dataset20

We initialize predicted landmark positions using the predicted IMU pose with the inverse observation model. Then, update the positions with observation in each time steps. In the following, Fig.4-6 show the results in update steps

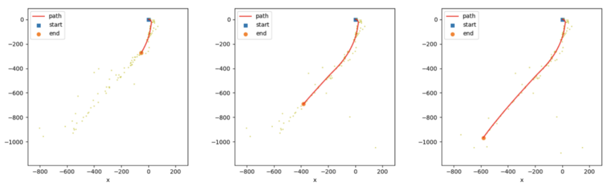

Fig. 4. The feature mapping of dataset42

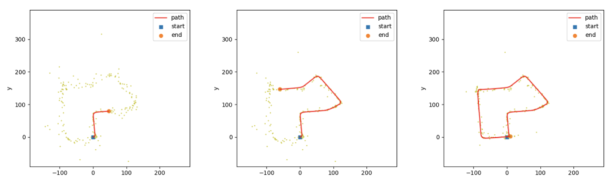

Fig. 5. The feature mapping of dataset27

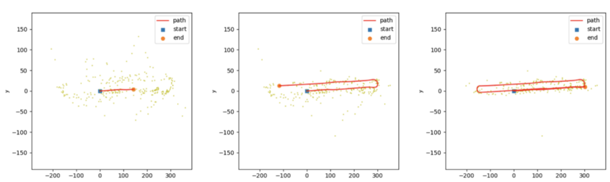

Fig. 6. The feature mapping of dataset20