Using Extended Kalman Filter for Visual-Inertial SLAM

Yushen Bai
University of California, San Diego
yub025@eng.ucsd.edu

Abstract: This report presents is to use the EKF equations with a prediction step based on SE(3) kinematics and an update step based on the stereo camera observation model to perform localization and mapping. In this work, we first implemented the EKF prediction step based on the SE(3) kinematics to estimate the pose of the IMU over time t. Then, according to the prediction, we did the landmark mapping via EKF update.

Index terms: mapping, extended Kalman filter, localization

I. Introduction

In robotic mapping and navigation, simultaneous localization and mapping (SLAM) is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it. It asks if it is possible for a mobile robot to be placed at an unknown location in an unknown environment and for the robot to incrementally build a consistent map of this environment while simultaneously determining its location within this map. Popular approximate solution methods include the particle filter, extended Kalman filter, and GraphSLAM.

In this report, the method of extended Kalman filter was used. The goal of the project is to use the EKF equations with a prediction step based on SE(3) kinematics and an update step based on the stereo camera observation model to perform localization and mapping. The mathematical method was introduced in Section II. The process of localization and mapping was described in Section III. Section IV shows the results and discussion.

II. Problem Formulations

A. Rigid Body Pose

Through an angle θ can be described by a rotation matrix R(θ):

$$s_w = \mathrm{R}(\theta) * s_B = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} * s_B$$

Where s_w is the position in the world frame, and s_B is the position in the body frame. θ, γ, φ are yaw, pitch, roll angels respectively.

Let B be a body frame whose position and orientation with respect to the world frame W are $\mathrm{p} \in R^3$ and $\mathrm{R} \in \mathrm{SO}(3)$, respectively. The rigid-body transformation is not linear but affine. It can be converted to linear by appending 1 to the coordinates of a points:

$$\begin{bmatrix} s_w \\ 1 \end{bmatrix} = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_B \\ 1 \end{bmatrix}$$

B. Lie Algebra

For θ∈so(3):

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}^{\wedge} = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

For ξ∈se(3):

$$\hat{\boldsymbol{\xi}} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}^{\wedge} = \begin{bmatrix} \hat{\boldsymbol{\theta}} & \boldsymbol{\rho} \\ \mathbf{0}^T & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

$$\overset{\curlywedge}{\boldsymbol{\xi}} = ad(\hat{\boldsymbol{\xi}}) = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}^{\curlywedge} = \begin{bmatrix} \hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\rho}} \\ \mathbf{0} & \hat{\boldsymbol{\theta}} \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

C. Localization via EKF Prediction

Assumptions: (i) the homogeneous coordinates in the world frame of the landmarks are known; (ii) the data association π_t: {1, ⋯, M} →{1, ⋯, Nt} stipulating which landmarks were observed at each time t is known or provided by an external algorithm.

Objective: given the IMU measurements $\{u_t\}_{t=0}^{T}$ with $u_t \coloneqq [v_t^T, \omega_t^T]^T$ and the visual feature observations $\{z_t\}_{t=0}^{T}$, estimate the inverse IMU pose: $T_t \coloneqq {}_wT_{i,t}^{-1}$ ∈SE(3) over time.

Assume $T_t \;/\; z_{0:t}, u_{0:t-1} \sim N\left(\mu_{t\;/\;t}, \Sigma_{\;\;t\;/\;t}\right)$ with $\mu_{t\;/\;t}$ ∈SE(3) and $\Sigma_{\;\;t\;/\;t} \in R^{6*6}$. The covariance is 6 × 6 because only the six degrees of freedom of $\mu_{t\;/\;t}$ are changing via a perturbation $\xi_t \in R^6$:

$$\boldsymbol{\mu}_{t+1|t} = \exp\left(\hat{\xi}_t\right)\boldsymbol{\mu}_{t|t} \approx (I + \hat{\xi}_t)\boldsymbol{\mu}_{t|t}$$

With time discretization τ and noise $w_t \sim N(0, W)$, the motion model is:

$$T_{t+1} = \exp\left((\tau(-u_t + w_t))^{\wedge}\right) T_t \qquad u_t := \begin{bmatrix} v_t \\ \omega_t \end{bmatrix}$$

Using the perturbation idea to separate the effect of the noise, we can re-write the motion model in terms of nominal kinematics of the mean of T_t and zero-mean perturbation kinematics:

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

The EKF prediction step:

$$\boldsymbol{\mu}_{t+1|t} = \exp\left(-\tau\hat{u}_t\right)\boldsymbol{\mu}_{t|t} \qquad u_t := \begin{bmatrix} v_t \\ \omega_t \end{bmatrix}$$
$$\Sigma_{t+1|t} = \mathbb{E}[\xi_{t+1|t}\xi_{t+1|t}^T] = \exp\left(-\tau\overset{\curlywedge}{u}_t\right)\Sigma_{t|t}\exp\left(-\tau\overset{\curlywedge}{u}_t\right)^T + \tau^2 W$$

D. Landmark Mapping via EKF Update

Assumptions: (i) the inverse IMU pose T_t over time is known; (ii) the data association π_t: {1, ⋯, M} →{1, ⋯, Nt} stipulating which landmarks were observed at each time t is known or provided by an external algorithm; (iii) the landmarks are static.

Objective: given the visual feature observations $\{z_t\}_{t=0}^{T}$ estimate the homogeneous coordinates $\mathrm{m} \in R^{4*M}$ in the world frame of the landmarks that generated the visual observations.

The calibration matrix M, extrinsics ${}_OT_I \in$ SE(3), (inverse) IMU pose $T_t \in$ SE(3), dilation matrix D, new observation $z_t \in R^{4*Nt}$ are all known. Compute the predicted observation based on μ_t and known correspondences:

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

Compute the Jacobian of it with respect to m_j evaluated at $\mu_{t,j}$:

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

Perform the EKF update:

$$K_t = \Sigma_t H_t^T \left(H_t \Sigma_t H_t^T + I \otimes V\right)^{-1} \qquad I \otimes V := \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix}$$
$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + DK_t\left(\mathbf{z}_t - \hat{\mathbf{z}}_t\right)$$
$$\Sigma_{t+1} = (I - K_t H_t)\Sigma_t$$

III. Technical Approach

The algorithms used in this project are discussed in this section.

A. Localization via EKF Prediction

Got data from IMU. According to the formula in Section II, implemented the EKF prediction step based on the SE(3) kinematics to estimate the pose of the IMU over time t. Then, drew the trajectory.

B. Landmark Mapping via EKF Update

Implemented an EKF with the unknown landmark positions as a state and perform EKF update step after every visual observation z_t in order to keep track of the mean and covariance of m.

IV. Result and Discussion

A. Set 27

Figure 1 shows the trajectory over time and the positions of landmarks of set 27. It looks great.

B. Set 42

Figure 2 shows the trajectory over time and the positions of landmarks of set 42. It looks fine.

C. Test Set

Figure 3 shows the trajectory over time and the positions of landmarks of test set. It looks well.

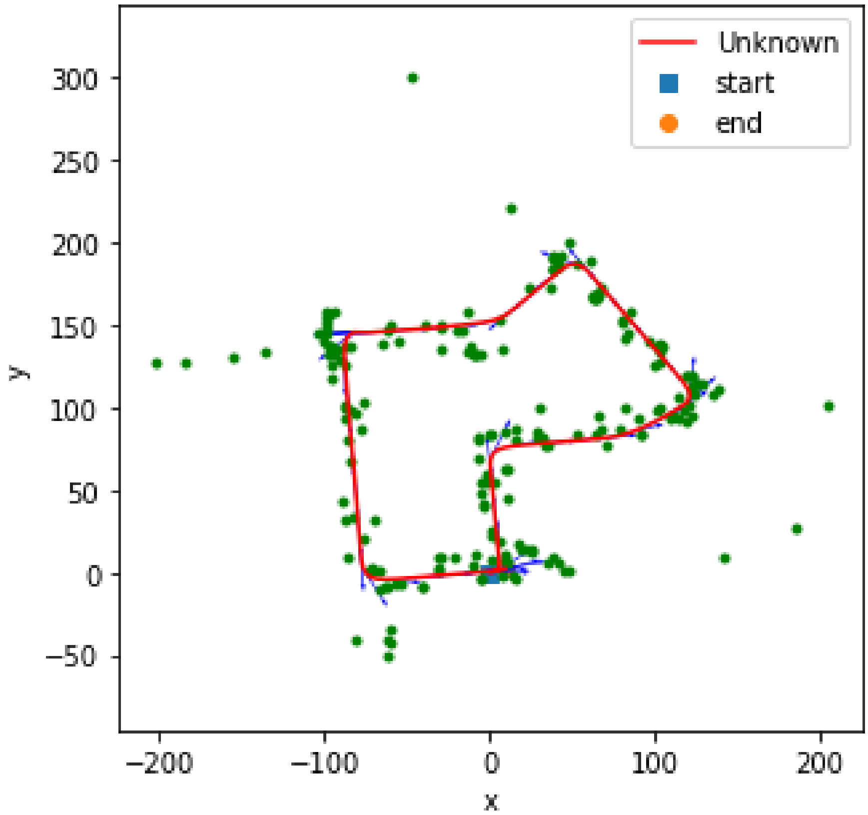

Figure1. the trajectory and the landmarks of set 27

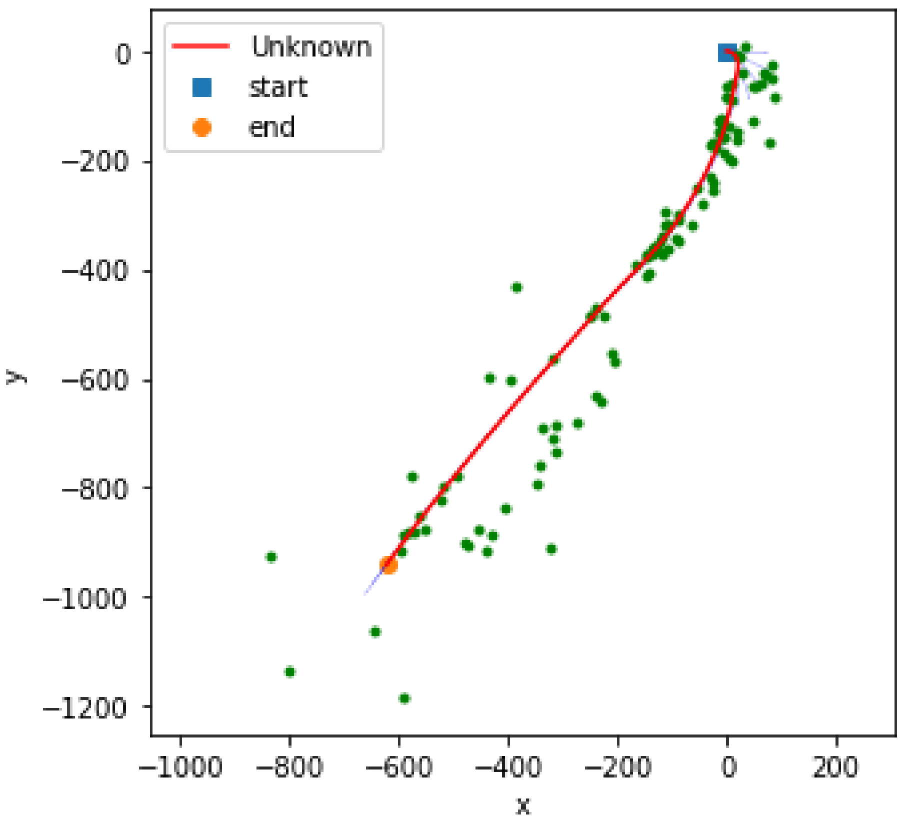

Figure2. the trajectory and the landmarks of set 42

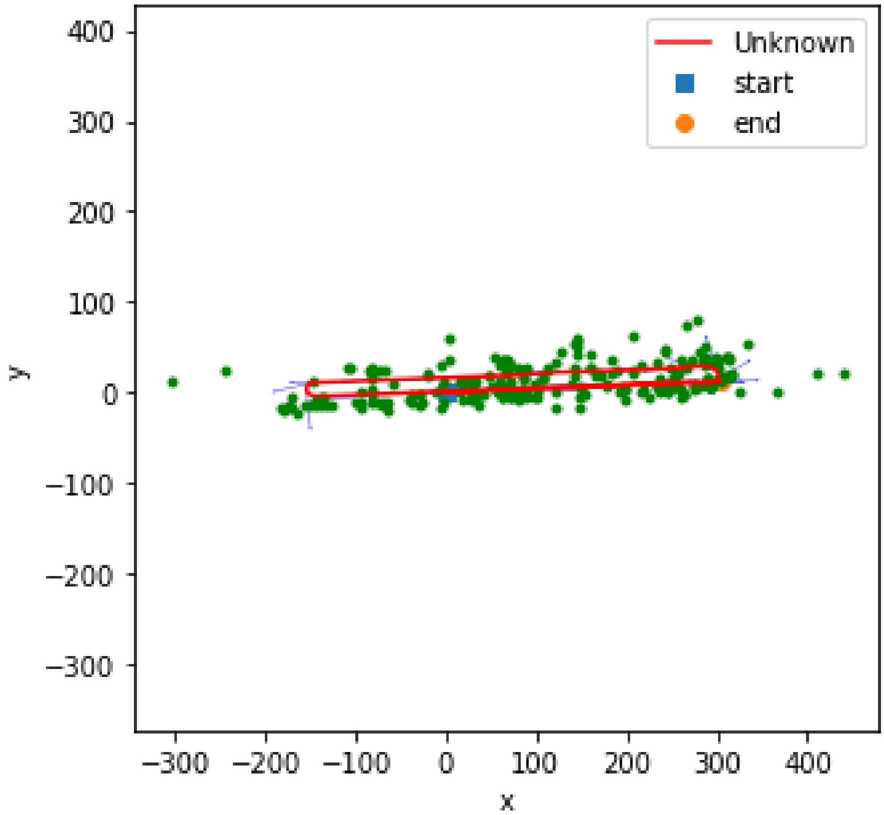

Figure3. the trajectory and the landmarks of test set