Visual SLAM with EKF Filter

Jiaming Lai[1]

I. INTRODUCTION

Visual SLAM contains three part: (1) IMU-based Localization via EKF Prediction; (2) landmark Mapping via EKF Update; (3) Visual-Inertial SLAM. In general, this is a simultaneous localization and mapping task. This problem is quite important for the following reasons: (1) an autonomous vehicle or robot should be able to localize itself given sensor data, for example, RGB-D camera etc; (2) the vehicle or robot should also generate environment map, since this map could help correct localization and navigation.

In our case, we propose EKF-based visual slam method to localize robot and generate environment map. In the first part, we implement the EKF prediction step to estimate the IMU pose over time. In the second part, we use the predicted IMU pose to initialize our landmark position. Then we perform EKF update step to track the mean and covariance of our landmark. In the third part, we combine IMU prediction step with landmark EKF udpate and IMU update step together, in order to generate complete visual-inertial SLAM algorithm.

The rest of paper is arranged as follows. First we give the detailed formulations of visual-inertial SLAM problem in Section II. Techinical approaches are introduced in Section III. And in the end we setup the experiment, results and discussion are presented in Section IV.

II. PROBLEM FORMULATION

A. Motion model

1) *Robot Body Pose:* In our project, we define IMU pose as our robot body pose, which is represented as ${}_WT_{I,t} \in R^{4\times4}$, the transformation matrix from IMU frame to world frame. Here t represents time index.
2) *Control Input:* At every time step, we define the control input as

$$u_t = [v_t^T, \omega_t^T]^T$$

where v_t is the linear velocity and ω_t is the angular velocity.
3) *Motion Model:* We then define our motion model as following:

$$U_{t+1} = exp(-\tau((u_t + w_t))^\wedge)U_t$$

where U_t is the inverse IMU pose and $U_t = {}_W T_{I,t}^{-1}$. τ is the discrete time stamps.

*This work was an project of ECE276A, University of California San Diego

[1]Jiaming Lai is with Master of Department of Electrical and Computer Engineering, University of California San Diego, 9500 Gilman Dr, La Jolla jil136@ucsd.edu

B. Observation model

The observation model we use in our project is as following:

$$z_t = M\pi({}_OT_IU_t\underline{m}) + v_t$$

where $\underline{m}$ is the homogeneous coordinates of landmark in the world frame. v_t is the noise with distribution

$$v_t \sim N(0, I \otimes V)$$

and z_t is the visual feature observations.

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

$$\pi(\mathbf{q}) = \frac{1}{q_3}\mathbf{q} \in R^4$$

$$\frac{d\pi}{d\mathbf{q}}(\mathbf{q}) = \frac{1}{q_3}\begin{bmatrix} 1 & 0 & -\frac{q_1}{q_3} & 0 \\ 0 & 1 & -\frac{q_2}{q_3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{q_4}{q_3} & 0 \end{bmatrix} \in R^{4\times4}$$

C. Task 1

Task 1 is IMU-based Localization via EKF Prediction. In this task, given $SE(3)$ kinematics $u_t = [v_t^T, \omega_t^T]^T$, we need to estimate the pose of IMU $T_t \in SE(3)$ over time t.

D. Task 2

Task 2 is Landmark Mapping via EKF Update. In this task, we assume the IMU pose $T_t \in SE(3)$ from task 1 is correct. Given a new set of pixel coordinates $z_t \in R^{4\times M}$, we estimate the unknown landmark position $m \in R^{3\times M}$ in world frame and update known landmark in order to track the mean and covariance of m. We need to use observation model to compute the Jacobian matrix H, and then compute Kalman gain K_t and finally update the mean and covariance of landmark.

E. Task 3

Task 3 is Visual-Inertial SLAM. In this task, we combine IMU prediction step from task 1 with landmark update from task 2 and IMU update based on the stereo camera observation model to obtain a complete visual-inertial SLAM algorithm. We need to compute the Jacobian matrix H, and then compute Kalman gain K_t and finally perform joint update to the mean and covariance of landmark and IMU inverse pose.

More detail about updating map please see techinical approaches in Section III.

III. TECHNICAL APPROACH

In this section, we will discuss about the algorithms and methods used in our project.

A. Task 1

In this part, we implement the EKF prediction step to estimate the IMU pose over time. In iteration, given a control input, we need to estimate the mean and covariance of IMU pose. Given our motion model,

$$U_{t+1} = exp(-\tau((u_t + w_t))^\wedge)U_t$$

the EKF predeiction step with noize $w_t \sim N(0,W)$ is as following:

$$\mu_{t+1|t} = exp(-\tau\hat{u_t})\mu_{t|t}$$

$$\Sigma_{t+1|t} = exp(-\tau\widehat{u_t})\Sigma_{t|t}exp(-\tau\widehat{u_t})^T + w_t$$

where $\mu_{t+1|t} \in SE(3)$ and $\Sigma_{t+1|t} \in R^{6\times6}$,

$$u_t = \begin{bmatrix} v_t \\ \omega_t \end{bmatrix} \in R^6$$

$$\hat{u_t} = \begin{bmatrix} \hat{\omega}_t & v_t \\ 0 & 0 \end{bmatrix} \in R^{4\times4}$$

$$\widehat{u_t} = \begin{bmatrix} \hat{\omega}_t & \hat{v}_t \\ 0 & \hat{\omega}_t \end{bmatrix} \in R^{6\times6}$$

Since we don't perform update step in this task, so

$$\mu_{t+1|t+1} = \mu_{t+1|t}$$

$$\Sigma_{t+1|t+1} = \Sigma_{t+1|t}$$

In our project, we initialize the covariance as $\Sigma_{0|0} = I \in R^{6\times6}$ and $\mu_{0|0} = I \in R^{4\times4}$. After we finish all the iteration, we get the IMU trajectory.

B. Task 2

In this part, we assume that the predicted IMU trajectory from task 1 is correct. So in every iteration, we can first project the new features into world frame and then update existing landmarks' mean and covariance. In every iteration, we record which landmark have been observed and projected,

1) *Project new landmark:* In this part, we need to project new features to world frame as initialized landmarks. Given a new observation $z_{t,i} = [u_L, v_L, u_R, v_R]^T$, the camera calibration matrix M and transformation from IMU frame to camera optical frame, suppose the new landmark in optical frame is $X_0 = [x_0, y_0, z_0, 1]^T$, we have:

$$\begin{bmatrix} u_L \\ v_L \\ u_R \\ v_R \end{bmatrix} = \begin{bmatrix} fs_u & 0 & c_u & 0 \\ 0 & fs_v & c_v & 0 \\ fs_u & 0 & c_u & -fs_ub \\ 0 & fs_v & c_v & 0 \end{bmatrix} \frac{1}{z_0} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}$$

So we can see:

$$z_0 = \frac{fs_ub}{u_L - u_R}$$

$$x_0 = z_0\frac{u_L - c_u}{fs_u}$$

$$y_0 = z_0\frac{v_L - c_v}{fs_v}$$

Then we can transform $X_0 = [x_0, y_0, z_0, 1]^T$ to landmark position $m \in R^3$ in world frame:

$$\underline{m} =_W T_{I,t} \ _IT_O \ X_0$$

where $_IT_O$ is the transform matrix from optical frame to IMU frame. $\underline{m}$ is the homogeneous coordinate of landmark in world frame.

2) *Update landmark:* In this part, given new observation $z_t \in R^{4\times N_t}$, we first compute the predicted observations based on existing landmarks' mean $\mu_t \in R^{3M}$, where M is the number of landmark.

$$\tilde{z}_{t,i} = M\pi(_OT_IU_t\underline{\mu_t}) \in R^4, \ for \ i = 1, 2, \ldots, N_t$$

In order to compute the Kalman gain, we need the Jacobian matrix. By using first-order Taylor series approximation to obesrvation i at time t, we could get the the Jacobian of $\tilde{z}$ with respect to m_j evaluated at $\mu_{t,j}$ is:

$$H_{t,i,j} = \begin{cases} M\frac{d\pi}{d\mathbf{q}}(_OT_IU_t\underline{\mu_{t,j}})_OT_IU_tP^T & if \ observation \ i \ corresponds \ to \ landmarkd \ j \ at \ time \ t \\ 0 & otherwise \end{cases}$$

where $P = [I \ 0]^T$. Then we could perform EKF udpate, given the landmark mean $\mu_t \in R^{3M}$ and covariance $\Sigma_t \in R^{3M\times3M}$, we have:

$$K_t = \Sigma_tH_t^T(H_t\Sigma_tH_t^T + I \otimes V)$$

$$\mu_{t+1} = \mu_t + K_t(z_t - \tilde{z}_t)$$

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

where $H_t \in R^{4N_t\times3M}$. In pratice, we only udpate those landmarks that already been projected in the previous steps and have corresponding observation in current new observation set.

C. Task 3

1) *Predict IMU pose:* In this part, we should first predict the IMU pose given the control input. This part is the same as task 1, so we can just skip here.
2) *Project new landmark:* The same as task 2, we project new features to world frame as initialized landmarks. Again, this part is the same as task 2, we just skip here.
3) *Perform joint update:* In order to udpate the mean and covariance of IMU pose and landmark, we perform joint update in this part. We first define the covariance matrix as following:

$$\Sigma_{t+1|t} = \begin{bmatrix} \Sigma_{t+1|t}^{LM} \in R^{3M\times3M} & 0 \\ 0 & \Sigma_{t+1|t}^{I} \in R^{6\times6} \end{bmatrix} \in R^{(3M+6)\times(3M+6)}$$

where $\Sigma_{t+1|t}^{LM}$ is the covariance matrix of landmark and $\Sigma_{t+1|t}^{I}$ is the covariance matrix of IMU. Then we define the Jacobian matrix as

$$H_{t+1|t} = \begin{bmatrix} H_{t+1|t}^{LM} \in R^{4N_t\times3M} & H_{t+1|t}^{I} \in R^{4N_t\times6} \end{bmatrix} \in R^{4N_t\times(3M+6)}$$

where $H^{LM}_{t+1|t}$ is the Jacobian corresponding to landmark and $H^{I}_{t+1|t}$ is the Jacobian corresponding to IMU. The compute method of $H^{LM}_{t+1|t}$ is the same as task 2, so we just skip here. For $H^{I}_{t+1|t}$, given IMU $\mu^{I}_{t+1|t} \in SE(3)$, $\Sigma^{I}_{t+1|t} \in R^{6\times 6}$ and the landmark position $m \in R^{3\times M}$, we have

$$H^{I}_{i,t+1|t} = M\frac{d\pi}{d\mathbf{q}}({}_OT_I\mu^{I}_{t+1|t}\underline{m_j})\,{}_OT_I(\mu^{I}_{t+1|t}\underline{m_j})^{\odot}$$

where m_j is the landmark corresponding to new observation $z_{t,i},\ i = 1,2,\dots,N_t$.

$$\begin{bmatrix} s \\ 1 \end{bmatrix}^{\odot} = \begin{bmatrix} I & -\hat{s} \\ 0 & 0 \end{bmatrix} \in R^{4\times 6}$$

The Kalman gain is computed as following:

$$K_{t+1|t} = \Sigma_{t+1|t}H^{T}_{t+1|t}(H_{t+1|t}\Sigma_{t+1|t}H^{T}_{t+1|t} + I \otimes V)$$

We note that

$$K_{t+1|t} = \begin{bmatrix} K^{LM}_{t+1|t} \in R^{3M\times(4N_t)} \\ K^{I}_{t+1|t} \in R^{6\times(4N_t)} \end{bmatrix} \in R^{(3M+6)\times(4N_t)}$$

For IMU update, given IMU $\mu^{I}_{t+1|t} \in SE(3)$, $\Sigma^{I}_{t+1|t} \in R^{6\times 6}$, we have

$$\mu^{I}_{t+1|t+1} = exp((K^{I}_{t+1|t}(z_t - \tilde{z}_t))^{\wedge})\mu^{I}_{t+1|t}$$
$$\Sigma^{I}_{t+1|t+1} = (I - K^{I}_{t+1|t}H^{I}_{t+1|t})\Sigma^{I}_{t+1|t}$$

For landmark update, given landmark $\mu^{LM}_{t+1|t} \in 3M$, $\Sigma^{LM}_{t+1|t} \in R^{3M\times 3M}$, we have

$$\mu^{LM}_{t+1|t+1} = \mu^{LM}_{t+1|t} + K^{LM}_{t+1|t}(z_t - \tilde{z}_t)$$
$$\Sigma^{LM}_{t+1|t+1} = (I - K^{LM}_{t+1|t}H^{LM}_{t+1|t})\Sigma^{LM}_{t+1|t}$$

IV. RESULTS

In this section, we will shown and discuss results generated from data set. The **red line** represents IMU trajectory. The initialized parameter is shown as below table.

A. Data_0022

1) *Task 1:* The IMU trajectory is shown in Fig. 1.

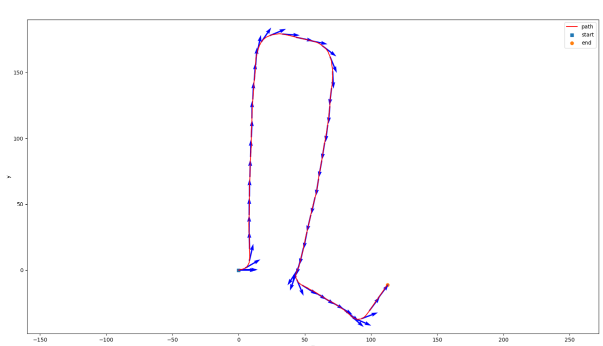

Fig. 1: Data_0022 task 1

2) *Task 2:* The IMU trajectory and updated landmark is shown in Fig. 2. Note that the **yellow** points are updated landmark and **blue** points are the first projected landmarks.

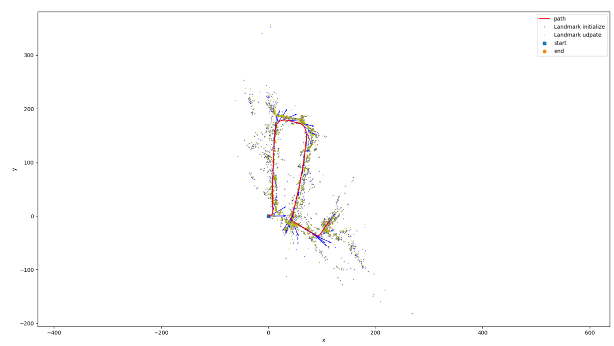

Fig. 2: Data_0022 task 2

3) *Task 3:* The IMU trajectory and landmark is shown in Fig. 3. Note that the **yellow** points are landmarks.

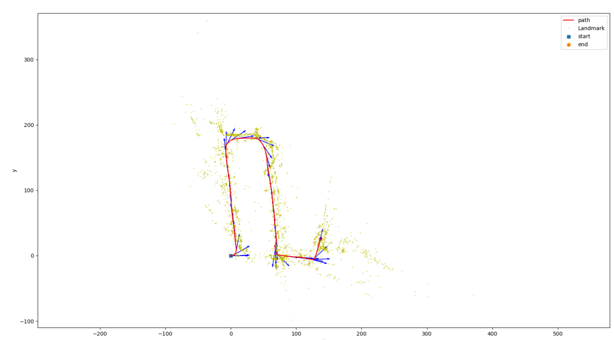

Fig. 3: Data_0022 task 3

B. *Data_0027*

1) *Task 1:* The IMU trajectory is shown in Fig. 4. We can see that the start point and the end point don't merge together, meaning our predicted IMU trajectory is incorrect.

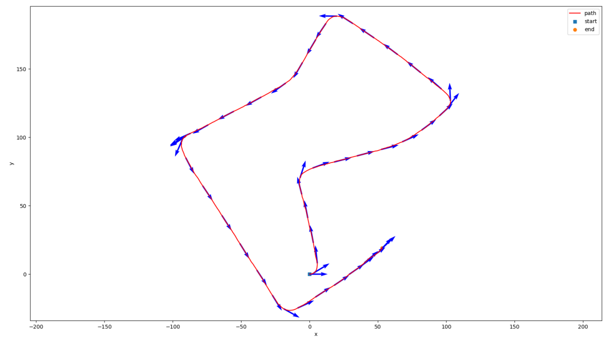

Fig. 4: Data_0027 task 1

2) *Task 2:* The IMU trajectory is shown in Fig. 5. Note that the **yellow** points are updated landmark and **blue** points are the first projected landmarks. We could still see that the start point and the end point of the IMU trajectory don't merge together. The projected landmarks and updated landmarks have small perturbation. It is intuitive because we perform udpate step.

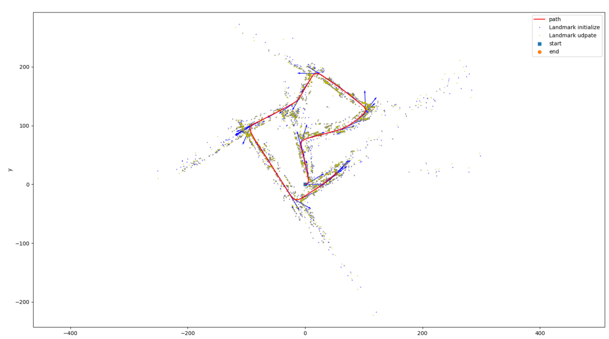

Fig. 5: Data_0027 task 2

3) *Task 3:* The IMU trajectory is shown in Fig. 6. We could see that by performing joint udpate, the start point and end point of IMU trajectory merge together, so here the IMU trajectory is correct.

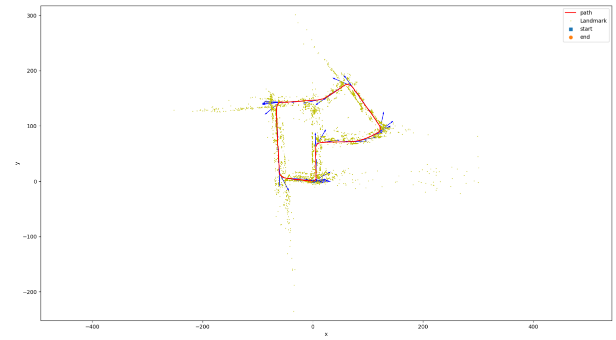

Fig. 6: Data_0027 task 3

C. *Data_0034*

1) *Task 1:* The IMU trajectory is shown in Fig. 7. We can see that the IMU trajectory have overlap where IMU perform large rotation.

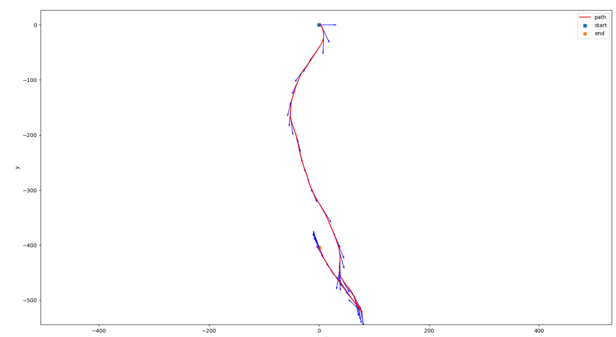

Fig. 7: Data_0034 task 1

2) *Task 2:* The IMU trajectory is shown in Fig. 8. Note that the **yellow** points are updated landmark and **blue** points are the first projected landmarks. We could still see that the IMU trajectory have overlap where IMU perform large rotation.
3) *Task 3:* The IMU trajectory is shown in Fig. 9. We could see that the IMU trajectory separates well, so here the IMU trajectory is correct.

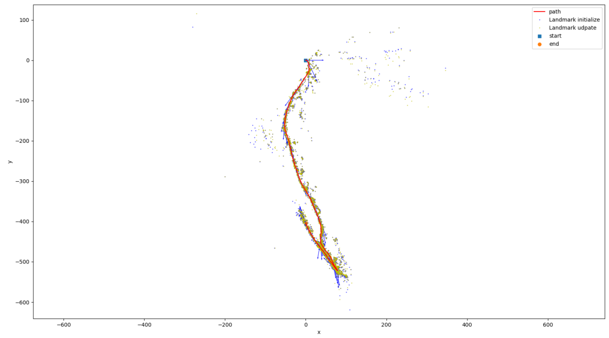

Fig. 8: Data_0034 task 2

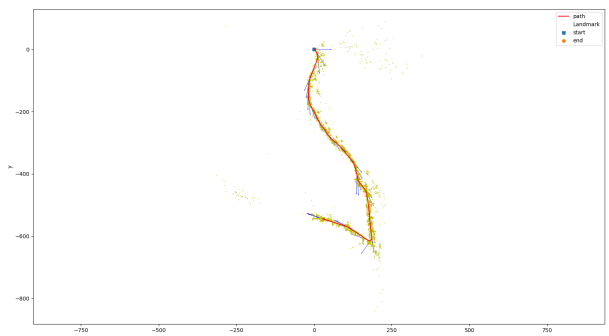

Fig. 9: Data_0034 task 3

V. DISCUSSION

Our method work pretty well in data_0022, data_0027, data_0034.

As we discuss above, if we just perform IMU pose prediction based on the kinematics input, our IMU trajectory is incorrect because of noise. In task 2, the updated landmarks are still incorrect because we assume the predict IMU trajectory is correct. But in fact, the predict IMU trajectory is noisy. In task 3, we could see that by performing joint update with IMU and landmarks, the IMU trajectory looks pretty make sense. It is intuitive because landmark information would help IMU to adjust its location and therefore, we get a better trajectory estimation.

VI. CITATIONS

The method of selecting features is inspired by Jiangeng Dong and Xinyu Liu.