IMU and 6 DoF Odometry (Stereo Visual Odometry) Loosely-Coupled Fusion Localization based on UKF

Preprint · February 2022

1 author:

Hongchen Gao
HJIMI Technology Co., Ltd.

IMU and 6 DoF Odometry (Stereo Visual Odometry) Loosely-Coupled Fusion Localization based on UKF

Hongchen Gao
cggos@outlook.com

Sunday 27th February, 2022

1 Introduction

It has been long known that fusing information from multiple sensors for robot navigation results in increased robustness and accuracy [1]. This paper presents a framework that make the loosely coupled fusion of IMU and Stereo Visual Odometry (VO) from ORB-SLAM2 [2] (Stereo Mode) based on ESKF (Error-State Kalman Filter) [3]. The code is available at: https://github.com/cggos/imu_x_fusion.

2 Unscented Kalman Filter (UKF)

2.1 Unscented Transformation (UT)

A simple example is shown in Figure 1 for a 2-dimensional system: the left plot shows the true mean and covariance propagation using Monte-Carlo sampling; the center plots show the results using a linearization approach as would be done in the EKF; the right plots show the performance of the UT (note only 5 sigma points are required). The superior performance of the UT is clear [4].

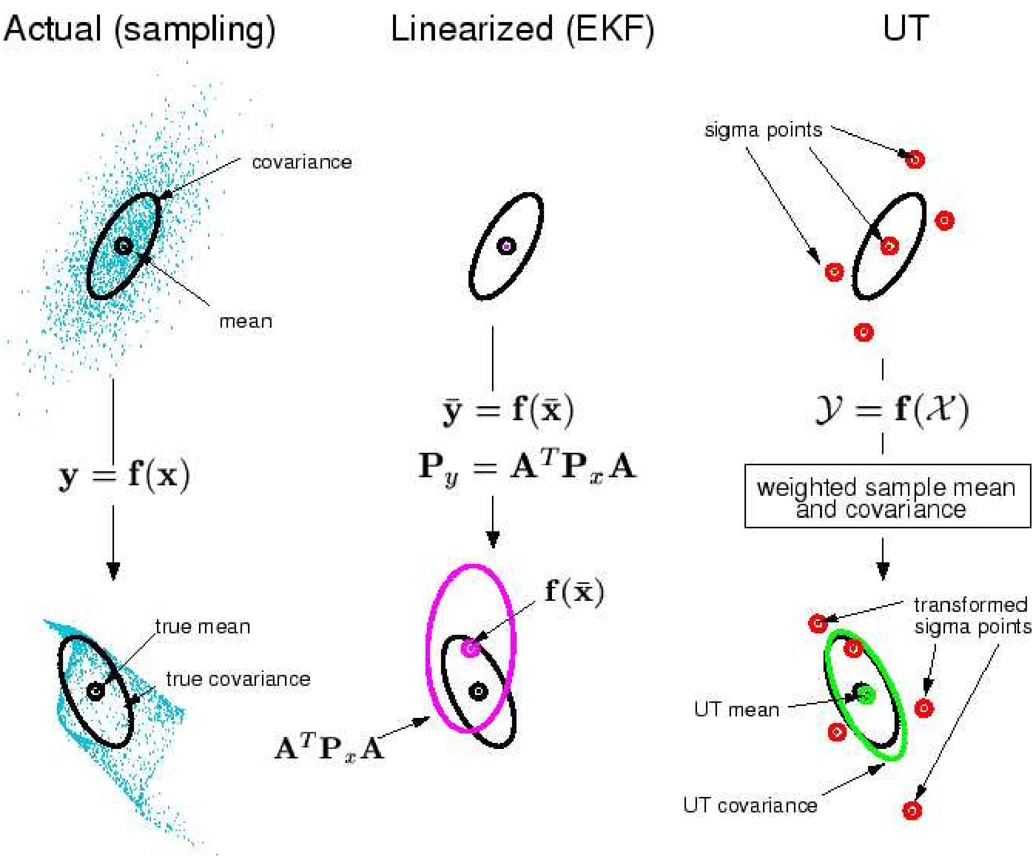

Figure 1: Example of the UT for mean and covariance propagation. a) actual, b) first-order linearization (EKF), c) UT.

The unscented transformation (UT) [5] is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. The n dimensional random variable x with mean $\bar{x}$ and covariance P_{xx} is approximated by $2n+1$ weighted points given by

$$\begin{aligned}
\chi_0 &= \bar{x} \\
\chi_i &= \bar{x} + \left(\sqrt{(n+\lambda)P_{xx}}\right)_i \quad i = 1, \cdots, n \\
\chi_i &= \bar{x} - \left(\sqrt{(n+\lambda)P_{xx}}\right)_i \quad i = n+1, \cdots, 2n
\end{aligned}$$

These sigma points are propagated through the function

$$\mathcal{Y}_i = f\left(\chi_i\right) \quad i = 0, \cdots, 2n$$

2.2 Mean and Covariance [6]

The predicted mean and covariance for the state estimate

$$\begin{aligned}
\hat{\mathbf{x}}_{k+1}^{-} &= \sum_{i=0}^{2n} W_i^{\text{mean}} \, \boldsymbol{\chi}_{k+1}(i) \\
P_{k+1}^{-} &= \sum_{i=0}^{2n} W_i^{\text{cov}} \left[\chi_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^{-}\right] \left[\chi_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^{-}\right]^T + \bar{Q}_k
\end{aligned}$$

and the mean and covariance for $\mathcal{Y}$ are approximated using a weighted sample mean and covariance of the posterior sigma points,

$$
\begin{aligned}
\hat{\mathbf{y}}_{k+1}^{-} &= \sum_{i=0}^{2n} W_i^{\text{mean}} \mathcal{Y}_i \\
P_{k+1}^{yy} &= \sum_{i=0}^{2n} W_i^{\text{cov}} \left[\mathcal{Y}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^{-}\right] \left[\mathcal{Y}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^{-}\right]^T
\end{aligned}
$$

We now define the following weights:

$$
\begin{aligned}
W_0^{\text{mean}} &= \frac{\lambda}{n+\lambda} \\
W_0^{\text{cov}} &= \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right) \\
W_i^{\text{mean}} = W_i^{\text{cov}} &= \frac{1}{2(n+\lambda)}, \quad i = 1, 2, \ldots, 2n
\end{aligned}
$$

where the composite scaling parameter, λ, is given by

$$
\lambda = \alpha^2(n + \kappa) - n
$$

The constant α determines the spread of the sigma points and is usually set to a small positive value (e.g., $1 \times 10^{-4} \leq \alpha \leq 1$). κ is a secondary scaling parameter which is usually set to 0. And β is used to incorporate prior knowledge of the distribution (a good starting guess is $\beta = 2$).

Then the innovations covariance is simply given by

$$
P_{k+1}^{vv} = P_{k+1}^{yy} + R_{k+1}
$$

Finally the cross correlation matrix is determined using

$$
P_{k+1}^{xy} = \sum_{i=0}^{2n} W_i^{\text{cov}} \left[\mathcal{X}_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^{-}\right] \left[\mathcal{Y}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^{-}\right]^T
$$

3 System Overview

We obtain the state vector $\mathbf{X}$ including $\mathbf{P}, \mathbf{v}, \mathbf{q}, \mathbf{b}_a, \mathbf{b}_g$ of the system by fusing the IMU data (linear acceleration and angular velocity) and the pose (3D position and rotation) from the Stereo Visual Odometry based on ESKF.

3.1 System Coordinate

This system mainly includes four coordinate systems, including IMU, camera, vision and world frame, as shown in Fig. 2 [1]. Need to align the two coordinate systems before information fusion.

3.2 System State [7]

3.2.1 system state vector

the nominal-state is

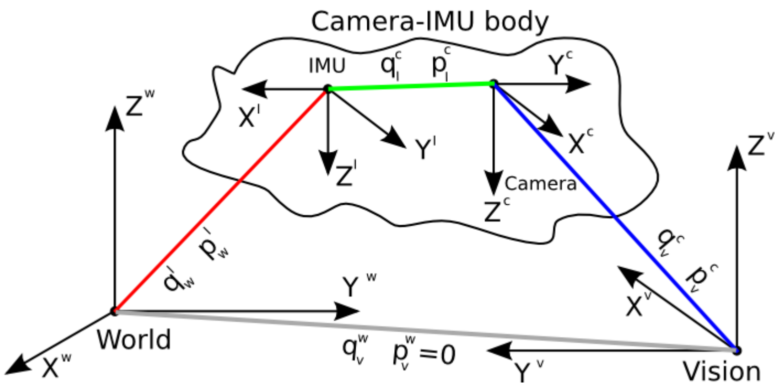

Figure 2: System Coordinate

$$\mathbf{x} = \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \\ \mathbf{q} \\ \mathbf{b}_a \\ \mathbf{b}_g \end{bmatrix} \in \mathbb{R}^{16\times 1}$$

the error-state is

$$\delta\mathbf{x} = \begin{bmatrix} \delta\mathbf{p} \\ \delta\mathbf{v} \\ \delta\boldsymbol{\theta} \\ \delta\mathbf{b}_a \\ \delta\mathbf{b}_g \end{bmatrix} \in \mathbb{R}^{15\times 1}$$

the true-state is

$$\mathbf{x}_t = \mathbf{x} \oplus \delta\mathbf{x}$$

3.2.2 system state covariance

$$\mathbf{P} \in \mathbb{R}^{15\times 15} \quad \text{with} \quad \mathbf{P}_0 = \mathbf{I}_{15}$$

4 Initialization

First, we need to initialize the Bias of the IMU and align the IMU coordinate system with the vision system in ENU frame with the z axis aligned with up which is define as the world frame **w**.

The gyroscope bias of IMU is computed with

$$\mathbf{b}_g = \frac{1}{n} \cdot \sum_{i}^{n} \boldsymbol{\omega}_i$$

The z axis aligned with the mean of accelerations

$$\mathbf{z} = \frac{1}{n} \cdot \sum_{i}^{n} \mathbf{a}_i$$

Through Gram–Schmidt orthogonalization [8], we can get x axis

$$\mathbf{x} = \mathbf{e}_x - \mathbf{z} \cdot \mathbf{z}^T \cdot \mathbf{e}_x \quad \text{with} \quad \mathbf{e}_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

And the y axis is get simply by the cross-product of x axis and z axis

$$\mathbf{y} = \mathbf{z} \times \mathbf{x}$$

Now, we can get the rotation from World frame to IMU frame

$$\mathbf{R}_{bw} = \begin{bmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \end{bmatrix}$$

As long as the system is initialized on timestamp m, we set the pose of IMU in World frame

$$\mathbf{T}_{wb} = \mathbf{T}_{b_0 b_m} = \begin{bmatrix} \mathbf{R}_{bw}^T & \mathbf{0} \\ 0 & 1 \end{bmatrix}$$

And the pose of stereo visual odometry on Visual frame is

$$\mathbf{T}_{vc} = \mathbf{T}_{c_0 c_m}$$

5 IMU-driven system kinematics in discrete time

According to the article [3] and system states in section 3.2, we can get the nominal-state and error-state kinematics.

The nominal-state kinematics is

$$\begin{aligned}
&\mathbf{p} \leftarrow \mathbf{p} + \mathbf{v}\Delta t + \tfrac{1}{2}\left(\mathbf{R}\left(\mathbf{a}_m - \mathbf{a}_b\right) + \mathbf{g}\right)\Delta t^2 \\
&\mathbf{v} \leftarrow \mathbf{v} + \left(\mathbf{R}\left(\mathbf{a}_m - \mathbf{a}_b\right) + \mathbf{g}\right)\Delta t \\
&\mathbf{q} \leftarrow \mathbf{q} \otimes \mathbf{q}\left\{\left(\boldsymbol{\omega}_m - \boldsymbol{\omega}_b\right)\Delta t\right\} \\
&\mathbf{a}_b \leftarrow \mathbf{a}_b \\
&\boldsymbol{\omega}_b \leftarrow \boldsymbol{\omega}_b
\end{aligned}$$

The error-state kinematics is

$$\begin{aligned}
\delta\mathbf{p} &\leftarrow \delta\mathbf{p} + \delta\mathbf{v}\Delta t \\
\delta\mathbf{v} &\leftarrow \delta\mathbf{v} + \left(-\mathbf{R}\left[\mathbf{a}_m - \mathbf{a}_b\right]_\times \delta\boldsymbol{\theta} - \mathbf{R}\delta\mathbf{a}_b + \delta\mathbf{g}\right)\Delta t + \mathbf{v}_\mathbf{i} \\
\delta\boldsymbol{\theta} &\leftarrow \mathbf{R}^\top\left\{\left(\boldsymbol{\omega}_m - \boldsymbol{\omega}_b\right)\Delta t\right\}\delta\boldsymbol{\theta} - \delta\boldsymbol{\omega}_b\Delta t + \boldsymbol{\theta}_\mathbf{i} \\
\delta\mathbf{a}_b &\leftarrow \delta\mathbf{a}_b + \mathbf{a}_\mathbf{i} \\
\delta\boldsymbol{\omega}_b &\leftarrow \delta\boldsymbol{\omega}_b + \boldsymbol{\omega}_\mathbf{i}
\end{aligned}$$

6 Generate Sigma Points

状态

$$x_k = \begin{bmatrix} p \\ v \\ \theta \\ b_a \\ b_g \end{bmatrix} \in \mathbf{R}^{15}, \quad N_x = 15$$

增广状态

$$x_{a,k} = \begin{bmatrix} p \\ v \\ \theta \\ b_a \\ b_g \\ n_a \\ n_{wa} \\ n_g \\ n_{wg} \end{bmatrix} \in \mathbf{R}^{27}, \quad N_{aug} = 27$$

Sigma Points 点数

$$N_\sigma = 2N_{aug} + 1 = 55$$

增广状态矩阵

$$X_{a,k|k} = \left[\ x_{a,k|k} \quad x_{a,k|k} + \sqrt{(\lambda + N_{aug})\, P_{a,k|k}} \quad x_{a,k|k} - \sqrt{(\lambda + N_{aug})\, P_{a,k|k}}\ \right]$$

增广状态协方差矩阵

$$P_{a,k|k} = \begin{bmatrix} P_{k|k} & 0 \\ 0 & Q \end{bmatrix} \in \mathbf{R}^{27\times 27}$$

7 Predict Process

7.1 Predict Sigma Points

通过 IMU 运动学方程

$$X_{a,k|k} \in \mathbf{R}^{N_{aug}\times N_\sigma} \quad \longrightarrow \quad X_{k+1|k} \in \mathbf{R}^{N_x \times N_\sigma}$$

7.2 Predict Mean and Covariance

权重

$$
\begin{aligned}
w_i &= \frac{\lambda}{\lambda + N_{aug}}, i = 1 \\
w_i &= \frac{1}{2\left(\lambda + N_{aug}\right)}, i = 2 \ldots N_\sigma
\end{aligned}
$$

预测均值

$$
x_{k+1|k} = \sum_{i=1}^{N_\sigma} w_i \mathcal{X}_{k+1|k,i} \in \mathbf{R}^{N_x}
$$

预测协方差

$$
P_{k+1|k} = \sum_{i=0}^{N_\sigma} w_i \left(\mathcal{X}_{k+1|k,i} - x_{k+1|k}\right)\left(\mathcal{X}_{k+1|k,i} - x_{k+1|k}\right)^T \in \mathbf{R}^{N_x \times N_x}
$$

8 Predict Measurement

Measurement Model

$$
z_{k+1} = h\left(x_{k+1}\right) + \omega_{k+1}
$$

Measurement function is

$$
\begin{aligned}
h(\hat{x}) \longleftarrow & \underbrace{T_{c_0 c_m} \cdot T_{cb} \cdot T_{b_0 b_m}^{-1}}_{T_{vw}} \cdot T_{b_0 b_n} \cdot T_{cb}^{-1} \\
=& T_{vw} \cdot T \cdot T_{cb}^{-1} \\
=& \begin{bmatrix} R_{vw} R R_{cb}^T & R_{vw}(t + R t_{bc}) + R_{c_0 c_m} t_{cb} + t_{c_0 c_m} \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

8.1 Predict Measurement Sigma Points

状态 sigma points 到测量 sigma points

$$
X_{k+1|k} \in \mathbf{R}^{N_x \times N_\sigma} \quad \longrightarrow \quad Z_{k+1|k} \in \mathbf{R}^{N_z \times N_\sigma}, \quad N_z = 6
$$

8.2 Predict Measurement Mean and Covariance

测量均值

$$
z_{k+1|k} = \sum_{i=1}^{N_\sigma} w_i \mathcal{Z}_{k+1|k,i}
$$

测量协方差

$$
S_{k+1|k} = \sum_{i=0}^{N_\sigma} w_i \left(\mathcal{Z}_{k+1|k,i} - z_{k+1|k}\right)\left(\mathcal{Z}_{k+1|k,i} - z_{k+1|k}\right)^T + R, \quad \text{with} \quad R = E\left\{\omega_k \cdot \omega_k^T\right\}
$$

9 Update State

Cross-correlation

$$T_{k+1|k} = \sum_{i=0}^{N_\sigma} w_i \left(\mathcal{X}_{k+1|k,i} - x_{k+1|k}\right) \left(\mathcal{Z}_{k+1|k,i} - z_{k+1|k}\right)^T$$

Kalman Gain

$$K_{k+1|k} = T_{k+1|k} S_{k+1|k}^{-1}$$

State update

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1|k} \left(z_{k+1} - z_{k+1|k}\right)$$

Covariance matrix update

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1|k} S_{k+1|k} K_{k+1|k}^T$$

References

[1] Simon Lynen, Markus W Achtelik, Stephan Weiss, Margarita Chli, and Roland Siegwart. A robust and modular multi-sensor fusion approach applied to mav navigation. In *2013 IEEE/RSJ international conference on intelligent robots and systems*, pages 3923–3929. IEEE, 2013.

[2] Raul Mur-Artal and Juan D Tardós. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, 33(5):1255–1262, 2017.

[3] Joan Sola. Quaternion kinematics for the error-state kalman filter. *arXiv preprint arXiv:1711.02508*, 2017.

[4] Eric A Wan and Rudolph Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pages 153–158. Ieee, 2000.

[5] Pifu Zhang, Jason Gu, Evangelos E Milios, and Peter Huynh. Navigation with imu/gps/digital compass with unscented kalman filter. In *IEEE International Conference Mechatronics and Automation, 2005*, volume 3, pages 1497–1502. IEEE, 2005.

[6] John L Crassidis. Sigma-point kalman filtering for integrated gps and inertial navigation. *IEEE Transactions on Aerospace and Electronic Systems*, 42(2):750–756, 2006.

[7] Hongchen Gao. Imu and vo loose fusion based on eskf. ResearchGate, 7 2021. doi: 10.13140/RG.2.2.28797.69602.

[8] Gram–Schmidt process. `https://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process`, 2021.