Sensor Fusion for State Estimation of a MAV using an Invariant Extended Kalman Filter

Blake Karwoski, Kun Huang, Yue Wu, Zhaoguo Wang
University of Michigan, Ann Arbor, MI

***Abstract*—As Micro Aerial Vehicles (MAVs) become more affordable and ubiquitous, their usage in complex urban environments will become more desirable for uses such as inspection, monitoring, and delivery. Navigation of these environments demands greater positional accuracy than can be provided by conventional GNSS systems. While MAVs typically contain inertial measurement units (IMUs), their integration-based state estimates are prone to drifting over time. We explore the usage of sensor fusion to combine these complementary sensors. In this project, we use an Invariant Extended Kalman filter (InEKF) to estimate the position of a MAV in a challenging urban environment. We evaluate our result by comparing the estimated position to the ground truth dataset.**

***Index Terms*—Invariant Extended Kalman Filter (InEKF), localization**

I. Introduction

The normal extended Kalman filter (EKF) estimates the covariance between the states by linearizing the dynamic equations. However, this leads to some limitations, such as incapability of guaranteed-convergence. In the meantime, InEKF can solve the problems above, and has strict mathematical derivation as guarantee.

Algorithm 1 Extended Kalman filter

Input: belief mean μ_{k-1}, belief covariance Σ_{k-1}, action u_k, measurement z_k;
Output: updated μ_k, Σ_k
1: $\mu_k^- \leftarrow f(u_k, \mu_{k-1})$ *%predicted mean*
2: $\Sigma_k^- \leftarrow F_k \Sigma_{k-1} F_k^T + W_k Q_k W_k^T$ *%predicted covariance*
3: $\nu_k \leftarrow z_k - h(\mu_k^-)$ *%innovation*
4: $S_k \leftarrow H_k \Sigma_k^- H_k^T + V_k R_k V_k^T$ *%innovation covariance*
5: $K_k \leftarrow \Sigma_k^- H_k^T S_k^{-1}$ *%filter gain*
6: $\mu_k \leftarrow \mu_k^- + K_k \nu_k$ *%corrected mean*
7: $\Sigma_k \leftarrow (I - K_k H_k)\Sigma_k^-$ *%corrected covariance*
8: return μ_k, Σ_k

II. Problem formulation

A. Transformation between Lie group and Lie algebra

•From $so(3)$ to $SO(3)$:
Assume $R \in R^{3\times3} in SO(3), \phi \in R^3$ in $so(3)$, $\phi = [\phi_1 \quad \phi_2 \quad \phi_3]^T$, then $\phi^\wedge$ is used to denote the skew transformation:

$$\phi^\wedge = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix}$$

Using exponential mapping:

$$R = \exp(\phi^\wedge) = \mathbf{I_3} + (\frac{\sin\theta}{\theta}\phi^\wedge) + (\frac{1-\cos\theta}{\theta^2})(\phi^\wedge)^2 \quad \in SO(3)$$

where $\theta \triangleq ||\phi||$,
$so(3)$ is mapped to the corresponding element in $SO(3)$. Reversely, using logarithm mapping, an element $R \in SO(3)$ can be recovered to its corresponding element in $so(3)$:

$$\phi^\wedge = \log(R) = \frac{\theta}{2\sin\theta}(R - R^T) \quad \in so(3)$$

where $\theta \triangleq \cos^{-1}\frac{tr(R)-1}{2}$

• From $se(3)$ to $SE(3)$:
Assume $H \in R^{4\times4} in SE(3), \xi \in R^6$ in $se(3)$, $\xi = [\phi \quad \rho]^T$, then $\xi^\wedge$is used to denote the skew transformation:

$$\xi^\wedge = \begin{bmatrix} \phi^\wedge & \rho \\ 0_{1\times3} & 0 \end{bmatrix}$$

Using exponential mapping:

$$H = \exp(\xi^\wedge) = \begin{bmatrix} \exp(\phi^\wedge) & J_l(\phi)\rho \\ 0_{1\times3} & 1 \end{bmatrix} \in \quad SE(3)$$

where $\exp(\phi^\wedge)$ and $J_l(\phi)\rho$ are the exponential map and the left Jacobian of SO(3).
$J_l(\phi)$ is able to be calculated as:

$$J_l(\phi) = \mathbf{I_3} + (\frac{1-\cos\theta}{\theta^2})\phi^\wedge + (\frac{\theta-\sin\theta}{\theta^3})(\phi^\wedge)^2$$

In this way $se(3)$is mapped to the corresponding element **H** in $SE(3)$. Reversely, using logarithm mapping, an element $H \in SE(3)$ can be recovered to its corresponding element in $se(3)$:

$$\xi^\wedge = \log(H) = \begin{bmatrix} \log(R) & J_l^{-1}(log(R))p \\ 0_{1\times3} & 0 \end{bmatrix} \in se(3)$$

B. Left-Invariant EKF Modal Derivation

- IMU dynamics:

$$\dot{R}_t = R_t(\tilde{\omega}_t - \omega_t^g)^\wedge$$
$$\dot{v}_t = R_t(\tilde{a}_t - \omega_t^a) + g$$
$$\dot{p}_t = v_t,$$

where g is the gravity vector.

- IMU Discrete Dynamics:

Assuming a zero-order hold on the incoming IMU measurement between t_k and t_{k+1}, we have:

$$R_{k+1} = R_k \Gamma_0(\bar{\omega}_k \Delta t) = R_k \exp(\bar{\omega}_k \Delta t)$$
$$v_{k+1} = v_k + R_k \Gamma_1(\bar{\omega}_k \Delta t)\bar{a}_k \Delta t + g\Delta t$$
$$p_{k+1} = p_k + v_k \Delta t + R_k \Gamma_2(\bar{\omega}_k \Delta t)\bar{a}_k \Delta t^2 + \frac{1}{2} g \Delta t^2$$

where $\bar{\omega}_t \triangleq \tilde{\omega}_t - \bar{b}_t^g$ and $\bar{a}_t \triangleq \tilde{a}_t - \bar{b}_t^a$ are the "bias-corrected" inputs. Here for the sake of simplicity, we assume there is no bias existing, which means $\tilde{\omega} = \bar{\omega}$.
The Γ_x functions are defined as following:

$$\Gamma_0(\phi) = I + \frac{sin(||\phi||)}{||\phi||}(\phi^\wedge) + \frac{1 - cos(||\phi||)}{||\phi^2||}(\phi^\wedge)^2$$
$$\Gamma_1(\phi) = I + \frac{1 - cos(||\phi||)}{||\phi||^2}(\phi^\wedge) + \frac{||\phi|| - sin(||\phi||)}{||\phi||^3}(\phi^\wedge)^2$$
$$\Gamma_2(\phi) = \frac{1}{2} I + \frac{||\phi|| - sin(||\phi||)}{||\phi||^3}(\phi^\wedge) + \frac{||\phi||^2 + 2cos(||\phi||) - 2}{2||\phi||^4}(\phi^\wedge)^2$$
$$\Gamma_m(\phi) = (\Sigma_{n=0}^{\infty} \frac{1}{(n+m)!}(\phi^\wedge)^n)$$

Note that $\Gamma_0(\phi)$ is simply the exponential map of $SO3$. $\Gamma_1(\phi)$ is also known as the left Jacobian of $SO(3)$.

- State definition:

For this MAV, the raw data include that captured by the IMU sensor, which is linear acceleration, gyro, and the raw gps data from the onboard gps sensor. We are here to set the state in Lie group as:

$$X_t = \begin{bmatrix} R_t & v_t & p_t \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{bmatrix} \in R^{5\times 5}$$

X_t is initialized as 5×5 identity matrix at the beginning.

- Using left-invariant EKF, define the error between the real state X_t and the estimated $\tilde{X}_t$ as:

$$\eta_t = X_t^{-1} \bar{X}_t$$

Now imitate things done in EKF, we start to derive the system function of this error:

$$\eta_t = X_t^{-1}\bar{X} = \begin{bmatrix} R_t^{-1}\bar{R}_t & R_t^{-1}\tilde{v}_t - R_t^{-1}v_t & R_t^{-1}\tilde{p}_t - R_t p_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\eta_t = X_t^{-1} = \begin{bmatrix} R_t^T\bar{R}_t & R_t^T\tilde{v}_t - R_t^T v_t & R_t^T\tilde{p}_t - R_t p_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Set} \quad R_t^T\bar{R}_t = \eta_{R_t}, s.t. \exp(\xi_{R_t}^\wedge) = \eta_{R_t}$$
$$R_t^T\tilde{v}_t - R_t^T v_t = \xi_{v_t}$$
$$R_t^{-1}\tilde{p}_t - R_t^T p_t = \xi_{p_t}$$

Therefore, using first order Taylor expansion, approximately gives out $\eta_{R_t} \approx \mathbf{I} + \xi_{R_t}^\wedge$

$$\dot{\eta}_t \approx \begin{bmatrix} \dot{\xi}_{R_t}^\wedge & \dot{\xi}_{v_t} & \dot{\xi}_{p_t} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

In which, $\dot{\xi}_{R_t}^\wedge$, $\dot{\xi}_{v_t}$, $\dot{\xi}_{p_t}$ are derived as following:

$$\begin{aligned} \dot{\xi}_{R_t}^\wedge \approx \dot{\eta}_{R_t} &= \dot{R}_t^T\bar{R}_t + R_t^T\dot{\bar{R}}_t \\ &= R_t^T\omega_t^\wedge\bar{R}_t - R_t^T\tilde{\omega}^\wedge\bar{R}_t \\ &= R_t^T(\omega_t^\wedge - \tilde{\omega}^\wedge)\bar{R}_t \\ &= -R_t^T(n_t^\omega)^\wedge\bar{R}_t \\ &= -R_t^T(n_t^\omega)^\wedge R_t R_t^T\bar{R}_t \\ &= (R_t^T n_t^\omega)^\wedge\eta_{R_t} \\ &\approx -(R_t^T n_t^\omega)^\wedge(\mathbf{I} + \xi_{R_t}^\wedge) \\ &\approx -(R_t^T n_t^\omega)^\wedge \end{aligned}$$

The H.O.T. in the last step is omitted.

$$\begin{aligned} \dot{\xi}_{v_t} &= R_t^T\omega_t^\wedge\tilde{v}_t + R_t^T\dot{\tilde{v}}_t - R_t^T\omega_t^\wedge v - R_t^T\dot{v}_t \\ &= R_t^T\omega_t^\wedge(\tilde{v}_t - v_t) + \eta_{R_t}\tilde{a}_t - a_t \\ \dot{\xi}_{p_t} &= R_t^T\omega_t^\wedge\tilde{p}_t + R_t^T\tilde{V}_t - R_t\omega_t^\wedge p_t - R_t V_t \end{aligned}$$

Rearrange the function above to get the format:

$$\xi_t = A_t^l\xi_t + \bar{Q}_t^l n_t$$

In which A_t^l and Q_t^l are:

$$A_t^l = \begin{bmatrix} -\tilde{\omega}_t^\wedge & 0 & 0 \\ -\tilde{a}_t^\wedge & -\tilde{\omega}_t^\wedge & 0 \\ 0 & I & -\tilde{\omega}_t^\wedge \end{bmatrix} \in R^{9\times 9}$$

$$\bar{Q}_t^l = Cov[\omega_t] \in R^{9\times 9}$$

Deterministic Nonlinear Dynamics:

$$f_{u_t}(\bar{X}_t) = \begin{bmatrix} \bar{R}_t\tilde{\omega}_t^\wedge & \bar{R}_t\tilde{a}_t + g & \bar{v}_t \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{bmatrix} \in R^{5\times 5}$$

- Invariant observation modal

If observation take a particular form, then the linearized observation model and the innovation will also be autonomous. This happens when the measurement, Y_{t_k}, can be written as

$$Y_{t_k} = X_{t_k} b + V_{t_k}$$

in left-invariant observation form, where b is a constant vector and V_{t_k} is a vector of Gaussian noise.
In our project, since X_{t_k} is set as a 5×5 matrix in Lie group and the target data is the position, and H a 9×3 matrix

$$H\xi_k^r = \quad \xi^\wedge$$

$$H = \begin{bmatrix} 0 & 0 & I \end{bmatrix}$$

- Propagation:

$$\frac{d}{dt}\bar{X}_t = f_{u_t}(\bar{X}_t), \quad t_{k-1} \leq t < t_k$$
$$\frac{d}{dt}P_t^l = A_t^l P_t^l + P_t^l A_t^{l^T} + Q_t$$

where $Q_t \in R^{9\times9}$is the action modal noise covariance and $P_{t_k} \in R^{9\times9}$is the state covariance, initialized as $\mathbf{I_{9\times9}}$ and the value from the reference respectively.

- Correction:

$$\bar{X}_{t_k}^+ = \bar{X}_{t_k} \exp(L_{t_k}(\bar{X}_{t_k}^{-1} Y_{t_k} - b))$$
$$P_{t_k}^{r+} = (I - L_{t_k}H)P_{t_k}^r(I - L_{t_k}H)^T + L_{t_k}\bar{N}_k L_{t_k}^T$$

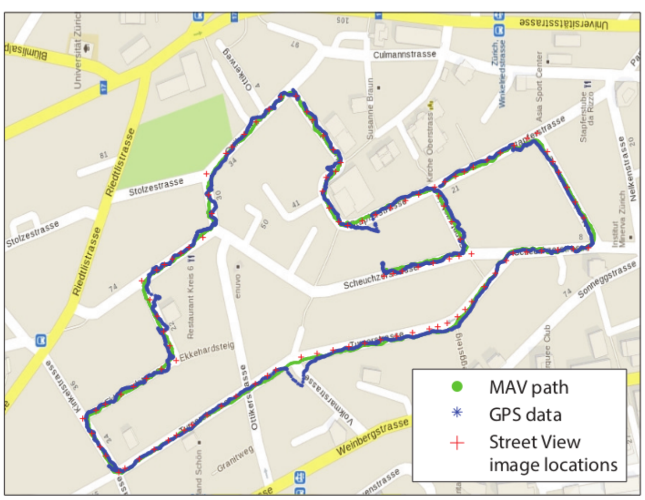
$$L_{t_k} = P_{t_k}^r H^T S^{-1}, \quad S = HP_{t_k}^r H^T + \bar{N}_k, \quad \bar{N}_k = \bar{X}_{t_k} Cov[V_k] \bar{X}_{t_k}^T$$

where $P_{t_k} \in R^{9\times9}$ is the belief covariance and $V_k \in R^{5\times5}$ represents the Gaussian noise during measurement process. P_{t_k} is initialized as a 9×9 identity matrix and V_k is chosen as a 5×5 identity matrix, according to the parameter reference.

III. Implementation

A. Dataset

We chose the Zurich Urban Micro Aerial Vehicle Dataset in order to apply and test our InEKF algorithm.



Aerial view

Fig. 1: Bird-eye view of the urban test area. The red plus signs mark the locations of the ground Google Street View images. The blue asterisks mark the GPS labels of the aerial MAV images measured on-board. The green dots represent the ground truth path of the MAV.

This dataset is the world's first dataset which was recorded on a Micro Aerial Vehicle (MAV) flying within urban streets at low altitudes (i.e., 5-15 meters above the ground). The 2 km dataset consists of time synchronized aerial high-resolution images, GPS and IMU sensor data, ground-level street view images, and ground truth data. The dataset is ideal to evaluate and benchmark our 3-D Invariant EKF algorithm by directly comparing our estimation result to the ground truth data.

B. Results and discussions

The InEKF produced a combined estimate of the MAV pose from the IMU and GPS data. Figure 2 shows the ground truth positions in blue, GPS positions in green, and the filter's positions and orientation shown the the red, green, and blue triad. After a poor initialization, the filter is able to follow the ground truth data reasonably closely.

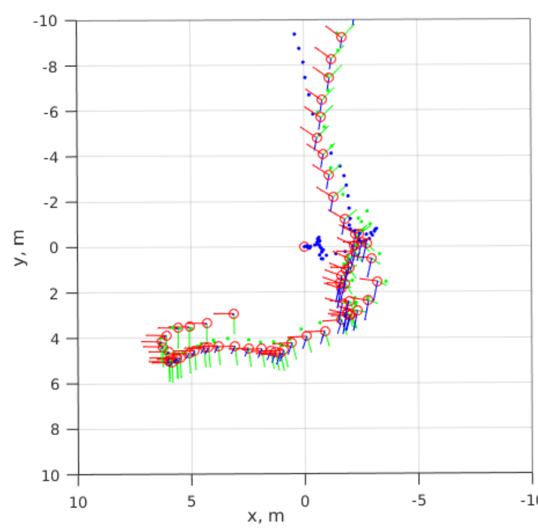


Fig. 2: Pose estimates for first 200 seconds, bird's eye view

This estimate was compared to the ground truth by finding the root mean squared pointwise distance between the corresponding ground truth and estimated poses. The uncorrected gps data had an rms distance of 7.0 m. Our best result with filter had a rms distance of 7.6 m. This result was unexpected, at it was anticipated that the additional IMU data should reduce, not increase, the rms error.

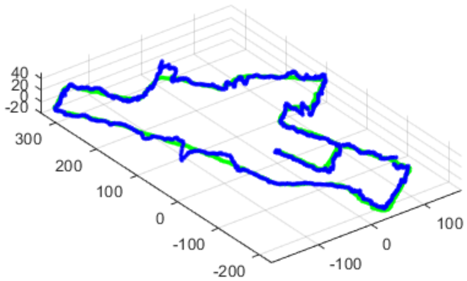


Fig. 3: Comparison between GT and raw GPS

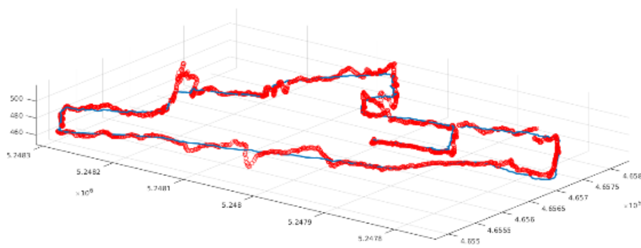


Fig. 4: Comparison between GT and InEKF estimated pose

Figures 3 and 4 show the raw GPS and filter estimate plotted with the ground truth data. The filter estimated pose closely follows the GPS estimate, giving little weight to the IMU provided data.

IV. Conclusion

We were able to successfully implement an Extended Invariant Kalman Filter to estimate the 3D position of an unmanned aerial vehicle by using IMU measurement data for dynamics prediction and GPS measurements for correction. We used a root mean square distance metric to evaluate the performance of our filter against the ground truth dataset. Unfortunately, we were unable to improve upon the baseline rms distance from the GPS data alone, which was 7.0 m. Our best rms distance with filtering was 7.6 m. In the future, this result could be improved by estimating IMU biases, and by using GPS measurements that contain measurement-wise variance estimation, rather than applying the same variance estimate to all collected GPS data.

References

[1] Ross Hartley, Maani Ghaffari Jadidi, Jessy W. Grizzle, and Ryan M. Eustice, "Contact-Aided Invariant Extended Kalman Filtering for Legged Robot State Estimation"

[2] Data Resource: The Zurich Urban Micro Aerial Vehicle Dataset for Appearance-based Localization, Visual Odometry, and SLAM, http://rpg.ifi.uzh.ch/zurichmavdataset.html

[3] Parameter Resource: https://dev.px4.io/v1.9.0/en/advanced/parameter_reference.html

[4] Andras L. Majdik Charles Till and Davide Scaramuzza, "The Zurich Urban Micro Aerial Vehicle Dataset", Journal Title XX(X):1–5 c The Author(s) 2016 Reprints and permission: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/ToBeAssigned www.sagepub.com/