A Machine Learning Approach to Inertial Reference Unit Calibration

Jeff Pinner - SUID 5215534
CS229 Final Project

1. Introduction

In spacecraft attitude control systems, the accuracy of the alignment and calibration of the various attitude and rate sensors is critical to achieving high pointing accuracy. An inertial reference unit (IRU) is an inertial sensor that is capable of measuring the change in angular direction about one or more 'sense axes.' The calibration of the IRU takes on even greater performance implications during those time periods when absolute orientation measurements are unavailable and the spacecraft must propagate its attitude based on IRU measurements alone. This project investigates the application of machine learning techniques to the estimation of IRU calibration parameters critical to achieving high spacecraft attitude control performance.

2. Attitude Measurement Model

It is presumed that the spacecraft is capable of directly measuring its absolute attitude with respect to inertial space, providing a sequence of attitude quaternions $q^{(i)}$ describing the orientation of the body frame in inertial space at discrete time intervals. The simplifying assumption is made that the spacecraft motion between the attitude measurements can be described by a simple rotation. This change in orientation in the body frame $\theta_b^{(i)}$ is described by:

$$\delta q^{(i)} = q^{(i)}(q^{(i-1)})^{-1}$$
$$\delta q^{(i)} = \cos\frac{\alpha^{(i)}}{2} + \vec{u}^{(i)}\sin\frac{\alpha(i)}{2}$$
$$\theta_b^{(i)} = \alpha^{(i)}\vec{u}^{(i)}$$

If the attitude measurements are produced, for example, by a star tracker, then they will contain errors due to misalignments, noise, errors in the star catalogue, etc. For simplicity, these errors are ignored and the measurements are assumed to be exact.

3. Inertial Reference Unit Calibration Model

The IRU measures the incremental angle change about its sense axis. Defining the change in orientation in the body frame over a discrete time interval $T = t^{(i)} - t^{(i-1)}$ as $\theta_b^{(i)}$, the IRU outputs a number of pulses $c^{(i)}$ given by the projection of orientation change onto the gyro's sense axis (given by the unit vector w):

$$sc^{(i)} = (1 - \lambda - \mu\,\mathrm{sign}(w\cdot\theta_b^{(i)}))w\cdot\theta_b^{(i)} - Tb^{(i)} - \int_T \eta(t)dt$$

where s is the assumed scale factor, λ and μ are the symmetric and asymmetric scale factor errors, $b^{(i)}$ is the bias over time interval T, and η is white noise.

4. Linear Dynamical System Model

The gyro alignment and calibration parameters (w, s, λ, and μ) are assumed to be constant. The angle white noise of the gyro output, η, is assumed to be Gaussian with a variance σ_η^2. The gyro bias is assumed to vary according to a Gaussian rate random walk process driven by white noise, η_{rrw}, with variance $\sigma_{rrw}^2 = \frac{2}{\tau_s}\sigma_s^2$ where σ_s is the bias stability and τ_s is the stability interval. We define the linear dynamical system model of the bias as:

$$b^{(i+1)} = b^{(i)} + \eta_{rrw}$$
$$y^{(i)} = Tb^{(i)} + \eta$$

where $y^{(i)}$ is the difference between the measured body rate and the estimated body rate about the sense axis from the IRU pulses $c^{(i)}$:

$$y^{(i)} = (1 - \lambda - \mu\,\mathrm{sign}(w\cdot\theta_b^{(i)}))w\cdot\theta_b^{(i)} - sc^{(i)}$$

5. Expectation-Maximization Algorithm for Parameter Estimation

The gyro alignment and calibration parameters can be estimated by applying an expectation-maximization algorithm to the linear dynamical system describing the bias perturbation. Given initial estimates for the alignment and calibration parameters (the assumed sense axis w, and the assumed scale factor s, λ, and μ), the following estimation and maximization steps are repeated until convergence.

Estimation Step:
Based on the linear dynamical system model, we can write the conditional densities for the state and output variables:

$$p(y^{(i)}|b^{(i)};w,s,\lambda,\mu,\sigma_\eta) = N(Tb^{(i)},\sigma_\eta^2)$$
$$p(b^{(i)}|b^{(i-1)};\sigma_{rrw}) = N(b^{(i-1)},\sigma_{rrw}^2)$$

From the Markov property implicit in the model, the joint probability of a bias and output sequence $p(\{b\},\{y\})$ is given by:

$$p(\{b\},\{y\}) = p(b^{(1)})\prod_{i=2}^{m} p(b^{(i)}|b^{(i-1)})\prod_{i=1}^{m} p(y^{(i)}|b^{(i)})$$

The sequence of biases that maximizes $\log p(\{b\},\{y\}|\{y\})$ is found by implementing a forward-backward algorithm over the linear dynamical system, using the Kalman filter equations as the forward step and Rauch's recursion as the backward step. Adopting the notation $\widehat{b}_{i|j} = E[b^{(i)}|y^{(1)},\ldots,y^{(j)}]$ and $P_{i|j} = \mathrm{var}(b^{(i)}|y^{(1)},\ldots,y^{(j)})$, the Kalman filter equations compute $\widehat{b}_{i|i}$ according to:

$$\begin{aligned}
\widehat{b}_{i|i-1} &= \widehat{b}_{i-1|i-1} \\
P_{i|i-1} &= P_{i-1|i-1} + \sigma_{rrw}^2 \\
K^{(i)} &= P_{i|i-1}T(TP_{i|i-1}T + \sigma_{awn})^{-1} \\
\widehat{b}_{i|i} &= \widehat{b}_{i|i-1} + K^{(i)}(y^{(i)} - T\widehat{b}_{i|i-1}) \\
P_{i|i} &= P_{i|i-1} + K^{(i)}TP_{i|i-1}
\end{aligned}$$

Rauch's recursion is now used to compute $\widehat{b}_{i|m}$ according to:

$$\begin{aligned}
\widetilde{K}^{(i)} &= P_{i|i}P_{i+1|i}^{-1} \\
\widehat{b}_{i|m} &= \widehat{b}_{i|i} + \widetilde{K}^{(i)}(\widehat{b}_{i+1|m} - \widehat{b}_{i|i})
\end{aligned}$$

Maximization Step:
Use the computed biases and gradient descent to estimate the parameters of the model. Defining the measurements as:

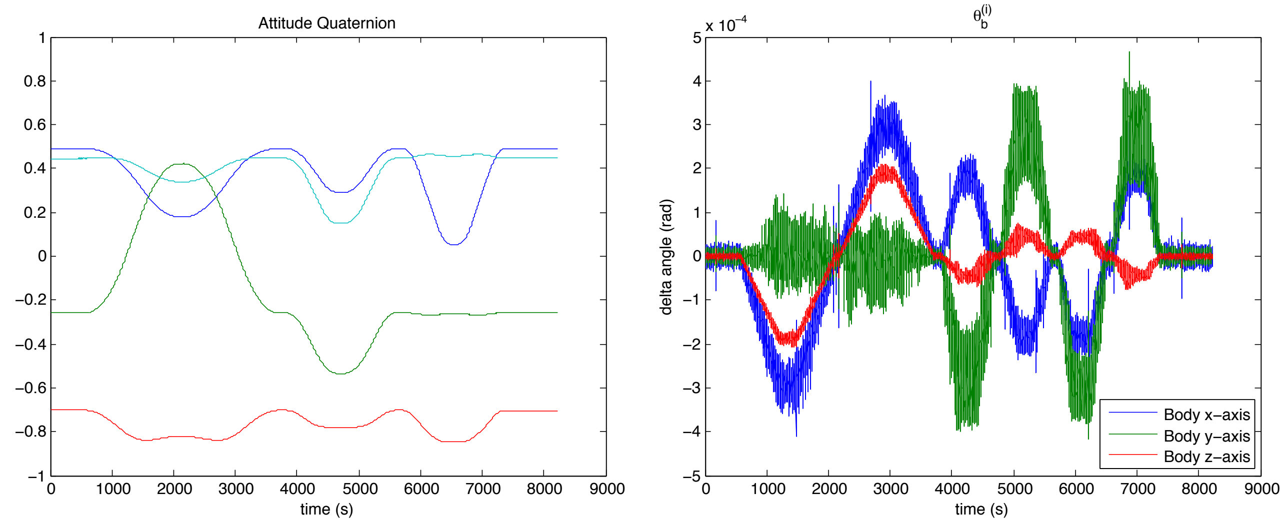
$$z^{(i)} = sc^{(i)} + T\widehat{b}_{i|m}$$

and the hypothesis of the model to be:

$$\begin{aligned}
\theta_g^{(i)} &= (w - \delta_u u - \delta_v v)\cdot\theta_b^{(i)} \\
h^{(i)} &= (1 - \lambda - \mu\,\mathrm{sign}(\theta_g^{(i)})\theta_g^{(i)}
\end{aligned}$$

where $[u,v,w]$ form a right handed orthogonal basis, the updates to the model parameters are given by:

$$\delta_u := \delta_u + \alpha\sum_{i=1}^{m}(h^{(i)} - z^{(i)})(1 - \lambda - \mu\,\mathrm{sign}(\theta_g^{(i)})u\cdot\theta_b^{(i)}$$
$$\delta_v := \delta_v + \alpha\sum_{i=1}^{m}(h^{(i)} - z^{(i)})(1 - \lambda - \mu\,\mathrm{sign}(\theta_g^{(i)})v\cdot\theta_b^{(i)}$$
$$s_+ := s_+ + \alpha\sum_{i=1}^{m}1\{\mathrm{sign}(\theta_g^{(i)}) = +1\}(h^{(i)} - z^{(i)})\theta_g^{(i)}$$
$$s_- := s_- + \alpha\sum_{i=1}^{m}1\{\mathrm{sign}(\theta_g^{(i)}) = -1\}(h^{(i)} - z^{(i)})\theta_g^{(i)}$$



(a) Body Frame Orientation in J2000

(b) Orientation Change in Body Frame

Fig. 1. Calibration Maneuver Attitude Profile

From the updates, the new alignment and scale factor errors can be computed:

$$
\begin{aligned}
s_+ &:= 1 - (1 - s_+)\,||w - \delta_u u - \delta_v v||_2 \\
s_- &:= 1 - (1 - s_-)\,||w - \delta_u u - \delta_v v||_2 \\
w &:= \frac{w - \delta_u u - \delta_v v}{||w - \delta_u u - \delta_v v||_2} \\
\lambda &:= \frac{s_+ + s_-}{2} \\
\mu &:= \frac{s_+ - s_-}{2}
\end{aligned}
$$

6. Results

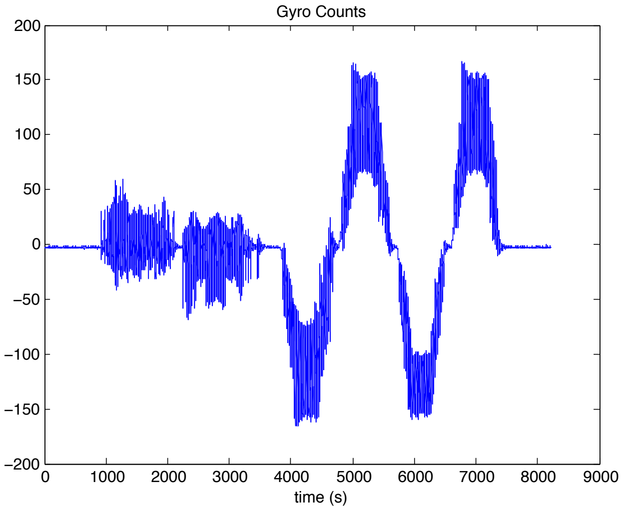
The EM algorithm was applied to attitude and IRU data. Calibration results are provided below for one gyro sense axes.

$$
\begin{aligned}
w &= \begin{bmatrix} 0.000983072276832 \\ 0.999999242473548 \\ -0.000740689698082 \end{bmatrix} \\
s &= 2.386393534388658 \times 10^{-6} \text{ rad / pulse} \\
\lambda &= 0.006789928965283 \\
\mu &= 7.437154236313304 \times 10^{-4}
\end{aligned}
$$

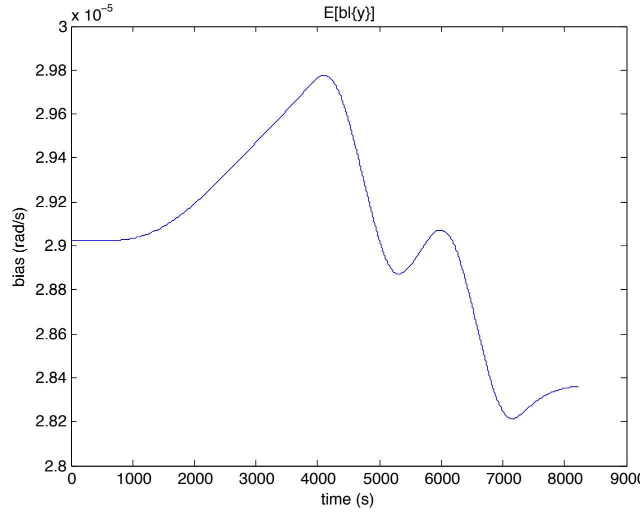
Figure 1 depicts the attitude quaternions and the derived change in orientation of the body frame, $\theta_b^{(i)}$, during .2 second intervals. Figure 2 shows the number of gyro pulses for the calibrated axis during the same intervals and the converged estimate of the gyro bias, $\widehat{b}_{i|m}$. The convergence behavior of the algorithm can be visualized by plotting the value of the cost function $J = \frac{1}{2}\sum_{i=1}^{m}(h^{(i)} - z^{(i)})^2$ at each iteration in the maximization step of the algorithm (given in Figure 3). The sharp discontinuities are due to the recalculation of $\widehat{b}_{i|m}$ during the expectation step. The algorithm has mostly converged after two iterations of the EM algorithm (corresponding to around 400,000 iterations of the maximization step).

7. Conclusions

This approach to IRU calibration provides an effective way of estimating the gyro alignment and calibration parameters without requiring complete IRU observations of the spacecraft rotation (measurements about three independent axes). This method also allows characterization of the bias stability and scale factor errors on a per axis basis, independent of how the spacecraft combines multiple gyro measurements within its attitude control system. This independence is helpful

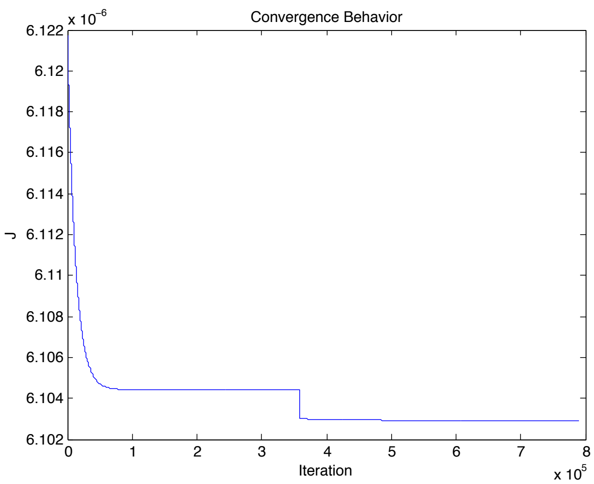


(a) Number of Gyro Pulses

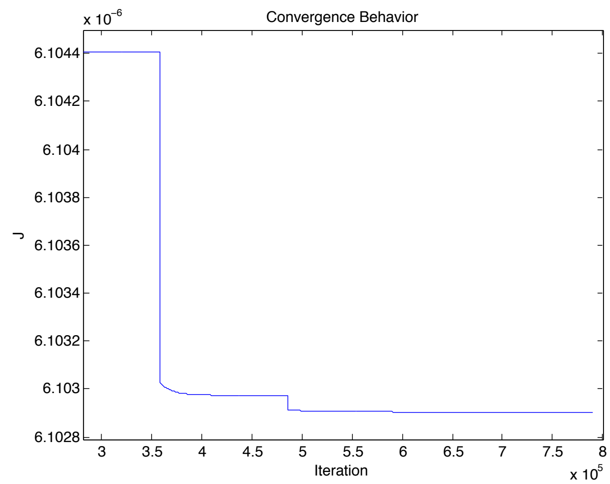


(b) EM Derived Gyro Bias

Fig. 2. Calibration Maneuver Gyro Behavior



(a) Convergence of Complete Algorithm



(b) Detail of Convergence after 300000 Iterations

Fig. 3. Parameter Estimation Algorithm Convergence Behavior

in decoupling performance characteristics of the hardware with those of the control system software.

8. References

Ng, Andrew, CS229 Lecture Notes, http://www.stanford.edu/class/cs229/materials.html, 2010.

Shumway, R.H. and Stoffer, D.S., An Approach to Time Series Smoothing and Forcasting Using the EM Algorithm, Journal of Time Series Analysis, Vol. 3, No. 4, pp 253 - 264, 1982.