Global Exponential Attitude and Gyro Bias Estimation from Vector Measurements

Philippe Martin, Ioannis Sarras

Global exponential attitude and gyro bias estimation from vector measurements

Philippe Martin[1] and Ioannis Sarras[2]

[1] Centre Automatique et Systèmes, MINES ParisTech, PSL Research University, 75006 Paris, France
philippe.martin@mines-paristech.fr
[2] ONERA – The French Aerospace Lab, F-91123 Palaiseau, France
ioannis.sarras@onera.fr

Abstract. We consider the classical problem of estimating the attitude and gyro biases of a rigid body from at least two vector measurements and a triaxial rate gyro. We propose a solution based on a dynamic nonlinear estimator designed without respecting the geometry of SO(3), which achieves uniform global exponential convergence. The convergence is established thanks to a dynamically scaled Lyapunov function.

1 Introduction

Estimating the attitude of a rigid body from vector measurements (obtained for instance from accelerometers, magnetometers, sun sensors, etc.) has been for decades a problem of interest, because of its importance for a variety of technological applications such as satellites or unmanned aerial vehicles. The attitude of the body can be described by the rotation matrix $R \in \mathrm{SO}(3)$ from body to inertial axes. On the other hand, the (time-varying) measurement vectors $u_1, \cdots, u_n \in \mathbb{R}^3$ correspond to the expression in body axes of known and not all collinear vectors $U_1, \cdots, U_n \in \mathbb{R}^3$ which are constant in inertial axes, i.e., $u_k(t) = R^T(t)U_k$. The goal then is to reconstruct the attitude at time t using only the knowledge of the measurement vectors until t. The solution to the problem would be very easy if the vector measurements were perfect and two of them were linearly independent: indeed, using for instance only the two vectors $u_1(t)$ and $u_2(t)$ and noticing that $R^T(x \times y) = R^T x \times R^T y$ since R is a rotation matrix, we readily find

$$\begin{aligned} R^T(t) &= R^T(t) \cdot \big(U_1\ U_2\ U_1 \times U_2\big) \cdot \big(U_1\ U_2\ U_1 \times U_2\big)^{-1} \\ &= \big(u_1(t)\ u_2(t)\ u_1(t) \times u_2(t)\big) \cdot \big(U_1\ U_2\ U_1 \times U_2\big)^{-1}. \end{aligned}$$

But in real situations, the measurement vectors are always corrupted at least by noise. Moreover, the U_k's may possibly be not strictly constant: for instance a triaxial magnetometer measures the (locally) constant Earth magnetic field, but is easily perturbed by ferromagnetic masses and electromagnetic perturbations; similarly, a triaxial accelerometer can be considered as measuring the direction of gravity provided it is not undergoing a substantial acceleration (see e.g. [13] for a detailed discussion of this assumption and its consequences in the framework of quadrotor UAVs). That is why, despite the additional cost, it may be interesting to use a triaxial rate gyro to supplement the possibly deteriorated vector measurements.

The current literature on attitude estimation from vector measurements can be broadly divided into three categories: i) optimization-based methods; ii) stochastic filtering; iii) nonlinear observers. Details on the various approaches can be found e.g. in the surveys [6, 17] and the references therein. The first category sets the problem as the minimization of a cost function, and is usually referred to as Wahba's problem. The attitude is algebraically recovered at time t using only the measurements at time t. No filtering is performed, and possibly available velocity information from rate gyros is not exploited. The second category mainly hinges on Kalman filtering and its variants. Despite their many qualities, the drawback of those designs is that convergence cannot in general be guaranteed except for mild trajectories. Moreover the tuning is not completely obvious, and the computational cost may be too high for small embedded processors. The third, and more recent, approach proposes nonlinear observers with a large guaranteed domain of convergence and a rather simple tuning through a few constant gains. These observers can be designed: a) directly on SO(3) (or the unit quaternion space), see e.g. [7, 10, 12, 16]; b) or more recently, on $\mathbb{R}^{3\times 3}$, i.e., deliberately "forgetting" the underlying geometry [2, 3, 8, 14]. Probably the best-known design is the so-called nonlinear complementary filter of [10]; as noticed in [11], it is a special case of so-called invariant observers [5].

In this paper, we propose a new observer of attitude and gyro biases from gyro measurements and (at least) two measurement vectors. It also "forgets" the geometry of SO(3), which allows for uniform global exponential convergence (notice the observer of [10] is only quasi-globally convergent). This observer is an extension of the observer of [14] (which is uniformly globally convergent), itself a modification of the linear cascaded observer of [3] (which is uniformly globally exponentially convergent). The idea of the proof is nevertheless completely different from the approach followed in [3]; it is much more direct, as it relies on a strict, dynamically scaled, Lyapunov function, see [1, 9].

2 The design model

We consider a moving rigid body subjected to an angular velocity ω. Its orientation matrix $R \in \mathrm{SO}(3)$ is related to the angular velocity by the differential equation

$$\dot{R} = R\omega_\times, \tag{1}$$

where the skew-symmetric matrix $\omega_\times$ is defined by $\omega_\times u := \omega \times u$ whatever the vector $u \in \mathbb{R}^3$.

The rigid body is equipped with a triaxal rate gyro measuring the angular velocity ω, and two additional triaxial sensors (for example accelerometers, magnetometers or sun sensors) providing the measurements of two vectors α and β. These vectors correspond to the expression in body axes of two known independent vectors α_i and β_i which are constant in inertial axes. In other words,

$$\alpha := R^T \alpha_i$$
$$\beta := R^T \beta_i.$$

Since α_i, β_i are constant, we obviously have

$$\dot{\alpha} = \alpha \times \omega$$
$$\dot{\beta} = \beta \times \omega.$$

To take full advantage of the rate gyro, it is wise to take into account that it is biased, hence rather provides the measurement

$$\omega_m := \omega + b,$$

where b is a slowly-varying (for instance with temperature) unknown bias. Since the effect of this bias on attitude estimation may be important, it is worth determining this value. But being not exactly constant, it can not be calibrated in advance and must be estimated online together with the attitude.

Our objective is to design an estimation scheme that can reconstruct online the orientation matrix $R(t)$ and the bias $b(t)$, using i) the measurements of the gyro and of the two vector sensors; ii) the knowledge of the constant vectors α_i and β_i. The model on which the design will be based therefore consists of the dynamics

$$\dot{\alpha} = \alpha \times \omega \tag{2}$$
$$\dot{\beta} = \beta \times \omega \tag{3}$$
$$\dot{b} = 0, \tag{4}$$

together with the measurements

$$\omega_m := \omega + b \tag{5}$$
$$\alpha_m := \alpha \tag{6}$$
$$\beta_m := \beta. \tag{7}$$

3 The observer

We want to show that the state of (2)–(7) can be estimated by the observer

$$\dot{\hat{\alpha}} = \hat{\alpha} \times (\omega_m - \hat{b}) - k_\alpha(\hat{\alpha} - \alpha_m) \tag{8}$$
$$\dot{\hat{\beta}} = \hat{\beta} \times (\omega_m - \hat{b}) - k_\beta(\hat{\beta} - \beta_m) \tag{9}$$
$$\dot{\xi} = l_\alpha(\omega_m - \hat{b}) \times (\hat{\alpha} \times \alpha_m) + l_\beta(\omega_m - \hat{b}) \times (\hat{\beta} \times \beta_m) \tag{10}$$
$$+ l_\alpha k_\alpha \hat{\alpha} \times \alpha_m + l_\beta k_\beta \hat{\beta} \times \beta_m \tag{11}$$
$$\dot{r} = -2\psi_1(r-1) + 2\big(l_\alpha |\alpha_m||\hat{\alpha} - \alpha_m| + l_\beta |\beta_m||\hat{\beta} - \beta_m|\big) r, \tag{12}$$

where

$$\hat{b} := \xi + l_\alpha \hat{\alpha} \times \alpha_m + l_\beta \hat{\beta} \times \beta_m \tag{13}$$

$$k_\alpha := k_1 + r\Big(\frac{1}{2\epsilon} + \frac{l_\alpha^2}{\epsilon_1} r\Big)|\alpha_m|^2 \tag{14}$$

$$k_\beta := k_2 + r\Big(\frac{1}{2\epsilon} + \frac{l_\beta^2}{\epsilon_1} r\Big)|\beta_m|^2. \tag{15}$$

$\hat{\alpha}, \hat{\beta}, \hat{b} \in \mathbb{R}^3$ are the estimates of α, β, b; $\xi \in \mathbb{R}^3$ is the state of the bias observer, and $r \in \mathbb{R}$ is a dynamic scaling variable; the (positive) constants $l_\alpha, l_\beta, \psi_1, k_1, k_2, \epsilon, \epsilon_1$ are tuning gains. Defining the estimation errors as

$$\begin{aligned} e_\alpha &:= \hat{\alpha} - \alpha \\ e_\beta &:= \hat{\beta} - \beta \\ e_b &:= \hat{b} - b, \end{aligned}$$

the error system reads

$$\dot{e}_\alpha = e_\alpha \times \omega - (\alpha + e_\alpha) \times e_b - k_\alpha(r, \hat{\alpha}) e_\alpha \tag{16}$$

$$\dot{e}_\beta = e_\beta \times \omega - (\beta + e_\beta) \times e_b - k_\beta(r, \hat{\beta}) e_\beta \tag{17}$$

$$\dot{e}_b = (l_\alpha \alpha_\times^2 + l_\beta \beta_\times^2) e_b + l_\alpha e_\alpha \times (\alpha \times e_b) + l_\beta e_\beta \times (\beta \times e_b) \tag{18}$$

$$\dot{r} = -2\psi_1(r-1) + 2\big(l_\alpha|\alpha||e_\alpha| + l_\beta|\beta||e_\beta|\big)r; \tag{19}$$

(18) is obtained thanks to the Jacobi identity $a \times (b \times c) + b \times (c \times a) + c \times (a \times b)$.
The main result is the global exponential convergence of the observer.

Theorem 1. *Assume $k_1, k_2, \epsilon, \epsilon_1 > 0$, $\psi_1 > \epsilon_1$, and l_α, l_β large enough so that $-(l_\alpha \alpha_\times^2 + l_\beta \beta_\times^2) > (\psi_1 + \epsilon) I$. Then the equilibrium point $(\bar{e}_\alpha, \bar{e}_\beta, \bar{e}_b, \bar{r}) := (0, 0, 0, 1)$ of the error system* (16)–(19) *is uniformly globally exponentially stable.*

Remark 1 (see [4, 15]). Since α and β are linearly independent, $-(l_\alpha \alpha_\times^2 + l_\beta \beta_\times^2)$ is a (symmetric) positive definite matrix when $l_\alpha, l_\beta > 0$; moreover, sufficiently large l_α, l_β yield $-(l_\alpha \alpha_\times^2 + l_\beta \beta_\times^2) > \mu I$ whatever the given constant μ.

Proof. First consider the candidate Lyapunov function for the (e_α, e_β)-subsystem

$$V(e_\alpha, e_\beta) := \frac{1}{2}|e_\alpha|^2 + \frac{1}{2}|e_\beta|^2.$$

Its time derivative satisfies

$$\begin{aligned} \dot{V} &= -\langle e_\alpha, e_\alpha \times \omega \rangle - k_\alpha |e_\alpha|^2 - \langle e_\beta, e_\beta \times \omega \rangle - k_\beta |e_\beta|^2 \\ &\leq -k_\alpha |e_\alpha|^2 - k_\beta |e_\beta|^2 + \big(\sqrt{r}|\alpha||e_\alpha|\big)\frac{|e_b|}{\sqrt{r}} + \big(\sqrt{r}|\beta||e_\beta|\big)\frac{|e_b|}{\sqrt{r}} \\ &\leq -\Big(k_\alpha - \frac{r|\alpha|^2}{2\epsilon}\Big)|e_\alpha|^2 - \Big(k_\beta - \frac{r|\beta|^2}{2\epsilon}\Big)|e_\beta|^2 + \frac{\epsilon |e_b|^2}{r}; \end{aligned}$$

where we have used $\langle a, a \times b\rangle = 0$ to obtain the first line, and Young's inequality $ab \leq \frac{a^2}{2\epsilon} + \frac{\epsilon b^2}{2}$ to obtain the second line.

Now, the obvious candidate Lyapunov function $V_b(e_b) := \frac{1}{2}|e_b|^2$ for the e_b-subsystem satisfies

$$\begin{aligned}\dot{V}_b &= \langle e_b, (l_\alpha \alpha_\times^2 + l_\beta \beta_\times^2) e_b\rangle + l_\alpha \langle e_b, e_\alpha \times (\alpha \times e_b)\rangle + l_\beta \langle e_b, e_\beta \times (\beta \times e_b)\rangle \\ &\leq -\mu |e_b|^2 + \big(l_\alpha |\alpha||e_\alpha| + l_\beta |\beta||e_\beta|\big)|e_b|^2,\end{aligned}$$

where we have used Remark 1. The term $\big(l_\alpha |\alpha||e_\alpha| + l_\beta |\beta||e_\beta|\big)|e_b|^2$ happens to be very difficult to dominate with a classical Lyapunov approach. To overcome the problem, we use instead the candidate Lyapunov function

$$\tilde{V}_b(e_b, r) := \frac{1}{2r}|e_b|^2,$$

obtaining by dynamically scaling V_b with r defined by (19). Notice $r(t) \geq 1$ for all positive t as soon as $r(0) \geq 1$. We then have

$$\begin{aligned}\dot{\tilde{V}}_b &:= \frac{\dot{V}_b}{r} - \tilde{V}_b \frac{\dot{r}}{r} \\ &\leq -\mu \frac{|e_b|^2}{r} + \big(l_\alpha |\alpha||e_\alpha| + l_\beta |\beta||e_\beta|\big)\frac{|e_b|^2}{r} - \frac{|e_b|^2}{2r}\frac{\dot{r}}{r} \\ &= -(\mu - \psi_1)\frac{|e_b|^2}{r},\end{aligned}$$

where we have used $\frac{r-1}{r} \leq 1$.

We next consider the candidate Lyapunov function for the r-subsystem

$$V_r(r) := \frac{1}{2}(r-1)^2.$$

Its time derivative satisfies

$$\begin{aligned}\dot{V}_r &= -2\psi_1 (r-1)^2 + \sqrt{2}(r-1)\sqrt{2} r l_\alpha |\alpha||e_\alpha| + \sqrt{2}(r-1)\sqrt{2} r l_\beta |\beta||e_\beta| \\ &\leq -2(\psi_1 - \epsilon_1)(r-1)^2 + \frac{r^2}{\epsilon_1}\big(l_\alpha^2 |\alpha|^2 |e_\alpha||^2 + l_\beta^2 |\beta||e_\beta|^2\big),\end{aligned}$$

where the second line is obtained by Young's inequality.

Finally, consider the complete Lyapunov function

$$W(e_\alpha, e_\beta, e_b, r) := V(e_\alpha, e_\beta) + \tilde{V}_b(e_b, r) + V_r(r).$$

Collecting all the previous findings, its time derivative satisfies

$$\begin{aligned}\dot{W} \leq& -\Big(k_\alpha - \frac{r|\alpha|^2}{2\epsilon}\Big)|e_\alpha|^2 - \Big(k_\beta - \frac{r|\beta|^2}{2\epsilon}\Big)|e_\beta|^2 + \frac{\epsilon |e_b|^2}{r} \\ & - (\mu - \psi_1)\frac{|e_b|^2}{r} \\ & -2\,(\psi_1 - \epsilon_1)(r-1)^2 + \frac{r^2}{\epsilon_1}\big(l_\alpha^2 |\alpha|^2 |e_\alpha|^2 + l_\beta^2 |\beta||e_\beta|^2\big) \\ =& -k_1 |e_\alpha|^2 - k_2 |e_\beta|^2 - (\mu - \psi_1 - \epsilon)\frac{|e_b|^2}{r} - 2(\psi_1 - \epsilon_1)(r-1)^2.\end{aligned}$$

Choosing $k_1, k_2 > 0$, $\psi_1 > \epsilon_1$, and l_α, l_β large enough so that $\mu > \psi_1 + \epsilon$ clearly guarantees the uniform global exponential stability of the equilibrium point $(\bar{e}_\alpha, \bar{e}_m, \frac{\bar{e}_b}{\sqrt{\bar{r}}}, \bar{r}) := (0,0,0,1)$, hence of $(\bar{e}_\alpha, \bar{e}_m, \bar{e}_b, \bar{r}) := (0,0,0,1)$. □

Remark 2. More than two vectors α and β can be used with a direct generalization of the proposed structure.

Remark 3. The observer does not use the knowledge of the constant vectors α_i and β_i. This may be an interesting feature in some applications when those vectors for example are not precisely known and/or (slowly) vary.

We then have the following corollary, which gives an estimate of the true orientation matrix R by using the knowledge of the inertial vectors α_i and β_i. Notice it is considerably simpler than the approach of [3], where the estimated orientation matrix is obtained through an additional observer of dimension 9.

Corollary 1. *Under the assumptions of Theorem 1, the matrix $\tilde{R}$ defined by*

$$\tilde{R}^T := \begin{pmatrix} \frac{\hat{\alpha}}{|\alpha_i|} & \frac{\hat{\alpha}\times\hat{\beta}}{|\alpha_i\times\beta_i|} & \frac{\hat{\alpha}\times(\hat{\alpha}\times\hat{\beta})}{|\alpha_i\times(\alpha_i\times\beta_i)|} \end{pmatrix} \cdot R_i^T$$

$$R_i := \begin{pmatrix} \frac{\alpha_i}{|\alpha_i|} & \frac{\alpha_i\times\beta_i}{|\alpha_i\times\beta_i|} & \frac{\alpha_i\times(\alpha_i\times\beta_i)}{|\alpha_i\times(\alpha_i\times\beta_i)|} \end{pmatrix}$$

uniformly globally exponentially converges to R.

Proof. By Theorem 1, $e_\alpha(t) \le C|e_\alpha(0)|e^{-\lambda t}$ and $e_\beta(t) \le C|e_\beta(0)|e^{-\lambda t}$ for some $C, \lambda > 0$. Therefore,

$$\begin{aligned} |\hat{\alpha}\times\hat{\beta} - \alpha\times\beta| &= |\alpha\times e_\beta + e_\alpha\times\beta + e_\alpha\times e_\beta| \\ &\le |\alpha||e_\beta| + |\beta||e_\alpha| + |e_\alpha||e_\beta| \\ &\le C|\alpha_i||e_\beta(0)|e^{-\lambda t} + C|\beta_i||e_\alpha(0)|e^{-\lambda t} + C^2|e_\alpha(0)||e_\beta(0)|e^{-\lambda 2t}; \end{aligned}$$

a similar bound is readily obtained for $|\hat{\alpha}\times(\hat{\alpha}\times\hat{\beta}) - \alpha\times(\alpha\times\beta)|$. As a consequence, all the coefficients of the matrix

$$\tilde{R}^T - \begin{pmatrix} \frac{\alpha}{|\alpha_i|} & \frac{\alpha\times\beta}{|\alpha_i\times\beta_i|} & \frac{\alpha\times(\alpha\times\beta)}{|\alpha_i\times(\alpha_i\times\beta_i)|} \end{pmatrix} \cdot R_i^T$$

globally exponentially converge to 0. The claim follows by noticing

$$\begin{aligned} \begin{pmatrix} \frac{\alpha}{|\alpha_i|} & \frac{\alpha\times\beta}{|\alpha_i\times\beta_i|} & \frac{\alpha\times(\alpha\times\beta)}{|\alpha_i\times(\alpha_i\times\beta_i)|} \end{pmatrix} \cdot R_i^T &= \begin{pmatrix} \frac{R^T\alpha_i}{|\alpha_i|} & \frac{R^T\alpha_i\times R^T\beta_i}{|\alpha_i\times\beta_i|} & \frac{R^T\alpha_i\times(R^T\alpha_i\times R^T\beta_i)}{|\alpha_i\times(\alpha_i\times\beta_i)|} \end{pmatrix} \cdot R_i^T \\ &= R^T R_i R_i^T \\ &= R^T, \end{aligned}$$

where we have used $R^T(u\times v) = R^T u \times R^T v$ since R is a rotation matrix. □

Of course, $\tilde{R}^T$ has no reason to be a rotation matrix (it is only asymptotically so); it is nevertheless the product of a matrix with orthogonal (possibly zero) columns by a rotation matrix. If a bona fide rotation matrix is required at all times, a natural idea is to project $\tilde{R}$ onto the "closest" rotation matrix $\hat{R}$, thanks to a polar decomposition.

References

1. A. Astolfi, D.Karagiannis, and R. Ortega. *Nonlinear and adaptive control with applications*. Springer-Verlag, London, 2008.
2. P. Batista, C. Silvestre, and P. Oliveira. A GES attitude observer with single vector observations. *Automatica*, 48(2):388–395, 2012.
3. Pedro Batista, Carlos Silvestre, and Paulo Oliveira. Globally exponentially stable cascade observers for attitude estimation. *Control Engineering Practice*, 20(2):148 – 155, 2012.
4. L. Benziane, A. Benallegue, Y. Chitour, and A. Tayebi. Velocity-free attitude stabilization with inertial vector measurements. *International Journal of Robust and Nonlinear Control*, 2015.
5. S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *IEEE Transactions on Automatic Control*, 53(11):2514–2526, 2008.
6. J. L. Crassidis, F. L. Markley, and Y. Cheng. Survey of nonlinear attitude estimation methods. *Journal of Guidance, Control, and Dynamics*, 30(1):12–28, 2007.
7. H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Attitude estimation using biased gyro and vector measurements with time-varying reference vectors. *IEEE Transactions on Automatic Control*, 57(5):1332–1338, 2012.
8. H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration. *Automatica*, 51:158–166, 2015.
9. D. Karagiannis, M. Sassano, and A. Astolfi. Dynamic scaling and observer design with application to adaptive control. *Automatica*, 45(12):2883–2889, 2009.
10. R. Mahony, T. Hamel, and J. M. Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, 2008.
11. P. Martin and E. Salaün. Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors. In *IEEE Conference on Decision and Control*, pages 1039–1045, 2007.
12. P. Martin and E. Salaün. Design and implementation of a low-cost observer-based attitude and heading reference system. *Control Engineering Practice*, 18(7):712–722, 2010.
13. P. Martin and E. Salaün. The true role of accelerometer feedback in quadrotor control. In *IEEE International Conference on Robotics and Automation*, pages 1623–1629, 2010.
14. P. Martin and I. Sarras. A global attitude and gyro bias observer from vector measurements. *IFAC-PapersOnLine*, 50, 2017. IFAC 2017 World Congress.
15. A. Tayebi, A. Roberts, and A. Benallegue. Inertial vector measurements based velocity-free attitude stabilization. *IEEE Transactions on Automatic Control*, 58(11):2893–2898, 2013.
16. J.F. Vasconcelos, C. Silvestre, and P. Oliveira. A nonlinear observer for rigid body attitude estimation using vector observations. *IFAC Proceedings Volumes*, 41(2):8599 – 8604, 2008.
17. M. Zamani, J. Trumpf, and R. Mahony. Nonlinear Attitude Filtering: A Comparison Study. *ArXiv e-prints*, arXiv:1502.03990 [cs.SY], 2015.