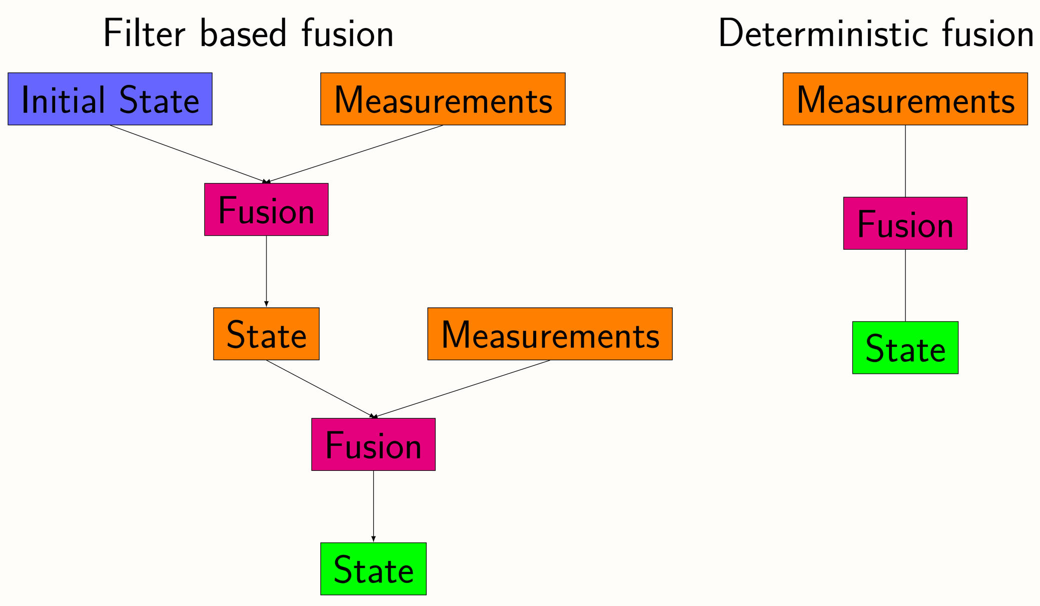
Absolute scale velocity determination combining visual and inertial measurements for drones

Jacques Kaiser, Agostino Martinelli

Team Chroma, INRIA

Introduction

State of the art approaches for visual-inertial sensor fusion are filter based algorithms. These methods are recursive by design and therefore require an initialization. A poor initialization can have a dramatic impact on the performance of the estimations. On the other hand, methods that work without initialization have been developed in computer vision but they can only determine physical quantities up to a scale.
Recently, a Closed-Form Solution that recovers the initial velocity, the roll and pitch angles and the absolute scale by fusing visual and inertial measurements has been derived [1, 2]. While mathematically sound, this method is not robust to noisy sensor data. We added a key step that makes the original method usable in practice.



The Closed-Form Solution

The Closed-Form Solution consists of a single equation that expresses the initial velocity, the distance to point-features and the gravity linearly with respect to the visual and inertial measurements. Note that the knowledge of the gravity in the local frame is equivalent to the knowledge of the roll and pitch angles.

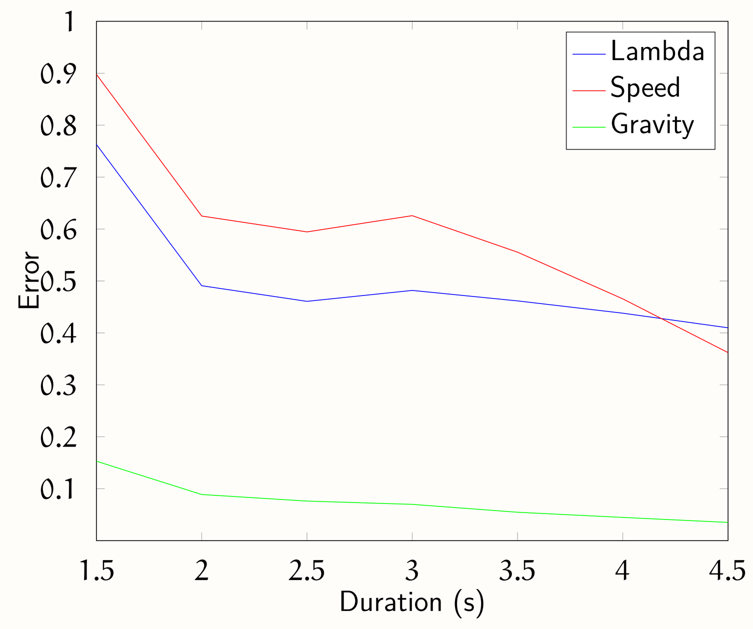
$$S_j = \lambda_1^i \mu_1^i - V t_j - G \frac{t_j^2}{2} - \lambda_j^i \mu_j^i$$

This equation is valid for every observed point-feature i at any time t_j. We therefore have an overconstrained linear system that we can write in matrix formulation:

$$\Xi X = S$$

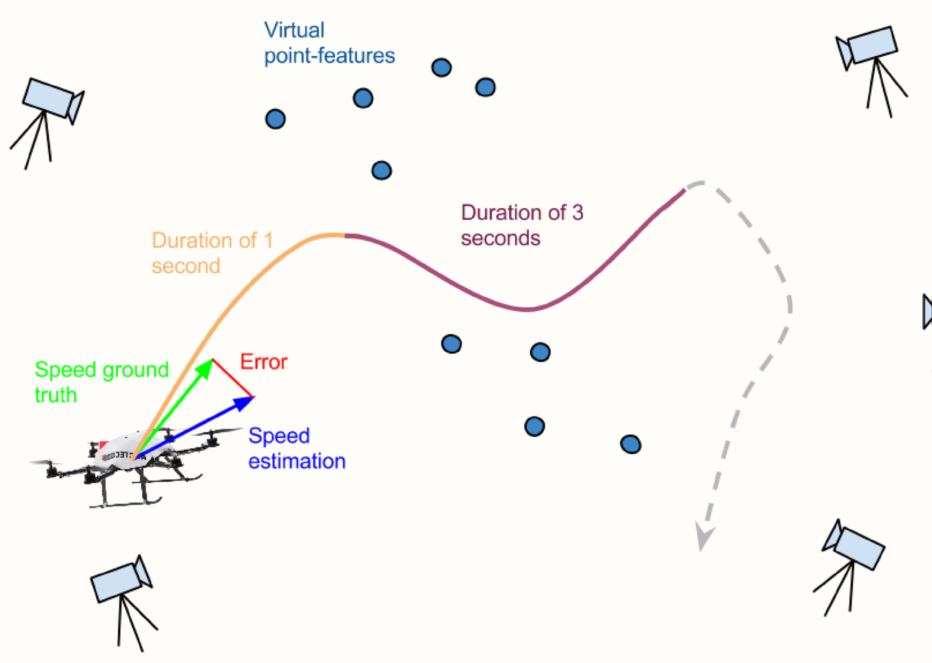
We can therefore recover the initial velocity, the distance to the point-features and the roll and pitch angle of the drone by finding the vector X that satisfies this equation. This method does not require any initialization.

Original performance



With around 45% of error on the speed and the distance to the features after 4 seconds of integration, the Original Closed-Form Solution does not perform very well on terrain data.

Test setup

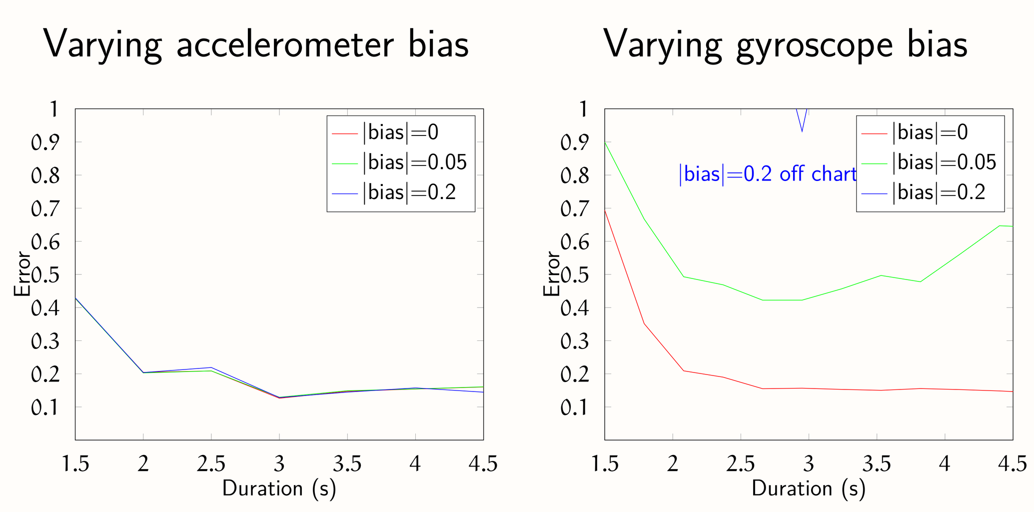


Impact of bias

We studied the impact of biased inertial sensors on the performance of the original Closed-Form Solution. by experimenting with real and synthetic data, we realized that the performance is:

- Strongly affected by the gyroscope bias;
- Weakly affected by the accelerometer bias.

Speed estimation error



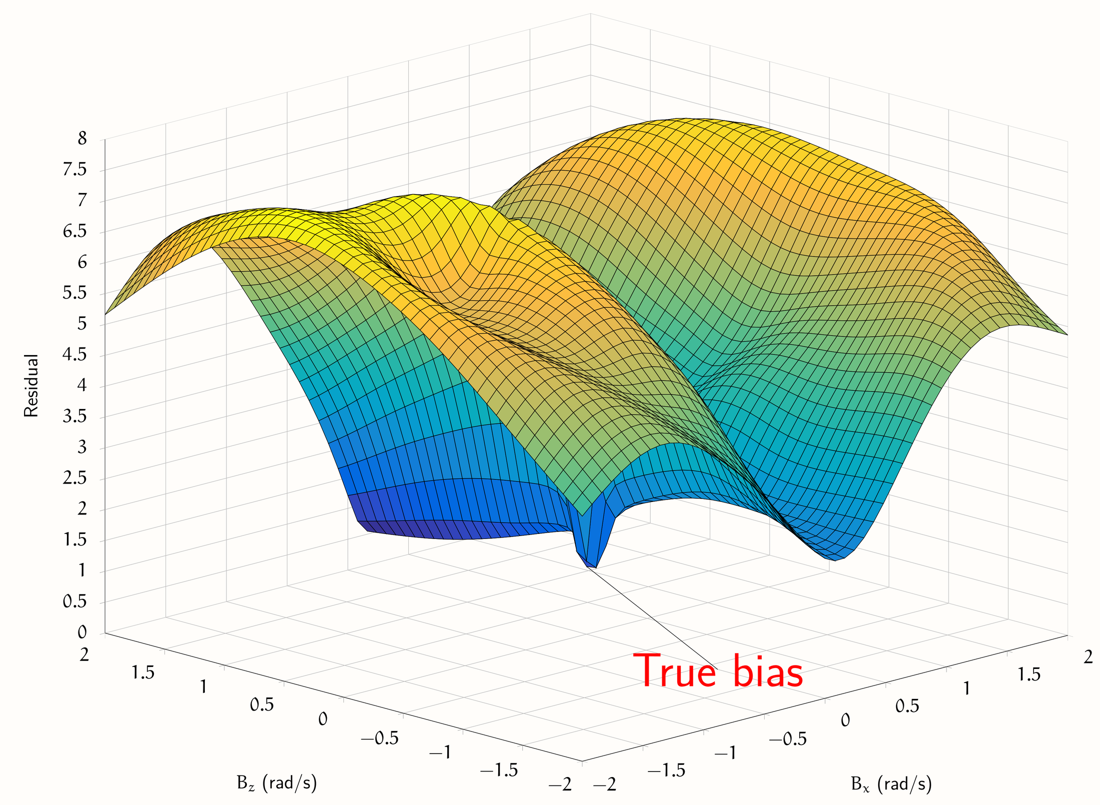
Our method

The goal is to factor out the gyroscope bias to improve the performance of the Closed-Form Solution. Optimally, we would add the gyroscope bias as an unknown term in X. Unfortunately we cannot express the gyroscope bias linearly with respect to the visual-inertial measurements. We propose a different approach: a non-linear minimization of the residual with respect to the gyroscope bias. In other words, we recover X by minimizing the cost function:

$$\text{cost}(B) = \underset{X}{\text{argmin}} \, \|\Xi X - S\|^2$$

With:
B the gyroscope bias,
Ξ and S computed with respect to B

The gyroscope bias is a 3D vector $B = [B_x, B_y, B_z]$. We can plot this cost function with respect to two components of the gyroscope bias:



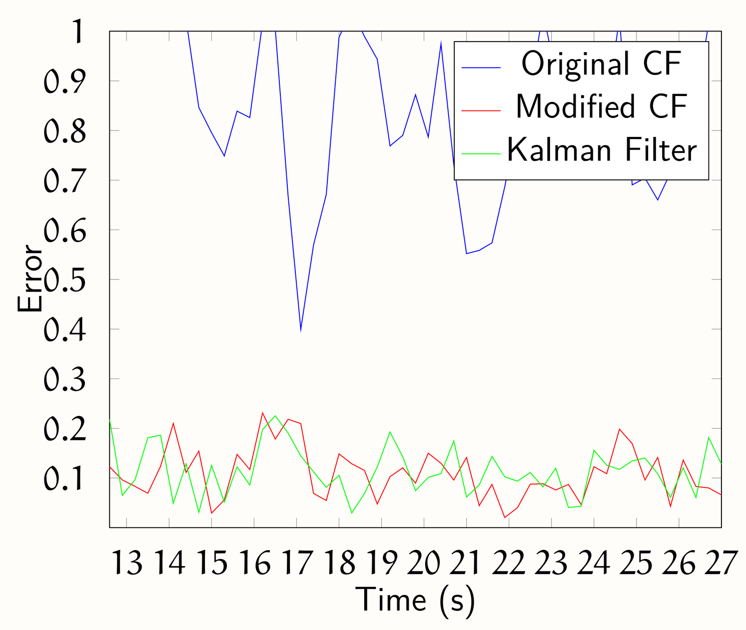
Cost function: residual with respect to B_x and B_z

Note that the residual is almost constant with respect to the third component of the gyroscope bias. This symmetry is induced by the strong weight of the gravity in the system. In order to prevent our optimization to converge towards high and unlikely values of the gyroscope bias, we add a regularization term to our cost function:

$$\text{cost}(B) = \underset{X}{\text{argmin}} \, \|\Xi X - S\|^2 + \lambda \times |B|$$

Results

Speed estimation error over a time sequence



The original method estimates the speed with approximately 80% error whereas our method can estimate it with around 10% error. The performance of our method is comparable to the one of state-of-the art Kalman Filter approach, although our technique does not require an initialization and now provides the gyroscope bias.

References

[1] Agostino Martinelli.
Closed-form solution of visual-inertial structure from motion.
International Journal of Computer Vision, 106(2):138–152, 2014.

[2] Agostino Martinelli and Roland Siegwart.
Vision and IMU Data Fusion: Closed-Form Determination of the Absolute Scale, Speed, and Attitude.
Handbook of Intelligent Vehicles, 28(1):1335–1354, 2012.

ARSO, 1 - 3 July 2015, Lyon, France