Stability-Based Scale Estimation for Monocular SLAM

MAVlab, Delft University of Technology, the Netherlands

Seong Hun Lee and Guido de Croon

Background

To enable autonomous robot navigation using monocular SLAM, accurate scale estimation is a critical requirement. Previous methods recover the scale by using additional metric sensors such as an IMU or a sonar altimeter. However, for small MAVs flying over an arbitrary ground structure, these sensors are often unreliable.

Objective

Solve the **scale ambiguity in monocular SLAM** without metric sensors, such as **IMU** or **sonar**.

Method

Theoretical analysis reveals a unique stability characteristics of **unscaled state feedback control** (Fig. 1).

We exploit this property to deduce the scale by **adaptively inducing and detecting vertical oscillations in hover.**

Contributions

1. We analytically show that when a proportional control system uses unscaled velocity feedback from monocular SLAM (Fig. 1), there is a **unique linear relationship** between the **absolute scale** of the SLAM system and the control gain at which instability arises, i.e., **critical gain** (Fig. 2).
2. We propose an **adaptive technique** to estimate the scale based on the hover stability of a quadotor MAV.

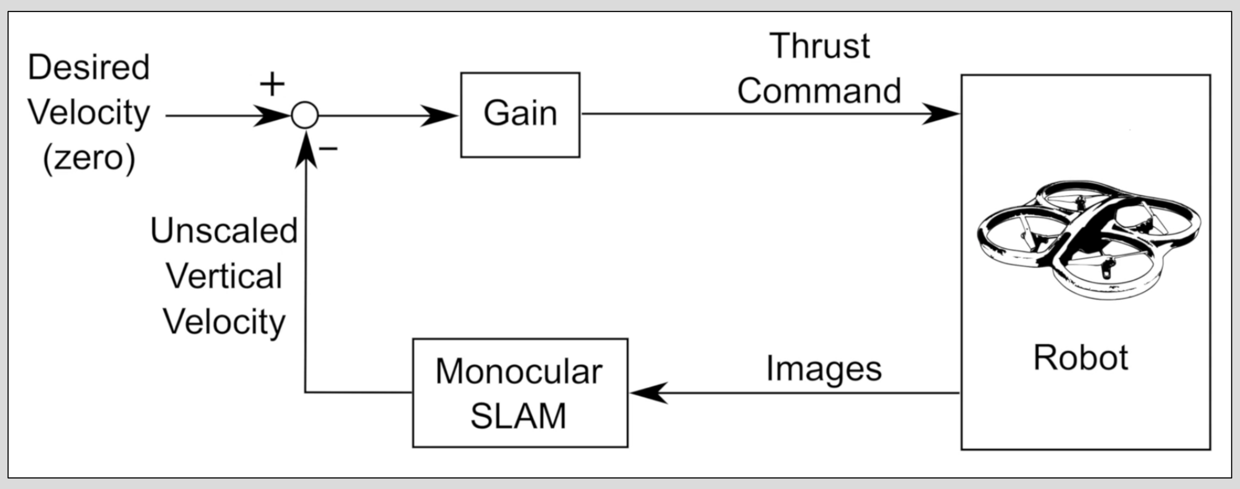
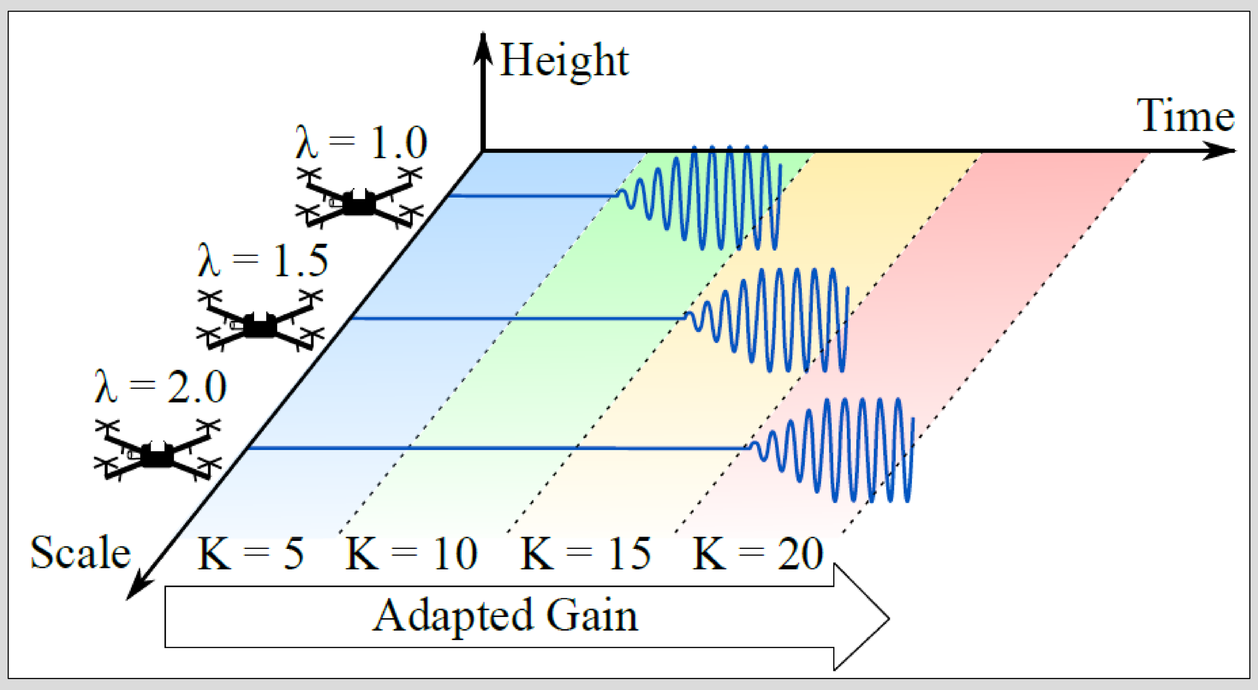


Figure 1. Unscaled Velocity Feedback Control



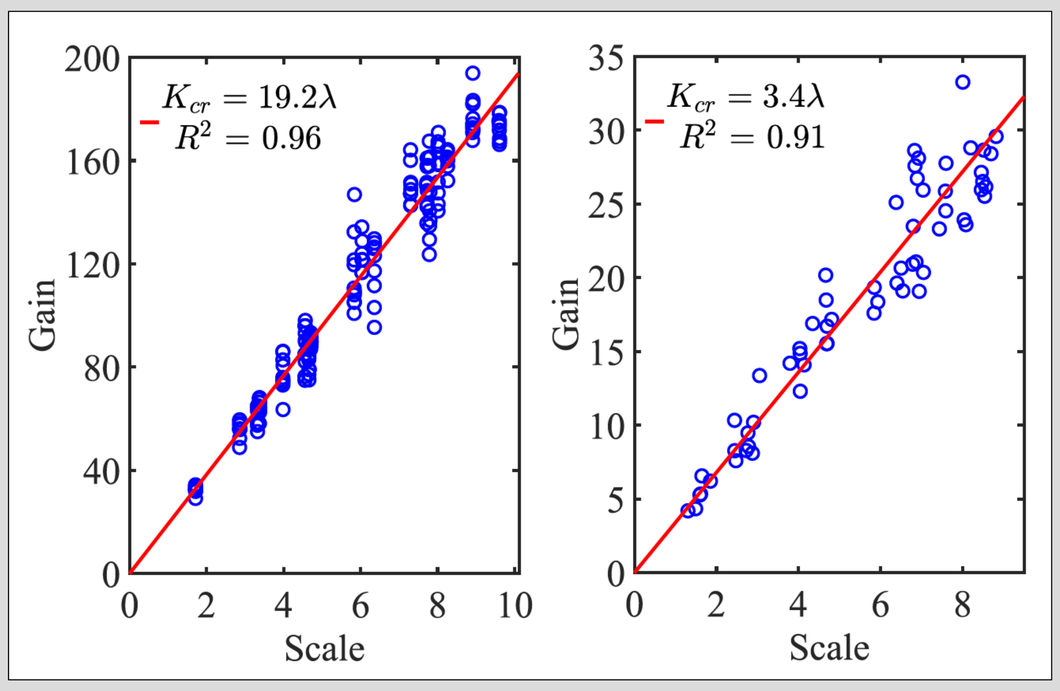
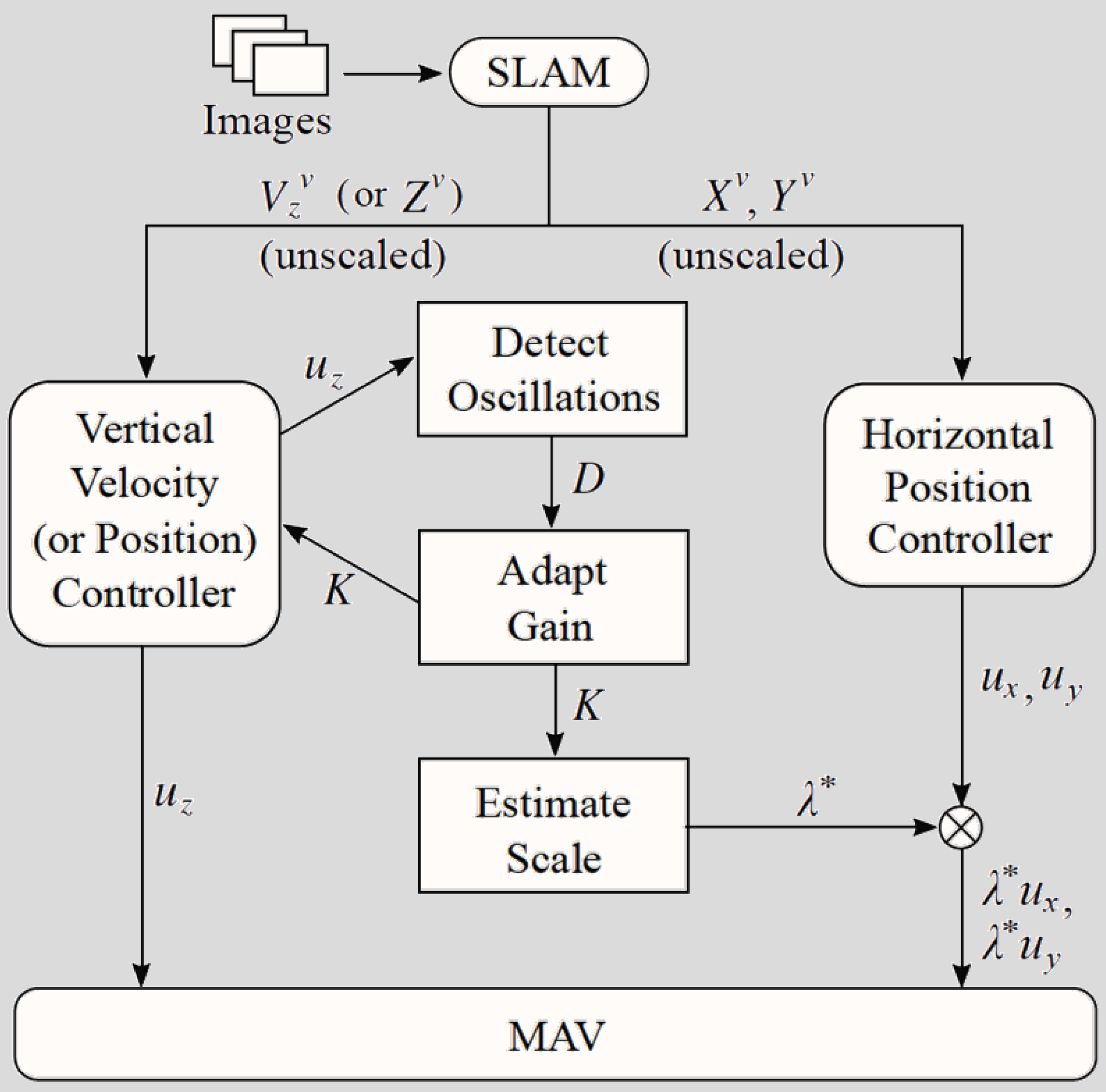


Figure 2. Linear relationship between the absolute scale and critical gain.
Left: Example illustration. Right: Verification through simulation and real-world experiments.

Adaptive Control



1. Detection of Self-induced Oscillations

We define a heuristic variable D as:

$$D(i, i_0, W) = \frac{1}{mW}\left(\max_{i' \in \{i_0, \cdots, i\}} \sum_{j=i'-W+1}^{i'} |u_z[j]|\right)$$

where m is mass, W is discrete time window size, and u_z is thrust signal. The onset of oscillations is detected when $D > D_{thr}$.

2. Adaptive Gain Control

$$K_{i+1} = K_i + PK_i\frac{(D_{thr} - D_i)}{D_i}$$

3. Iterative Scale Update for Horizontal Control

$$\lambda_i^* := \frac{K_i}{\alpha^*}$$

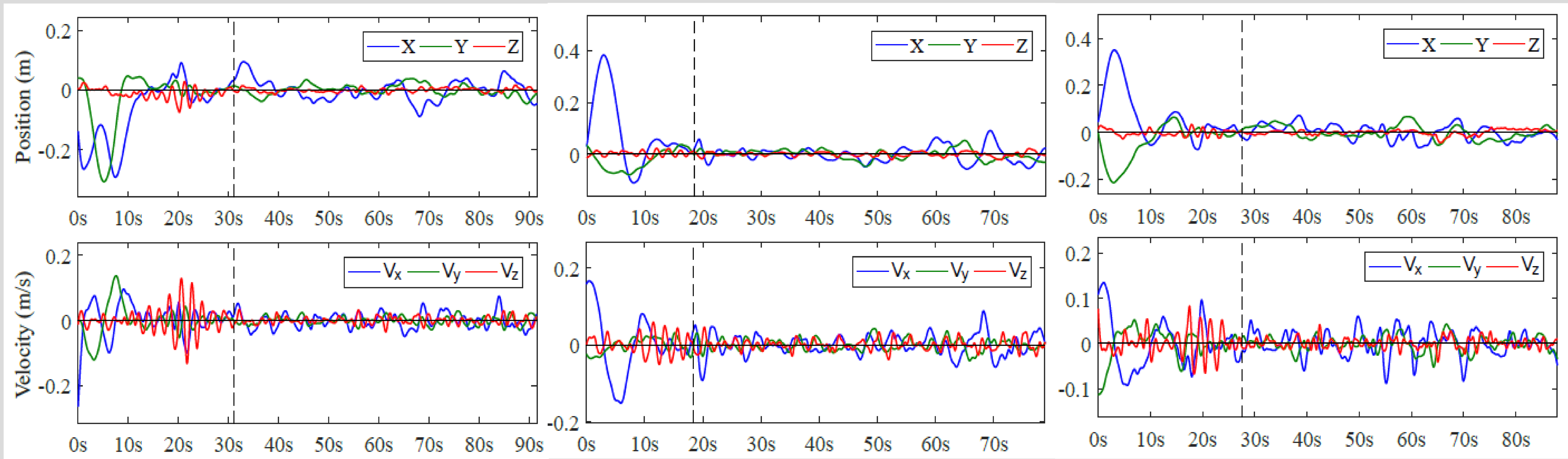
where α^* is to be empirically found in advance by performing a linear fit using the ground-truth.

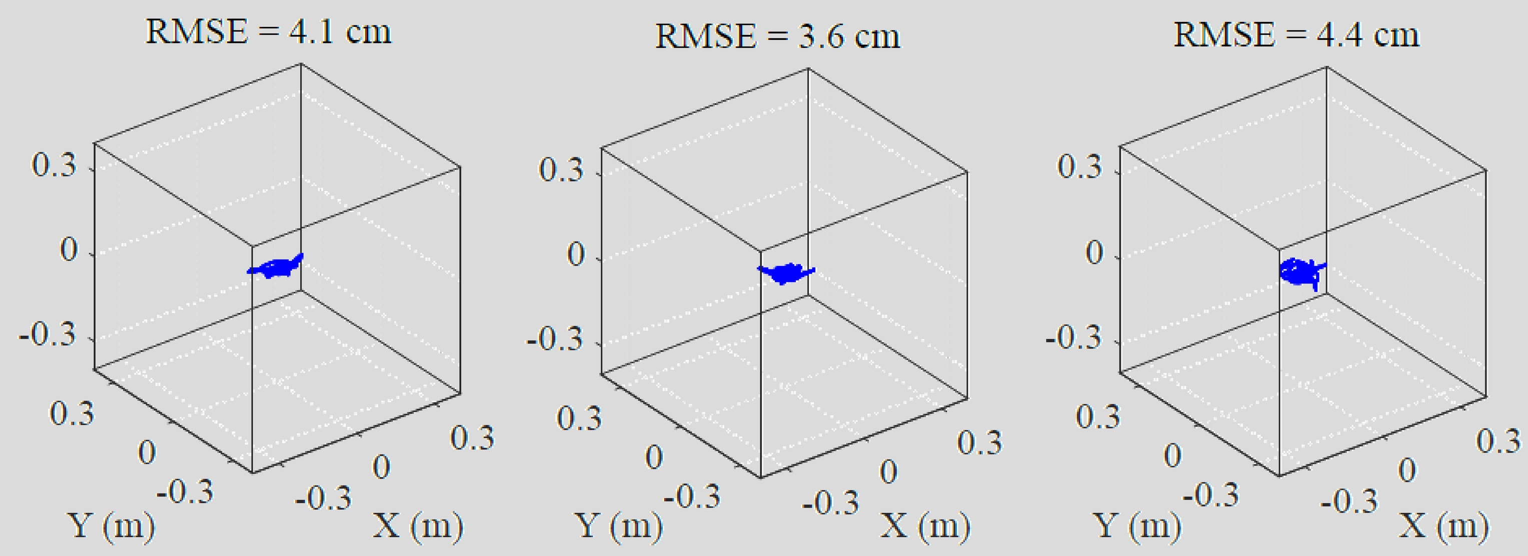
Results

1. Scale Estimation Accuracy:

Within the scene distance between 2m and 6m, the average scale estimation error is estimated to be **13.4%** and **16.0%** in simulation and in the real world, respectively.

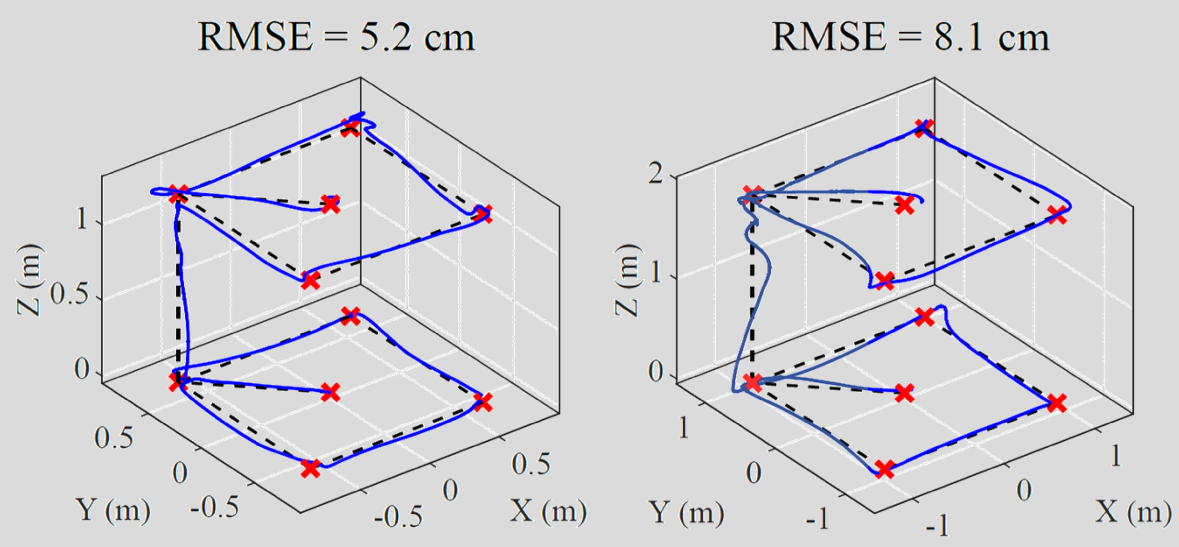
2. Hovering:





(a) Scene Distance = 2 m (b) Scene Distance = 4 m (c) Scene Distance = 6 m

3. Figure Flying



4. Waypoint Following

Target (x, y, z) (m)	$(0, 0, 2)$	$(0, 2, 0)$
Convergence time (s)	5.4 ± 0.9	2.7 ± 0.2
Peak speed (m/s)	0.86 ± 0.05	1.13 ± 0.05

5. Video Demo

We demonstrate hovering, figure-flying and 10m straight line following with Parrot AR.Drone.