

How to Properly Simulate Vision-Based Navigation in SITL Without a Real Camera?

Hi everyone,

We're currently working on a GPS-denied visual navigation system for a quadcopter and are trying to validate our pipeline using SITL before moving to hardware testing. We've made decent progress, but we're running into EKF stability issues and would really appreciate some guidance from people who have worked with ExternalNav or vision pipelines.

What We Are Building

We are developing a vision-based localization system that uses:

- Pre-stitched orthomosaic map of the operational area
- Downward-facing camera
- SIFT feature matching between live frames and orthomosaic
- Pose sent via VISION_POSITION_ESTIMATE MAVLink message
- EKF3 using ExternalNav as position source

Final Goal

1. Takeoff using GPS
2. Switch to vision-only navigation
3. Perform hover / waypoint navigation
4. Switch back to GPS and land

System Architecture

Pre-Flight: Generate orthomosaic map and extract SIFT keypoints.

In-Flight: Camera → Feature Matching → Homography → Vision Position → EKF3 → Controller

EKF Configuration

Vision Mode:

VISO_TYPE = 1

VISO_DELAY_MS = 50

EK3_SRC1_POSXY = 6

EK3_SRC1_VELXY = 6

EK3_SRC1_POSZ = 1

EK3_SRC1_YAW = 6

GPS Takeoff Mode:

EK3_SRC1_POSXY = 3

EK3_SRC1_VELXY = 3

EK3_SRC1_YAW = 1

What We Tried in SITL

Attempt 1: Constant fake vision position → EKF variance failsafe

Attempt 2: SITL motion delta added → Lane switching instability

Attempt 3: Reduced simulation noise → Slight improvement but unstable

Errors Observed

- EKF lane switch
- EKF primary changed
- EKF variance warnings
- EKF failsafe LAND trigger
- IMU stopped aiding
- SmartRTL bad position

Current Setup

- SITL: Mission Planner ArduCopter
- MAVLink: pymavlink (Python)
- Map: ~5437 x 6781 px (~5 cm/pixel)
- Localization: OpenCV SIFT matching

Desired Testing Tasks

Vision Hover and Vision Waypoint Navigation with GPS fallback and landing.

Questions for Community

1. Can vision navigation be tested in SITL without a camera feed?
2. What is correct method to simulate VISION_POSITION_ESTIMATE?
3. Recommended EKF tuning for ExternalNav only?
4. Is Gazebo or rendered camera simulation required?
5. Has anyone successfully achieved vision-only SITL hover/navigation?

Our assumption is SITL lacks camera realism causing EKF inconsistency between IMU and synthetic vision updates.

Tags: SITL, GPS-Denied, Vision Navigation, EKF3, ExternalNav, Visual Odometry