

ROBUST PERCEPTION FOR AUTONOMOUS AERIAL VEHICLES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Daniel Kim

May 2026

Unofficial template, not affiliated with or endorsed by Cornell University. The author, committee and results in this sample are fictional. Delete this line before you submit.

© 2026 Daniel Kim

ROBUST PERCEPTION FOR AUTONOMOUS AERIAL VEHICLES

Daniel Kim, Ph.D.

Cornell University 2026

Small aerial vehicles must perceive their surroundings with light, low-power sensors while moving fast, and their failures are expensive. This dissertation studies how onboard perception degrades under the conditions aerial robots actually meet, namely motion blur, low light and rapidly changing viewpoint, and how to make it degrade gracefully. Chapter 2 reviews visual odometry and learned object detection. Chapter 3 proposes a detector that predicts its own localization uncertainty and a planner that slows the vehicle when that uncertainty grows. Chapter 4 reports 214 flights in a forested test area, where the combination cut collisions from nine to one while lowering mean speed by only 12%.

BIOGRAPHICAL SKETCH

Daniel Kim received a B.S. in mechanical engineering and began doctoral study in robotics in 2020. His research concerns perception and planning for small aerial vehicles, and he has spent two summers testing drones in forests, fields and one very windy parking lot.

For my parents.

ACKNOWLEDGEMENTS

I thank my special committee chair for his guidance, and the members of my committee for their advice. The members of the robotics laboratory helped fly, fix and occasionally retrieve from trees every vehicle used in this work.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH	iii
ACKNOWLEDGEMENTS	v
1 Introduction	1
2 Background	2
3 Uncertainty-Aware Detection and Planning	3
4 Flight Experiments	4
5 Conclusion	5
A Calibration of Predicted Uncertainty	6

LIST OF FIGURES

3.1	Detection uncertainty flows into the planner's speed limit.	3
-----	---	---

LIST OF TABLES

4.1	Outcomes over 214 forest flights.	4
-----	---	---

Chapter 1

Introduction

An aerial vehicle flying through clutter has perhaps a second to detect an obstacle and react. Its perception pipeline, typically a camera feeding a visual odometry system and a learned detector, was usually developed on datasets recorded from cars or handheld cameras [1]. Aerial footage differs: the camera rotates quickly, blur is common, and light changes as the vehicle passes under canopy.

This dissertation asks what a planner should do when perception becomes unreliable, and answers that the perception system should report how unreliable it is.

Chapter 2

Background

Feature-based visual odometry tracks keypoints between frames and estimates camera motion by minimizing reprojection error [4]. Single-stage detectors predict object boxes and classes in one pass over the image [5], usually on a residual backbone [3]. Neither reports a calibrated uncertainty by default, and both can be fooled by perturbations that a human would ignore [2].

Chapter 3

Uncertainty-Aware Detection and Planning

We extend the detector to predict, for each box, a Gaussian over its corner coordinates. The predicted standard deviation σ inflates the obstacle when the planner checks for collisions, and the planner caps the commanded speed at

$$v_{\max} = \min\left(v_{\text{nom}}, \sqrt{2a_{\max}(d - k\sigma)}\right), \quad (3.1)$$

where d is the distance to the obstacle, $a_{\max}$ the braking limit and k a safety factor. Uncertain detections therefore slow the vehicle rather than being discarded.

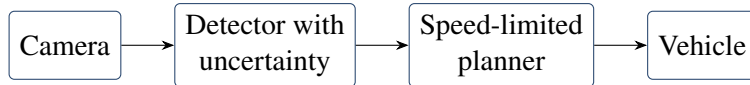


Figure 3.1: Detection uncertainty flows into the planner’s speed limit.

Chapter 4

Flight Experiments

We flew 214 missions through a forested test area at dawn, midday and dusk. Table 4.1 summarizes the outcomes.

Table 4.1: Outcomes over 214 forest flights.

Configuration	Collisions	Mean speed (m/s)	Missions completed
Standard detector, fixed speed	9	4.1	201
Uncertainty-aware, speed-limited	1	3.6	212

Most collisions with the standard detector happened at dusk, when motion blur hid thin branches. The uncertainty-aware system detected the same branches with large σ and slowed down.

Chapter 5

Conclusion

Perception for aerial robots does not need to be perfect if it knows when it is not. Feeding predicted uncertainty into the planner converts perception failures into slower flight rather than crashes, at a modest cost in speed.

Appendix A

Calibration of Predicted Uncertainty

We calibrated σ on 3,000 held-out frames by fitting a single temperature that minimized the negative log-likelihood of the true box corners. After calibration, 67.9% of true corners fell within one predicted standard deviation, close to the Gaussian value of 68.3%.

BIBLIOGRAPHY

- [1] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361, 2012.
- [2] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. In *International Conference on Learning Representations*, 2015.
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
- [4] Raúl Mur-Artal, J. M. M. Montiel, and Juan D. Tardós. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015.
- [5] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788, 2016.