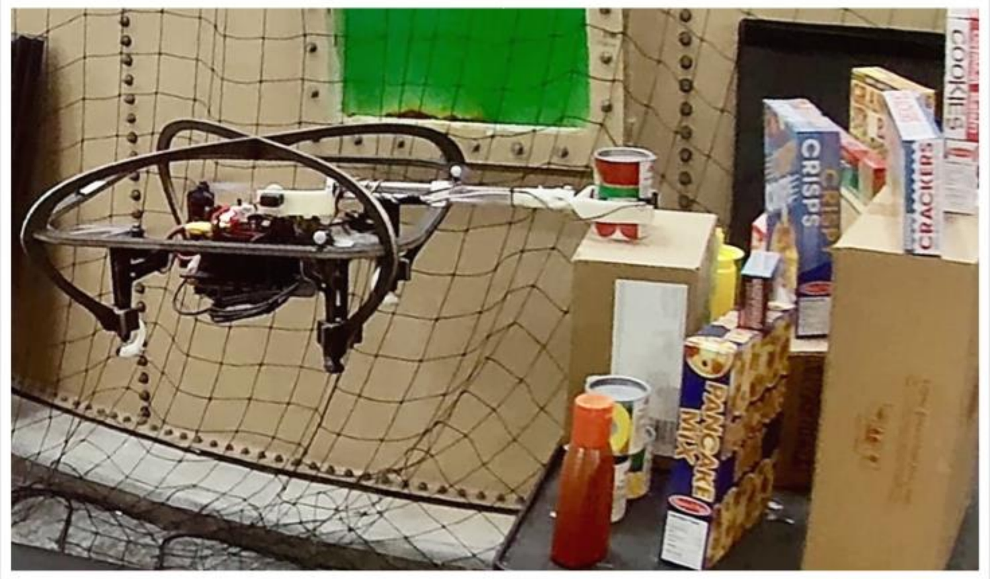
A Small Form Factor Aerial Research Vehicle for Pick-and-Place Tasks with Onboard Real-Time Object Detection and Visual Odometry

Marin Kobilarov, Johns Hopkins University
https://asco.lcsr.jhu.edu/



Abstract

This work introduces a novel, small form-factor, aerial vehicle research platform for agile object detection, classification, tracking, and interaction tasks. General-purpose hardware components were designed to augment a given aerial vehicle and enable it to perform safe and reliable grasping. These components include a custom collision tolerant cage and low-cost Gripper Extension Package, which we call GREP, for object grasping. Small vehicles enable applications in highly constrained environments, but are often limited by computational resources. This work evaluates the challenges of pick-and-place tasks, with entirely onboard computation of object pose and visual odometry based state estimation on a small platform, and demonstrates experiments with enough accuracy to reliably grasp objects. In a total of 70 trials across challenging cases such as cluttered environments, obstructed targets, and multiple instances of the same target, we demonstrated successfully grasping the target in 93% of trials. Both the hardware component designs and software framework are released as open-source, since our intention is to enable easy reproduction and application on a wide range of small vehicles.

Figure 1. Compact aerial vehicle research platform with custom components. Target objects of interest are displayed in front (toy cans and bottles) and a view of the open gripper, 9.5 cm across, around a 6.5 cm diameter can.

Contributions

Vehicle Design Contributions
Our first objective was to design modular hardware components to enable safe aerial grasping with small UAVs in indoor settings. The contributions include:

- Design of a light-weight collision tolerant cage and shock absorbing feet to enhance safety and reliability in constrained environments
- Design of a low-cost, lightweight fixed arm with an angular motion gripper, including grip detection circuit
- Open-source design files for easy reproduction

System Integration and Evaluation Contributions
Our second objective was to evaluate performance in pick-and-place experiments. Such analysis is useful since few other small form-factor aerial grasping vehicles have been reported with entirely onboard computation for image processing, 6D object pose estimation, state estimation, and motion planning. The contributions include:

- New real-time software that integrates established, computationally intensive algorithms on a small platform
- High-fidelity simulation for algorithm evaluation and development that can be seamlessly run on hardware
- Experimental results demonstrating the robustness of our proposed architecture
- Evaluation of challenges for robust pick-and-place in constrained environments of increasing levels of complexity in terms of clutter and occlusion

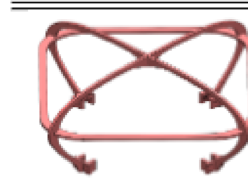
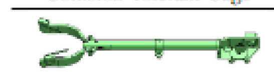
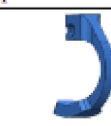
Compact Quadrotor Research Platform

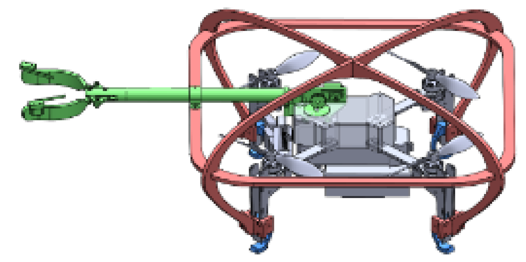
We modified the UVify IFO-SX quadrotor to be collision tolerant with a carbon fiber foam cage, shock absorbing feet, and a modular Gripper Extension Package, which we call GREP.

- Quadrotor frame size: 31cm
- Total modified vehicle mass: 1.67kg

Gripper Extension Package (GREP)
We designed GREP, a low-cost, easy to reproduce, lightweight 2-jaw angular motion gripper for picking up small objects that can be easily adapted for new vehicles.

TABLE I
CUSTOM MODULAR DESIGNS FOR ROBUSTNESS AND GRASPING

Component	Weight [g]	Dimensions (L×W×H) [mm]	Cost [$]
 Collision-Tolerant Cage	162	440 × 440 × 200	516
 Gripper Extension Package	91	440 × 120 × 45	43
 Shock Absorbing Foot	5	36 × 14 × 52	1



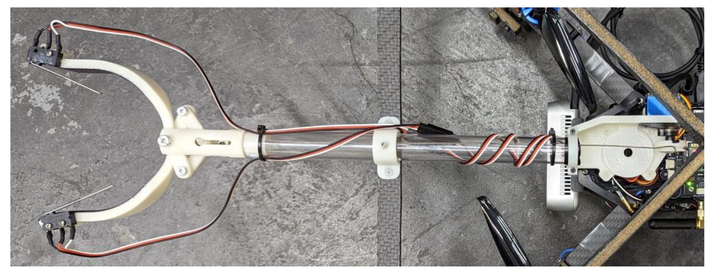


Figure 2. GREP: Fixed arm with angular motion gripper

Aerial Autonomy Software Framework

We design our software to be highly modular to integrate with different vehicles, manipulators, object detectors, and state estimators through our Aerial Autonomy software framework.

Key Components

- Aerial Autonomy fault-tolerance finite state machine
- Stereo Visual Odometry from NVIDIA's Isaac SDK for quadrotor pose estimation
- Deep Object Pose Estimation (DOPE) for object pose estimation
- High fidelity Gazebo simulation environment for rapid testing and prototyping of algorithms

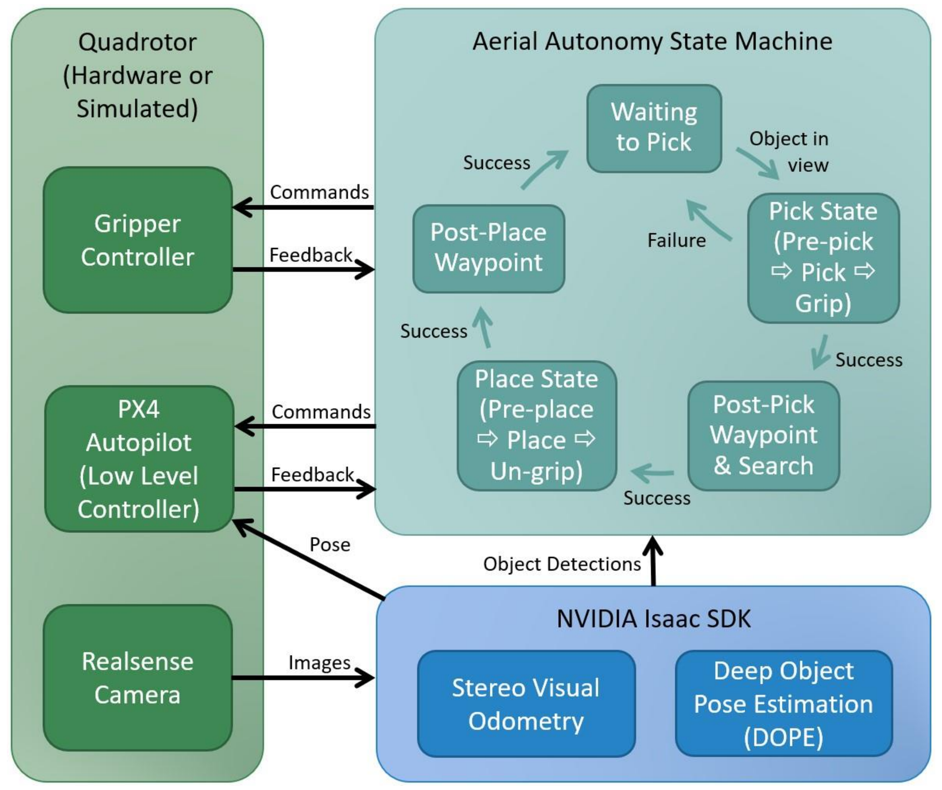


Figure 3. Software flow diagram showing a simplified pick-and-place state machine, begins at "Waiting to Pick" state.

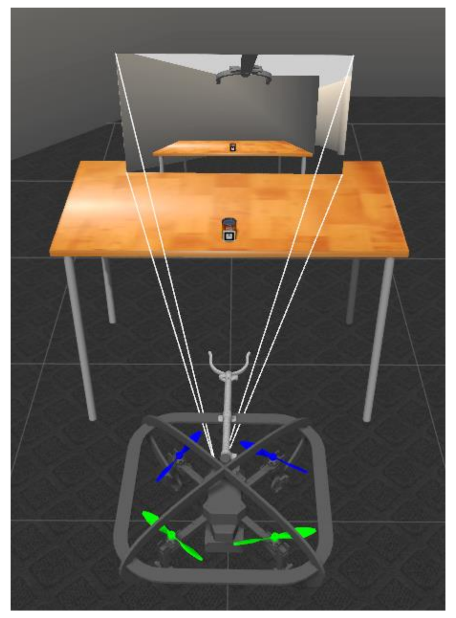


Figure 4. Quadrotor in Gazebo simulation in front of target object. The simulated camera's view is shown at the end of the camera's frustum.

Pick-and-Place Experiments

We performed hardware pick-and-place experiments to evaluate our system's performance. We placed a target object on a cart and a destination object on a table about 3 meters away. We manually controlled the system to a variable starting location facing the target object and then switched into autonomous mode, which follows the Aerial Autonomy state machine.

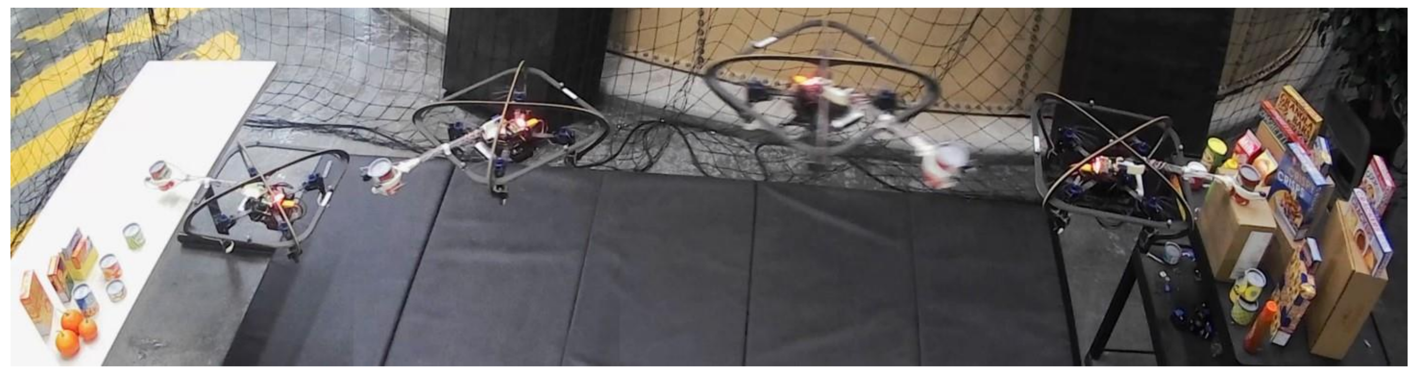


Figure 5. Aerial time lapse of a cluttered environment experiment.

We evaluated the system in 70 pick-and-place experiments in the following scenarios:

- Single object, i.e. bottle or can, in an uncluttered environment
- Cluttered environment with many items, some visually similar, surrounding the target
- An object obstructing 10-30% of the target when viewed from the vehicle's initial position
- Multiple instances of the same target object and picking and placing each

Experiment Results and Performance Metrics
With entirely onboard object detection and vehicle pose estimation, we demonstrated 93% pick accuracy and 86% place accuracy in a wide variety of challenging scenarios.

TABLE II
PICK-AND-PLACE PERFORMANCE METRICS

Scenario	Success Rates		Pick Resets
	Pick	*Place*	*Num Trials*
Single Bottle	18/20	16/18	3/20
Single Can	20/20	19/20	5/20
Can Cluttered Environment	9/10	8/9	3/10
Obstructed Can	8/10	4/8	3/10
Multiple Instance: First Can	5/5	5/5	1/5
Multiple Instance: Second Can	5/5	4/5	2/5
Overall	65/70	56/65	17/70

TABLE III
PICK ACTION OVERALL METRICS

Metric		Value
Starting Distance (m)	Min	0.78
	Mean	0.93
	Max	1.11
Pick Time (sec)	Min	13.9
	Median	18.2
	Mean	29.6
	Max	161.8

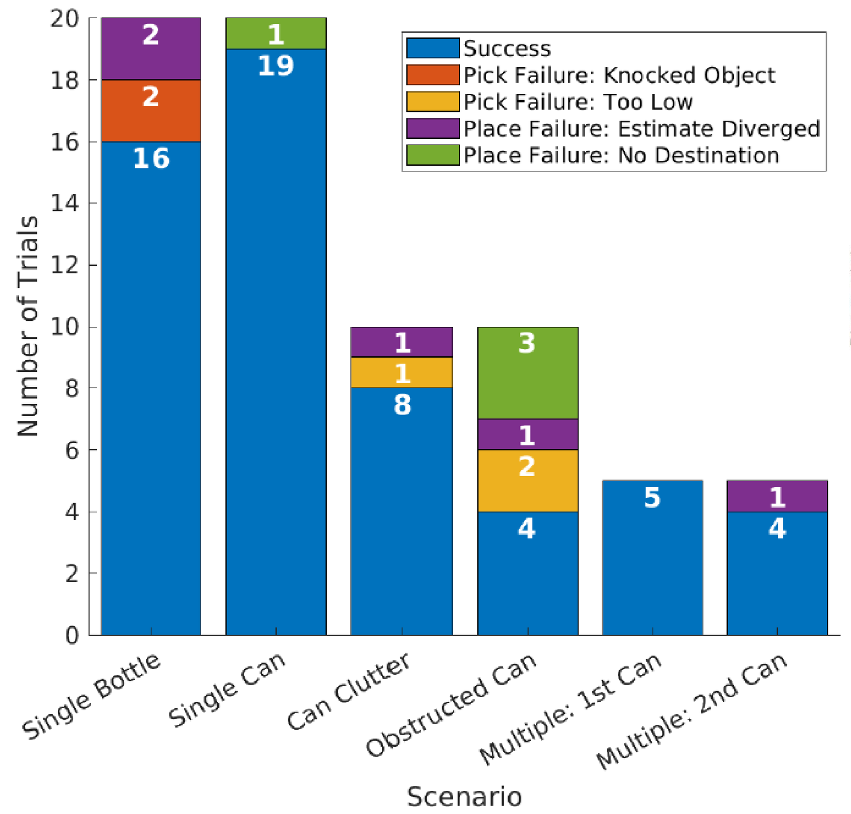


Figure 6. Failure modes across each experimental scenario.

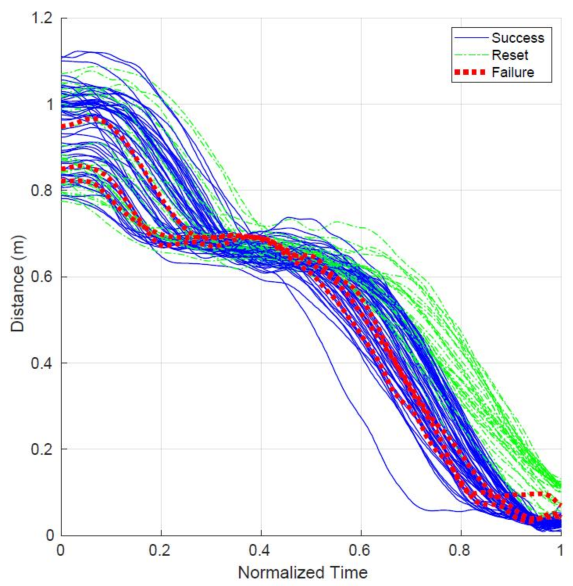


Figure 7. Distance from gripper center to target object during pick action versus normalized time across 60 trials with direct approaches (excludes obstruction trials).