THE UNIVERSITY OF WESTERN AUSTRALIA

FINAL YEAR PROJECT

Computer Vision Chase-Cam UAVs

Author:
Brett DOWNING

Supervisors:
Prof. Thomas BRAUNL
Chris CROFT

October 12, 2015

Abstract

Multirotors are here to stay, and may soon be expected to interact in a human environment. Commodity quadcopters are advertising capabilities to act as chase-cams and turn-key mapping solutions, but none of the current generation commodity uav chase-cams offer computer vision driven or even assisted flight modes to improve tracking, image framing or obstacle avoidance. Such vision assisted routines would also apply to autonomous or semi-autonomous inspection tasks for fixtures in remote or hazardous environments.

In this project, we build on the results of previous year groups and implement turn-key waypoint navigation and failsafe methods using the ardupilot software stack, and develop robust object tracking, data collection behaviours and exclusion zones on a computationally starved platform with an aim to integrate vision assisted behaviours in low-cost, lightweight UAVs.

Contents

1 Introduction

Background, Aims

1.1 Motivation

Recent Developments in power and control electronics has allowed small, consumer level UAVs to lift enough processing capacity to navigate at least partly by computer vision.

A number of Commercial Micro UAVs are beginning to advertise the ability to act as an autonomous chase-cam [1] [2]. Almost all of these systems rely on a tracking device on the user, and many navigate entirely by GPS. While the performance of GPS is improving, the performance of a chase-cam will be limited to the update rate and fix accuracy of the beacon and drone GPS modules. It also requires a lock to be achieved in both devices before tracking can commence.

The power distribution and plant monitoring sectors have recently been taking on small fleets of drones to perform routine inspection of remote, hazardous or otherwise difficult locations [3]. Many of these applications are reasonably repetitive and well defined enough to fully automate, but cannot be navigated by GPS alone due to the poor repeatability of the GPS system. During this project, we were approached by a mining safety startup interested in using UAVs for routine site inspections.

The UAVs we are targeting operate with small payloads, must react quickly to changes in the environment, and frequently feature camera systems. Despite falling costs of hobby and toy grade multirotor systems, Collision avoidance, Object tracking, mapping and inspection are not well catered for in the current market of UAVs. Part of the problem is that very few sensors are of the appropriate scale, or mass for low cost UAVs. Ground based vehicles have an advantage that a 1D sensor need only be swept in one axis to search for obstacles, but a UAV must collect data filling a volume ahead of it. About the only form of sensor capable of doing this is a 3D laser scanner which is a high-precision, high-cost, high-mass device. A simple camera collects data from an appropriately large volume at high enough speeds to navigate a multirotor, but the data is often difficult or computationally expensive to interpret.

With environmental data collected by an efficient computer-vision routine, the current generation of multi-rotor devices would be able to interact in a human environment.

Computer vision assisted control routines would permit common low cost UAVs to better frame extreme sport chase-cam film reels, automatically avoid obstacles, and stabilise GPS variances in remotely operated inspection tasks.

In this study we investigate the usability of computer vision alone to track and follow a moving target. It is hoped that this work can be used to improve the performance of camera tracking routines in autonomous chase-cam and cinematographic applications, and partially automate remote inspection

Much of the technology related to multi-copters is applicable to most other forms of UAV. The vast array of applications micro UAVs have found suggests this work may find use in Agriculture, mapping, targeted crop dusting, Cinematography, extreme-sport photography, Data collection, remote observation and inspection and hazard and environmental monitoring.

2 Literature Review

2.1 Optical Search

Search and Rescue is a field that is extremely appropriate to automation due to the difficulty in mobilising teams in remote, harsh or dangerous conditions. A micro UAV can be deployed quickly and commence the search operation before boots can be placed on the ground. The UAV Outback Challenge [4] is a regular competition to perform various search and rescue missions autonomously. The performance criteria are deliberately set very high, and the competition frequently goes uncompleted. In 2012, CanberraUAV [5], a UAV development team from Canberra completed the search aspect of the competition. After trying a number of image recognition algorithms, the search algorithm that they flew with simply looked for the blue of the jeans of Outback Joe [6]. This was sufficient to locate the target in a 4x6 km area. This goes to show that even a minimally complex routine can be effectively applied in appropriate conditions.

2.2 Terrain Estimation via Optical Flow Methods

Adding sensors to allow the copter to understand its environment is a surprisingly difficult task. Conventional contact methods operate at ranges far too close for UAVs, ultrasonics are buffeted by down-wash and most depth sensors are either too heavy or suffer under intense light. In terms of simplicity of algorithms, biomimicry has turned up some surprising results. In 2004, a French research group applied a number of optical flow algorithms to the navigation of a small helicopter inspired by insect vision [7]. The computer vision routines were used to inform the navigation loops, following the middle of urban tunnels and reducing speed in dense clutter, without necessarily computing the distance to the obstacles. These routines were relatively expensive in terms of computational power, but extremely simple and parallelizeable.

2.3 Position Estimation using Stereoscopic Methods

Any measurement will have an associated uncertainty, extracting the most information out of a collection of measurements rarely involves taking the most accurate measurements. It is possible to estimate the position of an object or feature using multiple images separated in either space or time, but making effective use the information requires an understanding of how the uncertainties behave. Error Modeling in Stereo Navigation [8] investigates a number of routines to estimate the position of a vehicle by tracking visible land-marks in a stereoscopic system, and notes the interaction between the geometry of the uncertainties, and the stability of the result.

2.4 Altitude Estimation and Odometry by optical flow

Elementary methods are still interesting for the sake of biomimicry, and indeed, a 20 element photo-array was demonstrated sufficient to control the altitude of an aircraft [9], coupling the altitude to the velocity. Dedicated, low resolution optical flow sensors have become exceedingly cheap since the market saturation of the optical mouse, and many commodity flight computers already include doppler information from GPS modules. Combining this data allows the UAV to estimate the distance to the terrain, [10]. These commodity sensors typically do not include circuitry for computing rotation, but given the cost of the sensors, that limitation can be overcome using two sensors [11]. With the availability of these parts, a number of UAV systems such as ArduPilot already include support for these devices in their code-base [12]. This support also covers pitch and roll compensation and position stabilisation using odometetry.

3 Project History

3.0.1 2013

The 2013 team, O'Connor [13] and Venables [14] established the project with the purchase of a DJI flame-wheel F550 Hexacopter with a NAZA-lite flight controller. This copter was fitted and tested, and finally converted into an autonomous platform with the addition of a Raspberry Pi single-board computer (RPi) and a circuit to switch the control channels from the radio receiver to the RPi GPIO outputs. The NAZA-lite at this time did not feature telemetry outputs or way-point inputs, but it was able to loiter in a stiff wind using a GPS fix. This team gathered location information for the RPi using a Qstarz GPS unit, and bearing information using an X-sens IMU. The sensor duplication did not exceed the maximum payload capacity of the platform, but it did suffer a short flight-time. Under direct control from the RPi, the drone was able to execute basic way-point navigation.

3.0.2 2014

The 2014 team, Baxter [15], Mazur [16] and Targhagh [17] continued the project adding an internet accessible web UI to control the various autonomous features of the platform, displaying the flight-path of the copter and a live feed of the camera. The navigation routines were improved and extended to permit operation without reliable heading information, and the computer vision routines were extented.

3.0.3 2015

The project is being continued by Allen [18], Mohanty [19], Tan [20] and Downing [21]. Because of the rapid pace that the market is adopting new features and technologies, we saw fit to undertake a critical review of all aspects of the platform, and begin improvements that facilitate the current typical feature set.

3.1 Platform

3.1.1 Description

This project centred around a hexacopter based on the DJI F550 frame. This frame allows for a generous lift capacity, and plenty of space to fasten flight assist hardware, and represents a low cost platform so that our work can be reproduced by a sufficiently motivated hobbyist.

In this work, the flight tasks are based around two physically separate processors; the Flight Controller, and the Server. This allows a stable, known safe build to be maintained on the flight controller, while experimental code runs on the server. If the server fails for any reason, the flight controller will maintain flight, and it allows the operator to engage various fail-safe behaviours without having to rely on the experimental code. At hand-over, the server was a Raspberry Pi model B+, and the fight controller was a NAZA lite. This combination did not go together smoothly. The NAZA lite was designed as an entry level free-flight controller and did not expose any interfaces for telemetry or data acquisition. As such, the previous teams had to duplicate the GPS and IMU sensors to have that data available to the server's algorithms.

The NAZA also did not feature waypoint navigation features, or secondary command channels so the previous teams used a relay board to physically switch the control inputs from the RC receiver to the server, and implemented their own flight control algorithms on the server. We were told that server lock-ups frequently caused the craft to power-dive.

The server has been upgraded to a Raspberry Pi V2 single board computer, which we consider to be a computationally starved platform. At the time of writing, the Raspberry Pi V2 has less computational power than a smartphone in the $200 price bracket.

Flight controller processor, gimbal

3.1.2 Platform Weaknesses

(fitness for purpose etc) We made some attempts to break into the data channel between the NAZA flight controller and its GPS module, but the raw GPS and compass data was not particularly useful without the accelerometer and gyro data. This channel had also been deliberately obfuscated by DJI, making it quite clear that this was beyond the design intent of the flight controller. A community RC group had reverse engineered this link, and we ported their code to the raspberry pi and removed the Qstar-z GPS unit.

The physical switch of control from the user to the server, and the ongoing development of a control loop meant that RC switch was never configured to surrender the throttle channel, and the server was left without altitude control. This meant that the server had continuous perturbations from the user who remained responsible for maintaining a fixed altitude by sight alone.

The NAZA is a robust, if basic free-flight controller, and did not feature telemetry. The server did not have access to an estimate of remaining flight time, and the user could not retrieve logs in the case of a crash.

Waypoints were handled on the server despite the NAZA featuring GPS assisted loiter behaviour, and a reliable return-to-launch failsafe feature. In general, control loops should be as tight as possible and routing them through an entirely separate processor is clearly an unnecessarily long loop.

3.2 Team Achievements

After spending some time with the hardware configuration of the previous teams, we prepared a report outlining the weaknesses and proposing a change. We swapped the NAZA flight controller out for a 3DR Pixhawk running the Ardupilot 3.3 firmware. This gave us Waypoints, altitude control, failsafes, telemetry, and a GPS-only driven follow-me mode out of the box. The open and configurable flight controller with built-in autonomous and hybrid modes, meant that we could fully surrender control to the autonomous platform, including altitude control. The built-in mode change behaviours meant that we could relinquish and resume manual control at the remote, or activate fail-safes just as reliably as with the NAZA. The telemetry link supports the MAVlink protocol, which defines various commands from interrogating the status to setting waypoints, or even setting the instantaneous velocity vector. This let us keep the high-speed components of the control loops firmly inside the flight controller, and direct it with the lower speed vision tasks. This also gave us linear control inputs in SI units. The new configuration allowed us to fully eliminate duplicated hardware, an opportunity to fully re-build the wiring harness, and significantly reduce the weight of the craft, extending the flight time to almost twenty minutes.

We saw fit to refresh the Web UI and HUD, implementing a more appealing and maintainable template system, We implemented a Waypoint modifier routine using polygonal no-fly zones, Glyph detection Capture and tagging of images to suit proprietary offline 3D reconstruction algorithms. Computer vision driven chase-cam behaviour

By using a fully-featured open-source flight controller, we also had access to the community developed simulation environment. The Ardupilot simulator is designed to simulate the behaviour of the flight controller against various hardware configurations and flight styles. As such, it incorporates a comprehensive inertial model and noise injection to the various sensors. We were able to compile our tests for our development machines couple them to the simulator over a local TCP socket, and test them against the same version of the flight control firmware as we had uploaded to the physical copter.

4 Object tracker

4.1 Position estimation

After experimenting with the rudimentary PID loop, we moved on to basic triangulation methods to estimate the location of the object with the assumption that it was on the ground. We were estimating the height based on the GPS altitude data streaming from the NAZA GPS, but without access to the NAZA's internal barometer, the uncertainty in height caused the navigation loop stability to vary significantly minute by minute. After the upgrade to the Pixhawk, we had access to the data from the Extended Kalman Filter in the flight controller. We had also made significant changes to the way the flight controller was guided by the server, and this prompted a near total re-write of the object tracking code-base.

4.2 Desired relative pose

Once the object is located in 3D space, the copter will calculate a vantage point from which to view it. This vantage point could include information regarding the terrain, the size and trajectory of the object, the field of view of the camera, locations of light sources etc. This is the point at which cinematographic style is included in the algorithm.

4.3 Motion Control

Moving the copter causes large changes to the pitch and roll of the copter

4.4 Observations and uncertainties

4.4.1 Structure from image

Resolving structure from image is a field unto itself. In general, the computational load of modelling the environment in three dimensions live is beyond supercomputing clusters that would fit in the mass restrictions of most multi-rotors. However, Simultaneous Localisation and Mapping (SLAM) systems have optimised components of this field to a vast array of algorithms of varying computational power requirements. One of the critical features of a modern SLAM systems is an acknowledgement that every measurement comes with some uncertainty. In many cases, the treatment of uncertainty can be very simple, carrying a one-dimensional confidence value, to the very complex, where every aspect of the measurement is recorded for post-processing where non-linear couplings can be accounted for.

Some systems are beginning to feature graph-traversal techniques where observations are stored as relative links in a graph, and networks are drawn through those links to optimise against the uncertainty of various aspects of the environment. These techniques allow the landmarks to be unloaded from memory very easily permitting a small live map, and can approach the detail of full-map covariance methods with a second, less-than-live graph traversal process.

4.5 Uncertainty Analysis

A good uncertainty model encodes everything that is known and nothing of what is unknown. Good treatment of uncertainties combines knowledge without losing information, or adding assumptions. In the case of structure from image, a system that extracts maximal information from a single camera should be automatically capable of full stereoscopy using only the uncertainty analyses that applied to the monocular case.

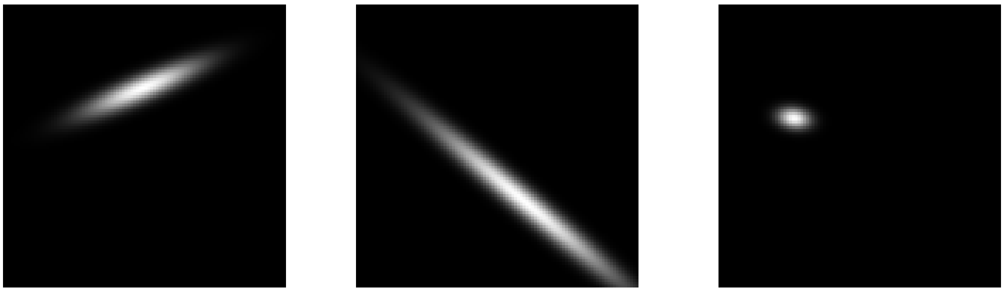
4.5.1 Ellipsoidal models as Covariance Matrices

For the sake of computational load, we've assumed the collected data has a three dimensional Gaussian distribution which we represent using covariance matrices. By permitting singular matrices, this gives us the flexibility to represent uncertainty distributions unbounded in length in any direction, and still combine them to clearly defined local distributions. These covariance matrix models require normalisation to behave as probability density functions, but since the normalisation constant is fully defined by the coefficients, it can be completely ignored until the distribution is to be rasterized and rendered.

4.5.2 Transformations

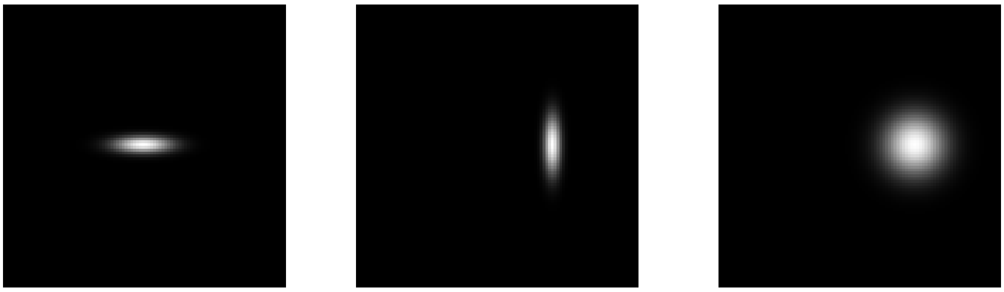
These covariance structures are well suited for use as vectors, automatically calculating the uncertainty of any combination of vectors.

Mean: Multiple observations of a single point are generally subject to averaging, but assigning weights to superior measurements is often difficult as is a common point of discussion in Kalman Filter design. If the subject of the measurements is known to be the same object, the object is likely to be at the centroid of the product of the probability density functions. This takes measurement accuracy into account. Using the form above, the product of density functions becomes the sum of polynomials. For brevity, the matrix form is shown below. This requires the result to be non-singular, because a singular matrix does not define a fixed centroid. However, a suitable centroid can be chosen. Vector sum: Treating the ellipsoids as vectors requires the definition of a vector sum. Again, the matrix form below does not permit singular Matrices as inputs nor outputs, but unpacking it into polynomial form allows suitable centroids to be chosen. In the same way as above, the vector difference can be computed. This allows the definition of a discreet time derivative. A positional derivative defined in this



(a) Measurement 1 (b) Measurement 2 (c) Combined Measurements

Figure 1: The Product of Two Uncertainty Distributions



(a) Object at the Origin (b) Uncertain Displacement (c) Final Object Location

Figure 2: The Vector Sum of Two Uncertainty Distributions

way has an appropriately large velocity uncertainty distribution that increases in size with shorter time-steps, but can be averaged with multiple samples.

Every sensor on our platform has some resolving power which favours particular directions. Our camera cannot locate a point in 3D space, but the direction is known within some small angular uncertainty. The Lidar Lite time of flight optical range sensor has a 3 degree conical beam, a 10mm starting aperture, 10mm resolution and a small amount of sample noise. It can locate an object in 3D space, but the uncertainty model favours depth to breadth. The GPS gets interesting, In general we can say it has at least 3m sigma = 1 radius. However, over short time periods it has a resolution much better than that.

4.5.3 The Hyperbolic case

With a negative determinant, the uncertainty model becomes hyperbolic, expanding to an asymptote with an elliptical conic cross section. This is interesting because it represents an angular uncertainty, with minimal changes to the uncertainty model. In order to represent the hyperbolic distribution, one of the diagonal terms of the covariance matrix must be negative. Here, positive numbers represent increasing certainty, zero represents total uncertainty. so mixing a hyperbolic distribution with an elliptical one would suggest that the result contains less information than the elliptical distribution alone. This stems from the hyperbola having a normalisation function over its area. Given that it integrates to infinity, the probability density function would be infinitesimal everywhere. A more useful description of its behaviour would be that its cross-section integrates to one. This model is similar to the definition of a Gaussian laser beam profile; the total energy in the beam is proportional to its length and is

(a) Negative Determinant (b) Normalised

Figure 3: A Hyperbolic Distribution Showing Laser-like normalisation

therefore unbounded, but the power density is fixed for any cross section through the beam. With a single image from a camera, the only valid fact about the location of an object is that it lies somewhere within an infinite cone.

4.5.4 Extensions

Using a vector difference of positions to estimate the velocity, models of objects can be evolved over time and updated with additional measurements. With additional statistical analysis, this can be extended to a form consistent with Kalman style filter for the motion of every visible landmark.

Storing a motion history of the objects would facilitate the generation of a time-dependent network of estimated locations and relative measurements. Here, the nodes represent the instantaneous locations of the objects, linked forward in time by velocity estimates, and linked in space by relative distance measurements. Assumptions can be added to this mix in the same form as the object locations, time evolution and relative measurements. A live process need only deal with the instantaneous positions of the robot and landmarks, but the models can be further refined less-than-live by a graph traversal. A SLAM process can be initialised by an assumption in the form of an artificial survey marker at boot. The assumptions can be later removed and replaced by real survey markers, with the map fully regenerated by the graph traversal process which need not be on-board the robot.

4.6 Current implementation

4.6.1 Strengths

Encodes appropriate information in a covariance matrix form Cleanly integrates multiple observations into an object model Makes assumptions explicitly

4.6.2 Weaknesses

Camera model does not account for uncertainty from angle of pixels, gimbal pose, IMU sample time (angular velocity), GPS drift Requires accurate calibration data for the image sensor

GPS drift is deliberately assumed to be negligible with short-term measurements. Cannot describe complex distributions (unbounded polynomial order) arbitrary geometry Uncertainty Images from SoggyDog

4.7 Future Work

The Object tracker Velocity interpolation not implemented. Graph traversal (SLAM) not implemented. Describe hyperbolic distribution and flaws Describe Laser beam necking case

5 Test Results

5.1 Simulation tests

We used the simulator built alongside the arducopter firmware to run our tests. This simulator is detailed enough to handle the inertial model of the platform and inject noise into the sensors. This simulator allows us to run our full software stack on any computer with a functioning compiler. The object tracking code estimates the location of the objects using data from the flight controller covering the Gimbal pose, the roll, pitch and yaw of the copter, and the GPS location as fused with the accelerometer and gyro data. Using a servo driven gimbal on the real copter we cannot perform mechanical compensation of the pitch and roll without injecting extremely large angular uncertainties. The object estimation code compensates for the pitch and roll using matrices representing body rotations under OpenCV. Because the the camera cannot resolve range, an assumption model was added to the code reducing the columnar vision model to a spot on the ground.

This compensation is designed to cleanly counter the non-linear second order signal injection, but with the camera mechanically coupled to a laptop instead of the simulated copter, the loop becomes completely unstable.

Commenting out the pitch and roll decoupling code makes the copter very well-behaved in the simulator, using a pose generator that attempts to place the object at a ray cast forward and down from the copter, the copter can be confidently led around the simulation environment. If the object appears behind the copter, the copter moves away from the object, turning around as it does so.

Camera uncertainty models combined with assumptions Pretty pictures Multiple observations with time-evolution

5.2 Live tests

Able to physically follow one object, while tracking multiple others. Acceleration limits, velocity limits, time-cutoffs. pitch and roll stability Basic colour matching limitations Not as robust as radio links, but

6 Future Work

Expand uncertainty analysis SLAM, end to end solution SIFT algorithm Optical Flow Recommendations: ROS, ROS, ROS.

7 Conclusions

We have greatly improved the capabilities of the UWA autonomous hexacopter platform, and have brought the code-base up to a level where it is feasible to implement a SLAM process.

References

[1] "Lily - The Camera That Follows You." https://www.lily.camera/, 2015.

[2] "Auto-Follow Drone for GoPro Camera." https://www.airdog.com/, 2015.

[3] "UAV Inspection." http://www.ropeaccess-wa.com/unmanned-aerial-vehicle-inspection/, 2015.

[4] "Uav ouback challenge." http://www.uavoutbackchallenge.com.au/.

[5] C. Anderson, "Open source civilian uav development." http://canberrauav.org.au/, 2012.

[6] Andrew Tridgel, "2012 UAV Outback Challenge." Personal discussion at Make-Hack-Void following Linux Conf Canberra, 2013.

[7] L. Murateta, S. Doncieuxa, Y. Briereb, and J.-A. Meyera, "A contribution to vision-based autonomous helicopter flight in urban environments," *Robotics and Autonomous Systems*, vol. 50, p. 195209, March 2004.

[8] L. Matthies and S. Shafer, "Error modeling in stereo navigation," *IEEE Journal Of Robotics And Automation*, vol. 3, pp. 239–248, June 1987.

[9] T. Netter and N. Franceschini, "A robotic aircraft that follows terrain using a neuromorphic eye," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2002.

[10] S. Griffiths, "Remote terrain navigation for unmanned air vehicles," Master's thesis, Brigham Young University, 2006.

[11] J. Kim and G. Brambley, "Dual optic-flow integrated navigation for small-scale flying robots."

[12] "Documentation for the installation of optical flow sensors on the ardupilot." http://copter.ardupilot.com/wiki/common-optional-hardware/common-optical-flow-sensors-landingpage/common-mouse-based-optical-flow-sensor-adns3080/.

[13] R. OConnor, "Developing a multicopter uav platform to carry out research into autonomous behaviours, using on-board image processing techniques," 2013.

[14] C. Venables, "Multirotor unmanned aerial vehicle autonomous operation in an industrial environment using on-board image processing," 2013.

[15] M. Baxter, "Autonomous hexacopter software design," 2014.

[16] A. Mazur, "Autonomous operation and control of a multirotor unmanned aerial vehicle through 4g lte using onboard gps and image processing," 2014.

[17] O. Targhagh, "Autonomous navigation of unmanned aerial vehicles," 2014.

[18] Richard Allen. https://github.com/20739367, 2015.

[19] Manish Mohanty. https://github.com/manishmohanty09, 2015.

[20] Jeremy Tan. https://github.com/jtanx/, 2015.

[21] Brett Downing. https://github.com/brettrd/, 2015.