Just Keep Flying: Machine Learning for UAV Thermal Soaring
CS 229 Final Project

Geoffrey Bower and Alexander Naiman

Introduction

Unmanned Aerial Vehicles (UAVs) have recently become popular for aircraft missions that would be dangerous or exceedingly boring for human pilots. One potential application of UAVs is for extremely long endurance missions, such as for military surveillance or for commercial use as atmospheric satellites. For this reason, extending the endurance of UAV flight is currently an area of major research interest.

One way to extend the duration of a UAV's mission is to take advantage of energy available in the atmosphere in the form of wind velocity gradients, or wind gusts. By adjusting an aircraft's attitude, speed, and direction of travel when it encounters a gust, a properly trained pilot can greatly extend the duration of a flight, even indefinitely if enough wind energy is available. Glider pilots, radio-control aircraft pilots, and birds all use these gust soaring techniques to reduce the amount of propulsive energy required to stay in the air.

Designing an automatic control system to take advantage of atmospheric energy for extending mission duration is a very difficult problem. The flight environment is not easily predictable, and gust detection equipment such as Doppler radar and LIDAR is prohibitively expensive in both dollars and power usage. These issues make this a perfect problem for utilizing a machine learning algorithm with simpler sensors inputs.

This project considered vertical thermal updrafts and how to best control the velocity and bank angle of the aircraft to extract the most energy from them. The goal of this project was to use reinforcement learning to determine control laws that lead to the longest duration flights on average. The control law inputs were based on sensors that can be placed on a small UAV. The aerodynamic properties and dynamics of a small UAV were modeled and used in a simulation with the control laws. A thermal model developed at NASA Dryden was used to model thermal generation based on the time of year, time of day and local atmospheric conditions.

Current research undertaken using the thermal model in maximizing duration of UAV's at NASA has used heuristics to search for thermals and then makes a number of assumptions about tracking and centering in the updrafts (30 sec to center, then fly at optimal radius). Through reinforcement learning, improved performance should be obtainable without relying on heuristics to search for and self center in thermals.

Thermal Model

For this project, choose to use a thermal updraft model developed by Allen at NASA Dryden (Ref. [2]) for a similar autonomous UAV soaring project. This model was developed using atmospheric data collected by the NOAA in Nevada using rawinsonde balloons released every 12 hours over the course of a year.

Thermal updrafts form when an air mass close to the ground is warmed and becomes unstable. This air then begins to rise and cool until it reaches equilibrium with the surrounding atmosphere. Conservation of mass dictates that this rising air be replaced by generally sinking air outside the thermal updraft. Allen's model provides the variation of the velocity and height parameters that characterize thermals, and also provides for the proper spacing and duration of thermal updrafts.

For this project, we used Allen's model to create a dynamic field of thermal updrafts of various characteristic heights, velocities, and durations inside a specified test area for one hour, which the UAV was constrained to fly inside. Figure 1 shows an example of such a field at a particular time during the simulation, where the wireframes indicate the extent of positive updraft velocities. Allen's model of the ground assumes a uniform radiative heat flux over the test area. In the real world, thermals are more likely to be formed over certain ground features that radiate a lot of heat to the air, for example, paved parking lots. A possible modification of this model would increase the probability of thermals forming over certain parts of the test area and decrease the probability over other parts. This would likely have a significant impact on the machine learning aspects of this problem, as the UAV would hopefully learn to look for thermal updrafts in more likely areas. This modification is beyond the current scope of this project, however.

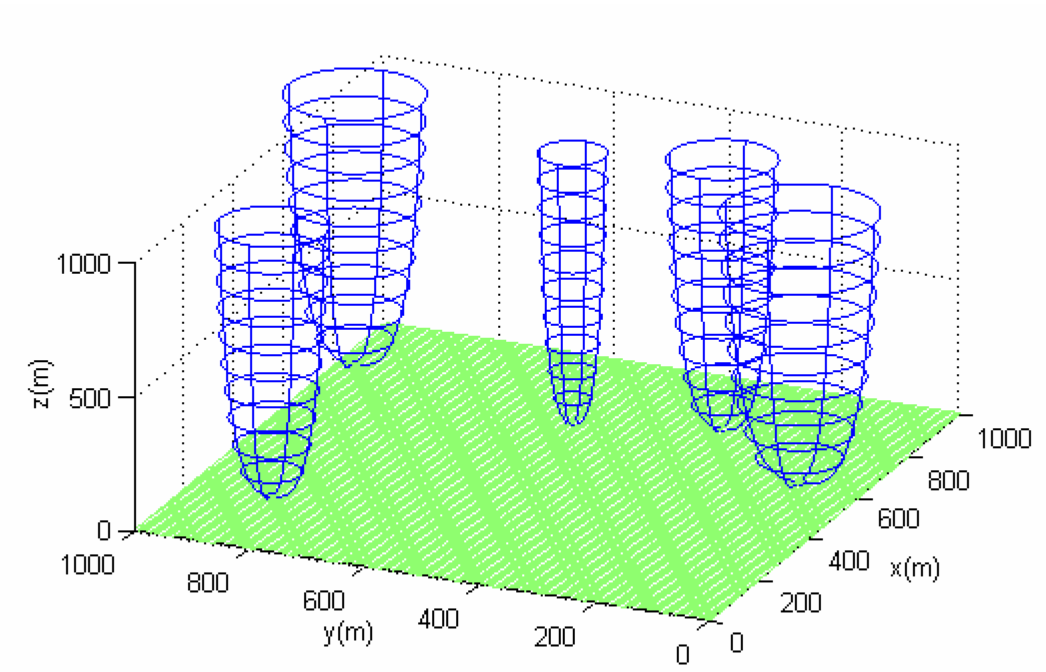

Figure 1. Sample Thermal Updraft Field

Dynamics Model

The airplane we modeled was based on a small unmanned glider built by a Stanford student, Chinmay Patel, for his research on autonomous soaring through high frequency gusts. This airplane was chosen because it is currently outfitted with an autopilot system and sensors, so as time permits the learned control algorithm could be tested on real hardware. The dynamic modeling of this airplane was simplified so that rapid simulation could be performed. The glider was modeled as a point mass with a simple parabolic drag polar. Based on the lift coefficient, bank angle, and updraft velocity, the flight path of the glider can be determined as well as the inputs to the control law at the next time step. The control outputs were the desired lift coefficient and the bank angle of the airplane.

Sensors present on the airplane that were used as inputs to the control laws are the GPS position (x,y,z), GPS velocity (u,v,w), airspeed (V) and bank angle (φ) all sampled at 4 Hz. Some constraints were placed on the changes in the control output that are related to the dynamics of the actual airplane. These limits relate to the maximum pitch rate and roll rate of the actual airplane. The important properties of the airplane are tabulated in Table 1.

This simple model captures the dynamics of the real airplane fairly accurately, but the assumptions create some short term dynamic differences from the real airplane. Unsteady and non-linear aerodynamics along with control transients limit the accuracy of the model in the regions of maximum control input, for this reason conservative constraints were placed on the pitch rate and roll rate.

Table 1 Airplane Properties

Mass	0.477 kg	C_{Lmax}	1.2
Wing Area	0.331 m^2	φ_{max}	30 deg
Effective Aspect Ratio (e_o*AR)	8.8	$(dC_L/dt)_{max}$	0.2 /sec
C_{D0}	0.025	$(d\varphi/dt)_{max}$	30 deg/sec

Control Strategy

Due to run time constraints for this project, we chose a simple objective that would allow rapid simulation. The selected objective was to maximize the altitude of the UAV at the end of an hour of simulation. This rewards control strategies that find the most thermals and center as quickly as possible in them to gain as much altitude as possible.

The control algorithm used in this project was based on a simple online learning strategy: the UAV attempts to find the strongest thermal updraft that it can and stays in it as long as possible, thereby gaining as much altitude as possible. To implement this strategy, we used a value function-based algorithm to estimate the vertical velocity at each point within the flying area. A separate control policy was then used to command the UAV to fly to the point within the flying area with the greatest expected vertical velocity.

More specifically, the online learning algorithm discretizes the flying area into 10-meter square cells. Initially, the expected vertical velocity in each cell is randomly set to a negative or very small positive value (between -1.0 and 0.01 m/s). The UAV can sense its own vertical velocity with GPS at each time step. As it flies, it updates its estimate of the vertical velocity of each cell that it enters to be the current measured velocity in that cell. If it reaches its target cell and finds a negative velocity there, it re-randomizes the estimates of the vertical velocities in all of the cells that it has not visited yet. Figure 2 shows an example of a value function field. Note that the path of the UAV is clearly visible among the noise of the randomized, un-visited cells, and that the location of thermals as plateaus of vertical velocity are also clearly visible.

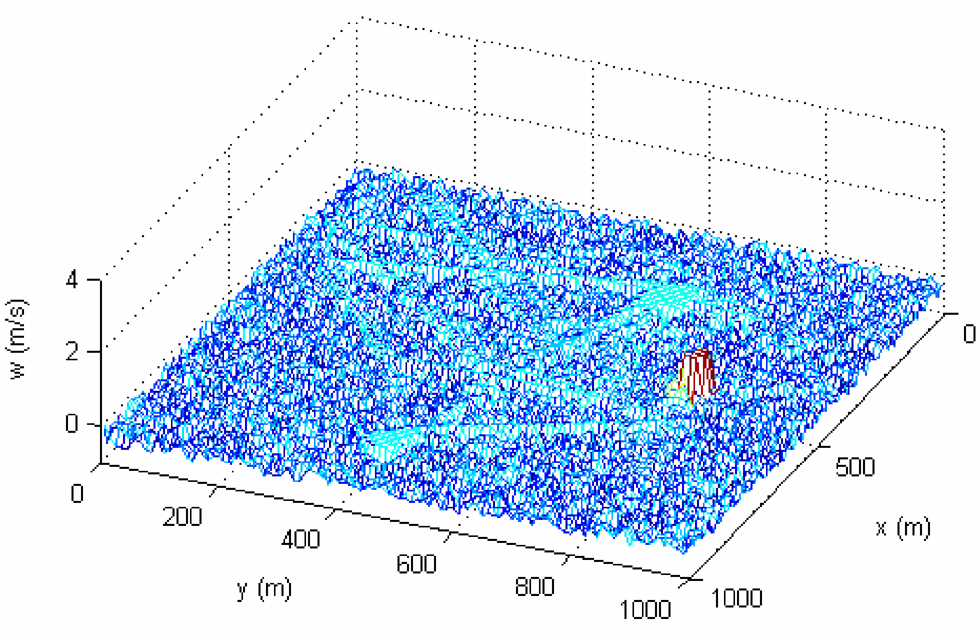

Figure 2. Sample Value Function Field

Because the updrafts have a finite lifetime, the value function estimate of a particular visited cell becomes less and less accurate depending on how recently the UAV has visited it. To reflect this increasing uncertainty with time, a discount factor of 0.995 was applied at each time step such that the estimated velocity of each cell gradually decays toward zero.

In our dynamics model, we have two control inputs to the flight path of the UAV: bank angle to control turning and lift coefficient to control speed and ascent/descent rate. The bank angle control policy used simple open-loop controller to control the heading, commanding the UAV to fly toward the desired point. The lift coefficient control policy, on the other hand, can greatly affect the performance of the overall algorithm. The optimal control policy will generally decrease the lift coefficient when outside a thermal updraft, increasing the flight speed and covering more area, which increases the chance of finding an updraft.

To find this optimal control policy, we implemented a policy search algorithm to maximize the altitude of the UAV after one hour. The policy family consisted of a linear combination of vertical velocity, bank angle, and a constant: $c_L = A_w * w + A_\phi * \phi + B$. We used the stochastic gradient ascent algorithm to optimize these three coefficients. Fifty simulations were performed for each policy at each step of the optimization, and the reward for a particular policy was taken to be the average final altitude of the fifty simulations.

Results

This control strategy worked very well in our simulation. Two important parameters characterized each simulated flight of the UAV: whether the UAV found a thermal updraft, and if it found an updraft, whether it was able to avoid hitting the ground for the entire hour. The chance that the UAV will find a thermal if it flies in a straight line across the flight area is approximately 33% based on the thermal model. The sink rate of the UAV is approximately 0.35 m/s, depending on its lift coefficient, so if it fails to find a thermal in its first five minutes from its initial altitude of 100 m, it will hit the ground. If it finds a thermal, however, its chances of staying aloft greatly increase. It is for these cases that the control policy has the greatest effect, since the policy affects how the UAV reacts when it finds a thermal updraft. Figure 3 shows an example flight path in red, with thermals in blue. The UAV has found two thermals, centered itself in their cores where the highest vertical velocities are found, and spiraled upward as high as possible.

We ran two of our control policies, the initial one from which stochastic gradient ascent was run and the final optimized policy, for 1000 simulations each in order to gauge their average performance. The initial and final policies are summarized in Table 2 below. For the initial policy, the UAV found thermals in 34.5% of the simulations. When it did find at least one thermal, it stayed aloft for the whole hour in 75.7% of the simulations.

Table 2 Control Policy

	A_w	A_φ	B
Initial	0.1	0.05	0.65
After 40 Iterations of Policy Search	0.073	0.084	0.745

For the optimized policy, the UAV found thermals in 37% of the simulations. When it did find at least one thermal, it stayed aloft for the whole hour in 92.7% of the simulations. We conclude from these results that our policy search greatly increased the efficiency of our control of lift coefficient, allowing the UAV to stay aloft for longer flights.

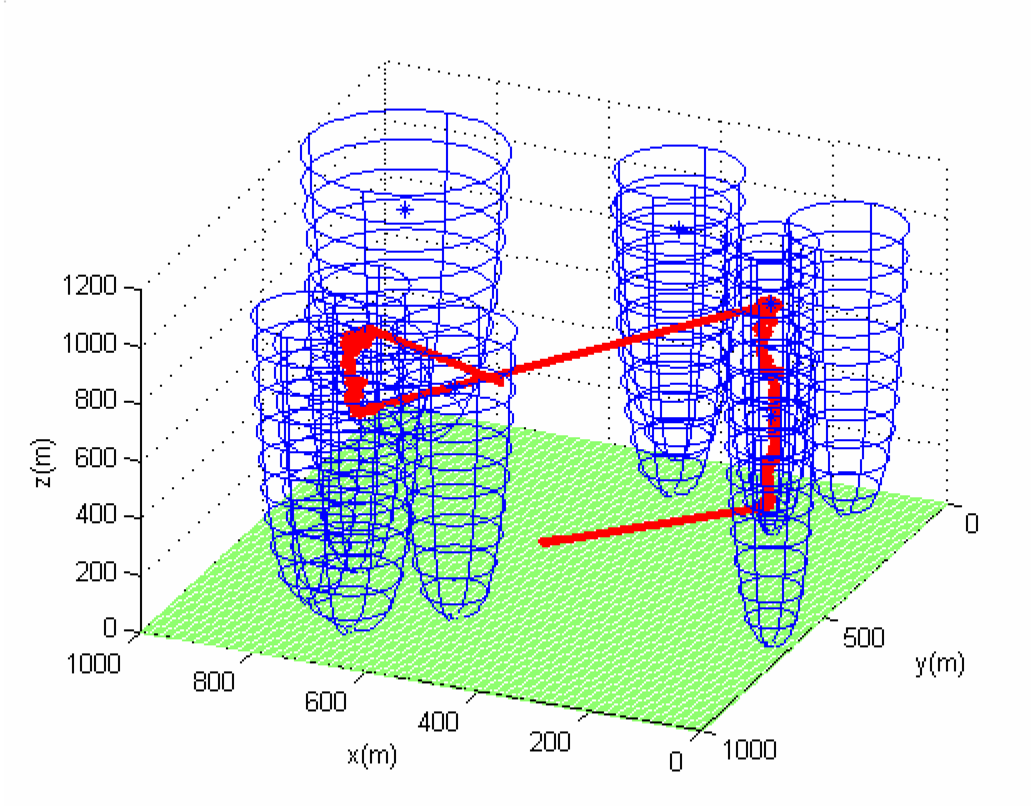

Figure 3. Sample Simulated Flight Path

Future Work

Several avenues are open for future work on the subjects that this project explored. As mentioned above, increasing the complexity of the thermal and aircraft dynamics models would help to increase the realism of the simulation. Additional control strategies could be explored, and the policy search could be refined. We also hope to eventually test the control algorithm on a real UAV platform.

Acknowledgements

Thanks are due to Professor Ilan Kroo for advice on how to proceed with this project, and to Chinmay Patel for providing his airplane as the basis for our model.

References

[1]Allen, M. J., "Autonomous Soaring for Improved Endurance of a Small Uninhabited Air Vehicle," AIAA 2005-1025, *43rd AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, 10-13 January, 2005.

[2]Allen, M. J., "Updraft Model for Development of Autonomous Soaring Uninhabited Air Vehicles," AIAA 2006-1510, *44th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, 9-12 January, 2006.

[3]Wharington, J., "Autonomous Control of Soaring Aircraft by Reinforcement Learning," PhD Thesis, Royal Melbourne Institute of Technology, Melbourne, Australia, November 1998.