Wind-Fields based Path Planning for UAV's using Markov Decision Processes

Motivation

Small, electric-powered, Unmanned Aerial Vehicles (UAVs) have been widely developed for use in both military and civilian applications. Such aircraft can be used for many applications such as coastal or border surveillance, atmospheric and climate research, as well as remote environment, forestry, agricultural, and oceanic monitoring and imaging for the media and real-estate industries. However, one of the main problems faced by small UAV's is their flight endurance regard to the limitations of the possible on board fuel/battery which can be carried by the UAV. Significant energy can be obtained from the environment if the energy sources can be exploited wisely. Glider pilots and birds frequently use winds to improve range, endurance, or cross-country speed.

There are three sources of wind energy available to exploit for this problem:
1) Vertical air motion, such as thermal instabilities, orographic lift or wave.
2) Spatial wind gradients, such as horizontal shear layers.
3) Temporal gradients, causing horizontal gusts.

Although we can exploit all of these, difficulty arises due to the high variability in wind magnitude and direction. This is compounded by the difficulty to precisely forecast wind magnitude and direction and at multiple altitudes at different times. The magnitude and direction of the wind significantly affects the onboard power. Thus, optimal path planning considering variable and uncertain environmental conditions (horizontal wind, vertical wind) is a high importance for these vehicles to increase their efficiency by maximizing flight duration and minimizing power consumption.

Problem Statement

We consider the motion of a UAV in a plane in the presence of wind fields. The UAV is supposed to go from an initial state to a finish/goal state using the wind-energy so as to minimise the power consumption and thus also obtain an optimal path. Here we consider the velocity of the UAV to be constant and equal to 20 m/s.

- A 7x7 Grid-World is constructed with a specified initial and goal state.
- A Wind-Field distribution is incorporated in the grid using the concept of vector fields which gives us a wind vector for each state/point in the 2D space.

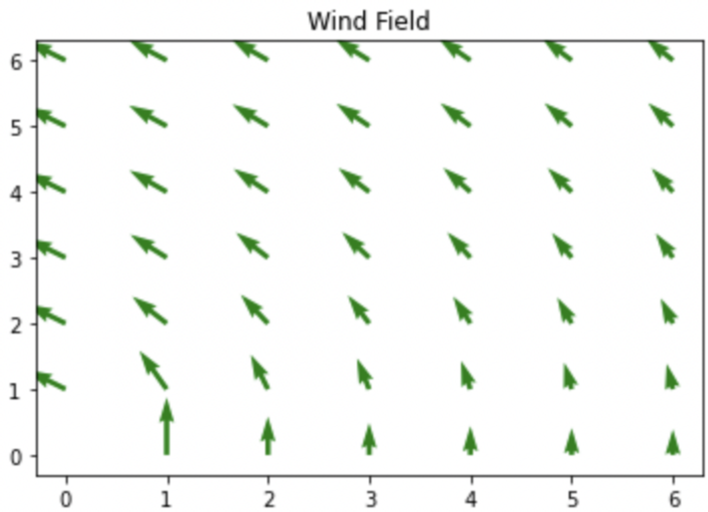

Figure 1: Wind Field Distribution

- Our objective is find the optimal wind-energy based path for the UAV in the presence of the wind-field distribution which will minimise the onboard electric power consumption of the UAV.

Approach

The motion-planning problem is to select the actions that minimise the power consumption of the UAV and minimise time-to-goal. This problem is thus naturally posed as a **Markov Decision Process** $(S; A; P; R)$, where:

- S denotes the set of possible states of the aircraft
- A is the set of actions available from each state
- P represents the transition probabilities $P_a(s_i; s_j)$ where (s_i) is the current state and (s_j) is the possible next states under action (a)
- R defines the expected immediate reward for each transition and each action (a).

MDP Formulation:

- *Possible states* (S)*:* the number of possible states will be equal to the number of cells in the discretized grid (7x7=49 in our case). The Cartesian coordinates of the state of the UAV at the centre of a cell will be denoted by $S_{i,j} = x_{i,j}, y_{i,j}$ where $x_{i,j}, y_{i,j}$ denote x position, y position cell (i,j) respectively. An important assumption is that the velocity of the aircraft is constant and equal to the Minimum Level-Flight Speed (V_{min}).
- *Actions from each state* (A)*:* we assume that the UAV can move in eight directions, $A = E,\ NE,\ N,\ NW,\ W,\ SW,\ S,\ SE$ as shown in Fig.(2). where taking the action E means the heading angle (ψ) is equal to zero degree.
- *Transition probabilities* (P)*:* The transition probabilities $P : P_{s,a}(s, \acute{s})$ manage the probabilities of what state $\acute{s}$ is entered after executing each action A from state s. We developed

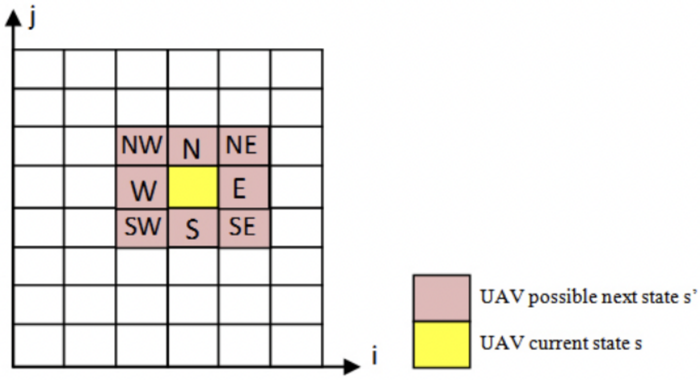

Figure 2: A graphical representation of the eight possible end locations for the eight given actions of the UAV from a starting point in yellow and an ending state in pink

a method based on Gaussian distribution to assign a realistic transition probability $P_{s,a}$ in the wind field to fit inside the MDP framework.

To determine the transition probabilities $P : P_{s,a}(s, \acute{s})$ the vector of the UAV velocity and the chosen vector of wind velocity at $cell_{i,j}$ are added. The summation result of the two vectors are represented by the magnitude $\overrightarrow{\mathrm{F}}$ and direction using the following equation :

$$F_x = V_{min}\cos(\psi_{i,j}) + W_{i,j}\cos(\theta_{i,j}) \tag{1}$$

$$F_y = V_{min}\sin(\psi_{i,j}) + W_{i,j}\sin(\theta_{i,j}) \tag{2}$$

$$\overrightarrow{\mathrm{F}} = \sqrt{F_x^2 + F_y^2} \tag{3}$$

$$\omega = \arctan(\frac{F_y}{F_x}) \tag{4}$$

$$\tag{5}$$

Figure 3 shows the normal distribution of transition probabilities (P) by setting ω from Eq. (4) as the mean value of a Gaussian distribution with standard deviation σ_ω in each cell. The Standard deviation is selected by us(1 in our case). The transition probabilities (P) will be represented by the area governed by the intersection between the curve and the range angle (Green line) for each state Eq. (6).

$$P : P_{s,a}(s, \acute{s}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\theta-\frac{\pi}{8}}^{\theta+\frac{\pi}{8}} e^{-\frac{1}{2}\frac{v-\omega}{\sigma}^2} dv. \tag{6}$$

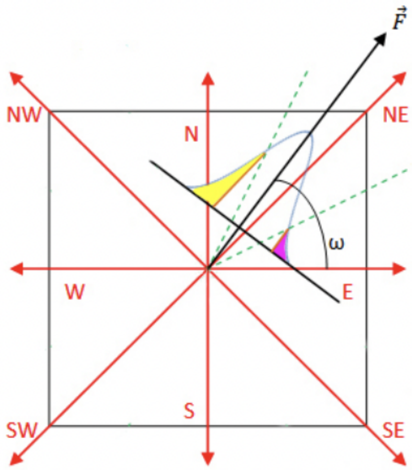

Figure 3: The normal distribution of transition probabilities

- *Reward for each transition and each action (R):* The direct reward value is calculated based on the wind component facing the target cell. C is the weight given by us as 30.

$$R_a(s_{i,j}) = (\frac{W_{i,j}\cos(\theta_{i,j} + \theta_T)}{W_{max}})C$$

- *Value Function:*

$$V(s_{i,j}) := E[R_a(s_{i,j}) + \gamma \sum P_{s,a}(s, \acute{s})V(\acute{s})]$$

  The optimal value function for a cell will be given by

$$V^*(s_{i,j}) := max_a E[R_a(s_{i,j}) + \gamma \sum P_{s,a}(s, \acute{s})V(\acute{s})]$$

  The factor γ represent the time ratio $(1 > \gamma > 0)$.

  Identifying the optimal values will lead to determining the optimal policy:

$$\pi^*(s) := arg\ max_a(R_a(s_{i,j})) + \gamma \sum_{\acute{s} \in S} P_{s,a}(s, \acute{s})V^*(\acute{s})$$

  Following the optimal policy will lead to the optimal path.

Results

Implementing Value Iteration, we obtain some fairly good results.

What was achieved:

After obtaining a policy by value iteration, the policy is played and we obtain quite intuitive results as displayed below:

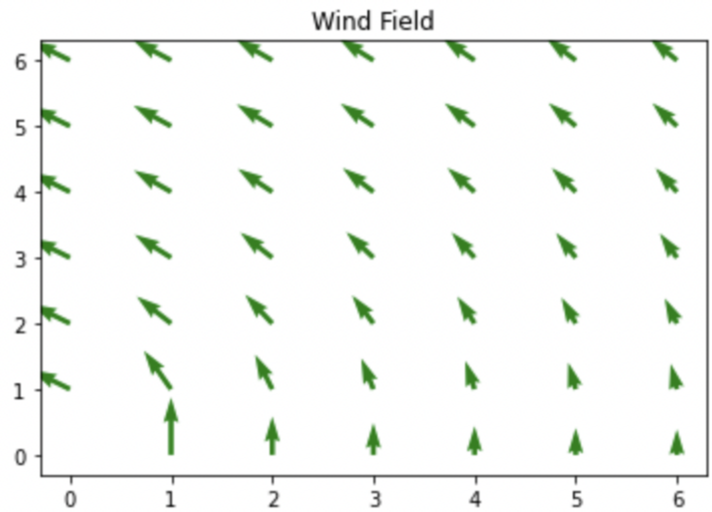

Figure 4: Wind Field Distribution

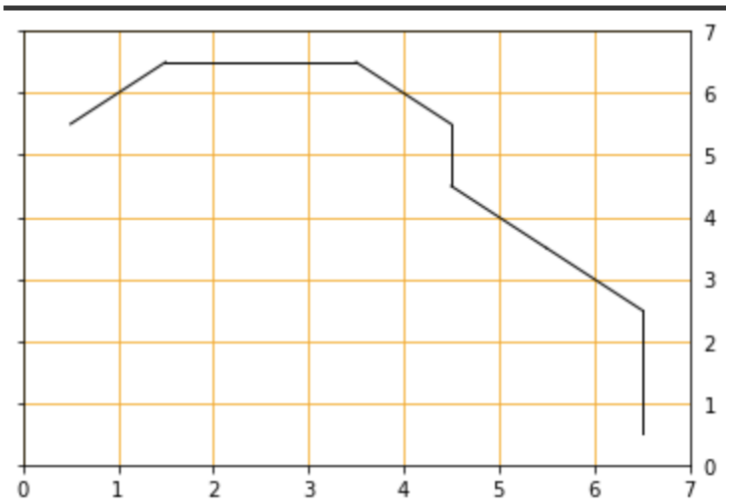

Figure 5: The Path obtained from (6,0) to (0,5)

What could not be achieved:

We used a static wind field and could not incorporate a time-varying wind field.