

eVTOLにおける遷移飛行を考慮した エネルギー効率化の研究

小川 裕生^{1,a)} 渡辺 魁 中沢 実^{1,b)}

概要：近年、垂直離着陸が可能な次世代のモビリティ「空飛ぶクルマ」が世界中から注目を集めている。本研究では「空飛ぶクルマ」の中でも飛行機のような固定翼とドローンのような回転翼を持った eVTOL 機の課題に着目した。eVTOL 機は 2 つの翼を使い分けて飛行するため、固定翼のみで飛行する状態、回転翼のみで飛行する状態、状態を切り替えるための遷移飛行の 3 つの状態を巡航やホバリングなど、状況によって使い分けて飛行している。現状では人の手によって経路と飛行状態を設定し飛行させる手法が一般的であり、すべて操縦者の経験や判断に頼った運用をしている。そこで本研究では、エネルギー効率を考慮したウェイポイントの選択を強化学習によって学習することで、効率の良い経路と飛行状態の選択を求めることを目的としている。シミュレーションにおける報酬は実際の eVTOL 機で用いているモーターで推力試験を行い、強化学習エージェントへ渡す報酬値へ反映した。

1. はじめに

近年、垂直離着陸が可能な次世代のモビリティ「空飛ぶクルマ」が世界中から注目を集めている。「空飛ぶクルマ」は正式名称を electric Vertical Take-Off and Landing Aircraft (eVTOL) と呼び、垂直に離着陸が可能かつ長距離の巡航を両立した回転翼と固定翼を持つハイブリッドな機体である。eVTOL 機が注目されつつあるのは、電池容量の増加と、固定翼を使用した効率のよい飛行によって運用に耐えうる航続距離を確保できるようになったからであると考えている。また実用化の動きも世界中の企業で進められており、日本でも「空の移動革命に向けた官民協議会 [1]」が定期的に行われ、国を挙げて eVTOL の開発が進められている。eVTOL 機は、垂直離着陸や航続距離の長さといった機体の特性を活かすことで、都市部の渋滞緩和や移動の困難な離島間や山岳部における交通手段、緊急時の物流など様々な分野に対する一つの解決策として期待されている。アメリカでは Uber や Joby Aviation, ヨーロッパでは Airbus, 中国では Ehang, 日本では SkyDrive など多くの企業が業界に参入している。また、こうした eVTOL を活用しようとする動きは企業のみならず、オープンソースソフトウェア (OSS) のコミュニティである PX4 や ArduPilot に

おいても注目されており、開発が進められている現状がある。しかし、多くの魅力的な特徴を持つ eVTOL にも欠点があり、ドローンのような回転翼、飛行機のような固定翼を状況によって使い分けて飛行しなければならないため、操縦や経路決定の複雑さが一般的なドローンや航空機に比べて高い。そのため、飛行状態を切り替える遷移飛行や、状況に合わせた飛行状態の選択など、ユーザーや操縦者にとって経験や知識が求められ、開発者や OSS のフライトコントローラ (FC) を使うユーザーにとって難しい状況になっている。そこで、本研究では回転翼と固定翼をどのように使い分ければエネルギー効率の優れたフライトが可能になるのかウェイポイントを用いた飛行と強化学習を用いることで検証した。また、本研究の貢献を以下にまとめる。

- (1) 遷移飛行を用いた eVTOL の状態の追加付けをエネルギー効率の観点から自動化する試みであること
- (2) ダイナミクスの正確性を求める制御や機体形状の検証などではなく、シミュレーションの特徴を活かしたユーザーの知見に役立つ結果を示すこと
- (3) 実機におけるホバリングと推進の定常飛行状態からそれぞれ消費電力の比を推定し、消費電力の概算が難しいシミュレーション環境内で報酬値として利用したこと

2. 関連研究

eVTOL に関する研究では、遷移飛行の実現へ向けたアプローチや実機における性能試験などは散見されるが、既

¹ 金沢工業大学大学院工学研究科情報工学専攻
Department of Information Engineering Kanazawa Institute of Technology

a) b6902041@planet.kanazawa-it.ac.jp

b) minoru.nakazawa.jp@ieee.org

存の eVTOL の運用や操作を含めた研究は少ない. ここでは eVTOL の遷移飛行に関する研究やドローンにおけるエネルギー効率の研究をまとめる.

2.1 Learning To Fly: Computational Controller Design For Hybrid UAVs With Reinforcement Learning

Jie[2] らは固定翼と回転翼を持つ機体のコントローラの生成を PPO という強化学習の手法で実現した. 学習環境には機体のモデルを自作で用意し, 単純なホバリングから遷移飛行を使った目標地点までの移動など実現の難しいとされていたタスクを達成している. 学習には位置, オイラー角, 速度, 角速度が 3 次元あり, 合計 12 次元の状態変数をもって学習には大規模なリソースが用いられた. また, この分野において固定翼や回転翼のようにダイナミクスの異なる機体形状のモデルに対してコントローラを作成するという点で優位性がある. 最終的にシミュレーション内でモデリングした 5 種の異なる機体を実機にコントローラを転用し, 試験飛行を行い同様の運動性能を実現した.

2.2 Energy-aware Goal Selection and Path Planning of UAV Systems via Reinforcement Learning

A.E.Niaraki[3] らは時間的に変化する風場下でスタートからゴールまでの経路計画と電力最適化問題を方策オフ型の Q-learning と方策オン型の SARSA アルゴリズムを用いて学習を行った. 学習環境は Gazebo 上で実装しており, 消費電力は独自に風の履歴を取り込むことで, 様々な条件下で現実的な電力応答を実現している. 学習に使った機体は Meyer らによって開発された Hector Quadrotor を使用しており, 4 つのモーターの電圧を元に 255 の範囲でパルス幅を与えて消費電力を求め, 正規化した. Gazebo シミュレーション内では, 消費電力などの再現がなされていないので, 独創的な手法を使って消費電力を表現したという点で優れていると考えた. 訓練では, エージェントは一定区間で区切られた x, y 座標を左右に移動できるのみで, 充電スポットやゴールも 1 マスとして表現されエージェントはマスの環境で実験を行った. 強化学習は疎なタスクに対して成果を上げることが困難になりやすいが, チャージングスポットや風場など探索の位置を調整することで収束しやすいエピソード数を増やす試みに成功している. 最終的に, 大規模なフィールドにおける様々な外乱条件下での強化学習の有用性を示した.

3. システム構成

本論文では, エネルギー効率の観点からあらかじめフライトモードと座標が設定されたウェイポイントをどのように選択してゴールまで到達するべきかを明らかにする. シ

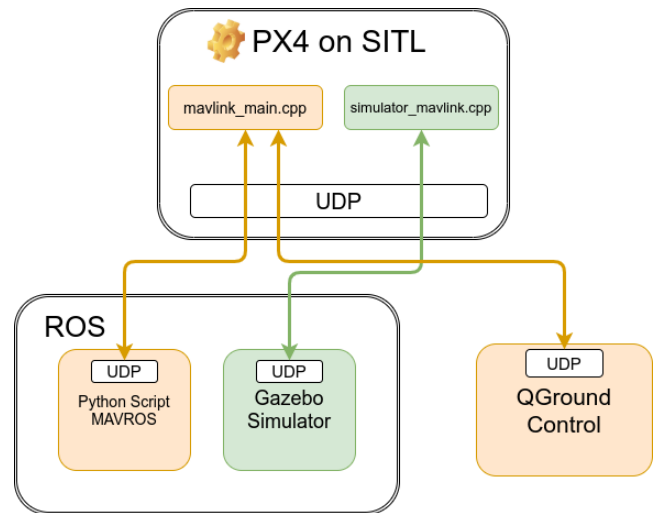


図 1 全体アーキテクチャ

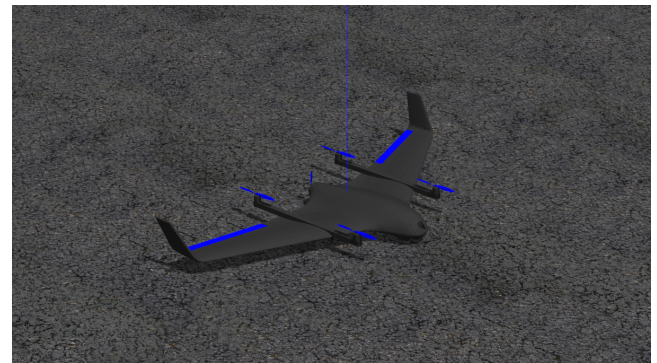


図 2 学習に使用する standardVTOL

ミュレーション内でドローンをウェイポイント航行させ, 消費電力の少ない方法でウェイポイントを選択していくよう学習させる.

3.1 全体の構成

学習に使用したシミュレーション環境のシステム構成を図 1 に示す. フライトコントローラ (FC) は OSS の PX4^{*1} を利用し, Software in the loop (SITL) で動作させている. そのため, シミュレーションから学習まで Ubuntu 上ですべて処理できる. ROS のノードである Python スクリプトは強化学習や機体操作を主に行っており, SITL でビルドされている FC に対して UDP で通信し接続している. 機体は基本的に Mavlink のプロトコルに従って動作しているため, Python で Mavlink を扱えるようになる MAVROS^{*2} によって操作が行われる. シミュレーションでは eVTOL 機の形状に近い, standardVTOL のモデルを学習に用いた.(図 2)

3.2 ROS

主に ROS 上で動作しているノードは 3 つある. 1 つは学習用のノードで, 2 つ目は MAVROS を用いて機体进行操作

*1 <https://px4.io/>

*2 <https://github.com/mavlink/mavros>



図 3 GroundControlStation

するノード, 3つ目は Gazebo シミュレーターである. 学習実行時にエピソードが終了した段階で環境の初期化をするため, 学習用のノードと Gazebo シミュレーター間で通信する必要があった. 本来, Gazebo は ROS 上で動かす必要はないが, ここでは ROS のノードとして実行している.

3.3 QGroundControl

QGroundControl^{*3}は機体と通信し, 図 3 のようにウェイポイントの設定など GUI で機体を操作できる地上管制用の OSS である. 本研究では, QGroundControl から機体を操作することはないが, SITL 上の FC と通信することができるため学習の様子を可視化するために使用した.

4. 提案手法

本節ではウェイポイントリストの生成手法と生成されたウェイポイントから, エネルギー効率のよいウェイポイントリストを取得するためのアルゴリズムの説明を行う.

4.1 ウェイポイントリストの作成手法

ここでは, 図 4 のようにあらかじめウェイポイントになりうる候補を用意する. 各ウェイポイントに 3 つの高度と 2 つのフライトモードが用意されているため, 全てで 120 箇所のウェイポイントが存在している. ウェイポイントリスト生成時はこれらのウェイポイントをスタートから隣接するウェイポイントを順番に選択していくことでゴールを目指す. また, ゴール方向へ近づく方向へしか移動はできない仕様となっている.

4.2 ウェイポイント飛行アルゴリズム

ウェイポイント飛行アルゴリズムでは強化学習手法の Q 学習を用いる.

強化学習のエージェントは収益の最大化を目指し, 自分の行動に対する環境の応答によって学習を進めていく. Q 学習は価値反復法のアルゴリズムであり, 状態 $S(t+1)$ の

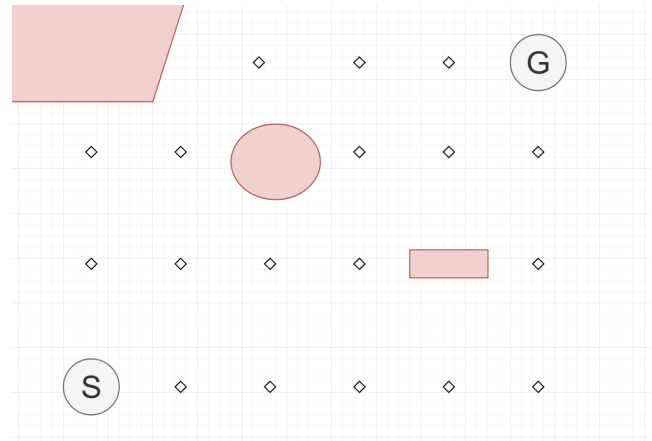


図 4 waypoint 候補のマップ

行動価値関数のうち, 最も大きいものを更新に使用する.[4] 行動価値関数 Q の更新が行動の決定方法に依存しないため, 方策オフ型と呼ばれ更新時に ϵ -greedy 法から生まれるランダム性が更新式に含まれない. そのため, 方策オン型の SARSA などに比べ, 行動価値関数の収束が早いという特徴がある.

ウェイポイント飛行アルゴリズムで用いる行動価値関数を以下の (1) 式に示す. S は状態, A は行動, R は報酬値, η は学習率, γ が時間割引率である. ここでは, 価値更新式と行動を決定するための方策は異なる.

$$Q(s_t, a_t) = Q(s_t, a_t) + \eta * (R(t+1) + \gamma \max_a Q(s_t, a) - Q(s_t, a_t)) \quad (1)$$

4.1 節で生成されるウェイポイントリストに対してステップごとに次のウェイポイントを決定しその場で反映させると機体の遷移飛行の都合上ウェイポイントリストが終了したと見なして, 一旦停止してしまう問題があった. そのため, 継続的に飛行を完遂するためウェイポイントをあらかじめゴールまで設定しておき, 1 エピソード終了の段階で報酬値による行動価値関数の更新を行っている.

行動決定のための方策は ϵ -greedy 方策とし, 確率 ϵ でランダムな行動を選択し, 確率 $1 - \epsilon$ で状態 S において最も価値の高い行動を選択する. 一般的に ϵ の値が高いほどランダムな要素が増え探索の幅が深まる. 逆に ϵ の値が低いほどこれまでの経験を重視するようになり, 局所的な解に陥ってしまう可能性が上がる.

5. 実験

本設ではホバリング用プロペラと推進用プロペラの性能試験と, 強化学習の結果について述べる.

5.1 pwm と推力の関係性

学習環境である Gazebo では, 電力の計算が行われていない. そのため, 学習に用いる報酬値返すために eVTOL の

*3 <http://qgroundcontrol.com/>

表 1 実機で使用しているプロペラの諸元

	Propeller diameter	Propeller Pitch
APC 10 x 5E	10	5
APC 12 x 5E	12	5

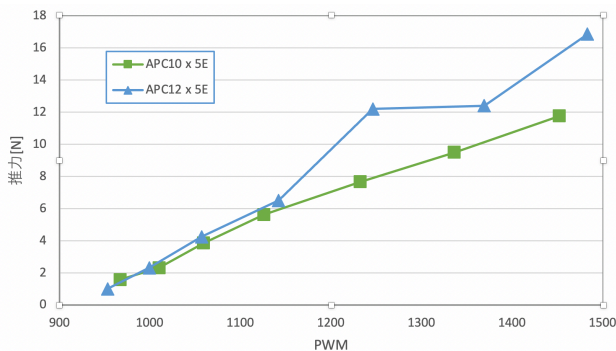


図 5 ホバーと推進用プロペラにおける PWM と推力の関係

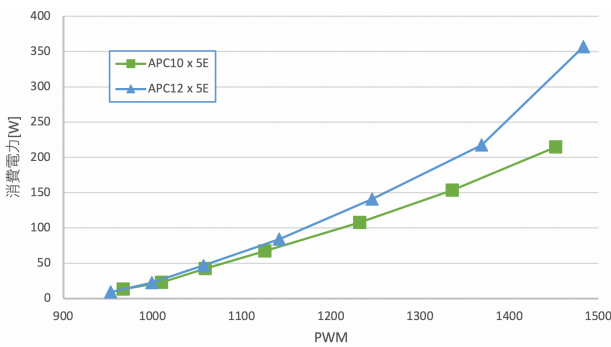


図 6 ホバーと推進用プロペラにおける PWM と消費電力の関係

各サーボで使用した PWM を信号の値を利用した。小型の eVTOL で実際に使用しているモーターを利用し、ホバリングと推進用のプロペラで PWM と推力、PWM と消費電力の関係を調査した。実機ではホバリング用と推進用で同じモーターを使用しているため、プロペラを変えててんびんを用いた推力実験を行った。ここでは、正確な消費電力を計算することが目的ではなく、ホバリング用のモーターを使った割合と推進用のモーターを出力したエネルギーの割合によって報酬値に比を設定することが目的である。実験で使ったモーターと、プロペラの諸元を示す。(表 1) PWM 信号と推力の関係を図 6 に、PWM 信号と消費電力の関係を 7 に示す。PWM が 1240 で高い推力の数値を取っているが、これは共振点ノイズが大きく乗ってしまったためだと考えられる。図 6、図 7 ともにプロペラの形が大きいと PWM が増加するほど径が大きいほど値が大きくなっている。

5.2 ウェイポイント飛行を用いた学習

ホバー用プロペラと推進用プロペラから得られた PWM と推力の比を報酬値に反映させ、学習を行う。図 6 のように 1 エピソードごとにスタートからゴールまでのウェイポイントリストを用意し、ウェイポイントを移動するごとに

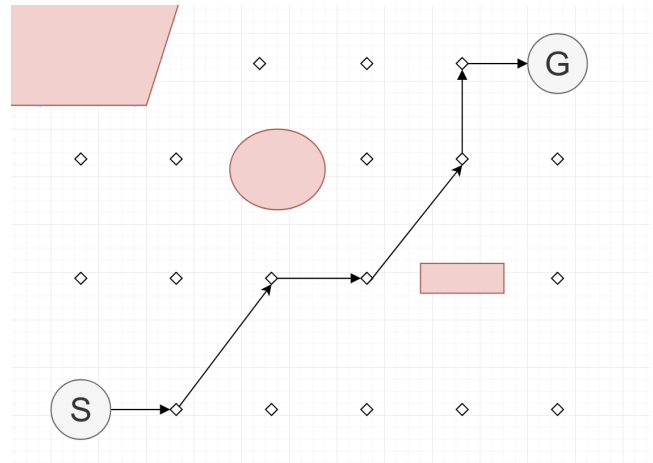


図 7 1 エピソードのウェイポイントリストサンプル

PWM の累計から計算した報酬が各ウェイポイントの価値観数を更新していく。スタートとゴールも含めると、2 つのフライトモードと 3 種類の高度があるため、実際にはウェイポイントは 120 箇所存在する。

5.3 実験結果

実験結果では、以下のことを予想している。

- 効率よく長い距離を飛行する場合、プレーンモードで飛行することが望ましい
- ホバリングモードで飛行することをなるべく避けるため、プレーンモードで飛行する距離を伸ばし可能な限り直線を飛行できるウェイポイントを設定するようになる
- 進入禁止エリアや障害物付近において、ウェイポイントを結んだ軌跡から外れることは可能な限り避けるべきであり、その点において、ホバリングモードを使い分ける必要がある
- 高度を稼いでからプレーンモードで遠くまで飛行する飛び方が最も効率が良くなる
- トランジションを用いることによって、急な角度でも旋回が可能になり操縦者の想定に近い操縦が可能になった

6. おわりに

本論文では強化学習の 1 種である Q 学習を用いてエネルギー効率が高くなるウェイポイントと飛行モードを選択するシステムの提案を行った。本システムはシミュレーションで行うため、詳細なダイナミクスや電力の計算を行っているわけではないが、制御の外側にある人の判断に関わるウェイポイントの設定など実用的な面で参考にされることを期待している。また、実験結果から遷移飛行を効果的に使うことでエネルギーを効率よく運用できることを予想している。基本的に eVTOL においてエネルギーを重視して飛行する場合固定翼を使ったプレーンモードで巡航すること

が適していると予想しているが、これは推進用のローターのみで飛行するほうがホバリングしながら飛行するよりもエネルギー効率がいいことから同様のことが言える。

今後の展望として、SARSA などの異なるアルゴリズムを使用することや ϵ -greedy 方策の調整を行い収束までの時間や学習結果に差が出るか検討していく必要がある。学習アルゴリズムの報酬値の計算では、可能な限り現実に則した値を用いることが正確な学習の要素であると考えているため、モーターの消費電力モデルの導入などシミュレーションの拡張を検討している。また、現在はウェイポイントを予め用意してリストを生成して学習を行っているため汎用性に乏しいが、避けるべき障害物や人口の集中地区を考慮したウェイポイントの生成の自動化も視野に入れていきたいと考えている。

参考文献

- [1] 経済産業省: 空の移動革命に向けた官民協議会, 経済産業省 (オンライン), 入手先 (https://www.meti.go.jp/shingikai/mono_info_service/air_mobility/index.html) (参照 2020-11-20).
- [2] Xu, J., Du, T., Foshey, M., Li, B., Zhu, B., Schulz, A. and Matusik, W.: Learning to Fly: Computational Controller Design for Hybrid UAVs with Reinforcement Learning, *ACM Trans. Graph.*, Vol. 38, No. 4 (online), DOI: 10.1145/3306346.3322940 (2019).
- [3] Asli, A. E. N., Roghair, J. and Jannesari, A.: Energy-aware Goal Selection and Path Planning of UAV Systems via Reinforcement Learning (2020).
- [4] 小川雄太郎: つくりながら学ぶ! 深層強化学習 PyTorch による実践プログラミング, 株式会社マイナビ出版 (2018).