

BIM と強化学習を用いたドローンナビゲーションに関する基礎的研究 Autonomous Drone Navigation Using Reinforcement Learning and BIM

○高添 祥太朗^{*1}, 志手 一哉^{*2}
Shotaro TAKAZOE^{*1} and Kazuya SHIDE^{*2}

^{*1} 芝浦工業大学大学院理工学研究科建築学専攻・修士課程

Master Course, Graduate School of Science and Technology, Shibaura Institute of Technology

^{*2} 芝浦工業大学建築学部建築学科 教授・博士（工学）

Prof., School of Architecture, Shibaura Institute of Technology, Dr.Eng

In this study, we aim to streamline the operations at construction sites by proposing a method for autonomous drone patrol and room-by-room photography in indoor spaces (non-GPS environments). The planning of drone movements poses a multi-objective and multi-constraint problem, and traditional rule-based approaches are inadequate due to the complexity of the environment. Reinforcement learning methods are gaining attention as a solution. However, learning in a single environment is insufficient for adapting to diverse indoor spaces such as factories and construction sites. Additionally, the development of a simulator capable of replicating various environments is costly. Our proposed method addresses these issues by using a virtual environment built on BIM as a basis for reinforcement learning, enabling the agent to acquire specialized action plans for specific indoor environments. In real-world environments, BIM is used as a prior map, and the agent observes a combination of the self-position estimated by monocular SLAM and the environmental map. The agent's output is then transmitted as control commands to the drone for operation. This enables highly autonomous flight using low-cost toy drones equipped with only cameras.

キーワード: BIM; 強化学習; ドローン; 屋内自律飛行; ROS

Keywords: BIM; Reinforcement Learning; Drone; Indoor Autonomous Flight; ROS.

1. 研究目的

近年、建設業における労働力人口の減少が深刻な問題となっている。政府は建設業の生産性向上によりこれらの課題の打開を画策しており、建設現場の生産性を 2025 年度までに 2 割向上を目指す指針を示している¹⁾。この目標に向け、国土交通省は 2016 年度より建設現場において ICT 活用等を進める「i-Construction」を推進し、労働力の減少を上回る生産性向上を試みている²⁾。

そのひとつの手段としてドローンの活用が注目されている。土木分野では、ドローンを用いた航空測量等が急速に普及しており、屋外自律飛行も実用化されつつある。建築分野では、高所点検等にドローンが用いられており、屋内自律飛行技術の研究も進められている。しかし、自由度の高い屋内自律飛行を行うには高価なセンサーを搭載した産業用機体が必要であるため、ドローンの幅広い活用には至っていない。また、自律飛行を行う際もドローンの飛行経路設定は人間が行う必要があり、業務省力化が十分にできていない。

そこで本研究では、内装工事から竣工に至る建設現場での業務省力化を目的とし、屋内空間におけるドローンの巡回飛行と撮影を自律的に行う手法を提案する。この

手法の実現により、建設現場での業務だけでなく、竣工後の維持管理業務の省力化にも貢献できると考える。

2. 既往研究

宮川は、Unity ML-Agents を用いて強化学習を行い、マルチ UAV 協調空輸において安全な飛行経路を獲得する手法を提案した³⁾。学習の際には、空間全体をセル化しエージェントに観測させている。また、シンプルなシミュレーション環境において提案方法の性能を評価し、リアルタイムで障害物を回避して安全な飛行経路を発見できることを示した。本研究では、BIM モデルを用いてより実環境に近い学習環境を構築し、特定の建物に特化した飛行制御を学習させる。

従来の屋内経路計画では、2 次元地図上で目的地を設定し、A*アルゴリズムのような古典的アルゴリズムによって算出した最短経路をロボットの移動経路としていた。しかし、近年は屋内経路計画に BIM (Building Information Modeling) を活用する手法が注目されている。Sina Karimi らは、IFC データから BIM セマンティクスを抽出し、経路計画用のグラフマップに変換する手法を提案した⁴⁾。この手法では、各部屋を頂点、ドアを辺としたグラフ構

造の地図を作成し、経路上の各部屋の状態も考慮して 2 点間の最適経路を探索している。本研究ではこの手法を拡張し、多数の部屋を訪れる最短巡回回路を求める。

3. 研究内容

本研究は、安価なドローンを用いて、屋内空間の巡回飛行と撮影を自律的に行うことを目標としている。そこで、ホビー用ドローン(Ryze Tech 製 Tello EDU)を用いて、未知の障害物が存在する屋内環境で自律飛行するシステムを開発する。研究の流れを図 1 に、システム概要図を図 2 に示す。機体の状態や周辺環境の状況等を考慮してドローンの動作をリアルタイムに決定するために、本システムでは SLAM、経路計画、障害物回避を実行する。なお Tello EDU の機体自体にはプログラムを実行する機能は備わっていないため、母艦となるパソコンでシステムを構築している。

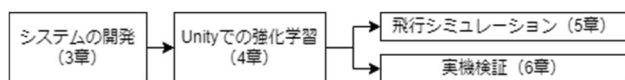


図 1 研究の流れ

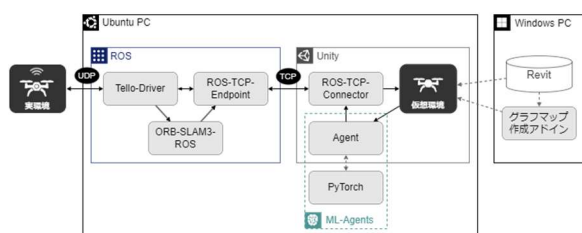


図 2 システム概要図

3.1. 機体

Tello EDU は機体重量 87g、大きさ 10cm×10cm と小型軽量の機体である。機体前方にカメラセンサーが搭載されており、映像をリアルタイムで送信することができる。また、機体下方に搭載された赤外線センサーと IMU によってある程度安定した飛行が実現されている⁵⁾。

3.2. SLAM

屋内空間では衛星電波が届かず、GPS を用いた自己位置推定ができない。そのため、屋内を飛行するドローンの自己位置推定技術として、SLAM (Simultaneous Localization and Mapping) が一般的に用いられる。SLAM とは、外界センサーから取得した情報を用いて自己位置推定と環境地図作成を同時に行う技術である。Tello EDU に標準装備で搭載されているのはカメラセンサー 1 台のみであり、LiDAR センサーや複眼カメラなどのセンサーを用いることはできない。そのため、1 つのカメラ映像のみで実行可能な単眼 SLAM (Monocular Visual SLAM) によって自己位置推定を行う。

Visual SLAM の古典的なアルゴリズムの多くは、画像全体の輝度値を用いる直接法と、画像内の特徴点を用いる非直接法に分類できる。直接法の代表的なものとして LSD-SLAM や DTAM、非直接法の代表的なものとして ORB-SLAM や PTAM が挙げられる。ORB-SLAM⁶⁾ は Raúl Mur-Artal らによって 2014 年に提案された特徴量ベースの Visual SLAM である。ORB-SLAM は比較的計算負荷が小さく、カメラの軌跡と周辺環境の点群を低遅延で取得できることから、リアルタイム処理を志向するロボットや自動運転車で広く利用されている。本研究では、ORB-SLAM の最新バージョンであり、オープンソースの SLAM 実装である ORB-SLAM3⁷⁾を採用する。

なお、単眼 SLAM では既知の長さ情報がないために絶対距離を取得できないスケール不定の問題が生じる。そこで本研究では、赤外線センサーで計測した機体高度の情報をを用いて単眼 SLAM の推定値を補正している。

3.3. 大局的な経路計画

経路計画は、事前地図と SLAM で得られた機体の位置姿勢および周辺環境の情報を用いて、現在地から目的地への飛行経路を計算する処理である。

はじめに、大局的な地図を用いて目的地となる部屋を巡回する経路を求める。本研究では、芝浦工業大学豊洲キャンパス本部棟の 8 階を対象とし、9 つの部屋を巡回する経路を求める。本部棟の BIM モデルからグラフ構造の地図を生成し、この地図から焼きなまし法で最短巡回回路を計算する。この機能は BIM ソフトウェアである Revit のアドインツールとして実装する (図 3)。



図 3 本部棟の BIM モデルとグラフ構造の地図

3.4. 局所的な経路計画と障害物回避

大局的な経路計画で計算した巡回路に沿って飛行するために、局所的な地図を用いてドローンを導く処理を実装する。この処理では、ドローンの機体モデルと制御アルゴリズムを利用し、障害物を回避して安全に飛行できるよう、制御コマンドを出力する。

本研究では、ドローンの制御に強化学習アルゴリズムの PPO (Proximal Policy Optimization) を採用する。PPO は John Schulman らによって 2017 年に提案されたモデルフリー・オンポリシーのアルゴリズムであり、実装とパ

ラメータ調整が非常に簡単でありながら高いパフォーマンスを発揮するアルゴリズムである⁸⁾。

3.5. プラットフォーム

物理シミュレーションを用いた強化学習環境を構築するために、機械学習フレームワークの Unity ML-Agents Toolkit⁹⁾を用いる。これにより、Unity のシーン（物理シミュレーション環境）と Python スクリプトを連携させて強化学習や模倣学習を実行できる。また、ドローンとの通信や SLAM 処理を実行するためのプラットフォームとして、オープンソースのロボティクス向けミドルウェアである ROS（Robot Operating System）を用いる。

ROS とドローンの通信は、ROS の「Tello-Driver」パッケージ¹⁰⁾を介して行う。このパッケージにより、Wi-Fi を介した制御コマンドの送信およびカメラ映像の受信が可能となる。また、ROS と Unity の通信は、ROS の「ROS TCP Endpoint」パッケージ¹¹⁾と Unity の「ROS TCP Connector」パッケージ¹²⁾を介して行う。

4. 強化学習

4.1 エージェントの設定

訓練シーンにおいて、エージェントは 0.02 秒毎にシーンの状態を観測し、次の行動を決定する。エージェントが 1 回行動する時間単位を 1step と定義する。エージェントの行動と観測、報酬条件を表 1 に、アルゴリズムのハイパーパラメータ設定を表 2 に示す。

エージェントの主要な観測であるグリッドセンサーは、縦横 30 セル、高さ 5 層からなり、格子状の単位でドローン周辺の障害物とゴールの位置を近似している。グリッドセンサーの見取り図を図 4 に、断面図を図 5 に示す。エージェントがより速くゴールへ辿り着くよう、step 毎に負の即時報酬を与える。過去 10 秒間で 1m 以上移動した場合は -0.0001、1m 未満の場合は -0.0002 を与える。即時報酬の値が大きすぎる場合、エージェントは「ゴールを目指すより、壁に衝突して即座にエピソードを終了した方が低損失だ」と判断してしまい、学習に失敗する。また、値が小さすぎると学習効率が悪くなる。そのため適当な値を経験的に設定した。

表 1 エージェントの行動・観測・報酬

行動	①前進 ②上昇・下降 ③左回転・右回転
観測	グリッドセンサー [30×30] × 5 層 ドローンとゴールの相対位置
遅延報酬	ゴールした場合 +1、 ゴールできなかった場合 -1
即時報酬	移動距離が 1m 以上の場合 -0.0001/step 移動距離が 1m 未満の場合 -0.0002/step

表 2 ハイパーパラメータの設定

パラメータ名	設定値
trainer_type	ppo
buffer_size	512
batch_size	10240
learning_rate	0.0003
beta	0.01
epsilon	0.2
lambda	0.95

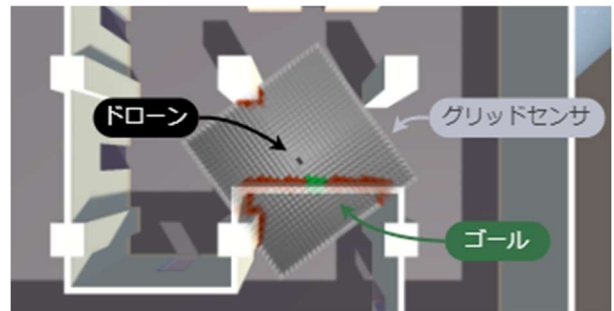


図 4 グリッドセンサーの上空視点の見取り図

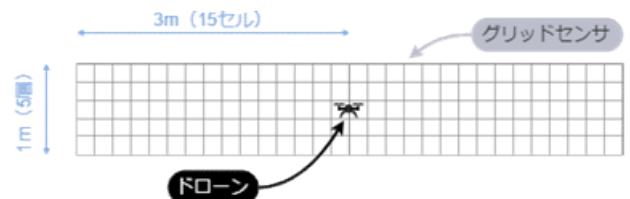


図 5 グリッドセンサーの断面図

4.2 訓練シーンの設定

5 段階の訓練シーンを用いてカリキュラム学習を行う。各シーンの上空視点からの見取り図 6 に示す。訓練シーンにドローンオブジェクトを配置し、安全にゴールまで飛行できるように学習させる。ドローンオブジェクトがゴールするまでを 1 エピソードと定義する。また、壁や床等に衝突した場合や、エピソード開始から 5000step 経過してもゴールできなかった場合、エピソードを中断しドローンオブジェクトの初期位置および方向をランダムに再設定する。

訓練シーン 1～訓練シーン 3 では、エピソード毎にランダムな位置に設置された球をゴールとし、エージェントに部屋空間とゴールという概念を獲得させる。はじめエージェントは機体をランダムに動かすだけだが、偶然「機体がゴールオブジェクトに接触する」という現象が起きた時、エージェントに正の報酬が与えられ、「ゴールすると報酬が得られる」というルールを学習できる。また、偶然「機体が障害物に接触する」という現象が起きた時、エージェントに負の報酬が与えられ、「壁や天井に触れてはいけない」というルールを学習できる。

訓練シーン 4 では、エピソード毎にランダムに選択したドアをゴールとし、ドアからドアへの飛行を学習させる。また、ゴールまで直進できない場合に、大きく迂回して飛行することを学習させる。

訓練シーン 5 では、空間を 3 層構造とし、階段を用いた移動を学習させる。

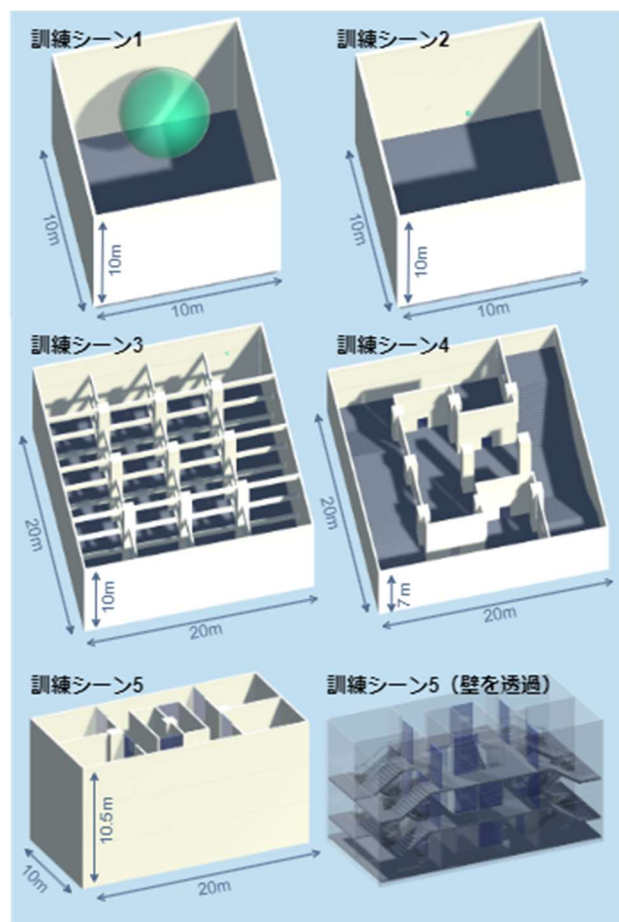


図 6 訓練シーン 1~5 の見取り図

4.3 学習結果

平均累積報酬が 0.99 を超えるまで学習を行い、その学習結果を次の学習の初期値として用いることを繰り返すことによってカリキュラム学習を行った。各訓練シーンの学習結果の推移を図 7 に示す。グラフの縦軸（左）はエージェントの平均累積報酬であり、報酬の平均値を 1 万 step 毎に示している。縦軸（右）は平均エピソード長であり、エピソードが終了するまでに要した時間の平均値を 1 万 step 毎に示している。横軸は step 数であり、百万単位 (M) で示している。グラフから、どの訓練シーンでも学習が進むにつれて平均累積報酬が増加し、平均エピソード長が減少していることが分かる。合計 8Mstep の学習を行い、学習完了までに約 19 時間を要した。

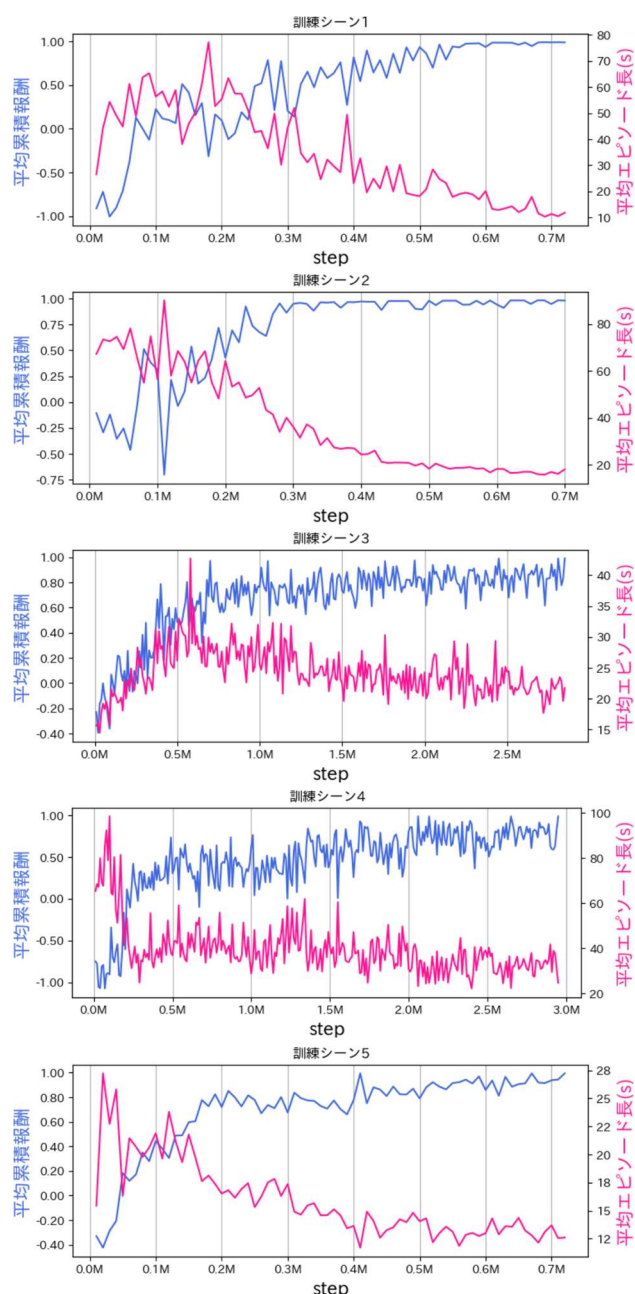


図 7 訓練シーン 1~5 の学習結果の推移

5. 飛行シミュレーション

本部棟を対象とし、3 章で求めた経路に沿って安全に飛行できるかを飛行シミュレーションで検証する。

5.1. テストシーンの設定

本部棟の BIM モデルを用いて図 8 に示すテストシーンを構築し、ドローンオブジェクトを配置する。また、ランダムな位置に障害物を配置する。障害物は直方体で、その各辺の大きさは 0.5~2m の範囲でランダムに決定される。検証では、本部棟 8 階のエレベーターホールを離着陸地点とし、3 章で求めた経路上の 9 個のドアを順にゴールとして設定する。ドローンが経路を一周して離着

The screenshot displays the ML-Drone application interface. The main window shows a 3D model of a building complex with numerous red cubes representing obstacles. A green line indicates the drone's flight path. The interface includes a top navigation bar with 'Training', 'Simulate' (selected), and 'Real' tabs. On the left, a 'Route' panel shows a progress bar at 20% and a 'Camera View' window. On the right, a 'Environment' panel lists parameters like 'max time: 01:20 s' and 'robotics: 128 s'. Below it, a 'CollectObservations' panel shows 'alpha: 1.430798' and 'beta: 0.394183'. A table at the bottom right compares 'Action' and 'Reward' values for different states.

Action	Reward
action: 0.000	reward: 0.1
action: 1	reward: 0.0
action: 1	reward: 0.0

5.2. 検証結果

次に、テストシーン内でランダムに選択した部屋を巡回するよう設定し、2Mstep の追加学習を行う。学習結果の推移を図 9 に示す。追加学習したエージェントを用いて 5 回試行すると、5 回全ての飛行に成功し、その平均飛行時間は 494 秒であった。検証結果を表 3 に示す。結果から、エージェントに追加学習を行うことで、より速く安全に飛行できるようになったことが分かる。

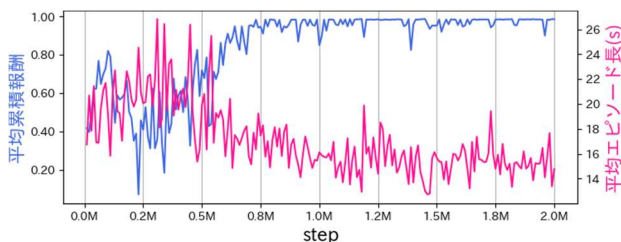


図9 追加学習の結果の推移

試行	飛行時間（追加学習なし）	飛行時間（追加学習あり）
1	記録なし	496 秒
2	541 秒	482 秒
3	記録なし	494 秒
4	記録なし	507 秒
5	記録なし	491 秒
平均	541 秒	494 秒

研究室を対象とし、実環境においてシミュレーションと同様に自律飛行できるかを検証する。実機の制御には

6.1. テストシーンの設定

図 10 研究室のテストシーン

ドローンの飛行経路を図 11 に示す。また、ドローンのカメラ映像に特徴点を重ねたものを図 12 に、自律飛行開始直後から 2 秒間隔で示す。図 11 および図 12 から、BIM モデルに含まれない家具等の障害物を点群として捉え、回避しながら飛行できていることが分かる。

なお、検証に成功するまでに複数回の試行を必要とした。ドローンから送信されるカメラ映像のコマ落ちがあり、単眼 SLAM での自己位置推定に失敗する事象が頻発したためである。原因については次節で考察する。

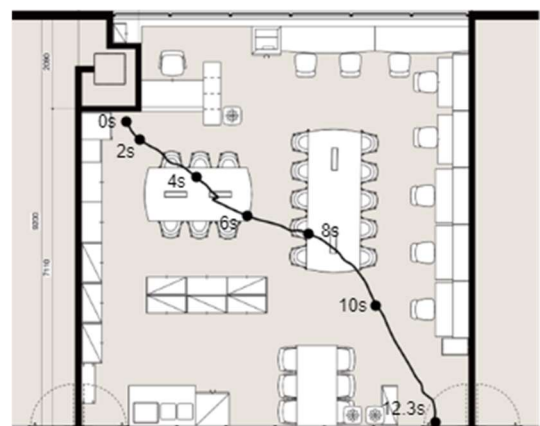


図 11 研究室の平面図とドローンの飛行経路



図 12 ドローンのカメラ映像と特徴点

6.4. 映像のコマ落ちについての考察

本研究で使用しているドローンは、Wi-Fi の 2.4GHz 帯で通信を行っているが、2.4Ghz 帯は一般的に 5GHz 帯より電波干渉が発生しやすい。とくに、実験を行った研究室には電波を発する多数の電子機器があるため、非常に電波干渉が発生しやすい状況であった。

また、ドローンはトランスポート層の通信プロトコルとして UDP (User Datagram Protocol) を使用している。UDP はリアルタイム性に優れているが信頼性の低いプロトコルであり、確認応答や再送制御などを行わない。そのため Wi-Fi の電波干渉によりパケットロスが発生しても、それを検知してデータを再送することはできない。

このように通信環境の問題とプロトコルの問題が重なったことで、ドローンから送信される映像のコマ落ちが頻発したと考えられる。

7. まとめ

飛行シミュレーションの結果から、シンプルな環境で学習したエージェントに、特定の建物に特化した飛行制御を追加学習させることで、より速く安全に飛行できることが分かった。また、実機検証の結果から、安価なホビー用ドローンを用いて屋内自律飛行を実現できることが確認できた。本研究はホビードローンのナビゲーションを目的としているが、本研究の成果は産業用ドローンの飛行経路の事前検討にも応用可能だと考える。

今後の課題としては、映像のコマ落ちへの対策と、自己位置推定および障害物検知の精度向上が挙げられる。

6 章 4 節で述べたように、Wi-Fi を用いた通信では映像のコマ落ちが発生し、自己位置推定に失敗する可能性がある。この問題の対策として、IMU 等内界センサーとのフュージョンによりロバスト性を高めることが考えられる。

また、センサーの測定値には誤差・ノイズがあるため、逐次的に SLAM を実行するだけでは誤差が累積して徐々に環境地図が歪み、自己位置を正確に推定することが困難になる。本研究では安全上の理由から研究室のみで実機検証を行ったため、SLAM の累積誤差は問題とならなかった。しかし、大規模環境においてはこの累積誤差の影響が大きくなり、シーン内のドローンオブジェクトの位置と実機の位置が大きくずれてしまう可能性がある。そのため、SLAM により得られた環境地図と BIM のジオメトリ情報をマッチングして環境地図の歪みを補正し、正確な自己位置を獲得することが必要である。

加えて、単眼 SLAM ではその原理上、動的障害物を正しく認識することが難しく、有人環境下で安全に飛行できないため、複眼カメラや LiDAR 等のセンサーの利用も検討する必要がある。

[参考文献]

- 1) 内閣府：日本再興戦略 2016, 2016
- 2) 国土交通省：I-CONSTRUCTION の推進, 2022
- 3) 宮川 勲：グリッドセンサを利用した強化学習に基づくマルチ UAV 協調空輸, 湘南工科大学学術リポジトリ, 2022
- 4) SINA KARIMI, RAFAEL GOMES BRAGA, IVANKA IO RDANOVA, DAVID ST-ONGE : SEMANTIC NAVIGATION USING BUILDING INFORMATION ON CONSTRUCTION SITE, 38TH PROCEEDINGS OF THE INTERNATIONAL SYMPOSIUM ON AUTOMATION AND ROBOTICS IN CONSTRUCTION, 2021
- 5) RYZE : TELLO EDU, <https://www.ryzerobotics.com/jp/tello-edu> (accessed 2023.4.10)
- 6) Raul Mur-Artal, J. M. M. Montiel, Juan D. Tardos : ORB-SLAM: a Versatile and Accurate Monocular SLAM System, IEEE Transactions on Robotics, 2015
- 7) GitHub : UZ-SLAMLab/ORB_SLAM3, https://github.com/UZ-SLAMLab/ORB_SLAM3 (accessed 2022.7.12)
- 8) 布留川英一：Unity ではじめる機械学習・強化学習 Unity ML-Agents 実践ゲームプログラミング, ボーンデジタル, 2018
- 9) Github : Unity-Technologies/ml-agents, <https://github.com/Unity-Technologies/ml-agents> (accessed 2022.8.20)
- 10) Github : appie-17/tello_driver, https://github.com/appie-17/tello_driver.git (accessed 2022.5.8)
- 11) GitHub : Unity-Technologies/ROS-TCP-Endpoint, <https://github.com/Unity-Technologies/ROS-TCP-Endpoint> (accessed 2022.8.20)
- 12) GitHub : Unity-Technologies/ROS-TCP-Connector, <https://github.com/Unity-Technologies/ROS-TCP-Connector> (accessed 2022.8.20)