

強化学習を用いた複数の建築物に対するドローン点検経路の自動生成

Automatic Generation of Drone Inspection Routes for Multiple Buildings Using Reinforcement Learning

○柴田 拓海^{*1}, 山田 悟史^{*2}
Takumi SHIBATA^{*1}, Satoshi YAMADA^{*2}

^{*1} 立命館大学 大学院 理工学研究科 博士課程前期課程

Graduate Student, Graduate School of Science and Engineering, Ritsumeikan Univ.

^{*2} 立命館大学 理工学部 建築都市デザイン学科 准教授・博士（工学）

Ph.D. Associate Professor, Dept. of Architecture and Urban Design, Ritsumeikan Univ., Dr.Eng.

キーワード：ドローン；維持管理；点検；強化学習；自動生成

Keywords: Drone; Maintenance; Inspection; Reinforcement Learning; Automatic Generation.

1. はじめに

1.1. 社会課題・背景

近年、温暖化などの気候変動が原因として、自然災害による被害が多く報告されている。我が国においても、世界有数の地震大国とされ、自然災害のリスクを多く抱えている。特に地震が建築物や橋梁やトンネルなどのインフラ構造物に与える影響は大きい。日本では首都直下型地震も予測されており、内閣府によればこの地震による被害は建物崩壊等と合わせると、焼失が約 61 万棟、死者が 2 万 3 千に上るとされる¹⁾。地震発生直後の建築物やインフラの早急の点検は、安全を担保する上で欠かせない課題となる。点検の遅れによる、建物倒壊、それによる新たな人的被害、建物倒壊による道路閉鎖と、人々の避難への影響などが想定される。さらに、その後の救助や消火、物資の運搬などにも深刻な影響を及ぼしうる。実際に 2016 年の熊本地震では、人員確保や点検調査体制などが原因で十分な被災状況が確認できず、被害状況の判定と通行可否の判断に時間を要したことが報告された²⁾。このような災害発生直後の建築物・インフラの点検は重要である一方で、人が行うことは、コスト・人材・安全性・迅速性といった困難な側面も多い。このような課題に対して、近年発達する AI やドローンの活用が求められている。AI による構造物の異常点検や³⁾、ドローンによる撮影の研究が多く行われている⁴⁾⁵⁾。また、点検のみならず、避難そのものへのドローンの活用も広がっている⁶⁾。

1.2. 研究目的

このような背景のもと、本研究では、災害発生時における、建築壁面のドローン点検の経路生成に関する研究を行った。災害時におけるドローンの迅速な建築壁面点検は被害を軽減することができると考える。予めドローン飛行経路を定めておくことは理想的である。しかし、すべての土地で経路を計画しておくことは困難である。さらに、飛行

経路を考える際に、様々な建物の形状や、敷地の条件、個々のドローンに応じて、その変化に対応するようにその都度プログラムに変更を加えることも現実的でない。既往研究においてこれを考慮した経路自動生成に関する研究は行われていない。そこで、構造物の壁面のドローン点検検査を想定して、飛行経路を強化学習により生成する手法について検討した。強化学習によるドローン点検の自動化は、日常的に行われるインフラ点検にも有効である。環境の構築などのシミュレーションにはゲームエンジンである Unity を用いた。強化学習を用いた自動化の例として、ドローン配送経路の最適化、クレーン揚重の自動化、車の駐車自動化などが挙げられる⁷⁾⁸⁾。筆者らは、過去に強化学習を用いて、一面、二面、三面壁面の全てを検査できることを示した⁹⁾¹⁰⁾。しかし、点検対象の複雑さと経路の最適化に課題が残った。本稿では、三面、四面、複数棟において、強化学習によるドローンの点検経路生成の手法について報告した。

2. 研究概要

2.1. 学習環境

強化学習¹¹⁾は機械学習の 1 つで、環境の状態と報酬に基づいて、最良の行動が取る方策を学習する手法である。図 1 に示すように強化学習は次の①~③の手順を繰り返し、長期的な収益値を最大化するように探索を進める。

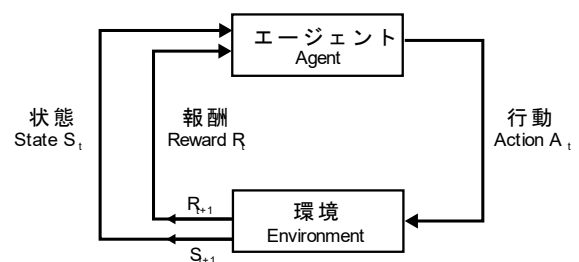


図 1（強化学習の概略）

- ①エージェントは時間ステップ t ごとに状態を取得
- ②取得結果に基づき、行動を選択
- ③選択した行動により、報酬を付与

本研究では、状態・行動・報酬の設計、学習・学習結果の確認は Unity を用いて行った。強化学習は Unity 向けの機械学習ツールキットである Unity Machine Learning Agent(ML-Agents v2.0.1)¹²⁾ を用いた。強化学習のアルゴリズムには学習アルゴリズムの 1 つであるオフポリシーの Proximal Policy Optimization(PPO)法¹³⁾を使用した。PPO はサンプル効率が低く、学習に必要なステップ数は多いが、学習が安定しているという特徴がある。また、学習に必要なハイパーパラメータの調整は個別に行った。

2.2. 環境設計

環境はゲームエンジンである Unity を使用した。ドローンと壁面については図 2 のように簡略化した。ドローンの動き方は、上下、左右、前後、左右への回旋とした。また、壁面は 1×2 の長方形(マス)の組み合わせで構成した。

壁面検査は次のような状況になることで検査をしたとみなすこととした。「ドローンのカメラを壁面に対して正面を向ける」かつ「ドローンと壁面との距離が一定以内になった時」の状況である。これは、高性能カメラの発達により、カメラが壁面をとらえることができたなら、検査は十分に行えるということに基づいて設定した。

2.3. 学習設計

ハイパーパラメータは表 1 のように設計した。1 マスを検査した時とすべてのマスを検査終了した時にプラスの報酬を与えた。範囲外に出た時、毎時にマイナスの報酬を与えた。また、直前に検査したマスの水平方向にあるマスを検査した時と、直前のマスから近いマスを検査するほど報酬が高くなるように設計した(筆者の過去の施行に基づき、学習の効率化、経路最適化の面から設定)。すべての壁面の検査を終了した時と、範囲外に出た時にエピソードを終了させ、ドローンを初期位置に戻し学習を再開させた。

観察については、ドローンの座標と旋回の回転角度を取得した。また、Unity のコンポーネントである Raycast を用いて壁面の位置を特定した。事前実験では、観察するマスの選定方法を考察した。

- A: 毎時ドローン位置から最も近いマスを観察する
- B: 直前にマスを検査したドローンの位置から最も近いマスを検査されるまで観察する
- C: 直前にマスを検査したドローンの位置から最も近いマスを観察する

以上の 3 つのマスの選定手法を用いて強化学習を行った。結果は、C の観察手法において、学習過程と学習後の経路が効率的であることが分かった。そこで本稿では C の観察マス選定手法を用いて、観察されるマスの情報を取得した。取得されるマスの情報はドローンとマスまでの距離とベクトルとした。他の具体的な数値については表 2 に示した。

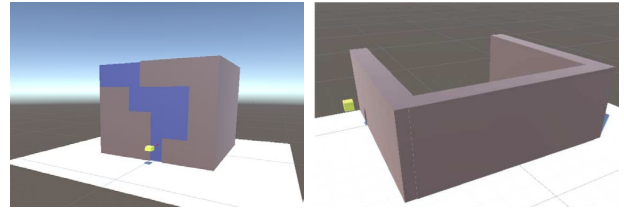


図 2 (ドローンと壁面の簡略化)

表 1 (学習時のハイパーパラメータ)

Hyperparameters	Network_settings
batch_size: 128	normalize: false
buffer_size: 2048	hidden_units: 256
learning_rate: 0.0003	num_layers: 2
beta: 0.01	vis_encode_type: simple
epsilon: 0.2	Reward_signals
lambda: 0.95	gamma: 0.99
num_epoch: 3	strength: 1.0
learning_rate_schedule: linear	

表 2 (学習環境の設計)

行動 Action	ドローン上下左右前後移動 (x, y, z 座標移動)	
	ドローン旋回 (y 軸まわりの回転)	
状態観察 Observation	ドローン x, y, z 座標	3 observations
	ドローン y 軸回転角度	1 observation
	ドローンとマスまでの距離・ベクトル	1 observation
	RayCast Observation	3 rays
報酬 Reward	垂直にある壁の検知をした時	$+1/(10^3 \cdot \text{distance}^{*1})$
	それ以外の壁面を検知した時	$+1/(10^3 \cdot \text{distance}^{*1})$
	全ての壁面を検知し終えた時	+1.0
	範囲外・毎時	$-1/10^{-3} \cdot -1/10^{-6}$

※distanceは直前に検査したマスと現在検査したマスの距離を示す(最小値=1.0)

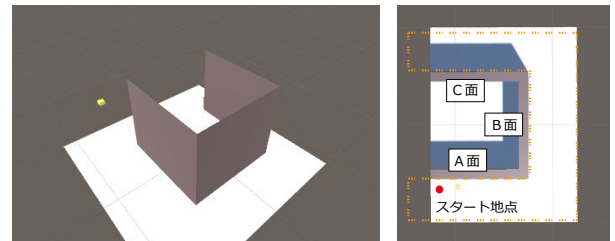


図 3 (三面壁面の環境, ドローンの行動範囲)

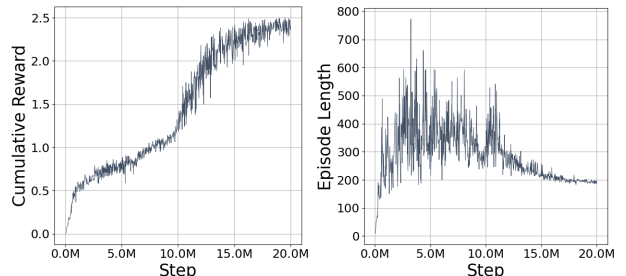


図 4 (平均累積報酬とエピソードの長さの推移)

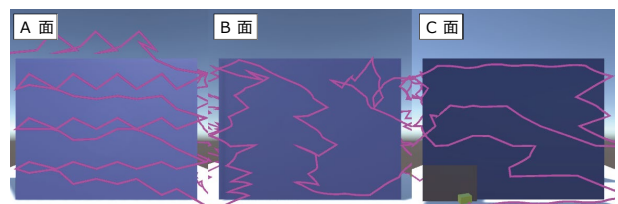


図 5 (学習後のドローン検査経路展開図)

3. 三面壁面における経路生成

3.1. 概要

三面壁面を「コ」の字型に配置しシミュレーションを行った。配置した壁面の位置と、ドローンの行動可能範囲は図3に示した。エピソード開始時のドローンの開始地点は赤点で示した。

3.2 学習結果

学習による平均累積報酬とエピソードの長さの推移を図4に示した。学習後のドローンの検査経路の1つを図5に示した。紫線はドローンが通った軌跡を表した。すべての壁面の検査が完了するまでに要した距離は、250.0 となった (単純最短距離は 128)。

4. 四面壁面における経路生成

4.1. 概要

四面壁面を「ロ」の字型に配置しシミュレーションを行った。配置した壁面の位置と、ドローンの行動可能範囲は図6に示した。エピソード開始時のドローンの開始地点は赤点で示した。報酬・状態観察については、3.の三面壁面と同様の設計とした。1回の学習では全面を検査できなかったため、カリキュラム学習を行った。

カリキュラム学習は、タスクの難易度を徐々に上げて、より複雑なタスクの攻略を可能にする学習手法である。ここでは、学習を2つのパートに分けた。Lesson1では0~20M stepsの区間に三面壁面までの学習を行い、Lesson2では20~30M stepsの区間四面目の壁面の学習を行った。

4.2. 学習結果

学習による平均累積報酬とエピソードの長さの推移を図7に示した。学習後のドローンの検査経路の1つを図8に示した。すべての壁面の検査が完了するまでに要した距離は、291.5 となった(単純最短距離は 171)。

5. 複数棟における経路生成

5.1.1. 概要 (三棟における学習)

先述の四面壁面と同様のものを3棟作成し、ランダムに配置した。配置した棟(A棟,B棟,C棟)の位置と、ドローンの行動可能範囲は図9に示した。1回の学習では全面を検査できなかったため、カリキュラム学習を行った。ここでは、学習を4つのパートに分けた。Lesson1では三面壁面までの学習を行い、Lesson2では四面目の壁面の学習、Lesson3では二棟目までの学習、Lesson4で三棟目までの学習を行った。それぞれは、0~20M steps, 20~30M steps, 30~70M steps, 70~300M stepsの区間で学習を行った。

5.1.2. 学習結果

学習による平均累積報酬とエピソードの長さの推移を図11に示した。学習後のドローンの検査経路の1つを図10に示した。三棟すべての壁面が残らず検査された。A棟、

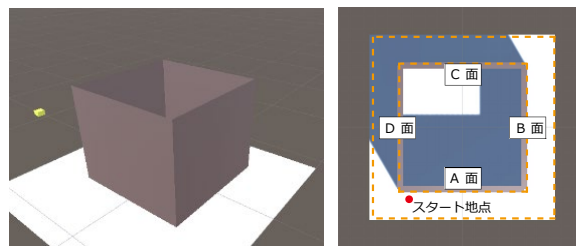


図6(四面壁面の環境,ドローンの行動範囲)

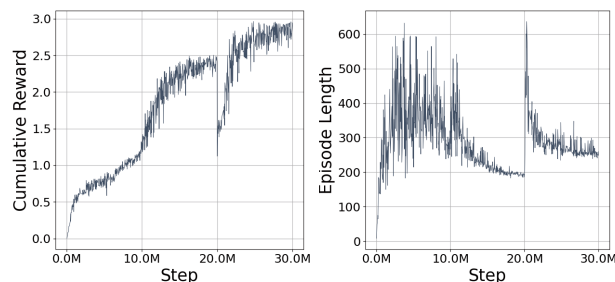


図7(平均累積報酬とエピソードの長さの推移)

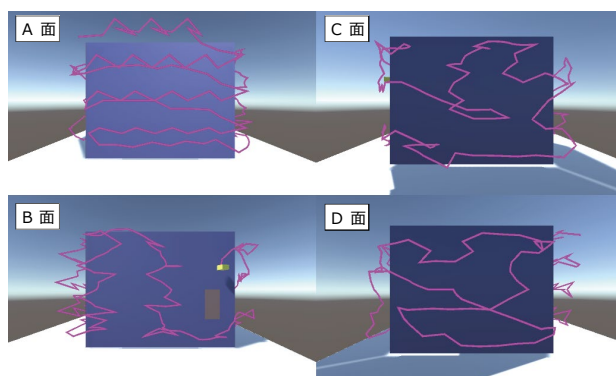


図8(学習後のドローン検査経路)

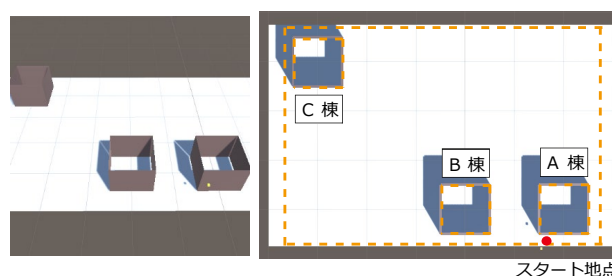


図9(三棟の環境,ドローンの行動範囲)

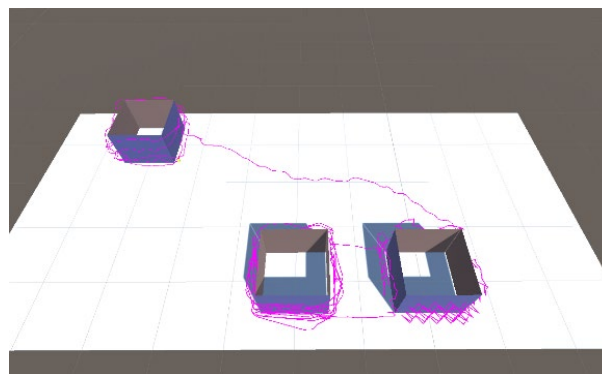


図10(学習後のドローン検査経路)

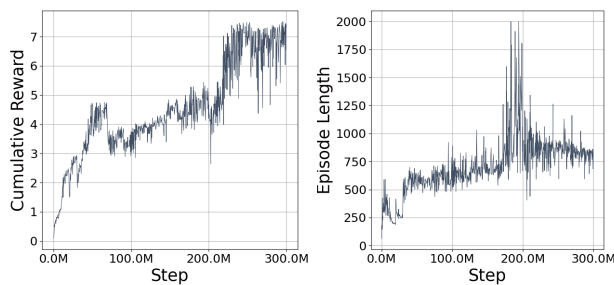


図 11(平均累積報酬とエピソードの長さの推移)

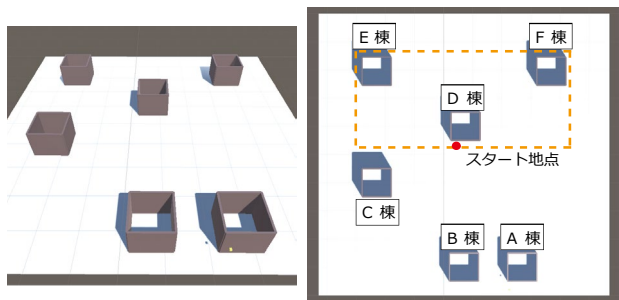


図 12(三棟の環境, ドローンの行動範囲)

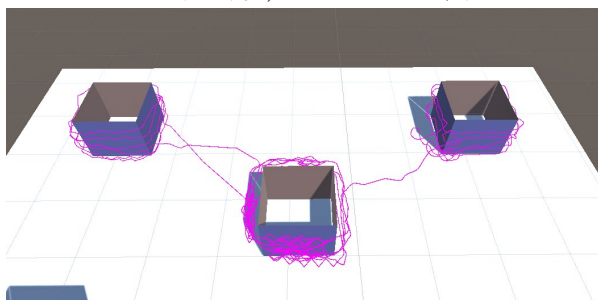


図 13(学習後のドローン検査経路)

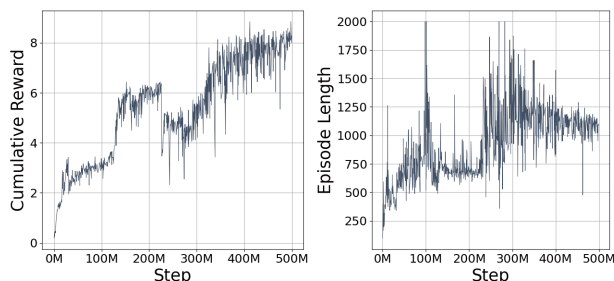


図 14(平均累積報酬とエピソードの長さの推移)

B 棟、A 棟、C 棟の順に検査をしており、A 棟に再び戻っての検査が見られた。また、棟を回周しながら検査をしていた。最善の経路とは言えず課題となった。

5.2.1. 概要

図 12 のように 3 棟(D, E, F 棟)を追加でランダムに配置した。4 棟目の学習に時間を要したため、5.1. のフィールド上に 3 棟を追加し、3 棟(D, E, F 棟)のみで学習をした。5.1. と同様に 3 つのパートに分けてカリキュラム学習させた。それぞれは、0~20M steps, 20~30M steps, 30~225M steps, 225~500M steps の区間で学習を行った。

5.2.2. 学習結果

学習による平均累積報酬とエピソードの長さの推移を図 14 に示した。学習後のドローンの検査経路の 1 つを図

13 に示した。D 棟、E 棟間を往復する経路で検査していた。

6. まとめ

強化学習によるドローン壁面点検経路の手法について検証した。三面、四面、三棟について、強化学習によりすべての壁面を検査できることを示した。しかし、学習後の検査経路距離が最適でないことや、次の棟の探索に対する学習時間が多くかかることに課題が残った。今後の展望として、より複雑な点検対象に対する強化学習や、効率の良いような学習環境の設計、学習時間を短くする方法の検討を進めたい。

[参考文献]

- 1) 内閣府, “特集 首都直下地震の被害想定と対策について (最終報告)”, https://www.bousai.go.jp/kohou/kouhoubousai/h25/74/special_01.html, (参照 2023-06-10)
- 2) 熊本県, “熊本地震での経験を踏まえた橋梁災害復旧のまとめ (災害査定編) (災害復旧工事編)”, <https://www.pref.kumamoto.jp/soshiki/101/939.html>, (参照 2023-06-10)
- 3) 中村 秀明, 山本 拓海, 青島 亘佐: コンクリート道路橋の点検省力化に向けた変状領域の自動抽出, インフラメンテナンス実践文集, 2022, 1 巻, 1 号, p. 386-393
- 4) 日経クロステック, “全自動ドローンが“現場監督”に、日々の出来高測量や安全巡視を無人化”, 2021.08.03 https://xtech.nikkei.com/atcl/nxt/column/18/00107/00171/?i_cid=nbnpnxt_reco_atype, (参照 2022-09-28)
- 5) 宮内 博之: 建築分野におけるドローンに係る基盤開発と建築物点検・調査へのドローン活用, 次世代移動体技術誌, 2022, 3 巻, 3 号, p. 27-42
- 6) 鈴木 学, 浜 克己: 中村 尚彦, 協調ドローンを用いた避難誘導支援システム, 計測自動制御学会論文集, 2020, 56 巻, 1 号, p. 24-30
- 7) 樽谷 葵, 石田 航星: 深層強化学習を用いた揚重計画の自動作成手法に関する研究—2 点間の揚重経路作成における衝突回避の実装—, 日本建築学会関東支部研究報告集, 91 号, p.417-420, 2021-03
- 8) Guillem Muñoz, Cristina Barrado, Ender Çetin, Esther Salami: Deep Reinforcement Learning for Drone Delivery, Drones 2019, 3(3), 72
- 9) 柴田拓海, 山田悟史: 強化学習を用いたドローン点検経路の自動生成, 第 45 回情報・システム・利用・技術シンポジウム: 報告 pp.68-71, 2022.12, 日本建築学会
- 10) 柴田拓海, 山田悟史: カリキュラム強化学習を用いたドローン点検に関する研究 —コ字型壁面に対する経路生成—, 日本建築学会大会 (関西) 学術講演会梗概集 (建築社会システム), pp.143-144, 2023.9, 日本建築学会
- 11) Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore.: Reinforcement learning: A survey. Journal of artificial intelligence research, 4:237-285, 1996.
- 12) Arthur Juliani, Vincent-Pierre Berges, Ervin Teng, Andrew Cohen, Jonathan Harper, Chris Elion, Chris Goy, Yuan Gao, Hunter Henry, Marwan Mattar, Danny Lange.: Unity: A General Platform for Intelligent Agents, arXiv:1809.02627
- 13) John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.