

シールドトンネル工事計画における深層強化学習を用いた シールドマシン制御

土橋俊紀[†] 加藤昇平^{†‡} 伊原滉也^{†‡} 増田湖一[§] 新宮康之[§]

[†]名古屋工業大学 大学院工学研究科情報工学専攻

[‡]名古屋工業大学 情報科学フロンティア研究院

[§]清水建設株式会社

1. はじめに

近年、シールドトンネル工事における生産性向上が求められており、AIによる施工計画の効率化に関する研究が盛んに行われている[1]。本研究では、都市部のトンネル工法の一つであるシールド工法に着目し、シールドマシン[2]制御の知能化を目指す。本稿では、掘進シミュレーション実験を行い、Deep Deterministic Policy Gradient (DDPG) [3]を用いて、シールドマシン制御のうちシールドジャッキ及び中折れ制御を学習する。シールドマシンの無人制御の実行可能性とそれに適した深層強化学習手法がシールドマシン制御に適用可能性を検討する。

2. シールドマシン

シールドマシンを図1に示す。シールドマシンとは、シールド工法で用いられる円筒状の掘削機である。事前に国土交通省から与えられた計画線に沿って様々な制約下で掘削を行う。シールドマシンは前胴部と後胴部、それらをつなぐ中折れ部を有し、それぞれが掘削に必要な要素を持つ。前胴部ではマシンの先端に装備されているシールドカッターによって掘削を行う。中折れ部の主な装備は中折れジャッキである。中折れジャッキはシールドマシンの前胴と後胴をつなぐもので、シールドマシンを曲線に沿って折り曲げることができる。後胴部の主な装備は、シールドジャッキである。シールド工法では、掘った部分が崩れてこないように後胴内部でセグメント（トンネルの外壁となるブロック）を組み立てていく。この組み立てたセグメントに後胴部の円周上に配置された独立して可動するシールドジャッキを押し、ジャッキを伸ばすことで反作用力を発生させ推進する。伸ばすジャッキ長に偏りをもたせることで曲進することもできる。



図1 シールドマシン

2.1 マシン制御

セグメントごとのシールドマシンの制御量は以下のように計画される。

- ・ J_L : 左ジャッキ屈伸長
- ・ J_R : 右ジャッキ屈伸長
- ・ J_U : 上ジャッキ屈伸長
- ・ J_D : 下ジャッキ屈伸長
- ・ $\Delta\theta_{LR}$: 掘削方向に水平方向の中折角
- ・ $\Delta\theta_{UD}$: 掘削方向に鉛直方向の中折角

例として左ジャッキ屈伸長を右ジャッキ屈伸長より長く伸ばすことによってマシンは右に向く動きをする。これらの制御を行った結果、マシンとセグメントが接触してしまう場合があり、掘削不能に陥るため、シールドマシンとセグメント間にクリアランスを持たせる必要がある。本研究では、クリアランスを一定上保ったまま掘削することを目的に強化学習を行う。

3. DDPG

DDPGは強化学習における Actor-critic という手法と Deep Q Network (DQN) [4]の構成要素である replay buffer を用いた方策勾配型のアルゴリズムである。actor に方策勾配を、critic に DQN を実装し、actor が決定した行動の善し悪しを critic が評価する。replay buffer に過去の状況・行動・報酬を保存しておき、それを用いて学習を行う。行動価値を学習する強化学習と違い行動が連続値でも学習できる利点がある。

Shield Machine Control Using Deep Reinforcement Learning in a Shield Tunneling Project

Toshiki TSUCHIBASHI[†], Shohei KATO^{†‡}, Ihara KOUYA[†]

[‡], Hiroichi MASUDA[§] and Yasuyuki SHINGU[§]

[†]Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

[‡] Frontier Research Institute for Information Science, Nagoya Institute of Technology

[§] Shimizu Corporation

^{†‡}Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

4. 深層強化学習によるマシンの制御

シールド工事で使用された計画線とセグメント，シールドマシンの実データを用いた計画線のうち1リング分のマシン制御のシミュレーションを行う．セグメント $\mathbf{R}=[R_1...R_T]$ は計画線に沿うように事前に割り付けられており，その割り付け方に応じたマシンの制御を学習する．ここでは R_i は i 番目のセグメントを表す．セグメント1つにつき1ステップ数とする．状態表現はシミュレータ上の観測情報を用いる．マシンの制御を以下のように定義する．

- ・ J_{LR} : 水平方向のジャッキストローク差
- ・ J_{UD} : 鉛直方向のジャッキストローク差
- ・ $\Delta\theta_{LR}$: 掘削方向に水平方向の中折角
- ・ $\Delta\theta_{UD}$: 掘削方向に鉛直方向の中折角

ジャッキは個別に屈伸させることができるが，制御の次元数を減少させる目的でジャッキストローク差としている．それぞれのジャッキの屈伸長さは以下のように決定できる．1ステップにつきセグメント1つ分の距離は掘削しなければならないのでセグメント長を1とすると以下のように定義できる．

$$J_L = 1 - \frac{J_{LR}}{2} \quad (1)$$

$$J_R = 1 + \frac{J_{LR}}{2} \quad (2)$$

$$J_U = 1 - \frac{J_{UD}}{2} \quad (3)$$

$$J_D = 1 + \frac{J_{UD}}{2} \quad (4)$$

報酬としてマシンとセグメントとのクリアランスに基づいて設計した．最も狭いクリアランス距離を基に設計する．

5. 実験設定

実験では，状態として28次元の情報を用いる．表1に t ステップ目の状態表現を示す． θ_{R_t} はシールドジャッキを伸ばす際に押し付けているセグメント R_n の方位とピッチングを表す． $\theta_{R_{t+1}}$ は新たに割り付けるセグメント R_{n+1} の方位とピッチングを表す． θ_m はマシンの前胴の先端 (FF)，後端 (FR) と後胴 (RF) の先端，後端 (RR) の方位とピッチング， \mathbf{d}_m はマシンの4部位と計画線からの垂直方向と水平方向の偏差量を表す． \mathbf{C}_m はマシンのRFとRRの上下左右とセグメントのクリアランス距離を表す．行動として，実マシンの制御量を基に表2のように設計した．報酬として最も小さいクリアランス距離を C_{min} とした時に以下のように与える．

$$r = \frac{2}{1 + e^{-C_{min} - 0.5}} - 1 \quad (5)$$

上記の状態表現，行動，報酬などを基に図2に示すDDPGの実験を行う．

表1 シミュレータで観測できる状態

情報	次元数
セグメント R_n の角度 θ_{R_n}	2
セグメント R_{n+1} の角度 $\theta_{R_{n+1}}$	2
マシンの各部位の角度 $\theta_{mFF}, \theta_{mFR}, \theta_{mRF}, \theta_{mRR}$	8
マシンの各部位の計画線からの偏差量 $\mathbf{d}_{mFF}, \mathbf{d}_{mFR}, \mathbf{d}_{mRF}, \mathbf{d}_{mRR}$	8
セグメントとのクリアランス \mathbf{C}_m	8

表2 マシンの行動範囲

$-75 [\text{mm}] \leq J_{LR} \leq 75 [\text{mm}]$
$-75 [\text{mm}] \leq J_{UD} \leq 75 [\text{mm}]$
$-0.5 [^\circ] \leq \Delta\theta_{LR} \leq 0.5 [^\circ]$
$-0.5 [^\circ] \leq \Delta\theta_{UD} \leq 0.5 [^\circ]$

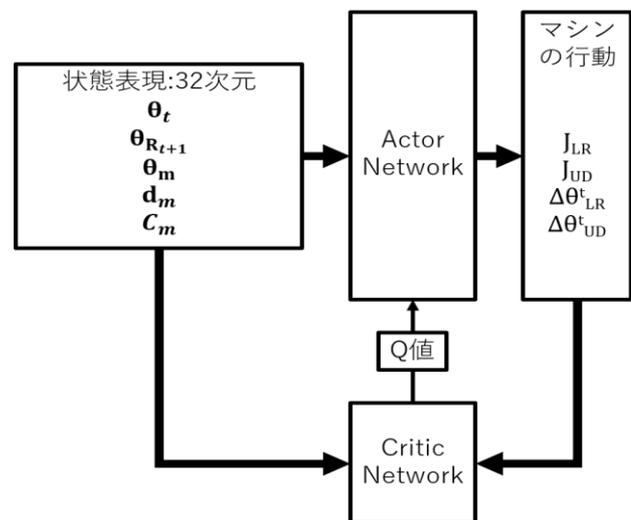


図2 DDPGによるマシン制御

6. まとめと今後の展望

DDPGによるシールドマシン制御を提案した．現在，DDPGに用いられるネットワーク構成の検討やハイパーパラメータなどのチューニングを行っており，有効性の検証を行う．今後の展望は，よりコストを削減するために掘削土量を考慮した報酬値を設計し，人手で行われたマシン制御と比較を行う．

参考文献

- [1] Koya Ihara, Shohei Kato, Takehiko Nakaya, Tomoaki Ogi, Hiroichi Masuda. (2019). Application of PSO-Based Constrained Combinatorial Optimization to Segment Assignment in Shield Tunneling. Agents and Artificial Intelligence, pp. 166–182, Springer.
- [2] 土木学会 2016 年制定トンネル標準示方書 [シールド工法編]・同解説, JSCE Publications
- [3] Timothy P, et al. (2016). “Continuous control with deep reinforcement learning”, ICLR2016.
- [4] Volodymyr Mnih, et al. (2015). Human-level control through deep reinforcement learning. Nature, Vol. 518, No. 7540, pp. 529–533.