

# 深層強化学習による五軸加工機の迂回動作の自動生成に関する研究

## A Study on Automatic Generation of Detour Operation of 5-axis Machine by Deep Reinforcement Learning

○藤岡 凌司\*<sup>1</sup>, 古庄 玄樹\*<sup>2</sup>, 加戸 啓太\*<sup>3</sup>, 平沢 岳人\*<sup>4</sup>  
Ryouji Fujioka\*<sup>1</sup>, Genki Furusho\*<sup>2</sup>, Keita Kado\*<sup>3</sup> and Hirasawa Gakuhito\*<sup>4</sup>

\*1 千葉大学大学院 融合理工学府 博士前期課程

Graduate Student, Graduate School of Science and Engineering, Chiba University

\*2 千葉大学大学院 融合理工学府 博士後期課程 修士 (工学) 日本学術振興会 特別研究員 DC

Graduate Student, Graduate School of Science and Engineering, Chiba University M.Eng., JSPS Research Fellow

\*3 千葉大学大学院 工学研究院 助教 博士 (工学)

Assistant Professor, Graduate School of Engineering, Chiba University Ph.D.

\*4 千葉大学大学院 工学研究院 教授 博士 (工学)

Professor, Graduate School of Engineering, Chiba University Ph.D.

キーワード : 深層強化学習; Unity ML-Agents; 五軸加工機; 迂回動作

Keywords: Deep reinforcement learning; Unity ML-Agents; 5-Axis Machine; Detour operation

### 1. はじめに

現在日本の建設業は少子化による人手不足問題を抱えており、機械加工による部材の大量生産によって現場での作業を省力化する取り組みが進められている。建築分野で用いられる自動加工機は規格品の効率的生産を主目的とするものが多いが、伝統木造建築の部品やコンピュータショナルデザインによって設計される部材は規格化できない形状が多く、加工機には多品種少量生産を可能にする性能が求められる。

このような問題を背景として、筆者らの所属する研究室では以前より人工技能と称される機械加工の研究に取り組んでいる。現在は五軸加工機 (図 1) と呼ばれる木材加工機の開発を行っており<sup>1) 2) 3)</sup>、昨年度までの研究でツールの自動交換による多様な形状の切削が可能となった<sup>4)</sup>。

五軸加工機では一つの加工動作ごとにツールの向きや角度を設定し、ツール先端部を指定した点まで移動させながら対象の加工を行っていく。このような加工を行うための運行軌跡プログラムを単位パスと呼ぶ。

また、形状が複雑な部材に対しては、ツール交換をしながら単位パスを連続的に組み合わせて加工を行うため、与えられた各単位パスの終点から次の単位パスの始点まで、部材に接触することなくツールを移動させる迂回動作が必要となる。この迂回動作を実行する運行軌跡プログラムを迂回パスと呼ぶ。

大量生産を行う機械加工では同一の加工動作を繰り返すため迂回パスは常に一定になるが、多品種少量生産を行う機械加工では各部材ごとに適切な迂回パスを設定し

なければならない。現在の五軸加工機では安全性を考慮し、十分にワークを避ける迂回パスを設定した上で、迂回動作に問題ないかを確認して作業を行う必要があり、パスの調整に加工機への慣れといった属人的な技能も必要であることから、確認に大きな労力を要している。また、動作範囲の一部を迂回用のスペースとして使用しているため、加工機本来の動作範囲を最大限に活用できず、加工機のサイズを活かした大型の材料加工が難しい状態となっている。

このような問題を解決するため、迂回スペースを必要としない迂回パスを自動生成する手法について検討する。迂回動作には加工ツールがワークに干渉しない安全性と、最短経路で移動を行う効率性が求められる。双方を成立させる最適な迂回パスを導出するには強化学習を用いた最適経路探索が有用であるが、状態空間と行動空間が共に連続値となる学習環境となるため Q-learning など古典的手法での導出は非常に困難である。

そこで本研究では、このような探索に向く深層強化学習によって合理的な迂回動作を生成する手法の開発を試み、検証を行う。

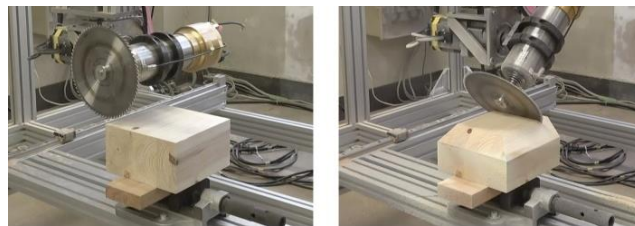


図 1. 五軸加工機での加工の様子

## 2. 研究概要

本研究では学習環境として五軸加工機のシミュレーターを制作し、深層強化学習によって安全かつ高効率な迂回行動を学習させる。

### 2.1. 学習環境

本研究の学習環境を図 2 に示す。五軸加工機シミュレーターの構築にはゲームエンジンの Unity を使用した。赤色の Agent パーツがツールの現在位置、青色の Target パーツがツールを移動させる目標地点、黄色の Work パーツが加工対象の位置を表している。

Agent は可動域内で自由に移動を行い、Target に到達した時のみ報酬を得ることによって、Work やフレームを避けながら Target の位置へ移動する動きを学習する。

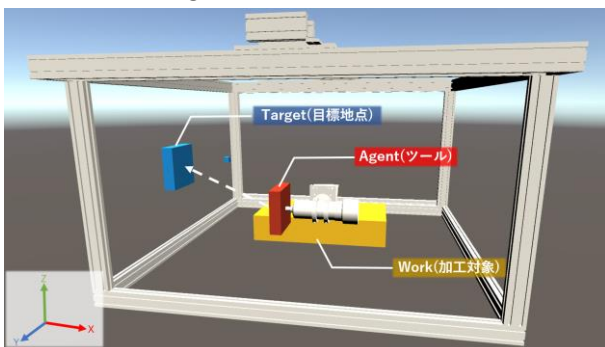


図 2. 五軸加工機シミュレーター

### 2.2. 学習フレームワーク

深層強化学習を行うフレームワークには Unity の ML-Agents を使用した。ML-Agents では Unity と Python 連携し、学習環境から得られる状態や報酬の観察値を Python スクリプトによって処理することで Agent の学習を行う。状態の観察には 3 つのデータ型(表 1)、Agent の行動には連続値と離散値の二つのデータ型 (表 2) を用いることができる。

表 1. 観察のデータ型

観察名	内容
Vector Observation	学習に必要な情報を格納した浮動小数点配列
Visual Observation	カメラやレンダリングテクスチャの画像
Raycast Observation	Agent から周囲に飛ばした Ray の衝突判定と衝突オブジェクトに対する距離(図 3)

表 2. 行動のデータ型

行動名	内容
Continuous	-1.0~1.0 の連続値
Discrete	設定範囲内の整数離散値

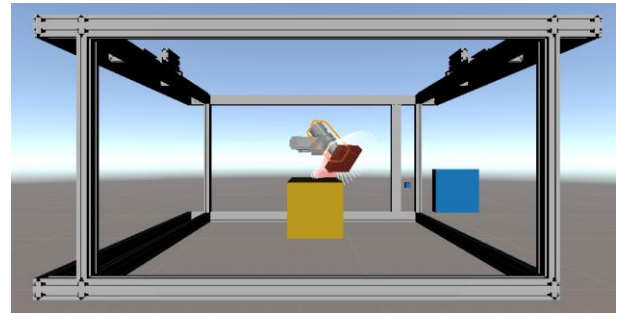


図 3. Raycast Observation の様子

## 3. 三軸加工環境での予備実験

五軸加工環境での学習を行う予備実験として、三軸加工環境での学習を行い、適切な迂回行動が獲得可能であるかを検証した。三軸加工ではツールの回転を含まないため、学習難易度が比較的低い環境である。観察データ型には Vector Observation、Visual Observation、Raycast Observation の 3 つを使用し、各観察状態における平均報酬と所要時間から学習結果の比較を行った。

### 3.1. 学習設定

エピソード開始時に Agent と Target それぞれの初期位置と、Work のサイズをランダムに決定する。Agent は学習から得られるポリシーに従って行動を選択し、毎 Step 移動を実行する。行動によって Agent が Target の位置に到達した場合には+1.0 の報酬を与えてエピソードを終了し、移動の途中で Work や加工機フレームに接触した場合には報酬を与えずエピソードを終了する。学習は 1000000 Step まで行い、平均報酬が不安定な場合や学習が十分に進んでいない場合にはその後も学習を継続させる。

各実験にて使用した観察設定の詳細を表 3、共通して使用した行動設定、報酬設定の詳細を表 4 に示す。

表 3. 各実験での観察設定

実験番号	観察内容
1	Vector Observation(サイズ 13) ・ 0,1,2 : Agent の三次元座標 ・ 3,4,5 : Target の三次元座標 ・ 6,7,8 : Work の三次元座標 ・ 9,10,11 : Work のサイズ ・ 12 : Agent と Target の距離
2	Visual Observation × 3 ・ 加工機上部からのカラー画像 ・ 加工機側面(X 軸方向)からのカラー画像 ・ 加工機側面(Y 軸方向)からのカラー画像
3	Raycast Observation ・ ツール周囲 360 度の Ray

表 4. 三軸加工環境での学習設定

項目	内容
行動	Continuous(サイズ 3) ・ 1: X 軸方向の移動 ・ 2: Y 軸方向の移動 ・ 3: Z 軸方向の移動
報酬	・ Agent が Target の位置に到達 →Reward = 1.0(エピソード終了) ・ Agent が Work やフレームに衝突 →Reward = 0(エピソード終了)

### 3.2. 学習結果・考察

学習中の様子を図 4、各実験における獲得報酬の推移を図 5、最終的な平均報酬と学習所用時間を表 5 に示す。

実験 1 では、Work を避けながら Target へと移動する基本的な方策を 200000 Step ほどで獲得し、安定して迂回動作を学習していることが確認できた。平均報酬は 0.8 前後で推移しており、学習の中盤では 0.85 付近で安定している部分も見られた。

実験 2 では、3 面の画像を元に学習を行っているため処理に時間がかかり、学習結果が安定しなかった。1000000 Step 経過後の段階で十分な学習ができていなかったため

2000000 Step まで学習を継続したが、最終的な平均報酬は 0.5 を下回っており、Visual Observation のみでの学習は困難であるということが確認できた。

実験 3 では、Ray の発射方向を球状に設定できなかったため、三次元空間に対して十分な認識を行うことができず、学習が成立しなかった。

上記三つの実験結果を踏まえ、追加実験として Vector Observation と Visual Observation を組み合わせた観察状態で学習を行ったが、平均報酬は Vector Observation 単体の時とほとんど変わらず、所要時間だけが增加する結果となった。

これらの結果から、本学習環境では Vector Observation を使った観察が最も適しているということが考えられる。

表 5. 三軸加工環境での学習結果

実験番号	観察内容	平均報酬	所用時間
1	Vector Observation	0.7813	01:46:01
2	Visual Observation	0.4807	17:58:15
3	Raycast Observation	学習不可	学習不可
追加実験	Vector Observation + Visual Observation	0.7773	15:33:04

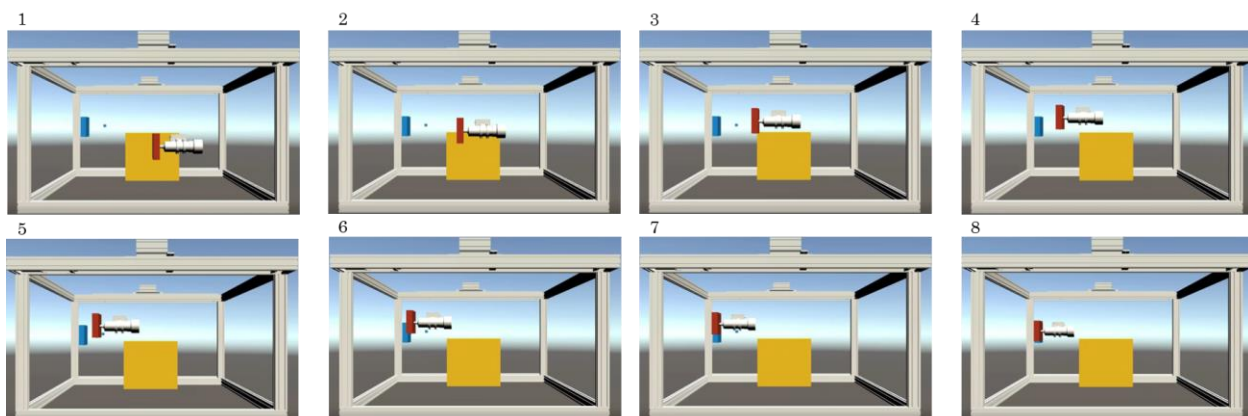


図 4. 三軸加工環境での迂回行動の様子

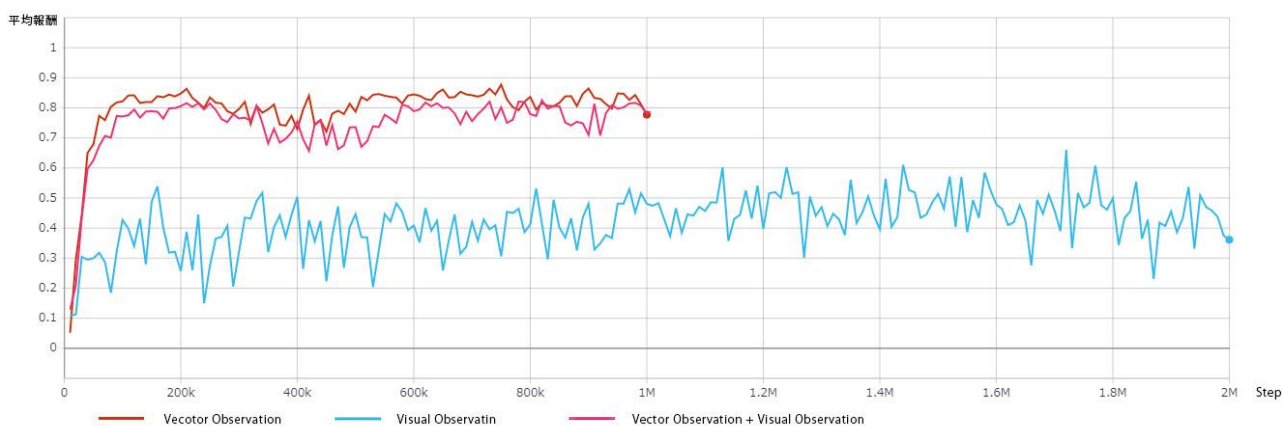


図 5. 三軸加工環境での平均報酬推移

#### 4.五軸加工環境での実験

予備実験の結果から、Vector Observation 単体での観察データを使用して五軸加工環境での学習を行った。

##### 4.1. 学習設定

三軸加工環境と同様の学習設定に、AB軸の回転制御プログラムとAgentの回転角度観察を加えて1000000Stepまで学習を行った。

五軸加工環境での学習設定を表6に示す。

表6. 五軸加工環境での学習設定

項目	内容
観察	Vector Observation(サイズ15) ・0,1,2: Agentの三次元座標 ・3,4,5: Targetの三次元座標 ・6,7,8: Workの三次元座標 ・9,10,11: Workのサイズ ・12: AgentとTargetの距離 ・13,14: AgentのA軸、B軸回転角度
行動	Continuous(サイズ3) ・1: X軸方向の移動 ・2: Y軸方向の移動 ・3: Z軸方向の移動
報酬	・AgentがTargetの位置に到達 →Reward = 1.0 (エピソード終了) ・AgentがWorkやフレームに衝突 →Reward = 0 (エピソード終了)

##### 4.2. 学習結果・考察

学習中の様子を図6、最終的な平均報酬と学習所用時間を表7に示す。

三軸加工環境と同様の基本方策を1000000Stepほどで獲得し、途中不安定な部分がありつつも5000000Step頃から比較的安定して学習をしていることが確認できた。

学習中には、ツール角度に合わせて細かい位置調整をする行動や、Workに当たらないギリギリの位置を探索する

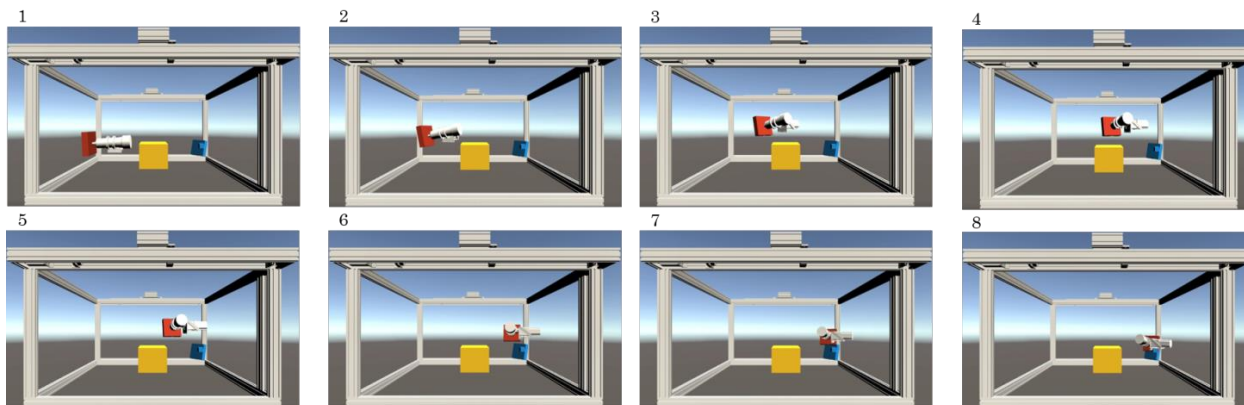


図6. 五軸加工環境での迂回行動の様子

行動など、適切な迂回行動を求めするためにAgentが試行錯誤している様子が数多く見られた。

しかし、Workへの接触を完全になくすことはできていないため、今後学習環境の改善や学習設定の調整によって学習の精度を高める必要がある。

表7. 五軸加工環境での学習結果

観察内容	平均報酬	所用時間
Vector Observation	0.76	04:14:27

#### 5. まとめと展望

五軸加工機による大型材料加工に向けて、ワーク加工時における適切な迂回動作を深層強化学習によって自動生成する手法について検討を行った。

実験では適切な迂回行動を学習し、AgentがWorkを避けながらTargetへと移動する様子が確認できたが、学習精度には改善の余地が見受けられた。

今後は学習設定の調整によって精度を高めると共に、実機での検証に向けたシステムの構築や、他の機械システムへの展開を行っていく予定である。

##### 【参考文献】

- 1) 中村優介、高橋雅生、戸田勇登、林真那、高林弘樹、加戸啓太、平沢岳人、五軸加工機による校倉構法の断面加工に関する研究 多軸加工によるあたらしい建築構法の創出研究その1、日本建築学会技術報告集、第24巻 第56号、pp.447-450、2018.2
- 2) 中村優介、漆山生羽、林真那、古庄玄樹、加戸啓太、平沢岳人、五軸加工機による丸ノコを用いた加工パス導出に関する研究、2017年度日本建築学会関東支部研究報告集、第88巻、pp.475-478、2018.3
- 3) 古庄玄樹、吉岡直希、大谷星輝、中村優介、加戸啓太、平沢岳人、ワークと機械情報を考慮した多品種少量生産に関する研究、2018年度日本建築学会関東支部研究報告集、2019.3
- 4) 古庄玄樹、吉岡直希、RANDI HANTORO、大谷星輝、中村優介、加戸啓太、平沢岳人、ATC付き五軸加工機のフォーリー制作を通じた可用性検討、第42回情報・システム・利用・技術シンポジウム論文集、pp.346-349、2019.12