深層強化学習によるロボット移動の視覚的説明と拡張現実による提示に関する研究

TR21003 尹 文韜　　指導教授：藤吉 弘亘

1. はじめに

深層強化学習はロボットの高い自律移動性能を獲得できる．一方で，学習したエージェントモデルは膨大な数のパラメータにより構成されており，モデルの内部処理を直接解析することができない．そのため，ユーザがエージェントであるロボットの行動選択に対する判断根拠を理解することは非常に困難である．そこで本研究では，自律移動ロボットと人の共存促進を目的とし，ロボットの行動選択に対する判断根拠の可視化，および拡張現実（AR）技術によるユーザに対するロボット動作の判断根拠を提示する手法を提案する．ロボットの高い自律移動性能を獲得するために，深層強化学習に Transformer [1] を導入する．Transformer の Attention をもとに深層強化学習モデルの注視領域を獲得する．そして，AR 技術を活用することで注視領域をユーザに提示する．これにより，ユーザがロボット動作の判断根拠をより容易に理解できるようになる．

2. 深層強化学習の視覚的説明

深層強化学習は様々なタスクで高い性能を獲得している反面，エージェントの意思決定に対する判断根拠が不明確である．そのため，この問題の解決を目的とした説明可能な強化学習 (XRL) 手法が研究されている．板谷らは深層強化学習に Transformer を導入した Action Q-Transformer (AQT) を提案している [2]．Transformer は入力トークンと他トークン間の類似度を算出して，入力に対する重要な情報を Attention として獲得している．この Attention を可視化することで，注視領域を把握することができる．AQT では Transformer Decoder に入力する Query を行動情報とすることで，各行動に対する注視領域を獲得できる．

3. 提案手法

本研究は深層強化学習手法 Deep Q-Network (DQN) [3] に Transformer を導入することで，ロボットの自律移動と同時にロボットの注視領域を示す Attention を獲得する．加えて，この Attention を AR により実空間に投影することでロボット動作の判断根拠を効率的にユーザへ提示する方法を提案する．

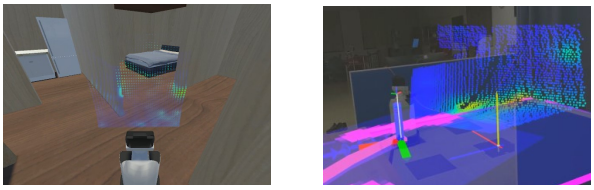
3.1 深層強化学習による自律移動能力の獲得

提案手法のネットワーク構造を図 1 に示す．ロボット動作に対する解釈性の高いネットワーク構造とするため，AQT をベースとした Transformer Encoder-Decoder 構造を採用する．以下で提案手法による行動算出までの流れを述べる．

はじめに，ロボット視点の画像上の物体を学習済みセマンティックセグメンテーションモデルにより壁，床，家具，人の 4 クラスに分類し，それぞれ緑，青，赤，灰色で表示する．これは，シミュレーション環境と実環境とのテクスチャの相違が生じるドメインギャップを吸収するためである．そして，この画像を 84×84 ピクセルにリサイズ後，12×12 の個のパッチ画像に分割する．つまり，パッチ画像のサイズは 7×7 ピクセルである．各パッチ画像は全結合層により埋め込みベクトルへ変換後，Positional Encoding により位置情報を付与し Encoder へ入力する．また，Decoder は Encoder の出力を Value と Key とし，Action query を Query とする．Action query は Query branch で算出した各動作とゴール情報を含む特徴ベクトルである．最後に，Decoder の各 Query に対する出力を Q Heads へ入力し，各動作の Q 値を算出する．ここで，エージェントであるロボットの動作は停止，前進，左右旋回の 4 種類である．

3.2 AR を用いたユーザへの提示

Transformer Decoder では行動情報を含む Action query を用いているため，ロボットの動作に関連した Attention を獲得している．そこで，出力した Q 値が最も高い動作に対応する Decoder の Attention を AR により可視化する．ここでの可視化方法は， Depth センサから獲得したロボット視点での深度情報を点群に変換し，その点群を Attention の値に応じて色付けする．ここで Attention はヒートマップで表現し，値が大きいほど赤く，小さいほど青く色付けする．シミュレーション環境と実環境における AR による Attention の可視化例を図 2 より示す．



(a) シミュレーション環境　　(b) 実環境

図 2: AR による Attention 可視化例

4. 評価実験

提案手法の有効性を確認するため，深層強化学習により獲得した自律移動性能の評価，および AR を用いたユーザへの提示に関する有効性調査を行った．

4.1 シミュレーション環境

本研究では，HSR ロボットによる屋内環境における自律移動タスクを対象とする．実環境でロボットを自律移動させて学習させるコストは高い．そのため，Unity 上で屋内環境を再現し，人や机等の障害物を設置した．また，ドメインギャップを対処するため，シミュレータ内のテクスチャをオブジェクトごとに色分けした．本環境では，エピ

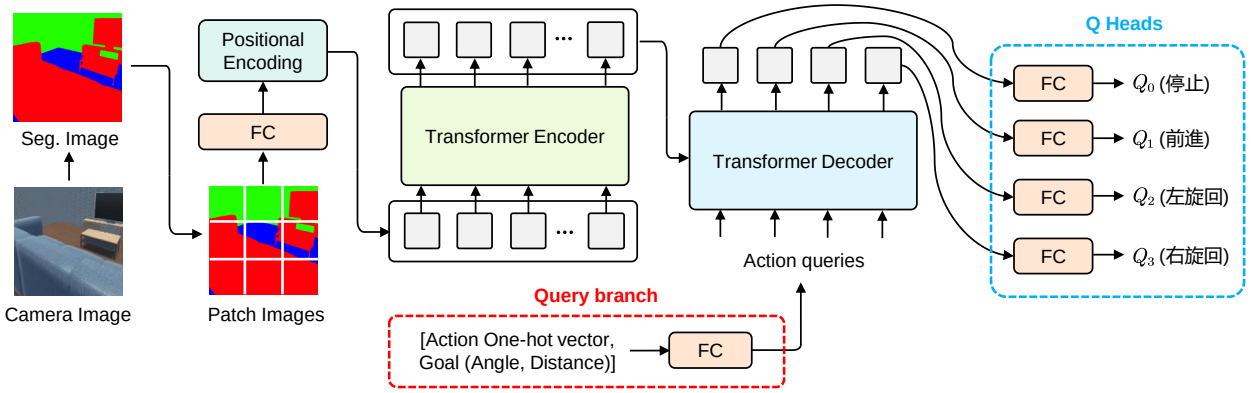


図 1: 提案手法のネットワーク構造

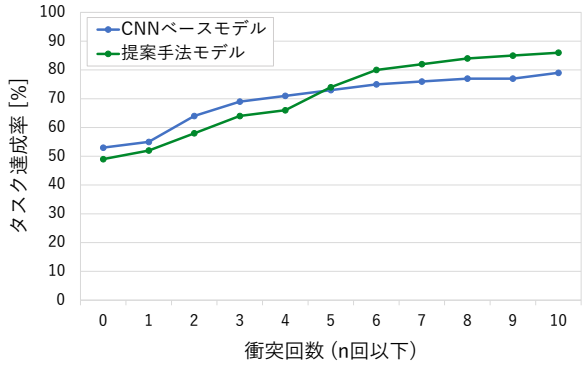
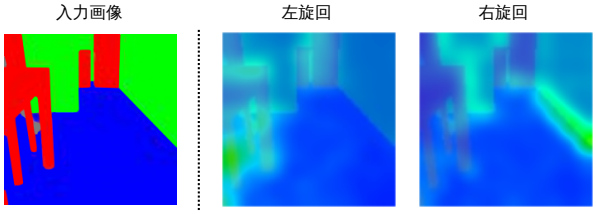
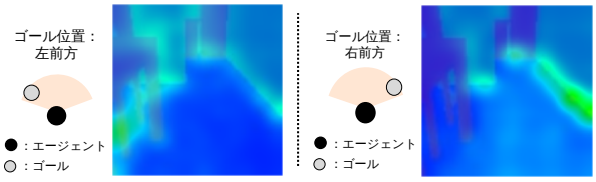


図 3: 100 エピソード間のタスク達成率



(a) 行動の変化による可視化: ゴール位置が左前方である時の “左旋回” と “右旋回” に対する Attention.



(b) ゴール位置の変化による可視化: ゴール位置のみ変更した時の “停止” に対する Attention.

図 4: Decoder の Attention 可視化例

ソード毎でランダムにスタート位置と最終ゴール位置を設定し，スタートと最終ゴール間にはサブゴールを生成する．

4.2 自律移動性能の評価

タスク達成率　提案手法の性能を評価するため，100 エピソード間のタスク達成率を比較する．ここでのタスク達成条件は，ロボットが 100 ステップ以内に最終ゴールに到達，かつ衝突が一定回数以下とした．比較手法は，従来の CNN ベースモデルと提案手法の Transformer ベースモデルである．各モデルは同じ学習条件により 1.0×10^6 ステップの学習を行った．図 3 から，提案手法の Transformer ベースモデルは CNN ベースモデルと比較し，エピソードごとの衝突回数がやや増加しているが，最終ゴールへの到達率が向上していることが分かる．このことから提案手法は高い自律移動性能を獲得できていると言える．

Attention の可視化　獲得した Attention が正しくロボットの注視領域を示しているか確認するため，Decoder の行動やゴール位置に対する注視領域の変化を確認した．図 4 (a) から，ロボットは，左旋回動作では左の家具を，右旋回動作では右の壁を注視していることが分かる．また図 4 (b) から，ゴール位置を左前方から右前方へ変化することで，ロボットの注視対象も左の家具から右の壁に変化している．これらのことから，Decoder の Attention は動作毎とゴール位置に対するロボットの注視領域を正しく示していることが分かる．

4.3 AR を用いた人への提示に対する有効性調査

本実験では 2 つの方法により，AR を用いた Attention 可視化がロボット動作に対するユーザの理解に有効かを調査した．

表 1: ユーザによるロボット動作の予測

被験者グループ	提示なし	提示あり
平均正答率 [%]	31.8	38.8

表 2: ロボット動作に関する理解度アンケートの集計結果

評価段階	定点肉眼観測 [%]	AR を介した観測 [%]
理解できる	14.1	**30.6**
ある程度理解できる	41.3	**45.8**
あまり理解できない	**22.8**	18.1
理解できない	**21.7**	5.6

ユーザのロボット動作予測による調査　本調査では，図 2 (a) に示すシミュレーション環境を使用し，33 名の被験者を対象とした．被験者を 2 グループに分け，それぞれ Attention 提示なしとありの状態でロボットが自律移動するデモ動画を視聴し，練習問題に回答することで，ロボットの動作について学んでもらう．その後，すべての被験者に Attention なしの状態でロボットの動作予測問題を 10 問回答してもらった．動作予測問題に対する全被験者の平均正答率を表 1 に示す．表 1 から，AR を用いた提示ありのグループが提示なしグループよりも正答率が 7pt 向上している．この結果から，AR を用いた Attention の提示は，ユーザがロボット動作を予測する上で有効と考えられる．

ユーザのロボット動作に対する理解度　ユーザのロボット動作に対する理解において，AR を用いた Attention 可視化が有効であるか確認するためにアンケート調査を行った．定点カメラによるロボット動作シーンの動画 (定点肉眼観測) と，AR による Attention 可視化を含むロボット動作シーンの動画 (AR を介した観測) を 23 名の被験者に視聴してもらった．その後，被験者に対しロボット動作に対する理解度アンケートを行った．本調査では理解度を 4 段階評価とした．また，動画は図 2 (b) に示す実環境下で 4 パターン撮影し，AR デバイスには Microsoft 社が開発した HoloLens2 を用いた．各動画に対する理解度アンケートの集計結果を表 2 に示す．表 2 から，AR を介した観測が定点肉眼観測と比較し，“理解できる” と答えた被験者の割合が増加している．この結果から AR を用いた Attention 可視化によるユーザへの提示は，ユーザがロボット動作を理解する上で有効と考えられる．

5. おわりに

本研究では，深層強化学習によるロボットの自律動作獲得と行動選択に対する判断根拠の可視化，およびユーザに対する AR を用いた効率的なロボット動作の理解を提案した．提案手法は Transformer の導入によりロボットの高い自律移動性能を獲得し，ロボット動作に対する視覚的説明を実現した．また，AR を用いてロボット動作の判断根拠を可視化することで，ユーザのロボット動作への理解促進を実現し，人とロボットの共存に貢献できる可能性を示した．

参考文献

[1] A. Vaswani, *et al.*, “Attention is all you need”, Advances in neural information processing systems, 2017.

[2] H. Itaya, *et al.*, “Action Q-Transformer: Visual Explanation in Deep Reinforcement Learning with Encoder-Decoder Model using Action Query”, arXiv preprint arXiv:2306.13879, 2023.

[3] V. Mnih, *et al.*, “Human-level control through deep reinforcement learning”, Nature, Vol. 518, No.7540, pp. 529–533, 2015.

研究業績

[1] 尹文韜，板谷英典，真野航輔，平川翼，山下隆義，藤吉弘亘,“Transformer モデルによる自律移動の視覚的説明と拡張現実による提示”, 日本ロボット学会学術講演会, 2023.