知識選択型転移強化学習を用いた移動ロボットによる動的障害物回避

高矢　空＊　　河野　仁　　須賀　哉斗　　鳥谷部　悠希（東京電機大学）池勇勳（北陸先端科学技術大学院大学）　　藤井　浩光（千葉工業大学）  
鈴木　剛（東京電機大学）

Dynamic Obstacle Avoidance by Mobile Robot Using Transfer Reinforcement Learning with Knowledge Selection

Takaya Sora＊, Kono Hitoshi, Toriyabe Yuki, Suga Kanato, (Tokyo Denki University)   
Ji Yonghoon, (Japan Advanced Institute of Science and Technology)   
Fujii Hiromitsu, (Chiba Institute of Technology). Suzuki Tsuyoshi, (Tokyo Denki University)

In recent years, machine learning technologies have been actively implemented in society. Especially technologies such as reinforcement learning and transfer learning are being implemented in intelligent robot systems. The authors have proposed knowledge-selective transfer reinforcement learning based on the spreading activation model, which is a knowledge in the field of cognitive science. In past research, although static obstacle avoidance has been achieved in mobile robots with knowledge-selective transfer reinforcement learning, dynamic obstacle avoidance has not been investigated. In this paper, it is realized that dynamic obstacle avoidance by tuning the hyper-parameters of transfer reinforcement learning with knowledge selection.

キーワード：転移学習，強化学習，知識選択，活性化拡散モデル

(Transfer learning, reinforcement learning, knowledge selection, spreading activation model)

1.　緒言

近年,自動運転技術の進展により,利用者が運転の負担から解放され,交通の安全性と効率性が向上する可能性が高まりつつある.自動運転車は機械学習技術を駆使して,リアルタイムの状況判断や障害物検知,交通ルールの遵守などを行い,自律的な運転を実現する.しかし,自動運転車が直面する課題の一つとして,動的な障害物の回避がある.現行の自動運転技術では静的な障害物への対応は従来の経路計画手法などで適応可能であるが, 動的な障害物の予測と回避は依然として課題とされている.このような課題に対して著者らは,知識選択型転移強化学習を用いた移動ロボットにおける動的障害物回避の実現を提案している(1) (2).

著者らの従来の研究においても同様に自動運転シニアカーにおける静止障害物の回避は実現しているが，動的障害部の回避は実現できていな．そこで，本研究では,知識選択型転移強化学習を用いた移動ロボットによる動的障害物回避について基礎的な検討と実験を行ったので報告する．知識選択型転移強化学習のハイパーパラメータを調整することで,動的障害物の回避が実現できることを確認し，知識選択型転移強化学習は,過去の学習経験を転移させることで学習速度の向上や新しい環境への適応度の向上を図る手法であることを示す．.

2.　提案システム

〈2･1〉 知識選択手法SAP-net

SAP-net(Spreading Activation Policy Network)とは、知識選択型の転移学習モデルのことを指す．転移学習前の手続きとして，予め学習した複数の強化学習の知識を距離と角度で知識として分類する．それらの複数の知識に類似度を計算した集合を指している．生成したカテゴリを用いて知識のネットワークを作成し保存する.

〈2･2〉 ハイパーパラメータチューニング

また、先述のSAP-net(Spreading Activation Policy Network)では、すでに静的障害物の検知及び回避には完了している。しかし、本研究ではSAP-netにあらたに角度情報と距離情報をタグとして各強化学習の知識を紐づけることで動的障害物の回避を可能にしている。

3.　提案手法

〈3･1〉 前提条件

前提条件を記載する.

〈3･2〉 提案手法の流れ

提案手法の流れを記載する.

〈3･1〉 それぞれの実験内容に合わせ記載

SAP-net(Spreading Activation Policy Network)とは、知識選択型の転移学習モデルのことを指す．転移学習前の手続きとして，予め学習した複数の強化学習の知識を距離と角度で知識として分類する．それらの複数の知識に類似度を計算した集合を指している．生成したカテゴリを用いて知識のネットワークを作成し保存する.

4.　実験と結果

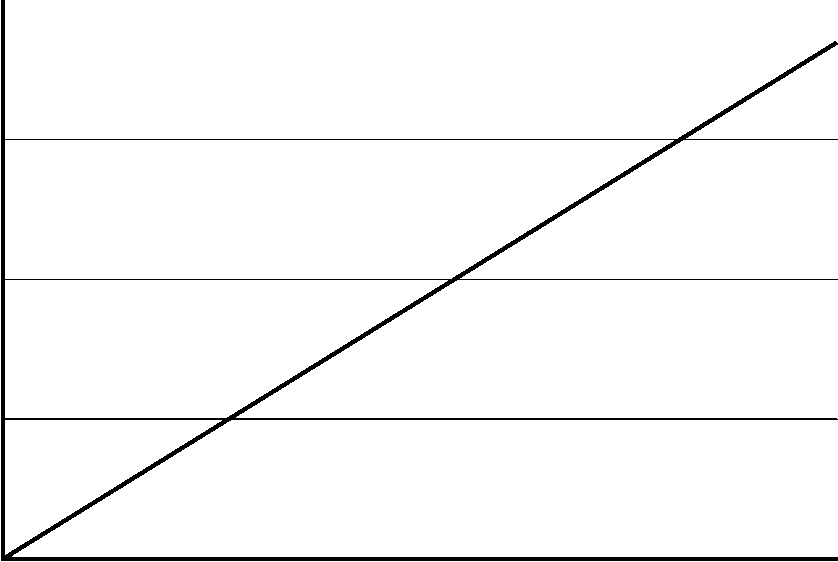
〈3･1〉 実験目的本提案手法における動的障害物回避を評価するために物理演算シミュレータ内に構築した．物理演算シミュレータにはCyberbotics社製のWebots 2023aを使用し，図〇のような環境を構築する．

□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□（表2参照）。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□（図1参照）。



(a)　グラフ1

図1　図表タイトル

Fig. 1.　Title.

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

本文□□□□□□□□□□□□□□□□□□□□□□□□□□□５□□□□10□□□□15□□□□20□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

0□□□□25□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□。

3.　結言

本論文では，自律移動ロボットの動的障害物回避を目的として，知識選択型転移強化学習であるSAP-netのハイパーパラメータを調整することで動的障害物回避が実現可能であることを示唆した．

今後の課題として，本論文では少ない行動価値関数の個数であることや動的障害物が単調な移動しか行わない条件でのシミュレーション検証であった．より複雑な条件でのシミュレーションや実際の移動ロボットへの実装も実施する．

謝　　　辞

　本研究の一部はJSPS科研費JP23K11276の助成を受けたものである．

文　　　献

(１) H. Kono, R. Katayama, Y. Takakuwa, W. Wen, and T. Suzuki: "Activation and Spreading Sequence for Spreading Activation Policy Selection Method in Transfer Reinforcement Learning", International Journal of Advanced Computer Science and Applications, Vol.10, No.12, pp.7-16 (2019)

(２) 河野仁，坂本裕都，温文，藤井浩光，池勇勳，鈴木剛，“知識選択型転移強化学習を用いたシニアカーの自律運転”，2022年電気学会電子・情報・システム部門大会，pp. 714-718，広島，2022．

(３) Name, Name, and Name : “Title Eng”, 雑誌名, Vol.巻数, No.号数 pp.000-000 (発行年)  
著書名・著書名・著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(４) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(５) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(６) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(７) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(８) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(９) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(10) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)

(11) 著書名：「タイトル」，雑誌名，Vol.巻数，No.号数 pp.ページ数 (発行年)