

制御学習法

工学研究科 山田 訓

Keywords: 自律的学習制御、強化学習、ロボット、モジュール型強化学習

1. 開発目的

人間と共存するロボットは、時々刻々と変化する環境の中で適切な制御を行なう必要がある。強化学習は、一連の制御を行った後で、その結果に対する評価（よかったか悪かったか）を与えることによって制御を学習していく学習法で、自律的に制御を学習できる。実際の世界の状況を理解し、適切な制御を行うためには、多数のセンサ情報を入力し、処理する必要がある。モジュール型強化学習は生物の脳のモジュール構造を模擬した学習法で、複数のセンサ情報を処理する必要がある実際の世界での学習が可能になるように提案している学習法である。

2. 技術の概要と特徴

赤外線センサで壁を検知し、CCDによって環境内にある対象物を認識し、壁を避けつつ適切な対象物に近づく制御を学習させた。赤外線センサの情報を処理するモジュールとCCDの情報を処理するモジュール及び、制御モジュールを選択する選択モジュールからなるモジュール型強化学習システム（図1）で学習を行った。環境内に黒塗りの棒と黒白に塗った棒を設置し、壁に衝突せずに、黒白の棒に近づくことが制御課題である。CCDセンサの情報は、目的物のエッジの位置と目的物の幅、エッジの数を抽出するという前処理を施し、強化学習モジュールに入力した。その結果、壁付近では、赤外線センサを処理するモジュールが、対象物のそばに来ると目的物到達制御モジュールが選択され、適切な対象物を選んで近づく制御が学習できた（図2）。

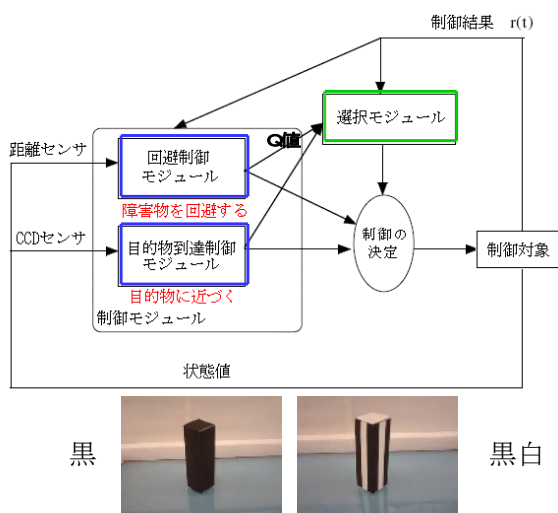


図1 モジュール型強化学習システムと環境内においた2種類の対象物

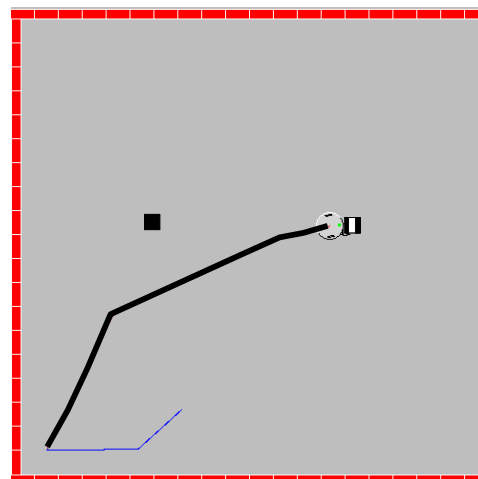


図2 学習された制御によるロボットの動き。太線ではCCDの情報を処理するモジュールを、細線では赤外線センサの情報を処理するモジュールを選択している。

3. 応用の可能性

複数のセンサ情報を処理した制御の自律的な学習が可能であるので、多様な動きの制御に適用できる。なお、この研究は文科省高度化推進事業・社会連携研究推進事業の一環である。