

Title	Self-construction of State Spaces of Single and Multi-Layered Learning Systems for Vision-based Behavior Acquisition of A Real Mobile Robot
Author(s)	高橋, 泰岳
Citation	大阪大学, 2002, 博士論文
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/352">https://hdl.handle.net/11094/352</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	高橋泰岳
博士の専攻分野の名称	博士(工学)
学位記番号	第 17102 号
学位授与年月日	平成14年3月25日
学位授与の要件	学位規則第4条第1項該当 工学研究科電子制御機械工学専攻
学位論文名	Self-construction of State Spaces of Single and Multi-Layered Learning Systems for Vision-based Behavior Acquisition of A Real Mobile Robot (実移動ロボットの視覚に基づく行動獲得のための単層および階層型学習システムにおける状態空間の自律的構成)
論文審査委員	(主査) 教授 浅田 稔  (副査) 教授 白井 良明 教授 北橋 忠宏

### 論文内容の要旨

本論文では、実ロボットによる行動学習のための状態空間の漸次的構成、複数の学習器の階層的構築による行動獲得、階層型学習機構における状態行動空間の構成方法を提案した。近年、強化学習を用いたロボットによる自律的な行動の獲得に関する研究がこれまで多くなされてきた。これらの手法には先見的な知識をほとんど必要とせずに適応的で即応的な行動を獲得できる利点が良い知られているが、状態空間をあらかじめ設計することが難しく、またその拡張性がないことが実環境におけるロボットタスクへの適用を妨げていた。そこで、状態空間を自律的に構成する手法、複数の学習器を用いて探索時間の爆発を抑える手法や、またその階層型学習機構の自律的獲得手法を提案した。

まず1章では、実環境におけるロボットのタスクに強化学習を適用する場合の問題点と、それに対する有効なアプローチについて述べた。

2章では、従来手法との比較について述べた。

3章では、ロボットが自身の経験に基づいて逐次的にセンサ空間を分割して現実的な時間内に目的の行動を学習する手法を提案した。すなわち、センサ出力に基づくセンサ変化と報酬信号の関数近似による局所モデルを構築し、このモデルに基づいてセンサ空間を分割することによって状態空間を構成する。本手法をボールをゴールにシュートするサッカーロボットに適用した。計算機によるシミュレーションと実機による結果を示した。

4章では、同一構造の学習器を複数用いて階層的に構築することによる行動獲得法を提案した。下位の層の学習器はそれぞれ異なるサブゴールを担当し、低レベルな行為を学習する。同時に上位の学習器は、下位の学習器を利用し、より高いレベルの行為を学習する。従来の手法とは異なり、それぞれの学習器が担当するサブゴール、タスクは自律的に決定され、また階層も自律的に構成される。提案する手法をナビゲーションタスクに適用した結果を示した。

5章では、階層型学習機構において異なる部分空間を状態空間として持つ下位の層の学習器を統合し、より多くの状態変数によって表現される状況における行為を上位の層で獲得する手法を提案した。提案する手法をロボカップに出場しているロボットを使用し、ボールをゴールにシュートするタスクに適用した結果を示した。

最後に6章では、本論文のまとめと今後の課題を述べた。

## 論文審査の結果の要旨

強化学習や遺伝的アルゴリズム等の手法を用いたロボットによる自律的な行動の獲得に関する研究がこれまで多くなされてきた。これらの手法には先見の知識をほとんど必要とせずに適応的で即応的な行動を獲得できる利点が多く知られている。しかし実環境におけるロボットタスクに強化学習を適用する場合、最も困難な問題の一つはスケラビリティがないことであろう。この問題を解決するために大きく分けて二つのアプローチがある。一つは状態空間をコンパクトに表現することで探索範囲を小さく抑えること。もう一つは与えられたタスクを複数のサブタスクに分解し、複数の学習器によりサブタスクに対する行為を獲得させ、それらを組み合わせることである。

本論文では以下の三つの手法を提案し、その有効性を実機を用いた実験により確かめている。

### (1)実ロボットによる行動学習のための状態空間の漸次的構成

ロボットが自身の経験に基づいて逐次的にセンサー空間を分割して現実的な時間内に目的の行動を学習する手法を提案している。センサ出力に基づくセンサ変化と報酬信号の関数近似による局所モデルを構築し、このモデルに基づいてセンサ空間を分割することによって状態空間を構成する。本手法をボールをゴールにシュートするサッカーロボットに適用している。計算機によるシミュレーションと実機による結果を示している。

### (2)複数の学習器の階層的構築による行動獲得

同一構造の学習器を複数用いて階層的に構築することによる行動獲得法を提案している。下位の層の学習器はそれぞれ異なるサブゴールを担当し、低レベルの行為を学習する。獲得された学習器に基づいて下位の層の状態・行動空間から状況・行為を定義し、これを上位の層の状態・行動空間として採用し、より抽象化された行為を学習していく。従来の手法とは異なり、それぞれの学習器が担当するサブゴール、タスクは自律的に決定され、また階層も自律的に構成される。提案する手法をナビゲーションタスクに適用した結果を示している。

### (3)階層型学習機構における状態行動空間の構成

階層型学習機構において異なる部分空間を状態空間として持つ下位の層の学習器を統合し、より多くの状態変数によって表現される状況における行為を上位の層で獲得する手法を提案している。提案する手法をロボカップに出場しているロボットを使用し、ボールをゴールにシュートするタスクに適用した結果を示している。

以上のように本論文は、実環境におけるロボットタスクに強化学習を適用する場合の一つの問題点である状態行動空間を構成する手法を提案するとともに、環境との相互作用を通して自分自身の行為を発達させていくロボットの実現に大きな示唆を与えており、実世界で動く知能機械の研究の発展に寄与することが大きい。よって本論文は博士論文として価値のあるものと認める。