

氏名	綿田 将悟
授与学位	博士(工学)
学位記番号	理工博甲第694号
学位授与年月日	平成28年3月17日
学位授与の要件	学位規則第4条1項
研究科, 専攻の名称	理工学研究科(博士後期課程) 情報・デザイン工学系専攻
学位論文題目	自律ロボットのための情動に基づく意思決定システムに関する研究
論文審査委員	主査 山口大学 教授 大林 正直 山口大学 教授 江 鐘偉 山口大学 教授 田中 幹也 山口大学 教授 松藤 信哉 山口大学 教授 浜本 義彦

【学位論文内容の要旨】

近年、人口減少や高齢化等により、医療や福祉の様々な分野における人的作業の補助を務める社会貢献型の高度な機能を有するロボットの開発が進んでいる。さらには、それらの自律的な動作を実現するため、人間の脳機能、特に、学習機能・記憶機能・適応機能をモデル化し、これらを利用した新たな制御システムの開発が盛んに行われている。その中で、人間の脳機能の一つである情動に焦点をあて、情動のモデル化やその工学的応用に関する研究が数多く存在する。生物の情動は、情動が持ち主の内部状態や外部環境に影響を受けて柔軟に表現することができる点や他者とのコミュニケーションを円滑にしている点で、種々の工学分野で応用可能であり、魅力的な研究分野である。ロボットの人工情動の研究においては、ロボットの外部刺激に対する情動的評価とそれに基づく情動的行動を実現するため、人間の情動を脳神経学や心理学の観点からモデル化した様々な情動モデルが提案されている。しかし、それらの研究のほとんどは人間が進化過程で既に獲得した情動の機能をモデル化したものである。一方、我々人間は情動発生メカニズムを進化の過程や後天的な経験で環境に適応して獲得した。自律的なロボットの開発の観点では、人工情動においても環境に適応的なシステムの構成方法が必要である。そこで、我々は、人間の情動反応を模倣させるのではなく、必要な情動反応をロボット自身が外部環境に適応して形成することが可能なシステムを提案する。

本稿ではまず、第1章で研究背景と目的について述べる。第2章では本研究におけるテーマである情動について、その概念と生体における働き、そして情動の工学分野への応用に関する先行研究について述べる。

次に、第3章では、マルコフ情動モデルに基づく自律ロボットの意思決定システムを提案する。システムに搭載するマルコフ情動モデルは有限オートマトンにより構成された確率モデルであり、心理学の観点から情動をモデル化されている。システムは2つの情動学習過程により、ロボットのスペックや対象とするタスクに適応した情動反応と情動行動に関するルールを自動的に構築可能で

ある。1つ目の情動形成学習は感覚刺激と情動反応の対応付けを行う学習過程である。この学習は自己組織化写像 (SOM) を用いた感覚刺激のクラスタリングによって行われる。2つ目の情動行動学習は情動反応と情動行動を関連付ける学習過程である。この学習はタスクの試行におけるシステムの設計パラメータを適切に調節することにより行われる。本研究ではパラメータの決定に遺伝的アルゴリズムを用い、個体の適応度は行動タスクの結果に基づいて決定される。システムはマルチロボットによる未知環境同定問題のシミュレーションによって性能評価された。結果より、適切な行動決定法がシステムの2つの情動学習過程により自動的に構築され、有効な行動決定が効果的に生成されることが確認された。しかし、提案システムの性能はその情動形成学習に SOM を用いず試行錯誤により手動設計したシステムに比べて劣る性能評価となった。この原因は、感覚刺激の教師無し学習時に、感覚刺激の発生確率を考慮しなかったため、SOM の競合層上に発生頻度の低い無駄な状態領域が構築されたためだと考えられる。

第4章では、SOM の競合層上における無駄な状態空間を減少させることを目的として、ロボットの経験に基づく情動の再形成学習を提案手法へ導入する。感覚刺激の予測値に基づく学習サンプルのみならず、ロボットがタスク中に得た感覚刺激をオンラインで学習可能となるようにシステムの部分的な改良を行う。改良手法の移動ロボットに関するシミュレーション結果より、情動の再形成学習により SOM 中の不必要な感覚刺激への対応付けが減少し、生成されるロボットの行動決定法の性能を向上させることが示された。これは、タスクにおける感覚刺激の頻度を考慮することによって、感覚刺激と情動反応の関係性がより効果的に修正されたことを意味している。また、それらの過程による SOM の学習結果や各設計パラメータを解析することで、情動反応と環境状態との関係性や、その行動決定への影響が意味的に理解可能であることが確認された。しかしながら、これらの研究では、システムにより設計される行動決定法は学習する行動タスクに依存するため、目的とする行動タスク毎に学習する必要がある。システムにおける学習過程は膨大な繰り返し計算が必要であるため、この問題点は実機ロボットにおける学習を困難とする場合があると考えられる。

第5章では、システムのさらなる汎用性を示すために、提案システムの強化学習への応用について述べる。システムに基本的な行動学習として強化学習を導入し、情動は学習を効率的に行うための補助的な役割をもつ学習戦略を提供する新たなシステムを提案する。ここで、学習戦略はタスクに依存せず、あらゆるタスクに再利用可能な知識の効果的な利用方法を指す。つまり、目的タスクに類似したタスクのシミュレーションにおいて学習戦略を獲得し、その学習戦略を実ロボットによる行動学習に利用することで学習を効率化することができると考えられる。本研究ではシステムにおける学習戦略を強化学習におけるメタパラメータの適応的制御とし、システムは環境変化を有する迷路探索問題に関する学習タスクのシミュレーションを用いて評価を行った。結果では、適切なメタパラメータの制御器がシステムにより自動的に設計され、獲得された制御器を用いることで強化学習手法による行動学習を効率化することが確認された。さらに、提案システムは従来のメタパラメータ制御手法よりも複雑な制御ルールを表現可能であることが示された。

第6章では、まとめと本研究の将来展望について述べる。本研究の成果として、我々は情動反応と情動行動のルールを自動的に生成可能なシステムおよびその学習手法を提案した。さらに、強化

学習への応用システムの提案により、実際のロボットへの対応におけるシステムのより汎用性の高い利用方法を示した。これらの成果は、ロボットのより自律的な動作の実現に貢献すると考えられる。

【論文審査結果の要旨】

近年、人口減少や高齢化等により、医療や福祉の様々な分野における人的作業の補助を務める社会貢献型の高度な機能を有するロボットの開発が進んでいる。その中で、人間の脳機能の一つである情動に焦点をあて、情動のモデル化やその工学的応用に関する研究が数多く存在する。生物の情動は、情動が持ち主の内部状態や外部環境に影響を受けて柔軟に表現することができる点や他者とのコミュニケーションを円滑にしている点で、種々の工学分野で応用可能であり、魅力的な研究分野である。

ロボットの人工情動の研究においては、ロボットの外部刺激に対する情動的評価とそれに基づく情動的行動を実現するため、人間の情動を脳神経学や心理学の観点からモデル化した様々な情動モデルが提案されている。しかし、それらの研究のほとんどは人間が進化過程で既に獲得した情動の機能をモデル化したものである。一方、我々人間は情動発生メカニズムを進化の過程や後天的な経験で環境に適応して獲得したものであり、自律的なロボットの開発の観点では、人工情動においても環境に適応的なシステムの構成方法が必要である。このような背景の下、申請者は、人間の情動反応を模倣させるのではなく、必要な情動反応をロボット自身が外部環境に適応して形成することが可能なシステムを提案している。具体的な論文構成は以下となっている。

第1章で研究背景と目的について述べている。

第2章では本研究におけるテーマである情動について、その概念と生体における働き、そして情動の工学分野への応用に関する先行研究について述べている。

第3章では、マルコフ情動モデルに基づく自律ロボットの意思決定システムを提案している。システムに搭載するマルコフ情動モデルは有限オートマトンにより構成された確率モデルであり、心理学の観点から情動をモデル化している。提案されたシステムは情動形成学習及び情動行動学習の2つの情動学習過程により、ロボットのスペックや対象とするタスクに適応した情動反応と情動行動に関するルールを自動的に構築可能としている。1つ目の情動形成学習は感覚刺激と情動反応の対応付けを行う学習過程である。この学習は自己組織化写像 (SOM) を用いた感覚刺激のクラスタリングによって行われる。2つ目の情動行動学習は情動反応と情動行動を関連付ける学習過程である。この学習はタスクの試行におけるシステムの設定パラメータを適切に調節することにより行われる。提案システムはマルチロボットによる未知環境同定問題のシミュレーションによって性能評価された。結果より、適切な行動決定法がシステムの2つの情動学習過程により自動的に構築され、有効な行動決定が効果的に生成されることが確認された。

第4章では、SOMの競合層上における無駄な状態空間を減少させることを目的として、ロボットの経験に基づく情動の再形成学習を提案手法へ導入する。感覚刺激の予測値に基づく学習サンプルのみならず、ロボットがタスク中に得た感覚刺激をオンラインで学習可能となるようにシステムの部分的な改良を行っている。改良手法の移動ロボットに関するシミュレーション結果より、情動の再形成学習により SOM 中の不必要な感覚刺激への対応付けが減少し、生成されるロボットの行動決定法の性能を向上させることが示された。しかしながら、これらの研究では、システムにより設計される行動決定法は学習する行動タスクに依存するため、目的とする行動タスク毎に学習する必要がある。システムにおける学習過程は膨大な繰り返し計算が必要であるため、この問題点は実機ロボットにおける学習を困難とする場合があると考察している。

第5章では、提案システムの応用方法及び汎用性を示すために、提案システムの強化学習への応用について検討している。システムにおける基本的な行動学習として強化学習を導入し、情動は強化学習を効率的に行うための補助的な役割をもつ学習戦略を提供する役割としての新たなシステムを提案している。学習戦略とはタスクに依存せず、あらゆるタスクに再利用可能な知識の効果的な利用方法を指す。つまり、目的タスクに類似したタスクの事前シミュレーションにおいて学習戦略を獲得し、その学習戦略を実ロボットによる行動学習に利用することで学習を効率化できると考察している。本章では提案のシ

システムにおける情動行動を学習戦略として扱い、この戦略を強化学習における（学習の効率を左右する）メタパラメータを適応的に制御することに利用している。そして、環境変化を有する迷路探索問題に関する学習タスクのシミュレーションを通して提案システムを評価している。結果では、適切なメタパラメータの制御器が提案システムにより自動的に設計・獲得された制御器を用いることで強化学習手法による行動学習が効率化されていることが確認されている。さらに、提案システムが従来のメタパラメータ制御手法よりも複雑な制御ルールを表現可能であることを示している。

第6章では、まとめと本研究の将来展望について述べている。本研究の成果として、情動反応と情動行動のルールを自動的に生成可能なシステムおよびその学習手法を提案したこと。さらに、提案した情動システムを強化学習への応用システムとして利用することの提案により、実際のロボットへの対応におけるシステムのより汎用性の高い利用方法を示している。これらの成果は、ロボットのより自律的な動作の実現に貢献すると考えられる。

公聴会には、学外者も含め 30 余名の参加者があり、活発な質疑応答がなされた。その主な質疑内容として、

- (i) ロボット（機械）にとって情動とはどのような意味を持つか
- (ii) 情動再形成の役割について。再形成でシンプルにすれば多様性が生まれなくなるのではないか
- (iii) 人間の情動を知る手段・分類について。例えば、情動に対して現れる行動の種類から情動を分類することも可能か
- (iv) 感覚刺激としてロボットへ自身のオドメトリーを入力しているが、時系列データを入力するとよりよい刺激・情動行動につながると思うがそうしていないのは何故か
- (v) 今後の課題で、ロボットに自律的に獲得させたい、人間がまだ獲得していない機能性情動とはどのようなものか

等の質問があり、いずれの質問に対しても申請者からの確かな回答がなされた。以上より、本研究は、新規性、信頼性、有効性、実用性ともに優れており、博士（工学）の論文に十分に値するものと判断した。

論文内容、審査会、及び公聴会での質問に対する応答などから総合的に判断して、最終試験は合格とした。

なお、査読のある関連論文の発表状況は以下の通りである。（下記以外 国際会議会議録 2編）

- (1) 綿田将悟, 大林正直, 呉本堯, 小林邦和, 間普真吾, マルコフ情動モデルを備えたロボットの行動決定法, 電気学会論文誌 C, Vol, 134, No.1, pp, 85-93, 2014
- (2) Shogo Watada, Masanao Obayashi, Takashi Kuremoto, Shingo Mabu, A Decision Making System of Robots Introducing a Re-construction of Emotions Based on Their Own Experiences, Journal of Robotics, Networks and Artificial Life, Vol. 1, No.1, pp. 27-32, 2014