

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 6 月 1 日現在

機関番号：17104

研究種目：基盤研究（C）

研究期間：2009 ～ 2011

課題番号：21500218

研究課題名（和文） 確率分布間距離に基づく好奇心概念の一般化とロボットの自律的発達

研究課題名（英文） A GENERALIZATION OF THE CONCEPT OF CURIOSITY BASED ON DISTANCE OF PROBABILISTIC DISTRIBUTIONS AND ITS APPLICATION TO AUTONOMOUS DEVELOPMENT OF ROBOTS

研究代表者

永松 正博 (Masahiro Nagamatu)

九州工業大学・生命体工学研究科・教授

研究者番号：70117307

研究成果の概要（和文）：好奇心や本能的欲求は、ロボットの自律的発達、学習、行動獲得、解空間探索に有用である。本研究では、好奇心や本能的欲求の定義を行い、これらが、自律ロボットの自己位置推定において頑健な行動選択と早期の推定に有効であること、追跡ゲームにおける行動獲得において早期の獲得を可能にすること、最適化問題やレイアウト問題において効率の良い探索を可能にすることを示した。

研究成果の概要（英文）：The curiosity and instinctive desires are useful for autonomous development, learning, behavior acquisition and space exploration for robots. In this study we define the curiosity and instinctive desires properly, and show how these can be used effectively for robust behavior selection and quick inference for the self-localization of autonomous robots, quick behavior acquisition in pursuit games, and solution space exploration in the optimization problems and layout problems.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,500,000	450,000	1,950,000
2010年度	1,100,000	330,000	1,430,000
2011年度	900,000	270,000	1,170,000
年度	0	0	0
年度	0	0	0
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：Surprise, 好奇心, 確率分布, 強化学習, 自己位置推定, 探索

1. 研究開始当初の背景

(1) SONY CSL Paris の Oudeyer らは予測誤差が減少する対象に興味を持つという Schmidhuber の好奇心第二仮説をロボット発達に適用した研究を行っている。しかしこの方法は、環境地図作成や意図推定などへの好奇心の一般化に際して予測誤差が求め難いことなど

の欠点を有する。マサチューセッツ大の Barto らは強化学習への内発的報酬の導入を提案したが、内発的報酬がタスク依存的であり、説得力に欠ける。他方、Baldi らは、事前確率分布と事後確率分布の間の距離を surprise と定義した。
(2) 申請者は、ロボットの自律的発達にとって、多様な世界で適切な行動をするため

には、環境からの報酬に加えて、特殊的好奇心、拡散的好奇心、新奇性、生存価値など複数の内発的報酬が必要であること、これらの内発的報酬がタスク非依存的事であること、科学的に興味ある行動が創発するだけでなく、工学的意義のある行動が創発することが重要であることを主張した。

(3) Baldi らの surprise を好奇心と結びつけた研究は知る限りではない。申請者は、工学的な好奇心研究と、数理統計の流れを汲む surprise を結びつけることを着想した。この統合により理論的基盤を持ち広範囲に適用可能な好奇心を定義でき、対象物体、環境地図、環境変化、行動観察に基づく意図推定という、好奇心対象のより抽象的方向への一般化が期待できる。これにより広範な場面での自律的発達を効果的に実現できると考えている。

2. 研究の目的

(1) バイズの考えに基づく surprise を好奇心に導入する。surprise の時間的減少は学習の進展を意味しているので、好奇心の仮説の第一候補として、学習の進展が最も大きな対象に対して好奇心を持つものとする。環境地図作成などにこの好奇心を使用し、従来の予測誤差に基づく好奇心との性能比較および有効性の検証を行う。

(2) 環境からの報酬に加えて、生存価値などの内発的報酬を用いると、強化学習を効率的に進めることができることを示す。バッテリー切れを起こさないようにしながら追跡ゲームを行う環境を考え、この環境において、内発的報酬の定義の方法が、どのように学習速度に影響を与えるが調査し、内発的報酬の定義方法を提案する。

(3) 特殊的好奇心、拡散的好奇心を解空間の探索に適用する方法を検討し、提案する。

3. 研究の方法

(1) まず surprise の定式化を行い、これを好奇心に導入する際の理論的・実際的問題点を解明する。引き続き、対象物体の予測、地図構築、環境変化、行動観察に基づく意図推定の4課題に対して、予測誤差に基づく好奇心、Kullback-Leibler 情報量に基づく好奇心、あるいは従来法を適用し、提案の確率分布間距離である Kullback-Leibler 情報量に基づく好奇心の有効性を、シミュレーション実験及び実機実験により検証する。

① 対象物体の予測、地図構築、環境変化、意図推定という4種類の課題に対し surprise の定式化を行う [1]。Kullback-Leibler 情報量は距離の公理

を満たさず、事前確率分布から事後確率分布までの距離と、事後確率分布から事前確率分布までの距離は一般に異なる。またこの距離は比例尺度ではない。どの距離を用いるかも含め、Kullback-Leibler 情報量を好奇心に導入する際の理論的・実際的問題点を洗い出し、解決策を求める。

② 静止物体、一定速度で移動する物体、確定的ダイナミクスに従って移動する物体、確率的ダイナミクスに従って移動する物体の4種類の対象がある場合、予測誤差に基づく好奇心および Kullback-Leibler 情報量に基づく好奇心を用いて、4対象物体の学習がどの順序で、それぞれどの程度進展するかを、シミュレーション実験により調べる。過渡状態および定常状態での性能比較を行うとともに、性能の違いの理由を明らかにする。なお好奇心に関して、Schmidhuber は「予測誤差の大きな対象に好奇心を持つ」、「予測誤差減少の大きな対象に好奇心を持つ」の2種類の仮説を提案している [2]。いずれの仮説を用いるか、あるいは第三の仮説を提案するかなども含めて検討する。

③ ロボットフィールドのグリッドベース地図を構築する際、移動ロボットの行動を、予測誤差に基づく好奇心の大きさにより決定、Kullback-Leibler 情報量に基づく好奇心の大きさにより決定、あるいは従来法により決定する際、地図の構築が如何に行われるかを、シミュレーション実験により調べる。過渡状態および定常状態での地図の精度比較を行い、精度の空間的な差異についても調べるとともに、精度の違いの理由を明らかにする。

④ ロボットフィールドの一部を突然変化させる場合、予測誤差に基づく好奇心の値が閾値を越える、あるいは Kullback-Leibler 情報量に基づく好奇心の値が閾値を越えるという基準に基づいて環境変化への気づきを判断する。これをシミュレーション実験により調べる。過渡状態および定常状態での性能比較を行うとともに、性能の違いの理由を明らかにする。なお一般的な環境変化を対象とすることは困難があることも予想されるので、予備的実験の結果によっては、環境変化が時間的・空間的に局限された比較的容易な課題を対象とすることもあり得る。方法論の確立にはこれでおおむね差し支えないが、どの程度実用的な意義のある課題に適用できるかに関しては不確定要素が残されている。

(2) 通常ロボットのミッションとしては、与

えられたタスクを実行すること以外に、危険を回避する、バッテリーが上がって動けなくなることを回避する等の、生存を維持することを要求される。このような環境をモデル化し、環境からの報酬に加えて、生存価値などの内発的報酬を用いると、強化学習を効率的に進めることができることをシミュレーションにより実験をおこなう。実験では、内発的報酬の定義の方法をさまざまに変え、どのように学習速度に影響を与えるが調査する。

- (3)まず最適化問題、レイアウト問題に特殊的好奇心、拡散的好奇心を導入する方法を検討する。次に、これらの好奇心の定義方法が解空間探索に影響を与えるか実験により調査を行い、好奇心導入方法のできるだけ一般的な手法を提案する。

4. 研究成果

- (1)これまで地図と位置・方位の同時推定 (SLAM) において、その性能を最大化するための経路最適化のために情報量増加を用いることが提案されていた。これに対し、情報量増加の代わりに **surprise**、すなわち事後確率分布から事前確率分布までの **Kullback-Leibler** 情報量を用いることを提案した。前者は事前確率に対しては事前確率で重み付けした情報量を定義し、事後確率に対しては事後確率で重み付けした情報量を用いる。他方後者は、事前確率および事後確率のいずれに対しても事後確率で重み付けした情報量を用いる。後者の方が性能が良いことを示した。

- (2)環境地図の構築への好奇心の導入。
自律ロボットの環境内での自己位置推定において、**surprise** の概念を付加することで早期の自己位置推定を実現する手法を提案した。提案手法の基本となるパーティクルフィルタは、パーティクルと呼ばれる粒子を探索空間に複数個配置し、その集合で自己位置の確立分布を表現するものである。本研究では、パーティクルをクラスタリングすることで、自己位置の候補点を絞る。候補点をより絞り込むために、各候補点からどのように動けばよいかを選択する際に **surprise** の概念を利用している。具体的には、次の動作によって、各候補点に対するセンサ情報の期待値の分散が最大になるような行動を選択する。これにより、早期の自己位置推定が可能となった。また、クラスタリングの際に、壁や障害物を考慮することにより頑健な行動選択を実現した。

- (3)ロボットの行動獲得における本能的欲求の導入。

ロボットにタスクを学習させる際の、本能的欲求の導入の方法について研究を行った。本能的欲求の大きさを本能的欲求の対象となる身体状態 (例えば、食欲という欲求の場合は、空腹度) の関数として表現する方法を提案し、行動獲得の速さを上げるためには、どのような関数であればよいかを実験的に調べ、最も良い関数を提案した。

- (4)最適化問題、レイアウト問題への好奇心の導入。

複雑な環境において複雑な拘束条件を持つロボットの経路計画は困難な問題になる。これに対して、始めは時間をかけて経路を探すが、慣れてくると短い時間で解決できるようになり、さらに、活用と探索のバランスを好奇心を使って解決する方法を提案した。また、最適化問題に対し、好奇心と同様な仕組みを入れることにより、解空間探索を効率良く行う方法を提案した。

また、選言的な制約を含む制約条件付き最適化問題や制約充足問題に対して、好奇心を導入して解探索における、多様化と集中のバランスをうまく取る方法を提案した。また、それをレイアウト問題に適用し、複雑な制約を持つ対象でも、局所最適解に陥ることなく効率良く解を求めることができることを示した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 15 件)

- ①三澤秀明, 松田充史, 堀尾恵一, 強化学習と情動学習に基づく意思決定法: 利己的な判断による協調行動の創発, 知能と情報 (日本知能情報フレンジ学会誌), 査読有, Vol.24, 2012, 582-591
- ②H. Misawa, K. Horio, N. Morotomi, K. Fukuda, H. Taniguchi, Extrapolation of Group Proximity from Member Relations Using Embedding and Distribution Mapping, IEICE Trans. on Information and Systems, 査読有, Vol.E95-D, 2012, 804-811
- ③佐藤肇展, 田中佑季, 永松正博, モーションサンプルと窓を使用した粒子フィルタによる複数移動物体の追跡, バイオメディカル・ファジー・システム学会誌, 査読有, Vol.13 2011, 41-48
- ④Muhammad Aziz Muslim, Masumi Ishikawa, Tetsuo Furukawa, Task segmentation in a mobile robot by mnSOM and clustering with spatio-temporal contiguity, International Journal of Innovative

Computing, Information & Control, 査読有, Vol.5, 2009, 865-875

- ⑤ Yuya Nishida, Masahiro Nagamatu, Trajectory Planning with Dynamics Constraints for an Underactuated Manipulator, Lecture Notes in Computer Science, Springer, 査読有, Vol.5506, 2009, 787-793
- ⑥ Mikio Morita, Masumi Ishikawa, Emergence of Behaviors based on the Desire for Existence by Reinforcement Learning, Lecture Notes in Computer Science, Springer, 査読有, Vol.5506, 2009, 762-769

[学会発表] (計 37 件)

- ① Y. Li, K. Horio, Stress State Estimation Based on Photoplethysmogram Dynamics, Int. Workshop on Nonlinear Circuits, Communication and Signal Processing, 2012年3月6日, ホノルル
- ② 上田敏雄, 永松正博, 選言的制約充足手法を用いたUMLクラス図の自動レイアウト, バイオメディカルファジィシステム学会, 2011年度年次大会, 2011年10月29日, 山口大学(山口)
- ③ 松田充史, 三澤秀明, 堀尾恵一, 強化学習と情動学習に基づいた意志決定法による自律エージェントの社会的行動の創発, 第1回日本知能情報ファジィ学会九州支部学生会研究発表会, 2011年3月5日, 佐賀大(佐賀)
- ④ Mikio Morita, Masumi Ishikawa, Masahiro Nagamatu, Using the Internal Rewards for a Mobile Robot to Survive and Perform a Task Effectively, Annual Conference of Biomedical Fuzzy Association (BMFSA 2010), 2010年10月9日, Kitakyushu(Japan)
- ⑤ 佐藤雅彦, 永松正博, 車両型ロボットの経路プランニングに関する研究, 第17回電子情報通信学会九州支部学生会講演会, 2009年9月30日, 九州工業大学(飯塚)
- ⑥ 蔡幼波, 石川眞澄, Performance Comparison of Surprise and Information Gain in Building a Grid-based Map with Rao-Blackwellized Particle Filters, 日本神経回路学会第19回全国大会, 2009年9月26日, 仙台

6. 研究組織

(1) 研究代表者

永松 正博 (MASAHIRO NAGAMATU)

九州工業大学・生命体工学研究科・教授
研究者番号：70117307
(H21:研究分担者)

(2) 研究分担者

堀尾 恵一 (KEIICHI HORIO)
九州工業大学・生命体工学研究科・准教授
研究者番号：70363413

(3) 連携研究者

石川 眞澄 (MASUMI ISHIKAWA)
九州工業大学・生命体工学研究科・教授
研究者番号：60222973
(H21:研究代表者)