［研究助成（A）］

人とロボットでの価値観の相互理解によるパートナーシップの創発

Emergence of partnership by mutual understanding of human and robot priorities

2191008

研究代表者　奈良先端科学技術大学院大学 情報科学領域　助　教　小　林　泰　介

［研 究 の 目 的］

ロボット技術の発展により，福祉・介護分野やサービス業などでよく見かける，人と触れ合う作業が求められる課題（例えば，歩行補助など）へのロボット導入が現実味を帯びてきている。このようなロボットは人の価値観を理解し動作に反映させるだけでなく，ロボットが提供したいサービス内容もある種のロボットの価値観として人に伝えることが望ましい。すなわち，人・ロボット間での相互理解を推進する技術を開発することで，ロボットが人との適切なパートナーシップを構築できると考えている。

このようなロボット技術として，本研究期間では大きく以下の２つの開発に取り組んだ。

1. 人の主観に応じたロボット制御則の最適化（ロボットによる人の価値観の理解）
2. ロボットが求める人の挙動からの異常検知（人へのロボットの価値観の呈示）

これら２つの基盤技術による，人・ロボット間での双方向な力覚的インタラクション実現が本研究の目的となる。

［研究の内容，成果］

0. ロボットの開発

本研究実施にあたって，人と力覚的インタラクションが可能な移動ロボットを開発した。当初は４脚ロボットを想定していたものの，力覚的インタラクションの実現可能性などを検討した末に，HEBI Robotics 社製が開発している作動駆動型モバイルベースに，２軸アクチュエータを搭載するポール部を取り付ける構成とした（図１参照）。

図１　力覚ベース駆動型移動ロボット

各アクチュエータは直列弾性アクチュエータ機構となっており，弾性要素であるバネ前後の角度差を２つのエンコーダで計測することによって外力がもたらすトルクを計測可能である。この計測値を利用して，ポール部およびモバイルベースのアドミッタンス制御を実装した。す

なわち，ポールに加わったトルクの一部をポールの（受動的な）駆動力とし，残りをポール部に関するヤコビアンを介して外力に変換した後，モバイルベースの駆動力として与える。これら駆動力は仮想的なインピーダンスパラメータで定義されるダイナミクス（および実際の移動台車のダイナミクス）に従って各アクチュエータの角速度指令値へ変換される。

1. 人の主観に応じたロボット制御則の最適化

上記で実装したアドミッタンス制御では，仮想的なインピーダンスパラメータ次第で応答が大きく変化するため，最適化が必要である。本研究では，人とのインタラクションを可能な限り円滑にすることが最適化目的となりえるが，これは定量的に定義しづらく，個人差が予想される。そこで，最適化目的を人の主観と定めて，そのブラックボックスな評価関数に関して経験的に最大化（あるいは最小化）を図るベイズ最適化をパラメータ設計に活用する。

具体的には，開発した移動ロボットと人とが行い得る基本動作として，8の字移動タスクをパラメータ最適化時のタスクとして設定した。このタスクの試行中において，人はリアルタイムで移動ロボットの操作感を3段階評価し続ける。この評価は手元の小型デバイスへの入力としてデータ化され，タスク終了時における平均値を実際の評価結果とする。なお，このときの3段階評価については陽に基準を設けず，あくまで人の主観に委ねるものとした。

ベイズ最適化では，評価結果を得るまでの1試行中はパラメータを固定しておき，固定されていたパラメータと評価値の対応から評価関数をガウス過程回帰を用いて推定していく。その推定された関数から，推定精度が不十分な，あるいは最も良いであろうパラメータを選択して次試行に用いる。この繰り返しによって，経験データのみから効率良く最適なパラメータを探索・決定できる。ただし，一般的なベイズ最適化では評価関数の分散は一定であると仮定するものの，本タスクのように人の主観を評価関数とする場合，そこには意思決定の揺らぎという到底分散が一定にはなり得ない要素が含まれてしまう。そのため，分散を可変とする異分散ベイズ最適化手法を採用することで，人の意思決定の揺らぎに対しても頑健にパラメータを最適化できるようにした。

評価実験および人の主観の分析として，実装した制御則が持つ4パラメータの最適化を合計8人に対して個別に実施した。なお，この実験は本学倫理委員会の承認の上実施された。

上述した8の字移動タスク（図2参照）を個々人で合計30試行繰り返した際の，各試行までにおける最大評価値を図3にプロットした。4パラメータでは，各パラメータの候補値をたかだか10パターン用意したとしても，総組み合わせ数は10,000通りとなる。そのような問題設定でありながら，どの被験者においても，30試行内で概ねより良い評価値を得るパラメータを発見できていることが確認できた。

次に，人の主観の個人差を改めて確認すべく，実験で得られたデータに対してクラスタリング

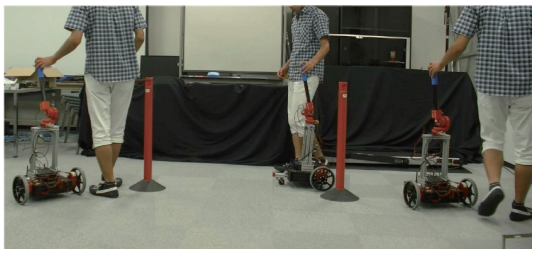

図2　8の字タスクの様子

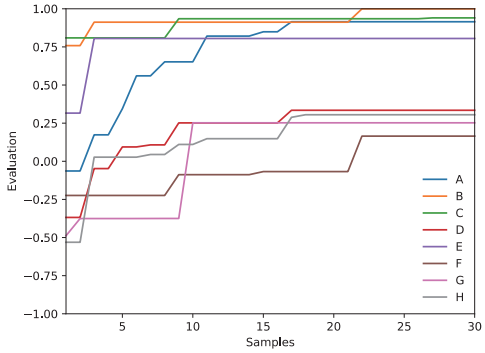

図3　ベイズ最適化の結果

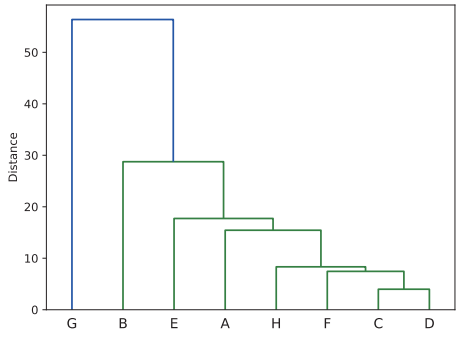

図 4　被験者の階層的クラスタリング結果

を図った。まず，人が試行中のどのような状況において良し悪しを評価していたのかを探るために，評価値と実験中の 19 次元の計測データ間の関係を重回帰分析により捉えた。ただし，実際には時系列要素の影響も想定できることから，時系列データに向けた拡張型の重回帰分析を用いた。得られた被験者ごとの偏回帰係数に対し，階層的クラスタリングを施した結果のデンドログラムを図 4 にまとめた。クラスタ数の自動決定手法を適用してみると，被験者 G のみが別クラスタとなり，その他は共通のクラスタに分類された。この結果から，今回対象とした，人と移動ロボット間の力覚的なインタラクションにおいては，多少の個人差はあるものの，概ね共通のパラメータで満足のいく操作感が得られることがわかった。すなわち，これらのパラメータは事前設計しておいて良く，その他に考えられる要素（仮想ばねの自然長や外力のポール部とモバイルベースへの分配法など）の変化を通じて，ロボットから人へ価値観を呈示する枠組みが望ましいといえる。

2. ロボットが求める人の挙動からの異常検知

基盤技術 1 の開発・検証を通じて，ロボットから人へ価値観を力覚ベースで呈示するには，例えば外力の分配法を状況に応じて調整することが肝要であることがわかった。そこで基盤技術 2 では，まずロボットが求める人の挙動を登録ユーザによる操作と定義し，非登録ユーザによる操作を異常と検知する機構を開発するとともに，異常検知の際に外力をモバイルベースに一切分配しないよう設計することで，非登録ユーザとみなされたことを力覚的に呈示する枠組みを開発する。

登録ユーザによる操作か否か判定するには，その操作履歴のパターンを適切に認識するような時系列データ処理技術が求められる。当初の計画では，これにリザーバコンピューティングと呼ばれる再帰型ニューラルネットワーク（RNN）の一種を活用する予定であったが，その特徴である出力層のみの学習では十分な予測・分類性能を発揮できないことが懸念された。そこで，先行研究で開発してきたリザーバコンピューティングの一種である少数階微分型のネットワークパラメータを一般的な RNN と同様に時間方向の誤差逆伝播で学習可能としたモデル RGaM を開発した。これは RNN のモデル構造を可能な限り簡素化した最新モデルである JANET の派生とみなせる設計となっている。

具体的には，RGaM は図 5 に示すニューロン（またはユニットと呼ぶ）構造を有する。JANET の開発時に指摘されたように，最も代表的で実際に優れた成果を収めている RNN モデルといえる LSTM の構造の中で最も重要な機構が入力ゲート機構，すなわち入力をどの程度内部状態に反映させるのか，である。JANET および RGaM はこの機構のみ LSTM から踏襲している。RGaM ではこれに加えて，少数階微分方程式を適切に近似することで得ら

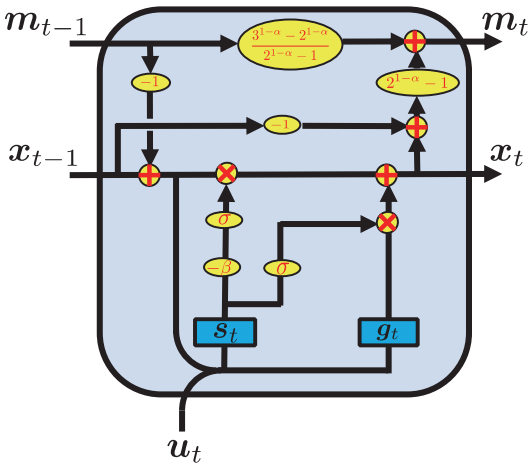

図 5　RGaM ユニットの構造

れる内部状態の履歴を保持する変数を持つ。これはLSTMにおける隠れ状態に対応しているとみなせるが，より具体的な履歴としての意味を持つことに注意したい。これを内部状態の更新に活用することでJANETよりも長期的な時系列データの表現に適したモデルとなっている。この履歴の保持期間を陰的に決定するパラメータは連続値で与えられていることから，誤差逆伝播で学習可能であることに注目し，ニューロンごとに個別に学習可能とする。これにより，対象とする時系列データに適した保持期間を自動的に決定できる。

前述の実験の被験者8人から明らかに傾向の異なった被験者Gを除き，新たに2人を加えた合計9人分のデータセットを用意した。その内の1人を登録ユーザと仮定して，登録・非登録ユーザの分類学習を交差エントロピー法により実施した。センサデータは合計26次元である。学習結果を図6にまとめる。開発したRGaMは全体的に従来のRNNモデルより優れた性能を示していることが確認できた。また，特に注目したいのが非登録ユーザを登録ユーザと誤認してしまったFN値であるが，RGaMでは0となっている。すなわち，RGaMを用いたロボットは，人に要求する挙動（登録ユーザによる挙動）を数値的に的確に表現することに成功したといえる。

このとき学習で得られたRGaMモデルを用いて，実際に非登録ユーザであることを力覚的に呈示する枠組みを開発し，そのデモンストレーション実験を行った。その結果，図7に示すように，登録ユーザがロボットを操作した場

		Model Prediction	
		Positive (Others)	Negative (Registered)
Actual classes	Positive (Others)	TP (True Positive)	FN (False Negative)
	Negative (Registered)	FP (False Positive)	TN (True Negative)

	Positive			Negative		
	LSTM	JANET	RGaM	LSTM	JANET	RGaM
Positive	87.8 (86.7-88.9)	86.30 (85.6-86.7)	**88.8** (88.8-88.8)	1.11 (0.00-2.22)	2.59 (2.22-3.33)	**0.00** (0.00-0.00)
Negative	4.81 (4.44-5.56)	5.19 (4.44-5.56)	**4.07** (2.22-5.56)	6.30 (5.56-6.67)	5.93 (5.56-6.67)	**7.04** (5.56-8.89)

図6　登録・非登録ユーザの分類結果

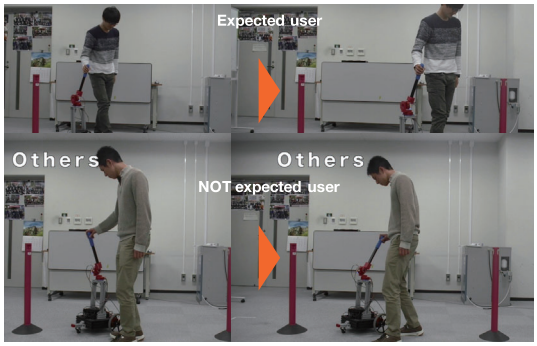

図7　ユーザ判定に基づく外力分配の切替

合には，適切にポール部とモバイルベースに外力が分配されて快適に操作できた一方で，非登録ユーザが登録ユーザとは異なる挙動で操作した結果，RGaMによる異常検知が働き，モバイルベースへの外力の分配を途絶えさせた。これにより，非登録ユーザにロボットの要求を力覚的に呈示することに成功したといえる。

また開発したRGaMのもう一つの利点として，省メモリであることが挙げられる。実際，上記の実験時のメモリ使用量はLSTMの約半分に抑えられており，制御用プログラムの使用量と比べると約10%となっていた。このため，計算資源が限定的な移動ロボットではRGaMは有効な時系列データ処理技術といえる。

［今後の研究の方向，課題］

本研究助成期間では，上記の2つの基盤技術を個別に開発を進めることができたものの，人・ロボットのインタラクションを扱うシステム全体としての開発は不完全に終わった。特に，当初の計画ではマルチエージェント強化学習を用いたインタラクションの最適化までを予定していたものの，今回の基盤技術1は強化学習における状態・行動空間の構築に，2は強化学習における報酬関数の設計に留まっているといえる。

今後は本助成期間で開発した基盤技術を基に人・ロボット強化学習環境を具体化し，より高水準な人・ロボット間の双方向性を有するイン

タラクション実現を目指す。また，そのときの課題として予想される，学習コストの低減に向けた強化学習技術の汎用的な改善，あるいは本問題に特化した手法の導出についても並行して進める予定である。

［成果の発表，論文等］

［1］ 小坂 麻人，小林 泰介，杉本 謙二：“車輪移動ロボットにおける力覚ベース操作の主観評価を尊重したベイズ最適化，”日本ロボット学会学術講演会，2C2-06，2020.

［2］ 馬渕 俊弥，小林 泰介，杉本 謙二：“RGaM：忘却ゲートと記憶履歴を持つ再帰型ユニット，”日本神経回路学会全国大会，P1-17，2020.