

平成 30 年 6 月 17 日現在

機関番号：34315

研究種目：基盤研究(B) (一般)

研究期間：2014～2017

課題番号：26280098

研究課題名(和文) 教示によって高速な全身運動を低速から逐次的に獲得する学習プロセスの計算論的理解

研究課題名(英文) Computational understanding of learning-by-demonstration process to acquire fast whole body motions sequentially from the low speed motions

研究代表者

玄 相昊 (Hyon, Sang-Ho)

立命館大学・理工学部・准教授

研究者番号：30344691

交付決定額(研究期間全体)：(直接経費) 12,300,000円

研究成果の概要(和文)：本研究では難易度の高い全身運動の教示方法を明らかにするため、強化学習による教示アルゴリズムを提案し、実際のヒューマノイドロボットを用いて検証することを目的とした。研究期間中にカルバック・ライブラー制御と呼ばれる確率最適制御理論に基づく新しい運動教示方法を考案した。提案手法の有効性は本プロジェクトで整備された双腕マニピュレータによって実証された。さらに、教師とロボットの共有潜在空間を利用することで低次元空間内で最適方策を高速に学習する方法を提案し、ヒューマノイドロボットのシミュレーションにおいて有効性を示した。

研究成果の概要(英文)：In this project, in order to clarify the teaching method of whole body movement with high degree of difficulty, we aimed to propose a teaching algorithm by reinforcement learning and to verify it using real humanoid robots. During the research period, a new teaching algorithm based on stochastic optimum control theory called Kullback-Leibler control was devised. We have experimentally proved the effectiveness of the proposed method using a dual-arm manipulator developed in this project. We also proposed to use a shared latent space between the human demonstrator and the robot, to quickly acquire the optimal control policy in the low-dimensional space. The method was also validated in dynamic simulations on a biped humanoid robot.

研究分野：ロボティクス

キーワード：ヒューマノイドロボット 運動制御 運動学習

## 1. 研究開始当初の背景

人はロボットのように教えられた通りの運動を正確に再現できないが、自らの身体と環境との相互作用の中で、教示された以上の複雑な全身運動を学習することができる。したがって、人はそもそも何を教示し、それをどのように全身運動の学習に活かしているか？教示は運動の上達の進度にどのように影響するか？という問いが自然に生じる。

近年、CG やゲーム技術の急速な発展に伴い、人の運動をロボットで実現する試みがなされている。ごく最近でも、環境とロボットの動力学モデルが既知であることを前提とし、モーションキャプチャデータや関節のキーポーズデータをそのままロボットの動力学に代入して運動指令(トルクや速度)を実時間で計算する試みがなされている。これに対し、本研究の興味の対象は、人やロボットが環境とのインタラクションによって徐々に運動を上達させる学習機構である。

我々は2008年から、学習器として、関節軌道の位相上の繰返し学習制御を提案し、高速なスクワット動作や、視覚を用いた野球バットイング学習への応用を試みていた。一方、2009年からは、予め与えられた多次元の関節時系列を運動ライブラリとして周期アトラクターや低次元空間に圧縮し、教示された運動から様々な似運動への補間・生成方法等について検討を行っていた。

さらに、ヒューマノイドロボット開発のノウハウを活用し、力制御による教示が可能で高速かつ軽量の油圧式ヒューマノイドロボットの開発を独自に行っていた。このロボットにモーションキャプチャーやビジョンセンサーを統合すれば、実際のハードウェアを用いて実世界の運動教示問題を研究する環境が整う。

以上の動機と準備のもと、本研究を申請するに至った。

## 2. 研究の目的

本研究では、難易度の高い全身運動の学習プロセスに教示がどのように関与するかを明らかにし、実際のヒューマノイドロボットに適用することによって検証することを目的とする。

教示は関節空間軌道やキーポーズで与えることができる。しかし、バランス維持のための軌道は他者から直接教示することはできず、学習者が自らの感覚を使って3次元空間における状態を認識し、その評価に基づいて運動指令を更新しなければならない。そこで本研究では、具体的に次の問題を実験を通じて明らかにする。

まず、教示データを活かした学習によって、高速全身運動をどの程度現実的な学習回数で獲得できるか、そして類似の運動にどの程度汎化的に学習の効果が活用できるかを明

らかにする。

次に、実際のヒューマノイドロボットを用いて、実環境における運動タスクを自らの視覚と力覚センサで評価し、教示データを用いた学習により、バランスのとれた高速全身運動がどの程度達成できるのかを定量的に明らかにする。

## 3. 研究の方法

- 全身ヒューマノイドロボットへの実装を最終ゴールに捉え、運動教示の理論的枠組みを検討し、アルゴリズムを開発する。
- 簡単なロボットモデルを用いて数値的に有効性を検証しながら徐々にスケールアップを図る。
- 学習の速度を高めるために、過去に実績のある低次元空間の中間表現を利用したアイデアを取り入れる。
- ロボットや運動計測システムの実験環境を並行して整備し、全身運動制御の方法を実装して最終実験に備える。

## 4. 研究成果

### (1)理論的な成果

理論面で特に大きな伸展があった。本研究においては、人から教示される運動を人型ロボットで実現する際に、必ず全身バランスを考慮する必要がある。この問題に対して従来は、「教示軌道の追従」と「バランス維持」の二つの目的を同時に満たす運動を学習、もしくは、最適化計算で求めるが、精度と計算量の面で実用には不十分であった。

この問題に対して本研究では、バランスを満たす状態空間の中で、教示軌道を追従するという発想に変えた。バランスを満たす空間を予め経験から獲得していなくてもはならないが、優先度が確保され、どのような教示軌道に対しても汎用的に適用可能となる可能性が高い。

様々な方法を検討する過程で2年次に進展があった。それは、近年発見された「カルバック・ライブラー制御(KL制御)」と呼ばれる確率最適制御理論に基づく、モデルベース強化学習による新しい教示方法である。KL制御とは終端の条件から最小化問題を時間の増加方向と逆向きに繰り返し解くことで各時刻の価値関数と最適方策を効率よく求める方法である。

特定のタスクにおけるロボットの状態遷移表現が行動と状態の対によって決まる確率分布で表現される(「制御ダイナミクス」という仮定をおき、そこから即時コストを表現する。即時コストの中には、制御ダイナミクスと別の状態遷移確率である自由ダイナミクスとの間の情報量である「KL距離」が含まれる。KL情報量の性質から、自由ダ

イナミクスにおいて遷移確率が存在しない状態遷移については、最適方策においても0となるため、実現不可能な遷移として事前除外できる利点がある。

終端時刻から初期時刻までの動作系列は、最初に価値関数を指数変換した「適合関数」を用いれば、ベルマンの最適性の原理は線形な更新式となり、解を高速に求めることができる。

本研究で提案された KL 制御を利用した教示方法は具体的には次の 2 ステップで構成される：

S1. 微小ノイズによってロボットを駆動し、動的バランスを満たす動作データのみを選定し収集する。そのデータから、ロボットの状態のマルコフ遷移モデルを同定する。

S2. 同定されたマルコフ遷移モデルを、“自由ダイナミクス”として、教示軌道との誤差に関するコスト関数を設計し、カルバック・ライブラー制御問題を解く。

各時刻のコスト関数は目標軌道誤差として設計する。また、状態は離散化し、状態遷移行列は動的バランスに関する制約条件(足先の接触点が床面から浮かないなど)を満たすようなものを用いる。これにより、獲得された最適方策を用いて、転倒しない範囲の教示動作を獲得することが可能となる。最終的に達成される目標状態を参照値として、下位レベルのコントローラでロボットの各関節を制御する。

最後に、本研究では3年次に、ヒューマノイドのような多自由度制御系への拡張方法を提案した。これは、教示者である人と、学習者であるロボットに類似タスクを行わせた際に現れる潜在共有空間をデータから抽出し(前処理)、それらの間の写像を利用して上記と同じ計算処理を低次元空間内で実行することにより、計算コストを大幅に削減する方法である。提案方法の有効性はヒューマノイドロボットのバランスを考慮した蹴り動作を動力学シミュレータで確認されたので、国際会議で発表した。

## (2)理論実装に関する成果

特にプロジェクト後半では提案手法を実機に適用するための検討を中心に行った。具体的には双腕アームと全身ヒューマノイドロボットへの適用方法を検討した。

前者はヒューマノイドロボットのバランスの制約を、それに似た制約条件に置き換えた場合の仮想実験であり、実機への適用可能性を検証するためのものである。具体的には、マニピュレータの手先の接触を維持しつつ、教示軌道に追従させるタスクを設定した。

マニピュレータ実機でノイズ駆動により手先の接触を維持するような状態遷移モデルを獲得し、新たに提示された教示軌道との距離を評価関数とした最適制御問題を解く

ことにより、本研究の目的である運動教示を達成した。この成果はロボット学会で発表した。

一方、後者への適用はシミュレーションに限定された。実ロボットに適用する際、低次元空間における状態遷移確率の取得のために、サンプリングの方法に工夫を要するため、現在、引き続き検討を行っている。この作業が完成すればシミュレーションで得られた結果を実ヒューマノイドロボットに実装することが可能と考えている。

## (3)実験システムに関する成果

本研究の成果を実証するための実験システム構築において大きな進展があった。まず、トルク制御可能な油圧駆動の双腕マニピュレータを開発し、欧文誌に掲載された。本研究の理論は主にこのロボットに実装され、様々な運動教示タスクにおける膨大な数の強化学習試行実験に有効活用された。

次に、ビジョンセンサーを搭載したロボットヘッドを構築し、FPGA によるパイプライン処理により高速に位置姿勢を計算することに成功した。これと並行して、FPGA を用いない視覚フィードバックによって両足で動的バランスを保つ実験に成功し、国内学会で論文発表した。

最後に、腕を除いたトルク制御型のヒューマノイドロボットハードウェアが完成し(油圧式としては世界で3番目)、英文誌に掲載された。

視覚フィードバックによる高速な運動制御は実装が間に合わなかったが、トルク制御による基本的なバランス制御を用いた不整地歩行実験に成功し(世界で3番目)、その論文を国際会議に投稿した。

## (4)まとめ

本研究の成果は4つである：

KL 制御を利用した新しい運動教示の枠組みとアルゴリズムを提案した。

実際のロボットマニピュレータの拘束条件付き教示タスクにおいて、提案手法の有効性を実証した。

提案手法を低次元潜在共有空間における学習に拡張し、多自由度の全身ヒューマノイドロボットのバランス制約付き運動教示シミュレーションにおいて有効性を確認した。

副次的な成果として、トルク制御型の等身大ヒューマノイドロボットが完成し、高度な運動学習の研究を推進する環境が整った。

これらの成果に関して国際会議2件、国内学会9件の学会発表を行った。また、アルゴリズム検証のために開発した双腕ロボットとヒューマノイドロボットについて英文ジャーナルペーパーに投稿し2件が掲載された。

今後、本研究のフォローアップとして、マニピュレータで得られた成果を論文にまとめ、ヒューマノイド実装に向けた研究を継続する予定である。

#### 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2件)

Sang-Ho Hyon, Daisuke Suewaka, Yuki Torii and Narifumi Oku, Design and experimental evaluation of a fast torque-controlled hydraulic humanoid robot, IEEE/ASME Transactions on Mechatronics, vol.22, no.2, pp.623-634,2017 (DOI: 10.1109/TMECH.2016.2628870) (査読有).

Kensuke Izawa, Sang-Ho Hyon, Prototyping force-controlled 3-DOF hydraulic arms for humanoid robots, Journal of Robotics and Mechatronics, vol.28, no.1, pp.95-103, 2016 (DOI: 10.20965/jrm.2016.p0095) (査読有).

[学会発表](計11件)

平山健太, 玄相昊, バランス制御実験のための簡易的なトルク制御型油圧ロボットの開発, 日本機械学会ロボティクス・メカトロニクス講演会, 2017.

Yuka Arika, Takamitsu Matsubara, Sang-Ho Hyon, Latent Kullback-Leibler control for dynamic imitation learning of whole-body behaviors in humanoid robots, IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids), 2016.

岡野雅史, 有木由香, 松原崇充, 玄相昊, 実ロボットによる KL 制御に基づく運動制御, 第 34 回日本ロボット学会学術講演会, 2016.

有木由香, 松原崇充, 玄相昊, 共有潜在空間を用いたカルバック・ライブラー動的見まね学習, 第 34 回日本ロボット学会学術講演会, 2016.

Sang-Ho Hyon, Daisuke Suewaka, Yuki Torii, Narifumi Oku, Hiroki Ishida, Development of a fast torque-controlled hydraulic humanoid robot that can balance compliantly, IEEE-RAS International Conference on Humanoid Robots, 2015.

薛咏, 玄相昊, トルク制御型ヒューマノイドロボットへのモデル予測制御の導入, 第 16 回計測自動制御学会システムインテグレーション部門講演会 (SI2015), 2015.

有木由香, 松原崇充, 玄相昊, カルバ

ック・ライブラー制御に基づく動的バランスを考慮した見まね学習法, 第 33 回日本ロボット学会学術講演会, 2015.

玄相昊, 永田光, 鳥居裕貴, 柔らかいヒューマノイドロボットにおける視覚フィードバック姿勢制御, ロボティクス・メカトロニクス講演会, 2015.

末若大輔, 井澤健祐, 北浦誠人, 玄相昊, 油圧駆動ヒューマノイドロボット RL-H2 の設計, ロボティクス・メカトロニクス講演会, 2015.

松本雅裕, 下ノ村和弘, FPGA によるロボットヘッドのリアルタイム位置姿勢計測の実装, デザインガイア 2015(電子情報通信学会 VLSI 設計技術研究会), 2015.

松本雅裕, 下ノ村和弘, ステレオビジョンを備えたロボットヘッドのリアルタイム位置姿勢推定, 日本機械学会ロボティクス・メカトロニクス講演会, 2015.

[その他]

ホームページ:

[www.humanoidsystems.org](http://www.humanoidsystems.org)

アウトリーチ活動:

ロボット動展示, 2015 国際ロボット展, 東京ビッグサイト, 2015/12/2-5.

#### 6. 研究組織

(1)研究代表者

玄 相昊 (HYON, Sang-Ho)

立命館大学・理工学部・准教授

研究者番号: 30344691

(2)研究分担者

松原 崇充 (MATSUBARA, Takamitsu)

奈良先端科学技術大学院大学・情報科学

研究科・准教授

研究者番号: 20508056

大塚 光雄 (OTSUKA, Mitsuo)

立命館大学・スポーツ健康科学部・助教

研究者番号: 20611312

下ノ村 和弘 (SHIMONOMURA, Kazuhiro)

立命館大学・理工学部・准教授

研究者番号: 80397679

有木 由香 (ARIKI, Yuka)

立命館大学・理工学部・助教

研究者番号: 80553239