

令和 4 年 6 月 20 日現在

機関番号：12601

研究種目：基盤研究(A) (一般)

研究期間：2018～2020

課題番号：18H04108

研究課題名(和文) 身体性共有と神経情報処理マッピングによる臨機応変な物体操作実現法

研究課題名(英文) Adaptive object handling by sharing embodiment and biological information

研究代表者

國吉 康夫 (KUNIYOSHI, Yasuo)

東京大学・大学院情報理工学系研究科・教授

研究者番号：10333444

交付決定額(研究期間全体)：(直接経費) 34,100,000円

研究成果の概要(和文)：深層模倣学習は、視覚を用いた物体操作を可能とするが、過学習の問題、特に無関係の情報の影響を受けやすいという問題がある。我々は、人がロボットを遠隔操作する際の視線情報を計測するプラットフォームを開発した。人の視線情報を深層模倣学習と組み合わせることで、タスクに無関係な情報に対してロバストな物体操作を実現することができる。Mixture density networkを用いて視線位置を再現し、その周辺の画像を切り出し、モータ出力の計算を行った。タスクに必要な情報を集中的に処理することで、針の糸通しやバナナの皮むきの際に臨機応変に状況判断することが可能となった。

研究成果の学術的意義や社会的意義

本研究は、自律ロボットの物体操作能力の向上のための基盤技術に関するものである。現状のロボット技術では、食材のように形状や固さなどのばらつきが多い対象物体はモデル化が困難であり、扱うことができなかった。我々は、深層模倣学習に着目した。人が直感的にロボットを遠隔操作するためのシステムと操作時の視線情報を計測するプラットフォームを開発した。人が物体操作時に重要な情報を持つ部分に注意を向けることを利用し、人の視線情報を模倣することで、複雑な環境や対象物体において、特に重要な情報のみを用いて模倣学習を行うことで、格段に性能向上可能であることを実証した。

研究成果の概要(英文)：Deep imitation learning enables the learning of complex visuomotor skills from raw pixel inputs. However, this approach suffers from the problem of overfitting to the training images. The neural network can easily be distracted by task-irrelevant objects. In this research, we use the human gaze measured by a head-mounted eye tracking device to discard task-irrelevant visual distractions. We propose a mixture-density network-based behavior cloning method that learns to imitate the human gaze. The model predicts gaze positions from raw pixel images and crops images around the predicted gazes. Only these cropped images are used to compute the output action. This cropping procedure can remove visual distractions because the gaze is rarely fixated on task-irrelevant objects. We evaluated our method on several manipulation tasks including handling multiple objects, needle threading, picking a small object, knot tying, and banana peeling.

研究分野：知能ロボティクス、機械学習、深層模倣学習

キーワード：ロボティクス 身体性 遠隔操作 深層模倣学習 視線計測

### 1. 研究開始当初の背景

人間の物体操作は、あらかじめ決められた動作の実行ではなく、状態変化に対して臨機応変に戦略を変えるダイナミックなプロセスである。このような能力をロボットで実現するために、単一動作ではなく、複数スキルを含む連続動作を学習する必要がある。われわれは、人の計測結果からスキルを抽出するために、動作の分散構造解析を行ってきたが、この手法は恣意的な動作の分節化やアライメントを必要とし、連続動作への拡張が困難であると言う課題があった。連続動作を扱うために、解析的なスキル抽出と人手による動作の分節化を排除し、動作全体の構造を直接ロボットに教示するフレームワークが必要である。本研究では、身体同型化による直接教示と生体情報を含む多相計測によって、連続動作スキル実現の基盤を確立する。

### 2. 研究の目的

本研究の目的は、物体操作における連続動作を人からロボットへ直接教示する方法を確立するとともに、大量の連続動作情報から臨機応変な戦略変更を再現する機械学習基盤を確立することである。

### 3. 研究の方法

従来研究の方法が連続動作の直接教示に適さない理由は、ロボットと人との間にある身体性の違いに起因すると考えられることから、教示データの質を向上することがもっとも重要な課題であると考えられる。本研究では、ロボットを人に近づけるのではなく、人をロボットに近づけるアプローチをとる。人は、高い機能性を有する身体を持つが、一方で、箸やトングなどの道具を介しても器用な動作を実現可能である。このように機能が制限された中でも適応的に振る舞う人の器用さを利用して、人の身体構造に制約を加える。具体的には、ロボットのようなリンク構造を持つメカニカル計測装置と、幾何学的に完全同型なロボットを用いることで、感覚情報の損失を最小限にするとともに、人の動作データを関節角レベルで一对一にロボットへ直接対応付け可能とする。これによって、教示データの質を最大限高める。結果として、人の動作をロボット動作へ変換する複雑なプロセスを排除し、変換に必要な動作の分節化と人手によるスキル抽出を不要とする。教示データの質を高めることで、動作意図や感覚運動の因果構造を自動抽出するという知的動作実現のための最重要課題を、最適な条件で取り組む。

連続動作から注意や動作の切り替え、感覚運動の因果関係を抽出するために、動作計測時には、視覚、力覚などのデータを同時計測する。このとき、視覚情報はロボット側カメラによって記録し、人には Head mount display を用いて与える。このようにすることで、人が物体操作するために必要な情報のすべてを多相記録可能となる。

人の注意の切り替えに関する情報を機械学習に用いることを考える。当初神経活動を機械学習に利用することを計画したが、視線情報が有効であることがわかったため、物体操作時の視線情報を計測可能な Head mount display を用いた開発を行った。人の視線は、物体操作時には必要な情報以外には稀にしか注視しないことが知られているために、視線情報をカメラ情報から推定する視線模倣と、注視領域以外を除くことで、タスクに不要な環境情報に対してロバストな深層模倣学習が可能であると考えた。

### 4. 研究成果

#### 4-1 メカニカルグローブと同型ロボットの開発

ロボットと同じ運動学パラメータを持つ機械式の姿勢計測システムの開発を行った。ロボットハンド部分は、計画通りトング型と多指型の開発を行った。多指型とトング型に関して、人が操作を行い、複数の物体操作タスクにおいて成功率と実現時間を評価する被験者実験を行った。結果としてほとんどのタスクにおいて有意な差が得られないという結果が得られた。多指型は自由度が多いため、一見するとよりタスクのパフォーマンスが上がると思っていたが、指先の滑り感覚などが硬い身体のために阻害されることと、自由度が多いことでかえって滑りやすいなどの問題があった。これらの実験結果に基づいて、模倣実験においてはトング型を活用することとした。

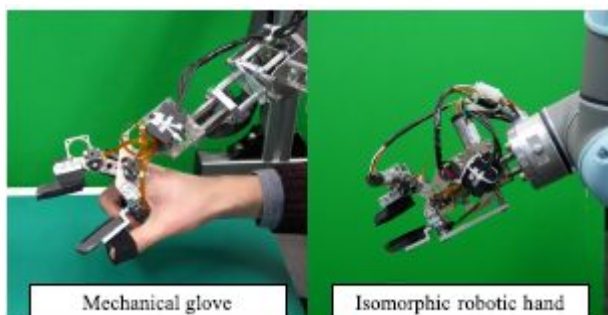


図. 開発したメカニカルグローブと同型ロボットハンド

#### 4-2 視線情報を活用した深層模倣学習

視覚情報を用いた深層模倣学習において、タスクと関係ない画像情報の変化に対する過学習を防ぐために、人が物体操作中にどの領域に目を向けているかを視線計測装置が内蔵された Head mount display を用いて計測した。開発した同型システムを用いてロボットを遠隔操縦し、関節運動とカメラ情報および視線情報の同時計測を行った。

人の視線情報は、タスク実行時には関係のある物体を注視することが知られている。実験結果から、対象物体、ロボットハンドなどの複数箇所に順次目を向けることがわかった。タスク似重要な領域を用いて運動出力を学習するためには、カメラ情報に対して人が注視する領域を複数抽出する必要がある。従来の視線予測のように注視領域の時間変化はそれほど重要ではない。そのため、カメラ情報から複数の注視領域を確率分布として出力可能な Mixture density network を用いて視線情報を推定した。

この推定結果を用いて、注視領域の画像を切り出し、この画像のみを用いて運動出力を学習した。この方法を用いて、複数物体が画面上にある場合に、順番に pick-and-place を行うタスクについて、視線を用いない方法と比較を行った。



図：複数物体の Pick-and-Place

	Sub-task	Mean	Best
Baseline	Apple Pick	13.9 (5.56)%	22%
	Apple Place	8.33 (5.56)%	11.1%
	Orange Pick	0.00 (0.00)%	0.00%
	Orange Place	0.00 (0.00)%	0.00%
Proposed method	Apple Pick	91.7 (10.6)%	100%
	Apple Place	91.7 (10.6)%	100%
	Orange Pick	91.7 (10.6)%	100%
	Orange Place	72.2 (42.1)%	100%

表：従来手法と提案手法の比較

視線を用いない場合と比較して、格段の性能向上が見られた。注目すべき点は、従来手法の場合、複数物体がある場合に Pick-and-Place という簡単な動作でさえも成功率がとても低いということである。一方で、提案手法では、適切に注視点領域を切り替えることで、タスクに無関係な情報を除去し、高い精度を実現している。一般に、困難とされている Long-horizon task は複数の物体を含むことが多い。物体が複数ある場面において、物体の配置のバリエーションは物体の数に応じて非常に多くなる。そのため、複数の物体が画面上に存在するだけでタスクが困難になるということが本研究より明らかになった。このようなときに注視メカニズムを導入することが極めて有効であることも同時に示された。

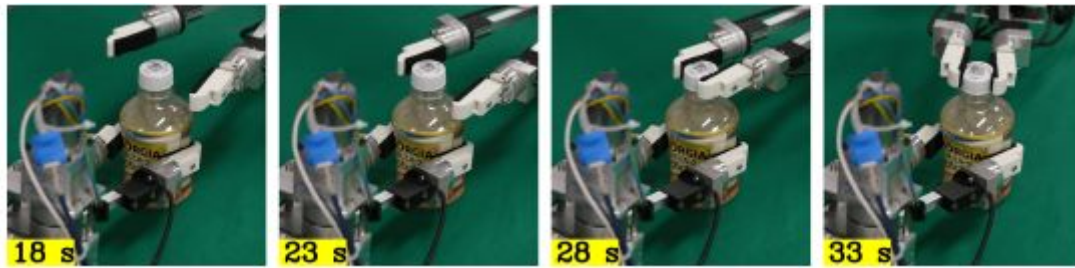
#### 4-3 力覚情報の利用

我々のシステムは、力の情報も同時に計測することができるという点が特徴的である。また、同型システムは遠隔操縦でロボットを操縦するのではなく、ロボットと同型の姿勢計測装置を直接操作することで、運動情報、手元のカメラ情報、力覚情報、視線情報の同時計測が可能である。

力制御が不可欠なタスクの場合、遠隔操縦やシミュレーション、VR システムなどでは模倣データを作成することが困難である。これに対して、バイラテラル装置などの高価な装置を利用す

る方法も考えられるが、我々は新しく同型システムを活用する方法を提案した。

同型システムを用いて直接物体操作したときのデータを用いて深層模倣学習することで、ペットボトルのフタを開けるなどの力制御が必要な動作を実現した。



図：ロボットによるペットボトルのフタ開け

#### 4-4 2重解像度による精密動作の実現

人の目は、中心が高解像度となっており、周辺視野の解像度は低い。そして、精密動作では中心視野の高解像度領域を利用することが知られている。この知見を利用してロボットによる精密動作の実現を行った。

一般にカメラが高解像度であるほど精密な動作が可能になるが、高い解像度は情報量が大きく過学習の原因となる。そのため必要な模倣データが膨大になるという問題があり、精密動作を模倣学習で実現することはできていなかった。

高解像度カメラ情報と視線予測をもちいることで、中心視野領域のみ高解像度の画像を切り出してモータ出力に用いる。精密動作が不要な場合には、周辺視野領域の低解像度の画像を用いてモータ出力を学習する。このような2重解像度システムによって、針の穴に糸を通すなどの精密動作が可能となった。



図：ロボットによる糸通し

図に示すように、糸は柔軟で時折変形することがある。このような場合も適切に状況判断を行い自律で糸通しが可能であることが示された。このような制御はモデルベースの制御では困難である。また、高解像度カメラ情報を用いて針の穴に糸が通っているか自動識別し、失敗したときにはやりなおすなどの臨機応変な行動を可能とした。

## 5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 2件/うちオープンアクセス 1件）

1. 著者名 金井嵩幸, 大村 吉幸, 長久保 晶彦, 國吉 康夫	4. 巻 39
2. 論文標題 メカニカルグローブによるロボットハンドデザインの第三者評価	5. 発行年 2021年
3. 雑誌名 日本ロボット学会誌	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kim Heecheol, Ohmura Yoshiyuki, Kuniyoshi Yasuo	4. 巻 6
2. 論文標題 Gaze-Based Dual Resolution Deep Imitation Learning for High-Precision Dexterous Robot Manipulation	5. 発行年 2021年
3. 雑誌名 IEEE Robotics and Automation Letters	6. 最初と最後の頁 1630 ~ 1637
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/LRA.2021.3059619	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Kim Heecheol, Ohmura Yoshiyuki, Kuniyoshi Yasuo	4. 巻 5
2. 論文標題 Using Human Gaze to Improve Robustness Against Irrelevant Objects in Robot Manipulation Tasks	5. 発行年 2020年
3. 雑誌名 IEEE Robotics and Automation Letters	6. 最初と最後の頁 4415 ~ 4422
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/LRA.2020.2998410	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Wakatabe Ryo, Morita Kohei, Cheng Gordon, Kuniyoshi Yasuo	4. 巻 -
2. 論文標題 Efficient Event-Driven Forward Kinematics of Open Kinematic Chains with $O(\log n)$ Complexity	5. 発行年 2018年
3. 雑誌名 2018 IEEE International Conference on Robotics and Automation (ICRA)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICRA.2018.8461211	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Keiko Fujii, Yoshiyuki Ohmura, Yasuo Kuniyoshi	4. 巻 -
2. 論文標題 Synaptic excitatory-inhibitory balance affect information integration via attractor dynamics	5. 発行年 2018年
3. 雑誌名 Association for the Scientific Study of Consciousness (ASSC2018)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yoshida Taketo, Kuniyoshi Yasuo	4. 巻 -
2. 論文標題 Adversarial Imitation Learning between Agents with Different Numbers of State Dimensions	5. 発行年 2019年
3. 雑誌名 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/AIKE.2019.00040	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Sekiya Kento, Ohmura Yoshiyuki, Kuniyoshi Yasuo	4. 巻 -
2. 論文標題 Generating an image of an object 's appearance from somatosensory information during haptic exploration	5. 発行年 2019年
3. 雑誌名 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)	6. 最初と最後の頁 8132-8137
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IROS40897.2019.8967795	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計13件 (うち招待講演 0件 / うち国際学会 4件)

1. 発表者名 Heecheol Kim, Yoshiyuki Ohmura, Yasuo Kuniyoshi
2. 発表標題 Transformer-based deep imitation learning for dual-arm robot manipulation
3. 学会等名 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (国際学会)
4. 発表年 2021年

1. 発表者名 Takayuki Komatsu, Yoshiyuki Ohmura, and Yasuo Kuniyoshi
2. 発表標題 Unsupervised temporal segmentation using models that discriminate between demonstrations and unintentional actions
3. 学会等名 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (国際学会)
4. 発表年 2021年

1. 発表者名 Takumi Takada, Yoshiyuki Ohmura, and Yasuo Kuniyoshi
2. 発表標題 Unsupervised Learning of shape-invariant Lie group transformer by embedding ordinary differential equation
3. 学会等名 2021 IEEE International Conference on Development and Learning (国際学会)
4. 発表年 2021年

1. 発表者名 Kento Sekiya, Yoshiyuki Ohmura, Yasuo Kuniyoshi
2. 発表標題 Learning to grasp multiple objects with a robot hand using tactile information
3. 学会等名 第26回ロボティクスシンポジウム
4. 発表年 2021年

1. 発表者名 K. Sekiya, Y. Ohmura, Y. Kuniyoshi
2. 発表標題 Generating an image of an object's appearance from somatosensory information during haptic exploration.
3. 学会等名 2019 IEEE/RSJ International Conference on Intelligent Robots and System (IROS) (国際学会)
4. 発表年 2019年

1. 発表者名 関谷研人、大村吉幸、國吉康夫
2. 発表標題 体性感覚情報に基づく手探り中の物体の外観の生成
3. 学会等名 ロボティクス・メカトロニクス講演会2019 (Robomech2019)
4. 発表年 2019年

1. 発表者名 伊藤龍一郎、金井嵩幸、大村吉幸、新山龍馬、國吉康夫
2. 発表標題 LSTMによる物体操作時の柔軟物変形予測
3. 学会等名 ロボティクス・メカトロニクス講演会2019 (Robomech2019)
4. 発表年 2019年

1. 発表者名 狩野泉実、田中一敏、新山龍馬、國吉康夫
2. 発表標題 報酬寄与率を考慮したパラメータノイズによる深層強化学習の探索と活用の調節
3. 学会等名 ロボティクス・メカトロニクス講演会2018 (Robomech2018)
4. 発表年 2018年

1. 発表者名 荻島諒也、米倉将吾、國吉康夫
2. 発表標題 自由エネルギー原理による生成モデル理解と環境認識に基づく適応的行動
3. 学会等名 ロボティクス・メカトロニクス講演会2018 (Robomech2018)
4. 発表年 2018年



1. 発表者名 伊藤龍一郎, 金井嵩幸, 大村吉幸, 新山龍馬, 國吉康夫
2. 発表標題 LSTMによる物体操作時の柔軟物変形予測
3. 学会等名 ロボティクス・メカトロニクス講演会2019 (Robomech2019)
4. 発表年 2019年

1. 発表者名 関谷研人, 大村吉幸, 國吉康夫
2. 発表標題 体性感覚情報に基づく手探り中の物体の外観の生成
3. 学会等名 ロボティクス・メカトロニクス講演会2019 (Robomech2019)
4. 発表年 2019年

1. 発表者名 皿海孝典, 狩野泉実, 國吉康夫
2. 発表標題 強化学習における報酬なしスキル獲得の階層化
3. 学会等名 第33回人工知能学会全国大会 (JSAI2019)
4. 発表年 2019年

1. 発表者名 小松 高歩, 金 東敏, 鈴木 裕真, 國吉 康夫
2. 発表標題 多自由度ロボットアームを用いた物体性質の自発的獲得
3. 学会等名 日本発達神経科学学会第8回学術集会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

國吉・新山研究室ホームページ  
<http://www.isi.imi.i.u-tokyo.ac.jp/?lang=ja>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	長久保 晶彦  (NAGAKUBO Akihiko)  (00357617)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員    (82626)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------