

令和 6 年 6 月 9 日現在

機関番号：32689

研究種目：若手研究

研究期間：2022～2023

課題番号：22K17979

研究課題名（和文）深層学習による複数触覚センサの統合を用いた多指ロボットハンドによる操り動作の実現

研究課題名（英文）Multi-Fingered In-Hand Manipulation Based on Multi-Type Tactile Information Using Deep Learning

研究代表者

船橋 賢（Satoshi, Funabashi）

早稲田大学・理工学術院・講師

研究者番号：00905151

交付決定額（研究期間全体）：（直接経費） 3,500,000円

研究成果の概要（和文）：本研究では、多指ロボットハンドの掌全体を使った巧みかつ安定した操り動作を実現するために、複数種類の触覚センサからの情報の統合とそれによる動作生成を目指した。まずは深層学習モデルの調査をした。触覚情報と関節情報を平面的に注意するように学習でき、未知の把持位置や物体との特性に合わせて適応的に動作を変えることに成功した。次にロボットハンドによるより器用な操り動作のために高分解能の光学式触覚センサを開発・改良した。人間の指先の面接触形状を取り入れ、操りが容易な指先形状でありながら微細な把持状態を取得できることを確認した。現在は、複数種類の触覚センサを統合したことによる繊細な動作生成を行っている。

研究成果の学術的意義や社会的意義

本研究では、各触覚センサの特徴や限界を逆手に取り、人の手上的固有受容器の分布を模倣するため、光学式触覚センサのような高い解像度を持つがスペースを取るセンサを指先に、センシング点は少ないが薄く容易に貼り付けられる磁気式の触覚センサを指腹や掌に搭載した。さらに重要なこととして、今まで深層学習は画像や関節角度などの情報をマルチモーダルに処理してきた実績があるが、触覚単体でも、様々なモダリティが混在する難易度の高い情報であり、しかしながら人間のように巧みな操りを行う上では欠かせない観点と考えた。これにより人間の作業を行うことを目指すことで学術的意義のみならず社会的意義も見出そうとした。

研究成果の概要（英文）：In the first year, we proposed a model that combines Transformer and LSTM, applying it to the tactile sensors and joint angle information of the multi-fingered hand. As a result, the model learned to focus on important modalities such as tactile and joint information, successfully adapting its motions to unknown grasping positions and object characteristics. In the second year, to achieve more dexterous manipulation with the multi-fingered hand, we developed and improved high-resolution optical tactile sensors. Human fingertips are characterized by a near-spherical shape at the tip, while the shape of the finger pads becomes flatter when closer to the first joint. We confirmed that the fingertip shape is easy to manipulate while still capturing fine-grained grasping states.

In the next six months, we will advance the research on a motion generation model that integrates the developed deep learning model and optical tactile sensors.

研究分野：知能機械学

キーワード：触覚センサ 多指ロボットハンド 深層学習

様式 C-19、F-19-1 (共通)

1. 研究開始当初の背景

少子高齢社会ではロボットが家庭や工場などで介護や協働などの作業支援を行うために人間が扱う様々な道具・物体を人間のように巧みに扱える必要がある。そのためには、人間のように物体との接触状態を常に検知できる触覚センサを搭載する多指ロボットハンドが必要である。多指ハンドの巧みさはその操り動作の多様さにある。それらは指先で物体を操る時の把持姿勢である「精密把握」と指腹や掌で握りこむ時の把持姿勢である「握力把握」と大きく分けられる。それぞれの把持姿勢において適切な把持力で操りを行うために触覚センサ情報が必要となる。人間の皮膚で考えると、圧力や振動など触覚を感じる器官として様々な機械受容野が存在する。例えば、指先にはサイズの小さいメルケル細胞が多数あり、指腹や掌にはサイズの大きいフィニ終末が広がっている傾向がある。このように部位によって異なるセンサ(機械受容野)が分布しており、人間はこれらの様々な触覚情報を統合的に処理し、結果として指先で精密な操りを、指腹・掌で安定した把持を行うことが出来ている。また、ドライバーでネジを回す、箸で物体を掴むといったタスクでは指先と指腹・掌を同時に使い、巧みに物体を操ることが出来る。ロボットハンドにおいてもこのような複数の触覚利用とその情報の統合的処理という方向性は人間が行っているより複雑なタスクの達成を目指すという点において非常に重要と考えられる。近年では深層学習や触覚センサによる多指ハンドの操りの研究が盛んになっているが、部位の違いによる複数の触覚センサの同時利用とその情報処理の統合方法は明らかになっていない。図1に示すようなロボットプラットフォームを用いて研究を進めた。

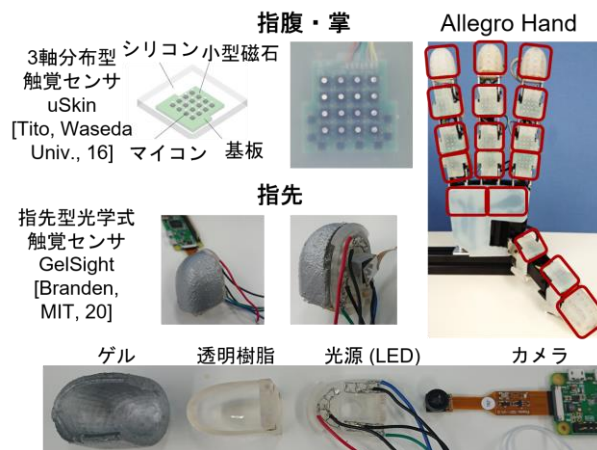


図1 Multi-fingered hand and two different tactile sensors (GelSight and uSkin).

2. 研究の目的

本研究提案では大きく二つの方向性で研究を進めた。1つ目は多量の触覚情報を処理しつつ、従来できなかった複雑な触覚情報を判断しながらでの多指ハンドの動作生成モデルの提案を行った。2つ目に、より巧みな操り動作生成を目指した光学式触覚センサのデザイン改善の研究を行った。当初の目的としては、初めから触覚情報を統合した動作生成であったが、各提案手法を改善することで、より確度の高い研究を進めることが出来ると判断した。下記に小項目を2つに分け、各研究の概要を示す。

2-(1).

従来研究では、指と掌に貼られた触覚センサをグラフ構造に見立てて GCN で処理することで指定された特性の物体に合わせた操り動作を実現している [1]. しかし、初期把持位置のズレや対象物体の特性を指定する必要があった事による未知物体の特性や姿勢などの未学習環境適応への限界があった。一方、触覚センサの情報を時系列方向に処理することで未知物体の内的特性を推測し滑り予測に成功した例や attention を用いて動作注目点を把握しながらロボットアーム2本の複雑な操り動作を成功させた例がある。そこで、本研究では、触覚センサを搭載した多指ロボットハンド、Allegro Handを使用して、未知の物体を認識し、把持位置のズレに対応した適応的な操作を生成することを目的とする。これらの点を踏まえた従来では行えなかった難度の高いタスクとして、離れた位置にあるランダムな姿勢の未知物体を触覚情報を用いて手の中に手繰り寄せる動作が考えられる。本実験ではその操り動作の実現を目指した。

2-(2).

従来の光学式触覚センサはそのセンシング面が平面形状をしており、これをそのままロボットハンドに組み込んで複雑な操りを行うには、ある程度の限界が存在した。人間の指の具体的な特徴として、指先の部分は球に近い形状だが、第一関節に近づくにつれて指の腹の形状は平たくなっている。この形状は一般的に面接触形状として知られるが、このメリットとして、指の角度により接触面積が変化することで、指と物体の間にはたらく圧力を調整できることが挙げられる。そのため安定した把持やダイナミックな操作が可能となる。先行研究において光学式触覚センサのセンシング面を人間の指先のように立体化する試みが行われたものの、人間の指の複雑な形状を完全に再現しているわけではない。そこで本研究では、人間の指先の幾何学的特徴をより詳細に模倣した面接触形状の光学式触覚センサの開発と、その形状の評価を目的とした。

3. 研究の方法

2. 同様、各研究内容の提案手法及びその評価方法を記載した。

3-(1).

・提案手法

本提案モデルは、図2に示すようなViT部とLSTM部とデコーダ部で構成される。

ある時刻tの触覚情報をViTで処理し、抽出した特徴量を時刻tの関節角度情報と結合したのちにLSTM部とエンコーダ部に入れることで時刻t+1の触覚情報と関節情報とリーダーの関節情報を予測する。各部詳細について以下に説明する。

Vision Transformer (ViT)

ViTは、画像処理の分野で有効性が示されている手法であり画像情報をパッチに分割して、各パッチと学習可能なCLSと呼ばれるパラメータをトークンとしてTransformerのエンコーダモデルに入力して、各タスクに固有な情報を予測することで学習できる。本実験では、この手法を触覚に応用する。図1のようにAllegroHandに搭載された触覚を18トークンに分割する。その後、従来のViTでは入力トークンに対して重み共有した線形変換を加えるが、本研究では以下の式(1)のように固有の線形変換を加える。この処理は、各触覚センサの特性の差を補正する働きを持つ。

$$z_0 = [x_{class}; x_p^1 E_p^1; x_p^2 E_p^2; \dots; x_p^{18} E_p^{18}] + E_{pos} \quad (1)$$

以上の線形変換を加えたのちは、3層のViT encoderで処理し、動作において重要な触覚センサの潜在表現をCLSトークンで抽出する。

LSTMとデコーダ部

未知物体の特徴を自律的に認識するには、物体の触覚情報の時系列的な推移を処理することで、頑健な特性認識ができる。本実験ではロボット分野の動作生成で広く使われるlong-short term memory(LSTM)を用いることで未知物体の特性情報や姿勢角を認識し、それに紐づいた動作と将来のセンサ値の予測に重要な隠れ状態を出力する。LSTMで隠れ状態を予測後、LeakyReLUを活性化関数としたデコーダ部(FCN)で時刻t+1の触覚情報やリーダー関節情報や関節情報を予測する。予測した時刻t+1のリーダー関節情報をAllegro Handに入力することでロボットを動作させる。

・実験設定

本研究では本手法の有効性を示すために以下の(a)~(c)の実験を行う。物体を手繰り寄せる動作を実行した後に、Allegro Handとテーブルを離れた際にハンドが物体を落とさずにしっかりと握りこんでいる状態である事を成功条件として成功確率を評価した。

- (a) 提案モデルとViTなしとLSTMなしのモデルの比較
- (b) 未知姿勢に物体に対する動作評価
- (c) 未知物体に対する動作比較

実験(a)は、提案手法であるViT-LSTMに含まれるViTとLSTMの有効性を評価する目的で行う。実験(b)では未知教示姿勢で置かれた物体に対する提案手法の動作のロバスト性を評価した。実験では長方形の箱について45°と90°と135°以外にも68°と113°の場合の動作の成功確率を評価した。

実験(c)では未教示物体に対する手繰り寄せ動作の成功回数を評価する。未教示物体としてサイコロ、大きいボール、ポテトチップスの缶、長いスポンジを採用した。

3-(2).

・提案手法

人型の指形状を取り込んだ光学式触覚センサを図3に示すパーツ構成で作成する。なお、シリコンの外側(カメラの反対側)の面には金属皮膜塗装を施しておく。LED光がシリコン内でランバート反射する。次にシリコンの表面に物体を接触させると、シリコンが弾性変形を起こす。このとき、シリコン内のLED光が乱反射し、カメラがその光線を捉えることで接触を知覚する。シリコンの変形の様子を直接的にカメラで撮影しているため高精細な触覚情報が得られる。

・実験設定

3.1. 角度に応じた接触面積の測定

点接触形状の指と面接触形状の指それぞれを図4のピン支持型カンチレバーの先端に装着す

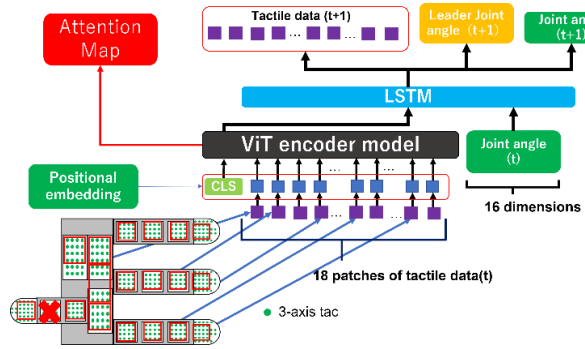


図2 Proposal model (ViT-LSTM)

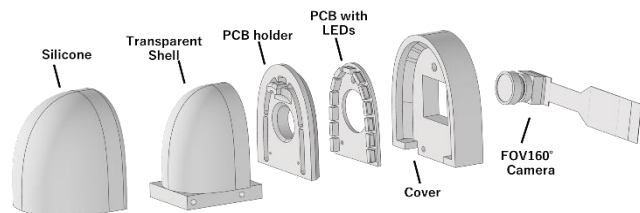


図3 Structure of the Human-Finger-Shaped(HFS) optical sensor.

る。面積測定の方法として、装置先端の角度と印加荷重を変えながら、指先をアクリル板に押し付けることで、各条件における接触面積を測定する。なお設置面積の測定はカメラ画像のピクセル数をカウントする手法を取るが、正確さを期すため、センサ内から撮影した画像でなく、アクリル板の反対側から高解像度の外部カメラで撮影した画像を用いる。

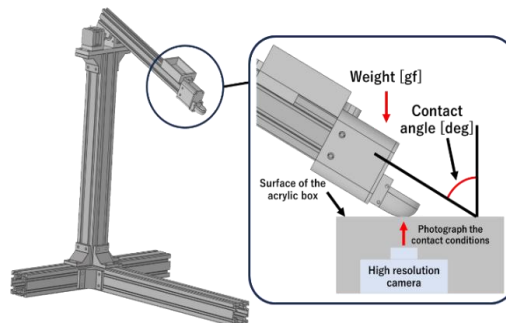


図4 Mechanism of the contact area measuring system.

3.2. マニピュレーション安定性の比較

Allegro Handの親指と人差し指を同一の指形状に設定して、筒状物体を回転する動作を行う。そして各形状の指で10回ずつ回転させたときの成功回数を比較することで、マニピュレーションの安定性を比較した。

4. 研究成果

4-1.

提案モデルとViTなしとLSTMなしのモデルの比較

表1のように提案モデルの把持の成功確率は95%となり、他のモデルよりも高くなった。本実験から提案手法におけるViTとLSTMの有効性が示された。

表1 Motion achievement of each model (1)

object	Box			Pet Bottle Bag		
	45°	90°	135°	45°	90°	135°
ViT-LSTM	5/5	3/5	5/5	5/5	3/5	5/5
LSTM only	2/5	0/5	5/5	1/5	0/5	1/5
ViT only	0/5	0/5	0/5	0/5	0/5	0/5

未知物体に対する動作比較

以下の表2は未知物体に対するモデルの引き込む時の動作成功確率を表している。表2から未知物体に対する成功確率は75%と高い確率で未知物体を認識して手繰り寄せることができると示された。

表2 Motion achievement of unknown object (1)

Object	Chips Can			Long Sponge		
	45°	90°	135°	45°	90°	135°
ViT-LSTM	3/4	4/4	1/4	4/4	4/4	2/4

Object	Large Ball	Dice	Success Rate
ViT-LSTM	3/4	5/5	75%

訓練されていない物体に対する手繰り寄せ動作は高い成功率で達成できたが、ポテトチップスの缶が135°に置かれたときと長いスポンジが135°に置かれたときの場合は動作成功率が低かった。ポテトチップスの缶では、物体の姿勢に合わせて正しい方法で物体を手繰り寄せることができたが、重心が手から離れすぎていたため、物体を掴んで持ち上げた後、物体が傾いて落下した。長いスポンジを引き込む際は、芝生との摩擦が強いため、物体を手繰り寄せることができなかった。今後の課題として、より多くの条件、すなわち摩擦力や物体の重心位置を考慮した把持戦略を学習できるようにすることが必要であると考えられる。

4-2.

角度に応じた接触面積の測定

それぞれの指形状について、いくつかの荷重条件(200g~1000g)を例に、接触角度を変化させた際の接触面積の変化を確認した。接触角度が小さい(=接触位置が指先に近い)ときは接触面積にあまり差はないが、接触角度を大きくしていくと、50°周辺から面接触形状の指の接触面積が急激に増加した。また、この傾向がいずれの荷重条件においても発生したことを確認した。

次に、面接触形状の指の接触面積を、点接触形状のそれと比した際の増加率を各角度と荷重ごとに1つのグラフにまとめたのが図5である。このグラフから、同じ接触角度で比較した際は荷重が大きいほど面積の増加率が小さくなる傾向にあることが新たに分かる。接触角度を80°としたとき、200gf印加時の接触面積増加率は67%程度であるが、1000gf印加時の接触面積増加率は42%程度にとどまっている。本研究で用いるシリコンの厚みは3mm程度であり、図3から分かるように、センサの構造的にシリコンの直下には硬い透明殻がある。そのため、厚さ3mmのシリコンを剛体で挟み込んで荷重を加えていることに等しいため、弾性限度に近づいているとみられる。荷重の増加に伴って面接触の優位性が小さくなる傾向は、1000gfを上回る荷重を印加させた際にもみられると考えられる。しかし一般的にロボットハンドで物体の操作をしたり、物体の表面テクスチャ取得が必要となる場面において、1指の先端に1000gfを上回る荷重が印加される状況は少ない。そのため、センサの実用上において弾性限度を考慮する必要はないといえる。

マニピュレーション安定性の比較

面接触形状の指では、5回以上連続的に回転動作を行っても物体を落とすことがなかったほどの安定性を示した。点接触形状の指で回転させた後は筒状物体が落下しそうで不安定になっていることが多かった。この結果は、提案した指先が操作性の向上に有効であることを示せた。

4-(3).

上記に示す通り、本研究提案では結果として2つの方向性の研究を行い、動作生成モデルと触覚センサの提案・評価を行った。これらの研究は、複数の触覚情報をどう統合して動作生成を目指すかという方向性に対し、重要なステップとして執り行われ、国内論文として発表されている[2][3]。国際的にも本提案者がAI系トップの国際会議で招待講演をするなど、注目を得られていると考える[4]。

今後の展望・現在の研究状況として、開発してきた深層学習モデルと光学式触覚センサを統合した動作生成モデルの研究を進め、複数種類の触覚センサを統合したことによる物体の状態を繊細に把握し、巧みな動作生成を行えるよう研究を進めている。

<引用文献>

1. Satoshi Funabashi, Tomoki Isobe, Fei Hongyi Atsumu Hiramoto, Alexander Schmitz, Shigeki Sugano and Tetsuya Ogata, "Multi-Fingered In-Hand Manipulation with Various Object Properties Using Graph Convolutional Networks and Distributed Tactile Sensors," in the IEEE Robotics and Automation Letters (RA-L), vol. 7, no. 2, pp. 2102-2109, April 2022, DOI: 10.1109/LRA.2022.3142417.
2. 上野貴久, 船橋賢, 伊藤洋, Alexander Schmitz, 尾形哲也, 菅野重樹, "時系列情報とAttentionを用いた未知物体のロバストな手繰り寄せ把持動作の安定化手法の提案", 日本機械学会, ROBOMECH2023, 1A2-F20.
3. 山田泰, 船橋賢, Alexander Schmitz, 尾形哲也, 菅野重樹, "器用な操り動作のための人の指形状を模した光学式センサ", 日本機械学会, ROBOMECH2024, 1P1-Q09.
4. Satoshi Funabashi, "Hand Morphology from Tactile Sensing with Spatial Deep Learning for Dexterous Tasks", NeurIPS | 2023 Thirty-seventh Conference on Neural Information Processing Systems, 1st Workshop on Touch Processing: a new Sensing Modality for AI.

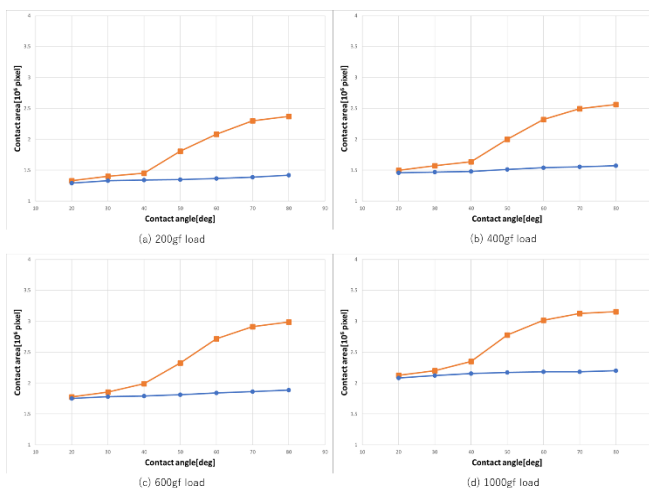


図5 Contact angle and contact area by some load. Orange(square) is surface contact shape and blue(circle) is point contact shape.

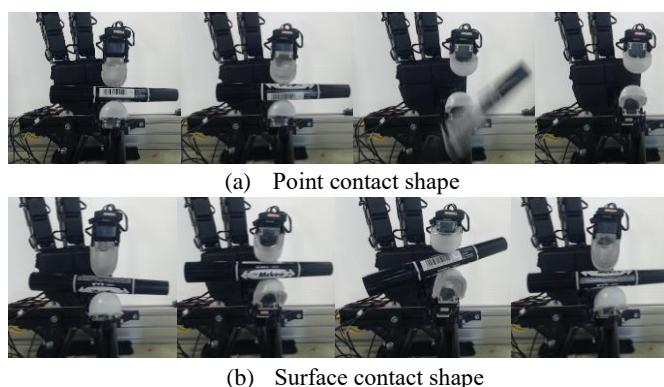


図6 Comparison of manipulation by each finger. (a) dropped the object, but (b) kept the object in its grasp.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計3件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 上野貴久, 船橋賢, 伊藤洋, Alexander Schmitz, 尾形哲也, 菅野重樹
2. 発表標題 時系列情報とAttentionを用いた未知物体のロバストな手繰り寄せ把持動作の安定化手法の提案
3. 学会等名 ROBOMECH2023
4. 発表年 2023年

1. 発表者名 山田泰, 船橋賢, AlexanderSchmitz, 尾形哲也, 菅野重樹
2. 発表標題 器用な操り動作のために人の指形状を模した光学式触覚センサ
3. 学会等名 ROBOMECH2024
4. 発表年 2024年

1. 発表者名 Satoshi Funabashi
2. 発表標題 Hand Morphology from Tactile Sensing with Spatial Deep Learning for Dexterous Tasks
3. 学会等名 NeurIPS 2023 Thirty-seventh Conference on Neural Information Processing Systems, 1st Workshop on Touch Processing: a new Sensing Modality for AI. (招待講演) (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
---------------------------	-----------------------	----

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
---------	---------