

令和 5 年 6 月 21 日現在

機関番号：13904

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K12168

研究課題名（和文）画像認識と筋電信号を用いた感覚統合義手システム

研究課題名（英文）Prosthetic Hand System Using Image Recognition and EMG

研究代表者

福村 直博（Fukumura, Naohiro）

豊橋技術科学大学・工学（系）研究科（研究院）・准教授

研究者番号：90293753

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：事故などにより手指など上肢の一部を失ってしまった患者のために、自分の意図通りに動く義手が求められている。以前より研究が進められている筋電義手は、筋電信号が微弱であることなどから複雑な制御が難しいが、AI技術を用いた画像認識に基づくロボットハンド制御を併用することで、精密な制御ができる義手が期待できる。本研究ではこのシステムの実現のために、深層学習モデルを用いた視覚-運動変換モデルを実ロボットハンドによる実験で検証し、コップの画像からその直径を認識し、そのコップの大きさに合った把持手形状を計算できることを示した。さらに画像と筋電信号を併用した義手システムの試作を行い、その有効性を検証した。

研究成果の学術的意義や社会的意義

本研究で目指している義手システムはユーザーの意図を筋電から読み取り、画像認識技術と併用することで、ロボットハンドを精密に制御できることを特徴としており、義手使用者のQOLの大きな改善が期待できると共に、音声入力を用いることで、高齢者や寝たきりの患者の生活サポートロボットなどにも応用できる。また、検証してきた視覚-運動変換モデルは多くの深層学習の手法と異なり、ロボット制御に必要な物体の特徴量を中間層に教師信号を使わずに抽出でき、そしてそれを介した拘束条件付き最適化問題を容易に解ける。さらに、この情報抽出は画像と姿勢情報の統合を通して実現しており、マルチモーダル情報統合の有用性も示している。

研究成果の概要（英文）：For patients who have lost a part of their upper limb, such as a hand, due to an accident, a prosthetic hand that operates according to their intention is essential to improve their quality of life. Electromyographic prosthetic hands, which have been studied for a long time, face challenges in complex control due to weak electromyographic signals. However, by combining robot hand control based on image recognition using AI technology, it is possible to achieve precise control of the prosthetic hand. In this study, we aimed to realize this system and verified the visual-motor transformation model using deep learning models through experiments with a real robot hand. We demonstrated that the model can recognize the diameter of a cup from its image and calculate the appropriate grasping hand shape based on the size of the cup. Furthermore, we developed a prototype of the prosthetic hand system that combines image and electromyographic signals and validated its effectiveness.

研究分野：計算論的神経科学

キーワード：筋電義手 画像認識 視覚 運動変換モデル 把持手形状決定 オートエンコーダ ニューラルネットワーク

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 1. 研究開始当初の背景

事故などにより手指など上肢の一部を失ってしまった患者にとって、自分の意図にしたがって動作する義手は QOL の向上に欠かせない。そのため、残存している腕の一部から得られる筋電信号を用いて制御する筋電義手の研究が以前から進められているが、皮膚表面から計測する筋電信号は微弱でノイズも多く、また筋疲労などの影響を受けて同じ運動を行っても信号が変化するなどの問題がある。そのため、多自由度な制御対象であるロボットハンドを精密に制御することは難しく、数種類の典型的な把持タイプを識別することが限度であり、多自由度で巧みに制御できる多指ハンドのメリットを活かせない。一方、深層学習による画像認識技術が近年めざましく向上し、画像処理から多指のロボットハンドを制御する研究も盛んに行われるようになった。この画像認識技術を用いれば、単に物体の種類を認識し、それに基づいて典型的な把持タイプを割り当てただけでなく、対象物の大きさなどの詳細な形状を認識し、それに合わせたロボットハンドの把持形状、すなわち指関節角の計算も可能であると期待できる。しかし義手用途を考えると、例えば同じコップであっても水を飲むために取手を持つ、あるいは運ぶためにコップ上部を持つなど、用途によって操作するための把持タイプが異なる場合があり、画像認識による情報だけでは、操作する意図を反映させることができない。そこで、画像と筋電信号を併用することで両者の利点を活かし、筋電信号でユーザーの意図を読み取り、画像情報から物体の種類、および形状の認識を行って指関節角を計算して精密な制御ができる、ユーザーにとって自然に扱える義手の実現できると期待できる。

この義手システム開発において重要なのは画像認識部分である。近年の深層学習を用いた技術発展により、物体の種類を識別する精度は高くなってきているが、精密な制御をする上で重要になるような、物体の大きさなどの詳細な形状の情報を認識するのは、教師信号を準備することが多大な手間がかかる。さらに形状が複雑な場合もあり、かつ日常的に用いる物体のどの情報が把持形状を決定する上で重要であるかはわからない場合が多い。そのため、物体の画像データと、それを把持した時の把持形状データのセットを用意して、その関係を学習することが考えられる。しかし、このような、対象物を撮影した画像情報から多指ロボットハンドの把持形状を計算することは多対多の変換である。まず画像情報はカメラおよび対象物との相対位置や角度などに応じて複数のパターンが考えられる。一方、先述のコップのように、物体によっては操作タスクに応じて異なる把持タイプが可能である上に、ある把持タイプを選択したとしても、各指の冗長自由度や把持点の変化により把持形状は一意に決まらない。そのため、対象物を撮影した画像データを学習モデルの入力、それを把持できた際の把持形状データを学習モデルの教師信号とした end-to-end の学習を行ったとしても、正しく視覚-運動変換を学習できない。このような、計算が困難である多対多の変換である把持動作を人は無意識のうちに行っている。人は物体把持を行う際、視覚情報だけでなくこれまでの様々な物体を把持したときの運動情報を統合して学習し記憶することで多対多の変換を実現する。これにより人は把持する前に視認した対象物の形や大きさに応じて把持形状を予測する。この情報処理過程を計算モデルで実現する場合、物体の画像から把持形状を決定するために必要な物体の形状に関する特徴量を計算し、その特徴量に合った把持形状の中から最適な把持形状を求め、という拘束条件付きの最適化問題として解く手順が考えられる。この考え方に基づき、物体の画像データとそれを把持した時の把持形状データを統合する把持運動モデルが提案され [1]、さらにこのモデルを改良した、複数の Auto-Encoder(AE) から構成される視覚-運動変換モデルが提案されている[2]。このモデルは多対多の関係にあるデータセットを学習でき、それぞれの情報に関する恒等写像と情報統合によって、両方の情報に共通に含まれる情報を教師信号なしに抽出できる。これを把持形状決定に使う場合、物体の画像データとそれを把持した時の把持形状データのセットを学習させると、両方の情報に含まれる物体の形状に関する情報が抽出できると期待できる。さらに、学習後のモデルに物体の画像データを入力すると物体形状に関する情報を求めることができるため、その情報に合った把持形状の中から把持形状に関する評価関数を最適化することにより、把持形状を一意に決めることができる。そして、この評価関数は操作タスクによって変化することが考えられるため、この評価関数を実現したい把持タイプごとに切り替えることで、操作タスクに依存した適切な把持形状を計算できると期待できる。このモデルは実ロボットを使った簡単な問題、あるいは把持運動の場合にはシミュレーションでのみ検証されているため、実ロボットハンドを用い、さらに日常的に使われる対象物を用いて検証することが必要になっていた。

## 2. 研究の目的

本研究では、筋電信号から得られるユーザーの意図を反映し、画像認識により精密にロボットハンドを制御することで巧みな把持動作を可能にする、図 1 のような義手システムの実現を目指す

す。まず、その基盤技術となる、画像から物体の形状に合わせた把持形状を計算する手法の確立を主な研究目的とする。研究代表者がすでに提案している、対象物の画像データとそれを把持した時の把持形状データを統合することで、物体の形状に関する特徴量を抽出することができ、さらに学習後には物体の画像データから把持形状を計算できるモデルの検証を、実ロボットを用いて試みる。さらにシステムのプロトタイプを作成して、提案システムの有効性を示す。

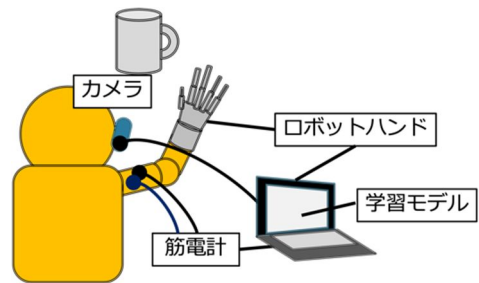


図 1 義手システムの概略図

### 3. 研究の方法

(1) ハンドをもつマニピュレータを制御して物体把持を行うためには、画像情報から対象物体の位置と形状の情報を別々に認識しなければならない。提案している、複数の感覚情報の統合から情報抽出を行う学習モデルを用いると、物体の画像情報とその物体を持った時のロボットアームの姿勢情報の統合により物体の位置の情報を、同じく画像情報と物体を持った時のロボットハンドの指関節角情報の統合により物体の形状情報を別々に抽出できることが期待できる。これを検証するため、図2のような4指で計4自由度をもち、さらに手首部分に2自由度をもつロボットハンドと、4自由度のロボットアームを組み合わせたマニピュレータを組み立て、位置が固定されたWebカメラで対象物体を撮影できる実験環境を構築した。



図 2 ロボットアームとハンドを用いた実験環境

(2) 筋電義手システムの場合は使用者の腕によって手先位置は制御できるため、人の手に近い多指ハンドの制御方法を中心に検討すべきであることから、図3のように、4指に各4自由度を持つ計16自由度のロボットハンドを導入して、人の腕に取り付けて操作できるようにした。さらにロボットハンドを装着した操作者の手指関節を計測するデータグローブを用い、このロボットハンドをリーダー・フォロワー制御して物体把持を実現するとともに、把持した時のロボットハンドの指関節角を計測できる実験環境を構築した。カメラについては、当初は位置を固定したWebカメラを用いたが、より一般的な条件においても対象物体の立体形状を取得するために深度画像を同時に取得できる深度センサを用いることとし、得られた深度画像を学習用、あるいはテスト用のデータとして利用した。対象物体として複数の持ち方が可能なコップを採用し、3Dプリンタを用いて直径と取手長さが異なるコップを15個作成した。

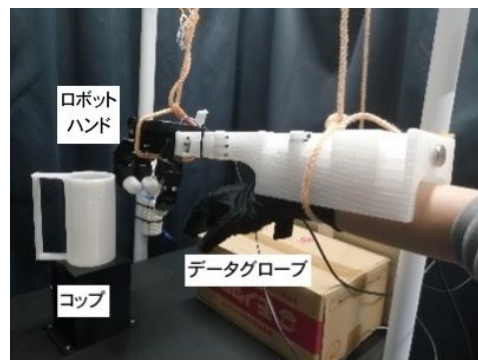


図 3 義手用ロボットハンドの実験環境

(3) 画像認識と筋電信号処理を組み合わせた義手システムの試作も進め、頭部にカメラを、上腕に筋電センサを取り付け、それぞれの信号をノートPCに取り込むシステムを構築した。

### 4. 研究成果

(1) 直径3cmと4cmの円柱を用意し、これを1cm間隔の格子状の25箇所の位置に順に置き、マニピュレータをマニュアルのコントローラで制御して円柱を把持して持ち上げ、卓上に戻す動作を行った。戻した時のマニピュレータのハンドおよびアームの関節角度を姿勢情報として取得し、その後ロボットを物体から離れた後の物体の画像情報をWebカメラによって取得した。これを各条件で5回繰り返し、合計250セットの学習データを取得した。このデータセットを用い、図4のように従来の視覚-運動変換モデルを拡張して、姿勢情報をアームとハンド用の異なるAEに入力することにし、それぞれのAEの中間層と画像AEの中間層の一部のニューロンを同じ値になるように拘束した。また、画像AE内の物体認識のためのエンコード部分にConvolutional Neural Network(CNN)を用いた。学習後に、再度学習データを入力してAEの中間層ニューロンの値を確認した結果、画像情報とアーム姿勢情報の統合により、円柱が置かれた格子状の位置が画像およびアームのAEの中間層の拘束した2つのニューロンにより表現されていることが確認できた。

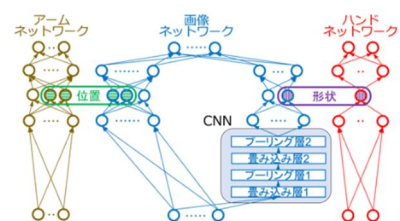


図 4 物体画像と、マニピュレータの姿勢情報を統合するモデル

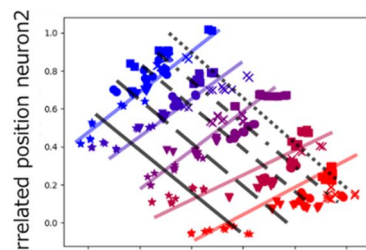


図 5 抽出された物体位置情報



(図5)。同時に、画像とハンドの指関節角情報の統合により、円柱の直径の情報を表現するニューロンが確認できた。この結果より、実環境において、このモデルを用いた画像情報とロボットの姿勢情報の統合により、対象物の位置や形状に関する情報を明示的な教師信号なしに抽出できることを示した。さらに適当な位置に置いた対象物の画像情報から、実際に到達把持制御ができることも確認した。

(2) 直径が 8cm から 12cm までの 5 通り、取手の長さが 3 通りの異なるコップ 15 個を 3D プリンタで作成し、対象物とした。コップの位置とカメラの位置を固定し、カメラから見たコップの向きを変えた画像データを取得した。さらに、4 指のロボットハンドをリーダー・フォロワー制御する実験環境を用いてコップの上部を持つ把持タイプ(TOP)と、コップの側面を持つ把持タイプ(SIDE)の 2 つの把持タイプで把持した時のロボットハンドの把持形状データを取得した。この画像データと把持形状データを組み合わせた合計 22500 セットのデータセットを用い、先述の AE を用いた視覚-運動変換モデルでコップの画像データとそれを持った時のロボットハンドの指関節角データを統合する学習を行なった。2 つの AE の中間層の 1 つのニューロンを同じ値になるように拘束することで、特徴量としてコップの直径の大きさの情報をそのニューロンに抽出することができ、かつ学習後のモデルを用い、画像データからコップの直径に合った TOP および SIDE の把持タイプを、把持タイプによる評価関数を切り替えることで、それぞれ計算できることを示した。最初の実験では 1 枚の画像から立体情報を得るために、対象物の位置とカメラの位置を固定していたが、より一般的な条件でも扱えるように、物体の立体形状に関する情報を取得できる深度画像が RGB 画像と同時に撮像できる深度センサを用いることとした。深度画像は深度が正確に取得できない部分が多く生じるため、RGB 画像で得られたエッジ情報によって推定された物体領域で深度センサ情報をマスキングし、物体内部のデータが取得できていない空間を補間することで、より正確な深度画像を作成した。さらに図 7 に示すように、モデルの画像 AE のエンコード部分に既存の学習済みの CNN モデルである ResNet50[3]を導入して学習実験を行なった。その結果、これまでの全結合型のニューラルネットワークを AE に用いた場合と比較して、AE の中間層ニューロンの値はコップ直径に対して極めて線形性が高い表現が得られ、さらに学習が早く進むことを確認した。また、汎化性能を評価するために学習に使わなかった直径のコップや、コップを置く角度を変えた場合の画像データをモデルに入力したところ、図 8 のように、コップサイズを表現するニューロンの値は非常に高い汎化性能を示した。また、画像データからそのコップを把持するのに適した TOP および SIDE の把持形状を計算できた(図 9)。以上の結果から、提案している視覚-運動変換モデルは、実環境で取得したデータでも、画像情報と手形状情報にも含まれている特徴量を教師信号なしに抽出することができることと、学習後のモデルを用い、把持形状決定のための評価関数を切り替えることで、異なる把持タイプの把持形状を計算できることを確認し、義手などのロボットハンドの精密な制御に利用できる可能性を示した。

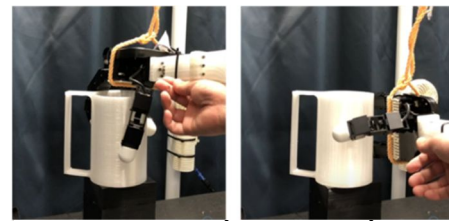


図 6 コップの把持タイプ  
(左: TOP, 右: SIDE)

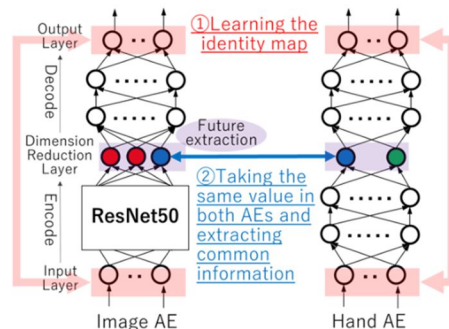


図 7 物体画像と把持形状情報を統合するモデル

と同時

に撮像できる深度センサを用いることとした。深度画像は深度が正確に取得できない部分が多く生じるため、RGB 画像で得られたエッジ情報によって推定された物体領域で深度センサ情報をマスキングし、物体内部のデータが取得できていない空間を補間することで、より正確な深度画像を作成した。さらに図 7 に示すように、モデルの画像 AE のエンコード部分に既存の学習済みの CNN モデルである ResNet50[3]を導入して学習実験を行なった。その結果、これまでの全結合型のニューラルネットワークを AE に用いた場合と比較して、AE の中間層ニューロンの値はコップ直径に対して極めて線形性が高い表現が得られ、さらに学習が早く進むことを確認した。また、汎化性能を評価するために学習に使わなかった直径のコップや、コップを置く角度を変えた場合の画像データをモデルに入力したところ、図 8 のように、コップサイズを表現するニューロンの値は非常に高い汎化性能を示した。また、画像データからそのコップを把持するのに適した TOP および SIDE の把持形状を計算できた(図 9)。以上の結果から、提案している視覚-運動変換モデルは、実環境で取得したデータでも、画像情報と手形状情報にも含まれている特徴量を教師信号なしに抽出することができることと、学習後のモデルを用い、把持形状決定のための評価関数を切り替えることで、異なる把持タイプの把持形状を計算できることを確認し、義手などのロボットハンドの精密な制御に利用できる可能性を示した。

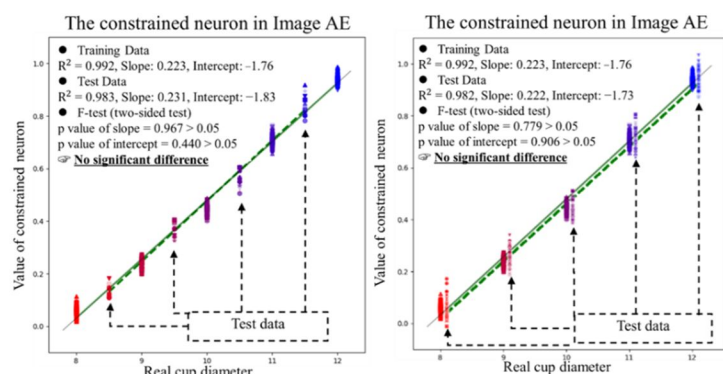


図 8 抽出されたコップのサイズ情報

	8 cm	8.5 cm	9 cm	9.5 cm	10 cm	10.5 cm	11 cm	11.5 cm	12 cm
TOP									
SIDE									

図 9 画像データから計算したロボットハンド形状

(3) 画像と筋電信号を同時に取得できるシステムを構築し、提案している画像認識と筋電信号処理を併用した義手システムの有効性を検証した。義手をスムーズに制御するためには、対象物に手を伸ばす到達運動中に把持形状を準備する Preshaping の段階での筋電信号により把持タイプを分類できることが望ましいが、到達運動の筋電信号成分が重畳するため、高精度での識別は難しい。本実験では、対象物をコップ、ボトル、文庫本、ハサミの4種類とし、それらを把持するための9種類の把持タイプを到達運動中に識別することを目的とする実験を行った。筋電信号のみで把持タイプを選ぶ実験を行ったところ、やはり識別は難しく、識別率が最大であった被験者でも57%であった。一方、画像認識を併用する手法として、まず対象物体の画像データからこれらの4種類の物体を識別できる YOLACT[4]によって把持運動の開始前に物体を識別し、次にその物体に合った把持タイプの中から筋電信号によって被験者が意図した把持タイプを選ぶ手法を用いた。YOLACT により物体の識別が正確にできると、最大で3つの把持タイプから筋電信号により選ぶことになり、識別は比較的容易になることが期待される。実際に検証実験を行ったところ、被験者によっては最大99%の認識率になった。この結果より、提案システムの有効性を示すことができた。

#### <引用文献>

- [1] Y.Uno, N. Fukumura, R. Suzuki and M. Kawato, "A Computational Model For Recognizing Objects and Planning Hand Shapes in Grasping Movements," Neural Networks, Vol.8, No.6, pp.839-851, 1995
- [2] N.Fukumura, K. Wakaki, Y. Uno, "A Modular Structure of Autoencoder for the Integration of Different Kinds of Information," 1<sup>st</sup> International Conference on Natural Computation(ICNC05), pp. 313-321, 2005.
- [3] K.He, X.Zhang, S.Ren, and J.Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 770, 2016
- [4] D.Bolya, C.Zhou, F.Xiao, and Y.J.Lee, "Yolact: Real-time Instance Segmentation," Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.9157-9166, 2019

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 松田基, 福村直博
2. 発表標題 視覚-運動変換の特徴抽出モデルによるロボットアームの到達把持運動制御
3. 学会等名 日本神経回路学会全国大会
4. 発表年 2019年

1. 発表者名 Motoi Matsuda, Naohiro Fukumura
2. 発表標題 Verification of Feature Extraction from a Visuomotor Integration Model with Convolutional Neural Network
3. 学会等名 International Conference on Advance Informatics: Concepts, Theory and Applications (国際学会)
4. 発表年 2020年

1. 発表者名 松田基, 福村直博
2. 発表標題 多指のロボットハンドによるサイズの異なるコップを把持するための視覚-運動変換モデルの検証
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2021年

1. 発表者名 大橋真愛, 松田基, 片山哲, 福村直博
2. 発表標題 CNNと深度画像を用いた視覚-運動変換モデルによるロボットハンド制御
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2022年

1. 発表者名 M. Matsuda, S. Katayama, M. Ohashi and N. Fukumura
2. 発表標題 Evaluation of Generalization Performance of Grasping Control with Robot Hand Using Visuomotor Integration Model
3. 学会等名 Neuro2022 (国際学会)
4. 発表年 2022年

1. 発表者名 M. Matsuda, S. Katayama and N. Fukumura
2. 発表標題 Grasp Control of a Multi-fingered Robot Hand for Cups of Different Sizes using a Visuomotor Integration Model
3. 学会等名 2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (国際学会)
4. 発表年 2022年

1. 発表者名 片山哲, 松田基, 福村直博
2. 発表標題 深度画像を用いた精密なロボットハンド制御のための把持対象物の特徴抽出
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2023年

1. 発表者名 荒木雄斗, 松田基, 福村直博
2. 発表標題 画像認識を併用した筋電義手システムの把持動作識別精度の検証
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

生体運動制御システム研究室ホームページ  
<http://www.bmcs.cs.tut.ac.jp>

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