

令和 6 年 6 月 25 日現在

機関番号：51201

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12079

研究課題名（和文）抽象的な概念を表す符号を自律的に学習するための自己符号化器の構成方法の開発

研究課題名（英文）Development of generating an auto-encoder for abstract target without explicit classification

研究代表者

松尾 直志（Matsuo, Tadashi）

一関工業高等専門学校・その他部局等・准教授

研究者番号：80449545

交付決定額（研究期間全体）：（直接経費） 2,100,000円

研究成果の概要（和文）：ロボットハンドでネギなどの粒状食材を把持した際の把持重量の推定は「食材の状況」という不確定要素を含む情報から「動作と把持される量の関係」を抽出する問題と考えれば複雑な対象から隠れた本質的な情報を抽出する問題と見なすことができる。しかし粒状食材の配置の仕方は無数にあるうえ、一度把持操作を行つとその食材の状況は破壊されてしまうため同じ状況での把持実験を繰り返すこともできない。そこで、食材の状況を表す画像から「動作と把持される量の関係」を把持量のばらつき自体も含めて予測する仕組みを開発し、実際の食材把持実験によって想定通りに予測ができることを確認した。

研究成果の学術的意義や社会的意義

本研究ではある状況下でのロボット動作によって得られる結果を確率付きで予測する方法を開発した。提案方法では状況を事前に分類する必要がなく、実際の状況とロボット動作情報、及び結果情報があれば予測モデルを組み立てることができ、人手による分析の手間を抑えることができる。ロボットハンドを用いて細かいネギの山から適量を把持する実験にこれを適用し、実際に把持される量の確率分布を予測できることを確認した。

研究成果の概要（英文）：When a robot hand grasps granular foodstuff such as green onion, the estimation of grasped weight can be considered as extraction of essential information from complex and indefinite situation such as a state of the foodstuff. Although the relation between a robot action and a grasped weight will depend on a state of the foodstuff, but we cannot repeat an experiment on the same situation because one action will destroy the situation. So, we proposed a framework for estimating the relation between an action and a grasped weight from an image of the foodstuff that represents the situation. The framework can also estimate the probabilistic distribution of weights and it can be applied to various robot actions with probabilistic results. We applied the framework to experiments of grasping green onion with a soft robot hand and we confirmed that the framework will work well.

研究分野：画像認識

キーワード：機械学習 教師なし学習 自己符号化器 コンピュータビジョン

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

近年の計算機は YOLO 等の技術を用いて画像内の物体種別を認識することができるようになってきている。物体種別を認識するためには事前にその種別を学習させる必要があるが、人は初めてみる物体であっても取手らしき部位があればそれを用いて把持を行うことができる。また、そのような典型的な持ち方の概念は、対象の物体の細かな形状や模様には依らない抽象的な概念であり、人は他の人がその物体を扱う様子を観察することで自律的に学習することができる。人はさらにこの持ち方の概念と取手のような形状との組を学習することで未知の物体に対する操作方法や種別を推測することができる。人の生活空間には日々新しい物体が導入されるため、そこで働くロボットを開発する際、事前にそれらの物体全てを学習することは現実的ではない。物体の典型的な部分形状と持ち方の関係を学習できれば、取手の形状を手掛りにそこを把持するといったことが可能となる。しかしそれらの関係を学習するには「持ち方」を計算機上で表現する方法が必要となる。画像の種類がごく少数であれば人手によりラベル付けして符号とすることも考えられるが、把持方法は本質的に連続であり、起こり得る全ての画像に対して人手でラベルを与えることは困難である。そこで教師ラベルを用いず画像群それ自体から自動的に符号を設計する自己符号化器(auto-encoder)の方法が考えられる。しかし、物体を持っている様子には様々な見え方があるため、従来の方法では画像からの情報だけで典型的な持ち方のグループを把握し自律的に符号化するのは難しい。

2. 研究の目的

本研究では、対象について与えられたある種の類似性に関する知識と整合する教師なし符号化を行う方法を開発する。知識としては例えば物体形状自体に着目する場合にその画像中の位置が多少変動しても対象としては同一と見なせるといった、自然ではあるがこれまで auto-encoder 構成の際には考慮されていなかったものを扱う。これによって事前に対象ごとの正規化を行うことなく対象データが持つ性質を反映した符号を得ることができる。本研究で導かれる auto-encoder を用いれば画像内の位置や向きに関係なく把持の方法自体に対応した符号が得られると期待できる。従来、人が物体を把持する際の把持方法はそれ自体を数値として表現することが難しく、物体形状とそれに対して人が取り得る把持の仕方との関係を記述するモデルを構成することは困難であったが、本研究に基づく符号を用いればそれが可能となる。ある物体の把持方法はその物体の持つ機能との深い関係を持つことが指摘されており、初見の物体に対する機能推定への発展も期待できる。物体形状に即した符号があれば、それらを組み合わせると初見の物体に対しても適切な持ち方を想起できることが期待される。

3. 研究の方法

本研究は、観察によって得られた多数のデータそれ自体から対象の状況を表現する概念を反映した符号を生成する auto-encoder を構成する枠組みの開発を目的とする。本研究では以下のサブテーマを設定し、事前知識を反映した計量を持つ符号を与える auto-encoder についての開発を行う。

- (1) 人による物体の把持方法や物体の配置それ自体を表現する符号が画像から得られるとすれば、その符号は画像撮影時の位置やアングルには依存しないはずである。しかし単に画像を再現する符号を考えると同種の把持方法でも書き方によって異なる符号となってしまう。ここから逆に、そのような変換によって符号は変化しないという拘束を評価関数に組み入れて符号を生成する方法が考えられるので、まずは回転や拡大縮小などの人工的な変換に関して不変な符号を与える auto-encoder を構成する方法を開発する。
- (2) 人工的な変換ではなく、ある把持方法を複数視点から撮影したデータや同一地点の昼夜の風景画像のように同値であることが分かっている観察データの組を手掛りにし、それらの関係を反映した auto-encoder を構成する方法を開発する。

4. 研究成果

本研究課題の成果としては、連続的なバリエーションのあり得る対象について本質的な情報を代表する符号を与える方法の開発と、それに関連して不確定要素を含む状況における適切なロボット動作を推定するための方法の開発がある。後者については当初予定していなかったものであるが、不確定要素を含む状況に隠れた本質的な情報を引き出すという意味では前者の方法の応用例の一種と見なすこともできる。それぞれについては以下の通りである。

- (1) 様々な見え方がある対象に自動的に符号を割り当てる方法の研究として、3次元物体を撮影した画像から、それと整合する3次元形状の候補を想起する方法を開発し、国際会議 (IROS2021) で発表を行った。

ある視点から撮影した画像1枚だけでは3次元形状が一般には定まらないため、その画像から想像できる矛盾しない3次元形状は無限に存在する。人はそれらの3次元形状を想像し、別の視点から撮影した画像を組み合わせる追加の視点から撮影した画像を組み合わせる3次元形状の候補を絞り込んでいくことができる。この能力を計算機上で実現するため、3次元形状に変換できる多次元ベクトルを考えて記述子とし、2次元画像を入力すると記述子空間上の集合を出力する深層学習モデルを構築した。

モデルが出力する集合内の記述子は入力された画像と矛盾しない形状を表しており、同一状況について複数視点から撮影した画像を

入力してそれぞれ集合を得たとき、それらの共通部分として絞り込まれた形状候補を得ることができる。発表では画像と記述子集合、及び3次元形状が見たすべき3つの論理的制約に基づく損失関数を新たに提案し、これを用いてモデルの訓練を行っている。この損失関数を最小化することによって、画像から記述子空間上の集合への変換や記述子から3次元形状への変換の両方が自動的に構成される仕組みとなっている。このため、画像と3次元形状のペアのデータセットさえあればモデルを構成でき、人手で設計する必要がない。このようにして得られる記述子は見え方と関連しており自動的に符号を割り当てたものと見なせる。また、この方法は部分的な情報から全体像を想像する問題に適用できる構造となっており、実際に手書き文字の一部の画像から全体形状を想像して絞り込む実験についても機能することが確認できた (図1)。

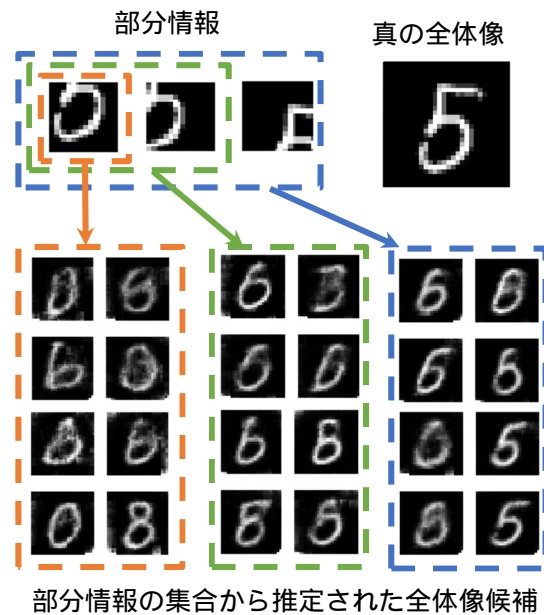


図1. 部分情報からの全体像候補の推定

- (2) 当初の計画では3次元立体形状を撮影した画像から3次元形状それ自体を表現する符号を生成する方法として、回転や拡大縮小などの人工的な変換に関して不変な符号を与える autoencoder を構成する方法について研究する予定であった。しかし、ロボットハンドによる粒状食材の把持重量の推定問題において、粒状食材の状況を表す符号を推定する手法の開発が進んだためこちらを優先した。

これは人工的な変換ではなく実際の実験を行うことによって収集したデータに基づいて推定を行う方法であり、また把持重量という対象 (粒状食材の状況) の類似性を表す手掛りを活用する方法であり当初予定において2022年度及び2023年度に取り組む予定であった内容に重なるものである。

ネギやコーンなどの粒状食材をロボットハンドによって把持する際、把持動作の結果実際に把持される食材の量はロボットハンドの動作だけでなく把持を行う場所での粒状食材の積み方にも依存する。同じ動作であっても食材が山のように積み上げられている場所では把持される量が多くなり、食材が凹みを形成している場所では把持される量は少なくなるという傾向はあるものの、その傾向自体を表現する数値を人が設計するのは困難である。加えて、食材の積み方が同様であってもロボットハンドの動作によって把持される量は異なる。つまり食材の状況によって「ロボットハンドの動作と把持される量の関係」がある程度定まるため、食材の状況から上記の関係を抽出できれば一定量を把持する動作を実現できる。しかし粒状食材の配置の仕方は無数にあるうえ、一度把持操作を行うとその食材の状況は破壊されてしまうため同じ状況での把持実験を繰り返すこともできない。

この問題は食材の状況として食材の色情報、深度情報を含む画像を考え、そこから「動作と把持される量の関係」を抽出する問題と考えれば複雑な対象から隠れた本質的な情報を抽出する問題と見なすことができる。また、似た状況であっても粒状食材の挙動は毎回異なるため、把持される量にはばらつきがある。そこで、食材の状況を表す画像から「動作と把持される量の関係」を把持量のばらつき自体も含めて予測する仕組みを開発 (図2) し、実際の食材把持実験によって想定通りに予測ができることを確認した。開発した方法を利用するシステムを企業で活用する話も進展しており展示会などでも発表する予定である。このシステムについての特許も出願中である (出願番号: 特願 2022-105871)。

開発した把持重量の確率分布を推定する方法では、粒状食材状況それぞれについての把持重量の確率分布という、食材状況を明示的に分類しなければ推定することが難しかったものを状況を明示的に分類することなく推定することに成功している。また、これによって一定量

の食材を把持するためのハンド動作量の推定も行うことができ、単に符号を生成するだけでなくそれをを用いた具体的な応用に近いものとなっている。加えて、開発したフレームワークは不確定要素を含む状況についての情報から、その状況でのロボット動作に対する結果の確率分布を推定するものであり、粒状食材の把持量推定だけでなく、色々なロボット動作への応用が期待できる。

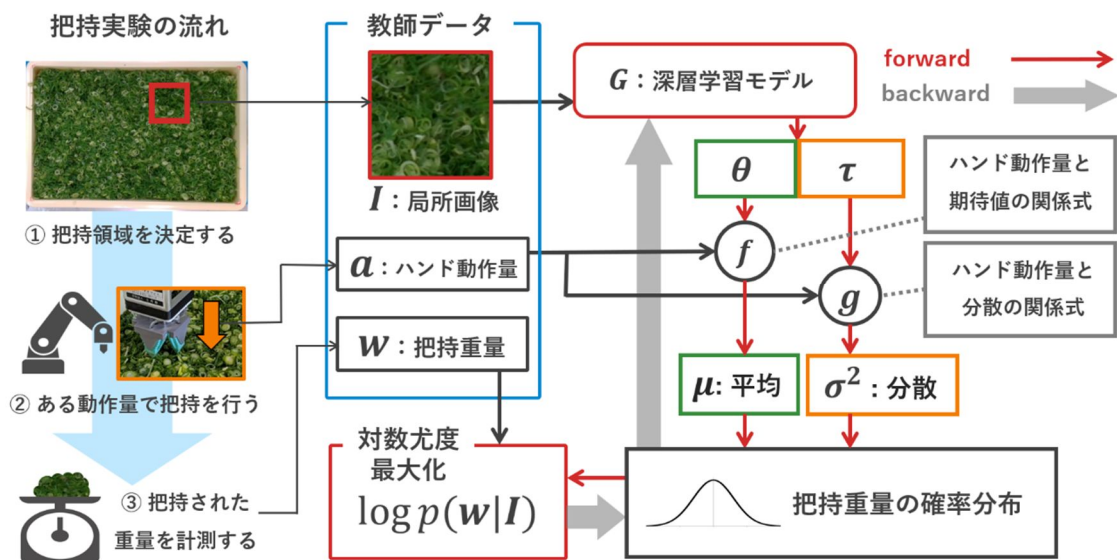


図2. ネギ把持実験におけるロボット動作量と把持重量の関係を推定するフレームワーク

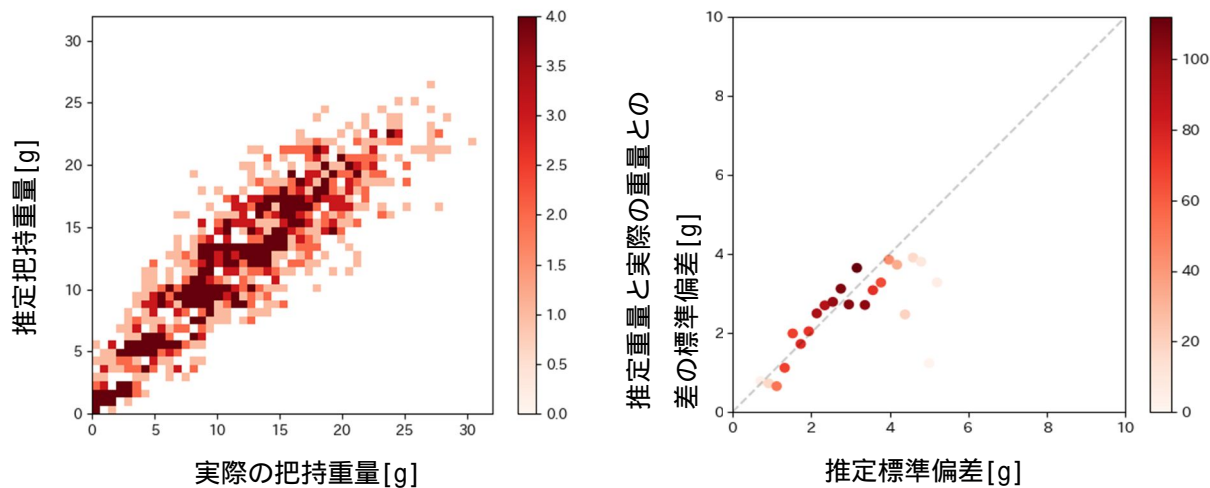


図3. ネギ把持実験において得られた把持重量に関する推定値と実際の値の関係

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 金子貴博, 松尾直志, 島田伸敬, 本田敦, 石田沙弥加
2. 発表標題 粒状食材の定量トッピング作業自動化に向けたRGBD画像からの領域提案モデルに基づく把持行動の生成
3. 学会等名 第25回画像の認識・理解シンポジウム(MIRU2022)
4. 発表年 2022年

1. 発表者名 深田 紘希, 松尾 直志, 島田 伸敬, 本田 敦, 吉浦 竜一, 王 忠奎, 平井 慎一
2. 発表標題 回帰係数を出力する深層モデルに基づく不確かさを考慮した粒状食材の定量把持制御
3. 学会等名 第29回画像センシングシンポジウム(SSII2023)
4. 発表年 2023年

1. 発表者名 Takaaki Fukui, Tadashi Matsuo, Nobutaka Shimada
2. 発表標題 Scene Descriptor Expressing Ambiguity in Information Recovery Based on Incomplete Partial Observation
3. 学会等名 The 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Hiroki Fukada, Tadashi Matsuo, Nobutaka Shimada, Atsushi Honda, Ryuichi Yoshiura, Zhongkui Wang, Shinichi Hirai
2. 発表標題 Uncertainty-aware quantitative grasping control of granular foodstuff based on a deep model that outputs regression coefficients
3. 学会等名 The International Workshop on Intelligent Systems(IWIS) (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 不定形の小型食材の集積場所への挿入深さを変えられるロボットハンドを制御するための制御情報を出力するシステム	発明者 島田 伸敬, 松尾 直志, 本田 敦, 石田 沙弥香	権利者 株式会社ニッポン
産業財産権の種類、番号 特許、特願2022-105871	出願年 2022年	国内・外国の別 国内

〔取得〕 計0件

〔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
研究 分 担 者	島田 伸敬 (Shimada Nobutaka) (10294034)	立命館大学・情報理工学部・教授 (34315)	

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