

令和元年度学内公募研究（萌芽型）  
〔研究論文〕

## 多脚クローラ型ロボットによる学習と 注視機能に基づく自律作業動作制御の研究

藤田 豊己<sup>1)</sup>, 相見 伸篤<sup>2)</sup>, 安田 裕一<sup>1)</sup>

### A Study on Autonomous Work by Multi-Legged Crawler Robot with Deep Learning and Visual Attention

Toyomi FUJITA<sup>1)</sup>, Nobuatsu AIMI<sup>2)</sup>, Yuichi YASUTA<sup>1)</sup>

#### Abstract

This study aims at developing functions for autonomous work by a crawler robot with four or six legs in dangerous site. We present a method for autonomous gripping and lifting of a box object based on the detection of object planes. The experimental result confirmed usefulness of the method. In addition, we consider a method for target-object detection using Convolution Neural Network (CNN) which is a kind of deep learning. Moreover, visual attention using captioning techniques by deep learning was also considered, which may be useful for recognition function of a robot.

#### 1 はじめに

災害現場や建設現場など、人間に危険な現場で活動するロボットへの期待が高まっている。そのようなロボットには、現場で自ら物体処理などの作業を遂行する能力が求められる。このことから、筆者の研究室ではクローラ機構と4本および6本の脚機構を有する多脚クローラ型不整地移動ロボットを開発してきた。これらのロボットは、脚を作業腕として使用することで把持持ち上げなどの対象物操作が可能となり、高い作業能力を有する。このような複雑な作業環境において、ロボットには状況に応じた適切な動作が求められるが、遠隔からの操縦はオペレータの負担が極めて大きくなる。そのため、ロボットが自律的に作業を遂行できることが望ましく、本研究プロジェクトでは、複雑環境下でのロボットの自律作業を可能とすることを目指している。そこで本研究では、深度センサ情報より箱状対象物を自律的に把持運搬する手法を考案した。また、対象物検出や複数ロボット作

---

1) 東北工業大学 工学部 電気電子工学科

Department of Electrical and Electronic Engineering, Faculty of Engineering, Tohoku Institute of Technology

2) セコム株式会社

SECOM Co. Ltd.

業時の他者認識のためには学習機能や注視機能も必要であり、それぞれ深層学習を応用した手法を検討した。

## 2 平面検出による箱型対象物の自律運搬

4脚を有するクローラ型不整地移動ロボットにおいて、脚を作業腕として使用し、多面体形状対象物の自律把持運搬手法を検討した。ロボットに搭載した深度センサを用いて対象物の平面を検出し、対象物の把持計画を行う [1]。

最初にオペレータによる遠隔操作などにより対象物にロボットを接近させているものとする。そのため、初期状態で対象物はロボットのほぼ正面に位置しているものとする。深度センサを用いて周辺環境の各平面を検出し、その中から2腕での把持が可能である平行な2面の組を逆運動学計算により抽出し、対象物の把持候補面とする。このとき対象物を把持する際のロボットの位置および作業腕の姿勢の可能な組が算出されるので、その中からロボットの移動が最小となる組を選択し、それに対応する把持面と最適な把持位置姿勢を決定する。そして得られた把持位置まで移動し、対象物を把持し運搬する。図1に把持位置の検出例を示す。直方体形状の対象物に対して得られた把持面と双腕の姿勢および手

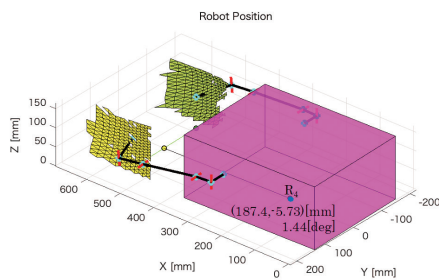


図1. 把持位置検出例

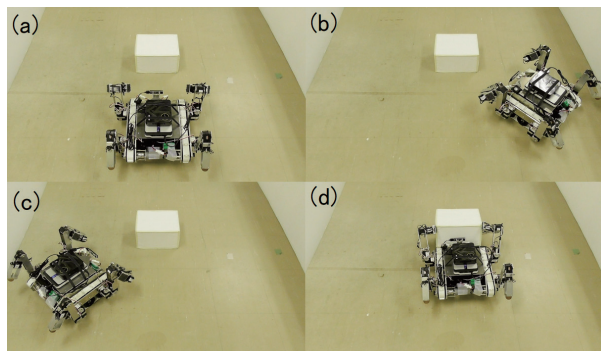


図2. 自律把持運搬実験

先位置が示されている。この手法は対象物の平面の数に制約なく適用することができるため、対象物の形状や姿勢に対して汎用性が高いという利点がある。

本手法を実機に適用して行った自律把持運搬実験を図2に示す。図中 (a), (b), (c) において異なる3点から対象物を観察し、平面情報を獲得して把持位置を検出した。そして、(d) にてその把持位置に移動し、求めた腕の位置・姿勢に双腕を動作させ、把持に成功した。この結果から、本手法が作業腕を有するロボットの自律把持に有効であることを確認した。

## 3 畳み込みニューラルネットワークを用いた対象物認識

ロボットによる運搬等の自律作業のためには、操作対象物の認識が必要である。そこで、災害時に火災や爆発の危険のため回収や発見の必要が考えられるカセットボンベを対象物として設定し、ロボットが、搭載された単眼カメラ画像を用いて、対象物を認識する方法

を検討した。図3にこの概念を示す。このために、深層学習の技術であり画像認識に特に有効である畳み込みニューラルネットワーク (Convolution Neural Network: CNN) を用いた。

認識対象のカセットボンベとその他の物体を識別する二項分類のCNNを構成した。画像データセットには、対象物としてスマートフォンカメラで撮影したカセットボンベ画像を拡張した200枚の画像と、その他の物体としてトロント大学より公開されている一般物体データセットCIFAR-100から3種の400枚の画像を用いた。それぞれから7割の画像を抽出してCNNの学習に用い、学習されたCNNを用いて残りの3割の画像を用いて対象物の検出をテストし、本手法の有効性を検証した。

本手法による学習曲線を図4.に示す。横軸は学習回数であり、青の実線が計算コストを示し、赤の実線および破線が学習用およびテスト用データに対するエラーを示す。学習を50回以上行うことで検出誤差が20%以下となったことがわかる。100回学習後のテスト用データに対する識別結果を表1.に示す。計180枚の画像において、入力対象物 ( $y = 1$ )、その他の物体 ( $y = 0$ ) それぞれで、対象物であると識別した ( $y' = 1$ ) か、そうでない ( $y' = 0$ ) かの検出結果数が記されている。Recallは各入力での正解率である。Accuracyは両者を合わせて対象物の有無を正確に判別できた割合を示しており、カセットボンベかその他かを96.6%の精度で認識することが出来たことがわかる。この結果から、本手法が対象物の認識に有用であることを確認した。

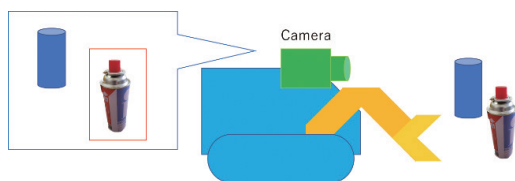


図3. 操作対象物認識の概念

表1 学習後の検出結果

Input	Detection		Recall[%]
	$y'=0$	$y'=1$	
$y=0$	112	4	96.55
$y=1$	2	62	96.87
Accuracy			96.66

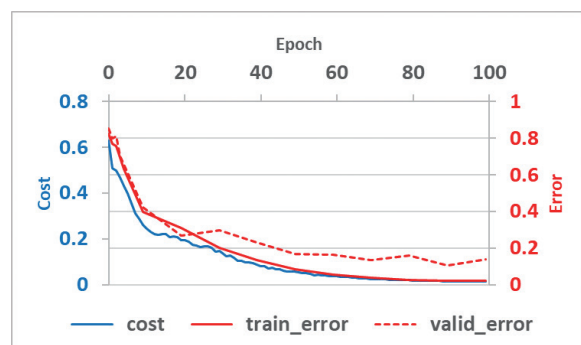


図4. 学習曲線

#### 4 深層学習を用いた視覚的注視機能

人間は注視により視野内の重要な領域を効率的に獲得し、認識している。そこでロボットへの視覚的注視機能の実装を目指し、深層学習のアテンション技術の応用を試みた。特に今回は他者ロボットの動作認識への発展可能性を検討する。そこで、既存のキャプション生成モデルをロボット動作画像に適用し、ロボット動作を表現できるかを検証した。

使用したキャプション生成モデルは、畳み込みニューラルネットワークに基づくエンコーダと、LSTM (Long and Short Term Memory) を用いたりカレントニューラルネットワークに基づくデコーダから成る。いずれも深層学習において近年よく用いられている技術である。入力画像をエンコーダで処理して画像特徴とそれに対応する注視領域を抽出する。そして、デコーダにて画像特徴と対応する単語を結びつけ、単語の時系列を生成す

る。学習データを用いてこのネットワークを学習することで、適切なキャプションを得ることが可能となる。

今回はマイクロソフト社より公開されている画像とキャプションの MSCOCO と呼ばれるデータセットを学習に用いた。ソフトウェアは Python および Tensorflow を用いて作成し、実際の学習は Google Colaboratory を利用して行った。その学習モデルをロボットがアームを動かす画像（図5）に適用したときの注視領域の変化を図6に示す。対応付けられた単語も付記している。ここで得られたキャプションは、「a toy truck is hanging from a table with a toy doll」となり、期待したものとは違いがあった。この理由として、学習データで用いた MSCOCO ではほとんどが一般的な画像を使用していたことが考えられる。ロボットの動作を対象とする学習に適切なデータを用いる必要がある。

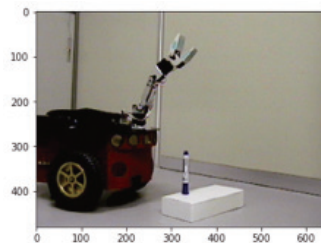


図5. ロボットアーム動作画像

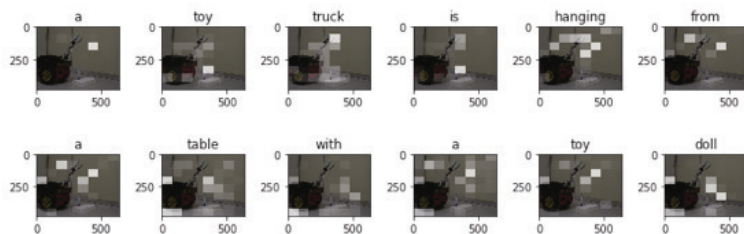


図6. 注視領域と対応単語の生成結果

## 5 おわりに

本研究では、複数の脚を有するクローラ型不整地移動ロボットによる自律的な作業動作の実現を目指し、そのための箱状対象物の平面検出に基づく自律作業の手法を考案し、実験により有効性を確認した。また、深層学習を用いた対象物認識や視覚的注視機能についても検討した。これらについては十分な結果を得るには至っておらず、今後の発展が必要となる。

### 参考文献

- [1] Toyomi Fujita and Nobuatsu Aimi, "Autonomous Gripping and Carrying of Polyhedral Shaped Object based on Plane Detection by a Quadruped Tracked Mobile Robot," In Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2019), Volume 2, pp.552-558, 2019