

# 自律ロボットの目標物認識能力の獲得について

兼松 明未 小松 芙美子 新出 尚之

我々は、人間の行為選択を模したエージェントモデルである BDI エージェントを用いて、実世界において自律的に目的を達成するロボットの実現を目指している。そのようなロボットに必要な能力の一つとして、目標物を認識してそこに到達する行動が挙げられる。従来研究では、目標物に関する大量の学習データの収集の困難さなどの制約のもとで、認識率を上げることが課題であった。そこで我々は、学習方法の検討により認識率を上げることを目指し、ディープラーニングを用いて物体を認識する能力の獲得を行うとともに、この能力を BDI エージェントが用いて行為決定を行うためのライブラリの実装を行った。これにより、ロボットが複数の経由地を経て目標物に到達するなどのプランで行動することが可能となることを示す。

## 1 はじめに

近年、自律的に目的を達成するロボットが普及しつつある中で、ロボットにも人間と同様に周囲の状況に応じて適切な行動をとる能力が求められている。そのようなロボットには、自ら思考して、目的を決め、それを達成するための手段を選択して行動する能力が必要となる。我々はそのようなロボットの基本行為の一つとして目標物を認識してそこに到達する能力の獲得を目指しており、この能力を持つ BDI エージェントの構築を行ってきた。

しかし、従来研究 [5] ではいくつかの課題があった。1 つ目は、ロボットの目標物認識能力が低かったことである。従来の学習方法では、しばしば目標物を誤って認識してしまっていたため、途中で目標を見失うことがあった。それにより、目標にたどり着くまでに遠回りをして時間を費やすことや、別の地点に到達してしまうことさえあった。そこで、認識率を上げるために、ディープラーニングによる物体認識法を利用する

ことを試みた。

2 つ目は、ロボットが行為の失敗を認識することができなかったということである。ロボットが行為の失敗を認識することができなければ、いつまでも同じ行動を繰り返してしまい、一向に目標を達成することができない。しかし、行為の失敗を認識することができるようになれば、失敗を回復するためのプランに切り替えて行動することができるようになり、目標を達成することが可能となる。このように、ロボットが状況に応じて柔軟にプランを切り替えて行動するためには、行為の失敗を認識する能力が重要となるため、これを実装しようと試みた。

本論文では、以上 2 点の解決によるロボットの目標到達能力の向上について述べる

## 2 目標物認識法

本節では、ロボットの目標物認識能力として利用するために用いた 2 種類の学習方法について述べる。

### 2.1 OpenCV による物体検出

画像認識に用いられる機械学習のアルゴリズムにはいくつかの種類がある [6]。従来研究 [3] [5] では、目標物認識の足がかりとしてオープンソースの画像処理ライブラリである OpenCV を用いて機械学習を行う方

Acquisition of object recognition ability of autonomous robot.

Akemi Kanematsu, Fumiko Komatsu, Naoyuki Nide,  
奈良女子大学人間文化研究科, Nara Women's  
University.

法を採用した [2]. なお、言語には Python を用いている. OpenCV では、識別する画像の特徴をとらえる特徴量として、Haar-like 特徴・LBP 特徴・HOG 特徴の 3 つから選択することが可能で、我々は Haar-like 特徴量を用いて学習を行った. その仕組みは、さまざまな特徴を持った Haar-like フィルタと学習サンプルを Adaboost に投入することで、どの特徴量を用いるとよいかを決定するというものである.

[3][5] では、学習によって作成した学習器を用いて目標物の認識実験を行った. すると、目標物でない物体を認識してしまう誤認識が多く見られ、認識率は約 60% 程度にとどまった. さらに、この学習器をロボットの目標物認識能力として利用し、ロボットが目標地点への到達行動を試みる検証実験を行った. この実験の結果は、10 回中 4 回の到達率となった.

#### 2.1.1 色情報を用いた認識の補完

達成率が低い要因は、誤認識の多さによってロボットが別の物体を追跡してしまうことにある. 学習器の性能を上げるために最も単純な方法は、学習に用いるデータ量を増やすことである. しかし、手作業での撮影では大量の学習データを収集することが困難な上に、学習に用いるデータ量が増大すると学習時間が増え、さらに学習結果を利用する際の処理時間も大きくなってしまいうためロボットの軽快な動きに影響を与える. 以上からこの方法は現実的ではなかった.

そこで我々は、色情報を用いて認識を補完することを考えた. 本研究で利用した Haar-like 特徴量とは、画像の一部分を切り出して局所的な明暗差の組み合わせにより画像を判別するアルゴリズムである. そのため、判別する画像は二値画像として処理されており、色の情報が失われてしまう. この色情報を用いて認識を補完することで、最も目標物に近い物体を選出することができるはずである. 手順としては、まず事前に目標物の一部を切り出した画像を用意する. つづいて、先の学習器を用いて物体検出を行い、検出された複数の領域のそれぞれについてカラーヒストグラムを計算する. そして、カラーヒストグラムが最も目標物に近似している領域を認識領域として判定し、ロボットに伝達する.

色情報による補間を行って、先と同じロボットの

目標物到達行動の検証実験を行った. この実験の結果は、10 回中 8 回の到達率となり、達成率を向上させることに成功した.

## 2.2 ディープラーニング

目標物の認識率を高め、更に到達率を上げることを目指して、我々は従来の学習方法を検討し直した. そして、画像認識の分野において非常に高精度の性能を発揮し、注目が高まっているディープラーニングを利用することを試みた [4]. なお、本研究で行ったのはニューラルネットワークによるディープラーニングである.

### 2.2.1 Caffe による画像分類

Caffe はオープンソースで提供されているディープラーニングのフレームワークである. 特徴としては、柔軟なアーキテクチャを保持していることで簡単に動作させることができる点や、GPU を利用した高速な計算処理が可能である点などが挙げられる. また、従来研究と同様の Python 言語を使用して、容易にコードを拡張することができる.

本研究では、Caffe がサポートしている CNN による多クラス分類を利用した. ネットワークの構成は、Caffe に付属の CIFAR-10 と呼ばれるデータセットを参考にしている. CIFAR-10 は、10 クラスの画像分類を行うデータセットで、2 層構造の畳み込みニューラルネットワークである.

本研究ではまず、6 クラスの画像分類器を作成した. 選択した 6 つのオブジェクトの画像をそれぞれ約 600 枚ずつ収集し、オブジェクトの領域を切り出して約 900 枚ずつの画像データを収集した. さらに、認識精度を高める手法として有効であることが知られている「データ拡張」を行った. データ拡張とは、学習データに反転、回転、拡大・縮小、輝度値の調整などさまざまな加工を施すことによって学習データ量を簡単に増幅させる手法である. これによって、各オブジェクトの画像データは 6 クラス合計で 34026 枚となった. これを学習用とテスト用に分け、その合計枚数は学習用が 29170 枚、テスト用が 4856 枚となった. これらの学習データを用いて画像分類器の学習を行った.

作成した画像分類器を用いて画像の分類テストを

行い、学習に用いるパラメータの調整を重ねた。その結果、この画像分類器の分類精度は、約 99.67%の精度を上げることができた。

クラス番号	収集枚数	サンプル数	拡張数
0	1000	1000	6000
1	605	861	5166
2	600	929	5574
3	800	1025	6150
4	600	913	5478
5	780	943	5658
合計	4385	5671	34026

表 1 6 クラス分類の画像データ枚数

### 2.3 ディープラーニングと物体検出

物体検出を行うためには、与えられた画像から物体の位置を特定し、クラス分類を行う必要がある。そこでまず、入力画像から候補領域の抽出を行わなければならない。これにはディープラーニングを用いる方法もあるが、本研究では従来研究で利用した OpevCV の traincascade を用いることにした。6 つのオブジェクトをまとめて学習すると、2.1 節と同様に物体検出器が作成される。この物体検出器で検出された複数の領域のそれぞれを、2.2.1 節で作成した画像分類器にかけることによって、検出物体の特定を行うことができる。

しかし、6 種の類似性の低いオブジェクトを 1 つの物体検出器で学習させたことによって、1 種のオブジェクトを学習させた従来の学習器よりはるかに誤認識の多い結果となった。そこで、Caffe で作成した画像分類器に 7 つ目のクラスとして「6 つのオブジェクトのいずれでもない不正解のオブジェクト」というカテゴリを設けて、再度画像分類器の学習を行うことにした。ここで、このクラスの学習データには、物体検出器で検出された誤認識領域を画像データとして保存して収集した 5400 枚を使用することとした。そして、先の学習器と同様に、このデータを拡張して得られた 39426 枚の画像を学習用とテスト用に分けると、

その合計枚数は学習用が 33799 枚、テスト用が 5627 枚となった。

こうして作成した 7 クラスの画像分類器を使用して画像の分類テストを行ったところ、98.2%の高い精度を上げることができた。そして、この画像分類器を使用することによって、誤認識領域を取り除いて物体検出を行うことが可能となった。さらに、入力画像として与えられるロボットに搭載したカメラの映像を 2 分の 1 サイズに変換して処理を行うことによって、処理の高速化を行うことができるようになった。今後、この物体検出法をロボットの目標物認識能力として利用し、ロボットが目標物到達行動を行う検証実験を行う。

クラス番号	収集枚数	サンプル数	拡張数
0	1000	1000	6000
1	605	861	5166
2	600	929	5574
3	800	1025	6150
4	600	913	5478
5	780	943	5658
6	5400	5400	5400
合計	9785	11071	39426

表 2 7 クラス分類の画像データ枚数

## 3 BDI エージェントの構築

BDI エージェントとは、信念 (Belief)、願望 (Desire)、意図 (Intention) の 3 つの心的状態を用いて熟考し、目標達成のためのプランを選択し、行動決定を行うエージェントである。これは、人間の行動決定方式を模倣したモデルであるため、人間同様の柔軟な思考を可能とする。

BDI エージェントでは、信念や願望から最終的に達成したい目標を持つと、いくつかの途中に達成する副目標が生成され、これをサブゴールと呼ぶ。このサブゴールを達成するための手段としてプランを用意しておく。そして、周囲の環境から知覚を得ると、これらのプランの中から状況に応じた最適なプランを

選択して行動する。このとき、意図が形成される。選択したプランが失敗した場合も、失敗を回復するためのプランを選び直すことが可能である。こうしてサブゴールを順に達成していくことによって最終目標の達成を目指すものである。

我々は、自律型のロボットエージェントを構築するにあたり、BDIの構築プラットフォームである Jason を使用した [1]。

### 3.1 Jason

Jason は AgentSpeak のインタプリタである。Jason で記述するプランはトリガリングイベント、コンテキスト、ボディで構成されている。論理型プログラミング言語である Prolog と同様の記述方式を用いて、条件を組み合わせることで複雑なプランを生成することができる。そのため、動的な環境に応じたさまざまなプランを用意しておくことで、目標達成に向けた幾多のアプローチが可能となる。

本研究において実装した「目標物に到達する」という目標もサブゴールのひとつである。このプランに失敗した際に、あらかじめ用意した失敗回復のためのプランを選択することによって、ロボットが目標地点への到達行動を達成することができるようになる。以降の節では、そのために実装した目標物到達行動の失敗の定義や、回復行動をとるために用意したプランなどについて述べる。

### 3.2 失敗の概念の追加

ロボットが目標物への到達行動を試みても、失敗する可能性がある。失敗が起こっても効率的に行為を遂行するために、まずロボットエージェントは行為の失敗を認識する必要がある。では、到達行動に失敗する場合はどのようなものが想定されるか。例えば、ロボットと目標物の間に障害物が存在していたり、目標物がロボットの認識可能圏内にないことが原因で、ロボットが目標物を見つけられない場合がある。また、ロボットが目標物を見つけても、目標物までの移動経路を乗り越えられない場合もある。さらに、ロボット自体の故障や不具合によって身動きがとれない場合などもある。このようにしてさまざまに起こりうる失敗

の概念について、定義を行う必要がある。

従来研究では、ロボットが目標物を認識している間は目標物への接近行動を行い、目標物を認識していない間は探索行動を行っていた。しかし、この探索行動は非常に単純なもので、単にその場で反時計回りに回転を行うというものであった。そのため、目標物が見つけれない場合はロボットがいつまでも同じ地点を回り続けてしまっていた。このような事態を避けるために、ロボットに搭載したコンパスセンサと呼ばれるセンサを利用することにした。コンパスセンサは地磁気を読み取ってロボットが向いている方角を検出するセンサである。このセンサを用いて、探索を開始する際の最初の方向を記憶しておく。そして、探索行動を繰り返したのちに、ロボットが 360 度回転したことを検知すると、その地点において目標物の認識に失敗したと判断し、行為の失敗をエージェントに伝える。これに加えて、時間がかかりすぎる行為についても失敗とみなすことにした。これには、Jason の内部アクションを利用して時間計測を行い、行為を開始して一定時間が過ぎるとその行為に失敗したと判断する。なお、失敗と判断するまでの時間は任意に設定することができる。

これらによって、「方位」と「所要時間」の側面からエージェントが行為の失敗を認識することができるようになった。

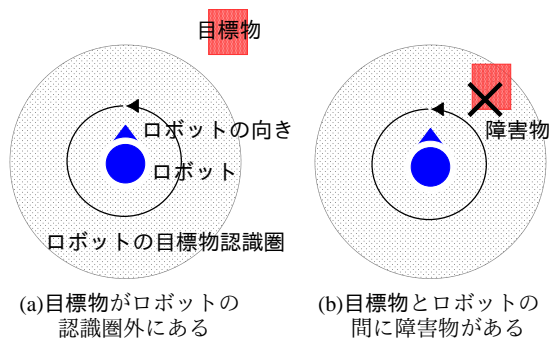


図 1 ロボットがある地点で目標物認識に失敗する

### 3.3 目標物認識の効率化

2.1 節でも述べた通り、従来の目標物認識法は精度が低かったため、一度目標物を検出しても、ロボットが目標物への接近行動を行っている間に目標物を見失ってしまうことがあった。一度目標物を見失うと、ロボットは目標物の方向から反れて探索行動を繰り返してしまうため、再発見できる可能性は低い。そのため、その地点での行為の失敗をエージェントに伝えることになる。そこで、一度見失った目標物を再発見する可能性を高めるために、探索行動の改善を行った。その内容は、先述のコンパスセンサを用いてロボットが以前目標物を認識した地点付近に戻って再探索を行うというものである。これによって、ロボットの目標物再発見率を高めることができるようになった。

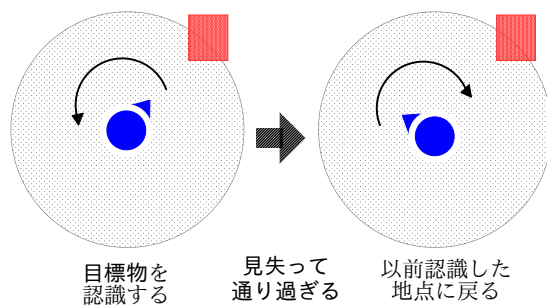


図2 ロボットの目標物再認識を効率化

### 3.4 失敗からの回復行動

3.2 節で定義した行為の失敗を検知したとしても、まだ目標を達成できないわけではない。3.1 節で述べた Jason の失敗回復行動をとるために、失敗回復のプランを2つ作成した。1つ目は、ロボットが目標物の探索行動を行った際にまったく目標物を認識せずに到達に失敗した場合に選択される。この場合、その地点には目標物が存在しない可能性が高いため、ロボットは別地点へ移動し、再び探索行動を開始する。2つ目は、ロボットが目標物の探索行動を行った際に目標物を認識した記録がある場合に選択される。この場合、目標物とロボットの間に障害物が生じたり、目標物認識能力の精度不足で目標物を途中で見失ってしまった可能性が高い。したがって、一定期間待機した

のちに同じ地点で再び探索行動を開始する。

このようにして失敗からの回復行動をとることができるようになった結果、効率的な目標物到達行動を実現することが可能となった。

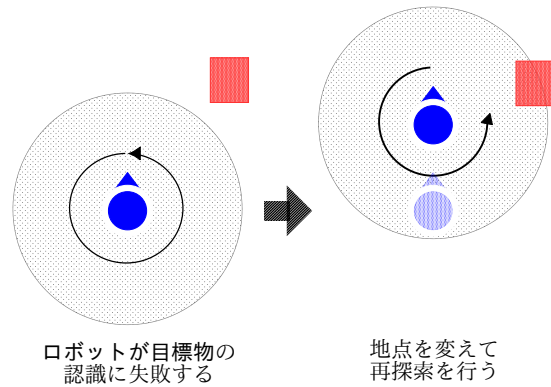


図3 ロボットが別の地点に移動して目標物の再探索を行う

## 4 まとめと今後の課題

本研究では、ロボットが自律的に目標物に到達する能力をより汎用的なものにするために、ディープラーニングを取り入れた目標物認識能力の高精度化と BDI ロボットエージェントの改良を行った。新規に獲得した能力を組み込んだ検証実験は未実施ではあるものの、各部で行ったテストでは好成績を上げている。ゆえに、ロボットの目標物到達行動の達成率も向上することが予測される。

今後の課題としては、さらに目標物の検出率ならびに認識率を向上させることによって、より正確な物体認識を行うこと。さまざまな状況に対応できるようにロボットエージェントのプランをより複雑に構成することなどが挙げられる。また、目標物到達行動以外の能力を順次獲得させることによって、自律ロボットの実用性を高めていくことができる。更なる発展を加え、より実世界に近い環境で目的を達成する自律ロボットの実現を目指していく。

謝辞 本論文の執筆にあたって、熱心にご指導いただいた新出尚之准教授と新出研究室の皆様深く感謝の意を表す。

## 参考文献

- [1] Bordini, R. H., Hübner, J. F., and Wooldridge, M.: *programming multi-agent systems in AgentSpeak using Jason*, Wiley, 2007.
- [2] 桑井博之, 豊沢聡, 永田雅人: 実践 *OpenCV2.4 for Python* 映像処理&解析, 株式会社カットシステム, 2014.
- [3] 兼松明未: ROS によるロボットの目標地点の到達行動の実現について, 2015 年度卒業論文, 奈良女子大学理学部情報科学科, 2016.
- [4] 斎藤康毅: ゼロから作る *Deep Learning Python* で学ぶディープラーニングの理論と現実, オライリー・ジャパン, 2016.
- [5] 樽井志織, 兼松明未, 新出尚之: 目標物到達機構を持つ自律的な小型ロボットの制御を行う BDI エージェントの構築, 日本ソフトウェア学会第 33 回大会, 2016.
- [6] 藤吉弘巨, 堀修, 三田雄志, 山内悠嗣他: デジタル画像処理 [改訂新版], 公益財団法人 画像情報教育振興協会, 2015.