

# インタラクションに着目した物体認識のための 人の動作による属性定義手法

石島 樹<sup>†</sup>, 石井 雅樹<sup>†</sup>

秋田県立大学<sup>†</sup>

## 1. 背景・目的

近年、人間の生活環境で共存することを目的としたロボットの開発が進んでいる。ロボットが人間と生活を共にするためには様々な変化に柔軟に対応し、作業を行うことが求められる。変化に柔軟に対応するための一機能として物体認識が挙げられる。従来の物体認識の手法では、入力画像に対して SIFT や SURF などを用いて特徴を抽出し、機械学習を行うことにより入力画像を識別している。また、近年では DeepLearning を用いた物体認識が行われている。DeepLearning では自動的に適切な特徴が抽出されるため、従来の物体認識よりも精度が高いことが特徴である。しかし、どちらの手法も汎化性を高めるために大量の学習データが必要となる。

人間は物体を認識する際、その物体の使われ方、すなわち、人と物体とのインタラクション(相互関係)を特徴として用いていると考える。本研究では人と物体との基本的なインタラクションを検出し、インタラクションが発生した際の相互関係(動作)を物体の属性情報として抽出し、その属性情報を物体認識の特徴として利用する手法の確立を目的としている。本稿では RGB-D センサーを用いて抽出した物体の形状と人間の動作を、物体の属性として付与する手法を提案する。

## 2. 提案手法

人と物体とのインタラクションとして本研究ではアフォーダンスの概念を用いた。アフォーダンスとは物体が人に情報を与え、その情報から人のどのような行動が誘発されるかといった考え方である。近年ではアフォーダンスの概念を物体認識に応用する研究が行われている[1]。物体が人に与える情報とは、その物体の本来の使用法であり、物体の形状によって表現されていると考える。この物体の形状を本研究では静的属性として定義した。また、その形状から実際に人が物体に対して行った動作を動的属性として定義した。本研究では、この二つの属性をもとに物体認識を行った。なお、本稿では対象物体を椅子と飲み物に限定した。また、RGB-D センサーとして Kinect for Windows v2(Microsoft 社)を用いた。

Attribute Definition Method Based on Human Action for Object Recognition Focused on Interaction

<sup>†</sup>Tatsuki Ishijima, Masaki Ishii : Akita Prefectural University

## 3. 形状認識

静的属性として物体の形状に着目した。例えば、椅子であれば平面と平面が垂直に交差していること、または人を支える座面を有すること、ペットボトルやマグカップであれば、液体を蓄える円柱形状を有することが挙げられる。すなわち、物体は単純な形状部分を有していることから、それらの特徴として用いることが有効と考えた。形状認識として椅子は座る平面形状の認識、飲み物は液体を蓄える円柱形状の認識を以下の手順で行った。

### (1) 平面認識

- ① 3次元点群取得
- ② クラスタリング
- ③ 各クラスターで RANSAC による平面認識

### (2) 円柱認識

- ① 点群を XY, YZ, ZX 平面に平行投影
- ② ZX 投影点群に対してノイズ除去
- ③ ZX 投影点群から凸包座標算出
- ④ 凸包座標から最小二乗法により円の半径を算出
- ⑤ 半径の閾値判定による円柱認識

## 4. 動作認識

動的属性として人の物体への動作に着目した。人が物体に対する動作は物体の形状から限定されると考えられるため、その動作を特徴として用いることが有効と考えた。本稿では認識動作として「座る」と「飲む」の二つの動作を対象とした。動作認識のための識別器は VisualGestureBuilder を用い、学習アルゴリズムには AdaBoost を選択した。「座る」動作の学習には 8 人分の動画像を用いた。また、「飲む」動作は右手に限定し、5 人分の動画像を用いた。テストデータは両者とも 1 人分の動画像を使用した。表 1 に認識結果を示す。どちらも RSME が 0 に近いので、動作認識に用いるには十分な精度と言える。よって、今後の物体認識の処理にも本章で作成した分類器を用いた。

表 1 動作認識結果

動作	適合率	RSME (Root Mean Square Error)
座る	0.97066	0.48801
飲む	0.96335	0.65591

5. 物体認識

形状認識による静的属性と動作認識による動的属性をもとに物体認識を行った。図 1, 2 に椅子の認識処理, 飲み物の認識処理の手順を示す。実験では認識対象として図 3 に示す 6 種の椅子と 10 種の飲み物を用いた。

図 4 に椅子, 飲み物として認識された点群を示す。椅子の認識では, 環境中の床という情報を用いて人が座れる平面を特定でき, 最終的に人が座ることで椅子を認識できた。飲み物の認識では, 人の骨格情報から手先の領域を抽出することで手によるオクルージョンを防いでいる。また, 椅子の認識と同様に床の情報を用いて飲み物が置かれている平面を除去することで飲み物の点群のみを分割できる。したがって, 分割した点群に対して円柱認識を行い, 最終的に人が飲む動作を行うことで飲み物の認識ができた。なお, 物体の本来の使用がなされない場合は認識が行われず, 誤認識が少ないことを確認している。

6. まとめ

本研究では物体の形状部分を静的属性, 人が物体に対して行った動作を動的属性として定義し, 二つの属性を用いた物体認識を試みた。実際に物体の使用目的を満たす形状部分を認識し, それに対して人間が行った動作を物体に付与することで椅子と飲み物の認識が可能であることを確認した。今後は静的属性, 動的属性に加え, 物体そのものが有するカラーや法線情報を組み合わせた物体認識手法について検討する予定である。

参考文献

[1] 秋月秀一, 飯塚正樹, 橋本学: 「アフォーダンスに着目した一般物体認識のための特徴量」, 第 21 回知能メカトロニクスワークショップ講演論文集, pp94-96(2016)

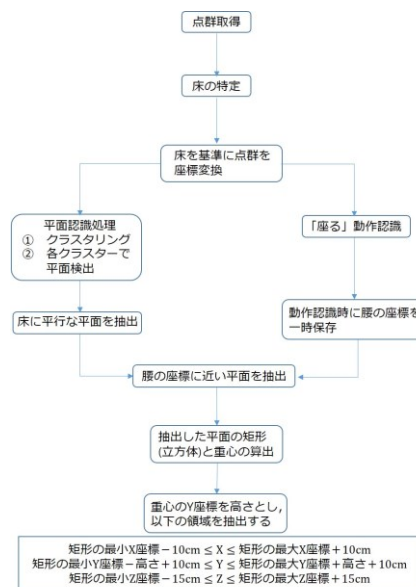


図 1 椅子認識処理手順

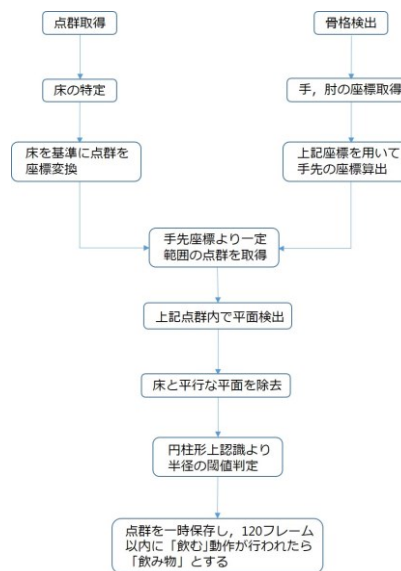


図 2 飲み物認識処理手順

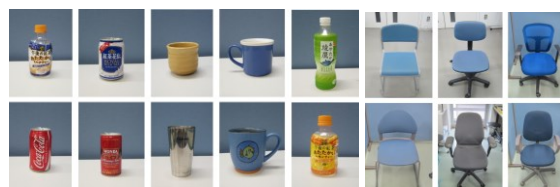


図 3 認識対象物体

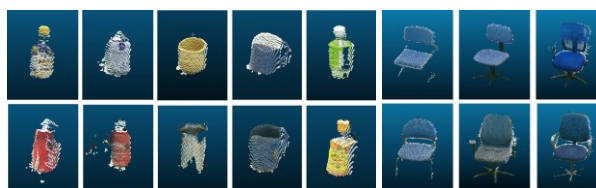


図 4 認識結果