

移動ロボットによる教師なし学習に基づく一般物体認識の研究

Generic Object Recognition Based on Unsupervised Learning Using Mobile Robot

2011017



研究代表者	秋田県立大学 システム科学技術学部	助教	間 所 洋 和
共同研究者	SmartDesign 株式会社	代表取締役	高 橋 淳 一

[研究の目的]

近年、人と生活し活動する人間共生型ロボットが、身近な存在になりつつある。このようなロボットにとって、環境中に存在する多種多様な一般物体を認識することは、人間とロボットの知覚レベルでの調和において必要不可欠な機能である。一般物体認識はコンピュータビジョンの分野で活発に研究されているが、ロボットへの応用を考えた場合、教示に伴う負荷が増大する。そこで本研究では、教師なし学習に基づく一般物体認識法を提案し、視覚センサを搭載した移動ロボットによる実環境での評価を目的とする。教師なし学習により、ユーザの負荷が軽減できるとともに、ロボットとのインタラクションを通じた、人間と機械の調和の実現を目指す。

[研究の内容、成果]

1. 教師なし学習に基づく物体認識法

一般物体認識において、画像の特徴表現から識別器の構築までを、単独の理論や手法で扱うことは困難な課題と言われている。このため本研究では、各段階の処理に対して複数の手法を組み合わせることにより、教師なし学習に基づく一般物体認識法を構築する。提案手法のネットワーク構成を図1に示す。

処理手順は6段階から構成され、1) SIFT

(Scale-Invariant Feature Transform) による特徴点の抽出と特徴量の記述, 2) OC-SVM (One Class-Support Vector Machines) による特徴点の選択, 3) SOM (Self-Organizing Maps) を用いた VW (Visual Words) と画像ヒストグラムの作成, 4) ART-2 (Adaptive Resonance Theory-2) を用いたラベルの生成, 5) CPN (Counter Propagation Networks) を用いたカテゴリマップの形成, 及び 6) U-Matrix を用いたカテゴリ境界の抽出となる。

画像の特徴表現に相当する 1)~2) は BoF (Bag of Features) に基づいている。SIFT の特徴点は OC-SVM により 2 分割され、物体が含まれる注目領域が絞り込まれる。VW の作成には、k-means ではなく SOM を用いる。ART-2 と CPN を組み合わせることにより、大

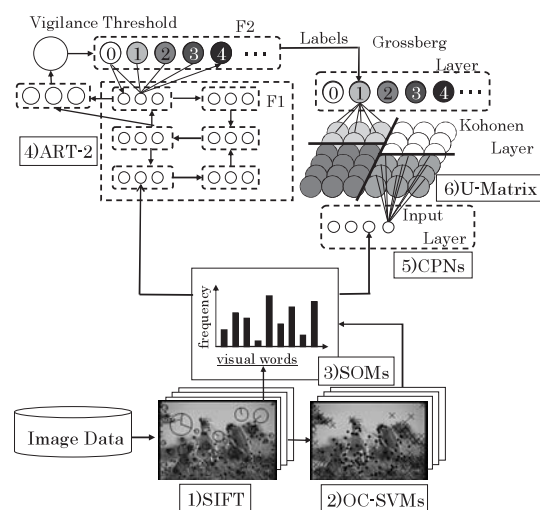


図1 提案手法の全体構成

量の画像から自動でラベルを生成するとともに、カテゴリマップを形成する。最後に、U-Matrix を用いてカテゴリマップ上から境界を抽出し、カテゴリ数を同定する。

2. 標準データセットを用いた評価実験

一般物体認識においてベンチマークテスト用の標準データセットとして広く普及している Caltech256 を用いて、提案手法の基本的特性を評価した。本実験では、256 種類のカテゴリから、データ数が多い上位 20 カテゴリを評価対象とした。各カテゴリからランダムに 5 枚の画像を取り出し、合計 100 枚の画像を用いて学習を行った。

CPN によるカテゴリマップの生成結果を図 2 に示す。カテゴリマップ上には、ユニットに写像された代表画像を表示している。5 カテゴリにおける分類では、同一カテゴリの画像は隣接したユニットに写像された。また、独立した領域が形成された。20 カテゴリにおける分類では、11 カテゴリの画像が隣接したユニットに写像された。また、ART-2 では複数のラベルに分割されていたカテゴリが、CPN では隣接したユニットに統合された。

続いて、5, 10, 20 カテゴリにおける学習とテストの認識率を図 3 に示す。ここでは、OC-SVM を適用した場合と未適用の場合を比較することにより、特徴点選択の有用性を検証した。OC-SVM を適用した場合は、学習データが 96%, 94%, 81%, テストデータが 76%, 42%, 45% となった。一方、未適用の場合は、学習データが 84%, 70%, 64%, テストデータ

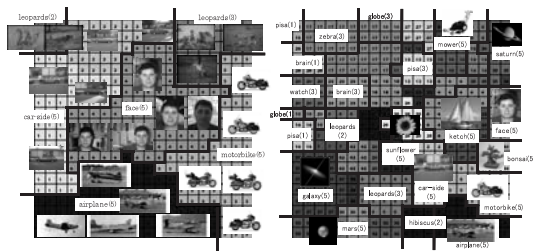


図2 カテゴリマップ
(左: 5 カテゴリ, 右: 20 カテゴリ)

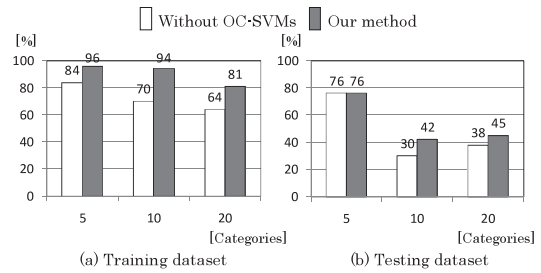


図3 認識率の比較結果

が 76%, 30%, 38% であった。以上の結果から、学習データでは 18%, テストデータでは 6% の精度向上がみられ、OC-SVM による特徴点選択の有用性が示された。

3. 移動ロボットの試作

視覚に基づく自律移動と物体認識のためのプラットフォームとして、図 4 に示す視覚機能を搭載した移動ロボット「MEGURI」を試作開発した。設計指針としては、人間と同じ目線で物体認識や環境理解が実現できるように、ロボットの全高を成人男性の平均身長と同程度の 170 cm にした。

ロボットの上部には、ステレオ視のためのカメラ Bumblebee2 と、全方位センサ VS-C14U を搭載した。また、下部に深度センサ XTION を搭載し、超音波センサとともに障害物検知を

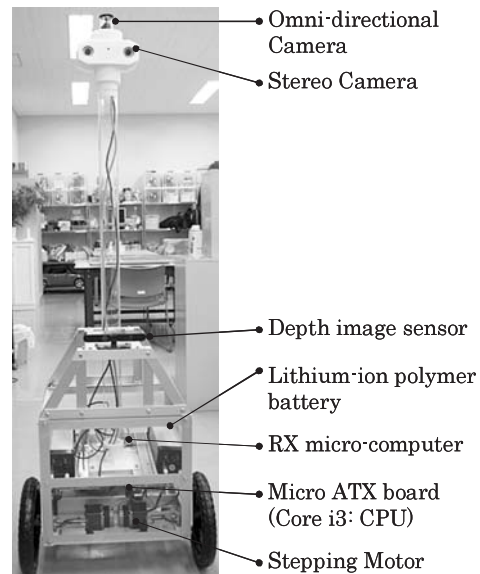


図4 試作した移動ロボット「MEGURI」

担う。駆動系は、ステッピングモータ 2 機による 2 軸独立駆動とした。また、安定性を考慮して補助輪は後部に 2 輪設置した。

計算機環境としては、モータの制御用に RX マイコン、画像処理用に CPU ボードを搭載した。データのストレージには、消費電力と耐振性を考慮して SSD を搭載した。電源は 20Ah のリチウムイオンポリマ電池を 2 機搭載し、約 2 時間の連続走行が可能となっている。ロボットの重量は約 20 kg、最高移動速度は 4.0 km/h である。

4. 移動ロボットを用いた評価実験

試作したロボットを用いて一般物体認識の評価実験を実施した。実験環境は、研究代表者が所属する研究室とした。2.5×2.5 m の格子状の区画に、4 カテゴリーの物体を配置し画像を取得した。認識対象とした物体は、事前にアンケートを実施し、研究室内で一般的な物体として選ばれた Chair, PC, Robot, Book の 4 物体とした。

ロボットは物体に向かって移動し、画像を取得する。なお、物体の大きさを考慮して、画像を取得はロボットの本体に搭載した XTION の RGB カメラを用いた。各物体の大きさは不揃いのため、画像中の物体が小さく、あるいは大きくなり過ぎないようにロボットの行動を制御した。また、視野内の認識対象は 1 物体とした。ロボットは環境内を 5 回走行し、走行毎に同一カテゴリの異なる物体に入れ替えた。このため、各カテゴリは異なる 5 種類の物体から構成される。これをデータセット D1～D5 として評価対象とした。取得した画像を図 5 に示す。

認識率の評価は、交差検定法によって全ての組み合わせに対して実施した。Caltech-256 などのコンピュータビジョン分野における一般物体認識用のデータセットでは、対象物体は 1 画像中に 1 物体のみであり、画像中央に位置するように撮影されている。したがって、本実験においても画像中に認識対象となる物体全体が写

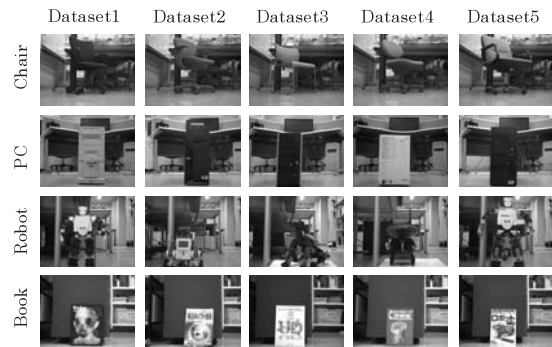


図 5 ロボットから取得した画像

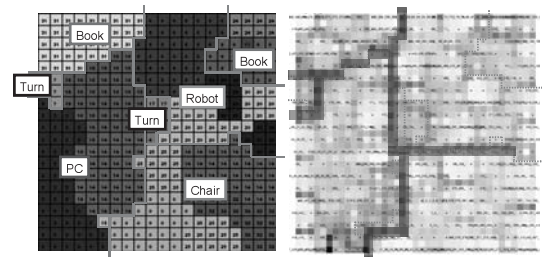


図 6 カテゴリマップと U-Matrix の結果

表 1 認識率

Dataset	Chair	PC	Robot	Book	ALL
D1	100	100	100	58	91
D2	97	96	100	88	96
D3	97	89	78	92	90
D4	100	100	74	88	92
D5	100	96.4	100	83	96
Ave.	99	96	90	82	93

るように制約条件を設けた。

認識率を表 1 に示す。なお、旋回時の画像はテスト対象から除いて認識率の算出し評価した。平均認識率は、Chair が 99%、PC が 96%、Robot が 90%、Book が 82% となった。また、データセット別では、D1 が 91%、D2 が 96%、D3 が 90%、D4 が 92%、D5 が 96% であった。全テストデータの平均認識率は、93% となった。

各データセットにおける混同対照表を表 2 に示す。横方向を基準として、識別に成功した画像枚数が対角線上に表示される。それ以外には、誤識別となった枚数が表示される。縦方向のカテゴリ名を参照することで、誤識別したカテゴリを特定することができる。最も低い認識率となった D1 の Book は、9 画像が Robot と

表 2 混同対照表

	Chair	PC	Robot	Book	Chair	PC	Robot	Book	Chair	PC	Robot	Book
Chair	38	0	0	0	37	0	1	0	37	0	1	0
PC	0	27	0	0	0	27	1	0	0	24	1	1
Robot	0	0	27	0	0	0	27	0	3	0	21	3
Book	1	0	9	14	0	0	3	22	0	0	2	22

Dataset1 Dataset2 Dataset3

	Chair	PC	Robot	Book	Chair	PC	Robot	Book
Chair	38	0	0	0	38	0	0	0
PC	0	28	0	0	1	27	0	0
Robot	2	5	20	0	0	0	27	0
Book	1	0	0	21	0	0	4	20

Dataset4 Dataset5

誤認識された。他のデータセットでは、D2で3画像、D3で2画像、D5で4画像がRobotと誤認識された。次に認識率の低かったD4のRobotは、2画像がChairと、5画像がPCと誤認識された。D3では、3画像がChair、3画像がBookと誤認識された。D1とD5のBookの表紙には、顔のイラストやロボットパーツのイラストが描かれていたため、Robotと誤認識されたと考えられる。

次に、見え方の変化に対するロバスト性を検証した。D5のChairに対して向きの異なる画像を取得し、D5a及びD5bとした。両データセットの画像を図7に示す。OC-SVMにより選択された特徴点は、向きが異なっても対象となる物体を中心に同心円状に分布している。背景領域は、椅子の背もたれ位置が異なるため、特徴点の分布に違いが現れた。

認識率は、D5aは100%であったものの、D5bは42%に低下した。D5bにおける誤認識は、Robotのラベルが影響を与えていると考えられる。ロボットは、D1～D4を学習データとした場合のCPNによるカテゴリマップで、Chairと隣接したユニットに写像された。本実験では、ロボットの行動パターンが直進と90度の右旋回に限られており、前後進以外の見え方の変化を学習していないため、認識率が低下

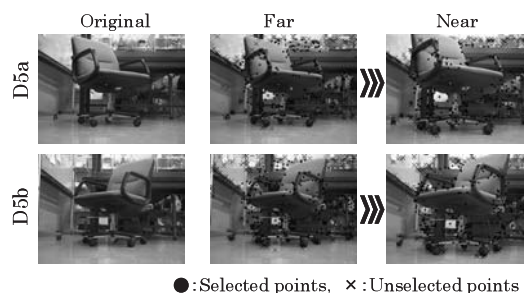


図 7 ロバスト性検証用の画像

したと考えられる。そのため、より柔軟な物体認識を実現するには、ロボットによる物体認識のための多様な行動様式の生成が必要であると考えている。

[今後の研究の方向性, 課題]

本研究では単一物体のみを対象としたため、今後は複数物体認識への拡張について提案手法の改良を進めたい。具体的には、画像中から複数の物体を検出するために、画像の輝度勾配や方位選択性等の特徴空間から視覚的に注意を引く領域を表現する顕著性マップの適用を検討している。更に、ロボットビジョンにおけるリアルタイム処理を実現するためのハードウェア環境の構築し、本試作ロボットに実装するとともに、実環境での評価実験を進めたい。また、本ロボットを情景分類にも応用したいと考えている。

[成果の発表, 論文等]

- [1] H. Madokoro, M. Tsukada and K. Sato: Unsupervised Feature Selection and Category Formation for Generic Object Recognition, Proc. 14th International Conference on Computer Analysis of Images and Patterns (CAIP), pp. 427-434, Aug. 2011.
- [2] H. Madokoro, M. Tsukada, and K. Sato: Unsupervised Feature Selection and Category Formation for Mobile Robot Vision, Proc. IEEE International Joint Conference on Neural Networks (IJCNN), pp. 320-327, Jul. 2011.