

把持動作による物体カテゴリの形成と認識

中村 友昭 † 西田 匡志 † 長井 隆行 †

† 電気通信大学 電気通信学研究科 電子工学専攻

1 はじめに

近年、物体カテゴリの教師なし学習が盛んに研究されている。このようなカテゴリを認識する能力は、未知物体を扱うために非常に重要である。さらに、人の円滑なコミュニケーションを行う上でも、人間の感覚に近いカテゴリを形成することが重要となる。現在、これらの研究は主に視覚情報を対象としており、色やテクスチャ等の情報を用いている。しかし、実世界にはこれらの視覚情報のみではカテゴリを形成できないものも数多く存在する。そこで本稿では、五感の1つである触覚に着目する。人は物体に触れただけで、その物体から様々な情報を取得しており、さらにその物体のカテゴリを認識することが可能である。ロボットがこのような能力を持つことで、今まで視覚情報のみでは形成できなかった物体カテゴリを得ることが可能となると考えられる。

本稿では3本指のロボットハンドに取り付けられた触角アレイセンサーを用いて、物体の把持により得られるセンサー情報から物体を分類する手法を提案する。提案手法は、物体把持時のセンサーの時系列データの近似を行い、その近似パラメータを各センサーの特徴ベクトルとして扱う。さらに、この特徴ベクトルを Bag-of-Features モデルとして扱うことで、把持位置に依存しない物体の特徴量を得ることができる。最終的にこれらの特徴量を教師なしで分類することで、物体カテゴリの形成を行う。分類には統計的手法である、Latent Dirichlet Allocation(LDA)[1]を用いる。提案手法の特徴として、汎用的なロボットハンドと触覚センサーを用いているため、様々なロボットへの適用が可能であること、またハンドの複雑な制御を必要としないため実装が容易である点等が挙げられる。

このような触覚情報を用いたカテゴリ分類の研究として、物体を周期的に振ることで発生する音声信号を基に分類を行っているものがある[2]。しかし、このシステムでは特殊なハードウェアが必要となり、汎用的なロボットアームやハンド等への適用は今のところ行われていない。さらに、触覚情報から形状や物体の認識を行う研究も行われている。文献[3]では、自己組織化マップを用いて、触覚情報を用いた物体の分類を行っている。文献[4]では、形状特徴のパターンを予め定義し、その出現頻度から物体の大きさや形状の認識を行っている。また、文献[5]では、触覚センサー上で物体を動かし、その触覚情報のパターンの変化で物体の形状の認識を行っている。しかし、これらの研究では、把持位置や把持の方法に制約があり、把持状態が変化すると正しく認識ができないといった問題がある。把持位置に依存しない物体認識の研究として、文献[6]がある。触覚情報のパターンを Bag-of-Features モデルとして扱うことで、把持状態に関わらず安定した物体の認識を実現している。しかし、この研究では、教師ありで物体の学習・認識を扱っており、物体のカテゴリ分類は行っていない。

Object Categorization and Recognition by a Grasping.
†Tomoaki NAKAMURA †Masashi NISHIDA †Takayuki NAGAI
†The university of electro-communications.

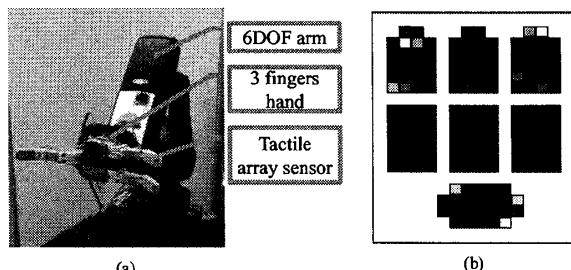


図 1 (a) ロボットの構成 (b) 触覚アレイセンサーの出力

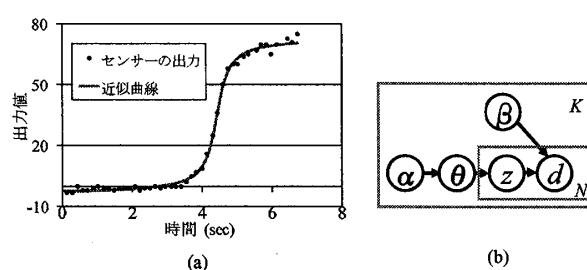


図 2 (a) センサーの出力と近似曲線 (b)LDA のグラフィカルモデル

2 提案手法

2.1 触覚センサー

今回実験で用いたロボットの構成を図 1(a)に示す。ハンドは、6DOF のアームの先に取り付けられ、アームを動かしハンドを物体位置まで移動させることができ。ハンドは、3 本の指を持ち、物体を把持することができ、各指の先と根元、さらに手の平には触覚アレイセンサーが取り付けられている。触覚アレイセンサーは、合計 162 個のセンサーから構成されており、各センサーに加わる圧力を取得することができる(図 1(b))。

2.2 触覚情報処理

ハンドを一定速度で閉じた際の、時系列の触覚アレイセンサーの各センサーの出力値を使用する。しかし、このままではデータ数も多く直接把持物体の特徴を得ることができないため、時刻 t でのセンサーの出力値を $p(t)$ とし、各センサーの時系列の出力値を以下の式へと近似を行う。

$$p(t) = a \tan^{-1} (b(t+c)) + d \quad (1)$$

センサーの出力値と、その近似曲線の例を図 2(a)に示す。この近似により、各素子から得られる時系列の変化を (a, b, c, d) という 4 つのパラメータで表現することができる。このうち、 b は物体に触れてから停止するまでの時間と、 a はその際の物体に掛かっている力と関係していると考えられる。また、 c は把持する物体の大きさに、 d は a に依存した値となるため、この 2 つの値を取り除き、 (a, b) の 2 つのパラメータのみを、各センサーの情報として用いることとする。最終的に、1 回の把持で 162 個の (a, b) が得られることになる。

2.3 Bag-of-Features モデル

Bag-of-Features モデルとは、物体がある特徴の発生回数のパターンで定義するものである。このモデルは、画像のカテゴリ分類で多く利用されており、局所的な特徴をその位置に関係なく、その発生回数のパターンを画像全体の特徴としている。この考え方を応用し、本稿では、触覚情報を Bag-of-Features モデルとして扱う。1つの物体を把持した際に、前節で計算した2次元のセンサーの特徴ベクトルが 162 個得られる。この2次元の特徴ベクトルを予め計算した 20 個の代表ベクトルを用いて、ベクトル量化を行う。最終的に、その 20 次元のヒストグラムを物体の特徴量として扱う。代表ベクトルは、多数の物体を把持することで得られる 2 次元の触覚特徴量を k-means により分類することで予め計算する。このように、触覚情報を Bag-of-Features モデルとして扱うことにより、多数・多次元の情報が、把持位置や形状に依存しにくい、20 次元の特徴量となる。

2.4 LDA による学習と認識

LDA は自然言語処理の分野において、文書の自動分類を行うための確率モデルであり、これを触覚特徴量の分類に利用する。図 2(b) が LDA のグラフィカルモデルである。物体のカテゴリ z は、ディリクレ事前分布 α によって決まる多項分布 θ によって決定され、物体の触覚特徴量 d は、多項分布 β から発生するモデルとなっている。 N は物体の数を、 K はカテゴリ数をそれぞれ表している。物体の分類は、このグラフィカルモデルのパラメータを学習することに相当する。LDA のモデルでは、与えられた学習データ集合に対する対数尤度が最大となる α と β を VB-EM アルゴリズムにより決定することになる。さらに、学習したモデルを用いることで、未学習物体の認識も可能となる。物体の触覚特徴量 d の、物体カテゴリ z は以下のように求めることができる。

$$\hat{z} = \operatorname{argmax}_z p(z|d) = \operatorname{argmax}_z \int p(z|\theta)p(\theta|d)d\theta \quad (2)$$

但し、 $p(\theta|d)$ は、学習した β を固定し、前述の VB-EM アルゴリズムを適用することで、 α を再計算することにより求める。

3 実験

実験では、図 3 に示した硬さの異なる 5 カテゴリ 24 個の物体を用いた。

3.1 学習

学習では、各物体を様々な方向から 5 回把持することで、前述の通りヒストグラムを作成し、LDA により分類を行った。作成されたヒストグラムの一部を、図 4 に示す。各カテゴリに共通した特徴が現れていることが分かる。分類結果は、図 3 に示した通りに分類された。このように、ぬいぐるみとゴム人形のように視覚特徴が類似している物体であっても、触覚情報を使用することで正しく分類することができる。また、ぬいぐるみのように形が様々に異なる物体や、ゴムボールのように大きさが異なる物体であっても、その触覚情報の類似性から正しくカテゴリを形成できていることが分かる。

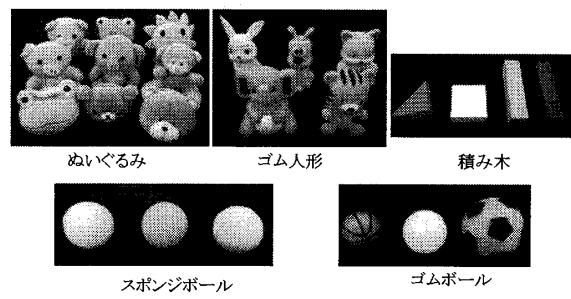


図 3 5 カテゴリ、25 個の物体

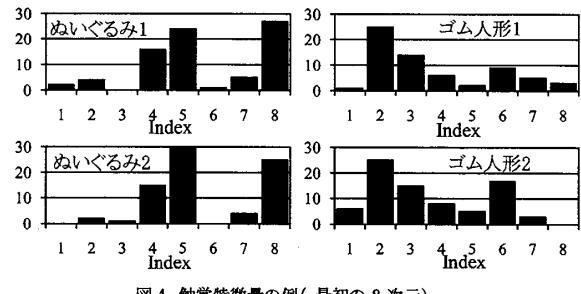


図 4 触覚特徴量の例(最初の 8 次元)

3.2 認識

24 個の物体のうち 23 個で学習を行い、残りの 1 つの物体の認識を行った。学習は前節と同様に行い、認識する物体を 1 回把持し、その触覚情報のみから認識を行った。以上の操作を 1 つの物体につき 5 回行うことで、合計 120 回の認識を行った。その結果、認識率は 91.7% となり、未学習の物体であっても正しく認識することが可能である。失敗は、主に学習時に局所解に陥ってしまうことが原因であった。

4 まとめ

本稿では、多次元の触覚の時系列情報を Bag-of-Features モデルとして扱うことで、把持位置・形状・大きさ等に依存しにくい触覚特徴量を計算した。さらに、それを統計的手法である LDA で学習することで、教師なしで触覚情報から物体のカテゴリの形成が可能であることを示した。今後の課題として、物体カテゴリの分解能の検証や、カテゴリ数・物体数を増やすこと等を考えている。

参考文献

- [1] D.Blei, "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, pp.993-1022, 2003.
- [2] 高橋慎也、細田耕、浅田稔“ダイナミックタッチに基づく物体カテゴリの形成”ロボティクス・メカトロニクス講演会'07, IA2-L09, 2007.
- [3] Lorenzo Natale, Giorgio Metta, and Giulio Sandini. "Learning haptic representation of objects." *Proc. of International Conference on Intelligent Manipulation and Grasping*, 2004.
- [4] 市川純章、原文雄“ロボット指による物体識別と人の識別能の比較”第 26 回日本ロボット学会学術講演会, 1E2-04, 2008.
- [5] R.Andrew.Russell. "Object recognition by a 'smart' tactile sensor." *Proc. of the Australian Conference on Robotics and Automation*, pp.93-98, 2000.
- [6] Alexander Schneider, Jurgen Sturm, Cyrill Stachniss, Marco Reisert, Hans Burkhardt, Wolfram Burgard. "Object Identification with Tactile Sensors using Bag-of-Features" *Proc. of International Conference on Intelligent Robots and Systems*, Molt10.3, 2009.