

迅速なマルチモーダルカテゴリ推定のための能動的行動選択法

Active Perception for Quick Multimodal Category Recognition

2041018



研究代表者

立命館大学 情報理工学部

准教授

谷口 忠大

[研究の目的]

ロボットが人間と共生するには、人間と似た物体概念をロボットが獲得する必要がある。マルチモーダルカテゴリゼーションはロボットが視覚、触覚、聴覚といった複数のモダリティ情報を統合し、物体概念を獲得し、認識するための理論である。Nakamuraらは一般性の高いマルチモーダルカテゴリゼーション手法である、マルチモーダルLDA及びマルチモーダルHDP(MHDP)を提案し、視覚、聴覚、触覚を持ったロボットが人間と極めて類似した物体カテゴリを形成しうることを示している。

しかし、先行研究ではロボットが物体の認識のプロセスにおいて、全てのモダリティ情報を取得することを前提としている。そのため、大規模な物体概念獲得と認識には、多大な時間がかかってしまう。マルチモーダルカテゴリゼーションでは、ロボットのモダリティ数が M であり、物体数が N であった時、全ての情報取得にかかる時間は $O(MN)$ であり、より多くのモダリティを持ち、より多くの物体を扱おうとするほど、多くの膨大な「情報取得時間」が求められる。

この情報取得時間を、短縮することができれば、ロボットによる人間の住環境下における様々な物体のマルチモーダルカテゴリゼーションをより現実的なものとする事が出来る。これは、ロボットが人間生活環境で共生する系を

考える上で重要な課題である。図1にタスクの概要図を示す。

マルチモーダルカテゴリゼーションについてはNakamuraやStoytchev, Takamukuなど世界的にもいくつかのグループの研究があるが、いずれの研究もロボットに与えるセンサ情報は固定しており、マルチモーダルカテゴリゼーションにおける知覚の行動選択を扱っている研究は極めて少ない。本研究は知覚には行動を必要とするという実世界的な制約を組み込み、問題としている点で特徴的である。

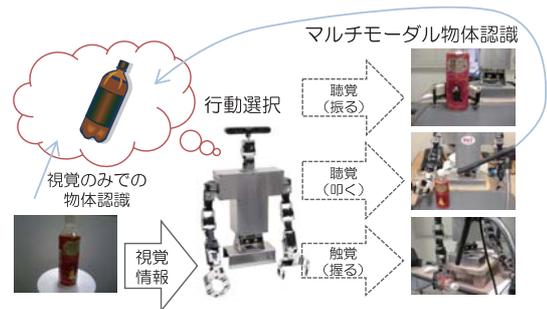


図1 タスクの概要図

[研究の内容, 成果]

本研究ではまず、図2に示すようなマルチモーダルカテゴリゼーション及びマルチモーダルカテゴリ推定のためのロボットシステムを構築した。ロボットは視覚情報、聴覚情報(叩いた時の音、振った時の音)、触覚情報を取得す

ることが出来る。ロボットはそれぞれの感覚情報から特徴量を抽出し、それをMHDPに渡す。図3に本実験で取り扱った認識対象の物体一覧

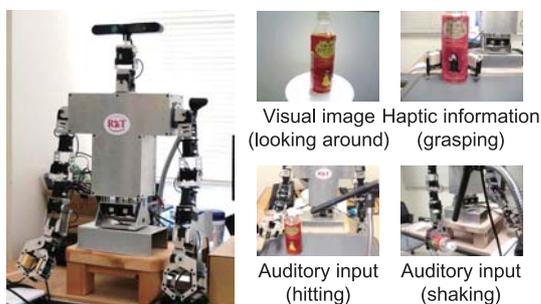


図2 構築したロボットシステム

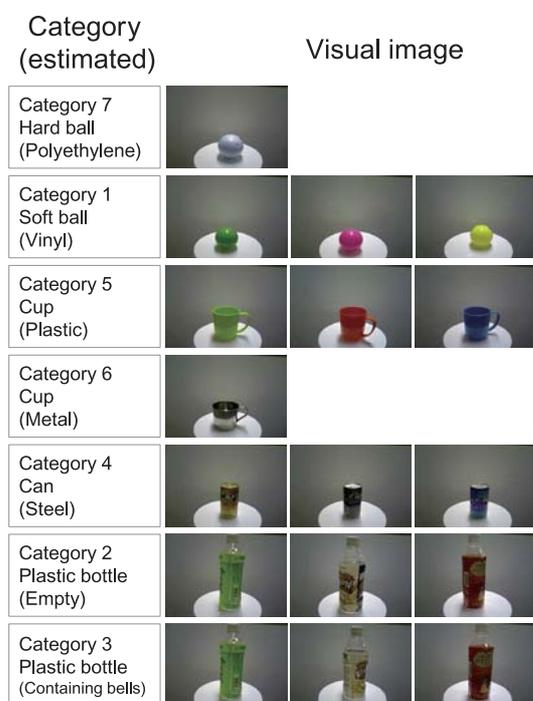


図3 認識対象物体の一覧

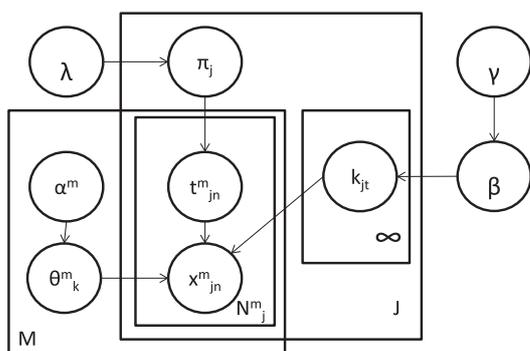


図4 MHDPのグラフィカルモデル

を示す。カテゴリ2 (空のペットボトル) とカテゴリ3 (ベルの入ったペットボトル) が示すように、視覚情報だけでは区別出来ない物体群になっている。

図4に一般化されたMHDPのグラフィカルモデルを示す。これはM個のモダリティを持つロボットのマルチモーダルカテゴリゼーションを表すグラフィカルモデルになっている。ここで、知覚のためのモダリティ数と知覚のための行動数は同一であるとしている。図3に示すカテゴリはロボットが推定したものであり、人間の区別と大筋で一致していることがわかる。

研究成果としては、まず、迅速なマルチモーダルカテゴリ推定のための能動的行動選択法として、物体カテゴリ事後分布のカルバックライブラー情報量 (KL 情報量) の期待値を最大化させるモダリティを選択する手法を提案した (成果発表1)。KL 情報量は確率分布間の距離を表す値である。提案した尺度は情報利得 (Information Gain) と等価である。また、そのような行動選択が全てのモダリティの情報を得た後の認識状態 (カテゴリの事後分布) とのKL 情報量の期待値を最小化させることを示した。これはある意味で本手法が最適であることを示している。

しかし、(成果発表1)では、次の一つの行動についてのみ扱っており、複数の行動を逐次的に取ることで、より確かな認識を得てくプロセスについて議論できていなかった。(成果発表2,3)ではMHDPのグラフィカルモデルを前提とした際には、情報利得が劣モジュラ性を満たすということから、効率的な逐次意思決定手法を提案した。しかし、これらの手法では、情報利得の計算に、複数回の繰り返し学習を必要としており、計算効率上、非効率な面があった。

そこで、(成果発表4)ではモンテカルロ近似を利用した効率的な逐次行動選択を提案した。これにより計算量を劇的に削減することに成功

した。

これら全ての成果を統合し、学術論文として(成果発表5)に投稿中である。

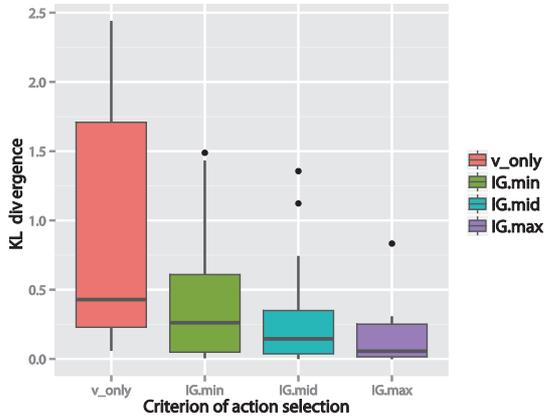


図5 各行動選択後の KL 情報量

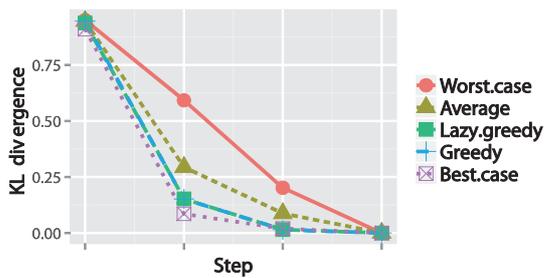


図6 逐次行動選択と KL 情報量

図5は各行動選択後の KL 情報量を示しており、v_onlyとした視覚情報のみを得た状態に比べて、情報利得最大の行動をとった場合、その最終的な認識状態までの KL 情報量が最も小さくなっていることが分かる。

図6は提案した能動的な逐次意思決定法の Greedy 法と Lazy greedy 法の性能を最適な行動 (Best case), 平均的なパフォーマンス (Average), 最悪な行動 (Worst case) の性能と比較している。全通りの試行を行い最適な行動を事後的に選んだ Best case に比べても提案手法は遜色ない性能を示していることが分かる。

この他にも人工データを用い、20以上のモダリティ数を持ったロボットのマルチモーダルカテゴリ推定についても実験を行った。これら

の実験を通じて、本手法はむしろ、そのような大規模なデータになればなるほど、顕著に有効であることが確認された。この視点より、スケラビリティに関しても問題はない。

[今後の研究の方向, 課題]

本研究期間を通して、迅速なマルチモーダルカテゴリ推定のための能動的行動選択法のための理論を確立することが出来たと考えている。また、マルチモーダルカテゴリゼーションを行うグラフィカルモデルの性質と情報利得を用いた能動的行動選択の間には極めて良好な関係があることも示唆される結果を得た。

一方で、マルチモーダルカテゴリゼーションを行う自律ロボットを家庭やオフィスに導入する際に、最も、時間がかかるのは、本研究で扱った認識プロセス以上に、学習プロセスである。この時間削減も同様に求められる。理論的洞察に基づけば、一見全く異なる学習プロセスと認識プロセスも隠れ変数の推論という視点から見れば非常に類似しており、本研究期間を通して確立された迅速なマルチモーダルカテゴリ推定のための能動的行動選択法のための理論を拡張することで、構築可能であると考えている。

マルチモーダルカテゴリゼーションにおける能動学習理論を構築し、実環境において、ロボットが自律的に物体概念学習を行える理論的整備を行いつつ、その、実世界での有効性検証を行うのが今後の課題である。

[成果の発表, 論文等]

1. 谷口忠大, 福田 一, 高野敏明, マルチモーダルカテゴリ能動知覚のための行動選択法と最適性, 2014年度人工知能学会全国大会, (2014)
2. 高野敏明, 谷口忠大, 劣モジュラ性に基づくマルチモーダルカテゴリ能動知覚の逐次的意思決定, 第32回 日本ロボット学会学術講演会, (2014)
3. 高野敏明, 谷口忠大, マルチモーダルカテゴリゼーションに基づく逐次的能動知覚, 計測自動制御学会

- システム・情報部門学術講演会 2014 (SSI2014). (2014)
4. 吉野 遼, 高野敏明, 谷口忠大, マルチモーダル能動知覚におけるモンテカルロ近似を利用した効率的な逐次行動選択, 第33回 日本ロボット学会学術講演会. (2015)
 5. Tadahiro Taniguchi, Toshiaki Takano and Ryo Yoshino, Active Perception for Multimodal Object Category Recognition Using Information Gain, IEEE Transactions on Autonomous Mental Development. (submitted)