

# 特徴点の依存関係ネットワークにもとづいたロボットの状態識別

State segmentation for robot based on dependency network of image features

小林 祐一<sup>\*1\*2</sup> 染井 貴之<sup>\*3</sup> 金子 透<sup>\*1</sup>  
 Yuichi Kobayashi Takayuki Somei Toru Kaneko

<sup>\*1</sup>静岡大学 Shizuoka University <sup>\*2</sup>理化学研究所 RTC RIKEN RTC <sup>\*3</sup>(株)ユーシン精機 Yushin Precision Equipment Co., Ltd.

This paper presents a construction of state space for a robot without predefined knowledge on its environment. Image features (keypoints) are clustered based on statistical dependencies with respect to their motions and occlusions. Estimation of conditional probability is used to evaluate statistical dependencies among configuration of robot and features in images. It was confirmed that part of the robot body was autonomously extracted without any a priori knowledge using conditional probability. In the generation of dependency network, different structures of networks were constructed depending on position of the robot hand relative to an object.

## 1. はじめに

ロボット・環境のモデルを事前に設定・解析することなくロボットの試行錯誤により適切な行動を生成させるアプローチは、強化学習のロボットへの適用や認知発達ロボティクス分野 [Lungarella 03] などで活発に研究されている。認知発達の視点からロボット知能を構築するにあたっての重要な課題の一つは、ロボットが動作を計画・生成・制御するための空間をどのように定義するかということである。動作計画・制御には、ロボットの関節角度や物体の世界座標系における位置・姿勢などから構成される Configuration 空間を用いることが一般的である。しかし、Configuration 空間の定義にもとづいたロボット制御の構築は、

- 状態変数の設定にタスクや環境に関する暗黙の仮定・知識が含まれる
- 設定された状態変数が常に正確に観測可能とは限らず、動作計画・制御に適しているという保証がない

という理由から、柔軟な知能を構築する上で必ずしも良い方法とは言えない。

認知発達ロボティクス研究には、ロボット自身の運動や知覚にもとづいて身体スキーマ [Stoytchev 08], 物体のクラス分類 [Sinapov 08], 身体と環境の接触 [Kobayashi 12] などを学習する手法が提案されている。一方、身体に関わる識別や物体の挙動に関わる識別を含んだより一般的な枠組みとして、身体・環境を含んだ系を一括して扱うボトムアップ型の識別法があれば、さらに設計者による事前知識の埋め込みを低減したロボットの動作生成法の構築に貢献することができる。

また、上記の関連する身体像・物体学習法においては、ロボット身体や物体が視覚中で完全に見えていることを前提としている。しかし実際の動作においては、身体・環境の一部が互いに隠されて見えない状態が必然的に発生する。このような遮蔽の問題を陽に扱える学習法が求められる。

そこで本研究では、ロボットの 3 次元の運動を扱い、ロボット身体と環境の間で奥行き方向の位置関係により遮蔽が発生するという前提のもとで、ロボットが環境に関する前提知識を

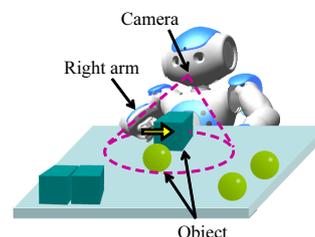


図 1: Robot and environment including objects

用いずに状態を識別するボトムアップな学習法の枠組みを提案する。

## 2. 問題設定

本研究で扱う問題設定を図 1 に示す。ロボットは単一のカメラと位置制御可能なアームを持つ。テーブル上に複数の物体があり、カメラはロボットアームとを視野に収める。アーム手先はカメラ光軸方向の運動を含んだ 3 次元空間で運動する。その動作に際し、アーム手先と物体が接触する場合、アーム手先が物体に隠される場合、物体がアーム手先に隠される場合があるものとする。手先と物体が接触する場合は、物体は手先の運動にともなって動く。

ロボットに物体の位置関係や形状、また身体の見え方といった視覚情報の知識を事前に与えられない。ロボットの目的は、物体・身体の動き方や隠れ方の傾向をもとに系の挙動を識別することである。そのために、画像内の特徴点の動き方および遮蔽についての依存関係を利用してネットワークを構成する。本研究では特徴点の抽出に SIFT [Lowe 99] を用いる。

## 3. 特徴の依存ネットワークとクラスタリング

アーム関節角度に微小な変位を与えてアームを動かす、その前後の画像を記録する。SIFT 特徴点の抽出と運動前後の画像間でのマッチングを行う。特徴点の間の依存関係を表すネットワークを 2 種類構成する。

運動に関する依存ネットワークの構築 身体運動指令に対して運動する特徴点同士をリンクで結ぶネットワークを構成する。

連絡先: 小林祐一, 静岡大学, 静岡県浜松市中区城北 3-5-1, tykobay@ipc.shizuoka.ac.jp

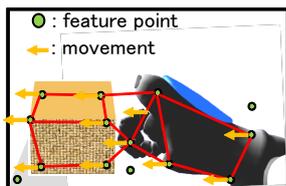


図 2: 運動ネットワーク

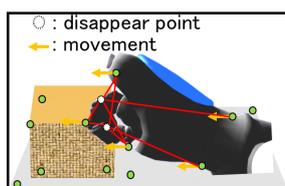


図 3: 遮蔽ネットワーク

画像中で身体のみが運動している間は身体上の特徴点がネットワークを構成し、身体と物体とが運動している時は図 2 に示すように身体上の特徴点と物体上の特徴点とがネットワークを構成する。特徴点  $j$  と  $j^*$  とがリンクを構成するための条件として、次式の条件を用いる。

$$-2 \sum_{i=1}^N \log\{p(\Delta q_j^{(i)} | \Delta q_{j^*}^{(i)})\} < -2 \sum_{i=1}^N \log\{p(\Delta q_j^{(i)})\} \quad (1)$$

ここで、 $N$  は特徴点数、 $q_j^{(i)}$  は画像  $i$  における特徴点  $j$  の位置を表し、 $\Delta$  は運動指令前後 (2 枚の画像間) の変位を表す。条件付き確率の推定には、LSCDE (Least-Squares Conditional Density Estimation) [Sugiyama 10] を用いる。

遮蔽に関する依存ネットワークの構築 画像  $i$  に対して、微小変位前後での特徴点  $j$  の出現・消失を表す変数を  $o_j^{(i)} \in \{-1, 0, 1\}$  とする。ここで、1 は出現、0 は変化なし、-1 を表す。これを用いて、遮蔽に関する依存を表すための条件として次式を用いる。

$$-2 \sum_{i=1}^N \log\{p(o_j^{(i)} | \Delta q_{j^*}^{(i)})\} < -2 \sum_{i=1}^N \log\{p(o_j^{(i)})\} \quad (2)$$

離散変数に関する条件付き確率の推定のために LSPC (Least-Squares Probability Classifiers)[Yamada 11] を用いる。

ネットワークのクラスタリング 上記の基準にもとづいて、状況によって多様なネットワークが構成される。これらのネットワークのうち似た構造をもつネットワーク同士をクラスタリングする。まずネットワーク間の距離計算をノードとリンクのマッチングにより行う。その距離情報を用いて多次元尺度構成法により、低次元空間における各ネットワークの座標を計算する。この座標値を用いて、Reciprocal nearest neighbors 法により近い点同士を凝集させる。

#### 4. 実験

提案するネットワーク構築およびネットワークのクラスタリング手法をヒューマノイドロボット NAO により検証した。運動に関する依存ネットワークをクラスタリングした結果を図 4 に示す。この例では 4 種類のクラスタが形成された。青い点に対応するクラスタは物体と手先が同方向に動いている状況に対応している。水色の点に対応するクラスタは、物体と手先との間の接触が起きていない状況に対応している。赤色の点はネットワークが構成されない場合に対応している。

また、遮蔽に関するネットワークをクラスタリングした結果を図 5 に示す。この例では 2 つのクラスタが形成された。図中左側のクラスタは手先が物体により隠される状況に対応している。このように、身体と物体の位置関係・運動の関係により異なるネットワーククラスタが形成されることが確認された。これらにより、ロボットと環境に関わる状態を識別することが可能であることが確認された。

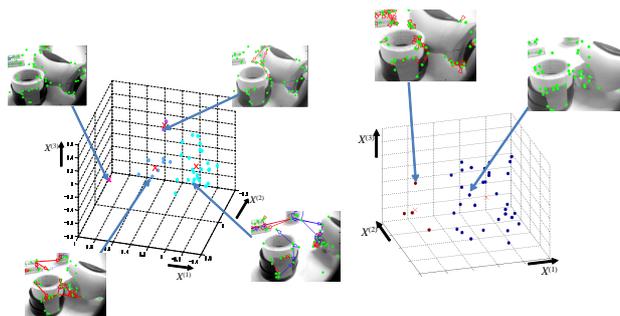


図 4: 運動ネットワークのクラスタリング結果 図 5: 遮蔽ネットワークのクラスタリング結果

#### 5. おわりに

本研究では、ロボット身体、環境に関する前提知識を用いずに視覚情報から状態を識別するための空間構成方法を提案した。手先のみの運動、手先と物体との運動、手先と物体間の遮蔽などの「関係性」に着目することで、設計者のラベル付けを経ないよりボトムアップな状態識別が可能であることが確認された。これにより、ロボットが行動を生成する際に「ロボットの動作が何を惹き起こすか」という可能性を考慮しながら状態を識別することが可能になる。

謝辞 本研究の一部は栢森情報科学振興財団の研究費助成の支援を受けて行われた。ここに謝意を表します。

#### 参考文献

- [Sinapov 08] J. Sinapov and A. Stoytchev, Detecting the functional similarities between tools using a hierarchical representation of outcomes, Proc. of 7th IEEE Int. Conf. on Development and Learning, pp. 91-96, 2008.
- [Lungarella 03] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, Developmental robotics: A survey, Connection Science, vol. 15, no. 4, pp. 151-190, 2003.
- [Stoytchev 08] A. Stoytchev, Learning the Affordances of Tools using a Behavior-Grounded Approach, In "Affordance-Based Robot Control," Springer Lecture Notes in Artificial Intelligence (LNAI) 4760, E. Rome et al. (Eds.), pp. 140-158, 2008.
- [Kobayashi 12] Y. Kobayashi, T. Okamoto and M. Onishi, Generation of Obstacle Avoidance Based on Image Features and Embodiment, Int. Journal of Robotics and Automation, Vol. 24, no. 4, pp. 364-376, 2012.
- [Lowe 99] D. G. Lowe, Object Recognition from Local Scale-invariant Features, Proc. of IEEE Int. Conf. on Computer Vision, 2, pp. 1150-1157, 1999.
- [Sugiyama 10] M. Sugiyama, et al., Least-Squares Conditional Density Estimation, IEICE Trans. on Information and Systems, E93-D, no.3, pp. 583-594, 2010.
- [Yamada 11] M. Yamada, et al., Improving the Accuracy of Least-Squares Probabilistic Classifiers, IEICE Trans. on Information and Systems, E94-D, no.6, pp. 1337-1340, 2011.