

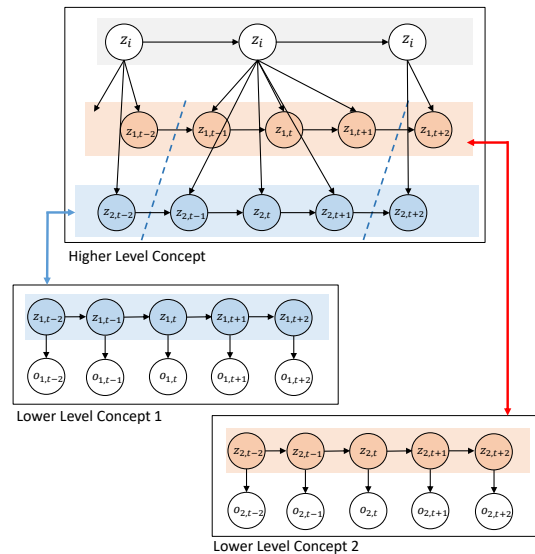
時系列マルチモーダル情報の分節・分類に基づく ロボットによる概念の学習

布川 遼太郎[†] 宮澤 和貴[†] 中村 友昭[†] 長井 隆行[†] 金子 正秀[†]

[†]電気通信大学

1 前書き

概念は人間の認知において重要な役割を担っている。人は概念を用いることで、知覚情報から最小限の認知的処理で様々な情報を獲得し、未知の事象にも対応している。ロボットも概念を用いることによって、一部の可観測情報から未観測情報を予想可能になり、様々な事物に対して柔軟に対応することが可能となると考えられる。本稿では概念を、知覚した情報をクラスタリングすることで形成されるカテゴリと定義し、ロボットが環境とのインタラクションを通じて取得可能なマルチモーダル情報を用いて概念形成を行う。しかし、実際に得られるセンサ情報は連続的であり、このままではクラスタリングすることが困難である。そこで、隠れセミマルコフモデルを用いてセンサ情報を分節・分類することで、1つの概念によって表現されるセンサ情報の時間的な範囲と、そのカテゴリを教師なしで推定する手法を提案する。



2 統合概念モデル

文献 [1] では、ロボットが何らかのタスクを行うことで、物体および動作の概念を統合した上位概念を獲得する手法を提案している。このモデルでは複数の下位概念が同時に共起するものとして、これらの下位概念の関係性を表現する上位概念を形成している。しかし、センサ情報は時系列情報であり、それらをクラスタリングした複数の下位概念の発生には時間的ズレが生じ、必ずしも同時に共起するわけではない。そのため、ロボットは下位概念のどこからどこまでを共起としてとらえるかを推定する必要がある。そこで、図1に示すような、時系列情報の分節・分類に基づいた統合概念モデルを考える。これは、観測情報 o_1 と o_2 の分節・分類によって下位概念を形成し、上位概念 z_i が下位概念 z_1^* と z_2^* の共起関係を表すモデルである。 z_i では、下位概念の共起関係と、上位概念で表される下位概念の時間的な範囲を同時に推定する。上位概念のカテゴリ系列 z は次式で推定することができる。

$$z \sim \iint P(z|z_1, z_2)P(z_1|o_1)P(z_2|o_2)dz_1dz_2 \quad (1)$$

しかし、この式を直接計算することは困難であるため、近似することで計算する。まず下位概念モデルで、下

Learning of Concepts by Robots Based on Segmentation and Classification of Time Series Multimodal Information

Ryotaro NUNOKAWA[†], Kazuki MIYAZAWA[†], Tomoaki NAKAMURA[†], Takayuki NAGAI[†], and Masahide KANEKO[†]

[†]The University of Electro-Communications, 182-8585, Chofu, Japan

図1: 時系列情報の分節・分類に基づく統合概念モデル

位カテゴリ系列 z_1^*, z_2^* を推定する。

$$z_1^* \sim P(z_1|o_1) \quad (2)$$

$$z_2^* \sim P(z_2|o_2) \quad (3)$$

これにより下位概念モデルは、推定された z_1^* と z_2^* を上位概念モデルに渡すことで式 (4) を Forward filtering-backward sampling によって計算することができる。

$$z^* \sim P(z|z_1^*, z_2^*) \quad (4)$$

Forward filtering では、動的計画法を用いて、時刻 t を終点とする分節長 k の下位概念の分節 $z_{*,t-k:t}$ が、上位概念 z になる確率 $\alpha[t][k][z]$ を計算する。

$$\alpha[t][k][z] = P(z_{1,t-k:t}|z)P(z_{2,t-k:t}|z) \sum_{k', z'} \alpha[t-k][k'][z'] \quad (5)$$

この確率を用いて Backward sampling を行い、下位概念の分節位置を決定することで z^* を決定することができる。

3 実験

前節で説明したモデルを用いて、実際にロボットが物体を操作し得られるマルチモーダル情報を分節・分

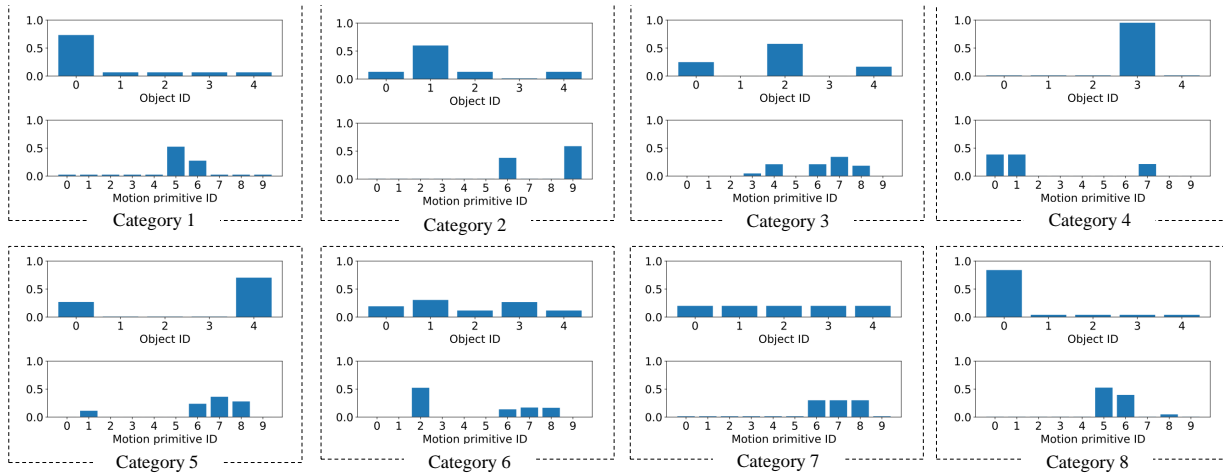


図 3: 各カテゴリにおける物体と動作プリミティブ系列の発生確率



図 2: 実験に用いた物体

表 1: 物体と動作の対応

物体 (ID)	動作プリミティブ系列
ぬいぐるみ (0)	7,8,7,8,6 または 5,5,6
ボタン (1)	9,9,6
おもちゃ(2)	7,3,6
ペットボトル (3)	7,1,0,1,0,7,8,6
スポンジ (4)	7,1,7,8,6
おもちゃ・スポンジ	7,4,7,4,8,6
全物体	7,2,2,2,8,6

類することで概念の学習を行った。ロボットは目の前に置かれた物体を認識し、その後、各物体に対応した動作をロボットに行わせた。このタスクを繰り返し行うことによって時系列マルチモーダル情報を取得した。下位概念としては物体概念と動作概念を用いた。操作対象の物体の画像を分類することで得られる物体 ID を物体概念、連続動作を GP-HSMM[2] によって分節することで得られる動作プリミティブを動作概念とした。実験には図 2 に示した物体を用い、表 1 に示す動作を各物体に対して行った。

図 3 が実際に学習された上位概念を分節・分類することで得られた、各カテゴリにおける物体 ID と動作プリミティブ系列の発生確率である。Category 1 にお

いて、物体 0 に対して正解である動作プリミティブ 5・6 によって構成される動作が行われることが学習されている。同様に Category 2・3・5・6 においても、各物体に対応する実行可能な動作が正しく学習されていることがわかる。一方で Category 1 と Category 8 の両方で、物体 0 に対してプリミティブ 5・6 で構成される動作が実行可能であることが学習されており、正解であるプリミティブ 6・7・8 で構成される動作が実行可能であることを表現する結果は得られなかった。これは、物体 0 に対して実行可能なプリミティブ 6・7・8 で構成される動作 (7,8,7,8,6) と、物体 4 に対して実行可能なプリミティブ 1・6・7・8 によって構成される動作 (7,1,7,8,6) が非常に似ているため、学習が正しくできなかったことが原因であると推察される。その結果、Category 5 に、プリミティブ 1・6・7・8 によって構成される動作が実行可能な物体として、正解である物体 4 に加えて正解でない物体 0 が学習結果として表れている。

4 結び

本稿では、ロボットが物体を操作し得られる時系列マルチモーダル情報を分節・分類することによって概念を学習する手法を提案した。実際にロボットを用いた実験を行い、提案手法によって上位概念の学習が可能であることを示した。今後の課題として、今回の実験結果で見られたような、似ている動作が行われた場合に学習が正しく行えない点を改善することで、より精度の高い学習を可能にすることが挙げられる。

参考文献

- [1] 宮澤 和貴, 青木 達哉, 日永田 智絵, Mata, D., 岩田 健輔, 中村 友昭, 長井 隆行: 多層マルチモーダル LDA を用いた概念形成と強化学習, 計測自動制御学会システム・情報部門 学術講演会, O1-2 (2016)
- [2] Nakamura, T., Takayuki, N., Daichi, M., Ichiro K., Hideki, A., Masahide, K.: Segmenting Continuous Motions with Hidden Semi-markov Models and Gaussian Processes, Front. Neurobot, doi: 10.3389/fnbot.2017.00067 (2017)