人間型ロボットによる内発的動機に基づく
物理理解のための行動選択

鈴木 智也† 鈴木 健嗣†
†筑波大学大学院システム情報工学研究科

1 はじめに
近年、ロボットが環境との相関作用を通じて新たな知見を獲得し、自身の行動計画を行うシステムモデルの研究が盛んに行われている。我々は、特に人間の乳幼児期において彼らの行動を引き起こし発達を促す、内発的動機の一つである好奇心の発動に着目し、人間型ロボットによる物理解のための行動への応用について検討している。Oudeyerらは予測誤差の減少を内部報酬として好奇心のモチベーションとして実ロボットによる実験を行った[1]。また、山下らは予測誤差の低減に加え、長時間探索を行っていない領域に価値を付加するモデルを提案し、シミュレーション上で有効性を示した[2]。本稿では、我々が提案したシステム（図1）[3]に物体間の類似性を評価する指標を導入し、物体特性及び把握行動に関する知見の獲得が可能であることを示す。

2 物体間非類似度の導入
過去に行われた探索行動の結果を、新規な物体の探索時における予測に活用することで、より効率の良い探索が期待される。そこで、ここでは静止物体間非類似度\(OD_k\)と物体間非類似度\(OD_d\)によって得まる物体間非類似度\(OD\)を導入する。\(OD_k\)はカメラ画像から得られる物体の静的特徴量\(s_k(j)\)より、\(OD_d\)は物体\(O_k\)に対する行動\(a_k\)によって最も生じやすいクラスターの平均センサパターン\(\bar{s}(a_k)\)より求める。その上で\(OD\)をパラメータ\(\alpha, \beta\)を用いて以下のように定義する\((\alpha = 1.0, \beta = 10.0)\)。なお、初見の予測時は\(OD_d(O_{j}, O_{k}) = 0\)とする。

\[
OD_d(O_{j}, O_{k}) = \left( |s_j(j) - s_k(j)|^2 \right)^{1/2} \tag{1}
\]

\[
OD_d(O_{j}, O_{k}) = \left( \sum |s_j(j) - \bar{s}(a_k)|^2 \right)^{1/2} \tag{2}
\]

\[
OD = \alpha OD_k + \beta OD_d \quad (\alpha < \beta) \tag{3}
\]

さらに、過去の経験を利用した対象物の模擬予測\(P(C_{i}|a_{n}, O_{k})\)を重み\(w_{j}\)から求める。

\[
w_{j}(a_{n}, C_{i}) = \sum_{k=(k)m\in a_{n}(a_{n}, O_{k})=C_{i}} e^{-OD(O_{j}, O_{k})} \tag{4}
\]

\[
P(C_{i}|a_{n}, O_{k}) = \frac{w_{j}(a_{n}, C_{i})}{\sum_{i} w_{j}(a_{n}, C_{i})} \tag{5}
\]

3 評価実験
3.1 システム構成
実験に用いる人間型ロボットGenie（図1左上）は7自由度の右腕、3指それぞれに角度センサ及び力覚センサを搭載したハンド。頭部にCCDカメラ、胸部にマイクロフォンを持つ。以下の実験では、これらのセンサから0.1秒毎に取得した時系列データを用いる。

3.2.1 実験環境
実験環境を図2に示す。ロボットは、ここでは水平面に置かれた積み木に対し、探索行動を行う。対象として、6種類の物体\(O_1 - O_6\)を用いる（図3）。\(O_1, O_3, O_5\)及び\(O_2, O_4, O_6\)はそれぞれ材質が異なり、全ての物体について形状は類似している。右腕の初期姿勢と物体の初期位置を予め定め、物体とは第3指のみが接触するものとした。探索時における右腕の挙動は水平面からの仰角\(\theta\)を180\(^\circ\)まで変化させた20段階の角度を\(\theta\)により定める。探索行動は、指先を\(\theta\)方向に1.0秒動か、元の方向に0.2秒戻り、0.8秒そこでどまる、という3段階の動作からなる。

物体の外観情報は、物理初見時にカメラ画像より算出する。物体の画像中心、主方向、面積及び色相とすると、一方、探索行動時には物体挙動等を表す時系列情報として、初期位置からの画像中心の差分、主方向の差分、面積の差分、音圧、第3指間節角度及び張力を利用する。

3.2.2 実験結果
上記の条件で各物体に対し順番に50回の探索行動を行った。物体\(O_1 - O_6\)探索後に物体\(O_k\)を対象とした探索時における\(O_k\)他の物体とは物体間非類似度\(OD\)の推移を図4に示す。探索行動を行う前の状態では、\(O_k\)の外観情報のみで非類似度の評価が行われる。今回、特に色相の違いが支配的であることが見てとれる。
図4：物体Oaと他の物体との物体間非類似度の推移

一方、探索開始後は得られたOaの物体挙動も考慮して、逐次的に類似性の評価が行われる。10回の探索を行った時点では、Oaと同じ材質からなるOa, Oa, その他、どの物体に比べて類似した物体であると見なされた。すなわち、物体とのインタラクションを通じて得られた情報を利用して、ロボットがその材質や重さの違いを踏まえた物体間の評価が可能であると言える。

3.3 把持行動に関する探索実験

ここでは、異なる取得センサ情報と行動パターンにおいても、提案手法によりロボットの行動に関する知見の獲得が可能であることを示す。

3.3.1 実験環境

ロボットは物体に対してハンドの姿勢を変え、第1指、第2指を用いた把持を続け行動の約束を通じて、各姿勢における把持の可能性についての知見の獲得を行う。初期姿勢は図5に示すような種類の姿勢とハンドの高さhによって一意に定める。今回、hは上記範囲で定め、ロボットの把持・持ち上げ行動は、4秒間の2指の屈曲方向への動作、1秒間の前直陣方向への持ち上げ、2秒間の静止、計7秒の3段階からなる、センサ情報として、前述の実験と同じ視覚情報に加え、2指分の関節角度、指の張力を用いる。

3.3.2 実験結果

100回の探索行動後に得られた行動空間内のセンサパターン予測分布を図6に示す。探索によって、空を切る、つまりで落ちる、持ち上げるの3種類の挙動を獲得した。図6(a)は行動時間全体における発生挙動予測分布を示す。横方向把持及び斜め把持においては一定の高さ以上で把持が出来ず、また下向き把持では一定範囲内の把持が出来ず、特に知見が得られた。図6(b),(c)は各探索行動のうち、挙動が発生した行動についての挙動予測分布を示す。このように探索行動を通して、把持・持ち上げ行動により発生する事象及び各行動との対応関係が獲得可能であることを示した。

4 まとめ

本稿ではこれまでに提案したシステムの概要を示す。提案手法においては、複数の物体を用いた探索から、ロボットが探索時に得た知見を用いて、物体間の材質の違いを評価評価することを示した。また、提案手法をロボットが物体の把持可能性を考慮する探索を行う従来に適用可能であることを実験により示した。本稿では探索により得られた知見を目的に応じて利用することを実証している。今後は、探索行動の連続化や自己の身体を対象とした探索行動の実現を目指す。

参考文献