

6B-4

# 物体静止画像から動的特徴を抽出する神経回路モデルの学習と解析

西出 俊<sup>†</sup> 尾形 哲也<sup>†</sup> 谷 淳<sup>‡</sup> 駒谷 和範<sup>†</sup> 奥乃 博<sup>†</sup>

<sup>†</sup> 京都大学 大学院情報学研究科 知能情報学専攻    <sup>‡</sup> 理化学研究所 脳科学総合研究センター

## 1. はじめに

近年、ロボットが能動的に環境に作用することで動的な情報を抽出する「アクティブセンシング」による物体認識が注目されている。高椋らは物体を押すことで動的な視覚特徴を [1]、尾形らは物体を叩くことで動的な視聴触覚特徴を抽出し [2]、抽出された動的特徴をもとに物体認識を行った。我々は学習物体に対する能動知覚経験をもとに、未知物体の挙動を予測する手法を提案している [3]。

前稿ではロボットの押し動作をもとに学習用物体の挙動を観察・学習し、未知物体に対する挙動連想について報告した。本稿では提案システムを解析することで、学習時にロボットが自動的に抽出した特徴と物体挙動の関係について調べる。解析の結果、エッジや縦横比など、物体挙動に大きく影響を与える静的特徴が抽出できたことを確認した。

## 2. 学習モデルと能動知覚実験

我々は谷らが提案した Recurrent Neural Network with Parametric Bias(RNNPB)[4] と階層型ニューラルネットを組み合わせた二段階学習モデルを提案している。RNNPB は複数のパターンを学習するために、Jordan 型 RNN に各パターンをエンコードする PB 層を付加したものである。PB は各パターンの類似度に応じて自己組織的に PB 空間にマッピングされ、この PB によって RNN の挙動は決定する。本手法ではまず物体操作時に観察される物体挙動を RNNPB で学習し、各挙動の PB 値を自己組織化する。得られた PB 値を教師信号とし、物体静止画像とロボットの動作を入力とした階層型ニューラルネットを学習する。提案モデルの概略図を図 1 に示す。

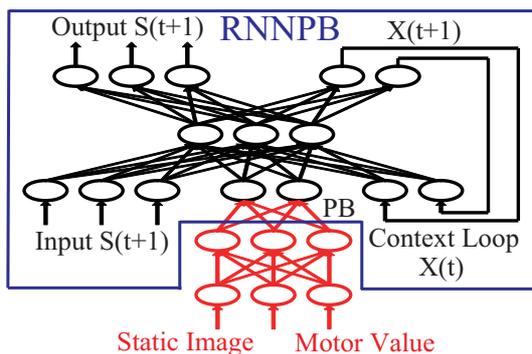


図 1: 二段階学習モデル

実験のテストベッドとして ATR 製の Robovie を改造した Robovie-IIs(図 2) を用いる。図 3 に示す学習物体に対して押し動作を行い、重心と慣性主軸を秒間 2.5 フレームで 7 ステップ分計測する。この時、物体はその形状に応じて「倒れる」、「滑る」、「転がる」挙動を生成する。

Training and Analysis of a Neural Network Model for Extracting Dynamic Features from Static Object Images Shun Nishide (Kyoto Univ.), Tetsuya Ogata (Kyoto Univ.), Jun Tani (RIKEN), Kazunori Komatani (Kyoto Univ.), and Hiroshi G. Okuno (Kyoto Univ.)

計測した挙動軌跡を [0,1] に正規化し、RNNPB で学習することで PB 値を自己組織化する。ニューラルネットの入力として、図 4 に示す 40×30 のグレースケール画像と物体を押す高さに相当するロボットの肩関節角を [0,1] に正規化したものを用いる。各ニューラルネットの構成は表 1 に示す。



図 2: Robovie II-s

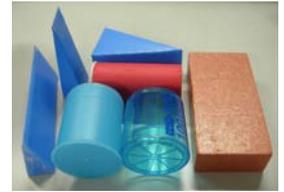


図 3: 学習用物体

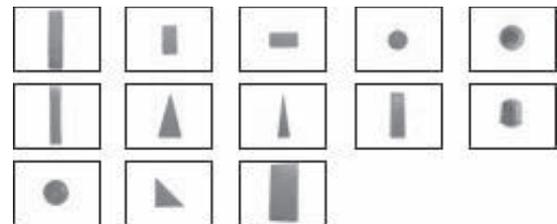


図 4: 入力物体画像

## 3. モデルの解析

学習した階層型ニューラルネットと RNNPB をそれぞれ解析することでニューラルネットが抽出した特徴と物体挙動の関係を調べる。以下にそれぞれの解析方法について述べる。

階層型ニューラルネットの解析については入力層から中間層への重みを調べる。画像入力部とロボット動作入力部を別々に評価する。画像入力部については、各ピクセルから中間層への重みを輝度値に変換し、抽出された静的特徴を画像として表し、評価する。ロボット動作入力部については、各静的特徴にどのような影響を付与するか調べる。

RNNPB の解析については PB 値がどの物体挙動に対応しているか調べる。学習によって生成された PB 空間を 25×25 に分割し、各領域を RNNPB の出力を入力にフィードバックする連想挙動によって評価する。また、階層型ニューラルネットの解析結果と統合し、物体形状・ロボット動作・物体挙動の関係を総合的に調べる。

表 1: ニューラルネットの構成

	RNNPB	Hierarchical NN
Input/Output Layer	3	40×30+1
Middle Layer	10	4
Context Layer	10	—
PB Layer	2	—
Iteration Times	1,000,000	30,000

## 4. 解析結果

### 4.1 階層型ニューラルネットの解析結果

学習した階層型ニューラルネットの重みを図5に示す。中間層のノード数は4であり、各ノードに1~4の層番号を付ける。

「抽出特徴」は画像部入力から中間層への重みを画像として描画したものである。「抽出特徴」の白い部分は中間層値に対して正に影響する箇所であり、黒い部分が負に影響する箇所である。各抽出特徴は図4に示した学習物体画像のうち、丸い・縦横比・上下面比など、挙動に大きく影響を与えるものを選出していることが分かる。すなわち、丸いほど転がりやすく、縦横比が大きいほど倒れやすく、上面のほうが下面より小さいほど滑りやすくなる。特に、縦横比の大きさに相当する箇所は黒く強調されており、倒れやすさは負の影響を与えることが分かる。また、物体領域のうち特にエッジ部分が強調されている。これより、物体挙動に大きく影響を与えるのはエッジ部分であることを学習したといえる。

「動作部」はロボット動作の入力から中間層への重みを表している。図5より、動作部は中間層のどのノードに対しても負の影響を与えている。これは「抽出特徴」内で黒く表示されている部分と同じ影響を与えることを意味する。よって、動作部が大きくなる、すなわち、押し高さが高くなるほど物体は倒れやすくなることを学習したといえる。

「PB1」と「PB2」は中間層から、二つのPB値、PB1とPB2、への重みを表している。中間層値は「抽出特徴」と「動作部」の和にシグモイド関数をかけた値であり、出力PB値も中間層値の重み付け和にシグモイド関数をかけたものである。シグモイドの入力が負になるほど出力は0に、正になるほど出力は1に近づく。これより、縦横比が大きい物体やロボット動作が大きいものは中間層の値が0に近づき、PB値は0.5に近づく。

### 4.2 RNNPBの解析結果

解析したPB空間を図6に示す。学習挙動は全て物体を右に押したものである。未学習の挙動は学習した二つの挙動の組み合わせ、あるいは奥に移動しているものとなっている。

階層型ニューラルネットの入力がRNNPBの物体挙動に与える影響を調べるため、入力物体画像として縦と横のピクセル数が、 $4 \times 28$ 、 $6 \times 24$ 、 $8 \times 20$ 、 $10 \times 16$ 、 $12 \times 12$ 、 $14 \times 8$ となる四角形を入力し、各中間層値とPB値を調べた。この時、ロボット動作として低いところを押す、0.1を入力した。その結果を表2に示す。

表2より、中間層の第1ノードでは倒れにくさ、第2ノードでは転がりやすさが抽出されていることが分かる。

層番号	1	2	3	4
抽出特徴				
動作部	-188.9	-153.7	-72.6	-143.6
PB1	-30.66	-52.61	39.62	45.44
PB2	-47.52	-5.99	-40.13	38.23

図5: 階層型ニューラルネットの重み

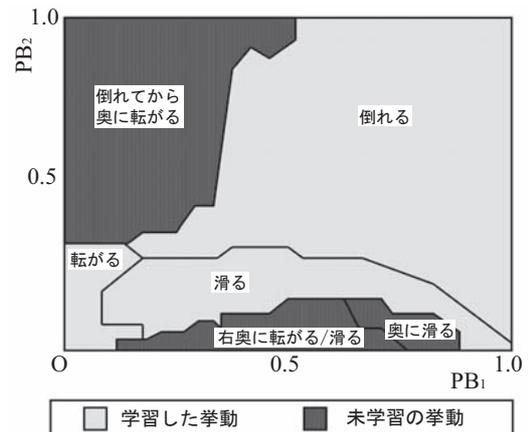


図6: PB空間の解析結果

表2: 物体画像とPB値の関係

	層1	層2	層3	層4	PB1	PB2
$4 \times 28$	0.178	0.570	0.227	0.981	0.712	0.695
$6 \times 24$	0.083	0.647	0.113	0.991	0.656	0.782
$8 \times 20$	0.959	0.964	0.924	0.999	0.524	0.075
$10 \times 16$	0.999	0.951	0.492	0.990	0.308	0.148
$12 \times 12$	0.999	0.976	0.020	0.826	0.101	0.245
$14 \times 8$	0.999	0.318	0.217	0.997	0.583	0.272

第3ノードでは $8 \times 20$ の時のみ大きな値をとっている。このノードではロボットの動作が物体挙動に影響を与えるものを抽出している。即ち、ロボット動作を高くすると、図5の「動作部」より、1, 2, 4ノードは大きく減少するが、第3ノードの減少量は小さい。これより、ロボットの動作が低いときには滑り、高くなると倒れるという対応付けが得られ、実際にロボットの動作を高くしたときにPB2が大きくなるという結果が得られた。第4ノードは全体的に大きな値をとっており、PB値を調整する役割を果たしている。

## 5. おわりに

本稿では物体挙動を予測するニューラルネットの解析を行った。解析の結果、「倒れる」「滑る」「転がる」物体挙動に大きく影響を与える丸い・縦横比・上下面の面積比などの静的特徴をニューラルネットが抽出していることが分かった。また、中間層の各ノードは物体画像から挙動を予測するフィルタの役割を果たしている。以上より、本システムの挙動予測器としての有効性が示された。謝辞

本研究は科研費若手研究A(No. 17680017)、栢森情報科学振興財団設立10周年記念特別研究助成、理研の支援を受けた。

### 参考文献

- [1] S. Takamuku, Y. Takahashi, and M. Asada. Lexicon Acquisition based on Behavior Learning. *Proc. IEEE ISDL, TALK 14*, 2005.
- [2] T. Ogata, H. Ohba, J. Tani, K. Komatani, and H. G. Okuno. Extracting Multi-Modal Dynamics of Objects using RNNPB. *Proc. IEEE/RSJ IROS*, pp. 160-165, 2005.
- [3] 西出俊, 尾形哲也, 谷淳, 駒谷和範, 奥乃博. “能動知覚経験に基づく物体静止画像からの挙動推定.” 計測自動制御学会システムインテグレーション部門学術講演会講演論文集 (SI), pp.39-40, 2006.
- [4] J. Tani and M. Ito. Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment. *IEEE Trans on SMC Part A*, Vol. 33, No. 4, pp. 481-488, 2003.