物体挙動の予測信頼性に基づく自律的な動作生成

○西出 俊（京大）　尾形 哲也（京大）　谷 淳（理研）　駒谷 和範（京大）　奥乃 博（京大）

Motion Generation based on Reliable Predictability of Object Dynamics

*Shun NISHIDE (Kyoto Univ.), Tetsuya OGATA (Kyoto Univ.), Jun TANI (RIKEN), Kazunori KOMATANI (Kyoto Univ.) Hiroshi G. OKUNO (Kyoto Univ.)

Abstract— In this paper, the authors present a technique for generating autonomous object pushing motions based on reliable predictability. The technique first creates an internal model of the robot and object through active sensing experiences using Recurrent Neural Network with Parametric Bias. Next, the technique searches through the model for the most predictable object dynamics and corresponding robot motion based on consistency of experiences. Finally, the static object image is linked to the acquired robot motion using a hierarchical neural network. Experiments with cylindrical objects resulted in generation of consistent rolling motions of the objects denoting the effectivity of the technique.

Key Words: Neural Networks, Neurorobotics, Humanoid Robots

1. はじめに

人間の知覚は環境変化の予測に基づいているといわれている [1]．すなわち，自身の経験をもとに脳内で構築された環境モデルで変化を予測し，予測通りに環境が変化するときに知覚することができる．環境の変化を正確に予測することは，信頼できる予測可能性（reliable predictability）と呼ばれており，本稿では**予測信頼性**が高いと定義する．高い予測信頼性は知覚のみならず行動においても重要であり，環境が予測通りに変化するように人間は行動を生成している．予測信頼性は経験の多さに依存し，同じ事象を何回も観測することで高い予測信頼性が得られる．

人間は未知の対象についても過去の経験から自身の動作と環境変化の関係を予測できる．これは対象の外観を脳内の環境モデルにあてはめ，類似した経験をもとに類推するのである．人間の脳は限られたニューロン数であらゆる事象を知覚することが必要であり，未知の対象については，これまで学習した経験をもとに予測し，その信頼性を評価することで行動を生成しているのである．

人間の行動獲得過程において，予測信頼性と同様に，環境応答の一貫性も重要な役割を果たしている．過去に一貫して起こる環境応答は，同様の未知対象についても起こるであろうと予測できることから，高い予測信頼性が得られるのである．一方，同じ主体動作に対して毎回環境応答が変わるような事象については環境モデルが構築できず，予測信頼性も低くなる．

本研究では上記のような人間の行動獲得過程をロボットに実装し，環境の状態に対して予測信頼性が高い行動，すなわち一貫した環境応答が得られる行動，を自律的に生成するロボットを開発することを目標としている．本稿では，円筒型物体に対する押し動作によってロボット自身の行動と物体応答の関係を学習し，予測信頼性を評価することで，物体の姿勢や大きさに適応し，未知の対象についても一貫性のある物体挙動をロボットが生成する手法を提案する．

以下，本稿の構成は次のようになっている．第 2 章では提案モデルと手法について述べる．第 3 章では評価用の動作生成実験について述べる．第 4 章では実験結果に関する考察を述べ，第 5 章で全体のまとめと今後の展望について述べる．

2. モデルの構築

本章ではロボットと環境モデルの構築法と予測信頼性が高い動作の探索方法について説明する．本手法では未知物体についても挙動予測，動作生成を行うため，汎化能力によって未知パターンが扱えるニューラルネットワークや再帰結合を持つリカレントニューラルネットワークを用いる．提案手法は三つの段階から構成されており，物体画像から一貫した物体応答が得られる押し動作生成を実現する．

第一段階では，ロボットが学習物体に対して押し動作による能動知覚を行い，観測される時系列の物体特徴とロボット動作をリカレントニューラルネットワークの一種である Recurrent Neural Network with Parametric Bias(RNNPB)[2] で学習する．RNNPB を用いることで複数の時系列を一つのモデルで学習でき，各時系列を PB に圧縮する形で，類似度に応じて自己組織的に各時系列を PB 空間内にマッピングする．すなわち，類似した系列の PB は PB 空間において互いの近傍に配置され，異なる系列の PB は互いに離れたところに配置される．これにより，押し動作によるモータダイナミクスと生成される物体ダイナミクスが PB 空間にクラスタリングされる．

第二段階では，第一段階で形成された PB 空間内を，予測信頼性と物体応答の一貫性を統合した評価関数に基づいて予測信頼性が最も高い PB 値を最急降下法で探索する．ここで得られた PB 値は，ロボットの経験をもとに自身の動作と物体挙動の関係が最も予測しやすいロボット動作と物体挙動の関係を表している．

第三段階では，三層のニューラルネットワークを用い，物体画像と第二段階で求めた予測信頼性の高い PB の関係を学習する．以上により，ニューラルネットワークに物体画像を入力することで予測信頼性の高い PB

が得られ，その PB を RNNPB に入れることで予測信頼性の高い動作を生成することが可能になる．提案手法の概略図を Fig. 1 に示す．

2·1 RNNPB の学習

物体に対する能動知覚をもとにロボットと物体のダイナミクスモデルを学習器 RNNPB で構築する．本研究では Jordan 型の Recurrent Neural Network (RNN)[3] の入力部に PB 層を付加した RNNPB をダイナミクスの学習器として用いる．RNNPB は，現時刻の物体特徴量とロボットの関節角 $S(t)$ を入力することで次時刻の物体特徴量とロボットの関節角 $S(t+1)$ を出力する予測器となっている．RNNPB の構成を Fig. 2 に示す．

RNNPB の学習は Jordan 型の RNN と同様のプロセスで行われる．すなわち，Back Propagation Through Time (BPTT)[4] アルゴリズムによって結合重みの更新と PB 値の更新を同時に行う．学習時，出力部から逆伝播される結合重みの誤差を，各時系列ごとに PB 層に積算することで PB が更新する．更新式は以下の通りである．

$$\Delta\rho = \varepsilon \cdot \sum_{t=1}^{T} \delta^{bp} \tag{1}$$

$$p = sigmoid(\rho). \tag{2}$$

ここで，ε は学習定数，p は PB 値，T は各パターンのステップ長である．まず (1) により，PB の内部値 ρ を更新するための Δ 値が計算される．各パターンのデルタ誤差 δ^{bp} は出力層の誤差を PB 層に逆伝播することで得られる．新しい PB 値 p は (2) により，デルタ値によって更新された内部値 ρ にシグモイド関数をかけることで得られる．

RNNPB を学習することによって，各時系列パターンは PB に圧縮され，互いの類似度に応じて自己組織的に PB 空間が形成される．また，(1) と (2) の更新は各パターンごとに行われるため，類似パターンが多い時系列は，PB 空間で広く分散される．学習後，PB を固定して出力を入力に再帰的にフィードバックすることで PB から時系列を復元することも可能である．この操作は連想と呼ばれる．

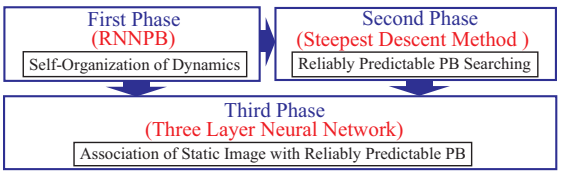

Fig.1 Schematic Diagram of Proposed Technique

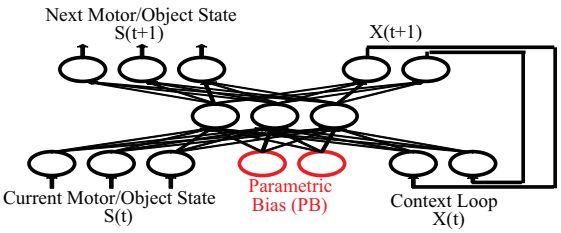

Fig.2 Structure of RNNPB

2·2 予測信頼性の計算法

本手法では形成された PB 空間内から予測信頼性が最も高いロボット動作を示す PB を探索するため，最急降下法を応用する．PB の微小変動に対する物体ダイナミクスの変動に関する評価関数を以下のように設定し，その局所解を最急降下法で探索することで予測信頼性が高いロボット動作を求める．

$$E(p) = \frac{\delta O^2}{\delta p}. \tag{3}$$

ここで O は物体ダイナミクス，p は PB 値である．この評価関数は予測信頼性を評価しており，最急降下法によって得られる局所解 (PB) は，その近傍領域において，物体ダイナミクスが最も予測しやすいロボット動作を表している．

式 (3) を解析的に解くことは困難であるので，本稿では数値計算で解く．式 (3) を離散化し，数値計算によって微分値の計算を行うことで，評価関数 E は以下と等価になる．

$$E = \frac{1}{\mu}\sum_{i,j,t}(O(p_1,p_2,t) - O(p_1+i\mu, p_2+j\mu, t))^2$$

$$(i,j=-1,0,1) \quad (i \cdot j = 0), \tag{4}$$

ここで t は系列のステップ数を，μ は離散幅を表す．

本手法では，予測信頼性と物体応答の一貫性を統一的に評価する．RNNPB の学習時，一貫して得られる応答パターンは PB 空間において広く分散されるため，求めるべき PB は最もたくさんの初期値から収束する局所解である．本手法ではこれを求めるため，[0,1] で定義されている PB 空間（シグモイド関数の値域）を格子分割し，各格子点を初期値として局所解を求める．最も多くの初期値が収束する局所解が予測信頼性の高い動作を圧縮した PB(p^*) である．本アルゴリズムの概略図を Fig. 3 に示す．

2·3 物体画像とロボット動作の連結

第三段階では三層のニューラルネットワークを用い，物体画像から第二段階で求めた予測信頼性の高い動作を生成する枠組みを付加する．Fig. 4 に示すように，ニューラルネットワークの入力部に初期物体静止画像を入れ，出力部を第二段階で得られた PB 層とする．これにより，物体の形状・姿勢に適応して，予測信頼性が高いロボット動作を生成することが可能になる．

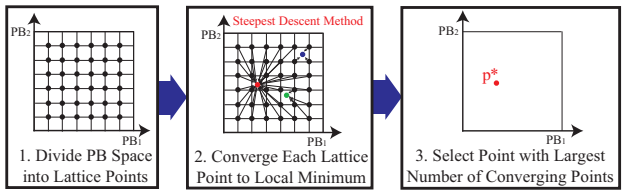

Fig.3 Calculation of Reliable Predictability

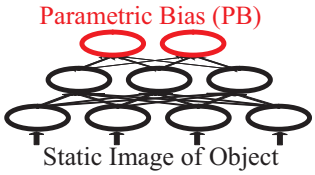

Fig.4 Construction of Neural Network

3. 動作生成実験

提案手法を評価するプラットホームとしてヒューマノイドロボット Robovie II-s[5] を用いた．Robovie II-s には首 3 自由度，両腕それぞれ 4 自由度ずつあるが，本実験では左腕の肩ロール軸と肘ヨー軸の 2 自由度を用いて実験を行った．また，頭部には 2 つのカメラが装着されており，うち左の一眼のみ用いた．

モデルの学習は以下の順序で行う．

1. 学習物体に対する能動知覚で物体特徴とロボット動作の系列を取得．
2. 取得した系列を RNNPB で学習．
3. 最急降下法で予測信頼性の高い動作を求める．
4. 三層のニューラルネットワークを学習して物体初期画像と動作 (PB) を連結．

また，動作生成は以下のように行う．

1. 物体初期画像を三層のニューラルネットワークに入力し，PB を求める．
2. 求まった PB を RNNPB に入れ，時系列を連想することで動作を生成する．

3·1 学習データの取得

本実験ではまず，学習用円筒物体に対する押し動作によってロボット動作と画像内の物体変化を観測した．肩ロール軸と肘ヨー軸を変えながら計 5 通りの押し動作を 5 つの姿勢で置かれた 2 種類の物体に対して行い，計 50 通りの動作系列を取得した．Fig. 5 に Robovie II-s が物体を押している風景を示す．

実験用物体として，Fig. 6 に示すような，2 つの学習物体と 2 つの動作生成用物体（未知物体）の計 4 つの物体を用いた．Fig. 6 の左の二つ（お菓子の容器と筆入れ）は学習用物体であり，右の二つ（スチール缶と貯金箱）は未知物体である．物体抽出作業の簡略化のため，上二つの物体は赤い画用紙を巻き，色情報で抽出できるようにした．物体特徴として，画像内の物体重心 (x, y) と慣性主軸の長軸方向 (θ)（計 3 自由度）を用い，各系列は 2.5 フレームで 8 ステップずつ計測した．学習データ取得時，物体の慣性主軸の短軸方向に押した時，一貫して物体が転がる動作が生成され，全 50 回中，計 33 回生成された．一方，他の動作については，物体挙動について一貫性が見られなかった．本実験で生成すべき予測信頼性が高い動作は物体を転がす動作である．

Fig.5 Robovie-IIs and Experiment Scene

Fig.6 Objects : Training (Left), Target (Right)

3·2 実験条件

RNNPB は計 42 個のニューロンによって構成されている：入出力層はそれぞれ 5 個ずつ，中間層とコンテキスト層はそれぞれ 15 個ずつ，PB 層は 2 個である．入出力層は 3 つのニューロンから成る物体特徴量と 2 つのニューロンから成るロボットの関節角によって構成されている．各学習データは $[0, 1]$ に正規化し，100 万回学習を繰り返すことでモデルを構築した．また，学習定数 ε は 0.03 と設定した．

本実験では PB 空間を 0.1 刻みで格子分割し，辺の部分を除いた 8×8 個の格子点を初期値として最急降下法で予測信頼性の高い動作を探索した．本実験では離散幅 μ を 0.001 と設定した．

三層のニューラルネットワークの設定は，入力層に 23×22 個のニューロン，中間層に 10 個のニューロン，出力層に 2 個のニューロン，とした．入力はロボットが物体を押す直前にカメラで取得された画像を縮小した 256 階調のグレースケール画像（解像度 23×22）の画素値を $[0, 1]$ に正規化したものによって構成されている．入力画像例を Fig. 7 に示す．出力層は RNNPB の PB 層の 2 つのニューロンから成っている．本実験では学習は 3 万回行った．

3·3 PB 空間の形成と未知物体に対する動作生成

RNNPB の学習により，各時系列を圧縮する PB が自己組織化され，PB 空間が形成される．形成された PB 空間を Fig. 8 に示す．Fig. 8 において，三角形の記号は転がる物体挙動の PB を表しており，正方形の

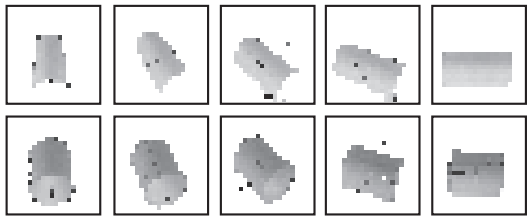

Fig.7 Examples of Input Images

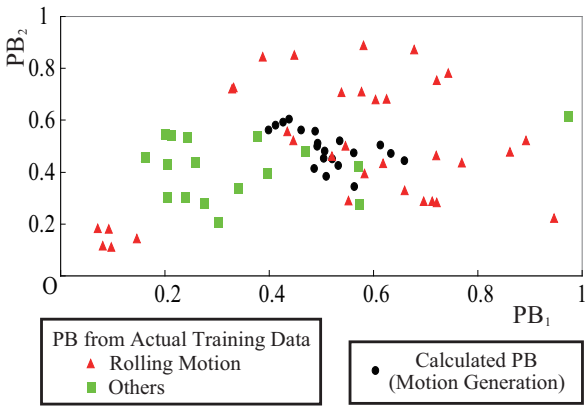

Fig.8 Calculated PB for Motion Generation

記号はそれ以外の物体挙動を表している．左下に分布している三角形は慣性主軸が -180° と 180° の間で不連続になるため，中央のクラスタから離れたところに分布している．また，三角形と正方形のクラスタは中央付近で多少重なるが，これは物体の初期姿勢によって転がる場合と転がらない場合と存在するためである．

動作生成実験は，Fig. 6 の 4 つの物体をそれぞれ 5 つの姿勢で置いて行った．20 通りの実験それぞれについて物体画像から PB を計算し，得られた PB から連想される動作を生成した．動作生成時に計算された PB を Fig. 8 において円の記号で示す．これらの動作生成時の PB は，学習時に得られた転がる挙動の PB の付近に分布していることが分かる．また，得られた PB をもとに生成した動作を Fig. 9 に示す．Fig. 9 から分かるように，全ての実験においてロボットは物体が転がる方向に押すことに成功した．

4. 考察

本実験結果に関する考察を以下に述べる．

4·1 特徴量の設定

RNNPB を学習することで形成される PB 空間は類似したダイナミクスをクラスタリングする．本実験で得られた PB 空間では転がる挙動に対応する PB のクラスタが二分されている．これは特徴量を恣意的に重心と慣性主軸と定めたためであり，慣性主軸の不連続性に対応することができないためである．また，転がる挙動とその他の挙動で重なる領域も存在することから，一つのカメラで抽出した重心と慣性主軸では特徴量が不十分であったと考えられる．より正確に物体挙動を記述するためには物体画像列から特徴量を自動抽出する枠組みを本手法に付加することが必要である．

4·2 転がる挙動の生成

本実験では，姿勢や形状が異なる円筒物体に対して，一貫して転がる挙動をロボットが生成することに成功した．これは学習データの中間的な状態も扱える汎化能力によるものである．しかし，貯金箱の形状が完全な円柱になっていないため，Fig. 9 の挙動は押した方向より右にずれながら転がっていくものになっている．生成した動作と物体挙動を新たな知識として追加学習する枠組みを付加することでオープンエンドな知識の獲得が可能になる．また，本実験結果はノーマンが提案する知覚されたアフォーダンス [6] に関して，円筒物体の姿勢や大きさが与える押し動作をロボットが自律的に獲得した結果になっている．

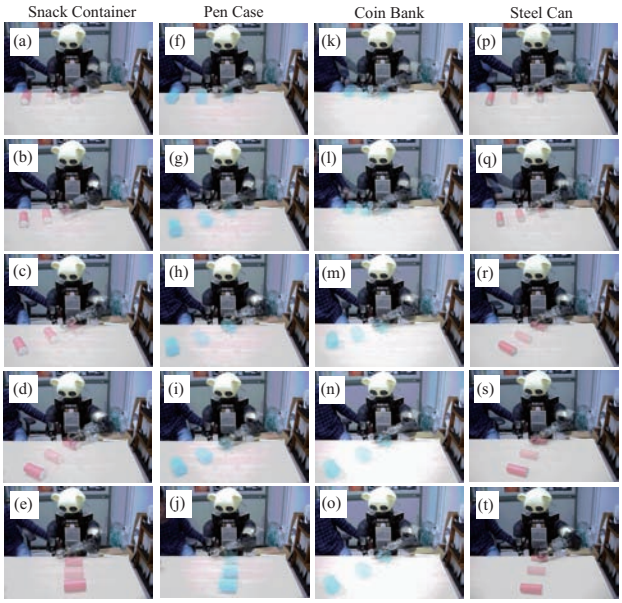

Fig.9 Generated Robot and Object Motions

5. おわりに

本稿では，物体挙動の一貫性・予測信頼性に基づいて動作生成する手法を提案した．提案手法は

1. ロボット－物体ダイナミクスを RNNPB で学習し，PB 空間を形成
2. 予測信頼性が高い PB（ロボット動作と物体挙動を圧縮）を PB 空間で探索
3. 物体画像と予測信頼性が高い PB を三層のニューラルネットワークで連結

の三つの学習段階によって成り立っている．動作生成は，物体画像を三層のニューラルネットワークに入力することで PB を計算し，得られた PB をもとに RNNPB でロボット動作を連想することで実現した．物体種類・物体姿勢が異なる 20 通りの実験全てについて一貫した「転がす」動作を生成することに成功した．

今後の課題として，物体特徴量を自動抽出することが挙げられる．本実験では物体挙動を記述する特徴量として重心と慣性主軸を用いたが，物体形状と挙動の複雑さに応じて，より多くの特徴量を用いることが必要になる．特徴量の自動抽出を実現することでより多種類の物体を扱えるようになることが期待できる．さらに，押し動作に加え，把持動作など，他の動作についても本手法の有効性を示していきたい．

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．

[1] J. Hawkins and S. Blakeslee, On Intelligence, *Times Books*, 2004. ISBN: 0805078533.
[2] J. Tani and M. Ito, Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment, in *IEEE Trans. on Systems, Man and Cybernetics Part A*, Vol. 33, No. 4, pp. 481-488. 2003.
[3] M. Jordan, Attractor dynamics and parallelism in a connectionist sequential machine, in *Eighth Annual Conference of the Cognitive Science Society*(Erlbaum, Hillsdale, NJ), pp. 513-546, 1986.
[4] D. Rumelhart, G. Hinton, and R. Williams, Learning internal representation by error propagation, in *D.E. Rumelhart and J.L. McLelland, editors Parallel Distributed Processing (Cambridge, MA: MIT Press)*, 1986.
[5] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, Robovie: an interactive humanoid robot, in *International Journal of Industrial Robotics*, Vol. 28, No. 6, pp. 498-503, 2001.
[6] D. A. Norman, The psychology of everyday things, *New York: Basic Books*, 1988. ISBN: 0465067093.