# 注意に基づく物理的因果性の獲得に向けた状況依存型予測器の学習

## Learning of the situation-dependent predictor for acquiring physical causality based on attention

○ 藤田 徹也 (阪大)　　正 荻野 正樹 (JST ERATO)
正 福家 佐和 (阪大)　　正 浅田 稔 (阪大，JST ERATO)

Tetsuya FUJITA, Osaka University, 2-1, Yamadaoka, Suita, Osaka
Masaki OGINO, JST ERATO Asada Project
Sawa Fuke, Osaka University
Minoru ASADA, Osaka University, JST ERATO Asada Project

Physical causality, like object permanence, is one of the most important concepts that robots working in our daily lives should have. The fundamental faculty enabling physical causality is prediction of objects' movements. If an object or the environment changes, the prediction should cope with this change. We propose a situation-dependent predictor which can change prediction depending on the object's shape and the environment. In order for a robot to obtain the concepts by itself, we propose an attention model with which a robot decides what to attend and when to serch a new object. This paper presents this situation-dependent predictor and shows that the system can predict object's flow depending on situation, and consequently robots acquires the physical causality with the attention model.

***Key Words:*** Anticipation, Restricted Boltzmann Machine, Object Permanence

## 1　はじめに

近年，ロボットが活躍する場面は多岐にわたり，実際の環境で行動できるロボットも強く期待されている．しかし，ロボットが複雑な環境の中で様々な行動を選択するためには，外界の事物に関する知識だけでなく環境を構成するルールを理解し，更にそれを行動選択に反映する能力が必要である．

人間の場合，幼児は生まれてすぐにはこのような知識は持っていないと考えられる．しかし，成長と共に積極的に外界に働きかけ，環境との関わりの中で得られた知覚情報を利用することによって，知識や概念を自律的に獲得していると考えられる．

認知発達ロボティクスの分野では，幼児の物体永続性に着目し，その振る舞いを実現するいくつかのモデルが提案されている [1] [2]．これらの研究は特定の環境内の特定の対象に対する注意のみを扱っている．しかし，注目対象の特徴が同じでも，環境が異なればその挙動は異なり，また逆に環境が同じでも対象が異なれば挙動は変化する．つまり実際の物体永続性，またより一般に様々な物理的因果性を獲得するためには多様な環境，多様な対象に対する予測ができることが必要である．

そこで本研究では，ニューラルネットワークの一種である Restricted Boltzmann Machine(以下 RBM) [3] [4] を用いた状況依存型予測器を提案する．状況依存型予測器とは，環境情報と注視対象の特徴を合わせた変数として「状況」を考え，状況と対象の挙動を統合する予測器である．本予測器が，各々の状況を踏まえ，様々な現象に対して，その予測に応じた注意が可能であることを示し，更に自律的に様々な因果性を獲得するための注意モデルを提案する．

## 2　RBM を用いた予測とその性能

### 2.1　RBM を用いた予測

RBM は表出層と隠れ層からなるニューラルネットワークの一種で，その特徴は連想記憶型ネットワークである点，そして自己組織化が可能である点にある．連想記憶型であるためノイズにロバストで，補完能力がある．この性質を利用し，予測器は現在の状況および注目物体の現在の状態，そして次状態のセットを学習する．これにより，現在の状況，状態のみから次状態を出力できるようになる．

状況に依存して予測を変化させるために，環境を学習する RBM(環境モジュール) と予測を学習する RBM(予測モジュール) の 2 種類の RBM を用意する．図 1 に予測器の全体像を示す．以下，予測器の各モジュールについて説明する．

### 2.2　環境モジュールと予測モジュール

注目物体画像（ここでは橙色のボール）と周辺画像をガボールフィルタにかけ，画像上のどの部分が際立っているかで 2 値化したものを入力 $\boldsymbol{v}^{<env>}$ とする．具体的には傾き $\psi = [0°, 45°, 90°, 135°]$ のガボールフィルタにより求められたエッジ画像をそれぞれ $5 \times 5$ ユニットに分割する．そして各ユニットについて，領域内の画素値を全て足し合わせた $p_j$ を最大の画素値 $p_{max}$ で除した $p_j/p_{max}$ をそのユニットの活性度 $a_j^{<env>}$ とし，環境値 $I_j$ は $a_j^{<env>}$ がある閾値 $th$ を超えた場合 1 を，超えない場合は 0 として求める．

$$I_j = \begin{cases} 1 & if \quad a_j^{<env>} > th \\ 0 & else \end{cases} \tag{1}$$

環境モジュールの入力 $\boldsymbol{v}^{<env>}$ は物体画像の環境値 $\boldsymbol{I}_{<obj>}$ と周辺画像の環境値 $\boldsymbol{I}_{<around>}$ を組み合わせた $\boldsymbol{v}^{<env>} =$

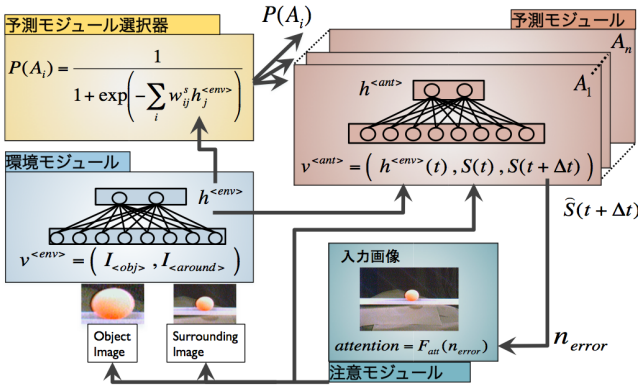



**Fig.1** Model of situation-dependent predictor

$(\boldsymbol{I}_{<obj>}, \boldsymbol{I}_{<around>})$ となる．

RBM において隠れ層は入力層の特徴抽出器の役割を果たす．よって物体画像と周辺画像を入力としたときの隠れ層の発火パターン $\boldsymbol{h}^{<env>}$ は，環境情報と物体情報の特徴を合わせた変数となる．以下これを「状況」と呼ぶ．

予測モジュールは前ステップから現在のステップへの挙動を学習するため，前ステップにおいて環境モジュールから得られた状況 $\boldsymbol{h}^{<env>}$ と，注目物体の位置と速度 $\boldsymbol{S}(t-1)$，そして現在のステップにおける位置 $\boldsymbol{S}(t)$ のデータセットを学習する．$\boldsymbol{S}(t)$ は $x$ 座標と $y$ 座標 $(\boldsymbol{x}, \boldsymbol{y})$，$x$ 方向速度と $y$ 方向速度 $(d\boldsymbol{x}, d\boldsymbol{y})$ を成分とする．

座標のノード表現について，画面分割が $n \times m$ のとき，$(\boldsymbol{x}, \boldsymbol{y})$ はそれぞれ $\boldsymbol{x} = (x_0, x_1, ..., x_{n-1})$，$\boldsymbol{y} = (y_0, y_1, ..., y_{m-1})$ で構成され，画面サイズが $W \times H$，注目物体の中心座標が $(x, y)$ とすると，

$$x_i = \begin{cases} 1 & if \quad \frac{x}{W/n} \leq i < \frac{x}{W/n} + 1 \\ 0 & else \end{cases} \tag{2}$$

$$y_j = \begin{cases} 1 & if \quad \frac{y}{H/m} \leq j < \frac{y}{H/m} + 1 \\ 0 & else \end{cases} \tag{3}$$

とすることで座標を表現する．

速度について，画面分割が同じく $n \times m$ のとき，$x$ 方向速度は $n$ 番目のノードを静止状態とし，$-x$ 方向に $d_n$ マス動く場合は $n - d_n$ 番目のノードで，$x$ 方向に $d_n$ マス動く場合は $n + d_n$ 番目のノードで表現する．ここで $d_n$ は $0 \sim n-1$ までの値をとるので $d\boldsymbol{x}$ は $2n-1$ 個のノードで構成される．$y$ 方向についても同様で，具体的には $d\boldsymbol{x} = (dx_0, dx_1, ..., dx_{2n-2})$，$d\boldsymbol{y} = (dy_0, dy_1, ..., dy_{2m-2})$ で構成され，1 フレーム間の注目物体の中心座標の変化が $(dx, dy)$ とすると

$$dx_i = \begin{cases} 1 & if \quad \frac{dx}{W/n} + n - 1 \leq i < \frac{dx}{W/n} + n \\ 0 & else \end{cases} \tag{4}$$

$$dy_j = \begin{cases} 1 & if \quad \frac{dy}{H/m} + m - 1 \leq j < \frac{dy}{H/m} + m \\ 0 & else \end{cases} \tag{5}$$

とすることで速度を表現する．

1 ステップ間のフレーム数を時間分割の粗さと呼ぶ．そして，画面分割や時間分割の粗さの異なる予測モジュールを多数用意する．その予測器群の中からいずれの予測器を用いるかは，環境に依存した予測器の信頼度によって決定する．

$i$ 番目の予測器の信頼度 $c_i$ は，環境モジュールの隠れ層 $\boldsymbol{h}^{<env>}$ で表される状況から想起され，

$$c_i = \sum_j w_{ij}^s h_j^{<env>} \tag{6}$$

と表される．ここで重み $w_{ij}^s$ は各予測器における予測の誤差 $\Delta r_i$ と状況 $h_j^{<env>}$ によって以下のように Hebb 則によって更新される．

$$\Delta w_{ij}^s = \epsilon(e^{-\Delta r_i} \times h_j^{<env>}) \tag{7}$$

ここで $\epsilon$ は学習率である．各予測器の信頼度 $c_i$ を用いて各予測モジュール $A_i$ の活性度 $P(A_i)$ を

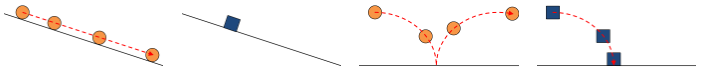
$$P(A_i) = \frac{e^{c_i}}{\sum_i e^{c_i}} \tag{8}$$

と表し，そのうち最大の $P(A_i)$ を持つモジュールをその状況で用いる予測モジュールとする．

### 2.3 実験と結果

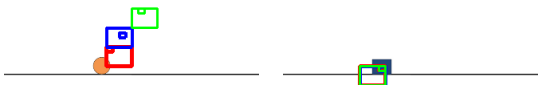
まずシミュレーションによって予測器が状況に応じて予測を変化させることを確かめた．ここでは環境 2 種類 (坂，水平)，物体 2 種類 (丸，四角) の計 4 状況を用意した．図 2 に挙動の概要を示す．各状況 20 フレームの移動を予測し，予測モジュールは画面分割 2 種類（$40 \times 40$, $10 \times 10$)，時間分割 3 種類（1, 2, 4 フレーム）の計 6 種類のモジュールを用意する．なお本実験では，注目する物体として予め丸と四角を設定している．

図 3 に予測結果の一例を示す．色枠が予測結果で，時間分割 [1, 2, 4] の順に [赤，青，緑] である．また枠の大きさの違いは画面分割の違いである．この 2 状況において，これ以前の物体の挙動は状況によらず同じであるにもかかわらず，物体形状に応じて予測を変化させていることがわかる．



(a) Environment1, Object1　(b) Environment1, Object2　(c) Environment2, Object1　(d) Environment2, Object2

**Fig.2** Situation of simulation

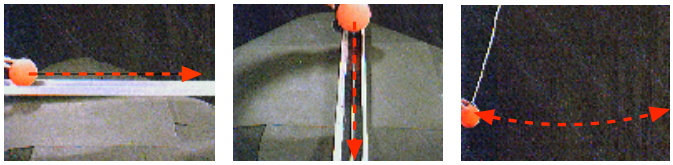


(a) Prediction of object1　(b) Prediction of object2

**Fig.3** Difference of predictions on object's form

次に実機による実験を行った．図 4 に示す 3 状況 (ボールの水平，鉛直，振り子運動) について予測実験をおこなった．IEEE1394 カメラを用い，33[msec/frame] で取得した 90 フレームの画像を用いる．予測モジュールは画面分割 2 種類（$40 \times 40$, $10 \times 10$)，時間分割 3 種類 (5, 10, 20 フレーム) の計 6 種類用意し，注目物体として予めボールを設定してある．

実験結果 (図 5) について，縦軸 Error Rate は予測と実際の位置のずれを正規化したもので，横軸 Learning Times は全状況のデータセットを全て学習し 1 回の学習としている．凡例については，時間分割を示す線色と画面分割を示す線種を組み合わせて参照されたい．これより全ての RBM で予測が学習できていることがわかる．また学習結果の例として図 6 に，画面分割 $10 \times 10$ の各予測器が，状況 1 (図 4(a)) のボールの動きを予測している様を 10 ステップごとに示す．色枠が予測結果で，時間分割 [5, 10, 20] の順に [赤，青，緑] である．各場所に応じてその先を予測していることがわかる．尚，緑枠は 20 ステップごとの予測なので 20 ステップ毎に表示される．



(a) (b) (c)

**Fig.4** Situation of experiment

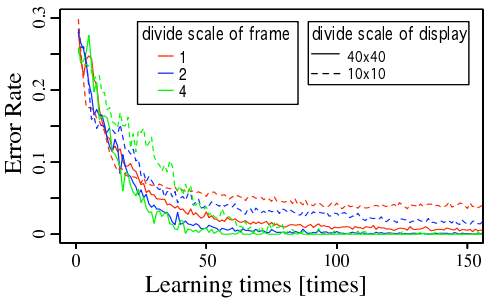



**Fig.5** Error rate of experiment

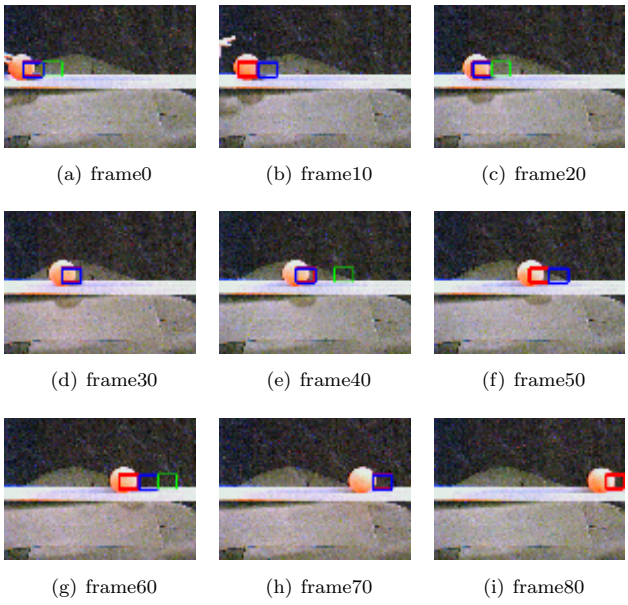
最後に追加学習の実験として，まず状況 1(図 7(a)) を充分学習した後，学習データに状況 2(図 7(b)) を追加する実験を行った．両状況とも 68 フレームずつで予測モジュールの種類は上と同じである．図 7(c) は状況 2 追加直後における各フレームの誤差である．壁に当たる前の 35 フレーム (図 7(d)) までは，状況 1 と大差無いため誤差は小さいが，壁に当たった 40 フレーム (図 7(e)) で大きな誤差を出す．図 7(f) に 40 フレームにおける予測結果を示す．このように環境の変化を無視し壁の向こう移動すると予測しているが，学習終了後は壁がなければボールの先を予測し (図 7(g))，壁がある時は壁を検知し，手前で停止すると予測し分けている (図 7(h))．

以上のように，本予測器は環境や物体によって予測を変化させていることがわかる．

## 3 注意に基づく予測とその性能

### 3.1 注意モジュールの設計

前述のような物体の運動予測モデルを学習するにあたり，事前に実験者が注目物体を与えずとも，ロボットが自発的に注意を向け入力値を求めることが望ましい．人において注意の遷移には解放，移動，保持の 3 つのステップがあり [5]，注意をモデル化するにあたり，注意の保持と解放の指針が重要となる．す

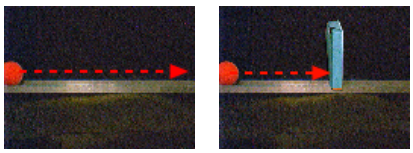


(a) frame0 (b) frame10 (c) frame20

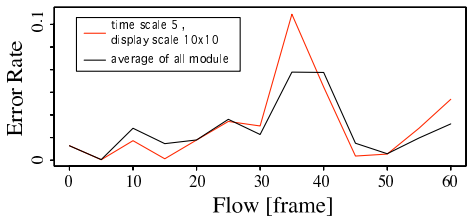
(d) frame30 (e) frame40 (f) frame50

(g) frame60 (h) frame70 (i) frame80

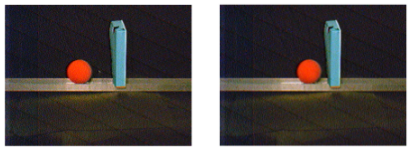
**Fig.6** Predictions of horizontal motion



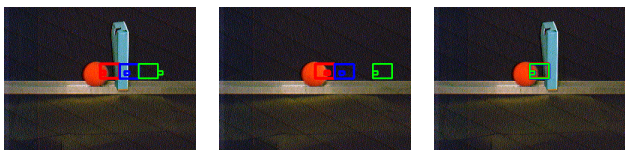
(a) Situation1 (b) Situation2




(c) Error rate of each flow



(d) Frame35 (e) Frame40



(f) Prediction before learning (g) Prediction of situation1 after learning (h) Prediction of situation1 after learning

**Fig.7** Adding learning

ぐに学習でき予測が可能となるもの（静止状態など）やすでに学習が終了したもの，また挙動が完全にランダムで全く予想できないものからは注意を外すことが望ましい．そして，その挙動が規則的で，かつまだ完全に学習できていない物体に対して注意を維持する必要がある．

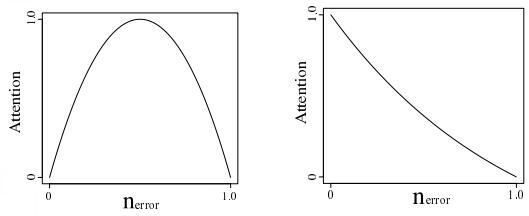
注意の保持について，注目物体の決定は顕著性マップ [6] を用いる．画面の中から色や明度，動きの点で顕著な部分に着目し，着目点と同色の点の集合を注目物体とする．こうして切り出した画像を物体画像，物体中心で大きさが入力画像の縦横半分の画像を環境画像とする．

注意の解放について，予測モジュール群の正誤から求められる注意値がある閾値を下回れば注意を解放するモデルを考案した．まず，式（9），図 8(a) のような関数で注意値を決定した．

$$attention = -\frac{e^{\frac{n_{error}}{n_{total}}} \times e^{\frac{n_{match}}{n_{total}}} + e + 1}{-2e^{0.5} + e + 1} \tag{9}$$

(ここで $n_{error}$ とは全ての予測モジュールについて，予測を失敗させたモジュールの数，$n_{match}$ は成功数，$n_{total}$ は予測モジュールの総数)．図 8 について横軸が予測の誤りであり，予測の正誤が半々になった時に注意値が最大になるのがわかる．これにより必要な注意を実現できると期待したが，本実験のようにそもそも注目し得る物体が 1 つしかない場合，試行直後の学習が進み過ぎてしまい，すぐに注意を外すため，結果一連の挙動を追うような予測器は学習できなかった．そこで，式（10），図 8(b) で表される関数を用い，予測が成功すれば注意値が大きくなるようモデルを変更した．

$$attention = \frac{e^{\frac{n_{match}}{n_{total}}} - 1}{e - 1} \tag{10}$$




(a) Attention function1　　(b) Attention function2

**Fig.8** Attention function

### 3.2　実験と結果

注意値を求める関数を変更したところ，学習するにつれ注意が外れにくくなり，ボールの挙動をある程度学習できた．結果の例を図 9 に示す．ここで黒枠が現在注意を向けている物体である．ボールの動く先を予測できていることがわかる．

## 4　結果と今後の課題

本研究では，環境情報と物体情報の特徴を抽出した状況と物体の挙動を統合する，「状況依存型予測器」を提案した．ロボットは複数の環境におけるボールの運動の予測を，状況に応じて変化させることが可能となった．更に，RBM への入力情報を選択する行動として注意メカニズムを考慮して実験をし，自発的に注目物体を決定しながら，その物理的因果性を獲得できる可能性を示唆した．

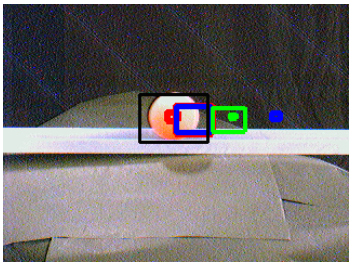


**Fig.9** A prediction of horizontal motion

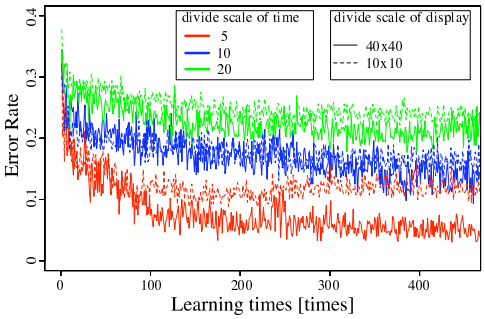



**Fig.10** Error rate of experiment with attention module

本研究では，実機による実験でも余計な物体は排除し，録画した数試行を繰り返し学習させるという比較的シミュレーションに近い状態で実験を行った．また常に一つの物体にしか注意を向けられないことから，試行開始直後の学習が完了すると注意を完全に別の物体に移動させてしまい，後半の挙動の学習が起こらなかった．そこで注意モデルを再設計したわけであるが，実環境においては注目しやすい物体が，常に因果性を獲得すべき対象とは限らない．慣れの度合いによって興味を失う habituation 関数を導入するなど実環境で用いるにはさらに踏み込んだモデルが必要である．

## 参考文献


[1] M. Schlesinger. A lesson from robotics: Modeling infants as autonomous agents. *Adaptive Behavior vol.11, no.2*, 2003.

[2] Andrew Lovett and Brian Scasselatti. sing a robot to reexamine looking time experiments. *4th International Conference on Development and Learning (ICDL).*, 2004.

[3] Geoffrey E. Hinton. Learning multiple layers of representation. *TRENDS in Cognitive Sciences. Vol.11 No.10, 428-434*, 2007.

[4] Geoffrey E. Hinton et al. A fast learning algorithm for deep belief nets. *Neural Comput. 18, 1527-1554*, 2006.

[5] 甘利俊一. 認識と行動の脳科学. 東京大学出版会, 2008.

[6] Laurent Itti, Nitin Dhavale, and Frederic Pighin. Realistic avatar eye end head animation using a neurobiological model of visual attention. *Proceedings of SPIE*, 2003.

[7] Brody L. R. Visual short-term cued recall memory in infancy. *Child Development, 52, 242-250*, 1981.

[8] A. Diamond and B. Doar. The performance of human infants on a measure of frontal cortex function, the delayed response task. *Developmental Psychobiology 22 (1989) (3), pp. 271 294*, 1989.

[9] Schwartz B. B. and Reznick J. S. Measuring infant spatial working memory using a modified delayed-response procedure. *Memory, 7, 1-17.*, 1999.

[10] J. Steven Reznick, Judy D. Morrow, Barbara Davis Goldman, and Jessica Snyder. The onset of working memory in infants. *Infancy, Volume 6, Issue 1 August 2004.*, 2004.