

移動ロボットを用いた物体識別のための探索行動の学習

Learning of exploratory behavior for object discrimination and its application to a mobile robot

郷古 学*¹ 小林祐一*² 金 天海*³
 Manabu Gouko Yuichi Kobayashi Chyon Hae KIM

*¹東北学院大学 *²静岡大学 *³ホンダリサーチインスティテュートジャパン
 Tohoku Gakuin University Shizuoka University Honda Research Institute Japan

This paper presents a active perception model for a robot. Proposed model can learn automatically an exploratory behavior for object recognition. To confirm the effectiveness of the proposed model, we carried out the mobile robot simulation. In the simulation, the robot was able to learn the exploratory behavior which can distinguish the objects of different form. Additionally, it has been confirmed that the acquired exploratory behavior can recognize the objects appropriately.

1. はじめに

現在、実環境下で生活支援を行う汎用型ロボットに関する研究が盛んに行われている。ロボットが環境中の様々な物体を認識し、それらを利用することができれば、より幅広い作業を行うことが可能となる。本研究では、このような物体操作の基礎となる、ロボットによる物体の知覚について扱う。本稿における「物体の知覚」とは、カテゴリ（類似する特徴を有する物体の集合）に関する知識をもとに、ロボットが知覚対象となる物体と、他の物体とを区別することを意味する。物体の知覚の実現には、どのようなカテゴリを形成するかが重要となる。

一般に、カテゴリの形成は二つに大別できる [池上 07]。一つ目は、知覚対象となる物体から、知覚主体が受動的に得る感覚情報をもとにカテゴリを形成する方法である。もう一方は、知覚主体自らが探索行動 (exploratory behaviors) と呼ばれる行動により対象の特徴を抽出し、カテゴリを形成する方法である (ダイナミカルカテゴリ)。探索行動を基礎とする知覚は能動知覚と呼ばれ、これまでの研究により、知覚主体の身体特性をより反映した表象の獲得が可能であることが示唆されている [Held 63, Turvey 96, Gibson 62]。

近年、ロボットによる能動知覚の研究が盛んに行われている。これまでに、物体に対して「振る」や「落とす」などの探索行動を行い、それにより生じる音響情報を用いて物体を知覚する方法 [Sinapov 09, Takamuku 08, Griffith 10] や、ロボットに物体を押させて、その際の物体の挙動 (転がる, 倒れるなど) の違いに応じて、カテゴリを形成する方法が提案されている [Nishide 08]。また、走行中の移動ロボットに搭載されたカメラの画像から、環境中の物体についてのカテゴリ形成を扱った研究 [Tsukada 10] や、単に能動知覚による物体の識別のみならず、識別された物体を道具として利用する手法 [Stoytchev 05] や、得られたカテゴリを利用して、探索行動を生成する方法も提案されている [Atil 10]。

これらの従来研究では、ロボットの探索行動は設計者によりあらかじめ設計されている。探索行動を設計するためには、その行動により、知覚対象となる物体からどのような特徴が抽出可能であるかを事前に想定する必要がある。また、探索行動により得られる特徴は、物体だけではなく、実際に知覚を行う

ロボットの身体特性 (センサの種類や身体機構) にも大きく依存する。つまり、探索行動を設計するためには、物体とロボット双方に関する十分な事前知識が必要となる。

以上の考察のもと、本研究では、探索行動の学習による獲得が可能な能動知覚モデルを提案する。提案モデルでは、ロボットは与えられる複数の物体に対して、それらの知覚 (区別) が可能な探索行動を学習により獲得する。提案モデルでは学習に際し、各物体がどのカテゴリに帰属するかは既知であるとする。これは、他者が自らの有するカテゴリと類似するカテゴリをロボットに獲得させる場合を想定している。ただし、各物体から得られる具体的な特徴は未知であり、探索行動に依存して決まる。そのためロボットは、異なるカテゴリに属する物体同士から得られる特徴の差異が、より大きくなるように探索行動を学習する。

探索行動の学習を扱った従来研究として、[Morimoto 04] では、移動ロボットを用いて床に描かれた図形 (四角形と三角形) を区別する探索行動の学習を行っている。ロボットは、四角形の上には長く滞在し、逆に、三角形の上は避けるような行動、つまり、各図形を区別する探索行動を学習により獲得している。また柴田 [柴田 01] は、視覚センサを対象に、複数パターンの識別を実現するセンサの移動 (行動) を強化学習により学習させた。これらの研究では、識別の正否に関する情報を探索行動の学習の基準として用いているが、提案モデルはカテゴリ形状を基準に学習を行っている点が異なる。

本稿では、移動ロボットを用いたシミュレーション実験により、探索行動の学習が可能であることを示すとともに、学習した探索行動を用いて物体の識別が可能であることを示す。

2. 提案モデル

2.1 モデルの概略

提案する能動知覚モデルを図 1 に示す。モデルはロボット、探索行動を生成するコントローラ、物体の特徴を一時的に記憶するバッファ、および特徴空間から構成されている。探索行動の学習は、ロボットに対して複数の異なる物体を、一定時間毎に順番に提示して行う。本研究では簡単のため、提示される各物体は、それぞれ異なるカテゴリに属するものとし、学習時に提示される物体の種類 (n =全カテゴリ数) も既知とする。

学習時には、まずロボットが物体に探索行動 (学習初期ではランダム) を行い特徴を得る。特徴空間上には各物体に対応す

連絡先: 郷古 学, 東北学院大学工学部 機械知能工学科,
 宮城県多賀城市中央 1-13-1,
 gouko@tjcc.tohoku-gakuin.ac.jp

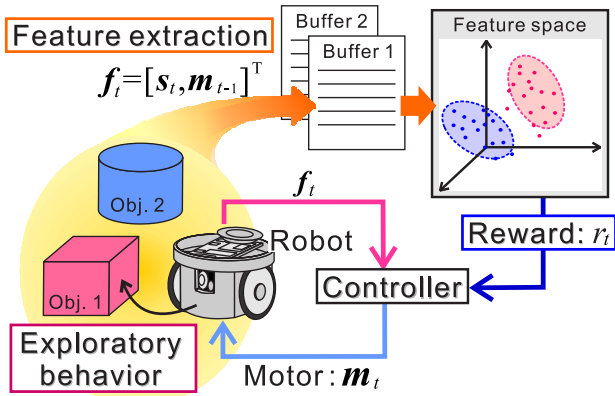


図 1: Proposed model.

る n 個のカテゴリが存在する。ただし、学習初期の各カテゴリの形状はランダムである。各カテゴリは、物体から得られた特徴にもとづき、その形状を逐次修正する。次に、各カテゴリの分離の度合いから報酬を求め、強化学習によりコントローラーの学習を行う。以上の学習を繰り返すことで、提案モデルは、それぞれのカテゴリがより分離するような、つまり、異なるカテゴリに帰属する物体同士からは差異の大きな特徴が得られるような探索行動の獲得が期待できる。次節以降では、特徴抽出とカテゴリ形成、探索行動の学習について詳述する。

2.2 特徴抽出とカテゴリ形成

提案モデルにおいて、ロボットは単位時間毎にセンサ情報を観測し、コントローラーから出力されるモータコマンドにもとづき動作を行う。時刻 t におけるセンサ情報を s_t 、モータコマンドを m_t とする。

ロボットは提示された物体 i (カテゴリ i に属する物体) に対して毎時動作を繰り返しながら、その物体の特徴を抽出する。時刻 t において得られる物体の特徴を、特徴ベクトル $f_t = [s_t, m_{t-1}]^T$ と表す。 f_t はバッファに一時的に記憶される。提案モデルは各カテゴリに対応したバッファを持ち、各バッファの長さ (記憶可能なベクトル数) は L で、記憶の古い順に新しい特徴ベクトルが上書き (更新) される。物体 i から得られた f_t は、カテゴリ i に対応するバッファ i に記憶される。

続いて、カテゴリ形成 (形状の修正) を行う。特徴空間上の各カテゴリは、対応するバッファに記憶されているデータにもとづき、その形状を修正する。本稿では、各カテゴリの形状は正規分布で表されるとし、カテゴリ i に対応する正規分布を $p_i(\mu_i, \Sigma_i)$ とする。ここで μ_i と Σ_i は正規分布の平均ベクトルおよび分散共分散行列である。これらのパラメータを、バッファ i が更新される度に、そこに記憶されている L 個の特徴ベクトルから最尤推定により求める。

2.3 探索行動の学習

探索行動の学習、すなわちコントローラーの学習には強化学習 [S. Sutton 98] を用いる。強化学習は、ロボットが試行錯誤的に行動を繰り返しながら、目的に応じた状態-行動マッピング (コントローラー) を獲得する学習手法である。強化学習では、ロボットは毎時行動する度に、その行動に対する評価として外部 (環境) から報酬を与えられる。学習に際し、設計者は目的の達成状況に応じた報酬を設定するだけでよく、ロボットは試行錯誤により、得られる報酬を最大化するようなコントローラーを学習していく。

本稿では、強化学習の一手法である Actor-Critic 法を用いた。Actor-Critic 法は行動空間が連続空間であっても適用可能な手法である。この手法では、観測される状態から、得られる報酬の期待値を推定する機構 (critic) と、得られる報酬が最大となるような状態-行動マッピング (actor, コントローラーに対応) の 2 つを同時に学習していく。本研究では、特徴ベクトル f_t を時刻 t における状態として用いた。また後述の実験では、critic および actor は共に 3 層構造のフィードフォワードニューラルネットワークによって構築した。

学習に用いる報酬の設定について述べる。本研究では、以下の二つの条件を満たす探索行動の学習を目指す。

1. 同一カテゴリに帰属する物体からは常に類似する特徴が得られる。
2. 異なるカテゴリに帰属する各物体から得られる特徴同士の差異が大きいの。

このような探索行動を学習するために、各カテゴリのカテゴリ内分散 (within-class variance) σ_W^2 と、カテゴリ間分散 (between-class variance) σ_B^2 を用いて報酬を設定した。

$$\sigma_W^2 = \frac{1}{nL} \sum_{i=1}^n \sum_{l=1}^L (d_i^l - \mu_i)^T (d_i^l - \mu_i) \quad (1)$$

$$\sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (\mu_i - \mu_{all})^T (\mu_i - \mu_{all}) \quad (2)$$

ここで μ_{all} は、各バッファに記憶されている全ての特徴ベクトルの平均であり、 d_i^l はバッファ i に記憶されている l 番目の特徴ベクトルである。 σ_W^2 は各カテゴリの平均的な広がりを表し、 σ_B^2 は各カテゴリ間の広がりを表している。これらの分散を用い、時刻 t における報酬 r_t を以下の式で設定した。

$$r_t = \sigma_B^2 / \sigma_W^2 \quad (3)$$

報酬は、クラス内分散 σ_W^2 が小さいほど、また、クラス間分散 σ_B^2 が大きいほど大きい値となる。このような報酬を用いることで、上述の条件を満たす探索行動の獲得が期待できる。

3. 実験

3.1 シミュレーション設定

移動ロボットを用いたシミュレーション実験により、提案モデルの有効性を検証する。実験では形状の異なる二つの物体を用いて、探索行動の学習が可能かどうかを確認する。

シミュレーション環境はロボットシミュレータ Webots6 (Cyberbotics 社製) を用いて構築した。実験の様子を図 2 に示す。移動ロボットは円筒形で、独立して回転する 2 つの車輪を持っている。また、周囲に 8 つの距離センサ (測定範囲 7cm) を有しており、各センサはロボットと物体との距離が小さいほど大きな値を出力する。本実験では、センサ s_t は、各距離センサの出力に対応した 8 次元ベクトルとし、モータコマンド m_t は、左右の車輪の出力に対応した 2 次元ベクトルとする。なお、 s_t および m_t の各要素は 0 から 1 の範囲になるように正規化を行った。モータコマンドは、0 および 1 が最大速度による車輪の後転と前転にそれぞれ対応する。各バッファのバッファ長 L は 100 とした。

実験では、図 2 右に示す直方体 (rectangular parallelepiped, object 1) と円柱 (cylinder, object 2) の 2 つの物体をロボット

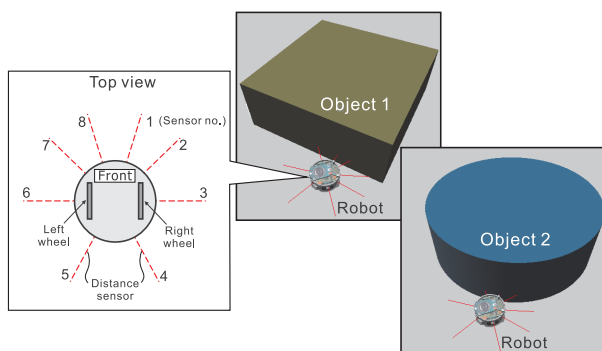


図 2: Experimental setup.

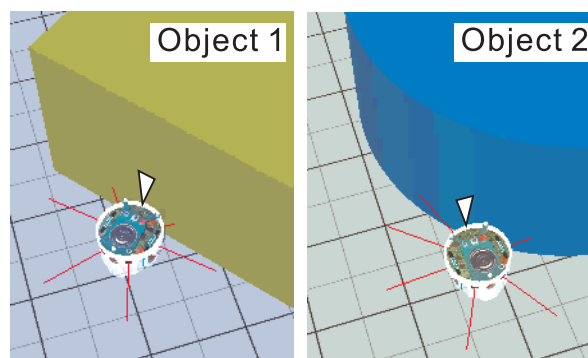


図 4: Behavior of the robot.

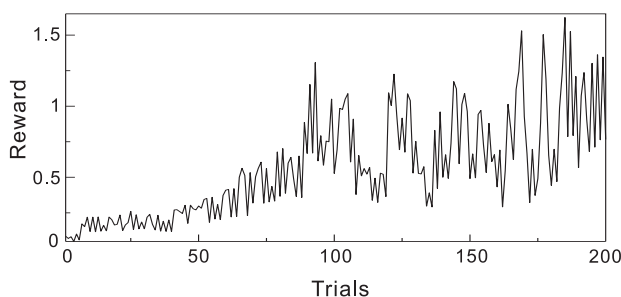


図 3: Change of average reward.

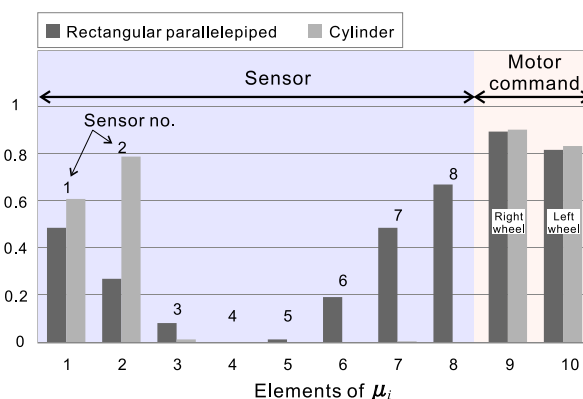


図 5: μ_1 and μ_2 .

に交互に提示して探索行動の学習を行った．直方体と円柱はそれぞれカテゴリ $i = 1, 2$ に帰属する．学習は物体 i およびロボットを初期位置 (図 2 右) に配置した後，以下の手順で行う．

1. s_t を観測し，特徴ベクトル $f_t = [s_t, m_{t-1}]^T$ を求める ($t = 0$ の場合， $m_{t-1} = 0$ とする) ．
2. 特徴ベクトル f_t をバッファ i に記録し，同バッファに記憶されている全データから μ_i と Σ_i を推定する ．
3. (3) 式より報酬 r_t を求め，actor と critic を更新する ．
4. f_t から m_t を求め，ロボットを動作させる ($t \leftarrow t + 1$) ．

1~4 までの手順をシミュレーションの単位時間 (1step = 360msec) 毎に行う．学習は，ひとつの物体に対して 100step 行う (これを 1 試行と呼ぶ) ．ある物体に対する学習が完了した後，物体を交換し，ロボットを初期位置に配置して同様の学習を行う．実験では，上記の学習を全 200 試行 (各物体に対して交互に 100 試行ずつ) 繰り返した ．

なお，学習においてロボットの初期位置は固定であるが，ロボットの向きは試行毎にランダムとした ．また，ロボットが物体から離れてしまった場合，つまり，いずれの距離センサの測定範囲内にも物体が存在しない場合には，その試行を打ち切り，物体を入れ替えて新しい試行を開始するとした ．

3.2 実験結果および考察

図 3 にロボットの獲得報酬 (試行毎の平均) と試行数の関係を示す ．試行を重ねるにつれて獲得報酬が上昇している様子が確認できる ．次に，学習後 (200 試行後) のロボットに各物体を提示し，その挙動を確認した ．その結果，ロボットは初期位置から回転動作などにより各物体に正対した後，直進し，図 4

に示すような姿勢で接触する様子が確認された (図中の三角形はロボットの正面を表す) ．

また，学習後の各カテゴリの平均ベクトルを図 5 に示す ．これを見ると，左右の車輪 (モータコマンド) に対応するベクトルの要素 (9, 10) は μ_1 と μ_2 で比較的類似している ．各要素はともに 0.8~0.9 で，これはロボットが前進することを意味している ．一方で，センサに対応する各要素 (1 から 8 まで) の差異が，比較的大きいことが分かる ．特に，7 番目と 8 番目の要素 (センサ No.7 と 8 に対応) をみると， μ_1 では 0.5~0.7 程度であるのに対し， μ_2 では，ほぼ 0 となっている ．

実験に用いたロボットは，全部で 8 つの距離センサのうち 4 つが正面付近に配置されている ．そのため，ロボットが直方体に対して正面から接触すると，図 4 左のように，ロボットの正面にある 4 つのセンサ (No.1, 2, 7, 8) すべてで，その測定範囲内に物体を捉える場合が多く生じる ．一方で，円柱の場合は，ロボットが正面から接触しても直方体のように 4 つのセンサすべてが同時にその測定範囲内に物体を捉えることができない ．つまり，ロボットは学習により，センサが密集している部分を物体に接触させるという探索行動を獲得し，それにより，各物体の形状を反映した特徴を抽出したと考えられる ．

3.3 学習された探索行動を用いた識別実験

探索行動を学習したロボットに対して，四角柱と円柱を提示し，それらをどのように識別するのかを確認した ．本識別実験では，ロボットは提示された物体に対して学習済みの探索行動を行い，特徴ベクトルを取得する ．次に，その特徴ベクトルがどのカテゴリに帰属するかを，学習済みの各カテゴリの正規

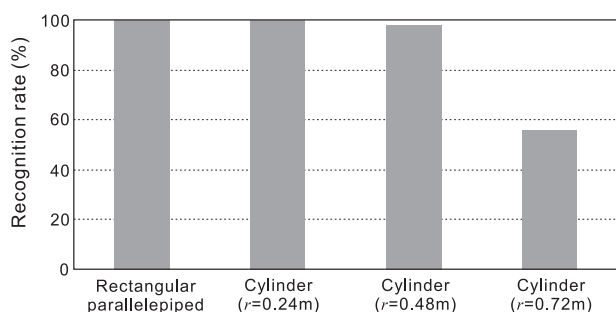


図 6: Recognition rate.

分布を用いてベイズ決定則により決定した。ロボットは提示された物体に対して最大で 100step 行動するものとし、毎時得られる特徴ベクトルに対して、カテゴリ 1, 2 のいずれかに属するかを求め、それらの多数決から最終的な識別結果を求めた。なお、以下で示す結果は、ロボットが識別開始時 (物体提示) から、50step 以上接触して (いずれかの距離センサ内に物体を捉えて) いる場合のみを対象としている。また、本実験中には、学習は一切行っていない。

識別実験では、学習時に用いた二つの物体のほか、学習に用いた円柱 (半径 $r=0.24\text{m}$) の半径を 2 倍, 3 倍とした円柱をロボットに与えた。半径の異なるこれらの円柱は、カテゴリ 2 に識別される場合を正解として識別率を求めた。

図 6 に識別率を示す。これらは各物体を 50 回ずつ識別した結果である。図より、学習で用いた二つの物体は 100% の割合で識別できていることが確認できる。半径を 2 倍とした円柱は 98% の割合で識別されており、学習により得られた探索行動が、未学習の物体 (半径の異なる円柱) に対しても妥当な識別が可能であると考えられる。さらに、円柱の半径を大きく (3 倍) にしたところ、識別率が 56% に低下した。これは、半径が大きくなることで、物体側面の曲率が小さくなり (平面に近づく)、その結果、識別精度が落ちたものと考えられる。

4. まとめと今後の課題

本研究では、ロボット自身が探索行動を学習により獲得可能な能動知覚モデルを提案した。移動ロボットによる物体識別シミュレーションを行った結果、形状の異なる各物体から異なる特徴を抽出可能な探索行動を獲得することができた。また、学習により得られた探索行動により物体の識別を行った結果、未学習の物体に対しても妥当な識別が可能であることを示した。

今後の課題として、学習後のロボットに、学習時に使用した物体と類似する物体を与え、探索行動によりそれらの物体を識別させて、学習した探索行動の有効性の検証を行う予定である。また、提案モデルでは、物体の特徴は各時刻のセンサとモータコマンドからなるベクトル (静的パターン) で表現できることが前提である。しかし、対象となる物体によっては、その特徴が時間的に変化する情報 (動的パターン) として表現する必要も考えられる。これについては、特徴空間上の各カテゴリの表現を、動的なモデルに置き換えるなどの対応が必要となる。この点に関しても今後の課題とする。

提案モデルの学習時には、対象物体がどのカテゴリに帰属するかが既知である。これは物体を区別するにあたり、教師にあたる存在が必要であることを意味する。例えば、人間が教師としてロボットに対して自らのカテゴリと類似するカテゴリを口

ポットに形成させるような場合には、提案モデルの方法で対応可能であるが、一方で、ロボット自身にオリジナルなカテゴリを形成させる場合には、教師無しでの学習が必要となる。この点については、各物体から得られる特徴を区別することなく自己組織的にクラスタリングし、その結果を探索行動に反映させる方法が考えられる。この点についても、今後の課題とする。

謝辞

本研究の一部は、科研費若手研究 (B) (24700196) の助成による行われました。ここに謝意を表します。

参考文献

- [Atil 10] Atil, İlkay, Dağ, N., Kalkan, S., and Şahin, E.: Affordances and Emergence of Concepts, in *Proceedings of the 10th International Conference on Epigenetic Robotics*, pp. 11–18 (2010)
- [Gibson 62] Gibson, J. J.: Observations on active touch, *Psychological Review*, Vol. 69, No. 6, pp. 477–491 (1962)
- [Griffith 10] Griffith, S., Sinapov, J., Sukhoy, V., and Stoytchev, A.: How to Separate Containers From Non-Containers? A Behavior-Grounded Approach to Acoustic Object Categorization, in *Proceedings of 2010 IEEE International Conference on Robotics and Automation*, pp. 1852–1859 (2010)
- [Held 63] Held, R. and Hein, A.: MOVEMENT-PRODUCED STIMULATION IN THE DEVELOPMENT OF VISUALLY GUIDED BEHAVIOR, *Journal of Comparative and Physiological Psychology*, Vol. 56, No. 5, pp. 872–876 (1963)
- [Morimoto 04] Morimoto, G. and Ikegami, T.: Evolution of Plastic Sensory-motor Coupling and Dynamic Categorization, in *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, pp. 188–193 (2004)
- [Nishide 08] Nishide, S., Ogata, T., Tani, J., Komatani, K., and Okuno, H. G.: Predicting Object Dynamics From Visual Images Through Active Sensing Experiences, *Advanced Robotics*, Vol. 22, pp. 527–546 (2008)
- [Sinapov 09] Sinapov, J., Wiemer, M., and Stoytchev, A.: Interactive Learning of the Acoustic Properties of Household Objects, in *Proceedings of 2009 IEEE International Conference on Robotics and Automation*, pp. 2518–2524 (2009)
- [S.Sutton 98] S.Sutton, R. and G.Barto, A.: *Reinforcement Learning: An Introduction*, The MIT Press (1998)
- [Stoytchev 05] Stoytchev, A.: Behavior-Grounded Representation of Tool Affordances, in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 3060–3065 (2005)
- [Takamuku 08] Takamuku, S., Hosoda, K., and Asada, M.: Object Category Acquisition by Dynamic Touch, *Advanced Robotics*, Vol. 22, pp. 1143–1154 (2008)
- [Tsukada 10] Tsukada, M., Madokoro, H., and Sato, K.: Unsupervised and Adaptive Category Classification for a Vision-Based Mobile Robot, in *Proceedings of The 2010 International Joint Conference on Neural Networks*, pp. 1–6 (2010)
- [Turvey 96] Turvey, M. T.: Dynamic Touch, *American Psychologist*, Vol. 51, No. 11, pp. 1134–1152 (1996)
- [柴田 01] 柴田克成, 西野哲生, 岡部洋一: Actor-Q アーキテクチャに基づく能動認識学習システム, 電子情報通信学会論文誌 D-II, Vol. J84-D-II, No. 9, pp. 2121–2130 (2001)
- [池上 07] 池上高志: 動きが生命をつくる 生命と意識への構成論的アプローチ, 青土社 (2007)