

視覚の精緻化が導くミラーニューロンシステムの発達モデル

○河合祐司(阪大院) 長井志江(阪大院) 浅田稔(阪大院, JST ERATO)

1. はじめに

ミラーニューロンシステム (MNS) とは他個体の運動を観察しているとき、その運動を自ら実行しているときに共通して活動する神経細胞群である [1]。MNS は脳科学にとどまらず、発達心理学やロボティクスなどの多くの分野で研究されており、MNS のもたらす認知機能が明らかになりつつある。しかし、その一方で MNS の発達メカニズムは未だ解明されていない。

Heyes らのグループ [2] は MNS 発達モデルとして associative sequence learning (ASL) モデルを提案している。乳児は運動表象と視覚表象の連合を学習するときに、他者に模倣されることによって、自己の運動表象とそれに対応する他者運動の視覚表象の連合、すなわち MNS を獲得するというものである。しかし、この ASL モデルは単純かつ概念的なモデルであり、高次の視覚表象、特に自他を区別した視覚表象がどのように構成されるのかについては考慮されていない。

従来の計算論的な MNS モデルとして、自他運動の等価性を用いたモデル [3] では、ロボットはあらかじめ自己の感覚運動マップを学習した後に、他者運動を自己運動と同一のものと認識することで模倣を行っている。しかし、このモデルでは自他運動を区別することができない。ロボット (乳児) の自他の区別はターンテイキングといった社会的振る舞いに必要不可欠であるため、自他運動の等価性のみ着目したモデルでは不十分である。また、視覚入力 that 自他別々に与えられるモデル [4, 5] では、ロボットがどのように自他を区別するのかという問題は説明できない。人の乳児は自他の境界があいまいな状態であるといわれている [6] ことから、乳児が自他を区別していく過程をも再現するモデルが望ましいと考えられる。

そこで本研究では乳児の視覚発達が自他の区別とその対応付け、すなわち MNS の発達をもたらし計算論的なモデルを提案する。ロボットは図 1 のような他者との対面インタラクションによって、自己の運動指令と視覚表象の連合学習を行う。これは ASL モデルと同様の枠組みであるが、我々はこのモデルに視覚表象の精緻化メカニズムを導入する。学習初期のロボットは低い時空間解像度の未熟な視覚を有し、自他の運動に関する視覚表象を形成する。このとき、ロボットは未熟な視覚のため他者の応答行動の遅れや視点の違いを検出できず、自他未分化な視覚表象を獲得すると考えられる。つまりこの期間に獲得される連合は、結果的に自己の運動指令と他者の運動を対応づけたものになる。そして、ロボットの視覚は視覚経験により発達し、徐々に自他の分化した視覚表象を形成する。しかし、ロボットは自他未分化期の経験や自他運動の類似性により、他者の運動と自己の運動指令の対応付け、すなわち MNS を獲得することができると考えられる。

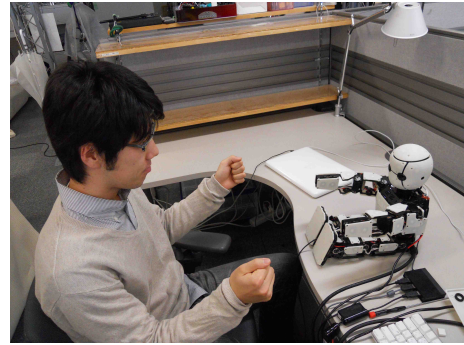


図 1 MNS 発達のための人-ロボットインタラクション

2. 前提条件と問題設定

図 1 のような乳児型ロボットと実験者の対面インタラクションを想定する。ロボットは実験者とのやりとりを通して、自己の運動指令と観察した身体運動の対応関係を学習する。以下に提案モデルの前提条件を挙げる。

- ロボットと実験者は相同の運動レパートリーを持つ。
- 実験者はロボットの運動に対して遅れのある随伴運動をし、一定の割合でロボットを模倣する。
- ロボットの視野は限られており、他者の運動の観測時にロボットは他者の上半身を見る。
- ロボットは自他の運動をオプティカルフローとして検出する。

以上の前提条件のもと、本研究では MNS 発達に関わる 2 つの問題を扱うこととする。1 つは自他分離問題である。ロボットは他者運動の遅れや視点の違いから、自他を分離した視覚表象を学習する。もう 1 つは自他の対応関係問題である。ロボットは自己の運動指令とそれに対応する他者の運動の連合を学習する。

3. MNS 発達モデル

3.1 モデルの概要

図 2 に提案モデルの概要図を示す。このモデルは 2 つの層で構成されており、上側が視覚表象層 (V)、下側が運動表象層 (M) である。ロボットの観測した運動情報は V に入力される。図中で赤い矢印が自己運動 (V_s)、青い矢印が他者運動 (V_o) を表しており、これらの視覚情報がクラスタを形成している。 M は運動指令のレパートリーで構成されている。

ロボットの視覚の精緻化と、 V と M の連合学習は同時進行する。この連合は V 中のクラスタと M 中の運動指令のヘブ学習により獲得される。図 2(a) は学習初期を表しているが、ロボットは未熟な視覚により他者運動の遅れなどの自他運動間の差を検出できず、 V

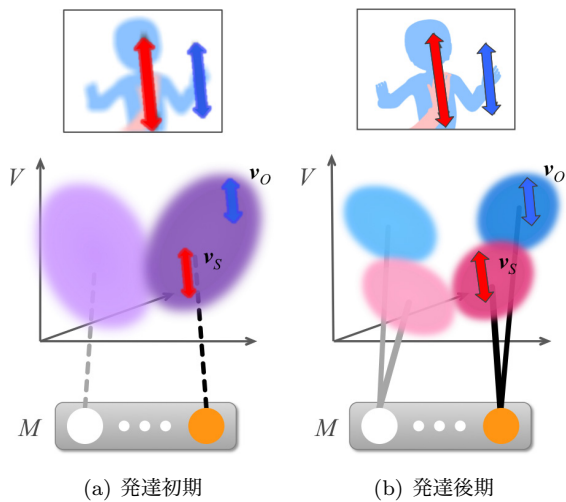


図2 未熟な視覚がもたらす MNS 発達モデル. 発達初期 (a) では運動指令と自他未分化な視覚クラスタの結合を獲得する. 発達後期 (b) では自他の分化したクラスタが得られるが, 対応した運動指令との結合は維持される.

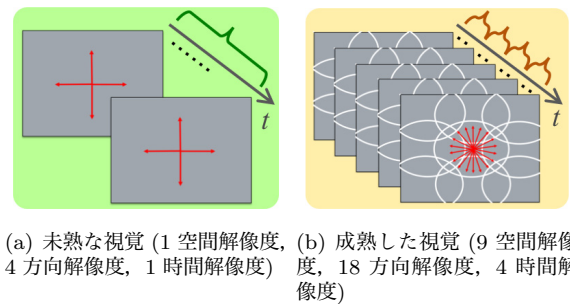
中の紫の領域のような自他が未分化なクラスタを形成する. この状態で運動指令とのヘブ学習を行うと, 自他の区別をすることなく他者運動と自己の運動指令を結び付けることができる. 図2(b)に学習後期の様子を示している. 視覚が発達することで自他運動のクラスタが分化する. 図中で赤い領域が自己, 青い領域が他者運動のクラスタを表している. このモデルの重要な点の一つとして, クラスタが分化するとき, 分裂したクラスタの初期結合荷重は分裂前のもののコピーを用いることとする. 自他未分化期に獲得された結合を保存することによって, 視覚クラスタが自他間で分化したとしても他者運動との対応を保持することができる. したがって, ASL モデルに視覚発達メカニズムを導入することで, MNS に必要とされる自他の区別と対応の両方を学習することができる.

3.2 視覚発達メカニズム

運動観察により得られた時系列のオプティカルフローに3種類の視覚解像度処理を施し, 視覚入力 \mathbf{v} を得る. このとき, 乳児の視覚発達を再現するために, ロボットの視覚の時空間解像度を変化させる. その符号化方法を図3に示す. 今回, 3段階の視覚発達過程を想定し, (a) はその発達初期, (b) は発達後期である.

1) 空間解像度: 一つ目は視覚受容野に関する解像度であり, 図3(b)中の白い楕円で表されている. それぞれの受容野内でオプティカルフローを積算し, オプティカルフローのヒストグラムを作成する. この受容野の増加が視覚発達を表現する. 初めは受容野を1つとし(図3(a)), 視覚情報を1つのヒストグラムとして表す. そして成熟した視覚では, 受容野を9つ(図3(b))設置する.

2) 方向解像度: 二つ目はオプティカルフローの角度を離散化するときの角度の解像度であり, 図3の赤い矢印で表されている. 発達初期, オプティカルフロー



(a) 未熟な視覚 (1 空間解像度, 4 方向解像度, 1 時間解像度), (b) 成熟した視覚 (9 空間解像度, 18 方向解像度, 4 時間解像度)

図3 視覚発達メカニズム. 3段階の発達過程のうち, 最初 (a) と最後 (b) を示している.

は 90° ごとに離散化され, 4つの方向選択性をもつヒストグラムが作成される(図3(a)). そして, 発達後期の方向解像度は 20° ごとの18とする(図3(b)).

3) 時間解像度: 三つ目はオプティカルフローを積算する時間窓の長さに関する時間解像度であり, 図3の}で表されている. ロボットは初め, 観測した運動の期間と同じ長さの時間窓でオプティカルフローを積算する. すなわち時間解像度1である(図3(a)). そして, その解像度を4(図3(b))まで増やしていく. 時間解像度が増加することで, ロボットは運動指令に対する, 他者運動の遅れを検出できるようになる.

今回, 3段階の視覚発達過程を想定し, \mathbf{v} の次元を増やしていく. 1段階目は $1 \times 4 \times 1 = 4$ 次元, 2段階目は $4 \times 8 \times 2 = 64$ 次元, 3段階目は $9 \times 18 \times 4 = 648$ 次元である. この視覚解像度は視覚経験に応じて発達させる. なお, この3つの視覚解像度は乳児の視覚の行動学的, 神経学的研究 [7, 8, 9] に着想している.

図4に視覚情報のコーディング例を示す. (a) は発達初期, (b) は発達後期であり, 左側はロボットが自己の左手の上下運動を観察したとき, 右側はロボットが他者の右手の上下運動を観察したときの視覚情報である. ロボットの視覚が未熟であると, その時空間解像度の低さから, 自他間の差違を検出できない(図4(a)). 一方, 成熟した視覚の場合, 他者のヒストグラムは自己運動に比べ小さくなっており, 他者の時間遅れを検出している(図4(b)). また, 他者運動では手以外の頭などの運動も検出しており, 自他の視点の違いが現れているといえる.

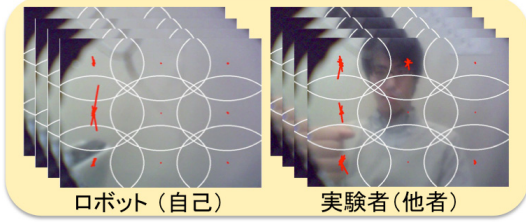
3.3 視覚情報のクラスタリングと連合学習

視覚情報 \mathbf{v} は視覚空間 V でクラスタリングされる. クラスタリング手法にはバタチャリヤ距離を用いた X-means[10]を使用する. X-means は自動的にクラスタ数を決定することができ, このクラスタの増加する過程が, 自他の視覚表象の細分化を表現する.

そして, 側抑制を考慮したヘブ学習により, 視覚クラスタと運動指令の連合を学習する. この学習則は最も反応した視覚クラスタ \mathbf{v}_{fire} だけではなく, その周辺のクラスタもその距離に応じて運動指令との結合を強化するものである. また, 学習を促進するために, \mathbf{v}_{fire} から遠く離れたクラスタの結合は抑制される. 今, $\mathbf{v}_i (i = 1, 2, \dots, N_v)$ と $\mathbf{m}_j (j = 1, 2, \dots, N_m)$ をそれぞれ視覚クラスタと運動指令とする. このとき \mathbf{v}_i と \mathbf{m}_j



(a) 未熟な視覚, 低い時空間解像度のため明らかな自他間の差異は存在しない。



(b) 成熟した視覚, 高い時空間解像度のため自他間の差異が表出している。

図4 視覚情報のコーディング例. 自己運動(左)と他社運動(右)を観測したときのオプティカルフローをコーディングしている. (a)と(b)の発達段階はそれぞれ図3(a), (b)に対応している.

の間の結合荷重 $w_{i,j}$ は以下の式で更新される.

$$w_{i,j}(t+1) = w_{i,j}(t) + \alpha(\mathbf{v}_i) \cdot \beta(\mathbf{m}_j) \quad (1)$$

$\alpha(\mathbf{v}_i)$ と $\beta(\mathbf{m}_j)$ はそれぞれ \mathbf{v}_i と \mathbf{m}_j の反応であり,

$$\alpha(\mathbf{v}_i) = a \exp(-a\pi d_b(\mathbf{v}_{fire}, \mathbf{v}_i)^2) - (a-1) \exp(-(a-1)\pi d_b(\mathbf{v}_{fire}, \mathbf{v}_i)^2) \quad (2)$$

$$\beta(\mathbf{m}_j) = \begin{cases} 1 & \text{if } \mathbf{m}_j \text{ is executed} \\ 0 & \text{else} \end{cases} \quad (3)$$

で表される. ここで a はガウシアン関数の先鋭度を決定するパラメータであり, $d_b(\mathbf{x}, \mathbf{y})$ は \mathbf{x} と \mathbf{y} の間のバタチャリヤ距離である.

視覚解像度の発達や視覚経験によって, 視覚クラスタは徐々に分化していく. あるクラスタが2つのクラスタに分裂した場合, 分裂後のクラスタと運動指令との結合荷重は分裂前のクラスタのものと等しいとする. 分裂前の結合荷重がコピーされることによって, 未熟な視覚で獲得された自他間の対応が保存される.

4. 実験

4.1 実験設定

図1に示されている乳児型ロボット (M3-neony[11]) を用いて, 提案モデルを評価した. このロボットには全身に22自由度, 頭部に2つのCMOS USBカメラ (640×480pixel) がある. ロボットは右手, 左手, 両手を上下, もしくは左右に振るという6種類の運動レパートリーを持ち, そのうち1つをランダムに選択し, 実行する. 実験者もロボットと同様の運動レパートリー

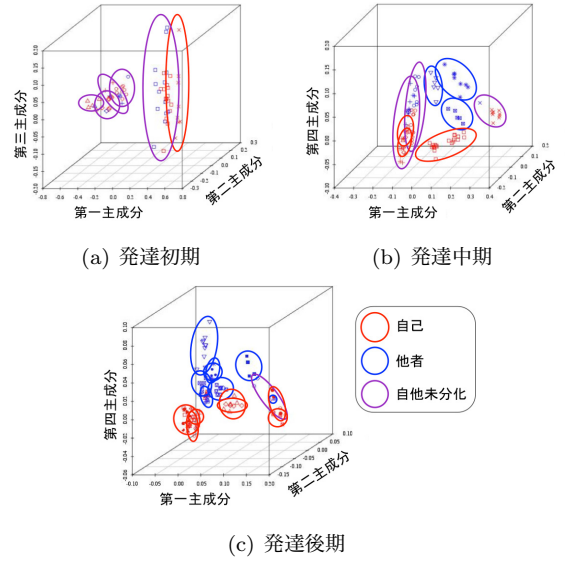


図5 視覚発達にともなう自他の区別. 自他の未分化なクラスタ(紫の楕円)が徐々に自己(赤の楕円)と他者(青の楕円)のクラスタに分裂していく.

を持つ. 実験者はロボットの運動に対して応答運動をするが, 30%の割合でロボットを模倣し, 70%の割合でランダムに運動を選択する. また, 実験者の応答はロボットの運動から2, 3秒の遅れがある.

学習はオフラインで行った. 120種類の視覚情報(自他の運動レパートリーをそれぞれ10回観察したもの)のうち, 1ステップの学習で, 自他の2つの視覚入力に与えられる. 視覚発達は3段階であるが, その入力された視覚情報が70種類と105種類になったときに視覚を発達させた. また, 式(2)のパラメータ a は予備実験により20とし, 学習回数を350回とした.

4.2 視覚発達にともなう自他の分化

図5に視覚情報を主成分分析した結果を示す. (a)は発達初期, (b)は中期, (c)は後期である. 図中の赤い点と青い点は, それぞれ自己運動と他者運動に対応している. 赤, 青, 紫の楕円はそれぞれロボットが獲得した自己, 他者, 自他未分化なクラスタを示している.

図5(a)より, 発達初期では, ほとんどの視覚クラスタが自他未分化状態であることがわかる. 時空間解像度の低い視覚では自他を区別できないためである. そして, 視覚発達にともない, ロボットは自他を分離した視覚クラスタを獲得していく(図5(b)). 最終的にロボットはほとんどの視覚クラスタの自他分離を完了した(図5(c)). 図5(c)においてグラフの上側に他者運動, 下側に自己運動が分布しており, 縦軸が自他を明確に分離していることがわかる. ロボットはこの軸を用いることで, 新奇な運動を観測したときでも自他を区別して認識することができると考えられる.

4.3 未熟な視覚がもたらす自他の対応

自他の対応関係の獲得を確認するために, 図6に学習後の結合荷重マップを示す. (a)は未熟な視覚から発達した提案モデルの学習結果で, (b)は初めから成熟した視覚で学習した場合の結果である. マップの行と列

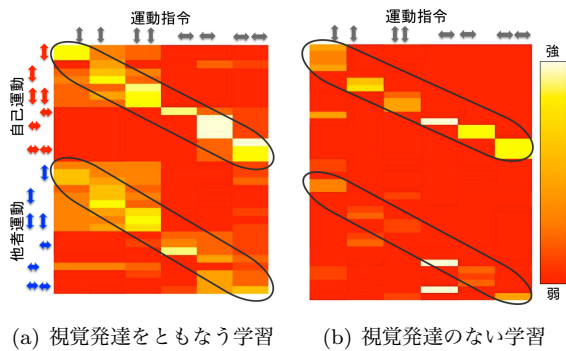


図6 学習により獲得された視覚-運動間結合荷重. 視覚発達のある提案モデル (a) にだけ自他の対応を示す左中央から右下にかけての強い結合が見られる.

はそれぞれ視覚クラスタと運動指令を表しており、マップの上部と左側の矢印は運動レパートリーを表す。例えば上部の最も左にある矢印は右手の上下運動を示している。視覚クラスタは便宜上、上側に自己運動、下側に他者運動となるように配置している。また、マップの色が白くなるほど、その行と列(視覚クラスタと運動指令)の間の結合荷重が強いことを示す。

図6(a)と(b)から、ロボットは視覚発達の有無にかかわらず、自己運動の観察と自己の運動指令の適切な対応関係(左上から右中央の強い結合)を獲得していることがわかる。これは自己運動の視覚クラスタと自己運動指令には強い随伴関係があるためである。一方、他者運動は必ずしも自己運動指令と対応しないため、正しい結合を獲得することは難しい。しかし、視覚発達のある図6(a)において、左中央から右下にかけて強い結合が得られており、自他の対応が獲得されたといえる。それに対して視覚発達のない図6(b)はその関係が見られない。したがってこれらの結果は、視覚発達が自他の対応関係の獲得に寄与していることを示している。

そして獲得された結合荷重を用いて、ロボットは他者運動を模倣することができる。模倣実験の結果、ロボットは時間遅れをとまないながら、他者の連続した2種類の手の運動を、運動間を内挿しながら再現した。

5. 考察と結論

本研究では、視覚の精緻化メカニズムをASLモデルに導入し、MNSの発達をロボティクスの見地から検証した。その結果、ロボットの視覚の低い時空間解像度が自他の運動の等価性を顕在化させ、自他の対応関係の獲得を促進することがわかった。また、視覚発達にともないロボットが徐々に自他を区別した視覚表象を獲得することを示した。したがって、提案モデルによって、ロボットは自他を区別し、その対応関係を認識する能力、すなわちMNSを獲得できたといえる。

一見不必要な乳児の視覚の未熟さが、MNSの発達を促すという仮説は、生物学・神経学的には未だ証明されていないが、我々のキーアイデアと同様の、乳児の未熟さによる制約が高次の認知機能の発達にかかわるといふ仮説は従来研究でも唱えられている。Newport[12]やElman[13]は乳児の言語学習における記憶能力の未

熟さの重要性を主張している。また、Nagaiら[14]は視覚発達が乳児の共同注意の学習を補助することをロボットを用いて示している。

今後の課題は、視覚発達だけではなく、運動発達や他のモダリティを考慮したMNSモデルの構築である。乳児の視覚表象だけでなく運動表象の発達も考慮することでさらにMNSの獲得が促進される可能性がある。さらに、体性感覚、触覚、聴覚といったモダリティの発達も考慮し、互いにその表象の精緻化に影響し合うメカニズムを追加することで、養育者のさまざまな応答形式からマルチモダリティなMNSを獲得するモデルの実現が期待できる。

謝辞 本研究の遂行にあたり、科学研究費補助金(基盤研究(S):課題番号22220002)の補助を受けた。

参考文献

- [1] G. Rizzolatti, et al. ミラーニューロン. 紀伊国屋書店, 2009. (柴田訳, 茂木監修).
- [2] C. Catmur, et al. Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B*, Vol. 364, No. 1528, pp. 2369–2380, 2009.
- [3] Y. Kuniyoshi, et al. From visuo-motor self learning to early imitation—a neural architecture for humanoid learning. In *Proc. of the 2003 IEEE Intl. Conf. on Robotics and Automation*, Vol. 3, pp. 3132–3139, 2003.
- [4] M. Asada, et al. Imitation learning based on visuo-somatic mapping. *Experimental Robotics IX*, pp. 269–278, 2006.
- [5] T. Minato, et al. A model of the emergence of early imitation development based on predictability preference. In *Proc. of the 9th Intl. Conf. on Development and Learning*, pp. 19–25, 2010.
- [6] S. M. Mahler, et al. 乳幼児の心理的誕生. 黎明書房, 1981. (高橋ら訳).
- [7] R. Wilson, et al. Development of spatiotemporal mechanisms in infant vision. *Vision Research*, Vol. 28, No. 5, pp. 611–628, 1988.
- [8] T. Banton, et al. Infant direction discrimination thresholds. *Vision Research*, Vol. 41, No. 8, pp. 1049–1056, 2001.
- [9] K. Hiraki. Detecting contingency: A key to understanding development of self and social cognition. *Japanese Psychological Research*, Vol. 48, No. 3, pp. 204–212, 2006.
- [10] D. Pelleg and A. Moore. X-means: Extending K-means with Efficient Estimation of the Number of Clusters. In *Proc. of the 17th Intl. Conf. on Machine Learning*, p. 727, 2000.
- [11] T. Minato, et al. A baby robot platform for cognitive developmental robotics. In *Proc. of the 2009 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2009.
- [12] Elissa L. Newport. Maturational Constraints on Language Learning. *Cognitive Science*, Vol. 14, No. 1, pp. 11–28, 1990.
- [13] Jeffrey L. Elman. Learning and development in neural networks: the importance of starting small. *Cognition*, Vol. 48, No. 1, pp. 71–99, 1993.
- [14] Y. Nagai, et al. Learning for joint attention helped by functional development. *Advanced Robotics*, Vol. 20, No. 10, pp. 1165–1181, 2006.