

下側頭葉ニューロンの想起活動に関する計算論的考察

末光 厚夫* 諸上 茂光** 森田 昌彦***

*筑波大学大学院工学研究科 **筑波大学大学院システム情報工学研究科 ***筑波大学機能工学系

Computational examination on the recall activity of inferior temporal neurons

Atsuo SUEMITSU* Shigemitsu MOROKAMI** Masahiko MORITA***

*Doctoral Program in Engineering, University of Tsukuba

**Doctoral Program in System Information Engineering, University of Tsukuba

***Institute of Engineering Mechanics and Systems, University of Tsukuba

Abstract: In the inferior temporal cortex of the monkey, 'pair-recall' neurons which exhibit prospective activity for the target during the delay period of a pair-association task have been found. To explain this and other physiological findings, we previously constructed a model of pair-association memory consisting of two interactive networks. The present paper reports that recent empirical data on the time course of the pair-recall activity accord very well with the prediction of our model. This strongly suggests that memories are embedded in trajectory attractors in area TE, implying that the learning signal necessary for forming them is sent backward from the perirhinal cortex.

1. はじめに

脳の視覚系において、刺激の形状に関する情報は後頭葉の V1 や V2 から V4, TEO 野を経て下側頭葉の TE 野へ送られる (Fig. 1)。さらにそこから側頭葉内側部の嗅周皮質、嗅内皮質、海馬体へと信号が伝達されるが、図形の認識は主に TE 野で行われると考えられている。同時に、TE 野およびその内側部は、視覚性の記憶に深く関与していることが知られている。

Sakai と Miyashita¹⁾ は、遅延対連合課題を実行中のサルの下側頭葉皮質を調べ、記憶のメカニズムを探る上で非常に重要な知見を得た。この実験は具体的に次のようにして行なわれた。彼らは、コンピュータで生成した 24 個の図形を適当に組み合わせて 12 組の図形対を作成し、これを用いてサルに遅延対連合課題を行わせた。これは、図形対の一方を cue として短時間提示し、数秒の遅延期間の後に提示した test 図形が cue の対図形 (target) であるかどうかを判断させるというものである。課題の正答率が十分高くなるまで訓練した段階で、微小電極を用いて IT 野のニューロン活動が測定された。

この実験の主な結果は、以下のようにまとめられる。

(1) 図形に対する反応選択性のあるニューロンの多くは、複数の全く異なる図形に反応する。ただし、

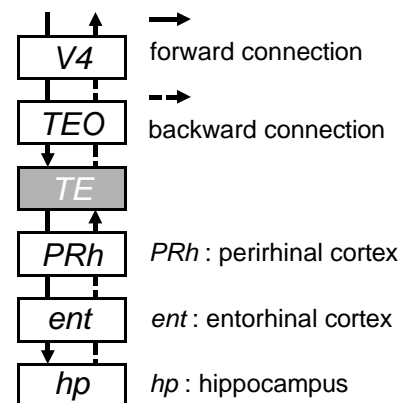


Fig. 1 Forward and backward pathways of visual signals in the temporal lobe.

強い反応を引き起こす図形はその中の少数である。

(2) 対となる二つの図形への反応の相関が全体的に高く、図形対の両方に反応するニューロン (pair-coding ニューロン) が比較的高い比率で観察される。これらの多くは、cue と target に対して強く反応するだけでなく、遅延期間中も持続的に活動する。

(3) 最も強い反応を引き起こす図形 (最適刺激) の対図形を cue として提示したとき、cue に対してはほとんど反応を示さないのに、遅延期間中に徐々に活動を高めていくニューロン (pair-recall ニューロン) が見られる。これらは、逆に最適刺激を cue とし

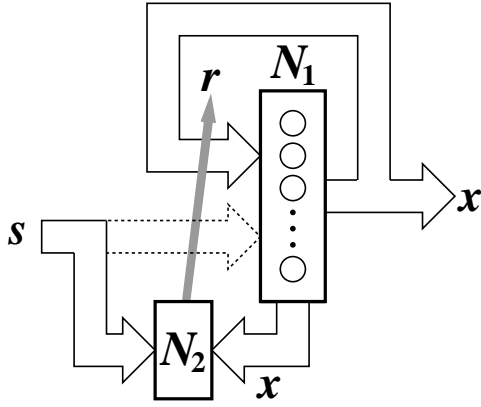


Fig. 2 Block diagram of a model of pair-association memory.

て提示すると、遅延期間中に活動が低下していく。

このうち、(3)の pair-recall ニューロンの活動は、点アトラクタを基礎とする従来の記憶モデル²⁾では説明が困難であり、そのメカニズムは不明であった。これに対して、著者ら^{3, 4)}は計算論的な考察により、軌道アトラクタ⁵⁾を基にした対連合記憶の神経回路モデルを構築した。このモデルは、生物学的制約条件に反することなく対連合課題の学習及び実行ができるだけでなく、対連合記憶関連ニューロンの活動をよく再現する。また、種々の生理学知見を説明すると同時に、TE野のニューロン活動などに関していくつかの現象を予言する。但し、これらの予言は、文献¹⁾の実験データでは十分に検証できなかった。

最近、Nayaら⁶⁾は pair-recall ニューロンの活動をより詳細に測定し、その時間特性を解析することによって非常に興味深いデータを得た。本論文では、上述のモデルに対して同様の解析を行い、モデルの予言を検証するとともに、下側頭葉における対連合記憶の形成及び想起のメカニズムについて考察する。

2. 対連合記憶のモデル

ここでは、著者らが構成した対連合記憶のモデル^{3, 4)}の概略を述べる。詳細な説明については文献³⁾を参照されたい。

2.1 構造とダイナミクス

このモデルは、Fig. 2に示すように、記憶の形成と想起が行われる回路網 N_1 と記憶形成に必要な学習信号を生成する回路網 N_2 とが相互に結合した構造をしている。 N_1 の出力 $x = (x_1, \dots, x_n)$ は

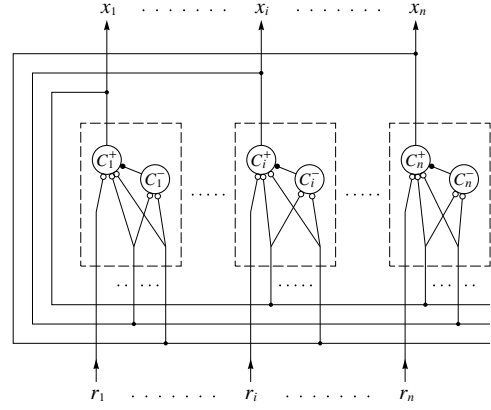


Fig. 3 Structure of the network N_1 .

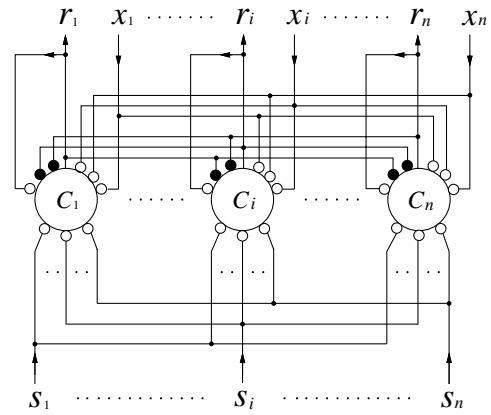


Fig. 4 Structure of the network N_2 .

N_1 に回帰的に入力されるとともに N_2 に送られ、 N_2 の出力 $r = (r_1, \dots, r_n)$ は学習信号として N_1 にフィードバックされる。外部からの入力パターン $s = (s_1, \dots, s_m)$ は本来 N_1 を経て N_2 に入力されるべきものであるが、このモデルでは簡略化のため N_1 への入力は省略し、直接 N_2 に入力されるようになっている。

各回路網はそれぞれ Fig. 3 及び Fig. 4 に示す構造をもち、以下の式に従って動作する。

$$y_i = f \left(\sum_{j=1}^n w_{ij}^- x_j - \theta \right) \quad (1)$$

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij}^+ x_j - w_i^* y_i + \lambda r_i \quad (2)$$

$$x_i = f(u_i) \quad (3)$$

$$\tau \frac{dv_i}{dt} = -v_i + \sum_{j=1}^n p_{ij} s_j + \sum_{j=1}^n q_{ij} x_j - \rho \sum_{j \neq i} r_j + \sigma r_i + h \quad (4)$$

$$r_i = f(v_i) \quad (5)$$

$$f(u) = \frac{1}{1 + e^{-cu}} \quad (6)$$

ここで、 w_{ij}^+ と w_{ij}^- は j 番目のユニットから i 番目のユニットの興奮性細胞 C_i^+ 及び抑制性細胞 C_i^- へのシナプス荷重、 w_i^* は C_i^- から C_i^+ への抑制性シナプスの効率、 u_i は C_i^+ の平均膜電位、 λ は r の入力強度を表す。また、 v_i は N_2 の i 番目の細胞 C_i の平均膜電位、 h はそのオフセットを表し、 p_{ij} と q_{ij} は C_i への入力シナプスの荷重であってランダムな値をとる。 $\theta, \tau, \rho, \sigma, c$ はいずれも正の定数である。

2.2 学習

学習は、各細胞が式 (1) ~ (6) に従って動作すると並行して、以下の式に従って N_1 内部のシナプス荷重を修正することによって行われる。

$$\tau' \frac{dw_{ij}^+}{dt} = -w_{ij}^+ + \alpha r_i x_j \quad (7)$$

$$\tau' \frac{dw_{ij}^-}{dt} = -w_{ij}^- - \beta_1 r_i x_j + \beta_2 x_i x_j + \gamma \quad (8)$$

ここで、 τ' は τ に比べて十分に大きい時定数、 α, β_1, β_2 は正の学習係数で $\beta_1 < \beta_2$ を満たす。 γ は正の定数で、ユニット間の一様な側抑制を表す。学習係数 α は x_i の減少関数とした方が学習性能が高いため、

$$\alpha = \begin{cases} \alpha'(\kappa - x_i) & (x_i < \kappa) \\ 0 & (x_i \geq \kappa) \end{cases} \quad (9)$$

としている ($\kappa \equiv \beta_1/\beta_2$, α' は正定数)。

このような学習を繰り返すことによって、 N_1 の状態空間において学習信号 r に「近い」(ベクトルの方向だけ考える) 点がアトラクタとなるが、 r を連続的に変化させることによって、 r の軌道にほぼ沿って連続したアトラクタが形成される⁵⁾。また、 r が x の変化を先導することによって、 x から r の向き、すなわち r の進行方向の緩やかな流れが生じる。このような流れをもった線状のアトラクタを軌道アトラクタと呼ぶ。

具体的にパターン間の対連合を学習する際には、これらのパターンを順次モデルに入力する。例えば、パターン A を入力する ($s = A$) と N_2 はそれを変換したパターン a を出力する ($r = a$) が、次いで B を入力すると、 N_1 と N_2 の相互作用によって N_2 の出力 r は Fig. 5 に模式的に示すような軌跡を描

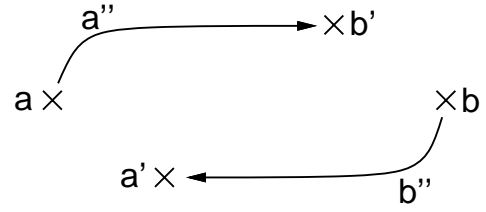


Fig. 5 Paths of the learning signal.

き、B のコードである b に近いパターン b' まで連続的に変化する。同様に、B, A の順に入力すると、 b から a' に向かう時空間パターンが N_2 から N_1 に送られる。このような r の軌跡に沿って軌道アトラクタが形成される結果、学習後のモデルに A を入力したとき N_1 の状態は a から b' まで自動的に遷移するし、B を入力すると a' に達することになる。

2.3 生理学的知見との対応と予言

学習後のモデルは、遅延対連合課題を正しく実行することができ、また下側頭葉の pair-recall ニューロンの活動をよく再現するが、それ以外にもこのモデルには重要な特徴がある。

まず、 N_1 と N_2 がそれぞれ TE 野と嗅周皮質に相当すると考えると、Fig. 1 のような解剖学的構造だけでなく、破壊実験の結果とも整合する。例えば、嗅周皮質を破壊したサルは、新たな刺激セットを用いた対連合学習は全くできないが、破壊前に学習した刺激セットに関しては、再学習すれば対連合課題を実行可能である⁷⁾。このことは、嗅周皮質は対連合学習に必要なものであるが、学習後の対連合課題の実行には不可欠ではないことを示唆しており、このモデルによってうまく説明される。

また、対連合課題における TE 野ニューロンの活動に関して、次のような予言が導かれる。

(1) 遅延期間中の活動変化のタイミングは、ニューロン毎に異なっているはずである。このことは、回路網 N_1 の状態が跳躍的に変化するのではなく、軌道アトラクタに沿って連続的に遷移することから帰結される。

(2) ある図形 A を cue として提示したときのニューロン群の活動パターンと、その対図形 B を cue 提示した後に想起される活動パターンとは少なからず異なる、すなわち遅延期間中に想起されるのは target の完全なコードではないはずである。これは、Fig. 5 において a と a' とが離れているからであ

るが、このことは反対方向に向かう2本の軌道アトラクタを相互干渉を抑えて形成するためには避けられない。

以下では、Naya らの実験データを基にこれらの予言を検証する。

3. 想起の時間特性の解析

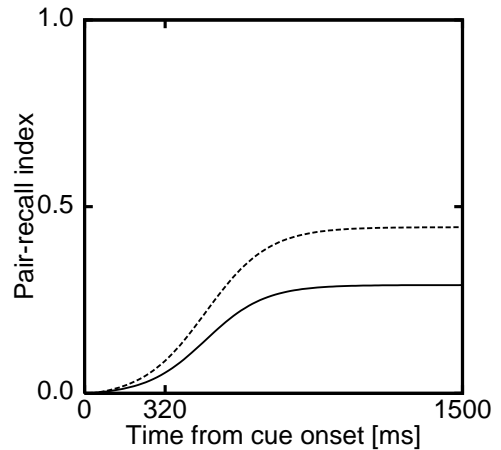
まず、Naya ら⁶⁾の解析方法と結果の概略を述べる。彼らは、Sakai と Miyashita¹⁾と同じ遅延対連合課題についてサルを十分に訓練した後、課題実行時のTE野及び36野(嗅周皮質の一部でTE野に隣接する領域)のニューロン活動を計測した。そして遅延期間中に選択的な活動を示したニューロンについて、それぞれ以下の式に従ってPRI (pair-recall index) と呼ぶ指標を計算した。

$$PRI(t) = \frac{\langle C_p | F(t) \rangle - \langle C | C_p \rangle \langle C | F(t) \rangle}{\sqrt{(1 - \langle C | F(t) \rangle^2)(1 - \langle C | C_p \rangle^2)}}$$

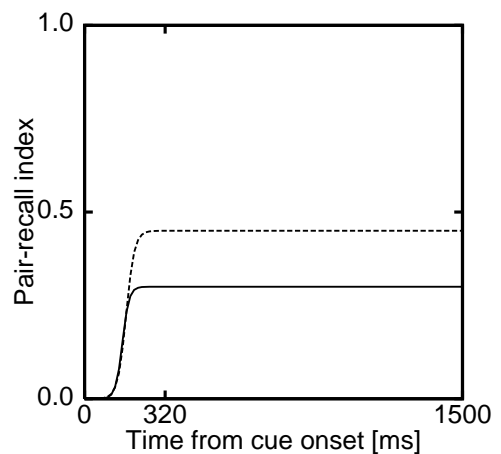
ここで $F(t) = [f_1(t), \dots, f_l(t)]$, $C = [c_1, \dots, c_l]$, $C_p = [c_{p(1)}, \dots, c_{p(l)}]$ であり、 l は学習させた図形の数(この実験では $l = 24$)、 $f_k(t)$ は k 番目の図形を cue としたときの時刻 t における活動(発火頻度)、 c_k は k 番目の図形を cue 提示したときの活動、 $c_{p(k)}$ は k 番目の図形の対図形を cue 提示したときの活動、 $\langle X | Y \rangle$ はベクトル X と Y の相関係数をそれぞれ表す。この指標は、時刻 t における各ニューロンの反応が、target に対する反応にどの程度近づいているかを表すものであり、最大で1、cue 期間中は $PRI(t) \simeq 0$ となるような正規化がなされている。また、cue 期間中と遅延期間中の活動レベルの違いによる影響は受けない。

PRIを計算したTE野ニューロンについて $PRI(t)$ を平均し、curve fittingを行った結果が Fig. 6(a) のグラフである。実線はすべてのニューロンの平均値、破線はPRIが有意(5%水準)に変化したニューロンに関する平均を表す。同様に、Fig. 6(b) は36野のニューロンに関する平均PRIを表したものである。ここで注目すべき点は、TE野における平均PRIが500ms以上にわたって徐々に増加していることである。また、36野の方がTE野よりもPRIの増加が早いというのも新奇かつ重要な結果である(Fig. 1からわかるように、cueに対する反応はTE野の方が早い)が、これについては4.で考察する。

次いで彼らは、各ニューロンについてTRT (transition time) と呼ぶ指標を求めた。これは $PRI(t)$ が



(a) Area TE



(b) Area 36

Fig. 6 Time courses of the averaged PRI for population of pair-recall neurons in the inferior temporal cortex (adapted from Naya et al.⁶⁾).

最大値の半分まで増加するまでの時間である。Fig. 7は、TRTの累積度数分布を示したもので、実線がTE野、破線が36野を表す。TE野の累積度数は少しずつゆっくりと増加しているが、このことはTE野ニューロンのPRIが増加するタイミングがかなり一様に幅広く分布していることを意味する。

さて、以上と同じ解析を、モデルの N_1 のユニットに対して適用した。具体的には、ユニット数 $n = 1000$ のモデルにランダムに作成した20組のパターンを入力し学習させた後、対連合課題を実行させ、ランダムに選んだ100個のユニットについて $PRI(t)$ を計算した。シミュレーションの方法やパラメータは、学習時の遅延時間を 3τ にしたことを除き文献³⁾と同じである。

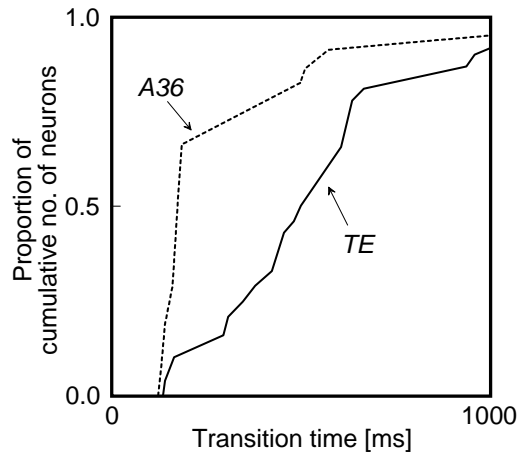


Fig. 7 Distribution of the TRT values for inferotemporal neurons (adapted from Naya et al.⁶⁾).

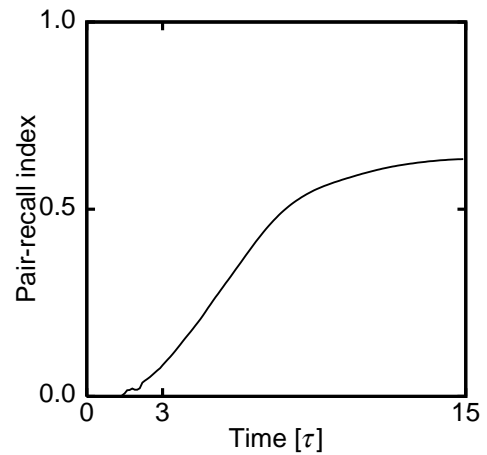


Fig. 8 Time course of the averaged PRI for the units of N_1 of the model.

Fig. 6 と同様に平均 PRI の時間変化をプロットした結果が Fig. 8 である。但し、横軸の単位は時定数 τ であり、curve fitting は行っていない。また、TRT の累積度数分布は Fig. 9 のようになった。

4. 考察

Fig. 6(a) と Fig. 8 を比較すると、TE 野とモデルの平均 PRI のグラフはよく似た形をしているが、前者の値は後者の半分程度であることがわかる。両者の値が異なる原因として、生理データには target の想起にほとんど寄与しないニューロンが含まれている可能性があるのに対し、モデルではそのようなユニットは全くないことが挙げられる。実際、Fig. 6(a) で $PRI(t)$ の増加が有意ではないニューロンを除いたグラフ (破線) は、定量的にも Fig. 8 にかなり近い。

一方、このような操作を行った場合でも、TE 野における遅延期間中の平均 PRI が 0.5 程度までしか増加しないということは、target をコードする活動パターンとその対図形を cue 提示したときに想起されるパターンとがかなり食い違うことを示している。このような食い違いは、test 図形が target に類似していると誤認識を起こしやすいという知見⁸⁾からも示唆されていたが、Naya らのデータはこれをより直接的に示すものであり、2.3 で述べた予言 (2) を実証するものと言えよう。

次に、Fig. 7 と Fig. 9 を比べると明らかなように、TRT の分布に関してモデルは TE 野のデータ

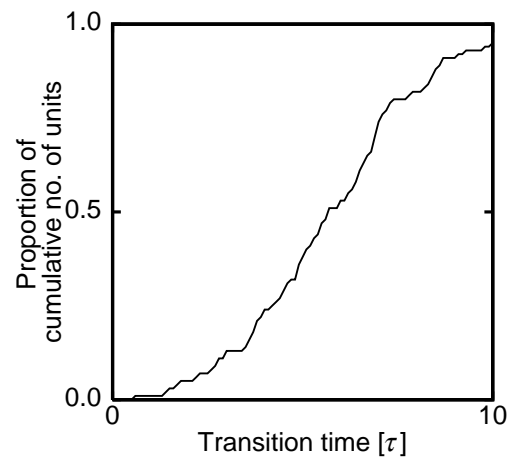


Fig. 9 Distribution of the TRT values in the simulation.

と非常によく一致している。実験条件を少し変えてシミュレーションをしても同様な分布となることから、これは単なる偶然ではなく、モデルの動作原理の反映だと考えられる。従って、Fig. 7 の生理データは、単に 2.3 の予言 (1) と合うというだけでなく、TE 野ニューロン群の活動パターンがモデルと同様な経過をたどって連続的に変化すること、さらにはその変化がモデルと同じメカニズム、すなわち軌道アトラクタによって生じていることを示唆する。

TE 野に軌道アトラクタが形成されているという主張には、もう一つ重要な根拠がある。それは、Fig. 6(b) や Fig. 7 の破線で示されているように、嗅周皮質 36 野の方が TE 野よりも先に target を予測する活動状態へシフトするということである。前述の

ように、軌道アトラクタの形成には回路網の状態遷移を先導する学習信号が必要であり、そのような信号が N_2 から N_1 に送られるというのがモデルの骨子の一つである。 N_2 が嗅周皮質に対応することを考慮すると、36野の状態遷移がTE野に先行するという事実は、嗅周皮質からのフィードバック信号を用いてTE野に軌道アトラクタが形成されるという仮説を支持する。

ここで、36野の活動によってTE野の状態遷移が引き起こされるのならば、TE野に軌道アトラクタがなくてもよいのではないかと疑問が浮かぶ。確かに、サルでは学習後もTE野の状態遷移が嗅周皮質にもある程度依存している(モデルで言うならば、学習信号の入力強度 λ が0でない)可能性が高い。しかし、対連合記憶は嗅周皮質にのみ形成されるという考えは、前述の破壊実験に関する知見と合わない。また、もしTE野には各図形をコードする点アトラクタしかなく、従ってtargetの想起は他の領域からの信号に完全に依存しているならば、一般に別のアトラクタへ遷移する際の過渡状態は不安定であり短時間で終わるから、TRTがFig. 7のように広く分布することはないはずである。従って、TE野に軌道アトラクタが形成されているという解釈が最も妥当だと考えられる。

5. おわりに

以上のように、TE野のpair-recallニューロンに関する生理データは、対連合記憶の計算論的モデルの予測と非常によく一致した。このことは、TE野に軌道アトラクタが形成されていること、及びそれに必要な学習信号が嗅周皮質から送られていることを強く示唆する。

但し、本研究で扱ったモデルは、主にTE野のニューロン活動を説明する目的で構築されたものであり、嗅周皮質のニューロン活動を説明するモデルとしては不十分である。実は、学習時の出力を用いて N_2 の細胞のPRIやTRTを求めると、Fig. 6(b)やFig. 7の点線と比較的よく似た形のグラフが得られる。しかし、現在のモデルの N_2 はtargetを想起するメカニズムを全くもっていないから、PRIが増加するのは N_2 にtargetを入力した後である。これに対して、36野ではcue提示の直後からPRIが増え始めているから、 N_2 と36野のグラフは横軸の原点が大きく異なる。

これに関連して、EricksonとDesimone⁹⁾は、嗅

周皮質ニューロンの遅延期間中の活動が対連合学習の初期段階ではcueを反映しているが、学習が進むとtargetを予測する成分を多く含むようになることを示している。また、Tokuyamaら¹⁰⁾はBDNF(脳由来神経栄養因子)遺伝子の発現を調べ、対連合学習の際に嗅周皮質のシナプス可塑性が高まることを示す結果を得た。これらの知見から、TE野での記憶形成に先行または並行して、嗅周皮質における学習がなされていると考えられる。

従って、36野のニューロン活動を完全に説明するためには、モデルの N_2 に関して構造の修正や可塑性の導入が必要と考えられるし、側頭葉のさらに内側部(嗅内皮質など)に対応する回路網を付加する必要があるかもしれない。これらを考慮したモデルの改良は今後の課題である。

また本研究は、TE野に軌道アトラクタを形成することがサルの課題遂行や脳の情報処理においてどのような意義をもつのか、という疑問に対して解答を与えるものではない。なぜなら、単に対連合課題を実行するだけなら、軌道アトラクタは全く必要ない(例えば、cueからtargetへの相互想起回路とcueまたはtargetを保持する自己想起回路があればよい)からである。これに対して、Nayaら¹¹⁾が行ったPACS(色スイッチ付き対連合)と呼ばれる課題は、軌道アトラクタを用いないと実行困難であると著者らは考えているが、本モデルへの適用は今後の課題である。

謝辞 有益な議論を賜った東京大学医学部宮下保司教授に感謝します。本研究は、文部科学省科学研究費補助金特定領域研究(#12050209)および科学技術振興調整費目的達成型脳研究「文脈主導型、認識・判断・行動機能実現のための動的記憶システムの研究」の補助を受けて行われた。

参考文献

- 1) Sakai, K. and Miyashita, Y. (1991) : Neural organization for the long-term memory or paired association, *Nature*, Vol. 354, pp. 152-155
- 2) Amit, D.J. and Fuji, S. (1997) : Paradigmatic working memory (attractor) cell in IT cortex, *Neural Computation*, Vol. 9, pp. 1071-1092

- 3) 末光厚夫, 森田昌彦 (2001) : 下側頭葉における対連合記憶の神経回路モデル, 電子情報通信学会論文誌 D-II, Vol. J84-D-II, pp. 718–727
- 4) Morita, M. and Suemitsu, A. (2002) : Computational modeling of pair-association memory in inferior temporal cortex, *Cognitive Brain Research*, Vol. 13, pp. 169–178
- 5) M. Morita (1996) : Memory and learning of sequential patterns by nonmonotone neural networks, *Neural Networks*, Vol. 9, pp. 1477–1489
- 6) Naya, Y., Yoshida, M. and Miyashita, Y. (2001) : Backward spreading of memory-retrieval signal in the primate temporal cortex, *Science*, Vol. 291, pp. 661–664
- 7) Murray, E.A., Gaffan, D. and Mishkin, M. (1993) : Neural substrates of visual stimulus-stimulus association in rhesus monkeys, *J. Neuroscience*, Vol. 13, pp. 4549–4561
- 8) Rainer, G., Rao, S.C. and Miller E.K. (1999) : Prospective coding for objects in primate prefrontal cortex, *J. Neuroscience*, Vol. 19, pp. 5493–5505
- 9) Erickson, C.A. and Desimone, R. (1999) : Responses of macaque perirhinal neurons during and after visual stimulus association learning, *J. Neuroscience*, Vol. 19, No. 23, pp. 10404–10416
- 10) Tokuyama, W., Okuno, H., Hashimoto, T., Li, Y.X. and Miyashita, Y. (2000) : BDNF upregulation during declarative memory formation in monkey inferior temporal cortex, *Nature Neuroscience*, Vol. 3, pp. 1134–1142
- 11) Naya, Y., Sakai, K. and Miyashita, Y. (1996) : Activity of primate inferotemporal neurons related to a sought target in pair-association task, *Proc. Natl. Acad. Sci. USA*, Vol. 93, pp. 2664–2669