

カテゴリー知識の獲得による視覚パターンの識別効率の向上

三崎 将也・乾 敏郎

We investigated the effect of category learning in visual pattern recognition by psychological experiment and neural network model simulation. In a psychological experiment, we used tachistoscopic presentation task, and we found that pattern discriminability was enhanced by category knowledge. We constructed a neural network model with three layers and reciprocal connections. We used Wake-Sleep algorithm for network learning and the network made its internal representation by the interaction of bottom-up and top-down processes. The network model can simulate the profit of having category knowledge observed in psychological experiment. Furthermore we considered the computational explanation of the profit of having category knowledge from the viewpoint of MDL (Minimum Description Length). Category knowledge helps the network to construct efficient (shorter description length) representation of patterns. In conclusion, category knowledge has a functional profit not only in visual object identification but also in efficient processing of pattern recognition.

Keywords: category (カテゴリー), neural network (ニューラルネットワーク), Wake-Sleep algorithm (Wake-Sleep アルゴリズム), MDL (最小記述長原理), top-down (トップダウン)

1. はじめに

視覚認識においては、対象の物理的な形態だけでなく、見る側が対象に対して持つ知識も重要な役割を果たしている (Medin, Goldstone, & Gentner, 1993; Humphreys, Riddoch, & Price, 1997; Leopold & Logothetis, 1999). また日常場面においても、よく知っている対象は見慣れない対象に比べてその認識が容易であることなどから、このような対象に対する知識が、視覚処理において機能的に有効な働きを持っていることが考えられる。本研究では視覚認識において対象の知識を持つことが、どのように有効な働きを持つかを心理実験を通じて調べ、その計算論的原理をニューラルネットワークモデルを通じて考察した。

Enhancement of Visual Pattern Discriminability by Category Knowledge, by Masaya Misaki (Brain Function Group, Kansai Advanced Research Center, Communications Research Laboratory) and Toshio Inui (Graduate School of Informatics, Kyoto University).

本研究では特にカテゴリー知識に着目して、カテゴリー知識が対象の認識においてどのように働き、そのような知識がない場合に比べて認識効率にどのような差が見られるかを調べた。

Livingston, Andrews, & Harnad (1998), 三崎・乾 (1999) では、対象に対するカテゴリー知識を獲得することが対象の認識の仕方をどのようにに変化させるかを、類似性評定課題を通じて調べている。そこではカテゴリー知識の獲得により、同一カテゴリー対象間の類似性が高まる結果が得られている。また、このようなカテゴリー知識の効果は、カテゴリー内において典型性の高い事例に対してより強く働くことも示されており (三崎・乾, 2001), カテゴリー知識の視覚認識に対する影響には、カテゴリー内での共通属性をカテゴリー知識からトップダウンに推測し、その検出を助けるという機能的な働きとの関連も示唆されている。

本研究ではカテゴリー知識を獲得することが、実際に対象の認識においてどのように有効に働いて

いるかを調べることを目的とする。カテゴリー知識が認識において機能的な働きを持つならば、カテゴリー知識を持つことで対象の認識において何らかの向上が見られるはずである。三崎・乾 (2001) で示されたように、カテゴリー知識の機能性が対象に存在するカテゴリーに関係する特徴の検出を助けることにあるとすれば、カテゴリー知識が最も有効に働くのは、対象に対する限られた情報しか得られないような場面であると考えられる。そこで本研究では、瞬間呈示課題という課題を用いて、カテゴリー知識を持つことにより認識率がどのように変化するかを調べる実験を行った。瞬間呈示課題とは、対象がはっきりとは確認できないくらい非常に短い時間だけ呈示された後、呈示されたものが何であったかを同定する課題である。このような課題では呈示時間が非常に短いため、対象に対する十分な情報は得られない。しかし、ここでわずかな情報からでも対象のカテゴリーが分かり、カテゴリー知識が機能的に有効な働きをすれば、対象の認識率が向上すると考えられる。実験ではカテゴリー知識の獲得により、実際にこのような認識率の向上が見られるかを調べた。

また、カテゴリー知識の獲得による視覚認識の脳内活動の変化を調べた研究に Reber, Stark, & Squire (1998) の研究がある。そこではドットパターンを用いて、カテゴリー知識を持つ対象を見ている時と、持たない対象を見ている時との脳活動の差を fMRI を用いて調べており、カテゴリー知識を持つ対象を見る方が初期視覚野における活性レベルが低くなることを示している。本研究では、心理実験で見られる認識率の向上や、このようなカテゴリー知識の獲得による活性レベルの低下が、知識からのトップダウンの効果により説明できることを、ニューラルネットワークモデルを用いたシミュレーションを通じて計算論的に考察した。

2. 心理実験の目的

カテゴリー知識を持つことが、瞬間呈示課題における対象の認識に対して、どのような有効性を持つかを調べる。実験では、カテゴリー知識を持つ条件と持たない条件を別々の被験者に対して割り当てる。これは瞬間呈示課題における認識率を調べるといふ課題の特性上、繰り返し課題を行うことによって練習の効果により認識率が向上していくことが考

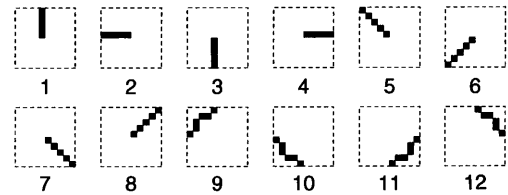


図 1 刺激パターンの要素特徴

えられ、同一の被験者で異なる課題経験の認識率を比較すると、それがカテゴリー学習の効果によるものか、練習の効果によるものかが分からなくなってしまうためである。

またカテゴリー学習は、瞬間呈示課題に用いるものとは別のパターンを用いるが、似たようなパターンを用いるため、認識率の差がカテゴリー学習によるものではなく、単に同じようなパターンを多く見たという視覚経験によるものとも考えられる。実際 Lawrence (1950) において、単に視覚経験を重ねるだけでも認識率の向上が見られることが示されている。そこで実験では、このような視覚経験をコントロールするために、カテゴリー学習を行わない被験者群に対してもカテゴリー学習に用いたものと同じパターンを見る課題を課した。

瞬間呈示課題では、同一カテゴリーに属する対象間の弁別 (Within category 条件)、異なるカテゴリーに属する対象間の弁別 (Between category 条件)、カテゴリーに属する対象と明確なカテゴリー知識が与えられない対象との間の弁別 (Categorical vs. Non-categorical 条件) で、どのような違いがあるかを調べた。

以上のように、本実験では同じ視覚経験のもとで、カテゴリー知識の有無のみが異なる 2 つの被験者群において、瞬間呈示課題における認識率がどのように異なるかを、対象のカテゴリー知識に関する条件ごとに調べた。

3. 方法

3.1 被験者

大学生、大学院生 16 人 (21 歳から 29 歳, 男性 13 人, 女性 3 人) が被験者として実験に参加した。

3.2 刺激図形の作成とカテゴリーの定義

実験では、10×10 のピクセルパターンを刺激パターンとして用いた。パターンは図 1 の 12 個の特



図2 瞬間呈示課題に用いた刺激パターン

徴から5個を選択して組み合わせることで作成した。また、パターンに対するカテゴリ知識が個々のパターンとラベルを対応付けるだけの個別学習になることを避けるため、カテゴリ学習に用いるパターンと瞬間呈示課題に用いるパターンとは別のパターンを用いた。パターン作成の細かな手順に関しては付録Aに記す。

付録Aの手順により、瞬間呈示課題に用いる5個のパターンをカテゴリ1から2つ、カテゴリ2から2つ、どちらにも属さないものから1つ選んだ。図2に瞬間呈示課題に用いた5つのパターンを示す。

また、カテゴリ学習にはこれら5個のパターンとは別の80個のパターン（各カテゴリ40個ずつ）を用いた（作成手順は付録Aを参照）。表1には、各カテゴリで示唆的な特徴がどのようなものになっていたかを、カテゴリ C_i における特徴 f_k の出現確率である $P(f_k|C_i)$ の値により示している。表1から、カテゴリ1では特徴7, 8, 10が、カテゴリ2では特徴5, 6, 11が各カテゴリで高い確率で存在する特徴になっていることがわかる（特徴の番号は図1の番号に対応する）。

3.3 手続き

被験者は瞬間呈示課題の前にカテゴリ学習を行うカテゴリ学習群と、同じパターンを見るがカテゴリ学習は行わない統制群に分けられ、各群の人数は8名ずつであった。カテゴリ学習群では瞬間呈示課題の後に、瞬間呈示課題に用いたパターンがどちらのカテゴリに判断されているかを調べるカテゴリテスト課題を行った。また実験に入る前に、瞬間呈示課題の呈示時間を決めるための練習課題を行っている。

カテゴリ学習群における実験の手順は、練習課題、カテゴリ学習課題、瞬間呈示課題、カテゴリテスト課題の順で行われ、統制群では、練習課題、統制課題、瞬間呈示課題の順で行われた。実験には制御用のパーソナルコンピュータ（DELL

Dimension XPS T600）、画像呈示用拡張ボード（Cambridge Research Systems VSG 2/5）、呈示用モニター（EIZO FlexScan F980）を用いた。以下にそれぞれの課題の手続きについて述べる。

3.3.1 カテゴリ学習課題

カテゴリ学習課題では、画面の中央にパターンが1つずつ呈示され、それが1のグループに属するものか2のグループに属するものかをキー押しにより判断し、その都度正誤のフィードバックが与えられた。80個の図形パターンをランダムな順序で呈示し、これを1セッションとして12セッションの課題を行った。図形パターンの大きさは1.5cm×1.5cmで、約143cmの距離から見ることで、視角にして約0.6度×0.6度の大きさで呈示された。以下の課題においても、呈示サイズは同じである。

3.3.2 統制課題

カテゴリ学習群と統制群との図形パターンに対する視覚経験を等しくするため、統制群に対してはカテゴリ学習とは別の課題を課した。ここでの課題は、連続して呈示されるパターンが1つ前のものと同じであるか違うものであるかを答える課題である。課題にはカテゴリ学習で用いたものと同じパターンを用いた。実験の手順は以下のとおりである。

まず初めに、モニターの中心から少し左にパターンが1つ呈示される。被験者がキーを押すと次のパターンがモニターの中心から少し右に呈示される。被験者の課題は、呈示されたパターンが1つ前のパターンと同じか異なるかを答えることである。同じなら1のキーを、異なる場合は2のキーを押す。被験者がキーを押してから2秒後に、再び次のパターンがモニターの中心から少し左に呈示され、このパターンが1つ前のパターンと同じであるかどうかを答える。これを繰り返して行っていく。

各パターンの呈示頻度はカテゴリ学習課題と同じにしており、全セッションを通じて1パターンあたり12回呈示されるので、全部で960回の課題となる。課題はカテゴリ学習課題と同様の手順になるように、80回ずつの12セッションに分けて行い、セッションの間隔には適宜休憩をはさんだ。各セッションの最後に呈示されるパターンは次のセッションの初めに呈示されるパターンとなる。また同じパ

表 1 カテゴリーの定義 (カテゴリー学習に用いたパターンセットから求めた $P(f_k|C_l)$ の値)

カテゴリー (C_l)	特徴 (f_k)											
	1	2	3	4	5	6	7	8	9	10	11	12
1	.55	.53	.50	.60	.15	.13	.65	.73	.40	.63	.00	.15
2	.55	.60	.50	.50	.73	.63	.10	.15	.15	.05	.68	.38

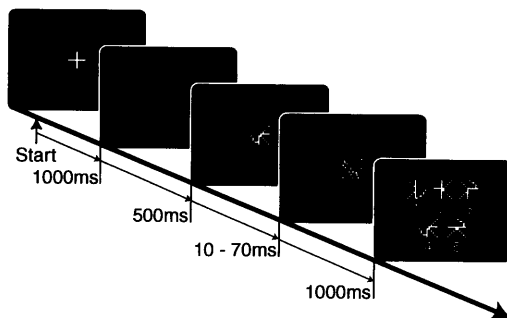


図 3 瞬間呈示課題の手続き

ターンが続けて出る確率と異なるパターンが出る確率は等しくした。

このようにして、統制群にもカテゴリー学習群と同等の視覚経験を与えつつ、パターンに対するカテゴリー知識は持たないようにした。

3.3.3 瞬間呈示課題

瞬間呈示課題は、両群の被験者ともに同じ手続きで行った。ただし被験者の瞬間視の能力を測るため、すべての課題の前に（カテゴリー学習課題または統制課題の前に）練習課題を行っている。練習課題では、20ms から 60ms まで 10ms ごとの 5 段階の呈示時間で 5 つのパターンを 4 回ずつ呈示する 100 回の瞬間呈示課題を行い、そこでの呈示時間ごとの正答率をもとに各被験者の呈示時間を決定した。

瞬間呈示課題の手続きは以下のようにして行った（図 3 参照）。まずモニターの中心に、十字の凝視点が呈示される。被験者がキーを押すと課題が開始し、キーを押してから 1 秒後に凝視点が消失、その 0.5 秒後に図 2 の 5 つのパターンの内の 1 つが瞬間だけ呈示され、その直後に正方形形状のランダムドットのマスクパターンが 1 秒間呈示される。マスクパターンは毎回変わり、その大きさは図形パターンと同じ大きさである。またランダムドットの一つのドットは、図形パターンのドットの 1/3 の大きさである。その後、5 つのパターンが番号付の一

覧となって呈示される。被験者は瞬間だけ呈示されたパターンが、一覧の中のどのパターンであったかを、一覧につけられた番号で答える。パターンの一覧には図 2 の 5 個のパターンのみを用いた。一覧の中でのパターンの位置は毎回ランダムに変化し、番号は左上から順に 1 から 5 の番号が付けられていたため、番号とパターンの対応は毎回変化した。パターンの一覧は、被験者の回答があるまで呈示され続けた。また回答に対する正誤のフィードバックは与えなかった。被験者の回答の後、再び凝視点が呈示され、被験者がキーを押すことで次の課題が始まった。

練習時の課題では、まったく見えなかった場合には 6 を押すように教示したが、実験時には必ず 1 から 5 の中から選ぶように指示した。また瞬間呈示の呈示時間は、練習時の正答率が最初に 30% を超えた呈示時間から 10ms だけ短いものを最短の時間とし、10ms 刻みで 6 段階設定した。実験時には 3 人の被験者（カテゴリー学習群で 1 人、統制群で 2 人）で 20ms から 70ms までの呈示時間を用い、それ以外の全ての被験者で 10ms から 60ms までの呈示時間を用いた。

実験の課題では呈示時間が 6 段階とパターンが 5 種類あり、それぞれの組み合わせを 15 回ずつ繰り返すことで合計 450 回の課題を行った。パターンの呈示順や呈示時間の順はランダムとしたが、同じパターンが続けて呈示されると視認性が上がるので、同じものが 2 回続けて呈示されることは無いようにした。

実験は被験者のペースで進められ、回答の制限時間等は設けなかった。また 90 回の課題ごとに休憩を挟んだ。

3.3.4 カテゴリーテスト課題

カテゴリー学習群に対しては、瞬間呈示課題に用いた 5 つのパターンに関して、そのカテゴリーを答える課題を行った。カテゴリー判断の基準には、

カテゴリー学習課題時に学習したものを用いるように指示した。実験の手続きはカテゴリー学習課題と同様であるが、正誤のフィードバックは与えなかった。ここでは、5 個のパターンを 10 回ずつランダムな順序で呈示する 50 回の課題を行った。

4. 実験の結果

カテゴリー学習群におけるカテゴリー学習の正答率を見ると、1 人を除いて全ての被験者で最終セッションで 92%以上の正答率を達成しており（1 人は最終セッションで 73.75%の正答率であった）、最終セッションにおける平均正答率は 93.91%であった。このことから、ほとんどの被験者でカテゴリー知識がほぼ学習できていると考えられる。またカテゴリーテスト課題の結果を見ると、1 から 4 のカテゴリーに属するパターンに関して、想定したカテゴリーに正しく分類される確率は、全被験者の平均で、パターン 1 が 100%、パターン 2 が 98.75%、パターン 3 が 98.75%、パターン 4 が 96.25%と、ほぼ想定したおりのカテゴリーに分類されていることが分かる。カテゴリー学習課題で正答率が低かった被験者も、カテゴリーテスト課題では想定した通りのカテゴリー判断ができていたので、データの分析に含めた。

またカテゴリー知識を与えないパターンとして設定したパターン 5 は、カテゴリー 1 と判断される確率が 16.25%、カテゴリー 2 と判断される確率が 83.75%で、多くの被験者で一貫してカテゴリー 2 と判断された。

次に瞬間呈示課題の結果を見ていく。まずはデータを、各被験者で呈示時間ごとに混同行列にまとめた。混同行列とは呈示されたパターンに対してどのパターンと答えたかの頻度を、呈示パターン × 回答パターンの行列に表したものである。この混同行列はパターン間の間違いやすさ、逆にいえばパターン間の弁別感度を表すものであり、混同が多いほどそのパターン間の弁別が難しく、少ないほど弁別が容易であったことを表す。混同行列をパターン間の弁別感度に展開するモデルはいくつか提案されているが、ここでは多くの場面においてデータへの適合が優れている Luce (1963) のモデルを用いた (Townsend, 1971a; Smith, 1982)。またモデルの適用の手続きは、Gilmore, Hersh, Caramazza, & Griffin (1979), Townsend (1971b) と同様にし、パ

ターン間の弁別感度として Luce (1963) のモデルにおけるパターン間の距離 d を求めた。

このような手続きで、各被験者で呈示時間ごとに弁別感度を求めるが、ここで被験者により瞬間視の能力に差があるため、同じ呈示時間でも同じ見えやすさとして同等に評価することはできない。練習課題においてある程度は呈示時間を調整したが、それでも同じ条件の被験者間でさえ、呈示時間ごとの正答率には大きな差が見られた。そこで、正答率が 30%以下の呈示時間は、ほとんど見えていない呈示時間と考え、最初に 30%を超えた呈示時間を 1 として、それより 1 段階 (10ms) 長い呈示時間を 2、次を 3 というようにしていき、正答率が 30%に満たない呈示時間は 1 段階ごとに 0, -1, -2 とすることで、最初に 30%を超えた呈示時間を基準とした尺度で各被験者ごとの呈示時間をそろえた。瞬間呈示課題が 5 個の中から 1 つ答える課題であったことから、でたために答えても 20%前後の正答率は達成できるが、30%以上の正答率はある程度見えていないと達成できないと考え、このような基準を定めた。これにより 0 以下の呈示時間は正答率が 30%以下の呈示時間となり、ほとんど見えていない呈示時間であると考えられる。

以上のようにして、被験者ごとに補正した呈示時間ごとの対象間の弁別感度を求め、グラフに表したものを図 4 に示す。図 4 では、Between category 条件での弁別感度（パターン 1 対パターン 3、パターン 1 対パターン 4、パターン 2 対パターン 3、パターン 2 対パターン 4）、Within category 条件での弁別感度（パターン 1 対パターン 2、パターン 3 対パターン 4）、Categorical vs. Non-categorical 条件での弁別感度（パターン 1 対パターン 5、パターン 2 対パターン 5、パターン 3 対パターン 5、パターン 4 対パターン 5）の 3 つに分けて、呈示時間ごとの弁別感度の平均値を条件ごとにグラフにしている。

図 4 から、カテゴリー学習条件（実線のグラフ）では統制条件（破線のグラフ）に比べ、呈示時間の増加による弁別感度の上昇が大きいことが分かる。

次に、これらの弁別感度の差を統計的に分析した。分析には呈示時間が 0 以上のデータのみを用いた。3 要因（カテゴリー学習、呈示時間、対象間の条件（Between category, Within category, Categorical vs. Non-categorical））の分散分析を行っ

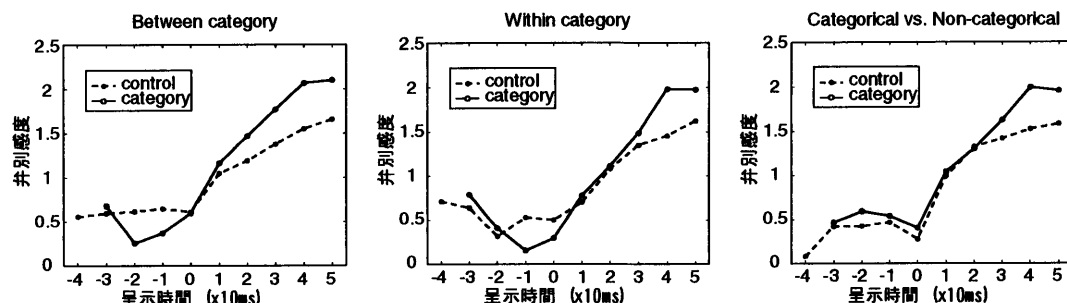


図4 弁別感度の呈示時間による変化

たところ、カテゴリー学習 ($F(1, 774) = 49.100$, $p < .001$), 呈示時間 ($F(5, 774) = 201.354$, $p < .001$), 対象間の条件 ($F(2, 774) = 11.467$, $p < .001$) の3つの要因全てで有意な主効果が見られた。

さらに対象間の条件ごとに、カテゴリー学習の単純主効果を調べたところ、Between category 条件 ($F(1, 774) = 39.532$, $p < .001$), Within category 条件 ($F(1, 774) = 5.584$, $p = .018$), Categorical vs. Non-categorical 条件 ($F(1, 774) = 19.229$, $p < .001$) のいずれにおいても有意な効果が見られた。

ここで瞬間呈示課題において、どちらのカテゴリーにも属さないものとして想定した図2のパターン5が、多くの被験者で一貫してカテゴリー2と判断されたため、Categorical vs. Non-categorical 条件をさらに細かく分析した。ここではパターンのペアごとに、カテゴリー学習の単純主効果を調べた。その結果、パターン1とパターン5 ($F(1, 690) = 9.729$, $p = .002$), パターン2とパターン5 ($F(1, 690) = 11.720$, $p = .001$), パターン3とパターン5 ($F(1, 690) = 4.380$, $p = .037$) で有意なカテゴリー学習の効果が見られたが、パターン4とパターン5 ($F(1, 690) = 0.008$, $p = .930$) ではカテゴリー学習の効果は見られなかった。

5. 心理実験の考察

カテゴリー学習によりパターン間の弁別感度が有意に上昇したことから、カテゴリー学習は対象の識別率を向上させることが分かった。これは統制群においても同じ視覚経験を持つことから、単に同種の対象を多く見たことによるものではないと考えられる。実際、統制課題の実験時間は、カテゴリー学

習課題に比べて少し長くなる傾向があったことから、統制課題において被験者が対象をあまり見えないということは無かったといえる。統制群とカテゴリー学習群との間で異なる点は、対象に対するカテゴリー知識を持つか持たないかにあり、この点が対象の識別率における差を生み出していると考えられる。

今回のカテゴリー学習に用いたパターンセットでは表1に示したように、カテゴリー1では特徴7, 8, 10が、カテゴリー2では特徴5, 6, 11が各カテゴリー内で共通に高い確率で持つ特徴になっている。また表2に示したように、ある特徴を検出したときにそれがどちらのカテゴリーのパターンであるかの確率を表す $P(C_i|f_k)$ の値を見ても、特徴5, 6, 7, 8, 10, 11がカテゴリー間での差異を最もよく表す特徴となっている。このことから、カテゴリー学習を行うことによりこれらの特徴の検出が容易になったことが、弁別感度の上昇につながったことが考えられる。

条件ごとに考察すると、Between category 条件では、これらの特徴の検出感度が上がったことにより、パターン間で異なる部位が検出しやすくなり、弁別感度が上昇したと考えられる。また Within category 条件でもパターン1とパターン2は特徴10により、パターン3とパターン4は特徴11により弁別可能であり、これらの特徴の検出感度の上昇により弁別感度の上昇が説明できる。実際、呈示時間が3以降のパターン間の混同率を調べると、パターン1が呈示されたときにパターン2と混同する確率は統制群で10.5%、カテゴリー学習群で7.1%であったのに対し、パターン2が呈示されたときにパターン1と混同する確率は統制群で14.1%、カテゴリー学習群で12.9%と特徴10を持つパターン1

表2 カテゴリー学習に用いたパターンセットから求めた $P(C_l|f_k)$ の値

カテゴリー (C_l)	特徴 (f_k)											
	1	2	3	4	5	6	7	8	9	10	11	12
1	.50	.47	.50	.55	.17	.17	.87	.83	.73	.93	.00	.29
2	.50	.53	.50	.45	.83	.83	.13	.17	.27	.07	1.0	.71

が呈示されたときのほうがカテゴリー学習の効果が大きい傾向が見られた。パターン3と4に関しても、パターン3が呈示されたときにパターン4と混同する確率は統制群で9.8%、カテゴリー学習群で6.3%であったのに対し、パターン4が呈示されたときにパターン3と混同する確率は統制群で7.1%、カテゴリー学習群で4.3%とわずかであるが特徴11を持つパターン3が呈示されたときのほうがカテゴリー学習の効果が大きかった。

パターン5とパターン1や2との弁別感度が上昇したのは、パターン5がカテゴリー2と判断されていたことから Between category 条件と同様の理由が考えられる。またパターン3とパターン5では特徴5や11の検出感度の上昇が弁別感度の上昇につながったと考えられる。パターン4とパターン5に関しては、パターン4にある特徴5やパターン5にある特徴7を検出すれば、パターン間の弁別が容易になると考えられるが、そのような効果は見られなかった。カテゴリー知識がトップダウンに働いて特徴の検出を助けると考えると、パターン5はどちらのカテゴリーの特徴も持つため、その働きが相対的に弱かったか、または特徴12が他の特徴に比べ強い影響をもっていたためパターン4とパターン5の弁別感度に変化がなかったことなどが考えられる。もともとパターン5はいずれのカテゴリーにも属さないものとして設定したものであるが、多くの被験者でパターン2と判断されてしまい、カテゴリー知識が働かないときの弁別感度の変化を見ることができなかった。これらは特徴が本来持つ顕著性の違いに起因するのではないか。本実験では付録Aに示したように、なるべく特徴間での顕著性の差が出ないように刺激を作成したが、それでも顕著性の違いを完全に制御することはできなかった。より正確な結果を得るためには特徴がもともと持っている顕著性までコントロールする必要があると考えられるが、特徴がどれだけ顕著であるかは他の特徴との組み合わせなどによっても変化し制御するこ

とが困難であった。この点に関しては今後検討する必要がある。

今回の実験で見られたように、カテゴリー知識はカテゴリーに関係する特徴の検出を助けるように働いていると考えられるが、このような働きには、カテゴリー知識がカテゴリー間の差を強調する形で働くことと、カテゴリー内での共通性を強調する形で働くことの二つが考えられる。今回の実験ではカテゴリー間の差を示す特徴と、カテゴリー内での共通性を示す特徴が同じになっていたため、これらを区別することはできなかったが、三崎・乾 (2001) では類似性評定課題において、カテゴリー知識の影響がカテゴリー内での共通特徴を強調する形で働いていることを示唆しており、今回の実験でも同様の機構でカテゴリー知識がトップダウンに働いて、対象のカテゴリーラベルから対象が持つ特徴を推測し、検出しやすくすることで視覚認識において機能的な役割を果たしていることが考えられる。ただし三崎・乾 (2001) では類似性評定課題であったのに対し、今回の実験では瞬間呈示課題を用いており、課題によってカテゴリー知識の働き方が変化することも考えられる。

以上、心理実験の結果からカテゴリー学習を行った被験者群の方が行わなかった被験者群に比べ、パターン間の弁別感度が上昇する傾向が見られ、カテゴリー知識の視覚認識における有効性が確認された。またカテゴリー学習の効果は、対象の持つ特徴の検出を助け、対象の識別を容易にする働きがあると考えられる。実際、カテゴリー知識の獲得による脳における視覚処理の変化を調べた Reber et al. (1998) の研究において、カテゴリー知識を持つ方が初期の視覚処理における活動が弱まることが示されており、これが視覚処理の負荷が小さくなっていることと関係しているとも考えられる。そこで次に、このようなカテゴリー知識の有効性が脳におけるどのような働きにより実現されているのかを、ニューラルネットワークモデルを通じて考察し、シ

ミュレーションを通じてカテゴリー知識の有効性を計算論的に考察する。

6. カテゴリー知識の獲得による脳活動の変化

Reber et al. (1998) はカテゴリー知識の獲得により、脳の活動がどのように変化するかを fMRI を用いて調べている。そこではドットパターンに対するカテゴリー学習を行い、その後学習したものとは異なる新たなテストパターンのカテゴリー判断を行う課題を被験者に課して、その課題を行っているときの脳活動を fMRI により測定している。その結果、カテゴリーに属するパターンを見ているときは、カテゴリーに属さないパターンを見ているときに比べ、初期の視覚情報処理に関係する部位である後頭葉 (BA 17/18) での活動の減少が見られた。この実験で用いた刺激はいずれもカテゴリー学習時には見ていないパターンであるため、これらの活動の減少が単に繰り返し呈示されたことによるものであるとは考えられず、カテゴリー知識の獲得が視覚情報処理に対して何らかの影響を与えていることが考えられる。しかし、事前に対象に対するカテゴリー情報をもつことが、どのような形で視覚情報処理に影響を与えているかは分かっていない。本研究のシミュレーションでは心理実験の結果の再現とともに、このようなトップダウンの情報による活動の減少に着目し、ネットワークシミュレーションを通じて、カテゴリー知識を持つことの認識における有効性と脳活動の変化との関係に関する計算論的な考察を行った。

7. ニューラルネットワークモデルの構成

7.1 ネットワークの構造

脳においては多くの部位が相互に結合を持っており、低次の感覚入力領域からボトムアップに高次の領域へ信号を送るだけでなく、高次の領域から低次の領域へ向けての結合も存在し、ボトムアップとトップダウンの相互の結合によって処理が行われていると考えられる。実際、Buckner, Koutstaal, Schacter, & Rosen (2000) では知識からの影響がトップダウンに働いていることが示唆されており、また高次の領域からのトップダウンのカテゴリー情報が存在することも Tomita, Ohbayashi, Nakahara, Hasegawa, & Miyashita (1999) において示

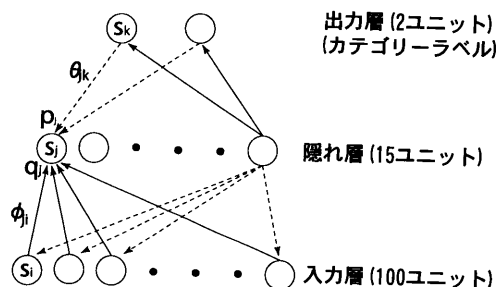


図5 ネットワークの構造

されている。本モデルにおいては、カテゴリー情報はトップダウンに働き、ボトムアップの視覚入力とトップダウンの影響の相互の影響過程により内部表現が形成されるネットワークモデルを構築した。

このような働きを実現するために、ネットワークモデルには相互の方向に結合を持つ階層的な構造を用いた。図5にネットワークの構造を示す。ネットワークは3層の階層構造を持ち、各層のユニットは上下の層と双方向の全結合を持ち、同一の層での結合は持たない。また、ボトムアップの結合とトップダウンの結合はそれぞれ異なる値を持つことができる。入力層には視覚入力に対応する画像パターンが入力され、カテゴリー学習条件では最上位の第3層にカテゴリーラベルが与えられる。

ネットワークの動作の目標は、隠れ層の内部表現から入力層の画像パターンを再生できるような内部表現を獲得することと、カテゴリー学習条件ではそれに加えて、入力パターンから正しいカテゴリーラベルを出力することである。

7.2 Wake-Sleep アルゴリズム

上記のような動作目標を達成するため、ネットワークの学習には“Wake-Sleep”アルゴリズム (Hinton, Dayan, Frey, & Neal, 1995) を用いた。Wake-Sleep アルゴリズムはもともと、相互結合を持つ階層的ネットワークの教師なし学習アルゴリズムとして考えられたものであり、入力パターンから自己組織化により、ネットワークに内部モデルを構成するための学習アルゴリズムである。

自己組織化のネットワークにおいては、どのような評価基準で内部表現を形成するかが問題となる。Dayan, Hinton, Neal, & Zemel (1995) によれば、Wake-Sleep アルゴリズムは、パターンが入力されたときに、ネットワークのボトムアップとトップダ

ウンの動作により再び入力パターンが再構成される尤度を最大化するような学習アルゴリズムになっている。このような尤度の対数の負値をネットワークのエネルギー関数と考えたときに、Wake-Sleep アルゴリズムはネットワークのヘルムホルツ自由エネルギーを結合荷重に関して最小化する学習アルゴリズムとなっていることから、Wake-Sleep アルゴリズムを用いたニューラルネットワークはヘルムホルツマシンとも呼ばれる (Dayan et al., 1995)。

また Frey & Hinton (1997), Hinton et al. (1995) では、Wake-Sleep アルゴリズムは、できるだけ簡潔でありながら、入力パターンをできるだけ正確に再構成できるような内部表現 (圧縮表現) を獲得するためのアルゴリズムであることを MDL 原理 (Rissanen, 1989) の観点から説明しており、Wake-Sleep アルゴリズムが入力パターンの再構成が可能な限りにおいて、データの記述長が最小となるようなネットワークの内部表現を近似的に獲得することが示されている。MDL 原理に関しては山西・韓 (1992), 山西 (1992) に詳しい説明がある。ただし、モデル選択の原理としての MDL 原理では、データの記述長だけでなく、モデルそのものの記述長も問題とすることで、より簡潔なモデルが選ばれるような基準となっているが、Wake-Sleep アルゴリズムは、モデルの複雑性に関してはなんら関与しておらず、それはネットワークの構造として決められたものとしていることを注意しておく。

Wake-Sleep アルゴリズムによる、ネットワークの細かな動作手順に関しては付録 B に記した。

8. シミュレーション

先の瞬間呈示課題と同様の課題を、上記のようなニューラルネットワークモデルに対して行った。シミュレーションではカテゴリー学習の過程は考慮せず、その後の瞬間呈示課題において見られたカテゴリー知識を持つことの優位性を再現できるかを調べた。

8.1 手続き

8.1.1 学習の手続き

ネットワークへの入力心理実験で用いた 10×10 のピクセルパターンを図の部分 1 、地の部分を 0 として、 100 次元の $\{0, 1\}$ のベクトルパターンにして入力パターンとした。よって入力ユニットの数は

100 個である。また中間層ユニットは 15 個とし、出力層のユニットは各カテゴリーに対応するものとして 2 個を設けた。

ネットワークの学習の手順には付録 B に記したものをを用いたが、ここで心理実験と同様に、学習の手順にカテゴリー学習条件と統制条件を設けた。

カテゴリー学習条件では、まず心理実験のカテゴリー学習に用いた 80 個のパターンを Wake-Sleep アルゴリズムにより学習していく。ただし Wake フェーズにおいて出力層のユニットの活性値は、入力パターンが属するカテゴリーに応じて $\{1, 0\}$ または $\{0, 1\}$ の活性値に固定する。また Sleep フェーズでは、出力層における始めのトップダウンの活性は、 $\{1, 0\}$ か $\{0, 1\}$ が等確率で出現するようにした。Dayan & Hinton (1996) において、このような教師ありの Wake-Sleep アルゴリズムに関して述べられており、カテゴリーラベルの出力を教師として与えることにより、ボトムアップのモデルにおいて入力パターンの特徴 f が与えられたときにカテゴリー C を出力する条件付確率 $P(C|f)$ を学習し、トップダウンのモデルにおいてカテゴリー C が与えられたときに特徴 f を出力する条件付確率 $P(f|C)$ を学習できることが示されている。これにより、カテゴリー知識からのトップダウンの影響が $P(f|C)$ に応じてカテゴリーと結びつきの強い特徴に対して与えられることになる。

統制条件でも同様に心理実験のカテゴリー学習に用いた 80 個のパターンを用いるが、Wake フェーズにおいて、出力層におけるカテゴリーラベルへの活性値の固定は行わず、また Sleep フェーズでは出力層における始めのトップダウンの活性は、出力ユニットへのバイアス項の入力のみから活性値を求めた。

このようにカテゴリー学習条件、統制条件ともに同じ 3 層構造のネットワークを用いるが、カテゴリー学習条件では出力層にカテゴリーラベルを与え、入力パターンからカテゴリーラベルを出力できるようにし、さらにカテゴリーラベルから内部表現にトップダウンに影響を与えるようになっているのに対し、統制条件では自己組織化により内部表現を学習するのみとなっている。

いずれの条件においても学習率は $\epsilon = 0.001$ とし、重みの初期値は全て 0 とした。各条件で別々のネットワークで学習を行い、 80 個のパターンをラ

ランダムな順序で入力することを1回の学習として、これを100,000回行った。

8.1.2 シミュレーションの手続き

学習が完了したネットワークに対して、瞬間呈示課題と同様の課題を行う。瞬間呈示課題において、被験者は対象に対する限られた情報しか得られず、それに基づいて図形の同定を行っていると考えられる。そこでシミュレーションにおいては、入力パターンの活性値を弱くして、さらにノイズを加えたものを入力とすることで入力パターンを劣化させ、瞬間呈示における視覚情報の制限をシミュレートした。

また、ここでは学習時とは異なり、入力パターンからのボトムアップの活性とボトムアップにより得られた出力層における表現からのトップダウンの活性を同時に行って、双方向の活性サイクルを繰り返し行い、活性パターンの変化が収束した時点での内部表現を入力パターンの内部表現とした。このようにすることで、ボトムアップとトップダウンの情報の両方の効果を受けた内部表現を求めることができる。双方向の活性サイクルはユニットへの入力値の変化が全て0.001以下になるまで行った。

シミュレーションではまず、瞬間呈示課題に用いた5つの図形の劣化していない状態を入力パターンとして入力し、双方向の活性を繰り返して安定状態に達した時点での隠れ層の活性パターンをそれぞれ求めておく。ここでの活性パターンは確率的に求めた $\{0, 1\}$ のパターンではなく、ユニットの活性確率そのものを用いた。

つぎに入力パターンの活性値を弱め、さらにランダムなノイズを加えたものを入力パターンとして、双方向の活性をネットワークの状態変化がほぼなくなるまで繰り返し行う。ネットワークの動作が安定した時点での隠れ層の活性確率のパターンに関して、あらかじめ求めておいた劣化していない状態でのパターンの隠れ層の表現の中から、ユークリッド距離が最も近いパターンを入力されたパターンと判断することで、パターン間の混同行列を求めた。

パターンの劣化のレベルは、もとの入力パターンに $e^{-6.5}$, $e^{-5.5}$, $e^{-4.5}$, $e^{-3.5}$, $e^{-2.5}$, $e^{-1.5}$ をそれぞれ掛けたものの6段階を用い、ノイズは全ての劣化レベルにおいて平均0, 標準偏差0.05の正規乱数を加えることで与えた。ただし値が負になった

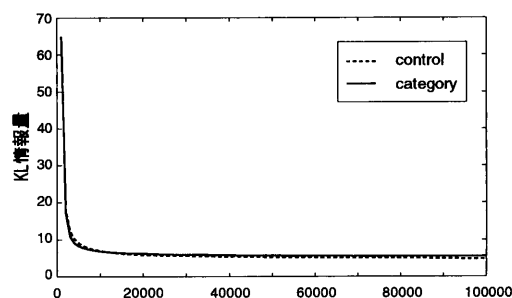


図6 ニューラルネットワークの学習誤差

ものは0とした。

このようにしてパターンと劣化レベルごとに、ランダムなノイズを変えて100回のシミュレーションを行い、全体で3000回の試行から混同行列を作成した。

8.2 結果

8.2.1 学習の結果

ネットワークの学習の進捗は、ボトムアップにより形成される各ユニットの活性確率 q と、トップダウンにより形成される各ユニットの活性確率 p との間のKullback-Leibler情報量(KL)を学習誤差として、その大きさにより測られる。ただし、カテゴリー学習条件においては出力層の活性がカテゴリーラベルに固定されるため、 KL の値は入力層と隠れ層においてのみ求めた。また統制条件においてもカテゴリー学習条件との比較のため、同様に入力層と隠れ層のみから KL の値を求めた。図6に学習経過における KL の変化を示す。図6から100,000回の学習により、どちらの条件でも KL が十分小さくなり、学習がほぼ収束していることが分かる。

また内部表現から入力パターンの再構成ができていないかを調べるために、入力パターンから隠れ層の活性をボトムアップに求め、得られた隠れ層の活性パターンから入力パターンをトップダウンに再構成し、もとのパターンとの誤差(パターンベクトル間のユークリッド距離)を求めた。図7に学習経過における誤差の変化を示す。図7から、学習により誤差は非常に小さくなり、内部表現からの入力パターンの再構成が問題なくできるようになっていることが分かる。

またカテゴリー学習条件においては、早い段階においてカテゴリー学習を達成し、248回目の学習以降は常に100%の正答率を達成した。

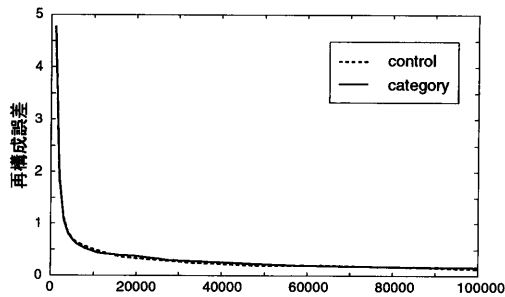


図7 ニューラルネットワークによる内部表現からの入力パターンの再構成誤差

8.2.2 心理実験のシミュレーションの結果

瞬間呈示課題のシミュレーションの結果を見る前に、まずはカテゴリー学習に用いなかった瞬間呈示課題用の5個のパターンが、カテゴリー学習条件のネットワークにより正しいカテゴリーに分類されていることを確かめる。各パターンを入力した時の出力層の活性確率のパターンを見ると、パターン1で{0.9998 0.0002}, パターン2で{0.9992 0.0001}, パターン3で{0.0000 1.0000}, パターン4で{0.0000 1.0000}, パターン5で{0.0001 0.9999}であった。このようにネットワークにおいてもカテゴリー知識の般化が正しくできており、さらに心理実験で見られた、パターン5がカテゴリー2と判断される傾向も再現された。

次に、瞬間呈示課題のシミュレーションの結果を見ていく。ここでは、シミュレーションにより得られた混同行列を心理実験と同じ手順により解析し、パターン間の弁別感度の劣化レベルによる変化を調べた。図8に弁別感度の劣化レベルごとの変化を示す。図8に見られるように、心理実験と同様に、パターンの情報が増えるにつれて弁別感度は上昇していくのが分かる。またカテゴリー学習条件の方が統制条件に比べ弁別感度が高くなる傾向も再現された。

8.2.3 ネットワークの内部表現

次にネットワークの内部表現（隠れ層ユニットにおける活性パターン）に関して、カテゴリーラベルが与えられた場合と与えられない場合との違いを調べた。ここでも内部表現を求める際には、双方向の活性を繰り返して安定状態における隠れ層ユニットの活性確率を求めた。またパターンセットには、カテゴリー学習に用いた80個のパターンを用

いた。

まず、パターンごとの隠れ層ユニットの活性確率の総和を求めることで、隠れ層の全体としての活性レベルを求めた。カテゴリー学習条件、統制条件のそれぞれで活性レベルのパターンセット全体での平均値を求めたところ、カテゴリー学習条件では6.8584であったのに対し、統制条件では7.5265であった。また、隠れ層のユニットの活性の分布がどのようなものであるかを調べるため、隠れ層ユニットの活性分布のエントロピー H を以下の式で求めた。

$$H = \sum_j -\bar{q}_j \log \bar{q}_j \quad (1)$$

ここで \bar{q}_j は、隠れ層のユニット j のパターンセット全体に対する平均活性確率を表す。カテゴリー学習条件では $H = 5.8945$ であり、統制条件では $H = 9.2258$ であった。

このようにカテゴリー学習条件においてエントロピーが低く、かつ全体の活性レベルが低い内部表現を用いていることから、内部表現における一つのパターンの表現に関して、統制条件では多くのユニットが中程度ずつ関与するが、カテゴリー学習条件ではより限定された少ないユニット数でパターンを表現するようになっていることがわかる。

ここで見られたような、カテゴリー学習条件における活性分布の偏りや活性ユニット数の減少は、カテゴリー知識が与えられたことによる効果であると考えられる。MDL原理に基づいてこのような符号化表現を考えると、まずカテゴリー情報が与えられない場合、パターンを最も効率的に符号化したとして、その平均符号長は、

$$\bar{l}_{NoCate} = \sum_j -P(f_j) \log P(f_j) \quad (2)$$

のようになると考えられる。ここで f_j はパターンに含まれる特徴を表し、 $P(f_j)$ はパターン空間全体におけるその確率分布を表す。

ここでカテゴリー知識が与えられた場合には、カテゴリーごとの特徴の確率分布をもとに符号化を行えば、より効率的な符号化が可能となり、その平均符号長は、

$$\bar{l}_{Cate} = \sum_k P(C_k) \left(\sum_j -P(f_j|C_k) \log P(f_j|C_k) \right) \quad (3)$$

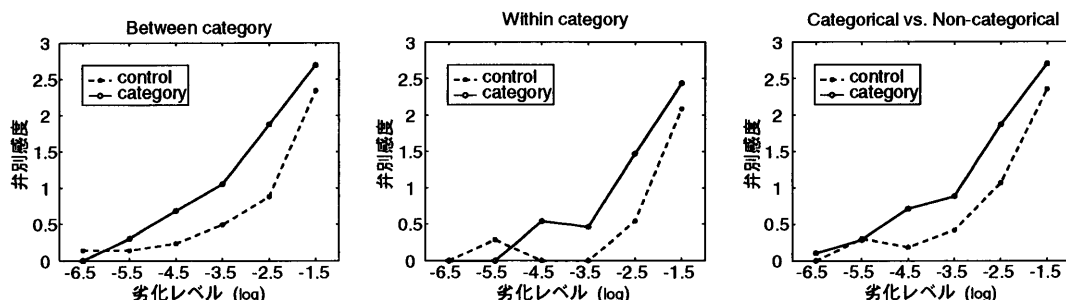


図8 ニューラルネットワークシミュレーションによる図形弁別率の変化

となる。カテゴリー分類が特徴空間内で分離できるようにうまく定義できていれば、 $P(f_j|C_k)$ は、 $P(f_j)$ よりも分散の小さい偏った分布を取り、そのエントロピーも小さくなることから、 \bar{I}_{Cate} は \bar{I}_{NoCate} よりも小さな値をとる。Wake-Sleep アルゴリズムは記述長を最小にする内部表現を求めるアルゴリズムになっており、カテゴリー学習条件のネットワークの方が隠れ層の内部表現において、よりエントロピーの低い表現が獲得できたのは、このような効果によるものと考えられる。

また実際に、カテゴリー学習条件のネットワークの方が統制条件のネットワークよりもカテゴリーごとに分離した表現を用いていることを確かめるために、隠れ層ユニットの活性確率の分布が、カテゴリーごとにどの程度異なるかを調べた。ここではカテゴリー学習に用いたパターンに関して、カテゴリーごとの隠れ層のユニットの平均発火確率 (\bar{q}_{j,C_1} , \bar{q}_{j,C_2}) を求め、以下の式により内部表現のカテゴリー間での KL 情報量を求めた。

$$KL = \sum_j \bar{q}_{j,C_1} \log \left(\frac{\bar{q}_{j,C_1}}{\bar{q}_{j,C_2}} \right) + (1 - \bar{q}_{j,C_1}) \log \left(\frac{1 - \bar{q}_{j,C_1}}{1 - \bar{q}_{j,C_2}} \right) \quad (4)$$

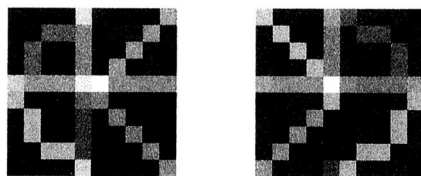
その結果、カテゴリー学習条件では $KL = 21.4619$ であったのに対し、統制条件では $KL = 13.4945$ であり、カテゴリー学習条件のネットワークの方が、内部表現においてカテゴリー間でより分離した表現を用いていることが分かった。

9. シミュレーションの考察

ニューラルネットワークシミュレーションにより、カテゴリー知識を与えることでパターンの識別率が向上する心理実験の結果を再現することができた。

このことから、カテゴリー知識の対象の識別における有効な働きが、カテゴリー知識からのトップダウンの効果により説明できることがわかった。ただし心理実験の結果では、異なるカテゴリーに属する対象間で最も強い効果が見られ、パターン5とパターン4の間の弁別感度のようにカテゴリー学習の効果が見られない部分もあったのに対し、ネットワークシミュレーションではいずれの対象間でもその傾向に大きな差はなく、常にカテゴリー学習条件のほうが高い識別率を示した。これはシミュレーションでは、わずかな差異でも厳密に判断して対象の識別を行うが、実際の人間の判断においては、認識における差がわずかであるときは、ある程度推測の要素が入るため、シミュレーションにおいてカテゴリー学習の優位性がより強調された形になったためと考えられる。

心理実験の考察では、カテゴリー知識の影響を特徴に対する影響として考察したが、ネットワークモデルでは隠れ層の内部表現においてパターンが分散的に表現されており、個別ユニットの活性と特徴との関係を明確に見ることはできなかった。しかし先に述べたように、今回用いたネットワークモデルではトップダウンの結合において、カテゴリー C が与えられたときに特徴 f を出力する条件付確率 $P(f|C)$ を学習することが Dayan & Hinton (1996) において示されており、ネットワークのカテゴリーラベル層のユニットの活性のみから入力層のパターンを再構成したときも、心理実験で考察したのと同様にカテゴリー1では特徴7, 8, 10がカテゴリー2では特徴5, 6, 11が強調されたパターンが再構成された(図9参照。図9では、正の強い結合ほど白く、負の強い結合ほど黒く表示している。)。このことからネットワークモデルにおいても、カテゴリーラベルからのトップダウンの影響がカテゴリー



カテゴリー 1

カテゴリー2

図9 カテゴリーラベルから再構成されたパターン

と強く関係する特徴の検出を助けることで入力パターンの識別を助けられていると考えられる。

またシミュレーションでは、カテゴリー学習条件の方が統制条件よりも活性レベルが低く、かつエントロピーの低い内部表現を獲得することができ、さらにカテゴリー間で分離した内部表現を獲得していた。これらの結果から、カテゴリー学習条件のネットワークにおいては、カテゴリー知識によりパターンの分布をより限定することで、効率的に入力パターンを表現していると考えられる。このことから、カテゴリー知識の獲得は、単に対象の識別において有効であるだけでなく、その処理の効率化においても有効であると考えられる。また、このような傾向は、実際の脳の活動で見られるカテゴリー知識の獲得による視覚野での活動の減少と何らかの関わりがあるのではないかと考えられる。すなわち、カテゴリー知識の獲得によって処理が効率的に行われ、それによって処理負荷が減少し、それが活動の減少となって現れたことが考えられる。ただし脳における活動の減少は何らかの抑制性の入力によるものとも考えられ、実際にどのような機構で活動の減少が起きているのかに関しては更なる検討が必要である。

10. まとめ

本研究の結果から、瞬間呈示課題を用いた心理実験では、カテゴリー知識を持つことの弁別感度における有効性が示され、またニューラルネットワークシミュレーションでは、トップダウンとボトムアップの相互作用により内部表現を形成するモデルで、心理実験の結果を再現し、さらに対象の内部表現の形成においても、カテゴリー知識が効率的な表現の獲得に寄与していることが示された。

本研究では、心理実験の行動指標において見られる現象と脳内における活動の変化をニューラルネットワークモデルを通じて結びつけることで、カテゴ

リー学習による視覚パターンの弁別効率の向上がどのような計算論的原理のもとで成立しているのかを考察した。心理現象のもととなっている脳内処理の機構を考察するときには、行動指標と脳活動の変化とを単純に対応付けることが困難である場合が多くあり、そのような場面においては今回のような計算論的原理とそれをもとにしたネットワークシミュレーションを用いた心理現象と脳活動の対応付けは有効な方法であると考えられる。

また、本研究ではボトムアップとトップダウンの相互の処理の機能的役割や計算論的意味を特にカテゴリー知識に関して考察したが、このような双方向処理は脳内において大きな役割を果たしていると考えられ (Humphreys et al., 1997; Tomita et al., 1999; Miyashita & Hayashi, 2000), 今回調べたような視覚認識におけるカテゴリー知識の影響だけでなく、認識のあらゆる場面に見られるものと考えられる。今後の研究においても、このような双方向処理がどのように行われているかを調べるとともに、その計算論的意味を明らかにしていくことが重要である。

謝 辞

本研究は文部科学省科学研究費補助金基盤研究(A)(2) (11309007), および文部省科学研究費補助金特別研究員奨励金 (00109739) 研究の一環として行われた。

文 献

- Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain*, 123, 620-640.
- Dayan, P. & Hinton, G. E. (1996). Varieties of Helmholtz Machines. *Neural Networks*, 9, 1385-1403.
- Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, 7, 889-904.
- Frey, B. J. & Hinton, G. E. (1997). A simple algorithm that discovers efficient perceptual codes. In L. Harris & M. Jenkin (Eds.), *Computational and Biological Mechanisms of Visual Coding*. New York: Cambridge University press.

- Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception and Psychophysics*, **25**, 425–431.
- Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The “wake-sleep” algorithm for unsupervised neural networks. *Science*, **268**, 1158–1161.
- Humphreys, G. W., Riddoch, M. J., & Price, C. J. (1997). Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Philosophical transactions of the Royal Society of London. Series B*, **352**, 1275–1282.
- Lawrence, D. H. (1950). Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, **40**, 175–188.
- Leopold, D. A. & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Science*, **3**, 254–264.
- Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 732–753.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol. 1, 103–189. New York: Wiley.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, **100**, 254–278.
- 三崎 将也・乾 敏郎 (1999). カテゴリー知識の獲得による認識空間の変化. 『認知科学』, **6** (2), 226–241.
- 三崎 将也・乾 敏郎 (2001). 視覚認識におけるカテゴリー知識の影響とその機能的役割. 『信学論DII』, **J84**, 2437–2447.
- Miyashita, Y. & Hayashi, T. (2000). Neural representation of visual objects: encoding and top-down activation. *Current Opinion in Neurobiology*, **10**, 187–194.
- Reber, P. J., Stark, C. E., & Squire, L. R. (1998). Cortical areas supporting category learning identified using functional MRI. *Proceedings of National Academy of Science USA*, **95**, 747–750.
- Rissanen, J. (1989). *Stochastic complexity in statistical inquiry*. Series in computer science. World Scientific.
- Smith, J. K. (1982). Recognition models evaluated: A commentary on Keren and Baggen. *Perception and Psychophysics*, **31**, 183–189.
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature*, **401**, 699–703.
- Townsend, J. T. (1971a). Alphabetic confusion: A test of models for individuals. *Perception and Psychophysics*, **9**, 449–454.
- Townsend, J. T. (1971b). Theoretical analysis of an alphabet confusion matrix. *Perception and Psychophysics*, **9**, 40–50.
- 山西 健司 (1992). MDL 入門：計算論的学習理論の立場から. 『人工知能学会誌』, **7** (3), 435–442.
- 山西 健司・韓 太舜 (1992). MDL 入門：情報理論の立場から. 『人工知能学会誌』, **7** (3), 427–434.

付 録

A. 刺激パターンの作成とカテゴリーの定義

実験で用いた刺激パターンは図1の12個の特徴から5個を選択して組み合わせることで作成した。ただしパターンの作成においては、特徴が重なる組み合わせ（特徴5と9、6と10、7と11、8と12がそれぞれ同時に存在するもの）、特徴が分離して存在するもの（9があつて1と2が無いもの、10があつて2と3が無いもの、11があつて3と4が無いもの、12があつて1と4が無いもの）、小さく閉じた部分があるもの（9と1と2、10と2と3、11と3と4、12と1と4がそれぞれ同時に存在するもの）などはそれ自体パターンとして特徴的であり、認識率を調べるにあたって平等に評価するのは難しいと考えられるので刺激パターンの候補から除外した。

上記のようなパターンを除くと、12個から5個をとってきて作成されるパターンは、全部で184個になる。この184個のパターンセットをもとに、各特徴が存在する事前確率 $P(f_k)$ を求めておく。ここで図1の12個の特徴をもとに、2つのカテゴリーを定めた。まず特徴 f_k が与えられたときに、カテゴリー C_l である条件付確率 $P(C_l|f_k)$ を $P(C_1|f_5) = P(C_1|f_{11}) = 0.1$, $P(C_1|f_8) = P(C_1|f_{10}) = 0.9$, $P(C_2|f_5) = P(C_2|f_{11}) = 0.9$, $P(C_2|f_8) = P(C_2|f_{10}) = 0.1$, その他の $P(C_l|f_k) = 0.4$ とし、

この $P(C_l|f_k)$ の値により、カテゴリー 1 の特徴として特徴 8 と 10 を、カテゴリー 2 の特徴として特徴 5 と 11 を定めた。ただし、この後で実際に使用するパターンを絞っていくので、ここでの値はあくまで初期値としての仮の値である。

この値をもとに、先の 184 個のパターンを 2 つのカテゴリーに分類する。分類の基準としては、以下のような式を用いた。

$$L(C_l|f^p) = \frac{\prod_k P(C_l|f_k^p)P(f_k^p)}{\sum_m \prod_k P(C_m|f_k^p)P(f_k^p)} \quad (5)$$

ここで f_k^p はパターン p が持つ特徴 k を表す。式 (5) は、各特徴ごとの事後確率 $P(f_k|C_l)$ ($= P(C_l|f_k)P(f_k)$) をパターン内で掛け合わせ、そのカテゴリーごとの尤度比をとってカテゴリー分類の基準としている。ここで尤度比が 0.5 のものはどちらのカテゴリーにも分類されないものとした。

このようにして 184 個のパターンが、カテゴリー 1 のパターン 61 個、カテゴリー 2 のパターン 64 個、どちらにも属さないもの 59 個に分類された。

次にカテゴリーに分類されたパターンの中から、実験に用いるパターンを選択する。まずは 2 つのカテゴリーのパターン数を同じにするため、数が多い方のカテゴリー（カテゴリー 2）から尤度比の低いものを除外して、2 つのカテゴリーのパターン数を 61 個ずつにした。これら 61 個のパターンの全てをカテゴリー学習課題に用いることも可能であるが、尤度比が低いものを用いると、特徴定義によるカテゴリー間の対比があいまいになり、またカテゴリーに分類されない対象と分類される対象との差も弱くなるため、ここではこれらのパターンから尤度比が高いものを選んで用いることとした。

カテゴリー学習には 40 個のパターンを用い、また瞬間呈示課題では各カテゴリーで 2 個ずつのパターンを用いるので、各カテゴリー 42 個ずつのパターンを尤度比の高いものから順に選び出した。ただしパターンを減らしていくと、全体としての特徴出現の事前確率 $P(f_k)$ や、事後確率 $P(f_k|C_l)$ などに変化するので、61 個の段階から最も尤度比が低いものを各カテゴリーで 1 個ずつ削除して、 $P(f_k)$ 、 $P(f_k|C_l)$ を再計算し、再び最も尤度比が低いものを削除するという手順を繰り返して、両カテゴリー 42 個ずつの 84 個のパターンを選択した。

次に、選択されたパターンから、瞬間呈示課題に

用いるパターンを選び出す。パターン選択の基準は、各カテゴリーに特有の特徴である $P(f_k|C_l)$ が高い特徴を持っていることと、瞬間呈示課題に用いる他のパターンが持っていない特徴がなるべく含まれないこと（その特徴さえわかればパターンが特定できてしまうようなことが無いこと）とした。

この時点で、再度 $P(f_k|C_l)$ を計算し、各カテゴリーにとって示唆的な特徴を調べたところ、カテゴリー 1 で特徴 7 ($P(f_7|C_1) = .667$, $P(f_7|C_2) = .095$)、特徴 8 ($P(f_8|C_1) = .738$, $P(f_8|C_2) = .143$)、特徴 10 ($P(f_{10}|C_1) = .619$, $P(f_{10}|C_2) = .048$) が、カテゴリー 2 で特徴 5 ($P(f_5|C_1) = .143$, $P(f_5|C_2) = .738$)、特徴 6 ($P(f_6|C_1) = .119$, $P(f_6|C_2) = .643$)、特徴 11 ($P(f_{11}|C_1) = .000$, $P(f_{11}|C_2) = .667$) が示唆的な特徴であった。

これらの基準をもとに図 2 の 1 から 4 のパターンを瞬間呈示課題に用いるパターンとして選んだ (1, 2 はカテゴリー 1 から、3, 4 はカテゴリー 2 から選ばれた)。表 1 では瞬間呈示課題に用いたパターンを取り除いた上で再度 $P(f_k|C_l)$ を計算したものを示している。

最後に、瞬間呈示に用いるカテゴリーに属さないパターンを作成する。このパターンはどちらのカテゴリーに特有な特徴も持たないことが望ましいが、そのようにすると瞬間呈示に用いる他のパターンとの見分けがつきやすくなってしまいますので、ここでは特徴 6 と 7 を同時に含み、その他の特徴はどちらのカテゴリーも示唆しないようなパターンを用いた (図 2 のパターン 5)。

以上のようにして、カテゴリー学習に用いるパターン 80 個 (各カテゴリー 40 個ずつ) と、瞬間呈示課題に用いるパターン 5 個 (カテゴリー 1 から 2 個、カテゴリー 2 から 2 個、どちらにも属さないもの 1 個) を作成した。

B. ネットワークの動作手順

ネットワークの各ユニットは確率的に活性値 0 または 1 をとり、その確率はユニットへの入力値により決まる。 i 層のあるユニットへの入力 x_i は以下の式により表される。

$$x_i = b_i + \sum_j s_j w_{ij} \quad (6)$$

ここで b_i はこのユニットのバイアス、 s_j はユニットに結合している j 層のユニットの活性値、 w_{ij} は

その結合荷重である。この x_i により、ユニットの活性が 1 の値をとる確率 p_i は、

$$p_i = \text{prob}(s_i = 1) = \frac{1}{1 + e^{-x_i}} \quad (7)$$

により決まる。

Wake-Sleep アルゴリズムではネットワークの学習は Wake フェーズと Sleep フェーズの 2 つのフェーズで行われる。ここで同じユニットの活性確率を、ボトムアップによる活性確率 q と、トップダウンによる活性確率 p とに区別しておく。またボトムアップの結合荷重を ϕ で表し、トップダウンの結合荷重を θ で表す (図 5 参照)。まず Wake フェーズではボトムアップの入力から各ユニットの活性確率 q と活性値 s^α を求める。次に上位の層の活性値 s^α をもとにトップダウンの入力から、ユニットの活性確率 p を求める。このように、ネットワークの活性はトップダウンとボトムアップで別々に求められ、双方向結合による活性の反響などは行わない。Wake フェーズで学習を行うのはトップダウンの結合 θ のみで、トップダウンの入力により得られる活性値がボトムアップで得られるものに近づくように、以下の式で θ を修正し学習を行っていく (ϵ は学習率)。

$$\Delta\theta_{jk} = \epsilon s_k^\alpha (s_j^\alpha - p_j) \quad (8)$$

Sleep フェーズではトップダウンの入力から各ユニットの活性確率 p と活性値 s^γ を求める。次に下位の層の活性値 s^γ をもとにボトムアップの入力からユニットの活性確率 q を求める。Sleep フェーズで学習を行うのはボトムアップの結合 ϕ のみで、ボトムアップの入力により得られる活性値がトップダ

ウンで得られるものに近づくように以下の式で ϕ を修正し学習を行っていく。

$$\Delta\phi_{ji} = \epsilon s_i^\gamma (s_j^\gamma - q_j) \quad (9)$$

この Wake フェーズと Sleep フェーズを 1 回の学習ごとに行い、入力からボトムアップに内部表現を構成し、内部表現から入力を再構成できるように結合荷重 θ , ϕ を修正していく。

(Received 9 April 2001)

(Accepted 15 Jan. 2002)



三崎 将也 (正会員)

1996 年京都大学文学部心理学専攻卒。1998 年京都大学大学院文学修士修了。2001 年博士 (情報学) (京都大学)。現在、独立行政法人通信総合研究所関西先端研究センター。日本心理学会、電子情報通

信学会各会員。



乾 敏郎 (正会員)

1976 年大阪大学大学院基礎工学研究科物理系修士修了。1987 年から 1991 年 (株) ATR 視聴覚機構研究所主幹研究員。1991 年から 1997 年京都大学文学部教授。現在、京都大学大学院情報学研究科教授。1999 年より日本学術振興会未来開拓研究プロジェクトリーダー (兼務)。視覚情報処理およびコミュニケーションメカニズムの研究に従事。文学博士。工学修士。日本心理学会、日本神経心理学会各会員。