

TR - A - 0108

神経回路モデルを用いた  
重なったパターンの分離

村上 由彦

1991. 3.25

ATR 視聴覚機構研究所

〒619-02 京都府相楽郡精華町乾谷 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

Inuidani, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411

Facsimile: +81-7749-5-1408

Telex: 5452-516 ATR J

## 目次

1. はじめに	1
2. 階層型神経回路モデル	1
2. 1 誤差逆伝播法	2
2. 2 恒等写像型連想記憶	3
2. 2. 1 3層型恒等写像連想記憶モデル	3
2. 2. 2 5層型恒等写像モデル	4
2. 3 誤差逆伝播型連想記憶	4
2. 3. 1 一対一学習	5
2. 3. 2 多対一学習	6
2. 3. 3 拘束条件	6
2. 3. 4 4層型神経回路モデル	7
2. 4 階層型神経回路モデルの考察	7
2. 4. 1 パターン学習	7
2. 4. 2 パターン想起	7
3. 相互結合型神経回路モデル	9
3. 1 学習則	9
3. 2 計算機シミュレーション	10
3. 3 考察	11
4. ネオ・コグニトロン	11
5. まとめ	12

## 1. はじめに

近年、多くの人々によって、神経回路モデルをパターン認識に応用しようとする試みがなされている。その多くの場合、認識を行ないたいパターンセットについて、それぞれのサンプルを入力し、出力に各カテゴリに対応するユニットを設けて、そのユニットのみが出力し、その他のユニットは出力しないように学習を行なう。この種の神経回路モデルは、パターンセットをカテゴリ分割する機能に優れており、文字認識や音声認識などに応用されている。しかし、物体の認識のような画像処理に神経回路モデルを適用した場合、工学的な応用を考えると、そこにあるパターンがどのカテゴリに含まれるかを知るだけでは不十分であり、シーンの中でそのパターンの位置や状態を知らなければ、部品検査やロボットの目になるようなシステムを構築することができない。

一方、人間の情報処理機構を考えると、このようなパターン抽出機構は「注意」という概念で与えられる。すなわち、ある物体（パターン）に注意を向けるとその物体だけがそのシーンの中から浮かび上がってくる。このような機構は、一つの考え方として上位の処理レベルから下位への処理レベルへの情報の伝達によって実現されていると考えられる。このような神経回路モデルを構築するには、入力パターンより記憶パターンを再現する機構（連想記憶）と神経回路自身を制御する機構（選択的注意）を持つ必要がある。

そこで、神経回路モデルに対して連想記憶の検討を行ない、モデルの変更によってこのような選択的な注意の機構を導入するにより、パターンの抽出を行なうことを検討する。特に、ここではパターンが重なり合った場合について、個々の図形を切り出すことに目標をおく。

## 2. 階層型神経回路モデル

階層型神経回路モデルの学習則について、Rumelhartらによって誤差逆伝播法が提唱されて以来 [1]、多くの研究者によってその応用が研究されてきた。誤差逆伝播法は教師つき学習として、出力層と結合していないものについても学習できる意味で強力な学習アルゴリズムである。この誤差逆伝播法を連想記憶に応用したモデルとして、大別すれば、恒等写像型と誤差逆伝播型に分けることができる。以下に、その誤差逆伝播法及びこれらの連想記憶モデルについて詳しく述べる。

## 2. 1 誤差逆伝播法 (Back-Propagation)

図 2. 1 のような 3 層型神経回路モデルの場合について、誤差逆伝播法による学習に以下を示す。むろん同様にして、4 層以上の階層型神経回路モデルについても学習可能である。

### 学習則

パターン  $p$  を与えた時、出力層ユニット  $k$  の誤差  $E_{pk}$  は、ユニットの出力  $O_{pk}$ 、教師信号  $t_{pk}$  とすれば、(2.1) 式のように定義する。そして、神経回路モデルに対しての全体の誤差は、(2.2) 式のように出力層ユニットの誤差の総和になる。

$$E_{pk} = \frac{1}{2} \cdot (T_{pk} - O_{pk})^2 \quad (2.1)$$

$$E_p = \sum_k E_{pk} = \frac{1}{2} \cdot \sum_k (T_{pk} - O_{pk})^2 \quad (2.2)$$

中間層・出力層の各ユニットは、前段の層から受けた信号を (2.3) 式に示すようなシグモイド関数を通して出力する。つまり、各ユニットへの入力を  $net_p$ 、外部からのバイアス入力を  $S$  とすれば、(2.4) 式のような関係になる。

$$y = \text{sigmoid}(x) = 1 / (1 + \exp(-x)) \quad (2.3)$$

$$O_{pk} = \text{sigmoid}(net_{pk}), \quad net_{pk} = \sum_j (w_{jk} \cdot O_{pj} + S_j) \quad (2.4)$$

これらの関係式より、 $W_{ij}$  を入力層から中間層へのウエイト、 $W_{jk}$  を中間層から出力層へのウエイトとすれば、各ウエイトに対するエラーの偏微分は、(2.5) 式のように計算できる。

$$\frac{\partial E_p}{\partial w_{jk}} = \frac{\partial E_p}{\partial O_{pk}} \cdot \frac{\partial O_{pk}}{\partial net_{pk}} \cdot \frac{\partial net_{pk}}{\partial w_{jk}} = -(T_{pk} - O_{pk}) \cdot O_{pk}(1 - O_{pk}) \cdot O_{pj} \quad (2.5a)$$

$$\begin{aligned} \frac{\partial E_p}{\partial w_{ij}} &= \frac{\partial E_p}{\partial O_{pj}} \cdot \frac{\partial O_{pj}}{\partial net_{pj}} \cdot \frac{\partial net_{pj}}{\partial w_{ij}} \\ &= \left[ \sum_k -(T_{pk} - O_{pk}) \cdot O_{pk} \cdot (1 - O_{pk}) \cdot w_{jk} \right] \cdot O_{pj}(1 - O_{pj}) \cdot O_{pi} \end{aligned} \quad (2.5b)$$

(2.5) 式は、誤差の最急降下方向を表しているので、(2.6) 式に従って誤差が十分小さくなるまでウエイトを変更を繰り返せば、出力層のユニットの出力は教師信号に近づく。

$$\Delta w_{jk}(t) = -\varepsilon \cdot \frac{\partial E_p}{\partial w_{jk}} + \alpha \cdot \Delta w_{jk}(t-1) \quad (2.6a)$$

$$\Delta w_{ij}(t) = -\varepsilon \cdot \frac{\partial E_p}{\partial w_{ij}} + \alpha \cdot \Delta w_{ij}(t-1) \quad (2.6b)$$

ここで、 $\varepsilon$  は学習係数、 $\alpha$  は慣性率である。

## 2. 2 恒等写像型連想記憶

出力層に入力層と同じ数だけのユニットを用意し、記憶パターンに対してそのパターン自身を出力するように学習を行なうもので、神経回路モデルは恒等写像を実現するように学習されるので、この種の連想記憶モデルを恒等写像型連想記憶と呼ぶことにする。

### 2. 2. 1 3層型恒等写像連想記憶モデル [2]

図2. 2に示すように、入力層と中間層と出力層の3層で構成される。学習は記憶パターンとして入力層に与えたパターンを出力層に教師信号として与える。図2. 3に学習に使用した数字パターンを示す。これらは、 $16 \times 16$ ドットで表現された0から9までの10個のパターン、計100パターンであり、以後の計算機シミュレーションでは、これらのパターンの全部又は一部を使用する。

#### 計算機シミュレーション1

0から4までの5つのパターンを5種類、計25パターンを入力層・出力層各256ユニット・中間層30ユニットの3層型神経回路モデルに対して、各500回呈示し学習させた。学習パラメータは、学習率：0.1 慣性率：0.9とした。学習後、それらの記憶パターンの2種類が重なったパターンを入力し、想起されたパターンの変化を図2. 4に示す。出力層の信号を入力層にフィードバックすることで、想起されたパターンが徐々に記憶したパターンに近づいていくことがわかる。

このことは、神経回路モデルに対しての誤差関数を考えた時、学習によって各記憶パターンがその誤差関数の極小解に埋め込まれていると解釈でき、さらにフィードバック処理を繰り返すことで、この誤差が徐々に減少し、記憶パターンに到達するものと考えられる。

しかし、このモデルでは記憶パターンのカテゴリ識別を行なうような機構がモデルに組み込まれていないので、パターン認識機能を持たない。さらに、想起パターンを外部から制御する機構が存在しないため、2つの重なったパターン以外の記憶パターンが想起されることもある(図2. 5)。

## 2. 2. 2 5層型恒等写像連想記憶モデル [3]

前節のモデルの問題点を改善するために、図2. 6に示すような5層型のモデルが提案されている。このモデルは、入力層・中間層・出力層・中間層・入力再現層の5層から構成されている。出力層の一部のユニットは、カテゴリを識別する役割を果たす。入力再現層のユニットは記憶パターンを再現する機能がある。これらのユニットのそれぞれに教師信号を与えて学習を行なう。

学習後には入力層にパターンを入力し、出力層に現れた状態によってカテゴリ識別を行なうことができる。また、記憶パターンの想起に関しては、入力層に初期のパターンを与えて、さらに想起したいパターンが含まれているカテゴリに対応する状態に出力層を固定する。この時、入力再現層に現れたパターンを入力層へフィードバックし、これらの手続きを繰り返して行なうことで、記憶パターンが想起できるものと考えられる。また、出力層では制御の対象とならない（カテゴリには関与しない）ユニットを設けているので、これらのユニットの状態によって同じカテゴリに含まれるパターンも分離することが可能であると考えられる。

しかし、人間の情報処理機構から考えて階層構造を持っているとしても、この神経回路モデルのように、上位層で下位層のパターンを再構成していると考えるのは不自然である。

## 2. 3 誤差逆伝播型連想記憶

一般に、階層型神経回路モデルをパターン認識に応用した場合、図2. 7のように、出力層に各パターンのカテゴリに対応したユニットを設けて、神経回路モデルに記憶パターンを入力した時に、そのパターンのカテゴリに対応したユニットのみが出力するように学習を行なう（以後、カテゴリ学習と呼ぶ）。この誤差逆伝播法によって学習の終了した階層型神経回路モデルに対して、ある目標となる出力を作成し、さらにそれに対する誤差を評価することにより、神経回路モデルの入力層に記憶パターンの再現をしようとするモデルをここでは誤差逆伝播型連想記憶と呼ぶことにする [4,5,6]。

この神経回路モデルでは学習時と想起時において、入力層と出力層に対する区別が混乱するので、改めて、学習時には記憶パターンを入力し想起時にパターンを出力する層を下位層と呼び、カテゴリ識別するパターン認識層を上位層と呼ぶことにする。

パターンの想起処理は以下のような手順によって行なう。

- (1) 記憶パターンセットに対して、カテゴリ学習を行なう。
- (2) 初期パターンを入力する。
- (3) 学習により獲得したウエイトを用いて、現在の出力層ユニットの出力を計算する。
- (4) 上位層において、最大の出力を出したユニットに対応したカテゴリのパターンが含まれていると仮定する。そして、その最大の出力を出したユニットのみが出力するような目標出力を作成する。
- (5) 初期パターンを入力した時の上位層の出力と設定した目標出力の誤差を逆伝播し、下位層の状態を変更する。

(6) 変更後の下位層の状態上位層の出力を計算し、同様な処理を誤差が十分小さくなるまで繰り返す。

この想起処理の流れを図2.8に示す。

上記のような処理を繰り返し行なえば、最終的に目標出力として設定したカテゴリに含まれるようなパターンを構成することができる。また、下位層に対してのパターンの変更は、学習の時と同様に誤差逆伝播法を用いて、(2.7)式のように下位層の出力に対する誤差の偏微分を求めることができ、(2.8)式に示したように変更すれば、その時の誤差が小さくなる。

$$\frac{\partial E_p}{\partial O_{pk}} = -(T_{pk} - O_{pk}) \quad (2.7a)$$

$$\frac{\partial E_p}{\partial O_{pj}} = -\sum_k (T_{pk} - O_{pk}) \cdot O_{pk} (1 - O_{pk}) \cdot w_{jk} \quad (2.7b)$$

$$\frac{\partial E_p}{\partial O_{pi}} = \sum_j \left[ -\sum_k (T_{pk} - O_{pk}) \cdot O_{pk} (1 - O_{pk}) \cdot w_{jk} \right] \cdot O_{pj} (1 - O_{pj}) \cdot w_{ij} \quad (2.7c)$$

$$\Delta O_{pi}(t) = -\varepsilon \cdot \frac{\partial E_p}{\partial O_{pi}} + \alpha \cdot \Delta O_{pi}(t-1) \quad (2.8)$$

ここで、 $t$ は変更回数であり、 $\varepsilon, \alpha$ は小さな定数とする。

また、このモデルにおいて、上位層をどのカテゴリに対応した状態に設定するかによって、想起パターンを制御することができる。

この操作を人間に当てはめてみると、あるシーンを見たとき、そこには何があるかを教示するとその物体が浮かび上がってくるような状況に似ており、注意の機構のような処理を実現しているとも考えられる。しかし一般には、想起パターンは初期パターンに依存するものと考えられる。

### 2.3.1 一対一学習 [4]

一つのカテゴリに対して、一つのパターンだけを学習させた。この場合、一つの教師信号に対して一つのパターンだけを記憶するので、この時の学習を一対一学習と呼ぶ。

#### 計算機シミュレーション2

0から9までの10個のパターンを下位層256ユニット・中間層15ユニット・上位層10ユニットの3層神経回路モデルに対して、0から9の順に2000回呈示し、学習を行なった。学習パラメータは、学習率：0.1 慣性率：0.9とした。その後、前述のアルゴリズムに従って上位層を固定し、下位層に現れるパターンを調べた。

一方、パターンの構造を学習する方法として、学習パラメータに忘却の係数を設定する方法がある。ここでは、一度も学習の起こらない結合に対して、ウエイトをゼロにする意味もあり、(2.9)式に従って設定した。

$$w_{jk}(t) = -\eta \cdot w_{jk}(t-1) + \alpha \cdot \Delta w_{jk}(t-1) \quad (2.9a)$$

$$w_{ij}(t) = -\eta \cdot w_{ij}(t-1) + \alpha \cdot \Delta w_{ij}(t-1) \quad (2.9b)$$

その結果、図2.9に示すように、上位層に与えた目標出力によって初期パターンに依存

せずに、下位層に記憶パターンを出力することができた。

このことは、双方向相互想起型連想記憶（Bidirectional Associative Memories）[7,8]にも関連するものと考えられる。

### 2. 3. 2 多対一学習

一つのカテゴリに対して複数のパターンを学習させた。この場合、モデルに与えた初期パターンによって、同じカテゴリ内に記憶されたパターンの中で一番近いと思われるパターンを想起することを目標におく。この際、記憶パターンと上位層での教師信号の誤差との関係は、想起処理アルゴリズムから図2. 10のような状態である必要がある。一般には、複数のパターンを一つの出力状態に学習した場合（多対一学習）、その逆変換は一意に定まらないため一種の不良設定問題となっている。そこで、何らかの拘束条件を用いて解空間を制限し、想起パターンを決定する必要がある。一方、上位層を目標出力に設定することで解空間を制限することになるが、カテゴリ内のパターンを決定するための条件にはならない。そこで、拘束条件を導入することで、想起方法に変更を加えることを考える。

### 2. 3. 3 拘束条件 [9]

拘束条件を神経回路モデルに組み込む際、一般的な手法としてはその拘束条件に相当する誤差を神経回路モデルの誤差に加えて全体の誤差と見なし、拘束条件を満たした時にその全体の誤差が小さくなるように定義する。

ここではカテゴリ内の想起パターンの精度を上げるために、以下の拘束条件を用いた。

(1) 入力パターンは、ストロークなどの連続性を持つ。

構造を取り込んだ形で、図形を再構成する。

(2) 記憶したパターンが重なっている状態を想定し、入力層に与えられた初期パターンには、記憶されたパターンが含まれていると仮定する。

初期パターンの論理値1で構成された部分空間内で想起パターンの探査を行なう。

この時、探査軌跡を何らかの方法で制御し、学習パターンを通過させるようにする。

### 計算機シミュレーション3

0から4までの5つのパターンを5種類、計25パターンを下位層256ユニット・上位層10ユニットとして、中間層30ユニットの場合と50ユニットの場合についての3層神経回路モデルに対して、各2000回呈示し学習させた。学習パラメータは、学習率：0.1 慣性率：0.9とした。学習後、重なったパターンを入力し、パターンの分離能力について調べた。その結果を図2. 11に示す。図より初期パターンに関係なく、同じカテゴリに学習したパターンの平均パターンが出力されることがわかる。また、拘束条件を入れた場合でも想起パターンの精度が良くないことがわかる。つまり、この結果を考えると学習時に、パターンを空間的な構造に則した形で記憶を行っていないことが一つの原因であると考えられる。



## 2. 3. 4 4層型神経回路モデル

パターンの構造を学習することを目標として、受容野を持つような形の神経回路モデルについて検討した。ここでは、4層構造のモデルを考え、第1層から第2層への結合領域を限ることで受容野を設定した。

### 計算機シミュレーション4

0から4までの5つのパターンを5種類、計25パターンを下位層256ユニット・下位中間層169ユニット・上位中間層30ユニット・上位層10ユニットの3層神経回路モデルに対して、各300回呈示し学習させた。学習パラメータは、学習率：0.1 慣性率：0.9とした。同様に、重なったパターンを入力し、パターンの分離について調べた。その結果、学習時にはパターンの空間的な構造を学習する前に、受容野内の論理値1の部分の頻度を学習してしまうことがわかった。

## 2. 4 階層型神経回路モデルの考察

### 2. 4. 1 パターンのカテゴリ学習

3層型神経回路モデルでは、パターンの空間的な構造を学習するためには、多くのパターンを記憶させる必要がある。ここで空間的な構造とは、線分の方向や屈折・分岐などの部分的な線分のつながりのことを指す。また、神経回路モデル自身はパターンの構造を抽出する機能はなく、ただその時の初期ウエイトから出発して、自分が含まれるカテゴリと他のカテゴリの区別をするだけの情報量を抽出し、中間層でコード化を行なうだけに過ぎず、構造を抽出する機能を持つためにはそのようなパターンセットを学習させる必要がある。いいかえるならば、記憶パターンセットが少ない場合、カテゴリ識別するための情報が空間的構造を反映したものである保証は全くなく、ましてや人間がカテゴリ識別するための評価とはかけ離れている場合も多い。

一方、4層型神経回路モデルにおいて、結合領域を限り受容野を設けることで、学習するパターンセットに余りとらわれることなく空間的な構造を学習する可能性がある。つまり、神経回路モデルは、その受容野内でパターンの識別を行なうことを強いられるので、空間的な構造に依存したようなコード化が行なわれるはずである。しかし、一つの受容野について一つのユニットしかない場合、そのユニットがどのような特徴を抽出すればよいかかわからないため、出力頻度を学習するような状態が起こる。ゆえに、中間層がそれぞれの空間的構造の特徴を抽出するように分化するためには、同じ受容野を持つユニットが多数存在し競合するような仕組みにする必要がある。

### 2. 4. 2 パターン想起

恒等写像型連想記憶では、記憶パターンを明示的に学習するので、想起したパターンも記憶したパターンに非常に近いものが得られる。しかし、想起パターンを制御することは難しい。

誤差逆伝播型連想記憶では、中間層の出力値が2状態（0または1に非常に近い状態）

に落ち着いている方が想起パターンは安定である。このことは、(2.7)式を見ても明らかである。また、一つのカテゴリに対して複数のパターンを学習させた場合、神経回路モデルが判断した距離（必ずしもパターン空間内のユークリッド距離ではないという意味）に近いものどうしが1つの内部表現として学習される。ここでの内部表現が同じと言う意味は、中間層ユニットにおける出力状態を2状態として見たときに同じという意味である。これは、ネットワークの汎化能力とも考えられ、その近傍で学習パターンが存在しない場合にもそのパターンはそのカテゴリに含まれるものと解釈される。ゆえに、このように学習された神経回路モデルでは、前述のアルゴリズムの想起処理を行なった場合、同じ内部表現で記憶されたパターンの平均的パターンが想起される。

記憶パターンを完全な形で想起するためには、以下の問題点を解決する必要がある。

- (1) カテゴリ分割に関与しない入力ユニットについて、いかなる処理を行なうか？  
入力を受けることが出力結果にほとんど影響しないユニットが入力層に存在する。この現象は中間層が少ない場合によく見られ、他のパターンとのカテゴリを識別するだけの必要な特徴だけを中間層で抽出するために起こるものと思われる。現象としては、逆伝播されてくる誤差が非常に小さいもしくは学習時にウエイトが小さくなってしまふことがあげられる。
- (2) 入力ユニットのオーバーフロー（アンダーフロー）をどのように対処するか？  
基本的には、(2.7)式において、一回のパターンを変更する処理ではウエイト比によって誤差が分配される。実際の変更においては、(2.7)式より入力層の値には無関係に誤差が逆伝播されるため、今、入力ユニットが0または1の2値をとるものとすれば、誤差が逆伝播したときに0以下（1以上）の値に変更するように要請される。このオーバーフロー（アンダーフロー）を抑え込むことによって、次回の誤差逆伝播の時にその誤差がウエイト比によって再分配される。
- (3) パターンの空間的構造（線分の端点・屈曲・分岐など）を内部表現として学習しないと想起パターンの精度が上がらない。
- (4) 記憶パターンに対して、誤差極小である保証がない。  
一つのパターンごとにウエイト変更を行なう場合、そのパターンについては誤差極小となる方向に学習されるが、すべてのパターンセットを呈示し学習を行なうことによって、最終的にはすべての記憶パターンに対しての誤差の和が極小になるように学習される。逆に各パターンについて誤差極小となるためには、中間層での発火状態が違った形で学習される必要がある。
- (5) 記憶パターンが持つ本質的な拘束条件を使用する。  
誤差極小を補う意味で、拘束条件を付け加えて検討した。しかし、記憶パターンそのものが持つ不変的な拘束条件を導入しないことには、得られたパターンには意味がないように思われる。例えば、川人らによって行なわれた運動軌道生成の問題 [10] に対しても、トルク変化最小という拘束条件を用いることによって期待する解が求めることができた。しかし、図形や物体についての一般的な拘束条件を求めることは、パターン認識において特徴とは何かという問題を解くことと同じぐらい困難な問題であると言ってもよい。

### 3. 相互結合型神経回路モデル

相互結合型神経回路モデルを用いた連想記憶の多くは、各ユニットどうしが相互に結合したモデルについて研究されている。一般に、連想記憶モデルは自己想起型と相互想起型とに分けることができる。しかし、これらの区別は、どのユニットを入力部や出力部として用いて学習を行なうかの違いだけで学習則自体は同じである。また、相互結合型神経回路モデルでは、明示的に記憶パターンと出力したパターンとの誤差が小さくなるように学習を行なうこと可能であり、記憶パターンに対しての誤差極小となる機構が存在する。

#### 3. 1 学習則 [11]

神経回路モデルの構造は、ここではパターン認識機能を持つように、ユニットを下位層部・中間層部・上位層部を持つモデル（図3. 1）について考える。下位層部では、パターン自体を入力として受け、上位層部ではカテゴリに対応づけたユニット（grandmother cell）を設ける。この種の神経回路モデルは、一般の相互結合型の連想記憶モデルに対して、隠れ層を持つところに違いがあり、この隠れ層がうまく働くことによって、上位層であるパターン認識層での複雑な情報処理が可能となる。この神経回路モデルに対してエネルギー関数を定義し、その値を小さくすることで神経回路モデルの定常解を求めることができる。学習は、2つの神経回路モデルの定常解の状態の違いによってウエイトの変更をする。

まず、神経回路モデルに対して (3.1) 式のようなエネルギー関数を定義する。

$$F(V) = E(V) - T \cdot S(V) \quad (3.1a)$$

$$E(V) = -\frac{1}{2} \cdot \sum_i \sum_j W_{ij} \cdot V_i \cdot V_j - \sum_j J_j \cdot V_j \quad (3.1b)$$

$$S(V) = -\sum_i [V_i \cdot \log V_i + (1 - V_i) \cdot \log(1 - V_i)] \quad (3.1c)$$

ユニットの内部ポテンシャルを  $U$  出力を  $V$  とすれば、(3.2) 式のような関係がある。

$$V_i = \text{sigmoid}(U_i) = 1 / (1 + \exp(-U_i)) \quad (3.2)$$

この回路の定常解は  $dF/dV = 0$  で求めることができ、内部ポテンシャルを (3.3) 式に従って変更すればよい。この時の内部ポテンシャルは、(3.4) 式の運動方程式に従う。

$$\frac{dU_i}{dt} = -\frac{\partial F}{\partial V_i} = T \cdot \frac{\partial S}{\partial V_i} - \frac{\partial E}{\partial V_i} \quad (3.3)$$

$$\tau \cdot \frac{dU_i}{dt} = -U_i + \left( \sum_j W_{ij} \cdot V_j + J_j \right) \quad (3.4)$$

実際のウエイト変更は、以下のようにして行なう。

記憶パターンセットに対して、学習を行ないたい状態に下位層ユニット・上位層ユニットを固定し、中間層を自由に動かした時に定常解に達した時の各ユニットの状態を

$V^*$ とする。同様に、下位層ユニットのみを固定した時の各ユニットの状態を  $V$  とすると、エラー評価関数  $E_{rr}$  を (3.5) 式のように定める。

$$E_{rr} = F(V^*) - F(V) \quad (3.5)$$

学習時に、このエラーが十分小さくなったということは、パターンを入力した時に神経回路モデルは、自動的に固定した上位層ユニットの状態を出力することができるようになっていたので、(3.6) 式に従ってウエイト変更すればよい。

$$\Delta W_{ij} = -\eta \cdot \frac{\partial E_{rr}}{\partial W_{ij}} = \frac{\eta}{2} (V_i^* \cdot V_j^* - V_i \cdot V_j) \quad (3.6)$$

下位層から上位層の双方向に学習するために、交互に固定してウエイト変更を行なった。また、神経回路モデルの構造より、中間層からは下位層と上位層との区別が与えられた教師信号だけに依存するために、初期ウエイトを入力層に対しての総和と出力層に対しての総和とをほぼ同じになるように定めた。

### 3. 2 パターンの想起

パターンの想起に関しては、以下の手順により行なう。

- (1) 下位層をその初期パターンに固定しておき、モデルを動作させてその時の上位層の出力を計算する。
- (2) この上位層の状態によって、初期パターンの中にどのカテゴリに属するパターンが含まれているかを決定する。
- (3) 決定したカテゴリの状態に上位層を固定しておき、さらに下位層のユニットに対しては初期パターンから出発し、また、外部からパターンが与えつづけられている (3.7) 式のように設定してモデルを動作させる。

$$J_i = \varepsilon \cdot X_i, \quad (X_i \text{ は初期パターン、} \varepsilon \text{ は小さな定数})$$

- (4) 最終的に落ち着いた下位層の状態を想起パターンとする。

上記の手順により、想起されたパターンと記憶パターンとを計算機シミュレーションにより比較検討を行なった。

### 計算機シミュレーション [12]

下位層 256 ユニット・中間層 20 ユニット・上位層 5 ユニットの相互結合型神経回路モデルに 0～4 までの各 5 種類の数字パターンをカテゴリ学習させた (一対一学習)。そのうちの記憶した 2 種類のパターンが重なったパターンを下位層に入力し、同様に上位層にそのカテゴリに対応した信号を与えて、神経回路モデルを動作させパターンの分離能力について調べた。その結果を図 3. 2 に示す。階層型神経回路モデルの時と同様に、初期パターンに依存せずに記憶パターンが出力することができた。さらに、上記の神経回路モデルに 0～4 までの 5 つのパターンを 5 種類、計 25 パターンをカテゴリ学習させた。この場合、多対一対応になっており初期パターンによってそれぞれの記憶パターンが分離できることが期待できる。図 3. 3 に結果を示す。

### 3. 3 考察

階層型神経回路モデルの場合と同様に、一対一学習の場合、記憶パターンを出力することができた。また、多対一学習の場合、記憶パターンの平均パターンが想起されることがわかった。しかし、モデル自体が下位層でのユニットが2値に近い状態の方が安定であるため、階層型神経回路モデルの場合に比べると、記憶パターンが2値パターンであるので、想起パターンは記憶パターンに近いものを出力することができた。

一方、パターンの構造的な学習については、階層型神経回路モデルの場合と同様に、かなり記憶パターンセットに依存するものと思われる。

### 4. ネオ・コグニトロン [13,14]

パターン認識機能と連想記憶能力を兼ね備えた階層型神経回路モデルであり、大きな特徴として選択的注意機構を持つ。しかし、神経回路モデルとしては構造と学習法則ともに複雑であるので、あまり応用的な研究がなされていない。一方、このモデルによれば、想起パターンは記憶パターンの空間的構造を反映したパターンになる。また、選択的注意機構は最終層であるパターン認識層に最大値検出器を設け、最終層の状態を固定して下位層に向かって処理を行なうことで実現している。その時の処理に対しては、遠心性結合と呼ばれる下位層に向かう結合が存在する。

以下に、ネオ・コグニトロンの特徴的な機構と思われるものについて示す。

#### (1) 受容野

下位層への結合は、結合領域を持つ。つまり、結合領域の中でパターン識別を行なう要請によって、パターンの部分的な構造を学習する。

#### (2) 側抑制

抑制性素子は、ほぼ同じ抑制領域（受容野）を持つ細胞の中で一つだけが出力を出すようにゲイン調整される。このことは、各細胞がそれぞれ違った特徴検出器として学習されることに重要な要素である。

#### (3) 細胞出力のゲイン調整

一定の値以上の出力している細胞に対して反応を抑制し、一定値に満たない出力をしている細胞に対して反応を促進する機能がある。遠心性結合に信号が流れる時には、求心性結合とは違った抑制性素子が働きゲイン調整が行なわれる。

#### (4) しきい値素子が、正の信号に対して線形に出力

遠心性結合の強さは、それと対応する求心性結合に比例した値になっており、上位層のユニットから下位層の複数のユニットへの信号は、結合比によって伝達される。ゆえに、線形に出力することは、比例配分による信号伝達によって、回路が動作することを補う一つの要因である。

以上のように、今まで検討した神経回路モデルの結果から見て、ネオ・コグニトロンは、パターンの構造の学習が容易になるような構造になっている。

## 5. まとめ

階層型神経回路モデルと相互結合型神経回路モデルについて、連想記憶を行ない、パターン認識モデルに注意の機構を導入することにより、パターンの抽出を行なうことについて検討を行なった。一つのカテゴリに複数のパターンを学習させた場合、想起パターンに対して完全な記憶パターンに一致させるためには、それぞれのパターンが内部表現の違った形で学習される必要があること、また、パターンの空間的な構造を中間層である程度学習する必要があることがわかった。そして、パターンの空間的な構造を学習するためには、いろんな空間的な構造を持つパターンセットを学習する必要があること、神経回路モデル自身にある程度の構造を持たせる必要がある [15] こともわかった。

しかし、基本的にパターン認識に応用した神経回路モデルは、記憶パターンセットをカテゴリ分割するだけの必要な情報を特徴として取り出すので、多くのパターンを学習させたからといって空間的な構造を学習するとは限らない。また、このような空間的な特徴をとらえるような記憶パターンセットを選出するのは非常に困難である。

一方、神経回路モデル自身にある程度の構造を持たせるとしても、どのような構造がパターン認識や抽出に適しているのか未だに明確ではない。本文で、ある程度生理学的な事実に基づいたモデルとしてネオ・コグニトロンを検討した。その結果、パターンの抽出機能においては、その他のモデルよりもこの構造が適していることが明らかになった。

今後、神経回路モデルをそれぞれの分野に応用していくことを考えると、その課題に適した構造を持つ神経回路モデルを採用することが、その性能に大きく左右することが考えられる。特に、パターンの認識・抽出に関しては、人間の視覚領などの情報処理機構に即した形で神経回路モデルを構成すれば、モデルの能力が向上するのは言うまでもない。これらの生理学の進歩とその知見をいかにモデルにうまく取り入れることが、神経回路モデルの応用に重要な鍵になるであろう。

## 参考文献

- [1] D.E Rumelhart and J.L McClelland : "Parallel Distribute Processing", MIT Press (1986)
- [2] 仁木、諏訪 : 「誤差訂正学習とフィードバックを持つ自己連想記憶 : MAC」 信学全大秋季全国大会 D-21(1988)
- [3] 金道 : 「ニューラルネットワークに於ける入力再現」 信学全大春季全国大会 D-32(1990)
- [4] 村上、森、尾田 : 「階層型ニューラルネットワークを用いた重なりパターン分離の一検討」 信学全大春季全国大会 D-483(1990)
- [5] 山田、津雲 : 「多層ニューラルネットを用いたパターンの意識的想起処理方法」 信学全大春季全国大会 D-540(1989)
- [6] 山田、天満 : 「多層ニューラルネットによる意識的想起方法と想起学習の検討」 NC90-22 (1990)
- [7] Bart Kosko : "Bidirectional Associative Memories" : IEEE Trans. Sys. Man and Cyber. Vol.18 No.1 pp.49-60 (1988)
- [8] Bart Kosko : "Competitive Adaptive Bidirectional Associative Memories" : IJCNN Vol.II 759-766 (1988)
- [9] 坂上、横矢 : 「弛緩法と正則化」 情報処理 Vol.30 No.9 pp1047-1057 (1989,9)
- [10] 前田、川人、宇野、鈴木 : 「ヒト上肢多関節運動軌道を学習・生成する多層神経回路モデル」 信学技報 MBE87-133 (1988)
- [11] C.Peterson and E.Hartman : "Explorations of the Mean Field Theory Learning Algorithm" Neural Networks Vol.2 pp.475-494 (1989)
- [12] 村上、佐藤 : 「相互結合型神経回路モデルを用いた重なりパターン分離の一検討」 信学全大秋季全国大会 D-371(1990)
- [13] 福島 : 「視覚パターン認識における選択的注意機構の神経回路モデル」 信学論D Vol.J69-D No.6 (1986.6)
- [14] 福島 : 「階層構造をもつ連想記憶の神経回路モデル」 信学論A Vol.J66-A No.11 (1983.11) p.1130~1137
- [15] 森田、吉澤、中野 : 「連想記憶の力学と記憶の神経回路」 第4回生体・生理工学シンポジウム 1A2-3 pp43-46
- [16] T.Kohonen : "Associative Memory", Springer-Varlag (1977)

## 謝辞

日頃、ご討論して下さった ATR 視聴覚機構研究所の諸氏に深く感謝致します。特に川人主幹研究員・佐藤主任研究員・森吉弘氏・入江文平氏をおかれては多大なるご助言をいただいたことに深く感謝いたします。また、本研究の機会を与えて下さった淀川英司社長及び上野圭一視覚研究室長に感謝致します。

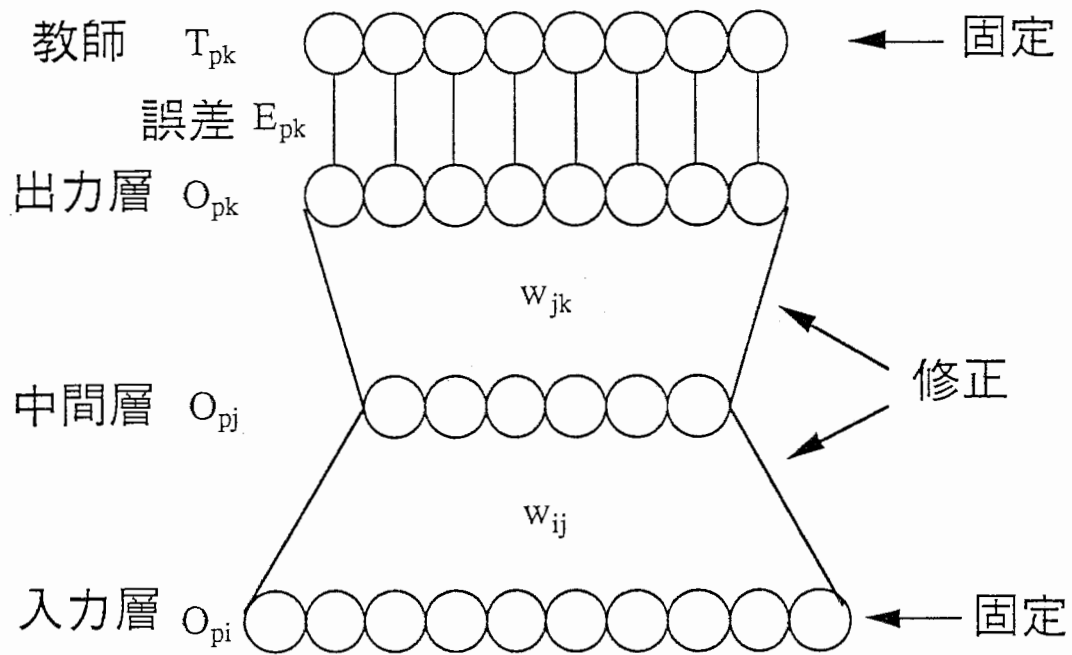


図2. 1 3層階層型神経回路モデル

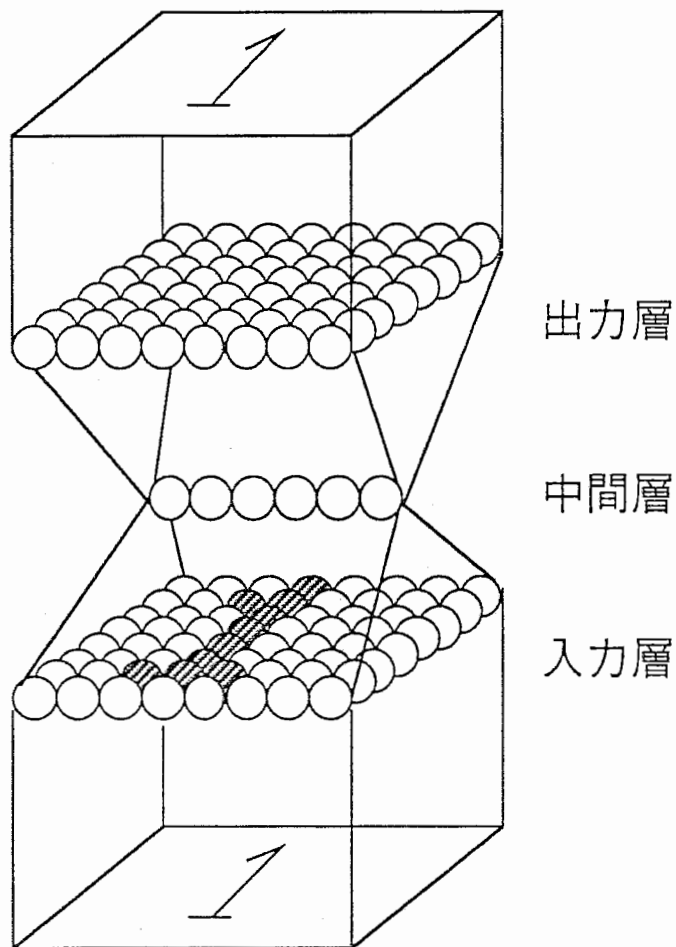


図2. 2 恒等写像型連想記憶



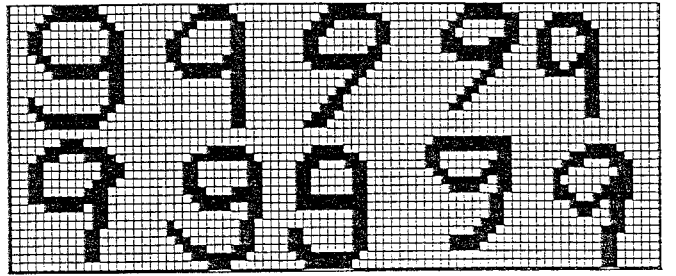
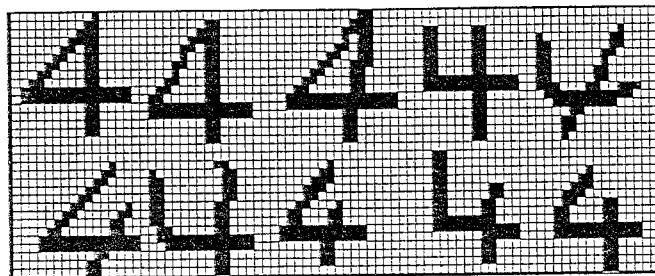
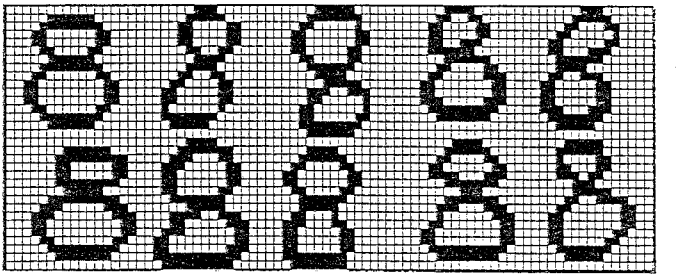
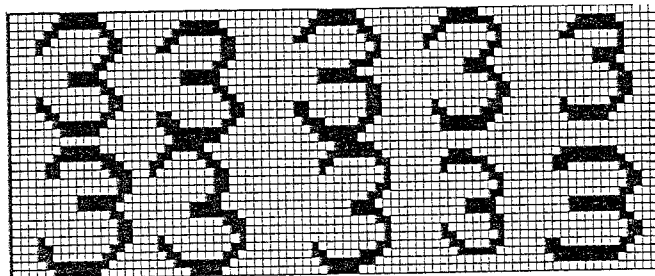
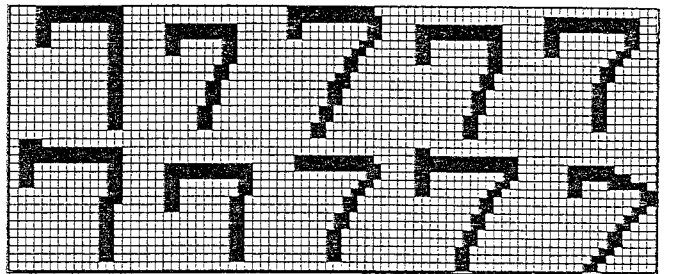
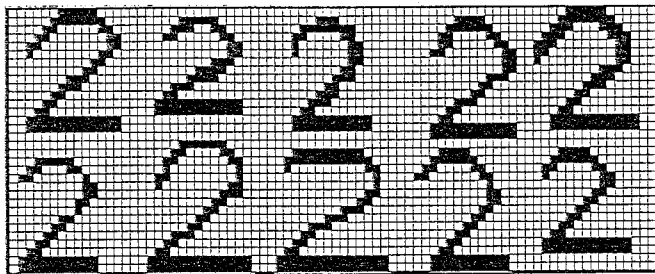
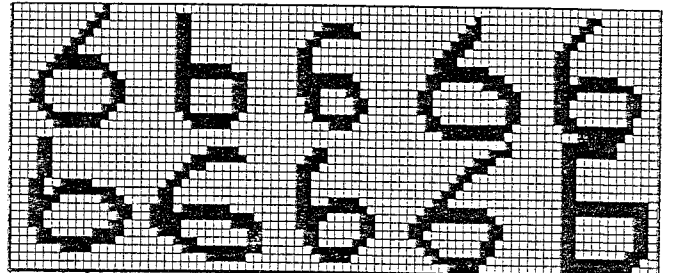
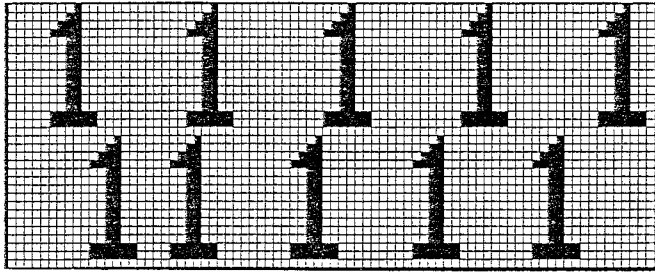
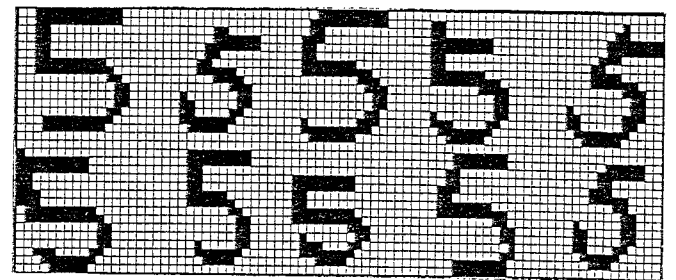
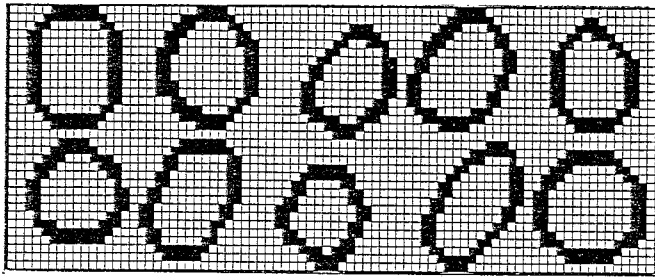


図2.3 シミュレーションで使った数字パターン

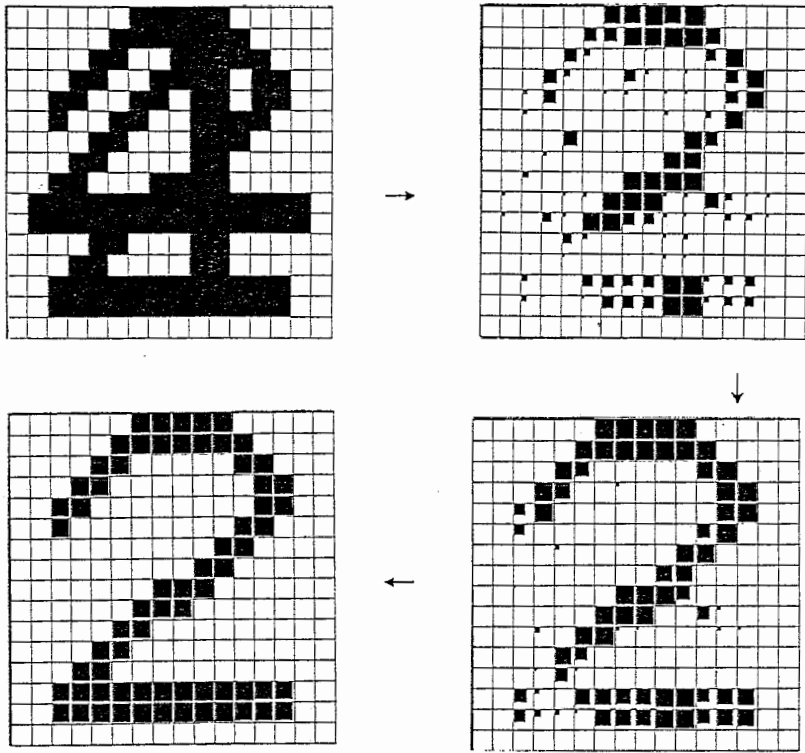


図 2. 4 正しく想起した例

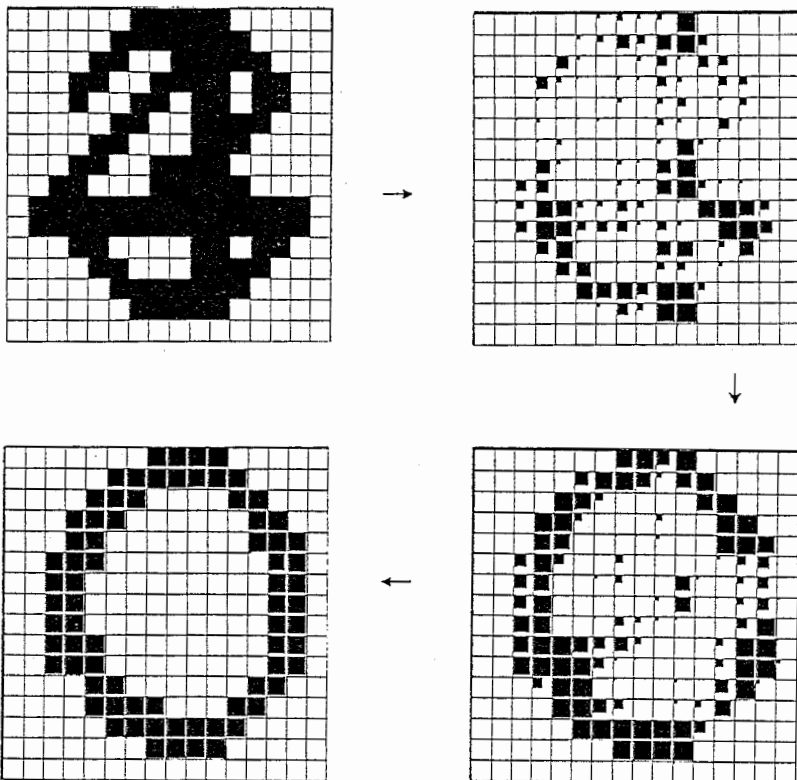


図 2. 5 誤って想起した例

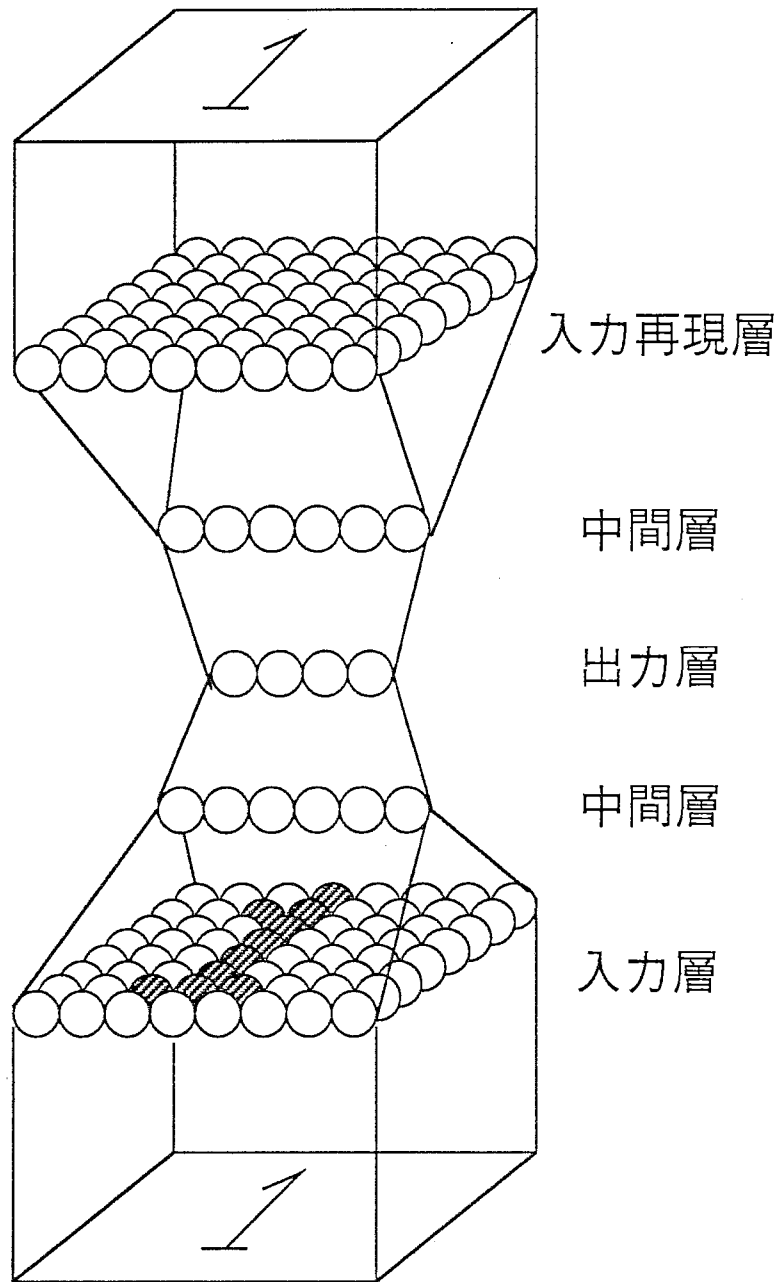


図 2. 6 5層階層型神経回路モデル

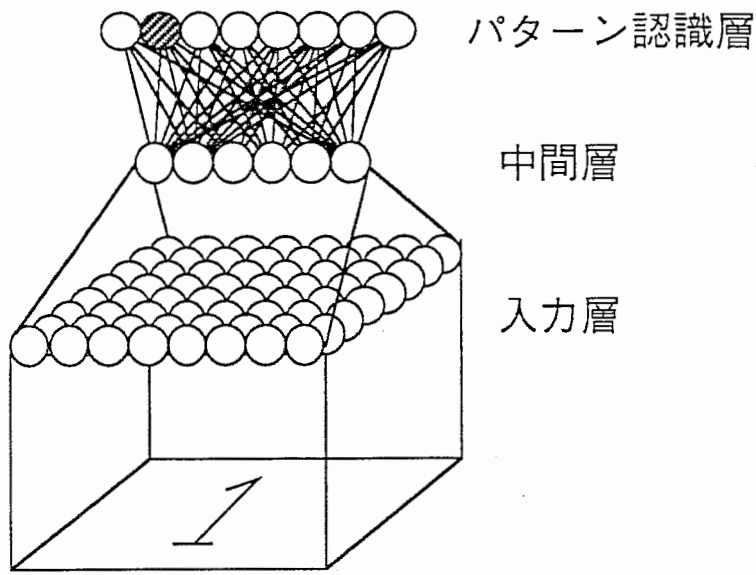


図2.7 パターン認識モデル

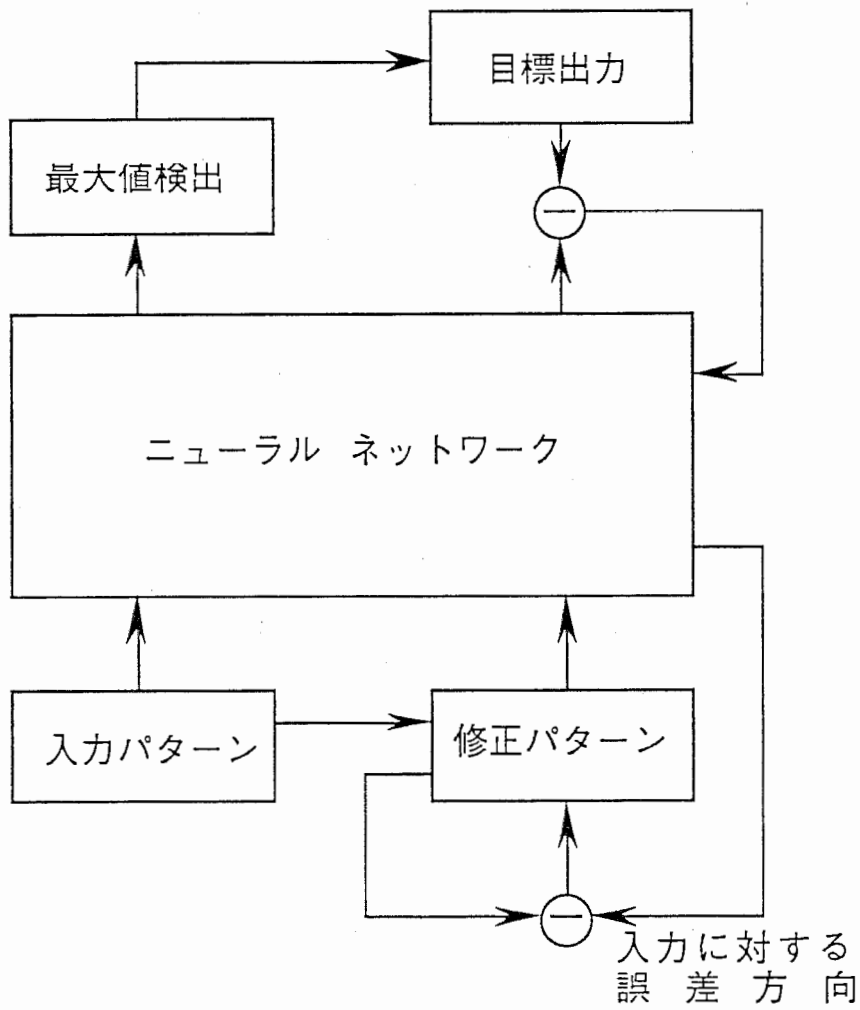
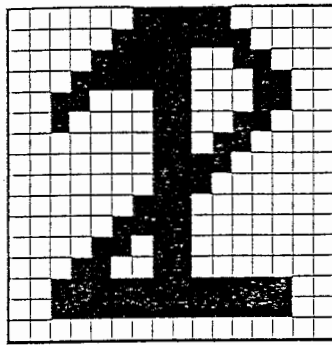
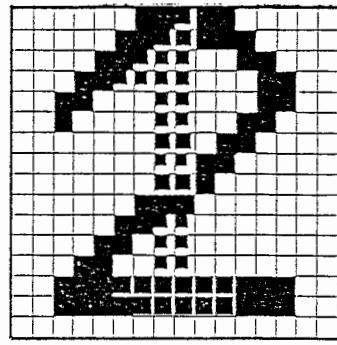


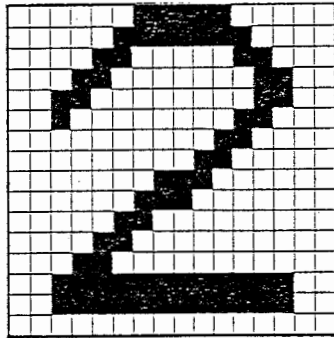
図2.8 想起処理の流れ



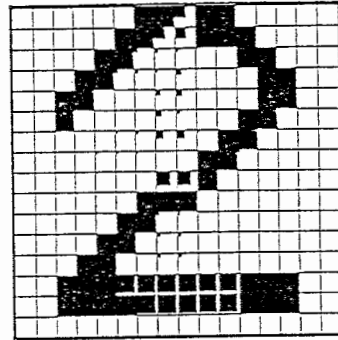
初期入力パターン



1000回修正  
E=0.015866



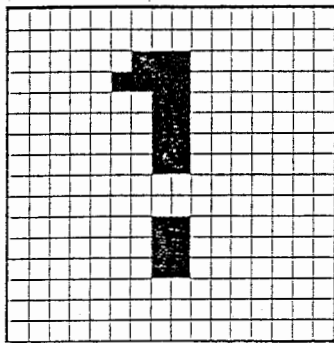
分離パターン



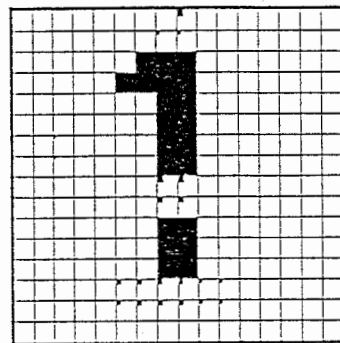
2000回修正  
E=0.005737

←  
しきい値  
処理

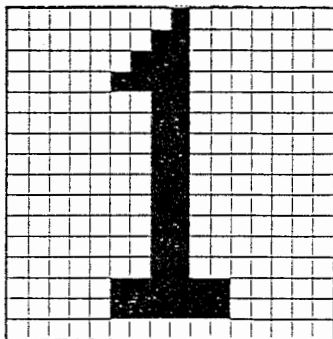
図2.9 (a) 想起パターン



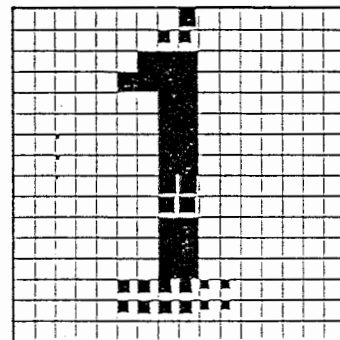
初期入力パターン



1000回修正  
E=0.010019



分離パターン



2000回修正  
E=0.005212

←  
しきい値  
処理

図2.9 (b) 差分後の想起パターン

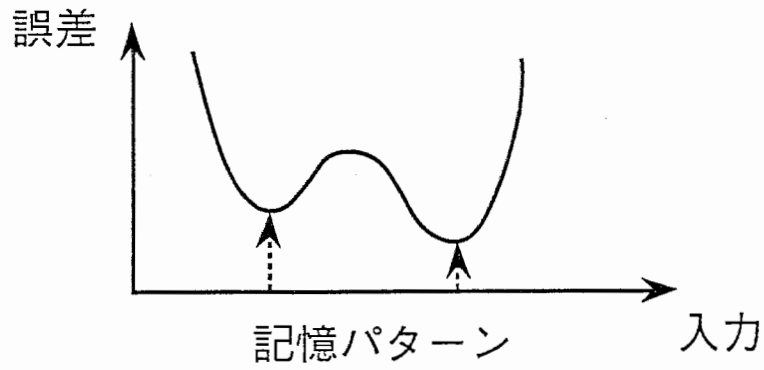
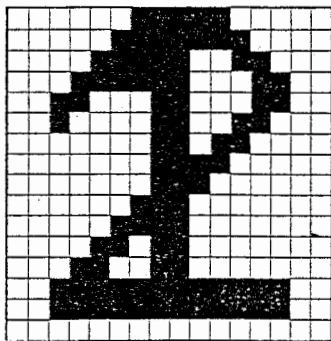
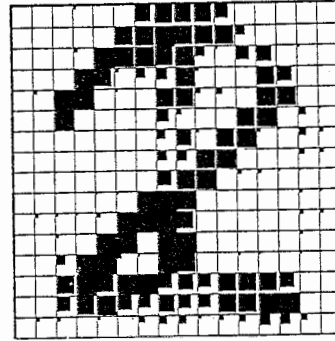


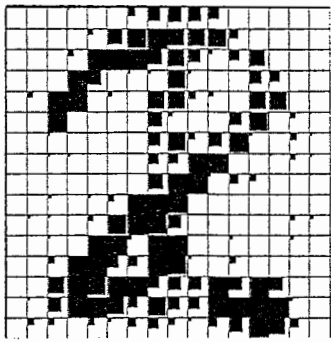
図2. 10 記憶パターンの状態



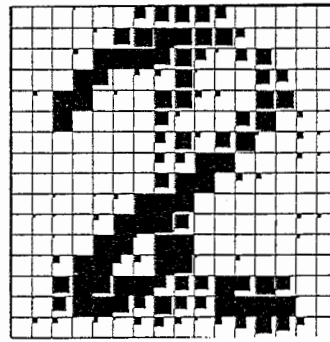
初期パターン



1 0 0 回変更 E = 0.002123

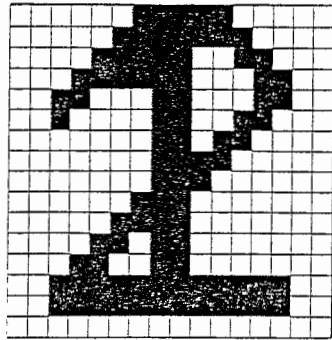


1 0 0 0 回変更 E = 0.001177

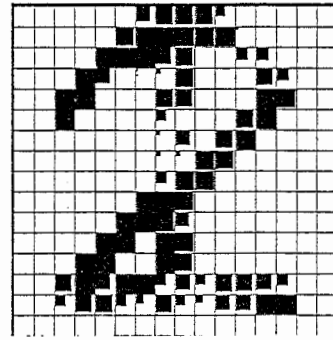


5 0 0 回変更 E = 0.001390

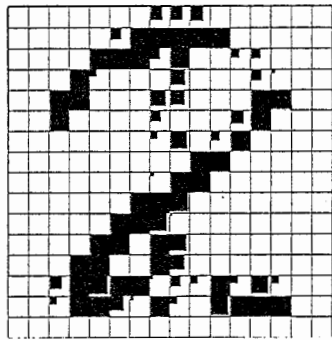
図2. 11 (a) 想起パターン (拘束条件なし)



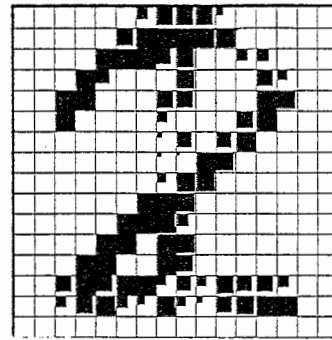
初期パターン



100回変更  $E = 0.003047$

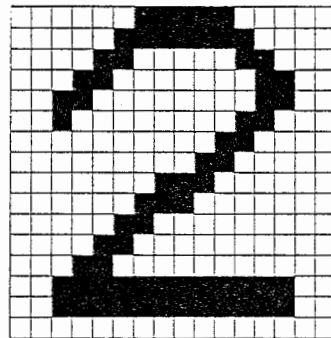


1000回変更  $E = 0.001798$

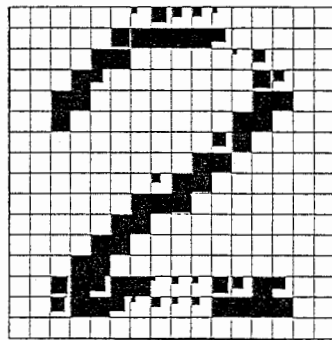


2000回変更  $E = 0.002523$

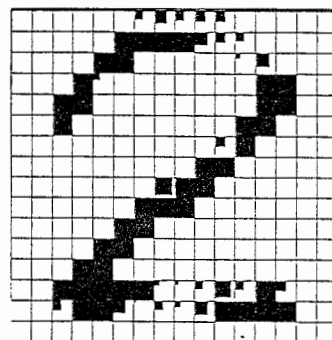
図2. 11 (b) 想起パターン (拘束条件あり)



記憶パターン



中間層20ユニット



中間層50ユニット

図2. 11 (c) 想起パターン (中間層の数の違い)

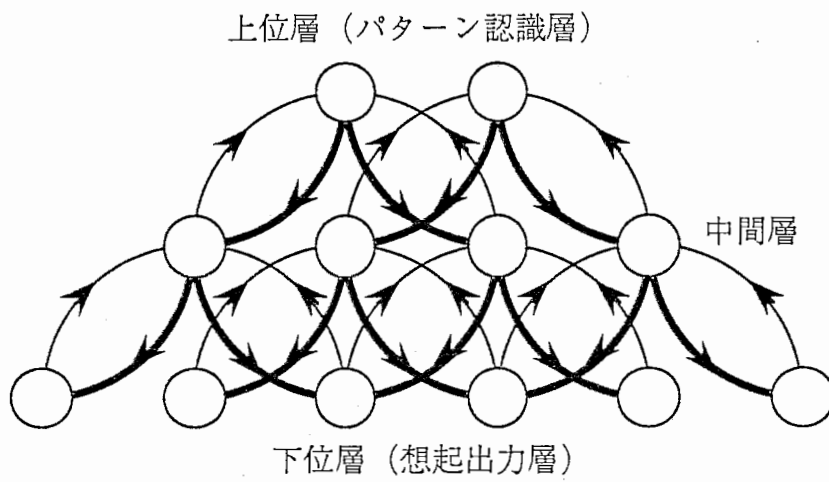
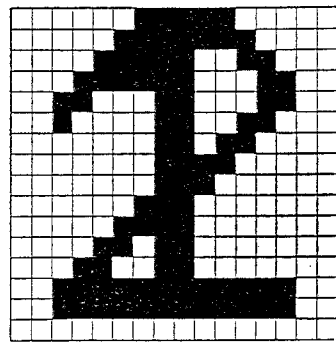
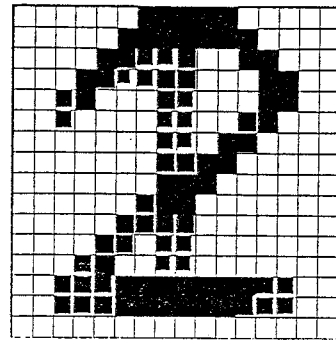


図3. 1 相互結合型神経回路モデル

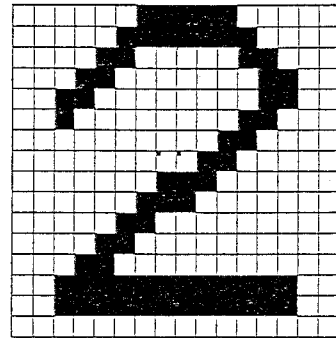




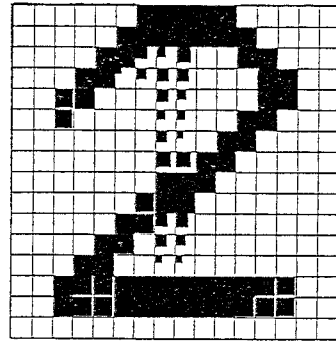
入力パターン



500回更新

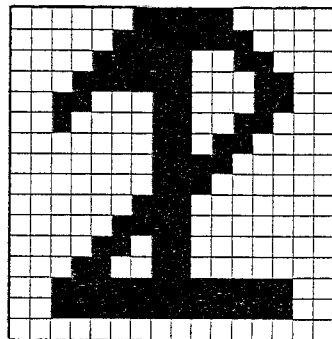


2000回更新

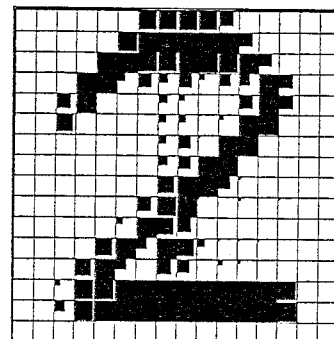


1000回更新

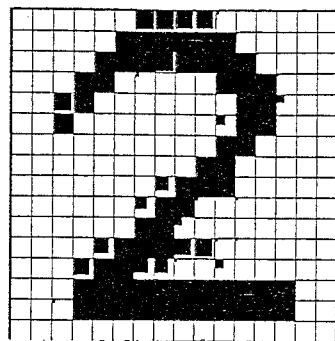
図3.2 想起パターン (一対一学習)



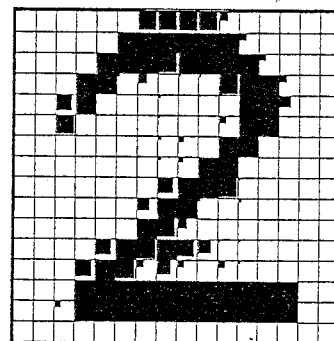
初期パターン



1000回変更



5000回変更



2000回変更

図3.3 想起パターン (多対一学習)