

ハイパーコラムモデルにおけるスパースコード学習

島田 敬士[†] 鶴田 直之^{††} 谷口倫一郎[†]

[†]九州大学大学院 システム情報科学府 〒816-8580 福岡県春日市春日公園 6-1
^{††}福岡大学 工学部 〒814-0180 福岡市城南区七隈 8-19-1
E-mail: †{atsushi,rin}@limu.is.kyushu-u.ac.jp, ††tsuruta@tl.fukuoka-u.ac.jp

あらまし ハイパーコラムモデルは自己組織化の競合学習を用いた階層型のニューラルネットワークである。ネオコグニトロン¹の *S* 層と *C* 層を階層型自己組織化マップで置き換えた構造をしており、物体の位置やサイズが変化する画像を認識することができる。従来のハイパーコラムモデルでは、各層において下位の層との結合のみが考慮されており、同層に配置された階層型自己組織化マップ間の結合については考慮されていない。そこで本論文では、階層型自己組織化マップ間の興奮性結合の学習法を提案する。ハイパーコラムモデルは、各自己組織化マップで活性化する勝利ニューロンの組み合わせを学習することができるため、認識フェーズでは、各階層型自己組織化マップから活性化するニューロンに制約を設けることができる。これにより、誤ったニューロンが活性化することを抑制することができるため、認識率の向上が見込まれる。我々は、本学習法をスパースコード学習と名づけた。

キーワード 画像認識, 人工神経回路網, ハイパーコラムモデル, スパースコーディング

Sparse Code Learning in Hyper-Column Model

Atsushi SHIMADA[†], Naoyuki TSURUTA^{††}, and Rin-ichiro TANIGUCHI[†]

[†] Department of Intelligent Systems, Kyushu University
Kasuga-koen 6-1, Kasuga, Fukuoka, 816-8580 Japan

^{††} Department of Electronics Engineering and Computer Science, Fukuoka University
Nanakuma 8-19-1, Jonan-ku, Fukuoka, 814-0180 Japan

E-mail: †{atsushi,rin}@limu.is.kyushu-u.ac.jp, ††tsuruta@tl.fukuoka-u.ac.jp

Abstract Hyper-Column Model (HCM) is a self-organized, competitive and hierarchical multilayer neural network. It is derived from the Neocognitron by replacing each *S* cell and *C* cell with a two layer Hierarchical Self-Organizing Map (HSOM). HCM can recognize images with variant object size, position, orientation and spatial resolution. Original HCM has connections only with its input layer or bottom layer and does not assume lateral connections among HSOMs. In this paper, we propose a new learning method with excitatory lateral connections. HCM can learn patterns of winner neurons which are activated in each HSOM by updating the excitatory connections, which improves the recognition accuracy. In recognition phase, HCM imposes constraints on the neurons which are activated in each HSOM by using the excitatory connections. We call this learning method “Sparse Code Learning” from the viewpoint that an entire activation pattern is described as multiple neurons.

Key words Image recognition, Artificial neural network, Hyper-Column Model, Sparse coding

1. はじめに

人工神経回路網は、入力データの次元数が比較的小さい文字認識やリモートセンシングにおける画素単位のクラスタリングなどで多く利用されてきた。近年では、顔画像認識をはじめとする一般の画像認識への人工神経回路網の応用が広がってきている。そこでの課題として次の3点が挙げられる。

【課題1】 一般の画像は特徴空間の次元数が高いが、多くの冗長性を含むので、画像空間に対して自然に得られる画像が分布する領域は小さい。この冗長性を取り除き、実際に画像が分布する領域を符号化することによって、入力データの次元数を減らす必要がある。

【課題2】 画像は、対象物の位置や照明条件など多くの独立した要因によって変化するので、実際に画像が分布する領域だけ

を符号化しても、その次元数は大きい。この次元数の本質的な増加に対して、システムの大きさが現実的な規模に収まるように、システムを構造化する必要がある。

【課題3】 認識対象の概念的な抽象度が高くなるので、画像空間におけるカテゴリ間の境界が複雑な形状になる。そのような場合、人工神経回路網では非常に多くの隠れニューロンを必要とするため、原理的には万能な人工神経回路網を用いたとしても、誤差逆伝播法など学習結果が結合係数の初期状態に大きく依存してしまう学習方式では安定した学習が行えない。従って、そのような場合でも安定した認識を可能にする学習方式を用いる必要がある。

鶴田らは、これらの課題を解決するために、Lampinen らの階層型自己組織化マップ [1] (Hierarchical Self-Organizing Map: 以下 HSOM) と福島らのネオコグニトロン [2], [3] (Neocognitron: 以下 NC) を統合したハイパーコラムモデル (Hyper-Column Model: 以下 HCM) を提案している [4], [5]。HCM の学習方式は Kohonen らの自己組織化マップ [6], [7] (Self-Organizing Map: 以下 SOM) を利用した教師なし学習であるため、結合係数の初期状態に依存することなく、HCM の構造上の利点を十分に引き出すことができる。また、HCM は、HSOM と NC の欠点を克服すると同時に、利点をそのまま継承できるので、一般の複雑な画像空間を極めて効果的に符号化することができ、HSOM や NC に比べて高い画像認識能力を持つ。

Tobely らは、HCM を用いた実時間での画像認識法を提案している [8]。高速化手法としてランダムマイズ法を取り入れ、ジェスチャ認識に応用している。また、我々は、HCM における教師あり学習法を提案している [9]。従来の入力信号および結合係数に対して連想信号と連想結合係数を付加することで教師あり学習が可能となり、認識能力が向上することが報告されている。オリジナルの HCM を含めてこれらのモデルに共通することは、HCM 内に配置された複数の HSOM は、個々で入力層との結合を持っているだけで HSOM 間には全く結合を持たないということである。そのため、認識フェーズでは個々の HSOM の受容野からの入力信号だけを参考にして活性化するニューロンが決定されてしまう。従って、いくつかの HSOM から活性化しているニューロンが誤りである場合、HCM は認識を誤ってしまう恐れがある。もちろん、HCM には多少の誤りであれば上位の特徴統合層でその誤差を吸収することができる。ここでの問題は、上位の特徴統合処理でも誤差を吸収できないほど多くのニューロンが誤って活性化してしまう場合である。

そこで本論文では、新たに HSOM 間に興奮性の結合を設け、その学習法を提案する。これらの興奮性結合を用いることで、HCM は各 HSOM から勝利ニューロンとして活性化するニューロンの組み合わせを学習することができるようになる。各 HSOM からひとつだけ活性化するニューロンの全体の組み合わせを学習するという観点から、提案する学習法を「スパースコード学習」と呼ぶことにする。認識フェーズでは、興奮結合により各 HSOM から活性化するニューロンに制約を設けることができるため、全体の活性化パターンを学習時のものに

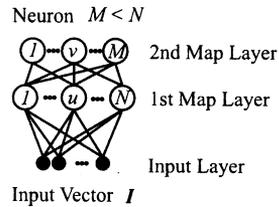


図1 HSOMの構造

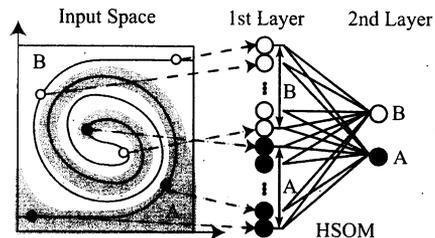


図2 HSOMが有効な問題例

近づけることができる。これにより、認識能力の向上が期待できる。

2. ハイパーコラムモデル (HCM)

2.1 階層型自己組織化マップ (HSOM) とネオコグニトロン (NC)

HSOM は、図3に示すように、SOM を二つ重ねた多層の人工神経回路網である。入力層の入力ベクトル I はマップ第1層のニューロン ($1 \dots u \dots N$) すべてと結合している。各ニューロン u は、入力ベクトルと同じ次元の結合係数ベクトル W_u を持ち、式 (1) を満足するニューロン c が活性化する。

$$\|I - W_c\| = \min_u (\|I - W_u\|) \quad (1)$$

学習時には、学習データが入力されるたびに、式 (1) により、最大活性化ニューロン c を求め、式 (2) により結合係数を更新する。

$$W_u(t+1) = W_u(t) + h_{cu}(I(t) - W_u(t)) \quad (2)$$

$$h_{cu} = \alpha(t) \cdot \exp\left(-\frac{\|r_c - r_u\|^2}{2\sigma^2(t)}\right) \quad (3)$$

SOM の特徴は、ニューロン間にも距離が定義されていることである。式 (3) の $\|r_c - r_u\|$ は、ニューロン c とニューロン u の距離である。学習時の繰り返し処理の各フェーズでは、最大活性化ニューロンの結合係数だけを更新するのではなく、最大活性化ニューロンの近傍のニューロンも同様の学習を行う。 $\sigma^2(t)$ は、最大活性化ニューロンと同時に学習する近傍範囲を制御するためのパラメータである。 $\alpha(t)$ は学習レートで、学習の過程で徐々に 0 に近づく。

マップ第2層のニューロン ($1 \dots v \dots M$) は、マップ第1層と結合を持ち、マップ第1層のニューロンの「添え字」がマップ

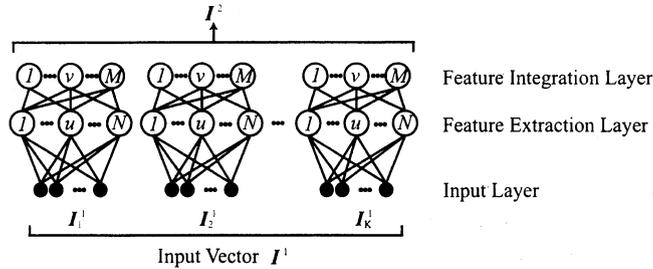


図3 HCM内に配置されたHSOMネットワーク

第2層の入力となる。マップ第2層のニューロンの数 M はマップ第1層のニューロンの数 N に比べ非常に小さいものとなっているため、マップ第1層で抽出された特徴が少数のニューロンにより再量子化される。これにより、HSOMは、図2に示すような複雑な識別問題を容易に識別することができる。

NCは、HSOMに類似したネットワークをピラミッド型に結合した構造になっている。HSOMのマップ第1層がNCの特徴抽出層 U_S に、マップ第2層がNCの位置ずれ吸収層 U_C に相当する。ひとつのHSOMに相当した部分は、特徴抽出層 U_S に複数のマップを持ち、 U_S 層の入力領域は、それぞれ独立しているのではなく、1画素あるいは1ニューロンずつずれるように重なりを持った構造になっている。位置ずれに不変な認識を実現するために、同一層のマップはその位置に関係なく同じである。

2.2 ハイパーストックモデル (HCM)

HCMは、HSOMとNCを統合したモデルで、その構造は、図1に示すようにHSOMのネットワークをNCと同様にピラミッド型に結合したものである。一つのHSOMの第1層(特徴抽出層)の入力領域もNCと同様に、図1の I_1^0 から I_k^0 のようにそれぞれ独立しているのではなく、1画素あるいは1ニューロンずつずれるように重なりを持った構造になっている。さらに、位置に不変な認識を実現するため、同一層のマップはその位置に関係なく同じである。HCMとNCの異なる点は、HSOMの第2層(HCMでは特徴統合層と呼ばれる)で、マップの再量子化と位置ずれの吸収を同時に行っていることである。これにより、NCよりも更に大幅な次元縮小が可能となり、大きな変形に対して不変な認識が可能である。

この構造により、HCMはHSOMとNCの欠点を克服すると同時に、利点をすべて継承することができるので、HCMによる画像認識は「対象物の位置と大きさ、変形、照明条件などの連続的な変化により画像空間中に張られる画像データの分布領域を効果的に階層的に量子化し、量子化によって得られた標本点を用いて最近傍法により認識する手法」と位置づけられている。

3. スパースコード学習

HCM内に配置されたHSOMは、各々が入力層との結合しか持っていない。そのため、認識フェーズでは個々のHSOM

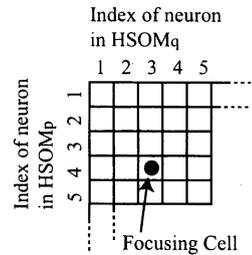


図4 HSOM_p, HSOM_q 間のニューロンの結合テーブル

の受容野からの入力信号だけを参考にして活性化するニューロンが決定されてしまう。従って、いくつかのHSOMから活性化しているニューロンが誤りである場合、HCMは認識を誤ってしまう恐れがある。もちろん、HCMには多少の誤りであれば上位の特徴統合層でその誤差を吸収することができる。ここでの問題は、上位の特徴統合処理でも誤差を吸収できないほど多くのニューロンが誤って活性化してしまう場合である。そこで、我々はこの問題を解決する新たな学習法を提案する。各HSOM内の全ニューロンと他のHSOM内のニューロン間に新たに興奮性の結合を設けることで、これまでのように入力層との結合のみではなくHSOM間の結合も考える。学習フェーズでは、HCMはこれらの興奮性結合を用いて各HSOMから活性化するニューロンの組み合わせを学習する。あるHSOMからひとつのニューロンが活性化したときにその他のHSOMからはどのようなニューロンが活性化したかという全体的な活性化パターンを記憶することができる。認識フェーズでは、興奮性結合を用いて各HSOMから活性化するニューロンに制約を設けることで、HSOM全体の活性化パターンを学習時の活性化パターンに近づけることができるようになる。このように各HSOMからひとつだけ活性化するニューロンの全体の組み合わせを学習するという観点から、提案する学習法を「スパースコード学習」と呼ぶことにする。以下では、スパースコード学習の設計、学習法、認識法について述べる。

3.1 設計

HCMに配置されたHSOMの数を H とする。このとき、新たに作成される興奮性結合の数 M は以下の式で表される。

$$M = {}_H C_2 \times N^2 = \frac{H(H-1)}{2} \times N^2 \quad (4)$$

ここで、 N は HSOM 内のニューロンの数である。例えば、 $H = 400$ で各 HSOM 内に 100 個のニューロンが存在する場合、新たに作成される興奮性結合の数は $M \approx 8 \times 10^8$ となる。同条件下での入力層との結合数は約 1×10^6 であるため、 8×10^8 という結合数は非常に莫大なものである。そこで、結合数を削減することを考える。上記の方法は、ある HSOM 内のニューロンに対して、その他の全ての HSOM 内のニューロンとの間に結合を設けている。これを小グループに分割された HSOM 内で考えると結合数を削減することができる。実際に小グループ数を G 、小グループ内の HSOM の数を H' とすると、興奮性結合の数 M は以下の式で表される。

$$M = {}_{H'} C_2 \times G \times N^2 = \frac{H'(H'-1)}{2} \times G \times N^2 \quad (5)$$

先ほどの例では、グループ数 G を 16 としたとき、 $M \approx 5 \times 10^7$ となり約 94% の結合数を削減することができる。

3.2 学習法

ここでは、スパースコード学習法について述べる。二つの HSOM (HSOM_p and HSOM_q ($1 \cdots p, q \cdots H'$ and $p < q$)) 間の結合を図 4 の結合テーブルを用いて表すことにする。例えば、同図中の黒丸が描かれたセルは、 HSOM_p の 4 番目のニューロンと HSOM_q の 3 番目のニューロン間の結合を表すことになる。以下、 HSOM_p の i 番目のニューロンと HSOM_q の j 番目のニューロン間の結合を $W_{i,j}^{p,q}$ と書き表すことにする。

全 HSOM で従来の自己組織化の学習終了後、各 HSOM で学習データに対する勝利ニューロンが活性化される。学習データ k に対して HSOM_p で活性化したニューロンを w_p^k 、 HSOM_q で活性化したニューロンを w_q^k とするとき、興奮性結合 $W_{w_p^k, w_q^k}^{p,q}$ の結合係数を更新する。このとき活性化ニューロン w_p^k 、 w_q^k の近傍の興奮性結合の結合係数も以下の式で更新する。

$$W_{w_p^k, w_q^k}^{p,q} = W_{w_p^k, w_q^k}^{p,q} + \beta \cdot \exp\left(-\frac{D^2}{2\sigma^2(t)}\right) \quad (6)$$

ここで、 D は w_p^k と w_q^k からの距離である。以下に学習法の疑似コードを示す。学習前に全ての $W_{i,j}^{p,q}$ は 0 に初期化されているものとする。また、 K は学習データの数を表す。

```

1: for  $g = 1$  to  $K$  do
2:   for  $k = 1$  to  $G$  do
3:     for all  $p, q$  such that  $p < q$  and  $p, q < H'$  do
4:       Weights are Updated by Eq. (6)
5:     end for
6:   end for
7: end for

```

上記学習後、全ての興奮性結合の結合係数はシグモイド関数により最大値 1、最小値 0 に正規化される。

$$W_{i,j}^{p,q} = \frac{1}{1 + \exp(\theta - W_{i,j}^{p,q})} \quad (7)$$

ここで、 θ は活性化度を決定するための閾値である。 θ が大きくなるにつれて、高い活性化度を持つ興奮性結合の数は減少する。

3.3 認識法

ここでは、スパースコード学習結果を用いた認識法について述べる。テストデータ k に対して、各 HSOM から活性化したニューロンはあらかじめ求まっているものとする。このとき、小グループ内でテストデータ k に対する活性化度 S_k を以下の式で定義する。

$$S_k = \frac{\sum_{p,q(p<q)}^{H'} W_{w_p^k, w_q^k}^{p,q}}{M} \quad (8)$$

活性化度 S_k が大きいほど、小グループ内で活性化しているニューロンのパターンは学習時のものに似ていることになる。活性化度 S_k が小さい場合は、 $W_{w_p^k, w_q^k}^{p,q}$ よりも高い活性化度を持つ興奮性結合を探す。これは、図 4 の注目セルをシフトさせることと等価である。高い活性化度を持つ興奮性結合が見つかった場合、HCM は活性化しているニューロンのラベルを強制的に書き換える。例えば、図 4 で注目セルをひとつ右にシフトさせた場合、活性化ニューロンのラベルは 3 から 4 に変更される。以下に認識法の疑似コードを示す。

```

1: for  $g = 1$  to  $K$  do
2:   for  $k = 1$  to  $G$  do
3:     Activation Value  $S_k$  is Calculated by Eq. (8)
4:     while  $S_k < TH_s$  and  $Area < TH_a$  do
5:       Search Neighbor Connection
6:       if Found Larger Activation Value then
7:         Change Index of Winner Neuron
8:         Recalculate  $S_k$ 
9:       end if
10:       $Area \leftarrow Area + offset$ 
11:    end while
12:  end for
13: end for

```

ここで、 $Area$ は探索範囲、 TH_s は活性化度 S_k に対する閾値、 TH_a は $Area$ に対する閾値である。

4. 実験結果

4.1 準備

まず、HCM に学習データを学習させた。学習データは図 5 に示す 10 カテゴリから構成される人の手形状画像であり、各カテゴリにつき 10 枚ずつの画像を用意した。画像サイズは 100×100 ピクセルで、256 階調の濃淡画像とした。HCM の構造は、図 6(a) に示すように 3 層とし、各層に配置された HSOM および HSOM 内のニューロンは表 1 のように設定した。HSOM の特徴抽出層内のニューロンは図 6(b) に示すように、環状に配置した。また、第 3 層のネットワークについては特徴統合処理が不要なので HSOM の代わりに SOM を選択した。さらに HCM 第 1 層においては、各 HSOM は局所的な部分画像と結合を持つようにした。我々は、HCM 第 1 層内の HSOM の特徴抽出層にスパースコード学習を適用した。HSOM を分割する諸グループの数 G は 16 とした。従って、各グループは縦横 5×5 個の HSOM 間の興奮性結合を学習することになる。こ

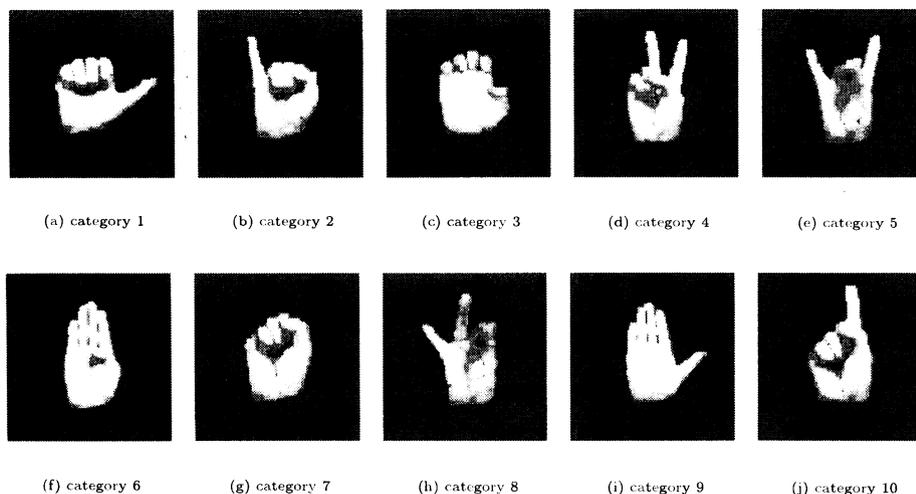


図5 実験で用いた手形状画像

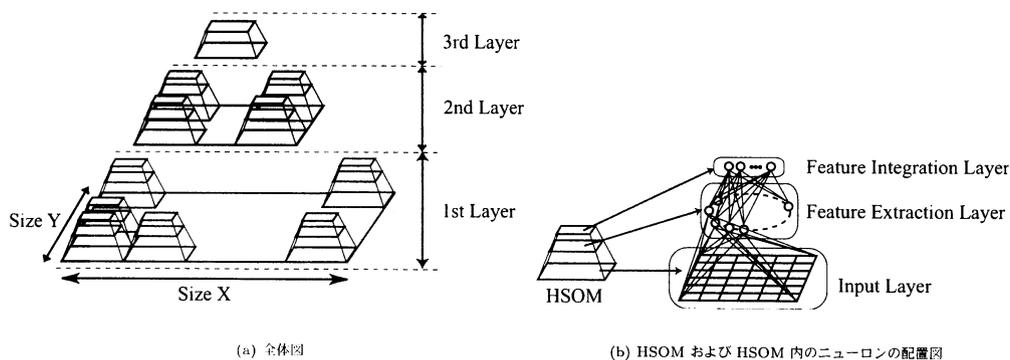


図6 実験で用いた HCM の構造図および HSOM の詳細図

の層は HCM 内の最下位層であるため、この層で活性化されるニューロンの誤りが減少すれば上位層での誤りを防ぐことができると考えた。

次に、各カテゴリ 500 枚、計 5000 枚のテスト画像を準備した。HCM はスパースコード学習結果を使わずに約 93% の画像を正しく認識することができた。参考に、一般的なパターン認識手法として知られている K-最近傍法 (K-Nearest Neighbor: K-NN) および多層パーセプトロン (Multi-Layer Perceptron: MLP) を用いて同テスト画像を認識させたところ表 2 のような結果が得られた。HCM の認識率は他の手法よりも高いことが確認できた。我々は、HCM が正しく認識することができなかったテスト画像を 300 枚取り出して、それらをスパースコード学習結果を用いた認識法の検証画像とした。

4.2 興奮性結合を用いた場合の認識精度

上記の 300 枚のテスト画像に対し、3.3 節で述べたスパース

コード学習結果を用いた認識手法を適用した。図 7(a) は興奮性結合を用いなかったときのあるテスト画像に対する各 HSOM からの活性化ニューロンのラベルを視覚化したものである。明るい画素ほど活性化したニューロンのラベルが大きいことを表すようにした。同図中に、部分画像に対する HSOM からの活性化ニューロンのラベルを表に表した。同様に、図 7(b) は興奮性結合を用いたときの活性化ニューロンを視覚化したものである。活性化ニューロンのラベルのうちいくつかはそのラベルが変更されていることが確認できる。実際に HCM は興奮性結合を用いることで、このテスト画像を正しく認識することができた。最終的に 300 枚のテスト画像のうち 147 枚の画像が正しく認識できるようになった。

5. おわりに

本論文ではハイパーコラムモデル (Hyper-Column Model :

表 1 The details of each layer of HCM.

		HSOM		Input Size	Neurons
3rd Layer (Top)	Size Y	1	FIL	-	-
	Size X	1	FEL	8 × 8	100
2nd Layer	Size Y	8	FIL	-	10
	Size X	8	FEL	5 × 5	100
1st Layer (Bottom)	Size Y	20	FIL	-	5
	Size X	20	FEL	5 × 5	100

FEL:: Feature Extraction Layer

FIL:: Feature Integration Layer

表 2 各認識手法における認識率の比較

手法	HCM	K-NN	MLP
認識率	93.3%	84.3%	77.5%

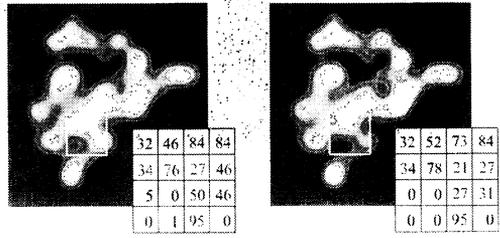
HCM)におけるスパースコード学習法を提案した。HCM内のHSOM間に興奮性結合を導入することでHCMは学習フェーズで学習データに対して各HSOMから活性化されるニューロンの組み合わせを学習することができるようになった。認識フェーズでは興奮性結合により各HSOMで活性化されるニューロンに制約を設けることで学習時と似た活性化パターンを再現しやすい状況を作り出すことができるようになった。実験では、従来のHCMが正しく認識できなかったテスト画像を興奮性結合を用いることで約50%の画像を正しく認識することができるようになった。この結果からも興奮性結合を用いることは非常に効果的であるといえる。しかしながら一方で、まだ50%は依然として誤認識されたままである。今回の実験では、スパースコード学習をあらかじめ全HSOMを小グループに分割したグループ内で適用した。そのため、認識フェーズでは小グループ内での各HSOMからの活性化ニューロンの組み合わせは学習時のものに近づけることができるが、グループ全体の活性化パターンが学習時のものに近づいたとは限らない。これはグループ間に結合を持たせなかったことが原因のひとつに挙げられるが、結合数を削減するためにグループ分割したために、グループ間結合を考慮するとなると再び同問題が生じると思われる。そこで今後は、スパースコード学習をHCM第1層の特徴統合層に適用することを考えている。この層はHSOMの数は特徴抽出層と同じであるが、HSOM内のニューロンの数が特徴抽出層と比べて非常に少ないので小グループに分割することなく、層全体にスパースコード学習を適用することが可能である。

謝 辞

本研究の一部は、21世紀COEプログラム「システム情報科学での社会基盤システム形成」の補助を受けた。

文 献

- [1] J. Lampinen and E. Oja. "Clustering properties of hierarchical self-organizing maps". *Journal of Mathematical Imaging and Vision*, Vol. 2, No. 3, pp. 261-272, 1992.
- [2] 福島邦彦. "位置ずれに影響されないパターン認識機構の神経回路モデル-ネオコグニトロン". 信学論(A), Vol. J62-A, No. 10, pp. 658-665, 1979.
- [3] K. Fukushima. "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaf-



(a) 興奮性結合なし

(b) 興奮性結合あり

図 7 活性化ニューロンの視覚化

ected by shift in position". *Biological Cybernetics*, Vol. 36, No. 4, pp. 193-202, 1980.

- [4] 鶴田直之, 谷口倫一郎, 雨宮真人. "ハイバコラムモデル: 階層化自己組織マップとネオコグニトロンを統合した画像認識モデル". 電子情報通信学会論文誌 J81-D-II/10, pp. 2288-2300, 1998.
- [5] Naoyuki Tsuruta, Rin-ichiro Taniguchi, and Makoto Amamiya. "Hypercolumn Model: A Combination Model of Hierarchical Self-Organizing Maps and Neocognitron for Image Recognition". *Systems and Computers in Japan*, Vol. 31, No. 2, pp. 49-61, 2000.
- [6] T. Kohonen. "Self-Organization and Associative Memory". Springer-Verlag, 1989.
- [7] T. Kohonen. "Self-Organizing Maps". *Springer Series in Information Science*, 1995.
- [8] Tarek El. Tobely, Naoyuki Tsuruta, Yuichiro Yoshiki, and Makoto Amamiya. "A randomized model of the hypercolumn neural network for gesture recognition". *the International Journal of Computers, Systems and Signal*, Vol. 3, No. 1, pp. 14-28, 2002.
- [9] Atsushi Shimada, Naoyuki Tsuruta, and Rin-ichiro Taniguchi. "Supervised Learning in Hyper-Column Model". *5th Workshop On Self-Organizing Maps*, pp. 481-488, 2005.