九州大学学術情報リポジトリ
Kyushu University Institutional Repository

ハイパコラムモデル : 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

鶴田, 直之
九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎
九州大学システム情報科学研究院知能システム学部門

雨宮, 真人
九州大学システム情報科学研究院知能システム学部門

https://hdl.handle.net/2324/5786

出版情報：電子情報通信学会論文誌. J81-D-II (10), pp.2288-2300, 1998-10. 電子情報通信学会
バージョン：
権利関係：

論　文

ハイパコラムモデル：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

鶴田　直之[†]　　谷口倫一郎[††]　　雨宮　真人[††]

Hypercolumn Model: A Combination Model of Hierarchical Self-Organizing Maps and Neocognitron for Image Recognition

Naoyuki TSURUTA[†], Rin-ichiro TANIGUCHI[††], and Makoto AMAMIYA[††]

あらまし　本論文では，一般の画像認識へ応用可能な画像認識モデルとしての人工神経回路網：ハイパコラムモデルを提案する．ハイパコラムモデルは，階層化自己組織マップとネオコグニトロンとを統合したモデルであり，それぞれの欠点を補完しつつ両者の利点を継承する．階層化自己組織マップは，自己組織マップを用いて入力空間を低次元の特徴空間に非線型にマッピングする．これにより従来の線形分離原理を用いたパターン認識手法や誤差逆伝搬法を用いた多層パーセプトロンでは認識が困難であった，データ空間内で複雑な形状のカテゴリー領域を形成する問題を容易に解く類別モデルとして注目されている．一方，ネオコグニトロンは，局所的な特徴抽出と，抽出した特徴ごとに解像度を減少させる位置ずれ吸収を階層的に行う構造になっており，次元数の高い入力データを認識するのに適した構造になっている．ハイパコラムモデルは，上で述べた階層化自己組織マップの特性とネオコグニトロンの構造との統合により，入力データの次元数が高く，データ空間内で複雑な形状のカテゴリー領域を形成する一般の画像の認識への応用が可能である．

キーワード　画像認識，人工神経回路網，ハイパコラムモデル，階層化自己組織マップ，ネオコグニトロン

1. まえがき

人工神経回路網は，入力データの次元数が比較的少な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で多く利用されてきた．これらの問題に対しては，人工神経回路網は，従来の線形手法である主成分分析法やノンパラメトリックな手法である k-nearest neighbor 法よりも優れているとする報告[21]がある．更に近年では，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まっている．そこでの課題は，次の3点である．以下，画像を画素数（画像サイズ）の次元をもつベクトルととらえ，画像が張る空間を画像空間と呼ぶとにする．

課題1　一般の画像は次元数が高いが，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，画像が分布する領域だけを符号化したとしても，その次元数は（画像サイズよりは小さいものの）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ー間の境界が複雑な形状になる．そのような場合，ニューラルネットワークでは非常に多くの隠れニューロンを必要とするため，原理的には万能なニューロルネットワークを用いたとしても，誤差逆伝搬法など学習結果が結合係数の初期状態に強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．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．

Poggio ら[16]は，顔領域の検出へ多層パーセプトロンを応用しているが，画像を直接多層パーセプトロ

† 福岡大学工学部電子情報工学科，福岡市
Department of Electronics Engineering and Computer Science, Fukuoka University, Fukuoka-shi, 814-0180 Japan
†† 九州大学大学院システム情報科学研究科，春日市
Graduate School of Information Science and Electrical Engineering, Kyushu University, Kasuga-shi, 816-8580 Japan

ンへ入力するのではなく，固有顔と入力画像のユークリッド距離および Mahalanobis 距離を計算し，これらを入力とすることで，課題1を解決しようとしている．また，課題2を解決するために，画像サイズおよび輝度の正規化を必要としている．以上二つの前処理によって大胆に入力データの次元数を低減しているため，最後の識別の段階では多層パーセプトロンを用いた場合と，単層パーセプトロンを用いた場合で差異が認められていない．従って，人物の識別など問題が複雑になった場合に，課題3に対応可能かどうか明らかではない．

また，Rosenblum ら [14] は，オプティカルフローをもとに，顔の特徴的な部分の動きを検出し，これを動径基底関数ネットワークに入力することによって，顔の表情認識を実現した．この場合も，課題1および課題2への対応として，顔のどの部分のオプティカルフローを用いるかはアドホックに決められており，一般性に欠ける手法となっている．また，文献 [14] では，課題3に対して，多層パーセプトロンよりは動径基底関数ネットワークが優れていることが示されているが，ここでも学習結果が結合係数の初期状態に強く依存してしまう学習方式が利用されており，課題3が解決されているとは言えない．動径基底関数ネットワークにおける課題3の解決法としては，次に述べる自己組織マップを用いた結合係数の初期状態に依存しない学習方法 [1] が良好な結果を得ている．

さて，生体の視覚大脳皮質には，ハイパコラムと呼ばれる特殊な構造が存在することが知られている．このハイパコラムをいくつかの側面からモデル化した人工神経回路網に，Kohonen らの自己組織マップ [8]〜[10]（Self-Organizing Map：以下 SOM と略す）と福島らのネオコグニトロン [2], [3], [6]（Neocognitron：以下 NC と略す）がある．

Lampinen [11] らは，課題1と課題3に対して，二つの SOM を重ねた階層化自己組織マップ（Hierarchical Self-Organizing Map：以下 HSOM と略す）が，誤差逆伝搬学習法を用いた多層パーセプトロンが苦手とするクラスのパターン認識問題を容易に解くことを示した．但し，単純な HSOM では，課題2が解決できないという問題が残る．実際，Lampinen らの事例では，課題2のための前処理として HSOM を局所的に適用し，最終的には HSOM の出力の主成分分析によって顔を識別している．

一方，NC は，局所的な特徴を抽出する層（Us 層）と，抽出した特徴ごとに解像度を下げる（位置ずれを吸収する）層（Uc 層）とを交互に積み重ねることによって，入力パターンの次元数を徐々に減らしながらパターン全体の位置ずれと拡大縮小に不変な認識を可能にしたパターン認識器である．このような分割統治型の構造は，課題2に対して効果的なものである．但し，NC の学習で用いられている競合学習法は学習能力が不十分であるため，各層を学習するごとに，それぞれに適したデータを人手で作成しなければならない．この段階で各モジュールがどのような特徴抽出を行うべきかを学習データ作成者が決めていることになり，課題1や課題3が解決できていない．

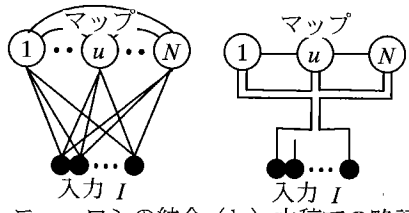
以上に述べた HSOM と NC は，互いの欠点を補完し合う関係にある．そこで，本論文では，NC の Us 層を HSOM の第1層で，また NC の Uc 層を HSOM の第2層で置き換えたモデル：ハイパコラムモデル（Hyper-Colunm Model：以下 HCM と略す）とその学習方式を提案する．学習方式は非常に単純でありながら，結合係数の初期状態に依存しない方式であり，HCM の構造上の利点を十分に引き出すことができる強力なものとなっている．また，HCM は，HSOM と NC の欠点を克服すると同時に，利点をそのまま継承することにより，一般の複雑な画像空間を極めて効果的に符号化することができ，HSOM や NC に比べて，非常に高い画像認識能力をもつ．

以下では，まず **2.** で HSOM と NC の利点・欠点を明らかにし，**3.** で HCM を提案する．**4.** で画像の認識実験を行い，HCM の能力を HSOM および NC と比較することにより，HCM が最初に述べた三つの課題を解決する画像認識モデルであることを示す．

2. 階層化自己組織マップ（HSOM）とネオコグニトロン（NC）

2.1 自己組織マップ（SOM）

Kohonen の SOM は，生体のハイパコラムの形成モデルとして知られると同時に，パターン認識における入力データ空間のベクトル量子化方式としても注目されている [10], [13]．SOM の構造を図1に示す．ニューロン $(1, \cdots, u, \cdots, N)$ はすべて，同一の入力ベクトル I を入力とする．そして，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 をもっており，I と W_u の類似度に応じて活性化する．また，すべてのニューロン間には強い抑制性の結合があり，最終的には，入力ベクトルに最も類似した



(a) ニューロンの結合 (b) 本稿での略記法

図1 SOMの構造
Fig. 1 Structure of SOM.

結合係数ベクトルをもったニューロンが他のすべてのニューロンを抑制して，一つだけ活性化する．単純化したモデルでは，例えば類似度にユークリッド距離を用いた場合，

$$\|I - W_c\| = \min_u(\|I - W_u\|) \tag{1}$$

となるニューロン c だけが活性化する．

学習時には，学習データが入力されるたびに，式(1)により最大活性ニューロン c を求め，

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

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

と学習する．SOMの特徴は，ニューロン間にも距離が定義されており，近傍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 u の距離である．そして，繰返し処理の各フェーズでは，最大活性ニューロンの結合係数だけを更新するのではなく，最大活性ニューロンの近傍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．$\sigma^2(t) = 0$ として近傍ニューロンの学習をせず，最大活性ニューロンのみを学習する方式を（狭い意味で）競合学習法と呼ぶ．$\alpha(t)$ は，学習レートであり，誤差逆伝搬学習法をはじめとする他の最急降下法を基本原理とする学習法と同様に，学習の過程で徐々に0に近づける．このような学習法は，入力データ I を標本点 c で表現することによる“量子化誤差”の平均

$$E = \int \|I - W_c\|^2 p(I) dI \tag{4}$$

但し，$p(I)$ は I の出現確率

を最小にするように標本点を最適配置するベクトル量子化問題の近似解法になっていることが知られている．

競合学習法によるベクトル量子化に比べ，Kohonenの学習方式によって得られる標本点すなわち結合係数の特徴は以下のように要約することができる[10]．

(1) 自己組織化の際に与える標本点の初期状態によらず，実際にデータが存在するデータ空間内の領域のみ，かつその領域全体を標本化するので効率が良い．

(2) 標本点の密度は入力データの密度をよく近似する．このことは，情報理論的な観点から見れば，ヒストグラムの均一化に相当し，量子化誤差最小の意味で最良のベクトル量子化をよく近似するものである．

(3) SOMの次元数は，入力データ空間の次元数と一致する必要はなく，トポグラフィック（位相保存）・マッピングと呼ばれる特性が保たれるので，データ空間の次元の縮小に適している．

ここで，トポグラフィックマッピングとは，入力空間での距離構造をネットワーク上で最大限に保存するマッピングのことを言う．このようなことから，SOMのネットワークを特徴「マップ」と呼ぶこともある．

図2にSOMを用いたベクトル量子化の簡単な例を示しておく．図2は，2次元のベクトル空間を1次元のSOMで量子化した例で，ら旋領域にデータが均一に分布している．図から，実線と白丸で示したSOMがデータの分布領域のみを一様に覆っているのがわかる．このように，入力ベクトル空間よりも実際にデータの分布する領域の次元数が低い場合に，SOMは興味深い特性を示すため，次節で述べるようにパターン認識の前処理としても注目されつつある．

2.2 階層化自己組織マップ（HSOM）

図3の左側に示した例は，図2と同様にら旋状に分布したデータを2分割し，領域Aと領域Bに分類する問題である．この問題は，誤差逆伝搬法を用いた多層パーセプトロンおよび従来の統計的手法にとっては複雑な問題である．これに対し，Lampinenらは，この問題がSOMを二つ重ねた多層の人工神経回路網で容易に類別できることを示している[11]．Lampinenらが用いたネットワークを図3の右側に示す．まず，十分な数のニューロンからなる1次元のSOM（1層目）に入力ベクトルを入力し，入力空間を量子化する．次に，1層目のニューロンの“添字”を2層目の入力とし，1層目のマップを二つに“再量子化”する．また，Lampinenらはいくつかの例題にHSOMを適応し，従来のSOMやK-mean法，Isodata法と比べて，類別問題におけるHSOMの優位性を示している．

HSOMの利点は，次の二つに要約でき，画像認識に極めて有効であると考えられる．

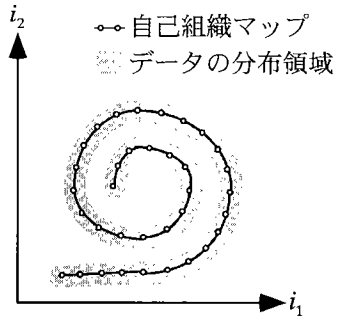


図2　SOMによるベクトル量子化の例
Fig. 2　An example of vector quantization by SOM.

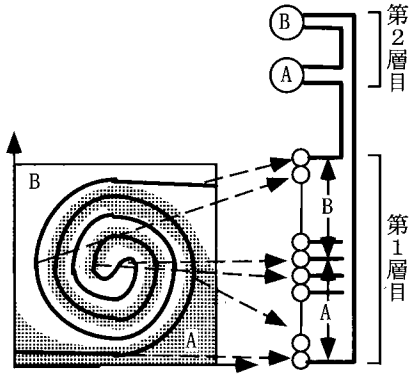


図3　Lampinenらの HSOM
Fig. 3　HSOM proposed by Lampinen.

（1）　入力データの連続的な変化に対して十分に密なサンプルデータが得られる場合には，1. の課題3で述べた，入力データの複雑な分布への柔軟な対応が達成できる．

（2）　データの分布領域の次元数がデータ空間の次元数よりも低い場合に，課題1を解消するパターン認識の前処理として，入力データの次元を縮小する効果をもつ．

しかし，HSOMをそのまま一般の画像認識に応用する場合は，次の三つの問題が生じると考えられる．

（1）　HSOMによる認識は，基本的にはメモリベースの手法であると考えられるので，データの分布領域の次元数が増え，空間のサイズが大きくなると，分布領域の“体積”のオーダでネットワークサイズが大きくなる．従って，課題2へ対応する別の仕組みが必要不可欠である．

（2）　マップの形成は，テンプレートマッチングによる画像間の距離（本論文の場合は式(1)）で定義される空間におけるサンプルデータの分布の連続性に基づいて行われる．従って，対象物の位置ずれと拡大縮小に対して画像間の距離が十分に滑らかに変化しない場合は，マップの形成が困難である．この場合は，画像をあらかじめぼかしておくなどのヒューリスティックを用いるか，十分に滑らかな画像間距離が得られる程度にサンプル数を増やす必要がある．

（3）　認識方法も，テンプレートマッチングによる画像間距離によるもので，自己組織化により得た標本点を用いた最近傍法と位置づけることができる．従って，画像をあらかじめぼかしておくなどのヒューリスティックを用いるか，非常に高い精度の対象領域の切出しと位置や大きさの正規化を行う必要がある．

2.3　ネオコグニトロン（NC）

NCは，図4に示すようにLampinen型の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．Lampinen型の1層目のマップがNCの特徴抽出層 U_S に，2層目マップがNCの位置ずれ吸収層 U_C に相当する．一つのHSOMに相当した部分は，1層目（特徴抽出層）に複数のマップをもち，特徴抽出層の入力領域は，図4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もった構造になっている．各位置のマップをハイパコラムと呼ぶ．位置ずれ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．

NCの「構造上の」利点は，次の二つに要約でき，HSOMの欠点を解消するものとなっている．

（1）　分割統治型の階層構造になっているため，各層におけるマップは比較的小さいものとなる．これは，HSOMの欠点1を解消する．

（2）　HSOMの欠点2および3に対して，NCは，位置ずれを吸収する機能をもっているため，ぼかしなどの前処理や精度の高い位置や大きさの正規化を必要としない．また，自己組織化の際に必要なサンプル数を軽減できる．

しかし，オリジナルのNCは，そのまま一般の画像認識に応用する場合，次の二つの問題点をもっている．

（1）　NCでは，近傍学習を行わない単純な競合学習を用いて特徴抽出層を学習するため，学習結果が各ニューロンに与える結合係数の初期状態に大きく依存し，十分な特徴抽出が行えない場合が頻繁に起こる．これは，HSOMの利点1によって解消可能である．

（2）　位置ずれ吸収層は，局所特徴の位置ずれだけ

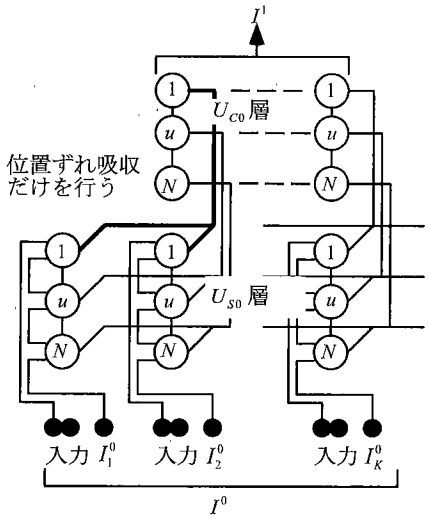


図4 ネオコグニトロンの構造
Fig. 4 Structure of Neocognitron.

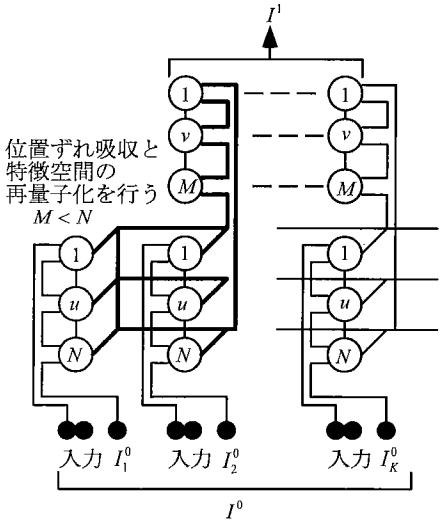


図5 HCMの構造
Fig. 5 Structure of HCM.

を吸収するものであり，HSOMのような特徴空間の再量子化は行わない．そのため，対象の変形に関しては十分な次元の縮小が行えていない．これは，HSOMの利点2によって解消可能である．

3. ハイパコラムモデル（HCM）

3.1 HCMの構造と利点

2.2および**2.3**から，NCの利点と欠点は，HSOMの利点と欠点の正に"裏返し"になっている．従って，HSOMとNCは，互いの欠点を利点で補完し合う関係にあると言える．そこで，本論文では，HSOMとNCを統合した画像認識モデルHCMを提案する．

提案するHCMは，図5のように，Lampinen型のネットワークをNCと同様にピラミッド型に結合したものである．また，一つのHSOMは，1層目（特徴抽出層）に複数のハイパコラムをもち，特徴抽出層の入力領域は，図の 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もった構造にする．また，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とする．HCMでは，NCと異なりLampinen型の2層目で，マップの再量子化と位置ずれの吸収を同時に行うので，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になる．HCMでは，この2層目を特徴統合層と呼ぶ．

この構造により，HCMは，HSOMとNCの問題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」と位置づけることができる．このような手法で高い画像の認識能力を得るためには，次の2点が重要であると考えられる．

（1） 最近傍法による認識であるから，対象物の位置や大きさ，形の連続的な変化（位相）を保存し，かつサンプルデータの分布密度をよく近似するマップをあらかじめ自己組織化によって形成しておくこと．

（2） 階層構造の要素としてメモリベースの手法であるHSOMを用いているので，特徴統合層において誤認識を起こさない程度に最大限にマップの次元数を減らすこと．

4.では，以上二つの観点からHCMの画像認識能力を評価する．

3.2 HCMによる認識

HCMによる画像認識は，次のように行う．まず，自己組織化の際に最上位層の各ニューロンに対応したサンプルデータの中で最も多数を占めたカテゴリーをそのニューロンが表すカテゴリーとして決めておく．認識の際は，画像データを入力したときに最上位層で最も強く活性化したニューロンのカテゴリーを入力デー

タのカテゴリーとして出力する．ニューロンの活性値は，以下の方法で最下位層から最上位層まで順に計算する．

HCM のある層の一つの HSOM を考える．この HSOM の特徴抽出層（第 1 層）には $K \times K$ 個のハイパコラムがあるとし，それぞれを C_{ij} : $(1 \leqq i \leqq K, 1 \leqq j \leqq K)$ と表す．C_{ij} 内のニューロン u_{ij} : $(1 \leqq u_{ij} \leqq N)$ を，下層の特徴統合層内の $P \times Q \times R$ のニューロンと結合し，それら $P \times Q \times R$ の活性化パターン I_{ij} を u_{ij} の入力とする．ここで，$P \times Q$ が位置ずれ方向の結合領域の大きさで，R が下層の特徴統合層のマップの大きさである．最下層では，入力は $P \times Q$ の部分画像である．また，すべてのニューロンの活性値は 0 から 1 に正規化されているものとする．u_{ij} の活性値 U_{ij} は，入力 I_{ij} と結合係数 W_u の類似度に基づいて計算し，$0 \leqq U_{ij} \leqq 1$ に正規化する．本論文では，

$$U_{ij} = 1 - \frac{\|I_{ij} - W_u\|}{\sqrt{PQR}} \tag{5}$$

とする．

次に，特徴統合層内のニューロン v : $(1 \leqq v \leqq M, M < N)$ は，次節で述べる自己組織化によってあらかじめ下の特徴抽出層の $K \times K \times N' : N' < N$ のニューロン集合 S_v に対応づけておく．このとき，v の活性値 V を，

$$V = \max_{u \in S_v}(U) \tag{6}$$

と計算する．以下，N'/N を特徴統合率と呼ぶ．

3.3 HCM の学習方法

以下に，HCM の学習方法を提案する．学習は，最下位層から順に 1 層ずつ学習する．同一層内ではその位置に関係なくマップは同じであるので，学習の際に用いるマップは各層で一つだけである．以下に，第 L 層の一つの特徴抽出層のマップの学習方法を示す．

```
σ，α の初期化
while( σ > 1.0) {
  for each サンプル画像 {
    L-1 層までの活性値を計算する．
    for each すべての HSOM の位置 {  --------A
      1.HSOM の入力部に L-1 層からのデータを
        入力し，1 層目全体で式 (1) を用いて
        最大活性化ニューロンを一つ求める．  ---B
      2. 式 (2) に従ってマップを更新する．
    }
  }
  σ を小さく．α を小さく．
}
最後の繰返しフェーズで，一度も最大活性化ニューロンとして選択されなかったニューロンを削除．
```

A のステップで，あらかじめ注目すべき特徴が選択できる場合は，必要な個所だけ処理をして計算量を軽減することができる．例えば，線画を扱う場合は，線を含む小領域だけを処理したり，一般の画像でもあらかじめ背景が除去できる場合は，背景を除く部分だけを処理すればよい．このような処理は，局所的な処理に基づくアルゴリズムの利点でもある．以下，注目すべき特徴領域を選択するフィルタを関心オペレータと呼ぶ．

また，B のステップで，1 層目全体で最大活性化ニューロンを一つ決めるため，$K \times K$ のハイパコラム内には同じ特徴の位置ずれパターンが同時に形成されることがない．従って，自己組織化の際，画像から部分特徴を切り出し，位置ずれの正規化を行うといったサンプルデータ生成のための前処理が不要である．

特徴統合層は，Lampinen らが示したように特徴抽出層で最大活性化ニューロンとして選択されたニューロンの添字を入力として特徴抽出層と同様の学習をすればよい．但し，後で示す実験では，特徴抽出層のマップを等分し，得られた各部分マップを特徴統合層の一つのニューロンに対応づけた．

4. 実　　験

3.1 で述べたように，HCM の認識能力を左右するポイントは，マップにおける位相の保存性とサンプルデータの分布密度の近似能力，および特徴統合率である．本章では，実験 1 でマップの特性を調べ，実験 2 で特徴統合サイズが及ぼす認識能力の変化を調べる．

4.1 実験 1：マップの特性

4.1.1 目　　的

自己組織化によって HCM の最上位層の特徴抽出層で形成されるマップの特性を調べる．具体的には，対象物の位置や形が連続的に変化する画像データの集合を入力とした場合に，対象物の位置や形の変化を連続的に写像したマップが形成され，かつデータの分布密度を近似した標本点集合によりベクトル量子化が行わ

れることを示す．また，その結果，最近傍法による良好な識別が可能になることを確かめる．

4.1.2 方　　法

a) 入力データ集合

マップの自己組織化に用いる画像は，次に示す四つのカテゴリーから無作為に生成した中心 (C_x, C_y)，半径 R の "1 画素幅の円周" を一つ描いた 187 × 187 画素の 2 値画像とした．

カテゴリー 1

$$55 \leqq C_x \leqq 80, 55 \leqq C_y \leqq 80, 25 \leqq R \leqq 50$$

カテゴリー 2

$$107 \leqq C_x \leqq 132, 55 \leqq C_y \leqq 80, 25 \leqq R \leqq 50$$

カテゴリー 3

$$55 \leqq C_x \leqq 80, 107 \leqq C_y \leqq 132, 25 \leqq R \leqq 50$$

カテゴリー 4

$$107 \leqq C_x \leqq 132, 107 \leqq C_y \leqq 132, 25 \leqq R \leqq 50$$

ここで用いたサンプルデータは，中心座標と半径に注目すれば，明らかに各カテゴリーに 1 対 1 に対応した，連続的に分布する四つのクラスタからなる．しかしながら，2 枚の画像間のユークリッド距離は，位置や大きさの連続的な変化に対して連続的に変化しない．例えば，図 6 (a) に示す同じカテゴリーに属する二つの円周間の距離と，図 6 (b) に示す異なるカテゴリーの円周間の距離とは同じになる．このことから，この画像集合のベクトル量子化は，画像全体を直接マップに対応づける HSOM にとっては，理想的なマップを形成することが困難な問題と言える．

また，円の大きさの変化は，円弧の曲率の変化を伴うものであるので，局所的に見れば，特徴の位置ずれだけでなく，特徴の変形を伴う．従って，局所的な特徴の位置ずれのみを吸収し，変形を吸収しない NC では，理想的なマップの形成が困難な問題と言える．

一方，HCM は位置ずれと拡大縮小（変形）を吸収する機構を備えているので，上記のいずれの問題も克服し，マップ内に四つの正しいクラスタを形成することができると考えられる．実験は，サンプル数を 100 個と 200 個，400 個と変化させて行った．

b) HCM の仕様

4 層からなる HCM を用いた．各層のニューロン数やマップの次元数，特徴統合の大きさ，関心オペレータは以下のように決めた．

[ニューロン数] 最上位層の特徴抽出層では 100 個のニューロンを用いて自己組織化を行った．各層のマップのニューロン数は，表 1 のとおりである．但し，表中のニューロン数は，**3.3** の学習アルゴリズムの最後のステップで不要なニューロンを削除した結果の数である．サンプル数が 200 個のときと 400 個のときは，サンプル数 100 個のときに形成した第 4 層（最下層）と第 3 層のマップをそのまま用い，上位 2 層だけを新たに自己組織化した．

ここで，目標となる認識精度を達成するために必要なニューロン数をあらかじめ決めることは，HCM に限らず困難な問題である．この問題に関して，自己組織化の過程で適応的にニューロン数を制御するベクトル量子化の研究が多くなされており [1], [7], [12], [15], [18]〜[20]，HCM でも，これらの技術を利用することにより，より適応的で柔軟なモデルを構築することができると考えられる．

[マップの次元数] 表示と解析の容易さからすべての層において 1 次元のマップを用いた．

ここで，円周の認識では，前にも述べたように局所

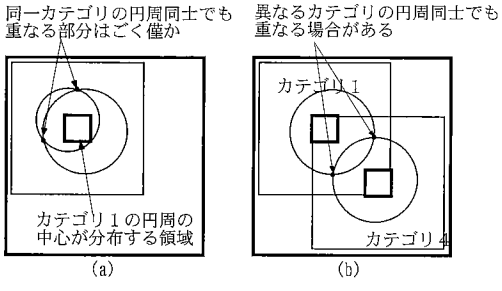


図 6 マップの生成に用いたサンプル画像例（円周）
Fig. 6 Examples of sample data (circle) for experiment 1.

表 1 円周のマップ生成に用いた HCM のニューロン数
(( ) 内は NC の場合)
Table 1 Number of neurons for each layer of HCM.

層番号/サンプル数		100	200	400
第 1 層	C_c 層	なし	なし	なし
最上位	C_s 層	82 (61)	97	99
第 2 層	C_c 層	10 (22)	10	10
	C_s 層	100 (22)	100	100
第 3 層	C_c 層	5 (100)	5	5
	C_s 層	100 (100)	100	100
第 4 層	C_c 層	3 (42)	3	3
最下位	C_s 層	42 (42)	42	42

的な特徴の変化は，位置ずれと曲率の変化を伴う．位置が異なる円弧は，それぞれ位置の異なるハイパコラム内のニューロンによって検出されるので，一つのハイパコラム内で検出しなければならないのは，二つの特徴量：曲率と向きが異なる円弧すべてである．従って，本来マップの次元数は2次元が適していると考えられる．但し，SOMのトポグラフィックマッピングの性質により，データ空間とマップの次元数の不一致が直ちに認識精度を低下させるものではない．また，HCMにおいて，マップの次元数を増やすことは極めて容易である．

［特徴統合の大きさ］ 特徴統合層では，マップの自己組織化による再量子化をせず，特徴抽出層のマップを等分して，それぞれの部分マップを特徴統合層の一つのニューロンに対応づけた．また，すべての層で共通に3×3のハイパコラム集合を一つに統合した．

各層で特徴抽出層から特徴統合層にかけて，どの程度データの次元を縮小させるかは，非常に難しい問題である．この問題については，実験2の考察で述べる．

［関心オペレータ］ 特徴抽出層のマップへの入力は5×5のハイパコラム集合とした．また，関心オペレータを用い，5×5の領域の中心にアクティブなニューロンが存在する場合のみ学習を行った．これにより，計算量を大幅に削減できた．

c）評価方法

HCMの最上位層の特徴抽出層で形成されたマップを，HSOMの1層目で形成されるマップ，およびNCの最上位層の特徴抽出層で形成されるマップと比較した．位相保存性は，一つのニューロンが最大活性化ニューロンとして選択されたカテゴリー数の平均（D1）と一つのカテゴリーに対する最大活性化ニューロンが形成する連続したニューロン集合（クラスタ）の個数の平均（D2）を用いた．D1とD2は，ともに1のときが理想的な値である．

また，標本点によるデータの分布密度の近似性能は，テストデータに対して各ニューロンが最大活性化ニューロンとして選択される回数（優勝回数）を求め，すべてのニューロンの優勝回数の平均が1回になるように，各ニューロンの優勝回数を正規化したときの，優勝回数の分散（VAR）を評価基準として比較した．ランダムに生成したテストデータに対しては，優勝回数が均一になり分散が0となる状態が理想である．

HSOMは，位置ずれや拡大縮小を吸収する機構をもっていないので，画像をあらかじめぼかしておく前処理を行った．3×3の平均化フィルタを0回適用した画像を入力とした場合（HCM0）と2回適用した画像を入力とした場合（HSOM1），4回適用した画像を入力とした場合（HSOM2）と比較した．

NCは，特徴の再量子化を行わないHCMの特殊な場合とし，前節で提案した自己組織化アルゴリズムを用いて構成した．オリジナルのNCのように，近傍学習を行わない単純な競合学習を用いた場合は，十分な特徴抽出が行われず，最上位層で複数のニューロンからなるマップが形成されないことが頻繁に起こり，比較が行えなかった．また，NCは特徴統合を行わないため，層間の結合数が非常に多くなり，学習時間が長くなる．実際，この実験の場合，HCMとNCでは，学習に要した時間は約20倍の差が生じている．そこで，NCは，サンプル数100個の場合のみ比較した．

4.1.3 実験結果

まず，各カテゴリーから無作為に25個ずつの円周を生成し，サンプルデータとした場合に形成されたマップの例を図7に示す．図の横軸はサンプルデータの番号で，1番から順に25個ずつのデータがぞれぞれカテゴリー1からカテゴリー4までに対応している．縦軸は，ニューロン番号である．HCMでは，四つのカテゴリーが，それぞれ連続した一つのニューロン集合（クラスタ）に対応しており，各ニューロン集合の大きさもほぼ同じものとなった．図中，カテゴリー3に対応したニューロンクラスタは二つのクラスタに分かれているように見えるが，これはマップが特徴空間上で環状をしているためであり，実際は連続した一つのクラスタを形成している．次に，各サンプル数における各モデルの評価値を表2に示す．最後に，サンプル数100個ので自己組織化したHCMに未学習のテストデータ100個を入力した場合のデータとマップの対応を図8に示す．このときの認識率は100%であった．

4.1.4 考　　察

HSOM0は，位置や大きさの変化を吸収する機構をもたないので，全く正しい類別が行えなかった．HSOM1とHSOM2は，前処理としての画像のぼかし操作によって，位置や大きさの変化に対する画像間の距離の連続的な変化が得られたため，比較的良好な結果を示したものの，いずれもHCMより劣った．HSOM1とHSOM2では，ぼかしの程度が大きいHSOM2の方がより良好であった．また，サンプル数を増やすことは，必ずしも性能の改善にはつながらなかった．ぼかしの度合を強めても，サンプル数を増やしても，D2の値

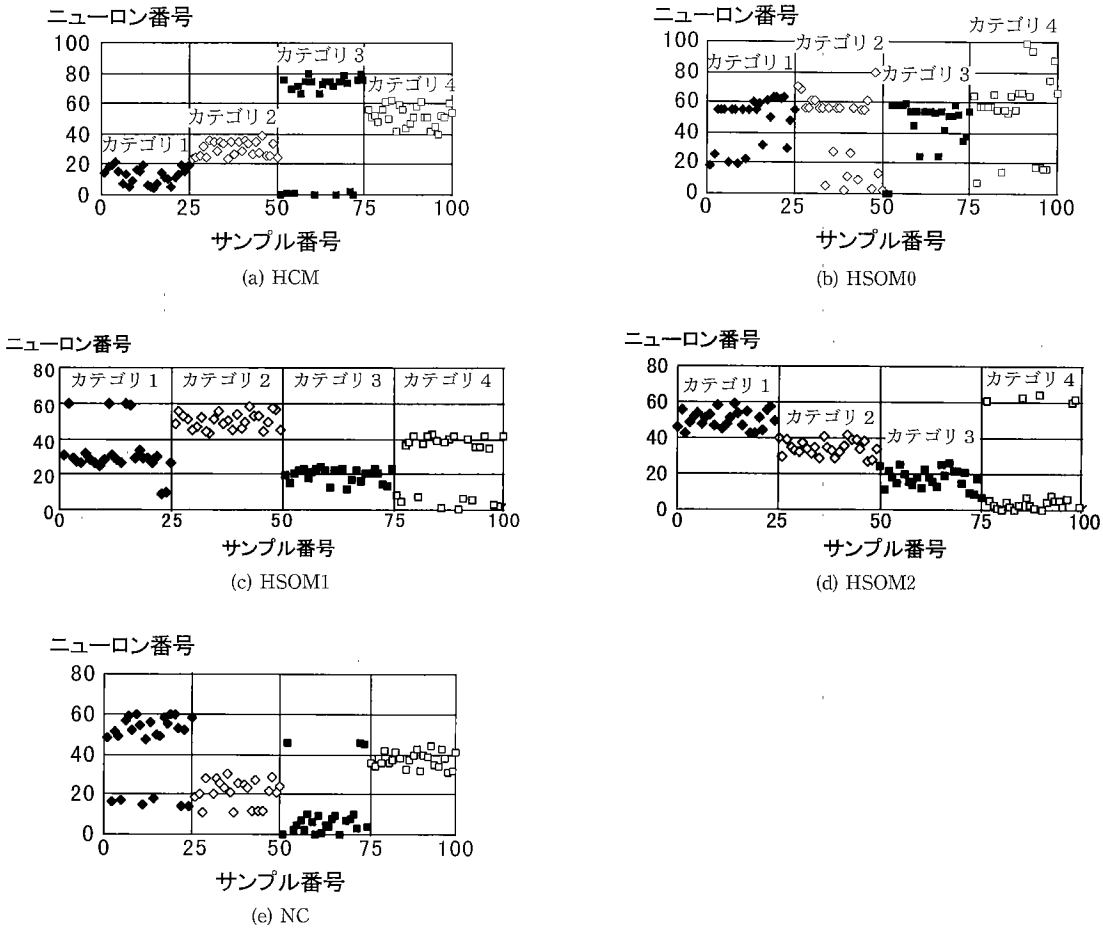


図7 各モデルのマップとサンプル画像の対応関係（サンプル数100の場合）
Fig. 7 Correspondence between maps of each model and sample data.

表2 各モデルのマップの評価
Table 2 Characteristics of maps of each model.
D1: num. of categories/neuron
D2: num. of neuron-groups/category
VAR: variance of normalized number of times to win

		評価値		
基準	モデルタイプ	100サンプル	200	400
D1	HCM	1.00	1.00	1.00
	NC	1.08	—	—
	HSOM1	1.00	1.01	1.01
	HSOM2	1.00	1.00	1.00
D2	HCM	1.00	1.00	1.00
	NC	3.50	—	—
	HSOM1	1.75	3.25	2.75
	HSOM2	1.00	1.75	1.75
VAR	HCM	0.13	0.14	0.14
	NC	1.08	—	—
	HSOM1	0.50	0.73	1.26
	HSOM2	0.26	0.69	0.62

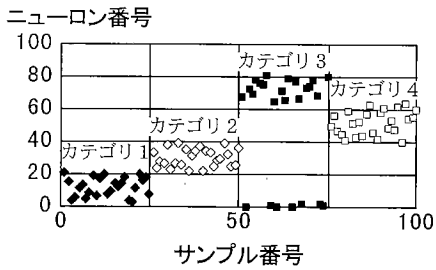
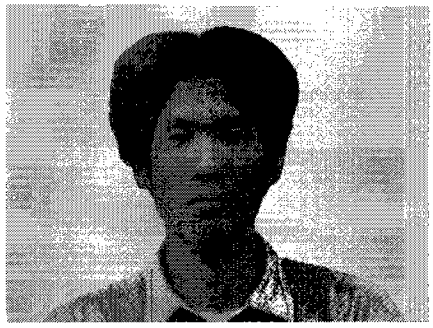


図8 HCMによる未学習データの認識（認識率は100%）
Fig. 8 Recognition of test-data by HCM. (the rate of correct is 100%)



(a) 人物 E の正面画像

(b) 人物 E の左 20 度画像

(c) (a) の canny filter 出力

(d) (b) の canny filter 出力

図 9　サンプル画像
Fig. 9　Examples of sample data.

が 1 に近づかない理由は，ぼかし操作によって得られる画像間距離が位置や大きさの変化を正しく反映していないためであると考えられる．このように，適切なヒューリスティックを発見することは一般には困難であると考えられる．この問題の他にも，一般の場合には，前処理としての画像のぼかし操作は，局所的な微細な特徴による類別・識別を不可能にしてしまうという問題点をもっている．

NC は位置ずれ吸収機構をもつにもかかわらず，良好な結果が得られなかった．D1 と D2 の値が理想値に一致しない理由は，ニューロン数の不足であると考えられる．NC は特徴空間の再量子化を行わないため，次元の縮小が十分に行えず，上位層の特徴抽出に必要なニューロン数が増大してしまう．

これらに対し，HCM は，すべてにおいて，極めて良好な結果を示している．このことから，HCM は，位置や大きさの変化を連続的に写像したマップを形成し，かつ，データの分布密度を近似した標本点集合によりベクトル量子化が行われると言える．実際，テストデータに対する認識実験でも，認識率 100%を示しており，極めて良好であった．

4.2　実験 2：特徴統合率と認識能力の関係

4.2.1　目　　的

HCM における特徴統合率とパターンのひずみに不変な認識能力の関係を調べる．同時に，実画像に対する HCM の認識能力を HSOM および NC と比較する．

4.2.2　方　　法

a）入力データ集合

5 名の人物を五つの方向（左 20 度，左 10 度，正面，右 10 度，右 20 度）から撮影し，320 × 240 画素，256 階調の濃淡画像とした．いずれの画像も人物の位置，大きさ，照明の条件をほぼ同じとした．左 20 度と正面，右 20 度の 3 枚ずつの画像を自己組織化のためのサンプルデータとし，残りの左 10 度と右 10 度の 2 枚ずつを認識時のテストデータとした．濃淡画像をそのまま入力とした場合と canny filter を用いてエッジ検出して入力とした場合の 2 通りの実験を行った（図 9）．

表3 顔画像の類別に用いたHCMのニューロン数
Table 3 Number of neurons for each layer of HCM.

層番号/タイプ		HCM1	HCM2	HCM3	NC
第1層	C_c 層	(なし)	(なし)	(なし)	(なし)
最上位	C_s 層	15	15	15	15
第2層	C_c 層	14	14	14	32
	C_s 層	15	18	15	32
第3層	C_c 層	8	8	45	33
	C_s 層	76	89	90	32
第4層	C_c 層	2	4	10	33
	C_s 層	100	100	100	33
第5層	C_c 層	2	4	10	33
	C_s 層	100	100	100	33
第6層	C_c 層	2	4	10	33
最下位	C_s 層	100	100	100	33

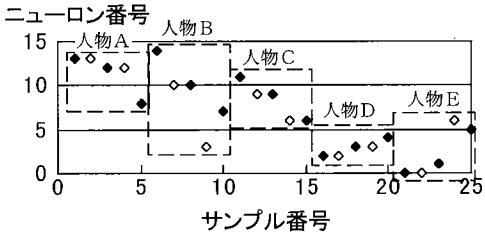


図10 マップとサンプル画像の対応（例：HCM2の場合）
Fig. 10 Correspondence between map of type-2 HCM and sample data.

b) HCMの仕様

6層からなるHCMを用いた．各層のマップの次元数と関心オペレータは実験1の場合と同じである．特徴統合率を1/2 (HCM1)，1/4 (HCM2)，1/10 (HCM3)と変化させた三つのタイプのHCMを用いた．それぞれのタイプのニューロン数は，表3のとおりである．

c) 評価方法

テストデータに対する認識率を評価基準として，特徴統合率と認識率の関係を調べた．また，実験1と同様の前処理を行った三つのタイプのHSOMおよびNCと認識率を比較した．前にも述べたように，NCはHCMに比べて計算時間が膨大なものになるので，ここでは，HCM3と同程度の計算時間（Pentiumプロセッサ166 MHzのPCで8時間程度）で学習が終了するニューロン数で実験を行った．

4.2.3 実験結果

図10にHCM2で形成されたマップと入力データの対応を例として示す．エッジ画像を入力としたときの各タイプに対する認識率を表4に示す．HSOM1以外は，すべて高い認識率を示している．また，濃淡画像を入力としたときの各タイプに対する認識率を表5に示す．HCM3とHSOM1の認識率は100%であった．

表4 エッジ画像の認識率
Table 4 Rates of correct recognition by each HCM for edge images.

モデルタイプ	認識率 (%)
HCM1	70
HCM2	80
HCM3	80
NC	70
HSOM1	40
HSOM2	90

表5 濃淡画像の認識率
Table 5 Rates of correct recognition by each HCM for gray images.

モデルタイプ	認識率 (%)
HCM1	80
HCM2	90
HCM3	100
HSOM1	100
HSOM2	80

4.2.4 考察

HSOMとHCMは，ともに高い認識率を示した．但し，HSOM1とHSOM2の認識率は，エッジ画像を入力とする場合と濃淡画像を入力する場合で逆転しており，前処理としてのぼかしの程度の決め方が難しいことを示す結果となった．また，もちろん，位置ずれや拡大縮小を伴うデータの場合は，HSOMの認識率は低下すると考えられる．

HCMに関しては，いずれのタイプも高い認識率を示し，認識率の変動は小さかった．傾向としては，特徴統合率を小さくした方が認識能力は高い．但し，NCのように，極端にこの割合を小さくしすぎると，段階的な次元の縮小が不十分であるため，マップの形成に必要なニューロン数が増加し，結果的に用意したニューロン数では十分な精度の特徴抽出が行えなくなり，認識率も低下する．特徴統合の最適な割合を求めることは非常に難しい問題である．経験的な指針としては，自己組織化後の「量子化誤差」が十分に小さいことを目安として，計算資源や計算時間が許す限り次元の縮小率を小さく抑える方法がよい．

5. むすび

Lampinenらの階層化自己組織マップと福島らのネオコグニトロンを融合することによって，画像認識を困難なものにしていた三つの要因，すなわち画像の冗長性とデータが分布するカテゴリー空間の高い次元数，

およびカテゴリー空間の複雑な形状を同時に緩和するハイパコラムモデルを提案した．階層化自己組織マップは，従来の線形投影を基本原理とした前処理と異なり，非線型の写像を基本原理とし，特徴空間を情報理論的に最適に量子化して構造化する．ハイパコラムモデルは，階層化自己組織マップをネオコグニトロン型のピラミッド構造に重ね合わせることによって，特徴空間の量子化を階層的に行い，適応的に入力画像の次元数を減らすことができる．

本論文で提案したHCMは，SOMの自己組織化を利用した教師なし学習による認識法であるにもかかわらず，実験では，非常に高い認識能力を示した．しかし，一般には，識別問題に適用する場合は，教師付き学習による方がよい．そこで，あるレベルまでHCMを用いて入力画像の次元数を縮小しておいて，最終的には，多層パーセプトロンのような教師付き学習が可能なモデルで識別する方法が考えられる．このような前処理としては，特に顔画像認識において主成分分析を用いることが一般になっている．しかし，主成分分析は，あくまでも線形な投影に基づく部分空間法であるので，非線型な写像を必要とする問題には本質的な解を与えない．この意味では，SOMを基本原理としたモデルは優位性を示すものと思われる．この問題については，今後，具体的な例題による比較評価が必要である．また，HCM自体に誤差逆伝搬に相当する教師付き学習を導入することも一つの方法であり，これについては現在研究を進めている．

謝辞　本研究は，文部省科学研究費補助金「基盤研究（C）(2)，課題番号（08680403）」の援助のもとに行われた．

文　献

[1] B. Fritzke, "Growing cell structures—A self-organizing network for unsupervised and supervised learning," Neural Networks, vol.7, no.9, pp.1441–1460, 1994.

[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, Oct. 1979.

[3] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biological Cybernetics, vol.36, no.4, pp.193–202, 1980.

[4] 福島邦彦，"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, pp.993–1003, June 1989.

[5] K. Fukushima, "A neural network model for selective attention in visual pattern recognition," Biological Cybernetics, vol.55, no.1, pp.5–15, 1986.

[6] 福島邦彦，"神経回路網と情報処理，" 朝倉書店，1989.

[7] S. Jockusch and H. Ritter, "Self-organizing maps: Local competition and evolutionary optimization," Neural Networks, vol.7, no.8, pp.1229–1240, 1994.

[8] T. Kohonen, "Self-organized function of topologically correct feature maps," Biological Cybernetics, no.43, pp.43–69, 1988.

[9] T. Kohonen, "Self-Organization and Associative Memory," Springer-Verlag, Berlin Heidelberg, 1989.

[10] T. Kohonen, "Self-Organizing Maps," Springer Series in Information Sciences, 1995.

[11] J. Lampinen and E. Oja, "Clustering properties of hierarchical self-organizing maps," Journal of Mathematical Imaging and Vision 2, pp.261–272, 1992.

[12] T. Martinetz and K. Schulten, "Topology representing networks," Neural Networks, vol.7, no.3, pp.507–522 1994.

[13] E. Oja, L. Xu, and P. Kultanen, "Curve detection by an extended self-organizing map and the related RHT method," Proc. of Int. Neural Network Conf., vol.1, pp.27–30, 1990.

[14] M. Rosenblum, Y. Yacoob, and L. Davis, "Human expression recognition from motion using a radial basis function network architecture," IEEE Trans. on Neural Networks vol.7, no.5, pp.1121–1138, 1996.

[15] D. Ruwisch, M. Bode, and H. Purwins, "Parallel hardware implementation of kohonen's algorithm with an active medium," Neural Networks, vol.6, no.8, pp.1147–1158, 1993.

[16] K. Sung and T. Poggio, "Example-based learning for view-based human face detection," Proc. of Image Understanding Workshop, vol.II, pp.843–850, 1994.

[17] 鶴田直之，谷口倫一郎，雨宮真人，"一般化ハイパーコラムモデルの画像認識能力について，" 画像認識・理解シンポジウム（MIRU'94），vol.1, pp.257–264, 1994.

[18] N. Ueda and R. Nakano, "A new competitive learning approach based on an equidistortion principle for designing optimal vector quantizers," Neural Networks, vol.7, no.8, pp.1211–1228, 1996.

[19] 宇津木明男，"トポロジー保存写像学習のベイズ学習，" 信学論（D-II），vol.J76-D-II, no.6 pp.1232–1239, June 1993.

[20] 宇津木明男，"ベイジアン自己組織化写像の側方相互作用，" 信学論（D-II），vol.J77-D-II, no.7 pp.1329–1336, July 1994.

[21] W.E. Weideman, M.T. Manry, H. Yau, and W. Gong, "Comparisons of neural network and nearest neighbor classifier via the numeric handprint recognition problem," IEEE Trans. on Neural Networks, vol.6, no.6, pp.1524–1530, 1995.

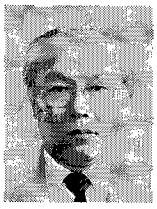
（平成9年10月22日受付，10年4月2日再受付）

鶴田　直之　（正員）

1988 九大・工・情報卒．1990 同大大学院総合理工学研究科修士課程了．同年同助手．1996 同大学院システム情報科学研究科助手．1998 福岡大学工学部電子情報工学科講師．工学修士．

谷口倫一郎　（正員）

1978 九大・工・情報卒．1980 同大大学院工学研究科修士課程了．同年同助手．1988 同助教授．1996 同大学院システム情報科学研究科教授．工博．



雨宮　真人　（正員）

1967 九大・工・電子卒．1969 同大大学院工学研究科修士課程了．同年，日本電信電話公社（現在 NTT）電気通信研究所入所．1988 九州大学大学院総合理工学研究科教授．現在，九州大学大学院システム情報科学研究科教授．工博．