Kohonen自己組織化特徴マップを応用した分類法

- 15種類のボタンと服地の色合せ問題への応用 -

正員 藤村 喜久郎 (鳥取大)

非会員 徳高 平 蔵 (鳥取大)

非会員 岸田 悟 (鳥取大)

A Method of Classification Using Kohonen's Self-Organizing Feature Maps
- Application of the Color Matching Problem in the Combination of Fifteen Buttons and Cloths -
Kikuo FUJIMURA, Member (Tottori University), Heizo TOKUTAKA, Non-member (Tottori University),
Satoru KISHIDA, Non-member (Tottori University)

We applied **Kohonen's SOMs** (Self Organizing feature Map) in the color matching problem in the combination of fifteen buttons and cloths. The chromaticity of cloth's color is measured by the method based on **CIE-Lab**. Each distribution of fifteen buttons is represented and formed by three **SOMs** on each two dimensional projected plane. From the observation of these **SOMs**, we confirm that each distribution of fifteen buttons has many mutually overlapping regions. In this situation, we propose a method of classification that have characteristics of ranking by adaptation in each button's **SOM**, and test the performance of our method using practical data. From the results, we obtained over 87% correct judgement, which is defined as correspondence of teacher's class with one of the class in the best three by our method. Also our method and **LVQ** obtained almost the same recognition rate about the best answer.

キーワード：クラス分類, CIE-Lab, 自己組織化特徴マップ, パターン認識, ニューラルネットワーク

1. まえがき

ニューラルネットワークはパターン認識分野における新しい方法のひとつとして近年脚光を浴びてきた。特にクラス分類問題に限れば、バックプロパゲーション学習則を適用した多層パーセプトロン (XOR問題などの非線型分離が可能), **Kohonen**型ネットワークによる学習ベクトル量子化 (**LVQ**) (1) (2), 適応共鳴理論 (**ART**) (3) といった有用な方法が提案され一部は実用化されている。

しかし、これらの方法が有効であるのはクラス間での境界が比較的明確に区別できることが必要となる。一般にこの条件を満たさない場合、特にクラス間での境界領域が重複している場合には、これらの分類法ではクラス分類の正しい結果を得る事が困難であることが知られている。例えば、色の分類に関して考えてみる。色度計を使って測定したある色をAさんは赤みがかった黄色、Bさんはオレンジ色と分類することがある。従来のクラス分類方法では、この色に関しては色度計では同じ値を示すので唯一のクラス

に分類することが要求される。したがって、この両者の分類のうちいずれか一方が間違っている事になるが、本来は両方ともある確率をもって正しいとするべきである。

本論文では、重複領域をもったクラス分類問題に対して有効なクラス分類方法を提案する。我々の提案する分類方法は複数の**Kohonen**の自己組織化特徴マップ⁽⁴⁾ (以降**SOM**と表記) を利用した方法である。**SOM**は入力に用いたデータの確率密度分布を写像する能力があることが知られており様々な数学的解析が行われている^{(5) (6)}。その特性を利用して各クラスにおける適合度を求め、その適合度が高い順に最適候補、第2候補、第3候補といった順位付けを行うことを可能とする。実際に色の分類問題のひとつである『ボタン選定問題』に適用し^{(7) (8)}、その有効性を検証した。

2. ボタン選定問題とは

『ボタン選定問題』とは無限の種類があるスーツの服地の色と15種類のボタンの色との最適な組み合わせを選ぶ

問題である。その選択基準は服地の色とほぼ同色のボタンを選ぶことである。

ボタンの色には大別して5種類あり、緑色系 (G-series), 青色系 (M-series), 灰色系 (N-series), 茶色系 (O-series), 紫色系 (V-series) がある。緑色系は"G2", "G4", "G6"の3種類, 青色系は"M2", "M4", "M8"の3種類, 灰色系は"N2", "N4", "N6", "N9"の4種類, 茶色系は"O2", "O4", "O6", "O8"の4種類, 紫色系は"V1"の1種類の計15種類の色が存在する。この記号表記において1番目の記号は色系統を意味する。それぞれ, 'G'は緑色系, 'M'は青色系, 'N'は灰色系, 'O'は茶色系, 'V'は紫色系を意味する。次の数字は色の明度を表しており大きいほど明度が明るいことを意味している。従って, 同じ緑色でも"G4"のほうが"G2"よりも明るく感じる。

図1 CIE-LAB色空間
Fig.1. CIE-LAB chromaticity space.

服地の色は色彩色度計を用いて測定されたCIE-LAB色空間 (JIS Z8729-80) (9) で表される L^* , a^* , b^* を使用した。各々の値は正および負の実数をとる。 L^* は明度に相当し, (+) 方向に明るい色, (0) 方向に暗い色に対応する。 a^* は (+) 方向に赤みを帯びた色, (-) 方向に緑みを帯びた色に対応する。 b^* は (+) 方向に黄色みを帯びた色,

(-) 方向に青みを帯びた色に対応する (図1を参照)。この色空間は3次元直交座標系で表すことができる。また, ある2つの色における色差についてはこの色空間内でプロットされる2点間のユークリッド距離で表現することができる。

我々はCIE-LAB色空間内で各ボタンに属する服地の色分布をSOMを用いて投影写像し表現することを考えた。図2にその基本概念を示す。CIE-LAB色空間が3次元の直交座標系であることを利用して3つの投影面 (L^*-a^* , L^*-b^* , a^*-b^*) 上に写像した平面地図 (SOMを意味する) を作ることを考案した。つまり, 投影面上での2つの座標軸に対応する数値データを用いてSOMを形成することにした。例えば, L^*-a^* 面では L^* および a^* の値を利用してSOMを形成する。 L^*-b^* 面, a^*-b^* 面も同様に処理した結果, 1つのボタンについて3投影面でのSOMが得られることになる。図2ではあるボタンに属する服地の色分布を"3D Object"と想定し, 投影面上で得られるSOMを" L^*-a^* Projection", " L^*-b^* Projection", " a^*-b^* Projection"としている。

図2 各平面上に投影されるSOM
Fig.2. Projected SOMs on each plane.

図3 SOMを用いたボタン選定システムの構成
Fig.3. The framework of the button selecting system using SOMs.

3. システム

我々の考案したボタン選定システムを図3に示す。このシステムは2つのソフトウェアとデータベースからなる。図中、左側がSOMを形成するソフトウェア (Soft 1)、中央がSoft 1により形成されたSOMの重みベクトル情報を蓄えたデータベースそして右側がそのデータベースから地形情報 (数値情報) を取り出して数値解析を行いボタンのクラス毎に適合度を計算しクラスの順位付けを行うソフトウェア (Soft 2) である。以下に処理手順を示す。

- (1) Soft 1により各クラス毎にSOMを作成する
- (2) Soft 2により与えられたデータについて各クラスに属する適合度を計算し、適合度の高いクラス順に順位付けする

(1) の処理が学習、(2) の処理が検証時の動作となる。(1) と (2) の動作は独立している。従って (1) でデータベースを更新した場合、(2) で有効にするためには更新されたデータベースを新たに指定してSoft 2を再起動する必要がある。このため、学習データの経時的変化をリアルタイムで反映することはできない。

<3-1> SOMの形成 Soft 1で使用したネットワーク構造を図4に示す。典型的なKohonen型ネットワーク構造をしている。1層目の入力 ξ_1 と ξ_2 の2ユニットで2層目はKohonen型の競合層になっている。競合層の大きさは $N \times M$ となっている。1層目と2層目のユニット間は全結合をしている。あるユニット U_{ij} について2つの重み w_1^{ij} と w_2^{ij} が存在する。学習のアルゴリズムは文献(4)の第5章に示す簡易型の重みベクトル更新則を使用している。重みベクトル $\mathbf{m}_{ij}$ を適当に初期化後、以下の処理を最初に決めた繰り返し回数 T だけ繰り返す。

図4 Kohonen型ネットワークの構造
Fig. 4. The structure of Kohonen's network.

処理1 データ提示

入力ベクトル $\mathbf{x}$ (ただし、 $n=2$)

$$\mathbf{x} = [\xi_1, \xi_2, \xi_3, \dots, \xi_n]^T \dots\dots\dots (1)$$

ユニット u_{ij} の重みベクトル $\mathbf{m}_{ij}$ (ただし、 $n=2$)

$$\mathbf{m}_{ij} = [w_1^{ij}, w_2^{ij}, w_3^{ij}, \dots, w_n^{ij}]^T \dots\dots\dots (2)$$

処理2 適合度検査

適合度の計算は(3)式に示すユークリッド距離で行い、その値が最小となるユニットが勝者ユニット U_c と定義する。ここで、 U_c の重みベクトルを $\mathbf{m}_c$ と表す。

$$\|\mathbf{x}(t_k) - \mathbf{m}_c(t_k)\| = \min_{i,j} \{\|\mathbf{x}(t_k) - \mathbf{m}_{ij}(t_k)\|\} \dots\dots\dots (3)$$

処理3 更新

最適適合度をもつ勝者ユニット U_c と近傍領域 N_c に含まれるユニットのみを更新する。

$i, j \in N_c$ のとき

$$\mathbf{m}_{ij}(t_{k+1}) = \mathbf{m}_{ij}(t_k) + \alpha(t_k) [\mathbf{x}(t_k) - \mathbf{m}_{ij}(t_k)] \dots\dots (4a)$$

$i, j \notin N_c$ のとき

$$\mathbf{m}_{ij}(t_{k+1}) = \mathbf{m}_{ij}(t_k) \dots\dots\dots (4b)$$

この繰り返し処理における近傍領域 N_c および更新係数 $\alpha(t_k)$ の計算を以下に記述する。

近傍領域 N_c の計算

勝者ユニット U_c の近傍領域 N_c の計算はパラメータ d_0 および T を用いて(5)式で表される。ただし、 N_c が整数である必要から(5)式における小数部を切り捨てて整数部を使用した。図5に近傍領域の形状を示しており、方形領域を使用した。 $N_c=1$ 、 $N_c=2$ の時それぞれ9個、25個のユニットが更新される。

$$N_c = N_c(t_k) = (d_0 - 1)(1 - t_k/T) + 1 \dots\dots\dots (5)$$

d_0 : 近傍距離の初期値 (任意)、整数値 ($d_0 > 0$)

T : 最大計算回数、整数値 ($T > 0$)

図5 $N_c=1$ 、 $N_c=2$ 、 $N_c=3$ の形状例

Fig. 5. The shapes of $N_c=1$ 、 $N_c=2$ and $N_c=3$ 、

更新係数 $\alpha(t_k)$ の計算

α_0 と p を用いた更新係数 $\alpha(t_k)$ の計算を(6) - (8)式で表す。(6)式の第1項 $a(t_k)$ は単調減少関数による更新

係数の時間 t_k における値を表している。第2項 $se(p)$ は側方相互作用関数の項であり、 U_c から距離が遠いほど更新係数が小さくなるような(8)式で表される単調減少関数を使用した。

$$\alpha = \alpha(t_k, p) = a(t_k) \cdot se(p) \quad \dots\dots\dots (6)$$

$$a(t_k) = \alpha_0 (1 - (t_k - 1) / T) \quad \dots\dots\dots (7)$$

$$se(p) = (N_c + 1 - p) / N_c \quad \dots\dots\dots (8)$$

α_0 : 更新係数初期値, 実数 ($0 < \alpha_0 < 1$)

p : U_c からの距離, 整数値 ($1, 2, 3, \dots, N_c$)

〈3・2〉 SOMの解析 Soft 2ではSOMから得られる適合度の計算に後述する3つの要素を考慮した。各クラスにおける総合適合度 A_{Point}^{Class} は(9)式で表される3要素の和とした。(9)式による評価点が高いクラスの順に順位付けを行う。

$$A_{Point}^{Class} = I_{Point}^{Class} + D_{Point}^{Class} + C_{Point}^{Class} \quad \dots\dots\dots (9)$$

I_{Point}^{Class} : 参照データがSOMによる特徴地図内に含まれることによる絶対評価 (IN-POINT)
パラメータ: P_{in}, P_{out}

D_{Point}^{Class} : 参照データと最近傍ユニットとその8近傍との平均誤差による相対評価 (DIFF-POINT)
パラメータ: P_{DIFF}

C_{Point}^{Class} : 参照データとSOMによる特徴地図の重心点からの誤差による相対評価 (CENTER-POINT)
パラメータ: P_{CENTER}

総合適合度に各3要素が影響する割合は上記の4つのパラメータで決定することができる。したがって、対象とする問題に応じて適切な値を設定する必要がある。

以下に各要素の定義を説明する。

(a) IN-POINT

あるデータ点が示された時に、あるクラスの3枚のSOMからの情報によりそのクラスでの状態が“IN”であるか“OUT”を決定する。“IN”状態とは、そのクラスに属する3つのSOM内にデータ点が入っている状態の事をいう。“OUT”状態とは“IN”状態以外の7つの状態の事を言う。その考えられる全ての状態について図6に示す。“IN”状態の場合には P_{in} (通常 $P_{in} > 0.0$ の条件を満たす) 点を与え、“OUT”状態の場合には P_{out} (通常 $P_{out} = 0.0$ の条件を満たす) 点を与える。

$$I_{Point}^{Class} = \begin{cases} P_{in} & (\text{reference data is "IN" condition}) \\ P_{out} & (\text{reference data is "OUT" condition}) \end{cases}$$

..... (10)

All Situation of IN-POINT				
L*-a*	L*-b*	a*-b*	Total Condition	POINT
			IN	Pin
			OUT	Pout
			OUT	Pout
			OUT	Pout
			OUT	Pout
			OUT	Pout
			OUT	Pout

x Reference data point (L*, a*, b*)

In SOM Area (IN condition)

Out SOM Area (OUT condition)

図6 IN-POINTについての説明

Fig.6. Explanation of IN-POINT.

(b) DIFF-POINT

15種類のClass = {G2, G4, G6, ..., O8, V1}のうち、あるClassでのL*-a*面上のSOMの最近傍ユニット U_c とその8近傍との平均2乗誤差 $D2_{La}^{Class}$ を(11a)式で求める。この式の中で $D2_{Lak}^{Class}$ は図7において示すように $k=0$ の時が最近傍ユニット U_c との2乗誤差 $D2_{La0}^{Class}$ を、 $k=1, 2, \dots, 8$ の時はその8近傍ユニットとの2乗誤差 $D2_{La1}^{Class}$ から $D2_{La8}^{Class}$ を意味する。 m は近傍ユニットの個数でありSOMの最外輪郭線上のユニット以外が最近傍となった場合には $m=9$ となる。同様にしてL*-b*面、a*-b*面の平均2乗誤差 $D2_{Lb}^{Class}$ および $D2_{ab}^{Class}$ は(11b), (11c)式で与えられる。

$$D2_{La}^{Class} = \sum_{k=0}^8 D2_{Lak}^{Class} / m \quad \dots\dots\dots (11a)$$

$$D2_{Lb}^{Class} = \sum_{k=0}^8 D2_{Lbk}^{Class} / m \quad \dots\dots\dots (11b)$$

$$D2_{ab}^{Class} = \sum_{k=0}^8 D2_{abk}^{Class} / m \quad \dots\dots\dots (11c)$$

各クラスにおける評価点数は(12), (13)式を用いて計算する。(12)式において誤差の逆数を取ることににより平均誤差が小さいものが D_{Lab}^{Class} での評価が高くなる。(13)式は(12)式の値を15のクラスでの総和で割り規格化し、パラメータ P_{DIFF} で(9)式で与えられる総合評価への影響量を調整している。

$$D_{Lab}^{Class} = 1/\sqrt{D2_{La}^{Class} + D2_{Lb}^{Class} + D2_{ab}^{Class}} \dots\dots\dots (12)$$

$$D_{Point}^{Class} = P_{DIFF} D_{Lab}^{Class} / \sum_{Class=G2}^{V1} D_{Lab}^{Class} \dots\dots\dots (13)$$

× Reference data point (L*, a*)

図7 L*-a*面上のSOMにおける2乗誤差の定義
Fig.7. The definition of square errors in L*-a* SOM.

(c) CENTER-POINT

前述の15種類のクラスのうち、あるClassでのL*-a*面上のSOMの重みの平均値(重心)とデータ点(L*, a*)の2乗誤差をC2_{La}^{Class}と定義する(図8参照)。同様にC2_{Lb}^{Class}, C2_{ab}^{Class}もそれぞれL*-b*面上およびa*-b*面上のSOMの重みの重心との2乗誤差と定義する。(b)の場合と同様に考えて、重心から2乗誤差が小さい順に高い評価点が付くように各クラスでの適合度を(14), (15)式を用いて計算する。この評価による総点数はパラメータP_{CENTER}とする。

$$C_{Lab}^{Class} = 1/\sqrt{C2_{La}^{Class} + C2_{Lb}^{Class} + C2_{ab}^{Class}} \dots\dots\dots (14)$$

$$C_{Point}^{Class} = P_{CENTER} C_{Lab}^{Class} / \sum_{Class=G2}^{V1} C_{Lab}^{Class} \dots\dots\dots (15)$$

図8 L*-a*面上のSOMにおける重心からの2乗誤差の定義

Fig.8. The definition of square error from center point of L*-a* SOM.

4. シミュレーション結果と検討

SOMの形成には全3965件の学習用データを用いた。学習用データは"m2"が最大で700件,"n9"が最小で33件と各クラス間でばらつきがあった。認識試験には全2405件試験用データを用いた。試験用データは件数が200件より少ないボタンである"G4", "M4", "N9", "O4", "O6", "O8", "V1"を除いて学習時と異なったものを選んだ。後に判明したことだが、これらのデータは約15%以上の境界領域が重複していた。

＜4・1＞ SOMの形成 学習用のデータを用いてボタンの種類毎にSOMの形成を行った。SOM形成の条件としてM=N=15, d_0=6, α_0=0.15, T=30,000を使用した。

M, Nは図4に示したネットワークの出力層(OUTPUT)の大きさを決めるパラメータでその大きさに比例して得られる情報量(特徴量)が増える平面、計算に要する時間が増加する。双方のバランスを考慮してM=N=15を使用した。d_0は初期の近傍半径の大きさである。一般にM/3=5からM/2=7.5≈8以上の値が推奨されており(10)その範囲内であるd_0=6を使用した場合に図4に示す競合層の全てのユニットが学習時に反応し、学習に成功することをあらかじめ確認した。したがって本実験ではd_0=6を使用した。α_0は更新係数の初期値であり0<α_0<1の範囲で設定する必要がある。この値の大きさに比例して1回の更新量が大きくなるため、特定のデータによる影響(SOMの変動)が大きくなることが考えられた。そこで、我々の実験においてはα_0=0.15を使用した。Tはデータの提示回数である。1つのデータが約100回程度提示される値(100×3965/15≈26433)以上のT=30,000に設定した。

これらのパラメータは適切に設定しないと目的とする特徴を持ったSOMを得ることができないが、それほど個々の値に敏感に影響を受けることはない。例えば、計算回数Tを多少増減しても、初期更新係数α_0を増減しても最終的に得られる特徴地図であるSOMには大きな違いはみられない。

図9 "G-series"のSOM (L*-a*面)

Fig.9. An example of "G-series" SOM (L*-a* plane).

得られたSOMの一例を図9に示す。図9は緑色系ボタンのクラスに属する服地の色分布に関するL*a*a*面上のSOMの重みベクトルの輪郭を描いたものである。“G2”と“G4”および“G6”の3つのクラスで重複する領域(L*が32~36付近, a*が0付近), “G2”と“G4”の重複領域(L*が25~35付近, a*が-2~+2付近), “G4”と“G6”の重複領域(L*が35~41付近, a*が-2~+2付近)が見られる。紙面の都合上で総ての分布を見ることができないが、この様と同じ色系統のボタンで重複した境界領域が見られるのみでなく、異なった色系統のボタンとも重複した境界領域が存在した。1つの面のみでなく複数の面において重複が存在する場合も見られた。以上、SOMの観察から本問題は複雑な重複した境界領域を持ったクラス分類問題であることが確かめられた。

＜4・2＞ 評価パラメータの最適化 評価パラメータ $P_{out} = 0.0$, $P_{DIFF} = 1.0$ または $P_{DIFF} = 0.0$ を固定し, P_{in} , P_{CENTER} の変化に対する認識率への影響を調べ, パラメータの最適値を求める実験を行った。その結果を図10に示す。図10は上から $P_{in} = 0.20$, $P_{out} = 0.0$, $P_{DIFF} = 0.0$ 固定で P_{CENTER} (図中の縦軸ではCenterと表記) を変化した6例, $P_{in} = 0.20$, $P_{out} = 0.0$, $P_{DIFF} = 1.0$ 固定で

P_{CENTER} を変化した6例, $P_{in} = 1.0$, $P_{out} = 0.0$, $P_{DIFF} = 1.0$ 固定で P_{CENTER} を変化した6例, $P_{out} = 0.0$, $P_{DIFF} = 1.0$, $P_{CENTER} = 1.0$ 固定で P_{in} (図中の縦軸ではinと表記) を変化した8例の計26条件の結果が示してある。ここで、正しい認識というのは我々の分類システムにより得られた第3位までの解答の中に教師クラスと一致するものがあつた場合と定義する。

図中, P_{in} の値を変えて認識率を求めた下の8例において $P_{in} > P_{DIFF}/15$ の条件を満たす時にほぼ認識率が飽和し(86.7~86.8%), $P_{in} < P_{DIFF}/15$ の条件では認識率が飽和値より悪化していることが判明した。 $P_{DIFF}/15$ という値は(10)式で表される距離要素 D_{Lab}^{class} の値が全てのクラスで等しい事を意味する。以上の事からボタン選定問題において P_{in} は $P_{DIFF}/15$ 以上が良いという条件が得られた。

他の18例は P_{CENTER} の値を変えて認識率を見た結果である。いずれの場合にも P_{CENTER} の値を大きくすると認識率が悪くなることから、色の分布が重心値を中心に均一に拡がっていないことが予想される。つまり、あるボタンに属する服地の色分布が局在している事を表していると考えられる。

図10 評価パラメータ P_{in} , P_{out} , P_{DIFF} , P_{CENTER} の値を適当に変化させて求めた認識結果。最適解, 第2位, 第3位までに正しいボタンが得られたものを正解と定義した。

Fig.10 The results of recognition test at the combinations of the parameters P_{in} , P_{out} , P_{DIFF} and P_{CENTER} . Here, correct recognition is defined as following, that one of the best, second and third solution by the system is the same as teacher's class.

図 1 1 我々の方法とLVQ法の最適解に関する認識率の比較
Fig. 11. Comparison of best recognition between our method and LVQ.

また、 $P_{DIFF}=0.0$ の時（上の6例）の認識率が $P_{DIFF}=1.0$ の時（他の20例）の認識率より2～3%程度悪いことから P_{DIFF} は認識率において非常に有効なパラメータである事が判明した。 $P_{in}=1.0$ 、 $P_{out}=0.0$ 、 $P_{DIFF}=1.0$ および $P_{CENTER}=0.0$ とした時に最も高い認識率87.7%が得られた。

〈4・3〉 LVQ法との比較 本システムにおける最適解（1位の解答）とLVQ法による最適解（判定結果）の認識率に関する性能比較を行った。本システムにおいて評価パラメータの最適値 $P_{in}=1.0$ 、 $P_{out}=0.0$ 、 $P_{DIFF}=1.0$ 、 $P_{CENTER}=0.0$ の値を使用した。LVQはKohonenらの配布しているLVQ_PAK V2.1の中からLVQ1を使用した。計算条件はコードベクトル数500個、更新係数 $\alpha=0.05$ 、学習回数30,000回であり、その結果を図11に示す。これ以外にコードベクトル数を増やした場合および学習回数を増やした時の条件下においても計算したが、LVQ法による総合的な認識率は最大で約65%であった。この理由については後に判明したデータのクラス間での重複がその原因と考えられる。特に、図11でLVQ法による認識率が低い（50%以下）クラス“G2”、“G6”、“N4”においてはクラス間の境界領域の重複が大きいことが推測されるが、その検証は今後の課題である。現時点ではおこなっていない。

まず、個々のクラスについての認識率について本システムとLVQによる結果を比較する。“G2”、“G6”のクラスで本システムがLVQ法を15%程度上回る認識率が得られている。一方、“N4”、“N6”、“O2”のクラスでLVQ法が本シ

ステムを10～15%程度上回る認識率が得られている。残りのクラスに関しては両者ともほぼ同等の認識率を示している。

つぎに、全クラスの平均認識率（Total）は両者とも62%前後であり特に有意差がみられない。このことから我々のシステムでLVQと同等のクラス分類能力があることがわかった。

最後に、本システムの特長はLVQ法では困難な2位以降のクラスの順位付けが可能な点である。LVQ法ではクラス間の境界を明確にすることが主目的であり、基本的に2位以降のクラスの順位付けは考慮されていない。そのため、本研究の対象であるボタン選定問題などのようにデータ空間においてクラス間の重複領域が多いような場合にはクラスの分類能力が上がらないし、その原因を見いだすことが困難である。本システムではこのような場合において、2位以降の順位付けをおこなうことで認識率に影響するクラス間の境界領域で重複に関する帰属割合を表現することにより、その原因を見いだすための情報を提供している。この順位付けの正当さについて、ボタン選定問題に限ってはデータ提供者による検証をしていただいたところ大まかには合っているとの回答を得た。

5. むすび

我々はSOMを応用したクラス分類方法を提案した。この方法を実際に境界領域が重複した分布を持つボタン選定問題に応用し、その問題に最適な評価パラメータ（ $P_{in}=1.0$ 、 $P_{out}=0.0$ 、 $P_{DIFF}=1.0$ 、 $P_{CENTER}=0.0$ ）を決定した。また、

最適解のみについてLVQと比較したところほぼ同等の認識率が得られたことから、提案したクラス分類方法の基本的な分類能力が十分にあることがわかった。また、順位付けの結果も大まかに正しいことから、順位付けをもった分類結果を得るクラス分類法の1つとして十分に有効な方法であることを確認した。

(平成6年7月1日受付, 同6年11月2日再受付)

謝辞

資料を提供して頂いたグッドヒル株式会社の波田野史郎氏に感謝いたします。

文献

- (1) T. Kohonen : "The self-organizing map", Proc. of the IEEE, **78**, 9, pp.1464-1480 (1990)
- (2) T.Kohonen, J.Kangas, J.Laaksonen, K.Torkkola : "LVQ_PAK: A program package for the correct application of Learning Vector Quantization algorithms.", IEEE. Proc. of the IJCNN '92 Baltimore, pp.725-730 (1992)
- (3) G.A.Carpenter and S.Grossberg : "The ART of adaptive pattern recognition by a self-organizing neural network", IEEE Computer, **21**, 3, pp.77-88 (1988)
- (4) T. Kohonen : "Self-Organization and Associative Memory", Springer-Verlag, Berlin-Heidelberg-NewYork-Tokio, 3 edition, (1989)
- (5) H. Ritter and K. Schulten : "On the Stationary State of Kohonen's Self-Organizing Sensory Mapping", Biological Cybernetics, **54**, pp.99-106 (1986)
- (6) 田中利幸・斎藤正男 : 「Kohonenのモデルが実現する写像について」, 電子情報通信学会技術研究報告 NC90-107, pp. 231-236 (1990)
- (7) 藤村喜久郎・徳高平蔵・岸田悟・西守克己・石原永伯・山根興・石原誠 : 「ボタン選定問題へのkohonen自己組織化特徴マップの応用」, 電子情報通信学会技術研究報告 NC92-141, pp. 103-110 (1993)
- (8) 藤村喜久郎・徳高平蔵・岸田悟 : 「無限の服地色から15種類のボタン色を選定する問題へのKohonen自己組織化特徴マップの応用」, 電子情報通信学会技術研究報告 NC93-146, pp. 261-268 (1994)
- (9) 川上元郎 他 : 「色彩の事典」, 朝倉書店 (1987)
- (10) J. デイホフ : 「ニューラルネットワークアーキテクチャ入門」, 第9章, 森北出版 (1990)

藤村 喜久郎

(正員) 1962年9月9日生まれ。87年鳥取大学大学院修士課程修了。92年鳥取大学電気電子工学科助手, 現在にいたる。ニューロコンピュータの研究に従事。電気学会, 電子情報通信学会, 情報処理学会, 日本神経回路学会などの会員。

徳高 平蔵

(非会員) 1937年9月28日生まれ。60年大阪大学理学部物理学科卒業。同年松下電器産業(株)入社。66年退社後, イギリス, カナダに留学。72年鳥取大学工学部助教授を経て, 現在, 同学部電気電子工学科教授。理学博士。Ph. D. 研究歴として, 電子デバイス材料の開発, 基礎研究に従事。現在, 高温超伝導材料のニューロコンピュータへの応用を目指して基礎研究をし, ニューロコンピューティングに興味を持ち研究している。

岸田 悟

(非会員) 1952年8月26日生まれ。77年鳥取大学大学院修士課程修了。同年同大学工学部助手を経て, 現在, 同学部電気電子工学科助教授。神戸大学学術博士。酸化物高温超伝導体の作製と評価, およびニューロコンピューティングの研究に従事。