

氏名(本籍)	わいずみ ゆうじ 和泉 勇治	(栃木県)
学位の種類	博士(情報科学)	
学位記番号	情博第189号	
学位授与年月日	平成13年3月26日	
学位の要件	学位規則第4条1項該当	
研究科、専攻	東北大学大学院情報科学研究科(博士課程)システム情報科学専攻	
学位論文題目	ニューラルネットワークを用いた 高精度な手書き文字認識システムに関する研究	
論文審査委員	東北大学教授 根元 義章	東北大学教授 牧野正三
	東北大学教授 阿曾 弘具 (工学研究科)	東北大学助教授 加藤 寧

## 論文内容要旨

### 1 緒論

手書き文字認識は、計算機への自動文字入力を実現するための重要な研究課題である。現在ニューラルネットワークを用いた文字認識の研究が盛んに行われており、数字認識などの小規模の認識問題に対し、非常に高い認識精度が報告されている [1]。しかし、日本語文字認識のように、多くの字種を扱う大規模分類問題に対しては、学習速度や収束性問題から、不十分な認識精度しか得られていない。

本研究は、大規模分類問題に対するニューラルネットワークを用いた文字認識システムの問題点を明確にし、高精度な大規模認識システムの構築法の確立を目的とする。

### 2 階層的領域分割による大分類システム

#### 2.1 階層型学習ベクトル量子化の提案

階層型学習ベクトル量子化(Hierarchical Learning Vector Quantization:HLVQ)は、図1に示すように、複数の参照ベクトルを有するニューロンを階層的に配置した構成で、ニューロンを節点とした木を構成する。HLVQの参照ベクトルは、それぞれ目標の分類字種数に達していない場合、下の階層に特徴空間を更に分割するためのニューロンを各々一個有している。ネットワークの構築法として、ニューロンの階層的自己増殖、クラスタリングの重要なパラメータである参照ベクトル数の調整のために、参照ベクトルの死滅を行う。個々のニューロンの学習は、LVQ2.1[2]を用いる。

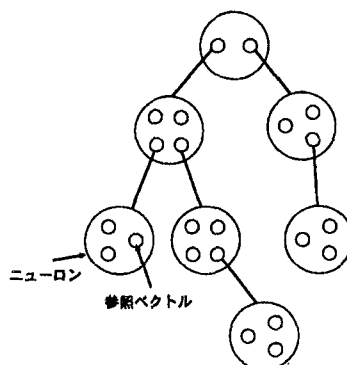


図 1: HLVQ の構成

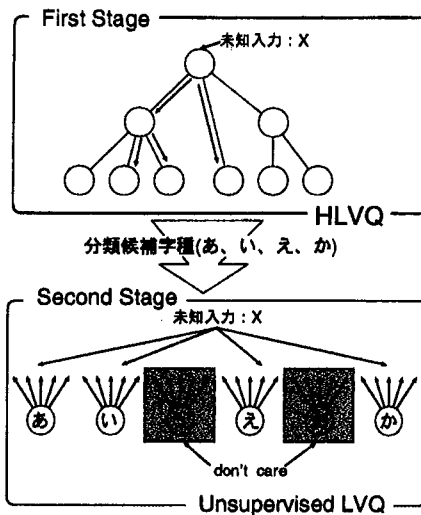


図 2: HLVQ を用いた大分類用ニューラルネットワーク

## 2.2 HLVQ の空間分割方法

HLVQ はベクトルの存在している特徴空間を、少数の参照ベクトルで階層的に分割して行く。少数の参照ベクトルで生成されるボロノイ領域は、非常に単純な境界面で構成されていると考えられる。そのため HLVQ では空間を分割する際、隣り合う領域間の境界付近に重なりを有して分割を行う。

## 2.3 HLVQ の分類アルゴリズム

分類時、未知入力 は HLVQ のルート・ニューロンへ投入され、HLVQ の木を葉に向かって探索を行いながら分類される。探索経路は木探索の途中、幾つかのニューロン内で分割される。探索経路へは、各ニューロン内で計算される帰属度を用いて得点付けが行われ、分割された経路の得点は減少して行く。また、未知入力のルート・ニューロンへの投入時、探索経路へ与えられるの得点の初期値は 1 である。探索経路の得点がある閾値未満になった時点で、その経路からの探索停止を行い、無駄な探索による分類候補字種数の増加と計算量の増加の抑制を行う。

## 2.4 ETL9B に対する分類実験

HLVQ の性能評価のために ETL9B3036 字種に対する分類実験を行う。大分類システムとしては、図 2 に示される HLVQ の下段に 1 字種 5 参照ベクトルの教師無し LVQ ネットワークを導入したものとする。比較手法として、平均ベクトルからのシテীবロック距離による分類手法を用いる。

表 1 に、上位 30 字種における分類率と分類速度(距離計算回数)を示す。表 1 に示されるように、本提案ネットワークは、平均ベクトルを用いたパターンマッチングと比較して高速・高精度な分類が可能である。

表 1: 上位 30 字種における分類率と分類速度

探索停止条件 StopScore	上位 30 位での 正解率	距離計算回数
0.88	99.83%	2326.67 回/文字
0.91	99.75%	1334.06 回/文字
0.93	99.60%	897.62 回/文字
0.95	99.05%	583.19 回/文字
平均ベクトル	99.65%	3036 回/文字

## 2.5 まとめ

本論文では、大規模分類問題に対し、高速・高精度な分類を実現するために新しいニューラルネットワークである HLVQ を提案した。ETL9B3036 字種分類実験では、従来手法の LVQ よりも高速・高精度な分類を実現し、本手法の有効性が確認できた。

## 3 局所的発火特性を実現するニューロンの結合法

文字認識問題の場合、各字種は特徴空間の局所領域に偏在していると考えられ、細分類部のモジュールは、特定の 1 字種にのみ反応する高い局所反応性が必要であると言える。本章では、従来の多層パーセプトロン [3] において、局所反応性の獲得を可能とする結合の提案を行う。

### 3.1 2乗結合多層パーセプトロン

提案する 2 乗結合を用いた多層パーセプトロンのネットワーク構成は、入力層を含めて 3 層構造で、隠れ層と出力層の間に伝達する信号を 2 乗する機能を持つ 2 乗結合を導入する。

順伝搬

入力ベクトルを  $\mathbf{x}$  とするとネットワークの出力  $o$  は以下の式で定義される。

$$h_i = \mathbf{x} \cdot \mathbf{w}_i + \theta_i \quad (1)$$

$$H_i = h_i^2 \quad (2)$$

$$O = \exp(-\mathbf{W} \cdot \mathbf{H} + \theta) \quad (3)$$

ここで、 $h_i$  は隠れ層ニューロン  $i$  の出力、 $\mathbf{H}$  は  $h_i$  を 2 乗した  $H_i$  を成分に持つベクトル、 $\theta_i$ 、 $\theta$  は bias ニューロンとそれに対する重みにより決定される値である。

逆伝搬

提案するニューラルネットワークの学習は、学習ベクトルとそれに対する教師信号の対を  $(\mathbf{x}^m, T^m)$ ,  $m = 1, \dots, M$  ( $M$  は定数) とした場合、式 (4) で定義される誤差に対し Back Propagation [4] を適用して行う。

$$E = \frac{1}{2} \sum_{m=1}^M (T^m - O^m)^2 \quad (4)$$

2 乗結合を導入したことにより学習則は次式のようにになる。

$$\begin{aligned} \frac{dW_i}{dt} &= -\alpha \frac{\partial E}{\partial W_i} \\ &= -\alpha \sum_{m=1}^M (T^m - O^m) O^m h_i^{m-2} \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{dw_{ij}}{dt} &= -\alpha \frac{\partial E}{\partial w_{ij}} \\ &= -2\alpha \sum_{m=1}^M (T^m - O^m) O^m W_i h_i^{m-1} x_i^m \end{aligned} \quad (6)$$

ここで、 $W_i$  と  $w_{ij}$  は、それぞれ、 $\mathbf{W}$  と  $\mathbf{w}_i$  の各成分である。また、式 (1)、(3) の  $\theta_i$ 、 $\theta$  は重み  $w_{ij}$ 、 $W_i$  の一成分として表現した。

### 3.2 文字認識への応用

提案手法を文字認識へ応用した際の認識精度を調べるため、ETL9B 平仮名 71 字種を対象に認識実験を行う。隠れ層ニューロン数、つまり、写像を行う軸の数を 5 個から特徴量の次元数である 196 個まで変化させ学習を行った。その時の認識率を図 3 に示す。隠れ層ニューロン数が 120 個の時最も高い 97.02% の認識率が得られている。

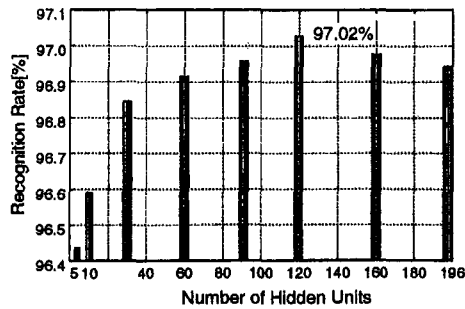


図 3: ETL9B 平仮名 71 字種認識実験:隠れ層ニューロン数と認識率の関係

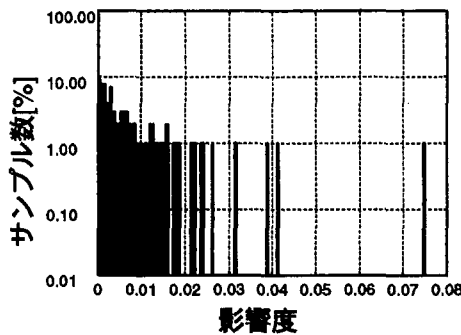


図 4: 「あ」のモジュールの発火用サンプルの影響度の分布

#### 4 サンプルの学習への影響度と汎化能力の関係を考慮した学習則

認識精度と学習サンプルのニューラルネットワークの学習への影響度の関係を考察し、高い汎化能力を獲得するための学習則の提案を行う。

##### 4.1 学習サンプルの学習への影響度

学習サンプル  $v$  の学習への影響度  $U_v$  を式 (7) によって定義する。

$$U_v = \sum_{t=0}^T \sum_{n \in N} |\Delta w_n|^2 \quad (7)$$

ここで、 $T$  は学習回数、 $n$  はニューラルネットワーク  $N$  に含まれるニューロン、 $w_n$  はニューロン  $n$  の重みベクトルである。式 (7) は、全学習過程において、学習サンプル  $v$  がニューラルネットワーク  $N$  に含まれるニューロン  $n$  の重みベクトルのどの程度変化させたかを定量的に表している。

図 4, 5 に、それぞれ、「あ」のモジュールの発火用サンプルと抑制用サンプルの影響度の影響度の分布を示す。図の横軸は式 (7) で定義された影響度を、縦軸はサンプル数を、それぞれの全学習サンプル数に対する割合を示している。図から分かるように、ほとんどの学習サンプルの影響度がゼロに近い値となっている。図に示した「あ」のモジュールの場合、影響度が 0.0004 未満のものが全体の 80% で、0.002 未満のものが全体の 98% となっている。最大の影響度は 0.075 であるので、全学習サンプルの 98% が最大の影響度の 38 分の 1 の影響度しか持っていないことになり、ネットワークの重みの決定には、全学習サンプルの 2% 程度しか寄与していないことを意味する。

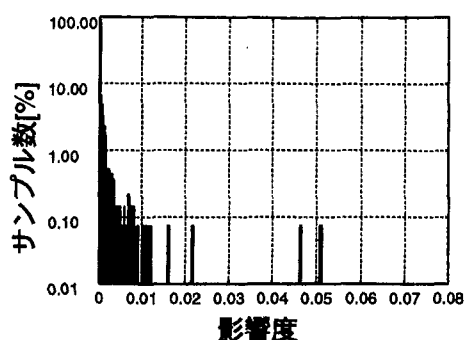


図 5: 「あ」のモジュールの抑制用サンプルの影響度の分布

#### 4.2 サンプルの学習への影響度を考慮した学習則の改良

文字認識問題のように、未知のサンプルに対する推定を行う場合、境界付近のサンプルに着目した学習よりも、学習サンプル全体の分布を捉え、境界付近の雑音や例外値などの影響を軽減すべきであると考えられる。そのため、従来の学習則よりも、サンプルの分布をより全体的に捉え、例外値などにロバストな学習則が必要であると言える。

式 (8) で定義される出力層ニューロンへの入力を小さくする項を誤差に付加することを提案する。

$$e = \frac{1}{2}(Tag - Out)^2 + \alpha H \quad (8)$$

ここで、 $Tag$  は教師信号、 $Out$  はネットワークの出力値、 $H$  は出力層ニューロンへの入力値、つまり、2乗結合によって2乗された隠れ層ニューロンの出力値をベクトル表現したものある。 $\alpha$  は、 $H$  を小さくする力の程度を表す定数である。

誤差へ出力層ニューロンへの入力値を小さくする項を付加したことで、隠れ層ニューロンの重み  $w$  の更新則は、以下のように定義される。

$$\frac{dw}{dt} = -\beta \left( \frac{\partial e}{\partial Out} \frac{\partial Out}{\partial H} + \alpha \right) \frac{\partial H}{\partial w} \quad (9)$$

$$= -\beta \left( -(Tag - Out) \frac{\partial Out}{\partial H} + \alpha \right) \frac{\partial H}{\partial w} \quad (10)$$

式 (10) の下線部は、(誤差) × (活性化関数の微分値) を表す項である。それに加算されている  $\alpha$  は、誤差がゼロとなる領域や、活性化関数の出力が平坦となる領域に存在している学習サンプルが投入された場合でも、学習が行われるように作用していると言える。

#### 4.3 文字認識への応用

提案した学習則の改良と  $\alpha$  の効果を調べるため、ETL9B 平仮名 71 字種認識実験を行い、 $\alpha = 0.75$  で隠れ層ニューロン数が 90 個のとき最も高い 97.90% の認識精度を得ることが出来ることが確認された。

### 5 大規模認識システムの構築と評価

本章では、前節までに提案した手法を用いて認識システムを構築し、ETL9B3036 字種の認識問題に適用することで、認識システムの構築・評価を行う。

その結果を、表 2 に示す。表から本提案システムが従来の認識システムより高い認識精度が得られていることが分かる。

表 2: ETL9B3036 字種認識実験

Method	Recognition rate
CombNET-II[5]	95.76%
ELNET-II[7]	95.84%
HoneycombNET[6]	96.00%
Alternative Method[8]	96.85%
Proposed Method	97.64%

## 6 結論

本論文では、大規模認識問題を対象に、ニューラルネットワークを用いた高精度な認識システムを構築することを目的とし、高速・高精度な大分類システムの構築、大規模認識問題に適したモジュラー型ニューラルネットワークモデルの構成、高い汎化能力を獲得し得る学習則の改良について研究を行い、大規模認識システムの構築を行った。その結果、従来よりも高い認識精度を得ることが可能となり、本研究により、大規模認識問題に適した認識システムの構築法を示すことが出来た。

## 参考文献

- [1] 井上・若林・鶴岡・木村・三宅, 自己想起回路による手書き数字認識, 情報処理学会論文誌, vol.39(8) (1998), pp.2476-2484.
- [2] T.Kohonen: "Self-organization and Associate Memory(2ndEdition)", Spring-verlag, pp.199-202 (1989).
- [3] Minsky M.L. and S.A. Papert(1969):"Perceptrons", Cambridge, MIT Press. Partially reprinted in Anderson and Rosenfeld[1988].
- [4] Rumelhart D.E.,Macclelland J.E. and PDP research group: "Parallel Distributed Processing",1,2 MIT Press, Cambridge, MA.(1986).
- [5] 猪野 浩, 岩田 彰: "Comb NET-II を用いた JIS 第 1 水準手書き漢字の認識", 信学技報, PRU93-151, pp.65-72 (1994).
- [6] 荒井 正之, 奥田 健三, 宮道 壽一: "Honeycomb ネット-II による多字種手書き漢字認識", 信学論 (D-II), J77-D-II, No. 9, pp.1708-1715 (1995-09).
- [7] Kazuki SARUTA, Nei KATO, Masato ABE and Yoshiaki NEMOTO, "High Accuracy Recognition of ETL9B Using Exclusive Learning Network-II(ELNET-II)," IEICE TRANS.INF. & SYST., VOL. E79-D, No.5, MAY 1996.
- [8] 石井 力, 和泉 勇治, 加藤 寧, 根元 義章: "ニューラルネットワークを用いた二者択一方式の手書き文字認識システム", 電子情報通信学会論文誌, VOL.J83-DII No.3 March 2000, pp.988-995

## 論文審査の結果の要旨

手書き文字入力システムはユーザフレンドリなマン・マシンインターフェースを提供し、かつ大量の文書を効率的にコンピュータに取込むことが可能であり、実用化を視野にこれまで多くの研究が行われてきているが、認識速度および認識精度の面で課題が残っている。著者は、ニューラルネットワークを用いた高速かつ高精度な手書き文字認識システムの実現を目指し、研究を行った。本論文はこれらの研究成果をまとめたもので、全編7章よりなる。

第1章は序論で、本研究の背景と目的を述べている。

第2章では、ニューラルネットワークを用いた数千字種におよぶ大規模認識システムを設計する際の問題点を整理し、認識システムは認識速度の速い大分類システムと認識精度の良い細分類システムで構成することが有用であることを述べている。

第3章では、階層的領域分割による大分類システムを提案している。学習サンプルの階層的なグループ分けを行い、グループ間の境界付近に重なる領域を持たせて、手書き文字のバラツキを吸収していることが本手法の特長である。これにより、約1/3程度の時間で従来法と同程度の分類率を達成できることを示している。

第4章では、2乗結合を用いた局所反応性に優れた多層パーセプトロンを提案し、細分類システムに適用している。認識実験により認識性能が大幅に向上することを確認している。これは、本手法が従来問題であった誤発火領域の抑制に有効であるという、著者の主張を実証したものであり、重要な成果である。

第5章では、個々のサンプルの学習への影響度を分析し、少数のサンプルがニューラルネットワークの重み決定に大きな影響を及ぼすことを明らかにするとともに、この分析結果に基づき、より多くの学習サンプルが重みの決定に利用される学習則を導いている。本学習則により、認識精度が向上すること、ニューラルネットワークが高い汎用性をもつことを認識実験により確認している。

第6章では、日本最大の手書き文字データベースを対象に、著者の提案した大分類システムと細分類システムを統合したシステム構築を行い、評価を行っている。その結果、これまでで最も高い認識精度を達成しており、高く評価できる。

第7章は結論である。

以上要するに本論文は、ニューラルネットワークを用いた高速かつ高精度な手書き文字認識システムの構築法を提案し、その有効性を実証したもので、システム情報科学およびニューラルネットワークを利用した認識工学の発展に寄与するところが少なくない。

よって、本論文は博士(情報科学)の学位論文として合格と認める。