

# ニューラルネットを用いた 手書き数字認識における特徴量削減について

新谷 素弘

吉川 敏則

長岡技術科学大学工学部電気系  
〒940-21 長岡市上富岡町1603-1

あらまし

ニューラルネットを用いて文字認識を行う際に、特徴量次元数が学習、認識時間に大きく影響する。本報告では、認識率を低下させずに、特徴量を削減する手法を提案する。提案法には白ランレングス特徴量を用い、この特徴量を用いた簡易細線化処理についても提案する。主成分分析を用いて、白ランレングス特徴量を新しい特徴量に変換し、その累積寄与率と特徴量次元数の関係を検討する。シミュレーションから、従来法と同等の性能を実現する累積寄与率を求め、それより、提案法の有効性を示す。

和文キーワード ニューラルネット、文字認識、特徴量削減、白ランレングス、主成分分析

## On the Reduction of Features for Hand-Written Numerals Recognition using Neural Networks

*Motohiro SHIN'YA*

*Toshinori YOSHIKAWA*

*Faculty of Engineering,  
Nagaoka University of Technology,  
Nagaoka-shi, 940-21, Japan*

Abstract

In the character recognition using neural network, the order of the feature gives the effect to the learning and the recognizing time. Without reducing the recognition rate, the method reducing the feature order is proposed. The proposed method is based on the White runlength feature. And the thinning algorithm is also proposed. Using the principle component analysis, the original feature is changed into the new one. The relation between the cumulative proportion and the new feature order is investigated. As a result, the cumulative proportion needed for the equivalent performance to the traditional method is shown. The effectiveness of the proposed method is shown with the simulations.

英文 key words neural network, character recognition, feature reduction, white runlength, principle component analysis

## 1 まえがき

近年、階層型ニューラルネット[1]による文字認識に関する研究が盛んに行われている。ニューラルネットを用いて文字認識を行う手法として、

- 入力層を2次元構成にして文字パターンをそのまま使う。
- 文字パターンから特徴を抽出し、その特徴量を使う。

の2種類がある。一般に、認識率が高いのは後者の特徴量を使う手法である。

しかし、このような特徴量はOCR技術で用いられているものが多く、特徴量の次元が高次となる。その結果、ニューラルネットの学習に膨大な手間と時間が必要となる。

そこで、ニューラルネットの学習を高速化し、更に、認識率を低下させない手法として、特徴量の次元削減について検討を行う必要がある。

本報告では、主に以下の2点について述べる。

1. 特徴量として白ランレングス[2]特徴量を用いること。
2. 特徴量次元削減法として主成分分析[3]を用いること。

白ランレングス特徴量は文字の外郭、内郭の情報を持つ特徴量である。そして、従来、マスク処理で行っていた細線化処理を、この特徴量のみで行うこと(以下、簡易細線化処理)が可能であるという特徴がある。また、多変量解析において主成分分析はよく使われる手法である。

対象とする文字は、比較的、変形度の大きい手書き数字とする。そして、文字パターンから白ランレングス特徴量を抽出し、それに主成分分析を行い、その結果から得られる特徴量を新たな特徴量として使用する。同時に、累積寄与率とその新しい特徴量の次元数との関係を調査し、ニューラルネットで学習させて、累積寄与率と認識率の関係を示し、その有効性について検討する。

本報告の構成は、次の通りである。2章で白ランレングス特徴量とその特徴量を用いた簡易細線化処理について述べる。3章では、特徴量削減の手法、ならびに累積寄与率と特徴量次元数の関

係について、4章では、各累積寄与率における学習時間、認識結果について述べる。

## 2 白ランレングス特徴量

### 2.1 特徴抽出法

図1に、シミュレーションで用いた手書き数字の一例を示す。大きさと位置の正規化を行った文字パターンにおいて、文字を構成している黒画素の部分を「文字部分」、それ以外の部分を「非文字部分」とする。この文字パターンを1画素単位で垂直、水平方向に走査し、最初に文字部分に到達するまでの非文字部分長を1次白ランレングス、そして続けて走査して、2度目に文字部分に到達するまでの非文字部分長を2次白ランレングスとする。また、1次白ランレングスを求めた後に出てくる文字部分長を黒ランレングスと定義する。図1(a), (b)はそれぞれ下方向から見た1, 2次白ランレングスを示す。

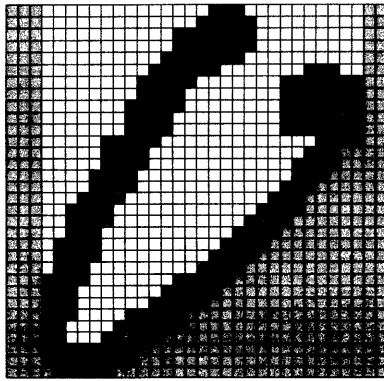
ただし、1次白ランレングスを求めるときに文字部分に到達しない場合は、文字枠の1辺の大きさ(この図の場合は32)を、また、2次白ランレングスの場合には0を与えることとする。これより、1次白ランレングスは文字の外郭、2次白ランレングスは文字の内郭を表すことになる。

### 2.2 簡易細線化処理

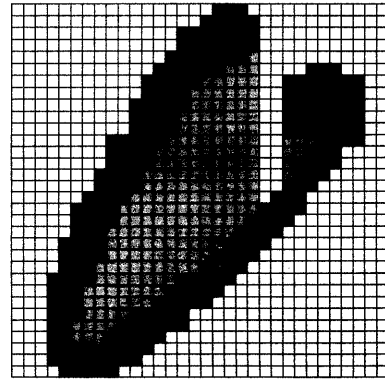
白ランレングス特徴量は2次までの特徴量を抽出するとき、黒ランレングスを抽出する。これを利用した簡易細線化処理について説明する。

1. 左方向から右方向へ水平に走査を行い、仮の白ランレングス、黒ランレングス特徴量を得る。
2. そのとき、黒ランレングスが2以上あり、1次、2次白ランレングスがあれば、文字の外郭から黒画素を2画素残し、そこから白画素を検出するまで黒画素を白画素にする。
3. 同様に右方向、上下方向は垂直に走査して簡易細線化を行う。

図2に、簡易細線化結果を示す。これより、前処理で行うべき細線化処理を特徴量抽出時に処理

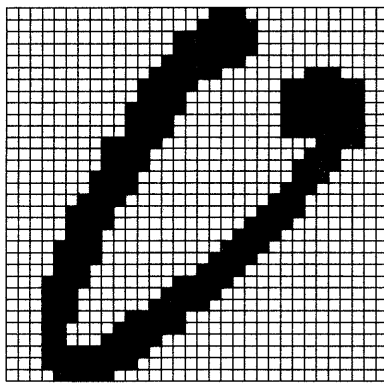


(a) 1次白ランレングス

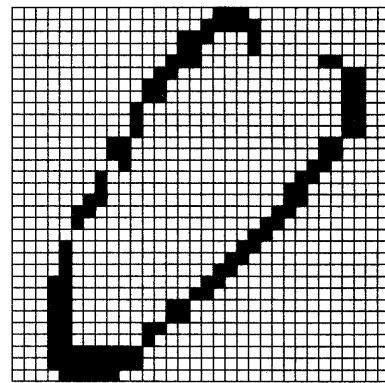


(b) 2次白ランレングス

図1 白ランレングス



(a) 処理前



(b) 処理後

図2 簡易細線化処理

できるので、前処理の時間が短縮可能である。しかし、この簡易細線化処理において、一部、文字欠けが起こる欠点がある。

### 3 特徴量次元の削減

#### 3.1 前処理

対象とする手書き数字は、郵政研究所のIPTP CDROM1の数字サンプル(400dpi, 256階調, 480×240(dot))を使用する。このサンプルには、郵便番号3桁がそのまま入っているため、均等に3分割する。その次に、2値化処理、大きさと位置の

修正を行う。このときの文字枠の大きさは32×32とし、文字の高さと幅から、数値の大きい方を文字枠に接するように拡大、縮小を行う。そして、その中から、人間が認識可能な数字を、各数字当たり80個を抽出する。そして、主成分分析、学習用に各数字20個を使用し、残りを検証用として使用する。

#### 3.2 主成分分析

学習サンプルに対して、主成分分析を行う。特徴量次元数を削減することで得られる効果は、次の通りである。

- ニューラルネットの規模を縮小する。
- 学習1回あたりの出力、誤差修正量の演算時間を短縮する。

その結果、ニューラルネットの学習を高速化することが可能である。このことは、文字認識時にニューラルネットの出力を得るまでの時間にも影響を与える。

文字枠の大きさを変えて特徴量の主成分分析を行った結果、もとの特徴量の次元が高くなっても、ある主成分までに寄与率が集中することが分かった。このことから、高次元の特徴量に主成分分析を用いて次元数を削減することは、高次になるにつれて特徴量次元削減率が高くなり、その結果、ニューラルネットの構成も縮小することが可能である。

本報告では、全数字の各方向、各次白ランレングス特徴量ごとに、主成分分析を行う。これは、主成分係数行列を求めるのに、行列演算を多用するため、高次の行列演算を行うことは、速度的に問題があるためである。その結果、左右上下の1, 2次白ランレングスに関する、8個の主成分係数行列が生成される。そして、次式で白ランレングス特徴量を変換する。

$$\begin{aligned}
 \mathbf{x}_{pri} &= \mathbf{T}_{direction} \mathbf{x}_{run} \\
 \mathbf{x}_{pri} &: \text{主成分} \\
 \mathbf{x}_{run} &: \text{白ランレングス} \\
 \mathbf{T}_{direction} &: \text{主成分係数行列}
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{x}_{pri}^T &= [x_{pri1}, x_{pri2}, \dots, x_{pri32}] \\
 \mathbf{x}_{run}^T &= [x_{run1}, x_{run2}, \dots, x_{run32}]
 \end{aligned}
 \tag{1}$$

上下左右, 1, 2次

各方向、各次白ランレングス特徴量ごとに、累積寄与率がそれぞれ70%, 75%, 80%, 85%, 90%, 95%を越えるときの主成分の個数を算出する。表1に、各累積寄与率における主成分数を示す。表より、2次白ランレングスに関係する主成分数が1次よりも多くなっていることが分かる。これは文字の特徴を、文字の外郭よりも内郭の方に注目していることを意味している。また、合計の特徴量次元数より、累積寄与率70%から80%までは単調増加するが、それ以降は指数関数的に増加することが分かる。

さらに、累積寄与率と相関係数からの主成分分析の固有値との関係を見ると、累積寄与率が

表1 各累積寄与率と特徴量次元数の関係

	70%	75%	80%	85%	90%	95%
下→上1次	4	4	5	7	9	14
下→上2次	5	7	8	10	13	18
左→右1次	4	5	7	10	13	19
左→右2次	5	7	9	11	15	20
右→左1次	5	6	8	10	13	19
右→左2次	5	7	9	11	15	20
上→下1次	4	5	6	7	10	16
上→下2次	6	7	8	10	13	18
合計	38	48	60	76	101	144

表2 シミュレーション条件

学習回数	10,000回
安定化定数 $\alpha$	0.9
学習定数 $\eta$	0.25
エラーリミット	$10^{-4}$
初期値幅	-1.0~1.0

80%のときの特徴量次元数が60に対し、固有値が1.0以上となる個数が53と類似していることが分かった。

## 4 シミュレーション

### 4.1 学習

3.2で得られた特徴量を用いて、ニューラルネットを学習する。比較のため、従来法についても学習を行う。ニューラルネットの構成は、予備実験の結果、中間層と出力層は従来法、提案法ともに固定で、それぞれ18, 10とする。また、入力層ユニットは特徴量の次元数に合わせて変化させる。従来法の入力層ユニット数は256である。各パラメータを、表2に示す。また、学習停止を防止するために、シグモイド関数の出力を0.01~0.99に制限した。

図3に提案法と従来法の学習誤差曲線を示す。これは、初期値8例中、認識率が最良のものであ

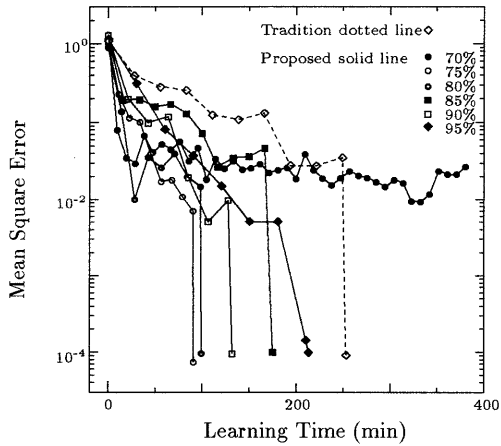


図3 累積寄与率と学習時間の関係

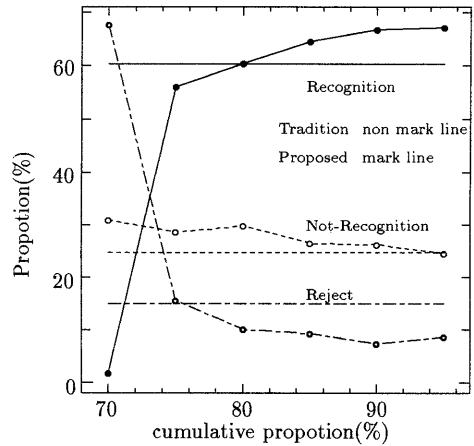


図4 累積寄与率と認識結果の関係

表3 平均, 最良, 最悪学習時間 (min)

		平均	最良	最悪
提案法	70%	196	190	198
	75%	91	19	178
	80%	113	53	205
	85%	143	70	190
	90%	185	101	382
	95%	227	141	326
従来法		447	251	923

る。ここで、入力となる次元数が異なるので、横軸は学習時間とする。これより、累積寄与率70%以外は従来法より速く収束することが分かる。さらに、表3に提案法、従来法に対する初期値8例の平均、最良、最悪学習時間を示す。

表より、平均学習時間においては、従来法の約2~5倍の速さで収束し、また、最良のときは、約12倍という結果が出ている。

## 4.2 認識

図4に認識結果を示す。これは、初期値8例の平均である。ただし、認識、誤認識、リジェクトは次のように定義する。

**認識:** ニューラルネットの出力が0.85より大きくなる出力層が唯一あり、入力パターンの希望

出力がネットワークの出力と一致するとき

**誤認識:** ニューラルネットの出力が0.85より大きくなる出力層ユニットが唯一あるが、入力パターンの希望出力がネットワークの出力と一致しないとき

**リジェクト:** 認識、誤認識の判定にあてはまらないとき

従来法と同等の認識率を得るためには、累積寄与率が最低80%以上必要であることが分かる。しかし、累積寄与率が増加することで認識率が向上するが、誤認識率はほとんど低下していない。

表4に、初期値8例の平均、最良、最悪認識率を示す。これより、累積寄与率70%のときは、学習が収束していないので、認識率は悪い。しかし、累積寄与率が大きくなると平均とともに、最良認識率が向上している。

また、認識率においては他の手法[4], [5]と比較して、かなり低下している。これは、学習サンプルの影響が強いものと思われる。

## 5 むすび

ニューラルネットの学習高速化のための、特徴量次元数削減のための手法について検討を行った。提案法では、前処理でマスク処理で行われる細線化処理を、白ランレングス特徴量を用いて行

表4 平均, 最良, 最悪認識率(%)

		平均	最良	最悪
提案法	70%	1.6	4.8	0.1
	75%	56.0	61.8	52.2
	80%	60.3	62.5	57.0
	85%	64.4	66.0	62.8
	90%	66.7	70.8	62.8
	95%	67.1	73.2	62.6
従来法		60.3	63.3	57.8

う。そして、この細線化処理を行った文字パターンに対し、主成分分析を行い、新たな特徴量とすることで次元を削減する。

簡易細線化処理では、前処理で行うべき細線化処理を特徴量抽出時にできるので、前処理時間を短縮することが可能である。しかし、文字の一部が欠けてしまう欠点があるため、今後の改良が必要である。

シミュレーションでは、学習時間において、提案法は従来法より平均で20%~50%高速になることが分かった。また、認識結果において、提案法は従来法と同等かそれ以上の結果が得られた。

さらに、特徴量次元数削減の目安として、累積寄与率が80%以上という結果が得られた。そして、ニューラルネットの入力層ユニットの個数の推定として、学習サンプルの相関係数を用いた主成分分析で、固有値が1.0となる主成分数を使うことができる。

しかし、他の手法と比較して、認識率が低いことが欠点としてあげられる。

今後の課題として、学習サンプルの検討と、更にサンプル数を増加した場合において、この手法の有効性を調べ、また、簡易細線化処理の改良の検討を行う予定である。

## 謝辞

手書き数字は郵政研究所のIPTP CDROM1のデータを使用しました。データの御提供に感謝いたします。

## 参考文献

- [1] M. J. Rumelhart D.E. and the PDP Research Group: "Parallel Distributed Processing", The MIT Press, Cambridge (1986).
- [2] 加藤誠巳, 水谷昭弘: "ランレングスを主入力とするニューラルネットワークによる手書き数字認識", 平2情処全大, pp. 2-125-126 (1990).
- [3] 奥野忠一他(編): "続多変量解析法", 日科技連出版社(1976).
- [4] 勘解由哲, 大西昇, 杉江昇: "ニューラルネットによる回転と大きさに対して不変な手書き数字認識", 信学論(D-II), **J75-D-II**, 11, pp. 1900-1908 (1992).
- [5] 山田敬嗣, 津雲淳, 天満勉: "学習停止回避型学習アルゴリズムを用いたニューラルネットによる文字認識", 信学論(D-II), **J75-D-II**, 11, pp. 1798-1808 (1992).