

3X-01

## テンプレートマッチングによる文字認識ニューラルネットワークの作成と評価\*

篠沢 佳久†

慶應義塾大学 インフォメーションテクノロジーセンター‡

大駒 誠一§

慶應義塾大学 理工学部 管理工学科¶

### 1 はじめに

現在文字認識の分野においては、多くの研究者によってさまざまな認識方法が提案されているが、最も基本的な方法は、認識したい未知文字の特徴と認識対象となるすべての標準文字の特徴との距離を一つ一つ求め、最も距離の短いものを認識結果とする総当たりのテンプレートマッチング方式である。

本報告においてはこのテンプレートマッチング方式をニューラルネットワークで実現する方法を提案する。すなわちある任意の文字の代表となる特徴と認識したい未知文字の特徴を入力層に同時に入力した時、この二つの文字が同じ場合は出力層が発火し、異なる場合は発火しないような挙動をさせることによって、認識を行なうニューラルネットワークを構築する。

本報告では、こうしたニューラルネットワークの構成、その学習方法、および ETL 9B を用いた認識実験について述べる。

### 2 構成

提案するニューラルネットワークはテンプレートマッチング方式と同様に特定のニューラルネットワークはそれと対応した一種類のみの文字の認識を行なう。従って認識対象数を  $N$  個とすると、同数の  $N$  個のニューラルネットワークが必要となる。このように一つのニューラルネットワークの認識対象を一文字とする考え方には、ELNET(排他的学習ネット)[1]によって実現され、その有効性も確認されているため問題はないはずである。例として図1に『あ』を認識対象としたニューラルネットワークを示す。一つ一つの

ニューラルネットワークは3層から構成される。入力層には、認識したい未知文字の特徴と比較対象となる文字の特徴を同時に入力する。従って入力層のニューロンの個数は特徴の総数分、必要となる。中間層のニューロンの個数は任意である。出力層のニューロンの個数は1個である。

そして入力層の比較対象部には標準となるテンプレートの『あ』の特徴を入力し、認識対象部には認識したい未知文字の特徴を入力し、入力した二つの文字の特徴が類似している場合は、出力層のニューロンが発火し、異なった場合は発火しない。

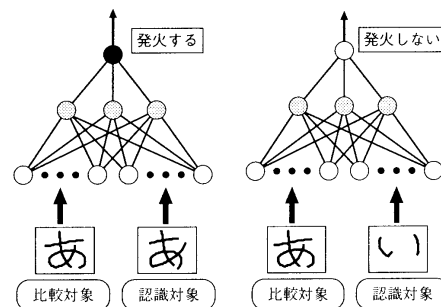


図1: ニューラルネットワークの構造

### 3 学習方法

最も重要なのは学習の方法である。認識対象数が  $N$  で、任意のニューラルネットワークの学習を行なう場合、教師信号はそのニューラルネットワークと対応させた文字1個であり、非教師信号は残り  $N - 1$  個の文字となる。しかし非教師信号として残り  $N - 1$  個の文字を学習に使用するの、非効率である。教師信号となる文字と類似している、すなわち特徴空間上で近い位置関係にある文字のみを非教師信号として使用すればよい。従ってあらかじめユークリッド距離などを用いて大分類を行ない、その上位に挙がった文

\*The Neural Networks for Character Recognition by Template Matching.

†Yoshihisa SHINOZAWA

‡Information Technology Center, Keio University

§Seiichi OKOMA

¶Faculty of Science and Technology, Keio University

字候補を非教師信号の候補として使用すればよい。

次にそのような類似した文字のみを使用してどのように学習させていけばよいかを論じる。例えば、『あ』『い』『う』『え』『お』の5文字を認識対象とした5個のニューラルネットワークを考える。仮に『あ』を認識させようとして、各ニューラルネットワークの出力層の発火の状況から『おえあうい』というような順番の認識結果になってしまった場合、

- 『お』および『え』を対象としたニューラルネットワークは『あ』を入力すると発火しやすいことから、非教師信号として『あ』を使用する。
- 『う』および『い』を対象としたニューラルネットワークは『あ』を入力しても発火しづらいことから、非教師信号として『あ』は使用しない。

といったようにあらかじめサンプル用の学習文字を使用して各ニューラルネットワークにおける非教師信号の対象を決めていけばよい。他の文字に関しても同じように行ない、各ニューラルネットワークにおける非教師信号の対象を決める。学習方法は以下の通りである。個々のニューラルネットワークの学習にはバックプロパゲーションアルゴリズムを用いた。

1. 大分類の上位候補からランダムに非教師信号を選び、学習する。

```
for( int i = 0 ; i < N ; i++ ){  
    第 i 番目のニューラルネットワークの  
    非教師信号をランダムに複数個選択  
    第 i 番目のニューラルネットワークの学習  
}
```

2. 教師信号の個数に変化がなくなるまで、認識、非教師信号の選択、学習を続ける。

```
while( 教師信号の個数の変動 < t ){  
    認識を行ない、非教師信号を決定する  
    for( int i = 0 ; i < N ; i++ ){  
        第 i 番目のニューラルネットワークの学習  
    }  
}
```

## 4 認識実験

電子技術総合研究所提供の ETL9B、3036 文字を利用し認識実験を行なった。一文字について学習文字として最初の一セット 40 パターンを利用し、認識実験で利用する未学習文字として二セット目の 40 パターンを利用した。

文字パターンには非線形正規化処理を施し  $48 \times 48$  の大きさにした後で、196 次元の輪郭線特徴を抽出した。実験は大きく分けて二通り、ユークリッド距離そして重み付けユークリッド距離によって大分類した第  $P$  候補までの認識結果から非教師信号を決定した。

実験には大分類の結果から第 10 候補 ( $P = 10$ ) もしくは第 20 候補 ( $P = 20$ ) までに出現する全ての文字を非教師信号とする方法、もしくは第 10 候補内で正解となった順位からさらに下位二つの候補までを用いた提案方法を行なった。

ニューラルネットワークの構造は入力層のニューロンは 392 個、中間層は 16 個、出力層は 1 個と固定した。学習の高速化のため入力層は 8 個の局所領域に分割し、入力層と中間層は局所結合化した。学習方法はモーメント法で学習時のパラメータの設定、学習回数は 100 回と全てのニューラルネットワークで同一とした。以上の条件で行なった認識結果を表 1、表 2 に示す。各表には一文字あたりの非教師信号の個数、学習文字の認識率、未学習文字の認識率を示す。

表 1: 大分類にユークリッド距離を使用した場合

	非教師信号	学習文字	未学習文字
大分類	—	94.95%	92.17%
提案方法	61.93	99.05%	92.97%
$P = 10$	81.06	99.08%	93.00%
$P = 20$	148.87	99.35%	93.04%

表 2: 重み付けユークリッド距離を使用した場合

	非教師信号	学習文字	未学習文字
大分類	—	97.92%	93.10%
提案方法	61.10	99.44%	93.52%
$P = 10$	82.18	99.44%	93.52%
$P = 20$	151.17	99.57%	93.57%

実験結果から、ニューラルネットワークを用いた場合、いずれの方法でも大分類での認識結果を上回った。さらに下位候補までを非教師信号として利用しても大きく認識率の向上に影響しないことも分かり、提案する学習方法では少ない個数の非教師信号を利用して効率的に認識率の向上に役立つことが分かった。

## 参考文献

- [1] 猿田 加藤 安倍 根元:排他的学習ネット (ELNET) を用いた手書き文字認識の細分類手法, 信学論 (D-II), Vol.J79-D-II, No.5, pp.851-859 (1996年5月).