

# ニューラルネットワークを用いたオンライン手書き文字認識 類似文字識別の一手法

1L-1

豊川和治, 藤田雅之†, 加藤 真

日本アイ・ビー・エム(株) 東京基礎研究所, (†)公共営業統括本部

## 1. はじめに

ニューラルネットワークをパターン認識に適用すると、識別するカテゴリー数があまり大きくない場合には、実現可能な学習プロセス・時間で、精度の高いパターン識別機能が、単一のネットワークで実現出来る [1]。しかしこの方法は、日本語の手書き文字認識のように多数のカテゴリーのパターン識別に適用した場合、実用に耐えるパターン識別機能を実現することは困難である。カテゴリー数の多いパターン識別で、従来から取られてきた方法は、全カテゴリーをまずいくつかのクラスに大分類した後、クラスごとに別のニューラルネットワークを用い、詳細識別を行う方法である [2,3]。大分類で選ばれたいくつかのネットワークの中で、最大の出力を与えるカテゴリーを正解候補とするものである。しかしこの方法は、学習が限られた数のデータで行われるため、必然的に各ネットワークに学習の程度の違いを生じ、複数のネットワーク間で単純に出力の大きさを識別結果を判定することに問題があった。

この問題を解決するための新しいニューラルネットワークの構成方法として、各クラスのネットワークを学習させた後、さらに各出力をそのカテゴリーが正解である尤度に変換することを考案した。この方法をオンライン手書き文字の類似文字識別に適用したところ、良好な結果が得られたので報告する。

## 2. オンライン手書き文字認識システムの構成

実験に用いたオンライン手書き文字認識システムの構成を図1に示す。タブレットより入力された筆記データはサイズの正規化、画数・ストローク重心位置情報などによる正解候補の予備選抜後、大分類部と、詳細識別部の構造特徴抽出ブロックへ入力される。大分類部は、筆記データの(x, y)座標を1ストロークあたり6点サンプルし、テンプレートとのパターン・マッチングによって正解候補を選ぶ。一位候補の距離で正規化した距離で域値内に残った文字カテゴリーを候補として残す。テンプレートは、常用漢字と非漢字、計2,384字をサポートしている。一方、テンプレート作成に使用したデータ(1文字当たり、約100サンプル)を用いて、正規化距離に対する正解確率を求めた。実験の結果、正解確率はあまり文字種に依存せず、正規化距離に対する一定の関数となる。

詳細分類部は'レ'と'ル'のような類似のパターンを持つ文字の識別に主眼を置いている。このような類似文字は、低画数の漢字及び非漢字に集中しているため、ここでは3画以下の334カテゴリーを詳細識別の対象とした。4画以上の文字は大分類の一位候補を正解とする。構造特徴抽出ブロックは、点や円弧の有無・方向、ストロークの開始・終端方向などの文字ストロークの形状情報の他、前後のストロークの相対位置などの情報を量子化し合計42成分の特徴ベクトルとする [4]。詳細識別部は、構造特徴抽出の後、1文字カテゴリーに対し1つのネットワークで構成され、各ネットワークの、入力・中間・出力層は夫々42, 6, 1ユニットである。出力ユニットの後段に、尤度変換ユニットが設けられている。

大分類部で選択された候補カテゴリーに対応するネットワークに、上記の特徴ベクトルが入力され、各カテゴリーが正解である尤度が出力される。一方、大分類部で得られた各候補の正解確率を事前確率とし、各候補の尤度とベイズ則を用いて事後確率を計算し、最も確からしいカテゴリーを正解とする。

## 3. ニューラルネットワークおよび尤度変換器の学習方法

(a) 教師データの作成方法: この類似文字識別部で対象とする334文字カテゴリーごとに、学習に用いる手書きデータ、約100サンプル/カテゴリーを用意する。このとき、次に示す手順で教師パターンデータを作成する。(1) 334文字のデータを前段のパターン・マッチング部に入力し、正規化距離が域値内にある候補文字を残す。(2) 候補文字コードで、入力パターン本来の文字コードと一致する場合を'正解'とし、それ以外の文字コードはすべて'不正解'と呼ぶ。各文字の100学習サンプルに対して、'不正解'文字コードのリストを作成する。(3) 各334文字に対して、元の学習用の手書きデータ100サンプルをとり

Verification of Online Handwritten Japanese Characters by Neural Network

Kazuharu Toyokawa, Masayuki Fujita†, Shin Katoh

Tokyo Research Laboratory, IBM Japan Ltd.

1623-14, Simotsuruma, Yamato, Kanagawa 242, Japan. (†) Public Sec. Ind. Mktg., IBM Japan Ltd.

正解パターンデータとし、'不正解'リストにある各文字のサンプル各々 100サンプルをとり不正解パターンデータとする。(4) 上記で作成した、各334文字ごとの正解・不正解パターンデータを、各ニューラルネットワーク学習のための教師データセットとする。

(b) ニューラルネットワークの学習方法： 上記の方法で集められた各文字ごとの正解・不正解パターンデータから、2.で述べた特徴ベクトルを抽出し、各文字毎に設けられたニューラルネットワークを学習させる。すなわち、正解パターンデータと不正解パターンデータを交互に取り出し、正解から取り出したときには教師信号として 0.9、不正解から取り出したときには教師信号として 0.1を3層目の出力層に与えBPにより学習を行う。十分な回数学習を行った後、正解・不正解パターンデータそれぞれに対する、教師信号と3層目の出力信号の自乗誤差の平均がともに十分小さくなったときに学習を終了する。

(c) 尤度変換テーブルの作成方法： 学習が終わった各ニューラルネットワークに、各文字ごとの正解・不正解パターンデータを別々に与え、3層目の出力の分布のヒストグラムを取り、出力に対する尤度変換テーブルを作成し、尤度変換ブロックの記憶領域に格納する。

4. 認識実験の結果

この実験では、未知の入力パターンに対して、大分類部で選ばれた複数の候補文字に対して、(a)従来法の尤度変換器をもたないニューラルネットワークにおいて、最大出力を与える文字カテゴリーを正解とした場合、(b)今回提案の尤度変換器を用いて、各ネットワークの出力を補正し最大値を与える文字カテゴリーを選んだ場合、さらに(c)応用として、大分類部で選ばれた各候補文字に対する正解確率と、各ネットワークの尤度との総合判定を行った場合の一位正解認識率を測定した。

実験データは、予め 141人の被験者より収集した手書き文字データ 2,384カテゴリー / 約 240,000 個の文字データのうち、認識テスト用に20人分 約32,000 個を除き、学習用データとして、パターン・マッチング用のテンプレートを作成した。またその中から、3画以下のデータ(334カテゴリー)を類似文字識別部の学習データとし、上記 3. で述べた方法で、334個の各ニューラルネットワークと尤度変換テーブルを学習させた。パターン認識部・類似文字識別部を含むオンライン手書き文字認識装置は、PS/55 ワークステーション上にソフトウェアで実現した。

実験結果を表 1 に示す。下記のように、従来方式(1)に比べて、今回提案の尤度変換を用いる方式(2)、(3)は類似文字の影響で認識率の低下が著しい英小文字、数字、記号に対して顕著な改善が認められ、本方式の有効性が確認された。

	(a) NN 尤度変換ナシ	(b) NN 尤度変換	(c) PM+NN 総合判定	認識率向上	
				(b) - (a)	(c) - (a)
ひらがな	87.5%	88.3%	89.5%	+0.8%	+2.0%
カタカナ	84.5%	84.7%	89.3%	+0.2%	+4.8%
英 大文字	87.0%	84.4%	88.2%	-2.6%	+1.2%
英 小文字	81.5%	83.1%	86.5%	+1.6%	+5.0%
数字	78.0%	87.3%	88.3%	+9.3%	+10.3%
記号	79.2%	82.3%	86.2%	+3.1%	+7.0%
総合(*)	84.2%	85.2%	88.4%	+1.0%	+4.2%

表 1. 類似文字識別による認識率評価結果 (\*上記6字種の総合評価)

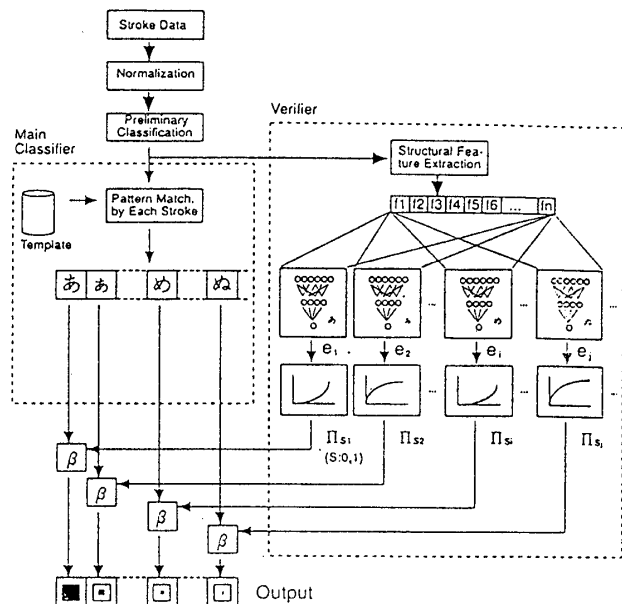


図 1. オンライン手書き文字認識システムの構成

参考文献

[1] I.Guyon et al.: Pattern Recognition, Vol.24, pp.105-119, 1991.  
 [2] 小島良宏 他: 90年電子情報通信学会全国大会予稿集 6-5.  
 [3] 岩田 彰 他: 電子情報通信学会論文誌 D-II Vol.73-D-II, pp.1261-1267, 1991.  
 [4] 加藤 真 他: 電子情報通信学会論文誌 D-II, to appear.