ニューラルネットワークによる手書き漢字生成の一手法*

張 貽彤**・重田 和夫**・志水 英二**

An Approach to Handwritten Chinese Character Generation Using Neural Networks*

Yitong ZHANG**, Kazuo SHIGETA** and Eiji SHIMIZU**

The purpose of this study is to construct a new model for handwritten character generation using neural networks. This model can generate not only trained characters as the handwritten-character-generation-models proposed formerly, but also untrained characters by combining radicals. The proposed model uses the neural networks with back-propagation training for generating the radicals of characters and also for estimating the features of untrained characters which signify the peculiarities of each individual handwriting. The basic experiments verified the validity of the model.

1. ま え が き

近年、パソコンやワープロなどのOA機器は急速に普及し、多くの事務書類や個人の手紙などもOA機器により印刷されるようになった。印刷された文字の書体の多くは明朝体である。読みやすさという実用的観点からは明朝体文字は適切であると思われるが、クールで事務的な印象を与える。ユーザのニーズの多様性や機器の多機能化や高機能化に伴い、より人間性のある文字、毛筆体や手書き風文字の生成手法も提案されるようになってきた^{6),7)}。

これまでに、手書き文字を解析および生成するため、様々な手書き文字再生モデルが提案されてきた^{1)~5)}。しかし、それらは、書字過程の関節運動のシミュレーションや手書き動作機構をモデル化したもので、学習文字しか再生できない。また、そのほとんどは、筆圧、運筆速度、加速度などの情報を利用したものである^{1)~4)}。

一方、人間は、書かれた文字を基に、その文字の形態などから個人の特徴を掴み、文字を真似ること、再生することができる。このことは、文字の形態に十分な個性情報が存在していることを意味し、これらの筆者固有の特徴を用いることによって、筆者らしい文字を再生することができると思われる。

本研究においてはこのような観点から、運筆速度などの情報を使わずに、手書き文字の形態情報（ストロークの位置、幅）を用い、部首を基本要素として、筆者らしい手書き文字生成を行う手法を提案する^{8)~10)}。

本手法は、従来の筆跡モデル及び文字生成手法と異なり、学習文字のほかに、学習した部首の組み合わせによる未学習文字も生成できる特徴を持つ。

2. 部首生成モデル

漢字は部首により構成されるので、文字認識の分野でもそれを階層的に利用して認識を行う手法が種々提案されてきた^{14),15)}。本論文においても、例えば“樹”は“木”と“尌”の二つの部首で構成されるように、文字が二つの部首により構成されるとする。したがって、部首間結合形態は三つに分類でき、それぞれは、左右（例：“河”、“住”など）、上下（例：“字”、“守”など）と内外結合（例：“団”、“聞”など）となる。部分ストローク数の少ない文字、例えば“人”、“二”などについて、その文字は一つの部首とし、もう一つの部首は空と考える。

そして、本手法では、まず、部首生成モデルによって部首を生成し、つぎに、生成された部首をさらに調整して、より手書き文字らしい文字を生成する。この手法の長所は、モデルの構成要素の全体量を軽減できるほかに、未学習文字についても部首の調整によって手書きらしい文字を生成できることである。

部首生成モデルでは、サンプル点を通る滑らかなストロークの生成が要求される。このために従来は、サンプル点を補間して曲線を再生する手法としてスプライン法

* 原稿受付 1993年12月10日

** 大阪市立大学 工学部 Faculty of Engineering, Osaka City University ; 3-138, Sugimoto 3-chome, Sumiyoshi ward, Osaka city, Osaka 558, JAPAN

Key Words : handwritten character generation, radical, feature estimation, back-propagation, neural networks.

が使われてきた^{3),11)}。

一方、バックプロパゲーション学習法に基づく多層ニューラルネットワーク（以下でBPモデルと呼ぶ）に関しては、3層以上のBPモデルによって、任意の N 次元空間から、任意の M 次元空間への連続写像が可能であることが証明されており¹²⁾、これらの知見から、滑らかな曲線、すなわち、ストロークの生成に3層BPモデルが利用できると考えられる。

今、ある部首が N 本のストロークで構成されるとする。この部首を生成する部首生成モデルは、入力層ユニット数1、中間層ユニット数30および出力層ユニット数 $3N$ のBPモデルである。学習の場合、 N 本のストロークをそれぞれ m 等分割し、学習パターン対 (d_i, T_i) ($0 \leq i \leq m$)を与えて行う。入力信号 d_i はストロークにおける i 番目サンプル点とストロークの始点間の相対距離

$$d_i = \frac{\text{始点とサンプル点 } i \text{ 間の長さ}}{\text{ストロークの長さ}}$$

である。各ストロークを同じように m 等分割するので、各ストロークにおける始点から i 番目サンプル点の相対距離も等しい。 T_i は教師信号である。

$$T_i = (t_{i1}, t_{i2}, \dots, t_{ij}, \dots, t_{iN})^T$$

ただし、 $t_{ij} = (X_{ij}, Y_{ij}, R_{ij})^T$ である。 X_{ij}, Y_{ij}, R_{ij} はそれぞれストローク j の i 番目サンプル点における x, y 座標とストロークの幅である。

学習後、部首生成モデルの入力層ユニットに相対距離 $d_i = i/n$ ($0 \leq i \leq n, n > m$)を入力すると、この d_i における N 本のストロークの x, y 座標と幅が出力層ユニットの出力として得られる。そして、これらの x, y 座標を中心、幅を直径とする N 個の円をプロットすると、この相対距離における各ストロークの位置と幅が再生される。このようにして、 n 個の相対距離を入力すると、各ストロークがそれぞれ n 個の円の重なりによって形成される。

Fig. 1 (a)は部首“可”の原画像、(b)は本手法の部首生成モデルにより生成した結果である ($m=10, n=100$)。

Fig. 1 Reconstruction of radical “可”

3. 特徴推定モデル

手書き文字の再現は、本方式では生成された部首を合成することによって行われる。学習文字を再生する場合、この文字を構成する各部首の位置やサイズなどの情報が部首生成モデルに保存されているため、そのまま部首を生成することによって、筆者の筆跡を忠実に再現することができる。しかし、未学習文字を生成する場合、単純に、学習された部首を重ね合わせて未学習文字を生成すると、筆者実際の文字との間には明らかな差が感じられる文字が生成されることがある。この原因は、同一部首でも文字によりこの部首の大きさや文字における位置が変わるためである。

Fig. 2 に示したのが中国元代の書道家趙松雪¹³⁾の文字で、Fig. 2 (a)における9個の文字の“さんずい”偏の大きさと位置は、その文字の右の部首の形状の影響によって異なる。すなわち、右の部首が縦長の場合、“さんずい”偏がやや小さく（沖、澹）、逆に、右の部首がやや短い場合、“さんずい”偏は大きく見える（泊、江、法）。ただし、これらの“さんずい”偏の形は互いに似ている。たとえば、中間の“点”は上の“点”よりやや左、下の“はね”はその点の真下、さらに、中間の“点”はほぼ真ん中にあるなどである。ほかのところ、例えば“流”、“江”における“さんずい”偏の“はね”は丸く見え、“波”、“法”、“湖”の“さんずい”偏の“点”と“はね”はつながっているなど異なる点もあるが、これは文字の構成によって起こったものではないと考えられる。すなわち、右の部首の形状により、“さんずい”偏の大きさ（縦幅、横幅）と文字における位置は変わるけれども、“さんずい”偏における各ストローク間の相対位置の変化は小さい。言い換えれば、部首“さんずい”の内部変動は少ない。

同じように、右の部首も左の部首に影響される (Fig. 2 (b)).

Fig. 2 Characters written by Zhao Song-Xue, the famous calligrapher in Yuan dynasty

したがって、本研究では、文字を構成する各部首内の変動を略し、部首の大きさ、部首間の位置関係のみに注目する。これらの関係情報をパラメータとして推定して、未学習文字の生成を行う。

3.1 特徴選択

部首の大きさ及び部首間の位置関係を推定するため、従来の筆跡分類の手法を参考に¹⁶⁾、以下16個の部首の縦横比および二つの部首間の関係パラメータから、比較的に有効なパラメータを特徴として選択する。

- (1) 両部首の重心の水平距離
- (2) 両部首の重心の垂直距離
- (3) パラメータ(1)と文字の横幅の比
- (4) パラメータ(2)と文字の高さの比
- (5) 両部首の高さの比
- (6) 両部首の横幅の比
- (7) 部首1の縦横比
- (8) 部首2の縦横比
- (9) 文字の縦横比
- (10) 文字の45°縦横比
- (11) 文字重心のX座標
- (12) 文字重心のY座標
- (13) 文字中心のX座標
- (14) 文字中心のY座標
- (15) パラメータ(13)と文字の横幅の比
- (16) パラメータ(14)と文字の高さの比

生成された文字が筆者の文字と類似するためには、選択される特徴は二つの条件を満たす必要がある。一つは、特徴値が特徴空間において筆者ごとに異なる部分空間に分布すること。もう一つは、同一筆者の同じ部首を持つ文字の特徴値は特徴空間で集中していることである。このことは、同じ筆者により書かれた文字“河”と“注”の部首“氵”に関する特徴が特徴空間内で近接していれば、“河”の“氵”は“注”の生成に有効利用できることを意味する。

それぞれのパラメータをこの二つの条件下で評価するために、パラメータ評価関数 F 比を用いる^{17), 18)}。条件1を評価する F 比 ($f_n(\text{ch})$) は多人数の筆者が書いた文字を ch 、パラメータを n として、

$$f_n(\text{ch}) = \frac{\text{筆者間平均の分散}}{\text{筆者内分散の平均}} \quad (1)$$

となる。

条件2を評価する F 比 ($f'_n(\text{sub})$) は筆者 sub が書いた同じ部首を有する文字(多字種)のパラメータ n に対して、

$$f'_n(\text{sub}) = \frac{\text{字種間平均の分散}}{\text{字種内分散の平均}} \quad (2)$$

となる。

筆者間平均の分散が大きくなるほど各筆者の特徴は特徴空間に広く分布し、筆者内分散の平均が小さくなるにつれ、パラメータ n は筆者ごとに異なる部分空間に分布する。すなわち、条件1を満たすパラメータ n は、 $f_n(\text{ch})$ が比較的大きいパラメータである。また、筆者 sub が書いた同じ部首を有する文字についてパラメータ n が空間の狭い幅に分布されるならば、 $f'_n(\text{sub})$ が小さくなり、条件2を満たす。それゆえ、本論文では、(1)式と(2)式を合わせた次式のような評価式を導入する。

$$F_n = \frac{\text{Avg}(f_n(\text{ch}))}{\text{Avg}(f'_n(\text{sub}))}$$

ここで、Avg は平均演算である。 F_n が大きいパラメータを有効なパラメータとして選択する。

Fig. 3 は、Table 1 のテストデータを用いた評価結果である。この結果から、左右結合形文字(データセット1, “○”印)のパラメータ(1), (2), (6), (7), (11)と(13)の F_n 値が相対的に大きいことから、これらのパラメータが左右結合形文字に有効であると考えられる。同様に、上下結合形文字(データセット2, “□”)と内外結合形文字(データセット3, “△”)にはそれぞれパラメータ(1), (2), (5), (6), (7), (12), (14)とパラメータ(1), (2), (5), (6), (7), (8)が有効である。以上3種類結合形の文字の F_n の和をパラメータ n の総合評

Fig. 3 The F_n values of 16 parameters about three kinds of characters being represented by ●, ■ and ▲ respectively.

Table 1 Test data set for parameters evaluation. (written by 5 subjects for 20 times)

data 1	河 注 泊 法 決 沖
data 2	宇 字 実 守 宋 宗
data 3	困 因 閑 聞 近 迂

価値とする。総合評価値が大きい六つのパラメータ (1), (2), (5), (6), (7), (8) を六つの特徴として選択する。

3.2 特徴推定モデル

未学習文字を生成する場合は、部首を変形して合成するため、学習文字の特徴 (3.1 で選択された六つのパラメータ) から、未学習文字の特徴を推定する。この推定は BP モデルによって行う。以下、この BP モデルを特徴推定モデルと呼ぶことにする。

特徴推定モデルは 3 層に構成され、入力層、出力層のユニット数はそれぞれ部首数 n 、特徴数 m ($m=6$) で、中間層のユニット数は 10 である。

学習の場合、 P 個の学習文字は、学習パターン対 (R_k, T_k) ($1 \leq k \leq P$) を与え、BP 法を用いて学習する。 $R_k = (r_{k1}, r_{k2}, \dots, r_{kn})^T$ は k 番目の学習文字を代表する入力ベクトルである。一つの文字が二つの部首 (i と j) で構成されるので、この二つの部首に対応する要素 r_{ki} と r_{kj} ($1 \leq i, j \leq n$) は 1 で、それ以外の要素は 0 である。教師ベクトル $T_k = (t_{k1}, t_{k2}, \dots, t_{km})^T$ は学習文字の特徴ベクトルで、 t_{ki} は k 番目の学習文字の i 番目の特徴である。学習は、平均 2 乗誤差が 0.01 以下になるまで行う。

未学習文字 A の特徴を推定する場合、学習されたネットワークに文字 A を代表するベクトルを入力し、出力層で得られる出力ベクトルが特徴ベクトルの推定値である。

4. 手書き文字生成シミュレーション

4.1 試作モデルの構成

提案する手書き文字生成試作モデルの処理は三つの部分で構成されている (Fig. 4)。まず、文字を構成する (二つの) 部首を部首生成モデルより生成し、つぎに、

Fig. 4 The process of handwritten character generation model

特徴推定モデルを用いて、文字の特徴を推定し、最後に、推定された特徴より、生成された部首の大きさと部首間の位置関係を決定して手書き文字を生成する。

(1) 部首生成モデル

2. で述べたように、部首生成モデルは 3 層 BP モデルである。一つの部首生成モデルは一つの特的な部首を生成する。本研究では、16 個の部首を生成するため、16 個の部首生成モデルを用いる。Fig. 5 (a) は二つの学習文字であり、Fig. 5 (b) は学習文字によって生成された部首である。

(2) 特徴推定モデル

入力層、中間層と出力層ユニットの数はそれぞれ、16 (部首数)、10 および 6 (特徴数) である。3.1 で選択された六つの特徴、パラメータ (1), (2), (5), (6), (7), (8) はそれぞれ特徴 1, 2, 3, 4, 5, 6 とする。

(3) 合成モデル

合成モデルは特徴推定モデルで推定された特徴を用いて、部首生成モデルで生成された部首の大きさや、部首間の位置関係を調整して、文字を合成する。その処理は以下の過程で行う。

- Step 1 部首生成モデルで生成された部首を長方形パターンと見なし (Fig. 5 (b)), 特徴 5, 6 の推定値により、部首の縦横比を調整する。
- Step 2 特徴 3, 4 の推定値より、二つの部首の高さの比及び横幅の比を調整する。
- Step 3 特徴 1, 2 の推定値により、二つの部首の重心の水平と垂直距離を調整する。

生成された部首 (Fig. 5 (b)) を合成モデルで調整し、その結果を Fig. 5 (c) で示す。また、特徴を推定せずに、Fig. 5 (b) の部首を直接重ね合わせると、Fig. 5 (d) のような結果が得られた。Fig. 5 (e) で同一筆者が実際に書いた文字 “何” を示した。明らかに、Fig. 5 (c) の文字

(a) Two training data

(b) Two radicals generated by radical generation model

(c) The result of combination model

(d) The result of combining radicals of (b) directly

(e) The character written by the same writer

Fig. 5 An example of generating untrained character

が Fig. 5 (d) の文字より筆者が書いた文字によく似ている。例えば Fig. 5 (d) の二つ部首が近すぎ、部首“可”の横幅が大きく、“人”偏の位置が低いなどである。ただし、これらの結論は目で“見”た結果で、量的な評価ではない。そのため、本研究では、簡単な評価を行う。特徴推定モデルで推定された未学習文字の六つの特徴を $f_1 \sim f_6$ とし、同一筆者が実際に書いた文字の特徴を $s_1 \sim s_6$ とする。筆者文字と合成された文字との誤差は

$$E = \sum_{i=1}^6 E_i = \sum_{i=1}^6 |s_i - f_i|$$

とする。 E_i は特徴 i の誤差である。

3.1 で議論したように、選択された六つの特徴は同一筆者では分散が小さく、異なる筆者では分散が大きいので、 $s_1 \sim s_6$ は筆者の“個性”を表わせるものと考えられる。そして、各特徴の誤差 E_i および文字の誤差 E は合成された文字の類似度の評価として使うことができると考えられる。文字“何”について、合成モデルで調整された文字 (Fig. 5 (c)) および調整せずに直接重ね合わせた文字 (Fig. 5 (d)) と同一筆者の文字 (Fig. 5 (e)) との誤差をそれぞれ Fig. 6 に示した。合成モデルで調整された文字の誤差のほうは、すべての特徴にわたって小さくなった。これは、目で“見”た結果と一致し、この評価手法は妥当であると考えられる。

Fig. 6 Comparison of two characters in feature errors (Fig. 5 (c), black bar and Fig. 5 (d), white bar)

4.2 シミュレーション

手書き文字生成試作モデルにおいて、学習文字として左右結合形 (“河”, “住”, “柏”), 上下結合形 (“字”, “茸”) および内外結合形 (“近”, “閑”, “因”) の計 8 個 (Fig. 7) を学習させ、学習文字と未学習文字を生成する。

Fig. 8 に学習文字の生成結果を示す。生成された文字は学習文字と類似し、筆者の個性が表わされている。

Fig. 9, Fig. 10 は、生成された 12 個の未学習文字お

Fig. 7 Training characters

Fig. 8 The reconstruction results of the training characters

Fig. 9 The reconstructed results of the untrained characters

Fig. 10 The total feature error E of untrained characters

よびこれら文字の誤差 E である。

誤差が一番大きい四つの文字“芹”，“宋”，“聞”と“迂”は，異なる結合形文字の部首の合成によって生成されたものである。たとえば，上下結合形文字“宋”の“木”は左右結合形文字“柏”の部首であり，内外結合形文字“迂”には上下結合形文字“字”の部首を利用している。反対に，文字特徴の誤差が最も小さい三つの文字“伯”，“柱”と“困”は，すべて同じ結合形文字の部首によって合成されたものである。一般的な書字常識から，このことを理解できる。たとえば“柏”の“木”偏と“宋”の“木”の書き方は全く違う。そのため，未学習文字は同じ結合形文字によって生成されると良い品質が得られると考えられる。

5. あとがき

本論では，文字の形態情報を用いて，手書き漢字生成モデルを提案した。そのモデルにより手書き文字生成の基礎的な実験を行い，生成された文字に筆者の個性が表わされていることを示した。

本モデルを用いる手法では，文字を直接に生成する代わりに，まず，部首生成を行い，生成された部首を調整して手書き文字を生成する。この二段階の文字生成により，一部の未学習文字も生成できる。

また，文字再生を可能にする特徴の抽出を試み，選択した六つの特徴によって，文字を構成する二つの部首の大きさと部首間の位置関係を調整することができる。

部首の生成については，少数のサンプル点によって部首を生成することができる BP モデルを提案した。また，文字生成の結果と特徴誤差の評価に示されているように，学習文字の特徴から未学習文字の特徴を推定する問題にも BP モデルが有効であることを明らかにした。

本研究では，小規模なテストデータと少数のパラメータに基づいて文字の特徴を選択した。より多くのテストデータおよびパラメータを選択すれば，より有効な特徴が得られると考えられる。

謝辞 本研究遂行上，有益な御助言を頂いた本学電子回路研究室長谷川浩史氏に深謝の意を表明します。

参 考 文 献

- 1) R. Plamondon and F. J. Maarse : An Evaluation of Motor Models of Handwriting ; IEEE Trans. Syst. Man. Cybernetics, Vol. 19, No. 5, pp. 1060~1072 (1989)
- 2) J. M. Hollerbach : An Oscillation Theory of Handwriting ; Biol. Cybern., Vol. 39, No. 2, pp. 139~156 (1981)
- 3) P. Morasso and F. A. Mussa Ivaldi : Trajectory Formation and Handwriting : A Computational Model ; Biol. Cybern., Vol. 45, No. 2, pp. 131~142 (1982)
- 4) M. Eden (P. A. Kolers and M. Eden (eds.)) : Handwriting

- ing Generation and Recognition ; Recognizing Patterns, MIT Press, pp. 138~154 (1968)
- 5) 田口, 増田, 田中 : 書字形成過程の相違による文字形態の多様性解析 ; 信学論 (D-II), Vol. J73-D-II, No. 5, pp. 756~766 (1990)
 - 6) 張, 真田, 手塚 : 漢字楷書体毛筆字体の計算機による生成 ; 信学論 (D), Vol. J67-D, No. 5, pp. 599~606 (1984)
 - 7) 塩野 : 非線形な幾何学的ひずみを用いた手書き風文字パターン生成の一手法 ; 信学論 (D-II), Vol. J74-D-II, No. 2, pp. 209~219 (1991)
 - 8) 張, 高橋, 松下, 志水 : ニューラルネットワークによる文字再生の一手法 ; 1991 年電子情報通信学会秋季大会講演論文集, p. 6-243 (1991)
 - 9) 張, 高橋, 松下, 志水 : ニューラルネットワークによる文字再生の一手法 (II) ; 1992 年電子情報通信学会春季大会講演論文集, p. 7-321 (1992)
 - 10) 張, 高橋, 松下, 志水 : ニューラルネットワークによる文字再生の一手法 (III) ; 1992 年電子情報通信学会秋季大会講演論文集, p. 6-276 (1992)
 - 11) 坂元, 高木 : 高品質明朝体ひらがな・カタカナフォントの計算機による生成 ; 信学論 (D), Vol. J68-D, No. 4, pp. 702~709 (1985)
 - 12) K. Funahashi : On the Approximate Realization of Continuous Mappings by Neural Networks ; Neural Networks, Vol. 2, No. 3, pp. 183~192 (1989)
 - 13) 趙 : 福神観記, 上海書画出版社 (1990)
 - 14) 大友, 原, 大内, 原 : 神経回路モデルを用いた手書き文字認識手法 ; 信学論 (D-II), Vol. J74-D-II, No. 2, pp. 158~165 (1991)
 - 15) 小川, 手塚 : 漢字の階層表現とその認識 ; 信学論 (D), Vol. J57-D, No. 12, pp. 700~707 (1974)
 - 16) 吉田, 高沢, 宮崎, 中島 : 文字の外形を利用した筆跡の分類に関する基礎的研究 ; 科学警察研究所報告法科学編, Vol. 40, No. 1, pp. 24~31 (1987)
 - 17) B. S. Atal : Automatic Recognition of Speakers from Their Voices ; Proc. of IEEE, Vol. 64, No. 4, pp. 460~475 (1976)
 - 18) J. J. Wolf : Efficient Acoustic Parameters for Speaker Recognition ; J. Acoust. Soc. Amer., Vol. 51, No. 2, pp. 2044~2055 (1972)