

# 類似文字認識のための統計的文字構造モデル構成方法

山口貴也<sup>†1</sup> 吉本尊人<sup>†1</sup> 岩満直紀<sup>†1</sup> 行天啓二<sup>†1</sup>

本論文では、統計的構造モデルを用いた類似文字認識手法について提案する。統計的構造モデルを用いた従来手法は、類似する文字に誤認識してしまう傾向がある。文字認識の性能を向上させるためには、文字認識を適用して幾つかの認識結果の候補を選んだ後に、より正確な類似文字認識を適用する必要がある。そこで我々は、類似文字認識用の統計的構造モデルと、その生成方法を提案する。提案手法は、類似文字間の異なる構造を表現する統計的構造モデルを、類似する2つの文字を構成するストロークの関係を調整することにより、学習用文字画像データから自動的に獲得することができることを特徴とする。

## Statistical character structure model for similar character recognition

TAKAYA YAMAGUCHI<sup>†1</sup> TAKAHITO YOSHIMOTO<sup>†1</sup>  
NAOKI IWAMITSU<sup>†1</sup> KEIJI GYOHTEN<sup>†1</sup>

This paper proposes a similar character recognition method based on the statistical character structure model. The original method sometimes confuses similar characters having the same structure in parts. We propose the statistical character structure model for the similar character recognition. In our model, the stroke pairs correspond to the partial difference between the similar characters and are obtained by calculating a maximum solution of an objective function that evaluates their dissimilar parts. From experimental results, we could observe that the obtained model had the appropriate pairs of closely related neighboring strokes that correspond to the dissimilar parts in the similar characters.

### 1. はじめに

文字認識は、研究分野としての歴史が長く、実用化された場合の社会への貢献度が高い点で、画像認識の分野における主要な技術であるといえる。文字認識の中でも、オフライン手書き漢字認識は最も難しいとされている。その理由として、漢字は数多くのストロークを複雑に配置することにより構成されている点、偏旁の合成により生成されているため、部分的に似た構造を持つ文字が多い点があげられる。また、手書きの場合、書き手による文字形状の差異も考慮しなければならないため、問題はさらに複雑になる。

一般に手書き漢字認識では統計的手法が主流となっており、HMM, Neural Network, SVM, MRFなどの統計的パターン分類器を利用して数多くの研究がなされ、良好な結果が得られている[1][2][3]。一方、構造的な手法による手書き漢字認識も長い間研究されている。構造的な文字認識手法についてはストロークベースのアプローチが多く、認識対象となる漢字画像から、漢字を構成する幾つかのストロークを獲得し、文字モデルとどの程度適合するかを算出する。この際、ストロークの位置や関係をどのように表現するかが問題となる。文字をストロークに基づいて認識することは、人間の文字認識手順に似ていることや、草書体文字認識への可能性を持つことから、様々な研究が行われている[2][4]。

本研究は、ストロークベースの文字認識手法である統計的構造モデルを用いた文字認識手法[5]に着目し、その

改良手法を提案する。統計的構造モデルは、文字を構成するストロークモデルの集合であり、ストロークモデルの存在位置をストロークの特徴点の分布として表現する。このモデルの特長として、ストロークモデル間の複雑な位置関係を、重要な隣接関係のみを選択して簡単に表現することができる点がある。統計的構造モデルを用いた文字認識手法は良好な結果を示しているが、類似文字に関して誤認識してしまう傾向を示すことが指摘されている。

そこで本論文では、類似文字認識のための統計的構造モデルを提案する。提案する類似文字認識用統計的構造モデルは、類似文字を構成する全てのストロークモデルから生成されており、ストロークモデル間の重要な隣接関係を適切に選択することにより、類似文字間の部分的に異なる構造を的確に表現することができることを特長とする。具体的には、学習用類似文字画像データを用いて、統計的構造モデルと、類似する2つの文字との各々の適合度が分離するように、ストロークモデル間の隣接関係を設定する。その結果、類似文字間において部分的に異なる構造を成すストロークモデルが、他のストロークと数多く重要な隣接関係を有することになり、類似文字間の差異に着目した、正確な類似文字認識が可能になる。実験では、提案手法を用いて、類似文字画像データから類似文字認識用の統計的構造モデルを生成した。その結果、類似文字間の異なる構造を表現するストロークモデル間の関係を、正しく獲得することができることを確認した。

<sup>†1</sup> 大分大学  
Oita University

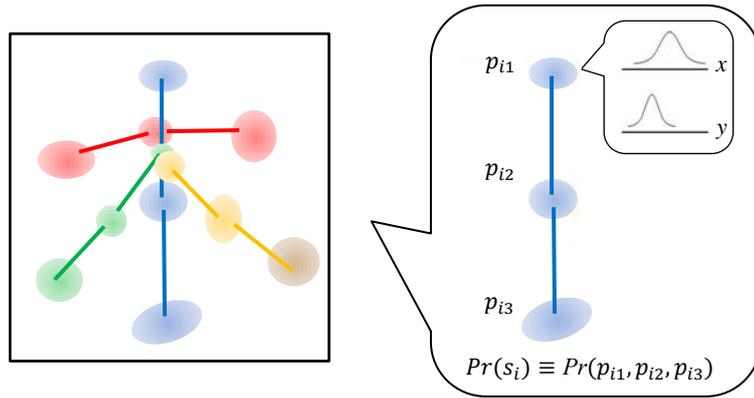


図 1 ストロークの表現  
 Figure 1 Representation of strokes.

以下、本論文の構成として、2.では、従来手法である統計的  
 文字構造モデルを用いた文字認識について説明する。3.では、  
 統計的  
 文字構造モデルを類似文字認識に拡張する  
 方法を提案する。4.では実験結果について述べ、5.でまとめ  
 と今後の課題について述べる。

## 2. 統計的 文字構造モデルを用いた文字認識

### 2.1 統計的 文字構造モデル

Kim らが提案する統計的  
 文字構造モデルは、図 1 に示す  
 ように、文字の構造をストロークの集合として表現する[5]。  
 ストロークは、始点・中点・終点を結ぶ線分であり、これ  
 ら 3 つの特徴点の座標から 6 次元ベクトルで表現される。  
 様々な書き手の様々な文字変形を表現するために、各文字  
 を、各ストロークの結合分布により統計的に表現する。例  
 えば、ある文字が  $n$  本のストロークから構成されるとした  
 場合、その文字モデルを、以下のような結合分布により表  
 現するものとする。

$$Pr(S) = Pr(s_1, s_2, \dots, s_n) \quad (1)$$

$S = \{s_i\}$  は、ある文字を表すストローク集合であり、各スト  
 ローク  $s_i$  は、平均と共分散行列によって表現される正規分  
 布に従うものとする。

式(1)は全てのストロークの関係を正確に表現している  
 一方、統計的に表現する上で重要でないストロークの関係  
 も含んでいる。そこで Kim らは、ストローク間の重要な関  
 係のみを用いて文字モデル  $Pr(S)$  を表現した。式(1)につい  
 て、

$$\begin{aligned} Pr(S) &= Pr(s_1, s_2, \dots, s_n) \\ &= Pr(s_1)Pr(s_2, \dots, s_n | s_1) \\ &= Pr(s_1)Pr(s_2 | s_1)Pr(s_3, \dots, s_n | s_1, s_2) \\ &\dots \\ &= \prod_{i=1}^n Pr(s_i | s_1, s_2, \dots, s_{i-1}) \end{aligned} \quad (2)$$

より、 $Pr(S)$  はあるストロークに先立つストロークを条件  
 とする条件付き確率の乗算となる。各ストロークの条件付  
 き確率  $Pr(s_i | s_1, s_2, \dots, s_{i-1})$  について、式(3)に示すように、  
 最も重要な唯一の他ストロークのみが条件となるように近  
 似する。

$$Pr(S) \approx \prod_{i=1}^n Pr(s_i | s_{nl(i)}) \quad (3)$$

ここで、 $s_{nl(i)}$  はストローク  $s_i$  と最も重要な関係を持つ隣  
 接ストロークである。

各ストロークの隣接ストロークについては、式(2)を式(3)  
 のように近似しても、元々の文字モデル  $Pr(S)$  の結合分布  
 をできる限り保持できるような隣接ストロークを選択する  
 ことが望ましい。Kim らは、式(2)に対する式(3)の近似の度  
 合いを評価するため、Kullback-Leibler measure を用いた。  
 その結果、各ストロークを表すノードと、ストローク間の  
 相互情報量を重みとしてもつエッジから成るグラフの、重  
 みつき極大全域木を獲得することによって、各ストローク  
 の隣接ストロークを獲得することができることを明らかに  
 した。以後、このようにして得られた  $Pr(s_i | s_{nl(i)})$  をス  
 トロークモデルと呼ぶ。

### 2.2 文字認識方法

2.1 で述べた文字モデルは、図 2 に示すような典型的な  
 構造的  
 文字認識手順によって、与えられた文字画像と照合  
 される。まず、細線化に基づく手法により、文字画像から、  
 各ストロークを抽出する。以後、文字モデル中に表現され  
 ているストロークモデルと区別して、文字画像から得られ  
 たストロークを入力ストロークと呼ぶ。

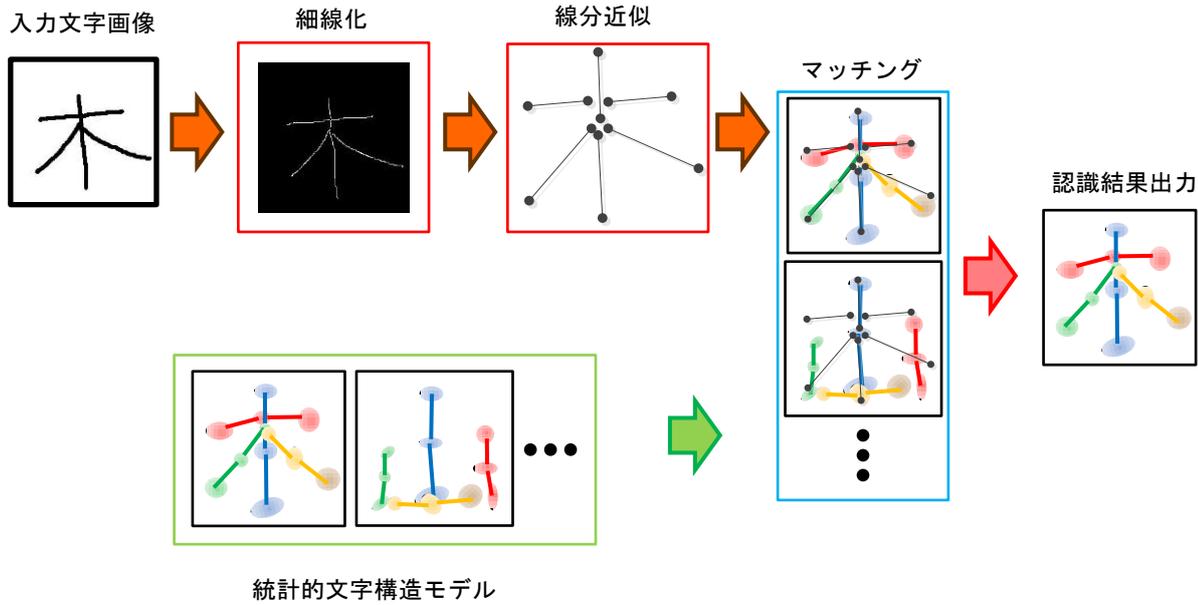


図 2 文字認識手順

Figure 2 Character recognition process.

次に、入力ストロークと各文字モデル中のストロークモデルとを対応付ける。図 3 に示すように、各ストロークモデルに対応するストロークは 1 本の入力ストロークではなく、複数の入力ストロークを数珠つなぎにしたものである場合もある。そのため、複雑な組み合わせ最適化問題として考えることができる。本手法では、入力ストロークとストロークモデルとの適合度が最も適切になるような組み合わせを、グラフ構造により表現される探索空間上で解く。探索空間におけるノードは、あるストロークモデルに対応する入力ストロークの組み合わせの一候補を表す。エッジは、互いに矛盾しない組み合わせ候補を表すノード同士を接続する。各ノードは、ストロークモデルと入力ストローク集合の適合度を値として持つ。ここでの最適解の探索は、探索空間上の最小コストパスを求めることと等価であり、ストロークモデルと入力ストロークの組み合わせは、求められた最小コストパス上のノードとして求められる。また、入力ストロークと文字モデルとの適合度は求められた最小コストと等価である。

以下のような処理を、全種類の文字モデルと入力ストロークに対して適用し、最も適合度が高い文字モデルを認識結果とする。

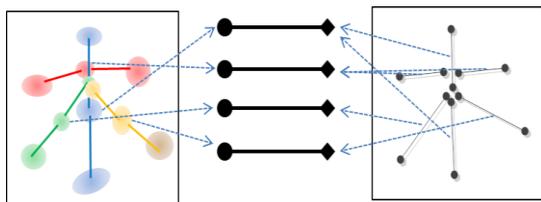


図 3 モデルストロークと入力ストロークの対応関係  
 Figure 3 Correspondence between model strokes and input strokes.

### 3. 統計的字体構造モデルによる類似文字認識

#### 3.1 類似文字の統計的字体構造モデル

2 つの類似文字を認識する場合、文字構造が異なる部分に着目して判断することが望ましい。2.1 において説明した統計的字体構造モデルでは、文字構造をストロークモデルの集合として表現する。構造が異なる部分に着目するということは、2 つの類似文字を構成するストロークモデル集合を比較して、異なる部分に着目するということである。そこで、2.1 で説明した文字モデルを拡張した、類似文字認識のための統計的字体構造モデル、およびその獲得方法を提案する。提案手法では、類似文字を構成するすべてのストロークモデルの結合分布を利用し、文字構造が異なる箇所に該当するストロークモデルを獲得する。このストロークモデルは、学習データから自動的に獲得される。得られたストロークモデルに着目することにより、2.1 で述べた従来手法に比べて、より正確な適合度の算出ができることが期待される。

類似文字判別のための統計的字体構造モデルを

$Pr(S^L | L_c)$  ( $c = 1, 2$ ) と表現する。  $L = \{L_1, L_2\}$  は類似文字  $L_1, L_2$  のペアであり、  $S^L = \{s_i\}$  は類似文字  $L_1, L_2$  を構成するすべてのストロークモデルである。式(1)と同様に、文字モデル  $Pr(S^L | L_c)$  は各ストロークモデル  $s_i$  の正規分布の結合分布によって表現できるものとする。文字モデル  $Pr(S^L | L_c)$  は、認識結果が類似文字ペア  $L = \{L_1, L_2\}$  のいずれかであると仮定した場合の、2.2 で述べた方法により文字画像から獲得された入力ストローク集合の文字  $L_c$  らしさを表す。以後、簡単のため  $Pr(S^L | L_c)$  を  $P_c(S^L)$  と表わす。類似文字  $L$  の認識では、入力ストローク集合と、文字  $L_1$  のモデル  $P_1(S^L)$  および文字  $L_2$  のモデ

ル  $P_2(S^L)$  との適合度をそれぞれ計算する。その後、適合度が大きい方のモデルを認識結果として出力する。

### 3.2 隣接ストローク選択

文字モデル  $P_c(S^L)$  は、式(2)と同様に、以下のように表すことができる。

$$P_c(S^L) = P_c(s_1, s_2, \dots, s_{|S^L|}) \\ = \prod_{i=1}^{|S^L|} P_c(s_i | s_1, s_2, \dots, s_{i-1}) \quad (4)$$

ここで  $|S^L|$  は類似文字  $L_1$  を構成するストロークモデル数と類似文字  $L_2$  を構成するストロークモデル数の和である。式(3)と同様に、各ストロークモデルの条件付き確率において、最も重要な唯一の他ストロークモデルのみが条件となるように近似する

$$P_c(S^L) \approx \prod_{i=1}^{|S^L|} P_c(s_i | s_{nl(i)}^L) \quad (5)$$

ここで  $s_{nl(i)}^L$  は、類似文字  $L$  を判別する際に、ストロークモデル  $s_i$  と最も重要な関係を持つ隣接ストロークモデルである。ここでの隣接ストロークモデル  $s_{nl(i)}^L$  については、式(3)中の  $s_{nl(i)}$  とは異なり、文字  $L_1$  と  $L_2$  における文字構造の違いを特徴づける重要なストロークモデルが選択されることが望ましい。

ここでは、学習データを用いて、類似文字  $L$  の認識に有効な各ストロークモデル  $s_i$  の隣接ストロークモデル  $s_{nl(i)}^L$  を選択する方法を提案する。学習データは  $\mathfrak{S} = \{(C_k, T_k)\}$  として与えられるものとする。 $C_k$  は学習用類似文字画像  $I_k$  から得た入力ストローク集合、 $T_k$  は学習用文字画像  $I_k$  の文字種類をそれぞれ表わすものとする。 $T_k$  の表す文字種類とは類似文字  $L$  中の文字  $L_1$ 、 $L_2$  のいずれかであり、 $T_k \in L$  である。学習データにおける入力ストローク集合  $C_k$  と、文字モデル  $P_c(S^L)$  を構成する複数のストロークモデルを、2.2 で述べた方法で対応付け、算出された適合度を  $P_c(S^L = C_k)$  と標記するものとする。

類似文字  $L$  の認識に使用する評価式  $F$  を、式(5)を用いて以下のように定義する

$$F(S^L) = \log P_1(S^L) - \log P_2(S^L) \quad (6)$$

このとき、学習データ  $\mathfrak{S} = \{(C_k, T_k)\}$  について以下の式が成立することが望ましい。

$$F(S^L = C_k) = \begin{cases} \infty & \text{if } T_k = L_1 \\ -\infty & \text{if } T_k = L_2 \end{cases} \quad (7)$$

つまり、以下の式をできる限り最大化するような文字モデル  $P_c(S^L)$  を構成すればよい。

$$V(S^L, \mathfrak{S}) = \text{avr}_{\{(C_k, T_k) | T_k = L_1\} \in \mathfrak{S}} \{F(S^L = C_k)\} \\ - \text{avr}_{\{(C_k, T_k) | T_k = L_2\} \in \mathfrak{S}} \{F(S^L = C_k)\} \quad (8)$$

ただし、 $\text{avr}\{\}$  は平均を表す。(8)を最大化するような文字モデル  $P_c(S^L)$  を構成するという事は、各ストロークモデル  $s_i$  について、他ストロークモデルを、隣接ストロークモデル  $s_{nl(i)}^L$  として1つずつ選択するという事である。

式(6)の定義に用いられている文字モデル  $P_c(S^L)$  の対数は、パラメータ  $\{w_{cij} | c=1,2 \ i,j=1,2,\dots,|S^L|\}$  を用いて、以下のように表現することができる。

$$\log P_c(S^L) \approx \sum_{i=1}^{|S^L|} \log P_c(s_i | s_{nl(i)}^L) \\ = \sum_{s_i \in S^L} \sum_{\substack{s_j \in S^L \\ s_j \neq s_i}} (w_{cij} \log P_c(s_i | s_j)) \quad (9)$$

$$\text{ただし } w_{cij} = 0 \text{ or } 1, \sum_{\substack{s_j \in S^L \\ s_j \neq s_i}} w_{cij} = 1$$

式(9)を式(8)に代入すると、

$$V(S^L, \mathfrak{S}) \\ = \sum_{s_i \in S^L} \sum_{\substack{s_j \in S^L \\ s_j \neq s_i}} (w_{1ij} (\text{avr}_{\{(C_k, T_k) | T_k = L_1\} \in \mathfrak{S}} \{\log p_1(s_i | s_j)\}) - \text{avr}_{\{(C_k, T_k) | T_k = L_2\} \in \mathfrak{S}} \{\log p_1(s_i | s_j)\}) \\ + w_{2ij} (\text{avr}_{\{(C_k, T_k) | T_k = L_2\} \in \mathfrak{S}} \{\log p_2(s_i | s_j)\}) - \text{avr}_{\{(C_k, T_k) | T_k = L_1\} \in \mathfrak{S}} \{\log p_2(s_i | s_j)\}) \quad (10)$$

となる。

式(9)における制約の下でこの値を最大化するには、モデル  $P_1(S^L)$  中の各ストロークモデル  $s_i$  について、

$$\text{avr}_{\{(C_k, T_k) | T_k = L_1\} \in \mathfrak{S}} \{\log p_1(s_i | s_j)\} - \text{avr}_{\{(C_k, T_k) | T_k = L_2\} \in \mathfrak{S}} \{\log p_1(s_i | s_j)\}$$

が最大となるストロークモデル  $s_i, s_j$  に対応する  $w_{1ij}$  のみを1、それ以外の  $w_{1ij}$  を0とすればよい。また、 $P_2(S^L)$  中の各ストロークモデル  $s_i$  について、

$$\text{avr}_{\{(C_k, T_k) | T_k = L_2\} \in \mathfrak{S}} \{\log p_2(s_i | s_j)\} - \text{avr}_{\{(C_k, T_k) | T_k = L_1\} \in \mathfrak{S}} \{\log p_2(s_i | s_j)\}$$

が最小となるストロークモデル  $s_i, s_j$  に対応する  $w_{2ij}$  のみを1、それ以外の  $w_{2ij}$  を0とすればよい。

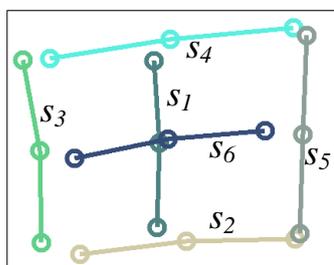
以上のようにして得られた、式(9)中のパラメータ  $\{w_{cij}\}$  は、式(5)で定義される文字モデル  $P_c(S^L)$  の隣接ストロークを表現している。 $\{w_{cij}\}$  が表す隣接ストロークは、類似文字  $L_1, L_2$  の文字構造の違いを特徴づける重要なストロークモデルとなる。

#### 4. 実験

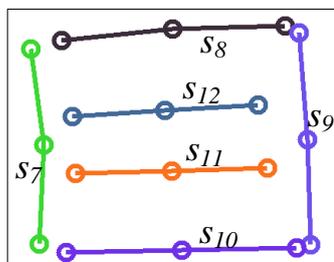
提案手法の評価実験を行った．学習用類似文字画像データとして，産総研手書き文字画像データベース ETL9B を利用した．細線化による入力ストローク獲得後，2. で述べた従来手法を用いて，入力ストロークとストロークモデルとの対応を獲得した．統計的文字構造モデルについては，ストロークフォント KST32B から得たストロークデータを初期データとして用い，学習用文字画像データから自動生成した．

まず類似文字  $L = \{L_1, L_2\}$  として「田」を  $L_1$ ，「目」を  $L_2$  とし，これらを認識する類似文字用統計的文字構造モデルを作成した．文字  $L_1$  と文字  $L_2$  それぞれ 30 枚の学習用文字画像から入力ストロークを獲得し，従来手法により入力ストロークとストロークモデルとの対応を獲得した．図 4 に，用いた統計的文字構造モデルを示す．図 4 中の  $s_i$  は，各ストロークモデルの始点・中点・終点の平均を表し，これらの集合が  $S^L$  に相当する．

次に，入力ストロークとストロークモデルの対応結果から，類似文字認識用統計的文字構造モデル  $P_1(S^L)$  と  $P_2(S^L)$  におけるストロークモデルの平均と共分散行列を求めた．図 5 に各ストロークモデルの平均を示す．例えば，図 2 中のストロークモデル  $s_1$  は，文字  $L_1$  中に存在し，文字  $L_2$  中に存在しないストロークを表現している．獲得さ



「田」のストロークモデル



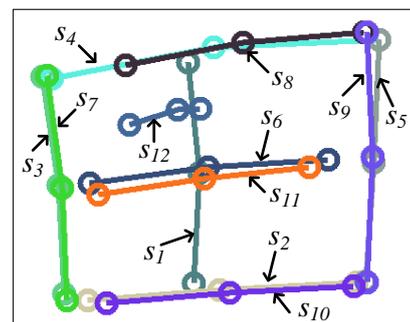
「目」のストロークモデル

図 4 統計的文字構造モデル

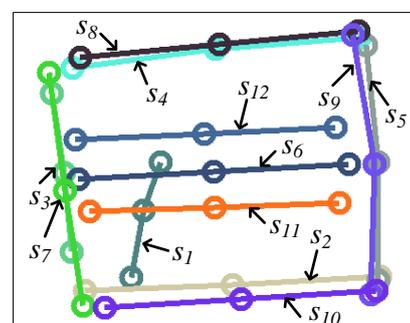
Figure 4 Stastical character structure model.

れた文字モデル内でも，ストロークモデル  $s_1$  は， $P_1(S^L)$  内では正しい位置に存在し， $P_2(S^L)$  では，意味のない位置に存在している．よってストロークモデル  $s_1$  は，類似文字  $L_1$  と  $L_2$  を判別する手がかりとなる極めて重要なストロークであると言える．

次に，得られた文字モデル  $P_1(S^L)$ ， $P_2(S^L)$  と前述した入力ストロークとストロークモデルとの対応関係から，3.2 に示した方法により，パラメータ  $w_{1ij}$ ， $w_{2ij}$  を求めた．その結果を図 6 に示す．このパラメータは，3.2 で述べたように，文字モデル  $P_1(S^L)$ ， $P_2(S^L)$  の隣接ストロークを表現している．この隣接ストロークの選択については，文字  $L_1$  と  $L_2$  の文字構造の違いを特徴づける，重要なストロークモデルが選択されることが望ましい．図 6 に示すように， $w_{1ij}$  の値は，文字  $L_1$  中に存在し文字  $L_2$  中に存在しないストロークモデル  $s_i$  が，隣接ストロークとして集中的に選択されていることを示している．また， $w_{2ij}$  の値は，文字  $L_2$  中に存在し文字  $L_1$  中に存在しないストロークモデル  $s_j$  が，隣接ストロークとして集中的に選択されていることを示している．以上より，提案手法により，学習用文字画像データから，類似文字間における部分的に異なる構造を，各ストロークモデルの隣接ストロークとして正しく獲得することができたと言える．



$P_1(S^L)$  のストロークモデル



$P_2(S^L)$  のストロークモデル

図 5 類似文字認識用統計的文字構造モデル

Figure 5 Stastical character structure model for similar character recognition.

	j=1	2	3	4	5	6	7	8	9	10	11	12
i=1	0	0	0	1	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0	0	0	0	0
9	1	0	0	0	0	0	0	0	0	0	0	0
10	1	0	0	0	0	0	0	0	0	0	0	0
11	1	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1

パラメータ $w_{ij}$

	j=1	2	3	4	5	6	7	8	9	10	11	12
i=1	0	0	0	0	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	1
4	0	0	0	0	0	0	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	0	1	0	0
6	0	0	0	0	0	0	0	0	0	0	1	0
7	0	0	0	0	0	0	0	0	0	0	0	1
8	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0
12	0	0	0	0	0	1	0	0	0	0	0	0

パラメータ $w_{2ij}$

図6 パラメータ $\{w_{cij}\}$ の獲得

Figure 6 Obtained result of parameter  $\{w_{cij}\}$ .

最後に、類似文字認識用統計的文字構造モデルが、類似文字認識に有効に働くかについて検証する。図7(a),(b)は、類似文字の組である「田」と「目」、「水」と「永」について、式(6)に示した評価式 $F(S^L)$ についての分布を示したものである。用いたデータは学習用文字画像データとは異なる、テスト用文字画像データ30枚である。図7(a),(b)中の○印は、テスト用文字画像データとして文字 $L_1$ を与えた時の評価値の分布であり、×印は、テスト用文字画像データとして文字 $L_2$ を与えた時の評価値の分布である。比較のため、従来手法による評価値も求めた。従来手法における文字 $L_1, L_2$ の統計的文字構造モデルを $P'_1(S^{L_1}), P'_2(S^{L_2})$ とし、

$$F'(S^L) = \log P'_1(S^{L_1}) - \log P'_2(S^{L_2}) \quad (11)$$

と定義する。ここで $S^{L_c}$ は文字 $L_c$ を構成するストロークモデル集合である。図7(c),(d)は、類似文字の組である「田」と「目」、「水」と「永」について、式(11)に示した評価値 $F'(S^L)$ の分布を示したものである。特に図7の(b)と(d)を比較すれば、提案手法である類似文字認識用統計的文字構造モデルを用いた方が、類似文字の評価値を明確に分離させることができることがわかる。

## 5. まとめ

本研究では、統計的文字構造モデルによる文字認識の改良手法として、類似文字認識用統計的文字構造モデルを提案した。文字を構成するストロークを、ストローク上の特

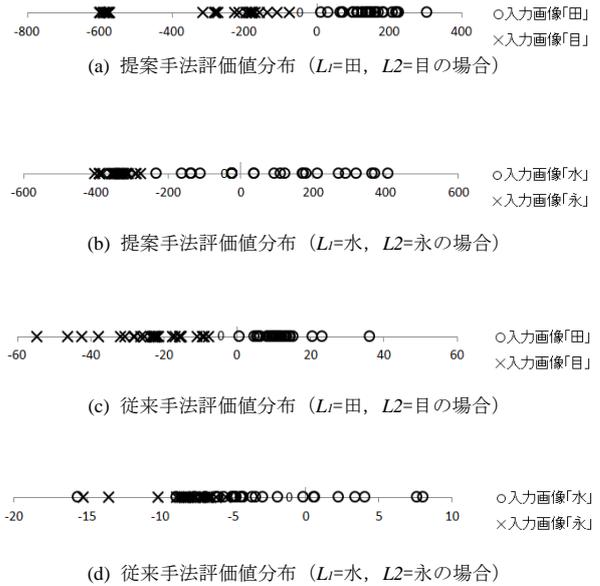


図7 従来手法との比較

Figure 7 Comparisons of results obtained from the proposed method and the original method.

徴点の分布であるストロークモデルの集合として表現する点は、従来の統計的文字構造モデルと同様である。類似文字認識用統計的文字構造モデルの特徴として、類似文字間において部分的に異なる構造を重要な隣接関係として表現する点がある。これにより、類似文字認識において、類似文字間の異なる構造に着目した評価が可能となった。さらに、学習用類似文字画像データを用いて重要な隣接関係を獲得する方法を提案した。実験により、類似文字間において部分的に異なる構造を、重要な隣接関係として自動的に獲得できる点と、提案手法により生成された類似文字認識用統計的文字構造モデルを用いれば、類似文字認識における評価値の信頼性が向上する点を示した。

今後の課題として、ストロークモデルと入力ストロークとの対応付けの高速化が挙げられる。現状では、各ストロークモデルに対応する入力ストローク集合が多数になると文字全体におけるストロークモデルと入力ストロークとの対応付けにおいて組み合わせの爆発が起こる。そのため、ストロークモデルと入力ストロークとの対応付け獲得に多大な時間を要するという問題が発生した。今後、ストロークモデルと入力ストロークとの対応付けの際に、ありえない対応付けを考慮しないアプローチを考え、高速化を目指していく。ストロークモデルと入力ストロークとの対応付けの高速化が実現されれば、従来の統計的文字構造モデルを用いた文字認識の後処理として、今回提案した類似文字認識手法を追加し、類似文字認識についても信頼性の高い、より高精度な手書き文字認識手法を目指す。

## 参考文献

- 1) Chou, C.-H. et al.: A prototype classification method and its use in a hybrid solution for multiclass pattern recognition, *Pattern Recognition*, Vol.39, Issue 4, pp.624-634 (2006).
- 2) Dai, R. et al.: Chinese character recognition: history, status, and prospects, *Frontiers of Computer Science in China*, Vol.1, Issue 2, pp.126-136 (2007).
- 3) Srihari, S.N. et al.: Offline Chinese handwriting recognition: an assessment of current technology, *Frontiers of Computer Science in China*, Vol.1, Issue 2, pp.137-155 (2007).
- 4) Zeng, J. et al.: Cascade Markov random fields for stroke extraction of Chinese characters, *Information Sciences*, Vol.180 Issue 2, pp.301-311 (2010).
- 5) Kim, I.-J. and Kim, J.-H.: Statistical character structure modeling and its application to handwritten Chinese character recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25, No.11 (2003).