

ストローク間の位置関係の文法記述に基づくオンライン手書き漢字認識

大田 郁実[†] 山本 遼[†] 酒向 慎司[†] 嵯峨山茂樹[†]

[†] 東京大学 大学院情報理工学系研究科

〒 113-8656 東京都文京区本郷 7-3-1

E-mail: †{ota,yamaryo,sako,sagayama}@hil.t.u-tokyo.ac.jp

あらまし 本論文では、オンライン手書き漢字認識において構造を文法記述する手法を検討する。本手法では、文字部品やストローク（画）で構成される漢字の階層構造に着目し、すべての漢字が確率文脈自由文法 (SCFG) で生成できるとした。ストローク認識結果と文字部品間の相対的な位置関係の両方を構文木を生成する過程で用いることによって、ストロークと位置関係の誤認識を相互に補償することが期待できる。さらに、位置関係を数種類の簡単なモデルで表現し、異なる漢字と共有することによって、少ない学習データでも十分な認識精度を出すことが可能になり、また、未学習の漢字に対しても、生成規則を書き込むことによって対応できるようになる。本手法を、JAIST（北陸先端科学技術大学院大学）のオンライン手書き漢字データベースに含まれる、筆順通りに比較的丁寧に書かれた新旧教育漢字 1,016 字のデータセットを用いて評価したところ、未学習文字に対しても既学習文字と比べて遜色ない認識率が得られた。

キーワード オンライン手書き文字認識, サブストローク HMM, 確率文脈自由文法, 文字部品間の位置関係

On-Line Recognition of Handwritten Kanji Characters Based on Grammatical Description of Relative Positions between Strokes

Ikumi OTA[†], Ryo YAMAMOTO[†], Shinji SAKO[†], and Shigeki SAGAYAMA[†]

[†] Graduate School of Information Science and Technology, The University of Tokyo

7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

E-mail: †{ota,yamaryo,sako,sagayama}@hil.t.u-tokyo.ac.jp

Abstract In this paper, we propose a new method for online handwritten Kanji character recognition. In this approach, we turn attention to the hierarchical structure of Kanji which consist of character-parts and strokes, and consider all character patterns to be generated by stochastic context-free grammar (SCFG). As we use both result of stroke recognition and position relationship between character-parts in process of parsing, we can expect that error in stroke and position can be mutually compensated. Furthermore, by expressing position in simple models and share with different Kanji, it enables to make effective results with a few training data, and it becomes possible to recognize untrained Kanji, just by writing down the rules for them. Through our experiment using proposed method, we indicated that the recognition error rate of non-trained Kanji is not inferior to that of trained Kanji, using 1,016 handwritten Kanji written carefully in correct order, which contained in JAIST (Japan Advanced Institute of Science and Technology) online handwritten Kanji database.

Key words Online Handwriting Recognition, Substroke HMM, Spatial Relationship of Subcharacters, Stochastic Context-Free Grammar

1. はじめに

オンライン手書き漢字認識技術は、PDA や携帯電話に搭載されているペン入力インターフェース等に広く用いられており、今後もより幅広い応用が期待される。一方、崩れ字に対する認

識精度は人間のそれには遠く及ばないのが現状であり、これを向上させるには漢字の構造に着目する必要がある。なぜなら、われわれはストローク形状と各ストロークや文字部品の位置関係を総合的に判断して漢字を識別しているのであり、崩れ字においてもそれらはある程度保存されると考えられるからである。

オンライン手書き漢字認識の研究分野においては、中井ら [1] が隠れマルコフモデル (HMM) を漢字認識に有効に応用した。ストロークをサブストロークという方向線分で表現するこの手法は優れた成果を上げたが、位置情報の利用が限定的なものに留まっていた。漢字を基本ストロークが一定の構造の下に組み合わせられたものとして考え、各ストローク間の位置関係を階層構造を持つグラフによって表現した研究として文献 [2] がある。ハングル文字に関する研究 [3] ではハングル文字と漢字の共通点に言及しつつ、ストロークの組み合わせが文字部品を、文字部品の組み合わせが文字を成すという階層構造を用いて、依存関係をグラフで繋いで表現している。これらの応用として徳野ら [4] は、「ベイジアンネットワーク」を用いる手法を提案したが、多量の学習データを準備する必要があるという問題点があった。

本研究では位置関係を漢字個々について定義するのではなく、ある一定数のラベルで表現することを考える。ただし、文字部品間の位置関係は筆者によってゆらぎが存在し、どのラベルに属するのか座標値のみから一意に定めるべきではない。このため各ストロークの組み合わせがどのような文字部品を成すかの判別は確率的に扱う必要があるが、これには数式認識における研究 [5] で用いられている SCFG が有効である。

われわれは、漢字の階層構造を有効に利用することに主眼を置き、SCFG に基づく新しい手書き漢字認識手法を提案する。以下では、まず関連する先行研究について述べ、SCFG の詳細と本手法での応用法、外接矩形を用いたストローク・文字部品間の位置関係の認識の手法、ストローク認識手法を詳述した後、CYK アルゴリズムを利用した構文解析手法を説明する。その後、評価実験と結論、今後の展望を述べることにする。

2. ストローク間の相対的な位置関係を利用した手書き漢字認識

2.1 サブストローク HMM による漢字認識

本項では、先行研究としてサブストローク HMM を用いた漢字認識手法の概要を述べる。詳細は [1] を参照のこと。

中井らは、漢字のストロークのカテゴリカルな区別には、その構成要素として図 1 のような 8 方向の長短の方向線分 (サブストローク) で十分である点に着目し、それらの連結でストローク形状を表現し、ストロークの組み合わせで漢字を表現した。ペンの軌道座標を等時間間隔でサンプリングして、時間的に連続する 2 点の座標の差を速度ベクトルと定義し、これを時間順に並べたものが HMM の特徴量となる。この手法では、漢字に含まれる複雑なストロークをコンパクトなサブストロークというモデルで表現したことによって、異なる漢字でモデルを共有することができ、少ない学習量で高性能な漢字認識を行うことが可能になった。また、筆者適応なども可能になるといった有利な点を持つ一方、ストローク間の位置関係に関する情報は pen-up モデルによってしか得ることができず、例えば「八」と「人」のように同一のサブストローク表現でストロークの位置関係のみが異なる漢字の判別は困難だった。

この問題を解決するため、徳野ら [4] はサブストローク HMM

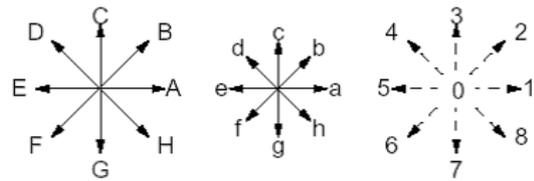


図 1 サブストローク HMM

中井ら [1] から引用、8 方向の長短 pen-down と pen-up サブストローク

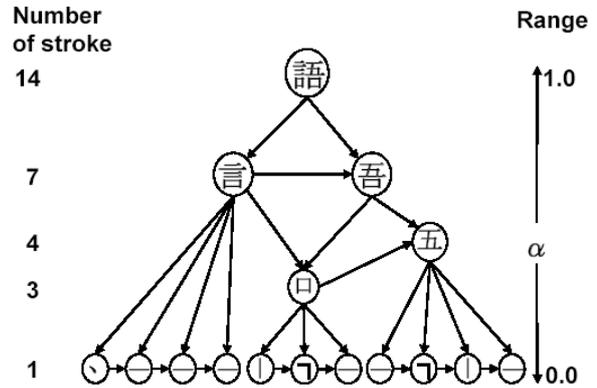


図 2 ベイジアンネットワークの例

徳野ら [4] から引用、「語」のベイジアンネットワーク表現

認識の後処理として「ベイジアンネットワーク」を利用して位置関係を評価する手法を提案した。これは、図 2 のように、位置に依存関係を持つすべてのストローク・文字部品を有向グラフで結び、それに基づいて位置関係を学習する。この手法は、文献 [1] の手法を改善したが、認識したい漢字すべてについて多量の学習を行わなければならない、また、未学習の漢字の認識は難しい。

2.2 漢字の階層構造

われわれは、異なる漢字と共有でき、学習を容易にするため、位置関係を限られた数のモデルで表現することを提案する。また、SCFG を漢字認識過程に導入し、漢字の階層構造を有効に利用することを考える。漢字は 1 本または複数のストロークの組み合わせである文字部品で構成される。例として、「位」という漢字は「イ (にんべん)」と「立」という文字部品で成り立っていて、これらはさらに小さい文字部品に分解でき、最終的にストローク単位に展開できる。漢字は多くの階層構造を内包していて、これを効率的に利用することで認識精度の向上が期待できる。

2.3 文脈自由文法による漢字構造記述

文脈自由文法は再帰的言語の生成と解析に適し、次の 4 項組 (V_N, V_T, P, S) で表現される。 V_T はそれ以上展開できない終端記号の集合、 V_N は非終端記号の集合、 S は始端記号、 P は以下のうちのいずれかで表される生成規則の集合である。

$$A \rightarrow BC \quad (A \in (V_T \cup V_N)^*, B, C \in (V_T \cup V_N))$$

$$A \rightarrow a \quad (A \in (V_T \cup V_N)^+, a \in V_T)$$

本研究では、各手書きストロークを終端記号、ストローク認識

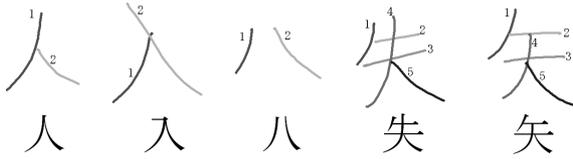


図3 ストローク表現が同一の漢字の例

3つの漢字「人」「入」「八」はストローク表現が同じであるため、位置情報を用いずに識別はできない。「失」と「矢」も同様で、第4ストロークが第2ストロークを上下に貫くか否かの違いしかない。

結果と各文字部品を非終端記号、認識漢字を始端記号とし、あらかじめ定義された生成規則に従って、漢字を解析する。

ストロークや文字部品間の相対的な位置関係の情報は高精度の認識を行う場合においてきわめて重要である。図3では位置関係の考慮なしでは識別が困難な漢字を示している。例として、「人」と「八」はストローク形状に大きな違いはなく、ストローク間の相対的な位置関係のみが異なる。また、両方の漢字において、第2ストロークが第1ストロークに対してどの位置に書かれるかは筆者によって位置関係にゆらぎが生じることから、確率的な枠組みにおいて考えるのが不可欠である。相対座標をとって「右」にあるか「右下」にあるかを解析する前から1つに限定して宣言するのは非効率である。そのため、われわれは上記の文脈自由文法に確率を付加したSCFGを導入する。

われわれは、図4のように $p = \langle A \rightarrow BC, s \rangle$ の形の非終端記号生成規則を並べた文法辞書を作成した。ここで、 A は漢字または文字部品、 B, C は漢字、文字部品またはストローク認識結果、 s は B と C の間の相対的な位置関係を表す。生成規則の適用確率は B と C が s の位置関係にある尤度（以下、位置尤度）で与えられる。終端記号生成規則は $q = \langle t \rightarrow \alpha \rangle$ の形で与えられ、 t はあらかじめ定義されている認識ストロークモデル、 α は手書きストロークである。適用確率は手書きストローク α がストロークモデル t である尤度（以下、ストローク尤度）と等しい。2つの尤度の具体的な計算方法については後述する。

次に解析において、 $p_n = \langle A_n \rightarrow B_n C_n, s_n \rangle$ を n 番目に適用された非終端記号生成規則、 $q_m = \langle t_m \rightarrow \alpha_m \rangle$ を m 番目に適用された終端記号生成規則、 N と M をそれぞれ適用された非終端記号、終端記号生成規則の数とする。このとき、漢字認識問題は以下のように定式化ができる。

$$X_0 = \arg \max_{X \in E_X} \prod_{n=1}^N P(p_n) \prod_{m=1}^M P(q_m) \quad (1)$$

上記の数式で、 X は仮説、 E_X は文法から生成可能な漢字の仮説集合、 X_0 は最尤仮説を表す。上式より、SCFGを用いた漢字認識問題はストローク尤度と位置尤度の積を最大にする導出を求める問題と等価であるといえる。

2.4 ストローク認識

文献[1][4][6]では、漢字全体を認識する過程でサブストロークHMMを用いているが、本手法ではストロークの認識のみを用いる。図5のように、pen-down サブストロークの結合で各ストロークモデルを記述することで、手書き漢字ストロークを

A	\rightarrow	B	$\langle s \rangle$	C
位	\rightarrow	イ	$\begin{matrix} \boxed{B} & \boxed{C} \\ \text{right} \\ B \rightarrow C \end{matrix}$	立
イ	\rightarrow	ノ	$\begin{matrix} \boxed{B} & \boxed{C} \\ \text{right-down} \\ B \rightarrow C \end{matrix}$	丨
立	\rightarrow	一	$\begin{matrix} \boxed{B} \\ \text{below} \\ \boxed{C} \\ B \rightarrow C \end{matrix}$	冫
人	\rightarrow	ノ	$\begin{matrix} \boxed{B} & \boxed{C} \\ \text{right-down} \\ B \rightarrow C \end{matrix}$	ノ
八	\rightarrow	ノ	$\begin{matrix} \boxed{B} & \boxed{C} \\ \text{right} \\ B \rightarrow C \end{matrix}$	ノ
入	\rightarrow	ノ	$\begin{matrix} \boxed{B} & \boxed{C} \\ \text{left-down} \\ C \rightarrow B \end{matrix}$	ノ

図4 確率文脈自由文法に基づく辞書の例

ノ	\rightarrow	G f	ノ	\rightarrow	F B
一	\rightarrow	A	フ	\rightarrow	F A G d
冫	\rightarrow	A G	ノ	\rightarrow	h

図5 ストロークモデルのサブストローク表現

各認識ストロークモデルをサブストロークの結合で表したものの。評価実験では、19種類のストロークモデルを採用した。

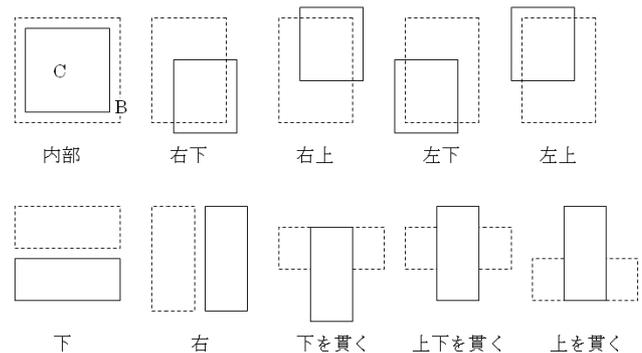


図6 採用した位置関係

サブストロークHMMの枠組みで認識する。HMMは確率モデルであり、認識結果はストロークモデル候補とそれぞれのストローク尤度である。このようにして得られるストローク認識結果をSCFGに組み込むことで、位置関係の認識結果と合わせて統一的に評価できることが期待される。

2.5 文字部品間の相対的位置関係

提案手法では、位置関係を「右」や「下」などの限られた数の単純なモデルとして表すことにする。これによって、異なる漢字間でモデルを共有でき、少ない学習でも優れた効果を期待できる。具体的には、本研究では位置関係モデルを図6の10種類に限定し、すべての位置関係をこれらのモデルで表現できるとした。

文字部品 X の外接矩形を、左端 X 座標 b_X^l 、上端 Y 座標 b_X^t 、右端 X 座標 b_X^r 、下端 Y 座標 b_X^b を用いて次のように表現する。

$$b_X \equiv (b_X^l, b_X^t, b_X^r, b_X^b)$$

次に手書き漢字 b_W は正方形の外接矩形を持っていると仮定で

きるため、以下のような正規化を行う。

$$b_W = (b_X^l, b_X^t, b_X^r, b_X^b) \Rightarrow b'_W \equiv (0, 0, 1, 1)$$

$$b_A = (b_A^l, b_A^t, b_A^r, b_A^b) \Rightarrow b'_A \equiv (b_A^l, b_A^t, b_A^r, b_A^b)$$

$$= \begin{pmatrix} b_A^l - b_X^l & b_A^t - b_X^t & b_A^r - b_X^r & b_A^b - b_X^b \\ b_X^r - b_X^l & b_X^b - b_X^t & b_X^r - b_X^l & b_X^b - b_X^t \end{pmatrix}$$

ここで、 b'_A は文字部品 b_A の線形変換の結果である。

位置関係を学習する際の特徴量は、以下のように2つの文字部品 B と C の外接矩形の差分 Δh_{BC} とする。

$$\Delta h_{BC} = (\Delta b_{BC}^l, \Delta b_{BC}^t, \Delta b_{BC}^r, \Delta b_{BC}^b)$$

$$= (b'_C - b'_B, b'_C - b'_B, b'_C - b'_B, b'_C - b'_B)$$

すべての位置関係 s について、特徴量が2次元正規分布にしたがうと仮定すれば、 B と C が位置関係 s にある位置尤度 $P_{pos}(B, C, s)$ は以下の式で求めることができる。

$$P_{pos}(B, C, s)$$

$$= \frac{1}{\sqrt{(2\pi)^4 |\Sigma_s|}} \exp\left(-\frac{1}{2}(\Delta h_{BC} - \mu_s)^T \Sigma_s^{-1} (\Delta h_{BC} - \mu_s)\right)$$

上式で、 μ_s は分布の平均、 Σ_s を共分散行列とする。

2.6 複数の筆順への対応

漢字には正規の書き順が存在するが、すべての漢字筆者がそれに従うわけではない。そのため、汎用性のある漢字認識の枠組みを構築するに当たって、複数の筆順に対応していくことは不可欠である。筆順は文字部品単位で変動する傾向がある。例えば、「認」、「刃」、「忍」の3つの漢字について考えると、各漢字に含まれる文字部品「刃」の書き順は各筆者ごとに同一であることが多い。このことから、書き順は文字部品ごとに継承するとして漢字の階層性を生かすことができる。

本手法で筆順違いに対応するには、辞書に生成規則を追加することで対処できる。この際、上述の理由から一つの漢字について、生成規則全てを書き直して追加する必要はなく、図7のように筆順違いが生じる文字部品についてのみ生成規則を追加すれば十分である。追加後は、複数の導出が許されるようになる。

[4]をはじめ多くの先行研究は筆順違いに対応しているが、筆順が異なる手書き漢字も学習に用いなければならない場合が多い。しかし提案手法では、漢字間で共有できる位置モデルを用いているため、各位置モデルについて十分学習がなされていれば、文法辞書に生成規則を追加するだけで対応可能である。このため、多様な筆順に対処することが期待できる。

2.7 構文解析

SCFGに基づく言語は、CYK(Cook-Younger-Kasami) アルゴリズムを用いて解析できる。

まず、手書き漢字をストロークごとに分解し、それぞれについてストローク認識を行う。結果は、ストロークモデル候補とストローク尤度の列で与えられるので、それをCYK行列の要素 $(1, j)$ $j = 1, 2, \dots, L$ に書き込む。ここで、 L はCYK行列の行数に相当し、また手書き漢字の画数に一致する。実際には、計算速度の問題からストローク候補のうち尤度上位 K 位までを書き込む。次に、要素 $(1, j)$ と $(1, j+1)$ ($j = 1, \dots, L-1$)

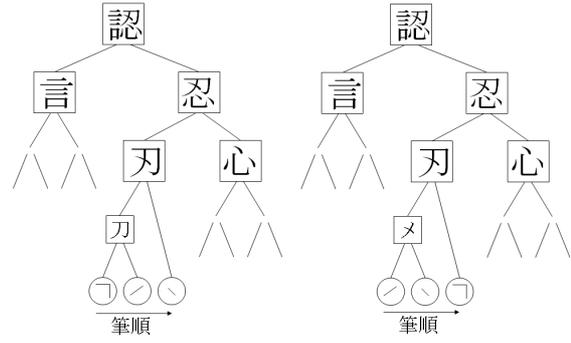


図7 筆順が異なる場合の構文木

筆順の変化は文字部品単位で発生すると仮定すると、上の漢字「認」について「刃」の部分の筆順が変化するとき、「刃」の生成規則を追加すれば十分であり、「刃」を持つ他の漢字にも適用される。

に適用可能な全ての非終端記号生成規則を探索する。生成規則 $p = \langle A \rightarrow BC, s \rangle$ が適合すれば、各ストローク尤度と位置尤度の積を、ストローク B と C が文字部品 A である尤度として A と尤度を要素 $(2, j)$ に書き込む。引き続き、要素 (i, j) $i = 2, \dots, L$, $j = 1, \dots, L-i$ に対しては、 (i, j) と $(i, j+1)$ に適用可能な非終端記号生成規則を探索し、2つの要素の尤度と位置尤度の積を A と共に $(i+1, j)$ に書き込む。このように続けていくと、最終的に認識結果が要素 $(L, 1)$ に書き込まれる。ここには結果が複数書き込まれることもあり、その場合は始端記号集合 S に属するもののうち、最も尤度の高いものを認識結果とする。

計算過程で尤度を掛け合わせていくため、尤度は非常に小さくなっていく。そのため、実際は以下のような対数尤度を用いて計算する。

$$\log P(A) = \log P(B) + \log P(C) + W \log P_{pos}(B, C, s)$$

上式で、 $P(A), P(B), P(C)$ は漢字、文字部品またはストロークがそれぞれ A, B, C である尤度、 W が重みを表す。後述する評価実験において、ストローク尤度よりも位置尤度を強調するため、われわれは経験的に $W = 5$ を採用している。

上述のCYKアルゴリズムは、実装が簡易である反面、自然言語解析等の膨大な情報を扱うには低速であるとされている。今後、さらに高速な構文解析アルゴリズムの実装を目指していく。

3. 評価実験

3.1 実験に用いたデータ

われわれは提案手法の評価をJAIST-IIPL(北陸先端技術大学院大学情報科学研究科知能情報処理学講座)で提供されているオンライン手書き文字データベース(図9)を用いて行った。このうち、 α 文字セットは正規の筆順と画数の下、比較的丁寧に書かれた漢字で構成され、 γ_1 文字セットは比較的丁寧に正規の画数で書かれているが筆順が正しいとは限らない漢字を含む。われわれは、学習と性能評価を α セット、筆順違いの漢字に対する評価を γ_1 セットに含まれる新旧教育漢字1,016字について行った。

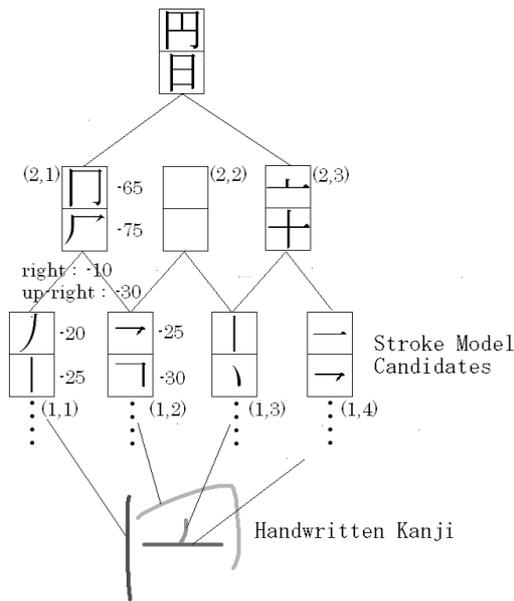


図 8 確率文脈自由文法の構文解析例

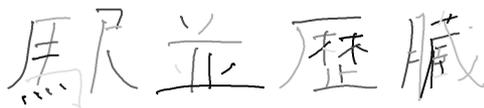


図 9 オンライン手書き漢字データベースの例

JAIST-IHPL が収集したオンライン手書き漢字データベースに含まれるデータの例。実際は、等時間間隔でサンプリングされたペン軌道の座標点と筆圧情報の配列である。本研究では筆圧情報は用いていない。

3.2 文法辞書

学習と筆順が正しい手書き漢字に対する評価実験に用いた辞書は、2,056 の生成規則を含み、正規の筆順通りに書かれた 1,016 漢字のみ生成可能である。つまり、各漢字、文字部品について 1 通りの導出しか許さない。これを以下、正規筆順辞書と呼ぶ。筆順違いへの対応を評価するために用いた辞書は、 γ_1 データセット数人分の筆順に対応するため、正規筆順辞書に 126 の生成規則を追加し、各漢字、文字部品について複数の導出を許し、1,016 漢字のみ生成できる。これを以下、複数筆順対応辞書と呼ぶ。

3.3 評価

評価実験は、複数筆者によって書かれた複数の手書き漢字データを前述の辞書を適用して認識させ、尤度 1 位の候補を認識結果とし、認識に用いられた手書き漢字のうち、誤認識された漢字の割合（誤認識率）を計算することで行った。

3.3.1 筆者への非依存性

われわれは、 α セットの 10 人分 1,016 漢字を用いて位置特徴量の学習を行った後、学習に用いたものと同じ筆者 10 人分 (closed) と異なる筆者 50 人分 (open) のデータについて 1,016 漢字を認識させて誤認識率を算出し、結果は表 1 のようになった。誤認識率の範囲を見ると、最悪値は open の場合の方が高い。一般に位置関係の崩れの度合いは筆者によって異なり、その度合いが学習データの分布から離れていれば認識精度は当然

カテゴリ	平均誤認識率	誤認識率の範囲
closed	1.92%	1.18% - 2.37%
open	2.45%	1.09% - 4.95%

表 1 筆者と誤認識率の関係

漢字	提案手法 (A)	中井ら (B)	誤認識例
人	6%	72%	八、入 (A,B)
相	0%	22%	查 (B)
八	0%	22%	人、入 (B)
夫	18%	16%	尺、天、井、木、王 (A,B)
文	88%	26%	六 (A,B)、夫 (B)
友	62%	10%	反 (A,B)

表 2 先行研究に比べて改善・悪化した漢字例

相対的な位置関係の評価を行った効果が顕著に現れているものを例示した。数字は誤認識率。

のことながら低下する。しかしながら、われわれは漢字の外接矩形が正方形になるよう正規化しているため、学習データ分布から極端に離れることは無く、認識率にも大きく影響しないようになっている。実際、誤認識率の平均値は closed と open で差は小さい。以上のことから、本手法は筆者に大きく依存しない手法であり、文字認識システムとして効果的に機能していると言える。

次に、サブストローク HMM のみを用いた先行研究 [1] と比較する。文献 [1] の手法は画数変動に対応しているため、提案手法と異なり、認識結果に画数が異なるものが出力される。そのため、以下では条件を近づけるために文献 [1] の手法に画数情報を付加することで、画数が正解と異なる漢字は結果から除外した。この条件下で HMM パラメータの学習に用いていない α セット 49 筆者について評価を行った結果、表 2 のように提案手法によって認識が改善した漢字、逆に悪化した漢字が存在した。これを見ると、「人」「八」のようにストローク形状がほとんど変わらず、かつ相互の位置関係が識別に重要な役割を果たす漢字は認識率が大きく改善した。「相」が正しく認識されていることから、位置関係の評価が大きな影響を与えていることがわかる。一方、「友」が「反」に誤認識される理由は、「友」の第 2 ストロークが第 1 ストロークを貫いているかどうかの判断を誤っているためである。ストロークがある文字部品の上下を貫いている位置関係は多様であり、1 つのモデルだけで対応するのは難しい。今後の課題として、位置関係モデルの再設定があげられることになろう。

3.3.2 学習漢字への非依存性

新旧教育漢字 1,016 漢字を JIS コード順に並べ、順番の偶奇で 2 つのグループに分けた。位置特徴量の学習は α セット 20 筆者について 508 漢字 (1,016 漢字の半分) を用いて行い、評価は 20 人分を、学習に用いた漢字 (closed)、学習していない漢字 (open) について両方行い、結果は表 3 のようになった。結果を見ると、位置関係が全く学習されていない漢字に対しても高精度の認識が期待できることがわかる。このことから本手法では、全ての漢字について位置情報を学習しなくても、生成

カテゴリ	平均誤認識率	誤認識率の範囲
closed	2.08%	1.18% - 3.08%
open	2.35%	1.58% - 3.16%

表 3 学習の有無と誤認識率の関係

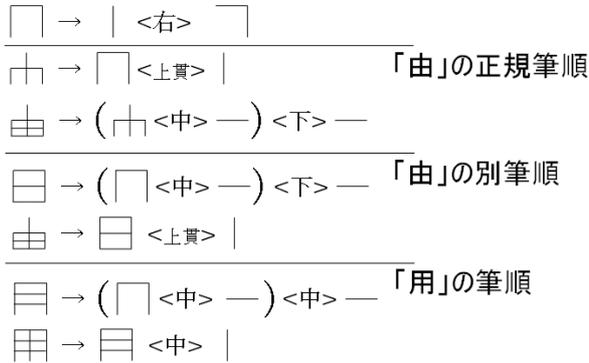


図 10 複数筆順辞書適用による性能の低下

正規の筆順では「用」と「由」に文法表現上大きな違いがあるが、「由」に別筆順を許すと、「用」とは文字部品間の位置関係の違いでしか区別できない。そのため、位置関係の誤認識が漢字の認識誤りに直結する。

規則を文法辞書に加えることで認識可能であることがわかる。この理由としては、位置モデルを限定し単純にしたことによって、異なる漢字間でモデルが共有され、少ない学習量でも効果的な学習が期待できるからであると考えられる。

3.3.3 筆順違いの漢字への対応

筆順違いの漢字を認識するため作成した複数筆順対応辞書は、 $\gamma 1$ データセット 2 筆者 (筆者 ID:0012,0052) 1,016 漢字において、認識に必要な 126 の生成規則を追加したもので、各漢字について複数の導出を許す。本辞書を適用したことによる、2 筆者の認識率の変化は表 4 上部のようになり、ともに、表 1 の水準に達している。このことから、本手法では生成規則を適宜追加することで、筆順が異なる手書き漢字に対応することが可能である。

次に、複数筆順対応辞書を用いて、位置学習に利用していない α セット 50 筆者について再度認識を行った。結果は、表 4 下部のようになり、認識率が正規筆順辞書を用いた場合に比べてやや悪化した。つまり、筆順違いのために生成規則を追加したことによって、正規の筆順で書かれた漢字のうち認識しにくくなった漢字が存在したことになる。これの顕著な例として、「用」は正規筆順辞書の下では誤認識率 4% だが、複数筆順対応の辞書の下では、88% に大きく性能が低下した。原因としては、図 10 のように、正規筆順しか認めない辞書では「用」と「由」はストローク表現が異なるため「用」と「由」が混同されることは無かったが、複数筆順を認める辞書では「用」と「由」でストローク表現が同一になる場合があるため、漢字の認識精度は位置認識の精度のみに依存するようになる。複数筆順対応により、上述の例以外にもストローク表現が同一になることによって認識が難しくなる漢字が出てくるため、認識率がやや低下したと考えられる。しかし、複数筆順への対応は不可欠であり、今後位置特徴量の改良が求められる。

データセット/筆者 ID	正規筆順辞書	複数筆順辞書
$\gamma 1/0012$	29.98%	3.75%
$\gamma 1/0052$	12.73%	3.26%
$\alpha/50$ 筆者	2.35%	3.66%

表 4 筆順違い対応による誤認識率の変化

各辞書を適用したときの誤認識率の変化。追加した規則数は 126 で、 $\gamma 1$ セットに対して認識率が上がる一方、導出が増えることから α セットに対してはやや低下した。

4. 結 論

本論文では、SCFG に基づき、ストローク認識と位置関係認識双方の結果を用いる新しい漢字認識手法を提案した。評価実験の結果としてこの枠組みが少ない学習データしかない条件下でも高い認識精度を示すことがわかり、かつ、未学習の漢字についても効果的であることを示すことができた。今後は、位置特徴量をより信頼性がおけるものに改良していくこと、また、JIS 第一水準、第二水準漢字まで認識できるよう文法辞書を拡張していくこと、画数が正規のものとは異なる手書き漢字に対する認識手法の提案を目指す。また、生成規則が増大してもリアルタイムでの認識を維持する必要があるため、CYK アルゴリズム以外の構文解析手法についても考察し、計算過程での枝刈りも含めてシステムとしての有効性もあわせて検討していく。

5. 謝 辞

本研究は、北陸先端科学技術大学院大学において開示された手書き漢字認識技術と膨大なデータベースを活用した。旧・知能情報学講座の方々に深謝する。特に、現・東京農工大学の徳野淳子助手には原稿を執筆するにあたって多大な助言を頂いた。また、現・富山県立大学の中井満講師からは助言のみならず、性能評価のための再実験をして頂き、その結果を頂いた。ここに記して両氏に深謝する。

文 献

- [1] M. Nakai, N. Akira, H. Shimodaira and S. Sagayama: "Substroke to HMM-based On-line Kanji Handwriting Recognition", Proc. of ICDAR'01, pp. 491-495 (2001).
- [2] J. Chen and S. Lee: "On-Line Chinese Character Recognition Via a Representation of Spatial Relationships Between Strokes", Int'l J. Pattern Recognition and Artificial Intelligence, 11, 3, pp. 329-357 (1997).
- [3] K. Kang and J. Kim: "Handwritten Hangul Character Recognition with Hierarchical Stochastic Character Representation", Proc. of ICDAR'03, pp. 212-216 (2003).
- [4] 徳野, 中井, 下平, 嵯峨山: "階層的な文字部品間の位置関係を選択的に利用したオンライン手書き文字認識", 電子情報通信学会技術研究報告, PRMU2005-79, pp. 95-100 (2005).
- [5] R. Yamamoto, S. Sako, T. Nishimoto and S. Sagayama: "On-Line Recognition of Handwritten Mathematical Expression Based on Stroke-Based Stochastic Context-Free Grammar", The Tenth International Workshop on Frontiers in Handwriting Recognition, pp. 491-495 (2006).
- [6] 徳野, 楊, G. P. da Silva, 末代, 中川: "構造化字体表現型 HMM に基づくオンライン手書き文字認識における座標特徴の利用法と効果", 電子情報通信学会技術研究報告, PRMU 2005-222, pp. 117-122 (2006).