

文字構造の文法記述に基づくオンライン手書き漢字列認識

大田 郁実[†] 山本 遼[†] 西本 卓也[†] 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,nishi,sagayama}@hil.t.u-tokyo.ac.jp

あらまし 本論文では、漢字の構造を文脈自由文法に基づいて記述し、オンライン孤立漢字認識、またそれを拡張した漢字列認識に対して用いる手法を検討する。本手法では、文字部品やストローク（画）が成す漢字の階層構造を利用し、全ての手書き漢字が、文字部品間の相対的な位置関係を考慮した確率文脈自由文法 (SCFG) で生成できるとした。HMM による手書きストローク尤度と位置関係尤度の両方を構文木を生成する過程で用いることによって、ストローク形状と位置関係ゆらぎを統合した尤度の精度を向上することが期待できる。漢字列に対しても漢字間の位置関係を確率的に扱うことで、文字境界確定と孤立漢字認識を統合的に処理することができ、境界確定誤りによる認識誤りを補償することができる。JAIST のオンライン手書き文字データベースの漢字 2～3 字で構成される熟語について評価したところ、言語モデルを用いずに文字列認識率 90.80% を得た。

キーワード オンライン手書き文字認識, 文字列認識, 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

Online Handwritten Kanji String Recognition Based on Grammar Description of Character Structures

Ikumi OTA[†], Ryo YAMAMOTO[†], Takuya NISHIMOTO[†], 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,nishi,sagayama}@hil.t.u-tokyo.ac.jp

Abstract In this paper, we discuss a method for online handwritten Kanji character recognition by describing the structure of Kanji using Context-Free Grammar(CFG), and extend it in order to recognize Kanji character strings. 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 from stochastic context-free grammar (SCFG). As we use both likelihoods of strokes and position relationship between character-parts in the process of parsing, it is expected that likelihoods of strokes and positions jointly improved the likelihood accuracy. Using this method on Kanji strings enables the system to deal with both of the main problems together, namely, the decision of boundaries between letters and isolated Kanji character recognition process. Therefore the recognition error due to the failure of letter boundary decision can be compensated. We evaluated the method using handwritten phrases consist of 2 or 3 Kanji characters contained in JAIST Online Handwriting Database, and gained the recognition rate of 90.80% without using language models.

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

1. はじめに

筆記された文字をテキストに変換する手書き文字認識技術は、煩雑なキーボード操作を介さずに情報をコンピュータに入

力できる手段として、現在幅広く用いられている。特に、ペンの時系列筆跡を利用して認識するオンライン手書き文字認識技術は、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利用され、今後もより広範な応用が期待され

る。近年では、筆記した文字列をリアルタイムで認識するデバイスが携帯機器向けを中心に開発されてきており、今後研究が加速すると見られている。

オンライン文字列認識手法はさまざまなものが提案されてきているが、それらは「枠あり」と「枠なし」型に大別できる。「枠あり」型文字列入力とは、筆記領域に複数のマス目や枠があらかじめ設定され、その枠内に一文字ずつ書くようにして認識を行う。これは、筆記者が文字境界を明示的に与える方法に相当し、孤立文字認識の手法がそのまま用いることができる利点がある。一方、「枠なし」型文字列入力は文字境界が筆記者によって与えられないが、枠がある場合と比較して筆記者への負担が少ないため、現在に至るまで盛んに研究されている。しかし、文字境界を認識システム側が求める必要があるため、より難しい問題とされている。

これまで、われわれはサブストローク HMM [1] の特徴を生かしつつ、認識過程で位置関係を積極的に用いる新しい手法を提案してきた [2], [3]。位置関係を少数のカテゴリで表現し、漢字の構造をストロークと位置関係の組み合わせで記述することによって、各カテゴリは異なる漢字間で共有され、位置の学習に必要な学習量を少なくすることができた。さらに漢字独特の階層構造を用いれば、漢字の構造は比較的簡便に記述できることから、未学習の筆順・字種への対応も容易に行うことができる。以上のことから、われわれは再帰的言語の生成に適しているといわれる文脈自由文法を確率の枠組みに拡張した、確率文脈自由文法を用いて漢字を記述することを提案した。本研究では、これまでの提案手法をさらに拡張し、漢字列の構成を文字間の位置関係を考慮した文法で記述した、漢字列認識システムを定式化し実現することを考える。

以下では、まず、確率文脈自由文法 (SCFG) に基づく漢字列認識に関する事項を詳述し、具体的な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. 確率文脈自由文法による漢字列認識

2.1 従来の漢字列認識問題

従来提案されてきた枠無し文字列認識手法は、文字領域切り出し (セグメンテーション) の後に個別文字を認識するという2段階を踏むものが主流である。セグメンテーションの方法としては空間的または時間的特徴を用いるものが多い。例えば、複数のストローク特徴を用いる方法 [4] では、学習によって文字の切り出しパラメータを最適化する拡張性の高い手法だが、漢字とかなで特徴が大きく異なる日本語での適用が難しい。また、文字の大きさを推定し確率的に用いる手法 [5] は、隣接文字のストロークが重なって書かれた場合に対して弱い。

2段階方式の場合、各文字を認識するためには、まず文字境界を高精度で確定させなくてはならない。しかし、座標値などから文字境界を事前に認識させ一意に定めた後、孤立文字認識を行うと文字の構成が困難になったり、有意でない文字列が得られてしまうなどの問題が出るのが予想される。例えば、図1のように文字部品間の距離で文字境界をあらかじめ決めてしまう場合、文字を構成しないストローク群が発生したり (図1

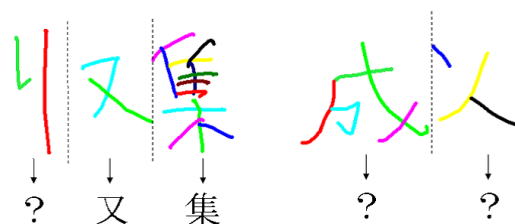


図1 単語境界付与の失敗による誤認識例

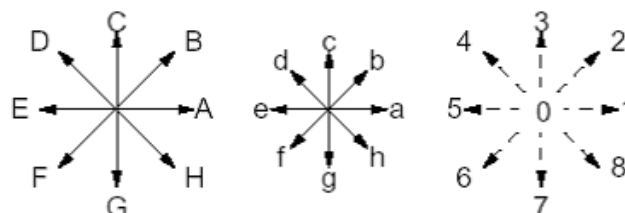


図2 サブストローク HMM [1] で用いられたペンアップ・ダウンベクトル

長短のペンダウンベクトル (左・中) とペンアップベクトル (右)。[1] より引用

左)、誤った境界が文字列全体に影響を及ぼす (図1右) こともあると考えられる。このことから、文字境界の決定と孤立文字認識を同時進行で行い、筆者の筆記自体に誤りがなければ、ストローク群が必ず有意な文字列を構成するような認識を行うシステムの構築が適当である。また、単語境界を確率的に扱うことで、文字境界決定と孤立文字列認識それぞれの誤りを相互に補償することが可能になる。

2.2 確率文脈自由文法による漢字認識

従来より漢字認識はパターンマッチングによる方法が主流であったが、必要な学習量とメモリが膨大になる上、筆者固有のストローク形状変形などへの対応が難しかった。そのため、ストローク変形を確率現象として捉え、音声認識分野で用いられてきた隠れマルコフモデル (HMM) の漢字への応用が近年研究されてきた。その中でペン筆跡を方向成分という細分化された単位でモデル化し、図2のようなペンダウン・ペンアップの長短25種類のベクトル (サブストローク) で漢字を表現したサブストローク HMM [1] が提案された。上記25種のサブストロークは、音声認識における音素に対応し、全ての漢字を区別する上で必要かつ十分なモデル数である。この手法は、モデル数の少なさ、サブストローク連結による漢字表現の容易さから未学習字種・筆順に対応できる点も利点で、成果を上げた。しかしながら、この手法はストローク変形には頑健だが、ストローク間位置関係をペンアップベクトルとしてでしか扱えないため、「人」と「八」のようにストローク形状がほぼ同一で位置関係のみが異なる漢字などは識別が困難であった。

2.3 文法の定義

漢字列認識の文法を以下の4項組による文法 G として定義する。

- (1) 終端記号集合:

$$V_T = \{ \text{各手書きストロークの時系列特徴量} \}$$

- (2) 非終端記号集合: $V_N = \{K_{str}, \mathcal{K}, \mathcal{P}, \mathcal{M}\}$
(3) 開始記号: S
(4) 生成規則集合: P (後述)

K_{str} は漢字列を表す記号、 \mathcal{K} は認識対象の漢字集合、 \mathcal{P} は文字部品・部首集合、 \mathcal{M} はストロークモデル集合である。生成規則は、下の4種を定義する。

a) 漢字生成規則

$$\langle S \rightarrow K, 1.0 \rangle \quad (K \in \mathcal{K})$$

K は認識システムが受容する漢字を表す。適用確率は常に1である。

b) 漢字列生成規則

$$(1) \quad \langle K_{str} \rightarrow K_1 K_2, p = P_{border}(b_{K_1}, b_{K_2}) \rangle$$

$$(K_1, K_2 \in \mathcal{K})$$

$$(2) \quad \langle K'_{str} \rightarrow K_{str} K, p = P_{border}(\tilde{b}_{K_{str}}, b_K) \rangle$$

$$(K_{str}, K'_{str} \in V_N, K \in \mathcal{K})$$

$$(3) \quad \langle S \rightarrow K_{str}, 1.0 \rangle \quad (K_{str} \in V_N)$$

(1) は漢字列が漢字2字を生成する様子を、(2) は漢字列が漢字列+漢字1字を生成する様子を表す。(3) は認識システムが漢字列を受容することを表しており、適用確率は常に1である。

(1) の適用確率は漢字 K_1 と K_2 の間に文字境界が存在する確率 P_{border} に相当し、各々の外接矩形に基づいて計算される。

(2) の適用確率も漢字列 K_{str} と漢字 K の間に文字境界が存在する確率 P_{border} に相当するが、漢字列 K_{str} の外接矩形ではなく、 K_{str} の末端の漢字の外接矩形 $\tilde{b}_{K_{str}}$ と漢字 K の外接矩形 b_K に基づいて計算される。つまり、新たに加えられる漢字とその左隣の漢字の2字についての位置関係を見ることになる。

以下、 P_{border} を境界尤度と表す。

認識システムの実現のためには、利用するストロークモデル集合 \mathcal{M} 、漢字の構成を表す非終端記号生成規則、認識システムが受容する漢字に対する漢字生成規則を記述することが不可欠である。その後、確率文脈自由文法に基づきボトムアップな構文解析を行い、手書きの筆跡が表す最尤の漢字を推定する。以下で、各尤度の計算方法、構文解析アルゴリズムについて詳しく述べる。

c) 非終端記号生成規則

$$\langle A \rightarrow BC, \ell, p = P_{pos}(b_B, b_C, \ell) \rangle$$

$$(A \in (\mathcal{P} \cup \mathcal{K}), B, C \in (\mathcal{K} \cup \mathcal{P} \cup \mathcal{M}), \ell \in \mathcal{L})$$

ℓ が B に対する C の位置関係を表すモデル、 \mathcal{L} が位置関係モデル集合を表す。漢字または文字部品が細分化された文字部品・ストロークモデルを生成する規則で、適用確率 p は B と C それぞれを囲む最小の長方形 (外接矩形) b_B, b_C が ℓ の位置関係にある尤度に等しく、以下これを位置尤度 $P_{pos}(b_B, b_C, \ell)$ と呼ぶ。

d) 終端記号生成規則

$$\langle M \rightarrow O, p \rangle \quad (M \in \mathcal{M}, O \in V_T, p = P(M|O))$$

ストロークモデル M 、筆跡特徴量の時系列ベクトル O を生成する規則で、ストローク形状認識に相当する。適用確率 p は手書きストロークがストロークモデル M である尤度と等しく、

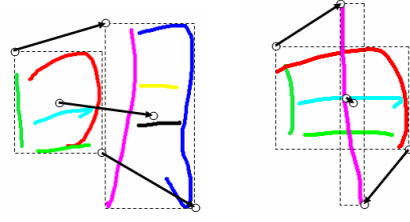


図3 明 → 日月, 申 → 日 | の位置関係の評価に用いる特徴量の例

これを以下ストローク尤度 $P(M|O)$ と呼ぶことにする。

2.4 位置尤度 P_{pos} ・境界尤度 P_{border} の計算

位置尤度計算における特徴量は図3のように、非終端記号 B, C に対応する終端記号集合をすべて包含する最小の長方形 (外接矩形) の差分をとる。具体的には、長方形 X を以下のように表現することを考える。

$$X \equiv (x_l, x_t, x_r, x_b) \quad \begin{cases} (x_l, x_t) : X \text{ の左上端点座標} \\ (x_r, x_b) : X \text{ の右下端点座標} \end{cases}$$

上に基づき、 B に対応する外接矩形を $B_b = (b_l, b_t, b_r, b_b)$ 、 C に対応する外接矩形を $C_b = (c_l, c_t, c_r, c_b)$ とおき、外接矩形の中心をそれぞれ $(b_x, b_y) = ((b_l + b_r)/2, (b_t + b_b)/2)$ 、 $(c_x, c_y) = ((c_l + c_r)/2, (c_t + c_b)/2)$ とおいたとき、以下の6特徴量と並べた特徴量ベクトル $\Delta(B_b, C_b) = (l_{lt}, l_{rb}, l_g, \theta_{lt}, \theta_{rb}, \theta_g)$ を考える。

- 左上点差分ベクトルの長さ $l_{lt} = \sqrt{(c_t - b_t)^2 + (c_l - b_l)^2}$
- 左上点差分ベクトルと水平右方向成す角
 $\theta_{lt} = \arctan(c_t - b_t / c_l - b_l)$
- 右下点差分ベクトルの長さ $l_{rb} = \sqrt{(c_b - b_b)^2 + (c_r - b_r)^2}$
- 右下点差分ベクトルと水平右方向成す角
 $\theta_{rb} = \arctan(c_b - b_b / c_r - b_r)$
- 中心点差分ベクトルの長さ $l_g = \sqrt{(c_y - b_y)^2 + (c_x - b_x)^2}$
- 中心点差分ベクトルと水平右方向成す角
 $\theta_g = \arctan(c_y - b_y / c_x - b_x)$

特徴量ベクトルが正規分布を成すと仮定すれば、以下の計算で尤度を求めることができる。

$$P_{pos}(B_b, C_b, \ell) = \frac{1}{\sqrt{(2\pi)^6 |\Sigma_\ell|}} e^{(-\frac{1}{2}(\Delta(B_b, C_b) - \mu_\ell)^T \Sigma_\ell^{-1} (\Delta(B_b, C_b) - \mu_\ell))} (1)$$

μ_ℓ, Σ_ℓ は、それぞれ学習によって得られたモデル ℓ の平均、分散である。分散は対角共分散を用いている。ただし、角度の正規分布を扱うため、 $\Delta(B_b, C_b) - \mu_\ell$ の計算結果の角度にあたる要素については次の補正を行う必要がある。 $\Delta(B_b, C_b)$ の角度要素の一つ θ と、対応する μ_ℓ の平均値要素 μ に対して、

$$\theta' = \begin{cases} \theta - 2\pi & \theta - \mu > \pi \\ \theta & |\theta - \mu| \leq \pi \\ \theta + 2\pi & \theta - \mu < -\pi \end{cases} \quad (2)$$

の変換を行い、 $\mu - \pi < \theta' < \mu + \pi$ の範囲に配置されるように変換を行う。





	→ F e G f F
	→ G A d
	→ a f h f o A
	→ G A c

図4 ストロークモデルのサブストローク連結による表現例

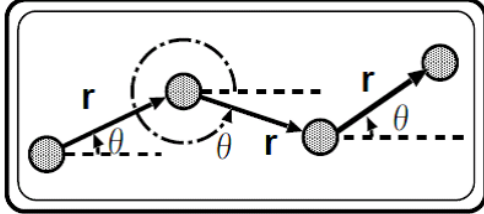


図5 HMM の特徴量 (r, θ) である筆跡の速度ベクトル ([6] より引用)

境界尤度については、漢字間の位置関係も一つの位置関係モデルとして扱うことによって、位置尤度と同様の計算で求めることができる。

2.5 ストローク尤度 $P(M|O)$ の計算

漢字の筆記に用いられるストロークモデル M を図4のように方向成分の連結として表すことにすれば、サブストロークHMMの枠組の応用としてストローク尤度の計算を行うことができる。

具体的には、ペンの筆跡座標を等時間間隔でサンプリングし、得られた座標を $(x_1, y_1), \dots, (x_t, y_t), \dots, (x_T, y_T)$ とおく。図5のように、時間系で隣り合う座標の差分を速度ベクトルとすると、速度ベクトルの長さ r_t と水平方向からの傾き θ_t は $t = 1, \dots, T-1$ について以下のように求めることができる。

- $r_t = \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}$
- $\theta =$ 水平方向とベクトル $(x_{t+1} - x_t, y_{t+1} - y_t)$ が成す角

$O = (r_t, \theta_t)$ を HMM の特徴量とし、時系列特徴量 $O = O_1 O_2 \dots O_T$ がストロークモデル M である尤度 $P(M|O)$ をストローク尤度と扱う。

2.6 漢字列認識のアルゴリズム

以下に、文脈自由文法の構文解析手法として一般的な CYK アルゴリズムを用いた漢字列認識について述べる。

まず、各ストロークに対してのストローク認識結果を尤度上位 K 番目までとり、ストロークモデル M 、ストローク尤度 p_M 、外接矩形 b_M の3項組 $[M, p_M, b_M]$ を CYK 行列の対角成分に保存する。2番目以降の対角線上の要素において以下の操作を行う。

a) 孤立漢字認識

$[B, p_B, b_B] \in T_{i, i+j-1}, [C, p_C, b_C] \in T_{i+j, i+n}$ ($j = 1, \dots, L-1$) であるような非終端記号生成規則 $p = \langle A \rightarrow BC, \ell \rangle$ を探索する。適合すれば、外接矩形を利用して位置尤度 $P_{pos}(b_B, b_C, \ell)$ を計算した上で、 $[A, p_A, b_A]$ ($p_A = p_B \cdot p_C \cdot P_{pos}(b_B, b_C, \ell)$) を $T_{i, i+n}$ に書き込む。尤度は以下のような計算で求める。

$$\log p_A = \log p_B + \log p_C + W_p \log P_{pos}(b_B, b_C, \ell) \quad (3)$$

W_p はストローク尤度に対する位置尤度の重みを表す定数で、実験的に高い認識性能が得られた $W_p = 2$ を採用している。

b) 漢字列認識

$K_B, K_C \in \mathcal{K}$ に対し、 $[K_B, p_B, b_B] \in T_{i, i+j-1}, [K_C, p_C, b_C] \in T_{i+j, i+n}$ であれば、次のような $[K_{str}, p_A, \tilde{b}_A]$ を求め、 $T_{i, i+n}$ に書き込む。

$$K_{str} \rightarrow K_B K_C$$

$$\log p_A = \log p_B + \log p_C + W_b \log P_{border}(b_B, b_C)$$

$$\tilde{b}_A = b_C$$

ここで、 W_b は境界尤度 P_{border} の重みで、評価実験では $W_b = 2$ を採用している。また、上式からもわかるように、 A の外接矩形 \tilde{b}_A として書き込むものは、漢字列自体の外接矩形 $b_B \cup b_C$ ではなく、末端に挿入した漢字の外接矩形 b_C とする。同様に、漢字列 K_{Bstr} と漢字 K_C に対し、 $[K_{Bstr}, p_B, \tilde{b}_B] \in T_{i, i+j-1}, [K_C, p_C, b_C] \in T_{i+j, i+n}$ であれば、次のような $[K_{str}, p_A, b_A]$ を求め、 $T_{i, i+n}$ に書き込む。

$$K_{str} \rightarrow K_{Bstr} K_C$$

$$\log p_A = \log p_B + \log p_C + W_b \log P_{border}(\tilde{b}_B, b_C)$$

$$\tilde{b}_A = b_C$$

ここでも、 $\tilde{b}_A, \tilde{b}_{Bstr}$ はそれぞれ漢字列の外接矩形でなく、終端の漢字の外接矩形であることに注意する。このように、漢字列と漢字の位置関係の評価が漢字列の末端との評価になっているのは、われわれが字を手で書いていく際の文字の配置が直近の文字との位置関係に最も影響を受けることに対応しており、妥当な仮定であると考えられる。また、漢字と漢字の位置関係の評価に帰着できることによって、境界尤度計算に必要な位置関係の学習において、漢字2字の間の位置関係を学習を行えばよいので、学習データが少量で済み、効率性が期待できる。

このように続けていき、最終的に $T_{N,1}$ に書き込まれた要素 $[A, p_A, b_A]$ に対して、生成規則 $S \rightarrow A$ が存在した場合、漢字または漢字列 A を認識結果とし、特にこのような A のうち p_A の最も高いものを一位認識結果とする。

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) のオンライン手書き文字データベース (図6) を用いて行った。このうち、 α セットは正規の筆順・画数で書かれた漢字、 β セットは自由な筆順・画数で書かれた JIS 第一・第二水準の漢字、 $\zeta-1$ セットは自由な筆順・画数で書かれた漢字・ひらがな・英数字を用いた文字列によって構成される。われわれは、学習と性能評価を α セット、多様な筆順・字種に対する評価を β セットを用いて行い、漢字列については $\zeta-1$ データのうち教育漢字のみで構成された漢字列 (2~3字) について評価を行った。

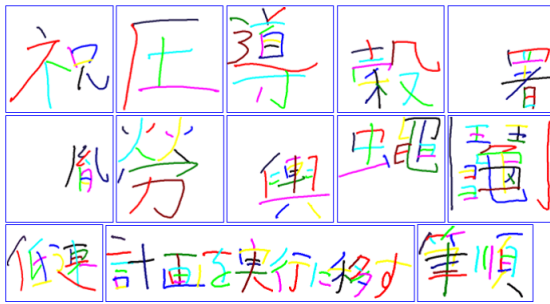


図 6 オンライン手書き漢字データベースの例

JAIST-IIPL 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。

表 1 作成した辞書の仕様

辞書番号	対応字種数	生成規則数	ストロークモデル数
(1)	1,016	1,885	17
(2)	1,016	2,507	34
(3)	2,965	5,315	34
(4)	6,353	9,610	34

3.2 漢字構造文法辞書

非終端記号生成規則 $\langle A \rightarrow B C, l, p \rangle$ に対応し、漢字、文字部品の構造を $[A, B, C, l]$ の形式で羅列したものを文法辞書と呼ぶ。性能評価のため、われわれは表 3.2 のように 4 種類の辞書を作成した。以下、簡単に各辞書の概略を述べる。

(1) 正規筆順辞書

新旧教育漢字の正しい筆順に基づく生成規則のみを並べている。位置関係の学習にも利用する。

(2) 異筆順辞書

新旧教育漢字の多様な筆順に対応している。 $\gamma - 1$ データの 20 筆者の筆順に対応している。評価実験では、提案システムの多様な筆順への対応の様子を示すために利用した。

(3) 第一水準辞書

われわれが書籍、新聞、雑誌等で日常生活見る頻度が比較的高い、JIS 第一水準漢字（新旧教育漢字も含まれる）までに対応した辞書。 β データ 20 筆者分の筆順に対応している。

(4) 第二水準辞書

第一水準漢字に加えて、われわれが日常生活で見る頻度が比較的低い JIS 第二水準漢字までに対応した辞書。 β データ 20 筆者分の筆順に対応している。

3.3 正しい筆順で筆記された新旧教育漢字の認識による基礎性能評価

最初に、辞書 (1) を用いて、 α データセットの新旧教育漢字を認識対象として、基礎的な認識性能について評価を行った。学習に使用したデータで評価を行ったもの (CLOSED) と、学習データとは異なる筆者で行ったもの (OPEN) についての認識率・10 位までの累積認識率は表 2 のようになった。

表 2 を見ると、提案手法によるシステムは、学習に用いた

表 2 新旧教育漢字 (α) データの累積認識率

評価データ	N 位累積認識率 (%)				
	1 位	2 位	3 位	5 位	10 位
CLOSED(10 筆者)	99.42	99.91	99.95	99.97	99.97
OPEN(30 筆者)	99.34	99.82	99.88	99.90	99.90

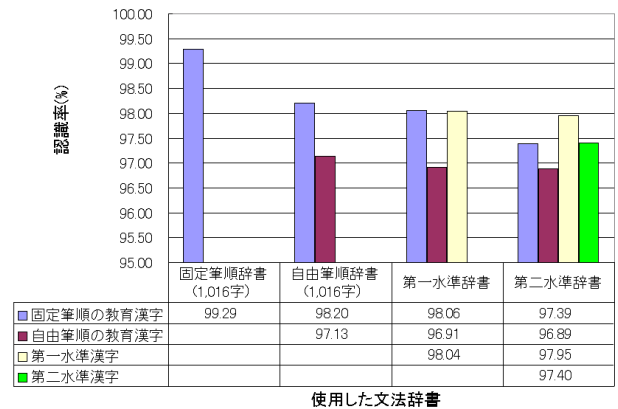


図 7 さまざまな文法辞書を適用したときの各字種集合の認識率

筆者かそうでないかにほとんど依存せず、高い認識率を得ることがわかる。CLOSED データと OPEN データを比較すると、OPEN データがわずかに低いものの、1 位認識率では 0.08% の差に留まっている。以上から、提案システムは筆者毎に異なる位置関係の揺らぎに対して比較的頑健であることが確認できる。

3.4 多様な筆順・字種の追加と認識精度の関係

次に、正しい筆順だけでなくさまざまな筆順を追加した辞書や、教育漢字以外の第一水準漢字や第二水準漢字を追加した辞書を用いて、性能評価を行った。評価に用いた素材については、固定筆順の評価には α セット、自由筆順漢字・第一水準漢字・第二水準漢字の評価には β セット、各 20 人分である。結果を図 7 に示した。実験で確認したい事項は以下の 2 点である。

- 提案手法のシステムで、字種が大きく増加した場合に十分な認識精度を得られるかどうか

- 字種・筆順の増加によって探索範囲が大幅に拡大し、曖昧性が増す。その際、性能にどのような変化があるか

図 7 を見ると、自由筆順の新旧教育漢字・新旧教育漢字以外の第一水準漢字・第二水準漢字について、生成規則を追加した新しい文法辞書を用いることによって、比較的高い認識率が得られていることがわかる。このことから、平易な漢字構造記述に基づいて字種を辞書に追加するだけで、未学習字種・筆順に対応できるとした、本手法の目的の一つを達成したと言えるだろう。ただし、自由筆順辞書を適用した際の認識率がやや低いため、今後、この辞書については生成規則の見直しを行う必要がある。

認識対象範囲を広げるため、多様な筆順・字種を追加すると構文解析時の探索範囲が広がり、文法の曖昧性が増す。これに伴い、追加前に認識できていた字種の認識精度の低下が予想される。これを調べるため、筆順の正しい新旧教育漢字の

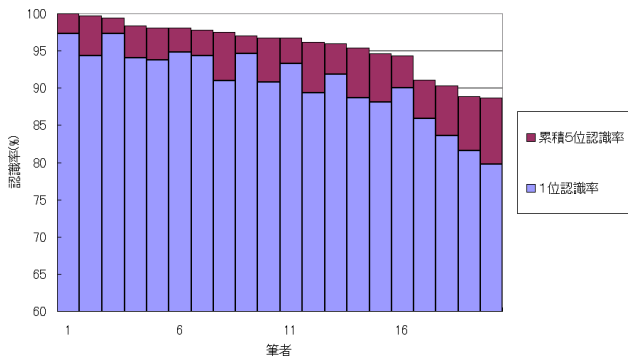


図 8 評価データ筆者毎の文字列正解率

データセットに対して、自由筆順辞書、第一水準辞書、第二水準辞書を適用して認識実験を行い、精度の変化を見た。グラフ 7 の結果を見ると、探索範囲が広がるに従い、認識率は 99.29% → 98.20% → 98.06% → 97.39% と低下していくのが確認できる。認識精度が大幅な低下を示していないため、提案手法は文法追加に対して比較的頑健と言えるが、今後はストローク間位置関係をより正確に表現、識別できるような特徴量の検討が不可欠であろう。

3.5 漢字列認識

この項では漢字列認識システムの認識率と誤認識例について評価して、提案システムの認識性能を確認した。今回、認識対象は $\zeta-1$ セットに含まれる文字列のうち、新旧教育漢字のみで構成された熟語（2～3字）で、誤字や余分な点を含むデータを除いたもの 20 筆者分である。辞書は自由筆順辞書を用いた。

結果文字列が正解文字列と同一である割合を文字列正解率として評価した。文字列正解率は平均 90.80% で、筆者毎の詳細は図 8、誤認識例を表 3 に示した。本実験では、評価漢字列データ全ての筆順を登録する作業を行っていないため、データ中頻繁に用いられている部首の筆順が辞書に登録されていない場合、認識精度が大きく下がることが想定された。各筆者文字列認識率（図 8）を見ると、筆者毎に認識率のばらつきがあるが、全体的に見ると文字列認識率 90% 前後で推移している。これは、サブストローク HMM において文字境界のためのペンアップモデルを採用することで文字列認識に拡張した須藤ら [6] の行った、教育漢字で構成された二字熟語の認識率 81.12% と比較して性能が向上したと言える。

主な誤認識例を示した表 3 を見ると、誤認識の原因はいくつかに分類できる。(1) 筆順がそもそも登録されていないことによる誤り、(2) 漢字単位でのストローク間位置関係の誤認識、(3) 漢字間位置関係の誤認識である。(2) の問題は、教育漢字の多様な筆順を追加し、曖昧性が増したためである。位置特徴量の妥当性について考える必要があるだろう。また、(3) の問題は、境界尤度重み W_b の値を変化させて実験を繰り返し、最適な W_b を求めることで、認識性能は向上すると考えられる。

4. おわりに

本論文では、確率文脈自由文法に基づき、文字境界と孤立漢

表 3 漢字列認識実験における誤認識例（20 筆者分）

正解漢字列	→	誤認識例	誤認識率
向上	→	向土、局土、白上	90%
工学	→	土学	75%
任意	→	件意、弁意	50%
感謝	→	感言射、感言身子	25%
線分	→	糸泉分、線欠	20%
不注意	→	友注意、木注意、不玉退八	25%
観測	→	声責測、考現測、勸科利、知現測	20%
音節	→	音建刀、音算、立由節、音食戸、専節	30%
一人	→	大	75%
一方	→	切	30%
中心	→	忠、中火	15%
帰納	→	帰糸分、白虫納、帰糸分、帰糸弓	20%
長短	→	長矢豆	10%

字認識を統合的に扱う漢字列認識手法を提案した。漢字間・漢字内部の文字部品間の位置関係を効果的に利用することによって横書き漢字列に対して高い認識率を確認することができた。

今後は、ひらがな・カタカナ・英数字など漢字以外の字種が使われた文字列に対しても本手法が適用できるよう生成規則を定義し、漢字認識システムを汎用性のあるものとしていきたい。また、今回は筆記方向を横方向のみに限定したが、これを縦書き、斜め書き、途中の改行にも対応させるため、新しい文字間位置関係モデルを考えて導入したい。

また、比較的長い文字列については、計算量が $O(n^3)$ の CYK アルゴリズムではリアルタイム認識が困難になるおそれがある。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きたい。

5. 謝 辞

本研究は、JAIST-I IPL(<http://iipl.jaist.ac.jp/>) の研究と手書き漢字データベースに基づいて行われた。われわれは、JAIST-I IPL のメンバーに深く感謝したい。

文 献

- [1] M. Nakai, N. Akira, H. Shimodaira and S. Sagayama: "Sub-stroke Approach to HMM-based On-line Kanji Handwriting Recognition," Proc. of ICDAR'01, pp. 491-495, 2001.
- [2] 大田郁実, 山本遼, 酒向慎司, 嵯峨山茂樹: "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," 電子情報通信学会技術研究報告, PRMU2007-106, pp. 49-54, 2007.
- [3] I. Ota, R. Yamamoto, S. Sako and S. Sagayama: "Online handwritten kanji recognition based on inter-stroke grammar," Proc. of ICDAR'07, 2, pp. 1188-1192, 2007.
- [4] 仙田修司, 濱中雅彦, 山田敬嗣: "切り出しパラメータが学習可能なオンライン手書き文字切り出し手法," 信学技報 PRMU97-219, 1998.
- [5] 福島貴弘, 中川正樹: "確率モデルに基づくオンライン枠なし手書き文字列認識," 信学技報 PRMU98-139, 1998.
- [6] 須藤隆: "隠れマルコフモデルに基づくオンライン手書き文字列認識に関する研究," 修士論文 北陸先端科学技術大学院大学, 2002.