

H-005

文字構造を利用した筆順探索の高精度化

Stroke-Order Search Algorithm Based on Hierarchical Kanji Structure

鈴木 綾[†]中井 満[†]

Aya Suzuki,

Mitsuru Nakai

1. はじめに

手書き入力インタフェースの登場により、コンピュータを用いた漢字学習が普及している。筆順の学習では手本をなぞる手法が主流である。普段の書き慣れた筆跡でも筆順判定できることが望ましいが、画対応を探索するため、多画数になると膨大なメモリを要することが問題となる [1]。本研究では、多画数文字の筆順探索精度を定量的に分析し、更には偏旁冠脚などの文字の構造に着目した高速化を図る。

2. 画の対応付けによる筆順探索

2.1. 筆順探索法の概要

基本的には全ての画の並び替えの中から最適な系列を探索する。本研究では筆跡の座標値と速度ベクトルからなる時系列パターンを特徴量とし、隠れマルコフモデル (HMM) を用いて画をモデル化し、尤度最大基準で探索する [2]。図 1 は画数 $N = 6$ の漢字“印”の例であり、それぞれの画 ($r1 \sim r6$) を left-to-right 型の HMM で学習する。図 2 はこの漢字の筆順探索ネットワークの一部であり、ノードは画の HMM を、番号は画番号を表す。図 1(b) の自由な筆順で書かれたパターンのように画数が変動することもあるので、探索過程ではペンの上げ下げ情報は使用しないことにする。つまり、ある時刻に筆記されている画が何画目であるかは事前に分からないので、全てのノードの状態にいる仮説を立てなければならない。そこで、本稿では探索ネットワークのノード数を探索空間の広さとする。なお、図 1(b) のパターンは尤度最大基準の Viterbi 探索によって、 $r1-r2-r4-r3-r6-r5$ と認識される。問題は漢字が多画数なことであり、教育漢字の最大の画数は 20 画である。事前に全ての画順をネットワークとして展開するのは非現実的なので、仮説を動的に展開し、常に上位 M 個 (ビーム幅) だけを残すビーム探索を使って、準最適解を求める。

2.2. 筆順探索実験

JAIST IPL オンライン手書き文字データベース中の自由な筆順で書かれた γ_1 セット [3] を用いて実験を行った。1,016 字種の新旧教育漢字について 68 名が筆記しているが、明らかな書き間違いを除外して合計 68,950 文字を使用した。このうちの 4 名分についてはビーム幅 $M = 10,000$ の条件で探索した結果とサンプルの筆順を視察で比較して、ほぼ正しく探索できていることを確認した。したがって、以降の実験ではビーム幅 $M = 10,000$ の条件で探索した筆順を正解として扱う。高速化を目的にビーム幅を小さくしたときの探索成功率と画数の

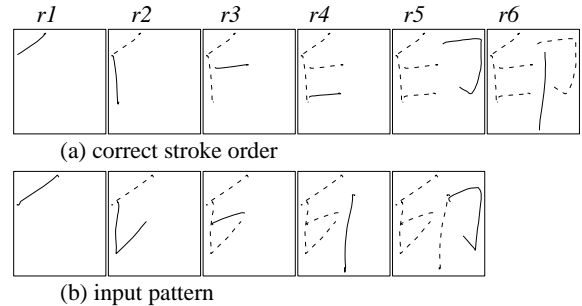


図 1: 漢字“印”の正しい筆順 (a) と自由な筆順 (b)

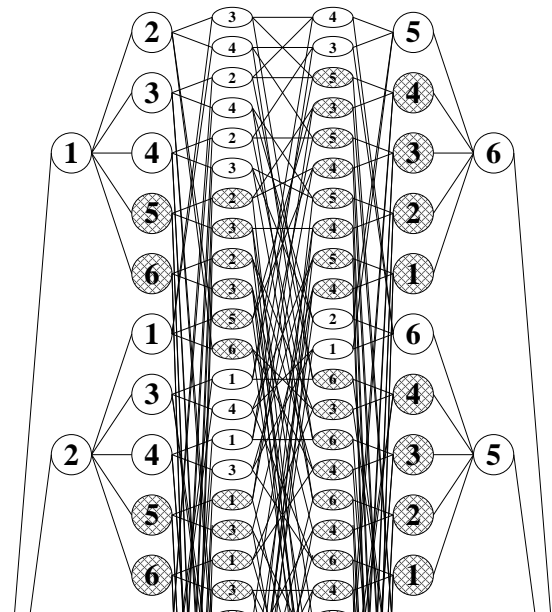


図 2: 漢字“印”の探索ネットワーク

関係が表 1 である。例えば 95% 以上の成功率を達成するためには、5 画以下の漢字では $M = 5$ でもよいが、17 画以上の漢字では $M > 5,000$ が必要となる。

3. 文字構造を利用した筆順探索

漢字は偏旁や傍のような基本的な文字部品で構成され、それらは図 3 のような階層的な構造を持つ。そこで、筆順誤りは文字の部品の中で起こり、一つの部品を書き終わる前に他の部品の画を書くことはないかと仮定する。例えば漢字“印”を偏と旁に分割した場合、初めに画 $r1$ が書かれると、次の系列として画 $r5$ 、画 $r6$ への展開は許されない。図 2 中の塗り潰したノードは展開の許されないノードであり、探索空間が著しく削減できる。

[†]富山県立大学, Toyama Prefectural University.

表 1: ビーム幅・画数毎の筆順探索成功率 [%]

画 \ 幅 M	5	10	50	...	1,000	5,000
:	:					
4	96.3	98.2	99.8	...	100.0	100.0
5	95.0	98.2	99.9		100.0	100.0
6	92.8	96.7	99.6		100.0	100.0
:	:					
16	55.0	68.8	84.2		95.8	98.5
17	45.6	60.7	80.5		93.5	97.1
18	47.1	66.1	82.0		94.7	97.7
19	40.3	58.2	77.1		92.6	96.8
20	41.6	52.5	73.8	...	91.1	97.5
計	79.6	87.5	95.4	...	99.2	99.8

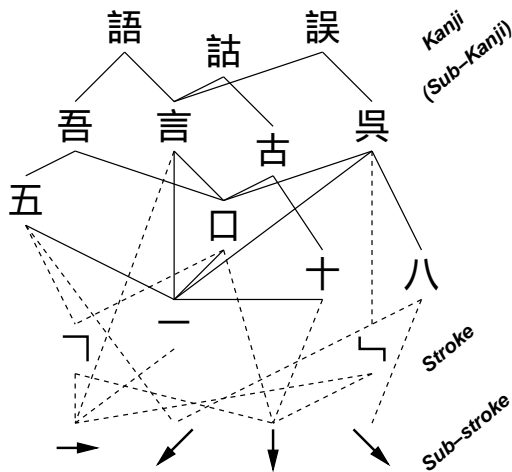


図 3: 漢字の階層的な文字構造

文字認識用に作成した文字構造 [4] を適用すると、20 画の漢字の場合では探索空間が 10,485,760 から 1,656 まで減少し、約 6,300 分の 1 にまで削減された。漢字“印”において $r1-r5$, $r1-r6$ などの系列で始まる筆順誤りが存在すれば文字構造に合致せず、探索不可能という結果になる。その文字数は全 68,950 文字のうち 3,176 文字存在する。したがって探索成功率は最大でも 95.4% にまでしかならないということが分かる。そこで、文字部品の画数 T を閾値とし、それ以下の文字構造を部品に分割しないという方法を用いる。

例えば画数 $N = 13$ の漢字“試”において閾値 $T = 6$ とすると、画数 $N = 7$ の部品“言”は分割されるが、画数 $N = 6$ の部品“式”はそれ以上分割されない(図 4)。部品の分割を増やす毎に探索空間が削減されるので、探索成功率は上昇すると期待される。しかし、例えば図 5 のような筆順誤りが生じた場合、閾値 $T = 5$ 以下では文字の部品をまたぐ筆順誤りとして探索不可能になるので、ある閾値 T を境に成功率も減少すると予想される。

文字構造を利用した実験結果は表 2 の通りであり、ビーム幅毎に最適な閾値があることが分かる。したがって、用途やコンピュータの性能に応じて文字構造の利用範囲を切り換えることが望ましいことが分かる。

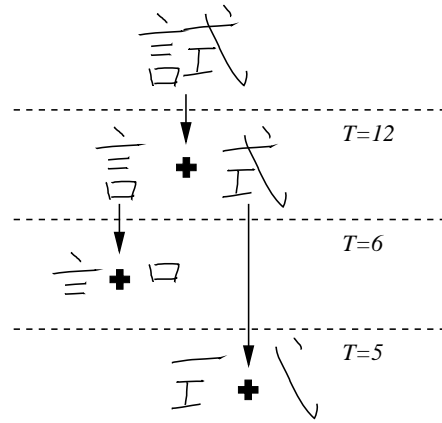


図 4: 部品画数 T を閾値とした構造の利用範囲

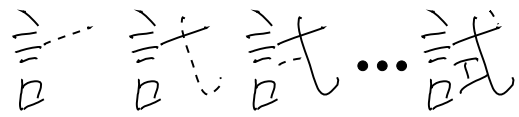


図 5: 漢字“試”の筆順誤り

表 2: ビーム幅 (M) 毎の最適な閾値 (T)

幅 M	5	10	50	...	1,000	5,000
閾値 T	5	7	9	...	12	17
成功率 [%]	82.3	89.2	96.3		99.3	99.8

4. おわりに

本研究ではビーム幅 $M = 10,000$ を基準としたが、画の挿入、脱落などがある文字に対しては正しく探索できていない。また、1つの画を2画に分けて間に他の画が書かれるというような筆順違いには対処できていない。今後は、このような入力パターンにも対応するよう検討が必要である。また、画として認識されていた余計な点などのノイズの影響は、文字構造の利用により改善が予想されるので、その効果についても調査する。

謝辞 本研究は平成 19 年度科学技術振興機構シーズ発掘試験課題 07-003 として行われた。

参考文献

- [1] 蔡文杰, 他, “部首単位標準パターンとキューブサーチに基づく筆順フリーなオンライン文字認識アルゴリズム,” 信学論, D-II Vol.J88-D-II, No.7 (2005-07).
- [2] 中井, 他, “ストローク HMM オンライン文字認識方式の階層構造辞書を用いた筆順違い対策,” 信学技報, PRMU2002-234 (2003-03)
- [3] 秋良, 中井, 他, “ストローク HMM を用いたオンライン非目視手書き文字認識の性能評価,” 信学技報, PRMU2000-206 (2001-03)
- [4] 中井, 他, “サブストローク HMM を用いたオンライン手書き文字認識,” 信学論, vol.J88-D-II, no.9 (2005-09)