

スタイルフリーな手書き文字列検索機能

織田英人[†] 伊藤禎宣[†] 中川正樹[†]

Style Free Handwritten Text Search

HIDETO ODA[†] SADANORI ITO[†] MASAKI NAKAGAWA[†]

1. はじめに

ペンで『書く(描く)』行為は有史以来、知的生産性の高い行為と捉えられてきた。これを電子化する手書き入力インタフェースの研究は、これまで TabletPC や PDA などハードウェア、ソフトウェアの両面から行われ、手書きによるインタラクティブな創造活動の可能性を示唆している。このような手書き入力インタフェースの普及に伴い、今後個々のコンピュータやネットワークサーバ上に手書きの情報(手書きボタン)が蓄積されていくことが予想される。これを効率的に参照したり再利用したりできるように、本稿では手書きボタンに対する全文検索手法を提案する。

2. 手書き文字列検索の設計

本稿では、図1に示すような、様々な筆記方向の文字列や図形が混在するスタイルフリーな手書きボタンを扱う。提案する手書き文字列検索機能では、このような手書き文書に対し、検索キーワードとして入力された文字コード列に該当する部分を出力する。

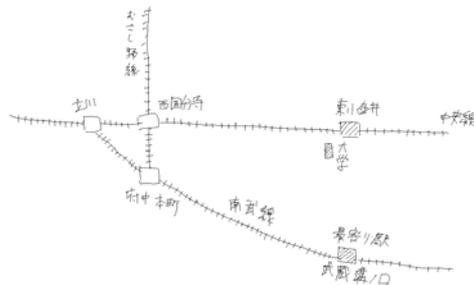


図1 手書きボタン
Figure 1 Handwritten Pattern.

2.1 文字列の抽出

本研究では、スタイルフリーな文書を収集する際、文字や図形、数式など手書きボタンの種類ごとに個別に筆記枠を設けない。よって、まず、手書きボタンから検索対象である文字列ボタンを抽出する必要がある。

我々は、既に、手書きボタンを文字、図形、数式に分離する手法について提案している[1]。筆記されたストローク(筆画)の大きさに着目すると、図形を構成するストロークは文字を構成するストロークよりも大きい、または、大きいストロークと接触しているという特性がある。これを利用し検索対象となる文字列ボタンを抽出する。

2.2 筆記方向の推定

西洋言語とは異なり、日本語には、横書きと縦書きの両方が存在する。また、黒板やホワイトボードの板書などは、文字列が斜めに傾く可能性を否定できない。

我々は、これまでに文字列の筆記方向を推定する手法を提案している[2]。筆記方向は個々の文字の向き(文字方向)と文字列の向き(行方向)から成る。ストロークの移動方向に着目し、これの分布を取ると、2つのピークが約90度離れた位置に出現する。この特性を用いることで、文字方向と行方向が推定できる(図2)。ピークが1つしか出現しない場合は、行方向と文字方向が一致していることを意味する。

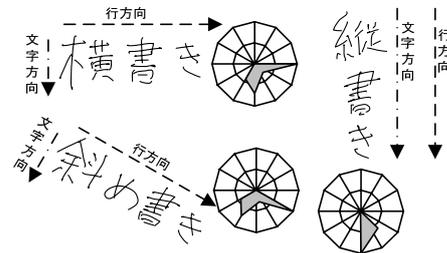


図2 文字方向と行方向
Figure 2 Character Direction, and Line Direction.

[†] 東京農工大学
Tokyo University of Agriculture and Technology

2.3 手書き文字列検索

検索の対象となる文字列パターンは、位置や方向特徴による時系列データであるため、文字コード列であるテキストと直接比較できない。そこで、手書き文字列認識技術を用いて、手書き文字列パターンを予めテキストデータに変換する必要がある[2]。しかし、文字列認識技術では、認識結果を一意に決定してしまうため、誤認識が生じた場合、検索対象が認識結果のテキストデータに含まれない危険性がある。

例えば、図3のような文字列を認識する場合、認識結果として「公太郎」と「ハム太郎」のどちらの可能性もあるため、それを一意に決定するのは困難である。また、多くの漢字は偏と旁から構成されるため「女子学生」が「好学生」などと誤認識される場合もある。

この問題に対し、我々は、図3に示すように、与えられた手書き文字列パターンの文字認識候補と分割候補を複数個生成し、これに対して検索を行うことで対処する。

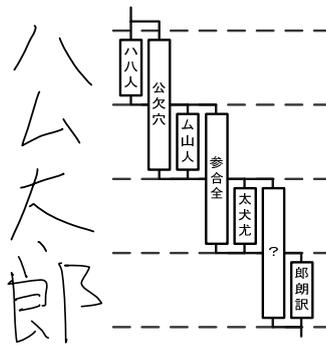


図3 認識候補ラティス

Figure 3 Recognition Candidate Lattice.

図3に示すネットワークを認識候補ラティスと呼ぶ。このノードとパスは、分割位置の候補を表現している。また、個々のノードには、各文字の認識候補と認識の確からしさを表す認識スコアが格納される。

認識候補ラティス内の、検索キーワードに適合する部分を探索することで、手書き文字列検索は実現される。これは、動的計画法の一種である Viterbi アルゴリズムによって実現される。

3. 性能評価

2章で提案したスタイルフリーな手書き文字列検索技術のうち、本稿では、2.3で述べた手書き文字列検索機能の性能についてのみ評価を行った。実験用データのうち、手書き文字パターンとして、当研究室が所有す

るオンライン手書き文字パターンデータベース TUAT Nakagawa Lab.HANDS-kuchibue (総筆記者数 120, 筆者一人当たり 11,962 文字パターン) [3]の文字パターンから生成した擬似的な横書きの文字列パターンを用いた。また、検索キーワードとして、無作為に抽出した 2~4 文字から成る単語 (各々 1,000 個) を用いた。実験結果を表1に示す。

表1 実験結果

Table 1 Experimental Result.

	再現率	適合率	F 尺度
2 文字	79.9%	75.0%	0.774
3 文字	86.1%	88.4%	0.872
4 文字	88.5%	94.2%	0.912

表1において、再現率は検索漏れの少なさを示し、検索が適合した文字列の総数と検索対象中に実際に含まれる適合する文字列の総数の商で求まる。適合率は検索ノイズの少なさを示し、検索が適合した文字列の総数と、検索ノイズを含む検索された情報の数の商で求まる。F 尺度はこれら相反する要素を総合的に評価する尺度で、両者の調和平均によって求まる。

実験の結果、全ての長さの検索キーワードで F 尺度が 0.75 を上回った。このことから、提案手法が実用にある程度、耐える精度を持っているといえる。

4. おわりに

本稿では、手書き文字列に対する全文検索手法を提案した。今後の課題として、様々な筆記方向の文字列や図形などが混在した実社会で筆記され得る手書きパターンを用いた評価が挙げられる。

謝辞 本研究の一部は、情報処理推進機構、2006年度下期未踏ソフトウェア創造事業の支援による。

参考文献

- 1) Mochida, K. and Nakagawa, M.: Separating Figures, Mathematical Formulas and Japanese Text from Free Handwriting in Mixed On-Line Documents, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 18, No. 7, pp.1173-1187, 2004.
- 2) Nakagawa M., Zhu B. and Onuma M.: A Model of On-line Handwritten Japanese Text Recognition Free from Line Direction and Writing Format Constraints, IEICE Trans. Inf&Syst., Vol.E88-D, No.8, pp.1815-1822, August 2005.
- 3) Nakagawa, M., Kaoru, M.: Collection of on-line handwritten Japanese character pattern database and their analyses, Int'l J. Document Analysis and Recognition, vol. 7, no. 1, pp. 69-81, 2004.