

6U-3 パターンマッチングを利用した筆順教示システム

郡司圭子† 横田登志美† 葛貫壮一郎† (株)日立製作所 日立研究所†

村尾龍之介‡ 齊藤幸夫‡ (株)日立情報システムズ‡

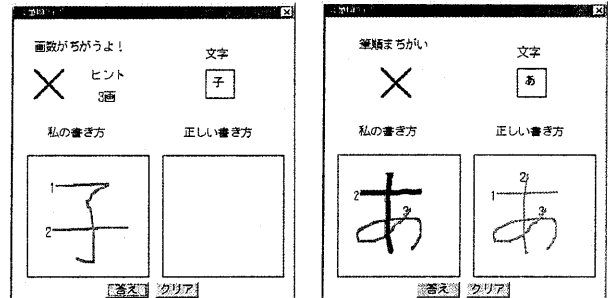
1. はじめに

文章を読み書きする能力は、小学校で漢字を習得した度合いに大きく影響されることが知られている。しかし、有限の授業時間を漢字の反復練習に使うことは難しく、習得にかかる時間も個人差が大きいため、個別に学習できるシステムが求められている¹⁾。

私どもは、オンライン手書き入力の特徴を生かし、紙に書かれた後では判定できない文字の筆順・画数を、パターンマッチングの技術を利用して教示するシステムを試作し、評価を行なった。

2. 筆順・画数教示システムのインタフェース

本システムでは、学習したい文字を、正しいと思う筆順・画数で手書き入力するだけで、特別な指定をせずに学習できるようにした。図1(a)は、本システムに画数誤りの文字を入力した例である。筆順を確かめたい文字「子」を手書き入力し、[答え]のボタンを押す。システムは、ユーザの書いた文字「子」を認識し、文字の画数・筆順をチェックし、判定結果を表示する。この例では、誤った画数(2画)で入力したため、正しい画数(3画)をヒントとして表示し、再入力するように促す。図1(b)は、本システムに筆順誤りの文字を入力した例である。筆順を確かめたい文字「あ」を手書き入力し、答えのボタンを押す。図の例では、文字「あ」を縦棒から先に[縦棒→横棒]と書いたため、システムは筆順誤りと判定する。そして、ユーザの筆跡を、筆順の誤り箇所の色を変えてトレース表示する。さらに、右側に正しい書き方をトレース表示する。本システムでは、ユーザの筆跡を、筆順の誤り箇所は色を変えてトレース表示するため、ユーザは自分の誤りを容易に認知できる。さらに、その後で、正しい筆順を繰り返してトレース表示することで、効果的に正しい筆順が覚えらる。



(a) 画数誤りの指摘例

(b) 筆順誤りの指摘例

図1 本システムの画面例

3. 筆順・画数の教示方式

図2に、筆順・画数の教示方式の処理フローを示す。まず、ユーザが手書きで文字を入力する(1)。次に、システムは、ユーザが手書きした文字の位置・サイズを正規化する(2)。そして、ユーザの手書きの筆跡を折れ線近似する(3)。折れ線近似したパターンを、予め、同様に位置・サイズを正規化し、折れ線近似されている正解パターンとパターンマッチングする(4)。すなわち、折れ線近似した入力の各線分に対応する正解パターンの線分を、各線分の始終点座標の差の絶対値の和(シティブロック距離)が最小となるように求める。このシティブロック距離が最小となった文字を、文字認識結果とする²⁾。入力パターンの画数を認識結果の正解パターンと比較し、画数が正しいならば(5)、正しい筆順と筆順の誤り箇所を教示するために、筆順、画の対応情報を生成する(6)。本方式は、筆順・画数が誤った入力も認識できるよう、画の情報は用いず、折れ線の各線分毎に対応する線分を求めるマッチング方式としている。そこで、入力の線分と正解パターンの線分の対応情報から、入力の画と正解パターンの画の対応情報を生成する。図の例では、「あ」の3画目の丸い部分の各線分は、正解の3画目の各線分と対応するため、入力の3画目は正解パターンの3画目と対応すると判

Teaching correct stroke order of KANJI characters using on-line pattern recognition techniques

† Keiko Gunji, Toshimi Yokota, Soushiro Kuzunuki, † Hitachi, Ltd. Hitachi Research Laboratory

‡ Ryunosuke Murao, Yukio Saito, ‡ Hitachi Information Systems, Ltd.

定する。同様に、左側の入力の1画目は正解の2画目と対応し、2画目は正解の1画目と対応する。従って、1・2画の筆順が誤っていると指摘し、正しい筆順を教示する(7)。一方、入力の画数が正しくないならば(5)、正しい画数をヒントとして教示し(8)、再入力を促す。

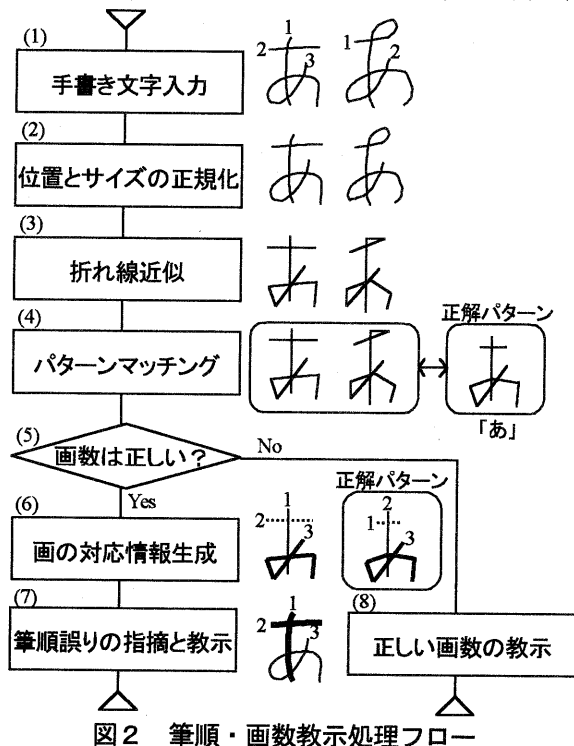


図2 筆順・画数教示処理フロー

4. 評価

本方式による筆順・画数の教示精度を、東京農工大学が収集したオンライン手書き文字データベースを用いて評価した。まず、比較的楷書で書かれた2人分のデータ(Mdb0066,Nky0101)から、小学校1～3年で学ぶ文字のうち200文字を抽出し、人手で(A)正しい画数・筆順で書かれたデータ、(B)正しい画数だが誤った筆順で書かれたデータ、(C)誤った画数で書かれたデータに分類した。その結果、各データの割合は、(A):68.5%、(B):10.5%、(C):21.0%であった(図3、最左)。

次に、本画数・筆順教示方式で上記データを判定させ、(A)～(C)がそれぞれ正しく判定できるか評価した。

その結果、(A)正しい画数・筆順で書かれたデータを、正しく判定できたのは78.8%であった。正しく判定できなかったうち、筆順誤りと誤判定したのが3.6%、別の文字に誤認識したのが17.5%であった(図3(A))。

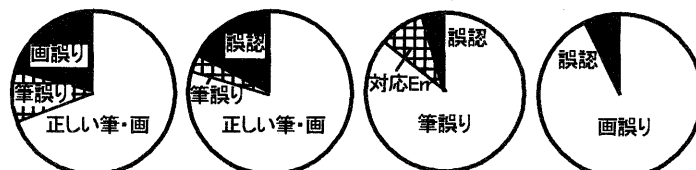
(B)正しい画数だが誤った筆順で書かれたデータを、正しく判定できたのは85.7%である。正しく判定でき

なかったうち、筆順教示に必要な画の対応情報が生成できなかったのが9.5%、別の文字に誤認識したのが4.8%であった(図3(B))。

(C)誤った画数で書かれたデータを正しく判定できたのは92.9%であった。正しく判定できなかったのは、別の文字に誤認識したデータで7.1%あった(図3(C))。

(A)～(C)総合では、筆順・画数を正しく判定できるのは82.5%と、実用的なレベルであることが分かった。

誤判定の原因を調査したところ、(A)正しい画数・筆順で書かれたデータを筆順誤りと誤判定したデータ、(B)正しい画数だが誤った筆順で書かれたデータで、筆順教示に必要な画の対応情報が生成できなかったデータとも、文字のバランスが正解パターンと異なるために、位置とサイズの正規化の際に位置がズレてしまい、本来、対応すべき折れ線の線分とは異なる線分を対応させる方が、パターン全体のシティブロック距離が小さくなり、線分の対応誤りを起こしたことが分かった。



データの割合 (A)正しい筆順・画数の判定結果 (B)筆順誤りの判定結果 (C)画数誤りの判定結果

図3 筆順・画数教示の判定精度

5. 今後の課題

正しく筆順・画数を教示できなかった原因が、誤認識と、正規化の際の位置ズレによる折れ線の対応誤りであったことから、正規化後の位置に依存しないパターンマッチング方式を立案し、筆順・画数の教示精度を向上してゆきたい。また、画数が異なる場合には、正しい画数のヒントのみでなく、画数の誤り箇所(続け書き・切り書き箇所)の指摘もできるようにしたい。

[謝辞]本研究に関して御指導いただいた東京農工大の中川教授、および、関係各位に感謝申し上げます。

[参考文献]

- 1)龍岡亮二, 吉本ミツ, “外国人・小学生を対象にした漢字教育支援システムの開発,” 信学技報 ET96-36, pp.103-110, Jun, 1996.
- 2)横田登志美, 郡司圭子ほか: 筆順画数フリー文字認識の開発: 第55回情処学全大, pp.2-188-189, Sep.1997.