

オンライン認識とオフライン認識の候補統合による 1D-2 ハイブリッド型手書き文字認識

田中 宏* 中島 健次* 石垣 一司* 秋山 勝彦** 中川 正樹**
(株)富士通研究所 東京農工大学 工学部**

1. はじめに

PDAやペンPCなどの環境でキーボードの代わりに文字入力をするためのオンライン手書き文字認識において、高精度な認識性能実現のためにストローク情報に基づくオンライン型文字認識とイメージ情報を用いるオフライン型文字認識を組み合わせるアイデアは従来から知られている（日本特許公開 昭 55-140970）[1]。また、オンライン認識の大分類としてオフライン的な特徴を使うやり方も多くの実績がある。しかしながら、両者を対等にハイブリッド化して実用的な速度、性能を実現した例はほとんど知られていない。

実用的なハイブリッド認識エンジンの実現には、まず個々の認識エンジンが単体でも高い認識性能を持つこと、認識速度が高速で並行動作しても実時間動作が可能であること、認識辞書が比較的コンパクトで辞書を二重に持つことが製品上障害にならないレベルであること、そして何よりも認識特徴が相補的で候補の統合処理が効果的に実現されやすいことが必要である。

今回、われわれは富士通研究所で開発された288次元圧縮特徴ベクトルによるオフライン認識エンジンと、東京農工大で開発された線形時間伸縮マッチング法による続け字対応のオンライン認識エンジンを統合し、速度・辞書サイズとも実用的で高い認識性能を持つハイブリッド型ペン入力手書き文字認識エンジンを開発した。

2. 認識エンジンの構成

図1に本認識エンジンの構成を示す。

オンライン型認識には、農工大の中川研究室で開発した線形時間伸縮マッチング(LTM)に基づく認識エンジン[2]を採用した。オフライン型

認識には、当社で開発した文書OCRエンジン[3]を採用している。

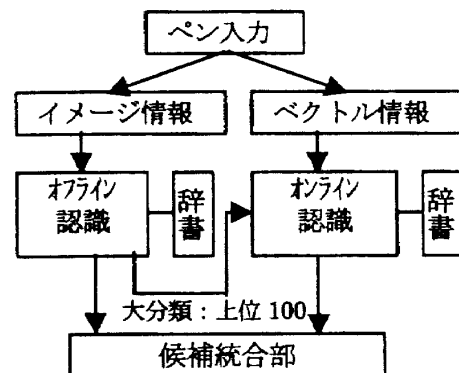


図1. 認識エンジンの構成

入力された文字パターンはイメージ情報に変換されてオフライン認識され、その上位100候補がオンライン認識部に渡る（大分類）。この100候補についてオンライン認識が行われて認識結果が得られる（続け字認識）。従来はこの結果が最終的な認識結果であった。

本認識エンジンでは、さらにオフライン認識の認識候補が結果に統合される（候補統合）ことにより、オンライン認識には不得手な文字変形（画の脱落、塗り潰しなど）を起こした文字でも認識することが可能となっている。

3. ハイブリッド候補統合

2つの異なる認識エンジンの結果を統合する方法として、[方法1] 認識スコア（尤度）は無視して候補の順位のみを信用する、[方法2] 認識スコアを正規化して候補を混ぜる、の2種類を検討した。

方法1では、オンライン認識の候補とオフライ

Hybrid Handwriting Character Recognition Based On Integration of Online-Offline Character Recognition Candidates

Hiroshi Tanaka, Kenji Nakajima, Kazushi Ishigaki, Katsuhiko Akiyama, Masaki Nakagawa

* Fujitsu Laboratories, Ltd.

64, Nishiwaki Ohokubo-cho, Akashi-shi, Hyogo 674-8555, Japan

** Department of Computer Science, Tokyo Univ. of Agriculture and Technology

2-24-16, Naka-cho, Koganei, Tokyo 184-, Japan

ン認識の候補を1位から順に交互に並べる(図2)。方法2では、オフライン認識のスコアをオンライン認識と比較できるよう正規化し、双方の候補を混ぜた(図3)。

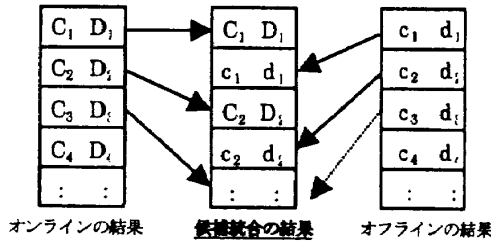


図2. 候補統合方法1

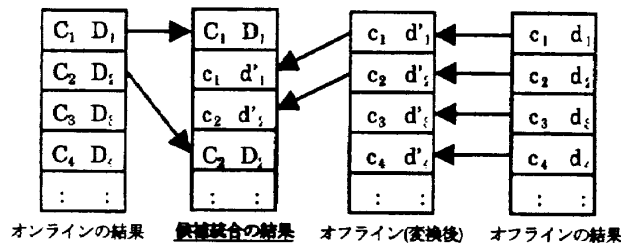


図3. 候補統合方法2

オフライン認識の認識スコアを d_i とし、上位5位までの平均値を m とすると、正規化後のスコア d'_i は以下の式で計算する。

$$m = (\sum d_i) / 5 \quad (i=1 \dots 5)$$

$$d'_i = a \times (d_i - m) + b$$

式1. スコア変換式

式1のパラメータ a 、 b は用いる認識エンジンによって異なるが、今回は概ね安定した結果が得られ得た値($a=0.7, b=680$)を用いた。

4. 認識実験

認識候補のハイブリッド統合の効果を調べるため、認識実験を行った(表1)。認識対象はJIS非漢字と第一水準漢字であり、東京農工大の中川研究室が中心となって収集したオンライン手書き文字DB[4]から、mdb0051、mdb0055、mdb0066の3つのデータセットを用いた(内訳: 非漢字18957文字、漢字16929文字)。

方法1と方法2では、全体に方法2が優れている。10位認識率で優位性が逆転しているが、1位認識率に1%以上の差があるのは魅力である。漢字のみの認識性能は特に高い値を示しており、本方式の有効性を示すものと考えられる。

表1. 認識実験結果

・全字種 [%]					
	1位	2位	3位	5位	10位
オンライン	65.4	78.7	83.9	88.7	92.9
オフライン	83.5	90.7	93.0	94.4	95.4
方式1	83.5	91.4	94.5	96.2	97.4
方式2	85.1	92.4	94.8	96.2	97.2

・漢字のみ [%]					
	1位	2位	3位	5位	10位
オンライン	78.6	87.3	90.7	93.4	96.0
オフライン	89.5	93.2	94.6	95.4	96.3
方式1	89.6	96.5	97.7	98.4	98.9
方式2	92.5	96.2	97.5	98.3	98.9

5. 成果

ハイブリッド化の結果、有意な認識性能の向上が認められた。本方式は比較的丁寧に書かれた文字なら筆順・画数・筆記方向とも完全に自由であり(オフライン認識)、かつ自然な筆順であればかなり崩れた字形でも認識できる(オンライン認識)。また、累積認識率の頭打ちが少ないことから、文脈処理の導入で更に高い認識性能が期待できる。

現在、オンライン辞書の作成はオフライン辞書に比べて多大な労力が必要である(文字のストローク構造を人手で記述するため)。本方式によれば、利用者がほとんど書く機会のないJIS第二水準の複雑な文字(「龜」など)はオフライン辞書に登録されていれば良いという利点もある。

6. おわりに

オンライン認識とオフライン認識の結果をハイブリッド統合する事により、それぞれの弱点を克服した精度の高い認識エンジンが実現できた。

参考文献

- [1] 赤松 茂、回陽 博史:「文字認識処理方式」、日本特許公開 昭55-140970
- [2] 秋山 勝彦、中川 正樹:「ストロークのつながりに寛容なオンライン手書き文字認識」、画像の認識・理解シンポジウム(MIRU'94)I, 145-152(1994.7)
- [3] 鎌田 洋、藤本 克仁、黒川 浩司、秋本 春雄:「圧縮文字特徴と原特徴を併用した高速かつ高精度な文字認識方式」、信学技報、PRU96-205(1997-03)
- [4] TUAT Nakagawa Lab. HANDS-kuchibue-97-01