

仙田修司<sup>†</sup>山田敬嗣<sup>‡</sup>NEC C&Cメディア研究所<sup>§</sup>

e-mail: shusie@ccm.cl.nec.co.jp

## 1 はじめに

漢字などカテゴリ数の多い文字を高速に認識するために、大分類で文字カテゴリを絞り込み、絞り込まれた文字カテゴリを対象として詳細分類を行うという2段階の認識が行われている。本報告では、大分類の精度向上を目的として、従来は詳細分類後のみ適用していた文字 bigram による言語処理 [1] を、大分類と詳細分類の間にも行う手法を提案する (図 1 参照)。

## 2 大分類に関する考察

大分類に求められる要件として、1. 高速であること、2. 出力する候補数が少ないこと、3. 高精度に正解を含んでいること、4. 実現に要する資源の増加が少ないこと、が挙げられる。従来からよく用いられている、詳細分類での特徴の次元数を圧縮した特徴によって大分類する手法は 1.4. の要件を満たすものである。しかしながら、3. の精度向上のためには従来は 2. を犠牲にして候補数を増やすしかなく、そのため詳細分類に要する時間が増加し大分類による高速化の効果が低下していた。

本提案では、大分類が出力した候補文字カテゴリを大分類用の言語処理により言語として適した候補に絞り込むことで、2. を犠牲にすることなく 3. を達成する。また、従来の言語処理と言語辞書を共用することで 4. の要件を満たし、大分類や詳細分類に比べれば言語処理は十分に高速であることから 1. も達成される。

従来の言語処理 [1] は、文字認識を全て終えてから文字列に対する評価値が最大となる候補を選択していたので、文字列全体を考慮した認識結果となっていた。それに対して、提案する大分類用言語処理は文字認識の途中で実行されるために、全体を考慮した評価を行うのが困難である。全ての文字について大分類だけを先に行うこ

\*Rough Classification with Character Bigram for Japanese Character Recognition

<sup>†</sup>Shuji SENDA

<sup>‡</sup>Keiji YAMADA

<sup>§</sup>C&C Media Research Laboratories, NEC Corporation

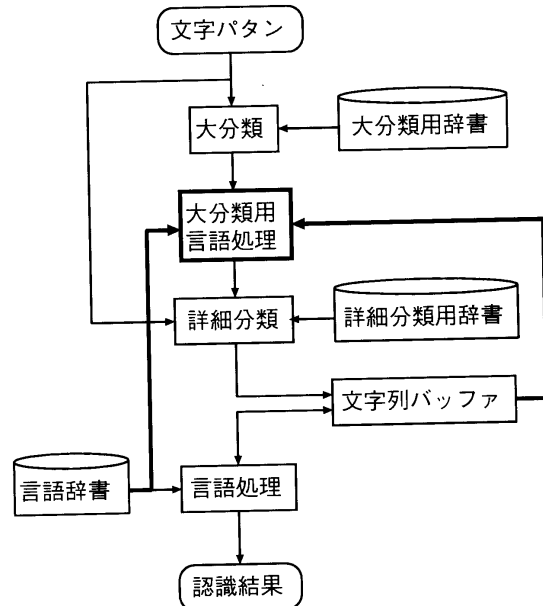


図 1: 提案手法の構成

太線は本提案による変更点を示している

とにより文字列全体を考慮することも可能ではあるが、その場合には個々の文字の特徴量といった、大分類から詳細分類に渡される中間情報を全て保持しなくてはならず、4. の要件を満たさない。そこで、既に文字認識が終わっている結果のみを利用することとした。つまり、ある文字の言語的なつながりやすさを評価するのに、その前後ではなくそれより前にある文字列のみを用いる。

## 3 手法

以下、具体的な手法について述べる。ここでは、文字切り出しが確定しており、言語知識として文字 bigram を用いるとする。文字列中の  $i$  文字目の文字パターン  $X_i$  に対して、大分類によって  $N$  個の候補文字カテゴリ  $c_j^i$ ,  $j = 1, \dots, N$  が出力されたとする。大分類用言語処理において、候補文字カテゴリ  $c_j^i$  の評価値  $\Lambda_j^i$  は、

$$\Lambda_j^i = P(c_j^i | X_{i-1}, X_i)$$

と定義できる。これは、直前位置の文字パターンが  $X_{i-1}$ 、注目位置の文字パターンが  $X_i$  であるという条件のもとで  $c_j^i$  が正解である確率を表す。文字パターン間の独立性を仮定すると、上記の評価値は

$$\begin{aligned}\Lambda_j^i &= P(c_j^i|X_i)P(c_j^i|X_{i-1}) \\ &= P(c_j^i|X_i) \sum_k P(c_j^i|c_k^{i-1})P(c_k^{i-1}|X^{i-1})\end{aligned}$$

と変形できる。候補文字カテゴリ  $c_j^i$  に対する文字パターン  $X_i$  の大分類での距離値  $d_r(c_j^i, X_i)$  と、詳細分類での距離値  $d_f(c_j^i, X_i)$  とから、以下のように確率を近似する。

$$\begin{aligned}\log P(c_j^i|X_i) &\simeq -a_r \cdot d_r(c_j^i, X_i) + b_r \\ &\simeq -a_f \cdot d_f(c_j^i, X_i) + b_f\end{aligned}$$

ただし、 $a_r, b_r, a_f, b_f$  は実験により求めた定数である。さらに、確率の和計算を最大値で置き換え、評価値の対数をとることにより、最終的に以下の評価式を得る。

$$\begin{aligned}\log \Lambda_j^i &= -a_r \cdot d_r(c_j^i, X_i) + b_r \\ &\quad + \max_{k=1}^L \{P(c_j^i|c_k^{i-1}) - a_f \cdot d_f(c_k^{i-1}, X_{i-1}) + b_f\}\end{aligned}$$

前半の項は従来の大分類による評価値、後半の項は直前の文字を考慮した評価値(高速化のため詳細分類の結果が良いものから上位  $L$  個のみを対象とする)となっている。以下の実験では  $L = 10$  とした。

## 4 実験

オンライン手書き文字パターンデータベース HANDS-kuchibue.d-97-06-10[2]の中から、大分類の精度が低めの mdb0051 の文章部分(先頭 10154 文字)を対象として認識実験を行った。文字認識には、オンラインデータを画像として認識する OCR 方式の手法を用い、大分類は 64 次元特徴による市街地距離、詳細分類は 256 次元特徴によるずらしマッチング有りの市街地距離で行った[3]。

実験では、大分類が出力する候補文字カテゴリ数を  $N$ 、大分類用言語処理が出力する候補文字カテゴリ数を  $M$  とし、従来手法については  $N(= M) = 10, 15, \dots, 60$ 、提案手法については  $M = 10, 15, \dots, 30$ 、 $N = M + 10, M + 20, \dots, M + 90$  と変化させた時の(詳細分類の入力となる)累積認識率と(詳細分類後の言語処理を含めた)最終的な文字認識率を求めた。ワークステーション EWS4800/460 (R10000 200MHz) による処理時間(ファイル I/O を除く)と認識率との関係を図 2、図 3 に示す。

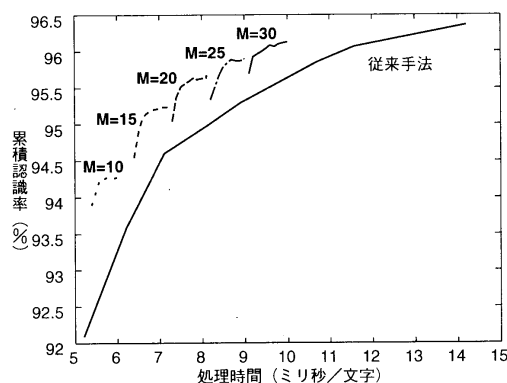


図 2: 累積認識率

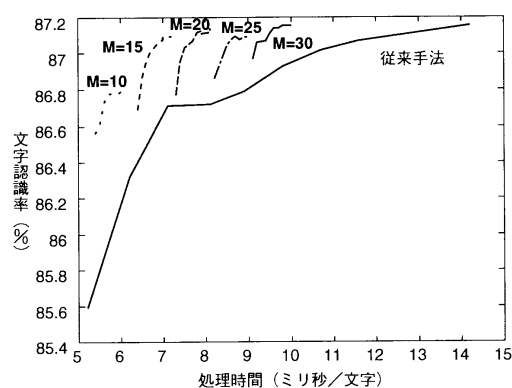


図 3: 文字認識率

図より、累積認識率、文字認識率ともに向上しており、提案手法は効果があることが分かる。特に文字認識率に関しては、 $M = 15$  程度の少ない候補数で限界に近い認識率を達成しており、処理時間を半減できている。

## 5 おわりに

文字認識における大分類と詳細分類の間で文字 bigram による言語処理を行う手法を提案し、その効果を実験により示した。今後は、文字 bigram だけでなく単語辞書を併用する手法についても検討していきたい。

## 参考文献

- [1] 仙田、濱中、山田: “切り出し・認識・言語の確信度を統合した枠なしオンライン文字列認識手法,” PRMU98-138, 1998.12.
- [2] 中川ほか: “文章形式字体制限なしオンライン手書き文字パターンの収集と利用,” PRU95-110, 1995.9.
- [3] 仙田、濱中、山田: “枠なし文字認識と手書き検索が可能なシステム手帳ソフトウェア～メモパッド～,” PRMU99-75, 1999.9.