

要 旨

大分類における距離を用いた 候補数の削減法

藤田 直幸

文字認識過程で処理時間の短縮が重要な課題となっている。本論文では大分類の候補数削減の方法を提案する。処理手順を次に述べる。まず、大分類結果の1位候補を参照する。次に、予め作成しておいたテーブルから1位候補に対応する距離比 r_f を読み出す。そして、1位候補の距離 d_f に距離比 r_f を乗じることにより距離 $d_f r_f$ を算出する。距離 $d_f r_f$ 以内に含まれる候補を大分類結果として出力する。また本手法は、カテゴリ毎に距離比を調整できるため、更なる候補数削減と分類率の向上が期待できる。漢字を含む3,036字種に適用した結果、従来の大分類法の候補数64個を73.3%削減することができた。また、分類率の低下も0.11%と軽微であった。これより、本手法の有効性が確認された。

キーワード 文字認識, 大分類, 候補数削減, 距離比

Abstract

Reduction methods of the number of candidates using distance in preliminary classification

Naoyuki Fujita

Shortening the processing time becomes an important problem in character recognition. This paper proposes the method of reducing the number of preliminary classification candidates. First of all, refer to 1st candidate of preliminary classification result. Next, ratio of distance r_f corresponding to 1st candidate is read from the predetermined table. Then, $d_f r_f$ is calculated by multiplying ratio of distance r_f by 1st candidate's distance d_f . The candidates included distance $d_f r_f$ are output as preliminary classification results. Since the proposed method can adjust the ratio of distance r_f for each category, a further reduction of the number of candidates and the improvement of classification rate can be expected. In the experiment using 3,036 different Kanji characters, 64 pieces, which is the number of candidates in traditional preliminary classification method, were able to be reduced by 73.3 %. The decrease of classification rate is 0.11 %, and is negligible. These results draw the effectiveness of the proposed method.

key words Character recognition, Rough classificatio, Reduction of number of candidates, Ratio of distance