

手書き英単語認識における文字傾き補正法の比較検討

渡辺 雄一郎* 平 英二* 内田 誠一** 迫江 博昭**

*九州大学大学院 システム情報科学府 **九州大学大学院 システム情報科学研究院

1 はじめに

手書き英単語認識において、文字傾きの補正処理は認識処理の精度向上を図る上で必須の技術である。従来は一律な傾き補正手法が用いられてきた。すなわち、文字列を構成する文字はすべて同一の角度で傾いていると仮定されていた。しかし、実際には、各文字の傾きは独立であると考えられ、加えて単文字内にも局所的な傾き変動が存在することもある。したがって、非一律な傾き補正処理が必要と考えられる。

本研究では、前処理により傾きを補正した手書き英単語画像についてマッチングにより認識を行ない、認識における非一律な傾き補正の有効性を示す。

2 非一律な傾き補正およびマッチングによる認識

本稿では、単語を構成する文字列全体を一つの文字画像のようにとみなして認識処理を行なう、いわゆる Holistic な認識の枠組み [1] で非一律な傾き補正の効果を確認する。Holistic な方法は、文字単位にセグメンテーションを行なう方法に比べ一般的に認識精度の点で劣る。しかし、セグメンテーションの影響を考えなくてよく、問題を単純化できるので、本稿ではこれを使用した。

本稿での認識の流れは以下の様である。まず前処理として、平らの手法 [2] を用いて入力パターンを非一律に傾きを補正する。具体的には、非一律な傾き補正を実現するために局所傾きの最適推定を行う。この局所傾きの最適推定は、画像中の長いストロークの傾きはそのストロークを含む文字の傾きを強く反映しているという考えと、できるだけ局所平均傾きに近づけようという考えに基づいて行なわれる。これにより、図 1 のように画像 A 中の各列の傾きの系列が求まり、これを画像 B のようにマッピングすることで非一律な傾き補正処理が行なわれる。

次に、その補正画像上で各列の特徴抽出を行なう。第 1 の特徴は、最上部・最下部輪郭線方向特徴である。第 2 は、各列における輪郭線交差数である。第 3 は、局所平均傾きである。これらの特徴は輪郭線画像において求めるものとする。この結果、入力パターンは 3 次元ベクトルの 1 次元系列として表現される。

最後に、標準パターンとのマッチングを行ないパターン間の距離を求める。具体的には、入力パターンと標準パターンの特徴ベクトル系列間に DP マッチングを施し、両者の距離を求める。この距離に基づく最短距離識別を行ない認識結果を得る。

3 実験

本実験では、(1) 傾き補正なし、(2) 一律な傾き補正、(3) 非一律な傾き補正、それぞれについて認識実験を行ない、認識に与える前処理の影響を検討する。

実験試料としてデータベース CEDAR CDROM 内の BD データベースの手書き英単語画像 (state names) より、入力パターンに 252 サンプル、標準パターンに 2490 サンプルを用いた。全カテゴリ数は 191 で、スペルの違うものは別カテゴリとした。これらの画像はすべて二値化を行ない、縦サイズを 64 に線形正規化した。

図 2 に補正の結果を示す。文字列の傾きが部分的に異なる

場合、一律な傾き補正結果には依然として傾きが残っている。しかし、非一律な傾き補正では文字の形状を崩すことなく適切に補正できていることが確認できる。このことから、文字列の非一律な傾き補正が一律な傾き補正より有効であることがわかる。

また、表 1 に認識率を示す。このように、認識時においても前処理に非一律な傾き補正を用いたものが最もよい結果を得ることができたことがわかる。なお、図 2 の 3 つの文字列はいずれも傾き補正を非一律にすることによって認識できるようになったものである。

4 まとめ

本稿では、認識する上での非一律な傾き補正の有効性を確認した。認識率としては、補正を行なわない時に比べ 10 % の向上、一律な傾き補正を行なった時に比べ 4 % の向上が得られた。

参考文献

- [1] S. Madhvanath and V. Govindaraju: IEEE Trans. Pattern Anal. Mach. Intell., Vol.23, No.2, pp.149-164, 2001.
- [2] E. Taira, S. Uchida, and H. Sakoe: IEICE Trans. Inf. & Syst., Vol.E87-D, No.5, pp.1247-1253, 2004.

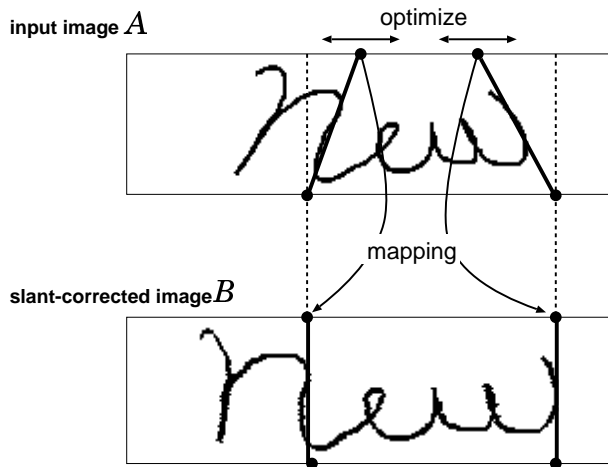


図 1: 非一律な傾き補正処理

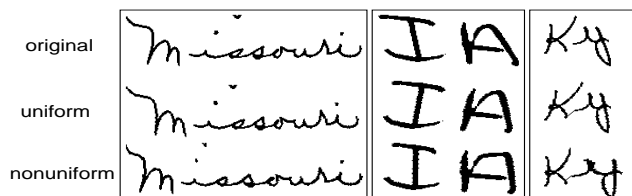


図 2: 傾き補正の例 (“missouri”, “IA”, “ky”)

表 1: 認識率

Slant Correction	No Correction	Uniform	Nonuniform
Recognition Accuracy	34%	40%	44%