

ストローク統計に基づいた柔軟なマッチングによる手書き数字認識

村松洋一, 小林隆, 杉山岳弘, 阿部圭一
静岡大学 大学院 情報学研究科

1 はじめに

現在、オフライン手書き文字認識の分野では、入力パターンから抽出した特徴量に基づいて認識を行う手法が主流である [1]。しかし、人間は文字の特徴ではなく構造を理解し認識していると考えられる。また、人間と同等の認識を単独の手法で担うことは難しいと考えられるため、特徴抽出による手法とは異なる手法を新しく確立することも重要である。

そこで、我々は手書き数字を対象とし、文字の構造と形状に注目した認識手法を提案する。手書き数字の多様な変形に対応するため、文字パターンを柔軟に変形させることでマッチングを行う。このとき、パターンの過剰な変形は、認識の精度に大きく関わる問題である [2]。本研究では、手書き数字の変形を、数字を構成するストローク単位で統計的に学習する。そして、その学習結果に基づき、数字の構造や形状を保持した変形を行う。

2 認識システム概観

本研究における認識システムの概観を図1に示す。本システムは、標準パターンとその変形範囲を統計的に学習する学習部と、学習部で作成した標準パターンを未知の入力パターンと柔軟にマッチングさせて認識を行う認識部とで構成されている。

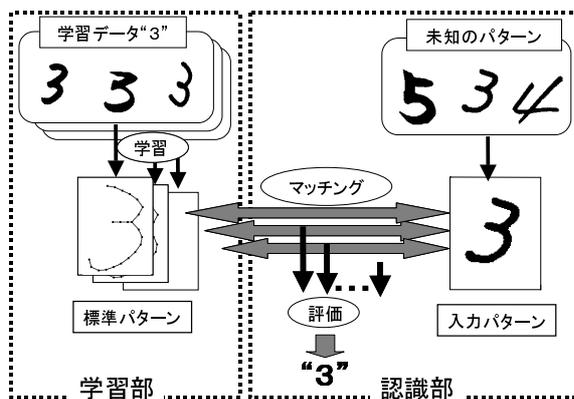


図1: 認識システムの概観

3 標準パターンの作成

標準パターンの学習は各文字クラスごとに行う。端点、分岐点、屈折点、曲率が最大となる点で区切られる線分をストロークとし、ストロークの方向、全ストロークの実効長 (ストロークの端点間の距離) の和に対するストロークの実効長の比、ストロークの実効長に対するストロークの高さ (ストロークの両端点を通る直線から最も離れたストローク上の点までの距離) の比、ストロークの端点間のベクトルと相対角度の統計をとる [3]。

これらのパラメータの最頻値を用いて標準パターンを作成する (図2)。標準パターンは文字幅1の図形であり、複数のストロークから構成される。また、各ストロークは学習したパラメータの統計データを持っている。

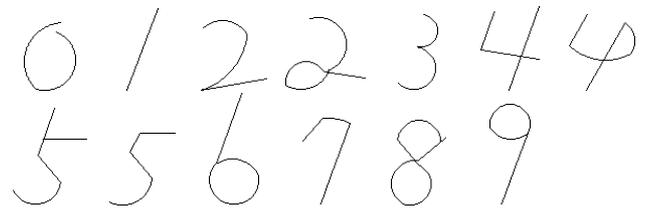


図2: 作成された標準パターン

4 認識方法

認識の流れを図3に示す。まず、学習部で作成された全ての標準パターンを、入力パターンの形へ近づけるように変形させる。標準パターンは、入力パターンの形へ変形する力 \vec{F}_{image} と、元の形状を保つ力 \vec{F}_{int} により、数字の構造を保ちながら柔軟に変形する。変形が終了した段階で、入力パターンの内部に存在している標準パターンを、入力パターン全体へと膨張させることで、標準パターンと入力パターンの形を完全に一致させる。最後に、標準パターンの変形度合と膨張にかかったコストを評価して認識結果を得る。以下、それぞれについて述べる。

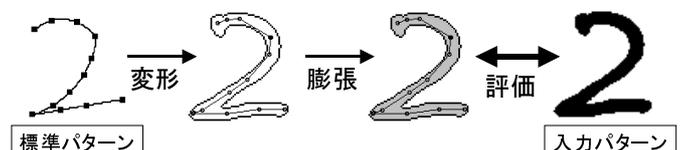


図3: 認識の流れ

4.1 標準パターンの変形

入力からの力 \vec{F}_{image} は、図 4 に示すような、標準パターン上のサンプル点と、入力パターンの輪郭線上の点における、方向と凹凸情報を利用して求める。サンプル点は、自分が持っている方向と似ている方向を持ち、かつ、距離が近い輪郭線部分から強く引きつけられて移動していく。ただし、サンプル点が凹凸属性付きの方向を持っている場合には、同じ凹凸属性を持つ輪郭線のみから引かれる。

標準パターンが元の形状を保とうとする内部的な力 \vec{F}_{int} は、標準パターンを作るさいに学習したストロークの各パラメータの統計量から内部エネルギーを求め、その内部エネルギーを元に力を算出する [4]。



(a) 標準パターン (b) 入力パターン

図 4: 方向と凹凸情報

4.2 膨張処理

変形した標準パターンと入力パターンの一致度と、ストロークの対応関係を調べるために、標準パターンの膨張を行う。標準パターンの各ストロークごとに異なるラベルを付け、そのラベルを距離変換に同期させて伝搬することで、入力パターンの内部から入力パターンと完全に一致するまで膨張させる (図 5)。

さらに、本研究では、膨張に指向性を持たせる。そのために、ストロークの方向として、水平 ()・垂直 (|)・右斜め (/)・左斜め (\) の 4 種類の属性を定義する。標準パターンのストロークの方向属性は、サンプル点間の線分ごとに与える (図 6(a))。入力パターンの方向属性は、輪郭線方向から 4 種類の方向属性を割り当て、それを輪郭線上から 3-4 距離変換に同期させて文字内部へと伝搬させることで与える (図 6(b))。

入力パターンと標準パターンの方向属性が同じ箇所では膨張しやすく、方向の差が大きく異なるにつれて膨張しにくくする。これにより、入力パターンと標準パターンが適切にマッチングしているほど、膨張が容易になる。この異方性膨張は、ラベル伝搬に同期する距離変換を、3-4、5-7、7-10 距離のいずれを用いるかにより実現する。



図 5: 膨張のイメージ

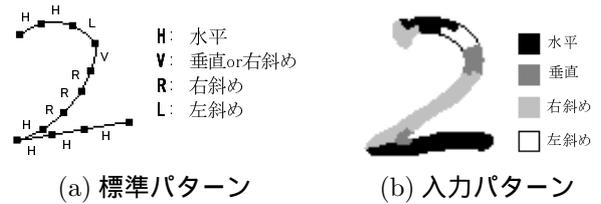


図 6: 膨張に用いる方向属性

4.3 マッチングの評価

マッチングの評価は、標準パターン R について、変形に関するコスト $E(R)$ 、膨張に関するコスト $D(R)$ 、ストロークの対応関係に関するコスト $S(R)$ の 3 つの和により行う。

$E(R)$ は、入力パターンに対する標準パターンの変形の大きさを、変形後の内部エネルギーから求める。 $D(R)$ は、膨張処理において距離変換によって得られた各画素の距離の総和、各ストロークにおける距離の総和の標準偏差、膨張開始時に入力パターンの背景に存在する標準パターンの画素の入力までの距離の総和の 3 つを用いる。 $S(R)$ は、ストロークの対応関係に不適切であった場合のペナルティ値である。対応関係は、膨張後の標準パターンの隣接関係が本来あるべき隣接となっているか、入力パターンと標準パターンのストロークが 1 対 1 の対応になっているか、という 2 点について調べる。

全標準パターンと入力パターンとのマッチングの結果、コストが最も小さかった標準パターンを認識結果とする。

5 実験結果とまとめ

標準パターンとその統計データの学習には手書き郵便番号データベース [5] のうち 5116 個を用いた。学習に用いていないデータから、各数字ごとに 300 個、合計 3000 個を対象に認識実験を行った結果、96.1% の認識率を得た。この認識率はまだ十分に高いとはいえず、さらに認識の精度を高めるための改良が必要である。今後は、現在誤認識しているパターンについて、その原因を究明し、それに対する改良を行う予定である。

参考文献

- [1] R. G. Casey and E. Lecolinet, "A survey of methods and strategies in character recognition", IEEE Trans. PAMI, Vol.18, No.7, pp.690-706, 1996.
- [2] 内田誠一, 迫江博昭, "固有変形の利用による手書き文字認識の高精度化", MIRU2002, vol1, pp.391-396, 2002.
- [3] T.Kobayashi, K.Nakamura, H.Muramatsu, T.Sugiyama, and K.Abe, "Handwritten numeral recognition using flexible matching based on learning of stroke statistics", Proc. 6th ICDAR, pp612-616, 2001.
- [4] 中村香織, 小林隆, 杉山岳弘, 阿部圭一, "輪郭線方向を利用したマッチングを用いた手書き数字認識", 信学技法, PRMU99-220, pp.9-16, 2000.2.
- [5] "IPTP CDROM1 : Database of Handwritten 3-digit Postal Code", 郵政省郵政研究所, 1994.4 .