

弾性膜モデルによるオフライン文字認識

1 P-5

柳田 正 永崎 健 中川 正樹

東京農工大学 工学部 電子情報工学科

1. はじめに

手書き文字認識の高精度化を図るには、筆記者による癖や、書き方の違いからおこるパターン形状の変動を吸収しなければならない。郭[1]は、未知パターンに対し、左寄せ、中心拡張および右寄せの3つの整形変換関数を適用し、得られた整形パターンの中から、最も整合性の良いものを選ぶ手法を提案した。若原[2]は、パターン全体の一様なアフィン変換成分を安定に抽出する大局的アフィン変換と、局所的な変形を近似する局所的アフィン変換を直列的に施す手法を提案し、その有効性を確認した。これらは、手書き文字認識において、手書き特有のパターン変形を矯正し、標準的なパターンに整形することが好ましいという考えに基づいている。我々も同様の考えに基づき、手書き文字の整形を強く指向する「力学的モデル」[3-5]を提案している。力学的モデルは、ポテンシャル場が整形を誘導する役割を担い、弾性体がカテゴリの異なるパターン同士の整合を阻止する役割を持つ。本研究の狙いは、両者の特性がうまく働きあうことで、字種に依存した正しい整形変換を行えることを期待するものである。

力学的モデルは包括的なモデルであり、弾性膜モデルはそのサブクラスとして位置づけられる。本稿では弾性膜モデルについて報告を行う。

2. 弾性膜モデルの概念

弾性膜モデルは、手書きパターンに現れる書き手による癖や位置的な変動の吸収過程を物体の力学的な運動としてモデル化した認識手法である。膜モデルの概念を次に示す。

入力パターンを質点の集合とみなし薄いゴム状の膜

に貼り付ける。それを標準パターンの形状をもとに生成した文字型の谷間(ポテンシャル場)に落とす。質点は膜の制約を受けながらポテンシャルの低い方へ移動する。今回は、この運動収束時の膜から得られるパターン形状を整形パターンとして扱う。

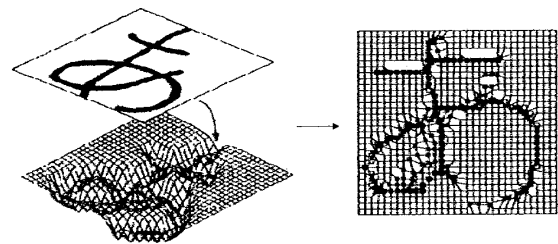


図1 膜モデルの概念

3. 認識システムの基本構成

本稿で提案する手書き文字認識システムは、図2に示すような5つの処理過程によって構成されている。

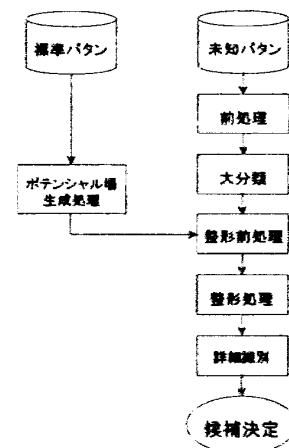


図2 認識システム

3.1 前処理

前処理は、パターン矩形抽出、正規化を行う。正規化では、領域内の走査線数を数えることによって、パターンの標本点数を調整し、線間密度を均一化する非線型正規化[6]を採用した。

3.2 大分類

計算時間を考慮して候補カテゴリを大分類の上位5位までとした。大分類アルゴリズムとして、方向線素

特徴量[7]を用いた認識システムを採用した。

3.3 整形前処理

整合前処理は、弾性体とそれを誘導するポテンシャル場を計算機上で表現する過程である。

膜のような連続的な弾性体の変形を解析するには、膨大な計算量と計算機リソースを必要とする。そのためここでは、バネ要素をメッシュ状に組み合わせたものによって膜を近似的に表現した。

3.4 整形処理

バネメッシュの時刻 t における i 行 j 列目の格子点の位置を、 $P_{ij}(n\Delta t)$ で表すと、格子点に関する運動方程式は、 $m\ddot{P}_{ij}(t) = F_{ij}(t)$ のような微分方程式となる。独立変数 t を $n\Delta t$ と表現し、差分的に反復演算を行い、格子点の位置を割り出す。

3.5 詳細識別

整形パターンを評価するにあたっては、重合法や方向特徴を使った識別法など、整形パターンそのものを評価する方法も考えられるが、本認識システムでは、膜モデルの評価値を含む次の 5 つの評価値を用いて行う。

- (1) 弾性膜内部エネルギー値
- (2) 弾性膜位置エネルギー値
- (3) パタン内包罰則値
- (4) 弾性分布評価値
- (5) 整形パターン方向特徴評価値

紙面分量の都合により各評価値についての詳細はここでは割愛する。

4. 認識実験

認識実験は、通産省電子技術総合研究所の ETL9B を使用した。本実験ではパターン整形の可能性を調べることを目的とし、平仮名 46 字種を認識対象とした。整形の見本となるパターンには、ETL9B のサンプルセットを、未知パターンセットとして 1~50 番までの 50 セット、辞書の学習セットとして 50~200 までの 150 セットを用いた。この条件における各評価値だけによる認識率と、それを標準化した値の和である総合評価値における認識結果を図 3 に示す。カッ

コ内の数字は 3.4 の各評価値に対応する。

評価値	(1)	(2)	(3)	(4)	(5)	総合
認識率 (%)	62.4	54.1	87.8	90.3	96.8	98.3

図 3 認識率

5. おわりに

本稿では、弾性膜モデルを手書き文字の整形に使うことの可能性を提案した。ETL9B のひらがな 50 セットを対象にして認識実験を行い、認識率は 98.3% を記録した。この結果から、弾性膜モデルが手書き文字の正しい整形を行えたかどうかは結論付けられないが、可能性は見出せたものと考えられる。

今後、さらに研究を進めることによってその有効性を立証していきたい。

本研究を行うにあたり、電子技術総合研究所の JIS 第一水準手書き漢字データベース ETL9B を利用させていただいた。また本研究は一部、情報処理振興事業協会による創造的ソフトウェア育成事業の補助による。ここに深く感謝する。

参考文献

- [1] 郭軍, 孫寧, 根元義章, 佐藤利三郎: “整形変換を用いた手書き文字データベース ETL9B の高精度認識”, 信学論(D-II), Vol. J76-D-II, No.5, pp.1015-1022(1993).
- [2] 若原徹, 小高和己: “GAT/LAT を用いた手書き文字の適応型整形”, 信学技法, PRMU96-186, (1997).
- [3] T.Nagasaki, T.Yamamoto, M.Nakagawa: “The Behavior of Dynamic Relaxation in an Elastic Stroke Model for Character Recognition”, Proceedings of Fourth International Conference on Document Analysis and Recognition, Vol. 1, pp.16-22, (1997).
- [4] R.Webster, M.Nakawaga: “An On-line/Off-line Compatible Character Recognition Method Based on a Dynamic Model”, IEICE Transactions on Information and Systems, Vol. E80-D, No.6 (1997).
- [5] 寺村健, 中川正樹: “手書き文字認識を指向した力学的認識モデルの一形態”, MIRU'94, I, 59-66(1994.7).
- [6] 吉田収志, 本郷保夫: “非線型正規化を用いた背景特徴パターンマッチング法による手書き漢字認識”, 信学技法, PRU92-34, (1992).
- [7] 孫寧, 安倍正人, 根元義章: “改良型マハラビスを用いた高精度な手書き文字認識システム”, 信学技報, PRU94-94, (1994).