

近年手書き文字認識技術は多大な発展をとげ、ほぼ実用に近い認識率が報告されている。しかし実際の運用は認識において得手不得手があり手修正なしでは実用的ではない。これは認識システムの構造が各々独立しており、各層での情報が結合されていないからである。本研究では現状の問題点を分析し、また複数の層で情報を交換する文字認識システムを提唱する。

1 はじめに

近年手書き文字認識に関して様々な手法が提案され、90%を超える認識率を持つものや、これを用い実用に供されるシステムも散見される。しかし実際の運用では活字の認識において得手不得手があり、修正なしでの変換は実用に至っていないと言えよう。これは一つには1文字あたりの認識速度に制限があるためと、認識精度がいまだ不十分なためである。

一方単一の手法の認識率に関しては、ほぼ頭打ちの傾向にあり、複数手法の組合わせ等の統計的アプローチが求められている。

本研究ではオフライン文字認識について本質的な解決方法として、楕円を認識単位とし、情報のフィードバックの機構を持つ文字認識を提唱する。

2 一般的な文字認識方式

オフライン文字認識に関しては大きく分けて2つの流れがある。一つはパターンマッチングで、もう一つは構造解析法である。

構造解析法はパターンの形状や構造を複数のプリミティブとそのプリミティブ同士の関係によって表現し、辞書として記述されている同様な表現と比較することによって認識を行うものである。またパターンマッチング法は単にパターンを重ね合わせるものに加え、特徴抽出した成分を統計的手法により分類するものも含まれる。

この両者の違いは前者がプリミティブ間の関係が密接であるのに比較して、後者がベクトル形式など要素間の関係が希薄である点にある。またこの点が両者の認識する文字の特徴となる。すなわちパターンマッチング法では基準となる文字フォントから変形が著しいものの認識が難しいが、構造解析法では位相が同じであれば認識可能である。

3 認識の流れ

以上の方式はFig. 1aのような認識の流れを前提としている。従ってある時点での切出しの失敗がそのあとに続く文字に影響する。このため文字切出しで失敗すれば認識率はきわめて低くなる。切出しは特徴量抽

出の前に行われる正規化の作業にも影響を与える。たとえば、非線形正規化においては、余計な情報を含めて正規化されるので全体としてはかなりの変形を余儀なくされる。

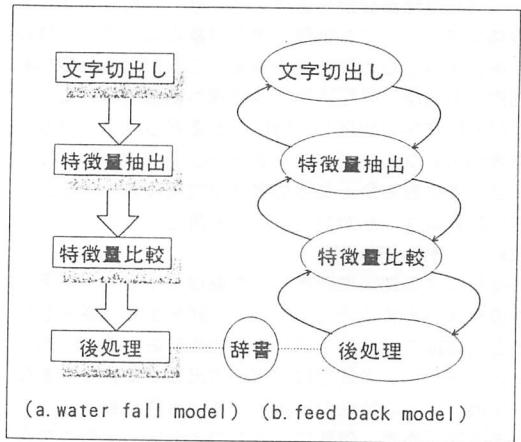


Fig. 1 Recognizing Process.

これらの結果より、従来の認識に関しては高精度の文字の切出しが前提条件となっているといえる。しかし単なるビットマップパターンからの情報のみでこれらを処理することは難しい。例えば「小青」と「情」とは位置的關係が異なるだけである。ここでは「小青」という言葉は一般に使用されないという知識が必要となる。

そこで、fig. 1bのようなフィードバックを持つモデルを考えてみる。ここでは1回目の認識は通常どおり行われるが、2回目以降は1回目によって得られた情報を元に次に接続すべき文字候補を持ってこることが可能である。これより認識で比較の候補となる文字の集合を縮小させることができる。この結果、認識速度、認識精度ともに大幅に向上させることができる。

4 後処理

認識により求まる文字候補と、使用する言語の辞書を使って次に認識すべき文字を正すことができる。

これは後処理と呼ばれ、すべての認識が終わった後に候補文字と突き合わせて行うものである。これを文字切出しや認識のフェーズにフィードバックすると、候補文字の大幅な削減が実現可能である。すなわち次に認識される可能性のある文字のみをテストするのである。特に日本語の場合には、文字の接続において極めて直交性が高いといわれている。本システムでは単語の有無だけが分る辞書を用意し、これを用いて接続関係をチェックした。しかし形態素分析等を用いてさらに候補文字列を絞り込むことも可能である。

5 認識

点素における認識は部分領域の類似度を基本としている。線素における認識では、線の情報を平均化して照合したり、弛緩法を用いるものもある。本システムでは楕円を基本単位として認識を進める。このフェーズには後処理部分から候補文字が与えられる。従って候補文字についての情報と現在対象となっている情報を照合することで類似度を得る。これがしきい値の範囲内であれば、処理結果を後処理へ転送する。そうでなければ文字の切出しに原因があるだろうと推測して文字切出しフェーズへバックトラックする。

ここで検査対象となる文字数が大幅に減少するので多少計算コストがかかる処理も実用となる。

5 楕円素

従来の文字認識の元となる画像はビットマップそのものを入力とする方式（点素）と線または輪郭を抽出して認識する方式（線素）とに二分される。しかしフィードバック方式では文字の切出しが柔軟にできないため、従来の認識の発展としては実現し難い。

そこで、点素、線素に変るものとして、点素の集合を楕円を使用してクラスタリングした楕円素を元として認識することとする。楕円素は、中心点、長軸、短軸からなり、この内部に分散する点群を代表するものである。すなわち認識は点群から楕円素へクラスタリングされた結果を元に行われる。

クラスタリングは、現状ではGAを用いて行われて

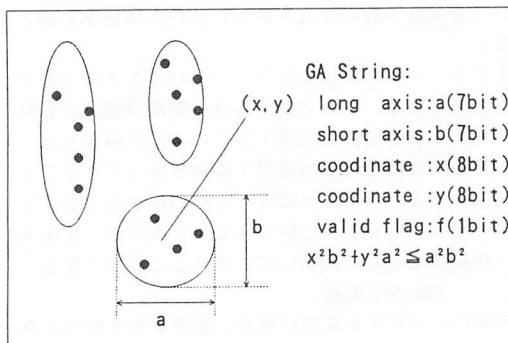


Fig. 2 Clustering with GA.

いる。GAを用いた場合精度的には満足できるが速度的には問題がある。しかしファジィクラスタリングやニューラルネットを使用すれば高速にクラスタリング可能であろう。点素から楕円へのクラスタリングについてfig. 2に示す。

GAは局所山乗り付きのものを用い、30秒程度でほぼ収束した。評価関数としては「楕円内に含まれない点数+楕円の面積」を重み付けして用いた。楕円の個数については各ストリングに有効であるかどうかの1ビットのフラグを設定して可変長とした。

6 認識結果

実験では従来方式によるもの（方式1）と、フィードバックを用いたもの（方式2）とを比較した。入力装置はマウスであり、入力サイズは画面64×64の長方形領域である。入力文字種は数字、ひらがな、アルファベット漢字である。また線分の情報は使用していない。

	第1位までの認識率	第2位までの認識率
方式1	72%	76%
方式2	90%	93%

全般に認識精度が低いのは、比較の対象となるフォントの解像度が低いこと、類似度の手法、そして入力装置がマウスであることなどがあげられる。

しかし人間が全くランダムな文字綴りを読む場合、認識率が低下することからも、文字認識に知識が必要なことは明らかである。また認識率、認識速度は辞書サイズに反比例して上がるので、特殊な文字、例えばある規則に基づいた記号列等は速度、精度的に極めて有効であろう。

7 あとがき

いままで別個に論じられてきた認識、後処理等について統合的な情報交換を持ったアプローチが必要であることを提唱した。また実験によりこのアプローチが有効であることを確認した。

文字認識に限らず、部分の認識精度が人間と同等になっている分野もあり同様なアプローチが重要になると思われる。

参考文献

- 1) 飯島 泰蔵 : パターン認識
 コロナ社、(1973)
- 2) 入江 文平 他 : 2×2近傍特徴による手書き漢字認識
 信学会全国大会、D-453、pp. 1-196(1988)
- 3) 坂井・森(東芝) : 漢字パターンの大分類-漢字パターンの性質
 PRL73-14、pp. 1-16(1973)
- 4) 岡 (電総研) : セル特徴を用いた漢字認識の試み
 PRL79-52、pp. 47-56(1979)