

漢字の階層性に注目した文字認識手法

2H-7

白石知之 田中英彦
東京大学工学部

1 はじめに

文字認識の研究は古くから盛んに行われてきた分野である。特にわが国では漢字を対象とした研究を中心に行われてきた。

漢字認識の分野では階層構造認識方式が提案され、また手書き文字に対する方向特徴の有用性が示されてからは、重ね合わせ法による認識手法が主に用いられるようになり、その結果現在では認識率の非常に高いシステムが実際に開発されている。

しかしそのため文字認識の研究は既に完了したものと思われがちであるが、現実には未解決の問題が数多く残されている。

その一つとして、現在の重ね合わせ法による認識手法では、常に知っている文字のどれか一番近いものに分類してしまうため、文字を正しく理解しているのではなく、当てているだけであるとの指摘がなされている。

文字を理解したと言えるためには、入力図形とその文字の系の内部での表現とを比較することが必要であり、そのためにも構造解析的な手法を用いることが不可欠であると考えられる。

本稿では漢字認識に構造解析的な手法を用いることを容易にするために、漢字の階層構造を利用することについて論じる。

2 漢字の階層構造

我々が漢字を分解するとき、一度に一画一画のストロークに分解するのではなく、ある程度まとまったブロックに分解する。それは例えば偏と旁であったりというように、その文字を構成する基本的な部品である。

「認」という文字を例にとっても考えると、まずこの文字は「言」と「忍」に分解することができる。そして

更に「忍」の部分は、「刃」と「心」に分解することが可能である。

こうして分解されていったブロックは、最終的には一つ一つがその組合せによって漢字を構成する基本字根となる。

これら基本字根間の関係は、次の四種類に分類することができる。

単独関係 一つの基本字根だけで成り立つ

離散関係 基本字根間が一定の距離を保っているもの

接続関係 基本字根と筆画の組合せで構成される

交叉関係 基本字根が重なり、交叉して構成される

このうち離散関係に注目すると、これは更に次の四種類に分類される。

左右型 要素と要素が左右に並んでいるもの

上下型 要素と要素が上下にならんでいるもの

外内型 要素の内部に要素があるもの

単体型 それ自身で構成されるもの

単独関係、接続関係、交叉関係によって成り立つ部位は、全て単体型であるとみなせば、全ての文字を離散関係の組合せで表現することが可能になる。

3 漢字認識への利用

漢字認識において、英数字の認識で用いられているような構造解析的な手法があまり用いられない理由の一つとして、漢字の構造が英数字のそれと比較して複雑であることが挙げられる。

そこで漢字をその構成要素に分解してやれば、それぞれの要素は文字全体と比較すると簡単になり、構造解析的な手法を用いやすくなるのではないかと考えられる。

ところがこれをボトムアップに行ったのでは、文字切り出し問題と同様の困難が生じるため、予め辞書に登録しておいた漢字の構造情報をトップダウンに与え

てやることにより、入力図形を図1に示すように構成要素単位の領域に分割する。

こうして分割された領域毎に弛緩法[1]などの構造解析を用い、それぞれの結果を全体として総合して判断するようなシステムを構築することで、漢字に対する構造解析を容易にすることが期待される。

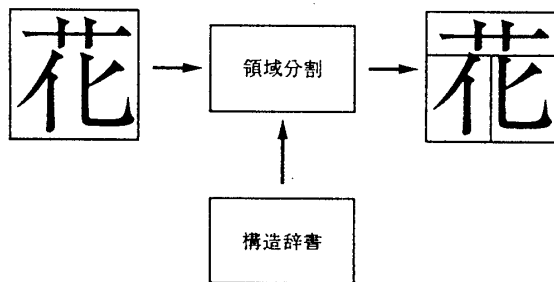


図1: 領域の分割

このシステムではトップダウンの予想として文字構造を与えてやる必要があるため、候補文字をあらかじめ限定しておかねばならない。

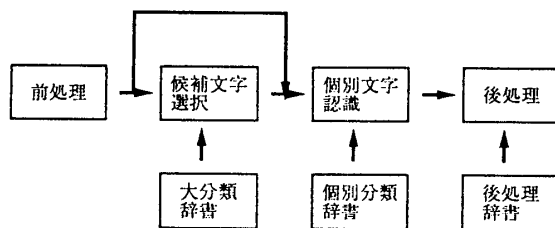


図2: 階層構造認識方式

現在の漢字認識システムでは、図2に示されるような階層構造認識方式が主に用いられている。そこで、この方式で個別文字認識と後処理の間にあたる部分、もしくは後処理に代わるものとして使用することによって、ある程度候補が絞られた状態にすることが可能になる。

実際に領域に分割する方法であるが、手書き文字の場合、印刷文字と異なり予想外の部分での接触や切断があるため、トポロジーを利用することは非常に困難である。

そこで統計情報を用いて分割することにする。トップダウンに構造情報が与えられているため、あらかじめどのように切断して分割すれば良いかは分かっている状態になる。

構成要素は更に細かい構成要素の組合せで成り立っているというように、入れ子構造になっていることを

利用して、最も粗い分割から順に行っていく。これによって同時に一種類の分割法しか用いられず、一度で全要素への分割を行うのに比較して、領域分割の切断線の決定が容易になる。

要素間の関係から切断線をどのように引くかを決定した後、切断する方向に対して垂直な方向成分を求め、それを横切る場合にペナルティを課す。その際切断される線分の中点に近い位置ほどペナルティが大きくなるように重みをかけておく。

こうして得られたペナルティの分布から、トップダウンに与えられた構造情報を元に切断する位置を判断し、領域の分割を行う。

分割された領域を更に細かく分割する場合は、それ以外の部分にマスクをかけてから、同じ処理を繰り返して用いてやることによって実現される。

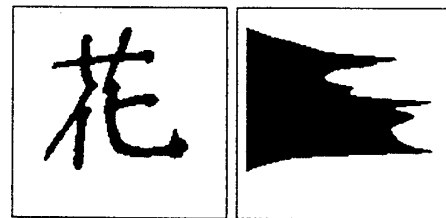


図3: 「花」の分解

例として図3に、「花」という文字を、くさかんむりと「化」という部分に分割する様子を示す。両者が接触しているためトポロジーを用いて分割することはできないが、この方法ではうまく切断する地点が求められていることがわかる。

4 おわりに

漢字の階層構造を利用して、入力図形に対してトップダウンに構造情報を与えることにより構成要素毎の領域に分解し、漢字の構造解析を行う補助手段を提案した。

今後この方法を利用して、漢字に対する構造解析が容易に行われることが期待される。

参考文献

- [1] 山本和彦. 弛緩整合法による手書き教育漢字認識. 電子通信学会論文誌, Vol. J65-D, No. 9, pp. 1167-1174, Sep. 1982.