階層型構造解析によるタイ語文字認識

1M－3

中村康弘・チャリアプロマイン・松井甲子雄
防衛大学校情報工学科

１．はじめに

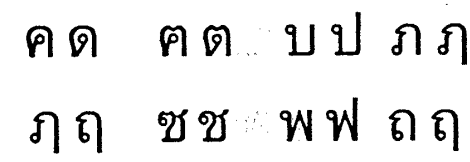
今日，日本語や英語の文字認識についてはすでに多くの報告があり，近年では手書き文字認識の高精度化に関する研究へと発展してきている[1]．しかしながら，その他の言語の文字認識に関する報告は少ない．この報告ではタイ語の文字の構造に着目し，従来の日本語や英語に対する手法の適用可能性について検討した．この結果，一部の文字ではその切り出しや認識において独特の考慮が必要となることが明らかとなった．この結果に基づいてタイ語の文字認識システムを構成・評価したので報告する．

２．タイ語の文字

タイ語の文字は，現在ほとんど使用されなくなった文字を除き，表１に示す子音，母音，声調記号および数字やその他の記号等から構成される．子音文字をベースライン上に書き，母音文字は付帯する子音文字の上下左右のいずれかに書く．声調記号はそれが作用する文字の直上あるいは右肩に書く．表１中の－はその母音文字が付帯する子音文字を表し，その高さをH，平均的な幅をWとする．

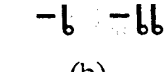
表1 タイ語の文字

子音文字	ก ข ฃ ค ฅ ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท น บ ป ผ ฝ พ ฟ ภ ม ย ร ฤ ล ฦ ว ศ ษ ส ห ฬ อ ฮ
母音文字	-ะ -ั -า -ำ -ิ -ี -ึ -ื -ุ -ู เ- แ- โ- ใ- ไ- -ๅ -็
声調記号	่ ้ ๊ ๋ ์
数字および特殊記号	๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙ -ๆ -ฯ -๎



(a)　(b)　(c)

図１ タイ語の類似文字



(a)　(b)

図２ 類似した母音

３．認識上の問題点

(1) 文字の切り出し

タイ語の文字は日本語よりも英語に近い，ベースラインに基づく表記をするが，声調記号は日本語の濁点・半濁点のように真上あるいは右上に付ける．また，母音は付帯する子音の上下左右のいずれかに書く．母音と声調記号が重なったときは，まず母音を書き，その上に声調記号を書く．このように，子音については英語とほぼ同様の手順で切り出し可能であるが，母音と声調記号については子音との相対位置関係を考慮して別に処理する必要がある．

(2) 文字の構成要素と字形

タイ語の文字はほぼ一筆書きなので，文字の構成要素間の位置関係などの構造解析的な手法は難しい．たとえば，図１(a)は小円から継続して筆記するときの方向の違いで識別される．(b)は文字上部中心の窪みによって識別される．(c)は右端の垂直線が文字の標準高Hより長いか，ベースラインより下までのびているかどうかで識別される．日本語でも士と土のように(c)に該当する例はあるが，(a),(b)のような例はない．

(3) 前後関係依存性

日本語や英語の文字は単文字識別が可能な構造となっているが，タイ語には図２に示す２組の母音が

Recognition of Thai Character by Hierarchical Structure
Yasuhiro Nakamura, Charia Promin and Kineo Matsui
The National Defense Academy
1-10-20 Hashirimizu, Yokosuka, Kanagawa 239, JAPAN

あり，文字を構成する線素の形状が同一である．したがって，連結領域に基づく単文字識別のみでは不十分であり，前後関係に依存した識別の後処理が必要となる．

4．文字認識処理

上記の問題点を解決するため，タイ語文字の識別処理手順を以下のように構成した．

(1) 行の切り出し処理

行を切り出すには，画素ヒストグラムを用いる方法・仮想カーソルを移動させる方法・小領域の統計的性質を用いる方法などがある．ここではヒストグラムによる方法を採用した．また，ここでは前述のHの範囲を行と定義し上下の母音は別に処理する．

(2) 文字の切り出し処理

タイ語のほとんどの文字は単一の連結成分と考えられるので，連結成分ごとに切り出せば容易に各文字に分離することができる．しかしながら，２つの文字が近接により連結してしまう場合があり，とくに子音文字と直上の母音文字が連結する場合が多い．そのような場合にはHあるいはWの大きさで分割し直して再度認識処理を行う必要がある．

(3) 文字の位置と大きさによる類別

文字切り出し処理の結果から得られる文字の大きさと位置情報をもとに，各文字を図３に示す9個のクラスに分割した．ここで，クラス６と７は図２(a)の母音のための一時クラスである．また，クラス８と９は子音の左右に付く母音用であるが，この時点ではすべてクラス３として処理し，後で分類する．

(4) クラス内文字の識別

文字の識別は，連結した画素を方向性を持つ線素として抽出し，その線素の集合としての円や直線等の各文字を構成する部分集合を得る．得られた部分集合を切り出された文字枠内の位置と大きさによる特徴量空間に射影し，あらかじめ用意した各文字との距離により識別した[2,3].

5．実験

イメージスキャナから8ビットモノクロ画像として原稿を読み取り，テストデータとした．読み取り解像度は300[DPI]である．原画像の一部を図４に示す．この画像から行を識別し，各文字を切り出した結果を図５に示す．さらに各行の位置と各文字の位置と大きさの関係から各文字の属するクラスを決定し，各クラスごとに識別処理を実行した．実験の結果，切り出し処理時に文字が連結してしまった場合に誤って識別してしまう場合があったが，他はおおむね良好な結果が得られた．

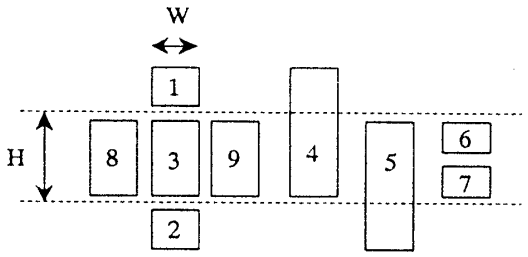


図３ 文字の位置と大きさによるクラス

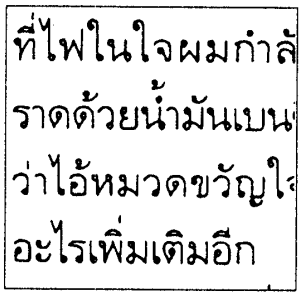


図４ 入力原画像

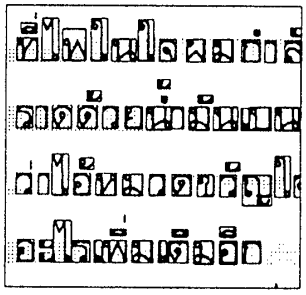


図５ 行と文字の切り出し

6．むすび

本報告では，タイ語の文字の特徴に基づく文字認識処理について検討し，認識処理アルゴリズムを提案した．実験の結果，文字の切り出し処理が正確に行えれば，正確な認識が可能となることが明らかとなった．さらに，連結した文字の再分割および手書き文字の場合について検討を続けたい．

参考文献

(1) たとえば，情処全大論文集, 1L (1993) など.

(2) 横塚，木田: 再帰型変数増減法による..., 信学論DII, J76, 9, pp.1994-2003 (1993).

(3) 孫，田原，阿曽，木村:方向線素特徴量を用いた..., 信学論DII, J74, 3, pp.330-339 (1991).