

Title	セグメント抽出方式を用いた手書き漢字認識に関する研究
Author(s)	馬場口, 登
Citation	大阪大学, 1984, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/2581
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

セグメント抽出方式を用いた 手書き漢字認識に関する研究

昭和59年5月

馬場口 登

序 文

本論文は、筆者が大阪大学大学院工学研究科（博士後期課程通信工学専攻）並びに愛媛大学工学部電子工学教室において行った、セグメント抽出方式を用いた手書き漢字認識に関する研究を5章に分けてまとめたものである。

第1章は、緒論であり、本研究の目的と本研究に関連した諸研究の現状を概説し、本研究の位置付け及び工学上の意義について述べる。

第2章では、手書き漢字認識系における有効な特徴抽出の一つであると考えられる方向セグメントの抽出方式を提案する。提案方式は漢字の特質の一つである構成要素の直線性に着目し、漢字パターンが近似的に方向を持つ線素（方向セグメント）の集合により構成されているという考えに基づいている。本方式の特徴は、デジタルノイズの影響を受けやすい細線化過程を含まないこと、原理的に任意の角度の方向セグメントを抽出できること、原パターンが保存可能なこと、アルゴリズムの簡単なことなどである。電子技術総合研究所編集の手書き文字データベースを用いて本方式の基本的特性を調べ、安定な抽出能力を有することを検証している。

第3章では、第2章で示した方向セグメントの抽出方式を用いてセグメント間の接続形状を記述し、交叉部・分岐部・屈曲部などの幾何学的特徴部をも同時に抽出する方式について詳述する。本方式は、漢字の持つ階層構造に着目したものであり、セグメントを基にして文字パターンの記述を試みている。接続している2本のセグメントの射影パターンからピークを検出することにより各特徴部が同一アルゴリズムのもとに抽出可能であり、セグメント間の構造を記述できる。漢字認識系で有効であるとされている部分パターン抽出への基礎的アプローチとして本手法の有効性を漢字部首の形状抽出実験により確かめ、部分パターン抽出への応用の可能性を示唆している。

第4章では、手書き漢字の認識方式として相関法と構造解析法の融合方式である矩形セグメント整合法を提案する。本認識方式は抽出された方向セグメントを矩形で近似し、基本処理単位を画素からセグメントを示す座標値に変換することにより処理の高速化を図っている。これは、漢字認識特有の問題点である字種の多さに伴う処理量の増加に対処しようとするものである。又、各セグメントが独立に取り扱えるため、文字認識において重要である各種技術、即ち、ずらし操作、相関計算、セグメント間の接続情報などが容易に且つ高速に実現できる。さらに近傍状態整合により局所的構造も考慮し、認識能力の向上を目指している。手書き教育漢字データベースを用いて認識実験を行い、セグメントの方向性という単一の特徴に着目した場合の認識能力を矩形セグメント整合法により検討し、認識系における方向セグメント抽出の有効性を確かめている。

第5章は結論であり、本研究で得られた諸結果を総括的に述べるとともに今後の課題について述べる。

関連発表論文

A 学会論文

- (A1) 馬場口, 北村, 塩野, 真田, 手塚: "細線化を用いない文字パターンの方向セグメント抽出の一手法", 信学論 (D), J 6 5-D, 7, PP.874-881 (昭和57-07).
- (A2) 馬場口, 相原, 真田, 手塚: "文字パターンにおける方向セグメントの接続形状と幾何学的特徴部の抽出", 信学論 (D), J 6 6-D, 5, PP.495-502 (昭和58-05).
- (A3) 馬場口, 正満, 真田, 手塚: "方形セグメント整合法による手書き漢字認識の検討", 信学論 (D), J 6 6-D, 10, PP.1209-1216 (昭和58-10).
- (A4) 塩野, 馬場口, 真田, 手塚: "方向性マッチングによる常用手書き文字の認識", 信学論 (D), J 6 3-D, 5, PP.402-409 (昭和55-05).

B 国際会議

- (B1) N.Babaguchi, T.Aibara, H.Sanada, Y.Tezuka: "Direction Segments Extraction and Its Application to the Recognition of Handprinted Chinese Characters", Proceedings of ICTP'83, PP.324-329 (1983-10).

C 学会研究会

- (C1) 馬場口, 塩野, 真田, 手塚: "特徴領域整合法による手書き文字の認識", 信学技報, PRL 81-19 (1981-06).
- (C2) 馬場口, 正満, 塩野, 真田, 手塚: "方向セグメントの矩形表現による手書き漢字の大分類", 信学技報, PRL 81-80 (1982-01).
- (C3) 正満, 馬場口, 真田, 手塚: "近傍状態を考慮した方向セグメントの矩形表現による手書き漢字の大分類", 信学技報, PRL 82-43 (1982-11).
- (C4) 馬場口, 相原, 真田, 手塚: "セグメント対応法による手書き漢字部分パターンの抽出と同定", 信学技報, PRL 83-59 (1984-01).
- (C5) 馬場口, 塚本, 相原: "手書き文字列における文字切り出しの基礎検討", 信学技報, PRL 83-58 (1984-01).
- (C6) 塩野, 馬場口, 真田, 手塚: "曲面補間を用いた文字サイズの正規化と方向性マッチングによる手書き文字の認識", 信学技報, PRL 78-66 (1979-01).

- (C7) 塩野, 馬場口, 真田, 手塚: "領域情報による大分類と方向性マッチングによる手書き文字の認識実験", 信学技報, PRL 79-25 (1979-07).

D 学会講演発表

- (D1) 馬場口, 塩野, 真田, 手塚: "テンプレートマッチングの拡張による手書き文字認識の一手法", 昭54関西連大, G8-15 (1979-11).
- (D2) 馬場口, 塩野, 真田, 手塚: "白領域の開閉状態を付加した方向性整合による手書き文字認識の一手法", 昭55信学総全大, 1312 (1980-03).
- (D3) 馬場口, 塩野, 真田, 手塚: "方向性と開閉状態を用いたパターン整合法による手書き文字認識の一手法", 昭55関西連大, G8-37 (1980-11).
- (D4) 馬場口, 塩野, 真田, 手塚: "特徴領域整合法による手書き文字の認識の一手法", 昭56信学総全大, 1379 (1981-04).
- (D5) 馬場口, 北村, 塩野, 真田, 手塚: "文字パターンの方向セグメント抽出の一手法", 昭56信学会情報システム部門全大, 72 (1981-10).
- (D6) 馬場口, 塩野, 真田, 手塚: "細線化を用いない文字パターンの特徴部分抽出法", 昭56関西連大, G8-14 (1981-11).
- (D7) 馬場口, 正満, 塩野, 真田, 手塚: "方向セグメントの矩形表現による手書き漢字の大分類の一手法", 昭57信学総全大, 1360 (1982-03).
- (D8) 馬場口, 相原, 真田, 手塚: "方向セグメントの形状判定と特徴部分抽出について", 情報処理学会第25回全大, 6B-2 (1982-10).
- (D9) 馬場口, 相原, 真田, 手塚: "方向セグメントの形状判定と幾何学的特徴部の抽出", 昭57四国連大, 12-12 (1982-10).
- (D10) 馬場口, 相原, 真田, 手塚: "セグメント対応法による手書き漢字の部分パターン抽出", 昭58四国連大, 12-3 (1983-10).

E 紀要

- (E1) N.Babaguchi, T.Aibara, H.Sanada, Y.Tezuka: "Some Applications to the Recognition of Handprinted Chinese Characters Using Direction Segments Extraction", Memoirs of the Faculty of Engineering, Ehime University, Vol.X, No.3, PP.163-174, February 1984.

目 次

第1章 緒 論	1
第2章 方向セグメント抽出	4
2.1 緒言	4
2.2 方向セグメント抽出方式の概要	5
2.3 距離の測定	6
2.4 方向コード化	8
2.5 高速方向コード化	11
2.6 方向画面	13
2.7 セグメント結合	13
2.8 抽出実験及び検討	17
2.8.1 センサー幅の検討	18
2.8.2 抽出方向の検討	20
2.8.3 4方向の方向セグメント抽出の検討	21
2.8.4 手書き漢字における抽出例	25
2.9 結言	25
第3章 方向セグメントの接続形状と幾何学的特徴部の抽出	26
3.1 緒言	26
3.2 提案方式の概要	27
3.3 セグメントラベル付け	27
3.4 接続セグメントリスト	28
3.5 射影パターン	31
3.6 ピーク検出	33
3.7 接続形状と特徴部の抽出	35
3.8 実験及び検討	39

3.8.1	片仮名の形状抽出実験と検討	39
3.8.2	細線化法との比較	41
3.8.3	漢字部首の形状抽出実験と検討	43
3.9	結言	45
第4章	矩形セグメント整合法	47
4.1	緒言	47
4.2	手書き漢字認識における矩形セグメント整合法の位置付け	47
4.3	矩形セグメント整合法の概要	49
4.4	矩形表現	50
4.5	相関計算	54
4.6	近傍状態整合	55
4.6.1	近傍状態の定義	55
4.6.2	セグメント間の対応付け	57
4.6.3	近傍状態の統合	58
4.6.4	近傍状態類似度	59
4.7	カテゴリー決定	60
4.8	認識実験及び検討	60
4.8.1	辞書の作成	61
4.8.2	セグメント数の統計的調査	61
4.8.3	ずらし幅・太め幅の検討	63
4.8.4	近傍状態整合の効果	63
4.8.5	分類特性の検討	65
4.9	結言	67
第5章	結論	68
	謝辞	71
	文献	72

第 1 章 緒 論

電子計算機によるパターン認識の研究が始められ、既に 20 年以上が経過している。当初、人間の持つ認識過程のメカニズムの解明という科学的要求とコンピュータへの入力機構の変革という社会的要求の 2 つから出発したパターン認識技術は日本のみならず世界的規模で活発な研究がなされた。数々の優れたパターン認識理論が展開されるのに伴い、^{(1)~(5)} 情報科学における最も重要な分野の 1 つとなるに至った。現在、認識対象は文字・図形、音声を始めとし画像^{(6)~(9)}、さらにはコンピュータビジョン⁽¹⁰⁾ と呼ばれる 3 次元物体にまで裾野は広がっている。パターン認識は画像処理や人工知能の諸分野^{(11)~(13)} とも深い関連性を持ち、次世代のコンピュータには重要不可欠な機能と目され、より知的な能力としてパターン理解⁽¹⁴⁾ にまで発展する気運にある。

日本におけるパターン認識技術の進歩の中で文字認識はパイオニア的役割を演じてきた。この理由には 2 つのことが考えられる。1 つは文字が 2 次元図形の中で比較的単純な図形的特徴を持ち、取り扱い易いと思われたこと、さらにもう 1 つはコンピュータへの入力手段の省力化として最も切実な要望があったことが挙げられる。特に 1970 年以降、通産省によるパターン情報処理大型プロジェクトの援助のもとに各機関で様々な研究がなされ、高い技術レベルを誇っている。

さて、最近の O A (Office Automation) 化の流行は計算機への日本文自動入力方式を極めて身近な問題とした。逐次入力方式である仮名漢字変換などの試みがなされているが、文字認識技術は伝票や原稿などの日本語文書を大量に取り扱う場合には極めて有効である。即ち、文書一括入力方式としての汎用文字認識技術の確立が望まれる。我が国においては欧米と異なり、タイプライターを日常的に使用する習慣性がないという特殊事情から、とりわけ手書き文字認識は重要であり、又、漢字処理は日本語処理における主要な課題と言えよう。

ここで文字認識方式を分類すると、まず印刷文字認識と手書き文字認識に分けられ、種類別には英数字、片仮名、平仮名、漢字に大別される。さらに入力形態別に分類するとオンライン文字認識とオフライン文字認識になる。印刷文字認識に関しては既に漢字まで読取対象になっており、OCR (Optical Character Reader) として商品化されている。又、オンライン文字認識では、ストローク (ペンダウンからペンアップまでを意味する) が抽出されていること、筆順が検出されるなどの利点から手書き漢字を含めて認識可能である。オフライン手書き文字認識においては、英数字、片仮名は実用レベルに到達しており、現在では手書き漢字 (漢字仮名混じり文を含む) 認識に研究の主眼が移ってきている。

手書き漢字認識には以下のような問題点が考えられる。第1に膨大な数の字種があることである。通常新聞などに用いる漢字で約2000字、教育漢字ですら881字と仮名の20倍程度の字種が存在する。これは漢字認識の量的問題と考えられ、それに伴う処理量の著しい増大がある。第2に構造の複雑さが挙げられる。仮名においては高々数ストロークであるが漢字では数十ストロークに及ぶものも存在する。第3の問題点は類似文字の多さである。漢字は長い歴史を持つ表意文字の代表であり、ストロークの有無が意味上における決定的な相違点となる。この顕著な一例が「大」「太」「犬」である。上述の第2、第3の問題は漢字認識の質的問題と考えられ、認識系における特徴抽出部及び識別部の高精度化が要求される。最後に第4の問題として、手書き文字認識全般に渡っていえるものであるが、筆記者の癖による文字パターンの手書き変形の多様性がある。極端な右上りの文字、曲線的な文字など筆記者の数に等しい変形の仕方があるといっても過言ではないであろう。故に構造的に安定な特徴の選択、抽出が必要である。このような背景のもとで手書き漢字は最も困難な認識対象と言えよう。これらに対処するため種々提案されており、例えば大分類の積み重ねによる段階的識別方式は量的問題を考慮したものであるし、相異

なる複数の相補的特徴を用いて認識能力の向上を図る方式は一層の高精度化の実現を目指すものである。

そこで、筆者は上述の問題点を解決するためには漢字の特質を十分考慮し、認識系の特徴抽出部や識別部に反映させる必要があると考え、本論文では漢字の特質を考慮した特徴抽出方式及び手書き漢字認識方式について検討する。まず、第2章では手書き漢字認識における有効な特徴抽出の1つであると考えられる方向セグメント抽出方式⁽¹⁵⁾について論じる。漢字の特質の1つである漢字構成要素の直線性に着目し、漢字パターンが近似的に方向を持つ線素（方向セグメントあるいは単にセグメントと呼ぶ）の集合により構成されているという考えに基づいている。第3章では第2章で示す方向セグメント抽出を用いてセグメント間の接続形状を記述し、交叉部・分岐部・屈曲部などの幾何学的特徴部をも同時に抽出する方式⁽¹⁶⁾について考察する。本方式は、漢字の持つ階層構造に着目したものであり、漢字認識における特徴抽出の1つであると共に、別の観点から見ると文字パターンの新しい記述法でもある。第4章ではセグメント抽出方式を利用し、矩形セグメント整合法⁽¹⁷⁾なる手書き漢字認識系を構成し、セグメントの方向性という単一の特徴に着目した場合の認識能力について検討する。第2章から第4章までの何れの手法についてもシミュレーション実験を行い、セグメント抽出方式の手書き漢字認識への有効性⁽¹⁸⁾を確かめる。

第2章 方向セグメント抽出

2.1 緒言

漢字の特質の1つとして漢字を構成する要素の直線性が挙げられる。漢字パターンの構造に着目した場合、漢字パターンは近似的に方向を持つ線素（方向セグメントと呼ぶ）の集合により構成されていると考えられる。これは漢字が平仮名などに比べて直線部が多く、大きい曲線は限られた字種にしか存在しないことから直観的に知ることができる。このことは方向セグメントを抽出することにより漢字パターンが近似表現されることを示唆している。従って、安定に方向セグメントを抽出することは手書き漢字に対して有力な特徴抽出法である。⁽¹⁹⁾

従来、この種の方向セグメントの抽出には細線化を施すものが多く^{(20),(21)}、チェーンコードによる抽出手法が提案されている⁽²⁰⁾。この方法は、細線化アルゴリズムの特性に大きく左右され、“ヒゲ”、“変形交点”、“小さい枝の消失”などの細線化雑音の影響が多い⁽²⁰⁾。これを解決するためには煩雑な局所的処理が要求されるので安定な抽出という観点からすれば細線化過程はあまり好ましくないといえる。又、チェーンコード手法は8近傍点の接続方向をコード化する手法であり、原理的に抽出できる方向は45°区切りにしかとれない難点がある。

本章では、細線化を用いなくて原パターンから直接に方向セグメントを抽出する方式について論じる。本方式の特徴は、上述の難点を克服するために細線化を含まないこと、任意の角度の方向セグメントを抽出できることなどである。文字パターンをマクロ的、アナログ的に取り扱うことにより種々の雑音の影響を軽減し、忠実且つ安定な方向セグメント抽出方式を実現している。

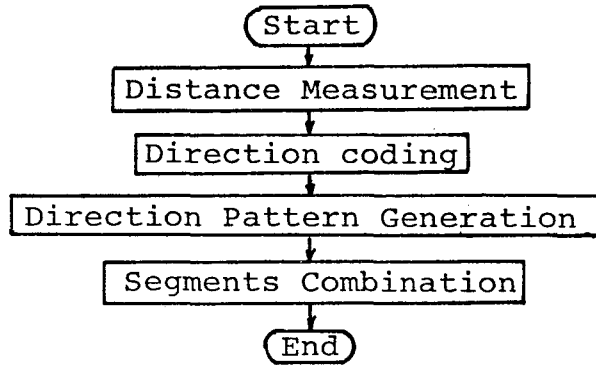


図2-1 本方式のフローチャート

2. 2 方向セグメント抽出方式の概要

本方式で対象とする入力文字パターンは、文字部を画素値1、背景部を画素値0とした2値パターンFであり、次式で表す。

$$F = \{ (i, j, f(i, j)) \mid f(i, j) \in \{0, 1\}, \\ i, j = 1, \dots, N \} \quad (2-1)$$

但し、 $f(i, j)$ は画素値、 i 軸、 j 軸は各々画面の縦方向、横方向を表す。又、 $f(1, 1)$ は画面の左上、 $f(N, N)$ は右下を示す。画素値1(0)の画素集合を黒(白)領域、黒(白)領域の各点を黒(白)点と呼ぶ。

入力文字パターンからの方向セグメントの抽出にあたって、細線化を用いなくて原パターンから直接抽出する方法として、以下に述べる方式を提案する。まず、2値入力パターンから抽出しようとするセグメントの方向を定める。次に黒点から各方向への距離を測定し、その最長距離を持つ方向のコードに画素値を変換する。変換された画素値に従い、数個の方向画面を作成する。方向画

面内の黒領域は抽出されうる方向セグメントの候補である。しかし、セグメントの交叉部で切れ切れになる場合があり、これを補正するためにセグメント結合処理を施す。セグメント結合処理を施した後の各方向画面における黒領域が文字パターンから抽出された方向セグメントであり、それらは各々の方向に最長距離を持つ画素の集合と考えることができる。

本方式のフローチャートを図2-1に示す。次節以下、フローチャート順に距離の測定 (Distance Measurement) , 方向コード化 (Direction Coding) , 方向画面の生成 (Direction Pattern Generation) , セグメント結合 (Segments Combination) の手順の詳細を示す。

2. 3 距離の測定

抽出する方向の角度 $\theta_k (0^\circ < \theta_k \leq 180^\circ, k=1, \dots, K)$ は向きを考えないので、 θ_k と θ_k+180° は同一方向であり、以後、 θ_k と θ_k+180° の角度を単に θ_k 方向と呼ぶ。ここで、 $\theta_k=0^\circ$ は j 軸方向であり、 K を抽出方向数と呼ぶ。

いま、簡単のため黒点を $f_b(i, j)$ と記し、1つの黒点 $f_b(i, j)$ での連結黒領域内部における θ_k 方向の幾何学的距離 $d_k(i, j)$ を求める。デジタル画面を取り扱う際、正方格子間に間隔が $\sqrt{2}$ 以内で等間隔に並ぶ方向は、 $\theta_k = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ 即ち水平、垂直、対角、逆対角の4方向のみである。従って、上記以外の方向の距離測定は、直接画素を探索することが不可能であるのでセンサーにより測定する。センサーは各方向の距離を測定するときに用いる幅 r の帯のことを言い、センサーの構成要素をセルと呼ぶ。図2-2に $f_b(i_0, j_0)$ でのセンサーを示す。

2値パターン F の定義域内で座標 (i_0, j_0) を通り、傾き $\tan\theta_k$ の直線における間隔 $\delta (1 \leq \delta \leq \sqrt{2})$ の点列を $T_n = (\tilde{i}_n, \tilde{j}_n), (n=-a, \dots, b;$
 a, b は整数) とする。点列 T_n はデジタル画面上の仮想的な点であり、通

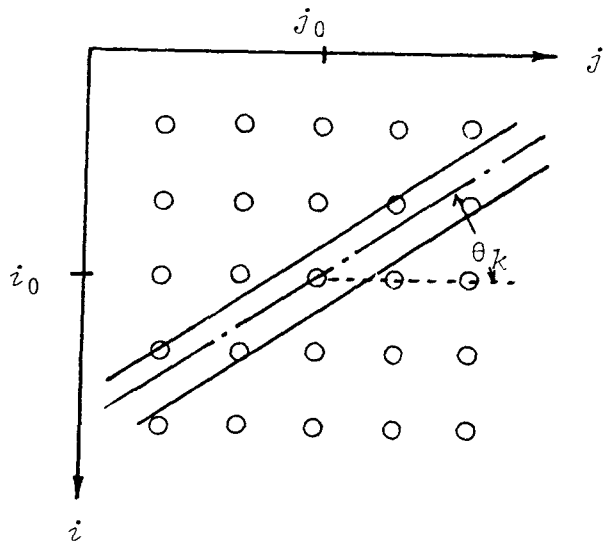


図2-2 センサー

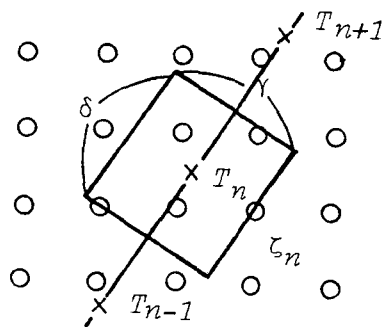


図2-3 セルと格子点

常、格子点とは一致しない。次にセンサーによる幾何学的距離 $d_k(i_0, j_0)$ の求め方を説明する。 $T_0=(i_0, j_0)$ を基準点とし、点列 $T_n(n=0, \dots, b)$ を考える。添字が正の点列は T_0 から見て角度 θ_k の向きに並ぶ点列である。点列の n 番目の点 T_n を中心として直線の傾きの方向と垂直方向に各々一辺 δ , r の矩形を第 n セル ζ_n とする。第 n セルと格子点の状態を図 2-3 に示す。

ここで第 n セル内の全ての格子点が白点であるとき黒領域の連結性が失われたと判断する。このように ζ_1 から順次判定を繰り返し、 $\zeta_{n'}$, ($n' \in n$) で連結性が失われたときの T_0 からの角度 θ_k の向きの幾何学的距離 $d_k^+(i_0, j_0)$ は、

$$d_k^+(i_0, j_0) = \sqrt{(i_{n'-1} - i_0)^2 + (j_{n'-1} - j_0)^2} \quad (2-2)$$

で得られる。

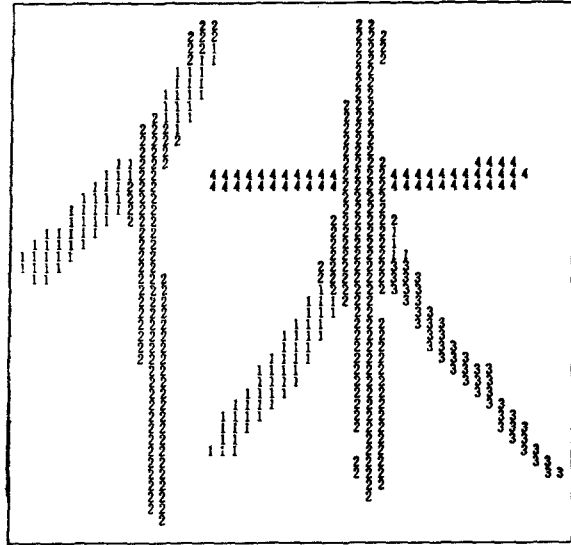
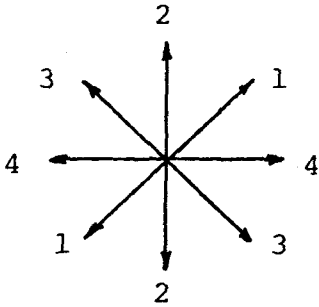
逆に、座標 (i_0, j_0) から角度 $\theta_k + 180^\circ$ の向きには添字が負の点列 $T_n(n=1, \dots, -a)$ が並び、 ζ_{-1} より順次判定を行い、上と同様に測定される距離を $d_k^-(i_0, j_0)$ とすると、 θ_k 方向の距離 $d_k(i_0, j_0)$ は

$$d_k(i_0, j_0) = d_k^+(i_0, j_0) + d_k^-(i_0, j_0) \quad (2-3)$$

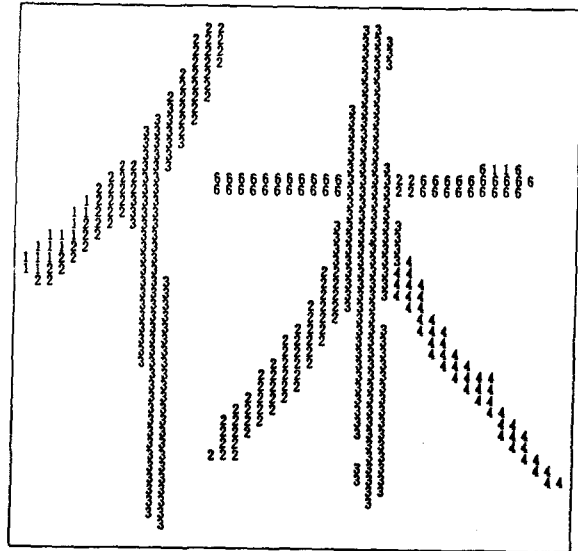
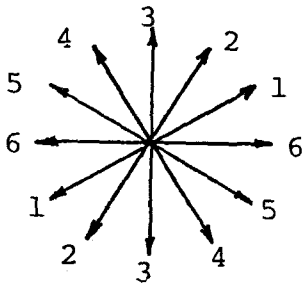
となる。この処理は仮想的なセルの積み重ねによりアナログ的に連結黒領域内部の距離を測定することに相当する。本方式では黒点から任意の角度の距離を測定でき、これ故に任意の方向のセグメントが抽出可能となる。センサーによる θ_k 方向の距離測定は 2 値パターン F のすべての黒点で実行される。

2. 4 方向コード化

θ_k 方向に対し、添字 k を方向コードと呼ぶ。又、抽出方向数 K に対しては $1, \dots, K$ までが各々方向コードになり、この数字の組を K 方向コードと

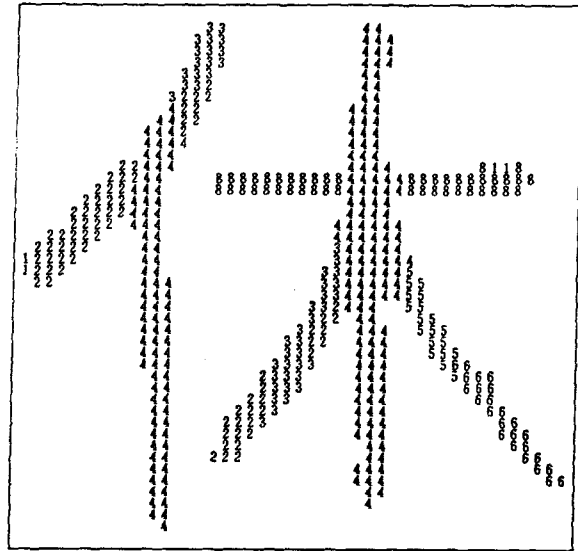
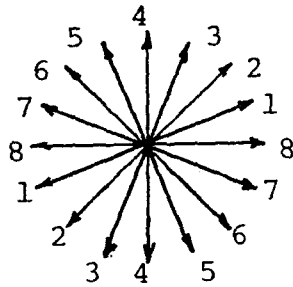


(a) 4方向



(b) 6方向

図2-4 方向コード化パターン



(c) 8方向

図2-4 方向コード化パターン

呼ぶ。

各黒点における θ_k 方向への幾何学的距離 $d_k(i, j)$ に基づき黒点 $f_b(i, j)$ の値を次式により方向コード k に変換する。

$$f_b(i, j) = k, (k=1, \dots, K) \quad (2-4)$$

$$\text{但し, } d_k(i, j) = \max_{1 \leq m \leq K} \{d_m(i, j)\}$$

式(2-4)の意味は注目する黒点における連結黒領域内部の最長の距離を持つ方向の方向コードをその黒点の値に変えることである。従って、 k の値はその黒点が何れの方法のセグメントに属しているかを示す。式(2-4)により変換されたパターンを方向コード化パターンと呼ぶ。図2-4(a)から(c)各々に4方向、6方向、8方向の抽出する方向の角度、方向コード及び方向

コード化パターンを示す。尚、方向コードにより各画素値を変換する手法には岡部らによる盛り上げ変換によるものも提案されているが⁽²²⁾本アルゴリズムの方が簡単である。又、水平、垂直、対角、逆対角の4方向コード化については直接、格子点を探索することによりセンサーを用いずに高速な方向コード化が実現できる。その詳細を次節に示す。

2. 5 高速方向コード化

画素 $f(r), r=(i, j)$ に対応させ、図2-5のマスクの8つの向きの画素数を記録する $G_m(r), (m=1, \dots, 8)$ を用意する。

第1に左上から右下への走査でマスクを移動させ、マスク中心点、即ち図2-5における0が黒点であるときマスクの1から4の点を調べ、次式を実行する。

$$\begin{cases} G_m(r)=0 & \text{if } f(r+r_m)=0 \\ G_m(r)=G_m(r+r_m)+1 & \text{if } f(r+r_m) \neq 0 \end{cases} \quad (2-5)$$

但し、

$$r_1=(-1, 1), r_2=(-1, 0), r_3=(-1, -1), r_4=(0, -1), m=1, 2, 3, 4$$

この処理を繰り返し、マスクが右下に達したとき、 $G_1(r)$ から $G_4(r)$ には点 $f(r)$ でのマスクにおける1から4の向きの画素数が記録される。

第2に、右下から左上への逆走査を行い、式(2-5)において

$$r_5=(1, -1), r_6=(1, 0), r_7=(1, 1), r_8=(0, 1), m=5, 6, 7, 8$$

とし、同様の処理を反復する。マスクが左上に達したとき、 $G_5(r)$ から $G_8(r)$ には点 $f(r)$ でのマスクにおける5から8の向きの画素数が記録される。

における各方向の距離 $d_k(r), (k=1, \dots, 4)$ は、

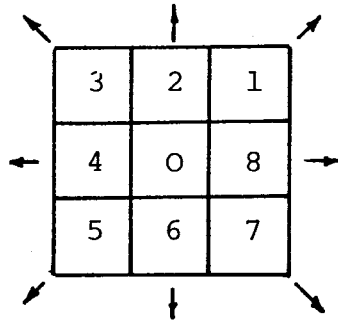


図2-5 マスク

$$\begin{cases}
 d_1(r) = \sqrt{2} (G_1(r) + G_5(r)) \\
 d_2(r) = G_2(r) + G_6(r) \\
 d_3(r) = \sqrt{2} (G_3(r) + G_7(r)) \\
 d_4(r) = G_4(r) + G_8(r)
 \end{cases} \quad (2-6)$$

で与えられる。次に、

$$\tilde{d}_k(r) = \max \{ d_k(r), d_k(r+r_2), d_k(r+r_4), d_k(r+r_6), d_k(r+r_8) \} \quad (2-7)$$

により $\tilde{d}_k(r)$ を求めるとセンサー幅2で求めた場合に近くなり、又、 $d_k(r)$ はセンサー幅1での測定値となる。 $d_k(r)$, $\tilde{d}_k(r)$ を基に式(2-4)により4方向コード化が施され、センサーによりコード化する場合の約5分の1の処理時間で実行される。

2. 6 方向画面

式(2-4)により黒点の値は1, ..., KのK値に変換される。この値は各黒点の連結黒領域内部の最長距離の方向を示す。次に方向コードに従い黒領域を分割し、次式によりK個の2値パターンを生成する。

$$F_k = \{(i, j, z) \mid \begin{cases} z=1 & \text{if } f(i, j)=k \\ z=0 & \text{otherwise} \end{cases}\} \quad (2-8)$$

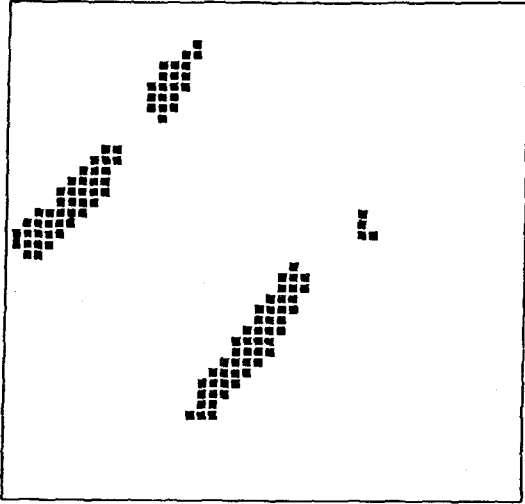
ここで F_k を第 k 方向画面と呼び、 θ_k 方向に距離が最長な点集合であり、抽出されうる方向セグメントの候補である。よって、黒領域からK個のサブパターンが生成され、 F を原パターンとすると、

$$F = \bigcup_{k=1}^K F_k, \quad \bigcap_{k=1}^K F_k = \emptyset \quad (2-9)$$

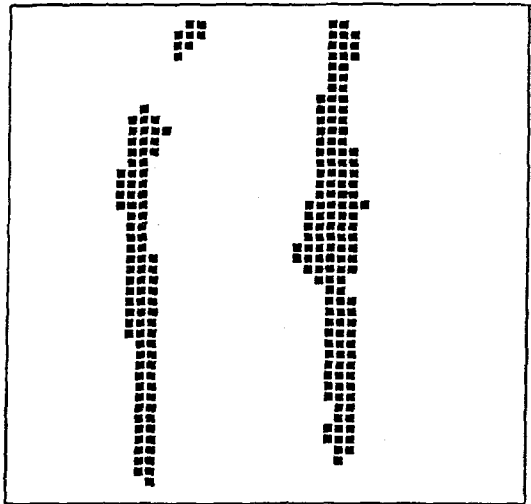
を満足する。図2-6に図2-4(a)の方向コード化パターンから生成される方向画面の実例を示す。

2. 7 セグメント結合

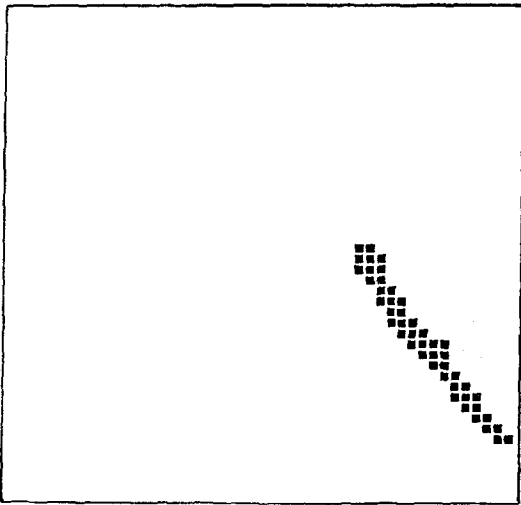
方向画面を生成した際に原パターンで連続であるべき方向セグメントが不連続になる不都合が生じるのでセグメント結合を施す。図2-7にその具体例を示す。ここでは簡単のため水平(方向コード4)、垂直(方向コード2)、対角(方向コード3)、逆対角(方向コード1)の4方向について述べる。図2-7(a)のパターンに4方向コード化を施すと領域A、Cは方向コード4、領域B、Dは方向コード2となる。次に、方向画面 F_2 、 F_4 を生成すると同図(b)、(c)になり、第4方向画面 F_4 における方向セグメントが領域Dで不連続になる。本来、同図(a)において方向セグメントは水平、垂直方向の1本ずつと思われるので第4方向画面における方向セグメントを結合すれば、



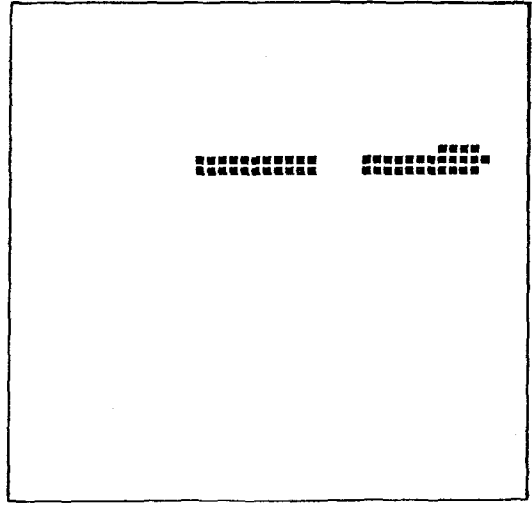
逆对角



垂直



对角



水平

图2-6 方向画面

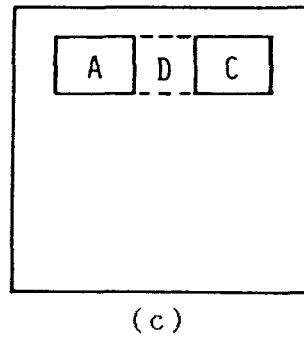
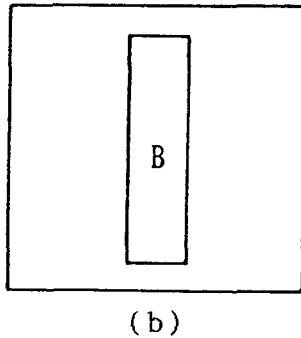
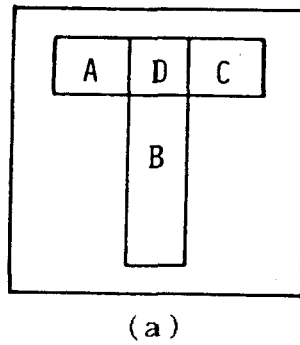
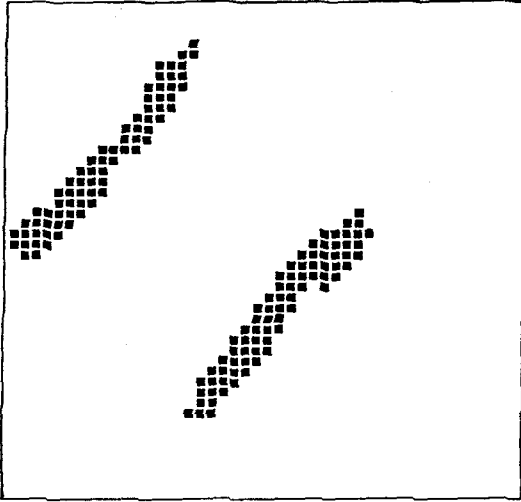


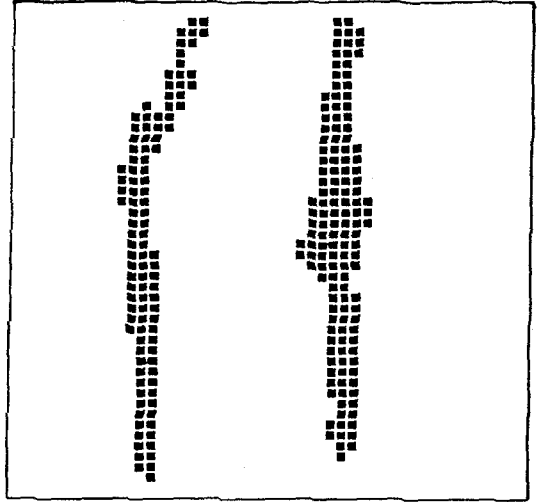
図2-7 セグメント結合

所望の結果が得られる。以下にセグメント結合のアルゴリズムを示す。

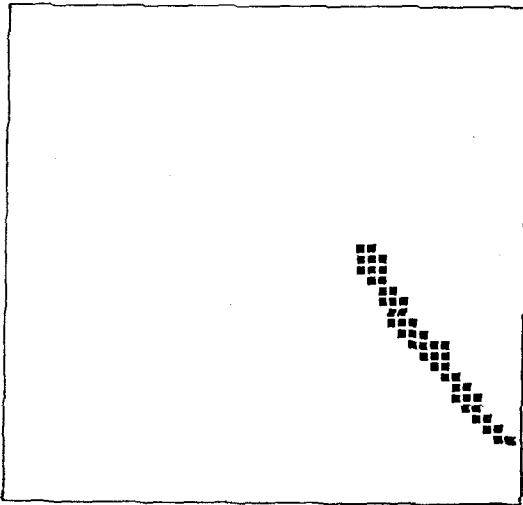
- (1) 第 k 方向画面 F_k 内を通常の走査（左上から右下）を行い、黒点のうち θ_k 方向の第1セル c_1 内に黒点を検出できない点（連結候補点と呼ぶ）を求める。
 - (2) 連結候補点より角度 θ_k の向きにセンサーを伸ばし、第 k 方向画面内に黒点の有無を調べる。黒点が存在しない場合には走査を続け、(1)の処理に戻る。存在する場合には(3)の処理を行う。
 - (3) 連結候補点と検出された黒点間のセンサーの連結性が原パターンにおいて保存される時、第 k 方向画面においてセンサー内部の各点のうち原パターンで黒点である点を黒点に変える。結合が完了すると走査を続け、(1)の処理に戻る。
- (1) から (3) の処理は生成されたすべての方向画面で実行する。図2-8



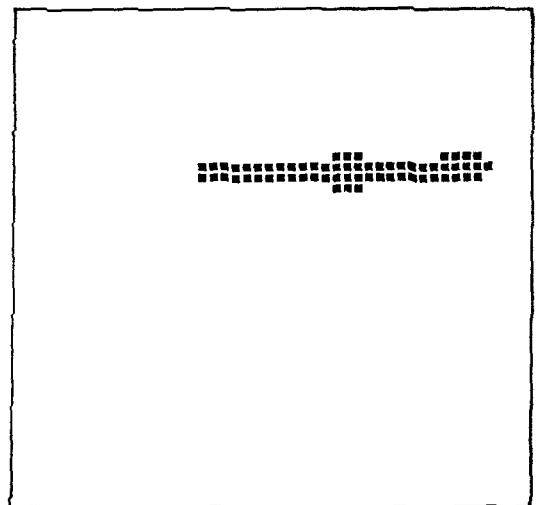
逆対角



垂直



対角



水平

図2-8 方向セグメント抽出の実例

の例では第4方向画面（同図（c））において領域Aの右端の点が連結候補点となり、その点から水平方向に黒領域が存在する。又、センサーを伸ばした部分、即ち点線の部分が原パターン（同図（a））では領域Dに相当し、黒領域であるのでセグメント結合が施される。

セグメント結合を施した後の各画面での黒領域が文字パターンより抽出された各方向の方向セグメントである。図2-4（a）の文字パターンより抽出された方向セグメントの実例を図2-8に示す。第3方向画面以外はセグメント結合が施されている。

2. 8 抽出実験及び検討

手書き文字に対する方向セグメント抽出方式の基本的特性を調べるため、漢字と同様に直線部の多い片仮名を用いて統計的実験を行った。一方、手書き漢字に対しては抽出例を示す。入力データには電総研旧常用手書き文字データベース（片仮名）⁽²³⁾と同手書き教育漢字データベース（ETL8）⁽²⁴⁾を使用した。いずれも2値化を施し、入力画面サイズは片仮名、漢字各々20、60（縦、横同じ）である。使用計算機システムは大阪大学大型計算機センターACOS-900、使用言語はFORTRANである。

統計的実験としては、各方向のセグメント数を1文字ごとに測定した。即ち、 θ_k 方向のセグメント数 n_k , ($k=1, \dots, K$) を検出し、その安定性並びに人間の評価との比較を行うことにより抽出性能を評価する。又、以下の説明では、4方向は（45°, 90°, 135°, 180°）、6方向は（30°, 60°, 90°, 120°, 150°, 180°）、8方向は（22.5°, 45°, 67.5°, 90°, 112.5°, 135°, 157.5°, 180°）である各々の方向のセグメントを抽出することであり、角度の小さいものから順に方向コードを1から割り当てる。

2.8.1 センサー幅の検討

図2-9にセンサー幅 $r = 1, 2$ の場合の4方向コード化を示す。 $r = 1$ の場合は文字輪郭部に雑音が顕著であり、 $r = 2$ の場合は雑音も少なく、忠実に方向コード化がなされている。この理由は $r = 1$ で距離を測定すると文字輪郭部の凹凸に大きく影響されるからである。この影響を除去するために文字輪郭部の平滑化も考えられるが、原パターンの保存を優先させセンサー幅を大きくすることによりほぼ解決した。表2-1に4方向の方向セグメント抽出とセンサー幅の関係を示す。「ア」から「ン」まで46カテゴリー50サンプルにおいて各方向のセグメント数 (n_1, n_2, n_3, n_4) が同じ値になったサンプル数とカテゴリーを示す。

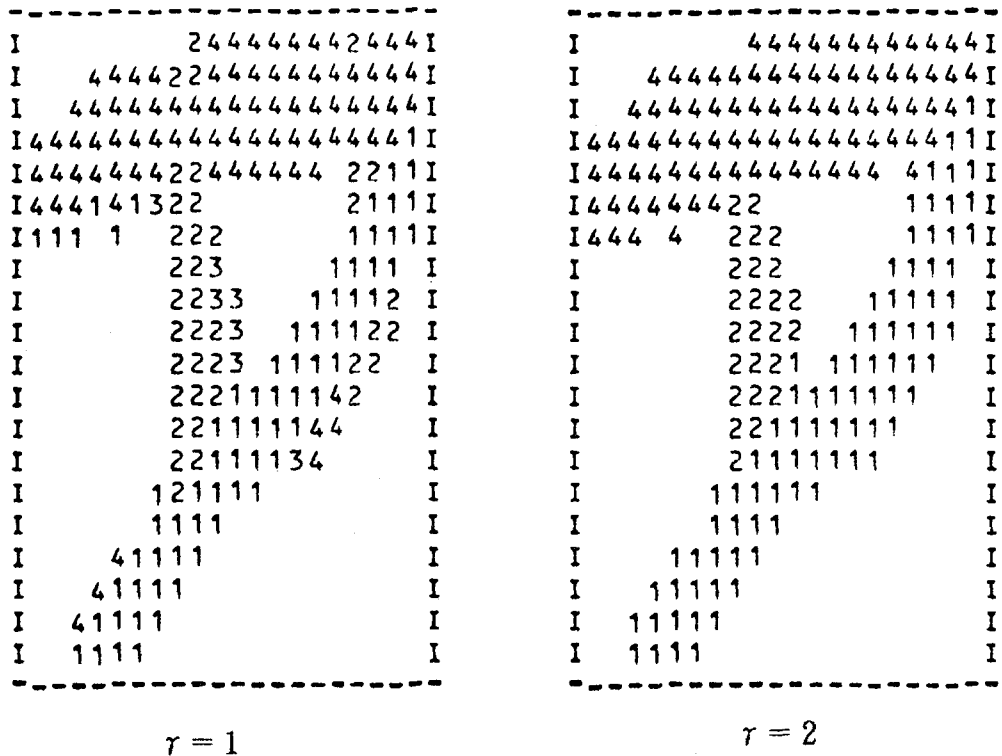


図2-9 センサー幅の比較

表 2-1 各センサー幅での抽出結果

サンプル数	$r=1$	$r=2$	$r=3$
40以上	エオコサセテナ モヨ (9)	エオコサスセテト ナニヒモユヨロワ (16)	アエオコサスセナ ニヒモユヨロ (14)
30以上	アイケスソニネ ハヒフホムヤユ ルレロワヲ (19)	アイケシソネノハ ヒマヤラリルヲ (15)	イシソテトネフマ ヤリルワ (12)
20以上	ウカキクシタチ トヌノヘマミメ ラリン (17)	ウカキクタチツヌ ヘホミムメレン (15)	ウカキクケタチヌ ノハヘホミムメラ レヨン (19)
10以上	ツ (1)	(0)	ツ (1)

表 2-2 6方向, 8方向の抽出結果

サンプル数	6方向	8方向
40以上	ニメ (2)	ノレ (2)
30以上	イクトヌノフホマヤ (6)	イケフヘムメ (6)
20以上	アウエオカキケコサシスセ ソタサナネハヒヘミムモユ ラリルレワン (30)	アウエオカクコサスセソタ チツトナニヌネハヒホマミ ヤユラリルロワン (32)
10以上	ツテヨロヲ (5)	キシテモヨヲ (6)

表 2-3 チェインコードによる抽出結果

40以上	コスセトニノヒミロ (9)
30以上	イウカヌハフマモヤユヨラルレワヲ (16)
20以上	アエオキタケサシソタチツテナネヘホムメリン (21)

同じ値になったサンプル数が多い程、良好に抽出されたことを示し、 $r = 2$ の場合が最も良好であった。これはセンサー幅2で距離を測定した場合に最も忠実に測定されたことを示唆している。一方、 $r = 3$ の場合は文字部の接近部において本来離れているセグメントがセンサーにより連結していると判定される弊害などにより劣下したものと考えられる。又、目視によると使用データの線幅はおよそ3、4画素幅であることから文字の輪郭線の1画素幅の凹凸は全体的形状には影響しない。ところが、2画素幅以上の凹凸は形状変化とみなされ、このことからセンサー幅2程度が適当であろう。文字線幅とセンサー幅の関係は、センサー幅を大きくすると文字パターンの微細な形状変化には鈍感であると推察される。ここでは実験的にセンサー幅2 ($r = 2$) がほぼ良好と考え、以下の実験を行った。

2.8.2 抽出方向の検討

6方向、8方向の方向セグメント抽出結果を表2-2に示す。表2-1と同様に46カテゴリー50サンプルでの各方向のセグメント数(n_1, \dots, n_6), (n_1, \dots, n_8) が同じ値になったサンプル数を表している。表2-1と比較すると特定のカテゴリーを除いて、セグメント数はばらつき、抽出安定性は劣下している。これは斜め方向のセグメントの角度が一定ではなく視覚的には同一視されるセグメントで方向コードが変化する場合が多いからである。従って、安定性の面からは6方向、8方向は現段階では有力であるとは言えないが、類似方向のセグメントの後処理などにより改善される余地が残されている。又、「メ」「ノ」「レ」などが4方向の場合より安定に抽出される理由は、逆対角の方向セグメントが安定に抽出できたためである。すなわち、 45° の角度より 30° 、 60° あるいは 22.5° 、 67.5° の角度で抽出する方が良好であることが分かる。よって、斜め方向の方向セグメントが安定に抽出されうる角度を調べる必要があると言えよう。

2.8.3 4方向の方向セグメント抽出の検討

最も汎用性の高いと考えられる水平、垂直、対角、逆対角の4方向の方向セグメント抽出について細線化を伴う手法との比較実験及び定量的評価を行った。ここでは比較手法として細線化とチェインコードによる手法を採用した。⁽²⁰⁾ 簡単にその手法について述べると、細線化はHilditchのアルゴリズム⁽²⁵⁾を用い、細線化パターンにおいて交叉数⁽²⁵⁾が0, 3, 4の点は取り除く。次に開始点から6点ずつとり、その間の方向の*i, j*成分の平均を取った後、3点目に平均した値を代入する。すなわち、移動平均により方向セグメントの平滑化を施す処理を行っている。さらに、本方式と同様に4つの方向画面を作成し、除去していた点や方向を決められなかった点を各画面に含め、方向セグメントの切れを無くす。長さ3以下のセグメントは雑音とみなし、方向セグメント数を求めている。表2-1, 表2-2と同様にチェインコードによる結果を表2-3に示す。

表2-1と表2-3を比較すると本方式の方が (n_1, \dots, n_4) が同じ値になるサンプル数が多く、かなり良好であると考えられる。しかし、各方向のセグメント数のうち最も安定性の良くない方向の結果に左右され、この評価基準だけで性能を判断することは難しい。たとえば、ある1つの方向のセグメント数がばらついていても他の方向は全て特定のセグメント数に集中している例が考えられ、このような場合は極めて多い。そこで、方向セグメント抽出の安定度を測る安定指数 S_F (Stability Figure) を定義する。各方向のセグメント数 (n_1, \dots, n_4) のNSサンプルについての度数分布を求めたとき、各方向について特定のセグメント数にサンプル数が集中している程、方向セグメントの抽出性能は安定していると考えられる。方向コード*k*についてセグメント数に関する最大サンプル数を N_k とすると S_F は次式で定義する。

$$S_F = \frac{1}{4} \sum_{k=1}^4 (N_k/NS)^\alpha, \quad (0 \leq S_F \leq 1) \quad (2-10)$$

但し、 $\alpha=10.3$ であり、 $0.8^\alpha=0.1$ を満足する α である。つまり、 N_k/NS が0.9から1.0であるもの、すなわち、ある特定のセグメント数に9割以上集中している方向は安定であるとして高い値を与え、反対に0.8以下のものには不安定であるとして0.1以下の値しか与えない。この性質は α を大きくする程極端になる。安定指数により4つの方向の平均的評価が可能となり、各文字による抽出の容易さを表す。

図2-10に本方式(This Method)とチェインコードによる方式(Chain Code)の50サンプルでのカテゴリー別安定指数と平均安定指数を示す。本方式がチェインコードによる方式より劣るものは6カテゴリーにすぎず、平均安定指数は約0.2上まわっている。又、サンプル当りの平均処理時間は、本方式、チェインコードによる方式各々47ms、50msとなり、ほぼ同様の結果が得られた。

次に、目視との比較を行った。まずカテゴリー別に想定される標準セグメント数の組を1個ないしは数個設定しておき、本方式で得られた結果と照合し、標準セグメント数の組以外の組合せが生じた場合にはうまく抽出できなかったものとする。たとえば、「エ」の標準セグメント数の組は水平方向2本、垂直方向1本として、(0, 1, 0, 2)としている。この場合、片仮名2300サンプルのうち85.6%が標準セグメント数の組と一致した。このように本方式は目視ともほぼ対応する結果が得られた。

従って、本方式では細線化を伴う方式よりも良好な特性が得られ、人間の直観ともほぼ対応する方向セグメントが抽出されることが分かった。この理由として、デジタルノイズの影響の受けやすい細線化過程を含まないこと、アナログ的に各点からの距離を測定するので文字パターンの雑音やひずみの影響を受けにくいことなどが挙げられ、有力な方向セグメント抽出方式の1つと言え

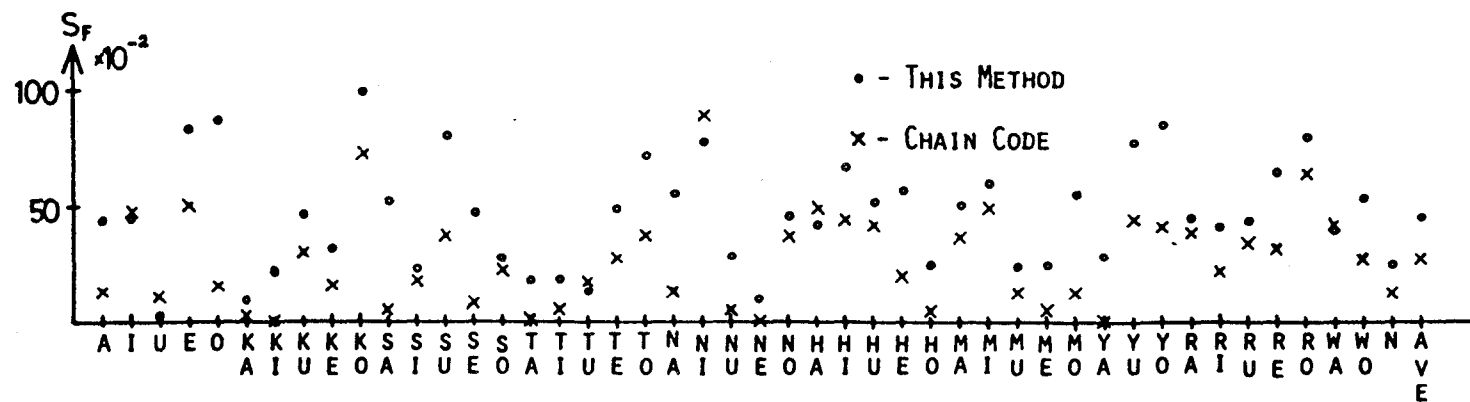


図2-10 カテゴリー別安定指数(片仮名)



図2-11 手書き漢字に対する抽出例

よう。

2.8.4 手書き漢字における抽出例

手書き漢字サンプルとして電総研手書き教育漢字データベース (ETL8) における第1データセットの冒頭の20文字を用い、4方向の方向セグメント抽出を行った実例を図2-11に示す。同図では左から原パターン、方向セグメント (逆対角 45° 、垂直 90° 、対角 135° 、水平 180°) を各々表す。手書き漢字においてもほぼ忠実に方向セグメントが抽出されているのが認められる。

2.9 結言

本章では、細線化を用いないで原パターンから直接に方向セグメントを抽出する方式を示し、抽出実験により細線化を伴う方式に比べ良好な結果が得られることを確認した。本方式の特徴を列挙する。

- (1)方向ごとに距離を測定するという簡単な処理により直観的に対応付け可能な方向セグメントが抽出できる。文字パターンをマクロ的、アナログ的に取り扱うことが好結果を与えたものと推察される。
- (2)任意方向のセグメントが抽出できる。抽出方向を予め決めてからセグメント抽出を行う場合のみならず、学習により良好に抽出されうる方向を選びだし、抽出を行う場合にも有効であろう。一例をあげると、手書きの代表的変形の1つである右上りパターンからセグメントを抽出する際に水平方向にわずかの角度を与えて抽出するなどの場合に利用できる。
- (3)原パターンを保存できる。方向セグメントが方向画面内の黒領域として表現されるので、手書き文字に有効とされる方向性を考慮した認識方式⁽³³⁾に適用が容易である。次章以下それらについて言及する。

第3章 方向セグメントの接続形状 と幾何学的特徴部の抽出

3.1 緒言

漢字パターンの構造に関する考察を加えると漢字パターンは上下、左右、内外などの複数の部分パターンから構成されているものが多い⁽²⁶⁾。各々の部分パターンは片仮名と同程度か、あるいはやや複雑な構造を持っているので部分パターンに対して手書き片仮名認識で有効な様々の考え方が利用できる。従って、部分パターン抽出は漢字の特質である階層構造に着目した有力な方式であると考えられる。さて、第2章で述べた方向セグメントを基本構成単位とした場合には、構成の小さい順に見て、方向セグメント、部分パターン、漢字パターンとなる。言い換えると方向セグメントの集合が部分パターンであり、部分パターンの相対的位置関係を記述することにより漢字パターンが表現される。そこで方向セグメントを1次の情報とすると高次の情報としてセグメント間の接続情報を抽出することは部分パターン抽出への基本的手法として重要であると考えられる。

本章では、方向セグメントの接続形状と交叉部、分岐部、屈曲部などの幾何学的特徴部（以後、特徴部と呼ぶ）の抽出方式^{(27)~(29)}について述べる。接続している2本のセグメントの形状が54種に分類された交叉形状、分岐形状、屈曲形状のいかなる形状に属するかを判定し、特徴部も同時に抽出する。特徴部抽出に関しては細線化を施し、3×3の論理マスクによる手法が一般的であるが⁽³⁰⁾、第2章で述べたように細線化雑音の影響で安定に抽出できるとは限らない。特に屈曲部の抽出には有効なアルゴリズムも明らかではないので様々な手法が試みられている^{(11), (31)}。ここでは、セグメントの接続形状を示す交叉部、分岐部、屈曲部はセグメント単位に処理を行うので重要な情報であるが端点部は不必要であり、抽出は行わない。本方式の特徴は接続している2本のセ

グメントの射影パターンの概形とセグメントの方向というマクロ的特徴にのみ依存することと各特徴部抽出のために個々の独立な処理を施していないことである。

3. 2 提案方式の概要

ここで取り扱う入力パターンは一定サイズに正規化された2値パターンであり、画素値が0, 1の画素を各々第2章と同様に白点, 黒点と呼ぶ。この入力パターンから、第2章で示した方向セグメント抽出により方向セグメントを得る。得られた方向セグメントについてセグメントごとに処理を施すため各々にラベルを与える。ここで対象とするのは2つのセグメントの接続形状であるので、文字パターンを構成するセグメント間の接続関係を表す接続セグメントリストを作成する。セグメント間の接続形状は射影パターンに反映され、特徴部の座標は射影パターンにおけるピークの位置により決定する。最終的に方向セグメントの接続形状を54種類の交叉・分岐・屈曲形状のいずれかに分類するとともに特徴部の座標を得る。

以上が提案する方向セグメントの接続形状と幾何学的特徴部の抽出方式の概要であるが、本方式をフローチャートの形で示したのが図3-1である。以下、方向セグメント抽出 (Direction Segments Extraction), セグメントラベル付け (Segment Labeling), 接続セグメントリストの作成 (Connected Segment List Generation), 射影パターンの作成 (Projection Pattern Generation), ピーク検出 (Peak Detection), 接続形状と特徴部の抽出 (Extraction of Connectivity Structure and Feature Region) の処理の詳細を説明する。

3. 3 セグメントラベル付け

本方式では、水平 (方向コード4), 垂直 (方向コード2), 対角 (方向コ

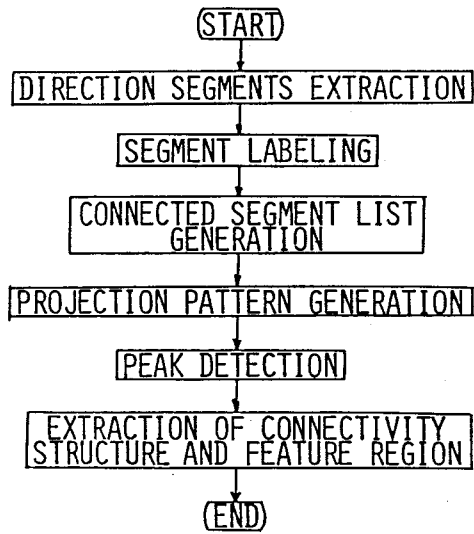
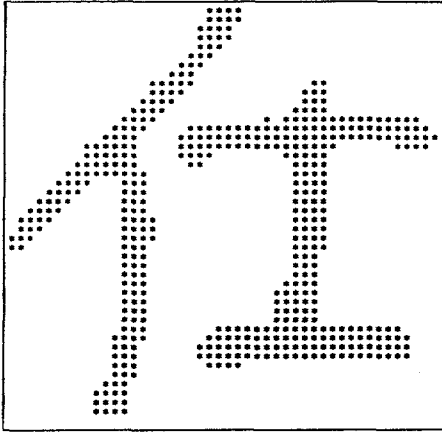


図3-1 抽出方式のフローチャート

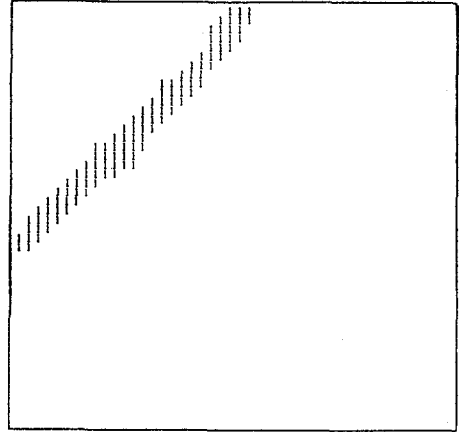
ード3), 逆対角(方向コード1)の4方向セグメント抽出を行う。方向セグメントは方向画面内の黒領域として得られるが, 各々のセグメントにセグメントラベル l , ($l=1, \dots, l_{max}$) を与える。但し, l_{max} は総セグメント数。ラベル付けは方向コード1から4の順, すなわち, 逆対角, 垂直, 対角, 水平の順に付け, 同一方向のセグメントには左上から右下への走査(ラスト走査)を行い, 出現する順にラベル付けし, 画素値を l に変える。このとき一定面積以下の微小セグメントは消去する。又, セグメントラベル l のセグメントの方向コード k_l , ($k_l \in \{1, 2, 3, 4\}$) も記録しておく。図3-2にセグメントラベル付けの実例を示す。同図(a)は原パターンであり, (b)から(d)はラベル付けされた逆対角, 垂直, 水平の各セグメントである。

3. 4 接続セグメントリスト

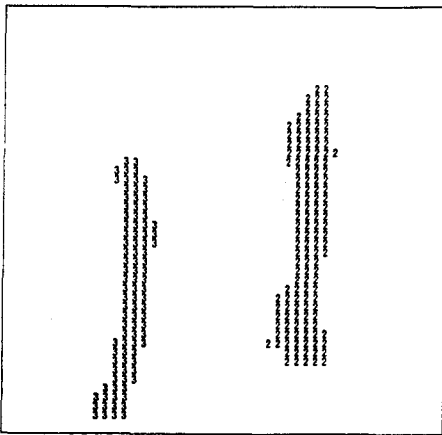
各セグメントの接続関係を表す接続セグメントリスト L_s を定義する。 L_s は $l_{max} \times l_{max}$ の正方リストであり, セグメントラベル p, q のセグメントが原パ



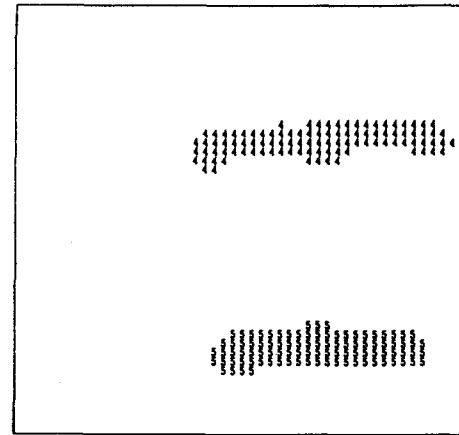
(a)



(b)



(c)



(d)

図3-2 セグメントラベル付けの実例

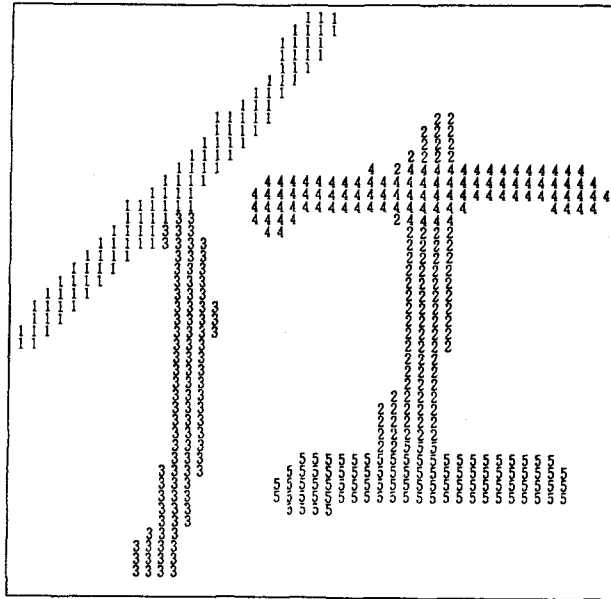


図3-3 重ね合せパターン

ターンで接続していれば $a_{pq}=a_{qp}=1$, 接続していなければ $a_{pq}=a_{qp}=0$ とする。但し, a_{pq} は L_S において p 行 q 列の値を示す。

接続しているか否かの判定は以下のように行う。ラベル付きセグメントが存在する4つの方向画面を1つに重ね合わせた画面 H

$$H = \{(i, j, h(i, j)) \mid h(i, j) = \max_{1 \leq k \leq 4} f_k(i, j)\} \quad (3-1)$$

但し, $f_k(i, j)$ は第 k 方向画面の点である。

を生成し, 3×3 の論理マスクを操作することにより求める。具体的には対象となる点がセグメントラベル p であるとき8近傍点を調べ, p 以外のセグメントラベル q が存在すれば $a_{pq}=1$ とする。4つの方向画面を1つに重ね合わせた画面 H を図3-3に示す。図3-3のパターンにおいて, L_S は,

$$L_S = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (3-2)$$

で表現される。

3. 5 射影パターン

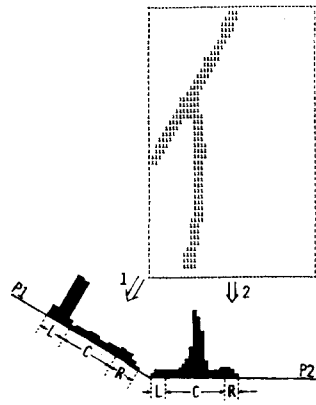
L_S において $a_{pq}=1$ なるセグメントの組 (接続セグメント対 S_{pq} と呼ぶ) , すなわち, 原パターンにおいて互いに接続している2つのセグメントを画面Fに順次, 発生させる. 画面Fに発生させる S_{pq} のパターンは

$$S_{pq} = \{ (i, j, z) \mid \begin{cases} z=1 & \text{if } h(i, j) \in \{p, q\} \\ z=0 & \text{otherwise} \end{cases} \} \quad (3-3)$$

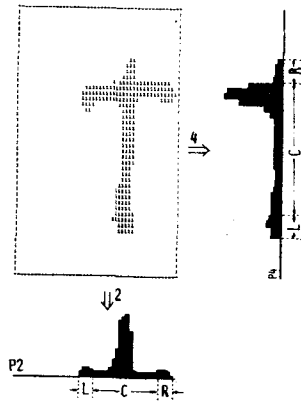
である. ここで $f(i, j)$ はFの画素値, N は画面サイズである.

いま注目しているセグメントに接続しているセグメントの状態は注目セグメントの方向への射影量に反映される. 従って, 接続セグメント対をなす2つのセグメントの方向へ接続セグメント対を射影することにより相互の接続関係を得ることができる. 接続セグメント対 S_{pq} の2つの方向コード k_p, k_q の向きに射影パターンを作成する. 方向コード1, 2, 3, 4に対し, 射影の向きは各々 ↘, ↓, ↙, → とする. 射影パターン $P_k(n), (k=1, 2, 3, 4)$ は

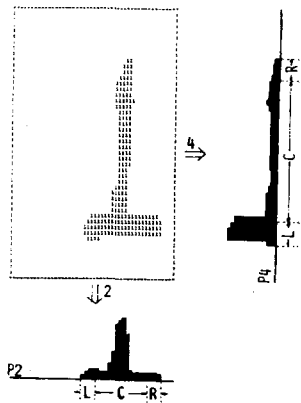
$$P_1(n) = \max \left(\sum_{i+j=2n} f(i, j), \sum_{i+j=2n+1} f(i, j) \right) \\ , (n=1, \dots, N-1) \quad (3-4)$$



(a)



(b)



(c)

図3-4 接続セグメント対と射影パターンの実例

$$P_1(N) = f(N, N) \quad (3-5)$$

$$P_2(n) = \sum_{i=1}^N f(i, n) \quad , (n=1, \dots, N) \quad (3-6)$$

$$P_3(n) = \max \left(\sum_{i-j=2n-N+1} f(i, j), \sum_{i-j=2n-N} f(i, j) \right) \\ , (n=1, \dots, N-1) \quad (3-7)$$

$$P_3(N) = f(N, 1) \quad (3-8)$$

$$P_4(n) = \sum_{j=1}^N f(n, j) \quad , (n=1, \dots, N) \quad (3-9)$$

で定義される。ここで方向コード 1, 3 の向きの射影パターン P_1, P_3 は斜め方向の 2 ラインの射影量の最大値より求める。図 3-4 (a) (b) (c) に接続セグメント対と射影パターンの実例を示す。

3. 6 ピーク検出

射影パターンを $P(n)$ と表すと、

$$\tilde{n} = \{n \mid P(n) \neq 0, (n=1, \dots, N)\} \\ \triangleq \{n_b, \dots, n_e\} \quad (3-10)$$

で定義される座標集合を射影パターン存在範囲 (P 存在範囲と略す) と呼ぶ。接続セグメント対の連結領域数が 1 であることから P 存在範囲数は 1 である。式 (3-4), (3-7) のように斜め方向の射影パターンを斜めの 2 ラインの射影により求めた理由は P 存在範囲数が 1 であることを保証するためである。例えば、対角方向に線幅 1 の直線が存在する場合に逆対角方向の射影パターンを 1 ラインずつ求めると 1, 0, 1, 0 と切れ切れになる不都合が生じるので 2 ラインごとに射影している。

次に存在範囲をさらに3つの範囲, L 範囲 \tilde{n}_L , C 範囲 \tilde{n}_C , R 範囲 \tilde{n}_R に分割する.

P_1 と P_2 に対しては

$$\tilde{n}_L = \{n_b, \dots, n_b + \Delta n\} \quad (3-11)$$

$$\tilde{n}_C = \{n_b + \Delta n + 1, \dots, n_e - \Delta n - 1\} \quad (3-12)$$

$$\tilde{n}_R = \{n_e - \Delta n, \dots, n_e\} \quad (3-13)$$

とし, 一方 P_3 と P_4 に対しては L 範囲と R 範囲が逆になり,

$$\tilde{n}_L = \{n_e - \Delta n, \dots, n_e\} \quad (3-14)$$

$$\tilde{n}_R = \{n_b, \dots, n_b + \Delta n\} \quad (3-15)$$

とする. L, C, R 範囲は存在範囲内で射影する向きに対して各々左側, 中央, 右側であることを示している. 又, Δn は文字線幅, 画面サイズ, 存在範囲に依存する正のパラメータであり, ほぼ文字線幅分の値が妥当である. 図3-4に射影パターンの L, C, R 範囲を示す.

ここで, \mathfrak{X} , ($\mathfrak{X} \in \{L, C, R\}$) 範囲のピーク値および射影量の平均を各々 $P_{\max}(\mathfrak{X})$, $P_{\text{mean}}(\mathfrak{X})$ とすると

$$P_{\max}(\mathfrak{X}) = \max_{n \in \tilde{n}_\mathfrak{X}} P(n) \quad (3-16)$$

$$P_{\text{mean}}(\mathfrak{X}) = \sum_{n \in \tilde{n}_\mathfrak{X}} P(n) / N(\mathfrak{X}) \quad (3-17)$$

により求まる. 但し, $N(\mathfrak{X})$ は \mathfrak{X} 範囲における座標点の数. さて, $P_{\max}(L)$, $P_{\max}(C)$, $P_{\max}(R)$ のうち最大のを $P_{\max}(X)$, 2番目に大きいものを $P_{\max}(Y)$ とし, $P_{\text{mean}}(L)$, $P_{\text{mean}}(C)$, $P_{\text{mean}}(R)$ のうち同様のものを $P_{\text{mean}}(X')$, $P_{\text{mean}}(Y')$, ($P_{\text{mean}}(X') \geq P_{\text{mean}}(Y')$) $X, Y, X', Y' \in \{L, C, R\}$

とする。

ピーク位置の検出及び検出不能を以下の4条件で判断する。

$$\begin{cases} P_{max}(X) - P_{max}(Y) > \lambda \\ X = X' \end{cases} \quad \text{のとき } X(X') \text{ 範囲をピーク位置} \quad (3-18)$$

$$\begin{cases} P_{max}(X) - P_{max}(Y) \leq \lambda \\ P_{mean}(X') - P_{mean}(Y') > \mu \end{cases} \quad \text{のとき } X' \text{ 範囲をピーク位置} \quad (3-19)$$

$$\begin{cases} P_{max}(X) - P_{max}(Y) > \lambda \\ X \neq X' \end{cases} \quad \text{のときピーク検出不能} \quad (3-20)$$

$$\begin{cases} P_{max}(X) - P_{max}(Y) \leq \lambda \\ P_{mean}(X') - P_{mean}(Y') \leq \mu \end{cases} \quad \text{のときピーク検出不能} \quad (3-21)$$

但し、 λ, μ は検出しきい値である。又、 $n_e - n_b + 1 < 3\Delta n$ のときには P 存在範囲を3等分して L, C, R 範囲を決定し、式(3-18)から式(3-21)により判断する。 $n_e - n_b + 1 \leq \Delta n$ のときには検出不能とする。

ピーク座標 n^* は

$$P(n^*) = H_{X(X')} \quad , (n^* \in \tilde{n}) \quad (3-22)$$

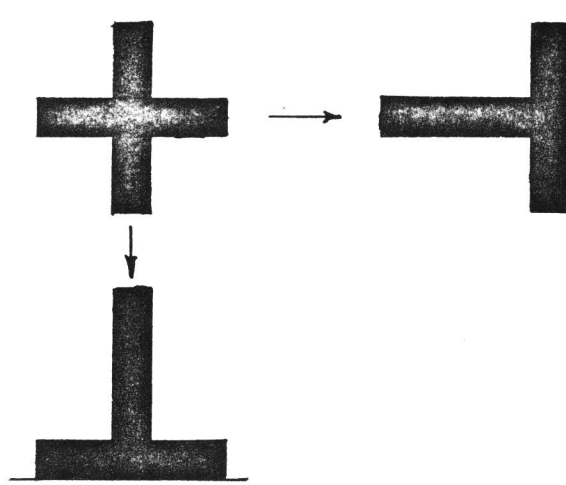
により定める。

3. 7 接続形状と特徴部の抽出

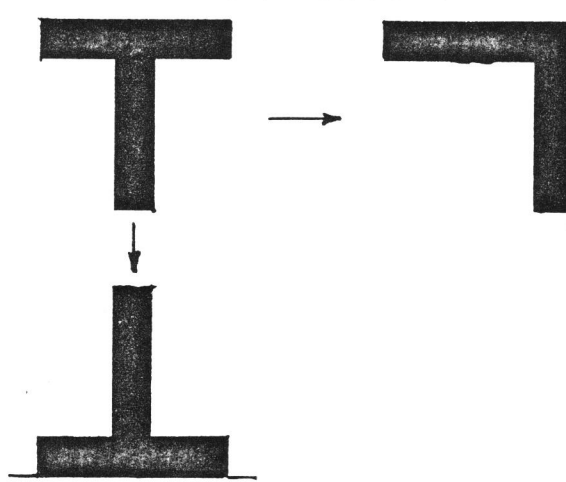
接続セグメント対 S_{pq} は

$$S_{pq} = (k_p, k_q, X_p, X_q, n_p^*, n_q^*) \quad (3-23)$$

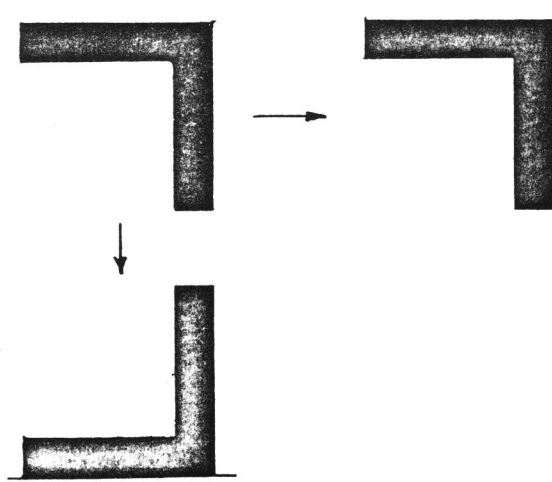
と表現される。ここで $k_{p(q)}$ はセグメントラベル $p(q)$ のセグメントの方向コード、 $X_{p(q)}, n_{p(q)}^*$ は各々方向コード $k_{p(q)}$ の向きの射影パターンのピーク位置とピーク座標である。 S_{pq} の接続形状は式(3-23)の方向コード k_p, k_q とピーク位置 X_p, X_q により一意に定められ、式(3-23)の最初



(a) 交叉部



(b) 分岐部



(c) 屈曲部

図3-5 接続形状と射影パターンの関係

の4つの要素を形状コードと呼ぶ。形状は交叉形状 C_r (十など), 分岐形状 B_r (Tなど), 屈曲形状 C_o (∟など) の3種に分類される。交叉形状は2つのピーク位置がいずれも C 範囲にあり, 分岐形状は2つのピーク位置のうち一方が C 範囲, もう一方が C 範囲以外にある。そして屈曲形状は2つのピーク位置が両方とも C 範囲に含まれない。形状コードの組合せにより交叉形状は6種類, 分岐形状, 屈曲形状はいずれも24種類あり, 接続形状は合計54種類に分類される。図3-5に接続形状と射影パターンの関係を示す。

次に, 射影パターンのピーク座標と特徴部の原パターンにおける座標との関係を示す。射影パターン P_1, P_2, P_3, P_4 のピーク座標 $n_1^*, n_2^*, n_3^*, n_4^*$ と特徴部の座標 (i^*, j^*) との間に

$$i^* + j^* - 1 = 2n_1^* - 1 \quad (3-24)$$

$$j^* = n_2^* \quad (3-25)$$

$$i^* - j^* + u = 2n_3^* - 1 \quad (3-26)$$

$$i^* = n_4^* \quad (3-27)$$

但し, $u = \{ (2n - 1) / 2 \} + 1$, $\{ \cdot \}$ はガウス記号。

なる関係がある。従って, n_p^*, n_q^* に対応する2式を式(3-24)から式(3-27)より選択し, 連立して解くことにより (i^*, j^*) が得られる。最終的に接続セグメント対 S_{pq} は

$$S_{pq} = (k_p, k_q, X_p, X_q, i^*, j^*) \quad (3-28)$$

として表現される。図3-4の各接続セグメント対の形状抽出結果を表3-1に示す。同表における形状モデルは形状コードの示す典型的な接続形状を表す。さらに特徴部抽出の実例を図3-6に示す。同図中, "F" は抽出された特徴部を表す。

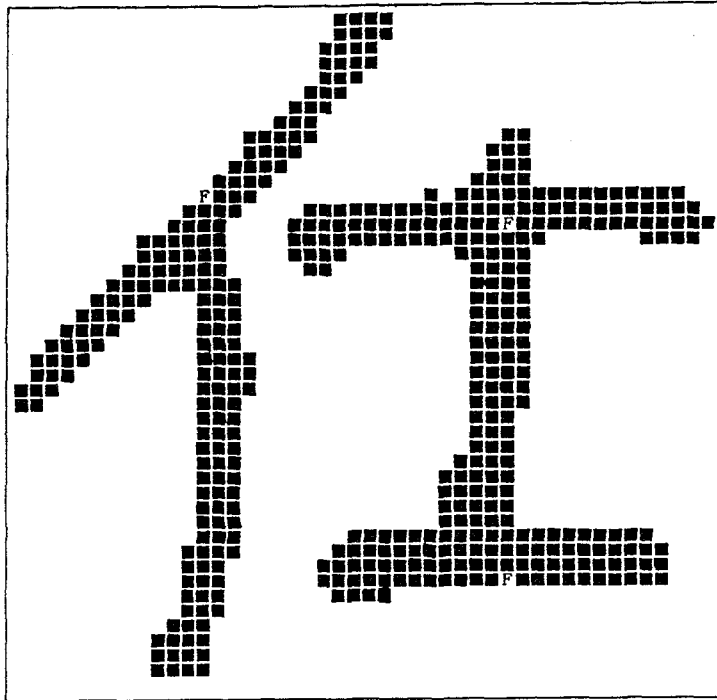


図3-6 特徴部抽出の実例

表3-1 形状抽出結果 (図3-4 (a) (b) (c))

	形状コード	形状モデル	特徴部	座標
(a) S_{13}	1, 2, L, C	イ	分岐	(14, 14)
(b) S_{24}	2, 4, C, C	十	交叉	(16, 34)
(c) S_{25}	2, 4, C, L	⊥	分岐	(40, 34)

3. 8 実験及び検討

本方式の基本的特性を調べるために電総研常用手書き文字データベース(片仮名)⁽²³⁾を用いて抽出実験を行った。又、漢字部分パターン抽出への本方式の可能性を確認するために同手書き教育漢字データベース⁽²⁴⁾を使用した。いずれも2値化、文字サイズの正規化を施している。入力画面サイズは各々22×22、50×50で周囲1画素幅の白枠をつけてある。本実験で用いるシミュレーションプログラムは約1000ステップのFORTRAN言語で作成され、使用計算機システムは阪大大型計算機センターACOS-1000並びに九大大型計算機センターM-200である。

パラメータの1つである Δn (式(3-11)～式(3-15)参照)は文字線幅の値が妥当であるので、目視により片仮名、漢字各々2、4とした。次に、 λ 、 μ (式(3-18)～式(3-21)参照)に関しては、値を大きくすると検出不能になるのが多くなり、反対に値を小さくすると射影パターンの変化に敏感となり過ぎ、誤りが多くなる傾向にある。 λ 、 μ を変化させたときの特性を調べた基礎実験により片仮名、漢字とも $\lambda=3$ 、 $\mu=1$ とした。いずれの場合もセグメントラベル付けの際に4画素以下の連結領域は消去した。

3.8.1 片仮名の形状抽出実験と検討

片仮名は直線から構成されていること、セグメント数が少ないこと、様々な接続形状を有していることなどから本方式の基本的特性を調べるには適当な対象であると考えられる。片仮名のうち実験対象としたカテゴリーは、通常、セグメントの接続が明らかに認められ、接続形状が文字パターン全体に大きな影響を及ぼし、且つ同一形状が1つであると考えられるものである。よって、「シ」「ソ」「ツ」「ニ」「ノ」「ハ」「ミ」「リ」「ン」は除外している。又、「キ」「サ」については同一形状が2箇所に出現すると思われるので詳細な検討は3.8.2節で行う。表3-2に片仮名35カテゴリー50サンプル、

表3-2 カテゴリー別形状抽出結果 (片仮名)

C_a	S_c	F_c	M_a	N_s	C_a	S_c	F_c	M_a	N_s	C_a	S_c	F_c	M_a	N_s	C_a	S_c	F_c	M_a	N_s
ア	1, 4, R, R	C_0	フ	39	コ	2, 4, R, L	C_0	リ	48	ネ	1, 2, L, C	B_T	イ	33	ニ	2, 4, R, R	C_0	リ	43
	1, 2, R, R	C_0	ノ	23		1, 3, L, C	B_T	ク	40		1, 4, R, R	C_0	フ	30		2, 4, C, L	B_T	レ	42
	1, 2, C, L	B_T	フ	19		1, 4, R, R	C_0	フ	35		1, 4, L, C	B_T	ク	12		2, 4, R, R	C_0	リ	50
イ	1, 2, L, C	B_T	イ	38	セ	1, 4, R, R	C_0	フ	33	ヒ	2, 4, L, C	B_T	フ	44	ヨ	2, 4, R, C	B_T	フ	40
	1, 2, L, R	C_0	イ	7		2, 4, C, C	C_T	フ	40		2, 4, L, L	C_0	リ	35		2, 4, R, L	C_0	リ	48
ウ	2, 4, R, R	C_0	リ	40	タ	2, 4, L, L	C_0	リ	42	フ	2, 4, R, R	C_0	リ	45	ラ	2, 4, R, R	C_0	リ	23
	2, 4, L, R	C_0	リ	22		1, 4, L, R	C_0	リ	26		1, 2, R, R	C_0	リ	30		1, 4, R, R	C_0	フ	21
エ	1, 2, R, R	C_0	リ	20	チ	2, 4, R, R	C_0	リ	25	ヘ	1, 3, L, R	C_0	リ	45	ル	1, 2, R, L	C_0	レ	49
	2, 4, C, R	B_T	フ	48		1, 4, R, R	C_0	フ	14		ホ	2, 4, C, C	C_T	+		45	レ	1, 2, R, L	C_0
オ	2, 4, C, L	B_T	レ	48	テ	1, 3, R, C	B_T	フ	33	マ	1, 4, R, R	C_0	フ	33	ロ	2, 4, L, R	C_0	リ	43
	2, 4, C, C	C_T	+	43		1, 2, L, C	B_T	イ	27		2, 4, R, R	C_0	リ	10		2, 4, R, R	C_0	リ	49
カ	1, 2, C, R	B_T	イ	34	ト	2, 4, C, C	C_T	+	45	ム	1, 3, C, L	B_T	フ	40	ワ	2, 4, L, L	C_0	リ	43
	2, 4, R, R	C_0	リ	45		2, 4, C, R	B_T	フ	41		1, 4, L, L	C_0	レ	15		2, 4, R, L	C_0	リ	43
ク	2, 4, C, C	C_T	+	33	ナ	1, 2, R, R	C_0	リ	23	メ	2, 4, L, L	C_0	リ	12	ヲ	2, 4, R, R	C_0	リ	46
	2, 4, R, R	C_0	リ	32		2, 4, L, C	B_T	フ	29		3, 4, R, C	B_T	ク	24		2, 4, L, R	C_0	リ	42
ケ	1, 4, R, R	C_0	フ	13	ヌ	2, 3, L, C	B_T	ク	17	モ	1, 3, C, C	C_T	×	36	ヲ	1, 2, R, R	C_0	リ	32
	1, 4, L, R	C_0	ク	28		2, 4, C, C	C_T	+	45		2, 4, C, C	C_T	+	40		2, 4, R, R	C_0	リ	45
	2, 4, L, C	B_T	フ	21		1, 2, R, R	C_0	リ	25		2, 4, C, R	B_T	フ	42		2, 4, R, C	B_T	フ	39
コ	1, 4, L, L	C_0	レ	16	ヤ	1, 4, R, R	C_0	フ	26	ヤ	2, 4, L, L	C_0	リ	40		1, 2, R, R	C_0	リ	31
	2, 4, C, R	B_T	フ	39		2, 4, R, R	C_0	リ	16		2, 4, C, C	C_T	+	30					
	2, 4, R, R	C_0	リ	49		1, 3, C, C	C_T	×	31		1, 4, R, R	C_0	フ	25					

合計1750サンプルについてのカテゴリー別形状抽出結果を示す。同表にはカテゴリー名 C_a ，形状コード S_c ，特徴部の種類 F_c ，形状モデル M_o ，出現頻度の高い形状の50サンプル中の出現数 N_s を記す。

実験結果より大部分が標準的と考えられる形状，例えば「エ」ならば「ㄣ」と「ㄥ」，「ス」ならば「フ」と「ノ」などに合致し，比較的良好な結果が得られている。出現頻度の低い形状は手書き変動の大きい部分であると考えられる。又，同一形状でありながら2つ以上の形状として抽出されたものも多い。たとえば，「ク」の右上の形状が「フ」あるいは「ㄣ」として抽出されたり，「ト」の形状が「卜」あるいは「ト」として抽出されたりする場合であり，類似形状を同一視するような処理も必要であろう。結果として一部のカテゴリーで満足な形状抽出がなされていないが，これは本抽出アルゴリズムによるところよりもその前段階である方向セグメント抽出によるところが大きい。なぜなら手書き文字に対して完全なセグメント抽出は困難であり，方向精度の向上，曲線部の処理などに改善の余地がある。

3.8.2 細線化法との比較

細線化法により特徴点抽出を行う際に一番問題となるのは変形交点の処理である。変形交点は交点であるべき箇所が2つの分岐点となっているもので細線化雑音の最も顕著な部分の1つである。分岐点の抽出は論理マスクにより比較的容易であると考えられ，「キ」「サ」を対象に交叉部（交点）抽出において比較を行った。又，細線化を用いて屈曲点を安定に抽出する決定的なアルゴリズムは見いだされていないので屈曲点抽出における本方式との比較は行わなかった。ここでは8連結性の保持されるHilditchの細線化アルゴリズム⁽²⁵⁾を用い，3×3の論理マスクにより三叉点と直交点を区別し，三叉点の変形交点を構成する点の1つであるか，あるいは単独のものであるかを距離により判別し，交点数を求める。本手法と細線化法を比較して50サンプル中の交点数の分布を

表 3-3 に示す。尚、本手法において検出不能の場合は交点数 0, 1 に含めている。

「キ」では細線化法が、「サ」では本手法が各々優れている。特に「サ」においては変形交点の処理がうまく行われなかったり、小さな枝の消失により安定に抽出できなかつたものと思われる。又、「キ」において本手法の結果が悪い理由として、縦のセグメントが斜めに傾く例が多く、射影パターンのピーク位置が明確に現れず、検出不能の点が多かつたことが挙げられる。

3.8.1 節, 3.8.2 節の結果を総合すると本方式は射影によりマクロ的に概形をとらえているので手書き変動には比較的安定に形状抽出されるが、反面、微細な形状変化には鈍感である。又、斜めのセグメントの方向と射影の方向が一致しないことにより忠実に形状抽出されない場合もあり、これらの対策が必要である。しかしながら、特徴部抽出方式としてはアルゴリズムが簡単であるなどの利点を有し、有効性が認められる。

表 3-3 細線化法との比較

交点数	キ		サ	
	本方式	細線化法	本方式	細線化法
0	4	0	5	10
1	7	8	10	21
2	39	42	35	19

3.8.3 漢字部首の形状抽出実験と検討

「イ」「人」「木」「サ」「口」「土」の各部首の形状抽出実験を行い、手書き漢字パターンから部分パターンを抽出する方式の基礎的アプローチとして本方式の有効性を確かめた。多くの部首のうちセグメントの接続が比較的保証され、様々な方向のセグメントを有するものとの考察から上に挙げた代表的な6部首を選択した。各部首を含むカテゴリー名は部首以外の複雑度が低いものから高いものまでほぼ一樣になるように決定した。表3-4に実験カテゴリー名を示す。各カテゴリー10サンプル、ETL8の第1から第10データセットの文字550サンプルを用いている。

まず、入力パターンに対しすべての接続セグメント対の形状抽出を行ったのち、各部首を構成していると考えられるセグメントの方向コード並びに特徴部の座標により各部首の形状を決定する。表3-5に部首別の形状抽出結果を示す。同表でand はセグメント対どうしが互いに接続していることを表し、and で結ばれているもの全体で1つの形状となる。又、 $(1, 2, L, \{C, R\})$ 、

表3-4 実験カテゴリー名

部首	カテゴリー名	数
イ	代保億仕住健倍使像俗	100
人	会合倉命舎念今食全令	100
木	極林機検材橋板標村横	100
サ	芸菜花蔵茶荷苦落暮英	100
口	囟園固団回国因圃	80
土	増境均塩坂場地	70

$([i_1, i_2], [j_1, j_2])$ は座標値範囲 $i_1 \leq i \leq i_2, j_1 \leq j \leq j_2$ に $(1, 2, L, R)$ あるいは $(1, 2, L, C)$ の形状が存在していることを表す。但し、座標値範囲は出現頻度の最も高いものについてのみ示す。

表 3-5 において “*” を付けた形状を標準形状と考えると 91.8% が標準形状と一致し、ほぼ正確な形状抽出がなされている。部首別に見ると「イ」「入」はセグメントの分離したものが生じ「木」ではセグメントの接続部が 1 箇所集中しているため形状の変化が多く見られた。又、「廿」では縦のセグメントが短いために射影パターンのピーク位置が明確に表れない。これからも微細な形状変化には鈍感であることが分かり、詳細な形状抽出処理が望まれる。「口」「土」に関しては極めて良好な結果を得た。

ここで形状抽出の対象とした部首はセグメントの接続構造を持つものであり、それ以外の構造の部首、例えば、「シ」などには本方式は適用できないので十分な実験結果とは言い難い。しかしながら、接続セグメントの形状を抽出、記述することにより高次の特徴抽出が実現されている。すなわちセグメント単独の情報ではなくセグメント間の関係を記述し、部分パターンの構造を明らかにしている。又、セグメントの接続関係は不安定であるという報告⁽³²⁾もあるが本方式の場合は接続条件を定義しなおすことにより若干の文字線の離れには対処できると考えられる。

3. 9 結言

本章では、部分パターン抽出への基礎的アプローチとして方向セグメントの接続形状と特徴部の抽出方式を与えた。手書き片仮名を対象に本方式の基本的特性を実験した結果、比較的忠実に形状抽出されていることが確認された。さらに漢字部首の形状抽出においても 91.8% が標準形状に一致し、有効性が認められ、部分パターン抽出への応用の可能性が示唆された。

本方式の特徴は、

(1)交叉部, 分岐部, 屈曲部の3つの特徴部が同一アルゴリズムのもとに抽出可能である。

(2)マクロ的処理である射影を用いていることにより手書き変動に対しても比較的安定である。

(3)方向セグメントを基本構成単位として文字パターンを記述する有効な手段となりうる。

などが挙げられる。一方, 問題点として

(1)斜めのセグメントを含む接続形状が忠実に抽出されない場合がある。これはセグメントの方向と射影の方向が一致していないことによりピークが明確に生じないからである。

(2)微細な形状変化には鈍感である。手書き変動に対する安定性と二者択一なものであり, これらはトレードオフ関係にある。

などが指摘される。

第4章 矩形セグメント整合法

4.1 緒言

第2章において特徴抽出手法としての方向セグメント抽出のアルゴリズムと基本的抽出特性について述べた。本章では方向セグメント抽出を用いて矩形セグメント整合法なる手書き漢字認識系を構成し、認識系における方向セグメント抽出の貢献度及び矩形セグメント整合法の認識能力、適用範囲について考察する。まず、手書き漢字認識方式の概説をし、本方式との関連性を説明する。次に本認識方式の詳細並びに実験結果と検討を示す。

4.2 手書き漢字認識における矩形セグメント整合法の位置付け

従来、文字認識手法は手書き文字に適するとされる構造解析法と印刷文字に適するとされる相関法に分類されてきた。前者は文字がいかなる特徴からどのように構成されているかを解析し、構造を記述することにより認識する手法である。特徴抽出の際に文字のどの部分に注目するかにより線解析、輪郭解析、背景解析に分けられ、各々文字の黒地、黒地と白地の境界、白地に注目するものである。構造解析法は文字線の切れや接触などには弱い、概してかなりの変形にも強いとされている。しかしながら、大量のデータの解析の必要があり、手書き漢字のような量的にも質的にも問題点の多い対象への発展性や拡張性には課題もあると考えられる。一方、相関法は2次元的重ね合せにより認識する手法で最も基本的形態を持つものといえよう。相関法は直観との対応もよく、2次元処理に適し、さらに標準パターンの作成、更新が容易であるなどの利点を持つので拡張性には富んでいる。ただ、文字部の点であるか背景部の点であるかという位置的な情報のみを持つ場合の最も単純な重ね合せにおいては1次元系列の相関と何ら変わりはなく、単なる点整合の和となるため位置ずれには極めて弱い。

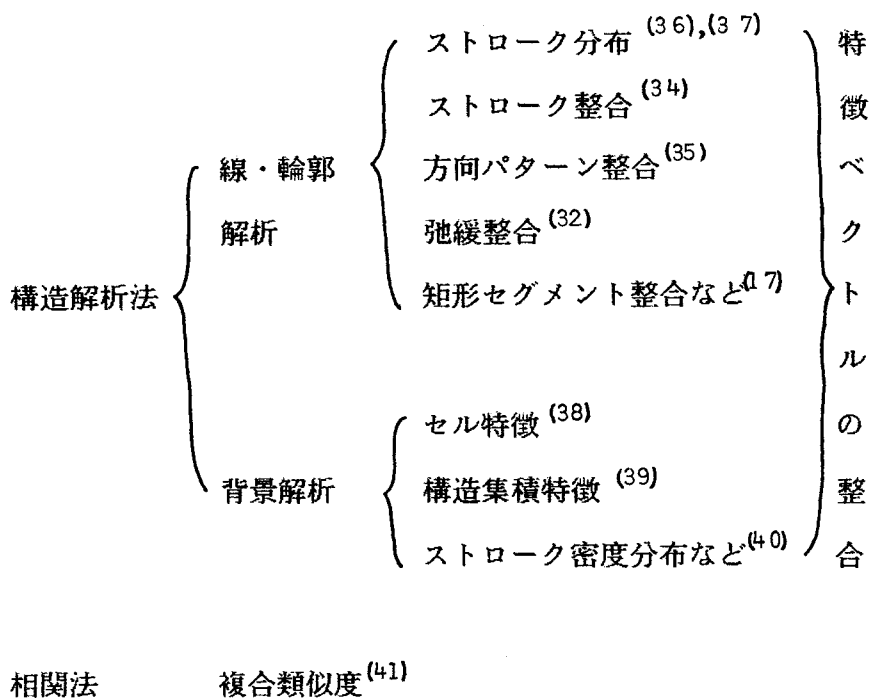


図4-1 手書き漢字認識方式の分類

従って、最も困難な認識対象である手書き漢字に対処するために構造解析法と関連法の得失を考え、両者の融合手法が現在の研究の主流である。思想的には、2次元パターンとして文字をとらえ、構造解析法の基礎をなす特徴抽出により得られる特徴ベクトル場の相関である⁽⁹⁾。すなわち、位置情報のみを持つスカラ量での点整合とは異なり、文字の2次元特性から得られる特徴（方向性、位相構造など）を各点に反映させ、ベクトル量で表現し、整合を行うものである。この概念は安田の線素整合法⁽³³⁾により最初に導入されたものであり、図4-1に手書き漢字認識方式の分類を示したもののうち大部分は特徴ベクトル場の相関と考えられる。何れも漢字ストロークの方向性特徴や背景部に周囲の文字線の状況を反映させる位相構造特徴などを利用している。一般に文字パターンからストローク、輪郭線、部分パターン⁽⁴²⁾などを抽出し、辞書と対応付けする手法は個別・詳細認識に適し、ストローク分布や概形特徴に着目した

り、背景構造解析はカテゴリー数を減少させる大分類に適しているとされている⁽⁴³⁾。

本章で述べる矩形セグメント整合法 (Rectangular Segment Matching ; R S M法と略記する) も線構造解析における方向性特徴に着目したものであるが、さらに漢字認識の量的問題を克服するべく、処理の高速化に重点を置いているのが特色である。又、認識能力を向上させるために相異なる分類特性を有する複数の特徴を組み合わせた方式が多い^{(36),(37)}。高性能手書きOCR開発の実用的側面から認識システムに様々な特徴を組み入れるのも一手段であろうが、認識技術を確立するためには各特徴ごとの分類特性を調べ、適用範囲を明確に示すことも重要であると考え、そこで本方式では従来方式とはやや立場を変え、漢字パターンを構成するセグメントの方向性という単一の特徴に着目し、方向セグメントに関する情報だけでどの程度の認識能力を有するかを検討する。

4. 3 矩形セグメント整合法の概要

図4-2にRSM法のフローチャートを示す。入力パターン (Input Pattern) 2値パターン (文字部1, 背景部0) とする。前処理 (Pre-Processing) では雑音除去と正規化を施す。雑音除去は孤立した2画素以下の黒点集合を除去する処理であり、正規化は入力パターンの文字部を囲む最大正方形を相似形のまま一定サイズの正方画面に変換する処理である。

方向セグメント抽出 (Direction Segments Extraction) では水平 (第4方向とする)、垂直 (第2方向とする)、対角 (第3方向とする)、逆対角 (第1方向とする) の4つの方向の方向セグメント抽出を行う。ここで第 k 方向 ($k=1, 2, 3, 4$) の方向セグメントは k 方向のDS (Direction Segment) と略記する。次に抽出されたDSを矩形表現し、位置を示す4つの座標値にデータ圧縮を行う。特徴抽出を次元の低減化と考えた場合、矩形表現は文字パターンそのものをDSを基にデータ圧縮するものであり、文字パターンの概形は

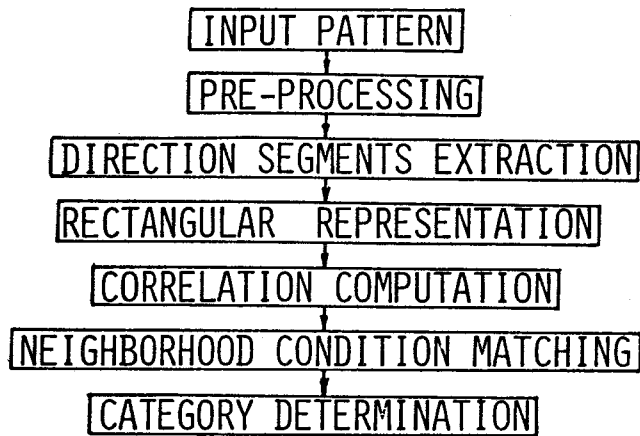


図4-2 RSM法の概要

ほぼ保存され、文字パターンの新しい表現法でもある。矩形表現により処理単位を画素単位からセグメント単位に変換し、処理の高速化を図る。又、各セグメントが独立に取り扱えるため以下の各種応用が容易である。相関法で有効であるが操作や太め操作及び辞書セグメントとの相関計算が簡単な数値計算により得られる⁽⁴⁴⁾。さらに矩形セグメントの周囲における各方向の矩形セグメントの分布を調べ、位置的に対応していると考えられるセグメントとの近傍状態⁽⁴⁵⁾整合⁽⁴⁵⁾を行う。即ち、矩形セグメントに他の矩形セグメントとの相対的位置情報を付加することにより文字パターンの局所的構造を反映させようとするものである。

次節より矩形表現から順に説明する。

4. 4 矩形表現

水平・垂直方向（2・4方向）のDSの表現には通常用いられるIJ座標に

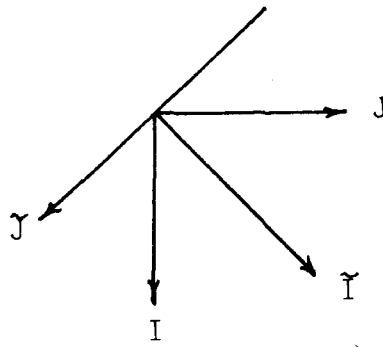


図 4 - 3 座標系

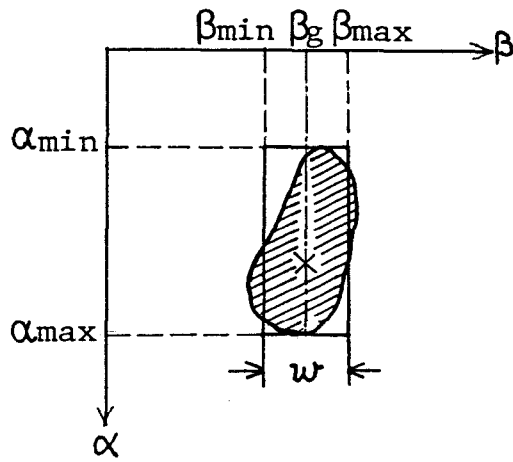


図 4 - 4 矩形表現

表 4 - 1 各方向の DS の長辺方向と短辺方向

	逆対角	垂 直	対 角	水 平
長辺方向 (α)	J	I	I	J
短辺方向 (β)	I	J	J	I

より行い、斜め方向（1・3方向）のDSの表現には図4-3に示すような
[I]座標により行う。ここで、

$$\tilde{I} = I + J \quad (4-1)$$

$$\tilde{J} = I - J \quad (4-2)$$

である。但し、 \tilde{J} 座標は負値もとりうる。

いま、 k 方向のDSは一般には k 方向に長い形状であると考えられるので表
4-1のように各方向のDSの長辺方向 α 軸、短辺方向 β 軸を定義する。

各DSについて表4-1の座標軸を用いて面積 M 、長辺方向の最大、最小
座標 $\alpha_{min}, \alpha_{max}$ 、重心座標 (α_g, β_g) を測定し、それらを基に各DSを矩形
で近似表現 (Rectangular Representation) する。このときDSの長辺方向の
長さは保存し、矩形表現されたDS (Rectangular Segment 以下RSと略記
する) の短辺方向の最大、最小座標 β_{min}, β_{max} を次式で定める。

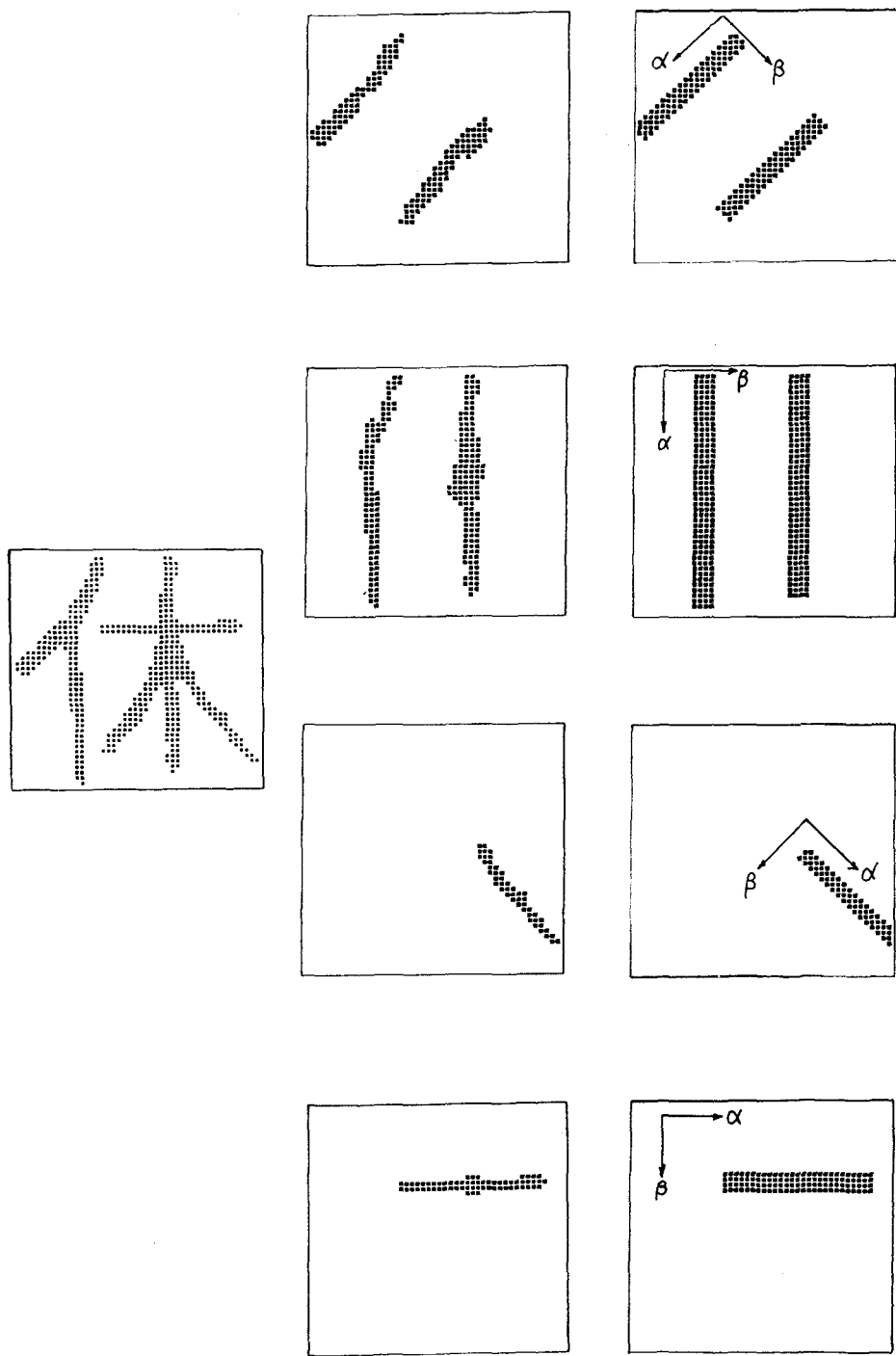
$$\beta_{max} = \beta_g + w/2 \quad (4-3)$$

$$\beta_{min} = \beta_g - w/2 \quad (4-4)$$

$$w = M / (\alpha_{max} - \alpha_{min}) \quad (4-5)$$

但し、 w はRSの短辺方向の幅である。図4-4に矩形表現の例を示す。

このようにRSは各々 $(\alpha_{min}, \alpha_{max}, \beta_{min}, \beta_{max})$ なる位置を示す4つの座標
値により表現される。矩形表現により1本のDSは4つの数値の組に情報圧縮
され、処理単位が画素からセグメントに変換されたことに相当する。これ故に
各セグメントが独立に取り扱え、高速な処理が可能となる。又、わずかな屈曲
やひずみを補正する効果も期待している。図4-5にRSの例を示すが、同図
(a) (b) (c)は各々漢字パターン、方向セグメント(DS)、矩形セグ
メント(RS)を表す。



(a)

(b)

(c)

図4-5 漢字パターン，方向セグメント，矩形セグメントの実例

4. 5 相関計算

入力パターンのRSを $(\alpha_{min}, \alpha_{max}, \beta_{min}, \beta_{max})$, 辞書のRSを $(\alpha_{min}^*, \alpha_{max}^*, \beta_{min}^*, \beta_{max}^*)$ とするとき, 2つのRS間の相関 c は,

$$c = \{\max(E_\alpha, 0)\} \times \{\max(E_\beta, 0)\} \quad (4-6)$$

$$E_\alpha \triangleq \min(\alpha_{max}, \alpha_{max}^*) - \max(\alpha_{min}, \alpha_{min}^*) \quad (4-7)$$

$$E_\beta \triangleq \min(\beta_{max}, \beta_{max}^*) - \max(\beta_{min}, \beta_{min}^*) \quad (4-8)$$

で得られる。図4-6に相関の例を示す。斜線部の面積が c に相当し, E_α , E_β は各々斜線部の長辺, 短辺を表す。

ここで相関を求める際, 手書きひずみの1つであるセグメントの位置変動を吸収するために入力パターンのRSにずらし幅を与え, 辞書との最大相関を求める。ずらし幅を長辺方向に $\Delta\alpha$, 短辺方向に $\Delta\beta$ としたときのRSのずらし操作は,

$$(\alpha_{min}, \alpha_{max}, \beta_{min}, \beta_{max}) \rightarrow (\alpha_{min} + \Delta\alpha, \alpha_{max} + \Delta\alpha, \beta_{min} + \Delta\beta, \beta_{max} + \Delta\beta) \quad (4-9)$$

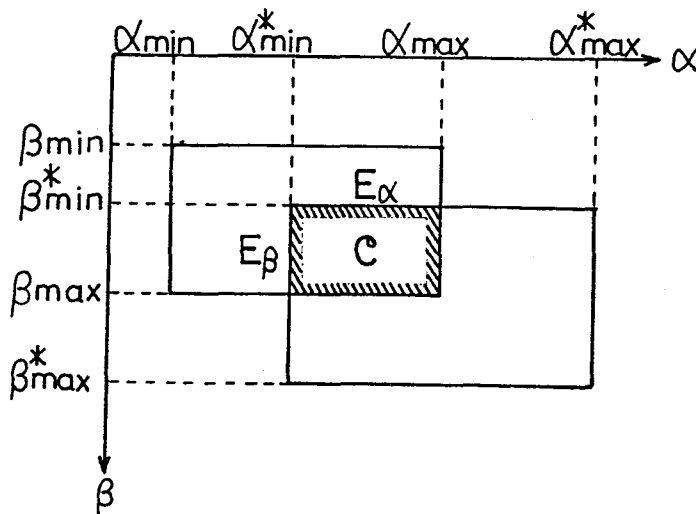


図4-6 相関

のように座標値の加減算により容易に実現でき、画素単位のずらしに比べて大幅な処理時間の短縮が可能である。又、太め操作も同様に実現可能である。

式(4-6)の相関計算(Correlation Computation)を用いて入力パターンと辞書の全てのRS間の類似度(セグメント類似度と呼ぶ) S_p を

$$S_p = \sum_{k=1}^4 c_k / \sqrt{A \cdot A^*} \quad (4-10)$$

で定義する。但し、 c_k は k 方向($k=1, 2, 3, 4$)のRSの相関の総和、 A, A^* は各々入力パターン、辞書の全RSの面積の総和である。セグメント類似度は単純類似度にはほぼ近い意味を持つ。

4.6 近傍状態整合

4.6.1 近傍状態の定義

図4-7のようにRSの周囲に位置する8つの矩形領域とRS自身の領域に番号を付け、近傍領域とする。あるRSに着目した場合、近傍領域における方向別の他のRSの分布状況を調べ、各近傍領域に存在する方向別のRSの本数を近傍状態と定義する。複数の領域にまたがるRSが存在する場合は領域毎に独立に計数する。各RSの近傍状態は近傍状態行列 N^b なる 4×9 の行列で表し、 k 行 l 列の値は第 l 近傍領域における k 方向のRSの本数である。全ての

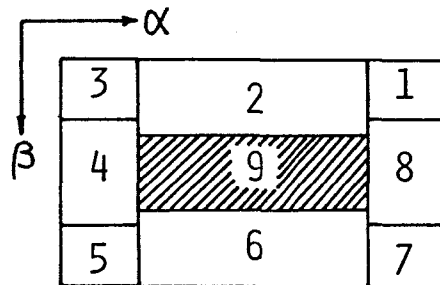


図4-7 近傍領域

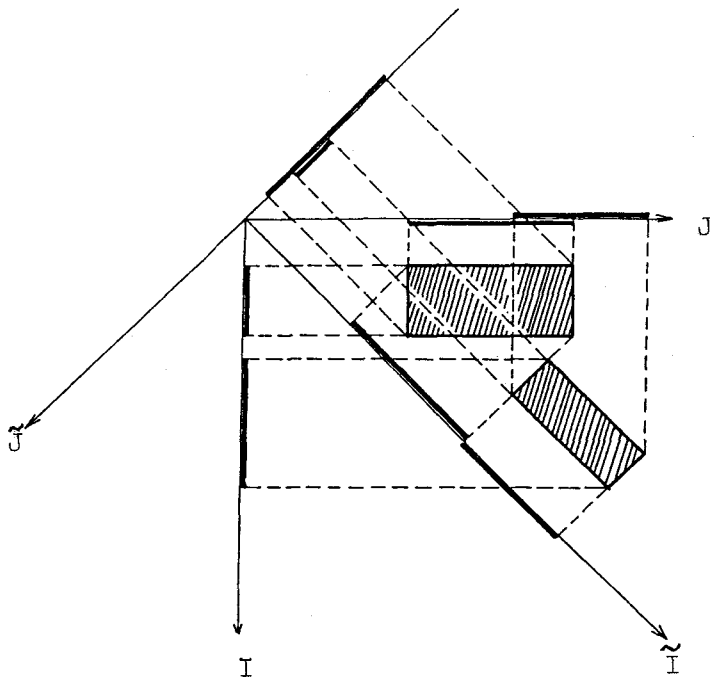


図4-8 各軸への投影と接続関係

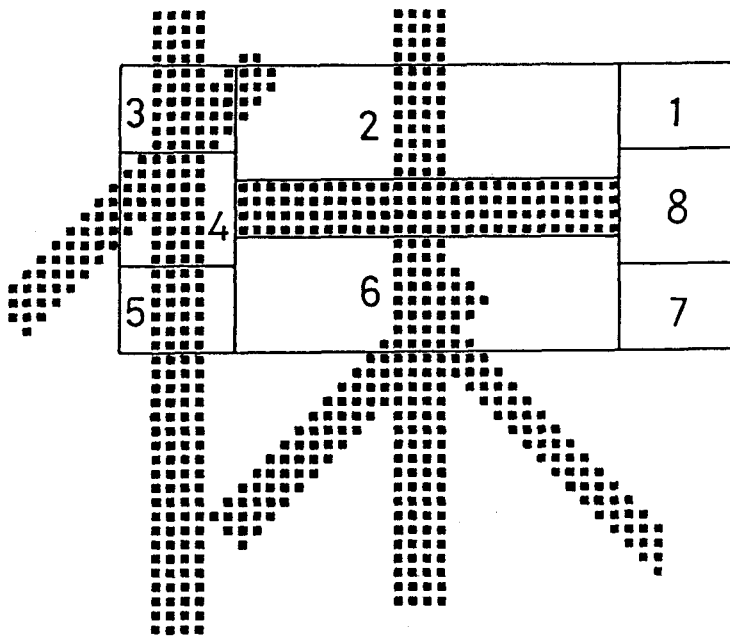


図4-9 近傍領域の実例

RSについて近傍状態行列を作成する。

さて、2つのRS間において、 $I \setminus J \setminus J$ 軸上への投影の重なりが全ての軸上に存在する場合のみ接続関係がある。図4-8に接続していない場合の各軸への投影を示す。いま、2つのRSで一方のRSを図4-7の近傍領域の1つと考えれば、RSと近傍領域の接続関係を調べることにより近傍状態を検出することができる。図4-9にRSより構成された漢字パターン「休」の水平方向のRSについての近傍領域の実例を示す。又、同じRSの近傍状態行列 N^b は次式となる。

$$N^b = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4-11)$$

4.6.2 セグメント間の対応付け

近傍状態整合 (Neighborhood Condition Matching) には入力パターンと辞書のRS間の適当な対応付けが必要である。ここでは相関計算による位置的対応付けを行う。すなわち、入力パターンと辞書の同一方向におけるRS間の相関を調べ (ずらし操作を含む) , 次式が成り立つときに位置的に対応しているとする。

$$c / \min(a, a^*) > t \quad (0 < t \leq 1) \quad (4-12)$$

但し、 c は両RS間の相関値、 a, a^* は各々注目している入力パターン、辞書のRSの面積、 t はしきい値 (面積の割合) である。

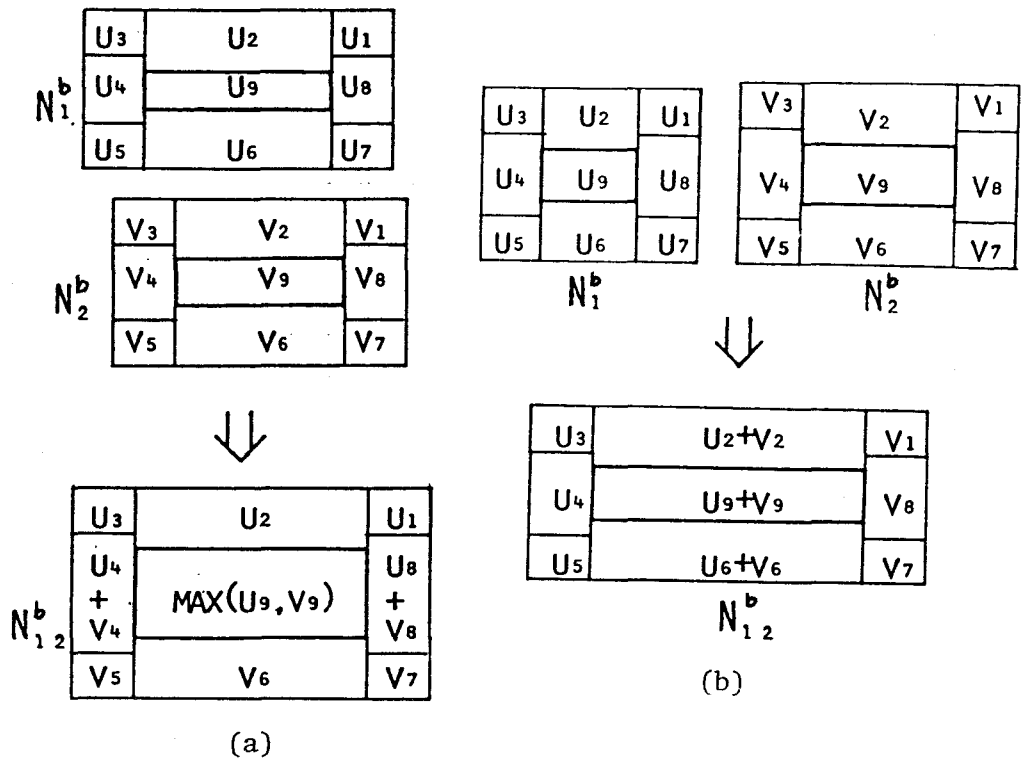


図4-10 近傍状態の統合

4.6.3 近傍状態の統合

手書き文字パターンより得られるRSでは切れや不必要な接触が多く見られる。従って、入力パターンのRSから見て1対1の対応のみならず1対複数あるいは複数対1の対応付けが生じると考えられる。このような場合、対応付けに関係する複数のRSの近傍状態を統合し、1対1の対応に帰着させる。

本方式では1対2及び2対1の対応の場合にのみ近傍状態の統合を行い、入力パターンと辞書において対応に関与しているRSの総数が4以上の場合には近傍状態の整合は行わない。1対2（2対1）の対応に伴う近傍状態の統合は図4-10（a）上下（b）左右の2通りが考えられる。2つのRSの近傍状態行列を N_1^b, N_2^b とし、 U_l, V_l を第 l 近傍領域での各々の近傍状態を表す 4×1 の列ベクトルとすると、

$$N_1^{\phi} = (U_1, U_2, U_3, U_4, U_5, U_6, U_7, U_8, U_9) \quad (4-13)$$

$$N_2^{\phi} = (V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9) \quad (4-14)$$

であり、 N_1^{ϕ} と N_2^{ϕ} を統合した近傍状態 N_{12}^{ϕ} は以下の2つの場合に対して、

(a) 上下の場合

$$N_{12}^{\phi} = (U_1, U_2, U_3, U_4+V_4, V_5, V_6, V_7, U_8+V_8, \max(U_9, V_9)) \quad (4-15)$$

(b) 左右の場合

$$N_{12}^{\phi} = (V_1, U_2+V_2, U_3, U_4, U_5, U_6+V_6, V_7, V_8, U_9+V_9) \quad (4-16)$$

とする。

4.6.4 近傍状態類似度

入力パターンと辞書において1対1の対応ないしは近傍状態の統合により1対1の対応に変換されたRS間の近傍状態 N_1^{ϕ}, N_2^{ϕ} の類似度 r_N は $N_1^{\phi} = (U_{kl}), N_2^{\phi} = (V_{kl})$ とするとき、次式で定義される。

$$r_N = \frac{1}{1 + (\gamma d_N)^2} \quad (\gamma = \text{Const}) \quad (4-17)$$

$$d_N = \frac{1}{36} \sum_{k=1}^4 \sum_{l=1}^9 |U_{kl} - V_{kl}| \quad (4-18)$$

次に、全てのRSにおける近傍状態類似度 S_N は、

$$S_N = W_C \times D_N \quad (4-19)$$

$$W_C \triangleq \frac{\hat{m} + \hat{m}^*}{m + m^*} \quad (4-20)$$

$$D_N \triangleq \frac{1}{M_C} \sum_{M_C} r_N \quad (4-21)$$

で定義される。ここで M_C は対応の総数（統合による対応も1個とし、1対2・2対1を越える対応は考慮しない）である。又、 W_C は対応関与率と呼ぶ重み関数であり、 m, m^* は各々入力パターン、辞書のRSの総数、 \hat{m}, \hat{m}^* は各々入力パターン、辞書における対応の付いたRSの総数である。 W_C は対応付けに関与したRSの数の割合を示し、対応付けされなかったRSを含むすべてのRSについて D_N を平均させる重み関数である。すなわち、対応付けされなかったRSの近傍状態の類似度を0、対応付けされたが1対2あるいは2対1を越える対応であったため近傍状態整合が行われなかった近傍状態の類似度を D_N に等しくすることにはほぼ相当する。

4. 7 カテゴリー決定

入力パターンのカテゴリーはセグメント類似度と近傍状態類似度の組合せで決定する（Category Determination）。カテゴリーごとに2つの類似度を計算し、類似度の最大となる辞書のカテゴリーを入力パターンのカテゴリーとする。

4. 8 認識実験及び検討

RSM法の分類特性を調べるために電子技術総合研究所編集の手書き教育漢字データベースETL8⁽²⁴⁾中の漢字881字種を用いて認識実験を行った。ETL8には各カテゴリー当たり160サンプルあり、最初から順に各カテゴリーについて1サンプルを取り出した881サンプルのデータ集合をデータセットと呼び⁽³⁵⁾、第1から第160データセットまで存在する。本実験では第1から第20データセットまでの17,620サンプルを対象とした。オリジナルデータは128×127の画面サイズで濃淡レベルは16レベルであるが、ここでは

60×60にデータ圧縮を施し、固定しきい値（しきい値2）により2値化した。又、48×48のサイズに大きさの正規化を行っている。尚、使用計算機は大阪大学大型計算機センターACOS-1000（15MIPS）であり、使用言語はアセンブラ（データ入力段階）とFORTRAN（処理段階）である。

4.8.1 辞書の作成

第1から第10データセットまでを辞書作成サンプルとし、1カテゴリ当たり1辞書を作成した。入力パターンよりDSを抽出し、各DSを矩形表現する。得られたRSを方向別に10サンプル重ね合わせ、総和をとり、適当なしきい値で2値化を行う。つぎに再び矩形表現を施し、各RSを座標値で表して辞書のRSとする。

一方、カテゴリごとの辞書のRSと設計サンプルのRSとの対応付けを行い、対応付けされた設計サンプルのRSの近傍状態を調べ、サンプル数により平均して辞書の各RSの近傍状態とする。

4.8.2 セグメント数の統計的調査

RSM法において処理単位であるRSの本数が、処理速度や辞書の規模などに大きく影響するので辞書と入力サンプル（第11から第15データセット）について4つの方向別のRSの本数とRSの総本数に関して、各々平均及び分散を調査した。

セグメント数の調査結果を表4-2に示す。同表に示すように辞書と入力パターンでは方向別の本数も総本数もほぼ一致することが分かる。入力パターンのRSでは切れが生じやすいが、辞書では重ね合せ、2値化などにより切れが生じにくく、平均の総本数はやや少なくなっている。又、RSの総本数が約12であることから、文字パターンを矩形表現により平均60次元（1本のRSは方向と4つの座標値の合計5つの要素で表される）に圧縮される。よって、

表4-2 セグメント数の平均と分散

		逆対角	垂 直	対 角	水 平	総本数
辞	平均	2.0	4.0	1.3	3.9	11.2
	分散	1.8	2.9	1.4	1.3	13.0
書	平均	2.1	4.0	1.4	4.4	11.9
	分散	2.3	3.2	1.5	4.4	16.5

表4-3 ずらし幅の比較

短辺方向	長辺方向	変 化
0	0	基 準
±2	±2	+4%
±4	±4	-5%
±2	0	+5%

表4-4 太め幅の比較

短辺方向	長辺方向	変 化
0	0	基 準
1	0	+1%
2	0	+2%
3	0	-2%
2	2	±0%

他の認識方式と比べて比較的コンパクトに辞書構成が成されている。

4.8.3 ずらし幅・太め幅の検討

RSM法の基本的特性を調査し、種々のパラメータを決定するために881字種中の最初の200字種に関して以下の基礎実験を行った。⁽⁴⁴⁾

(i) ずらし幅の比較

RSの形状から考え、長辺方向より短辺方向のずらしの方が相関値に対する寄与が大きい。そこで本方式のずらし操作は、短辺方向に、あるステップ幅で最大幅までずらし、この中で最大相関値を与える位置から、あるステップ幅で長辺方向にずらし、最大相関値を求める。表4-3にずらしなしの場合の認識率を基準としたとき、各々のずらし幅での認識率の変化を示す。ずらし過ぎた場合には合せ過ぎの生じること、長辺方向のずらしは効果のないこと、ステップ幅には影響されないことなどが判明し、以後の実験ではずらし幅を短辺方向に±2とした。

(ii) 太め幅の比較

RSについて短辺方向のみに太めを施した場合、短辺・長辺両方向に太めを施した場合の比較を行った。表4-4に太めなしの場合の認識率を基準にしたとき、各々の太め幅での認識率の変化を示す。同表より各RSは短辺方向に2だけ太め操作を行う。

また、近傍領域の範囲はRSの存在範囲より長辺、短辺方向に±8の範囲とする。次に、セグメント間の対応付けの際のずらし幅は短辺方向に±2とし、式(4-12)のしきい値は $t=0.5$ とする。

4.8.4 近傍状態整合の効果

第11から第15データセットまでの4405文字の未知サンプルを対象として以下の \mathcal{S}_1 、 \mathcal{S}_2 の類似度の組み合わせを基に近傍状態整合の効果調べた。

$$S_1 = S_P \quad (4-22)$$

$$S_2 = S_P + S_N \quad (4-23)$$

S_1 はセグメント類似度のみでカテゴリー決定を行ったもの、 S_2 はセグメント類似度と近傍状態類似度の和によりカテゴリー決定を行ったものである。表 4-5 に各々の累積分類率を示す。

表 4-5 に示すように近傍状態整合により約 7% 認識率が向上し、セグメント類似度を単独で用いるよりもセグメント類似度と近傍状態類似度を組み合わせて用いる方が良好な結果が得られる。セグメント類似度は各 RS を独立にずらし、最大相関を求めることにより得られるが、各 RS の相対的位置関係などが考慮されていない難点がある。一方、近傍状態類似度は各 RS の周囲の状況を反映し、いわば局所的な構造解析を行っていることに相当するのでセグメント類似度とは異なる分類特性を持つと推察される。又、セグメント類似度では考慮されない対応付けられなかった RS の情報が近傍状態類似度により考慮されることも認識率向上の一因といえよう。

以上のように近傍状態整合の効果が認められ、局所的な構造解析による改善が見られた。しかしながら、本方式における近傍状態の定義は、着目する RS

表 4-5 類似度の比較

順位	1	5	10	25	50
S_1	80.2	95.2	97.2	98.8	99.4 (%)
S_2	87.5	97.5	98.6	99.3	99.6 (%)

の周囲における方向別のRSの分布状況を表す。従って、文字パターンを構成するRSの本数を正確に示すものではなく、厳密な意味での構造の記述とは言い難いので近傍状態の記述法の改良も必要と考えている。又、現段階ではセグメント類似度と近傍状態類似度の重みを1対1にしているが、最適な重みを決定することも重要であろう。

4.8.5 分類特性の検討

ETL8の第11から第20データセットの手書き漢字データ8810サンプル（未知サンプル）を対象としてRSM法による認識実験を行った。表4-6にRSM法による累積分類率を示す。但し、第11から第15データセット、第16から第20データセットの平均認識率は各々87.5%、80.5%で、1文字当りの平均処理時間は約1.0秒であった。

各カテゴリについて40%以上同じカテゴリに誤った例を以下の3つに分類して示す。矢印の左側のカテゴリが右側のカテゴリに誤読されたものである。

(1)共通の偏、旁、冠などの部首を持つカテゴリに誤読された例

選→運, 験→險, 際→除, 国→固, 晴→情, 討→計, 得→待, 輪→諭,
 低→体, 視→祖, 識→講, 柱→住, 池→地, 調→講, 覚→党, 間→問,
 婦→婦, 聞→開, 便→使, 規→現, 退→返, 雜→推, 愛→夏

表4-6 RSM法による累積分類率

順位	1	5	10	25	50
分類率	84.0	96.1	97.8	99.1	99.5 (%)

(2)同一の部分形状（部首を除く）を含むカテゴリーに誤読された例

土→上, 大→犬, 犬→大, 州→川, 系→糸, 士→十, 止→上, 止→正

(3)全体形状の類似したカテゴリーに誤読された例

貴→黄, 未→未, 后→信, 可→町, 集→美, 筆→里, 野→輕, 乘→集,

不→下, 我→牧, 魚→負, 要→早, 觀→勤, 積→横, 破→服, 方→力

上記の例のうち(1)は部首以外の部分でセグメントの抽出, 対応付けが不良のために類似セグメント集合を有する字種に誤まったもの, (2)ではセグメントの抽出, 対応付けは良好に行われ, 対応の付いた部分の類似度が高すぎたために残りの部分の情報が反映されなかったもの, (3)は(1), (2)両方の要因により生じたものであると考えられる。全体的に見て, 水平方向のセグメントが密集している文字において誤りが顕著である。これは複数の同一方向のセグメントが接触し, 1つのセグメントと判定されたからであろう。

以上のように, 漢字パターンの方向性に着目する本方式では, 大分類に対して有効であると結論される。現段階では, 個別認識にそのまま適用するのは難しいが, セグメントの方向性特徴という単一の特徴を基にした場合でも比較的良好な認識能力が得られることが確かめられた。

本方式の第1の問題点として辞書の不備が挙げられる。今回の実験では設計サンプルが各カテゴリー当り10サンプルと非常に少なく, 辞書が平均的な特徴を十分示しているとは言えない。従って, 設計サンプルの増加あるいは辞書作成法の改良が望まれる。

第2の問題点は複雑度の高い漢字に対してセグメント間の対応付けが完全ではないことである。たとえば, 弛緩整合法⁽⁸²⁾のように文字パターン全体の線分のつじつま合せによる整合法とは異なり, 本方式はセグメントの位置的な対応に局所的な構造の類似性をサポートさせる方式であり, 近傍状態も局所的な文字構造を正確に記述することができない。手書き変形の多い文字に, 完全なセグメント対応付けアルゴリズムを開発するのは困難であろうが, これらを改善

し、さらに柔軟性の高い整合法を目指したいと考える。

最後に、誤読例が示すように偏、旁などの部分パターンや同一部分形状を含む字種に誤る例が多い。これは逆に考えると、部分パターン抽出によりさらに詳細な識別が可能であることを示唆している。方向セグメントを基に部分パターンを記述し、認識方式に用いることも必要であろう。

4. 9 結言

本章では漢字パターンを構成するセグメントの方向性という単一の特徴に着目した場合どの程度の認識能力を有するかを検討した。矩形セグメント整合法なる手書き漢字認識系を構成し、手書き漢字 8 8 1 字種 8 8 1 0 文字に対して認識率 8 4. 0 %, 第 5 0 位累積分類率 9 9. 5 % を得、本方式が手書き漢字大分類に適用できることを示し、認識系における方向セグメント抽出の有効性を確かめた。

矩形セグメント整合法の特徴として、

- (1)漢字パターン自体を方向セグメントの矩形表現により平均 6 0 次元に圧縮可能であり、文字パターン表現法の新しい試みである。
 - (2)矩形表現により文字認識における相関法及び構造解析法で重要な各種技術の応用が容易に実現される。すなわち、高速な相関計算、ずらし・太め操作、セグメント間の接続状態、近傍状態などが容易に得られる。
 - (3)近傍状態整合により局所的構造解析を行い、セグメント類似度と近傍状態類似度により良好な特性が得られる。
- などが挙げられる。

第 5 章 結 論

文字認識において最も困難な認識対象である手書き漢字に対処するためには漢字の特質を十分考慮して認識系の特徴抽出部、識別部を設計する必要があるとの観点から手書き漢字認識の問題を考察した。本論文では、漢字パターンを構成する要素の直線性に着眼し、漢字パターンから方向セグメントを抽出する方式、方向セグメント抽出を用いた文字パターンの記述方式及び手書き漢字認識方式を提案し、シミュレーション実験により各方式の有効性について論じた。以下に本研究で得られた諸結果をまとめる。

第 2 章では、手書き漢字認識系における有力な特徴抽出の 1 つである方向セグメントの抽出方式を示した。提案方式は、手書き変形にも比較的安定であること、任意角度のセグメントが抽出できること、原パターンが保存されること、アルゴリズムが簡単なことなど数々の利点を持ち、手書き文字における抽出実験の結果、細線化を伴う従来手法より優れた特性を有することを確認した。又、微小セグメントが密集する部分においては忠実な抽出がなされない場合があることも判明した。

第 3 章では、セグメント間の接続形状と交叉部・分岐部・屈曲部などの幾何学的特徴部の抽出方式を示した。この方式は、方向セグメントを基に文字パターンの構造を記述するとともに各特徴部が同一アルゴリズムのもとに抽出されるという他方式にはない特長を持つ。漢字部首の形状抽出実験においては 91.8% が標準形状に一致し、部分パターン抽出への基礎的アプローチとして本方式の有効性を示した。特性上の問題には微細な形状変化に鈍感であることが挙げられる。

第 2 章、第 3 章で述べた方式は手書き漢字認識において有効と考えられる特徴抽出、あるいは別の見方をすれば漢字パターンの表現、記述の新しい試みである。両方式の背景に流れる思想は、手書き変形の影響を排除するためのマク

口的処理の重視といえよう。反面、いずれにおいても図形的に微細な形状である部分、言い換えると、ミクロ的処理を要する部分において問題点を含んでいる。マクロ的処理とミクロ的処理においては、特性の面でトレードオフ関係が存在し、両者を巧妙に併用することが肝要である。提案した2つの方式にも改良の余地が多く、上述のことを考慮してより完全な抽出を実現することが今後の課題である。

第4章では方向セグメント抽出を用いた手書き漢字認識系である矩形セグメント整合法を提案した。矩形セグメント整合法は、漢字認識の量的問題を克服するために、処理単位を画素からセグメントに変換し、相関やずらし操作などの処理の高速化に重点を置いている。又、認識精度の向上を図り、近傍状態整合なる局所的構造解析も導入している。手書き漢字881字種8810文字に対し、第50位累積分類率 99.5%を得、本方式が手書き漢字大分類に適用可能であることを確かめた。

以上のようにセグメント抽出方式が手書き漢字認識系に寄与するところは極めて大きく、本論文で提案された各方式が、有効であることを示した。次に本研究を含む手書き漢字認識技術における今後の課題を挙げる。

まず、第1に柔軟なマッチングを可能にする構造的セグメント整合アルゴリズムの開発である。矩形セグメント整合法の適用範囲を拡大するには不可欠であると考えられる。この概念をくむ1つのパイロットモデルが弛緩マッチングであるが⁽³²⁾、反復計算などを含むので処理時間がかかる難点もある。故に、手書き変形にも十分安定な特徴の選択、抽出及びそれらの適当な記述法などを再び検討しなおすとともに、ある種のあいまいさを許容するマッチング過程に注意を払うべきであろう。

第2に、段階的識別の導入である。これは相異なる特徴で候補字種数を限定していく量的な段階識別のみならず漢字構造に立脚した質的な段階識別の検討も必要である。漢字パターンは上下・左右などに分割される構造のものが多い。

そこで1つの解決策が第3章でも触れた部分パターン抽出を含む認識系である。たとえば、第1段目で部分パターンの部分とマッチングし、抽出した後、第2段目で部分パターン以外の部分とマッチングするなどが推測される。第1段目により字種が限定され、なおかつ構造の複雑さも大きく緩和される一石二鳥の方法といえよう。部分パターン抽出を含む方式の有効性は矩形セグメント整合法の誤読傾向からも明らかであり、今後、本論文で述べた諸方式を総合し、考察することが急務であると考えられる。

第3に、本論文の範囲をややはずれるが、言語知識・文脈情報などの利用である。すなわち、文章としての文字の読み取りである。日常生活においても1文字では判別できない場合でも文章中であれば結構認識できる場合が多い。言語知識あるいはさらに知的処理を用いてパターン認識を試みる研究は端緒に終わったばかりであり、問題点は数多く残されている。しかしながら、これらの技術は、今後、文字認識における重要な分野になると予想される。⁽⁷⁾

現在、文字認識技術や音声認識技術のレベルはまだ低く、人間のレベルに匹敵するなどというには程遠い。文字認識技術に対する期待には、単に読み取り機としてではなく、人間のパターン認識機能の解明の一手段として見るものがある。これまで、文字認識技術は実用的側面に偏してきた傾向がない訳ではない。未だ、パターン認識一般、さらに文字認識においても解明されていないこと、考察しなければならないことは山積している。今後、社会的期待に応えるためにも基礎的研究の積み重ねが益々必要であると考えられる。文字認識技術の将来に、本研究の成果が多少なりとも役立てば筆者の望外の喜びとするところである。

謝 辞

本研究の全過程を通じ、直接懇切なる御指導、御鞭撻を賜った大阪大学工学部通信工学教室手塚慶一教授に心より御礼申し上げます。

大阪大学工学部、同大学院において御指導、御教示賜り、又、本研究に関し、御審査頂いた大阪大学工学部通信工学教室熊谷信昭教授、中西義郎教授、滑川敏彦教授並びに同学産業科学研究所角所収教授に厚く御礼申し上げます。

筆者が愛媛大学に勤務以来、御指導、御支援頂く愛媛大学工学部電子工学教室相原恒博教授に深謝する。

本研究に関し、終始有益な御助言、御討論頂いた大阪大学真田英彦助教授に心から感謝する。

筆者の所属していた手塚研究室の中西暉講師、打浪清一助手をはじめ、研究室の諸氏には御世話になった。さらに、研究分野への最初の手ほどきを与え、御指導下さった岡山理科大学塩野充助教授、本研究に際し、いろいろの御示唆を頂いた愛媛大学村上研二講師の両先生に御礼申し上げたい。特に、大阪大学大学院生の北村義弘氏、正満峰夫氏には、本研究を遂行するにあたり、多大なる御助力を頂いた。又、筆者の研究仲間であった手塚研究室のパターン認識グループ、服部哲郎氏（現、東芝エンジニアリング）、荒木吉雄氏（現、日本IBM）、小山孝司氏（現、松下電器）、田中直樹氏ほか同グループの諸氏には様々の面で御協力を仰いだ。

最後に、本研究の実験において使用させて頂いた手書き文字データベースを作成、編集された電子技術総合研究所の関係各位に心から感謝する。

ここに記して、以上の皆様に深甚なる感謝の意を捧げる。

文 南

- (1) Fukunaga, K. : "Introduction to Statistical Pattern Recognition", Academic Press (1972) .
- (2) Duda, R.O. and Hart, P.E. : "Pattern Classification and Scene Analysis", John Wiley & Sons (1973) .
- (3) 飯島泰蔵 : "パターン認識", コロナ社 (昭48) .
- (4) Pavlidis, T. : "Structural Pattern Recognition", Springer-Verlag (1977) .
- (5) Fu, K.S. : "Syntactic Pattern Recognition and Applications", Prentice-Hall (1982) .
- (6) 中田和男 : "パターン認識とその応用", コロナ社 (昭53) .
- (7) 橋本新一郎 : "文字認識概論", 電気通信協会 (昭57) .
- (8) 長尾 真 : "画像認識論", コロナ社 (昭57) .
- (9) 森 俊二 : "文字・図形認識技術の基礎", エレクトロニクス文庫23, オーム社 (昭58) .
- (10) 白井良明 : "コンピュータビジョン", 昭晃堂 (昭55) .
- (11) Rosenfeld, A. and Kak, A.C. : "Digital Picture Processing", Academic Press (1976) .
- (12) Gonzalez, R.C. and Wintz, P. : "Digital Image Processing", Addison-Wesley (1977) .
- (13) 長尾, 白井 : "人工知能", 培風館 (昭55) .
- (14) 坂井利之 : "情報基礎学", コロナ社 (昭57) .
- (15) 馬場口, 北村, 塩野, 真田, 手塚 : "細線化を用いない文字パターンの方向セグメント抽出の一手法", 信学論 (D), J65-D, 7, PP.874-881 (昭57-07) .

- (16) 馬場口, 相原, 真田, 手塚: "文字パターンにおける方向セグメントの接続形状と幾何学的特徴部の抽出", 信学論 (D), J66-D, 5, PP. 495-502 (昭58-05) .
- (17) 馬場口, 正満, 真田, 手塚: "方形セグメント整合法による手書き漢字認識の検討", 信学論 (D), J66-D, 10, PP.1209-1216 (昭58-10) .
- (18) Babaguchi, N., Aibara, T., Sanada, H. and Tezuka, Y. : "Direction Segments Extraction and Its Application to the Recognition of Handprinted Chinese Characters", Proceedings of ICTP'83, PP.324-329 (1983-10) .
- (19) 馬場口, 北村, 塩野, 真田, 手塚: "文字パターンの方向セグメント抽出の一手法", 昭56信学会情報システム部門全大, 72 (1981-10) .
- (20) 塩, 小森: "手書き漢字認識におけるストローク抽出の一手法", 信学技報, IE80-14 (1980-05) .
- (21) 木村, 吉村, 三宅, 市川: "細線化パターンに対する非決定的ストローク抽出法", 信学論 (D), J61-D, 7, PP.496-503 (昭53-07) .
- (22) 岡部, 吉村, 三宅, 市川: "拡張距離関数と線形フィルタによる手書き文字の特徴抽出", 信学論 (D), J59-D, 12, PP.858-865 (昭51-12) .
- (23) 山田, 森: "手書き文字データベースの変動解析 (I)", 電総研彙報, 39, 8, PP.580-599 (昭50) .
- (24) 森, 山本, 山田, 斎藤: "手書教育漢字データベースについて", 電総研彙報, 43, 11/12, PP.752-773 (昭54) .
- (25) Hilditch, C., J.: "Linear Skelton from Square Cupboards", Machine Intelligence IV, New York (1969) .
- (26) 渡辺 茂: "漢字と図形", 日本放送出版協会 (昭51) .

- (27) 馬場口, 塩野, 真田, 手塚: "細線化を用いない文字パターンの特徴部分抽出法", 昭56関西連大, G8-14 (1981-11) .
- (28) 馬場口, 相原, 真田, 手塚: "方向セグメントの形状判定と特徴部分抽出について", 情報処理学会第25回全大, 6B-2 (1982-10) .
- (29) 馬場口, 相原, 真田, 手塚: "方向セグメントの形状判定と幾何学的特徴部の抽出", 昭57四国連大, 12-12 (1982-10) .
- (30) 安居院, 中嶋: "コンピュータ画像処理", 産報出版 (昭54) .
- (31) 小山, 塩野, 真田, 手塚: "細線化図形の屈折点検出の一手法", 信学技報, P R L 80-107 (1981-03) .
- (32) 山本和彦: "弛緩整合法による手書き教育漢字認識", 信学論 (D) , J 65-D, 9, PP.1167-1174 (昭57-09) .
- (33) 安田, 藤沢: "文字認識のための相関法の一改良", 信学論 (D) , J 62-D, 3, PP.217-224 (昭54-03) .
- (34) 高橋, 増田: "筆点の生起順序を利用したストローク抽出法による手書き教育漢字認識の基礎検討", 信学論 (D) , J 65-D, 10, PP.1294-1301 (昭57-10) .
- (35) 斎藤, 山田, 山本: "手書き漢字の方向パターンマッチング法による解析", 信学論 (D) , J 65-D, 5, PP.550-557 (昭57-07) .
- (36) 萩田, 梅田, 増田: "三つの概形特徴を用いた手書き漢字の分類", 信学論 (D) , J 63-D, 12, PP.1096-1102 (昭55-12) .
- (37) 萩田, 内藤, 増田: "大局的・局所的方向寄与度密度特徴による手書き漢字認識方式", 信学論 (D) , J 66-D, 6, PP.722-729 (昭58-06) .
- (38) 岡 隆一: "セル特徴を用いた手書き漢字の認識", 信学論 (D) , J 66-D, 1, PP.17-24 (昭58-01) .
- (39) 赤松, 川谷, 小森: "手書き漢字認識のための構造集積特徴", 信学論

- (D), J65-D, 5, PP.542-549 (昭57-07) .
- (40) 内藤, 小森: “手書き漢字を対象としたストローク密度分布による候補選択法”, 信学論 (D), J64-D, 8, PP.765-772 (昭56-08) .
- (41) 黒沢, 前田, 麻田, 坂井: “複合類似度による手書き漢字認識実験”, 昭56信学会情報システム部門全大, 79 (1981-10) .
- (42) 木村, 斐, 吉村, 三宅: “手書き漢字の部分パターン抽出と認識”, 信学論 (D), J64-D, 11, PP.1037-1044 (昭56-11) .
- (43) 日経エレクトロニクス: “大分類の段階がほぼ完成した手書き漢字認識の研究”, No.279, PP.148-167 (1981) .
- (44) 馬場口, 正満, 塩野, 真田, 手塚: “方向セグメントの矩形表現による手書き漢字の大分類”, 信学技報, P R L81-80 (1982-01) .
- (45) 正満, 馬場口, 真田, 手塚: “近傍状態を考慮した方向セグメントの矩形表現による手書き漢字の大分類”, 信学技報, P R L82-43 (1982-11) .