

タブレット入力による手書き文字・図形認識

n-Line Recognition Algorithms for
Handwritten Chinese Characters and Hand-drawn Line Figures

若原 徹 村瀬 洋 梅田三千雄

Toru WAKAHARA, Hiroshi MURASE and Michio UMEDA

日本電信電話公社 武蔵野電気通信研究所

Musashino Electrical Communication Laboratory, N.T.T.

1 まえかき

計算機の導入による高度な情報処理システムが広く社会に浸透しつつあり、今後ますます計算機と人間との円滑な対話実現への希求が強まると予想される。特に、素人でも扱える簡便な文字・図形入力法を確立する必要がある。

タブレット上に手書きした文字・図形を計算機が認識して入力する方式は、手書きという自然な入力形態を取る点および誤認識の発見や修正等在那場で対話的に行える点で有利である。筆者らは、文字・図表が混在する日本語文書を容易に手書き入力できるシステムの実現を目指し、これまで手書き線図形および手書き文字の認識法を検討してきた。将来はこれらをさらに高度化し融合していくつもりである。

本論文では、まず、接続ルールを考慮した候補ラティス法によるオンライン手書き線図形認識法を報告する。線図形認識では、比較的単純な構造を持つ少数のシンボル群を扱うが、それらシンボルを筆記する際の画数・筆順を無理に指定したり、シンボル毎のセグメンテーションをシステムに指示することは筆記者に大きな負担となる。本手法は、それら筆記制限を取り除き、任意の画数・筆順で筆記されたシンボルを含み一気に書かれた自由書式の図面認識を実現した。手法は、入力図形中から対象とするシン

ボルの候補図形を抽出し、候補ラティスと呼ぶ表に登録する段階と、候補ラティスの中から矛盾のないシンボル系列を選択し、それを認識結果とするという2段階から構成される。認識対象としては、今回JIS C6270で定められた28種類とシンボルと直線から構成されるフローチャートを選び、100枚のフローチャートの認識実験により、95.9%の認識率を得た。

次いで、選択的ストローク結合による画数・筆順に依存しないオンライン手書き文字認識法を報告する。文字認識では、漢字など多字種を扱うため、文字毎のセグメンテーションは前提とした上で、多様な手書き変形に対処し認識精度の向上を狙う。本手法は、楷書体での自然な筆記によるストロークの分断・統合という画数変動および任意の筆順変動を許容するために、画数・筆順の異なる入力および標準パターンの中で、ストローク対応を決定し、選択的ストローク結合で画数をそろえ、DPマッチングによりパターン間距離の算出を行う。当用漢字1851字種に対し認識率98.9-99.8%を得た。また、装置化の話題として、筆順自由の認識法をハードウェア化したオンライン手書き文字認識装置およびEレディスプレイ一体型タブレット装置について簡単に報告する。

以下、章2で手書き線図形の認識、章3で手書き文字の認識および装置化の話題を述べる。

2 手書き線図形の認識

線図形認識では、比較的少数のシンボル群を扱うが、それらの筆記法に一定のルールはなく、認識を困難にしていた。このため、従来のタブレット入力型の図形認識システムでは、画数と筆順を制限してシンボルのセグメンテーションも指示させたり^[1]、方眼紙上に筆記させていた^[2]。本章で提案する候補ラティス法は上記制限を取り除き、自由書式の図面認識を可能とする。

2.1 認識対象の設定

フローチャート、ブロック図、回路図など通常の技術文書に現れる線図形を扱う。今回は特に、JIS C 6270^[3]で定められた28種類のシンボルと直線から構成されるフローチャートとを認識対象とした。

2.2 認識アルゴリズム

認識アルゴリズムの流れを図1に示し、以下順に説明する。

(1) 手書き図形入力

データタブレットからストロークの系列として、図形情報を入力する。

(2) 候補図形の抽出^[4]

シンボル形状の位相構造を利用し、候補図形を抽出する。まず、入力図形の部分図形(例えば図2(a))に着目し、その図形は或るシンボル

(例えば磁気ディスク、図2(b))だと仮定する。部分図形を構成するストローク列の端点を、その位置情報を用いて、「仮定したシンボル」上の候補端点に対応付ける(図2破線)。その後、端点の接続関係をもとに、「仮定したシンボル」のすべての線分を通過する径路を探索する。径路の探索にはDepth First法による木探索を用いた。探索不能場合には仮定を棄却するが、可能な場合は着目した部分図形を候補図形とし、探索結果に基づき候補ストローク列を生成する。図2の例では、図2(c)に示す2個の候補ストローク列が生成される。他の画数の場合でも同様の処理を行うことにより、任意の画数・筆順に応じた候補ストローク列が生成される。すべてのシンボルの存在を仮定し、上記処理を種々の部分図形に対して行うことにより、候補図形の抽出処理を終了する。

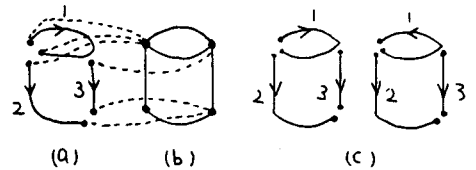


図2 候補ストローク列の生成

(3) 相異度の計算^[4]

生成された候補ストローク列と候補図形との間でマッチングを行い、候補図形が仮定したシンボルであるとしたときの相異度を求める。相異度には、ストロークの位置と方向系列のDPマッチングを採用する。

(4) 候補ラティスの探索^[4]

候補ラティスは入力図形のストローク番号に対応して、候補図形のシンボル名と相異度を記述した表である。図3の入力例に対する候補ラティス(一部分)を図4に示す。候補ラティスの中から、すべてのストロークをおおう最適な図形系列を選択することにより、入力図形に含まれるシンボルのセグメンテーションを決定し、同時にシンボルの名称も識別され、すなわち入力図形が認識される。ここでは、

$$\text{評価関数} = \sum_{\text{候補図形} \in \text{図形系列}} (\text{相異度}) / (\text{ストローク数})$$

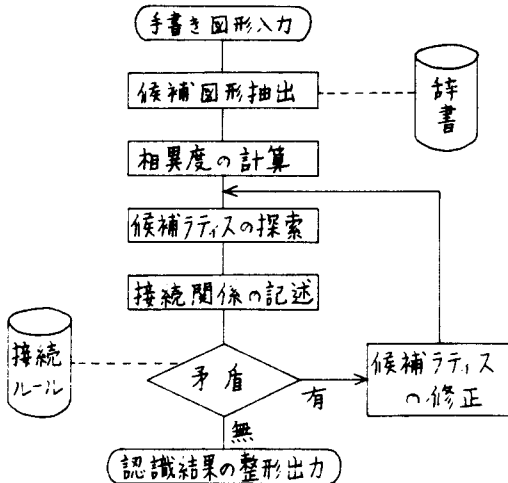


図1 線図形認識アルゴリズム

を最小とする候補図形の系列を選択し、この図形系列を「仮の認識結果」とする。図4の例では、正解は太枠で示した候補図形であるのに対し、図中井印を付した候補図形が抽出される。

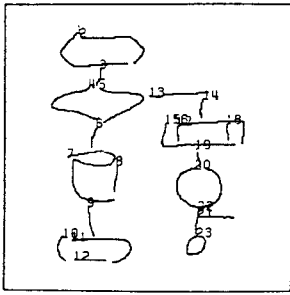


図3 入力例

この「仮の認識結果」を入力図形に対応させてそのまま表示したものが図5である。

(5) 接続関係の記述^[5]

入力図形は複数のシンボルとそれを接続する直線群から構成されている。各シンボルが直線を経由して他のシンボルに接続されているかを記述する。例えば、図5において「田」は「◇」と「○」に接続されている。

(6) 候補ラティスの修正^[5]

フローチャートには「端子記号はフローの中間には存在しえない」等のいくつかの接続ルールが存在する。これらの接続ルールに対して、(5)で記述した接続関係が不適切な場合には、関係する候補図形を候補ラティスから除去する。図5の例に対し、除去した候補図形を図4中に斜線で示す。その後、この修正された候補ラティスを再度探索する。

(7) 認識結果の整形出力^[5]

上記処理を反復実行し、接続関係に矛盾がなくなれば、それを最終認識結果として、整形して出力する。図3の例では、3回の反復実行の後に、図6の出力結果が得られた。

スロープ番号	15	16	17	18	19	20	21	22	23
	0.187	0.170	0.025	0.012	0.019	0.610	0.008	0.025	0.707
	0.344				0.132		0.097		
	0.070		0.038		0.101		0.140		
	0.076		0.113						

除去された候補図形
 仮の認識結果

図4 候補ラティス

2.3 図形データと認識実験結果

データ収集では、5名の筆者に5~10個程度のシンボルと直線から構成されるフローチャート(図3が一例)を書かせ、その際に①1つのシンボルは連続して書くもの、システムにセグメンテーションを通知する必要はない、②各シンボルの画数・筆順は自由である、とし、合計100枚の図形データを収集した。

上記データの認識実験結果を表1に示す。

表1 認識実験結果

接続ルール	認識率	セグメント誤り	シンボル誤り
無	89.2%	2.9%	7.9%
有	95.9%	0.6%	3.5%

接続ルールを導入したことにより、セグメンテーションの誤り(シンボルの包含関係加原因、例えば「Q」→「○」+「-」)、シンボル誤り(シンボル形状の類似)に対し、大幅な改善が見られた。4.1%の誤読の例には、接続ルールで修正できなかったもの(「○」→「◇」0.5%、「○」→「□」0.4%、...)、候補図形として抽出されなかったもの(0.4%)等がある。

接続ルールを考慮した候補ラティス法は、任意の画数・筆順で筆記されたシンボルを含み一気に書かれた自由書式の図面認識を可能とした。

今後の課題には、接続ルールを最適化すること、汎用性のある本手法をフローチャート以外の一般線図形に適用していくこと、等がある。また、文字・図表が混在した文書の扱いに対応するためには、文字と図形部分を分離する技術を確立する必要もある。

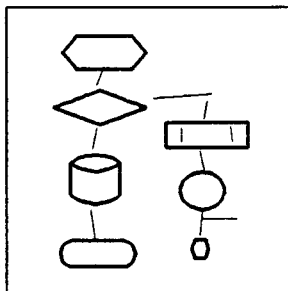


図5 第1回目認識結果

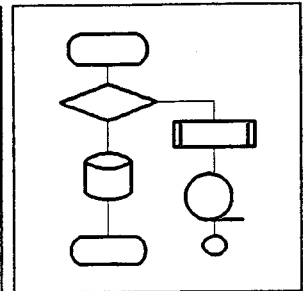


図6 最終認識結果の整形出力

3 手書き文字の認識

オンライン手書き文字認識を用いた漢字入力法を真に使い易いものとするには、画数・筆順遵守という筆記制限を緩和する必要がある。これまでは一方のみ^[6,7,8]あるいは両者の緩和については簡単な報告^[9]にとどまっている。本章では、認識精度を低下させることなく、楷書体での自然な筆記によるストロークの分断・統合という画数変動および任意の筆順変動を許す手書き文字認識法を提案する。但し従来通り、文字のセグメンテーションは前提とした。

3.1 認識アルゴリズム

入力パターンは、ストロークの分断・統合による画数変動および任意の筆順変動を含み得る。一方、標準パターンとしては、正しい画数と筆順で筆記した各文字カテゴリーの平均パターンのみを登録しておく。従って、画数・筆順が変動した文字をすべて登録する方法に比べて辞書規模ははるかに小さい。ここで、入力パターンと着目する標準パターンのうち、画数の多い方をM画、少ない方をN画とする($M \geq N$)。また、M画パターンはストローク群 T_1, T_2, \dots, T_M から、N画パターンはストローク群 S_1, S_2, \dots, S_N から構成されるとする。但し、添字は筆順を表す。このとき、パターン間距離 D を決定する。

認識アルゴリズムの流れを図7に示し、以下順に説明する。

(1) ストローク間距離算出

ストロークを構成する筆点列から、予め定めた等間隔刻み値 Δ で特徴点列を抽出する。ストロークTの特徴点列を (t_1, t_2, \dots, t_m) とすると、引き続く特徴点 t_i と t_{i+1} 間のストローク長が Δ となる。特徴点 t_i はXY座標値 (t_{ix}, t_{iy}) で表現する。2本のストローク $T = (t_1, \dots, t_m)$, $S = (s_1, \dots, s_n)$ に対し、次の2種のストローク間距離を定義する。 $m \geq n$ として示す。

$$df(T, S) = \frac{m}{n} \cdot \sum_{i=1}^n \{ |t_{ix} - s_{ix}| + |t_{iy} - s_{iy}| \} \quad (1)$$

$$dDP(T, S) = \min_{\tau} \sum_{j=1}^n \{ |t_{\tau(j)x} - s_{jx}| + |t_{\tau(j)y} - s_{jy}| \}$$

但し、 $\tau(1) = 1, \tau(n) = m$ (2)

df はストローク対応決定に、 dDP は選択的ストローク結合とパターン間距離算出に用いる。

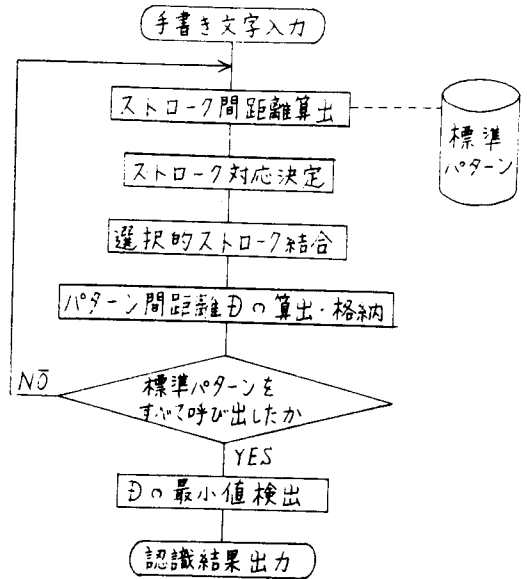


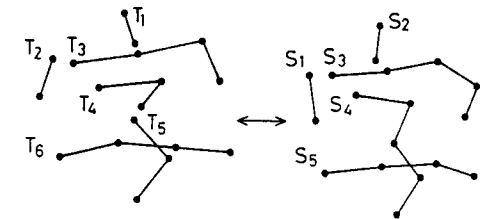
図7 文字認識アルゴリズム

(2) ストローク対応決定

M画とN画のパターン間で、式(1)の df によるストローク間距離和を最小とするN個のストローク対応決定し、筆順変動吸収を実現する。すなわち、対応決定条件は次のようになる。

$$\sum_{j=1}^N df(T_{i_j}, S_j) \rightarrow \min \quad (3)$$

かつ $T_{i_1}, T_{i_2}, \dots, T_{i_N}$ はすべて異なる



	S ₁	S ₂	S ₃	S ₄	S ₅
T ₁		○			
T ₂	○				
T ₃			○		
T ₄				○	
T ₅					
T ₆					○

○ … ストローク対応

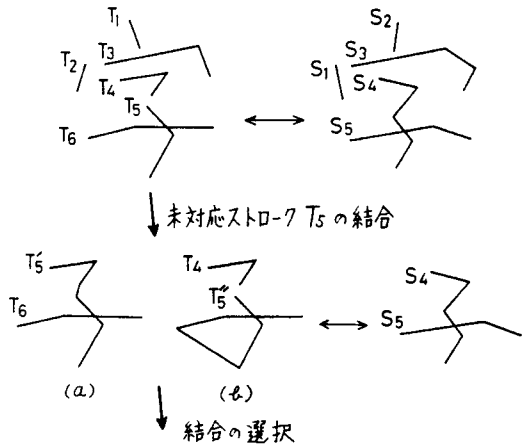
T₅ … 未対応ストローク

図8 ストローク対応決定

図8に、互いに画数・筆順の異なる「字」の2パターン間でのストローク対応結果を示す。図より、第1、第2ストロークの筆順変動が吸収され、画数変動が生じた「子」部分では、 S_4 に、その書き出し部分で重なる T_4 が対応している。

(3) 選択的ストローク結合

M画パターン中のM-N本の未対応ストロークの結合相手を決定する。それぞれ、筆順で前後する2本の既対応ストロークのいずれかと結合することとし、N画パターン中の対応ストロークとの重なりが良くなる方の結合を選択する。この処理により、M画パターンをN画に変換し、画数変動吸収を実現する。図8に示した「字」の例で、未対応ストローク T_5 に対する選択的ストローク結合の過程を図9に示す。結合相手の選択には、式(2)のdopによるストローク間距離の大小による判定を用いる。



$\alpha' \cdot dop(T_5, S_4) < \alpha'' \cdot dop(T_5, S_5) \rightarrow (a)$ を選択。
但し、 α', α'' : パターンの特徴点数比 (≥ 1)

図9 選択的ストローク結合

(4) パターン間距離算出

パターン間距離Dの算出は次式により行う。

$$D = \sum_{\text{結合ストローク}} \alpha_j \cdot dop(T_{ij}^+, S_j) + \sum_{\text{非結合ストローク}} dop(T_{ij}, S_j) \quad (4)$$

但し、 T_{ij}^+ : 選択的ストローク結合を施されたもの。
 α_j : T_{ij}^+ と S_j の特徴点数比 (≥ 1)

上記Dが最小となる標準パターンが属する文字カテゴリを認識結果として出力する。

3.2 文字データと認識実験結果

楷書体での自然な筆記で生じる画数変動の種類について、成人10名を対象に調査した結果を表2に示す。

表2より、丁寧な筆記でも極めて生じ易い1~5の画数変動を選択し、それら画数変動が生じた126カテゴリを収集した。一方、画数の正しい文字データとして、当用漢字1851カテゴリを収集した。両文字データとも任意の筆順変動を含む。また収集規模はともに40サンプル/カテゴリである。

表2 基本的画数変動の種類 (): 画数

	種類		種類		種類
1	子(3)→子(2)	6	い(2)→い(1)	11	ク(2)→ク(1)
2	又(3)→又(2)	7	レ(2)→レ(1)	12	コ(2)→コ(1)
3	β(3)→β(2)	8	ル(2)→ル(1)	13	么(3)→么(4)
4	之(3)→之(2)	9	ル(2)→ル(1)	14	弓(3)→弓(4)
5	水(4)→水(3)	10	了(2)→了(1)		

認識実験では、入力パターンの画数Kに対し、 $K-2 \sim K+2$ の画数範囲の標準パターン内で認識を行うこととした。表3に得られた認識率および累積分類率を示す。

表3 収集データの認識結果 単位: %

文字データ	認識率	累積分類率		
		第2位	第5位	第10位
画数変動5種 126カテゴリ	98.88	99.62	99.69	99.92
画数の正しい 1851カテゴリ	99.75	99.97	99.99	99.99

更に本手法の適用限界を調べるため、表2に掲げた6~14の9種類の画数変動を含むデータを画数の正しい当用漢字データより各々10カテゴリずつ生成して認識実験を行った。表4に得られた認識率および累積分類率を示す。

表4 生成データの認識結果 単位: %

文字データ	認識率	累積分類率		
		第2位	第5位	第10位
画数変動9種 90生成カテゴリ	98.96	99.69	99.83	99.86

誤認識例として、牛→午，兄→冗，各→右などがあり、字体変形を含む画数変動加画数の少ない文字に生じた場合に、本手法のように楷書体の標準パターンを用いた重ね合せ法は限界を持つことが明らかとなった。

今後はくすし字を目標に、より自由な手書き文字の認識法へ発展させていく。また、文字・図表混在の文書を扱うには、文字のセグメンテーションの問題に対処する必要がある。

最後に、試作の話題を述べる。筆者らは、筆順自由の文字認識法をハードウェア化してオンライン手書き文字認識装置を試作し、これを組み込んだ手書き入力日本語ワードプロセッサ「AESOP」を実現した^[10]。当用漢字・平仮名を含む2057字種を0.5秒/字以下で手書き入力できる(写真1)。更に、AESOPにおいて、透明電極により電磁結合型タブレットを透明化し、平板型ELディスプレイと一体化した。書きつつある文字の筆跡および認識結果の活字パターンを表示し、手書き入力時の操作性を著しく向上させた(写真2)。

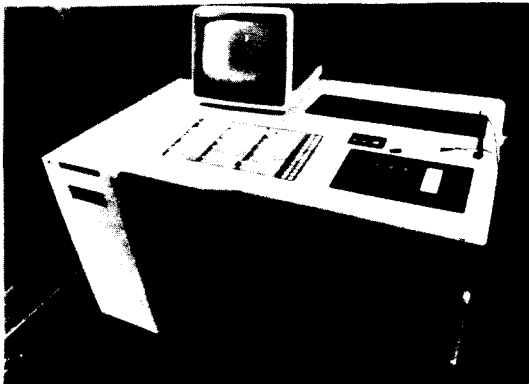


写真1 AESOP全景

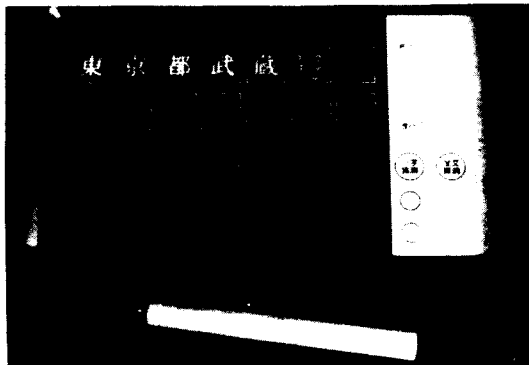


写真2 EL表示一体型タブレット

4 むすび

筆者らが研究を進めているオンライン手書き文字・図形認識法について報告した。

手書き線図形認識では、接続ルールを考慮した候補ラティス法を提案した。シンボル毎のセグメンテーション指示を不要とし、各シンボルの画数・筆順も自由とし、更にシンボル間の接続ルールの導入により類似シンボル間の認識精度を高めた。100枚のフローチャートデータに対して95.9%の認識率を達成した。

手書き文字認識では、楷書体での自然な画数変動および任意の筆順変動を許す認識法を提案した。画数・筆順の異なる入力および標準パターン間で、ストローク対応を決定し、選択的ストローク結合で画数をそろえ、パターン間距離の算出を行う。当用漢字1851字種に対して98.9-99.8%の認識率を達成した。

今後は、文字・図表の混在した日本語文書を容易に手書き入力できる操作性の良いシステムの実現を目指し、認識アルゴリズムを高度化し融合していくつもりである。

謝辞 日頃御指導頂く野柳基礎研究部長、島田統祐役、増田第二研究室長、及び熱心に討論して頂いた諸氏に深謝する。

文献

- [1] W.C.Lin, J.H.Pun: "Machine Recognition and Plotting of Hand-sketched Line Figures", IEEE, SMC-3 (1978).
- [2] Kato, Iwase, Yoshida, Tanahashi: "Interactive Hand-drawn Diagram Input System", PRIP pp544-549 (1982).
- [3] JIS: "情報処理流丸国記号, C6270".
- [4] 村瀬, 若原, 梅田: "タブレット入力による手書き線図形認識アルゴリズムの検討", 信学技報PRL81-69 (昭57).
- [5] 村瀬, 若原, 梅田: "接続ルールを考慮した候補ラティス法による手書きフローチャートのオンライン認識", 信学技報PRL82-71 (昭58).
- [6] 小高, 荒川, 増田: "ストロークの点近似による手書き文字のオンライン認識", 信学論(D), J63-D, 2, p.153 (昭55-2).
- [7] 小高, 若原, 増田: "筆順に依存しないオンライン手書き文字認識アルゴリズム", 信学論(D), J65-D, 6, p.679 (昭57-6).
- [8] 吉田, 迫江: "ストローク数に依存しないオンライン手書き文字認識の一手法", 昭57 信学全大, no.1357.
- [9] 小高, 橋本: "画数に依存しないオンライン文字認識法の一検討", 昭56 信学全大, no.1397.
- [10] 小高, 若原, 橋本: "オンライン手書き文字認識装置", 信学論(D), J65-D, 8, p.951 (昭57-8).

若原 徹 武蔵野通研基礎研究部第二研究室

〒180 武蔵野市経町3-9-11 0422-59-2226

村瀬 洋 同上 0422-59-2226

梅田三千雄 同上 0422-59-2227