

画数, 筆順, 回転, 区切りに依存しないオンライン 手書図形認識方式†

正 嶋 博†† 横 山 孝 典††
葛 貫 壮 四 郎†† 福 永 泰††

本論文は、オンライン手書図形認識方式について述べている。本図形認識アルゴリズムは、「線分セグメント法」を用いている。まず、入力ストロークを2種の線分要素（直線及び円弧）に分解し、それらを1方向にトレースする。このとき、隣合う線分要素間の角度変化を辞書図形のそれと比較することにより、最終的にマッチングする。この方式により、入力した手書図形に対し、画数フリー、筆順フリー、図形フリーを許容する。一方、図形を入力する際、文字認識における文字枠に相当するものが無いため、個々の図形を切り出す処理が必要となる。そこで、多数の線分要素から図形を自動的かつ高速に抽出する高速部分パターン抽出法を用いる。この結果、接続線と続け描きした図形等も、線分要素を入力順の両端から順次切り落とし、辞書図形と対応させることにより、効率良く切り出しできる。この方式は、従来の総当たり法と比較して、4~10倍高速に処理できる。認識率は、単純平均で82.7%、また技術文献中の対象図形と一致した図形2,626個中96.4%の作成が可能であることがわかった。認識速度は、16ビットマイクロプロセッサを使用して、約0.5秒/図形である。

1. ま え が き

手書入力、計算機の利用を多数の非専門家に広げるための有効なマンマシンインタフェース（マンマシン I/F）であり、文書作成、図面清書等の広い応用範囲を持つ。手書入力装置としては、OCR¹⁾に代表されるオフライン認識装置や手書ワープロ等のオンライン認識装置が従来から積極的に研究開発されている²⁾。

本論文では入力制限の少ないオンライン手書図形認識方式について述べる。オンライン認識を採用した理由は、計算機のパーソナル化に伴って、その対話性がマンマシン I/F の重要な要因の一つになると考えたからである。

ところで従来のオンライン手書認識方式は、文字認識が中心であり、画数、筆順、入力位置等に制限を持つのが普通である³⁾。入力データが文字列の場合には、上記の制限は致命的ではない。一方、入力データが図形の場合に、図形には標準的な画数、筆順がなく、入力位置も任意であり、さらに同一図形を傾けて入力したい場合もあるため、これらの自由度を制限する従来の方式をそのまま適用することはマンマシン性の点で適当でない。

そこで、本論文では、画数、筆順、回転、区切りに依存しない手書線図形の認識方式について述べる。

2. 認 識 対 象

今回認識の対象とした図形は、複数の雑誌、文献で用いられている図面から抽出した。抽出条件は以下のとおり。

- ①図形内にたかだか一つの閉ループを持つ。
- ②他の図形に対し包含関係に無い。
- ③使用頻度が高い。

この結果、表1に示すような図形を選択した。これらの図形を認識できれば、調査した文献内図面の約70%以上を作成することが可能となる。

3. 辞 書 構 造

各図形は360°を32分割したベクトルのチェーン構造で表現する。この分割数は、表1の認識対象図形を表現できる最小角度により決定している。本方式で“円弧”は、その弦の方向と弦に対する右回り/左回り方向フラグで表現する。このほか、図形の線分要素数及び開図形と閉図形を区別するための開放端数も大分類処理のために合わせて登録しておく。

図1は「プリンタ」図形の認識辞書構成である。

4. 認 識 アルゴリズム

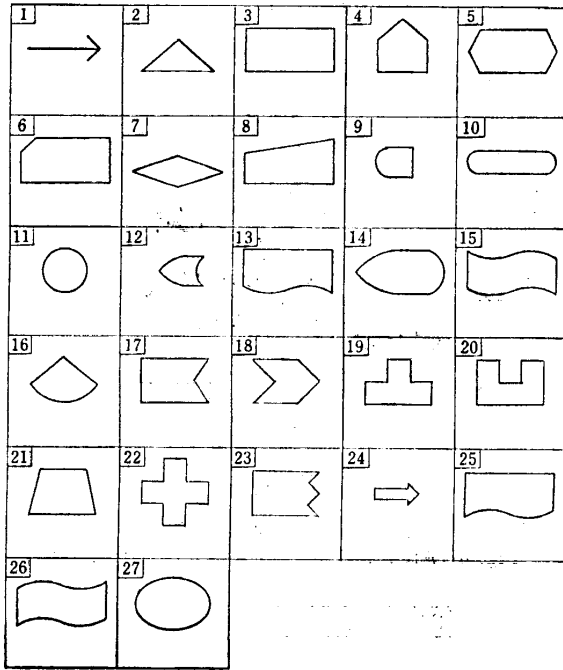
4.1 概 要

第3章で述べたように、本アルゴリズムでは図形をベクトルのチェーン構造で表現している。チェーン構

† Stroke Number and Order, Figure Rotation and Division Free Online Handsketched Figure Recognition Algorithm by HIROSHI SHOJIMA, TAKANORI YOKOYAMA, SOSHIRO KUZUNUKI and YASUSHI FUKUNAGA (Hitachi Research Laboratory, Hitachi Ltd.).

†† (株)日立製作所日立研究所

表 1 認識対象図形
Table 1 Figures for recognition.



プリント図形辞書

n	← 図形番号
5	← 線分要素数
2	← 円弧数
0	← 開放端数
8	← 第1線分要素方向コード
0	← " " 円弧方向フラグ
0	← 第2
0	← 第3
24	← 第4
0	← 第5
16	← 第6
1	← 第7
16	← 第8
-1	← 第9

図 1 辞書の構成例

Fig. 1 Structure of dictionary.

4.2 処理フロー

図 2 に本方式による手書図形認識の処理フローを示す。本処理は1ストローク入力ごとに実行され、図形を認識すると直ちに結果を出力する。以下、同図 a で示す入力例を用いて各処理について説明する。

(1) 線分要素化

入力されたストロークを2種の線分要素(直線/円弧)に分離する。本方式では折れ線近似後、円弧を抽出する。

(i) 面積法による折れ線近似

ストロークの点列を線分群に変換する手法は従来から検討されている⁹⁾が、ここでは、積分効果によりストローク中の手ぶれを効果的に吸収できる“面積法”を取用した。

図 2 a のストロークは図 3 a) のようなサンプル点列として取込まれる。これらのサンプル点列の任意の2点 P, Q に弦を張り、ストローク PQ と弦 PQ

造解析による図形認識は従来から検討されている⁹⁾。

ここではオンラインの特徴を生かして、①手書きストローク入力ごとにこれを直線と円弧の線分要素列に展開し、②1方向に線分要素をたどって1次元列を作成し、③1次元列の2線分要素間の角度変化を辞書図形と比較してパターンマッチする。なお、図形の区切りを自動化するため、④高速部分パターン抽出法(High speed Partial pattern Cutting: HPC 法)を立案した。以下、その詳細について述べる。

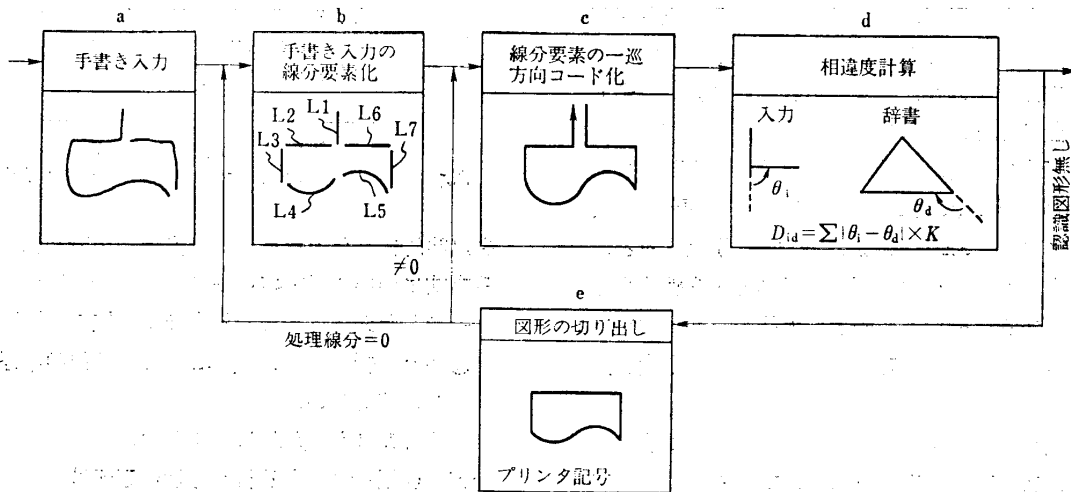


図 2 手書図形認識処理フロー

Fig. 2 Flow chart of handsketched figure recognition.

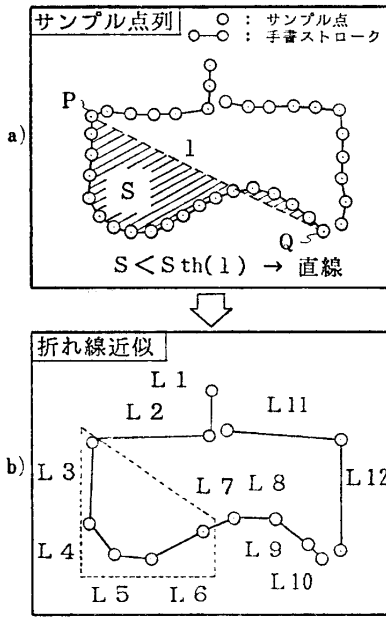


図3 手書きストロークの折れ線近似
Fig. 3 Polyline approximation of hand-drawn stroke.

で囲まれた面積 S を求める。これに対し、弦 \overline{PQ} の長さ l の関数である面積しきい値 $S_{th}(l)$ を設定し、 $S < S_{th}(l)$ (1) が成立した時、ストローク \overline{PQ} を一つの線分に変換する。上記(1)式を満足しない場合には、 \overline{PQ} 間をせばめて同じ処理を繰り返す。

ここで、面積は長さの2乗に比例することから、上記 $S_{th}(l)$ は l の2次関数になることが予想される。さらに、線が短い部分での手ぶれを抑えるため、定数項を正の値にすることが効果的と考えられる。この考えに基づいて、11人の被験者に総数440本の直線を描いてもらい、その長さとして面積との関係を調べたのが、図4である。この図で、領域 a は、曲線を描いたものと見なせる。したがって、残りの部分の上端に接するような関数を面積しきい値 $S_{th}(l)$ とすれば良い。ところが、線分長 L が長くなるにつれ、 $S_{th}(l)$ の値が2次的に増大すると、同図右上に示すようなカギ形の図形をも直線化のおそれがある。そこで、一定の線長 l_0 (図4参照) から先は、1次関数に置きかえ、 $S_{th}(l)$ の増大を抑えた。

この結果、ストロークの手ぶれを効果的に吸収した上で、折れ線近似を行うことができる。

(ii) 半径/線分法による円弧の分離

前記(i)の処理により折れ線近似したストローク

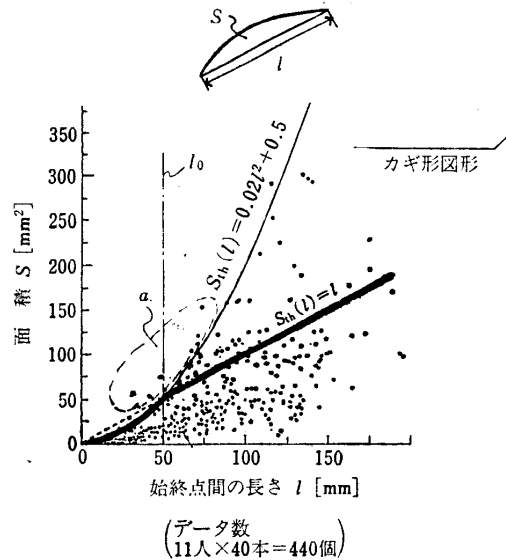


図4 直線ストロークの解析
Fig. 4 Area analysis for hand-drawn lines.

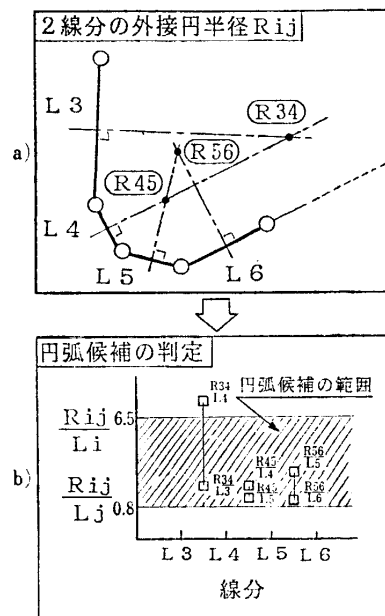


図5 R/Lによる直線/円弧分離
Fig. 5 Line/curve division by parameter R/L.

を最終的に2種の線分要素(直線/円弧)に分離する方式について述べる。一般に円弧を折れ線 i, j で近似した図形は次の特徴を持つ(図5a)参照。

- ① 隣合う折れ線 L_i, L_j に外接する円の半径 R_{ij} がほぼ等しい。
- ② 線分長 L_i, L_j がほぼ等しい。

一方、前記(i)による折れ線近似方式は

- ①' ゆるやかな部分は長い線分列で近似する。
- ②' 急峻な曲がりは短い線分列で近似する。

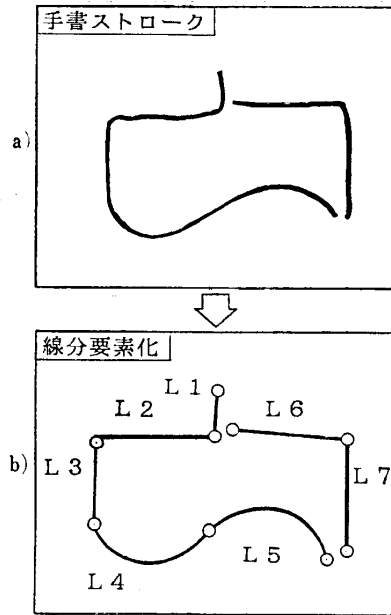


図6 手書ストロークの線分要素化例
Fig. 6 Line-segments of handsketched strokes.

という特徴を持つ。すなわち相似形の入力ストロークは、ほぼ相似形の折れ線に近似されるため、円弧の部分では上記の外接円半径 R_{ij} と折れ線長 L_i, L_j の比 (R_{ij}/L_i 及び R_{ij}/L_j) がほぼ等しくなる。したがって図5 b) で示すように R/L の範囲を適当に取れば、円弧部を抽出することができる。この R/L の上限と下限は約 300 個の円弧を含む図形の認識結果をもとに、

$$6.5 \geq R/L \geq 0.8$$

と設定し、連続した 3 本以上の折れ線に対する R/L がこの範囲内にある時、この部分を円弧としている。図6に円弧分離の例を示す。同図 a) は入力ストロークで、上記(i)(ii)の処理を施した結果、同図 b) のように 5 本の直線 (L_1, L_2, L_3, L_6, L_7) と 2 本の円弧 (L_4, L_5) が抽出される。特に円弧については、弦の方向に対し、右回り/左回りを 1 及び -1 のフラグを付ける。

(2) トレース

(1) で求めた線分要素群を一定方向にたどって一次元の線分要素列を作成する。この結果は入力時の筆順の影響を受けない。図2 c) の例は上端の線分から右回りにトレースした結果を示しており、線分要素列 $\{L_1, L_6, L_7, -L_5, -L_4, -L_3, -L_2, -L_1\}$ が得られる。ここで負号は、入力時と逆の方向にトレースされたことを示す。

表2 角度変化の差の平均値による類似図形間の有意差 (線分要素数: 4)

Table 2 Distance between similar figures by meaning of difference of angle between adjacent line-segments. (4 line-segments)

No	1	2	3	4	5
図形	長方形	キーボード	ひし形	台形	端子
1		2.0	4.0	4.0	2.0
2	2.0		2.0	2.0	4.0
3	4.0	2.0		4.0	6.0
4	4.0	2.0	4.0		6.0
5	2.0	4.0	6.0	6.0	

※) 数字は角度変化 (方向コード差) を示す。

(3) マッチング

上記(1), (2)の処理で、入力ストロークの前処理が終了し、辞書図形とのマッチングが行われる。辞書図形から、入力ストロークと同線分要素数のものを選択し、対応する角度変化を求め、上記(2)で得た 1 次元線分要素列の角度変化との差を算出し、それらの和を相違度 F として評価すると次式のようになる。

$$F = \alpha \times \frac{\sum |d\theta_i - d\theta'_i|}{N} + \beta \times \sum |C_i - C'_i| \quad (3)$$

ここで α, β : 係数

N : 線分要素数

$d\theta_i$: 辞書図形の i 番目角度変化

$d\theta'_i$: 入力図形の " "

C_i : 辞書図形の i 番目線分要素の円弧方向フラグ (-1, 0, 1)

C'_i : 入力図形の " " "

ここで $\alpha=2, \beta=1$ とすると、各図形間の有意差は、四つの線分要素を持つ図形を例にとると表2のようになり、十分に判別できることがわかる。

回転して描かれた図形に対しては、上記(3)式をそのまま適用すると入力図形と辞書図形との線分の対応関係がずれてしまう。これによる誤認識を避けるため、辞書図形の角度変化列を 1 ずつずらして発生させ、それぞれ入力図形と比較し最小の値をその辞書図形に対する相違度としている。

(4) 図形の切り出し

図形は、文字認識における文字ワクに相当するパターンの切り出し領域を設定できないが、マンマシン

性を考慮すると、図形の切り出しを自動化する必要がある。そこで、高速部分パターン抽出法 (HPC 法) による自動切り出しを行う。

筆者らの経験によれば、図形間を結ぶ接続線は、前述の図2 a の入力例に示すように、図形入力の前後に発生することが多い。したがって、入力制限として、

「図形入力の途中で接続線を描いてはいけない」という項目を付加することによるマンマシン性の低下は小さいと考えられる。そこで、本 HPC 法は、上記条件のもとで、新たなストロークが入力される度に、その最新の入力ストロークに対応する線分要素と、以前入力した線分要素を合わせ、それらの時間的な入力順に対し、その両端から順次線分要素を切り落とすことにより、図形を効率良く切り出している。これらの切り出した部分パターンに対し、第4.2節(2)及び(3)の処理により、辞書とのマッチングを行う。ここで、切り出しのある時点で辞書図形とのマッチングが成功した場合は、残りの線分要素を以下の2種に分ける。

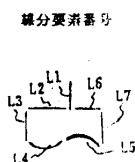
①辞書図形に対応する線分要素よりも前に入力した線分要素は、接続線として処理完了とする。
 ②一方、後に入力した線分要素は、新たな部分パターン切り出し対象とする。

したがって、手書図形の入力終了後に残った上記②に類する線分要素は、一定時間経過後、上記①の接続線として処理する。

表3は、上述の切り出し処理の具体例を示す。同表は図2の入力(第1ストローク {L1, ..., L5}, 第2ストローク {L6, L7}) に対し、プリンタ図形の部分パターン {L2, ..., L7} を切り出すまでの切り出し処理

表3 高速部分パターン抽出法による図形切り出し
 Table 3 Figure segmentation by high speed partial pattern cutting method.

処理回数	切り出し図形の線分要素番号
1	L1 L2 L3 L4 L5
2	L1 L2 L3 L4
3	L2 L3 L4 L5
4	L2 L3 L4
5	L1 L2 L3
6	L2 L3 L4
15	L5
16	L1 L2 L3 L4 L5 L6 L7
17	L1 L2 L3 L4 L5 L6
18	L2 L3 L4 L5 L6 L7



回数と、各時点で切り出した部分パターンを構成する線分要素群 {Li, ..., Lj} を示している。

④第1ストローク {L1, ..., L5} に対する切り出し

処理

{L1, ..., L5} のどの部分パターンも表1に示す辞書図形とマッチングしないため、線分要素数が1になるまで切り出し処理を進める(処理回数 1...15)。

⑤第1+第2ストローク {L2, ..., L7} に対する切り出し処理

上記④で {L1, ..., L5} は辞書図形とマッチングしなかったため、第2ストローク {L6, L7} を加えた線分要素群 {L1, ..., L7} に対し、切り出し処理を行う(処理回数 16...18)。

この結果、トータルで18回目の切り出し処理でプリンタ図形に相当する線分要素 {L2, ..., L7} の切り出し/マッチングが成功する。この切り出し/マッチ

表4 認識率と図形作成率
 Table 4 Recognition rate and drawing rate of figures.

No.	図形	個別認識率	出現頻度
1	矢印 1	100%	1,368個
2	三角形	100	33
3	長方形	100	684
4	五角形 1	95	0
5	六角形	95	2
6	カード	100	1
7	ひし形	90	22
8	キーボード	60	7
9	A N D	100	47
10	端子	75	104
11	円(楕円)	85	290
12	O R	65	43
13	プリンタ 1	65	2
14	C R T	5	3
15	紙テープ 1	60	0
16	扇形	82.5	1
17	PAD 判定 1	90	0
18	矢印 2	95	0
19	凸形	100	3
20	凹形	95	0
21	台形	95	6
22	白抜き十字	80	0
23	PAD 判定 2	100	0
24	矢印 3	90	10
25	プリンタ 2	65	0
26	紙テープ 2	65	0
	図形数	—	2,626個
	単純平均認識率	82.7%	—
	図形作成率	96.4	—

ング成功までの処理回数は, 従来の総当り法と比較して, 数分の1から10数分の1に減少しており, 本切り出し処理は認識時間の短縮に効果がある。

5. 認識結果と考察

表1に示した認識対象図形について, 各々の認識率を調べ, さらに文献内の頻度分布による重み付けを行って評価した。

テスト方法は, タブレット上の90×220mmのエリアに図形を一つずつ入力させた。この時,

- (i) 大きさ
- (ii) 画数
- (iii) 筆順
- (iv) 回転

については, 自由, とした。図7にサンプル例を示す。

表4に示すのが, 20人のサンプルによる認識結果と頻度分布による評価である。

なお, 同表では, 三角形と扇形については, 頂角が45度/90度の2種を用意したが, 被験者による入力図形のあいまいさから, 各々第2候補以内に入っている場合も可とし, かつ両方の認識率の平均をその認識率としている。

表4の個別認識率から言えることは, 一般に直線で構成された図形(三角形, 矢印, 四角形等)は認識率が高く, 円弧を含む図形(端子, 円, OR, プリント等)は認識率が低い。これは, 本方式が, 折れ線を基本としたアルゴリズムであることに起因すると考えられる。同表中, CRT図形の認識率が極端に低いのは, 上記の原因のほかに, 辞書データの不適切さも考えられる。

今回頻度分布調査に用いた文献は, 電子通信学会論文誌を初めとする学術論文が中心であったため, 対象図形中一度も用いられなかった図形もあった。文献中表1の認識対象図形に一致したものの2,626個の分布を表4出現頻度の欄に示す。同表より, 三角形, 矢印, 円で全体の約90%になり, かなり出現頻度に片寄りがある。本評価で, 本認識方式の有用性を結論付けるのは危険ではあるが, 少なくとも, 電子情報関係の図面に関し, かなり高い作成率(約96.4%)を示しており, 今後分野別に認識対象図形を選択する等の手段を講ずることにより実用的なシステムになると考えられる。

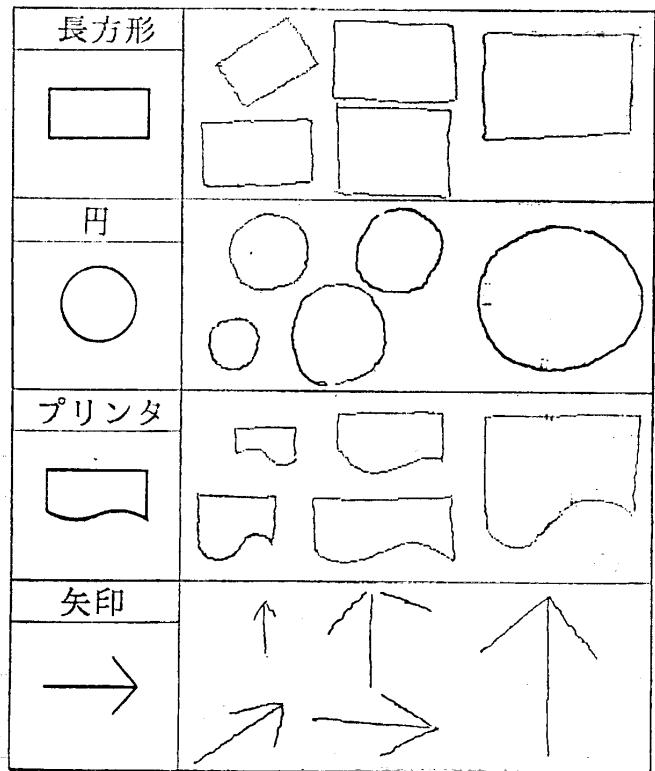


図7 図形認識サンプル例
Fig. 7 Example of samples for figure recognition.

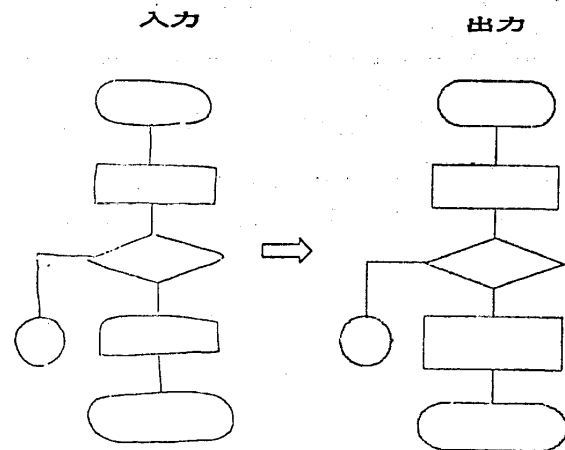


図8 手書図形認識例
Fig. 8 Example of handsketched figure recognition.

図8に, 実際に一つの図面としてよく用いられるフローチャート図面の認識例を示す。同図〔入力〕のように比較的丁寧に描かれた図形に対しては, 各図形とも平均0.5秒で認識し, 直ちに消書表示できる。また, 同図〔出力〕に示すように, 各図形とも, ほぼ描いた位置に描いた大きさで消書表示されていることがわかる。

6. む す び

以上、図形を手書入力する際、その画数、筆順、回転、区切りを意識せずに入力可能なオンライン図形認識方式と、被験者20名による認識実験及び使用頻度による図形作成率の評価について述べた。

入力制限の緩和という点では当初の目標を達成していると考えられるが、認識実験で得られた認識率は、単純平均 82.7%、技術文献図形の作成率 96.4% であり、昨今の文字認識技術を下回っている。これは予想された以上の変形があったため、今後、変形の吸収方式の改良、及び被験者の習熟度の検討を進める必要がある。

謝辞 最後に今回の研究にあたり、お世話になった研究室内外各位に深謝する。

参 考 文 献

- 1) 小川秀夫ほか：正読率が95%以上の手書漢字読み取りシステム，日経エレクトロニクス，1974年8月12日号，pp. 82-119 (1974).
- 2) くずし字など筆記制限を緩和する方向に進むオンライン手書き漢字認識，日経エレクトロニクス，1983年12月5日号，pp. 115-133 (1983).
- 3) 若原 徹ほか：EL表示付透明電極タブレットの手書き入力日本語ワードプロセッサへの応用，電子通信学会論文誌，Vol. J67-D, No. 9, pp. 981-988 (1984).
- 4) 長尾 真：画像認識論，p. 149，コロナ社，東京 (1983).
- 5) 岩瀬洋道ほか：手書き図形入力におけるストローク認識方式，電子通信学会総合全国大会，p. 1343 (1981).

(昭和60年7月24日受付)

(昭和61年2月20日採録)



正嶋 博 (正会員)

昭和32年生。昭和54年九州大学工学部電気工学科卒業。昭和56年同大学院総合理工学研究科修士課程修了。同年(株)日立製作所日立研究所入社。以来、ワークステーションソフトウェアの開発に従事。電子通信学会会員。



横山 孝典

昭和34年生。昭和56年東北大学工学部通信工学科卒業。昭和58年同大学院工学研究科修士課程修了。同年(株)日立製作所入社。以来日立研究所にてパターン認識、知識工学を応用したユーザインタフェースの研究開発に従事。電子通信学会、日本音響学会各会員。



葛貫壮四郎 (正会員)

昭和19年生。昭和53年3月茨城工業短期大学電気工学科卒業。昭和37年3月(株)日立製作所日立研究所入社。以来、電力システムの安定度向上の研究、エレベータ群制御の研究、現在ワークステーションのマンマシンインタフェースの研究に従事。電子通信学会会員。



福永 泰 (正会員)

昭和25年生。昭和50年京都大学大学院修士課程電気系工学研究科修了。同年(株)日立製作所入社。現在同社日立研究所研究員。計算機アーキテクチャ、グラフィックス、マンマシンインタフェースの研究に従事。電子通信学会、ACM 各会員。