

外周境界・孔境界を区別した対境界追跡による流れ図の構成要素の分離

Discrimination of Components in a Flowchart by Border Pair Tracing
with Distinction of Outer/Hole Borders

東谷 易 阿部圭一

Yasushi AZUMATANI and Keiichi ABE

静岡大学工学部

Faculty of Engineering, Shizuoka University

A method for discriminating the components in a flowchart image into symbols, flow lines, and characters is described. First, it extracts connected components containing symbols and/or lines based on their topologies and sizes, by tracing the borders. Meanwhile it dichotomizes the borders into outer and hole borders; the latter is further divided into three subclasses according to their sizes. The gist of the method is to follow a pair of borders of the extracted connected components simultaneously, separating different components. As the types of the pair borders is different for symbols, lines, and characters, they guide the separation and discrimination.

1. はじめに

図とそれに付随した短い文からなる図面は、情報をわかりやすく、かつ正確に表現することができるので、様々な場面で使用されている。このような図面の自動入力、CADの原データ作成等の観点から強く要求されている。そのためには図面からその表現している情報を抽出する必要があり、図面認識として広く研究が行なわれている[1][2]。

一般に図面認識は、図面の構成要素を分離・識別して、その構成要素相互間の関係を抽出することにより行なわれる。本報告では、図面の中から、ソフトウェアのドキュメントの一部を構成する流れ図を取り上げ、それを認識する最初の段階として、流れ図をその構成要素である文字、流れ図記号、流れ線に分離する方法を提案する。対象とする流れ図は、オフラインで入力し、文字と図形(流れ図記号と流れ線)が接触していてもかまわないが、図形には線の途切れがないことを前提とする。

本報告で提案する方法では、まず境界追跡により、図形の持つべきトポロジカルな性質と大きさを持つ連結成分を抽出して、これを図形の候補とする。このとき、抽出した図形候補の境界を外周境界と孔境界に区別し、さらに孔境

界を大きさにより分類する。次に、図形候補を構成する線の両側の境界を同時に追跡し、流れ図記号と流れ線ではその両側の境界の種類が異なることを利用して、流れ図記号と流れ線を別々に抽出する。このとき、流れ線は、境界の形状を利用して、1本ずつ切り離して抽出する。

本方法の処理は、境界追跡に基づいている。境界追跡は、2次元的な広がりを持つ画像において、扱うデータ量を減らすことができるので効果的である。特に図面のような背景の白地が多い大きい画像では効果的である。さらに、一度境界追跡を行うと以後の処理は、それによるコード化された情報を利用して行なえるので、処理時間の短縮が期待される。

2. 基礎的概念

本報告で必要となる基本的な概念を定義する。

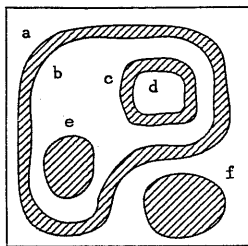
定義1 流れ図

流れ図は文字と図形から構成され、さらに図形は流れ図記号と流れ線から構成されている。従って、流れ図の構成要素は、流れ図記号、流れ線、文字であり、このうち図形を構成する流れ図記号と流れ線が図形要素である。そして、

本報告で対象とする流れ図記号は、記号の内側や外側に線のない1つの閉曲線からなる流れ図記号とし、定義済み処理記号や磁気テープ記号は除外する。

定義2 最も外側の外周境界、孔境界、連結成分

黒画素の連結成分とそれを直接囲む白画素の連結成分の間の8連結の境界を外周境界と呼ぶ。また、黒画素の連結成分とそれに直接囲まれる白画素の連結成分の間の8連結の境界を孔境界と呼ぶ。なお、どちらの境界とも、白画素に隣接する黒画素上にとる。そして、他の外周(孔)境界に囲まれていない外周(孔)境界を最も外側の外周(孔)境界と呼ぶ。ここで、本報告では以後、単に連結成分と言った場合は、黒画素の連結成分を意味することにする。連結成分についても同様に、他の連結成分に囲まれていない連結成分を最も外側の連結成分と呼ぶ。図1に例を示す。



- ・外周境界 : a, c, e, f
- 最も外側の外周境界 : a, f
- ・孔境界 : b, d
- 最も外側の孔境界 : b
- ・最も外側の連結成分 :
- 境界 a を持つ連結成分
- 境界 f を持つ連結成分

図1. 最も外側の外周境界、孔境界、連結成分の例

定義3 境界の外接長方形

境界上の全ての画素を含み、水平方向と垂直方向の辺を持つ、最も小さい長方形を境界の外接長方形と呼ぶ。

3. 流れ図の形状の特徴と構成要素分離の考え方

本報告で提案する方法で使用した流れ図の形状の特徴を述べる。そして、それらの特徴をどのように使用して、流れ図からその構成要素を分離するのかを述べる。

流れ図の形状の特徴

【特徴1】 図形は最も外側の連結成分である。言い換えれば、図形の中に図形があることはない。

【特徴2】 図形には、流れ図記号が1つ以上あり、流れ図記号は孔を持つ。このため図形は孔を持つ。

【特徴3】 流れ図記号の中には本文や識別子などの文字がある。

【特徴4】 孔の相対的な大きさは表1のように分けられる。なお本報告では、流れ図記号、流れ線による閉曲線の形状から考えて、孔の大きさは孔境界の外接長方形の縦辺の長さで評価する。

孔の種類	大きさ	大きさのばらつき
文字、結合子記号	小	中
標準的な流れ図記号	中	小
流れ線による閉曲線	大	大

表1. 孔の大きさによる分類

この表の分類に従って、本報告では、文字の大きさを持つ小さい孔境界を「文字の孔境界」と、標準的な流れ図記号の大きさを持つ中くらいの孔境界を「流れ図記号の孔境界」と、流れ線の作る閉曲線の大きさを持つ大きい孔境界を「流れ線による孔境界」と、それぞれ呼ぶことにする。

【特徴5】 流れ図記号は1個の閉曲線であり、流れ線は1本の線である。

【特徴6】 流れ線同士の接触点はT型(合流)か十型(交差)のいずれかの形状をしている。また流れ線と流れ図記号および流れ線同士の接触点には矢印があることがある。

以上の特徴を次のように利用して処理を行うと、構成要素を分離することができる。

【特徴1】から【特徴3】は、流れ図における図形のトポロジカルな性質を述べている。これより一般に流れ図における図形は、ある連結成分を囲む最も外側の連結成分であると言える。まずこの性質を利用して図形の候補を絞る。それは次のようにして行う。最も外側の連結成分は、最も外側の外周境界を持つ連結成分である。さらに、これで孔を持つものは、最も外側の外周境界と孔境界を持つ連結成分である。そこで、境界の包囲関係を考えた境界追跡を行い、最も外側の外周境界と孔境界を抽出する。ある連結成分を囲むかどうかは、TV走査を行い最も外側の孔境界の内側に黒画素があるかを調べることにより知ることができる。

【特徴4】は大きさに関する特徴である。。図形が持つべきトポロジカルな性質を持つ連結成分が、さらに標準的な流れ図記号の孔以上の大きさの孔を持てば、この連結成分は図形を含んでいる。このことは、流れ図記号が結合子記号だけである流れ図はないことを利用している。

このように【特徴1】、【特徴2】、【特徴4】により図形を含む連結成分を抽出することができる。ここで文字と図形が接触していない場合は、抽出される連結成分が図形であり、残りを文字であることにによって、文字と図形を容易に分離することができる。さらに【特徴3】でその分類をより確からしくすることができる。

また【特徴4】は、図形を含む連結成分の孔は、その大

きさにより孔の種類がわかることを示している。流れ図の各構成要素を形成する線の両側の境界は、表2のようにまとめることができる。この表を見ると、流れ線と流れ図記号を構成する線の両側の境界は異なり、流れ図記号だけが片側に流れ図記号の孔境界を持つことがわかる。従って、図形を含む連結成分の線の両側の境界を調べることで、図形から流れ図記号と流れ線を別々に抽出することができる。ただし、流れ図記号の外側と流れ線に接触した文字は流れ線となってしまう。

構成要素	形成する線の両側の境界
流れ線	共に外周境界 外周境界と流れ線による孔境界 共に流れ線による孔境界
流れ図記号	流れ図記号の孔境界と外周境界か流れ線による孔境界
文字	流れ図記号のパターン以外の全ての組み合わせ

表2. 流れ図の構成要素を形成する線の両側の境界

しかし、これでは合流や交差している流れ線は1つの流れ線として抽出されてしまうので、1本ずつ切り離す必要がある。そこで次のように行う。連結成分の両側の境界を同時に追跡し、境界の組み合わせにより、流れ図記号か流れ線かを判断して、それぞれを抽出する。このとき【特徴5】と【特徴6】により、接触している流れ線同士を切り離す。これにより流れ図記号と流れ線が、それぞれ1つの閉曲線、1つの線として抽出される。

以上が本報告で提案する構成要素の分離方法の考え方である。本方法で使用する形状の特徴は、トポロジー、大きさ、記号と接続線はそれぞれ閉曲線と線であること、そして接続点の形状であり、この順に使用している。手書きによる変動に最も弱い接続点の形状の使用を最後にすることにより、確からしさを高め、処理を単純化している。

以下、このような考え方にに基づき、4節で、文字と図形が接していない場合、文字と図形が簡単に分離できるアルゴリズムを説明する。そして5節で、このアルゴリズムをさらに発展させた、文字と図形の接触を許した流れ図をその構成要素である流れ図記号、流れ線、文字に分離するアルゴリズムの説明を行う。

4. 文字と図形が接していない場合の

流れ図の文字と図形の分離

文字と図形が接していない流れ図を対象として、図形を抽出することにより文字との分離を行う簡単なアルゴリズムを説明する。なお、本アルゴリズムは向田ら[3]の方法を整理したものである。

アルゴリズム1

初めに図形と考えられる連結成分を全て抽出し、それから順々に候補を絞ることにより、図形を抽出する。従って、各ステップの処理対象は、1つ前のステップで抽出した連結成分である。

- (1) 原画像から、最も外側の連結成分を抽出する【特徴1】。これにより図形の内部の文字は削除される。
- (2) 孔のある連結成分を抽出する【特徴2】。これにより孔のない文字は削除される。
- (3) 孔の中に黒画素のあった連結成分を抽出する【特徴3】。これにより内部に黒画素のない文字は削除される。

以上の処理は、最も外側の外周境界と孔境界を抽出するアルゴリズム(5.3節参照)を利用して1回のTV走査と境界追跡によって行なうことができる。

- (4) 図形には、標準的な流れ図記号が1つ以上ある。そこで孔を大きさにより、①文字と思われる小さい孔のグループ、②標準的な流れ図記号と思われる中くらいの大きさでそろっている孔のグループ、③流れ線による大きい孔のグループの3つに分類する【特徴4】。詳細は5.4節で述べる。そして②か③のグループの孔を持つ連結成分を抽出する。これにより文字程度の大きさの孔しか持たない連結成分は削除されて、図形だけ残る。

図2に流れ図からこの方法で図形だけを抽出した実験例を示す。ステップ(1)から(3)を一緒に実行した結果が図(b)である。そして図(b)の2つの連結成分ともステップ(4)で抽出されるので、この図が最終的に得られた図形を表わしている。なお参考として、図(c)、(d)に、ステップ(1)から(3)を別々に行なった場合に抽出される最も外側の連結成分と孔のある最も外側の連結成分を示しておく。

5. 文字と図形の接触を許す流れ図の構成要素の分離

文字と図形の接触を許した流れ図を対象として、流れ図記号と流れ線を抽出することにより、流れ図をその構成要素に分離するアルゴリズムの説明を行う。まず5.1節で中心となる考え方を述べ、5.2節でそのアルゴリズムの概要を示す。そして以後の節で順にアルゴリズムの各処理について説明を行う。

5.1 外周境界・孔境界を区別したペア境界追跡

図形を形成する線の両側の境界を、外周境界と孔境界に区別し、さらに表1の分類に従って孔境界を3種に分類すると、表2に示したように、流れ図記号と流れ線を形成する線の両側の境界は異なる。従って、それぞれの両側の境界を持つ線を抽出することにより、流れ図記号と流れ線を別々に抽出することができる。

ところが、構成要素同士が接触している部分をどう扱うかが問題となる。本方法では、図形を形成する線の両側の境界を同時に追跡することを行う[4]。追跡した2つの境界の間に新たに境界が現れたら、そこは分岐していることがわかる。流れ図記号は1つの閉曲線であるし、流れ線は1本の線であるから、これらには分岐がない。従って、分岐は他の構成要素が接触しているために生じる。このため、新たに現われた境界は、本来はペアで追跡した2つの境界の内の種類が同じ境界とつながっているはずである。そこで同じ種類の境界の間で切断を行い、接触している他の構成要素を切り離す。図3は、外周境界と流れ図記号の孔境界のペアで追跡しているときに新たに外周境界が現れた例である。このときは、2つの外周境界の間で切断を行い、右に伸びている線を切り離す。そして再び、外周境界と流れ図記号の孔境界のペアで追跡を続ける。

確かに境界の種類による切断は、流れ図記号を追跡しているときはうまくいく。しかしながら、流れ線同士が接触している部分ではだめなときがある。このため、その場合

は、接触点の形状の特徴を利用する。接触点の形状は境界に表れている。流れ図では、接触点の形状は単純なので、境界のなめらかさ、言い換えると折れ曲がっているかいないかにより接触を解析することができる。

5.2 アルゴリズム

アルゴリズム2

- (1) 最も外側の外周境界と孔境界を抽出する。そして抽出した外周境界と孔境界上の画素に異なったラベルを付け、それぞれを「外周ラベル」、「流れ図記号ラベル」と呼ぶ。【特徴1、2】
- (2) 孔境界を大きさにより3つのグループに分類する。【特徴4】
- (3) 孔境界のラベルの付け替えを行う。文字による最も小さい孔境界のグループに属する孔境界の画素に別のラベルを付け、これを「文字ラベル」と呼ぶ。さらに流れ線による最も大きい孔境界のグループに属する孔境界の画素に、各孔境界ごとに異なるラベルを付け、これらを「結線ラベル」と呼ぶ。結線ラベルはn種類(n:流れ線による閉曲線の数)になる。【特徴4】
- (4) 境界を区別したペア境界追跡を行い、追跡した部分を別々の閉曲線と線に分割する。このとき、流れ図記号の内側などに接触した文字は分離される。【特徴1、2、4、5、6】
- (5) 内部に黒画素のある閉曲線を抽出する。これが流れ図記号である。【特徴3、5】
- (6) (5)で抽出した閉曲線の間を結ぶ線を抽出する。これが流れ図記号間を結ぶ流れ線である。【特徴5】
- (7) (5)で抽出した閉曲線と(6)で抽出した線を結ぶ線を抽出する。これが流れ図記号と流れ線を結ぶ流れ線である。【特徴5】

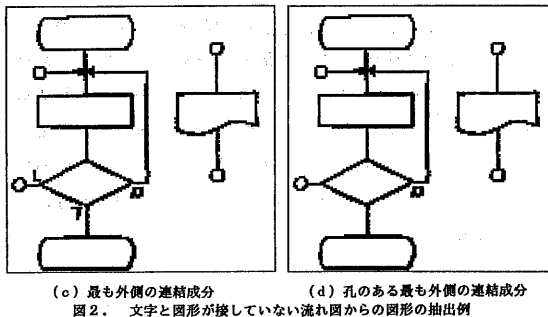
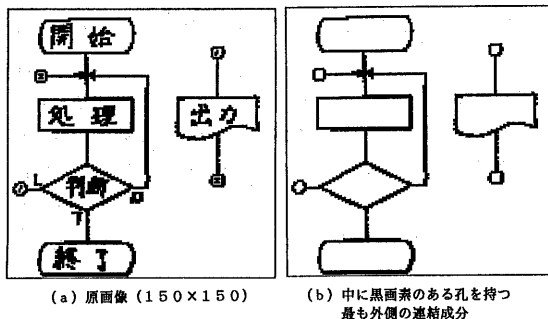


図2. 文字と図形が接していない流れ図からの図形の抽出例

図4に本アルゴリズムによる処理例を示す。

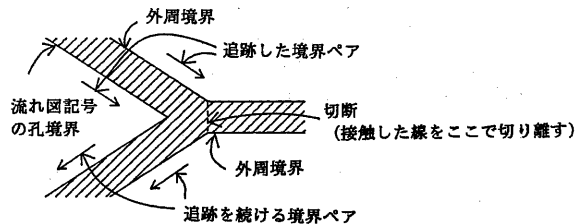


図3. 外周境界と孔境界を区別すると分岐点での処理が簡単になる例

5.3 最も外側の外周境界と孔境界を抽出する

境界追跡アルゴリズム

文献[5]に述べられている、最も外側の外周境界だけを抽出するアルゴリズムを修正して作成した。具体的なアルゴリズムとしては、文献[5]のアルゴリズム1を変更する部分だけを述べる。

①境界追跡の始点の条件

最も外側の外周境界の始点は、

$$f(i, j-1) = 0 \quad \text{かつ} \quad f(i, j) = 1$$

$$\text{かつ} \quad \text{LNBD} = 0 \quad \text{or} \quad -2$$

を満たす画素 (i, j) とする。また最も外側の孔境界の始点は、

$$f(i, j) \geq 1 \quad \text{かつ} \quad f(i, j+1) = 0$$

$$\text{かつ} \quad \text{LNBD} \geq 2$$

を満たす画素 (i, j) とする。

②境界のマーク付け

最も外側の外周境界を追跡するときは、“NBD”と“-NBD”はそれぞれ“2”と“-2”にする。また、最も外側の孔境界を追跡するときは“NBD”と“-NBD”はそれぞれ“3”と“-3”にする。その他のマーク付けの方針は変わらない。

③LNBD値の設定

TV走査の各行の始めて0にクリアする。そしてTV走査をして最後に出会った0、1以外の値を持つ画素値にする。

5.4 孔境界を大きさにより3つのグループに

分類するアルゴリズム

K平均アルゴリズムを用いて、全ての孔境界をその外接長方形の縦辺の長さにより、3つのクラスに分類する。

3個の初期クラスタの中心を小さい方から順に Z_1 、 Z_2 、

Z_3 とし、これらを次のように決める。

Z_1 = (孔境界の外接長方形の縦辺の長さの最小値)

Z_3 = (孔境界の外接長方形の縦辺の長さの最大値)

Z_2 = (次のステップで求められる $\bar{h}(S)$)

(1) 全ての孔境界における外接長方形の縦辺の長さの集合を S とする。

(2) S の要素の平均 $\bar{h}(S)$ と標準偏差 $\sigma_h(S)$ を求める。

$$\sigma_h(S) / \bar{h}(S) \leq \theta_1$$

を満たせば、終了する。なお θ_1 はあらかじめ与えた閾値である。

(3) S の他の全ての要素 h_j に対して

$$|h_i - \bar{h}(S)| \geq |h_j - \bar{h}(S)|$$

を満たす h_i を求める。 $S = S - \{h_i\}$ として (2) へ行く。

文字による小さい孔境界のクラスタと流れ線による大きい孔境界のクラスタが存在しない場合は、そのクラスタを除いて分類する。

5.5 ペア境界追跡アルゴリズム

本アルゴリズムは、図形の一部と思われる線の両側の境界を同時に追跡し、境界の種類と形状により、接触している線を切り離して、追跡した領域を閉曲線と線に分割して抽出する。図5にアルゴリズムの大まかな流れを示す。まず流れ図記号の線の両側の境界上にペア境界追跡（略してペア追跡ともよぶ）の開始点を設定する。そこから追跡している2つの境界の間に新しい境界がないか調べながら、追跡を行う。新しい境界が見つければ、そこは他の線が接触したために分岐していることがわかる。そこで境界間で切断を行って接触している線を切り離す。そしてさらに追

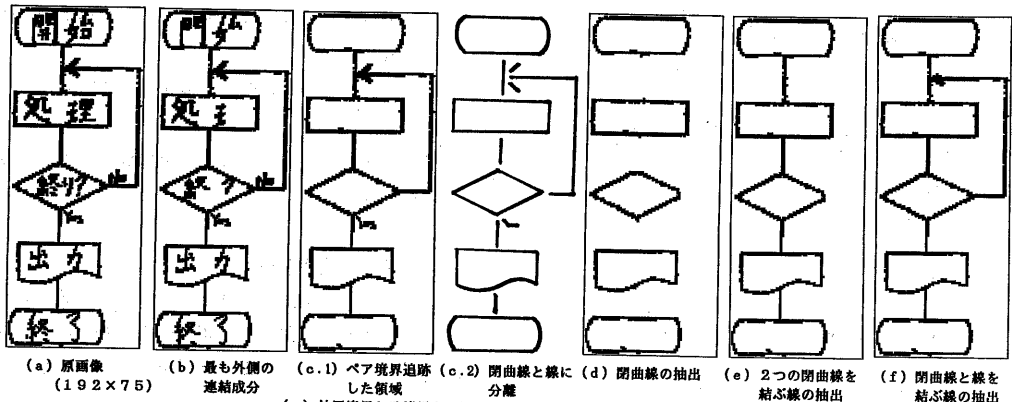


図4. アルゴリズム2による流れ図記号と流れ線の抽出処理例

跡を続ける。流れ図記号の閉曲線を追跡し終わったならば、今度はさきほど切り離れた線の追跡を行う。

次に、個々の処理を説明していく。

(a) 追跡始点ペアの検出

ペア境界追跡を開始する2つの境界上の一对の画素を「始点ペア」と呼ぶ。この始点ペアは、図形要素の境界上に設定したい。表2を見ると、流れ図記号の孔境界と外周境界が流れ線による孔境界のペアが現れるのは、流れ図記号だけである。その他の組み合わせは、文字の境界線上にも現れることがある。そこで一对の流れ図記号の孔境界と外周境界が流れ線による孔境界との上に始点ペアを設定する。アルゴリズムは次のようになる。

- (1) 画像をTV走査して、流れ図記号ラベルを持つまだペアで追跡していない画素 P_h を見つける。 P_h の4近傍のうち黒画素が連結している全ての方向へ、別の境界上の画素 P が現れるまでたどる。外周ラベルか結線ラベルを持つ P の中で、 P_h との距離が最も小さい P と P_h とを始点ペアとする。全ての P が外周ラベルも結線ラベルも持たなければ、 P_h の次の画素からTV走査を再開して P_h を見つけ直す。TV走査が画像の右下の端に達したならば、始点ペアはもはやないの

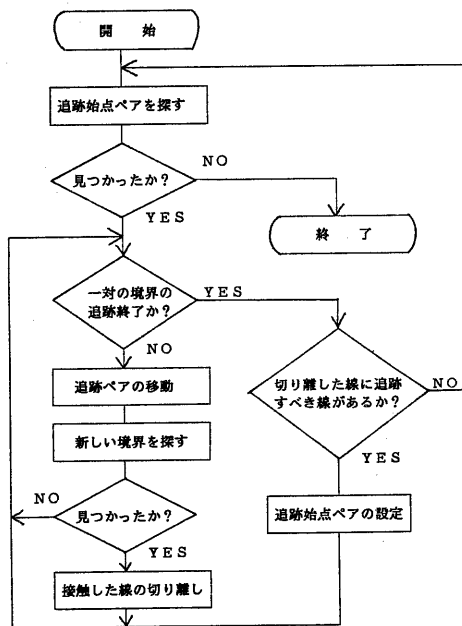


図5. ペア境界追跡アルゴリズムの概要

で、ペア境界追跡アルゴリズムは終了する。

- (2) この始点ペアからのペア追跡が終了したら、TV走査を P_h の次の画素から再開して同様に繰り返す。

このように始点ペアの一方を流れ図記号ラベルを持つ境界上に設定すると、標準的な流れ図記号の大きさの孔を持たない連結成分は追跡されない。これはそのような連結成分は文字とみなされて削除されることを意味する。

(b) 一对の境界の追跡終了

次のいずれかの状態になったとき、現在追跡している一对の境界の追跡を終了して、次の始点ペアを見つけに行く。

- ① 閉曲線を一周した
- ② 線の途切れた端に達した
- ③ 線の接触点において、後で述べる処理(e)により、現在追跡している線が端に達したと判断されて、切り離された。

また、このとき、その追跡において切り離れた線に追跡すべき線（流れ図記号の内側に接していた線以外）があれば、その線の切り離れた端を始点ペアとして、再び一对の境界の追跡を行う。

(c) 追跡ペアの移動

線の両側の境界上にある現在追跡している2つの画素を「追跡ペア」と呼ぶ。追跡ペアは、始点ペアから出発し、その間の距離を常に最小に保ちながら、1画素ずつ移動させる。次の3つの移動の仕方があり、このうち、新しい追跡ペア間の距離が最も小さい移動を選ぶ。

- ① 追跡ペアの両方ともそれぞれの境界上の次の画素へ移動する。
- ② 追跡ペアの片方だけその境界上の次の画素へ移動する。
- ③ 追跡ペアの別の片方だけその境界上の次の画素へ移動する。

(d) 新しい境界の検出

追跡ペアを移動させながら、それらの境界の間に白画素がないか調べる。調べた画素には調査済みのマークを付けていく。白画素が見つければ、その画素の近傍には新しい

境界上の画素がある。探すべき領域は、図6に示すように、追跡ペアを結ぶ直線の傾きにより、次の2つの場合に分けて、設定する。

- ① 追跡ペアを結ぶ直線とx軸、y軸とのなす角が共に 22.5° 以上の場合は、追跡ペアを対角頂点とする水平方向と垂直方向の辺を持つ長方形内で、境界の間の画素を調べる。
- ② その他の場合は、追跡ペアを対角頂点とする 45° 方向の辺を持つ長方形内で、境界の間の画素を調べる。

図7に新しい境界の検出例を示す。

(e) 接触した線の切断

流れ図記号は閉曲線である。また流れ線は流れ図記号間か流れ図記号と流れ線の途中を結んでいる。従って、図形要素による分岐は、ある流れ線の端が他の図形要素の途中に接触したために生じる。言い換えれば、分岐点から伸びている3つの線の内、2つは同一の図形要素であり、他の1つは別の図形要素である。このため後者と前者の図形要素を切り離せばよい。また流れ線は交差する場合がある。従って、交差は2本の線がクロスしたために生じたものであり、これも2つの図形要素から構成されている。

処理(d)により新しい境界が見つかったと、2つの境界の間で切断を行ない、追跡している線からそれに接触している別の線を切り離す。そして追跡してきた線の延長であるとみなされる、新たに選ばれた2つの境界でペア追跡を続ける。

切断する境界の対と新たに追跡する境界の対は、まず追跡した2つの境界と新たに検出された境界のラベルによって決める。ラベルだけでは決まらない場合には、これら3つの境界のなめらかさを調べる。ここで「なめらか」とは、現在注目している領域(分岐点付近)において、その境界

が屈折点を持たないことをいう。新しい境界がなめらかならば、そこは追跡している線が他の図形要素の途中に接触している分岐点であり、線はそこで終わりである(図8.(a)の場合)。そうでなく、追跡した境界の一方がなめらかならば、そこは追跡している線から別の線が出ている分岐点であり、追跡している線はなめらかな境界に沿って続いている(図8.(b)の場合)。新しい境界も追跡した2つの境界もなめらかでないならば、交差点であり、追跡している線は交差を飛び越えて続いている(図8.(c.1)の場合)。また矢印のところも、この場合になり、交差点と同様に矢印を飛び越えて続くと考えることができる(図8.(c.2)の場合)。

具体的には、表3を上段から順に調べて行き、当てはまる欄を探す。その欄に従って、切断する境界の対と新たに追跡する境界の対を決める。

表3の(I)と(II)は、図形に文字が接触したために小さい孔が発生したところで現われる。文字を切り離して、削除する。(流れ線として文字を追跡しているときに現われることもある。)(III)と(IV)は、流れ図記号を追跡しているときに現われる。(III)で、流れ図記号の内側に接触した文字を削除する。(IV)で、流れ図記号の外側に接触した線(流れ線か文字)を切り離す。切り離した線は後で追跡する。(V)から(K)は、流れ線を追跡しているときに現われる。(V)と(VI)では、追跡している線はそこで終わりとなる。(V)は、流れ図記号に達した場合である。(VI)は、他の流れ線の途中に達した場合である。(VII)では、同じラベルを持つ境界は切断を介して連続すべき境界であるとみなすことができる。(VIII)は、分岐点であり、追跡している線から別の線が出ている場合である。(IX)のときだけは、さらに近くを調べる必要がある。交差点か矢印のところであり、近くにもう一つなめらかでない新しい境界があるはずである。(VII)、(VIII)、(IX)で切り離した線は後で追跡する。

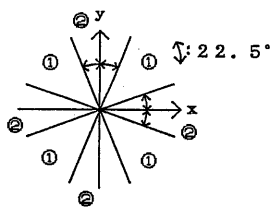


図6. 追跡ペアを結ぶ直線の傾きによる場合分け

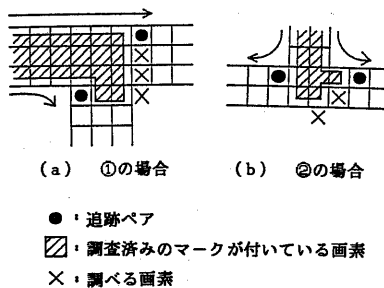


図7. 新しい境界の検出例

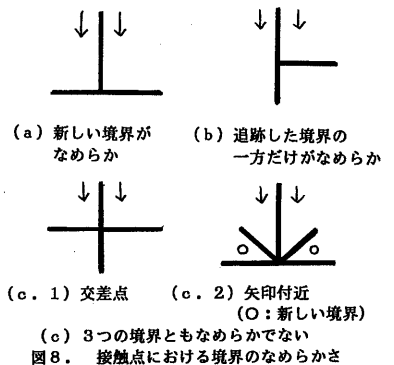


図8. 接触点における境界のなめらかさ

6. おわりに

本報告では、連結成分のトポジカルな性質と大きさを使用すると、流れ図から図形を含む連結成分を抽出できることを示した。そして、外周境界と孔境界を区別し、さらに孔境界を大きさにより分類して、境界をペアで追跡することにより、流れ図を個別の流れ図記号、流れ線、文字に簡単に分離できることを示した。本方法の処理の多くは、流れ図記号の一般的な構造に基づく特徴を利用し、接続点の形状は最後に利用している。このため、本方法は、手書きの変動に対して強いと言える。

本方法は、構成要素の分離の他に、流れ図の認識において有用な次のような情報も抽出している。流れ図記号と流れ線を分離できたので、流れ線による流れ図記号間の接続関係を知ることができる。また、境界を追跡するときに、流れ図記号の形についての情報を得ることができる。さらに、接触している線を切り離すことにより、流れ図記号から出ている流れ線の数、磁気テープ記号の横棒と思われる線を抽出できる。

最後に、今後の課題としては、

- 1) 孔の大きさだけで、それが何の孔であるか判断している。このため、結合子記号の小さい孔は、文字による孔と判断されてしまうので、結合子記号は抽出されない。この対策としては、孔の形を識別すること、孔境界の隣接関係と結合関係 [6] による区別を利用することなどが考えられる。
- 2) 内側や外側に線のある流れ図記号や重ねて書いた記号を抽出できるようにする必要がある。それには、流れ図記号をペア追跡したときに切り離した線を全て追

跡して調べることが考えられる。

- 3) 1つの閉じた閉曲線を流れ図記号として抽出し、両端が流れ図記号が流れ線に接触している線を流れ線として抽出している。このため、図形の線が途切れている場合は、その部分は抽出されない。従って、途切れを接続しなければならないが、それには前処理としての接続を行う方法と構成要素の分離を行ないながら接続する方法の両方から考える必要がある。などがある。

参考文献

[1] 棟上、佐藤：“図面の自動認識と理解”、情報処理、Vol. 24, No. 9, pp. 1086-1094 (1983)。
 [2] 岡崎、恒川：“サーベイ：線図形認識技術のE.A. O.Aへの応用”、情処研資、CV34-8 (1985)。
 [3] 向田、鈴木、阿部：“境界追跡を利用した流れ図中の文字と図形の分離”、信学技法、PRL83-70 (1984)。
 [4] T. Agui and T. Furukawa: "A method for extracting and restoring of contour lines on a contour map", Trans. IECE of Japan, Vol. E63, No. 8, pp. 581-587 (1980).
 [5] S. Suzuki and A. Abe: "Topological structural analysis of digitized binary images by border following", Computer Vision, Graphics, and Image Processing, Vol. 30, No. 1, pp. 32-46 (1985).
 [6] 鈴木、向田、阿部：“孔境界の隣接・結合関係の抽出とその流れ図認識への応用”、情処第27回全大、3Q-9 (1983)。

/\	ペア追跡した境界		新しい境界		切断する境界の対	新たにペア追跡する境界	付加処理	
	ラベル	形状	ラベル	形状				
(I)	——	—	文字	—	新しい境界となめらかでない境界	新しい境界となめらかな境界	なし	
(II)	一方は文字		——					
(III)	流れ図記号と外周が結線		流れ図記号		—	新しい境界と流れ図記号ラベルの境界		新しい境界と外周が結線ラベルの境界
(IV)			外周が結線			新しい境界と外周が結線ラベルの境界		新しい境界と流れ図記号ラベルの境界
(V)			流れ図記号			追跡した境界の組		(追跡ストップ)
(VI)	——		——		なめらか			
(VII)	互いに異なる(2つの結線か外周と結線)		追跡した境界と同じ	—	新しい境界とそれと同じラベルの境界	新しい境界とそれと違うラベルの境界		
(VIII)	——		一方だけなめらか		新しい境界となめらかでない境界	新しい境界となめらかな境界		
(IX)	——		——		新しい境界と追跡した境界で2組	新しい境界同士		もう一つなめらかでない新しい境界を見つける

表3. 接触点における切断する境界の対と新たにペア追跡する境界 (上段から順に当てはめる。 — : don't care.)