

複数の画像を用いて状態を記憶する図面認識の制御法

猪飼秀隆*

阿部圭一

静岡大学工学部

図面に書かれた文字や図形の認識を行なうには、部分的な認識処理においても、図面の記述に関する経験則のような知識を利用し、大域的な認識結果との整合性を計る必要がある。このために、本報告では、図面認識過程に対する次のような制御方法を提案する。

(1) 部分的な認識処理の結果のそれぞれを画像として記憶し、このような複数の画像全体が総合的な認識結果であると考え、(2) 図面のもつ意味に対する仮説を画像として生成し、総合的な認識結果を利用して検証を行なう。本文では、これらの提案を具体的に、文字と図形の分離アルゴリズムとして実現する。そして、流れ図図面を入力とした実験を通して、提案の有用性と可能性を示す。

A Control Method of Recognition of Diagrams
with Multiple Interpreted Images

Hidetaka IKAI⁺

Keiichi ABE

Faculty of Engineering, Shizuoka University
Johoku 3-5-1, Hamamatsu, JAPAN

We propose a method of controlling the recognition process of diagrams, such as flowcharts, utilizing the result of global recognition as well as the heuristic rules of drawing diagrams. Our proposal consists of:

- (1) to memorize the state of each partial recognition subprocess as an image;
- (2) to generate a hypothesized image for a possible local interpretation of the diagram and to test its compatibility with the global recognition result.

In order to show the effectiveness and possibility of the method, we implemented it as an algorithm for segmenting flowchart images into characters and figure components.

* 現在 久留米工業大学

+ Current address: Kurume Institute of Technology
Kamitsu 2228, Kurume, Fukuoka, JAPAN

1. はじめに

図面に書かれた文字や図形の認識を行なう場合、画像上の画素の並びや距離といった部分的な情報を手がかりに処理するだけでなく、図面の構成のような総合的な認識の結果を利用することが望まれる。我々は、部分的な認識処理においても、総合的に行なわれた認識の結果を利用するために、複数枚の画像を用いて認識状態を記憶する制御方法を提案する。

これまで、線画やシーンを対象とした画像認識処理（画像理解）における総合的な制御に関するいくつかの研究が行なわれてきた。線画の解析(1)では、多面体に関するヒューリスティックな知識を利用し、線要素を予想した処理を行なうものがある。航空写真の解析(2)では、シーンに関する知識の表現方法や利用方法に関する研究が行なわれている。そこでは、部分的な画像解析から得られた情報をもとに対象の物体をボトムアップに予測して、その物体に関する知識を取り出し、画像上に存在すると考えられる対象物をトップダウンに探索する。ここでの部分の認識結果は（画像の形ではなく）シンボルの形で記憶され、多くの可能な結果が非決定的に並行処理されて淘汰されていく。また最近では、図面を対象とした研究報告(3)もある。そこでは、曖昧さを考慮した図面の構造解析を行なっている。

画像認識処理の制御には、種々のアプローチがある(4)。物体の画像上の最適な配置をダイナミックプログラミングにより求める方法、フィードバックを含む制御によって逐次的に処理を行なう方法、可能な解釈の組合せから矛盾する解釈を消去していく方法、対象に関する知識を確率で表現して確率的弛緩法を用いるもの、などが挙げられる。

これまで、論理回路図やソフトウェア流れ図を対象とした図面認識処理では、図面の構成要素を分離した後、文字の認識処理と図形の認識処理に分けて考えることが多かった。そして、2値化処理やとぎれた線の接続を認識処理の前処理として考えた。ところが、例えば文字と図形の分離処理について考えると、分離結果は文字の認識と図形の認識に影響を与える。分離

に誤りがあれば、それに基づく認識はもっと大きな誤りを犯す恐れがある。また、分離処理の見方を変えると、文字や図形をそれぞれ認識できることが文字と図形を分離できることと言える。とぎれた線の接続処理についても、画素の並びや距離といった部分的な情報を手がかりに処理するだけでは、要求を満たす接続を行なうことは困難であることがわかっている。図面の構成のような総合的な認識の結果を利用することが必要であろう。つまり、このような認識処理は互いに密接に関係しているので、切り離しては考えることのできない処理である。従来の研究でも、このように相互に関係した総合的な認識の制御の必要性が言われていた。我々は、そのような制御方法として、部分的な認識結果を画像の形で記憶する方法を考え、これを図面認識における文字と図形の分離処理に対して具体化した。

本文では、図面認識処理を行なう上で有効であろうと思われる解釈画像という記憶を基にした認識の制御方法に関する提案を示す。そして、提案に基づく「文字と図形の分離アルゴリズム」を示し、具体的な入力に対しての処理結果を示す。実験を通して、提案の有用性と多くの問題点がわかった。

2. 解釈画像の提案

認識過程の記憶の実現に複数の画像を用いて、この複数の画像全体をまとめたものが認識状態を記憶していると考え、この複数の画像を解釈画像と呼ぶことにする。図面認識処理では、入力画像、文字画像（文字であると認識された領域）、記号画像（記号であると認識された領域）、仮定画像などを考えればよい。図1に解釈画像のモデルを示す。

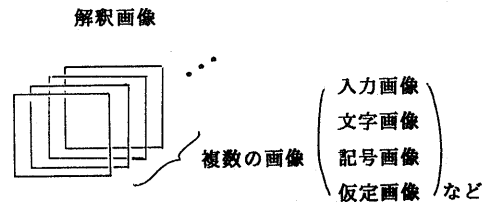


図1 解釈画像

解釈画像を用いた認識処理では、解釈画像の更新手続きに弛緩法を利用することができる。そして、「画像のある領域に解釈を与える」という動作が認識の状態決定であると考えられることができる。文字と図形の分離処理では、これまで「画像を文字を表わす領域と図形を表わす領域に分割する」としていた処理とは異なる考え方ができる。そして、領域（画素）に複数の解釈を与えることもできる。さらに、解釈画像を用いれば、認識の解釈関係を操作する処理を実現しやすい。例えば、流れ図図面を対象とした認識処理では、流れ図記号の中に書かれた文字とか、逆に文字を囲んでいる流れ図記号というものを考えて、そこに画素を探していく処理がある。流れ図記号のうち判断記号がわかったとき、その出口付近に条件を表わす文字が書かれているのではないかという予想にもとづき、文字を構成する画素を発見する処理もある。

認識処理の過程において、ある部分が明確に認識されていないために他が処理できないということがある。これを解決するために、認識が未処理の部分に対して都合のよい解釈を与えた画像（仮説）をつくることにより、認識処理を暫定的に進めることができると考えた。このつくられた画像を仮定画像と呼ぶことにする。図面認識では、いく通りかに解釈できる部分について何か解釈を決めてしまうことであり、図面のその箇所の意味を仮定して表現したものであると考える。

3. 文字と図形の分離アルゴリズム

弛緩法により解釈画像を更新する「文字と図形の分離アルゴリズム」を示す。以降では、処理の概略フローを示し、そして、処理の詳細と具体的な入力画像に対する処理結果を示す。実験は、2値化された比較的低解像度の流れ図図面を入力として、それから文字領域部分、図形領域部分の分離を行なうことを目標とした。なお、対象とする図面に次の制限をつけた。

- (1) 流れ図記号は、端子、処理、判断の3種類とする。
- (2) 判断記号は、入力の流れ線が1本、出力の流れ線が2本のものに限る。
- (3) 流れ線は交差しない。

3.1 概略フロー

分離アルゴリズムの概略フローを図2に示す。解釈画像として入力画像、文字画像（文字領域であると認識された領域）、図形画像（図形領域であると認識された領域）、図形の解釈に対する仮定画像を考える。処理は、解釈画像を解釈の状態を用いて更新する弛緩法により実現した。まず始めに、弛緩法の初期状態である文字画像と図形画像の初期像をつくる。次に、図形画像から得られる流れ図の意味を表現した仮定画像の生成を行ない、これを解釈画像の更新に利用する。弛緩法は文字と図形の分離が矛盾なく行われた場合に終了する。

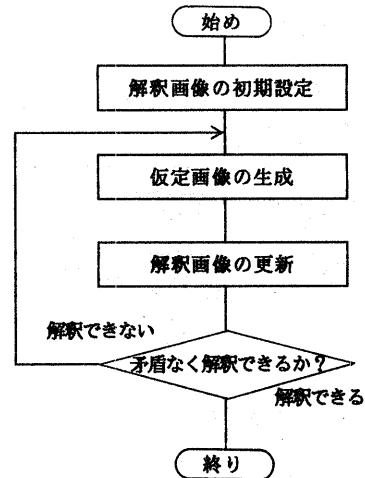


図2 概略フロー

3.2 アルゴリズム

まず、アルゴリズムの中で用いた概念を説明する。

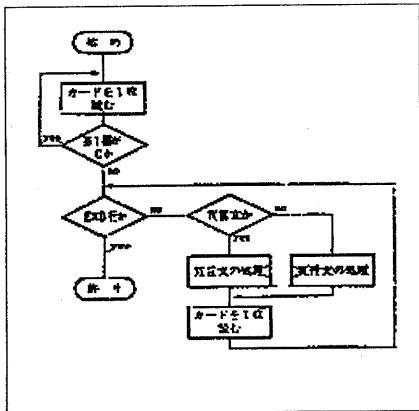
(1) 流れ図記号のポート

流れ線が流れ図記号に入るところを「入力ポート」、流れ線が出るところを「出力ポート」と呼ぶ。流れ図記号のポートの位置は経験的に定まっている。

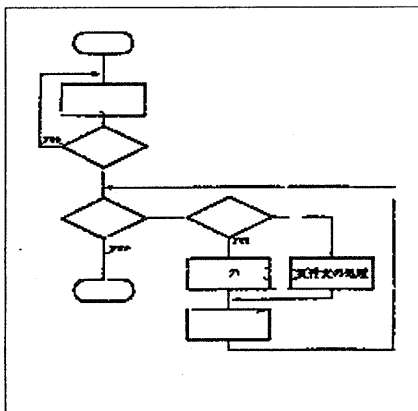
(2) 2値画像の近傍濃度

2値画像上の任意の点について、その近傍の黒画素の数を近傍濃度と定義する。2値画像が数値 0/1 で表現されていれば、単なるフィルタリング操作となる。そして2値画像上のすべての点について近傍濃度を求めた結果の濃淡画像を近傍濃度画像と呼ぶ。

入力原稿は、活字による文字とテンプレートをを用いて描かれた図形からなる。画像1は、TVカメラにより入力した。解像度が 256 × 256 で 8bit 階調の濃淡画像を、シェーディング補正した後、2値化したものである。そして、いくつかの図形のとぎれ、および文字と図形の接触を付加した。以降に、処理の詳細と結果を示す。



画像1 入力画像



(a)

画像2 図形画像と文字画像の初期像

処理 1-1

入力画像から文字領域部分であろうと思われる画素を抽出するために、次の処理を行なう。

- (1) 入力画像の連結領域にラベル付けを行なう。
- (2) 大きな連結領域を図形画像候補とする。
- (3) 小さな連結領域を文字画像候補とする。

処理 1-2

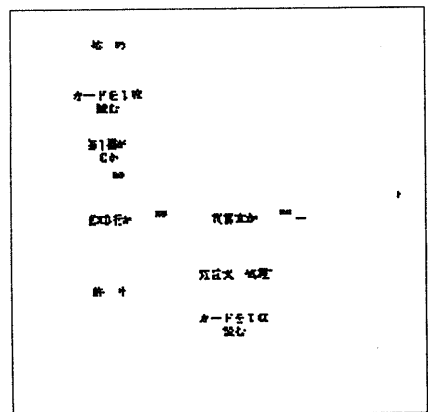
文字画像候補と図形画像候補の情報を使って、次の処理を行なう。

- (1) 文字画像候補中の連結領域で横方向に隣接しているものをグループ化する。
- (2) グループ化した領域の左右の領域に対して図形画像候補から塊状画素を探す。
- (3) 塊状画素が新たに見つかったときは、見つけた塊状画素を文字画像候補に加えて、(1) から処理を繰り返す。

塊状画素の検出により、本来文字領域であるが図形と接触しているため、小領域として検出できなかった領域の一部が発見できる。

処理 1-3

入力画像から文字画像候補を除いたものを図形画像候補とする。これで、図形画像と文字画像の初期像が生成できた。処理結果を画像2 (a) と画像2 (b) に示す。



(b)

処理 2-1

図形の構成要素を取り出すために、次の処理を行なう。

- (1) 図形画像に分割化演算を施し、図形を直線線分などのセグメントに分割した画像を得る。
- (2) その画像を細線化する。
- (3) 線分を検出し、線分の属性（位置，方向，長さ）を得る。

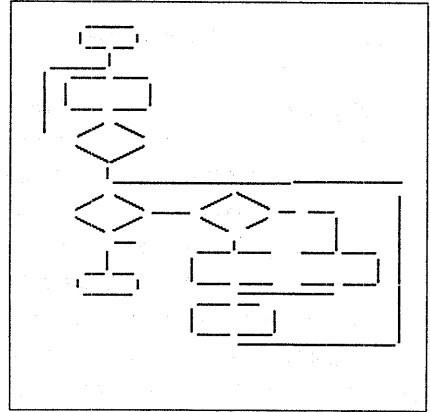
分割化演算は，図形画像を細線化した画像に対して近傍濃度画像を求め，近傍濃度の大きい領域を除く。そうして得られた画像中で元の画像にも存在する画素だけを残す。近傍濃度の大きい領域は，図形画像の中で，流れ図記号の角，流れ図記号と流れ線の交点，および抽出漏れとなっている文字領域部分の周辺，などに存在している。さらに，これから直線線分を取り出した状態が画像3 (a) である。

処理 2-2

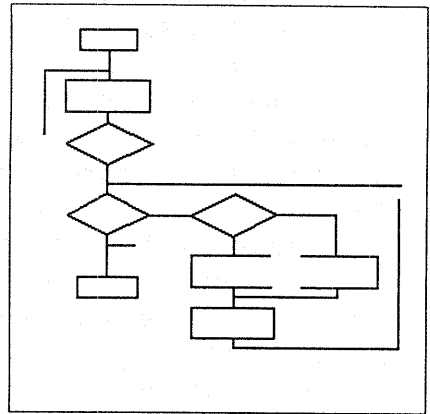
局所的知識による図形の接続を，次の (1)～(3) に従ってこの順に行なう。

- (1) 流れ図中の判断記号による閉ループは構成線分が斜めの方向の線分 / と \ によるものであることから，確実に検出することができる。
- (2) 方向が同じ（水平または垂直）線分を2本考えたとき，その一方の始点と他方の終点が近くに存在しているとき，その両方の線分を適切に延長する。さらにこのとき，2本の線分が1本の線分に並ぶように互いに座標を調整する。
- (3) 方向が異なる2本の線分の各々の端点が近くに存在していて，それぞれの端点の延長によって2本の線分が交わる時，その両方の線分を適切に延長する。

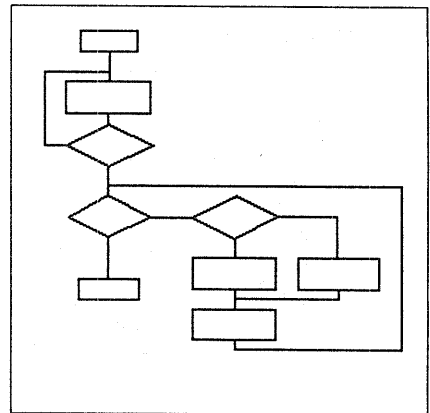
局所的知識による線分の延長はいずれの場合も確実に接続すべき箇所の一部となっている。すなわち，これらの延長によって接続されてはならない箇所が誤って接続されることはない。



(a)



(b)



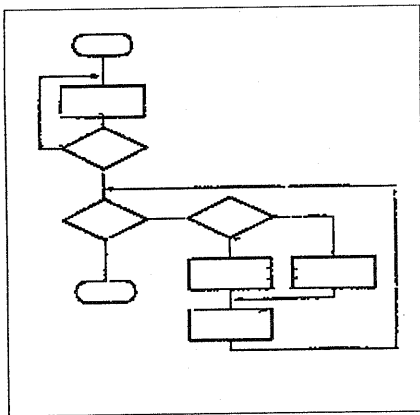
(c)

画像3 仮定画像の生成

処理 2-3

大局的知識による図形の接続を、次の (1)～(5) に従ってこの順に行なう。

- (1) 処理 2-2の(1) によりいくつかの判断記号が見つけられている。これから各判断記号のポートを求める。処理記号または端子記号のポートは、処理 2-1で分割化された後に、処理 2-2の(2) で接続されていると考えられる。これを処理記号または端子記号のポート候補としておく。これと先の判断記号のポートから、ポートの並び方の知識を用いてポートの集合を決定する。
- (2) 判断記号のポートに向かう線分があつて線分の始点または終点がポートに近いとき、適切な線分を追加する。またこのとき、ポートに向かう線分の位置を適切に調整する。
- (3) ポートの集合の中で3本の線分が交わっていない点を未処理のポートとして検出する。線分の中でその始点または終点が他のどの線分とも接続していない線分を未処理の線分として検出する。
未処理の線分の未処理である端点と未処理のポートが近くがあれば、適切に線分を追加して接続する。このとき、未処理であった線分の端点の座標を調整する。
- (4) 未処理のポートに向かう線分があつて線分の始点または終点がポートに近いとき、適切な線分を追加する。このとき、ポートに向かう線分の位置を適切に調整する。



(a)

画像4 新しい図形画像と文字画像

- (5) 未処理のポートと他の未処理のポートが近くにあつて並んでいるとき、この2つのポートをつなぐ適切な線分を追加する。このとき、ポートの座標の調整をする。

このときの仮定画像の状態は画像3(b)である。

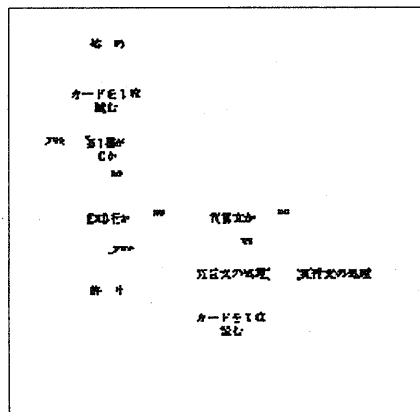
処理 2-4

未処理のポートと未処理の線分の各々について、考えられる接続と削除によって得られる組合せを調べる。流れ図記号と流れ線の接続状況について矛盾のない流れ図が得られれば、仮定画像の生成が完了する。仮定画像の生成結果を画像3(c)に示す。

処理 3-1

次の (1)～(3) により新しい図形画像と文字画像を生成する。

- (1) 仮定画像を拡散したものと図形画像の共通部分を取り出す。これを文字画像から除いたものを拡散する。
- (2) 文字画像から (1)の画像を除いたものを、図形画像に加える。文字画像から (1)の画像を除く。
- (3) 図形画像を拡散した画像を入力画像から除いた画像を得る。それを入力画像に沿って拡散する (これは、拡散した画像と原画像との共通部分を取る算を繰り返す)。これにより得られた画像を文字画像に加える。



(b)

ここでは文字画像中にあった図形部分であろうと考えられる画素を図形画像に移す操作 (1), (2) を行なう。さらに、図形画像中にあった文字部分であろうと考えられる画素を文字画像に移す操作 (3) を行なう。これにより、新しい図形画像と新しい文字画像を得る。

この処理の結果を画像 4 (a) と画像 4 (b) に示す。

処理 3-2

認識処理の収束判定は、ここまでで得られた文字画像と図形画像の整合性をチェックすることにより行なう。次の (条件 1) と (条件 2) をともに満たすことを収束条件とした。

(条件 1) 認識されたそれぞれの流れ図記号の中に文字を見つけることができる。

(条件 2) 判断記号から出る流れ線の近くに文字を見つけることができる。

収束していればアルゴリズムは終了、そうでなければ処理 2-1より繰り返す。

実験では、2つの条件は満たされているから、処理は終了する。

4. アルゴリズムの評価

分離アルゴリズムによる実験をとおして、いくつかの問題点、課題が明らかになった。

4.1 仮定画像の生成について

仮定画像の生成においては、図面の意味（とくに、流れ図記号と流れ線の大局的な連結関係）から、仮定をパーズングできるはずである。ここで、図面の記述に関する知識を与えて、それから一つの正しい（知識に矛盾しない）解釈を見つけることを「パーズング」と呼んでいる。具体的な実現は、線画にラベル付け可能な集合（ラベル付けの条件）を与えて、矛盾しないラベル付けを見つける処理などが考えられる。

今回のアルゴリズムでは、仮定画像の生成を、図形画像を分割化処理した後で得られる図形構成線分の位

置と方向の関係を用いて行なった。それは生成過程での線分の延長、追加、削除の処理が十分に経験的であることと、入力を解釈に曖昧さのない図面に限ったことによる。ところが、仮説となる仮定画像は、大局的な認識により整合性を検証されなければならない。さらに、仮定画像は一つとは限らない。このような仮定の生成の制御は今後の課題の一つである。

4.2 認識の収束判定について

総合的な認識処理を考えると、認識の収束（弛緩法の収束）や認識の目標は、外からの処理要求によるものでなければならないと考えられる。そのような、要求を明確に定義できることが必要である。

認識処理の応用を考えると、文字と図形の分離抽出処理はこれだけで閉じたものではない。分離抽出処理は、分離の結果を利用したい認識処理からサブルーチンの使われ方をする。このことを考えると、認識（分離）の収束条件は分離抽出処理を起動した認識処理の要求により決定されるべきであると思われる。今回の実験では、解釈に曖昧さのない図面を入力して、文字領域と図形領域の分離結果を得ることを目標としている。そのため、収束条件は「満たされるであろう条件」となっている。ところが、実際の認識処理への応用では、「条件を満たすことはできそうにない」という状態での終了を定義できる必要があると考えられる。また、条件を全部満たすべきなのか、一部の条件が成り立てば良いのかというような点も認識の進展にともなって動的に変更される必要があろう。

4.3 パラメータの値の決定

アルゴリズムの中にはいくつかのパラメータが現れる。2値化処理のしきい値、距離や面積に関する述語で使われる値、拡散・収縮の繰り返し回数、などである。そこでは、入力画像中にある流れ図記号の大きさ、流れ図の図形部分の線の太さ、1文字の大きさ、等をもとに決めているが経験による値も多い。本来ならば、入力画像の縦および横方向の画素数のヒストグラムなどを用いて、いくつかの基本となる定数を決める処理を行なうべきであろう。そして、この基本の定数の調

整に関するループが必要となる。本研究では、この部分の処理は割愛した。

4.4 入力画像の参照の問題

仮定画像の生成は入力画像を参照せずに行なった。これにより画素レベルのデータから離れて抽象化したデータに対する操作として処理を行なうことができる。処理は柔軟に行なえるが、入力画像中の情報を利用する可能性を失っている。考えられるすべての組合せに対して仮説を生成するのではなく、入力画像中の画素の情報と照合して適切なものだけを生成することが望ましい。この実現のためには、画像の画素レベルのデータと認識の状態の関係を的確に表現できる必要があるだろう。

5. おわりに

本報告では、認識処理の総合的な制御を行なうために、次の2つの提案をした。

- (1) 解釈画像を用いた認識状態の記憶法の提案
- (2) 仮説である仮定画像の生成とその検証による制御法の提案

本研究ではこれらの提案を文字と図形の分離アルゴリズムとして具体的に実現した。そして、流れ図図面を入力とした実験をとおして、仮定画像の生成や処理の終了に関して問題点が明らかになった。ただし、解釈画像の更新である弛緩法の図面認識処理に対する有効性の検討は今後の課題である。

謝辞

本研究に関して御助言、御討論を頂きました静岡大学中谷広正助教に、感謝いたします。

なお、本研究の一部は文部省科学研究費特定研究（多元知識情報）No. 61102005，および一般研究No. 61550259 によった。

参考文献

- (1) 白井：“シーンに関する知識を用いて明るさの配列を解析する”，「コンピュータビジョンの心理」，第3章，産業図書（1979年）。
- (2) 松山，ハング：“画像理解システムSIGMA — ボトムアップ，トップダウン解析の統合 — ”，情報処理学会論文誌，Vol. 26, No. 5, pp. 877-889（1985年）。
- (3) 森本，有木，坂井：“図面の構成知識を用いた図面構造解析”，電子通信学会技術報告，PRU86-12（1986年）。
- (4) 白井：“コンピュータビジョン”，第7章，昭晃堂（1980年）。