Osaka University Knowledge Archive

Title	認識戦略の自動生成に基づく図面解釈
Author(s)	加藤, 博一
Citation	大阪大学, 1996, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3110151
rights	
Note	

Osaka University Knowledge Archive : OUKA

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

Osaka University

認識戦略の自動生成に基づく図面解釈

加　藤　博　一

1995年11月

大阪大学基礎工学部

認識戦略の自動生成に基づく図面解釈

加 藤 博 一

１９９５年１１月

大阪大学基礎工学部

要　約

図面解釈は，2次元画像情報として入力された図面データから，その図面に描かれている情報を計算機が自動的に読みとるための方法を考える問題である．これは，人間の視覚情報処理過程の模倣を試みる画像理解研究としても位置付けることができる．画像理解研究には解決すべき多くの課題が残されているが，その中に，知識に基づく知的で柔軟な認識機能の実現がある．本論文は，図面に対象を限定することで，画像理解研究の課題からこの点を抜き出し，その課題の抱える問題点の本質の明らかにし，その解決策を検討するものである．

まず，多くの知識を積極的に利用した楽譜解釈システムを構築した．知識を有効に活用するために黒板法に基づく階層型の処理形態を導入した．実験結果として，これまでの楽譜解釈システムでは取り扱うことのできなかった複雑な楽譜に対しても，満足できる結果が得られ，知識を積極的に利用したトップダウン処理の有効性を示すことができた．

次に，処理の汎用化と画質の劣る図面への対応を目指し，処理と知識を分離したシステムに関して検討した．記号で表現された知識の取り扱いは，知識工学の分野で数多く研究されている．そのような研究の成果を利用することにより，処理の汎用化を目指した．画質劣化に伴い，パターン処理に対する性能向上が要求されるが，そのためにも知識の積極的な利用が必要となる．しかし，パターン処理，特に切り出し処理において知識を積極的に利用した方法論は確立されていない．そこで，パターン処理に対しての積極的な知識の利用を諦め，記号処理において対処することを考えた．実験では，画質の劣化により意味的に曖昧さを含む図面に対しても良好な結果が得られた．しかし，大きな図面においては取り扱う仮説が膨大になり，記号処理において組み合わせ爆発が生じ，動作不能となることが多かった．

この問題を解決するために，対象の特性に応じた処理が可能な高性能な切り出し処理に関して検討を行い，弾性モデルに基づく新たな切り出し手法を提案した．さらに，この切り出し手法は，パターンの構造解析において重要な役割を果たすことができた．パターンの構造解析は，対象とするパターン毎にそのパターンの特性に応じた処理を行うための方法論であるが，処理と知識の分離を考えた場合，様々な要求に対応できる汎用的な構造解析の処理の枠組みを準備しておかなければならない．このような問題に弾性モデルによる切り出し手法が非常に有効に機能することを示した．

最後に，弾性モデルによる切り出しと構造解析を利用した汎用的図面解釈の方法論の提案を行った．記号処理部において動的に認識戦略を生成し，それに従った認識処理をパターン処理部において弾性モデルに基づく汎用的構造解析処理が行うという，人間の注意制御機構を模倣した構成を採用した．実験により，この手法が画質の点においても，種類の点においても汎用性があることを示した．

以上の実験・検討結果より，画質の劣化した図面に対応できる汎用性のある図面解釈の方法論を示したと共に，画像認識・理解の研究に対して，

・問題固有の知識を積極的に利用できる汎用的パターン処理に関する検討

・部分問題としての様々な認識処理に対し，最適な問題設定のための方法論の確立

の2点の重要性を明らかにした．

目　次

1章　序　論 1
1.1　パターン認識と図面解釈 1
1.1.1　パターン認識 1
1.1.2　パターン理解と解釈 2
1.1.3　図面解釈 3
1.2　工学的側面から見た図面解釈 3
1.2.1　図面解釈の歴史 3
1.2.2　図面解釈の基本的処理 4
1.2.3　現状と問題点 5
1.3　科学的側面から見た図面解釈 6
1.3.1　コンピュータ・ビジョンにおける物体認識と図面解釈 6
1.3.2　知識の有効性 7
1.4　本論文の取り扱う問題 8

2章　トップダウン処理による図面解釈 11
2.1　トップダウン処理による図面解釈 11
2.1.1　トップダウン処理とボトムアップ処理 11
2.1.2　前向き推論と後ろ向き推論 13
2.1.3　図面解釈におけるトップダウン処理の有効性 13
2.2　楽譜認識システム 14
2.2.1　楽譜認識技術の動向 14
2.2.2　トップダウン処理による楽譜認識 16
2.2.3　処理の概要 16
2.3　前処理 17
2.3.1　画像入力 17
2.3.2　五線検出 17
2.3.3　小節線検出 19
2.4　小節単位での認識処理 19
2.4.1　五線消去 20
2.4.2　音部・調子・拍子記号の認識 20
2.4.3　音符・休符・臨時記号の認識 21
2.4.3.1　階層ワーキングメモリ 22
2.4.3.2　処理モジュール 24
2.4.3.3　認識戦略 26
2.5　後処理 26
2.6　実験 27
2.7　まとめ 29

3章　知識処理に基づく図面解釈 31
3.1　認識・解釈処理と知識の分離 31
3.1.1　汎用化の２つの軸 31
3.1.2　認識・解釈処理と知識の分離 32
3.2　論理回路図面解釈システムのコンセプト 33
3.2.1　論理回路図面の特徴 33
3.2.2　解決すべき問題点 33
3.2.3　階層的知識の相互利用 34
3.3　処理の概要 34
3.3.1　階層的知識構造 34
3.3.2　概要 36
3.4　シンボル仮説の生成 37
3.4.1　プリミティブ仮説の生成 37
3.4.2　シンボル仮説の合成 39
3.5　図面仮説の生成 39
3.5.1　前向き推論処理 39
3.5.2　後向き推論処理 40
3.6　図面仮説の検証 41
3.6.1　矛盾検証 41
3.6.2　確信度計算 42
3.7　実験と検討 42
3.8　まとめ 45

4章　弾性モデルに基づく形状切り出し 47
4.1　汎用的切り出し手法 47
4.1.1　図面解釈における形状切り出し 47
4.1.2　領域抽出と領域分割 48
4.1.3　汎用的切り出し手法 48
4.1.4　パターンマッチング手法の利用 49
4.2　弾性モデルによる切り出し 51
4.3　弾性モデル 52
4.3.1　モデル表現 53
4.3.2　弾性モデルへの変換処理 53
4.4　ポテンシャル場の生成 54
4.5　落下運動シミュレーション 55
4.6　領域の抽出 56
4.7　実験と検討 58
4.7.1　歪んだ図形への対応付け 58
4.7.2　ノイズを含んだ図形の切り出し 58
4.7.3　汎用性に対する考察 58

4.7.4 面領域への対応 59
4.7.5 初期位置の決定問題 59
4.8 まとめ 59

5章 パターン構造の表現と解析 61
5.1 トップダウン型パターン処理の問題点 61
5.2 パターン構造の表現と解析 62
5.2.1 パターン表現を用いた構造記述 62
5.2.2 切り出し処理の改良 64
5.3 実験 66
5.3.1 部分構造の切り出し実験 66
5.3.2 部分構造の解析 67
5.3.3 仮想領域の効果 69
5.4 まとめ 70

6章 認識戦略の自動生成に基づく図面解釈 71
6.1 汎用的図面解釈システムにおけるパターン処理 71
6.2 注意制御に基づく認識処理 73
6.2.1 人間のパターン認識 73
6.2.1.1 受動的認識処理 73
6.2.1.2 注意制御による能動的認識処理 74
6.2.2 図面解釈における注意制御 75
6.3 前向き推論型パターン処理による注意仮説の生成 76
6.3.1 注意仮説 76
6.3.2 基本アルゴリズム 76
6.3.3 テンプレートマッチングの改良 79
6.3.4 類似度法の検討 80
6.3.5 実験 82
6.3.6 検討 82
6.4 状況適応型認識処理 84
6.4.1 注意仮説に基づく認識処理 84
6.4.2 競合する注意仮説情報に基づく認識戦略 84
6.4.3 形状起因度に基づく認識戦略 85
6.5 認識戦略の自動生成 85
6.5.1 シンボルの記述 85
6.5.2 認識戦略の生成アルゴリズム 86
6.5.2.1 競合の解消 87
6.5.2.2 起因度に応じた処理 89
6.5.3 認識戦略の実行アルゴリズム 89
6.6 シンボルのモデル記述 90
6.6.1 注意仮説生成用のモデル記述 91

6.6.2 注意仮説検証用のモデル記述 91
6.6.3 その他の情報の記述 94
6.7 実験 94
6.7.1 処理対象図面 94
6.7.2 モデル 94
6.7.3 実験 95
6.7.4 検討 101
6.8 まとめ 101

7章 結 論 103

1章　序　論

図面解釈は，2次元画像情報として入力された図面データから，その図面に描かれている情報を計算機が自動的に読みとるための方法を考える問題であり，パターン認識研究の1つの応用分野である．様々な情報メディアの電子化が要求されるマルチメディア社会に向かっている現在，図面は文字と同様，現代社会における必要不可欠な情報伝達メディアであるため，その成果に対する期待は大きい．それ故にこれまで大規模な研究開発が産業界を中心に行われてきた．また，図面解釈は画像理解研究の前哨戦として位置付けることもでき，コンピュータ・ビジョンや人工知能の分野からの関心度も高い．本論文は主に後者の立場から，図面解釈を捉えたものである．まず1章では，図面解釈の定義を行い，このような図面解釈問題の背景に関して説明する．それに基づいて本論文で取り扱う問題領域と意義について述べる．

1.1　パターン認識と図面解釈

1.1.1　パターン認識

パターン認識は，計算機科学における重要な研究分野の1つとしてすでに定着しており，工学としても多くの実績を残してきた．まず，このパターン認識という言葉の意味について考え，その周辺の用語について本論文での意味を定義する．

パターンとは，計算機科学においては，空間的，時間的に分布したデータとして定義される[Anzai89]．また，パターンと対比して使用される言葉に，シンボル（記号）という言葉がある．この対比で説明すると，パターンとは複数の要素で構成される情報において個々の要素が重要な意味を持たず，それら複数の要素の構造に重要な意味を持つ情報のことで，シンボルとは実体や概念を表現するための個々の要素のことであり，その要素自身が重要な意味を持つものである．この定義を用いると，パターンとシンボルは背反する概念ではない．計算機で取り扱う情報の要素はすべてシンボルであり，パターンはそのシンボルの集合として構成されていると言ってもよい．そのとき，個々の要素の意味を重視するのか，それともその構造の意味を重視するのかにより，シンボルとパターンの使い分けが行われる．つまり，シンボル処理とは個々の要素の意味を重視した処理であり，パターン処理とは複数の要素の持つ構造を重視した処理といえる．

認識とは，recognitionの訳として使われる言葉であり，re（再び）＋cognition（認知）という成り立ちである．認知とはある対象をそれとして認めることであり，理由や原因などに関係なく，実体や概念それ自体の存在を認めることである．つまり，認識とはある与えられた対象を一旦認知した実体や概念（これらは知識として記憶されている）のどれかと同一のものと見なすことである．

パターン認識とは，この両者の定義を組み合わせたものであり，パターンとしての対象を認識することである．また，次に説明する理解や解釈，さらにはそれらの共通問題である切り出し問題をも含めて，パターン認識という言葉で研究領域全体を示す広義の意味もある．

1.1.2 パターン理解と解釈

パターン理解という言葉は,パターンを理解することという直接的な意味で用いられているが, それ以上に深い定義はなされていない. そもそも, 理解という言葉はパターン認識, コンピュータ・ビジョン, 人工知能に関連する分野において頻繁に用いられているが, その標準的な定義はない [Miyake91]. 一般には, 従来のパターン認識を知的化したものとしてパターン理解という言葉を利用する場合が多い. これは画像認識と画像理解という言葉の対にも共通する [Yachida89]. ここでは, 理解という言葉と共に, 本論文の題目としても使用されている解釈という言葉の意味を考えてみる.

1.1.1で述べたように, 認識とは入力情報を既知の知識と同一のものと見なすことである. 人間は視覚や聴覚からの入力信号からなんらかの情報を認識し, それに基づき行動している. それら入力信号には多くの情報が含まれていて, それらに対して様々な認識が行われている. 例えば, 多少の音楽知識のある人が図1-1のような楽譜を見たときを考えると, まず音符や休符など楽譜の構成要素である個々の記号の認識が行われるであろう. それと同時にそれらの位置的な (おもに5線を基準とした) 関係も認識するであろう. このような個々の記号の認識とそれらの関係の認識の両者を実行できてはじめて, 音符や休符の持つ時系列的な演奏情報を認識できる. さらにそれらの音楽的関係を認識することにより, フレーズ認識など楽曲構造の認識が行える. つまり, 人間の認識過程においては, 入力されたデータから様々な情報の認識が構造化・階層化されながら実行される. このような認識を大きく2つに分類すると, 対象自体の認識, 複数の対象間の関係の認識に分けることができる. 本論文では, 対象自体の認識を狭義の認識, 対象間の関係の認識を解析と呼ぶ. さらにこの認識と解析を行うことで情報の構造化を行う過程を解釈と呼ぶ. 階層的に見た場合, この解釈過程が一つの認識になる場合もあれば認知となる場合もある.

理解という言葉について考えてみると, その直感的な意味としては対象の本質の認識という捉え方ができる. これは今述べた用語を用いると, 階層的な解釈過程を通じて, 深い階層にお

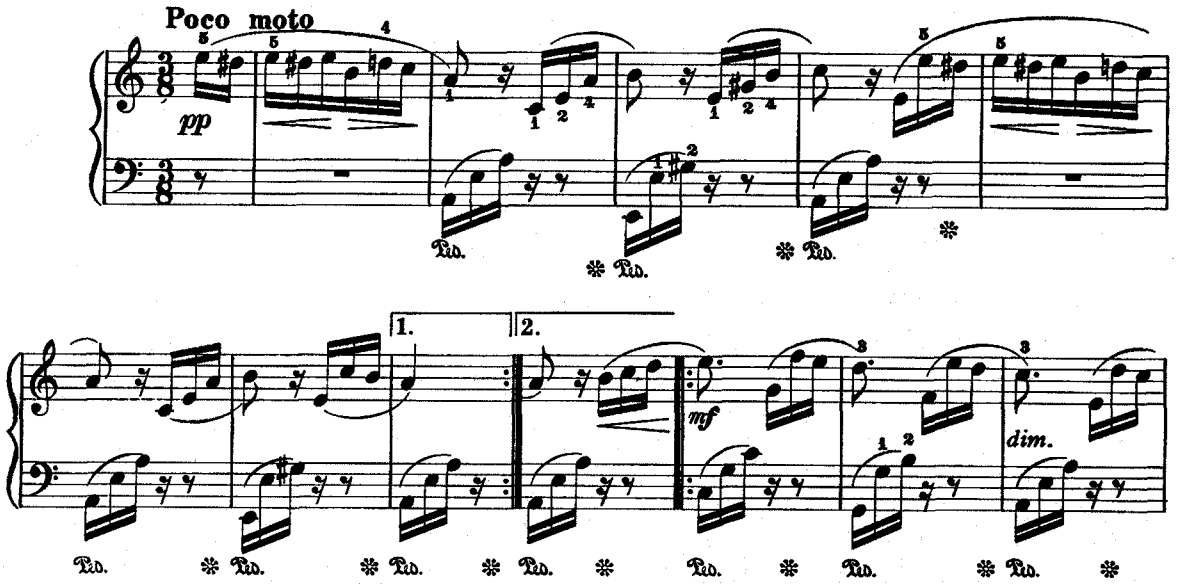


図1-1 楽譜図面の例

ける情報の構造化を行うことと定義できる．ここで深いという相対的な表現を用いたのは，理解という言葉自身が実際に主観的な用語だからである．例えば，図1-1において音の高さと長さという演奏情報を解釈することが理解なのか，その楽曲の音楽的構造の解釈ができて理解なのか，またその作曲家のその曲に込めた意図までもが解釈できて理解なのか，その基準は主観的である．この主観性が，理解が認識の知的なものという相対的な概念として用いられる原因である．

1.1.3 図面解釈

本論文の題目として使用している図面解釈という言葉における解釈には，1.1.2で述べた解釈の定義をそのままあてはめることができる．つまり，図面解釈とは，入力として与えられた図面から個々の要素の認識を行い，それら要素間の関係を解析することで，図面に含まれる情報の構造化を行うことである．このように図面解釈はパターン認識の一つの応用問題として捉えることができる．しかし，それ以外にも図面解釈の重要な捉え方がある．1.2, 1.3では，異なる側面から見た図面解釈について説明する．

1.2 工学的側面から見た図面解釈

図面は，文字と同様現代社会における必要不可欠な情報伝達メディアである．産業社会から情報社会への移行の過渡期である現在では，様々な情報の電子化が推し進められており，図面メディアの電子化も盛んに行われている．この状況において図面解釈研究の成果に対する期待も非常に大きくなり，大規模な研究開発が産業界を中心に行われてきた．ここでは，そのような図面解釈研究に関しての説明を行う．

1.2.1 図面解釈の歴史

計算機による図面の自動入力を目指した図面解釈の研究が始まったのは1970年代である [Dezawa76] [Jarvis77]．しかし本格的に研究が行われるようになったのは1980年代半ばからであった [Tsunekawa85]．イメージスキャナなどの画像入力装置や大きな画像を処理するのに必要な大容量メモリ・高速処理を備えた計算機が一般に普及しはじめたのが1980年代に入ってからであり，それと同時に計算機を使ったCAD/CAMシステムの導入が始まった．その結果様々な図面に描かれた設計情報の自動入力を目的とする図面解釈のニーズが1980年代半ばになって急速に高まり，このようなニーズに応えるべく多くの研究が行われるようになったのである．

これまで研究の対象となっている図面は多種多様にわたり，それらの特徴により分類すると図1-2のようになる [Shirai87]．個別に見ていくと，地形図では図中に表されている線をトレースしベクトル化するという処理が行われている．同じ形状図である市街図ではそのベクトル化されたデータを用い，さらに道路や家屋の抽出といったより高度で困難な処理を含んでいる．機械図面では，形状の認識だけでなく寸法線・寸法補助線なども認識し，それらの数値情報との関連付けを行う必要性がある．天気図や楽譜では，シンボルと背景図形との位置関係が非常に重要であり，シンボルの認識を行う際には背景図形をあらかじめ除去しておかなければなら

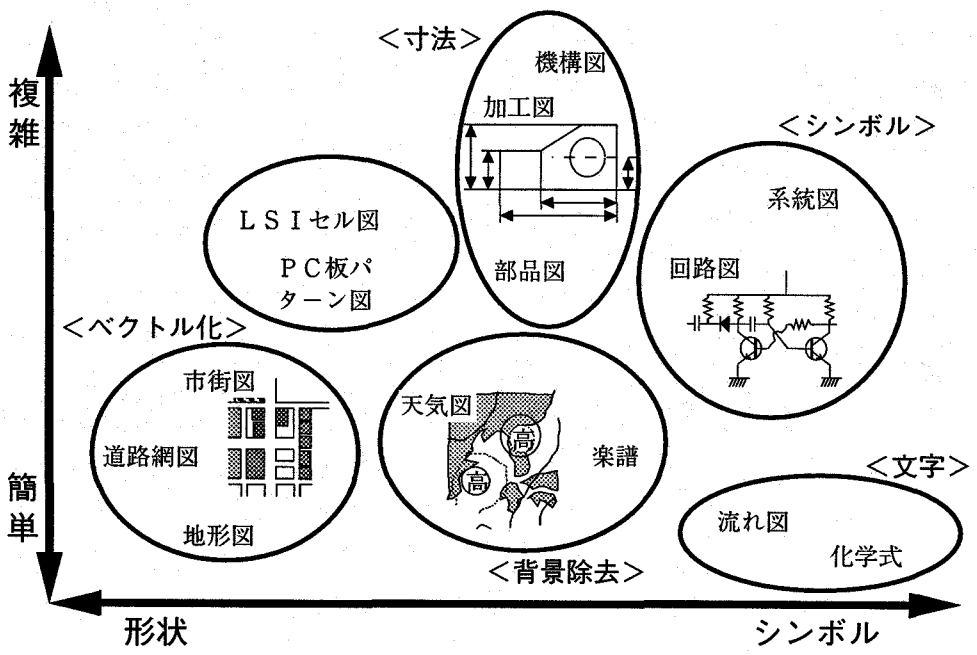


図 1-2 図面の種類 [Shirai87] より

ないという問題もある．回路図のように，主として部品の接続関係の記述を目的とする図面では，各種のシンボル図形の認識が課題となる．さらに流れ図などでは，文字の分離抽出および文字認識が主な課題となる．このような多種類の図面に対して，それぞれの持つ問題の解決を図りながら，研究が進められたのが1980年代であった．

ところが1990年代になると図面解釈の研究はその数が減少するようになった．これは，図面解釈の研究が行き詰まったこと，手動入力による人件費の増加を考慮しても採算のとれるほど図面の電子化のニーズが高まったこと，さらにはヒューマン・インタフェース技術の進歩により手動での図面入力自体の効率も改善されたことなどが考えられる．

1.2.2 図面解釈の基本的処理

1.2.1で述べたように図面解釈には様々な対象が存在するが，全体の処理の流れを一般化すると図1-3のようになる．その処理の中心となるのは，次の3つである．

1) 文字・シンボル・線（構造線，接続線）の3要素の分離抽出処理
2) 3要素それぞれの認識処理
3) 3要素の統合処理

1,2が図面の構成要素の認識，3がそれらの関係解析に対応する．これらの処理を実現するためには，画像処理を始めとする様々な処理を組み合わせる必要がある．特にいろいろな処理方法をどう組み合わせて制御するかが，図面解釈の効率・能力を決定する大きな要因となる．また知識に基づき不完全な情報の内挿を行うことや，いろいろなレベルでの処理を統合することなど，図面解釈の制御に関わる問題は数多く存在する．

実際の図面解釈処理では1,2,3の処理が手順通りに行われることはまれで，多くの場合は誤りを訂正するために何度も処理を繰り返しながら正しい認識結果を導き出すことになる．

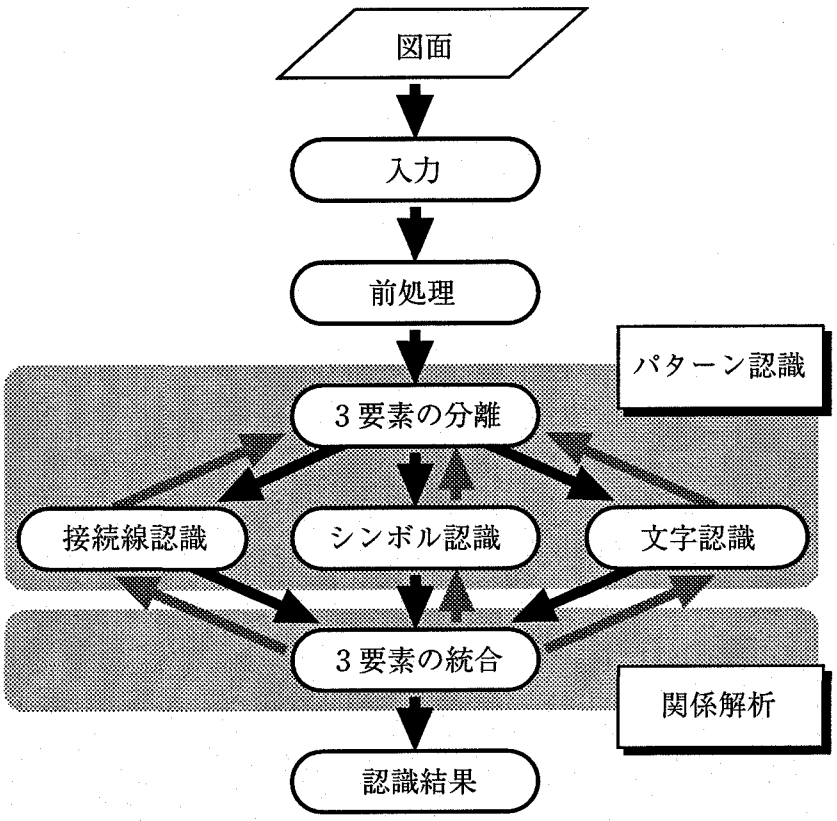


図 1-3 図面解釈の処理手順

1.2.3 現状と問題点

図面は人間と人間の情報伝達に用いられるものであった．それ故に，図面は，読み手の解釈しやすさを意識して描かれる．図面解釈システムにおいても，精度よく解釈できる図面とそうでない図面とがある．ところが，人間の解釈しやすさと計算機による解釈しやすさの基準は，全く異なっている．例えば，人間なら形状の多少の変形や線の途切れに対してはなんの問題もなく対処できるが，図面の構図の複雑さや図面自体の大きさにより，解釈に掛かる負担は増加する．つまり，質の変化より量の変化に弱い．逆に，図面解釈システムの場合では，図面の構図の複雑さや図面自体の大きさの精度に対する影響は少ない．しかし，形状の歪みや線の途切れなどに対しては敏感である．つまり，量の変化より質の変化に弱い．このような人間と図面解釈システムの特性の違いが，利用者のシステムに対する期待と実際の性能の間に大きな隔たりを生み，システムの実利用の普及が停滞する要因となっている．

ところで，図面情報の電子化に対して要求される処理に，図面解釈が必ずしも必要とは限らない．構成要素の認識ができれば，それらの関係解析が必要でない場合もある．さらには，単に線分の折れ線近似や円弧近似（そのような処理をベクトル化と呼ぶ）ができればよいという場合もある．特に近年，図面の電子化処理の中で最も注目されている地図情報処理システム [Kamata89] においては，図面の高精度なベクトル化が最大の要求事項であり，それ故，図面解釈の研究でなくベクトル化と構成要素の認識に関する研究が多く行われるようになった．しかし，それでもなお上述の問題があり，完全な自動化が行えずに，手動による入力や修正が重要になっている [Sakauchi92].

1.3 科学的側面から見た図面解釈

計算機による人間の視覚機能の模倣という問題は,非常に大きな問題であり簡単には解決できない. そこで, 様々な部分問題へと分割し, その部分問題を解くことがまず必要である. 図面は, 人間の視覚機能のうち知性に支えられる機能, つまり知識に基づくトップダウン的な視覚機能の研究に対する格好の素材である. ここでは, そのような立場から図面解釈について考えてみる.

1.3.1 コンピュータ・ビジョンにおける物体認識と図面解釈

コンピュータ・ビジョンにおける物体認識とは, 画像中から特定の対象を探し出したり, その位置姿勢を決定する問題である. これは図面解釈における, 構成要素の認識問題に通じる. 物体認識問題では, 知識を積極的に利用したモデルベース認識手法が最も有効とされている. そこでは,与えられた物体モデルからどのようにして効率的な認識アルゴリズムを生成するかが重要である. これには以下の3種類の方式が考えられる [Ikeuchi89].

1) **ハンドコンパイル方式**

プログラムの設計者が効率的なアルゴリズムを考案し, プログラムに埋め込む方式. 性能はよいが, 汎用性はない.

2) **インタープリタ方式**

汎用的な認識アルゴリズムを設計し, 処理と物体モデルを分離する方式. 多くの問題に対して対応可能であるが, 対象となるモデル毎に最適な処理は行えず, 性能はよくない.

3) **コンパイル方式**

与えられた物体モデルに基づき, その物体の認識のために最適なアルゴリズムをオフラインで自動生成し, そのアルゴリズムを用いた認識処理を行う方式. 汎用性と性能を両立している.

コンパイル方式としては, [Ikeuchi87] をその代表的なものとして上げることができる. 与えられた物体モデルを基に物体の見え方の解析を行い, その物体の姿勢決定に有効な特徴を利用した解釈木を事前に生成する. 実行はその解釈木に従い行われる. このような考え方は, 汎用的なシステムにおける認識精度の向上において非常に重要であり, 図面解釈における構成要素の認識問題に通じるものである. しかし, [Ikeuchi87] においては対象物体が1つであり, 同一形状物体が山積みされたビン・ピッキング問題における姿勢決定を取り扱っていた. この状況においては, 実行時において問題設定が変更することはないので, 事前にオフラインでアルゴリズムを生成するコンパイル方式は有効である. ところが, 図面解釈問題では, 複数の構成要素が存在すること, 図面に対して多くの拘束条件が利用できることから, 解釈過程において構成要素の認識問題の設定が動的に変更する. つまり事前のオフラインでの認識アルゴリズムの生成は適さず, 解釈過程の中で動的に生成する方式が適する. このことから考えると上記の3つの分類では説明できず, 表1-1に示す分類を行い, その上で上記方式および図面解釈における構成要素の認識に適した方式を位置付けることができる.

表1-1 認識アルゴリズム生成方式の分類

	特殊アルゴリズム		汎用アルゴリズム
	プログラム埋め込み	自動生成	
オンライン処理		図面解釈問題で有効な方式	インタープリタ方式
オフライン処理	ハンドコンパイル方式	コンパイル方式	

1.3.2 知識の有効性

図面には，多種多様な知識が関係している．人間が図面を解釈する場合においても，このような知識の有無により認識精度が異なる．知識のある人間の方が優れていることは，直感的に理解できることではあるが，ここでは知識の有無と認識精度の関係に関して考えてみる．

1.1で述べたように，図面解釈は構成要素の認識とそれらの関係の認識により構成される．この認識問題での知識の利用法について「ナス」と「トマト」の認識を例に考えてみる．まず，

知識1：ナスは紫色である．
知識2：トマトは赤色である．

という知識の利用を考える（青いトマトもあるというような，この知識の実世界における正当性の議論はここではしない）．この知識を利用した認識のためのルールとして，

認識ルール1：紫色であればナスである．
認識ルール2：赤色であればトマトである．

が利用できる．これは上記の知識の逆の命題である．一般には命題の逆は成立しない場合があるので，このようなルールが常に利用できるわけではない．例えば，この認識に「ニンジン」の知識として，

知識3：ニンジンは赤色である．

を追加すると，認識ルール2が成立しないのがわかる．対象を規定する概念には，外延的概念と内包的概念があり [Ansai89]，上記知識のような外延的概念で対象を規定する場合には，この問題が必ず発生する．一般に画像認識の対象を規定する概念は外延的概念であり，この問題を回避するために，

1) 対象を限定することで，全集合を明確にする．
2) 逆の命題の真偽を検証する．

という作業が事前に必要となる．上記認識問題がナスとトマトで閉じていれば，認識ルール1,2は真となり，その利用が可能である．しかしニンジンを加えた世界で問題を考えると知識3が有効になり，認識ルール2が利用できなくなる．人間が認識アルゴリズムを設計する場合には，常にこの2つの処理を意識している．計算機による自動認識アルゴリズム生成を行うためには，この処理，特に1の処理のアルゴリズム化が重要となる．

この例が示すのは，単に知識が増えれば認識が容易になるという考えを否定するものである．

言い換えると，認識対象が増加するほど対象を規定する外延的概念を表す知識も増加し，認識が困難となる．つまり，認識の精度を上げるには，対象を限定することが重要であり，それを可能とするための知識と方法論の確立が必要である．図面解釈過程では，様々な局面で認識対象が限定される．そのような局面を把握し，その状況における認識対象の枠を決め，その枠内において最適な認識ルールの決定することが必要である．このような処理を可能にするための知識が精度の向上にとって重要なのである．

1.4 本論文の取り扱う問題

ここでは，これまでに述べてきた背景事情の中での本論文の取り扱う問題を定義する．入力対象は2値画像データとしての図面である．図面の種類としては，回路図のような複数のシンボルにより構成され，それらの接続関係が重要な意味を持つ図面を対象とする．出力結果は図面解釈の結果である．つまり，個々の構成要素の認識結果とそれら要素間の関係の解析結果により構成される構造記述である．そのような入出力関係を成立させる方法を提案することが本論文の問題である．中でも，

1) 知識利用の有効性の実証
2) 汎用性のある方法論の提案
3) 人間の解釈のしやすさの基準と類似した性能を発揮する方法論，つまり画質の変化に対してロバストな方法論の提案

に重点を置く．この問題設定は，単に図面解釈システムとしての性能の向上を目指すものではなく，画像理解研究全般において共通する問題点を打開するきっかけになることを期待するものである．

2章において知識を用いたトップダウン処理の有効性を実証する．3章では知識処理による処理と知識の分離による汎用化への試みについて述べる．3章で提案する方法論は実際には満足のいく結果には至らなかったが，そこで明らかなった問題点が最終的に本論文で提案する方法論の重要な手がかりとなった．その問題点を解決すべき新たな要素技術の提案を4，5章で述べ，6章においてそのアイデアをまとめる．7章では本論文での達成度と今後の展望について述べる．

参考文献

[Anzai89] 安西祐一郎：認識と学習，岩波書店，p.2 (1989)

[Dezawa76] 出澤正徳：設計図面の処理，精密機械，42, 8, pp.733-739 (1976)

[Ikeuchi87] K.Ikeuchi : Generating an Interpretation Tree from a CAD Model for 3D-Object Recognition in Bin-Picking Tasks, Int. J. Computer Vision, 1, 2, pp.145-165 (1987)

[Ikeuchi89] 池内克史：物体認識と認識プログラムの自動生成，人工知能学会誌，4, 1, pp.30-42 (1989)

[Jarvis77] J.F.Jarvis : The line drawing editor ; Schemantic diagram editing using pattern recognition techniques, Computer Graphics and Image Processing, 6, 5, pp.452-484 (1977)

[Kamata89] 鎌田靖彦 (編) : 地図情報システム入門 , 日刊工業新聞社 (1989)

[Miyake91] 三宅芳雄 : 理解と知識 , 人工知能学会誌 , 6, 6, pp.834-842 (1991)

[Sakauchi92] 坂内 , 角本 , 大田 , 林 : コンピュータマッピング , 昭晃堂 , pp.63-64 (1992)

[Shirai87] 白井良明 (編) : パターン理解 , オーム社 , p.122 (1987)

[Tsunekawa85] 恒川 , 下辻 : 図面読取装置 TOSGRAPH, 電子情報通信学会論文誌 , J68-D , 4, pp.466-472 (1985)

[Yachida89] 谷内田正彦 : 画像処理から画像理解へ , 人工知能学会誌 , 4, 1, pp.5-12 (1989)

2章　トップダウン処理による図面解釈
ーピアノ楽譜の自動認識ー

図面は書き手から読み手へ何らかの客観的な情報を伝達するメディアの1つである.そこには,書き手と読み手の間に図面を解釈するための共通の規則が存在する.それらは明文化されているものから暗黙のものまで多種多様なものである．そのような規則を積極的に利用したトップダウン処理による認識手法は,初期の図面解釈研究から利用されている.この章では図面解釈におけるトップダウン処理の有効性を楽譜認識システム*という具体例を用いて説明する.

2.1　トップダウン処理による図面解釈

2.1.1　トップダウン処理とボトムアップ処理

パターン認識は,入力されたパターンがあらかじめ記憶している情報のどれと対応するかを見つける処理である.このとき,記憶している情報の表現形式が入力データの表現形式と一致していないなら,情報の比較を行うためになんらかの情報変換が必要となる.一般には,入力パターンから特徴抽出を行うことでパターン情報を記号情報へ変換した後,特徴間の局所的な関係を解析することで入力パターンの大局的関係記述を作成する.その大局的関係記述という情報の表現形式の上で知識との比較を行うという手法がよく利用される.つまり,パターン認識は,（1）情報変換,（2）情報比較，の2つの処理で構成される.

ある具体的なパターン認識の問題が設定されたとき,その問題固有の知識を処理プログラムに与える必要がある.基本的には,入力データと比較するための情報,通常はそれが結果として得たい情報を表現するものであるが,それを処理プログラムに登録することにより2の情報の比較処理が可能となる.例えば,テンプレートマッチング手法では,テンプレートという入力データと同一の表現形式によって情報を記憶しているので,1の情報変換処理を必要とせず比較処理のみでパターン認識処理を実現している.しかし,一般に処理プログラムに与える知識は,この比較処理に用いられる情報だけでなく,情報変換のために与えられる知識も多い.例えば画像中から人物像を切り出し,その人物判定をする問題を考える.人物判定には顔の特徴が有効であるので,認識プログラムは顔の特徴記述を用いて比較処理を行う方針で設計されていたものとする.いま,処理のある段階で胴体部分の検出が成功したとする.このとき,知的な認識プログラムであれば,顔は胴体の上部にあるので,切り出しの成功した胴体領域の上方に注目して顔の切り出し処理を行うことを考えるだろう.

この例の中で認識プログラムは,比較のための顔の特徴記述を得る段階で,顔は胴体の上方に位置するという知識を用いて情報変換を行ったことになる.当然,この知識は机の認識には役に立たず,問題に固有の知識といえる.このような問題固有の知識を用いて情報変換を行うことの特徴として,次のことが上げられる.

＊ 図面解釈と対応させ楽譜解釈とすべきかもしれないが,この関連の国内文献では楽譜認識という用語を採用しているものが多いので，ここでも楽譜認識という用語を用いた.

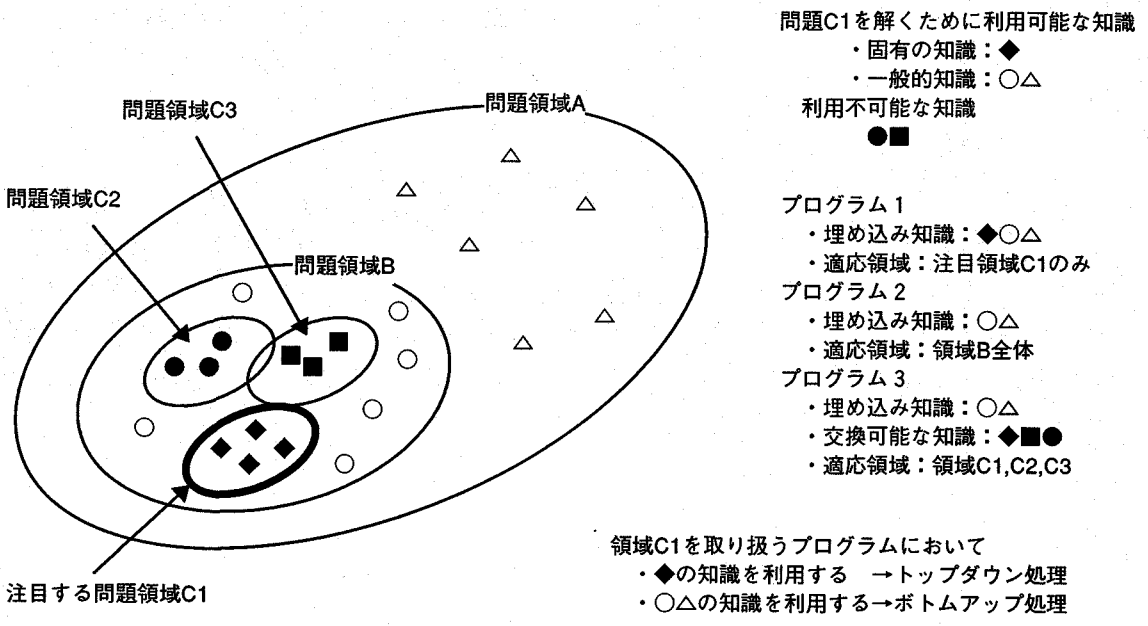


図2-1　トップダウン処理とボトムアップ処理に関連した領域知識と処理の汎用性の問題

1) 情報変換の精度や効率，つまりは認識そのものの精度や効率の向上に非常に有効である．
2) 問題固有のものであるため知識利用のための汎用的な枠組みが存在せず，問題毎に利用法を設計者が考える必要がある．
3) 知識がプログラムに直接的に埋め込まれる場合が多く，ある問題に対して開発された処理を別の処理にそのまま流用するには，埋め込まれた知識を取り替える必要がある．これは非常に効率が悪く，欠陥の潜んだプログラムになる恐れもある．

2,3のような問題が存在するものの，1の貢献は非常に大きなものである．

ここで処理方式という点から考えたとき，情報の変換処理において問題固有の知識を用いたものをトップダウン処理，問題固有の知識は用いず汎用的な枠組みでの処理を採用したものをボトムアップ処理と本論文では定義する．これらの言葉は広く用いられているものの，その定義は研究者によって異なる [Anzai89] [Yachida90].

ところで，問題固有の知識という言葉の対比として，問題普遍の知識という言葉を上げることができる．しかし，知識はこのように二分できるものではなく，知識の特殊性により何重にも包含関係を持つものである．対象とする領域を設定したときにはじめてその輪郭の内外で固有の知識とそれ以外の知識とに分けることができる．つまり問題固有の知識でなくともそれが普遍的な知識とはいえない．

プログラムの設計において，設計者は個別の問題領域に固有の知識と，それ以外の知識とを意識しておかなければならない．画像を入力として用いることは設計者が具体的な問題を想定したからこそ決定できた方針であり，その方針決定には何らかの知識が利用されたのは明らかである．例えば，食べ物の味を認識するのに画像を入力に使用しようとは思わない．しかし，手紙に書かれた文字の認識をするには当然のように画像入力を考える．これは，プログラムの

設計者が具体的問題が設定された時点で,その問題に関連する知識を引き出して判断を行ったからである．このときの「○○○の認識には画像処理により得られる情報が有効である.」という知識は,問題固有の知識とはここでは考えない.このような知識は確かに限られた領域でしか適応できないものであるが,その領域は実際に作成するプログラムが対象とする問題領域と比較するとより広い範囲を覆っていると考えられるからである.認識プログラムにおいて,そのプログラムに直接的に埋め込まれたすべての知識の個々の適応領域の交わりの領域が基本的にそのプログラムの汎用限界に対応する.

上述の問題固有の知識とは,そのプログラムの汎用限界に対応する領域知識を指すのではなく,実際にそのプログラムが動作している状況において取り扱っている問題領域を基準に考えている．その問題領域の外に位置する知識を用いた処理はボトムアップ処理と見なす．図2-1にこの考え方の概念図を示す.

2.1.2　前向き推論と後ろ向き推論

トップダウン処理とボトムアップ処理という分類と類似した処理の分類方法に,前向き推論と後ろ向き推論というものがある.この定義は人工知能研究において,問題空間の探索の方向について明確に定義されている言葉である.つまり,初期状態から最終状態に向かって探索を行う推論を前向き推論,最終状態から初期状態に向かって探索を行う推論を後ろ向き推論と呼んでいる [Rich83].

図面解釈処理は,処理過程そのものを入力画像という初期状態から,構成要素の関係記述という最終状態を推論する一種の探索処理と見なすことができるので,この前向き推論と後ろ向き推論という言葉で,処理を分類することができる.具体的には,入力画像から結果として出力すべき解釈結果へと向けて情報変換を行う処理の流れが前向き推論,最終的な解釈結果や中間的な記述情報からその妥当性を検証するために入力画像へ向けて情報変換を行う処理の流れが後ろ向き推論に対応する．一般には,ボトムアップ処理と前向き推論,トップダウン処理と後ろ向き推論が同一であるかのように見なされる場合もあるが,これらは異なる分類の方法である.

2.1.3　図面解釈におけるトップダウン処理の有効性

前述の処理の分類方法に沿って，各処理の特徴を以下にまとめる.

トップダウン処理＆前向き推論：

手続きにより表現された知識埋め込み型のプログラムで採用される場合が多い．実システムの開発においては，ほとんどがこの方式である．問題毎に個別に設計を行う必要があるので，開発効率は悪いが，実行性能は最も優れている.

トップダウン処理＆後ろ向き推論：

知識埋め込み型のプログラムでもよく採用されるが,この分類に属する特徴的な処理形態として，知識処理に基づく仮説検証型処理がある．問題固有の知識を知識ベースとして処理プログラムから独立させ,推論エンジンが知識を参照しながら仮説の生成と検証を進める方式である.

ボトムアップ処理＆前向き推論：

入力画像に対する前処理などがこれに該当する．入力画像の正規化，平滑化，特徴抽出処理などである．

ボトムアップ処理＆後ろ向き推論：

問題の最終状態というのはその問題固有のものなので，後ろ向き推論を行うには通常問題固有の知識を必要とする．つまり，ボトムアップ処理での後ろ向き推論は一般には考えにくい．強いて上げると，3次元幾何形状をシステムが取り扱う場合に，ある形状仮説からコンピュータグラフィックスの技法を用いて画像生成を行い，画像情報として入力画像とマッチングを行うような処理方法が上げられる．この場合，形状仮説から画像生成には問題固有の知識は必要とされず，汎用的な処理によって行われる．

実行性能の点からは，トップダウン方式を利用すべきであるというのは明らかである．つまり，問題固有の知識を積極的に利用すべきということである．しかし，知識をプログラムに埋め込むことにより処理の汎用性が制限されるという問題点もある．この章では，問題固有の知識を積極的に利用した図面解釈システムとして，ピアノ楽譜の自動認識システムを取り上げ [Kato88] [Kato90] [Kato92]，まず図面解釈におけるトップダウン処理の有効性を実証する．

2.2 楽譜認識システム

2.2.1 楽譜認識技術の動向

楽譜認識は図面解釈研究の対象であると共に，音楽情報処理研究の対象でもある．コンピュータ技術の発展は，音楽の分野にも非常に大きな影響を及ぼしてきた [Inokuchi81]．電子楽器を用いての自動演奏は既に実用化されており，音楽様式の分析・比較，自動作曲・編曲システムなど，さらには，これらすべてを統合した音楽情報処理システムの研究も行われている [Inokuci86]．これらの研究を行うに当たっては，データベースは不可欠なものであるが，一般にその入力方法が問題となっている．音楽情報をコンピュータに入力する方法として主に次の4つが上げられる．

1) 音響データからの自動採譜
2) 楽譜の自動認識
3) 鍵盤型キーボードによる入力
4) 対話的手法を用いた手動入力

民族音楽などの楽譜化されていない曲をデータベース化するには，絶対音感のある専門家が必要になる．そこで専門家でなくとも採譜作業を行うことができるような自動採譜システムや採譜支援システムの開発が行われている．また，楽譜として存在する曲の入力も，その効率を考え，自動認識システムの開発が行われてきた [Ohteru85]．しかしこれらのシステムは，現在も研究段階であり，完全に実用化されているものはなく，現状では，鍵盤型キーボードによる入力や対話的手法を用いた手動入力がほとんどである．

楽譜の自動認識はこのような音楽情報のデータベース化の1つの手段であるが，その研究は

これまでにも多数報告されている．プロジェクションプロフィールを用いた認識手法として [Nakaura79] [Fujinaga89] などの報告がある．これは，パターン内の特定点の位置情報が重要な意味を持つことを利用して，垂直・水平方向のプロジェクションにより記号の位置情報と形状情報に関する特徴を抽出して認識を行う手法である．この手法は簡単ではあるが，シンボルが縦方向に並んだ場合や互いに接触した場合などの認識が困難であり，童謡程度の単旋律の簡単な楽譜にしか適応できない．青山らは，切り出された記号を適当な特徴パラメータで分類する手法を報告している [Aoyama82] [Tojo82]．この手法は，シンボルの外接四角形の大きさにより記号空間上で大分類を行い，記号領域の縦・横方向の黒ラン長に基づいて各グループ毎に固有の構造解析を行うことにより認識を行う．この手法ではシンボルの分類は比較的簡単に行えるが，それ以前の五線消去とセグメンテーションにおいて煩雑な処理が必要となる．大照らは，オンライン自動演奏を目的とした自動認識システムを報告している [Matsusima85]．このシステムは簡単なエレクトーン楽譜を対象としており，十数秒程度の短い時間での認識を実現するためシンボルの検出をハードウェア化し高速化を図っている．

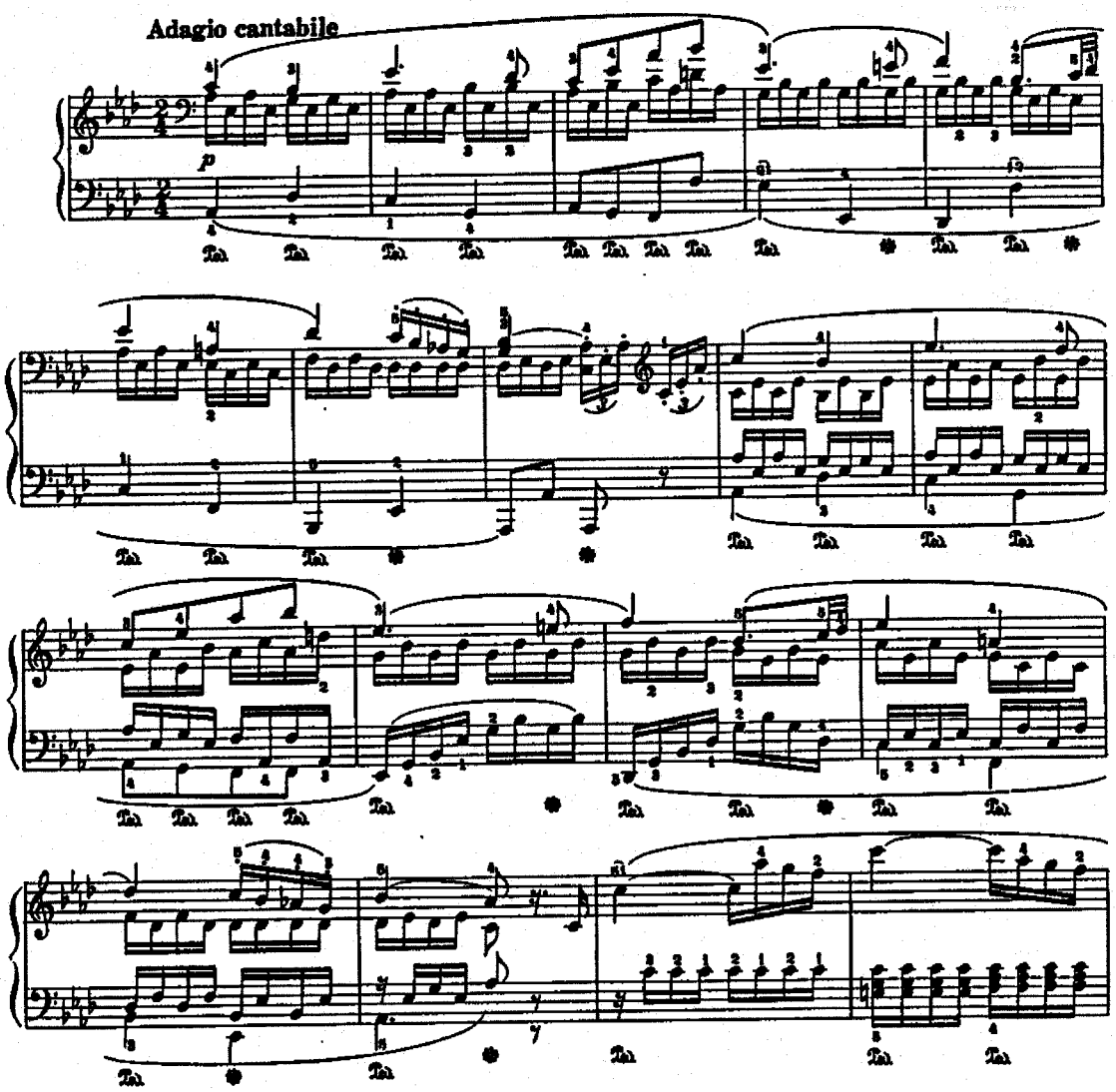


図2-2　ピアノ楽譜の例（ベートーベン作曲：ソナタ悲愴第2楽章）

楽譜は曲目ごとに認識における難易度が大きく異なる．単純な楽譜では，記述されるシンボルの密度が低く，シンボル間の接触は考慮しなくても良い．また，意味的にも一意な解釈ができる．複雑な楽譜では，高い密度でシンボルが記述され，シンボル相互の接触・交差・入り込みなどが頻繁に起こる．意味的にも，複数の解釈が可能な曖昧な記述や規則に反した記述が行われる場合もある．そこで，複雑な楽譜に対処するためには，知識を積極的に用いた処理が必要となる．これまでの楽譜認識の手法では音楽や楽譜に関する知識を積極的には用いておらず，複雑な楽譜に対応できるものはない．

2.2.2 トップダウン処理による楽譜認識

ここで提案する手法は，複雑な楽譜への適応を試みたものである．対象楽譜には，一般的で，曲目により難易度の差が大きいということからピアノ楽譜を選んだ．図2-2にピアノ楽譜例を示す．認識における基本的な考え方としては，楽譜や音楽に関する知識を積極的に利用することにより，図形的複雑さや意味的曖昧さに対処する．特に，音楽的規則を有効に利用するために小節という単位に注目し，その単位毎の認識処理を行う．処理はパターン認識処理と意味解析処理に大別でき，パターン認識処理では，シンボル相互の接触などに対処するため，それらの構成要素となるプリミティブを抽出し，それを再合成することにより個々のシンボルの抽出を行う．このことにより，画像から抽出すべき図形の形状が非常に簡単になり，接触や途切れ等のノイズへの対処が比較的容易に行える．また，プリミティブ相互間の連結関係の情報を積極的に使うことにより仮説検証処理や仮説間の矛盾の検出も可能となる．意味解析処理では，パターン認識処理で得られた結果の意味的な無矛盾性を検証するとともに，もし矛盾が検出されればその原因となる仮説を棄却する．また音楽的規則を優先順位の高い方から順に適用することにより，シンボルの意味やその接続関係の解釈を行い実際に演奏可能なデータを生成する．

2.2.3 処理の概要

図2-3に処理の流れを示す．入力された楽譜画像に対し，前処理として，五線の検出，小節線の検出を行い，楽譜を小節単位に分割する．認識処理は分割された小節単位に行われ，各小

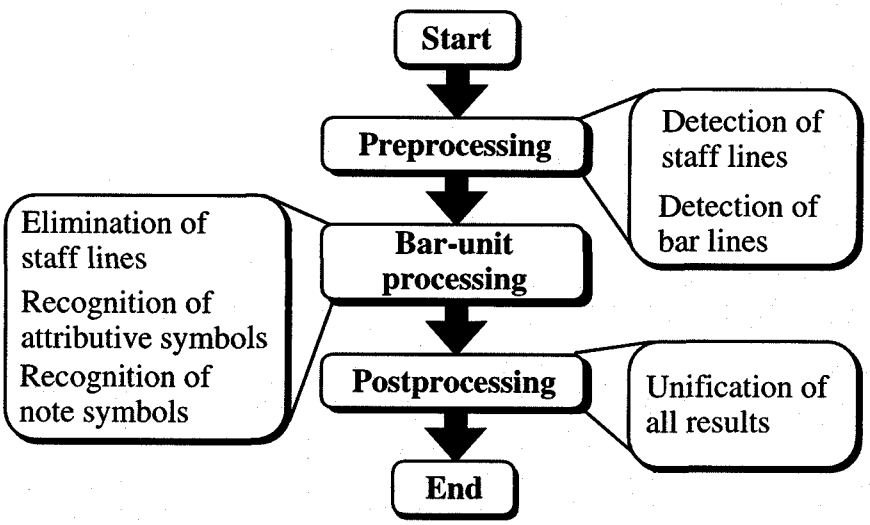


図2-3 処理の流れ

節ごとに音楽的知識を用いた処理が実行される．後処理では，各小節ごとの認識結果を反復記号を考慮しながら結合し最終的に演奏データを生成する．

ここで，楽譜における音の強さの記述について考えてみると，その記述は非常に曖昧で，記述されている情報のみで実際の音の強さを割り出すのは極めて困難と考えられる．そこで，今回は音の強さについての認識は行わないことにした．また，音の強さを表す記号などの認識対象外記号についての考察は，2.6で述べる．

2.3　前処理

楽譜は五線を基準に構成されており，記号の記述される位置も限定されている．また，楽譜は時系列データを記号化したもので，五線に沿って時間的な経過が表現されている．2.1で述べたように認識処理を行う場合，問題固有の知識を基に限定された範囲内に特定の処理を施すというトップダウン処理が，精度的にも処理時間の上でも有効である．小節は，拍子という単位で五線を時間的に等間隔に分割したものであり，その中では拍数が必ず一定になるという強い拘束条件がある．つまり，小節という単位を認識における処理単位とすることは，音楽的にも意味があり，音楽知識を用いた処理が容易に行える．そこでまず，前処理として五線の検出と小節線の検出を行い，楽譜を小節単位に分割する．

2.3.1　画像入力

楽譜入力にはイメージスキャナを用い，Ａ４版１ページ単位に数回に分けて取り込む．解像度は200DPIとし，その画像は１ページ当り，横１９６０，縦２６４０の大きさになる．

2.3.2　五線検出

入力された楽譜画像においては，五線の水平性は保証されてなく，ある程度傾きを持っているものと考えられる．また五線の線間隔や１ページに書かれる段数も決っていない．そこで，五線の各線においてその始点と終点の座標を求め，その２点を結ぶ直線として五線を表す．

まず，楽譜画像を等間隔に１０ヶ所縦方向に走査して，黒画素と白画素のラン長を測定し，それぞれヒストグラムをとる(図2-4)．これらのヒストグラムにおいて最大値をとるラン長を

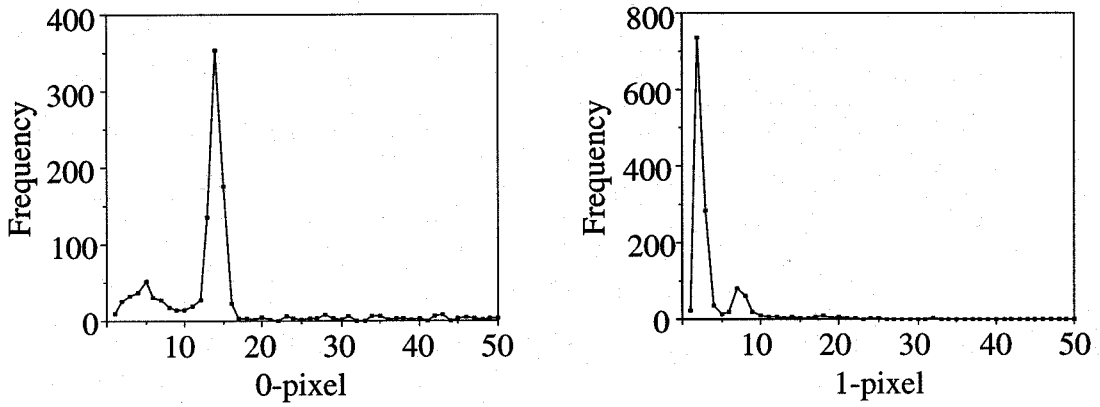
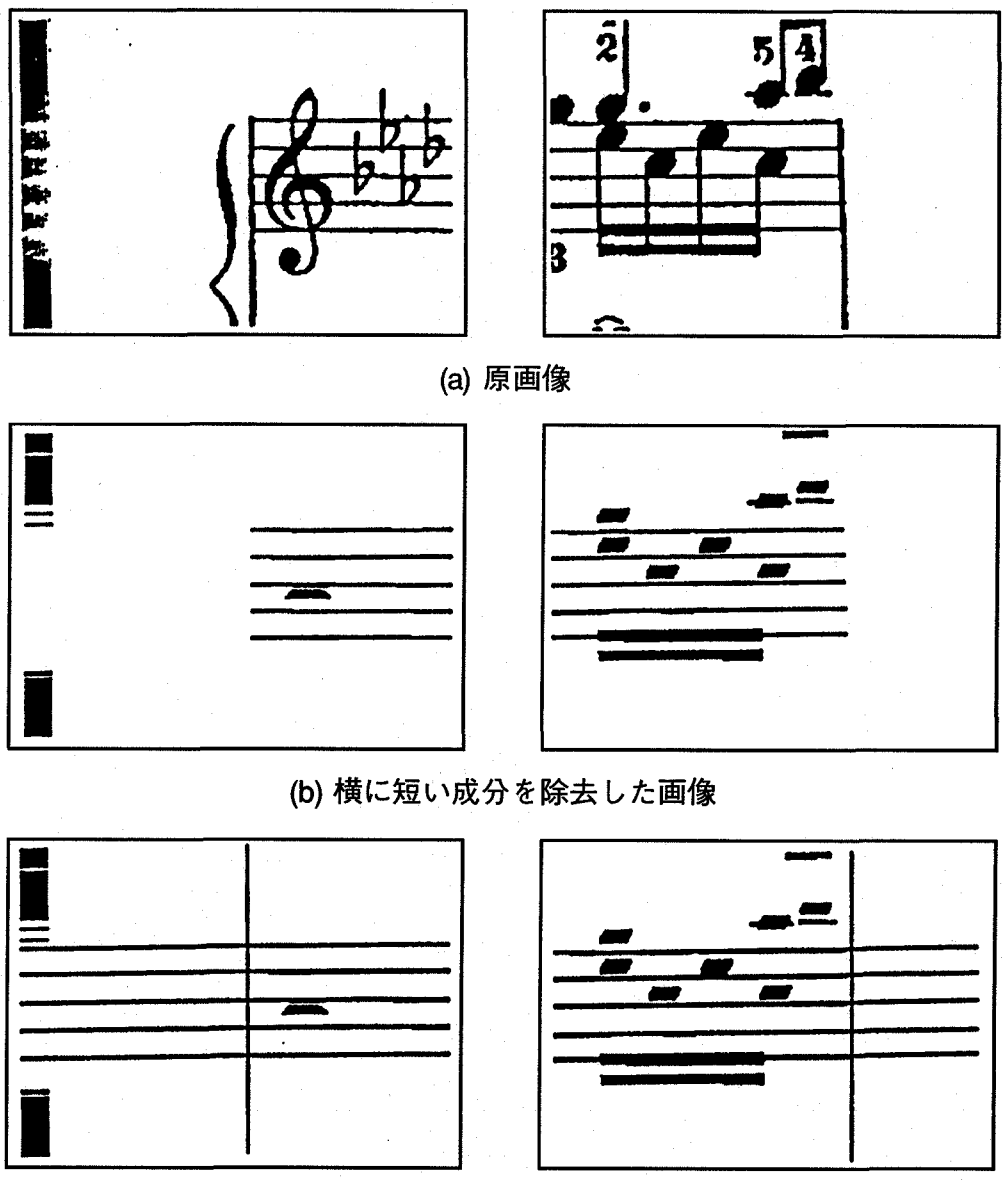


図2-4　線幅・線間隔のヒストグラム

それぞれ五線の線幅・線間隔と考えた.

次に, 画像を先と同様に縦方向に１０カ所走査し, いま求めた線幅・線間隔の値をもとに各走査ごとに五線候補を選ぶ. すべての走査における五線候補をまとめ, その頻度の高いところを五線とする. 線幅から推定される五線の大きさの約３倍のサイズのウインドウを五線位置に掛け, それを左右に動かし, ウインドゥ内で横方向に短い成分を除去した後, 縦横両方向の投影を利用することにより各五線の始点と終点の正確な位置を求める. 図2-5にその処理過程の例を示す. ピアノ譜は一般に高音部譜表と低音部譜表の２段で構成されており, 五線検出は各段を単位として並列に処理される.



(a) 原画像

(b) 横に短い成分を除去した画像

(c) 五線検出画像

図 2-5　五線検出の処理の例（図 2-2 の 1,2 段目）

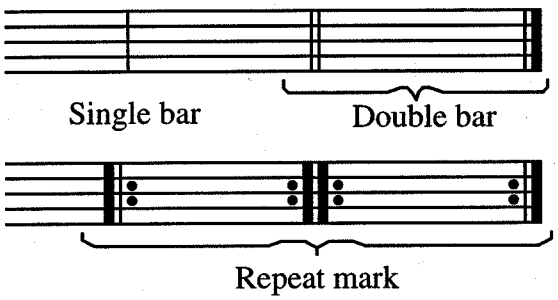
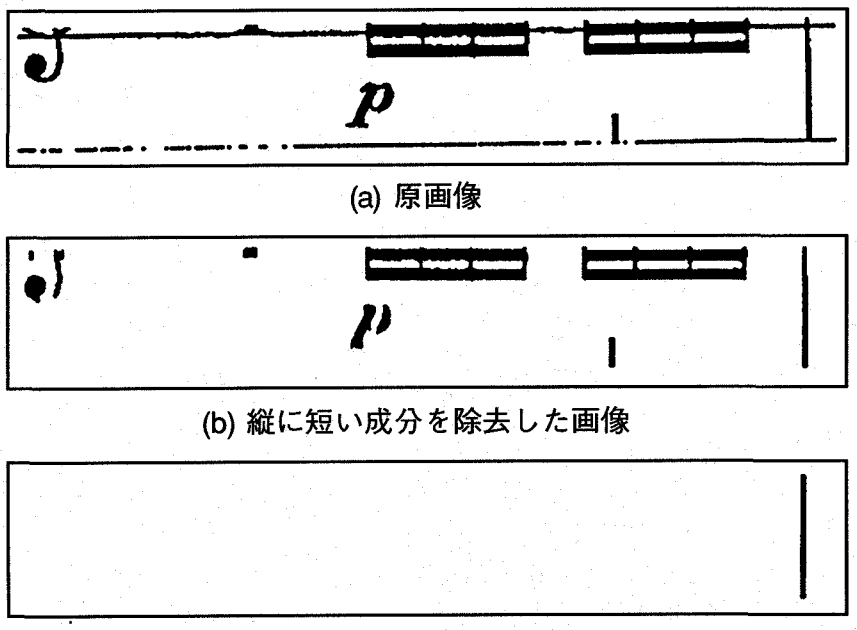


図2-6　小節線の種類



(a) 原画像

(b) 縦に短い成分を除去した画像

(c) 小節線検出画像

図2-7　小節線検出の処理の例（図2-2の1段目と2段目に挟まれた領域の左側部分）

2.3.3　小節線検出

本システムで小節線として認識されるシンボルを図2-6に示す．ピアノ譜は先にも述べたように高音部譜表と低音部譜表の2段で構成され，小節線はその2つの領域を貫いて書かれる．そこで，高音部・低音部及びそれらの譜表に挟まれた中間部の3ヶ所の領域において，まず縦方向に短い成分を除去した後，縦方向の投影から各領域での小節線の候補を抽出する．次にそれらの候補の中から縦位置が一致する組を捜し，それを小節線として検出する．そして相互位置関係から単縦線と複縦線を識別し，さらに線幅，第2間・第3間の点の有無により反復記号の識別を行う．図2-7に小節線検出の処理過程の例を示す．

2.4　小節単位での認識処理

認識処理は，前処理により分割された小節単位に行われ，次の3つの処理部に分けられる．

1) 五線消去
2) 音部・調子・拍子記号の認識
3) 音符・休符・臨時記号の認識

これらの処理においては，認識された記号は原画像から順に消去され，処理の終了時には認識対象外の記号のみの画像が残る．

2.4.1　五線消去

五線はすべての記号の位置情報の基準となるもので，非常に重要なものであるが，認識を行う過程では，他の記号と重なり合い，むしろ邪魔なものとなる．そこで，あらかじめ五線を消去した後，記号の認識を行うという方針を採用した．前処理によって検出された五線の始点と終点の位置から，各小節内での五線の端点の座標値を推定する．この推定値をもとに，小節の左端から五線を追跡する．このとき，五線の線幅も前処理によって検出されているので，この値を利用し，線幅がしきい値以下のところでのみ消去を行う（図2-8）．この場合，本来連結すべき記号が分割される場合も生じるが，この問題に関しては，各々の記号の認識処理において対処する．

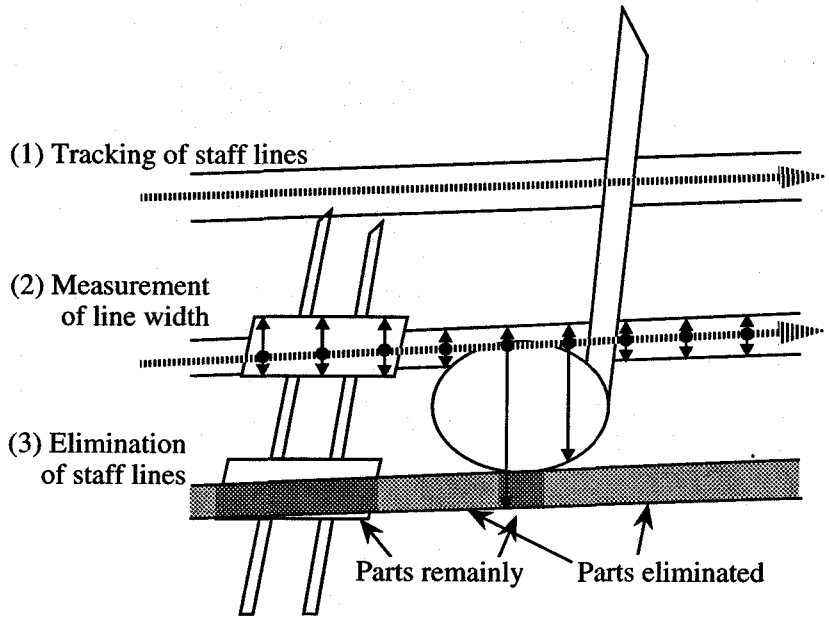


図2-8　五線消去の説明

2.4.2　音部・調子・拍子記号の認識

これらの記号は，記述される位置が限定されている．つまり，

1) 音部記号は，各段の五線の最初に必ず書かれる．
2) 調号の記入は，音部記号の後，または，複縦線の後に限られる．
3) 拍子記号は，曲の最初の五線の音部記号の後に必ず書かれ，また，複縦線に続けて書かれる場合もある．

これらの規則から，小節の位置またはその小節の直前の小節線の種類により，認識処理を行わなければならない小節を限定することができ，その限定された小節でのみ認識を行う．これにより，曲の途中での拍子・調子の変化への対応が可能となる．

<u>1) 音部記号</u>

対象とする音部記号は，ト音記号とヘ音記号である．五線との位置関係より音部記号と推定される連結成分を抽出し，その外接四角形の大きさにより識別する．

また，音部記号は，図2-2の楽譜例の第1小節や第8小節目にみられるように，上で示した規則に反して各五線の途中に現れることもある．このような音部記号の認識処理はすべての小節で行う必要がある．しかし，これらの音部記号は五線に対して常に一定の縦位置に現れ，また次に記入される可能性のある音部記号は，その直前に記入されているものとは必ず異なる．このことを利用し構造解析を行うことにより認識を行う．

2) 調号

調号は，シャープまたはフラットにより構成され，図2-9のようにその違いにより記入される位置が異なってくる．また上下各段で，調号は同様な記述がなされる．この2つの規則により種類と個数を判別する．

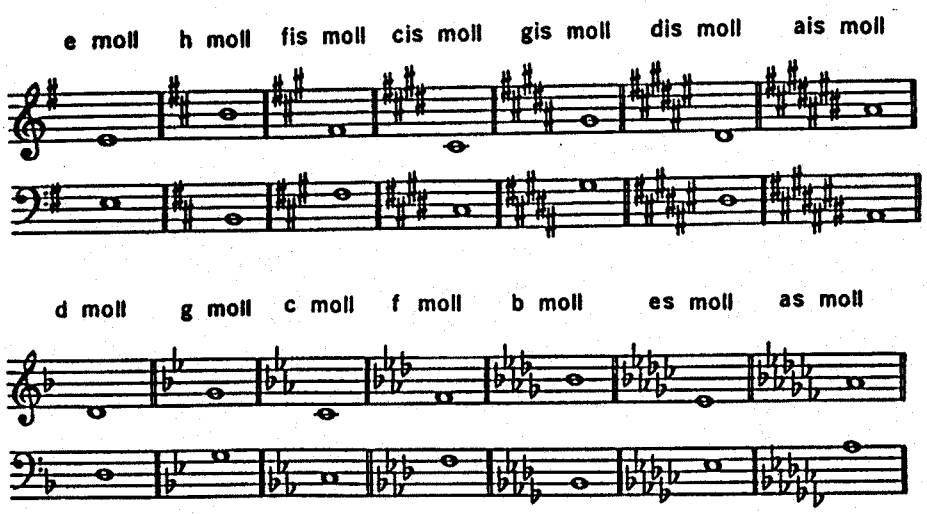


図2-9　調号の位置　[Shimoda82] より

3) 拍子記号

拍子記号は，分子・分母の2つの数によって構成されているが，その数値は限られている．そこで投影パターンの特徴を比較することにより認識を行う．

これらの記号は，それ自体が音を表すものではなく，後続するすべての記号に影響を与える．この限界は，小節内に留まらないので，その情報は，後の小節へと引き継がれなければならない．

2.4.3　音符・休符・臨時記号の認識

これらの記号は，音を表すもっとも基本的な記号である．形状は非常に単純であるが，認識においては次のような問題点があげられる．

1) 他の記号との接触・交差・入り込みなどが頻繁に現れ，容易に切り出しが行えない．
2) 同一の記号でも，形や大きさが一定でない．
3) 五線により本来空白であった部分が埋められたり，また，五線消去時に，本来連結していた部分が分割される場合もある．
4) 細い成分が多いため，記号の途切れやかすれが現れる．
5) 他の記号の影響で意味が変化するものもある．
6) 複数の記号間の意味的なつながりが曖昧である．

1から4は，パターン処理上の問題で，5,6は，意味解析上の問題である．本システムでは，記号を複数のプリミティブに分解し，それらプリミティブのパターン認識と，意味解析によるフィードバックを取り入れた処理を行っている．図2-10にその概念図を示す．中間結果を記述するための5層の階層構造を持つワーキングメモリと，それを自由にアクセスすることのできる4つの処理モジュールから成り立っている．初期状態では，最下位の画像階層にのみデータが存在する．処理モジュールが起動されると，ワーキングメモリがアクセスされ，データがその中を上下の階層へと移動する．そして最上位の階層に書かれるデータが，認識結果となる．

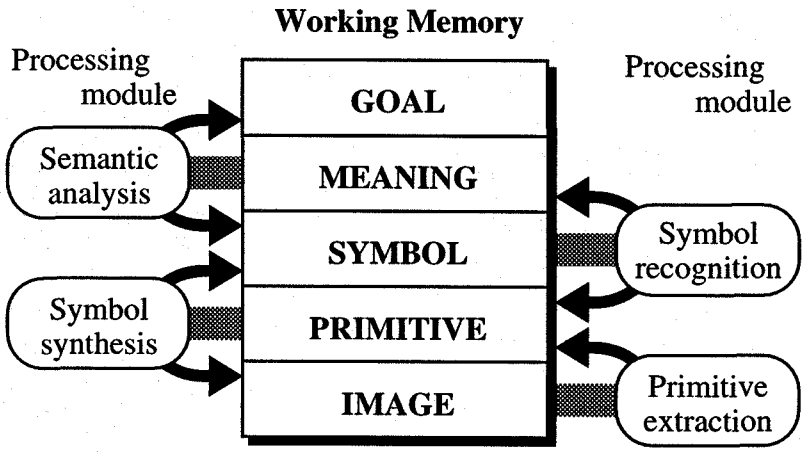


図2-10　認識処理の概念図

2.4.3.1　階層ワーキングメモリ

a) IMAGE

これは最下位の階層であり，認識が行われる画像データである．図2-11に示すように，初期状態から処理が進むにつれ，抽出されたものが消去されていき，最終的には認識対象外の記号のみの画像となる．

b) PRIMITIVE

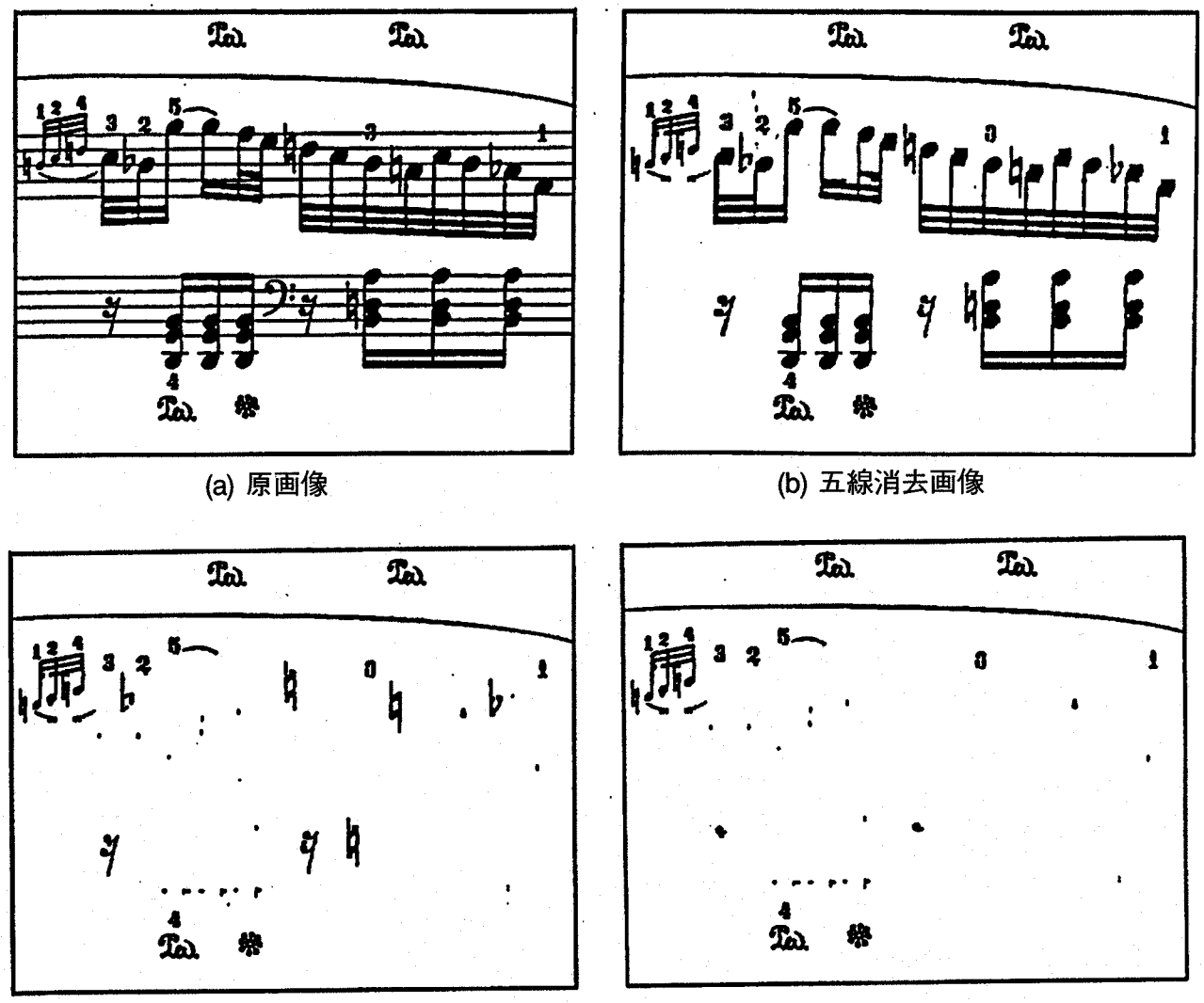
ここにはIMAGEの階層から抽出された音符・休符の構成要素となるプリミティブが記述される（図2-12）．ここでは，臨時記号についても付点などと同様に，音符の構成要素の一部と考える．主なデータ内容は，プリミティブの種類・その外接長方形の位置・原画像におけるラベルである．

c) SYMBOL

この階層に記述されるものは，PRIMITIVEの階層に記述されている構成要素を連結することにより合成される音符・休符の図形的情報である．データはPRIMITIVE階層へのリンクで表される．

d) MEANING

SYMBOLの階層に記述されている個々の音符・休符の意味，つまり，音の高さと長さの情報が記述される．また各々の記号の位置情報も記述される．この位置情報は，後段の処理にお



(a) 原画像　(b) 五線消去画像

(c) 途中経過　(d) 処理終了時の状態

図2-11　ＩＭＡＧＥ階層の変化の様子

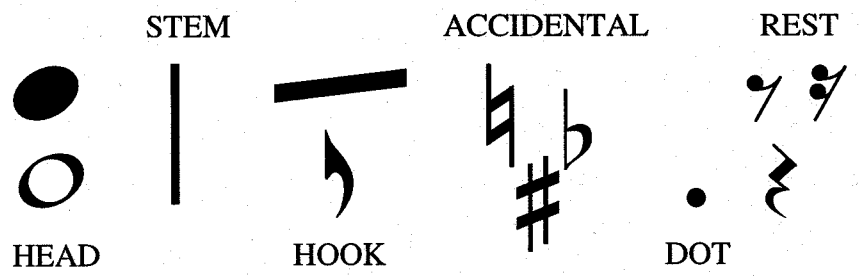


図2-12　プリミティブの例

いて各記号間の時間的なつながりを解析する際に利用される．この階層では，音符と休符は同じ取り扱いをされ，休符は単に音高のない音符と見なされる．

e) GOAL

MEANINGの階層のすべてのデータを連結して得られる結果が記述される．このデータが，対応する小節における最終的な認識結果となる．

2.4.3.2　処理モジュール

処理モジュールは，ワーキングメモリの対応する階層にあるデータをその上下の階層へ移動させる．これらのモジュールには次の4つがあり，それぞれ複数のユニットにより構成されている．

1) プリミティブ抽出モジュール

このモジュールは，図2-12に示すプリミティブを画像データから抽出するユニットから構成される．各々のプリミティブを抽出するためのユニットが各々用意され，それらは独立した処理を行う．各々のユニット内での処理は，基本的には，対応するプリミティブの形状的な特徴や位置情報を利用した構造解析に基づいている．このとき，PRIMITIVEの階層に既に記述されている情報を利用し処理を行うユニットもある．例えば，臨時記号を抽出するユニットは，PRIMITIVEの階層に記述されている符頭の位置情報を利用し処理を行っている．候補として抽出されたものは，PRIMITIVEの階層に書き込まれ，同時に，IMAGEの階層からは，消去される．

PRIMITIVEの階層の情報を利用するモジュールが存在することと，画像データが消去されることの2つの理由により，ユニットの起動順序が処理結果に影響を与える．また，各々のユニットは，しきい値をパラメータとして持っており，この値を変えることにより抽出されるプリミティブの信頼度が変わる．つまり，しきい値が厳しいときは，各ユニットが持つモデルに対し正確なものが抽出されるが，これが緩いときは，途切れを起こしていたり，形状が歪んだものも抽出される．

これらのことから，このモジュール内のユニットは，あらかじめ決められた順序で起動され，まずしきい値を厳しいものにして抽出を行い，残ったものに対してしきい値を徐々に緩めながら繰り返し抽出を行うという方法をとっている．

2) シンボル合成モジュール

このモジュールには，音符合成ユニットと休符合成ユニットの2つがある．それぞれ位置情報をもとにプリミティブを合成することにより，意味のある記号を抽出する．音符は，基本的に符頭と符尾から成り立つ．この組合せをすべて捜し，付加的なものとして符鉤・付点・臨時記号の有無を調べることにより音符の合成が行われる．また位置的に矛盾するものを表す規則として次のものがある．

1) 符尾の両端に符鉤は付かない．
2) 単鉤には符尾が1本のみ付く．
3) 連鉤は小節線を越えて書かれたもの以外は2本以上の符尾が付く．
4) 符尾の両端に符頭は付かない．
5) 符頭に付く符尾は上下に1本づつである．

その他，符頭の大きさ・数と符尾の長さの関係などプリミティブの相対的な大きさを制限する規則もある．これらの規則に矛盾するプリミティブは，IMAGEの階層に戻され，残された矛盾のないプリミティブを用いて音符が合成されSYMBOLの階層へ移動する．休符はそれ自体で1つのシンボルを構成するので，そのままSYMBOLの階層に移動する．また付点が付くものは，付点休符として合成される．

これらのユニットもしきい値をパラメータとして持っており，その値により合成されるシンボルの信頼度が変わる．例えば，四分音符のような簡単な音符を考えると，まず‘符尾’と‘符頭’の距離が測定される．そしてその値がしきい値より小さな場合には，1つの音符として抽出されSYMBOLの階層に移動するが，大きな場合には，別々の記号を表すものとして扱われ，再度IMAGEの階層へ戻される．

3) シンボル認識モジュール

このモジュールは，SYMBOLの階層に記述されている音符・休符の意味，つまり音の高さと長さを検出するものである．ここでは休符を音高のない音符と見なすことによって両者を統一して考えている．音の高さは，符頭と五線・加線との位置関係及び音部記号により一意に決定できる．しかし音の長さに関しては，複数の可能性が考えられ，それは基本的解釈と例外的解釈に分けられる．例外的解釈には，音楽的要因から起こるものと，図形的要因から起こるのもがある．音楽的要因とは，三連符など連符に関するものである．図形的要因とは，ノイズ等の影響によるもので，そのため一意な意味解析が行えない場合がある．具体的には‘符鉤’の数を検出するときに起こるのだが，複数の符鉤を持つ音符ではそれら符鉤に挟まれた白領域に五線が重なったり，印刷時のインクのにじみのため，複数の符鉤が一体化する場合もあり，良好な検出を行うことが困難である．そこで縦・横方向の連結成分のラン情報を用いた検出方法など複数の方法で検出を行い，その結果を集計することにより優先順位付きの異なる複数の結果を生成する．

各々の音符・休符に対して解析を行う際，まず基本的な解釈がなされ，MEANINGの階層に移動する．そしてMEANINGの階層から棄却されたデータに対して次々と例外的な解釈がなされる．

4) 意味解析モジュール

このモジュールは，MEANINGの階層に記述されているデータを連結し，その小節における最終的な認識結果を生成する．ここでは，次のことが問題となる．

1) 同時刻に打鍵される音でも，横方向にずれて記譜される場合がある．
2) 多旋律の曲の場合，注目している音がどの音に続くのかが曖昧に記譜される場合もある．

この2つを考慮に入れ連結を行い，解を生成する．

この解の正当性は，拍数によって検証され，もし拍数が規則に反する場合には，原則として解は棄却される．しかし弱起の場合やその小節を区切る前後の小節線が反復記号の場合には，拍数が異なる可能性もある．この場合，解の正当性の検証には，他の小節の結果との比較が必要となる．そこで，もし予想される拍数と一致する解が存在しない場合には，得られたすべての解を保存しておき，後処理においてその1つを選択する．

2.4.3.3 認識戦略

認識処理は,処理モジュール内のユニットを逐次起動することにより各階層で中間結果を生成し，最終的にGOALの階層に至る結果を生成することによってなされる．下位階層のモジュールは，そのほとんどがラベリング等の2値画像処理を含んでおり，シンボル化されたデータを扱う上位階層のモジュールと比較すると，処理時間を要するものである．そこで，これら下位階層のモジュールに対しては，あらかじめ定められた起動順序に従い処理を行う．また上位階層のモジュールに対しては，その使用する規則に優先度を定め，解が求まるまでその優先度の高い方から総当たりに処理を行う．解が存在しなかった場合，下位階層のモジュールにおいてしきい値を緩め，再度同様の処理を行う．ただし一度SYMBOLの階層に移動したデータは，PRIMITIVEの階層まで棄却されることはあっても，IMAGEの階層まで棄却されることはない．

この手法は，画質のよい画像に対しては初期の段階で解が求まるが，画質の悪い画像では，解が求まるまで時間がかかる．しかし安定な記号から順に抽出されるので，しきい値を緩めることによるノイズの影響を抑えることができる．

ところで，処理過程において完全な解に到達する以前に不完全な解が求まる場合もある．例えば，図2-2の楽譜例の第1小節目の場合，図2-13に示す記号のみが抽出されMEANINGの階層に記述されたとする．このとき，これらは拍数の条件を満足するので，解としてGOALの階層に移動する．ところが，この解は図2-2から分かるように4分音符が2つ認識されていない．つまり，解が得られたとしても，それが完全な解であるとは限らない．そこで，実際は処理過

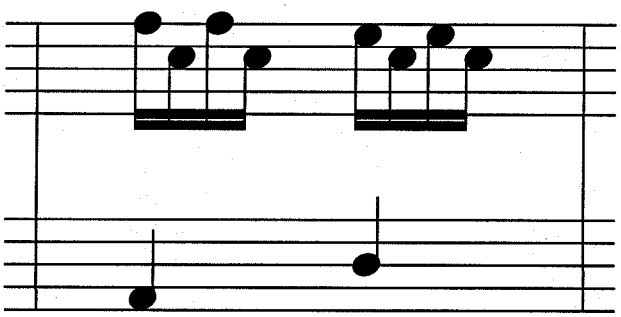


図2-13 不完全な解の例（図2-2の第1小節目）

程で解が得られた場合，その解を保存しておき，解が得られなかったものとしてさらにしきい値を下げ，一定の範囲内で処理を繰り返す．このとき，もし新たな解が得られたならそれを解として更新し処理を続け，得られなかった場合には，保存していた解を真の解として処理を終了する．

2.5 後処理

この処理では，反復記号を考慮にいれることにより，小節単位に得られた結果を演奏される順序に連結する．また，弱起の場合などで複数の解が存在している小節があれば，対応する小

表2-1　図2-2の楽譜例から生成された演奏データの一部

```
Flat : 4
2/4
==<1>== : 4 * 2
C3  0 4 (right, slur,
         p, pedal)
Ab2 0 1 (right)
Ab1 0 4 (left, slur)
Eb2 1 1 (right)
Ab2 2 1 (right)
Eb2 3 1 (right)
Bb2 4 4 (right,
pedal)
G2  4 1 (right)
Db2 4 4 (left)
Eb2 5 1 (right)
G2  6 1 (right)
Eb2 7 1 (right)
```

```
==<2>== : 4 * 2
Eb3 0 6 (right,
pedal)
Ab2 0 1 (right)
C2  0 4 (left)
Eb2 1 1 (right)
Ab2 2 1 (right)
Eb2 3 1 (right)
Bb2 4 1 (right,
pedal)
G1  4 4 (left)
Eb2 5 1 (right)
Bb2 6 1 (right)
Db3 6 2 (right)
Eb2 7 1 (right)
```

節との相互関係よりその1つを選択する．これにより，演奏データが生成される．このデータは，MIDI (Musical Instrument Disital Interface) 規格対応の電子楽器における演奏情報を表現するための統一フォーマットであるスタンダードMIDIファイルとして保存される．表2-1に生成された演奏データの例を示す．

2.6　実験

今回，実験対象としては，難易度の異なる次の6種類の楽譜を選んだ．プログラムはすべてC言語で記述し，UNIXワークステーション上で実験を行った．

1) バイエル57番
2) バイエル67番
3) エリーゼのために　(ベートーベン作曲)
4) トルコ行進曲　(モーツァルト作曲)
5) 別れの曲　(ショパン作曲)
6) 悲愴第2楽章　(ベートーベン作曲)

1,2の楽譜は，ピアノの入門者向きの楽譜で非常に簡単なものである．3,4の楽譜は，初心者向きで，5,6の楽譜は，中・上級者向きの楽譜である．これらの楽譜に対する認識結果を表2-2に示す．

表2-2　認識結果

曲　目	総音数	誤音数	認識率(%)
バイエル57番	101	0	100
バイエル67番	119	0	100
エリーゼのために	903	52	94.2
トルコ行進曲	1412	127	91.0
別れの曲	1860	119	89.3
悲愴第2楽章	1639	277	83.1

認識率はデータベース化する際に人間が修正しなければならない音の数の割合で，楽譜上のすべての音に対しての誤認識とリジェクトされた音の数の合計から求めたものである．1,2の楽譜は非常に判り易く記述されており，記号間の接触もなく良好に認識が行えた．3,4の楽譜では連符が現れ臨時記号も多い．特に4では非常に多くの装飾音符が記譜されており，誤認識の原因となっている．5,6の楽譜は非常に多くの記号が高い密度で記譜されており，記号間の接触などが頻繁に起こっている．特に和音や臨時記号が非常に多いことが特徴として上げられ，それらの記号の取りこぼしが目立った．

改善点としては，ほとんどの誤認識の原因となった下位階層のモジュールにおける各ユニットの処理の強化が挙げられる．また，意味解析処理で使用される音楽的規則として現在は音の長さに関する情報しか取り扱ってないが，和音構造などを使用することにより音の高さに関する精度の向上も期待できる．

1ページ当りの認識時間は，そのページに含まれる小節の数及び画質により異なるが，MIPS R8000 (75MHz) CPUの搭載された計算機を使用して，平均して約2分であった．本システムは小節単位で処理を行っているが，この処理において，各小節間での情報の受渡しは音部・調子・拍子記号に関するものに限られており，これらの記号の処理が終了した時点で次の小節の処理が可能となる．このことから，小節単位での部分的な並列処理が可能となり認識時間の短縮が行える．例えば3) の楽譜におけるCPU数と処理時間との関係は，図2-14のようになった．CPU数が増加するほど処理時間は短縮しているが，約30秒当りに漸近している．これは，前処理・後処理に要する時間に加え，並列化不可能な音部・調子・拍子記号の認識に起因するものである．

今後の課題として次の2つが考えられる．1つは認識対象記号の拡張である．現時点での認識対象記号は音高・音長情報を表す記号のみで最も基本的なものであり，これから拡張される記号としてはアクセント・フォルテ・クレシェンドなどの強弱記号，タイ・スラー・スタカート・テヌートなどの奏法記号が挙げられる．パターン処理の立場から見ると，本手法では認識された記号が消去されており，しかもそれら認識済みの記号の情報を用いたトップダウン処理を行うことにより，残ったものに対しての認識処理は比較的容易に行えると考えられる．しか

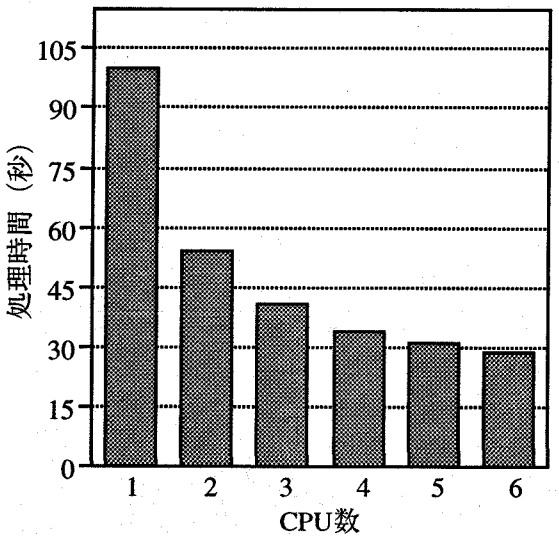


図2-14　CPU数と処理時間の関係（エリーゼのために）

し,音楽的な立場から見ると,これらの記号に対し定量的な解釈を行うのは非常に困難であり,意味解析処理に対する音楽的検討が必要となるであろう．現在，タイ・スラー・ペダル記号など幾つかの記号は，後処理において処理を行っている．しかし，記号の位置情報を検出したにとどまり，演奏情報にそれら記号の意味を反映させるには至っていない．

２つめには，他の楽譜への適用である．現在のシステムはピアノ楽譜を対象にしているが，より多くのパートからなる楽譜,例えば弦楽四重奏などの楽譜への適用も考えられる．この場合，中音部記号や弦楽器特有の記号も現れてくるが，基本的な記譜の様式は同じなので，アルゴリズムの大幅な変更の必要はないであろう．また弦楽四重奏に限れば，どの段の五線にも単一の旋律しか記譜されていないのでピアノ譜より良好な認識結果が得られる可能性もある．

2.7　まとめ

本章では，ピアノ楽譜の自動認識システムについて述べた．小節単位での認識処理を行うことにより音楽的知識を有効に利用し，音楽的に高度な楽譜の認識も可能となった．また階層的な認識処理を行うことにより，楽譜上の記号の図形的・音楽的知識の利用が可能となった．

現段階では，初心者向きの楽譜においては，認識率も９０％を上回り，認識結果の修正も容易に行える．中上級者向きの楽譜では認識率は落ちるが，既存の対話形式の楽譜データ編集ツールにより修正を行うことで，実用に耐え得るものとなっている．また，上級者用の楽譜には，非常に複雑なものや，その楽譜固有の表現を使用しているものなど，認識が不可能な楽譜も多く存在する．しかし，これまで対象とされていなかったような複雑な楽譜に対しても実用に耐えうる認識率を収めた点，単なるシンボルの認識に留まらず，その意味解析を行い楽典規則に矛盾しない音楽情報を抽出している点などに，本手法の有効性が示されている．

図面解釈システムとして,このシステムは音楽や楽譜の知識を積極的に利用したトップダウン処理指向のものであり，トップダウン処理の有効性を示している．しかし，すべての知識がプログラムに埋め込まれているために,他の楽譜への適用というような変更作業にもかなりの作業が必要とされる．この問題を解決するには，知識の交換を容易にする必要がある．そのためには，処理と知識を分離したシステム構成を検討しなければならない．

参考文献

[Anzai89]　安西祐一郎：認識と学習，岩波書店，pp.183-185　(1989)

[Aoyama82]　青山，棟上：印刷楽譜の自動読み取り，画像電子学会誌，11, 5, pp.427-435　(1982)

[Fujinaga89]　Fujinaga, I. el al : Issues in the Design of an Optical Music Recognition System, Proc. of 1989 ICMC, pp.113-116　(1989)

[Inokuci81]　井口征士：音楽情報の処理とデータベース，日本音響学会誌，37, 11, pp.563-568 (1981)

[Inokuci86]　井口，今井：音楽情報の認知，ＡＩジャーナル，5, pp.68-73　(1986)

[Kato88] 加藤博一, 井口征士: 小節単位処理に基づいたピアノ楽譜の自動認識, 電子情報通信学会論文誌, J71-D, 5, pp.894-901 (1988)

[Kato90] Hirokazu Kato and Seiji Inokuchi : A Recognition System for Printed Piano Music Using Musical Knowledge and Constraints, Pre-proc. of SSPR90, pp.231-248 (1990)

[Kato92] Hirokazu Kato and Seiji Inokuchi : A Recognition System for Printed Piano Music using Musical Knowledge and Constraints, Structured Document Image Analysis, Springer-Verlag, pp.435-455 (1992)

[Matsusim85] 松島, 金森, 大照: 楽譜の自動認識システム（ＷＡＢＯＴ－２の視覚系）, ロボット学会誌, 3, 4, pp.354-361 (1985)

[Nakaura79] 中村, 進藤, 井口: 音楽情報の入力とデータベースの作成, 信学技報, PRL78-73, pp.41-50 (1979)

[Ohteru85] 大照, 松島, 金森: 楽譜の自動認識, 日本音響学会誌, 41, 6, pp.412-415 (1985)

[Rich83] Elaine Rich : Artificial Interigence, McGraw-Hill (1983) , 廣田, 宮村訳: 人工知能Ｉ, マグロウヒル, pp.67-72 (1984)

[Shimoda82] 下田和男: 楽典 －理論と実際－, 共同音楽出版社 (1982)

[Tojo82] A. Tojo, H. Aoyama : Automatic Recognition of Music Score, Proc. of 6th ICPR, p.1223 (1982)

[Yachda90] 谷内田正彦: ロボットビジョン, 昭晃堂, p.124 (1990)

3章　知識処理に基づく図面解釈
―認識・解釈処理と知識の分離―

前章では，図面解釈におけるトップダウン手法の有効性を示した．しかし，多種多様な知識を処理の至る所で使用するために，それら知識がプログラムに埋め込まれていた場合，システムの仕様変更や汎用化が難しいという問題点も明らかとなった．この章では，汎用的図面解釈システムの構築に向けて，処理と知識の分離，シンボル切り出しの問題の２点から新たな認識手法の提案を行う．

提案手法の有効性を実証するために，ラフに手書きされた図面への対応を図った論理回路図面の解釈システムのプロトタイプを構築した．しかし，十分満足のいく結果は得られなかった．そこで明らかとなった問題点について述べ，その解決策について検討を行う．

3.1　認識・解釈処理と知識の分離

3.1.1　汎用化の２つの軸

例えば音声，文字，図面など，本来人間と人間の間でのコミュニケーションのために使用されるメディアを対象とした認識･解釈システムにおいては，そのシステムが取り扱う領域を表現する２つの軸がある．パターンの構造と意味との関係の違いによる分類である，いわゆる種類がその１つの軸である．音声であれば「日本語」「英語」「フランス語」，文字であれば「アルファベット」「漢字」「アラビア文字」，図面であれば「楽譜」「電子回路図面」「建築図面」などの分類を表す軸である．もう１つはパターンの揺らぎやノイズの混入度合いのようなメディアの質を表す軸である．この軸は信号処理論的にはS/N比という定量的な数値で表現できるかもしれない．しかし人間の個人性によるパターンの揺らぎは，そもそも標準というものが定義できないために，定量化しにくい．汎用システムについて検討するには，これら２つの軸に着目しなければならない．

種類を表す軸での汎用化は，２つのスケールに分けることができる．小さなスケールでの汎用化とは，例えば文章解釈システムにおいて「第一水準文字」対応のシステムを「第二水準文字」対応のシステムへと拡張する汎用化である．ここでは，構成要素の違いに対処するだけでよく，関係解析の変更は必要とされない．大きなスケールでの汎用化とは，「名刺解釈システム」を「新聞解釈」にも対応可能にする問題である．ここでは，関係解析部における知識の大幅な変更が必要となる．前者はパターン処理に関連した問題であるが，後者は記号処理にその重点が置かれる．本論文は，パターン認識，画像認識に関する問題解決を検討しているものであるために，前者に関する検討が中心となる．

パターンの揺らぎに対しての対応は，パターン処理の手法自体から見直す必要があるために非常に難しい．これは，文字認識が印刷英数字と印刷漢字を対象とした場合の比較と，印刷英数字と手書き英数字を対象とした場合の比較において，後者の方が技術的な格差が大きいのが明らかなことでわかる．

3.1.2　認識・解釈処理と知識の分離

知識を用いたトップダウン手法において処理の汎用化を目指すためには,種類の軸における観点からでは,汎用的な処理系の構築とその処理系とは独立した知識を準備しておく必要がある．一般にパターン認識・理解における処理系は，図3-1のように，パターンを入力信号として処理するパターン処理系と記号を入力信号として処理する記号処理系に分類できる.知識処理は記号処理の側面が非常に強いので，記号処理系において，処理と知識の分離，処理の汎用化は比較的容易である．しかし，パターン処理系においては，これまでに深く考えられていない．

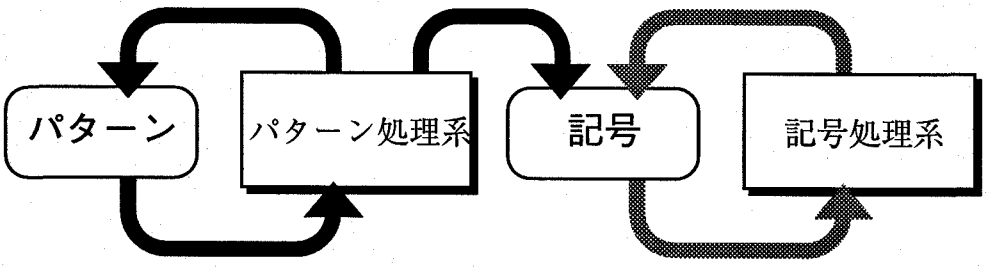


図3-1　パターン認識・理解の処理系

パターン処理系は,切り出し処理と認識処理に分けて考えることができる.認識処理に関しては,テンプレートマッチングをはじめとするマッチング手法において,マッチングエンジンとモデル（知識）の分離が実現できている．しかし，切り出し手法においては，汎用的切り出しエンジンとモデルという考え方は一般的な処理系としてあまり考えられていない．

従来の図面解釈システムにおいては,文字やシンボルのもつ普遍的な特徴を利用した手法では質の悪い図面に対処できないために,図面固有の知識が積極的に用いられていた.それでもなお,文字やシンボルが互いに接触してはいけない,大きさが一定でなければいけないなどという制約を必要とした.構成要素の認識や関係解析においては知識を有効に利用する方法を構成しやすいが,切り出し処理においては,知識を有効に利用するための方法論が確立されていない．そのため，切り出し処理の性能を向上させるために，上記入力図面に対する制約が必要とされた．2章で示した楽譜認識の場合は,各シンボル毎にそのシンボルの切り出しに適した手法を開発したために，画質の変化に対しても対応できるシステムが構築できた．つまり，切り出し処理においては,処理と知識の分離以前に知識を有効に利用した切り出し手法自体が確立されていないという大きな問題が残されている [Osawa89].

つまり,汎用的図面解釈システムを構築するためには,解釈システム全体としての知識と独立した汎用的処理系の考案と汎用的な切り出し処理の確立という2点で検討しなければならない.この章では,論理回路図面を例として汎用的処理系の枠組みと切り出し処理に対する工夫を説明する [Kato90] [Kato91].

3.2　論理回路図面解釈システムのコンセプト

3.2.1　論理回路図面の特徴

論理回路図面は，複数のシンボルとそれらの接続関係を表現する図面である．この図面によって伝達される情報は，論理的・客観的な意味情報である．個々のシンボルの形状の歪み方とか傾き具合などのような書き手の個人的特徴を反映する主観的情報を読みとる必要はない．論理的意味情報の世界では，シンボルの接続関係に何らかの制約条件が存在するため，非常にラフに描かれた図面でも，人間はそれらを利用することにより図面上に書かれた情報を完全に読みとることができる．

ラフに手書きされた簡単な論理回路図面の例を図3-2に示す．この図面は，幾つかの不正な接続線のとぎれやシンボルの変形を含んでいる．しかし，我々は，次の論理式で示される情報をこの図面から容易に読みとることができる．

$$D = A + B \times C \qquad \text{(式3.1)}$$

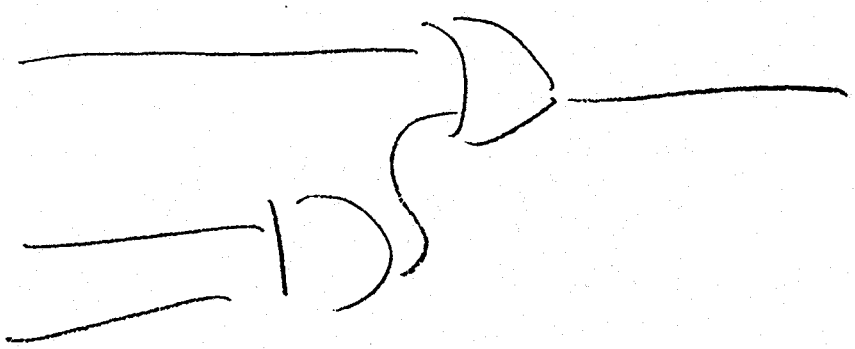


図3-2　ラフに手書きされた簡単な論理回路図面の例

3.2.2　解決すべき問題点

3.1において基本的な問題点は説明したが，ラフに書かれた図面を対象とする場合に，さらに派生する問題点がある．それは，形状に関してのあいまいさは形状に関する情報のみで解消することができないということである．これに関して，図3-3を用いて説明する．図中のA地点において線の途切れが存在する．もしこの線の途切れが正式なとぎれであるなら，この図面の意味する情報は次の論理式により表させる．

$$\begin{cases} X_1 = A + D \\ X_2 = B \times C \end{cases} \qquad \text{(式3.2)}$$

もしこの部分が不正な途切れなら，この図面の意味する情報は次の論理式により表させる．

$$X = A + B \times C \qquad \text{(式3.3)}$$

ほとんどの人間は後者の解釈を選択すると思われる．このとき途切れている接続線の端点間の正確な距離はこの判断に重要かもしれないが，この情報のみを判断基準とするのは非常に危

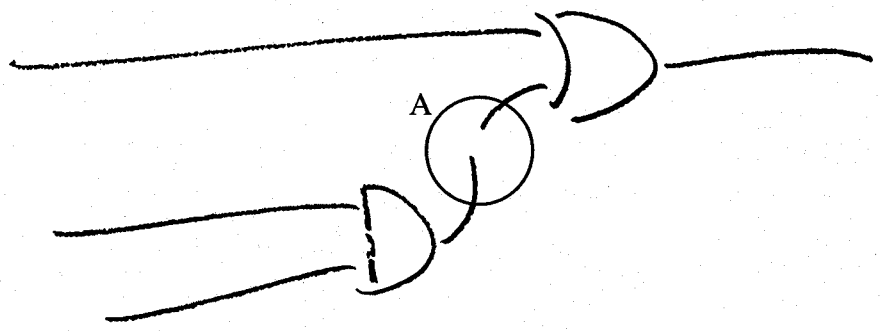


図3-3　接続線の途切れを含む図面

険である．

ANDゲートとORゲートを表す幾つかのシンボル図形を図3-4に示す．この図の中に，いくつANDゲートがあるであろうか．確かな答えは，それを書いた本人しかわからない．なぜならシンボルの類似度の評価基準は，個々の人間によって異なる主観的なものだからである．そのためにシステムが図面中のある部分に対して，形状特徴のみからその部分が表す意味を決定するのは困難である．これは文字認識において古くからいわれてきた問題点である．形状特徴の他にも多くの知識を利用する必要がある．

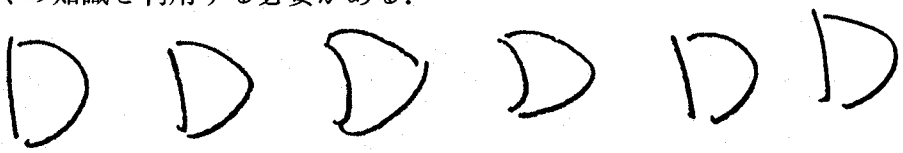


図3-4　ANDゲートとORゲート

3.2.3　階層的知識の相互利用

前節で述べた問題点を解決するために，次の2つのコンセプトに基づく新しい認識手法を提案する．

1) 対象の切り出し処理に厳密な結果を要求しない．多くの仮説を確信度情報を伴う形で生成する．
2) 多様な知識を用いた処理を導入する．それら知識を利用することで形状のあいまいさを解消する．

次節では，これらのコンセプトに基づく認識手法について述べる．

3.3　処理の概要

3.3.1　階層的知識構造

処理系とは独立した多様な知識を利用するために，まずその分類がシステム設計において重

Symbol name : AND

Function for calculation of the certainty factor:
CF_{and} (line1, line2, line3, arc1);

Elements:　LINE, 3 ; ARC, 1

Fundamental positionship

○: Connecting point between the primitives

Position to connect with the other symbols

●: Output point

: Input point

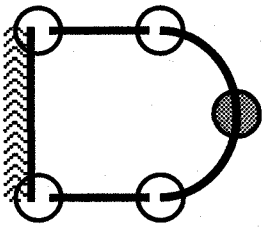


図3-5　ANDゲートの形状表現

要となる．ここで提案する認識手法は，次の7種類の知識を使用する．

1) シンボルの形状を表現する知識

このカテゴリーに属する知識は，図形世界の形状と意味世界の記号の対応を表現するものである．ANDゲートの例を図3-5に示す．この記述は，シンボルの構成要素，それら要素間の位置関係，他のシンボルとの位置関係を表現するために用いられる特徴点，確信度を計算するための関数からなる．

2) 図面の書き方に関する知識

図面の書き方を拘束する規則がある．例えば，接続線の連結する場所では黒丸を書きなさい，というようなものである．一般にこのカテゴリーの規則は，図面を拘束するものであるが，実際の図面では違反して書かれることも多い．

3) 意味世界での知識

論理図面は意味情報を伝達する．この情報の中には多くの拘束条件が存在する．図面として書かれた情報は，これらの拘束条件を満足しなければならない．

4) 経験的に学習した形状の典型的パターン

人間は繰り返し同じ形状パターンを見ると，そのパターンを容易に認識できるように

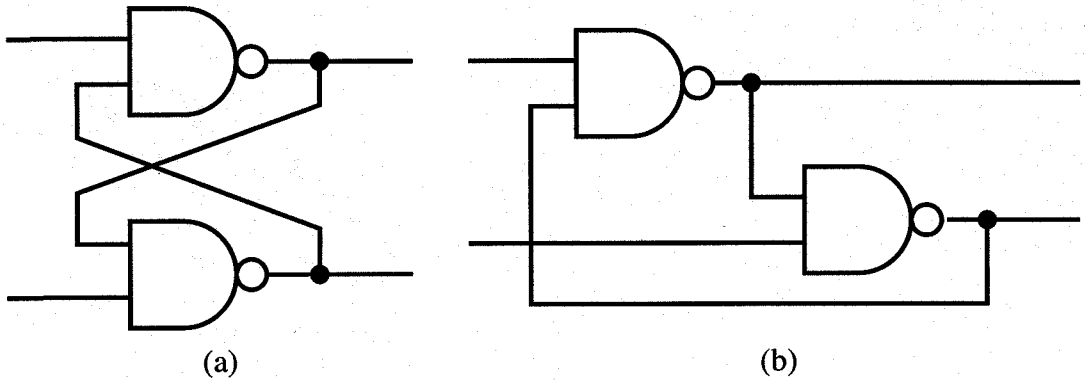


図3-6　同じ意味を表現する図面

なる．同じ意味を持つ２つの図面を図3-6に示す．２つの図面の差異は，シンボルの配置のみである．もし，(a)の配置のパターンをこれまでに何度も見たことのある人であれば，(b) の配置の図面より，より容易にその図面の意味する情報を読みとることができると思われる．この場合，(a) においては，個々のシンボルの接続関係を個別に調べているのではなく，図面全体を１つのパターンとして認識していると考えることができる．

5) 意味解析のための知識

図面によって表現される意味情報には，より高次の情報に変換可能な部分がある．図3-6が示す図面はRS-F.F.として解釈でき，その解釈結果が図面全体としての解釈に安定性を与えると考えられる．このカテゴリーに属する知識による意味解析によって，シンボルの接続関係に意味的な安定性を与える．

6) 主観的な知識

意味解析のための知識の他にも，図面に対し安定性を与える知識がある．これらの知識は，主観的な感覚に起因する．一般的に，丁寧に書かれた図面はラフに書かれた図面より安定とみなすことができる．しかし，丁寧とかラフの定義は個人によって異なる．そのような知識も認識には重要である．きれいな図面を表現する主観的知識の一例を以下に示す．

すべてのシンボルの大きさや方向は同じである．

7) 一般的語彙

前述の１～６の知識はすべて宣言型知識として表現される．一般的語彙は，それら宣言型知識を表現するために用いられる知識の集合であり，そのほとんどが手続き型知識である．図面からの抽出処理や検証処理のすべては，一般的語彙の手続き型知識によって処理される．

１から６の知識は処理と独立になるが，7の知識は手続きとしてプログラムに埋め込まれる．また，1から6の知識は拘束性のあるものとないものに分けることもできる．

3.3.2　概要

提案する認識・解釈手法の概念図を図3-7に示す．シンボル仮説の生成処理（ＧＳＨ）は，前処理に当たる．3.2.3で

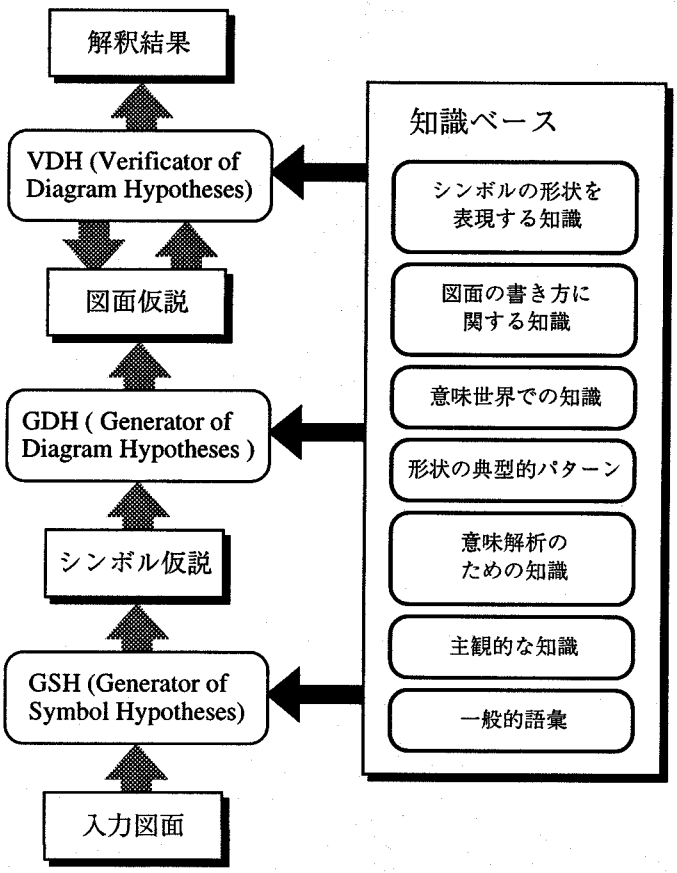


図3-7　認識処理の構成

述べたコンセプトに基づき，多くのシンボル仮説がこの処理により生成される．

処理の中心的なものは，図面仮説の生成処理（ＧＤＨ）と図面仮説の検証処理（ＶＤＨ）である．この２つの処理により取り扱われる情報が，図面仮説である．それは，シンボル仮説をノードとし，接続関係仮説をアークとするネットワークとして表現される．各ノードやアークは形状や位置に基づく確信度を持っている．図面仮説の例を図3-8に示す．

ＧＤＨは，ＧＳＨにより生成された多くのシンボル仮説の中から幾つかの仮説を抽出し，さらにそれらの接続関係を検証することで，図面仮説を生成する．もし，ＧＤＨが適切な戦略によって処理をしなければ，組み合わせの爆発が起きる．ＧＤＨは，前向き推論と後ろ向き推論の組み合わせた抽出手法を採用することで，この問題に対処している．

ＶＤＨは，ＧＤＨによって生成されたすべての図面仮説の検証を行う．図面仮説が幾つかの矛盾を含む場合，ＶＤＨはその原因を特定する．もし，矛盾を発見することができなければ，ＶＤＨは，ある評価値を計算する．最も大きな評価値の与えられた図面仮説が認識結果となる．

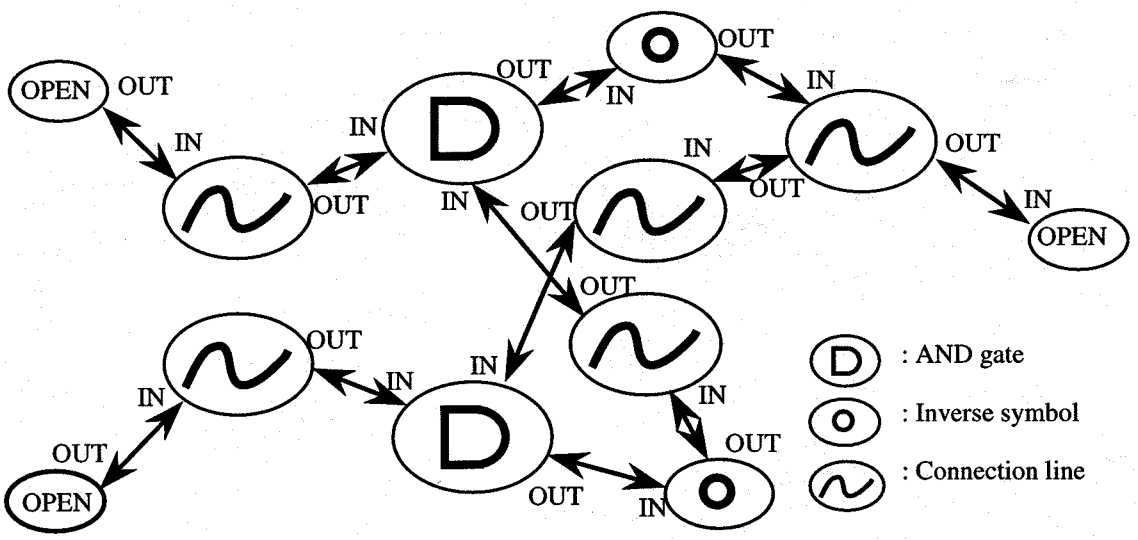


図3-8　図面仮説の例

3.4　シンボル仮説の生成

3.4.1　プリミティブ仮説の生成

すべてのシンボルは，より簡単な形状の部分要素を意味するプリミティブによって構成される．図3-9に示されるシンボルは，３種類のプリミティブ，"STRAIGHT LINE", "CIRCLE", "ARC" から構成される．論理回路図面におけるシンボルのほとんどは形状の簡単なものであり，これら３つのプリミティブの組み合わせで表現できる．もしこれら３つのプリミティブの組み合わせで表現できないシンボルを対象とするなら，そのシンボルを表現するための適当なプリミティブ抽出手続きを追加しなければならない．

入力図面は，二値化・細線化・折れ線近似処理を通してベクトルデータに変換される．ベクトルデータとは，二地点間を結ぶ単線分をその両端の座標値によって表現したデータである．幾つかのベクトルデータの組み合わせに対し，図3-10のようにプリミティブのフィッティン

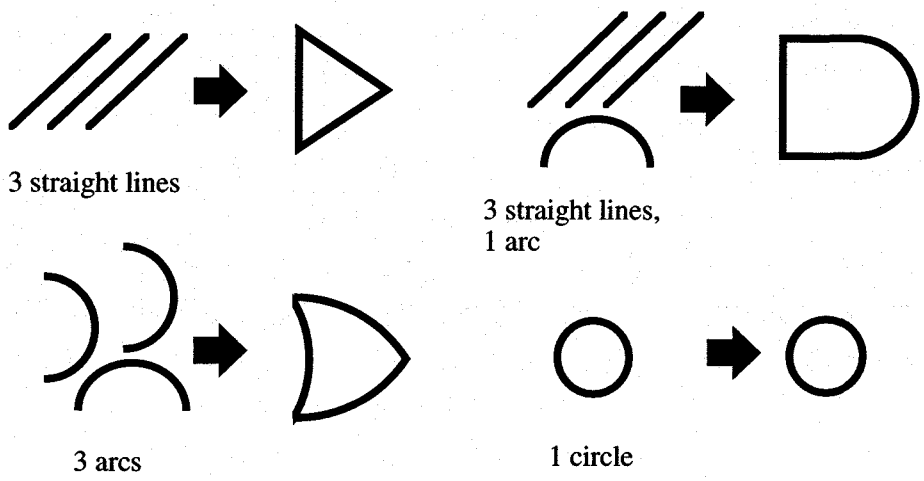


図3-9　シンボルとプリミティブの関係

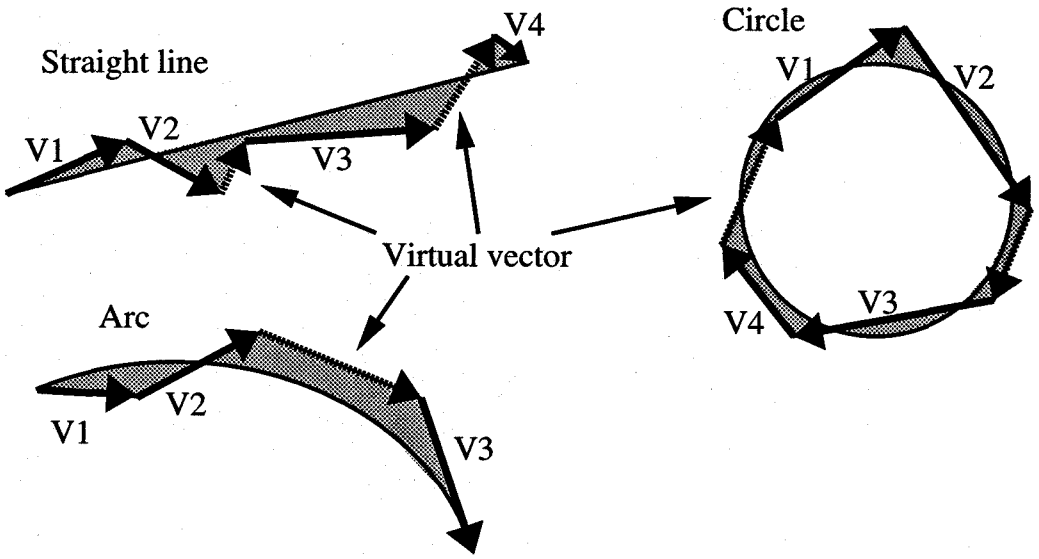


図3-10　プリミティブのフィッティング

グを行い，その類似度から確信度を計算する．このフィッティングされたプリミティブがプリミティブ仮説となる．

ベクトルデータの選択，プリミティブのフィッティング，確信度の計算の3つの処理は各プリミティブ毎に経験的に調整した処理を用いる．フィッティングすべきベクトルデータの選択に用いる閾値は，多くの可能性を保持するために非常に低い値に設定している．しかし，フィッティングの結果，確信度があまりにも低い仮説は棄却する．

この処理があまりにも多くの時間を費やすという問題に関しては，神経質になる必要はない．ベクトルデータの組み合わせごとの処理は完全に独立しているので，並列処理手法が容易に適用できる．

3.4.2　シンボル仮説の合成

3.3.1で述べたシンボルの形状を表現する知識は，シンボル仮説の合成において使用される．この処理の基本的な考え方は，前節で述べたプリミティブ仮説の生成処理と同一である．構成要素とそれらの位置関係を表す知識に基づき，適切なプリミティブ仮説の組み合わせが選択される．シンボル仮説の確信度は，使用するプリミティブ仮説の確信度とそれらの位置関係から知識として表現される確信度関数により計算される．この処理における計算時間の問題もプリミティブ仮説の生成処理の場合と同じく重要でない．前節，およびここで用いる確信度計算関数は，経験的に決定したものである．

3.5　図面仮説の生成

図面仮説の生成処理は，いくつかのシンボル仮説を抽出し，それらの間の接続関係を検査することである．組み合わせの爆発を避けるために，できるだけ早い段階で正しい図面仮説を生成できることが期待される処理手法が望まれる．処理の構成を図3-11に示す．前向き推論処理と後ろ向き推論処理が同時に独立にシンボル仮説を抽出し，それらの間の接続関係を調べる．ここで，それぞれの処理に対する確信度が計算され，より大きな確信度を返す処理の持つ仮説が図面仮説に追加される．以下では，各処理について説明する．

3.5.1　前向き推論処理

基本的には，高い確信度の与えられているシンボル仮説は，正しいシンボルを表すものと考えられる．この考えに基づき，ボトムアップ処理では高い確信度を持つ順にシンボル仮説を抽出する．この処理では，接続関係仮説は生成されない．この処理によって返される確信度は，抽出するシンボル仮説の持つ確信度と同一である．事前にすべてのシンボル仮説を確信度順に並べ替えておくことで，この処理は非常に高速に実行できる．

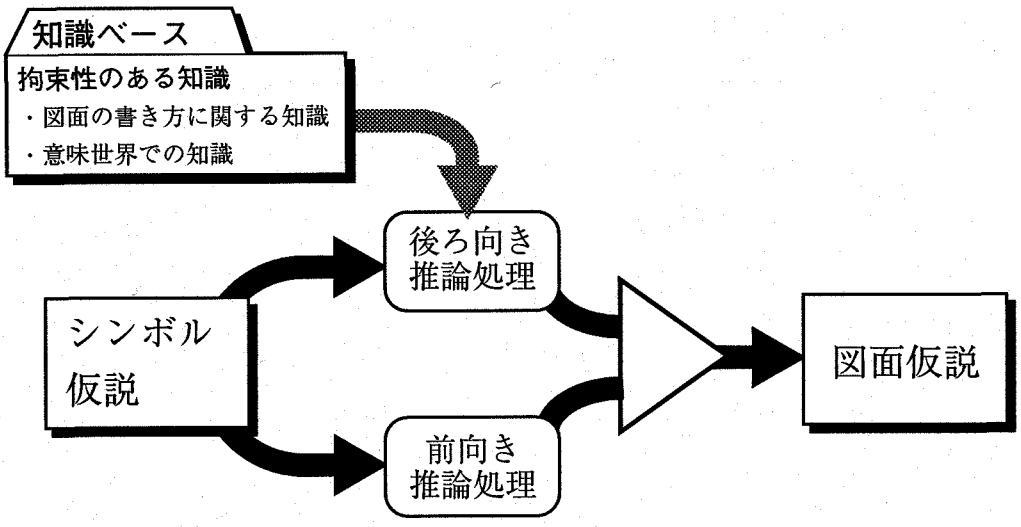


図3-11　図面仮説生成処理の構成

3.5.2 後向き推論処理

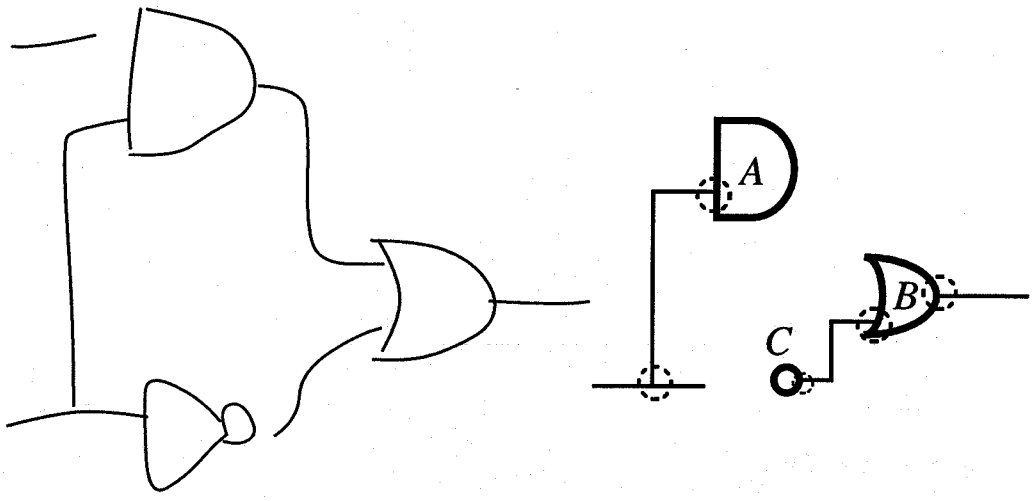
もし正しいシンボル仮説に対し低い確信度が与えられていれば，前向き推論処理のみでは，うまく対応できない．後向き推論処理は接続関係を持つシンボル間の大きさの比率や意味世界における局所的な接続関係によって表現される拘束条件を利用する．これらの知識は，図面の書き方に関する知識，意味世界での知識の中に含まれるものである．図3-12を用いて，この処理の概念を説明する．入力図面を(a)に示す．すでに図面仮説内に抽出されているシンボル仮説と接続関係仮説（破線円で表示）を(b)に示す．

シンボル仮説A（ANDゲート）に注目すると，Aの入力は1つの接続線とすでに接続関係にある．この接続関係の他にAには他の接続関係はない．もし仮説Aが正しい仮説であれば，Aは入力にあと一つ以上の，出力に1つの接続関係を持たなければならない．さらにAと接続関係を持つシンボルの種類は限定され，その大きさも予想できる．具体的なルールは，記述される知識の内容に依存する．仮説B，Cに関しても同様の考えが適応できる．トップダウン処理はこのような考え方に基づき，シンボル仮説と接続関係仮説を抽出する．つまり，シンボル仮説と接続関係仮説の間には依存関係が形成される．もし該当するシンボル仮説がすでに図面仮説内に抽出されていた場合には，接続関係仮説のみが新たに抽出される．この処理により返される確信度CF_rは次の式によって計算される．

$$CF_r = (CF_{symbol} \times CF_{size} + CF_{connect})/2 \qquad (式3.4)$$

CF_{symbol}は，抽出されるシンボル仮説の持つ確信度である．$CF_{connect}$は接続関係に与えられる確信度である．CF_{size}は抽出されるシンボルの大きさの評価として計算される値である．この関数も経験的に決められたものである．

この処理は低い確信度を持つ正しい仮説にも十分対処可能である．しかしもし図面仮説の中に高い確信度を持つ間違った仮説がすでに存在していた場合，それを取り除くことはこれまでの処理では不可能である．次節で述べる処理は，この問題に対処するものである．



(a) (b)

図3-12 後ろ向き推論処理の説明

(a) 入力画像，(b) すでに抽出済みのシンボル仮説と接続関係仮説

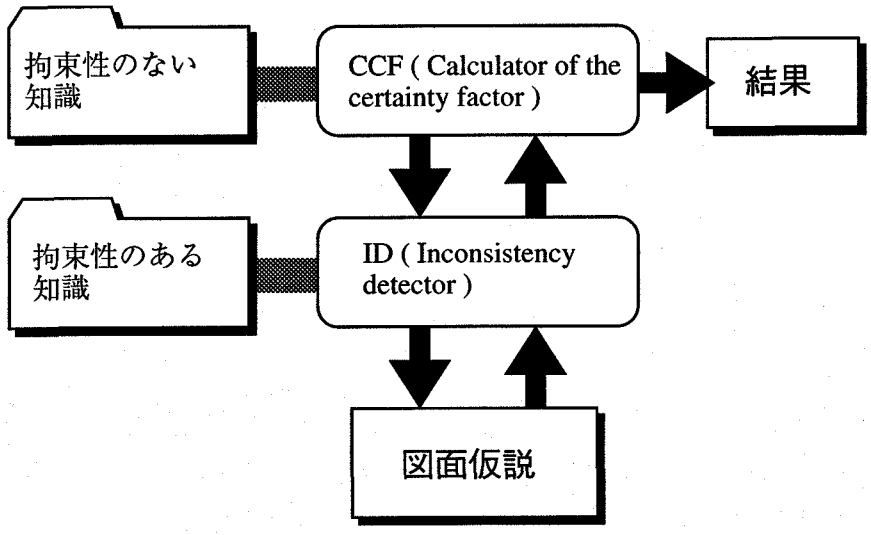


図 3-13　図面仮説検証処理の構成

3.6　図面仮説の検証

図面仮説の検証処理の構成を図 3-13 に示す．新たなシンボル仮説や接続関係仮説がＧＤＨによって図面仮説に追加されたとき，矛盾検証処理（ＩＤ）は拘束性のある知識を用いて図面仮説を検証する．いくつかの矛盾が図面仮説内に存在した場合，ＶＤＨはその原因を特定する．原因となるシンボル仮説や接続関係仮説は棄却され図面仮説内の整合性が維持される．この目的のためにＴＭＳ (Truth Maintenance System) を利用する [Doyle79]．図面仮説内に矛盾が存在しない場合，確信度計算処理（ＣＣＦ）に処理が引き継がれる．ＣＣＦは図面仮説内のすべてのシンボル仮説や接続関係仮説が持つ確信度，および拘束性のない知識を用いて，図面仮説の大局的な確信度を計算する．最も高い対極的確信度を持つ図面仮説が最終的な認識結果となる．

3.6.1　矛盾検証

矛盾検証処理は３段階の処理により構成される．最初の処理は図面仮説内の矛盾を調べる．3.5 で述べたように，もし高い確信度を持つ間違ったシンボル仮説が存在した場合，間違った図面仮説が生成される．図面仮説内の間違ったシンボル仮説や接続関係仮説の存在を調べるために，拘束性のある知識が利用される．図面の書き方に関する知識や意味世界での知識がこれらの知識に該当する．この処理においては矛盾の存在が確認できるが，その原因まではわからない．

２番目の処理は，矛盾の原因を特定する処理である．この処理は，3.6.2で説明するＣＣＦの計算結果を利用する．矛盾関係にあるシンボル仮説や接続関係仮説とネットワークトポロジー的に一定の距離内に存在する仮説集合が，図面仮説の中から取り出される．これらの集合に対する確信度がＣＣＦにより計算される．そのとき最も低い確信度をとる仮説が図面仮説から棄却される．

このとき図面仮説内の整合性が維持されなければならない．ＴＭＳがこの目的に使用される．ＧＤＨによるトップダウン処理は，シンボル仮説と接続関係仮説の間に依存関係を形成する．それ故，ある仮説が棄却された場合，その仮説によって指示されていたすべての仮説も棄却されなければならない．ＴＭＳの利用は，それがこのような処理の実現に適しているからである．

ＧＤＨによって抽出される仮説がなくなったとき，この処理は特に低い確信度を持つ仮説を棄却する．そうすることで，さらにいくつかの図面仮説を生成する．これが３番目の処理である．

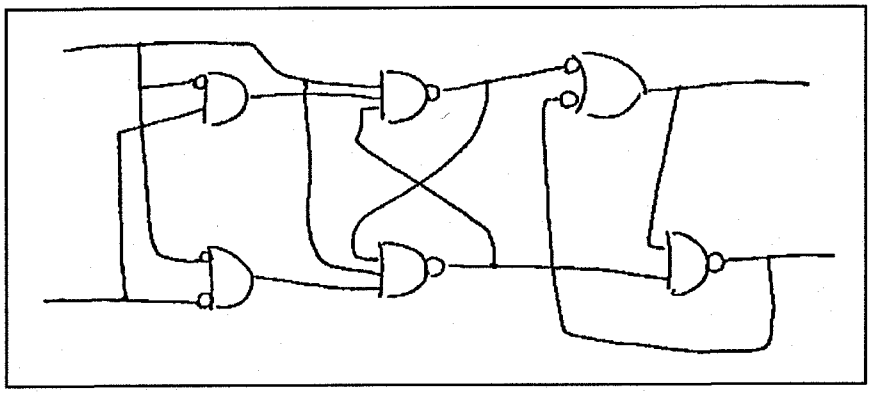
3.6.2 確信度計算

確信度計算処理（ＣＣＦ）は，ＩＤから渡される図面仮説の大局的確信度を計算する．図面仮説中のすべてのシンボル仮説や接続関係仮説は局所的な確信度を保持している．これらすべての局所的確信度を統合することにより，大局的確信度を計算する．しかし，この考え方では形状の確信度しか反映されない．それ以外の知識を大局的確信度に反映させるために，この処理では図面の書き方の規則，経験的に学習した形状の典型的パターン，意味解析のための知識，主観的な知識に含まれる拘束を伴わない知識が利用される．これらの規則はパターン（形状のパターン，意味的なパターン）とそれに与えられる得点からなる．肯定的なパターンには正の得点が，否定的パターンには負の得点が与えられる．パターンのいくつかが図面仮説の部分に一致したとき，一致した部分の持つ確信度に得点に対応した重み付けが行われる．具体的な確信度の計算方法やパターンに与えておく得点は，経験的に設定したものである．

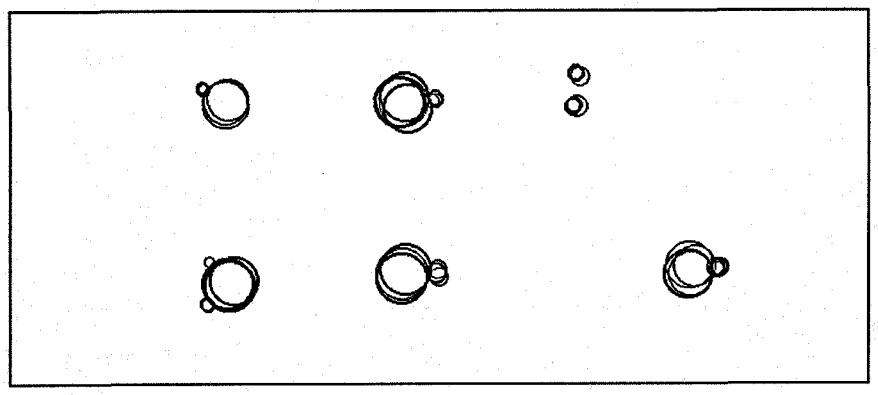
3.7 実験と検討

ここでは，いくつかの例を示す．プログラミングにはC言語およびルール記述言語OPS83を用いた．このシステムは図3-9に示す４種類のシンボルを取り扱う．図面は120DPIの解像度で入力した．ＧＳＨの実験結果を図3-14に示す．(a) は入力図面である．プリミティブ"Circle"の抽出結果を (b) に示す．シンボル"AND"の抽出結果を (c) に示す．この処理において使用した切り出しのための閾値は非常に低く設定した．そのために大量の仮説が生成された．しかし，正しいシンボルを表現する仮説がこの中に含まれているのも，図から明確に読みとることができる．認識結果を表3-1に示す．これは，いくつかのシンボルのラベルとそれらの接続関係を表す情報からなる．この情報から再構成された図面を図3-15に示す．

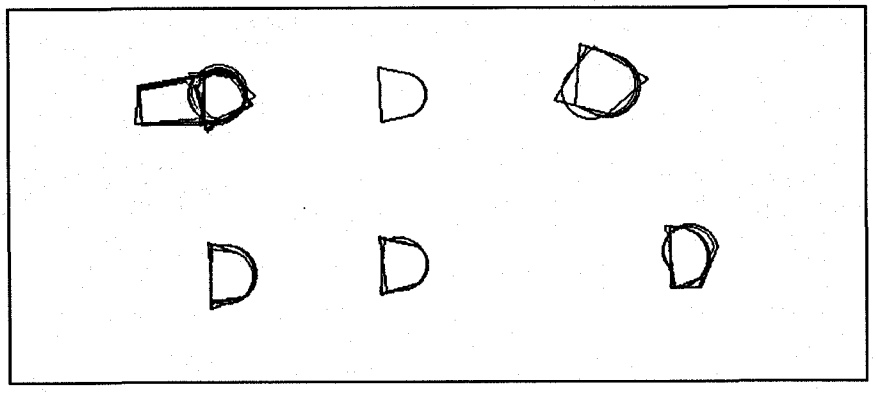
別の実験例を図3-16に示す．(a) は入力図面を示すが，非常にラフに書かれているのがわかる．(b)，(c) は図面仮説の再構成図面である．円Ａにおける線の途切れが正当な途切れと判断した場合，(b) の図面仮説が生成された．逆にその途切れを不正な途切れと判断した場合，(c) の図面仮説が生成された．形状の確信度においては，(b) は (c) より高い値をとる．なぜなら，(b) は実際の線の途切れを途切れているものとしてそのまま取り扱うのに対し，(c) は線の途切れを接続されているものと読み変えて取り扱っているからである．しかし，最終結果としては (c) が選択された．これは，主観的知識の利用の効果である．



(a)　入力画像



(b)　プリミティブ“円”の抽出結果



(c)　シンボル“AND”の抽出結果

図3-14　実験例（その1）

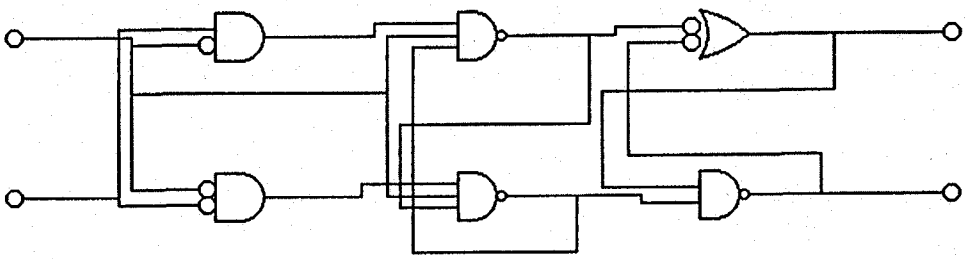
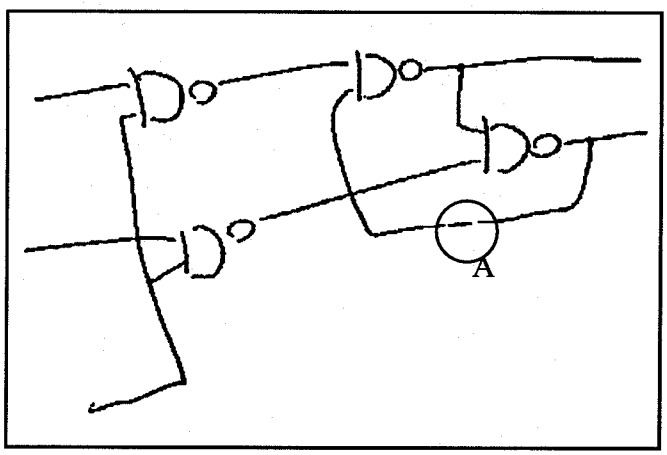


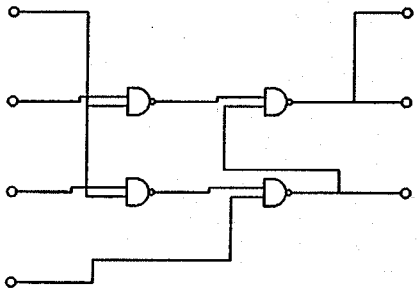
図3-15　認識結果より再構成された図面

表3-1　図3-14の認識結果

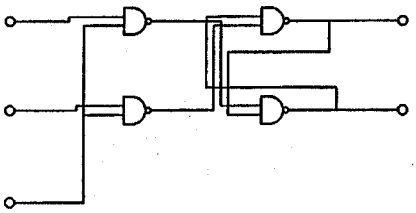
1::AND	3,OUTPUT <--> 9,OPEN
2::AND	8,OPEN <--> 1,INPUT
3::OR	3,OUTPUT <--> 5,INPUT
4::NAND	2,OUTPUT <--> 6,INPUT
5::NAND	1,OUTPUT <--> 4,INPUT
6::NAND	7,OPEN <--> 4,INPUT
7::INPUT	7,OPEN <--> 6,INPUT
8::INPUT	7,OPEN <--> 2,INPUT(INVERSE)
9::OUTPUT	7,OPEN <--> 1,INPUT(INVERSE)
10::OUTPUT	8,OPEN <--> 2,INPUT(INVERSE)
	5,OUTPUT <--> 10,OPEN
	4,OUTPUT <--> 6,INPUT
	4,OUTPUT <--> 3,INPUT(INVERSE)
	5,OUTPUT <--> 3,INPUT(INVERSE)
	6,OUTPUT <--> 4,INPUT
	6,OUTPUT <--> 5,INPUT



(a)　入力画像



(b)　棄却された図面仮説



(c)　採用された図面仮説

図3-16　実験例（その1）

3.8 まとめ

この章では，ラフに手書きされた論理回路図面の解釈を例に，知識処理に基づく図面解釈手法を提案した．汎用性があり画質の劣る図面にも対応できることを目的とし，処理と知識の分離を図った．切り出し処理における知識の積極的な利用が困難なので，画質の劣る図面に対処するために従来の方法では入力図面に制約を課していた．ここで提案する手法は，厳密な切り出しを避け多くのシンボル仮説を生成し，記号処理系において対処することを基本コンセプトの１つとした．認識結果は，この手法がラフに書かれた論理回路図面の認識に有効であることを示していた．

この手法の汎用性を評価することを目的とした実験は行われていないが，利用される知識がその種類により分類され，宣言的に表現されているために，汎用性はあると思われる．実際に認識シンボルの追加や電気回路図面のような類似した図面への対応は容易である．

画質の劣る図面への対応に関しては，最も問題となる切り出し手法に対してその問題を避ける方向での手法とした．これにより切り出しによる解釈誤りはなくなったものの，生成される仮説が非常に多くなった．この非常に多くの仮説から正しい解を得るために知識を積極的に利用することを考えた．実験例で示したような小規模な図面に対しては良好な結果が得られたものの，大きな図面になるほど生成される仮説の数やその計算コストが指数的に増加し，記号処理系においても組み合わせの爆発が生じ，満足のいく結果は得られなかった．つまり，切り出し処理における問題が先送りされただけであり，記号処理系においてその問題が表面化した．

この章で提案した手法の実験結果から，図面の種類の軸における汎用化には処理と知識を分離した手法が実際に有効であるが，画質の軸に関する汎用性，つまり画質の劣る図面への対応を考えた場合，パターン処理系，特に切り出し処理の問題点をすべて記号処理系で補うことはできなかった．記号処理系の性能の向上も重要であるが，高精度な切り出し処理を考案する必要がある．

参考文献

[Osawa89]　大沢，山川他：図面の認識と理解，昭晃堂，pp.94-95 (1989)

[Kato90]　H.Kato, S.Inokuchi : The recognition method for roughly hand-drawn logical diagrams based on hybrid utilization of multi-layered knowledge, Proc. of 10th ICPR, vol.2, pp.578-582 (1990)

[Kato91]　加藤，井口：あいまいさを考慮した手書き論理回路図面の認識，北村，片井 (編)：画像と制御，朝倉書店，pp.123-137 (1991)

[Doyle79]　J.Doyle : A truth maintenance system, Artificial Intelligence, Vol.12, pp.231-272 (1979)

4章　弾性モデルに基づく形状切り出し

人間の知識には，記号的なもののみではなく，パターン的な知識も非常に多い．言葉を持たない動物や乳児・幼児の知識のほとんどはパターン的なものと考えられる．そのようなパターン的な知識をトップダウン処理においてどのように取り扱うかを考える必要がある．その1つがパターン的知識に基づく汎用的切り出し処理の実現である．

3章では，切り出し処理の問題を解決するために，切り出し処理に対する積極的な知識の利用を避けた．形状モデルの記述を基に単純な探索処理により形状候補を切り出すこととした．これには，精度が悪く，計算コストも高いという問題点が残る．精度の点に関しては，続く記号処理系での図面仮説生成・検証処理によりその解決を図るという方針であった．しかし，記号処理系が基本的に組み合わせ探索処理であることから，間違いを多く含む仮説を対象とした場合，組み合わせ爆発に陥ることもあり，うまく機能しなかった．

つまり，知識を積極的に利用した優れた記号処理系が存在しても，やはり優れたパターン認識処理も必要とされる．そのために，本章では汎用的切り出し処理について再検討を行う．

4.1　汎用的切り出し手法

4.1.1　図面解釈における形状切り出し

1章でも述べたように優れた認識・解釈能力を有するシステムはこれまでに登場していない．これは，手書きによる変動や様々なノイズの影響に対応させることが非常に難しい問題として残されているからである．図面中のシンボルの認識手法として，大きく2種類の方法が採用されてきた．1つはテンプレートマッチングに代表されるシンボル形状を2次元パターンとして画像配列で表現し，形状全体の類似性から認識する方法である．もう1つは，シンボル形状の特徴を幾何学的概念を用いて記号や手続きで表現し，それら特徴の類似性から認識する方法である．後者の手法の方が詳しい特徴表現が可能なために認識の能力は一般に高い．

しかし，図面解釈においてはシンボルの切り出し手法まで含めて同時に議論しなければならない．前者の認識方法は，シンボル形状を表現する画像配列（モデル画像）を入力図面を表現する画像配列（入力画像）の上で走査し，逐次類似度を計算する．ここでは，モデル画像の重なり合う入力画像の領域自体が切り出し領域であり，認識と切り出しが同時に行われる．ただし，拡大・縮小変換や回転変換をうけたシンボルの認識を可能にするためには，走査方法やモデルに多少の工夫が必要とされる．

後者の方法は切り出しに関しての対処方法は手法自体に含まれておらず，問題に応じて別途検討しなければならない．これにも，大きく2つの方法がある．1つは2章の楽譜認識システムで採用した方法のように，シンボルの切り出しの手法を各シンボル毎に考案し，モデルに埋め込む方法である．この方法は，認識すべきシンボルの数が少ない場合は有効であるが，方法に一般性がなく，開発コストが非常に高い．もう1つは，すべてのシンボルに共通な特徴を捕らえ，それに基づき切り出しを行う方法である [Mukouda83]．この方法は，設計者が共通な特徴を見つけることが，認識の精度の最大の決め手となる．しかし，現実問題としてそのような

都合の良い特徴がいつも存在するとは限らない．現状のシステムでは，切り出しで使用する共通特徴の点でシンボルをグループ化し，各グループ毎に切り出し手法を検討するという両者の組み合わせが用いられる．例えば，文字のグループ，小シンボルのグループ，大シンボルのグループと分割し，それぞれのグループ毎に切り出し手法を考案し，それを用いた切り出し，さらにそれに続く認識処理という処理の流れを採用している．それでも画質の劣る図面にも対応できるようなシステムは存在しない [Ohsawa89].

4.1.2 領域抽出と領域分割

切り出しは，その目的から領域分割と領域抽出に分けられる．領域分割とは画像を互いに素な領域集合に分割する問題であり，領域抽出とは画像からある特定の領域を抽出する問題である [Wada95]．領域分割には画像中の対象に依存しない，画像自体の普遍的な特徴を利用し行われる汎用的切り出し手法が要求される．いわゆるボトムアップ処理にあたり，対象に関する知識を積極的に利用しないという特徴がある．対象に関する情報を利用することのできない初期視覚の研究で通常利用される．つまり，図面解釈という対象が明確に定められた問題に利用するのは不適切である．

領域抽出はなんらかの知識に基づき，意図する領域を画像中から抽出する切り出し問題である．図面解釈におけるシンボル領域や文字領域の切り出しは，この領域抽出にあたる．

4.1.3 汎用的切り出し手法

領域分割は対象に関する知識を利用しないボトムアップ処理であるために，汎用性の高い手法が多く提案されている [Ohta78] [Wada95]．しかし，領域抽出処理には対象に関する知識を積極的に利用したトップダウン処理が要求されるために，汎用化には処理エンジンと対象に関する知識の分離が必要である．そのような手法の1つにSNAKES [Kass88] がある．以下では，このSNAKESに関して説明する．

SNAKES

SNAKESは曲線の持つ内部エネルギーと画像等の外部のエネルギーの和を最小化するモデルである．内部のエネルギーは曲線の滑らかさなどの拘束力から生じ，SNAKESの形状により決定される．外部のエネルギーは画像中のエッジ等の特徴へSNAKESを引きつける力から生じる．図4-1はSNAKESによる唇の輪郭抽出を行った例である．実際にはSNAKESの扱うエネルギー関数は多くの極小値をもつので対象の輪郭に意図通りに収束させるには，適切な初期位置を与えなければならない．この初期位置を初期値とし，変分法によって弛緩法の枠組みで最小化問題を解く方法である．

つまり，SNAKESはエネルギー関数の最小化問題を解く汎用エンジンであり，エネルギー関数の中に対象に関する知識を導入することもできる [Sakaguchi90]．しかし，これを図面解釈のシンボル切り出しに利用することを考えた場合，自由形状をSNAKESが取り扱うことのできる評価関数として表現するのは困難であり，精度の高い切り出しは望めない．

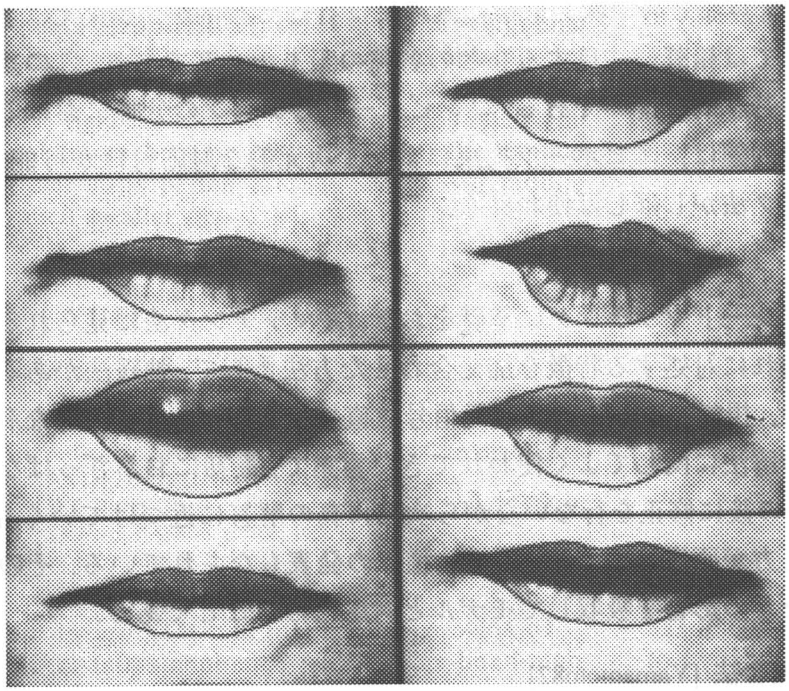


図4-1　SNAKESによる唇の抽出　[Kass88] より

4.1.4　パターンマッチング手法の利用

高精度な切り出し手法を実現するために対象の知識を積極的に利用した手法を考える必要がある．さらに汎用的手法とするために，利用する知識と独立した処理エンジンが必要となる．この手法に要求される性能をまとめると，

1) 各形状毎にその形状の特性に応じた切り出しが可能
2) 切り出しの手法に汎用性があり，モデルの構築が容易
3) 手書き変動やノイズの影響に対しロバスト

である．領域抽出は，知識として与えられたモデルに対応する部分を入力データから抜き出すわけだから，記号処理として捕らえれば探索問題，パターン処理として捕らえれば対応付け問題と考えることができる．3章での失敗は，パターン処理による知識の利用を避け，早い段階で記号化を行い記号推論による探索問題として切り出しを行おうとしたところである．ここでは，パターン処理による対応付け問題として，汎用的切り出し手法を考える．以下では，既存の汎用的対応付け手法について説明する．

DPマッチング法 [Funakubo91]

評価関数中の全変数が同時に互いに影響し合うわけではないという場合の最適化問題を解く一つの手法である．たとえば，パターンの局所的な変形をコストとして表現し，その総和が最小となるような対応付けを探索する場合に用いられる．このとき解を得るためにはいろいろな方法があるがその中でもより少ない計算量で解を求めるためにDPマッチング法が使用される．一般にDPは多段階の決定過程を扱うのに有効な手法であり，全過程の最適化がある段階までの最適化とその段階での最適値の和の再帰的過程から求められるので漸化的定式を取る．問題点は1次元系列の対象しか取り扱えないこと

である．音声などの1次元信号の処理には有効であるが，2次元以上の視覚パターン情報に対しては輪郭に注目するなどの各種の工夫が必要である．この点があとに述べる弛緩法との基本的な違いである．コンピュータビジョンの分野におけるステレオ画像間の対応付け問題は，エピポーラ拘束により1次元的な対応付けが可能であるのでこのDPマッチング法がしばしば使われる．

弛緩法

テンプレートマッチングに基礎をおき，DPマッチング法に見られるようなパターンの変形や切れに対処する特徴を加え合わせた方法として，弛緩法がある．DPマッチング法とは違い，視覚パターンの2次元信号を直接使えることが，大きな利点である．弛緩は方程式の解法に関してよく使われてきた言葉で，最初におおざっぱな形より解を求めておき，その後に修正を行ないながら次第に真の解答に近づけていくものである．このような考え方が画像処理に導入され，種々の発表もなされている．その目的は画像の変形や切断などの影響を吸収してモデルと対応付けることである．

a) 弛緩照合法 [Yamamoto82]

弛緩照合法は手書き漢字の認識における入力画像とモデルとの対応付けを行う手法である．入力の線分と辞書線分の対応付けの程度を局所的に定め，局所的に決められない場合，いくつかの対応の可能性を残しておき，弛緩法を用いて全体と個々との関係の修正を施す方法である．図4-2に例を示す．(a) は漢字「己」の入力線分，(b) は辞書線分とする．9番の辞書線分は，最初，入力線分の4，6，8番に対応するが，数回の反復処理の後辻褄合わせが行われ，(9，6) の結果，つまり入力の6番の線分が選択される．

b) 力学的リラクセーション [Shizawa88] [Webster93] [Teramura94]

図4-3に示すような弾性的な拘束を与えられた質点の集合で表される陽パターンと，重力ポテンシャル場で表される陰パターンという2つの力学的モデルを用い，力学法則に従って2つのパターンの対応付ける手法である．2つのパターンがなるべく重なりあるようにリラクセーションにより少しずつパターンを変形させていくことで形状の変動を吸収するため，変形に強い認識が可能である．一般に変形を受けるのは陽パターンであ

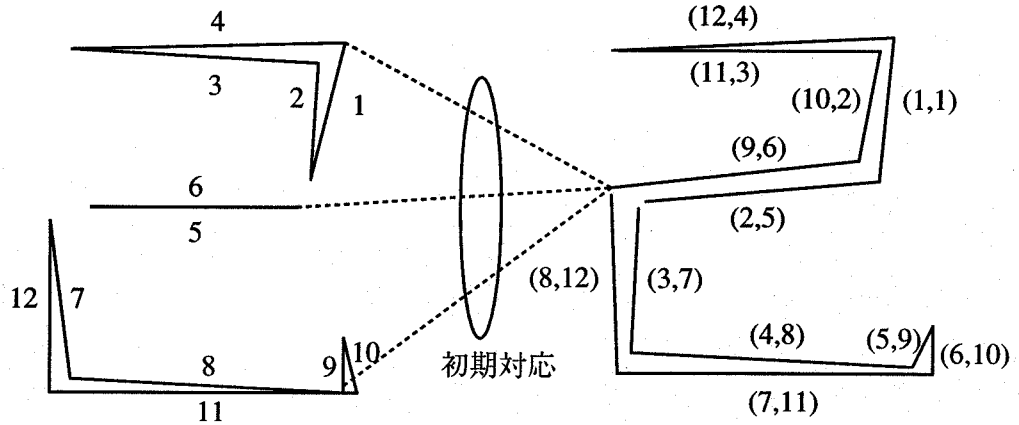


図4-2　弛緩照合法による線分の対応付け [Yamamoto82]

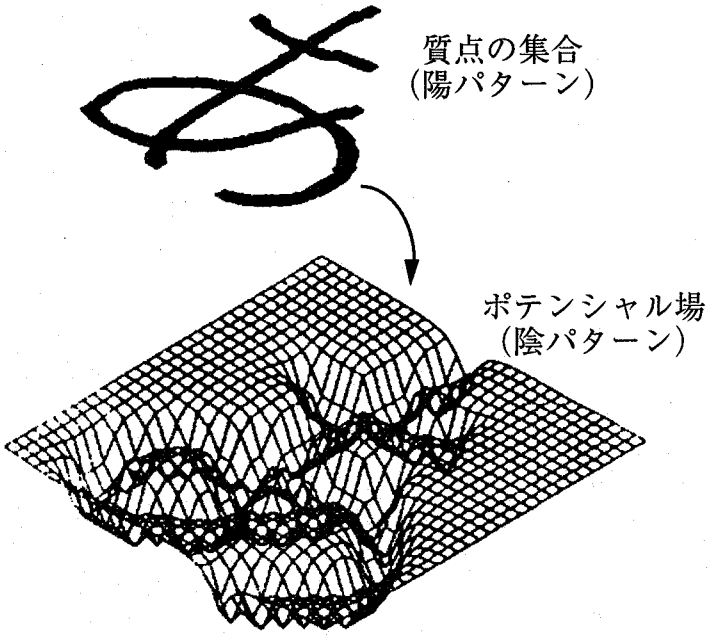


図4-3　力学的リラクセーションの概念図　[Teramura94] より

り，陰パターン上での陽パターンの持つエネルギー（質点の位置エネルギーおよび質点間の弾性エネルギーの和）が極小値に収束するまで変形が行われる．収束した状態での陽パターンの持つエネルギーや陰パターンの充填度などが相違度として用いられ，カテゴリ判断に用いられる．

これらパターンの対応付け手法は，対応付けの結果自体を得ることを目的とするものと，対応状態から計算される類似度を得ることを目的とするものに分けられる．DPマッチング法は前者に当たり，弛緩照合法，力学的リラクセーションは後者に当たる．DPマッチング法は汎用的な手法ではあるが，図面に表れるシンボル形状を表現するには定式化が難しい．そこで，後者のパターンとしてモデルを表現する手法の応用を考える．しかし，ここでは対応付けの結果自体を得ることを目的とする手法を考えるものであり，後者の手法の本来の目的とは異なることに対する工夫が必要となる．

ここでは，力学的リラクセーションに基ずく対応付け手法の切り出しへの応用を考えた[Uchiyama94]．この手法は，バネによる弾性モデルを用いる点で，[Teramura94] などにある対応付け手法と基本的に同一である．しかし，それらが類似度計算を目的としているのに対し，ここでは切り出しを目的とするところが異なる．また，繰り返し計算によりモデルを適当に変形させることにより切り出す手法としては，SNAKEに類似しているが，形状が評価関数内に数式として埋め込まれるスネークに対し，非常に記述が容易できめ細かなモデルが作れる．

4.2　弾性モデルによる切り出し

図4-4に処理の概念図を示す．入力画像に対し黒画素が低くなるように溝が彫られたような2次元的なポテンシャル場を想定する．その上から切り出したいシンボル形状を表すバネの連

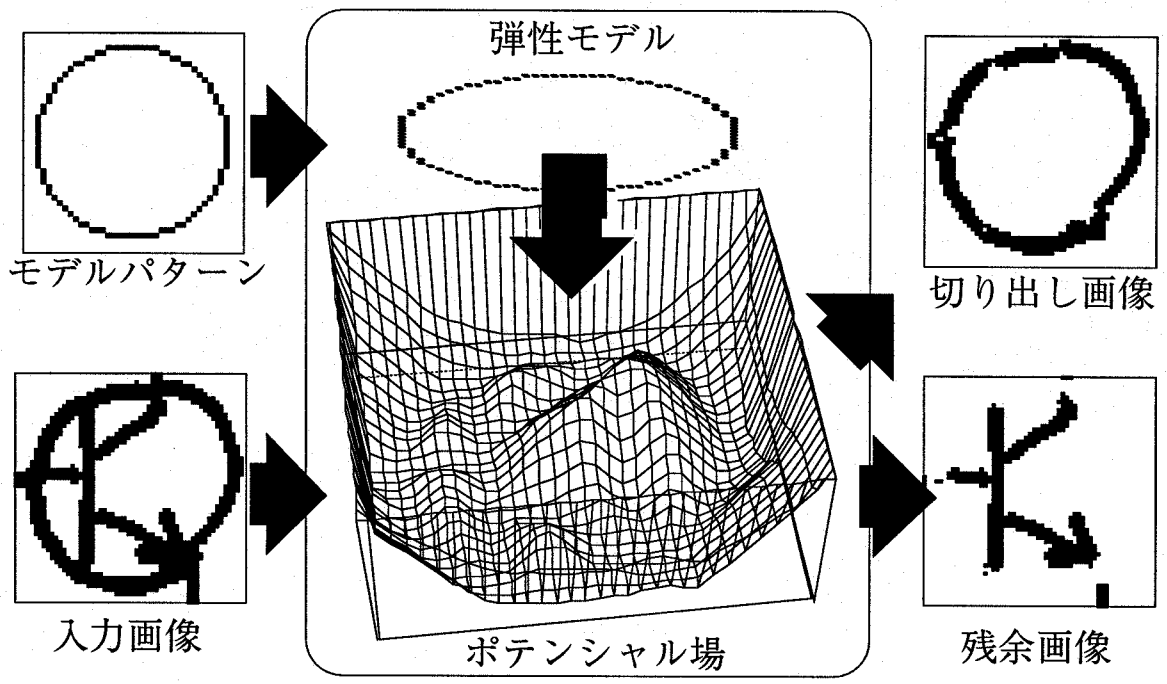


図4-4 弾性モデルによる切り出しの概念図

結したモデルを投下し重力により変形しながら溝に落ちていくという落下運動シミュレーションを行う．重力とバネの伸縮力のみではエネルギー保存則により運動が静止しないので，速度に比例する抵抗力を用いて系全体のエネルギーを減少させ，最終的にモデルが静止状態に収束するようにする．静止した時点で，溝に落ちたバネがその部分と対応しているとみなす．

4.3 弾性モデル

切り出しを行うための知識としてパターン表現された形状モデルを用いる．このモデルは切り出し処理とは独立しており，モデルを交換することで異なる形状に対する切り出しが可能である．つまり処理自体は汎用的なものであるが，その処理とモデルとを一体として見た場合はそのモデルの切り出しに対して有効に機能するトップダウンな切り出し処理となる．

図4-5に提案する弾性モデルの概念図を示す．形状を表現するための特徴点とそれらを連結するバネで構成される．特徴点には質量があるものと考え，それがポテンシャル場の中をポテンシャルの低い方向へと移動する．ばねは弾性モデルの形状の極端な変化を抑制させる目的で使われる．このバネには3種類のものを想定している．図4-5において，バネ1はポテンシャル場に投下した直後の特徴点の位置（初期位置）と特徴点を結合するバネである．入力画像とモデルの位置は大きく異ならないという前提条件のもと，特徴点が初期位置から大幅にずれてしまうのを防ぐ効果を期待する．バネ2は隣接特徴点を結ぶもので初期状態での隣接特徴点との距離を保持する効果を期待する．バネ3も隣接特徴点とを結ぶもので初期位置においての隣接特徴点との角度を保持する効果を期待する．つまり，バネ1でモデルの移動を抑制し，バネ2，3でモデルの形状変化を抑制する．

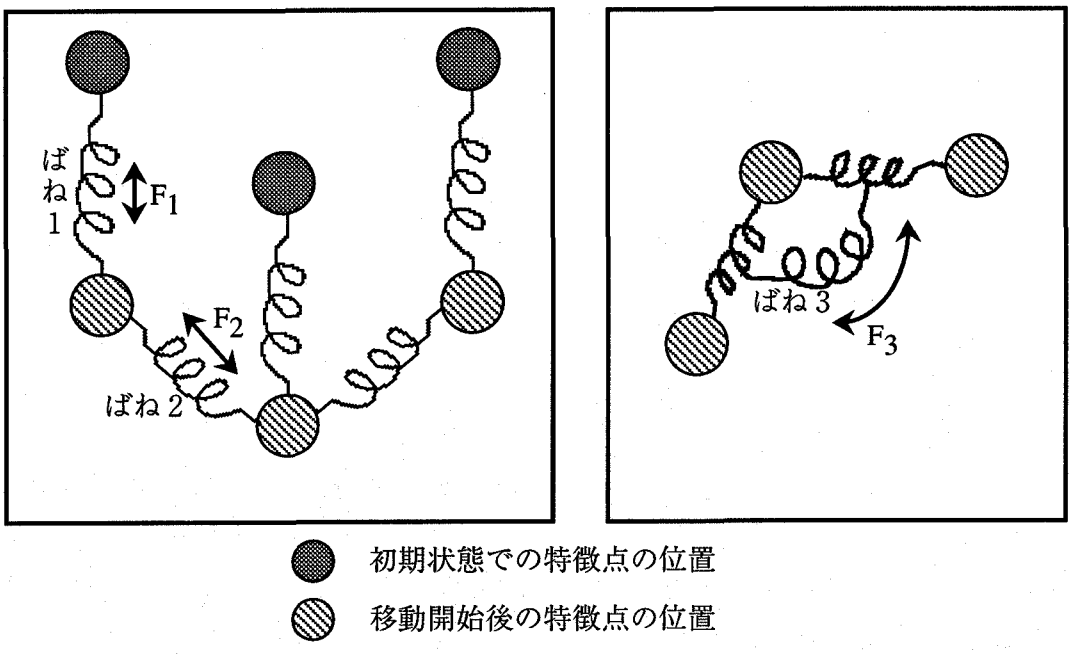


図4-5　弾性モデルの概念図

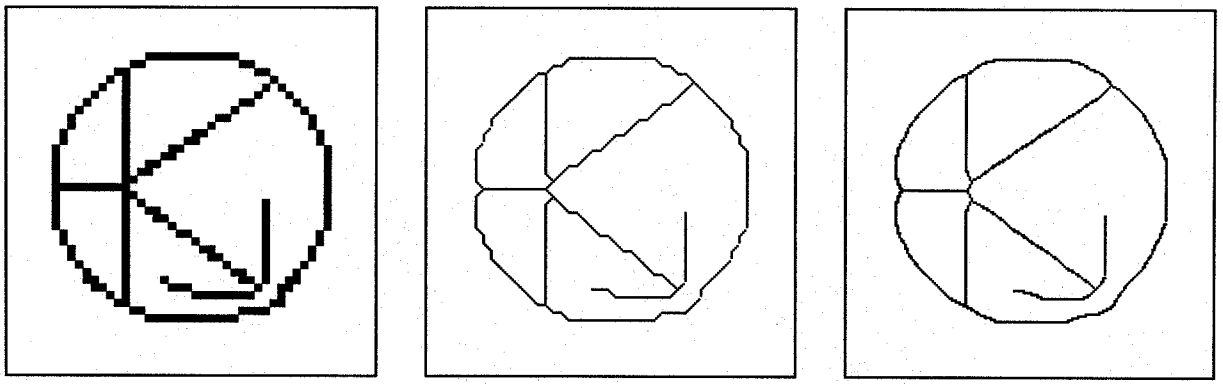
4.3.1　モデル表現

4.1.4で述べたように，システムに登録するモデルの構築はユーザーにとって容易なものでなければならない．ここでは形状を直接的に表現する2値画像をシステムに登録するモデルとして用いる．これは，テンプレート・マッチングと同一のモデル表現であり，ユーザーにとって非常にその構築が容易である．

4.3.2　弾性モデルへの変換処理

システムは以下の流れで2値画像として与えられたモデルを弾性モデルへと変換する．

1) モデル画像に8連結細線化処理を施し，細線化画像を得る．つまりこの章で述べる切り出し手法は，細線化処理が形状情報を保存するような線図形しか基本的に対応できない．面図形に対する検討は4.7.4で述べる．
2) 細線化画像から各画素の座標，隣接画素の番号の情報を表現するテーブルを作成する．
3) 各画素を特徴点し連結画素間にバネを張ることを想定する．このとき，ディジタル画像の特性から，連結画素の角度情報は8方向に量子化されている．これは本来の局所的形状情報を反映していないので，落下運動シミュレーションにおいて精度の点で問題となることが実験で確かめられた．そこで，各特徴点の位置を，その特徴点と直接連結する特徴点の位置の重心位置に置き換える，いわゆる平滑化を行う．ただし端点となっている特徴点では大きな位置変化がおきるため，その特徴点に2，連結する特徴点に1の重みを与えた平滑化を行った．この時点で，画素位置はピクセル単位の量子化精度ではなく，サブピクセル単位の表現になる．落下運動シミュレーションにおいても，サブピクセル単位で計算を行う．図4-6に平滑化処理の効果を示す．
4) 各特徴点において，連結画素との距離，角度情報を求める．図4-7に最終的な弾性モデルの表現形式を示す．



(a) 入力画像　　(b) 平滑化なし　　(c) 平滑化あり

図4-6　弾性モデルへの変換

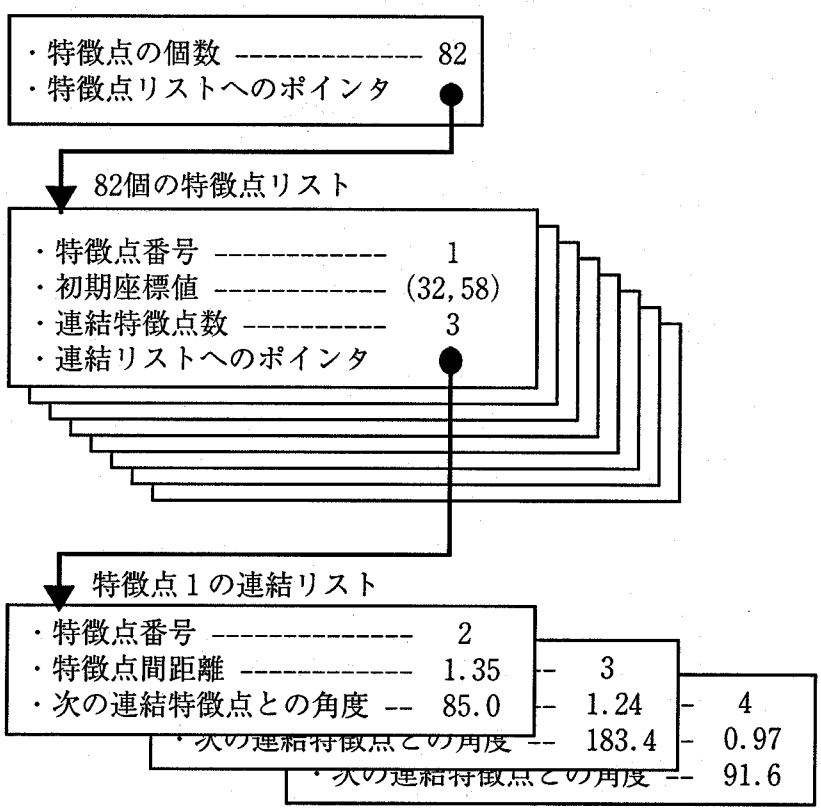


図4-7　弾性モデルの表現

4.4　ポテンシャル場の生成

入力画像からは，ポテンシャル場を以下の手順で生成する．

1) 入力画像$I(x,y)$の1-画素に距離変換を施し$D_1(x,y)$を生成する．
2) 入力画像$I(x,y)$の0-画素に距離変換を施し$D_2(x,y)$を生成する．
3) $D_2(x,y)$の画素の最大値$d2_{max}$を$D_2(x,y)$の外周画素全てに代入する．

4) if $\mathrm{I}(x_1, y_1) = 1$

then $\mathrm{D}_3(x_1, y_1) = -\mathrm{D}_1(x_1, y_1)$

else $\mathrm{D}_3(x_1, y_1) = \mathrm{D}_2(x_1, y_1) \times 10$

5) $\mathrm{D}_3(x, y)$に対してサイズ3×3の平滑化フィルタを施し，ポテンシャル場$\mathrm{P}(x, y)$を生成する．

図4-8に生成されたポテンシャル場の例を示す．

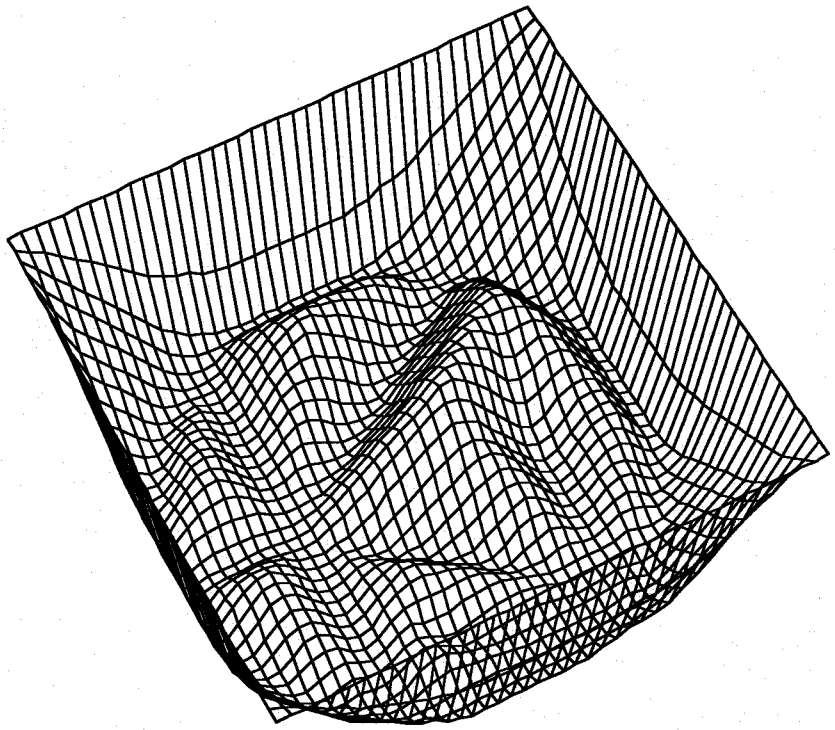


図4-8　ポテンシャル場（図4-6の画像を使用）

4.5　落下運動シミュレーション

各特徴点は接続されているバネの力とポテンシャル場からの力を受けている．この合力の大きさと方向で次の移動位置が求められる．i番目の特徴点の時刻nTにおける位置ベクトルを$\mathbf{P}_\mathrm{i}(nT)$で表し，その特徴点がバネとポテンシャル場から受ける力の合力の方向ベクトルを$\mathbf{F}_\mathrm{i}(nT)$とすると，以下の式で特徴点の位置が更新する．

$$\mathbf{P}_\mathrm{i}(nT) = \mathbf{P}_\mathrm{i}((n-1)T) + \dot{\mathbf{P}}_\mathrm{i}((n-1)T) \cdot T \qquad \text{(式4.1)}$$

$$\dot{\mathbf{P}}_\mathrm{i}(nT) = \dot{\mathbf{P}}_\mathrm{i}((n-1)T) + \ddot{\mathbf{P}}_\mathrm{i}((n-1)T) \cdot T \qquad \text{(式4.2)}$$

$$\ddot{\mathbf{P}}_\mathrm{i}(nT) = \frac{1}{m} \cdot \left\{ \mathbf{F}_\mathrm{i}(nT) - D \cdot \dot{\mathbf{P}}_\mathrm{i}(nT) \right\} \qquad \text{(式4.3)}$$

ここで，Tは離散的シミュレーションにおけるサンプリング周期，mは各特徴点の仮想質量，Dは各特徴点の移動速度に比例した減衰力を表す減衰係数である．減衰力を導入することにより運動シミュレーションの発振を防止し，収束性が保証される．

4.6 領域の抽出

前節までの処理で，入力画像とモデルの対応付けが可能となる．ここでの目的は入力画像からのモデルに対応する領域の抽出なので，この対応付け結果を基に以下の手順で領域抽出を行う．

1) 入力画像の小領域への分割

入力画像を細線化した画像に対し折れ線近似を行い，短い連結成分であるチェインデータを生成する．通常の折れ線近似 [Tamura85] では，近似誤差による評価に基づき繰り返し分割を行い，誤差が許容誤差E_{max}以下になったとき分割処理を停止するが，ここでは図4-9に示すように，分割されたチェインデータの数がC_{max}以下になるまで分割を繰り返す．パラメータの具体的な数値として，E_{max}=1.0, C_{max}=7としている．ただし，精度を向上させるために，端点や交差点に接するデータにおいては，C_{max}=3としている．

入力画像の各チェインデータに対応する画素に対しチェインデータ単位でラベル付けを行う．このとき交差点のチェインデータは取り除いておく．そのラベルを入力画像の1-画素の領域内で膨張させることで，入力画像が小領域に分割される．

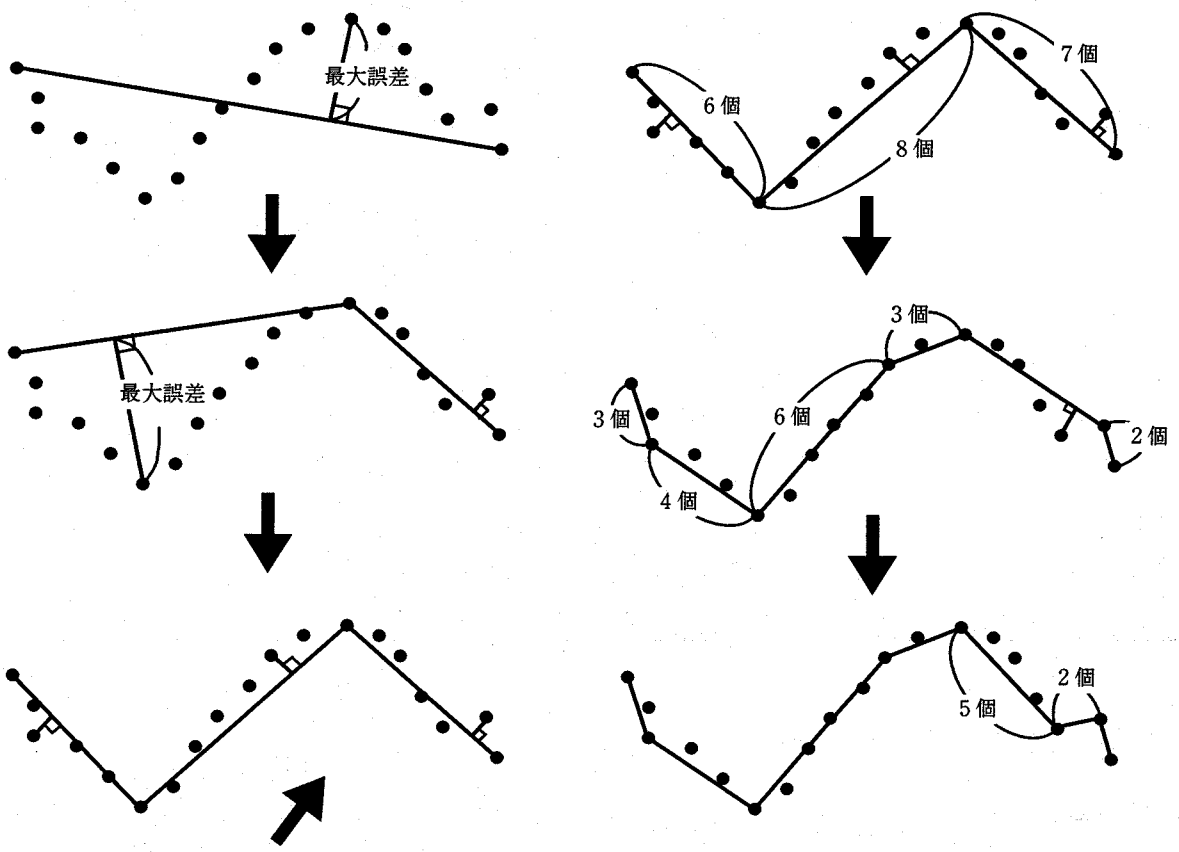


図4-9 区分的直線近似法

2) 小領域とモデルの対応付け

落下運動シミュレーションにより変形したモデルの各特徴点が1つでも重なる小領域をすべて抽出する．これにより，モデルに対応する領域の抽出が終了する．領域抽出処理の様子を図4-10に示す．

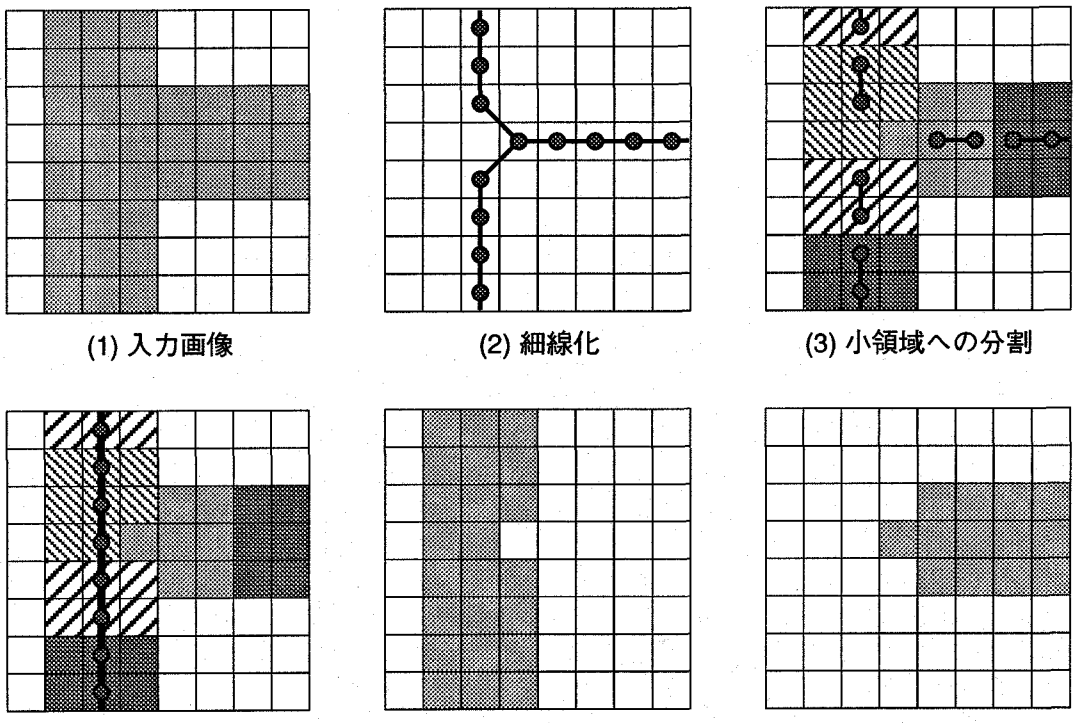


図4-10　領域抽出処理

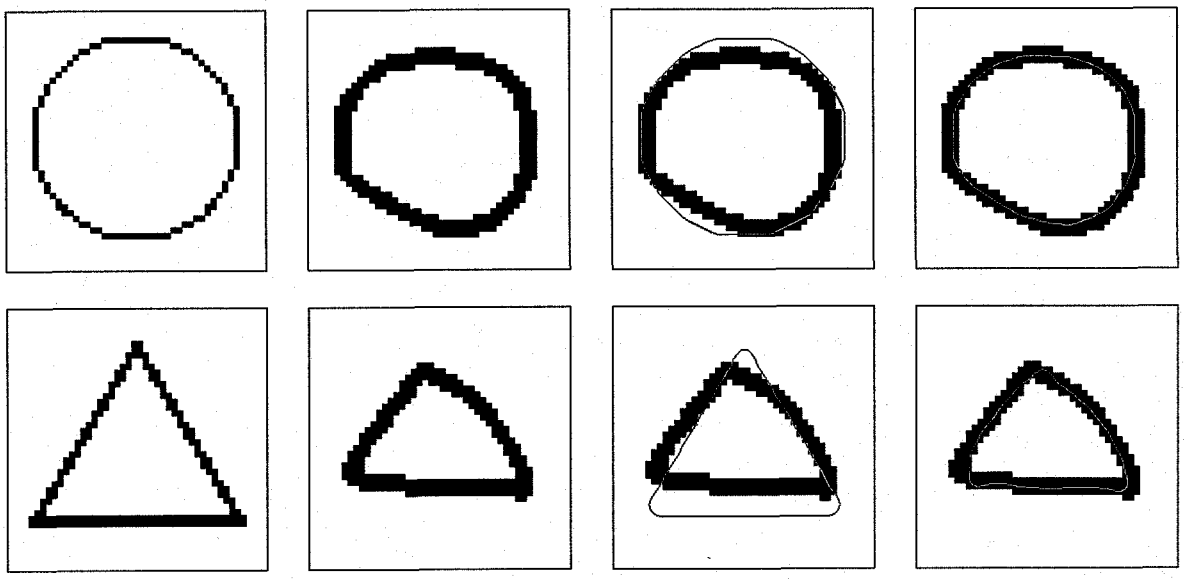


図4-11　歪んだ形状への対応付け

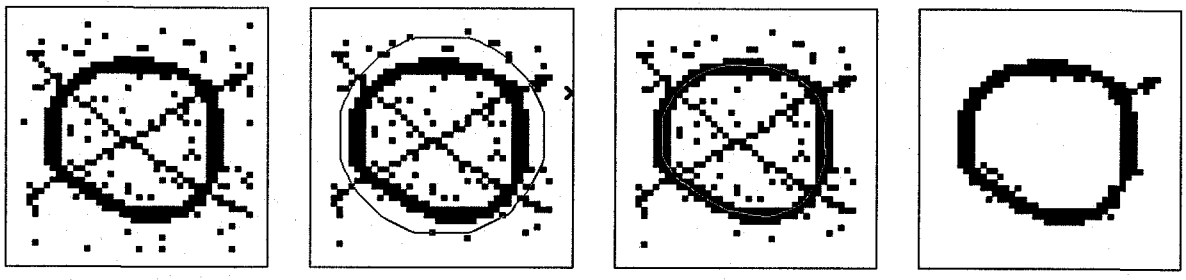
4.7 実験と検討

4.7.1 歪んだ図形への対応付け

まず，落下運動シミュレーションが歪んだ形状の対応付けに有効であることを示すために，歪んだ手書き図形の切り出し実験を行った．図4-11に例を示す．弾性モデルが入力画像の歪みに沿って変形し，良好に対応付けが行えていることがわかる．切り出し結果は1-画素領域がすべて抽出された．

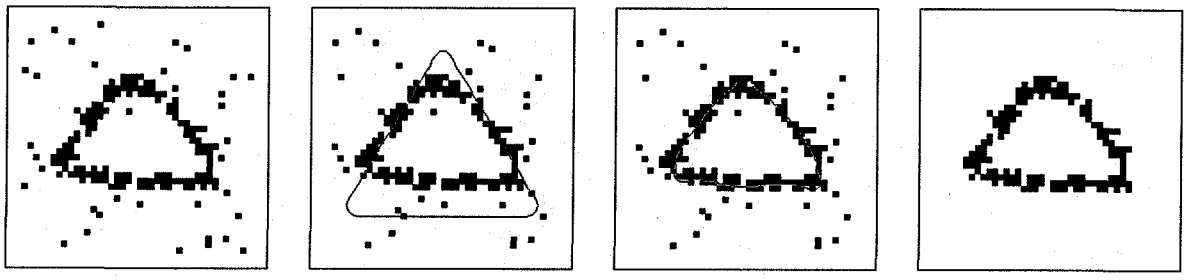
4.7.2 ノイズを含んだ図形の切り出し

図面におけるノイズとしては，余分な成分の混入と形状の途切れやかすれがある．図4-12に余分な成分の混入した場合の例を示す．モデルと入力図形の大きさが多少異なっているにも関わらず，対応付けは良好に行えている．切り出しにおいてはノイズ成分を多少拾っているものの，ほぼ期待通りの切り出しが行えていると言える．図4-13に形状のかすれや途切れを含む場合の例を示す．良好に領域抽出が行えていることがわかる．



(a) 入力画像 (b) 初期状態 (c) 収束状態 (d) 切り出し結果

図4-12 余分な成分の混入した画像からの切り出し（モデルは図4-10と同一）



(a) 入力画像 (b) 初期状態 (c) 収束状態 (d) 切り出し結果

図4-13 かすれや途切れを含む画像からの切り出し（モデルは図4-10と同一）

4.7.3 汎用性に対する考察

ここでは2つの簡単な図形での実験例を示したが，モデルを変えることで異なる形状の切り出しを安定に行えたことがわかる．この手法には調整が必要とされるパラメータが，対応付けにおいてバネ係数，仮想質量，重力係数，減衰係数，サンプリング周期などに7個，領域抽出において小領域への分割パラメータに3個ある．これらは予備実験において調整された後，すべての実験で同一の値を使用した．またモデル毎の調整も必要としなかった．形状の歪み，余

分な成分の混入，かすれや途切れという性質の異なるノイズに対しても全く調整を必要とせず対応できた．これは，この手法が画質に関して広く対応可能であり，画質の点においても汎用的な手法になっているといえる．

4.7.4　面領域への対応

ここで提案した手法は線図形にのみ対応したもので，面領域を持つ形状へは対応できない．そこで図4-14に示すように，輪郭線を抽出し，その画像において対応付けを行うことで，対処することが考えられるが，現時点では十分な検討が行われておらず，今後の課題となっている．

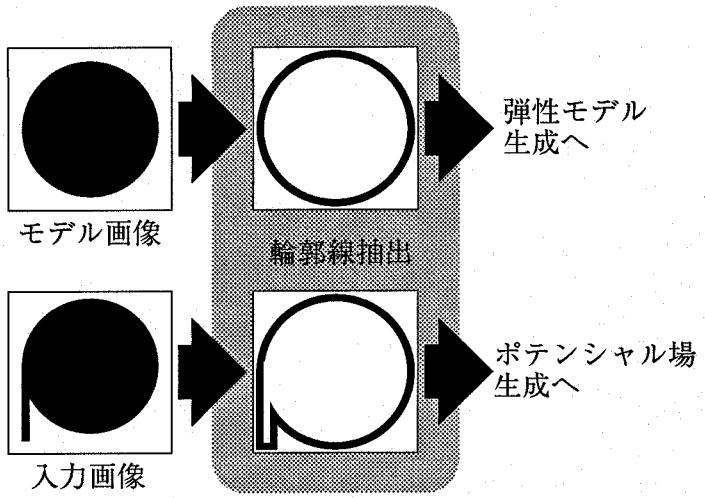


図4-14　面領域への対応手法

4.7.5　初期位置の決定問題

本手法の精度に大きく関与する問題として，弾性体の初期位置の決定問題が上げられる．初期位置があまりにも不適切であれば，誤った対応付けが行われることは予想できる．そもそも切り出し処理は，形状の認識処理の前段階の処理として位置づけられるのが通常であるために，ここで提案する特定の形状のモデルを使用し，さらに位置もほぼ一致させた状態での切り出し処理の有効性は疑問視されるかもしれない．この点に関しては6章で詳しく説明するが，これまでの「切り出し」から「認識」という処理の流れをやめ，「認識」から「切り出し」という順序での処理を考えている．

4.8　まとめ

この章では，弾性モデルによる対応付けに基づく切り出し手法を提案した．この手法は，対象の形状パターンを利用した対応付けを行っているために，切り出しの精度が高い．また，モデルと処理が独立しており汎用的な手法にもなっている．

この手法では，入力画像に投下する弾性モデルの初期位置問題が非常に重要である．これに関しては，この手法の利用目的と密接に関係があり，実際にこの手法を組み込んだ図面解釈手法を提案する6章で詳しく説明する．

ところで，この手法について知識表現の点から見ると，これまでの図面解釈システムにおける切り出し手法が使用していたヒューリスティックな知識は使用されていない．また，記号的に表現された知識も使用されていない．2値画像パターンとして表現されたモデル形状のみが使用されている．「記号処理」と「パターン処理」の融合は重要な問題とされてきたが，効果的に機能するメカニズムは提案されていない．この行き詰まりは知識表現に起因すると思われる．パターン処理においては，切り出し処理が重要な要素となるが，ここで提案するような汎用的かつ知識と独立した処理エンジンを採用した手法はなく，処理に埋め込まれた手続き型知識を用いるのが通常であった．このような処理埋め込み型の手続き型知識は，その開発コストが高く自由かつ容易に登録できるようなものではない．そのため，記号処理部において知識を充実させ高度な処理が可能になっても，パターン処理との情報の受け渡しのための表現力が制限され，十分に性能が発揮できなかった．ここで提案する汎用的切り出し手法は，記号処理との融合が極めて簡潔な形で実現でき，しかもその表現力が高いという性質を持っている．5章では，この話題に関して説明する．

参考文献

[Funakubo91] 船久保登：パターン認識，共立出版，pp.65-67 (1991)

[Kass88] M. Kass, A. Witkin and D.Terzopoulos : SNAKES : Active Contour Models, Int. J. of Computer Vision, 1, 4, pp.321-331 (1988)

[Mukouda83] 向田，鈴木，阿部：協会追跡を利用した図中の文字と図形の分離，信学技報，PRL83-70, pp.35-42 (1983)

[Ohta78] 大田，金出，坂井：領域分割処理によるカラー画像情報の構造化，情報処理，19, 12, pp.1130-1136 (1978)

[Osawa89] 大沢，山川他：図面の認識と理解，昭晃堂，pp.94-95 (1989)

[Sakaguchi90] 坂口，美濃，池田：SNAKEパラメータの設定についての検討，信学技報PRU, 90, 21, pp.43-49 (1990)

[Shizawa88] 志沢雅彦：弾性体モデルを用いた図形のマッチング法，情処研資CV, 53-1, pp.1-8 (1988)

[Tamura85] 田村秀行：コンピュータ画像処理入門，総研出版，p.93 (1985)

[Teramura94] 寺村，中川：力学的な認識モデル「膜モデル」による手書き文字認識実験，信学全大，D-467, 7, p.215 (1994)

[Uchiyama94] 内山，加藤，井口：パターンマッチングによる二値画像中のシンボルの切り出し，信学春大，7, p.203 (1994)

[Webster93] R.G.Webster, M.Nakagawa : The Feasibility of a Parallel Processing Oriented Character Recognition Method Based on a Dynamic Model, Proc. of 2nd ICDAR, pp.714-717 (1993)

[Wada95] 和田，野村，松山：分散協調処理による画像の領域分割法，情処論文誌，36, 4, pp.879-891 (1995)

[Yamamoto82] 山本和彦：弛緩整合法による手書き教育漢字，信学会論文誌，J65-D, 9, pp.1167-1174 (1982)

5章　パターン構造の表現と解析

図面解釈に限らずパターン理解，画像理解においては，記号処理とパターン処理という異質な処理により構成されるが，これらを融合するためには多くの特化された手続きのプログラミングが必要とされた．汎用的な図面解釈システムを構築する上では，これは大きな問題となっている．4章で提案した汎用的切り出し手法は，記号処理との融合が極めて簡潔な形で実現でき，しかもその表現力が高いという性質を持っている．この章ではこの点に関して説明する．

5.1　トップダウン型パターン処理の問題点

トップダウン的な認識・解釈システムにおいては，後ろ向き推論方式での検証処理が通常採用される．まず解釈過程において，それまでに採用された仮説と知識から検証すべき事項が決定される．この処理は，記号表現された知識を用いた記号処理として実行される．次にその検証事項に対する実際の検証処理が実行される．図面解釈ばかりでなくパターンを対象とした問題においては，この検証処理は記号処理の範疇に留まらず，パターン処理の問題となる．そこではパターンの属性が調べられるわけだが，その属性を表現する記号と処理が予め関係づけられておく必要がある．

3章のシステムでは，このようなパターン処理と直接関係づけられた記号を一般的語彙と呼び，それらを用いてパターン形状を表現していた．そこでは，一般的語彙の量と質が表現力を左右する非常に重要な問題であった．汎用的システムを構築する場合，多様なパターン形状を表現できるだけの一般的語彙を準備しておかなければならない．この一般的語彙はパターンを表現するための概念の集合として捕らえることができるが，それらは大きく2種類に分類できる．1つはパターン自体の性質を示す概念であり，例えば「円」「直線」「点」「面」というようなものである．もう1つは，それらの関係や属性を表現する概念であり，「距離」「角度」などである．

パターン構造を記号で表現する場合，これら概念を組み合わせることにより宣言的に表現できる [Kawagoe87]．しかしそれだけでは実際の検証処理は機能しない．例えば，「三角形」の仮説が存在したとする．次にその三角形が「正三角形」かどうかの検証が必要になったとする．記号表現された知識を用いて，「3つの角度が60度であることを確かめなさい．」とか「3つの辺の長さが等しいことを確かめなさい．」という検証事項を発見するのは容易であろう．では，次にどのようにすれば，「角度」や「辺の長さ」の値を調べることができるであろうか．もし「三角形」仮説の頂点座標を示す属性値がすでに埋まっていれば簡単である．これだと記号処理の範囲で検証処理が終了する．しかしその属性値が空の場合，実際にパターンからその属性値を見つけてこなければならない．この手続きにはパターン処理が必要となり，「三角形の頂点を検出するパターン処理」を準備しておかなければならない．このことは，「正方形」の場合でも同様であり，「四角形の頂点を検出するパターン処理」を準備しておかなければならない．汎用化するためには「多角形の頂点を検出するパターン処理」を準備しておけばよい．

ところで，このようなパターン処理をどれだけ準備しておけば足りるのであろうか．実際に

は，システムの取り扱う形状を制限し，その中での形状表現に必要なパターン処理を準備するのが効率がよい．しかし，このような方法では汎用化は望めない．

つまり記号処理部での表現能力を高めるためには，その表現に対応したパターン処理の手続きが準備されていることが要求される．つまり記号表現において直接パターンを参照する必要がある概念を１つ追加する度に，その概念に対応したパターン処理を準備しなければならず，この処理の準備のコストが表現力の限界となる．

5.2 パターン構造の表現と解析

４章では，２値画像配列としてパターン表現された知識と，弾性モデルに基づく汎用切り出しエンジンによるパターンの切り出し手法について述べた．この手法は単に切り出し処理の開発コストを抑えるばかりでなく，5.1で述べた問題点であるパターン構造の記号表現における表現能力を高めるための汎用的パターン処理を担うこともできる [Kumagai95].

5.2.1 パターン表現を用いた構造記述

「三角形」を例に説明する．図5-1に意味ネットワーク表現での表現例を示す．「三角形」は，($x $y $z) の３つの座標値を引数として表現でき，「多角形」に属し，３つの「辺」を持つという簡単な表現である．これを基にパターンの構造解析を行うために必要な手続きを考えてみる．

```
(triangle   argument:   ($x $y $z)
            isa:        polygon
            haspart:    (side $x $y);
                        (side $y $z);
                        (side $z $x))
```

図5-1 「三角形」の意味ネットワーク表現

まず，座標値がすでに決まっていた場合，２つの座標値 ($x $y) で表現される線分がパターンとして存在するかどうかを検証する手続き verify_side ($x $y) を準備しておく必要がある．次に１つの座標値 ($x) が未知数の場合にそれを埋めるための手続き find_vertex ($y $z) も必要となるであろう．さらに２つの座標値($x $y)が未知数の場合にそれを埋めるための手続き，３つ座標値とも未知数の場合にそれを埋めるための手続きと必要になる．さらに「正三角形」を取り扱うような場合には辺の長さを調べる手続き，「直角三角形を」取り扱う場合には角度を調べる手続きなどが必要となる．システムの取り扱うすべての形状に対しこのような手続きをすべて準備するのは非常に難しいが，これらを次の３種類に分類して考えてみる．

1) 形状評価

形状を表現する概念と実際のパターン形状が対応していることを調べる手続きのことである．例えば，side ($x $y) は，($x $y) という２つの座標値で表される「辺」を表現する概念であり，

実際にその概念に対応するパターンが存在することを検証する手続きが必要である.「円」や「四角形」などの形状を表現する概念に直接付随する手続きである.

2) 位置検出

次に説明する幾何学的属性評価のための位置を指定する手続きのことである. 例えば,「三角形」を表す座標 ($x $y $z) の実際の座標値を見つける手続きである. 複雑な形状を表現しようとする場合, この手続きの構成が最も難しい問題となる.

3) 幾何属性評価

幾何学的属性を評価する手続きのことである.「距離」「角度」「直交」などの概念に対応する.

「形状評価」に関しては, 形状を表現する概念に対して1つの手続きが考えられる. しかし, 例えば「四角形」などの形状概念は, その部分構造を表す「辺」という形状概念を用いて表現することもできるので, それらに評価を引き渡すという手法を採用することにより「四角形」を直接評価する手続きは必ずしも必要ではない.「形状評価」の手続きは最も基本的な形状概念に対してのみ存在すればよい. しかし, 処理の精度や高速化の点からはすべての形状概念に対しての評価が行えることが望ましい.

「位置検出」の手続きは, 形状の部分構造を表現するのに必要となる. この手続きは一般化が困難で, 複雑な形状ほど多くの数が必要とされ, その処理も複雑になる. 例えば,「トランジスタ」の部分構造表現を例にすると, その部分構造である「矢印」の先端を指定する記号表現やその位置をパターンから検出する手続きを考えれば, その困難さが想像できる.

「幾何属性評価」に関しては, パターン処理とは切り放して考えることもでき, その種類も少なく, 構成も容易である.

つまり,「形状評価」「位置検出」の手続きに関して, 汎用的な手続きを利用することができれば, 形状の表現力, 解析力の向上が期待できる. これには4章で提案した切り出し手法が利用できる. この切り出し手法では, 形状モデルとして2値画像配列を用いた. その配列で表現されたパターンが入力画像に対応づけられる.「形状評価」に関してはこのときのパターンの歪みの度合いを評価する関数を利用することができる. また, 形状モデルを2値の画像配列として表現するのではなく, 構成要素となる部分形状毎にラベル付けをしたラベル画像配列として表現し, そのラベル毎に領域の抽出を行うことで, 部分構造の抽出とその「位置検出」が可能となる. 図5-2に構造記述の例を示す. 弾性モデルによる切り出し処理では切り出しと同時に,

1) 弾性モデル全体の歪み率の計算:　warp_all
2) 各ラベル毎の歪み率の計算:　warp ($label)
3) 各ラベル領域の座標値の検出:　coord ($label)

を行う. ラベル x, y, z は三角形の「辺」に対応する. ラベル a, b, c は「頂点」に対応する. 記号表現のみで形状を表現しパターンの検証処理に利用する場合, 図5-1で示した記述の他に, 部分構造の関係記述も必要となる. しかしこのモデルでは, パターン表現の中に部分構造の関係がすべて含まれていると考えることができる. このようなパターンモデルを用いた切り出し

```
(triangle  argument:  ($x $y $z $a $b $c)
           haspart:   (side $x);
                      (side $y);
                      (side $z);
                      (virtex $a);
                      (virtex $b);
                      (virtex $c))
```

							a	a							
							a	a							
						x			z						
						x			z						
					x	x			z	z					
					x					z					
				x	x					z	z				
			x	x							z	z			
			x									z			
		x	x									z	z		
	b	b	y	y	y	y	y	y	y	y	y	y	c	c	
	b	b	y	y	y	y	y	y	y	y	y	y	c	c	

図5-2　記号表現とパターン表現の併用

を行うことで，部分構造の形状評価と位置検出が完了する．幾何属性評価においてもこの処理で抽出された情報を利用すれば，新たにパターン処理手続きを考える必要はない．複雑な形状を表現する場合でも，そのために特別なパターン処理のための手続きを準備する必要はない．また部分構造の関係記述も不要である．

5.2.2　切り出し処理の改良

1) 歪み率の算出

弾性モデルの歪みが切り出し領域のモデルとの類似を表現することになる．まず，弾性モデルの対応付けの結果から次の２つの評価値を求める．

a. バネに蓄えられているエネルギー

図4-5に示す３種類のバネのうち，形状の変形に影響されるバネ２，３に蓄えられているエネルギーを計算し，同一ラベルのすべての特徴点に接続されるバネに関して平均値を求める．この値を実験的に調整した関数を用いて０～１の間の実数値に正規化する．

b. 領域対応の成功率

特徴点が実際に入力画像の1-画素の領域に対応づけられているかを調べる．同一ラベルのすべての特徴点のうち，入力画像の1-画素の領域に対応づけられている割合を計算する．

この２つの値の積を歪み率とした．

2) ラベル領域の抽出

4.6の処理をラベル毎の領域が抽出できるように改良する．この処理の概念図を図5-3に示す．単純な方法としては，4.6,2) での処理において特徴点に対応する領域の抽出において，各小領域に対応する特徴点のラベルを割り振ればよい．しかし，これでは小領域に単位での精度でしか分割が行えない．そこで，一旦従来方法で全領域を抽出し，特徴点と重なる画素にのみ

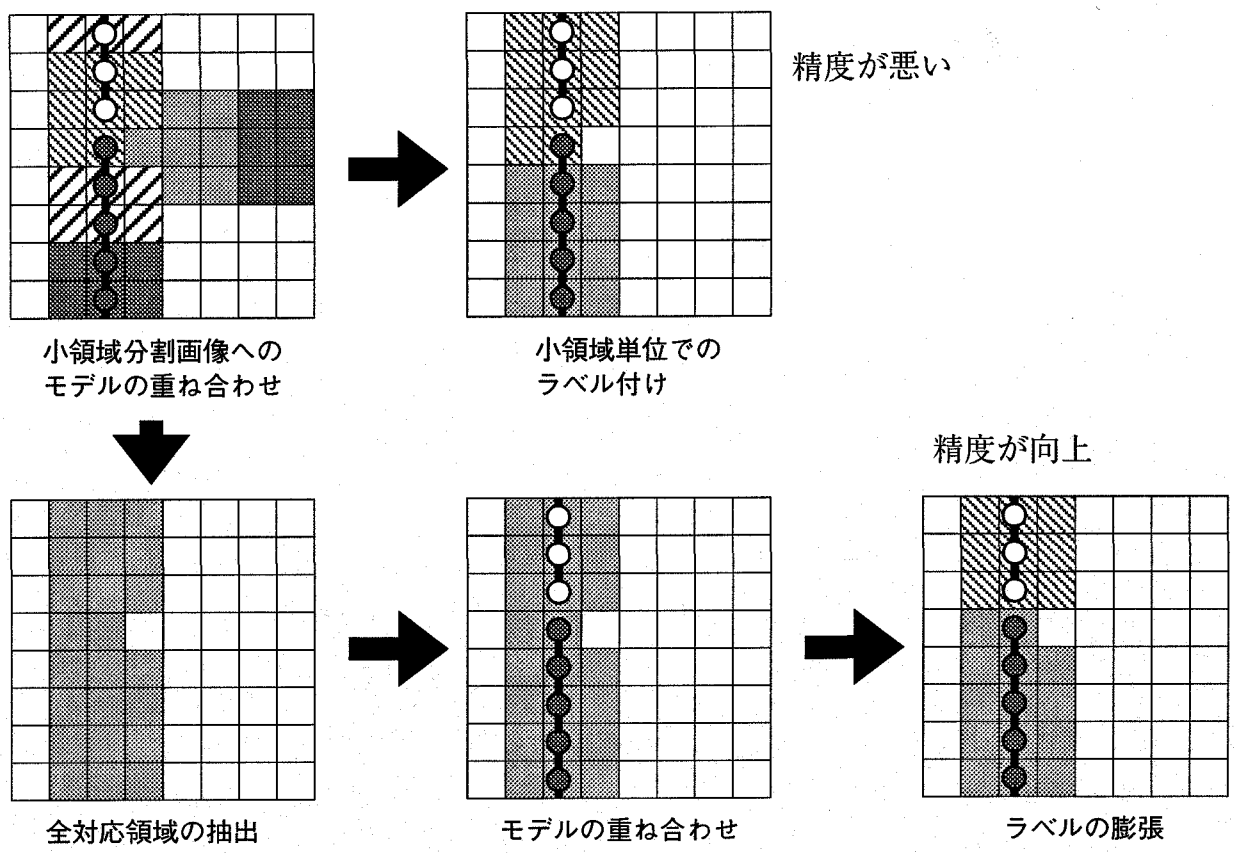


図 5-3　ラベル領域の抽出

```
(circle   argument:   ($a $b $c)
          haspart:    (shape $a);
                      (virtual radius $b);
                      (virtual radius $c);
                      (virtual radius $d);
                      (virtual radius $e);
                      (virtual center $f))
```

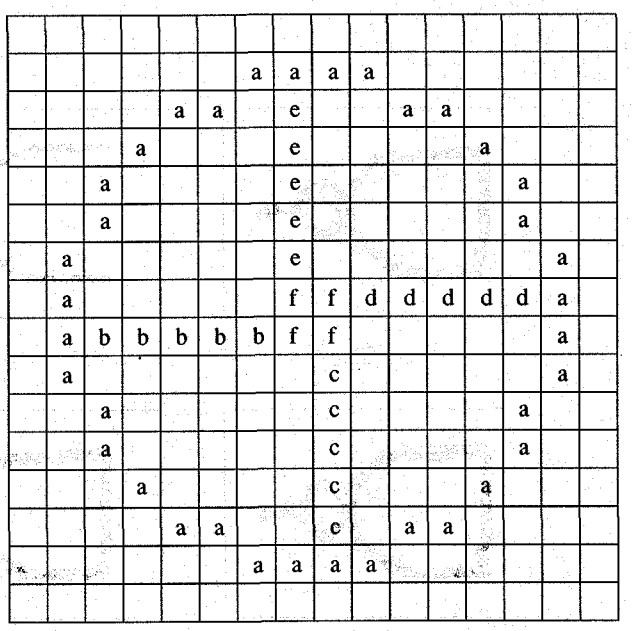


図 5-4　仮想領域を用いた形状記述

ラベルを付け，それを抽出された領域内で膨張させることでより精度の高い分割を行った．

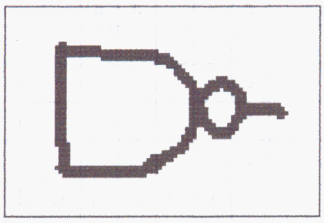
3) 仮想領域の導入

形状の記号記述を行うときに，パターン形状に存在しない特徴点を利用したい場合がある．例えば，円の中心座標を用いようとした場合，そこには1-画素領域が存在しないために，ラベルを割り振ることができない．そこで仮想領域という概念を導入する．これは特徴点を表現するための領域であり，実際のパターンは存在しない．図5-4にその例を示す．ラベルb,c,d,e,fを仮想領域として定義している．切り出し処理において，仮想領域の特徴点では重力質量を0として計算する（慣性質量は通常通りに取り扱う）．また，図4-5のバネのうちバネ1を設定しない．これにより仮想領域の特徴点の効果は全体形状を維持することにのみ表れる．

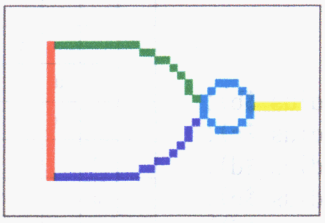
5.3　実験

5.3.1　部分構造の切り出し実験

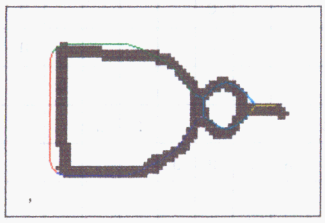
弾性モデルを用いた切り出し手法の改良で，部分構造の切り出しが可能であることを確かめる．論理回路図で使用されるNADNゲートの例を図5-5に示す．(a) は入力画像，(b) はモデルとなるラベル画像である．5つの部分にラベル分けされている．(c) は弾性モデルの初期位置，(d) は収束後の状態である．期待通りの収束結果が得られているのがわかる．(e) はこれをもとに部分構造に分割を表すラベル画像である．(b) のモデル画像と比較して良好に分割が行われているのがわかる．同一のモデルを画質の劣化した画像に適応した例を図5-6に示す．かなり画質が劣化しても弾性モデルの対応付けは良好に行えているのがわかる．しかし，部分領域への分割においては，ノイズを抽出したり，必要部分を取りこぼしているのがわかる．これは4.6で説明した小領域への分割処理が細線化処理に基づいているが，ノイズが混入したり，途切れが発生することで，この細線化画像が非常に複雑な形状となったのが原因であった．全体的には，良好な部分構造の切り出し処理が行えている．



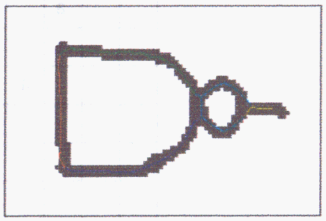
(a) 入力画像



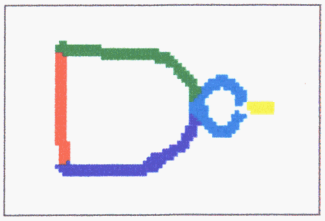
(b) モデル画像



(c) 弾性モデルの初期位置

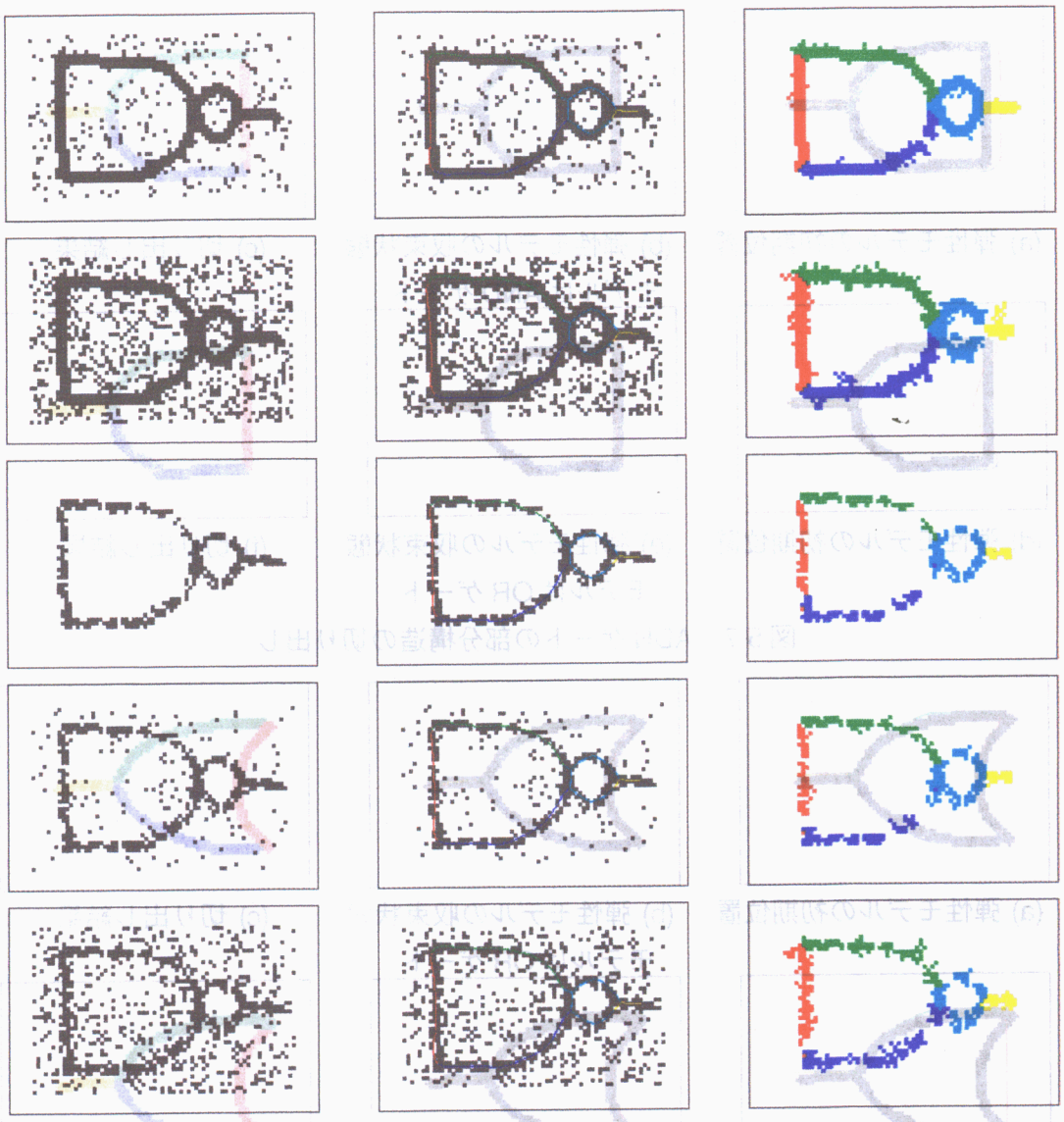


(d) 弾性モデルの収束状態



(e) 切り出し結果

図5-5　NADNゲートの部分構造の切り出し



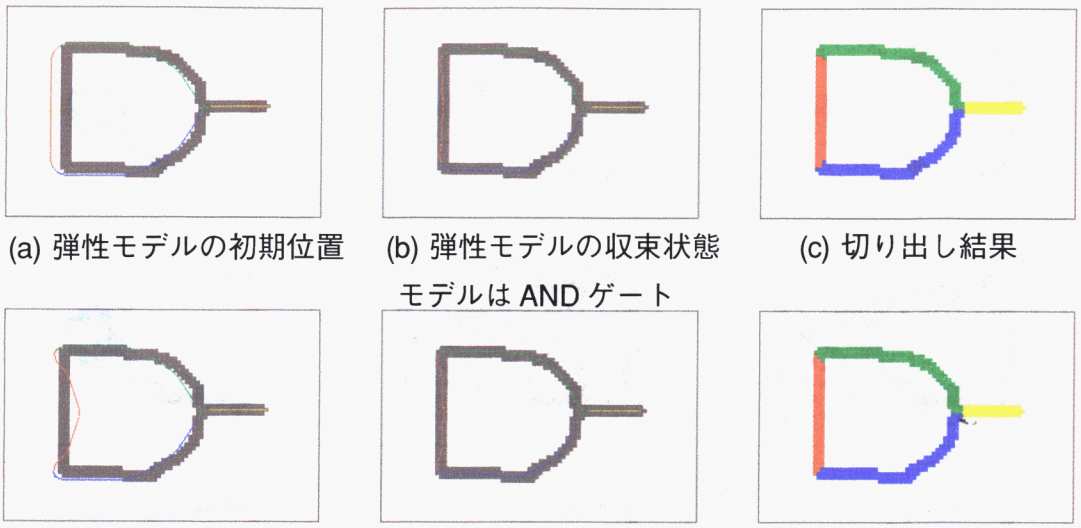
(a) 入力画像　(b) 弾性モデルの対応結果　(c) 切り出し結果

図5-6　画質の劣る画像に対する切り出し実験

5.3.2　部分構造の解析

提案手法が部分領域の構造解析にも有効であることを示すために，形状の類似したシンボルであるANDゲートとORゲートを例に，歪み率の変化を観察した．図5-7は，入力図面としてANDゲートを与え，モデルとしてANDゲート,ORゲートの両者に対して切り出しを行った例である．また，図5-8は，同様の実験をORゲートの入力図面に対して行った例である．どの結果をみても，弾性モデルが柔軟に変形し，どのモデルを用いても同一の切り出し結果となった．このときの各ラベルの歪み率の値を表5-1に示す．ラベル1がゲートの入力側に当たる辺に対応する．この値を見ると，モデルと入力画像の形状の類似性が明確に反映されていることがわかる．

このことは，この手法を用いることにより，部分構造の類似性の検証のために特別な手続きのプログラミングを必要とせず，記号表現の中に直接的に部分形状の類似性を表現できることを表している．



(a) 弾性モデルの初期位置　(b) 弾性モデルの収束状態　(c) 切り出し結果

モデルは AND ゲート

(d) 弾性モデルの初期位置　(e) 弾性モデルの収束状態　(f) 切り出し結果

モデルは OR ゲート

図 5-7　ADN ゲートの部分構造の切り出し

(a) 弾性モデルの初期位置　(b) 弾性モデルの収束状態　(c) 切り出し結果

モデルは OR ゲート

(d) 弾性モデルの初期位置　(e) 弾性モデルの収束状態　(f) 切り出し結果

モデルは AND ゲート

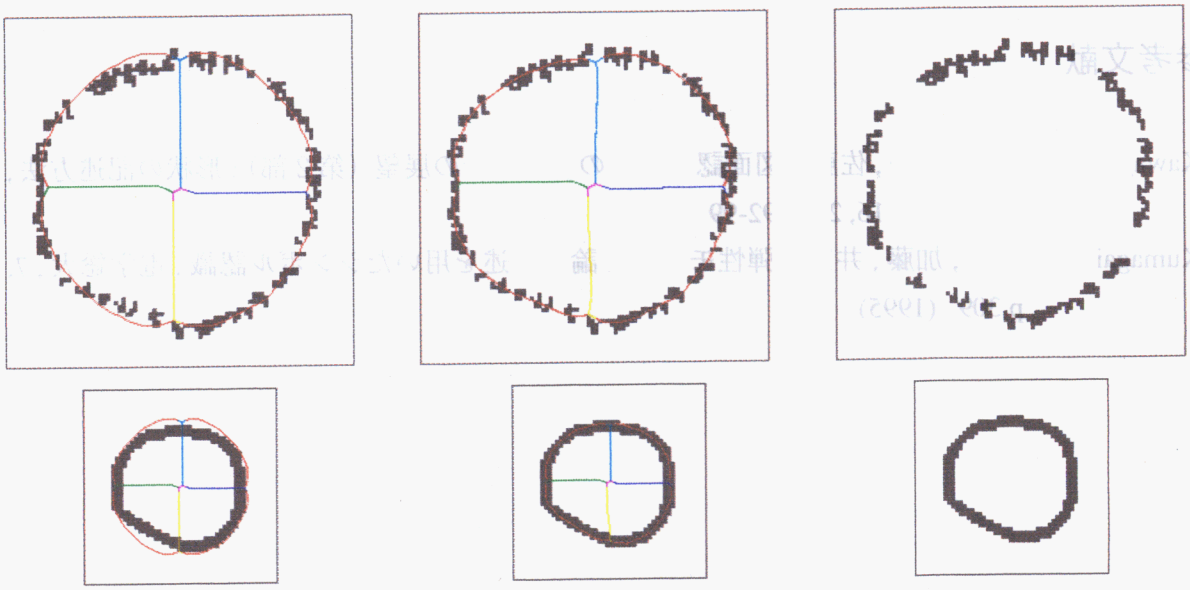
図 5-8　OR ゲートの部分構造の切り出し

表 5-1　図 5-7,8 の切り出しにおける各ラベル領域の歪み率

	Model	AND		OR	
	Input	AND	OR	OR	AND
Label No.	1（赤）	0.748097	0.295189	0.898678	0.273154
	2（緑）	0.699640	0.701070	0.798606	0.704445
	3（青）	0.716694	0.657335	0.770549	0.601059
	4（黄）	0.735775	0.736922	0.749632	0.745681

5.3.3　仮想領域の効果

仮想領域の有効性を調べるために，図5-4のモデルを用いて円の切り出し実験を行った．切り出しの際に，モデル画像から生成された弾性モデルの特徴点について，各ラベル毎にその重心位置，最大矩形領域を求める処理を行った．これにより，弾性モデルの対応付け終了後に，モデル画像の各ラベルの存在領域が検出できる．図5-4の場合，ラベルfの重心位置が入力画像の円の中心座標，ラベルb,dの横軸方向の領域幅とラベルc,eの縦軸方向の領域幅が半径を示す．大きさの異なる2つの入力画像に対して切り出しを行った例を図5-9に示す．その時の重心位置，最大矩形領域を表5-2に示す．対応付けが良好に行えているために，これらの値の信頼度は高い．つまりこのことは，各ラベルの割り当てられた部分形状がどんな形状であれ，その位置情報の検出手続きのプログラミングを必要としないために，記号表現における表現力の向上に貢献するものである．



(a) 弾性モデルの初期位置　(b) 弾性モデルの収束状態　(c) 切り出し結果

図5-9　仮想領域を持つモデルによる形状の切り出し

表5-2　図5-9の切り出しにより得られた各部分領域の位置情報

		x_min	y_min	x_max	y_max	center(x,y)
かすれた円（直径約78画素）	a（赤）	(3.24,	5.12)-(	74.85,	74.30),	(39.27, 39.45)
	b（緑）	(4.80,	38.00)-(	35.90,	38.77),	(20.61, 38.09)
	c（黄）	(37.93,	42.10)-(	38.64,	72.31),	(38.30, 57.00)
	d（青）	(41.96,	39.45)-(	72.94,	40.10),	(57.22, 39.79)
	e（水）	(39.93,	7.18)-(	40.72,	36.17),	(40.32, 21.77)
	f（紫）	(37.92,	38.62)-(	39.93,	39.62),	(38.93, 39.12)
歪んだ円（直径約38画素）	a（赤）	(2.87,	4.41)-(	35.55,	35.53),	(19.66, 19.70)
	b（緑）	(3.54,	18.96)-(	18.02,	19.32),	(10.87, 19.00)
	c（黄）	(18.90,	20.92)-(	19.71,	34.57),	(19.26, 27.73)
	d（青）	(20.86,	19.74)-(	34.88,	20.09),	(27.84, 20.04)
	e（水）	(19.85,	4.95)-(	20.03,	18.15),	(19.95, 11.52)
	f（紫）	(18.96,	19.32)-(	19.92,	19.75),	(19.44, 19.53)

5.4 まとめ

この章では，弾性モデルに基づく切り出し処理が，パターンの部分構造の表現や解析のために，有効な情報を供給する役割を果たすことについて述べた．従来からの手法では，記号によるパターンの特徴表現のためにパターン処理の手続きプログラミングが必要であったが，この困難さのためにパターン構造の柔軟な記号表現が困難であった．ここで提案した手法は，パターンの形状に依存せずに部分形状の特徴抽出が行える汎用的なメカニズムを与えるものであり，この問題を解決するものである．

次の章では，このような汎用的解決手法が利用できることにより可能となった，動的な認識プログラムの生成とそれに基づく柔軟な図面認識・解釈のための方法を提案する．

参考文献

[Kawagoe87] 川越，喜多，佐藤：図面認識技術の現状とその展望（第2部）：形状の記述方法，画電学誌，16, 2, pp.92-99 (1987)

[Kumagai95] 熊谷，加藤，井口：弾性モデルと論理記述を用いたシンボル認識，電学総大，7, p.309 (1995)

6章　認識戦略の自動生成に基づく図面解釈

3章では知識処理を導入することで,処理と知識とを分離し汎用的な図面解釈システムの構築を試みた．しかし，パターン処理系における処理と知識との分離に問題があり，期待すべき成果が得られなかった．4，5章では，この問題点を解決するための要素技術として，弾性モデルに基づく切り出し手法の提案を行った．ここでは，この要素技術を用いて再度汎用的図面解釈のための方法論について検討を行う．記号的知識による認識戦略の生成と、パターン的知識による認識処理の実行という、2種類の知識を有効に利用した汎用的なトップダウン図面解釈の手法を提案する。

6.1　汎用的図面解釈システムにおけるパターン処理

図面解釈をプログラムの設計の観点から検討してみる．図6-1は2章で示した楽譜認識のシステムで採用された設計方針である．プログラムの設計者が実際の図面解釈の実行プランを設計し，それをそのまま計算機に手続きプログラムとして埋め込む方式である．この場合，対象を変更するためにはプログラムの設計者が実行プランを組み立て直さなければならず,汎用性がない．

図6-2は3章で示した論理回路図の認識システムで採用した方式である．知識処理に基づき記号処理系において処理と知識の分割を試みた．これにより，記号処理系の汎用化が可能になったことに加え，記号処理自身の能力も向上した．しかし新たな問題点が持ち上がった．これを図6-3を用いて説明する．前向き推論型処理においては，パターン処理により生成された仮説に対し，記号処理により多様な解析や矛盾検証が可能であり，パターン処理では解決できない曖昧さを取り除くことができる．しかし，多様な解析処理の可能性を裏返すと，パターン処理の生成する仮説の数が増加することで探索における組み合わせの爆発を招くことになる．そこで，純粋な前向き推論型処理だけではなく，後ろ向き推論型の処理を組み合わせることにより，探索空間の削減を図ることが考えられる．つまり，記号処理部において仮説の生成を行

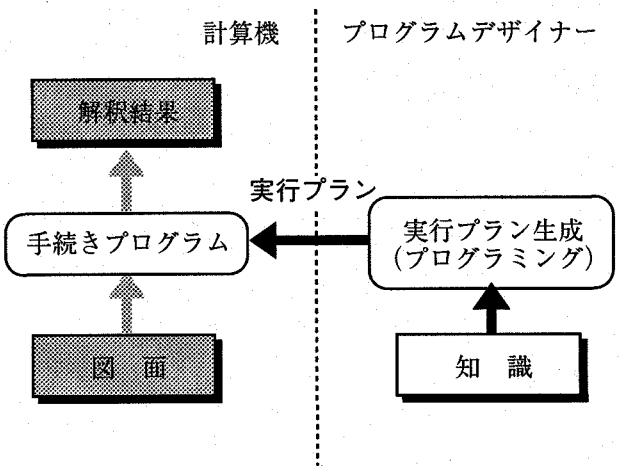


図6-1　知識埋め込み型図面解釈システム

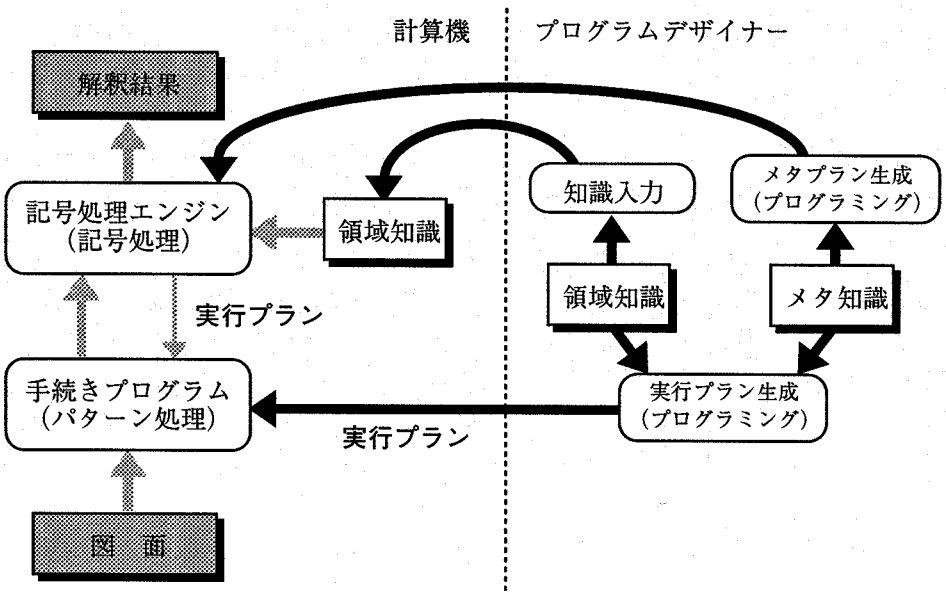


図6-2　知識処理を導入した図面解釈システム

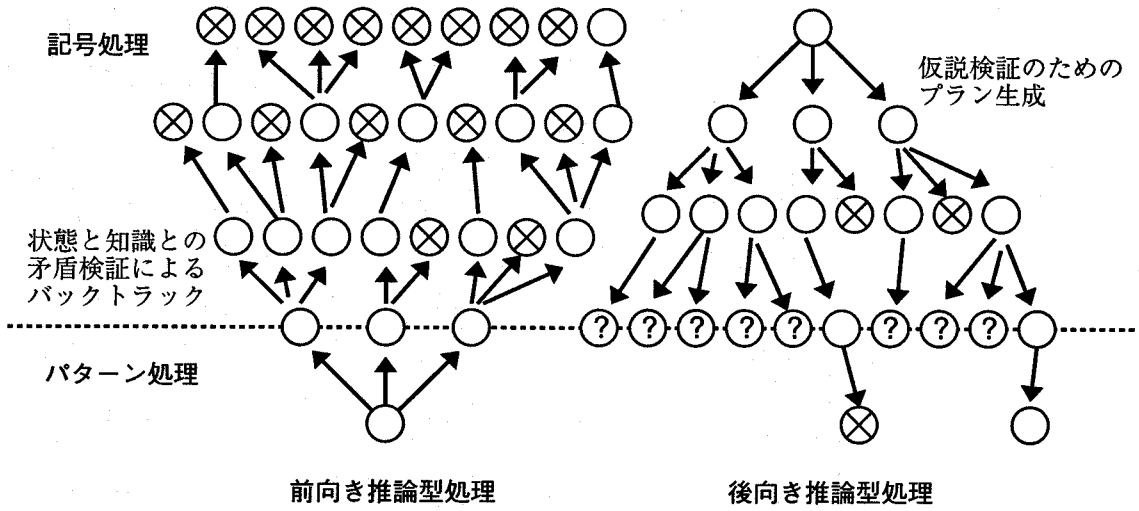


図6-3　記号処理部の役割

い，その妥当性を実際の入力画像で検証するという方式である．このとき，複雑な表現や推論が可能な記号処理系では検証のために多くのプランを生成することが可能である．しかし，そのプランを実行する能力がパターン処理部に備わっていなければ，実際の検証を行うことは出いない．図6-2でわかるように，パターン処理部は手続きプログラムが埋め込まれているために，その処理の柔軟性は低い．つまり，後ろ向き推論型の処理においては記号処理の優れた能力を生かすことができない．

このような問題点を解決するには，パターン処理部においても処理と知識の分離，処理の柔軟性の向上を目指す必要がある．図6-4はこのような処理方式を示したものである．パターン処理系においても処理と知識の分離を行い，その実行プランは記号処理部からの動的な要求に応えられるような柔軟なものである必要がある．4，5章で提案した弾性モデルを用いた汎用切り出し手法は，この要求を満たすものである．しかし，弾性モデルによる切り出し手法は，

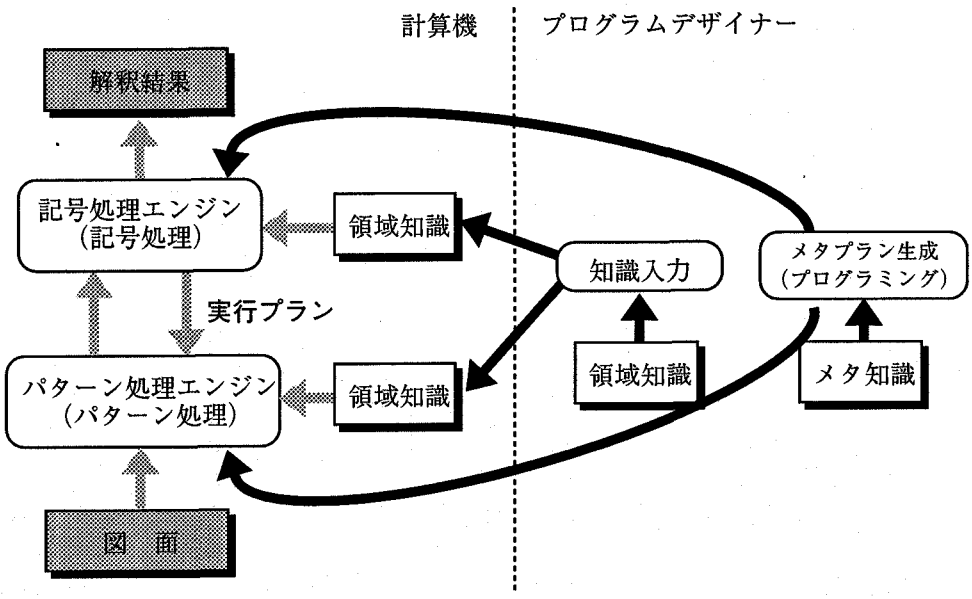


図6-4　汎用図面解釈システム

おおよその位置情報が既知でない限り適用できないので,後ろ向き推論型処理では利用できるが，前向き推論型の処理での利用は困難である．そこで，別途前向き推論型処理に適したパターン処理を検討しなければならない．

汎用図面解釈システムの構築において，さらに検討すべき点として，

1) 前向き推論型のパターン処理
2) 後ろ向き推論による実行プランの生成（認識戦略の自動生成）

がある．6.2では図面解釈の処理の全体像に関して，6.3では1.に関して，6.4，6.5で2.に関して説明する．

6.2　注意制御に基づく認識処理

6.2.1　人間のパターン認識

6.2.1.1　受動的認識処理

目に視覚刺激を与えると脳の中の神経細胞である視覚ニューロンはいろいろな反応を示す．図6-5のようにそれぞれの視覚ニューロンは網膜の特定の位置「受容野」に対応していて，そこに入ってくる光のパターンにより様々に反応する．それら視覚ニューロンは図6-6に示す脳の中の視覚領と呼ばれる領域に存在する．ネコやサルの視覚領には，特定方向の傾きを持った線分や輪郭など，図形の部分的な特徴に反応する神経細胞が見つかってい

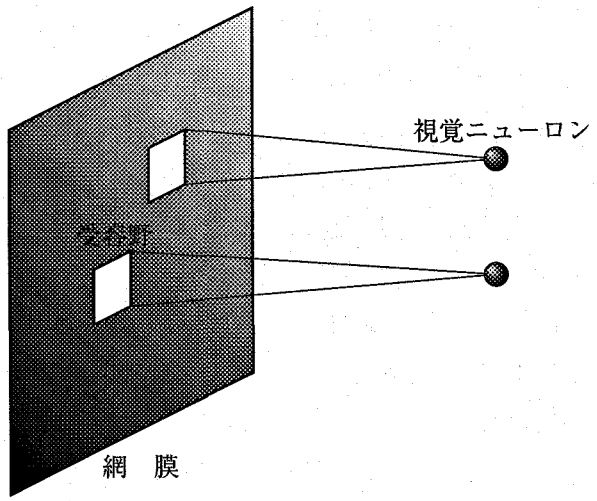


図6-5　視覚ニューロンと受容野

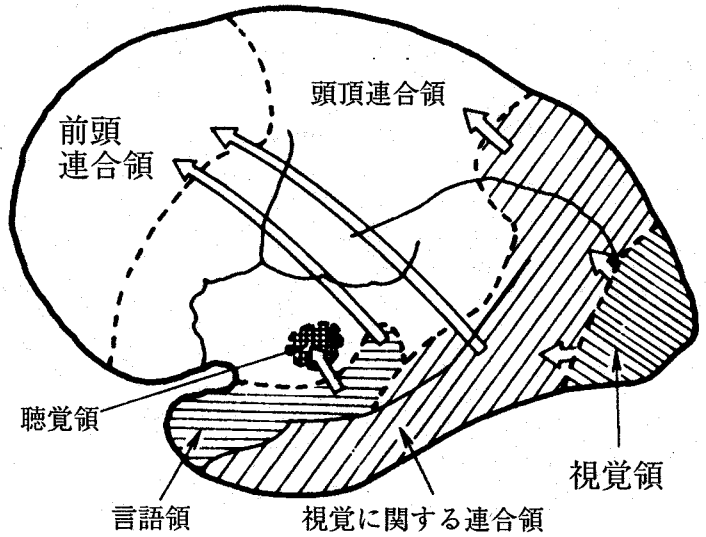


図6-6　脳の知覚処理領域　[Shirai87] より

る．視覚領よりもさらに上位の中枢には，丸や三角といった図形や，人間の顔のような複雑なパターンに選択的に反応する細胞の存在も確認されている．最終的にはこれらの情報は前頭連合領で意識下に処理されることになる．生理学で得られたこのような事実から，視覚神経系は，簡単なものから複雑なものへ，情報が処理され抽象化されていくような階層的な構造を持つものと考えられる [Fukushima89].

このような視覚ニューロンによる受動的な認識処理は，極めてパターン処理的に低次の視覚細胞レベルで行われる無意識的な処理である．この処理に関わる細胞は膨大な数に上り，それらの細胞が様々な認識処理を分担し並列的に実行している．

6.2.1.2　注意制御による能動的認識処理

人間の認識処理は受動的な処理だけではなく，能動的な処理もある．その能動的な情報処理機能の1つに「注意機能」がある．注意にはいくつかの種類があるが，その中に“何かある特徴や部分に意識を集中する”という注意があり，こういう注意を「集中的注意 (focal attention)」という．この注意の対象としては色や形状といったある属性を持つものに対して行われる．それらの属性としては前述の低次の受動的認識処理によって得られるものを利用することになる [Inui93]．この能動的な認識処理は記号処理的な側面を多分に持つ処理である．

たとえば，人混みの中から水色の服を着た人を探す，といったタスクを考える．このとき人間は，受動的な処理により抽出される色の情報から水色の特徴を発見する．そして発見すれば，そこに注意を向けて確かにそこに水色の服を着た人がいるかどうかを確かめる処理をすることになる．この後者の処理が「集中的注意」である．

一般に集中的注意の要求がおこる対象というのは1つではなく数多く存在する．ところが意識的な思考を伴う能動的な認識処理では思考そのものがもつ逐次性により，並列性をもつ受動的な認識処理とは異なり複数の対象について並列的に処理することができない．そういった複数の注意の要求に対して焦点をあてる注意を効果的に選択する制御を行う働きが「注意制御」である．

このような受動的な認識処理と記号処理的な側面も持つ能動的な認識を融合させた処理機構が，人間の持つ柔軟かつ優秀なパターン認識能力を支えていると思われる．

6.2.2　図面解釈における注意制御

今述べてきたような人間の注意制御機構に基づき，図6-7に示す図面解釈機構を提案する．記号処理部とパターン処理部の間の情報の流れで構成をまとめている．パターン処理部における仮説生成は前向き推論型処理として実行され，その結果は注意仮説として記号処理部に渡される．これは，人間の受動的認識処理に対応する．この注意仮説は一般に精度が悪く，量も多いので，直接記号処理部の前向き推論処理に渡すと組み合わせ爆発を起こすために，必ず仮説検証処理が行われる．記号処理部において仮説検証のためのプランが生成され，その検証プランがパターン処理部へ渡される．ここで実際にプランに沿った検証処理が実行され，検証に成功した仮説が図面仮説に追加される．この図面仮説に対し矛盾検証や意味解析が行われ，その結果が解釈結果として出力される．意味解析により新たな仮説生成が期待される場合は，それを直接図面仮説に登録するのではなく，一旦注意仮説として検証処理を通過させた上で図面仮説に登録する．

この図面解釈機構は，4つの処理部分から構成されるのがわかる．パターン処理部における仮説生成処理に関しては6.3で述べる．仮説検証処理は，4，5章で述べた弾性モデルによる切り出し処理が利用できる．記号処理部における仮説検証プラン生成処理とともに6.4, 6.5で述べる．矛盾検証・意味解析処理に関しては，3章で述べた処理の記号処理部を利用したのでその説明は3章を参照していただきたい．

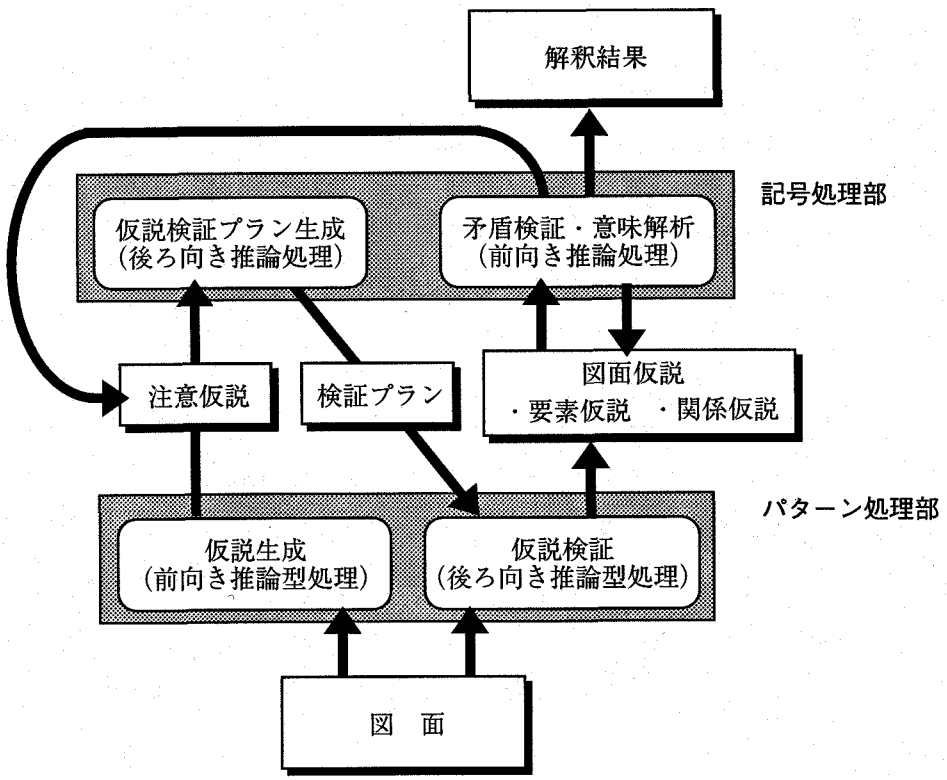


図6-7　注意制御機構を模倣した図面解釈システム

6.3 前向き推論型パターン処理による注意仮説の生成

6.3.1 注意仮説

注意仮説はパターン処理部における前向き推論型処理で生成される仮説である.これは次のような情報を持つ.

1) 注意を向ける対象の属性（種類情報）
2) 注意を向ける位置・大きさ（領域情報）
3) 注意を向けた理由（起因度情報）

1,2は「どこにどんなものがありそうなのか」という情報である.3は図6-7において注意仮説がパターン処理から生成されたものか,記号処理により生成されたものかを表す情報である.「それらしいものが見えたから」という理由と「隣に～というものがあったから」という理由では,注意仮説の性質が異なる.前者は画像から前向き推論により判断されたものであるが,後者は図面仮説に基づく知識処理から後ろ向き推論のゴールとして生成されたものである.

注意仮説において,形状情報に起因する度合いを形状起因度,意味情報に起因する度合いを意味起因度とし,それぞれ0.0～1.0の実数で表す.

6.3.2 基本アルゴリズム

注意仮説生成のための前向き推論型処理は,知識と独立した汎用的処理でなければならない.知識表現としては,パターン表現された形状モデルの利用が考えられる.そのような手法として,テンプレートマッチング法やニューラルネットの利用が考えられる.まず,それらに関して個別に検討する.

テンプレートマッチング

パターン表現された形状モデルに基づく認識手法として,2次元の画像パターンを直接テンプレートとして用い,画素単位のマッチングをとるテンプレートマッチングがある.一般に用いられる評価法に,2次元画像パターンをその画素数に等しい次元の空間上に投影し,2つのパターンに対応する点の距離や角度から類似度を計算する手法がある.2点間の距離により類似度を算出する前者の方法を最短距離法,多次元空間の原点から2点へ伸ばしたベクトルを考えてその2つのベクトルの角度により類似度を算出する後者の方法を類似度法という.

この手法の利点としては形状モデルの作成・追加が容易であることが挙げられる.問題点としては,類似形状の識別能力が低いことなどが挙げられる.

ニューラルネットワーク

ニューラルネットワークは,脳内のニューロンレベルの細胞のつながりを模倣して提案された認識手法である.そのため,人間の受動的認識処理の模倣を考えている本手法に適していると思われる.

ニューラルネットワークにも多数の提案がされている.バックプロパゲーション法を学習に用いたパーセプトロン型ニューラルネットワークは3層パーセプトロンの限界を

打ち破るべく提案された手法で [Amari88]，現在ではあらゆる分野で応用例が報告されている．汎用的なニューラルネットワークであると言える．

福嶋の提案するネオコグニトロンは，視覚神経系の簡単なものから複雑なものへ情報が処理され抽象化されていくような階層的な構造をモデル化したもので，ある範囲内での位置ずれや形状の変化の影響をほとんど受けることなく入力パターンを認識できる優れたニューラルネットワークである[Fukushima89]．さらに認識結果を入力画像にフィードバックして，重なりあうパターンから1つのパターンを抽出したり，形状の欠落を補完したりすることができる選択的注意機構を追加することも可能で，単体のパターン認識手法としてはかなり完成されたものとなっている．

ニューラルネットワークを利用した手法の利点は，適切に学習が行われればかなり高精度な認識結果が得られることである．しかし学習用に非常に多くのパターンを用意しなければならないこと，学習用の計算コストも決して低くなく，学習の際に局所解に落ち込んでしまう危険性があることなどが挙げられる．

ここで生成すべき注意仮説に要求される条件をもう一度考えてみる．注意仮説の仮説としての信頼度や精度は高いほどよいが，検証処理が続けて行われるので，この時点で高い信頼度・精度は要求されない．検証終了後に高い信頼度や精度が得られればよい．問題となるのは棄却率である．シンボルの可能性がある部分を，余すことなく注意仮説として生成できるほうが重要である．またシステムの汎用性の観点から，モデルの作成・追加が容易でなければならない．

以上のことから，単体のパターン認識性能としては優れているがモデルの登録に問題のあるニューラルネットワークを利用した認識手法を避け，上記の条件を満たしているテンプレートマッチング法を用いて注意制御機構の実現を考える．

図面の中にはあらゆる位置に様々な大きさ・種類のシンボルが存在する可能性がある．注意仮説を生成するにはそのような位置と大きさの変化に対応しなければならない．テンプレートマッチング法は，比較的小さな位置変化や大きさ変化には対応できるが，大きな変化には対応できない．これはテンプレートマッチング法が持つ根本的な問題である．このため次の手順で注意仮説を生成する．

1) 入力図面を異なる解像度へ変換した変換画像を生成する．
2) 解像度変換した画像からモデルパターンに応じた大きさの画像を切り出し，テンプレートマッチングを行う．
3) 画像を走査し，切り出す位置を変えながら2を繰り返す．
4) システムに登録されているすべてのモデルパターンに対して2,3の処理を繰り返す．
5) 1に戻る．

2のマッチング処理において，モデル毎に定められた閾値を越える評価値を得られたものを注意仮説として生成する．その注意仮説には種類情報としてそのモデルパターンの種類，領域情報としてそのときの入力図面上での絶対座標による位置・大きさ，形状起因度としてマッチング処理の評価値を与える．形状起因度v_{shape}の算出は認識処理での評価値v_{sim}とモデルに記述された閾値v_{thresh}により次のように行う．

$$v_{shape} = \frac{v_{sim} - v_{thresh}}{1.0 - v_{thresh}} \qquad (式6.1)$$

ここで，v_{sim}は1.0以下の実数とする．

1の解像度変換は図面上のパターンとモデルパターンの大きさの違いを吸収するためである．このとき本来2値画像である入力画像から2値画像へと変換したのでは，解像度を下げたとき，つまり画像の標本化レベルを下げたときに大きく情報量が欠落し，認識精度の大幅な低下が懸念される．そこで変換後の画像を濃淡画像とし量子化レベルを上げることで，解像度変換による情報量の欠落を緩和する．

解像度変換は整数分の1あるいは整数倍の解像度に限られない．このとき画像の1画素を単に「点」とみなして変換処理を行うと変換前と変換後の画素間で対応する画像数の違いから，エイリアシングが生じる．そこで図6-8に示すように画素を「面」とみなし，その面積を基準に変換を行う．変換前と変換後の画像で各画素を面積換算で対応を取ることによってエイリアシングのない解像度変換が可能になる．

このような解像度変換画像を用いるため，モデルパターンも濃淡値のパターンを用いたものであることが望ましい．また5における解像度の変更は，「粗いものから始めて徐々に細かいものへ」という処理戦略に基づいて行う．

以上が前向き推論型処理による注意仮説の生成方法であるが，この処理の計算コストは高い．しかしアルゴリズムは並列化可能であり，近年開発が急ピッチで行われている超並列計算機には適した処理である．

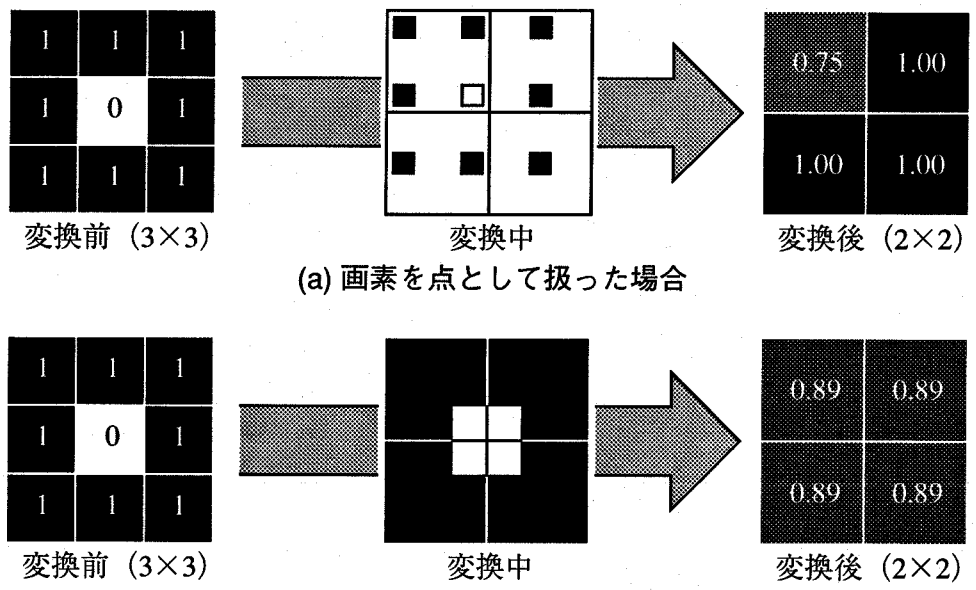


図6-8　入力画像の解像度変換

6.3.3　テンプレートマッチングの改良

6.3.2で述べた基本アルゴリズムにより，シンボルの位置・大きさの変化には対応できるが，シンボルには形状の変形の問題がある．これは，画質の劣化や手書きによる変形を許容するロバストなシンボル認識を行う場合には大きな問題となる．そこで，シンボルの形状の変形に対応するために以下の工夫を行った．

1) 低解像度のモデルパターンを用いる．
2) 評価の閾値を低く設定する．
3) 濃淡値のモデルパターンを用いる．
4) 類似度法による評価を行う．

1と2は認識の精度を意図的に落とすことで形状の変形を吸収しようとするものである．これは単独の認識処理として用いる場合には問題があるが，注意仮説の生成という役割は十分に果たせる．しかし，精度を維持することも大切であり，解像度変換画像と同様に量子化レベルを上げて粗い標本化による情報量の欠落を少しでも補う．そのための措置が3である．図6-9に「NPN型トランジスタ」のテンプレートパターンを示す．4は入力パターンとテンプレートとのパターン全体としての濃度差を吸収するために用いる．図6-10に示すように最短距離法

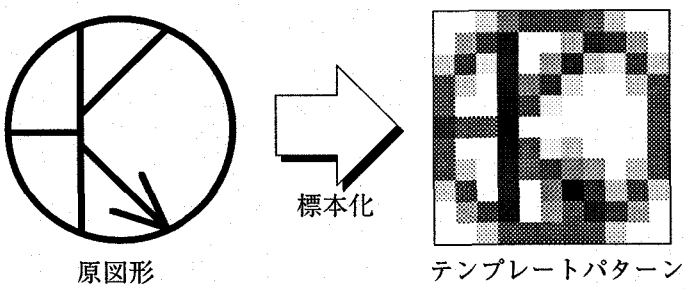


図6-9　モデルパターン

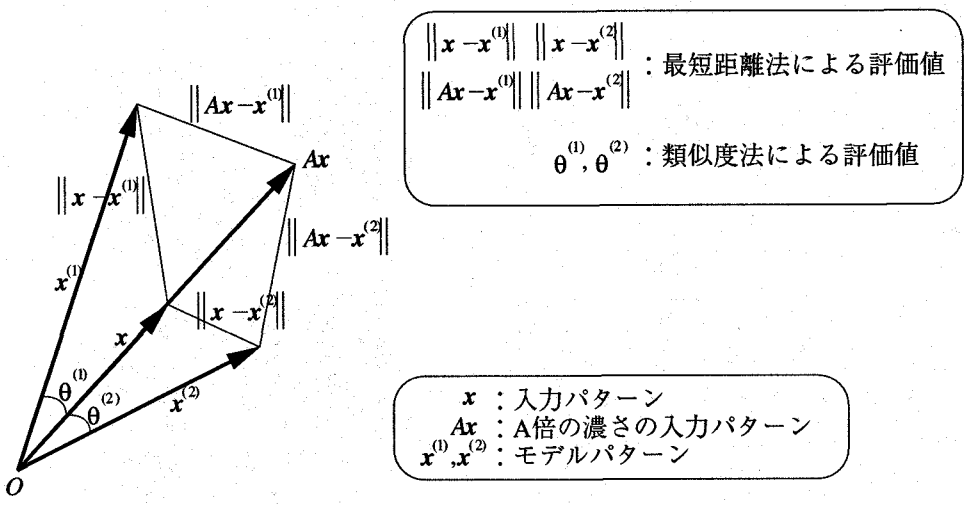


図6-10　最短距離法と類似度法　[Funakubo91] より

ではパターンの濃度変化に影響を受けるが，類似度法ではパターンの全体的な濃度変化に対して不変である [Funakubo91].

6.3.4　類似度法の検討

類似度法では，入力パターンとモデルパターンをその画素数に等しい次元の空間上に投影し，それぞれの方向ベクトルを求めそのベクトル間の角度により類似度を算出する．仮にパターンが$m \times n$のパターンであれば，mn個の軸を持つmn次元超空間上に投影される．入力パターンの方向ベクトルを$\mathbf{a}$，モデルパターンの方向ベクトルを$\mathbf{b}$とすると，その類似度sは次のように表される．

$$s = \cos\theta = \frac{(\mathbf{a},\mathbf{b})}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} \qquad \text{(式 6.2)}$$

θはベクトル$\mathbf{a},\mathbf{b}$のなす角である．方向ベクトルの各軸方向の成分には，その軸に対応する画素の画素値がそのまま用いられるのが一般的であり，通常0.0以上1.0以下の値をとる．例えば図6-11のような画像パターンがあれば，そのパターンを表すmn次元超空間上のベクトル$\mathbf{a}$は，その画素値を1次元配列した

$$\mathbf{a} = (a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, \ldots, a_{mn}) \qquad \text{(式 6.3)}$$

で表されるわけである．

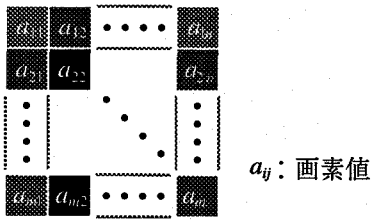


図6-11　ベクトル$\mathbf{a}$の画像パターン

ところがそのように成分の決定を行った場合，通常は背景（白画素）にあたる0.0の値を持つ画素が全く成分を持たないことになり，その結果図6-12のような4つのパターンに対して次のような不合理な認識結果を引き起こす．(a) (b) のような白画素の中に1画素だけ黒画素が混入した場合には，2つのパターンを表す方向ベクトルは直交し，類似度は0.0となる．ところがパターンとしては単に白黒が反転しただけの (c) (d) の場合には，2つの方向ベクトル

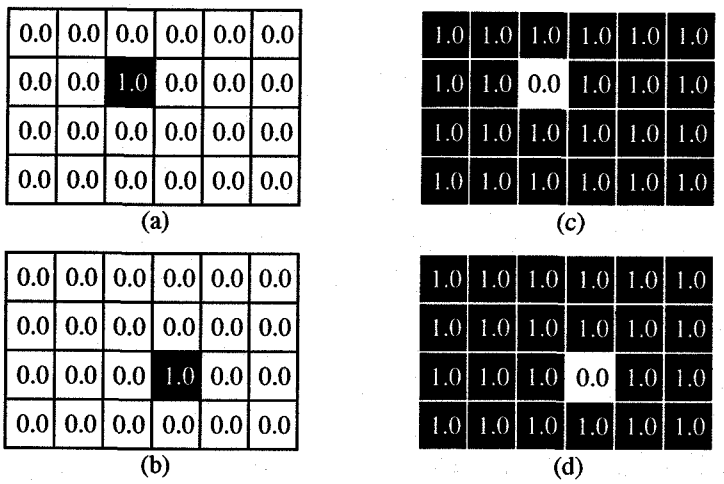


図6-12　4種類の画像パターン

は直交せず，むしろそのベクトル間の角度は0に近いものであり，類似度も1.0に近い値を示す．このような白画素と黒画素の取り扱いが異なるという不都合は，塗りつぶし領域を持つシンボルに対する注意仮説と，そうでないシンボルの注意仮説の生成感度に違いが生じるなどの問題を引き起こす．

そこで，次のようにベクトル$\mathbf{a}$の成分を変換してベクトル$\mathbf{a}'$を生成する．

$$\bar{a} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} \qquad (式6.4)$$

$$a'_{ij} = a_{ij} - \bar{a} \quad (i = 1 \sim m, j = 1 \sim n) \qquad (式6.5)$$

ベクトル$\mathbf{b}$に対しても同様な成分変換を行う．変換後のベクトル$\mathbf{a}'$はその成分として負の値も持つようになる．変換前の成分が0.0以上1.0以下である場合，変換後の成分は-1.0以上1.0以下の範囲にその値が写像される．この結果2つのベクトル間の角度の範囲も0度から180度までと広がる（図6-13）．

（式6.2）の類似度の式をそのまま用いると，類似度sは-1.0以上1.0以下の値域を持つ．類似度sが-1.0となるのは2つのパターンの画素値が完全に反転した場合であり，類似度sが0.0となるのは2つのパターン間に全く関連性がない場合である．図6-12の例では (a) (b) の類似度は (c) (d) の類似度と同じく1.0に近い値を示す．また (a) (c) および (b) (d) の類似度は-1.0となる．

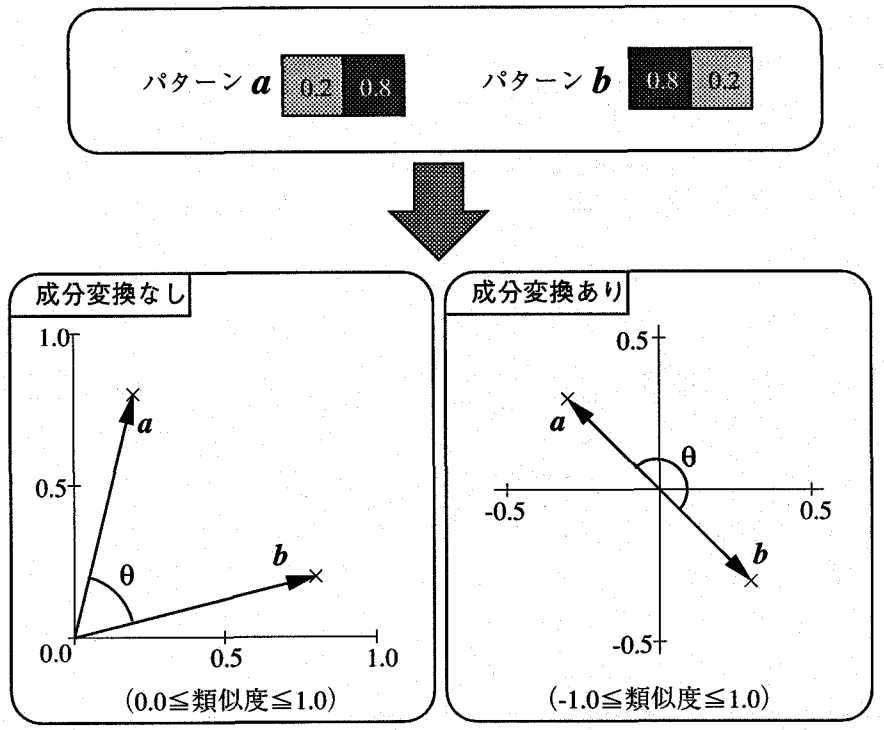
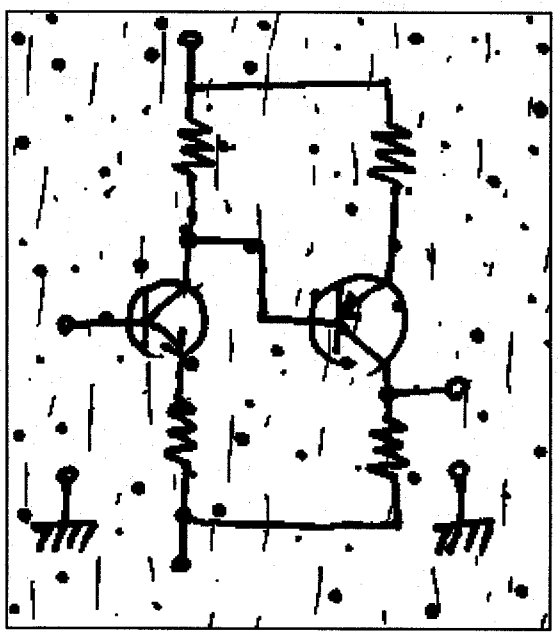


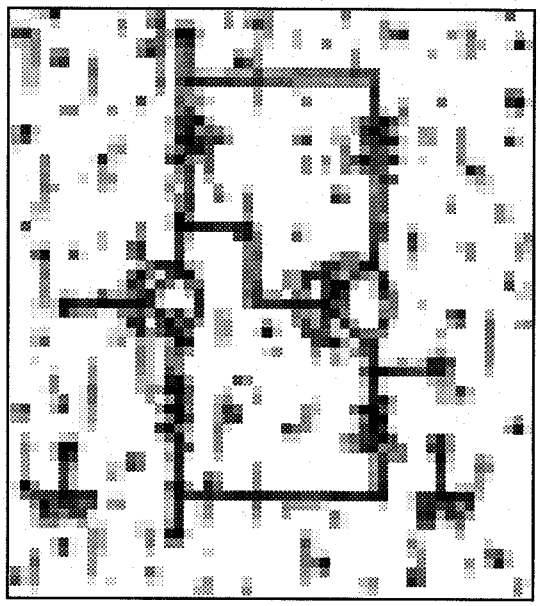
図6-13　ベクトルの成分変換

6.3.5　実験

注意仮説の生成例を示す．図6-14に入力画像とその解像度変換画像の例を示す．図6-15には4種類のシンボルについてそのモデルパターンと生成された注意仮説の領域を示す．同一の領域に対し，異なるシンボルの注意仮説が生成されていることがわかる．



(a) 入力画像



(b) 解像度変換画像

図6-14　解像度変換の例

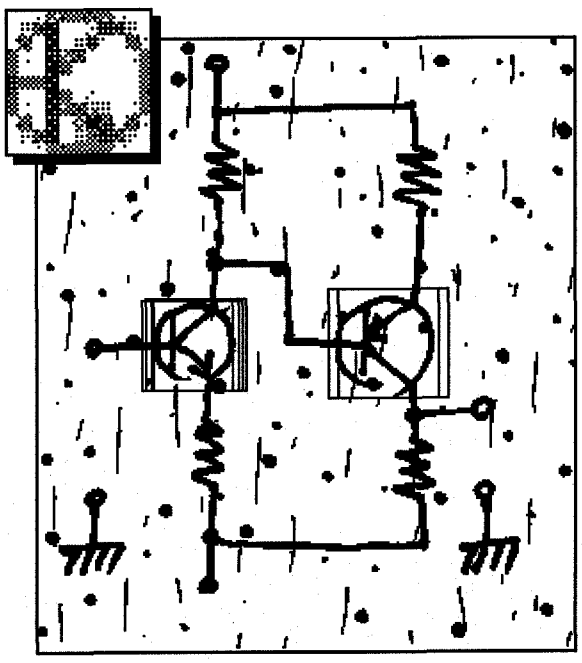
6.3.6　検討

ここで採用した注意仮説生成のアルゴリズムは，形状の変形や図面の質の劣化等に対してロバストな仮説生成が可能であるが，注意仮説の精度が犠牲になっており，生成される仮説に冗長性が生じる．具体的には図面上の同一の領域に対して重複して注意仮説が生成される場合が生じる．実験例からもわかるように，この重複には次の2つの場合がある．

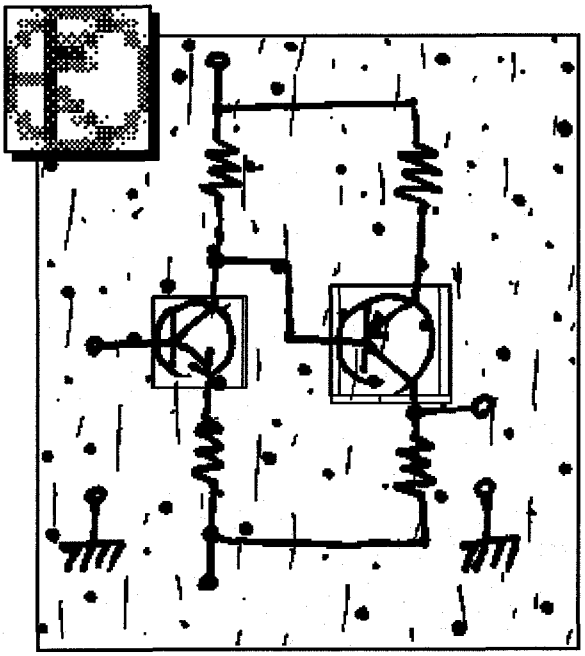
1) 対象となるシンボルは同じで，領域がわずかに異なる複数の注意仮説が生成される場合
2) 同じ領域に，異なるシンボルに対する複数の注意仮説が生成される場合

1は解像度変換において画像の位相がわずかにずれたパターンが切り出された場合に生じ，領域がわずかに異なる仮説が生成される．このような注意仮説は，処理の冗長性をなくすために1つ統合することが望ましい．統合の基準としては，各注意仮説についての矩形領域を基にそれらの重なり具合を用いる．現時点では領域の50%以上がお互いに重なりあっている仮説を統合している．また統合処理では単に余分な注意仮説を削除するだけではなく，起因度情報の更新も行う．この処理によって生成された注意仮説は，形状に起因した起因度情報である形状起因度を持つが，統合処理によりこの形状起因度を増加させる．

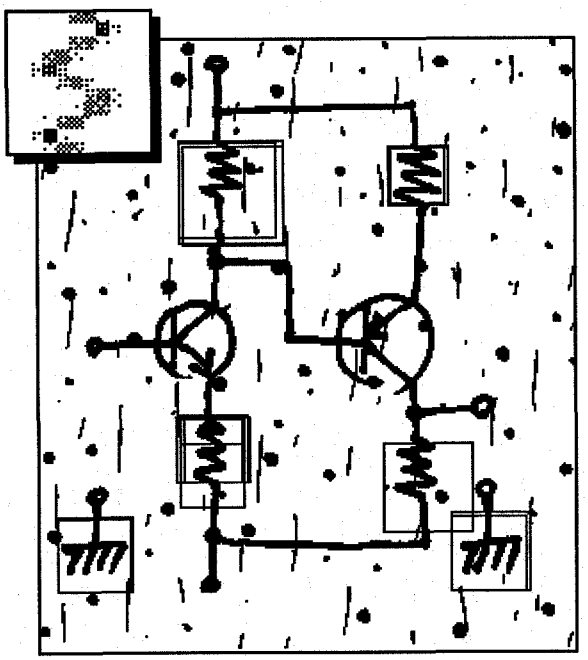
また低解像度のモデルパターンの利用や低い閾値の影響で，2のように間違った注意仮説も同時に生成されてしまう．このような注意仮説の競合は，仮説検証処理で積極的に利用される．これに関しては6.4で説明する．



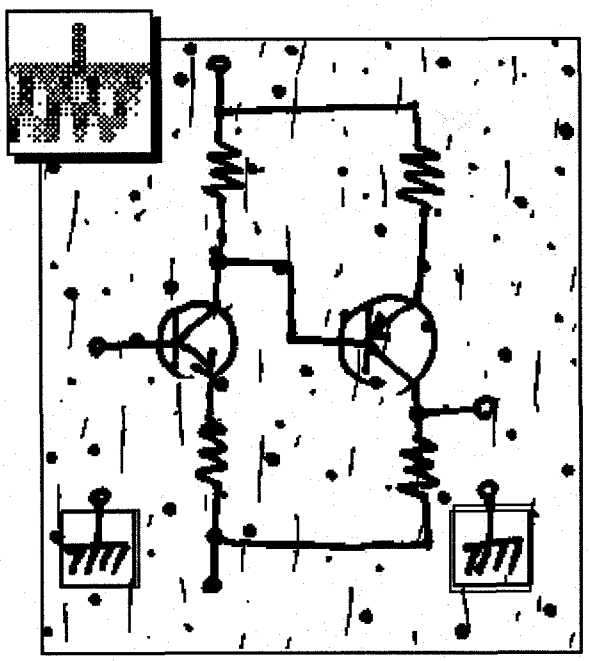
(a) NPN型トランジスタ



(b) PNP型トランジスタ



(c) 抵抗器（縦方向）



(d) アース

図6-15　モデルパターンと注意仮説の生成された領域

6.4 状況適応型認識処理

注意仮説は一般に精度が悪いので，必ず検証が行われなければならない．仮説検証処理は，注意仮説に基づく後ろ向き推論型の認識処理のことであり，前向き推論型認識処理にはない利点がある．

6.4.1 注意仮説に基づく認識処理

注意仮説に基づく認識処理では，注意仮説の持つ領域情報とシンボル情報を利用することができる．前向き推論型処理では図面全体を処理対象とし汎用的な認識アルゴリズムを採用したが，後ろ向き推論型の認識処理においては対象が限定されるため，知識を積極的に利用することで状況に応じた適切な処理を行うことができる．ここでは，次の3つの方針について検討する．

1) 処理におけるパラメータの最適化
2) 競合情報の積極的な利用
3) 起因度情報の利用

1は注意仮説により処理の対象と領域を限定できることから，それに応じたパラメータの最適化を図ることができるということである．例えば，弾性モデルによる切り出しを行う場合には，モデルの初期位置が精度を決める非常に重要な要因となるが，注意仮説の持つ領域情報を基に初期位置や大きさを決定できる．2, 3に関しては，以下で詳しく説明する．

6.4.2 競合する注意仮説情報に基づく認識戦略

注意仮説の競合情報を積極的に利用することで，認識処理の精度や効率を上げることを考える．前向き推論型処理により同一の領域に対して生成された異なる注意仮説間では，シンボル構造の形状特徴が類似している場合が多い．例えば6.3.5で示した例では，NPN型トランジスタとPNP型トランジスタという構造的に非常に類似した形状の描かれた領域に対し，互いのシンボルとしての注意仮説が生成されている．図面という対象を取り扱う場合，これら注意仮説は互いに競合していると見なすことができ，正しい注意仮説は高々1つである．ときにはすべてが誤っている可能性もある．

競合注意仮説に対して認識処理を独立に行う場合，それら注意仮説の形状差異が非常に小さければ，評価値にその差を判断するのに十分な違いが現れないことがある．これは同一形状部分の評価が大部分を占め，実際に識別の決め手となる形状部分での評価がほとんど反映されずに打ち消されてしまうときなどに生じる．そもそも前向き推論型認識処理において競合仮説が生成される要因となるのも，その認識手法がパターンを全体として評価しているからである．

そこで，競合する注意仮説間の特徴的な差に注目し比較処理を行うことを考える．つまり，競合注意仮説間の同一形状部分と同一形状でない部分区別し，後者の部分の評価に重点を置いた認識プランを作成・実行することで，微細な違いにも対応できる認識を行う．

また競合の中には，本来シンボルなど存在しない領域にも誤って競合する注意仮説が生成される場合もある．これはノイズの影響が大きい場合に生じやすい．このような場合には，同一

部分形状に対する評価値が著しくなることが予想されるので，そのときは認識処理を打ち切り，それらの注意仮説を全て棄却する．冗長な処理を回避して処理の効率化を図るためである．

6.4.3　形状起因度に基づく認識戦略

注目する注意仮説が形状あるいは意味的関連性にどの程度起因して生成されたかによって，処理を変更することが可能である．図6-7で示すように，前向き推論型認識処理以外にも記号処理部における意味解析処理によっても注意仮説は生成される．前者の処理により生成される注意仮説は形状に起因したもの，後者の処理により生成される注意仮説は意味的関連性に起因したものである．注意仮説情報の中でそれぞれへの起因度情報を表すものが形状起因度と意味起因度である．

図面解釈過程の初期の段階においてパターン処理によって生成された注意仮説は意味起因度を持たない．これはその注意仮説がパターン処理によって生成されたからで，入力画像とモデルパターンの形状類似度のみがその注意仮説の正当性を保証しているからである．また図面解釈過程の進行と共に，図面仮説に対する意味解析処理が行われ，その結果，シンボル間の意味的関連性などに基づいて生成される注意仮説においては，形状起因度は特になく，意味起因度を持つことになる．しかし，注意仮説の起因度は6.3.6で述べたように注意仮説の統合処理によって変動する．同じ種類の注意仮説が同一領域に生成された場合，その起因度は統合処理により増加する．

注意仮説に基づく認識処理では，これらの起因度によって認識戦略を変更すべきである．形状起因度，意味起因度の両者の値により，形状の検証処理の詳細さのレベルを変更する．基本的には，起因度が低い場合にはより詳細な検証を，起因度が高い場合には簡単な検証を実行する．

6.5　認識戦略の自動生成

6.5.1　シンボルの記述

6.4で述べた認識処理機構を導入するためには，各シンボルのモデル表現が重要となる．仮説の競合状態において形状の違いに注目した処理を行うためには，認識対象シンボルのモデルとして，部分形状による構造記述を行っておく必要がある．競合注意仮説の示すモデルの構造記述を比較することで，同一形状とそうでない部分を区別する．そのためにどのような部分形状に分解して記述するかも重要なことである．起因度情報による認識の詳細さのレベルの変更に関しては，各シンボルの部分構造記述の中にその構造特徴の重要度を表現しておく必要がある．

そこで，モデル表現として部分形状を形状特徴の重要度の点で階層的に表現することを基本とし，それぞれの部分形状毎に次のような情報を記述する．

1) 部分形状名
2) 階層レベル
3) 相対位置とその基準

4) 形状寄与率
5) 評価基準値

部分形状名はその部分形状に関するさらに詳しい情報が必要なとき,例えばその部分形状の切り出し処理に関するパラメータ情報が必要になったときなどに,その部分形状名を冠するモデルを参照し必要な値を得ることになる.

階層レベルとはシンボル全体を階層記述する上でその部分形状がどの階層に位置するのかを表すものである.これは形状特徴としての大きさを表現するものでもあり,階層レベルが高い部分形状ほど早い段階で処理が行われる.1つの階層に複数の部分形状が存在してもかまわない.

相対位置とはその部分形状がシンボル領域の中でどの部分に存在するのかを表すものであり,部分形状毎の認識処理を行う際に処理領域を指定するために用いられる.実際に描かれたシンボルの形状が変形していることを考慮すると,複雑なシンボルの部分形状の切り出しを弾性モデルにより一度に行うと,細かな部分でエラーがおきる可能性が高い.そこで,大きな構造要素から段階的に切り出しを行うことを考え,先に切り出された部分形状の位置情報から,次に切り出すべき部分形状の位置情報の修正を行う.そのために,位置情報の修正において基準となる部分形状とそれに対する相対位置という2つの情報を表現しておく.

また形状起因度に基づく認識処理を実現するために,各部分形状に次の情報を形状寄与率として記述する.これはシンボルの形状の中でそれぞれの部分形状がどの程度「そのシンボルらしさ」に寄与しているかをしめす値であり,その部分形状の評価値に対する重みとみなすこともできる.この値はそのシンボル全体の合計で1.0となるように,0.0～1.0の範囲の実数値で表す.

評価基準値は,部分形状毎の形状評価値を正規化するため値である.ここで各シンボルの評価法を示しておく.i番目の部分形状の形状寄与率をr_i,評価基準値をb_i,認識処理による評価値をe_iとすると,そのシンボルの評価値eは次の式で表される.

$$e = \frac{\sum_{i=1}^{n} r_i \frac{e_i}{b_i}}{\sum_{i=1}^{n} r_i} \qquad \text{(式 6.6)}$$

nをある階層の部分形状の番号とすれば,その階層以上でのシンボルの評価値を求めることができる.図6-16に階層記述の概念図を示す.

6.5.2　認識戦略の生成アルゴリズム

認識処理には,競合の解消と起因度への対応の2つが求められる.この2つを満たす認識戦略を後ろ向き推論型の記号処理の中で生成する.まず競合があれば競合の解消を行い,その結果1つに絞られた注意仮説に対し起因度に応じて検証処理の詳細レベルを決定する.

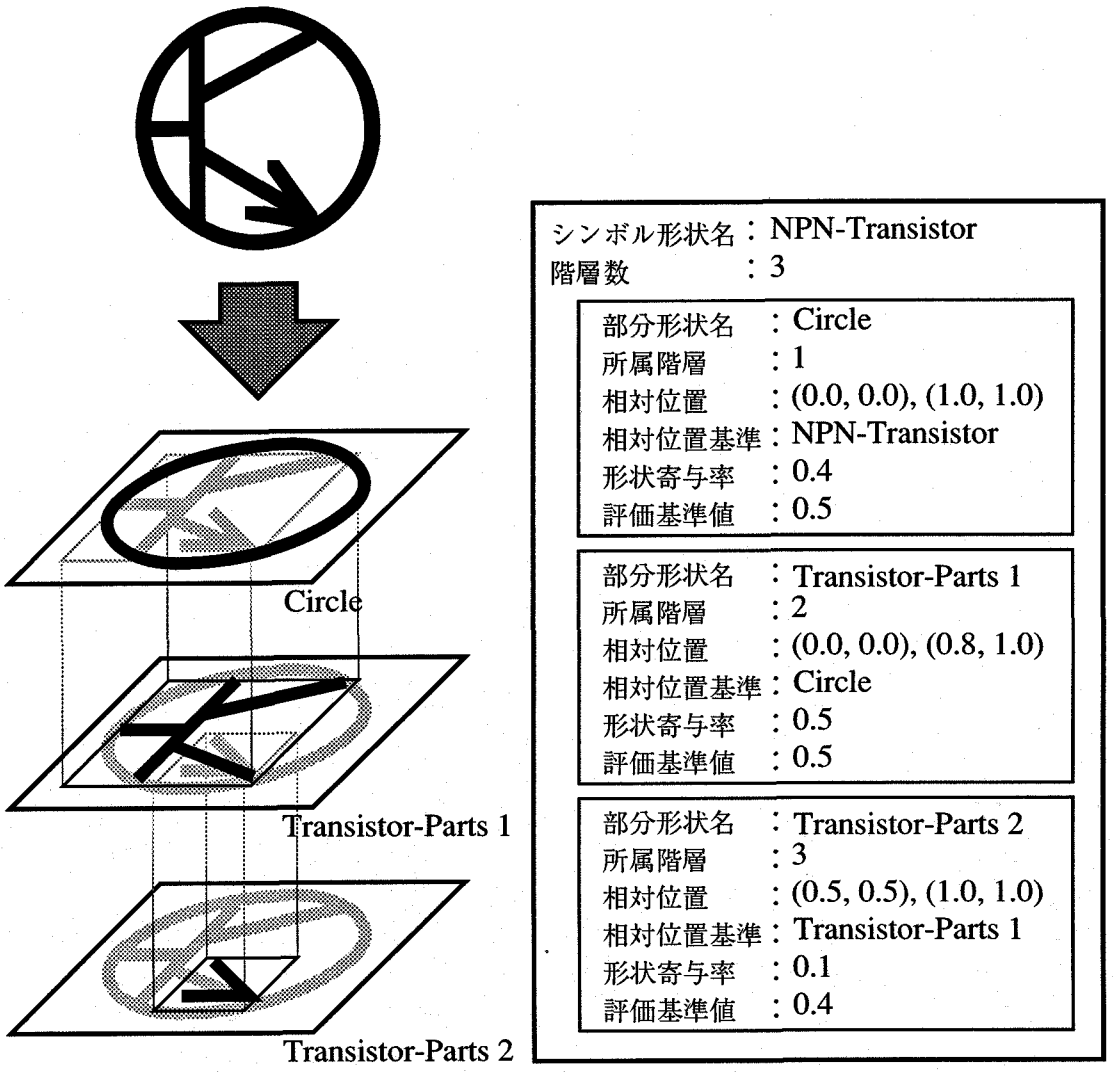


図6-16　部分形状による階層記述の概念図

6.5.2.1　競合の解消

競合解消を効率的に行うために，競合している注意仮説の示すモデルの部分形状を比較し，それに基づいて認識戦略を動的に生成する．その認識戦略は認識木の形で表わされ，部分形状毎に検証処理が行われる．部分形状の位置・大きさに関わるパラメータは，処理過程の中で動的に設定される．

認識木の生成は次のような手順で行う．トランジスタなど3つのシンボルが競合した場合の認識木の生成の様子を図6-17に示す．

1) 領域の競合している注意仮説を選び出す．
2) 各注意仮説に対応するモデルの形状記述を読み出す．
3) 形状記述の最上位の階層に注目する．
4) 注目する階層において部分形状の種類と位置の同一性を確かめる．
5) 4の同一性を満たす注意仮説の集合を形状グループとして構成する．
6) 5以外の注意仮説は単独の形状グループを構成する．

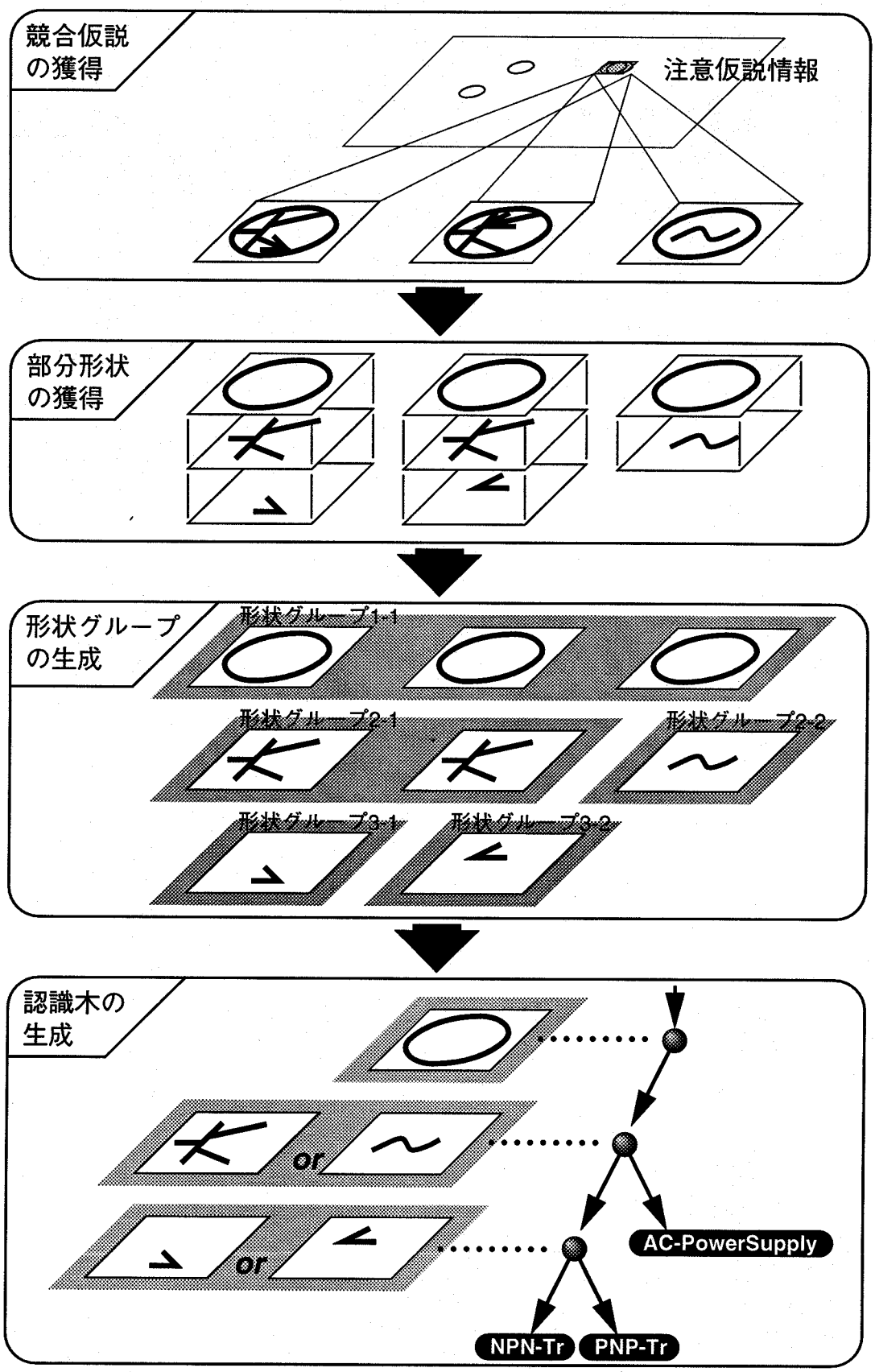


図6-17　認識木の生成例

7) それぞれの形状グループをノードとして認識木とする.
8) 次の階層に注目し，各形状グループ毎に4から7の処理を繰り返す.
9) 形状グループが1つの注意仮説しか含まない場合，その階層で処理を停止する.

6.5.2.2 起因度に応じた処理

これまでの処理で，認識木の末端のノードに含まれる注意仮説はすべて1つである．それらに対して，注意仮説の持つ2つの起因度（形状起因度と意味起因度）に基づいて認識戦略を生成する.

まず2つの起因度から，その形状を認識するうえで必要とされる形状寄与率必要値r_{min}を算出する．形状起因度をv_{shape}，意味起因度を$v_{meaning}$とすると，形状寄与率必要値r_{min}は次のように算出する.

$$r_{min} = 1.0 - \left\{ v_{shape} + \left(v_{meaning} \right)^{\beta} \right\}^{\alpha} \qquad (式6.7)$$

パラメータα, βはそれぞれ正の実数値であり，ここでは$\alpha = 2.0, \beta = 0.5$の値を用いる．これは実験により求めたものである.

次に検証済みの全部分形状の累計形状寄与率r_{sum}を算出する．累計形状寄与率r_{sum}が形状寄与率必要値r_{min}を越えるまで，最上位の階層からまだ検証の対象になっていない部分形状の検証ノードを追加する.

6.5.3 認識戦略の実行アルゴリズム

生成された認識木に従って，以下の手順で認識処理が行われる.

1) 認識木の最上位ノードに注目する.
2) 注目するノードで対象となる各形状グループの切り出し処理を行う.
3) 切り出しにおける歪み率をその形状グループの評価値とする.
4) 注目するノードにおける累積評価値を計算する.
5) 累積評価値が基準値を満たさない場合，そのパスの処理を停止する.
6) 残ったパス間で累積評価値の相対的評価を行う.
7) 相対評価値が基準より低い場合，そのパスの処理を停止する.
8) 1つ下の階層に注目し，2から繰り返す.

4ではその注意仮説が示すシンボルの記述に基づき，形状寄与率と評価基準値とから6.5.1に示した式により評価値の累積を行う．5は認識処理の評価値が著しく低いパスを早い段階で除外するためである．7によって他のパスに対してその評価値が大きく劣るパスを除外し，処理の効率化を図っている．これは単に累積評価値の大小関係だけに注目するのではなく，評価値間の差についても考慮する．つまり評価値同士が近似している場合は，その階層でのわずかな大小の違いで判断せずに，先の階層での認識結果を待って，より明確な差として現れた段階で判断する.

すべての処理が終了した時点で，残ったパスに属する注意仮説が認識結果となる．すべての

パスの処理が途中で停止した場合，すべての注意仮説が棄却される．複数のパスが残った場合は，最大の評価値をもつノードの注意仮説を認識結果とする．

6.6　シンボルのモデル記述

ここで提案する手法を実現するためには，シンボルの記述としてパターン表現と記号表現を兼ね備えたモデル記述が必要となる．図6-18にモデル記述の概念図をNPN型トランジスタを例として示す．モデル記述についてこの例を用いて説明する．

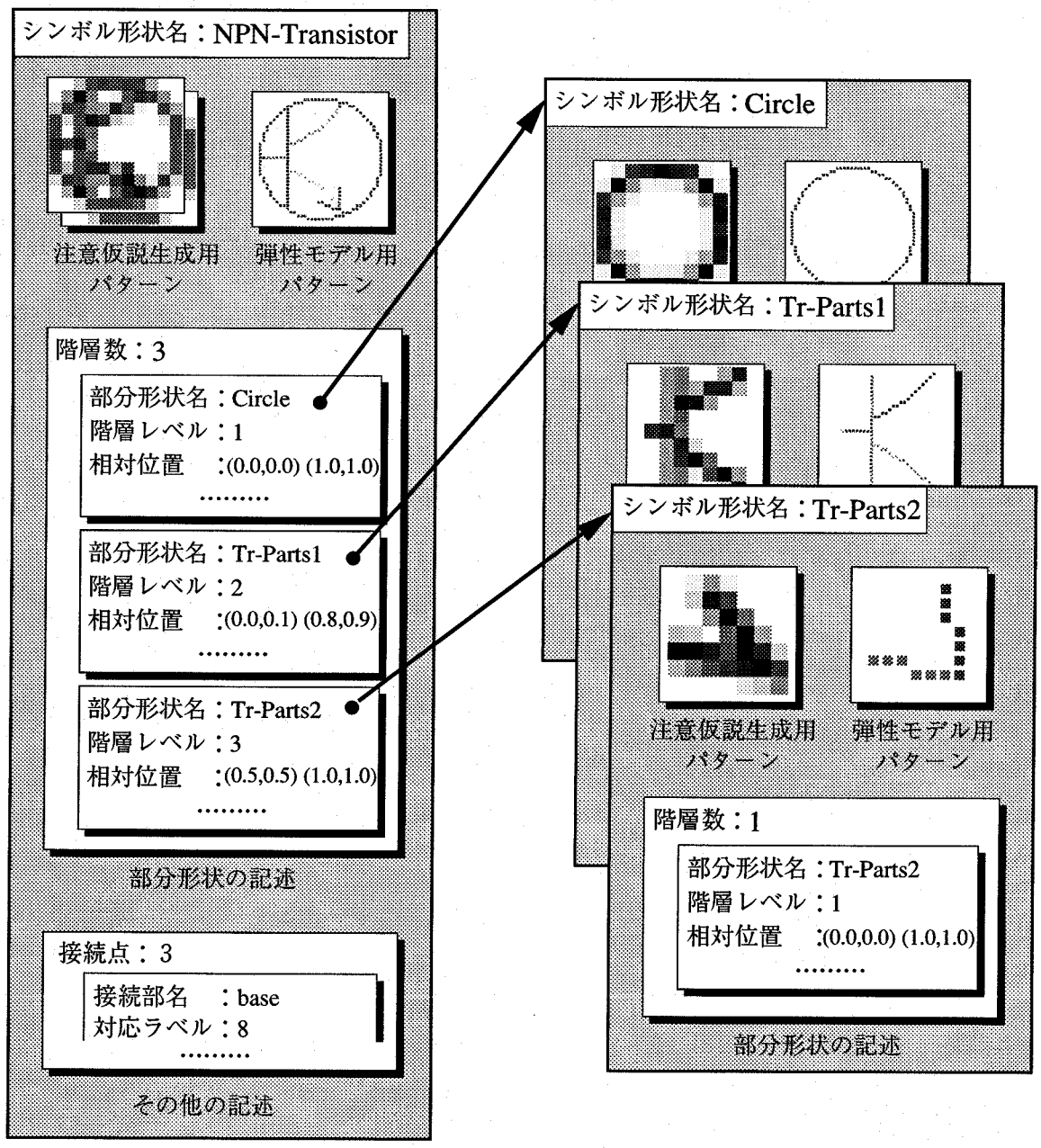


図6-18　モデル記述の概念図

6.6.1　注意仮説生成用のモデル記述

注意仮説の生成のためには，低解像度の濃淡値パターンを使用する．低解像度のパターンを使用するだけでは吸収しきれない変形等は，複数のパターンを準備する．回転等の変形に対しても，予めそのような変形を加えたパターンを準備する．NPN型トランジスタの場合は2種類のパターンを使用した．その2つのパターンを図6-19に，また1つのパターンについて実際の記述形式であるテキスト形式で図6-20に示す．

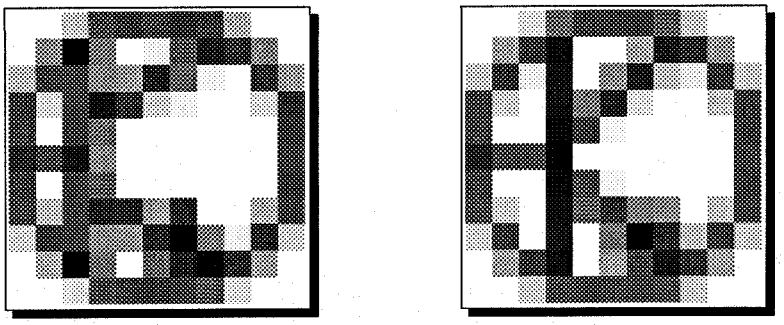


図6-19　注意仮説生成用パターン

```
#### NPN-Transistor pattern for Attention Generation
11 11    -- パターンサイズ
0.8      -- 注意仮説生成時の変換解像度基準値
0.5      -- テンプレートマッチングの閾値
0.00000 0.00100 0.09423 0.51087 0.61445 0.53698 0.64384 0.59101 0.13163 0.00100 0.00000
0.00100 0.28148 0.93345 0.33321 0.00100 0.01093 0.49071 0.74536 0.59235 0.27784 0.00100
0.11427 0.60340 0.59344 0.30917 0.21603 0.74013 0.40425 0.01275 0.00100 0.59903 0.14633
0.56563 0.06618 0.58579 0.87553 0.63619 0.06910 0.01000 0.00100 0.00100 0.03947 0.61773
0.53698 0.00100 0.58579 0.41336 0.00100 0.00100 0.00000 0.00000 0.00000 0.00100 0.59842
0.75070 0.53698 0.83424 0.30917 0.00100 0.00000 0.00000 0.00000 0.00000 0.00100 0.53831
0.59976 0.00100 0.58579 0.47359 0.00255 0.00100 0.00100 0.00000 0.00000 0.00100 0.59575
0.52823 0.12228 0.58579 0.75653 0.72848 0.22119 0.66181 0.00100 0.00000 0.12228 0.50285
0.06217 0.62344 0.59344 0.30917 0.26393 0.74281 1.00000 0.33182 0.00765 0.65282 0.06217
0.00100 0.16794 0.92678 0.38664 0.00100 0.46108 0.79526 0.92702 0.66084 0.16794 0.00100
0.00000 0.00100 0.04748 0.50553 0.59575 0.65452 0.59575 0.47614 0.04748 0.00100 0.00000
```

図6-20　注意仮説生成用パターンの実際の記述

6.6.2　注意仮説検証用のモデル記述

注意仮説の検証処理では，シンボルの部分形状を階層的に表し，それを利用して認識戦略を生成する．そのための階層構造の記述をとして，1つのシンボルの各部分形状に次の情報を記述する．

1) 部分形状名
2) 階層レベル
3) 相対位置とその基準
4) 形状寄与率
5) 評価基準値

階層レベルは全体の階層構造の中でのその部分形状の階層を整数値で表す．相対位置は外接長方形領域で表現する．その基準となる部分形状の外接長方形について，その左上点の座標を(0.0, 0.0) 右下点の座標を (1.0, 1.0) とみなした座標平面上で，同じく左上点と右下点の2点の座標で示す．形状寄与率ならびに評価基準値はともに0.0～1.0の値をとり，そのうち形状寄与率は各部分形状の合計値が1.0になるように設定する．

各部分形状はそれぞれを1つのシンボルとみなし，詳しい情報はその部分形状についてのモデル上に記述する．部分形状のみ分離抽出を行う際には，その部分形状のモデルの分離抽出用のパターンを用いる．

NPNトランジスタの階層記述の概念図は図6-16および図6-18で示した．図6-21に実際の記述形式であるテキストデータを示す．

```
##### NPN-Transistor
10          -- モデルIDnumber
npn1        -- モデル名
3           -- 部分形状数

 1          -- 部分形状IDnumber
 circle     -- 部分形状名
 1          -- 階層レベル
 1          -- 基準形状参照用IDnumber
 0          -- 相対位置基準形状IDnumber
 0.0  0.0   -- 相対位置（左上点）
 1.0  1.0   -- 相対位置（右下点）
 0.4        -- 形状寄与率
 0.5        -- 評価基準値
 1          -- 弾性モデルラベル数
   1          -- ラベル変換用テーブル

 14         -- 部分形状IDnumber
 trl-in     -- 部分形状名
 2          -- 階層レベル
 2          -- 基準形状参照用IDnumber
 1          -- 相対位置基準形状IDnumber
 0.0  0.1   -- 相対位置（左上点）
 0.8  0.9   -- 相対位置（右下点）
 0.5        -- 形状寄与率
 0.5        -- 評価基準値
 7          -- 弾性モデルラベル数
   2 3 4 5 8 9 10  -- ラベル変換用テーブル

 15         -- 部分形状IDnumber
 trl-yl     -- 部分形状名
 3          -- 階層レベル
 3          -- 基準形状参照用IDnumber
 2          -- 相対位置基準形状IDnumber
 0.5  0.5   -- 相対位置（左上点）
 1.0  1.0   -- 相対位置（右下点）
 0.1        -- 形状寄与率
 0.5        -- 評価基準値
 2          -- 弾性モデルラベル数
   6 7        -- ラベル変換用テーブル
```

部分形状1

部分形状2

部分形状3

図6-21　階層構造の記述

部分形状の評価には，弾性モデルによる切り出し処理を用いる．図6-22にNPNトランジスタの各部分形状についての弾性モデル生成用のパターンを示す．また図6-23にその1つについて実際の記述形式であるテキスト形式を示す．

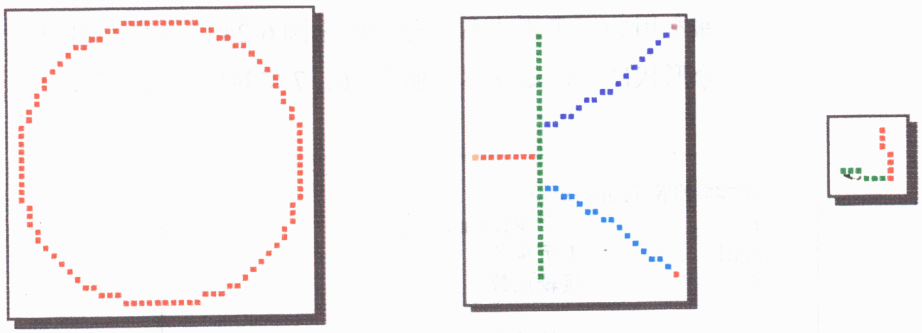


図6-22　弾性モデル生成用パターン

```
#### Transistor-Partsl pattern for Elastic Model Generation
27 35    -- パターンサイズ

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 6
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 3 3 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 3 3 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 5 1 1 1 1 1 1 1 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 4 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 4 4 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 4 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 4 4 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 4 4 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 4 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 4 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 4 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 4 0 0 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 4 4 0 0 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 4 0
0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 7
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

図6-23　弾性モデル生成用パターンの実際の記述

6.6.3　その他の情報の記述

意味解析を行うには，接続線の接続情報や他のシンボルとの接触情報などを記述しなければならない．例えば電気回路図面におけるシンボルの場合は接続線がつながる部位を特定する必要がある．それぞれの動作まで考慮した意味解析を行う際には，動作に関する各種知識に関連して，様々な情報をシンボル毎に記述しておく必要が生じる．そこで，切り出し処理のための部分形状の表現のラベル情報を用いて表現する．その例を図6-24に示す．図中のラベル番号8,9,10は図6-22で示した部分形状におけるラベル値5，6，7に対応している．

```
##### NPN-Transistor
10          -- モデルIDnumber
npn1        -- モデル名
3           -- 接続点数

  1           -- 接続点IDnumber
  base        -- 接続点名
  8           -- 対応ラベル番号

  2           -- 接続点IDnumber
  collector   -- 接続点名
  9           -- 対応ラベル番号

  3           -- 接続点IDnumber
  emitter     -- 接続点名
  10          -- 対応ラベル番号
```

図6-24　接続情報の記述

6.7　実験

6.7.1　処理対象図面

電気回路図面を対象にシンボル仮説の生成実験を行った．図質に対するロバスト性を示すために，手書きによる変形やノイズ等の加わった図面を用いた．また，汎用性を示すために，楽譜を対象に実験を行った．

6.7.2　モデル

1) 電気回路図面

認識の対象とするシンボルのモデルとして，トランジスタ（NPN，PNP），交流電源，抵抗器（縦方向，横方向），コンデンサ（縦方向，横方向），アースの5種類8パターンを準備した．

2) 楽譜

認識の対象とするシンボルのモデルとして，2分音符（上向き，下向き），4分音符（上向き，下向き），8分音符（上向き，下向き），4分休符，8分休符，ト音記号，ヘ音記号の6種9パターンを準備した．

6.7.3 実験

システムはUNIXワークステーションHP9000/712上に構築した．記号処理部はルール記述型言語OPS83とC言語を用いて記述し，その他はC言語を用いた．画像入力はイメージスキャナーを用いて，100dpiで取り込んだ．

1) 電気回路図面

図質の異なる様々な電気回路図面に対して実験を行った．ここでは代表的な4種類の図面に対する実験結果を示す．各種処理パラメータはいづれの図面に対しても同一の値を使用している．図6-25に印刷図面に対する結果，図6-26にノイズ・かすれ等は含まない手書きによる図面に対する結果，図6-27にノイズの混入した手書き図面に対する結果，図6-28にかすれの生じた手書き図面に対する結果を示す．

手書きの変動に対してはおおむね良好な結果が得られた．トランジスタの部分に対して特に注意仮説の競合が見られたが，動的に生成された認識戦略により効果的に競合の解消が行われている．

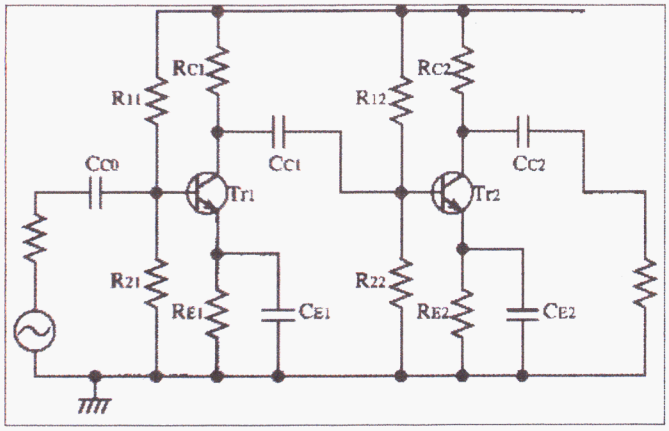
ノイズの混入に対してもおおむね良好な結果が得られた．通常はノイズの混入した画像からノイズを含まないように形状を抽出するのは難しいと言われているが [Ohsawa89]，弾性モデルを利用した分離抽出処理により比較的精度良くシンボル形状の抽出が行われている．

今回の実験ではかすれによる影響が最も大きく，生成されない仮説が多くあった．手書きの際に生じる線のとぎれに関しては，形状としての欠落が少ないので十分に対応可能であるが，かすれの影響で形状の一部が大きく欠落した場合は，注意仮説の生成が難しい．また検証処理での部分形状の切り出しミスも発生した．このようなかすれに完全に対応するには意味解析処理からの後ろ向き推論による仮説生成・検証が有効と考えられる．提案する手法ではそれが可能ではあるが，意味解析部ので知識が現状では充実しておらず，この実験では期待通りの動作にはならなかった．

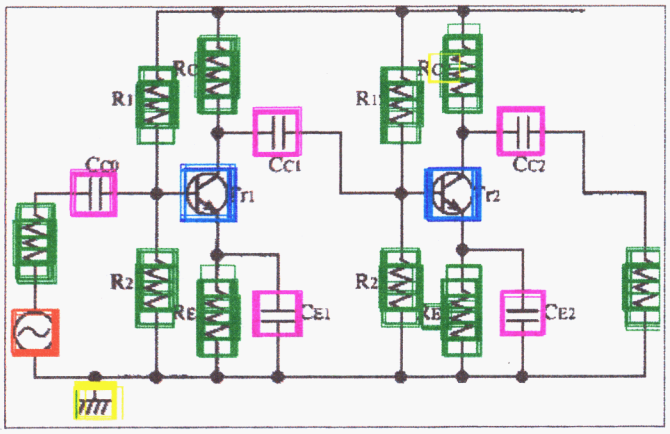
2) 楽譜

印刷楽譜に対して要素仮説の生成を試みた．図6-29にその結果を示す．2分音符と4分音符の間で音符の種類の誤認識が見られた．4.7.4で述べた切り出し手法の塗りつぶし領域への対応を試みてはいるが，その性能が満足のいくもでではなく，それが原因となっている．連符に関しては，固定形状でないために1つの形状としてモデル化できないので，今回はシンボルとして登録していない．そのため連符の部分はその部分形状である4分音符としてシンボル仮説の生成が行われている．

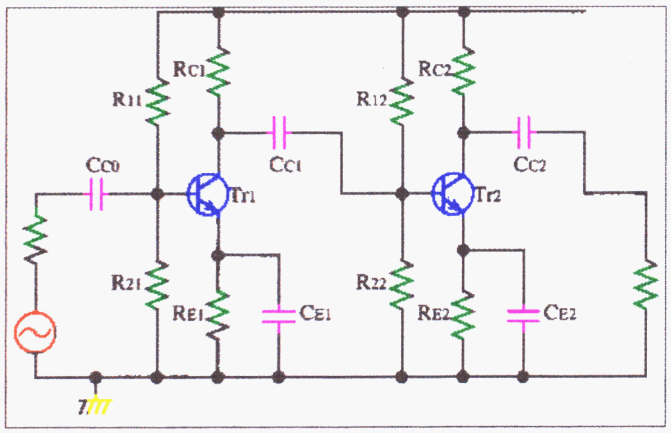
楽譜の場合には記述上の制約条件が比較的厳しいため，領域知識を導入した意味解析は行うことで誤って生成された要素仮説の棄却は容易だと思われる．2章で述べたシステムでは，前処理として5線の消去を行った．しかし，ここの実験ではそのような特殊な前処理は行わず，5線が存在する状況下で認識を行った．その状況においてもシンボルの認識が可能であったことは評価できる．



(a) 入力図面画像



(b) 生成された注意仮説

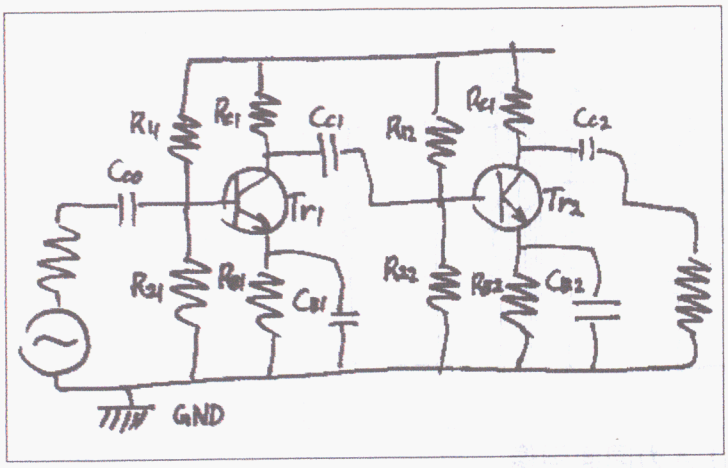


(c) 生成された要素仮説

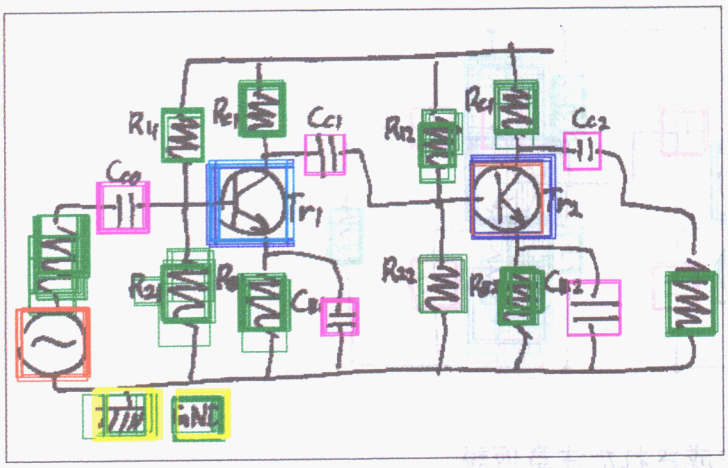
凡例

- NPN型トランジスタ
- PNP型トランジスタ
- 交流電源
- 抵抗器
- コンデンサ
- アース

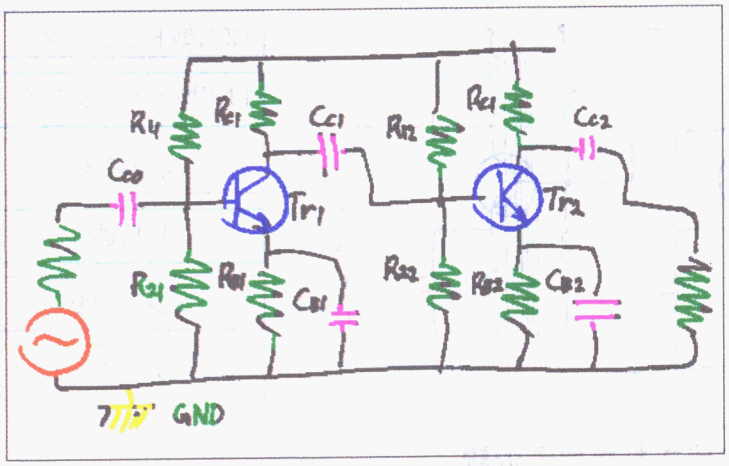
図 6-25　印刷図面に対する結果



(a) 入力図面画像



(b) 生成された注意仮説

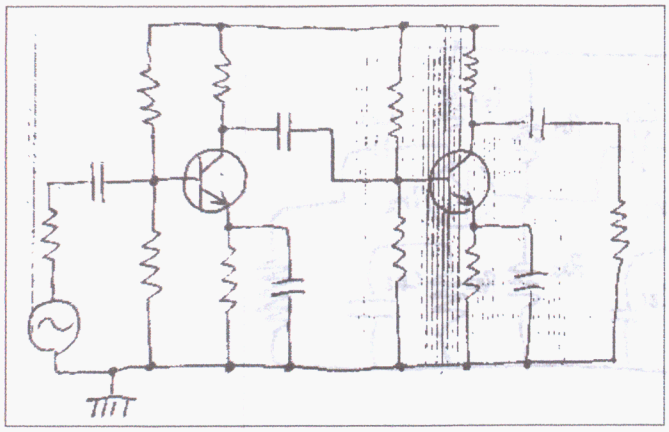


(c) 生成された要素仮説

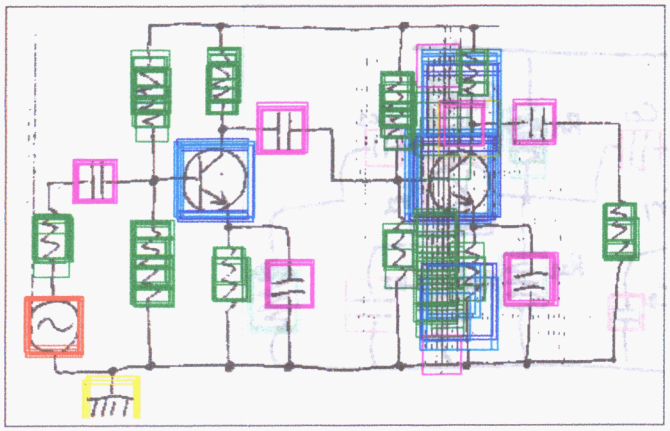
凡例

- NPN型トランジスタ
- PNP型トランジスタ
- 交流電源
- 抵抗器
- コンデンサ
- アース

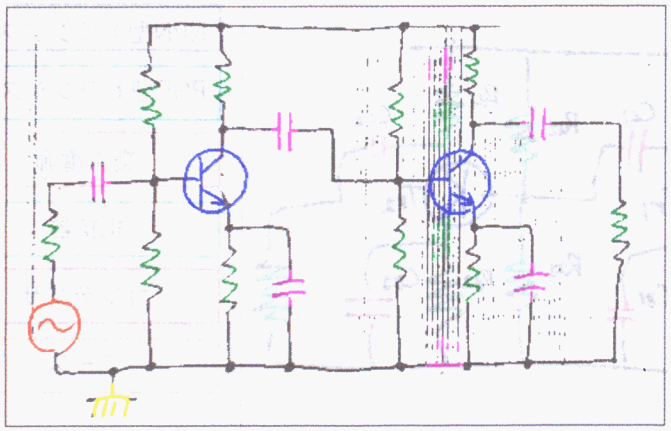
図 6-26　手書き図面に対する結果



(a) 入力図面画像



(b) 生成された注意仮説

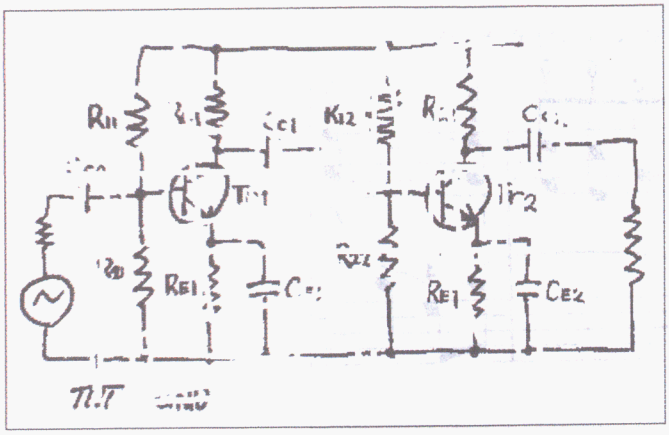


凡例

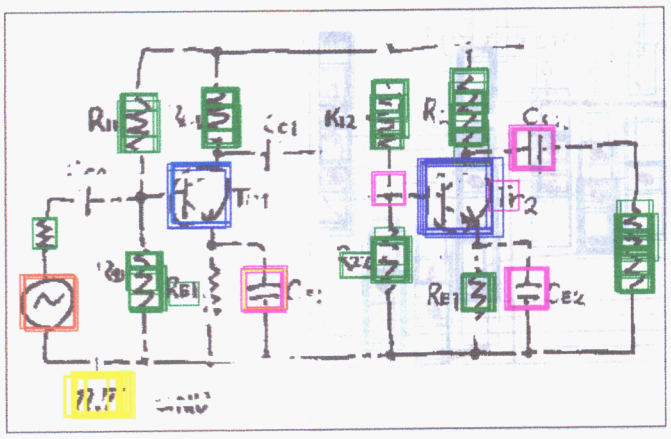
- NPN型トランジスタ
- PNP型トランジスタ
- 交流電源
- 抵抗器
- コンデンサ
- アース

(c) 生成された要素仮説

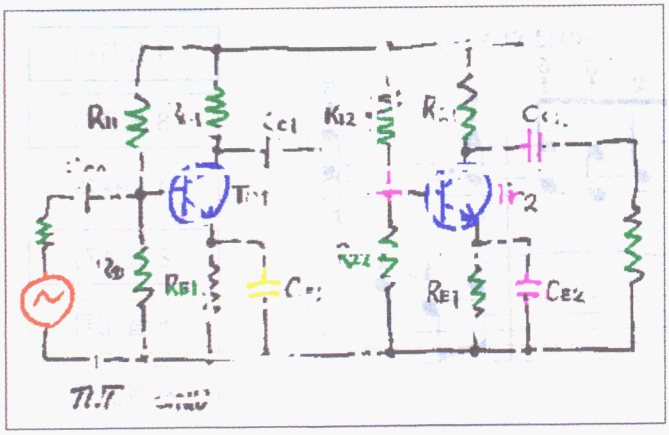
図6-27　ノイズを含む手書き図面に対する結果



(a) 入力図面画像



(b) 生成された注意仮説

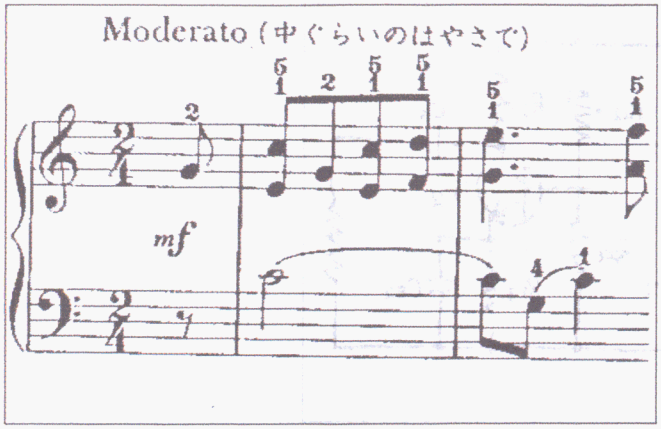


(c) 生成された要素仮説

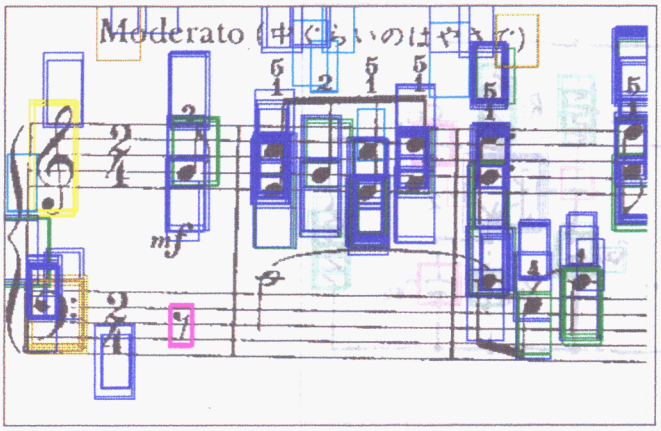
凡例

- NPN型トランジスタ
- PNP型トランジスタ
- 交流電源
- 抵抗器
- コンデンサ
- アース

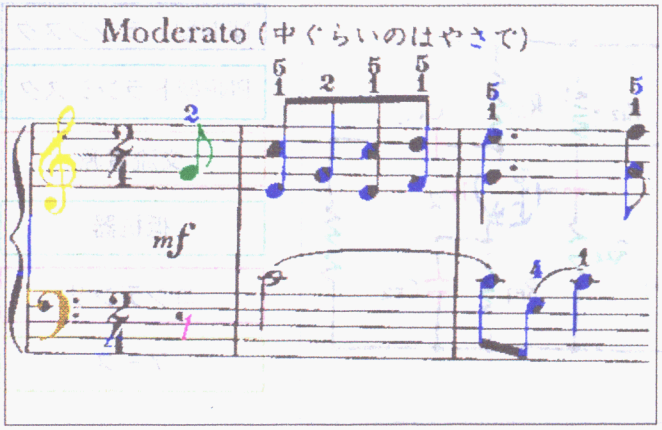
図 6-28　かすれた手書き図面に対する結果



(a) 入力図面画像



(b) 生成された注意仮説



(c) 生成された要素仮説

凡例

- 2分音符
- 4分音符
- 8分音符
- 4分休符
- 8分休符
- ト音記号
- ヘ音記号

図6-29　楽譜に対する結果

6.7.4 検討

1) 注意仮説検証時の形状評価について

現在は注意仮説検証時の形状評価は，その形状を切り出した際の弾性モデルの歪み率を用いているが，他の評価量の導入が考えられる．5章で述べたように，この切り出し手法には記号処理との融合による多くの評価量を導入することができる．そのためには，形状モデルへの属性情報の追加とそれらを有効に利用できる認識戦略の自動生成法を再検討する必要がある．しかし，これは純粋な記号処理の問題であり，既存の研究成果を用いることで可能と思われる．

2) シンボル以外の図面要素の認識について

今回はシンボルについてのみ注意制御機構に基づく認識処理を考えたが，シンボル以外の図面要素でも原理的には同様な認識処理を適用できる．

文字についてはモデルの数が著しく増大するのは避けられないが，基本的な処理はシンボルと同じでよい．意味解析では文字列の解析を行う．しかし現実問題として注意仮説の生成処理に並列処理の導入が必要となる．

線については，線を区分的に捕らえてシンボルの一種としてモデル化し，シンボルとして仮説の生成を行うことが考えられる．意味解析においてそれらの仮説の統合を行い，線として認識することが可能である．シンボルと同様ノイズ等の影響を受けにくい認識処理の実現が可能であると思われる．楽譜における五線も同様に区分的にシンボルと見なして図面要素として抽出が可能であると思われる．

3) 部分形状からのシンボル形状の統合

現在の処理では，生成された注意仮説に基づき，その部分形状を検証することで認識処理を行っている．これは注意仮説そのものが各シンボルの全体形状に起因して生成されていることにもよる．別の方法として部分形状に起因して，上記の方法とは逆に部分形状からシンボル全体形状を推定し，その検証を行うことで認識するという方法も考えられる．この方法は処理対象となるモデル数が多くなるにつれて有効性を増すと考えられる．

6.8 まとめ

この章では，パターン的知識と記号的知識という2種類の知識を有効に利用した図面解釈手法について提案した．図面解釈においては知識の積極的な利用が有効ではあるが，従来の記号処理における知識の利用だけではその有効性が発揮されない．パターン処理においても知識を利用できるような枠組みを考え，その上で図面解釈としての方法論を構成しなければならない．

弾性モデルによる切り出し手法は，パターン処理における知識の有効利用を可能とする．この手法を採用した図面解釈の提案した．記号処理により自動生成された認識戦略に沿って，この弾性モデルによる切り出しに基づく検証処理が実行される．

解釈手法の全体的な構成としては，人間の注意制御機構を模倣した．テンプレート・マッチングに基づく手法が前向き推論的にシンボル仮説を生成する．また，意味解析処理によっても後ろ向き推論的にシンボル仮説が生成される．これらシンボル仮説を注意仮説と呼んだ．この

注意仮説に対し，その認識戦略が生成される．

実験においては，意味解析処理が不完全ではあるが，様々な画質の図面に対しても全くパラメータなどの調整なしで，安定したシンボルの認識が行えた．また，電気回路図面，楽譜という異なった種類の図面に対しても，モデルを交換するだけで処理の変更なしに対応できた．本手法が画質の点においても，種類の点においても汎用性があることを示した．

後ろ向き推論におけるシンボルの検証処理においてこの手法の有効性が最も発揮できる．3章で述べたシステムは記号処理部の充実を図ったシステムなので，そのシステムとの融合により図面解釈システムとしての完成度を高め，その上でこの手法の有効性を検討しなければならない．これが，今後すべきこととして残されている．

参考文献

[Amari88] 甘利俊一: ニューロコンピューティングの数学的基礎，コンピュートロール，24, pp.2-14 (1988)

[Fukushima89] 福嶋邦彦：神経回路と情報処理，朝倉書店，p.128 (1989)

[Funakubo91] 舟久保登：パターン認識，共立出版，p.23 (1991)

[Inui93] 乾敏郎：脳と視覚，サイエンス社，p.209 (1993)

[Osawa89] 大沢，山川他：図面の認識と理解，昭晃堂，pp.94-95 (1989)

[Shirai87] 白井良明 (編)：パターン認識，オーム社，p.2 (1987)

7章　結　論

本論文では，図面解釈問題において，画質の劣る図面への対応，多種類の図面への対応，という2つの課題を実現すべき方法論の確立を目指してきた研究をまとめたものである．図面解釈のみならず，画像認識・理解の研究に共通して，S/N比の悪い入力への対応や汎用性のある方法論に対する研究が行われている．このような大きな問題領域での基本的な検討項目に対し，図面というように入力を限定することにより，その問題の本質を解明し，それを打開する方法論の提案を試みたものである．

知識の積極的な利用は，以前からその重要性は強調されている．2章で述べた楽譜認識システムは，まさにこの考え方に従い設計された．ピアノ楽譜という非常に限定された対象を取り扱うために，利用可能な知識が明確になり，処理への知識の導入が容易であった．ここでは，知識の積極的な利用が図面解釈において有効であることを示すことが目的とされ，図面解釈として一般性のあるの方法論の提案はしていない．そのため，知識は手続きとして処理の中に埋め込まれ，その開発コストも非常に大きかった．しかし，実行性能としては，結果自体，処理速度の点において満足できるものであり，図面解釈における知識の積極的な利用の有効性が示せた．

汎用性を目指す上で，多種類の図面への対応や同一種類の図面でもシンボルの追加性などを容易とするためには，処理と知識の分離が必要となる．また，画質の劣化した図面を対象とするためには，知識の積極的な利用と共にパターン処理における問題点の解決を行わなければならない．つまり，画質劣化の直接的な影響を受ける切り出し処理に対する工夫である．3章で提案した手法は，知識と処理の分離を行った．そのためにまず，利用可能な知識に対する検討を行い，7種類に整理した．また，処理の中での利用においては，制約を伴う知識と制約とならない知識に分けた．また，画質の劣化した図面に対処するための工夫として，パターン処理に対しての積極的な知識の利用を諦め，多くの可能性を残した仮説を汎用的な手法で生成し，記号処理において対処することを考えた．実験では，画質の劣化により意味的に曖昧さを含む図面に対しても良好な結果が得られた．しかし，これら実験が成功したのは，図面の規模が小さな場合のみであり，大きな図面においては生成される仮説が膨大になり，記号処理において組み合わせ爆発が生じ，動作不能となることが多かった．この原因は，切り出し処理の問題点を記号処理により解決するという方針自体にあった．しかし，記号処理部における処理と知識を分離した汎用的な図面解釈機構自体の有効性は示すことができた．

4章では，高性能な切り出し処理に関して検討を行い，弾性モデルに基づく新たな切り出し手法を提案した．高性能な切り出しを行うためには，やはり知識を積極的に利用し，対象の特性に応じた処理をする必要があると考えた．2値画像パターンという対象の形状モデルを知識として利用し，入力画像と形状モデルの対応付けを行うということを考えた．弾性モデルを用いた力学的運動シミュレーションによる対応付けを行い，その結果を使用した切り出しのアルゴリズムを提案した．人間の手書きによる歪みやノイズの混入やかすれを含む入力画像に対して実験を行い，その性能を実証した．

この弾性モデルによる切り出し手法は，単に高精度な切り出しを可能にするのみでなく，パターンの構造解析において非常に重要な役割を果たすことができる．パターンの構造解析は，対象とするパターン毎にそのパターンの特性に応じた処理を行うための方法論であるが，処理と知識の分離を考えた場合，具体的な構造解析の処理自体は事前に決まっていないので，様々な要求に対応できる汎用的な構造解析の処理の枠組みを準備しておかなければならない．また，その処理に適したパターン構造の表現法に関しても検討しなければならない．5章はこのような問題に弾性モデルによる切り出し手法が非常に有効に機能することを示し，そのためのアルゴリズムの改良と知識表現法に関して述べた．パターンの構造を記号のみで表現するのではなく，パターンによる表現と融合した．パターンによる形状表現能力の高さと記号による論理表現能力の高さが有効に活用できる汎用的な構造解析処理が実現できた．

このような処理を導入することにより，先に述べたパターン処理における問題点を解決できる．6章では，この弾性モデルによる切り出しと構造解析を利用した新たな図面解釈の方法論に関して提案した．人間の注意制御機構を模倣し，パターン処理において前向き推論的に注意仮説を生成する．また，意味解析においても後ろ向き推論的に注意仮説を生成する．それら注意仮説の検証を行うために，記号処理部において動的にその検証処理プラン，つまり認識戦略を生成し，それに従った認識処理をパターン処理部において弾性モデルに基づく汎用的構造解析処理が行う．図面解釈における構成要素の認識処理が，前向き推論と後ろ向き推論の両者においてパターン処理と記号処理を境界を越えて実現されている．実験により，この手法が画質の点においても，種類の点においても汎用性があることを示した．

しかし，この図面解釈の方法論においても，まだ問題は残る．人間の認識・解釈を考えたとき，それが実生活において常に正しく行われている保証はない．人間が視覚を用いるのにはその目的があり，その目的が達成できればよい．その時，視覚に対し高精度な結果を要求する場合もあるであろうが，低い精度でもよい場合もある．つまり，人間の視覚はそれ単体での性能評価はあまり重要ではない．一般的視覚システムを追求するコンピュータ・ビジョンは視覚システムとして独立したシステム単体においての高性能化を目指しているが，特定目的での視覚システムを追求するマシン・ビジョンの複合体として考えるべきと思う．

図面解釈においても，その図面の解釈結果を何に利用するのかという1つ上の階層の要求が定義されてこそ，評価が出来るものと考える．通常は，暗黙のうちにその目的が設計者によりシステムに埋め込まれるが，より汎用的なシステムを目指すためには，そのような目的に応じて，出力すべき情報の構造が最適化されたり，評価基準が動的に設定され，その評価基準を最もよく満たす処理を自動生成するような技術も必要になると思われる．例えば，地形状況が知りたい場合と道路網を知りたい場合で，地図という同一の入力からでも出力すべき情報構造をことなるが，それに応じて処理も最適化されるようなシステムや，認識率優先とか認識速度優先とかに応じて処理形態が自動的に変わるようなシステムに関しても検討する必要が出てくるであろう．

以上のような本論文での実験・検討結果より，画像認識・理解の研究に対して以下の事項の重要性を主張することができる．

1) 知識の積極的な利用，処理と知識の分離は重要であるが，その検討範囲を記号処理の枠に留めるべきでなく，問題固有の知識を積極的に利用できる汎用的パターン処理に関して検討すべきである．

2) 解釈や理解を行うシステムにおいては，その部分問題として様々な認識処理が発生する．その部分問題を汎用的な枠組みで解くのではなく，問題自体を単純化すること，つまり最適なパターン認識問題の動的設定のための方法論を検討すべきである．

本論文の直接的な成果として，パターン認識，画像理解の研究分野に対しての貢献度は高いものとはいいがたい．しかし，このような主張が今後この分野の発展のきっかけになれば幸いである．

謝　辞

本論文は，著者が１９８５年に大阪大学基礎工学部制御工学科の卒業研究として取り組んで以来今日まで，井口征士教授のご指導のもとに行ってきた研究をまとめたものです．井口征士教授には学生として４年間，助手として６年間という長期にわたり，パターン認識，画像処理研究のいろはから最先端の内容まで，適切かつ細やかなご指導を賜りました．決して順調には進まなかった本研究課題に対し，途中で見捨てることなく，辛抱強くご討議いただきましたことが本論文の完成に繋がりました．また，広い分野にわたる多くの優秀な研究者と自由に討論ができる機会を数多く与えて下さいましたこと，常に最先端の実験環境を整備して頂きましたことが，著者の発想に大きな影響を与えました．心から感謝の意を表します．

谷内田正彦教授には，本研究の遂行にあたり，多くの貴重なご示唆を賜りました．また，ロボット・ビジョン研究において世界的に著名な研究者である先生が，駆け出しの研究者である著者に対しても，常に笑顔で朗らかに接して頂きましたことが，著者の研究活動の励みとなりました．心からお礼申し上げます．また，本論文をまとめるにあたり，詳細なご検討を賜りました西田正吾教授に厚くお礼申し上げます．

本研究の遂行にあたり，数々の有益なご討議を頂きました和歌山大学沼田卓久教授，奈良先端科学技術大学院大学今井正和助教授，佐藤宏介助教授に厚くお礼申し上げます．

大阪大学基礎工学部在学以来，学生として，また大学教官としての研究活動に様々なご指導，ご鞭撻を賜りました桜井良文教授（現大阪工業大学），白江公輔教授（現，摂南大学），坂和愛幸教授（現，近畿大学），辻三郎教授（現，和歌山大学），須田信英教授，田村坦之教授，笠井健教授，木村英紀教授，藤井隆雄教授に厚くお礼申し上げます．

大阪大学知識科学研究会での先輩，同僚諸氏との討論は，本研究の基本的な構想を考えるにあたり，非常に有意義でした．深く感謝いたします．

原潤一氏（現，リコー），多羅尾雅之氏（現，コナミ），北村聡氏（現，ＪＲ西日本），八田曜尚氏（現，ファナック），内山浩一氏（現，野村総研），大阪産業大学熊谷聡子氏には，本研究の共同研究者として多大なる協力を頂きました．また，井口研究室のヒューマン・インタフェース研究グループの村井稔弥氏，金原玄氏，蔦田広幸氏，金井良充氏，木村朝子氏には，論文作成にあたって，献身的なご協力を頂きました．ここに感謝の意を表します．また，真鍋佳嗣助手，山口証助手をはじめとする井口研究室のスタッフ，学生の温かいご援助，ご協力に深く感謝いたします．

最後に，このような研究活動に従事する機会を与えて下さった上，優しく見守って頂きました両親と，研究生活，家庭生活の両面において様々な援助をして頂きました弟に，心から感謝の意を表します．

本論文に関連した発表

[学術論文]

1. 加藤博一, 井口征士: 小節単位処理に基づいたピアノ楽譜の自動認識, 信学論文誌, J71-D, 5, pp.894-901 (1988)
2. 原潤一, 加藤博一, 井口征士: だ円膨張法を用いた動物シルエット画像の記述と認識, 信学論文誌, J74-D2, 3, pp.366-375 (1991)
3. 北村聡, 加藤博一, 井口征士: 図面理解におけるシンボル認識戦略の動的生成, 情処論文誌 (査読中)
4. 加藤博一, 井口征士: 弾性モデルに基づく形状切り出しと構造解析, 情処論文誌 (投稿準備中)

[著書]

1. 加藤博一, 井口征士: あいまいさを考慮した手書き論理回路図面の認識, 北村, 片井 (編): 画像と制御 ーニューロ・ファジィ・ＡＩによるアプローチー, 朝倉書店, pp.123-137 (1991)
2. Hirokazu Kato and Seiji Inokuchi : A Recognition System for Printed Piano Music using Musical Knowledge and Constraints, Baird, Bunke, Yamamoto (Eds.) : Structured Document Image Analysis, Springer-Verlag, pp.435-455 (1992)

[国際会議]

1. Haruhiro Katayose, Hirokazu Kato, Masakazu Imai and Seiji Inokuchi : An Approach to an Artificial Music Expert, proc. of '89 Int Computer Music Conf., pp.139-146 (1988)
2. Hirokazu Kato and Seiji Inokuchi : The Recognition Method for Roughly Hand-Drawn Logical Diagrams Based on Hybrid Utilization of Multi-Layered Knowledge, proc. of 10th International Conference on Pattern Recognition, 578-582 (1990)
3. Hirokazu Kato and Seiji Inokuchi : A Recognition System for Printed Piano Music Using Musical Knowledge and Constraints, pre-proc. SSPR90, pp.231-248 (1990)
4. Hirokazu Kato, Satoshi Kitamura and Seiji Inokuchi : Symbol Recognition Based on Attention Control Mechanism in Drawing Interpretation, 2nd ACCV (accepted)

[口頭発表]

1. 加藤博一, 広野淳之, 井口征士: 規則性を重視した楽譜認識システム, 信学全大, 6, p.245 (1986)
2. 加藤博一, 井口征士: 音楽理解システムにおけるピアノ楽譜の自動認識, 情処全大, pp.1475-1476 (1986)
3. 今井正和, 加藤博一, 井口征士: 音楽分析のためのソフトウェアツール, 情処全大, pp.2781-2782 (1987)

4. 加藤博一,井口征士:曖昧さを考慮した手書き論理回路図面の自動認識,信学全大,1,p.319 (1988)
5. 加藤博一,井口征士:曖昧さを考慮した雑な手書き論理回路図面の認識,シンポジウム「あいまい情報処理と知的システム制御」,pp.97-102 (1988)
6. 加藤博一,井口征士:意味解析を利用した雑な論理回路図面の認識,信学春大, (1989)
7. 加藤博一,井口征士:知識を利用した雑な図面の認識手法について,シンポジウム「知的システム構築の実践と将来」,pp.121-124 (1989)
8. 多羅尾雅之,加藤博一,井口征士:ニューラルネットを利用した図形パタンの概念学習,画像と認識・理解シンポジウム,1,pp.57-64 (1992)
9. 内山浩一,加藤博一,井口征士:パターンマッチングによる二値画像中のシンボルの切り出し,信学春大,7,p.203 (1994)
10. 北村聡,加藤博一,井口征士:相補的モデルを利用した注意制御による図面理解,信学春大,7,p.208 (1994)
11. 多羅尾雅之,加藤博一,井口征士:パターン認識における特徴利用方法についての検討,通学春大,7,p.222 (1994)
12. 北村聡,加藤博一,井口征士:注意制御による図面理解,画像の認識・理解シンポジウム,2,pp.17-24 (1994)
13. 熊谷聡子,加藤博一,井口征士:弾性モデルと論理記述を用いたシンボル認識,信学総大,7,p.309 (1995)
14. 北村聡,加藤博一,井口征士:,初期仮説に基づく認識戦略の動的生成,信学総大,7,p.315 (1995)
15. 金原玄,加藤博一,井口征士:統計的手法を用いた犬と猫の顔画像の判別,シ制情学会講演会,pp.301-302 (1995)

その他の発表

[学術論文]

1. Toshinori Yamasaki, Masahiro Yamamoto, Hirokazu Kato and Seiji Inokuchi : Knowledge Based Training System for Handwriting of Japanese Characters, Trans. IEICE, E73, 3, pp.341-348 (1990)
3. 山本勉,加藤博一,佐藤宏介,井口征士:機能モデルを用いた3次元物体認識,信学論文誌,J74-D2, 5, pp.601-609 (1991)
4. Yasuo Sakaguchi, Hirokazu Kato, Kosuke Sato and Seiji Inokuchi : Acquisition of Entire Surface Data Based on Fusion of Range Data, Trans. IEICE, E74, 10, pp.3417-3422 (1991)
5. 橋本理,加藤博一,森脇耕介,佐藤宏介,井口征士:カラー画像を用いた材質感の判別と相互反射の検出,信学論文誌,J74-D2, 6, pp.727-735 (1991)
6. 汪増福,加藤博一,佐藤宏介,井口征士:鏡面物体の3次元形状計測,信学論文誌,J75-D2, 7, pp.1177-1186 (1992)

7. 汪増福，加藤博一，佐藤宏介，井口征士：単眼観測画像による鏡面形状の復元・推定，信学論文誌，J76-D2, 3, pp.575-585 (1993)
8. 和気早苗，加藤博一，才脇直樹，井口征士：テンション・パラメータを用いた協調型自動演奏システムーＪＡＳＰＥＲー，情処論文誌，35, 7, pp.1469-1481 (1994)

[国際会議]

1. Yasuo Sakaguchi, Hirokazu Kato, Kosuke Sato and Seiji Inokuchi : An Integration of Measured 3-D Patch for CAD Modeling, proc. of JAPAN-USA Sympo. on Flexible Automation, pp.1213-1216 (1990)
2. Yasuo Sakaguchi, Hirokazu Kato, Kosuke Sato and Seiji Inokuchi : Generation of 3-D Models Based on Image Fusion of Range Data, Proc. of MVA'90, (1990)
3. Tomoharu Maeda, Hirokazu Kato and Seiji Inokuchi : Image Fusion System for Object Tracking, Proc. of Japan-USA Sympo. on Flexible Automation, pp.365-368 (1992)
4. Hirokazu Kato, Sanae Wake and Seiji Inokuchi : Cooperative Musical Partner System : JASPER (Jam Session Partner) , Advances in Human Factors/Ergonomics, Vol.19B Human-Computer Interaction : Software and Hardware Interfaces, pp.250-255 (1993)
5. Kosuke Sato, Tomoharu Maeda, Hirokazu Kato and Seiji Inokuchi : CAD-Based Object Tracking with Distributed Monocular Camaras for Scurity Monitoring, Proc. of 2nd CAD-Based Vision Workshop, pp.888-899 (1994)

[口頭発表]

1. 白沢寿夫，辻本敬一，加藤博一，井口征士，山崎敏範：オンライン文字認識手法を用いた書写学習システム，第１回ヒューマン・インタフェース・シンポジウム，425 (1985)
2. 山崎敏範，山本雅弘，加藤博一，井口征士：書写技能知能を組み込んだＣＡＩシステム，第５回ヒューマン・インタフェース・シンポジウム，pp.141-146 (1989)
3. Apivat Vacheravothan, 加藤博一，佐藤宏介，井口征士：距離画像計測のためのカラー光による空間コード生成法，信学春大，D-3 (1989)
4. Kaset Sirisantisamrid, 加藤博一，佐藤宏介，井口征士：トランスピュータによる並列処理システムの試作と信号処理への応用，信学春大，D-585 (1989)
5. 加藤博一，坂口康雄，山本勉，佐藤宏介，井口征士：距離画像と濃淡画像を併用したビンピッキング，インテリジェントFAシンポジウム，pp.69-71 (1989)
6. 汪増福，加藤博一，佐藤宏介，井口征士：鏡面物体の３次元形状計測に関する一考察，自動制御連合講演会，pp.533-534 (1990)
7. Kaset Sirisantisamrid, 汪増福，加藤博一，井口征士：構造可変型並列処理システム及びその応用，信学春大 (1990)
8. 橋本理，森脇耕介，加藤博一，佐藤宏介，井口征士：カラー画像・距離画像からの質感（反射特性）抽出に関する基礎研究，センシングフォーラム (1990)
9. 橋本理，加藤博一，佐藤宏介，井口征士：イメージフュージョンによる物体の質感認識ー

カラー画像と距離画像の融合ー, シ制情講演会, pp.205-206 (1990)
10. 真鍋佳嗣, 加藤博一, 佐藤宏介, 井口征士: 距離画像・カラー画像のイメージフュージョン, SICE パターン計測部会研究会 (1990)
11. 内山晋二, 加藤博一, 井口征士: 距離画像計測による足の3次元形状モデリングー人体の動的形状モデルの構築ー, シ制情講演会, pp.227-228 (1990)
12. 内山晋二, 加藤博一, 佐藤宏介, 井口征士: 距離画像計測による足の3次元形状モデリングー布状ばねモデルによる皮膚表現ー, 自動制御連合講演会 pp.491-492 (1990)
13. 汪増福, 加藤博一, 佐藤宏介, 井口征士: 鏡面物体の3次元形状計測, 信学春大, D-641 (1991)
14. 真鍋佳嗣, 佐藤宏介, 加藤博一, 井口征士: カラー画像からの相互反射成分の抽出, 信学春大, D-683 (1991)
15. 内山晋二, 加藤博一, 佐藤宏介, 井口征士: 距離画像計測による人体の3次元形状モデリング, 信学春大 (1991)
16. 前田智治, 加藤博一, 井口征士: 複数台のカメラによる移動物体の追跡, シ制情講演会, pp.349-350 (1991)
17. 内山晋二, 加藤博一, 佐藤宏介, 井口征士: 距離画像による人体の3次元形状の計測とモデリング, 情処研資, 91-CV-74, pp.1-9 (1991)
18. 前田智治, 加藤博一, 井口征士: イメージ・フュージョンによる動物体の追跡, 信学春大, 7, p.411 (1992)
19. 汪増福, 加藤博一, 佐藤宏介, 井口征士: アクティブステレオに基づく鏡面の形状抽出, 信学春大, 7, p.399 (1992)
20. 和気早苗, 加藤博一, 井口征士: JASPER ー協調型演奏システムー, 信学春大, 1, 307 (1992)
21. Zengfu Wang, Hirokazu Kato, Kosuke Sato and Seiji Inokuchi : Obtaining shape of specular surfaces from monocular observing image, Proc. of 31sr SICE Annual Conf., pp.1285-1288 (1992)
22. 和気早苗, 加藤博一, 才脇直樹, 井口征士: 演奏者の感情を考慮した協調型演奏システム---Jasper---, 音楽情報科学研究会会報, pp.43-46 (1992)
23. 村井稔弥, 加藤博一, 井口征士: テレビ会議システムにおける対話者識別法についての検討, SICE Human Interface N&R, pp.177-182 (1995)
24. 盧存偉, 張　逸芳, 加藤　博一, 井口　征士: 大比重の泥水で満ちた立坑形状の超音波計測, シ制情講演会, pp.325-326 (1995)
25. 千葉雅裕, 加藤博一, 井口征士: 格子投影による三次元計測の高精度化, シ制情講演会, pp.407-408 (1995)
26. 蔦田広幸, 加藤博一, 井口征士: 生理指標に基づく意外感と覚醒水準／精神的負担との関係, 第11回ヒューマン・インタフェース・シンポジウム, pp.219-224 (1995)

[その他]

1. 加藤博一, 井口征士: 画像処理の基礎, 調理科学, 24, 1, pp.62-66 (1991)