

Title	多重解像度モデルと仮説ネットワーク上での並列空間推論による視覚情報処理
Author(s)	佐藤, 嘉伸
Citation	大阪大学, 1988, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/1192
rights	
Note	

Osaka University Knowledge Archive : OUKA

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

Osaka University

多重解像度モデルと
仮説ネットワーク上での並列空間推論
による視覚情報処理

1988年 2月

大阪大学大学院基礎工学研究科

佐藤 嘉伸

多重解像度モデルと
仮説ネットワーク上での並列空間推論
による視覚情報処理

1988年 2月

大阪大学大学院基礎工学研究科

佐藤 嘉伸

内容梗概

本論文は、筆者が大阪大学大学院基礎工学研究科（物理系専攻情報工学分野）在学中に行ったコンピュータビジョンに関する研究をまとめたものである。本研究では、視覚情報処理の問題として、画像フィルタリング、両眼立体視といった初期処理(early process)から物体認識といった高次処理(high-level process)までを扱っているが、表現と制御の手法として、特に、多重解像度モデルと仮説ネットワーク上での並列空間推論に焦点を当てている。

1970年代以来、ピラミッド構造を基にした階層的処理、及び、ネットワーク構造を基にした協調的処理(cooperative process)は、コンピュータビジョンにおいて代表的な表現と制御の手法として、Rosenfeldに代表される工学的な立場からも、Marrに代表される視覚の計算モデルの立場からも注目され続けてきた。これは、解像度軸に沿った階層的表現と制御、及び、ネットワーク構造による表現と制御が、階層性と空間的な広がりを持つ視覚データの処理に本質的なものであるということに起因しており、現在でもスケールスペース、及び、神経ネットワークといった最もホットな分野を形成している。

本論文は、それらに関する3つの話題を含んでいる。1つめは、多重解像度モデルとして性質のよいガウスピラミッドを効率よく、かつ、精度よく構成する方法に関するもので、2つめは、ガウスピラミッドを利用した多重解像度構造とネットワーク構造を結合した両眼立体視モデルに関するもので、3つめは、仮説ネットワーク上での並列空間推論によるパラメータ化3次元物体の認識に関するものである。

第1章では、コンピュータビジョン研究の流れの中で、本研究を位置づけその意義を明確にする。

第2章では、複数の解像度（標準偏差）のガウスフィルタリング、及び、その微分を、ガウス関数の性質を生かした再帰的な構造のフィルタで高速に精度よく行なう手法について述べる。ここでは、過去において提案されたガウス関数のたたみ込みの繰り返しによる方法を一般化し、微分、及び、スケールファクタ（標準偏差の拡大比）の可変性を許して再定式化を行い、十分な精度を得るための重み係数の設定条件を明確にすることを試みる。まず、連続系におい

て、指定したスケールファクタでのガウシアンフィルタおよびその微分フィルタ系列を得るための定式化を行う。次に、これを基に離散系において、同一の重み係数でのたたみ込みの繰り返しによるフィルタ系列の生成式を導く。そして、離散化に伴う誤差の発生する要因を解析し、その誤差を十分小さく抑える条件を明確化する。また、この議論は、Hartmannによつて提案された網膜受容野の再帰的構造モデルを理論的に説明するものであり、人間の視覚モデルの立場からも興味深いものである。

第3章では、複数の表面が互いに交錯しているような不連続の多いシーンに対しても効果的に働く両眼立体視モデルについて述べる。Prazdnyは、視差の不連続をペナルティとせず、類似性による促進作用に基づき曖昧さ解消を行うことで、不連続に対処できると指摘したが、基本的にシーンの制約を弱くする分だけ曖昧さ解消能力の低下を招くという問題点がある。ここでは、様々な観点からみた類似性の促進を最大にする視差場を選択するため複数の対応規則と対応単位を組み合わせ、これらをHopfield型神経ネットワークのエネルギ最小化モデルに統合することにより、強力な曖昧さ解消能力を得ている。特に、複数解像度レベルの結合に対して、各レベルが並列に処理を行いながら、同一の物理現象に対応する画像特徴のみに対し、粗レベルから密レベルに情報を伝播することで、不連続に強い処理を実現している。これらの考察から、Marr-Poggioステレオ第1アルゴリズムの自然な拡張が導かれる。フェンス越しのシーンなどに対する実験結果によつて、ネットワークに埋め込まれた各相互作用が有効に働いていることを示す。

第4章では、単眼画像からパラメータ化3次元物体の認識を行なう手法について述べる。物体モデルのパラメータ化を許すことで、解の拘束条件が弱まり、それだけ強力な探索機構を必要とする。その解決は、逐次的な探索法、或は、Hough変換等では難しいと考えられる。ここでは、並列認識に適したパラメータ化3次元物体モデルの表現方法について述べ、Hopfield型神経ネットワークモデルに仮説間の興奮と抑制の相互作用を埋め込む。認識過程は、3段階からなる。

- (1)画像特徴とモデルプリミティブのすべての可能な対応により仮説生成を行なう。
- (2)物体モデルと画像生成過程の拘束を利用して仮説間の相互作用を行ない、画像特徴と物体モデルの対応関係を決定する。
- (3)決定した対応関係より変換パラメータとモデルパラメータを推定する。

実験結果によって、雑音、隠れの多い状況においても安定に認識が行えることを示す。

第5章では、本研究のまとめを行ない、さらに、今後の問題点について論じる。

多重解像度モデルと仮説ネットワーク上での並列空間推論
による視覚情報処理

目 次

第1章	序論	1
第2章	多重解像度ガウスフィルタの計算機構	5
第1節	序言	5
第2節	連続系のカスケードガウスフィルタ	6
第3節	離散系のカスケードガウスフィルタ	9
第4節	離散化に伴う誤差の解析	13
第5節	関連研究	24
第6節	結言	26
第3章	多重解像度両眼立体視モデル	27
第1節	序言	27
第2節	基本方針	30
第3節	対応仮説の生成	36
第4節	仮説ネットワークによる対応づけ	39
第5節	実験結果	46
第6節	結言	55

第4章	パラメータ化モデルを用いた三次元物体認識	57
第1節	序言	57
第2節	並列認識のための物体モデル表現	59
第3節	認識過程の概要	67
第4節	対応仮説の生成	69
第5節	仮説ネットワークによる物体の分離と認識	70
第6節	変換パラメータとモデルパラメータの推定	76
第7節	実験結果	77
第8節	結言	81
第5章	結論	83
謝辞		85
参考文献		87
付録		93
1.	エッジセグメントの傾きから視差勾配の計算	93
2.	鉛直軸回りの回転角 θ_{pw} の推定	94
3.	可変方向辺の方向の推定	97
4.	平行移動パラメータとモデルパラメータの推定式の導出	98
発表論文		101
1.	学会誌論文	
2.	国際会議発表	
3.	研究会発表	
4.	全国大会発表	

第1章 序論

コンピュータビジョン研究の目的は、画像情報を基に、ある実体が外界(3次元世界)を行動したり、或は、操作するのに必要な情報を獲得するアルゴリズムを開発し、またその開発を通して視覚機能の根本原理を解明することにある。視覚系は、表面の知覚、動きの理解、物体認識など様々な問題を解決しなければならないが、どのレベルにおいても視覚データ特有の階層性、及び、空間的広がりを利用した表現形式、制御構造が必要になる。

1970年代以来、ピラミッド構造を基にした階層的処理、及び、ネットワーク構造を基にした協調的処理(cooperative process)は、コンピュータビジョンにおいて代表的な表現と制御の手法として、Rosenfeldに代表される工学的な立場からも、Marrに代表される視覚の計算モデルの立場からも注目され続けてきた。これは、解像度軸に沿った階層的表現と制御、及び、ネットワーク構造による表現と制御が、階層性と空間的な広がりを持つ視覚データの処理に本質的なものであるということに起因しており、現在でもスケールスペース[1.1]、及び、神経ネットワーク[1.2]といった最もホットな分野を形成しており、従来のヒューリスティックな方法から、より理論的な確立が行なわれつつある。

多重解像度解析は、

1. 粗レベルでの解析が密レベルでのプランを与える(注視点の制御)。
2. シーンの事象(event)は本質的に階層構造を持ち、そのシーンを反映する画像においても、種々の解像度レベルでその情報が明確になる。

という点において、視覚情報処理において重要な役割を果たす。

多重解像度解析の研究の進展において重要な点は、解像度軸としてガウス関数の標準偏差を選ぶべきであるというMarrの提案[1.3]であると思われる。多重解像度解析の理論的研究の多くはガウス関数の種々の数学的な性質のよさに基づいている。これは、すでに20年近く前に、わが国の飯島[1.4]が提案していたことであるが、心理学的、生理学的、さらには、計算論的にガウスフィルタの妥当性を唱えたMarrの功績は大きい。

ガウスフィルタの問題点は、標準偏差が大きくなった場合の計算コストの増大である。第2章では、ガウスピラミッド、或は、ラプラシアンも含めたその線型微分ピラミッドをいかに効率よく、かつ、精度よく構成するかという問題に関して議論する。この議論は、Burtの研究[1.5]を一般化したものであるが、インプリメンテーションの面からもガウス関数のよい性質が生かされることが理論的に示され、網膜受容野の再帰的構造を示唆するものになっている。

一方、弛緩ラベリング処理(Relaxation Labelling Process, 以下R L Pと略す)は画像処理特有の曖昧さを近傍との相互作用によって解消する枠組みとして、1970年代においては、ヒューリスティックな方法[1.6]で応用されてきたが、Hummel & ZuckerがR L Pの数学的意味を明確に定式化した[1.7]。彼らはR L Pを適合係数によって定まる支持関数を極大化する過程と定式化した。これはHopfieldの神経ネットワークのエネルギ極小化モデル[1.8]と本質的に等価なものである。これらの研究は、R L Pを明確に最適化過程であると定式化し、従来のヒューリスティックな方法に比べ、適用の際の見通しをよくした。また、Hopfieldのモデルは、R L Pのようにラベルの割当の可能性を確率と考えるのではなく活性度(仮説の信頼度)と考えることによって、R L Pの問題点とされてきた“多ラベル問題”，及び，“ラベル無し問題”を自然に解決できるという利点がある。例えば、群化特徴を対応単位としたステレオ対応を考えると、一方の画像特徴が他方の画像特徴と1対1対応するとは限らず(多ラベルとなる)、また隠れのある場合は、他方の画像での対応は存在しない(ラベル無しとなる)。3次元モデルとの照合においても、すべての画像特徴がモデルの一部と対応するとは限らず(ラベル無しとなる)、これらの問題の扱いは重要になる。両眼立体視においても、物体認識においても、解決すべき問題は、画像から得られる可能な3次元空間解釈(仮説)の中で、3次元中の空間配置が最も理にかなったものを選択するという組合せ最適化問題として定式化される。このとき、空間配置の選択基準は、両眼立体視の場合は表面の性質を表わす拘束条件、物体認識の場合は物体モデルの拘束条件より得られる。第3章と第4章では、個々の仮説をユニットとして、拘束条件をユニット結合関係とする仮説ネットワーク上での並列分散処理によってこれらの問題を解決しようというものである。この仮説ネットワークはHopfieldの神経ネットワークモデルで具体化される。

第3章では、Hopfieldモデルと多重解像度モデルを結合した両眼立体視モデルについて述べる。従来の弛緩法と多重解像度モデルの結合[1.9]は、主として1段階のみで行なった場合の収束の遅さを解決するため利用されている。本研究では、処理効率の問題よりも、曖昧さ解消のための1つの要因として多重解像度結合を捉え、本質的な能力の向上を目指し、表面に不連続の多い場合にも多重解像度結合が効果的に働くことを示す。

第4章では、Hopfieldモデルを用いた単眼画像とパラメータ化3次元物体モデルとの照合について述べる。雑音・隠れが多く、かつ、複数の物体が画像中に投影されている場合を考え、対象を分離識別するようにネットワークをプログラムする。モデルはパラメータ化されているので、モデルプリミティブの空間的配置の拘束条件は弱く、解釈の可能性の適合関係に基づく仮説ネットワー

ク上での並列空間推論によって、弱い証言を並列に集約することによってロバストな処理を実現している。

第3章、第4章ともに、不連続の多い表面、或は、パラメータ化モデルといった、より一般性のある弱い拘束条件のもとでの問題解決を行っており、仮説ネットワーク上での並列集約的な処理の能力を実験的に実証するものとなっている。

第2章 多重解像度ガウスフィルタの 計算機構

第1節 序言

視覚処理および信号解析において、微分処理は非常に重要である。このとき、微分による高周波強調に伴うノイズ強調をふせぐためローパスフィルタを組み合わせることが多いが、ガウスフィルタはローパスフィルタとして良好な低域通過特性を持つ。また、多重解像度解析の観点からもガウス関数は重要で、標準偏差(σ)を解像度パラメータとする多重解像度モデル[2.1],[2.2]はよい性質を持つことが知られ[2.3]-[2.6]、複数あるいは連続した σ のガウスフィルタリングおよびその微分により得られた画像(信号)系列は、構造記述、coarse-fine戦略において重要な役割を果たす[2.7]-[2.9]。しかし、 σ が大きくなるにつれ計算コストが莫大になるという問題点がある。そこで、これを克服するための手法が開発されてきた。これらの研究は、ガウス関数のたたみ込みの繰り返しによるもの[2.10],[2.11]と、中心極限定理に基づく一様フィルタの繰り返しによるもの[2.12],[2.13]にわけられる。本章では、前者の手法により複数の σ のガウスフィルタおよびその微分フィルタを設計する場合の理論的基礎を与え、過去の研究[2.10]-[2.13]との関連について述べる。

本手法の理論的根拠となるのは、線型微分およびたたみ込み積分に関する結合性および可換性とガウス関数に関する以下に示すよく知られた性質である。

[性質1] ガウス関数同志のたたみ込みはガウス関数である。(再生性)

[性質2] ガウス関数は空間および周波数領域での2次モーメントの積を最小にする関数である。

[性質3] 多次元ガウス関数は1次元ガウス関数の積に分解できる。(可分性)

性質1は、連続系での議論の基礎であり、性質2は、離散化したときの、いかに少ない計算コストで誤差を小さくするかという兼ね合いを最適化し、性質3は、多次元フィルタを構成する際に重要になる。

本章での議論は、Burt[2.10],Crowley[2.11]の手法の一般化であるが,Castan & Shen[2.14]が試みたようにBurtの手法の理論づけというより、所望のスケールファクタ(σ の拡大比)で、所望の微分オペレータを、所望の精度で実現するための原理を明確にするものである。

第2節 連続系のカスケードガウスフィルタ

連続系において、ガウス関数のたたみ込みの繰り返しにより、指定したスケールファクタでのガウスフィルタ系列を得るための定式化を行う。

ガウス関数は、

$$g(x; \sigma) = (1/\sqrt{2\pi} \sigma) \exp(-x^2/2\sigma^2) \quad (2.1)$$

と表される。そして、ガウス関数のたたみ込みの繰り返しにより得られる系列すなわち連続系カスケードガウスフィルタ(Continued Cascaded Gaussian Filters, 以下CCGFと略す)を次式で定義する。

$$\left. \begin{aligned} h_0(x) &= g(x; \sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-t)g(t; r_n \sigma_0) dt \quad (n=1, 2, 3, \dots) \end{aligned} \right\} \quad (2.2)$$

性質1は、

$$g(x-t; \sigma_a)g(t; \sigma_b)dt = g(x; \sqrt{\sigma_a^2 + \sigma_b^2}) \quad (2.3)$$

と表されるので、式(2.2)において、次式が成立する。

$$h_n(x) = g(x; \sigma_n) \quad (2.4)$$

ただし、 $\sigma_n = \sigma_0 \sqrt{1 + \sum_{i=1}^n r_i^2}$

ここで、 $\sigma_n = s \sigma_{n-1}$ の関係を満たしたいとき(ただし、 s はスケールファクタで $s > 1$ とする)[†]、 $\sigma_n = s^n \sigma_0$ 、 $\sigma_{n-1} = s^{n-1} \sigma_0$ であるので、これらを式(2.3)より導かれる

$\sigma_{n-1}^2 + (r_n \sigma_0)^2 = \sigma_n^2$ に代入すると、

$$r_n = s^{n-1} \sqrt{s^2 - 1} \quad (2.5)$$

が得られる。よって、スケールファクタ s のCCGFは式(2.5)を式(2.2)に代入することにより得られ、次式で表される。

[†] 文献[2.6]において、解像度(σ)軸を一様な情報量を持つように標本化するためには、対数スケールでなければならないことが示されている。

$$\left. \begin{aligned} h_0(x) &= g(x; \sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-t)g(t; \sigma_0 s^{n-1} \sqrt{s^2-1}) dt \end{aligned} \right\} \quad (2.6)$$

さらに、式(2.6)において、 $t = s^{n-1}t'$ とおくと、
 $g(t; \sigma_0 s^{n-1} \sqrt{s^2-1}) = (1/s^{n-1})g(t'; \sigma_0 \sqrt{s^2-1})$ となり、これを式(2.6)に代入し、
 t' を t とおきなおすと次式が得られる。

$$\left. \begin{aligned} h_0(x) &= g(x; \sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)g(t; \sigma_0 \sqrt{s^2-1}) dt \end{aligned} \right\} \quad (2.7)$$

線型微分を考慮に入れると、式(2.7)は次式のように拡張される。

$$\left. \begin{aligned} h_0(x) &= g^{(m)}(x; \sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)g(t; \sigma_0 \sqrt{s^2-1}) dt \end{aligned} \right\} \quad (2.8)$$

ただし、 $g^{(m)}(x; \sigma) = \frac{d^m}{dx^m} g(x; \sigma)$ ($m=0, 1, 2, \dots$)である。

すなわち、 D を線型微分オペレータ、 $*$ をたたみ込み積分とすると、結合律

$$(Dg_1)*g_2 = D(g_1*g_2) \quad (2.9)$$

が成立するので、レベル0(ただし、 n を $h_n(x)$ のレベルと呼んでいる)で微分フィルタを設定することにより、各レベルで線型微分ガウスフィルタが構成される。式(2.8)が、離散系において、各レベル n ($n \geq 1$)で同一の重み係数のたたみ込みの繰り返しにより、スケールファクタ s のガウスフィルタあるいはその微分フィルタ系列を構成するための基本式となる。また、結合律およびたたみ込み積分の可換性より導かれる次の関係

$$f*(Dg) = (Df)*g = D(f*g) \quad (2.10)$$

によって、式(2.8)と1次元信号とのたたみ込みも微分と同様に、レベル0でその信号とのたたみ込みを設定することで、各レベルでその信号に対するガウスフィルタリングあるいはその微分出力が得られる。

次に、多次元CCGFを考える。等方性のぼけを持つガウスフィルタリングを仮定すると、性質3は、

$g_2(x, y; \sigma)$ を2次元ガウス関数として、

$$\frac{\partial^{m+1}}{\partial x^m \partial y} g_2(x, y; \sigma) = g^{(m)}(x; \sigma) \cdot g^{(1)}(y; \sigma) \quad (2.11)$$

と表される。また、式(2.11)を画像にたたみ込む場合、x方向とy方向の2回の1次元たたみ込みを逐次的に行うことで実現できる。よって、2次元CCGFは次式で表すことができる。

$$\left. \begin{aligned} h_0(x, y) &= \frac{\partial^{m+1}}{\partial x^m \partial y} g_2(x, y; \sigma_0) \\ &= g^{(m)}(x; \sigma_0) \cdot g^{(1)}(y; \sigma_0) \\ h_{x_n}(x, y) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t, y) g(t; \sigma_0 \sqrt{s^2-1}) dt \\ h_n(x, y) &= \int_{-\infty}^{\infty} h_{x_n}(x, y-s^{n-1}t) g(t; \sigma_0 \sqrt{s^2-1}) dt \quad (n=1, 2, 3, \dots) \end{aligned} \right\} (2.12)$$

性質3の利用により、離散化した場合、多次元化に伴う1標本点あたりの計算量の増加は、フィルタの幅をN、フィルタの次元をpとして性質3を用いない場合の $O(N^p)$ から $O(pN)$ に抑えることができる。また、レベル0で所望の線型微分オペレータを設定すればよいので、たとえば2次元ラプラシアンガウスフィルタ[2.15]を構成する場合は、レベル0において、

$$\begin{aligned} h_0(x, y) &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) g_2(x, y; \sigma_0) \\ &= g^{(2)}(x; \sigma_0) g(y; \sigma_0) + g(x; \sigma_0) g^{(2)}(y; \sigma_0) \end{aligned} \quad (2.13)$$

と設定すればよい。これを画像にたたみ込む場合、レベル0において4回の1次元たたみ込み、レベルn($n \geq 1$)においては微分オペレータの種類に依存せず、2回の1次元たたみ込みで実現できる。3次元以上の多次元CCGFも同様に構成される。

第3節 離散系のカスケードガウスフィルタ

式(2.8)の離散化に対し、2種類の標本化パターンを考える。一方は標本化密度を解像度(σ_n)に無関係に各レベルで一様としたもの(Discrete Cascaded Gaussian Filters. 以下, DCGFと略す)で、他方は、解像度に応じて標本化密度を設定するもの(以下, Reduced DCGFと呼ぶ)である。以下では、1次元でのみ定式化を行うが、第2節で述べたように多次元への拡張は、性質3の利用により容易に行える。

3.1 DCGF

1次元DCGFは次式で定義される。

$$\left. \begin{aligned} h_0(x) &= g^{(m)}(x; \sigma_0) \cdot \sum_{k=-w_1}^{w_1} \delta(x-k) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t) \{g(t; \sigma_0 \sqrt{s^2-1}) \cdot \sum_{k=-w_2}^{w_2} \delta(t-k-x_0)\} dt \end{aligned} \right\} \quad (2.14)$$

ただし、 $\delta(x)$ はディラックのデルタ関数、 w_1, w_2 はフィルタの半径を表す正整数、 x_0 は各レベル間の標本点の位置関係を調節する定数で、次レベルでの標本位置の移動量を表し、 $-1/2 \leq x_0 \leq 1/2$ とする。さらに、式(2.14)に対する離散変数 i を用いた計算式は、次式のように表される。

$$\left. \begin{aligned} h_0(i) &= \begin{cases} g^{(m)}(i; \sigma_0) & (-w_1 \leq i \leq w_1) \\ 0 & (i < -w_1, i > w_1) \end{cases} \\ h_n(i) &= \sum_{k=-w_2}^{w_2} h_{n-1}(i - \lfloor x_n \rfloor - \lfloor x_{n-1} \rfloor - s^{n-1}k) \cdot g(k-x_0; \sigma_0 \sqrt{s^2-1}) \end{aligned} \right\} \quad (2.15)$$

ただし、 $x_n = \sum_{k=1}^n x_0 s^{k-1} = x_0 (s^n - 1) / (s - 1)$ 、 $\lfloor a \rfloor$ は a を越えない最大整数を表すとする。このとき、レベル n の標本点 i における実際の空間位置 x は次式で与えられる。

$$x = i + x_n - \lfloor x_n \rfloor \quad (2.16)$$

図2.1に $s=2, x_0=1/2$ のときの計算の流れを示す。すなわち、各レベル n における計算は、標本化間隔を s^{n-1} にした同一の重み係数 $g(i-x_0; \sigma_0 \sqrt{s^2-1})$ のたたみ込みにより行われ、計算コストは各レベルにおいて一定である。なお、DCGFでは s は標本化パターンにおける各標本点間のとりうる距離に限られるので、1次元では整数しか許されない。しかし、図2.2に示すように2次元正方格子では整数以外にも $\sqrt{2}, \sqrt{5}$ など、六角格子では $\sqrt{3}, \sqrt{7}$ などのスケール

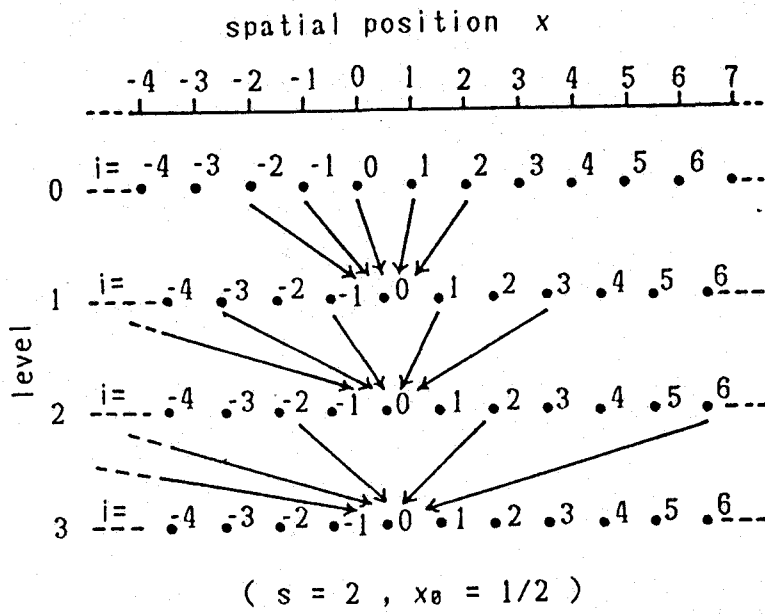


図 2.1 DCGF

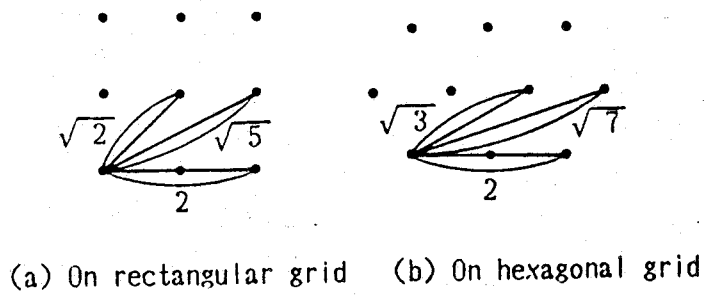


図 2.2 各格子における標本点間の距離

ファクタも可能になる。正方格子では、 $n \geq 2$ の各レベルの計算は、2回の1次元たたみ込みの方向を、その距離(s^{n-1})をとる標本点の方向とその直角方向に合わせ回転すればよい。ガウス関数の円対称性(等方性)により、結果は回転には不変である。六角格子では性質3をそのまま適用できずやや考慮を要するのでここでは触れない。

3.2 Reduced D C G F

1次元Reduced D C G Fは次式で定義される。

$$\left. \begin{aligned} h_0(x) &= g^{(m)}(x; \sigma_0) \cdot \sum_{k=-w_1}^{w_1} \delta(x-k) \\ h_n(x) &= \int_{-(w_2+\frac{1}{2})}^{w_2+\frac{1}{2}} h_{n-1}(x-s^{n-1}t) g(t; \sigma_0 \sqrt{s^2-1}) dt \cdot \sum_{k=-\infty}^{\infty} \delta(x-x_n-s^n k) \end{aligned} \right\} (2.17)$$

式(2.17)に対する離散変数 i を用いた計算式は次式のように表される。

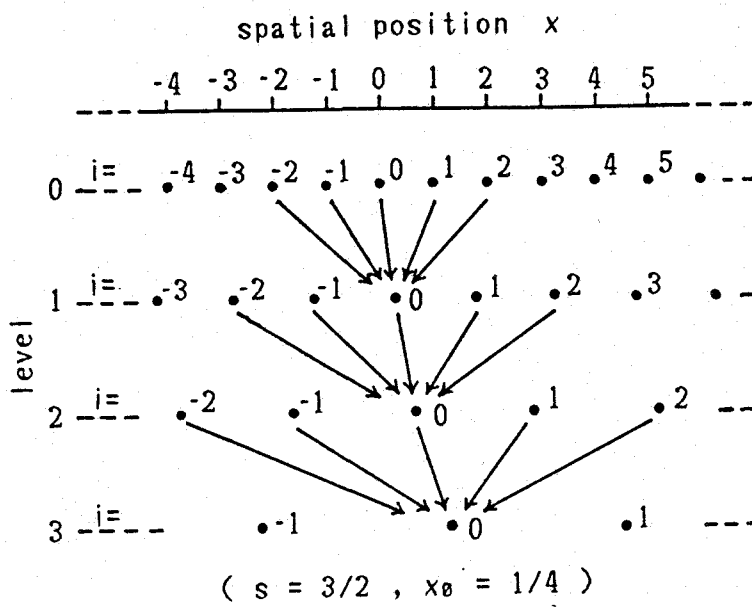
$$\left. \begin{aligned} h_0(i) &= \begin{cases} g^{(m)}(i; \sigma_0) & (-w_1 \leq i \leq w_1) \\ 0 & (i < -w_1, i > w_1) \end{cases} \\ h_n(i) &= \sum_{k=-w_2}^{w_2} h_{n-1}(\lfloor si+x_0+1/2 \rfloor - k) \\ &\quad \cdot g(k - (si+x_0 - \lfloor si+x_0+1/2 \rfloor); \sigma_0 \sqrt{s^2-1}) \end{aligned} \right\} (2.18)$$

このとき、レベル n の標本点 i における実際の空間位置 x は、次式で与えられる。

$$x = s^n i + x_n \quad (2.19)$$

図2.3に $s=3/2, x_0=1/4$ のときの計算の流れを示す。

Reduced D C G Fにおいては、各レベル n での計算は全く同一で、計算すべき標本数はレベル0の標本数に比べ $1/s^n$ になる。また、 s は整数である必要はないが、たたみ込みの重み係数のパターンは s を既約分数で表したときの分母の値だけ必要になり、 s が無理数の場合は各標本点 i にそれぞれ別の重み係数の設定が必要になる。ただし、先に述べたのと同様に、2次元正方格子での $\sqrt{2}$ 、 $\sqrt{5}$ など、および六角格子での $\sqrt{3}$ 、 $\sqrt{7}$ などのスケールファクタに対しては、重み係数の設定は1通りとすることができる。なお、図2.3において、 x_0 は各レベル間の標本点の継承関係を対等にするように設定されている。



☒ 2.3 Reduced DCGF

第4節 離散化に伴う誤差の解析

離散化したときの誤差，すなわちDCGFおよびReduced DCGFのCCGFに対する情報の損失を解析する。直感的に，各レベル毎に高域が減衰されていくので，それに応じて標本化間隔を大きくしても情報の損失は少ないであろうことは容易に想像される。そこで，その誤差を計算誤差，表現誤差という2つの誤差に着目し解析する。計算誤差は，計算された各標本値とその標本点での所望の連続関数の値との誤差を表し，前レベルにおけるなるべく少ない標本点から(すなわち，なるべく少ない計算コストで)次レベルでの値を計算したいという要求に起因する。これは，DCGFおよびReduced DCGFの計算過程より生ずる本手法に固有の誤差である。表現誤差は，各レベルでの標本化間隔がそのレベルでの所望の連続関数を表現するのに十分小さいかどうかを測る測度であり，計算誤差とは独立である。これは， σ が大きくなるにしたがい，標本化間隔を大きくするReduced DCGFにおいて問題になる。以下では，それらの誤差を十分小さく抑えるための条件を考察する。

4.1 計算誤差の解析

基本的に σ_0 と w_1, w_2 の間に計算コストと精度に関する兼ね合いがある。折り返し誤差を十分小さくするためには σ_0 を大きくとる必要があり，そうすればフィルタの幅の打ち切り誤差を十分小さくするため w_1, w_2 も大きくとらねばならず，計算コストが大きくなる。性質2は，ガウス関数がこの兼ね合いを最適化する関数であることを示し，これらの値をうまく選べば少ない計算量で誤差を十分小さくおさえることができる。

4.1.1 折り返し誤差の解析(σ_0 の設定条件)

まず， w_1, w_2 は理想的(無限大)とし，計算誤差を十分小さくするためには σ_0 をどう設定するべきかを考える。なお，以下の議論では一般性を失わず $x_0=0$ と仮定する。Reduced DCGFでの計算は式(2.17)に示されるようにレベル $n-1$ での標本点から，ガウス関数を内挿関数として次式によって $h'_n(x)$ を復元し，

$$h'_n(x) = \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)g(t;\sigma_0\sqrt{s^2-1})dt \quad (2.20)$$

それを再標本化したものと等価である。そこで，式(2.20)の周波数特性と所望の周波数特性の差を折り返し誤差関数として定義する。DCGFにおいても，Reduced DCGFにおける内挿関数による復元，そして再標本化という処理がそ

それぞれの標本点で並列に行われていると考えることができるので、Reduced DCGFと同じ誤差関数を考えて問題ない。

式(2.17)および式(2.20)のフーリエ変換は次式のようになる。

$$\left. \begin{aligned} H_0(u) &= \sum_{k=-\infty}^{\infty} G^{(m)}(u-k; \sigma_0) \\ H_n(u) &= \sum_{k=-\infty}^{\infty} H'_n(u-k/s^n) \\ H'_n(u) &= H_{n-1}(u)G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1}) \end{aligned} \right\} \quad (2.21)$$

ただし、 $G^{(m)}(u; \sigma) = (2\pi ju)^m \exp(-u^2/2(1/2\pi\sigma)^2)$ 。

図2.4に式(2.21)の計算の流れを示す。すなわち、内挿フィルタ $G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1})$ と $H_{n-1}(u)$ におけるsidelobeの干渉が折り返し誤差となる。よって、レベル n における折り返し誤差関数 $f_n(u)$ を次式で定義すると、

$$f_n(u) = H'_n(u) - G^{(m)}(u; s^n \sigma_0) \quad (2.22)$$

その主要成分は次のようになる。

$$\left. \begin{aligned} f_0(u) &= 0 \\ f_n(u) &= \left\{ \left(\sum_{k=-1}^1 f_{n-1}(u-k/s^{n-1}) \right) + G^{(m)}(u-1/s^{n-1}; s^{n-1} \sigma_0) \right. \\ &\quad \left. + G^{(m)}(u+1/s^{n-1}; s^{n-1} \sigma_0) \right\} G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1}) \end{aligned} \right\} \quad (2.23)$$

ここで、レベル0の誤差を便宜上0としているが、レベル0の誤差は扱いが異なるので後で述べることにする。

式(2.23)において、次式が誤差の源であり、各レベル毎に発生し、伝播する。

$$f'_n(u) = \left\{ G^{(m)}(u-1/s^{n-1}; s^{n-1} \sigma_0) + G^{(m)}(u+1/s^{n-1}; s^{n-1} \sigma_0) \right\} G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1}) \quad (2.24)$$

そこで、次式で表される $f'_n(u)$ の平均2乗誤差の相対量(相対誤差) F'_n が十分小さくなるように σ_0 を設定することを考える。

$$\begin{aligned} F'_n{}^2 &= \frac{\int_{-\infty}^{\infty} |f'_n(u)|^2 du}{\int_{-\infty}^{\infty} |G^{(m)}(u; s^n \sigma_0)|^2 du} \\ &= 2 \frac{\int_{-\infty}^{\infty} G^{(m)}(u-1/s^{n-1}; s^{n-1} \sigma_0) G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1})^2 du}{\int_{-\infty}^{\infty} |G^{(m)}(u; s^n \sigma_0)|^2 du} \end{aligned} \quad (2.25)$$

$m=0, 1, 2$ の場合について式(2.25)を計算整理すると、レベル n に依存せず次のよ

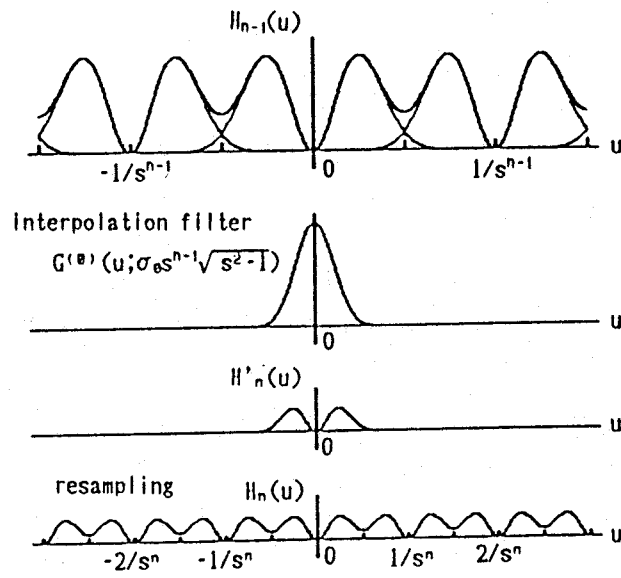


図 2.4 周波数領域における Reduced DCGF

うになる。

$$\left. \begin{aligned}
 m=0 \text{ のとき } & F'_n{}^2 = 2 \exp(-A\sigma_\theta^2) \\
 m=1 \text{ のとき } & F'_n{}^2 = (4A\sigma_\theta^2 + 2) \exp(-A\sigma_\theta^2) \\
 m=2 \text{ のとき } & F'_n{}^2 = (8A^2\sigma_\theta^4/3 + 8A\sigma_\theta^2 + 2) \cdot \exp(-A\sigma_\theta^2)
 \end{aligned} \right\} \quad (2.26)$$

ただし, $A = 4\pi^2(1 - 1/s^2)$.

実際には, その誤差が各レベル毎に式(2.23)に従い累積されていくので次式で表される各レベルでの相対誤差 F_n はレベル毎に増加する傾向にある。

$$F_n^2 = \int_{-\infty}^{\infty} |f_n(u)|^2 du / \int_{-\infty}^{\infty} |G^m(u; s^n \sigma_\theta)|^2 du \quad (2.27)$$

F_n の解析的取扱は困難であるので, 数値計算により, $s=2$ および $s=\sqrt{2}$ のときの誤差蓄積の傾向を調べると図2.5のようになる。また, s が小さくなるにつれ, 誤差蓄積の傾きが大きくなる傾向にあることが確かめられている。よって, $F'_n = F'_1 = F_1$ であるので $s \geq \sqrt{2}$ で4オクターブ程度の帯域幅の計算を行う場合, 精度 F_n を得るには F'_n が,

$$\left. \begin{aligned}
 m=0 \text{ のとき } & F'_n \leq F_n/8 \\
 m=1 \text{ のとき } & F'_n \leq F_n/4 \\
 m=2 \text{ のとき } & F'_n \leq 10^{-2} F_n
 \end{aligned} \right\} \quad (2.28)$$

を満たすように式(2.26)の σ_θ を設定すれば十分である。特に, $s=2$ および $\sqrt{2}$ の場合には, 直接図2.5より, たとえば $F_n = 1/256 = 3.9 \times 10^{-3}$ (8bit)の精度を得るためには, 次のような σ_θ の設定条件が得られる。

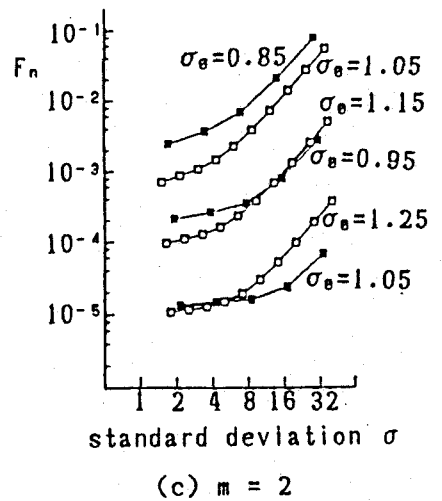
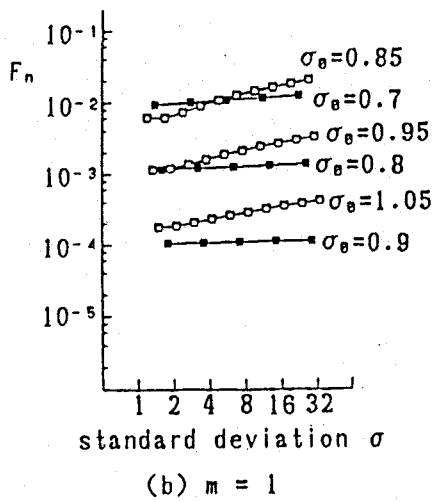
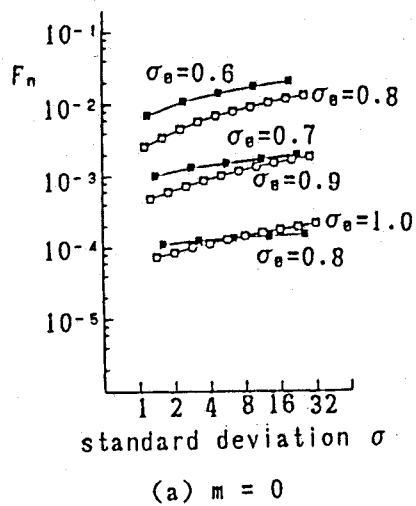
$$\left. \begin{aligned}
 m=0 \text{ のとき } & s=2: \sigma_\theta \geq 0.7, \quad s=\sqrt{2}: \sigma_\theta \geq 0.9 \\
 m=1 \text{ のとき } & s=2: \sigma_\theta \geq 0.8, \quad s=\sqrt{2}: \sigma_\theta \geq 0.95 \\
 m=2 \text{ のとき } & s=2: \sigma_\theta \geq 0.95, \quad s=\sqrt{2}: \sigma_\theta \geq 1.15
 \end{aligned} \right\} \quad (2.29)$$

4.1.2 打ち切り誤差の解析(w_1, w_2 の設定条件)

4.1.1に従い, 折り返し誤差を十分小さく抑えるように設定された σ_θ に対して, フィルタの幅の打ち切り誤差を十分小さく抑えるための w_1, w_2 の設定条件を考える。折り返し成分が十分小さいとして, 標本化により生ずる周波数領域でのsidelobeを無視すれば(すなわち, 連続系で考えれば), w を

$$(2w_1 + 1)/\sigma_\theta = (2w_2 + 1)/(\sqrt{s^2 - 1} \sigma_\theta) = w \quad (2.30)$$

とおいたときの各レベルの周波数特性は次式のようにになる。



(■ $s = 2$, □ $s = \sqrt{2}$)

図 2.5 各レベル毎の F_n の誤差蓄積

$$\left. \begin{aligned}
 H_0(u; \sigma_0) &= G^{(m)}(u; \sigma_0) * \sigma_0 w \operatorname{sinc}(\sigma_0 w u) \\
 H_n(u; \sigma_0) &= H_{n-1}(u; \sigma_0) \{ G^{(0)}(u; \sigma_0 s^{n-1} \sqrt{s^2-1}) \\
 &\quad * \sigma_0 s^{n-1} \sqrt{s^2-1} w \operatorname{sinc}(\sigma_0 s^{n-1} \sqrt{s^2-1} w u) \}
 \end{aligned} \right\} \quad (2.31)$$

ただし、 $\operatorname{sinc}(u) = \sin(\pi u) / (\pi u)$ とし、 $*$ は u についてのたたみ込み積分を表す。
打ち切り誤差関数 $t_n(u; \sigma_0)$ を次式で定義し、

$$t_n(u; \sigma_0) = H_n(u; \sigma_0) - G^{(m)}(u; s^n \sigma_0) \quad (2.32)$$

$u' = \sigma_0 u$ とおくと、 $t_n(u; \sigma_0) = (1/\sigma_0^m) t_n(u'; 1)$ 、 $G^{(m)}(u; \sigma_0) = (1/\sigma_0^m) G^{(m)}(u'; 1)$ となるので、レベル n における相対誤差 T_n は

$$\begin{aligned}
 T_n^2 &= \int_{-\infty}^{\infty} |t_n(u; \sigma_0)|^2 du / \int_{-\infty}^{\infty} |G^{(m)}(u; s^n \sigma_0)|^2 du \\
 &= \int_{-\infty}^{\infty} |t_n(u'; 1)|^2 du' / \int_{-\infty}^{\infty} |G^{(m)}(u'; s^n)|^2 du'
 \end{aligned} \quad (2.33)$$

となり σ_0 に依存しない。すなわち、式(2.30)のように w をおくことにより、ある σ_0 においてフィルタ半径を w_1, w_2 に設定したときの誤差を w のみに着目して解析できる。式(2.33)の数値計算により、 $s=2$ および $s=\sqrt{2}$ のときの誤差蓄積の傾向を調べると図 2.6 のようになる(ただし $m=2$ の場合、 $s=2$ のときの誤差蓄積が $s=\sqrt{2}$ のときとほとんど等しいので $s=\sqrt{2}$ のときのみ示している)。たとえば 4 オクターブ程度の帯域幅の計算を行う場合、8bit の精度を得るには、

$$\left. \begin{aligned}
 m=0 \text{ のとき } & s=2: w \geq 6.8, \quad s=\sqrt{2}: w \geq 7.0 \\
 m=1 \text{ のとき } & s=2: w \geq 7.4, \quad s=\sqrt{2}: w \geq 7.4 \\
 m=2 \text{ のとき } & s=2: w \geq 10.6, \quad s=\sqrt{2}: w \geq 10.6
 \end{aligned} \right\} \quad (2.34)$$

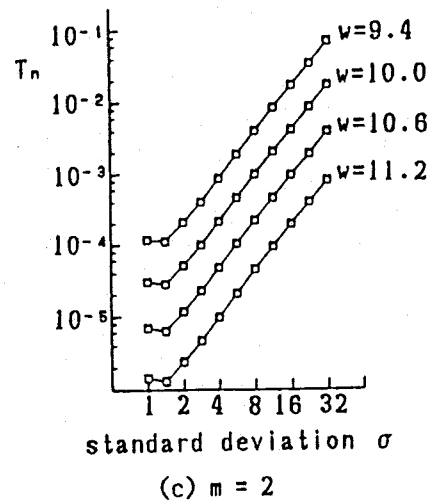
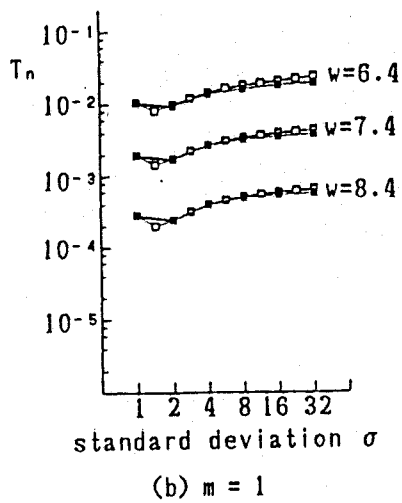
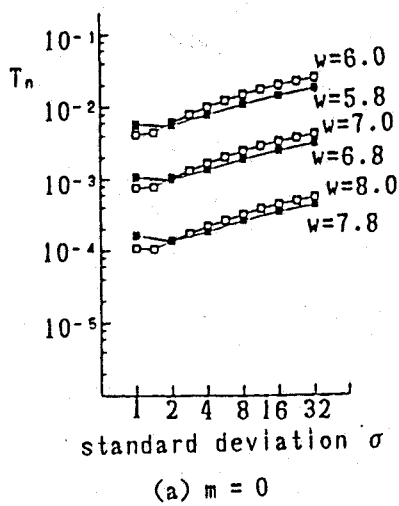
と設定すれば十分である。また、2 あるいは $\sqrt{2}$ 以外のスケールファクタでもほぼ同じ設定条件でよいことが確かめられている。これより、 w_1, w_2 は、 x_0 も考慮にいれると、次式によって決定される。

$$\left. \begin{aligned}
 w_1 &= \lceil (\sigma_0 w - 1) / 2 \rceil \\
 w_2 &= \lceil (\sigma_0 \sqrt{s^2-1} w - 1) / 2 + |x_0| \rceil
 \end{aligned} \right\} \quad (2.35)$$

ただし、 $\lceil a \rceil$ は a より小さくない最小整数を表す。

4.2 表現誤差の解析

4.1 での議論は、レベル $n-1$ での標本点から、レベル n での所望の連続関数を十分精度よく復元するための条件を導いた。しかし、この条件は各レベルの標



(■ $s = 2$, □ $s = \sqrt{2}$)

図 2.6 各レベル毎の T_n の誤差蓄積

本値が、次レベルの計算をするための十分な情報を持ち、かつ、その標本点における所望の連続関数の値と十分近いことは保証するが、各レベルでの所望の連続関数を表現するために標本化間隔が十分小さいかどうかは保証しない。信号あるいは画像に処理を施す場合、 σ に比べ標本化間隔がある程度以上広いと、処理結果がエイリアシングの影響を大きく受けることになる。たとえば、ラプラシアンガウスフィルタを適用した画像から零交差輪郭を抽出する場合、エイリアシングの影響を受けるとそのトポロジカルな連結関係も大きく変化してしまい、定性的な特徴すら捉えられなくなる。そこで、周波数領域において折り返し成分として現れる偽の情報およびナイキスト境界外に失われる情報を誤差と考え、標本化間隔を Δx として次のような表現誤差関数 $r(u)$ を定義する。

$$r(u) = \begin{cases} \sum_{k=-\infty}^{\infty} G^{(m)}(u-k/\Delta x; \sigma) - G^{(m)}(u; \sigma) & (|u| < 1/2\Delta x) \\ -G^{(m)}(u; \sigma) & (|u| \geq 1/2\Delta x) \end{cases} \quad (2.36)$$

そして、その相対誤差 R は次式で近似できる。

$$R^2 = 4 \int_{1/2\Delta x}^{\infty} |G^{(m)}(u; \sigma)|^2 du / \int_{-\infty}^{\infty} |G^{(m)}(u; \sigma)|^2 du \quad (2.37)$$

表現誤差はDCGFではほとんど問題にならないので、以下ではReduced DCGFについてのみ考える。式(2.37)において、 $\sigma = s^n \sigma_0$, $\Delta x = s^n$ として、 $m=0, 1, 2$ のときについて計算整理すると、相対誤差 R は n, s に依存せず次のようになる。

$$\left. \begin{aligned} m=0 \text{ のとき } R^2 &= (1/\sqrt{\pi})(1/\pi \sigma_0) \exp(-\pi^2 \sigma_0^2) \\ m=1 \text{ のとき } R^2 &= (1/\sqrt{\pi})(4\pi \sigma_0 + 1/\pi \sigma_0) \cdot \exp(-\pi^2 \sigma_0^2) \\ m=2 \text{ のとき } R^2 &= (1/\sqrt{\pi})(8\pi^3 \sigma_0^3/3 + 4\pi \sigma_0 + 1/\pi \sigma_0) \exp(-\pi^2 \sigma_0^2) \end{aligned} \right\} (2.38)$$

ただし、次に示す $x > 2.5\sigma$ の範囲における近似式[2.16]

$$\int_x^{\infty} g(t; \sigma) dt = (\sigma / (2\sqrt{2\pi} x)) \exp(-x^2/2\sigma^2) \quad (2.39)$$

を用いている。

4.1.1に従い設定された σ_0 に対して R が十分小さくならない場合、 σ_0 を大きくして R を十分小さくするか、各レベルで標本点を加える(たとえば、標本化密度を倍にした場合、式(2.38)において σ_0 を倍にしたのと等価である)必要がある。

4.3 検討

上で触れなかったいくつかの問題点について検討を加えておく。

4.3.1 レベル0での誤差の解析

レベル0における計算誤差(折り返し誤差)は式(2.36)において $\sigma = \sigma_0, \Delta X = 1$ と
 いた場合と定義でき、表現誤差と同様にその大きさは式(2.38)で表される。

4.1.1に従い設定された σ_0 に対しRが十分小さくならない場合、 σ_0 を大きくする
 か、あるいはレベル0での計算結果を用いず、次式のように直接レベル1から計
 算を行うようにすればよい。

$$\left. \begin{aligned} h_1(x) &= g^{(m)}(x; s\sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)g(t; \sigma_0\sqrt{s^2-1})dt \quad (n=2, 3, 4, \dots) \end{aligned} \right\} \quad (2.40)$$

ただし、レベル1におけるフィルタの半径 w_1 は式(2.35)において σ_0 を $s\sigma_0$ に置き
 換えた大きさとする必要がある。

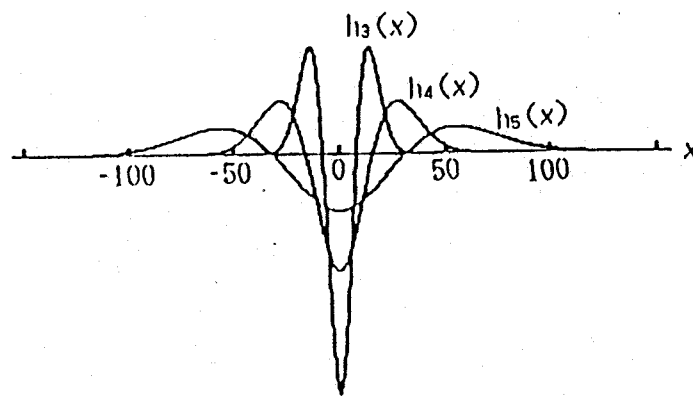
4.3.2 まるめ誤差の対策

これまでの議論では、有限語長計算によるまるめ誤差については論じなかつ
 た。浮動小数点で計算を行う場合はほとんど問題はないが、固定小数点の場合、
 その影響を考慮する必要がある。さらに、微分を行う場合 σ が大きくなるにつ
 れ信号の電力が非常に小さくなり、真の情報が各レベル毎にまるめ誤差に埋も
 れていくので、その影響は深刻になる。この問題は次式のように各レベルで信
 号を増幅することにより回避できる。

$$\left. \begin{aligned} h_0(x) &= C\sigma_0^m g^{(m)}(x; \sigma_0) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)\{s^m g(t; \sigma_0\sqrt{s^2-1})\}dt \end{aligned} \right\} \quad (2.41)$$

Cをある程度大きく定めれば、フィルタを信号にたたみ込んだとき、各レベルの
 信号値に十分な桁数を与え、次レベルの計算におけるまるめ誤差の影響を小さ
 くすることができる。 σ_0^m と s^m は一種の正規化係数で、微分による信号電力の減
 衰を防ぎ、各レベルにおけるフィルタの周波数特性のピーク値を σ_0 に依存せず
 一定にするものである。このように設定すると、Reduced DCGFでは各レベ
 ルのフィルタの電力と標本化間隔との積が等しくなり、その意味で各レベルの
 各標本点は対等な情報量を担っていることになる。また、重み係数は十分大き
 な桁数で表現されねばならず、たたみ込み計算の後、正規化を行う必要がある。

なお、式(2.41)に基づくDCGFで $m=2, s=2, x_0=0, \sigma_0=1.0, w_1=5, w_2=8$ と設定して合成されたフィルタ系列($n=3\sim 5$)を図2.7に示す。この場合、計算誤差は $n=5$ で $F_5=4.4\times 10^{-4}, T_5=2.4\times 10^{-3}$ である。図2.7ではDCGFによらず直接計算を行ったフィルタ系列を重ね合わせているが、ほぼ完全に重なっている。



($m=2, s=2, x_0=0, \sigma_0=1.0, w_1=5, w_2=8$)

図 2.7 DCGF により合成されたフィルタ系列

第5節 関連研究

5.1 Burt[2.10], Crowley[2.11]の研究との関連

まず、本研究とBurtの研究の関連について述べる。本章における $m=0, s=2$ でのDCGFが、BurtのHDCに相当し、 $x_0=0$ のときodd HDC、 $x_0=1/2$ のときeven HDCに相当する。また、Reduced DCGFは、BurtのReduced HDCおよびFractional HDCを包含するもので、それぞれ本章での $s=2, s=3/2$ の場合に相当する。ただし、Burtがフィルタの係数を半ば実験的に設定したのに対し、本章では、このようなフィルタが構成できる理論的根拠を与えたといえる。なお、Burtの定式化では式(2.7)に相当する式が次式で表せるので、

$$\left. \begin{aligned} h_0(x) &= \delta(x) \\ h_n(x) &= \int_{-\infty}^{\infty} h_{n-1}(x-s^{n-1}t)g(t;\sigma_0)dt \end{aligned} \right\} \quad (2.42)$$

厳密には、各レベルにおいてスケールファクタは一定ではない。

Crowleyは、本章における $m=0, s=\sqrt{2}$ での2次元Reduced DCGFによるDOG(Difference of Gaussians)フィルタの構成法を述べている。Crowleyは実験により $\sigma_0=\sqrt{2}, w_1=w_2=4, (x_0=0)$ の場合良好な結果が得られたとしているが、本章の解析からもこれらの値が十分な精度を与えることは容易にわかる。図2.5および式(2.38)より、 $\sigma_0=1.0$ でも計算誤差 2×10^{-4} 、表現誤差 2×10^{-3} の精度を得ることも可能であるので、Crowleyの設定は十分余裕があるといえる。

Burt, Crowleyの研究が、固定したスケールファクタで、微分を考慮に入れてなかったのに対して、本章での議論は、彼らの手法を一般化し、微分オペレータ、スケールファクタの変性を許して再定式化を行い、さらに、誤差の発生する要因を解析し、十分な精度を得るための条件を明確化したものである。よって、所望の線型微分、所望のスケールファクタでガウスフィルタが構成でき、さらに、第4節で述べた設定条件を満たす範囲で任意の σ_0 の設定ができるので微妙な解像度の調節も行える。

5.2 一様フィルタの繰り返し法[2.12], [2.13]との比較

複数解像度のガウスフィルタを高速に実現するもう一つの方法に一様フィルタの繰り返し法(Cascaded Uniform Filters. 以下、CUFと略す)がある。CUFは、中心極限定理に基づき、一様フィルタ(移動平均)を数回繰り返すとガウスフィルタの非常によい近似が得られるというもので、スケールファクタもある程度調節可能である。

DCGFあるいはReduced DCGFが1標本点につき $2w_2+1$ 回(レベル0のみ $2w_1+1$)の積和計算が必要であるのに対して, CUFでは1標本点につき1回ずつの加算と減算で済む。よって, DCGFとCUFを比べると明らかにCUFのほうが計算効率が高い。しかし, CUFでは解像度に応じた標本化密度(標本数)で結果を得たい場合でも, 各レベルでレベル0での標本数の計算が必要である。Reduced DCGFでは, フィルタの次元を p , レベル0での標本数を N^p とすると, 計算すべき標本数は約 $(s^p/(s^p-1))N^p$ となり, さらにレベル1から結果を得たい場合は式(2.40)をReduced DCGFに定式化すると約 $(1/(s^p-1))N^p$ の標本数の計算でよいことがわかる。一方, CUFでは計算すべき標本数は, 計算したいレベル数を q として qN^p となる。よって, 解像度に応じた標本化密度で結果を得たい場合はReduced DCGFのほうが計算効率のよい場合がかなりある。

しかし, 計算効率の面以外ではDCGF, Reduced DCGFは本章において微分などいくつかの変要素も含めた設計原理を確立したので, より柔軟で精密な設計が行えるという点でCUFより有利である。たとえば, 1次元信号の連続した σ のガウスフィルタリングあるいはその微分を行う(すなわち, scale space画像[2.2]を得る)場合, $s=2$ において, $\sigma_0=a, 2^{1/b}a, 2^{2/b}a, \dots, 2^{(b-1)/b}a$ (ただし, $b=2, 3, 4, \dots$)と設定した b 個のDCGFを適用すれば, かなり小さいスケールファクタ $2^{1/b}$ でscale space画像を高速に得ることができる。また, 式(2.20)に示されるように, レベル $n-1$ からレベル n の標本値を任意の位置で得ることも可能であるので精密な位置合わせへの応用等にも有効である。

第6節 結言

コンピュータビジョン研究において重要な役割を果たしている複数解像度のガウスフィルタおよびその微分フィルタを，ガウス関数の性質を活した再帰的構造のフィルタで，高速に精度よく実現するための手法を明確にした．本手法の計算量の詳しい解析[2.10], [2.11]およびFFTとの比較[2.10]については参考文献から容易に類推されるので参考にされたい．

また，本章での議論は，網膜受容野の再帰的構造モデル[2.17]を理論的に説明するものでもある．多重解像度モデルとしてのガウスフィルタが，その実現面においてもよい性質が活かされることが明確に示されたことは，人間の視覚モデルの立場からも興味深い．

第3章 多重解像度両眼立体視モデル

第1節 序言

ステレオ対応アルゴリズムにおいて重要な問題の一つは、曖昧さを解消するための対応規則である[3.1]. 対応規則としては、なめらか拘束(smoothness constraint)[3.2]-[3.4]や順序拘束(ordering constraint)[3.5]がよく用いられる. しかし, 例えば, フェンス越しにテクスチャ表面を見た場合では, いたるところが不連続で視差場(disparity field)全体のなめらかさは成立せず, 図3.1に示すように, 対応点の順序が入れ替わる場合も多くなる. ここで, 重要な点は, 人間の両眼立体視能力を鑑みると, なめらか拘束や順序拘束が成立しない状況は, 例外的な場合というよりも, むしろ, その様な場合にこそ, 真に, 両眼立体視の能力が発揮されるべき状況であることが多い, ということである. 即ち, これらの拘束は, 曖昧さの解消能力を強力にするために, シーン, 或は, シーンとカメラの関係に課している制約が一般性を欠くものになっている.

Prazdnyは, 表面が互いに交錯したような場合にも成立する, より一般性のあるシーンの制約として凝集性(coherence)に着目し, 対応規則の満たすべき基本原理を示唆した[2.6]. 即ち, 視差の不連続をペナルティとせず視差の類似性による促進作用に基づいて曖昧さ解消を行うことで, 不連続に対処できることを指摘した. このことは当然のことに思えるが, 従来の多くのアルゴリズムが, 視差の不連続をペナルティとして, これを画面全体で最小にする視差場を得るような定式化を行っている[3.3], [3.4]. これらのアルゴリズムは, 不連続部における対応付けが不安定なものになるので, いたるところが不連続である透過性の表面, 或は, フェンス越しのシーンなどでは, うまく働かない.

Prazdnyは, 対応によって生じる3次元点間の結合の強さを表す支持関数を定義して, 対立する対応候補の中で, 近傍領域からの支持の総和が最も大きいものを対応点とすることで, 凝集性に基づくアルゴリズムを実現した[3.6]. しかし, 基本的にシーンに課している制約を弱くする分だけ曖昧さの解消能力が低下するので, Prazdnyの用いた局所的な判断基準では, 満足な性能を得ることは難しいと考えられる.

本研究では, あらゆる視差分布の可能性の中で類似性による促進を最大にする視差場を選択するという1つの最適化アプローチを示す. 最適化過程に, 複数解像度の結合を含む複数の対応規則と対応単位を統合することによって, 制

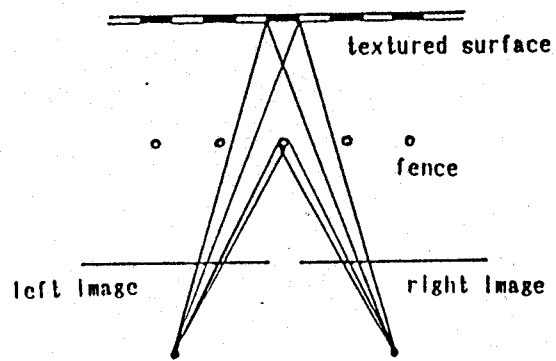


図 3.1 順序拘束の反例

約の弱いシーンに対して強力な曖昧さ解消能力を実現することを目的とする。

本研究では、最適化をHopfieldの神経回路網類似ネットワークのエネルギー最小化モデル[3.7]によって実現する。エネルギーは、ネットワークの結合関係により定まり、対応の最適性、即ち、類似性による促進を反映するものである。まず、あらゆる対応の可能性(対応仮説)を列挙し、それぞれをネットワークの各ユニットとする。対応規則は、各対応仮説間の融合により生じる3次元空間での群化規則と捉えることができ、正しい対応仮説に対して、より群化が促進されるものでなければならない。よって、信頼性の高い対応を得るためには、様々な観点からみた群化規則の統合が必要になる。このモデルでは、対応規則はネットワークのユニット間の結合規則として表されるので、複数の対応規則の組合せがネットワークにおける結合関係の重ね合わせとして、自然に、しかも、柔軟に表現することができ、並列分散処理によりエネルギーを最小化することによってこれらの規則を同時に最もよく満足する解を得ることができる。本研究では、凝集性に加えて、輪郭の連続性(figural continuity)[3.8]を対応規則に組み込み、複数解像度の結合に関しても不連続シーンに対して効果的に働く結合規則を提案する。また、相補的な異なる対応単位の相互作用も試みる。これらの統合によって、強力な曖昧さ解消能力の実現が期待できる。

以下では、まず第2節で、ステレオ対応問題の従来の研究を踏まえて、対応単位と対応規則について考察し、アルゴリズムの設計方針を述べる。次に第3節で、対応単位の抽出法と局所的な対応規則に基づく対応仮説の生成について述べる。さらに第4節で、ネットワーク収束の原理と、ネットワークへの対応規則の埋め込みについて述べ、アルゴリズムを定式化し、第5節では、実験結果を示す。

第2節 基本方針

2.1 対応単位

対応単位として、まず、エッジ輪郭を細かく直線近似したエッジセグメントを考える。Medioni & Nevatiaは、対応単位としてエッジセグメントを用いることの基本的な利点を強調しているが[3.3]、本研究では、より明示的に、各対応規則においてエッジセグメントの方向性を有効に活用する。

エッジセグメントの対応は密な対応結果が期待できるが、一方かなりの曖昧性がある。そこで、本研究では、もう1つの対応単位として、特徴点(interest point)を考える。特徴点は、画像明度分布の極めて個性的な点であり、まばらにしか存在しないが、対応の曖昧性は非常に少なくなる。本研究では、信頼性の高い特徴点の対応仮説がエッジセグメントの対応仮説に対して相互作用を及ぼすようにモデルを構成する。これにより、エッジセグメントの対応仮説の曖昧さ解消に有効に働くことが期待できる。

2.2 対応規則

対応規則は、カメラモデルの制約、或は、対応単位の局所的な特徴量に基づき、対応の可能性を絞るための局所対応規則と、局所対応規則に従うあらゆる対応の可能性から、最も好ましい対応関係を選択するための大局対応規則に分けられる。具体的には、局所対応規則は、対応仮説、即ち、ネットワークの各ユニットを生成するための規則として、そして、大局対応規則は、ネットワークの各ユニットの結合関係を定めるための規則として実現される。

局所対応規則としては、以下の規則を用いる。

- (1)左右の走査線整合と最大視差の限定。
- (2)エッジセグメントの最大視差勾配(disparity gradient)の限定：視差勾配は、融合によって生じる3次元解釈の奥行き(視差)方向の傾きであり、Burt & Juleszが、人間の両眼融合能力の視差勾配への依存性を示している[3.9]。視差勾配は、左右のエッジセグメントの傾きから計算でき、対応仮説を両眼融合が起こり得る視差勾配の限界に基づいて生成する。
- (3)エッジセグメントのコントラスト極性。
- (4)特徴点の型と方向性。

大局対応規則としては、以下の規則を用いる。

- (i)一意性の拘束(uniqueness constraint)[3.2]：一方の画像中の1点は、他方の画像中の高々1点に対応すると仮定する。一般に、透過性表面では、画像の

あらゆる点でこの拘束が成立するわけではない。しかし、エッジ点、或は、特徴点は、シーンの明確な物理現象に対応していると考えられるので、特殊な視点から見た場合以外は、一意性の拘束は成立すると考えて問題ない。図3.2は、左右画像の各走査線の対応を表し、各交点に対応仮説となる。一意性の拘束により、図3.2の対応仮説Aに対して、左右の同一視線上の対応仮説はすべて両立しない。対応仮説の活性度を考えると、両立しない対応仮説同士は互いに活性度を弱め合うことになり、ネットワークにおいてそれぞれ抑制性結合される。この関係は、WTA(winner-take-all)ネットワーク[3.10]、或は、基本競合系[3.11]と呼ばれる、ネットワーク構成における基本的な関係である。基本競合系を構成する対応仮説同士が、それぞれ、類似性の促進(群化の促進)を競い、この並列相互作用の繰り返しを通して、最終的に、類似性の促進を最大にする仮説群が両立しない他の仮説を抑制して活性化する。以下に示す(ii)~(iv)の対応規則は、類似性の促進作用を規定する融合空間における群化規則である。

(ii)凝集性：Prazdnyは、近傍からの結合の強さを、視差の差 ΔD と画像上での距離 d によって、次に示す支持関数 $s(\Delta D, d; c)$ で定式化した[3.6]。

$$s(\Delta D, d; c) = (1/cd\sqrt{2\pi}) \exp(-\Delta D^2/2c^2d^2) \quad (3.1)$$

ここで、 c は定数で、式(3.1)は図3.3のような形状をしている。このとき、視差勾配 $|\Delta D|/|d|$ が結合の強さに反映される。本研究では、各ユニット間の興奮性結合の重みを、式(3.1)に基づいて設定している。Prazdnyが、繰り返し処理によらず近傍領域の支持の総和のみにより対応を決定したのに対し、本研究では最適対応を得るため、ネットワーク内の繰り返し過程を通して信頼性の高い情報が伝播して曖昧さを解消していくので、アルゴリズムの能力はより向上する。

(iii)輪郭の連続性：凝集性は等方性であったのに対して、輪郭の連続性はエッジセグメントの方向性を考慮して、その3次元空間でのなめらかなつながり具合を優先する群化規則である。この規則の埋め込みも式(3.1)に基づき、視差の差 ΔD を、エッジセグメントの対応仮説間の方向のなめらかさを測る尺度に置き換える。

(iv)複数解像度の結合：複数解像度レベルの結合も曖昧さの解消に重要な役割を果たすと考えられる。しかし、Prazdnyが指摘するように[3.13]、透過性表面、或は、交錯した表面を含むシーンでは、各解像度レベルが本質的に異なる視差情報を含んでいる場合がある。このような場合、従来の粗レベルから密レベルへ順次精密化を行う処理形態[3.14]では密レベルに含まれている視差情報を抽出することはできない。この問題に対処するため、Nishimoto & Shiraiは、ウイ

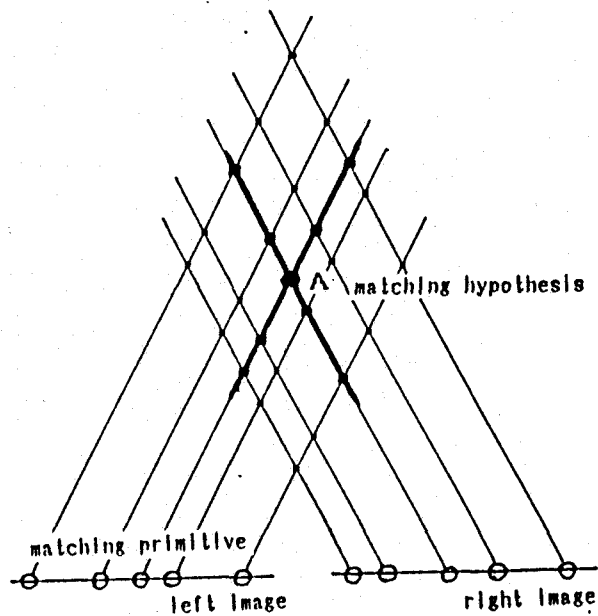


図 3.2 一意性の拘束

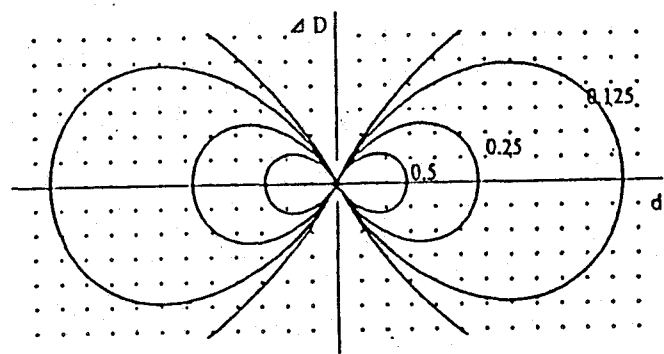


図 3.3 支持関数の形 (文献[3.12]より)

ンドウ内の視差ヒストグラムを複数の解像度でとり、最も信頼のおける解像度の情報を足掛りに対応を定めていくことで、有効に多重解像度情報を利用して[3.15]。これに対して本研究では、次のようなアプローチをとっている。視差の抽出は、おのおの各解像度レベルで並行して行う。但し、隣接するレベルにおいて、ほぼ同じ位置と方向を持つ対応単位はシーンにおいて同じ物理現象に起因し、等しい視差を持つという仮定(空間的一致の仮定(spatial coincidence assumption)[3.16])に従い、その仮定に適合する対応単位だけに対して、密レベルが粗レベルから支持を受ける。即ち、視差の継承に関する確実な証言が得られた場合のみ、より曖昧さの少ないと思われる粗い解像度の情報を有効に利用しようという考え方である。

2.3 融合座標系の設定

両眼融合により生じる3次元解釈を表現するため、融合座標系(U, V, D)(cyclopean coordinate system)を図3.4のように設定する。左画像の点(U_L, V_L)と右画像の点(U_R, V_R)の対応から生じる3次元解釈は、融合座標系において次のように表されたとする。

$$\left. \begin{aligned} U &= (U_L + U_R) / 2 \\ V &= V_L = V_R \\ D &= U_L - U_R \end{aligned} \right\} \quad (3.2)$$

対応仮説の情報は融合座標系で表現され、対応仮説間の相互作用は融合座標系における関係に基づく。

2.4 処理の概要

処理の概要を図3.5に示す。まず、左右各画像から対応単位を抽出する。次に、局所対応規則に基づき対応仮説を生成する。そして、各対応仮説をそれぞれユニットと見なし、大局対応規則に基づきユニット間の結合の重みを決定し、ネットワークを構成する。ネットワークは複数の解像度において同時に構成され、各解像度レベル間の結合が行われる。このように構成されたネットワークにおいてエネルギーを最小にする並列分散処理が行われ、収束時に活性化しているユニットが対応結果を与える。

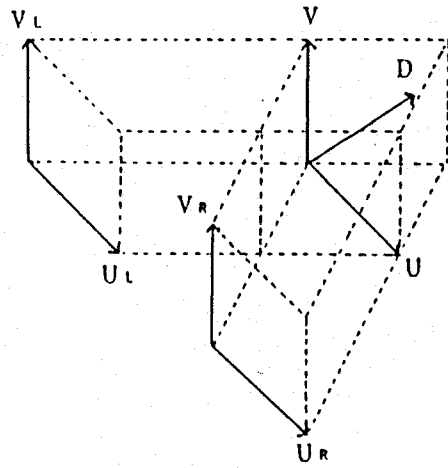
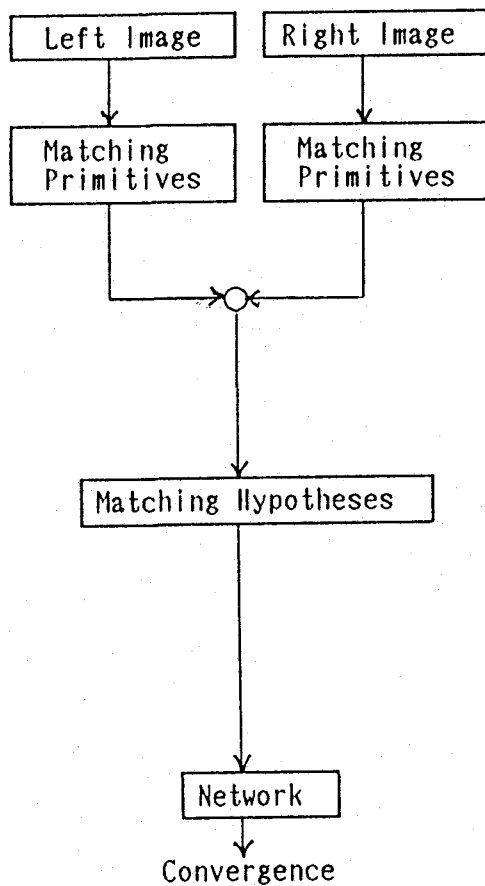


图 3.4 融合坐标系



● Extraction of edge segments and feature points

● Generation of matching hypotheses by local matching rules

- Restriction of search space
- Disparity gradient limit (edge segment)
- Polarity of contrast (edge segment)
- Similarity of feature (feature point)

● Network construction by global matching rules

- Uniqueness constraint
- Coherence principle
- Figural continuity (edge segment)
- Combining multi-resolution channels

● Energy minimization

図 3.5 処理の概要

第3節 対応仮説の生成

3.1 対応単位の抽出

対応単位の抽出は、ガウシアンフィルタを適用した画像の微分値に基づいて行われる。以下の各微分処理は、すべてガウシアンフィルタと組み合わせて複数の解像度(標準偏差)で行われ、各解像度レベルでの標本化間隔は標準偏差に比例して大きくする。複数解像度のガウシアンフィルタと組み合わせた微分値は、カスケードガウシアンフィルタの手法[3.17]により高速に精度よく計算することができる。

3.1.1 エッジセグメントの抽出

ラプラシアンフィルタを適用して、その零交差輪郭をエッジ点として抽出し、各零交差輪郭を多角形近似により分割する。ここで、分割された各エッジ点列毎に各点の明度勾配の大きさを計算し、その平均値があるしきい値以下の点列を取り除く。そして、各エッジ点列を最小自乗法により直線当てはめを行う。このとき、ある長さ以上のセグメントは所定の長さに分割する。これは、個々のセグメントの担う情報量の均一化を行うためである。得られたセグメントは、端点の座標、方向、及び、コントラストの極性を情報として持っている。

3.1.2 特徴点の抽出

画像を $f(x, y)$ として、その偏微分を f_x, f_{xx}, f_{xy} などと表す。特徴点の抽出、及び、その特徴量の計算をヘシアンに基づいて行う。ヘシアン H は次のように定義される。

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \quad (3.3)$$

個性の強さを示す尺度として、 $\det H = f_{xx}f_{yy} - f_{xy}^2$ を用いる。これは、画像濃度表面の鋭い角、或は、鞍点に反応する。まず、 $\det H$ の極大点と極小点を特徴点の候補とし、さらに、その絶対値があるしきい値以上の点を特徴点とする。 H の2つの固有値の符号から鞍点、凸点、凹点の型分類がなされる。特徴点は、型、及び、方向性を表す H の固有ベクトルと明度勾配ベクトルを情報として持つ。

3.2 対応仮説の生成

3.2.1 エッジセグメントの対応

走査線整合と最大視差の限定に従って，左画像のエッジセグメントに対する右画像における探索範囲は，最大視差を $\pm D_{max}$ として，図3.6のように表される．探索範囲内の各セグメントに対して対応を試み，各対応において，コントラストの極性が一致し，しかも，視差勾配の値があるしきい値以下のとき，対応仮説が生成される．たとえば，図3.6において，SegLとSegRの対応を仮定すると，それらの対応部分の midpoint のU座標の差により視差を得る．各セグメントの傾きを図3.7に示す角度で表し，左右それぞれ ψ_L ， ψ_R とすると，融合座標系での視差勾配は次式で表される．（導出過程は付録1．参照）

$$\left(\frac{1}{\tan \psi_R} - \frac{1}{\tan \psi_L} \right) / \sqrt{1 + \left(\left(\frac{1}{\tan \psi_R} + \frac{1}{\tan \psi_L} \right) / 2 \right)^2} \quad (3.4)$$

対応により，融合座標系における3次元セグメントが定まる．セグメントの対応仮説は，対応により得られる3次元セグメントの融合座標系での midpoint の座標，視差勾配，及び，その長さを，情報として持つ．但し，走査線に水平のセグメントは視差が決定できないので，対応単位から省いている．

3.2.2 特徴点の対応

走査線整合と最大視差の限定に従うが，探索範囲を同一走査線上に限るのはやや制限が厳しいので，縦視差を上下1画素ずつ許す．その探索範囲において型の整合性，及び，方向の類似性を基に対応仮説を生成する．まず，型の整合性を調べ，次に対応を仮定した各特徴点に対しHの固有ベクトルと明度勾配ベクトルから方向性の有無を調べ，どちらかが方向性を持つ場合は方向の類似性も調べる．エッジセグメントの対応仮説に対して，できるだけ信頼のおける情報を伝播することを目的としているので，対応仮説生成の条件はやや厳しく設定している．特徴点の対応仮説は，対応により定まる融合座標系での座標を情報として持つ．

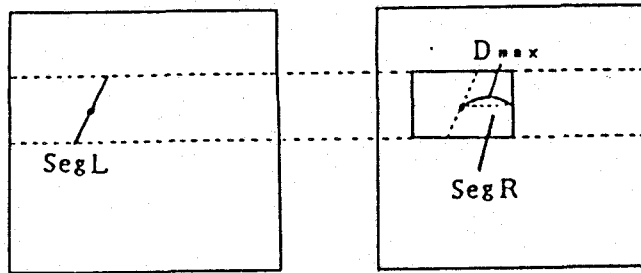


図 3.6 探索範囲

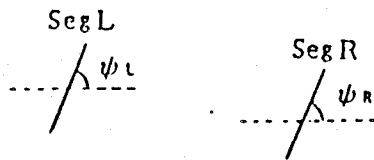


図 3.7 エッジセグメントの傾き

第4節 仮説ネットワークによる対応付け

4.1 ネットワークのエネルギー極小化

ある解像度レベルにおいて生成された対応仮説ユニットの数を、エッジセグメントの対応仮説、及び、特徴点の対応仮説を共に含めて n 個とし、それらの活性度を $p = (p_1, p_2, \dots, p_n)$ と表す。但し、 p_i はユニット i の活性度で、 Max を活性度の上限として $0 \leq p_i \leq \text{Max}$ の範囲の実数とする。また、各ユニット間の結合の重みとなる適合関係を $R = \{r_{ij}\}$ で表す。但し、ユニット間の適合関係の対称性を仮定して R は $n \times n$ 対称行列とし、各要素は実数で、対角成分は 0 とする。4.2 で述べるように、大局対応規則が R に埋め込まれる。このとき、ネットワーク全体のエネルギーは次式で定義される [3.8]。

$$E(p) = (-1/2) \sum_{i,j} r_{ij} p_i p_j - \sum_i (\eta_i - \theta_i) p_i \quad (3.5)$$

ここで、 η_i は外部入力で、本研究では、粗レベルからの継承があるユニットに関しては、粗レベルのユニットの活性度を外部入力とする。 θ_i は抑制定数で、十分な支持が得られるまでユニットを活性化させないためのものであり、時期尚早な決定を避ける働きがある。

E を最小にする p が適合関係 R を最もよく満足する活性化パターンであり、一意性の拘束を満たし、かつ、類似性による促進を最大にする。そこで、 E をなるべく小さくする p を求めるため、最急降下することを考える。 $\text{grad} E(p)$ を $E(p)$ の勾配ベクトルとして、 $q = (q_1, q_2, \dots, q_n) = -\text{grad} E(p)$ とおくと、

$$q_i = -\partial E / \partial p_i = \sum_j r_{ij} p_j + \eta_i - \theta_i \quad (3.6)$$

となる。さらに、 $0 \leq p_i \leq \text{Max}$ の制約条件より実際の探索方向 $u = (u_1, u_2, \dots, u_n)$ は q を次式のように修正したものとなる。

$$u_i(t) = \begin{cases} 0 & \text{if } (p_i(t) = 0 \wedge q_i(t) < 0) \vee (p_i(t) = \text{Max} \wedge q_i(t) > 0) \\ q_i(t) & \text{その他} \end{cases} \quad (3.7)$$

即ち、ユニット i の活性度の変化量は、図 3.8 に示すようにユニット i に結合するユニットの活性度と結合の重みの積和 $\sum_j r_{ij} p_j$ 、外部入力 η_i 、及び、抑制定数 θ_i に基づき決定される。以上のことから、ネットワークの極小化アルゴリズムは図 3.9 のように定式化される。

式(3.4)が凸関数でない限り、極小値は最小値であるとは限らないので、最適性の保証はない。しかし、曖昧さの解消能力の大きな対応規則は、エネルギー

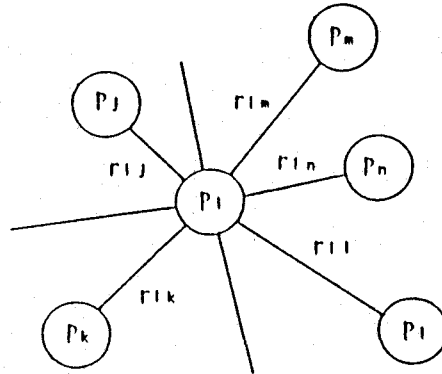


図 3.8 ユニットの結合

Initialize:

- 1) Set $t=0$. Start with an initial assignment $p_i^{(0)}$ ($0 \leq p_i^{(0)} \leq \text{Max}$).

Loop

- 2) $q_i^{(t)} = \sum r_{ij} p_j^{(t)} + \eta_i - \theta_i$
- 3) $u_i^{(t)} = \begin{cases} 0 & \text{if } (p_i^{(t)} = 0 \wedge q_i^{(t)} < 0) \vee (p_i^{(t)} = \text{Max} \wedge q_i^{(t)} > 0) \\ q_i^{(t)} & \text{otherwise} \end{cases}$
- 4) if $\|u^{(t)}\| < \delta$ (δ is a small positive value.) Stop.
- 5) $p^{(t+1)} = p^{(t)} + \alpha (u^{(t)} / \|u^{(t)}\|)$
(α is a step size.)
- 6) if $p_i^{(t+1)} > \text{Max}$ $p_i^{(t+1)} = \text{Max}$
if $p_i^{(t+1)} < 0$ $p_i^{(t+1)} = 0$
- 7) $t = t+1$

End Loop

図 3.9 ネットワークのエネルギー極小化アルゴリズム

関数を凸関数に近づける働きをすると考えられる。

ネットワークのエネルギー極小化アルゴリズムは、Hummel & Zuckerが行った対象物(object)へのラベル付け問題を解くための弛緩法の定式化(Relaxation labeling process)[3.18]と類似しているが、前者は、後者において問題となるラベルなし問題(対象物にどのラベルも対応しない)や多ラベル問題(対象物に複数のラベルが対応する)、或は、対象物とラベルが対称関係にある場合[2.19]を、自然に扱えるという点で優れている。これらは、対象物とラベルが左右各画像の対応単位に相当するステレオ対応問題には、特に重要である。

4.2 ネットワークへの対応規則の埋め込み

ここでは、2.2で考察した対応規則に基づき、対応仮説間の相互作用を、具体的に、ネットワークの各ユニット間の結合の重み、即ち、適合関係 R にどう埋め込むか、粗レベルからの継承があるユニットに関しては、1段粗い解像度レベルを外部入力 η_i としてどのように結合するかについて述べる。以下では、一意性の拘束、凝集性、及び、輪郭の連続性を表す関係をそれぞれ R_u 、 R_c 、及び、 R_f とする。各行列は $n \times n$ 対称行列で各成分は実数とし、対角成分は0とする。

4.2.1 一意性の拘束

対応仮説ユニットの中で、一意性の拘束に反する対応仮説間に抑制性結合を行う。エッジセグメント間の対応に関しては、必ずしも、1つのエッジセグメントは高々1つのエッジセグメントにしか対応しない、という法則は成り立たない。たとえば、図3.8において SegL_0 と SegR_0 の対応と、 SegL_0 と SegR_1 の対応は矛盾するが、 SegL_0 と SegR_0 の対応と、 SegL_0 と SegR_2 の対応は矛盾するものではない。また、 SegL_0 と SegR_0 の対応と、 SegL_0 と SegR_3 の対応も積極的に適合することはないが、矛盾するとは言えない。特徴点間の対応に関しては、1つの特徴点は高々1つの特徴点にしか対応しないという基準に従えばよい。

そこで、次の条件に従って、一意性の拘束に基づく仮説間の不適合関係を定める。

エッジセグメント SegL_0 と SegR_0 の対応仮説($\text{SegL}_0, \text{SegR}_0$)は、次の条件1)、或は、2)を満たす対応仮説(SegL, SegR)には不適合である。

1) $\text{SegL} = \text{SegL}_0$, かつ、 SegR と SegR_0 は走査線方向にずらすと重なる部分がある。

2) $\text{SegR} = \text{SegR}_0$, かつ、 SegL と SegL_0 は走査線方向にずらすと重なる部分がある。

特徴点PtL₀とPtR₀の対応仮説(PtL₀, PtR₀)は、次の条件1), 或は, 2)を満たす対応仮説(PtL, PtR)には不適合である。

$$1) PtL = PtL_0.$$

$$2) PtR = PtR_0.$$

よって、一意性の拘束を表す不適合関係R_uの各成分r_{u_ij}は次のように定められる。

$$r_{u_{ij}} = \begin{cases} 1 & \text{if ユニット}i\text{とユニット}j\text{が不適合} \\ 0 & \text{その他} \end{cases} \quad (3.8)$$

このように、関係を定めることによつて、エッジセグメントの1対多対応がうまく処理される。

4.2.2 凝集性

凝集性に基づく2つの対応仮説間の適合性は、エッジセグメントに対しては融合座標系での中点、また、特徴点に対してはその融合座標系での座標に基づく。R_cの各要素r_{c_ij}は、対応単位に関係なく、2つの対応仮説の座標をそれぞれ(U_i, V_i, D_i), (U_j, V_j, D_j)、視差の差をΔD=D_i-D_j、距離をd=√(U_i-U_j)²+(V_i-V_j)²として、式(3.1)に基づき、次のように定める。

$$r_{c_{ij}} = \begin{cases} s(\Delta D, d; c_0) & \text{if } |\Delta D| \leq f \wedge d \leq w \\ 0 & \text{その他} \end{cases} \quad (3.9)$$

ここで、fとwは融合空間における支持近傍を示している。

4.2.3 輪郭の連続性

輪郭の連続性は、エッジセグメントの対応仮説に対してのみ適用される。2つの対応仮説間の適合性は融合座標系での2つの3次元セグメントの連続性に基づき、次のような基準を用いる。

図3.9(a)は3次元セグメントの融合座標系のU-V平面への正射影で、α_i+α_jを2つのセグメントの画像面での連続性の測度とする。図3.9(b)は図3.9(a)の直線A-Bを縦軸にとり視差を横軸にとつたもので、各セグメントの視差勾配をG_i, G_jとして、β_i=|Tan⁻¹G_i|, β_j=|Tan⁻¹G_j|と近似する。そして、β_i+β_jを2つのセグメントの視差方向の連続性の測度とする。R_fの各要素r_{f_ij}は、α_i+α_jとβ_i+β_j、及び、式(3.1)に基づき、次のように定める。

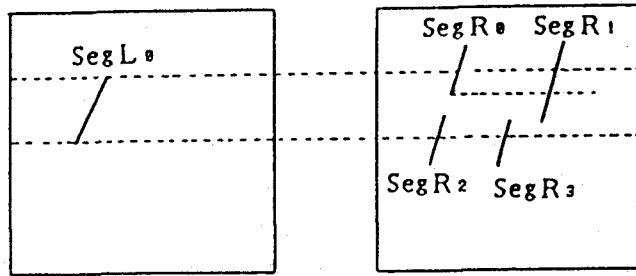


図 3.10 エッジセグメントの対立する対応仮説

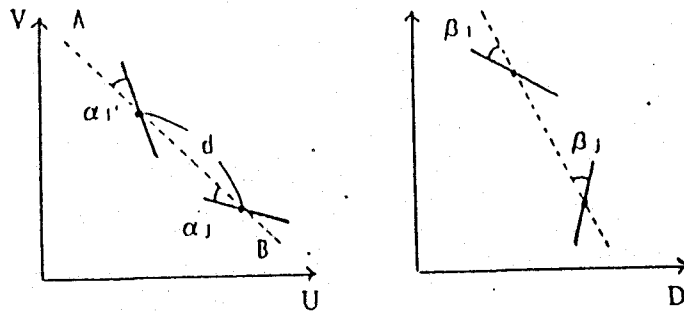


図 3.11 輪郭の連続性

$$r_{fij} = \begin{cases} s(\alpha_i + \alpha_j, d; c_f) \cdot s(\beta_i + \beta_j, d; c_f) \\ \text{if } \alpha_i + \alpha_j \leq \alpha_{\max} \wedge \beta_i + \beta_j \leq \beta_{\max} \wedge d \leq w \\ 0 \quad \text{その他} \end{cases} \quad (3.10)$$

4.2.1~4.2.3のそれぞれの対応規則は、次式のように4.1の適合関係Rに埋め込まれる。

$$R = -\varepsilon^- R_u + \varepsilon^+ (\varepsilon_c R_c + \varepsilon_f R_f) \quad (3.11)$$

ここで、 ε^- と ε^+ は、抑制性結合と興奮性結合の重みの割合を調節する係数で、 ε_c と ε_f は、各促進作用の重みの割合を調節する係数である。

4.2.4 複数解像度の結合

2.2で述べたように、粗レベルから密レベルへの継承の条件は、空間的一致の仮定に基づく。なお、以下に示す処理は左画像の対応単位に対してのみ行われる。図3.10において、1段粗い解像度のエッジセグメントSegCは次の条件1), かつ, 2)を満たすエッジセグメントSegFに対して継承関係を結ぶ。

1) SegCの方向を ψ_c , SegFの方向を ψ_f として,

$$|\psi_c - \psi_f| < \Delta \psi.$$

2) SegFの midpoint から SegC までの最短距離を h として,

$$h < \Delta h.$$

また、1段粗い解像度の特徴点PtCは次の条件1), かつ, 2)を満たす特徴点PtFに対して継承関係を結ぶ。

1) PtCとPtFの型が同じ。

2) PtFからPtCまでの距離を h として,

$$h < \Delta h.$$

但し、ある対応単位が1段粗い解像度の2つ以上の対応単位から継承関係を結ばれた場合は、より関係の強いものを1つだけ選択する。

継承関係が結ばれた対応単位に対する対応仮説は、粗レベルの活性度が密レベルのユニットの外部入力となるように結合される。左画像のレベル n における対応単位FLと1段粗いレベル $n+1$ の対応単位CLが継承関係を持つ場合、FLを含む対応仮説(FL, FR)に相当するユニット i の外部入力 η_i は次式で表される。

$$\eta_i = \varepsilon^+ \varepsilon_f \sum_{D_j \leq f} p_j s(D - D_j, 1; c_f) \quad (3.12)$$

ここで、 ε_f は粗レベルからの情報の重みを調節する係数、 D は対応仮説(FL, FR)の視差、 p_j はCLを含む対応仮説の活性度で D_j はその視差とする。また、外部入

力も興奮性結合の一種と考え、興奮性結合の重み ε^+ をかけている。

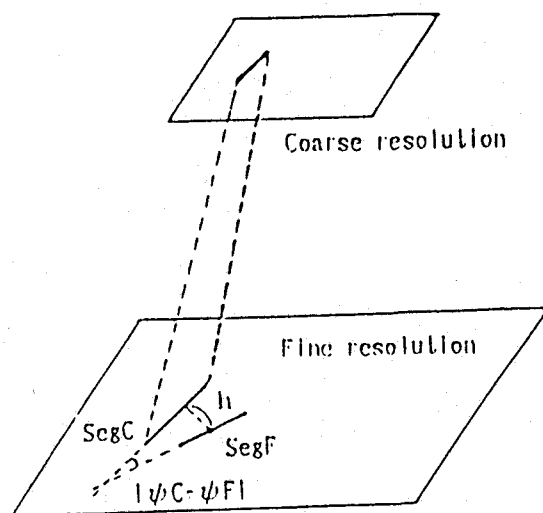


図 3.12 複数解像度の結合

第5節 実験結果

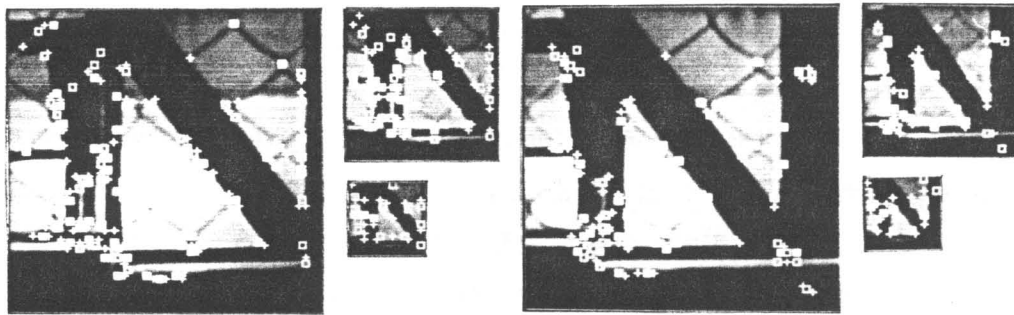
本研究で構成した両眼立体視モデルの能力を確認するために、いくつかの自然画像を対象として実験を行った。実験では、特に不連続の多いシーンに対する能力を調べるため、フェンスやネット越しに物体がある場合を主に扱っている。なお、活性度の上限は1000、エッジセグメントの対応仮説ユニットの初期値は全て200、特徴点の対応仮説ユニットの初期値は全て300としており、収束時に700以上のユニットを対応結果とした。

まず、図3.11(a)のようなフェンス越しのシーン(抽出した特徴点を重ね合わせて表示している)に対して、対応の誤り率と見逃し率により、特徴点、輪郭の連続性、及び、複数解像度の結合の効果の定量的評価を行った。図3.11(b)がセグメント画像で、これらの正しい対応をマニュアルで求め、これを基に、各種条件で実験を行った結果を表3.1に示す。各条件とも、抑制性結合と興奮性結合の重みを調節するパラメータは、その条件において、誤り率、及び、見逃し率が最良になるように調節している。図3.12に各条件での、誤り、及び、見逃し箇所を示す。図3.12から分かるように、各相互作用の組合せにより、左側の2層の表面が重なっている部分で大きな改善が見られる。このことから、より曖昧さの大きい状況において、本研究で組み込んだ各種相互作用が有効に働いていることが確かめられる。また、図3.13に、すべての相互作用を組み込んだ条件での活性度の変化の様子を示す。図3.13は、左画像の各エッジセグメントの対応仮説の中で最も活性度の高い仮説の活性度を明度に変換して表示している。即ち、左視線方向の基本競合系の中で、最も活性度の高い仮説の活性度が表示されている。曖昧さの少ない右側部分では早々に収束しているが、曖昧さの大きい左側部分では、このシーンの場合、特に、粗い解像度からの情報が有効に影響を及ぼしている。図3.13から分かるように、密レベルの左側において、繰り返しの初期では同じレベルの近傍からの情報では判断ができず基本競合系を構成するどの対応仮説も活性化しないが、粗レベルからフェンス越しの物体に関する信頼性の高い情報が伝播するにしたがって、順次活性化していくという最良優先探索的な動きを見せている。即ち、これは、注意の集中が予め与えられたヒューリスティックによるのではなく並列相互作用を通して適応的に行われる、という並列分散処理の利点を実証するものとなっている。また、このとき、手前のフェンス部分の一部(図3.12対応結果に於てフェンス越しの縦長物体の下方、及び、中程やや上にある)は、それに影響を受けず正しい対応に導かれている。

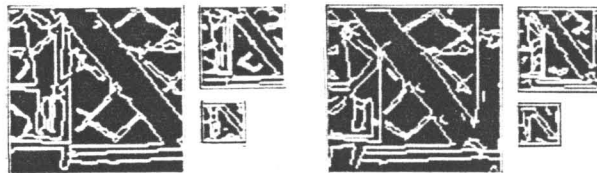
しかし、基本的に交錯した部分に於て手前のフェンス部分の対応はあまり得

られていない。これは、特徴抽出の段階で、左右で対応すべき特徴を同時に捉えられない部分が多いためである。この問題は、この様な複雑なシーンを扱う場合、ある程度避けることのできない問題であり、積極的に解決策を検討する必要がある。この問題の解決は、単眼手がかりの群化規則が重要な役割を果たすと考えられる。群化規則は同一の物理現象に対応する可能性の高い特徴を関係づけるもので、対応の定まった特徴からの視差情報伝播(内挿)の法則を与える。その様なメカニズムも、ネットワークの相互作用に自然に統合されるものと考えられる。

図3.14と図3.15にそれぞれ、兎の置物、及び、ネット越しの葉っぱの各画像の実験結果を示す。また、表3.2に実験結果の要約を示す。但し、兎の置物については、コントラストの極性の整合性が成立しない部分(左下部分)が多く、この規則を対応規則から省いて実験を行った。さらに、図3.16、及び、図3.17にフェンス越しのシーンと兎の置物について、検出された視差の立体表示を示す。なお、システムはVAX11/780上にインプリメントされ、C言語でかかっている。



(a) 入力画像と特徴点
 (□:凸点, ■:凹点, +:鞍点)



(b) エッジセグメント画像

図 3.13 フェンス越しのシーン

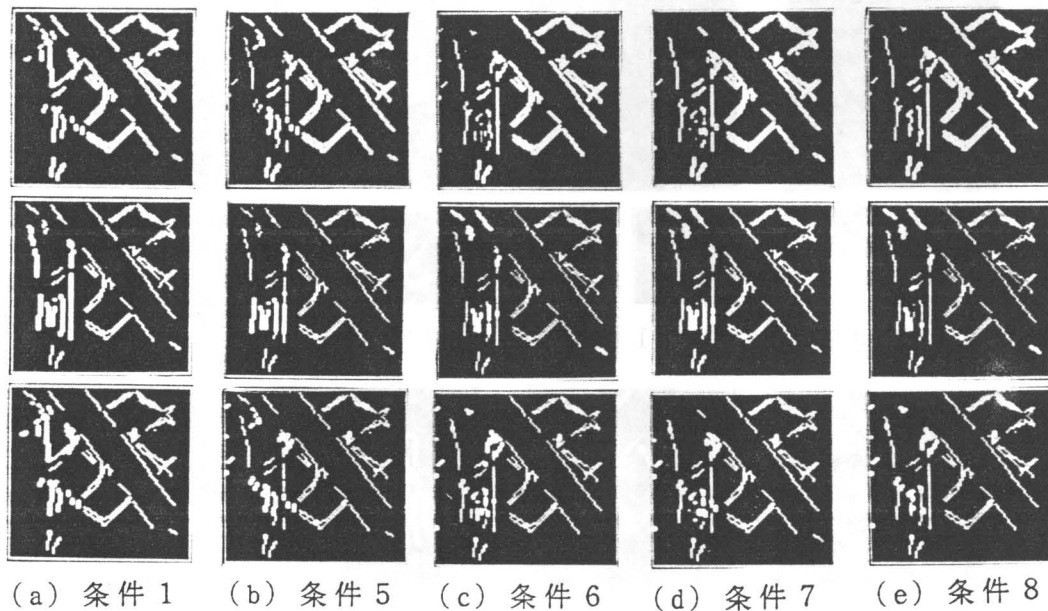


図 3.14 各条件での対応結果と見落とし，誤り箇所

(上から，対応結果，見落とし箇所，誤り箇所．対応結果は視差を明るさで表示し，見落とし箇所は，正しい対応結果のうち，見落とし箇所を太線で表示し，誤り箇所は，求められた全ての対応のうち，誤り箇所を太線で表示している)

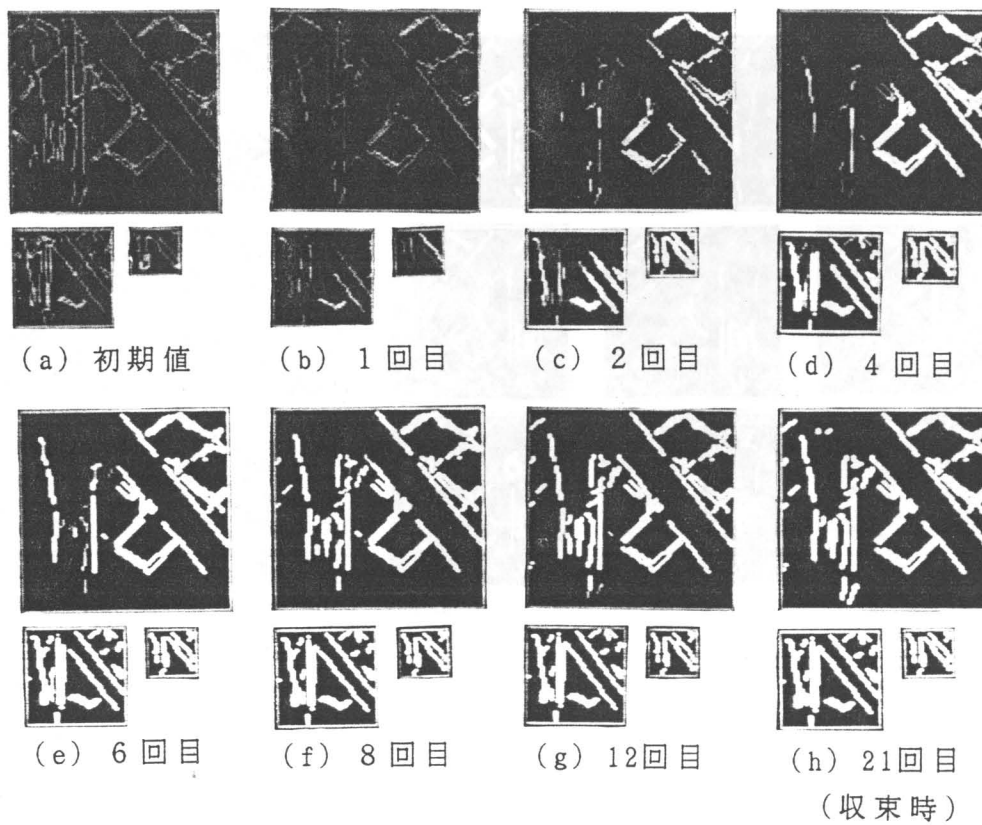
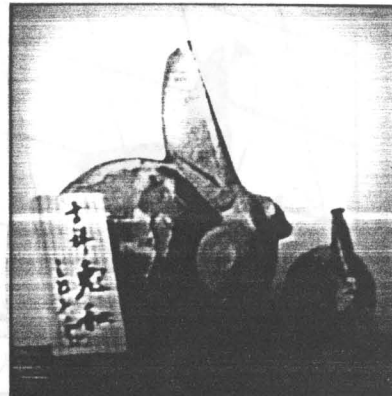
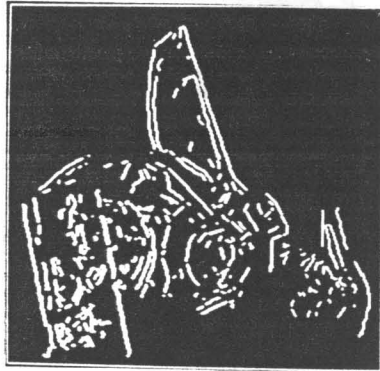


図 3.15 活性度の変化

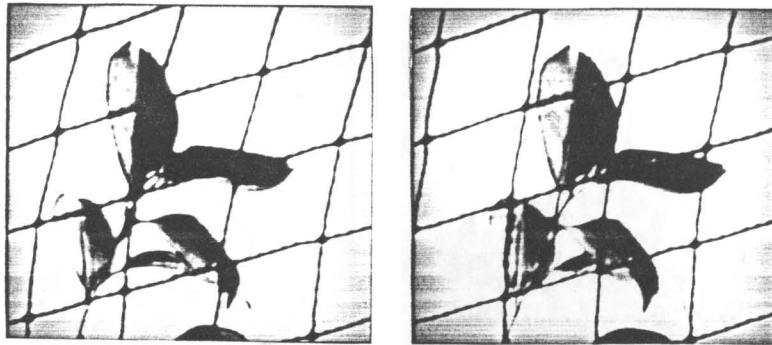


(a) 入力画像

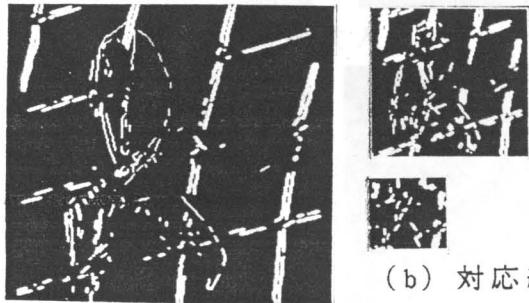


(b) 対応結果

図 3.16 兎の置物



(a) 入力画像



(b) 対応結果

図 3.17 ネット越しの葉

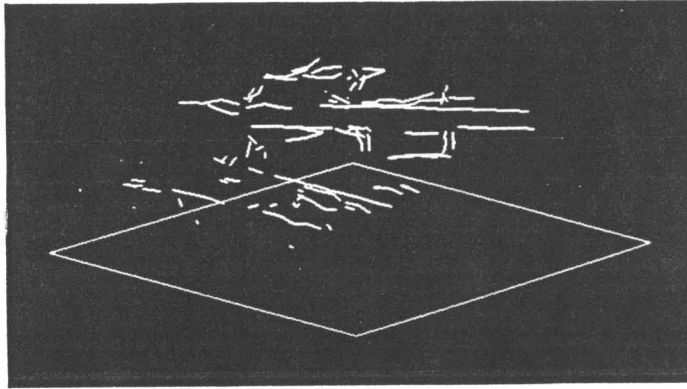


図 3.18 “フェンス越しのシーン” 立体表示 (左下から)

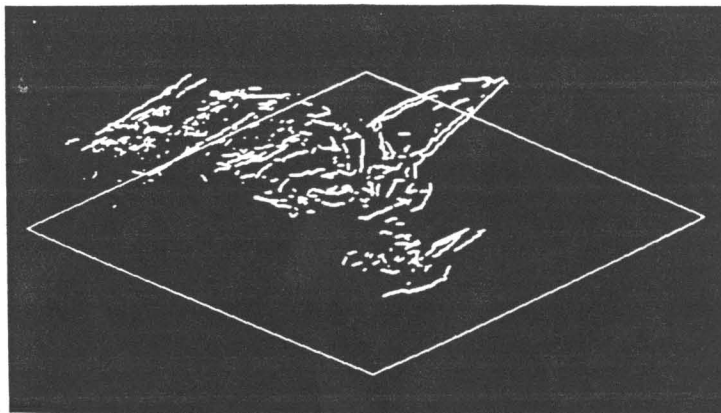


図 3.19 “兎の置物” 立体表示 (右下から)

表 3.1 フェンス越しのシーンに対する実験結果

条件番号		1	2	3	4	5	6	7	8
対応 単位	エッジ	●	●	●	●	●	●	●	●
	特徴点		●			●	●		●
対応 規則	一意性	●	●	●	●	●	●	●	●
	密着性	●	●	●	●	●	●	●	●
	連続性			●		●		●	●
	解像度				●		●	●	●
見落し率(%)		24.3	19.8	15.3	16.0	15.3	10.5	11.2	8.6
誤り率(%)		22.2	17.6	18.5	19.6	17.8	15.5	15.0	13.4

ただし, ●印はその作用を組み込んでいることを表す。
 見落し率=(見落した正しい対応の数)/(正しい対応の数)
 誤り率=(誤った対応の数)/(求められたすべての対応の数)

表 3.2 実験結果の要約

	フェンス	兎の置物	ネット越し
画像サイズ(縦,横)	100, 100	300, 300	280, 270
視差範囲	17	38	62
許容最大視差	±20	±30	±40
許容最大視差勾配	±30°	±20°	±20°
支持の近傍(w, f)	20, 3	30, 3	30, 3
エッジ数(左,右)			
解像度レベル1	304, 308	889, 959	861, 808
解像度レベル2	106, 117	286, 298	348, 343
解像度レベル3	45, 39	104, 102	153, 130
特徴点数(左,右)			
解像度レベル1	181, 129	487, 586	1012, 985
解像度レベル2	115, 73	206, 209	308, 280
解像度レベル3	54, 60	96, 102	159, 134
活性化仮説数/全仮説数(%) レベル1 (エッジ,特徴点)	42.9, 26.0	22.2, 77.0	43.9, 66.1
CPU時間(分)			
特徴抽出	5	33	26
対応付け	15	112	24

各パラメータの設定値：
 $c_c=0.6, c_r=6, c_f=0.6, \alpha_{max}=30, \beta_{max}=10, \Delta h=2, \Delta \psi=10$
 各解像度レベルに於けるガウシアンフィルタの標準偏差値：
 レベル1： $\sigma_1=2.0$, レベル2： $\sigma_2=4.0$, レベル3： $\sigma_3=8.0$

第6節 結言

本章では、互いに交錯した表面、或は、透過性表面などを含む不連続の多いシーンに対しても効果的に働く両眼立体視モデルについて述べた。ネットワークのエネルギー最小化モデルの利点は、種々の対応規則を容易に付加できることであり、これにより、様々な角度から対応の最適性を定義することができ、曖昧さの解消能力を強力にできることを示した。特に、複数解像度レベルの結合に対して、隣接するレベルで同一の物理現象に対応する特徴のみ情報を伝播することにより有効に多重解像度情報を用いる方法を提案した。また、実験により、モデルに埋め込んだ、特徴点がエッジセグメントに与える曖昧さの解消効果、輪郭の連続性に基づく相互作用の効果、及び、複数解像度の情報伝播の効果を確認した。

本章で述べたモデルは、結局、Marr-Poggio第1アルゴリズム[3.2]の自然な拡張となっている。但し、本研究では、各ユニットはアナログ素子で、ネットワークの挙動をエネルギー極小化の大局的最適化過程と捉えている点で異なる。Marr & Poggioは彼らの論文[3.14]の中で神経回路網の協調性処理の自然画像への有用性を否定したが、本研究は、対応単位と対応規則に関して自然画像の群化特性に適したものを用いることにより、複雑な自然画像に対して強力なモデルとなることを示している。

また、工学的にみた場合、本モデルにおける問題点は、対応規則の重みを定める各パラメータの、わずかな変動が、結果に大きな影響を及ぼす危険性である。実際、実験では、抑制性結合と興奮性結合の重みの割合を調節する係数 ϵ^- と ϵ^+ の設定が結果に少なからず影響している。 ϵ^- を小さくすると時期尚早な解釈に陥る傾向にあり、逆に大きくすると正しい対応を見逃す消極的な解釈に陥る傾向にある。1つの方法としてネットワークの動きを監視しつつ、適宜、この重みの調節を行うプロセスの導入が考えられる。また、解の探索法の観点からみると、必ずエネルギー関数を最小にする解を求めることができれば、重みの調節はそれほど大きい問題ではない。ボルツマンマシン[3.20]は、その問題を解決する1つの方向であろう。

第4章 パラメータ化モデルを用いた 3次元物体認識

第1節 序言

本章では、雑音や隠れの多い状況において、単眼の2次元画像から3次元物体認識を行うための並列分散処理アプローチについて述べる。但し、従来の研究の多くが、物体モデルの大きさ、形状が固定であったのに対して、物体モデルには、いくつかのモデルパラメータによって表される可変性を許している。これにより、類似の構造を持つ物体の集合を1つの物体モデルで認識することが可能になる。

認識を目的とした可変性を含む3次元物体モデルの表現法の検討は従来からなされているが、従来の研究では、認識実験は2次元においてのみ行われているにすぎない[4.1],[4.2]。基本的に、可変性を許すことによって、解の拘束が弱くなり、それだけ強力な探索機構が必要になる。並列分散処理は、逐次的手法に比べ、弱い証言しか利用できない場合にもそれらを集約することによって強力な認識が期待できる。

物体認識の問題は、画像特徴と物体モデルの対応関係を見だし、同時に、物体の位置・姿勢を表す変換パラメータ、及び、物体モデルの可変要素を定めるモデルパラメータを決定することと捉えられる。画像特徴と物体モデルの一部との対応を仮定すると、変換パラメータとモデルパラメータを拘束する連立方程式が得られる。連立方程式が解を持つとき対応が可能で、対応仮説が生成される。このとき、対応部分が多いほどパラメータを強く拘束し、対応の信頼性も高くなる。認識過程は、整合の取れる対応を統合によって証言を蓄積し、同時に矛盾を取り除いていく過程である。既存の認識手法は、証言の蓄積法の違いによって分類できる。

1つの方法は、信頼性の高い対応仮説を足掛りに、画像特徴の存在を予測して逐次的に証言を蓄積する方法である。認識の効率は初期仮説の信頼性に依存するので、群化した画像特徴と構造を持つモデルプリミティブとの対応によりパラメータを強く拘束する初期仮説を生成する必要がある。Lowelは、視線の変化に不変、かつ、あらゆる物体モデルに含まれる比較的大きな構造を反映する偶然に起きそうにない画像構造の群化過程が、物体認識において重要な役割を果たすことを強調している[4.3]。しかし、隠れや雑音が多い場合には、物体モ

デルに1体1対応するような画像構造の安定な抽出は期待できない。

証言の蓄積を並列に行う方法として、一般化Hough変換が物体認識に適用され成功を収めてきた[4.4],[4.5]。対応単位は比較的単純なものが用いられるので、雑音や隠れに対して強くなる。しかし、あらゆる可能な対応候補(仮説)から得られる拘束をパラメータ空間へ投票するという単純な証言の蓄積法では、モデルパラメータを多く含む物体モデルの認識に対しては、莫大なパラメータ空間を必要とし、種々の場当りの方法でそれを回避する必要がある。また、物体が複数個存在する場合のパラメータ空間のピークの判断基準、各仮説間の矛盾関係が明確化されない等という問題点がある。

弛緩法(Relaxation labelling processes)[4.6]は、仮説間の並列相互作用によって証言の蓄積するという点で、Hough変換よりも強力なメカニズムであると考えられる。しかし、従来行われた物体・形状認識への応用[4.7],[4.8]では、認識に先立つ認識対象のセグメンテーションが仮定されている。認識とセグメンテーションは同時に解決されるべき問題であり、弛緩法によるアプローチではこれをどう解決するかが問題となる。また、入力データとモデルの次元整合が仮定されており、さらに、モデルパラメータを扱った研究は行われていない。

近年、多くの初期視覚(early vision)の問題が拘束満足ネットワークのエネルギー関数を最小化する最適化問題として定式化されている[4.9]。Hopfieldの提案した“エネルギー最小化の原理”[4.10]は、視覚の問題特有の、最適性自体が漠然とした基準でしか表せない最適化問題を、神経回路網類似のネットワークを用いて解く強力な並列計算モデルである。本質的には、弛緩法と等価である[4.6]が、より簡潔、かつ、洗練された定式化となっている。本研究は、初期視覚だけでなく物体認識においても、このモデルが強力な探索手段となることを実証するものである。Hopfieldのモデルは局所的な画像特徴から得られる可能な解釈(即ち、仮説)間の組織的な相互作用のメカニズムを与える。逐次型の探索法と比べた場合、探索の順序を制御する必要がなく、あらゆる情報が同時に相互作用を及ぼし合うことによって、雑音や隠れに強く、かつ、各場合に適応した最良優先探索的な処理が行えるという特長がある。本章では、拘束満足ネットワークによって、認識対象の認識と分離を同時に行なう方法について述べる。

以下では、まず、第2節で、並列分散処理に適した物体モデルの表現、及び、物体についての制約について述べ、第3節で、認識過程の概要を述べ、第4節～第6節において、認識過程の各段階について詳述し、第7節では、実験結果を示す。

第2節 並列認識のための物体モデル表現

2.1 モデルプリミティブの選択

物体モデルは、画像特徴と照合を行う際に対応の基本単位となるモデルプリミティブ、及び、それらの間の関係構造によって表現する。このとき、モデルプリミティブとして何を用いるかは、雑音・隠れに対する強さ、認識効率、及び、認識可能な物体のクラスにおいて重要な要因となる。本研究では、モデルプリミティブとして、コーナーを用いる。コーナーは物体の2辺の接続部分であり、1点とその点を始点とする2方向ベクトルで表現される。コーナーの画像投影であるL型接続点は、局在化した特徴であるので雑音・隠れに対して安定に抽出が行え、3D from 2Dの対応を拘束する比較的多くの位置・方向情報を含んでいる。また、コーナーを用いることは、必ずしも認識可能な物体を多面体に限定しない。即ち、認識に重要な主特徴が、コーナーとその関係で表現できればよい。

2.2 物体モデルの可変性とその表現

物体モデルの幾何学的拘束は、モデルプリミティブ自身の拘束と各プリミティブ間関係の拘束によって表される。形状の可変性は、これらの幾何学的拘束に対する自由度の与え方によって表現する。

並列認識過程による効率のよい利用を考慮して、各モデルプリミティブ毎にそれを中心としたプリミティブ座標系 (x_p, y_p, z_p) を導入する。各プリミティブ座標系毎に

- (1)プリミティブ自身の拘束,
- (2)近傍のプリミティブ座標系との関係に関する拘束,
- (3)物体座標系との関係に関する拘束

が格納され、物体モデルはそれらの集合として冗長に表現される。Hough変換では、(1)と(3)の情報を利用して認識を行うが、本研究では、対象の切り出しと認識において(2)の情報が最も重要な役割を演じる。以下では、各拘束条件の表現と可変性の制約について述べる。

(1)プリミティブ自身の拘束：図4.1に示すように、コーナーを定義する2ベクトルの交点をプリミティブ座標系の原点とする。2つのベクトルは方向が固定、或は、1自由度の方向の可変性を許す。可変方向ベクトル1は次式のように2つの固定方向ベクトル l_1 と l_2 の線形和で表されるものとする。

$$l = \lambda l_1 + (1-\lambda)l_2 \quad (4.1)$$

このとき、 λ の可変範囲が割り当てられる。

(2)近傍のプリミティブ座標系との関係に関する拘束：本研究において、プリミティブ座標系間の可変性は、平行移動のみに制約する。あるプリミティブが、物体モデル中の他のプリミティブの中で、関係を記述するプリミティブ(以下、結合プリミティブと呼ぶ)は、例えば“平行移動が2つの固定方向ベクトルの線形和、或は、それ以内の自由度で記述できる”プリミティブ(以下、近傍プリミティブと呼ぶ)とすることができる。本研究では、近傍プリミティブとの関係を以下の3種類に分類する(図4.2)。

(a)平行移動 t が2つの固定方向ベクトル t_1, t_2 の線形和で表現できる。即ち、平行移動が、 α, β をモデルパラメータとして、 $t = \alpha t_1 + \beta t_2$ と記述できる。

ここで、 α / β の可変範囲の割当が行えるものとする。

(b)平行移動 t が1つの固定方向ベクトル t_1 の定数倍で表現できる。即ち、平行移動が、 α をモデルパラメータとして、 $t = \alpha t_1$ と記述できる。ここで、 α の可変範囲の割当が行えるものとする。

(c)同じ原点を共有する。即ち、 $t = 0$ と記述できる。

本研究における認識過程の立場からは、あるプリミティブに対して、必ずしも全ての近傍プリミティブを結合プリミティブとする必要はないが、関係記述の対称関係(プリミティブ p_1 がプリミティブ p_2 の結合プリミティブであるならば、同時に、プリミティブ p_2 はプリミティブ p_1 の結合プリミティブである)は保たれていることが望ましい。

(3)物体座標系との関係に関する拘束：(2)の制約から、物体中心座標系から各プリミティブ座標系への変換 T_{OP} は、平行移動成分 t_{OP} のみからなり、次式のようにモデルパラメータ($P_1, P_2, P_3, \dots, P_n$)の線形結合によって表される(図4.3)。

$$t_{OP} = (P_x, P_y, P_z) = \left(\sum_i a_i P_i, \sum_i b_i P_i, \sum_i c_i P_i \right) \quad (4.2)$$

(1)~(3)の記述は、図4.4のようなワイアフレームモデルから容易に構成することができる。また、認識に先立つ視覚による学習過程から獲得することも考えられる。

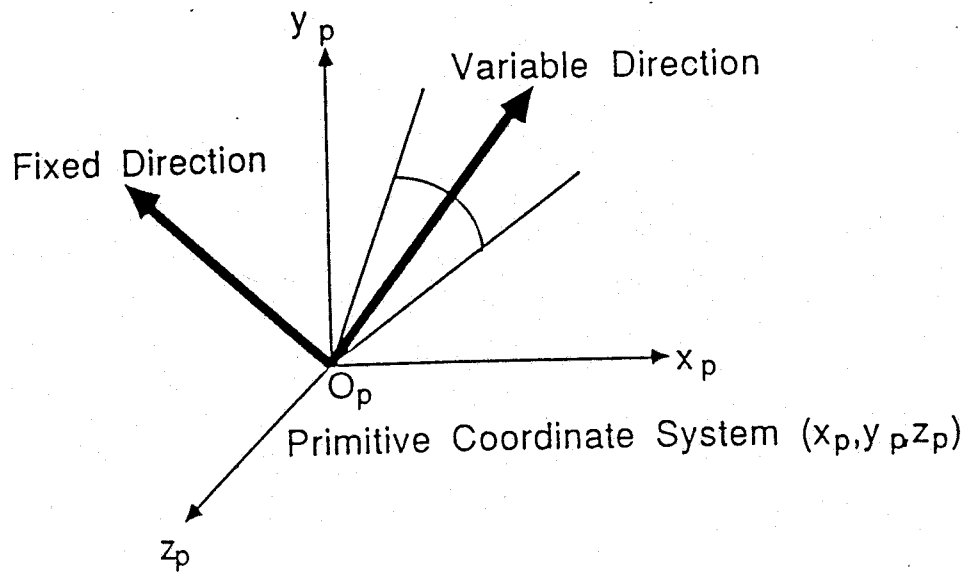


図 4.1 モデルプリミティブの可変性

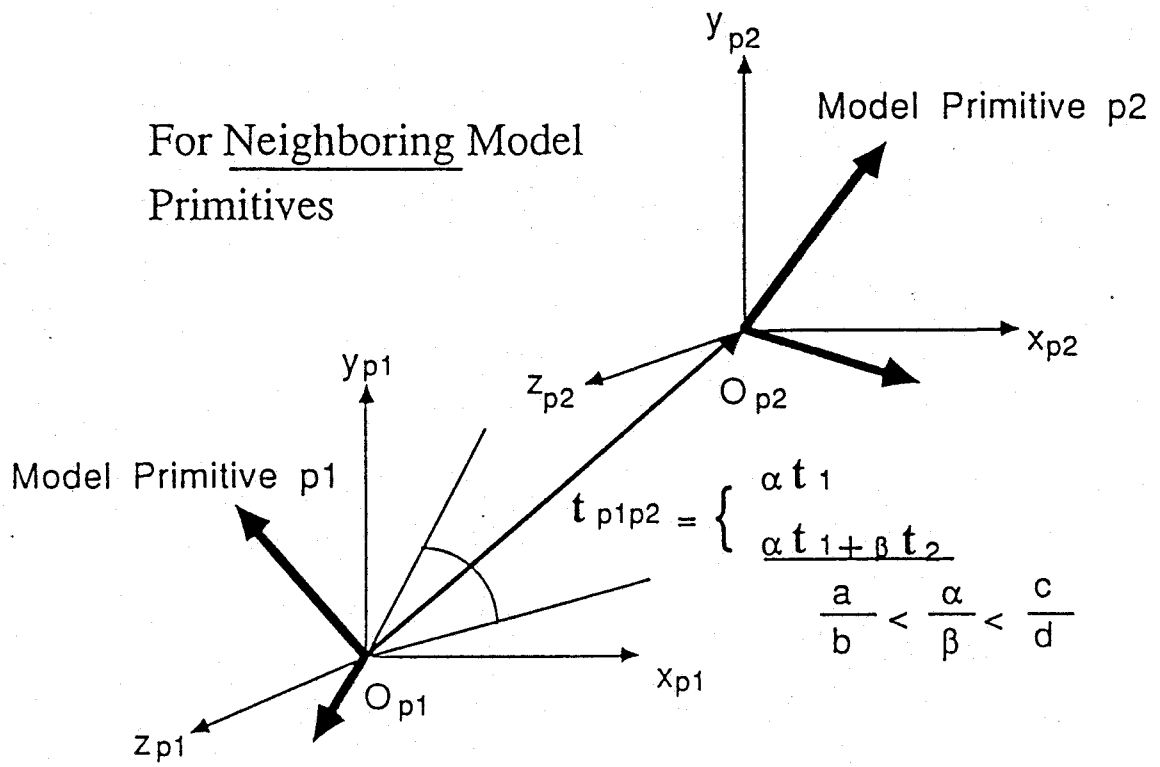


図 4.2 近傍プリミティブとの関係の可変性

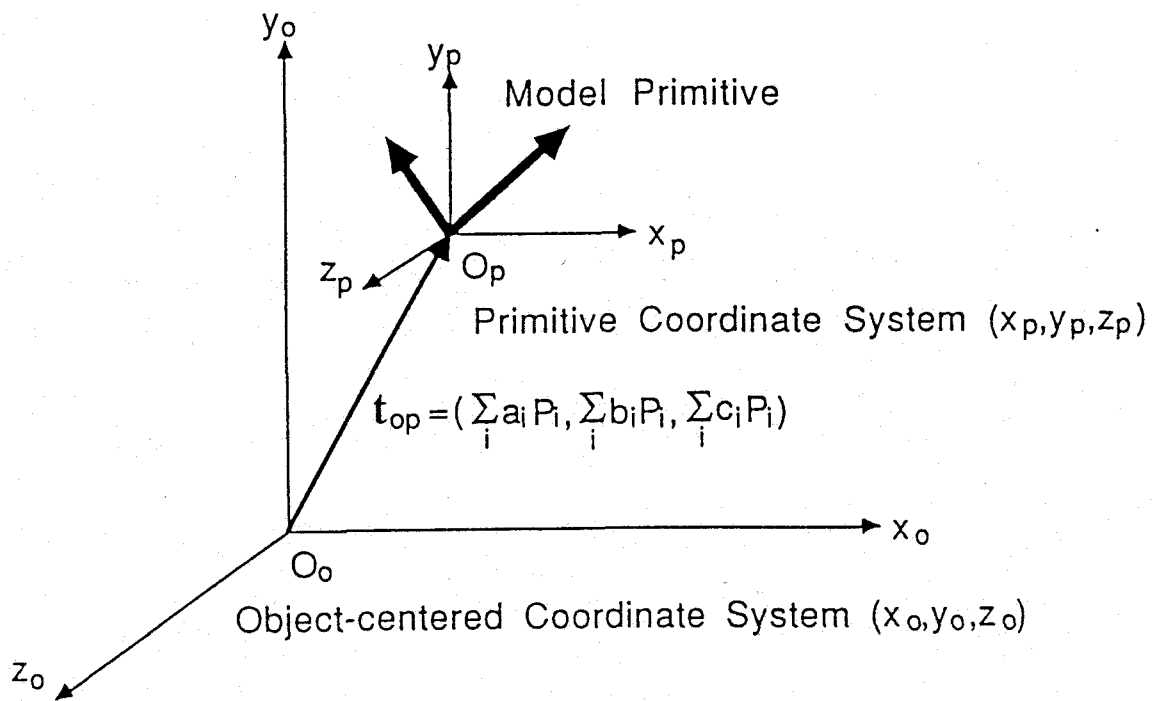


図 4.3 プリミティブ座標系と物体中心座標系の関係

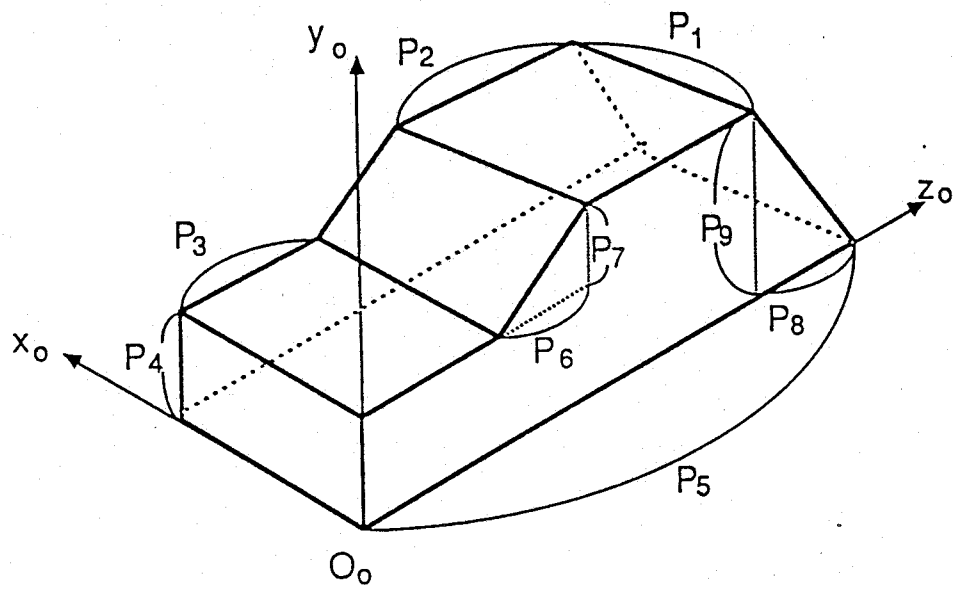


図 4.4 パラメータ化モデルの例

2.3 物体の姿勢の制約

一般に、物体の3次元空間中の位置・姿勢は、6自由度(平行移動3、回転3)あるが、物体のとりうる安定姿勢を考慮にいれ、回転軸を鉛直方向に限ると平行移動とともに4自由度で表現できる。この拘束は、現実世界での物体の位置・姿勢を考えてもそれほど厳しいものではない。そこで、簡単化のため変換パラメータはこれら4つのパラメータで記述する。しかし、人間が、画像から得られる物体の姿勢の推定値と物体の安定姿勢の知識から重力方向を推論する能力を持っている(即ち、視覚が重力方向を感知するセンサになっている)ことを考慮すると、この制約は理にかなったものではない。

ワールド座標系 (x_w, y_w, z_w) 、物体中心座標系 (x_m, y_m, z_m) 、画像平面座標系 (U, V) を図4.5のように設定する。 y_w 軸を鉛直方向とすると、物体中心座標系からワールド座標系への変換 T_{ow} は平行移動 $t_{ow}=(t_{owx}, t_{owy}, t_{owz})$ と y_w 軸回りの回転 θ_{ow} の4つのパラメータで記述できる。ワールド座標系から画像平面座標系への変換 T_{wi} は中心投影を仮定し、レンズの中心(視点)はワールド座標系の原点、光軸は y_w-z_w 平面上とし、俯角 α 、及び、焦点距離は既知とする。画像平面座標系は、焦点距離が1となるようにスケーリングされているとする。

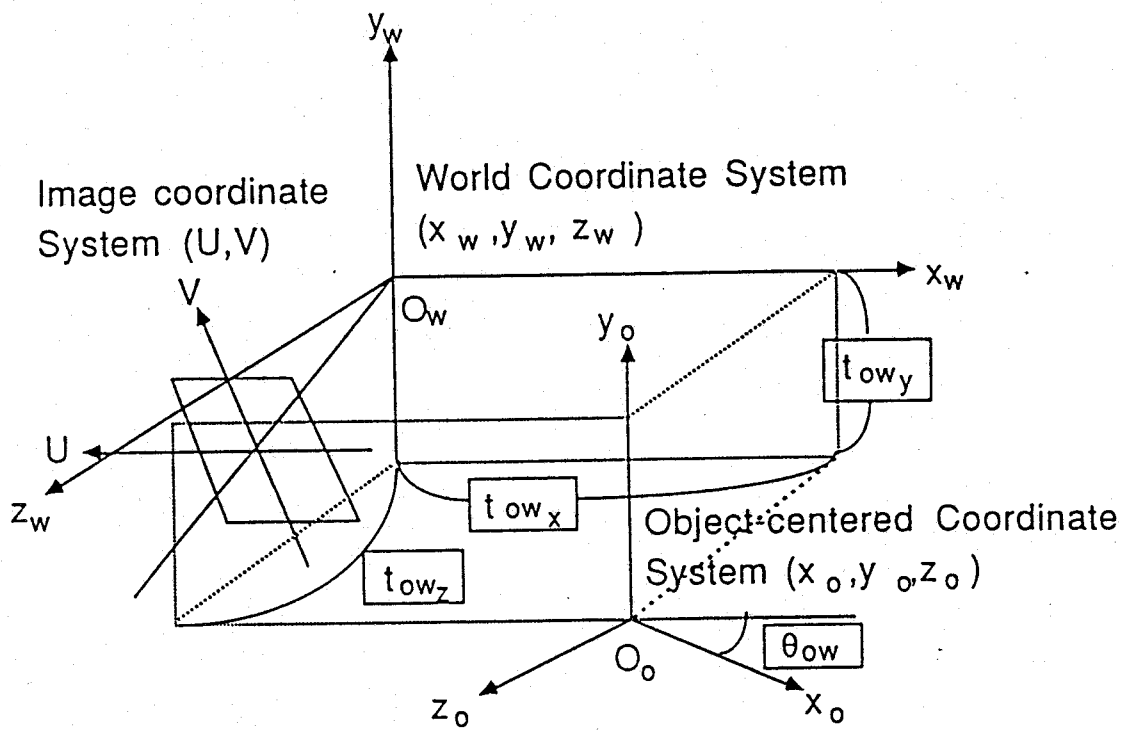


図 4.5 物体の姿勢の制約

第3節 並列認識過程の概要

本研究では、複数の物体モデルを仮定し、さらに画像中に認識可能な物体が複数個投影されているものと仮定する。

問題は、画像特徴と物体モデル間の最適対応を見だし、認識された各物体に対して T_{ow} の4つのパラメータとモデルパラメータを決定することである。最適対応を見出す問題は、画像特徴とモデルプリミティブの全ての可能な対応候補の中で、最適集合を選択するという組合せ最適化問題として定式化される。個々の物体に相当する対応の部分集合から、変換パラメータとモデルパラメータの推定が行える。

認識過程は、まず、画像特徴とモデルプリミティブのあらゆる対応の組合せを考え、物体モデルと画像生成過程の拘束を用いて、グルーピングとフィルタリングを行う過程であり、次の3段階に分けられる。(図4.6)

(1)対応仮説の生成：画像特徴とモデルプリミティブのあらゆる対応の可能性を調べ、対応が可能な場合、対応仮説が生成される。この段階で、モデルプリミティブの拘束と画像生成過程の局所的な拘束により対応候補が絞られる。

(2)仮説間の相互作用による拘束フィルタリング：各対応仮説の中から、各プリミティブ間の関係から得られる拘束条件を最もよく満足し、互いに整合のとれる対応仮説クラスタを抽出する。ここで、1つの対応仮説クラスタが、1つの物体モデルのインスタンスに相当する。

(3)変換パラメータとモデルパラメータの推定：対応仮説クラスタに含まれる対応関係より、各プリミティブと物体中心座標系との拘束を利用して変換パラメータとモデルパラメータを推定する。

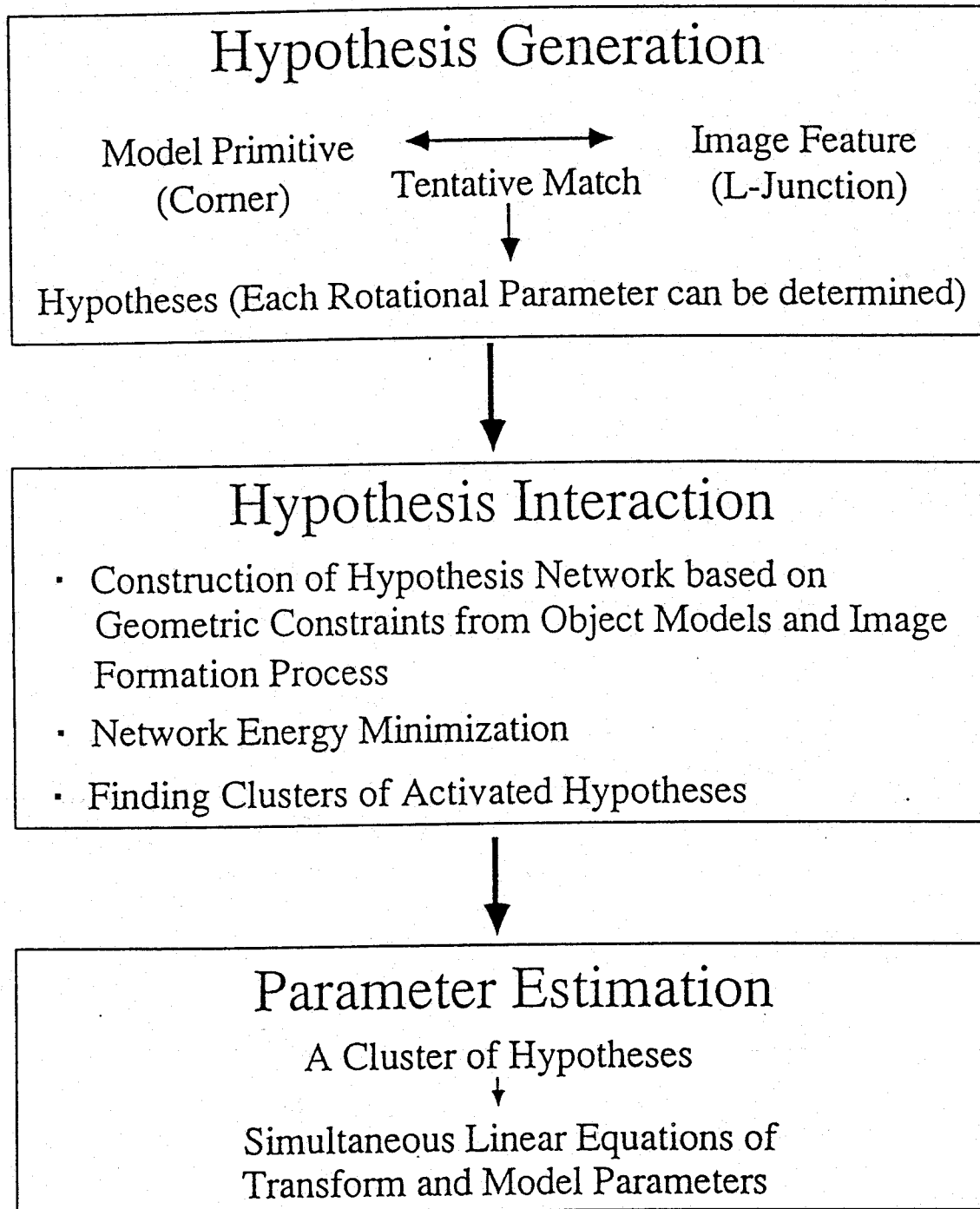


図 4 . 6 認識過程

第4節 対応仮説の生成

本節では、3Dコーナーと2DL型接続点との間の対応の拘束について考察する。付録2よりコーナーとL型接続点との1点と1方向の対応が可能な場合、プリミティブ座標系からワールド座標系への変換 T_{PW} の回転成分 θ_{PW} は特殊な場合を除いて2通りに定まる。また、 T_{OP} は平行移動成分のみであるので、 θ_{PW} は θ_{OW} に等しい。特殊な場合で重要な場合は、モデルプリミティブの辺がx-z平面に垂直(以下、単に垂直と呼ぶ)な場合であり、これは現実世界に多く存在する。辺が垂直な場合は θ_{PW} は定まらないが対応可能な条件は示されている。付録3より、方向が可変の辺は θ_{PW} が決定するとその方向が決定する。

以上より、モデルプリミティブは次の5種類に分類される。

- ① 2辺とも方向が固定でしかも垂直でない。
- ② 2辺とも方向が固定であるが1辺は垂直。
- ③ 1辺の方向が可変で他辺は垂直でない。
- ④ 1辺の方向が可変で他辺は垂直。
- ⑤ 2辺とも方向が可変。

①では、2辺のそれぞれが対応可能で、しかもそれぞれに対して得られた θ_{PW} の値が等しい(十分近い)ときのみ対応可能である。②では、2辺のそれぞれについて解が存在するとき対応可能で、垂直でない辺の対応から θ_{PW} が得られる。③では、方向が固定の辺の対応が可能な場合 θ_{PW} が得られるので、その θ_{PW} よりもう一方の辺の方向が決定する。但し、その方向が予め決められた可変範囲内にある場合のみ対応可能とする。①②③の場合について、対応が可能なとき、

対応仮説 $H_i(ML_j, MP_k, IF_l, \theta, \lambda)$

が生成される。ここで、 ML_j はモデルラベル、 MP_k はモデルプリミティブ、 IF_l は画像特徴、 θ は対応によって決定された T_{PW} の回転パラメータ、 λ は MP_k の1辺が可変の場合の対応によって決定された方向である。

④、及び、⑤では θ_{PW} の推定が行えないので、モデルプリミティブとして考えない。即ち、物体モデルにおいて④や⑤のパターンを多く含む場合、それだけ認識能力が落ちることになる。

第5節 仮説ネットワークによる対象の分離と認識

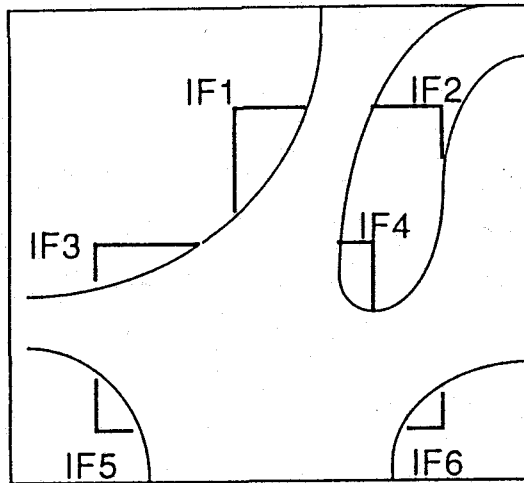
生成された仮説間の並列相互作用によって、仮説のグルーピングとフィルタリングを行う。処理は2段階からなり、各段階でHopfieldモデルに基づく拘束フィルタリングを行う。

まず、1段階目において、認識された各物体に相当する仮説のクラスタを抽出する。この処理は、処理単位として“仮説対”を用いて全ての仮説に対して適用される。処理単位として単独の仮説ではなく“仮説対”を用いるのは、画像中に複数の認識可能な物体が存在する場合、単独の仮説間の関係、即ち、2項関係では、それらの分離が不可能な場合があるためである。単純な2次元マッチングの例として図4.7(a)の画像と(b)のモデルが与えられ回転を考えない場合、画像中の認識可能な物体が高々1つという仮定のもとでは、図4.7(c)のようなネットワークの構成が考えられる。しかし、シーン中の複数個の認識可能な物体を仮定した場合、ネットワーク中に抑制結合されている各仮説は矛盾しているとはいえない。また、異なる物体の対応仮説が興奮性結合され、これらを生離することは不可能である。この場合、図4.7(d)のようにグルーピングの可能性のある“仮説対”間の関係を用いなければ、2つの長方形を生離できない。このことは、認識過程において、仮説を選択するだけではなく、仮説間のグルーピングの可能性を選択する過程が必要であるということを示す。

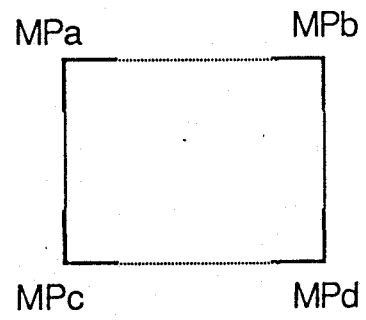
次に、2段階目において、1段階目で抽出された各仮説クラスタの検証を行う。この処理は、各仮説クラスタに対して別個に行われる。この段階では、各仮説クラスタは高々1つの物体を含むと仮定することができ、図4.7(c)と同様の拘束関係を適用できる。この段階で、仮説クラスタ中の未だなお矛盾する仮説の除去、及び、証言が不十分な仮説クラスタの除去が行われる。

5.1 “不適合仮説対”の設定

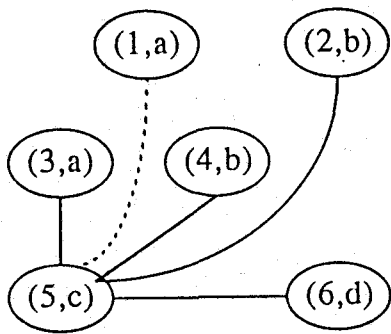
1つのL型接続点から複数の対応仮説が生成された場合(大部分の場合に当てはまる)、その中で正しい解釈は高々1つである。よって、それらの仮説は互いに両立せず、“不適合仮説対”と設定する。これは、いわゆる“一意性の拘束(Uniqueness Constraint)”である。



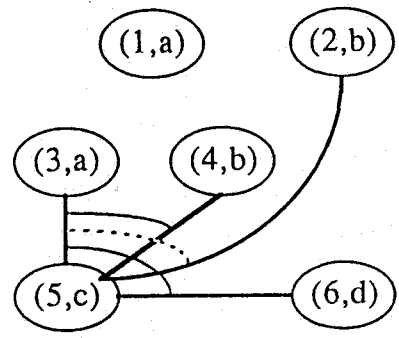
(a) Image



(b) Model



(c) Using hypothesis
(Connections for hypothesis(5,c))



(d) Using hypothesis pair
(Connections for hypothesis pair
((3,a), (5,c)))

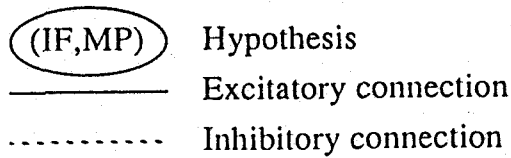
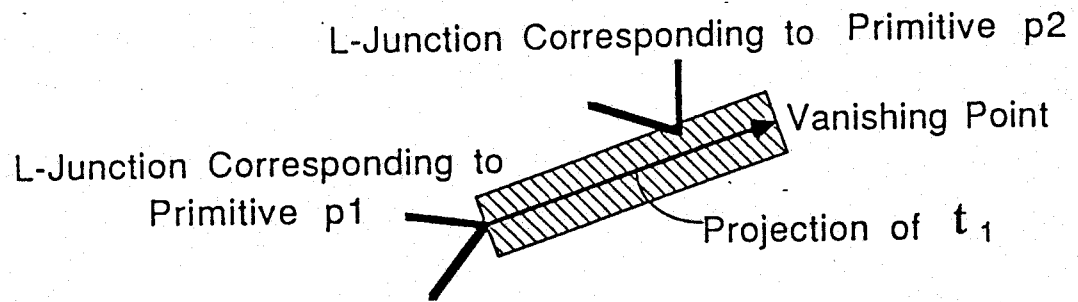
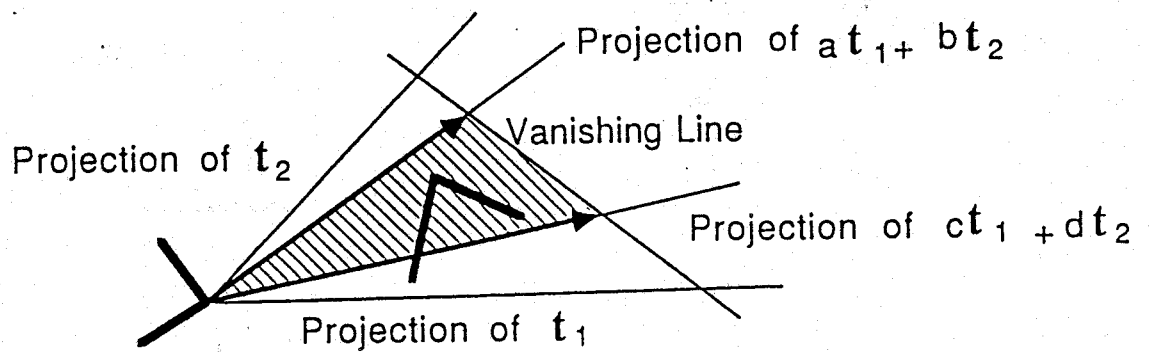


図 4.7 仮説ネットワークによる対象の分離と認識



(a) プリミティブ間が自由度1のとき



(b) プリミティブ間が自由度2のとき

図4.8 仮説間の適合条件

5.2 “適合仮説対”の設定

物体モデルにおいて近傍プリミティブの2つの対応仮説

$$H_{i1}(ML_{j1}, MP_{k1}, IF_{l2}, \theta_1, \lambda_1) \text{と}$$

$$H_{i2}(ML_{j2}, MP_{k2}, IF_{l2}, \theta_2, \lambda_2)$$

が以下の条件を満たすならば、それらは“適合仮説対”と設定する。“適合仮説対”はグルーピングの可能性を示すものである。

- (1) $ML_{i1} = ML_{i2}$
- (2) $IF_{l1} = IF_{l2}$
- (3) MP_{k1} と MP_{k2} が互いに近傍プリミティブである。
- (4) $|\theta_1 - \theta_2| < d\theta$
- (5) IF_{l1} の接続点と $(\theta_1 + \theta_2)/2$ を用いて、2.2で述べた MP_{k1} から MP_{k2} への3D平行移動の拘束を画像平面上に投影し、 IF_{l2} がその投影範囲内(ある程度の許容誤差を含む)に存在する。図4.8に、平行移動の自由度が、1、或は、2の場合の条件を示す。

5.3 “適合仮説対”間の相互作用

問題は、“適合仮説対”の中での最適な組合せを選択する問題として定式化できる。各“適合仮説対”の適合関係は、以下の条件にしたがって設定する。

(図4.9)

- (1) 2つの“適合仮説対”の仮説の中に、5.1で設定した“不適合仮説対”が存在すれば“不適合”。
- (2) もし2つの“適合仮説対” $HP_{ij}(H_i, H_j)$ と $HP_{ik}(H_i, H_k)$ が同じ対応仮説 H_i を含むならば、適合関係は次の条件により定める。
 - (a) “適合仮説対” $HP_{jk}(H_j, H_k)$ が存在すれば適合。
 - (b) H_j と H_k が同一プリミティブでなく、かつ、互いに近傍プリミティブの対応仮説でなければ適合。
 - (c) (a),(b)以外の場合、即ち、 H_j と H_k が同一プリミティブ、或は、互いに近傍プリミティブで、しかも“適合仮説対” $HP_{jk}(H_j, H_k)$ が存在しなければ不適合。
- (a)の場合は、3つの対応仮説 H_i, H_j, H_k が互いに適合する場合である。
- (b)の場合は、 H_j と H_k が互いに近傍プリミティブの対応仮説でないためそれらの適合性は問題にしていない。(c)の場合は互いに近傍プリミティブの対応仮説である H_j と H_k が適合していないということは明らかに矛盾である。

上述の適合関係に基づく相互作用をHopfieldのモデルに埋め込む。“適合仮説対”の数を n 個とし、それらの活性度を $p = (p_1, p_2, \dots, p_n)$ と表す。但し、 p_i は実数で、 Max を活性度の上限として $0 \leq p_i \leq \text{Max}$ とする。“適合仮説対”間の“適合”，及び“不適合”関係を表す $n \times n$ 対称行列を、それぞれ $R^+ = \{r^+_{ij}\}$ ， $R^- = \{r^-_{ij}\}$ とする。 R^+ の各要素は“適合”のとき 1，それ以外のとき 0，また R^- の各要素は，“不適合”のとき 1，それ以外のとき 0 と設定する。次式で示されるエネルギーを最小にする活性化パターン p が最適な対応仮説とそのグループングを表す。

$$E(p) = (-1/2) \sum_{i,j} (\epsilon^+ r^+_{ij} - \epsilon^- r^-_{ij}) p_i p_j + \sum_i \zeta p_i \quad (4.3)$$

ここで、 ϵ と ϵ^- は興奮性結合と抑制性結合の割合を調節する係数で、 ζ は大局的抑制定数である。

5.4 検証

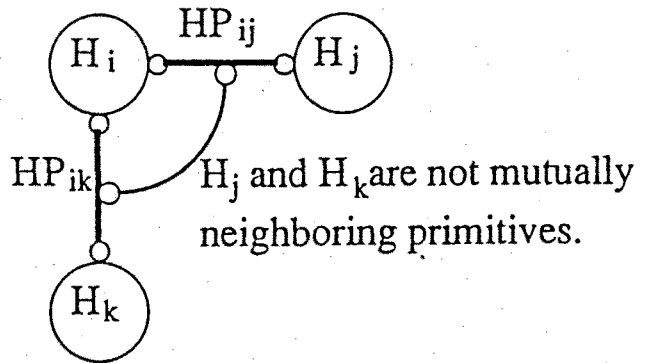
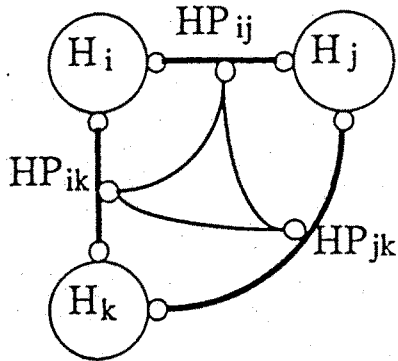
収束したネットワークから、活性化し、かつ、互いに興奮性結合している仮説対に含まれる仮説を仮説クラスタとして取り出す。このクラスタには高々 1 つの物体が含まれていると仮定できるが、未だなお矛盾する仮説が含まれていたり、証言の不十分な仮説クラスタも含まれる。そこで、個々の仮説クラスタについて“仮説”を処理単位として、再度、拘束フィルタリングを適用する。

各仮説クラスタそれぞれについて、以下の条件にしたがって仮説間の適合関係を設定する。

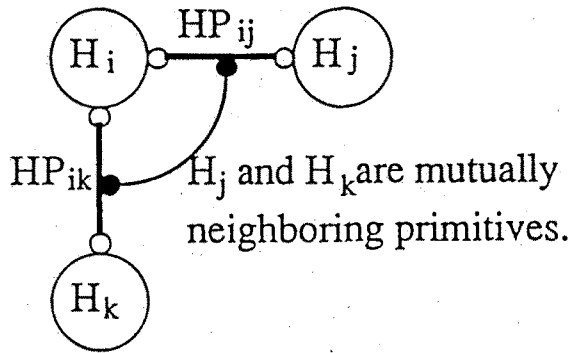
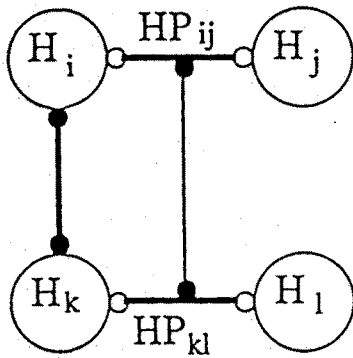
- (1) 5.1 の条件に従い“不適合”関係を設定する。
- (2) 5.2 の条件に従い“適合”関係を設定する。
- (3) 同一プリミティブの対応仮説同士に対して“不適合”関係を設定する。
- (4) 近傍プリミティブであり、かつ、“適合”関係が設定されていない仮説同士に対して“不適合”関係を設定する。

上述の適合関係に基づき、各仮説クラスタについてネットワークを構成し、式(4.3)のエネルギー関数の最小化を行う。これによって、大局的に矛盾のない解釈を導く。

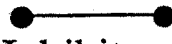
- Excitatory Connection



- Inhibitory Connection



Excitatory Connection



Inhibitory Connection

図 4.9 仮説対間の適合条件

第6節 変換パラメータとモデルパラメータの推定

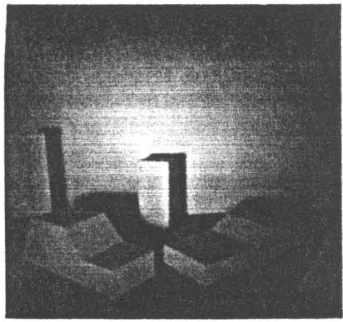
仮説クラスタの各対応仮説におけるモデルプリミティブとL型接続点の対応から変換パラメータとモデルパラメータの推定を行う。それぞれの対応仮説から T_{PW} の拘束が得られ、また各モデルプリミティブ毎に物体中心座標系からの変換の拘束 T_{OP} が記述されている。 T_{OW} の拘束は $T_{OW} = T_{OP} T_{PW}$ により得られる。

T_{OW} の回転成分 θ_{OW} に関しては各対応仮説においてすでに定まっているのでその平均をとればよい。付録4より1組のL型接続点と物体モデルの頂点との対応から、 T_{OW} の平行移動成分 $t_{OW} = (t_{OWx}, t_{OWy}, t_{OWz})$ とモデルパラメータ P_1, P_2, \dots, P_n に関する2つの線型方程式が得られる。よって、仮説クラスタのすべての対応仮説から得られる線型方程式を連立して解くことにより各パラメータが得られることになる。しかし、実際は単眼での対応からは、それらのパラメータ間の比は求まるが絶対量は求まらない。この連立方程式を解く場合の1つの問題点は、あるパラメータは拘束が不十分(underconstrained)であり、また別のパラメータは拘束が過剰(overconstrained)であるということである。この問題は、線型連立方程式に対して線型計画法における実行可能解を得る手続きを適用することによって解決される。このとき、同時に、各パラメータの変換範囲(たとえば、 $a \leq P_n \leq b, c \leq t_{PWz} \leq d$ など)、モデルパラメータの関係式(図4.4では、 $P_2 + P_3 + P_6 + P_8 = P_5, P_4 + P_7 = P_9$)も、線形式である限りその手続きに組み込むことができ、その条件を満たす実行可能解を得ることができる。得られた実行可能解は各パラメータ値の1組の可能な値を示し、それから各パラメータ間の比が得られる。

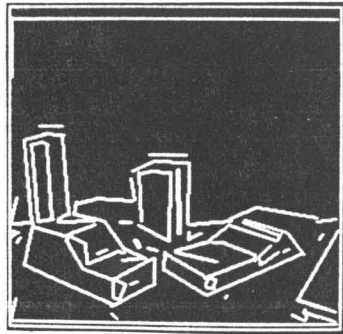
第7節 実験結果

入力画像(256×256)に対してDOGフィルタを適用し、その零交差輪郭の直線近似を行い、L型接続点を抽出する。図4.4に示したモデルについて入力画像から得られたL型接続点との対応を試み、変換パラメータとモデルパラメータの推定を行った結果を示す。図3.9のアルゴリズムを適用し、各パラメータはそれぞれ $\varepsilon^+=0.05$, $\varepsilon^-=0.5$, $\zeta=0.1$, $\text{Max}=100$, $p_i^{(0)}=10$, $\alpha=20$, $\delta=1$ とした。

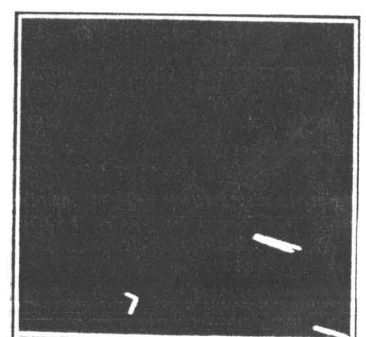
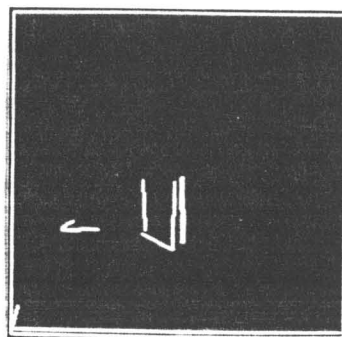
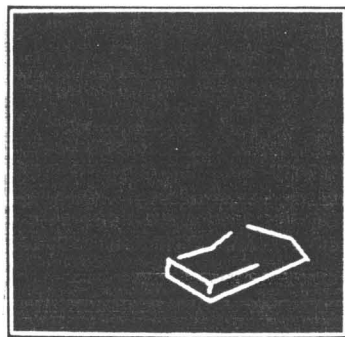
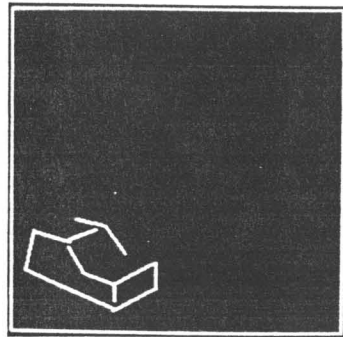
図4.10～図4.12に入力画像、直線近似を行った結果、検証前の仮説クラスタ(対応の付けられたL型接続点を表示している)、検証後の仮説クラスタを示す。図4.10では、パラメータ化された1つの物体モデルにより、形状の異なる2つの物体が認識されている。図4.11では、隠れのある状況において、認識が行われている。さらに、図4.12では、2つの物体モデルが用いられている。また、変換パラメータとモデルパラメータについても良好な推定値が得られている。



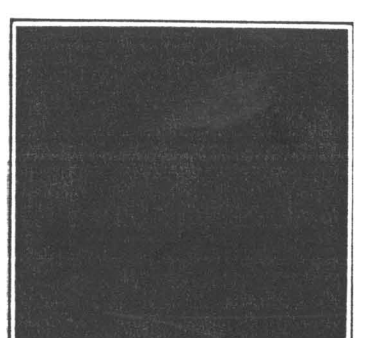
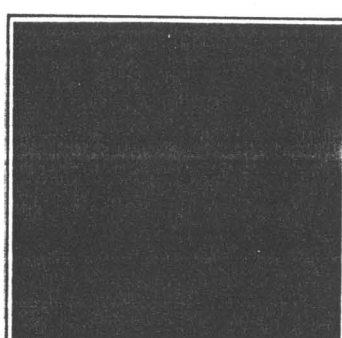
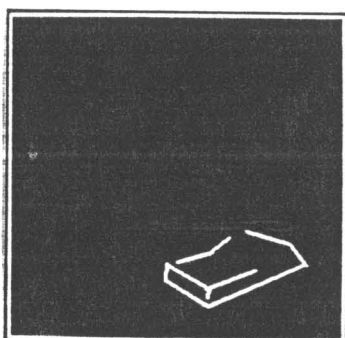
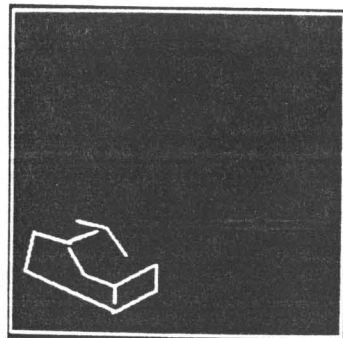
(a) 入力画像



(b) セグメント画像

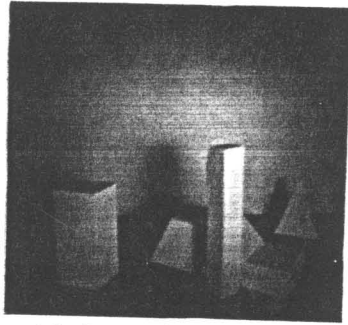


(c) 認識結果 (検証前)

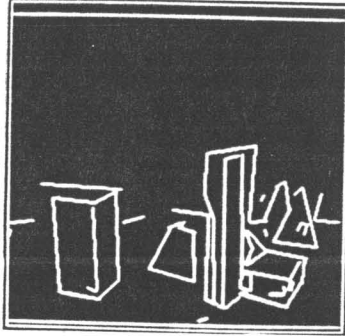


(d) 認識結果 (検証後)

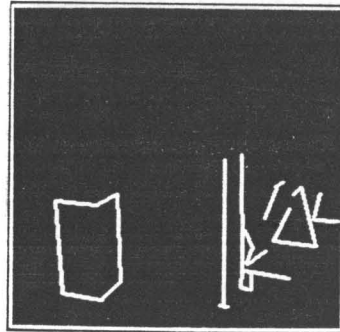
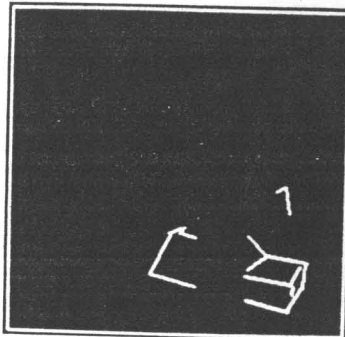
図4.10 1つの物体モデルを用いて2つの形状の異なる物体の認識



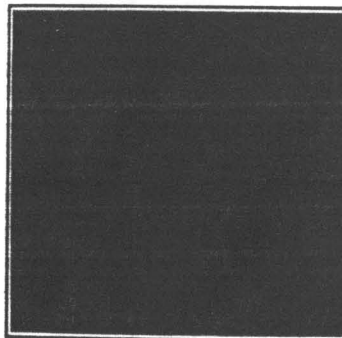
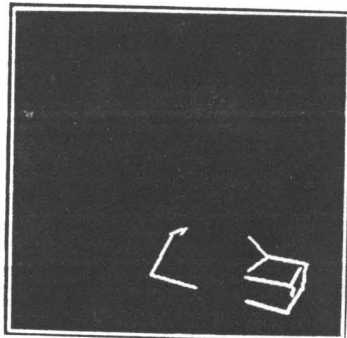
(a) 入力画像



(b) セグメント画像

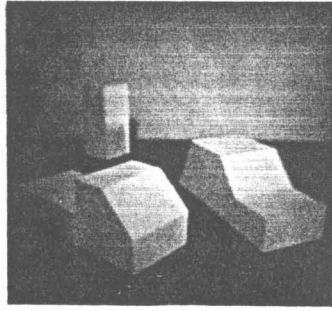


(c) 認識結果 (検証前)

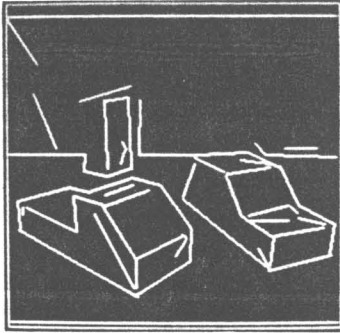


(d) 認識結果 (検証後)

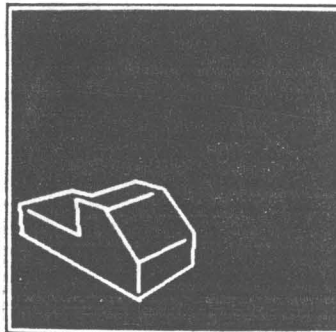
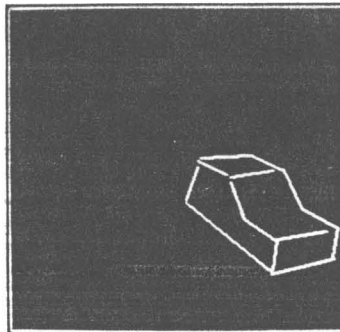
図 4.11 隠れの多い場合



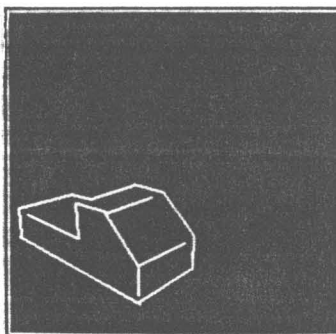
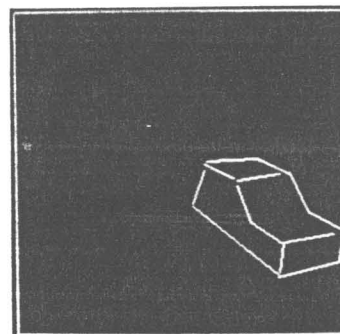
(a) 入力画像



(b) セグメント画像



(c) 認識結果 (検証前)



(d) 認識結果 (検証後)

図4.12 2つの物体モデルを用いた場合

第8節 結言

本章では、仮説間の並列相互作用によるロバストな3次元物体認識手法を示した。物体の認識と分離を2段階の拘束満足ネットワークによって行った。これによって、パラメータ化された物体モデルといった弱い制約しか用いることができない場合でも、証言を並列に集約することによって安定に認識が行えることを示した。

本章で述べたネットワークの大きな特長は対象を分離する能力である。物体認識を含め多くの視覚の問題（例えば、surface reconstruction, structure from motion等）は、個々の処理課題と同時に、画像要素の群化／分離（segregation/aggrigation）—セグメンテーション—の問題を解決しなければならない。視覚の問題は並列ネットワークを用いてうまく解決される場合が多いが、このとき画像要素間の関係だけではなく、その関係の関係をを用いる（即ち、関係を選択する過程が）必要があると考えられる。画像要素間の関係の拘束は、3D from 2Dの問題を解く場合の拘束条件の不足を補う。しかし、関係づけられた画像要素がシーン中で同一の事象を反映していると暗黙に仮定されている。ここで、同一の事象とは、物体認識の場合は“同一の物体に属している”，表面再構成の場合には“同一の表面に属している”ということである。これは、予めセグメンテーションが仮定されていることに他ならない。これをうまく解決しているのは、Kochら[4.9]のライン過程（line process）の考え方である。画像要素を単位としたネットワークと画像要素間の関係を単位としたネットワーク（ライン過程）を相互作用させることによって、なめらかな表面再構成とセグメンテーションを同時に解決している。本章で述べたアプローチもこの方法と類似している。ただし、本研究では2つの過程が逐次的に適用されている。“仮説対”のネットワークは関係の選択によってセグメンテーションを行なうものである。これらのアプローチは、従来の弛緩法など並列ネットワークによる処理の能力を本質的に高めるものと考えられる。

第5章 結論

本研究では、まず、ガウス関数の性質を生かした再帰的構造のフィルタでガウスピラミッドを高速に精度よく計算する手法について述べた。さらに、仮説ネットワークによる並列空間推論による、多重解像度両眼立体視モデルとパラメータ化3次元物体の認識手法について述べ、不連続の多い表面、或は、パラメータ化モデルといった、より一般性のある弱い拘束条件のもとで、仮説ネットワーク上での並列集約的な処理の能力を実験的に実証した。

ここでは、本研究で取り上げた話題に関する今後の問題について触れる。

1. ガウスフィルタの多重解像度モデルとしての妥当性.

ガウス関数の性質のよさからスケールスペースの種々の興味深い性質が発見されてきたが、画像表現としての妥当性の議論よりも性質のよさからくる数学的な解析が先行しているように思われる。視覚データの階層性をガウス関数の標準偏差という1つの軸で表わそうというのは無理があると考えられる。スケールスペースは等方性拡散過程で線型フィルタリング結果のスタックであるが、視覚データの性質を考えると“不連続”という本質的な要素があり、そこで生じる非線型効果を考慮する必要がある。不連続点では、情報の拡散がせき止められる。文献[5.1]では、“不連続”の感度(sensitivity)という別の軸を示唆している。即ち、“細かく見る”、“粗く見る”という概念の定式化として、“いろいろなスケールでものを見る”だけでなく、“いろいろな感度でものを見る”という要素を考慮する必要があると考えられる。

2. ネットワークの制御について.

現在のネットワークの多くは、各ユニット間の相互作用は入力データを基に設定される結合の強さによって定まる。しかし、第4章の結言で述べたようにセグメンテーションが仮定されていない場合、あるユニットが協調(抑制)しあうべきか、そうでないかを決定するのもネットワーク全体の動的な振舞いによって決定されねばならない。画像要素の属性(ラベル、シーンパラメータ)を決定するためには、それらの間の関係の拘束を利用しなければならないが、その相互作用は画像要素間のシーンでの因果関係に依存する。ライン過程[5.2]のような画像要素間の因果関係(種々のシーンパラメータの不連続など)を明確にする過程をいかに組み込みこんでいくかが重要な課題であると考えられる。

謝 辞

本研究を行うに際して、多くの方々から有益な示唆、助言、及び、暖かい励ましを頂いた。これらの方々に対して、心より感謝する。

嵩忠雄教授からは、常に暖かい励ましを賜り、また多くのお世話を頂き、研究生活を支えて頂いた。

故田中幸吉教授からは、学部・大学院を通して御教示頂き、本研究を暖かく見守って頂いた。

産業科学研究所の北橋忠宏教授からは、専門的な事柄から精神的な面まで、洞察深く、かつ、暖かい助言を賜った。

情報工学科の鳥居宏次教授からは、精神的な面、及び、研究者のあり方に関して有益な助言を頂いた。

産業科学研究所の豊田順一教授からは、学部・大学院を通して御教示、御指導、励ましを頂いた。

制御工学科の辻三郎教授、ならびに、井口征士教授からは、学部・大学院を通して御教示頂き、また学会誌等を通して多くの専門的な知識を賜った。

情報工学科の藤沢俊男教授、都倉信樹教授、故高島堅助教授、谷口健一教授、大型計算機センターの宮原秀夫教授、ならびに、産業科学研究所の角所収教授には、学部・大学院を通して御教示頂いた。

田村進一助教授は、直接の指導教官であり、日頃から筆者の様々な我がまを聞いてくださり、筆者の研究を様々な面から支えてくださった。

広島大学の田中稔助教授、ならびに、産業科学研究所の小川均講師からは、技術的、専門的な面、及び、研究生活において様々な助言を頂いた。

京都教育大学の唐沢博講師からは、人工知能研究の魅力について教えて頂き、筆者がコンピュータビジョン研究を行う動機づけを与えて頂いた。

関浩之助手からは、博士課程の後半において研究生活における様々な有益な助言を賜った。

大学院生の玉野隆一氏（現NEC）、ならびに、栗屋英司氏からはインプリメンテーションと実験に際して、多大な貢献を頂いた。また、彼らとの種々の議論は、非常に有益なものであった。

大学院生の岡本恭一氏（現東芝）、ならびに、川崎伸吾氏（現日本楽器）の作成した画像処理ツールの利用は、本研究の推進に非常に効果的であった。

その他、嵩研究室、及び、旧情報基礎論講座の多くの方々にお世話になった。

さらに、博士課程の後半において、(株)ATR視聴覚機構研究所の多くの方々にお世話になった。

参考文献

第1章

- [1.1] Witkin, A. : "Scale space filtering", Proc. of 8th IJCAI, pp. 1019-1022(1983).
- [1.2] Koch, C. Marroquin, J., and Yuille, A. : "Analog "neuronal" networks in early vision", Proc. Natl. Acad. Sci. USA, 83, pp. 4263-4267(1986).
- [1.3] Marr, D. and Hildreth, E. : "Theory of edge detection", Proc. R. Soc. Lond., B 207, pp. 187-217(1980).
- [1.4] 飯島泰蔵 : "パターン認識", 電気電子工学体系, 43, コロナ社(1973).
- [1.5] Burt, P. J. : "Fast filter transforms for image processing", CGIP, 16, pp. 20-51(1981).
- [1.6] Zucker, S. W. : "Relaxation labelling and the reduction of local ambiguities", Proc. of 3rd IJ CPR, pp. 852-861(1976).
- [1.7] Hummel, R. A. and Zucker, S. W. : "On the foundation of relaxation labelling processes", IEEE Trans., PAMI-5, 3, pp. 267-283(1983).
- [1.8] Hopfield, J. J. and Tank, D. W. : "Neural" computation of decisions in optimization problems", Biol. Cybern., 52, pp. 141-152(1985).
- [1.9] Terzopoulos, D. : "Image analysis using multigrid relaxation methods", IEEE Trans., PAMI-8, 2, pp. 129-139(1986).

第2章

- [2.1] 飯島泰蔵 : "パターン認識", 電気電子工学体系, 43, コロナ社(1973).
- [2.2] Witkin, A. : "Scale space filtering", Proc. of 8th IJCAI, pp. 1019-1022(1983).
- [2.3] 趙, 飯島 : "図形観測における視点, 視野の決定法に関する理論", 信学論(D), J68-D, 4, pp. 508-514(昭60-04).
- [2.4] 趙, 飯島 : "安定視点木法による画像の特徴抽出理論", 信学論(D), J68-D, 5, pp. 1125-1132(昭60-05).
- [2.5] Yuille, A. L. and Poggio, T. A. : "Scaling theorems for zero crossings", IEEE Trans., PAMI-8, 1, pp. 15-25(1986).

- [2. 6] Koenderink, J. J.,"The structure of image". Biol. Cybern. , 50, pp. 363-370(1984).
- [2. 7] Crowley, J. L. and Parker, A. C. : "A representation for shape based on peaks and ridges in the difference of low-pass transform", IEEE Trans. , PAMI-6, 2, pp. 156-169(1984).
- [2. 8] Asada, H. and Brady, M. : "The Curvature Primal Sketch", IEEE Trans. , PAMI-8, 1, pp. 2-14(1986).
- [2. 9] Mokhtarian, F. and Mackworth A. : "Scale-based description and recognition of planar curves and two-dimensional shapes", IEEE Trans. , PAMI-8, 1, pp. 34-43(1986).
- [2. 10] Burt, P. J. : "Fast filter transforms for image processing", CGIP, 16, pp. 20-51(1981).
- [2. 11] Crowley, J. L. and Stern, R. M. : "Fast computation of the difference of low-pass transform", IEEE Trans. , PAMI-6, 2, pp. 212-222(1984).
- [2. 12] Wells, III, W. M. : "Efficient synthesis of Gaussian filters by cascaded uniform filters", IEEE Trans. , PAMI-8, 2, pp. 234-239 (1986).
- [2. 13] Burr, D. J. : "A fast filtering operation for robot stereo vision", Proc. of 7th ICPR, pp. 669-672(1984).
- [2. 14] Castan, S. and Shen, J. : "Fast filter transform theory and design for image processing", Proc. of CVPR' 85, pp. 420-422(1985).
- [2. 15] Marr, D. and Hildreth, E. : "Theory of edge detection", Proc. R. Soc. Lond. , B 207, pp. 187-217(1980).
- [2. 16] Boll, M. (弥永, 矢野訳) : "万能数值表", pp. 578-581, 白水社(1952).
- [2. 17] Hartmann, G. : "Recursive features of circular receptive fields", Biol. Cybern. , 43, pp. 199-208(1982).

第 3 章

- [3. 1] Mayhew, J. : "Stereopsis", in Physical and biological processing of images, Braddick, O. J. and Sleigh, A. C. ed., Springer, Berlin, pp. 204-216(1982).
- [3. 2] Marr, D. and Poggio, T. : "A cooperative stereo algorithm", Science, 194, pp. 283-287(1976).

- [3. 3] Medioni, G. and Nevatia, R. : "Segment-based stereo matching", CVGIP, 31, pp. 2-18(1985).
- [3. 4] Barnard, S. T. : "A stochastic approach to stereo vision", Proc. of AAAI' 86, pp. 676-680(1986).
- [3. 5] Ohta, Y and Kanade, T. : "Stereo by intra- and inter-scanline search using dynamic programming", Technical Report, CMU-CS-83-162(1983).
- [3. 6] Prazdny, K. : "Detection of binocular disparity", Biol. Cybern., 52, pp. 93-99(1985).
- [3. 7] Hopfield, J. J. and Tank, D. W. : "Neural" computation of decisions in optimization problems", Biol. Cybern., 52, pp. 141-152(1985).
- [3. 8] Mayhew, J. E. W. and Frisby, J. P. : "Psychophysical and computational studies towards a theory of human stereopsis", Artif. Intell., 17, pp. 349-387 (1981).
- [3. 9] Burt, P. J. and Julesz, B. : "A disparity gradient limit for binocular fusion", Science, 208, pp. 615-617 (1980).
- [3. 10] Feldman, J. A. and Ballard, D. H. : "Connectionist models and their properties", Cognitive Science, 6, pp. 205-254 (1982).
- [3. 11] 甘利俊一 : " 神経回路網の数理" , 産業図書, pp. 195-229 (1978).
- [3. 12] Szeriski, R. and Hinton, G. : "Solving random-dot stereograms using the heat equation", Proc. of CVPR '85, pp. 284-288 (1985).
- [3. 13] Prazdny, K. : "On the coarse-to-fine strategy in stereomatching", Bulletin of the Psychonomic Society, 25, 2, pp. 92-94(1987).
- [3. 14] Marr, D. and Poggio, T. : "A computational of human stereo vision", Proc. R. Soc. Lond. B. 204, pp. 301-328 (1979).
- [3. 15] Nishimoto, Y. and Shirai, Y. : "A parallel matching algorithm for stereo vision", Proc of 9th IJCAI, pp. 977-980 (1985).
- [3. 16] Marr, D. and Hildreth, E. : "Theory of edge detection", Proc. R. Soc. Lond. B. 207, pp. 187-217 (1980).
- [3. 17] 佐藤, 田村 : " カスケードガウスフィルタの設計法" , 信学論(D), J70-D, 10, pp. 1911-1920(1987).
- [3. 18] Hummel, R. A. and Zucker, S. W. : "On the foundation of relaxation labeling processes", IEEE Trans., PAMI-5, pp. 267-286 (1983).

- [3. 19] Ibison, M. C. and Zapalowski, L. : "On the use of relaxation labelling in the correspondence problem", Pattern Recognition Letter, 4, pp. 103-109 (1986).
- [3. 20] Hinton, G. E. and Sejnowski, T. J. : "Optimal perceptual inference", Proc. of CVPR '83, pp. 448-453 (1983).

第 4 章

- [4. 1] Brooks, R. : "Symbolic reasoning among 3-D models and 2-D images", Artificial Intelligence, 17, 1-3, pp. 285-348(1981)
- [4. 2] Grimson, W. E. L. : "Recognition of object families using parameterized models", Proc. of 1st ICCV, pp. 93-101(1987).
- [4. 3] Lowe, D. G. : "Visual recognition from spatio correspondence and perceptual organization", Proc. of 9th IJCAI, pp. 953-958(1985).
- [4. 4] Ballard D. H. and Sabbah D. : "On shapes", Proc. of 7th IJCAI, pp. 607-612(1981).
- [4. 5] 田中, 辻, Ballard, D. H., Curtiss, M. : "多面体認識の並列処理", 情処CV研資, CV34-3(1985).
- [4. 6] Hummel, R. A. and Zucker, S. W. : "On the foundations of relaxation labelling processes", IEEE Trans., PAMI-5, 3, pp. 267-283(1983).
- [4. 7] Davis, L. : "Shape matching using relaxation techniques", IEEE Trans., PAMI-1, 1, pp. 60-72(1979)
- [4. 8] Bhanu, B. : "Representation and shape matching of 3-D objects", IEEE Trans., PAMI-6, 3, pp. 340-351(1984).
- [4. 9] Koch, C., Marroquin, J., and Yuille, A. : "Analog "neuronal" networks in early vision", Proc. Natl. Acad. Sci. USA, 83, pp. 4263-4267(1986).
- [4. 10] Hopfield, J. J. and Tank, D. W. : "Neural" computation of decisions in optimization problems", Biol. Cybern., 52, pp. 141-152(1985).

第5章

- [5.1] Blake, A. and Zisserman, A. : "Visual Reconstruction", The MIT Press, Cambridge Massachusetts(1987).
- [5.2] Koch, C, Marroquin, J., and Yuille, A. : "Analog "neuronal" networks in early vision", Proc. Natl. Acad. Sci. USA, 83, pp. 4263-4267(1986).

付 録

1. エッジセグメントの傾きから視差勾配の算出

図3.7に於て、SegLとSegRの対応を仮定すると、融合座標系に於て3次元セグメントSeg3Dが定まる。Seg3Dの端点の融合座標をそれぞれ (U_1, V_1, D_1) , (U_2, V_2, D_2) とする(但し一般性を失うことなく $V_1 > V_2$ と仮定)と端点間の視差の差 $\Delta \text{disp} = D_1 - D_2$ は、次式で与えられる。

$$\Delta \text{disp} = \frac{V_1 - V_2}{\tan \psi_R} - \frac{V_1 - V_2}{\tan \psi_L} \quad (\text{A.1})$$

また、Seg3DのU-V平面への正射影の長さXは、次式で与えられる。

$$X^2 = (V_1 - V_2)^2 + \left(\left(\frac{V_1 - V_2}{\tan \psi_R} + \frac{V_1 - V_2}{\tan \psi_L} \right) / 2 \right)^2 \quad (\text{A.2})$$

Seg3Dの視差勾配は、次式で与えられるので、

$$\Delta \text{disp} / X \quad (\text{A.3})$$

式(A.3)に、式(A.1)、及び、式(A.2)を代入することにより、式(3.4)が得られる。

2. 鉛直軸回りの回転角 θ_{PW} の推定

図 A. 1 のように 1 点と 1 方向の対応を基に θ_{PW} の推定式を導く。画像上の点の同次座標を $i = (u, v, t)$ ，モデルプリミティブの点の同次座標を $p = (x_p, y_p, z_p, 1)$ とすると，次式が成立する。

$$i = p T_{PW} T_{WI}, \quad (U, V) = (u/t, v/t) \quad (A. 4)$$

ただし，

$$T_{PW} = \begin{bmatrix} \cos \theta_{PW} & 0 & \sin \theta_{PW} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_{PW} & 0 & \cos \theta_{PW} & 0 \\ t_{PWx} & t_{PWy} & t_{PWz} & 1 \end{bmatrix},$$

$$T_{WI} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (s/f) \\ 0 & 0 & 0 \end{bmatrix}.$$

(U, V) は，画像平面座標系での画像点。

図 A. 1 におけるプリミティブ座標系の $(0, 0, 0) - (px_1, py_1, pz_1)$ と画像平面上の $(U_0, V_0) - (U_0 + u, V_0 + v)$ の対応を仮定すると，次の 4 つの方程式が得られる（下線部は未知数）。ただし， $p > 0$ とする。

$$\underline{t_{PWx}} - a \underline{t_{PWy}} + b \underline{t_{PWz}} = 0 \quad (A. 5)$$

$$-c \underline{t_{PWy}} + d \underline{t_{PWz}} = 0 \quad (A. 6)$$

$$\underline{t_{PWx}} - a' \underline{t_{PWy}} + b' \underline{t_{PWz}} + p(b' x_1 - z_1) \underline{\sin \theta_{PW}} + p(b' z_1 + x_1) \underline{\cos \theta_{PW}} = pa' y_1 \quad (A. 7)$$

$$-c' \underline{t_{PWy}} + d' \underline{t_{PWz}} + pd' x_1 \underline{\sin \theta_{PW}} + pd' z_1 \underline{\cos \theta_{PW}} = pc' y_1 \quad (A. 8)$$

ただし， $a = sU_0 \sin \alpha / f$ ， $b = sU_0 \cos \alpha / f$ ， $c = sV_0 \sin \alpha / f + \cos \alpha$ ，

$d = sV_0 \cos \alpha / f - \sin \alpha$ ， $a' = s(U_0 + u) \sin \alpha / f$ ， $b' = s(U_0 + u) \cos \alpha / f$ ， $c' = s(V_0 + v) \sin \alpha / f + \cos \alpha$ ， $d' = s(V_0 + v) \cos \alpha / f - \sin \alpha$ 。

$\{((A. 7) - (A. 5)) * v - ((A. 8) - (A. 6)) * u\} / p$ より次式が得られ，

$$\{(bx_1 - z_1)v - dx_1 u\} \underline{\sin \theta_{PW}} + \{(bz_1 + x_1)v - dz_1 u\} \underline{\cos \theta_{PW}} = (av - cu)y_1 \quad (A. 9)$$

さらに，(A. 9) 式と $\sin^2 \theta_{PW} + \cos^2 \theta_{PW} = 1$ より， $\sin \theta_{PW}$ に関する 2 次方程式が得ら

れる。よって、図 A. 1 の対応から θ_{PW} の推定は次式を解くことに帰着できる。

$$\begin{aligned} & ((bv-du)^2+v^2)(x_1^2+z_1^2)\sin^2\theta_{PW} \\ & -2((bv-du)x_1-vz_1)(av-cu)y_1\sin\theta_{PW} \\ & +((av-cu)^2y_1^2-((bv-du)z_1+vx_1)^2) = 0 \end{aligned} \quad (A.10)$$

ただし、 $x_1=0$ かつ $z_1=0$ 、あるいは、 $d=0$ かつ $v=0$ のとき解は不定または不能となり、(A.10)式の判別式 $D < 0$ のとき、解は得られない。

1) $x_1=0$ かつ $z_1=0$ のとき、物体の辺が x - z 平面に垂直であることを表し、 $av-cu=0$ のとき対応は可能で解は不定になり、それ以外のとき解は不能でどのように θ_{PW} を選んでも対応は不可能なことを表す。

2) $d=0$ かつ $v=0$ のとき、 (U_0, V_0) に対応する視線が x - z 平面に平行であることを表し、 $y_1=0$ のとき対応は可能で解は不定になり、それ以外のとき不能となる。
 $y_1=0$ のときは、 x - z 平面に平行な辺が偶然、高さ h のところにある場合に相当する。

3) 1), 2) 以外のときで、 $D < 0$ のとき、どのように θ_{PW} を選んでも対応は不可能で、 $D = 0$ のときは 1 個、 $D > 0$ のときは 2 個それぞれ解が定まる。ただし、 $t_{PWz} > 0$ となる場合のみを真の解とする。

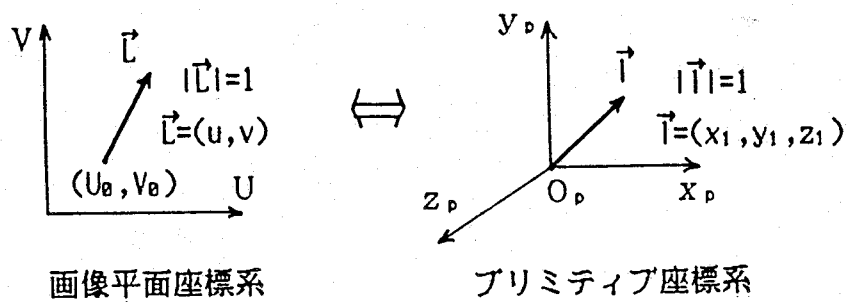


図 A. 1 θ_{pw} を推定するための仮対応

3. 可変辺の方向の推定

方向が可変の辺 $l_v = (x_1, y_1, z_1)$ を 2 つの既知ベクトル $s_1 = (x_{s1}, y_{s1}, z_{s1})$, $s_2 = (x_{s2}, y_{s2}, z_{s2})$ の線型和 $q s_1 + r s_2$ で表し, θ PW が得られている場合について, q および r を推定する.

式(A. 9)を変形すると次式が得られる.

$$K_1 x_1 + K_2 z_1 = K_3 y_1 \quad (\text{A. 11})$$

ここで, $K_1 = (bv - du) \sin \theta + v \cos \theta$, $K_2 = -v \sin \theta + (bv - du) \cos \theta$, $K_3 = av - cu$.

式(A. 11)と以下の条件と合わせると,

$$l_v = q s_1 + r s_2 = (q x_{s1} + r x_{s2}, q y_{s1} + r y_{s2}, q z_{s1} + r z_{s2}) \quad (\text{A. 12})$$

$$|l_v| = |s_1| = |s_2| = 1 \quad (\text{A. 13})$$

q, r はそれぞれ次式で推定できる.

$$q = \frac{B}{\sqrt{(B x_{s1} - A x_{s2})^2 + (B y_{s1} - A y_{s2})^2 + (B z_{s1} - A z_{s2})^2}} \quad (\text{A. 14})$$

$$r = \frac{A}{\sqrt{(B x_{s1} - A x_{s2})^2 + (B y_{s1} - A y_{s2})^2 + (B z_{s1} - A z_{s2})^2}} \quad (\text{A. 15})$$

ここで, $A = K_1 x_{s1} + K_2 z_{s1} - K_3 y_{s1}$, $B = K_1 x_{s2} + K_2 z_{s2} - K_3 y_{s2}$.

ただし, $t_{PWz} > 0$ となる場合のみを真の解とする.

4. 平行移動パラメータとモデルパラメータの推定式の導出

T_{OW} , T_{OP} , T_{PW} をそれぞれ同次変換行列として表すと次式が成立する.

$$T_{OW} = T_{OP} T_{PW} \quad (A. 16)$$

ただし,

$$T_{OW} = \begin{bmatrix} \cos \theta_{OW} & 0 & \sin \theta_{OW} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_{OW} & 0 & \cos \theta_{OW} & 0 \\ t_{OWx} & t_{OWy} & t_{OWz} & 1 \end{bmatrix},$$

$$T_{OP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -P_x & -P_y & -P_z & 1 \end{bmatrix},$$

$$T_{PW} = \begin{bmatrix} \cos \theta_{PW} & 0 & \sin \theta_{PW} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_{PW} & 0 & \cos \theta_{PW} & 0 \\ t_{PWx} & t_{PWy} & t_{PWz} & 1 \end{bmatrix}.$$

ここで, (P_x, P_y, P_z) は, モデルプリミティブの頂点の物体座標系における座標で, 次式のように n 個のモデルパラメータの線型結合で表される.

$$(P_x, P_y, P_z) = \left(\sum_i a_i P_i, \sum_i b_i P_i, \sum_i c_i P_i \right) \quad (A. 17)$$

各プリミティブには, $a = (a_1, a_2, \dots, a_n)$, $b = (b_1, b_2, \dots, b_n)$, $c = (c_1, c_2, \dots, c_n)$ が記述されている.

式(A. 16)より次式が得られ,

$$\left. \begin{aligned} t_{OWx} &= -P_x \cos \theta_{PW} + P_z \sin \theta_{PW} + t_{PWx} \\ t_{OWy} &= -P_y + t_{PWy} \\ t_{OWz} &= -P_x \sin \theta_{PW} - P_z \cos \theta_{PW} + t_{PWz} \end{aligned} \right\} \quad (A. 18)$$

これらをプリミティブ座標系の原点と画像上の (U_0, V_0) の対応により得られる式(A. 5), 式(A. 6)に代入することにより次に示す平行移動パラメータ $t_{OW} = (t_{OWx}, t_{OWy}, t_{OWz})$ とモデルパラメータ P_1, P_2, \dots, P_n の推定式が得られる.

$$\left. \begin{aligned} t_{OWx} - a t_{OWy} + b t_{OWz} + (b \sin \theta_{PW} + \cos \theta_{PW}) P_x \\ - a P_y + (b \cos \theta_{PW} - \sin \theta_{PW}) P_z = 0 \\ -c t_{OWy} + d t_{OWz} + d \sin \theta_{PW} P_x - c P_y + d \cos \theta_{PW} P_z = 0 \end{aligned} \right\} \quad (A. 19)$$

また、式(A.19)とは別に可変辺を含むモデルプリミティブとの対応から定まった可変辺の方向について、モデルパラメータの推定式が得られる。付録2より可変辺の方向は $l_v = q s_1 + r s_2$ (ただし、 $|l_v| = 1$) と表される。 s_1, s_2 それぞれの方向の長さを表すモデルパラメータを P_i, P_j とすると、次に示すモデルパラメータの推定式が得られる。

$$q/r = P_i/P_j \quad (A.20)$$

発表論文

1. 学会誌論文

・本論文に関して

- 1)佐藤, 田村: “カスケードガウスフィルタの設計法”, 電子情報通信学会論文誌 (D), Vol. J70-D, No. 10, pp. 1911-1920(1987).
- 2)佐藤, 栗屋, 田村: “両眼立体視のエネルギ最小化アプローチ: 複数の交錯する表面の知覚”, 電子情報通信学会論文誌 (D), Vol. J71-D, No. 2 (1988).
- 3)佐藤, 玉野, 田村: “並列分散処理による三次元物体認識: 単眼画像とパラメータ化モデルの照合” (投稿準備中).

・その他

- 4)佐藤, 河合, 田村: “CCD撮像素子を用いたマルチスペクトル画像入力装置”, 画像電子学会誌, Vol. 14, No. 2, pp. 112-120(1985).

2. 国際会議発表

- 1)Y. Sato, R. Tamano and S. Tamura: “Connectionist Approach to 3-D Object Recognition: Matching of Parameterized Models and Monocular Image”, IEEE Computer Society Workshop on Computer Vision, Miami Beach, pp. 252-254(1987).

3. 研究会発表

・本論文に関して

- 1)玉野, 佐藤, 田村: “エッジ対応に基づく3D物体認識”, 情報処理学会研究会報告, 86-CV-44-2(1986).
- 2)佐藤, 玉野, 田村: “局所仮説の統合による3D物体認識”, 電子通信学会技術研究報告, PRU86-105(1987).
- 3)佐藤, 栗屋, 田村: “エッジセグメントに基づく協調性ステレオアルゴリズム”, 電子情報通信学会技術研究報告, PRU87-14(1987).

・その他

- 4)佐藤, 河合, 田村: “CCD撮像素子を用いたマルチスペクトル画像入力装置” 電子通信学会技術研究報告, IE83-78(1983).

4. 全国大会発表

・本論文に関して

- 1) 佐藤, 田村: “カスケードガウシアンフィルタの設計法”, 第33回情報処理学会全国大会, pp. 1597-1598(1986).
- 2) 佐藤, 玉野, 田村: “一般化Hough変換による内部パラメータを持つ3D物体認識”, 第33回情報処理学会全国大会, pp. 1613-1614(1986).

・その他

- 3) 佐藤, 河合, 田村, 田中: “マルチスペクトル画像入力装置の試作”, 昭和57年電気関係学会関西連合大会, G251(1982).
- 4) ルー, 佐藤, 河合, 田村: “眼底画像による両眼眼球運動の測定”, 昭和60年電気関係学会関西連合大会, (1985).