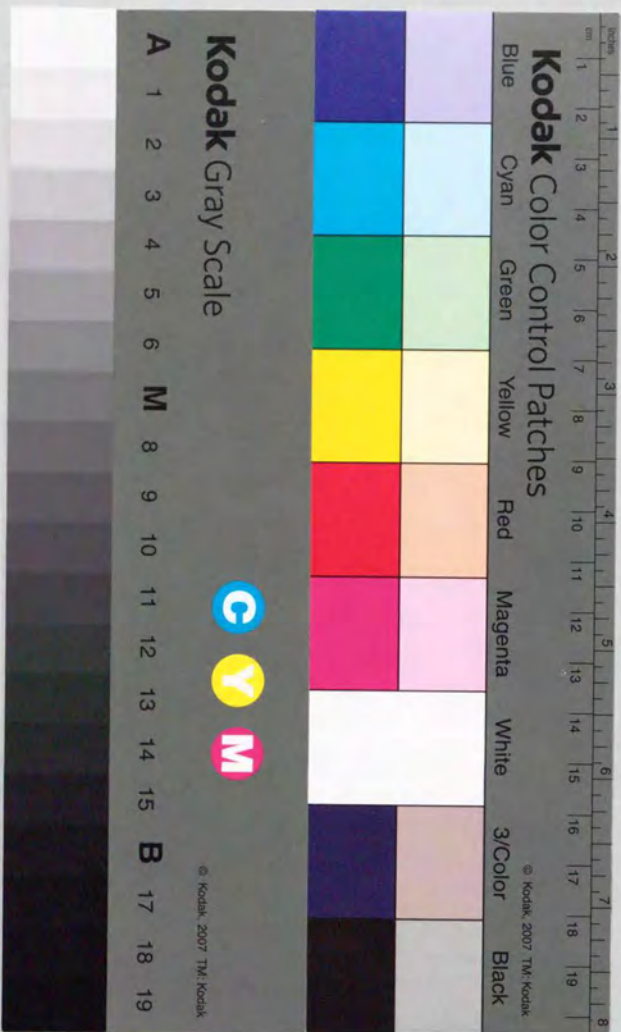


多面体全体形状の視覚認識および  
ロボット制御への応用に関する研究

金子 俊一





①

多面体全体形状の視覚認識および  
ロボット制御への応用に関する研究

金子 俊 一

## 目 次

第1章 序論 .....	1
1.1 本研究の目的 .....	1
1.2 関連研究のまとめ .....	5
1.3 本論文の構成 .....	9
第2章 基本的アプローチ .....	11
2.1 前提およびアプローチの概要 .....	11
2.2 成果の概要 .....	15
第3章 直接像および鏡像を利用したステレオ撮像による位置計算法 ....	17
3.1 概要 .....	17
3.2 特徴点の位置計算の定式化 .....	18
3.3 対応点探索問題に関する考察 .....	23
3.4 隠れ面上にある特徴点の位置計算 .....	25
3.5 座標変換行列の変換 .....	31
3.6 実験 .....	33
3.7 まとめ .....	40
第4章 多視野のステレオ画像対を利用した多面体物体の認識 .....	41
4.1 概要 .....	41
4.2 特徴点の位置計算の定式化 .....	43
4.3 回転角度の設定 .....	47
4.4 画像処理およびグラフの生成 .....	49
4.5 対応点探索のための処理 .....	59
4.6 全体形状の再構成 .....	63
4.7 実験 .....	70
4.8 まとめ .....	78
第5章 視覚認識を利用したロボットアームの制御 .....	79
5.1 概要 .....	79
5.2 視覚的捕捉 .....	84
5.3 対象物の認識 .....	90
5.4 対象物の把握 .....	106
5.5 実験 .....	112
5.6 まとめ .....	124



第6章 提案した視覚認識法の評価 .....	125
6.1 問題点と解決の見通し .....	125
6.2 ステレオ撮像による位置計算法の比較・評価の方法 .....	128
6.3 計算精度の検討 .....	133
6.4 まとめ .....	143
第7章 結論 .....	144
7.1 本研究の成果 .....	144
7.2 今後の課題 .....	147
謝辞 .....	149
参考文献 .....	150

## 第1章 序論

### 1.1 本研究の目的

視覚認識，すなわち画像情報を利用して対象を認識するための技術を確立することは，工学的および産業の見地から大きい意義をもつ．中でも3次元物体の認識は，人工知能<sup>33)</sup>，ロボット工学<sup>46)</sup>，心理学，生理学および生体工学<sup>73)</sup>などの分野において基礎的な研究が行われている．一方では産業的要請をうけて実際の応用を指向した手法の開発も強く望まれている．特に，ロボットシステムなどを多目的に利用するという目的のために，いわゆるロボットビジョンを中心とするセンサシステムに対する機能の高度化および柔軟化の要求は強い．たとえば，現在，一般のロボットシステムはオフラインのプログラミングあるいは教示によって設定された作業を行い，それ以外の柔軟性は期待できない．その原因の1つは，現在オペレータの行っている対象物の識別，位置合わせ，作業の監視などの仕事をロボットシステム自体が代行できないことである．本研究で提案する手法はそのような機能の自動化の基礎技術をめざしている．すなわち，多面体の全体形状（構造）の認識や，おおまかな姿勢認識処理などの認識機能を実現するためのものである．しかしながら，そこでの位置計算の精度は後述するようにたかだか数ミリメートルであるために，これらの手法のみを利用してたとえば精密組立などの作業を行うことはできない．また，ロボットシステム全体としての目的に応じて，他の高精度センサなどを併用しなくてはならない場合もある．そのような場合における視覚認識手法の果たすべき役割は，視野全体にわたる環境情報や対象物の全体的形状情報の獲得など，リアルタイムでかつ大域的な処理を行うことであると考えている．そして，その大域的情報を利用して，専用高精度のセンサ群を駆動しそれらの出力を統合することにより，ロボットシステム全体の柔軟性が確保できると考えている．

本研究の目的は，ロボット制御を指向しつつ，視覚情報を用いた3次元物体認識の手法を検討することである．後の展望において触れるが，ロボット工学における画像情報の利用，あるいは視覚認識技術の利用は近年検討され始めてきており，2次元の対象物の輪郭形状に基づく位置・姿勢の認識技術などは実用化へ向かってい



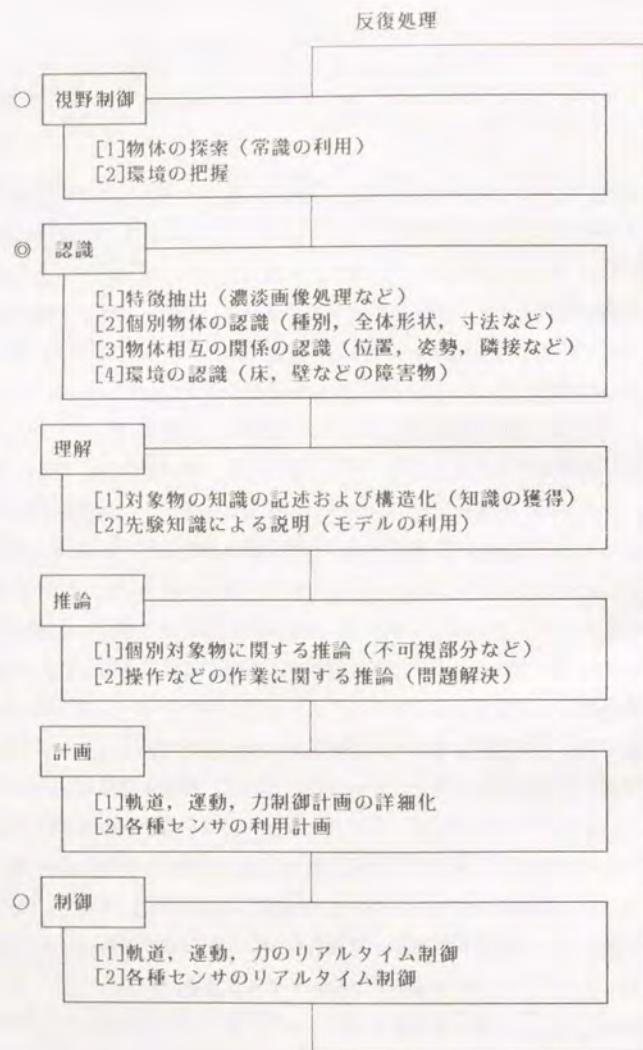


図 1.1 視覚認識機能を中心としたロボットシステムの構成

る。しかしながら，本研究の目的とする3次元対象物の認識技術については，未だ解決すべき問題が多く，基礎的検討を必要としている段階である。その中で本研究では特に，通常の照明下で撮像された濃淡画像を用いて，3次元対象物の全体形状を知るための新しい手法を試み，実験を通してその評価を行う。また，視覚情報を利用したロボットアームの制御方法の試みを行う。

本研究の意義を述べるために，図1.1に本研究で指向するロボットシステムの構成を示す。この構成は視覚認識に関連する機能を中心として，実際の稼働時に必要な機能をまとめたものであり，事前に行われるべき全体的な作業計画，軌道計画，校正などのオフライン機能は含まない。この図において本研究が取り扱うのは，視野制御，認識および制御の部分であるが，中でも認識に関して中心的に考察する。まず，対象物や環境を探索するための視野制御に続いて，対象物を視覚的に捉えて認識処理を行う。本研究は通常的环境下における方法を指向するので，濃淡画像を処理して対象物の特徴を抽出することを考える。理解においては，対象物に関して得られた知識を記述し，構造化する。また，先験的知識による説明機能なども必要となる。推論を通して，軌道，運動計画を実際の環境に適合するように詳細化あるいは修正する。動作時には各種センサの情報を統合し制御する機能が必要である。図中の各項目に関する解説は次節の展望において行うが，このようなシステムを実現するという視点から，現在の3次元視覚認識のための様々な手法を調べてみると，次のような具体的な問題点がある。

- ①対象物の全体的形状をその画像情報に基づき，再構成し認識するためのアプローチが少ない。
- ②濃淡画像の処理を含めた特徴抽出のための具体的方法の提案が少ない。
- ③認識された視覚情報の利用法，とくに認識系と操作系（ロボットアームなど）との結合様式およびその具体的可能性が明確でない。

まず，対象物をロボットハンドなどによって操作するためには，その全体形状を認識することが必要である。これまで提案されている方法においては，対象物の3次元モデルを用意しておき，可視面のみからの何らかの特徴を用いて対象物をモデル中から識別するやり方が多い。そのような照合モデルを用いない場合には，まず



対象とする3次元物体の全体的な形状情報を再構成することが必要である。本研究では、なるべく少ない手間に必要な画像情報を得るための撮像方式（第3章）、および複数の視野からの情報を統合して全体形状を再構成するための手法（第4章）について検討する。これらの手法を用いることにより、図1.1の全体構成における視野制御および個別物体の全体形状の認識が可能となると考える。

次に、特別な照明装置や環境を利用せずに、人間の生活空間のような通常環境で稼働することを指向するときには、その濃淡画像を扱うことが必要である。これまでも、画像処理、情景解析などの分野ではさまざまな濃淡画像処理の手法が提案されてきた。しかしながら、ここで主に必要とするような頂点や輪郭上の折れ曲がり点などの幾何学的な特徴をもつ部分の抽出技法や、濃淡エッジへの直線あてはめ技法などは少なく、さらにステレオ撮像による位置計算の精度などをも考慮した画像処理手法となると、ほとんどないと思われる。特に、特徴点などを安定に抽出するための画像処理手法はシステム実現の鍵となる技術であり、図1.1の全体構成での認識における特徴抽出の問題を解決するための試みである。

また、獲得した視覚情報をロボットアームなどの操作系の制御に利用することを考える。本研究で扱う視覚認識の手法においては、位置精度はたかだか撮像距離の1%程度のオーダーであり、測定精度よりもむしろ処理の柔軟性に期待するところが大きい。したがって、全体システムの中でのその機能は、情景における特定の対象物の位置決めや姿勢判定、面数や頂点の配置などに依存するようなおおまかな形状認識、複数対象物の相互の位置関係の把握などであると捉えている。そのためには、視覚認識独自の適用方法を検討する必要がある。さらに、視覚認識機能とハンドリング機能の結合様式や、視覚センサ、近接センサ、接触センサを含めた機能分散及び統合の方法など解決されるべき問題は多い。本研究では、学習された物体モデルなどの予め知られている情報に基づいてセンサをどのように利用したらよいかという問題について考察する。

なお、ここに挙げた3つの問題点は次節で展望する多くの方法あるいはシステムにおいて共通に存在すると考えられ、これらの解決を計ることは3次元物体認識の多くの手法の実用化へ向けてとくに必要であると考えられる。

## 1.2 関連研究のまとめ

この節では本研究に特に関連するような他の研究について、ステレオ撮像法を基礎とする手法を中心に展望する。

本論文では、「ステレオ撮像」ということばを使用するが、これは複数（狭い意味では2枚）の可視側面を撮像すること、およびそのための撮像方式を表すものとする。一方、いわゆる「ステレオ視」はステレオ撮像をも含めた認識あるいは認知に至る過程と考えることにする。

まず、ステレオ撮像を次のように分けて展望する。

- (1) 受動的ステレオ法： 普通照明下で撮像した2枚の画像を用いて、3次元対象物の奥行きを算出する方法である。主に2台のカメラによる2枚の画像を用いるので、画像間での対応点問題の解法が課題となる。人間の両眼視の直接的なモデルを想定し、設計される場合が多い。
- (2) 能動的ステレオ法： 被測定物体に既知の構造的照明パターンを投影し、その歪みから奥行きを得る方法である。主に1台のカメラと投光器の構成であり、対応点探索が比較的容易である。

そして最後に、

- (3) 3次元物体の認識手法

について対象物の記述方式なども含めながら展望する。

### (1) 受動的ステレオ法

通常2台のカメラによって構成される撮像系<sup>36)</sup>が使われるが、1台のカメラと平面鏡で構成する場合もある<sup>4)</sup>。空中ステレオ写真による実体測定<sup>19)</sup>やSEMステレオ写真からの対象物表面の凹凸の再構成<sup>31)</sup>などは代表的なものである。後に説明する能動的ステレオ法と異なり、対象とする情景に対して構造的照明パターンなどを投影しないので、対応点探索処理の効率的実現が最も大きな課題の1つである。一般に、位置測定の精度を向上させるためには、2台のカメラの幅角あるいは平行移動量を大きくとる必要があるが、大きな視差は必然的に対応点付近の見え方を変化させたり、ときには点を遮蔽してしまうので対応点問題をより困難にする。このトレードオフをできる限り改善することが対応点探索処理の主な目的である。



カメラの位置・姿勢を含めた撮像パラメータが既知であれば、epipolar line (2台のカメラのレンズ中心と画像上の一点によって決まる平面と、他の画像面との交線。以下では対応点の位置を拘束するという意味で拘束直線と呼ぶ) 上を探索すればよい<sup>36)</sup>。しかしながら、拘束直線上に対応点があるという条件は必要条件であるので、この条件だけから対応点が一意に判定できるとは限らない。実際には、画像中に含まれるノイズ成分や校正誤差にも影響されてしまう。

そこで、これまでに多くの対応点探索法が提案されてきている。特徴点を中心とした微小領域内の微分値の相関関数を用いる方法<sup>4)</sup>、原画像の照合を行う方法<sup>31)</sup>、画素の灰色度の分散値およびその方向を合わせもつ演算子を用いる方法<sup>21)</sup>などは計算は単純であるが、その局所性から特徴点の見え方の変化やノイズなどによる誤対応の可能性を多く残している。Marr および Poggio<sup>9)</sup>の神経回路網的な相互作用に基づく協調的アルゴリズム、Nishihara の零交差の符号の相関値に基づく粗密照合法<sup>44)</sup>においては比較的良好な結果が得られるが、計算時間が莫大であり、また通常の情景のように対象物が遮蔽されている場合などは誤対応が生じる。領域抽出処理による signature 関数を探索する方法<sup>20)</sup>は対象物が遮蔽されていても一部の形状(輪郭など)の一致を手がかりにして対応付けを進める。しかしながら、このようないわゆる領域法の欠点は領域分割処理の結果に多く依存することである。すなわち、対応付けすべき特徴の規模が大き過ぎることに問題がある。動的計画法を用いた同一走査線上の特徴点の組合せによる区間対応法<sup>60)</sup>は走査線全域(1次元)にわたる最適対応の組合せを探索するので、誤対応は減少するが非常に多くの計算時間を使用する。他には、対象物を線分などにより単純化し、プリミティブ表現した全体的配置を構造的に照合する方法<sup>81)</sup>などがあるが、プリミティブ表現の巧拙にかなり依存すると思われる。これらの探索方法はすべて2枚の画像間の特徴点の対応を求めるのであるが、3枚めの画像を加え誤対応を削減する方法<sup>39, 59)</sup>も提案されている。また、対象物の濃淡分布が局所的に連続であるという制限を加え、左右画像の視差と物体面の凹凸との局所的な線形関係を仮定して奥行きを再現する手法<sup>74)</sup>も報告されており、これは対応点探索の必要はないが、処理が局所的であり濃淡分布に関する仮定が成立するか否かが問題となる。

受動的ステレオ法の特徴は第1にその柔軟性および情報量の多さにある。したがって、将来のロボットビジョンなどにおいては有望であると思われる。またステレ

オ撮像による位置計算と併せて、濃淡画像処理手法を用いた情景解析<sup>33)</sup>などを併用することも容易である。ただしその実用化のためには、対応点探索処理をより信頼性のあるものにしていく必要がある。現在の対応点探索処理の問題点は、①ノイズ耐性、②見かけの変化に伴う誤対応、③処理速度などである。画像ノイズの除去は対象物の特徴を弱める効果を常に併せもつので、その設計には注意を要するが、対象物に関する知識(多面体など)を利用する画像処理について検討する必要がある。面が隠れたり現れたりするような見かけの変化に影響されない手法は、特にロボットビジョンなどにおいて、比較的近くにある対象物を認識する場合などには重要である。この場合、見かけの変化の程度は大きく、面や稜線などの関係を解析することや、対象物に対する事前知識を利用することなどが必要となろう。処理速度については、データ圧縮処理とアルゴリズムおよびハードウェアの問題と考えられる。このうちデータ圧縮に関しては、よく設計された画像処理によって画像内の2次元的特徴を必要に応じて選抜することによりかなり効果的に実行できると思われる。柔軟性の確保とのバランスにおいて、冗長な特徴をどのくらい保持できるかという選択が重要となるであろう。

## (2) 能動的ステレオ法

スリット光などの3次元的に構造化された照明光パターンを対象物に投影し、対象物と照明パターンとの干渉によるパターンの歪みから3次元情報を獲得する方法である。SRI International で開発された振幅変調したレーザー光の振幅および位相差から明るさと距離を同時に測定する方法<sup>12)</sup>は走査ミラーによる情景の走査とデータ処理にかなりの時間(1点あたり0.5秒)を要する。位置・姿勢の既知である投光器から縦長のスリット光<sup>40)</sup>やグリッド光<sup>35)</sup>を投影し三角測量する方法、カメラに対して水平方向の2枚の光平面を投影する方法<sup>38)</sup>などは非常に多くの変形・改良を伴って実際に使われている。対応点探索が非常に容易になっていることに加えて、面の傾きなどの情報もスリット光の密度分布から推定できるという利点もある。しかしながら、普通照明下での測定が難しいこと、測定点の密度が対象面の傾きによって異なること、距離測定のためだけに専用化されたデータであるので、他の情報を組み合わせて副次的に利用することが難しいことなどの欠点をもつ。レーザー光を利用した光切断法<sup>41)</sup>、モアレ縮投影法<sup>76)</sup>なども同様な特徴をもつ。

能動的ステレオ法の第1の利点は、対応点探索処理が不要あるいは非常に容易で



あることである。ただし、現在の能動的ステレオ法はそれぞれの目的に沿って専用化されており、安定性および高速性<sup>58)</sup>を確保するためには有効であると思われるが、3次元物体の柔軟な認識および多面的な情報処理を行うためには、受動的ステレオ法に比べて不向きであると思われる。

### (3) 3次元物体の認識法

計算機を応用するために適した3次元物体の表現あるいは記述を利用することは処理の効率やプログラミングの容易さを確保するために重要である。Binfordの一般化円筒 (Generalized Cylinder) 表現<sup>7)</sup>はスリット光を用いたレンジファインダ (測距装置) によって得られた距離データを圧縮して簡潔に表現するために導入された。特に、曲面物体の3次元表現として有効である。Marr<sup>15, 34)</sup>は階層的表現構造を提案した。そこでは画像の濃淡値の変化を捉えた原始スケッチ、両眼ステレオ視などによって得られた3次元情報を付加した2 1/2次元スケッチ、それらの連結による円筒表現という階層が提案されている。表現の各階層に従って、零交差<sup>34)</sup>、微小面素法線分布<sup>15)</sup>などが実現手段として提示されている。零交差やその符号は濃淡値の変化を捉えることに優れているが、緩やかな濃淡の傾斜をもつ部分では無効となるので、ランダムドットパターンを投影してステレオ計算を行う場合もある<sup>44)</sup>。また、微分(あるいは差分)計算に基づく特徴であるのでノイズ耐性が問題となる。

さて、実際の3次元物体の認識手法は数多いので、すべてを網羅するのではなく、代表的と思われる手法や、特徴のある手法を中心に展望することにする。1970年代中ごろ、3次元物体の2次元画像上の特徴量だけを用いる手法<sup>8, 16)</sup>が提案された。物体どうしの重なりを許さない場合<sup>8)</sup>と遮蔽されていてもある程度認識を可能とする場合<sup>16)</sup>がある。いずれも平面的な輪郭を記述する signature 関数 (代表的なものに輪郭に沿った距離対接線角度の関数がある) に基づいている。一方、実際の産業的応用における視覚認識<sup>24)</sup>を見てみると、ボルトの締緩用の識別システム、トランジスタペレットの位置決めシステムなどほとんどの視覚システムは種類の既知である対象物の位置決め用であった。70年代後半から80年代ははじめにかけて、それまでにパターン認識や画像処理の技術的基盤<sup>6, 11)</sup>があったことにも助けられて、弛緩法<sup>18, 29)</sup>、Hough変換の拡張<sup>30)</sup>、構造的パターン文法の応用<sup>22)</sup>などその後

利用される3次元物体認識や情景解析における重要な手法が提案されている。時を同じくして、距離画像を物体認識に応用する研究が始められた<sup>38, 40)</sup>。それらは専用のレンジファインダを用いたシステムであったが、距離画像上でのセグメンテーションの手法、距離データからの面の内挿法などに成果を残した。画像解析およびパターン認識の手法による物体認識は形状認識が主であるが、面情報を利用する方法<sup>25)</sup>、曲率を用いた輪郭の詳細な記述に基づく手法<sup>28)</sup>、簡単な形状モデルとの照合による手法<sup>45)</sup>などが提案されている。その後も、2次元のアプローチではあるが部分的に隠された対象物の分散した特徴に基づく方法<sup>63)</sup>などがある。

モデル内の情報からの推論の結果を利用する認識法を総称してモデル駆動型物体認識と呼ぶ<sup>50)</sup>が、特に3次元物体の形状に注目する場合には幾何モデルを利用する方法<sup>69, 70)</sup>が有効である。これらの方法は人工知能の分野における知識工学の手法と同じ動機をもつ。すなわち、対象物を顕著な特徴を用いて識別した後は逆にモデル内の情報を用いて推論を行い、処理の効率を高めるというやり方である。

## 1.3 本論文の構成

本論文の構成は次のようである。

第1章では、本研究において設定した問題の当該研究分野における位置付けについて述べる。

第2章では、前章で導入した問題を解決するための考え方、およびそのための基本的アプローチについて具体的に展開する。

第3章では、対象物の全体形状を認識するための手法として、通常の像 (直接像) および鏡像 (視野内に平面鏡を設置して撮像する) の両方を同時に用いる位置計算法について論ずる。

第4章では、対象物の異なる見え方を統合して全体形状を再構成する手法を提示する。ここでは、濃淡画像から頂点、稜線および面などの2次元の特徴量を抽出する画像処理手法に関しても検討し、実験を通して一連の処理の有効性について検討する。また、画像情報を有効に無駄なく取り込むための視野制御の手法についても検討する。



第5章では、図1.1で示した本研究の指向するロボットシステムの実験のプロトタイプについて述べる。認識系と操作系との結合方法を検討するための例として、ロボットアームに物体を把握させる作業を選び、実験システムを設計・試作した。ここでは、対象物の視覚的捕捉（視野制御）、輪郭頂点を利用した特徴抽出および3次元的特徴の利用の仕方、予め学習しておいた物体モデルを利用した認識法、センサを利用した把握制御法などを検討する。

第6章では、提示した各手法の問題点および解決の見通しについて考察する。撮像方式の評価および位置計算の誤差解析などの検討を行う。

第7章では、本研究で得られた成果についてまとめる。

## 第2章 基本的アプローチ

### 2.1 前提およびアプローチの概要

第1章で議論した方針に従って、本研究における前提について述べる。まず、3次元位置の計算法として、受動的ステレオ撮像法を拡張して利用する。その理由は大きな柔軟性および汎用性にある。次に、受動的ステレオ撮像法を構成する要素として、①対象物、②背景、③照明、④撮像装置、および⑤それらの相互関係、が考えられる。それぞれに関する具体的な前提条件を次のように設定する。

#### ①対象物およびモデル

- ・基本的には、多面体（凹凸）を対象とする。箱型機械部品などの近似物体として有効と考えられる。視覚認識の分担すべき役割から、本研究で扱う手法の精度を考慮すれば、多面体程度の近似が適当と思われる。実験では、多面体状の実物体も試みる。図2.1に対象物の例を示す。

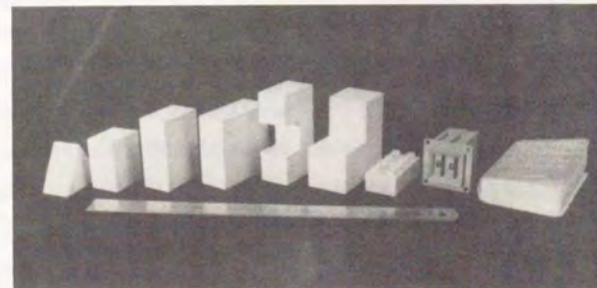


図2.1 対象物の例



- ・安定な姿勢で設置されている。
- ・単一対象物とする。複数の対象物の場合には、セグメンテーションを必要とする。
- ・基本的には、つや消し面とする。ただし、鏡面反射などの起こりにくい照明法を使用することでプラスチックなどの対象物も扱う。
- ・背景とのコントラストは充分大きいとする。

#### ②背景

- ・一様であるとする。
- ・対象物と充分なコントラストをもつ。

#### ③照明

- ・生活照明に近い可視光とする。
- ・スリット光などように構造化されたものではないとする。

#### ④撮像装置

- ・通常のITVカメラ、あるいはCCDカメラを用いる。
- ・撮像パラメタ（カメラの内部パラメタおよび外部パラメタ）は既知とする。

#### ⑤相互関係

- ・多面体の写り方において、面や稜線の縮退や重なりはないものとする。
- ・撮像装置の解像力は、被写界分解能で1mm以下とする。撮像距離は焦点距離などのパラメタとこの分解能から設定される。
- ・物体の影が写らないような照明方向を選ぶ。すなわち、雰囲気照明およびカメラと同軸の照明用電球とを併用する。

つづいて先に述べた3つの問題に対する具体的アプローチの内容について述べる。

#### (1) 全体的形状へのアプローチ

対象物の全体形状を認識するためには、まず、対象物の多くの可視側面を効率よく撮像することが重要であるとする。本研究ではそのために次の2つのアプローチを試みる。

①視野内に鏡面を設置して、対象物の前面と同時に裏側面の画像をも撮像し処理する方法。

②物体とカメラの相対的位置を制御して不可視の側面を能動的に撮像する方法。

この2つのアプローチはそれぞれ第3章、第4章において具体的に展開される。再構成するときの目標精度のオーダーは被写界分解能と等しいことを目安とする。

①の鏡面を利用するアプローチはこれまでにステレオ法とは異なる方法のなかで試みられている<sup>41)~43)</sup>。本研究では、直接像および鏡像の幾何学的関係から三角測量計算が可能であり、さらに双方に現れる可視部分を統合することにより対象物のかなりの形状が再構成できることに注目する。また一般的な座標系の設定に従った連立方程式による定式化を行い、応用における便宜を計る。

②のアプローチにおいては、カメラと対象物の相対的位置を変化させる手段としてターンテーブルを導入する。これはさらに一般的な手段である移動カメラによる撮像方式の基礎的段階を実現するものであると考える。この設定に基づいて、多面体構造の見かけの変化（面が現れたり隠れたりする）を能動的に解析することにより、適度な冗長度をもたせながら、用いる画像データの量を削減し処理を効率化する方法を検討する。

さらに、既対応点と未対応点の双方を含む対応面の解析を中心的な手続きとし、見かけの構造の変化に対処できるような対応点探索法を提案する。対象物の全体にわたる特徴点（頂点）の3次元位置を算出するためには、従来のステレオ撮像法におけるように2つの画像が大きく違わない設定を保持し続けることはできず、見かけの構造が大きく変化するような画像を扱う必要がある。そのためには、まず、構造の変化に対処可能な対応点探索処理が必要となる。第4章では、基本的な幾何学的対応関係により容易に対応の付けられる頂点を用いて、稜線および面の対応関係を調べ、その結果から逆に頂点の未知対応を判定する方法を考案する。

対象物の計算機内表現として、グラフ表現を用いることにする。このことにより、部分的形状の併合や、隠れ面を含む多面体形状の統一的扱いが可能となり、その一貫性の検査などを容易にすることができるようになる。



## (2) 画像処理および特徴抽出手法

多面体の濃淡画像から、面、稜線および頂点の連結や隣接関係などの線画構造を抽出するための画像処理手法を試みる。一様な色をもつ対象物の画像から幾何学的特徴を抽出する処理は、実対象物に対する処理の基礎となるものとして重要であり、実験によりその拡張性、および問題点などを明らかにしておく必要がある。

第4章では平滑化、微分、点列の直線化などの基本的な画像処理技術を組み合わせ、多面体の頂点、稜線および面の抽出過程を設計し実験により評価する。特に、画像処理手法を用いて抽出した頂点、稜線および面という幾何学特徴の隣接構造や連結構造を解析することにより、それらの3次元構造が推測でき、後の対応点探索処理を効率化できることを示す。このことは画像処理を伴う受動的ステレオ法の大きな利点の1つであると考えられる。

一方、第5章では、2値化した輪郭形状に基づく多面体認識法を試みる。これは先の一連の画像処理に比べて、より安定に稼働することを目的として設計されている。まず、濃淡画像の灰色度ヒストグラムに基づく自動閾値判定法により2値化、抽出された対象物の輪郭形状から頂点を検出する。この輪郭頂点を主な特徴とし、ステレオ法を用いてその3次元位置を算出し、位置・姿勢などを決定する。輪郭頂点を特徴として採用した理由は、①一般的に背景と対象物間の灰色度コントラストは対象物内部の面間のコントラストより高いことが期待でき、したがって、内部頂点より安定に検出できること、②輪郭形状の凸包を調べることにより、凸および凹多面体を一貫して扱うことが可能となること、などが挙げられる。

## (3) 認識系と操作系との結合様式

第5章では、物体の視覚認識系とハンドリングのための操作系との結合における問題点を明確にするために、実験的プロトタイプシステムを設計・試作する。その具体的な目的は小型CCDカメラを装備したロボットアームによる対象物の把握である。

まず、3次元対象物の姿勢の検出のために、座標値の共分散行列の固有ベクトル組という新しい特徴を提案する。

認識系と操作系の一般的な結合のために、対象物のモデルを導入する。ここにおけるモデルは対象物の識別用の特徴量（主に2次元形状の記述）、3次元位置・姿

勢を記述する特徴量、および把握操作時の制御に必要な対象物の寸法値の3種類の項目を含むものである。さらに、そのモデルを学習によって作ることを検討する。すなわち、モデル内の識別用および位置・姿勢記述用特徴量をすべて認識プログラムによって算出する。こうすることによって、まず、識別処理におけるパラメタである特徴量の統計的分散などを、モデルおよび対象物にわたって等しくすることができる。次に、ハンドの位置決め誤差を見かけ上減少することができる。一般にロボットアームの較正は難しい問題の1つであり、第5章の実験システムにおいては、認識系で算出した対象物の位置・姿勢と、操作系の位置決めのための座標系のずれが問題となる。認識プログラムの算出した位置・姿勢を用いてロボットアームの較正を行えば、固定的な系統誤差による位置ずれのような影響は排除することができる。

モデルを用いるアプローチでは、常にモデル数の有限性の問題がある。たとえば、選ばれた最適モデルと対象物との姿勢のわずかな差は常に存在する。ここでは、最適モデルの姿勢を対象物の姿勢に一致させるために、対象物の姿勢行列と最適モデル内に登録された姿勢行列との間の補正変換を行う方法を試みている。さらに、ロボットアームを利用したカメラの視野制御や、ステレオ画像の撮像、モデルを利用した対象物の把握制御などについて実験により検討する。

## 2.2 成果の概要

本研究で得られた成果について、その概要をまとめて述べる。

(1) 全体形状の認識手法に関しては次のとおりである。

- ①対象物の直接像および鏡像を利用する新しい手法を定式化し、目視による実験により有効性を確認した（第3章）。
- ②複数のステレオ画像の組を解析して得た部分形状を統合して全体形状を再構成する新しい手法を提示し、画像処理を含む一連の計算機処理により方法の有効性を確かめた（第4章）。
- ③多面体の面の対応関係や法線情報を利用する新しい対応点探索処理を試み、実



際の濃淡画像から抽出した線画構造に適用し良好な結果を得た（第4章）。

(2) 画像処理手法に関しては次のとおりである。

- ①多面体の濃淡画像から頂点、稜線および面を抽出する基本的な独自の手法を試み、良好な結果を得た（第4章）。
- ②多面体状の物体の輪郭を抽出するために、自動閾値決定法を試み、濃淡画像を2値化して得た輪郭画像から特徴点として頂点を抽出する一連の処理を提示した（第5章）。

(3) 認識系と操作系との結合様式の検討に関しては次のとおりである。

- ①3次元物体の姿勢を推定するために、輪郭頂点座標値の主軸ベクトルに加えて、共分散行列の第2および第3の固有ベクトルをも同時に用いる独自の方法を試み、良好な結果を得た（第5章）。
- ②対象物の形状および寸法を表す特徴を保持するモデルを利用し、対象物画像から算出した特徴との照合によって認識を行う独自の手法を試みた。照合時には、2次元特徴および3次元特徴を同時に利用する（第5章）。

その他の成果としては次のとおりである。

- ①ステレオ撮像による手法の撮像力を比較するための指標を試算した（第6章）。
- ②ステレオ撮像による位置計算法に伴う、画素量子化誤差、校正誤差などの系統誤差について定式化し、設計における基礎資料を得た（第6章）。

### 第3章 直接像および鏡像を利用したステレオ撮像による位置計算法

#### 3.1 概要

一般にステレオ撮像による3次元対象物の位置、姿勢および形状の測定法や認識法においては、対象物の可視の部分のみの処理に関するものが多い。通常のステレオ撮像法における撮像方式においては、ステレオ画像の対を構成する2枚の画像は、非常に小さな視差をもたせたものであり、したがって、そこに写し込まれる対象物の見え方（面の構成など）はほとんど変化しない。その理由には、まず、同一点の同定という対応点探索処理をなるべく容易に解決するということと、人間の両眼視からの素直な拡張であるという2つのことが考えられる。しかしながら、たとえば機械部品の自動識別などを指向するためには、ある決まった視線方向からでは見えない部分をも処理の対象として考慮する必要がある。本章では、そのための新しいステレオ撮像方式として、単一の画像により多くの視覚情報（画像情報）を取り込む手法について考える。

本章で述べる手法<sup>48,53,54,61,64)</sup>の基本的考え方は、まず対象物のより多くの部分を画像情報として取得し、それらの各々表す部分的形状を統合して対象物の全体形状を再構成しようとするものである。そのような部分の情報を得るためには、対象物と撮像装置の相対的位置関係を変えて、その部分の見える画像を得なければならない。そのために本手法の撮像方式においては、視野内に1枚（もしくは複数枚）の平面鏡を設置して対象物およびその鏡像の双方を単一の画像内に撮像する。鏡像側には直接像には映らない部分が現れるように平面鏡を設置する。対象物の直接像および鏡像の特徴点の対応関係に基づいて特徴点の3次元座標を計算する系統的方法を示す。具体的な画像処理方法は示さないが、計算機処理を実現する場合、対応点問題に関しては、無限消失点を利用して、画像上の対応する特徴点間の幾何学拘束を明らかにしている。本手法は、従来のステレオ撮像による位置測定法など比較して次のような特徴をもつ。

- (1) 1枚の画像を用いてステレオ撮像を行うことができる。
- (2) 対象物の隠れ面（もしくはその面上の点）に関する3次元情報を得ることが



できる（ただし、後述する計算可能性の問題が付随する）。

- (3) 撮像パラメタの較正は、1台のカメラのパラメタと1枚の平面鏡の空間位置・姿勢であるから、2台のカメラの場合に比べて容易である。

また、平面鏡を利用する手法はこれまでもある<sup>41)~43)</sup>が、本手法における使い方は異なるものである。さらに、光本ら<sup>82)</sup>は本節で述べる手法を応用すると直接像と鏡像との間に面対称性（鏡面に関して）が成立することに着目し、本章で述べる手法を拡張して、鏡像を利用した3次元物体の復元法を考案している。

### 3.2 特徴点の位置計算の定式化

三角測量の原理は古くから知られており、様々に応用されてきているが、本節では、ロボットビジョンとしての系統利用を容易にするために、同次座標変換などを用いて定式化する。

本方法において用いられる座標系およびその構成を図3.1に示す。世界座標系（World Coordinate System, WCS）は対象物および他の座標系の位置決め基準となる。実際の撮像装置のモデルとして、レンズ中心前方 $f$ （実際に使用するカメラの焦点距離に相当する）の位置に仮想画像面をもつピンホールカメラモデルを採用する。その仮想画像面と $X_c, Y_c$ 座標面が一致するようにカメラ座標系（Camera Coordinate System, CCS）を設定する。CCSの $Z_c$ 軸をカメラの光軸と一致させ、その正の向きをカメラの前方とする。カメラはWCSの原点側を見込む姿勢をとる。画像座標系（Image Coordinate System, ICS）として、実際の画素幅で量子化された座標値をもつ2次元座標系を仮想画像面上に設け、画像上の各画素の位置を pixel 単位で定める。

次節以下で用いる座標変換行列を説明する。WCSからCCSへの座標変換行列を $\hat{T}_{0c}$ と表す。さらに、CCSにおける透視変換行列を $\hat{T}_c$ として次式で定義する。

$$\hat{T}_c = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & r \end{pmatrix}, \quad r \triangleq \frac{1}{f}$$

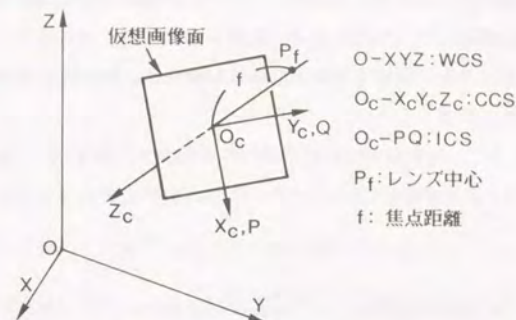


図3.1 座標系の設定

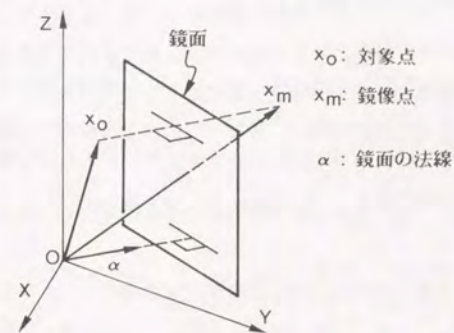


図3.2 鏡面による反射の幾何



$\hat{T}_p$ は光軸上( $z_o = -f$ )のレンズ中心 $P_l$ を中心とした単点透視変換を表す。なお記号 $\hat{\cdot}$ は同次変換表現を示す。

ここでは視野内に1枚の平面鏡を設置する場合を扱うが、複数枚設置する場合の扱いも原理的には同様である。

図3.2に示すように、対象物を見込む視野内に平面鏡を設置する。鏡面の方程式を次式で定義する。

$$x^t \alpha = k$$

ここで、 $x$ はWCSにおける鏡平面上の点を表す位置ベクトル、 $\alpha$ は鏡平面の法線ベクトル、 $k$ は次のような定数である。

$$k = \begin{cases} 0 & (\text{鏡平面がWCSの原点を通る場合}) \\ 1 & (\text{鏡平面がWCSの原点を通らない場合}) \end{cases}$$

多面体対象物の頂点のような3次元の特徴点を対象点 $x_o$ と呼ぶ。鏡面により対象点 $x_o$ は鏡像点 $x_m$ に次のように変換される。

$$x_m = \left[ I_{3 \times 3} - \frac{2}{|\alpha|^2} \alpha \alpha^t \right] x_o + k \frac{2}{|\alpha|^2} \alpha$$

ここで、 $I_{3 \times 3}$ は $3 \times 3$ 単位行列、 $|\alpha|$ は $\alpha$ のユークリッドノルムを表す。さらに、同次座標表現を用いて上式は、

$$\hat{x}_m = \hat{T}_m \hat{x}_o$$

となる。ここで、

$$\hat{x}_o \triangleq [x_o^t, 1]^t$$

$$\hat{x}_m \triangleq [x_m^t, 1]^t$$

$$\hat{T}_m \triangleq \left( \begin{array}{c|c} I_{3 \times 3} - \frac{2}{|\alpha|^2} \alpha \alpha^t & k \frac{2}{|\alpha|^2} \alpha \\ \hline O_{3 \times 1} & 1 \end{array} \right)$$

ステレオ撮像による位置計算を行うためには対象点 $x_o$ とその鏡像点 $x_m$ を画像上で検出し、それらの2次元位置座標値(画像座標系)を用いて、 $x_o$ の3次元位置座標を変数とする連立方程式を導かなければならない。このとき、以下の仮定をおく。

【仮定1】変換行列 $\hat{T}_{cu}$ 、 $\hat{T}_{m0}$ および $\hat{T}_p$ は既知とする。

【仮定2】ICS上の各画像点の座標は既知とする。

【仮定3】対象点とその鏡像点の画像上の対応関係は既知である。

【仮定1】は、いわゆる撮像パラメータを含むパラメータが既知であることを示す。変換行列 $\hat{T}_{cu}$ 、 $\hat{T}_{m0}$ はそれぞれ世界座標系におけるカメラおよび平面鏡の位置と姿勢を示している。行列 $\hat{T}_p$ は、カメラの内部パラメータの焦点距離 $f$ が正確にわかっていることを示している。【仮定2】はカメラの画面上において、特徴点が識別され、その位置が求められていることを示している。実際には、計算機による画像処理を用いるか、あるいは目視により座標を読み取るかはここでは問わない。【仮定3】は直接像と鏡像にわたる同一点の同定がすべて既知であるということを示す。計算機によるステレオ計算などにとって、この仮定を実現することは重要である。これに関する対応点探索方法については第3.3節で述べる。

連立方程式の導出手順を次の8段階に分けて導く。

① 対象点およびその鏡像点をWCSからCCSへ座標変換し、さらに透視変換を施した点を $\hat{x}_{co}$ および $\hat{x}_{cm}$ と呼ぶ。

$$\hat{x}_{co} = \hat{T}_p \hat{T}_{cu} \hat{x}_o \quad (1)$$

$$\hat{x}_{cm} = \hat{T}_p \hat{T}_{cu} \hat{T}_{m0} \hat{x}_o$$

②  $\hat{x}_{co}$ および $\hat{x}_{cm}$ の第4要素をそれぞれ $w_1$ および $w_2$ とする。それらは次式で表される。



$$w_1 = (r(\hat{T}_{cu})_3 + a) \hat{x}_0 \quad (2.a)$$

$$w_2 = (r(\hat{T}_{cu})_3 + a) \hat{T}_{no} \hat{x}_0 \quad (2.b)$$

ここで、 $a \triangleq (0, 0, 0, 1)$ 、 $(\hat{T}_{cu})_3$  は行列  $\hat{T}_{cu}$  の第3行ベクトルを表す。

③  $p$  を ICS 上に透視投影された画像点を示す2次元位置ベクトルとする、 $\hat{x}_0$  および  $\hat{x}_n$  に対応する画像点をそれぞれ  $p_0$  および  $p_n$  とすると、式(1)および(2.a)(2.b)から、

$$w_1 p_0 = I_{2 \times 4} \hat{T}_p \hat{T}_{cu} \hat{x}_0$$

$$w_2 p_n = I_{2 \times 4} \hat{T}_p \hat{T}_{cu} \hat{T}_{no} \hat{x}_0$$

となる。ここで、 $I_{2 \times 4} = [I_{2 \times 2} \mid O_{2 \times 2}]$  である。

④ 上式を行列方程式の形式にまとめると、

$$W s = \begin{pmatrix} I_{2 \times 4} \hat{T}_{cu} \\ I_{2 \times 4} \hat{T}_{cu} \hat{T}_{no} \end{pmatrix} \hat{x}_0$$

となる。ここで、 $W \triangleq \text{diag}(w_1, w_2, w_3, w_4)$ 、 $s \triangleq (p_0^t, p_n^t)^t$  である。標記を簡単にするために、上式の右辺の係数行列を新たに  $A_1$  と表す。すなわち、

$$A_1 \triangleq \begin{pmatrix} I_{2 \times 4} \hat{T}_{cu} \\ I_{2 \times 4} \hat{T}_{cu} \hat{T}_{no} \end{pmatrix} \quad (3)$$

⑤ 第4座標  $w_1$  および  $w_2$  は  $\hat{x}_0$  の要素を含むので、それらから  $\hat{x}_0$  の要素を未知数としてくり出さなければならない。

$$W s \triangleq A_2 \hat{x}_0$$

ここで、

$$A_2 \triangleq \begin{pmatrix} p_0 (r(\hat{T}_{cu})_3 + a) \\ p_n (r(\hat{T}_{cu})_3 + a) \hat{T}_{no} \end{pmatrix} \quad (4)$$

⑥ 以上で  $\hat{x}_0$  に関する次の同次連立方程式を得る。

$$(A_1 - A_2) \hat{x}_0 = O_{4 \times 1}$$

⑦  $\hat{x}_0$  の第4座標は常に1であるから、係数行列  $(A_1 - A_2)$  から定数ベクトル  $b$  を分離できる。すなわち、

$$A_1 - A_2 \triangleq [A \mid b]$$

⑧ 次の連立方程式を得る。

$$A x_0 = -b$$

上式は、たとえば最小2乗法を用いて次のように解くことができる。

$$x_0 = -(A^t A)^{-1} A^t b \quad (5)$$

### 3.3 対応点探索問題に関する考察

#### (1) 画像点間の幾何学的関係

対象点および鏡像点の画像上の座標が求められると、次に、それらの画像点間の対応付けを ICS 上で行うための幾何学的拘束を導く必要がある。一般のステレオ撮像による位置計算法では、この対応点問題の解法のために、epipolar line (2台のカメラのそれぞれのレンズ中心と抽出された画像点のうちの片方を含む平面と、もう一つの画面との交線) を幾何学的拘束を表す直線として用いる。すなわち、この直線は既知であるカメラパラメタ (カメラの空間的位置・姿勢、レンズ光学系の焦点距離、主点位置、中心点位置など) と、投影された特徴点の2次元位置だけから計算可能であり、対応すべき点の対(組)は原理的にはこの直線上にあるはずであるので、それらの探索に利用できる。

鏡面は平面であるから、任意の対象点とその鏡像点を結ぶ直線はすべて鏡平面に直交し、互いに平行である。したがって、それらを ICS 上に透視投影すれば、図 3.3 に示すように、1つの無限消失点  $p_v$  に収束することになる。拘束ベクトル  $\delta$  を



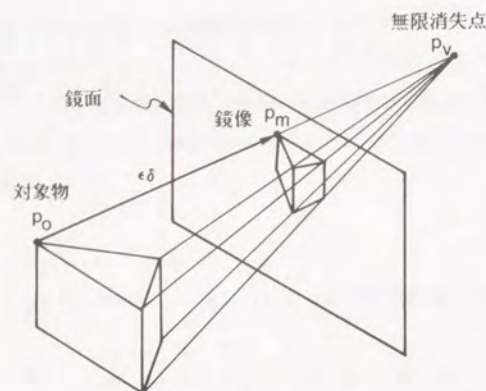


図3.3 対応点対と無限消失点の幾何学的関係

$$\delta \triangleq p_v - p_o$$

と定義すると,

$$p_n = p_o - \varepsilon \delta, \quad 0 \leq \varepsilon \leq 1$$

となり,  $p_n$  の位置は  $p_o$  と  $p_v$  を結ぶ直線上に拘束される. ICS上で  $p_o$  もしくは  $p_n$  の座標がわかれば, 上式を用いて残りの画像点を探索し, 対応付けすることができる.

## (2) 無限消失点の位置

無限消失点  $p_v$  の位置は簡単な幾何学的考察から, ICS上で次のように求まる.

$$p_v = \left[ \frac{(\hat{T}_{cw})_1 \alpha'}{r (\hat{T}_{cw})_3 \alpha'}, \frac{(\hat{T}_{cw})_2 \alpha'}{r (\hat{T}_{cw})_3 \alpha'} \right]^t \quad (6)$$

ここで,  $\alpha' \triangleq [\alpha^t, 0]^t$  である.

$p_v$  は特徴点の位置によらず, カメラと鏡平面それぞれの位置および姿勢だけから計算できる.

## 3.4 隠れ面上にある特徴点の位置計算

対象物の隠れ面を鏡面を介して撮像し, その面上の特徴点の3次元位置を計算できる場合について考える. 図3.4に示すように, 対象点  $x_{o4}$  は隠れ面上にあり, 第3.2節で述べた方法では計算できない. しかし, 平面  $F_1$  については, 3個の鏡像点  $x_{ni}$  ( $i=1,2,3$ ) の位置が第3.2節の方法で求められるので, その位置と姿勢を決定できる. 何らかの方法で対応する平面が決定できたとすれば, 鏡像点  $x_{n4}$  の位置は平面  $F_1$  と視線  $P_1 p_{n4}$  (点  $P_1$  と  $p_{n4}$  で示される点とを通る直線) の交点として計算でき,  $x_{n4}$  に逆変換  $T_{n0}^{-1}$  を施し,  $x_{o4}$  の位置が得られる.



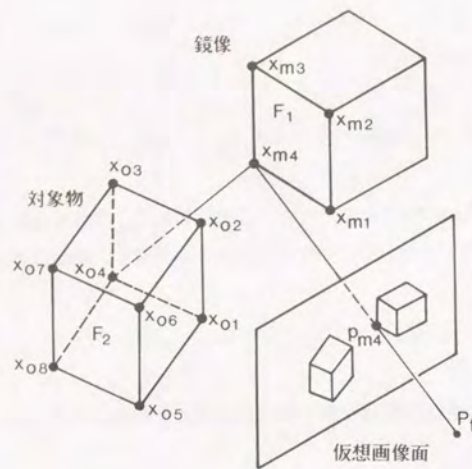


図3.4 隠れ面上の頂点の位置決定

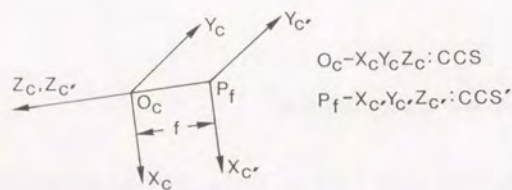


図3.5 補助座標系

#### (1) 平面の法線の決定

レンズ中心  $P_f$  を含む平面は、透視投影により  $ICS$  上で線に縮退する。したがって、対象物およびその鏡像はカメラに対して一般的位置にあり、それらを構成するいかなる平面ないしその延長もレンズ中心  $P_f$  を含まないと仮定する。多面体対象物の場合、このような設定は常に可能である。

図3.5に示すように、 $CCS$  を  $Z_c$  軸に  $-f$  だけ平行移動した補助的な3次元座標系  $CCS'$  を設定する。 $CCS'$  における任意の平面の方程式は  $\gamma$  をその法線ベクトルとすると、

$$(p^t, f) \gamma = t$$

となる。ここで、 $p$  は  $ICS$  上の2次元位置ベクトル、 $t$  ( $0 \leq t \leq 1$ ) はレンズ中心からの奥行きを表すスケール因子である。 $CCS'$  における位置ベクトル  $x$  とその画像点  $p$  の関係は次式で表される。

$$t x = (p^t, f)^t$$

上式を用いて、位置の既知である鏡像点の座標とその画像点の座標の組  $\{x_{mi}, p_{mi}\}$  ( $i=1, 2, \dots, n, n \geq 3$ ) を代入して得られる次の連立方程式を解いて  $\gamma$  を求める。

$$\begin{pmatrix} p_{m1}^t, f \\ p_{m2}^t, f \\ \vdots \\ p_{mn}^t, f \end{pmatrix} \gamma = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} \quad (7)$$

#### (2) 平面と視線の交点の決定

法線ベクトル  $\gamma$  をもつ平面  $F_1$  と視線  $P_f p_{m4}$  の交点として鏡像点  $x_{m4}$  が次式でめられる。



$$x_{m4} = \frac{\begin{pmatrix} p_{m4} \\ f \end{pmatrix}}{(p_{m4}^T, f)^T} \quad (8)$$

これを用いて、対応する対象点  $x_{o4}$  は次式で求められる。

$$\hat{x}_{o4} = \hat{T}_{eo}^{-1} \hat{T}_{cu}^{-1} \hat{T}_{co}^{-1} \hat{x}_{m4} \quad (9)$$

ここで、

$$\hat{T}_{co}^{-1} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & f \\ & & & 1 \end{pmatrix}$$

はCCSからCCSへの座標変換行列である。

ここで述べた方法は、図3.4の  $x_{o8}$  のように鏡像対をもたない対象点に対しても適用できる。この場合、既知の3個以上の対象点、たとえば図中の  $x_{oi}$  ( $i=5,6,7$ ) を用いて直接対象物の平面  $F_2$  を決定し、以下同様の手続きで  $x_{o8}$  の位置を求める。

### (3) 特徴点の計算可能性に関する考察

これまでに説明してきた鏡像を利用した多面体頂点の位置計算法においては、次の2種類の計算可能性がある。

- ① 直接像および鏡像の両方において可視である場合。
- ② 上記①項は満足しないが、どちらか一方において可視であり、かつ、本節で述べたような位置の決定可能な平面上にある場合。

1つの対象物の形状を計測し再構成するためには、なるべく多くの頂点が計算可能となるようにカメラおよび鏡平面の位置・姿勢、対象物のカメラおよび鏡平面に対する相対的位置・姿勢を選ぶ必要がある。しかしながら、その問題について一般

的に議論するためには、対象物の形状に対する一般的な取扱いを必要とする。そこで本節では、本手法を実際に利用する場合に、どのような（どのように見える）特徴点が計算可能であるかについて、より具体的に検討することにする。議論を始める前に、ここでの基本的前提を述べる。

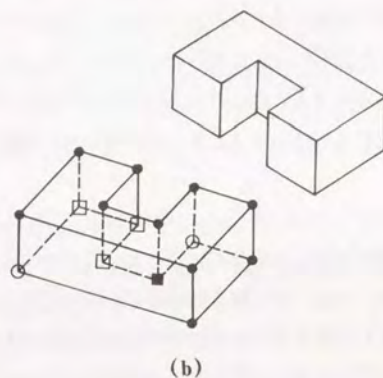
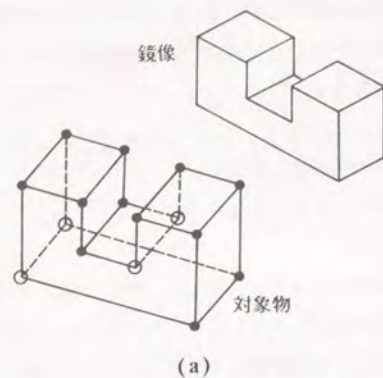
【前提1】対象物は1つの多面体とする。

【前提2】対象物の特徴点として、頂点を選ぶ。

頂点の3次元位置座標値が前節までに説明した一連の方法によって計算されうる時、その頂点は計算可能であるとする。この計算可能性は各頂点に固有の性質ではなく、その見え方に依存し、対象物だけではなく、カメラや鏡平面の位置および姿勢によっても変化する性質である。したがって、それらすべてについて考察することが望ましい。しかしながら、位置計算法が対象物の画像に基づく方法であるので、画像における対象物見え方による計算可能性の変化を検討する。2台のカメラを使用する一般のステレオ撮像による位置計算法においては、頂点の計算可能性は2枚の画像間での対応頂点の見え方に依存する。それに対して、1枚の画像内に対象物の直接像と鏡像を含む本手法の場合には、直接像および鏡像における頂点の見え方によっている。

図3.6に具体的な例を示す。図3.6(a)の場合、すべての頂点が計算可能である。すなわち、‘●’の印の付いた頂点は直接像および鏡像の両方において可視であるために計算可能である。‘○’印の頂点は直接像あるいは鏡像のどちらか一方で不可視であるが、いずれも本節で述べた位置の決定可能な平面上にあり、したがって視線との交点として計算できる。一方、図3.6(b)は同じ形状の対象物を異なる姿勢（転倒した状態）に置いた場合を示す。直接像および鏡像の双方ともにおいて不可視の場合（‘■’の頂点、1個）は勿論計算不可能であるが、このほかに、鏡像において平面の位置決定ができないために計算不可能な頂点（□）も存在する。これらの頂点は鏡像においては可視であるけれども、それらを含む平面が3点以上の既知頂点をもたないためである。興味あることは、(b)の場合に‘■’の頂点の位置が与えられれば、すべての頂点が計算可能になるということである。このような多面体のような立体構造では、その計算可能性の伝播の仕方が稜線や面を介した





注) (a) ‘●’ 頂点は直接像および鏡像の双方で可視である (計算可能)．‘○’ 頂点は平面と視線の交点として計算できる (計算可能)．(b) ‘■’ 頂点は直接像および鏡像の双方で不可視である (計算不可能)．‘□’ 頂点は鏡像では可視である (計算不可能)．

図3.6 頂点の計算可能性

頂点どうしの関係に依存している。

ここで述べた計算可能生は頂点の可視・不可視に依存するものであるが、上に述べたようにある頂点座標がまわりの頂点の座標からどのように拘束されているかというその多面体固有の性質にも依存する。直接には同一平面上の頂点からの拘束があり (平面が決まれば未知の頂点はその平面と視線の交点として拘束され任意の位置には存在できない)、間接的にはその他の頂点からも拘束される。このような意味で、一般に多面体の線画を与えるとき、最低幾つの頂点の3次元座標値を与えればその多面体の3次元形状が決定するかという問題に対して、杉原<sup>17)</sup>は多面体線画の“階数”を定義しその自由度について考察している。

### 3.5 座標変換行列の校正

この節では、カメラおよび鏡面の校正について考察する。本手法では、2台のカメラではなく1台のカメラと1枚の平面鏡を使用する。平面鏡については外部寸法と位置および姿勢を決めればよく、そのために通常のステレオ撮像による位置計算法に比べて、校正に要する手間は少ない。これは本手法の特長の1つである。

ここでは、特に測定の難しいカメラパラメタの校正法について考察する。ここでは、レンズ系の焦点距離  $f$  は既知であるとする。この仮定は固定焦点のレンズを使用すれば通常成立する。したがって、校正手続きはWCSからCCSへの座標変換行列  $\hat{T}_{cw}$  を求めることに相当する。行列  $\hat{T}_{cw}$  の各要素を

$$\hat{T}_{cw} \cong \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (10)$$

と定義する。このとき、校正点  $x_i = [x_i, y_i, z_i]^T$  を透視変換したときの同次座標は

$$w_i \cong r a_{31} x_i + r a_{32} y_i + r a_{33} z_i + r a_{34} + 1$$



となる。この式の同次座標は一意に定まらないので、次のように正規化する。

$$\begin{aligned}\bar{w}_i &\triangleq \frac{w_i}{r a_{34} + 1} \\ &\triangleq r \bar{a}_{31} x_i + r \bar{a}_{32} y_i + r \bar{a}_{33} z_i + 1\end{aligned}\quad (11)$$

ここで、 $\bar{a}_{ij} = a_{ij} / (r a_{34} + 1)$  である。 $x_i$  を ICS 上に透視変換して得られる画像点を  $p_i \triangleq (p_i, q_i)^t$  とすると、

$$\bar{w}_i p_i = \bar{a}_{11} x_i + \bar{a}_{12} y_i + \bar{a}_{13} z_i + \bar{a}_{14} \quad (12.a)$$

$$\bar{w}_i q_i = \bar{a}_{21} x_i + \bar{a}_{22} y_i + \bar{a}_{23} z_i + \bar{a}_{24} \quad (12.b)$$

を得る。式(11)を式(12.a)および(12.b)に代入して未知数 $\bar{a}_{ij}$ について整理すると次式を得る。

$$h_{pi}^t \theta = -p_i \quad (13.a)$$

$$h_{qi}^t \theta = -q_i \quad (13.b)$$

ここで、

$$\begin{aligned}h_{pi} &\triangleq (-x_i, -y_i, -z_i, -1, 0, 0, 0, 0, r p_i x_i, r p_i y_i, r p_i z_i)^t \\ h_{qi} &\triangleq (0, 0, 0, 0, -x_i, -y_i, -z_i, -1, r q_i x_i, r q_i y_i, r q_i z_i)^t \\ \theta &\triangleq (\bar{a}_{11}, \bar{a}_{12}, \bar{a}_{13}, \bar{a}_{14}, \bar{a}_{21}, \bar{a}_{22}, \bar{a}_{23}, \bar{a}_{24}, \bar{a}_{31}, \bar{a}_{32}, \bar{a}_{33})^t\end{aligned}$$

n 較正点について式(13.a), (13.b)を作り、まとめると次の連立方程式を得る。

$$H \theta = -u \quad (14)$$

ここで、

$$H \triangleq (h_{p1}, h_{q1}, h_{p2}, h_{q2}, \dots, h_{pn}, h_{qn})^t$$

$$u \triangleq (p_1, q_1, p_2, q_2, \dots, p_n, q_n)^t$$

式(14)を $\theta$ について解く。式(14)において、形式的な未知数の数は 11 個であ

るが、式(10)の行列が座標変換行列であることを考慮するとその自由度は6であり、いずれか5個の未知数は冗長であることになる。したがって、式(14)を解く場合にもその冗長性を排除するような拘束条件を必要である。いま、

$$K \triangleq r a_{34} + 1$$

とすれば、行列 $\bar{T}_{cu}$ が正規直交行列であることから、係数 $\bar{a}_{ij}$ について次のような条件がある。

$$\sum_{i=1}^3 \bar{a}_{ij}^2 = 1/K^2 \quad (j=1, 2, 3)$$

$$\sum_{j=1}^3 \bar{a}_{ij}^2 = 1/K^2 \quad (i=1, 2, 3)$$

新たな未知数 $K$ を1個導入したが、係数 $a_{ij}$ を $\bar{a}_{ij}$ に変換したときに $a_{34}$ を削除してあるので、上の6つの拘束条件の導入により冗長性が排除できる。たとえば、ラグランジュの未定乗数法などにより計算することができる。

得られた $\bar{a}_{ij}$ により構成される行列を $\bar{T}_{cu}$ とすると、次式が成り立つ。

$$\bar{T}_{cu} = (r a_{34} + 1) \bar{T}_{cu} \quad (15)$$

ただし、行列 $\bar{T}_{cu}$ の3行4列の要素は1とする。 $\bar{T}_{cu}$ を用いる場合には、先に式(3)および(4)で定義した係数行列 $A_1$ および $A_2$ を次のように変更すればよい。

$$A_1 \triangleq \left( \frac{I_{2 \times 4} \bar{T}_{cu}}{I_{2 \times 4} \bar{T}_{cu} \bar{T}_{no}} \right) \quad (16)$$

$$A_2 \triangleq \left( \frac{p_o (r \bar{a}_{31}, r \bar{a}_{32}, r \bar{a}_{33}, 1)}{p_o (r \bar{a}_{31}, r \bar{a}_{32}, r \bar{a}_{33}, 1) \bar{T}_{no}} \right) \quad (17)$$

### 3.6 実験

ここで提示した位置計算法を用いて、基本的な実験を行った。表3.1は使用し



た装置の仕様を示す。表3.2は校正された変換行列 $\hat{T}_{cu}$ および鏡平面の法線ベクトル $\alpha$ を示す。校正点として多面体の20個の頂点を選び、別に開発した画像処理システムPDMS<sup>52)</sup>を用いて、目視により画像上の頂点の位置を測定した。図3.7は位置計算の実験に使用した濃淡画像および頂点の番号付けを示す。表3.3は頂点位置の測定値と計算値を示す。‘\*’付きの頂点は直接像あるいは鏡像の一方で可視でないために、決定された平面と視線との交点として計算された。これらの点の計算誤差は他の可視頂点に比べて大きい。これは隠れ面を計算するときに用いられた頂点の位置誤差が平面を位置決めする場合に累積したためと思われる。表3.4は誤差の平均値と標準偏差を示す。位置の最大誤差13.5mm(点22のY座標)、標準偏差の最大値4.60mm(Y座標)は撮像距離696.5mmのそれぞれ1.94%および0.66%であった。図3.8は再構成された多面体を示す。

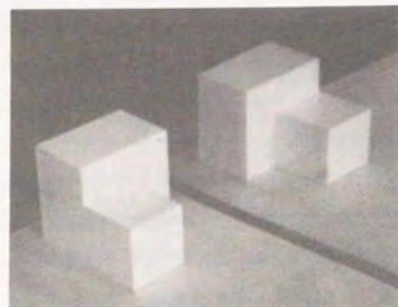
表3.1 実験の仕様

焦点距離	35 mm
撮像距離	696.5 mm
撮像装置	ITVカメラ(1インチ・ビジコン)
有効画面	9.6×12.8 mm
分解能	25.6 pikel/mm, 256×384×8ビット×4ページ

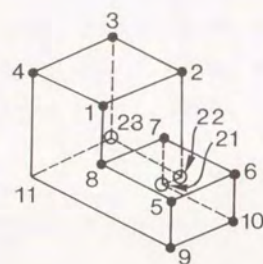
表3.2 校正された変換行列 $\hat{T}_{cu}$ および鏡面の法線 $\alpha$

$\hat{T}_{cu}$	0.355	0.340	-0.831	1.96
	-0.836	0.620	-0.153	1.37
	-0.610	-0.653	-0.477	696.5
	0.0	0.0	0.0	1.0
$\alpha$	$(a, 0, 0)^t, a \neq 0$			





(a) 原画像



(b) 頂点番号

図3.7 実験に用いたステレオ画像

表3.3 頂点位置座標値の測定値と計算値の比較

頂点番号	x 座標値	y 座標値	z 座標値
1	74.0	33.0	51.5
	73.5	32.5	53.5
2	23.0	33.0	51.5
	22.0	29.5	51.5
3	23.0	2.0	51.5
	22.0	-8.0	50.0
4	74.0	2.0	51.5
	72.0	-5.0	53.0
5	74.0	65.0	31.0
	75.5	68.5	31.5
6	43.0	65.0	31.0
	43.5	65.0	30.0
7	43.0	33.5	31.0
	43.0	27.0	27.5
8	74.0	33.5	31.0
	73.5	30.5	29.5
9	74.0	65.0	0.0
	76.0	61.5	-5.5
10	43.0	65.0	0.0
	44.0	59.0	-6.5
11*	74.0	2.0	0.0
	73.5	-1.5	-2.0
21*	42.0	33.5	0.0
	42.5	21.0	-8.0
22*	23.0	33.5	0.0
	20.0	20.0	-7.0
23*	23.0	2.0	0.0
	11.0	-6.0	0.0

注) 上段は測定値, 下段 は計算値を示す.



表3.4 計算誤差

評価項目	x 座標値	y 座標値	z 座標値
絶対値平均誤差	0.95 1.86	4.27 5.79	2.32 2.89
平均誤差	-0.05 -1.07	-3.64 -5.29	-1.59 -2.32
標準偏差	1.14 3.29	3.56 4.60	2.57 3.14

注) 上段の数値は頂点1~10までのデータから計算した値.

下段の数値は1~23までのデータから計算した値.

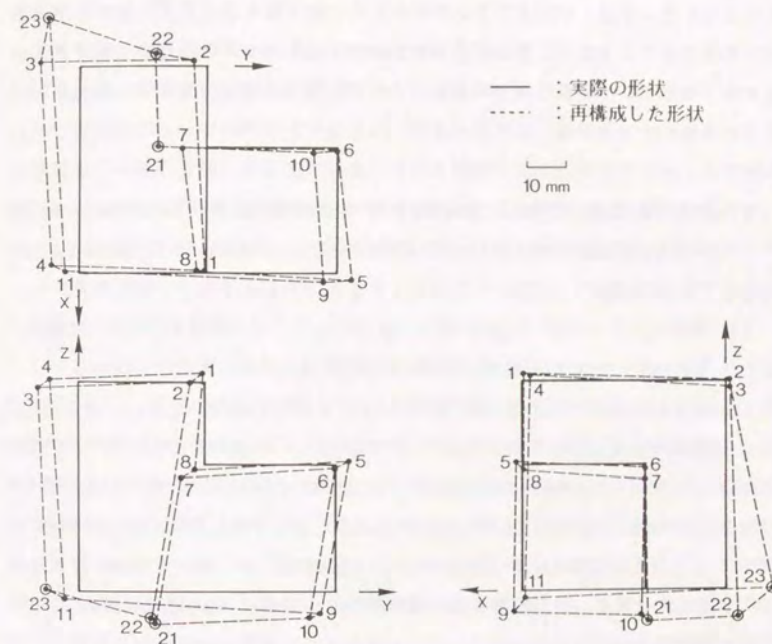


図3.8 再構成した形状



### 3.7 まとめ

本章では、3次元多面体対象物の頂点の位置計算のために、その直接像と鏡像を写し込んだ1枚の画像を用いて等価的にステレオ視計算を実現する方法を定式化した。そこにおいては、対応点問題の解決のために無限消失点に関する幾何学的拘束が利用できることを示し、多面体対象物に関する隠れ面上の特徴点の位置計算法も示した。あわせて、凹多面体を対象として実験を行い本手法の有効性を確かめた。以下に結果をまとめる。

- (1) 視野内に鏡面を設置し、多面体対象物の直接像と鏡像を撮像した単一画像を解析することにより、対象物の特徴点(頂点)の3次元位置が計算できることを示した。
- (2) 画像上の2つの対応点は無限消失点に関して一定の幾何学的拘束を満たし、その拘束は対応点の探索に利用できることを示した。
- (3) 対象物およびその鏡像の双方において、可視の特徴点の位置は、連立方程式を解き、直接求められることを示した。また、直接像あるいは鏡像の一方でしか可視でない特徴点の位置は、それを含む平面が先に計算された特徴点を3つ以上含み、その法線ベクトルが決定できる場合に計算可能である。
- (4) 多面体を対象物とした実験において、撮像距離 696.5mm、対象物 51×51×63mm の場合、最大位置誤差は撮像距離の 1.94%、最大標準偏差は 0.66% であった。

## 第4章 多視野のステレオ画像対を利用した多面体物体の認識

### 4.1 概要

本章では、普通照明下で撮像した通常の濃淡画像に基づいて3次元多面体の「全体形状」を再構成する体系的な手法<sup>79, 86)</sup>について述べる。併せて、画像処理を含めた自動認識のための基礎的な実験を行い、提示する手法の精度等について確かめる。第3章で述べた直接像と鏡像を用いたステレオ撮像による位置計算法も、対象物の部分的形状のみならず、全体形状の再構成を指向するものである。本章で述べる手法は更に進んで、対象物とカメラの相対的位置および姿勢を変化させて、欠落している未知の部分の形状を再構成するために画像データを取り込み位置計算するためのものである。

ステレオ撮像の原理に基づいて3枚以上の画像を利用して3次元の位置計算を行う法はいくつか提案されている<sup>39, 59)</sup>。それらの方法の主に意図するところは、画像間での対応点探索処理や面の再構成を容易にすることである。しかしながら、対応付けに際して画像間で面や稜線の2次元的トポロジーが変化しない場合(多面体の頂点が見えなくなったり、新しい頂点が現れたりしない場合)のみを扱うものや、凸多面体を対象とし隠れ面等を扱わないものも多い。ある特定の視野において可視である部分だけでなく、対象物の全体形状を捉えるためには、それを取り囲む方位のうちから必要な複数の視野を含む画像を得て、それらから得られる部分的な情報を統合することが必要であり、このことと構造的照明による撮像方法を併用して全体形状を復元する試みが行われている<sup>18, 41)</sup>。

それに対して、本手法の特色は、構造的照明などを用いない通常の撮像環境で得られる濃淡画像を用い、かつなるべく少ない枚数の画像データにより、効率良く3次元対象物の全体形状を再構成することにある。すなわち、全体形状を再構成するために、視野の異なる複数枚の画像(以下、多視野画像と呼ぶ)を撮像し、それらを処理して得られる対象物の異なる側面の形状情報を統合する手法について検討する。

ここでは、カメラの視軸(光軸)方向を変化させるためには、対象物を固定して



カメラがその周囲を運動するのが一般的であるが、ここではカメラを固定し、対象物をターンテーブルの中央付近にのせて適宜回転させることにする。カメラを任意に運動させる場合に比べると、カメラの傾斜角度（チルト角度）および撮像距離が固定となるために、視軸方向と視野の広さの自由度は減少するが、対象物とカメラとの相対的な位置・姿勢にある程度の拘束を付加することによって、計算を単純化でき、更に視野を決定するパラメタをターンテーブルの回転角度だけとすることができる（この場合には、カメラの実際の視野は固定であるが、対象物の異なる側面を観測できるという意味で多視野画像ということばを使用する）。カメラを任意移動させる場合との比較考察は後に述べる。多視野画像は1組2枚ずつのいわゆるステレオ画像の対として処理され、各視野において可視である部分の3次元情報（頂点の3次元位置、面の法線方向等）を求める。次に、各視野にわたって「共通に」含まれる頂点を探索し、それらを手がかりとして、共通稜および共通面の統合を行い全体形状を再構成する。ターンテーブル上に多面体を設置してその寸法を計測する手法についてはすでに発表された研究<sup>14)</sup>があるが、本章で述べる手法と比較して、画像処理を行っていないこと、頂点を目視により人手で与えていること、全体形状に統合していないことなどの違いがある。本章で述べる手法の特徴を以下に示す。

- (1) ターンテーブルを回転させ、固定の単一視野では不可視となる部分の画像をも能動的に撮像する。
- (2) 画像から得られる情報を利用して、次の撮像位置を動的に決定し、できるだけ少ない枚数の画像データから対象物の全体形状を再構成する。
- (3) ステレオ画像対における対応頂点の探索処理は、2次元的トポロジー（面および稜の見かけの連結構造）の対応関係に基づいた処理であり、凹多面体に現れ得る隠れ面を含む場合も扱える。

本章では、3次元位置計算法の定式化、ターンテーブルの回転角の設定方法、画像内の頂点等を求めるための画像処理、対応点の探索処理および全体形状の再構成の手順、計算機による自動認識を指向した基礎的実験の結果について述べる。

## 4.2 特徴点の位置計算の定式化

本節では2枚のステレオ画像の対（以下、ステレオ画像対と呼ぶ）における頂点の2次元画像座標値と、対応点関係が知られている場合に、頂点の3次元座標を未知数とする連立方程式を導く。定式化は第3.2節におけるものと基本的には同一であるが、ターンテーブルの回転という新しい座標変換が含まれている。

図4.1では本節で用いる座標系の設定を示す。ターンテーブル上にはその回転軸上に原点をもつターンテーブル座標系（Turn-table Coordinate System: TCS）を設定する。これは対象物および撮像系の位置決めを行うための参照座標系である。撮像装置であるITVカメラを、レンズ中心の前方に仮想画像面をもつピンホールカメラによって近似し、それに対して画面の中心を原点とし光軸をZ軸とするカメラ座標系（Camera Coordinate System: CCS）と、仮想画像面上に同じく画面の中心を原点とする2次元画像座標系（Image Coordinate System: ICS）を設定する。更に、カメラの視軸（光軸）がターンテーブル座標系の原点（回転中心）を通るように設置する。TCSからCCSへの座標変換行列を $\hat{T}_{ct}$ 、TCSでのZ軸回りの回転行列を $\hat{T}_r$ 、ICSでの頂点の2次元位置ベクトルを $p=[p, q]^t$ と表す。行列の大きさは特に指定しないときには $4 \times 4$ とし、記号 $\hat{\cdot}$ はベクトル、行列の同次座標表現を示すものとする。

図4.2にステレオ画像対の撮像の様子を示す。ターンテーブルを角度 $\phi$ だけ回転し、対象物の頂点 $x_1$ を $x_2$ に移動する。

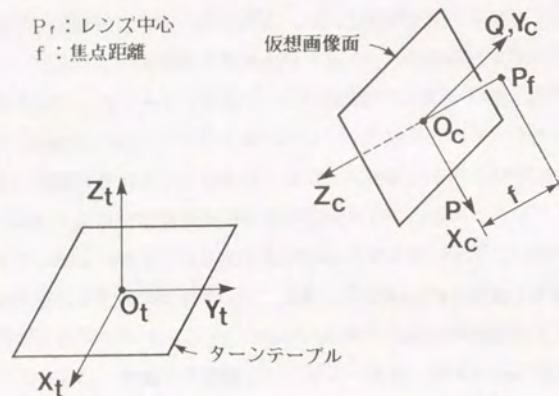
$$\hat{x}_2 = \hat{T}_r \hat{x}_1 \quad (20)$$

頂点 $\hat{x}_1$ および $\hat{x}_2$ の画面上の投影像をそれぞれ $p_1, p_2$ とし、レンズ中心とそれらを結ぶ直線に沿うベクトルをそれぞれ $\hat{u}_1, \hat{u}_2$ とすると、次式が成立する。

$$d_1 \hat{u}_1 + c = \hat{T}_{ct} \hat{x}_1 \quad (21.a)$$

$$d_2 \hat{u}_2 + c = \hat{T}_{ct} \hat{x}_2 = \hat{T}_{ct} \hat{T}_r \hat{x}_1 \quad (21.b)$$

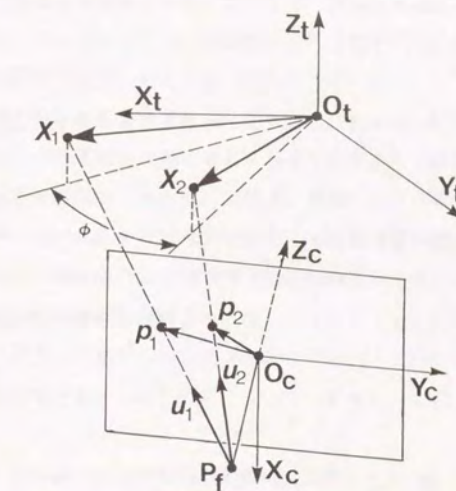




$P_f$ : レンズ中心  
 $f$ : 焦点距離

$O_t-X_t Y_t Z_t$ : ターンテーブル座標系 (TCS)  
 $O_c-X_c Y_c Z_c$ : カメラ座標系 (CCS)  
 $O_c-PQ$ : 画像座標系 (ICS)

図4.1 座標系の設定



$x_1, x_2$ : 対象点,  $p_1, p_2$ : 投影点 (画像点)  
 $u_1, u_2$ : 視線ベクトル,  $\phi$ : 回転角度

図4.2 ステレオ画像対の撮像



ここで,

$$d_1 \triangleq [0, 0, 1/f, 0] \hat{T}_{ct} \hat{x}_1 + 1$$

$$d_2 \triangleq [0, 0, 1/f, 0] \hat{T}_{ct} \hat{x}_2 + 1$$

$$\hat{u}_1 \triangleq [p_1^t, f, 1]^t$$

$$c \triangleq [0, 0, -f, 0]^t$$

以上の式を用いて頂点  $\hat{x}_1$  の3次元座標を未知数としてまとめると次の連立方程式を得る.

$$(A_1 - A_2) \hat{x}_1 = 0$$

ここで,

$$A_1 \triangleq \left( \frac{I_{2 \times 4} \hat{T}_{ct}}{I_{2 \times 4} \hat{T}_{ct} \hat{T}_r} \right) \quad (22)$$

$$A_2 \triangleq \left( \frac{p_1 \{ r (\hat{T}_{ct})_3 + a \}}{p_2 \{ r (\hat{T}_{ct})_3 + a \} \hat{T}_r} \right) \quad (23)$$

第3.2節と同様に, 定数項を分離して移項すると,

$$A x_1 = -b$$

ここで,

$$A_1 - A_2 \triangleq [A \mid b]$$

最小2乗法を用いて, 次式の解を得る.

$$x_1 = -(A^t A)^{-1} A^t b$$

#### 4.3 回転角度の設定

対象物の全体形状を捉えるためには, 対象物に対するカメラの相対的位置および姿勢を変えて撮像する必要がある. ここでは, 視野を相対的に変化させるために, 対象物をターンテーブルの中央付近にのせて, それを回転する方法を採用する. 本節では, その回転角度の設定方法について述べる.

まず, 視野と回転角度の関係について考える. いま, 図4.3に示すように, 2組のステレオ画像対を得た状態を考える. すなわち, 1つの頂点の回転軌跡で示すと,  $x_1$ と $x_2$ とで画像対1を構成し, 同様に $x_3$ と $x_4$ とでもう1つの画像対2を構成する. このとき, 各画像は対象物の異なる4つの側面を捉えているが, 1つのステレオ画像対から対象物の部分的3次元構造が作られるので, その画像対を構成する2枚の画像は同じ視野を表現すると考える. したがって以後, 視野ということは対象物のまとまった部分的な3次元構造と対応させて用いることにする. 次の2種類の回転角度を定義する.

(1)  $\alpha$ : 隣り合う2組の画像対間の回転角度

(2)  $\phi$ : ステレオ画像対を構成する2枚の画像間の回転角度

角度 $\alpha$ が上記の意味における視野を特徴づけている. 本手法では, 画像を2枚1組の画像対として扱い, 全体形状を再構成するために複数組の画像対によって対象物全体を撮像する. その際に, それぞれに作られた部分的3次元構造を1つに統合していくためには, 連続する画像対間に共通の頂点が少なくとも1つ存在することが必要である. したがって, 角度 $\alpha$ は次のように設定される. 図4.3に示すように, 1組めの画像対における1枚めの画像の最も左(画像座標系のQ軸の負方向)にある頂点( $x_1$ )が, 2枚目の画像対における2枚目の画像の最も右( $x_4$ )に位置して見えるような限界の回転角度をととする. まず, カメラを下方に傾けずに, ターンテーブル座標系の $X_t Y_t$ 平面上に視軸を置いた場合について考える. 限界の回転角度を求めるために, 頂点 $x_1$ を含み, 画像面上で縮退する面 $s$ を想定する. テーブルを回転し,  $x_1$ が $x_4$ の位置に来るときの回転角度 $\alpha$ は次式で求まる.

$$\alpha = \pi - 2\phi \quad (24)$$



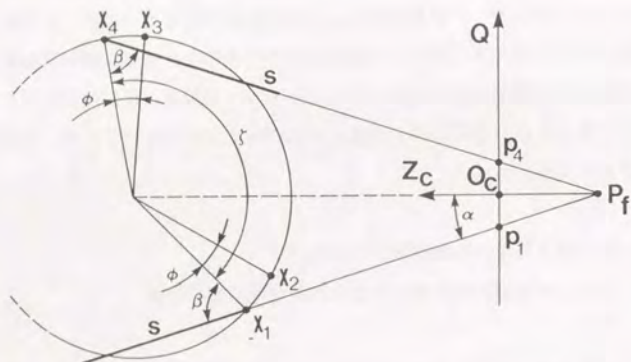


図4.3 ターンテーブルの回転角度の算出

ここで、 $\alpha$ は面 $S$ と $Z_c X_c$ 平面とのなす角度であり、頂点 $x_1$ の投影点 $p_1$ の座標のみから計算できる。カメラの画像面がターンテーブルに対して角度 $\psi$ だけ下方に傾いている場合には、角度 $\zeta$ は

$$\zeta = \pi - 2 \tan^{-1} (\tan \alpha / \cos \psi) \quad (25)$$

となる。

角度 $\psi$ はステレオ撮像における適当な大きさの視差を与えるための輻輳角度である。この角度の大きさは普通、対応点検出の困難さと位置誤差の大きさとのバランスにおいて決められる。すなわち、一般的に輻輳角度が小さければ、対応付けは容易になり反対に位置誤差は大きくなる。一方、輻輳角度が大きければ、対応付けは難しくなり反対に位置誤差は小さくなる。

#### 4.4 画像処理およびグラフの生成

本節では濃淡画像を処理し、頂点、稜線、面という要素を検出し、それらの連結関係を解析し、画像対ごとの部分的な幾何構造を表すためのグラフにまとめるまでの一連の処理について説明する。図4.4に具体的な処理の流れを示す。

これらの処理は対応点問題、頂点の3次元位置計算、さらに全体形状への統合のための基礎的なデータ構造をつくる段階である。したがって、3次元物体の視覚的認識技術の実際の問題への適用を考えると、まず、解決しなければならないものである。本手法では、各段階の処理は全体のなかで明示的ではないが階層的に結合されており、扱うデータ構造も各階層ごとに異なり、よりまとめられた形になっていく。一般的に、初期のデータ処理による誤りをより高い階層において発見し、再度修正してやり直すことは、バックトラックの操作を必要とし、そのためには各階層における過去の途中結果を必要に応じて保持しておくようなデータ構造がなければならない。本手法の場合、この途中結果には画像処理された画像データそのものも含まれるために、かなりの困難がある。したがって、本手法ではバックトラックの操作は考えず、それに対応できるようなデータ構造も特に考えていないので、当



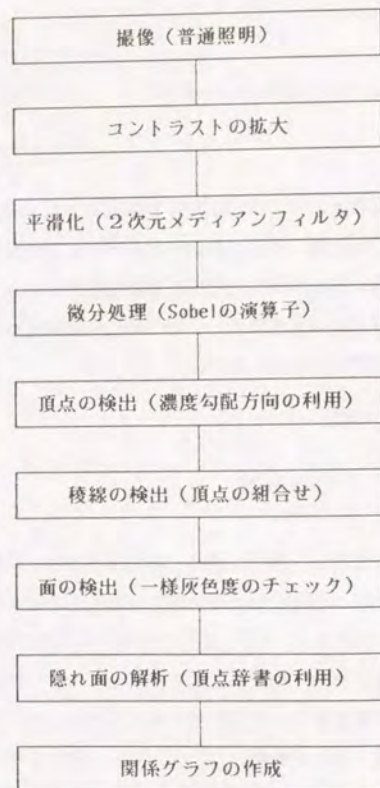


図4.4 画像処理の手続き

然のことではあるが各段階において、確実に処理を実行しなければならない。この要求は1回の解の候補が得られるまでの計算時間の増加につながるが、最終的判断（得られた結果をよしとして、処理を終了するか、条件を変えて再実行するかの判断）を再実行に対してよりルーズに設定し、バックトラックを行い解の収束を計るような処理に比べて計算時間において一概に不利ではないと思われる。

受動的ステレオ法においては、照明条件が普通のいわゆる生活照明に近いことが望まれる。ここでは、原濃淡画像を撮像するとき、とくに対象物の影が写らないようにするために、①雰囲気照明、②同軸照明（照明方向とカメラの光軸が一致する）とを併用している。

- (1) 撮像：普通照明下で画像を撮る。照明は白熱電球であり、影が見えないように、カメラの直下に設置している。
- (2) コントラストの拡大：濃淡原画像の灰色度を最小値0、最大値255レベルに拡大する。
- (3) 平滑化処理：雑音除去のために空間メディアンフィルタ（3×3オペランド）をかける。
- (4) 微分処理：対象物の稜線の近傍画素を抽出するために Sobel の空間勾配オペレータを適用する。
- (5) 頂点の検出：微分処理された画像を2値化し、稜線を含めたその近傍を白画素によって表す。白画素のみに再度 Sobel の勾配オペレータを施すことにより、灰色度の2次勾配を求める。対象物が多面体であるので頂点は直線稜線の交わりと考えられ、頂点近傍では各稜線の両側の灰色度により定義される勾配の方向が顕著に変化することが予想される。したがって、その方向成分の極大値をクラスタリングして頂点の近傍画素を求め、その重心を頂点として検出する。
- (6) 稜線の検出：得られた頂点のすべての組合せに対して仮想の直線線分を作り、それぞれの線分上の画素と2値化画像の白画素との一致を調べ、一致数が多いければ稜線として連結する。
- (7) 面の検出：幾つかの連結する稜線で閉じられた一様な灰色度の領域を面として検出する。



- (8) 隠れ面の検出：隠れ面を検出し面および稜線の正しい隣接構造を作る。
- (9) 2次元関係グラフの作成：上記の一連の処理によって得られた頂点、稜線および面の隣接および連結構造を記述するためのグラフ<sup>54)</sup>をつくる（これを2次元関係グラフと呼ぶ）。

(1) ~ (4) はよく用いられる基本的処理である。(5) ~ (7) は2次元線画の解釈のための処理であり、そこまで得られる構造は、たとえば、図4.5に示すように、面  $s_3$  は面  $s_1$ ,  $s_4$ ,  $s_7$  によって部分的に隠されており、稜線  $e_9$ ,  $e_{12}$ ,  $e_{16}$  は面  $s_2$  を実際に構成する稜線ではない。そこで処理 (8) においてこのような隠れ面の構造を検出し、面に隣接しない稜線をその面の記述から削除する。隠れ面の構造を解析するために、稜線に関する次の3つの仮定を設ける。

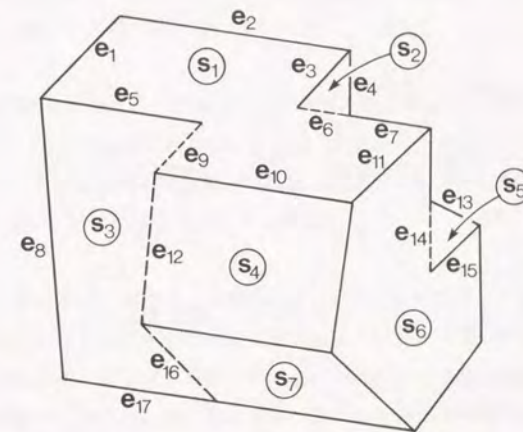
【仮定1】対象物は1つであり、背景によって取り囲まれている。従って、最外側の輪郭をつくる稜線は隣接する面に連結する。

【仮定2】隣接する面の連結は必ず1本の稜線による。

【仮定3】T型に交わる3本の稜線のうち、共線上の2稜線は面を連結しない。

【仮定1】はターンテーブル上には対象物が1つしかなく、背景と対象物とを分離する稜線は必ず対象物側の面に隣接することを示している。したがって、外輪郭稜線は、それを含む面に必ず隣接する（稜線が面を構成する）。【仮定2】は画像処理の結果得られる稜線構造に関して次の処理を可能にする。最外側稜線でない内部の稜線上に2つの稜線しか交わらない頂点がある場合には、その頂点が誤りであるか、あるいはその頂点で連結するはずの稜線が検出されなかったかのいずれかである。したがって、ここでは、2稜線がほぼ共線上にあれば頂点を削除し稜線を併合する。【仮定3】は隠れ面あるいは面の縮退した場合に特徴的に現れるいわゆる「T型頂点」に関するものである。T型頂点では3つの稜線が交わるが、そのうち、共線上にある2稜線は本来1つの稜線であり、他の1稜線を遮蔽していると考ええる。

次に、図4.6のような隠れ面によく現れる稜線構造（頂点周り）を選び、次の5つの処理を再帰的に適用する。ここでは、稜線のうち、その両側の面を連結するもの（実際に両側の面の交線であるもの）を連結稜（Connecting Edge, CE）、連結しないものを非連結稜（Non Connecting Edge, NCE）と呼ぶことにする。ここで、

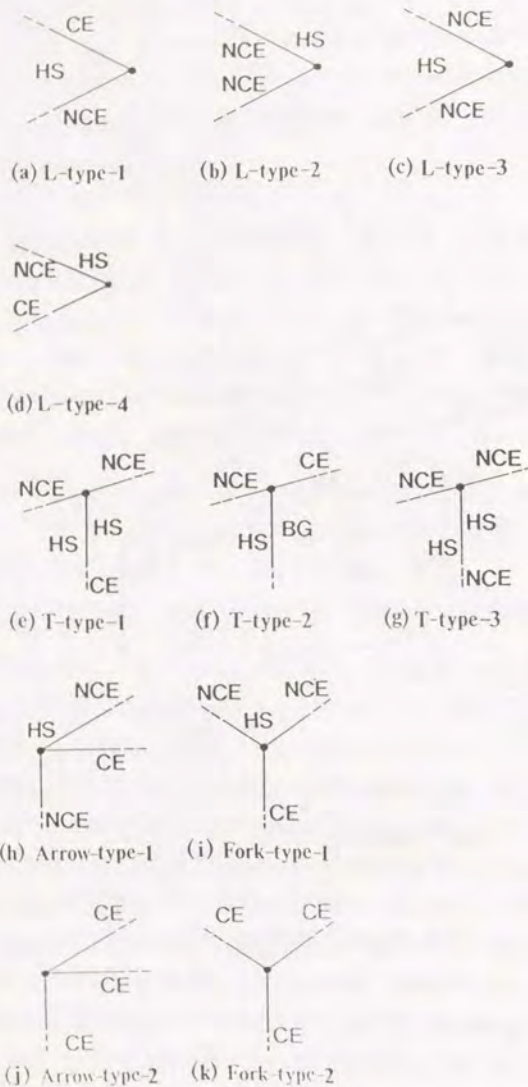


$s_i$  : 面番号

$e_j$  : 稜線番号

図4.5 隠れ面の構造





注) (j) および (k) は隠れ面を含まない。

図4.6 隠れ面を解析するための頂点辞書

外輪部稜線は、2重の性質をもつことに注意する。すなわち、内部の面に必ず隣接するという連結稜としての性質があり、一方、背景を常に隠すという非連結稜としての性質である。したがって、外輪部稜線はNCEとラベル付けされるが、それを含む内部の面の記述から削除されることはない。

- (a) L, T型頂点を含む面、および2本以上の稜線で連結している面を検出する。
- (b) 外輪部稜線をNCE, ARROW型頂点の中心稜線をCEとする。
- (c) FORK型頂点の3稜線のうち、任意の2本がCEであれば、他の稜線もCEとする。
- (d) T型頂点で交わる3稜線のうち、共線上の2本をNCEとする。
- (e) 図4.6の型に照合するものを選び、稜線にラベル付けする。
- (f) (a) で選ばれた面に隣接する稜線のうち、任意のNCEに挟まれ、かつ属性の未定である稜線をNCEとする。

これら6つの処理は、仮定1, 2および3とともに、いわゆる3面頂点多面体(いかなる頂点でも3つの面が交わるような多面体)に関する頂点辞書<sup>2,3)</sup>におけるラベル付けの規則を一部改良して用いている。 Huffmanの頂点辞書は連結性の他に凸稜線、凹稜線も含んでいる。ここでは、抽出した線画像における稜線の両側(最外側の稜線であれば、内側のみ)にある面とその稜線との関係を調べており、連結か非連結かが問題であるので、それに関する規則に書き直し、さらに隠れ面に関するT型頂点を加えて使用している。さらに、L-type-1およびL-type-4においては、CEとNCEを交換した頂点ボタンが必要であり、実際のラベル付け処理では辞書ボタンをそのように変換して用いている。

ここでは、図4.6においても、3面頂点以外は考慮していない。実際の3次元立体の形状を多面体で近似する場合にも、ほとんどの頂点は3面頂点によって有効に近似できると考えている。しかしながら、4つ以上の面の交差であるような頂点も存在する。上に述べた処理はいずれも3面頂点を選んで稜線のラベル付けを行うものであるが、4面頂点などが現れてもそれに隣接する3面頂点に関するラベル付けの結果から合理的に稜線のラベル付けが行われる場合には問題はない。実際の実験でも4面頂点を含む場合を行っているが稜線のラベル付けは完了している。



さらに、稜線のラベル付けが終了すると、隠れ面の決定を行わなければならない。  
隠れ面の判定のための条件は、図4.6にあるHS（隠れ面）を定めるものであるが、  
まとめると次のようになる。

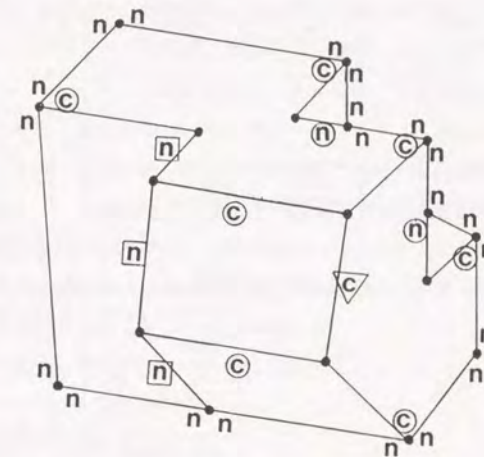
- (1) CEとNCEとがL型頂点を構成するとき、2稜線によってつくられた鋭角側にある面が隠れ面である。
- (2) T型頂点の共線上にない稜線を含む面は隠れ面である。
- (3) FORK型頂点の2本のNCEを含む面は隠れ面である。

隠れ面の判定後、面および稜線の正しい隣接構造をつくる。すなわち、隠れ面と判定された面に隣接していたNCEをその面を構成する稜線群から取り除き記述を書換え一貫性を保持する。図4.7にこのアルゴリズムの処理結果を示す。

ここでは面および稜線の縮退（面がちょうど線に見えたり、稜線が点に見えたりする場合）や、頂点どうし、或は頂点と稜線との重なりなどがある場合は考慮していない。また、図4.6に示す稜線構造だけでは表せないより複雑な隠れ面もあるがここでは扱わない。隠れ面構造を扱うとき、影の処理は実際上問題であるが、ここでは、影が顕著に現れないように雰囲気照明を用いるなどの工夫をしている。複雑な面構造および影の問題などは、本論文で提示するような線画を中間表現とする認識手法において解決しなければならない課題ではあるが、その問題は画像処理における分解能や多面体近似の問題と不可分であると考えられ、実際の対象物を多面体としてどのくらい詳細に記述できるのか、どのくらい正確に記述すればよいのかといった観点が必要になる。そうすれば、たとえば、多面体近似をある程度の分解能で行い、さらに詳細な測定は他の補完的な手法を用いるといったシステム構成が可能となる。この場合にも効率化を計るためには機能分散が必要になる。

処理(8)で得られる面および稜線の隣接構造および連結構造に基づき、処理(9)で各画像ごとにつくられる2次元関係グラフ（したがって、1組の画像対で2個のグラフがつくられる）は、その画像ごとに可視である対象物の一部分の2次元のトポロジー（隠れ面による非連結を含んでいるので、完全な2次元トポロジーではない）を表現する。これは図4.8のようなリスト構造によって表現される。

以上(1)～(9)の一連の処理は各原画像ごとに行う。



注) n: 処理(b)による. c: 処理(b)による. n: 処理(d)による.  
n: 処理(e)による. c: 処理(e)による.

図4.7 頂点ラベル付けアルゴリズムの結果



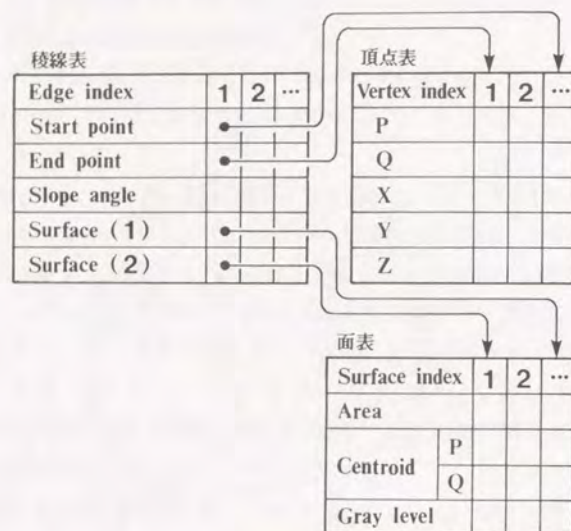


図4.8 関係グラフのリスト表現

#### 4.5 対応点探索のための処理

第4.3節で述べた方法によって頂点の位置計算を行うためには、画像対をなす2枚の画像内の頂点のうち、どの頂点が同一頂点を示すかという対応点問題を解く必要がある。本手法では、対象物を回転しつつ撮像した画像対を用いるために、2枚の画像において対象物の見え方が変化する場合を考慮しなければならない。後の章で検討するが、デジタル画像を用いるステレオ撮像に付随する画素量子化誤差を減少させるためには、幅角をある程度大きくしなくてはならず、両画像にわたるトポロジーの変化はやはり避けられないと考えられる。したがって、対応点検出においても、多面体の見かけのトポロジー（稜線、面等の連結関係）が変化する場合にも対処可能な、面や稜線の連結の仕方や見かけの形状に左右されにくい対応付けの方法が必要である。本節では、幾何学的に容易に決定可能な頂点の基本的な対応関係に基づいて稜線および面の対応関係を調べ、その結果を利用して未知の頂点の対応をさらに決定していくという方法を試みる。

そのために次の3つの処理によって対応点検出を行う。

##### (1) 拘束直線による頂点对応の探索

まず、明かに判定可能な頂点の対応を求めておく。

カメラの位置・姿勢およびターンテーブルの回転角度が既知であるので、対応点間には次に説明する幾何学的拘束がある。図4.9に示すように、頂点 $x_1$ とレンズ中心 $Q_0$ とを結ぶ視線 $L_1$ を、もし仮に頂点 $x_1$ に固定して回転角度 $\phi$ だけ回転すると視線 $L_2'$ となる。 $L_2'$ をレンズ中心に関して透視投影すると、画面上の投影像として拘束直線 $L_c$ が得られる。頂点 $x_2$ は視線 $L_2'$ 上についているので、 $x_2$ の投影点 $p_2$ がこの拘束直線上に必ずあることがわかる。拘束直線の方程式はカメラパラメータ、回転角度 $\phi$ およびどちらかの画像上の頂点座標から計算できる。実際の計算機処理では、この拘束直線に沿ってある幅を設定して探索することにより対応点の候補を検出する。図4.10に拘束直線の様子を示す。このとき、T型頂点は見かけの頂点とみなして対応付けを行わない。この幾何学的拘束は頂点が対応するための必要条件であるから、1つの頂点に2つ以上の対応点候補が現れることもしばしばあ



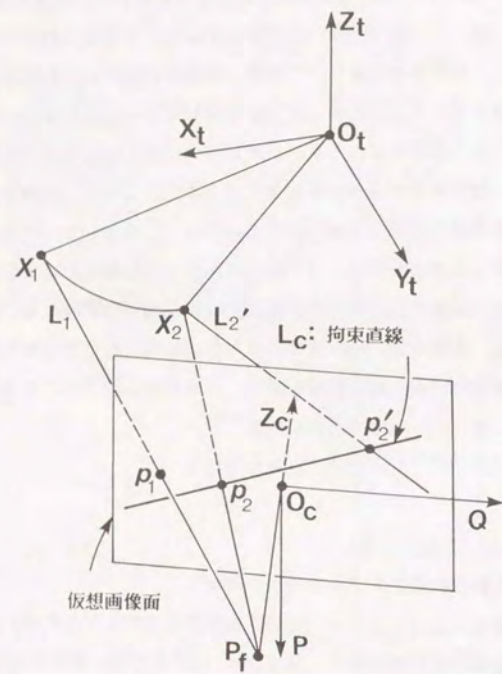
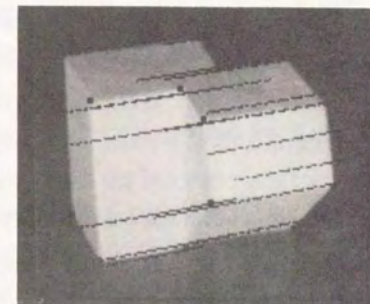


図4.9 拘束直線の幾何学



(a) 画像1-②



(b) 画像2-②

図4.10 拘束直線



り得る。ここでは、1対1の対応が得られる頂点のみを対応頂点として判定し対応表に登録する。

#### (2) 対応面を用いた対応付け

拘束直線による探索によって得られる対応関係のうち、一意に対応していると判定できる頂点对を選ぶことにより、それぞれの画像においてそれらを含む稜線および面の対応を判定できる場合がある。そしてそれらの対応から逆に未定の頂点の対応を求めることができる。たとえば、面を対応付けるためには、両面に隣接している3点以上の頂点の対応があればよい。各画像における頂点と面の隣接関係は2次元関係グラフとして整理されているので、前節で求められた対応する頂点对のデータを加えて3組以上の頂点对をもつ面の組合せを調べればよい。このことによって、少なくとも3組の頂点对の対応が保存されているような変化であれば、今まで不可視であった稜線や面が現れたり、面が隠れたりするトポロジーの見かけの変化にも対処できると考えられる。

面どうしの対応が得られた場合、その面に隣接する未だ対応のついていない頂点のうち、拘束直線による対応付けでは複数候補が現れ一意に判定できなかった頂点を対応面の組合せから一意に絞り込めるときには、判定された頂点对として対応表に登録する。

#### (3) 仮の3次元位置による対応付け

上の2つの対応点検出処理を経てもなお対応の付かない頂点も存在する。それらの未知の頂点は、

- ①何らかの原因で拘束直線による対応点候補の選択にはかからなかったもの
- ②同一の対応面でのなかに複数の対応点候補をもつもの

の2種類と考えられる。

上記の2つの対応点検出処理により対応の判定した頂点については、3次元位置が算出される。1つの面上の3点以上の頂点の3次元位置がわかる場合には、その面上にある対応関係の未知である頂点の3次元位置をその面と視線の交点として計

算できることとそのための具体的方法も、第3.4節からわかっている。各画像内の対応の付いていない頂点のうち、決定可能な面に隣接する頂点の3次元位置を求め、お互いに近傍にある頂点を対応点对として登録する。

上記の処理のうち(2)および(3)は、新しく判定される対応点对が現れなくなるまで反復して適用される。以上の処理において、頂点の対応関係(対応点表)のみならず稜線および面の対応関係、面の法線なども算出されている。図4.1.1に実際の対応付けの結果を示す。そこには、対応点とともに、対応稜線および対応面も示されている。これら3つの対応点検出処理を他の対応点問題の手法と比較すると、次の2つの特徴をもつ。

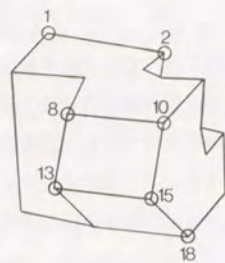
- ①確実性の向上： 多面体に固有の構造を調べることにより、これまで不可能であったような特徴点をも対応付けできる。
- ②付加的な情報の獲得： 画像内の単なる特徴点としての対応だけでなく、これらの処理を経たデータは連結構造および稜線や面という高次の対応関係をもっている。

## 4.6 全体形状の再構成

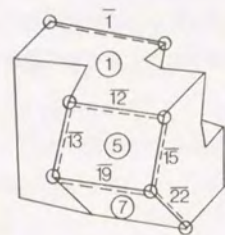
2組の画像対を処理して得られた2次元関係グラフ、対応点表、稜線および面の対応関係から、それらを統合して対象物の全体形状の再構成を試みる。基本的な進め方について述べておく。まず、2組の画像対(4枚の画像)の処理を終了した段階で、再構成を試みる。これは本手法の目的の1つであるが、なるべく少ない枚数の画像データを扱うためである。すなわち、まず最小限のデータで再構成を行い、そこに不足や不合理があればそれを解消するように次の処理に進む、という方法である。

再構成処理の基本的な方法は、頂点、稜線および面の対応関係に基づいた各階層におけるグラフの併合を行うものである。図4.1.2は記述グラフの併合の様子を実例を用いて示している。2次元関係グラフを各画像対内の対応関係に基づいて併合したグラフを「部分構造記述グラフ」と呼ぶ。さらに、それらを共通頂点、共通

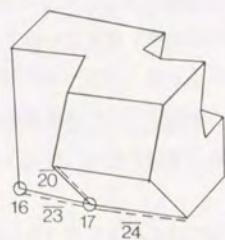




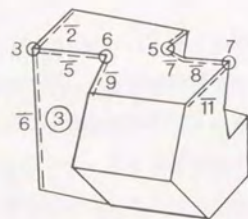
(a) 拘束直線による対応付け  
(画像1のみを示す)



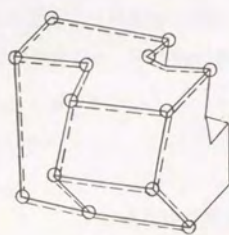
(b) 対応稜線および対応面



(c) 対応面による対応付け

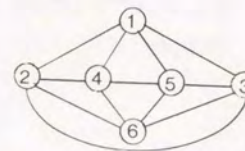
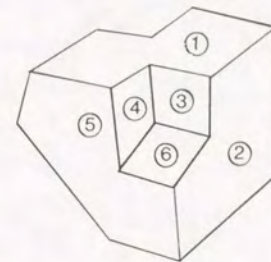
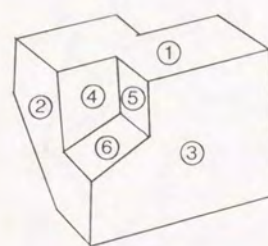


(d) 仮の3次元位置による対応付け

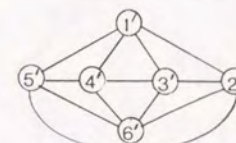


(e) 最終結果 (対応頂点: 13,  
対応稜線: 15, 対応面: 4)

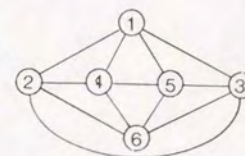
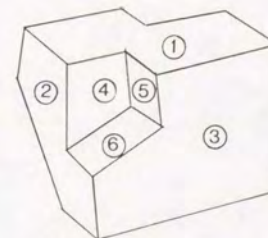
図4.11 対応点の判定手順



関係グラフ1



関係グラフ2



部分構造記述グラフ1  
(再番号付け済み)

図4.12 (a) 部分構造記述グラフ1 (画像対1)



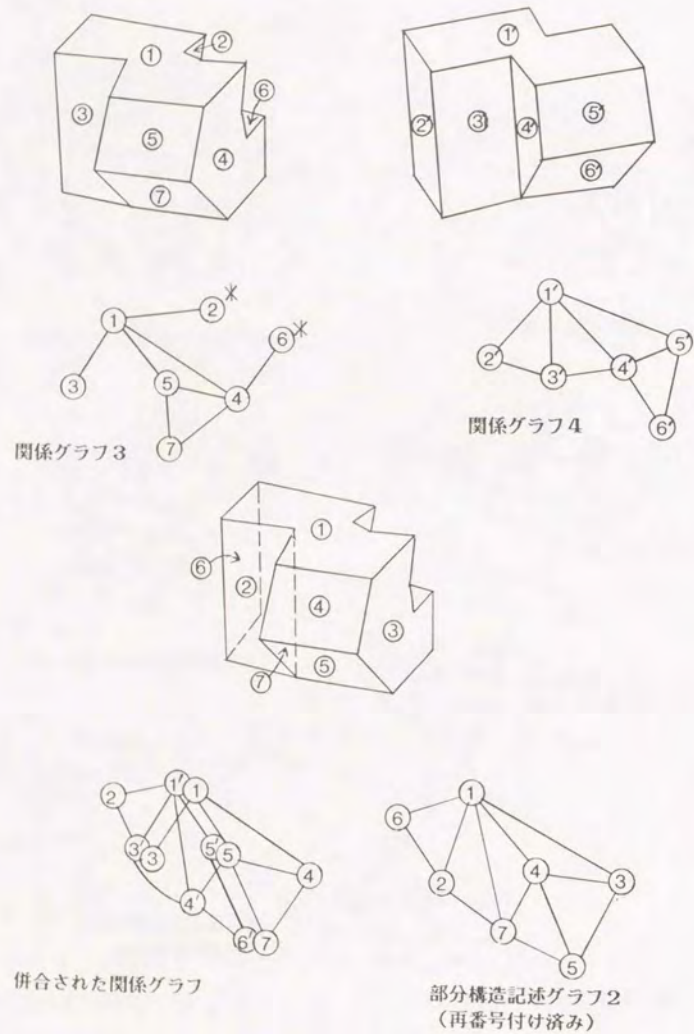


図4.12 (b) 部分構造記述グラフ2 (画像対2)

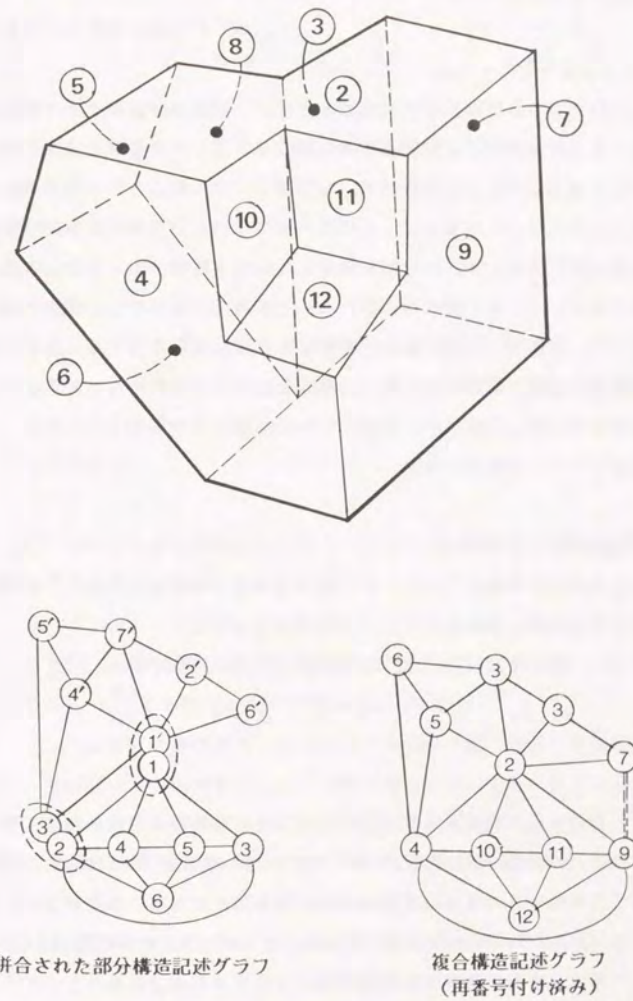


図4.12 (c) 複合構造記述グラフ



稜線および共通面の関係に基づいて併合したものを「複合構造記述グラフ」と呼ぶ。これが対象物の全体形状を表現すべき記述グラフである。

#### (1) 部分構造記述グラフの作成

各画像について得られた2次元関係グラフを頂点、稜線および面それぞれの対応関係に基づいて、各画像対ごとに併合する。さらに、ステレオ撮像による位置計算により得られた頂点の3次元位置座標値、面法線方向余弦などの幾何情報を付加して部分構造記述グラフを作成する。この部分構造記述グラフは、2個の2次元関係グラフのそれぞれにおいて、面法線の決定している（隣接する3つ以上の頂点の3次元位置が決定している）面を併合したものである。まず面の対応関係に基づき併合し、その後、対応関係のない面も連結関係に基づいて組み入れる。したがって、このグラフは各画像対（視野）から得られる部分的形状を表現するものである。図4.1.2 (a)において、画像1内の面②および⑥は面法線が算出されなかったため、部分構造記述グラフには含まれない。

#### (2) 複合構造記述グラフの作成

対象物の全体形状を各画像対において可視である部分的形状の併合として再構成する。そのために各部分構造記述グラフの共通部分を判定し、それに基づいてグラフの併合を行い、複合構造記述グラフを作成する。その実例を図4.1.2 (c)に示す。

まず、共通頂点、稜線、面の判定のためにそれぞれ次の条件がある。

【条件1】 共通頂点の判定：3次元座標の近いものを共通の頂点と判定する。第4.3節で述べた角度 $\alpha$ の設定方法からわかるように、連続した画像対は少なくとも1個の共通の頂点をもつことになる。またさらに、図4.2に示した撮像環境では対象物の最上面に含まれる頂点および稜線は常に見えているので、各画像対にわたって共通である。

【条件2】 共通稜線の判定：両端点がそれぞれ共通頂点である稜線を共通稜線と判定する。

【条件3】 共通面の判定：共通稜線を共有するという条件に加えて、2つの面の

面法線（面の傾き）の類似により判定する。

具体的な手順は次のようである。

- ① 各々の部分構造記述グラフにわたる頂点のすべての組合せについて、その3次元的な距離を算出し、ある閾値以内の距離にある頂点どうしを共通頂点と判定する。
- ② 共通頂点の判定を行う。
- ③ 2本以上の共通稜線をもつ面どうしを共通面と判定する。1本の共通稜線をもつ面どうしの対応も調べる。
- ④ 共通面により部分構造記述グラフを併合する。
- ⑤ ③項で調べた共通稜線を1本もつ面は、複合構造記述グラフにおいて、連結する面であるので、連結する。（図4.1.2 (c)の面⑦と⑨との連結はこの処理による。）

さらに、併合された複合構造記述グラフが矛盾なく立体を表現しているかどうかの検定を次の2項目について行う。

【条件1】 すべての面を構成する稜線が閉じている。

【条件2】 すべての稜線が2つの面を連結している。

ただし、底面はいずれの画像対でも不可視であるので、底面に含まれる稜線は除いている。底面を構成する稜線の判定は両端点のZ座標値によって行う。すなわち、頂点のZ座標値がある閾値より小さければその頂点が底面に含まれると考える。

上の2つの条件を満足すれば再構成処理を終える。しかしながら、一般的には2組の部分構造記述グラフからいつも全体形状が完全に再構成されるわけではない。そのときは上の2つの条件に合致しない稜線や面が現れるので、さらに、複合構造記述グラフに対して、以下順次新たな画像対に対応する部分構造記述グラフを取り込みながら以上の処理を繰り返す。



#### 4.7 実験

実験に使用した装置の仕様および撮像パラメタを表4.2示す。ターンテーブルは手動で回転した。図4.13に示す2組の画像対(a), (b)および(c), (d)に対して実験を行った。回転角度 $\phi$ は画像対1で $30^\circ$ 、画像対2で $40^\circ$ とした。回転角度 $\phi$ の計算値は $171.8^\circ$ となり、 $170^\circ$ に設定した。図4.13の原画像に対する画像処理の例を図4.14に示す。他の画像においても、可視の頂点、稜線および面がすべて検出された。図4.15にステレオ計算で求めた頂点の3次元座標の誤差を示す。平均位置誤差6.1mm(撮像距離の0.76%)以内で頂点を位置決めすることができた。図4.16は複合構造記述グラフをもとに再構成した対象物のワイヤーフレーム表示である。表4.3に計算時間の内訳を示す。使用計算機はACOS-600(0.69 MIPS)のTSSであり、CPU使用時間を測定した。前半の画像処理にほとんどの計算時間を費やしている。後半のグラフに対する処理では効率よく処理が進んでいる。

表4.2 実験の仕様

撮像装置	ITVカメラ サイズ: 1 インチ・ビジコン 撮像距離: 800 mm
フレームメモリ	画素数: 256 × 384 分解能: 25.6 pixel/mm グレイレベル: 256 レベル
対象物	サイズ: 90 × 90 × 80 mm 色: 白色 材質: 紙
ターンテーブル	サイズ: 560 × 560 × 270 mm 分解能: $0.5^\circ$

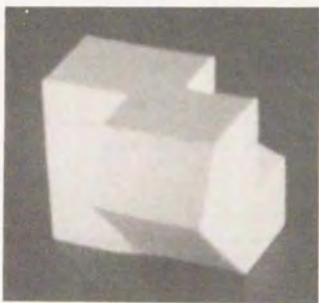




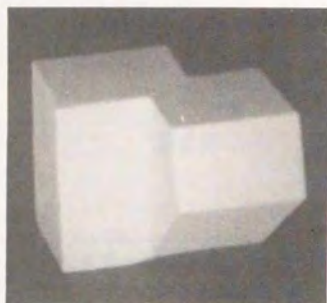
(a) 画像対 1-①



(b) 画像対 1-②



(c) 画像対 2-①



(d) 画像対 2-②

図 4.13 ステレオ画像対

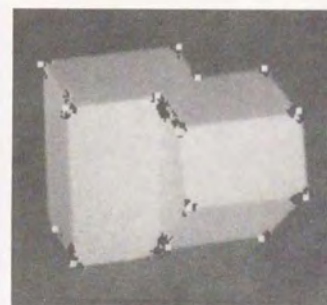
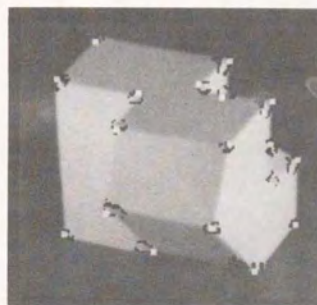
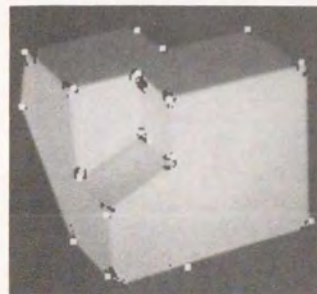


図 4.14 (a) 画像処理：頂点の検出



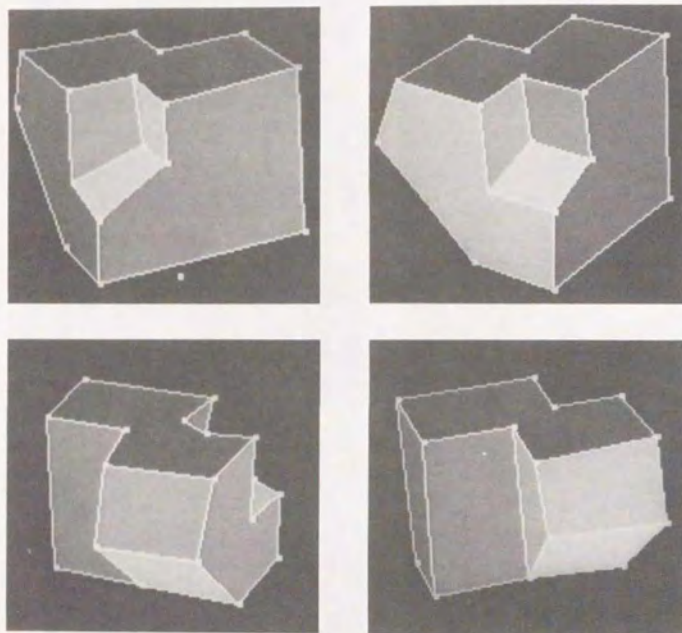


図4.14 (b) 画像処理：面の抽出

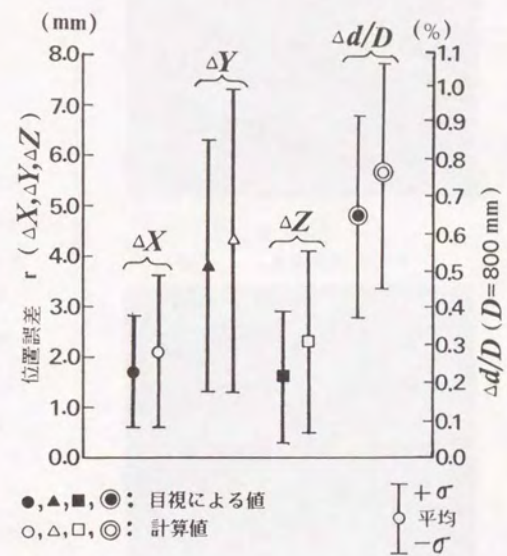
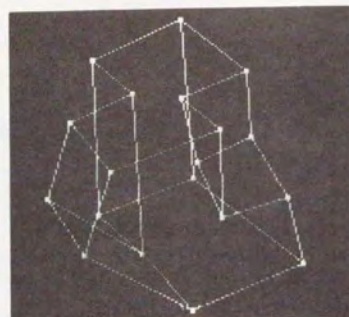
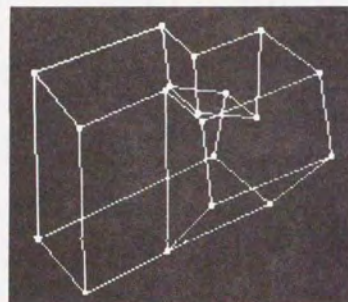


図4.15 頂点の位置誤差





(a) 回転角度：90°



(b) 回転角度：180°

図4.16 再構成した多面体

表4.3 計算時間（画像1枚について）

処理	CPU時間 (s)	備考
平滑化	30.231"	2次元メディアンフィルタ
微分処理	14.357"	Sobel演算子
頂点候補検出	13.511"	
クラスタ	13.498"	逐次的距離計算
リング		
頂点検出	12.814"	
稜線検出	7.828"	
面検出	1.589"	
頂点ラベル	0.034"	
付け		
隠れ面構造の	0.052"	
検出		
対応点探索	0.359"	画像対1の2枚分の和
	0.440"	画像対2の2枚分の和
再構成と	0.804"	画像対1および2の4
解析		枚分の和
計（平均）	94.3"	すべての計算時間を画像枚数 で平均した。



#### 4.8 まとめ

多面体対象物の全体形状を認識するための新しい方法について述べた。以下に結果をまとめる。

- (1) 3次元多面体の異なる側面を多視野の画像対として撮像し、部分的に再構成された構造を統合して対象物の全体形状を再構成する方法を示した。
- (2) ターンテーブルの回転という設定において、その回転角度を工夫することにより、なるべく少ない画像枚数から全体を系統的に再構成するための手法を示し、実験により有効性を確かめた。
- (3) 画像対内の対応検出処理において、頂点だけでなく稜線および面の対応をも考慮し、隠れ面や見かけの2次元的トポロジーの変化にも対処できる手法を示した。
- (4) 実験例では、11面から成る多面体について、その2組4枚の濃淡画像から、平均 6.1mm (撮像距離の 0.76%) 以内の位置誤差で頂点座標を求め、全体形状を再構成することができた。

#### 第5章 視覚認識を利用したロボットアームの制御

##### 5.1 概要

以前の2つの章で述べた位置計算法は、いずれもステレオ撮像法を拡張した多面体対象物の全体形状を捉えるためのアプローチである。本章では、そこで得た知見を基礎として、より実際の応用のレベルに近付くために、ロボットアームの把握操作を自律的に制御するために視覚情報を利用する方法<sup>72,84)</sup>について考察する。ここでは、主に多面体物体の把握を目的とした実験的なプロトタイプシステムの設計および試作を行いながら、視覚認識系から得た3次元空間に関する情報から、操作系の駆動情報を生成あるいは変換する場合の諸問題<sup>55,83)</sup>について検討する。とくに、次の3つの項目について、新しい方法を提示し、実験システムでの検討を行う。

- (1) 対象物の視覚的な自動探索手法
- (2) 2次元輪郭形状の凸包による記述および特徴抽出
- (3) 輪郭頂点座標の共分散行列の固有ベクトル組の姿勢判定への応用

また、処理の高速化についても検討する。そのために、識別力との兼ね合いを考慮しながら、なるべく単純な特徴量を採用することでシステムの高速性を得ることを検討する。

まず、プロトタイプシステムの仕様を次のように定める。

- (1) ロボットアームは6軸多関節(RPPRPR)型とする。
- (2) 視覚センサとしてTVカメラ(小型CCDカメラ)を用い、それをロボットアームの第5軸に固定する。したがって、視軸の方向は任意であるが、視軸周りの回転の自由度はない。
- (3) 対象物は多面体とするが、凹凸は問わない。また、ロボットハンド(2指、1自由度)で把握可能とする。
- (4) 対象物はロボットアームが設置されているテーブル面上にあるものとする。



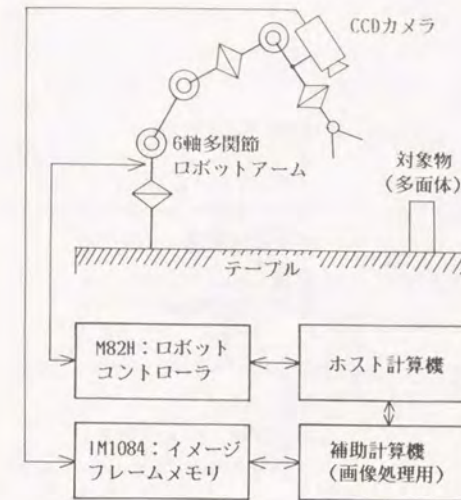
- (5) テーブル（色：黒、一様色）と対象物（一様色）とは画像上である程度のコントラストをもつ。
- (6) 照明は実験室内の通常のもの（雰囲気照明）のほか、カメラの上部に小型白熱電球（同軸照明）を固定する。撮像は普通照明下で行う。

(2) において、カメラの視軸周りの回転角度を自由に設定することはできないが、基礎的な予備実験からこのような設定で対象物を視覚的に捉えることが可能であることを確認した。(3) では、対象物の凹凸に依存しない特徴量によって認識を行うことを考える。(5) については、ある程度理想的な多面体のための画像処理の検討は第4章で行った。普通のテーブル上の通常の対象物が処理されるべきではあるが、この点については簡単化し画像処理の手間を減らすことを考えた。一般の3次元対象物の画像処理（情景解析と呼ばれる分野）は難しい問題を多く含んでおり興味ある分野であるが、ここでは扱わないこととした。(6) については、特殊な照明器具は用いないが、影の影響を少なくするために常にカメラの方向から照明するようにした。

実験システムは次の3つのサブシステムを統合して構成されている。

- (1) 視覚的捕捉サブシステム：ロボットアーム第5軸に固定されたカメラの画面内に対象物の像を捉える操作。テーブル上にあるべき対象物をカメラを移動させながら探索し、対象物を画面の中央付近に映すようにカメラの移動（センタリング）を行う。
- (2) 視覚的認識サブシステム：対象物の2次元の特徴量およびステレオ撮像による位置計算法によって求めた3次元の特徴量を用いて対象物の同定を行う。そのためのモデルデータは前もって準備する。
- (3) 把握制御サブシステム：視覚的に対象物を同定し、対象物の位置・姿勢を判断し把握するための処理である。ステレオ撮像による位置計算法に含まれる位置誤差に対処するために、ハンドの指先に接触センサを装着し、その情報を基に把握し直すことも考える。

図5.1に実験システム全体のハードウェア構成およびシステムの全景を示す。小



(a) ハードウェア構成



(b) システム全景

図5.1 自律的ロボットアームシステム



表5.1 実験システムの仕様

ロボットアーム	6軸多関節型 全長: 1125 mm 精度 (関節角度): 0.12 (位置): 34 $\mu$ m
CCDカメラ	2/3インチ 焦点距離: 16 mm (固定) 有効撮像面積: 8.8 mm $\times$ 6.6 mm
照明	カメラ上部: 50 W 白熱球 ロボット後方: 100 W 白熱球
フレームメモリ	画素数: 縦 256 $\times$ 横 384 分解能: 縦 26.8 $\mu$ m, 横 26.0 $\mu$ m
計算機	ホスト計算機: PC9801VX4  補助計算機: CPU:MC68000 (8 MHz)
撮像距離	500~770 mm

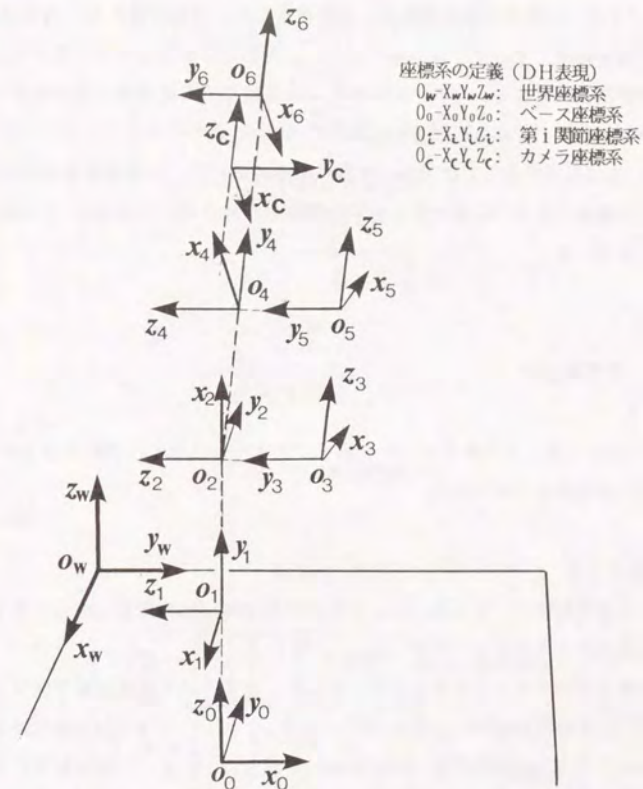


図5.2 ロボットアームシステムの座標系の設定



型CCDカメラは6軸多関節(RPPRPR)型ロボットアームの第5軸に固定されている。カメラからのビデオ信号はイメージフレームメモリ(256×384画素×8ビット×4ページ)でA/D変換され画像として保持される。この装置には、画像処理用の補助計算機が接続している。ロボットアームは専用のロボットコントローラによってサーボ制御されており、位置・姿勢などの状態はホスト計算機からのコマンドにより制御している。これらの装置の仕様を表5.1にまとめる。

図5.2に本章を通して使用する座標系の設定を示す。座標系はカメラ座標系以外は右手座標系であり、ロボットアームの各軸の座標系はいわゆるD-H表現に従って設定されている。

## 5.2 視覚的捕捉

図5.3に視覚的捕捉処理の流れを示す。大きく分けると、探索処理とセンタリング処理とから構成されている。

### (1) 探索処理

ここでは対象物は一様な黒いテーブル上に置かれ、その背景に対してある程度以上のコントラストをもつものとした。

対象物はテーブル上にあるはずであるから、まず予め全探索領域を図5.3に示すように10個の探索領域(設定画面)に分割しておく。分割された各視野を実現するロボットアームの関節角度(設定移動)を記憶しておく。次に各視野ごとに対象物の有無を判定する。その探索順序は任意ではあるが、とりあえず図5.4に示された領域番号順に進める。

処理時間を短縮するために、図5.5のようにサンプリング用格子点列を設け、各視野での対象物の有無の検出は、サンプル点の画素値がある閾値以上になるかどうかで判定する。ここでは、サンプル点は1画素で構成されており、閾値以上の画素があればその近傍領域(3×3画素の正方領域)における8画素のうち5個以上の画素が同様に閾値以上であるかどうかを調べ、ノイズによる誤動作を防いでいる。また各設定画面内においてテーブルの外側にはみ出すサンプル点は処理しない。

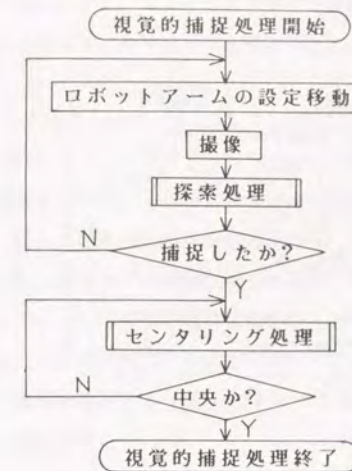


図5.3 視覚的捕捉の手順

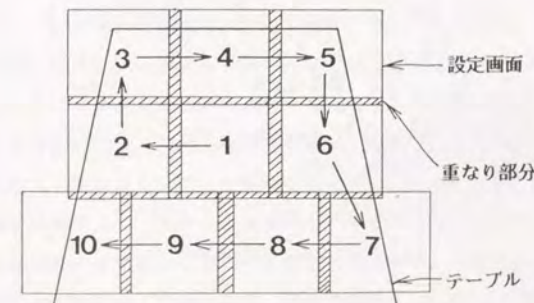


図5.4 設定移動による画面構成



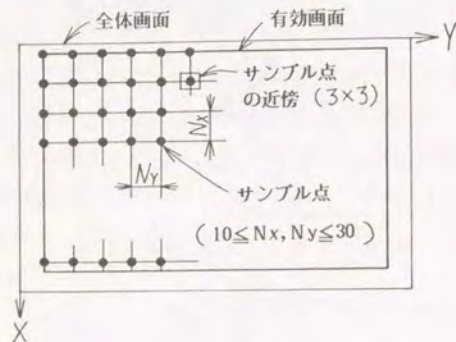


図5.5 サンプルング格子点列 (探索処理)

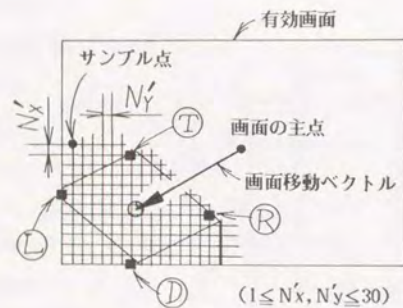


図5.6 画面移動ベクトルの算出

本プロトタイプシステムでは、カメラの撮像位置・姿勢を変化させる目的でロボットアームにカメラを装着し、全作業空間を見渡す広角カメラなどは使用していない。

## (2) センタリング処理

対象物の有無が検出されると、その対象物を画面の中央で撮像するためのセンタリング処理を行う。ここでも処理の高速化を考慮して、図5.6に示すように画面上にサンプル格子点列を設定する。ただし格子間隔 $N_x$ ,  $N_y$ は探索処理の場合より小さく設定する。各格子点の画素値を閾値とを比較しながら、対象物の占める領域内の画素を探索する。その領域内に含まれる最上(T), 最下(D), 最右(R), 最左(L)の4個の画素を検出する(図5.5)。それらの座標から次に撮像すべき画面の中心を次式で算出する。

$$V_x = \frac{T_x + D_x}{2} \quad (26.a)$$

$$V_y = \frac{R_y + L_y}{2} \quad (26.b)$$

ここで、 $T_x$ ,  $D_x$ はそれぞれ点T, DのX座標、 $R_y$ ,  $L_y$ は点R, LのY座標である。すなわち、点T, Dを通る水平線および点R, Lを通る垂直線によってつくられる長方形の中心を新たな画面の中心とするものである。画面移動ベクトルは点( $V_x$ ,  $V_y$ )と画面の主点(視軸と画像面の交点)を結び、画面の移動すべき方向と移動量を示している。新しい主点を実現するために、ロボットアームの第4および第3関節角を調節する。一般的には画面移動ベクトルによって示される位置と姿勢を実現する関節角度をいわゆる逆問題を解いて求めるが、可動範囲の広い軸の関節角度を直接求めた方が速く容易である。第4関節角 $\theta_4$ および第3関節角 $\theta_3$ のそれぞれの増分は次式で与えられる。

$$\Delta\theta_4 = \frac{-Z_{4a} V_y}{f(165 - f + Z_{4a})} \quad (27.a)$$



$$\Delta\theta_3 = \frac{Z_{40} V_x}{f \sqrt{290^2 + (165 - f + Z_{40})^2 + 580(165 - f + Z_{40}) \cos \theta_5}} \quad (27.b)$$

ここで、 $f$ ：焦点距離、 $Z_{40}$ ：第4軸座標系原点から $X_u Y_u$ 平面までの視軸に沿って計った距離である。この $Z_{40}$ の値は座標変換行列 $\hat{T}_{cw}$ および $\hat{T}_{cs}$ から容易に求められる。その他の定数は一般化することも可能であるが、ここでは本システム固有のものを使用している。センタリング処理は画面移動ベクトルの大きさがある程度小さくなったとき、終了する。

### (3) ステレオ画像対の撮像

センタリング処理によって視覚的に捕捉された対象物の画像とステレオ画像対を構成するために、もう1枚の画像が必要となる。次の撮像位置および姿勢を決定しロボットアームを移動させる必要がある。図5.7にあるように輻輳角を設定する。対象物の近傍で視軸が交差することによって、後に述べる対応点問題が容易になる。図においてカメラの視軸の交点位置は対象物の近傍が望ましいが、その3次元位置は未知であるので、視軸と $X_u Y_u$ 平面( $Z_u = 0$ )との交点とする。ロボットアームの移動すべき位置・姿勢を示す座標変換行列 $\hat{T}_{wh2}$ は次式で与えられる。

$$\hat{T}_{wh2} = \hat{T}_{cw1}^{-1} \hat{T}_{zc}^{-1} \hat{T}_d \hat{T}_{zc} \hat{T}_{cw1} \hat{T}_{wh1} \quad (28)$$

ここで、 $\hat{T}_{wh1}$ 、 $\hat{T}_{wh2}$ ：世界座標系におけるハンドの位置および姿勢、

$\hat{T}_d$ ： $X_c$ 軸周りの輻輳角度分の回転行列、

$\hat{T}_{cw1}$ ：世界座標系からカメラ座標系への座標変換行列、

$\hat{T}_{zc}$ ：カメラ座標系の $Z_c$ 軸に沿った並進行列、

$$\hat{T}_{zc} \triangleq \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & r_z \\ & & & 1 \end{pmatrix}$$

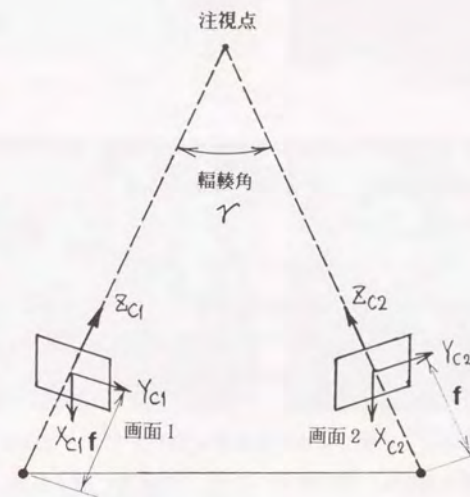


図5.7 輻輳角の設定



$$r_z \triangleq (\hat{T}_{cu1})_{34} / (\hat{T}_{cu1})_{33}$$

$(T)_{ij}$ : 行列  $T$  の  $i, j$  要素

算出された  $\hat{T}_{uh2}$  から逆問題を解いて関節角度を求め、ロボットアームを駆動する。もし逆問題が解けない場合には、予め与えてある関節角度を使用している。

図5.8に視覚的捕捉の実験例を示す。

### 5.3 対象物の認識

3次元対象物の認識はその種類および位置・姿勢の2つに行う。そのために、本実験システムでは2次元輪郭形状に関する特徴量および計算された位置座標に基づく3次元特徴量を用いている。

3次元物体の認識において最も本質的な問題は、物体の見え方の変化に由来するものである。たとえば、遮蔽、隠れ面、縮退などはその典型例である。物体とカメラとが標準的な位置関係にあるにしても、そこにはさまざまな見え方の変化がある。そのような認識対象に対して、次のようなアプローチが考えられる。

①見え方の変化の中から（ときには複数の画像を用いて）、3次元構造に依存する不変な特徴量を抽出し、識別を行う<sup>48)</sup>。そのとき、推定あるいは推論に基づく必要もある。

②見え方を特徴付ける2次元特徴量を抽出し、それによって識別を行う<sup>28)</sup>。

①のアプローチでは、特徴抽出過程に難しさがあり、そこでは推論手法も用いられる場合もある。しかしながら、一旦、3次元特徴量が算出されれば、ある方向において同じ見え方をもつ異なる対象物を識別する能力を期待できる。②のアプローチでは、特徴抽出過程はそれほど複雑ではないが、特徴量はあくまで見え方に依存しており、異なる対象物の同じような側面を識別するのは難しい。識別力の不足を特徴量の種類と速さで補っている。マッチングに必要なモデル数という観点からは、①のアプローチが優っている。なぜなら②では、1つの対象物の異なる見え方に対してもモデルを用意しなければならないからである。しかし、②のアプローチ

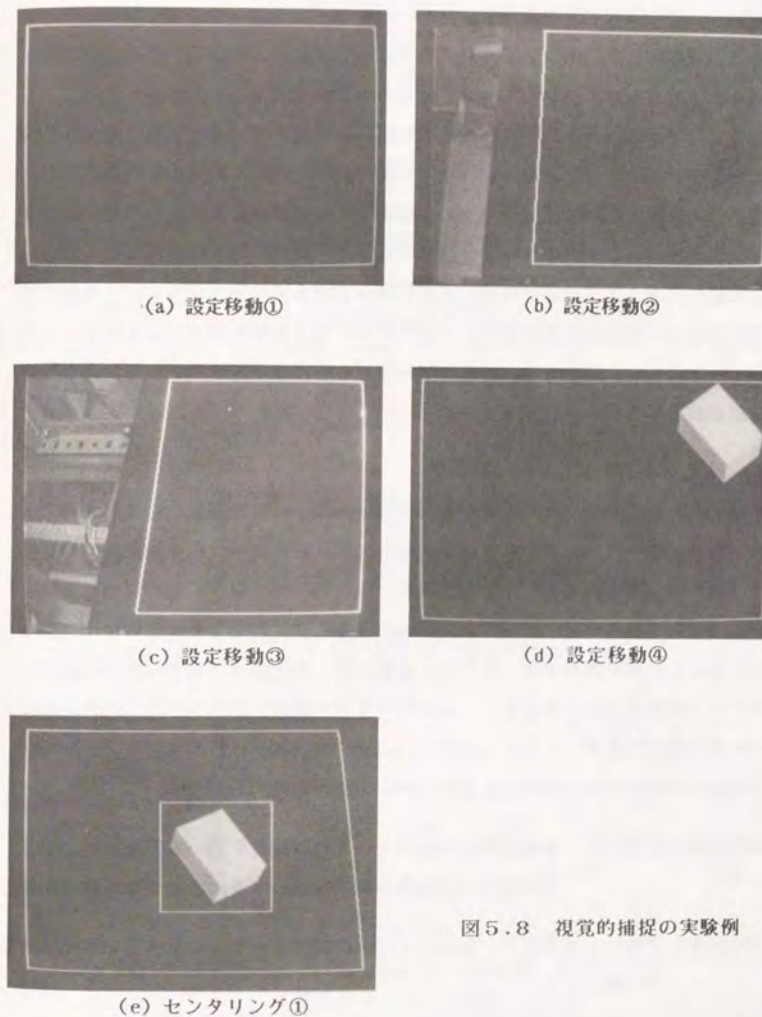


図5.8 視覚的捕捉の実験例



では、特徴抽出計算の簡明さによる特徴の信頼性は高いと思われる。したがって、ここではより実際の応用を指向するために、両方の特徴量を用いて認識を行うことにする。すなわち、物体の見え方はその2次元輪郭形状で記述し、3次元的位置・姿勢を輪郭頂点の3次元位置を用いて記述するものである。

認識におけるクラスに相当するモデルを定義、作成する必要がある。モデルの定義は対象物の識別するための具体的な特徴量の設計を通して構成的に明らかになる。しかし、一般的にその作成はかなりの手間を要する仕事となる。ここでは、認識プログラムそのものをモデル作成にも利用する学習手続き<sup>8,16)</sup>を採用した。ただし、この実験システムでは認識と同時に学習は行っておらず、同一プログラムを利用するにとどめた。認識対象物と同じようにモデルとなる物体もテーブル上におき、同じカメラで撮像し、以下で述べる種々の特徴量の算出を認識プログラムにより行い、その計算結果をモデルファイルに格納していく。

#### (1) 2次元輪郭形状に関する特徴量

本システムの目的に沿って、利用すべき特徴の望ましい性質について考えると、まずロボットアームの位置・姿勢制御のためには、対象物の3次元的位置・姿勢の算出しやすい特徴であることが必要であり、次にハンドリング機能の自律化を目指すためには、照明条件などの外部の環境の小さな変化に対しても安定に抽出可能な特徴であることが重要となる。そこで、本節では、多角形（多面体の投影）の凸包近似という特徴について検討する。これは対象物の輪郭形状を包み込む凸包を構成する頂点の配列であり、ステレオ計算もこれらの頂点にだけ行う。図5.9に2次元輪郭形状の特徴抽出処理の手順を示す。次に項目別に説明を加える。

①濃淡画像の2値化： 画像対を構成するそれぞれの画像において対象物と背景とを分離する灰尺度の閾値を選択する。次の2つを仮定する。

【仮定1】背景は一様であり、画面上でのその面積は対象物より大きいか相当する程度である。

【仮定2】対象物は平均的に背景より明るい（灰尺度は大きい）。

この仮定に基づきながら、画像の灰尺度のヒストグラムにおける「谷」を探す。一般に背景の面積は大きいので、ヒストグラムにおいてもかなり顕著なピークと谷を

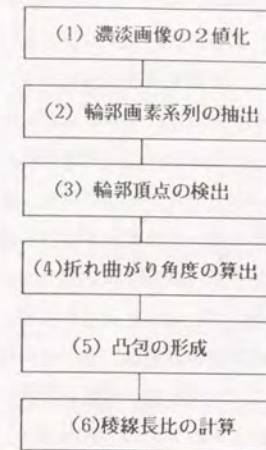


図5.9 2次元特徴量の算出手順

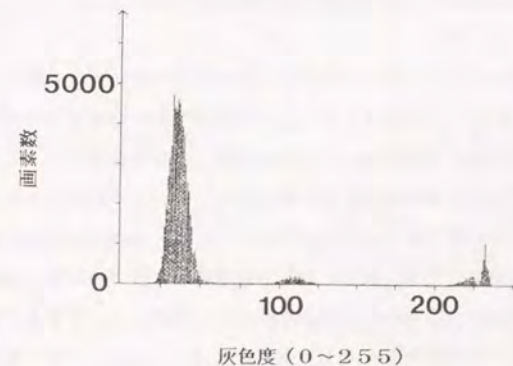


図5.10 ヒストグラムの実例（図5.22(a)の処理例）

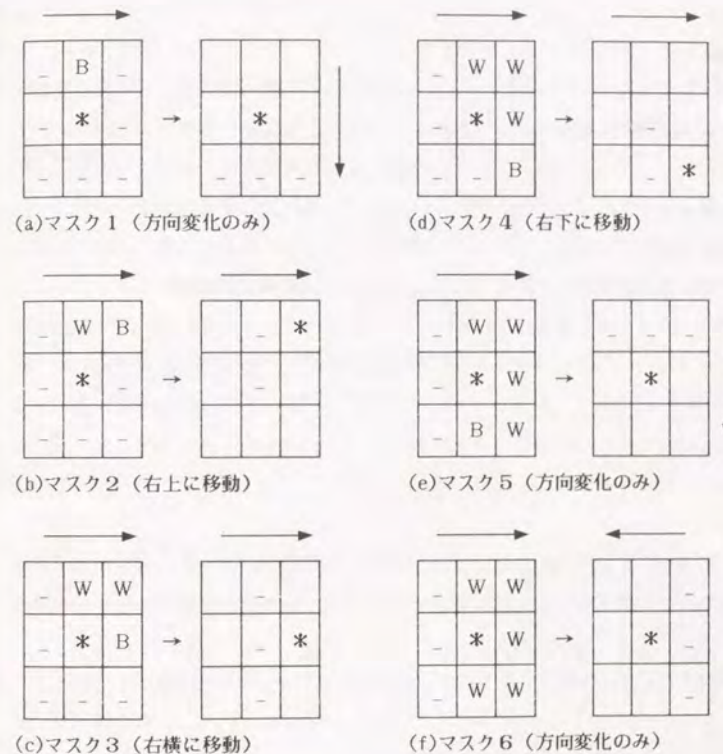


もつ。図5.10にヒストグラムの例を示す。ヒストグラムから累積度数分布に変換し単調増加関数にしておいてから、最初の谷の探索を行う。図5.10のヒストグラムの実例を観察してもわかるように、背景を代表するピークに関しては、ピークの大きさも充分であり、かつその他の小さなピークからの相対位置もある程度離れており分離することは困難ではない。しかしながら、対象物の各面に対応するピークは小さく、しかも互いに近接（低コントラスト）している。さらに、その状態は照明条件の微妙な変化に依存し、背景のピークほど安定には検出されない。

ここでは、対象物の各面を分離せずに、最外側の輪郭を基本特徴データとする。したがって、画像において、対象物の中にある頂点（内部頂点と呼ぶ）は検出されない。第4章において採用した複雑な画像処理は、そこにおける設定がターンテーブル上の対象物の認識であり、ある程度恣意的に環境を設定することもできる。その場合、照明条件などの細かい調整も可能である。たとえば、組立ラインの途中で固定の認識ステージを組み入れるような場合である。本節で述べる手法は環境条件の変化に対してより安定に稼働することを目的とするものである。

②輪郭画素系列の抽出： 輪郭画素系列は輪郭追跡プログラムによって抽出される。図5.11に示すマスク行列を用いて、登録画素および次の探索方向を決定していくものである。この追跡アルゴリズムは単一のループ形状をもつ輪郭を対象としている。また、輪郭によく現れる細い小さな突出部分などはこの段階で除去している。

③輪郭頂点の検出： 輪郭画素系列を折れ線近似するために特徴的な折れ曲がり点を検出する。まず、図5.12(a)に示すように、輪郭画素系列の始点および終点を決定する。このいわゆる頭出しの処理は、もとの輪郭画素系列を巡回的に入れ替え（再番号付け）、最初に現れる曲率の極大点を始点とし、その1つ前の点を終点と定義するものである。次の手続きを再帰的に繰り返す。始点と終点を結ぶ直線線分を考え、その間の輪郭画素のうち直線線分に対して最も遠い点までの距離が閾値を超えればそれを新しい終点とし、その1つ前の点を新しい終点とし、いままでの部分的な画素系列を2つに分けて手続きを適用する（図5.12(b)）。もし超えなければ、輪郭頂点候補として登録する。



注) (1) ‘\*’は現在点, ‘-’は無関係点, ‘W’は白画素(対象物), ‘B’は背景点.  
(2)追跡方向は0:右, 1:下, 2:左, 3:上.

図5.11 輪郭画素系列追跡のためのマスク行列



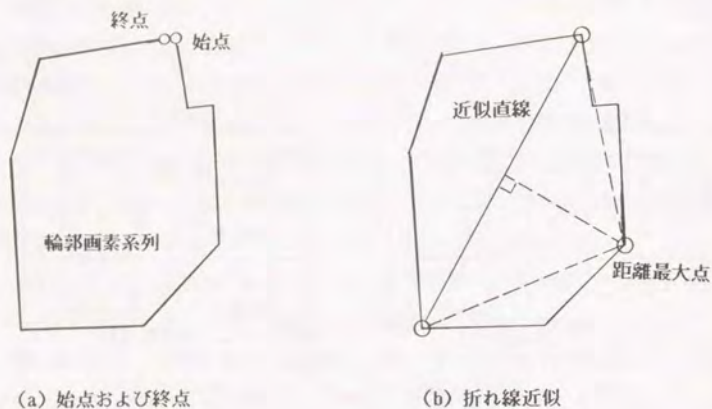


図5.1.2 輪郭頂点の検出

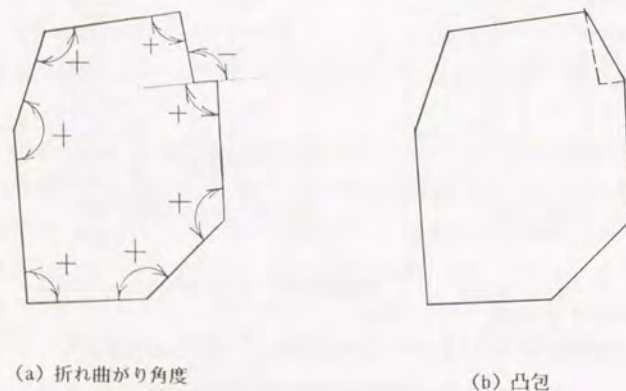


図5.1.3 凸包の抽出

④折れ曲がり角度の算出： 図5.1.3 (a) にあるような折れ曲がり角度を計算する。時計回りに進むとして、進行方向に向かって右に曲がる時には正、左に曲がる時には負の角度と定義する。このような角度を定義すると、その正負が多角形の凸凹に対応する。

⑤凸包の形成： 負の折れ曲がり角度をもつ点（凹の頂点）を除去して前後の点を連結していくと凹の部分が埋められた凸包を構成できる（図5.1.3 (b)）。最終的な特徴点としてはこの凸包を構成する頂点を選ぶ。

2次元輪郭を多角形近似するためにその凸包を選ぶ理由は、凹の部分をもつ対象物をも認識の対象とするためである。凹多面体には自らの面による遮蔽が見られるが、そのとき、一部遮蔽された部分の輪郭には見かけの頂点（第4章で考察したT型頂点）が現れる。そのような点はステレオ撮像による位置計算が不可能である。そのような見かけの点は輪郭上の特徴点から削除しなければならない。そのために、見かけの点は必ず凹となるので、凹頂点をすべて除去する。勿論、凹頂点であるからといって、必ずしも見かけの点ではないが、それらの区別は困難であると思われる。

⑥稜線長比の計算： 凸包を構成する稜線の長さを求め、その最大長で正規化し、降順にソートして特徴量とする。折れ曲がり角度も凸包に対して再計算しておく。

以上の6つの処理を画像対の2枚の2値画像に施し、2次元輪郭形状に関して次の特徴量を得る。

- [1]稜線数（頂点数）： $E$
- [2]折れ曲がり角度： $\{\alpha_i\}$  ( $i=1, 2, \dots, E$ )
- [3]稜線長比： $\{RE_i\}$  ( $i=1, 2, \dots, E$ )

これらのマッチング用特徴量は画像対の中の1枚めの画像に関して算出する。頂点数および頂点位置については2枚めの画像からも算出される。

図5.1.4は2次元特徴量の算出例である。





(a) 検出された輪郭画素系列および輪郭頂点

```

*****_Total_Number_of_Edges_*****
7  ← 輪郭頂点数
*****_List_of_Edge_lengths_*****
88.074
52.887
51.971
51.788
33.838
29.682
23.022
} 稜線長リスト(未正規化)
*****_List_of_Outer_Angles_*****
122.618
162.730
145.678
120.062
91.362
170.031
96.516
} 折れ曲がり角度リスト
*****_P,Q_Coordinates_of_vertices_*****
41, 218, 1
66, 234, 1
117, 248, 1
149, 237, 1
159, 186, 1
73, 167, 1
50, 166, 1
} 輪郭頂点座標リスト

```

(b) 2次元特徴量 (データファイル)

図5.14 2次元特徴量の算出例

## (2) 3次元的位置, 姿勢, 寸法および形状に関する特徴量

3次元の特徴量はステレオ計算により3次元位置の得られている外輪郭頂点に基づいて計算される。まず, 対象物の位置を代表する点として輪郭中心を用いる。これは多面体の通常重心とは異なり, 輪郭頂点の中心であるので, 選ばれた輪郭頂点の組合せに依存する。次に, 対象物の姿勢を記述するために, 輪郭頂点の共分散行列の固有ベクトルを用いる。3本の固有ベクトル組は輪郭頂点の空間的分布, すなわち, 対象物の形状に依存して変化する。また, 3個の固有値組はそれぞれ対応する固有ベクトル(軸)に沿った頂点座標の広がり程度を表している。対象物の大きさ(寸法)に関する特徴量と見なすことができる。

前項で得られた2次元頂点位置に基づいて位置計算を行い, 3次元位置, 寸法および形状に関する特徴量を計算する。図5.15に処理手順を示す。

①対応頂点の同定: 対応点問題は次の2つの条件を考慮して解く。

⑧拘束直線上に存在する,

⑨上の条件で複数候補が選択されたならば, より距離の近い方を選ぶ。

これらの条件は, 高速処理, 多関節型ロボットアームの可動範囲を考慮して設定し

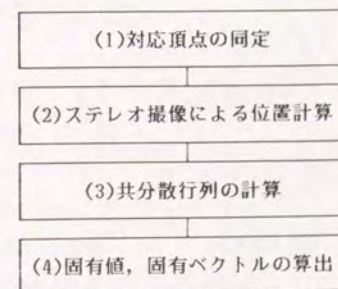


図5.15 3次元特徴量の算出



たものである。②の条件については、ロボットアームの第5軸にカメラが固定されているので、カメラの位置・姿勢パラメタはロボットアームの関節角度から求まり、拘束直線が算出できる。③の条件については、拘束直線上で1対1の対応がつかないときに考慮されるが、対応頂点が互いの近傍に存在するためには、さらに、幅角が小さいこと、対象物を中心に視軸が回転する傾向をもつこと、対象物上の頂点間の距離が視差によるずれより充分大きいことなどの条件が必要である。このうち、幅角は位置精度を保つため、あまり小さくはできないが、他の条件は本実験システムにおいては満足されていると思われる。

②ステレオ撮像による位置計算：次にステレオ撮像による位置計算を行い、輪郭頂点の3次元位置が得られる。

③共分散行列の計算：対象物の姿勢を決定する特徴量として、3次元位置座標の主軸方向を算出する。算出された3次元位置座標ベクトルは世界座標系で表現されているので、まず、カメラ座標系に変換する。

$$\bar{x}_{ci} \triangleq T_{cw2} \bar{x}_i$$

共分散行列C (3×3行列) を次式で定義する。

$$C \triangleq \frac{1}{E} \sum_{i=1}^E (\bar{x}_{ci} - \bar{x}_c) (\bar{x}_{ci} - \bar{x}_c)^t \quad (29)$$

ここで、

$$\bar{x}_c \triangleq \frac{1}{E} \sum_{i=1}^E \bar{x}_{ci}$$

$$\bar{x}_{ci} \triangleq (\bar{x}_{ci}^t, 1)^t$$

ここで、 $\bar{x}_c$ は輪郭頂点の中心であるので、対象物の位置を示す特徴量とみなす。

④固有値、固有ベクトルの算出：共分散行列の固有値および固有ベクトルの組をべき乗法に Hotellingの方法を併用して求める。それらを固有値の大きな順に番号付けて次のように定義する。

$$\{(\lambda_i, e_i)\} \quad (i=1,2,3)$$

ここで、

$\lambda_i$ ：固有値 ( $\geq 0$ )、

$e_i$ ：固有ベクトル。

固有値は対応する固有ベクトルの方向における座標の拡がりを示す量であるので、対象物の空間的拡がり(寸法)を表す特徴量とみなせる。一方、固有ベクトル $e_i$ は対象物の輪郭頂点のみを取り出した場合の主軸方向を示している。また、固有ベクトルの3つ組 $\{e_i\}$  ( $i=1,2,3$ )は縮退していない場合(共分散行列の階数が3の場合)には、3次元空間の基底となりえるので、対象物の形状および姿勢を表す特徴量とみなすことができる。輪郭頂点がある軸周りに回転対称である場合などには、この固有ベクトルの値は一意に決まらない場合もある。実際的には、輪郭頂点の空間的な拡がりがある角度に関して回転対称に近い場合、位置計算の少しの誤差によってその回転角度分の回転が起こり得るなどという問題がある。これらの問題についてはさらなる検討が必要である。今回の実験システムにおいてはこのような場合について特別な処理はしていない。

以上4つの処理を通して、3次元位置、姿勢、寸法および形状に関する特徴量として次のものが得られる。

[1]輪郭中心： $\bar{x}_c$ 、

[2]共分散行列の固有値： $\{\lambda_i\}$  ( $i=1,2,3$ )

[3]同じく固有ベクトル： $\{e_i\}$  ( $i=1,2,3$ )



### (3) 認識のためのモデルとマッチング

この項では主に認識のためのモデルの構造とマッチングの方法について述べる。

#### ① 認識のためのモデルの構造

本システムというモデルは3次元物体のある見え方を表している。すなわち、同じ物体の異なる見え方は異なるモデルとして登録され、扱われる。勿論、同じ物体の異なる見え方であるということをシステムは知ることができる。こうすることの理由は、物体の見え方に依存する情報を表現し易くするためである。本プロトタイプシステムでは、カメラはロボットアームの第5軸に固定されて移動しているので、物体の見え方によって把握できる姿勢に変化がある。したがって、見え方ごとに把握位置・姿勢などを変えて登録することは、人間などが物を把握するときに行う判断の仕方を実装しているといえるであろう。

図5.16に実際の認識用モデルファイルの内容を示す。認識のためのモデルは前2項において算出した特徴量をすべて含んでいる。すなわち、

- [1] 稜線数 (頂点数) :  $E$
- [2] 稜線長比リスト :  $\{RE_i\}$  ( $i=1,2,\dots,E$ )
- [3] 折れ曲がり角度 :  $\{\alpha_i\}$  ( $i=1,2,\dots,E$ )
- [4] 輪郭頂点 (3次元) の数 :  $V$
- [5] 固有値 :  $\{\lambda_i\}$  ( $i=1,2,3$ )
- [6] 固有ベクトル :  $\{e_i\}$  ( $i=1,2,3$ )

この他に次の2つの変換行列および寸法値を格納しておく。

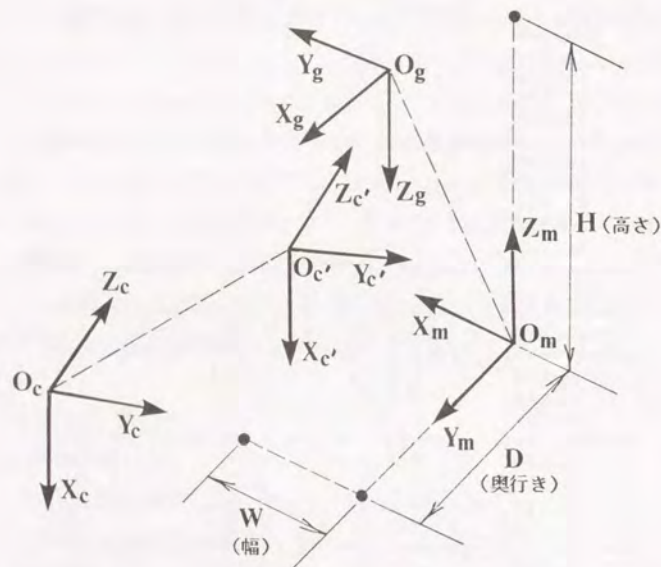
- [7] カメラ座標系からモデル座標系への変換行列 :  $T_{nc}'$
- [8] 把握位置および姿勢を指定する変換行列 :  $T_g$
- [9] 対象物の寸法 :  $W$  (幅),  $D$  (奥行き),  $H$  (高さ)

図5.17にこれらの関係を示す。変換行列  $T_{nc}'$  は輪郭中心  $\bar{x}_c$  へ原点移動したカメラ座標系  $x_c - X_c' Y_c' Z_c'$  からモデル座標系  $O_n - X_n Y_n Z_n$  への座標変換行列である。対象物の3次元空間内の位置・姿勢を指定するためには、通常、対象物に固

```
*****_Model_Number__[ 1000]*****
1000
Comment:[ '89.01.13(Friday) ]
*****_Number_of_Edges_____*****
5
*****_Edge_Lengths__[w.r.t._Image_1]*****
89.828
77.833
46.011
37.537
18.028
*****_Bending_Angles__[w.r.t._Image_1]*****
129.736
134.939
91.884
88.708
94.723
*****_Number_of_3D_Points_____*****
5
*****_Eigenvalues_____*****
1218.969, 315.916, 136.955
*****_Eigenvectors_____*****
0.623, 0.295, 0.724
0.011, 0.923, -0.385
-0.782, 0.248, 0.572
*****_Elements_of_Tmc_____*****
0.564, -0.819, -0.108, 12.346
0.812, 0.573, -0.111, 29.525
-0.153, 0.025, -0.988, 36.000
0.000, 0.000, 0.000, 1.000
*****_Elements_of_Tg_____*****
0.000, 1.000, 0.000, -25.000
1.000, 0.000, 0.000, -15.000
0.000, 0.000, -1.000, 40.000
0.000, 0.000, 0.000, 1.000
*****_W(idth),D(epth),H(eight)*****
30.000, 50.000, 60.000
```

図5.16 モデルファイル





$O_c-X_cY_cZ_c$ : カメラ座標系  
 $O_g-X_gY_gZ_g$ : 把握座標系  
 $O_m-X_mY_mZ_m$ : モデル座標系  
 $O_c'-X_c'Y_c'Z_c'$ : 輪部中心に移動したカメラ座標系

図5.17 モデル内の座標系の関係

定の座標系の位置・姿勢で表現する。上で定義したモデル座標系もそのような座標系であり、普通はモデルとなる実物体の底面（もしくは基準面）にある1つの頂点に原点をおき、寸法が計り易いように座標軸を割り付ける。後の把握位置・姿勢もこの座標系における指定とする。

本プロトタイプ実験システムの目的の1つは、対象物の把握にある。行列  $\hat{T}_g$  は把握位置・姿勢を示す。図5.17にある座標系  $O_g-X_gY_gZ_g$  とロボットアームのハンド座標系（図5.2参照）を一致させるように把握する。一般的には、1つの見え方に対して複数の把握位置・姿勢がある得るが、ここではそのうちの代表的な位置・姿勢を1つ登録する。複数ある場合に拡張するのは容易である。

最後に、モデルの幅、奥行き、高さを登録する。このうち、幅（W）および高さ（H）は次節で述べる把握制御で用いる。

この他に、実際のモデルファイルは管理用のコメント欄やモデル識別子（モデル番号）などの付随情報も含んでいる。

## (2) 対象物データとモデルデータとのマッチング

対象物データとモデルデータとを区別するために次のような記法を用いる。

$f$ : 対象物の特徴量

$jf$ :  $j$  番めのモデルの特徴量

特徴空間内で対象物に最も近いモデルを識別する。ここでは、拡張した近似マハラノビス距離に基づく識別を行っている。モデルファイルを作成するときも認識プログラムによって特徴を抽出しているので、モデルの特徴間の共分散行列  $\Sigma_m$  と、対象物の特徴量間の共分散行列  $\Sigma_o$  は等しいと仮定でき、かつここでは対角行列として近似している。共分散行列は普通対角優勢となる場合が多いが、ここでは異なる特徴量間の相関は0と近似している。さらに、全体的な調整のための加重係数を導入して次式により距離計算を行う。



$$d_j \triangleq \frac{{}^j W_{RE}}{\sigma_{RE}^2} \sum_{i=1}^{E(=jE)} (RE_i - {}^j RE_i)^2 \quad (30)$$

$$+ \frac{{}^j W_a}{\sigma_a^2} \sum_{i=1}^{E(=jE)} (\alpha_i - {}^j \alpha_i)^2$$

$$+ \frac{{}^j W_\lambda}{\sigma_\lambda^2} \sum_{i=1}^3 (\lambda_i - {}^j \lambda_i)^2$$

$$+ \frac{{}^j W_e}{\sigma_e^2} \sum_{i=1}^3 \min \{ |e_i - {}^j e_i|^2, |e_i + {}^j e_i|^2 \}$$

ここで、 ${}^j W_{RE}$ 、 ${}^j W_a$ 、 ${}^j W_\lambda$ 、 ${}^j W_e$ 、および $\sigma_{RE}^2$ 、 $\sigma_a^2$ 、 $\sigma_\lambda^2$ 、 $\sigma_e^2$ はそれぞれ対応する特徴に関する加重係数および分散であり、 $j$ はモデルの番号である。固有ベクトルの評価項では、ベクトルが反転している場合を考慮している。すなわち、モデルおよび対象物の固有ベクトルの3つ組は各々右手系か左手系を構成するが、その組合せは任意であるからである。

まず、稜線数 $E$ および輪郭頂点数 $V$ の等しいモデルを選び、そのなかで上の距離 $d_j$ を計算し、最小のものを最適モデルとして選択する。その意味は、最も見え方が似ていて、かつ、輪郭頂点の拡がりのパタンの似ているモデルを選択するということである。したがって、見え方なり、輪郭頂点としては違わないが、検出されなかった内部の頂点や凹の部分をもつ対象物である可能性がある。

#### 5.4 対象物の把握

認識された対象物をロボットハンドによって把握するための処理およびロボットアームとハンドの制御について述べる。

##### (1) 姿勢の補正および把握位置・姿勢の算出

第5.3節の(3)項で説明した全モデルは、種々の対象物をテーブル上に任意に置き、認識プログラム(特徴抽出ルーティン)によって作成されたもので、その設置の仕方に拘らず、有限の代表的な標本パターンである。一方、対象物は任意の位置・姿勢でテーブル上に置かれるので、最適モデルであっても位置・姿勢のずれは免

れない。位置ずれに関しては、対象物の輪郭中心位置を用いることで、原理的には解決する。しかし、最適モデルに登録されている把握姿勢は補正しなければならない。本システムの目的の1つは対象物の把握であるので、把握姿勢の補正ができればよいのであるが、ここでは一般性を考慮して、対象物座標系の補正問題を考える。ただし、そこには難易の差はない。

図5.18(a)は固有ベクトルの3つ組および物体座標系を示している。補正行列 $\hat{T}_{000}$ を求めるためには対象物と最適モデルとの姿勢差を知らなければならない。そのために、固有ベクトルの3つ組を用いる。固有ベクトルは輪郭頂点位置の拡がりを表しているので、物体の姿勢に固有である。したがって、モデルの固有ベクトルの3つ組を対象物の固有ベクトルの3つ組へ変換する変換行列が補正行列となる。図5.18(b)は変換の様子を示している。まず、モデルの ${}^j e_1$ ( $j$ は最適モデルの番号)を対象物の $e_1$ に一致させる回転変換を $\hat{T}_1$ とする。これは回転軸を ${}^j e_1 \times e_1$ とし、回転角度 $\mu$ を

$$\mu \triangleq \tan^{-1} \left\{ \frac{|{}^j e_1 \times e_1|}{{}^j e_1 \cdot e_1} \right\} \quad (31)$$

とする回転行列である。次の回転行列 $\hat{T}_2$ 、 $e_1$ を回転軸とし、回転角度 $\nu$ を

$$\nu \triangleq \tan^{-1} \left\{ \frac{|{}^j e_2 \times e_2'|}{{}^j e_2 \cdot e_2'} \right\} \quad (32)$$

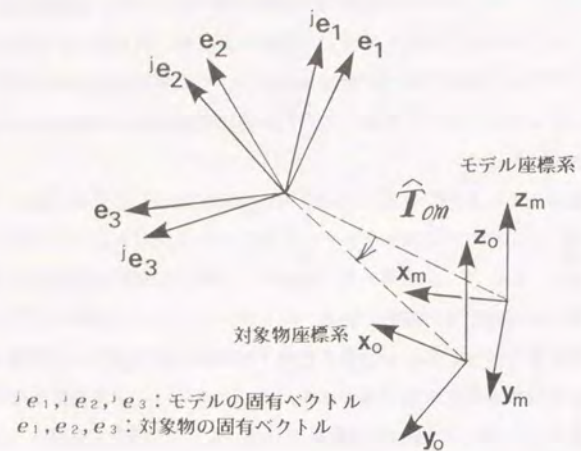
ここで、 $e_2'$ は第1の回転行列によってすでに回転された固有ベクトルを示す。したがって、固有ベクトルの回転行列は

$$\hat{T}_{0000} \triangleq \hat{T}_2 \hat{T}_1 \quad (33)$$

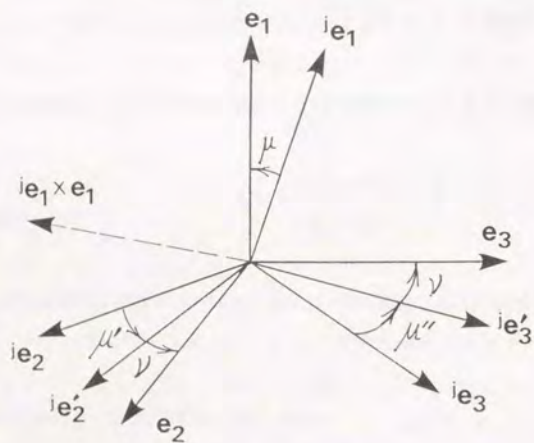
となる。モデルの姿勢補正行列は次式で与えられる。

$$\hat{T}_{00} = \hat{T}_{000}' \hat{T}_{0000} \hat{T}_{000}^{-1'} \quad (34)$$





(a) モデルおよび対象物の固有ベクトル



(b) 2つの補正角度

図5.18 モデル座標系の補正

把握位置・姿勢を示す変換行列  $T_{gu}$  は次式で与えられる。

$$\hat{T}_{gu} = \hat{T}_g \hat{T}_{om} \hat{T}_{mc}' \hat{T}_c' \hat{T}_{cu2} \quad (35)$$

ここで、変換行列  $\hat{T}_c'$  はカメラ座標系  $O_c-X_c-Y_c-Z_c$  を輪郭中心  $\bar{x}_c$  へ並進させるための座標変換行列である。

## (2) 把握処理およびロボットアームの制御

前項で求めた把握位置・姿勢は対象物に依存したものである。その位置・姿勢に関して、ロボットアームが到達可能であるかどうかを判定する必要がある。もし、逆問題の解が得られなければ、把握座標系の  $Y_o$  軸（図5.17参照）周りに微小回転させながら到達可能性を調べるといった簡易的な方法を採用している。

到達可能な把握位置・姿勢と現在のロボットアームの位置・姿勢の間に中間位置・姿勢を設定する。中間位置・姿勢の条件は、そこから把握位置・姿勢まで関節空間における線形補間（関節角度の等分割して駆動する）でも、対象物に触れずにロボットアームが移動可能であることである。実験から、把握座標系の  $Z_o$  軸（アプローチベクトル）の負方向に40～60mm平行移動した位置・姿勢を採用する。実現不可能な姿勢であれば予め準備したオフセット関節角度を使用する。

ロボットアームを最低速度で現在位置から中間位置を通して、把握位置まで移動し、次の把握処理に入る。

ここで論ずる把握処理は次に述べる前提条件をもつ。

- ①ロボットハンドは平行リンク機構による2本指であり、最大幅は55mmである。最大可搬重量は1.5kgである。
- ②力制御は行わないが、指先に装着した接触センサによりある程度弾性をもった把握が可能である。
- ③指先の接触センサは片指12個（縦3行×横4列）の計24個のセンサ配列を構成する（図5.19）。
- ④指面積の1/2以上の面積をもつ平行面を把握の対象とする。



ここでは、図5.19に示すように、各指の接触センサの出力を加重係数として修正並進ベクトル  $g$  を次式で与える。

$$g \triangleq \sum_{i=1}^3 \sum_{j=1}^4 r_{ij} {}^r v_{ij} + \sum_{i=1}^3 \sum_{j=1}^4 l_{ij} {}^l v_{ij} \quad (36)$$

ここで、

$$r_{ij} \triangleq \begin{cases} 1 & \text{(右指センサ } i, j \text{ がオンのとき)} \\ 0 & \text{(右指センサ } i, j \text{ がオフのとき)} \end{cases}$$

$$l_{ij} \triangleq \begin{cases} 1 & \text{(左指センサ } i, j \text{ がオンのとき)} \\ 0 & \text{(左指センサ } i, j \text{ がオフのとき)} \end{cases}$$

であり、 ${}^r v_{ij}$  および  ${}^l v_{ij}$  は右指および左指の各センサの中心位置ベクトル（ハンド座標系における）である。修正並進ベクトル  $g$  の Y 座標値は強制的に 0 としている。図5.19のようなセンサの配置ではハンド座標系の Y<sub>0</sub> 軸方向の位置ずれを計ることが難しいためである。

$$g' = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 1 \end{pmatrix} g$$

この  $g'$  によりハンドの位置・姿勢を修正して把握する。修正並進ベクトルの大きさがある閾値より小さくなるか、オンのセンサ数がある閾値を超えたところで終了する。

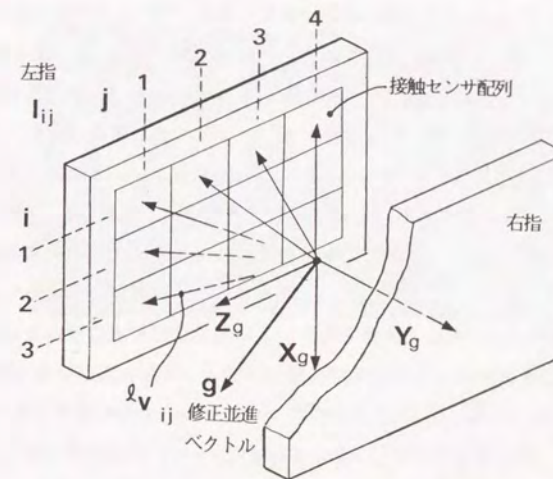


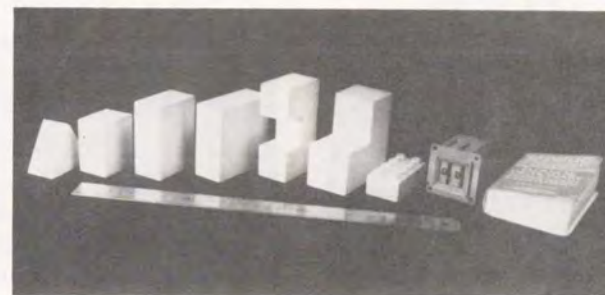
図5.19 接触センサ配列と修正並進ベクトル



## 5.5 実験

図5.20に実験においてモデルとして使用した3次元物体を示す。これらはいずれもロボットハンドで把握可能なものである。モデルファイル内には、6種類の模型の多面体の異なる見え方 40 パターンと、実物の例としてレゴブロック、スイッチボックスおよび小型の本の異なる見え方15パターン（各5パターン）を登録した。

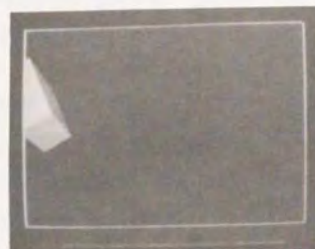
図5.21から図5.28までが一連の実験例である。図5.21は凸多面体④を任意の位置・姿勢に置いた実験の様子である。(a)は設定移動①における画面で、対象物を左端に捉えている。(b)および(c)はそれぞれ1回め、2回めのセンタリング処理の結果である。この例では2回のセンタリング処理で収束した。(d)は画面中央における捕捉を確認した様子であり、この画像がステレオ画像対の1枚めの画像となる。(e)は幅角を $5^\circ$ としてロボットアームを移動させて撮像した2枚めのステレオ画像である。この例では画像①において可視であった右面が画像②では見えなくなっている。(f)および(g)はそれぞれの画像を2値化して輪郭画素系列を抽出し輪郭頂点を求めた結果である。この例では凹頂点は現れない。図5.22および図5.23はそれぞれ2次元特徴量データファイルおよび3次元特徴量データファイルの内容である。図5.24はロボットアームの運動の様子である。(a)は設定移動①、(b)はセンタリング②に対応する。(c)は幅角( $5^\circ$ )を実現するための移動である。(d)から(g)まではそれぞれ中継点、把握、持ち上げおよび移動(ハンド間)の様子を示す。この例では把握位置の修正の必要はなかった。表5.2に処理時間を示す。表5.3に3次元位置の計算誤差および姿勢誤差を示す。姿勢角度の誤差が大きい標準偏差が小さいことから系統誤差であると推測される。



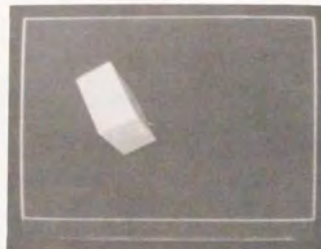
モデル番号	寸法 (mm)	種類
①	30×52×45	三角柱
②	30×40×50	直方体
③	30×50×70	直方体
④	30×70×70	直方体
⑤	40×80×80	L型柱
⑥	30×50×90	T型柱
⑦	30×25×60	レゴブロック
⑧	50×50×60	スイッチボックス
⑨	80×115×27	小型辞書

図5.20 学習されたモデル





(a) 設定移動①



(b) センタリング①



(c) センタリング②



(d) ステレオ画像①



(e) ステレオ画像②



(f) 輪郭頂点の検出 (画像①)



(g) 輪郭頂点の検出 (画像②)

図5.21 凸多面体④に対する実験

```

*****_Total_Number_of_Edges_*****
6
*****_List_of_Edge_lengths_*****
64.885
63.063
42.485
37.577
33.838
33.302
*****_List_of_Outer_Angles_*****
173.218
87.821
113.449
155.937
91.972
97.592
*****_P,Q_Coordinates_of_vertices_*****
81,      184,      1
137,     213,      1
153,     179,      1
131,     154,      1
72,      127,      1
53,      165,      1
*****_Total_Number_of_Edges_*****
5
*****_List_of_Edge_lengths_*****
93.606
63.891
43.012
37.643
30.017
*****_List_of_Outer_Angles_*****
88.940
102.989
167.319
90.289
90.453
*****_P,Q_Coordinates_of_vertices_*****
143,     228,      1
154,     192,      1
128,     177,      1
67,      158,      1
54,      199,      1

```

画像  
①

画像  
②

図5.22 2次元特徴量データファイルの内容



***** Elements_of_Tcw1 *****				世界座標系 から 各カメラ座標 系への 変換行列
0.967,	-0.083,	-0.241,	-626.222	
0.028,	0.975,	-0.222,	-285.916	
-0.253,	-0.208,	-0.945,	687.537	
0.000,	0.000,	0.000,	1.000	
***** Elements_of_Tcw2 *****				頂点数
0.963,	0.065,	-0.262,	-664.471	
-0.103,	0.986,	-0.134,	-192.848	
-0.249,	-0.156,	-0.956,	669.985	
0.000,	0.000,	0.000,	1.000	
***** Number_of_Corresponding_Points *****				頂点の 3次元 座標
5				
***** Coordinates_of_Contour_Points *****				
650.299,	291.458,	5.387,	1.000	
660.069,	268.374,	6.777,	1.000	
656.295,	264.245,	57.803,	1.000	
618.782,	249.561,	62.841,	1.000	
609.190,	273.562,	63.535,	1.000	
***** List_of_Eigenvalues *****				固有値, 固有ベクトル
1070.123,	209.679,	77.153		
***** List_of_Eigenvectors *****				
0.737,	0.314,	0.598		
-0.624,	0.656,	0.424		
0.259,	0.686,	-0.680		
***** List_of_corresponding_points *****				対応点 リスト
0:	-1,	-1,	-1, -1	
1:	0,	-1,	-1, -1	
2:	-1,	0,	-1, -1	
3:	1,	-1,	0, -1	
4:	-1,	-1,	-1, 0	
5:	-1,	-1,	-1, 0	

図5.23 3次元特徴量データファイルの内容



(a) 設定移動①



(b) センタリング②



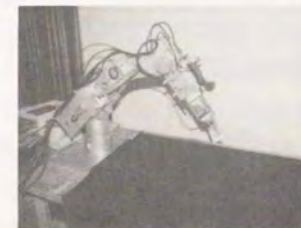
(c) 幅狭角の移動



(d) 中継位置



(e) 把握位置



(f) 持ち上げ位置



(g) 移動

図5.24 ロボットアームの  
運動の様子



表5.2 処理時間(モデル③)

視覚的捕捉	設定移動	8.00" (直立姿勢から) 3.26" (平均)
	センタリング	3.49" (平均)
	ステレオ撮像	4.34" (2枚)
特徴算出	輪郭頂点	18.40" (2枚)
	ステレオ計算	5.73" (固有値, 固有ベクトルの計算も含む)
モデル照合	モデル ファイル入力	7.74" (モデル数: 55)
	照合	0.056" / モデル (平均)
把握	位置探索	0.30" / 回 (10.88" / 36回)
	把握	7.00"
	移動	5.30"
合計		1' 6"

表5.3 計算位置誤差および姿勢誤差

	位置誤差 (mm)				姿勢誤差 (°)			
	x座標	y座標	z座標	距離	主軸	2軸	3軸	全軸
絶対値平均	3.58	4.41	14.3	-----	----	----	----	-----
平均	-0.75	4.08	-3.25	16.4	11.2	13.5	15.0	13.2
標準偏差	4.10	4.03	10.2	7.5	6.1	8.8	2.1	5.2

注) モデル③についての実験. 標本点数: 24 頂点. 幅角: 5°.

姿勢誤差は固有ベクトル間の偏角の大きさ.

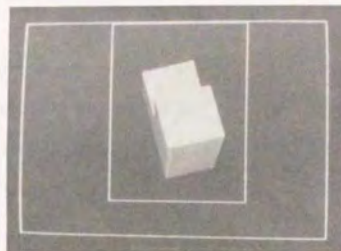
図5.25は凹対象物の処理例である。(a)および(b)はステレオ画像対を示す。(b)および(c)は検出された輪郭頂点を示す。(e)は把握,(f)は修正された把握位置を示す.

図5.26は把握位置の修正の実験例である。(a)および(b)はそれぞれ中継位置および把握位置である. 把握位置は任意に設定した.(c)に接触センサ配列の値を示す. このときはオン・センサ数の閾値を8と設定しているため, 把握位置の修正を行う.(d)は修正後の把握位置を示している.(e)および(f)はそれぞれ修正後のセンサ配列の値および修正並進ベクトルを示す.

図5.27は小型の本を対象物として実験した例である.(a)から(d)に示すように画像処理および特徴量の算出がなされた.(e)および(f)それぞれは中継位置, 把握位置を示す. 把握位置においてハンドが少し傾いたため本の表紙部分に力加わっているのがわかる. 図5.28も実対象物の処理例である.

実験の結果として, 視覚的捕捉処理の成功率は90%以上である. まれに見られる失敗例はセンタリング時にロボットアームの可動範囲を超える場合である. 輪郭頂点の検出では90%以上は安定に抽出できる. 影を捉えないようにカメラ上部に小型照明を取り付けたが, そのことによってカメラ視軸に対して平行に近い面が暗くなり, 2値化画像では稜線が検出できないという場合が見られた. 頂点の3次元位置計算は約80%の成功率である. この原因は主に輪郭頂点の検出ずれである. 画面制御の必要から幅角を比較的小さく設定したために対応処理ではほとんど問題はない. 認識におけるモデルとの照合自体の成功率は, 90%以上である. したがって, 総合的な認識率は65%程度となる. 把握処理においては成功率は未だ低く, 認識された場合には約70%の成功率である. したがって, 完全に把握した場合のみを数えると約45%程度となる. その中には, ロボットアームの可動範囲外に姿勢解が与えられる場合を含む. 把握失敗の主な原因は頂点の位置誤差に起因する把握位置・姿勢のずれである. 表5.3で示したように, xおよびy座標値(x y座標面はテーブル面と一致する)の平均, 標準偏差はいずれも3~4mm程度であるが, z座標値が16mmとやや大きい. これはロボットアームの手先が下向きであることによる. 位置誤差を減少するためには, ロボットアームの可動範囲を考慮すると, 幅角を増やすよりも, ズーム機構をもつカメラなどを利用することが有効であると思われる. 量子化誤差などの誤差解析は第6章で行う.





(a) ステレオ画像①



(b) ステレオ画像②



(c) 輪郭頂点①



(d) 輪郭頂点②

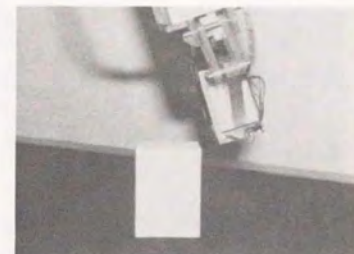


(e) 把握

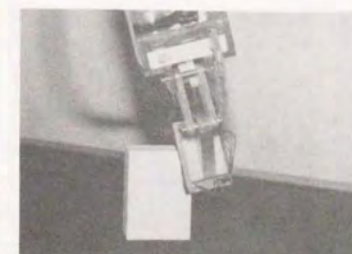


(f) 把握位置の修正

図5.25 凹多面体③の実験例



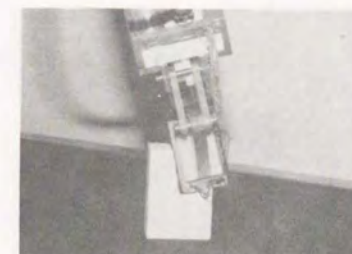
(a) 中継位置



(b) 把握位置

	左	右	
指	0 0 0 0	0 0 0 0	指
	1 0 0 0	0 0 0 1	
先	0 1 1 0	0 0 0 0	先

(c) 接触センサ配列 (修正前)



(d) 修正後の把握位置

	左	右	
指	0 1 0 0	0 0 1 0	指
	1 1 1 0	0 0 1 1	
先	0 1 0 0	0 1 1 0	先

(e) 接触センサ配列 (修正後)

$$\begin{bmatrix} 10.0 \\ 0 \\ -5.0 \\ 1 \end{bmatrix}$$

(f) 修正並進ベクトル

図5.26 把握位置の修正実験





(a) ステレオ画像①



(b) ステレオ画像②



(c) 輪郭頂点の検出①



(d) 輪郭頂点の検出②



(e) 中継位置

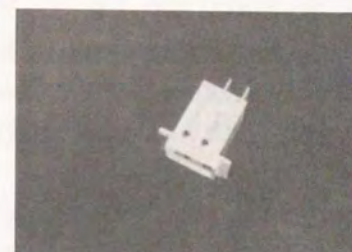


(f) 把握

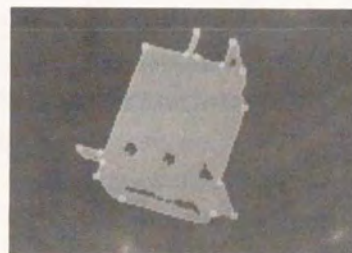
図5.27 実対象物（小型辞書）の実験



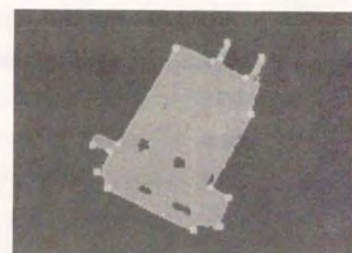
(a) ステレオ画像①



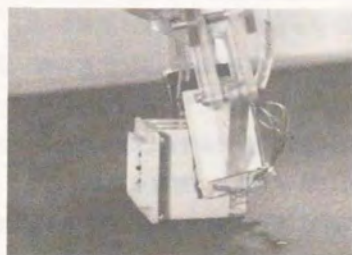
(b) ステレオ画像②



(c) 輪郭頂点の検出①



(d) 輪郭頂点の検出②



(e) 把握

図5.28 実対象物（スイッチボックス）の実験



## 5.6 まとめ

本章では、視覚認識のより実際的应用を可能にするために、ロボットアームの位置・姿勢の自律的制御のために視覚情報を利用する方法について考察した。

次に得られた成果をまとめる。

- (1) ロボットアームに取り付けられたCCDカメラの主に位置・姿勢を制御することによって、3次元対象物の像を自動的に画面中央部に収めるような視覚的捕捉を可能にした。
- (2) 対象物の濃淡画像のヒストグラムから閾値を自動決定し、変換した2値画像に基づく輪郭頂点が安定に検出できることを示した。特に、2次元輪郭形状の凸包近似に基づく形状の特徴量を定義し、その有効性を確かめた。また、3次元特徴量として特に、物体の姿勢を判定するために共分散行列の固有ベクトル組を用いた。実験では、対象物9種類、55パターンを学習させた場合に、認識率約65%という結果を得た。内訳は、2次元特徴量の検出率が90%、3次元特徴量の検出率が80%、認識率が90%である。
- (3) 輪郭頂点の3次元位置の計算精度は、距離平均16.4mm、標準偏差7.5mm（xy平面内では距離平均5.7mm、標準偏差2.7mm）の精度であった。算出された姿勢の角度誤差に関しては、平均13.2°、標準偏差5.2°で再構成できた。
- (4) モデル照合のための特徴量と物体操作のための情報（ハンドの把握位置・姿勢など）を結合してモデル内に学習させることにより、頂点情報の欠落しやすい場合などにも、この方法により認識および操作が可能になることを確認した。実験では、認識が成功した場合には、約70%の把握成功率を得た。
- (5) ロボットハンドによる把握において、指先に取り付けた接触センサ配列の情報から把握位置・姿勢を修正することが有効であることを確認した。

## 第6章 提案した視覚認識法の評価

これまでの3つの章で示した方法に共通する特徴は、次のような点である。

- (1) ステレオ撮像による位置計算法を基礎とする
- (2) 多面体対象物を対象とする
- (3) 特徴点として頂点を検出する

これらの共通特徴をもつと同時に、カメラ、平面鏡、対象物などを含む撮像系の設定や、画像処理、対応点問題の解法などにおいては、それぞれ異なった特徴をもっており、各々に個別の問題点が考えられる。本章では第3、第4および第5章において提案した方法の問題点について検討する。さらに、共通の問題として撮像力、計算精度の問題などの検討を行う。

### 6.1 問題点と改良の見通し

本節では問題点およびその改良の見通しについて議論する。

#### (1) 直接像および鏡像を利用した位置計算法（第3章）

①対応点問題： 視野内に平面鏡を設置することによって、通常のステレオ撮像の効果を得ることができる。そこにおいては、対象物のいわゆる裏側をも同時に1枚の画像の中に捉えることが可能になる。通常ステレオ法や立体視などといわれている方法においては、一般的位置（General position）<sup>2)</sup>から対象物を撮像するものが多い。すなわち、その位置からわずかに動いても見えている対象物の見かけのトポロジーが変化しないような位置である。一方、鏡像を利用する方法では、物体をほぼ横側もしくは裏側から眺めることになるので、直接像と鏡像とは異なるトポロジーをもつことの方が多い。したがって、対応点（頂点）問題を解くことに関しては、通常のステレオ法がより容易である。しかしながら、筆者らの検討<sup>57)</sup>によれば、多面体のような対象物であれば、頻繁に現れる頂点構造があり、それらの鏡



像における変化の仕方を参照表に登録しておき対応付けに利用できる。3面頂点に関する参照表を表6.1に示す。

②撮像時の被写界深度：空間的には鏡像は対象物から鏡面までの距離の2倍近く対象物より遠くにある。したがって、カメラの被写界深度は通常のステレオ撮像に比べて深くなければならない。この要求は普通大きな絞り値によって満たされるが、そうすると画面が暗くなりコントラストが低下する。したがって、比較的強い雰囲気照明が必要となる。

③背景の写し込み：視野内に鏡面があるので、鏡面の置き方如何によってはその反射によって余計な背景が写る場合がある。視野内の背景処理の問題は通常のステレオ撮像をも含めた視覚認識における大きな問題（情景解析と呼ばれる分野）である。鏡像を利用する方法の場合には、鏡面は不必要に大きいものは使用せず、カメラのチルト角を比較的大きく設定することで背景をきれいにすることは可能である。

## (2) 多視野の画像対を利用する方法（第4章）

①チルト角固定の問題：第4章で述べた設定では、ターンテーブルに対するカメラのチルト角は固定であった。したがって、カメラの視軸に対してより鋭角に交わるような面はターンテーブルを回転しても撮像することができない場合がある。しかし、そのような面は最後に再構成されていないことが判定できるので、もし、カメラ側に計算機制御の可能な上下移動機構および縦の首振り機構があれば、位置を選択的に絞り込んで撮像し、局所的な処理で再構成可能である。

②面、稜線の縮退や重なりの問題：多視野法では画像枚数の増加に伴い、対象物の面や稜線が画面上で縮退したり、重なったりする確率が明らかに増加する。その場合、この方法では面が再構成されない。したがって、全体形状の一貫性のチェックのときに、【条件2】に違反して再構成できたということにならず、再度画像対の撮像から行われることになる。

表6.1 3面頂点の鏡像反転ボタン

No.	対象物	鏡像	No.	対象物	鏡像
1			10		
2			11		
3			12		
4			13		
5			14		
6			15		
7			16		
8			<p>注) + : 凸稜線 - : 凹稜線 --- : 遮蔽稜線 ■ : 背景</p>		
9					



### (3) 2値画像の輪郭頂点のステレオ計算法(第5章)

②視覚的捕捉法のリアルタイム視覚フィードバック制御への拡張性: 視覚的捕捉の方法は、たとえば、移動物体の追跡などの問題へ拡張可能であると思われる。しかし、画像処理、ステレオ計算などを含む認識処理の高速化は、リアルタイム視覚フィードバック制御へ拡張するためには是非必要である。そのための工夫の1つとして、モデル情報のより積極的な利用が考えられる。ひとたび、最適モデルが決定できれば、後の制御はモデル中心に行うことが有効であると思われる。たとえば、ロボットアームの軌道の変化に伴い変化する対象物の見え方をモデルに対する簡単な透視変換から推測することができよう。さらに、画像処理の高速化を計るために、予想される視野に処理を限定することは有効である。

③モデルの有限性の問題: このシステムでは、識別をはじめ対象物座標系の設定や把握位置・姿勢の決定などを行うために、扱う対象物はすべてモデルファイル内に登録する必要がある。この制限を緩和するために、より細かい特徴の組合せによってモデルおよび対象物を記述する手法<sup>6,3)</sup> およびより抽象化された表現法に則って対象物を抽象度の高いレベルで記述する方法<sup>7)</sup> などがある。ただし、本プロトタイプシステムの目的の1つでもあるが、定量的認識を行うことが重要であると考えられる。すなわち、直方体であればその寸法、位置・姿勢などを定量的に測定し、操作系(ロボットアームなど)への制御入力とすることを目的としているので、記述をあまり抽象化することはできない。しかしながら、特徴をより細かい特徴要素の組合せとして表現する手法は利用可能であると思われる。また、物体を把握することだけが目的であれば、把握可能な部分(たとえば、大きな面)を中心に探索するやり方などがあると思われる。

## 6.2 ステレオ撮像による位置計算法の比較・評価

第3, 4, 5章では、それぞれに特徴をもった位置計算手法を提案してきたが、本節では、撮像力を表す定量的な指標に基づく比較考察を行う。

図6.1(a)に示すようなピンホールカメラモデルによる撮像範囲を示す四角錐を考える。これを撮像四角錐と呼ぶ<sup>2,3)</sup>。この四角錐はレンズ中心を頂点とし、視軸

に沿って画像平面を線形に拡大掃引したものである。ここでは、その高さDおよび底面の各辺の長さX, Yを次の式により定める。

$$D = f \frac{\delta_s}{\delta_i}$$

$$X = X_0 \frac{\delta_s}{\delta_i}$$

$$Y = Y_0 \frac{\delta_s}{\delta_i}$$

ここで、 $\delta_s$ : 実空間における分解能、

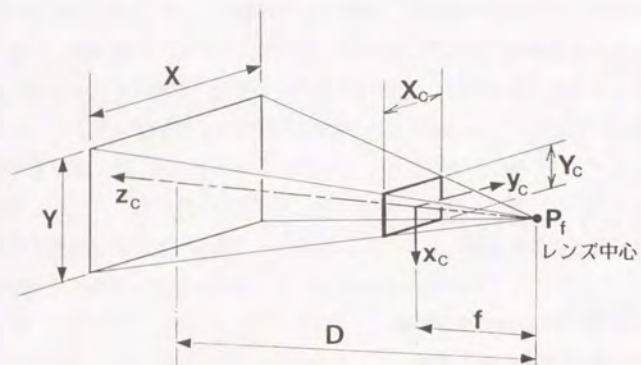
$\delta_i$ : 画像平面上の分解能、

f: カメラの焦点距離

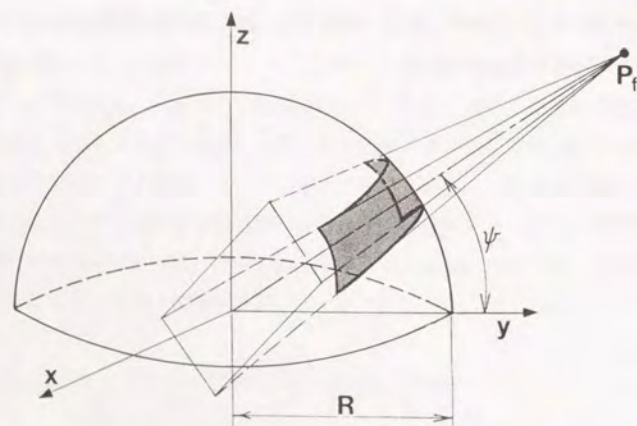
$X_0, Y_0$ : カメラの有効画面の縦および横の寸法

対象物をカメラのよって見込むことは、幾何学的にはこの撮像四角錐によって対象物を覆うことに対応する。





(a) 撮像四角錐



(b) 撮像四角錐の切り取る表面積

図 6.1 撮像能力指標の算出

次に、図 6.1 (b) に示すような対象物を包み込む半径  $R$  の半球を考える。上で述べた画像情報の獲得能力は、それぞれの方法におけるカメラの設定方法によって、撮像四角錐がこの半球のどれほどの表面積を切り取るかによって判断できる。半球を仮定したのは、対象物を下から仰ぎ見ることをしないという前提によっている。したがって、指標  $\kappa$  として半球の全表面積 ( $2\pi R^2$ ) に対する撮像四角錐の切り取る表面積と定義する。

$$\kappa = \frac{\text{撮像四角錐の切り取る表面積}}{2\pi R^2} \times 100 (\%) \quad (19)$$

いま、撮像四角錐の底面の中心が半球の中心を通ると考え、次のようなパラメタを選ぶ。

$$\delta_i = 0.0391 \text{ mm}$$

$$\delta_o = 1.00 \text{ mm}$$

$$f = 35.0 \text{ mm}$$

$$X_o = 9.60 \text{ mm}$$

$$Y_o = 12.8 \text{ mm}$$

$$R = 200 \text{ mm}$$

$$D = f \frac{\delta_o}{\delta_i}$$

すなわち、画素幅  $39.1 \mu\text{m}$  の 1 インチビデオカメラ (有効画面  $9.60 \times 12.8 \text{ mm}$ ) に、焦点距離  $35 \text{ mm}$  のレンズを装着し、対象面上で約  $1 \text{ mm}$  の分解能を得るような設定である。そのとき、単一のカメラの指標  $\kappa_o$  は約  $12.7 \%$  となる。

カメラが任意に移動可能であれば、その指標  $\kappa$  は  $100 \%$  となる。すなわち、対象物をいかなる方位からでも撮像可能である。本節で述べる多視野画像による方法では、カメラがチルト角  $\psi$  だけ傾いた状態で撮像を行うので、その指標  $\kappa$  は

$$\kappa = 75.6 \cos \psi (\%)$$



となる。  $\phi = 45^\circ$  のとき、  $\kappa = 53.5\%$  である。したがって、カメラを任意に移動させる場合と比較すると、約半分強の撮像能力をもっていることになる。

ちなみに、通常のステレオ撮像による位置計算法および第3章で述べた平面鏡を利用する方法について算出しておく。通常のステレオ法では、撮像四角錐は球上で重なりあうので、その重なり率を  $\varepsilon$  ( $0 \leq \varepsilon \leq 1$ ) とすると、指標は

$$\kappa = 12.7 \times (2 - \varepsilon) (\%)$$

となる。一方、第3章の平面鏡を視野内に設置して撮像する場合には、普通重なりは起こらない。したがって、指標は

$$\kappa = 12.7 \times 2 (\%)$$

となるであろう。比較の結果を表6.2にまとめておく。

ここで導いた指標  $\kappa$  は効率を評価する意味ではまだ不十分である。指標で示した領域をどの位の手間(回数)で処理できるかという問題が残る。そのことは撮像系の設定だけの問題ではなく対象物の形状およびアルゴリズムに多く依存するので、それらをも複合的に検討することが必要となる。

表6.2 撮像能力指標  $\kappa$  の比較

方法	指標 $\kappa$ (%)	備考
カメラの任意移動の場合	100	
多視野画像を利用する場合	53.4	$\phi = 45^\circ$
平面鏡利用のステレオ撮像の場合	25.4	重なり率: $\varepsilon = 0$
通常のステレオ撮像の場合	19.1	重なり率: $\varepsilon = 0.5$
単眼の場合	12.7	

### 6.3 計算精度の検討

ステレオ撮像による位置計算法を含んだ画像計測的手法に固有の測定誤差の原因には次のようなものが挙げられる<sup>37, 68)</sup>。

- (1) 画素量子化誤差(画素分解能)
- (2) 撮像系の校正誤差
- (3) 特徴点の検出ずれ(レンズ系の歪, 画像処理)

次に上記の誤差原因それぞれについて基本的検討を加える。

#### (1) 画素量子化誤差

有限の大きさをもつデジタル画像の画素分解能による量子化の影響は最も基本的なものである。図6.2に画面上の1軸(Y座標)における画素量子化誤差の影響を示している。X座標における影響も同様である。3次元空間においては、X座標およびY座標の影響は厳密には点の位置によって異なり独立ではないが、ここでは独立と考えて取り扱う。算出された点の位置には、カメラの画素サイズ  $\delta_i$  の影響で図中の斜線部分の面積に比例する誤差が生じうる。Z座標値およびY座標値に関する分解能を算出する。

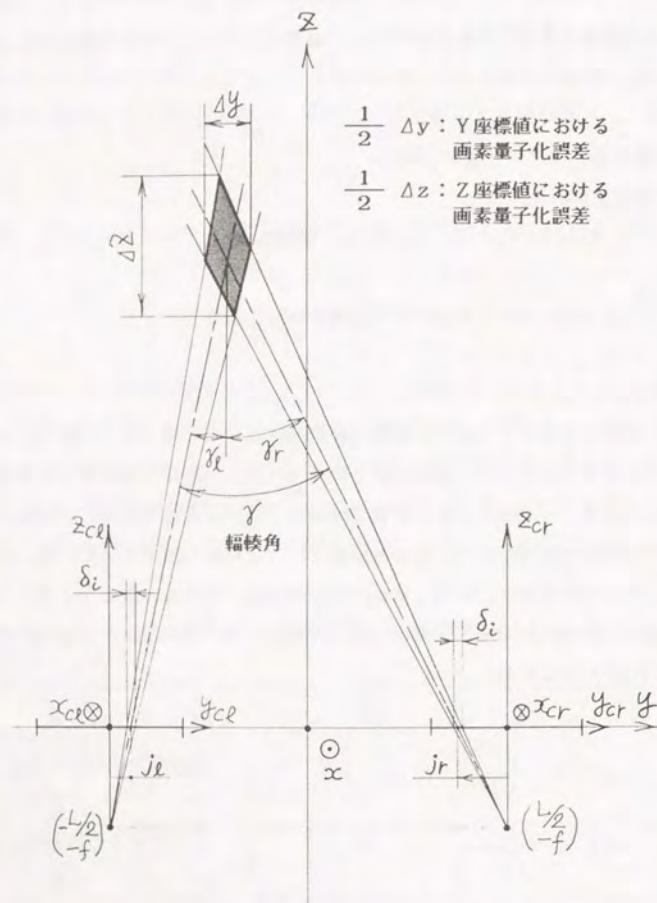
$$\Delta Z \approx \frac{fL}{\Delta j^2 \delta_i} \quad (37.a)$$

$$\Delta Y = \frac{L}{2 \Delta j} \quad (37.b)$$

$$\Delta j = j_i - j_r \quad (37.c)$$

ここで、 $f$ : 焦点距離,  $L$ : 基線長(Base line),  $\delta_i$ : 画素サイズ,  $j_i, j_r$ :





注)  $\gamma$  : 幅収角,  $f$  : 焦点距離,  $L$  : 基線長,  $\delta_i$  : 画素サイズ,  
 $j_l, j_r$  : 画素座標値 (整数:  $-128 \leq j_l, j_r \leq 127$ )

図6.2 画素量子化誤差の算出

画素座標値 (整数:  $-128 \leq j_l, j_r \leq 127$ ) である。さらに, 図6.2に示すように, 幅収角  $\gamma$  との関係を求める。幅収角と画素座標値との関係は次式のようになる。

$$\begin{aligned}
 \gamma &= \gamma_r + \gamma_l \\
 &= \tan^{-1} \frac{-j_r \delta_i}{f} + \tan^{-1} \frac{j_l \delta_i}{f} \\
 &\approx \frac{\delta_i}{f} \Delta j
 \end{aligned} \tag{38}$$

したがって,

$$\Delta Z \approx \frac{\delta_i L}{f \gamma^2} \tag{39.a}$$

$$\Delta Y \approx \frac{\delta_i L}{2 f \gamma} \tag{39.b}$$

となる。

第3章, 第4章, 第5章において述べた手法の分解能を表6.3にまとめる。

表6.3 各手法における画素量子化誤差

	$\Delta z$ (mm)	$\Delta y$ (mm)	備考
鏡像利用法 (第3章)	0.32	0.33	$\delta_i = 0.0391$ mm, $f = 35$ mm $L = 1188$ mm, $\gamma = 118^\circ$
多視野法 (第4章)	1.7	0.45	$\delta_i = 0.0391$ mm, $f = 35$ mm $L = 418$ mm, $\gamma = 30^\circ$
輪郭頂点法 (第5章)	9.32	0.41	$\delta_i = 0.0268$ mm, $f = 16$ mm $L = 44$ mm, $\gamma = 5^\circ$



画素量子化誤差を支配するパラメタは

- ①焦点距離 ( $f$ )
- ②基線長 ( $L$ )
- ③画素サイズ ( $\delta_i$ )
- ④輻輳角度 ( $\gamma$ )

の4つである。誤差は画素サイズに比例する。焦点距離は撮像倍率に關係しており、 $1/f$  (ジオブタ)を減少させることは、画素サイズを小さくする、すなわち、細かい量子化を行うものと同じ効果をもつ。基線長と輻輳角度とはカメラから対象物までの撮像距離が予めある程度決められている場合、ほぼ比例する。したがって、式(39.a)および(39.b)においては( $L/\gamma$ )に量子化誤差が依存している。

ここで導いた画素量子化誤差の見積りのための公式は、三角測量の原理に則った2次元的なものである。3次元空間において厳密に解析するためには、第6.2節で導入した撮像四角錐の交わりを考える必要がある。ここでは、2次元の場合の公式の $\Delta Y$ が $\Delta X$ に等しいと仮定する。また、画素量子化誤差は対象物の寸法を定義する座標系での値に変換されて位置誤差となるので、実際の位置誤差の現れ方は座標変換に依存する。

## (2) 撮像系の校正誤差

ステレオ撮像による位置計算法を行うためには、カメラのパラメタ測定を含めた撮像系を校正することが必要である。写真測量学ではこの校正問題を「写真測量の主問題」と呼び<sup>19)</sup>、画面の中点および焦点距離を求めることを「内部定位」、カメラの位置・姿勢を求めることを「外部定位」と呼んで区別している。ここでは、それらの校正(決定)方法<sup>5)</sup>は述べないが、これらの中に含まれる校正誤差が位置計算に及ぼす影響について考察する。

通常のステレオ撮像による位置計算の定式化は第3章の定式化を参考にして、次のようになる。

$$Ax = -b$$

$$x = -(A^t A)^{-1} A^t b$$

ここで、 $A_1 - A_2 \triangleq (A \mid b)$

$$A_1 \triangleq \left( \frac{I_{2 \times 4} \hat{T}_{cu1}}{I_{2 \times 4} \hat{T}_{cu2}} \right)$$

$$A_2 \triangleq \left( \frac{p_1 \{ r(\hat{T}_{cu1})_3 + a \}}{p_2 \{ r(\hat{T}_{cu2})_3 + a \}} \right)$$

$p_1, p_2$  : 画像1および画像2における対応点位置(2次元)

$\hat{T}_{cu1}, \hat{T}_{cu2}$  : カメラ1およびカメラ2の座標変換行列

したがって、計算される位置座標の誤差 $\Delta x$ は世界座標系においては、次のように表現できる。

$$\Delta x = \Delta x(\hat{T}_{cu1}, \hat{T}_{cu2}, \hat{T}_0, p_1, p_2)$$

すなわち、上の式に現れるパラメタまたは観測値に依存して3次元位置が求められる。しかしながらこの項では、校正誤差の影響を調べるのが目的であるので、まず、観測値である $p_1$ および $p_2$ の影響はないものとする(画素座標値の影響は次項で扱う)。

$$\Delta x = \Delta x(\hat{T}_{cu1}, \hat{T}_{cu2}, \hat{T}_0)$$

次に、変換行列 $\hat{T}_{cu}$ は世界座標系におけるカメラ座標系の位置・姿勢を指定するものであり、ステレオ撮像による位置計算に固有の影響は及ぼさない。なぜならば、カメラの光軸に沿った奥行き方向と、前頭面方向(光軸に垂直な方向)の双方における誤差がステレオ撮像に固有であり、それらがカメラの位置・姿勢に従って座標変換されて世界座標系における位置誤差となるからである。そこでここでは、次のように一方のカメラの世界座標系における位置・姿勢を残し、他方のカメラは1台めのカメラに対して位置決めされているとする。

$$\hat{T}_{cu2} = \hat{T}_{21} \hat{T}_{cu1}$$



したがって、

$$\Delta x = \Delta x (\bar{T}_{21}, \bar{T}_p) \quad (40)$$

となる。 $\bar{T}_p$  は焦点距離  $f$  のみに依存する。 $\bar{T}_{21}$  はステレオ画像対におけるいわゆる輻輳角を与えるパラメタである。画像内の各対応点ごとにその視差角度は異なるが、ここでは図6.3に示すようにカメラの光軸の輻輳角度  $\gamma$  で近似する（実際に、焦点距離が大きいか、あるいは撮像距離に対して物体の大きさが充分小さい場合にはこの近似が有効である）。これは双方のカメラが対象物を画面の中央付近に捉えている場合に対応する。式(40)は、

$$\Delta x = \Delta x (f, \gamma, L) \quad (41)$$

となる。具体的に検討するために、図6.3に示す2次元の場合を考える。図より、距離比  $\beta$  を求めると次のようになる。

$$\begin{aligned} \beta &\triangleq \frac{Z}{Z_1} \\ &= \frac{f \sin \gamma}{f \sin \gamma - q \cos \gamma} \cos \left\{ \gamma - \frac{q}{f} \right\} + \\ &\quad 2(1 - \cos \gamma) - \frac{f^2 \sin^2 \gamma}{(f \sin \gamma - q \cos \gamma)^2} \sin \left\{ \gamma - \frac{q}{f} \right\} \end{aligned} \quad (42)$$

となり、その全微分  $d\beta$  を次のように定義する。

$$d\beta = \frac{\partial \beta}{\partial \gamma} d\gamma + \frac{\partial \beta}{\partial f} df \quad (43)$$

それぞれの感度係数の変化の様子を図6.4に示す。

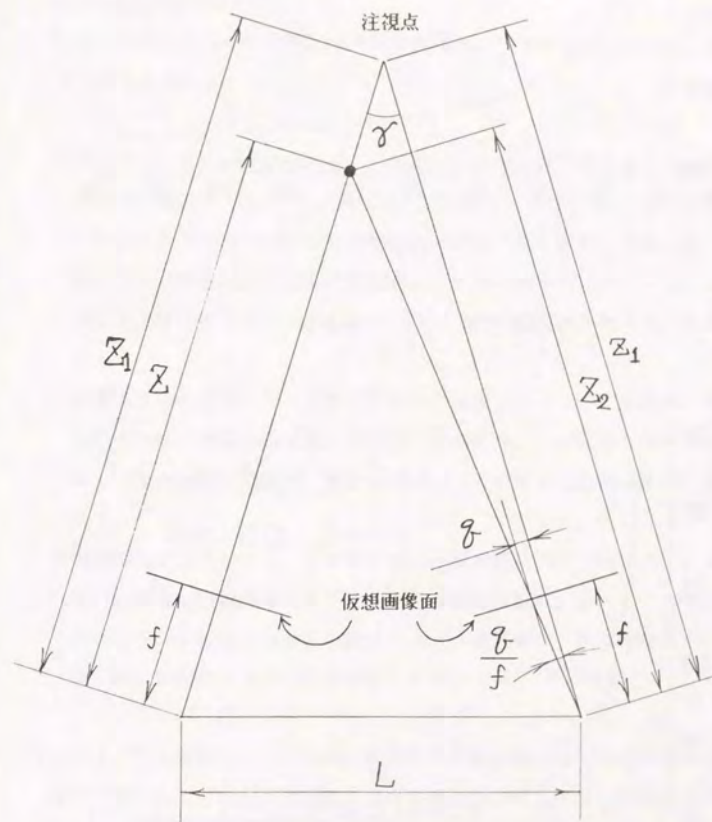
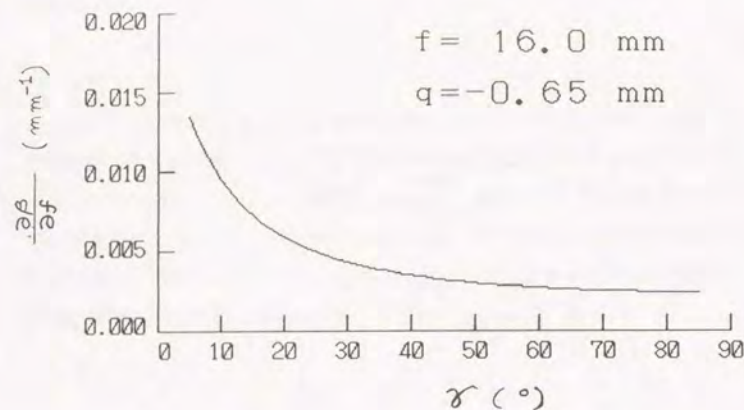
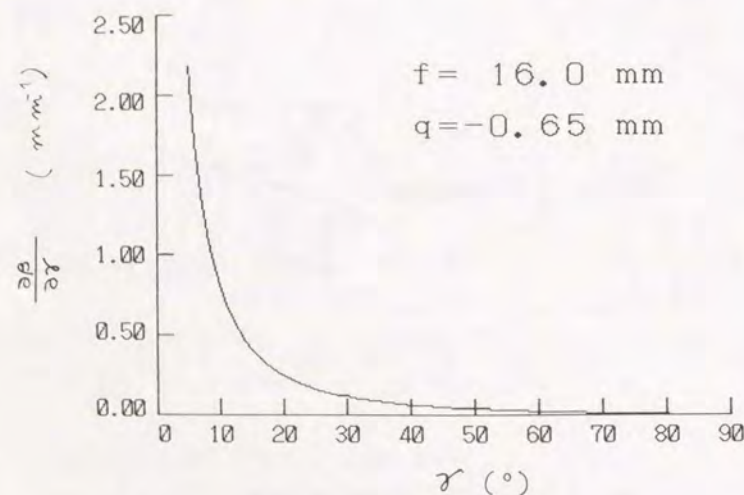


図6.3 2次元空間における奥行き計算





(a) 感度係数  $\frac{\partial \beta}{\partial f}$  の変化



(b) 感度係数  $\frac{\partial \beta}{\partial \gamma}$  の変化

図 6.4 距離比  $\beta$  の感度係数の変化

基線長および幅角は大きい値が望まれるが、対応点問題が複雑・困難になる傾向があるので、その兼合いが重要である。焦点距離は対象物の必要な撮像倍率により決められる。

### (3) 特徴点の検出ずれ

特徴点の検出ずれも位置誤差の大きな要因の 1 つである。その原因には次のような項目が考えられる。

- ①カメラレンズ系の光学的な歪： 通常のカメラレンズにおけるレンズ歪曲の大きさが撮像面サイズの 0.2 % 程度といわれている<sup>32)</sup>。第 5 章の実験で用いた小型 CCD カメラでは、水平分解能約 8.8mm/384 画素であるので、レンズ歪曲により画素位置がずれることはない。しかしながら、たとえば、この倍の分解能のカメラおよびフレームメモリシステムを使用する場合は、考慮する必要がある。
- ②画像処理によるもの： 第 4 章においても述べたように、画像内の頂点などの特徴点を検出するためには、平滑化、勾配計算、2 値化などの画像処理を用いる。特徴点の検出位置は、画像ノイズによるだけでなく画像処理特性にも依存する。
- ③撮像環境によるもの： 第 5 章で述べた輪郭頂点の検出法においては、2 値化閾値を画像の灰色度ヒストグラムから自動決定している。そのとき、原画像においてコントラストが小さい部分（カメラの光軸に平行な面など）の 2 値化では、稜線が欠けるなどの変形が起こり易い（第 5.5 節参照）。

ここではそれらの効果をすべて単に画面上の位置ずれとしてその影響を考察する。

前項で導いた式 (42) より距離比  $\beta$  の特徴点の 2 次元位置  $q$  に関する感度係数を求める。図 6.5 に感度係数の幅角  $\gamma$  に対する変化の様子を示す。本節および前節で議論した感度解析の定式化は、ステレオ撮像法を設計するための誤差評価方式の 1 つの指針を与えるものであると考えられる。

表 6.4 に提示した 3 つの方法における感度係数を示す。



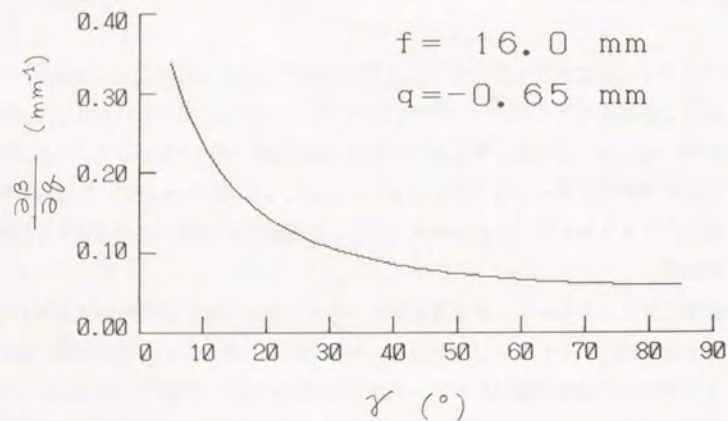


図6.5 感度係数  $\frac{\partial \beta}{\partial q}$  の変化

表6.4 各手法の感度係数

	$\partial \beta / \partial f$	$\partial \beta / \partial \gamma$	$\partial \beta / \partial q$
鏡像法	-0.00183	0.0286	0.0320
多視野法	0.00098	0.0577	0.0532
輪郭頂点法	0.0135	2.180	0.337

鏡像法における  $(\partial \beta / \partial f)$  が他に比べて負になっているが、等価的に幅狭角が  $118^\circ$  と大きいためである。輪郭頂点法ですべての感度係数が大きいのは、方法自体の特性ではなく幅狭角が  $5^\circ$  と他に比べて小さいことによる。

#### 6.4 まとめ

本章では次に示す結果を得た。

- (1) 提案した3つの方法の問題点についてその改良あるいは解決の見通しについて述べた。
- (2) 撮像力指数を提示し、それを用いてステレオ撮像方式の比較が行えることがわかった。
- (3) ステレオ撮像による位置計算法における①画素量子化誤差、②撮像系の校正誤差、および③特徴点の位置ずれに関して一般的な定式化を行い、座標値の計算誤差に対する上記3因子の感度解析を行った。



## 第7章 結論

本研究においては、ロボットビジョンや機械部品の自動組立などへの応用を指向して、（構造的照明などを用いることなしに）通常環境で稼働する3次元物体の認識手法を提案し、実画像の撮像法、画像処理およびグラフ的構造データの処理を含めた計算機による情報処理のより具体的な方法を提示し、基礎的な実験を通してその評価を行った。そこにおいて得られた成果をまとめる。

### 7.1 本研究の成果

第3章で述べた直接像および鏡像を利用した位置計算法に関しては次のような結果を得た。

- (1) 多面体対象物の直接像と鏡像を撮像した単一画像を解析することにより、対象物の頂点の3次元位置が計算できることを示した。
- (2) 対応点の探索問題に無限消失点に関する対応頂点間の幾何学的拘束が利用できることを示した。
- (3) 対象物及びその鏡像の双方において、可視の特徴点の位置は、連立方程式を解き、直接求められることを示した。また、直接像あるいは鏡像の一方でしか可視でない特徴点の位置に関しては、それを含む平面が決定できる場合がありそのとき計算可能であることを示した。
- (4) 多面体を対象物とした実験において、撮像距離 696.5mm、対象物 51×51×63mm の場合、位置測定精度は撮像距離の 1.94% であり、最大標準偏差は 0.66% という結果を得た。

第4章で述べた多視野の画像対を用いた多面体形状の再構成法に関しては次のような結果を得た。

- (1) 3次元多面体の異なる側面を2枚ずつの画像対による多視野の画像として撮像し、それぞれの画像対から再構成された部分的構造を統合して対象物の全

体形状を再構成する方法を示した。

- (2) 必要な画像対の数を減少させるような撮像位置を動的に決定する手法を示した。
- (3) 特徴点の探索処理のために、対応稜線および対応面というより上位の対応関係を考えるが有効であることを示した。また、そのことによって隠れ面や見かけの2次元的トポロジーの変化にも対処できる手法となることがわかった。
- (4) 実験例では、11面から成る多面体について、その2組4枚の濃淡画像から、平均 6.1mm（撮像距離の 0.76%）以内の位置誤差で頂点座標を求め、全体形状を再構成することができた。

第5章で述べた2値画像に基づく輪郭頂点の位置計算および多面体の認識法に関しては次のような結果を得た。

- (1) ロボットアームに取り付けられたCCDカメラの主に位置・姿勢を制御することによって、3次元対象物の像を自動的に画面中央部に収めるような視覚的捕捉を可能にした。
- (2) 対象物の濃淡画像のヒストグラムから閾値を自動決定し、変換した2値画像に基づく輪郭頂点が安定に検出できることを示した。特に、2次元輪郭形状の凸包近似に基づく形状の特徴量を定義し、その有効性を確かめた。また、3次元特徴量として特に、物体の姿勢を判定するために共分散行列の固有ベクトル組を用いた。実験では、対象物9種類、55パターンを学習させた場合に、認識率約 65% という結果を得た。内訳は、2次元特徴量の検出率が 90%、3次元特徴量の検出率が 80%、認識率が 90% である。
- (3) 輪郭頂点の3次元位置の計算精度は、距離平均 16.4mm、標準偏差 7.5 mm（ $x-y$  平面内では距離平均 5.7 mm、標準偏差 2.7 mm）の精度であった。算出された姿勢の角度誤差に関しては、平均 13.2°、標準偏差 5.2° で再構成できた。
- (4) モデル照合のための特徴量と物体操作のための情報（ハンドの把握位置・姿勢など）を結合してモデル内に学習させることにより、頂点情報の欠落しやすい場合などにも、この方法により認識および操作が可能になることを確認



した。実験では、認識が成功した場合には、約 70%の把握成功率を得た。処理時間は平均的に約 1 分である (16 ビットマイクロプロセッサの場合)。ステレオ撮像時に対象物をより拡大して撮像する処理 (Zooming) によって位置測定精度を増し、把握成功率を上げることが可能であると思われる。

- (5) ロボットハンドによる把握において、指先に取り付けた接触センサ配列の情報から把握位置・姿勢を修正することが有効であることを確認した。

第 6 章では、提示した各々の方法における問題点および解決の見通し、撮像力指数の導入、ステレオ撮像による位置計算の計算精度の検討などについて述べた。すなわち、次のような結果を得た。

- (1) 鏡像を利用した位置計算法に関しては、まず、対応点探索処理を参照表などを利用することにより実用化が可能である。また鏡面を用いることに起因する被写界深度や背景の問題は、照明の位置およびカメラと鏡面との相対位置・姿勢を変化させることにより対処可能である。
- (2) 多視野画像を利用する多面体の認識法では、チルト角固定による視野の固定化の問題は、計算機制御可能な上下移動機構および雲台機構によって撮像位置・姿勢を可変にすることによって解決できる。
- (3) 輪郭頂点による物体認識法に関しては、内部頂点を特徴として用いないことの認識力への影響は勿論存在するが、特徴を増加した場合の処理の複雑化、処理コストの上昇などを検討すると、現方式が有効であると思われる。
- (4) 撮像力指数を導入し、それを用いてステレオ撮像方式の比較が可能であることがわかった。
- (5) ステレオ撮像法の位置誤差の解析を行い、画素量子化誤差、カメラパラメタの較正誤差、および画像処理などによる位置ずれに関する感度係数を導出した。

以上の結果より、3 次元物体の全体形状に関する視覚認識については、凸および凹多面体を中心として新しい手法を示すことができた。鏡面を用いて対象物の裏面をも撮像・処理する方法 (第 3 章)、能動的に視野を変化させ全体形状を統合する

方法 (第 4 章)、モデルを用いた認識法 (第 5 章) を示し、全体的画像情報を得るための撮像方法および全体形状を統合する方法を試みた。本研究で提示した手法により、多面体などの視覚認識における基本的問題が解決できたと思われる。また、これらの手法は独立に使用することも、他の既存の手法 (第 1.2 節参照) と組み合わせることも可能である。このことによって、本研究で提示した手法は全体形状の認識の具体的な可能性を拓くものであると考える。

さらに、頂点、稜線および面などの特徴を抽出するための独自の画像処理手法を試みた (第 4 章)。そこにおいては、普通照明下で撮像した画像から、上記の特徴量を濃淡情報を用いて安定に抽出する手法を試み、実験によりその有効性を確認した。それに基づき算出した 3 次元位置に関しても十分な精度を得た。

最後に、認識系および操作系の結合様式に関して、モデルに基づく認識および把握制御の新しい方法を試みた (第 5 章)。そこにおいては、2 次元輪郭形状の凸包近似に基づく特徴量を新しく定義し、有効性を確かめた。また、モデルの有限性に対処するために、モデルの補正方法 (姿勢行列) を提示した。また、特徴量およびその照合過程の安定性を確保するために、学習方式を採用し、実験によりその有効性を確認した。

以上本研究で得られた成果を利用することによって、3 次元物体の視覚認識の応用性および実用性が得られると思われる。

## 7.2 今後の課題

今後解決すべき全体的な課題として、次のような項目が挙げられる。

- (1) 誤り率の低下あるいは認識率の向上: 3 次元形状に関する知識を明確し、計算機内に実現することにより、たとえばモデル内の知識を利用してステレオ撮像による位置計算の結果についてデータの信頼性を判定するような処理が必要であると思われる。第 5 章で述べたロボットアームの制御においても、誤った対応から算出した頂点座標に基づく操作は危険でさえあるからである。
- (2) 処理速度の向上: リアルタイムのロボットビジョンへの応用にはさらなる



高速化が必要である。そのためには、より効率的な情報処理様式を確立するとともに、処理の並列化、基本処理のハードウェア化を進める必要がある。最近の画像処理用計算機<sup>67)</sup>の実用化などによって環境は整い始めたと思われる。

- (3) より多くの対象物の処理： このためには第5章で述べた認識用モデルを拡張することや、幾何モデリングサブシステムの利用、特徴抽出機能の充実などが重要になると思われる。このことは対象物体を多面体（輪郭においては多角形）で近似する場合の表現力あるいは記述力の問題をその具体的方式とともに考察する必要がある。
- (4) 視覚認識と他のセンサとの融合的利用： 視覚センサは非常に高機能であり、汎用性に富んでいるが、さらにその機能を高めるためには、他のセンサを用いて環境の粗い奥行きデータを得る<sup>58)</sup> ことにより、処理の局所化を行うなどの必要があろう。また、逆に視覚情報に基づいた推論により、他の個別のセンサを計画的に使用することも重要である。

## 謝辞

本研究に関して御指導を賜り、本論文をまとめるにあたって種々御教示を戴いた東京大学工学部木村文彦教授に深甚の謝意を表します。

また、本論文を御査読のうえ多くの御助言を下された東京大学工学部木下夏夫教授、大園成夫教授、新井民夫教授、および杉原厚吉助教授にお礼を申し上げます。

折にふれ御指導を戴いた東京大学名誉教授および元東京農工大学工学部教授、明星大学工学部土井康弘教授に感謝の意を表します。

本研究に協力戴いた東京農工大学工学部電子情報工学科（旧数理情報工学科）本多研究室の卒業生並びに在学生の方々に感謝致します。

最後に、本研究を遂行するにあたって常に適切な御指導を賜り、また本論文をまとめる際に御指導と御助言を戴いた東京農工大学工学部本多庸悟教授に謝意を表します。



参考文献(年代順)

- [ 1 ] L.G.Roberts: Machine Perception of Three-Dimensional Solids, Optical and Electro-optical Information Processing, MIT Press, pp.159-197 (1965).
- [ 2 ] D.A.Huffman: Impossible Objects as Nonsense Sentences, Machine Intelligence, Vol.6, pp.295-323 (1971).
- [ 3 ] M.B.Clowes: On Seeing Things, Artificial Intelligence, Vol.2, pp.79-116 (1971).
- [ 4 ] 安江利一, 白井良明: 物体認識のための両眼立体視, 電子技術総合研究所彙報, Vol.37, No.12, pp.1101-1119 (1973).
- [ 5 ] I.Sobel: On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes, Artificial Intelligence, Vol.5, pp.185-198 (1974).
- [ 6 ] L.Kanal: Patterns in Pattern Recognition: 1968-1974, IEEE Trans., Vol.IT-20, pp.697-722 (1974). Also in "Digital Image Processing for Remote Sensing" R.Bornstein ed., IEEE Press, pp.369-394 (1978).
- [ 7 ] G.J.Agin and T.O.Binford: Computer Description of Curved Objects, IEEE Trans., Vol.C-25, No.4, pp.439-449 (1976).
- [ 8 ] 谷内田正彦, 辻 三郎: 機械部品の学習・認識システム, 電子通信学会論文誌, Vol.J59-D, No.3, pp.149-156 (1976).
- [ 9 ] D.Marr and T.Poggio: Cooperative Computation of Stereo Disparity, Science, Vol.194, 15 Oct. 1976, pp.283-287 (1976).
- [ 10 ] A.Rosenfeld, R.A.Hummel and S.W.Zucker: Scene Labeling by Relaxation Operations, IEEE Trans., Vol.SMC-6, No.6, pp.420-433 (1976).
- [ 11 ] K.S.Fu and A.Rosenfeld: Pattern Recognition and Image Processing, IEEE Trans., Vol.C-25, No.12, pp.1336-1346 (1976). Also in "Digital Image Processing for Remote Sensing" R.Bornstein ed., IEEE Press, pp.358-368 (1978).
- [ 12 ] D.Nitzan, A.E.Brain and R.O.Duda: The Measurement and Use of Registered Reflectance and Range Data in Scene Analysis, Proc. IEEE, Vol. 65, No.2, pp.206-220 (1977).
- [ 13 ] D.A.Huffman: Realizable Configurations of Lines in Pictures of Polyhedra, Machine Intelligence, Vol.8, pp.493-509 (1977).
- [ 14 ] 菊池日出男, 辻三郎: 回転台上の多面体パラメータの測定, 電子通信学会論文誌, Vol.J60-D, No.5, pp.339-346 (1977).
- [ 15 ] D.Marr: Representing Visual Information, Computer Vision Systems, A.R.Hanson and E.M.Riseman ed., Academic Press, New York, pp.61-80 (1978).
- [ 16 ] W.A.Perkins: A Model-Based Vision System for Industrial Parts, IEEE Trans., Vol.C-27, No.2, pp.126-143 (1978).
- [ 17 ] 杉原厚吉: 多面体線画のいくつかの特徴量とその意味, 電子通信学会研究会資料, Vol.PRL78-57, pp.33-40 (1978).
- [ 18 ] 佐藤幸男, 渡辺 勉, 小野雅一, 藤田広一: 運動立体視を利用した多面体の測定, 電子通信学会論文誌, Vol.J61-D, No.8, pp.573-579 (1978).
- [ 19 ] 尾崎幸男: 写真測量, 森北出版 (1978)
- [ 20 ] W.N.Martin and J.K.Agarwal: Computer Analysis of Dynamic Scenes Containing Curvilinear Figures, Pattern Recognition, Vol.11, No.3, pp.169-178 (1979).
- [ 21 ] H.Moravec: Visual Mapping by a Robot Rover, Proc. of 6th Int. Joint Conf. on Artificial Intelligence, pp.598-600 (1979).
- [ 22 ] K.C.You and K.S.Fu: A Syntactic Approach to Shape Recognition Using Attributed Grammars, IEEE Trans., Vol.SMC-9, No.6, pp.334-345 (1979).
- [ 23 ] W.M.Newman and R.F.Sproull: Principles of Interactive Computer Graphics, 2nd edition, McGraw-Hill (1979).
- [ 24 ] 烏野 武: 自動組立て用視覚認識装置, システムと制御, Vol.23, No.7, pp.355-361 (1979).
- [ 25 ] 鈴木義規, 木下夏夫, 東村公一, 福井雅彦: 面情報を利用した物体認識(第1報), 昭和55年度精機学会秋季大会学術講演会論文集, pp.644-645 (1980).
- [ 26 ] 白井良明: コンピュータビジョン, 昭晃堂 (1980) (Book).
- [ 27 ] M.Brady: Preface -- The Changing Shape of Computer Vision, Artificial Intelligence, 17, pp.1-15 (1981).
- [ 28 ] T.P.Wallace, O.R.Mitchell and K.Fukunaga: Three-Dimensional Shape Analysis Using Local Shape Descriptors, IEEE Trans., Vol.PAMI-3, No.3, pp.310-323 (1981).
- [ 29 ] W.S.Rutkowski, S.Peleg and A.Rosenfeld: Shape Segmentation Using Relaxation, IEEE Trans., Vol.PAMI-3, No.4, pp.368-375 (1981).
- [ 30 ] D.H.Ballard: Generalizing The Hough Transform to Detect Arbitrary Shapes, Pattern Recognition, Vol.13, No.2, pp.111-122 (1981).
- [ 31 ] 小堀哲雄, 河野隆之, 篠崎志朗: ステレオ画像の三次元自動解析, 情報処理, Vol.22, No.9, pp.846-855 (1981).
- [ 32 ] 田村 稔, 新川幸信: 写真講座, 1.カメラとレンズ, カメラ毎日別冊 (1981).
- [ 33 ] H.Ballard and C.M.Brown: Computer Vision, Prentice-Hall (1982) (Book).
- [ 34 ] D.Marr: VISION, W.H.Freeman and Company (1982) (Book).
- [ 35 ] E.L.Hall, J.B.Tio, C.A.McPherson, C.S.Draper and F.A.Sadjadi: Measuring Curved Surfaces for Robot Vision, Computer, 12, pp.42-54 (1982).
- [ 36 ] S.T.Barnard and M.A.Fishler: Computational Stereo, Computing Surveys, Vol.14, No.4, pp.553-572 (1982).
- [ 37 ] E.S. McVey and J.W. Lee: Some Accuracy and Resolution Aspects of Computer Vision Distance Measurements, IEEE Trans., Vol. PAMI-4, No. 6, pp.646-649 (1982).
- [ 38 ] J.Albus, E.Kent, M.Nashman, P.Mansbach, L.Palombo and M.Shneier: Six-dimensional Vision System, SPIE, Vol.336 Robot Vision, pp.142-153 (1982).
- [ 39 ] 大森隆司, 森下 敏: 多視点ステレオ画像を用いた物体検出, 計測自動制御学会論文集, Vol.18, No.1, pp.716-722 (1982).
- [ 40 ] 大島正毅, 白井良明: 3次元情報を用いた物体認識, 電子通信学会論文誌, Vol.J65-D, No.5, pp.629-636 (1982).
- [ 41 ] 吉澤 徹, 田代発造: 三次元物体の断面測定と形状の復元化, 精密機械, Vol.48, No.9, pp.1187-1192 (1982).
- [ 42 ] 岩田一明, 森脇俊道, 川野常夫, 三原毅史: 荷役作業における作業者の運動



- 解析と評価の研究(第2報), 精密機械, Vol.48, No.11, pp.1465-1470 (1982).
- [ 43] 堤 江美子: 鏡を用いた生体モデル計測法, 昭和57年度精機学会秋季大会学術講演会論文集, pp.63-65 (1982).
- [ 44] H.K.Nishihara: PRISM: A Practical Realtime Imaging Stereo Matcher, Proc. of the 3rd Int. Conf. on Robot Vision and Sensory Controls, pp. 121-129 (1983).
- [ 45] T.Walter and H.Tropf: 3-D Recognition of Randomly Oriented Parts, Proc. of 3rd Int. Conf. on Robot Vision and Sensory Controls, pp. 193-200 (1983).
- [ 46] 白井良明: ロボットと視覚技術, 日本ロボット学会誌, vol.1, No.4 (視覚特集号), pp.244-245(1983).
- [ 47] 柿倉正義, 坂根茂幸: シーンの特徴抽出と環境の理解, 日本ロボット学会誌, Vol.1, No.4, pp.246-254 (1983).
- [ 48] 金子俊一, 本多庸悟: 鏡像を利用した3次元物体の位置計算, 第26回自動制御連合講演会, pp.419-420 (1983).
- [ 49] 杉原厚吉: 線画の代数構造を利用した単眼画像からの立体復元法, 電子通信学会論文誌, Vol.J66-D, No.5, pp.541-548 (1983).
- [ 50] 辻 三郎: コンピュータビジョンへの道, 情報処理, Vol.24, No.12, pp. 1414-1420 (1983).
- [ 51] 池内克史: 単眼視による2次元画像からの3次元情報の再構成, 情報処理, Vol.24, No.12, pp.1421-1428 (1983).
- [ 52] 本多庸悟, 金子俊一, 吉岡和則, 小松敏夫, 土井康弘: 画像図形データ管理システム(PDMS)の一構成, 情報処理学会第26回全国大会論文集, pp.933-934 (1983).
- [ 53] Shun-ichi Kaneko and Tsunenori Honda: Geometric Reconstruction of 3D Object from A Single View of Both Direct and Reflected Images, Proceedings of the 5th Int. Conf. on Production Engineering (ICPE), pp.738-743 (1984).
- [ 54] 金子俊一, 本多庸悟, 北嶋克寛: 鏡像を利用した位置計算にもとづく多面体の再構成法, 昭和59年度精機学会秋季大会学術講演会論文集, pp.475-476 (1984).
- [ 55] D.Nitzan: Development of Intelligent Robots: Achievements and Issues, IEEE Journal, Vol.RA-1, No.1, pp.3-13 (1985).
- [ 56] W.E.L.Gimson: Computational Experiments with a Feature Based Stereo Algorithm, IEEE Trans., Vol.PAMI-7, No.1, pp.17-34 (1985).
- [ 57] Tsunenori Honda(2), Shun-ichi Kaneko and Hidehiko Takeyama(1): Pattern Recognition of Part and/or Workpiece for Automatic Setting in Production Processings, Annals of the CIRP, Vol.34, No.1, pp.29-32 (1985).
- [ 58] 竹中正明, 加藤純一, 大園成夫, 高増 潔, 平野 剛: 半導体レーザを用いた環境センサの製作と応用, 昭和60年度精機学会秋季大会学術講演会論文集, pp.657-658 (1985).
- [ 59] 渡辺正規, 大田友一, 池田克夫: 3眼ステレオ視による位置決め改善とオクルージョン処理, 情報処理学会コンピュータビジョン研究会資料, Vol.38, No.2 (1985).
- [ 60] 大田友一, 正井康之, 池田克夫: 動的計画法によるステレオ画像の区間対応法, 電子通信学会論文誌, Vol.J68-D, No.4, pp.554-561 (1985).
- [ 61] 金子俊一, 本多庸悟: 三次元物体の位置測定法, 公開特許昭60-102504, (1985).
- [ 62] 本多庸悟: 接触検知素子, Pat-pend.昭60-061704 (1985).
- [ 63] T.F.Knoll and R.C.Jain: Recognizing Partially Visible Objects Using Feature Indexed Hypotheses, IEEE Trans., Vol.RA-2, No.1, pp.3-13 (1986).
- [ 64] 金子俊一, 本多庸悟: 直接像及び鏡像を利用した多面体対象物の位置計算, 精密工学会誌, Vol.52, No.1, pp.149-155 (1986).
- [ 65] 洪 海, 川嶋稔夫, 青木由直: 照度差ステレオ法による三次元物体の形状と位置情報の復元, 電子通信学会論文誌, Vol.J69-D, No.3, pp.427-433(1986).
- [ 66] 岡田拓史, 泰 清治, 高木裕治: 立体形状の高速認識方式, 精密工学会誌, Vol.52, No.6, pp.1014-1018 (1986).
- [ 67] 前田 明: 画像処理マシン, 情報処理, Vol.28, No.1, pp.19-26 (1987).
- [ 68] S.D.Blostein and T.S.Huang: Error Analysis in Stereo Determination of 3-D Point Positions, IEEE Trans., Vol.PAMI-9, No.6, pp.752-765(1987).
- [ 69] 高井憲二郎, 木村文彦, 佐田登志夫: 幾何モデルにもとづくロボットの視覚のための物体認識, 精密工学会誌, Vol.53, No.1, pp.131-136 (1987).
- [ 70] 池内克史, 越川和忠: 幾何モデルより導出された解釈木によるピンピッキングタスク中の対象物体の位置姿勢決定, 電子情報通信学会論文誌, Vol.J70-D, No.1, pp.127-138 (1987).
- [ 71] Tsunenori Honda, Shun-ichi Kaneko and Hitoshi Inazumi: 3-D Shape Recognition of Metal Products, Proceedings of the 2nd Int. Conf. on Computer-Aided Production Engineering, Vol.2, pp.9-15 (1987).
- [ 72] Tsunenori Honda, Shun-ichi Kaneko, Satoshi Igarashi, Koichi Murata and Zhan-zhong Cui: A Method for Autonomic Robot Arm to Manipulate Polyhedral Parts, Proceedings of the 6th ICPE, Vol.6, pp.772-777 (1987).
- [ 73] 生体・生理工学シンポジウム論文集(第2回), 計測自動制御学会生体・生理工学会部会, (1987).
- [ 74] 安藤 繁: ステレオ画像の微分恒等式を用いた形状再現システム, 計測自動制御学会論文集, Vol.23, No.4, pp.319-325 (1987).
- [ 75] 安居院 猛, 齊藤知人, 中嶋正之: 背面鏡を用いた3次元物体データ入力システム, 電子情報通信学会論文誌D, Vol.J70-D, No.5, pp.995-1002 (1987).
- [ 76] 出澤正徳: ロボットのための距離検出法, 計測と制御, Vol.26, No.2, pp. 5-12 (1987).
- [ 77] 木村文彦: 形状モデリングとCAD/CAM, 精密工学会誌, Vol.53, No.3, pp.361-364 (1987).
- [ 78] 松下俊夫, 柿倉正義: 3次元物体認識のための距離センサと走査技術, 計測と制御, Vol.26, No.4, pp.59-68 (1987).
- [ 79] 本多庸悟, 金子俊一: 三次元物体の全体形状の視覚的決定法, 公開特許昭62-87807, (1987).
- [ 80] K.S.Fu, R.C.Gonzalez and C.S.G.Lee: Robotics, Control, Sensing, Vision, and Intelligence, McGraw-Hill (1987) (Book).



本多庸悟監訳，本多庸悟，遠山茂樹，古屋信幸，金子俊一訳：ロボティクス，  
日刊工業新聞社（1989）。

- [ 81] K.L.Boyer and A.C.Kak: Structural Stereopsis for 3-D Vision, IEEE  
Trans. of PAMI, Vol.PAMI-10, No.2, pp.144-166 (1988).
- [ 82] 光本浩士，岡崎耕三，田村進一，梶見直樹，副井 裕：複数の鏡像を含む単  
眼画像からの面対称復元手法による隠れのない3D物体復元，電子情報通信  
学会論文誌D，Vol.J71-D，No.11，pp.2300-2310 (1988).
- [ 83] N.Fujita: Assembly of Blocks by Autonomous Assembly Robot with  
Intelligence(ARI), Annals of The CIRP, Vol.37, No.1, pp.33-36 (1988).
- [ 84] Shun-ichi Kaneko and Tsunenori Honda: Visually Controlled Autonomic  
Robot Arm System for Manipulation of Polyhedral Parts, Proceedings  
of the 4th ICCAPE (Edinburgh), Vol.4, pp.121-126 (1988).
- [ 85] 小島徹也，新井民夫：複数の近接覚センサを用いた把握作業の実現，昭和63  
年度精密工学会春季大会学術講演会論文集，pp.45-46 (1988).
- [ 86] 金子俊一，本多庸悟，池田恵一：多視野のステレオ画像対を利用した多面体  
物体の認識，精密工学会誌，Vol.55, No.4, pp.721-728 (1989).



