



Title	動物体のモデリングと追跡に関する研究
Author(s)	日浦, 慎作
Citation	大阪大学, 1997, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.11501/3129131">https://doi.org/10.11501/3129131</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

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

The University of Osaka

# 動物体のモデリングと追跡に関する研究

日 浦 慎 作

1997年1月

大阪大学基礎工学部

## 要約

昨今の急速な計算機性能の向上によって、画像による動的シーンの解析が可能となってきた。また、LSI技術の発展等によりレンジファインダなど画像計測装置の速度も飛躍的に向上している。本論文は、対象物体の運動に伴い得られる画像列を解析することによって、物体の全周三次元形状と表面各点の反射率をモデリングするための二つの手法に関するものである。また、この処理に必要となる対象物体の運動を実時間的に追跡する手法についても詳しく検討した。

まず、対象物体をターンテーブルにより回転させ、それについて変化する画像を一台のカメラにより撮影することで対象物体の形状を求めるシステムを構築した。対象物体とカメラの配置に拘束条件を設け時空間画像処理を用いることにより、受動的計測原理によてもCG生成に利用可能な程度の形状計測が可能であることを示した。また、これに追加する形でCGモデルに必須である表面反射率を求める手法を導入した。円環状の光源を用い、対象物体の運動に伴う明度の変化を解析することによって拡散反射率と法線ベクトルを同時に求める手法を提案した。

次に、物体の運動を実時間的に追跡するシステムについての研究を行った。レンジファインダは濃淡画像センサに較べ低速であり動物体の計測には向かないと考えられてきたが、LSI技術によって飛躍的な高速化を図った機器が出現している。そこでセンサに見合った高速処理として、距離画像と濃淡画像を極力単純にモデルベース処理することにより直接的に幾何学的形状物体の運動が求められることを示した。

さらに、自由曲面物体の追跡を行うためのより高度な処理として、モデルから仮想的な距離画像を複数枚生成し、これらを総合的に用いることにより6軸運動を追跡する手法を提案した。距離画像を生成する過程にCG描画ハードウェアを用いて高速化を図った。また、レンジファインダから得られる形状情報を対象物体の回転に応じてモデルへ統合していくことで、追跡に同時平行して対象物体の全周形状を得る手法を導入した。これによりモデルベース処理によりながらも計測に先立ち対象の形状モデルを用意する必要がなくなった。

最後に、回転対称な物体の計測を行うために濃淡画像を融合し、テクスチャ情報を用いて追跡を行う手法を導入した。距離画像と同様にモデルから濃淡画像を生成する処理にCGハードウェアを用いて高速化を図った。モデルの表面反射率を求める処理にもまたCGハードウェアを用いることでリアルタイム処理を実現した。これらの処理により、追跡に伴いモデルを拡大していくことで最終的に対象物体全周の形状と反射率に関するモデルを得ることができた。

以上の研究により、受動的・能動的計測原理双方において明確なタスク設定とそれに適した機器を構成することで、より高度な画像計測システムが構築可能であることを示した。

# 目次

1 章 序論 .....	1
1.1 画像計測におけるタスクの意義 .....	1
1.1.1 ビジョンパラダイムの発展とタスク .....	1
1.1.2 画像計測機器の構成とタスク .....	4
1.1.3 画像計測の発展と新たな応用 .....	5
1.2 本研究の意義と論文の構成 .....	5
2 章 受動的原理に基づく回転物体の形状計測 .....	9
2.1 対象物体の運動に基づく形状計測 (Shape from Motion) .....	9
2.1.1 運動が未知である場合の形状計測 .....	9
2.1.2 運動が既知である場合の形状計測 .....	11
2.1.3 運動が1軸回転である場合の形状計測 .....	12
2.2 対象物体の回転による全周形状の計測 .....	13
2.2.1 対象物体の回転と特徴点の軌跡 .....	13
2.2.2 時空間螺旋画像 .....	15
2.2.3 時空間螺旋画像を用いた処理の流れ .....	15
2.2.4 シルエット像による形状復元 .....	17
2.3 実験 .....	18
2.3.1 実験装置の構成 .....	19
2.3.2 円筒の計測による精度評価 .....	19
2.3.3 実物体の計測例 .....	20
2.3.4 シルエット像による形状復元 .....	23
2.4 考察 .....	24
2.4.1 時空間螺旋画像による計測結果について .....	24
2.4.2 シルエット像の解析による計測結果について .....	25
2.4.3 計測時間について .....	25
2.4.4 計測装置について .....	26
3 章 円環光源を用いた回転物体の反射率・法線ベクトル計測 .....	27
3.1 各種の反射率計測法 .....	27
3.1.1 一様光を仮定する方法 .....	27
3.1.2 法線ベクトルと光源方向を既知とする方法 .....	27
3.1.3 照度差ステレオ法 .....	27
3.2 反射率と法線ベクトルの計測原理 .....	30
3.2.1 計測点の追跡 .....	30
3.2.2 円環光源による照明 .....	31
3.2.3 放射強度の変化とパラメータ推定 .....	32
3.3 反射率・法線ベクトルの計測実験 .....	34
3.3.1 計測装置の構成 .....	34

3.3.2 円錐の計測による精度評価.....	34
3.3.3 実物体の計測例.....	37
3.4 考察 .....	38
3.4.1 反射率・法線ベクトルの計測結果について .....	38
3.4.2 計測時間について .....	38
3.4.3 装置の構成について .....	39
3.5 今後の課題 .....	39
3.5.1 鏡面反射率の復元 .....	39
3.5.2 カラー計測 .....	39
3.5.3 法線ベクトルによる形状の高解像度化 .....	40
<b>4 章 動距離画像と濃淡画像の融合による幾何学形状物体の実時間追跡 .....</b>	<b>41</b>
4.1 距離画像による物体の姿勢推定と追跡 .....	41
4.1.1 表面曲率を用いる方法 .....	41
4.1.2 多面体の面方程式を用いる方法 .....	41
4.1.3 距離画像間の距離最小化による方法 .....	42
4.1.4 距離画像の固有空間照合法 .....	42
4.2 シリコンレンジファインダ .....	44
4.2.1 シリコンレンジファインダの原理 .....	44
4.2.2 シリコンレンジファインダのキャリブレーション .....	45
4.3 距離画像と濃淡画像による連結物体追跡の原理 .....	48
4.3.1 距離画像と濃淡画像による並列処理 .....	48
4.3.2 座標変換とキャリブレーション .....	50
4.3.3 モデル記述 .....	52
4.3.4 変位の検出 .....	54
4.3.5 追跡パラメータによるラベル画像・距離画像の生成 .....	58
4.3.6 パーツ相互作用とモデルの移動 .....	59
4.3.7 追跡処理とカメラ座標系 .....	60
4.4 実験 .....	61
4.4.1 実験装置の構成 .....	61
4.4.2 固体の位置・姿勢の追跡実験 .....	62
4.4.3 連結物体の位置・姿勢・変形の追跡実験 .....	62
4.4.4 精度評価実験 .....	62
4.5 考察 .....	64
4.5.1 位置・姿勢の追跡について .....	65
4.5.2 変形の追跡について .....	65
4.5.3 計測精度について .....	66
4.5.4 計測速度について .....	66
<b>5 章 複数距離画像の生成による自由曲面物体の実時間追跡 .....</b>	<b>67</b>
5.1 自由曲面の位置合わせ .....	67
5.1.1 ICP .....	68

5.1.2 4章で提案した手法の問題点	70
5.2 CG ハードウェアによる距離画像生成の高速化	71
5.2.1 陰面消去と Z バッファ法	72
5.2.2 Z バッファと同次座標変換	73
5.2.3 CG ハードウェアとレンジファインダ座標系の一致	77
5.2.4 距離値の読み出し	78
5.3 複数の生成距離画像を用いた移動量算出	80
5.3.1 距離画像比較による姿勢推定	81
5.3.2 微小変位量の決定と背景画素の処理	83
5.4 モデリング	84
5.4.1 距離画像によるモデル構築	84
5.4.2 複数距離画像の統合	84
5.5 実験	87
5.5.1 実験装置	88
5.5.2 平行移動の追跡精度評価	89
5.5.3 回転運動の追跡精度評価	89
5.5.4 モデリング結果	92
5.6 考察	94
5.6.1 位置・姿勢追跡の精度について	94
5.6.2 構築された形状モデルの精度について	95
5.6.3 計測時間について	95
<b>6 章 濃淡画像と距離画像の融合による自由曲面物体の実時間追跡</b>	<b>97</b>
6.1 濃淡画像による運動推定	97
6.1.1 勾配法によるオプティカルフローの算出	97
6.1.2 特徴のモデル照合に基づく対象物体の追跡	99
6.2 濃淡画像と距離画像の融合による動物体追跡	100
6.2.1 濃淡画像を用いた移動の検出	100
6.2.2 対象物体の姿勢に関する明度勾配法	101
6.2.3 CG ハードウェアによる明度変化率の計算	101
6.2.4 剛体性の拘束条件に基づく姿勢変化量の算出	103
6.2.5 距離画像との融合	105
6.3 CG ハードウェアによる実時間反射率推定	106
6.3.1 実画像と CG の比較による反射率推定	106
6.3.2 明度値のサンプリングと法線ベクトル	108
6.4 センサ計測方位の回転による対象像の追跡	108
6.4.1 対象像の追跡とカメラ制御手法	108
6.4.2 並列処理	110
6.5 実験	112
6.5.1 実験装置	112
6.5.2 回転運動の追跡精度評価	112
6.6 考察	115

6.6.1 動物体の位置・姿勢追跡について .....	115
6.6.2 形状・反射率モデルについて .....	116
<b>7 章 結論 .....</b>	<b>117</b>
<b>参考文献 .....</b>	<b>119</b>

# 1章 序論

視覚は人間の五感の中で最も情報量の大きな感覚器官であると言われている。特に、外界と身体の位置関係など三次元空間の知覚に果たす役割は大きい。そこで、計算機により人間の視覚システムと同等の機能を実現することで、労働・福祉分野におけるロボット応用をはじめ、人間の知的活動を支援する道具としての可能性を拡大することが期待される。これまで多数の革新的な研究と計算機性能の向上により生産設備の自動化などの分野では実用に供されるまでになったが、人間の視覚のようにシーン中のさまざまな情報を獲得しうる単一のシステムとしての完成にはほど遠い。それゆえシーンが持つ全ての情報を同時に得ようとするのではなく、遂行すべき最終的な目的に適した情報をシーンから抽出することがビジョン研究の主たる目標である。また、用途によっては投光装置など人間には備わらない機構を用いて信頼性を確保することが有効であることを無視すべきではない。本論文は、環境に対して位置や姿勢が変化する、いわゆる動物体に関してその運動を求める、またその形状や反射率を計算機内に獲得することを目的に、独自の装置構成により効果的に計測を行うシステムに関するものである。まずこの章では、視覚システム構築に関する枠組みとその発展について述べ、本論文で提案する各手法の位置づけとその意義について明らかにする。

## 1.1 画像計測におけるタスクの意義

画像にはさまざまな情報が含まれている。タスクを明確にすることなしに、画像から抽出すべき情報の種類やその手段について論じることはできない。ここでは計測機器の構成や計算機内処理の形式とタスクの関わりについて考察するとともに、タスクそのものの変遷についても論述する。

### 1.1.1 ビジョンパラダイムの発展とタスク

我々人間は視覚によりさまざまな情報を得ている。すなわち書物を読んだり自動車を運転したり、また表情からとっさの感情を読み取ったりと、一口に視覚情報とくるのがはばかられるほど多岐にわたる情報を両眼の画像から抽出している。計算機により同等の機能を実現するためには、人間が実現している情報処理過程を解明し模倣する方法がまず考えられる。このような動機に基づく研究は、進化の過程で生物が獲得してきた視覚システムに対する科学的興味という点でまず評価されるべきである。また、画像フィルタやニューラルネットワークなどの実用的な手段の発見における功績も大きい。しかし、人間の脳と同等の機能を果たす情報処理機械の開発は非常に困難であり [Shirai87]、純粹な意味でこの手法が実現可能であるとは言いたい。

ビジョン研究、つまり画像から有用な情報を抽出しようとする研究は、基礎的研究に主眼をおくか、技術の応用性を重視するかによりある程度分類することができる [Yachida90]。その中でも、狭義のコンピュータビジョン (computer vision, 以下 CV) と言われる領域は、基礎的

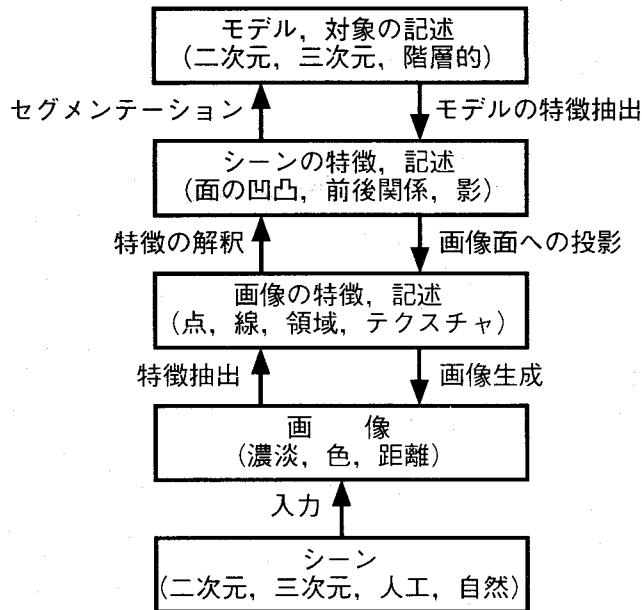


図 1.1 視覚情報処理過程

で科学的な側面を重視する傾向が最も強い。この研究領域では画像やシーンの持つ定性的な性質から演繹的に数学的事実を導出することによって画像を解析する手段を提供することが目的とされており、一般性のないシーンや特殊な計測装置は排除される傾向にある。画像センサや高性能な計算機の普及とともに活発な研究が行われた結果、さまざまの有用な理論が生み出された。代表的な成果として、カメラ系の幾何学的性質から導き出されるエピポーラ拘束条件やキャリブレーションに関する理論、shape from X ( $X = \text{shading, motion, texture, focus, etc.}$ ) に代表される三次元構造の復元理論などがある。しかしその反面、数学的に高度ではあるが一般性や実用性に欠けた研究例も多く見受けられる [Tamura96]。十分汎用的な手法については抽出された情報の利用目的、つまりタスクの設定についても一般的でよいと思われるが、シーンに対する条件や仮定が多い手法については、その手法がどのような応用分野に利用可能かを明らかにする必要があると思われる。

CV 研究の始まりとして、直方体や三角柱などの単純な形状の物体が組み合わされたシーンを対象とし、その位置や種類を決める手法に関する研究 [Roberts65] を挙げることができる。この研究以降、シーンから情報を抽出・解析していくことでシーン記述を得ることを目的とした研究 [Walts75] [Mackworth73] が活発に行われた。しかし、画像からボトムアップ的にシーンに含まれる情報全てを記述しようとする考え方は、単純な「積み木の世界」ではうまく動作するが、より一般的なシーンでは情報量の爆発により実行不可能であることが明らかとなってきた。冒頭の例では、読書をするときにはその本の形状を意識することはないし、また逆に運転時には前車のナンバープレートを読み取ることもしない。なぜなら、それらに関する情報は必要でないからである。そこでこのボトムアップ的ビジョン方法論にモデルの概念を導入した一般的な視覚情報処理過程が Marr により体系化された [Marr82] [Shirai87] (図 1.1)。この過程ではモデルに関する知識を用いて情報の取捨選択を行うことができる事を示しているが、依然として情報の主たる流れは画像からモデルへ向かうボトムアップ形式を保持している。モ

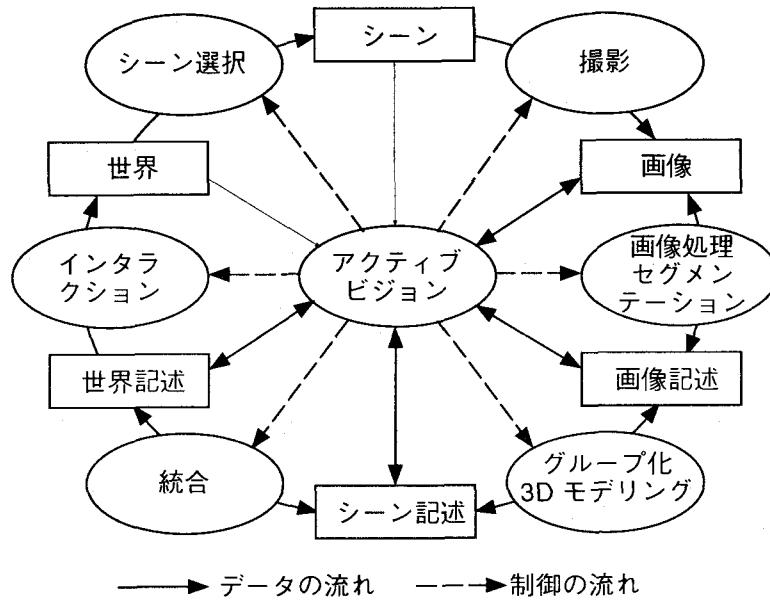


図 1.2 アクティブビジョン

モデル表現に視覚システムの目的、つまりタスクに関する情報を陰に含めることは可能であると考えられるが、モデルに記述しうる情報を全て抽出しようとする点では変化がない。しかし冒頭の例では、読書時にはすでに読み終わった字句を認識することができるように、人間がそれぞれのタスクを実行しているときにはそのタスクに必要な情報を選択的に画像内から抽出していると考えることができる。この考え方をアクティブビジョン (active vision) [Aloimonos87] [Kuno95] と呼び、個々の処理過程において研究がなされている（図 1.2）。もちろんむやみにアクティブ化するのがよいのではなく、固定的なタスク設定のもとではその必要はないし、またアクティブ化の形態とタスク遂行戦略との関連は密接であるために、より高次のタスクを明確に意識することが重要である。これはビジョン研究全体についても同様である。

アクティブビジョンの考え方によれば、タスクによる処理領域の限定などにより処理すべき情報の絶対量が減少するために、モデルと入力を直接照合するような単純な手法が利用可能となる。結果的にモデルの導入による情報選択の効果がより発揮されることになり、複雑な中間記述を用いることなく処理を進めることができる。そこで、より良好で限定的な入力を前提とし、モデル構造をより直接的に用いた処理形態を model based vision と呼ぶ。例えば机の上にあるべきコップをシーン中から探すとき、Marr のパラダイムのようにシーンの三次元的な記述を得てからコップを探すのが良い方法といえるだろうか。実際にはコップの色彩や二次元画像テンプレートを用いてコップらしき部分を探すことでも十分事足りるはずである。もちろんコップを手にとるときにはその位置を把握する必要があるが、この段階ではすでにコップのおおよその位置が分かっているはずであるから、さらに複雑だが精度の高い手法を用いればよい。このように、アクティブビジョンの考え方と、タスク遂行戦略・モデル記述は密接な関係にあることを意識する必要がある。

### 1.1.2 画像計測機器の構成とタスク

人間の視覚はそもそも生命を維持するための情報源として発達してきた機構であり、根元的なタスクは種の保存にあると言える。二眼球運動機構、三原色知覚や網膜構造などは個体ごとに備わる機構として実現可能な範囲で最適化されていると考えられる。その点で計算機はいかに発展しようとも人間が利用する道具であるから、全ての機器が同一である必要もなく、また多目的である必要もない。つまり、人間の視覚と同様の機器構成により高度な機能を実現しようすることは自然科学的に興味深くても工学的な価値は薄い。人間が持つ視覚情報処理機構は多機能実現の観点でも非常に洗練されているが、工学的には目的達成のために最適であると言える場合にのみ同様の構成を採るべきである。例えば物体の形状を測定する手法としての二眼ステレオ法は比較的簡便な方法としての意義はあるが常に最適な手法でないことは明らかである。

人間は両眼立体視により立体感を知覚することはできても、対象の正確な寸法を得ることができないように、実応用においては人間よりも機械による方が容易なタスクも数多く存在する。寸法・反射パラメータなど物理量の精密な計測は人間が最も不得意な分野であると言われているし、人間の感覚としては備わっていない不可視光領域にも有用な情報が存在する。狭義の画像計測やマシンビジョン・ロボットビジョンと呼ばれる分野はこのような領域を主として扱い、生産システムなどの閉鎖的環境では大きな成果を上げている。これに較べ福祉支援機器や自動車の自動走行装置など社会環境における応用が立ち後れているのはタスクや環境がより複雑であるからであり、レーザ投光装置を用いた計測が安全性の低下を招くために応用できないからではない。言い換えると、人間がこのような環境下で両眼からの情報のみによって目的を達成できるからといって、同じ手法が最適であるという理由にはならない。より複雑な環境下で信頼性の高い計測を行うためには、さまざまな計測・処理手法を目的に応じて選択・融合することが必要であり、また作用装置など視覚システム以外の構成要素との関連についても考慮されなければならない。

画像センサはテレビ放送のために開発・改良されてきたセンサで、これは伝送された画像を鑑賞するための機器であり計測のためには最適でないことにも注意を払う必要がある。例えば、明るさが極端に異なる日中野外と室内では明度値を一定の範囲内に圧縮しなければ鑑賞に耐えないが、計測の観点からはセンサ出力が物理的な放射強度と対応付けられていることが望ましい。放送帯域幅の制限や人間の視覚特性などにより輝度と色相の情報量には差異が設けられているし、三原色もまた人間の色知覚にとって必要十分ではあるがマルチバンド計測による物体認識のためには不十分である。ビジョン研究の拡大の背景には画像入力装置の低価格化が大きく寄与しているが、純粋な計測機器として見た場合これらの機器は不十分であり、さらに高性能だが高価な機器が必要である場合が多い。この点でレンジファインダや遠赤外カメラ、マルチスペクトル画像センサ等が特に高価であるとは言いかたい。放送用画像センサで実現可能なタスクも数多く存在するが、これらから得られる画像の品質が不十分である場合には特殊画像センサの利用についても議論されるべきである。

### 1.1.3 画像計測の発展と新たな応用

前に述べたように、画像計測は生産設備における自動化の要請に応えて発展してきた。目視による計測・検査、ワークの位置決めなどを自動化するために、専用ハードウェア化を指向した二値画像処理 [Takano84] を初めとして、三次元構造の復元を伴わない画像計測が広く応用されている。レンジファインダを用いた鋼板プロファイルの計測など三次元計測を用いた例も見られるが、二次元画像処理に比べ少ない。三次元計測を用いなければ実現し得ない生産工程は、全く別の生産方式や機構・材質により計測の必要性をなくした工程に置き換えるほうが生産の安定性やコスト軽減につながることも多いことが実応用例の少ない理由である。多品種小量生産化による設備投資の削減や雇用確保の問題などもあいまって、画像計測による自動化ニーズは80年代にくらべ落ち着きを見せており、ビジョン研究の方向性は次なる応用に向けての準備段階という色合いが濃い。

一般生活環境に視覚システムを備えたロボットを導入することで福祉介護・家事労働の補助や一般オフィスの自動化を図ることはビジョン研究の夢であるが、タスクや環境の複雑さから実用に耐えうるシステムを構築することはまだ困難である。そこで、工場などの閉鎖的環境と一般環境の中間的な存在として、宇宙や原子力発電設備のための極限作業ロボットの研究が通産省大型プロジェクトを中心として精力的に行われ大きな進展を見た。現在ではやや一般環境にシフトした研究として高速道路上で自動走行を行うシステムの研究が盛んに行われており、画像による車線検出とレーザレーダによる距離計測を用いて車間距離を一定に保つシステムが実用化・市販に至っている。このように、画像計測応用はしだいに複雑な環境へと浸透していくだろう。

90年代に入り、一般家庭やオフィスへ高性能なコンピュータやネットワークが普及するにしたがって、実環境への作用を伴わない用途へのビジョン応用が急速に発展してきた。いわゆるマルチメディアの中核をなすものは画像情報であるが、この画像の解釈・生成の双方についてビジョン技術が応用されることで、知的活動や教育・娯楽ニーズを原動力にさまざまな研究が進んだ。代表的な例としては、映画やテレビ広告など高付加価値な映像制作のための画像計測が挙げられる。滑らかで自然なCGを制作するためには人体を計測することにより形状モデリングを行い、役者の動作・表情を入力することによりCGを動かすが、これらのほとんどは工業応用のために開発された技術が基盤となっている。また、画像からの三次元情報復元はCG生成とは逆向きの処理であるということなどを勘案すると、ビジョン技術は文字や図形の認識技術と共にメディア技術と不可分な存在であると言うこともできる。ネットワーク上を氾濫する情報にビジョン技術が応用されることで、画像から意味のある情報を選択・統合し脳内表現に帰着するというコンピュータビジョンが元来目指していたものが、我々が持つ両目の処理能力を拡張するための手段として開花するのではないだろうか。

## 1.2 本研究の意義と論文の構成

この論文では、対象物体が運動しているか、もしくは動きを与えた時に得られる画像列を解析することによって、物体の全周三次元形状と表面各点の反射率をモデリングする二つの手法

について述べる。また、対象の運動が既知でない場合には、モデリングに先立ち運動を求めておく必要がある。そこで、対象物体の6軸運動を実時間的に追跡する手法についても詳しく述べている。

通常、金型製作やCG生成などに用いる三次元形状の入力のためには、スリット光投影法に代表されるアクティブ計測法を用いることが多い。また、単一のセンサにより対象物体の全周形状を得るために、ターンテーブルなど対象物体の向きを変更する機構が必須となる。しかしこの機構を用いれば、対象物体の移動に応じて変化する画像を連続的に取り込むことで、動きによる形状復元 (shape from motion) [Ullman79] と呼ばれる原理により投光装置なしで形状に関する情報が得られるはずである。そこで2章では、いずれにせよ必要となるターンテーブルとカメラのみを用い、投光装置を廃した単純な機器構成により形状を復元する手法について述べている。ここでは形状計測結果をCG生成に用いることを主目的とし、対象物体とカメラの配置などに強い拘束条件を設けることにより受動的原理に基づく方法であっても実用的な精度で形状計測を行うことが可能であることを示す。

3章では、2章で述べた手法を拡張する形でCGモデルに必須である対象物体表面各点の反射率を求める手法について述べる。計測された全周形状がアクティブ計測に較べて劣るために、あえて法線ベクトルを用いた単純な逆演算に頼らずに、対象物体の運動に伴う明度の変化を解析することによって拡散反射率と法線ベクトルを同時に求める手法を提案する。対象物体の運動が単純であるために、点光源ではなく円環状の光源を用いることで安定に反射率を求めることができる。2・3章の独自性は、CG生成に用いる程度のモデルであれば、受動的原理に基づく手法であっても実用的な精度を得ることが可能なことを示したことである。

4章以降では逆に、距離画像センサ (レンジファインダ) を用いて動物体の運動を実時間的に追跡するシステムについて述べる。通常レンジファインダは濃淡画像センサにくらべ低速であり、動物体の計測にはそぐわないと考えられてきた。しかし、LSI技術によって飛躍的な計測速度の向上を図ったシリコンレンジファインダを用いることによって、濃淡画像を用いた場合に必要な奥行き情報の復元処理を行うことなく、モデルベース処理により直接的に対象の運動が求められることを示す。単純な処理による高速性を追求するために、濃淡画像を二値画像的にハードウェア処理することにより距離画像では得られない方向の運動についても追跡を行う。

4章の手法では、処理の単純さゆえに幾何学的な形状の物体しか追跡することができなかつた。しかし、レンジファインダでは対象物体表面のうち滑らかな部位についても距離情報を得ることができるのが利点である。そこで5章では、モデルから仮想的な距離画像を複数枚生成し、これらを総合的に用いることによりモデルベース的に自由曲面物体の6軸追跡を行う手法を提案する。距離画像を生成する過程では、CG描画ハードウェアによりレンジファインダをシミュレートする手法を導入することで高速化を図る。また、距離画像センサから得られる形状情報を対象物体の回転に応じてモデルへ統合していくことで、追跡に同時平行して対象物体の全周形状を得る手法についても述べる。

5章の方法では距離画像のみにより6軸追跡が可能であったが、球や円筒などの回転対称な物体については運動を検出できない自由度が存在する。そこで4章とは全く異なる手法により

濃淡画像を融合し、対象物体のテクスチャ情報を用いてより安定に追跡を行う手法を導入する。5章と同様にモデルから濃淡画像を複数枚生成する処理にCGハードウェアを用いることで高速化を図る。物体の運動に伴う陰の影響を除去するために、モデルには反射率パラメータを格納する必要があるが、像から反射率を求める処理にもまたCGハードウェアを用いることにより実時間的なモデリングを行う。追跡に伴いモデル情報が拡充していき、最終的に対象物体全周の形状と反射率に関するモデルが構築される。4章以降の独自性として、動物体の実時間運動計測にレンジファインダを用いることそのものが挙げられるが、それ以外にもCGハードウェアによる距離画像の生成・反射率の復元や、高速に距離画像同士の位置合わせを行う新手法を導入したことが挙げられる。また、アクティブビジョン研究の一環として、センサ計測方位を対象物体の方向へ制御する実験なども行った。

最後に7章では、動物体計測と動画像処理研究に対する本論文の成果と寄与を大局的に評価することで結論とする。

## 2章 受動的原理に基づく回転物体の形状計測

昨今の計算機の能力（速度、記憶容量）の目覚ましい向上により、一般ユーザレベルにおける CG アート・アニメーション制作が現実味を帯びてきた。しかし、パーソナルコンピュータ用のレンダリングツールが数多く市販されているにも関わらず、個人による CG 制作が広く普及することなく専門家による手工業的制作にとどまっている大きな理由の一つに、対象物体の数値モデルデータの取得が困難であることが挙げられる。実際に表現したい対象は幾何学的な形状をしていない不定形である場合が多く、CAD アプリケーションを用いた作業には限界がある。このような対象については計測により直接的にモデリングする方法が用いられ、計測装置も市販されているが、複雑な構成のために高価であり普及していない。ここでは極力簡便な装置構成とするため、レーザ等を用いずに受動的な計測手法によって対象物体の全周三次元形状を計測することを目的とする。

### 2.1 対象物体の運動に基づく形状計測 (Shape from Motion)

直接的にシーンの三次元情報を含んでいない二次元画像から対象物体の形状を再構築する手法の1つに、運動に基づく形状計測 (Shape from Motion) がある。これは、対象物体の運動に伴い変化する画像列を解析し、対象物体の形状を求める手法である。対象物体を剛体であると仮定すると、対象物体の移動を考える代わりにカメラが対象の周囲を運動していると考えることでも対象とカメラの相対位置を同一にすることができる。そのときカメラの運動軌跡上の2箇所の像を用いることは、両眼立体視により形状復元することとほぼ等価であると言える。また対象物体の運動が既知である場合、カメラの位置関係が既知であるとの等価である。両眼立体視や多点ステレオ法との相違点としては、

1. 像の数（カメラ位置の数）が多数であると仮定することができる
2. カメラ位置は連続的な軌跡上に存在する
3. 対象物体と照明の位置関係が変化する
4. 実際にカメラ位置を移動しない限り、カメラと照明の位置関係・背景は変化しない

などが挙げられる。ここでは特に1.に注目し、人間の視覚システムとの対比など CV 的興味が中心となる研究ではなく、多数の画像を統合してより信頼性の高い形状計測を行う手段としての Shape from Motion と位置づける。そのため対象の運動として回転運動を主に取り扱い、また画像の連続性に着目した手法を紹介する。

#### 2.1.1 運動が未知である場合の形状計測

対象物体の形状が既知であり、運動が未知である場合、その運動を求めることが追跡 (tracking) と呼ぶ。また、運動が既知である物体の形状を求めることが Shape from Motion に含

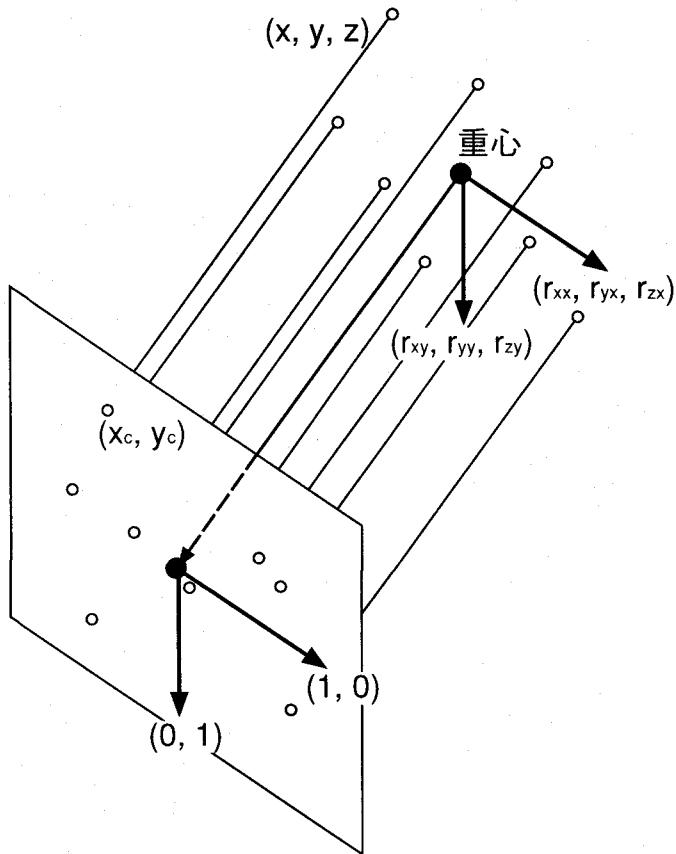


図 2.1 点群の正射影と重心の像

まれることは先に述べた。tracking / Shape from Motion 双方において、それぞれ前提としてきた既知情報をより少なくする研究が進められてきた。しかし、これらの双方が未知であっても、それらを求めることは可能である。Tomasi らは、形状と運動を全く同時に推定する Factorization Method と呼ばれる手法 [Tomasi92] を提案した。この手法ではシーンの投影として図 2.1 のように正射影が仮定されており、以下のように計算される。

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} r_{xx} & r_{yx} & r_{zx} \\ r_{xy} & r_{yy} & r_{zy} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (2.1)$$

この関係が画像上の  $P$  個の特徴点について連続的に  $F$  フレームの間観測されたとすると、それらをまとめて以下のように表すことができる。

$$\begin{bmatrix} x_{c11} & x_{c12} & \cdots & x_{c1P} \\ y_{c11} & y_{c12} & \cdots & y_{c1P} \\ \vdots & \vdots & \ddots & \vdots \\ x_{cF1} & x_{cF2} & \cdots & x_{cFP} \\ y_{cF1} & y_{cF2} & \cdots & y_{cFP} \end{bmatrix} = \begin{bmatrix} r_{xx1} & r_{yx1} & r_{zx1} \\ r_{xy1} & r_{yy1} & r_{zy1} \\ \vdots & \vdots & \vdots \\ r_{xF} & r_{yxF} & r_{zxF} \\ r_{xyF} & r_{yyF} & r_{zyF} \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \cdots & x_P \\ y_1 & y_2 & \cdots & y_P \\ z_1 & z_2 & \cdots & z_P \end{bmatrix} \quad (2.2)$$

この関係を  $W=RS$  と表すと、 $W$  のランクは 3 であるために、 $W$  を特異値分解することによって  $R$  と  $S$  とを同時に求めようとするものである。この手法では、それぞれの画像の原点は世界座標系における原点に対応しているとしているため、平行移動成分を扱うことができない。そこで、世界座標系における全ての点群の重心位置を原点と考えると、この原点の像は二次元画像上に分布する特徴点位置の重心位置となる。すなわち、行列  $W$  を構築する際にはフレームごとに全特徴点の重心を求め、その重心位置をそのフレームの全ての特徴点位置から減じることで仮想的に原点まわりの画像群に変換する。

この手法の第一の問題点は、正射影および疑似正射影（弱透視変換：weak perspective）下でのみ成立するということである。これを透視変換に拡張する試みは様々な研究者によって行われている [Quan96]。第二の問題点は、特徴点の対応付け問題と、その出現・消滅である。文献 [Tomasi92] でも既に一定の観測区間において常に観測される点のみを用いることで出現・消滅に対応する手法が提案されており、また対応付け問題については別に様々な研究が行われているが、常に“点”同士の対応が得られる必要があり、エッジ情報を用いることができないという問題点が依然残存している。

### 2.1.2 運動が既知である場合の形状計測

対象物体の運動に応じて連続的に画像を計測した場合、像面の 2 軸と時間軸の 3 次元的なデータが得られる。これを時空間画像と呼ぶが、その容量が膨大であるために全体を同時に処理することは難しい。そこで、通常は 1 画像のみで行われた処理を時間軸方向に統合するが、Bolles ら [Bolles87] はこの三次元空間について別の断面を取ることによって時間軸方向の連続性を重視した処理が可能であることを示した。

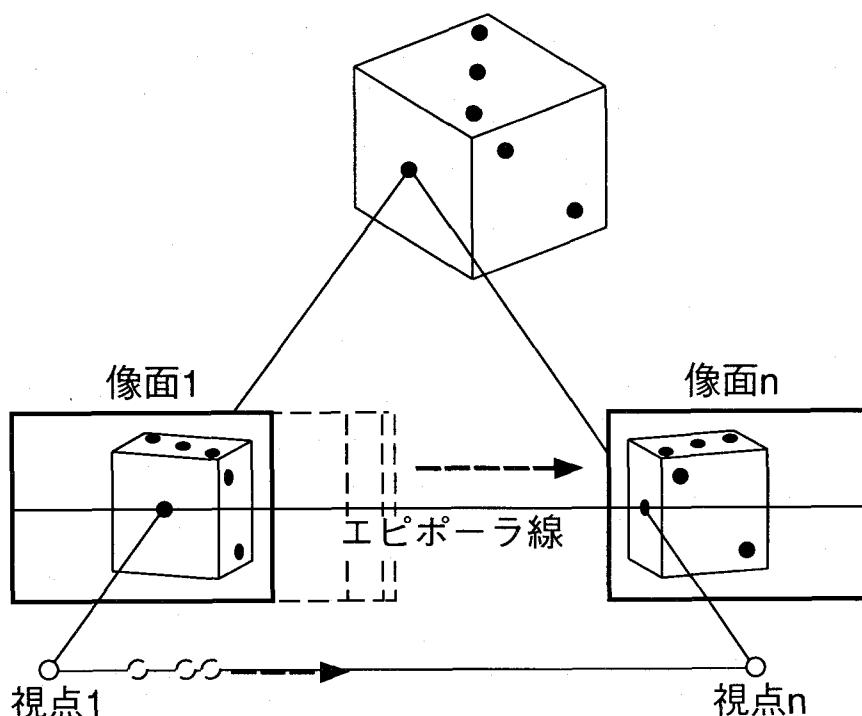


図 2.2 カメラ並進時のエピポーラ線

対象物体の移動を考える代わりに、カメラが対象の周囲を運動していると考えると、図2.2のようにカメラの運動が直線上にあるときにはエピポーラ線が全ての画像対について共通となる。つまり対象物体上の点は一定のエピポーラ線上に拘束されるため、時間軸方向にエピポーラ線内の画素を全て集めた二次元画像を構築する。これをエピポーラ平面画像 (EPI : Epipolar-Plane Image) と呼び、全ての特徴点はこの画像上で曲線をなす（図2.3）。カメラの運動が等速直線運動である場合には、EPI 上で特徴点は直線をなすため、特徴点の追跡問題は EPI 上の直線抽出問題に置き換えられる。本章で述べる手法はこれを回転運動に拡張したものである。

### 2.1.3 運動が1軸回転である場合の形状計測

対象物体の運動が回転運動であり、かつその軸が固定されていると仮定して問題を簡単化し、より安定に形状復元を行おうとする試みはさまざまな研究者により行われている。なかでも、[Szeliski91] では、カメラによる撮影を一般的な透視変換とし、対象物体の回転角と回転軸の位置を既知として全周形状が計測されている。対象物体を撮影するカメラの光軸は必ずしも回転軸に垂直ではないので、円運動する特徴点の軌跡は透視変換の結果画面上では橢円として現れる。図2.4のように、特徴点の三次元位置によってこの橢円の直径や位置は様々に重なり合いながら変化するため、2.1.3で紹介した EPI 解析を用いることはできない。そこでまず画像上から特徴点を抽出・追跡し、この軌跡から特徴点位置を抽出する際にカルマンフィルタを用いることで精度の向上を図った。しかし特徴点の追跡によるために、コーナー特徴量を含む対象でなければ用いることはできず、テクスチャのない物体やエッジ特徴量を主とする対象には応用が困難である。また、形状はそれぞれの特徴点の位置の集合として得られるため、表面形状モデルに統合する処理が別に必要となる。

対象物体のシルエットを解析することで形状を求める手法をはじめとして、回転物体の形状計測についての様々な手法が提案されている。まず、Vaillant はシルエットの位置のみならず、シルエットの画像面上での移動速度から、対象物体表面の曲率やオクルージョンの認識を行う手法を提案 [Vaillant92] した。ただし、この手法では回転軸に垂直な平面で断面をとったときに凹となる部分の形状を求めることはできない。また Zheng は同様の手法 [Zheng94] 以外に

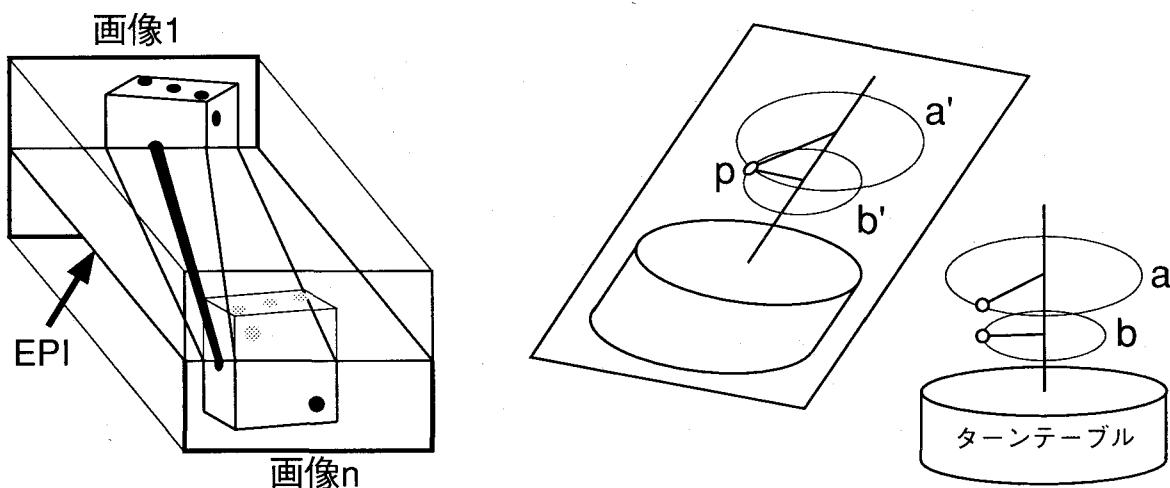


図 2.3 Epipolar Plane Image

図 2.4 一般的な回転体上の特徴点の軌跡。p の軌跡が a' 上であるか b' 上であるかは追跡しないと不明

も、表面の反射特性が鏡面反射であるような物体について、円環状の光源の映り込みの回転に伴う変化を解析して形状・法線方向を解析する手法 [Zheng96a]、対象物体表面全体に渡って一定の反射率を有する拡散反射物体の濃淡値を用いて形状を復元する手法 [Zheng96b]などを提案している。これらは全て光軸に垂直な方向から正射影を近似的に満たす撮像系により計測することを仮定しており、その点が本章で述べる手法との共通点である。

本研究では、受動的計測手法により安定に形状を復元することを目標としているため、ある程度環境を強い拘束条件下におく。すなわち、対象物体の回転は1軸に関する回転運動とし、その軸の位置と回転角は既知であるとする。撮像系は正射影を仮定し、回転軸に垂直な方向から撮影されるものとしている。この条件下でEPI解析を行い、エッジ特徴の追跡により全周形状を求める。また、シルエット解析による形状計測手法を融合することにより、テクスチャのない対象の計測も可能とする。

## 2.2 対象物体の回転による全周形状の計測

本節では、スリットレーザ光等を投影することなく、言わば受動的に対象物体の全周形状を計測する手法を説明する。受動的計測手法はしばしば不良設定問題となり、計測の信頼性が低下したり処理量が大きくなるという欠点があったが、ここでは対象物体の運動をカメラ光軸に垂直な回転運動に規定することで問題の簡単化を図っている。この方法では対象物体の回転軸を中心とした円筒座標系において形状が求められるため、各姿勢ごとの形状計測データを後から統合する必要がない。また、受動的計測手法によっているために、次章で述べる反射率計測と親和性が高い点が利点として挙げられる。本研究では、特徴点追跡を基本とする方法に時空間断面画像処理を利用することで簡単かつ高速な計測を可能とした。より一般的な対象物体の計測を可能とするために、シルエット像を利用する方法を融合することも試みた。

### 2.2.1 対象物体の回転と特徴点の軌跡

前節で述べたように、全周形状を得るためにターンテーブルを用いる方法がもっとも簡便である。ターンテーブルは計算機により制御されるため、対象物体の運動は一般的な回転運動として取り扱う必要はなく、対象物体の回転軸と回転角は既知であるとしても良い。しかし、現在多く見受けられる任意回転運動を取り扱った運動・形状計測は、計測精度や速度よりも解析手法的一般性を重視した研究が多く、そのままモデリングに用いるには最適であるとは言い難い。ここでは回転軸・回転角が既知であるメリットを最大限に利用することで高速かつ高精度に形状を計測することを目指す。

対象物体が図2.5のようにターンテーブルに固定され、計算機により任意の角度に回転することができるものとする。ビデオカメラをその光軸が回転軸と垂直になるように設置し、その焦点距離は正射影と見なすことができるぐらい長いものとする。物体形状は回転軸を中心とした円筒座標系により表現するものとし、計測すべき特徴点の座標を $(r, z, \theta)$ とする。ターンテーブルにより物体が $z$ 軸を中心に $\phi$ だけ回転した場合、その点は $(r, z, \theta + \phi)$ へ移動する。前述したように、カメラによる撮影は平行投影と仮定しているので、その点の投影面上の

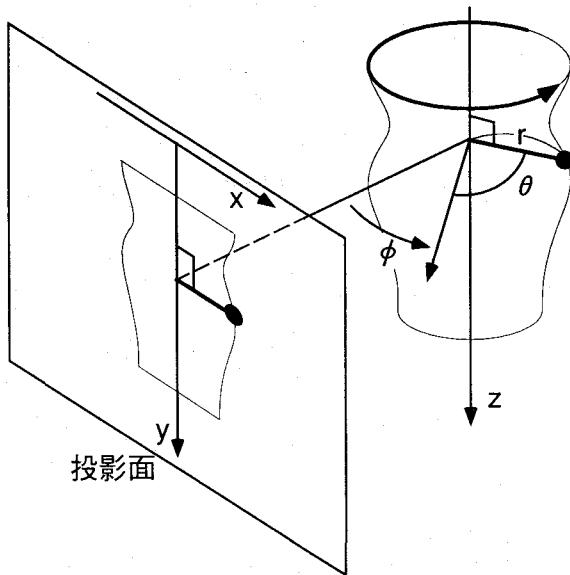


図 2.5 対象物体の回転とその正射影像

像は次式のように表すことができる。

$$\begin{cases} x = r \sin(\theta + \phi) \\ y = z \end{cases} \quad (2.3)$$

この式で、未知数  $r, z, \theta$  は一定なので、 $y$  は一定であり、また  $x$  は  $\phi$  に関する  $\sin$  関数となる。つまり、対象物体の像是対象物体の回転について回転軸の像に垂直な直線上を単振動し、その振幅が計測点の回転軸からの距離  $r$  に、位相が回転軸回りの方位角  $\theta$  に対応する。また、特徴点位置の  $z$  軸成分は、画像上では  $y$  座標と対応する。観測値  $(x, y, \phi)$  から未知数  $(r, z, \theta)$  を推定するには最小二乗法を用いるが、実際には回転軸の像是観測面上の  $y$  軸と一致しているわけではないので、そのずれの量（オフセット）を同時に推定する必要がある。すなわち

$$x = A \sin \phi + B \cos \phi + O \quad \text{ただし} \quad A = r \cos \theta, B = r \sin \theta \quad (2.4)$$

について残差平方和を最小化するような  $A, B, O$  を算出する。そのためには  $n$  個の計測値の組  $(\phi, x)$  について次式

$$\begin{bmatrix} \sum_n \sin^2 \phi & \sum_n \sin \phi \cos \phi & \sum_n \sin \phi \\ \sum_n \sin \phi \cos \phi & \sum_n \cos^2 \phi & \sum_n \cos \phi \\ \sum_n \sin \phi & \sum_n \cos \phi & n \end{bmatrix} \begin{bmatrix} A \\ B \\ O \end{bmatrix} = \begin{bmatrix} \sum_n x \sin \phi \\ \sum_n x \cos \phi \\ \sum_n x \end{bmatrix} \quad (2.5)$$

において逆行列を解くことで各パラメータが推定される。

### 2.2.2 時空間螺旋画像

特徴点の軌跡に最小二乗法を適用することで形状を求めていたため、高い計測精度を得るには極力長期間にわたり特徴点を追跡し、計測値の数を増すことが望ましい。安定に特徴点追跡を行うためには特徴点周辺の画素情報を利用することが多いが、そのためには各回転角ごとの画像をメモリ中に確保する必要がある。十分小さいステップ幅で対象物体を回転させると画像の総数が大変多くなるために、現在の計算機の容量では実現が難しい。しかし、前節で示したように特徴点の像は水平な直線上を運動することが分かっているので、実際に参照すべき画素データは1スキャンライン上にのみ存在する。そこで、回転角ごとに時系列に得られる画像群を一旦画像ディスク等のビデオ機器に記録し、計測を行いたいスキャンライン上の画素を逐次取り出し、それらを順次平行に並べることで関係するデータのみが含まれる一つの二次元画像を得ることができる。このように、横軸が空間軸に、縦軸が時間軸に対応する画像は時空間画像（画像の時系列的变化を追った三次元的な画像）の断面に相当するので、時空間断面画像と呼ぶ（図2.6）。対象物体の回転に関する時空間断面画像はちょうど螺旋のように見えるので、本論文ではこれを時空間螺旋画像と呼ぶ。一つの時空間螺旋画像からは、回転軸に垂直な平面で対象物体を切断した1断面の形状を得ることができる。

これまで時空間断面画像は、モーションステレオ等の分野で利用されてきた。例えばカメラを等速直線運動し、対象物体の像を連続的に取得した場合、物体上の特徴点軌跡は直線となり、直線の傾きが特徴点の奥行きに比例する[Bolles87]。この場合はカメラの移動軌跡に対して相対的に物体の位置・形状が計測されるが、対象物体を回転させた場合、特徴点の軌跡はsin曲線を描き、その振幅が回転軸からの距離に、位相が方位角に相当するため、物体の中心近くに座標系原点が位置するためにより高精度な計測を行うことができる。

### 2.2.3 時空間螺旋画像を用いた処理の流れ

通常の特徴点追跡処理では、まず画像から特徴抽出フィルタを用いてコーナー点等を抽出し、それらを連続するフレームの前後で対応付け、連続した運動としてとらえるという手順を踏む。

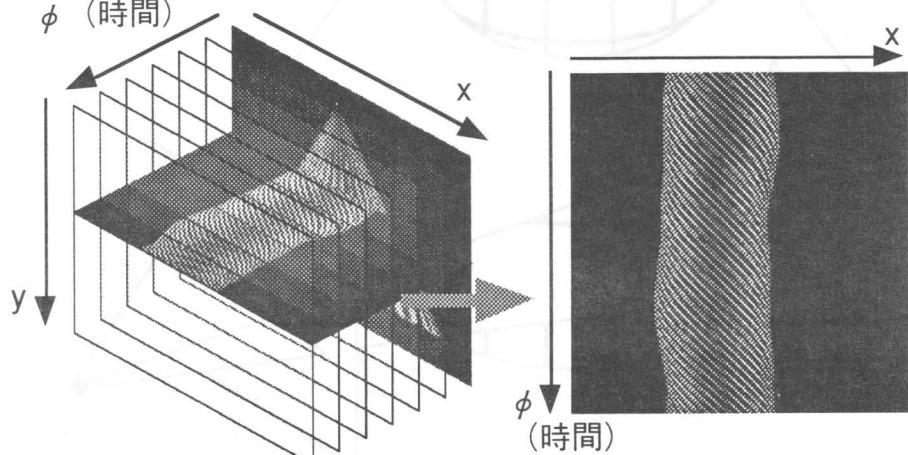


図2.6 時空間螺旋画像の合成

また、特徴点の出現・消滅の判定などの処理も必要となる。しかし、時空間画像では、特徴点はエッジ曲線として現れるので、それらの問題は曲線抽出処理に置き換えられる。前述したように、前後フレームの画素値を利用することでより高感度な特徴点抽出処理を行なうことが多いが、時空間螺旋画像は縦軸・横軸がそれぞれ空間・時間方向の連続性を保つ画像となっているために、通常の二次元微分オペレータを用いることで時空間的な連続性を利用した微分処理を行うことができる。

ここまでカメラによる撮影を正射影と仮定してきたが、実際にはレンズの焦点距離は有限であり、また装置の大きさにも制限を受けるので、撮影は透視変換となる。この場合、カメラの光軸から上下に離れた特徴点は直線上を運動せずに、像上を橢円形に運動する。特徴点の追跡処理は時空間螺旋画像を生成した後の段階で行い、時空間螺旋画像上のエッジはスキャンラインの観測する平面と物体上のエッジ曲線の交点の軌跡を反映していると考えられる。そのため、時空間螺旋画像上で追跡されたエッジ曲線は実際には一つの特徴点を追跡しているのではなく、物体上のエッジをわずかに上下しながら追跡していることになる（図2.7）。物体上のエッジが垂直である場合、この問題は無視することができるが、エッジの角度が水平に近づくにつれ真の特徴点軌跡からかけ離れた軌跡を描くようになる。そのため、得られた画像から直接時空間螺旋画像を生成するのではなく、得られた画像から鉛直に近いエッジをまず検出し、そのエッジ画像に関して時空間螺旋画像を生成するようとする。鉛直に近いエッジを検出するためにはソーベルフィルタを用い、 $45^\circ$ よりも水平に近いエッジ曲線はあらかじめ排除するようにした。

次に、時空間螺旋画像の微分処理の結果得られるエッジ特徴量を利用して、曲線追跡を行う。一つの時空間螺旋画像は対象物体の1回転分の画像から生成するので、回転開始時点の画像と回転終了時点の画像は連続している。そこで、螺旋画像の上端と下端がループ状に接続している。

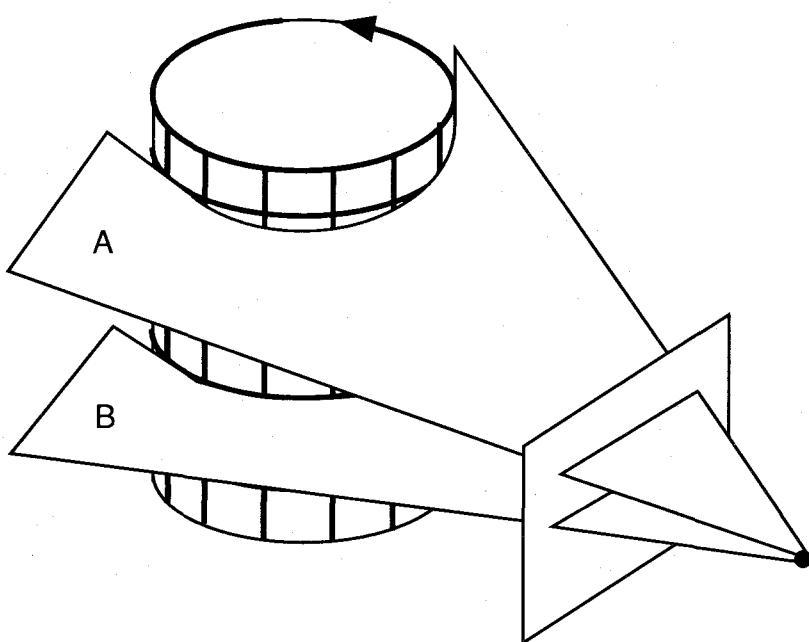


図2.7 透視変換によるエピポーラ拘束の不成立

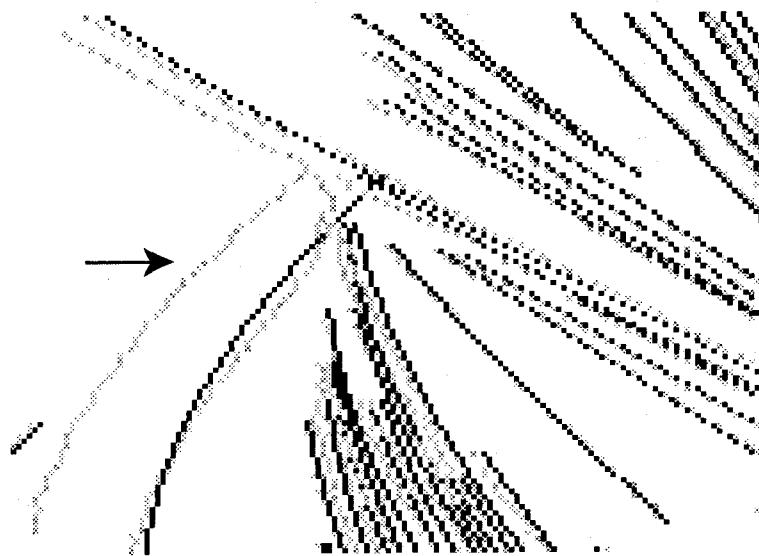


図 2.8 御追跡エッジの除去. 矢印部分が御追跡の発生したエッジであり,  $\sin$  曲線が推定されていない. (灰色: 抽出されたエッジ 黒色: 推定された  $\sin$  曲線)

るものとして処理を行う必要がある. また, 対象物体のシルエットにより生成されたエッジはここで除いておく.

最後に, 曲線追跡によって分離された個々の曲線について, 最小二乗法を用いて各パラメータを求める. 時空間螺旋画像には, 特徴点による曲線以外に, 物体自身の影や光源の写りこみなど, 実際には存在しない像が含まれることがある. また, 曲線追跡段階で誤った追跡結果が発生することもある. しかし, そのような回転運動と直接関係のないエッジ曲線は  $\sin$  曲線をなさない場合が多く, 最小二乗法でパラメータを推定する段階で残差平方和が大きくなることを利用して除去することができる (図 2.8).

一つの時空間螺旋画像からは対象物体の 1 断面の形状しか得られないため, 複数の断面について形状を求め, それらを統合して全周三次元形状を再構成する.

#### 2.2.4 シルエット像による形状復元

時空間螺旋画像に現れるエッジのうち, 最も外側のものは対象物体の輪郭によるものである. そこで図 2.9 のように, 左右のエッジ領域よりもさらに外側には対象物体が存在しないことを利用して, 丁度彫刻を制作するように物体存在領域を切り出すことで形状を復元する. この手法ではエッジのない領域に関して形状復元することができる半面, 凹部については形状を得ることができないため, 特徴点追跡による形状復元手法と組み合わせることで高い汎用性を得ることができる.

最後に, シルエットを利用した形状復元と, 時空間螺旋画像による形状復元とを併用するため, それらの計測結果を統合する必要がある. 一般に, 凸部ではシルエットによる形状復元の方が時空間螺旋画像による方法よりも高精度な結果を得ることができるため, シルエットにより形状が復元されなかった部分に時空間螺旋画像による形状を追加することで統合する. 図 2.10 に統合処理の流れを示す. 図中 A では, 時空間螺旋画像による形状 (□と■) のうち, □

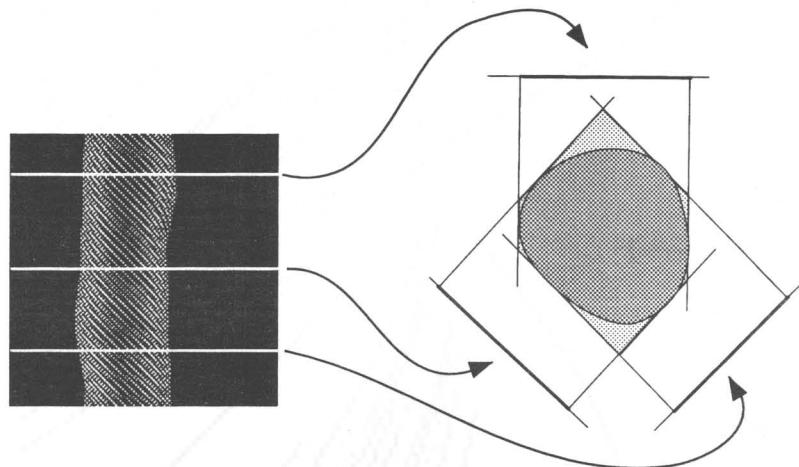


図 2.9 シルエット像による凸形状復元

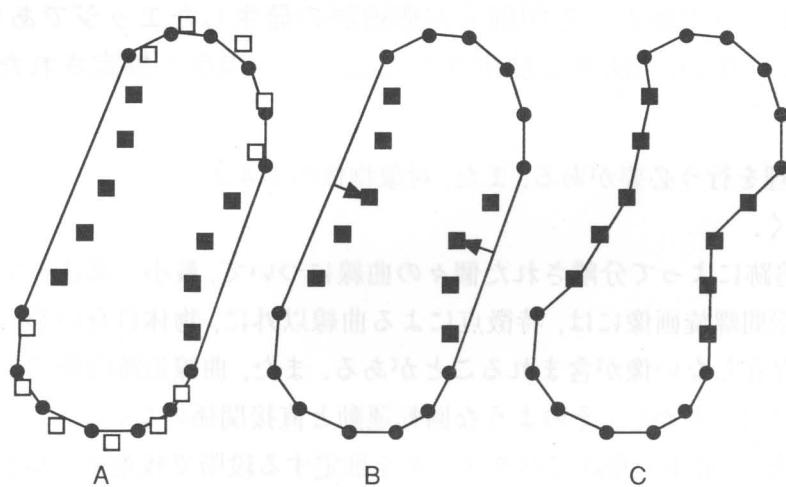


図 2.10 シルエットと時空間螺旋画像による形状の統合

がシルエットによる形状計測点（●）の近傍にあるため削除される。次に、B のようにシルエットによる輪郭（閉曲線で表示）から最も離れた位置にある■を矢印のように輪郭へ統合し、統合されていない■がなくなるまで繰り返す。結果として、C のように統合された輪郭形状を閉曲線として得ることができる。

### 2.3 実験

この節では、2.2 で述べた全周形状の計測手法に基づき実際に計測装置を製作し、計測実験を行った結果を示す。また、計測処理の中間段階で用いる画像も併せて示した。ここで示す最初の 2 例ではシルエット像による形状復元手法を用いず、時空間螺旋画像のみによる形状計測手法に関する評価を行った。まず、コントラストのはっきりした既知の形状を計測することで測定精度の評価を行った。次にコントラストが弱く光沢のある物体に関する計測を行い、実応用が可能であることを示した。最後の例では、時空間螺旋画像とシルエット像による形状計測

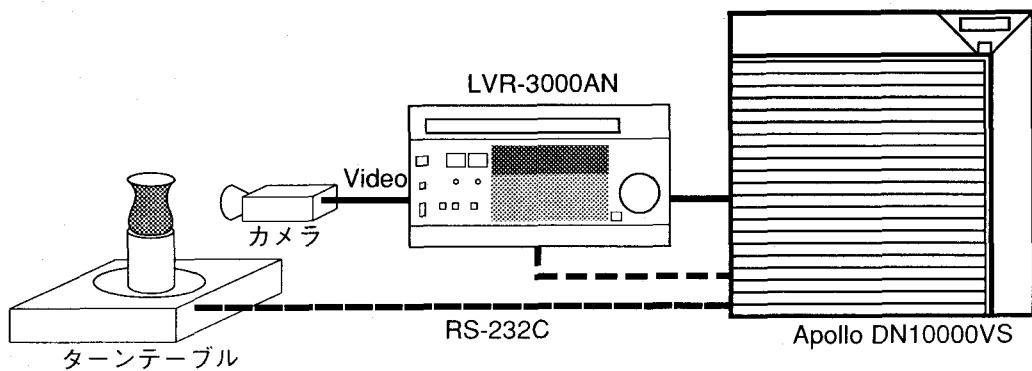


図 2.11 形状計測の機器構成

の結果が互いの弱点を補いあつてゐることを示す。

### 2.3.1 実験装置の構成

2.2で述べたように、本研究で提案する計測手法ではターンテーブルや光画像ディスク装置などいくつかの特殊な装置を用いる。実験装置の接続図を図2.11に示す。ターンテーブルと光画像ディスク装置（LVR-3000AN）はそれぞれRS-232Cにより制御され、対象物体の回転とその像の記録・再生を行う。ターンテーブルは $0.036^\circ$ の角度分解能（1回転10,000ステップ）を有するが、今回は $0.72^\circ$ 刻み（1回転500ステップ）で回転した。光画像ディスク装置は画像の記録・検索を0.5秒以内で行うことができる。ホスト計算機Apollo DOMAIN DN10000VSに搭載した画像取り込みボードはフィルタ処理等をハードウェアにより高速に実行する機能を持つが、今回は利用していない。

### 2.3.2 円筒の計測による精度評価

円筒形の対象物体の形状を測定し、真円度を求めて計測精度を評価した。ここでは安定に特徴点の追跡を行うため円筒にコントラストのはっきりした縦縞を配した物体を用いた。

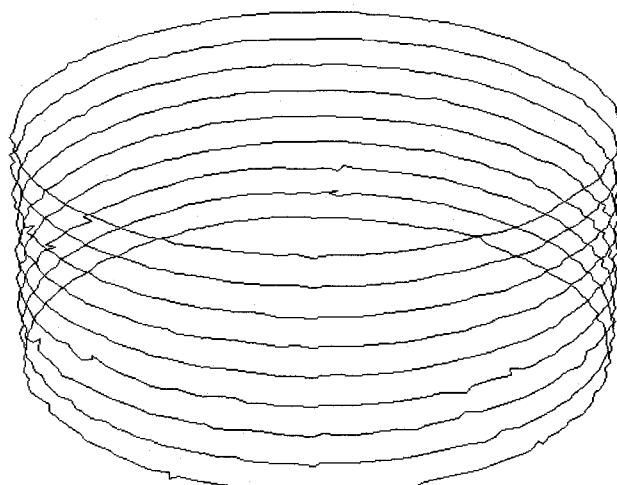


図 2.12 円筒の形状復元結果

測定に用いた画像の大きさは 512\*480 ドットで、1 回転を 500 ステップに分割して時空間螺旋画像を生成した。全体で 1325 個の特徴点の座標が得られ、半径  $r$  の誤差の標準偏差は半径の平均の 0.9% だった。また、ドット単位における誤差の標準偏差は 0.88 ドットであったことから、特徴点軌跡を長期間追跡することによってサブピクセル精度が得られていると考えられる。また、この測定精度はこの装置の限界であると考えられる。復元された形状を図 2.12 に挙げる。

### 2.3.3 実物体の計測例

比較的コントラストが弱く、かつ光沢を持つ物体として恐竜の模型をとり、自然照明のもとで測定した。測定対象の写真を図 2.13 に示す。この物体は複雑に入り組んだ形状をしており、また腹部にあたる所はほぼ完全に拡散反射であるが、背中に当たる部分は上葉がかけられているために光沢を有する。この物体を 1 回転 500 ステップで回転させ、下方 1/5 の部分から生成した時空間螺旋画像を図 2.14 に示す。中央やや右寄りの部分などに物体自身の光沢による光源の写り込みが見られる。

抽出された時空間螺旋画像上の曲線群に当てはめられた  $\sin$  曲線を図 2.15 に示す。また、求められた断面形状を図 2.16 に示す。得られた特徴点の座標同士がある程度以上近い場合に線分で接続して表示している。追跡特徴点となるエッジがない部分で形状が復元されていない部分が見受けられる。得られた複数の断面形状を復元した断面数だけ積層することで生成した 3 次元形状を図 2.17 に示す。影になる部分や模様のない部分で形状データに欠落が見られるが、そのほかの部分では正しく形状が得られている。また、頭の部分に、回転軸から同一の方向に複数の半径が得られた部分があるが、そのような凹形状が測定可能な点も本手法の特徴である。



図 2.13 計測対象（恐竜の模型）の外観

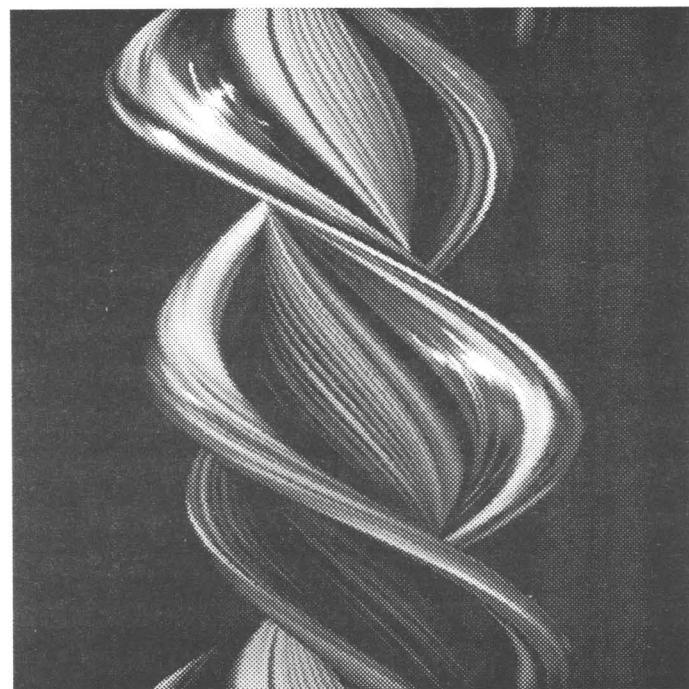


図 2.14 時空間螺旋画像

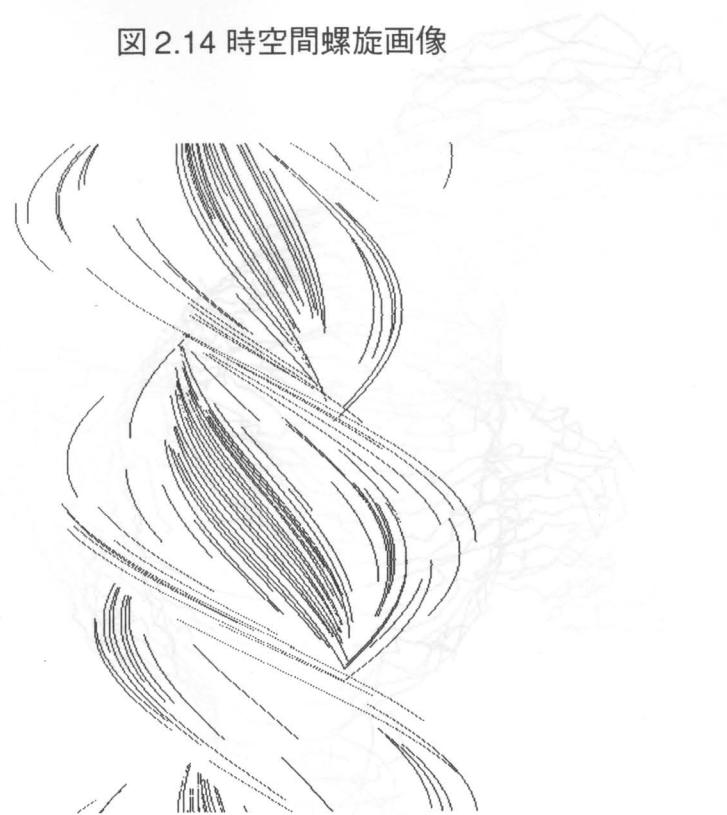


図 2.15 推定された sin 曲線

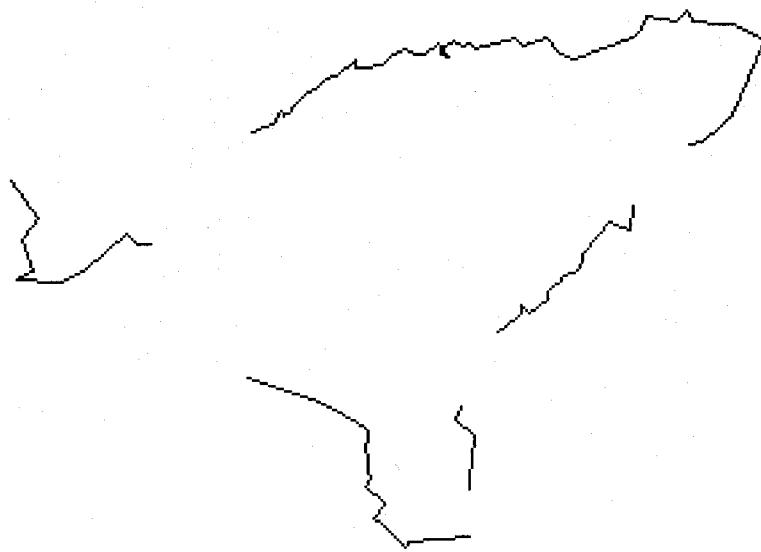


図 2.16 復元された断面形状

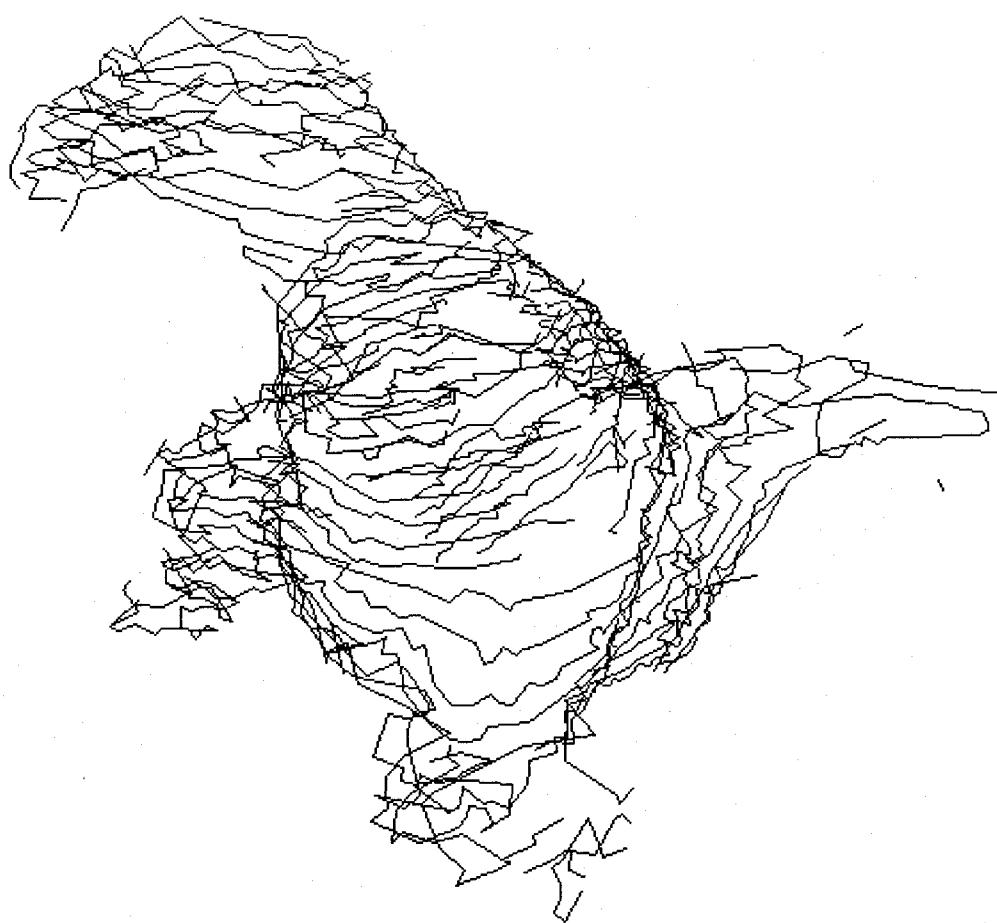


図 2.17 復元された恐竜の模型の全周形状

### 2.3.4 シルエット像による形状復元

ここまでに示した時空間螺旋画像による形状復元に加え、シルエット像の解析による形状復元を併用した実験結果を示す。図 2.18 に示すような陶器を用いて実験を行った。シルエットによる形状復元を行うことができた領域は時空間螺旋画像による形状を行わず、凹部分の形状復元を補うためだけに時空間螺旋画像を用いた。得られた計測点についてそれぞれいずれの方法により求められた点であるかを示したのが図 2.19 である。取っ手や注ぎ口により隠されてしまう部分が時空間螺旋画像により形状復元されていることがわかる。



図 2.18 実験に用いた陶器

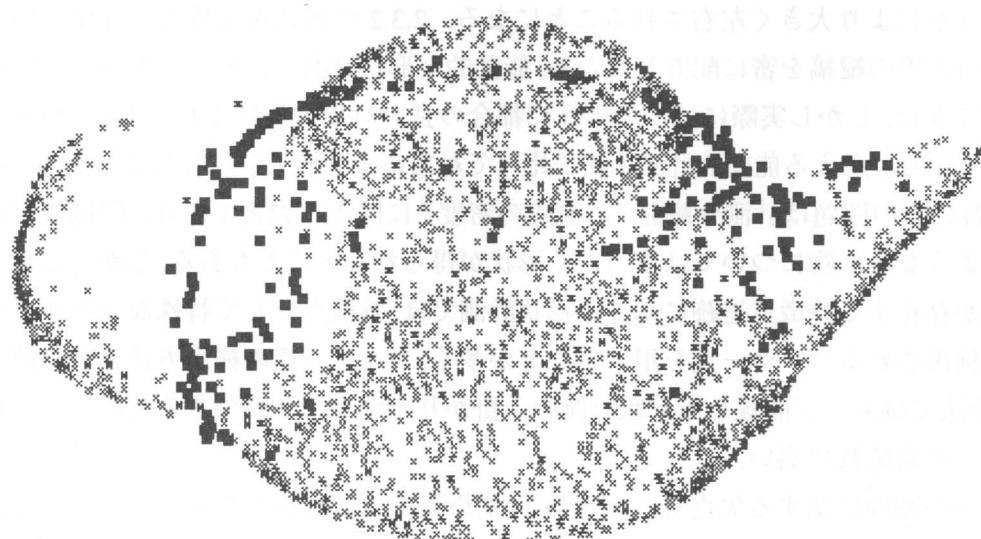


図 2.19 2つの形状復元手法と計測点の配置

×：シルエットによる計測点

■：時空間螺旋画像による計測点

## 2.4 考察

本研究では対象物体の運動とカメラの配置の関係をあらかじめ規定することにより高速・安定に、かつ簡便に対象物体の全周形状を計測することを目的とした。ここでは本研究の手法をCG制作の全周形状入力として用いることを前提にその計測特性・計測時間・計測装置の構成を中心に考察する。

### 2.4.1 時空間螺旋画像による計測結果について

対象物体をCGにより表現するために用いる三次元形状計測は、金型製作などの工業応用のための計測と比較して絶対精度が必要とされることが少ない。なぜなら、計測されたデータは光源計算や透視変換などの光学的計算を施すことで二次元画像を再構成するための中間データであり、最終的な出力の品質そのものが評価の尺度であるからである。そのため、他の手法との比較のような相対的評価を別として、絶対的な計測結果の精度や分解能に対する評価はその計測結果を用いて再構成されたCG画像をもとにして評価すべきであると考えられる。本研究は三次元計測のみを問題として取り扱っているが、CG制作システム全体としての性能を向上させるためには本研究で提案した手法に適合する計測データ後処理やCG計算モデルなどにも改良の余地が残されている。ここでは計測された三次元形状に対してのみ評価を行うが、CG計算において反射モデルと形状モデルはそもそも不可分な性質のものであるため限られた絶対評価しか行うことができない点を指摘しておく。

#### a. 計測点の密度について

時空間螺旋画像から形状を復元する際には画像中のエッジ検出を行っているため、おのずと対象物体表面における反射率の変化部分や形状としてのエッジに関して三次元形状が計測される。そのため得られる三次元形状の分解能は対象物体表面に検出の容易なエッジが多数存在するかどうかにより大きく左右されることになる。2.3.2の形状復元精度の評価で用いた円筒は故意に白と黒の縦縞を密に配置したため、能動的な形状計測と比較しても劣らない結果を得ることができた。しかし実際にはこのように都合の良い対象物体はまれであり、積極的に特徴を付加していると言える能動的計測手法に比べて対象により結果が左右してしまうのは避けられない。特にこの問題は計測分解能（計測点の密度）に関して顕著であり、印刷などが施されていないような人工物についてはまったく形状が得られないこともある。しかしこれまで多数の研究例が存在する特徴点追跡に基づく形状計測では、特徴点として特殊なフィルタを用いるなどして検出されるコーナー点を用いることが多く、コーナー点の検出方法も多数研究されている。本手法ではエッジ曲線をも形状計測の手掛かりとすることができるためより一般的な対象についての適応性は高い。

特徴点の密度に関する欠点も、実際のCG制作現場では利点となる可能性もある。なぜなら、計測された特徴点の分布は対象物体の特徴（反射率と形状のエッジ）が反映されていると考えられるため、形状の修正や変形、また反射率・色の設定を人手により行う場合の境界線として用いることができるからである。また、単に計測された特徴点を接続することで物体形状をワ

イヤフレーム表示する場合はワイヤが計測物体のエッジに対応するために自然な表示が可能であることが特徴として挙げられる。

#### b. 計測精度について

各計測点の位置精度については2.2で述べたとおり、計測画像の解像度により制限を受けていると考えられる。スリット光とターンテーブルを用いたような能動的計測についても同様に撮影系の解像度により精度は制限を受ける。また、計測される値は回転軸上的一点を原点とした座標系において得られるため、カメラ平行移動による計測のようにカメラ系を中心とした形状を計測する手法に比べ高い精度を得ることが容易である。しかし、対象物体の撮影を正射影であると仮定しており、近似的に正射影を実現するためには物体の大きさに比べカメラと物体の距離を大きくとる必要がある。透視変換が正射影からかけ離れるに従い2つの問題が生じる。その一つは特徴点の軌跡がsin曲線ではなくなるという点であり、もう一つは時空間螺旋画像上のエッジが対象物体上的一点の像ではなくなるという問題である。特に後者が実際の計測に大きく悪影響をもたらしており、水平に近いエッジを多く含むような対象物体では大きな誤差を生じてしまう。そのため、時空間螺旋画像を元画像から直接生成するのではなく、回転軸とエッジのなす角が45°より小さなエッジのみをソーベルフィルタを用いて検出することで、誤差を生じる可能性のある特徴点を排除するようにした。

#### 2.4.2 シルエット像の解析による計測結果について

計測対象に対する適応性を高めるため、本研究ではシルエット解析による形状計測を融合することも試みた。この方法は簡単であるにも関わらず非常に安定・高精度に対象物体の形状を計測することができる。しかし2.2でも述べたように、回転軸に垂直な平面に関する対象物体の断面が凹形状である場合凹部分に関する形状を得ることができない。そのため、凹部分に関してのみ時空間螺旋画像による形状計測を併用することで、より対象物体を選ばない計測システムを構築した。

シルエット像による形状復元のもう一つの利点として、計測点の三次元位置がそれぞれ独立に得られるのではなく、計測点の連結関係も同時に得られる点である。大きな石膏の塊から彫刻を削り出すように、十分大きな正方形から物体が存在しない領域を削り出すことで多角形として形状を得ることができるため、最終的には線分の連結として形状が得られるからである。こうして得られた凸多角形に時空間螺旋画像による形状計測結果を追加する形で凹部分の形状を反映するため、CGモデルとして必要な面の集合としての三次元モデルを得ることができる。

#### 2.4.3 計測時間について

対象物体による形状計測は大きく二つの段階に分かれている。まず対象物体を回転させその像を光画像ディスク装置に記録し、つぎにディスクから断面画像を合成・解析して特徴点の位置を算出する。まず撮影段階では、対象物体を $0.72^\circ$ づつ回転させ、1回転分の画像を光画像ディスクへ1フレームづつ記録する(1回転は500フレームである)。1フレームの記録時間は約1秒であるため、約8分で撮影は終了する。

一つの時空間螺旋画像から1断面の形状を復元する時間は特徴点の個数により大きく左右さ

れるが、約5～10秒で処理が終了する。しかし、光画像ディスクから時空間螺旋画像を生成する処理がデータ量が大きいため長時間かかる。そこで、一回に数十の時空間螺旋画像をまとめて生成しておくことで計算機のディスク容量と処理時間のバランスを向上している。もちろん、全体の処理時間は計測する断面数によるが、100枚程度の断面について計測した場合、撮影・処理をすべて行っても1時間前後で全周三次元形状が得られる。

#### 2.4.4 計測装置について

計測装置として我々は画像の取り込みが可能なワークステーションとビデオカメラ、計算機から制御可能なターンテーブルおよび光画像ディスクを用いた。このうち光画像ディスクは研究の効率化を目的に導入したものであり研究初期の段階では用いておらず、原理的にもなくてもよい機器である。ターンテーブルは一般的に普及している機器ではないが、たとえばレコードプレーヤのように一定速度で回転するような機器を用い、マーカー等を用いて回転角度を計測により求めることでも計測は可能であると考えられる。また、画像の取り込みについてもビデオカメラや画像キャプチャの行えるパソコンは急速に普及しているため、工夫すれば非常に簡単に計測システムを構成することは可能であると考えられる。

## 3章 円環光源を用いた回転物体の反射率・法線ベクトル計測

CGによって物体を合成表示するためには、物体の形状に関する情報とともに、物体の色や陰影情報を決める反射率に関する情報が必要となる。自然に存在する物体のみならず人工物においても、物体の表面全体で反射率が同一であるということはまれであり、自然なCG表現のためには対象物体表面の各点における反射率を設定することが望ましい。これらは今まで手作業により入力され物体形状と対応付けられてきたが、実対象から直接取り込むことが省力化・品質の向上の両面から切望されてきた。ここでは第2章で述べた形状計測と平行して反射率計測を行う手法を解説する。また、CGにおいて反射率と組み合わせて放射強度の計算に用いられる法線ベクトルの算出を同時に行う。

### 3.1 各種の反射率計測法

物体の反射率を計測するためには、反射率計と呼ばれる専用の装置を用いる。しかしこの装置では通常一点に関する計測しか行うことができない。ここでは物体表面の各点において同時に・独立に反射率を計測する手法をいくつか挙げる。

#### 3.1.1 一様光を仮定する方法

対象物体を特に一定の方向から照明せず、どの方角からも一様に照明されていると仮定すると、拡散反射率が画像強度に比例することを利用して計測する方法である。しかし、完全な一様光を実現するのは困難であるばかりでなく、対象物体が凹部を持つ場合、その部分が実際よりも小さな反射率として計測されてしまう欠点がある。

#### 3.1.2 法線ベクトルと光源方向を既知とする方法

物体の形状計測結果を用いて計測面の法線ベクトルを計算し、光源方向との角度を計算することによって反射率を復元する。大槻はスリット光切断法による形状計測と組み合わせることで鏡面反射率も算出した [Otsuki93]。計測された形状データから法線ベクトルを算出し、対象物体の回転にあわせて計測点をトラッキングすることで対象物体表面各点について多方向から照明された反射光を計測する。反射率パラメータの推定は形状を記述する全ての三角パッチに対して行われるのではなく、模様の解析により領域分割された各領域ごとに行うことで処理の効率化を図っている。この方法では精度よく法線ベクトルが計測されている必要があるが、図3.1のように形状計測を密に行うほど法線ベクトルの誤差が大きくなるというトレードオフが発生し、形状や法線ベクトルに対して平滑化などを行う必要が生じることが多い。

#### 3.1.3 照度差ステレオ法

照度差ステレオ法 [Woodham89] [Hayakawa93] はカメラと物体の位置を固定し、光源を異なる

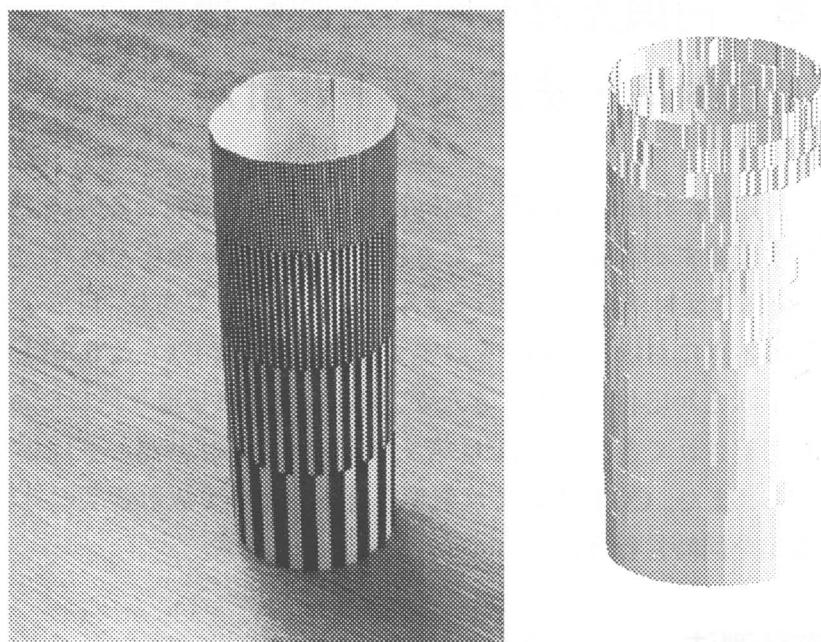


図 3.1 計測点の密度と法線ベクトル精度の関係。上方ほど縞模様を密に配置した円筒（左）の形状計測結果を一定の反射率でフラットシェーディング表示（右）した。計測点の密度が高いほど法線ベクトルの精度が落ちることが分かる。

る 3 方向から照射することにより、3 枚の陰影画像から物体の面の向きを推定する方法である。

まず、図 3.2 に陰影画像生成の基本的な原理である光学反射の計算モデルを示す。方向  $L$  からの入射光は物体表面で反射され、画像面上で画像強度  $I$  として観測される。この画像強度  $I$  は、面の法線ベクトル  $N$ 、光源方向ベクトル  $L$ 、観測方向ベクトル  $V$  の相対的な関係および光源照度  $i$ 、そして面の反射特性により決定される。面の反射特性を記述する反射率関数を  $\tau$  で表せば、

$$I = i \cdot \tau(N, L, V) \quad (3.1)$$

と書くことができる。ここで、光源は物体から十分離れた点光源であるとし、観測者の位置は物体から十分離れていると仮定すれば、光源照度  $i$ 、光源方向ベクトル  $L$  および観測方向ベクトル  $V$  は画像平面上でそれぞれ一様な値となる。そして、3 方向から照射する光源の方向ベクトルが同一平面上にはないという条件のもとで、観測方向ベクトル  $V$  と三つの光源方向ベクトル  $L$  の相対的な関係およびそれらの光源照度  $i$  が既知であれば、未知の  $N$  を求めることができる。これが、照度差ステレオ法の基本的な原理である。

また、物体表面が反射率  $d$  の拡散反射面であると仮定することによって、この手法はさらに簡単化される。この場合、観測者の位置は物体から十分離れているという仮定は必要がなくなる上、反射率関数は簡易な次式で置き換えられる。

$$I(x, y) = d(x, y)i(N(x, y), L) \quad (3.2)$$

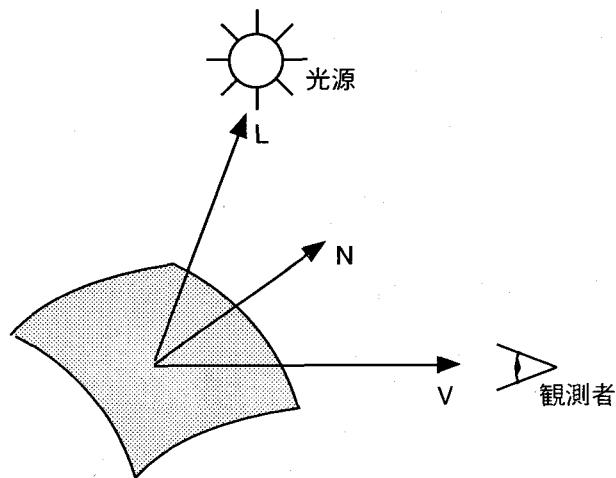


図 3.2 対象物体と光源・視点の位置関係

なお、 $(x,y)$  は画像平面座標を表している。ここで 3 方向の光源照度が一定もしくはそれらの比が既知であるとし、その値で画像強度を正規化したものを  $\Gamma(x,y)$  とすれば、

$$I'_k(x,y) = d(x,y)(N(x,y) \cdot L_k) \quad k=1,2,3 \quad (3.3)$$

となる。さらに、これを行列表現で書き直せば、

$$\mathbf{J}(x,y) = d(x,y)(N(x,y))^T \mathbf{M} \quad (3.4)$$

となる。但し、

$$\begin{aligned} \mathbf{J}(x,y) &= [I'_1(x,y) \quad I'_2(x,y) \quad I'_3(x,y)] \\ \mathbf{M} &= [L_1 \quad L_2 \quad L_3] \end{aligned}$$

である。ここで、三つの光源方向ベクトルが同一平面上にはないという仮定より、行列  $\mathbf{M}$  の逆行列が存在することは明らかである。したがって、式 3.4 に基づいて  $d$  と  $N$  について解くことができる。まず、 $N$  が単位ベクトルであることから、

$$d(x,y) = \|\mathbf{J}(x,y)\mathbf{M}^{-1}\| \quad (3.5)$$

となり、続いて  $N$  について解けば

$$N(x,y) = \frac{1}{d(x,y)} \mathbf{J}(x,y)\mathbf{M}^{-1} \quad (3.6)$$

となる。もちろん、この手法が使えるための前提条件として、観測している面上の点がどの光

源に対しても陰に入らないことが必要である。

照度差ステレオ法では、対象物体とカメラを固定し、光源を移動することで対象物体各点の法線ベクトルと拡散反射率を算出する。しかし、対象物体が完全拡散反射である場合、対象物体の放射強度は物体とカメラの位置関係には依存しない。そこで、カメラと光源を固定し、対象物体を運動させた場合にも照度差ステレオ法を用いることができると考えられる。但しこの場合、カメラと対象物体各点の位置関係が変化するため、対象物体上の計測点の像の軌跡を追跡する必要が生じる。本論文では、対象物体の形状が既知であるとしたときに、対象物体を回転させることで物体表面各点の拡散反射率と法線ベクトルを求める手法について解説する。

### 3.2 反射率と法線ベクトルの計測原理

この節では、光源とカメラを固定し、物体を回転させることによって各点の拡散反射率と法線ベクトルを計測することを目的とする。物体から反射される光の強さ（放射強度）は、完全拡散反射面の場合、面の反射率と、光源方向と面の法線方向の間の角度によって決定されるため、一方の推定精度が他方の推定精度に影響を及ぼす。そのため、これらのパラメータは同時に推定することが望ましい。そこで、照度差ステレオ法を応用することで法線ベクトルをあらかじめ計測することなしに対象物体各点の拡散反射率を推定する原理を示す。

#### 3.2.1 計測点の追跡

図3.3に、形状計測を含めた全体の処理の流れを示す。形状計測の結果から時空間螺旋画像

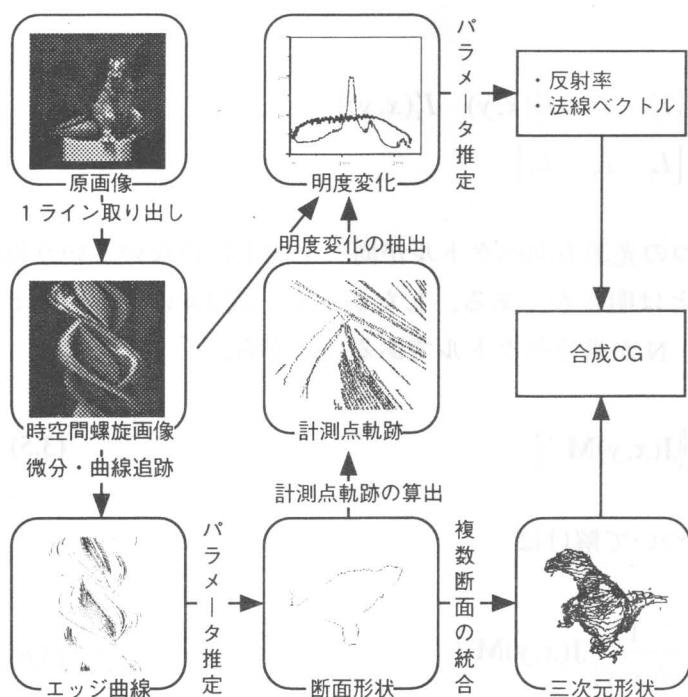


図3.3 形状・反射率計測の処理の流れ

上の計測点の軌跡が求められ、軌跡上の画像強度変化を抽出し解析することにより反射率を推定する。形状計測の際の計測点は時空間螺旋画像上のエッジに対応し、それらは反射率の変化点と考えられるため、反射率の推定は時空間螺旋画像上のエッジとエッジとの間で求めることが望ましい。実際には隣同士の計測点座標の中点を求め、その軌跡を時空間螺旋画像上において追跡することでエッジ間の明度変化を取り出す。この処理はよりもなおさず反射率の変化によって計測領域のセグメンテーションを行っていることに他ならず、均一に反射率を計測する場合と比較して大きく計算量を減少させることができる。人工的な物体で単純な模様を持つ物体の多くは塗料により着色されており、塗料によって反射特性が変化することが予想されるので、この仮定は必ずしも不適切ではない [Otsuki93]。形状計測段階においては特に一方光からの強い照明を行っているわけではないので、鏡面反射による影響は少ないと考えられ、その場合画像強度と拡散反射率の相関は高いと考えられるため自然材についても問題はないと考えられる。

### 3.2.2 円環光源による照明

図3.4に示すように、対象物体の回転軸を含み、カメラの光軸と直行するような平面内に円環状の光源を配置する。図3.5において、対象物体は凸形状をしていると仮定すると、法線ベクトル  $N$  を  $xy$  平面上に射影したベクトルを  $N'$  としたとき微小面を照明する光源の方向は  $N'$  からの角  $\psi$  が  $-\pi/2$  から  $\pi/2$  の範囲の部分のみである。微小面の拡散反射率を  $d$ 、光源の単位

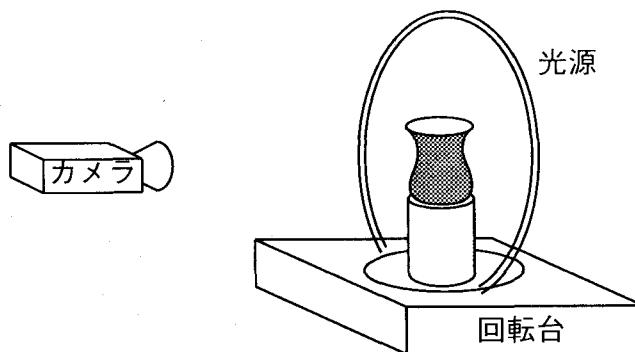


図3.4 対象物体とカメラ・回転台・光源の配置

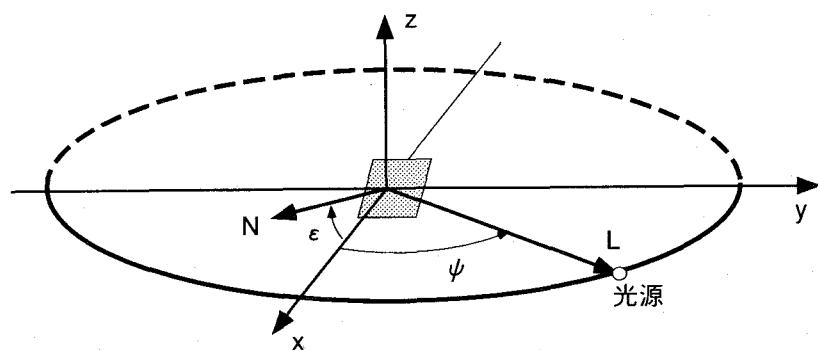


図3.5 円環光源による照明と法線ベクトルの方向

角あたりの光源照度を  $i$  とすると、微小面の放射強度  $I$  は単位角あたりの光源について積分することで求められる。

$$I = \int_{-\pi/2}^{\pi/2} d i (N \cdot L) d\psi \quad (3.7)$$

この式において、微小面の法線ベクトル  $N$  と光源方向ベクトル  $L$  はそれぞれ

$$\begin{aligned} N &= (\cos \varepsilon, 0, \sin \varepsilon) \\ L &= (\cos \psi, \sin \psi, 0) \end{aligned} \quad (3.8)$$

であるため、 $\psi$ に関する積分を計算することによって

$$I = 2 d i \cos \varepsilon \quad (3.9)$$

と簡略化することができる。この式の意味するところは、円環光源による照明は微小面の法線ベクトル  $N$  を円環光源を含む面へ投影した方向に単位角あたりの放射強度の 2 倍の照度の光源が存在するのと等価であるということである。これにより照度差ステレオ法の前提条件である、三つ以上の同一面内にない同一照度の光源を配置したのと同様の効果を得ることができる。それに対し、单一光源を用いた場合では、光源の位置と法線ベクトルの向きの組み合わせによっては悪条件となることがある。

### 3.2.3 放射強度の変化とパラメータ推定

図 3.6において、光源の形状は  $Z$  軸に対して対称なので、計測する微小面の放射強度は光源の含まれる平面  $XY$  と、微小面の法線ベクトル  $N$  のなす角  $\varepsilon$  によって決定される。つまり、円環状の光源による反射は、 $N'$  方向に  $2i$  の点光源があるのと等価である。ここで 3.9 式の  $\cos \varepsilon$  が  $N'$  の長さであることを利用すると、次のように展開することができる。

$$I = 2 d i \sqrt{1 - \cos^2 \beta \sin^2 \alpha} \quad (3.10)$$

この式で、 $d, i, \beta$  は定数であるが、 $\alpha$  は対象物体の回転に応じて変化する。微小面の回転軸からの方方位角を  $\phi$  とし、方位角からさらにどれだけ法線ベクトルが回転しているかを表す変数を  $\gamma$  とすると、 $\alpha = \phi + \gamma$  となる ( $\phi$  は既知、 $\gamma$  は未知である)。ここで、明度の二乗値について最小二乗法を適用する。なぜなら、凹形状の部分を計測する場合、明度の低いところでは物体自身の陰に入ることがあり、その影響を減少させるためである。

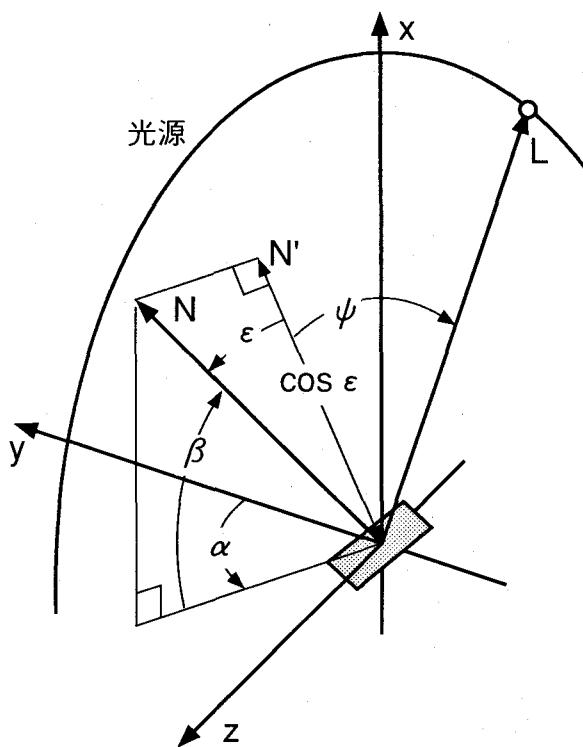


図 3.6 回転に伴う法線ベクトルの回転と明度変化

$$\begin{aligned}
 I^2 &= 4 d^2 i^2 \left\{ 1 - \cos^2 \beta \sin^2(\gamma + \phi) \right\} \\
 &= 4 d^2 i^2 \left\{ 1 - \cos^2 \beta \frac{1 - \cos(2\gamma + 2\phi)}{2} \right\} \\
 &= 2 d^2 i^2 \cos^2 \beta \cos(2\gamma + 2\phi) + 2 d^2 i^2 (2 - \cos^2 \beta) \\
 &= A \sin 2\phi + B \cos 2\phi + O \quad (3.11) \\
 &\quad \left( \begin{array}{l} A = -2 d^2 i^2 \cos^2 \beta \sin 2\gamma, \quad B = 2 d^2 i^2 \cos^2 \beta \cos 2\gamma, \\ O = 2 d^2 i^2 (2 - \cos^2 \beta) \end{array} \right)
 \end{aligned}$$

この式に関する最小二乗近似は形状計測部分と同一の形式であるため、同様に各パラメータを算出することが可能である。この式において

$$\begin{aligned}
 \sqrt{A^2 + B^2} &= 2 d^2 i^2 \cos^2 \beta \\
 O &= 2 d^2 i^2 (2 - \cos^2 \beta) \quad (3.12)
 \end{aligned}$$

より  $\beta$  と  $di$  が求められ、また  $A$  と  $B$  の比率から  $\gamma$  が求められる。

微小面の運動は形状復元の結果追跡可能であり、それは時空間螺旋画像上では  $\sin$  曲線を描く。 $\sin$  曲線上の明度の変化に 3.11 式を当てはめると微小面の法線方向と反射率係数が同時に計測される。変化させるパラメータは回転角のみなので、形状復元の場合と同じく、時空間螺

旋画像内で全ての処理が完結する。ただし、法線ベクトルの仰角  $\beta$  の正負は判別できないため、形状計測結果を回転軸に平行な方向へ微分することで法線ベクトルの符号を決定する必要がある。

### 3.3 反射率・法線ベクトルの計測実験

この節では、第2章で述べた全周形状の計測手法にこの章で述べた拡散反射率・法線ベクトル計測手法を融合し、実際に計測実験を行った結果を示す。CG制作を目的とした計測であるため、計測結果を用いたCG画像の合成結果も示す。

#### 3.3.1 計測装置の構成

実験装置の構成はほぼ2.3.1の図2.11と同一である。ただし、対象物体を照明する円環状の光源が図3.4のように設定されている点が異なる。計測装置の外観を図3.7に示す。光源は40Wのサークル蛍光灯の内周付近をスリット状に約1cm幅残して遮光塗装したものを用いた。円環光源の直径は345mm、回転軸とカメラの距離は640mmである。カメラには白黒CCDカメラを用いた。

#### 3.3.2 円錐の計測による精度評価

図3.8に示すような円錐を用い、円環光源を用いて拡散反射率と法線ベクトルの推定実験を行った。図3.9に、形状と法線ベクトルの推定結果を示す。推定された法線ベクトルはそれぞれ放射状の短線分として表示されている。法線ベクトルの仰角の符号は断面画像のみでは判定

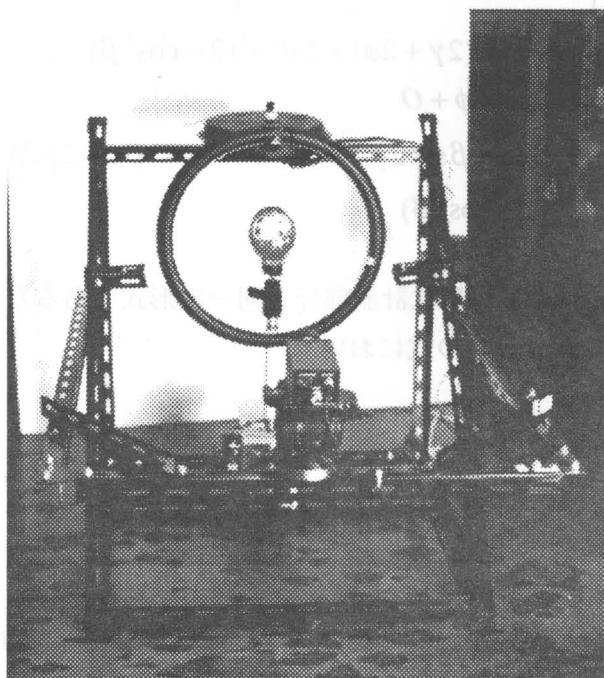


図3.7 計測装置の外観

不可能なので、ここでは一律に正方向とした。符号が未知である場合には、前後の断面形状を利用して判定する必要がある。

得られた拡散反射率を用いて再合成した画像を図3.10に示す。3角パッチそのものの法線ベクトルでなく、推定された法線ベクトルを用いて明度を計算している。推定結果がおおむね妥当であることが確認できる。法線ベクトルの推定誤差の標準偏差は方位角・仰角ともに5.5 degであった。法線ベクトルと反射率の分布を図3.11に示す。明度が物体の白黒に応じて分離していることと、仰角が一定の値の付近に集中していることがわかる。

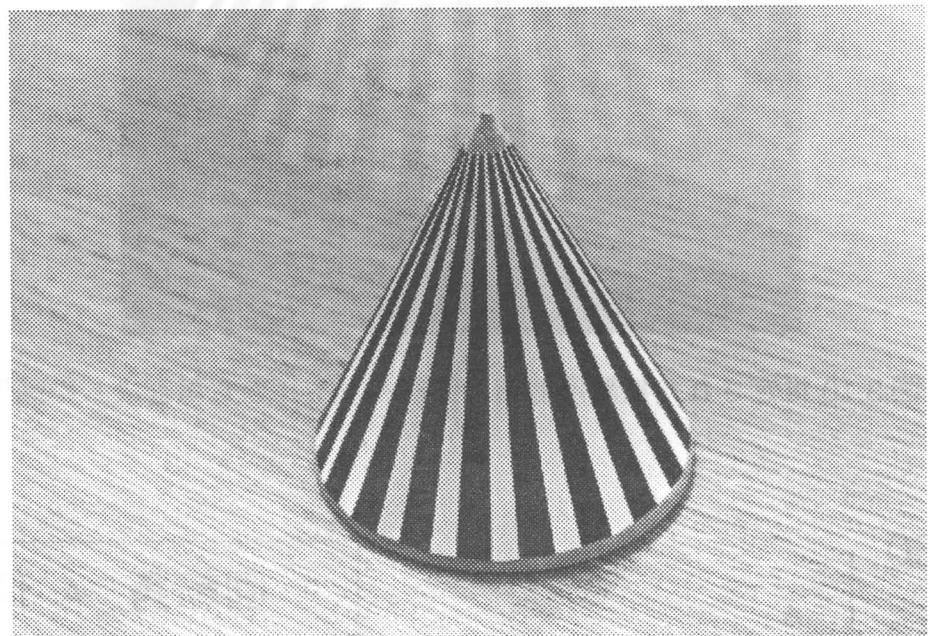


図3.8 精度評価に用いた円錐

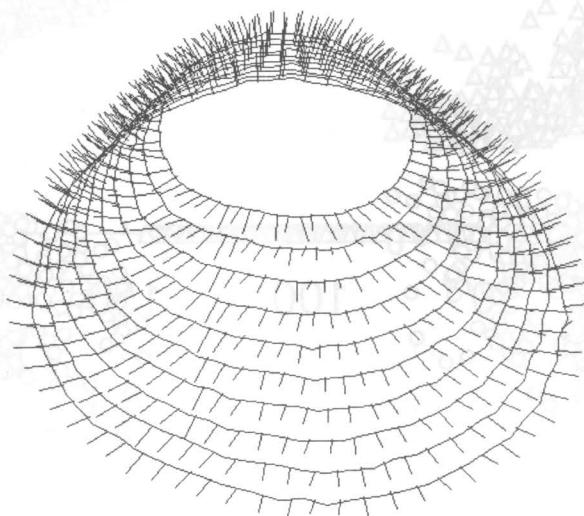


図3.9 形状と法線ベクトルの推定結果

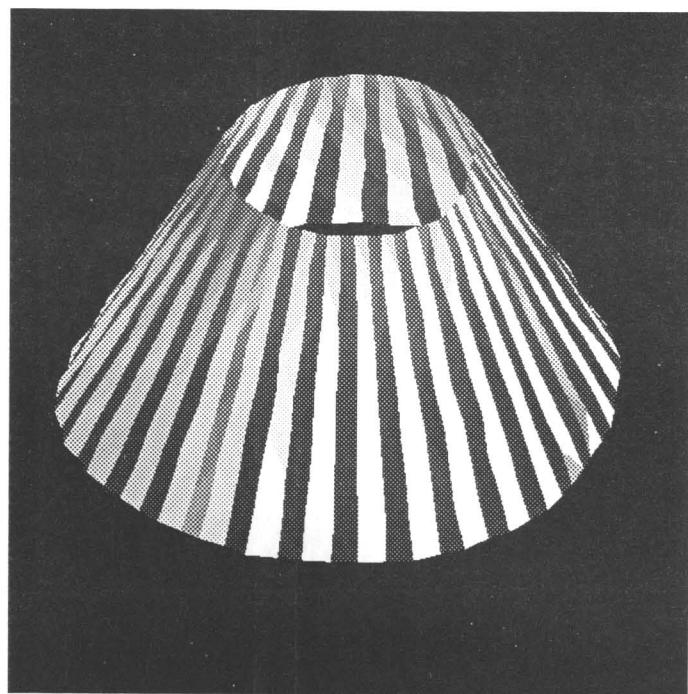


図 3.10 推定された拡散反射率と法線ベクトルから再合成した画像

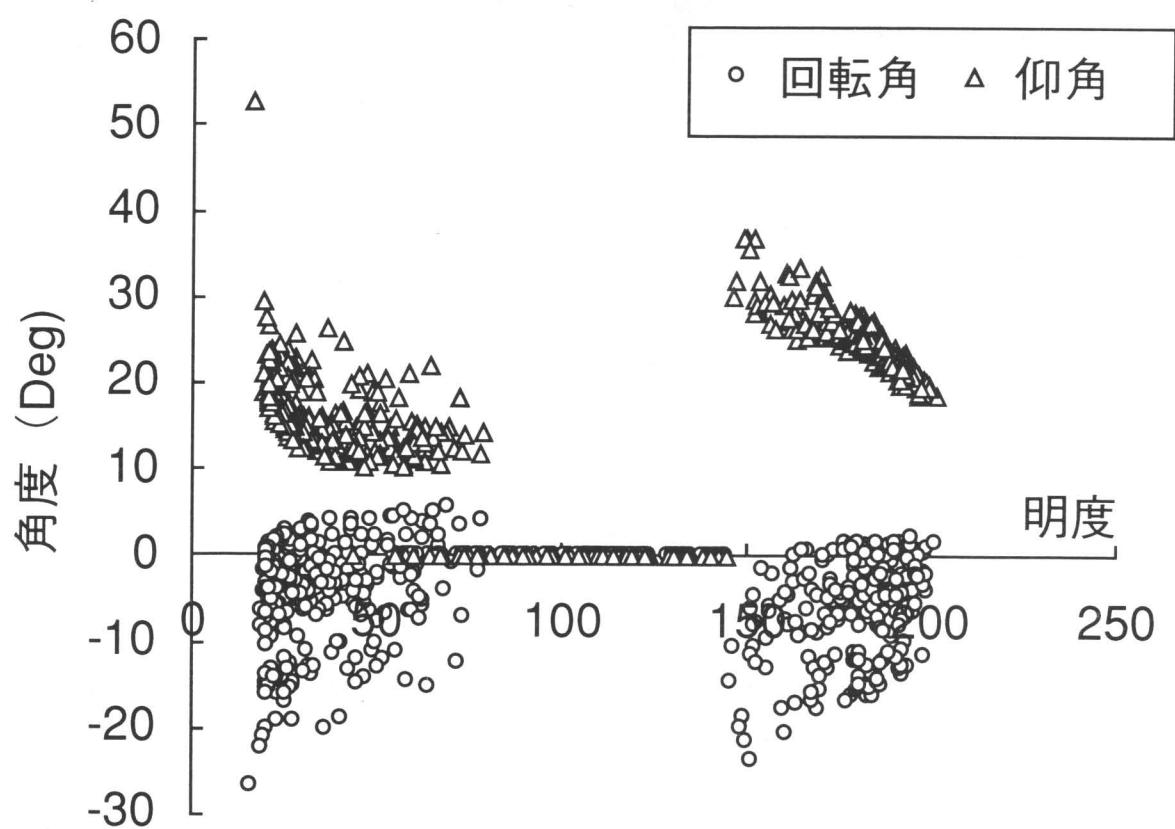


図 3.11 法線ベクトルと明度値の分布

### 3.3.3 実物体の計測例

図3.12に示すようなダルマの模型を用いて形状・反射率計測の実験を行った。対象物体は顔の周辺部で凸形状となり、また顔以外の領域では比較的強い鏡面反射が観測される。形状計測段階では、時空間螺旋画像上で特徴点を追跡する方法に加え、シルエット像を解析する方法を組み合わせて汎用化を図った。計測された形状と反射率を用いてシェーディング表示を行った結果を図3.13に示す。

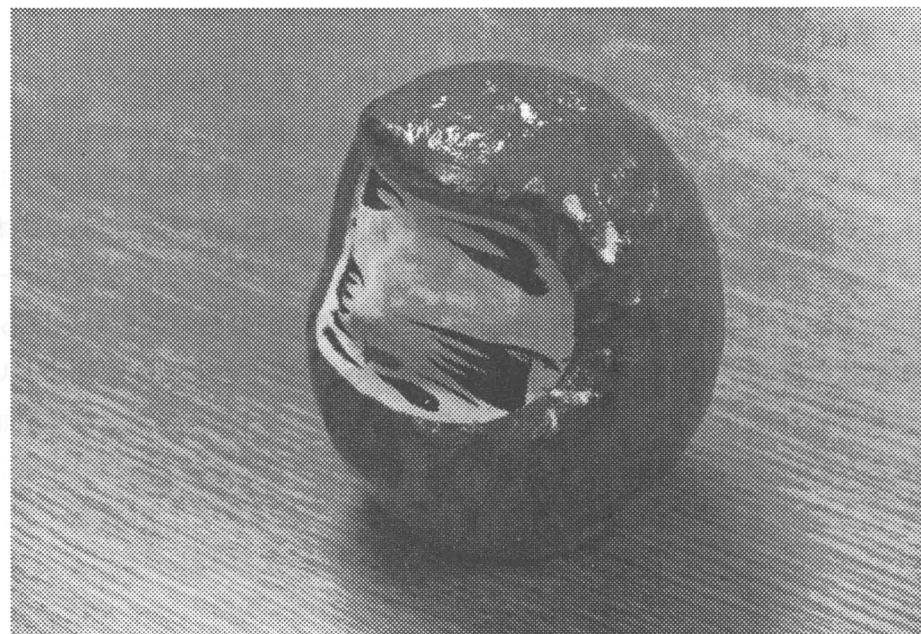


図3.12 実験に用いた対象物体の外観

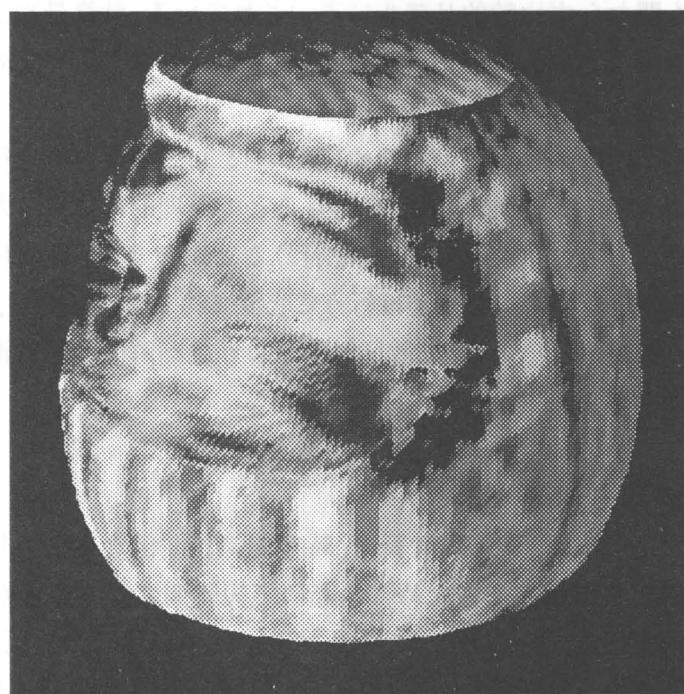


図3.13 計測された反射率と法線ベクトルを用いてシェーディング表示した画像

### 3.4 考察

CG モデルにおいて物体の明るさや色を決定付ける反射率パラメータと、リアルな陰影を表現する法線ベクトルは形状に比べ緻密なデータが要求される。ここでは得られた反射率・法線ベクトルについて CG 制作に用いることを前提に考察する。また、この技術が CG 製作現場で用いられることを念頭に置きながら計測速度や計測装置の簡便性について考察を行う。

#### 3.4.1 反射率・法線ベクトルの計測結果について

日用品などの工業製品はそのたいていが滑らかな形状にさまざまな文字や絵柄が印刷により配されている物である。このような物体を CG により表現する場合、もっともよく用いられる手法は CAD により定義された形状に、別に用意した二次元画像をテクスチャマッピングすることで表示する方法である。自然物の表現においても同じく定義した形状にありあわせのテクスチャデータを選択し配置することで表現することが多い（大理石のような模様データなどが多数商品化されている）。これは物体の形状が模様に比べ変化が少ないと挙げられるが、もう一つは物体形状の細部が少々実物体と異なっていても、合成された CG 画像において明暗が正しく表現されればリアルに感じるという事実に基づいている。また、物体表面の細かい形状を表現するために、直接形状データを細かく定義するのではなく、单一の平面について法線ベクトルを各点定義することで波打っているような表現を行うことがある。これをバンプマッピングと呼び、直接形状を定義する場合に比べ計算量やデータ量が少なくてすみ、かつ簡便であるため広く用いられている。本手法で得られた反射率・法線ベクトルはテクスチャ・バンプマッピングを用いた CG 計算モデルに適合するデータであり、3.3 で示した CG 例もこの手法により作成した。3.3.2 で示したダルマの表示例はこれらの反射率・法線ベクトルの解像度が不十分であるにも関わらず比較的リアルに表示されている。計算量や記憶容量の制限からさらに高解像度な反射率復元は行わなかったが、この CG における質的な問題はほとんど解像度の不足に起因すると考えられる。

実験にはモノクロビデオカメラを用いたため、カラーとしての反射率を復元していないという問題も挙げられる。しかし各色同様の処理を行うことで原理的にはまったく手を加えることなく赤・青・緑の三原色についての反射率計測を行うことができる。鏡面反射率を計測していないため、光沢のある物体のリアルな表現は行うことができない。しかし、鏡面反射成分は物体の材質により大きく影響されるため、拡散反射率のように高解像度で計測する必要があるかどうか検討する必要がある。

#### 3.4.2 計測時間について

一点についての反射率・法線ベクトル計算は、時空間螺旋画像から抽出した明度値の集合について最小自乗法を用いるだけであるため、きわめて短時間で行うことができ、計算量は形状復元とほとんど同一である。現在は一つの三角パッチについて一つの反射率・法線ベクトル復元しか行っていないため、形状復元処理と組み合わせても計算時間がおよそ倍になるだけであったが、実際の CG 制作ではさらに高解像度のデータが要求されるため、長時間の処理が

必要になることが考えられる。計算量の減少のためには、現在のように照度差ステレオ法を全ての計測点について行うのではなく、間隔をあけて計測した法線ベクトルを用いて周辺のさらに高解像度の拡散反射率を求めるなどの工夫が考えられる。

### 3.4.3 装置の構成について

円環状の光源は一般的な装置ではないが、正確に反射率を解析する手法としては光源や撮影方向を機械的に移動する変角光度計に比べ簡便であると言える。回転による形状計測と組み合わせて反射率を計測する手法としては簡潔かつ強力な方法である。

## 3.5 今後の課題

現在の計測システムに対する評価をもとに、2章の実験結果も踏まえてこの計測手法を今後どのように拡張・発展するべきか考察する。

### 3.5.1 鏡面反射率の復元

対象物体の質感表現は微細なテクスチャとともに、光源の映り込みなどによるハイライトが大きく影響する。特にCGアニメーションでは、物体の運動とともに移動するハイライトが材質感や微妙な形状の歪みに対する情報を与えるために重要である。鏡面反射率を求める研究は大槻 [Otsuki93] が行っているが、法線ベクトルを既知とした上で単一光源によるシェーディング解析により求めているために正確に鏡面反射率が求められているか疑わしい。正確に鏡面反射率を計測するためには、鏡面反射が生じているときの反射強度を得ることが重要であり、妥当な鏡面反射モデルを用いたとしても、反射光分布の“裾野”から全体を推定するのは非常に困難である。

円環光源を用いることでハイライト時の反射光を得ることは非常に簡単となる。我々が用いた装置のように、カメラ光軸と垂直な面内に円環光源を配置した場合、対象物体表面の法線ベクトルと回転軸のなす角が $45^\circ$ 以上であれば必ずハイライトが計測される。円環光源をカメラ光軸と物体回転軸の張る平面内に配置した場合は法線ベクトルの向きに関わらず必ずハイライトが観測される。しかしこのような配置はカメラの撮影範囲と光源が干渉するために実現が不可能である。なお、拡散反射率の推定に関してはカメラの光軸と円環光源の配置とは本質的に無関係であるため、どちらの方法をとっても同じ手法を用いることができる。

鏡面反射率の推定については、反射モデルが複雑であり円環光源についての線積分の展開が複雑である点、またその線積分を簡単に行うことができる反射モデルの調査が進んでいない点が問題として残っている。

### 3.5.2 カラー計測

カラー計測については前にも述べたとおり、反射率の推定処理を赤・青・緑各色独立に行うことで実行が可能である。または、モノクロで求められた拡散反射率をカラー値の比率に分割することで求めることもできる。しかしカラーカメラはモノクロカメラと比較して解像度が劣

り、またNTSCの性質として縞模様などを撮影するとクロスカラーなどが発生することから、現在は機器の都合によりカラー化を見送っている。3CCD方式・RGB出力のカメラを用いるとこれらの問題は解決されるであろう。

### 3.5.3 法線ベクトルによる形状の高解像度化

反射率推定段階で得られた法線ベクトルは対象物体の細部構造を反映していると考えられる。本手法の形状計測原理は特徴点追跡に基づく方法であるため、エッジのない部分については形状を得ることができない上、特徴点の検出の困難さから、入りくんだ形状について計測することは困難である。そこで、形状計測結果に法線ベクトルのデータを付加することで細かな形状を復元することが可能であると考えられる。照度差ステレオ法では得られた法線ベクトルを積分することで全体形状を得るが、ここでは形状計測により得られた特徴点の間を区分割するだけで良いので、ジャンプエッジを含む形状についても絶対的な精度を高く得ることができる。また、得られた断面形状同士を連結する際の情報としても用いることができると考えられる。

## 4 章 動距離画像と濃淡画像の融合による 幾何学形状物体の実時間追跡

従来静的なシーンに対する解析が中心であったコンピュータビジョン研究は、近年の飛躍的な計算機性能の向上によって計算量的問題から開放され、動的シーン研究分野へ拡大しながらシフトしつつある。なかでも、空間内を移動する物体の運動を画像により計測・追跡する問題（tracking）は、実時間視覚システムの基礎的研究テーマとして非常に重要性が高い。また、情報の入力部たる画像計測装置も長足の進歩を遂げ、濃淡画像のみならず距離画像をも実時間的に計測する装置が出現するようになってきた。そこでこの章から6章にかけては、距離画像を用いて動物体の運動を実時間追跡するシステムについて行った研究を紹介する。

### 4.1 距離画像による物体の姿勢推定と追跡

濃淡画像により動物体の運動を実時間的に追跡する研究は多数行われているが、距離画像を用いた研究例は非常にわずかである。なぜなら、動物体を追跡し得るほど十分高速なレンジファインダは開発途上であり、世界的にも数例しか稼働しているシステムがないからである。そこで、この節では距離画像を用いて物体の姿勢を推定する研究例も含めて紹介する。

#### 4.1.1 表面曲率を用いる方法

距離画像は対象物体の表面形状に関する情報であるから、対象物体の姿勢を求める方法としてはその表面形状から特徴量を求め、モデルと対応付けする方法が考えられる。藤田・佐藤[Fujita90]は曲面物体の認識・姿勢の算出を行うために、表面曲率を用いる方法を提案した。この手法では、密に得られた距離画像の隣接点同士を調べることによってガウス曲率と平均曲率の二つの特徴量を求め、この符号や絶対値の大小によって面をセグメンテーションする。次に、その隣接関係や特徴の大小を用いて階層構造的なシーン記述を得、物体種類の認識・モデルとの比較を行い姿勢を求めた。この手法は、ボトムアップ的なMarrのパラダイム[Marr82]を距離画像に適用したものであると考えることができ、シーン記述の構築段階までにモデルに関する情報を陽に用いていないために、高品質な距離画像が得られる必要がある。

また池内は、法線ベクトル分布のヒストグラムである拡張ガウス像を用いて姿勢を検出する手法に関する研究を多数行った[Ikeuchi90]。またその発展として、球体状のメッシュモデルを収縮させることにより対象物体を包含し、各メッシュごとに表面曲率を格納したようなデータ構造の回転量を求ることにより曲面物体の姿勢を求める手法も提案している[Hebert95]。

#### 4.1.2 多面体の面方程式を用いる方法

池内は、対象物体を多面体と仮定したときに、複数の距離画像同士の位置合わせを精度良く行う方法としてPCAMDを用いた解法を提案している[Ikeuchi95]。距離画像はあらかじめエッジを境界線としたセグメンテーションと平面近似により多面体モデルに変換されているとする。

また、扱う全ての距離画像について、各平面同士の対応付けはなされているものと考える。各平面からは領域の大きさや形状などの特微量を抽出することができるため、対象物体が連続的に運動している場合には各平面同士の対応付けは比較的簡単に行うことができる。

各距離画像の視点の位置関係を算出する手法としてまず考えられるのは、各平面を表す方程式同士を比較する方法がある。視点の位置関係が求められれば、それぞれの距離画像を座標変換して統合することが可能である。しかしこの方法を全周に渡って用いた場合、誤差がしだいに蓄積し、全体としてはひずんだ形状が得られやすい。そこで視点位置の算出をこのような逐次外挿的手法で行わずに、主成分分析法により一度に全ての視点位置と面方程式を求めることを考える。すなわち、2.1.1で述べた Factorization Method と同様に、観測される各面の方程式は、原座標系における面方程式に対して同次座標変換を乗じることにより得られたものであると考え、この観測行列を面方程式行列と座標変換行列に特異値分解することによりそれぞれの平面の方程式と視点位置を正確に得る。観測方向の数と面の数をそれぞれ  $m$  と  $f$  とすると、座標変換行列は  $4m$  行  $4$  列、面方程式行列は  $4$  行  $f$  列であるために、観測行列のランクは  $4$  となり、それぞれの行列に分解することが可能である。ただし、観測行列中の面方程式は視点の位置によって虫食い状であるために、このような行列に対して適用可能な特異値分解として Wiberg により提案された PCAMD [Wiberg76] を用いる。

この手法は明らかに安定にセグメンテーション可能な多面体についてのみ用いられるが、清水は複雑な凹凸を持つ物体についても三次元凸包を用いることにより同様の手法を応用可能であることを独自に提案した [Shimizu96]。三次元凸包とは、任意の物体を内包する最小の凸多面体のことであり、この凸包多面体同士の位置合わせによって対象物体の位置合わせを行う。この手法は滑らかな凸曲面が表面の大部分を覆う対象について用いることはできない。

#### 4.1.3 距離画像間の距離最小化による方法

得られた二つの距離画像同士が一致するように、距離画像同士の相違度を表す規範を定め、その値が最小となるように運動パラメータを直接最適化する手法である。この手法のうち、多数の研究者によって多数の発展が試みられている手法として、Besl と McKay によって提案された ICP (Iterative Closest Point) アルゴリズム [Besl92] がある。この手法では、距離画像間の距離として、一方の距離画像上の各点からもう一方の距離画像上の最近傍点までの距離を求め、この値の自乗和が最小となるように最適化を繰り返す方法である。Simon はこの手法に様々な高速化アルゴリズムを組み込み、高速レンジファインダ [Kanade91] を用いて動物体の実時間追跡を行った [Simon94]。この手法は次章との関連が深いため、次章 5.1.1 で詳細な検討を行う。

#### 4.1.4 距離画像の固有空間照合法

距離画像は濃淡画像と比較してより直接的に対象の形状を表していると考えられるため、解析的手法により様々な応用が可能である。しかしこれを敢えて単なるパターンデータと捉え、パターン照合的手法により対象物体の姿勢を検出することも可能である。すなわち、対象物体の様々な姿勢に関する距離画像（学習距離画像）を多数用意しておき、得られた距離画像をこ

これらと比較することにより最も近い姿勢をもつて対象の姿勢とする手法である。この手法では多数の学習距離画像を蓄え、さらに比較するための容量的・計算量的コストがかかるため、主成分分析を用いて記憶容量と計算量の削減を実現する手法を筆者の共同研究者である天野が開発した [Amano96]。

この方法では、あらかじめ学習距離画像から主成分分析法により寄与率の高い情報のみを取り出し、固有ベクトルの一次結合として学習画像を圧縮しておく。姿勢を検出すべき新たな距離画像が得られた時には、これも固有ベクトルの一次結合として分解し、この結合係数同士を比較することにより最も近い距離画像を探査する（図4.1）。また、その姿勢の周辺の学習距離画像のうち次に近い距離画像も探し、これらの類似度と姿勢を用いて線形補間により連続的な姿勢値を求めた。

この手法ではあらかじめ各姿勢に対する距離画像を計測しておく必要があるが、対象の形状モデルから学習距離画像を計算機内で仮想的に生成することも可能である。この仮想学習法では、濃淡画像に対する固有学習法と異なり照明条件を考慮する必要がないため、よりロバストに環境内の物体姿勢を求めることが可能である。また、対象物体の一部が他の物体により遮蔽されたり、対象物体の反射率などにより距離値が得られない画素（画素欠損）が現れた場合でも安定に姿勢検出が可能である手法も開発した。全探索を行う性質上、初期条件なしに一度で最適解を得ることができる利点を有するため、他の逐次最適化手法に基づく姿勢推定の初期値生成に用いることが可能であると考えられる。ただし、学習距離画像を圧縮しているとはいえ、6軸追跡に応用するには容量的・計算量的コストは依然大きすぎると思われる。

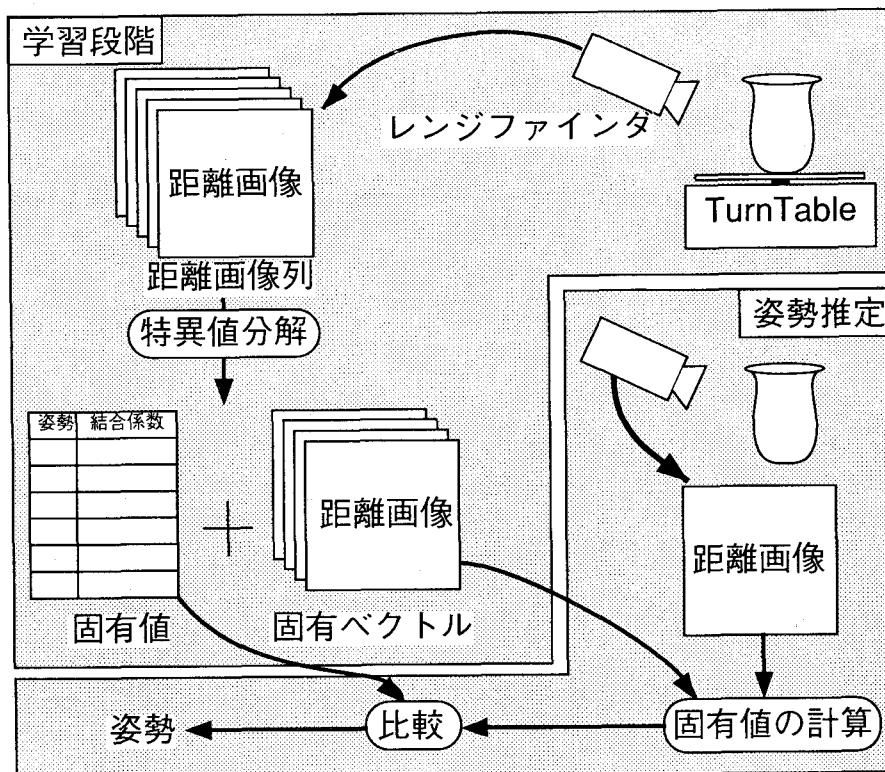


図 4.1 固有空間照合法による姿勢検出

## 4.2 シリコンレンジファインダ

昨今の LSI 技術の発展により、飛躍的な高速化を図ったレンジファインダが出現するようになった [Kanade91] [Kang95]。シリコンレンジファインダ [Yokoyama96] もそのようなセンサの1つであり、受光素子に専用フォトLSIを用いることにより毎秒30フレーム以上の距離画像を計測することができる。この節では、このセンサの高速化原理と、そのキャリブレーション手法について解説する。

### 4.2.1 シリコンレンジファインダの原理

シリコンレンジファインダのシステム構成を図4.2に示す。レーザ光源によるスリット光を、ガルバノミラーにより偏向することで投光器とする。シーンに投影されたスリット光の像を受光系により撮影し、ミラーの角度と像の位置関係から対象物体の形状を計測する。言い換れば、スリット光により対象物体を切断し、その切断面の形状を計測することで形状を得ているとも言えるため、これをスリット光切断法とも呼ぶ。スリット光切断法による計測時間は、一つのスリット光の像を取り込む時間と、断面を取るスリット光の数(奥行き分解能を決定する)の積で決定する。通常のビデオカメラでは一つの画像を取り込むのに1/30秒必要とするため、例えば300枚のスリット光により計測を行うためには最低でも10秒を要する。つまり、奥行

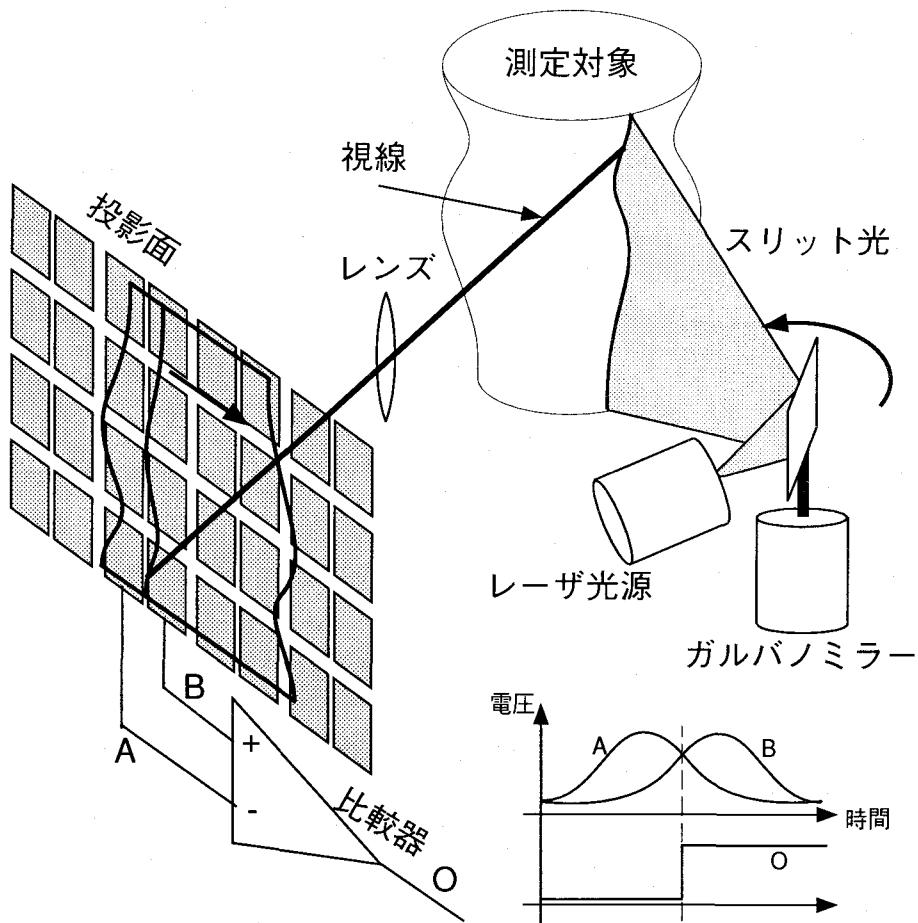


図 4.2 シリコンレンジファインダの計測原理

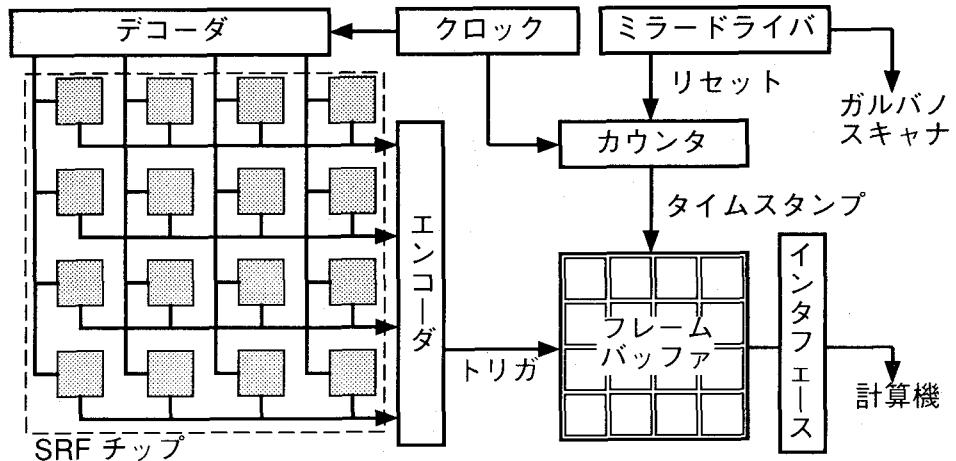


図 4.3 SRF の読み出しとカウント値の計算

き分解能を確保しながら計測速度を向上させるためには、受光系の取り込み速度を高速化する必要がある。そこで、通常のビデオカメラのように、各画素の受光量を順次取り出す方式を取らずに、並列的に読み出すことで計測速度を改善することを考える。

スリット光の像を検出するためには、背景の明度に比べ十分に強力なレーザを用い、閾値を適切に設定することで行うことができるが、これでは環境にあわせて設定を変更する必要が生じる。また、レーザスリット光の検出位置がレーザ光の強度によって変化することもある。そこでシリコンレンジファインダでは、左右に並べた2つの受光素子の光量を比較することで、高精度かつ安定にレーザ光の通過時刻を算出することができる。また、光量の比較回路とその比較結果を保持するフリップフロップを受光素子とともにチップ上に置くことで回路の動作速度を向上させ、レーザ光の通過時刻（スリット番号に対応する）を最高  $10 \mu\text{秒}$  単位で計測することができる。画素数は縦横ともに  $24 \times 24$  画素であるために、この比較結果は専用の信号処理装置から順次読み出され、スリット光が通過したときにスリット光平面の位置に対応するカウント番号が各画素専用の 16bit メモリに書き込まれる（図 4.3）。ホストコンピュータからはスリット光のスキャンが終了してからこれらのカウント番号メモリを読み出すことによって、各画素ごとにスリット光が通過したタイミングを得ることができる。

#### 4.2.2 シリコンレンジファインダのキャリブレーション

スリット光切断法によるレンジファインダ光学系の幾何学的表現として、システムパラメータ表現がある [Amano95]。図 4.4 のように、世界座標系  $(X, Y, Z)$ 、カメラ系の投影面の座標系を  $(X_c, Y_c)$ 、プロジェクタ系の投影面の座標系を  $(X_p)$  とする。これらの座標間の関係を同次座標変換を用いることで以下のように表現することができる。

$$\begin{bmatrix} h \cdot x_c \\ h \cdot y_c \\ h \end{bmatrix} = \mathbf{C} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad \begin{bmatrix} h \cdot x_p \\ h \end{bmatrix} = \mathbf{P} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.1)$$

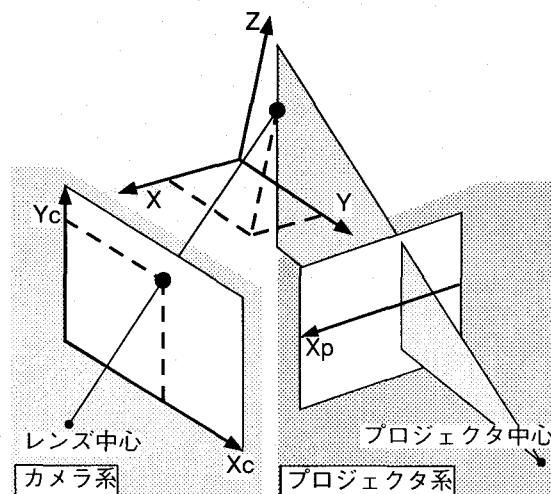


図 4.4 システムパラメータ表現

この 2 式において、3 行 4 列の行列  $C$ （カメラパラメータ）と、2 行 4 列の行列  $P$ （プロジェクタパラメータ）が既知であるとすると、 $(X_c, Y_c, X_p)$  を与えることで  $(X, Y, Z)$  を求めることができる。これをレンジファインダのシステムパラメータ表現と呼ぶ。この方法を用いることで、例えば世界座標系を再定義する場合には、新たな世界座標系への同次変換行列を  $C$  と  $P$  の後ろから乗算することで簡単に新しいパラメータを得ることができる。

カメラパラメータ  $C$  は、世界座標  $(X, Y, Z)$  と像座標  $(X_c, Y_c)$  の組を 6 つ以上与えることで求めることができる。同様に、プロジェクタパラメータ  $P$  は世界座標  $(X, Y, Z)$  とプロジェクタ座標  $(X_p)$  の組が 8 つ以上与えられることにより求められる。シリコンレンジファインダは受光素子として専用 LSI を用いているため、ピクセル解像度の問題から像座標  $(X_c, Y_c)$  を高精度に求めることができない。そこで、通常の CCD とプリズムを用いて同一レンズを通した濃淡画像を撮影する構造とし、濃淡パターンを Z ステージにより平行移動し、多数撮影することによりカメラパラメータのキャリブレーションを行った（図 4.5）。SRF は極端に反射率の低い部分については正常なカウント値を得ることができないため、この画素欠け位置と CCD から得た濃淡画像を比較することにより CCD と SRF の各画素の位置関係を得た。

また、プロジェクタ系に回転ミラーを用いているために、透視変換を表すプロジェクタパラ

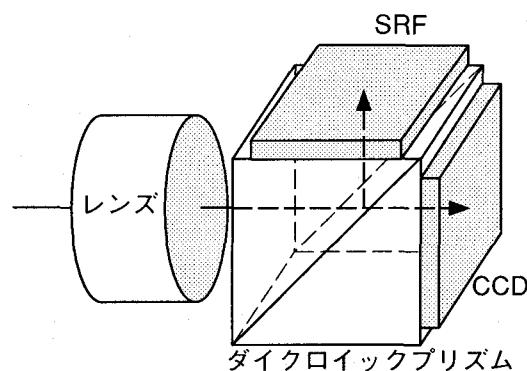


図 4.5 プリズムによる濃淡画像の同軸計測

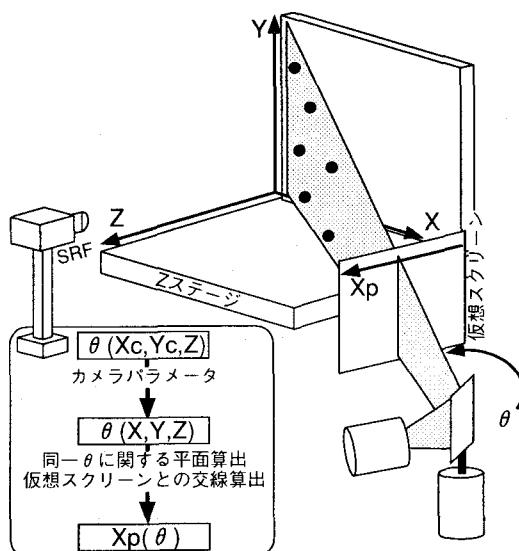


図 4.6 プロジェクタ補正関数の算出

メータをそのまま用いることができない。そこで、仮想スクリーンという概念を導入し、ミラーの回転角をプロジェクタ座標 ( $X_p$ ) に変換する（図 4.6）。プロジェクタ座標  $X_p$  と世界座標系の関係は透視変換であるため、プロジェクタパラメータを用いることで表現することができる。結果的に受光系から得られた SRF の画素番号を CCD 上の座標に変換し、またミラー回転角をプロジェクタ座標に変換することでシステムパラメータ表現によりシリコンレンジファインダによる計測を行うことができる（図 4.7）。

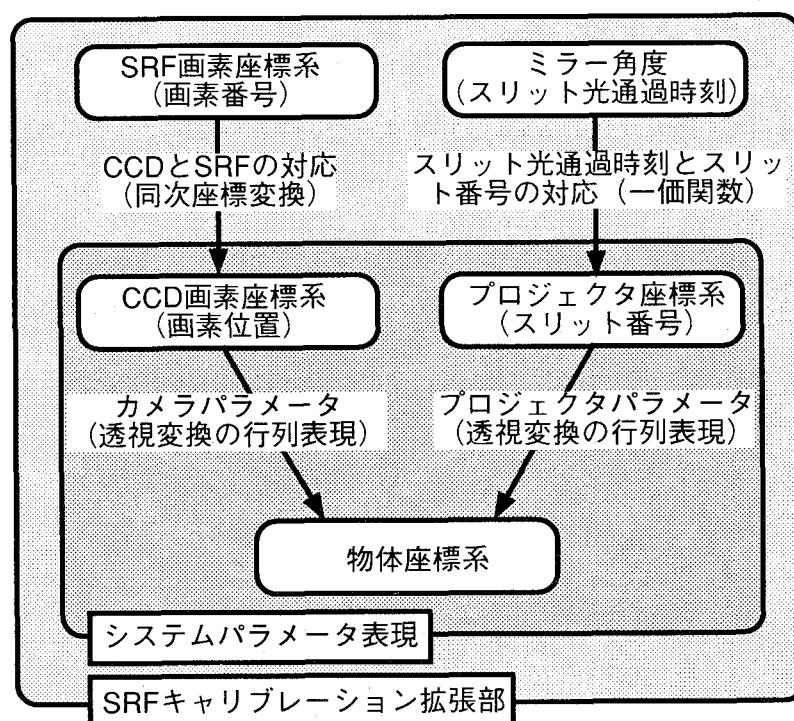


図 4.7 座標系とシステムパラメータ表現の関係

### 4.3 距離画像と濃淡画像による連結物体追跡の原理

物体の三次元位置を計測するにはレーザ測距儀などのように一点について精密に距離計測する装置を用いることが考えられる。しかしこれは計測点と装置の間の距離がわかるだけであり、対象物体のどの点を計測しているのかわからないため、物体の位置・姿勢を計測したことにはならない。対象物体の大きさが計測分解能に比べ無視できないくらい大きな場合は、物体の中心を正しく得るために物体の姿勢の検出が不可欠である。計測装置と対象物体の距離を計測することにより物体の位置・姿勢を検出するには、一般的には以下の条件が満たされる必要がある。

A：一つの剛体につき三箇所について距離が得られること

B：計測点の物体表面における位置がわかっていること

物体位置はあらかじめわかっているわけではないから、条件Aを満たすためには物体の大きさに比べ十分細かいピクセル解像度の距離画像が得られている必要がある。これについては次の節で論じる。

条件Bを満たすためには、より多数の距離値と対象物体の形状の関係を解析することで得る方法がある。しかしこのような方法は対象物体の形状やモデリング方法に条件が課せられることが多く、自由に物体を定義することはできない。また、むやみに解析的手法で計算しようとすることは計算量から考えても現実的でない場合が多い。

位置・姿勢を解析的に導出する方法の他に、計測値を良く満たすパラメータを逐次最適化手法により求める方法が考えられる。逐次最適化手法で高速に解を発見するためには、1ステップごとのパラメータ変化量を高速かつ正確に求める必要がある。以下ではこれらの条件を念頭に置いて計測手法とその優位性を詳説する。

#### 4.3.1 距離画像と濃淡画像による並列処理

物体の各点への距離を求ることで対象物体の位置・姿勢をリアルタイムに追跡することを考える。動物体追跡では、物体がどこへ移動しても物体への距離値が得られる必要があるため、物体の移動範囲をカバーする距離画像が得られていることが望ましい。動物体追跡に距離画像を用いるためには十分高速なレンジファインダが必要となるが、現在ビデオカメラと同等以上の計測速度を持つレンジファインダのピクセル解像度は非常に低く、シリコンレンジファインダもその例外ではない。しかし、条件Aはわずか3点についての距離値を要求しているだけであるため、対象物体によってはシリコンレンジファインダの解像度である縦横 $24 \times 24$ 画素でも条件を満たす場合は多いと考えられる。逆にこの条件を満たさないような小さな物体については、姿勢が正確に検出されなくても物体の位置検出誤差は小さい。

条件Bについては、逐次最適化手法を用いて解決を図るため、推定すべき位置・姿勢の変化量を正確に求める必要がある。位置・姿勢は全部で6変数により表されるため、求めるべき変化量も6個である。これらの変化量の正確さは均等であることが理想的であるが、シリコ

シリコンレンジファインダは奥行き解像度が数千以上ありながらピクセル解像度が低いため、カメラ光軸に垂直な方向の変位を求めることは困難であると考えられる。

他方、ビデオカメラは放送技術の発展とともに高性能化を遂げてきており、計測速度・解像度ともさまざまな性能の機器を選択することができる。そのため、ビデオカメラは動物体の位置・姿勢を検出する研究に多数用いられてきた。しかし、通常の濃淡画像からカメラ光軸方向の運動を検出することは、カメラ光軸に対して垂直な運動を検出することに比べ非常に困難な問題である。そのため、これら三次元計測研究の多くが対象物体の奥行き方向の運動をいかにして検出するかを主題とするものであった。また、単一のカメラにより奥行き方向の位置を検出するには対象物体の運動や大きさなどに何らかの制約を設ける必要が生じることが多い。しかし、FA分野においては、プリント基板への部品挿入やLSIのボンディングのように、視覚センサによる二次元的な位置合わせが多用されている。このように、ビデオカメラは光軸に垂直な方向の変位を精密に計測することができると考えられる。

総合すると、平行移動成分については、シリコンレンジファインダとビデオカメラは互いに精密計測の可能な軸方向が異なっていることがわかる。そこで、これら二つのセンサを統合することでより高性能な動物体追跡センサを構築することができると考えられる。さらに、対象物体の回転についても考察することで、一般的には図4.8のように検出可能軸を3つずつ分担することがよいと考えられる。

逐次最適化手法を高速に行うもう一つの条件として、各パラメータの独立性が考えられる。それぞれのパラメータが完全に独立である場合、それぞれのパラメータを順番に最適化することでも解を得ることができる。しかし、最適化するパラメータのうち回転の三つは互いに独立ではない。しかし、微小な回転については各軸の回転をベクトル的に加算することで近似できることが証明されている。よって、物体の運動に比べ高速に計測を行うことで、各変位パラメータをそれぞれ独立に算出することが許される。これは、連結物体の追跡における剛体間の相互作用も同時に合成するために必要な条件である。

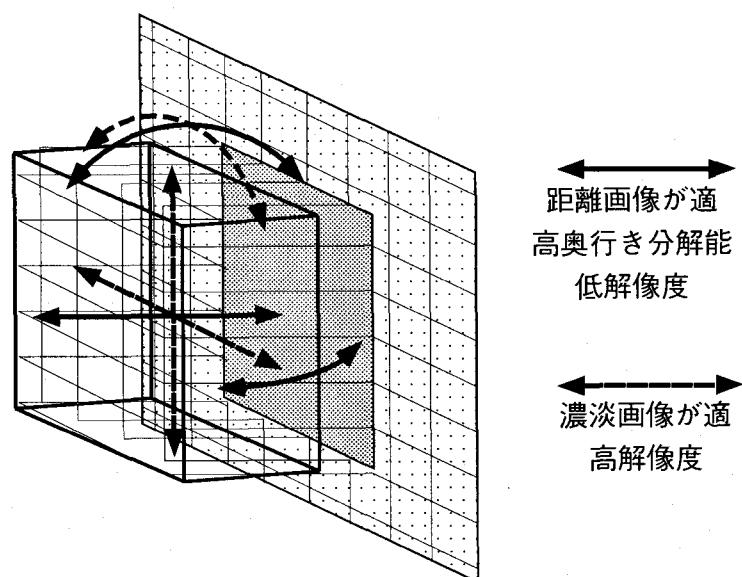


図4.8 距離画像と濃淡画像による変位検出の成分分担

#### 4.3.2 座標変換とキャリブレーション

ビデオカメラにより得られる像とシーンとの関係は一定である。すなわち、レンズの焦点を原点とし、カメラ光軸を  $Z_D$  軸、像面の  $X_c$ ,  $Y_c$  軸と平行に  $X_D$  と  $Y_D$  軸をそれぞれシーン中にとると、カメラによる座標変換（透視変換）は次のように表される（図 4.9）。

$$\begin{cases} x_c = f \cdot x_D / z_D \\ y_c = f \cdot y_D / z_D \end{cases} \quad (4.2)$$

他方、シリコンレンジファインダのキャリブレーションは、シーン中に設置したスライドテーブルにより行うため、その装置の配置が世界座標系を定義し、計測値としては世界座標系  $X_w Y_w Z_w$  における値が得られる。前節で述べたように、シリコンレンジファインダとビデオカメラにより検出される変位は互いに直行している必要があるが、 $X_D Y_D Z_D$  座標（カメラ座標系）と  $X_w Y_w Z_w$  座標（世界座標系）は一致していないため、このままでは変位を合成することができない。そこで、シリコンレンジファインダにより計測される値をカメラ座標系に変換する必要がある。

世界座標系・カメラ座標系ともデカルト座標系であるため、相互の変換は回転と平行移動により表すことができる。これらを直接回転成分と平行移動成分に分けて計算しても良いが、線型変換ではないため不便である。そこで、同次座標変換を用いることで線型化する。

$$\begin{bmatrix} x_D \\ y_D \\ z_D \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.3)$$

ここで、 $r_{rc}$  ( $r,c=1,2,3$ ) は回転成分を表し、 $t_r$  ( $r=1,2,3$ ) は平行移動成分を表す。この 4 行 4

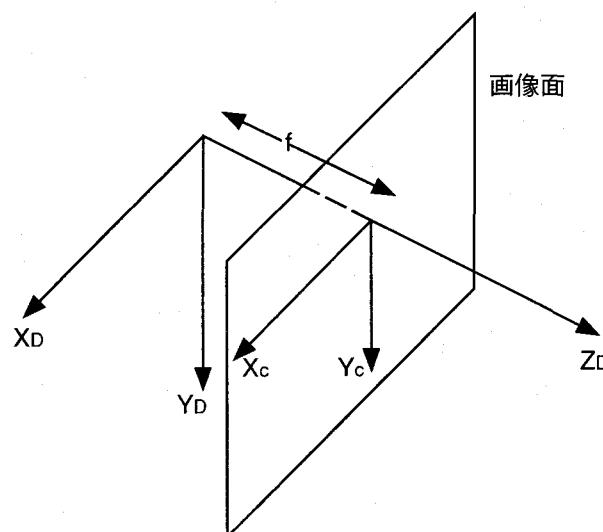


図 4.9 カメラ座標系と透視変換

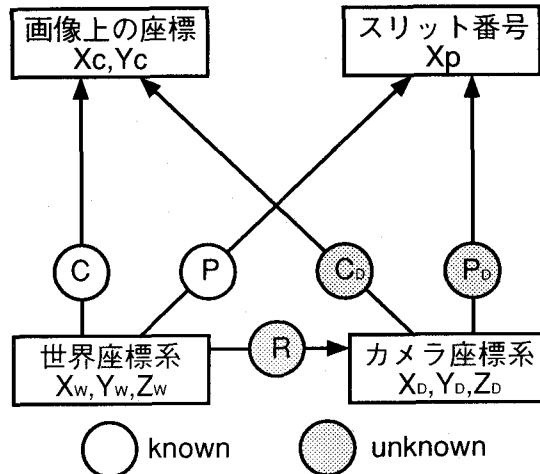


図 4.10 座標系とシステムパラメータの関係

列の行列  $R$  を用いることで、シリコンレンジファインダより得られた座標値をカメラ座標系における値に変換することができる。しかし、これでは変換が余計に一つ増えることになり速度の低下が免れない。そこで、シリコンレンジファインダのシステムパラメータを  $R$  を用いて変換することで、直接カメラ座標系における値を出力するような新しいシステムパラメータを生成する。システムパラメータは、世界座標系から画素番号への変換  $C$ （カメラパラメータ）と、同じく世界座標系からスリット番号への変換  $P$ （プロジェクタパラメータ）からなっているので、それらの変換の前にカメラ座標系から世界座標系への変換を合成すればよい（図 4.10）。 $R$  はその定義から逆行列が存在することがわかっているので、 $R^{-1}$  を  $C$  と  $R$  の前から掛けることでカメラ座標系におけるシステムパラメータ  $C_D$  と  $P_D$  を生成することができる。

$R$  を算出するには、世界座標系をカメラ座標系において計測する必要があるが、これは計測が含まれるため手順が複雑である。そのかわり、カメラ座標系におけるカメラパラメータ  $C_D$  を先に算出することで  $R$  を算出する手順を用いた。つまり、カメラパラメータ  $C$  にはカメラの配置に関する情報（外部パラメータ）と、カメラそのものに関する情報（内部パラメータ）が含まれているが、このうち内部パラメータが  $C$  と  $C_D$  双方で共通であることを利用する。

まず、 $C$  から内部パラメータ（焦点距離、画像中心、アスペクト比）を算出する。求めるべきカメラパラメータ  $C_D$  の外部パラメータは、カメラ座標系とカメラの配置の関係が明らかであることから算出可能であり、これに  $C$  から求めた内部パラメータを合成することで  $C_D$  を生成する。

次に、 $C$  と  $C_D$  を比較することで  $R$  を求める。 $C_D$  は、その性質上第 4 列がすべて 0 であるため、次のように左 3 列と第 4 列を分解して考える。

$$C = [C_{33} | C_T], \quad C_D = \begin{bmatrix} 0 \\ C_{D33} \\ 0 \\ 0 \end{bmatrix}, \quad R = \left[ \begin{array}{ccc|c} R_{33} & & & T_{31} \\ 0 & 0 & 0 & 1 \end{array} \right] \quad (4.4)$$

また  $C = C_D R$  であるため、次のように展開することができる。

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{\mathbf{D}_{33}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \left[ \begin{array}{c|c} \mathbf{R}_{33} & \mathbf{T}_{31} \\ \hline 0 & 0 \\ 0 & 0 \end{array} \right] = [\mathbf{C}_{\mathbf{D}_{33}} \mathbf{R}_{33} | \mathbf{C}_{\mathbf{D}_{33}} \mathbf{T}_{31}] \quad (4.5)$$

よって、 $\mathbf{R}$  の成分はそれぞれ次のように計算することができる。

$$\mathbf{R}_{33} = \mathbf{C}_{\mathbf{D}_{33}}^{-1} \mathbf{C}_{33}, \quad \mathbf{T}_{31} = \mathbf{C}_{\mathbf{D}_{33}}^{-1} \mathbf{C}_T \quad (4.6)$$

また、カメラ座標系における計測を行うためのプロジェクタパラメータ  $P_D$  は  $P_D = P R^{-1}$  により計算することができる。

実際にこのまま計算を行うと  $R_{33}$  の行列式が 1 にならない。なぜなら、透視変換はスケールに関する情報を持たない（世界座標系の計測単位が cm 単位か m 単位かはカメラによる撮影では区別がつかない）ため、カメラパラメータやプロジェクタパラメータは全体をスカラ倍しても同一の変換を表してしまう。これでは、新たに生成したシステムパラメータによる計測単位がわからなくなってしまう。そこで、計算に用いる  $C_{D_{33}}$  や  $C_{33}$  の行列式が 1 になるようあらかじめ正規化しておき、それにより  $R$  を求めることで、もとの座標系と同一のスケールで計測が可能なシステムパラメータを生成する。

#### 4.3.3 モデル記述

対象物体の位置・姿勢を計測するためには、対象物体の形状モデルが必要であることは 3.2.1 節で述べた。また、計測しようとする対象は剛体が多数連結されたような連結物体であるため、剛体相互の連結情報についてもモデル記述が可能である必要がある。物体モデルに要求される条件には以下のような事柄が挙げられる。

- A : 対象物体の形状を正確に表現できること
- B : 対象物体の形状定義に自由度が高いこと
- C : モデルの定義が簡単に行えること
- D : 剛体同士の連結が自然に表現できること
- E : モデル内部に矛盾を含まないような構造であること
- F : 計算機により高速に処理が可能なこと

条件 A・B については、三角パッチモデルを用いることで解決できるが、その反面、複雑な形状であるとパッチ数が多数になり、定義が煩雑となる。そこで、モデル記述を階層構造にすることでこの問題を解決した。すなわち、各剛体の形状を定義するパーツ定義と、剛体の大きさ・位置・姿勢、接続情報を記述したモデル定義に分けることで、複雑な形状定義が何度もモデル定義に現れることを避けた。円筒や直方体などの単純な形状をパーツ定義に記述しておけば、モデル定義において拡大縮小することで何度も利用することができる。パーツはその高さや直径などを 1 とすることで拡大縮小率がそのままパーツのサイズとなるようにした。

条件Dについては、剛体同士の接続を木構造に限定することで混乱を避けた。おのののパーツについて、親パーツの有無と、親パーツと子パーツそれぞれの接続点を記述することで分岐数可変の木構造を定義するようにした。矛盾が起こらないようにするために、子パーツの位置情報は無視され、親パーツと子パーツの接続点が一致するように子パーツを平行移動することで初期位置を定義するようにした。

#### リスト 4.11 モデル記述の例（手のモデル）

[PALM]		パート：手のひら	
MODEL : CYLINDER		種類：円筒	
MAG (1, 3)		拡大率	
POSITION (4, 4)	4.8	2.5	
1	0	0	
0	1	0	
0	0	1	
0	0	0	
JOINT=	0	接続先なし	
[THUMB1]		パート：親指（第一関節以下）	
MODEL : CYLINDER			
MAG (1, 3)			
POSITION (4, 4)	1	1	4
0	-0.70711	-0.70711	-6.14132
0	-0.70711	0.70711	0.58071
-1	0	0	0
0	0	0	1
JOINT=	1	接続先あり	
PARENT : PALM		種類：円筒	
MJOINT (4, 1)		このパートの接続位置	
0			
0			
-2			
1			
PJOINT (4, 1)		親パートの接続位置	
-4.7271			
-0.83351			
0			

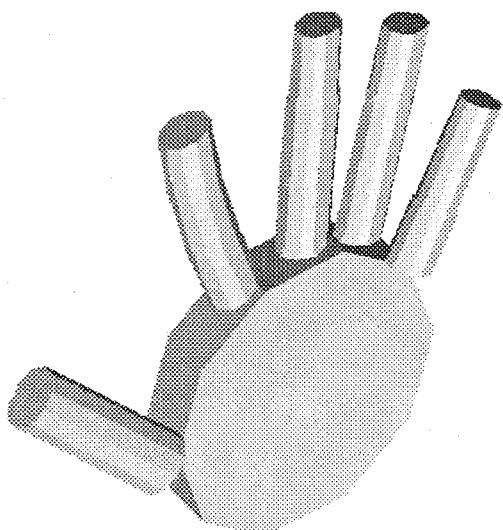


図 4.12 モデル記述を元にレンダリング表示した手の形状

計算機により高速に処理を行うため、パーツの位置・姿勢の定義は同次座標系表現をそのまま記述した。また、パーツの拡大縮小も同次座標変換により表現することができるため、計算機にモデル定義ファイルを読み込んだ直後にこれらを合成し、モデリング変換とした。対象物体の追跡処理において物体に施される回転・平行移動も同次座標変換により表されるため、これらの処理は4行4列の行列の積を取ることで実行される。また、モデリング変換とカメラパラメータ（ビューリング変換を表す）の積をあらかじめ計算しておくことで、これら二つの変換を統合することも可能である。この場合、パーツを定義する三角パッチの頂点座標を透視変換するのに、3行4列の行列との積を一度とるだけで実行することができる。

モデル定義の実例をリスト4.11に示す。著者の左手をモデリングした例である。これは15個の円筒を組み合わせることにより定義されている。また、このモデルをシェーディング表示した画像を図4.12に示す。

#### 4.3.4 変位の検出

計算機内に保持されている物体の位置・姿勢・変形を逐次近似により計測値に近づけるためには、これらの追跡パラメータの移動量を算出する必要がある。3.2.1節で述べたように、パラメータ移動量はその成分ごとに距離画像と濃淡画像で分担して検出するため、ここでは双方の処理は完全に独立している。

センサより得られる計測量は、距離画像もしくは濃淡画像であり、このままでは計算機内の追跡パラメータと直接比較することはできない。両者の違いを検出するためにはどちらかのデータに変換を施す必要がある。計測値に何らかの変換を加えて追跡パラメータと対応付けるのは困難であるため、逆に追跡パラメータから距離画像・濃淡画像を生成することにする。言い換えれば、それぞれのセンサの動作を計算機によりシミュレートすることで計算機内モデルの計測像を得ることである。図4.13のように、実際の計測系と、計算機内の画像生成系を対称な関係であると見做し、それぞれ得られた画像の差を元の実物体と計算機内モデルの差とする。ここではそれらの画像の差から、いかに移動量を算出するかを説明し、変形パラメータか

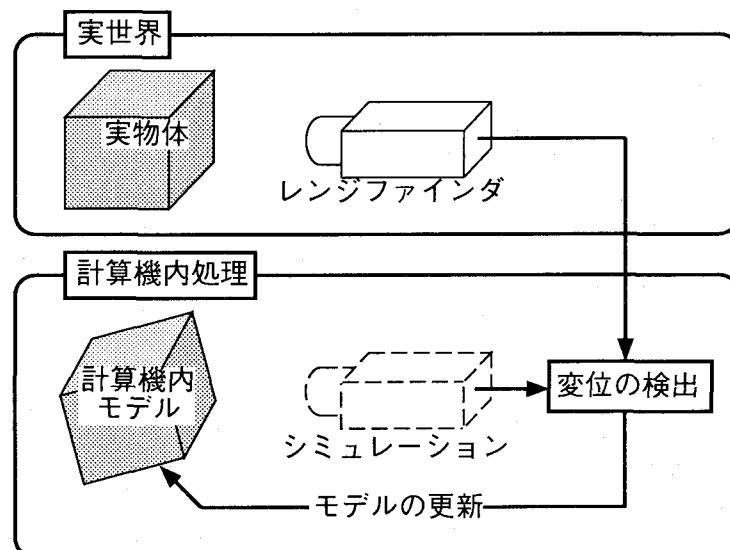


図4.13 モデルに基づく運動追跡

ら距離画像・濃淡画像を生成する方法は後述する。

### a. 濃淡画像による変位の検出

3.2.1 節では、濃淡画像によりカメラ光軸に垂直な方向の平行移動成分と、カメラ光軸周りの回転を検出するとしていた。しかし、対象物体の形状によっては回転成分の検出は困難であるため、現在は平行移動成分のみを検出している。

単一物体の平行移動成分を算出することは非常に簡単である。物体の占める領域を検出し、領域の重心の位置を検出するなどの方法が考えられる。しかし、連結物体を追跡するためには、物体の占める領域のうちどの部分がどのパーツに対応するか判断することが必要である。物体の位置があらかじめわかっているわけではないため、正確な対応付けは不可能である。しかし、計算機内モデルの追跡パラメータは処理時点において実対象とほぼ一致していると仮定し、また物体の運動速度と比較して計測速度は十分に速いと仮定しているため、計算機内モデルのパーツの配置情報をを利用して大まかに領域分割することができる。つまり、計算機内モデルから生成する濃淡画像に、どのパーツによる像であるかを示すラベルを含めておき、それを計測画像と重ねあわせることで領域分割を行う。以下では生成された画像をラベル画像と呼ぶ。

ラベル付けされた濃淡画像を用いて各パーツの変位を求めるには、各ラベルと濃淡画像の重なり具合を調べることにより検出することができる。例えば対象物体が右に移動した場合、ラ

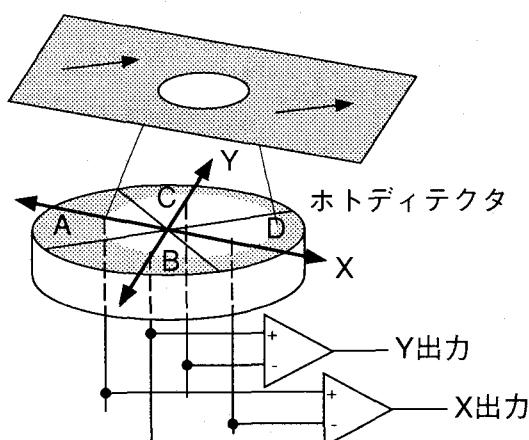


図 4.14 4 分割フォトディテクタによる変位の検出

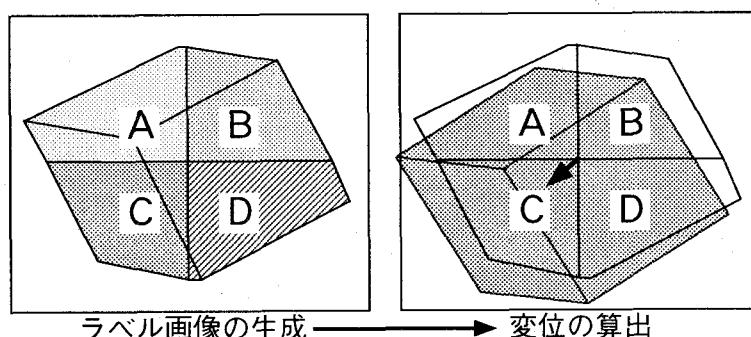


図 4.15 ラベル画像による変位の検出

ベル付けされた画像はラベル領域中の右に偏り、ラベル領域の左方には背景が撮影された部分が生じる。これらの偏りを検出する方法は各種考えられるが、ここでは高速化のため画像処理ハードウェアによる実行が可能な手法を考案した。

位置の変位を検出する素子にホトディテクタがある（図4.14）。これは受光した光量を電気信号に変換して出力する素子であり、通常検出領域が4分割されておりそれぞれの受光量の違いを比較することで変位を検出する。通常は受光素子の出力を単に加減算することで比較を行う。そこで同様に、一つのラベル領域を一つのホトディテクタのように考え、4分割した領域ごとに対象物体の像の占める面積を求めて移動量を検出した。重複領域の検出および面積計算はハードウェアにより高速に実行することが可能であるため、画像処理装置から原画像を取り出すことなく、直接各受光素子ごとの重複面積を得ることができる。

受光領域の配置は、パーツの重心を中心に4つの象限に分割する（図4.15）。元のパーツ定義は原点を中心にして構築されているため、単に原点座標をモデリング・ビューアング変換することでラベル画像におけるパーツ中心を求めることができる。象限ごとに分割された検出領域を用いて左右の移動量を求めるには、右半分の検出領域（BとD）の受光量から左半分の検出領域（AとC）の受光量を引き算する。上下についても同様に求める。実際には各検出領域の面積が異なっていることがあるため、各領域における受光面積比を用いて演算する。

ハードウェアによりここまで処理を行う手順を図4.16に示す。計測に先立ち、計測領域の背景画像を取り込んでおく。物体存在領域を計算するには、得られた画像から背景画像を減算し、その絶対値が閾値を超えた領域を物体の占める領域としている。次に、物体存在領域の画素値を0x80に、物体の存在しない領域を0x00に設定する。他方、生成されたラベル画像は、各受光領域ごとに画素値を割り振る。具体的には、第1パーツには0x01から0x04、第2パーツに0x05から0x08というふうに各パーツごとに画素値が4つづつ割り当てられる。受光領域のない領域は0x00が割り当てられる。

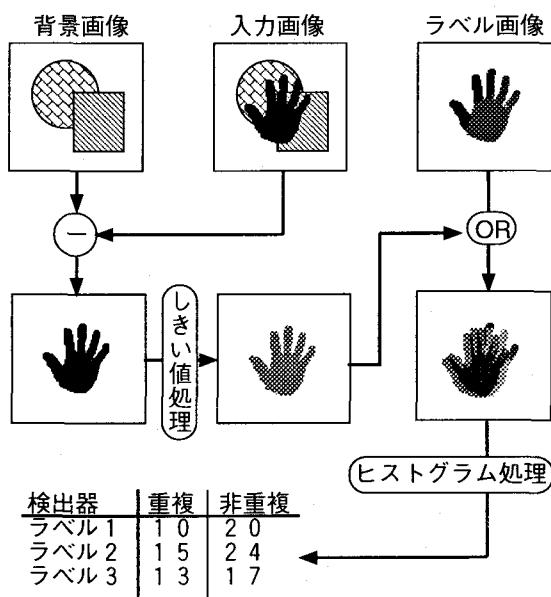


図4.16 ハードウェアによる重複面積の計算

次に、計測画像とラベル画像の論理和を各ビットごとに取る。物体存在領域は最上位ビットに1が立っており、背景領域は最上位ビットが0となる。例えば、ラベル番号0x25の領域のうち、物体が重複している領域は0xA5となり、背景が撮影されている領域は0x25のままである。

最後にこの画像に対して、全ての画素値についてヒストグラム処理を行う。結果として、各受光領域ごとに、物体が重複した領域と背景が撮影された領域の画素数がそれぞれ得られる。これを取り出してホストCPUにより面積比率を算出する。現在用いている画像処理ハードウェアは、画像の減算に1ステップ、画像の閾値処理に1ステップ、論理和の計算とヒストグラム処理に1ステップの合計3ステップで計算を行い、これらの処理に30msしか必要としない。

### b. 距離画像による変位の検出

距離画像により求めるべき追跡パラメータの変化量は、カメラ光軸方向の平行移動と、カメラ光軸に垂直な軸周りの回転である。これらの運動を力学的に考えると、剛体に対してカメラ光軸方向の力が加えられた場合に発生する一般的な運動がこれらの運動と対応する。そこで、対象物体の各パーツについて、計測された距離画像と計算機内モデルとの差分から、パーツに及ぼすべき力を算出することにする。濃淡画像の場合と同様に、計算機内の追跡パラメータと計測された距離画像を直接比較することはできないため、追跡パラメータと形状モデルを利用することにより距離画像を生成し、計測により得られた距離画像と比較することで変化量を算出する。

剛体に作用する力の大きさは、単に計測された距離画像と計算機内で求められた距離画像の距離値の差をそのまま用いた（図4.17）。それぞれの力がどのパーツに及ぼされているのかの判定は、濃淡画像による変位検出で用いたラベル画像を利用して行なった。一つのパーツに作用するいくつかの力は、0次モーメント（平均値）と、XY軸方向それぞれの1次モーメントを求ることで、前者からは平行移動成分に対応する変化量を、後者からは各軸周りの回

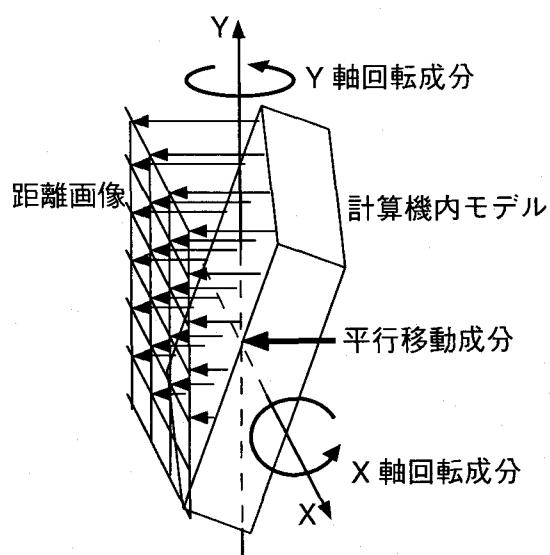


図4.17 距離画像による変位の検出

転成分に対応する変化量をそれぞれ求めることができる。

#### 4.3.5 追跡パラメータによるラベル画像・距離画像の生成

前節で述べた、画像の比較による変化量算出を行うためには、計算機内モデルによるラベル画像・距離画像の生成が必要となる。画像の生成方法にはレイトレーシングに基づく方法などもあるが、ここでは生成速度を重視するため、ポリゴン描画による画像生成を行った。

生成する画像は計測する画像と同一の視点による像でなければならないため、透視変換のパラメータはカメラパラメータを直接用いる。3.2.3節で述べたように、モデリング・ビューリング変換は行列の積を一度とるだけで実行が可能なため、パーツ定義における頂点座標をピクセル座標に変換することは簡単である。二次元画像上に展開された三角パッチは、デジタル微分解析(DDA)を用いた方法により描画される。このとき同時に、パーツ重心を原点とした象限ごとに異なるラベル番号を描画する。描画する内容はパーツ番号に対応したラベル番号であるため、一つのパーツにつき隠面消去を行う必要はないが、パーツの表裏を判定することにより描画ポリゴン数を半分に減らすことができる。パーツ同士のオクルージョンを解決するには、単にカメラからパーツの重心への距離を算出し、遠くにあるパーツから描画することで行われる。

距離画像生成の代表的な手法は、主にグラフィックスワークステーションにおいて利用されているZバッファ描写がある。これは、まず三角パッチの各頂点の奥行き値を求め、パッチを描画する際に線形に補間することでピクセルごとの奥行き値を求める方法である。しかし、この手法は計算量が膨大であり、実用的な速度でZバッファを描画するには、線形補間ハードウェアが必要となる。そこで、濃淡画像による方法とは逆に、レイトレーシング的手法により各画素における視線についての奥行き値を求めた。得るべき距離画像はシリコンレンジファインダと同じ解像度 $24 \times 24$ 画素であるため、かえって線型補間法に比べ計算量を減少させることができる。以下では、この処理を高速化するための式展開について詳説する。

カメラ座標系において、ポリゴンは以下のように定義される。 $v_3$ は三角パッチの一つの頂点の座標であり、 $v_1$ と $v_2$ は頂点 $v_3$ から別の二つの頂点への相対ベクトルである。三角形の境界条件より、パラメータ $s$ と $t$ には三つの条件式が成立する必要がある（図4.18）。

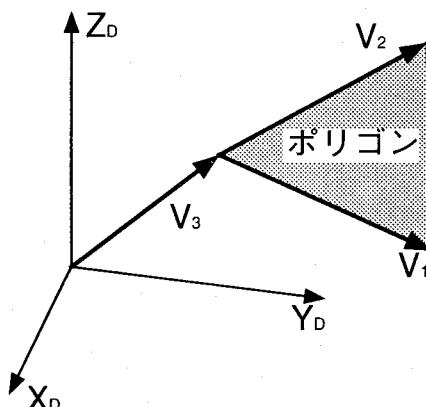


図4.18 ベクトル方程式によるポリゴンの表現

$$\begin{bmatrix} x_D \\ y_D \\ z_D \end{bmatrix} = v_1 s + v_2 t + v_3, \quad \begin{cases} s+t < 1 \\ s > 0 \\ t > 0 \end{cases} \quad (4.7)$$

ここで、同次座標表現を用いるため、 $v_i$  を次のように 4 行 3 列の行列  $\mathbf{V}$  にまとめる。

$$\mathbf{V} = \begin{bmatrix} v_1 & v_2 & v_3 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.8)$$

レンジファインダによる撮影は、カメラパラメータ  $\mathbf{C}_D$  により次のように表現されるため、三角パッチの頂点とピクセル座標との関係は以下のようになる。

$$\begin{bmatrix} h \cdot x_s \\ h \cdot y_s \\ h \end{bmatrix} = \mathbf{C}_D \begin{bmatrix} x_D \\ y_D \\ z_D \\ 1 \end{bmatrix} = \mathbf{C}_D \mathbf{V} \begin{bmatrix} s \\ t \\ 1 \end{bmatrix} \quad (4.9)$$

行列  $\mathbf{C}_D \mathbf{V}$  は 3 行 3 列となるため、逆行列を求めることにより、ピクセル座標からパラメータ  $s, t$  を求めることができる。

$$\begin{bmatrix} s \\ t \\ 1 \end{bmatrix} = h \cdot (\mathbf{C}_D \mathbf{V})^{-1} \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} \quad (4.10)$$

ここで、 $s$  と  $t$  が式 4.7 の条件を満たさない場合、この三角パッチはこのピクセルにより撮影されていないことになる。 $s$  と  $t$  が条件を満たす場合、再度  $\mathbf{V} [s \ t \ 1]^T$  を求めることにより、世界座標系における視線とポリゴンの交点を算出することができる。交点座標の  $x_D$  が奥行き値である。

シリコンレンジファインダのピクセル  $(x_s, y_s)$  ごとに全てのポリゴンを計算するのは非効率的であるため、まずラベル画像を参照し、どのパーツについて計算すべきか判定する。つぎに、このパーツの全てのポリゴンについて交点を求め、最もカメラに近い奥行き値をそのピクセルの距離値とする。

#### 4.3.6 パーツ相互作用とモデルの移動

計測された濃淡・距離画像を利用することにより、各パーツの追跡パラメータの変化量は 6 個の成分ごとに求めることができる。パーツ同士の相互作用についても同様に、剛体運動の 6 成分として変化量を計算する。計算モデルは単に、モデル同士の接続点をバネで接続したような力学的モデルを利用する（図 4.19）。パーツの重心は既知であるため、バネによる外力について、モーメントを求めて回転成分が、そのまま加算することで並進成分がそれぞれ求

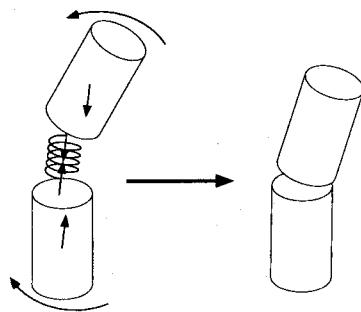


図 4.19 バネモデルによる連結部分の表現

められる。これらの力を計測による変化量に加算し、剛体の運動を力学的に求めることで最終的なパーツ運動を算出する。

#### 4.3.7 追跡処理とカメラ座標系

3.2.2 節では、追跡パラメータの変化量をカメラ座標系において算出する利点を述べた。また、3.2.5 節では、高速な距離画像の生成にカメラ座標系を用いた。しかし、実際の計測では、物体の運動を表現する座標系を任意に定めることができる方が都合がよい。そこで、追跡物体の初期位置の定義や、追跡結果は世界座標系において入出力することにする。計算機内で処理に用いる追跡パラメータを世界座標系で保持することは、処理の度にカメラ座標系への変換が発生して速度の低下が発生するため、入出力のみにおいて座標変換を行い、内部処理はすべてカメラ座標系で統一した。処理の流れと座標系の関係を図 4.20 に示す。

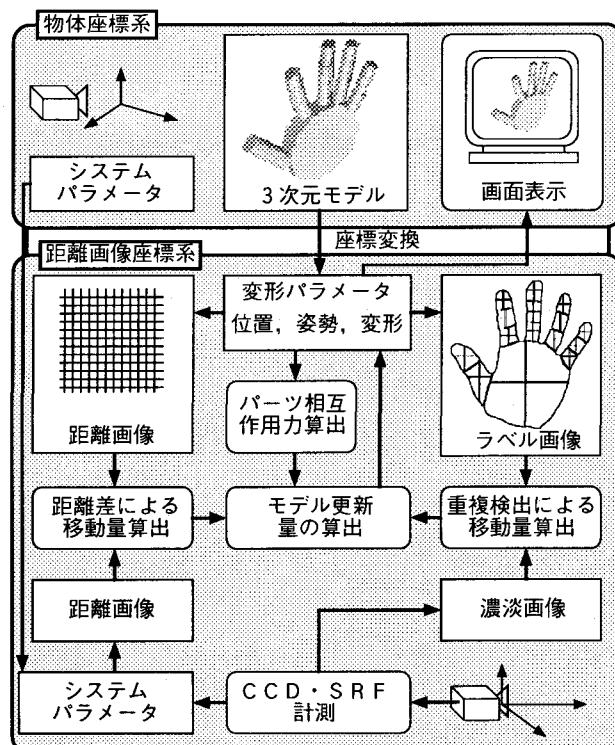


図 4.20 処理の流れと座標系の関係

## 4.4 実験

この節では、4.3で説明した動物体追跡の原理をもとに、シリコンレンジファインダを用いて計測した動距離画像と濃淡画像による剛体・連結物体の追跡実験を示す。実験の厳密な評価のためには、物体を構成する剛体それぞれについて剛体運動6成分の真値と計測値を比較する必要があるが、これらの成分をすべて動的に位置決めすることは困難である。そのため、動物体の追跡がリアルタイムに行われていることを示す実験と、いくつかの運動成分についての精度評価を行う実験に分けてその結果を示す。

### 4.4.1 実験装置の構成

実験に用いた機器の構成を図4.21に示す。シリコンレンジファインダにより取り込まれた濃淡画像と距離画像をApollo DOMAIN DN10000VSに取り込み、計測・追跡処理を行う。Ethernetで接続されたグラフィックスワークステーション SiliconGraphics IRIS Indigo XSにより、追跡結果をリアルタイムに表示する。

処理の流れと各機器の関係を図4.22に示す。距離画像は直接10000VSのパラレルインターフェースボードにより取り込み、CPUで処理する。濃淡画像は、画像処理ハードウェアTAICHI-COREにより面積算出までの処理を行うため、取り込まれた画像をCPUにより処理することはない。TAICHI-COREは、8近傍についての加減算処理などを1画素あたり40nsで行うことができる。取り込まれた画像の大きさは $512 \times 480$ 画素であるため、1フレームの

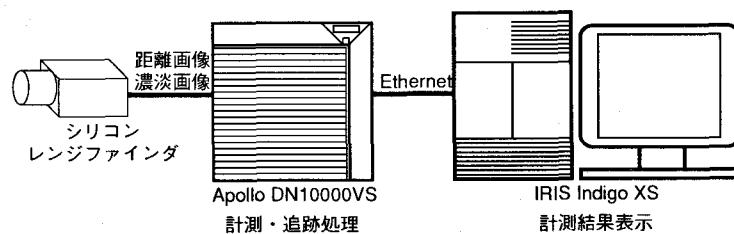


図4.21 動物体追跡システムの機器構成

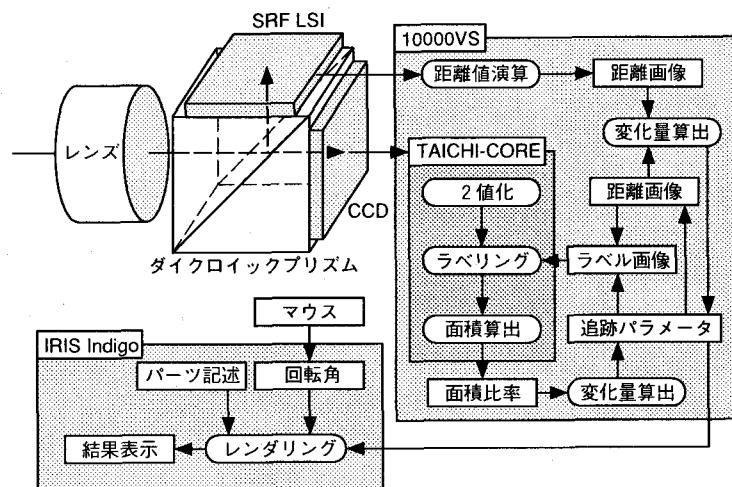


図4.22 各機器と処理・データの関係

処理を 10ms で行うことができる。

処理の進行に従い連続的に得られる追跡パラメータは, Ethernet を介して Indigo に転送される。Indigo では, 世界座標系内を移動する物体をリアルタイムレンダリングにより表示を行い, マウスにより視点を変更することができる。

#### 4.4.2 剛体の位置・姿勢の追跡実験

単一物体の追跡の例として, 直方体の追跡の例を挙げる。本論文で述べた手法では, 光軸周りの回転を検出することができないため, ここでは平行移動と光軸に垂直な軸に対する回転による追跡実験を行った。直方体の大きさは, 7cm × 4cm × 1.5cm である。図 4.23 のように, 画面上にリアルタイムに計測物体の姿勢が 3-D 表示される。追跡出力のうちの 9 フレームを図 4.23 に示す。追跡処理を 1 ステップ進めるために要する時間は約 0.3 秒である。物体の移動速度が限られた範囲内にある限り, 正しく追従することが確認された。

#### 4.4.3 連結物体の位置・姿勢・変形の追跡実験

連結物体の例として, 前出の直方体を二つ連結したオブジェクトを用いて実験した。追跡出力のうちの 9 フレームを図 4.24 に示す。現在は連結情報として 3 自由度を許しているために二つのオブジェクトがねじれてしまうことがあったが, 特に速く動かさない限り正しく追従することができた。追跡処理を 1 ステップ進めるために要する時間は約 0.6 秒である。

#### 4.4.4 精度評価実験

追跡の結果得られたパラメータについて, 真値と比較することによりその精度を評価した。

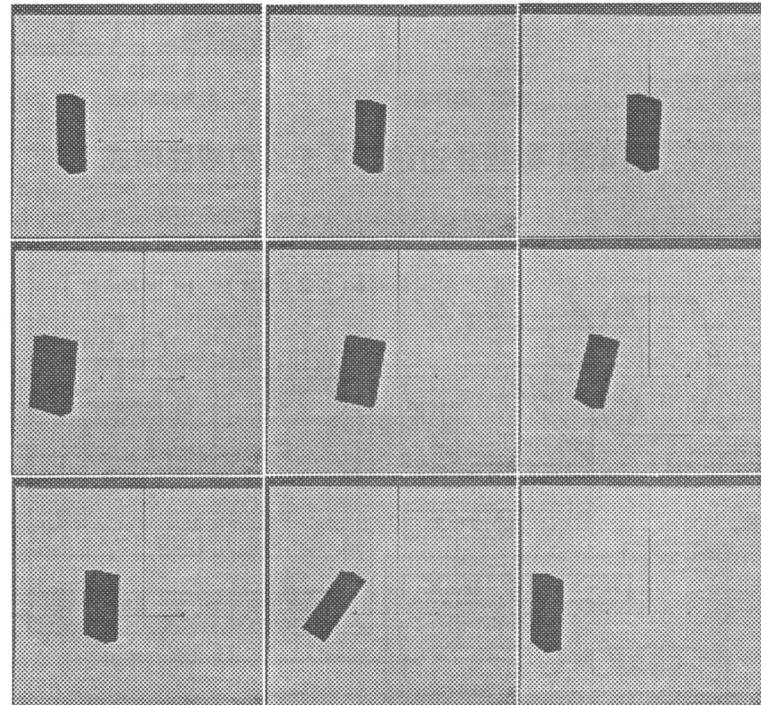


図 4.23 直方体の追跡実験結果

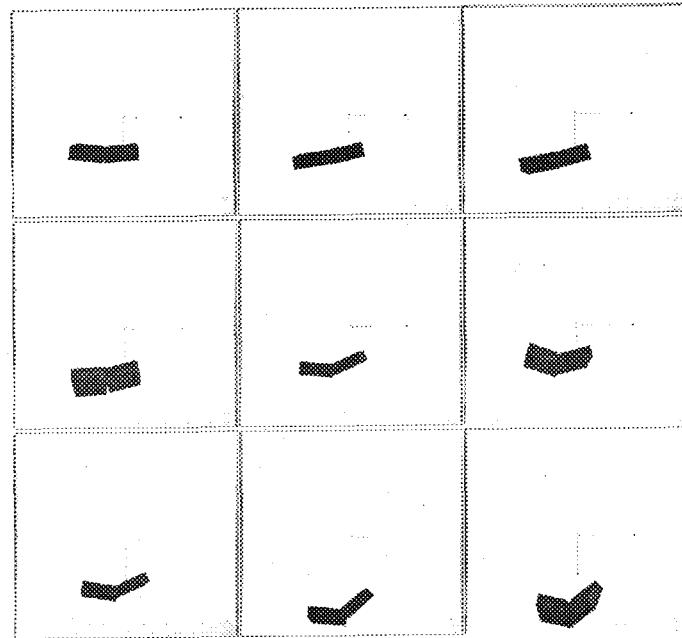


図 4.24 連結物体の追跡結果

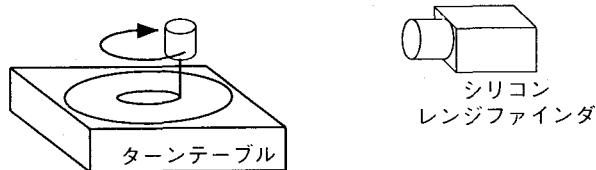


図 4.25 物体位置追跡実験の機器構成

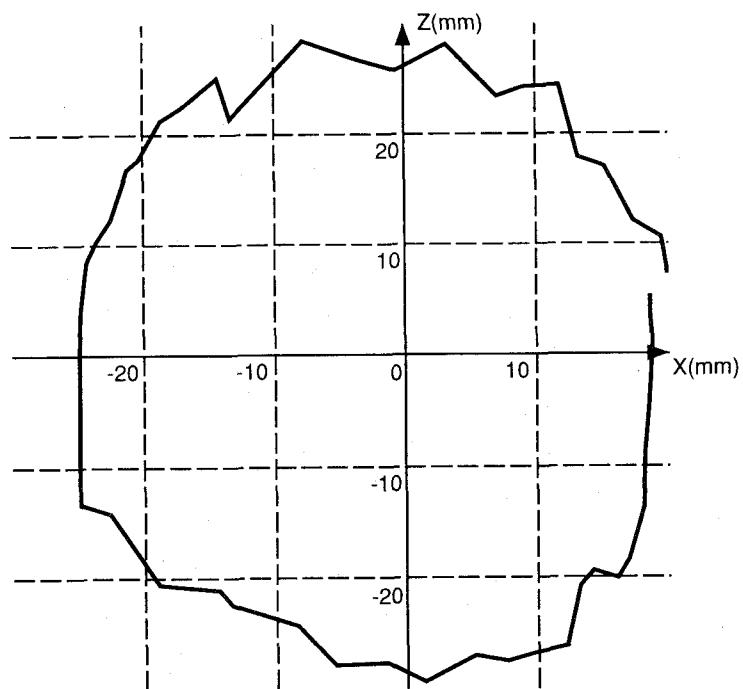


図 4.26 物体の軌跡の追跡結果（Z：光軸方向）

対象物体の移動・回転には共にターンテーブルを用い、物体の中心の軌跡が円になることと、物体の回転角が正しく得られているか検証した。

まず、ターンテーブルに円筒形の物体（半径 16mm、高さ 50mm）を、回転軸から 25mm ずらして固定し、その運動軌跡を調べることで位置精度の評価を行った（図 4.25）。追跡の結果得られた円筒の軌跡を図 4.26 に示す。得られた物体位置と円（真の軌跡）との距離の誤差の標準偏差は 3.2mm であった。

次に、ターンテーブルに薄板（幅 72mm、高さ 47mm）を固定し、ターンテーブルの回転角と追跡の結果得られた回転角を比較した（図 4.27）。真値と計測値、誤差の関係を図 4.28 に示す。誤差の標準偏差は 3.2 deg であったが、誤差の分布がランダム性ではないため、キャリブレーションパラメータの精度等に起因する系統誤差であると考えられる。

## 4.5 考察

本節では、連結物体の実時間追跡について、追跡精度や計測速度等全般について考察を行う。また、現在のシステムの問題点や改良すべき点、これからの研究の指針についての考察も行う。



図 4.27 回転角追跡実験の機器構成

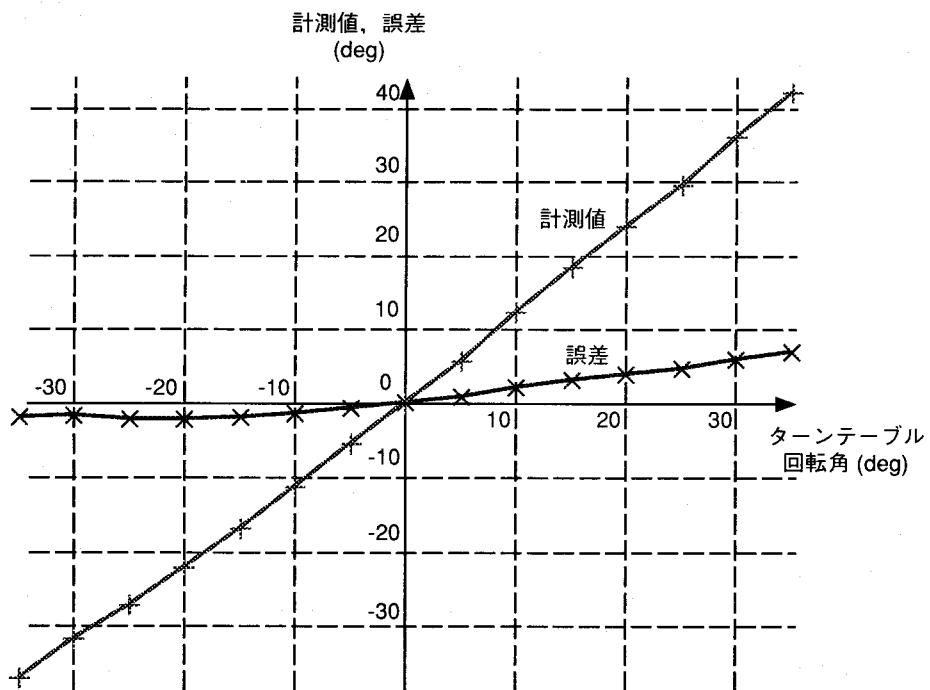


図 4.28 回転角の計測結果

#### 4.5.1 位置・姿勢の追跡について

対象物体を単に平行移動するだけの実験では、濃淡画像による二次元的な物体の追跡と、レンジファインダによる距離の単なる組み合わせであると考えることもできる。しかし、物体の絶対位置を計測結果から導出するのではなく、物体の変位を検出することで物体の追跡を行っているため、濃淡画像のみでは世界座標系における物体の移動量は求められない。なぜなら、物体が画像上で右に 10 ピクセル移動することが、カメラから 10cm の距離にある物体が右に 1cm 移動することに相当するのであれば、カメラから 20cm の距離にある物体が右に 2cm 移動することも同様に画像上では 10 ピクセルの移動として現れるからである。この問題を解決するために、前のステップまでの処理により推定されている物体の奥行き位置を利用し、実際の空間における物体の移動量を算出している。また、正しくラベル画像を生成するためにも、物体の奥行き値が正しく得られている必要がある。物体の奥行き方向の推定位置が異なっていれば、物体の像の大きさが実際の像の大きさと異なってしまうためである。

物体の距離を同定するためには、距離画像においてどの画素がどの剛体と対応するのかを判断する必要がある。そこで、距離画像の各画素に対応するラベル画像上の画素値（ラベル値）を用いて距離画像をセグメンテーションしている。このように、濃淡画像による追跡と距離画像による追跡は前ステップまでの追跡結果を互いに参照し合っている。一つのステップの処理は互いに独立であるが、その結果はステップの終了時点で一つのパラメータに統合されるため、完全に独立した評価を行うことはできない。

物体がカメラ光軸と垂直な方向に高速に移動した場合、前ステップまでの追跡結果より生成したラベル画像との共通部分がなくなってしまうことがある。このような現象は物体が小さいほど発生しやすいため、物体の像が大きいほど追跡の安定性は増す。物体の移動速度が高速すぎる場合、光軸と垂直方向の追跡が正しく行われなくなるだけではなく、物体の奥行き推定位置が背景と同じ位置になってしまることがある。このような現象を防ぐため、物体の像とラベル画像が重複していない領域に含まれる距離画像は用いないなどの処理を行っている。

物体の回転角計測については、計測安定性は物体の形状に大きく左右される。例えば、物体の形状が薄い板状のものである場合、物体の面がカメラ方向を向いているときには安定に回転角を計測することができるが、物体面の法線ベクトルがカメラ光軸に対して垂直に近づくにしたがって計測の安定性が減少する。また、円筒のような物体では、円筒の軸に関する回転は距離画像により計測することができない。

物体の光軸周りの回転は、現在では検出していなかったため、理論的には推定値もカメラ光軸周りの回転を生じないと考えられる。しかし、カメラ光軸に垂直な軸周りの回転速度が高速すぎたり、計測安定性が減少した場合、これらのベクトルの合成順序の影響で物体がカメラ光軸周りに回転してしまうことがある。カメラ光軸周りの変位を検出可能にすることで、これらによる誤差を減少させる方向に変位が検出され、計測安定性が向上すると考えられる。

#### 4.5.2 変形の追跡について

連結物体の変形は、その関節の両側の物体の位置・姿勢の関係として得られる。関節両側の物体が極端に離れてしまわないので、両側の物体が関節部分でバネにより接合されているも

のとしている。しかし、両側の物体の姿勢に制約を設けていないため、物体同士が実物体では起こり得ないねじれを起こしてしまうことがある。人体のように複雑に変形する物体でも3自由度の関節は数少なく、ロボット等の人工物ではほとんど見られないため、関節の自由度に制限を設けることは計測安定性の向上に大きく寄与すると考えられる。

#### 4.5.3 計測精度について

4.4.4で示した円筒の位置検出結果は、物体の移動範囲の大きさ（半径25mmの円）に対して大きな誤差を含んでいると思われる。しかし、ここで行った実験は、物体の中心の軌跡を求めており、物体が大きいほど精度の確保が困難になると思われる。実験に使用した円筒の半径は16mmであり、求められた軌跡と円との距離の標準偏差が3.2mm程度であったことから考えると、物体の中心座標が精度よく求められていると考えることもできる。

物体の回転角の精度評価実験では、全体に滑らかな計測結果を得ることができた。しかし、全体的に計測値の絶対値が真値と比べて大きい傾向が見られる。原因としては、レンジファインダのシステムパラメータと距離画像生成パラメータの相違により、生成された距離値のスケールがやや大きいことが考えられる。

#### 4.5.4 計測速度について

現在のシステムでは、単純な物体の追跡のために要する時間は約0.3秒である。しかし、濃淡画像・距離画像による変位の検出は、約40msしか要しておらず、処理時間の大部分を占めているのが、計測値との比較に用いるラベル画像と距離画像の生成である。これらの処理を高速化することで1ステップの処理を0.1秒以内に行うことができると考えられる。

## 5章 複数距離画像の生成による 自由曲面物体の実時間追跡

4章では、距離画像と濃淡画像を融合することによって、幾何学形状物体の姿勢を実時間追跡する手法を紹介した。しかし、対象物体の形状によっては安定に追跡できないなどの問題点があった。この章では、これらの問題点を明らかにするとともに、距離画像単体でも6軸全ての変位を検出することのできる手法を導入する。そのためには複数の距離画像を計算機内モデルから生成する必要があるが、速度の低下をきたさないためにCGハードウェアを用いた距離画像生成を行う。追跡と同時平行して対象の形状をモデリングすることによって、あらかじめ対象の形状モデルを用意せずとも対象の6軸運動を求める手法についても詳説する。

### 5.1 自由曲面の位置合わせ

距離画像を用いて物体の運動を追跡する問題 (tracking) は、複数の表面形状モデル同士を位置合わせする問題 (registration) と共通する点が非常に多い。なぜなら、距離画像は対象の三次元形状に関する情報を直接的に保持しており、簡単な処理によって様々な形状モデルへ変換することが可能である。また、物体がレンジファインダに対して運動するということは、相対的には計測対象に対してレンジファインダの視点が移動することと等価である。同一の表面形状に対して互いに異なる視点から距離画像が観測された場合、双方の視点から共通に計測された領域に関しては、同一の形状が計測されるはずである。それら複数の距離画像の共通領域が、それらの形状の同一性を根拠として位置合わせされれば、それぞれの距離画像を計測した時点の視点位置を求めることが可能であり、すなわちレンジファインダに対する対象の相対位置の変化が計測されたと考えることができる。

動物体の姿勢を連続的に追跡する問題と、物体の全周形状をモデリングするために位置合わせを行う問題とでは、得られる情報として一組の距離画像以外にも共通のものが存在する。その一つとして、位置合わせに先立ちおおよその相対的な位置関係が既知であると言うことが挙げられる。動物体の運動による姿勢の変化は連続的であり、瞬時に姿勢が変化することはありえないから、十分に短い間隔で計測された距離画像同士については、その間の姿勢の変化は十分に小さいという仮定をおくことができる。そのため、相前後する2つの距離画像同士の姿勢の相違は微小であると言える。また、全周モデリングについては、レンジファインダの位置・計測方位を変化させる装置や、物体の姿勢を変化させる装置は計算機により制御されることが多い。そのような場合、移動装置そのものや、移動装置とレンジファインダの位置関係の精度が十分でない場合にのみ位置合わせが必要となる。また、人手によりレンジファインダや対象の姿勢を変更したような場合でも、人手による入力によりおおよその位置を既知とすることはそれほど非現実的であるとは言えない。

実際、距離画像同士の位置合わせを行うアルゴリズムの大半は、逐次最適化に基づく手法である。この手法では、局所解に陥りやすいと言う欠点を避けがたい。しかし、図5.1に示すよ

うに、周期的形状が大部分を占めるような対象について、最適解を自動的に発見する手段を探ることが良いとは一概に言えない。なぜなら、前述したように最適解に近い初期位置が与えられる場合が多いこと、また十分よい初期位置を与えるコストは最適解を全探索するコストに較べ大きいとは言えないからである。よって、十分よい初期位置から逐次的に最適化を繰り返すことによって最適解を得ようとする手法は十分に一般的であると言える。

### 5.1.1 ICP

逐次最適化により三次元空間内の形状モデル同士を位置合わせする手法として、最も有名なものに Besl と McKay による ICP (Iterative Closest Point) アルゴリズムがある [Besl92]。この手法では、以下の処理を繰り返すことによりモデル同士の相対位置を求める（図 5.2）。

1. 一方のモデル上を、適当な間隔でサンプリングする。
2. 全てのサンプル点に関して、もう一方の形状モデル上で最も距離が短い点を対応付ける。
3. 全ての対応付けられた点同士の距離の自乗和が最小となる姿勢パラメータを求める。
4. 2 以下を繰り返す。

文献 [Besl92] では、形状モデルとして、距離画像処理で用いるような点群や三角パッチモデルの他に、1 値関数や陰関数、パラメータ表現などの数学的形状表現手段についても言及しているが、これらの相違が影響するのは 2 の処理のみである。3 の処理は、3 次の回転行列を  $R$ 、平行移動ベクトルを  $t$  とすると、これらを最小化する指標は

$$\min_{R,T} \sum_i |\mathbf{R}\mathbf{x}_i + t - \mathbf{p}_i|^2 \quad (5.1)$$

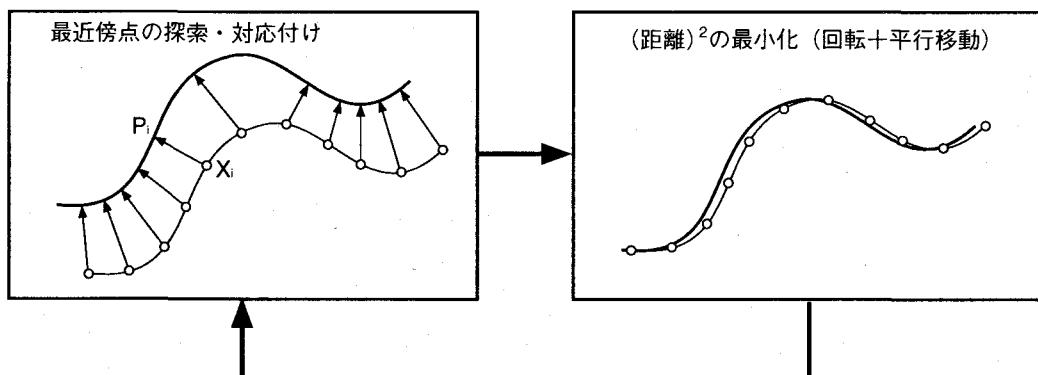


図 5.2 ICP Algorithm

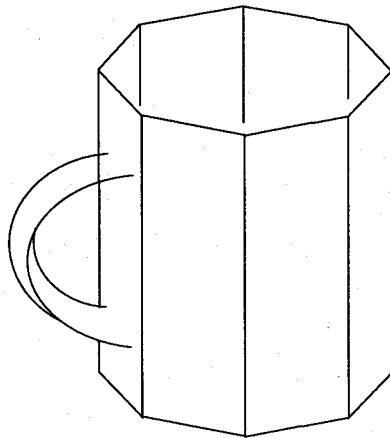


図 5.1 局所解の多い物体

と表される。ここで、 $X_i$  はサンプル点、 $P_i$  はそれに対応付けられた最近傍点である。この式の最小化手法の解析は様々な研究者により行われており、絶対収束であることが分かっている。つまり姿勢パラメータの最適化が行われた結果距離の自乗和は減少するが、それぞれの対応付けについてはもはや最近傍点ではなくなっている可能性がある。そこで、2 の処理を再度行うことにより、前回よりもさらに距離が短いような対応付けを調査する。結果的に、2, 3 の処理双方で距離の自乗和が増加することはありえないため、全体の処理としても絶対収束が保証されている。

ICP は非常に一般性の高い方法であるため、様々な研究者により改善・発展が行われてきた。ICP ではすべてのサンプル点について対応付けを決定するが、実際には距離画像を撮影する視点の相違によって、一方ではまったく計測されていないような点について生成された偽の対応付けが結果に深刻な影響をもたらす。そこで増田は、自乗和に基づく最適化ではなく、中間値 (median) を最小化することにより、50% 以上の共通領域が計測されれば安定に姿勢が求められる手法を提案している [Masuda96]。また、3つ以上の距離画像を位置合わせする手法としては、Stoddart らが行った研究がある [Stoddart96]。この手法では、全ての距離画像同士の組み合わせについて最近傍点を求めておき、それらの全ての距離の自乗和が最小となるような全ての距離画像間の運動パラメータを力学モデルにより求める手法である。運動パラメータは距離画像の枚数から 1 を減じた数だけ同時に求められる。

ICP を動物体追跡に用いるために最も障害となるのは、2 の計算量である。距離画像の計測点数が  $n$  である場合、最近傍点の組み合わせを全て求めるためには  $O(n^2)$  の計算が発生する。また、計測点の数が比較的少ない場合、点群同士の対応付けのみでは対応づけられた点同士の距離が十分に短くならないため、安定に最小化が行えない。そこで、一方を三角パッチモデルに変換してから対応付けを求めるが、一般に点との距離が最小となるような三角形上の点の位置には三角形内、3辺、3頂点の7つの場合があり、計算量が多い（図5.3）。また、このような処理はハードウェア処理に適しておらず、メイン CPU により逐次的に行うほかはない。Simon は、シリコンレンジファインダと同等の解像度・計測速度を有するレンジファインダについて ICP を適用すべく、Kd-Tree 法や巧妙なキャッシングなど様々な計算量削減手段を講じて Sun の SparcStation10 上で最大 10Frame/秒のパフォーマンスを得た [Simon94] が、これは

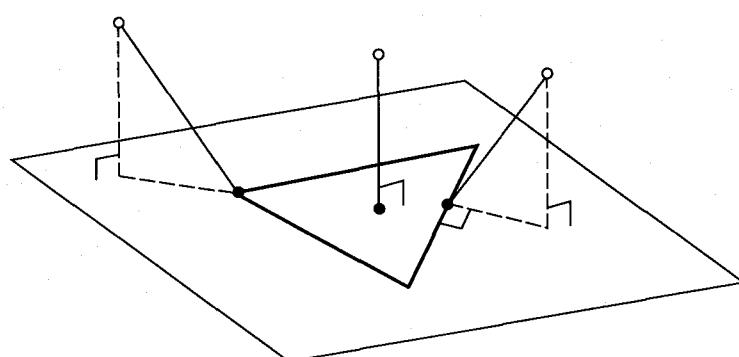


図 5.3 三角形と点の最近傍点検索における場合分け

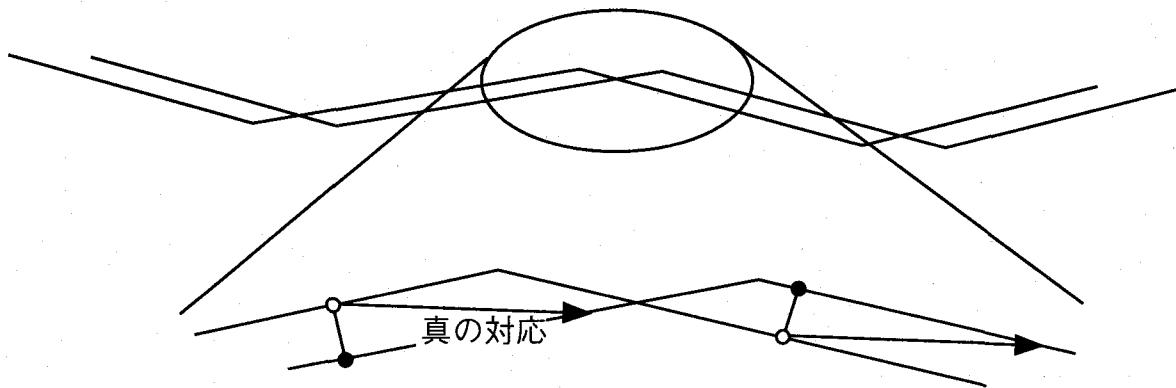


図 5.4 ICP では収束が遅い形状の例

あくまでも膨大な前処理や2回目以降の処理におけるキャッシングが有効に働いた場合であり、當時安定にこの性能が発揮されるとは考えにくい。

ICP は絶対収束が保証されているが、これは必ずしも収束速度が速いということを意味するものではない。性質上、対象の形状によって収束性能が極端に低下する場合が存在する。例えば、図 5.4 のように、面に平行な方向へのずれが残存しており、最近傍点が真の対応点と著しく異なるような場合は、多数の繰り返し処理を要するため、微小な面の歪みなどを手掛かりに物体の姿勢を追跡するような用途には向かない。

### 5.1.2 4 章で提案した手法の問題点

4 章では、濃淡画像と距離画像を融合することにより、比較的簡単かつ高速な処理で対象の位置と姿勢を求める手法を提案した。しかし、これをそのまま自由曲面物体へ応用することはたいへん困難である。なぜならば、4.3.1 で述べた、距離画像と濃淡画像の役割分担は、対象の形状に関する依存性がたいへん高いからである。例として、球や円筒、立方体についてこれら

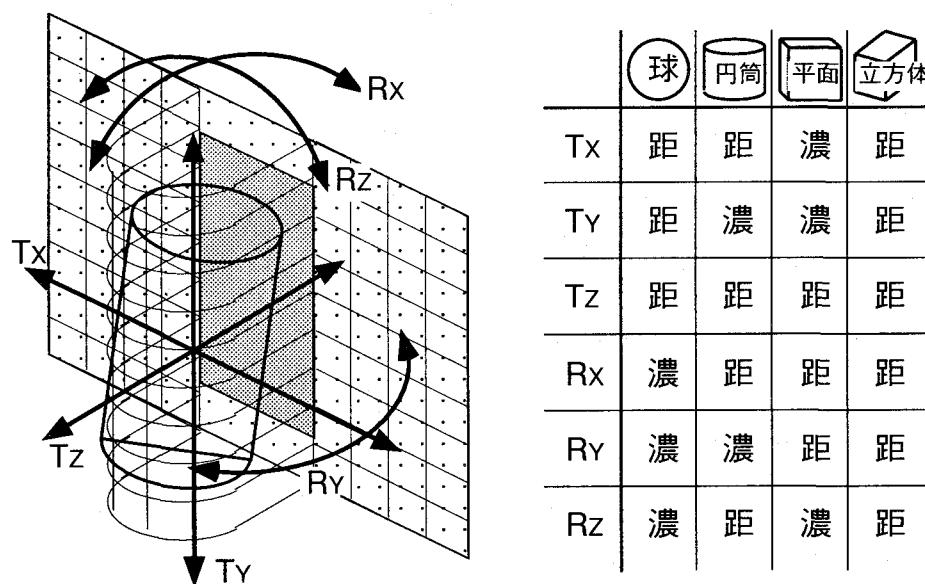


図 5.5 各自由度の検出における濃淡画像・距離画像の適・不適

の姿勢を求めるために、距離画像と濃淡画像のどちらが有効かを図5.5に示す。立方体の例で分かるように、対象の形状だけではなく、その姿勢によっても濃淡画像と距離画像の有効性の順位が入れ替わることが分かる。また、球体の場合では、テクスチャ情報がない限り姿勢の推定は不可能である。このように、濃淡画像と距離画像のどちらを用いるのが良いかは、対象物体の形状や姿勢、物体表面のテクスチャの種類や方向によって変化する。特に、自由曲面物体についてはこれらの順序関係がたいへん複雑であるとともに、濃淡画像についても輪郭形状とテクスチャのどちらが有効かが判然としない。そこで、濃淡画像の利用については5章では扱わないこととし、距離画像単体により対象の6軸姿勢を求める手法を開発することとする。しかし、4章で述べた方法では、距離画像の相違として光軸方向の距離差を用いたが、この手法では光軸に垂直な方向の平行移動及び光軸まわりの回転を扱うことができない。他方、ICPのように光軸方向以外の部位へ対応付けを拡張していく方法では、計算量が大きくなるという問題があることは既に述べた。これらはともに、1枚の計測距離画像に対して、唯1枚の距離画像を生成・比較することに起因する問題点である。そこで図5.6のように、1枚の計測された距離画像に対して、複数枚の距離画像を生成し、それらを総合的に比較した結果として6軸姿勢を求める手法を開発する。そのためにはまず、モデルからの距離画像生成の速度向上を図る必要がある。次節では、まずその問題の解決手法から解説する。

## 5.2 CG ハードウェアによる距離画像生成の高速化

エンジニアリング分野では現在、計算結果等を視覚的に表現する必要性が高いために、高速に三次元CGを表示するためのハードウェアを搭載したワークステーションが普及してきている。通常、CGとは光源の配置や物体表面の反射率・テクスチャなどから計算される輝度情報であるが、三次元で表現されているモデルを二次元ビットマップへ矛盾なく変換するために、様々な座標計算を行う必要があり、これらの高速化が三次元CGハードウェアの要となってい

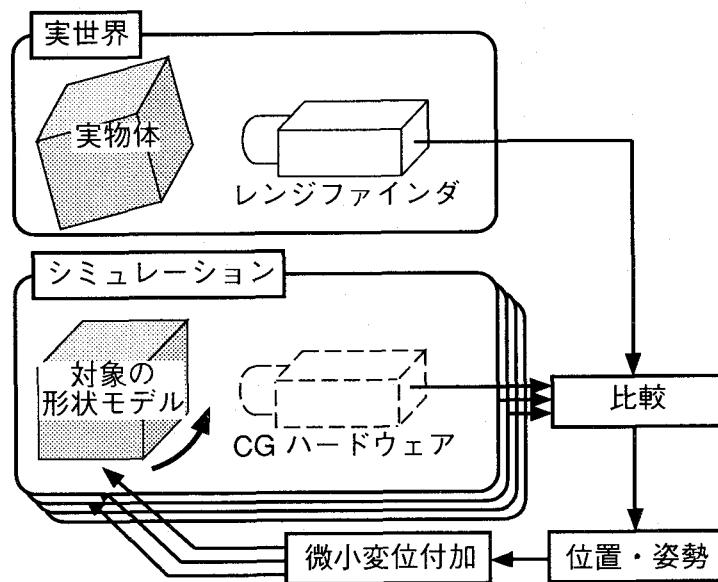


図5.6 複数距離画像の生成による動物体追跡の原理

る。ここでは、三角パッチモデルにより表現された対象物体の形状モデルから距離画像を生成する処理に、このCGハードウェアを流用して高速化する手法を提案する。

### 5.2.1 陰面消去とZバッファ法<sup>1)</sup>

CGを生成する手法は様々提案されているが、そのうちレイトレーシング法やラジオシティ法などは、photo-realisticなCGを時間をかけて計算するための手法であり、主として複雑な表面反射モデルや光源の配置などによって決定される輝度値を正確に算出するために用いられる。つまりこれらの手法は、画像の品質が計算時間よりも優先される場合にのみ用いられ、リアルタイムCG生成には向かない。それらの背景に加え、ハードウェア化が困難であることが相まって、ハードウェアによって高速実行される内容にも制限がある。CGハードウェアによって高速に実行されるCG生成手法は、主としてポリゴンシェーディング法と呼ばれるものである。この手法では、全ての三次元モデルはその表面形状を三角形などの多角形モデル（ポリゴン）を組み合わせたものとして扱われる。それぞれの物体はまず個々のポリゴンに分解され、各ポリゴンごとにその頂点の三次元座標を透視変換によって二次元画像上の座標に変換する。ポリゴンの内部は二次元画像上で塗りつぶされ、面として表示される。テクスチャマッピングなどの例外を除き、面上の輝度は各画素ごとに計算されることなく、代わりに各頂点ごとに計算された輝度を補間処理することによりなめらかな輝度変化を得る。

このポリゴンシェーディング法を正常に実行するためには、ポリゴン同士の前後関係の判定が非常に重要である。なぜなら、より視点より遠くにあるべきポリゴンを二次元画像上で上書きすると、表示されたCGは矛盾を含んだものとなるからである。このように、ポリゴンの前後関係を判定・解決する処理を陰面消去と言う。従来はこの処理を行うために、Zソート法やBSP (Binary Space Partitioning) 法など、ポリゴン同士の前後関係をあらかじめ算出しておき、より奥に位置するポリゴンから順に描画することで解決する方法が主流であった。しかし、現在ではハードウェアの進化・低価格化によって、より高価だが柔軟で高速な手法として、Zバッファ法が普及してきた<sup>2)</sup>。この手法では、あらかじめポリゴン同士の前後関係を判定することは行わない。ポリゴン頂点座標の透視変換までは通常通りを行い、その内部を二次元画像上で塗りつぶす際に、各画素ごとに以前描画したポリゴンよりも手前にあるかどうかを逐一判定する手法である。この処理を行うためには、これまでに描画したポリゴンまでの奥行き値を精度良く保持するためのビットマップ（Zバッファ）と、新たに描画するポリゴンの奥行き値を計算しZバッファと比較するハードウェアが必要となる。図5.7のように、カップの中にスプーンを描画する例で考えてみると、スプーン全体のうち目に見える部分はカップから外に出た部分のみである。そこでまず、カップを描画する際に、Zバッファに視点からカップ表面までの距離値を画素ごとに格納しておく。背景はZバッファに格納し得る最大の値を格納しておく。次に、スプーンを描画する際には、スプーン全体について視点からスプーン表面までの距

<sup>1)</sup> Zバッファには世界座標系のZ値が格納されるわけではないため、depth bufferと呼ぶほうがよいが、こちらの用語の方が普及しているため敢えてZバッファとした。

<sup>2)</sup> 低価格化の一例として、最も新しい世代の家庭用ゲーム機である Nintendo 64 はZバッファを搭載しており、それ以前に発売された PlayStation / Saturn は搭載していない。

離を計算し、それを Z バッファの値と画素ごとに比較する。このとき、スプーンまでの距離のほうが近ければ RGB バッファに輝度値を格納し、またスプーンまでの距離値を Z バッファへ格納する。スプーンまでの距離が Z バッファ値よりも遠い場合は何も行わない。これらの処理はスプーンとカップの位置関係だけでなく、カップの外側面と内側面の位置関係など単一のオブジェクト内の位置関係を解決するためにも行う必要があるため、実際には個々のポリゴンごとに全て行う。

Z バッファ法の利点としては、全てのポリゴンに関する情報が得られなくても描画を開始することが可能であるため、描画開始から表示完了までの時間遅れ (latency) が少ないとある。それと比較して BSP 法などでは、描画するポリゴンの指定を全て完了してから前後関係を判定し、その後に実際の描画開始となるために時間遅れが大きくなりがちである。これは、主として Virtual Reality などの実時間応用分野では非常に重大な問題である。また、プログラミングの容易さと言う観点でも、いついかなる時に描画してもまったく意識することなく陰面消去が行える Z バッファ法の方が使いやすい。その反面、BSP 等のように頂点座標を用いて判定するのではなく、画素ごとの計算が必要となるため、実用的な速度を得るためにハードウェアによる高速化が必須である。

### 5.2.2 Z バッファと同次座標変換

ポリゴンは平面であるため、三次元空間においてポリゴン上の全ての点の座標値はそれぞれ線形な関係にあるのは自明であるが、それを二次元画像に投影した場合、二次元画像の画素位置とその奥行き値の関係は非線型である。これは、図 5.8 のような消失点を含む画像を考えると明らかである。このことはつまり、画素ごとの奥行き値を二次元画像上で線形補間することによって計算することは不可能であることを示している。また、この非線型な関係は焦点距離

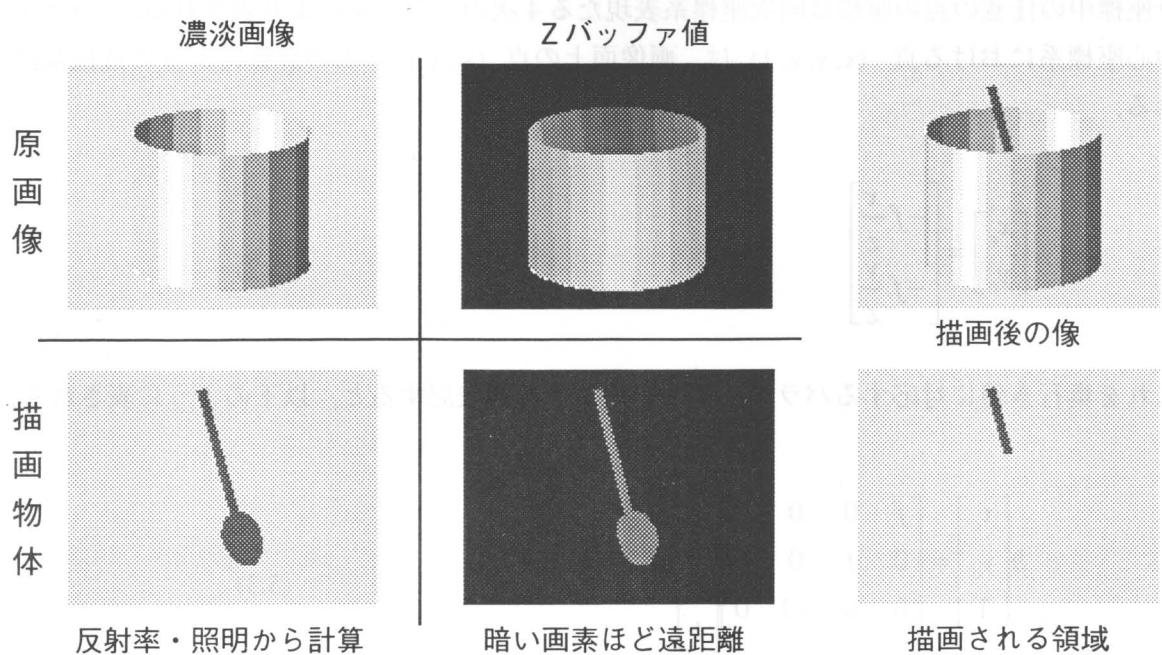


図 5.7 Z バッファによる陰面消去

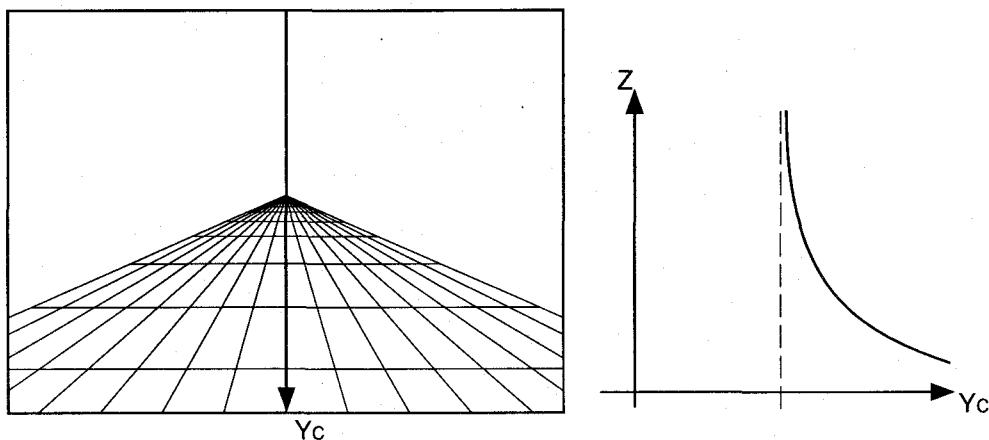


図 5.8 画像面上での奥行き値の非線型性

など透視変換のパラメータによって変化するため、透視変換による効果を無視して計算することはできない。そこで、透視変換後も奥行きに関する情報が失われないように通常の透視変換に若干の改良を施した変換を用いる。この変換は前述した非線型な奥行き値変化を線型化する働きを有しており、二次元画像上で線形補間をも行うことが可能となる。以下にその原理を詳述する。

図5.9のように、外側レンズ主点にカメラ中心座標系の原点をとる。画像面は本来シーンに対してレンズを挟んで反対側にあるべきであるが、それをレンズ主点に対して点対称な位置、すなわちZ軸に垂直で $z=-f$ を通過する平面として配置する。ここで、 $f$ は像面の大きさを単位とした焦点距離であり、像面の大きさはX, Yともに-1から1の範囲であるとする。Z軸の向きを通常と逆向きにとったのは、座標系を右手系に統一するためと、以下の説明に用いるCG描画ライブラリ IrisGL に統一する必要があるためである。世界座標系には適切な同次座標変換をあらかじめ施すことによって、カメラ中心座標系に移されているものとするので、カメラ座標中の任意の点の座標は同次座標系表現たる4次のベクトルにより表される。このカメラ中心座標系における点  $(x, y, z, 1)$  は、画像面上の点  $(x_c, y_c)$  へ次の計算によって透視変換される。

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} -f \frac{x}{z} \\ -f \frac{y}{z} \end{bmatrix} \quad (5.2)$$

これを奥行き値に対応するパラメータ  $h$  を用いて行列表記すると、以下のように表される。

$$\begin{bmatrix} x_c \\ y_c \\ h \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.3)$$

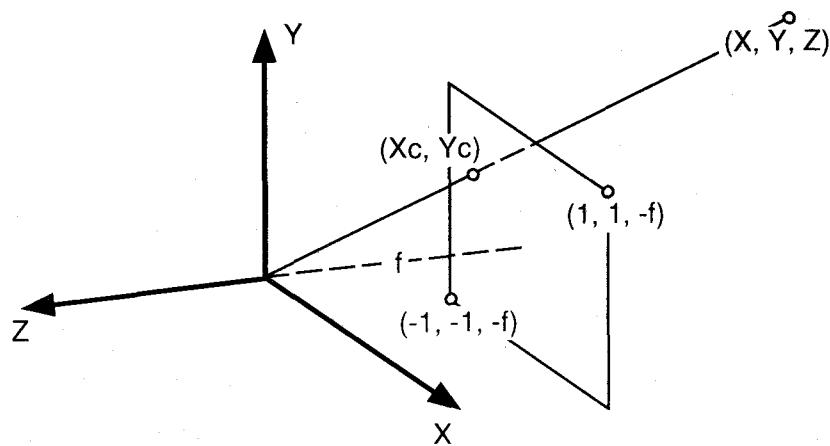


図5.9 カメラ中心座標系と像面の配置

この計算により、画像面上の座標  $(x_c, y_c)$  が決定する視線上の点は全てこの一点に投影されるため、視線内の位置に関する情報、つまり奥行きに関する情報は失われることになる。これは透視変換の性質として当然の事実であり、通常の CCD カメラ等で環境を撮影しても奥行きに関する情報が直接に得られない理由である。しかしこれでは前述したように透視変換後に陰面消去に関する処理を行うことができない。そこで、3行4列の透視変換行列の3行目に次のような行を挿入し、奥行きに関する情報  $z_c$  を保持する。

$$h \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.4)$$

この変換によれば、いくつかのカメラ中心座標の点  $(x, y, z)$  がある直線上にあるとき、それに対応する  $(x_c, y_c, z_c)$  もまた直線上にある。ただし、透視変換の効果により、 $(x, y, z)$  を直線上に等間隔にとったとしても、 $(x_c, y_c, z_c)$  は不等間隔に直線上へ並ぶことになる。これを確かめるには、 $k$  をパラメータとする直線の方程式

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = k \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} \quad (5.5)$$

をカメラ中心座標系に代入すればよい。透視変換パラメータ  $h$  と直線を表すパラメータ  $k$  を消去すれば、 $x_c$  と  $y_c$ ,  $z_c$  の関係が線形であることが確かめられる。この性質により、ポリゴン内部の各点について距離値を計算する必要がなくなる。その代わりに、ポリゴンの各頂点に対して透視変換を行い求めた  $(x_c, y_c, z_c)$  を二次元画像上で線形に補間することでポリゴン内部の全点に対して正確な  $z_c$  を求めることができる。通常、Zバッファを搭載した CG ハードウェ

アに搭載される画素塗りつぶしハードウェアには線形補間を行うための回路が含まれており、非常に高速に各点の  $z_c$  値を計算しながら塗りつぶしを行うことができる。この線形補間塗りつぶしハードウェアはスキャンライン方向にのみ動作するため、まずは三角形の辺上の点全てについて  $z_c$  を求めておき、次いで同一スキャンライン上の左方の辺と右方の辺の間をハードウェアによって塗りつぶす。同時に Z バッファと比較を行い、もしも Z バッファの値がより視点に近いならばその点についての描画は行わない。Z バッファの 1 画素のビット長は 24bit 以上のものが多いため、この線形補間には乗算を用いず、加減算と正負の判定のみにより正確に実行可能な DDA (Digital Differential Analyzer) の一種である Bresenham の方法が用いられる。同様のハードウェアは、グローシェーディングにより各頂点の輝度値を線形補間するためにも搭載されている。

ここで Z バッファに任意の実数を格納することが可能であれば、 $a$  と  $b$  は任意の実数で良い。しかし実際には Z バッファの 1 画素には有限長のビット列が格納されるため、奥行き方向の精度を確保するためにはバッファに格納可能な値の範囲全体を有効に利用することが望ましい。逆に言えば、Z バッファに格納することのできない値に対する処理を考慮する必要がある。そこで、ここでは Z バッファに格納可能な値の範囲を  $-1 < z_c < 1$  とし、ハードウェア依存となる実際の Z バッファへの格納値はこの値を線形にスケーリングすることで用いることとする。さらに、ある座標が与えられたとき、その変換値が -1 から 1 までの範囲を超えたとき、このような物体は可視範囲を越えたと考え、表示を行わないこととする。この処理は  $z_c$  のみならず  $x_c, y_c$  についても行い、画面からはみ出るようなポリゴンの描画を避けることとする。このような処理を clipping と呼ぶ。ここで、 $-1 < \{x_c, y_c, z_c\} < 1$  に収まるようなカメラ中心座標系中の範囲は、図 5.10 のような四角錐台となる。これを clipping volume と呼ぶ。

この四角錐台の底面と頂面は XY 平面と平行であり、この平面の Z 座標をそれぞれ -F, -N とすると、 $a, b$  は以下のように決定される。

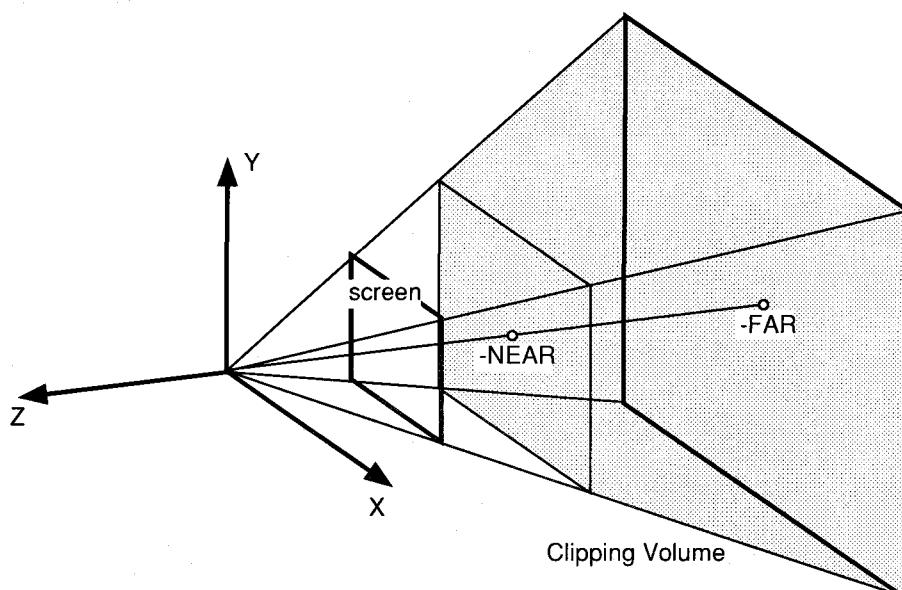


図 5.10 clipping volume

$$a = -\frac{F+N}{F-N}, b = -\frac{2 \cdot F \cdot N}{F-N} \quad (5.6)$$

clipping volume 中の座標は、この変換により  $-1 < \{x_c, y_c, z_c\} < 1$  の範囲へ変換される。

実画面上の表示領域の大きさは機種やプログラムの設定により様々に変化するため、それに応じて表示領域全体へ  $x_c, y_c$  は拡大・平行移動される（図 5.11）。また、 $z_c$  は Z バッファに格納し得る値に変換される。現在一般的なワークステーションに搭載されている Z バッファはほとんどが 24bit か 32bit である。

### 5.2.3 CG ハードウェアとレンジファインダ座標系の一致

CG ハードウェアにより高速に距離画像が生成されても、これがレンジファインダで計測された値と対応がつかなければなんの意味も持たない。そこで、この節ではシリコンレンジファインダの座標系と CG ハードウェアの座標系を一致させ、同一座標系を基準とする距離画像が得られるような設定について解説する。

4.2 で述べたように、シリコンレンジファインダの撮像系は、世界座標系に対するキャリブレーションにより 3 行 4 列のカメラパラメータ行列として表現されている。この行列は解析（付録 A を参照）によりカメラ中心座標系におけるカメラパラメータ C と、世界座標系からカメラ中心座標系への同次座標変換 T に分解することができる。C にはカメラの内部パラメータに関する情報が、また T にはカメラの外部パラメータに関する情報が格納されていると言うことができる。

ここで C は実際に得られる濃淡画像の解像度が決める範囲への透視変換を表すため、これを -1 から 1 の範囲に変換するような三次元の同次座標変換を C の前から乗じる。濃淡画像の大きさは横が  $X_{CCD}$ 、縦が  $Y_{CCD}$  とする。

$$\mathbf{C}' = \begin{bmatrix} 2/X_{CCD} & 0 & -1 \\ 0 & 2/Y_{CCD} & -1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{C} \quad (5.7)$$

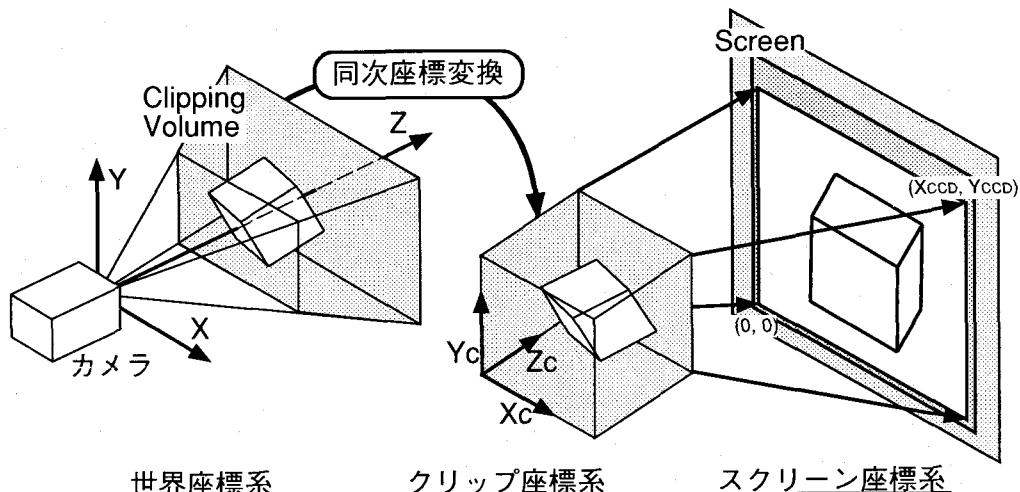


図 5.11 clipping volume とスクリーンの対応

次に、この行列  $C'$  に 3 行目を追加する。これは (5.6) の  $a, b$  の式を用いる。この時、 $F$  と  $N$  には、シリコンレンジファインダにより計測可能な範囲を包含する値とする。

$$\mathbf{C}_2 = \begin{bmatrix} c'_{11} & c'_{12} & c'_{13} & c'_{14} \\ c'_{21} & c'_{22} & c'_{23} & c'_{24} \\ 0 & 0 & a & b \\ c'_{31} & c'_{32} & c'_{33} & c'_{34} \end{bmatrix} \quad (5.8)$$

この行列を直接 CG ハードウェアへ設定する命令を用いることで CG ハードウェアの設定は完了する。例えば、シリコンレンジファインダにより計測した対象表面の座標（カメラ中心座標）をそのまま描画するだけで、CCD カメラにより撮像された画像と同一の位置に CG が現れる。また、CG ハードウェアには透視変換行列とは別に、視点を移動したりモデル内の構造を記述したりするための同次座標変換を行うためにモデルビュー行列を保持しているため、これへ  $T$  を設定すると世界中心座標において実世界と CG を一致させることができる。

#### 5.2.4 距離値の読み出し

5.2.2 で述べた方法により、世界座標系における点が透視変換の結果濃淡画像上のどの点へ変換されるのかを CG ハードウェアによって正確にシミュレートすることができる。ここでは、全ての描画の後に、Z バッファを読み出すことによって距離画像を復元する手順を解説する。

あらかじめ用意される対象物体の表面形状モデルは、世界座標系もしくはカメラ中心座標系に基づく値により定義する。また、シリコンレンジファインダにより得られる対象物体の表面形状もまたこれらの座標系に基づいている。これらはそのまま描画されるだけでなく、回転や平行移動する必要も生じるが、これらの変換もまた描画の前に CPU によって行う必要はない。なぜなら、CG 描画ライブラリは、透視変換とは別に視点の移動や対象物体の移動を実現するための同次座標変換を独立に扱うことが可能であり、これを用いることで余計な CPU 処理を避けることができるからである。例えば、計算機内モデルを基準姿勢から 10deg 回転した姿勢で描画したい場合は、まず CG 描画ライブラリに 10deg の回転を行う変換を指定しておき、その後に基準位置に関する座標値を次々と描画していくだけで良い。この、透視変換とは独立した座標変換をモデルビュー変換と呼ぶ。

全ての描画が終了すると、Z バッファには視点から描画されたオブジェクトまでの距離に関する情報が各画素ごとに格納されている。Z バッファに格納される値はモデルビュー変換に関わらず一定であるため、描画時に行ったモデルビュー変換を考慮する必要はない。ただし、Z バッファ中の値は実際の距離値に対して非線型な値であるため、カメラ中心座標系へ逆変換する必要がある。この逆変換には、CG ハードウェアに設定した透視変換行列  $C_2$  の逆行列を用いる。 $C_2$  は 4 行 4 列の正則な行列であるため、容易に逆変換を求めることができる。

$$\frac{1}{h} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{C}_2^{-1} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (5.9)$$

ただし、 $x_c, y_c, z_c$  はともに -1 から 1 までの値であるが、これらは実際にハードウェアへ書き込まれるときにスケーリングが施されていた。つまり Z バッファを読み出すと、各画素ごとに  $z_c$  に対応する値が求められるが、この値は Z バッファのビット長に依存した大きさとなっている。また、 $(x_c, y_c)$  に対応すべき読み出す画素位置も画像サイズに依存する値となっている。そこで、 $\mathbf{C}_2^{-1}$  の後ろから次のような行列をかけることにより、画素番号  $(x_2, y_2)$  とその点の Z バッファ値  $z_2$  からカメラ中心座標系における値  $(x, y, z)$  を求めることができる。ここで、 $Z_{MAX}$  と  $Z_{MIN}$  は Z バッファがとり得る最大の値と最小の値である。

$$\frac{1}{h} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{C}_2^{-1} \begin{bmatrix} 2/X_{CCD} & 0 & 0 & -1 \\ 0 & 2/Y_{CCD} & 0 & -1 \\ 0 & 0 & \frac{2}{Z_{MAX} - Z_{MIN}} & \frac{-Z_{MAX} - Z_{MIN}}{Z_{MAX} - Z_{MIN}} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ 1 \end{bmatrix} \quad (5.10)$$

これらはもちろん、もとのポリゴンモデルの頂点に一致する座標である必要はない。もしも  $(x_2, y_2)$  がポリゴンの頂点に一致しなくても、読み出された  $(x, y, z)$  はその  $(x_2, y_2)$  が決定する視線上に存在するポリゴンと視線が交差する点の座標が求められることになる。完全な陰面消去が行われているため、実世界におけるオクルージョンを正確にシミュレートしていると言うこともできる。CG 描画ハードウェアやライブラリに搭載された様々な高速化手法やユーティリティをそのまま用いることができる点も利点として挙げられる。例えば、面の表裏を判定し、視点からは裏側が観測されるポリゴン (backfacing polygon) を除去することや、個々のポリゴンに分割する代わりにメッシュデータとすることで座標変換の計算量を削減するなどの機能が IrisGL には搭載されているが、これらは全て利用可能である。

Z バッファ値の逆変換をテーブル化により高速化することも可能である。なぜなら、カメラ中心座標系における座標 Z と、Z バッファに格納されている値  $z_c$  は 1 対 1 に対応しており  $x_c, y_c$  には依存しないからである。ただし、行列で逐一計算する方法ではその点の X, Y, Z 座標が全て求められるが、テーブルを利用する方法では Z 値しか求められない。

この手法により計算された値の精度は Z バッファのビット長により左右され、それが整数値であることから、十分な精度が得られるかどうかが問題となる。しかし実際には、Z バッファで表現され得る奥行き値の範囲は自由に設定することができ、これはレンジファインダの計測可能領域を含んでいればよい。レンジファインダの分解能は、8bit から多くても 16bit 程度であり、これに較べて Z バッファのビット長は 24 もしくは 32 bit であるから、十分な精度を有すると考えられる。例えば、計測可能領域の奥行きを 1m と考えても、24 bit の Z バッファは

$1 \mu\text{m}$  以下の精度を有するので、精度的な問題は全く生じない。

### 5.3 複数の生成距離画像を用いた移動量算出

この節では、対象物体の形状モデルから生成された距離画像と、実際に計測された距離画像を比較することによって、計算機内モデルと実物体の姿勢の差を検出する方法について解説する。前述したように、ICP では計測された距離画像の各点について、計算機内モデル上の全ての点のうち最も近い点（最近傍点）を対応付け、全ての対応付けされた点同士の距離の自乗和が最小になるような姿勢を求める。もしも最近傍点が正しい対応付け（モデルと対象物体上の同一点同士の対応付け）であれば理想的な収束性を示すが、対象物体の形状によっては正しい対応付けとはかけ離れた対応点が選択されることにより、必ずしも速く収束しないことは 5.1.1 で示した。しかも、対象物体の形状モデルは三次元空間中に複雑に分布するので、最近傍点求めるためには探索する必要があり速度低下を免れない。高速な動物体追跡を実現するには、このような探索を行うことは避けるべきである。

4 章で述べた手法では、最適化を行うための距離画像同士の差の情報として、光軸方向の奥行きの差を用いた。しかしこの方法では、対象物体の光軸方向の平行移動と、光軸に垂直な軸廻りの回転について heuristic な計算法により運動が求められたが、光軸に垂直な方向の平行移動や光軸廻りの回転については検出が困難であった。また、この heuristic な計算法は、対象物体の形状や姿勢によっては正常に動作しない場合が存在する。このように、距離画像の光軸方向の差のみを用いる方法では、対象物体固有の形状情報を無視しては安定かつ高速な追跡を行うことはできない。

対象物体の形状情報を陽に用いて姿勢を求める方法（抽出された特徴点同士の対応付けなど）は、複雑で時間のかかる処理となる場合が多く、また利用する特徴量によっては対応できない形状が存在することが多いため、ここでは特徴点の抽出等を行うことも避ける。そこで、対象物体の形状情報を直接用いるのではなく、対象物体が移動したときに、距離画像上にどのような変化が生じるのかを各自由度ごとにあらかじめ調べ、実際に得られた計測距離画像と生成距離画像の差をこの変化に分解することによって実際の変位を求めることを考える。この方法では、4 章で述べた方法のように距離画像の奥行き方向の差を用いて姿勢の変化を検出するが、その変化がどのような姿勢変化から生じたものであるかを知ることができるために、あらゆる姿勢変化に対応することができる。例えば、対象物体の光軸まわりの回転については、回転に伴う距離画像の変化が形状依存のために 4 章の方法では検出が難しかったが、あらかじめ対象の各軸まわりの回転に対する距離画像の変化の様子が分かっていればそれらに分解することで検出することができる。ICP のように、距離画像同士を比較するために最近傍点を探索する必要がないため、姿勢検出時間もたいへん短い。そのかわり、距離画像の変化を各運動自由度ごとに調べる必要があるが、これは 5.2 節で述べた方法により距離画像生成の時間が飛躍的に高速化されているために問題とならない。

### 5.3.1 距離画像比較による姿勢推定

対象物体と同一の形状を表現する対象物体モデルが計算機内に既に用意されており、計算により任意の位置・姿勢について距離画像の生成が可能であるとする。この距離画像は、精密なシミュレータによって計測点の配置や計測値の単位系などが実在のレンジファインダと一致しているものとする。計測・生成により得られた距離画像は、各画素の計測値のうち、奥行き値のみを取り出し、それらを順に並べたベクトルとする。ベクトルの長さは最大で全画素数となる。求めるべき変数は計算機内モデルの位置・姿勢に関するパラメータで、原理的にはいくつでも良いが、こでは剛体運動パラメータの6自由度をそれぞれ $t_x, t_y, t_z, r_x, r_y, r_z$ とする。これらのパラメータから生成された距離画像が、計測された距離画像と一致したときに、これらのパラメータを求めるべき対象物体の姿勢とする。これらのパラメータから距離画像ベクトルを生成する操作を $f$ とする。

計測された距離画像ベクトルを $m$ とする。また、初期位置に関して生成された距離画像 $s$ を

$$s = f(t_x, t_y, t_z, r_x, r_y, r_z) \quad (5.11)$$

とする。これらのベクトル $m$ と $s$ は、距離画像ベクトルと等しい次元数を有する距離画像空間上に表すと、図5.12のようになる。この空間内において、 $s$ は明らかにパラメータ空間 $f(t_x, t_y, t_z, r_x, r_y, r_z)$ が表し得る部分空間上にある。計測された距離画像ベクトル $m$ もまた最終的に求められるべき姿勢を $f$ に与えて得られる距離画像と同一であるためにこの部分空間上に存在する。 $f$ の変化は各自由度に対して非線型な変化を示すため、姿勢推定問題はこの部分空間内において、各6自由度を独立変数とし、 $s$ と $m$ の間の距離を従属変数とする非線型最適化問題であると言うことができる。

このような最適化問題を解くためには最急降下法や滑降シンプレックス法など様々な方法が考えられる。しかしこれらは従属変数の値の増減のみを利用して最適化を行うために収束速度が速いとは言えない。幸いなことに、この部分空間は区別的に連続かつ滑らかであるため、微小領域に関しては線形であると考え、線形最適化により姿勢を検出することが可能である。そこで、各自由度のパラメータに対して微小な変位量をそれぞれ与えて得られる距離画像を用いて、この部分空間の基となるベクトルを生成する。つまりこれらのベクトルは互いに独立であり、これらを線形結合した空間は $f$ がとり得る部分空間全体を張る。例えば、 $t_x$ に対する微小変位を $\Delta t_x$ とすると、その微小変位により生じる距離画像ベクトルの変化量 $v_1$ は

$$v_1 = f(t_x + \Delta t_x, t_y, t_z, r_x, r_y, r_z) - s \quad (5.12)$$

となる。同様の操作を $t_y, t_z, r_x, r_y, r_z$ にも施し、それぞれ $v_2$ から $v_6$ とすると、この部分空間は各自由度に関する変位係数 $a_1 \dots a_6$ をパラメータとする線形6次元空間であると考えられ、次の式によって表される。

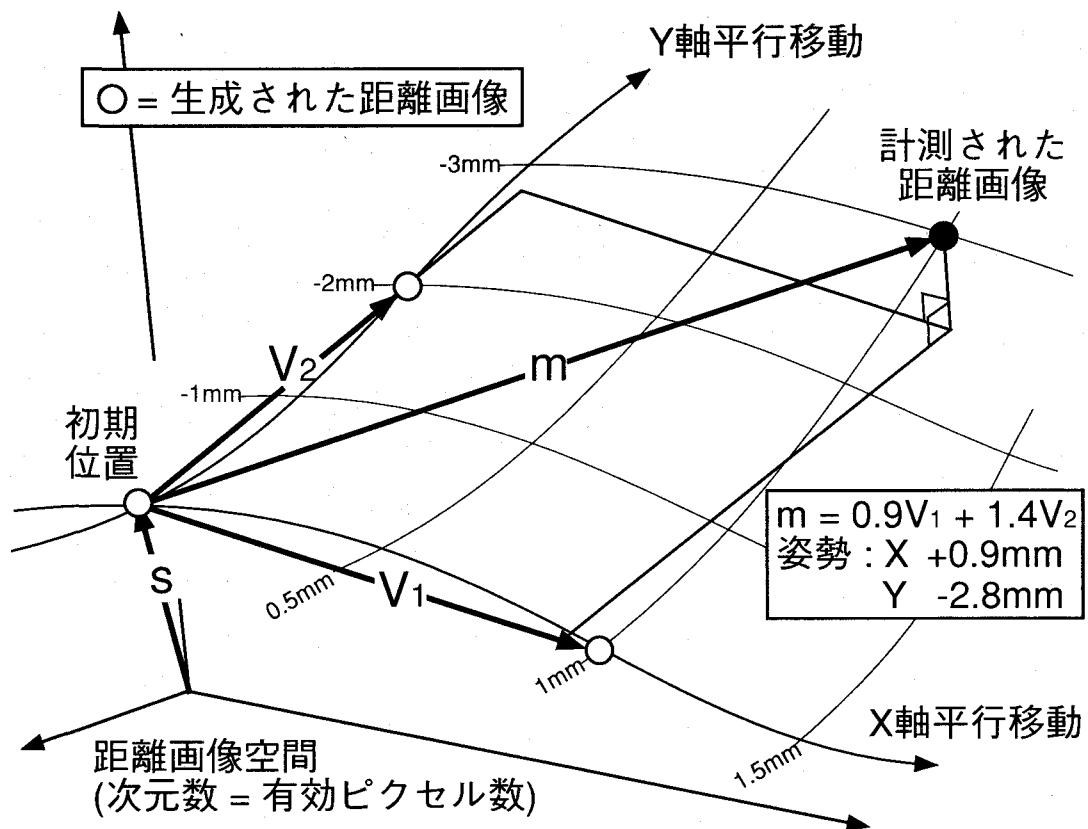


図 5.12 距離画像空間における線形近似

$$\begin{aligned}
 & f(t_x + \Delta t_x a_1, t_y + \Delta t_y a_2, t_z + \Delta t_z a_3, r_x + \Delta r_x a_4, r_y + \Delta r_y a_5, r_z + \Delta r_z a_6) \\
 &= s + \sum_{i=1}^6 a_i v_i
 \end{aligned} \tag{5.13}$$

$m$  もまた  $f$  が張る空間上に存在するので、これらの線形結合パラメータ  $a_1 \dots a_6$  を求めることによって姿勢が求められる。

$$\min_{\alpha} |m - \mathbf{V} \cdot \alpha|^2 \tag{5.14}$$

ここで、 $\alpha$  と  $\mathbf{V}$  は以下のように定義される。

$$\alpha = [a_1 \ a_2 \ \dots \ a_6]^T, \quad \mathbf{V} = [v_1 \ v_2 \ \dots \ v_6] \tag{5.15}$$

実際には非線型性によって  $m$  は  $v_1 \dots v_6$  の張る空間上に乗ることは限らないため、次の最小自乗法によって結合係数を求める。

$$\alpha = (\mathbf{V}^T \cdot \mathbf{V})^{-1} \cdot \mathbf{V}^T \cdot m \tag{5.16}$$

この計算によって得られた結合係数  $a_1 \dots a_6$  と、あらかじめ与えた微小変位量との積をとることによって、各自由度ごとに実際の変位が求められる。この計算では線形性を仮定しており、この仮定の下では繰り返し計算なしに一度で変位量が求められる。しかし、実際の変位と距離画像の変化の関係は非線型であるため、この計算を繰り返し行うことによって真の変位量を求める。

### 5.3.2 微小変位量の決定と背景画素の処理

前節で述べた方法では、仮想的な変位に対する距離画像の変化を調べるために微小変位を与えたが、この変位量の決定は追跡性能に大きな影響を与える。微小変位量が大きすぎると線形性の仮定から外れた領域を調べることになり、正常な近似計算を行うことができなくなる。また、微小変位量が小さすぎると、生成された距離画像同士の変化が小さくなるために距離画像生成時の誤差が問題化してくる。実際には、距離画像生成の精度は非常に高いために、微小変位量を小さく設定したときに発生する誤差はわずかである。しかし、実物体の姿勢が大きく変化した場合は、もはや初期姿勢周辺からは線形的に表すことができないような距離画像に変化しているような場合も生じる。理想的には、対象物体の運動が大きな場合には、実際の変位量と微小変位量が一致し、線形結合パラメータが 1 となる場合に精度が高い。そこで、前回の最適化で得られた変位量  $a_1 v_1 \dots a_6 v_6$  を次回の微小変位量とすることによって、連続的に運動する物体への追従性を高める。この方法では微小変位量が異常に大小することがあるために、閾値を設けて微小変位量の絶対値がある範囲から逸脱した場合は標準的な大きさに設定し直す。

前節では、微小変位によって距離画像の全ての画素が連続的に変化すると仮定したが、実際には対象物体の回転によって画素をジャンプエッジが横切った場合などに不連続が生じる。また、背景に対応する画素や、物体の反射率が低いような場所では距離値が得られないことがある。これらの現象は計測距離画像  $m$ 、生成距離画像  $s$ 、 $v_1 \dots v_6$  についてそれぞれ生じるため、全ての画素位置に対して、これら 8 つの距離画像を全て調べ、どれか 1 つでも未計測であったり異常に大きな変化が現れたりしている場合にはこの画素を無視する。結果的に、全ての距離画像について有効かつ連続的な変化が見られるような位置の画素を選択し、これらを順に並べることによって距離画像ベクトルを生成する。対象物体の姿勢や環境、微小変位量によって選択される画素数は変化するため、距離画像ベクトルの大きさは計算ごとに変化する。

これまで、各自由度に関する変位量は独立に扱ってきたが、実際の描画などの処理ではそれらが結合された上で利用される。回転に関する 3 つのパラメータは、それぞれをどの順序で演算するかによって結果は異なるが、変位量が微小である場合にはこの問題は無視できることが知られている。そこで、これらの演算は単に順番に施され、実際の変位が比較的大きい場合には繰り返し演算することによって誤差が吸収されるものとする。つまり、各軸まわりの回転量は最小自乗法によって求められるごとに回転行列へ反映・統一され、次の最適化計算ではこの回転行列によって表される姿勢を初期位置として計算を行う。

## 5.4 モデリング

対象物体の絶対位置・姿勢を求めるためには、あらかじめ計算機内に對象の形状に関する情報を構築しておく必要がある。つまり逆説的には、形状が既知でない物体については絶対的な姿勢は定義できない。なぜなら、絶対的な位置や姿勢はその形状との関係から決定するからである。例えば、対象物体の形状をモデリングした時の座標系を絶対位置として定義し、モデリング時からの相対的な姿勢の相違が絶対的な姿勢であるとするような場合である。しかし実際には、計測開始時点からの相対的な姿勢の変化を検出するような応用分野も多い。このような場合には、計測開始段階では対象の形状が未知であっても、最初に得られた距離画像との相違から相対的な姿勢情報が得られると考えることができる。このような計測を実現するためには、計測初期に得られた距離画像から、対象物体の形状をモデル化する機構が必要となる。また、距離画像には対象物体表面の一部に関する情報しか含まれていないため、対象物体の回転に応じて得られた距離画像とモデル間の共通の領域が減少した場合、正常に追跡を実行することが不可能となる。この問題を解決するために、対象物体の回転や移動に応じて新たに出現する部分をモデルへ追加することが必要である。

### 5.4.1 距離画像によるモデル構築

運動追跡に先立ち、対象物体の形状モデルを構築するために、レンジファインダにより最初に計測された距離画像が用いられる。この距離画像は単なる奥行き値の配列ではなく、対象物体表面上に適当な間隔で分布した計測点についての三次元座標が得られているものとする。対象物体のモデルは最終的にCG ハードウェアにより描画することができるモデルならどのような形式のものでも良いが、シリコンレンジファインダの計測点の密度などから考え、単に三角パッチモデルとした。さらに計測点数が多いレンジファインダを利用した場合では、描画速度の向上などの観点から、平面の当てはめやセグメンテーションを行うことも考えられる。また、次節で述べる複数距離画像の統合のために適したモデルとすることも考えられる。

距離画像の計測点は画像上では等間隔に配列しており、隣同士の位置関係は明確であるため、隣の計測点同士を用いて三角パッチを張る。ここで、極端に奥行き値が異なるような場合はジャンプエッジであると見做し、パッチ生成は行わない。得られた三角パッチモデルの原点は、距離画像の原点となるが、原点位置は三角パッチモデルの重心付近であることが望ましい。なぜなら、運動追跡の際に与える微小移動量のうち、回転に関しては原点まわりの回転として与えられるが、原点が三角パッチモデルから離れた場合には、図 5.13 のように回転運動によるモデルの移動量が大きくなり、平行移動による距離画像の変化と近似的に同一の距離画像ベクトルが生成されることになるからである。そこで、三角パッチモデルを構成する全ての頂点の重心を求め、その点を原点としたモデルに変換する。また、モデル初期位置の平行移動成分をカメラ中心座標系における重心の位置とする。

### 5.4.2 複数距離画像の統合

最初の距離画像を用いて対象物体の形状をモデリングした場合、対象物体表面のうちレンジ

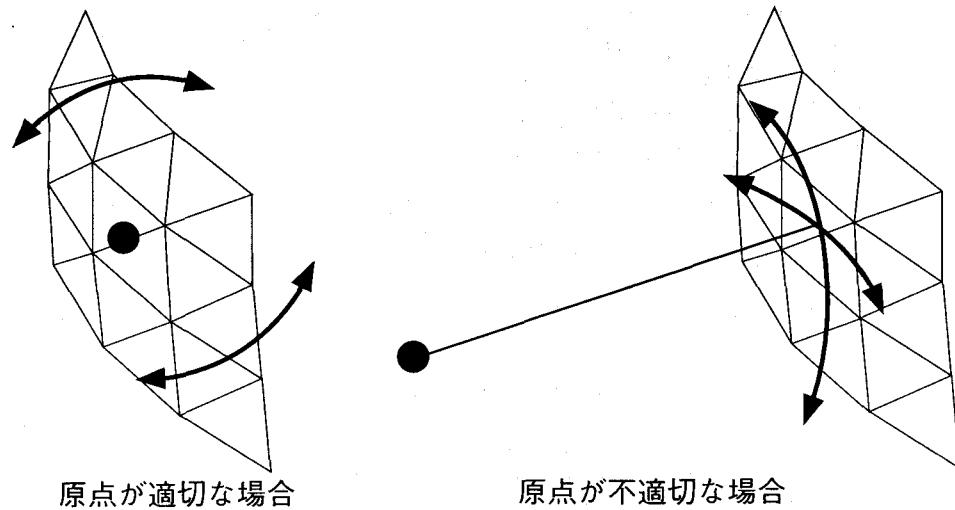


図 5.13 モデル原点の位置と微小回転の効果

ファインダにより計測が行われた領域以外については情報が得られない。追跡すべき対象物体の回転角が十分小さい範囲に収まることがあらかじめ分かっている場合を除き、対象物体が回転するにつれて得られる距離画像と計算機内モデルとの共通領域は減少し、追跡の安定度や精度が低下する。そこで、対象物体の回転につれて現れる、モデルに含まれていない領域を逐次的にモデルへ追加することにより、対象物体の回転角に関する制限を取り除く。また、対象物体が1回転するなど、対象物体上の全ての点について一度でも計測が行われれば、対象物体の全周形状が追跡の結果副次的に得られることになる。

モデル追加の方法として、対象物体が一定角度回転する度に得られた距離画像全体を三角パッチモデルに変換し、既存のモデルと単に併用する方法がまず考えられる。しかしこの方法では、距離画像と形状モデルが重なり合った部分は二重もしくはそれ以上に重なったパッチモデルにより表現され、ポリゴン数がたいへん多くなる。正確に姿勢を抽出するためには距離画像と形状モデルの重なりの割合を一定以上確保しておく必要があるため、モデル追加を判定する角度が大きすぎると正確な追跡処理が行えなくなり、また角度が小さすぎるとポリゴンの重なりが多くなる。

この問題を避けるためには、円筒座標系モデルなど单一の形状モデルを逐次的に更新していく方法が考えられる。しかし、入り組んだような形状を含むあらゆる全周形状を表現し得る適当なモデルが存在しないことや、計算量が高く追跡速度の低下を来す可能性があるために、この方法は用いない。そこで、単に三角パッチモデルの集合体として形状を表すが、互いに重なり合った部分を削除してポリゴン数の不当な増加を抑える方法が考えられる。この方法を実現するためには、計測された距離画像のうち、どの領域が既にモデル化されており、どの部分が新規に現れた領域なのかを判定する必要がある。代表的な方法としては、三次元空間中に分布する計測点の座標が、形状モデルから一定の距離以上離れている場合に新たな計測点とする方法などがある。しかしこの方法では、ICPと同様に点からモデルまでの距離を計算する必要があり、十分な速度を得ることができない。そこで図5.14のように、二次元画像面上で、レン

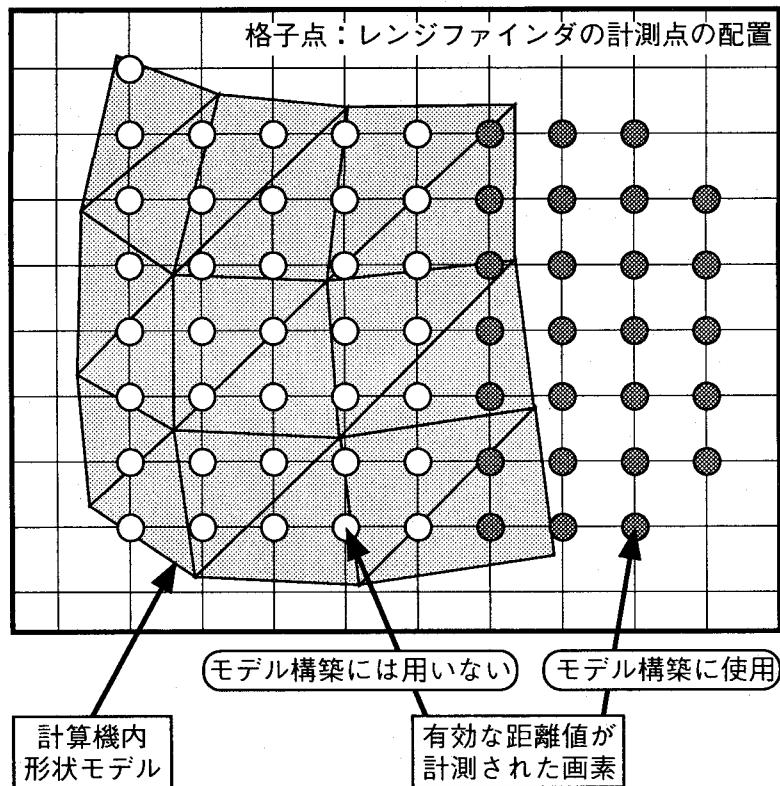


図 5.14 新規出現部分に相当する画素の判定

ジファインダにより計測された領域と、形状モデルの占める領域の重なりを判定し、重なっていない部分のみをモデルへ追加することを考える。追跡処理の段階で、すでに形状モデルは距離画像へと変換されているので、モデルが画像上を占める領域は既に分かっている。この情報を用いて、距離画像の計測点のうちモデルと重なっていない点の数が一定の閾値を超えた場合に、これらの点を三角パッチモデルに変換し、以前の三角パッチモデルへと追加する。より正確には、継ぎ目の部分に隙間がないように、計測点のうちモデルと重なっていない点と、その隣接点を全て取り出し、それらの点を用いて三角パッチモデルを生成する。一定回転角ごとに形状モデルを追加する方法では、対象物体の凸部などで十分な数の計測が行われない場合などが生じるが、形状モデルを追加するタイミングを決定するために重なっていない計測点の数を利用することで、対象の形状の複雑さや凹凸に影響されず安定に追跡とモデル追加を行うことができる。

対象物体の回転に応じてモデルが次々と追加され、回転角が  $180\text{deg}$  に達すると、対象の裏側までモデリングが進行し、最初にモデル化された領域の裏側が未計測領域へと侵入してくる（図 5.15）。この現象が起こると、未計測領域の判定が正しく行われず、モデルの追加が停止してしまうので、面の表裏を判定する必要が生じる。幸い、CG ハードウェアは描画の高速化や輝度の計算のために表裏を判定する機能を有しているので、この機能を利用して単にカメラから裏側が観測されるポリゴンの描画を停止すればよい。これは追跡に用いる距離画像生成の高速化にも有効である。

5.4.1 ではモデルの原点として、モデルを表す全点の重心座標を用いたが、このままでは対

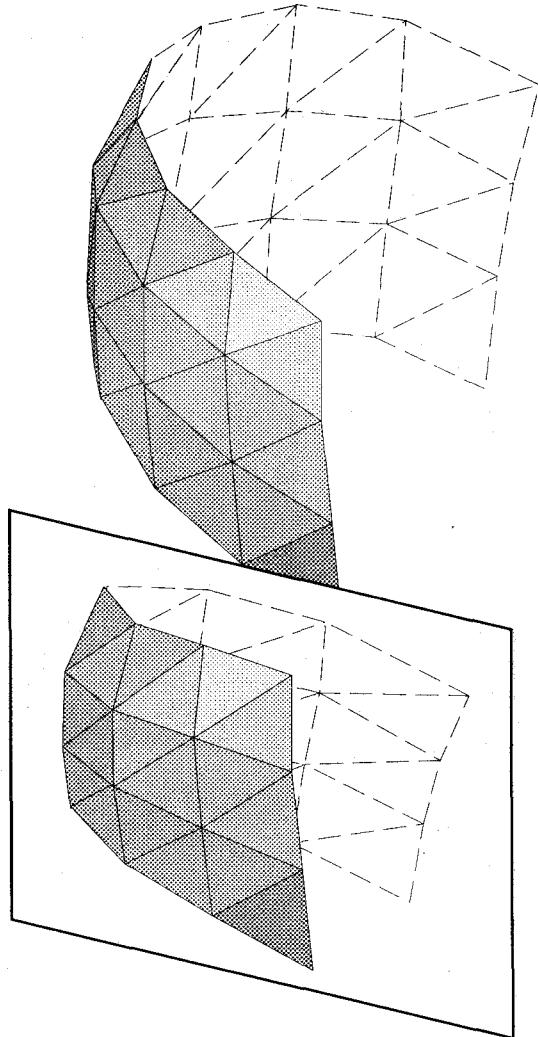


図 5.15 裏面描画の停止

る。むしろ他の shape registration / fusion 研究が前提とするデータ群を与えるシステムと/or ことができるだろう。

## 5.5 実験

この実験を通して用いる対象物体として、図 5.16 に示すような日本人形を用いた。人形の全高は約 180mm で、頭部の直径は約 60mm である。頭部は主に目や鼻、耳などの凹凸が、また胴部は和服の袖や衿、帯などにより、対象全周にわたり形状のうねりがあるが、特徴としては弱いため特徴抽出による姿勢の推定は比較的困難であると考えられる。ここではこの対象物体に対して、回転と平行移動に関する精度評価と、その結果得られたモデルの精度を評価する。形状のモデリングには、5.4 で述べた同時モデリングを用い、あらかじめ形状モデルを与えることは行わなかった。

象物体が回転したときに偏芯することになり、裏側の追跡性能が落ちる。この問題を避けるためには、あらかじめおおよそのモデルの大きさを指定しておき、最初に設定する原点をずらすか、モデル更新の度に原点位置を再設定する方法がある。

これまで述べた一連の処理を繰り返すことにより、単に対象物体の距離画像を連続して計測するだけで、対象物体の連続的な位置・姿勢の変化と全周形状を得ることができる。つまり、この処理は動物体の実時間追跡であり、かつ実時間形状モデリングもある。生成された形状モデルの形式はエレガントとは言えないし、物体の運動が速く追跡の精度が十分でない場合には三角パッチモデル同士に継ぎ目のようなものが看取されることもある。これを解決するためには、様々な研究者によって提案されている時間はかかるがさらに精度が高い手法を用いて位置合わせを行い、また単一の形状モデルへ統合をすることができる。この点において、本研究が提案する手法はリアルタイム性を追求したものであり、これまでの研究とは問題定義や結果は似ているが互いに干渉しない性質のものである。

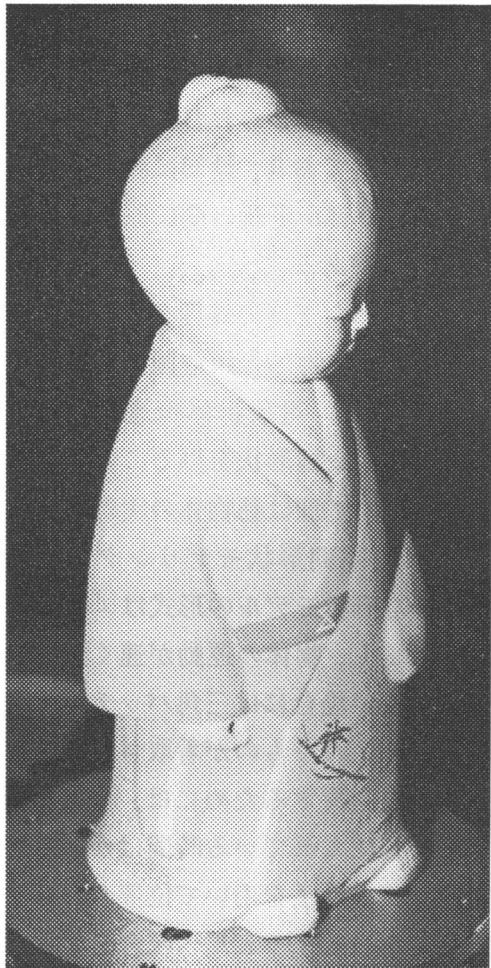


図 5.16 対象物体

### 5.5.1 実験装置

実験装置には 2 台のワークステーションを用い、6 章で述べる方法により並列的な処理を行った。図 5.17 に示すように、シリコンレンジファインダは SGI 社製ワークステーション Iris Indy に接続された VME 拡張ボックスを経由し、パラレル IO により駆動・読み出しが行われる。また、キャリブレーションに必要な濃淡画像の読み出しも同様に VME 拡張ボックスに搭載した画像入力・処理ボードにより行われる。シリコンレンジファインダから読み出されたカウント値は Indy 内で世界座標値に変換され、Ethernet を経由してもう一台の SGI 社製ワークステーション Iris Indigo<sup>2</sup> Extreme に転送される。このワークステーション中では、CG ハードウェアによる距離画像生成・距離画像の比較による姿勢検出など大部分の処理を行っている。このワークステーションに搭載された CG ハードウェアの Z バッファはその深さが 24bit である。また、描画速度はグローブシェーディング、ライティング、Z バッファ有効とした時の三角形メッシュの描画速度が 415,000 個／秒である。

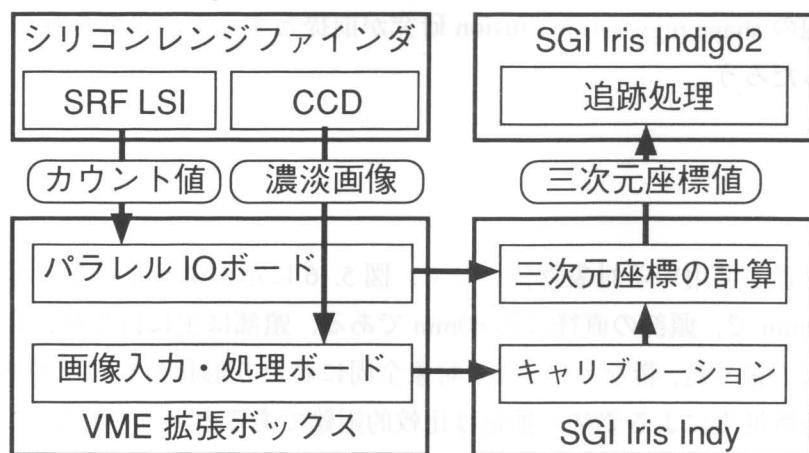


図 5.17 機器構成

### 5.5.2 平行移動の追跡精度評価

対象物体をスライドステージ上に設置し、対象物体を平行移動させたときの位置検出精度を評価する。図5.18に示すように、スライドステージはカメラ光軸であるZ軸にはほぼ並行に設置した。また、角軸まわりの回転角をそれぞれ $R_x, R_y, R_z$ とする。追跡処理そのものには全く変更を加えずに、6自由度全てを常に追跡する動作としているため、一定に変化すべき値 $T_z$ 以外は一定値となるはずである。対象物体は回転しないため、形状モデルとしてはスライドステージのほぼ中間位置で取り込んだものを用い、その後にスライドステージを一度端まで移動してから精度評価を行った。追跡の結果得られた各パラメータの変化と、スライドステージ位置の関係を図5.19に示す。 $T_z$ 値はスライドステージの位置とよく一致している。ここで、スライドステージが0の点は、カメラからは約200mm離れた点に対象物体がある状態であるため、画像上における対象物体の像は最初と最後で約2倍に変化している。また、 $T_x$ と $T_y$ は常に0となるはずであるが、スライドステージの移動軸とシリコンレンジファインダの光軸方向がややすれていたために、 $T_y$ 値は線形にずれている。この時、スライドステージの移動量と、 $T_z$ 値の誤差の標準偏差は0.17mmであった。これはシリコンレンジファインダの奥行き精度(1mm以下)に対して向上しているが、これは多数の計測点から総合的に位置を算出することにより、誤差の平滑化効果が表れたことによると考えられる。また、一定であるべき各軸まわりの回転角については、ノイズ的なばらつきがどの軸回りにもみられるものの、全体としては2deg以内に収まっており、平均的には0.5deg程度の誤差であった。

### 5.5.3 回転運動の追跡精度評価

次に、同一の対象物体に対して、回転角の検出精度を評価する。図5.20のように、ターンテーブルの軸はシリコンレンジファインダの光軸と垂直であり、Y軸とほぼ並行に設置した。対象物体のモデルを先に与えることはせず、回転の追跡とともに新たに出現する部分をモデリングし、追跡に用いた。ターンテーブル回転角と、各軸まわりの回転角の関係を図5.21に示す。

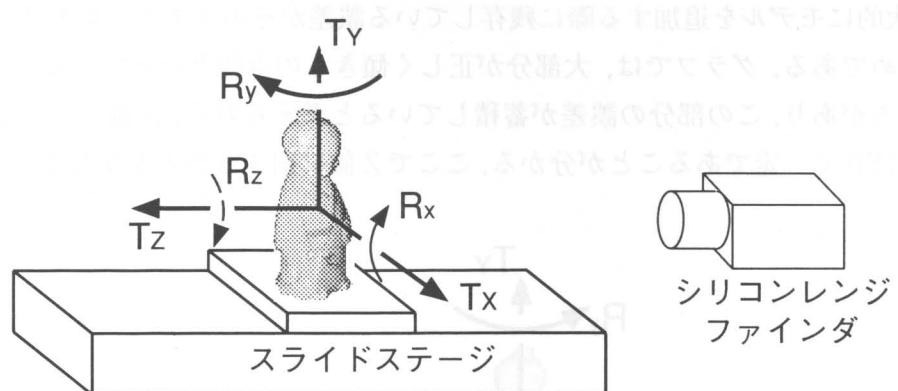


図5.18 平行移動精度評価のためのスライドステージの配置

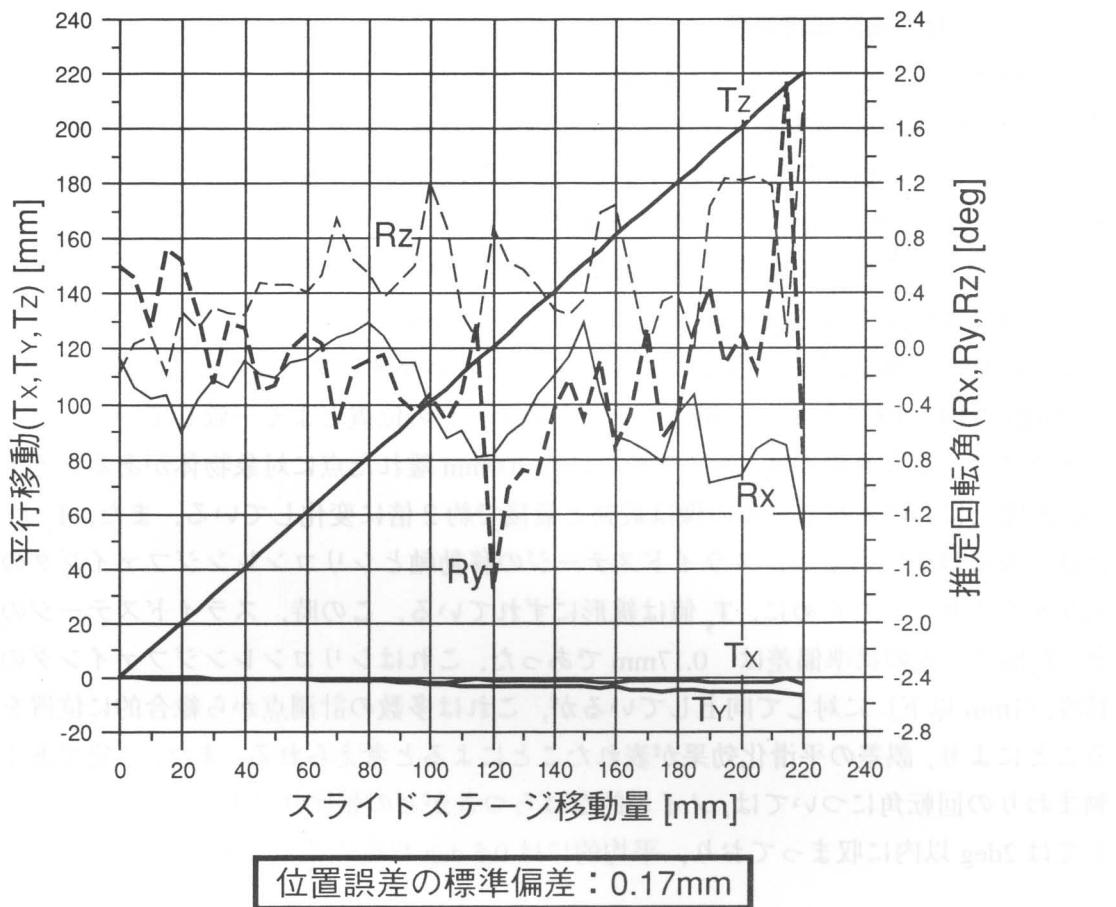


図 5.19 平行移動時の 6 軸運動推定結果

す。Y 軸まわりの回転角はターンテーブルの回転に応じて直線的に検出されているが、最終的にターンテーブルが 240deg 回転したところで検出角は 230deg と、若干ドリフトしている。これは、逐次的にモデルを追加する際に残存している誤差がそのままモデル形状に反映され蓄積されたためである。グラフでは、大部分が正しく傾き 1 の直線となっているが、所々傾きが小さいところがあり、この部分の誤差が蓄積していると考えられる。X 軸および Y 軸に関する回転角はほぼ 0 で一定であることが分かる。ここで Z 値に関する誤差が生じると、対象物体が 1

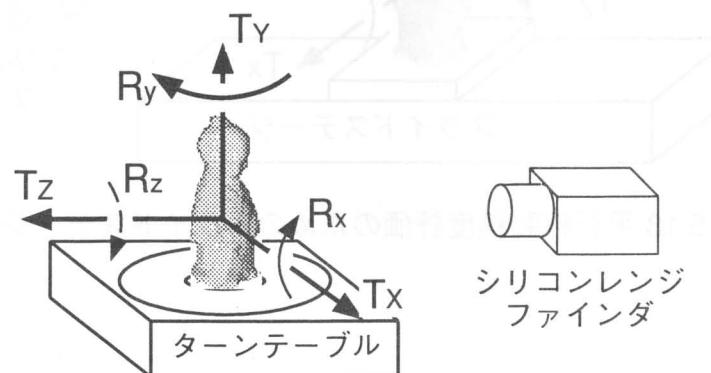


図 5.20 回転運動追跡精度評価のためのターンテーブルの配置

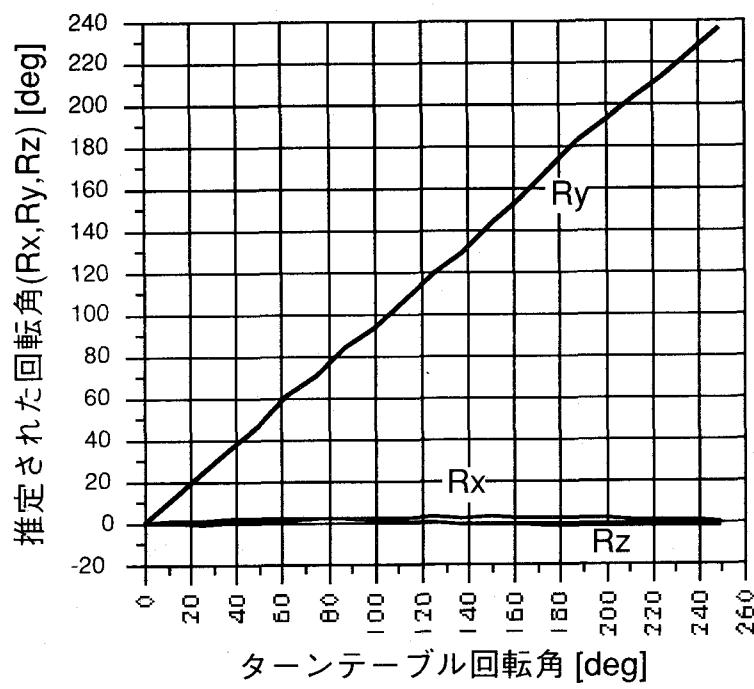


図 5.21 回転運動時の回転成分推定結果

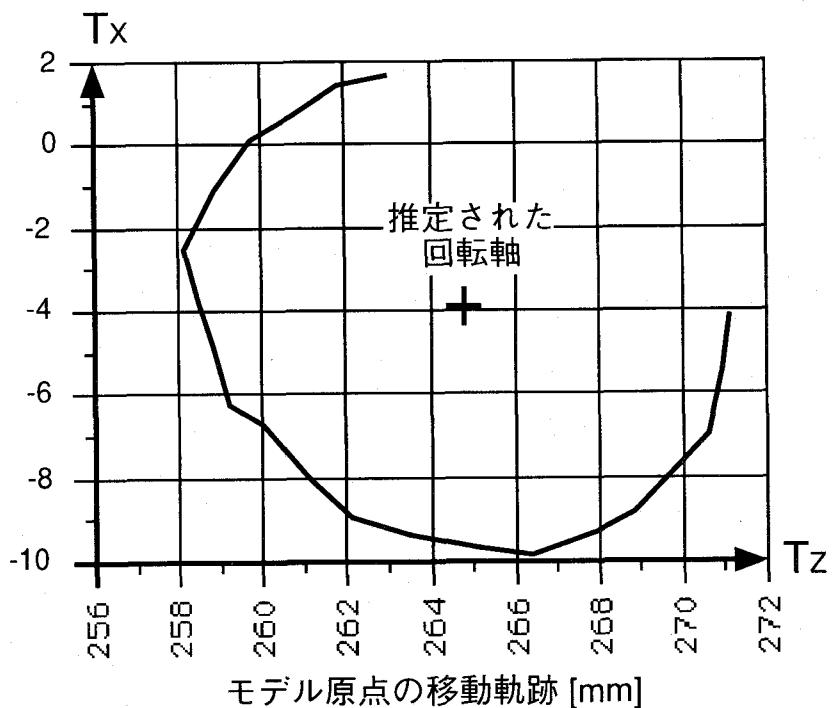


図 5.22 回転運動時のモデル原点の軌跡

回転した時点で生成されたモデルにねじれが見られるが、次節で示すようにこのような現象は見られなかった。Y軸まわりの回転角と、ターンテーブル回転角の差の標準偏差は3.24degであった。

ターンテーブルの回転軸の位置に関する情報は全く利用しておらず、また対象物体の運動が回転であるという情報も用いられていない。そのため、最初の距離画像からモデルを構築する際に自動的に設定されたモデル原点はターンテーブルの回転軸上にはない。また、対象物体は回転運動を与えられるため、このモデル原点はY軸まわりを円運動するはずである。そこで、モデル原点のX Z座標上の軌跡を描いたグラフが図5.22である。モデル頭部の直径は約60mmであるが、モデル原点の軌跡の直径は約12mmであり、円の軌跡からの誤差は1mm以内である。対象物体が回転しているにも関わらず、またそのモデルが逐次的に構築されているにも関わらず、たいへん正確に対象物体の位置が求められていることが分かる。また、モデル原点のY座標はほぼ一定であり、誤差は1mm以下であった。

#### 5.5.4 モデリング結果

ここでは、5.5.3の実験時に得られた全周形状についての評価を行う。図5.23において、(a), (b)は対象物体が初期位置にあるときの画像である。(a)はシリコンレンジファインダに内蔵されたCCDカメラによる濃淡画像であり、その時のワークステーション上の表示が(b)である。(b)において、白色のワイアフレームモデルは計測された距離画像そのものを表し、またフラットシェーディングされたポリゴンモデルはその距離画像から構築された対象物体の表面形状モデルである。また、左上から右下に向かう白い直線は、カメラ光軸を表す。距離画像生成時には、(a)と同一の視点からの画像がCGハードウェアによって描かれるが、ここでは対象物体の回転の様子を見やすくするために、対象の右側下方に視点を設定して表示している。形状モデルは計測された距離画像から直接生成されるため、対象物体の移動が行われない間は(b)のように距離画像と完全に一致している。(c)では対象物体が左へ回転し、それによって生じた距離画像の変化から対象物体モデルが実物体と同じ方向へ回転した様子である。距離画像の変化はわずかであるため、ワイアフレームの形状は大きく変化しているように見えないが、正しく運動が検出されていることが分かる。(c)では新規出現画素の数が閾値以下であり、まだ新規出現部分はモデル化されていないが、(d)のようにさらに回転すると新規出現画素の数が閾値を超え、右側に観察される出現部の距離画像が三角パッチモデルに変換され、形状モデルに追加される。この処理の繰り返しによりモデリングが行われる。

対象物体が1回転し、最終的に得られた全周三次元形状モデルを図5.24に示す。対象物体

表5.23 構築されたモデルの精度 [mm]

誤差の種類	初期姿勢	後退100mm	回転30deg.
最大値	0.741	5.158	5.002
最小値	0.0005	0.0015	0.0026
絶対値平均	0.200	0.647	0.585
標準偏差	0.216	0.901	0.899

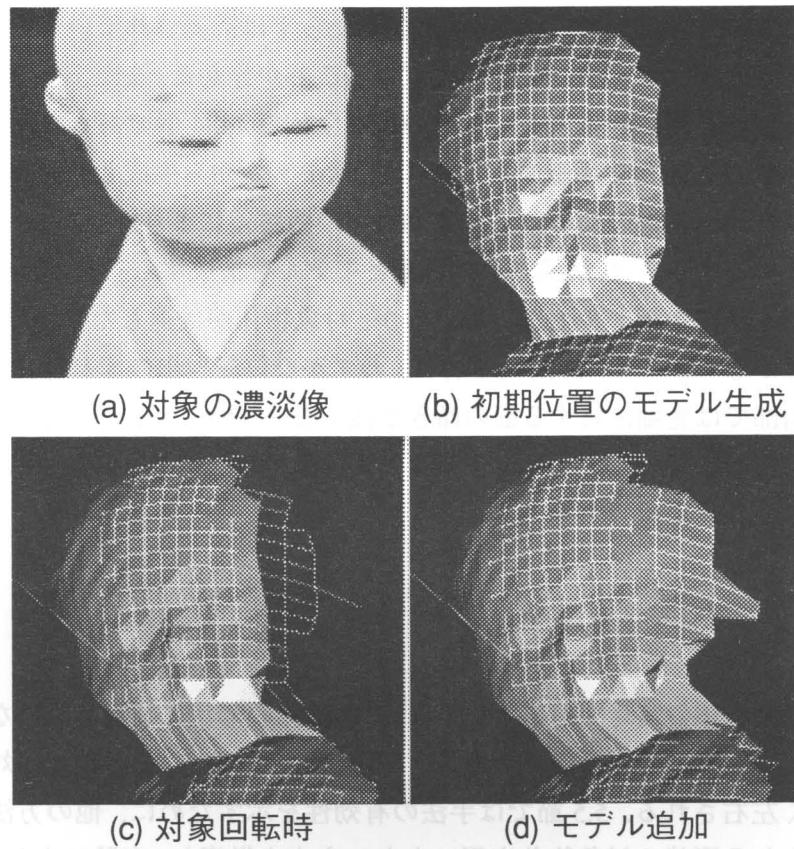
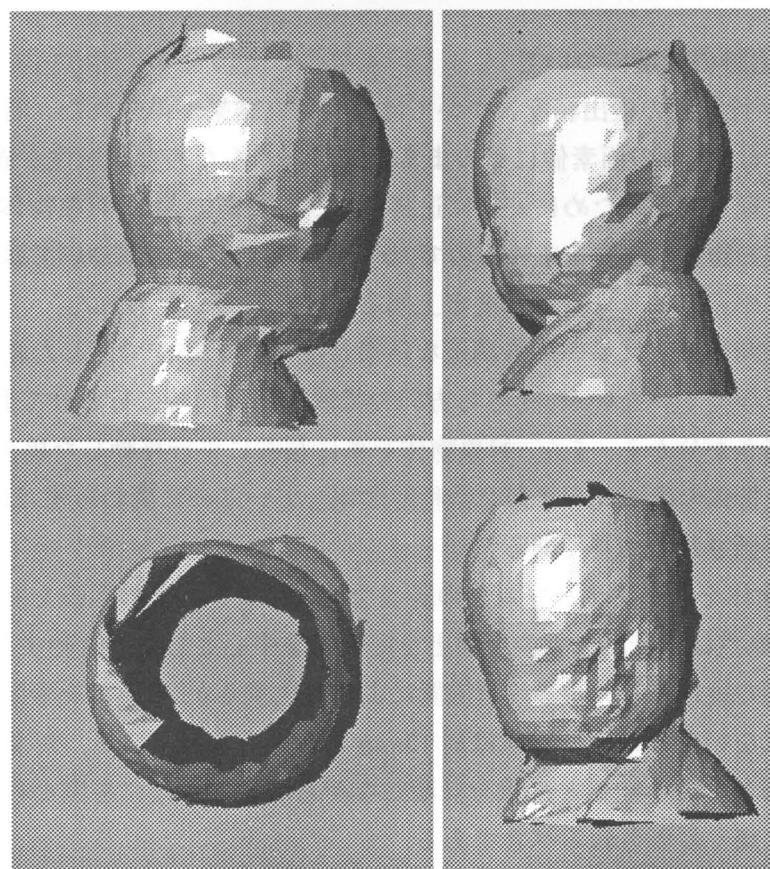


図 5.23 モデル構築と追加の様子



は単に回転されただけであるために、頭頂部及び胴から下は計測されていないが、目や鼻・口・耳など顔面上の細かな凹凸がモデリングされていることが分かる。また、衿の部分の形状も見てとれる。このモデルを構成する三角パッチの数は全体で 4484 個であった。

このモデルの形状の正確さを評価するために、形状モデルから生成した距離画像と、実物体を測定して得られた距離画像の間の誤差を光軸方向の距離差により評価した（表 5.23）。対象物体が初期位置にあるときには、計測された画素とモデルを構成する三角パッチの頂点が一致するためについへん正確なモデルとなっているが、対象物体が後退したり回転した場合には、それらの対応関係がなく誤差が増大する。特に、わずかな位置検出のずれ等によって、カメラから見た時の側面部では光軸に対する面の傾きが浅くなり、奥行き方向に発生した大きな誤差が誤差の最大値に影響を及ぼしているが、誤差の標準偏差は 1mm 以下と、ほぼレンジファインダの精度と同等の精度が得られた。

## 5.6 考察

この章で述べた動物体の実時間追跡手法では、対象物体の形状に制限を設けなかった。実際、ついていの物体について追跡が可能であることが確かめられたが、その精度は対象物体の形状にたいへん大きく左右される。5.5 節では手法の有効性を示すために、他の方法では追跡が困難であると考えられる形状の対象物体を用いたという点も勘案し、実験により得られた位置・姿勢の精度、生成されたモデルの精度と、追跡速度について考察する。

### 5.6.1 位置・姿勢追跡の精度について

5.5.2 で示した平行移動の検出精度については、満足な結果が得られた。特に奥行き方向の平行移動については、各々の画素値にそのまま変位が現れるために線形性が成立しやすく、また距離画像の変化量も大きいために安定に追跡を行うことができる。対象物体上の全ての計測点の情報を統合して物体の位置を求めるので、ランダム性のノイズについては大数の法則によりセンサそのものの精度よりも計測精度が上回るという現象が見られた。

光軸に垂直な方向の精度についても満足な結果となった。5.5.2 で示した結果では光軸に垂直な方向の運動を与えていないが、そちら方向の位置検出は同時に行われており、その直線性から検出精度の高さが分かる。また、対象物体を単に左右に平行移動した場合とは異なり、画面中における対象物体の像の大きさが 2 倍程度変化しているのに関わらず、ほとんど誤差が看取されない点は評価されるべきである。

5.5.3 では対象物体をターンテーブル上の適当な位置に置き、その回転を計測しているために、形状モデルの原点は回転軸に垂直な面上を回転軸を中心として円運動する。得られた円の直径はモデル原点の設定とモデルの設置位置により変化するが、モデルを意図的に回転軸からずらしたわけではないためにその直径は約 12mm と小さくなっている。対象物体の頭部の直径は約 60mm であり、この回転と形状を計測しながら測定されたにも関わらず、円形が復元されていることは高く評価されるべきである。

5.5.2 で示した実験結果では、対象物体は平行移動されているために回転角は一定であるべ

きだが、平均的に  $0.5\text{deg}$  程度、最大でも  $2\text{deg}$  以下の誤差が見られる。回転角の精度は平行移動に較べ劣るように見えるが、これは対象物体の形状が球形に近く、回転に起因する距離画像の変化がわずかであるからである。対象物体表面のうねりは大きな部分でも  $5\text{mm}$  程度と小さく、これにより回転が検出されたということだけでも評価に値する。また、5.5.3 の実験では姿勢推定と同時平行して対象物体のモデルを構築したために、回転角推定の誤差がモデルに蓄積することによって次第に角度の推定値がドリフトする現象が見られた。最終的に蓄積した誤差は  $10\text{deg}$  程度と大きいが、当然最初にモデルを与えて実験を行えば誤差は遙かに減少する。しかし、モデルを逐次的に構築する手法でも角度検出が可能であることを示すために敢えてこの実験結果を掲載した。

### 5.6.2 構築された形状モデルの精度について

モデル化された形状の誤差の要因は、センサ計測誤差と姿勢検出誤差の二つである。センサ解像度の問題などから細部構造の復元には弱いが、全体の形状としては非常に正確にモデリングされていると言える。最大の問題は追跡に伴う姿勢推定誤差の蓄積であるが、これを追跡実行時に補償することはたいへん困難である。この誤差をさらに減少させるためには、多数の研究成果が発表されている、時間はかかるがさらに高精度な位置合わせ手法や3つ以上の距離画像を同時に位置合わせする手法など [Masuda96] [Stoddart96] を用いてオフライン的に最適化を図るのがよいと思われる。

距離画像は物体の姿勢変化に応じて次々と得られるため、これらの重なり領域を統合することにより精度の向上・ノイズの除去を行うことも可能であると考えられる。これについても多数の研究が行われているため、用途に応じた様々な手法が利用可能であろう。

### 5.6.3 計測時間について

5.5.1 で述べたように、追跡処理には2台のWSを並列動作させることにより高速化を図った。今回の実験ではシリコンレンジファインダの計測速度を毎秒15フレームに設定した。これにより計測・計算された距離画像はEthernetを介してもう一台のWSに転送されるが、センサ画素数が少ないためにこれによる速度低下は見られない（次章で詳しく論じる）。2台目のWSにおいて、最も時間がかかる処理は距離画像の生成である。このWSにより距離画像の生成処理のみを実行した場合、毎秒100枚の距離画像が生成可能である。1回の追跡ステップには7枚の距離画像が必要であり、最終的な繰り返し速度は毎秒約7回であったため、1秒間に49枚の距離画像が生成されていることになる。つまり距離画像生成に費やされる時間は全体の約半分を占めている。残りの半分は、距離画像の比較による姿勢推定や通信・画面表示などである。センサ計測速度にはやや劣るが、その高速性を生かすことはできたと自負する。

## 6 章 濃淡画像と距離画像の融合による 自由曲面物体の実時間追跡

前章では、距離画像のみを用いて対象物体の6軸追跡を実時間的に行う手法について述べた。しかしこの方法では対象物体の形状が平面・円筒面、もしくは球面である場合には運動を検出することのできない自由度が存在する。このように距離画像に変化が生じない自由度に関しては、テクスチャ情報などの濃淡画像を用いるほかはない。そこでこの章では、5章の方法を濃淡画像情報を用いるように拡張することによりこのような問題を解決する手法について述べる。

濃淡画像は各点の反射率と照明条件によって決定されるので、対象物体上的一点の明度値は回転につれて変化する。つまり、濃淡画像をモデルに含んだシステムでは、表面反射率の復元を行わない限り一様光源の仮定が成立する環境下でしか用いることはできない。通常、表面反射率の復元計算は計算量が多いため、実時間的に復元を行うシステムはほとんど存在しない。そこで、5章で述べた手法と同様に、CG ハードウェアを用いて高速に反射率計算を行う手法を導入する。この手法によって、動物体の追跡と同時平行して対象物体の形状と反射率が逐次的にモデリングされていき、追跡が終了した段階で CG モデルが得られる。

LSI 技術を用いた高速レンジファインダは、静止物体の形状計測に用いられる高密度のレンジファインダと異なり、画素数が極端に少ないことが欠点として挙げられる。そのうえ、物体の運動を固定センサで計測するには視野を大きくとる必要があり、物体上の計測点数がさらに減少する。そこで、レンジファインダの計測方位を回転させることによりよって計測範囲の拡大を図る。また、センサの回転制御には姿勢の検出処理とは別に並列処理を行うことによりよって追従速度の向上を図る。

### 6.1 濃淡画像による運動推定

濃淡画像を用いて対象物体の運動を求める方法には多数の研究例があるが、ここでは特に本章で述べる方法に関連性が深いものを中心に取り上げる。他には、2.1.1 で述べた特徴点追跡に基づく方法などが代表的である。

#### 6.1.1 勾配法によるオプティカルフローの算出

対象表面の明度の分布が滑らかに変化していると仮定すると、対象物体の移動につれて観測される画像上の明度は連続的に変化する。勾配法は、この明度の変化量と、対象物体の移動量との関連性から画像面内の移動量（オプティカルフロー）を求める手法である [Yachida90]。

いま、画像上の点  $(x, y)$  の時刻  $t$  における明るさを  $I(x, y, t)$  とし、微小時間  $\Delta t$  後に物体が  $\Delta x, \Delta y$  だけ移動したとする。ここで、物体上の点の明るさが移動後も変わらないと仮定すると、

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (6.1)$$

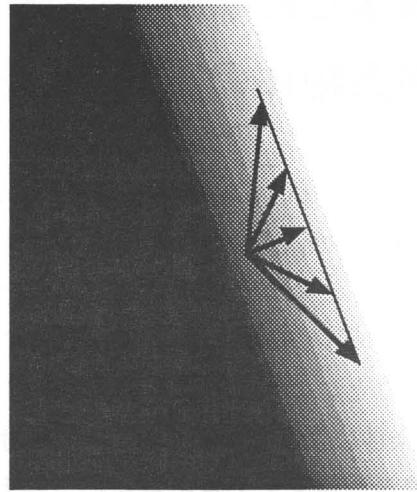


図 6.1 式 6.3 が定める直線

が成り立つ。右辺を Taylor 展開すると,

$$I(x, y, t) = I(x, y, t) + \Delta x \frac{\partial I}{\partial x} + \Delta y \frac{\partial I}{\partial y} + \Delta t \frac{\partial I}{\partial t} + e \quad (6.2)$$

となる。ここで  $e$  は  $\Delta x, \Delta y, \Delta t$  の高次の項であるが、これを無視して、両辺を  $\Delta t$  で割り  $\Delta t \rightarrow 0$  とすると,

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (6.3)$$

が得られる。 $\partial I / \partial x, \partial I / \partial y, \partial I / \partial t$  は画像から求められるから、この式は  $dx/dt$  についての 1 つの拘束条件を与える。すなわち、オプティカルフローは図 6.1 のようにある直線上に拘束されるということが得られるだけである。実際のオプティカルフローを求めるには他の拘束条件がさらに必要となり、このような拘束条件はいくつか提案されている。その中でも、Horn が提案した滑らかさの拘束 [Horn81] を用いる方法などが有名であるが、それらのほとんどは Marr のパラダイム [Marr82] たるボトムアップ的ビジョンの範疇にとどまるものであり、対象物体に関する知識を陽に用いる、いわゆるモデルベースビジョンへの発展がなされたものは少ない。しかし実際には対象物体が剛体であると仮定したとき、1 つの剛体に関するオプティカルフロー群は互いに強い拘束を受けているはずであり、これを用いることにより勾配法からも有用な情報を得ることは可能であると考えられる。

もう一つの問題は、オプティカルフローからいかにして対象物体の実運動を求めるかという問題である。例えば、右向きのオプティカルフロー群が得られたとき、これは対象物体が右へ平行移動したことによるものか、それとも対象物体が縦軸まわりに向かって右へ回転したことによるものかを判定する必要があるということである。この問題は勾配法に限らず、エッジ追跡や特徴点追跡などの他のオプティカルフロー解析についても同様に発生する。これについて

も、対象物体の剛体性に基づく拘束条件を用いることにより解決を図ることが必要であると考えられる。

### 6.1.2 特徴のモデル照合に基づく対象物体の追跡

画像から得られた特徴（点、エッジ等）と、対象物体のモデルとを逐次的に対応付けながら対象物体の位置・姿勢を求める方法である。対象物体の形状や大きさを表す三次元モデルが与えられたとき、一般的に3点の対応がモデルと画像間で取られれば、対象物体の位置と姿勢が一意に決定する。また、エッジの場合には、3本のエッジがモデルと対応付けられれば同様に位置と姿勢が定まる。静止物体の姿勢を求めるためには、この特徴をシーンから得られる構造的情報を手掛かりにモデルと対応付ける方法が考えられるが、動物体の追跡では前回の追跡処理の結果モデルと実対象の位置・姿勢が近いと仮定することができ、その周辺を探索することにより対応付けることが可能である。対象物体の形状情報によりスケールファクタが与えられるために、一眼のみによっても奥行き方向の運動を推定することが可能である。

Starkは、Active Contourと呼ばれる輪郭モデルを用いて対象物体の運動を追跡する手法を提案している [Stark96]。この手法では、図6.2のように対象物体の三次元モデルから作成した遮蔽輪郭 (occluding contour) を、実画像から抽出した輪郭にフィッティングすることにより、モデルのエッジ線と実物体のエッジ線を対応付けて姿勢を検出する。また Armstrongは、形状に起因するエッジのみならずテクスチャによるエッジもモデルに含め、モデルを仮想的に画像に投影したときのエッジと、実物体の像から抽出したエッジ間の距離が二次元画像上で最小となるように運動パラメータを最適化するという手法で6軸運動を実時間的に追跡することに成功した [Armstrong95]。この手法では、図6.3のように対象物体が他の物体に隠されたような場合でも、その領域ではエッジ間の対応がつかず結果的に無視されることにより正常に追跡を続

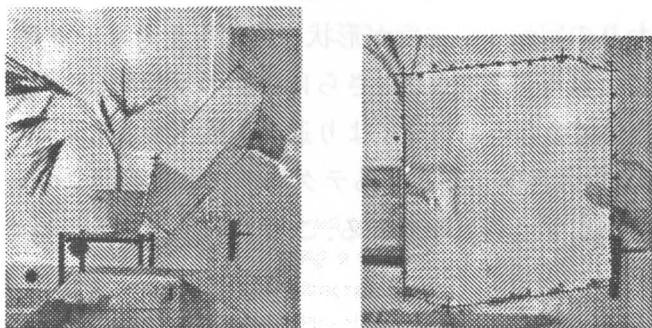


図 6.2 Active Contour による動物体追跡 [Stark96]

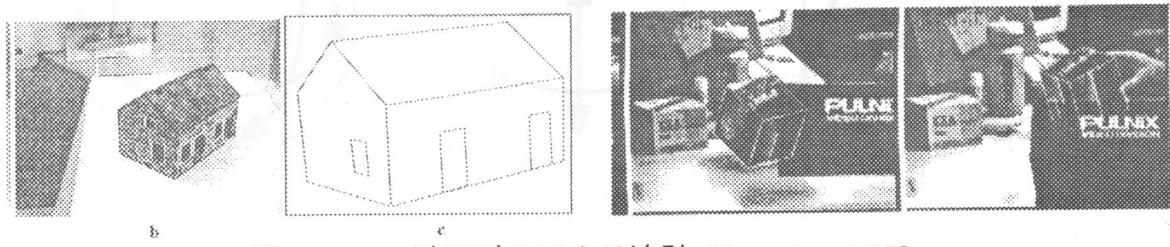


図 6.3 エッジモデルによる追跡 [Armstrong95]

けることができる。また、複数の画像について同時に距離の評価を行い、総合的に最適化することにより、そのまま簡単にステレオ視や三眼視に拡張することができるため、奥行き精度が重視されるような場合にも用いることが可能である。このように、モデルとの照合による追跡法では、対象物体の一部が他の物体により遮蔽されたり、背景に他の物体のエッジが存在するような場合でも、追跡に必要な情報を選択的に抽出・照合することが可能であり、安定度・処理速度に対する効果が高い。その反面、特徴量を抽出しているために自由曲面物体に応用することは比較的困難であると考えられる。

## 6.2 濃淡画像と距離画像による動物体追跡

5章で述べた距離画像による動物体追跡では、多数の生成された距離画像を用いて姿勢の検出を行っており、単一の画像面内の処理はほとんど行わないことにより高速化を図っていた。またこのことにより、モデル内に陽に特徴に関する知識を埋め込まなくても追跡が可能となり、同時並行的にモデリングを行うことが可能となった。また、特徴の抽出が困難な複雑な形状の物体に対して適用可能となった。ここでは、この性質を保ったまま濃淡画像による情報を融合することでより安定な追跡を実現する手法について述べる。

### 6.2.1 濃淡画像を用いた移動の検出

距離画像によって運動が抽出できない物体形状の代表例は平面・円筒面・球面である（図6.4）。ここで平面の大きさ、および円筒面の軸方向の長さが計測範囲に較べて十分大きい場合などを考え、ここでは境界に関する情報を用いないと仮定する。この仮定のもとでは、平面の面内平行移動（2自由度）及び平面に垂直な軸まわりの回転は形状に変化が現れないために追跡することができない。また、円筒については軸まわりの回転と軸方向の平行移動が、また球面については中心まわりの回転3自由度が形状に変化を生じない運動であり、これらは距離画像によって追跡することは不可能である。さらに表面の反射率が一定であり、濃淡画像にも変化が生じないような物体はそもそも画像により運動を検出することが不可能であるのでここでは扱わない。つまり対象物体表面に存在するテクスチャの移動により距離画像情報では抽出が不可能な方向の運動を抽出することを考える。このような状況は特殊であるかのように思われ

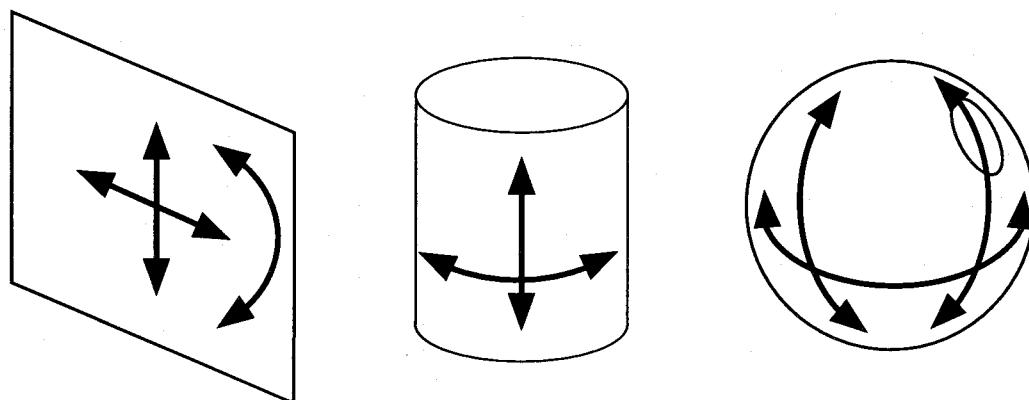


図 6.4 距離画像により運動が検出できない形状とその自由度

るが、自然物・人工物を問わず、全体的もしくは部分的にこの三形状やそれに近い形状をなす物体が多い。また一部に特徴的な部分が含まれていても、その領域の面積が比較的小な場合はやはり安定に運動を追跡することができない。そこで、常に濃淡画像による情報を併用することにより、対象物体の形状に関する手法の一般性をより向上させることを目指す。

### 6.2.2 対象物体の姿勢に関する明度勾配法

6.1.2で紹介した特徴のモデル照合による手法では、画像面内を探索することにより対応付けを行う必要があるために特徴種類の限定により一般性が高いとはいえない。また、6.1.1の勾配法では、モデルに関する知識を陽に用いないことによる限界が明らかである。そこで、勾配法を、画像面内の移動とそれに伴う明度値の変化と捉えるのではなく、対象モデルの運動に伴う明度変化と考えることにより、直接的に明度値の変化から運動量を算出することを考える。

いま、時刻  $t$ において対象物体の姿勢が  $n$  個のパラメータ  $p_1 \dots p_n$  によって決定されるとする。このとき、対象物体上のある画素の明度が  $I(p_1, \dots, p_n)$  で与えられるとする。対象物体の像面上の移動量（オプティカルフロー）は姿勢パラメータの変化に対して連続的かつ滑らかである。また、対象物体表面の明度値の変化も滑らかであると仮定すると、姿勢の変化に起因する像面の明度変化はまた滑らかである。そこで、時刻  $\Delta t$  後に姿勢が  $\Delta p_1, \dots, \Delta p_n$  だけ変化したときの個の画素の明度は、式 6.2 と同様に Taylor 展開によって

$$I(p_1 + \Delta p_1, \dots, p_n + \Delta p_n) = I(p_1, \dots, p_n) + \Delta p_1 \frac{\partial I}{\partial p_1} + \dots + \Delta p_n \frac{\partial I}{\partial p_n} + e \quad (6.4)$$

となる。高次の項  $e$  を無視して一次近似し、右辺第一項を左辺に移項すると、

$$I(p_1 + \Delta p_1, \dots, p_n + \Delta p_n) - I(p_1, \dots, p_n) = \Delta p_1 \frac{\partial I}{\partial p_1} + \dots + \Delta p_n \frac{\partial I}{\partial p_n} \quad (6.5)$$

となる。この式により、時刻  $\Delta t$  後の明度値の変化と各姿勢パラメータに関する一つの拘束条件が与えられる。この拘束条件により、対象物体の運動が 1 自由度であれば  $n=1$  となり明度変化から姿勢の変化量が一意に決定することになるが、6 軸運動を仮定するならば  $n=6$  となるため、通常の勾配法に較べより不利な条件であるかのように思われる。しかし実際には、対象物体の剛体性の仮定により、一つの対象物体の像に関し姿勢パラメータ  $p_1 \dots p_n$  が全画素について同一であるという強い拘束条件が存在する。これはちょうど 2 次元画像の勾配法について、ある領域全体が変形せずに平行移動していると仮定することと同程度の効果を持つ。この剛体性の拘束条件を用いるためには、画像から対象物体についての像領域を切り出す必要があるが、この手法については後に述べる。

### 6.2.3 CG ハードウェアによる明度変化率の計算

勾配法では像面内の画像処理により  $\partial I / \partial x, \partial I / \partial y$  が求められたが、式 6.5 の  $\partial I / \partial p_i$  ( $i =$

1...n) は单一の像から計算することは難しい。そこで、距離画像の時と同様に、対象物体モデルから仮想的に生成することを考える。 $\partial I / \partial p_i$  は次式のように変形することができる。

$$\frac{\partial I}{\partial p_i} = \frac{\partial I}{\partial x} \frac{dx}{dp_i} + \frac{\partial I}{\partial y} \frac{dy}{dp_i} \quad (6.6)$$

この式で、左辺は各自由度の変化に対する明度値の変化であり、これは各自由度の変化に起因するオプティカルフロー ( $dx/dp_i, dy/dp_i$ ) と、勾配 ( $\partial I / \partial x, \partial I / \partial y$ ) から計算することができる。これらのうち後者は勾配法と同様に計算することができるが、前者は画素位置に対応する対象物体上の点の座標が得られる必要があり、計算量が高い。そこで、これらを式 6.6 のように分解して求めることなく、5.3 節と同様に対象物体の姿勢に微小変位を与えたときの明度値の変化を仮想的に求め、これにより姿勢を検出する。対象物体はその形状と表面の明度値が既にモデル化されており、任意の姿勢値から実環境内における撮影画像と同等の画像を生成することが可能であるとする。すなわち、任意のパラメータに対する  $I(p_1, \dots, p_n)$  を計算することができるとして、 $\partial I / \partial p_i$  を偏微分の定義

$$\frac{\partial I}{\partial p_i} = \lim_{a \rightarrow 0} \frac{I(p_1, \dots, p_i + a, \dots, p_n) - I(p_1, \dots, p_n)}{a} \quad (6.7)$$

の極限をとらない形、すなわち  $a$  に有限の値を与えたときの差分値をもって微分値に替えることができる。

最初に対象物体表面の明度値の変化は滑らかであると仮定したが、一般的にこのような仮定は成立しない。しかし、勾配法では対象の像にぼかしフィルタを施し像面の明度値を滑らかにすることでエッジ状の特徴についてもオプティカルフローを計算することができる。同様にここで提案する手法でもぼかしフィルタを用いることができるが、計算量が大きいために実時間処理の妨げとなる。そこで、対象物体表面の明度値をモデリングする段階で先にモデル上ではぼかしを施しておくことで、単にその像を CG 的に生成することにより滑らかな明度変化がある像を生成することを考える。

対象物体表面の各点の明度値を十分滑らかに表現し得るビットマップをテクスチャマッピングする手法がまず考えられるが、この方法では画像生成コストが高い上、ビットマップの容量が大きくなるという問題がある。そこでここでは 5.2 節と同様に、CG ハードウェアにより高速に描画可能である手法として、グーローシェーディング (Gouraud shading) を用いた。通常、ポリゴンモデルに光源モデルを用いてシェーディング表示するとき、それぞれのポリゴンの法線方向と光源方向・視線方向等の情報を用いて明度値を算出し、その明度値をポリゴン全体にわたって適用する。これをフラットシェーディング (flat shading) といい、この手法では姿勢の微小変位に関する明度値の変化を求めることができない（図 6.5）。グーローシェーディングは、各頂点ごとに法線ベクトルや反射率を独立に設定し、各頂点ごとに求めた明度値を用いて

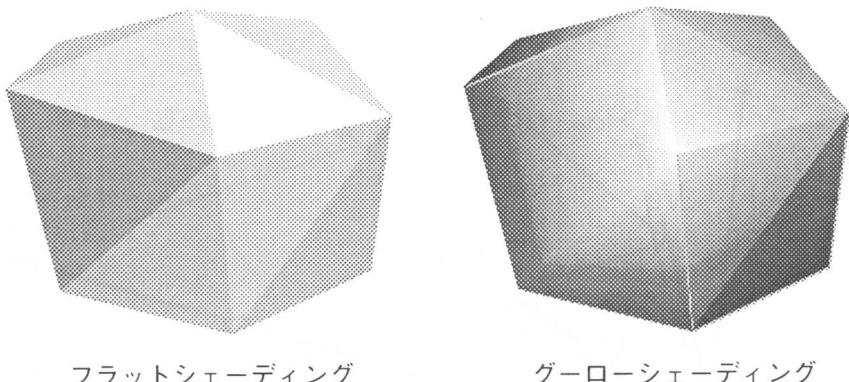


図 6.5 シェーディングモデル

ポリゴン内部を線形補間する手法である。この方法ではポリゴンモデルがある程度粗であっても、各頂点間の明度値はなめらかに表示されるため、ポリゴンモデルにありがちなエッジ感がない画像を生成するために用いられる。この特性はそのまま非常に勾配法と適合度が高い。すなわち、対象物体上の表面明度を適当な間隔でサンプリングし、それを各頂点の明度とするだけで、その間が線形に補間された画像が得られるため、モデルに微小変位を与えて得られる明度値の変化率を、微小変位の大きさによらずに安定して得ることができるからである。

式 6.5において、右辺は直接画像から得られる情報であり、微小時間を経た前後の画像から用いられる明度値の差分である。これらをそれぞれ実画像とする方法もあるが、この方法では得られた移動量  $\Delta p_i$  ( $i = 1 \dots n$ ) をモデル姿勢に反映しても前回の処理によって得たモデル姿勢の推定誤差がそのまま残存することになる。そこで、右辺第一項は実画像とし、第二項は前回の推定によって得られたモデル位置について生成された画像を用いることとする。これによってモデルに残存する誤差は明度差に反映され、その誤差を減少するように移動量が求められることになる。この場合、明度値の差をとることができると想定される画素は、モデルから生成された画像領域内に限定されるが、これはシーン中から対象の存在する領域を切り出していることに他ならず、先に問題となった剛体性が成立する領域のみを抽出することができるという利点がある。

#### 6.2.4 剛体性の拘束条件に基づく姿勢変化量の算出

勾配法に基づくオプティカルフローの抽出では、局所的情報のみから完全なフローベクトルを得ることができなかったため、滑らかさの仮定など他の拘束条件を用いることは 6.1.1 で述べた。ここでは、単一剛体の像領域内の明度勾配を統合して、この剛体の姿勢変化量を算出す方法について述べる。

式 6.5において、左辺の 2 項および右辺の微分係数  $\partial I / \partial p_i$  ( $i = 1 \dots n$ ) の算出が可能であることは 6.2.3 で述べた。ここで、領域内の 1 個の画素に番号  $j$  を振り、その画素の明度変化量とそれとの微分係数を次のように変数  $m_j, v_{ij}$  とおく。

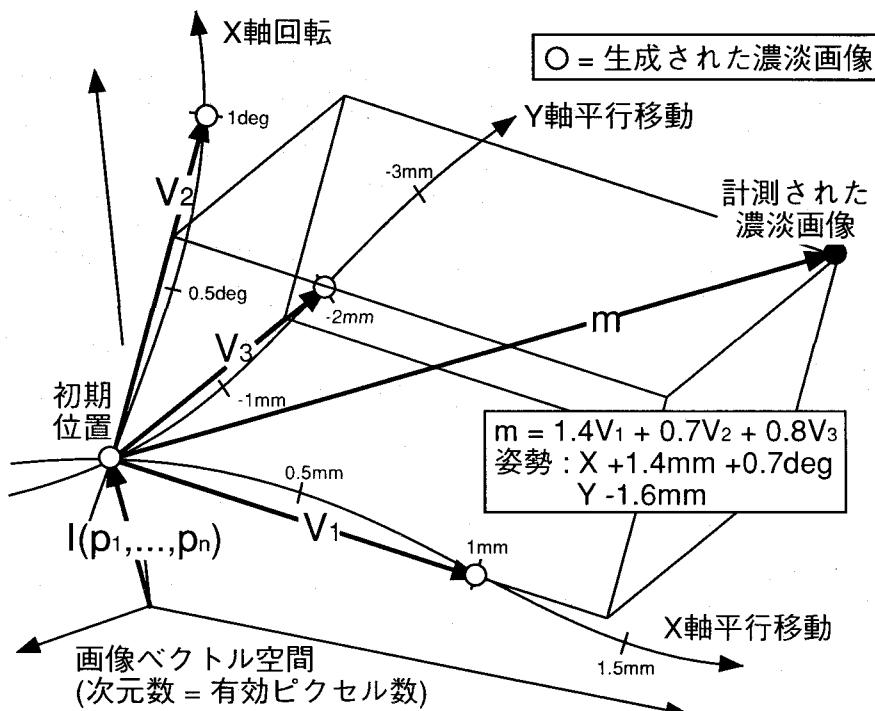


図 6.6 画像ベクトル空間における線形近似

$$m_j = I_j(p_1 + \Delta p_1, \dots, p_n + \Delta p_n) - I_j(p_1, \dots, p_n) \quad (6.8)$$

$$\nu_{ji} = \frac{\partial I_j}{\partial p_i}$$

このとき、6.5 式は

$$m_j = \sum_{i=1}^n \nu_{ji} \Delta p_i \quad (6.9)$$

と書き表すことができる。このうち剛体性の拘束条件より、 $\Delta p_i$  は全ての画素について共通であることから、全ての画素に関する情報を統合し、

$$m = V \cdot \alpha \quad (6.10)$$

と書き表すことができる。ここで、

$$m = \begin{bmatrix} m_1 \\ \vdots \\ m_l \end{bmatrix}^T, \quad V = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{j1} & \cdots & v_{jn} \end{bmatrix}, \quad \alpha = \begin{bmatrix} \Delta p_1 \\ \vdots \\ \Delta p_n \end{bmatrix} \quad (6.11)$$

となる。6.11 式は明度勾配の線形性、姿勢と像面移動量の線形性などの近似が含まれているた

めに、実際には両辺の自乗誤差を最小化する規範として 5.14 式と全く同一の式

$$\min_{\alpha} |m - \mathbf{V} \cdot \alpha|^2 \quad (6.12)$$

における  $\alpha$  を求める。 $\alpha$  は 5.16 式と同一の式

$$\alpha = (\mathbf{V}^T \cdot \mathbf{V})^{-1} \cdot \mathbf{V}^T \cdot m \quad (6.13)$$

によって計算することができる。参考に図 5.10 とほぼ同一の図 6.6 を示す。

### 6.2.5 距離画像との融合

5.3 節とこの節の内容を較べれば、これらの手法が非常に似通っていることが分かる。6.2.1, 6.2.2 では勾配法との対比のために偏微分モデルを用いて解説した。しかし、明度の勾配である微分係数を求めるために 6.7 式の極限を取らず、有限の微小移動量に関する明度値の変化を用いて濃淡画像の変化率を求めた。このため、実際に行われる処理は 5.3 節で述べた距離画像を用いた方法と同じである。また、距離画像における計測点の間にポリゴンを張るということは、隣接点間に平面を張るということであり、これは計測点の間を線形補間していると言うことができる。この点で、滑らかに明度値が変化する濃淡画像を頂点の明度値を用いてグローチェーディングすることと等価である。

プログラムの実装という点でも、これらの処理は同時並行的に行われる。すなわち、対象物体の計算機内モデルとしてはポリゴンモデルの各頂点に反射率を与えた形式とし、これをグローチェーディングする。この結果得られた Z バッファと濃淡画像バッファの双方を読み込むことでそれぞれ仮想的な距離画像と濃淡画像を生成することができる。濃淡画像を生成することによるオーバーヘッドはフラットシェーディングではなくグローチェーディングとすることによるものと、濃淡画像の読み出し時間であるが、これらはハードウェアの構成にもよるが無視できるぐらい大変に短い。

姿勢推定段階では、対象物体の移動量を距離画像と濃淡画像からそれぞれ独立に推定するのではなく、双方の誤差を最小化する共通の姿勢パラメータ変化率を求める。つまり、得られた距離画像と濃淡画像をベクトル化し、それらをそれぞれ 5.14 式と 6.12 式の  $m, V$  とするが、これらを単に縦に連結して演算することで共通の  $\alpha$  を計算することができる。ここで、この行列に格納される値は距離画像の部分は距離値、また濃淡画像の部分は明度値である。仮にこれらの値のとり得る範囲が大きく異なると、自乗誤差は大きな値をとる方に大きく左右されるため、双方の情報がバランス良く反映されない。例えば、距離画像の単位系が mm であり、濃淡画像の値が 8bit であると、濃淡画像の明度が白から黒に変化することの効果が距離値にして 255mm 移動したことと等価とみなされることになる。そこで、距離画像と濃淡画像の値の比率を決定する係数  $k$  を与える必要があるが、これは単位系・センサ精度・対象物体表面のコントラストなどによって調整するほかはないと思われる。

$$m = \begin{bmatrix} m_r \\ km_i \end{bmatrix}, V = \begin{bmatrix} V_r \\ kV_d \end{bmatrix}, \alpha = (V^T \cdot V)^{-1} \cdot V^T \cdot m \quad (6.14)$$

距離画像の時に行われる他の処理については、ほぼ濃淡画像についても同様の処理を行うか、もしくは共通化できるものが多い。例としては、[5.2.3 CG ハードウェアとレンジファインダ座標系の一致], [5.3.2 微小変位量の決定と背景画素の処理], [5.4.2 複数距離画像の統合] で述べた内容がそれらに相当する。濃淡値のモデリングについては次節で詳しく述べる。

### 6.3 CG ハードウェアによる実時間反射率推定

これまで、対象物体のある一点の明度は対象物体の姿勢によらず一定であると仮定してきた。しかしこれは一様光源（環境光）のもとでのみ成立することであり、実際には光源や視点の位置によって変化する。勾配法によるオプティカルフローの算出では、短い時間間隔をおいて前後する2枚の画像を用いて計算するため、その2枚の画像に撮影されている対象物体の姿勢は大きく変化していないという仮定をおくことができる。しかし6.2節で述べた方法では計算機内モデルと実画像の差からモデル移動量を算出するため、追跡に従い対象物体の姿勢が変化していくと計算機内モデルの推定姿勢はモデリング時の姿勢と大きく異なるために生成された濃淡画像の明度値が実際とは大きく異なることになる。例えば、対象物体が向かって左側から照明されている場合では、対象物体の右側のある点の明度は相対的に暗いが、この物体が左へ回転していくにつれこの点に光源からの光が届くようになり次第に明度が増す。このような場合では、単にモデルに画像上で得られた明度値を格納する方法では正常な追跡ができなくなる。そこで、モデルとして対象物体の表面反射率を用い、画像生成時には照明条件と反射率から明度値を計算することにする。

#### 6.3.1 実画像と CG の比較による反射率推定

反射率を用いた明度計算には、濃淡画像の生成段階において光源計算が余計に必要となるが、我々が用いたものを含め CG 描画ハードウェアは光源計算も高速化することができるため、この計算時間は問題とはならない。ただし、どのような光源計算モデルをも高速化されるわけではないため、個々の CG ハードウェアによって実行可能な計算モデルに従った反射率をモデリングする必要がある。また、CG ハードウェアによる光源計算は、最終的に人間が表示物体の形状を把握するための手段として提供されているものであり、物理的に正確な輻射強度が計算できるものではない。より具体的に言うと、画面に表示される明度値は実世界と同等の dynamic range を持たないために、CG ハードウェアを駆動するために用いられるさまざまなパラメータは、lux や candela といった物理的な単位系に即したものではなく、0 から 1 までの範囲の数値などの比例系で表される。そのため、今回の目的のように CG ハードウェアによって生成される画像と実画像との明度値を比較するような場合には、一度生成された CG 画像と実画像を比較してその比例定数を求めておく必要が生じる。これを人手で調整することは避けたいので、ここでは自動的にこの比を吸収することができるよう、CG ハードウェアによっ

て生成された画像とモデリングすべき実画像を比較することによって反射率を算出する。

ここでは仮にCGハードウェアには実世界と同等の照明条件と反射モデルが設定されており、その比率のみが分かっていないとする。また、対象物体の形状モデルは距離画像から変換されることにより既に完成しているものとする。ここでまず、仮の反射率をモデルに設定し、一度CGハードウェアにより描画を行う。これにより、この仮の反射率に対応する明度値が得られる。これは当然実画像の明度値とは異なるために、実画像の明度値と仮の明度値の比率を仮の反射率に乗じる。この計算により、もう一度CGハードウェアにより描画を行うと実画像と同一の明度値がCGとして得られるはずである(図6.7)。ここで注意せねばならないことは、仮の反射率によって描画された濃淡画像がフレームバッファによって表現し得る値の範囲を超えてしまい飽和した場合、補正計算をしても正確な反射率が求められないことである。光源位置の設定については、CGハードウェアは5.2.3で述べた方法により実座標系と一致するように初期化してあるために、実在の環境を測定した値をそのまま用いることができる。

この計算では、CGハードウェアによって実行される反射モデルに依らず、常に実画像と同一の明度値を生成するような反射率を求めることができる。しかし、この反射モデルに与えることができる反射率パラメータが2つ以上ある場合、その比率を求めることはできない。例えば、対象物体の表面に鏡面反射が見られ、鏡面反射と拡散反射の比率がいたるところ一定であれば、この比率をあらかじめCGハードウェアに与えておくことで反射率を求めることができが、この鏡面反射率と拡散反射率の比率が対象物体の部位によって異なる場合にはその比率を求めることはできない。通常、CGハードウェアによって計算が可能な反射モデルは、光源として環境光(一様光源)成分と平行光源を与えることができる。また、対象物体表面の反射率として、拡散反射率と鏡面反射率、それに鏡面反射の鋭さに関するパラメータを与えることができる。これらのうち、環境光源と平行光源の比率はあらかじめ測定することにより求められる。また、拡散反射率と鏡面反射率の比率、また鏡面反射の鋭さについては、今回は一定で

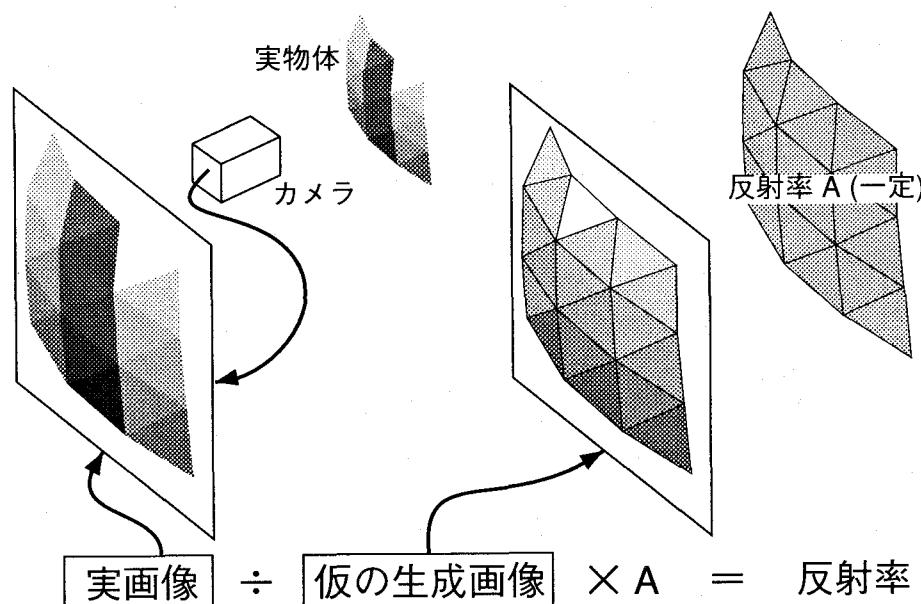


図6.7 反射率推定の流れ

あるとした。これらのパラメータが実際と大きく異なる場合、対象物体の姿勢が変化しない間は実画像と同一の明度値が生成されても、対象物体の姿勢が変化するにつれて誤差が生じてくる。

### 6.3.2 明度値のサンプリングと法線ベクトル

CG ハードウェアによって描画されるポリゴンモデルでは、各ポリゴンの頂点についてのみ反射率パラメータを設定することができる。そこで、各頂点に対応する実画像上の明度値を求め、これを用いて 6.3.1 の補正計算を行う必要がある。幸い、5.2.3 の計算により描画される CG の大きさと位置は実画像と一致している。また、シリコンレンジファインダの計測点が実画像上でどの画素にあたるのかは、4.2.2 のキャリブレーションによって得られている。距離画像からモデリングが行われるとき、距離画像の計測点同士を結んでポリゴンを生成するために、各々のポリゴンの頂点は距離画像の計測点と一致する。そこで、ポリゴンの頂点に対応する距離画像の計測点の位置を実画像上の位置に変換し、その画素位置の明度値を利用して補正計算を行うことにより全てのポリゴン頂点の反射率を求めることができる。

計測された距離画像の精度が十分でない場合は、個々のポリゴンの法線方向に乱れが生じることがある。これは [3.1.2 法線ベクトルと光源方向を既知とする方法] で示した図 3.1 にあるように、計測点の密度が高いほど問題となる。シリコンレンジファインダは計測点の画素密度が低いために大きく法線方向が乱れることはないが、さらに高密度なレンジファインダを用いた場合には簡単な平滑化処理によってなめらかなポリゴンモデルにしておく必要がある。

## 6.4 センサ計測方位の回転による対象像の追跡

動物体の回転運動を正確に求めるためには、画像内に対象物体の像をできるだけ大きく撮影する必要がある。しかし、これでは対象物体が大きく移動したときに、視野から対象物体が外れてしまう。そこでセンサの計測方位を対象物体の方向へ追従させることで、対象物体の姿勢検出精度を高く保ったまま移動範囲を広くすることを行う。対象物体への方位を決める方式として、6 軸追跡の結果を用いるのではなく、距離画像に簡単な処理を施すことで高速に実行する手法について述べる。また、この追跡処理や距離画像の生成処理を 2 台のワークステーションを用いて並列的に実行する手法についても解説する。

### 6.4.1 対象像の追跡とカメラ制御手法

我々が用いている距離画像センサであるシリコンレンジファインダ [Yokoyama96] は、奥行き方向の精度はたいへん高いものの、画像面内の密度、すなわち画素数は通常のレンジファインダに較べたいへん少なく、縦・横ともに 24 画素である。これは、LSI 素子上に比較的面積の大きなフォトダイオードペアと、その信号の処理回路が組み込まれているからである。これはセンサの動作速度に影響の大きい感度を確保するためでもあり、技術の発展により解決する可能性があるもののやはり画素密度の問題は残存すると思われる。これと同様の問題は他の LSI 技術を用いた高速レンジファインダ [Kanade91] [Kang95] についても存在する。しかしこ

これらのレンジファインダが画素密度を犠牲にしてまで高速性を追求したのは、動的シーンの解析を距離画像で行うために開発されたからであり、対象物体の形状を精密にモデリングするためではない。また、一般的なシーンにおいてはほとんどの物体が静止しているということを考えると、ALV (Autonomous Land Vehicle) のようにセンサが移動体に取りつけてあるような場合にこのセンサの性能が発揮されると考えられる。さらに昨今のアクティブビジョン研究の隆盛に見られるように、センサの位置やパラメータ、さらには計測形態まで変化させることによりシーンの情報を積極的に獲得することが有効であると考えられるようになってきた。アクティブビジョンにはタスクの実行に必要な情報のみを効率よく取得することにより高い精度や高速性を得る目的があり、この点でタスクを無視してセンサを単に移動するだけでは意味がない。この点で、高速レンジファインダを生かした応用分野を開発するためには、このようにタスクに応じたセンサ計測方位の制御が重要であると考えられる。ここでは移動物体の6軸運動を実時間的に追跡するというタスクを考慮してセンサ計測方位の制御方法について考察する。

シリコンレンジファインダの計測点数は比較的少ないために、対象物体の回転運動を正確に検出するためにはできるだけ対象物体を画像中に大きく撮影し、対象物体上の計測点数を確保する必要がある。そのためには計測範囲を狭くする必要があるが、これでは対象物体が平行移動した場合に物体が計測範囲から逸脱しやすくなる。結果として、対象物体の回転運動の検出精度と平行移動範囲がトレードオフの関係となる。そこで、対象物体の個数が1つである場合、対象物体が計測範囲から逸脱する事がないようにセンサの計測方位を制御することでこの問題を解決する。

対象物体を追跡するためには、センサから対象物体までの方位を求める必要がある。対象物体の運動は6軸追跡の結果得られているため、この追跡結果を用いて方位を求めることがまず考えられる。しかしこの方法では、姿勢計算処理が終了してからセンサの姿勢制御が行われるために、そのむだ時間だけ制御が遅れることになる（図6.8）。そこで、ここでは高速化のために、より直接的な手法を用いることにした。この方法では、距離画像が得られた時点で対象物体表面の各計測点の座標値が分かっているので、それらの重心を求めこの重心に向けてセンサの方位を回転させる。この方法では、距離画像が得られた直後にセンサの回転方位が決定する

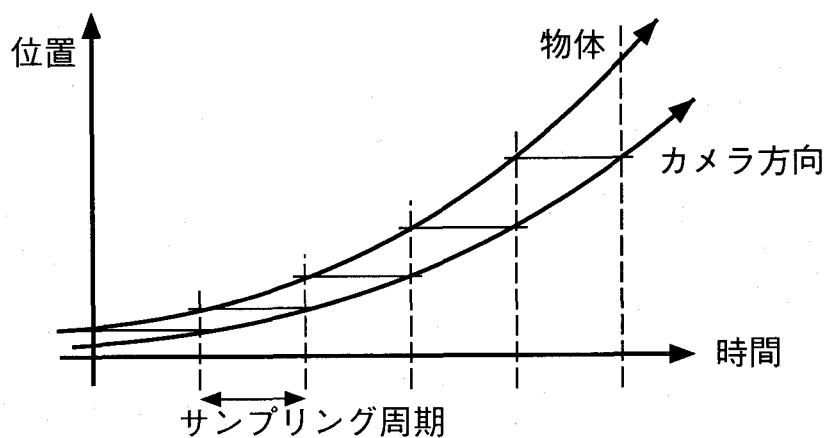


図6.8 センサ計測方位の遅れ

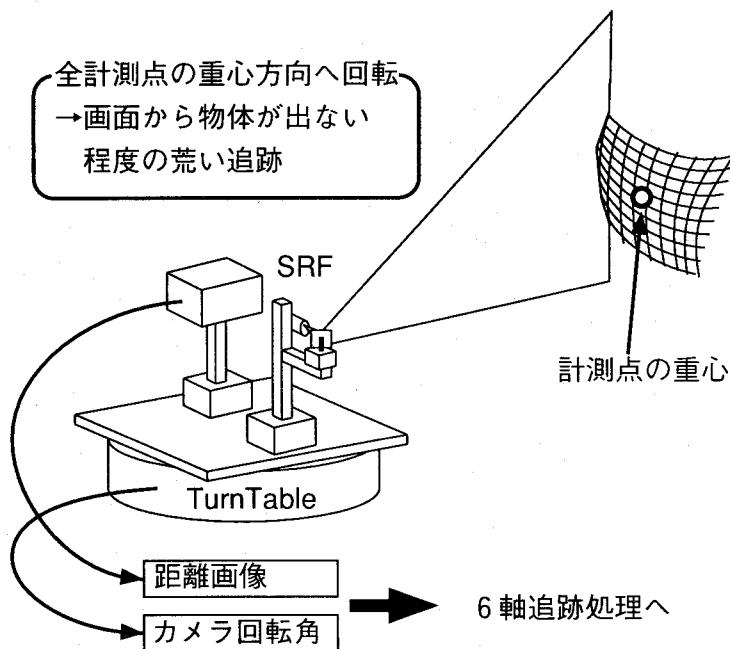


図 6.9 像の重心を用いた追跡

ために、時間遅れがたいへん短い。実際、距離画像を用いた姿勢推定では毎秒 7 回の速度で姿勢が検出されるが、距離画像の計測速度はこれより速く最低でも毎秒 15 回計測されるため、より細かい時間間隔でセンサ方位の制御を行うことができる。センサ方位の回転角は計算により求められているので、この回転角を距離画像と組にして 6 軸追跡処理部へ転送することで、世界座標系内の対象物体の絶対位置を知ることができ、カメラの回転角に多少誤差があったとしてもこれは 6 軸追跡の結果相殺される（図 6.9）。

これら二つの方式は主に追跡結果の応用目的によって良否が決定する性質のものである。対象物体の運動に応じて追跡とモデリングを繰り返し、最終的に CG モデルを得ることが目的であるならば、後者の像座標フィードバックがよいと思われる。なぜなら、対象物体が予期するよりも大きく、画面からはみ出しているような場合では、センサの回転が行われない。対象物体の移動について輪郭部が画面の端に現れたときのみ、この輪郭部を減少させる方向にカメラが制御される。結果としてセンサの回転は最小限となる。他方、対象物体の形状モデルが既に与えられており、シーン中のモデル化された部分のみを切り出して追跡するようなタスクでは、前者の物体座標フィードバックが良いと考えられる。例えば、ロボットアームがワークを掴んで移動するような場合、そのワークが落下したりずれて姿勢が変化したりしていないかどうかを検証するシステムでは、ロボットアームの像に影響されずにワーク部のみを正確に追跡する必要がある。また、運動している対象物体表面の特定の点にレーザを照射するような目的のためには、6 軸追跡の結果からその点の位置を参照し、その点が中央に配置されるように制御を行うことなどが考えられる。これらを整理すると図 6.10 のようになる。

#### 6.4.2 並列処理

これまでに述べた全ての処理を单一のワークステーション（以下、WS）により行うことは

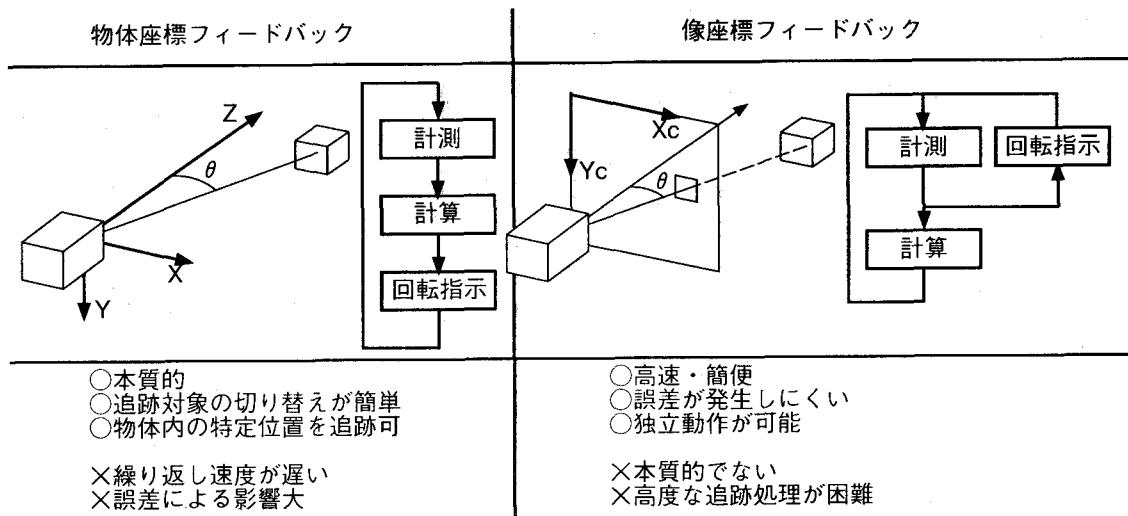


図 6.10 センサ方位算出方式とその性質

もちろん可能である。しかし、手法の並列性を明らかにし、複数の計算機による高速化の可能性を示しておくことは有意義である。著者が用いた機器構成では、シリコンレンジファインダが接続してある WS と、グラフィックス描画の速い WS が別であったために必然的に並列化されたという面もある。この機器構成から、図 4.13 に示したように計測処理と生成処理の対象性を用いて、それぞれの処理を分担するという考え方を並列化の基準とした。

2 台の WS における処理の分担を図 6.11 に示す。WS1 (SGI Iris Indy) にはシリコンレンジファインダとセンサ計測方位を変更するためのターンテーブルが接続されている。この WS では、シリコンレンジファインダの駆動とデータからの距離画像算出を行う。また、距離画像の計測点に対応する濃淡画像の画素から明度値をサンプリングする処理も行い、これらを Ethernet 経由でもう一方の WS へ転送する。距離画像は各々の計測点に対して double 値が XYZ

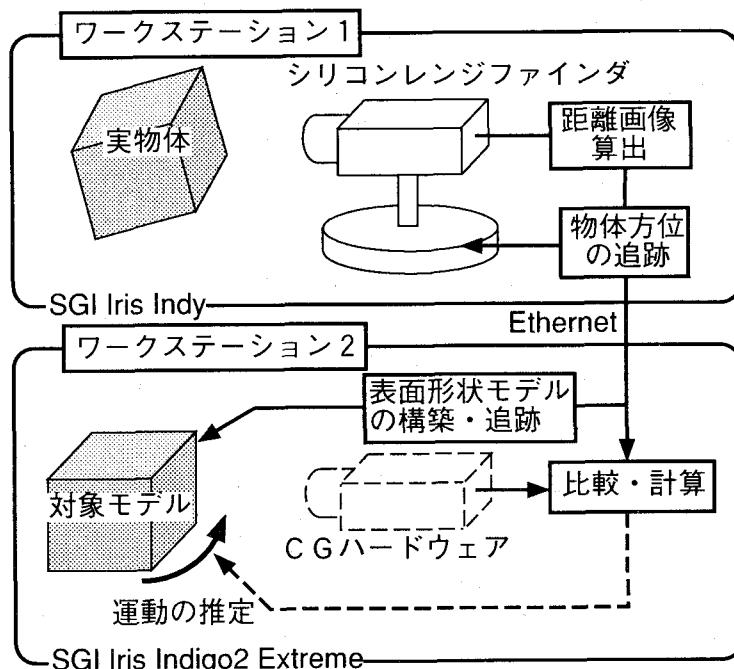


図 6.11 2 台の WS による並列処理

の3つであり、画素数が $24 \times 24$ であることから13824バイトである。また濃淡画像は既にサンプリングされているので同じく1画素1バイト、 $24 \times 24$ 画素であることから576バイトとなり、合計で15KBにも満たないためにEthernetで十分リアルタイムに転送が可能である。このWSでは距離画像を用いたセンサ計測方位の制御も行っている。WS2(SGI Iris Indigo2 Extreme)は高速なCGハードウェアを搭載している機種であり、距離画像の生成処理を行う。変位量の算出処理も行うが、単に6次の最小自乗法を行うだけであるからこの処理時間はたいへん短い。また、このWSでは適宜モデルの追加更新処理も行う。

これら2台のWSは同期をとらず完全に非同期で動作している点が特徴である。すなわち、WS1は次々と距離画像を計測してネットワークへ送出するので、WS2の処理速度に無関係に動作している。WS2はネットワークを介して送られてくる距離画像のうち、常に最新の距離画像を用いて追跡処理を行う。つまり、WS1の方が高速である場合には、WS2における1ステップ間に複数の距離画像が送られてくることがあるが、この場合最新のものだけを用いて古い距離画像は破棄する。逆にWS2の方が速い場合には、1ステップの処理の間に新しい距離画像が到着していない場合が生ずる。この場合は前回の距離画像を用いて同じ処理を繰り返することで姿勢推定値の精度を向上させる。

## 6.5 実験

この節では、5章の方法では追跡を行うことができなかった円筒の表面にテクスチャを配し、その軸まわりの回転角と軸方向の平行移動が正しく計測されるかどうか調べる。また、形状と反射率をモデリングした結果をシェーディング表示した画像も示す。

### 6.5.1 実験装置

実験装置の構成は5.5.1で述べた構成と同一である。ただし、5章ではキャリブレーションのために濃淡画像を用いていたが、この機器を用いて追跡実行時の濃淡画像計測を行う。濃淡画像はSGI Iris Indyに接続されたVME拡張ボックスに搭載した画像処理ボードにより取得する。このボードには画像処理ハードウェアTAICHI COREが搭載されているが、これは全く用いていない。このWSにより、シリコンレンジファインダの計測点に重なる画素の濃淡値をサンプリングし、もう一台のWSであるSGI Iris Indigo2 Extremeに転送する。このWSのCGハードウェアは、Zバッファを利用状態にし、かつダブルバッファ表示した場合にはフレームメモリが不足するために濃淡値がディザリング表示される。これにより濃淡値計算における量子化誤差が増大するため、ここではダブルバッファをOFFとした。

### 6.5.2 回転運動の追跡精度評価

5.5.3と同様、図6.12に示す装置構成により回転運動の精度評価を行った。ターンテーブル上に市松模様状のテクスチャを配した円筒を設置し、Ty軸回りに回転させた場合の運動を検出する。光源としてシリコンレンジファインダの背後約50cm、高さ30cmの位置に点光源を置いた。また、暗室を用いず、通常の室内程度の環境光のもとで計測を行った。点光源と環境光

の明度比率は4:1に設定した。シリコンレンジファインダと対象の間の距離は約55cmである。シリコンレンジファインダはある程度反射率の低い面についても計測を行うことができるが、完全な黒色では計測できないので、テクスチャの暗い部分は黒色ではなく灰色とした。円筒の直径は約8cmである。対象物体を計測装置のCCDカメラで撮影した像を図6.13に示す。点光源を用いたため、中央部に較べ左右がやや暗くなっている。対象物体のモデルをあらかじめ与えることはせず、追跡と同時平行してモデリングを行っている。

対象物体の回転に応じて得られた姿勢の変化の様子を図6.14に示す。対象物体は1degづつ回転させ、5degごとに追跡途中の姿勢を抜き出してグラフ化した。回転軸たるY軸まわりの回転については、誤差の標準偏差が5.21degであった。5章の方法と同様に、モデルに回転角の推定誤差が蓄積することにより次第に真値とずれていく様子がグラフから分かる。対象物体をほぼ1回転させたため、最初に計測された部分が元の位置に戻ってくることにより回転角の推定誤差が修正され、最終段階では誤差がまた減少していることも分かる。回転とは関係のないX,Z軸まわりの回転角はほぼ一定である。平行移動成分については大きな変化が見られる

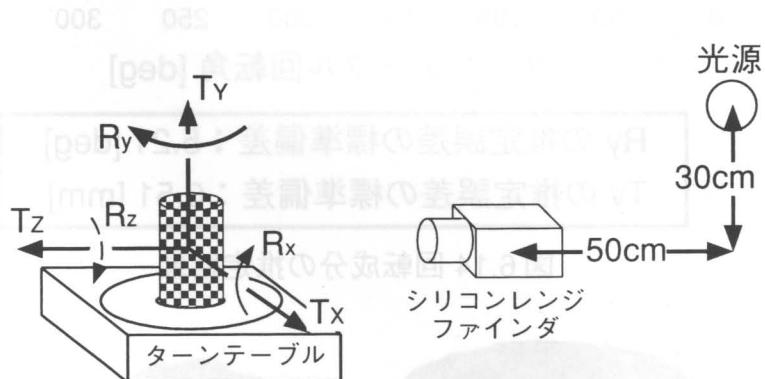


図6.12 回転運動追跡精度評価のためのターンテーブルの配置

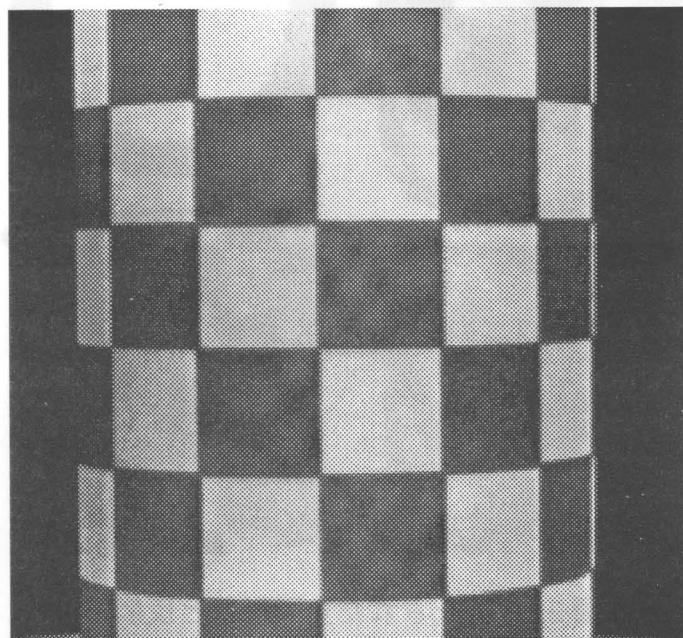
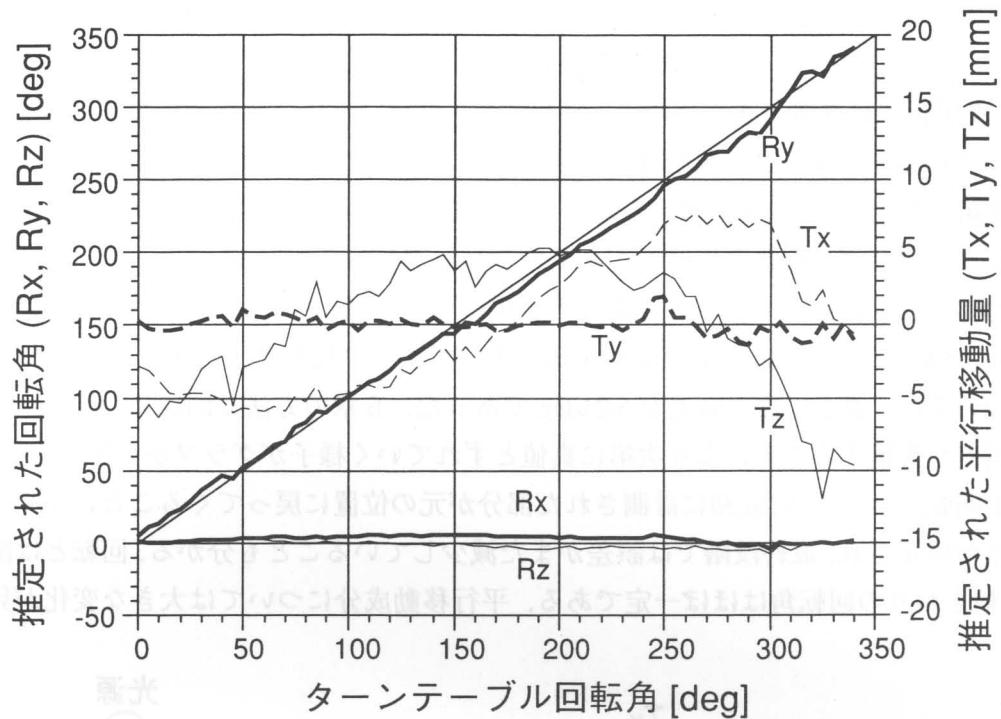


図6.13 センサから得られる濃淡画像



Ry の推定誤差の標準偏差 : 5.21 [deg]  
Ty の推定誤差の標準偏差 : 0.51 [mm]

図 6.14 回転成分の推定結果

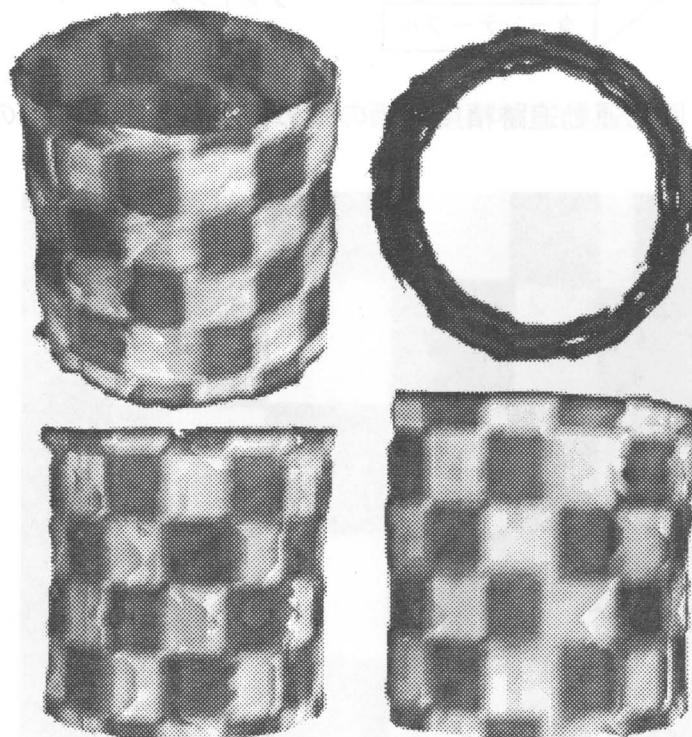


図 6.15 復元された全周形状と反射率

が、これは 5.5.3 と同様に自動的に設定されたモデル中心が回転軸から外れていたことによるもので、この影響を除外すると全体的に 2mm 程度の推定誤差であった。距離画像のみにより計測を行うと姿勢検出不可能なパラメータが存在するために、円筒を静止させていても計算機内モデルは回転しながら軸方向に平行移動してしまうが、この実験では Y 軸方向の平行移動量の標準偏差は 0.51mm とほぼ一定であった。

最後に、復元された対象物体の形状と反射率モデルを用いてシェーディング表示した結果を図 6.15 に示す。環境光が実際には一様でないことなどから誤差が発生し、それにより形状モデルの継ぎ目では反射率の相違が見られるが、全周に渡っておおよそ均一な反射率が求められたことが分かる。追跡に用いる濃淡画像の計測点が少ないために画像がぼやけているが、より微細なテクスチャが必要な場合にはセンサから得られる濃淡画像をそのまま蓄積しておき、後に解析することにより位置合わせ精度の向上やより正確で細かな反射率の計算が可能である。しかしこれらは今回の研究には含まれていない。

## 6.6 考察

この章における問題設定はたいへん一般的であり前提条件をほとんど必要としない。動物体追跡の側面では、対象物体に関する情報を与えることなく 6 軸運動パラメータを実時間的に計測することができるシステムと言える。また逆にモデリングシステムとしては、全周形状と各点の反射率を得るために対象物体の運動にはなんらの制約条件を設ける必要がない。光源モデルは CG として生成できる形式であればどのようなものでも用いることができ、直接的にポリゴン各頂点の反射率が求められる。これらは既存の追跡・モデリングシステムにはなかった特質であり、それゆえこのシステムは全く新しい応用を切り拓く可能性がある。この節では実験結果の評価に加え、このシステムによりどのような応用が可能となるのかについても考察する。

### 6.6.1 動物体の位置・姿勢追跡について

この章では対象物体のモデルを先に与えることはせず、モデルを逐次的に構築した場合の精度評価しか行っていないために、計測誤差がモデルに蓄積する現象が多く見られる。レンジファインダでは対象物体表面のうちカメラとプロジェクタの双方から観測可能な領域のみの像が得られるため、ベースライン長の設定にもよるが 6.5.2 の実験に用いた対象物体については周のうち 120deg 程度しか一度に得られない。濃淡画像が得られる領域はさらに広いが、この章で述べた手法では距離画像が得られた領域内の情報のみを用いている。この領域のみの情報により追跡が行われていることを考えると、濃淡画像による姿勢推定の安定度はかなり高いと思われる。円筒の軸に垂直な方向の平行移動と回転については濃淡画像と距離画像双方の情報が用いられるために安定度は非常に高く、その点で回転角に大きな推定誤差が生じたとしても形状そのものに歪みは見られないはずである。しかしこの実験では濃淡画像のみにより追跡が行われている軸まわりの回転と軸に沿った平行移動についても正しく追跡されていることが図 6.14 からわかる。また、構築されたモデルの模様が大きくずれていないことからも、正しく追跡が行われたことが分かる。

濃淡画像による移動量の算出は、テクスチャの性状により結果が大きく変化する。実験例で示した物体は素性のよいモデルであるが、市松模様の境界線をぼかすことによりさらに安定度や精度が増す。逆に細かすぎるテクスチャや全体にグラデーションのかかった物体の追跡は不安定である。また、入り組んでおり影が生じる対象では、CG ハードウェアが影付け計算をサポートしていないためにその部分の反射率推定が正しく行えない。勾配法は明度の変化が不連続な部分では成立しないので、画像をぼかすことが有効であるとされている。しかし画像をぼかすコストはたいへん高いので計測速度の問題から今回は行っていない。

### 6.6.2 形状・反射率モデルについて

形状モデルの精度についてはほぼ 5 章で述べたことと同じである。濃淡画像を用いてはいるが、形状が変化する方向の運動パラメータは距離画像の情報により追跡されるために、姿勢精度の誤差により表面が浮いたようになることは少ない。全ての軸について正確な運動パラメータが求められた場合には生じないが、6.5.2 で示した例のように濃淡画像のみにより復元されたパラメータが存在する場合にはそのパラメータ方向にテクスチャがずれ、統合した二つの距離画像同士の継ぎ目が不自然になることがある。互いに重なり合った面の反射率を平均することなどによってこの継ぎ目を目立たなくすることは可能であるが、これは後処理によって行うべきであり、今回提案した手法とは独立した研究テーマである。

後処理は主としてモデルの品質を高めるために有効であり、距離画像を統合することにより形状の精度を高めることも同様に可能である。これよって、オンラインに推定された対象物体の運動パラメータを修正することも可能である。これらで必要となる技術は、距離画像や濃淡画像を用いた位置合わせ、また複数のテクスチャマップされたパッチモデルの統合などが挙げられる。多数の研究がこの分野についてなされており、そのほとんどでおおよその初期値が与えられていることが前提となっている。本手法は、これらの研究が前提としているデータを与えるためのオンライン処理と位置づけることができ、この観点からも独自性が高い。

反射率の推定については、計測された濃淡画像をそのまま保存しておき、オフラインに解析することによって photometric stereo 法などにより法線ベクトルや光源方向の推定を行うことも可能である。3 章で述べたように、photometric stereo では対象物体とカメラの位置を固定し、光源の位置を変化させて像を取り込むことにより対応点問題を避けている。ここでは物体が回転するために、相対的には物体に対して光源とカメラの双方が運動していることになり、対応点の追跡問題が発生する。2 章・3 章の手法と同様に、対象物体の形状と運動の双方が分かれているので点の追跡が可能となり、それにより反射率を求めることが可能である。反射率解析の研究も多数行われているために、鏡面反射や相互反射の解析も可能であると思われるが、これはオンラインで行うべきでないだろう。

## 7 章 結論

本論文は、画像計測の一応用として運動する物体の全周形状と表面各点の反射率をモデル化する二つの手法についてまとめたものである。また、物体の運動が未知である場合には、その運動を求める必要があるため、6軸運動パラメータを実時間的に追跡する手法についても詳述した。特に、目的に合致した機器構成や処理形態を追求することでよりよい計測結果を得ることができることを示した。

計算機性能の向上に伴い Virtual Realityなどの分野で三次元的情報が広く扱われるようになつたにも関わらず、その入力手段たる三次元形状計測装置の普及は進んでいない。通常、物体形状を計測する場合にはスリット光源などを用いたアクティブな計測原理に基づく手法が用いられるが、この光源の特殊性が普及の遅れる一因であると考えられる。そこで投光装置を用いない受動的計測原理により形状計測を行う手法について研究を行つた。全周形状を得るためにには物体の向きを変える必要があるため、このために用いるターンテーブルで対象物体を運動させることにより連続的に変化する画像を得、これを解析することにより形状を得た。カメラ光軸を回転軸と垂直に設定することで時空間画像処理を導入することが可能となり、計測精度や安定性が向上した。また、シルエット解析による形状復元手法を融合することにより、様々な形状・表面特性の物体が計測可能となった。しかし依然アクティブ計測手法に較べると計測精度が劣り、形状計測結果から計算により求めた法線ベクトルは精度的に信頼できない。そこで円環状の光源を用い photometric stereo に類似した手法により反射率と法線ベクトルを求めた。平滑化などの後処理が必要となる場合も多いが、CG制作のための形状入力に利用可能な程度のモデルを得ることができた。この研究により、計測目的に適した装置構成を用いることにより様々な効果的手法が利用可能となり、受動的計測では極限的な精度を得ることができることが明らかとなった。

他方、対象物体の運動を求めるためには、アクティブ計測、特にレンジファインダが用いられるることは少ない。なぜなら一回の計測のために多数の画像を処理する必要があるために計測速度が低くなるからである。しかし、専用 LSI によりスリット光の像を撮影・処理することで、通常の CCD カメラと同等の計測速度を達成したレンジファインダが出現してきた。そこでこのようなセンサの一つであるシリコンレンジファインダを用いることにより動物体の運動を実時間的に計測し、かつ形状や反射率をモデリングする手法についての研究を行つた。まず 5 章では、力学モデルに基づいた単純な処理によって幾何学的な形状の物体の運動が計測可能であることを示した。濃淡画像も同時に用いるが、単に二次元的な二値画像処理をハードウェアにより行うことにより位置検出の補助とした。力学モデルに基づくため、関節を有する物体の運動を求めることが可能であることも示した。

4 章の方法では幾何学的な形状の物体しか計測することができなかつた。そこで、5 章ではより高度な追跡手法として様々な姿勢変化に関する距離画像の変化を計算機内でシミュレートし、この結果を複数統合することにより自由曲面物体の追跡を行う手法を示した。計算機内に用意した対象の形状モデルから距離画像を生成する処理に CG ハードウェアの Z バッファを用いることで高速化を図る技術について特に詳しく解説した。姿勢推定の代表的な手法である

ICPでは必要な速度を達成することは不可能であるため、7枚の生成距離画像を統合することにより効果的に6軸姿勢推定を行う新たな手法を開発した。レンジファインダから得られる距離画像は対象物体の形状に関する情報を直接的に保持しているため、これにより対象物体の形状モデルを逐次的に構築していく手法を導入し、追跡に先立ち対象の形状モデルを与える必要をなくした。これらにより、単に対象物体をセンサの前で様々に運動させるだけで対象物体の全周形状を得ることができるシステムを構築した。

次に、距離画像のみでは原理的に姿勢を検出し得ない形状の物体に対応するため、対象物体表面のテクスチャ情報を用いて姿勢変化を検出する手法を融合した。距離画像のときと同様に、反射率パラメータを持つ対象物体モデルから濃淡画像を生成する処理にCGハードウェアを用いることにより高速化を図った。これは距離画像を生成する場合に較べてより常識的であるが、計測系と生成系の座標系を完全に一致させるためにレンジファインダのシステムパラメータを用いている。生成された濃淡画像から姿勢を検出する処理は距離画像の場合と全く同じであるが、これは勾配法に基づくオプティカルフロー算出原理に類似した手法であることを示した。勾配法では他の拘束条件を併用する必要があるが、本論文で提案した手法では対象物体の剛体性を直接拘束条件とすることことができ、計測の安定性に寄与している。

像から得られる明度値を直接モデル化する手法では、光源の配置により生じる陰影に影響され正しく姿勢推定することができないため、シェーディング解析により表面反射率を求める必要がある。これを行う処理にもまたCGハードウェアを用いて高速化を図った。あらかじめ光源モデルを設定したCGハードウェアにより一定の反射率を設定したモデルを描画することで純粋な陰影効果のみを得、この結果と実画像を比較することにより反射率を推定する。これにより追跡処理の性能をほとんど低下させることなく形状と反射率からなる対象物体モデルを逐次的に構築していくことが可能となった。画像の計測処理と生成処理に対称性を持たせた処理形態であるため、ワークステーションを2台用いて並列化を図ることが容易であることを示した。また、対象物体が画面から逸脱してしまうことがないように、カメラ計測方位を対象物体方向へ制御することも行った。

これら全ての処理により、対象物体に関する情報を全く与えなくても、対象物体をセンサの前で自由に動かすことによりその運動が実時間的に求められ、また形状と反射率に関する三次元モデルが構築される。これまでリアルタイム計測にはほとんど用いられなかったレンジファインダを効果的に用いることにより、これほどまでに柔軟で利用範囲の広いシステムが構築可能であることを実証したことがこの研究の最大の成果である。また、対象物体モデルを陽に用いることにより安定な計測が可能であること、しかしモデルは必ずしも先に準備しておく必要がないことなど距離画像にモデルベース処理を用いることの利点を示した。これはちょうど二次元画像処理においてテンプレート照合を行うこととほぼ同様であるが、レンジファインダはそれと同程度のコストで三次元世界を扱うことのできるセンサであると言ふこともできる。

本論文の冒頭において、タスクを明確にした上で最適な計測機器を選択すること、またタスクと対象に関する知識を陽に用いてトップダウン的処理を行うことが有効であると述べたが、この研究ではそれらを具現し示すことができた。ビジョン・画像計測の新たな可能性をこの論文に感じていただければ、それは筆者の幸いである。

## 参考文献

- [Aloimonos87] J. Aloimonos, I. Weiss, and A. Bandyopadhyay : Active Vision, Proc. 1st Conf. on Computer Vision, pp.35-54 (1987)
- [Amano95] 天野敏之, 日浦慎作, 佐藤宏介, 芦ヶ原隆之, 横山敦, 井口征士 : シリコンレンジファインダのキャリブレーション, 第5回インテリジェントFAシンポジウム講演論文集, 103 (1995)
- [Amano96] 天野敏之, 日浦慎作, 山口証, 井口征士 : 距離画像の固有空間照合法に基づく物体の姿勢検出 - 画素欠損画像に対するロバスト姿勢検出手法 -, MIRU'96講演論文集1, 205-210 (1996)
- [Armstrong95] M. Armstrong, and A. Zisserman : Robust Object Tracking, Proc. of 2nd ACCV, pp.I-58 (1995)
- [Besl92] P. J. Besl and N. D. McKay : A Method for Registration of 3-D Shapes, IEEE Trans. of PAMI, Vol.14, No. 2, pp.239-256 (1992)
- [Bolles87] R. C. Bolles, H. H. Baker, and D. H. Marimont : Epipolar-Plane Image Analysis : An Approach to Determining Structure from Motion, International Journal of Computer Vision, Vol.1, No.1, pp.7-56 (1987)
- [Fujita90] 藤田武洋, 佐藤宏介, 井口征士 : 局所曲面形状解析に基づくビンピッキングのためのビジョンシステム, 電子情報通信学会論文誌, Vol.J73-D2, No.1, pp.46-53 (1990)
- [Hayakawa93] 早川秀樹 : 複数濃淡画像からの形状・面反射率・光源方向の推定, 信学技報 PRU92-129 (1993-01)
- [Hebert95] M. Hebert, K. Ikeuchi, and H. Delingette : A Spherical Representation for Recognition of Free-form Surfaces, IEEE Trans. of PAMI, Vol. 17, No.7 (1995)
- [Horn81] B. K. P. Horn and B. G. Schunck : Determining optical flow, Artificial Intelligence, Vol.17, pp.185-203 (1981)
- [Ikeuchi90] 江尻正員, 大田友一, 池内克史 : マシンビジョン, 昭晃堂 (1990)
- [Ikeuchi95] 池内克史 : 視覚による仮想現実感システムのためのモデル生成, 第一回画像センシングシンポジウム講演論文集, pp.219-222 (1995)
- [Kanade91] T. Kanade, A. Gruss, and L. R. Carley : A Very Fast VLSI Rangefinder, Proc. IEEE Int'l Conf. on Robotics and Automation, pp. 1322-1329 (1991)
- [Kang95] S. H. Kang, Y. D. Seo, K. S. Hong, and O. H. Kim : A High Speed Digital Range Imaging Sensor, Proc. of 2nd ACCV, pp.III-490-494 (1995)
- [Kuno95] 久野義徳 : アクティブラビジョンの展開 カメラが動くと何が良いのか ?, 画像ラボ12月, pp.35-38 (1995)
- [Mackworth73] A. K. Mackworth : Interpreting pictures of polyhedral scenes, Art. Intel. 4, 121-137 (1973)

- [Marr82] D. Marr : Vision - A Computational Investigation into the Human Representation and Processing of Visual INformation, W. H. Freeman & Co. (1982)
- [Masuda96] 増田健, 坂上勝彦, 横矢直和 : モデル生成のための複数の距離画像の位置合わせと統合, MIRU'96 講演論文集 II, pp.247-252 (1996)
- [Otsuki93] 大槻正樹, 三浦淳, 佐藤幸男 : 多方向測定による物体形状と表面反射特性の計測, 信学論 Vol.J76-D-II No.8 pp.1536-1543 (1993)
- [Quan96] L. Quan : Self-Calibration of an Affine Camera from Multiple Views, International Journal of Computer Vision, Vol.19, No.1, pp.93-105 (1996)
- [Roberts65] L. Roberts : Machine Perception of three-dimentional solids, Optical and Electro-Optical Information Processing (J. Tippett, ed.) MIT Press, pp.159-197 (1965)
- [Shimizu96] 清水郁子, 出口光一郎 : 位置関係が未知の複数視点からの距離画像の統合, MIRU'96 講演論文集 I, pp. I-199-204 (1996)
- [Shirai87] 白井良明 (編) : パターン理解, オーム社 (1987)
- [Simon94] D. A. Simon, M. Hebert and T. Kanade : Real-time 3-D Pose Estimation Using a High-Speed Range Sensor, Proc. IEEE Int'l Conf. on Robotics and Automation, pp.2235-2241 (1994)
- [Stark96] Katrin Stark, Siegfried Fuchs : A Method for Tracking the Pose of Known 3-D Objects Based on an Active Contour Model, Proc. of ICPR'96, pp.I-905-909 (1996)
- [Stoddart96] A. J. Stoddart, A. Hilton : Registration of Multiple Point Sets, Proc. of 13th ICPR, pp.B-40-44 (1996)
- [Szeliski91] R. Szeliski : Shape from Rotation, CVPR-91, pp.625-630 (1991)
- [Takano84] 高野英彦 : 形状パターンの認識技術, 情報調査会 (1984)
- [Tamura96] 田村秀行 : 知能情報 (さいばあ) メディア時代と PR / IP / CV 技術, MIRU'96 講演論文集 1, pp.7-12 (1996)
- [Tomasi92] C. Tomasi and T. Kanade : Shape and Motion from Image Streams under Orthography : a Factorization Method, International Journal of Computer Vision, Vol.9, No.2, pp.137-154 (1992)
- [Ullman79] S. Ullman : The interpretation of structure from motion, Proc. of R. Soc. Lond, B203, 405-426 (1979)
- [Vaillant92] R. Vaillant, O. D. Faugeras : Using Extremal Boundaries for 3-D Object Modeling, IEEE Trans. on PAMI, Vol.14, No.2, pp.157-173 (1992)
- [Walts75] D. Walts : Understanding line drawings of scenes with shadows, The Psychology of Computer vision, P. H. Winston, ed., pp. 19-91, McGrawHill (1975)
- [Wiberg76] T. Wiberg : Computation of principal components when data are missing, Proc. of 2nd Symp. Computational Statistics, pp. 229-236 (1976)
- [Woodham89] R. J. Woodham : Photometric method for determining surface orientation from multiple images, in B. K. P. Horn and M. J. Brooks (ed.) , Shape from Shading, pp.513-532, The MIT Press (1989)

- [Yachida90] 谷内田正彦：ロボットビジョン，昭晃堂（1990）
- [Yokoyama96] 横山敦，佐藤宏介，芦ヶ原隆之，井口征士：無調整型フォトVLSIセンサを用いた実時間距離画像計測 -シリコンレンジファインダ-, 信学論 Vol. J79-D-II, No.9, pp.1942-1500 (1996)
- [Zheng94] J. Y. Zheng : Acquiring 3-D Models from Sequences of Contours, IEEE Trans. on PAMI, Vol.15, No.2, pp.163-178 (1994)
- [Zheng96a] J. Y. Zheng, A. Murata, Y. Fukagawa and N. Abe : Reconstruction of 3D Models from Specular Motion Using Circular Lights, Proc. of ICPR'96, pp.A869-873 (1996)
- [Zheng96b] J. Y. Zheng, H. Kakinoki, K. Tanaka and N. Abe : Computing 3D Models of Rotating Objects from Moving Shading, Proc. of ICPR'96, pp.A800-804 (1996)

## 謝辞

本論文は、著者が大阪大学基礎工学部システム工学科井口研究室に初の飛び級の大学院生として所属した1993年4月から、通常より1年早く修了することになった現在までの4年間に行った研究をまとめたものです。これほどまでに早く学位取得にこぎつけることができたのは、ひとえに多数の方々の援助によるものです。ここにお世話になった方々へ心からの謝意を表します。

井口征士教授には研究の方向付けから学会発表の細部に至るまで、あらゆる面で多くの指導を賜りました。また、最新の設備をはじめとした快適な研究環境を用意していただいたことで、先端的な研究に取り組むことができました。数多く与えていただいた発表活動の機会で得た刺激や他機関の研究者との出会いは著者の大きな成長につながりました。心から感謝の意を表します。

谷内田正彦教授には、本論文をまとめるにあたり詳細なご検討を賜りました。また、本研究について先進的な示唆を賜りました。心からお礼申し上げます。西田正吾教授には、本論文に対し建設的なご意見・ご検討を賜りました。厚くお礼申し上げます。

加藤博一先生には、研究全般にわたって何度もたいへん熱心にご討議いただきました。また、労を惜しまず研究設備の整備などをしてくださいり、研究者としての考え方・取り組み方を身をもって教えてくださいました。心から感謝いたします。

奈良先端科学技術大学院大学の佐藤宏介助教授には、研究テーマの設定や研究の進め方など多くのご指導を賜りました。また、快くシリコンレンジファインダを利用させてくださいました。心からお礼申し上げます。和歌山大学の沼田卓久教授には、数々の有益なご討議を頂きました。深くお礼申し上げます。システム工学科井上雄二郎教授、藤井隆雄教授、笠井健教授、豊田順一教授、田村坦之教授、溝口理一郎教授にはたいへんお世話になりました。厚くお礼申し上げます。

井口研究室の真鍋佳嗣先生、山口証先生には研究生活全般にわたりたいへんお世話になりました。深く感謝いたします。また、天野敏之氏には共同研究者として多大な協力を頂きました。深く感謝いたします。また、三次元計測グループの三田雄志氏には、同じく共同研究者の青木利道氏、貝原輝昌氏とともに研究の遂行にご協力いただきました。ここに感謝の意を表します。また、設備の使い方を教えていただいた上級生の方々、研究室を楽しくしてくれた同級生の方々、そして快適な研究生活を支えてくれた下級生の方々など、井口研究室のみなさんには、たいへんお世話になりました。深く感謝いたします。

名古屋工業大学の佐藤幸男教授と佐藤研究室のみなさんには、さまざまなお討議を賜りました。厚くお礼申し上げます。また、谷内田研究室のみなさんには、弟共々たいへんお世話になりました。深くお礼申し上げます。

最後に、この機会を与え支えてくれた両親と、全てに協力してくれた弟に、心から感謝の意を表します。ここに書ききれませんが、私の身のまわり全ての方々に感謝いたします。

# 業績表

## [学術論文]

1. 日浦慎作,佐藤宏介,井口征士:対象物体の回転による形状と反射率の同時計測,情報処理学会論文誌, Vol.36, No.10, pp.2295-2302 (1995)
2. 日浦慎作,山口証,佐藤宏介,井口征士:動距離画像の計測と生成による任意形状物体の実時間追跡,電子情報通信学会論文誌 (条件付採録, 再投稿中)
3. 日浦慎作,山口証,佐藤宏介,井口征士:距離画像と濃淡画像の融合による任意形状物体の実時間追跡,電子情報通信学会論文誌 (投稿中)
4. 天野敏之,日浦慎作,山口証,井口征士:固有空間照合法に基づく距離画像からの物体の姿勢検出,電子情報通信学会論文誌 (採録決定)

## [国際会議]

1. Shinsaku Hiura, Kosuke Sato and Seiji Inokuchi : Tracking Jointed Object using Sequence of Range Images, Proc. of ACCV95 Vol.3, 499-503 (1995)
2. Shinsaku Hiura, Kosuke Sato and Seiji Inokuchi : Range Image Processing using Graphics Accelerator, Proc. of Japan-USA Symposium on Flexible Automation, 955-958 (1996)
3. Shinsaku Hiura, Akashi Yamaguchi, Kosuke Sato and Seiji Inokuchi : Real-Time Object Tracking by Rotating Range Sensor, Proc. of 13th ICPR, Vol.I 825-829 (1996)
4. Toshiyuki Amano, Shinsaku Hiura, Akashi Yamaguchi, Seiji Inokuchi : Eigenspace approach for a pose detection with range images -Robust pose detection method for pixel lacks of range images-, Proc. of 13th ICPR, Vol.I 622- 626 (1996)

## [口頭発表]

1. 日浦慎作,佐藤宏介,井口征士:対象物体の回転による三次元形状の計測,電気関係学会関西支部連合大会講演論文集, G114 (1993) (奨励賞受賞)
2. 日浦慎作,佐藤宏介,井口征士:対象物体の回転による形状と反射率の同時計測,電子情報通信学会春季大会講演論文集, 7-284 (1994)
3. 日浦慎作,佐藤宏介,井口征士:対象物体の回転による形状と反射率の同時計測,画像の認識・理解シンポジウム講演論文集 II, 279-286 (1994)
4. 日浦慎作,佐藤宏介,井口征士:動距離画像による多関節物体の実時間追跡,電子情報通信学会総合大会講演論文集, 7-411 (1995)
5. 天野敏之,日浦慎作,佐藤宏介,芦ヶ原隆之,横山敦,井口征士:シリコンレンジファインダのキャリブレーション,電子情報通信学会総合大会講演論文集, 7-413 (1995)
6. 日浦慎作,佐藤宏介,井口征士:リアルタイム距離画像センサによる動物体の多視点観測,マルチメディア・仮想環境基礎研究会報告書, 35-40 (1995)

7. 日浦慎作,佐藤宏介,井口征士:動距離画像による連結物体の実時間追跡,第1回画像センシングシンポジウム予稿集,315-318 (1995)
8. 天野敏之,日浦慎作,佐藤宏介,芦ヶ原隆之,横山敦,井口征士:シリコンレンジファインダのキャリブレーション,第5回インテリジェントFAシンポジウム講演論文集,103 (1995)
9. 天野敏之,日浦慎作,井口征士:距離画像の固有空間照合報に基づく物体の姿勢検出,研究会報告書,パターン認識と理解,145-150 (1995)
10. 日浦慎作,山口証,佐藤宏介,井口征士:距離画像センサの回転を用いた動物体追跡,電子情報通信学会総合大会講演論文集,7-345 (1996)
11. 三田雄志,日浦慎作,山口証,井口征士:リアクティブレンジファインダ,電子情報通信学会総合大会講演論文集,7-375 (1996)
12. 日浦慎作,山口証,佐藤宏介,井口征士:距離画像センサの回転を用いた動物体追跡,画像の認識・理解シンポジウム講演論文集II,253-257 (1996)
13. 天野敏之,日浦慎作,山口証,井口征士:距離画像の固有空間照合法に基づく物体の姿勢検出-画素欠損画像に対するロバスト姿勢検出手法-,画像の認識・理解シンポジウム講演論文集I,205-210 (1996)
14. 三田雄志,日浦慎作,山口証,井口征士:リアクティブ・レンジファインダ,画像の認識・理解シンポジウム講演論文集II,229-234 (1996)
15. 日浦慎作,三田雄志,加藤博一,井口征士:ビデオプロジェクタを用いた簡易距離画像計測装置,第39回自動制御連合講演会 前刷,423-424 (1996)
16. 天野敏之,日浦慎作,山口証,井口征士:固有空間照合法に基づく距離画像からの物体の姿勢検出,第39回自動制御連合講演会 前刷,421-422 (1996)
17. 三田雄志,日浦慎作,加藤博一,井口征士:リアクティブ・レンジファインダ,第39回自動制御連合講演会 前刷,419-420 (1996)