The University of Osaka
Institutional Knowledge Archive

| | |
|---|---|
| Title | 固有空間法による距離画像照合と認識 |
| Author(s) | 天野, 敏之 |
| Citation | 大阪大学, 2000, 博士論文 |
| Version Type | VoR |
| URL | https://doi.org/10.11501/3169478 |
| rights | |
| Note | |

*The University of Osaka Institutional Knowledge Archive : OUKA*

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

The University of Osaka

工甲 7550

# 固有空間法による距離画像照合と認識

天 野 敏 之

2000年1月

大阪大学大学院基礎工学研究科

# 固有空間法による距離画像照合と認識

天　野　敏　之

2000年1月

# 要　約

来世紀では、人々の夢であった自律移動ロボットが完成し、実社会の至る所で活躍する時代が来るといわれている。現在においても学術機関や企業において、さまざまな自律移動ロボットが試作され、その兆しが見え始めている。しかし、自律移動ロボットの手や足を実現するマニピュレータや二足歩行などの研究に対して、自律移動ロボットの目を実現するためのコンピュータビジョンの研究は解決すべき課題が多く、いまだ実用の域に達していない。本論文は、自律移動ロボットの視覚を実現するための物体認識手法として、距離画像を用いた固有空間法からのアプローチについて検討するものである。

本論文では、物体の様々な見かけを多数記憶し、画像を照合することにより物体を認識するアピアランスベースの物体認識において、濃淡画像のかわりに距離画像を用いる”距離画像の固有空間法”を提案した。この手法では、照明条件の変化に伴う画像の変化が生じないため、学習サンプル数を大幅に削減することができる。そのうえ、距離画像の固有空間法はアピアランスベースの方法論でありながら、形状モデルから学習を行うことが可能な”仮想学習”を提案した。しかし、実際のレンジファインダで撮影される距離画像撮影では、オクルージョンなどの理由により欠損が生じ、これを固有空間法でどのように扱えば良いかという問題が生じた。

次に、欠損を含む画像を固有空間照合する方法として、欠損画素により生じた固有空間上の投影誤差を逆変換することにより取り除く方法と、欠損パターンに応じて固有空間を再構築する手法を提案し、比較検討を行った。また、このような欠損を含む画像の固有空間照合法を応用して、固有ベクトルを知識としたデータの時空間補間法を提案した。この手法において、固有ベクトルが学習サンプル列に対する知識の一つであると言う考えを示した。

さらに、距離画像の固有空間法において、物体の奥行き方向の移動と姿勢変化による距離画像の変化を直交化して表現することが可能な”拡張固有空間”を提案した。これにより、従来のアピアランスベースの物体認識では実現することが困難な奥行き方向の位置検出も可能となった。このような奥行き方向の位置検出は自律移動ロボットの視覚機能で必要とされる機能であり、距離画像の固有空間法の優位性を示すことができた。

最後に、距離画像の固有空間法による物体認識・姿勢検出および奥行き検出手法をパンステージで回転するレンジファインダに応用し、平面上に配置された認識対象の探索を実現した。また、探索ではパンステージの回転とともに距離画像の固有空間法で認識された結果を仮説として扱い、仮説の投票と検証により実現した。この手法を用いることにより、環境中に複数の認識対象が混在するような環境においても、それぞれの物体ごとに物体認識・位置および姿勢検出が可能であることを示した。

# 目 次


第１章　序論 ..... 1
1.1　自律移動ロボットのためのコンピュータビジョン ..... 1
1.1.1　自律移動ロボットの必要性 ..... 1
1.1.2　自律移動ロボットの構成要素と現状 ..... 2
1.2　コンピュータビジョンにおける物体認識 ..... 3
1.2.1　モデルベースとアピアランスベース ..... 3
1.2.2　モデルベースの物体認識 ..... 4
1.2.3　アピアランスベースの物体認識 ..... 5
1.3　コンピュータビジョンのための画像と方法論 ..... 7
1.3.1　物体認識のための画像 ..... 7
1.3.2　距離画像を用いた固有空間法 ..... 8
1.4　本研究の意義と論文の構成 ..... 8

第２章　距離画像の固有空間法 ..... 11
2.1　濃淡画像と距離画像 ..... 11
2.1.1　見た目の姿と物理的形状 ..... 11
2.1.2　距離画像の計測原理と照明条件変化 ..... 12
2.2　距離画像による固有空間生成と認識手法 ..... 14
2.2.1　距離画像による固有空間生成 ..... 14
2.2.2　物体認識・姿勢検出手法 ..... 17
2.3　距離画像の固有空間法に対する考察 ..... 19
2.3.1　照明条件不変による利点欠点 ..... 19
2.3.2　物理的な形状表現による利点 ..... 20
2.3.3　欠損画素問題 ..... 20
2.3.4　距離値を表現していることによる利点欠点 ..... 21
2.3.5　その他の問題点 ..... 22
2.3.6　形状モデル照合と距離画像の固有空間法の比較 ..... 23
2.4　まとめ ..... 24

第３章　仮想学習による固有空間生成 ..... 25
3.1　固有空間法における学習サンプル獲得コスト ..... 25
3.1.1　自由度の変化に対する学習サンプル数の変化 ..... 25
3.1.2　固有空間上の補間による学習サンプル獲得コストの低減 ..... 25
3.1.3　視覚サーボにおける学習サンプル獲得コスト ..... 27
3.2　仮想学習 ..... 27
3.2.1　ポリゴンモデルによる距離画像の生成 ..... 27
3.2.2　レンダリングハードウェアによる距離画像の生成 ..... 29
3.2.3　ビジュアルベース仮想学習 ..... 30
3.3　仮想学習実験 ..... 30
3.3.1　認識対象と形状モデル ..... 30
3.3.2　仮想学習結果 ..... 32
3.3.3　物体認識・姿勢検出 ..... 35

3.4 仮想学習考察 ..... 36
3.4.1 学習サンプル獲得コストと距離画像生成時間 ..... 36
3.4.2 ポリゴンモデルによる距離画像の生成誤差 ..... 36
3.4.3 欠損画素問題 ..... 37
3.5 まとめ ..... 37

**第4章 欠損したデータの固有空間法と固有空間を知識とした補間法 ..... 39**
4.1 欠損画素の意味と固有空間照合での問題 ..... 39
4.1.1 欠損画素とは ..... 39
4.1.2 固有空間照合での問題点 ..... 41
4.2 欠損画素を含む画像の固有空間照合 ..... 43
4.2.1 欠損画素逆行列を用いた固有空間照合法（BPLP） ..... 43
4.2.2 固有空間の再構築による固有空間照合法（ARLP） ..... 45
4.3 欠損画素を含む画像の固有空間照合実験 ..... 46
4.3.1 ARLPとBPLPの比較シミュレーション ..... 46
4.3.2 BPLPによる逆射影 ..... 47
4.3.3 欠損画素を含む画像に対する物体認識・姿勢検出 ..... 48
4.4 BPLPの応用 ..... 48
4.4.1 BPLPを用いた欠損データの補間 ..... 49
4.4.2 濃淡画像の補間実験 ..... 51
4.4.3 BPLPを用いた欠損データの補間考察 ..... 52
4.4.4 他の補間法との比較 ..... 53
4-5 まとめ ..... 55

**第5章 仮想平行光学系と拡張固有空間による奥行移動量の検出 ..... 57**
5.1 仮想平行光学レンジファインダ ..... 57
5.1.1 物体の奥行き移動による見た目の大きさ変化 ..... 57
5.1.2 仮想平行光学RF ..... 60
5.2 拡張固有空間 ..... 61
5.2.1 奥行き検出に対する問題点 ..... 61
5.2.2 拡張固有空間による奥行き移動成分の直交化 ..... 62
5.2.3 拡張固有空間に対するBPLP ..... 65
5.2.4 物体認識・奥行き位置姿勢検出方法 ..... 66
5.3 奥行検出実験 ..... 67
5.3.1 実験環境 ..... 67
5.3.2 仮想平行光学系 ..... 68
5.3.3 仮想学習結果 ..... 68
5.3.4 実験結果 ..... 70
5.4 仮想平行光学RFおよび拡張固有空間に対する考察 ..... 72
5.4.1 仮想平行光学RFにおけるポリゴン近似誤差 ..... 72
5.4.2 拡張固有空間を用いた直交化の利点 ..... 73
5.5 まとめ ..... 73

第6章　回転レンジファインダによる物体探索 ..... 75
6.1　物体探索 ..... 75
6.1.1　自律移動ロボットで必要とされる視覚機能 ..... 75
6.1.2　ビンピッキングと ImageSpotting ..... 76
6.1.3　本章で扱う問題 ..... 77
6.2　回転レンジファインダ ..... 78
6.2.1　回転レンジファインダによる探索 ..... 78
6.2.2　装置構成 ..... 79
6.2.3　キャリブレーション ..... 81
6.3　探索アルゴリズム ..... 83
6.3.1　認識と信頼性 ..... 83
6.3.2　仮説の投票 ..... 83
6.3.3　仮説の検証 ..... 84
6.4　物体探索実験 ..... 85
6.4.1　認識対象が複数存在する場合 ..... 85
6.4.2　認識対象が一つだけ存在する場合 ..... 87
6.4.3　未知の物体が含まれる場合 ..... 88
6.4.4　背景が存在する場合 ..... 89
6.5　物体探索考察 ..... 90
6.5.1　未知物体を含む場合への拡張 ..... 90
6.5.2　背景問題について ..... 91
6.5.3　オクルージョンについて ..... 92
6.5.4　自律移動ロボットへの応用 ..... 92
6.6　まとめ ..... 93

第7章　結論 ..... 95

参考文献 ..... 97

謝　辞 ..... 99

# 第1章　序論

来世紀では、人類の夢であった自律移動ロボットが完成し、実社会の至る所で活躍する時代が来るといわれている。現在においても、学術機関や企業においてさまざまな自律移動ロボットが試作され、いくつかの簡単な自律移動ロボットが製品化されている。しかし、自律移動ロボットの手や足を実現するマニピュレータや二足歩行などのハードの研究は比較的実用の域に達しているのに対して、自律移動ロボットの目を実現するためのコンピュータビジョンの研究は解決すべき課題が多く、いまだ実用の域に達していない。コンピュータビジョンを実現するためには、画像から物体を認識する物体認識が必要不可欠であり、物体認識はコンピュータビジョン研究の最大の課題である。本論文では、自律移動ロボットの視覚を実現するための物体認識手法として、固有空間法による距離画像照合からのアプローチを提案する。まず第1章では、物体認識の方法論を説明し、その利点・欠点を述べる。また本論文の位置づけを明らかにするとともに、各章で提案する手法の概略について述べる。

## 1.1　自律移動ロボットのためのコンピュータビジョン

### 1.1.1　自律移動ロボットの必要性

ENIAC誕生から50年、弾道の軌道計算や科学技術計算などを目的として作られた計算機は、パソコンという一般家庭で使われる道具のひとつと考えられるまでに普及した。現在において計算機は、流体や構造力学、天体物理などの複雑なシュミレーションを瞬時に行なうことができる超並列スーパーコンピュータや、映画製作などのCG画像生成を行うグラフィカルワークステーション、また煩雑な事務作業を人の変わりに行なう業務用計算機や、一般家庭で文章作成や家計簿管理、またインターネットやゲームを楽しむパソコンなど、計算機の用途と種類は多種多様であり、我々現代人の生活において必要不可欠な道具である。このように計算機が世の中に普及した要因としては、近代社会のニーズによるところも多いが、機械の手を借りてより複雑で高度な仕事を行いたいという人類が昔から抱いていた夢によるところもある。このような夢の一つに、SF小説などで昔から描かれていた、”自分のために働いてくれる機械によるパートナー(ロボット)”を持ちたいというものもある。

20世紀もあとわずかとなった現在、学術や産業分野においてこのようなロボットの誕生の兆しが見え始めている。原子炉内作業や惑星探査のように、人が直接行くことが困難な場所で人に代わって作業を行なう極限作業ロボットや、工場の作業ラインで人に代わって効率的な様々な作業を行なうFAロボットなどは、1980年代後半から盛んに研究され実用化されている。1990年代に入り、特定の環境のみで運用するロボットに対して、ユーザーのニーズに基づいたロボット開発の必要性が唱えられており[Tatuno 93]、人間の作業方法をまねてタスクレベルの指令で基本的な作業を作業を行う自律移動ロボットや、自由度の高い多指ハンドの開発が行わ

れている。さらに、このようなFAで用いるような実用的なロボットではなく、実社会に導入することを目的としたロボット開発が行われている。特に、近年発表されたHONDAのP3は、二足歩行ロボットで長年実現することが困難であった動歩行を実現して話題となっている。またSONYが商品化したAIBOは、人間のために何かの作業をしてくれるのではなく、仕草や動きで人間を楽しませてくれる機械によるペット（ペットロボット）であり、今まで考えられてきた人間とロボットの関係とは異なる新しい考えを提案している。

このようなHONDAやSONYが開発したロボットは、そのまま何かの役に立つものではないかもしれないが、ENIACができた時代において現在のような計算機が日常生活に必要不可欠な道具となる時代を想像できなかったように、自律移動ロボットが来世紀にどのような役割を果たしているのかということは想像もつかない。しかし、近い将来自律移動ロボットは、日常生活に必要不可欠な人間のパートナーとなることであろう。

### 1.1.2 自律移動ロボットの構成要素と現状

人の手を借りずに、ロボット自身の判断により知的な作業を行なう自律移動ロボットを実現するためには、実環境の状況を取り込むセンサーとそれを判断する人工的な知能、そしてこれらに基づいて行動するためのアクチュエータが必要となる。これらの要素を組み合わせて自律移動ロボットを構成する方法はシリアル型やパラレル型、ハイブリッド型が考えられてきた。しかし、これらの方式はそのままでは柔軟性などの欠点があり、図1-1のようなシリアル型に準じた自律移動ロボットの構成が提案されている[AI 90]。

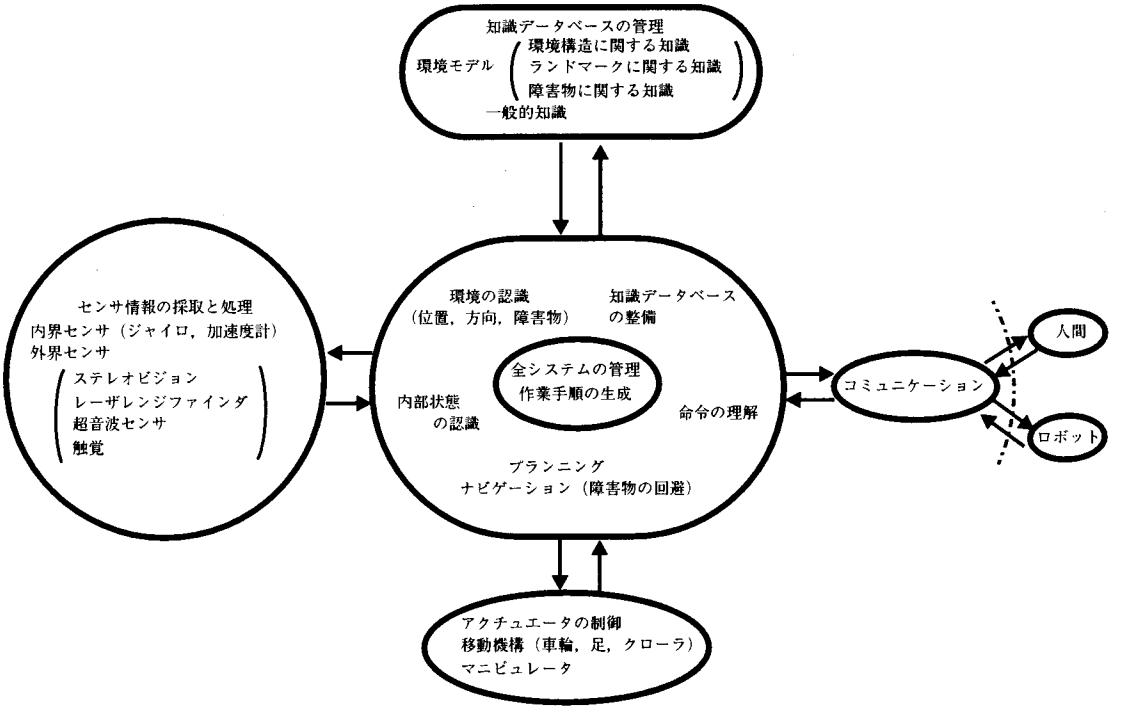



図1-1 自律移動ロボットのシステム構成例[AI 90]

図1-1で示したシステム構成の要素技術はどれも難しい課題であり、そのため様々な研究分野に細分化されている。外界から情報を得るための視覚(ロボットの目)を実現するためには、コンピュータビジョン(もしくはロボットビジョン)が必要であり、視覚から得た様々な情報から次の行動を決めるためには、人工知能が必要である。また、行動を起すためには何らかの移動機構やマニピュレータが必要であり、人間とロボットとのコミュニケーションを実現するためにはヒューマンインターフェースが必要である。この中で特にコンピュータビジョンは、自律移動ロボットが自律的に知的な作業を行う際に欠かせない重要な要素技術の一つであり、実現することが切望されている。なぜならば、人間の五感により外界から得られる情報のうち視覚情報は8割程度の情報量を占めるといわれており、もし自律移動ロボットに人間が行っている日常的な作業を要求するのであれば、視覚情報は最も重要な情報になり得るからである。また、コンピュータビジョンでは画像を計測して適切な処理を施す(ImageProcessing)だけで良いわけではなく、視覚によりとらえられた画像を理解(ImageUnderstanding)することにより初めて有用な情報となり、視覚機能を実現したといえる。また、撮影された画像を理解するためには、まずカメラで撮影された物体を認識(ObjectRecognition)することが必要になる。

## 1.2 コンピュータビジョンにおける物体認識

### 1.2.1 モデルベースとアピアランスベース

自律移動ロボットの視覚を実現するためには、まず計測された画像の中に何が写っている物体を認識することが必要になる。このような物体認識は学習により与えられた何らかの知識との照合により実現される。物体認識を照合する知識のレベルで大別すると

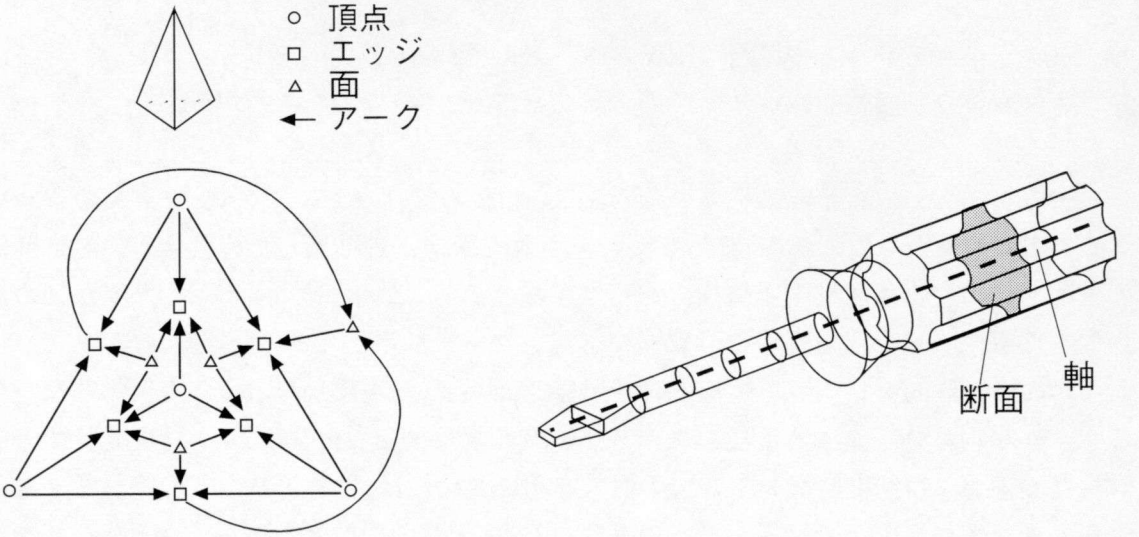
*1) 三次元特徴モデルでの照合(モデルベースの物体認識)*
*2) 画像および画像特徴レベルでの照合(アピアランスベースの物体認識)*

に分類できる。三次元特徴を用いた手法は、物体の幾何モデルやCADモデルのような幾何学的に表現された対象物体の特徴点と、計測された画像から抽出した特徴量との対応関係を求めることにより物体認識や姿勢検出を行う手法である。このような手法は、人間が初めて見た物体を理解する時の認識行程を計算機に応用した解析的手法である。本論文では、頂点や辺、面などの幾何学的な特徴量を積極的に活用し、計測された画像から求めた幾何学的な特徴量との照合から物体を認識する手法をモデルベースの物体認識と呼ぶ。また、三次元特徴モデルに対して画像および画像特徴レベルでの照合を用いた手法は、物体形状などの特徴量を抽出して幾何学的な情報を用いて解析を行うのではなく、計測した画像とあらかじめ記憶しておいた画像データベースとの単純な比較を行い、どの画像と最も似ているかということにより認識を行う手法である。画像レベルでの照合とは、画像をそのまま知識として持ち、テンプレートマッチングに代表される単純な相関演算を行う手法である。また画像特徴レベルでの照合とは、あらかじめ学習した画像をそのままテンプレートとして用いるのではなく、それらの学習画像よ

り固有空間やスペクトル空間、またはエッジ情報など画像を識別する上で有用な特徴量をあらかじめ求めておき、与えられた画像をこれらの特徴量により照合する方法である。画像特徴レベルでの照合は、画像レベルでの照合をより効率的に行う手法と考えることができ、本質的には同じ方法論である。そのため、本論文では画像レベルの照合は画像特徴レベルの照合に含むものとして、これらをアピアランスベースの物体認識と呼ぶ。また、形状モデルを用いた物体認識手法であっても、単に形状モデルとのマッチングより物体を認識するなら、知識として距離画像列であるか距離画像列を組み合わせたものを用いるのかの違いであり、方法論としては画像レベルでの照合と考えることができる。そのため、形状モデルを用いた物体認識でも、画像照合を用いた認識手法であればアピアランスベースの物体認識に含まれるものとする。

### 1.2.2　モデルベースの物体認識

モデルベースの物体認識のプロセスは、ボトムアップ解析による解析的な認識手法が例にあげられる。ボトムアップ解析とはdata-drivenな解析とも言われる。ボトムアップ解析では、まず与えられた入力画像に対し、第一段階として線や領域といった画像特徴を抽出する。しかし、このようなプリミティブな線や領域といった特徴量は照合において記述レベルが低く、扱いにくい。そこで、第二段階としてこれらの特徴量を組み合わせて構造化（グループ化）する。さらに第三段階では、構造化された線や領域といった二次元特徴量をエッジや面などの三次元情報へ変換する。そして、あらかじめ知識として与えられたモデルと照合を行い、物体を認識したり姿勢を求める。しかし、画像特徴の抽出が完全に行えなかったり、特徴との照合のコスト




(a) 幾何モデル　　(b) 一般円筒モデル

図1-2　認識対象のモデル記述

よりボトムアップ解析だけでは十分とは言えない。そこで、ボトムアップ解析によりある程度認識対象の具体的なモデルが予測できたあとに、モデルの記述に従い解析を行う手法が必要となる。この解析をトップダウン解析という。このような認識手法で用いられるモデルの代表例として、幾何モデル（図 1-2(a)）や一般化円筒モデル（図 1-2(b)）があげられる。

*◆幾何モデル*

物体形状を表現する上で、頂点、エッジ、面、物体をノードとし、それらの間の関係をアークで表現するデータ構造を用いる。このような幾何学的な特徴を用いたモデルは、対象物体のCAD データを直接変換することにより作成することもできるが、形状が複雑になると膨大な記述になる。また、認識において特徴点の組み合わせ問題が生じる可能性がある。

*◆一般化円筒による表現*

一般化円筒とは、中心軸を滑らかな自由曲線として、中心軸に沿って変化する断面形状で表現したモデルである。初期の研究では断面形状を円で近似していたが、その後断面形状として楕円や多角形も使われるようになった[Agin 76]。Nevatiaらはこの考えをさらに発展させて、ひとつの一般化円筒のみでは記述できないような複雑な物体を、いくつかの一般化円筒の組み合わせとして表現した[Nevatia 77]。

このような三次元特徴モデルでの照合は。

*1) 幾何学的に複雑な認識対象のモデル記述は困難*
*2) 画像から特徴量を安定に抽出することが困難*
*3) モデルが複雑になると計算コストが莫大になる*

という欠点がある。しかし、認識対象の見た目の大きさ変化や位置・姿勢変化に対して柔軟に認識を行うことができる。また、認識対象の見え方（アスペクト）の変化に対して容易に対処できるという利点がある。

### 1.2.3 アピアランスベースの物体認識

アピアランスベースの物体認識は、認識対象となる物体の画像を全て学習する必要がある。濃淡画像を用いた場合、照明条件による画像の変化や物体の姿勢による画像の変化など考慮すべき点が多く、学習サンプル数は積算的に増加する。このことより、記憶容量、照合時間の点で困難とされていた。しかし、SirovichやTurkらにより固有空間を用いた画像照合法が提案され、少ない記憶容量で学習画像を記憶し、固有空間上で照合することで高速に照合を行うことが可能となった[Sirovich 87][Turk 91]。近年この手法を応用して、物体の認識・姿勢検出を行うパラメトリック固有空間法[Murase 94a]がNTTの村瀬らにより提案され、複雑背景中の物体の位置検出、動画像認識、視覚サーボ、照明計画、shape from shading、距離画像への適用や局所領域への適用と様々なところで応用されている[Matuyama 98]。

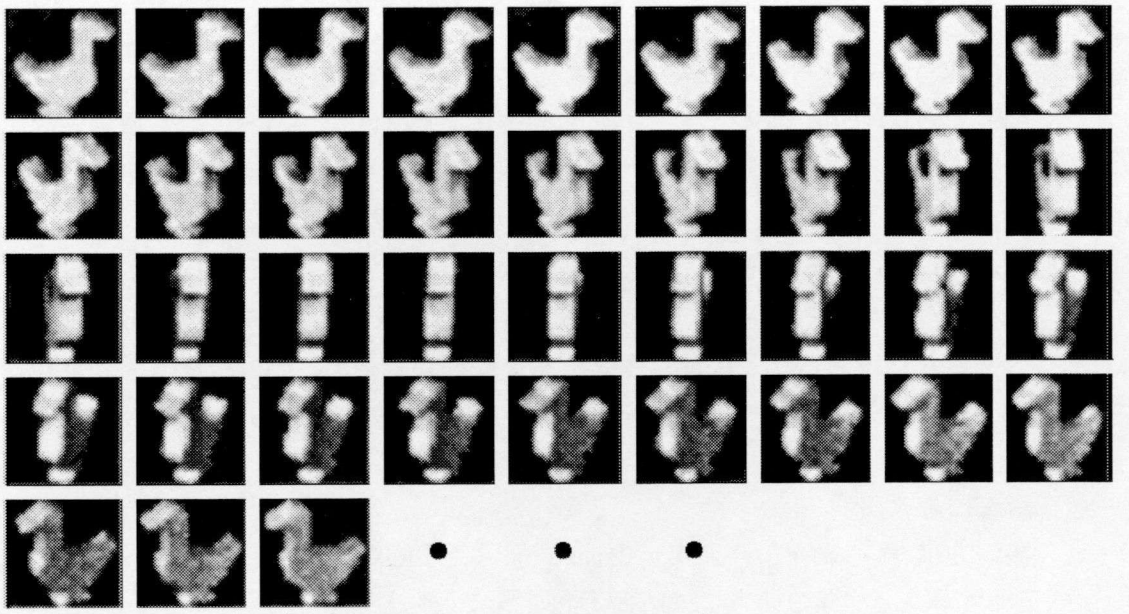


図 1-3　見かけ画像の変化

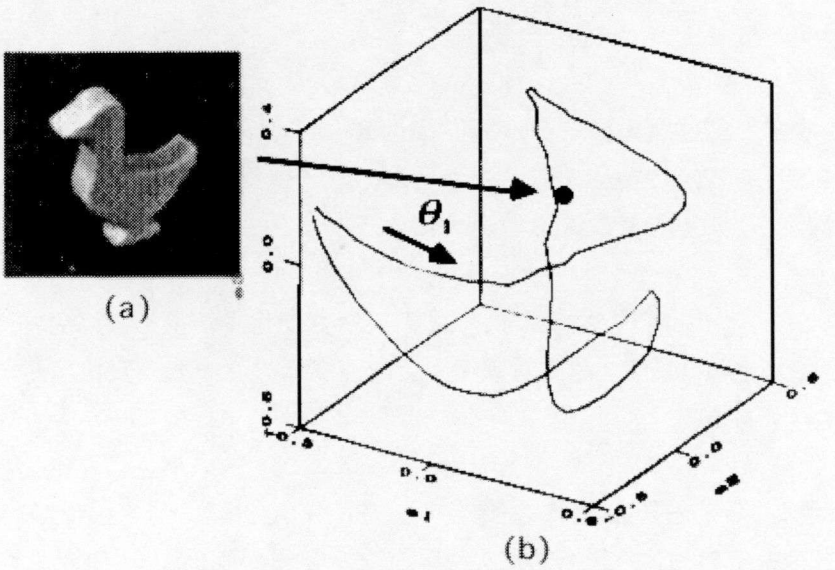



図 1-4　パラメトリック固有空間法

◆*パラメトリック固有空間法*

固有空間への学習サンプルの投影において、学習サンプルは相関の強い学習サンプルの近くに投影される性質がある。このため、物体の運動に応じて連続的に画像を撮影して固有空間を生成すると、学習サンプルの投影点により固有空間上に軌跡ができる。パラメトリック固有空間法とは、図 1-3 のように連続的に変化する画像列を学習サンプルとして、図 1-4 のように認識対象の姿勢変化や照明条件の変化をパラメータとする固有空間上の多様体で学習サンプルを

表 1-1　モデルベースと画像照合ベースの物体認識手法の利点と欠点

| | | |
|---|---|---|
| モデルベース | 利点 | 位置ずれや見た目の大きさ変化に対して柔軟に対応。<br>アスペクト変化などに容易に対応できる。 |
| | 欠点 | 幾何学的に単純な物体に限られる。<br>特徴量を安定して抽出することは困難。 |
| アピアランスベース | 利点 | 幾何学的に複雑な形状の認識も容易に行うことができる。<br>ヒューリスティックな特徴量検出は不要である。 |
| | 欠点 | 多量の学習サンプルが必要。<br>位置ずれや、見た目の大きさ変化に弱い。 |

表現する方法である。このような学習結果より物体認識・姿勢検出を行うには、固有ベクトルを基底として固有空間上に計測された画像を投影した後、最も近い学習サンプルの投影点を求めればよい。

このようなアピアランスベースの物体認識では、三次元特徴モデルの照合とは異なり、特徴モデルを用意する必要がなく、幾何学的に複雑な形状の認識や、顔画像認識なども容易に行うことができる。しかし、画像照合を基本とするため、以下の点で問題がある。

*1)　認識対象の運動や照明の変化などの自由度に応じて多数の学習サンプルを用意する必要がある。*
*2)　微妙な位置ずれや回転、見た目の大きさ変化などにより誤認識を生じやすい。*

以上より、モデルベースとアピアランスベースの物体認識手法の利点と欠点をまとめると表1-1のように表すことができる。

## 1.3　コンピュータビジョンのための画像と方法論

### 1.3.1　物体認識のための画像

人間の視覚は左右の両眼による立体視を行なっており、認識対象の色や陰影と同時におおよその距離計測も行なっている。このような人間の視覚を模倣して、自立移動ロボットの視覚システムでは2台のCCDカメラを用いたステレオ画像計測を用いたものが多い。人間のように3次元物体を認識する視覚を実現するために、まず人間と同じ計測方式により画像を撮影するという考えは自然な発想と言えるが、このような考えは必ずしも賢明ではない。なぜならば、人間が自立移動ロボットを必要とするのは、自立移動ロボットなら人間では行うことができないような複雑な作業を、正確かつ効率的に行うことができると期待しているからであるからである。また、人間の視覚では見ることができないような遠赤外線や、照明がほとんど照射されていない暗やみでも見ることができる視覚機能をコンピュータビジョンで実現できれば、人間では生物的な限界により実現することができないようなことも可能となるからである。このようなことは現在用いられている計算機からも言えることである。もしも、計算機が人間のように柔軟で適応性の高い計算ができたとしても、その計算結果が人間のように計算間違えが多

かったら使い物にならなかったことであろう。そのため、コンピュータビジョンで用いる画像としては濃淡画像にこだわる必要はなく、距離画像や赤外線画像、サーモグラフやスペクトル画像などを積極的に用いるべきである。また現在では、距離画像計測のためのレンジファインダ、サーモグラフやスペクトル画像を計測するための特殊なカメラは、CCDカメラに比べて高価な装置であり実用化の点で問題を指摘される場合が多い。しかし、これらの装置の価格は生産技術や需要によるところが大きく、その必要性が高まれば価格の問題は解決するものと思われる。

### 1.3.2 距離画像を用いた固有空間法

1.2.1で述べたように、画像計測より物体を認識する方法論としてはモデルベースとアピアランスベースの物体認識に大別することができる。これらの手法はそれぞれ長所短所があり、どちらの手法が優れていると一概に言うことはできない。しかしどちらの方法論においても、CCDカメラを用いた認識ではCCDカメラで撮影される2次元的な画像から実世界の3次元構造を理解するという不良設定問題を抱えていることが多い。そのため、レンジファインダで計測される距離情報を積極的に活用したアプローチも研究されている。このような距離画像を用いた認識は、Horandらが提案した3DPO's[Horand 84]のように”距離画像＝形状”という考えに基づいて三次元特徴モデルでの照合を行う手法が多い。しかし、モデルベースの物体認識では幾何モデルなどのモデル表現ができる必要性がある上、認識において必要な特徴量を検出する必要がありこの点で問題が生じる。そこで本研究では、画像として距離画像を用い、方法論として固有空間法を用いたアピアランスベースの物体認識を扱うものとする。濃淡画像は、認識対象の物理的な形状と同時に光学的な特性を合わせ持つ画像であり、認識を行う際に光学的な特性をどのように除去するかということが問題となるが、距離画像を用いることで物理的な形状のみを用いた認識が可能となり、認識の簡便さや信頼性の点で有利になる。また、方法論として固有空間法を用いることにより三次元特徴レベルの照合で問題となる特徴量検出やモデル記述の問題を回避することができる。

## 1.4 本研究の意義と論文の構成

本論文では、コンピュータビジョンの研究において最大の課題である物体認識を実現するための手法として、画像として距離画像を、方法論として固有空間法を用いたアプローチ（距離画像の固有空間法）について述べる。

第2章では距離画像の固有空間法の手法について説明し、従来一般的に用いられてきた濃淡画像を用いた固有空間法との対比を行い、距離画像の固有空間法の利点欠点を述べる。距離画像の固有空間法は、距離画像を用いることにより濃淡画像で本質的に含まれる光学的な特徴を切り離すことが可能になる。また距離画像の固有空間法では、アピアランスベースの物体認識で問題となる学習サンプル獲得コストを削減することもできる。アピアランスベースの物体認識では物体の姿勢変化などで生じる画像の見た目の変化を全て学習する必要があり、多自由度

の位置･姿勢検出を行うことを考えると学習サンプル数が指数関数的に増加するという問題がある。この学習サンプル数の増加はアピアランスベースの物体認識において、本質的な問題である。しかし、距離画像の固有空間法はアピアランスベースの認識手法でありながら、形状モデルによる学習を行うこと（仮想学習）ができる。この仮想学習の詳細については第3章で述べる。

以上のような利点がある反面、距離画像は欠損画素が生じるという特徴があり、この欠損画素を固有空間照合でどのように扱えばよいかということが問題となる。第4章では、距離画像の固有空間法における欠損画素問題について述べ、欠損画素問題を解決するための二つの手法を提案する。また、これら二つの手法の比較を行う。固有空間は直交部分空間であるため、投影の逆の操作である一次結合を行うことにより画像を復元することができる。この性質と欠損画素問題の解決手法を組みあわせることにより、固有空間を知識とした補間を行うこともできる。第4章では、この補間法についても述べる。

また、距離画像の固有空間法では距離画像を用いていることにより、認識対象の表面の凹凸に対して敏感であるため、精度の高い姿勢検出や物体認識を行うことができる。しかし、この利点は逆に認識対象の微妙な奥行き方向の位置変化には弱いということである。第5章では、仮想平行光学レンジファインダと拡張固有空間を提案し、認識対象の奥行き変化に対する問題を解決するとともにアピアランスベースの物体認識で問題となる計測された画像上での認識対象の見た目の大きさ変化の問題を解決する。また、これらの問題解決と同時に認識対象の奥行き方向の移動量検出も実現する。

第6章では、第2章から第5章で提案してきたさまざまな手法をパンステージで回転するレンジファインダに応用し、環境中に配置された物体を探索する手法を述べる。

最後に第7章では、本論文で達成した成果をまとめ、コンピュータビジョンに対する本論文の有用性について吟味するとともに今後の展望を述べる。

# 第２章　距離画像の固有空間法

顔画像認識で用いられるアイゲンフェースや、画像の連続的な変化を固有空間上の軌跡で表現したパラメトリック固有空間法などで知られる固有空間法は、2次元の画像より3次元物体の認識や姿勢検出などを行う手法であり、画像処理やパターン認識の様々な分野で応用されている。固有空間法とは、画像集合の共分散行列から得られる固有ベクトル列より部分空間を形成し、その部分空間上で画像を照合・比較する手法である。しかし、従来の固有空間法では濃淡画像を用いているため、安定した認識を行うためには照明条件の変化を考慮する必要があった。そこで、この章では照明条件の変化に影響されない距離画像を用いた固有空間法を提案する。また、濃淡画像と距離画像の性質の違いより、距離画像を用いた固有空間法の利点と欠点について述べる。

## 2.1　濃淡画像と距離画像

### 2.1.1　見た目の姿と物理的形状

濃淡画像は物体の見た目の姿（Appearance）を示すものである。ここで言う見た目の姿とは、物体が置かれている環境の照明光が物体表面の凹凸により遮られることにより生じる陰影のことであり、照明条件の変化により撮影される画像は変化する(図2-1)。そのうえ環境中に移動物体がある場合を考えると、照明の一部が遮られたり二次反射光が複雑に生じるため、撮影される画像はさまざまに変化する。そのため、濃淡画像による固有空間法では照明条件の変化を考

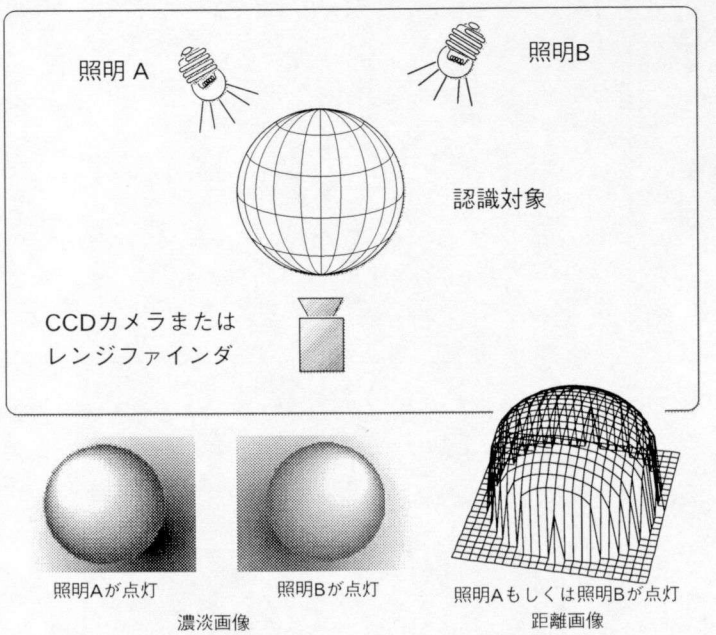



図2-1　照明条件の変化による濃淡画像の変化

慮して学習を行う必要がある。村瀬らは、工場内などの限られた環境を想定して照明の変化を単光源の一軸移動と仮定したが、環境内に遮光する要因となる人間、搬送車、ロボットアーム等が複数ある場合など様々な場合を想定すると、単光源の一軸移動では不十分であると思われる。このとき、あらゆる照明条件の変化を想定し、このすべてについて学習することでこの問題は解決できるが、すべての照明条件を想定することは困難である上、学習サンプル数が指数的に増える。これに対し距離画像は物体の幾何学的な形状を示す画像であり、照明条件の変化を考慮する必要はない。

### 2.1.2　距離画像の計測原理と照明条件変化

CCDカメラなどで撮影される濃淡画像とは、実環境に存在する照明により物体表面に生じた陰影を画像として表現したものである。そのため照明条件が変化すると、濃淡画像は変化する。これに対して距離画像とは、実環境に存在する物体とカメラまでの距離を格子状に配置した画像である。そのため距離画像は照明条件の変化に対して不変な画像である。濃淡画像と距離画像で計測している物理量は、濃淡画像は幾何学的な特徴の上に光学的な特性を畳み込んだ物であり、距離画像は幾何学的な特徴量のみを計測していると考えることができる。画像計測より物体認識を行うためには、認識対象の幾何学的な形状を計測することが重要であるが、濃淡画像を用いた認識では、どのように光学的な特性を取り除くかということが問題となる。これに対して距離画像は、光学的特性を持たないため認識に有効な幾何学的形状を画像から直接

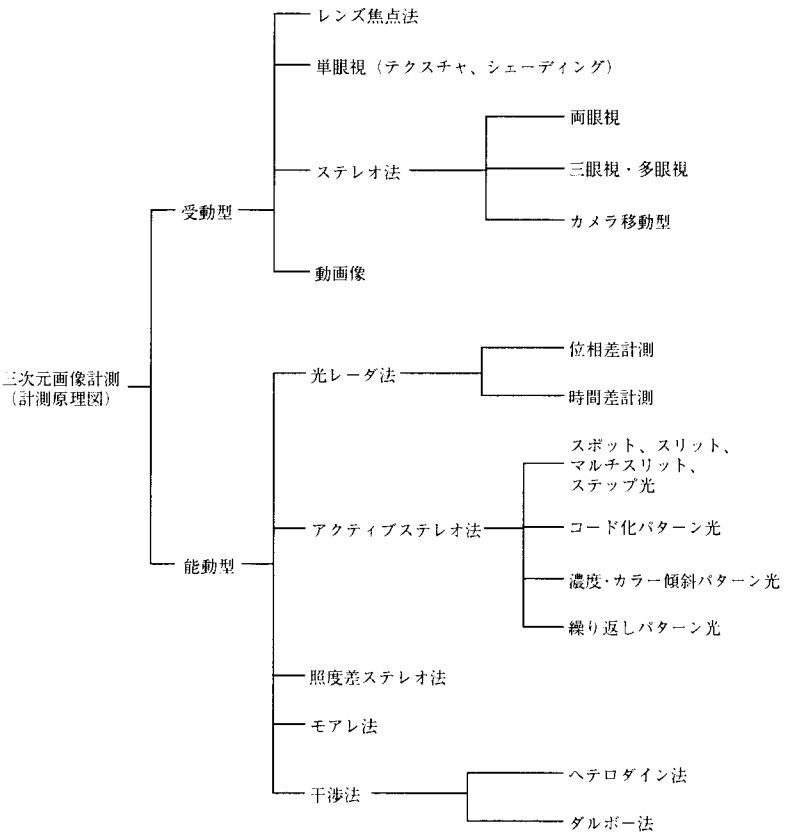



図2-2　三次元画像計測 [Inokuchi 90]

得ることができる。しかし、距離画像の計測原理によっては濃淡画像の撮影を行っているものもあり、最終的に生成される距離画像で照明条件の変化により計測誤差を生じる事も考えられる。以下では、距離画像の計測原理を説明するとともに、それぞれの計測原理での照明条件変化や認識対象の種類に対する安定性について考える。

距離画像を計測する装置（レンジファインダ）を計測原理により大別すると、受動型計測と能動型計測に分けることができる(図2-2)。受動型計測のうち代表的なものとしては、ステレオ法、レンズ焦点法などがあげられる。

### ◆ステレオ法

ステレオ法は、人間の両眼のように左右にカメラを配置し、左カメラと右カメラで撮影される二枚の画像よりカメラ間での視差を求め、視差より距離画像を計測する方式である。このとき、照明条件が変化すると左右のカメラで撮影される画像も変化するが、両方のカメラで同じように変化するため照明条件の変化はキャンセルされる。このような計測方式では、左右の画像の対応を求めることが必要であるため、認識対象表面にテクスチャ成分が存在しないようなものでは距離画像を生成できない。

### ◆レンズ焦点法

レンズ焦点法はカメラのフォーカスのボケから距離を計測する方式である。十分に浅い焦点深度をもつレンズを用いたカメラでは、ピントのあっている部分とあっていない部分は撮影される画像のボケ量から判定することができる。また、ピントを合わせるために調節したフォーカス位置は対象物体までの距離を表しているため、ピントのあっているところを検出したりボケ量を検出することにより距離を得ることができる。この計測法において、画像のボケが検出できなくなるような照明条件としては、CCD などの受光素子が飽和するような強力な光を当てた場合などが考えられる。しかし、CCD カメラは通常 Auto-Gain-Control（AGC）やシャッタスピードなどで光量を自動的に調整するため、照明条件の変化に対して不変な距離画像を得ることができる。

能動型計測は、光や超音波又は電磁波などを計測対象に照射して、その様子を撮影することにより距離画像を得る計測方法である。能動型計測のうち代表的な計測方式としては、光レーダ法、アクティブステレオ法があげられる。

### ◆光レーダ法

光レーダ法は地形測量などで一般的によく使われる計測方法で、計測器から対象物体に照射された光が反射して計測器に帰ってくるまでの時間により距離画像を得る方式である。照射する光としては、環境光に影響されにくい赤外の変調光を用いるのが一般的である。また AGC を用いて微弱な反射光も安定して検出することができるため、反射光を検出することができれば照明条件の変化に影響することなく距離画像を計測できる。

◆*アクティブステレオ法*

アクティブステレオ法は、スリットやパターン光などをプロジェクタにより投影した様子をカメラで計測し、三角測量の原理により距離画像を得る方式である。このアクティブステレオ法ではCCDカメラなどを用いて構成されることが多い。撮影される画像は照明条件の変化に影響を受けるため、強度の環境光が照射される場合CCDのチャージが飽和し、照射したパターン光が検出できなくなることも考えられるが、CCDカメラ自体にAGCを備えていることや、距離画像を計算する処理過程で差分二値化などを用いて環境光を除去する処理を行なうことが多く、照明条件の変化に影響されない。

以上のように特殊な環境での距離画像計測では、認識対象や計測方式によっては距離画像を計測できなることもあり得る。しかし図2-2で示した多種の計測原理は、様々な環境で距離画像計測を実現するために開発された手法である。そのため、環境に応じた計測原理を用いることにより、環境中に存在する認識対象の幾何学的な特性のみを計測することが可能であり、安定した物体認識を容易に実現することができる。

## 2.2 距離画像による固有空間生成と認識手法

### 2.2.1 距離画像による固有空間生成

ここでは学習サンプルを距離画像として、物体認識と回転による1自由度の姿勢検出のための固有空間を生成する手法を説明する。固有空間法は認識対象の見た目の変化に対して多様に変化する画像を全て記憶・照合する認識手法であるため、学習において様々な見え方の画像を計測する必要がある。そのため学習では、図2-3のようにそれぞれの認識対象をターンテーブルで回転させ、ステップごとに距離画像を撮影して学習サンプルを得る。距離画像では各々の画素に対し物体座標系上の座標値(X,Y,Z)が得られるが、計測される三次元座標が各々の受光素子の視線上の点であることを考慮するとこの表現は冗長である。そこで、それぞれの受光素子に対して奥行き方向の値をラスタスキャンし、式(2-1)のようにベクトルとして表現した(図2-4)。このようにベクトルで表現された画像を画像ベクトルという。また、$R$は回転の刻み数、$S$は認識対象の数、画像ベクトルの次元$N$は画像の大きさ$sizeX \times sizeY$である。

$$\begin{aligned} &\boldsymbol{x}'_{r,s}=\left[x_1, x_2, \cdots, x_N\right]^T_{r,s}, \quad N= sizeX \times sizeY \\ &where \quad r=1,2,\cdots,R \quad s=1,2,\cdots,S \end{aligned} \tag{2-1}$$

濃淡画像に固有空間法を適用して画像照合を行う場合は、得られる濃淡画像の画素値はセンサの感度や光源の光量に依存するため、濃淡画像の濃淡値の絶対的な大きさよりも濃淡値の分布に意味がある。そこで濃淡画像では画像ベクトルの長さを正規化し、センサの感度や光量による影響を除くことにより固有空間照合の安定性を向上させている[Murase 94a]。しかし、距

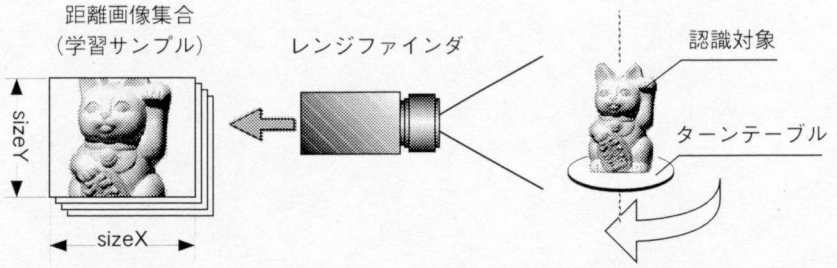



図2-3　学習サンプルの獲得

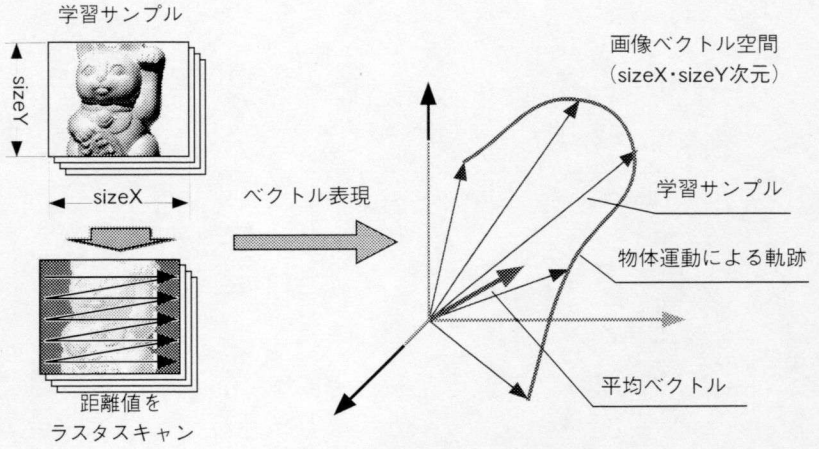



図2-4　画像ベクトルの生成

離画像では各画素の値は物体形状を直接表しているため、ベクトルの正規化を行う必要はない。また、濃淡画像を用いた固有空間照合では、それぞれの画像ベクトルから全画像の平均ベクトルを引く処理を行う。この平均ベクトルを引く操作を分散行列の数式的な意味から考えると、各々の学習サンプルの投影点の分散を大きくする意味があり、より正確な画像識別が行なえる。この操作はN次元の画像ベクトルを表現する空間において、すべての学習サンプルの投影点を平均ベクトルにより並行移動することを意味している。この操作では学習サンプルの投影点の相対的な位置関係は不変であり、距離画像を用いた固有空間照合でも矛盾を生じることはない。そこで距離画像における固有空間照合では、濃淡画像で行うときに用いる画像ベクトルの長さの正規化処理を行わず、平均ベクトルを引く処理を用いて分散行列を生成する。式(2-2)に平均ベクトル、式(2-3)に平均ベクトルを引いて表現した画像ベクトル列$\boldsymbol{X}$を示す。

$$\boldsymbol{c}=\frac{1}{R}\frac{1}{S}\sum_{r=1}^{R}\sum_{s=1}^{S}\boldsymbol{x}'_{r,s} \tag{2-2}$$

$$\begin{aligned}\boldsymbol{x}_{r,s}&=\boldsymbol{x}'_{r,s}-\boldsymbol{c}\\ \boldsymbol{X}&=\left[\boldsymbol{x}_{1,1},\boldsymbol{x}_{2,1},\cdots,\boldsymbol{x}_{R,S}\right]\end{aligned} \tag{2-3}$$

この画像ベクトル列$\boldsymbol{X}$の分散行列を求め、対角化により固有ベクトルと学習サンプルの投影点を求める。固有空間における学習サンプルの投影点は、

$$\boldsymbol{p}_{r,s}=\left[p_1, p_2, \cdots, p_D\right]^t_{r,s}=\boldsymbol{E}^{\mathrm{T}}\boldsymbol{x}_{r,s} \tag{2-4}$$

と示される。ただし、$\boldsymbol{E}$は生成された固有ベクトル列

$$\boldsymbol{E}=\left[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_D\right],\ D<R\times S<N \tag{2-5}$$

である。このとき、$D$は学習により生成する固有空間の次元数（固有空間の生成次元数）で、投影点と固有ベクトルの線形結合より学習サンプルを復元するときの精度に影響する。そのため、学習サンプルを復元する際の近似度の尺度である累積寄与率

$$Contribution\ Ratio=\frac{\sum_{i=1}^{D}\lambda_i}{\sum_{i=1}^{RS}\lambda_i} \tag{2-6}$$

より固有空間の生成次元数を決定する。ただし式(2-6)において$\lambda$は画像ベクトル列$\boldsymbol{X}$により生成された固有ベクトルの固有値である。

複数物体の認識と対象物体の一軸回転による学習を考えると、固有空間に投影される学習サンプルの軌跡は$S$個の閉曲線群になる。この学習結果と認識を行なう画像の投影点との距離関係より、物体認識・姿勢検出を行なう。このとき、学習により得られた固有ベクトルのうちどの次元まで用いて認識するかという認識次元数の決定は、物体認識・姿勢検出の正解率・精度を大きく左右する。特に、累積寄与率が高くなるように固有空間の次元を高くすると、パターン認識の分野でよく知られている”the curse of dimensionality”（次元の呪い）という問題が生じる。固有空間の認識次元数の決定手法としては、固有空間の近似度の尺度である累積寄与率やAICによる次元数の決定法[Sakano 97]などがある。累積寄与率は固有空間を学習サンプルの圧縮記憶としてとらえたとき、蓄積した学習データの品質と圧縮率の尺度としては有効であるが、学習サンプルに含まれるノイズなどにより累積寄与率が高ければ良い認識結果が得られるわけではない。また、同じ学習サンプルを用いる場合でも、物体認識を行なうのか、物体認識と姿勢検出を行なうのかというタスクの違いにより最適な次元数が変化する。そのため、AICによる次元数の決定法を用いても不十分であると思われる。よって、学習サンプルを用いるか実際に撮影を行うなどして物体認識および姿勢検出の最適な次元を決定することが望ましい。この次元数の決定手法の詳細については第4章で述べる。

### 2.2.2 物体認識・姿勢検出手法

学習した固有空間により、物体認識・姿勢検出を行う手法について説明する。固有空間とは、画像ベクトルを表現する$N$次元直交空間（画像ベクトル空間）を回転・縮退した単なる部分空間であり、固有空間上での演算と画像ベクトル空間上での演算はほぼ等価である。たとえば画像の類似性を求めるために、固有空間上で画像の相関を求めるということは、画像ベクトルの相関を求めることと同じである。また、画像ベクトルの距離をユークリッドノルムで求めることは、固有空間上でユークリッドノルムを求めることと等価である（図2-5）。文字認識などでは類似性の尺度として正規化相関を用いるが、固有空間法を用いた物体認識では一般的にユークリッドノルムを用いる。一般的な認識の流れとしては、

*1) 計測された画像を固有ベクトルにより固有空間に投影する。*
*2) 学習サンプルの投影点列と比較して最も近いものを見つける。*

により物体認識・姿勢検出を行う（図2-6）。このように単に学習サンプルとの距離を求めることにより物体認識・姿勢検出を行うことができるが、物体数が100体程度になると物体の姿勢変化によるクラス内の分散よりも異なる物体間でのクラス間の分散が非常に大きくなる。そのため、物体認識のための固有空間と個々の認識対象に対して姿勢検出を行なう固有空間を用意して認識を行なう手法が提案されている[Murase 94a]。またこの方法において、クラス間分散

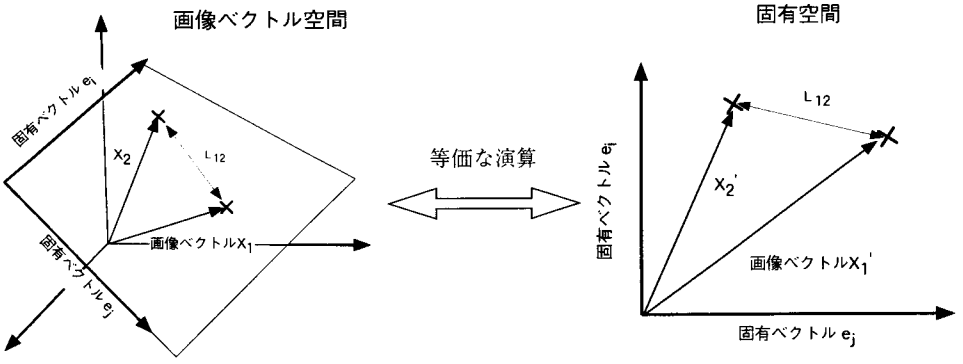



図2-5　画像ベクトル空間と固有空間上の尺度

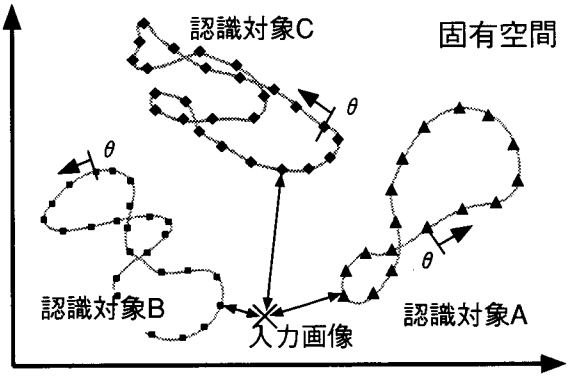



図2-6　物体認識手法

を大きくする固有空間（MEFs）とクラス内分散を大きくする固有空間（MDFs）[Daniel 96]を用いて、物体認識をMEFs、姿勢検出をMDFsで行うことにより、認識率と姿勢検出精度の向上が期待できる。しかし、本研究では扱っている認識対象数が少ないため、このようなクラス間分散とクラス内分散に関しては考慮しないものとし、一つの固有空間により物体認識と姿勢推定を行うものとする。2.2.1で述べたように複数物体の物体認識と回転による1自由度の姿勢検出を行う固有空間を生成すると、固有空間上に複数の超閉曲線が生成される。実際には、これらの超閉曲線群は学習サンプルの投影点列の集合であるため、本研究では図2-7に示すように投影点列を折れ線近似してその折れ線で表現された軌跡との距離を求めることにより物体認識を行った。このようにして物体認識を行えば、軌跡が隣接するような場合においても正しく認識を行うことができる。姿勢検出に関しては、図2-8に示すように認識時に画像の投影点と最も近い隣り合う2点の学習サンプルの投影点の組を見つけ、

$$\theta_e = \frac{l_s\theta_f + l_f\theta_s}{l_s + l_f} \tag{2-7}$$

として線形近似により姿勢を推定した。ただし、$l_s$、$l_f$は投影点$p_e$から学習サンプル$p_s$、$p_f$までの距離である。このような線形近似はパラメトリック固有空間法の前提条件である、学習サンプルの投影点の連続性を考えると学習サンプルの投影点の局所では線形性が成立すると考えられるため、妥当な近似であると思われる。

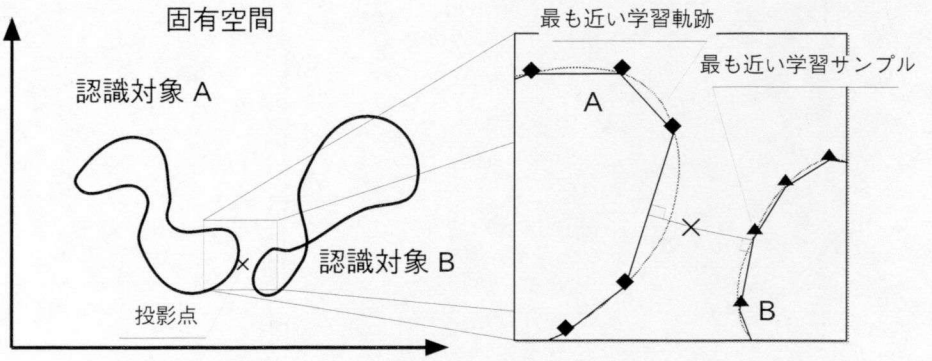



図2-7　点列比較による誤認識

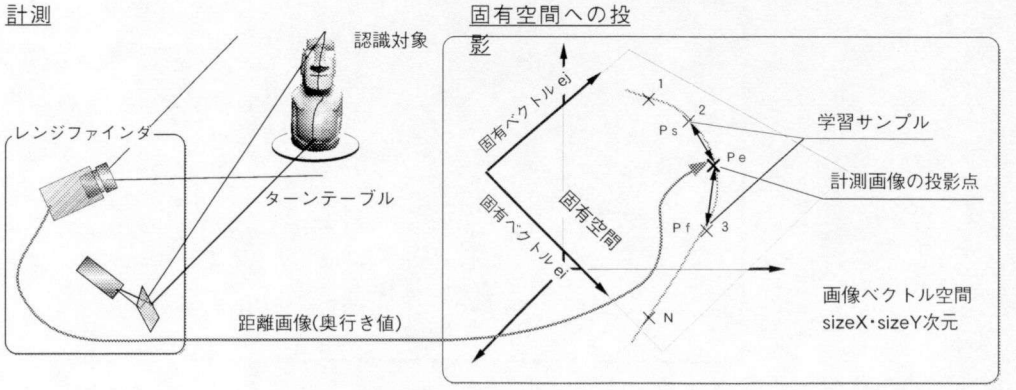



図2-8　線形近似による姿勢推定

表 2-1　距離画像を用いた固有空間法の利点欠点

| 距離画像の特徴 | 利点 | 欠点 |
|---|---|---|
| 光学的な特性を持たない | 学習サンプル数の減少 | テクスチャ成分の利用は不可 |
| 物理的な形状 | 形状モデルから学習 | ―――――― |
| 欠損画素が生じる | ―――――― | 欠損画素の考慮が必要 |
| 距離値を表現 | 表面の凹凸を利用可能<br>見た目の大きさが不変<br>奥行き位置検出 | 奥行き位置変化で誤認識 |

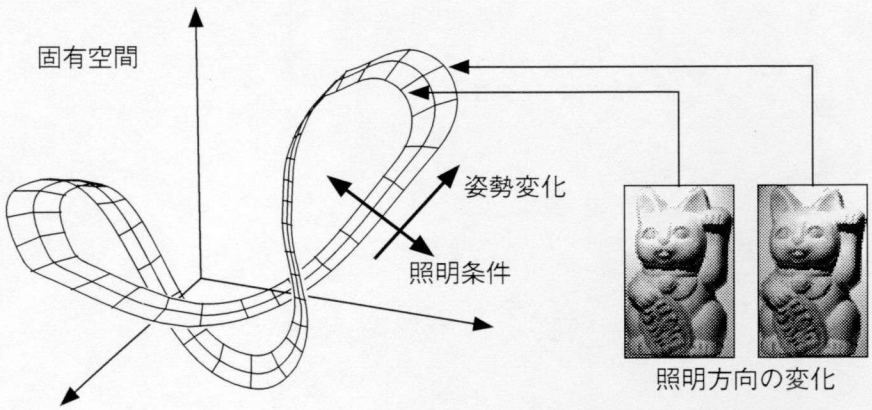



図 2-9　照明条件の推定図

## 2.3　距離画像の固有空間法に対する考察

この節では2.2.2で述べた距離画像を用いた固有空間法を濃淡画像を用いた固有空間法と比較･検討する。距離画像を用いた固有空間法は、距離画像の特徴より表 2-1 に示す利点と欠点がある。2.3.1 から 2.3.4 ではこれらの利点と欠点について述べる。

### 2.3.1　照明条件不変による利点欠点

濃淡画像を用いた固有空間法において、照明条件の変化は誤認識の要因になるため、学習においてあらかじめ照明条件の変化を考慮する必要が生じる。これに対して距離画像は照明条件の変化に依存しない対象物体の幾何学的な形状を表現しており、学習サンプル数が少なくできる。そのため、濃淡画像を用いる場合に比べて自由度の高い姿勢検出を比較的容易に実現できる。また学習サンプル数が少なくなることにより、精度の高い認識が期待できる。この反面、濃淡画像を用いた固有空間法では図2-9のように照明条件の変化も学習すれば照明条件を推定することができる。また、図 2-10 に示すように幾何学的な形状と照明条件の変化による濃淡画像の変化を固有空間上で対応づけることにより陰影からの形状推定（Shapefromshading）を

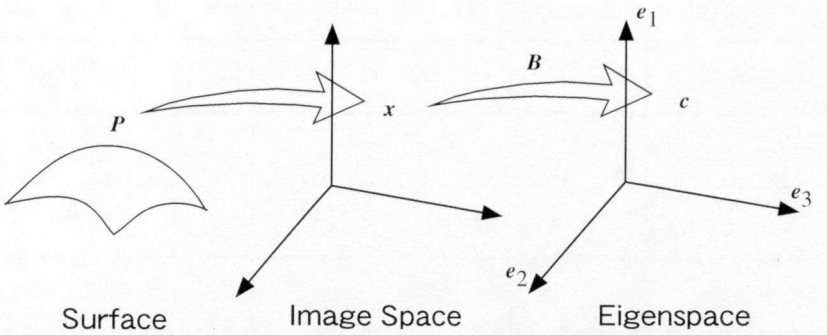



図2-10　固有空間法による Shape From Shading

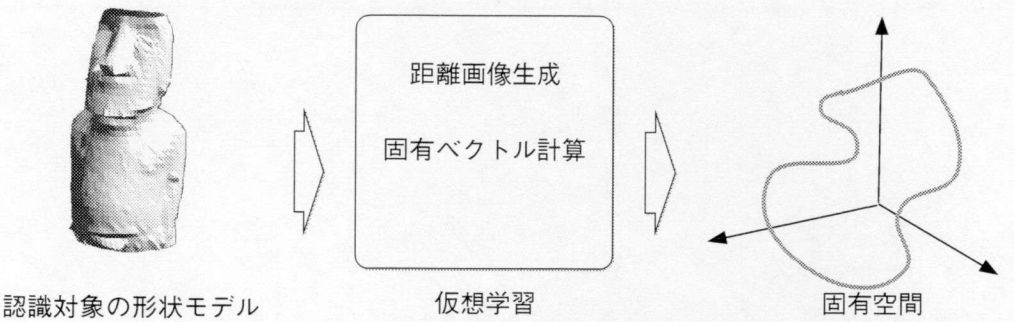



図2-11　仮想学習概念図

行うことができる[Okatani 95]。距離画像は直接形状を表現しているため形状推定を行う必要はないが、照明条件の推定や、物体認識においてテクスチャ成分を利用することは本質的に不可能である。

### 2.3.2　物理的な形状表現による利点

固有空間法のような画像照合に基づいた物体認識では学習サンプルを収集するコストが問題となる。濃淡画像やカラー画像のような見た目の画像は形状モデルやテクスチャから生成することは比較的困難であるが、距離画像では実環境で撮影される距離画像を認識対称の形状モデルから容易に生成することが可能である。このように形状モデルから仮想的に学習サンプルを生成し学習を行う手法を仮想学習と呼ぶ（図2-11）。この仮想学習により距離画像の固有空間法において学習コストを大幅に減少することができる。この仮想学習については第3章で述べる。

### 2.3.3　欠損画素問題

2.1.1で述べたようにレンジファインダには様々な計測方式があるが、いずれの手法も計測対象の幾何学的な形状を得ることを目的としており、照明条件の変化に対して不変な画像を計測できる。また、形状モデルから仮想的に学習を行なうことができる。しかし、距離画像にはデータの一部が欠けることがあり、この欠損をどのように扱うかが問題になる。

図2-2の受動型・能動型においてステレオ法・アクティブステレオ法を用いたレンジファインダでは三角測量に基づく計測を行っているため、オクルージョンを生じることがある。このオクルージョンを生じた部分では、距離値を計測することができず画像の一部が欠損する。またその他の能動型の計測方式でも、計測対象に何らかのエネルギーを照射してその様子を観測することにより距離画像を得るため、対象物体表面のテクスチャなどにより十分な反射強度が得られず計測できなくなる場合もある。本論文ではこれらの理由により計測できなかった画素を欠損画素と呼ぶ。この欠損を生じた距離画像を固有空間法へ用いるときにどのように扱えば良いかと言う問題が生じる。欠損画素を生じると固有空間へ投影する際の内積演算は厳密には行うことができない。そのため、学習画像集合の平均を求めて欠損画素に代入することも考えられるが、欠損画素が画像の大半を占めるときなどはこの方法は妥当ではない。しかし固有空間法では、画像をラスタスキャンした数百～数万次元のベクトルを高々数十次元の部分空間で表現しているため、極端な縮退写像になっている。この性質を利用すれば距離画像で生じる欠損画素に影響されない固有空間照合が可能となる。この問題の解決方法は第4章で述べる。

### 2.3.4 距離値を表現していることによる利点欠点

濃淡画像では認識対象がカメラから遠ざかったり、反対に近づいたりすると撮影される画像において認識対象が見た目に小さくなったり大きくなったりする。このような変化は固有空間法のような画像照合に基づく物体認識において誤認識を生じる要因になる。これらの問題に対し村瀬やHorstらは学習段階で見た目の大きさが変化した学習サンプルも用意することで解決しているが、見た目の大きさ変化に対応するために学習サンプルが積算的に増えるという問題がある[Murase94b][Horst 98]。2.2.1で示した距離画像を用いた固有空間生成では、距離画像のうち奥行き方向の成分のみを用いているので画像において撮影された認識対象は画像上で見た目に小さくなったり大きくなったりする。しかし、距離画像は対象物体までの距離を表現しているため、簡単な変換により見た目の大きさが変化しない画像を生成することができ、この問題を容易に解決できるという利点がある。また距離画像では、濃淡画像で捕らえにくいような対象物体表面の奥行き方向の微妙な凹凸を利用することができ、比較的精度の高い姿勢検出を行うことが可能となる。しかし、奥行き方向の変化に敏感であると言うことは、物体の奥行き方向の微妙な位置ずれにより誤認識しやすくなると言える（図2-12)。このような変化は、撮影される距離画像の全ての画素に同じ値が足し引きされるオフセット成分によるものである。正しい認識を行うために、計測された距離画像のピークや平均値を用いて除去する方法も考えられるが、2.3.3で述べた欠損画素が生じた場合を考えると必ずしも正しく求めることが期待できない。しかし、認識対象の奥行き移動が距離画像のオフセット成分になることを固有空間に組み込むことにより、物体認識・姿勢検出と同時に物体の奥行き位置を検出することも可能となる。以上の問題点とその解決方法の詳細については第5章で述べる。

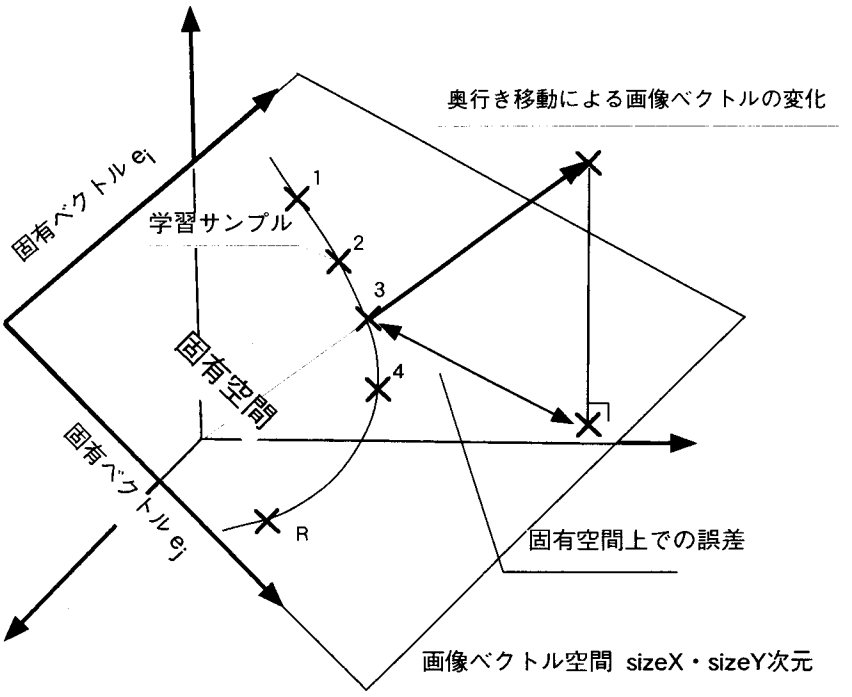



図2-12　距離画像のオフセット成分による誤認識

### 2.3.5　その他の問題点

画像計測より姿勢検出を行う全ての手法について言えることであるが、物体の姿勢変化に伴う画像の見た目の変化において、類似する見え方が存在する場合は正しく姿勢を求めることができなくなる。例えば濃淡画像であれば、表面にテクスチャを持たない単一色の立方体は姿勢検出においてどこの面が撮影されているのか判断することができないということである。これは距離画像でも同様である。しかし、立方体の各面に異なるテクスチャがある場合やそれぞれの面が異なる面で塗られている場合、濃淡画像やカラー画像では姿勢を検出することができる。立方体のように幾何学的に単純な形状の物体認識・姿勢検出はモデルベースの物体認識の方が容易であり、距離画像の固有空間法を用いる必要性はないが、濃淡画像を用いた固有空間法で識別できるテクスチャを識別できないという点で不利になる。森岡らはこの問題を指摘し、距離画像と色相情報を用いた固有空間法を提案している。具体的には図2-13に示すようにカラー画像より生成した色相画像により距離画像を分割し、色相ごとに距離画像による固有空間を複数作成する手法である。カラー画像もやはり照明条件の変化に依存するが、光源色の変化がなければ色相成分を用いることにより照明条件の変化に対してロバストな固有空間照合を実現している。

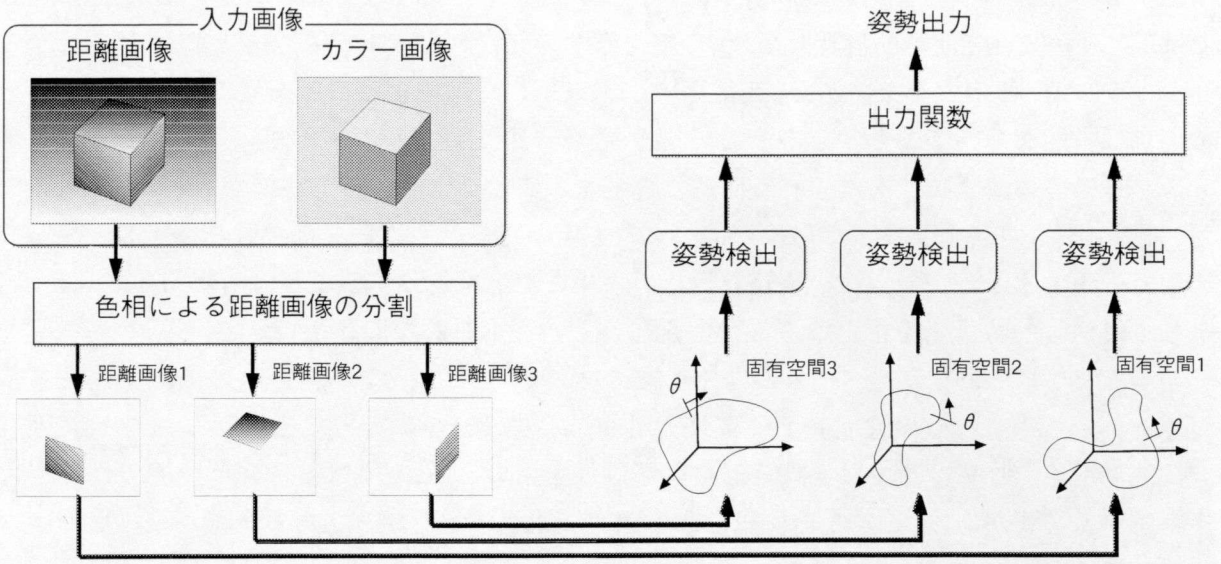



図2-13　距離画像と色相情報を用いた固有空間法

### 2.3.6　形状モデル照合と距離画像の固有空間法の比較

距離画像計測を用いたアピアランスベースの認識を行なう方法としては、距離画像列を学習サンプルとする距離画像の固有空間法のほかに、あらかじめ用意した形状モデルと計測された距離画像を単純に照合し認識を行なう方法が考えられる。これらの手法を比較すると、距離画像の固有空間法で用いている学習サンプルは形状モデルの断片の集合と考えることもできる。また、どちらの手法も方法論としては画像照合であり、本質的には等価な手法である。
トラッキングなどのようにある姿勢付近で数回の照合を行なうのであれば、距離画像と形状モデルを直接照合しても計算コストはあまり問題とはならない。しかし、物体認識・姿勢検出のように全探索を必要とするような場合では照合回数は膨大であり、直接照合する手法は計算時間の点で問題が生じる。これに対して距離画像の固有空間法では、様々な見え方の距離画像集合を用意して固有空間を生成する処理が必要になるが、認識に有効な画像特徴のみを用いた高速な距離画像照合を実現できるという利点がある。

認識に許される時間は、物体認識を用いる分野やタスクにより変化するため、一概にどれぐらいの時間で認識できればよいということはできないが、物体認識を環境理解のための要素技術などと考えれば、現実的な時間以内での環境理解のためにはより高速な物体認識を実現することが望まれる。

## 2.4 まとめ

この章では距離画像を用いた固有空間法について述べた。従来、距離画像を用いた物体認識では”距離画像＝幾何学的形状”であり、幾何学的形状から幾何モデルなどとの照合を行うことが多かった。このような発想は、距離画像が本来形状を計測することを目的としている事を考えれば自然な発想である。そのため、”距離画像＝距離値を表現する画像”としてアピアランスベースの物体認識である固有空間法に適応することは異端的な方法論かもしれない。しかし、距離画像、濃淡画像に関わらず幾何モデルを用いたモデルベースの物体認識では、幾何モデルで表現することができない複雑な形状の物体を認識対象にすることは困難であるうえ、安定して幾何学的特徴量を抽出することも難しいという本質的な問題を持っている。

近年、このようなモデルベースの物体認識に対して、アピアランスベースの固有空間法に基づく物体認識が計算機の性能向上により実用的な手法の一つになってきている。しかし、固有空間法を用いた物体認識では画像として濃淡画像を用いるのが一般的であり、照明条件の変化を考慮する必要がある。このため作成される固有空間上の学習サンプルの表現が複雑な形状の多様体になり照合コストが大きくなる。また、照明条件の変化を考慮している分だけ認識精度が低下するなどの問題がある。そこで、本研究は距離画像を固有空間法に応用した。距離画像を固有空間法に応用することにより、濃淡画像を用いた固有空間法で問題となっていた照明条件の変化を考慮する必要がなくなる。そのため、照明条件の変化により複雑になっていた固有空間表現が簡単に表現できるようになる上、物体表面の微妙な凹凸を利用できることから濃淡画像に比べて精度の高い照合が可能となる。また、認識対象の形状モデルを用いることにより仮想的な学習が行えることなどの利点がある。しかし、その一方で距離画像特有の欠損画素であるとか認識対象の奥行き方向の移動で誤認識を起こす恐れがある。これらの問題点の詳細と解決方法については第4章、第5章で述べる。

# 第3章　仮想学習による固有空間生成

固有空間法による物体認識は、学習サンプルを圧縮して表現することができるため効率の良い物体認識を行なうことができる。この固有空間法は原理的には画像照合を行なっていることと等価であり、濃淡画像、距離画像に関わらず、あらかじめ物体の姿勢変化による様々な見え方の変化を学習しておく必要がある。このため、自由度の高い姿勢検出を扱うと学習サンプルが積算的に増え、学習サンプルの獲得コストが問題となる。この章では、固有空間法で多自由度の認識を行なう際に問題となる学習サンプル獲得コストについて説明し、その解決法として形状モデルから仮想的に固有空間を生成する手法である仮想学習†について述べる。

## 3.1　固有空間法における学習サンプル獲得コスト

### 3.1.1　自由度の変化に対する学習サンプル数の変化

固有空間照合では、固有空間上の学習サンプルの投影点と入力された画像の投影点を照合することにより物体認識や姿勢検出を行なう。そのため、多自由度で位置・姿勢検出を含めた物体認識を考えると、様々な見えかたの組み合わせによるすべての画像を学習する必要がある。たとえば、対象対象の数を $S$、回転や平行移動などの運動の1自由度に対して刻み数を $R$、その自由度を $DOF$ とすると、総学習サンプル数は、

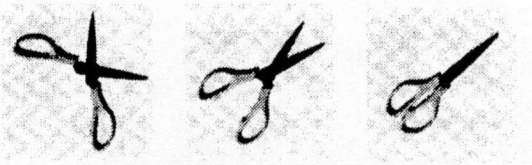
$$The\ number\ of\ learning\ samples = S \cdot R^{DOF} \tag{3-1}$$

となる。式(3-1)において学習サンプル数は認識対象の数 $S$ に対しては比例の関係にあるが、位置・姿勢などの自由度 $DOF$ に対しては指数的な関係にある。このため、固有空間法などの画像照合に基づく手法での認識において、より自由度の高い位置・姿勢検出を行うことを考えると、学習サンプル数が膨大になり、学習サンプルの計測は容易ではない。また、認識対象を多自由度で動かす装置を構成することも困難となる。このような学習サンプル数の爆発は、固有空間法のみならず画像照合に基づく認識の大きな問題である。

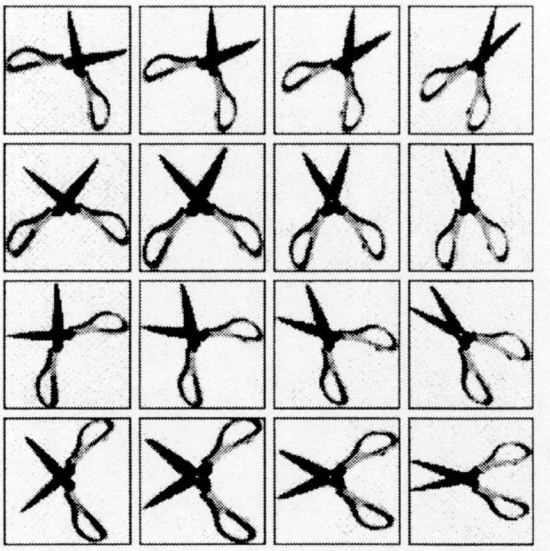
### 3.1.2　固有空間上の補間による学習サンプル獲得コストの低減

固有空間法を用いた多自由度の位置・姿勢検出では、様々な位置・姿勢で撮影される学習サンプルを必要とする。この問題を解決する方法として、少数の学習サンプルを用いて固有空間を生成し、得られた軌跡を何らかの仮定を用いて補間することも考えられる。しかし、対象物

†仮想学習では形状モデルより学習サンプルを生成しているが、距離画像の固有空間法は認識の方法論としては固有空間を用いた画像特徴レベルの照合法である。そのため、本論文では仮想学習を用いた場合でも、距離画像の固有空間法はアピアランスベースの物体認識と考えている。



(a)　学習サンプル



(b)　生成された画像

図3-1　モーフィングによる学習サンプルの生成［Murase 96］より

体の姿勢変化に伴い描かれる固有空間上の軌跡は連続的な曲線になるとは言えるが、少数の投影点列からその軌跡の形を予測することは困難である。なぜならば、軌跡を推定することを幾何学的に考えると、固有空間において何らかの制約条件を用いて任意の姿勢における学習画像を生成していることと等しいからである。また、少量の学習サンプルのときと、多数の学習サンプルのときでは画像ベクトル空間での投影点列の分布形状が変化する。そのため、生成される固有空間自体が変化する。このとき、学習サンプル数の違いにより固有空間が変化するということは、固有空間で着目する特徴量が変化するということであり、少量の学習サンプルで得られた固有空間は、学習サンプルで示される姿勢以外については最適ではないといえる。そのため、少数の学習サンプルより得られた固有空間を補間することは望ましくない。そこで、3.2では距離画像の固有空間法において認識対象の形状モデルより距離画像を仮想的に生成して学習する手法（仮想学習）を述べる。NTTの村瀬らは、関節を有する鋏やペンチなどのようにあらかじめ全ての学習画像サンプルを用意することが困難である認識対象に対し、少数のサンプル画像からモーフィングを用いて様々な学習サンプルを生成する手法[Murase 96]を提案しており、やはり固有空間法の問題点として学習サンプルの獲得コストを指摘している（図3-1）。また村瀬らは、モーフィングによる画像サンプルの仮想生成と固有空間上での補間の比較も行っており、画像サンプルの仮想生成の有効性も確認している。

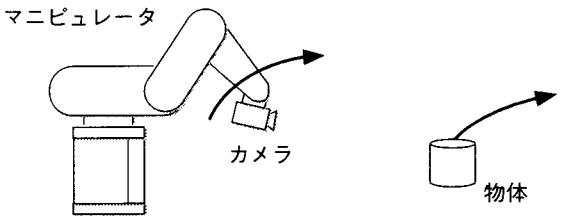



図3-2　視覚サーボ

### 3.1.3　視覚サーボにおける学習サンプル獲得コスト

固有空間法を用いて多自由度で物体の位置姿勢を求める位置・姿勢検出に対し、ある目的の見えかたになるように認識対象またはカメラを移動させるような視覚サーボでは、必ずしもすべての状態を学習する必要はない（図3-2）。阪大の横山らはこの点に着目し、すべての組み合わせによる学習ではなく、基準位置からの1自由度の姿勢・位置変化による学習サンプルをそれぞれの自由度について用意し、固有空間を生成する手法を提案している[Yokoyama 99]。また、固有空間上の投影点の分散により固有空間を階層化を行い、効率の良い探索を提案している。この手法では、学習サンプル数は物体運動の自由度に対して積算的な関係になるため、多自由度の視覚サーボも容易に実現することが出来る。このような固有空間を用いた視覚サーボは東大の出口ら[Deguchi 97]も提案しているが、出口らの手法に比べ横山らの提案した手法では大きな姿勢変化を扱うことが出来るうえ、学習サンプル数の問題も解決しており実用的な手法であると思われる。しかし多自由度の位置・姿勢検出では、どのような位置・姿勢にあってもその時の位置・姿勢を検出することが必要であり、多自由度の位置・姿勢検出における学習サンプル獲得コスト問題をこの手法により解決することは出来ない。そのため、位置・姿勢検出では全ての組み合わせが必要となり、学習サンプル獲得コストを低減する手法が必要となる。

## 3.2　仮想学習

### 3.2.1　ポリゴンモデルによる距離画像の生成

距離画像は、レンジファインダから物体までの距離情報を二次元格子状に配列した表現した画像である。このとき、物体の大きさ及び形状が既知であれば、形状モデルとレンジファインダのモデルから距離画像を仮想的に生成することができる。このような点も距離画像を用いた固有空間法の利点の１つである。

レンジファインダにおいて、物体座標系とカメラ座標系の対応はカメラパラメータで表現され、これはレンジファインダのキャリブレーションにより既知である。またカメラパラメータは、撮像素子の任意の画素における視線を示す直線（視線方程式）を表現していると考えることができる（図3-3）。この関係を用いて、図3-4(a)に示すように計算機内に配置した認識対象の形状モデルと視線方程式との交点を計算することにより容易に距離画像を生成することができる。図3-4(b)において、カメラ座標系に投影された三角ポリゴンの頂点$P_1(X_1,Y_1,Z_1)$、

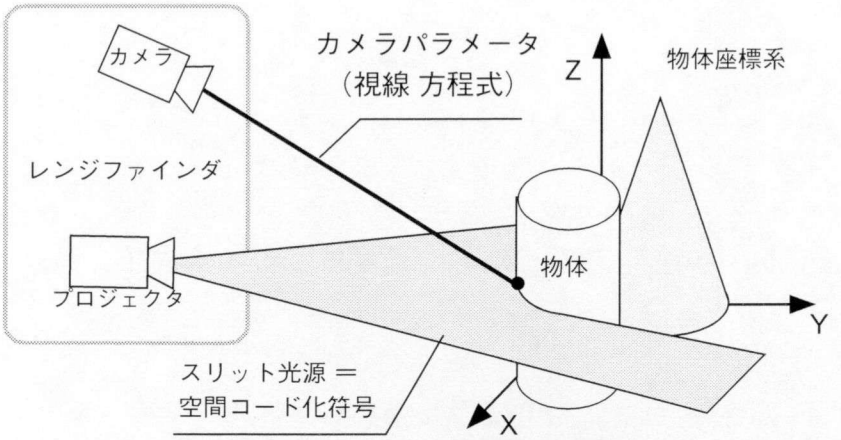



図3-3　アクティブステレオ計測

$P_2(X_2,Y_2,Z_2)$、$P_3(X_3,Y_3,Z_3)$をベクトル $v_1$、$v_2$ および $m$ で示すと3行4列のカメラパラメータ $C$ により

$$\left[v_1+m,\ v_2+m,\ m\right]=\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{bmatrix}\begin{vmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \\ 1 & 1 & 1 \end{vmatrix} \tag{3-2}$$

と表現される。このとき画素 $X_c$、$Y_c$とベクトル $v_1$ 、$v_2$ および $m$ の関係は

$$\left[v_1,\ v_2,\ m\right]\begin{bmatrix} s \\ t \\ 1 \end{bmatrix}=u\begin{vmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{vmatrix} \tag{3-3}$$

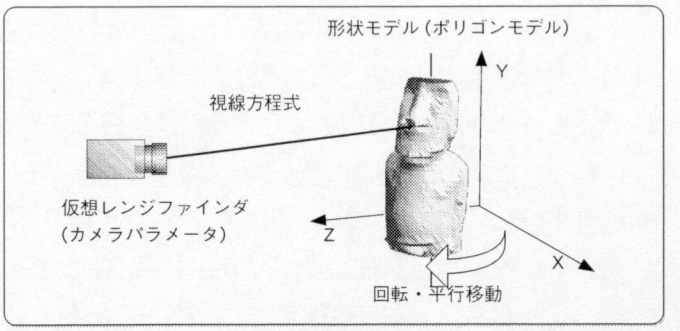
となる。このとき係数 $s$ 、$t$ を求めることにより、視線と撮像素子面の交点をもとめることができ、交点は $\boldsymbol{P}$

$$\boldsymbol{P}=s\,v_1+t\,v_2+m \tag{3-4}$$

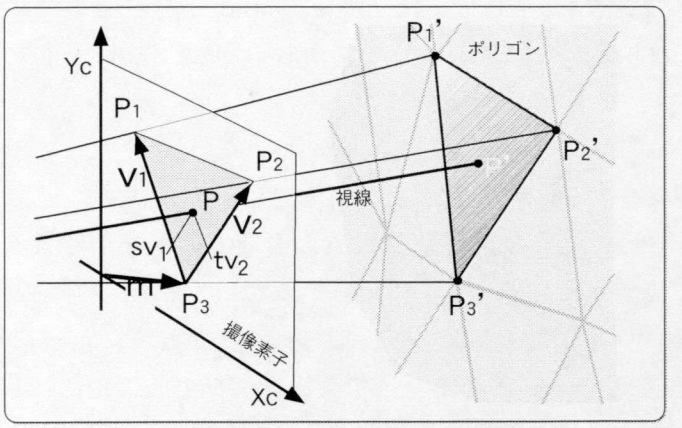
となる。また、求められた交点が $\boldsymbol{P}$ について

$$\begin{aligned} &s>0\,,\,t>0 \\ &s+t\leq 1 \end{aligned} \tag{3-5}$$

により交点がポリゴンの内部にあるかを判断する。陰面消去については、すべてのポリゴンに対する交点のうち最もカメラに近い結果を用いればよい。この操作を全ての画素に対して行えば、距離画像を仮想的に生成することが出来る。このような距離画像生成により様々な姿勢による距離画像を生成すれば、実際に学習サンプルを計測することなく固有空間を生成することが可能になる。




(a)　概念図




(b)　ポリゴンモデルとの交点

図 3-4　仮想学習

### 3.2.2　レンダリングハードウェアによる距離画像の生成

3.1.1 で説明したポリゴンモデルから距離画像を生成する手法は、幾何学的には簡単な手法である。しかし、ポリゴン数が多くなったり、自由度が高くなると計算時間が膨大になる。特に自由度が上がるとその計算コストは積算的に増える。固有空間法では、学習と照合の二つの流れにより認識を行なう手法であり、高速に学習を行う必要はないが、学習に要する時間があまりにも長すぎると実用的ではない。そのため、より高速に距離画像を生成する方法が必要とされる場合も考えられる。このような場合、阪大の日浦らが行なったレンダリングハードウエアを応用した距離画像の生成法を応用することができる[Hiura 96]。具体的には、レンダリングハードウエア上で形状モデルを描画し、その後レンダリングハードウエアのZバッファ（も

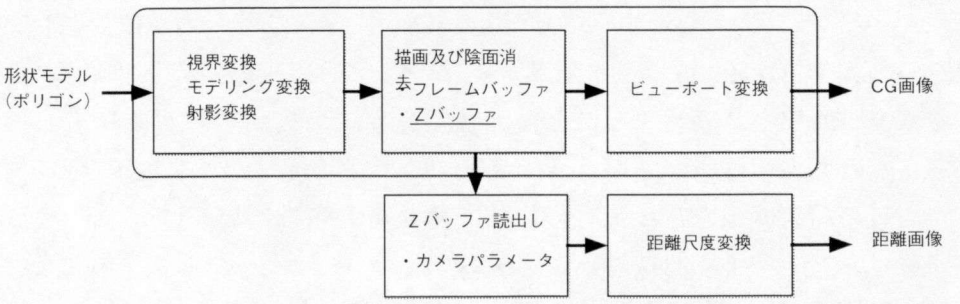



図 3-5　レンダリングハードウエアによる距離画像生成

しくはデプスバッファ）を読み出すことにより距離画像を生成する手法である。通常Zバッファは、3Dモデルの描画において陰面消去を行うためにポリゴンの重なりの関係を記憶しておくためのものであるが、これは簡単な変換により距離値を得ることが可能であり、レンジファインダのカメラパラメータが既知であれば、レンジファインダの座標系における距離画像へ変換することができる（図3-5）。また、レンジファインダの距離分解能が通常8bitから16bit程度であるのに対して、Zバッファの分解能は通常24bitもしくは32bitであり、距離画像の生成精度の点でも問題を生じることはない。距離画像を生成する時間はレンダリングハードウエアがポリゴンモデルを描画する時間に相当する。近年の高速なハードウエアを用いることを考えれば、DOS/V機の拡張ボード程度のもので毎秒100万個のポリゴン描画が可能なものもあり、総ポリゴン数5000程度の形状モデルであれば、毎秒100枚程度の距離画像を生成することが期待できる。

### 3.2.3 ビジュアルベース仮想学習

距離画像を用いた固有空間法では、認識対象の形状モデルを用意することにより仮想学習を行うことができる。仮想学習において、形状モデルのほかに認識対象表面のテクスチャモデルも用意すれば、濃淡画像を用いた固有空間法においても仮想学習を実現できる。しかし、濃淡画像を用いた固有空間法で仮想学習において形状モデル生成のためにレンジファインダも用意する場合、CCDカメラのみで3次元物体の学習・認識ができるという濃淡画像の固有空間法の簡便性が失われてしまう。しかし、はじめから認識対象の形状テクスチャモデルが得られているなら濃淡画像での仮想学習も有効であると思われる。電総研の王らは、モーションステレオと固有空間法を組み合わせ、モーションステレオににより生成されたテクスチャを持つ形状モデルより濃淡画像による仮想学習（ビジュアルベース仮想学習）[Wang 98]を実現している。この手法において、モーションステレオでは様々な見え方を計測しているため、仮想学習を用いる利点は少いように思える。しかし、一度生成された形状モデルから任意の見え方を再構成できる上、モーションステレオで計測した画像数よりも多くの画像を仮想的に生成できることもあり、このような利点を考えると有効な手法である。

## 3.3 仮想学習実験

### 3.3.1 認識対象と形状モデル

仮想学習の実験では図3-6に示す4体の認識対象を用いた。これらの認識対象は、高さが13cm程度の陶器製の人形である。これらの認識対象は幾何学的に複雑であり、形状から特徴量を求めて拡張幾何モデルなどで表現することは困難である。また、どのモデルにも人形の首の所にくびれがあるなど類似する所が多いうえ、幾何学的な特徴量は膨大な数になる。そのため、幾何学的な特徴量の対応関係からこれらの物体を認識するのは簡単ではない。

形状モデル生成では、まず認識対象を図3-7のようにターテーブルにより回転させ、全周を60deg.おきに全部で6枚の距離画像をレンジファインダにより撮影した。通常多視点距離画像

(a) 招き猫　(b) 花っ子　(c) タヌキ　(d) シーサー

図 3-6　認識対象

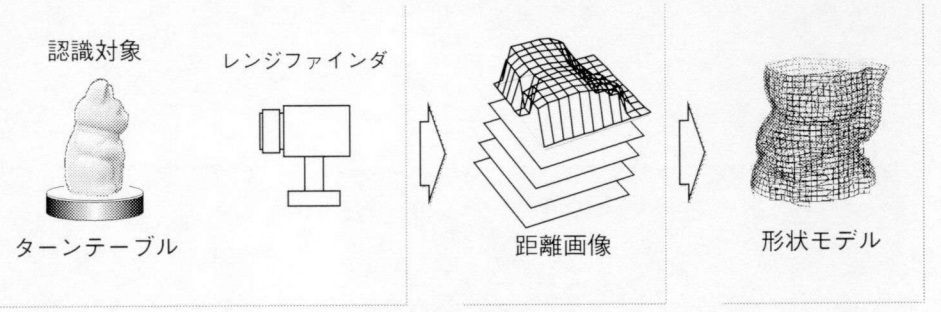



図 3-7　全周形状の計測

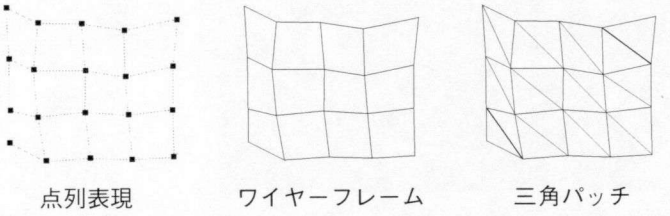
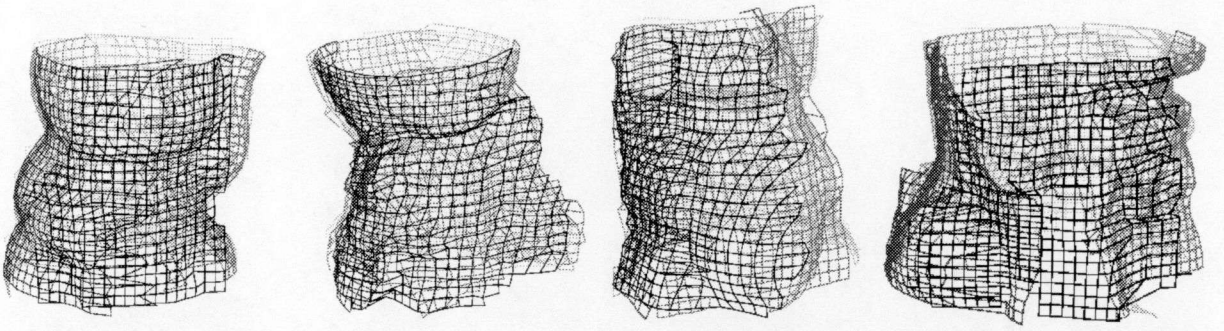


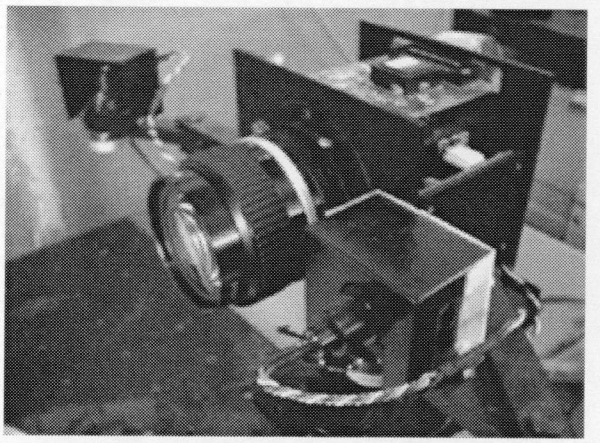

図 3-8　三角パッチによる形状の表現

から全周形状モデルを作成するときには、メッシュ化、位置あわせ、重なったパッチの消去、縫い合わせなどの操作により 1 つに統合された形状モデルを作成する。しかし、仮想学習では多視点で計測された距離画像が統合されている必要はない。そのため、ここでは計測された6枚の距離画像の位置合わせを行ない、図3-8のように三角パッチによる表現を用いた。この三角パッチによるモデル記述は、仮想学習を行なう際に距離値の読み出し方法を工夫すれば特別なモデル記述は必要としない。具体的には、距離画像の枚数、距離画像の縦横の大きさを記述したヘッダの後に距離画像の枚数分の座標値が羅列されたものを形状モデルとしている。以上のモデル化を図 3-6 で示した 4 体の認識対象について行なった結果、図 3-9 に示す形状モデルが生成された。ただし、距離画像の位置合わせを行なう際に必要となるターンテーブルの回転軸は、あらかじめキャリブレーションにより求めている。この形状モデルを生成するにあたり、レンジファインダとして図 3-10 に示すシリコンレンジファインダ(SRF )[Sato 94] を用いた。SRF は、画像解像度 24 × 24 、奥行き分解能 1mm 程度のレンジファインダである。

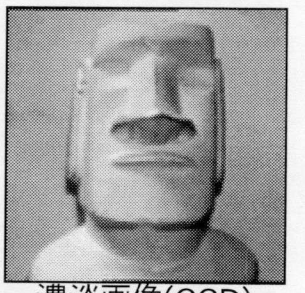


(a) 招き猫　　(b) 花っ子　　(c) タヌキ　　(d) シーサー

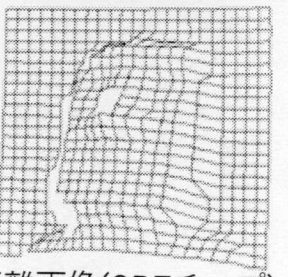
図3-9　形状モデル



・距離画像と濃淡画像の撮影が可能
・距離画像計測速度：毎秒30枚以上
・距離計測精度1mm程度
・画像解像度24×24画素



濃淡画像(CCD)



距離画像(SRFチップ)

図3-10　シリコンレンジファインダ(SRF)[Sato 94]

### 3.3.2　仮想学習結果

仮想学習の実験は、SiliconGraphics Octaneにより行なった。3.3.1で示した形状モデルはそれぞれ6300ポリゴン程度で構成されている。この形状モデルを用いて、計算機上で図3-11に示すように認識対象の鉛直軸周りでの1自由度の回転について距離画像を生成し、得られた学習サンプル群を用いて固有空間を生成した。また距離画像の生成は、認識で用いるSRFのセンサのスペックより画像解像度を24×24画素としている。生成された距離画像の一例を図3-12に

示す。仮想学習の結果、図3-13に示す平均ベクトルおよび固有ベクトルが得られた。また、固有ベクトルにより学習サンプルを固有空間へ投影した結果、学習サンプルの軌跡は図3-14のように生成された。仮想学習において、学習サンプルは鉛直軸周りの1回転に対して64枚の距離画像を生成しており、4物体すべてに対する学習サンプル数は256枚である。このとき、仮想学習に要した時間は、256枚の距離画像生成に2分52秒、固有空間の生成に19秒であり仮想学習全体では3分11秒であった。

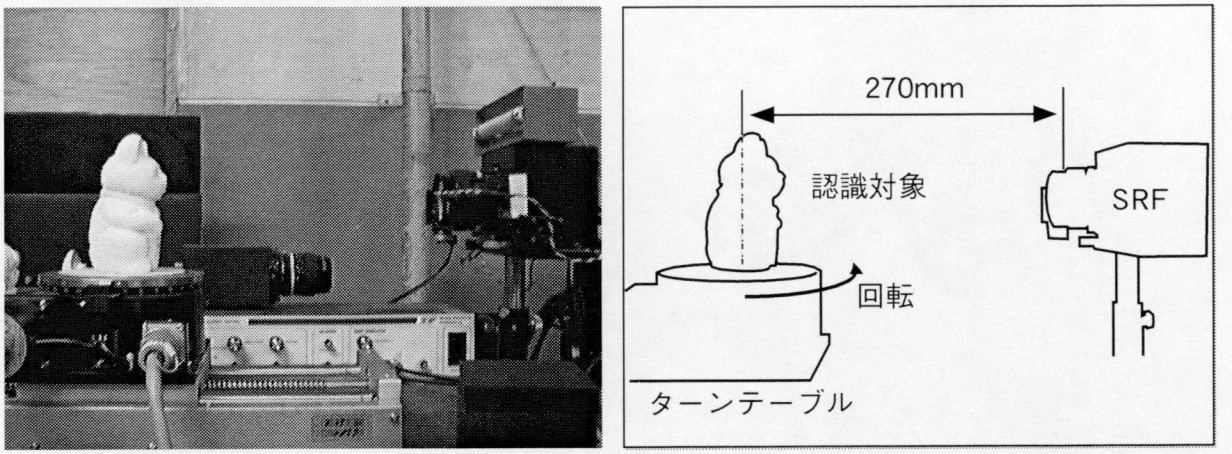



図3-11　実験環境

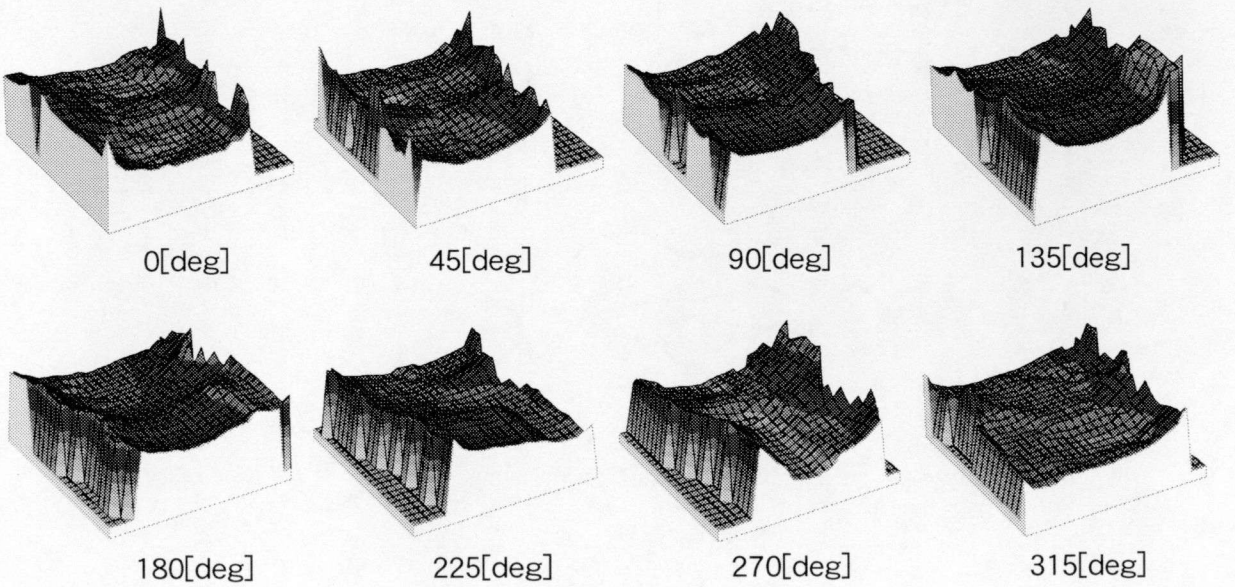



図3-12　生成された距離画像

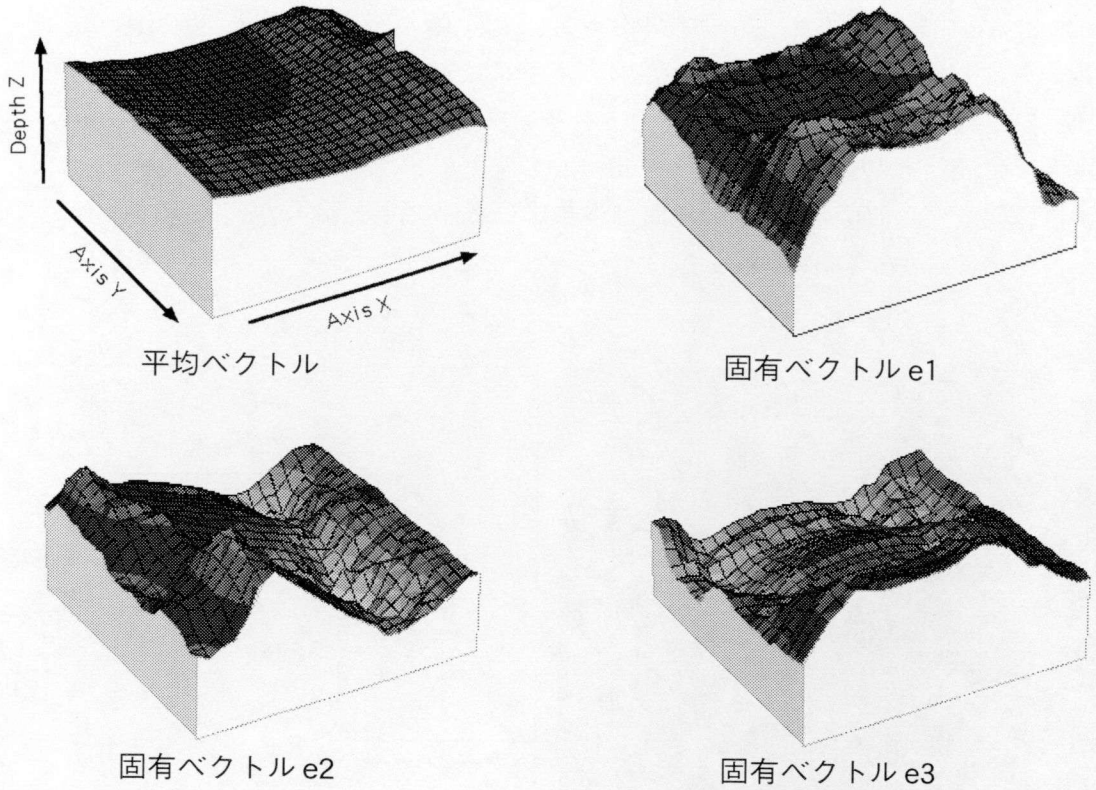



図 3-13　平均ベクトルと固有ベクトル

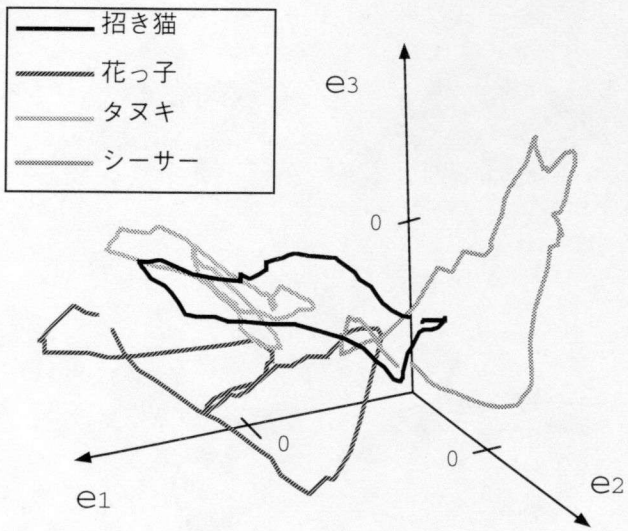



図 3-14　固有空間上での軌跡

### 3.3.3 物体認識・姿勢検出

仮想学習により得られた固有空間を用いて、実際のレンジファインダで物体認識・姿勢検出を行なった。実際のレンジファインダで計測される距離画像では、オクルージョンやその他の理由により欠損画素が生じる。欠損画素は物理的量を持たないため、固有空間法においてどのように扱えばよいかが問題となる。この問題については欠損画素を0として距離画像を固有空間に投影し、欠損画素による投影誤差を固有空間上で補正を行うBPLPと、欠損画素のパターンに応じて固有空間を再構築するARLPにより解決した。また、認識次元数の決定は仮想学習を応用したシミュレーションにより物体認識、姿勢検出とも7次元とした。欠損画素問題の解決手法であるARLPとBPLPの詳細と、認識次元の決定手法については第4章で述べる。物体認識・姿勢検出実験を以下に示す流れで行なった。

*1) 認識対象の距離画像をレンジファインダにより計測する。*
*2) 計測された距離画像に対してBPLPを用いて7次元で物体認識を行なう。*
*3) 認識された物体の学習結果のみを用いてARLPにより7次元で姿勢検出を行なう。*

この手法を用いて、図3-6で示した4体の認識対象をそれぞれ3deg.刻み(120ステップ)で距離画像を計測し、物体認識・姿勢検出を行なった。また、距離画像計測はSRFで行った。その結果、それぞれの物体の認識において誤認識は120回の認識のうち2回以下であり、物体認識率は98%程度であった。表3-1に全物体での物体認識率と姿勢検出誤差を、また図3-15に姿勢検出結果の一例を示す（認識対象：招き猫）。図3-15を見ると姿勢角度210deg.から250deg.付近で大きな誤差が生じており、249deg.のところでシーサーと誤認識した。これは、招き猫の姿勢角度249deg.のとき計測される距離画像が、シーサーの学習サンプルと類似していたため、SRFの計測誤差や欠損画素で生じた投影誤差により物体認識を誤ったものと思われる。また、表3-1を見るとシーサーの物体認識・姿勢検出結果は他の認識対象に比べると1.5倍程度の誤差が生じている。これはシーサーを撮影したときの距離画像にジャンプエッジがほとんど含まれていなく、画像のエッジは固有空間照合において重要な特徴量となっているためであると思われる。また、全物体全姿勢角において、欠損画素数は平均で30%程度であった。

**表3-1 認識率・姿勢検出誤差**

| 認識対象 | 認識率[%] | 姿勢検出誤差<br>標準偏差[deg.] |
|---|---|---|
| 招き猫 | 99.16 | 6.44 |
| 花っ子 | 99.16 | 8.04 |
| タヌキ | 99.16 | 6.49 |
| シーサー | 98.33 | 10.41 |

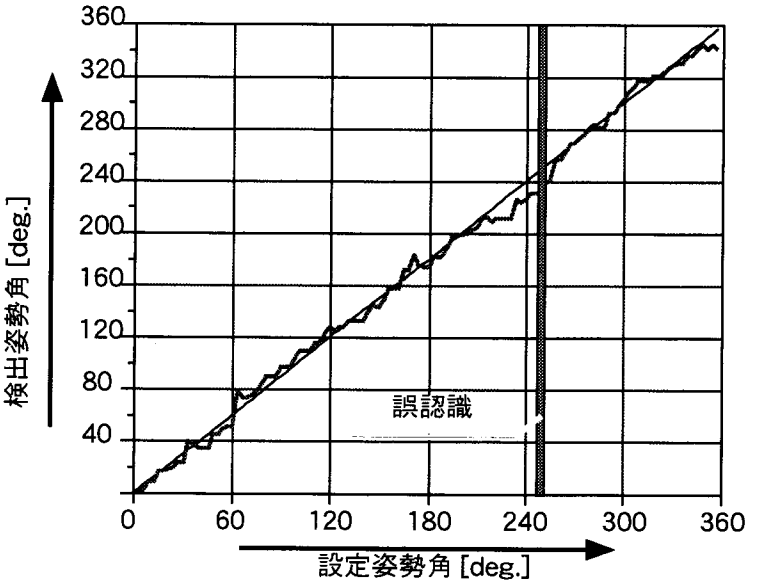



図3-15 姿勢検出結果（招き猫）

## 3.4 仮想学習考察

### 3.4.1 学習サンプル獲得コストと距離画像生成時間

実験では、4物体についてそれぞれ認識対象の鉛直軸周りの仮想学習を行なった。仮想学習を行なうにあたり形状モデルとして認識対象の6方向からの距離画像を組み合わせたものを用いているため、4物体に対して24枚の学習サンプルから256枚の学習サンプルを生成していることになる。そのため学習サンプル獲得コストとしては10分の1以下に抑えることができたと考えることができる。このような1自由度の学習は、実際に画像を計測して学習を行なってもそれほど学習サンプル獲得コストが問題となることは少ないが、学習サンプル獲得コストは、自由度に対して指数的に増加する傾向があり、このように一自由度の学習であっても、学習サンプル獲得コストを下げることを示すことができれば、多自由度の学習への可能性を示すことができる。本章では6枚の距離画像を張り合わせたものを形状モデルとして用いているが、FAや日常環境に存在する工業製品などはその設計過程においてCADデータなどを持っていることが多い。仮想学習はポリゴンモデルのみならず、CADデータを用いて行なうこともでき、もしもCADデータがあれば学習サンプルを計測する必要はない。

### 3.4.2 ポリゴンモデルによる距離画像の生成誤差

ポリゴンモデルは、物体の形状を多面体の集合で表現したものである。このようなポリゴン表現は、形状を区分的に線形近似していることと等価である。そのため、仮想学習で生成している距離画像にはポリゴン近似による誤差を含んでいることになる。しかし、固有空間はLowPassフィルタの性質をもっており、ポリゴン近似による誤差が平滑化されるため生成される固有ベクトルでは問題にはならない。LowPassフィルタの性質は、累積寄与率を低くすると強いLowPassフィルタとなる傾向にあり、固有空間法による物体認識では一般的ノイズ除去を

行うために累積寄与率をある程度の値以下に押さえることが多い。しかし、何らかの理由で累積寄与率の高い固有空間を用いる場合は、認識対象のポリゴン近似による誤差を無視できなくなる可能性がある。このような場合は、より精密なポリゴンモデルを用いる必要があり、認識に用いるレンジファインダよりも十分に解像度の高いものが必要となる。

### 3.4.3 欠損画素問題

仮想学習は、カメラパラメータと形状モデルを用いて幾何学的な関係から距離画像を生成して学習を行う手法であり、実際のレンジファインダによる計測で生じる欠損画素を考慮していない。そのため、ここで説明した仮想学習の結果では欠損画素をどのように扱えばよいかが問題になる。欠損画素数がわずかであれば欠損画素に適当な値を代入しても大きな問題とはならないが、欠損画素数が多くなると無視できない問題となる。この問題を解決するには以下の二つの方法が考えられる。

*1) 仮想学習において実際のレンジファインダで生じる欠損画素も考慮する。*
*2) 固有空間照合で欠損画素を含む距離画像に対しても正しく照合できる手法を用いる。*

1)の方法はレンジファインダのプロジェクタパラメータや物体表面のテクスチャも考慮することにより、三角測量に基づくレンジファインダで生じるオクルージョンやアクティブレンジファインダのセンサ感度により欠損を予測することにより実現することができる。2)の方法を用いるとき、固有空間照合において欠損をどのように扱えばよいかが問題になるが、固有空間法が極端な縮退写像であることを考えれば全ての要素を用いない固有空間照合が可能である。これら2つの方法論の考察と固有空間法における欠損画素問題の解決手法については第4章で述べる。

## 3.5 まとめ

固有空間法はアピアランスベースの物体認識手法であり、多自由度の位置・姿勢検出を考えると学習サンプル獲得コストが問題となる。本章では、この問題の解決法として距離画像の固有空間法において形状モデルから仮想的に固有空間を生成する手法を説明した。3.3では、認識対象の鉛直軸での回転による1自由度の仮想学習の実験結果を示した。この実験では24枚の学習サンプルから256枚の学習サンプルを生成しており、学習サンプル獲得コストを10分の1に抑えることができた。そのため、学習サンプル獲得コストを低減するという点において有効性を示すことができた。しかし、多自由度の仮想学習を行なう際には距離画像生成時間が問題となると予想される。この問題に対しては、レンダリングハードウェアによる距離画像の生成を用いることができ、多自由度の仮想学習も比較的容易に実現できることが期待できる。また、電総研の王らの研究のように形状モデルがあらかじめ用意されている状況であれば、仮想学習は濃淡画像を用いた固有空間法においても有効な手法であると思われる。

# 第4章　欠損したデータの固有空間法と固有空間を知識とした補間法

第3章では学習サンプル獲得コストの問題について述べ、解決策として仮想学習を提案してその有効性を検討した。実験の結果、学習サンプル獲得コストの削減において仮想学習の有効性が確認された。しかし、仮想学習では実際の距離画像計測で生じる欠損画素を考慮していないため、欠損画素を含む画像をどのように扱ったら良いかという問題が生じる。この章では、まず距離画像計測において欠損画素が生じる仕組みについて説明し、欠損画素の物理的意味から固有空間照合での問題点を述べる。次に欠損画素を含む画像を固有空間照合する方法として、固有空間上での逆射影による方法と固有空間の再構築による方法を提案し、これらの比較検討を行う。最後に、欠損画素を含む画像の固有空間照合法を応用した欠損データの補間法を提案し、濃淡画像の補間実験よりその有効性を検討する。

## 4.1　欠損画素の意味と固有空間照合での問題

### 4.1.1　欠損画素とは

距離画像は、レンジファインダと対象物体の位置関係を示す距離値が格子状に配列された画像である。一般的に用いられる距離画像の計測方式は2.1で示したような光学的な計測である。そのため、何らかの理由により撮像素子に十分な受光量が得られない場合、距離値が計測できなくなる。このような現象が生じると距離画像の一部の画素においてデータが得られず欠損となる。このような画素を欠損画素という（図4-1）。この欠損画素が生じる主な要因としてはオクルージョンによる遮光とテクスチャによる投影光の減衰が挙げられる。

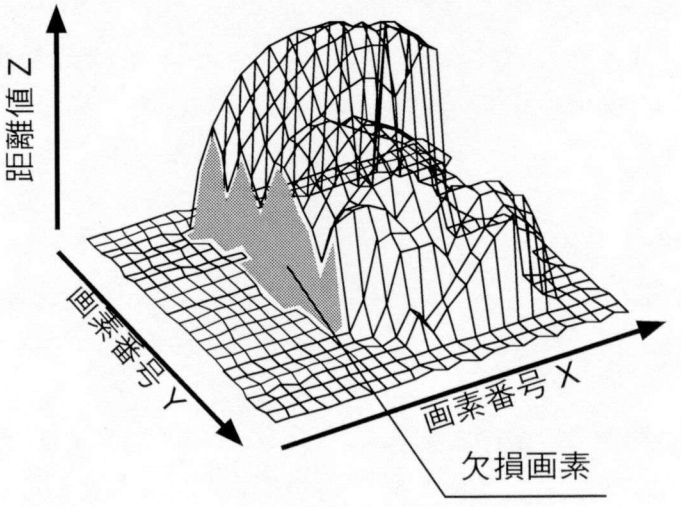



図4-1　欠損画素

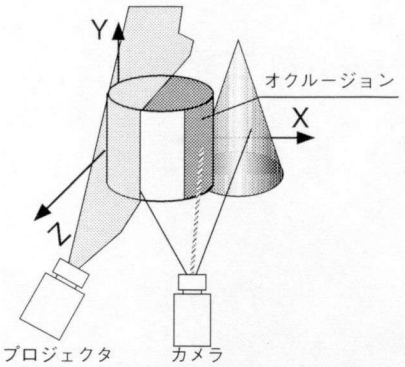



図4-2　アクティブステレオ法での欠損

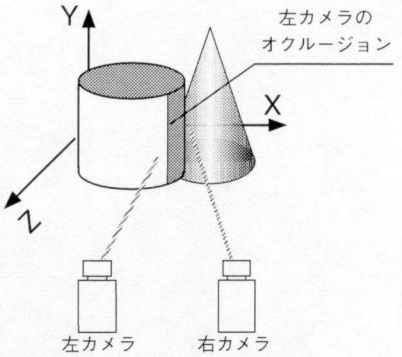



図4-3　ステレオ法での欠損

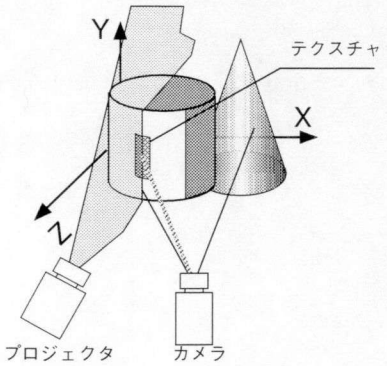



図4-4　テクスチャによる欠損

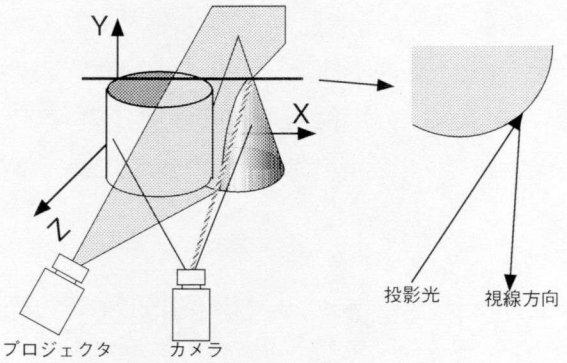



図4-5　対象物体表面の傾斜による欠損

アクティブステレオでは、カメラで撮影される画像においてプロジェクタの投影光が照射されないところでは、空間コードが得られないため距離値を得ることができなくなる（図4-2）。また、ステレオ法では、左右のカメラで同時に観測できないところでは、視差量を求めることができないので距離値を得ることができない（図4-3）。このようにアクティブ、パッシブに関係なくステレオ法、アクティブステレオ法、モアレ法のように三角測量に基づくレンジファインダではオクルージョンが生じることがあり、これにより欠損画素が生じる。またアクティブ計測では、計測方式によりパターン光や、変調光、干渉光などの光を対象物体に照射しその様子を観測することにより距離値を得ている。そのため、いずれの場合でも図4-4のように対象物体表面に反射率が低いテクスチャがある場合などは、対象物体に照射された光を観測することができず欠損画素となる。オクルージョン、テクスチャ以外の要因としては、物体表面の形状や光沢による欠損なども考えられる。アクティブレンジファインダにおいて図4-5のように物体表面の法線方向がプロジェクタの視線に対して垂直に近い場合、撮像素子に十分な光量が入射されないため距離値を得ることができなくなる。以上の理由により、距離画像では欠損画素が生じる。

これに対して、濃淡画像では十分な光量が得られなかった画素では、熱雑音との比率を考えるとあまり有用な情報ではない。そのため、十分な光量が得られなかった画素を欠損画素と考えることもできる。しかし、その画素に相当する物体表面上の領域では、黒いテクスチャがあ

るか又は照明が遮られているため十分な輝度が得られなかったということを意味しており、物理的な意味を持っている。このような特徴を用いることでshape-from-shadingであるとかテクスチャによる物体認識などを行なうことができ、有用な情報となる。そのため、濃淡画像では撮像素子の欠陥がない限り距離画像で生じる欠損画素に相当する現象はない。従来固有空間法を用いた画像認識では、濃淡画像を用いるのが一般的であり、欠損した画素を含む画像をどのように固有空間法に用いるかということは考えられてこなかった。

### 4.1.2　固有空間照合での問題点

距離画像の固有空間法では、形状モデルから仮想的に学習を行なうことができ、固有空間法のような画像照合に基づく物体認識の問題である学習サンプル獲得コストを削減することができる。しかし、実際のレンジファインダにより計測される距離画像において欠損画素をどのように扱えばよいかが問題となる(図4-6)。

仮想学習を用いない場合でも、照合において欠損画素を含む距離画像を固有空間に投影することが必要となる。しかし、このような場合は学習サンプル・照合画像はどちらも実際のレンジファインダで撮影される距離画像であるため、学習サンプルと照合画像で同じ画素が欠損画素となる。そのため、欠損画素は生じているものの学習サンプルと照合画像は全く同一の画像であり、欠損画素に適当な値を用いれば問題は生じない。そのうえ、欠損画素はランダムに生じるわけではなく、認識対象のオクルージョン領域であるとか、テクスチャにより生じるため、欠損画素にいくらの値を用いるかにより精度の高い認識を行うこともできる。例えば、欠損画素に欠損しなかった画素の値とかけ離れた値を用いれば、認識において欠損画素が生じるパターンを積極的に使うことができる。反対に、欠損画素を欠損しなかった画素の平均値とすれば欠損パターンを重視しない照合を行うこともできる。このように、学習サンプルと照合画像で同じように欠損画素が生じている場合は、欠損画素を有効に利用することができる。しかし、第3章の仮想学習で説明した方法では距離画像生成で欠損画素を考慮していないため問題が生じる。この問題を解決するためには、二つの方法が考えられる。

*1)　欠損画素を考慮した固有空間を生成する。*
*2)　欠損画素に依存しない固有空間照合を行なう。*

1)の手法を実際に行なうには、仮想学習の距離画像生成において、プロジェクタパラメータや

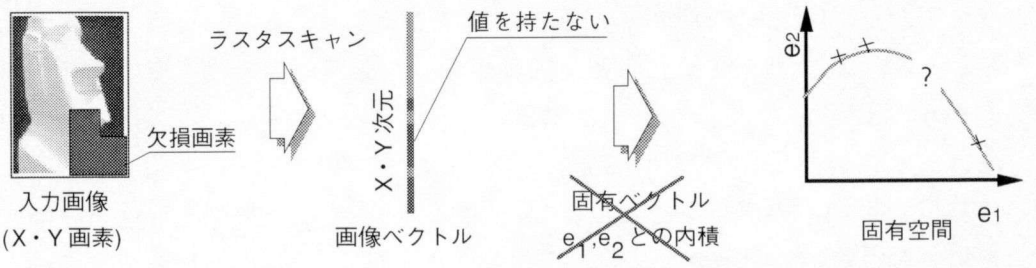
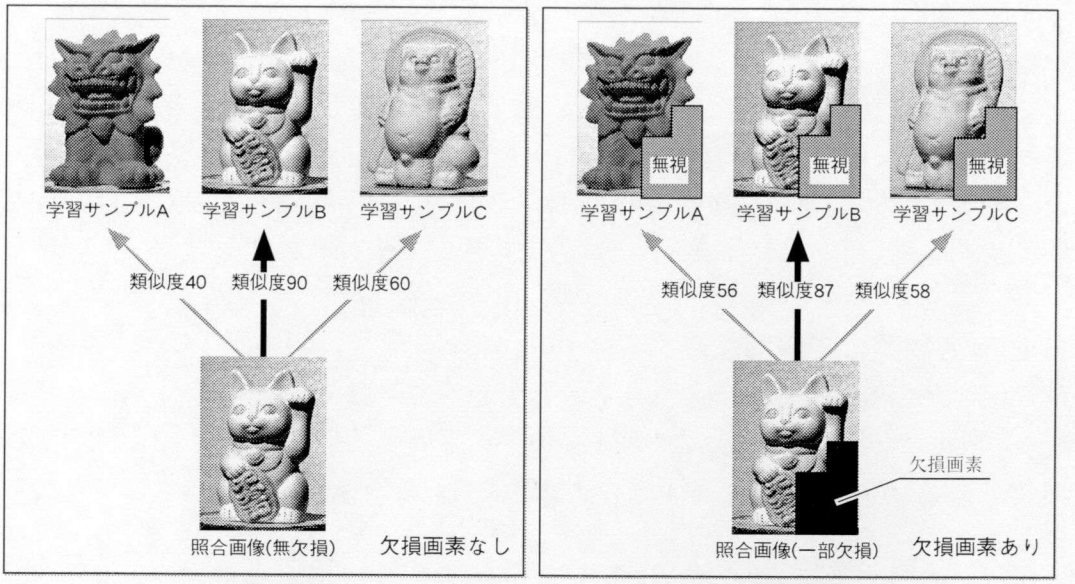



図4-6　欠損画素を含む画像の固有空間照合




(a)　欠損画素がない場合　　(b)　欠損画素がある場合

図4-7　画像照合での欠損画素問題

認識対象表面のテクスチャモデルも用いて欠損画素も考慮した距離画像生成を行なえばよい。しかし、ステレオ計測におけるオクルージョンやテクスチャによる投影光の減衰を正確に評価することは難しい。そのうえ、テクスチャモデルを用意する必要があり、距離画像の固有空間法の利点が半減してしまう。また、カメラパラメータとプロジェクタパラメータから欠損画素のなりやすい領域をあらかじめ求めておき、その領域を用いないで固有空間生成・照合を行なうという手法も考えられる。しかし、この方法では認識対象の姿勢変化において偏心があると画像領域の大部分が欠損が生じる領域となり、姿勢変化により欠損する恐れのある領域の和集合を用いないとするなら、照合に有効な領域が無くなってしまうという問題がある。そのため、2)の方法により解決することが望まれる。

2)の手法において簡単に解決するには、照合において欠損画素の画素値を適当な値で表現する方法が考えられる。例えば、欠損画素の周囲の画素値により補間する方法や学習サンプルの平均画像の値を用いる方法などが考えられる。欠損画素の補間は、距離画像から形状モデルを生成する際に一般的に行われる処理であり、欠損の周囲画素を用いて線形補間やスプライン補間を行うことが多い。しかし、このような補間を物理的に考えると認識対象の形状を平面であるとか滑らかな曲面であると仮定することと等しい。そのためこのような方法は、距離画像の固有空間法において認識対象の形状について制約条件をつけることになり、あまり望ましくない。平均画像の画素値を用いる方法は、欠損画素に無難な値を用いるという点では筋が良い方法とは言える。しかし、欠損数が多くなると欠損画素を含む距離画像の投影点は固有空間の原点に次第に向かうことになり、正確な認識が困難になる。このため、単純に欠損画素を適当な値とする方法はあまり望ましくない。

固有空間法を用いない単純な画像照合では、図4-7のように画像の類似性を評価する際に欠損画素を用いないで照合を行なえば問題を生じることはない。固有空間上でもこのような方法を用いれば欠損画素に依存しない照合が可能となるが、このときどのようにすれば固有空間に

投影することができるかが問題となる。しかし、固有空間法では画像をラスタスキャンした数百～数万次元のベクトルを高々数十次元の部分空間で表現しているため極端な縮退写像になっており、このような縮退写像の特性を考えれば欠損画素が生じても正しく投影することが可能になる。

## 4.2　欠損画素を含む画像の固有空間照合

距離画像計測で生じる欠損画素は、固有空間法でどのように扱えばよいかが問題となる。この節では欠損画素に依存しない固有空間照合に基づき、欠損画素逆行列による解決法と固有空間の再構築による解決法を提案する。また、欠損画素逆行列による解決法をBPLP(Back Projection for Lacked Pixels)、固有空間の再構築による解決法をARLP(Adaptive Reconstruction for Lacked Pixels）と命名する。

### 4.2.1　欠損画素逆行列を用いた固有空間照合法（BPLP）

ここでは、欠損画素を含む画像の投影点と欠損画素がないときの投影点の位置関係から、欠損画素を含む画像の投影点を欠損画素がないときの投影点へ近似的に逆投影する変換行列を考える（図4-8）。

アクティブレンジファインダで撮影される距離画像では、計測より得られた空間符号化コードの値より欠損画素を容易に見つけることができる。この欠損画素情報より、欠損した画素に学習サンプルの平均値を代入したとすると、欠損画素を含む画像ベクトル$\hat{x}$は欠損画素がない画像ベクトル$x$を用いて式(4-1)のように示される。

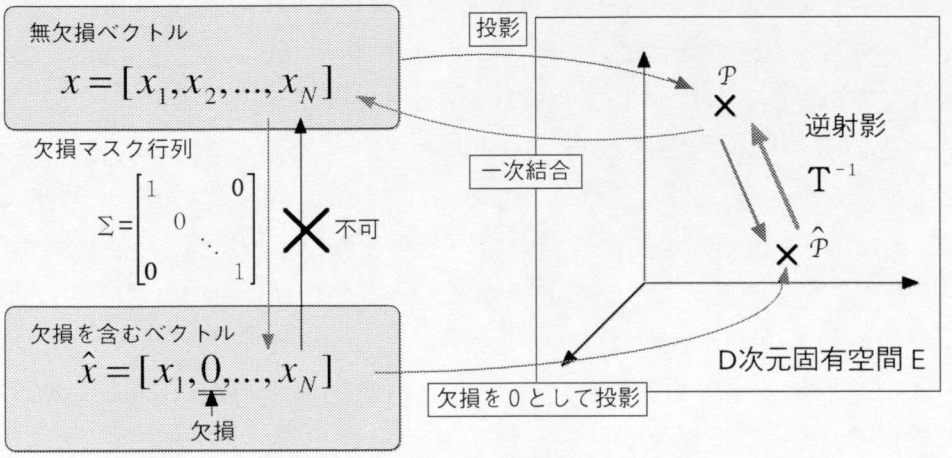



図4-8　BPLP を用いた固有空間照合法

$$\hat{\boldsymbol{x}}=\left[x_1, x_2,\cdots,0,\cdots,0,\cdots, x_N\right]^T$$

$$=\begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & \mathbf{0} & \\ & & \ddots & & & & & \\ & & & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & & \\ & \mathbf{0} & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}\left[x_1, x_2,\cdots, x_N\right]^T=\boldsymbol{\Sigma x} \tag{4-1}$$

このとき、欠損画素がない画像ベクトルは仮想学習により得られた学習サンプルに相当する。また式(4-1)において$\boldsymbol{\Sigma}$は欠損画素の有無をしめす行列であり、画素欠損マスクと呼ぶ。この画像ベクトルは仮想学習により得られた固有空間$\boldsymbol{E}$では、

$$\hat{\boldsymbol{p}}=\left[\hat{p}_1, \hat{p}_2,\cdots, \hat{p}_D\right]^T=\boldsymbol{E}^T\boldsymbol{\Sigma x} \tag{4-2}$$

へ投影される。このとき、(4-1)式より欠損画素を含む画像の投影点と欠損画素がない画像の投影点の対応関係は、

$$\begin{aligned}&\hat{\boldsymbol{p}}=\boldsymbol{E}^T\boldsymbol{\Sigma x}\cong\boldsymbol{E}^T\boldsymbol{\Sigma E p}=\boldsymbol{T p}\\&\boldsymbol{T}=\boldsymbol{E}^T\boldsymbol{\Sigma E}\end{aligned} \tag{4-3}$$

で示される。この変換行列Tの逆行列を求めることにより、欠損画素を含む画像の近似的な投影を求めることができる。式(4-3)では、欠損画素を含む画像を学習で得られた固有ベクトルと学習サンプルの投影点から近似的に学習サンプルを復元していることに相当する。また、この近似は固有空間の累積寄与率が十分大きいときに成立する。

この式(4-3)で示される逆行列は、学習により生成された固有空間の次元数（生成次元数）を$D$とするとき、$D\times D$次元の対称行列となる。このため、計算コストは学習サンプル数に依存しない。また、この逆行列は欠損画素がないときは単位行列となる。逆行列が存在するためには、変換行列Tがフルランクである必要があり、またこれは欠損画素の要素を0とした固有ベクトル列のランクが固有空間の生成次元数$D$となることを意味する。もしこの条件を満足するなら、24×24画素の学習サンプルに対して20次元の固有空間を作成した場合、画像の96%程度が欠損しても固有空間照合が可能となる。実際には学習サンプルで変化しない領域などもある。また、計測誤差による誤差もあるため、90%程度の画素が欠損すると安定した照合は難しいが、実験において70%程度の欠損画素が生じた画像に対しても比較的安定した固有空間照合が行なえることが確認されている。

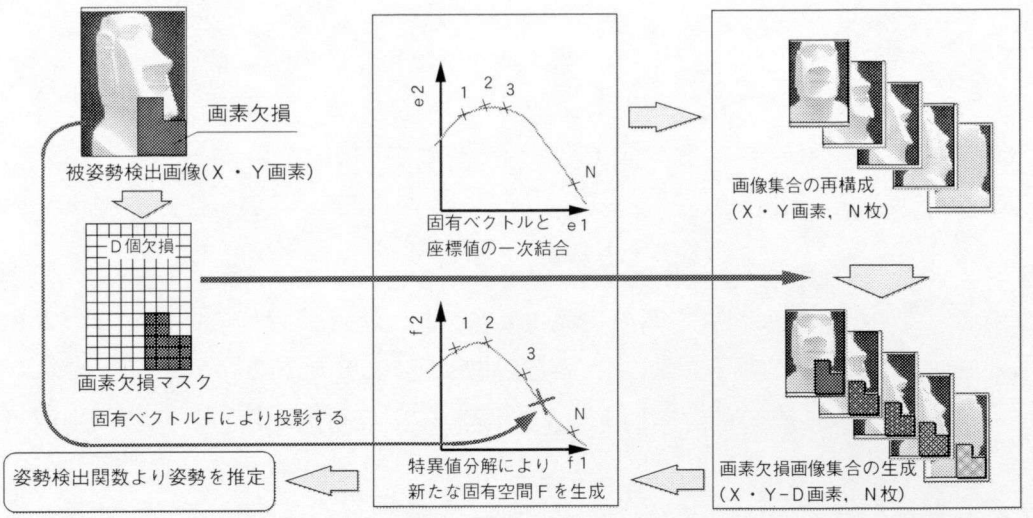



図4-9　ARLPによる固有空間の再構築

### 4.2.2　固有空間の再構築による固有空間照合法（ARLP）

BPLPでは、欠損画素を含む画像の投影点から欠損画素がない画像の投影点への近似的な逆投影を用いた。この投影では欠損画素に平均値を代入し、欠損画素を考慮していない固有空間$\boldsymbol{E}$に投影したうえで補正をかけていることと等しい。このとき、欠損した画素が固有空間生成において特徴的な画素であった場合を考えると、その画素を除外して生成した固有空間と欠損画素がないときの固有空間ではそれぞれの画素に対する重みが変化する。このことより、欠損画素を含む画像の扱いとしては、欠損パターンに応じて固有空間を生成することが望ましい。そこで、学習サンプルを学習結果から近似的に復元し、欠損画素となった要素を用いない固有空間を再構築する方法を考える。

式(4-1)で示した欠損画素マスク行列を用いることにより、欠損画素を考慮したときの分散行列は式(4-4)のように表現される。

$$\widehat{\boldsymbol{Q}}=\boldsymbol{\Sigma}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma} \tag{4-4}$$

この分散行列を対角化することにより固有空間が再構築されるが、一般的に画像ベクトルの次元$N$の方が学習サンプル数Rより大きくなるため画像列のランク

$$rank(\boldsymbol{X})\leq R<N \tag{4-5}$$

を考慮し、特異値分解より固有空間を再構築する。ここで、特異値分解を

$$\begin{aligned}&\boldsymbol{\Sigma X}=\boldsymbol{USV}^T\\&\boldsymbol{U}_{N\times D}\quad \boldsymbol{U}^T\boldsymbol{U}=\boldsymbol{I}_{D\times D}\\&\boldsymbol{V}^T{}_{D\times R}\quad \boldsymbol{V}^T\boldsymbol{V}=\boldsymbol{I}_{D\times D}\\&\boldsymbol{S}_{D\times D}\quad SingularValue\end{aligned} \tag{4-6}$$

とすると、この結果より再構築した固有空間の固有ベクトル列$\boldsymbol{F}$と投影点列は

$$\begin{aligned}\boldsymbol{F}&=\left[\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_D\right]=\boldsymbol{U}\\ \hat{\boldsymbol{P}}&=[\hat{\boldsymbol{p}}_1,\hat{\boldsymbol{p}}_2,\cdots,\hat{\boldsymbol{p}}_R]=\boldsymbol{S}\boldsymbol{V}^T\end{aligned}\tag{4-7}$$

で示される。ただし、$D$は固有空間の生成次元数、$\boldsymbol{S}$は特異値行列である。このように欠損画素となった要素を0として再構築した固有空間$\boldsymbol{F}$は、欠損画素を除くベクトルが直交基底となっている。そのため、この固有空間上に欠損画素を含む画像を投影すれば欠損画素に依存しない固有空間照合が可能となる。以上のARLPの流れを図4-9に示す。この手法では、それぞれの欠損画素パターンに応じて最適な固有空間で照合を行なうことができるため、精度の高い固有空間照合が期待できる。しかし、特異値分解で対角化する分散行列の大きさは*学習サンプル数×学習サンプル数*となるため、QR法などの対角化手法の計算コストは行列の行数（または列数）の二乗に比例するため、ARLPの計算コストは学習サンプル数の二乗に比例して増加する。そのため、学習サンプル数が多いときには効率的でないという欠点がある。

## 4.3　欠損画素を含む画像の固有空間照合実験

### 4.3.1　ARLPとBPLPの比較シミュレーション

第3章の仮想学習で得られた固有空間（学習サンプル数256枚=認識対象4体×姿勢変化64姿勢、固有空間の生成次元数=16次元）を用い、4.2節で示したBPLPとARLPの精度と計算コストを計算機上のシミュレーションにより比較した。また、シミュレーションにおいて認識に用いる固有空間の次元数（認識次元数）の変化に対する両手法の精度も比較した。

シミュレーションでは仮想学習で用いた距離画像生成手法により、認識対象4体に対しそれぞれ3deg.おきに距離画像を生成し、生成された距離画像に対して欠損画素をランダムに与え

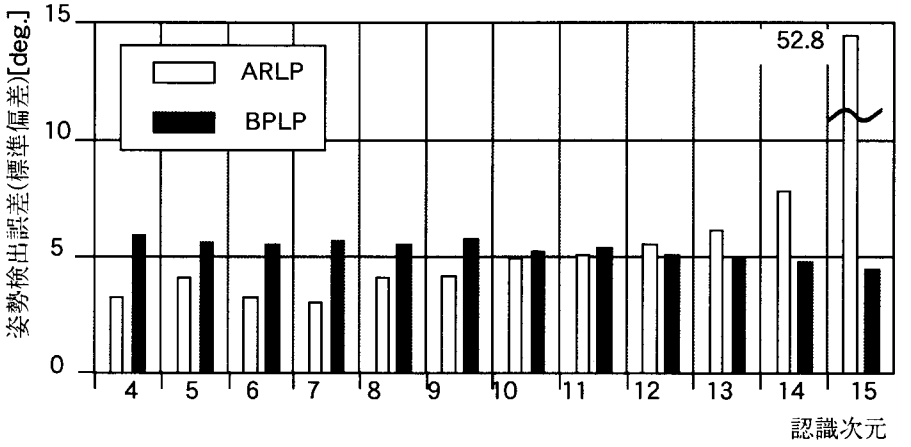



図4-10　ARLPとBPLPによる姿勢検出結果

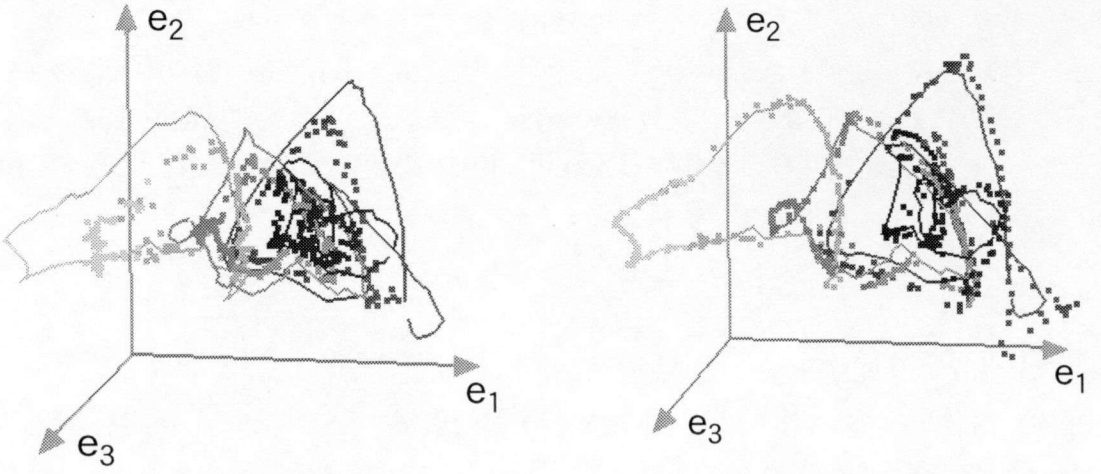
表4-1　ARLPとBPLPの比較

| 手法 | 計算コスト | 精度 | 次元設定 |
|---|---|---|---|
| BPLP | 低い（固有空間の次元に依存） | △ | 容易 |
| ARLP | 高い（学習サンプル数に依存） | ◎ | 困難 |

た。ただし欠損画素数は10%から30%とした。このような距離画像に対しBPLP、ARLPによる姿勢検出を行ない認識次元数4次元から15次元について両手法で姿勢検出の二乗平均誤差を求めた。この結果を図4-10に示す。図4-10を見ると、BPLPでは全域に渡り安定した結果が得られており、全域での平均5.30deg、最小4.48degであった。これに対し、ARLPでは認識次元数を10次元以上としたとき、計算誤差により姿勢検出誤差が増大していくが、認識次元数が4次元から11次元ではBPLPよりもARLPのほうが誤差が小さい。また、ARLPでは認識次元数を7次元としたとき誤差が3.04degで最小となる。以上の結果より、適切な認識次元数を設定すればARLPの方がBPLPよりも高精度な固有空間照合ができることが分かった。二手法において計算コストを比較すると、BPLPでは学習サンプル数に関わらず計算コストは16×16次元の逆行列の計算量である。これに対しARLPでは学習サンプル数に依存し、全対象物体で再構築を行なうと256×256次元の対称行列の対角化の計算量が必要となる。そのため計算コストはARLPの方がBPLPよりも圧倒的に大きくなる。表4-1に両手法の利点欠点をまとめる。

### 4.3.2　BPLPによる逆射影

ARLPでは欠損パターンに応じて固有空間を再構築しているため、欠損パターンごとに異なる固有空間で照合を行なっている。そのため、固有空間上の挙動を解析することは困難であり、4.3.1のように物体認識や姿勢検出の結果より評価するのが妥当である。しかし、BPLPでは学




(a)　欠損画素を含む画像の投影　　(b)　逆射影結果

図4-11　BPLPによる逆射影

習時に得られた固有空間上での逆射影を行っているため、学習で得られた軌跡と比較することで、手法の有効性を確認することができる。

図4-11にBPLP逆射影による投影点の変化を示す。図4-11(a)はBPLPの計算において欠損画素を0として投影したときの結果で、4-11(b)はその後逆射影を行なった結果である。また、このとき欠損画素数は全画素の30%とした。距離画像の固有空間生成では平均ベクトルを引いているため、欠損画素を0とすることは欠損画素に平均画像の画素値を代入していること等価である。そのため、図4-11(a)は4.1.2で述べた、欠損画素に平均画像の画素値を代入して固有空間照合を行なう方法を意味している。このように欠損画素に単に平均画像の値を用いると、投影点が原点方向に投影されるため、中心部の軌跡が重なっているところでは誤認識する恐れがある。また、このような誤差が生じると正確な姿勢検出も期待できなくなる。しかし、BPLPを用いて投影点の補正を行なうことにより、図4-11(b)のように投影点が学習サンプルの軌跡に近づき、欠損画素による投影誤差がキャンセルされる。このようにBPLPを用いることにより、欠損画素に依存しない固有空間照合が可能となる。

### 4.3.3　欠損画素を含む画像に対する物体認識・姿勢検出

4.3.1の結果から考えれば、物体認識・姿勢検出の全てにおいて欠損画素問題をARLPで解決すれば精度の高い固有空間照合を行なうことができると言える。しかし、ARLPで行っている固有空間の再構築は学習サンプル数の二乗に比例するため、物体認識のように多数の学習サンプルに対して行うことを考えると計算コストが膨大になり現実的でない。また、BPLPはARLPに比べ認識精度は劣るものの、計算コストは学習サンプル数に依存しないという利点がある。

物体認識では、計測画像の投影点と学習サンプルにより得られた投影点との正確な投影は必要とせず、どの軌跡と近いかだけを判断できればよいため、物体認識を行なう際にはBPLPを用いたほうが効率的である。これに対して姿勢検出では、計測画像の投影点がどの学習サンプルと近いかを正確に見つけることが必要である。もし、物体認識の結果により認識対象が特定されていれば、ARLPを用いる際に認識された一物体の学習サンプルだけについてARLPを用いれば、比較的少ない計算コストで姿勢検出を行なうことができる。

上記の計算コストと要求される精度の関係を考慮し、第3章の実験では物体認識をBPLPにより行い、その結果より認識された一物体の学習サンプルだけについてARLPを用いて姿勢検出を行なった。また、第3章の実験でのARLPとBPLPの認識次元数（ARLP:7次元、BPLP:7次元）は本章のシミュレーション結果によるものである。

## 4.4　BPLPの応用

BPLPは欠損画素が含まれる画像を固有空間上で逆射影を行うことにより、欠損画素が生じていない画像の投影点を求めることができた。ここではBPLPの逆射影を応用し、欠損したデータを補間する方法について述べる。また濃淡画像を用いた実験を示し、その有効性について検討する。

### 4.4.1 BPLP を用いた欠損データの補間

固有空間法は学習により得られた固有空間を用いて、学習サンプルを固有空間上の軌跡やクラスタで表現することで認識を行う手法である。このとき、固有ベクトルは学習サンプルや計測された画像を投影するための基底ベクトルとなっている。この投影は、学習で得られた固有ベクトル列を$\boldsymbol{E}$、画像ベクトルを$x$、投影点を$p$とすると、

$$\boldsymbol{p} = \boldsymbol{E}^T \boldsymbol{x} \tag{4-8}$$

となる。また、一次結合を用いれば画像ベクトルは

$$\boldsymbol{x} \cong \boldsymbol{E}\,\boldsymbol{p} \tag{4-9}$$

で表される。この投影において基底ベクトルの数は固有空間の生成次元数に相当し、通常画像ベクトルの次元数よりも十分小さい。そのため、一次結合による画像ベクトルの表現は近似的なものとなるが、固有空間上での分散の自由度が$D$程度であるならこの一次結合は妥当といえる。また、この近似は固有空間生成で得られる累積寄与率により、どの程度の画質の画像を復元できるかという補間の妥当性の目安を得ることができる。

ARLPでは欠損画素に依存しない固有空間照合を行うために欠損画素となった要素を用いない固有空間を再構築しているため、一次結合により学習サンプルは復元できない。しかし、BPLPでは学習で得られた固有空間の上で変換を行うため、欠損画素を含む画像を式(4-3)に示す変換行列を用いて BPLP により逆射影し、その後一次結合を行えば

$$\boldsymbol{x} \cong \boldsymbol{E}\,\boldsymbol{T}^{-1}\boldsymbol{E}^T\hat{\boldsymbol{x}} \tag{4-10}$$

となり、欠損画素を含む画像から学習サンプルを近似的に復元することができる。ただし、式(4-10)において、$\hat{x}$は欠損画素を含む画像である。

この式により画像を復元した後、欠損画素を生じていない部分と合成することにより欠損画素の補間を行うことができる(図 4-12)。

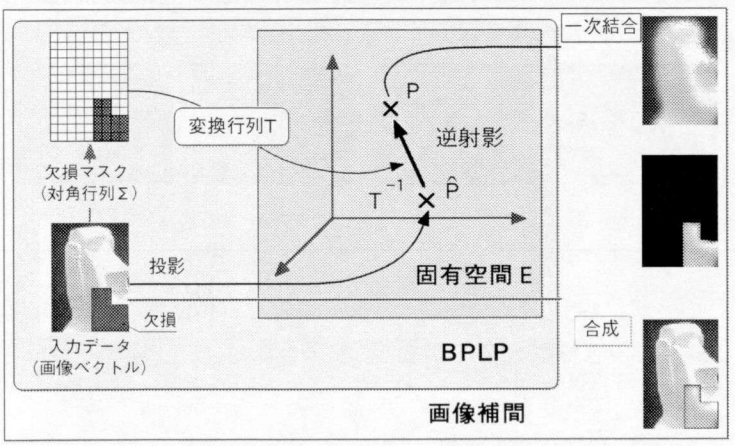
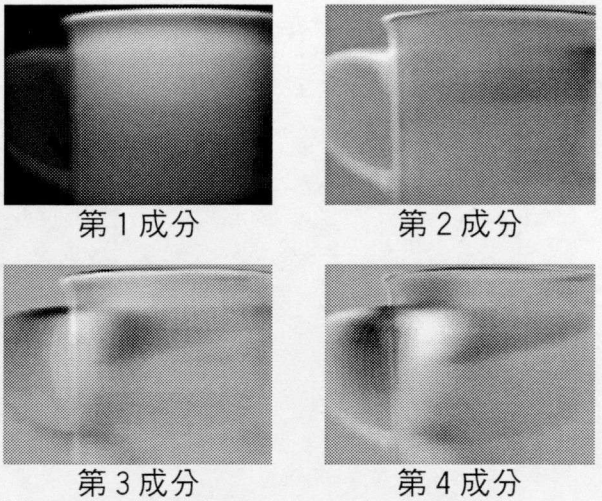
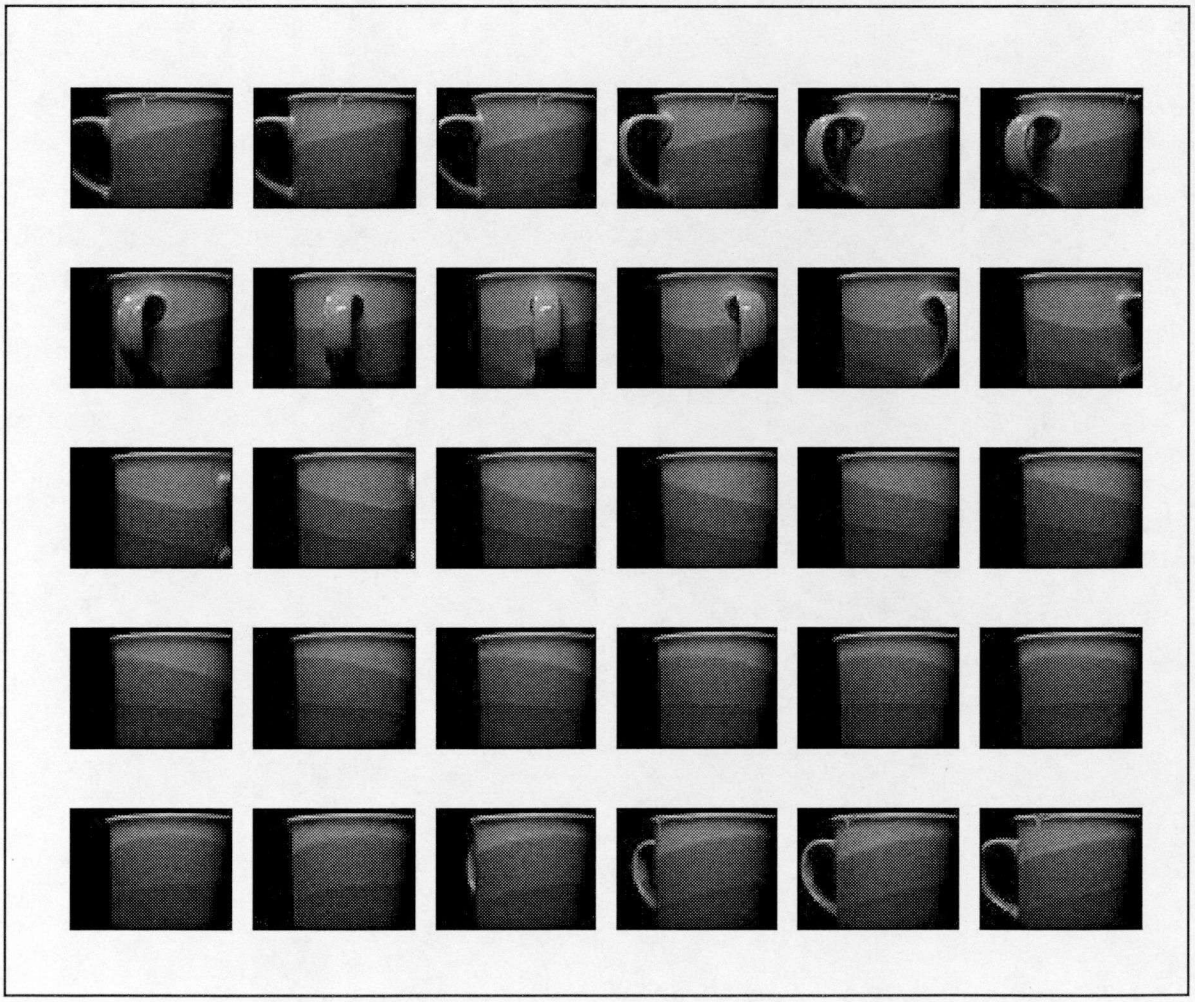



図 4-12　BPLP を用いた欠損データの補間法

(a)　対象物体




(c)　固有ベクトル



(b)　学習サンプル

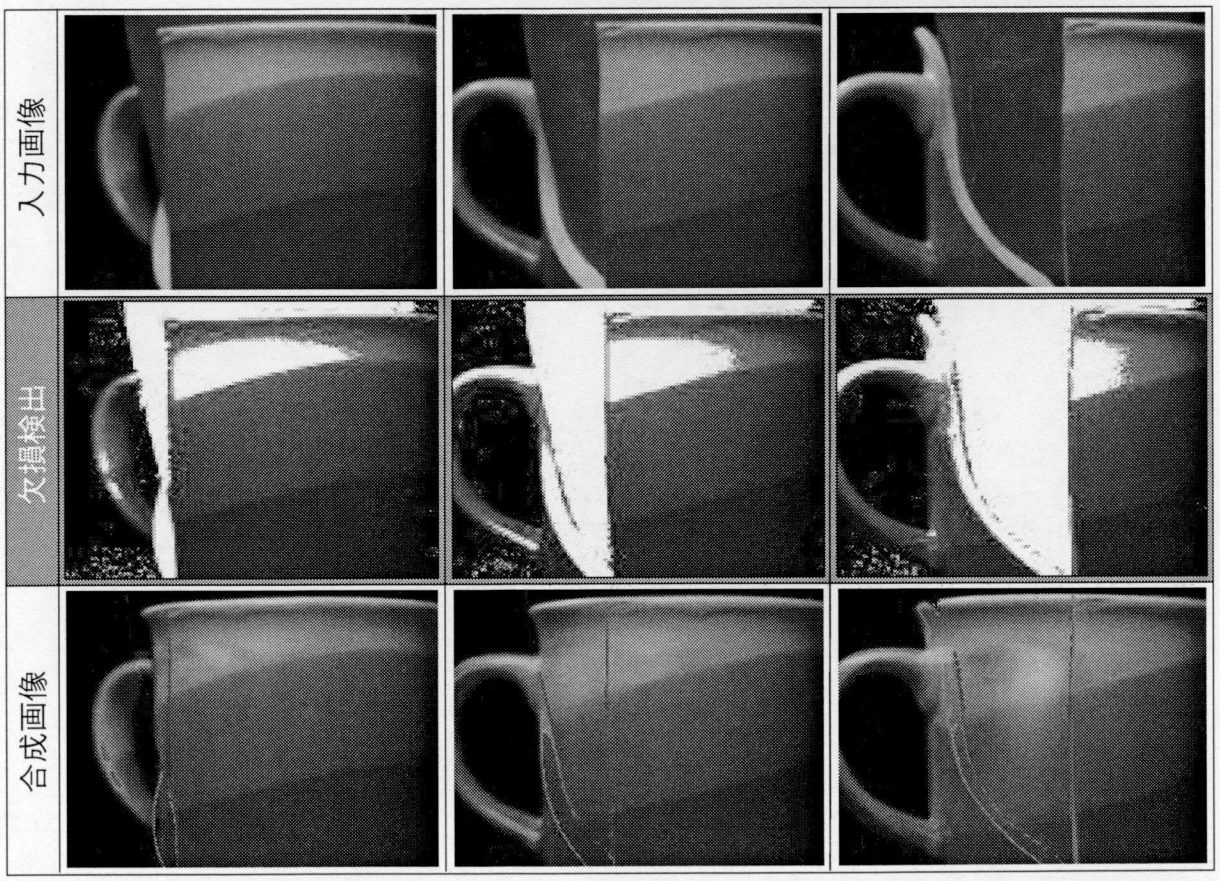
図4-13　対象物体と学習サンプルおよび固有ベクトル

### 4.4.2　濃淡画像の補間実験

BPLPは距離画像の固有空間法の欠損画素問題のために考案した手法であるが、距離画像のみならずベクトル化されたデータの固有空間照合ならば、どのようなデータについても適応可能である。ここでは、濃淡画像の固有空間法についてBPLPを用いた画像の補間について実験を行なった。

実験では図4-13(a)に示すコップの様々な姿勢で得られる濃淡画像（図4-13(b)）を学習した。

(a)　補間対象物体




(b)　補間結果

図 4-14　補間対象と補間結果

この学習サンプルは鉛直軸による回転での見た目の変化であり、固有空間生成では60枚の学習サンプルを用いており、固有空間の生成次元数は累積寄与率70%を目安として20次元とした。学習により生成された固有ベクトルを図4-13(c)に示す。

以上の学習を用いて図4-14(a)のように割れたコップのかけた個所を欠損としてBPLPを用いた補間で画像を復元した。割れた個所の検出は様々な方法が考えられるが、ここでは簡単にブルーバックと輝度値により判定した。欠損画素の検出結果と、画像補間結果を図4-14(b)に示す。補間結果を見ると、多少ぼけた画像となって補間されていることがわかる。これは、BPLPにおいて変換行列$T$や、一次結合による画像生成が固有空間の累積寄与率に依存するためである。もし、固有空間の累積寄与率が100%であれば、式(4-10)の一次結合により完全な画像が生成が出来ることになる。そのため、累積寄与率を目安に固有空間の生成次元数を決定すれば、補間精度を変化させることができる。

### 4.4.3　BPLPを用いた欠損データの補間考察

一般的に用いられる固有空間法による認識や姿勢検出では、学習結果として固有ベクトルと同時に学習サンプルの投影点列を必要とする。このとき、固有ベクトルは固有空間へ投影するときの単なる基底ベクトルであり、学習サンプルの投影点列の方を辞書（＝知識）と考える傾向が強い。これに対して、BPLPによる欠損データの補間法では固有空間上で照合を行わないため、学習結果として学習サンプルの投影点列を必要としない（図4-15）。また、固有空間への投影は固有ベクトルを用いた直交変換である。一般的によく知られている直交変換としてはフーリエ変換が挙げられる。どちらも同じ直交基底を用いた変換であるので、この節で示した補間において基底ベクトルとして固有ベクトルの代わりにsin、cos関数を用いることもできる。しかし、基底ベクトルにsinやcosを用いると、画像の復元性が悪くなるため固有空間を用いた場合よりも遥かに多くの基底を必要とする。なぜならばsinやcosが汎用な基底であるのに対して、固有ベクトルが学習サンプルに特化した基底ベクトルであるからである。このとき、固有ベクトルが学習サンプルに特化しているということは、学習が行われたということに他ならない。そのため、固有ベクトルは学習サンプルに対する知識の1つと考えることができる。

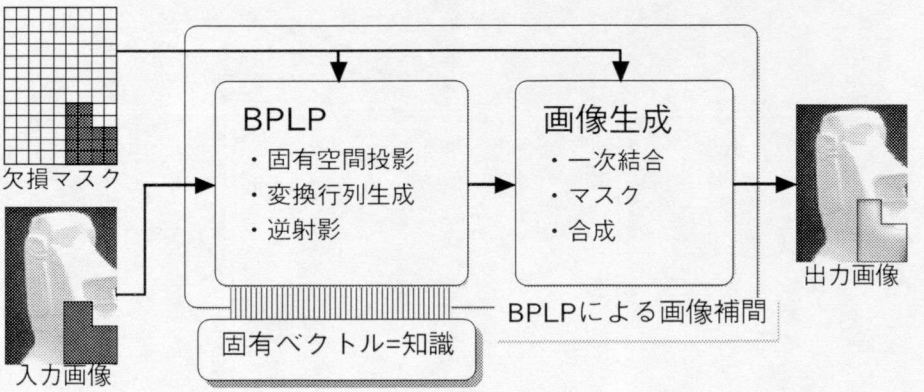



図4-15　BPLPを用いた欠損データの補間ダイアグラム

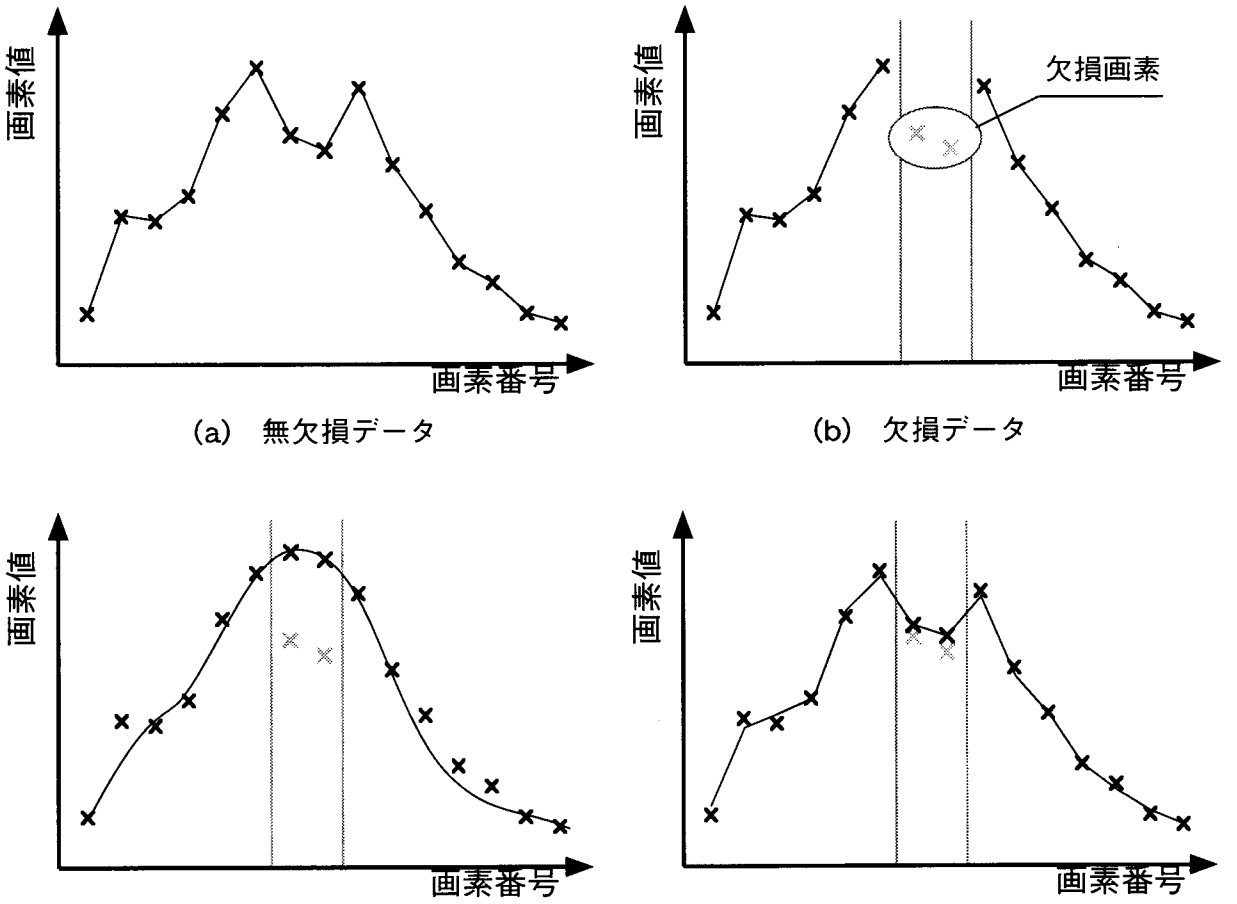



図 4-16 汎用的な補間と BPLP を用いた補間

また、固有空間法では、連続的に変化する画像を用いれば学習サンプル間の未学習画像を表現できる性質がある。そのうえ、BPLPでは画素欠損を補間することができる。このとき画像面を空間、姿勢などのパラメータ変化を時間と考えれば、BPLPを用いた欠損データの補間は固有ベクトルを知識とした時空間補間法と考えることができる。

### 4.4.4 他の補間法との比較

一般的に用いられる汎用的な補間法としては、多項式を用いた線形補間や$N$次補間、スプライン関数を用いたスプライン補間などが挙げられる。これらの補間法は$N$次関数やスプライン関数に基づく補間法であり、もし被補間データがこれらのモデルに従うことが理論的に示されるのであれば妥当な補間である。しかし、画像の補間などは対象とする画像に依存するため、汎用的なモデルは存在しない。BPLPを用いた欠損データの補間は、これらの汎用的な補間法とは異なり学習を必要とする欠点がある。しかし、この学習は被補間データの性質に特化するという意味があり、これにより汎用的な補間法にはない補間が可能となる。例えば、図4-16(a)に示すデータに対して図4-16(b)のような欠損を補間するとき、スプライン関数や多項式を用いた場合図4-16(c)のように補間される。これに対し、BPLPを用いた欠損データの補間では学習サンプルより得られる固有ベクトルを補間の規則としているため、図4-16(d)のように学習サン

プルの性質を反映した補間を行なうことができる。このような学習サンプルに特化された補間を行なうことができることは、汎用的な補間法と異なる所であり、BPLPを用いた欠損データの補間の利点である。また、固有ベクトルのかわりにsinやcos関数を基底ベクトルとすれば汎用的な補間法となり、これはFourier変換を用いた補間となる。

学習サンプルを与えて画像を認識・復元する方法としては、連想記憶を用いた方法がある[Itou 94]。伊藤らは反応拡散方程式を用いた自己想起型連想記憶を提案し、実験において5枚の学習サンプルを与え、テストパターンとして学習サンプルの画像の一部分を与えて認識を行ない、画像を復元している。BPLPを用いた補間法とは方法論は異なるが、

*1)　学習サンプルを用いる*
*2)　画像の一部を与えて画像を復元する*
*3)　出力画像は学習サンプルから得られる画像群の一次結合で生成される*

という点で類似することが多い。また、連想記憶では高周波ノイズを加えた画像を入力として与えると、ノイズを含まない学習サンプルに収束することが一般的に知られている。固有空間法では固有ベクトルによる投影が一種のローパスフィルタの効果を持つため、やはり高周波ノイズに対して安定している。このような所も固有空間法と連想記憶の類似する所である。異なる点としては、

反応拡散方程式を用いた自己想起型連想記憶では、
*1)　入力された刺激を学習サンプルで得られた安定点に収束させる*
*2)　自律分散系を用い繰り返し計算をおこなう*

BPLPを用いた補間法では、
*1)　入力された画像を固有空間に投影し逆射影を行なう*
*2)　繰り返し計算を必要としない*

というところが挙げられる。BPLPを用いた補間法ではあらかじめ欠損領域を示す必要があるが、繰り返し計算を必要とせず高速に補間を行なうことができるうえ、収束のためのチューニングなどを必要としないという利点がある。

## 4-5 まとめ

この章では欠損画素に対する固有空間法を説明した。欠損画素に対する解決法としては、欠損画素に特定の値を用いる方法と、欠損画素を照合に用いない方法が考えられる。ここでは欠損画素を用いずに固有空間照合を行うという考えに基づいて、欠損画素に対して固有空間を再構築するAPLPと固有空間上での逆射影を行うBPLPを提案した。ARLPとBPLPを比較すると、認識精度ではARLP、計算コストではBPLPが有利になることがわかった。これらの手法は物体認識・姿勢検出において精度と計算コストを考慮して使い分けることにより効率的でかつ正確な認識を行なうことができると思われる。また、BPLPを応用した欠損データの補間法も提案した。このような補間法は、あらかじめ補間のルールを得るために学習サンプルを必要とするが、補間する画像が学習サンプルと近いものであれば学習サンプルの特性を反映するため、スプラインや多項式近似などの一般的な補間法よりも有効であると思われる。これらARLP、BPLPにより距離画像の固有空間法における欠損画素問題を解決することができた。しかし、距離画像の固有空間法のもう一つの問題である距離画像のオフセット成分による誤認識を解決していない。次の章では、距離画像のオフセット成分が距離画像の固有空間法に及ぼす問題点とその解決法を述べる。また、この解決法を用いて物体認識・姿勢検出と同時に、認識対象の奥行き移動量の検出方法も述べる。

# 第５章　仮想平行光学系と拡張固有空間による奥行移動量の検出

距離画像を用いた固有空間法では、照明条件の変化に対して不変な認識、形状モデルによる仮想的な学習が行うことが出来るという利点がある。しかし、欠損画素や奥行き変化による距離画像のオフセットをどのように扱えば良いかという問題が生じた。第４章ではこれらの問題点のうち欠損画素問題について解決している。ここでは、アピアランスベースの認識において問題となる見た目の大きさ変化について説明し、距離画像の固有空間法での解決手法を述べる。また、物体の奥行き移動により生じる距離画像のオフセット成分より、認識対象の奥行き移動量を検出する手法を述べる。さらに、実際のレンジファインダでの実験からその有効性を検討する。

## 5.1　仮想平行光学レンジファインダ

### 5.1.1　物体の奥行き移動による見た目の大きさ変化

物体の奥行方向の運動をカメラで観測するとき、図5-1のように撮影された画像での見た目の大きさは変化する。カメラでの撮影は、幾何学的に考えると３次元空間から２次元平面への透視変換であり、見た目の大きさが変化するということは通常のカメラ撮影では当然の現象である。しかし、このような現象は固有空間法などの画像照合法に基づく認識では大きな問題となる。たとえば、画像の相関を用いた照合では、図5-2のように撮影された画像における物体のディテールよりもむしろその大きさに依存する傾向が強く、そのため単に画像相関を用いた認識では誤認識となることが多い。このような現象は、画像の類似度を画像ベクトルの距離により評価する場合もやはり同様である。この問題を解決するには、照合時に画像中の見た目の大きさを等しくすればよい。見た目の大きさを等しくする方法としては、

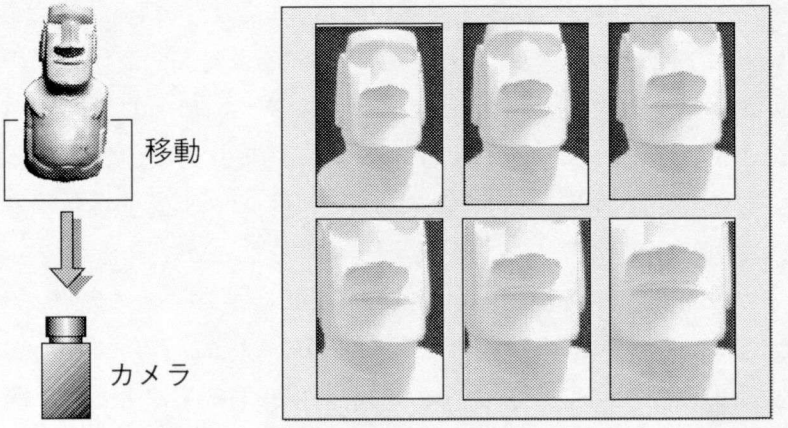



図5-1　画像の見た目の大きさ変化

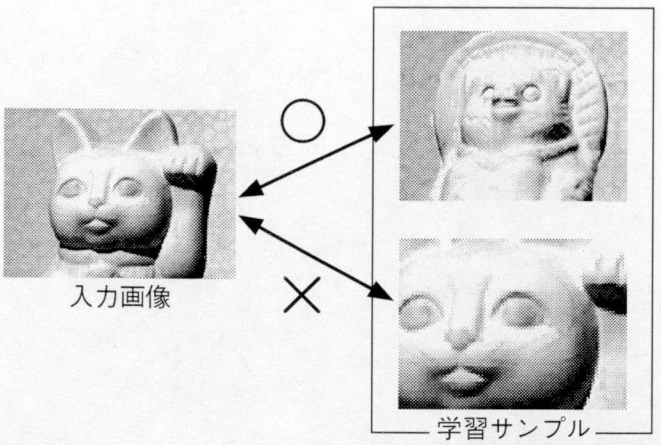
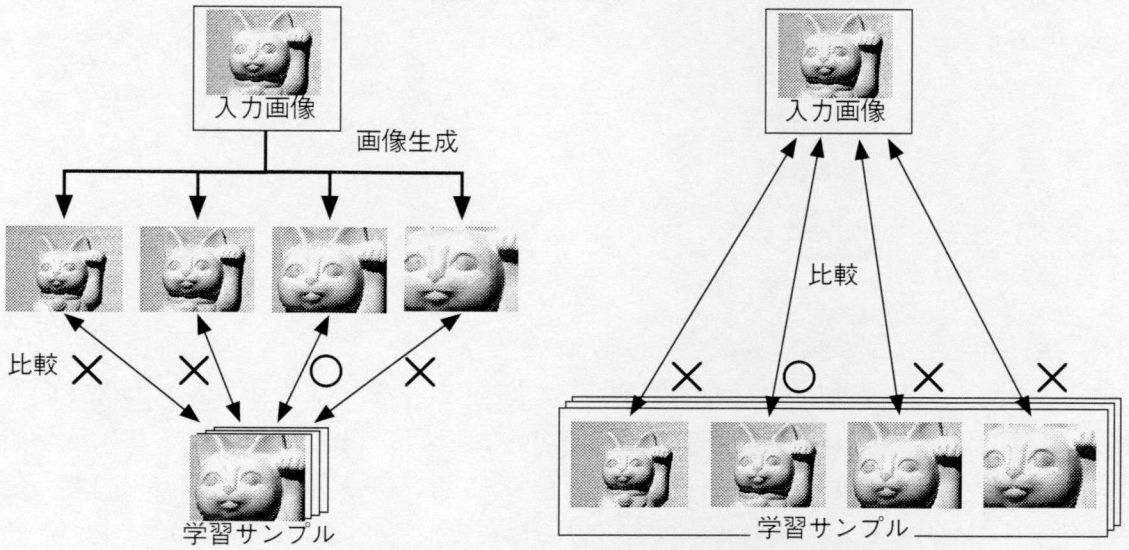



図5-2　見た目の大きさ変化による誤認識




(a)　照合画像の拡大縮小　　(b)　大きさ変化を考慮した学習サンプル

図5-3　見た目の大きさ変化に対する画像照合

*1)　照合時に照合画像を拡大縮小する（図5-3(a)）*

*2)　学習時に拡大縮小した学習サンプルを複数用意する（図5-3(b)）*

のどちらかを用いれば良い。1)の方法は、学習サンプルでは大きさ変化を考慮する必要がなく、学習サンプルの記憶容量を小さく出来るという利点がある。しかし、照合時に入力画像を拡大縮小する必要がある。これに対して2)は入力画像を拡大縮小する必要はなく、照合を高速に行うことが出来る。しかし、学習時に拡大縮小した学習サンプル群を複数用意する必要がある。固有空間法などは学習により学習サンプル群を圧縮することができるので2)の方法を用いて高速な照合を行うのが一般的である[Murase 94b][Horst 98]。このような画像の見た目の大きさ変化に対する画像照合は、方法論自体は簡単ではあるが、実際に行うには学習コストや探索コストが大きくなり実用的ではない。このような問題を解決する方法として、図5-4のようなテレセントリック・ゲージングレンズを用いる方法が考えられる。

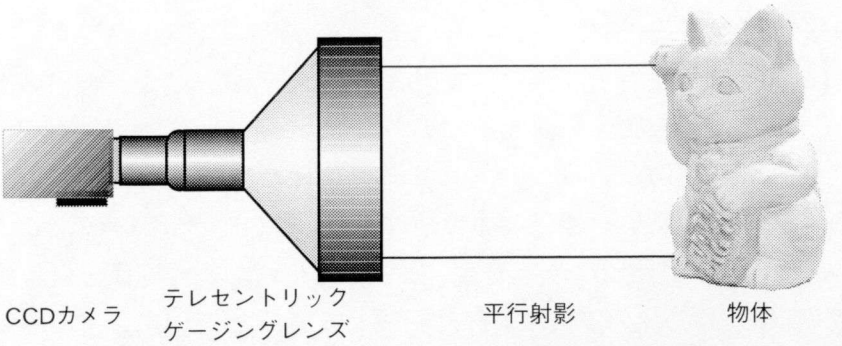



図5-4　テレセントリック・ゲージングレンズ

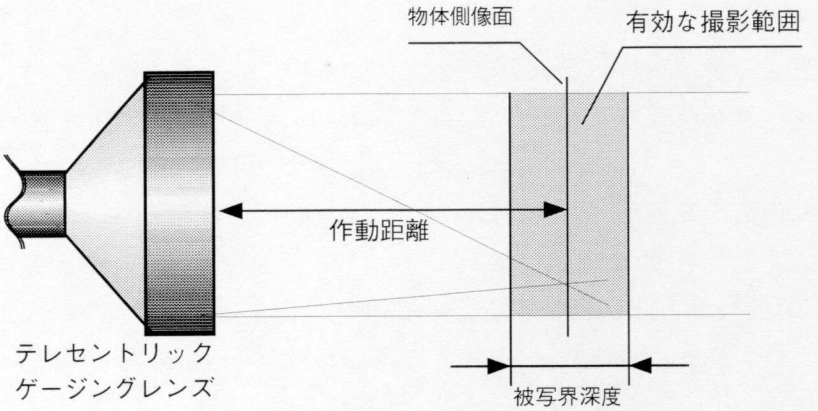



図5-5　作動距離と焦点深度

テレセントリック･ゲージングレンズは平行射影を実現することができる特殊なレンズであり、撮影された画像において見た目の大きさ変化は生じない。しかし、撮影する物体の大きさよりも大きな口径のレンズが必要となるため大掛かりな装置となるという欠点がある。例えば、現在製造されているテレセントリック・ゲージングレンズは、大きなもので口径が145mmのものがあるが、このレンズで撮影できる最大物体サイズは縦横67×89mm程度であり、あまり大きな物体を撮影することができない。このような、テレセントリック・ゲージングレンズはFA分野における小さな部品の精密検査などでは有効な製品ではあるが、環境中に存在する物体の撮影など比較的大きな物体を撮影する用途には適していない。また、テレセントリック・ゲージングレンズは光学的に結像を行うため、図5-5に示すように通常のカメラのレンズと同様に作動距離と被写界深度がある。作動距離とはレンズから物体側で焦点を結ぶ位置（物体側像面）までの距離であり、また被写界深度とは作動距離を中心に像を結ぶ範囲のことである。この作動距離や被写界深度は通常のレンズと同様に、フォーカスにより作動距離を変えることができ、また絞りを絞れば焦点深度を大きくすることができる。しかし、絞りを絞ることやフォーカスを合わせておくということは、照明条件として明るい環境であるとか、認識対象の存在する位置がどの範囲であるかという制限を与えることになる。そのため、見た目の大きさ変化を防げたとしてもフォーカスや強い照明が必要となり新たな問題を生じることになる。距離画像計測方式によっては、レンズの焦点深度による同様の問題が生じる恐れはあるが、ある程度のボケ量であればその後の処理において問題とはならない。また、光学的な結像が不要である光レーダなどを用いればこのような問題は生じない。

### 5.1.2　仮想平行光学 RF

距離画像は計測対象の絶対的な大きさを得ることが出来るが、距離画像の固有空間法では距離画像のうち奥行き方向の成分を画像としているため、濃淡画像を用いた固有空間法と同様に見た目の大きさ変化が問題となる。しかし、距離画像はカメラパラメータが既知であるなら物理的な大きさを表現する世界座標系の座標値$(x,y,z)$を得ることができる。このことを利用して図5-6のように仮想的な平行光学系を考え、一度計測された距離画像を仮想的な平行光学レンジファインダ（仮想平行光学RF）により計算機上で計測しなおすことで画像中の見た目の大きさが変化をなくすことができる。このような平行光学変換は第3章の仮想学習で用いた画像生成法において、

*・計測された距離画像を三角パッチで表現して形状モデルとする。*
*・カメラパラメータを視線方向を奥行き移動方向にあわせた正射影とする。*

とすることにより実現できる。以上により、見た目の大きさ変化が生じない仮想的な撮像系を実現することができるが、距離画像はレンジファインダから物体までの距離を表現する画像であるので、物体が奥行き方向に移動すると距離値が変化するという問題が生じる。距離値の変化を幾何学的に考えると、認識対象の奥行き方向の移動量が全ての画素に均一に加算されるオフセット成分と考えることができる。
この仮想平行光学RFにおいて、対象物体が奥行方向運動を画像ベクトル空間で考えると、奥行き移動を含む画像ベクトル$x$は

$$\begin{aligned} &\boldsymbol{x}=\boldsymbol{x}^{*}+v_Z\,\boldsymbol{e}_v \quad ,N=sizeX\times sizeY \\ &\boldsymbol{x}=col\left(x_1,x_2,\cdots,x_N\right),\ \boldsymbol{e}_v=col\left(\frac{1}{\sqrt{N}},\frac{1}{\sqrt{N}},\cdots,\frac{1}{\sqrt{N}}\right) \end{aligned} \tag{5-1}$$

と表すことができる。ただし、$V_Z$は奥行移動量、$x^*$は奥行移動量がないときの画像ベクトルであり$N$は画像ベクトルの次元（画像サイズ$sizeX\times sizeY$）である。このため対象物体が奥行方向に移動したとき、図5-7(a)に示すように$e_v$方向に対応点を探索するか、もしくは図5-7(b)のように奥行き移動方向による画像ベクトルの変化を直線で表現した投影線を考え、この直線と学習

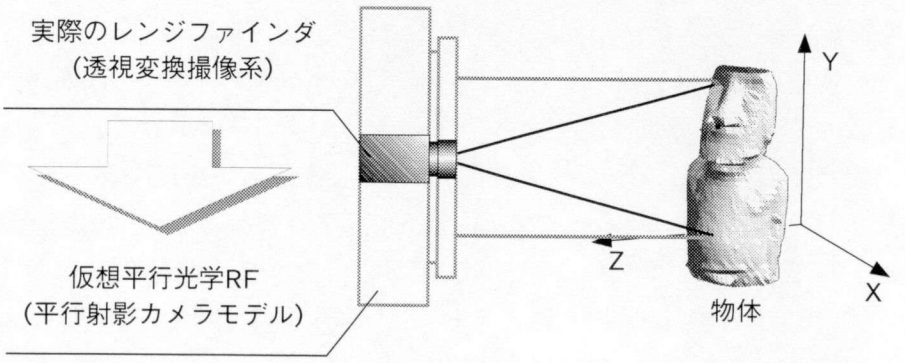
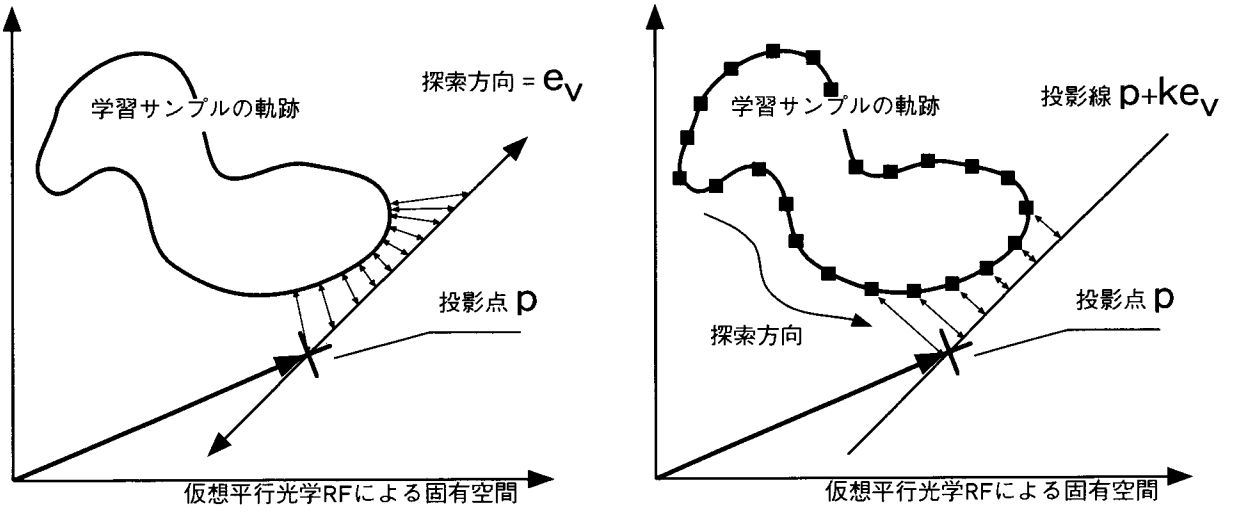



図5-6　仮想平行光学 RF




(a) 探索による方法　　(b) 直線との距離計測

図5-7　認識対象の奥行き移動に対する認識

サンプルの投影点の距離を求めることで物体認識・姿勢検出を行うことができる。また探索により、奥行き移動量$V_z$を検出することもできる。

## 5.2 拡張固有空間

### 5.2.1 奥行き検出に対する問題点

5.1.1で示したように、仮想平行光学RFを用いることにより奥行方向移動に対する画像ベクトル空間上の投影点の変化は直線的な移動となる。固有空間とは、画像ベクトルを表現する画像ベクトル空間を回転・平行移動および縮退して得られる部分空間であるので、固有空間でも同様に直線移動となる。そのため、5.1.2で述べたように照合時に撮影された画像の投影点から$e_v$方向に探索するかもしくは照合画像の投影点を$e_v$方向に通過する直線を考え、その直線と学習サンプルの投影点の距離を求めれば良いことになる。しかしこの探索を幾何学的に考えると、すべての学習サンプルに対して多次元空間中の直線と投影点との距離を求めることになり計算コストが高い。もし、計算コストが問題とならない場合であればこのような方法を用いることも可能であるが、エレガントな方法ではない。

このような考え方のほかに、何らかの方法により認識対象の奥行き方向移動による画像ベクトルの変化分を正規化することにより取り除くことも考えられる。具体的には、画像の平均値又はピーク値を0にする方法である。このような手法により学習・認識時に同様の正規化を行えば、奥行き方向移動による画像ベクトルの変化がキャンセルされるため、正しく認識を行うことができる。しかし、実際のレンジファインダで計測される距離画像は、欠損画素が生じる。第3章で述べたように、この欠損画素は必ずしもレンジファインダのオクルージョンのような幾何学的な関係だけで生じるわけではなく、認識対象表面のテクスチャ成分により欠損が生じることもある。そのため、計測される距離画像のピークのところが欠損したり、偏った欠損に

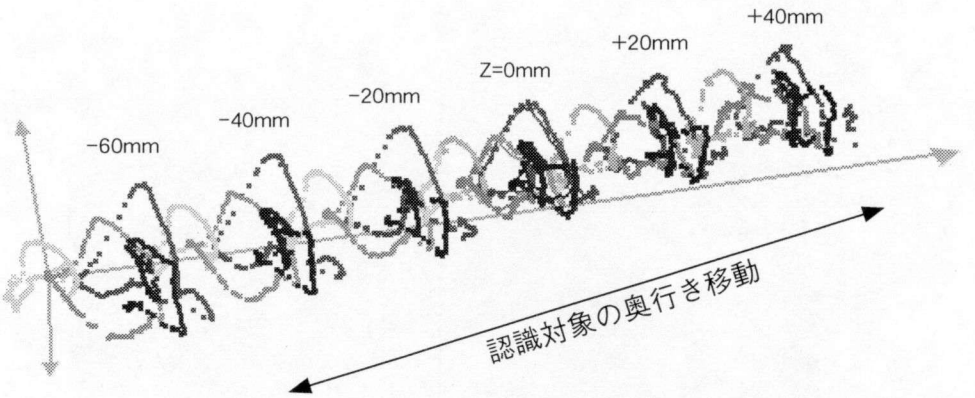



図5-8　奥行き移動に対する投影点の分散

より距離値の平均がずれることは十分考えられる。そのため、何らかの方法により計測された距離画像の定性的な性質から正規化を考える必要が生じるが、撮影された距離画像が何らかの定性的な特徴を持つということは認識対象の形状を制限していることにほかならない。また奥行移動に対する画像ベクトルの分散は図5-8†に示すように、回転等による画像ベクトルの分散に比べ大きくなる傾向にあるので、計測された距離画像の正規化の誤差がわずかでも生じると物体認識を安定して行うのは難しくなる。そこで、このような問題を解決するために直交性を用いることを考えた。

フーリエ変換などでは、変換を行う画像にスパイク状のノイズなどが含まれていても低周波成分のパワーや位相は正しく求めることができる。これは、フーリエ変換の直交基底が、スパイク状のノイズの周波数成分が高く低周波の基底ベクトルと直交しているからである。もし、このような直交性を距離値のオフセット成分に適応できれば正規化などの操作を必要とせず、オフセットを含む距離画像を単に投影するだけで独立化できることになる。また、直交化された奥行き移動の基底ベクトルと固有空間を生成すれば、全体を新たな固有空間として第4章で述べたBPLPやARLPにより欠損画素問題を解決できる。

### 5.2.2　拡張固有空間による奥行き移動成分の直交化

仮想平行光学RFを用いれば認識対象の奥行き方向の移動は、画像ベクトルにおいて単一のベクトル方向の変化として表現される。このような単一ベクトル方向の移動で変化を記述できれば、固有空間のような線形写像でも容易に扱うことが可能となる。なぜなら、もしも固有空間の第1成分をこの奥行き移動を表すベクトルとすることができれば、奥行き移動検出と物体認識・姿勢検出を独立化することができるからである。

この直交独立化ができれば、5.2.1で述べた物体ごとの姿勢変化と、奥行き位置変化での投影点分散の大きさのちがいによる誤認識を防ぐことができる。そのため、ここでは奥行移動のための部分空間と物体認識のための部分空間を直交させた固有空間を生成し、奥行移動量を含む距離画像をそのまま固有空間に投影することにより、物体認識及び姿勢角・奥行移動量検出を同時に行うことを考える。

†図5-8では奥行き方向移動と認識対象の鉛直軸回転で得られた学習サンプル群の固有空間を生成した後に、固有空間上の投影点を3次元で表現している。

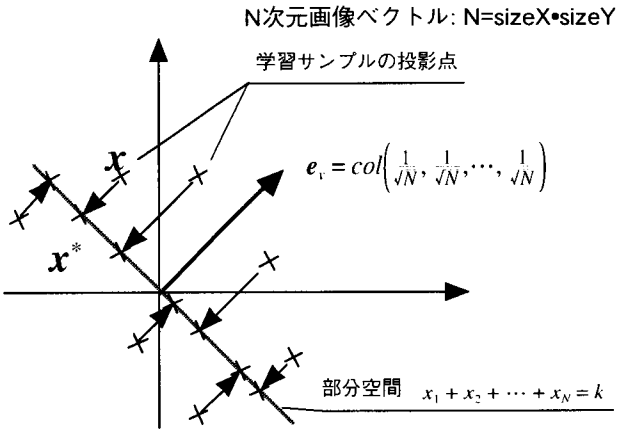



図5-9　奥行き移動不変投影

固有空間を幾何学的に考えれば、固有空間とは学習サンプルから生成された画像ベクトルの分散の大きい方から直交基底を生成した部分空間であるが、画像ベクトルの分散と認識対象の物理的な運動は直接結びつかない。そのため、生成された固有ベクトルに物理的な意味を持たすことは困難である。そこで、KL展開などの固有空間生成法の分散が大きい方向から基底ベクトルを選ぶという固有空間の性質を逆手にとり、学習サンプルの特定方向の分散をなくすことにより、意図的に固有ベクトルの第1成分が相対的な奥行方向の移動量になる固有空間を生成することを考える。

まず、すべての学習サンプルを奥行き移動ベクトルと直交するように変換することを考える。式(5-1)の奥行移動量の基底ベクトルであると直交する部分空間は$N$次元画像ベクトル空間中の超平面

$$x_1 + x_2 + \cdots + x_N = k \tag{5-2}$$

で示される（図5-9）。特にここでは$k$=0とする。学習サンプルをこの超平面に射影すれば、奥行移動量に対する分散が0となる学習サンプルが生成される（式(5-3)）。この射影を画像処理で考えると、一枚の学習サンプルに対して画素平均を0とする操作と等価である。そこで、全ての学習サンプルに対して平均値が0となるように、全ての画素値から一定の値を引いて画像ベクトルを生成する。また、平均値を学習サンプルごとに記憶し、姿勢角検出の結果より奥行方向の絶対位置を求める際の辞書$M$を作成する。

$$\begin{aligned}
&\boldsymbol{x}_{r,s} = col\left(x_1, x_2, \cdots, x_N\right)_{r,s} \\
&\boldsymbol{x}^*_{r,s} = col\left(x_1 - m, x_2 - m, \cdots, x_N - m\right)_{r,s} \\
&m_{r,s} = \frac{1}{N}\sum_{i=1}^{N} x_i \Big|_{r,s} \\
&M = \left[m_{1,1}, m_{2,1}, \cdots, m_{R,1}, \cdots, m_{1,S}, m_{2,S}, \cdots, m_{R,S}\right] \\
&\qquad\qquad r=1,2,\cdots,R \quad s=1,2,\cdots,S
\end{aligned} \tag{5-3}$$

ただし、$R$は回転の刻み数、$S$は認識対象の数である。一枚の画像の平均値を引くことにより得られた画像ベクトル列を生成する。さらに全画像ベクトルの平均ベクトル$c$を引いた画像ベクトル列より分散行列$Q$を生成する。その後分散行列$Q$を対角化し、固有ベクトル列を求める。

$$
\begin{aligned}
&\boldsymbol{c} = \frac{1}{RS}\sum_{r=1}^{R}\sum_{s=1}^{S}\boldsymbol{x}^*_{r,s} \\
&\boldsymbol{X}^* = \left[\boldsymbol{x}^*_{1,1} - c, \boldsymbol{x}^*_{2,1} - c, \cdots, \boldsymbol{x}^*_{R,S} - c\right] \\
&Q = \boldsymbol{X}^*\boldsymbol{X}^{*T} \\
&\boldsymbol{E}^* = \left[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_D\right],\ D < R \times S < N
\end{aligned}
\tag{5-4}
$$

式(5-4)において$D$は固有空間の生成次元で、累積寄与率を参考に決定する。また、固有ベクトル列で学習サンプルを投影して投影点列を求める。物体認識･姿勢角検出ではこの投影点列との比較を行う。学習結果により得られた固有ベクトル列に式(5-1)に示した奥行移動成分を含む距離画像を投影すると、

$$
\begin{aligned}
\boldsymbol{p}_E &= \boldsymbol{E}^{*T}\boldsymbol{x} = \boldsymbol{E}^{*T}\boldsymbol{x}^* + v_Z\boldsymbol{E}^{*T}\boldsymbol{e}_v \\
&= \boldsymbol{E}^{*T}\boldsymbol{x}^* \quad , \boldsymbol{e}_i^T\boldsymbol{e}_v = 0\ , i = 1,2,\cdots,D
\end{aligned}
\tag{5-5}
$$

となり、奥行移動量に対して不変な投影が可能となる。また、奥行移動量の基底ベクトルに式(5-1)を投影すると

$$
\begin{aligned}
\boldsymbol{p}_{ev} &= \boldsymbol{e}_v^T\boldsymbol{x} = \boldsymbol{e}_v^T\boldsymbol{x}^* + v_Z\boldsymbol{e}_v^T\boldsymbol{e}_v \\
&= v_Z \quad , \boldsymbol{e}_v^T\boldsymbol{e}_v = 1
\end{aligned}
\tag{5-6}
$$

となるため、物体や姿勢に依存しない相対的な奥行移動量を検出できる。このため、物体認識･姿勢角検出の結果に対応する辞書$M$の値と相対位置を加えれば奥行方向の位置を得ることができる。これら奥行移動量の基底ベクトルと固有ベクトル列を組合せた固有空間を固有ベクトル列$\boldsymbol{E}$

$$
\boldsymbol{E} = \left[\boldsymbol{e}_v, \boldsymbol{E}^*\right]
\tag{5-7}
$$

で表現し、固有ベクトル列Eで表現される固有空間を拡張固有空間とよぶ。この拡張固有空間により物体認識・姿勢検出と同時に奥行き位置検出を実現することができる。

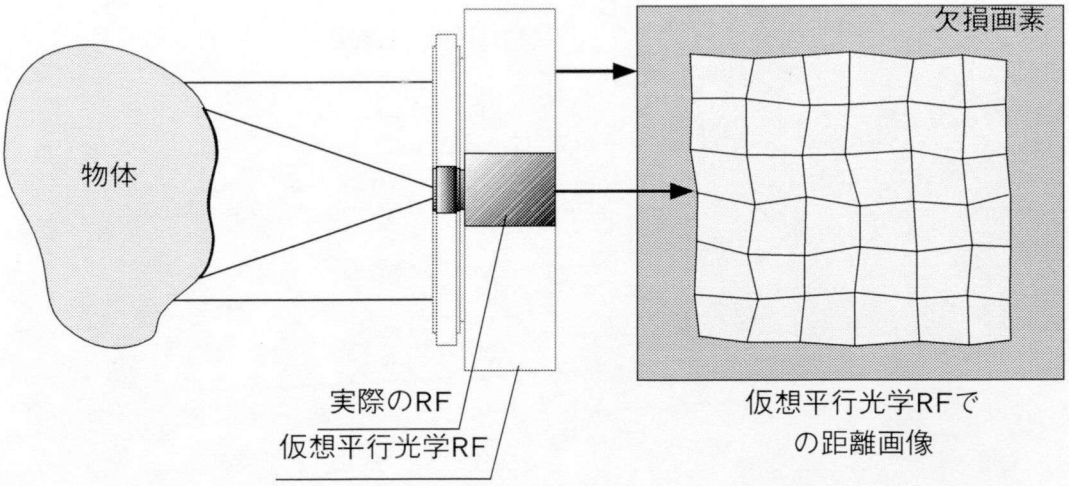



図5-10　仮想平行光学変換で生じる欠損画素

### 5.2.3　拡張固有空間に対するBPLP

拡張固有空間を用いて物体認識・姿勢検出と奥行き位置検出を固有空間上で直交化すれば各パラメータの分散の大きさが極端に異なる場合でも安定した認識が期待できる。しかし、実際のレンジファインダで撮影される距離画像では欠損画素を生じるため、拡張固有空間においてもやはり欠損画素を考慮する必要が生じる。特に姿勢検出、位置検出のようにパラメータ間での分散が極端に異なる場合は大きな問題となる。

拡張固有空間における第1成分は、認識対象の奥行き移動を表現しているため、その他の固有ベクトルと異なり意図的なものとなっている。そのため、従来の単純に分散方向より基底ベクトルを求める固有空間とは若干意味が異なるが、拡張固有空間を構成するベクトル列は全て直交しておりまた正規化されている。そのため、固有ベクトル列の性質としては通常の固有ベクトル列となんら変るところはない。ゆえに、拡張固有空間に対しても第4章で示した欠損画素を含む画像に対する固有空間照合法をそのまま適応できる。4.2.1で示したBPLPをそのまま拡張固有空間へ適応すれば、式(4-3)の変換行列$\boldsymbol{T}$は拡張固有空間では

$$\boldsymbol{T} = \boldsymbol{E}^T \boldsymbol{\Sigma} \boldsymbol{E} = \begin{vmatrix} \boldsymbol{e}_v^T \boldsymbol{\Sigma} \boldsymbol{e}_v & \boldsymbol{e}_v^T \boldsymbol{\Sigma} \boldsymbol{E}^* \\ \boldsymbol{E}^{*T} \boldsymbol{\Sigma} \boldsymbol{e}_v & \boldsymbol{E}^{*T} \boldsymbol{\Sigma} \boldsymbol{E}^* \end{vmatrix} \tag{5-8}$$

となる。この変換行列$T$を用いて拡張固有空間上で逆射影を行なえば、拡張固有空間を用いた物体認識・奥行き位置姿勢検出でも欠損画素問題を解決することができる。欠損画素問題は距離画像計測の時に生じるものだけでなく、仮想平行光学RFでも生じることがある。具体的には、図5-10のように認識対象がカメラの近くにあるような場合、この距離画像を仮想平行光学RFで撮影しなおすと仮想平行光学RFで得られる距離画像の周辺が欠けてしまうことになる。このような欠損画素が生じても、レンジファインダによる計測で生じた欠損画素と同様に扱えばBPLPによりキャンセルすることが可能であり、固有空間照合で問題とはならない。

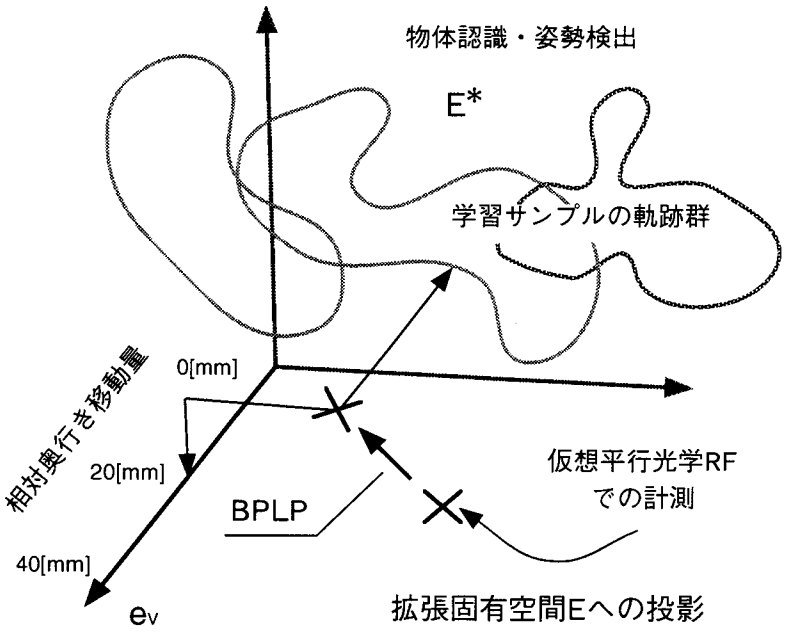


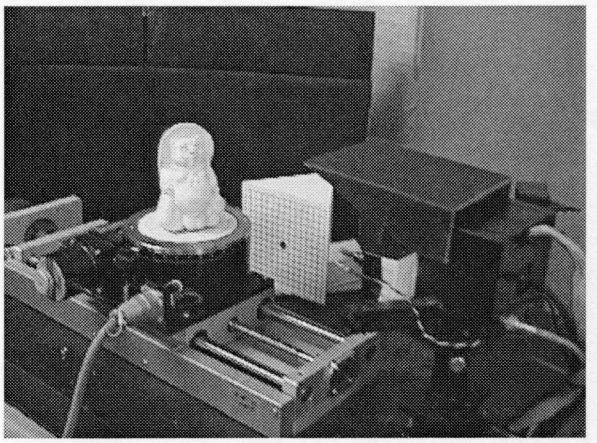

図5-11　仮想平行光学RF・拡張固有空間を用いた認識

### 5.2.4　物体認識・奥行き位置姿勢検出方法

仮想平行光学RFと拡張固有空間を用いた物体認識・奥行き位置姿勢検出において、実際のレンジファインダで生じる欠損画素を考慮すると処理の流れは図5-11のようになる。具体的には仮想平行光学RFにより得られた距離画像に対して、

*1) 平均画像を引き、欠損画素の画素値を0とする。*
*2) 欠損画素マスクよりBPLPの変換行列Tを求める。*
*3) 固有ベクトル列Eにより固有空間へ投影する。*
*4) 変換行列Tを用いて投影点を逆射影する。*
*5) 投影点の成分より相対的な奥行移動量を求める。*
*6) 投影点の成分と学習サンプルの固有空間上の軌跡を比較し物体認識・姿勢角検出を行う。*
*7) 物体認識・姿勢角検出結果より辞書Mを参照し、絶対的な奥行移動量を求める。*

となる。以上の行程により、認識対象が回転及び奥行方向に平行移動したときに撮影された距離画像から、物体認識・奥行き位置姿勢検出ができる。この手法では奥行き検出のために探索や繰り返し計算を必要としてない。そのため、新たに必要となる計算としては仮想平行光学RFで行なう平行光学変換と、辞書Mより奥行き位置の絶対値を求める計算であり、固有空間照合での計算コストは奥行き検出を考慮していないものと同じである。仮想平行光学変換は一度計測された距離画像をポリゴンモデルで表現し、その後平行光学変換するため比較的計算コストは大きい。しかし、形状モデルを用いた仮想学習とは異なり正射影カメラパラメータの一本の視線と交差するポリゴンはただ1つである。そのため、ポリゴンと視線の交点を求める際に



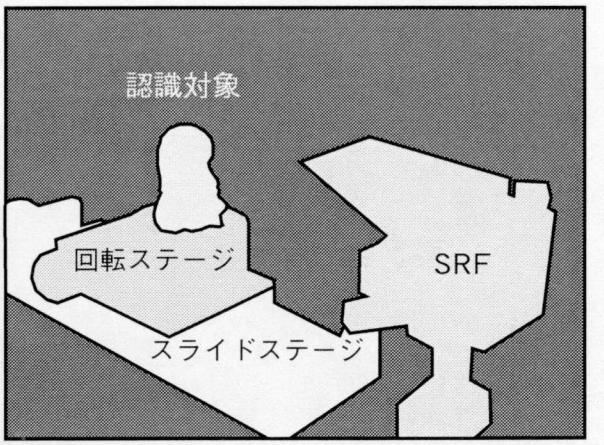



図 5-12　機材構成

交差するポリゴンを1つ見つければ良い。また、ポリゴン縦横の位置関係と、交差する視線の位置関係は対応している。そのため、

*・交差するポリゴンが1つ見つかれば全てのポリゴンについて探索する必要はない。*
*・ある視線で交差するポリゴンは、隣りの視線で交差するポリゴンの周囲にある。*

ということを利用すれば、高速に平行光学変換を行なうことができる。以上の手法でも変換速度が問題となるならば、3.2.2で紹介したレンダリングハードウェアによる距離画像の生成を用いることにより高速化することもできる。

## 5.3　奥行検出実験

### 5.3.1　実験環境

仮想平行光学RFと拡張固有空間を用いた物体認識・奥行き位置姿勢検出を評価するために図 5-12 に示すような実験機材を構成した。実験装置において認識対象を回転ステージにより鉛直軸周りの回転を、スライドステージによりレンジファインダの奥行き方向の平行移動を行なうことができるようになっている。また、回転ステージの基準位置はレンジファインダから300mmとしている。実験環境においてレンジファインダはSRF（画像解像度 24 × 24、奥行分解能 1mm 程度）、認識対象としては、第3章の実験で用いた4体の人形（花っ子、招き猫、たぬき、シーサー（高さ 120mm 程度））を用いた。

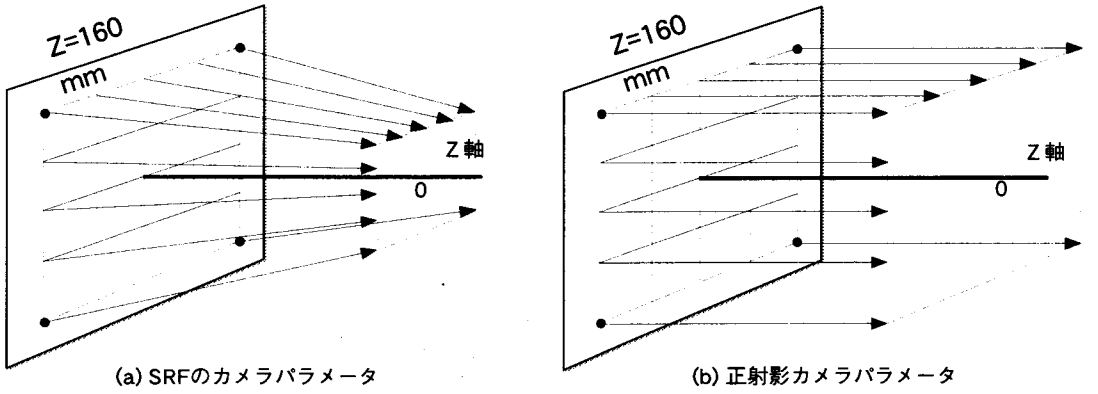



図5-13　正射影カメラパラメータの生成

### 5.3.2　仮想平行光学系

仮想平行光学RFを実現するために実際のレンジファインダのカメラパラメータ

$$\begin{bmatrix} h_s x_s \\ h_s y_s \\ h_s \end{bmatrix} = \begin{vmatrix} -0.71752 & -0.00492 & 0.07275 & 14.30576 \\ 0.00420 & -0.73470 & 0.11068 & 17.35869 \\ 0.00000 & 0.00000 & 0.00726 & 1.00000 \end{vmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{5-9}$$

を用いて正射影のカメラパラメータ

$$\begin{bmatrix} h_s x_s \\ h_s y_s \\ h_s \end{bmatrix} = \begin{vmatrix} -0.33195 & -0.00219 & 0.00000 & 11.99871 \\ 0.00182 & -0.33968 & 0.00000 & 16.21748 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{vmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{5-10}$$

を生成した。ただし、$x_s, y_s$は受光素子上の座標値、$X,Y,Z$は世界座標系の座標値である。式（5-10）を見ると世界座標の$Z$の係数は全て0となっており、奥行き変化により画像の拡大縮小が起こらないことが確認できる。また、この正射影のカメラパラメータは、シリコンレンジファインダのキャリブレーション[Amano 95]で得られたカメラパラメータより距離画像の四隅の点での姿勢方程式と基準平面（奥行き位置160mm）との交点を求め、その4点を通りスライドステージの移動方向に平行な視線となるように生成した（図5-13）。この正射影のカメラパラメータを用いて仮想平行光学RFを実現した。

### 5.3.3　仮想学習結果

奥行き検出実験においても学習サンプル獲得コストを考慮し、仮想学習を用いて学習を行なった。仮想学習は、第3章で示した形状モデルと同様に6方向から撮影して得られた認識対象の形状モデルを用いた。ただし、実際の認識では仮想平行光学RFを用いて認識を行なうため、距離画像生成においてカメラパラメータは式（5-10）で示した正射影のものとした。

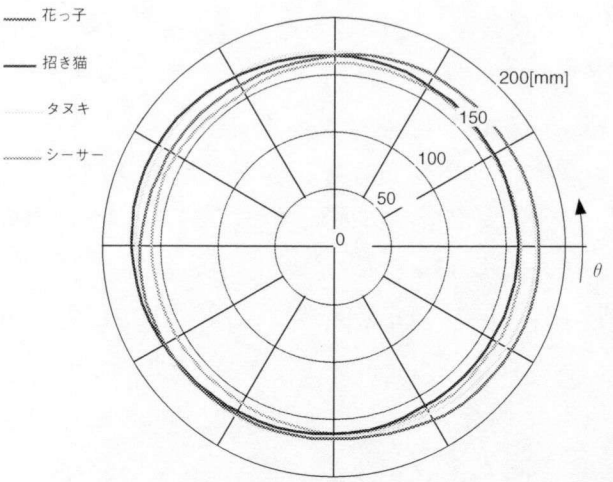



図 5-14　姿勢変化に対する平均距離

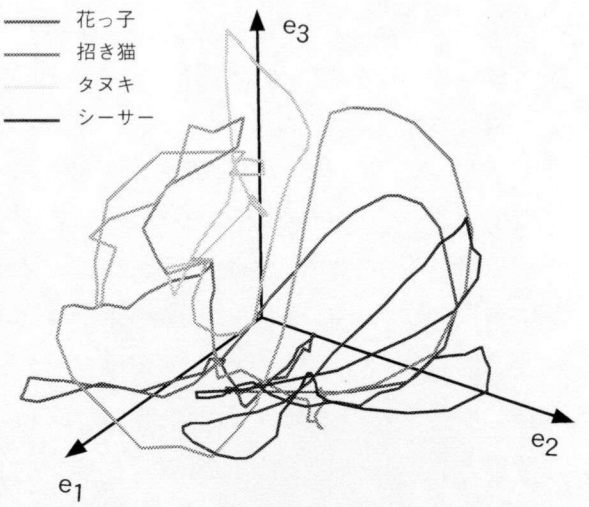
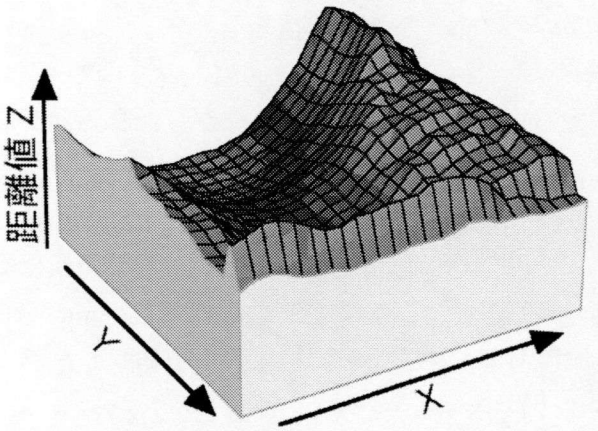
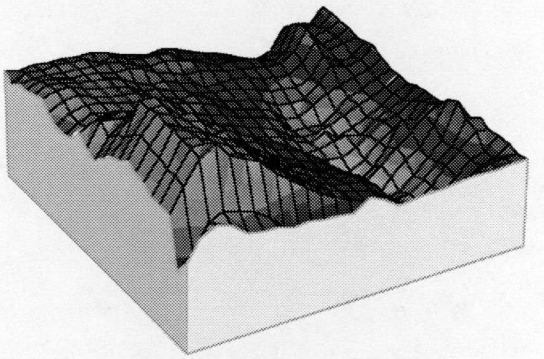


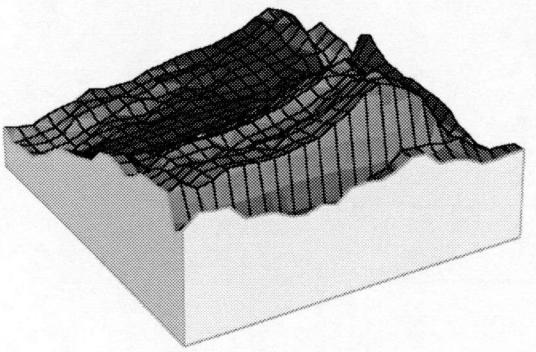

図 5-16　学習サンプルの軌跡



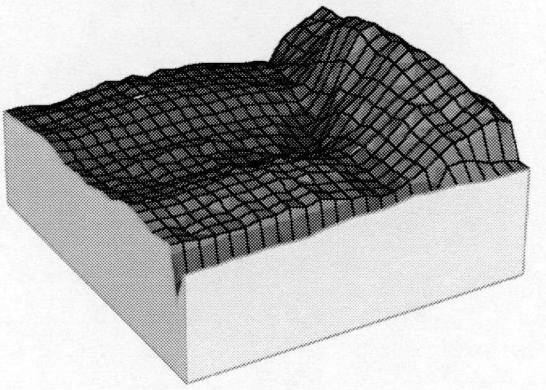

(a) 平均ベクトル



(b) 固有ベクトル e1



(c) 固有ベクトル e2



(d) 固有ベクトル e3

図 5-15　平均ベクトル・固有ベクトル

4体の人形について1物体あたり64ステップで鉛直軸周りで回転させた学習サンプル（4物体×64ステップ=256枚）を用い、式（5-3）で示した辞書$M$を生成した結果を図5-14に示す。この図では極座標で偏角を認識対象の姿勢角として、半径で認識対象の距離値のオフセット成分を示している。その後、全ての学習サンプルより奥行き移動成分を除去して固有空間ベクトルの累積寄与率0.8を目安に31次元で生成した。ただし、認識において固有ベクトルを第何成分まで用いるかという認識次元は、シミュレーションにより得られた結果を用いるものとする。仮想学習により生成された平均ベクトルと固有ベクトルの一例を図5-15に示す。このとき、図5-15で示されている固有ベクトルは式（5-4）で示される$E^*$であり、拡張固有空間はこれらに加えて奥行き移動を示すベクトル$e_v$が第1成分として構成される。これらの固有ベクトルにより投影された学習サンプルの固有空間上の投影点を図5-16に示す。

### 5.3.4　実験結果

認識対象を図5-17のようにレンジファインダから300mmを基準位置として前後±60mmの範囲で7ステップで動かした。また、回転は認識対象の中心で鉛直軸まわり120ステップとし、1物体あたり840回の認識を4物体について行った。また認識次元は事前に行なったシミュレーション結果より、物体認識の認識次元を6次元、姿勢角検出では認識次元を7次元とした。表5-1に認識結果を、図5-18に奥行位置検出結果の1例を示す（対象物体=花っ子）。表5-1の物体認識結果の全ての位置・姿勢および認識対象での平均は91.5%である。また姿勢角検出誤差は標準偏差値で33.2degから51.2deg程度、奥行位置検出誤差は標準偏差値で2.67mmから3.94mm程度であった(表5-2)。ただし、この結果は物体認識を誤ったときの検出結果を含めた値である。

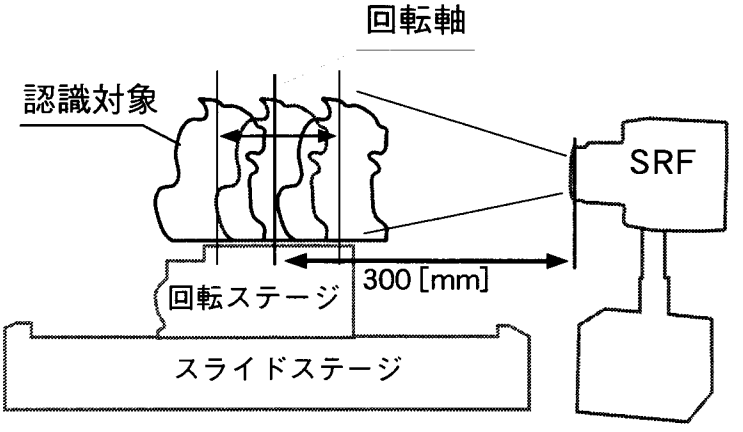



図5-17　実験環境

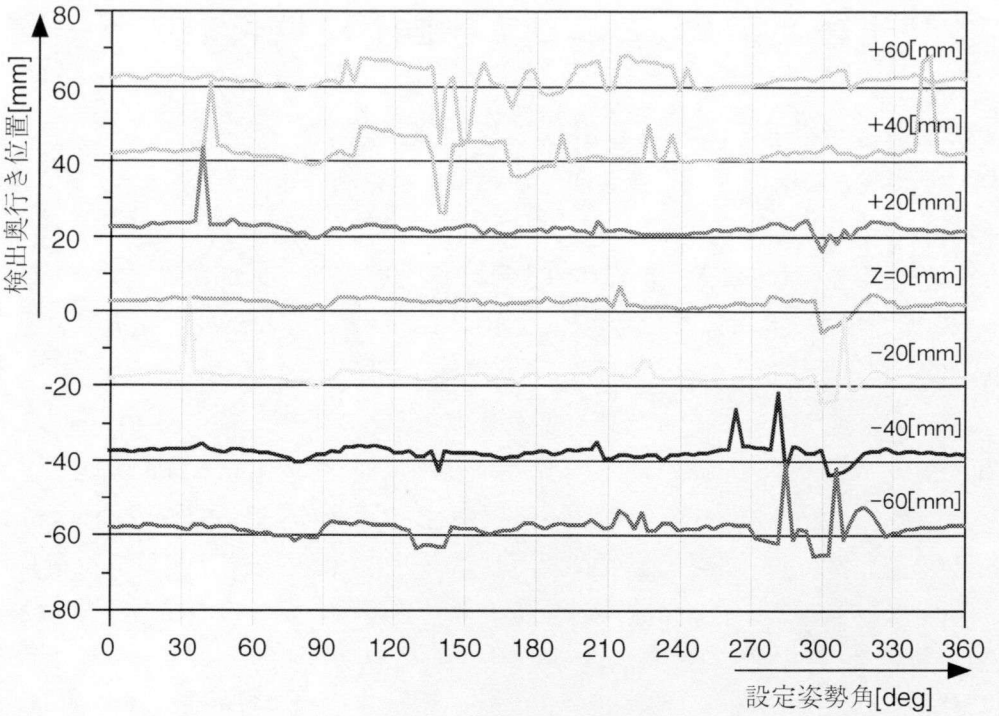



図5-18　奥行き位置検出結果（花っ子）

表5-1　物体認識結果

| 対象物体 | | 花っ子 | 招き猫 | たぬき | シーサ |
|---|---|---|---|---|---|
| 誤認識数/認識回数 | Z =-60mm | 7/120 | 1/120 | 11/120 | 9/120 |
| | Z =-40mm | 3/120 | 2/120 | 5/120 | 5/120 |
| | Z =-20mm | 2/120 | 3/120 | 3/120 | 2/120 |
| | Z = 0mm | 0/120 | 2/120 | 9/120 | 4/120 |
| | Z = 20mm | 4/120 | 27/120 | 7/120 | 7/120 |
| | Z = 40mm | 24/120 | 42/120 | 12/120 | 4/120 |
| | Z = 60mm | 31/120 | 42/120 | 12/120 | 7/120 |
| 認識率 [%] | | 91.5 | 85.8 | 93.0 | 95.5 |

表5-2　奥行き位置および姿勢角度検出結果

| 対象物体 | 花っ子 | 招き猫 | たぬき | シーサー |
|---|---|---|---|---|
| 奥行き位置誤差 SD[mm] | 3.77 | 3.04 | 2.67 | 3.94 |
| 姿勢検出誤差 SD[deg] | 41.93 | 33.27 | 44.13 | 51.28 |

## 5.4　仮想平行光学 RF および拡張固有空間に対する考察

### 5.4.1　仮想平行光学 RF におけるポリゴン近似誤差

仮想平行光学RFは見た目の大きさ変化がおこらない画像を生成することができるため、認識対象の奥行き変化に対して画像ベクトルの変化が単純な直線移動になるという利点がある。しかし、仮想平行光学RFは一度計測された距離画像をポリゴンで表現して、正射影のカメラパラメータで仮想的に撮影しなおすため、図5-19に示すようなポリゴン近似による誤差が問題となる。このような誤差の大きさは認識対象の表面形状の起伏とレンジファインダの画像解像度に依存する。図5-18を見ると認識対象の奥行き位置が基準位置より40mm以上遠くなると奥行き位置検出誤差が大きくなっている。この基準位置から+40mmは世界座標系の絶対位置で考えると奥行き位置340mmであり、平行光学変換のための基準平面をZ=160mmから大きくはなれている。そのため、平行光学変換でポリゴン近似誤差が生じ、奥行き位置検出誤差が大きくなっている。表5-2に示した認識対象毎の奥行き位置・姿勢検出誤差を見ると、シーサーの姿勢検出誤差は、他の認識対象に比べあまり良好な結果が得られていない。これはシーサーの表面形状の起伏が他の物体よりも大きいため、やはり仮想平行光学RFでの距離画像のポリゴン近似による誤差が大きくなっているからである。このポリゴン近似誤差は、レンジファインダの画像解像度と表面形状の起伏・空間周波数とのトレードオフの関係にある。そのため、より画像解像度の高いレンジファインダを用いればレンジファインダから遠く離れた場所に認識対象が配置されている場合や、表面形状の起伏が激しい認識対象でも正確な認識を行うことができると思われる。

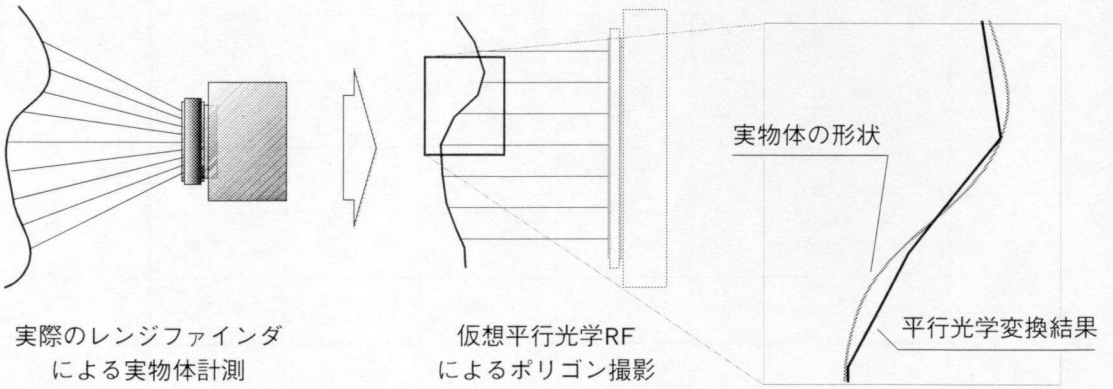



図5-19　ポリゴン近似による誤差

### 5.4.2 拡張固有空間を用いた直交化の利点

仮想平行光学RFと拡張固有空間を用いることにより、見た目の大きさ変化に対する学習サンプルを必要としないだけではなく、探索コストも削減している利点がある。学習コストを問題としないなら、単純に様々な奥行き位置による学習サンプルも用意して固有空間を生成すれば本手法と同様に奥行き位置検出を行なうことができる。しかし、固有空間は画像ベクトルの分散を特徴量とするため、図5-8に示したような奥行き方向の分散が極端に大きい場合は物体認識や姿勢検出の検出を正確に行なうことができなくなる。このように分散が極端に違うような場合、大局的な固有空間（ユニバーサル固有空間）と局所的な固有空間（オブジェクト固有空間）の2種類を用意して認識を行なう方法[Murase 94a]や、クラス間分散を大きくした固有空間とクラス内分散を大きくした固有空間により分散の大きさの違いに対応する方法[Daniel 96]を用いることも考えられる。しかし、もっとも大きく分散する要素は奥行き位置変化であり、奥行き変化は連続量である。そのため、単純に大局と局所に分けることができない。そのため、本手法のように直交化は、パラメータが同じ連続量同士である場合は有効な方法と言える。もちろんこのような拡張固有空間を用いるためには、パラメータ変化による画像ベクトル変化が線形的な関係で表現することが必要であり、画像の平行移動や回転など、様々な画像変化を画像ベクトルの線形的な変化で記述することは不可能に近いが、もし認識対象の奥行き位置変化に対する距離画像のオフセット成分のように線形化ができるような対象であれば、同様に拡張固有空間を適用することは可能である。

## 5.5 まとめ

距離画像の固有空間法では画像として距離画像を用いているため、奥行き位置の変化に対して敏感であり、それにより濃淡画像と比べて高精度な姿勢検出を実現できた。しかし、その反面奥行き位置のわずかな誤差が認識誤差を生じる原因となっていた。また、距離画像の固有空間法も画像照合に基づく方法論であるため見た目の大きさ変化も問題となっていた。この章では仮想平行光学RFを提案し、固有空間法における照合画像の見た目の大きさ変化の問題を解決した。さらに、拡張固有空間を用いることにより認識対象の奥行き位置の検出も可能にした。本章で述べた手法では、奥行き方向の位置検出のために新たな学習サンプルを必要としておらず、検出パラメータ数の増加に対する学習サンプル獲得コストの問題についても解決しており、距離画像の固有空間法の特性を生かした有効な手法である。しかし、自律移動ロボットの視覚として用いる場合を想定すると、認識対象がカメラの前にある場合だけでは不十分であり、環境中に存在する認識対象を探しだすことが必要になる。認識対象を探し出すためには物体認識が必要となるが、アピアランスベースの認識では認識対象が画像の中央に存在しないといけない。そのため、認識対象をどのように探しだせば良いかという問題が生じる。
第6章では自立移動ロボットの視覚などを想定し、今までに述べてきた距離画像の固有空間法を物体探索に拡張する手法について述べる。

# 第6章　回転レンジファインダによる物体探索

第2章から第5章にかけて距離画像の固有空間法に関する問題点を解決し、また仮想平行光学系や拡張固有空間を用いることにより物体認識・奥行き位置姿勢検出を実現した。しかし、自律移動ロボットなどの視覚システムへ応用するためには、認識対象がカメラの前にある場合だけでは不十分であり、環境中から認識対象を見つけ出す必要がある。このとき、環境中に存在する認識対象の位置を知るためには物体認識を行う必要があり、物体認識を実現するためには認識対象の位置を知る必要が生じる。そのため、物体探索では物体認識と位置検出によるフレーム問題を解決する必要がある。この章では、物体探索におけるフレーム問題を仮説の投票と検証を用いて解決する手法について述べ、実際のレンジファインダでの実験結果を示してその有効性を検討する。

## 6.1　物体探索

### 6.1.1　自律移動ロボットで必要とされる視覚機能

FA環境では、生産ライン上に流れる様々な機械部品などはパレット上に配置されていることが多く、物体の位置や姿勢は既知である。このように、もしも物体の位置や姿勢が既知であるれば認識は不要であり、指示された座標にアームを移動して物体をつかめばよい。しかし、人間の生活環境で自律移動ロボットを運用することを考えると、物体の位置や姿勢が既知であることは稀である。そのため人間の生活環境などの自律移動ロボットに要求されるタスクに対しては、視覚やその他のセンサーで物体を探索する必要がある。さまざま考えられるタスクに対して自律移動ロボットで必要とされる視覚機能としては以下の3つの機能が挙げられる。

*1)　既知の物体だけが存在する環境において既知の物体を見つける*
*2)　視覚で捕らえた物体が未知の物体であるか既知の物体であるかを判別する*
*3)　物体の背後に存在する既知の物体を見つけだす*

1)の機能は最も基本的な機能である。たとえば、自律移動ロボットにタスクを与える際に

*”コーヒーカップを食器棚に収納せよ”*
*”机の上の本を自分のところに持ってきなさい”*

というような”XをYせよ”というような命令は最も必要とされるタスクである。この”Xを”という表現は”X”というものがどういう物であるかという知識を持っているから可能であり、また詳細な位置を教えずに行うことが望まれているからである。もし、環境中に存在しうる全

ての物体を学習し認識できれば自律移動ロボットの視覚の機能としては十分である。しかし、全ての物を学習することは現実的ではないし、また人間の生活するような環境は常に変化する。そのため2)の機能が必要となる。具体的には、

*"床に散らばっている紙屑をごみ箱に捨てなさい"*

というようなタスクである。このような場合、紙屑一つ一つに対するモデルを用意することは現実的ではなく、本やノート以外の物体であれば紙屑であると認識することが望ましい。3)の機能が必要となるのは、仕事机に本やノートが山積みになって置かれていて、その背後にコーヒーカップが置かれているとき、

*"机の上にあるコーヒーカップを持ってきなさい"*

というタスクを与えるときなどである。これは、本やノートの認識は目的ではないため2)の機能で目的の物体でないことを知れば良い。その後アームで物体を移動するか、ロボット自身が移動することにより1)の機能で目的のものを見つければ良い。そのため、自律移動ロボットに必要とされる視覚機能としては1)と2)があればおおよその作業を行うことが出来るものと思われる。その中でも特に"既知の物体を見つけ出す"ことは視覚機能の基本であり、最も切望される機能である。本章では"既知の物体を見つけ出す"ことを物体探索として、以下では物体探索に的を絞って述べる。

### 6.1.2　ビンピッキングとImageSpotting

計測された画像から物体を探し出し、その位置姿勢を検出するという研究は古くから行われており、その先駆けとなった研究としてはPerkins[Perkins 77]やIkeuchi,Horn[Ikeuchi 86]によるビンピッキングシステムが挙げられる。このようなビンピッキングの研究では箱の中に工業部品が乱雑に配置されているシーンを対象としていた。これらの研究では認識対象を幾何モデルで表現し、撮影された濃淡画像から抽出された幾何学的特徴量との比較を行うことにより物体の位置や姿勢を検出していた。一般に工業部品は幾何学的に容易に表現することが出来ることが多く、また認識対象のCADモデルから幾何モデルを生成することも出来る。そのため、このような幾何モデル照合に基づく物体探索はFA環境では実用的な手法である。しかし、人間が生活するような環境で必要となる物体探索において、認識対象が常に幾何モデルで表現することが出来るとは限らない。また、幾何モデル表現が出来たとしても、計測された画像から安定して幾何特徴を抽出できるとは限らない。そのため、幾何モデルで表現することが困難な物体の探索を行う際には、固有空間法のようにアピアランスベースの物体探索が有利になる。

アピアランスベースの物体認識は文字認識や顔画像認識で古くから用いられている固有空間法が良く用いられる。NTTの村瀬らは、パラメトリック固有空間法を用いて環境中に存在する物体を探索するImageSpottingを提案している[Murase 94b]。この方法では、パラメトリック固有空間法を用いて観測された画像上でPixel-by-Pixelで照合を行い、物体探索を実現している。

このImage Spottingではパラメトリック固有空間法を用いているため、人形や人物の顔などのような幾何学的に複雑な形状の物体も容易に検出することができる。しかし、このImage SpottingではPixel-by-Pixelで行った認識結果のうち最も確からしい認識結果を探索結果として出力するため、複数物体が同時に存在する場合はどれか一つの物体しか検出することができない。もしも環境中に存在する認識対象の個数が既知であれば、信頼性の高い認識結果のうちいくつかを出力とすることで解決することができるが、外部から知識としていくつ存在するか与えられないのであれば、物体探索のみでいくつ存在するのか判断する必要がある。

### 6.1.3　本章で扱う問題

自律移動ロボットに対してより高度な作業を要求する場合は、視覚以外の知識と連携することが必要不可欠である。人間が日常生活で探し物をするときをを考えると、過去の記憶からおおよその位置や何個存在するのかを事前に知識として持っていることが多く、これらの知識を物体探索に用いていることも多い。しかし、このような前提知識がないこともあるうえ、たいていの場合このような前提知識も視覚から得られたものである。そのため、自律移動ロボットの視覚としても物体探索においても前提知識を用いずに物体探索を実現することが期待される。この章では、このような前提知識を用いないということを

*1)　認識対象の個数は未知である（認識対象がない場合も含む）*
*2)　同じ認識対象が複数ある場合もある*
*3)　異なる認識対象が同時に複数存在する場合もある*

と考え、自律移動ロボットの視覚機能として図6-1のような状況において

*”環境中に何がどこにあり、どちらを向いているか？”*

というタスクに答えるための物体探索を距離画像の固有空間法により実現することを目的とする。ただし、問題を簡単化するために

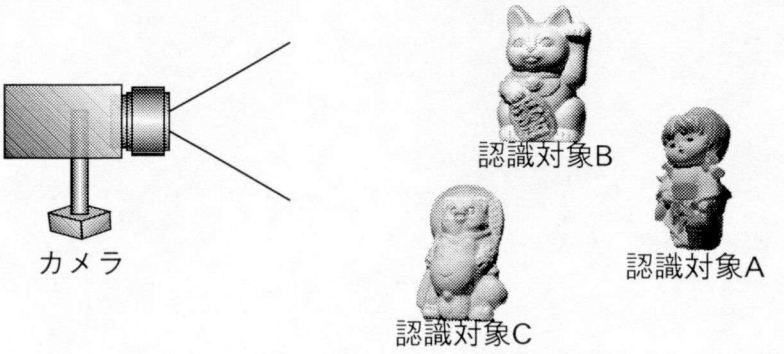



図6-1　本章で扱う物体探索

*1)　認識対象は平面上に配置されている（上下方向の移動はない）*
*2)　認識対象は傾いていない*
*3)　環境中に存在する物体は既知物体のみである*

という条件を用いることとする。段差やスロープの無い部屋や廊下は平面であり、また認識対象が安定した形であれば、認識対象が斜めに傾くことはない。このよう条件は、オフィスや住宅などの屋内環境であれば至る所で成立する。そのため上記の条件は、屋内利用を目的とした自律移動ロボットで物体探索を実現する際に非現実的なものとはなっていない。

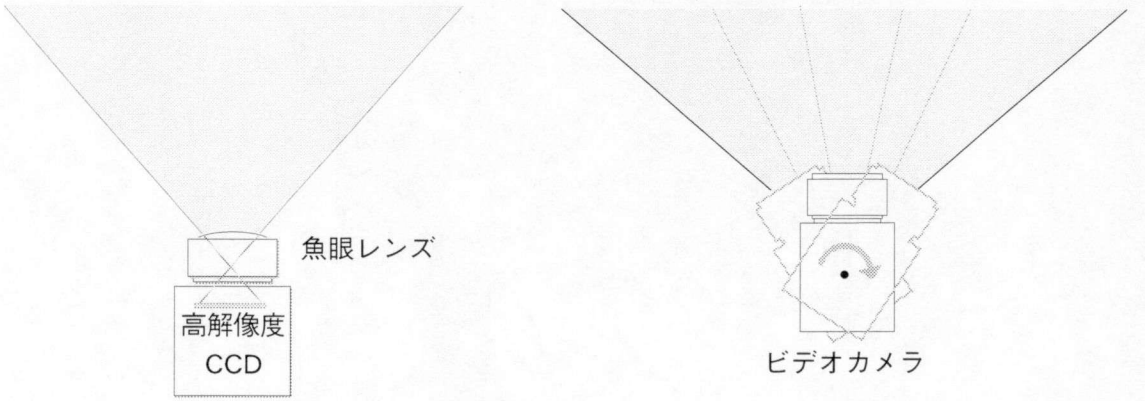
## 6.2　回転レンジファインダ

### 6.2.1　回転レンジファインダによる探索

近年市販されているカメラやレンジファインダは、年々安価で高画質になってきており、これにより正確な認識を実現しやすくなったが、画像解像度は依然としてNTSCやPALなどのビデオ信号により制限されていることが多い。物体探索を行うにあたり、どれぐらいの範囲をどれぐらいの画像解像度で撮影すれば良いかということは、自律移動ロボットに要求される作業の種類により決定されるため、一概に撮影範囲と解像度を言うことはできないが、通常のビデオカメラやレンジファインダでは画角に対する画像解像度は不十分であると思われる。このような問題を解決するためには、より広い範囲を高解像度で撮影する装置を用いることが必要となるが、そのための方法論としては以下の二つが考えられる。

*1)　魚眼レンズなどのように広角のレンズと高解像度のカメラを用いる（図6-2(a)）*
*2)　カメラをパン・チルトステージで回転させ広域を捜査する（図6-2(b)）*

上記の1)の方法論では、広域を撮影するために特殊な機構を必要としないという利点はある




(a)　魚眼レンズを用いる方法　　(b)　パンステージを用いる方法

図6-2　広範囲の撮影手法

が、高解像度のカメラを用意する必要がある。これに対して2)の方法論では、カメラは既存のものを用いればよいが、カメラを回転させるための機構を必要とする。このように、それぞれ利点欠点はあるが、1)の方法で計測された画像において特定の領域を捜査したり、また2)の方法で計測された画像群を統合して認識を行うとすれば、どちらの方法でも認識段階では等価な撮影装置として考えることが出来る。距離画像の固有空間法を用いて物体探索を実現するためには、広域を撮影することが出来るレンジファインダを用意する必要があるが、汎用のレンジファインダではあまり高解像度のものはない。特に本論文で扱ってきたSRFは画像解像度が24×24画素とあまりにも低い。そのため、広域を撮影するためのレンジファインダとしては2)の方法論を用いることにした。

### 6.2.2　装置構成

環境中の広域を撮影するために図6-3に示すようにレンジファインダとパンステージを組み合わせた装置（回転レンジファインダ）を構成した。また、レンジファインダはSRFを用いた。この回転レンジファインダではステージの鉛直軸周りの回転により左右方向に対しては広い範囲を撮影することができるが、上下方向はレンジファインダの画角の範囲しか撮影することができない（図6-4）。しかし、6.1.3で設けた”認識対象が平面上に配置されている”という条件の基ではこのようなパンステージのみの回転だけで十分である。また、この回転レンジファインダに第5章で述べた仮想平行光学RFと拡張固有空間を用いることで平面上に配置された認識対象の奥行き方向の位置も検出することができる。

この回転レンジファインダを探索に用いるにあたり、環境の座標系を示す絶対座標系とレンジファインダ上の座標系を用いて図6-5のように設定した。この座標系は、レンジファインダの回転角度が0のときのレンジファインダの座標系$X_0 - Y_0 - Z_0$を絶対座標系として考えており、レンジファインダの座標系$X - Y - Z$はステージの回転によって回転する座標系を用いている。

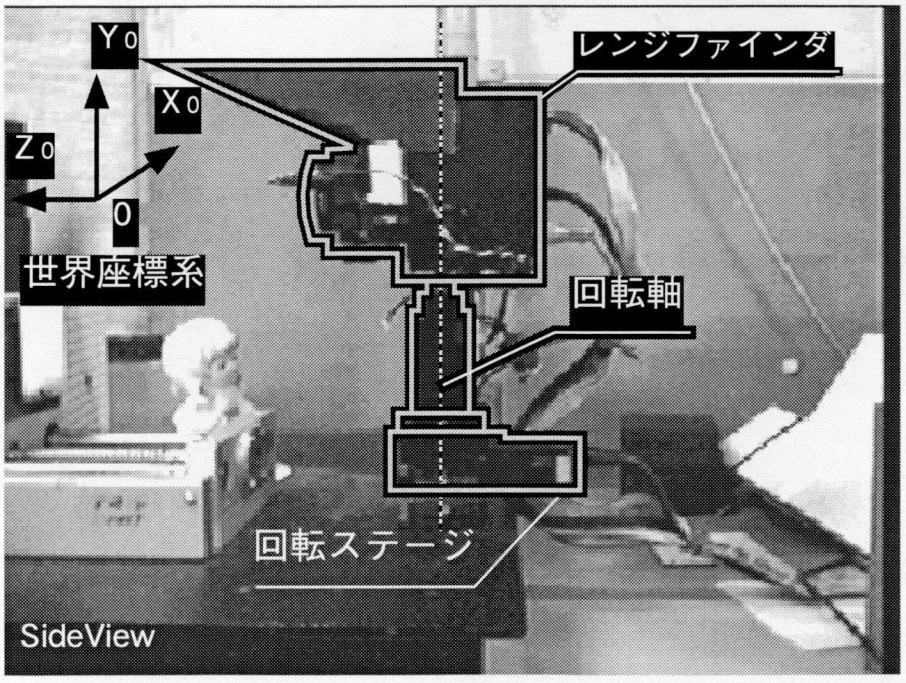



図6-3　回転レンジファインダ

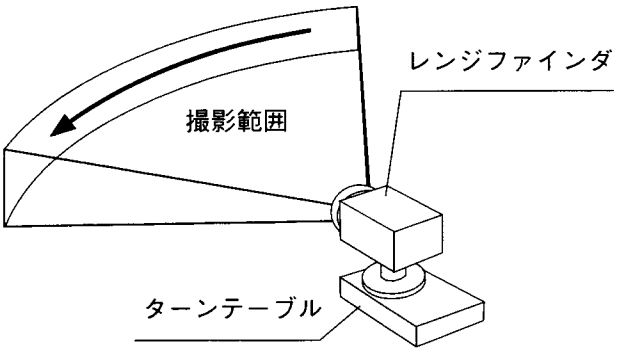



図6-4　回転レンジファインダの撮影範囲

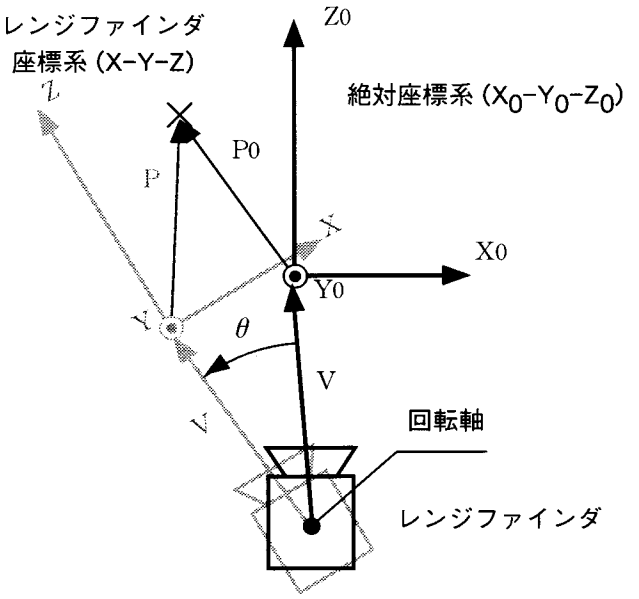



図6-5　絶対座標系とレンジファインダ座標系

ただし、レンジファインダの回転は$X-Z$平面に直交する回転軸であるものとして、座標変換は$X-Z$平面上のみで考えるものとする。また、図6-5において$V$は絶対座標系上でレンジファインダの回転軸（パンステージ）の位置を示すベクトルである。以上より、レンジファインダが$\theta$だけ回転したときに計測された点$P$の絶対座標系での位置は

$$P_0 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \left[ P + V \right] - V = \boldsymbol{R} \left[ P + V \right] - V \tag{6-1}$$

と表すことができる。

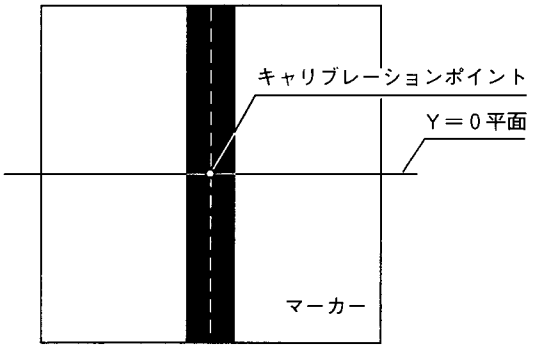



図6-6　マーカーの位置計測

### 6.2.3　キャリブレーション

図6-4に示した回転レンジファインダを用いることにより広域を撮影することができ、環境の広範囲に対して物体探索を行うことが可能となる。ここでは、図6-5で示した回転レンジファインダの座標系においてレンジファインダの回転軸の位置を求めるキャリブレーションについて述べる。

式(6-1)で示した回転レンジファインダにおける座標変換の式より、レンジファインダの回転軸の位置を示すベクトル$\boldsymbol{V}$は簡単に

$$V = \left[\boldsymbol{R} - \boldsymbol{I}\right]^{-1}\left[P_0 - \boldsymbol{R}P\right] \tag{6-2}$$

と示すことができる。そのため、異なるレンジファインダの回転角において観測された同一の点$P_0$と$P$が一組あれば$V$は決定される。また、任意の回転角$\theta_1$、$\theta_2$における対応点$P_1$、$P_2$から$V$を求めるときは

$$V = \left[\boldsymbol{R}_1 - \boldsymbol{R}_2\right]^{-1}\left[\boldsymbol{R}_2 P_2 - \boldsymbol{R}_1 P_1\right] \tag{6-3}$$

とすれば良い。そのため、何らかのマーカーを用いて異なる回転角で対応する2点を計測することによりレンジファインダの回転軸の位置$V$を知ることができる。

距離画像に対するマーカーは適当な凹凸を持った基準物体を座標系に配置すればよいが、SRFで撮影される距離画像は奥行き解像度は高いものの、画像解像度は24×24画素と非常に少ない。そのため、距離画像のみでは精度良くマーカーの位置を計測することはできない。そこで、SRFで同時に撮影することができる濃淡画像も用いてマーカーの位置計測を行った。具体的には図6-6に示すような平板に黒い線を引いたマーカーを用いて、CCD画像からキャリブレーションターゲットーの視線平面を、また距離画像からマーカーの平面を計測してその後これらの平面と絶対座標系の$Y$=0平面との交点を求めることによりマーカーの位置を計測した。これにより、マーカーの位置を高精度に計測できるが、図6-3に示した回転レンジファインダはカ

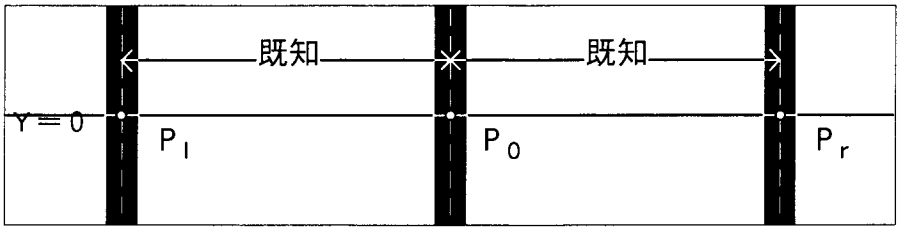



図6-7　キャリブレーションターゲット

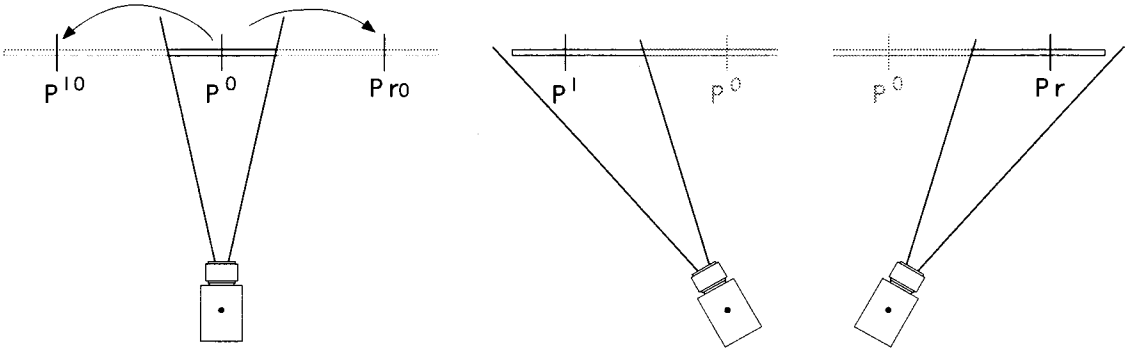



図6-8　キャリブレーション方法

メラの画角が狭いために用いる機構であり、レンジファインダを回転することにより同一のマーカーを撮影することができなくなる。そこで、図6-7に示すキャリブレーションターゲットを用いて対応点を求めることにした。ただし、キャリブレーションターゲットは絶対座標系のZ軸と直交するように配置することにして、マーカーの間隔は既知としている。キャリブレーションの流れとしては（図6-8）、

1) *中心のマーカーをの位置を計測して、レンジファインダの回転角$\theta_0$におけるレンジファインダの座標系において左右のマーカーの位置$P_{l0}$,$P_{r0}$を求める。*
2) *レンジファインダを左方向$\theta_l$回転させレンジファインダの座標系において左のマーカーの位置$P_l$を計測する。*
3) *同様に右のマーカーの位置$P_r$とその時のレンジファインダの回転角$\theta_r$を得る。*

以上の計測により得られた2組の対応する座標値とレンジファインダの回転角のデータ

$$\begin{aligned}(P_{l0}, \theta_0) &\leftrightarrow (P_l, \theta_l)\\ (P_{r0}, \theta_0) &\leftrightarrow (P_r, \theta_r)\end{aligned} \tag{6-4}$$

をそれぞれ式(6-3)に代入して$V$を求め、これらの平均を用いることにした。

## 6.3 探索アルゴリズム

### 6.3.1 認識と信頼性

回転レンジファインダを用いることにより広域を解像度を落とさずにスキャンすることが可能となる。このようなレンジファインダを用いて6.1.3で示した物体探索を実現するためには、どのようにレンジファインダで走査すれば効果的であろうか。この答えとしては、様々な方法が考えられるが、ここでは簡単に探索範囲の端から端までを等間隔のステップで認識を行うことにした。個々のステップでの認識は、第5章で述べた仮想平行光学RFと拡張固有空間を用いた物体認識・姿勢検出と同時に奥行き位置検出を用いる。これにより、図6-4で示した撮影範囲で様々な位置に配置された認識対象を検出することができる。以上のように探索を行うことで、探索の結果としてレンジファインダの回転角ごと認識結果として物体とその時の位置・姿勢角度を得ることができる。しかし、これらの認識結果はすべてが正解ではなく、誤認識や重複した認識結果を含むものである。ここでは、このように誤認識や重複した結果を含む認識結果を仮説と呼ぶ。このような仮説から正解を見つける手がかりの一つとして、認識で得られた仮説が認識結果を $s$ 番目の認識対象であると判断したとき、仮説の信頼性を

$$
\begin{aligned}
&reliability(s) = 100 \times \frac{l_s^{-1}}{\sum_{i=1}^{S} l_i^{-1}} \ [\%] \\
&l_s = \min_r \left\| p_{r,s} - p_m \right\|
\end{aligned}
\tag{6-5}
$$

により評価した。ただし、$S$は認識対象の数、$p_m$は仮想平行光学RFで計測された距離画像$E^*$上の投影点であり、$p_{r,s}$は学習サンプルの軌跡上の点である。この評価関数は、拡張固有空間の物体認識・姿勢検出を行うための部分空間$E^*$上において認識対象ごとの学習サンプルの軌跡と投影点の距離の逆数の比率を信頼性としている。このままでは、認識対象ごとの学習サンプルの分散の違いにより、信頼性に偏りが生じるため、影点$p_m$が部分空間$E^*$の原点であったとき

$$
reliability(s) = 100 / S \ [\%] \ \ for\ s = 1,2,\cdots,S \tag{6-6}
$$

となるように認識対象ごとに距離尺度を正規化した。

### 6.3.2 仮説の投票

式(6-5)で示した評価関数を用いることにより、仮説の信頼性を測ることができるため、正解を見つけるための有用な手がかりの一つとなる。もし、環境中に存在する認識対象が常に一つであるならば、信頼性の最も高いものを探索結果として出力すれば良い。しかし、本章で目的

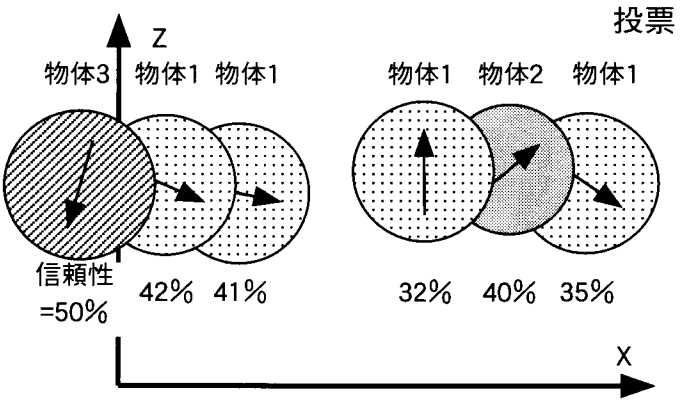



図6-9　仮説の投票

とする物体探索は、複数の認識対象が同時に存在するような場合も考えており、認識の信頼性のほかに何らかの情報が必要となる。そこで、正解を見つけるためにもう一つの手がかりとして認識対象の大きさを用いることにした。認識対象の大きさを与えるためには、あらかじめ認識対象の大きさを計測しておく必要があるが、仮想学習では認識対象の形状モデルを用いているため、この形状モデルから容易に認識対象の大きさを知ることができる。以上の認識結果の信頼性、認識対象の大きさの二つを手がかりとして第5章で述べた奥行き位置検出手法を用いれば、認識によって得られた仮説は図6-9のように絶対座標系上に投票することができる。ただし、仮説の投票は認識対象の大きさに応じた円を用いた。また、このような仮説の投票において計測された距離画像の画素のすべてが欠損していた場合は、レンジファインダの前に物体が存在しないということであるのでこのような時は投票しないことにした。これにより環境中に物体が存在しない場合は、物体が存在しないという探索結果を出力することができる。また、$sizeX \times sizeY$の距離画像、$D$次元の固有空間で平均ベクトルを用いる際に、欠損画素率が

$$Lacked\ Pixel\ Allowance = 100 \times \frac{N-(D+2)}{N}\ [\%] \tag{6-7}$$

以上であるときは、BPLPによる逆射影が成立しないため、計測および仮想光学変換された距離画像において欠損画素の数が式(6-7)より多い場合は仮説を投票することができない。そのため、この条件を満たしていない距離画像については仮説の投票を行わないことにした。

### 6.3.3　仮説の検証

絶対座標系に円で投票された仮説において、正解は物理的な観点や仮説の信頼性から

- *ほかの仮説と重なることはない。*
- *信頼性が低い物ほど不確かな仮説である。*

ということができる。このような考えに基づき、仮説から正解を導くために、

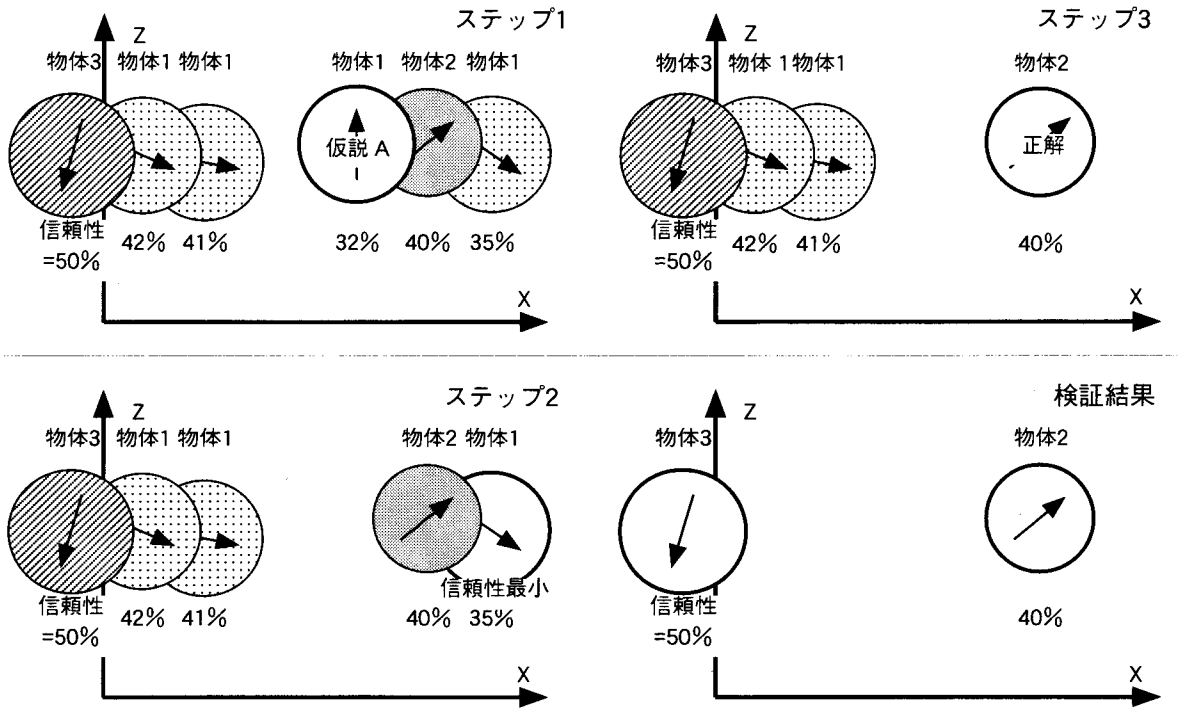



図 6-10　仮説の検証

*1)　信頼性が最小の仮説（仮説 A）を見つけ、その他の仮説との重なりを調べる。*
*2)　仮説 A が他の仮説と重なっていなければ仮説 A を正解とする。*
*3)　仮説 A が他の仮説と重なっている場合は仮説を削除する。*
*4)　1)から 3)を繰り返し仮説の重なりがなくなったら終了する。*

という手法を用いた（図 6-10）。この手法では、認識対象が複数存在する場合、同じ認識対象が存在する場合も対応することができる。

## 6.4　物体探索実験

物体探索の実験では、認識対象として図 6-11 に示す 4 体の人形を用いた。探索を行うためには認識対象の奥行き方向の位置検出も必要となるので、認識は第 5 章で仮想平行光学 RF と拡張固有空間を用いた。また、学習サンプル獲得コストを軽減するために、仮想学習を用いた。

探索範囲は回転レンジファインダの回転角 -35deg. から 25deg. で撮影される範囲として、レンジファインダの回転ステップを 0.5deg. 刻み（全 120Steps)とした。

### 6.4.1　認識対象が複数存在する場合

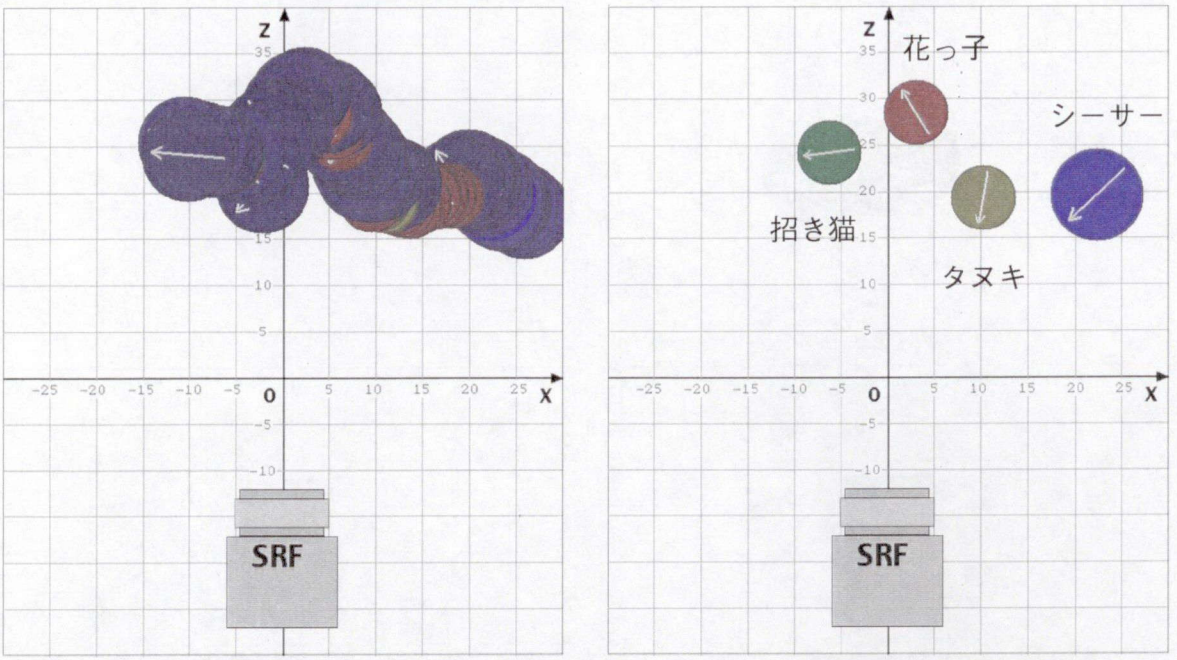
6.3で述べた物体探索手法では複数の認識対象が環境中に配置されている場合にも対応することができる。この探索手法の有効性を検証するために、図 6-12 に示すような状態について物体探索を行った。探索の過程において、絶対座標系上に仮説を投票した結果を図 6-13(a)に、

(b) 花っ子　　(a) 招き猫　　(c) タヌキ　　(d) シーサー

図6-11　認識対象

図6-12　認識対象が複数存在する場合

検証した結果を図6-13(b)に示す。図6-13において、円は認識により得られた仮説を示しており、円の色で認識対象の対象物体の種類、円の中の矢印は検出された認識対象の姿勢を示している。この実験では、認識対象は位置決めをしていないため、探索結果の精度評価はできないが、結果を見ると、それぞれの認識対象の位置関係やおおよその姿勢は正しく検出されていることがわかる。また、本手法で用いている固有空間法による認識は、第5章で述べた認識手法を用いているため、おおよその精度としては第5章の実験結果より確認された程度の精度が得られているものと思われる。また、探索時間は全120Stepsで58Sec.であり、1Stepあたり0.48Sec.程度であった。この時間は、パンステージによる回転とレンジファインダでの距離画像計測、仮想平行光学変換、物体認識を全て含めた時間である。実験に用いた計算機は、物体認識にSGI社Octane、距離画像計測にSGI社Indy、ステージコントロールにSGI社Indigo2を用いており、これらをEtherNetでつないで分散処理を行なっている。




(a) 投票結果　　(b) 検証結果

図 6-13　複数物体の探索結果

### 6.4.2　認識対象が一つだけ存在する場合

図 6-14 に示すように認識対象が一つしか存在しない場合について実験を行った。図 6-15(a) に仮説の投票結果を、図 6-15(b)に仮説の検証結果を示す。6.3 で示した物体探索手法は環境中に一つしか存在しないときであっても、同様に検出することができる。

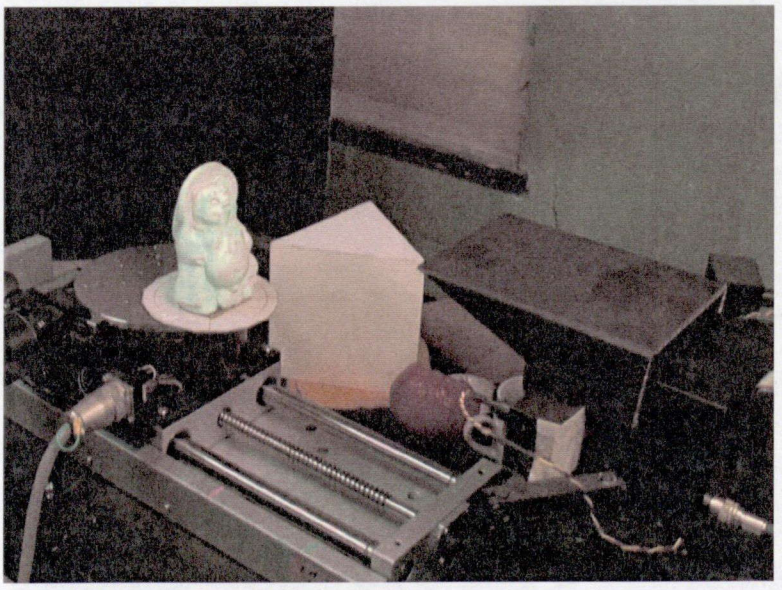
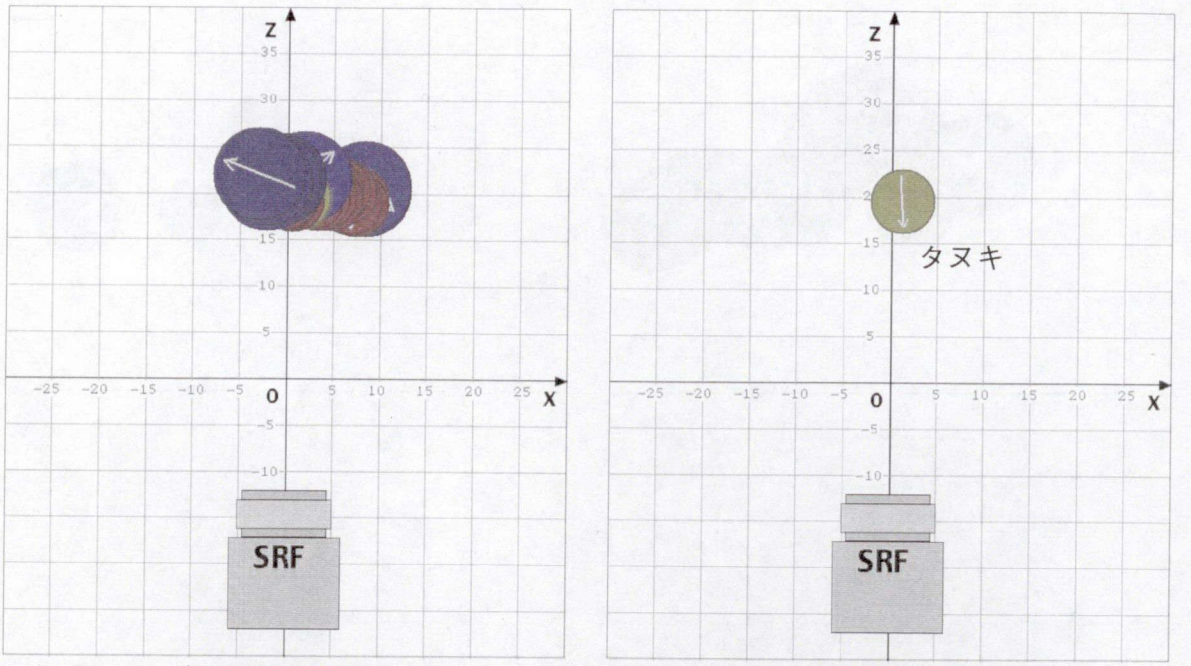


図 6-14　認識対象が一つ存在する場合



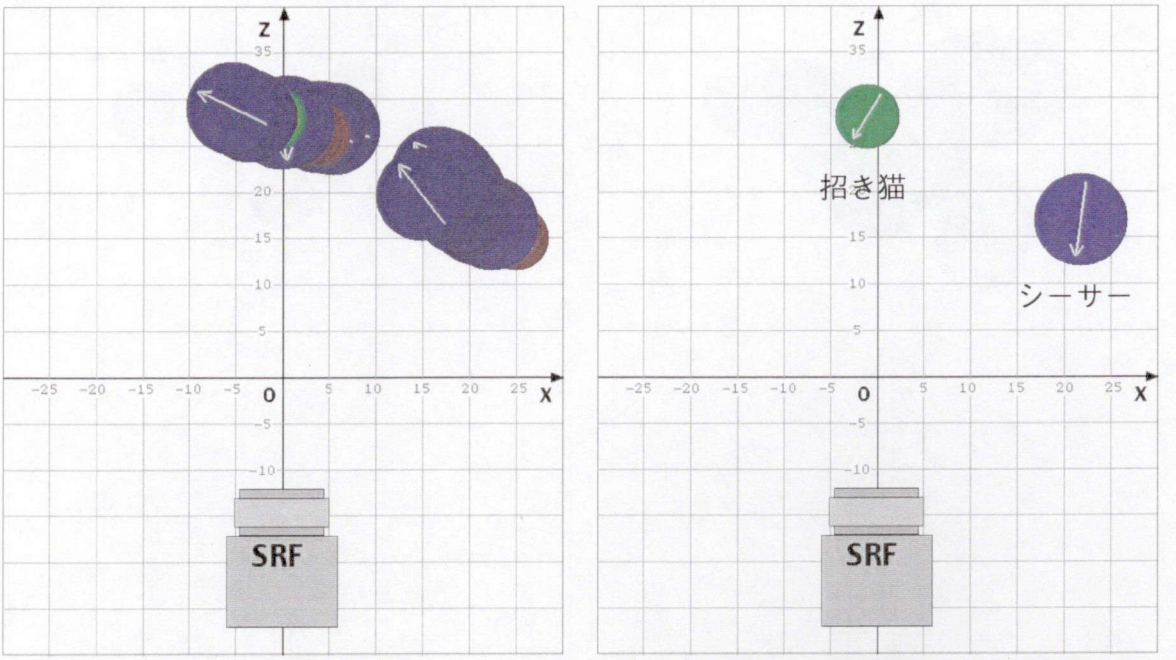

(a) 投票結果　　(b) 検証結果

図6-15　単一物体の探索結果

### 6.4.3　未知の物体が含まれる場合

この章で述べた手法では、6.1で述べたように既知の物体のみが存在する場合を扱っているため、未知の物体が含まれている場合は物体探索を行うことができない。しかし、未知の物体が同程度の大きさであるなら、認識は正しく行えないものの探索によりどこ物体が存在する場所は検出することができる。図6-16のように、既知の物体と未知の物体が存在する状況で物体探索を行った結果、図6-17に示すような探索結果が得られた。図6-17からわかるように、未知の物体は図6-11に示した既知の物体として検出されているが、検出された位置はおおよそ正しい位置になっている。

図6-16　未知の物体が含まれる場合




(a) 投票結果　　(b) 検証結果

図6-17　未知物体を含む場合の探索結果

### 6.4.4　背景が存在する場合

図6-18に示すように背景が存在する場合について探索を行った結果を図6-19に示す。背景であるのか物体であるのかという議論は、画像のセグメンテーションで最も基本となる問題であるが、もし背景という概念がなければ背景も物体の一つである。この章で述べた物体探索は、既知の物体のみが存在する環境を扱っているため、このような背景も未知物体の一種となりこのように誤った探索結果が得られた。しかし、中央には位置された既知の物体の検出は正しい結果が得られている。

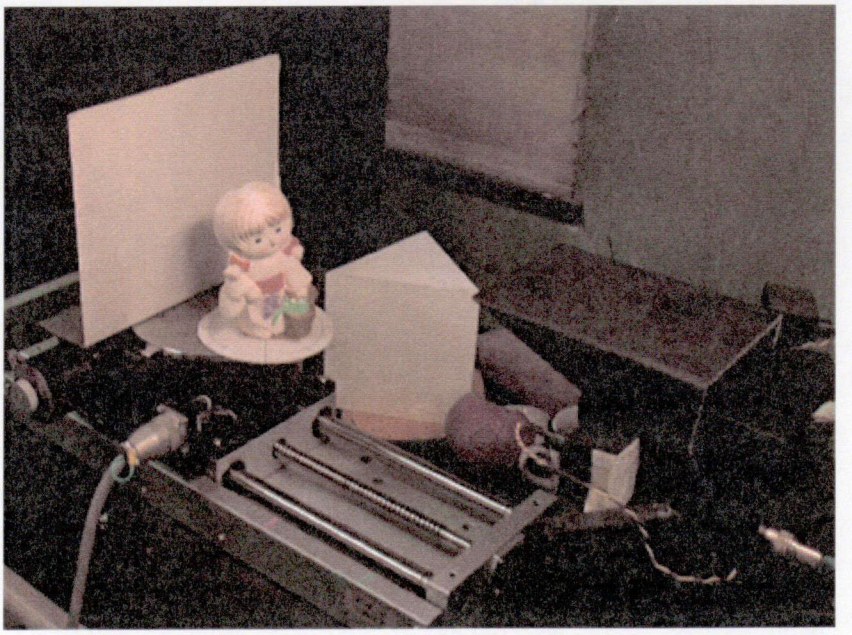
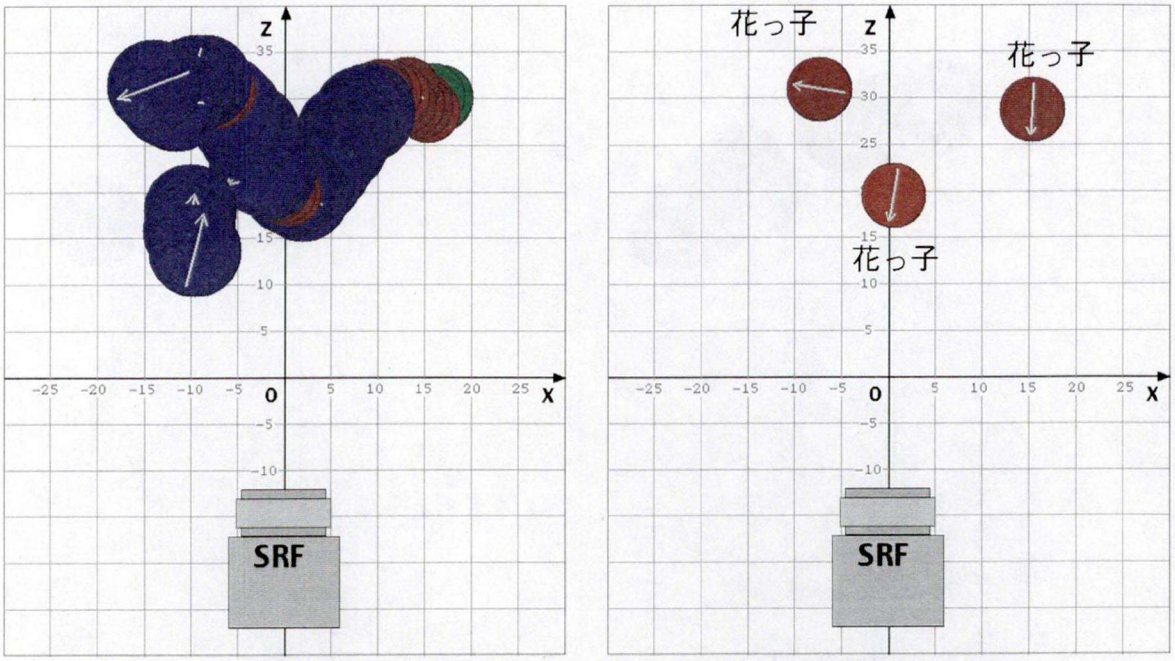


図6-18　背景が存在する場合




(a) 投票結果　　(b) 検証結果

図6-19　背景が存在する場合の探索結果

## 6.5　物体探索考察

6.4の実験結果より、既知の認識対象が複数存在するような場合についても物体探索が行えることができた。しかし、背景が存在する場合であるとか、未知の認識対象が存在する場合についていくつかの問題点が残った。ここでは、これらの問題点について考察しその解決手法を提案する。また、本手法を自律移動ロボットの視覚として応用する際の考察も述べる。

### 6.5.1　未知物体を含む場合への拡張

この章で述べた手法では未知物体が含まれるような場合は考慮されていない。そのため、このままでは未知物体を含むような環境の探索は行うことができない。人間が生活するような環境において、全ての物体を学習するということは現実的な解決方法ではなく、未知の物体が同時に存在するような環境も扱えることが望ましい。本手法では仮説の検証方法において認識対象の大きさを用いているため、そのままでは未知物体を含む環境に対応することができない。しかし、固有空間法において未知の物体であることを認識するのは比較的簡単であり、たとえば物体認識において図6-20に示すように固有空間上の原点との距離も考えればよい。このような未知の物体をすべて原点で代用させる方法は、必ずしもうまく行く保証はない。しかし、固有ベクトルは学習サンプルに対しては相関が強くなっており、未知の画像に対しては大きな相関を持たない傾向にあるため妥当な手法である。もしもより正確な認識が必要であれば、投影点と軌跡の距離を求めて経験的に定めた閾値で未知であるか既知であるかを決定すれば良い。

以上のように固有空間法での認識は比較的容易に対応することができるが、仮説の投票において未知物体の大きさをどのように扱えば良いかということは問題である。もし大きさがおお

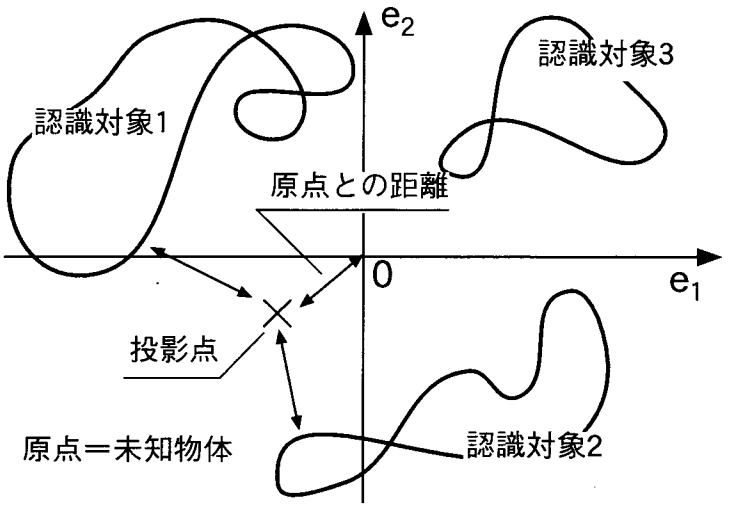



図 6-20　未知物体であることの認識

よそ既知であればその大きさに応じた円で投票すればよいが、大きさも未知であれば本章で述べた方法論では扱うことができなくなる。もちろん、未知物体が存在することやその位置を知る必要がないのなら仮説を投票しなければ未知物体に対する問題は解決する。

### 6.5.2　背景問題について

本章で提案した手法では、既知物体のみが存在する環境での物体探索を課題として扱っている。壁は物体とは別の概念で考える場合もあるが、アピアランスベースの認識手法において、物体とは物体の見た目の姿の集合であり、壁を撮影したときに得られる未知画像は、未知物体を意味している。そのため、背景が存在する場合は本章の問題設定の域を越えた課題である。しかし、構造物内で物体探索を行う際には、周囲には至るところで壁が存在するため、背景に影響されない物体探索を実現することが望ましい。このような背景は撮影される画像において未知物体と壁画像は学習していない画像ということでは同じであり、6.5.1で述べた手法で解決することができる。また、距離画像は奥行き方向の位置関係を知ることができるため、距離情報より認識の前処理として何らかのセグメンテーションで背景を除去することができればこのような背景が存在するような場合にも対処することができると思われる。

別の解決手法としては、背景の平面性を仮定することが考えられる。背景とは、人間の主観や認識を行う際の都合により定義されるものであり、物体の背後に存在する平面を背景とすることは常に正しいわけではない。しかし、建物内の廊下や部屋は、壁で仕切られており壁は平面近似することができることが多い。もし、このように背景を平面で仮定できるなら、固有空間法の学習段階において壁をモデル化した平面の様々な姿勢を学習すれば、本手法で背景を扱うことができる。また、このように壁を検出することができれば、物体認識と同時に環境マップを生成することもできる。

図6-21　オクルージョンが生じている場合の認識

### 6.5.3　オクルージョンについて

画像照合に基づく認識の問題点の一つとして、図6-21に示すように認識対象が何らかの物体により一部が隠されている場合の認識が挙げられる。このような問題は一般にオクルージョン問題と呼ばれている。本章で述べた物体探索法は、物体認識において固有空間法を用いており、やはりこのような問題が生じる。正しく照合を行うためには、オクルージョン領域をあらかじめ知っておく必要があるが、オクルージョン領域を知るためには認識が必要である。そのため、この問題は、認識が先か、オクルージョン検出が先かという難しい問題になり容易に解決することはできない。しかし、距離画像では奥行き方向の位置関係を知ることができるため、何らかの手法によりオクルージョン領域を知ることができる可能性がある。また、もしも手前にある物体が既知の物体であるなら、手前の物体を認識できればオクルージョン領域を知ることができる。もし、以上の手法によりオクルージョン領域を知ることができれば、オクルージョン領域を用いずに認識を行えば良い。オクルージョン領域を用いない認識は、オクルージョン領域を欠損画素として扱えば第4章で述べたARLPやBPLPにより実現することができる。

### 6.5.4　自律移動ロボットへの応用

この章で述べた物体探索では、認識対象が平面上に配置されている場合のみしか扱うことができない。そのため高さ方向の位置検出は不可能である。また、物体の姿勢角検出も鉛直軸での姿勢角しか検出できない。しかし、構造物内を走行する自立移動ロボットや産業用ロボットの視覚システムでは、検出を必要とする対象物体が平面上に設置されていることが多く、このような場合では平面内の位置と鉛直軸回転での姿勢検出で十分であると思われる。もしも認識対象の高さ方向の位置ずれや傾きに対応する必要があるならば、このような変化もパラメータとして固有空間法で学習すれば良い。学習サンプルを単に増やして解決することは、学習サンプル獲得コストを増大させるためあまり妥当な解決策ではないが、本章で述べた物体探索で用いた固有空間のパラメータは個々の物体に対する鉛直軸周りの姿勢変化のみである。そのため、高さ方向の位置検出を新たなパラメータとして扱っても学習サンプル獲得コストは非現実的なものとはならない。

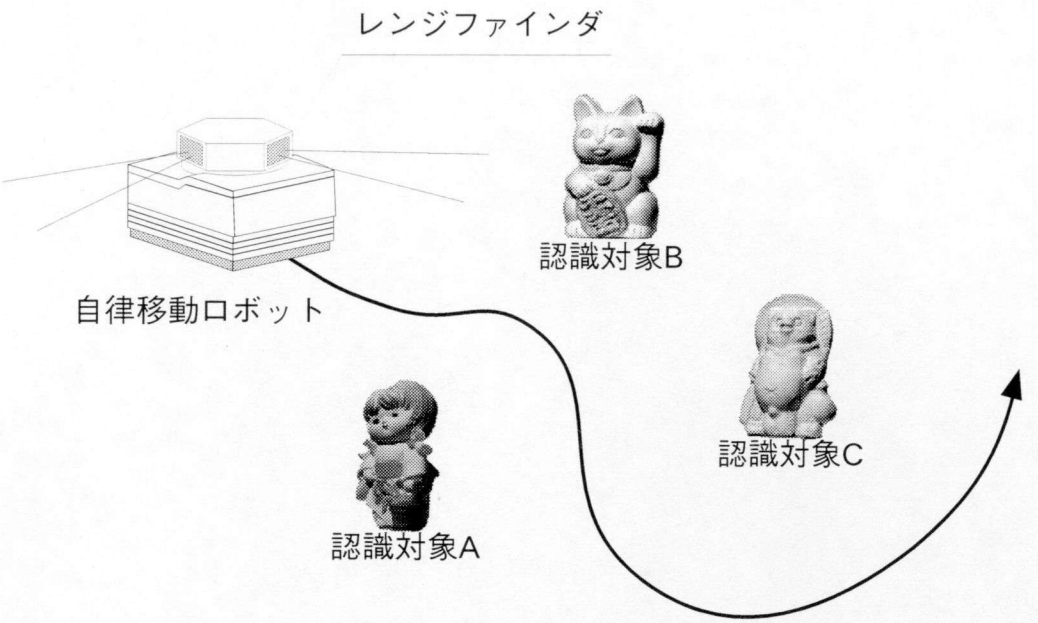



図6-22　自立移動ロボットへの応用

また、物体探索では広い視野を確保するために、レンジファインダを回転させるパンステージを用いたが、自律移動ロボットなどに実際に搭載する際には、自律移動ロボット自体が移動することができるため、このような回転機構を用いる必要はない。そのため、図6-22のようなレンジファインダを搭載している自律移動ロボットなどに対しては本手法をそのまま応用することができると思われる。

## 6.6　まとめ

本章では、第5章までに述べた距離画像の固有空間法の様々な手法を用いて物体探索を実現した。物体探索では、広い視野を確保するためにパンステージにより回転するレンジファインダを用いた。また、物体探索において物体の位置は物体を認識することにより決まり、物体を認識するためには位置を知る必要があるという問題に対しては、認識結果を仮説として表して検証を行うことにより解決した。実験の結果、本手法では複数の既知物体が混在するような場合でも物体探索を行うことができることが確認された。また、本手法では探索において環境中に存在する既知物体の個数を制限したり、個数を外部から与える必要はないという利点がある。ここで行った物体探索は、既知物体が複数存在する環境を仮定していたため、実験において未知物体が含まれるような場合は誤認識が生じる。また、背景が存在する場合についても誤認識が生じることが確認された。このような問題を解決するためには、6.5.1で述べたように既知の物体であるか、未知の物体であるかを判断する手法により解決することができる。この機能の実現については今後の課題としたい。

# 第7章　結論

本論文は、コンピュータビジョンの基本的な課題である物体認識において、距離画像を用いた固有空間法による方法論を確立することを目標に研究してきた成果をまとめたものである。物体認識の歴史は長く、1970年代後半から様々な研究が行われている。しかし、ビンピッキングや3DPO's が開発されてから10数年経つ現在においても、物体認識は依然として難しい課題である。濃淡画像を用いたビジョンシステムや幾何学的な特徴量の照合などの人間模倣型の物体認識に固執した考えは、物体認識を難しくしている要因の一つであると思われる。もちろん、人間と同じ計測原理を用いて人間と同じ方法論で物体認識を実現することは、コンピュータビジョンの研究のみならず認知科学の分野でも興味深い研究である。しかし、2次元画像から3次元情報を生成したり、撮影される画像において光学的な特性と幾何学的な特性を分離する必要が生じる。また、幾何モデル照合や線画の解釈は明確であるが、実環境の計測ではモデルで記述されている特徴量を必ずしも抽出できるとは限らない。そこで本研究では、このような人間を模倣したアプローチではなく、計測的な立場から物体認識を実現することを心がけてきた。

距離画像は形状計測を目的とする画像であるため、3次元幾何モデルと照合により物体認識を行うことは自然な考えである。しかし、距離画像を用いた固有空間法では距離画像を距離値という値が並んだ画像と考えていると固有空間照合を行っている。そのため、認識において距離画像の意味を積極的に用いた手法ではないが、距離画像を固有空間法に用いても、距離画像は照明条件に依存しないとか、カメラからの物体表面の形状を直接反映しているということに違いはない。そのため、究極的にはどちらの方法論を用いても、同等の物体認識を行うことができる。しかし現実的な点を考えると、幾何学的に複雑な形状の物体を3次元幾何モデルで表現することは困難である。これに対して、距離画像の固有空間法では3次元幾何モデルなどの表現を必要としないため、自由曲面物体などの認識も容易に実現できる。しかし、このような利点に対して距離画像の固有空間法では、見え方の変化による多量の学習サンプルを用意する必要が生じる。そこで、この問題に対して認識対象の形状モデルから距離画像を仮想的に生成して学習を行う”仮想学習”を提案した。この仮想学習を用いると、画像照合ベースの物体認識手法でありながら、学習サンプルを必要としない認識手法となり、認識システムに与える知識としてはモデルベースの物体認識と同等なものとなる。このとき異なる点としては、モデルベースの物体認識では、物体認識を実現するために辺や頂点などの特徴量をを組み合わせてどのような認識木をつくるかという、認識プログラムを作成する必要があるのに対して、仮想学習では、形状モデルから仮想的に距離画像を生成して、固有空間をつくるというところである。第3章の実験では、4体の対象物体の形状モデルからそれぞれ64枚ずつ距離画像を生成した(256枚)。このとき、形状モデルは6枚の距離画像を張り合わせたものを用いた。そのため、この仮想学習では24枚の距離画像から256枚の距離画像を生成したと等価であり、学習サンプル獲得コストの低減に有効な手法であることが確認できた。

しかし、このような仮想学習を実際のレンジファインダに応用する際には、実際のレンジファインダで計測される距離画像で生じる欠損画素をどのように扱えばよいかが問題となる。

このような欠損画素は距離画像のみで生じる現象であり、この点で濃淡画像を用いた固有空間法と比べて不利になる。最も簡単な解決法としては、欠損画素に適当な値を代入すればよい。しかし、欠損画素の持つ物理的な意味からすると、このようなことはすべきではない。なぜならば、欠損とは値を持たないと言うことであり、0と言うわけではないからである。このような問題に対して、第4章では欠損画素に対する固有空間上の逆射影（BPLP）と欠損パターンに応じた固有空間の再構築（ARLP）を提案した。これらどちらの手法でも、欠損画素に依存しない安定した認識を行うことができることが確認された。また、BPLPを応用して、固有ベクトルを知識とした欠損データの時空間補間法を提案した。この補間方法ではスプライン関数やn次関数による補間とは異なり、学習サンプルに特化された補間を可能にしている。このような特化は、一般的な補間手法にはない特徴でありBPLPを用いた補間法の評価できるところである。

パラメトリック固有空間法は、2次元画像のみで3次元物体を認識できる手法であり、その簡便さから注目を浴びている。しかし、この認識はパターン認識の手法としては有効な手法ではあるが、自律移動ロボットのための物体認識としては不十分である。なぜなら、自律移動ロボットにおいて物体認識とは、その後に物体をつかむとか物体を避けてぶつからないように軌道計画を行うなどの目的があるからである。このような目的のためには、認識された物体が何かと言うことも重要であるが、その物体までの距離を知る必要があるが、濃淡画像を用いた固有空間法では認識対象までの正確な距離を知ることができない。これに対して距離画像を用いた固有空間法では、第5章で述べた仮想平行光学RFと拡張固有空間により、物体認識・姿勢検出と同時に物体の奥行き位置検出もでき、マニピュレータで物体をつかむタスクなどにおいて有用な情報を与えることができる。このように奥行き方向の位置検出も行うことができるという特徴は、濃淡画像を用いた固有空間法では実現することが困難であり、距離画像を用いた固有空間法の最大の利点である。このように奥行き検出も可能な距離画像の固有空間法をパンステージにより回転するカメラに応用すれば、第6章で示した実験のように平面上に既知の物体が複数配置された環境においても物体を探索を行うことが可能であり、複数物体を同時に扱えるということに関しては有用な手法である。しかし、人間が日常暮らしている環境において、全ての物体を学習させるということは現実的ではない。そのため、この論文で述べた物体探索の次の段階として、カメラで撮影された物体が既知の物体であるのか未知の物体であるのかということを判断する機能が必要になる。

本論文で提案した手法は、現状では自律移動ロボットの視覚機能として用いる際に、背景問題であるとか、オクルージョン、物体の傾きへの対処など解決すべき問題は多いが、画像照合ベースの認識において距離画像を用いるという本論文の提案が、今後のコンピュータビジョン研究において新たな展開のきっかけになれば幸いである。

# 参考文献


[Agin76] G.J.Agin and T.O.Binford : "Computer description of curved objects", IEEE Transaction on Conputers, Vol.25, pp. 439-440. (1976)

[AI90] "人工知能ハンドブック", 人工知能学会編, オーム社, pp.343-929, (1990)

[Amano 95] 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, pp.103, (1995)

[Daniel 96] Daniel L. Sweets and John (Juyang) Weng : "Using Discriminant Eigenfeatures for Image Retrieval", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 8, pp.831-836,(1996)

[Deguchi 97] 出口, 野口 : "固有空間法による視覚サーボにおける制御に適した固有部分空間の再構成", 信学論, Vol. J80-D-II, No.6, pp. 1522-1529, (1997)

[Hiura 96] Shinsaku Hiura, Kosuke Sato and Seiji Inokuchi : "Range Image Processing using Graphics Accelerator", Proc. of Japan-USA Symposium on Flexible Automation, pp.955-958 (1996)

[Horand84] P.Horand and R.C. Bolles : "3DPO's strategy for matching three-dimensional objects in range data", International Conference of Robotics, pp.78-85, (1984)

[Horst 98] Horst Bischof, Ales Leonardis, "Robust Recognition of Scaled Eigenimages Through a Hierarchical Approach", Proc. of Computer Vision and Pattern Recognition, pp.664-670,(1998)

[Ikeuchi 86] Ikeuchi, K., Nishihara, K., Horn, B. K. P., Sobalvarro, P., and Nagata, S. : "Determining grasp points using photometric stereo and the prism binocular stereo system", The International Journal of Robotics Research, Vol. 5, No. 1, pp. 46-65, (1986)

[Inokuchi 90] 井口征士, 佐藤宏介: "三次元画像計測", 昭晃堂, (1990)

[Itou 94] 伊藤聡、湯浅秀男、伊藤正美: "反応拡散方程式を用いた事故連想記憶による画像認識", 計測自動制御学会論文集, Vol.30, No.1, pp.97-103, (1994)

[Matuyama98] 松山隆司、久野義徳、井宮淳(編) : "コンピュータビジョン", 新技術コミュニケーションズ, pp.204-218, (1998)

[Murase 94a] 村瀬洋, Shree Nayar : "２次元照合による３次元物体認識　－パラメトリック固有空間法－", 信学論, vol.J77-D-II, No.11, pp.2179-2187, (1994).

[Murase 94b] 村瀬　洋, S. Nayar, "パラメトリック固有空間法による３次元物体の認識とス

ポッティン グ," MIRU'94, II,pp.49-56, (1994)

[Murase 96] 村瀬　洋, S.Nayar : "アピアランスに基づく物体認識のための生成学習法 Learning-by-generation", 第3回画像の認識・理解シンポジウム, vol.I,pp.I79-I84, (1996).

[Nevatia77] R.Nevatia and T.O. Binford : "Description and recognition of curved objects", Artificial Intelligence, Vol.8, No.1, pp.77-98, (1977)

[Okatani 95] 岡谷, 出口 : "固有空間法を利用した陰影からの曲面の形状復元", 情報研資料, CV93-1, (1995)

[Perkins 77] Perkins, W. A. : "Model-based vision system for scene contacting multiple parts", Proc. 5th International Joint Conference on Artifical Intelligence, pp.678-684, (1977)

[Sakano 97] 坂野鋭, 武川直樹 : "AICによる部分空間次元数の決定法"PRMU97-173,pp.159-165, (1997)

[Sato 94] K. Sato, A. Yokoyama, and S. Inokuchi, "Realtime range Imaging using Adjustment-Free Photo-VLSI - Silicon Range FInder -, "Proc. of Custom Integrated Circuit Conference 94, pp.339-342, (1994)

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

[Sirovich87] Sirovich l. and Kirby M. : "Low dimensional procedure for the characterization of human faces", Journal of Optical Society of America, 4, 3, pp.519-524 (1987)

[Tatuno93] 辰野恭市 : "ロボットをもっと世に送り出すために -企業の一開発者の自戒-", 日本ロボット学会誌 Vol.11 No.1, pp.85, (1993)

[Turk91] Turk M. A. and Pentland A. P. : "Face Recognition Using Eigenfaces", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp.586-591 (1991).

[Wang 98] 王彩華、坂上勝彦 : "3Dモデルを用いたView-based仮想学習と3次元物体の認識", 電子情報通信学会PRMU研究会報告書, pp.25-32, (1998)

[Yachida90] 谷内田正彦 : "ロボットビジョン", 昭晃堂, (1990)

[Yokoyama 99] 横山太郎, 八木康史, 谷内田正彦 : "階層的パラメトリック固有空間構成法", 電子情報通信学会PRMU研究会報告書, pp.59-67 (1999)

# 謝　辞

本論文は、私が学部４年の終わりから博士課程前期、後期にわたり研究した成果をまとめたものです。主査をしていただいた井口征士教授には、学部4年から現在に至るまで日頃より研究内容および方針について終始細やかなご助言、ご指導を賜りました。また、御厚意により様々な発表の機会を与えていただき、少なからぬ成果を挙げることができましたことも感謝いたします。

副査をしていただいた谷内田正彦教授、西田正吾教授、佐藤宏介助教授には、公聴会にて貴重なご意見をいただきました。どうもありがとうございました。谷内田正彦教授が著した「コンピュータビジョン」は、私の研究活動全般において基礎的な知識として活用させていただきました。直接研究の指導を受ける機会がございませでしたが、佐藤宏介助教授が開発されたSRFは本論文で述べた研究の基礎となる装置であり、距離画像の固有空間法で利用することを快く承諾していただき本論文を執筆することができました。どうもありがとうございました。

加藤博一助教授（現　広島市立大学）には、私が学部4年から修士にかけてプログラミングなどの技術的な側面などで、様々な御支援をいただきました。また、私の研究に対する好意的な意見も頂き、研究活動の大きな励みになりました。どうもありがとうございました。日浦慎作助手には、学部4年次に固有空間法に距離画像を用いるという基本的なアイデアを頂きました。また、本研究に対する、肯定的な意見のみならず否定的な意見も多数いただきました。どうもありがとうございました。

パラメトリック固有空間法の発案者であるNTTの村瀬洋氏には、学会などの様々な機会で私の研究に対する的確な御指示をいただきました。また、会議報告書や書籍において私の研究を紹介していただきました。どうもありがとうございました。

電子技術総合研究所の王彩華氏、大阪大学の横山太郎氏、北陸先端大の中村知彦氏には、学会発表等で私の研究を参照して頂きました。参照して頂いたことにより、研究活動において大きな意義とやりやりがいを得ることができました。どうもありがとうございました。

同じ研究グループの井下哲夫氏には、私自身で行うことができなかった固有空間法によるトラッキングを実現していただきました。井下哲夫氏の研究成果から、固有空間の違った一面を見ることができ、固有空間法に対する新たな知見を得ることができました。また、森岡義貴氏には、距離画像と色相情報を用いた固有空間法について研究していただき、距離画像の固有空間法の欠点の指摘と、その改善方法を示していただきました。どうもありがとうございました。

その他、この場で具体的に御名前を挙げていない井口研究室のスタッフや学生の皆様にも、研究活動やその他で、大変お世話になりました。特に、夕食を御一緒させていただいたことなどは、研究室生活の楽しみの一つで、研究活動の大きな活力となりました。どうもありがとうございました。

最後に、私が博士課程後期に進学することに対して、理解と献身的な協力を頂きました両親に感謝いたします。

# 謝辞

# 業績表

[学術論文]

1. 天野敏之, 日浦慎作, 山口　証, 井口征士: "固有空間照合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D-II, Vol.J80-D-II , No.5, pp.1136-1143 (1997)
2. 天野敏之, 山口証, 井口征士: "固有空間照合を用いた距離画像の仮想学習による物体認識・姿勢検出", 電子情報通信学会論文誌D-II, Vol.J82-D-II , No.2, pp.250-258 (1999)
3. 天野敏之, 井口征士: "距離画像の固有空間法による物体探索", 電子情報通信学会論文誌D-II, (採録決定)
4. ToshiyukiAMANO, Akashi Yamaguchi, Seiji Inokuchi: "Eigenspace Approach for Object Recognition and Its Pose Detection", John Wiley & Sons, Inc., Systems and Computers in Japan, SCJ0539, (採録決定)

[解説記事]

1. 天野敏之, 山口証, 井口征士: "固有空間照合に基づくく距離画像からの物体の姿勢検出 －形状モデルを用いた仮想距離画像生成－", 画像ラボ, Vol.9, No.1, pp.10-14 (1998)

[国際会議]

1. Toshiyuki Amano, Shinsaku Hiura, Akashi Yamaguti, 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)
2. Toshiyuki Amano, Shinsaku Hiura, Kosuke SATO and Seiji Inokuchi: "Object Search and Pose Detection based on Eigenspace Approach", 2000 Japan-USA Symposium on Flexible Automation, (投稿中)

[口頭発表]

1. 天野敏之, 美濃部賢一, 澤洋一郎: "自律移動ロボットの動力学解析と走行制御(LQIおよびFUZZYによる速度・姿勢角制御)", 日本機械学会[No.930-40]　ロボティクス・メカトロニクス講演会'93講演論文集　p901-906 (1993)
2. 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)
3. 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)
4. 天野敏之, 日浦慎作, 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", 電子情報通信学会パターン認識と理解研究会報告書, 145-150 (1995)

5. 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 1 , 205-210 (1996)

6. 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)

7. 福本広次, 天野敏之, 山口証, 井口征士: "光レーダ型レンジファインダに関する一考察", 第41回システム制御情報学会研究発表講演会講演論文集, 655-656 (1997)

8. 井下哲夫, 天野敏之, 山口証, 井口征士: "エッジ情報を用いた画像圧縮", 1998年電子情報通信学会総合大会講演論文集情報・システム講演論文集 2 , 28 (1998)

9. 天野敏之, 山口証, 井口征士: "固有空間照合を用いた距離画像の仮想学習による物体認識・姿勢検出", 電子情報通信学会PRMU研究会報告書, 71-78 (1998)

10. 井下哲夫, 天野敏之, 山口証, 井口征士: "エッジ情報を利用した画像圧縮", MIRU'98講演論文集I, 129-134 (1998)

11. 井下哲夫, 天野敏之, 井口征士: "微少移動を投影した固有空間を用いた物体追跡", 電子情報通信学会総合大会論文集情報・システム 2 , pp298 (1999)

12. 天野敏之, 井口征士: "固有空間照合を用いた距離画像の仮想学習による物体認識", 電子情報通信学会総合大会論文集情報・システム2, pp299 (1999)

13. 森岡義貴, 天野敏之, 井口征士: "距離画像とカラー画像を用いた固有空間法による姿勢検出", 電子情報通信学会総合大会論文集情報・システム 2 , pp301 (1999)

14. 井下哲夫, 天野敏之, 井口征士: "微少運動固有空間を用いた物体追跡", システム制御情報学会研究発表講演会論文集, pp479-480(1999)

15. 森岡義貴, 天野敏之, 井口征士: "距離画像と色相情報を用いた固有空間法による姿勢検出", システム制御情報学会研究発表講演会論文集, pp481-482(1999)

16. 天野敏之, 井口征士: "仮想平行光学レンジファインダと拡張固有空間による物体探索", システム制御情報学会研究発表講演会論文集, pp483-484(1999)

17. 天野敏之, 井口征士: "欠損したデータの固有空間照合とその応用", 1999年電子情報通信学会ソサイエティ大会講演論文集, pp176 (1999)

18. 井下哲夫, 天野敏之, 井口征士: "背景がある環境での固有空間を用いた動物体追跡", 1999年電子情報通信学会ソサイエティ大会講演論文集, pp171(1999)

Z
(
15