

Title	移動ロボットの能動的視覚
Author(s)	石黒, 浩
Citation	大阪大学, 1991, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.11501/3054381">https://doi.org/10.11501/3054381</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

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

Osaka University

# 移動ロボットの能動的視覚

Active Vision for Mobile Robot

石黒 浩

大阪大学大学院基礎工学研究科

平成3年2月

大阪大学大学院基礎工学研究科

博士論文

移動ロボットの能動的視覚

**Active Vision for Mobile Robot**

石黒 浩

平成3年2月

## 要旨

本論文は、視覚移動ロボットの能動的視覚とその効果について述べたものである。

従来の視覚移動ロボットでは、視覚として用いるTVカメラはロボット上に固定されていた。しかし、視野の限られたTVカメラを用いる場合、環境を柔軟に観測するためには、ロボットの運動と独立なカメラ運動が必要不可欠である。また、カメラがロボット上に固定されている従来の視覚移動ロボットの研究では、ロボットの目的や環境に応じて、環境認識の方法を考えるとより、いかなるロボットの目的にも、いかなる環境にも適用可能な、一般的な環境認識の方法を開発しようとする傾向にあった。しかし、そのような一般的な環境認識の手法を開発することは不可能であり、むしろ、ロボットの目的や環境に応じたいくつかの環境認識の手法を用意し、状況に応じてそれらを使い分けるといった戦略が必要であり、それらの手法では、異なるカメラ運動が用いられるべきである。

本論文では、視野の限られたTVカメラを用いることを前提として、ロボットの目的や移動する環境から、どのようなカメラ運動が必要とされるかを考え、そのカメラ運動として

- (1) 環境の一部を注視する動作
- (2) 全環境を見渡す動作

の2つについて考察する。

環境の一部を注視するカメラ運動は、ロボットが環境に対してある程度の知識を持っており、全環境の精密な計測を必要としない場合に有効である。例えば、そのような状況として、ロボットが廊下を移動する場合が考えられる。

## 要旨

カメラは環境内より詳しく観測する必要がある部分を選択し、ロボットの移動中、その部分(注視点)に注視する(注視制御)。注視することによって、その注視点を中心とした正確な局所地図を得ることができる。この動作は、最近の話題である能動的視覚(Active Vision)のパラダイムに深く関係するものである。本論文ではカメラによる注視制御を、ロボットの環境認識に取り入れることを提案し、注視制御による局所地図の獲得方法、注視制御によって獲得される局所地図の誤差解析、従来の方法との比較を行う。また、実験では、連続的観測を前提とした局所地図を計算する方程式をもとに、カルマンフィルタを用いて、観測結果を融合する方法を提案する。

一方、全環境を見渡す動作は、ロボットが未知環境にさしかかった時や、環境内で物体を探索する場合に必要となる動作である。ロボットは、全環境を見渡しながらか、環境の正確な全体地図を作成する。

全環境を見渡す動作によって得られる全方位画像は、正確な方位角情報と全方位の距離情報を持つ。正確な方位角情報は、カメラを回転させて得られる全方位画像の重要な特徴の一つである。また、全方位の距離情報は、単眼全方位ステレオと名付ける方法により得られる。

全方位画像を用いて環境の全体地図を作成する手順は以下のものである。まず最初に、ロボットの移動経路に沿って全方位画像を離散的に撮像する。次に、隣合う2つの全方位画像を用いて2眼全方位ステレオを行い、局所地図を作成する。最後に、得られた複数の局所地図を統合することで、全環境の正確な全体地図を作成する。ここで、局所地図の作成と全体地図の作成をより正確に行うために、全方位画像からの視覚フィードバックを用いた、ロボットの移動方法を提案する。また、この全体地図を獲得する過程では、ロボット中心、経路中心、環境中心の3つの環境表現が用いられる。

視覚移動ロボットの独立なカメラ運動は、ロボットの目的や移動環境に応じた環境認識の戦略を導くだけでなく、観測(局所地図や全体地図の獲得)において、ロボットの正確な運動パラメータの計測を必要としないという利点を持つ。すなわち、完全ではないが(環境の一部に注視するカメラ運動

## 要旨

ではロボットは直進運動をしなければならない。また、全環境を見渡すカメラ運動では、より完全な全体地図が得られるように、ロボットは環境内を移動しなければならない)、ある程度、ロボットの移動と、環境の観測を切り放せることにある。

視覚に用いるTVカメラの運動を考えることは、より柔軟な環境認識の実現をめざす視覚移動ロボットの研究の重要な課題であり、本論文はその一部をなしたと確信する。

## 目次

要旨

目次

第1章 序論	1
第2章 ロボットの視覚に関する研究	11
2.1 能動的センシング	13
2.2 受動的センシング	18
2.2.1 屋内環境	18
2.2.2 屋外環境	27
2.2.3 画像認識に共通する問題	30
第3章 環境の一部を注視するカメラ運動	39
3.1 注視制御に関する研究	40
3.2 注視制御を用いた環境認識	42
3.3 ロボットとカメラの運動に関する仮定	45
3.4 連続的観測における局所地図の獲得	46
3.4.1 局所地図の計算方法	46
3.4.2 誤差解析	48
3.4.3 実験結果	51
3.4.4 考察	53
3.5 離散的観測における局所地図の獲得	59
3.5.1 局所地図の計算方法	59

## 目次

3.5.2 誤差解析	59
3.5.3 実験結果	61
3.5.4 考察	62
3.6 注視点選択に関する考察	68
3.7 まとめ	73
第4章 環境を見渡すカメラ運動	75
4.1 環境の全体地図作成に関する研究例	76
4.2 全方位画像を用いた環境認識	78
4.3 ロボットと環境に関する仮定	80
4.4 全方位画像	81
4.4.1 全方位画像に関する研究例	81
4.4.2 全方位画像の撮像機構	82
4.4.3 正確な方位角情報の獲得	84
4.4.4 全方位の距離情報の獲得	96
4.4.5 全方位画像への投影	105
4.4.6 考察	108
4.5 全体地図の作成	110
4.5.1 能動的全方位視覚を用いた移動	110
4.5.2 局所地図の作成方法	112
4.5.3 全体地図の作成方法	114
4.5.4 実験結果	115
4.5.5 考察	116
4.6 まとめ	124
第5章 結論	125



## 目次

謝辞	127
文献	129
研究業績	137
著者略歴	139
付録	141
A. 一般的なロボットの運動における注視制御	141
B. ロボットの運動とカメラの運動の関係	146
C. 注視制御におけるカメラの回転角速度	148
D. 注視制御におけるオプティカルフロー	149
E. 円弧運動における注視制御の誤差解析	151
F. カルマンフィルタの設計	153
G. 全方位画像の解析	155
H. 回転中心の位置合せについて	160

## 第1章 序論

### ロボットの研究について

人間は非常に優秀なロボットであり、人類が初めて道具を使うようになって以来、いつかはこの優秀なロボットと同じものを機械で実現したいと考えてきた。ロボットの研究は、自動機械を作ることを目標に始まり、コンピュータが発明され、人工知能が一つの研究分野として独立してから、2つの流れに分かれていった。一つは自動機械を作ることを目標とするそれまでの流れに従う、主に機械工学の分野の研究である。もう一つは、人工知能の分野において実行例に用いられるロボットであり、知能をインプリメントする対象としてのロボットである。さらに、人工知能の研究は人間の知能を機械で実現しようとする工学的立場と、人間の知能のメカニズムを解明しようとする科学的立場にわけられる。

人工知能の研究が始まると同時に、その一部である視覚機能の研究も始まった。人間は環境からの情報の90%を視覚によって獲得するといわれるように、人間の機能を目指すロボットにとっても、視覚は非常に大切な機能である。ロボットの視覚の研究も、視覚を一つのセンサとして考える工学的立場と、人間の視覚機能そのものを解明しようとする科学的立場にわけることができる。ロボットビジョンとコンピュータビジョンという2つの言葉があり、2つの言葉の区別は明確ではないが、筆者自身はそれぞれ工学的な立場と科学的な立場からの研究を代表する言葉と考えている。

このように、ロボットと視覚の研究はそれぞれ工学的立場と科学的立場にわけることができ、視覚を兼ね備えたロボットを考えた場合も同様である(以後ロボットには視覚の機能が含まれるものとして議論を進める)。この2つの立場には大きな隔たりがあるように思う。

工学的立場の研究では、まず最初に研究者が目指す目標、目指すロボットの機能があり、ロボットはその機能を実現するのに必要な情報をセンサより取り込み、その機能を実現するのに必要十分な内部表現を用いる。一方、科学的立場の研究では、目標は常に人間であり、人間のように情報を入力し、人間と同等の内部表現を実現しようとするのである。

科学的立場のロボットの研究は、知能に対する大きな仮説(例えば画像認識における Marr のパラダイム [Marr 82])に始まり、それに基づいて作り上げられたシステムと人間の機能との比較を行ったり、そのシステムにより知能のメカニズムを検証することが目的であり、最終的にある仕事を達成するロボットを作り上げることを直接の目的としない。これに対して、工学的研究のヒントの多くは科学研究より生まれ、工学的立場のロボットの研究は、センサ、環境計測の方法、ロボットの制御方法といったものを、それらのヒントを基に具体化し、最終的に何かの機能を果たすロボットを作り上げる。

しかし、近年の研究ではこれら両者の関係が、かなり接近してきていると思われる。環境認識の機能がすぐれ、様々な目的を果たす便利なロボットを作ろうとした場合、手本となるのはやはり人間であり、工学的立場からのロボット研究も、科学的立場で人間の機能を研究した後に、工学的に応用できる部分をロボットに実装するといったアプローチが必要である。また、科学的立場で人間の機能を解明しようとした場合にも、実際にロボットに実装し、ロボットを動かして初めて次の段階の研究に移れるということもあるだろう。

#### 視覚移動ロボット研究の問題点

視覚移動ロボットの研究は、視覚により環境を認識し、環境内を移動するロボットの実現を目標とする研究であり、特に工学的研究と科学研究の接触が必要とされる分野である。

視覚移動ロボットの手本を人間と考えた場合、これまでの視覚移動ロボ

ットは、人間からあまりにもかけ離れたものであり、まだまだ、人間に習う部分が多いと思われる。人間との違いは知識情報処理のレベルはさておき、運動機能における違いも大きい。人間において、身体性が環境認識や環境表現を制約するように、ロボットにおいても、その運動性能により環境認識の方法や環境の表現方法が制約される。

人工知能の研究が進む中で、ロボットによる環境認識(例えば[Faugeras 90, Thrope 88]など)や、環境の表現方法(例えば[Elfes 87, Yeap 88]など)に関する研究も数多くなされてきた。これらの研究に用いられた移動ロボットでは、視覚に用いるカメラはロボット上に固定され、ロボット本体は床面上での2次元運動が許されているといった形態をとっていた。しかし、より柔軟で効率の良い環境認識を実現するには、ロボットの身体性の拡張、すなわち、より自由度の高い運動能力をロボットに与える必要がある。

#### 体と首の独立な運動

視野の限られた目を持つ人間にとって、環境認識における重要な機能の一つは、体と首の独立な運動である。

人は、歩行とは独立に、首を運動させながら環境を認識する。例えば、廊下を移動する場合、廊下に関する知識はあらかじめ持っており、その単純な構造のため、廊下自体の形状は頻繁に観測する必要はない。視覚はむしろ、廊下内にある移動体や未知の物体、または廊下の折れ曲がる地点でコーナなどを注目するため、体は常に前進しても、首だけは廊下の各部分に注視する運動を繰り返す。一方、人が未知な部屋に入った場合はどうだろうか。人は部屋全体を見渡すことで環境全体の構造を知ると同時に、その部屋の中で自分がどこにいるのかを知る。このときも、首の運動は360°全方向を見渡そうと動くのに対し、体はさほど動かない。

このように、人が環境内を移動するためには、首と体の独立な運動は必要不可欠である。人において首と体の運動が独立に行われるのは、次の理由によると考える。

一つは、方向感覚が狂うのを防ぐためである。環境内をくまなく探索する視覚にあわせて体を運動させていたのでは、すぐに方向感覚を失ってしまうことは明かであろう。もう一つは、環境内の移動と視覚による環境の探索は、必ずしも同時に行う必要がないことである。もちろん環境内を移動するには、視覚からの情報が必要となる。しかし、移動の際に用いる視覚は、移動可能な自由領域を離散的に探索すればよく、常に移動する地点を見つめ続けるといった動作は必要ない。それよりも視覚は、環境内の他の部分の観測にほとんどの時間を費やしていると考ええる。

### 移動ロボットの能動的視覚

上述の人間における体と首の独立な運動は、そのまま視覚移動ロボットに取り入れることができる。この運動を取り入れることで、視覚移動ロボットは、より柔軟な環境認識の機能を持つことになる。視覚移動ロボットにおいて、人間の体と首に相当するものは、車輪で移動するロボット本体と、回転台によって回転するTVカメラに相当する(図1.1)。本論文は、視覚移動ロボットに、車体とは独立なカメラ運動を許した場合の効果について述べる。独立なカメラ運動を取り入れる利点は、環境を探索する際に、車体の運動を最小限に抑えられるため、車体の運動パラメータの累積誤差が少なくなることと、ロボットの移動と環境の観測が切り放せるため、従来よりも柔軟な環境認識が実現できることである。

ロボットの場合も、カメラの運動は

- (1) 環境の一部に注視する運動
- (2) 全環境を見渡す運動

の2種類に分類することができる(図1.2)。これら2つのカメラ運動は、ロボットに与えられる命令とロボットの移動する環境によって、人間の場合と同様に使い分けることができる。環境の構造が比較的簡単で、環境内での探索をあまり必要としない命令が与えられている場合、ロボットは移動に必要な情報(自由領域等)はカメラから離散的に入力すればよい。むしろ、

カメラは他の注目し認識する必要のある物体を見るべきである(環境の一部に注視する運動)。これに対して、ロボットに環境内での探索命令が与えられたり、ロボットが未知環境にさしかかった場合には、ロボットは全環境を見渡し、環境を綿密に調べることで目的物を探し出したり、環境の正確な構造を調べ、自分と環境の位置関係を明確にする必要がある(全環境を見渡す動作)。本論文ではこのような2つのカメラ運動についてその効果と、それをを用いた環境認識の方法について述べる。次にそれらの概要を説明する。

#### 環境の一部に注視する運動 (図1.2a)

3章では、環境の一部に注視するカメラ運動について述べる。ここで述べる研究は、近年注目を集めている能動的視覚(Active Vision)のパラダイムに基づくものである。後に3章で説明するように、能動的視覚のパラダイムに基づく研究は、既に幾つか報告されているが、本論文では、ロボットが直進や円弧などの拘束された運動を行う際、ロボット上に取り付けられたカメラ回転運動により注視制御を行う場合について述べる。

3章ではまず、視覚移動ロボットの環境認識に、注視制御を用いる戦略を提案する。この戦略は、ロボットは移動中に環境内の興味ある部分に注目し、その部分のみの正確な局所地図(注1)を作成することで、ロボットに必要な情報のみを記録するものである。ここでの中心的话题は、局所地図を得るための連続的な観測をもとにした方程式の導出、カルマンフィルタを用いた局所地図の獲得方法、注視制御により得られる局所地図の精度に関する考察等である。

注視制御を移動ロボットの視覚に取り入れる利点は、(1)注視点を中心とする物体中心の座標系で局所地図を獲得することができる。(2)十分長い

---

(注1) 本論文の実験では、局所的または全環境に渡って存在する、特徴点の2次元位置を示したものを、局所地図または全体地図と呼ぶ。

間、注視点付近を観測できる。(4)注視制御を取り入れた視覚移動ロボットの新たな環境認識の方法を生むことである。

#### 全環境を見渡す運動 (図1.2b)

次に4章では、全環境を見渡すカメラ運動について述べる。全環境を見渡すカメラ運動より獲得される全方位画像は、全方位の正確な角度情報と、全方位の距離情報を含む。この全方位画像は、環境の正確な構造を復元するために用いられる。

全方位画像を用いて、環境の正確な構造を復元する手順は次のようである。まず最初に、ロボットの移動経路に沿って、全方位画像を離散的に撮像する。このとき、全方位画像に含まれる距離情報をもとに、ロボットは、移動可能な自由領域を探索しながら移動する。次に、隣合う2つの観測点で得られた全方位画像を用いて、ステレオを行うことで、局所地図を作成する。最後に、得られた局所地図を統合して、全体地図を獲得する。

局所地図の作成と、局所地図の全体地図への融合においては、ロボットの正確な運動パラメータが要求される。このため、全方位画像からの視覚フィードバックを用いて、ロボットの運動を正確に制御する方法を提案する。この方法では、環境に対するロボットの向きを正確に知ることができる。

またここでは、単眼全方位ステレオと2眼全方位ステレオの、2種類のステレオ法を提案する。先に、全方位画像には、正確な方位角情報と、全方位の距離情報が含まれると述べたが、単眼全方位ステレオはこの全方位の距離情報を、1台のカメラを回転させることにより獲得する方法である。一方、2眼全方位ステレオは、局所地図を得る際に2つの全方位画像を用いて実現されるステレオ法である。

さらに、全方位画像を用いて環境の全体地図を得る過程では、観測者中心、経路中心、環境中心の3つの環境表現を用いる。

環境の一部を注視するカメラ運動と全環境を見渡すカメラ運動のどちらにおいても、視覚移動ロボットの能動的なカメラ運動は、環境の観測とロボットの移動の直接的なつながりを切り放し、より柔軟な環境認識を実現する基本的なロボットの機能である。



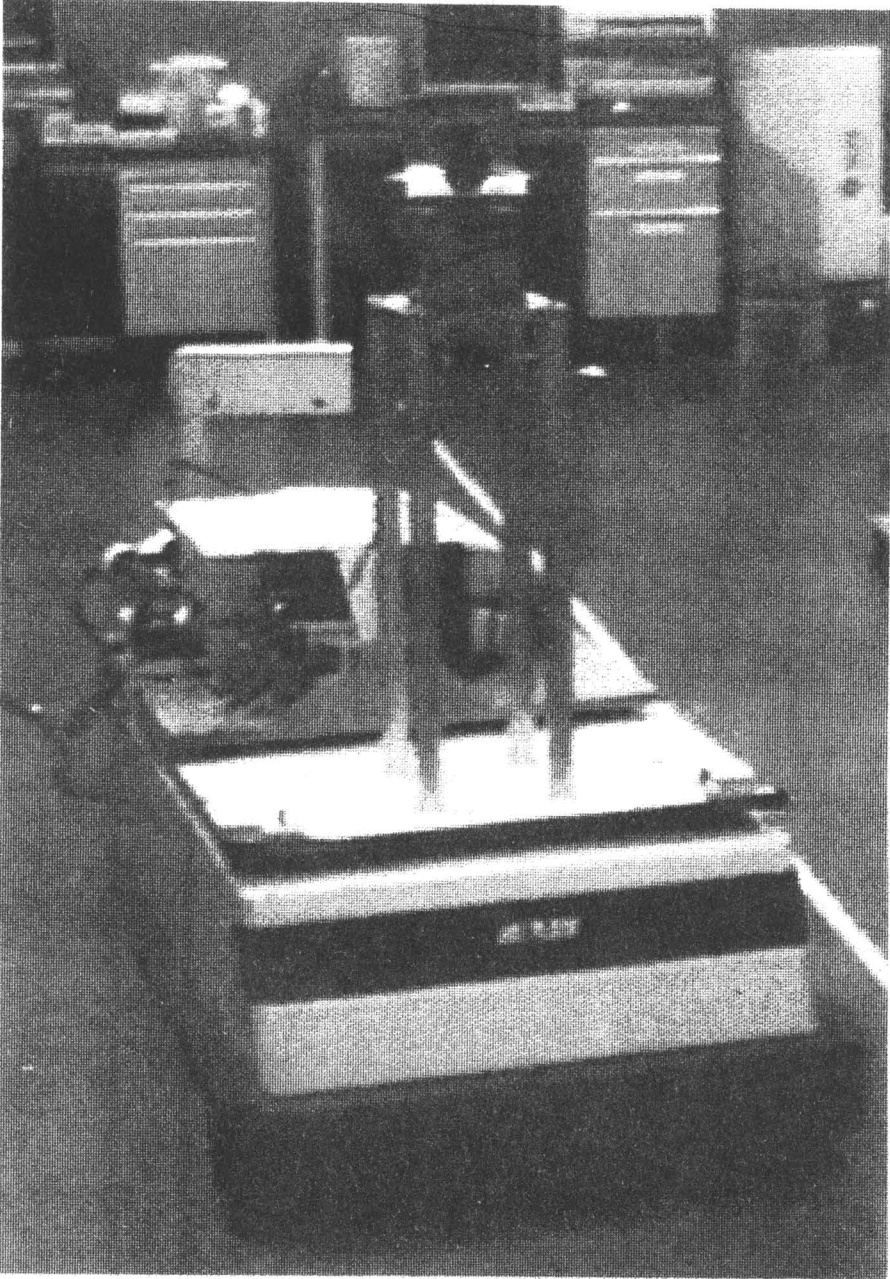
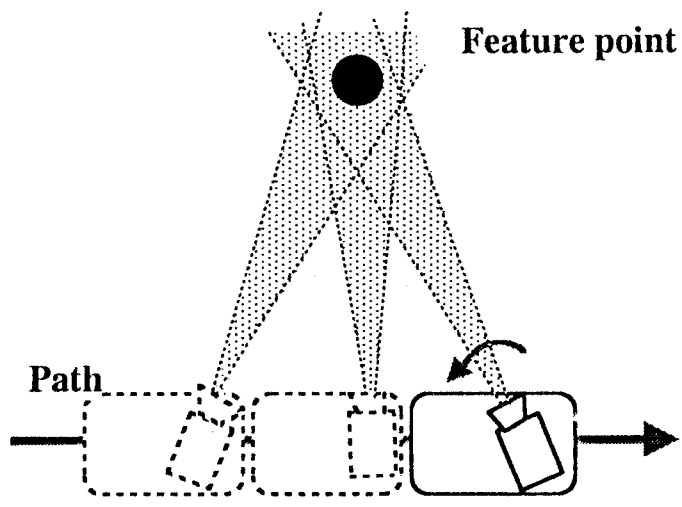
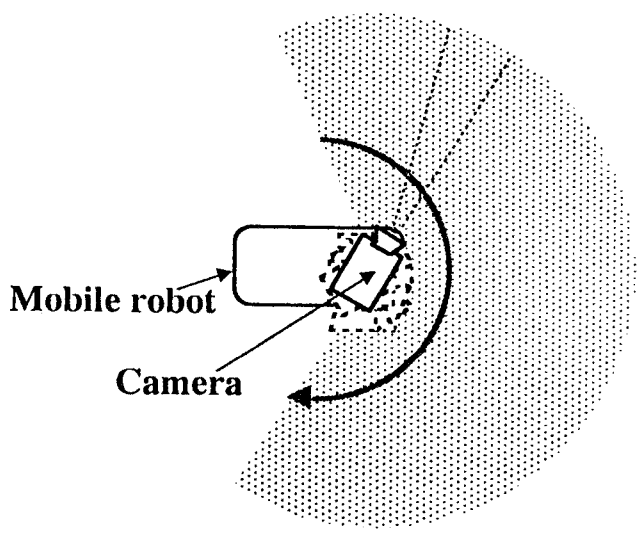


図1.1 視覚移動ロボット



(a) 環境の一部に注目するカメラ運動



(b) 全環境を見渡すカメラ運動

図1.2 2種類のカメラ運動



## 第2章 ロボットの視覚に関する研究

本章では、移動ロボットの視覚に関する研究を傍観し、研究の流れを把握するとともに、これまでの研究における問題点を考察する。

ロボットビジョン研究の最近の特徴は、移動ロボットの研究の進展と、画像処理装置の進歩により、単なる物体認識から、移動ロボットのための環境認識へと、研究目標が移ってきたことである。従来は、物体を対象とする画像認識のアルゴリズムの研究が多かったが、最近の画像処理装置の高速化、大容量化により、実際の環境での実験が行われるようになってきた。実際の環境下での研究は、移動ロボットの視覚の研究につながる。

視覚移動ロボットの研究では、それまでの物体認識にはない多くの問題点がでてきた。例えば、

- (1) カメラは環境内を自由に運動するため、高速な動画像処理の機能が必要となる。
  - (2) カメラの運動パラメータを正確に求めることが難しい(移動ロボットの運動制御は、車輪のスリップなどにより、必ずしも正確ではない)。
  - (3) 広範囲の環境認識を行わなければならないので、より抽象的で操作性のよい環境表現が必要となる。
  - (4) ロボットの視覚としての機能を満たすには、画像認識に多大な時間がかかるのは好ましくない。そのため、環境に対する知識を使用し、より頑強で処理の簡単な画像認識アルゴリズムを開発しなければならない。
- 等である。以下では、これらの問題点をふまえながら、移動ロボットの視覚に関する研究を傍観する。

移動ロボットの視覚に関する研究は、いろんな基準で分類することがで

きるが、ここでは、光などを対象物体に照射する能動的センシング (Active Sensing) と、カメラから得られる情報のみを用いる受動的センシング (Passive Sensing) に分けて述べる。受動的センシングはさらに、ロボットの移動する環境によって、屋内環境と屋外環境にわけて述べる。また、全てに共通する問題については、後半まとめて討論する。なお、注視制御と全方位画像に関係の深い研究については、3章、4章に新たに節を設けて説明をする。

## 2.1 能動的センシング

人間は、視覚によって環境の3次元構造を巧みに理解することができる。3次元環境で生活する人間にとって、視覚は3次元構造を理解するために、なくてはならないものである。人間と同様に、ロボットにおいても、環境の3次元構造を理解する機能は、3次元環境で行動するために重要なものであり、多くの研究が行われてきた。

1章で述べたように、視覚の研究は工学的立場の研究と、科学的立場の研究にわけることができる。人間の視覚にとらわれずに、様々な手段により環境の3次元環境を理解しようとする研究は、工学的立場の研究と言える。工学的立場の研究には、対象物体に何らかのパターンを投光したり、超音波センサや、レンジファインダを用いて、環境の3次元構造を直接計測する能動的センシング(Active Sensing)と呼ばれる方法がある。

能動的センシングは、物体に光やレーザを投光するため、環境を十分整える必要があり、そのため、作業台上での物体認識など、範囲の限られた環境への応用が多かった。しかし最近では技術の進歩により、屋外移動ロボットの視覚[Thorpe 89]にも用いられるようになってきた。

能動的センシングは主に、(1)物体にパターン光を投光するもの、(2)超音波を利用するもの、(3)レーザ光を利用するものの3通りがある。

物体にパターン光を投光する方法は、物体に投光されたパターン光をTVカメラで撮像し、投光したパターン光までの距離を計測するもので、画像中でパターンが正確に抽出できなければならない。このため、日中の太陽の照明下や、屋内でも明るい照明に照らされている場合には不可能な方法である。そこで、残る(2)、(3)の方法がロボットの視覚として用いられる。

### 超音波の利用

超音波センサを用いた距離計測方法は、機構が単純で安価なため、多くの研究で用いられている。

超音波センサの特徴としては

- (1) ロボットから遠い物体までの距離は測定できない。
- (2) 距離測定の精度はあまりよくない。
- (3) 距離測定の精度は測定対象の形状に大きく依存する。

等があげられる。このため、応用範囲はかなり限定され、超音波センサによりロボットが移動できる環境は、屋内の壁などで覆われている場所に限られる。実際のロボットに搭載され、役立っている例としては、廊下の大まかな形状を計測し、あらかじめ与えら得ている廊下の地図をもとに、自らの位置決めを行う研究 [Matumoto 86] や、ロボットの移動可能な自由領域を発見する研究 [Matthies 88] などがある。

しかし、超音波センサは限られた環境でしか使用できないにしても、機構が安価で単純なことは、とりあえずロボットを作成するには便利であり、実際の環境で動くロボットを実現するというよりも、環境から得た情報を処理するロボットの形態を考える研究等に多いに用いられている。例えば、Elfes [Elfes 87] は、超音波センサによる観測結果をもとに、抽象度、分解能、地域性の3つの尺度において、それぞれ3つのレベルで異なる環境表現を提案し、移動ロボットの環境認識への利用を示している (Elfesの研究については2.2.3で述べる)。

#### レーザ光の利用

レーザ光を用いたレンジファインダ [Lougheed 88] は、その機構の改良により、屋外環境でも用いることができるようになってきた [Thorpe 89]。レーザ光を用いたレンジファインダは、超音波センサより距離測定の精度は高く、対象物の形状に影響されることは少ない。しかし、いくつかの問題点が残されている。

最も大きな問題は、装置全体が大がかりなことであろう。現在では小型化のための研究がいろいろ行われているようであるが、TVカメラほどではない。第二に、密な距離画像が得られない点である。ELIMのレンジファイ

ンダ [Lougheed 88] では、1/2秒で  $256 \times 64$  点の距離画像が得られるが、移動ロボットの視覚として考えた場合、この値は決して密でもなければ高速でもない。第3にはレーザ光を吸収する物体(水など)にはもちいることができない点である。

しかし、特別な環境では有利な方法となる。例えば、火星探査ロボットを考えてみよう。火星探査ロボットは、高速に移動する必要はなく、十分な時間をかけて、火星表面を移動し、岩石などのサンプルを収拾すればよい。CMUのKanadeを中心とするグループは、火星探査ロボットの視覚としてレーザレンジファインダを用いることを提案している。彼らはロボットの複数の地点で得られた距離画像を高さ地図(elevation map)に変換、融合する方法を示した [Kweon 90]。図2.1.1は複数の地点からの観測を融合して得られた、全体的な高さ地図である。レンジファインダを用いて、地形を観測した場合、距離計測の誤差は、レンジファインダの光軸方向に広がるのであるが、彼らは、これをある平面からの高さの誤差に変換し、複数の地点からの観測結果をつなぎ合わせている。

レーザを用いたレンジファインダを用いて距離画像を作成する場合と、2つのカメラで撮像したステレオ画像をより距離画像を作成する場合を比べると、

- (1) 距離画像の密度はレンジファインダが劣る。
- (2) 2枚の画像を用いるステレオ法と同程度に密な距離画像を求めるなら、レンジファインダの処理速度は必ずしも速くない。
- (3) 2枚の画像を用いるステレオ法は、対応の曖昧さ、画像の角度分解能の粗さ等より、一点の距離測定精度は、レンジファインダの方が安定である。
- (4) レンジファインダから距離画像を作成する計算量は、2枚の画像を用いるステレオ法を用いて同じ大きさの距離画像を作成する場合よりも、多いと言われている。

上記の理由より、レンジファインダは移動ロボットの視覚として用いるに



は、まだまだ改良する必要がある。

能動的センシングの例として、超音波センサとレーザ光を用いたレンジファインダについて述べたが、どちらも移動ロボットの視覚に用いるには不十分な点が多い。環境を整えることができる工業的な応用では能動的センシングは有効であるが、一般的な環境を移動するロボットの視覚には、TVカメラから入力される画像のみを処理の対象とした受動的センシングの手法が期待される。TVカメラから入力される画像からは、明度、彩度、色相等の情報や、図形の2次元的な形の情報の他に、動画像処理やステレオ法により、環境の3次元構造を復元することができる。次は、受動的センシングについて述べる。

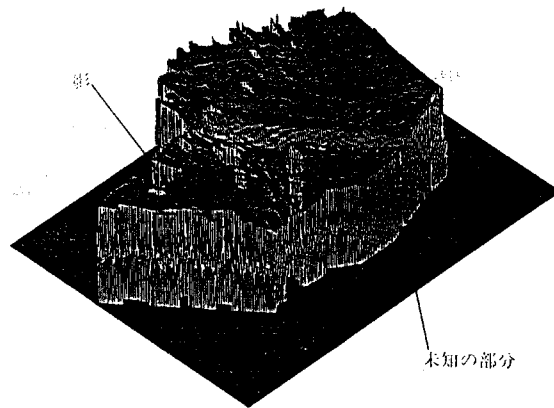


図2.1.1 レンジファインダにより獲得された  
地形の高さ地図 [Kweon 90]

## 2.2 受動的センシング

画像理解の研究が進み、最初は個々の物体を認識対象としていた研究が、環境認識に応用されるようになってきた。それらの研究は、ある特定の環境(能動的センシングの環境よりもずっと広く、変化の激しいものであるが)に依存した手法を用いることが多いため、ここでは、屋内環境と屋外環境分けて述べる。また、屋内環境を対象とした研究は、(A)物体認識の研究を環境認識へ拡張したものと、(B)移動ロボット固有の研究に分けられる。

### 2.2.1 屋内環境

#### (A) 物体認識から環境認識へ

屋内環境の認識は、物体認識に用いた手法がそのまま使えることが多く、研究例も数多くある。それらは、(1)幾何学的拘束、(2)時空間画像、(3)オプティカルフロー、(4)環境モデルをそれぞれ利用したものに分類することができる。

#### 幾何学的拘束

移動ロボットの Passive Sensing では、距離測定の原理に、異なる観測点で撮像した複数の画像によるステレオ視を用いるものが多い。ステレオ視の最も大きな問題は、画像間での対応をいかに正確に求めるかということである。連続画像を用いる場合はともかく、2つの固定されたカメラから入力された画像や、ロボットの移動中に長い時間間隔をおいて撮像された、視差の大きな画像間の対応はかなり難しい。

この問題を最も確実に解決する方法は、3つのカメラを用いた3眼視ステレオ法 [Kitamura 87] であろう。図 2.2.1 に 3眼視ステレオの原理図を示す。3眼視ステレオでは、任意の2つの画像間での対応探索結果を、第3のカメラから得られた画像を用いて検証する。まず最初に3つのカメラを設置し、キャリブレーションを行う。次に、2つのカメラで従来の方法と同様にエピ

ポーラライン上で対応探索を行う。このとき、オクルージョンや画像上のノイズなどの影響により、一点Pに対する対応候補は、時として $p_1$ 、 $p_2$ 、 $p_3$ のように複数個現れる。そこで、対応候補それぞれについて、3次元位置を計算し、計算した3次元位置をもとに、第3の画面上での位置 $q_1$ 、 $q_2$ 、 $q_3$ を求める。このとき、真の対応候補については、第3の画面上で、計算された位置の近くに同じような特徴を持った点が存在する。これを繰り返すことで、2枚の画像間での正確な対応を求めることができる。この3眼視ステレオ法は、カメラの配置や、対応単位とする画像内の特徴により、多くの利点を生む [Ejiri 90]。

3眼視ステレオ法は、ロボットの移動中に撮像された3枚の画像を用いても行える。ロボットの視覚に用いた場合は、逆に対応探索の結果を用いて、ロボットの運動パラメータの正確な評価ができる [Suzuki 89]。

### 時空間画像

複数画像間での対応問題を解決するもう一つの方法は、画像を連続的に撮像し、時空間画像を作成する方法である。画像を十分短い時間間隔で撮像すれば、連続する2枚の画像間での物体の動きは微小となり、簡単な相関法 (SSDA 等) を用いて容易に追跡することができる。Matthies ら [Matthies 89] は、時空間画像全体の作成はしないが、カメラを微小に動かすことでステレオの対応問題を解決している。彼らの方法は一般にスライダステレオと呼ばれる。

カメラの運動中に撮像された時空間画像は、2次元の画像情報だけでなく、時間情報も含む。すなわち、カメラにより獲得できる情報の全てを含んでいる。この時空間画像の性質を明らかにしたのは Bolles ら [Bolles 87] である。彼らは、時空間画像を、カメラの直進運動から決定されるエピポーラ平面で切断し、その平面上に描かれた物体の軌跡より、物体の3次元構造が復元できることを示した。カメラが直進運動する場合には、エピポーラ平面上の物体の軌跡は直線となる。静止環境内でカメラのみが運動してい

ると仮定すると、この直線の傾きはカメラの運動速度を表す。カメラの運動がわかると、物体の相対的な3次元位置は一意に決定される。

#### オプティカルフロー

Bollesらの時空間画像の解析 [Bolles 87] では、カメラの直進運動から生じる、物体の画面速度よりカメラ運動と環境の3次元構造の関係を求めたが、この問題をより一般的なカメラ運動に拡張したのが、Passive Navigation [Bruss 83] または Structure from Motion と呼ばれる方法である。静止環境内でカメラを移動させると(回転運動と並進運動のどちらも含む)、カメラの運動に伴って、画面上の物体の投影も運動する。この画面上での物体の動きをオプティカルフローまたは物体の画面速度と呼び、Passive Navigation ではこれを計測することにより、環境の相対的な3次元構造を復元する。逆に、環境構造が既知である場合は、カメラの運動パラメータを求めることができる。

Passive Navigation とモーションステレオを比較すると、前者は連続的な観測をもとに、物体の画面速度より、ロボットの移動速度と環境の3次元構造の関係を求める。これに対し後者は、離散的な観測をもとに、画面上の物体の運動視差より、ロボットの絶対的な移動量と環境の3次元構造の関係を求める。

3章で述べる注視制御では、局所地図を得る方程式を Passive Navigation [Bruss 83] をもとに導く。Passive Navigation は本論文と深く関係する手法であるので、以下に文献 [Horn 87] より引用する。

静止環境内でカメラが運動すると仮定する。カメラを中心とし、カメラの光軸とZ軸が一致する座標系 O-XYZ を想定する(図 2.2.2)。一般に剛体運動は、平行移動の成分と、座標原点を中心とした回転運動に分解できる。そこで、カメラの運動の並進速度成分を  $\mathbf{t}$ 、回転かく速度成分を  $\boldsymbol{\omega}$  で、環境内の一点 P のカメラ中心の座標系における座標を  $\mathbf{r}=(X,Y,Z)^T$  で表す。

このとき、点PのXYZ座標系に対する速度は

$$\mathbf{V} = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{r} \quad (2.2.1)$$

$\mathbf{t}, \boldsymbol{\omega}$  をそれぞれ

$$\mathbf{t} = (U, V, W)^T, \quad \boldsymbol{\omega} = (A, B, C)^T \quad (2.2.2)$$

と定義すると、点Pの角速度成分は

$$\begin{aligned} X' &= -U - BZ + CY \\ Y' &= -V - CX + AZ \\ Z' &= -W - AY + BX \end{aligned} \quad (2.2.3)$$

となる。ここで'は時間微分を表す。

画像中の各点のオプティカルフローは、その点の明るさの瞬間的な変化である。画像上での点の位置を $(x, y)$ で表し、物点Pが画像上の点pに中心射影により投影されると仮定すると、点pの各座標は

$$x = \frac{X}{Z}, \quad y = \frac{Y}{Z} \quad (2.2.4)$$

で表される(ここでカメラの焦点距離は1とする)。また、点 $(x, y)$ のオプティカルフローを $(u, v)$ で表すと

$$u = x', \quad v = y' \quad (2.2.5)$$

である。式(2.2.5)は、式(2.2.4)の $x, y$ を時間微分することで、式(2.2.3)を用いて式(2.2.6)で表せる。

$$u = \frac{X'}{Z} - \frac{XZ'}{Z^2} = \left(-\frac{U}{Z} - B + Cy\right) - x \left(-\frac{W}{Z} - Ay + Bx\right) \quad (2.2.6)$$

$$v = \frac{Y'}{Z} - \frac{YZ'}{Z^2} = \left(-\frac{V}{Z} - Cx + A\right) - y \left(-\frac{W}{Z} - Ay + Bx\right)$$

さらに、式(2.2.5)は

$$\mathbf{u} = u_r + u_t, \quad \mathbf{v} = v_r + v_t \quad (2.2.7)$$

のように表すことができる。ここで、 $(u_r, v_r)$ はオプティカルフローの並進成

分、 $(u_r, v_r)$ は回転成分であり、それぞれ式(2.2.8)のように表せる。

$$\begin{aligned} u_l &= \frac{-U+xW}{Z}, & u_r &= Axy-B(x^2+1)+cy, \\ v_l &= \frac{-V+yW}{Z}, & v_r &= A(y^2+1)-Bxy-Cx \end{aligned} \quad (2.2.8)$$

式(2.2.8)より、画像上の5点のオプティカルフローが既知であれば、

$$\frac{U}{W}, \frac{V}{W}, A, B, C \quad (2.2.9)$$

の5つの値を求めることができる。

オプティカルフロー $(u, v)$ の求め方は、2枚の画像間で対応探索を行って求める対応法と、画像の時間微分を用いる勾配法の2通りがある。

対応法は2枚の画像間で点単位の対応を求めるため、2枚の画像間の時間間隔より、直接 $(u, v)$ を決定することができる。さらに、2枚の画像間の撮像時間間隔を長くすれば、長いオプティカルフローが求められるので、オプティカルフローの計測精度はよい。しかし、対応探索を行う必要があるのが欠点である。

一方、勾配法は、十分短い時間間隔で撮像された画像に対して適用できる方法であり、画像の時間微分を求めるため(エッジ検出の1次微分オペレータと同様のものを時間軸方向に適用する)、対応問題がない。しかし、 $(u, v)$ の2つの成分を直接求められない。これについて以下に説明する。

画像が空間的、時間的になめらかな変化をすると仮定し、画像強度関数 $f(x, y, t)$ を考える。ここで、 $x, y$ は画像の空間的位置を表し、 $t$ は時間を表す。この関数をテイラー展開し、高次項を無視すると式(2.2.10)を得る。

$$f(x+dx, y+dy, t+dt) = f(x, y, t) + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial t} dt \quad (2.2.10)$$

ところで、ある時刻  $t+dt$  における画像が時刻  $t$  における原画像を、 $dx$  および  $dy$  だけ平行移動したものであると考えると、式(2.2.11)が成り立つ。

$$f(x+dx, y+dy, t+dt) = f(x, y, t) \quad (2.2.11)$$

したがって、式(2.2.10)と式(2.2.11)より、式(2.2.12)が得られる。

$$-\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \quad (2.2.12)$$

ここで、 $\partial f/\partial t$ 、 $\partial f/\partial x$ 、 $\partial f/\partial y$ 、はそれぞれ画像の  $x$  方向の微分、 $y$  方向の微分、時間方向の微分であり、全て測定可能な量である。また、 $dx/dt$ 、 $dy/dt$  は、オプティカルフローの2つの成分  $u$ 、 $v$  であり、式(2.2.12)は、

$$-\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} u + \frac{\partial f}{\partial y} v \quad (2.2.13)$$

となる。

この式(2.2.13)はオプティカルフローの拘束条件であり、勾配法はこの式をもとにしている。

さて、問題は式(2.2.13)にあり、この式からは直接  $u$ 、 $v$  を決定することはできない。この問題の解決法として、Brussらは、式(2.2.10)を得る際に無視したテイラー展開の高次項を誤差項として、最小自乗法により最適な  $u$ 、 $v$  を決定する方法を提案した [Bruss 83]。

さらに、Hornらは式(2.2.13)を、画像の明るさの変化と、カメラの運動  $t$ 、 $\omega$  と直接結び付ける式

$$\frac{\mathbf{s} \cdot \mathbf{t}}{Z} + \mathbf{v} \cdot \boldsymbol{\omega} + \frac{\partial f}{\partial t} = 0 \quad (2.2.14)$$

$$\mathbf{s} = \left[ -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right]^T$$



$$v = \left[ \frac{\partial f}{\partial y} + y \left( x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right), -\frac{\partial f}{\partial x} - x \left( x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right), -x \frac{\partial f}{\partial y} \right]^T$$

に変形し、画像の明るさの変化  $\partial f/\partial x$ 、 $\partial f/\partial y$ 、 $\partial f/\partial t$  から、カメラの運動または環境構造を、直接評価する方法を導いた。なお、 $Z$  は物体までの距離を表す。

### 環境モデル

これまでは、カメラより取り込まれた画像情報と、ロボットの運動パラメータを用いて、環境の3次元構造を知る方法のいくつかについて述べた。しかし、実際の移動ロボットでは、環境に関する知識が与えられる場合もあり、先に述べた方法に加え、環境モデルからの知識を有効に利用すれば、移動ロボットの環境認識の機能はより充実する。

環境モデルの利用としては、環境内での正確なロボットの位置の発見や、環境内に存在する物体の認識が考えられる。屋内環境を移動するロボットに、あらかじめ環境モデルを与えておけば、ロボットは視覚センサによる環境の計測結果を、環境モデルと照らし合わせることで、環境モデル内での自らの位置を発見したり [Yachida 83]、環境内にある物体が何であるかを知ることができる。

環境モデルの利用は、ロボットの状態や移動する環境に応じたものでなければならない。また、環境表現の方法も常に同じものを用いるのもよくない。これらの問題は、移動ロボットの研究の枠を超えた、人工知能全体の研究課題である。

### (B) 移動ロボットによる環境認識

(A) では、物体認識に用いられるいくつかの原理が、移動ロボットの環境認識にも利用されていることについて述べたが、ここでは、移動ロボット特有の研究について述べる。比較的新しい研究の中から、画像の入力方法、自由領域の認識、面情報の復元に関するものを取り上げた。この内、画像

の入力方法は本論文と深く関係するものである。

#### 画像入力の方法

環境全体の構造を知るには、できるだけ視野の広い画像入力装置を用いた方が都合がよい。そのために、いくつかの方法が提案されている。視野の広い画像を入力する主な方法として、円錐鏡を用いる方法、カメラを回転させる方法、魚眼レンズを用いる方法の3つがある。円錐鏡を用いる方法と、本論文で用いるカメラを回転させる方法は、ある地点から水平方向に見渡した全方位画像を構成し、利用するものである。これらについては、4.1で述べる。

魚眼レンズを用いた方法 [Morita 88] は、 $180^\circ$  の視野を持つカメラを使って、球面投影された画像を直接撮像し、環境内の平行線の消失点を Hough 変換により求めるものである。屋内環境を特徴づけるのは、柱や天井と壁の境界などであり、それらはほとんど直線である。これらの直線からなる平行線の消失点は、環境の構造を特徴づける重要な要素である。特に構造の簡単な廊下などでは、消失点を用いた環境認識の方法が幾つか提案されている。魚眼レンズを用いると、ガウス球への投影が直接得られるという特徴より、得られる円形の画像の周辺が無遠点となる。この性質を利用し、Hough 変換することで、その消失点の位置を正確に求めることができる。

#### 自由空間の認識

移動ロボットが移動可能な自由空間を発見する研究は、すでに多くのものが報告されている。しかし、近年においても未だ多くの研究発表がなされているのは、移動ロボットを研究する機関が増えてきているためであろう。自由空間の認識に用いられる手法の中で最も研究例の多いのは、床面への逆透視変換を用いた方法である。一般のカメラで撮像される画像は3次元空間の中心射影によって得られるため、画像から元の3次元空間を復元することは難しい。しかし、もし3次元空間内に一つの平面のみが存在すると

仮定し、その平面に対するカメラの位置が与えられていれば、得られた画像から元の3次元空間内の平面を再現できる。この画像の3次元空間内の平面への変換を逆透視変換という。3次元空間内の平面としては主に床面が用いられる。床面への逆透視変換が用いられる理由は、他の距離計測方法に比べアルゴリズムが簡単で、床面に関する知識が正確に与えられている場合、安定な方法となるためである。次に、その一例を紹介しよう。

Storjohann[Storjohann 90]らは、移動ロボット上に左右に取り付けた2台のカメラにより撮像された画像を、それぞれ床面上へ逆透視変換し、それらの画像の差分をとることにより、床面上に高さのある障害物が存在するかどうかを判断している。すなわち、床面上に高さのある物体が存在しない場合、左右のカメラで撮像された画像の逆透視変換は、正確に一致するはずであり、一方、高さのある物体が存在する場合には、その物体のみが左右の画像の逆透視変換で一致しない領域として検出される。彼らは専用のハードウェアを開発し実時間の処理を実現している。

逆透視変換を用いている例は他にも数多くある[例えば Asada 89]。

### 面情報の復元

面情報の復元は環境認識の大きな課題の一つである。認識に時間をかけることが許されるなら、緻密な距離画像を作成し面情報を得ることができる。しかし、ロボットの環境認識のように、処理時間をできるだけ短縮したい場合には、環境内の特徴点や、線分までの距離情報を基に、何らかの方法でそれらの間に存在する面を確認したい。

Faugerasら[Faugeras 90]はデローニ三角法(Delaunay Triangulation)と呼ぶ方法を用いて、移動ロボット上に取り付けられた2台のカメラによるステレオ視によって得られる観測結果を融合し、ロボットの移動に伴い、徐々に環境全体の正確な表現が完成されていく方法を示した。一般に、ステレオ視によって得られる情報は、特徴点や線分の3次元位置であり面情報が含まれていない。そこで、デローニ三角法では、まず最初に、ステレオ視によっ

て位置づけられた3次元環境内の線分とその端点を利用して空間を三角形で分割し、それぞれの三角形は面であると仮定し、図2.2.3(a)に示す立体を得る。次に、違う観測地点よりもう一度観測し、視線の拘束条件から、存在しえない面を表す三角形を取り除いていく。この処理をカメラの移動に伴いながら繰り返していくと、最後に、同図(b)に示す真に存在する面だけから構成された立体を得ることができる。これは4つの観測地点での観測結果を用いたものである。面の存在を認知する方法として興味深い。

### 2.2.2 屋外環境

屋外環境の移動ロボット研究は、道路などの整地環境下を移動するロボットの研究と、荒野などの未整地環境を移動するロボットの研究に分けられる。未整地環境を移動するロボットの研究は既に1節でも触れているため省略する。

整地環境下を移動するロボットの研究は、DARPAのALV(Autonomous Land Vehicle: 自律走行車)プロジェクトを皮切りに盛んになり、現在では米国を初めとして、日本や西ドイツでも研究されている。整地環境下では、環境に関する知識を有効に利用することができるため、ほとんどの研究は道路環境のモデルを用いている。

#### 道路の認識

道路環境に関する知識を有効に用いている代表的な研究例としてはThropeら[Thrope 88]の研究があげられる。彼らは、道路の色情報と道路の幾何モデルを用いたSCARF(Supervised Classification Applied to Road Following)と呼ぶシステムを開発した。SCARFでは数種類の道路領域や非道路領域を表す色に関する知識を用いて、ベイズの定理により道路領域を検出し、検出された道路領域をHough変換することで、道路の向きと道路境界の推定を行っている(図2.2.4)。なお、屋外では、明るさの変動が著しい。この広い変動幅に対しても色情報を安定して撮るために、カメラを2台使って

等価的にカメラのダイナミック・レンジを拡げている。また、道路は連続しているので、前画面の処理結果から次画面における道路の位置を予測できる。この予測を用いて2画面以降を能率的に処理できる [Waxman 87]。

高速道路のように整備された道路だと、道の両端に白線が描かれている。この白線を検出することにより、走行可能な領域を発見しようとする研究もある。Dickmannsら [Dickmanns 88a, 88b]は、白線を検出する窓をいくつか画像上に出し(1本の白線に対し3個)、それぞれの窓が連続する画像上で白線を追跡していくという方式をとっている。追跡にはカルマンフィルタを用いて次画面における位置を予測している。この方式では、誤差範囲が予測できるので誤った検出を防げる、また、どれかの窓で白線の検出に失敗しても、成功した窓を用いて走行可能な道の位置を知ることができるので信頼性が良いといった利点がある。なお、一つの窓内の処理を一つのプロセッサが担当するというマルチプロセッサ型のハードウェアを開発することにより、高速道路上で時速100Kmという高速な自律走行を可能にしている。現在、白線のない一般道路への拡張が検討されている。

ニューラルネットの利用も新しい流れの一つである。Pomerleauら [Pomerleau 90]は、道路領域と非道路領域の判別にニューラルネットを用いた。これをナビゲーションシステムを組み込むことにより、道路境界が曖昧であるアスファルト道路以外の道路においても、安定に走行できることを実際の移動ロボットによって確かめた。変化に富む屋外環境では、事前に完全なアルゴリズムを設計することは難しいので、種々の環境を見せることにより学習することのできるニューラルネットを利用しようという考え方である。

### ランドマークの認識

移動ロボットが道路を認識できたとすると、次に問題となるのは経路上での自分自身の位置決めである。このためにはランドマークと呼ぶ目印となる特徴をあらかじめ教えておき、それらとの位置関係から位置決めをす

るという方法がよく用いられる。

ランドマークとしては、道路標識や道路表示等が考えられる。これらは自動車の運転の際に用いられるものであるため、道路環境内の他のものよりもはっきりとした視覚的特徴を持っている。しかし、屋外環境では照明の変化が激しかったり、背景が非常に複雑であるため、これらのランドマークを安定に検出することは難しい。

Blancard ら [Blancard 90] はあらかじめ与えられている道路標識モデルと、入力画像の標識と思われる領域とをニューラルネットを用いて照合を行って認識している。

小森谷ら [Komoriya 84] は、道路環境沿いにある構造物をランドマークとする方法を提案した。この方法では、電柱やビルの壁面などの構造的特徴を複数の超音波センサを用いて検出するため、カメラを用いた場合に問題となる照明の変化等の影響を受けない。彼らはこの方法を街中を移動する盲導犬ロボットのナビゲーションに利用している。

人工物をランドマークとするこれらの研究に対し、自然の地形をランドマークとして利用する研究がある。テキサス大 [Talluri 90] では、環境のモデルが高さ地図の形で与えられたと仮定して、2次元の画像からロボットの位置決めを試みている。地平線の見え方を鍵となる特徴として用いている。すなわち、画像から地平線を抽出した後、ロボットの位置を変えたときの地平線の見え方モデルを高さ地図から作成し、それが画像から得られたものと一致するロボットの位置をヒューリスティックに探す。また、Levitt ら [Levitt 87] はある地点において全方位を見渡したときのランドマークの見える順序となす角度から、ランドマークの位置を示した地図をもとに、ロボットの位置をある程度推定する定性的ナビゲーション (Qualitative Navigation) と呼ぶ方法を提案している。

一方、Zheng ら [Zheng 90B] は、環境内の個々の物体をランドマークとするのではなく、移動中に撮影された連続的な画像より、色情報に代表されるいくつかの特徴量を取り出し、それらをランドマークとして用いる方法

を提案した。彼らの方法は、ロボットの移動方向と垂直な方向に取り付けられたカメラから、ロボットの移動中に撮像される画像の画像中心を通る幅1画素の画像データを横一列に整列させ(図2.2.5)、得られた画像をそのままテンプレートとして記録しておく。ロボットの移動経路に沿ってこのような視覚的な履歴を作成してしまえば、もう一度同じ経路を通る際に、同様の方法で画像を撮像し、動的計画法を用いて対応をとることで、その経路におけるロボットの位置を知ることができる。この方法は、個々の物体をランドマークとしないため、種々の環境に用いることができる。

### 2.2.3 画像認識に共通の問題

最近のロボットの視覚に関する研究全体に共通する問題としては、能動的視覚、環境表現、計測誤差の収束方法(カルマンフィルタ)等がある。能動的視覚は本論文と深く関係するため、3.1でも取り上げる。以下、これらの問題について、現時点における研究成果と今後の課題について述べる。

#### 能動的視覚(Active Vision)

最近、物体認識の研究において、Marrの提唱した画像認識におけるパラダイムにかわる能動的視覚(Active Vision)と呼ばれるパラダイムが提案され、注目を集めている。能動的視覚は物体認識に注視行動を取り入れたパラダイムであり、Marrの提唱したパラダイムとの違いは、注視行動により、直接、物体中心の座標系で物体の記述を得ることである。主に Aloimonos と Ballard の2人の研究者によってその枠組みが築き上げられた。

Aloimonos ら [Aloimonos 87, 89] は Structure from Motion のみならず、一般に Shape from X と総称される問題について、特徴点を画像中心に固定しながら、その特徴点を中心とした円弧運動を行なうカメラ動作(注視制御)の有効性について理論的な考察を行った。注視制御を用いることで、これらの問題において不良設定(ill-posed)なものが良設定(well-posed)になったり、非線形の問題が線形となる等の利点がある。Aloimonos の研究の内、Structure

from Motion を次に紹介する。

従来の Structure from Motion [Bruss 83] では、図 2.2.2 に示すように、カメラ運動の並進速度成分を  $(U, V, W)$ 、回転速度成分を  $(A, B, C)$  とすると物体のオプティカルフローを観測することにより決定できるカメラの運動パラメータは、 $U/W, V/W, A, B, C$  の 5 つである。一方、カメラが注視制御される場合、オプティカルフローを観測して決定できるカメラの運動パラメータは、注視点までの距離を  $Z_0$  として、 $U/Z_0, V/Z_0, W/Z_0, C$  の 4 つになる。すなわち、カメラを能動的に運動させることによって(注視点を中心とした円弧運動)、カメラの光軸に直交するカメラ運動の回転成分を消去できる。また、対象物体までの距離  $Z_0$  を何らかの方法により測定すれば、カメラの並進速度成分を求めることができる。

一方、Ballard は実時間で物体を追跡するシステムを用いた Animate Vision と名付ける研究の枠組みを提案した [Ballard 89]。輻輳運動が可能な 2 台のカメラをマニピュレータの先端に取り付け、人間の眼球運動に似た機能を実現している。環境構造を復元する計算式は、中心窩における近似 (Retina Approximation) を用いてより単純化されているが、基本的に Aloimonos らのものと差はない。

これら以外の能動的視覚の研究としては、拘束されたカメラ運動のもとに注視制御の利点を利用する研究 [Sandini 90]、人間の網膜の構造に似た撮像素子を利用する研究 [Tistarelli 90]、考えられる限りのヒューリスティックを直接適用することで注視点の選択基準を与える研究 [Kuniyoshi 90] 等がある。

Active Vision のパラダイムに基づく研究の課題は、注視点の選択方法であるが、これについての理論的な研究はほとんどない。

### 環境表現

視覚センサから得られる情報より環境を認識するには、環境の表現方法が重要な問題となる。環境は、距離センサの情報から作成される高さ地



図(図 2.1.1)や、ステレオ視から得られる距離画像により直接的に表現することができる。しかし、より柔軟な環境認識を行なうためには、さらに工夫された形式で環境を表現する必要がある。そのような環境の表現方法の一つとして、屋内環境を床面上の2次元平面で表し、その平面を正方形や長方形のグリッドで区切ることにより、環境を表すグリッド表現がある。ここでは、グリッド表現を用いた幾つかの研究例を取り上げる。

Kambhampatiら [Kambahampati 86]はグリッド表現を階層的に用いた4分木法(quad-tree)により、全環境を少ないデータ量で表わすことを試みた。4分木法では、まず最初に全環境を一つの正方形グリッドで表す。次に、グリッド中に物体が存在する場合はグリッドを4つの正方形グリッドに分割し、グリッド内に物体が存在しない場合はグリッドは分割しない。この処理を繰り返すことにより、物体の存在しない広い自由領域は大きなグリッドで、複雑な物体は小さなグリッドの集合として表わされる。

さらに、Matthiesら [Matthies 88]は、グリッド表現を複数センサ情報の融合に用いた。彼らは、環境を格子状に区切り、各々のグリッドに障害物が存在する確率変数を割り付けた(占有グリッド(Occupied Grid))。これらの確率変数は、ステレオ視による物体の距離情報と超音波センサによる障害物の有無の確率的な判定結果を用いて、ベイズの定理により更新される。同一のグリッドに対する観測を繰り返すほど、そのグリッドにおける障害物の有無は明らかになる。図 2.2.6(a)は超音波センサによる観測を、同図(b)はステレオ視による観測をそれぞれ10回行った結果を示し、同図(c)はそれらを融合した結果を示す。この環境表現方法の利点は、新たに別のセンサを用いた場合、それから得られる情報を容易に融合できることである。

一方、Elfes[Elfes 87]はグリッド表現をもとにした環境表現の様々な形態について考えている。図 2.2.7は、超音波センサによる観測から、抽象度、分解能、地域性の3つの尺度において、それぞれ3つのレベルで異なる環境表現が得られることを示している。これらの環境表現はロボットによる観測にとまって更新され、ロボットの状態や目的に応じて最も適当な環境表

現が利用される。

Elfesの行なった抽象度における様々なレベルでの環境表現は、認知地図(Cognitive Map)の研究につながる。認知地図に関する研究例は少ないが、その中で、Yeap[Yeap 88]は環境内を移動するロボットのセンサより2.5D表現が得られることを仮定し、認知地図を作成するシミュレーションを行なっている。今後、実環境下での認知地図作成の研究が期待される。

以上に述べた環境表現は、まず最初にセンサ中心の座標系で表現され、その後、座標変換により物体中心や環境中心の座標系に変換することで得られる。しかし、最近の研究では先に述べた能動的視覚のパラダイムが提案され、直接、物体中心の座標系で物体の記述を得る方法と、その座標系での計算の有効性が明らかになってきた。このように、各種の座標系における表現はそれぞれ利点を持ち、環境を表現する座標系について考えることも重要な課題である。本論文の4章ではこの問題についても取り上げる。

#### 計測誤差の収束方法

従来、ステレオなどの計測誤差を収束させる方法として、最小自乗法が最も多く用いられてきたが、最近の研究では、最小自乗法にかわり、信号処理の分野で有名なカルマンフィルタが用いられるようになってきた。カルマンフィルタの特徴は、最小自乗法に比べ収束速度が早いことである。

ロボットの視覚において最初にカルマンフィルタが用いられたのは、ステレオ法による計測において、カメラの運動パラメータと、特徴点の位置を正確に決定する研究においてである[Ayache 87]。後に、多くの研究者によってカルマンフィルタの有効性が実験的に確かめられた。特に、Matthiesら[Matthies 89]は誤差の収束に対する理論的な考察を行い、最小自乗法ではデータ数の2乗に比例し誤差が収束するが、カルマンフィルタでは3乗に比例することを明らかにした。

カルマンフィルタはハードウェア化が容易であり、今後も多くの実験に用いられると思われる。

本章では、本論文を位置付けるこれまでのロボットの視覚研究について述べた。

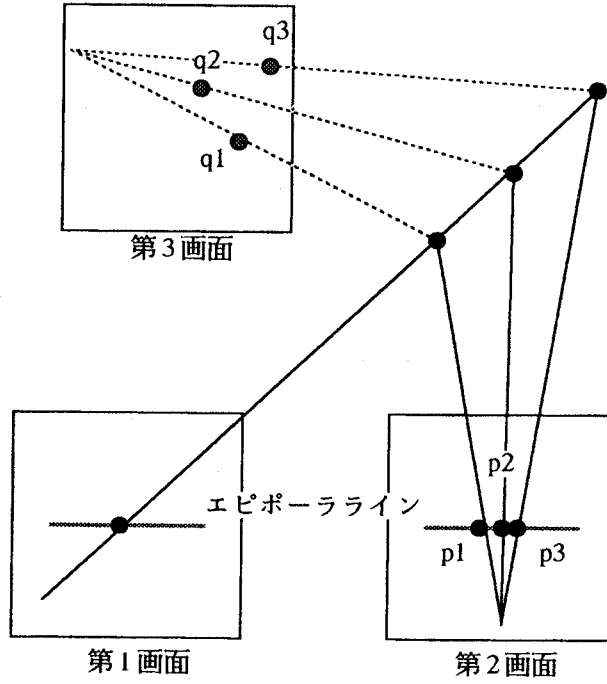


図2.2.1 3眼視ステレオ法の原理

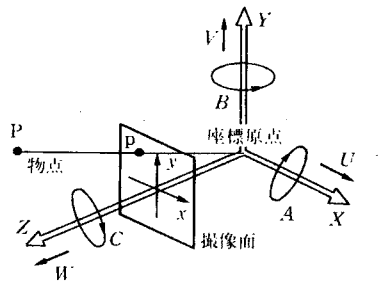
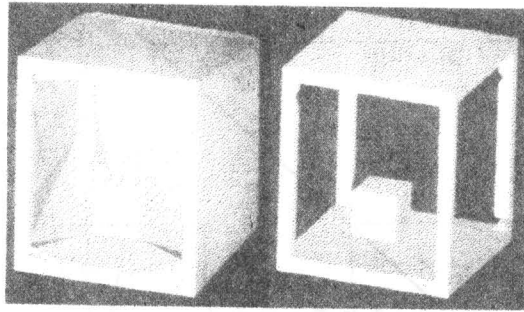


図2.2.2 カメラ中心の座標系



(a)1回の観測結果

(b)4回の観測結果の融合

図 2.2.3 デローニ三角法による面情報の復元 [Faugeras 90]

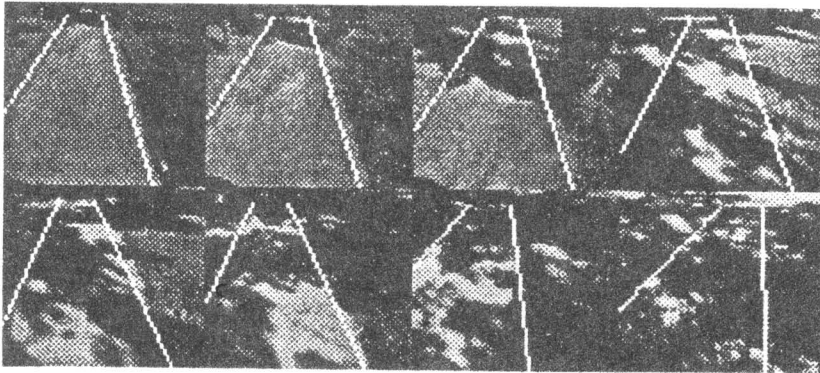


図 2.2.4 影のある道路の認識 [Thrope 88]

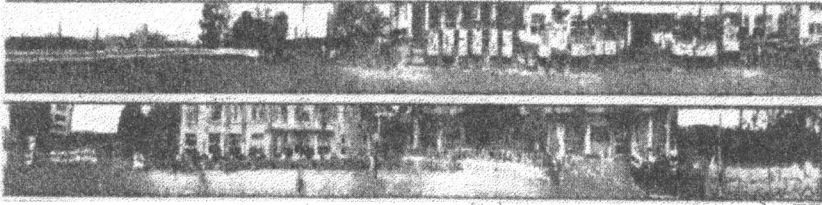
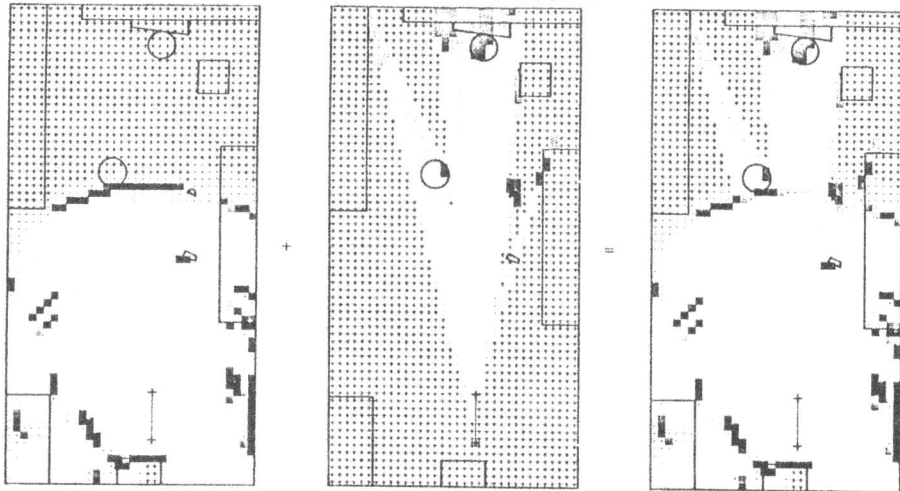


図 2.2.5 ロボットの移動中に撮像された視覚的履歴 [Zheng 90]



(a)超音波センサによる観測 (b)2眼ステレオによる観測 (c)融合結果

図 2.2.6 グリッド表現を用いた環境認識 [Matthies 88]

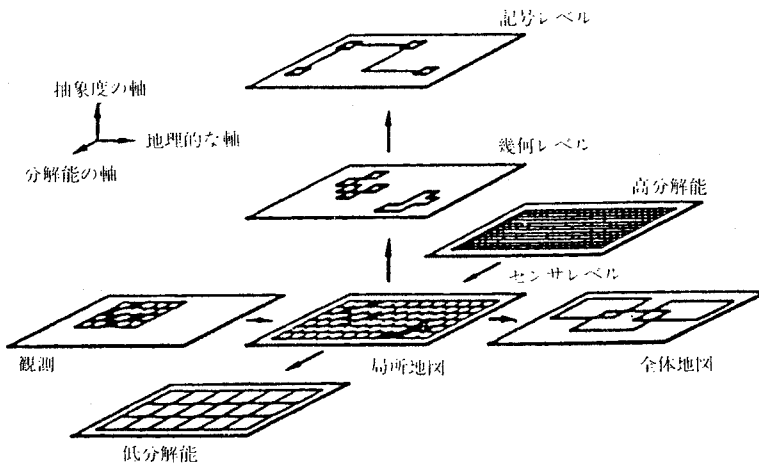


図 2.2.7 環境表現の様々な形態 [Elfes 87]

### 第3章 環境の一部を注視するカメラ運動

従来の視覚移動ロボットでは、視覚センサ(TVカメラ)は移動ロボット上に固定されていた。しかし、より柔軟に環境を認識するには、視覚センサを能動的に動かす必要がある。本章では、注視制御を用いた移動ロボットの環境認識における戦略を提案し、その実現における問題の一つである局所地図獲得について考察する。まず最初に、連続的な観測を前提とした局所地図を得るための方程式を導く。次にその方程式の誤差解析を行うことにより、注視制御の利点を明らかにする。注視制御により局所地図を獲得する利点は、直接物体中心の座標系で環境記述を得られるだけでなく、注視点付近の物体の位置決め誤差が、注視制御を用いない方法に比べ、小さくなることにある。実験では、連続的な観測結果をカルマンフィルタを用いて融合し、正確な局所地図を得る方法を提案する。



### 3.1 注視制御に関する研究

近年、画像理解や移動ロボットの視覚に、能動的カメラ移動の効果を用いる研究が注目されている [Aloimonos 87, Ballard 88, Sandini 90]。これらの研究は、Marrの提唱した画像認識におけるパラダイム [Marr 82] にかわる、注視行動を取り入れた能動的視覚 (Active Vision) と呼ばれるパラダイムに基づくものである。

Aloimonos [Aloimonos 87] らは Structure from Motion (Passive Navigation)、Shape from Shading、Shape from Texture、Shape from Contour の問題について能動的視覚のパラダイムに基づく理論的な考察を行った。カメラの円弧運動の中心に特徴点 (注視点) を固定することにより、先の Shape from X の問題がより良設定 (well-posed) となることを示した。特に Structure from Motion の問題では、注視点を中心とする物体中心の環境記述を直接得ることができ、さらに、注視点までの距離を何らかの方法で計測することにより、カメラの運動パラメータの全てを求めることができる。

一方、Ballard [Ballard 89] は、能動的視覚と人間の視覚行動との共通点を重要視し、物体の一部への注視を繰り返すことで、物体や環境を認識する人間の視覚行動をもとにした Animate Vision と名付ける研究枠組を提案した。Animate Vision では、2眼ステレオ視を用いた実時間に物体を追跡するシステムを用いている。彼は、このシステムを用いて、注視制御により3次元構造が復元できることを示すだけでなく、注視点の選択についても考察している。

Aloimonos や Ballard の研究では、カメラは空間を自由に移動できることを前提としている。例えば、Ballard が実験に用いたシステムでは、カメラはマニピュレータの先に固定されており、マニピュレータの運動できる範囲ではカメラは自由な姿勢を取ることができた。このため、カメラの運動は注視点を中心とする円弧運動をする (2.2.3 で説明したように、Aloimonos の研究では、カメラを円弧運動させる必要がある)。これに対して、カメラが

直線の移動経路をたどりながら、回転運動により環境の一部に注視する場合について考察したのがSandiniらである。

Sandiniら [Sandini 90]は、カメラが回転運動により注視点に注視しながら、直線の移動経路を移動する間に、離散的に撮像された画像より、注視点を中心とする環境構造が復元できることを示した。

本章で述べる研究も、Sandiniらと同様のカメラ運動を前提としているが、彼らとの違いは、

- (1) Sandiniらが離散的観測を前提としたのに対し、連続的観測を前提とした環境構造の復元方法を導き、カルマンフィルタより観測結果を融合する方法を提案する。
- (2) 注視制御における誤差解析を行う。
- (3) 注視制御を用いた移動ロボットの環境認識の戦略を提案する。

等があげられる。

## 3.2 注視制御を用いた環境認識

従来の移動ロボットを用いて環境の全体地図を作成する研究例としては、環境内の複数の地点で、ロボット本体に固定された2台のカメラにより撮像された画像よりステレオ視を行い、得られた複数の局所地図をロボットの車軸に取り付けられたエンコーダの情報をもとにつなぎ合わせる研究がある[Ayache 87]。この研究のように、ロボットの移動経路を正確に再現することで環境の正確な全体地図を得る研究に対して、ロボットに必要な情報は環境内で局所的に存在し(局所地図)、それら局所地図間の位置関係は正確に知る必要がないと考える研究[Asada 88]もある。

筆者も後者と同様の考えを持つ。例えば、屋内環境であれば、ロボットに必要な情報は廊下の曲がり角などの環境構造を特徴づける部分や、ドアなどのロボットのランドマークとなるものであり、ロボットが移動可能な自由空間を見つけ、自律的に移動する機能を持つ場合には、正確な全体地図を作成する必要な少ないと考える。本節では、このような移動ロボットの環境認識に注視制御を取り入れた次の戦略を提案する。

図3.2.1は我々の戦略に基づく視覚移動ロボットの観測行動を示す。移動ロボットは1自由度の回転可能な(床面と垂直な軸を中心として)一台のカメラを搭載しており、この直線と円弧(進路変更のための運動)を組み合わせた経路を移動しながら、直線運動を行う際に環境内より興味ある一部分を選び、それを注視する。注視はその興味ある一部分の中心に位置する一点(注視点)を、視覚フィードバックの機能によって画像中心に固定することで実現される。このような注視行動を行った後、注視点を中心とする物体中心の座標系で表わされる局所地図が得られる。具体的には、最初A点(図3.2.1)を選択し、A点を注視しながら移動して行くが、A点周辺の構造を十分復元した後は、新たな注視点Bを選択し、再びB点の回りの構造を探索する。いくつかの局所地図を得た後、環境の全体的な記述を得るために、各注視点間の大まかな関係を利用して、得られた局所地図を全体地図に統合

する。

各注視点間の大まかな関係を得るためには、注視制御を行うカメラシステム以外に、ロボット上に固定された視野の広いカメラが必要である。その視野の広いカメラでモーションステレオを行うことで、各注視点間の大まかな関係を得ることができる。[Stelmaszyk 88]。このように注視制御を用いて環境認識を行うロボットには、注視制御を行うカメラシステムと、ロボット上に固定された視野の広いカメラシステムが必要である(図3.2.2)。

本章では、上記の戦略の実現における問題の一つである局所地図の獲得を中心に述べる。

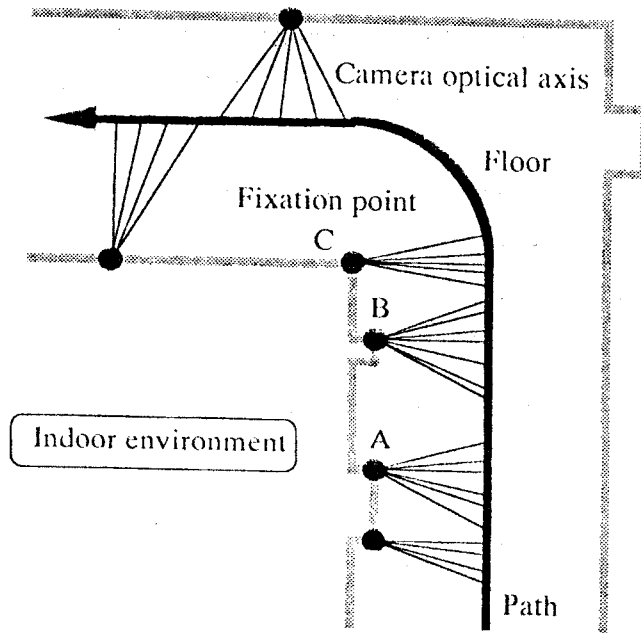


図 3.2.1 注視制御を用いた移動ロボット的环境認識

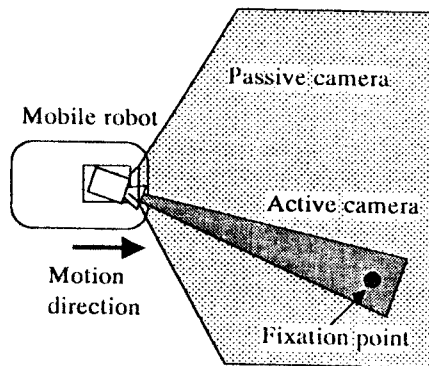


図 3.2.2 2つのカメラシステム

### 3.3 ロボットと環境に関する仮定

本章では以下の仮定のもとに局所地図を得る方法を導く。

- (1) ロボットの運動は正確に制御でき、直線運動を組み合わせた屋内環境の床面上の経路を移動することができ、特に直線運動を行う場合には、そのカメラの光軸に対するロボットの運動方向が測定可能である。
- (2) ロボットには正確に位置制御できるカメラ回転台が搭載されており、カメラはその回転台によって床面と平行な面上で回転運動を行うことができる。注視制御に用いるカメラは1台であり、そのカメラの焦点は回転台の回転軸上に位置する。
- (3) ロボットの移動中、注視点は視覚フィードバックの機能により常に画像中心に固定される。
- (4) 環境を構成する特徴点や注視点として、環境内の垂直線を用いる。屋内環境内には多くの垂直線が存在し、それらの垂直線はロボットの移動にとって必要十分な情報を含んでいると考える。

以上の仮定のもとに、局所地図は床面に平行な平面上での垂直線の位置を求めることで表わされる。この際、局所地図は画面上の垂直線の水平座標と画面速度の水平成分を用いて計算される。

### 3.4 連続的観測における局所地図の獲得

#### 3.4.1 局所地図の計算方法

本節では、ロボットが直線運動を行う際に、注視制御により注視点中心の座標系で物体の位置決めを行う方程式を導く(一般的なロボットの運動における注視制御については付録Aを参照)。同様の方法はすでに Sandiniら [Sandini 90] により報告されているが、彼らが不連続な観測をもとにした方程式を導いたのに対し、本節では、連続的な観測をもとにした方程式を導く。その理由は、連続的な観測結果をカルマンフィルタにより融合し局所地図を得る方法を実現するためである。また、本章では2次元の局所地図を求める方程式を導くが、画面上の垂直線の両端の垂直座標を用いることにより、3次元空間での位置決めを行う方程式に拡張することは容易である。

図3.4.1に示すように、カメラ中心の座標系O-XYZを考え、そのZ軸をカメラの光軸とする。カメラの並進速度成分を  $T=(U,V,W)^T$ 、回転速度成分を  $R=(A,B,C)^T$ 、3次元環境内の点を  $P=(X,Y,Z)$ 、点Pの3次元環境内でのX,Y,Z軸方向の並進速度成分を  $X',Y',Z'$  とすると、画面上の点pのx座標とその画面速度  $u$  とカメラ運動の関係は、式(3.4.1)で表わすことができる [Bruss 83]。

$$u = \frac{fX'}{Z} - \frac{fXZ'}{Z^2} = f \left( -\frac{U}{Z} - B + \frac{Cy}{f} \right) - x \left( -\frac{W}{Z} - \frac{Ay}{f} + \frac{Bx}{f} \right) \quad (3.4.1)$$

ここで、 $f$ はカメラの焦点距離である。

ところで、先に述べた仮定より、ロボットが直線運動する間、注視点を画像中心に固定するようにカメラの回転が制御されると仮定すると、カメラの運動パラメータは以下のように表すことができる。

$$\begin{cases} U = U_R \cos \varphi \\ V = 0 \\ W = U_R \sin \varphi \end{cases} \begin{cases} A = 0 \\ B = \omega \\ C = 0 \end{cases} \quad (3.4.2)$$

ここで、 $\omega$ はカメラ回転台より直接計測でき、ロボットの運動パラメータとは独立な値である。 $\varphi$ はカメラの光軸とロボットの移動経路間の角度であり、カメラ回転台より直接計測できる。 $U_R$ は直線運動するロボットの速度である。

式(3.4.1)と式(3.4.2)より、物体の画面速度の水平成分  $u$  は式(3.3)で表される。

$$u = \frac{U_R(x \sin \varphi - f \cos \varphi)}{Z} - \omega \left( f + \frac{x^2}{f} \right) \quad (3.4.3)$$

一方、注視点は注視制御により常に画像中心に固定されるため、その画面速度と画面上での  $x$  座標はどちらも 0 となり、これより、カメラの焦点から注視点までの距離  $ZF$  を式(3.4.4)で計算することができる。

$$ZF = -\frac{U_R \cos \varphi}{\omega} \quad (3.4.4)$$

また、カメラの焦点から物体までの距離  $ZS$  も、式(3.4.3)より、式(3.4.5)として与えられる。

$$ZS = \frac{U_R(x \sin \varphi - f \cos \varphi)}{u + f\omega + \frac{\omega x^2}{f}} \quad (3.4.5)$$

次に、式(3.4.4)と式(3.4.5)より、ロボットの速度  $U_R$  を消去するために、 $ZS/ZF$  を求める。

$$\frac{ZS}{ZF} = \frac{f - x \tan \varphi}{\frac{u}{\omega} + f + \frac{x^2}{f}} \quad (3.4.6)$$

式(3.4.6)と、図 3.4.2 に示される幾何的關係より、注視点に対する環境内の一点の位置は、極座標を  $(\theta, L)$  を用いて式(3.4.7)で表すことができる。



$$\theta = \tan^{-1}\left(\frac{xK}{f(1-K)}\right), L = \frac{D}{\cos\phi} \sqrt{(1-K)^2 + \frac{K^2 x^2}{f^2}} \quad (3.4.7)$$

ここで、 $K=ZS/ZF$ である。また、 $D$ は注視点とロボットの移動経路間の距離であり、得られる局所地図のスケールファクタとなる。我々の提案する環境認識の戦略においてはスケールファクタの厳密な決定は必要ではないが、その大まかな評価は必要である。この測定は、3.2節の環境認識のための戦略で述べた、ロボット上に固定されたもう一つのカメラによりモーションステレオによって行う。

式(3.4.7)に示されるように、ロボットが直線運動を行う場合には、ロボットの運動パラメータ  $U_r$  は用いず、かわりに、ロボット上に取り付けられたカメラ回転台の運動パラメータ  $\omega$  を用いる(カメラ回転台の回転角速度  $\omega$  とロボットの速度  $U_r$  の関係は付録Bで述べる。また、カメラ回転台の回転角速度  $\omega$  の性質は付録Cで述べる)。カメラ回転台は環境から影響を受けずに制御できるため、車輪のスリップなどにより誤差の積算し安いロボットの運動パラメータに比べて、その回転角度は高い精度で測定できる。

また、注視制御における物体の画面速度が0となる点(Focus Of Expansion)は、注視点を通りロボットの移動経路に接する円となる。これについては付録Dで述べる。

### 3.4.2 誤差解析

注視制御を用いて局所地図を獲得する場合、従来の方法と比べてどのような利点があるのだろうか。この問題を明らかにするために、前章で導いた局所地図を計算する方程式の誤差解析を行った(付録Eでは、Aloimonosのカメラが注視点を中心とした円弧運動を行う場合の誤差解析について述べる)。

式(3.4.8)を用いて  $\Delta\theta, \Delta L$  を評価したのち、Jacobi行列式を用いて、原点が注視点上にあり  $X$  軸がロボットの移動経路に平行な直交座標系  $XY$ (図3.4.3)

における位置決め誤差 $\Delta X, \Delta Y$ に変換し、 $\sqrt{\Delta X^2 + \Delta Y^2}$ によって誤差を表わす。

$$\Delta\theta = \left| \frac{\delta\theta}{\delta\omega} \right| \Delta\omega + \left| \frac{\delta\theta}{\delta u} \right| \Delta u + \left| \frac{\delta\theta}{\delta x} \right| \Delta x + \left| \frac{\delta\theta}{\delta\phi} \right| \Delta\phi + \left| \frac{\delta\theta}{\delta f} \right| \Delta f \quad (3.4.8)$$

$$\Delta L = \left| \frac{\delta L}{\delta\omega} \right| \Delta\omega + \left| \frac{\delta L}{\delta u} \right| \Delta u + \left| \frac{\delta L}{\delta x} \right| \Delta x + \left| \frac{\delta L}{\delta\phi} \right| \Delta\phi + \left| \frac{\delta L}{\delta f} \right| \Delta f$$

#### 注視制御の位置決め誤差

以下に示す各パラメータで表わされる状態において、局所地図獲得の際の誤差評価を行う(図3.4.3)。

$$\phi = 45.0 \text{ degrees}$$

$$D = 5.0 \text{ m}$$

$$f = 600.0 \text{ pixels}$$

$$U_r = 0.5 \text{ m/time-unit}$$

また、その時の各パラメータの計測誤差を以下のように仮定する。

$$\Delta\omega = 0.1 \text{ degrees}$$

$$\Delta u = 1.0 \text{ pixel}$$

$$\Delta x = 1.0 \text{ pixel}$$

$$\Delta\phi = 0.1 \text{ degrees}$$

$$\Delta f = 10.0 \text{ pixels}$$

この時、図3.4.3中の注視点を中心とする正方形の領域内での位置決め誤差 $\sqrt{\Delta X^2 + \Delta Y^2}$ は図3.4.4に示される。図3.4.4は各点での位置決め誤差の大きさを、図中下方の正方形からの高さとして等高線を用いて表わしたものである。図中中央の矢印は注視点の位置を表わしている。

図3.4.4に示すように、誤差はロボットから離れるに従い大きくなるが、注視点付近での誤差の増加する割合は小さく、注視点から離れるに従い、増加の割合が増す。

位置決め誤差の要因となる各パラメータの計測誤差について、その性質を個々に調べてみると、 $\Delta\omega$ のみを考慮した場合の位置決め誤差(その他のパラメータの計測誤差を0と仮定する)と $\Delta u$ のみを考慮した場合の位置決

め誤差は、それぞれ注視点に最も近い位置とロボットに最も近い位置で最小となり、この2つのパラメータの誤差により、図3.4.4の形がほぼ決定されていることがわかった。

#### 注視制御を行わない場合との比較

注視制御の利点を明らかにするために、ロボットの直線運動において注視制御を行わずに局所地図を得る方法 (Passive Navigation) と比較する。

注視制御を行わない場合、カメラをロボットの進行方向に対して角度  $\pi/2 - \varphi$  で固定し、注視点が画像中心を通過する瞬間の情報より注視点を中心とする極座標で局所地図を計算する (図3.4.3参照)。カメラの運動パラメータは、ロボットの直線運動において注視制御を行う場合は式(3.4.2)で表わされたが、注視制御を行わない場合は式(3.4.9)で表わされる。

$$\begin{cases} U = U_R \cos \varphi \\ V = 0 \\ W = U_R \sin \varphi \end{cases} \begin{cases} A = 0 \\ B = 0 \\ C = 0 \end{cases} \quad (3.4.9)$$

注視点に対する物体の位置は、物体の画面上での位置  $x$ 、その画面速度  $u$ 、注視点の画面速度  $u_f$  を用いて式(3.4.6)を式(3.4.10)におきかえ、式(3.4.7)で計算できる。

$$\frac{ZS}{ZF} = \frac{u_f(f - x \tan \varphi)}{uf} = K \quad (3.4.10)$$

注視制御を行わない場合について、先に述べた注視制御の誤差解析と同様の条件における誤差解析を図3.4.5に示す。ただし、この場合はカメラの回転角速度のかわりに注視点の画面速度  $u_f$  を用いるため、その計測誤差  $\Delta u_f = 1.0 \text{ pixel}$  を仮定している。

図3.4.4と図3.4.5を比較すると、注視点の近傍では、注視制御を用いる場合 (図3.4.4) が、注視制御を用いない場合 (図3.4.5) よりも位置決め精度が良いことがわかる (注視制御の誤差解析においてカメラ回転台の計測誤差  $\Delta \omega$

=0.1degreeを用いたが、焦点距離  $f=600.0\text{pixels}$  のカメラの  $1.0\text{pixel}$  は画角にして約  $0.1\text{degree}$  であるため、この比較は妥当である)。

注視点を正確に画像中心に固定できる機能があり、カメラの回転角速度が十分精度よく計測できるとき(カメラ回転台の制御は環境から独立しているため、その精度を向上させるのは容易である)、注視制御は、注視制御を行わない場合に比べ、注視点の付近で信頼性の高い位置決めができる。このことから、物体中心の表現を行う場合、注視制御は精度的に有効な方法であると言える。

### 3.4.3 実験結果

#### 実験環境

実験には、村田精機製の屋内搬送車を移動ロボットとして用いた。この屋内搬送車は、駆動輪かつステアリングである前輪1輪と、自由輪である後輪2輪を持つ3輪車である。移動の機構として、床面に設置された白線を追跡する機能を持つ。移動ロボットに搭載されるカメラ回転台は、オプティカルエンコーダの信号によりフィードバック制御され、減速ギアの機構と合わせて  $0.005$  度毎に制御できる。用いたカメラは焦点距離が  $16\text{mm}$  のレンズを備えた松下通信工業製のカラー CCD カメラである。画像処理には Nexus6800 を用い、メインコンピュータには Sun4 ワークステーションを用いた。

画像の大きさは  $512 \times 512$  画素、焦点距離はキャリブレーションの結果  $591$  画素であった。

視覚フィードバックを実現するための注視点追跡は、SSDA法を用いて行った。画像中心からのずれが検出されると、直ちに回転台に角度を補正する命令が送られ、注視点は常に画像中心に固定される。

注視点回りの特徴点の画像間での対応は、Stelmazykらによって開発された Token Tracker [Stelmazyk 88] を用いて行った。Token Trackerは連続した画像間で線分の対応を行う高速なアルゴリズムである。原画像よりエッジ検

出を行い多角形近似した後、抽出された線分は4つのパラメータで表現される。それらパラメータはカルマンフィルタを用いて予測され、その予測は次の画面内の線分との対応に用いられる。

#### カルマンフィルタによる局所地図の獲得

一般に物体の画面速度情報を用いた計測は誤差を伴いやすく、何らかの方法で複数の観測結果を融合し、位置決め精度を向上する必要がある。その方法としてカルマンフィルタが多く用いられるようになってきたが、我々の実験においても、カルマンフィルタを用いて、ロボットの複数の位置から計算した局所地図を融合する。

カルマンフィルタは3.4.1で導いた連続的な観測をもとにした方程式を用いて、計測する各パラメータ( $\omega, u, x, \phi, f$ )に誤差が含まれることを前提として設計した。(カルマンフィルタの設計については付録Fを参照)。

測定結果を融合する場合、予測される誤差の範囲から、融合できる場合と融合できない場合がある。この判定はMahalanobis距離[Ayache 87]を用いて行った。

ロボットは視覚フィードバックの機能により注視点を画像中心に固定しながら、直線上の経路を0.05m毎に移動する。図3.4.6はロボットの移動中に撮影された画像である。図の中心に矢印で示される垂直線が注視点である。

ロボットは0.25m移動する毎に局所地図を評価する。図3.4.7は、ロボットが2.0m移動する間、局所地図を12回評価し、カルマンフィルタによって融合した結果を示す。実際の点の位置と、計測された点の位置を比較するために、ロボットの経路と注視点間の距離1.04mを実測し、スケールファクタとして用いた。

図3.4.8は、図3.4.7中の一点について、カルマンフィルタにより誤差が収束していく様子を示す(図3.4.8では点の実際の位置からの距離を誤差としている)。

図3.4.8より、実際の実験においても、注視点付近で位置決め誤差が小さく、注視点から離れれば離れるほど誤差が大きくなることがわかる。実験の結果、4回以上カルマンフィルタで融合された点に関しては、注視点を中心とする半径1.0mの円内において、誤差は最大で0.018mであった。

#### 3.4.4 考察

本節で述べた、注視制御により局所地図を獲得する方法の利点は、ロボットの運動パラメータを用いずに、ロボットの運動とは独立に制御可能な、カメラ回転台のパラメータを用いて局所地図を計算する点にある。すなわち、カメラを回転させ、正確に注視点画像中心に固定することができれば、局所地図は、注視点付近だけでなく、より広範囲で正確なものとなる。

しかし、問題は正確な局所地図を得るためには、ロボットが正確な直進運動をしなければならないことである。ロボットは必ずしも正確な直進運動をするとは限らないので、何らかの方法によりこの問題を解決しなければならない。この問題について、本節では、カルマンフィルタを用いて、連続的に得られる観測結果を融合することで対処した。

ところで、この問題を解決するもう一つの方法は、ロボットが長い距離を移動する間の、離散的な観測結果を用いることで、相対的にカメラの回転角度の計測誤差を小さくするという方法である。この問題について、次節の連続的観測における局所地図の獲得で述べる。

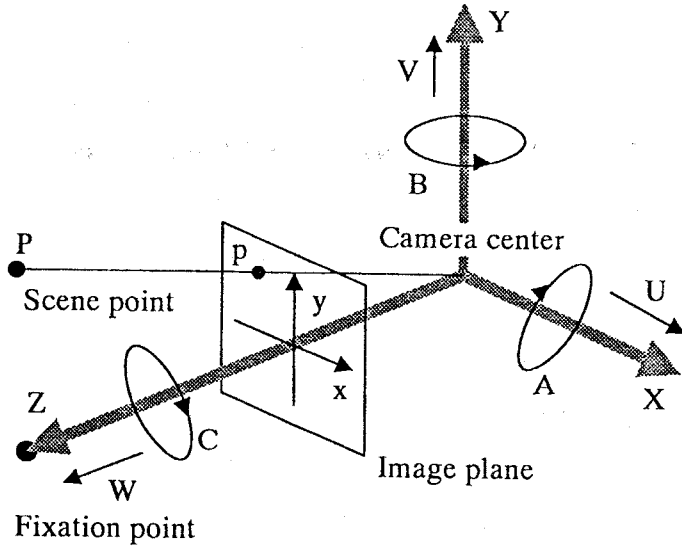


図3.4.1 座標系

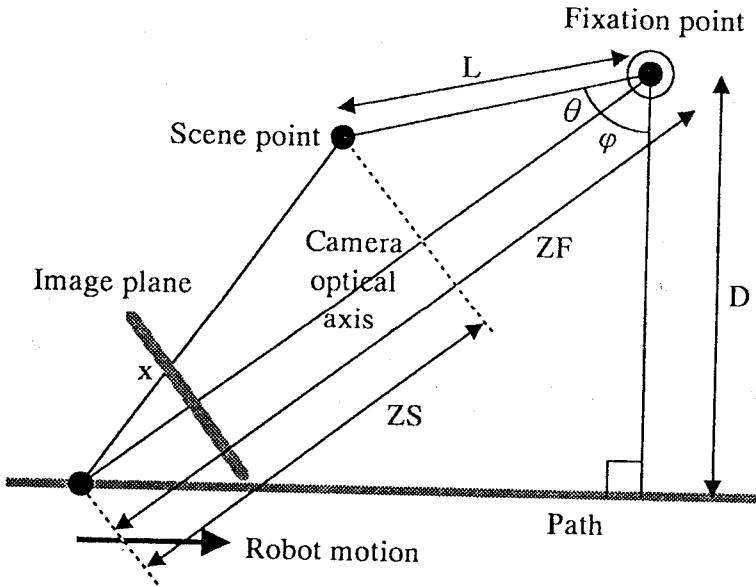


図3.4.2 ロボットの直進運動における注視制御

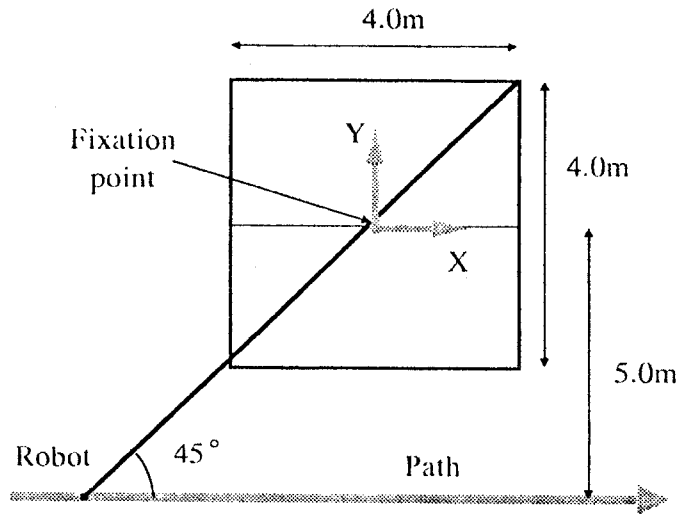


図3.4.3 誤差解析におけるロボットの状態

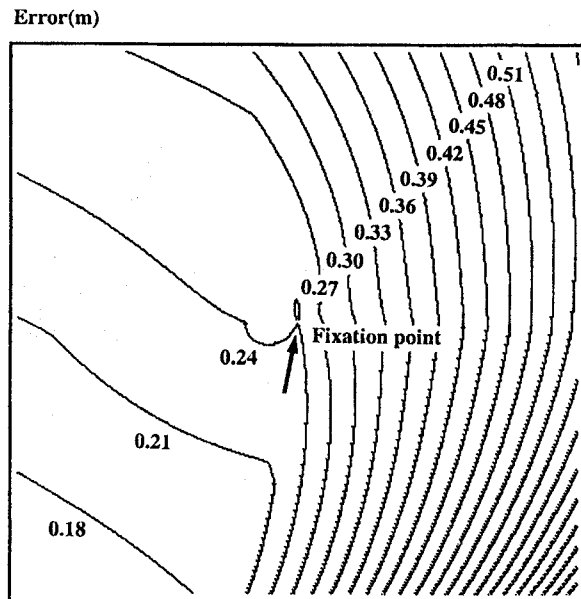


図3.4.4 注視制御における位置決め誤差



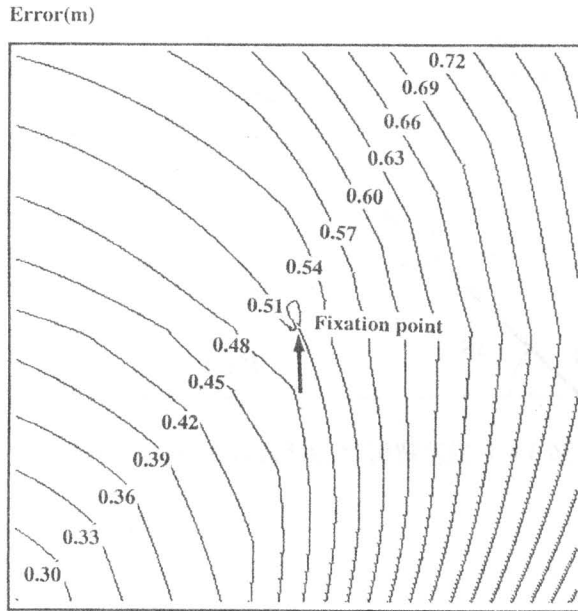


図3.4.5 注視制御を行わない場合の位置決め誤差

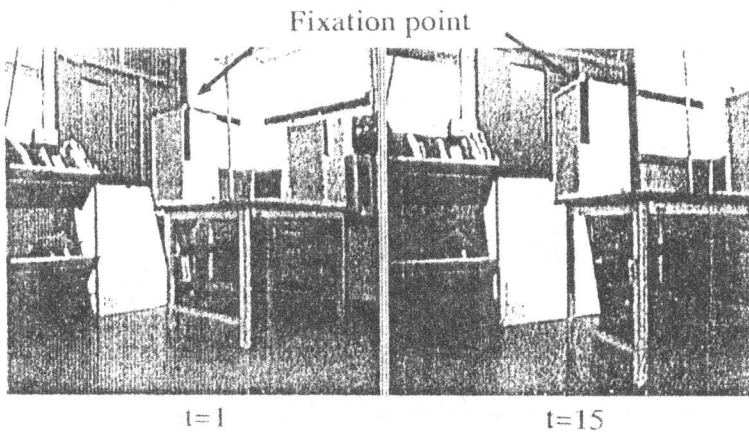


図3.4.6 注視制御中に撮像された画像

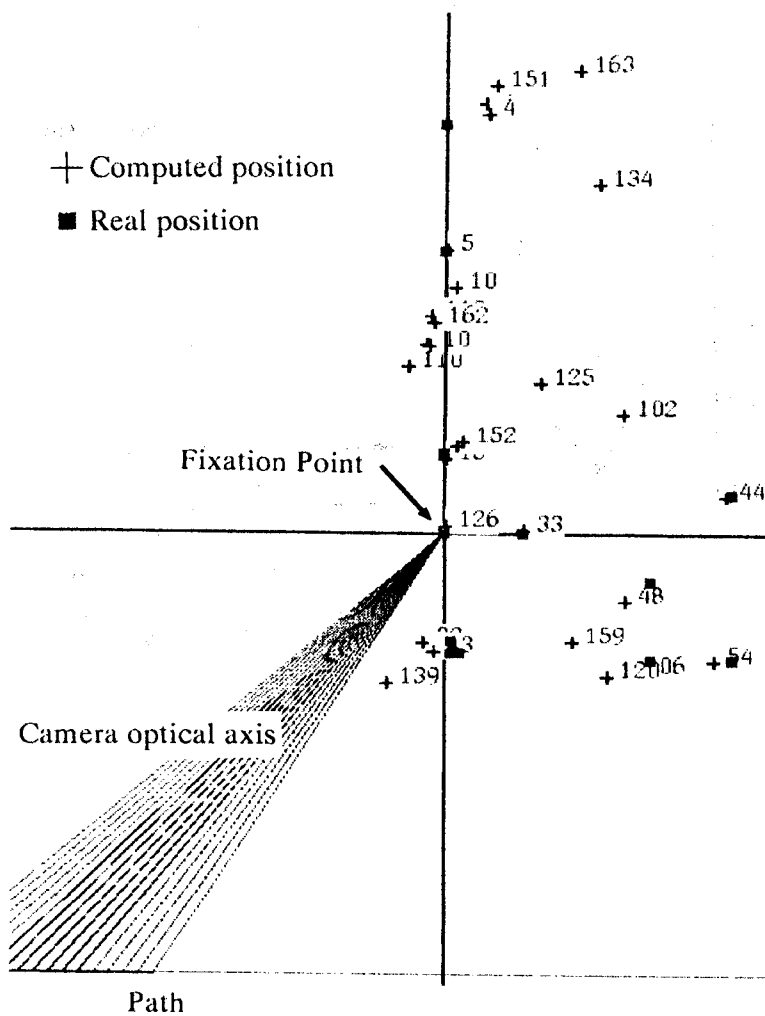


図3.4.7 注視制御によって獲得された局所地図

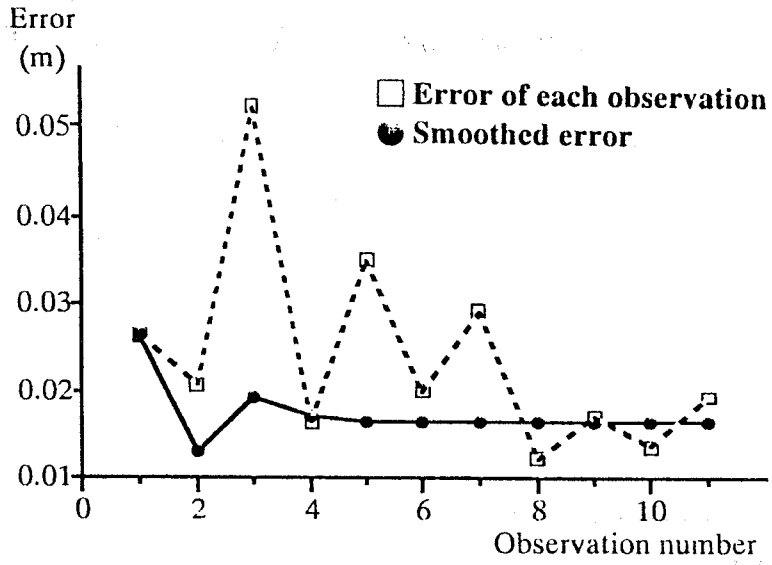


図 3.4.8 カルマンフィルタによる誤差の収束

### 3.5 離散的観測における局所地図の獲得

3.4では画面速度の連続的な値が観測できると仮定して計算式を導いたが、ここでは、離散的な観測に基づく計算式を導き、その計算式の誤差解析と、誤差解析に基づく観測計画について述べる。

#### 3.5.1 局所地図の計算方法

ロボットは直線運動しながら、回転台の機能を用いて注視点を画像中心に固定する。図3.5.1に示す幾何学的関係より、式(3.5.1)に示すように、局所地図は注視点を中心とする極座標 $(\theta, L)$ で表すことができる。

$$\theta = \frac{\sin\beta_1 \cos(\varphi - \Delta\alpha) \sin(\Delta\alpha - \beta_2 + \beta_1)}{\sin\beta_1 \cos(\varphi - \Delta\alpha) \cos(\Delta\alpha - \beta_2 + \beta_1) - \sin\beta_2 \cos\varphi} - \beta_1 \quad (3.5.1)$$

$$L = \frac{D \sin\beta_1}{\sin(\theta + \beta_1) \cos\varphi}$$

ここで、 $\varphi$ はロボットの直線運動の方向、 $\Delta\alpha$ はロボット2点間のカメラ回転台の回転角度、 $x$ は物体の画面上での位置、 $\Delta x$ はその画面速度である。また、 $\beta_1$ 、 $\beta_2$ はカメラの焦点距離 $f$ を用いて、それぞれ、 $\tan^{-1}(x/f)$ 、 $\tan^{-1}((x+\Delta x)/f)$ として計算できる。

この計算方法の利点は、3.4と同様に、ロボットの移動速度を用いずに、局所地図を得ることができる点である。

#### 3.5.2 誤差解析

次に、この注視制御を用いた局所地図の計算方法の利点を明らかにするために、式(3.5.1)の誤差解析を行う。誤差の評価方法は、3.4.2で述べた誤差解析と同様に、式(3.5.1)を各計測パラメータで偏微分する方法を用いる。

式(3.5.1)における各パラメータとその計測誤差を次のように仮定する( $U$ はロボットの移動距離、画像の大きさは $512 \times 512$ 画素)。

$\varphi = 45.0$ degrees	$\Delta\alpha$ の計測誤差 = 0.1 degrees
$D = 5.0$ m	$\Delta x$ の計測誤差 = 1.0 pixel
$f = 600.0$ pixel	$x$ の計測誤差 = 1.0 pixel
$U = 1.0$ m	$\varphi$ の計測誤差 = 0.1 degrees
	$f$ の計測誤差 = 10.0 pixel

このとき、物体の位置決め誤差は図 3.5.2 のようになる。図 3.5.2 では各点の位置決め誤差の大きさを高さとして表している。この図よりわかるように、連続的観測をもとに場合 (図 3.4.4) と似た形しており、位置決め誤差は、注視点付近ではその増加の割合は小さく、注視点から離れるほど大きくなる。離散的観測においても、注視制御を行わない場合に比べ、注視制御を行う場合の方が注視点付近で物体の位置決めを正確に行うことができる。

いろいろなパラメータの値を設定して誤差解析した結果、注視制御を用いた計算方法は、注視点付近で、ロボットの移動方向を示すパラメータ  $\varphi$  の計測誤差に影響されることは少なく、影響される場合でも、ロボットの移動距離  $U$  を大きくすることで位置決め誤差を小さくできることがわかった。これについては、カメラの回転角度  $\Delta\alpha$  についても、同様のことが言える。

### 観測計画

誤差解析では、注視制御を用いることで注視点付近の位置決め誤差が小さくなることを示したが、次に、物体の計測において必要なだけの精度を得る方法について考える。絶対的な位置決め精度を向上させるには、

- (1) 焦点距離の長いカメラを用いて、カメラの回転角度の計測精度、物体の視差  $\Delta x$  の計測精度を向上させる。
- (2) エッジの位置をサブピクセルのオーダで測定する方法など、画像処理の手法により、物体の視差  $\Delta x$  を正確に計測する。
- (3) ロボットができるだけ長い距離を移動する。

などの工夫が考えられるが、ここでは (3) を調整することにより位置決め精度を向上させる。

絶対的な位置決め精度は、カメラの視野と注視点の位置から決定される円内(注視制御により常に観測できる範囲)の平均の位置決め精度として考える。

注視点に対するロボットの移動方向を示すパラメータ $\phi$ を決定すると、図3.5.3に示すようなグラフを描くことができる。このグラフは、異なるカメラの焦点距離 $f$ に関して、ロボットの移動距離(図3.5.1における2つの観測点間の距離)と絶対的な位置決め精度の関係を示すものであり、これを利用することにより、必要な精度を得るための移動距離を決定することができる。もちろんこのグラフは $\phi$ に依存するため、複数の $\phi$ に対して計算する必要がある。

### 3.5.3 実験結果

3.4.4と同様の実験設備を用いて、図3.5.4に示す模型を対象として実験を行った。図3.5.4(a)は対象物を正面から見た図であり、図3.5.4(b)は対象物を上から見おろした図である。図3.5.4(b)の矢印は注視点の位置を示す。図3.5.5は図3.5.4(b)に示すロボットの位置AB間での観測結果である。図3.5.5は図3.5.4(a)に示す物体全面の特徴点の位置を、地面に平行な2次元平面で表したものである。図3.5.4(b)と比較すると、ほぼ物体の形を反映していることがわかる。

図3.5.5の実験において、ロボットの移動距離をさらに長くした実験結果が図3.5.6である。図3.5.6では図3.5.4(b)に示すロボットの位置AC間で観測した結果である。図3.5.6と図3.5.5を比べると、例えば、ラベル12から17までの特徴点は、本来一直線上にあるべきものだが、図3.5.6ではほぼ一直線に並んでいるのに対し、図3.5.5では一直線上に並んでいない。他の部分を比較しても、明らかに、図3.5.6の方がより正確な対象物の形を復元している。すなわち、ロボットの移動距離が長いほど正確な環境構造が復元できる。

次に、ロボットの移動方向を表すパラメータ $\phi$ と、カメラの回転角度 $\Delta\alpha$

に計測誤差があった場合、観測結果がどのように変化するかという実験を行ってみた。

図3.5.7は、図3.5.5の実験において、 $\varphi$ の値を $\pm 5^\circ$ 変化させた結果である。この結果から、 $\varphi$ の値が多少違ってても、観測結果全体が回転するだけで、対象物の形そのものには影響を与えないことがわかる。

一方、図3.5.8は $\Delta\alpha$ の値を $\pm 5^\circ$ 変化させた場合である。この場合には、 $\Delta\alpha$ の値により、対象物の形が大きく変化することがわかる。

この実験結果から言えることは、対象物の形の復元のみを考えれば、 $\varphi$ の値はそれほど正確である必要がないということである。すなわち、ロボットの進行方向に対するカメラ光軸の方向を正確に求めなくとも、対象物の形状は復元できることを意味する。

#### 3.5.4 考察

離散的観測をもとにした局所地図の獲得においても、連続的観測をもとにした場合と同様注視点付近での位置決め誤差が、他の部分よりも小さくなることがわかった。さらに、ロボットが長い距離を移動することで、相対的にカメラの回転角度の計測誤差を小さくでき、その結果、正確な対象物の形状が復元できることを確かめた。

以上のように、離散的な観測の利点は、ロボットが長い距離を移動して観測を行う点にあるが、欠点としては、立体的な対象物の角を注視点として選んだ場合、ロボットの移動に伴い、見え隠れする部分が多く、全ての特徴点を十分長い間観測することが難しい場合もある。

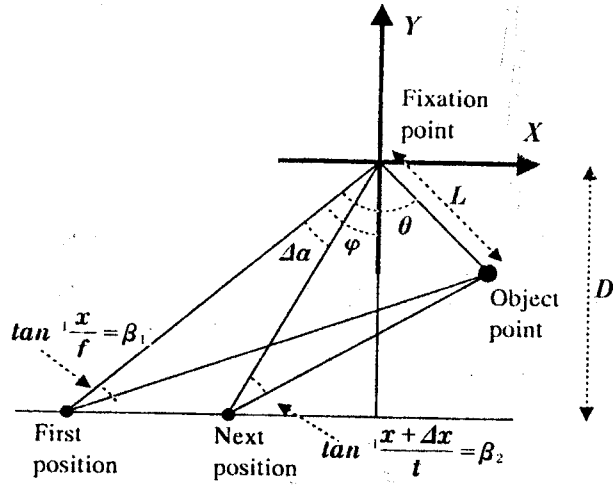


図 3.5.1 離散的観測における局所地図の獲得

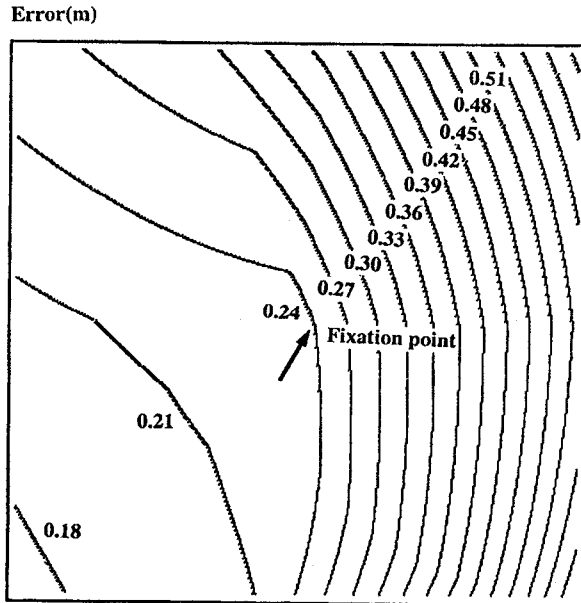


図 3.5.2 離散的観測の誤差解析



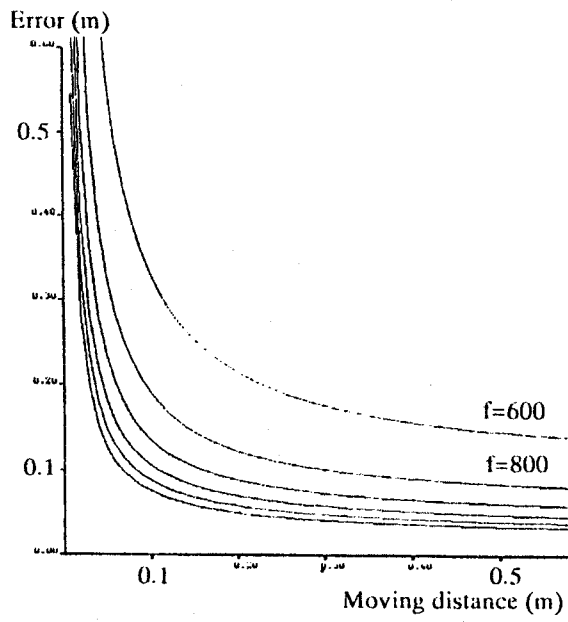


図 3.5.3 観測計画

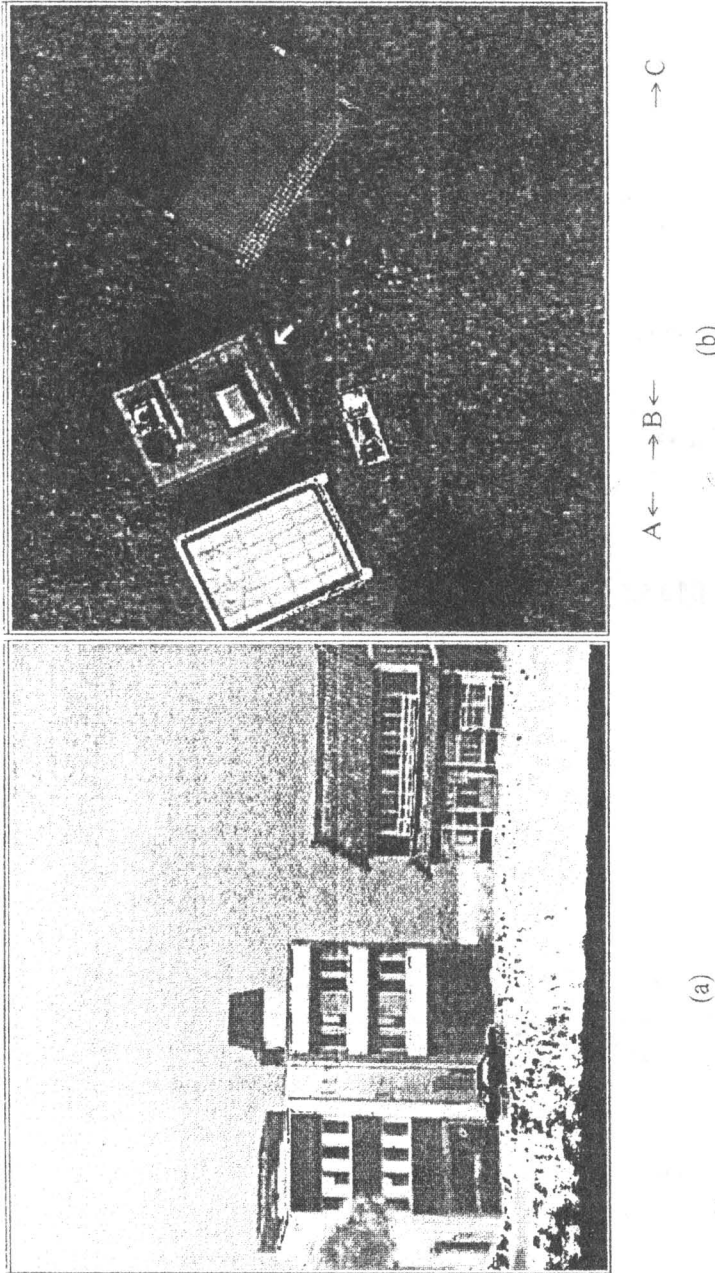


図 3.5.4 実験に用いた模型

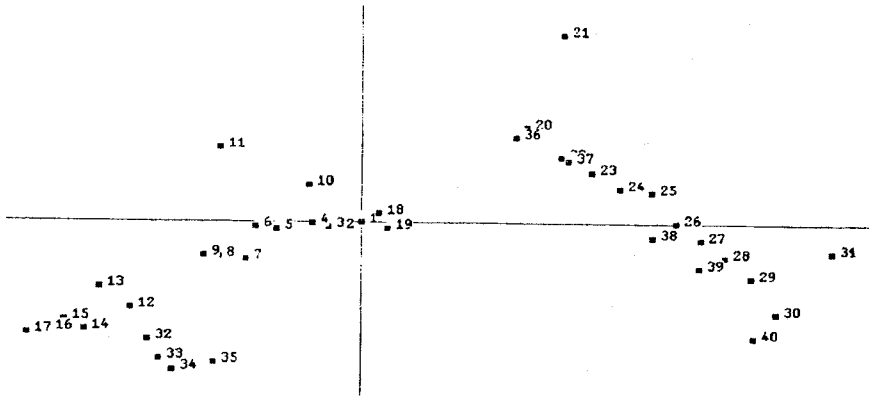


図 3.5.5 図 3.5.4(b)における AB 間での観測

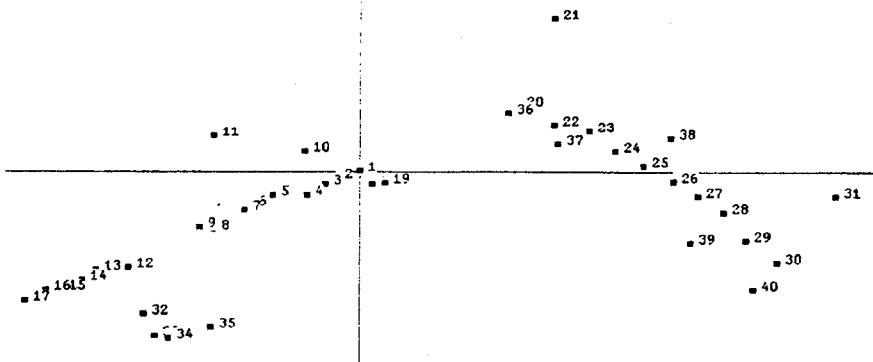


図 3.5.6 図 3.5.4(b)における AC 間での観測

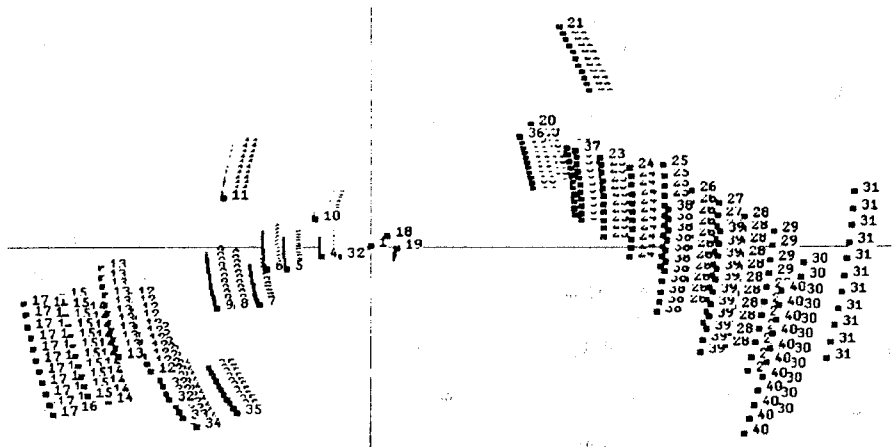


図 3.5.7  $\phi$  の誤差による局所地図の変化

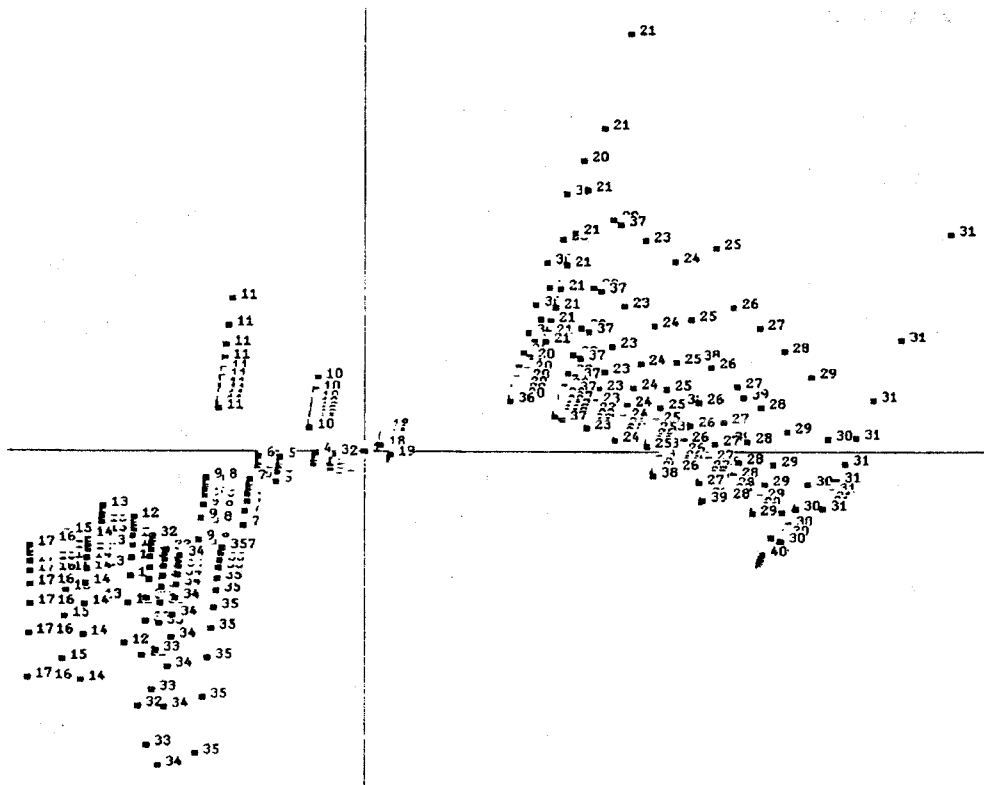


図 3.5.8  $\Delta\alpha$  の誤差による局所地図の変化

### 3.6 注視点選択に関する考察

注視制御に関する研究において、注視点の選択は重要な問題である。注視点の選択は、ロボットの目的や移動する環境などによって左右される。また、それだけではなく、環境の認識のレベルにも依存する問題である。しかし、それらを全て満たすような注視点の選択方法を考えることは難しい。

ここでは、屋内環境を移動するロボットが、3.2で述べた戦略を実行する際に、注視点とすべき対象や、その注視点の選択方法について考える。このためには、3.2でも述べた、注視制御を行うカメラシステムとは別の、第二のカメラシステムが必要となる。このカメラシステムは、ロボット上に固定された視野の広いものであり、そこから得られる情報をもとに注視点を選択する。

#### モーションステレオによる環境構造の復元

ロボットは、ロボット上に固定された視野の広いカメラを用いて、モーションステレオにより環境の構造を認識しながら環境内を移動する。これにより、正確ではないが、環境全体の大まかな構造を知ることができ、簡単な構造を持つ環境なら移動可能である(例えば障害物のない単純な廊下を例にあげることができる)。

このモーションステレオから得られる環境構造に関する情報は、ロボットの基本的な移動の実現と、注視点選択の基準を得るために用いる。

モーションステレオで得られる局所的な環境構造の評価は、ロボットの車輪に取り付けられたエンコーダの情報を用いて、移動経路に沿って結合される。

エンコーダによる計測は常に正確ではなく、モーションステレオで得られる環境構造の評価は、局所的には比較的正確であっても、全体としてはゆがんだものになる可能性がある。さらに、視野の広いカメラを用いてい

るため、環境構造の評価は比較的粗いものである。

しかし、我々の目的は環境構造の正確な評価を得ることではなく、ロボットが環境内を移動するための情報を得ることであり、モーションステレオによって得られる環境構造の評価は我々の目的をほぼ満たすものと考えられる。もちろん、環境やロボットの目的によって正確な構造を知る必要がある場合もある。そのような場合には、後に述べる注視制御を用いた観測により、環境の正確な構造を調べるのである。

ロボット上に固定された視野の広いカメラからの情報は、モーションステレオによる環境構造の復元に利用するだけでなく、色情報やオプティカルフロー等の獲得にも用いる。

#### 環境認識を目的とした注視点の選択

環境の大まかな構造を観測しながら移動するロボットにとって、注視し、詳しく観測する必要のある部分(注視点)は、ロボットの目的と環境に応じて選択する必要がある。

屋内環境を移動するロボットの注視対象としては、

- (1) ランドマーク
- (2) 複雑な形状の静止物体
- (3) 環境の構造を特徴づける部分
- (4) 移動物体

の4つが考えられる。

(1)のランドマークはロボットの位置決めに必要な情報であり、あらかじめ記憶されているランドマークのモデルとの照合を正確に行うには、ランドマークを注視し、ランドマークが最も正確に観測できる位置に移動する必要がある。

(2)の複雑な形状の静止物体は、先に述べたモーションステレオではその構造を正確に測定できないものをさしており、これらは焦点距離の長いカメラを用いて注視することで正確な構造を評価することができる。

(3)の環境の構造を特徴づける部分としては、コーナなどが考えられる。これらは、あらかじめ与えられている環境モデルとの照合において重要な特徴となるため、できるだけ正確な構造を知る必要がある。

(4)の移動物体は、ロボットにとって障害物となる可能性がある。そのため、ロボットは移動物体を発見したら直ちに静止し、注視しながら移動物体の移動速度、移動方向、形状などを計測し、障害物となるか否かの判断を行わなければならない。

注視点の選択方法は、注視対象に依存するものであり、それぞれの注視対象はそれぞれ異なった基準で選択する必要がある。先に述べた注視対象ごとに考察する。

(1)のランドマークは視覚的な特徴がはっきりしているものが用いられるため、その特徴をキーに選択すればよい。それらの特徴としては赤や黄色などの目立つ色情報、正方形や三角形のようにはっきりとした形状の情報と考えられる。特に色情報を注視対象の選択基準に用いる研究は、すでにいくつか報告されている [Swain 90]。

(2)の複雑な形状を持つ物体は、エッジ画像を作成した際に、局所的にエッジが多く存在する部分を抽出することで、注視対象を選択することができる。図 3.6.1 は、ロボット上に固定された視野の広いカメラで撮像された画像を示す。図中に注視対象の例を示す。

(3)の環境の構造を特徴づける部分は、コーナなどの壁と壁の接する部分である。これらは、ロボット上に固定された視野の広いカメラの観測により、その候補を選択することができる。

ロボット上に固定された視野の広いカメラでモーションステレオを行った場合、正確ではないが大まかな構造は復元できる。復元された環境構造にはその環境を特徴づけるコーナなどもある程度再現されており、少なくとも注視対象の候補の選択は可能である。

注視制御を用いてコーナを観測する利点は、コーナの形状を正確に復元できるばかりではなく、能動的にコーナの形状を探索することにある。

すなわち、コーナに注視することにより、コーナの向こう側に広がる通路を早く発見するには、コーナを注視する動作が最も効率が良い。

(4)の移動物体は、ロボットが静止している場合にはある一定時間おいて撮像された2つの画像の差分を求めることで発見でき、ロボットがモーションステレオを行うために直線運動をしている場合には、オプティカルフローを観測することにより発見できる。ロボットが直線運動を行う場合、静止物体のオプティカルフローはある一点のFOE(focus of expansion)を持つ。しかし、そのなかで運動する物体のオプティカルフローは静止物体とは違ったFOEを持ち、これを利用して移動物体を発見することができる。



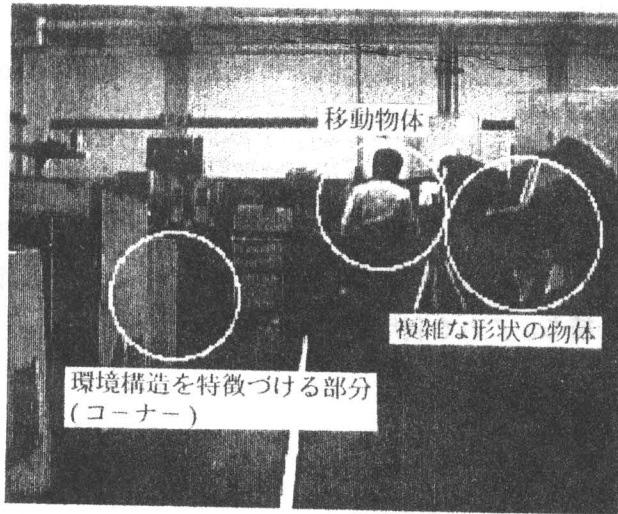


図3.6.1 移動ロボットの注視点

### 3.7 まとめ

注視制御を用いた視覚移動ロボットの環境認識における一戦略を提案した。ロボットが直線運動する際に、連続的な観測をもとに、注視制御により直接物体中心の座標系で表わされる局所地図を獲得する方法を示した。

誤差解析を行うことにより、注視制御の位置決め精度における利点を明らかにした。誤差解析は、連続的観測をもとにした方法だけでなく、離散的観測をもとにした方法においても行った。

実験においては、カルマンフィルタにより正確な局所地図を獲得する方法を提案した。

注視制御を移動ロボットの環境認識に用いる利点は以下の4点である。

- (1) 物体中心の環境記述を直接得ることができる。
- (2) 物体の画面速度を用いて環境構造を復元する方法において、注視点付近の特徴点の位置決め精度が良い。
- (3) ロボットの移動中に、環境内の興味ある部分を十分長い間観測できる。
- (4) 我々が提案したように、視覚移動ロボットの環境認識において新たな戦略を生む。

今後の課題は、我々の戦略を実現するために、ロボット上に固定されたもう一つのカメラシステムとの統合を行い、ロボットが環境内を移動するために十分な情報を持つ全体地図を獲得することである。このためには、注視点の選択方法を具体化するだけでなく、ロボットの観測結果を記録、更新する全体地図の構築方法について十分検討しなければならない。



## 第4章 環境を見渡すカメラ運動

本章では、我々が提案した全方位画像の撮像方法を用いて、環境の全体地図を作成する方法を述べる。我々の方法によって得られる全方位画像は物体への正確な方位角情報と距離情報を持つ。この全方位画像を用いて環境の全体地図は次の手順で得られる。ロボットは全方位画像から得られる距離情報(単眼全方位ステレオ)をもとに自由領域を発見し、環境内を移動する。次にロボットの2つの地点で得られた全方位画像を用いて局所地図を作成する(2眼全方位ステレオ)。局所地図を作成するには、ロボットの2点間の関係を正確に知る必要がある。このために、我々は全方位画像からの視覚フィードバックを用いて、ロボットの運動を制御する方法を提案する。得られた局所地図はそれらの位置関係を決定することで全体地図に統合される。また、この方法では、ロボット中心、経路中心、環境中心の座標系で表される3つの環境表現が用いられる。

## 4.1 環境の全体地図作成に関する研究

環境内を移動するロボットの目的は、与えられた目標点に到達したり、与えられた目標物を探索することにある。このような目的を達成するために、移動ロボットは環境構造を理解する能力が必要とされる。移動ロボットの研究は、あらかじめ与えられる知識によって、以下の3つに大別できる。

- (1) 環境の正確な幾何モデルが与えられている場合
- (2) ランドマークの位置関係によって表される地図を持つ場合
- (3) 未知環境を移動する場合

環境モデルが与えられている場合、ロボットは視覚センサから得られる情報[Matumoto 85]や、幾何モデルから得られる拘束条件[Onoguchi 88]を用いて、与えられた幾何モデルに対する自分の位置を発見できる。問題点は、環境の一部が何らかの要因で変化した場合、与えられた幾何モデルを柔軟に更新していかなければならないことである。

ランドマークは幾何モデルよりも抽象的な表現であり、環境はランドマークを基準とした位置関係によって表わされる。幾何モデルを用いる場合に比べ、あらかじめ行う環境計測の手間を大幅に軽減できる[Komoriya 84, Mori 87]。ロボットはランドマークを発見することで環境内での自らの位置を発見するのであるが、このためには、移動可能な自由領域を発見し、ランドマークを探索する機能が必要となる。

未知環境内を移動する場合、ロボットは移動可能な自由領域を発見しながら環境内を移動し、環境地図を作成しなければならない。本論文では、移動ロボットに搭載された視覚センサ(TVカメラ)を用いて、未知環境の全体地図を作成する新たな手法を提案する。

全体地図を作成する研究としては、ロボットの移動経路に沿ってステレオ視を行うための2枚の画像を撮像し、各観測点においてステレオにより得られた局所地図を統合し全体地図を得る研究[Ayache 87]や、移動経路に沿って連続的に撮像された画像列を用いる研究[Bolles 87]がある。

しかし、これらの研究の焦点は、局所的な構造をモーションステレオや2眼ステレオによって復元することであり、各観測点で得られた局所地図をどのように全体地図に統合するか、また、その過程でどのような環境表現が必要となるかといった、全体地図作成における本質的な問題についてはあまりふれていない。

## 4.2 全方位画像を用いた環境認識

本章では、全方位画像より得られる情報をもとに全体地図を作成する方法を提案し、特に、(1)能動的視覚を用いた移動方法、(2)局所地図の獲得方法、(3)全体地図の作成方法の3つを考察する。

全方位画像の撮像方法としては、魚眼レンズ[Morita 87]や円錐鏡[Yagi 90]を用いる方法が提案されているが、より簡単な方法は、TVカメラを回転させて全方位画像を構成する方法である[Sarachik 89]。カメラを回転させながら、画像中心を通る垂直線上の画像データを横一列に整列させると、その観測点から見える全ての物体を含んだ全方位画像(図4.4.2)を構成することができる。この全方位画像は環境の全体地図を作成するロボットの視覚に適している。我々の提案する撮像方法では、正確な方位角情報と全方位の距離情報を得ることができる。

全方位画像から得られる距離情報は、その撮像機構上の問題により、あまり精度はよくない(実験では、撮像時間を短縮するために、かなり荒い角度幅でカメラを回転させるためである。しかし、この角度幅を小さくすれば正確な距離情報が得られる)。このため、異なる観測点で撮像された複数の全方位画像を用いて、基線距離の大きなステレオ視を行い、より正確な環境地図を作成する必要がある。環境の全体地図は、複数の観測点で撮像された全方位画像から、以下の手順で作成する(図4.2.1)。

- (1) ロボットはある地点で全方位画像を撮像する。
- (2) 全方位画像より得られる距離情報をもとに、移動可能な自由領域を発見する。
- (3) ロボットは全方位画像からの視覚フィードバックを用いながら、次の観測点に移動する。
- (4) 2つの観測点で撮像された全方位画像から局所地図を作成する。
- (5) (1)から(4)の処理を繰り返し、得られる局所地図を順次、全体地図に統合していく。

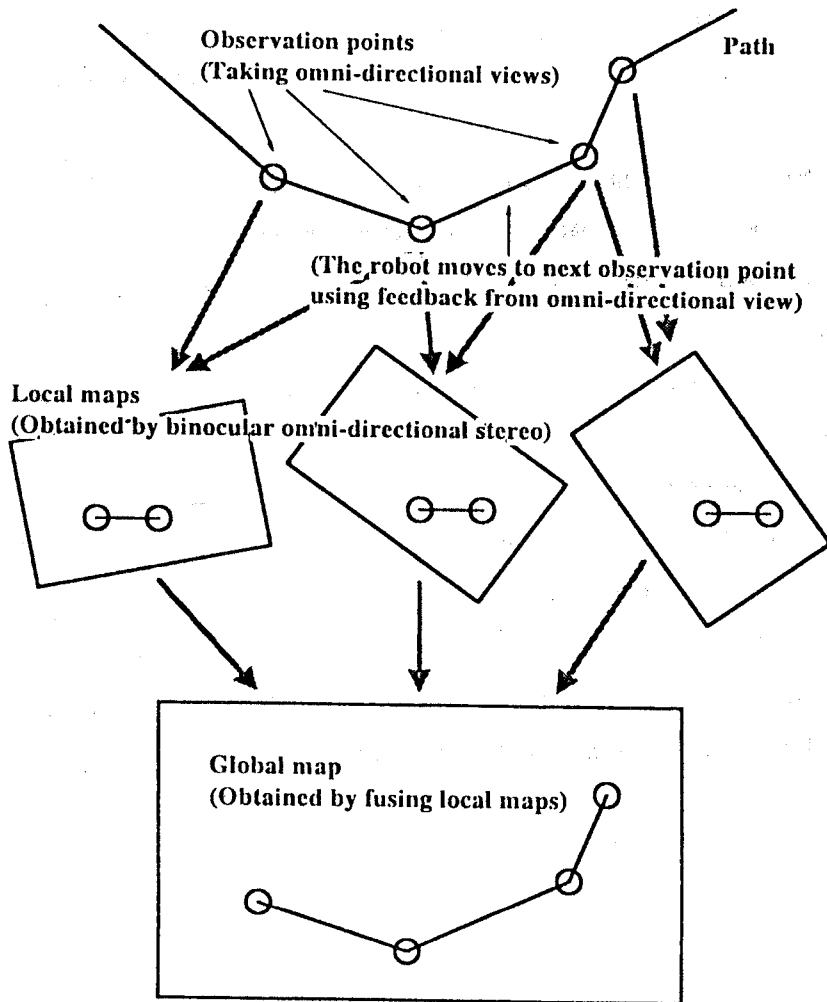


図4.2.1 全方位画像から全体地図まで



### 4.3 ロボットと環境に関する仮定

本章では、ロボットと環境に関して、以下の仮定を設ける。

- (1) ロボットは屋内環境の平らな床面上を移動する。このためロボットの運動は、床面上での位置を表す2つのパラメータと1つの回転角度によって表す。
- (2) 移動ロボットの運動は、車輪に取り付けられたエンコーダで測定するため、車輪のスリップ等により正確な制御は難しい。このため、ロボットが長い距離を移動する場合、誤差の累積を防ぐ何らかの方法が必要となる。
- (3) ロボットは停止して全方位画像を撮像し、撮像を終えてから次の地点へ移動する。
- (4) 全方位画像を撮像するカメラの回転軸は常に床面と垂直であり、カメラの光軸は常に床面と平行に保たれる。
- (5) カメラは正確に制御できるカメラ回転台によって回転する。
- (6) 環境を構成する特徴として垂直線を用いる。屋内環境には多くの垂直線が存在し、ロボットの環境理解に必要な十分な情報を含んでいると考える。また、環境は垂直線の床面上での位置を示すことで表現する。

## 4.4 全方位画像

本節では、一台のカメラを回転させることにより、正確な角度情報を持つ全方位画像と全方位の距離情報を獲得する方法について述べる。全方位画像は、カメラの回転中に、画像面上に設けられた垂直スリットを通して得られる画像データを横一列に整列させることで構成でき、その全方位画像はカメラの回転精度に等しい方位角情報を持つ。一方、距離情報は、カメラを回転台上で回転中心から一定の距離をおいて設置し、カメラの回転中に、画像面上に設けられた2本の垂直スリットを通して得られる2つの全方位画像から獲得できる(全方位ステレオ)。一般の2眼ステレオでは距離情報に画像の量子化による誤差が含まれるが、この方法では、回転台の角度分解能による誤差が含まれる。

この撮像方法から得られる全方位の視覚情報と距離情報を持つ環境表現(我々はパノラマ表現と呼ぶ)は移動ロボットの視覚として有効である。

### 4.4.1 全方位画像に関する研究

視覚移動ロボットにとって、ある地点から全環境を見渡す動作は重要である。ロボットが未知環境に到達した時、ロボットはその環境の正確な地図を作成する必要がある。このような場合、視野の限られたカメラを用いて、環境内をくまなく探索するよりも、全方位の視野情報を利用した方が効率が良い。しかし、一般のカメラは視野が限られているため、ロボットを中心とする全方位の情報を獲得するためには何らかの工夫が必要となる。

八木ら[Yagi 90]は円錐鏡に全方位の環境を映し、それを円錐の軸方向からカメラで撮影して全方位画像を得ている。この方法は実時間でイメージングが可能で、ロボットが直進する時、環境内の静止点のオプティカルフローが円弧を描く性質を利用して解析を行なっている。しかし、全方位の環境を一つの画像面に投影するため、全方位画像に含まれる方位角情報の分解能はあまり高くない(512×512画素の画像面で、全方位画像の最小角度

分解能は約0.2度)。さらに、円錐の先端に近い投影ほど画像情報が粗になるという欠点がある。

Sarachik [Sarachik 89] は、ロボット上に上下方向に取り付けた2つのカメラを回転させることにより、全方位画像を得ると同時に、上下2つのカメラによるステレオ視によって、天井と壁の境界の水平線 (ceiling edge) までの距離情報を獲得する方法を考案した。このシステムでは、室内の壁と天井を同時に撮像できるように、カメラの回転軸は傾けてある。彼女は、天井と壁の境界線までの距離を計測することにより、長方形の室内におけるロボットの位置決めと、回転軸の床面に対する角度を測る方向を提案した。

一方、これら受動的な方法に対して、計測対象に何らかのパターンを投光し、そのパターンまでの距離を測定する能動的な方法がある。Blais [Blais 89] らは、ステレオ法の原理により、物体に投影された光までの距離を計測するコンパクトなシステムを作成し、それを  $360^\circ$  回転させることにより全方位の距離情報を得ている。能動的な方法の欠点としては、パターンを投光するため環境に何らかの制約(とくに照明条件など)を設けなければならないことである。

#### 4.4.2 全方位画像の特徴

本節で提案する方法は、パターンを投光しない受動的な方法であり、Sarachik と同様に回転台を用いてカメラを回転させて全方位画像を撮像する。しかし、彼女が2つのカメラを用いて距離情報を獲得したのに対し、本手法では一台のカメラで距離情報を獲得する。この方法によって撮像される全方位画像には、360度の視野の正確な方位に対する視覚情報と距離情報が同時に含まれる。この視覚情報と距離情報を持つカメラ中心の2-1/2D表現をパノラマ表現 (Panoramic Representation) と呼ぶ [Zheng 90]。Blais らも全方位の視覚情報と距離情報を同時に得る方法を示しているが、彼らとの違いは、パターン投光以外に、本手法では全方位画像に含まれる方位角情報と距離情報はカメラ回転台の回転精度に依存するのに対し、彼らの方法で

は画像の分解能に依存することである。

ロボットの視覚にパノラマ表現を用いる利点を以下にまとめる。

#### (A) 全方位の視覚情報

全方位の視覚情報には次のような特徴がある(付録Gに、全方位画像の詳しい解析を記した)。

(1)長方形の室内で撮像した全方位画像には、90度毎に4つの消失点が現れる。

(2)全方位画像を撮像するカメラの回転軸に垂直な平面上を直線運動する時、環境内の物体のオプティカルフローは、全方位画像上で180度毎に2つのFOE(Focus Of Expansion)が現れる。

(3)複数の全方位画像を用いることで、全方位の環境構造を知ることができる。

(1)の特徴を拘束条件として用いることで、全方位画像内上での消失点の発見はより正確なものとなり、移動中のロボットの回転角度を求めるのに役立つ。(2)の特徴を利用し、ガウス球面上への投影により、ロボットの回転角度と移動方向を求めることができる。(3)を利用する具体的な手段として異なる位置で撮像された2つの全方位画像を用いるステレオ視が考えられる。一般のステレオ視では2つの画像間の対応発見が問題となるが、全方位画像を用いたステレオ視では、全方位画像の構造的な特徴を用いた循環動的計画法(circular dynamic programming)[Zheng 90a]を用いることによって対応を求めることができる。

#### (B) 正確な角度情報

本手法で撮像された全方位画像はカメラ回転台の回転精度に等しい精度の角度情報を含む。すなわち、正確に制御できるカメラ回転台を用いることで、一般の撮像方法の画像に比べて正確な角度情報を獲得できる。正確な角度情報が得られることはたいへん有利な特徴である。例えば、(A)の(3)の特徴で述べた2つの全方位画像によるステレオ視では、全方位の環境構造を知ることができるだけでなく、精度的にも正確な位置決めが可能と

なる。

### (C) 全方位の距離情報

提案する撮像方法では、全方位の視覚情報を撮像すると同時に、撮像された物体までの距離情報を獲得することができる。距離情報は、ロボットが移動する際の自由領域の発見や、2つの全方位画像間での対応探索、ロボットの運動パラメータの決定に利用することができる。

本節では、我々の提案するパノラマ表現が、(B)、(C)の特徴を持つことを中心に述べる。

#### 4.4.3 正確な方位角情報の獲得

図4.4.1は全方位画像を撮像する機構を示す。カメラはレンズの焦点を通る軸を中心に回転し、その光軸は回転軸に直交する。

全方位画像は、カメラの回転中に、画像中心を通る画面上に設けられた幅1画素の垂直スリットの画像を、横一列に整列させて作成する。図4.4.2の全方位画像は計算機室内でカメラを0.4度毎に回転させて得た $900 \times 256$ 画素の画像である。

カメラをさらに細かい角度で回転させることにより、全方位画像は非常に正確な角度情報を含むことができる。以下、この方法について説明する。

一般に、カメラレンズは理想レンズではないために、各種の収差を持っており、正確な焦点面は決定できず、そのため、画面上の物体の投影は必ずある程度のボケを持つことになる。しかし、もし画面上にぼやけて投影された物体の、ぼけたエッジの極大点を検出できるなら、その物体への方位角をサブピクセルのオーダーで求めることができる。この方法はすでにMerickelらのLinear Mixing Model [Merickel 83] で用いられている。

しかし、エッジのボケ全体が画像の1画素中におさまってしまう場合には、正確なエッジの方位角を検出することができない。そこで、カメラを回転させることにより、ぼけたエッジの形を正確に検出する方法を提案する。この方法では、カメラ回転の回転精度でエッジの方位角を測定できる。

実験に用いたTVカメラから得られる画像の角度分解能は $0.1^\circ$ であるのに対し、カメラ回転台は最小 $0.005^\circ$ 毎に回転させることができる。その結果、エッジの方位角はカメラ回転台の分解能すなわち $0.005^\circ$ の精度で測定可能である。

Linear Mixing Modelでは、鋭いエッジや、2本のエッジが1画素の幅におさまってしまうような場合には、正確な方位角が検出できなかったが、提案する方法では、そのような場合にも正確な方位角が検出できる。

#### 正確な方位角情報の獲得方法

正確にカメラを回転させることで、正確な方位角情報が得られる理由はカメラレンズの性質にある。これまでのコンピュータビジョンの研究ではカメラレンズの性質については、あまりふれられてこなかった。しかし、カメラレンズの性質を考えることは重要である。まず、最初にカメラレンズの性質と、実際のカメラレンズを通して得られた画像の基本的な性質について、光学的な立場から考察し、その後で提案する方法について詳しく述べる。

光学的には、理想レンズの焦点面はガウス平面と呼ばれる。一方、実際の実験に用いられるのは理想レンズではないため、各種の収差を持つことになる。それゆえ、実際にはレンズの正確な焦点面の位置を求めることはできない。図4.4.3は実際に用いるレンズによる焦点付近の収束拘束を示す。図4.4.3に示されるように、実際のレンズを用いた場合は、理想レンズの焦点面であるガウス平面、最も明るい像が得られる最大輝度平面、最も解像度の高い像が得られる最大解像度平面の3つの平面が定義できる。

このように、実際のレンズを通して点光源を画像面上に投影した場合、いくらレンズの焦点面の位置を調節しても、画像面上に点光源を点として結像させることは不可能なのである。要するに、実際のレンズを用いれば、画像面上の投影は幾らかボケを持つことになる。

このことを、光学的に説明しよう。画像面上の明るさの分布は、レンズ

の収差や焦点のずれ等のレンズの光学的性質を表す関数と、対象物体の明るさの畳み込み積分として表される。

レンズの収差から生じるレンズの応答関数は、Optical Transfer Function とよばれ、フーリエ変換を用いて表すことができる。図 4.4.4(a) は Optical Transfer Function の例を示す。この関数は光学的な手法を用いて決定することができる。対象物体の明るさの分布関数を  $F(x,y)$ 、Optical Transfer Function を  $G_o(x,y)$  で表すと、画像面上の明るさの分布関数は、式 (4.4.1) で表される。

$$I_o(x,y) = \iint_S F(x-x',y-y')G_o(x',y')dx'dy' \quad (4.4.1)$$

ここで、 $(x,y)$  は画像面上の位置を表し、積分範囲  $S$  は画像面全体を表す。

次に、焦点面の光軸方向の位置ズレによって生じるボケは、図 4.4.4(b) に示すような円柱状の強度分布関数によって表すことができる。この関数を  $G(x,y)$  と表すと、最終的に得られる画像面上の明るさの分布関数  $I(x,y)$  は、式 (4.4.2) で表すことができる。

$$I(x,y) = \iint_S I_o(x-x'',y-y'')G(x'',y'')dx''dy'' \quad (4.4.2)$$

このように、実際のレンズを通して得られる画像は常に幾らかぼけている。しかしながら、この性質を利用することで、正確なエッジの方位角を検出することができる。

図 4.4.5(a) に示すような垂直エッジを持つパターンを CCD カメラで撮像する場合について考えてみよう。図 4.4.5(b) は図 4.4.5(a) の直線 AB に沿った明るさを示す。図 4.4.5(a) にエッジ検出オペレータを適用し、図 4.4.5(a) の直線 AB に沿った位置のエッジ強度を示したのが図 4.4.5(c) である。しかしながら、実際のカメラレンズを通して得られた画像は、上述のように、いくらかぼけているため、実際の撮像面上では図 4.4.5(d) のような明るさの分布を

示し、エッジ強度も図4.4.5(e)のようになる(この形は、撮像面がガウス平面である場合、Optical Transfer Functionを表す)。

ここで、図4.4.5(e)においてその極大値を検出すれば、エッジの位置を正確に求めることができる。しかし、実験した結果、鋭いエッジのボケは図4.4.5(f)に示すように画像の1画素の幅におさまってしまうが多かった。すなわち、単にCCDカメラで撮像した画像からは、図4.4.5(e)におけるエッジ強度の極大値を常に検出することはできない。

そこで、図4.4.5(e)の極大値を検出するために、カメラを量子化された画像の1画素よりも小さい角度幅で回転させる方法を提案する。まず、画面上に幅1画素のスリットを設定する。次に、カメラを微少な角度で回転させながら、スリットを通して得られる画像を横一列に並べ、図4.4.6(b)に示すような画像を構成する。この画像の横方向の角度分解能は、カメラの回転角度分解能によって決定される。図4.4.6(c)は図4.4.6(b)の直線CDに沿ったエッジ強度を示す。図4.4.6(c)には、はっきりとした極大値が現れており、この極大値を検出することにより、少なくともカメラの回転角度分解能に応じた精度でエッジの方位角を決定することができる。一般にカメラを正確に回転させることは容易であり、提案した方法はサブピクセルのオーダでエッジの方位を決定できる。

#### 実験結果

カメラは、DCモータによって回転する回転台上に取り付けた。DCモータは、オプティカルエンコーダのフィードバックを用いて、 $0.005^\circ$  毎に正確に回転する。

図4.4.7(a)はカメラを $0.0^\circ$  から $5.0^\circ$  の間 $0.01^\circ$  毎に回転させながら構成した画像である。画像中には垂直エッジが映っていることがわかると思う。この画像にSobelのエッジ検出オペレータを適用して得られたエッジ画像を図4.4.7(b)に示す。図4.4.7(c)は図4.4.7(b)中の直線ABに沿ったエッジ強度を示す。図4.4.7(c)の極大値を検出することにより、正確なエッジの方位が



決定できる。

図4.4.8は模型を撮像した例で、図4.4.8(b)に2次方程式による近似を用いてエッジの極大値を検出し、正確なエッジの方位を求めた結果を示す。

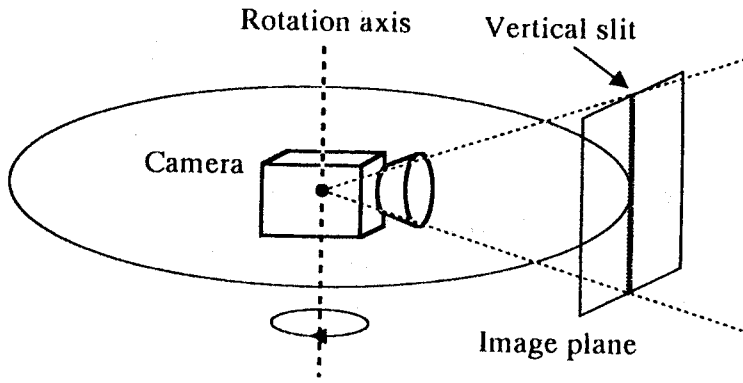


図 4.4.1 撮像機構

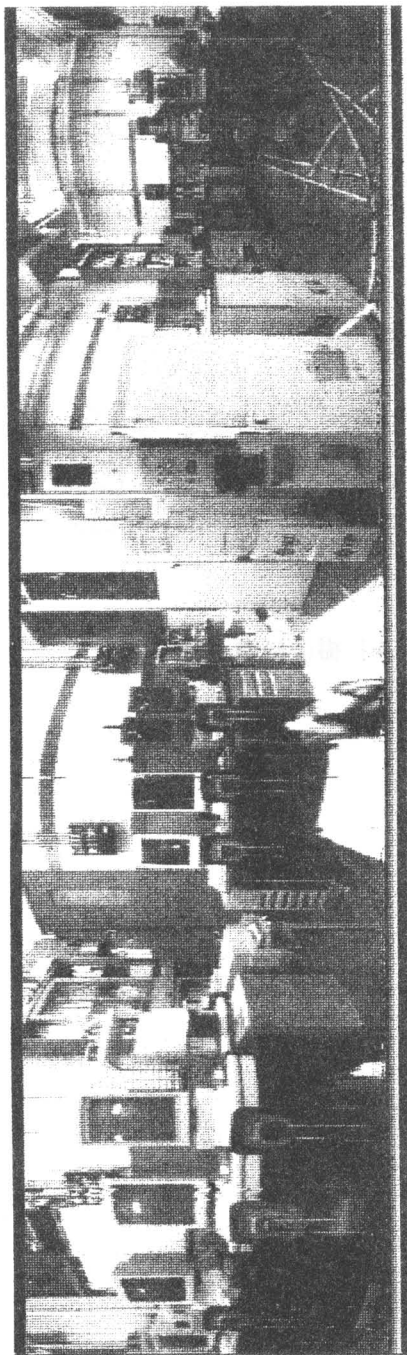


図 4.4.2 全方位画像

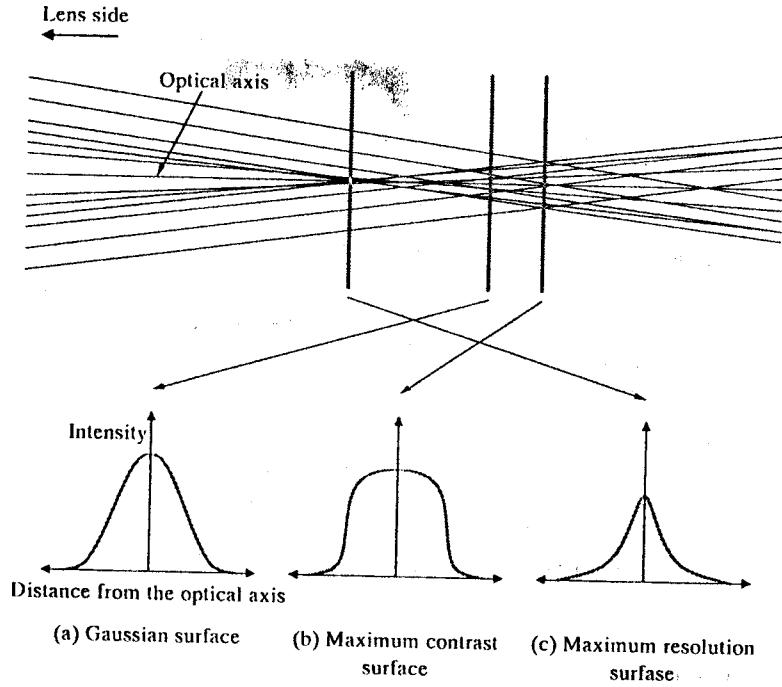


図 4.4.3 焦点付近の収束光束

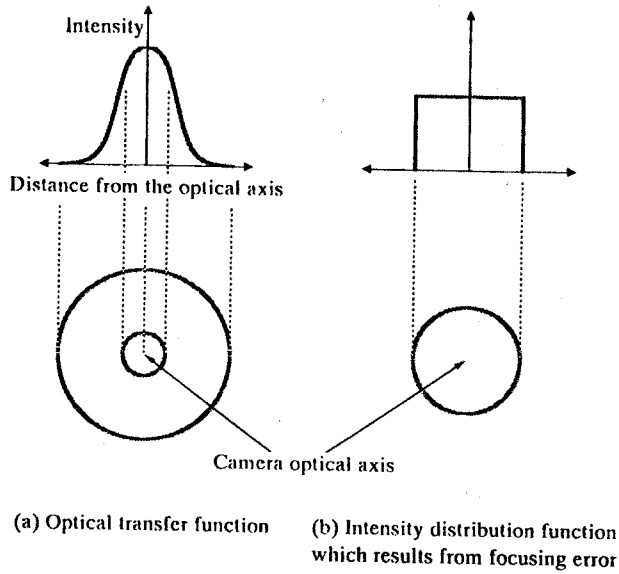


図 4.4.4 レンズの性質

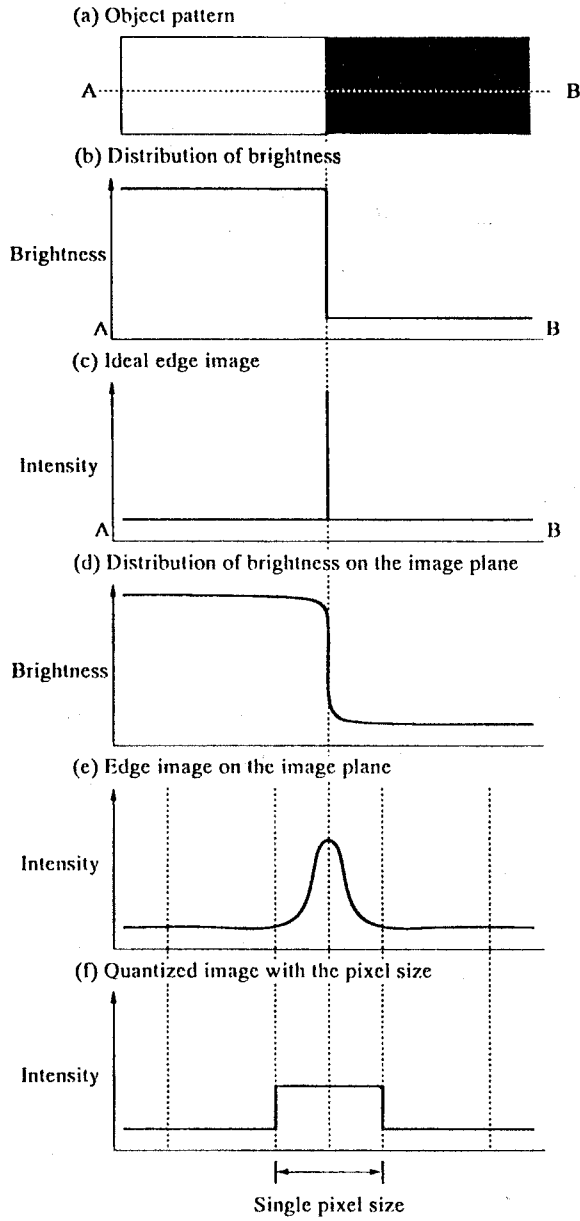
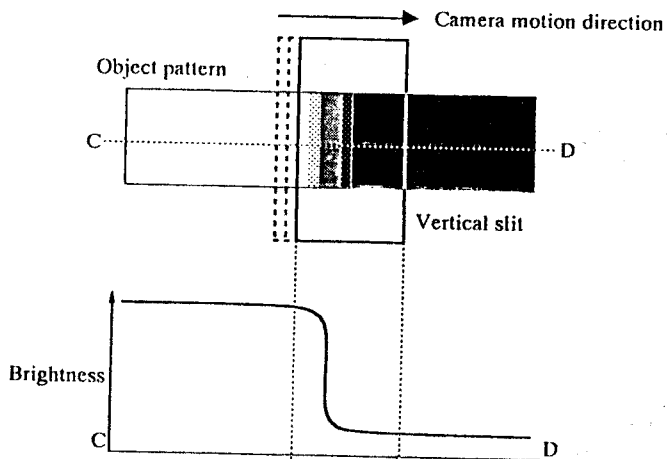
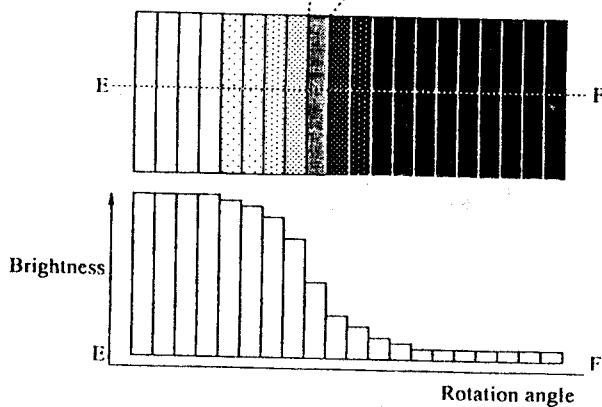


図4.4.5 ぼけの効果

(a) Object pattern and the vertical slit



(b) Image sequence taken through the vertical slit while the camera rotation



(c) Intensity of the dege image of (b)

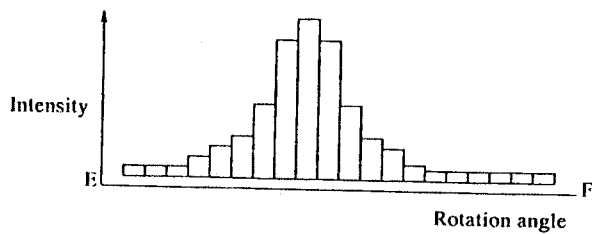


図 4.4.6 カメラ運動によるエッジのぼけの検出

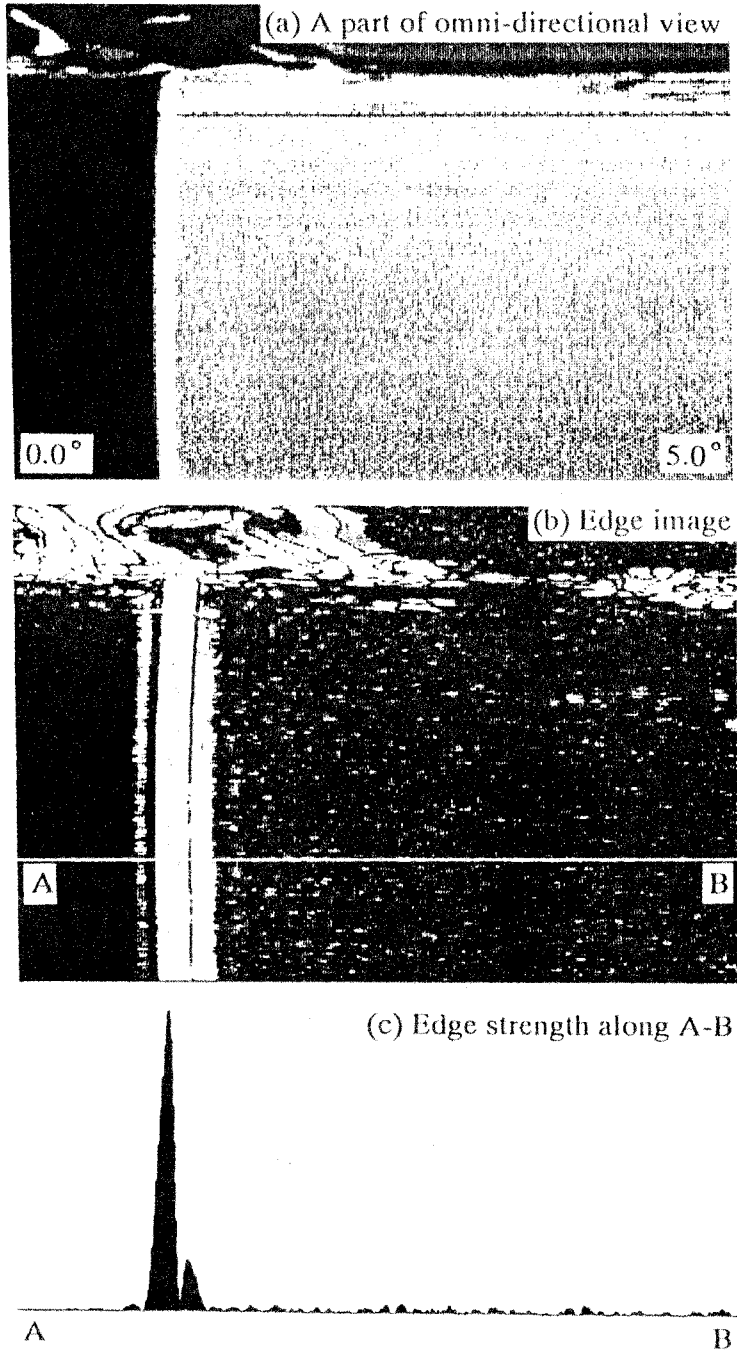


図4.4.7 実験結果1

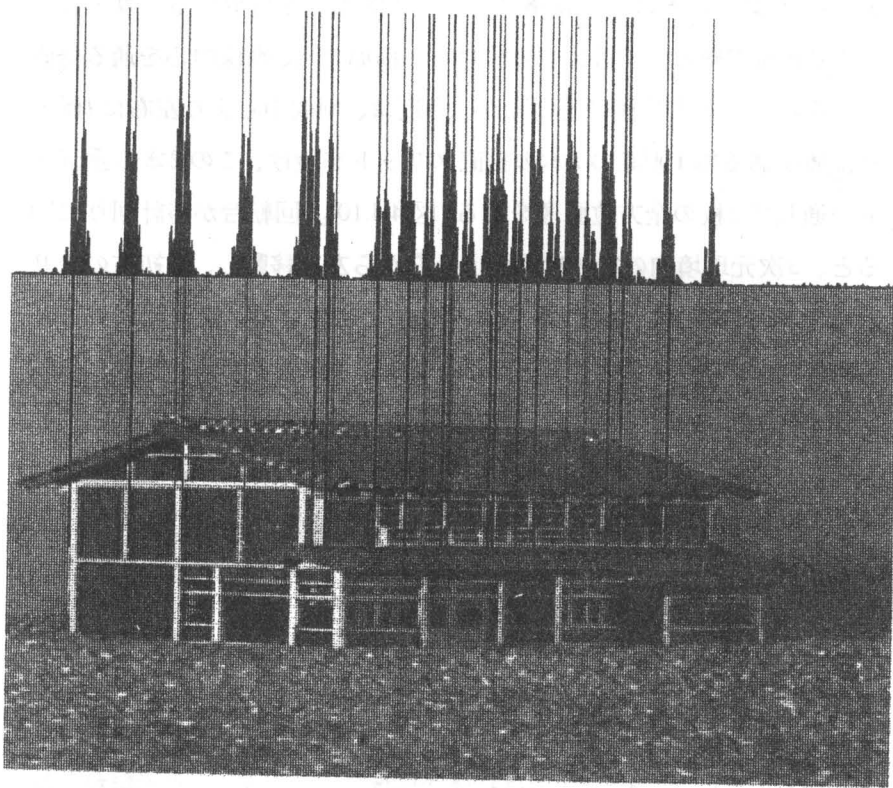


図 4.4.8 実験結果 2



#### 4.4.4 全方位の距離情報の獲得

次に、先に述べた全方位画像を撮像する機構を改良し、全方位の視覚情報と距離情報を同時に得る方法を述べる。

距離画像を得る一般的な方法は、2つのカメラを用いるステレオである。Sarachikは回転軸の上下に取り付けた2つのカメラを同時に回転させて2枚の全方位画像を撮像し、それらの画像間の対応付けを行なって距離情報を獲得した。

一方、我々は図4.4.9に示す機構によって距離情報を獲得する。カメラを回転軸より距離R離れた位置に取り付ける。前節では、画像中心を通る垂直スリットを通して全方位画像を得たが、今度は、画像中心より左右に1画素離れた位置を通る幅1画素の2本の垂直スリットを設け、この2本の垂直スリットを通して2枚の全方位画像を得る(図4.4.10)。回転台が時計回りに回転すると、3次元環境内の物体は画面上を右から左に移動し、最初右のスリット上に現れた物体は、やがて左のスリット上に現れる。この間のカメラの回転角度 $\theta$ を計測することにより、図4.4.11に示される幾何学的関係より式(4.4.3)が得られ、これを用いて物体までの距離を知ることができる。一般のステレオでは2枚の画像間での物体の視差情報より物体までの距離を計測するが、2枚の全方位画像間での物体の角度差情報が視差に対応する。

$$L = \frac{R \sin \varphi}{\sin(\varphi - \theta)} \quad (4.4.3)$$

ここで、 $\varphi$ はカメラの光軸と画像中心より1画素離れた画面上の位置と焦点を結ぶ直線との角度( $\varphi = \tan^{-1}(l/f)$ )、 $\theta$ は物体が画像中心より右に1画素離れた位置から、画像中心まで移動する間のカメラ回転台の回転角度、Lはカメラの回転中心から物体までの距離である。

また、その物体の高さHは式(4.4.4)で計算できる(図4.4.12)。

$$H = (L \cos \theta - R) \frac{y}{f} \quad (4.4.4)$$

高さを計測するには、物体の角度差情報 $\theta$ の他に全方位画像上の縦の位置情報 $y$ と、カメラの焦点距離 $f$ が必要となる。

2つの全方位画像間の対応は、カメラの回転に伴い左のスリットに現れた物体が再び右のスリットに現れるまでの間、全方位画像を作成するために連続的に撮像される画像上で、2つのスリット間を移動する物体を追跡することで容易に発見できる。物体の画面上での移動方向が既知であり、しかも移動量は微少であるため追跡は確実に行える。追跡する物体が他の物体に隠され、追跡の途中でその物体を見失ったとしても、見失うまでのカメラ回転台の回転角度より距離を計測できる。ここでの対応探索の方法は、カメラ運動に伴うオプティカルフローを検出する従来の方法をそのまま用いることができる。

さらに、この全方位ステレオの利点の一つとして、カメラ設置の容易さがある。一般の2眼ステレオでは2台のカメラの間の多くのパラメータを校正しなければならない。その点、一台のカメラを用いる全方位ステレオではその手間が軽減される。以下、この一台のカメラによる全方位ステレオを単眼全方位ステレオと呼ぶ。

#### 単眼全方位ステレオの誤差

一般の2眼ステレオ法では、測定距離には画像の量子化による誤差が含まれる。一方、単眼全方位ステレオでは回転台の回転角度の分解能 $\Delta\theta$ が一般のステレオ法における画像の分解能に相当する(図4.4.13)。回転角度に $\Delta\theta$ の誤差が含まれるとき、特徴点までの距離の誤差 $\Delta L$ は式(4.4.5)で表すことができる。

$$\Delta L = \frac{\sqrt{1 - \left(\frac{R \sin \varphi}{L}\right)^2}}{R \sin \varphi} L^2 \Delta \theta \quad (4.4.5)$$

式(4.4.5)より、物体までの距離が十分遠い場合( $L \gg R \sin \phi$ )、誤差 $\Delta L$ はカメラの回転中心からの距離 $L$ の2乗に比例し、 $R \sin \phi$ (一般の2眼ステレオでは2つのカメラ間の基線距離に相当する)に反比例することがわかる。

図4.4.14は回転台の角度分解能 $\Delta\theta$ を0.4度に設定したときの違った $R$ に対する物体までの距離とその測定誤差の関係を表している。言うまでもなく、距離測定誤差を小さくするには $R$ を大きくすればよい。しかし、この撮像系を移動ロボットの視覚に用いる場合、ロボットに搭載できる回転台の大きさは限られ、 $R$ の大きさには限界がある。後に示す実験ではカメラの焦点距離 $f=600$ 画素、 $R=0.2\text{m}$ に設定し、画像中心より左右に200画素の位置に設置した2本の垂直スリットを通して撮像された2つの全方位画像を用いてステレオ視を行っている。

しかし、上述の理由から $R$ の長さが制限されるにも関わらず、我々の撮像方法では環境内の物体の方位角がカメラ回転台の分解能に等しい精度で決定できるため、カメラ回転台の分解能 $\Delta\theta$ を非常に小さくでき、その結果、距離計測誤差を少なくすることができる。我々が実験に用いたカメラ回転台は0.005度の分解能を持っており、カメラ回転台を0.005度毎に回転させ、距離計測誤差が回転台の分解能のみに依存すると仮定すると、カメラの回転中心より1.0mの距離にある物体の距離計測誤差は約0.0014mである。

### 実験結果

実験は屋内環境で行った。カメラはその光軸が常に床面と平行になるように、床面から1.0mの高さで回転運動を行う。用いたカメラは松下通信工業製のカラーCCDカメラで、焦点距離が16mmのレンズを備えている。回転台は減速機構としてウォームギアを持つDCモータによって回転し、DCモータはそのモータ軸に直結されたオプティカルエンコーダによりフィードバック制御され、0.005度毎に正確に制御できる。

カメラから入力される画像の大きさは $512 \times 512$ 画素である。キャリブレーションの結果、画像の焦点距離は591画素であった。画像入力装置とし

てNexus6000、回転台の制御装置としてエプソン製のラップトップコンピュータ、ホストコンピュータとしてSun4ワークステーションを用いた。

単眼全方位ステレオにより、カメラを中心とする全方位の距離情報を獲得する実験を行った。

カメラの光軸上に回転台の回転中心が常にあるようにカメラを配置することが必要である。このため、回転台の回転中心を通る直線上の回転中心をはさむ2個の特徴点が、カメラを180度毎に回転させた時、常に画像中心にくるようにカメラの方向を調節することで、正確にカメラを回転台上に設置した。また、カメラの回転半径 $R$ は距離が既知の特徴点を用いることで較正な値を測定した。

図4.4.10は、図4.4.9に示す撮像方法から得られた2つの $900 \times 512$ 画素の全方位画像(縦に1/2縮小している)である。カメラを半径 $R=0.2\text{m}$ の円上で0.4度毎に360度回転させ、画面上で画像中心から左右に200画素離れた位置に設けた2つの垂直スリットより撮像した。2つの全方位画像間の対応は2つの垂直スリット間を移動する環境内の垂直線を追跡することで求めた。

図4.4.15は単眼全方位ステレオで測定された環境内の垂直線の床面上での位置を表す。実験結果より、カメラの回転中心を中心とする半径約1.0mの円内では、約0.05mの誤差で垂直線の位置が測定されている。しかし、回転台は最大0.005度の精度で制御できるため、回転台の精度を向上させることで、より正確な距離情報が獲得できると考えられる。

図4.4.16は単眼全方位ステレオによって、全方位画像上に投影される垂直エッジの下端の高さを測定したときの誤差を示している。全体的に誤差は約0.1m以内である。高さ測定に用いる式(4.4.4)には、回転台の精度で測定できるパラメータ $\theta$ の他に、画像の分解能によって精度の決まるパラメータ $y$ が含まれているため、床面からの高さ測定の精度は、床面上での位置測定の精度より悪い。

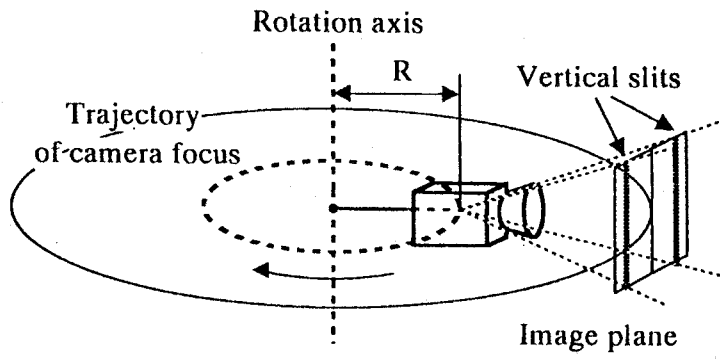


図 4.4.9 撮像機構

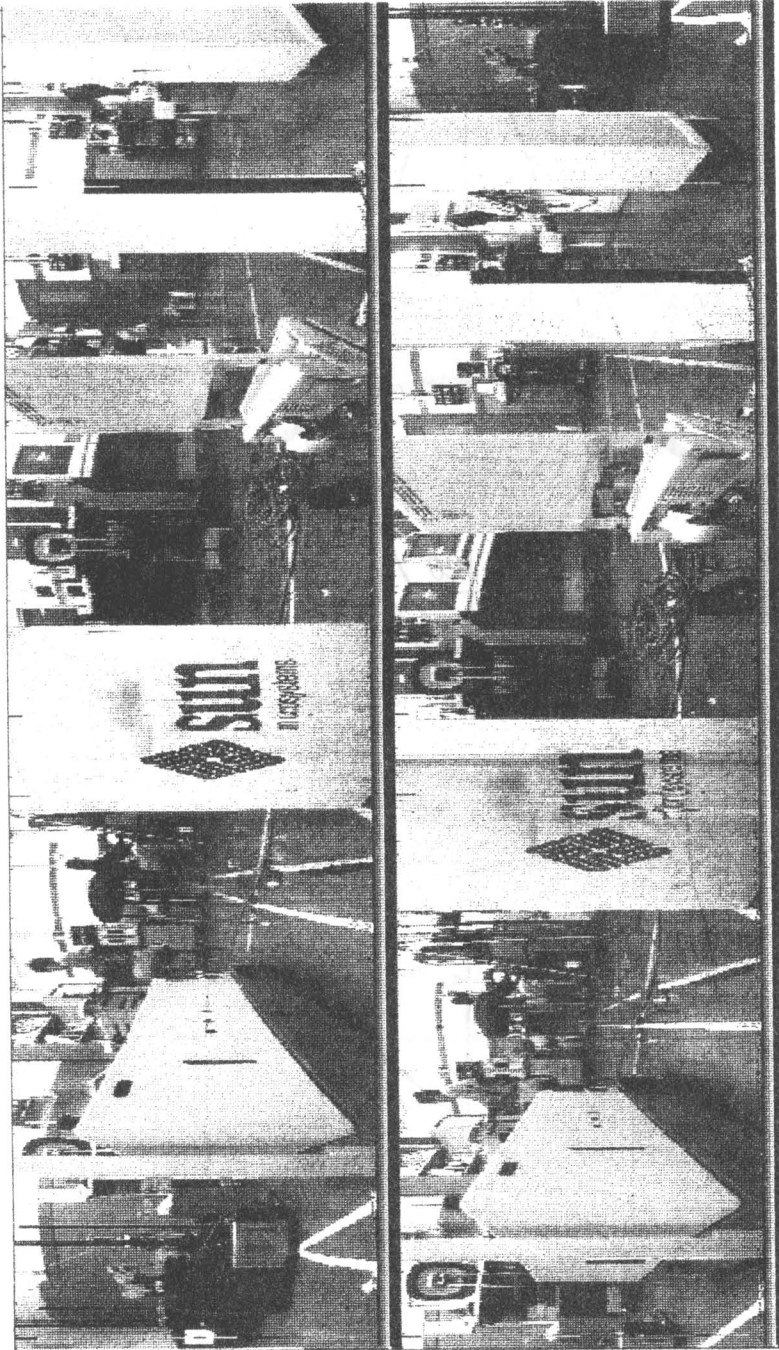


図 4.4.10 距離情報を得るための2つの全方位画像

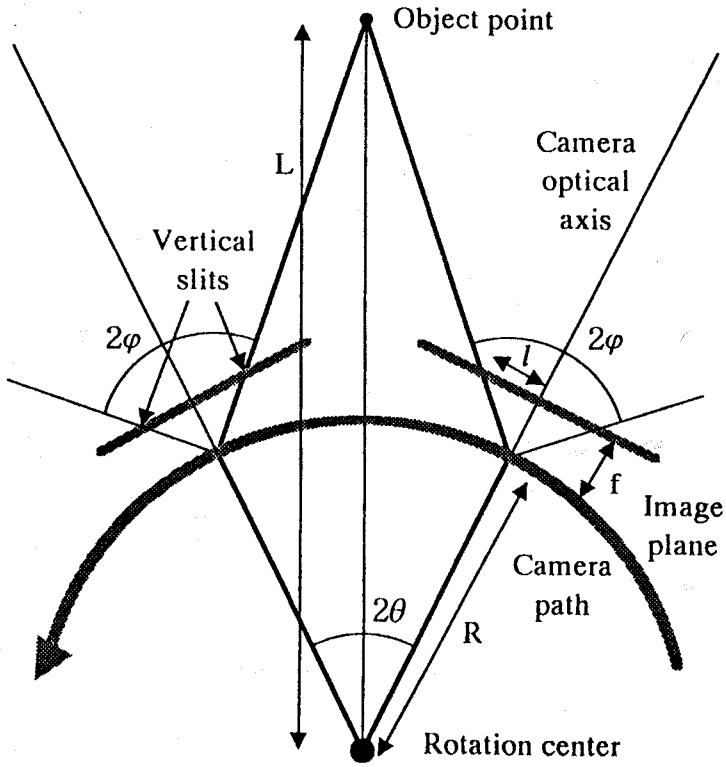


図4.4.11 単眼全方位ステレオによる距離計測

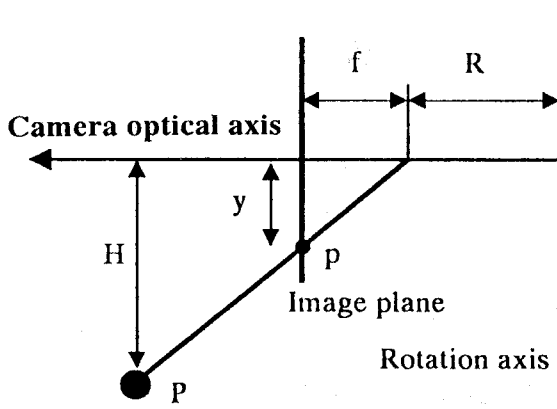


図4.4.12 単眼全方位ステレオによる高さ計測

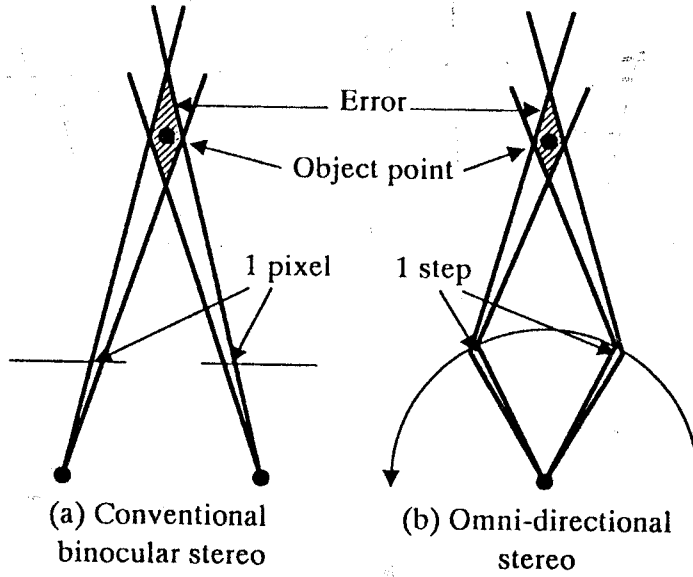


図 4.4.13 位置決めにおける不確定性の比較

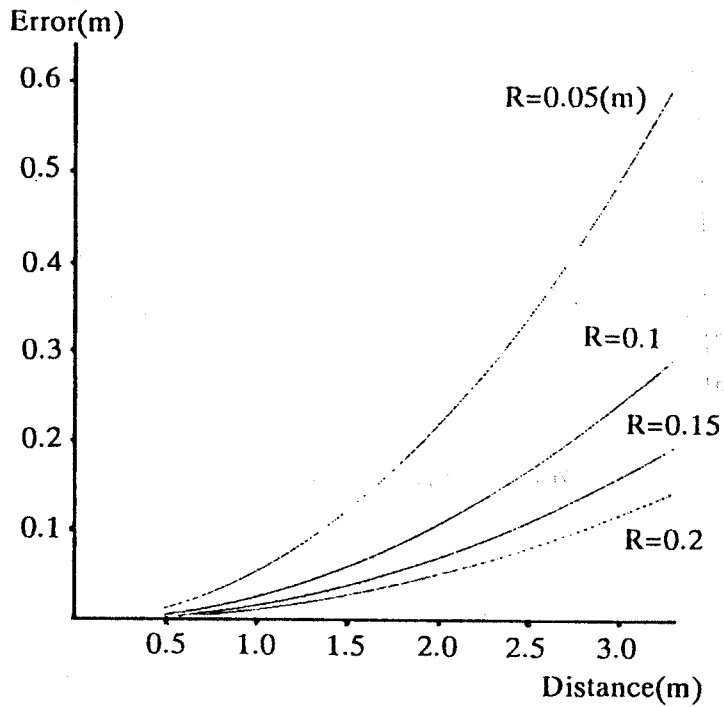


図 4.4.14 単眼全方位ステレオの位置決め誤差



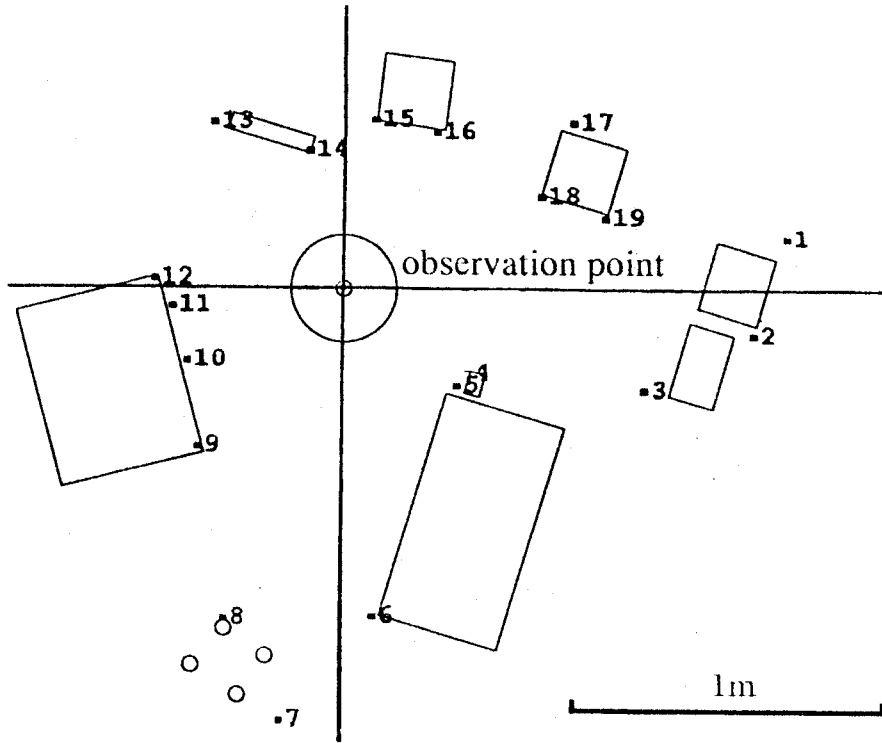


図 4.4.15 距離測定の実験結果

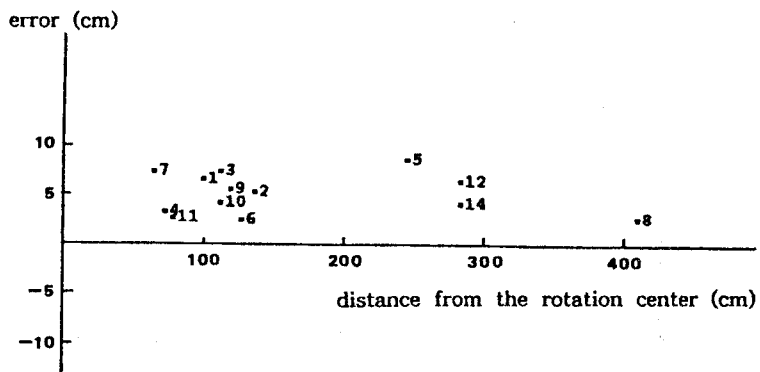


図 4.4.16 高さ測定の実験結果

#### 4.4.5 全方位画像への投影

ここでは、全方位画像上への投影の数学的な性質について述べる。より詳しい考察は付録Xにある。全方位画像は3次元空間の円柱射影として特徴付けられる。3次元空間を円筒座標系で表わす。ただし、 $\zeta$ 軸を回転軸に取り、原点Oをその上でカメラ焦点の高さの点とする。また $\theta$ は、任意の参照軸からの角度である。(図4.4.17、図4.4.18)。

##### 図4.4.1の撮像方法による全方位画像

図4.4.17は4.4.3で述べた撮像方法による全方位画像への投影のモデルを表す。3次元環境内の一点 $P=(\rho, \theta, \zeta)$ は式(4.4.6)によって全方位画像面の点 $p=(x, y)$ に投影される。

$$x = \theta, \quad y = \frac{f\zeta}{\rho} \quad (4.4.6)$$

ここで、 $f$ はカメラの焦点距離である。

$$\rho = \frac{D}{\cos\theta}, \quad \zeta = H \quad (4.4.7)$$

3次元空間内の直線は、座標原点Oと直線を含む平面で円柱を切断した際に現れる楕円上に投影され、それを平面に展開した全方位画像では正弦曲線となる。簡単な例として、参照軸に直交し、 $\zeta$ 軸からの距離がD、高さがHの水平な直線(式(4.4.7))は、式(4.4.8)で示される正弦曲線として投影される。

$$y = \frac{fH \cos x}{D} \quad (4.4.8)$$

しかし、環境内の垂直線は全方位画像上でも垂直線として投影される。

図4.4.2に示されるように、この撮像方法では3次元空間内の物体は全方

位画像上に歪んだ形に投影される。しかし、この全方位画像からもとの画像を復元することは容易であり、この意味において、この方法で撮像された全方位画像は、環境内のある地点でカメラを回転させることによって得られる情報の全てを含んでいる。

#### 図4.4.9の撮像方法による全方位画像

図4.4.18は図4.4.9の撮像方法において、画面上の画像中心を通る垂直スリットを通して得られる全方位画像への投影を表す。3次元環境内の一点  $P=(\rho, \theta, \zeta)$  は式(4.4.9)によって全方位画像面の点  $p=(x, y)$  に投影される。

$$x = \theta, \quad y = \frac{f\zeta}{\rho - R} \quad (4.4.9)$$

ここで、 $f$ はカメラの焦点距離である。

3次元空間内の直線は、円柱面上の網膜の楕円として投影されるのではなく、多少複雑になる。例えば、式(4.4.7)で示される環境内の水平な直線は、式(4.4.10)で示される曲線として投影される。

$$y = \frac{fH \cos x}{D - R \cos x} \quad (4.4.10)$$

しかし、環境内の垂直線は全方位画像内でも垂直線として投影される。

この撮像方法の場合、カメラの運動によってもたらされる情報の全てではないが、ほとんどの情報を含み、それを何らかの方法により抽出することができる。例えば、2.2節で述べた全方位ステレオはその方法の一つであり、全方位画像よりカメラの運動により生じる角度差(一般のステレオでは視差)情報を抽出し、距離情報に変換するものである。

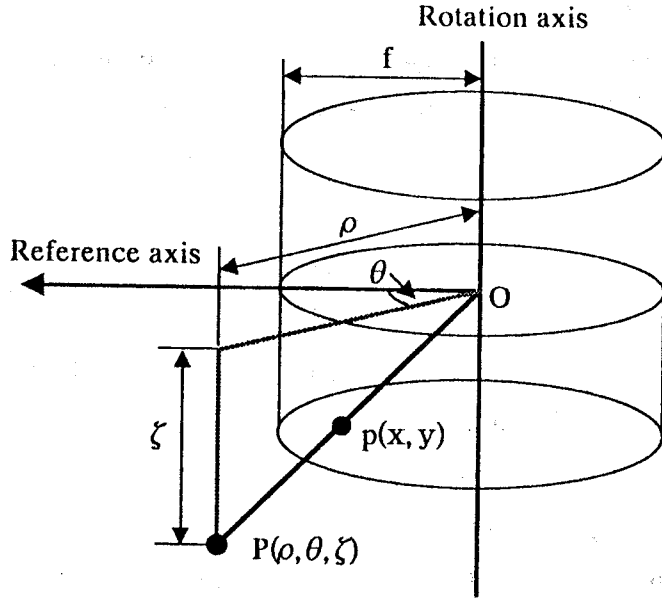


図4.4.17 撮像モデル1

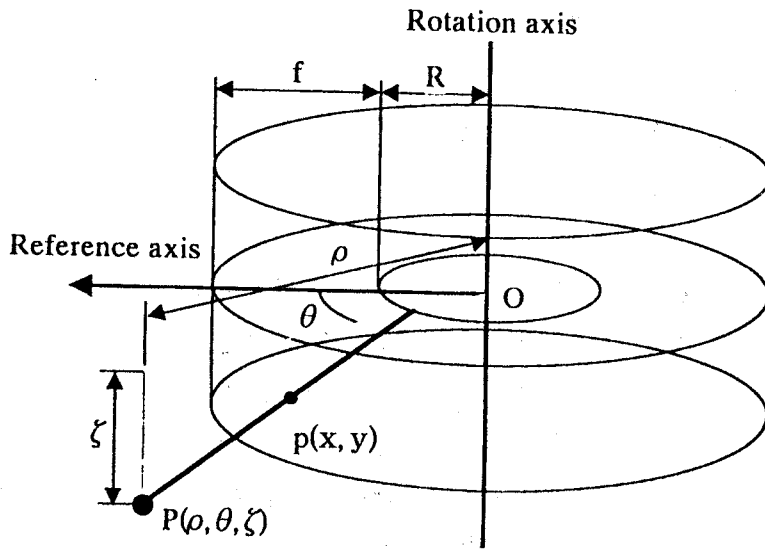


図4.4.18 撮像モデル2

#### 4.4.6 考察

カメラを回転台を用いて回転させることで、正確な角度情報を持つ全方位の視野情報と、全方位の距離情報を同時に獲得する方法を考案した。得られる距離情報の誤差解析、環境内の物体の投影などを考察し、提案する方法の特徴を明らかにした。また、実際に実験を行い、その有効性を検証した。全方位の距離測定を行う単眼全方位ステレオの特徴は以下の3点である。

- (1)一般の2眼ステレオでは計測距離には画像の量子化による誤差が含まれるのに対し、単眼全方位ステレオではカメラ回転台の分解能による誤差が含まれる点である。カメラ回転台の分解能は容易に高くすることができ、その結果、単眼全方位ステレオでは、角度差情報(一般のステレオでは視差)が小さい場合でも、かなり精度の高い距離測定が可能である。
- (2)2つのスリット間で物体を追跡することにより、対応問題を解決できる。
- (3)カメラの設置が容易である。

ここで述べた撮像方法は正確な方位角情報が獲得できるが、これは、従来の2眼ステレオに応用しても有効な方法である。

最近の研究では、誤差を含んだデータを取り扱うのに、最小自乗法に代わり、カルマンフィルタが用いられるようになってきた [Ayache 87, Matthies 89]。例えば、Matthisらはスライダステレオにカルマンフィルタを用いて、正確な物体の3次元位置を出している [Matthies 89]。彼らの研究の本質は、画像の量子化による位置決め誤差を、カルマンフィルタによって収束させることにある。言い替えると、画像より正確な物体の方位角が検出できれば、彼らのようにカルマンフィルタを用いる必要がなくなる。一方、我々が提案した、カメラを回転させて物体の正確な方位角情報が得られる。これを従来の2眼ステレオに用いる2つのカメラに用いれば、物体の正確な3次元位置が検出できるのである。

一方、この方法の問題点は撮像にかなりの時間を要することである。今

後、専用のハードウェアを開発したり、目的に応じて部分的に全方位画像を撮像することで処理の高速化をはからなければならない。

## 4.5 環境の全体構造の復元

本節では、4.4で述べた全方位画像を用いて、環境の全体地図を獲得する戦略について述べる。この戦略により、未知環境の環境構造を知ると同時に、ロボットの移動経路を復元することができる。また、この戦略では、ロボットは全方位画像からの視覚フィードバックを用いて環境内を移動し、観測者中心、経路中心、環境中心の3つの環境表現を用いて全体地図を獲得する。

### 4.5.1 能動的全方位視覚を用いた移動

ロボットはある地点で全方位画像を得た後、移動可能な自由領域を移動し、次の観測点に到達する。ここでも同様に全方位画像を撮像し、得られた2つの全方位画像を用いて、ステレオ視により局所地図を獲得する。ここでのステレオ視を4.4で述べた単眼全方位ステレオと区別するために、2眼全方位ステレオと名付ける。単眼全方位ステレオではロボットの大きさの制約により、正確な距離情報を得ることが難しいが、2眼全方位ステレオでは、十分大きな基線距離をとることにより、正確な距離測定ができる。

2眼全方位ステレオを行う際に問題となるのは、2つの観測点間の位置関係を正確に計測しなければならないことである。しかし、4.3で仮定したように、我々が用いるロボットは正確に制御できない。そのため全方位画像からの視覚フィードバックを利用した移動方法を提案する。

ロボットが、環境内の2点の特徴点を結ぶ直線上を正確に移動し、その移動経路と全方位画像の角度参照軸との関係がわかれば、2眼全方位ステレオは容易に実現できる。2点の特徴点を結ぶ直線上にロボットが位置することを知るには、全方位画像上でその2点が $180^\circ$ の角度を隔てて投影されていることを確認すればよい。すなわち、2点の特徴点が全方位画像中で常に $180^\circ$ の角度を隔てて投影されるように、ロボットの運動を制御すれば、ロボットは2点の特徴点間を正確に直線運動できる。この2点の特徴点を結ぶ

直線は、そのまま2眼全方位ステレオの基線となり、2つの全方位画像の角度参照軸と2眼全方位ステレオの基線との角度は、全方位画像の角度分解能(カメラ回転台の精度)の単位で測定できる。2つの観測点の位置関係を表すパラメータの内、残るパラメータは、2つの観測点間の距離であるが、これはロボットの車輪に取り付けられたエンコーダの情報を用いて評価しなければならない。しかし、この情報は正確でないため、2眼全方位ステレオによって得られる局所地図のスケールファクタとして残す。

上述のように、全方位画像からの視覚フィードバックを用いれば、ロボットは正確な直線運動ができる。しかし、環境内を自由に移動し、様々な観測点から局所地図を作成するには、ロボットは自由に移動方向を変えることができなければならない。移動方向を変えた場合、これまで移動してきた直線経路と、新たに移動経路として選択した直線経路の関係を知る必要がある。そのため、ロボットが移動経路を変更する地点は、2点の特徴点を結ぶ直線が交わる点、すなわち、その地点で得られる全方位画像中に $180^\circ$ の間隔をおいて投影される2点の特徴点の組が2つあるような地点を探索し、その地点でロボットの移動方向を変更する。2つの直線経路の関係は、全方位画像上に投影される2組の特徴点間の角度より直接計測できる。

図4.5.1は環境内に存在する特徴点、それらを結ぶ仮想的な直線とロボットの移動経路の例を示す。最初ロボットは環境内のある地点 $O_0$ にいる。次にロボットは、単眼全方位ステレオにより得られる環境構造の荒い計測結果をもとに、環境内より適当な2点A、Bを選び、この2点が全方位画像中で $180^\circ$ の間隔で投影される位置 $O_1$ に移動する。この2点は、(1)ロボットからの距離が小さい(単眼全方位ステレオの結果を利用する)、(2)2点の間隔があらかじめ $180^\circ$ に近い等の基準により選択する。 $O_1$ に到達したロボットは2眼全方位ステレオを行うために、第二の観測点 $O_2$ に移動する。このとき、全方位画像の視覚フィードバックにより2点間の角度を $180^\circ$ に保つことで、線分AB上を正確に移動する。 $O_2$ に到達したロボットは、 $O_1$ と $O_2$ で撮影された全方位画像をもとに2眼全方位ステレオを行い、局所地図を得



る。O<sub>2</sub>は2点A、Bを180°の間隔で観測すると同時に、2点C、Dも180°の間隔で観測する位置である。次にロボットは移動経路を変え、線分CD上を移動する。O<sub>3</sub>に到達した時点で再びO<sub>2</sub>、O<sub>3</sub>間で局所地図を作成するのであるが、これら2つの局所地図の関係(2眼全方位ステレオの2つの基線間の関係)は、観測点O<sub>2</sub>で撮像した全方位画像上での点A、C間(またはB、D間)の角度を直接計測することで得られる。

一般に環境内には数多くの特徴点が存在するため、本論文の提案する移動方法で、観測点となりうる2つの特徴点を結ぶ直線の交点は、十分密に存在する。すなわち、ロボットは十分自由に環境内を移動することができる。また、全方位画像からの視覚フィードバックを実現するために、全方位画像全体を撮影する必要はない。特徴点が存在する付近の角度のみを撮影すればよく、撮像範囲を限定することで、4.4で述べた撮像方法の欠点である撮像時間の問題を克服できる。

#### 4.5.2 局所地図の獲得

##### 全方位画像間の対応

局所地図獲得における問題は、2つの全方位画像間での対応探索である。全方位画像間の対応を求める方法には、全方位画像の円型構造を利用した循環動的計画法 [Zheng 90] がある。

本論文の場合も、循環動的計画法を用いるができる。しかし本論文の場合、4.5.1で説明したように、観測点間の位置関係を正確に知ることができ、その結果、対応探索の強力な拘束条件であるエピポーララインを求めることができる。

エピポーララインは式(4.5.1)で示される。

$$y = \frac{fH \sin x}{D - R \sin x} \quad (4.5.1)$$

ここで、(x,y)は全方位画像上での物体の投影の座標(xが方位角、yが縦の座

標)で、全方位画像の方位参照軸は2眼全方位ステレオの基線方向にとる。また、 $f$ はカメラの焦点距離、 $R$ はカメラ焦点の回転半径、 $H$ はカメラ光軸から物体までの高さ、 $D$ は2眼全方位ステレオの基線から物体までの距離を表す。

式(4.5.1)は、未知数である $H$ と $D$ を一つの変数として扱うことができないため、全方位画像上の一点の座標 $(x,y)$ だけでは、その点に対するエピポーララインを決定できない。そこで、4.4.4で述べた単眼全方位ステレオで得られる距離情報より $D$ を決定し、エピポーララインを求める。

図4.5.1は、ロボットが全方位画像からのフィードバックを用いながら環境内の2つの特徴点間を移動する間の、幾つかの特徴点の全方位画像上での軌跡(エピポーラライン)を示す。

### 2眼全方位ステレオ

2眼全方位ステレオによって得られる局所地図は、どちらか一方の観測点を原点とし、2つの観測点を結ぶ直線を方位参照軸とする座標系によって表す(図4.5.3)。この座標系で表される局所地図を経路中心の環境表現と呼ぶ。これまでに求めた全方位画像間の対応と観測点間の位置関係より、物体の位置 $(X,Y)$ は式(4.5.2)で計算できる(図4.5.3)。

$$X = \frac{\cos \theta_1 \sin \theta_2}{\sin(\theta_2 - \theta_1)} D, \quad Y = \frac{\sin \theta_1 \sin \theta_2}{\sin(\theta_2 - \theta_1)} D \quad (4.5.2)$$

ここで、 $\theta_1$ と $\theta_2$ は基線方向に対する物体の角度であり、 $D$ は基線距離である。ロボットの移動パラメータは正確に計測できないため、 $D$ は局所地図のスケールファクタとして残す。

### 2眼全方位ステレオの誤差

2眼全方位ステレオの誤差は式(4.5.3)で表される。

$$\Delta X = \frac{Y^2}{D} \left\{ \frac{\cos \theta_1}{\sin \theta_1 \sin^2 \theta_2} \Delta \theta_2 + \frac{\cos \theta_2}{\sin^2 \theta_1 \sin \theta_2} \Delta \theta_1 \right\}$$

$$\Delta Y = \frac{Y^2}{D} \left\{ \frac{1}{\sin^2 \theta_1} \Delta \theta_1 + \frac{1}{\sin^2 \theta_2} \Delta \theta_2 \right\}$$
(4.5.3)

ここで、 $\Delta \theta_1$ 、 $\Delta \theta_2$ は2眼全方位ステレオの基線方向に対する、2つの全方位画像の方位角測定誤差である。基線距離Dの測定誤差は考慮していない。

式(4.5.3)において、 $\Delta \theta_1$ 、 $\Delta \theta_2$ をそれぞれ $0.5^\circ$ 、 $0.5^\circ$ とすると、観測点の回りの位置決め誤差は図4.5.4に示す形になる。図4.5.4に示されるように、経路中心の表現では、基線方向にある物体の位置決め精度はよくない。この問題は、局所地図を統合し全体地図を得る過程において克服する。

ここで行った誤差評価は、基線距離(局所地図のスケールファクタ)の関数で表し、局所地図を構成する各点に関して計算され、全体地図を獲得する際に用いる。

### 4.5.3 全体地図の獲得

一つの局所地図を得た後、ロボットはさらに移動し、新たな局所地図を得る。得られた2つの局所地図は、一つの全体地図(環境中心の環境表現)に統合される。ロボットが移動する毎にこの処理を繰り返し、全体地図をより正確なものにしていくと同時に、ロボットの移動経路を復元する。

全体地図を得る際に気をつけなければならないのは、1つ1つの局所地図の基線方向が同じにならないように、観測計画をたてることである。前章で述べたように、局所地図の作成で用いる2眼全方位ステレオではその基線方向の位置決め精度がよくない。これを克服するには、複数の局所地図の基線方向が異なるように観測を行う必要があり、例えば、ロボットが正三角形の移動軌跡をとるのが理想である。

局所地図間の位置関係は平行移動、回転、スケールファクタの比の3つによって表される。

#### 平行移動

図4.5.1において、ロボットが観測点  $O_1$  と  $O_2$  の間で局所地図を得た後、観測点  $O_2$  と  $O_3$  の間で次の局所地図を得たとすると、共通の観測点  $O_2$  を重ね合わせることで、2つの経路中心の環境表現間の平行移動は一意に決定できる。

### 回転

観測点  $O_1$  と  $O_2$  から得られた局所地図と、観測点  $O_2$  と  $O_3$  から得られた局所地図間の回転成分は、 $O_2$  において、全方位画像上での基線方向を示す特徴点  $A$  と  $C$  の角度を計測して求められる(4.5.1参照)。

### スケールファクタの比

各局所地図は、スケールファクタを未知数として表現していたため、全体地図に統合するには、局所地図のスケールファクタの比を計算しなければならない。スケールファクタの比  $S$  は、式(4.5.4)によって計算する。

$$S = \frac{1}{\sum_{i=1}^n \frac{1}{E_i E_i'}} \sum_{i=1}^n \frac{1}{E_i E_i'} \frac{\sqrt{X_i'^2 + Y_i'^2}}{\sqrt{X_i^2 + Y_i^2}} \quad (4.5.4)$$

ここで、 $X_i$ 、 $Y_i$ 、 $X_i'$ 、 $Y_i'$  は2つの局所地図のスケールファクタを等しくした場合(どちらも1とした場合)の局所地図を構成する各点の座標であり、 $E_i$ 、 $E_i'$  はその点の誤差評価である。

各局所地図間の大きさの比は求めるが、その絶対的な大きさは求めない。もし必要とあらば、ロボットの車輪に取り付けられたエンコーダの情報と、ここで求めたスケールファクタの比より最小自乗法を用いて容易に求めることができる。

局所地図間の関係が求まれば、ロボットの移動経路を復元することができる。ロボットの移動経路が復元された後、全体地図を構成する各点の位置をカルマンフィルタによって再評価する。

#### 4.5.4 実験結果

移動ロボットには、屋内搬送車の Murata100 をコンピュータから制御できるように改良して用いた。移動ロボット上に搭載されるカメラ回転台は、オプティカルエンコーダ信号によりフィードバック制御され、減速ギアの機構と合わせて  $0.005^\circ$  毎に制御できる。画像処理には Nexus6800 を、ホストコンピュータには Sun4 ワークステーションを用いた。

移動ロボットは環境内の複数の観測点間を移動し、各観測点においては静止して全方位画像を撮像する。各全方位画像はカメラ回転台を  $0.4^\circ$  毎に回転させて撮像した。

図 4.5.5 は計算機室内の 7 つの観測点で撮像された全方位画像を示す (大きさは  $900 \times 512$  画素)。これらの観測結果をもとに、各観測点間で合計 6 つの局所地図を作成し、それらを統合して得られた全体地図を図 4.5.7 に示す。

図 4.5.6 で、正方形の点は、得られた全体地図を構成する特徴点の位置を、円形の点は計算されたロボットの位置を、直線は実際の計算機室内の構造を示す。実際の環境では、全体地図を構成する特徴点は、図 4.5.6 中の直線上に位置するものであり、この実験結果においても、計算された特徴点は環境構造を示す直線上に位置していることがわかる。

この実験ではカメラを  $0.4^\circ$  毎に回転させて全方位画像を撮像したが、我々の用いたカメラ回転台は  $0.005^\circ$  の精度で回転させることができ、より正確な計測が可能である。

#### 4.5.5 考察

環境内の複数の観測地点で撮像された全方位画像より局所地図を作成し、それらを統合して環境の全体地図を得る方法について述べた。

特に、局所地図と全体地図の獲得では、移動ロボットの運動パラメータを正確に知るために、全方位画像からの視覚フィードバックを用いる方法を提案した。この方法により、単眼全方位ステレオから得られる荒い環境構造の計測結果を用いて移動計画を立て、2眼全方位ステレオにより正確な

環境構造を知るという戦略が実現できる。

また、全体地図は、ロボット中心、経路中心、環境中心の3つの環境表現を用いて獲得される。それぞれの環境表現について以下にまとめる。

- (1) 単眼全方位ステレオによって獲得されるロボット中心の環境表現は、ロボットの移動可能な自由領域の発見に用いられる。
- (2) 2眼全方位ステレオによって獲得される経路中心の環境表現(局所地図)は、経路上を移動するロボットに十分な情報を与える表現方法であり、長い基線距離のステレオによって、より正確に環境の位置決めを行うことができる。
- (3) 複数の経路中心の表現を融合することによって得られる環境中心の環境表現(全体地図)は、環境の正確な構造と、ロボットの移動経路を再現する最終的な環境表現である。

今後の課題は、面情報の復元を行いより完全な全体地図を作成することである。

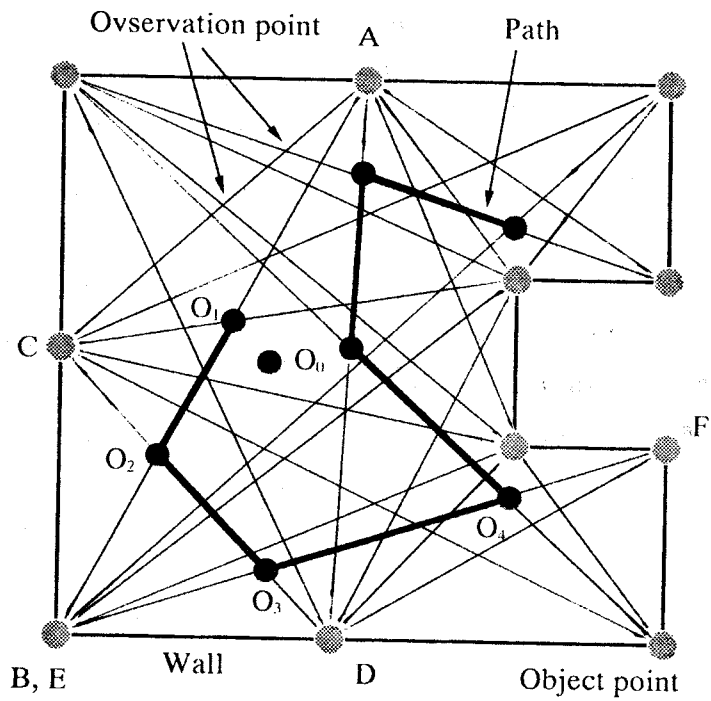
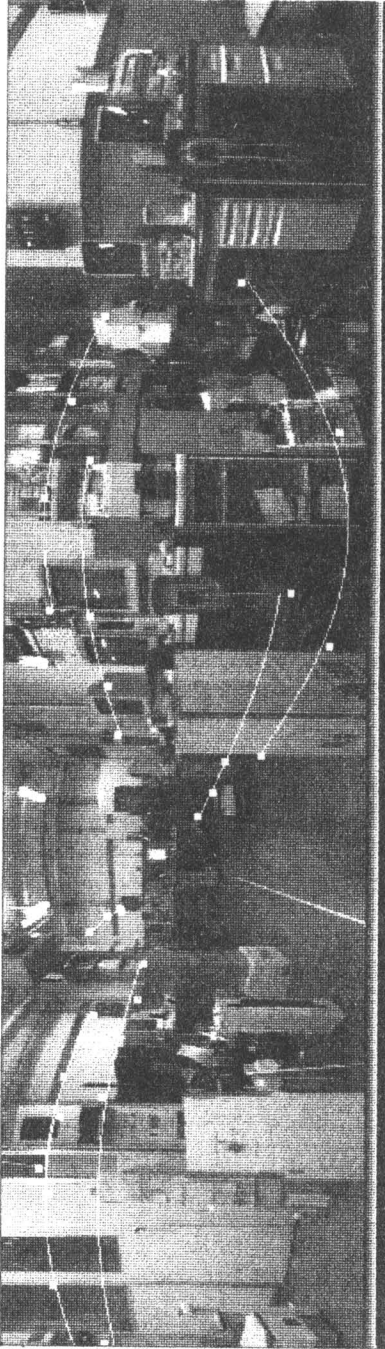


図4.5.1 視覚フィードバックを用いた移動



☒ 4.5.2 エピポーラライン



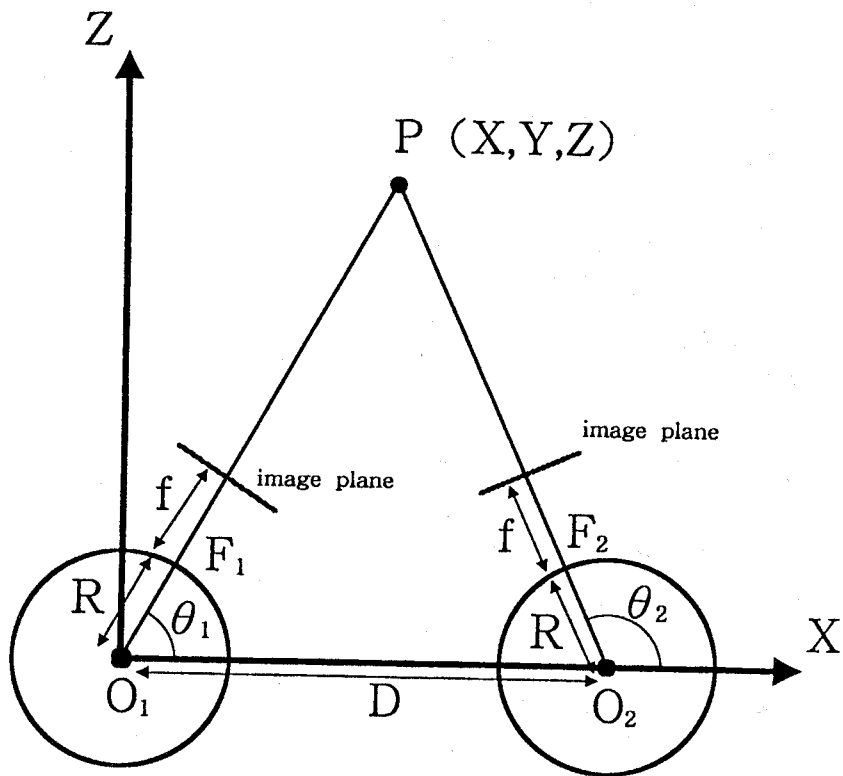


図4.5.3 2眼全方位ステレオ

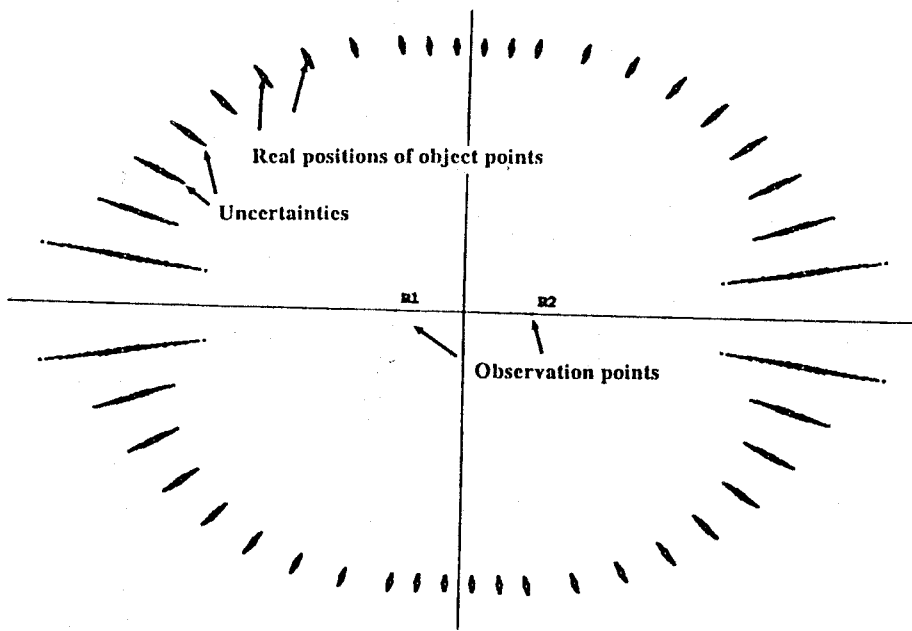


図4.5.4 2眼全方位ステレオの誤差



図 4.5.5 撮影された7枚の全方位画像

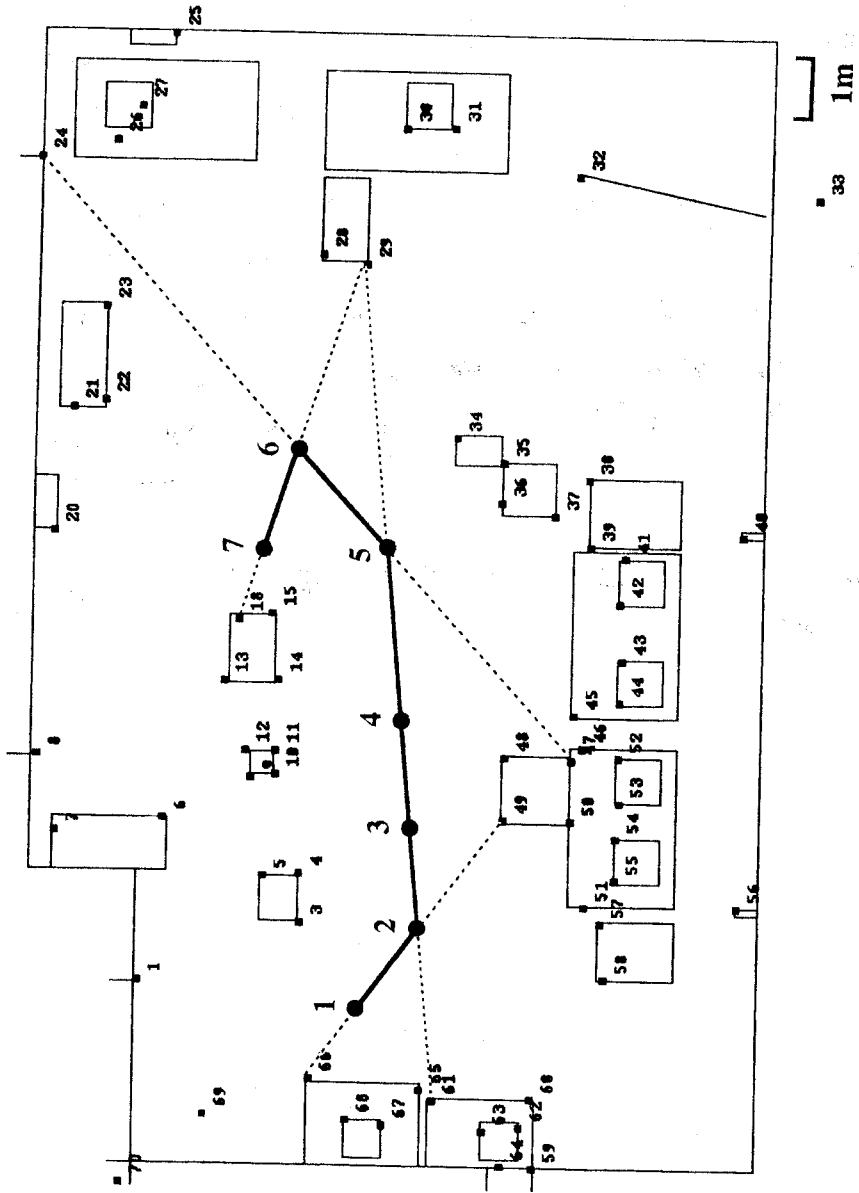


図 4.5.6 全体地図

## 4.6 まとめ

本節では、全環境を見渡すカメラ運動について述べた。全環境を見渡すカメラ運動は、ロボットが未知環境にさしかかった場合など、全環境の正確な構造を知る必要がある場合に有効である。

ここで提案した環境の全体地図を得る方法の特徴を以下にまとめる。

(1) 環境内のある地点で、ロボットはカメラを回転させることにより、全方位の正確な方位角情報(カメラ回転台の角度分解能に等しい方位角分解能を持つ)と、距離情報を持つ、全方位画像を撮像する方法を提案した。

(2) 全方位画像からの視覚フィードバックを用いて環境内を移動する方法を提案した。この方法により、ロボットは環境に対する自らの移動経路の方向を正確に保つことができる。

(3) 環境の全体地図は、観測者中心、経路中心、環境中心の3つの環境表現を用いて獲得される。

## 第5章 結論

本論文では、移動ロボットの能動的視覚について述べた。能動的視覚として、環境の一部に注視するカメラ運動と、全環境を見渡すカメラ運動について考察した。本論文の独創性を以下に記す。

### (1) 環境の一部に注視するカメラ運動

- ロボットの直線運動中に、注視制御により局所地図を得る方程式を導いた。導いた方程式は、連続的観測を前提としている(離散的観測については Sandini ら [Sandini 90] によって報告されている)。
- 注視制御により局所地図を獲得する方法の誤差解析を行った。
- 注視制御により局所地図を得る方程式よりカルマンフィルタを導き、それにより局所地図を得る実験を行った。
- 注視制御を移動ロボットの環境認識に用いることを提案した。

### (2) 全環境を見渡すカメラ運動

- カメラを正確に回転させて、正確な方位角情報を得る方法を提案した。
- 一台のカメラを回転させて、全方位の距離情報を獲得する方法を提案した(単眼全方位ステレオ)。
- 全方位画像からの視覚フィードバックを用いて、環境内を移動する方法を提案した。この方法により、2つの地点で得られた全方位画像を用いた、正確な2眼ステレオ(2眼全方位ステレオ)が実現できる。
- ロボットの移動に伴って得られる複数の局所地図を統合し、正確な全体地図を得る方法を示した。また、全体地図を獲得する過程では、観測者中心、経路中心、環境中心の3つの環境表現を用いる。

能動的視覚の効果は、ロボットの移動とカメラによる観測の直接的なつながりを切り放せることにある。これにより、ロボットの無駄な運動が省かれると同時に、ロボットの正確な運動制御の必要性が小さくなる。

また、移動ロボットの能動的視覚は、環境認識の方法を特徴付ける基本的な機能である。環境の一部に注視するカメラ運動からは、正確な環境の局所的な構造を記録する局所地図と、その大まかな関係より、環境を表現する方法が導かれる。全環境を見渡すカメラ運動からは、全環境の正確な構造を記録する環境表現が導かれる。

ロボットの運動性能や身体性の拡張をはかり、それらに応じた環境表現を用いることは、ロボットの環境認識における重要な課題であり、本論文はその一部を成しえたと確信する。

## 謝辞

本論文は多くの人の助言や協力を得て執筆することができました。末筆ながら、ここに心から感謝の意を表します。

まず最初に、本論文選考の主査であり、指導教官である辻三郎教授に深く感謝いたします。博士過程に入学して以来、教授は常に理解ある御指導と励ましの言葉を与えて下さいました。私が教授から教えて頂いたことは、これから研究者として歩もうとする私にとって、たいへん貴重なものであります。

本論文に関して適切な御指摘、御指導を頂いた、井口征士教授と谷内田正彦教授(基礎工学部情報工学科)に感謝します。また、谷内田教授からは、普段の研究活動において様々な助言や、解説執筆の機会を与えて頂いたことを感謝します。

笠井健教授と浅田稔助教授(工学部電子制御機械工学科)からは、普段の研究において多くの御意見、御指摘を賜りました。笠井教授からは、注視制御に関する問題や、カメラ運動による正確なエッジの方位角検出に関する問題において、生体工学の立場からの有意義な御教示を賜ったことを感謝します。また、浅田助教授には、普段より研究に関する有意義な助言を頂いたことを感謝します。

今井正和助手、徐剛助手、山田誠二助手、早瀬敏夫技官には、研究室での活動において常に励まし、助けて頂いたことを感謝します。今井助手、早瀬技官には計算機システムを中心に多くの便宜をはかって頂いたことを感謝します。徐助手からはコンピュータビジョンに関する有意義な意見を頂いたことを感謝します。山田助手には、人工知能に関する御教示を頂いた



と共に、普段の生活面で多くの助言を頂いたことを感謝します。

Dr. Patrick Stelmaszyk(I.T.M.I(France)勤務)と Dr. Matthew Barth(研究員)とは、注視制御に関する研究をともに行ってきました。両氏と研究を共にできたことを光栄に思います。また、両氏との出会いは私にとって大変貴重なものでありました。

山本雅史君(修士課程2回生)、加藤浩仁君(修士課程1回生)、宮脇一路君(学部4回生)は、私と共に研究を進めてきた同じグループの仲間であります。山本君とは、3年間に渡り全方位画像の研究を行いました。加藤君とはカメラ運動による正確な方位角情報の検出について研究を行いました。宮脇君とは注視制御に関する研究を行いました。3人の辛抱強い努力に感謝します。特に、山本、加藤両君の本論文への貢献に感謝します。

鄭絳宇氏(ATR勤務)、陳謙氏(博士課程3回生)、李仕剛氏(博士課程1回生)らには、共通の研究分野で多くの討論が持てたことを感謝します。

本論文の印刷にあたり協力をしてくれた寺内敦君、磯田佳徳君を初めとする研究室の皆さんに感謝します。

最後に、私を移動ロボットの研究に導き、博士過程進学の機会を与えて頂いた森英雄助教授(山梨大学工学部電子情報工学科)と、9年間に及ぶ大学生活を支えてくれた両親に感謝します。

## 文献

[Aloimonos 87] J. Aloimonos, Isaac Weiss and A. Bandyopadhyay: "Active Vision", Proc. Image Understanding Workshop, pp. 552-573 (1987).

[Arimoto 77] 有本卓: "カルマン・フィルター", 産業図書 (1977).

[Aloimonos 89] J. Aloimonos and D. Shulman: "Integration of Visual Modules, An Extension of the Marr Paradigm", Academic Press, pp. 143-181 (1989).

[Asada 88] M. Asada, Y. Fukui and S. Tsuji: "Representing Global World of a Mobile Robot with Relational Logical Maps", IEEE Int. Workshop Intell. Robots and Systems, pp. 199-206 (1988).

[Asada 89] M. Asada and Y. Shirai: "Building a World Model for a Mobile Robot Using Dynamic Semantic Constraints", 11th Int. Joint Conf. Artificial Intelligence, pp. 1629-1634 (1989).

[Ayache 87] N. Ayache and O. D. Faugers: "Building, Registrtrng and Fusing Noisy Visual Maps", Proc. 1st Int. Conf. Computer Vision, pp. 73-82 (1987).

[Ballard 82] D. H. Ballard and C. M. Brown: "Computer Vision", Prentice-Hall, Inc. (1987). 邦訳「コンピュータ・ビジョン」(福村晃夫他訳, 日本コンピュータ協会)

[Ballard 88] D. H. Ballard and A. Ozcandari: "Eye Fixation and Early Vision : Kinetic Depth", Proc. 2nd IEEE Int'l Conf. Computer Vision, pp. 524-531 (1988).

[Ballard 89] D. H. Ballard: "Reference Frames for Animate Vision", Proc. 11th Int'l Joint Conf. Artificial Intelligence, pp. 1635-1641 (1989).

[Blais 88] F. Blais, M. Rioux, J. Domey and J. Beraldin: "A Very Compact Real Time 3-D Range Sensor for Mobile Robot Applications", Proc. SPIE, Vol. 1007,

pp. 330-338 (1988).

[Blancard 90] M. de Saint Blancard: "Road Sign Recognition: A Study of Vision-Based Decision Making for Road Environment Recognition", Proc. Roundtable Discussion Vision-Based Vehicle Guidance, pp. 6.1-6.9 (1990).

[Bolles 87] R. C. Bolles et al.: "Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion", Int. Journal Computer Vision, 1, No.1, pp. 7-56 (1987).

[Brooks 86] R. A. Brooks and J. H. Connell: "Asynchronous Distributed Control System for a Mobile Robot", Proc. SPIE, Cambridge, MA, pp.77-84 (1986).

[Bruss 83] A. R. Bruss and B. K. P. Horn: "Passive Navigation", Computer Vision Graphics and Image Processing, 21, pp. 3-20 (1983).

[Cutting 86] J. E. Cutting: "Perception with an Eye for Motion", The MIT Press (1986).

[Dickmanns 88a] E.D. Dickmanns and V. Graefe: "Dynamic Monocular Machine Vision", Machine Vision and Application, 1, pp. 223-240 (1988).

[Dickmanns 88b] E.D. Dickmanns and V. Graefe: "Applications of Dynamic Monocular Machine Vision", Machine Vision and Application, 1, pp.241-261 (1988).

[Ejiri 90] 江尻正員, 太田友一, 池内克史: "マシンビジョン", 昭晃堂, pp. 70-75 (1990).

[Elfes 87] A. Elfes: "Sonar-Based Real-World Mapping and Navigation", IEEE Journal Robotics & Automation, RA-3, 3, pp. 249-265 (1987).

[Faugeras 90] O. D. Faugeras, E. Le Bras-Mehlman and J. D. Boissonnat: "Representing Stereo Data with the Delaunay Triangulation", Artificial Intelligence, 44, pp. 41-87 (1990).

[Graefe 90] V. Graefe: "An Approach to Obstacle Recognition for Autonomous

Mobile Robots”, Proc. IEEE Int. Workshop Intell. Robots and Systems, pp. 151-158 (1990).

[Horn 86] B. K. P. Horn: “Robot Vision”, The MIT Press, pp. 401-403 (1986).

[Horn 88] B. K. P. Horn and E. J. Weldon: “Direct Methods for Recoreing Motion”, Int. Journal Computer Vision, Vol. 2, pp. 51-76 (1988).

[Ikeuchi 88] K. Ikeuchi and T. Kanade: “Towards automatic generation of object recognition program”, Proc. IEEE, Vol 76, No. 8, pp. 1016-1035 (1988).

[Ishiguro 90a] H. Ishiguro, P. Stelmazyk and S. Tsuji: “Acquiring 3-D Structure by Controlling Visual Attention of a Mobile Robot”, Proc. IEEE Int. Conf. Robotics and Automation, pp. 755-760 (1990).

[Ishiguro 90b] H. Ishiguro, M. Yamamoto and S. Tsuji: “Analysis of Omni-Directional Views at Different Locations”, Proc. IEEE Int. Workshop Intell. Robots and Systems, pp. 659-664 (1990).

[Ishiguro 90c] H. Ishiguro, M. Yamamoto and S. Tsuji: “Omni-Directional Stereo for Making Global Map”, to appear in Proc. IEEE Int. Conf. Computer Vision (1990).

[Kambhampati 86] S. Kambhampati and L. S. Davis: “Multiresolution Path Planning for Mobile Robots”, IEEE Journal Robotics & Automation, Vol. RA-2, No. 3, pp. 135-145 (1986).

[Kakikura 89] 柿倉正義: “知能ロボット入門”, オーム社 (1989).

[Kitamura 87] 北村, 来海, 谷内田: “エッジの連続性と画面の局所的性質を利用した3眼視”, 電子情報通信学会論文誌, Vol. J72-D, No. 9, pp. 1757-1764 (1987).

[Kuniyoshi 90] Y. Kuniyoshi, H. Inoue and M. Inaba: “Design and Implementation of a System that Generates Assembly Programs from Visual Recognition of Human Action Sequences”, Proc. IEEE Int. Workshop Intell. Robots and Systems, pp. 567-574 (1990).

[komoriya 84] 小森谷清, 館章, 谷江和雄: “移動ロボットの自律誘導の—方法”, 日本ロボット学会誌, Vol. 2, No. 3, pp. 46-55 (1984).

[Kweon 90] I. S. Kweon and T. Kanade: “High Resolution Terrain Map from Multiple Sensor Data”, IEEE Int. Workshop Intell. Robots and Systems, pp. 127-134 (1990).

[Levitt 87] T. S. Levitt, D. T. Lawton, D. M. Chelberg, P. C. Nelson: “Qualitative Navigation”, Proc. DARPA Image Understanding Workshop, pp. 447-465 (1987).

[Lougheed 88] R. M. Lougheed and R. E. Sampson: “3-D Imaging Systems and High-Speed Processing for Robot Control”. Machine Vision and Applications, Vol. 1, No.1, pp. 727-733 (1988).

[Matthies 88] L. Matthies and A. Elfes: “Integration of Sonar and Stereo Range Data Using a Grid-Based Representation”, Proc. IEEE Int. Conf. Robotics and Automation, pp. 727-733 (1988).

[Matthies 89] L. Matthies and T. Kanade: “Kalman Filter-Based Algorithms for Estimating Depth from Image Sequences”, Int. Journal Computer Vision, 3, pp. 209-236 (1989).

[Matumoto 86] 松本他: “自律ロボット「山彦」9号の屋内環境走行方式”, 第3回知能移動ロボットシンポジウム予稿集 (1986).

[Marr 82] D. Marr: “Vision”, W. H. Freeman and Company (1982). 邦訳「ビジョン」(乾, 安藤訳, 産業図書)

[Mori 87] 森英雄他: “構内道路を走行する自律移動ロボット(移動ロボットの—戦略-晴信4号機)”, 日本ロボット学会誌, Vol. 5, No. 5, pp. 29-41 (1987).

[Morita 89] T. Morita, Y. Yasukawa, Y. Inamoto, T. Uchiyama and S. Kawakami: “Measurement in Three Dimensions by Motion Stereo and Spherical Mapping”, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 422-428 (1989).

[Nelson 87] R. C. Nelson and J. Aloimonos: "Finding Motion Parameters from Spherical Flow Fields", Proc. Workshop on Computer Vision, pp. 145-150 (1987).

[Nilsson 84] N. J. Nilsson: "Shakey the Robot", SRI AI center Technical Note 323 (1984).

[Olson 88] T. J. Olson and R. D. Potter: "Real-Time Vergence Control", Technical Report 264, Computer Science Dept. Univ. Rochester (1988).

[Onoguchi 88] 小野口一則他: "移動ロボットの位置認識システム", 情報処理学会研究報告, CV 57-3 (1988).

[Pomerleau 89] D. A. Pomerleau: "ALVINN: An Autonomous Land Vehicle in a Neural Network", In Advances in Neural Information Processing Systems, Vol. 1, Morgan Kaufmann (1989).

[Raviv 90] D. Raviv and M. Herman: "Towards an Understanding of Camera Fixation", Proc. IEEE Int. Conf. Robotics & Automation, pp. 28-33 (1990).

[Sandini 90] G. Sandini and M. Tistarelli: "Active Tracking Strategy for Monocular Depth Inference Over Multiple Frames", IEEE Trans. Pattern Anal. and Machine Intell., PAMI-12, 1, pp. 13-27 (1990).

[Sarachik 89] K. B. Sarachik: "Characterising an Indoor Environment with a Mobile Robot and Uncalibrated Stereo", Proc. IEEE Int. Conf. Robotics and Automation, pp. 984-989 (1989)

[Simons 84] G. L. Simons: "Introducing Artificial Intelligence", The National Computing Centre Limited (1984) 邦訳「人工知能その限界と挑戦」(田村, 佐藤訳 近代科学社)

[Stelmaszyk 88] P. Stelmaszyk, C. Disscours and C. Chehikian: "A Fast and Reliable Token Tracker", Proc. IAPR Workshop on Computer Vision (1988).

[Storjohann 90] K. Storjohann, Th. Zielke, H. A. Mallot and W. von Seelen: "Visual Obstacle Detection for Automatically Guided Vehicles (1990).

[Suzuki 89] 鈴木, 谷内田: “移動視における対応付けと3次元情報の獲得”, 電子情報通信学会誌, Vol. J72-D, No. 5, pp. 686-695 (1989).

[Swain 90] M. J. Swain and D. H. Ballard: “Indexing Via Color Histograms”, Proc. 1990 Int. Conf. Computer Vision, pp. 390-393 (1990).

[Talluri 90] R. Talluri and J. K. Aggarwal: “Position Estimation for a Mobile Robot in an Unstructured Environment”, IEEE Int. Workshop Intell. Robots and Systems, pp. 159-166 (1990).

[Tistarelli 90] M. Tistarelli and G. Sandini: “Robot Navigation Using an Anthropomorphic Visual Sensor”, Proc. IEEE Int. Conf. Robotics and Automation, pp. 374-383 (1990).

[Thorpe 89] C. Thorpe and T. Kanade: “1988 Year End Report for Road Following at Carnegie Mellon”, Carnegie Mellon Univ. the Robotics Institute Technical Report, 89-5 (1989).

[Yeap 88] W. K. Yeap: “Towards a Computational Theory of Cognitive Maps”, Artificial Intelligence, 34, pp. 297-360 (1988).

[Yachida 83] M. Yachida, T. Ichinose and S. Tsuji: “Model Guided Monitoring of Building Environment by a Mobile Robot”, Proc. 8th Int. Joint Conf. Artificial Intell., pp. 1125-1127 (1983).

[Yachida 90a] 谷内田正彦: “環境理解の展望”, コンピュータビジョンシンポジウム (1990).

[Yachida 90b] 谷内田正彦編: “コンピュータビジョン”, 丸善株式会社 (1990).

[Yachida 90c] 谷内田正彦, 石黒浩: “ロボットの視覚”, 人工知能学会誌, Vol. 5, No. 6, pp. 720-730 (1990).

[Yagi 90] Y. Yagi and S. Kawato: “Panorama Scene Analysis with Conic Projection”, IEEE Int. Workshop Intell. Robots and Systems, pp. 181-190 (1990).

[Yamada 89] 山田誠二：“問題解決における戦略知識の学習”，大阪大学大学院基礎工学研究科博士論文 (1989)

[Zheng 90a] J. Y. Zheng and S. Tsuji: “Panoramic Representation of Scenes for Route Understanding”, Proc. 10th Int. Conf. Pattern Recognition, pp. 161-167 (1990).

[Zheng 90b] J. Y. Zheng and S. Tsuji: “From Anorhoscope Perception to Dynamic Vision”, Proc. IEEE Int. Conf. Robotics & Automation (1990).

[Zheng 90c] J. Y. Zheng and M. Barth: “Qualitative Route Description Using Autonomous Landmark Detection”, Proc. IEEE Int. Conf. Computer Vision, pp. 558-562 (1990).





## 研究業績

## 論文等

H. Ishiguro, P. Stelmaszyk and S. Tsuji: "Acquiring 3-D Structure by Controlling Visual Attention of a Mobile Robot", Proceedings of 1990 IEEE International Conference on Robotics and Automation, pp. 755-760 (1990).

H. Ishiguro, M. Yamamoto and S. Tsuji: "Omni-Directional Stereo for Making Global Map", Proceedings of Third International Conference on Computer Vision, pp. 540-547 (1990).

石黒浩, P. Stelmaszyk, 辻三郎: "注視制御による局所地図の獲得", 電子情報通信学会論文誌(1990). (採録決定)

石黒浩, 山本雅史, 辻三郎: "全方位視野の距離情報獲得", 電子情報通信学会論文誌(1990). (採録決定)

H. Ishiguro, M. Yamamoto and S. Tsuji: "Omni-Directional Stereo", IEEE Transactions on Pattern Analysis and Machine Intelligence (1990). (採録決定)

H. Ishiguro, M. Yamamoto and S. Tsuji: "Acquiring Precise Range Information from Camera Motion", Proceedings of 1991 IEEE International Conference on Robotics and Automation (1990). (採録決定)

M. Barth, H. Ishiguro and S. Tsuji: "Determining Egomotion by Tracking the Focus of Expansion with an Active Camera", Proceedings of 1991 IEEE International Conference on Robotics and Automation (1990). (採録決定)

石黒浩, 山本雅史, 辻三郎: "能動的全方位画像を用いた環境構造の復元", 日本ロボット学会学会誌 (1990). (投稿中)

S. Tsuji, H. Ishiguro and K. Kato: "Locating Edges in Images on a Moving Retina", Proceedings of Twelfth International Joint Conference on Artificial Intelligence (1991). (投稿中)

### 口頭発表等

P. Stelmaszyk, H. Ishiguro, P. Roland and S. Tsuji: "Local Processing as a Cue for Decreasing 3-D Structure Computation", Proceedings of Roundtable Discussion on Vision-Based Vehicle Guidance '90, pp. 3.1-3.10 (1990).

H. Ishiguro, M. Yamamoto and S. Tsuji: "Analysis of Onmi-Directional Views at Different Locations", Proceedings of IEEE International Workshop on Intelligent Robots and Systems '90, pp. 659-664 (1990).

## 著者略歴

氏名 石黒浩(いしぐろひろし)

生年月日 昭和38年10月23日

本籍 〒520-12 滋賀県高島郡安曇川町大字下古賀1119番地

現住所 〒563 大阪府池田市空港1丁目9番7号クレベ空港405号室  
☎06-845-1862

## 在学先住所

〒560 大阪府豊中市待兼山1丁目1番  
大阪大学基礎工学部制御工学科辻研究室  
☎06-844-1151内(4619)

学歴 昭和57年3月 滋賀県立高島高等学校卒業  
昭和57年4月 山梨大学工学部計算機科学科入学  
昭和61年3月 山梨大学工学部計算機科学科卒業  
昭和61年4月 山梨大学大学院工学研究科修士課程入学  
昭和63年3月 山梨大学大学院工学研究科修士課程修了  
昭和63年4月 大阪大学大学院基礎工学研究科博士課程入学

所属学会 情報処理学会

電子情報通信学会

日本ロボット学会

人工知能学会

IEEE

## 付録

### A. 一般的なロボットの運動における注視制御

3.4では、ロボットの直視運動における注視制御について、局所地図を得る方程式を導いたが、ここでは、ロボットの円弧運動を含めた一般的な運動に対する方程式を導く。

ロボットの運動と注視点の位置関係を図A.1に示す。一般にロボットは円弧運動を行い、注視点を注視しながら移動する。その特別な場合として、ロボットの円弧運動の中心に注視点がある場合と、ロボットが直進運動を行う場合が考えられる。

各パラメータの説明を以下に記す。環境内の点の位置は $L$ と $\theta$ を用いて表される。

N: ロボットの円弧運動の半径

M: ロボットの円弧運動の中心から注視点 (fixation point) までの距離

$\alpha$ : ロボットの移動方向とカメラの光軸に垂直な直線のなす角

$U_R$ : ロボットの移動速度

$Z_o$ : ロボットから環境内の一点 (object point) までの距離

$Z_f$ : ロボットから注視点までの距離

$x$ : 環境内の一点の画像面への投影 ( $x$  軸方向)

$\theta$ : 環境内の一点と注視点を結ぶ直線とカメラの光軸のなす角

$L$ : 注視点から環境内の一点までの距離

ロボットの並進速度成分  $U, V, W$  と回転速度成分  $A, B, C$  は、先に述べた仮定より、式3.1のように表せる。ここで、 $\omega_c$  はカメラ回転台の回転角速度、

$\omega_R$  はロボットの運動による回転角速度、 $\omega$  はそれらの和である。

$$\begin{cases} U = U_R \cos \alpha \\ V = 0 \\ W = U_R \sin \alpha \end{cases} \begin{cases} A = 0 \\ B = \omega = \omega_c + \omega_R \\ C = 0 \end{cases} \quad (\text{A.1})$$

環境内の一点の画面上への投影を  $x$  とすると、そのオプティカルフロー  $u$  は式 (A.2) で表される [Bruss 83]。

$$u = \frac{-f U_R \cos \alpha + x U_R \sin \alpha}{Z_o} - \omega \left( f + \frac{x^2}{f} \right) \quad (\text{A.2})$$

ここで、 $\alpha$  はロボットの移動に伴い変化することに注意されたい。

注視点のオプティカルフロー  $u_f$  は、視覚フィードバックの機能により、画像中心に固定されるため  $u_f = 0$  となり、これより式 (A.3) を得る。

$$u_f = 0 = -\frac{f U_R \cos \alpha}{Z_f} - f \omega \quad (\text{A.3})$$

ここで、 $Z_f$  は注視点までの距離である。式 (A.3) より、カメラ回転台の回転角速度とロボットの回転角速度の和  $\omega$  は

$$\omega = \omega_c + \omega_R = -\frac{U_R \cos \alpha}{Z_f} \quad (\text{A.4})$$

である。ロボットの回転角速度は、その回転半径と移動速度より式 (A.5) で表される。

$$\omega_R = \frac{U_R}{N} \quad (\text{A.5})$$

また、 $\theta$  と  $Z_o, Z_f$  ( $Z_o$  は環境内の一点までの距離) の関係より式 (A.6) を得る。

$$\tan \theta = \frac{x}{f} \left( \frac{Z_f}{Z_o} - 1 \right)^{-1} \quad (\text{A.6})$$

ここで、 $Z_i/Z_o$ は式(A.2)と式(A.4)を用いて式(A.7)のように表され、

$$\frac{Z_o}{Z_i} = K = \frac{f - x \tan \alpha}{\frac{u}{\omega} + f + \frac{x^2}{f}} \quad (\text{A.7})$$

式(A.6)と式(A.7)より、 $\theta$ は式(A.8)で表すことができる。

$$\theta = \tan^{-1} \left( \frac{xK}{f(1-K)} \right) \quad (\text{A.8})$$

$L$ については、注視点までの距離を計算することにより式(A.9)で表わされる。

$$L = (N \cos \alpha \pm \sqrt{M^2 - N^2 \sin^2 \alpha}) \sqrt{(1-K)^2 + \frac{K^2 x^2}{f^2}} \quad (\text{A.9})$$

以上の結果より、環境内の一点の位置は注視点を中心とした極座標( $L, \theta$ )で表わされ、オブジェクトセンターの表現を直接得ることができる。しかし、ここで述べた一般的な場合には、環境内の一点の位置は注視点までの距離に比例しない。そこで、次には特別な場合に付いて考える。

ロボットが円弧運動を行う場合（円弧の中心に注視点がある）

この場合は Aloimonos [Aloimonos 87] らの計算と同じである。 $\theta$ を求める方程式は式(A.8)、(A.9)より、

$$K = \frac{Z_o}{Z_i} = \frac{f}{\frac{u}{\omega} + f + \frac{x^2}{f}}$$

$$\theta = \tan^{-1} \left( \frac{xK}{f(1-K)} \right) \quad (\text{A.10})$$



$$L = N \sqrt{(1-K)^2 + \frac{K^2 x^2}{f^2}} \quad (\text{A.11})$$

ここで、 $\omega = \omega_R$ であり、カメラはロボットの円弧運動の中心を向いて固定される。また、 $N$ はロボットの円弧運動の中心(注視点)までの距離である。

ロボットが直線運動を行う場合

この場合も、先に述べた一般的な場合の特別な場合であるが、移動ロボットにおいては最も重要な運動である。この場合に、カメラ回転台の回転角速度は有効に利用することができる。環境内の一点の位置( $L, \theta$ )は式(A.12)、(A.13)で表わすことができる。

$$K = \frac{Z_o}{Z_r} = \frac{f - x \tan \varphi}{\frac{u}{\omega} + f + \frac{x^2}{f}}$$

$$\theta = \tan^{-1} \left( \frac{xK}{f(1-K)} \right) \quad (\text{A.12})$$

$$L = \frac{D}{\cos \varphi} \sqrt{(1-K)^2 + \frac{K^2 x^2}{f^2}} \quad (\text{A.13})$$

ここで、 $D$ はロボットの移動経路と注視点の距離であり、 $L$ はこれに比例して決まる。また、直交座標( $X, Y$ )には式(A.14)を用いて変換することができる。

$$\begin{cases} X = L \cos(\theta - \varphi - \pi/2) \\ Y = L \sin(\theta - \varphi - \pi/2) \end{cases} \quad (\text{A.14})$$

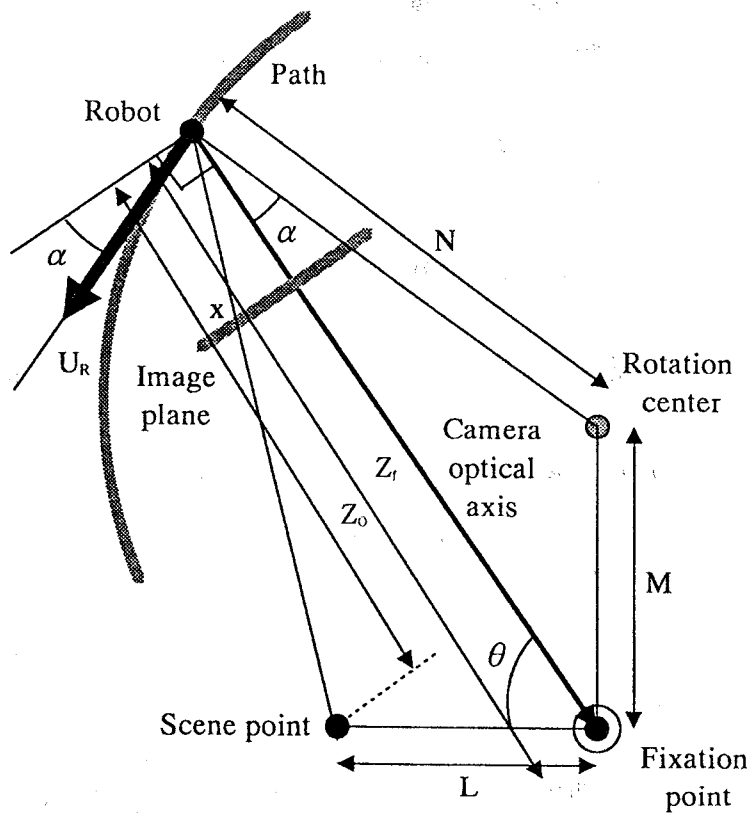


図 A.1 一般的なロボットの運動における注視制御

## B. ロボットの運動とカメラの運動の関係

ロボットが直線運動を行う場合には、ロボットの回転速度成分は、注視点を追跡するカメラ回転台の角速度のみによって決まり、カメラ回転台の角速度情報を用いることができる。すなわち、カメラの回転角速度を回転台の機能を用いて直接計測し、ロボットの速度は用いない。ここでは、その関係を明らかにするためにロボットが直進運動を行う場合の、カメラの回転角速度とロボットの移動速度の関係について述べる。

図B.1に示す幾何学的関係より、式(B.1)で示される関係が得られる。

$$U \Delta t = D \tan \varphi - D \tan \left( \varphi - \frac{d\varphi}{dt} \Delta t \right) \quad (\text{B.1})$$

これを微分することにより、 $\varphi$ の角速度を得る。

$$\frac{d\varphi}{dt} = \frac{U}{D} \cos^2 \left( \varphi - \frac{d\varphi}{dt} \Delta t \right) \quad (\text{B.2})$$

さらに、図B.1より以下の関係が得られる。

$$\left( \alpha + \frac{d\alpha}{dt} \Delta t \right) + \left( \varphi - \frac{d\varphi}{dt} \Delta t \right) = \pi \quad (\text{B.3})$$

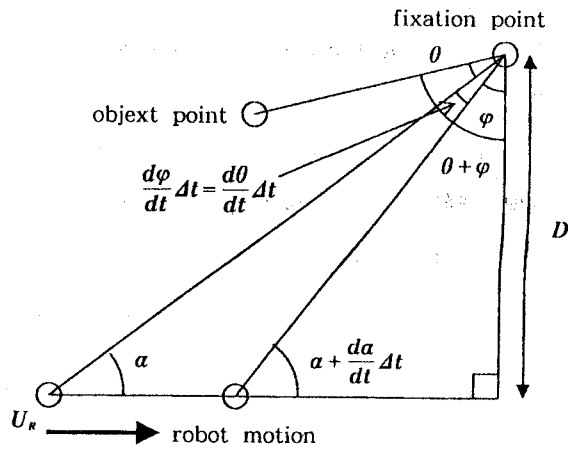
$$\omega = \frac{d\alpha}{dt} = \frac{d\varphi}{dt} = \frac{d\theta}{dt} \quad (\text{B.4})$$

最終的に、カメラの回転角速度 $\omega$ と、ロボットの移動速度 $U_R$ の間には以下の式で表わされる関係が成り立つ。

$$\omega = \frac{U_R}{D} \cos^2 (\varphi - \omega \Delta t) \quad (\text{B.5})$$

すなわち、ロボットが直進運動を行う場合には、カメラの回転角速度は

ロボットの速度情報を含むことになる。



図B.1 ロボットの速度とカメラの回転角速度の関係

### C. 注視制御におけるカメラの回転角速度

ロボットが直進運動を行いながら、カメラを回転させて環境内の一点に注視した時の、カメラの回転角速度  $A$  は、注視点に最も近いロボットの移動経路上の一点から、ロボットのまでの距離  $L$  を用いて式 (C.1) で表わすことができる。

$$A = \frac{K}{1+K^2L^2} \quad (\text{C.1})$$

ここで、 $K$  はロボットの速度  $U$  と、ロボットの移動経路と注視点間の最短距離  $D$  の比  $U/D$  である。

式 (C.1) をグラフで表わしたのが図 C.1 である。図 C.1 では、縦軸がカメラの回転角速度、横軸が経路上での移動距離を表わす。カメラの回転角速度が最も大きい地点が、注視点に最も近い位置である。

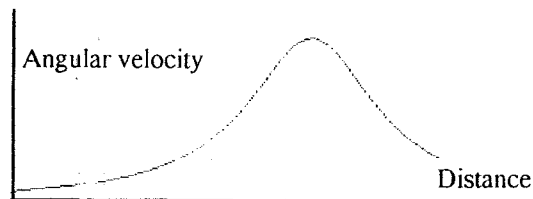


図 C.1 カメラの回転角速度

## D. 注視制御におけるオプティカルフロー

環境内の点の位置決めを行う際に、その点の画面速度の情報を用いたが、この画面速度の符号は興味深い性質を持つ。

図D.1(a)に示すように、ロボットが直線運動を行う場合、カメラの運動に伴う物体のオプティカルフローは、ロボットの移動経路に接し、注視点とロボットの位置を通る円上で0となり、その円の内側ではロボットの運動と逆の方向の、その外側ではロボットの運動と同じ方向の画面速度が観測される。同様に、ロボットが円弧運動を行う場合にも、図D.1(b)に示す円上で物体の画面速度は0となる。ただし、ロボットが直線運動を行う場合には、ロボットの位置によってその円の大きさは変わるが、ロボットが円弧運動を行う場合には常に一定の大きさである(円弧運動における注視制御については付録Aを参照)。これらの円は式(D.1)によって表すことができる。

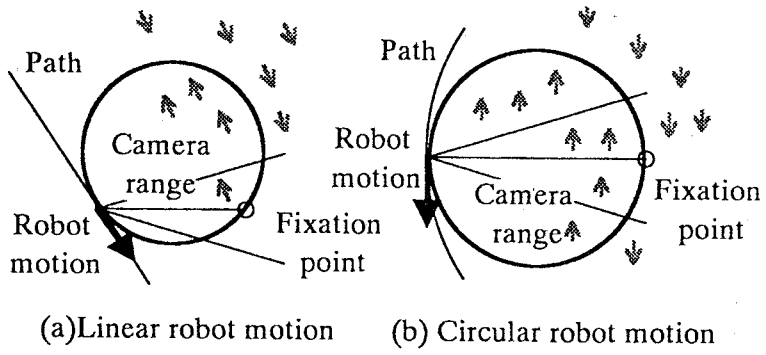
[ロボットが直線運動を行う場合]

$$D\cos(\theta - \varphi) - L\cos^2\theta = 0 \quad (D.1)$$

[ロボットが円弧運動を行う場合]

$$N\cos\theta - L = 0$$

注視制御におけるオプティカルフローに関する一般的な考察は文献[Raviv 90]にある。

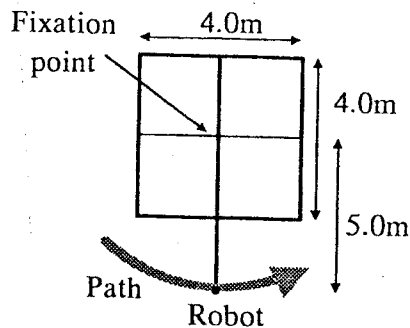


図D.1 注視制御におけるFOE

## E. 円弧運動における注視制御の誤差解析

ロボットが注視点を中心とした円弧運動を行う場合の注視制御について誤差解析を行う。誤差解析は、連続的観測をもとにした付録Aの式(A.10)、式(A.11)を用いて、3.4.2の誤差評価法を用いて行った。図E.2(b)は図E.1に示す正方形の領域での位置決め誤差を、高さとして表わしたものである。比較のために図E.2(a)に直進運動における誤差(図3.4.4と同じである)を示しておく。誤差評価を行うロボットの状態は3.4.2と同じである。ただし円弧運動の場合には、ロボットの移動方向を示すパラメータ $\phi$ が含まれない。

円弧運動においても、直進運動と同様に、注視制御による位置決め誤差は、注視点付近で最も小さくなるのがわかる。



図E.1 誤差解析を行う状態



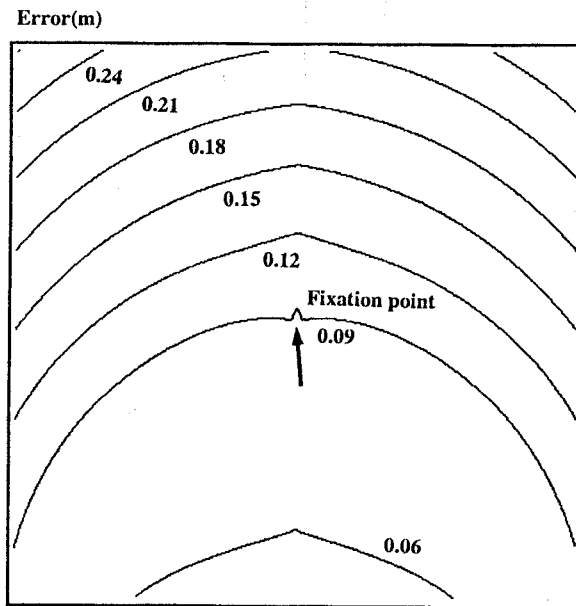
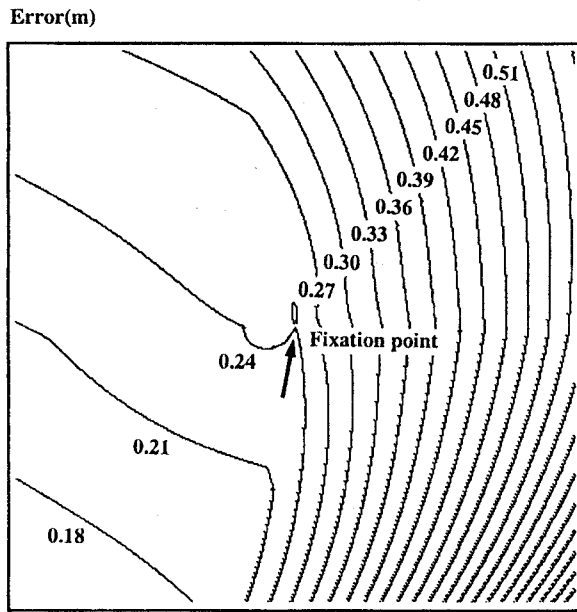


図 E.2 直進と円弧運動における位置決め誤差

## F. カルマンフィルタの設計

一般にオプティカルフローを用いた計測は誤差を伴いやすく、何らかの方法で複数の観測結果を融合し、位置決め精度をあげる必要がある。その方法としてカルマンフィルタが多く用いられるようになってきたが、我々の注視制御を用いる場合においても、カルマンフィルタを用いて、ロボットの複数の位置から計算した結果を融合する。

位置決め誤差に大きく影響を与えるのはカメラの回転角速度 $\omega$ とオプティカルフロー $u$ であり、ここでは2つの値の誤差のみを考えカルマンフィルタを導くが、他のパラメータの誤差を考えた場合も容易に拡張できる。

環境内の一点の位置 $(X, Y)$ は $\omega$ と $u$ を用いて式(F.1)で表わされる。

$$\begin{aligned} X &= f_1(\omega, u) \\ Y &= f_2(\omega, u) \end{aligned} \quad (\text{F.1})$$

また、 $X$ 、 $Y$ の誤差を表わす共分散行列はJacobianを用いて式(F.2)で表わされる。

$$\Lambda = \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y \\ \sigma_y \sigma_x & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial \omega} & \frac{\partial f_1}{\partial u} \\ \frac{\partial f_2}{\partial \omega} & \frac{\partial f_2}{\partial u} \end{pmatrix} \begin{pmatrix} \sigma_\omega^2 & \sigma_\omega \sigma_u \\ \sigma_u \sigma_\omega & \sigma_u^2 \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial \omega} & \frac{\partial f_1}{\partial u} \\ \frac{\partial f_2}{\partial \omega} & \frac{\partial f_2}{\partial u} \end{pmatrix} \quad (\text{F.2})$$

ここで、 $\sigma_\omega$ と $\sigma_u$ は $\omega$ と $u$ の考えられる測定誤差の最大値である。

測定結果を融合する場合、予測される誤差の範囲から、融合できる場合と、融合できない場合がある。2つの測定結果の予測される誤差の範囲が重なり合っている場合は融合可能であるが、重なり合わない場合は融合することができない。この判定は式(F.3)に示すMahalanobis距離を用いて行う。

$$(\mu_1 - \mu_2)^T (\Lambda_1 + \Lambda_2)^{-1} (\mu_1 - \mu_2) \quad (\text{F.3})$$

$$\mu_1 = (X_1, Y_1)^T, \quad \Lambda_1 = \begin{vmatrix} \sigma_x^2 & \sigma_x \sigma_y \\ \sigma_y \sigma_x & \sigma_y^2 \end{vmatrix}$$

$$\mu_2 = (X_2, Y_2)^T, \quad \Lambda_2 = \begin{vmatrix} \sigma_x^2 & \sigma_x \sigma_y \\ \sigma_y \sigma_x & \sigma_y^2 \end{vmatrix}$$

この Mahanobolis 距離が 1 より小さいときのみ融合可能である。実際の実験において融合できない場合は、新たな候補として一旦記録し、数回の観測の後、さらに融合されなければ間違いと判定する。

Mahanobolis 距離の判定により融合可能な場合は、カルマンフィルタを用いて融合を行い、計測結果を更新する。カルマンフィルタは式 (F.4) で表わされる。

$$\begin{aligned} K &= \Lambda_p (\Lambda_p + \Lambda_N) \\ \Lambda_{\text{UPD}} &= \Lambda_p - K \Lambda_N^T \\ \mu_{\text{UPD}} &= \mu_p + K (\mu_p - \mu_N) \end{aligned} \tag{F.4}$$

ここで、 $K$  はカルマンゲイン、 $\Lambda$  は共分散行列、 $\mu$  は点の位置を表わす。

## G. 全方位画像の解析

### G.1 全方位視野像への投影

図G.1に示すように、カメラの光軸と回転軸が直交すると、3次元空間から全方位視野像への投影は円柱写影となる。全方位視野像上の座標は、Z軸(基準のカメラの光軸方向)とY軸(回転軸)の交点を原点として、縦方向をy、横方向を $\theta$ とする。また、図中のU,W,Bは、それぞれX,Z軸方向のロボットの並進速度成分、Y軸まわりの回転速度成分を示す。

カメラの回転面が床面と平行である場合、3次元空間内の点 $P=(X,Y,Z)$ の全方位視野像への投影は式(G.1)によって示される。

$$Z \geq 0: \theta = \tan^{-1} \frac{X}{Z}, \quad Z < 0: \theta = \tan^{-1} \frac{X}{Z} + \pi \quad (G.1)$$

$$y = f \frac{Y}{\sqrt{X^2 + Z^2}} \quad \left( = f \frac{Y}{Z} \cos \theta \right)$$

この場合、式(G.2)で表わされる床面に平行な直線は、全方位視野画像内では式(G.3)に示されるように、正弦曲線として投影される。

$$Z = aX + b, \quad Y = c \quad (G.2)$$

$$y = \frac{c}{b} f \sqrt{a^2 + 1} \cos \left( \theta + \tan^{-1} \left( -\frac{1}{a} \right) \right) \quad (G.3)$$

カメラの光軸が床面と $\alpha$ の角度をなす場合には、3次元空間内の点 $P=(X,Y,Z)$ の全方位視野像への投影は式(G.4)によって示される。

$$\theta = \tan^{-1} \frac{X}{Z} \quad (G.4)$$

$$y = f \frac{Y}{\sin \alpha (Y \cos \alpha + \sqrt{X^2 + Z^2} \sin \alpha)}$$

この場合、3次元空間は全方位視野像に円錐写影され、床面に平行な直線は、正弦曲線ではなく円錐曲線に投影される。以後の議論は、カメラの光軸が床面と平行な場合についてのみ行う。

## G.2 Hough変換

前述のように、3次元空間内の床面に平行な直線は正弦曲線として投影される。この正弦曲線は、Hough変換を用いることにより抽出できる。正弦曲線を式(G.5)で表現した場合、全方位視野像内の点(x,y)は、パラメータ空間a-bにおいて、式(G.6)で示される直線として投影される。

$$y = a \sin \theta + b \cos \theta \quad (\text{G.5})$$

$$b = -a \tan \theta + \frac{y}{\cos \theta} \quad (\text{G.6})$$

また、正弦曲線を式(G.7)で表現した場合には、全方位視野内の点(x,y)は、パラメータ空間a-bにおいて、式(G.8)で示される曲線として投影される。

$$y = \sin(\theta - b) \quad (\text{G.7})$$

$$b = \theta - \sin^{-1}\left(\frac{y}{a}\right) \quad (\text{G.8})$$

## G.3 エピポーラライン

基線距離*l*だけ離れた2つの位置で全方位視野像を撮影して、ステレオ視を行なう場合、そのエピポーララインがどの様になるかについて考える。図G.2に示されるように、3次元空間内の点P=(X,Y,Z)と2つの全方位視野像の焦点O<sub>1</sub>,O<sub>2</sub>を含むエピポーラ平面は、2つの全方位視野像を示す円筒を角度αで切断する。このため、エピポーララインは全方位視野像内で、正弦

曲線となってあらわれる。全方位視野像内のエピポーララインは式(G.9)によって示される。

$$y = f \tan \alpha \cos \theta = f \frac{Y}{Z} \cos \theta \quad \left( \alpha = \tan^{-1} \frac{Y}{Z} \right) \quad (\text{F.9})$$

また、座標軸原点  $O_r$  における全方位視野像への P 点の射影を  $(\theta_r, y_r)$ 、 $O_r$  より X 軸方向に  $l$  だけ離れた点  $O_i$  における P 点の射影を  $(\theta_i, y_i)$  とすると、点 P の座標は式(G.10)で示される。

$$\begin{aligned} X &= \frac{l \cos \theta_i \sin \theta_r}{\sin(\theta_i - \theta_r)} \\ Y &= \frac{y_l \cos \theta_i \cos \theta_r}{f \sin(\theta_i - \theta_r)} \\ Z &= \frac{l \cos \theta_i \cos \theta_r}{\sin(\theta_i - \theta_r)} \end{aligned} \quad (\text{G.10})$$

#### G.4 オプティカルフロー

カメラの光軸が床面に平行で、ロボットが並進運動を行う場合、3次元空間内の一点の、ロボットの移動にともなうオプティカルフローの全方位視野像中の軌跡は、そのエピポーララインと同じ形である。

X, Y, Z 軸に平行なロボットの並進速度成分をそれぞれ U, V, W、X, Y, Z 軸回りの回転角速度成分をそれぞれ A, B, C とすると (図 G.1 参照)、3次元空間の点  $P = (X, Y, Z)$  の速度  $V = (X', Y', Z')$  は式(G.11)で表わされる [Bruss 83]。

$$\begin{aligned} X' &= -U - BZ + CY \\ Y' &= -V - CX + AZ \\ Z' &= -W - AY + BX \end{aligned} \quad (\text{G.11})$$

前述したように、ロボットは床面上で、Y軸回りの回転と、X,Z軸方向への移動のみが許されているので、 $V=0, A=0, C=0$ 、従って式(G.11)は式(G.12)となる。

$$\begin{aligned} X' &= -U - BZ \\ Y' &= 0 \\ Z' &= -W + BX \end{aligned} \tag{G.12}$$

全方位視野内に現れる点 $(\theta, y)$ におけるオプティカルフロー $(u, v)$ は、式(G.1)と式(G.12)を用いて式(G.13)で表わされる。

$$\begin{aligned} u &= \frac{ZX' - XZ'}{X^2 + Z^2} \\ v &= \frac{2Y}{\sqrt{X^2 + Z^2}} \frac{XX' + ZZ'}{X^2 + Z^2} \end{aligned} \tag{G.13}$$

式(G.13)に式(G.1)を代入し、整理すると式(G.14)がえられる。

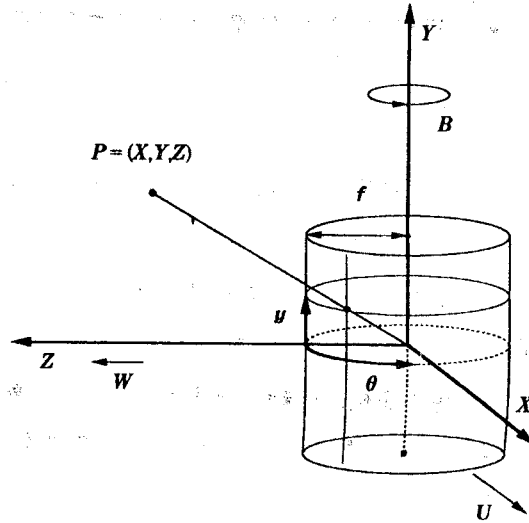
$$\begin{aligned} u &= f \frac{\cos \theta}{Z} (W \sin \theta - U \cos \theta) - B \\ v &= -2y \frac{\cos \theta}{Z} (W \cos \theta + U \sin \theta) \end{aligned} \tag{G.14}$$

これより、ロボットが並進運動のみを行う場合には、FOEは式(G.15)で示される。

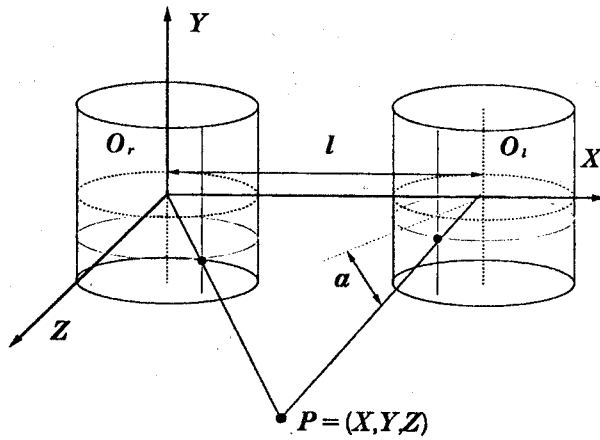
$$(\theta, y) = \left( \tan^{-1} \frac{U}{W}, 0 \right) \tag{G.15}$$

式(G.14)より、全方位視野内に現れるオプティカルフローの軌跡は、ロボットが並進運動のみを行なう場合には、FOEを原点とした正弦曲線となっ

て現れることがわかる。またFOEはそれらの正弦曲線が一点で交わる点である。



図G.1 撮像モデル



図G.2 エピポーラライン



## H. 回転中心の位置合せについて

カメラを回転させ物体の方位を測定する方法では、カメラの回転中心とカメラ焦点の間に多少の位置ずれがあっても、方位角の測定精度にはあまり影響しない。

まず、カメラの焦点が、カメラの回転軸上にある場合について考えてみる。図H.1に示すように、カメラの回転中心とカメラの焦点の間のずれを $l$ 、回転中心とカメラの焦点を結ぶ直線に対するカメラの光軸のずれを $\alpha$ とすると、回転中心を基準とした物体A、B間の角度計測誤差 $E$ は、式(H.1)で表わすことができる。

$$E = \left| \theta_R - \theta_E \right| = \left| \sin^{-1} \left( \frac{l}{D_A} \sin \alpha \right) - \sin^{-1} \left( \frac{l}{D_B} \sin \alpha \right) \right| \quad (\text{H.1})$$

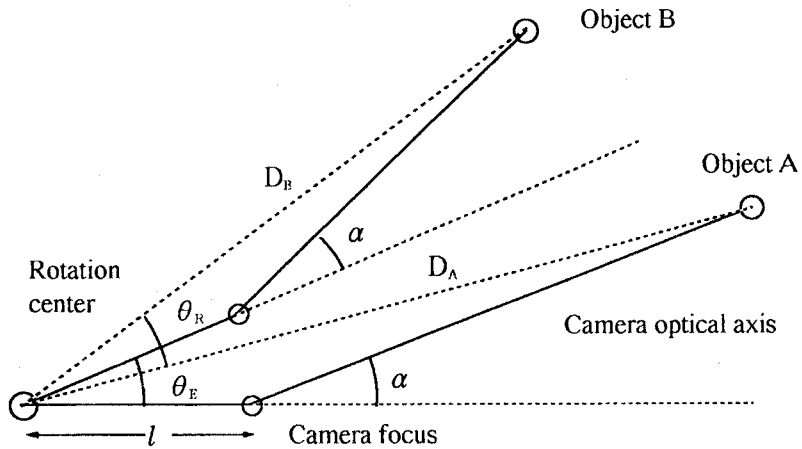
式(2)は $l \ll D_A, D_B$ ならば0に近くなる。実際に

$$l \leq 0.05\text{m}, \alpha \leq 1.0\text{degree}, D_A, D_B \geq 1.0\text{m} \quad (\text{H.2})$$

の場合、誤差 $E$ は最大で $0.0050^\circ$ 程度であり、特別な微調整用の装置なしに、 $l, \alpha$ を式(H.2)の値以下に保つことは可能である。

次に、カメラの焦点が回転軸より $R$ だけ離れている場合(すなわち、カメラの焦点が半径 $R$ の円弧運動を行う場合)について考えると、全方位画像に含まれる方位角情報は、カメラの焦点が回転軸上に位置する場合と同様に、特別な微調整装置なしに得られるが、距離情報に関しては、カメラの光軸の延長線上にカメラの回転軸が正確に位置する必要がある。ここで、調整すべき部分はカメラ光軸の方向のみである。この調整は4.3で述べた簡単な方法でおこなっている。この方法でも、2つの特徴点がカメラの回転中心から十分離れているか、特徴点の位置を正確な方位角情報を持つ全方位画像を撮像して測定すれば、後は、カメラ光軸の方向を微少に変えて調整することで十分正確な調整が可能である。カメラ光軸の微調整はマイクロメータ

付きの小型手動回転台にカメラを取り付けることで簡単にできる。しかし、この方法は大変労力を必要とするため、カメラ光軸の向きを含んだ距離計算式をたて、キャリブレーションにより、カメラ光軸の向きを得る必要がある。



図H.1 回転軸のずれによる方位測定の誤差

