Osaka University Knowledge Archive

| | |
|---|---|
| Title | 単眼視動画像からの関節物体の三次元形状と姿勢の推定に関する研究 |
| Author(s) | 島田, 伸敬 |
| Citation | 大阪大学, 1997, 博士論文 |
| Version Type | VoR |
| URL | https://doi.org/10.11501/3129070 |
| rights | |
| Note | |

*Osaka University Knowledge Archive : OUKA*

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

Osaka University

# 単眼視動画像からの関節物体の三次元形状と姿勢の推定に関する研究

1996 年 12 月

島田 伸敬

## 論文概要

本論文では単眼視動画像から関節物体の３次元形状と姿勢を推定する手法について論じた．１章では，単眼視動画像から物体モデルを用いて関節物体の形状と姿勢を推定する場合の困難と，本論文で提案する解決について概説した．

単眼視動画像から物体モデルを用いて関節物体の形状と姿勢を推定する場合，１）オクルージョンや物体の速い動きによって画像とモデルの対応づけに曖昧さが生じる．２）関節物体は形状と姿勢の自由度が高いため，探索空間が膨大になる．３）奥行きが観測できないため，形状と姿勢の間に曖昧さが生じる，といった困難がある．そこで本論文では推定問題を以下の２つの段階に分けて論じた．

まず，２章では大まかな形状モデルを用いて画像とモデルの対応づけと物体姿勢の大まかな推定を行なう問題について論じた．ここでは典型的な関節物体である人間の手指について，シルエット動画像からその３次元姿勢を推定する手法について述べた．大まかな形状モデルや，特徴とモデルの対応が曖昧な状況で姿勢が推定できるように，手指の三次元形状モデルから見え方候補を生成して単眼視入力の画像と照合する方法を提案した．また確率に基づく候補の評価や動きに基づく予測とビーム探索を利用した効率的でロバストな探索法を提案した．章末では，これらの手法を手指の動きが速く頻繁に自己隠蔽を生じる実画像列に対して適用し，その有効性を示した．

次に３，４，５章では，対応づけと大まかな姿勢があたえられたもとで，形状と姿勢の双方を同時に推定する手法について論じた．ここでは，関節角度や各部の長さ，太さに関する制約知識を利用して，単眼視画像における形状と姿勢の曖昧性を解決する手法を提案した．まず，考えられる形状と姿勢をパラメータ空間における領域として表現し，制約知識や画像から得られる観測を統合する枠組を考えた（３章）．次にこの推定の枠組を拡張カルマンフィルタを用いて実現すると，観測の非線形性や奥行きの曖昧性のために誤推定が生ずることを実験によって明らかにした（４章）．そこで，観測の非線形性を考慮した楕円近似法と，複数の推定候補を生成して一時的な奥行きの曖昧性を適切に取り扱うことで，形状と姿勢の範囲を同時に推定する手法を提案した（５章）．章末では，最初に大まかな形状と姿勢の範囲を与えた上でさまざまな時系列観測を入力すると，単眼視画像から物体の形状と姿勢が正しく推定できることを，実験例を使って示した．

最後に６章において，本論文の成果のまとめと課題について述べた．

# もくじ

**1 はじめに** **1**

1.1 コンピュータビジョンによる３次元物体の認識 . . . . . 1

1.2 モデルを用いた関節物体の形状と姿勢の推定 . . . . . 2

1.3 本論文の構成 . . . . . 3

**2 手指のシルエット動画像からの姿勢推定** **6**

2.1 関連研究と本研究の関わり . . . . . 6

2.2 シルエットマッチングに基づく姿勢推定と処理の概要 . . . . . 9

2.3 手指シルエット特徴の抽出 . . . . . 12

2.4 手指モデルの初期化 . . . . . 15

2.5 姿勢候補の生成 . . . . . 16

2.5.1 階層的な姿勢候補の生成 . . . . . 18

2.5.2 姿勢パラメータの適応的量子化 . . . . . 18

2.5.3 動きに基づく事前確率順探索 . . . . . 20

2.6 姿勢候補の評価 . . . . . 21

2.6.1 掌姿勢候補の評価 . . . . . 22

2.6.2 突起状特徴と指モデルとの対応づけ . . . . . 23

2.6.3 指姿勢候補の評価 . . . . . 24

2.6.4 手指全体の評価 . . . . . 26

2.7 複数の姿勢候補保持とビーム探索法 . . . . . 29

2.8 実験結果 . . . . . 32

2.9 まとめ . . . . . 36

**3 関節物体の姿勢と形状の同時推定の原理** **38**

3.1 関連研究と本研究の関わり . . . . . 38

3.2 形状と姿勢に関する制約知識の利用による奥行きの曖昧性の除去 . . . . . 40

3.3 形状と姿勢ならびに制約知識の領域表現 . . . . . 47

**4 カルマンフィルタを用いた形状と姿勢の推定とその問題点** **50**
4.1 システムの方程式 . . . . . 50
4.2 拡張カルマンフィルタ . . . . . 51
4.3 制約による推定分布の修正 . . . . . 52
4.4 線形な制約による分布切りとり法 . . . . . 53
4.5 実験１ . . . . . 55
4.6 複数候補の生成による非線形性の補償 . . . . . 58
4.7 実験２ . . . . . 60
4.8 拡張カルマンフィルタの問題点 . . . . . 64
4.8.1 リンク長の推定値が誤っていても分散が小さくなる問題点 . . . . . 64
4.8.2 複数の推定候補が生成できない問題点 . . . . . 66

**5 楕円近似法を用いた形状と姿勢の推定** **67**
5.1 楕円近似法 . . . . . 67
5.2 観測の非線形性の見積もりと補償 . . . . . 70
5.3 楕円近似による予測 . . . . . 72
5.4 実験１ . . . . . 75
5.5 モード分解による計算量の減少 . . . . . 75
5.6 複数候補の生成 . . . . . 78
5.7 実験２ . . . . . 80
5.7.1 楕円近似法による形状と姿勢の推定例 . . . . . 80
5.7.2 楕円近似法における複数候補の生成例 . . . . . 81
5.7.3 大きさの異なる物体に対する推定の比較 . . . . . 87
5.7.4 楕円近似法による EKF の問題点の改善 . . . . . 92
5.7.5 楕円近似法による推定の精度上の問題点 . . . . . 97
5.8 まとめ . . . . . 99

**6 おわりに** **100**

**A 投影行列と剛体回転行列** **106**
A.1 投影写像行列 . . . . . 106
A.2 ３次元の剛体回転行列 . . . . . 106

**B EKF における分布切り取りによる制約の統合** **109**

**C 楕円近似法におけるトレース最小解と体積最小解** **111**

**D 楕円近似法における適切な線形化位置の計算** **113**

**E 楕円近似法における観測の相関性を考慮したモード分解** **115**

# 第 1 章

# はじめに

## 1.1　コンピュータビジョンによる３次元物体の認識

古くから人間はさまざまな機械をつくり利用してきた．すべての機械は労力を必要とする仕事を簡単に解決するための道具である．かつて機械は肉体的な力や技を補助するための自分の手足の延長であったが，現在では情報の収集や判断といった知性的な能力を補助する機械が中心的役割を果たすようになった．また，機械が複雑になるにつれて，機械をより簡単に操作するために，自動的に人間の意図を汲みとる知能的インターフェース装置の開発が競われている．これらの機械，装置はいわば人間の目や耳そして脳の延長である．

人間の目（視覚）は外界の情報を収集する機能に長けており，視覚のもたらすきわめて多くの情報が脳による判断の材料に使われている．この人間の視覚による情報収集と判断処理を代行する機械をテレビカメラとコンピュータを用いて構築しようとするのがコンピュータビジョンである．コンピュータビジョンには多くの課題があり，さまざまな研究が行なわれているが，我々や機械が存在する世界は３次元空間であるから，３次元物体を認識する問題がもっとも中心的な課題である．コンピュータビジョンによって人間のような３次元物体が自動認識できれば，人間がジェスチュアで機械に指示するようなインターフェースを構築することができる．また，馬や人間などの動きを画像から計測して，コンピュータグラフィックスで仮想的な動物を動かすこともできる．

一般にコンピュータビジョンによる３次元物体の認識とは，カメラ画像に写っている３次元空間中の物体が１）何であるか，２）どのような形，姿勢，動きをしているか，を自動的に推定することである．一般的なシーンの場合は画像にさまざまな物が写っている可能性があるため，１）を推定するのは難しい問題である．

しかし，画像による機械とのインターフェースのように，注目する特定の物体の形状や姿勢を計測するという問題を考えた場合，画像に写っている物体に関する知識が利用できるの

で，未知の物体よりも形状や姿勢が正確に推定できる．

## 1.2 モデルを用いた関節物体の形状と姿勢の推定

既知の物体知識を３次元の構造モデルとして表現し，その３次元形状と姿勢，動作を動画像から推定する研究が従来から行なわれてきた (2.1節， 3.1節で詳述)．剛体物体の推定についてはこれまでさまざまな研究がなされているが，ヒューマンインターフェースのような応用を考えるときには，人間や動物のような各部が可動する物体の３次元推定が必要になる．

関節物体の３次元形状と姿勢の推定には以下にあげるような問題がある．

第一は，画像とモデルの対応づけが難しいことである．モデルを用いて形状や姿勢を推定するには，画像中のどの領域がモデルのどの部分に対応するかを決定しなければならない．しかし，人間の動きは速く，またオクルージョンのため形が複雑に変化するため，安定に抽出できる画像特徴が少ない．しかも，対応づけに何通りもの可能性がある場合が多い．従来の研究では簡単な対応処理に基づくものが多く，実際の画像列に対しての頑健さにかける．

第二は，関節物体の姿勢の自由度が高く，推定に多大な計算量を必要とすることである．人間の手指は関節の角度や物体の向きだけで23自由度になり，関節の可動範囲などの制約を考慮しても，可能な姿勢の空間の大きさはきわめて膨大である．画像から得られる限られた特徴だけでは完全に可能性を絞れない上に，関節構造は非線形性が高いため，部分的に可能性を絞ることも簡単ではないため，従来の手法ではきわめて大きな計算量を必要とする．

第三は，単眼視画像では形状が正確にわかっていなければ，姿勢は正確に推定できないことである．人体の形状には個人差があり，前もって正確に知ることができない場合が多い．形状と姿勢を同時に推定しようとすると，物体の自由度がさらに大きくなるのに加えて，単眼視画像では奥行きが観測できないために長さと角度の曖昧性が生じる．従来の研究にはステレオ視からの３次元情報を利用した推定手法もあるが，人間は単眼でも姿勢を推定できることから，機械にもそれが可能であると考えられる．

モデルと画像の対応付けの曖昧性や画像情報の不足のために，物体姿勢の可能性が複数生じる場合は，基本的にそれらをすべて考慮する必要があり，それには多量の計算量を必要とする．しかし，あらかじめ大まかな初期推定をあたえておけば，その初期推定に基づくの可能性の高さを評価することができる．したがってその可能性の高い順に候補を検証していけば，すべてを探索しなくても正しい推定候補が見つかると考えられる．最後にいくつかの可能性が残る場合は，ひとつに限定せず，やはり可能性の高いいくつかの候補のみを残していけば，無駄な探索を減らしつつ頑健な推定が可能になる．

また，正確な形状がわかっていなくても，初期推定の大まかな形状（長さ，太さなど）を

用いれば大まかな姿勢（関節角度や向き）は推定できる．そこでまず，形状を固定して姿勢のみを粗く推定し，可能性を限定する．次に形状を推定しようとすると，単眼視画像では奥行きがわからないので角度と長さの間に曖昧性が生じるが，人体など既知の関節物体は，関節角度や長さなどに関する制約知識がわかっているから，これを用いれば形状と姿勢の可能性を限定できる．

そこで本研究では次のような枠組を用いる (図 1.1).

1. 大まかな形状と姿勢の初期推定もしくは前時刻での推定を最初の手がかりとして用いる．
2. 1 での手がかりを利用して画像特徴とモデル各部の対応づけを仮定する．
3. 仮定された対応づけのもとで，大まかな姿勢の推定を行なう．
4. 大まかな推定が得られたら，次に物体に関する制約知識を用いて，粗く得られている形状と姿勢を修正する．
5. 次々に入力される動画像に対してこのプロセスを繰り返し，次第に正確な推定に近付けていく．

本論文では，このように大まかな推定を次第に修正していくことによって，適切かつ効率的に可能性を限定し形状と姿勢を推定する枠組を提案する．

## 1.3 本論文の構成

本論文では前節で述べた枠組に基づき，単眼視動画像から関節物体の３次元形状と姿勢を推定する問題を

1. 画像とモデルの対応づけがわからないもとでの，姿勢の大まかな推定
2. 対応づけと大まかな姿勢がわかったもとでの，形状と姿勢双方の精密な推定

の２段階にわけて論ずる．本論文の本論は以下の４つの章によって構成される．

まず２章で前者の問題について論ずる．ここでは典型的な関節物体である人間の手指について，シルエット動画像からその３次元姿勢を推定する手法について述べる．大まかな形状モデルや，特徴とモデルの対応づけの曖昧さがある状況で姿勢が推定できるように，手指の三次元形状モデルから見え方候補を生成して単眼視入力の画像と照合する方法を用いる．また確率に基づく候補の評価や動き基づく予測とビーム探索を利用した効率的でロバストな探

## 1.ROUGH MATCHING Phase

画像とモデルの対応づけ

大まかな精度の姿勢推定 (関節角度 のみ)

## 2.REFINEMENT Phase

大まかな姿勢推定の修正 (形状と関節角度)

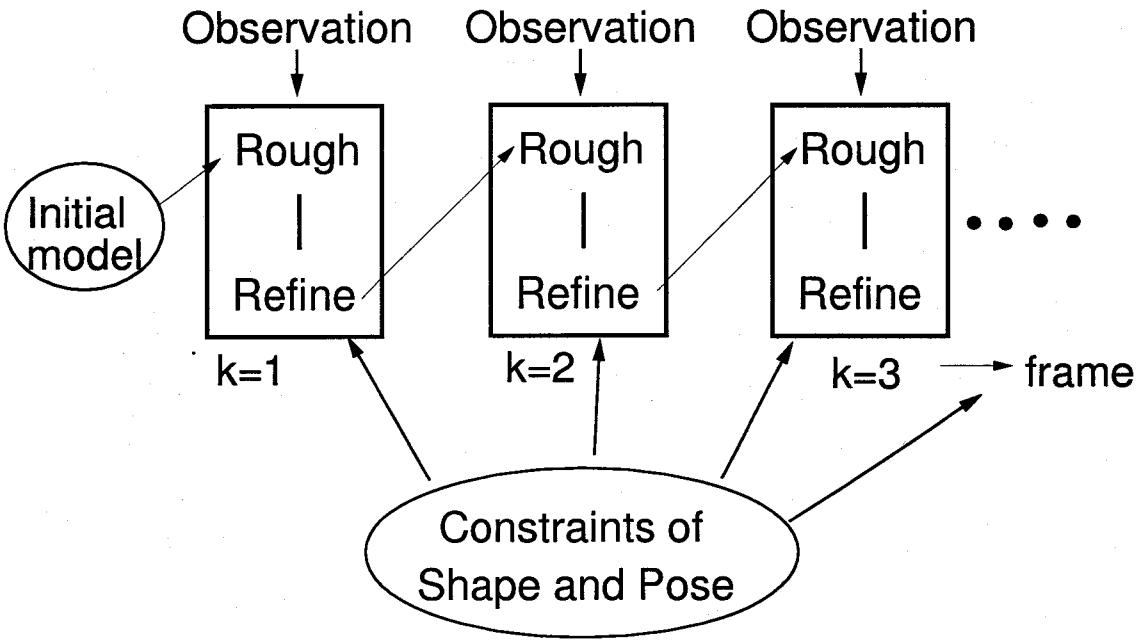



図 1.1: 関節物体の形状と姿勢推定における粗い推定と精密な推定の関係

索法を提案する．これらの手法を手指の動きが速く頻繁に自己隠蔽を生じる実画像列に対して適用し，その有効性を示す．

つづく３，４，５章では，対応づけと大まかな姿勢があたえられたもとで，形状と姿勢の双方を同時に推定する手法について論ずる．まず３章では，関節角度や各部の長さ，太さに関する制約知識を形状と姿勢の推定に利用する枠組について論ずる．ここでは考えられる形状と姿勢をパラメータ空間における領域として表現し，制約知識や画像から得られる観測を統合する手法を提案する．４章では，３章で述べる推定の枠組をカルマンフィルタを用いて実現する手法とその実験例を示し，観測の非線形性や奥行きの曖昧性によって正しく推定できない問題点があることを明らかにする．そこで５章では，３章の枠組を楕円近似法を用いて実現し，また複数の推定候補を生成して一時的な奥行きの曖昧さを適切に扱うことによって，４章で述べる問題を解決する．最後に本手法を用いた実験例を用いて，その有効性を示す．

# 第 2 章

# 手指のシルエット動画像からの姿勢推定

本章では，もっとも身近な関節物体として人間の手指の３次元姿勢を単眼視動画像から推定する手法について論じる．

## 2.1 関連研究と本研究の関わり

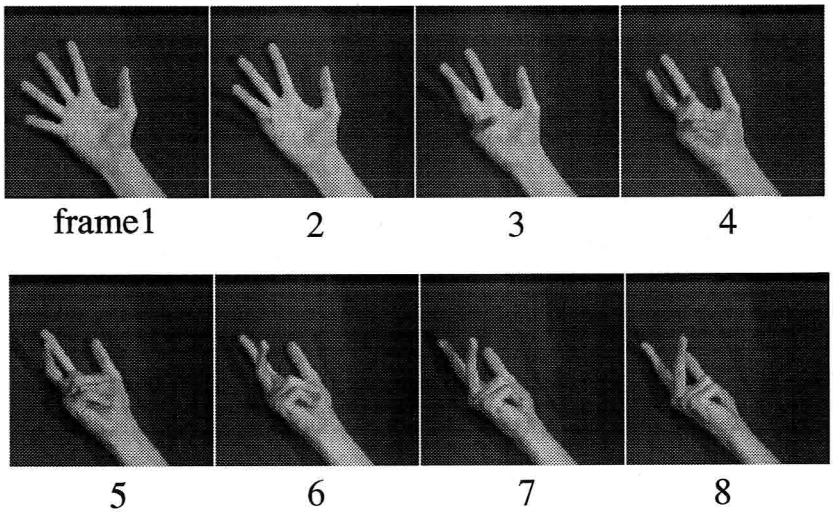
手振りは我々がものを指示したり意思を表示するときにもっとも頻繁に用いられる手段のひとつなので，手振りを機械に自動認識させることができれば，使いやすい機械のインターフェースが構築できる．

指文字のような簡単な命令を識別するだけならば，２次元の画像パターンの識別が有効である [1][2]．しかし詳細な手振りの動作を識別するためには手指の三次元姿勢の推定が必要である．

手指の３次元姿勢の推定には大きく分けて二つの手法がある． ひとつはデータグローブのような器具を装着する手法 [3] であり，そしてもうひとつはコンピュータビジョンの技術を利用した手法である．前者の場合，まずグローブを装着する必要があり，利用するときのオーバーヘッドが大きい．これは不特定のユーザが頻繁に入れ替わりつつ使用されるような環境では望ましくない．この点は後者のコンピュータビジョンの方法によって改善することができる．

画像を用いて手指の３次元の姿勢を推定する場合は，手指の３次元形状モデルを知識として利用することが一般的である．このとき，画像から得られる特徴とモデルの各部分との対応づけが必要になる．

簡単に対応づけを行なうために，マーキングを施した指先，指関節や手首をカメラで観測する手法 [4][5][6][7] などがあるが，マーキングを用いることは先のデータグローブを利用する手法と同様の問題点がある．




指の動きは速く、かつオクルージョンが起きる。

図 **2.1**: 手指の入力画像列の例

画像のサンプリング時間内に手指はわずかしか動かないという仮定を用いて画像特徴を追跡し，前時刻での推定結果と現在の入力の差から関節角などの姿勢パラメータを最小二乗法による修正する手法 [8][9][10] がある．またオプティカルフローを利用して人体の３次元の動きを推定する手法 [11][12][13][14] も提案されている．しかし通常の手振りをビデオレート程度の間隔で撮像する場合には，この仮定はかならずしも成立しない．図 2.1に示すように，手指の形は複雑に変化する．いくつかの指が他の指や掌に隠されることによって，その指に相当する特徴が画像中から急に消失してしまうことが頻繁に発生するため，画像中からモデルの各部に対応する部分を探すことは難しくなる．そのような場合には特徴の追跡に失敗することがあり，オプティカルフローの精度も低くなるので，これらの手法では推定が不正確になる．

オクルージョンに対して効果的な手法として，形状モデルから考えられる姿勢候補を積極的に生成し，その中で 入力シルエット (図 2.1) に対する適合度がもっとも高い候補を探索するシルエットマッチング法 [15][16][17] がある．この手法は入力シルエットと形状モデルの対応があらかじめわからなくても姿勢推定ができるので，オクルージョンがおきる場合に姿勢推定を行なうには有望なアプローチである．しかしこの手法には図 2.2に示す問題点がある．

まず，人間の手指の自由度は大きいために探索空間は広く，一般に全探索は現実的でない．[18] は掌から指先へむかって順に姿勢を決定するヒューリスティックを利用して探索

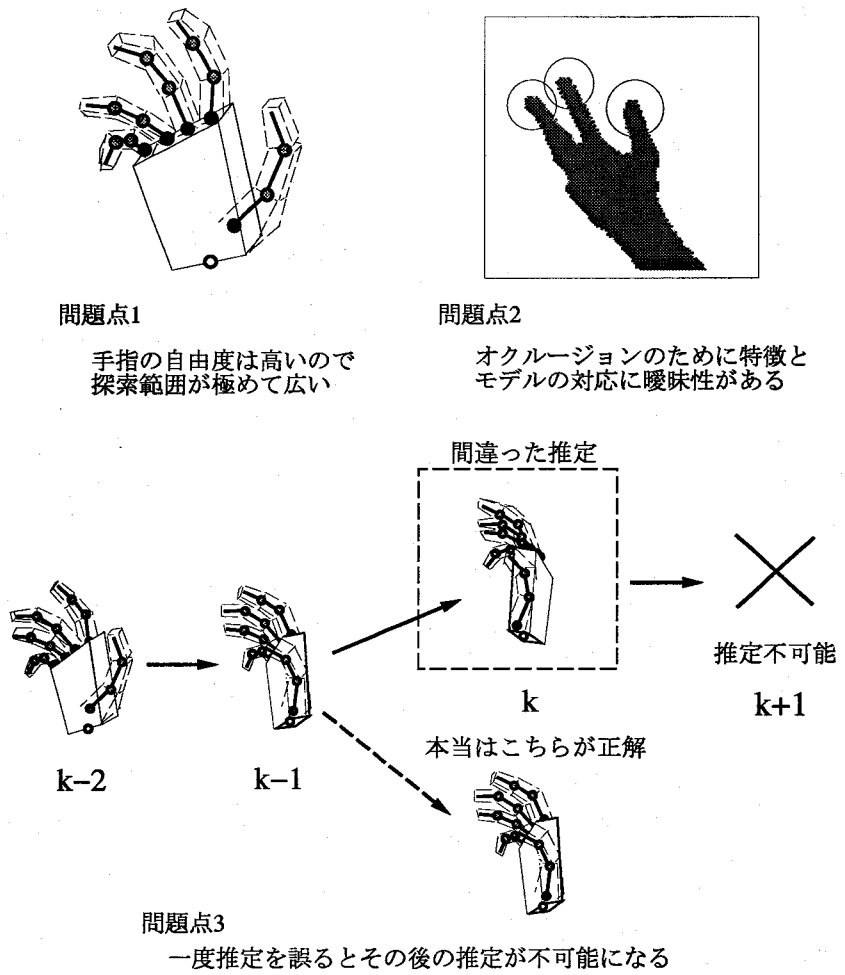



図 2.2: 手指の姿勢推定における問題点

空間を削減したが，常に最初に見つかった解を採用すると，オクルージョンによって複数の解釈が可能なシルエット形状の場合には局所最適に陥ることがある．Genetic Algorithm を応用してこれに対処した研究 [19] があるがやはり計算量の大きさが問題になる．

[18][20] は二つのカメラを用いて曖昧性を除去しているが，オクルージョンがおきた場合には一方のカメラからは観測できない部分が生じるので，単眼視と同じ問題になる．人間は単眼でも姿勢を推定できることから機械にもそれが可能であると考えられる．また，カメラの数は少ない方がコストが低くできるメリットがあるので，単眼視による推定の研究価値は大きい．

また，現在の入力に対して最良な推定がその後の時系列を観測すると必ずしも最適ではないという場合があることも問題となる．この場合には一時的に推定に失敗するので，何らかの対処をしなければそれ以降の推定を続けることができない．機械とのインターフェースへ

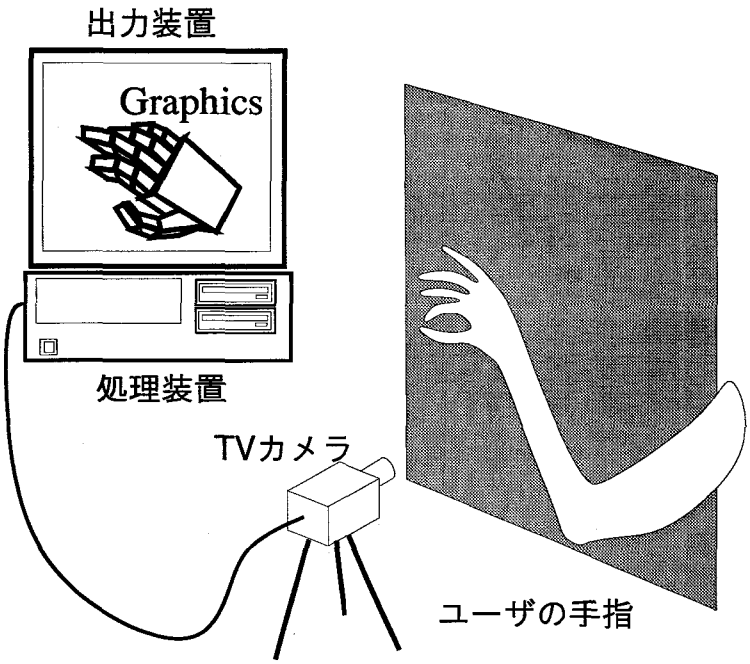



図 2.3: 取り込みシステム

の応用を考えると，ロバスト性の確保は重要な課題である．

本研究は，オクルージョンに対処できるようにシルエットマッチングに基づいて，単眼視シルエット動画像から手指の三次元姿勢を推定する．本研究では上で述べたシルエットマッチングの問題点を解決するために，以下の手法を提案する．

まず，動きの予測から可能性の高い候補を優先的に探索，照合を行なうことで，探索量を削減する．また，画像中に観測される形状特徴と姿勢候補の各部位の対応づけをいくつか仮定し，部位ごとに適した姿勢候補の評価を行って対応づけを検証する．評価に際しては，観測された特徴のもとでの姿勢候補の条件つき確率を用い，異なる部位の評価値を統合して正確な照合を行う．さらに，ビーム探索法 [21] を利用して複数の推定候補を保存しながら推定を行う．これによって，現在の入力に対して最良の推定が誤りであっても推定を続けることができ，必要な場合にはその誤りを後で訂正することができる．

## 2.2 シルエットマッチングに基づく姿勢推定と処理の概要

本研究では図 2.3に示すような画像入力システムを用いる．手指のシルエットを抽出しやすいように黒板を背景とし，手指全体が画面内に映るようにする．また一般に時系列による推定では初期値の与え方が重要になるが，ここでは初期推定を与えやすいように，単純な

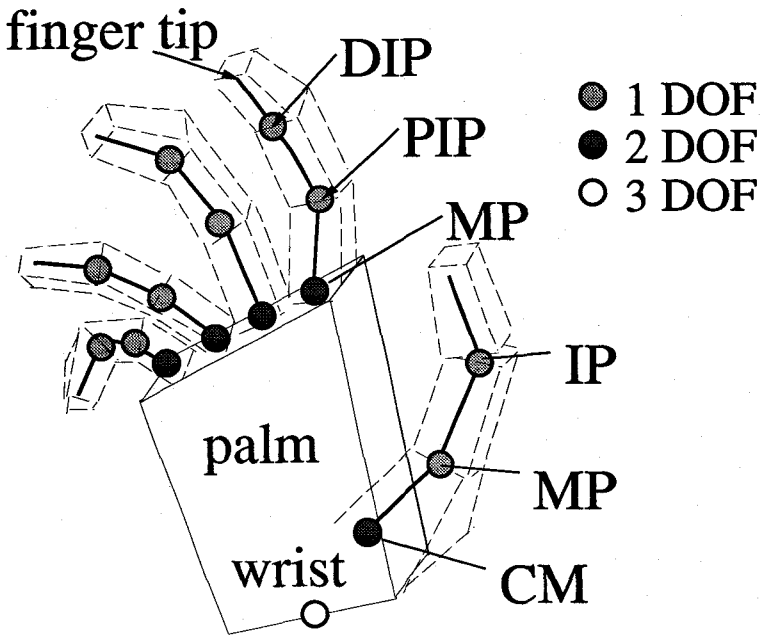



図 2.4: 手指モデル

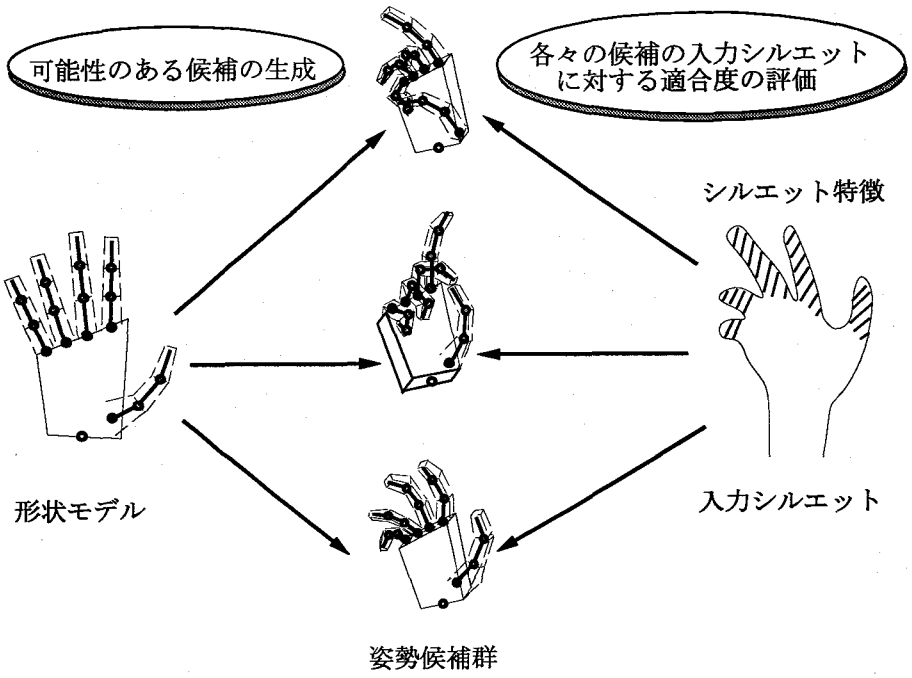



図 2.5: シルエットマッチングに基づく姿勢推定

姿勢（たとえば各指を伸ばして掌を画像平面に平行に配置した姿勢）から姿勢入力をはじめる．

我々の手法で用いる手指モデルは，図 2.4に示すように，手首の位置，16 個の各剛体部分の形状と合計 23 自由度 (DOF) の関節，さらに剛体部分の接続関係の記述からなっている．掌モデルは厚みをもった板状のポリゴンで表現される．指モデルは，観測できる指の姿勢推定に用いる中心軸モデルと，掌に隠れて観測できない指の推定に用いる箱状のポリゴンからなる．手指の姿勢は手首位置と各関節角度のセットによって記述され，このモデルを用いて手指のとりうるあらゆる姿勢の投影画像を直交射影を用いて生成することができる．

本手法では，可能性のある候補のシルエット画像を形状モデルから次々に合成し，入力シルエットとの適合度を評価して適合度の高い候補をその時刻の推定姿勢とする (図 2.5)．

以下に姿勢推定の処理手順を示す．

1. 現フレームの画像を入力する．
2. 手首の位置と形状特徴を抽出する．最初のフレーム以外は 4 へ．
3. モデルの形状と姿勢の初期化を行う．
4. 前フレームの推定から考えられる掌姿勢候補を生成し，入力に対する評価が良いものを一つ選ぶ．
5. 4 で得られた掌候補について，前フレームから考えられる指姿勢候補を生成し，入力に対する評価がしきい値以上のものを選んで，掌候補とともに手指の推定候補とする．もししきい値以上のものが得られないときは 4 で選んだ掌候補を棄却する．
6. 得られた手指の推定候補が規定の個数に達していなければ 4 へ戻る．
7. 得られた手指の推定候補の手指全体としての評価をし，順位づけた上で保存する．
8. 時刻を 1 フレーム進めて 1 へ戻る．

2 では，手指のシルエット画像から指らしい領域と手首を抽出する (2.3節)．3 では，2 で抽出された突起状の特徴と手指についての知識を用いて，手指モデルの形状の拡大率と各関節角度を初期化しておく (2.4節)．

4,5 の入力シルエットに対して適合度の高い姿勢候補を探索する処理は，

**(A)** 評価の対象となる姿勢候補群の生成 (2.5節)

**(B)** 生成された候補の，入力シルエットに対する適合度の評価 (2.6節)

の二つに分けられる．2.1節で述べたように，可能性のある候補の探索領域は膨大である．また姿勢候補の投影画像を生成する処理のコストは高いので，探索中に投影を生成する回数を極力減らす必要がある．そこで (A) において，

**(A-1)** 関節角度の変動に対してシルエットの変形度合を均一にする適応的な探索空間の量子化 (2.5.2節).

**(A-2)** 動き予測に基づく事前確率の高い候補順の投影画像生成 (2.5.3節).

の各処理を行う．

次に，(B) の処理時には従来手法のように重なり面積だけを利用するのではなく，

**(B-1)** 掌については，はみ出し形状を考慮した評価 (2.6.1節)

**(B-2)** 指については，シルエット中の突起状の特徴に基づいた評価 (2.6.3節)

を行う．この時特徴とモデルの各指との照合 (2.6.2節) を行うことで，シルエットが重ならなくてもシルエット形状の類似度を正しく評価する．

7 では，手指全体の適合度を評価するために，手指の各部位に対する評価値を統合する (2.6.4節). 手指全体の評価が最も高い候補をそのフレームでの暫定の推定姿勢とする．各部位の評価値の，入力シルエットのもとでの条件付き確率をベイズ推定により求め，正規化された適合度とする．手指全体の評価が最も高い候補をそのフレームでの暫定の推定姿勢とする．このとき一時的に推定を誤った場合や，オクルージョンによって解釈に曖昧性がある場合に対応するために，適合度の高い候補をいくつか保持しておき，それらから予測される姿勢候補の近傍のみを探索するビーム探索法を用いる (2.7節).

以下の節では各処理の内容について論ずる．

## 2.3 手指シルエット特徴の抽出

[16][18] は人体の形状モデルから姿勢の候補となる合成シルエット画像を生成し，モデルの画像と入力シルエットの重なり面積を用いて候補の適合度を評価した．この手法はいかなる三次元モデルについてもその投影領域を生成するだけで評価が可能であるという一般性を持つが，可能性のある候補の探索領域は膨大なので，計算量が非常に多くなってしまう．

一般に階層構造を持つモデルでは，根にあたるパーツの姿勢は全体の大局的な姿勢 (どこにあるかなど) を決定するのに対し，枝にあたるパーツの姿勢は局所的な輪郭特徴を作り出

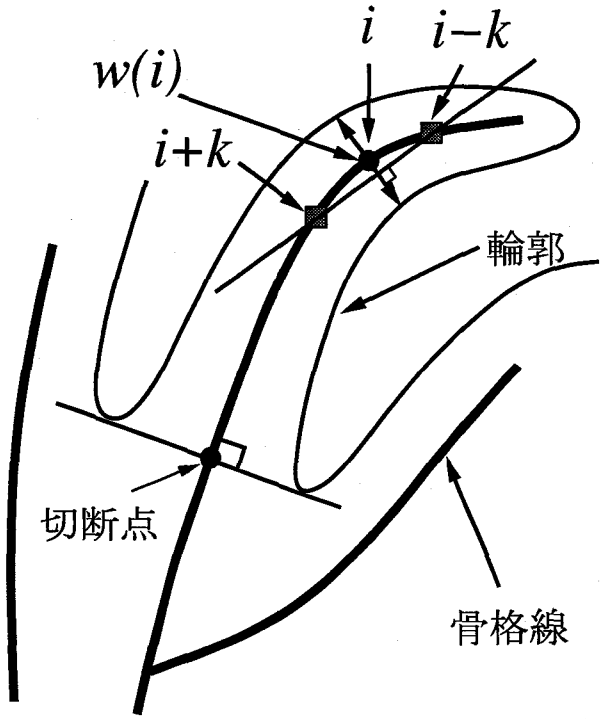



図 **2.6**: シルエット特徴の抽出

す場合が多いと考えられる．したがって良好な局所的輪郭特徴が得られる場合には，その情報を用いて候補の可能性を限定し，また枝にあたるパーツの姿勢を評価することができる．そこで本手法では，指らしい領域として手指シルエット画像から突起状の特徴を抽出する．

まず手指のシルエット画像を細線化することでその骨格線を抽出する．中心線を端からたどりながら腕の幅が明確に変化する場所を次の処理によって見つけ，骨格線を切断する．

1. 骨格線の端点から分岐点までを分離する．

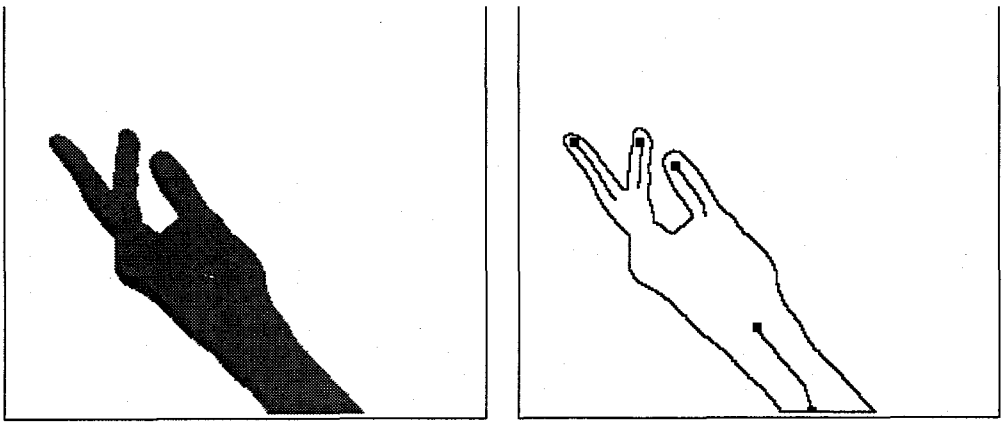
2. 分離した骨格線を端点から順にたどりながら，骨格線上の各点 $i$ に対して垂直にもとの二値図形の幅 $w(i)$ を計算する．

3. 現在点での図形の幅とそれまでの平均との比がしきい値以上なら (式 2.1)，現在点を切断点とする．そうでなければ 2 へ戻る．

$$(i-1)\cdot w(i)/\sum_{j=1}^{i-1} w(j) > threshold \tag{2.1}$$

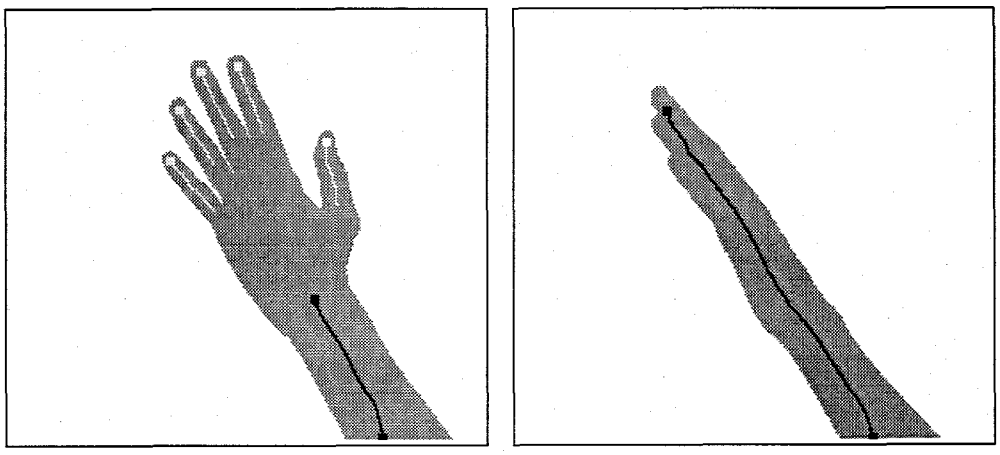
$i$ 番目の点 $(x(i), y(i))$ に対する垂直方向は図 2.6に示すように

$$\theta = \begin{cases} \tan^{-1}\left(-\dfrac{x(i-k)-x(i+k)}{y(i-k)-y(i+k)}\right) & (y(i-k)-y(i+k) \neq 0) \\ \frac{\pi}{2} & (y(i-k)-y(i+k) = 0) \end{cases} \tag{2.2}$$



(a) 入力シルエット　　(b) 抽出された特徴

図 **2.7**: 抽出された突起状特徴



(a) 抽出できる場合　　(b) 抽出できない場合

図 **2.8**: 手首位置の抽出

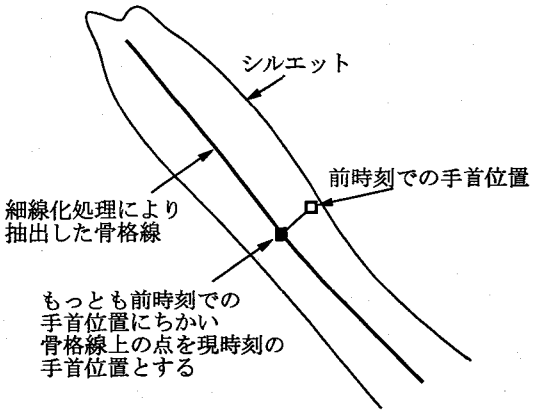



図 **2.9**: 手首位置が抽出できない時の処理

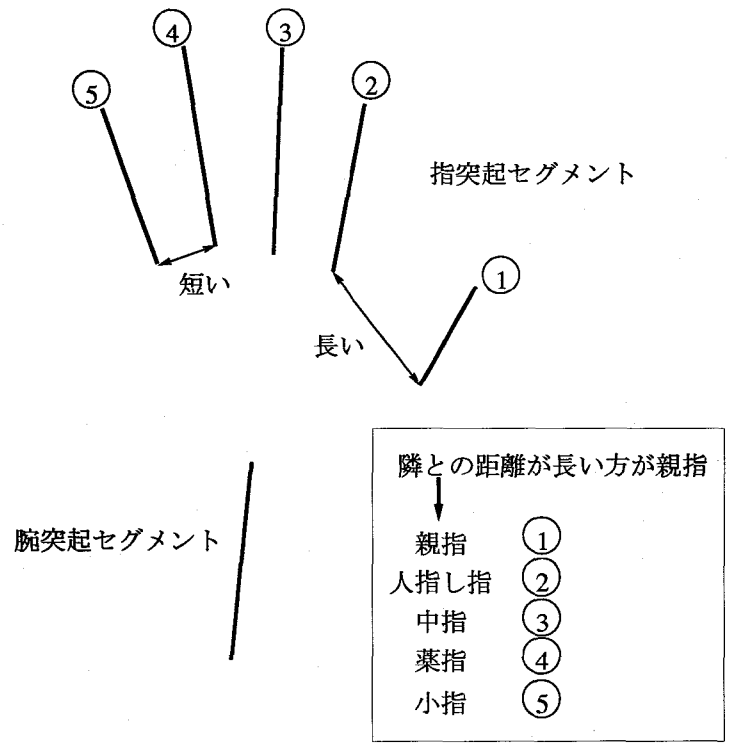



図 **2.10**: モデルの初期化における突起状特徴と指モデルの対応

で与えられる．以上の処理によって，図 2.7に示す突起状の特徴が得られる．

「手指は完全に画面の中に入っている」という仮定から，端点が画面の縁にかかっているものは腕に相当する特徴であると仮定する．腕だと判定された骨格線では式 2.1のしきい値を指のものより小さくして切断点が得られやすいようにする．こうして得られた切断点を手首位置とし，腕以外に抽出された突起状の領域は指に対応する形状特徴とする．場合によっては図 2.8(b) のように，手首位置に相当する切断点がえられない場合もある．そのときは図 2.9に示すように，前のフレームにおける手首位置に最も近い骨格線上の点を手首位置として推定する．

## 2.4 手指モデルの初期化

人間の手指には個人差があり，指の長さ，太さ，関節位置，掌の大きさなどが異なる．また画像を取り込むカメラの焦点距離やカメラと手指の間の距離により，撮影される手指の大きさが変化する．このため三次元モデルを用いる際にはこれらの変化を吸収するために，モデルを初期化する必要がある．

そこで，本研究では単純な初期姿勢を時系列の最初に用いることで，手指の三次元モデルの初期化を行う．初期姿勢として図 2.12(a) のような掌を正面にして各指を伸ばした姿勢を

採用した．

本研究では画像中の手は右手であると仮定した上で，モデルに関する知識として

1. 親指以外の各指の MP 関節 (図 2.4参照) の重心，手首位置，親指の CM 関節位置 (図 2.4参照) の 3 点を結んだ三角形の 3 つの内角

2. 各指における各節の長さの比

3. 親指以外の各指のつけ根における太さの平均と掌の厚さの比

を利用することで，不明な関節位置や掌の厚さを推定する．

モデルを初期化する際に，まず 5 つの突起状特徴がどの指に対応するかを決定する必要がある．これは図 2.12(a) の姿勢だけを初期姿勢に考えるので，親指，人指し指，中指，薬指，小指がこの順番にならぶと仮定できる．従って一番端にある突起状特徴に対応するのが親指か小指かを決定すればよい．親指の付け根と人指し指の付け根の距離は，薬指の付け根と小指の付け根の距離に比べて長いことから，図 2.10のように両端の突起状特徴について，そのとなりの突起状特徴と手の平側端点の距離を求め比較する．距離の長いほうを親指であるとする．残りの指と突起状特徴の対応は先の仮定から順に決定できる．

つぎに手指の各関節位置を求める．手首位置と，親指以外の MP 関節位置は画像より抽出される．そこで，4 指の MP 関節位置の重心と手首位置を用いて，図 2.11のように手首位置を通ってカメラの光軸に平行な軸まわりの掌の回転角度を求める．またこの重心と手首位置を結んだ線分を掌の中心軸にとり，掌座標系の座標軸とする．掌の中心軸の長さは掌の大きさを代表しているのでこれをスケールパラメータ $C_{scale}$ とする．本研究ではスケールパラメータは時系列全体に渡って一定と仮定する．

親指の CM 関節位置は，4 指の MP 関節位置の重心位置，手首位置，知識 1 できまる相似な三角形から求める．親指以外の指の DIP，PIP 関節の位置は対応する突起セグメントの端点どうしを結ぶ直線を，知識 2 を用いて内挿する．掌の厚さは各指の MP 関節位置における太さの平均と知識 3 を用いて求める．

以上のようにして初期化されたモデルが図 2.12(b) である．

## 2.5 姿勢候補の生成

手指が関節が多く自由度が大きいので，姿勢候補の探索領域はきわめて広大になる．本節では可能性のある手指の姿勢候補を効率的に生成するための手法について述べる．

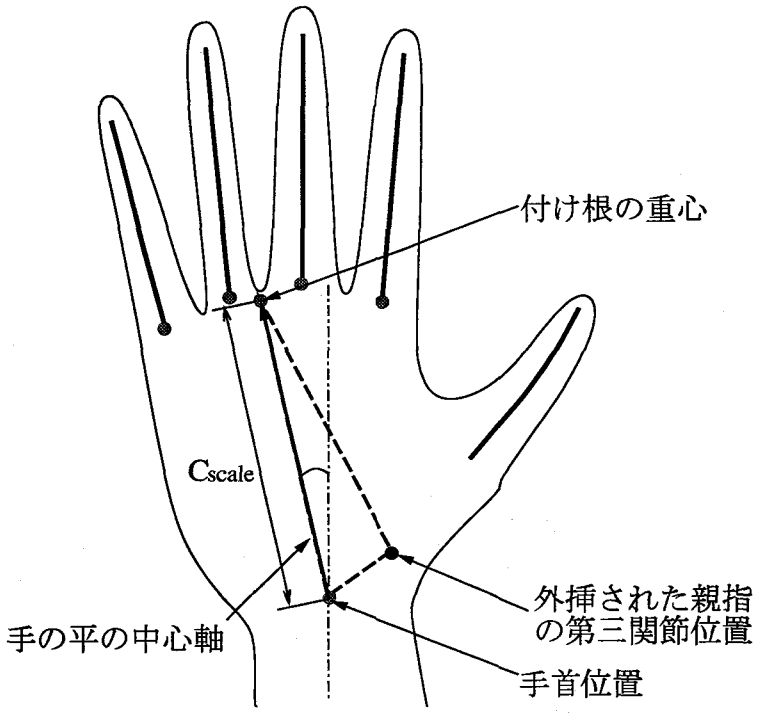
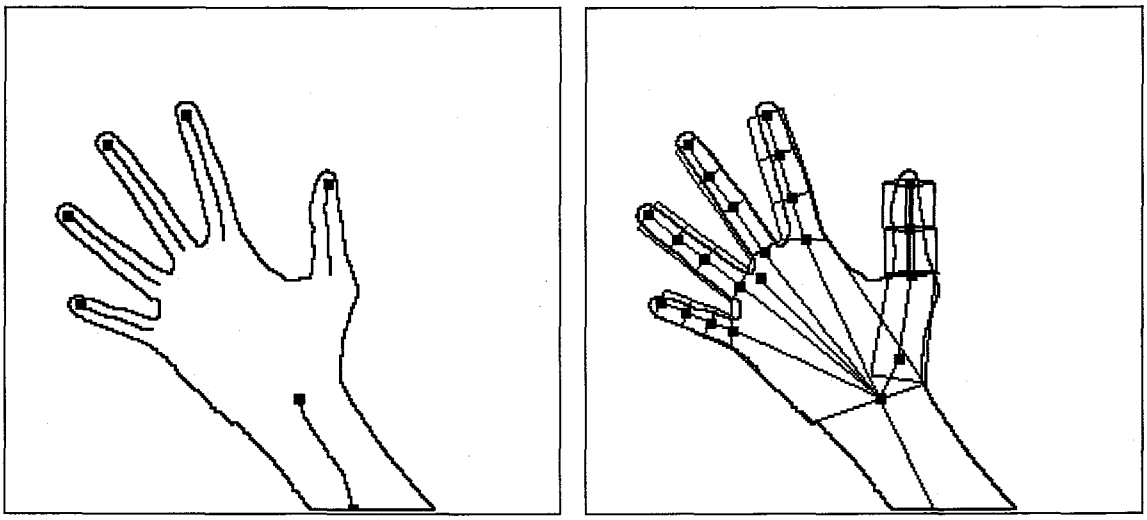



図 **2.11**: 手の平の中心軸と親指の CM 位置の推定



(a) 初期姿勢　　(b) 初期化された手指モデル

図 **2.12**: 手指モデルの初期化

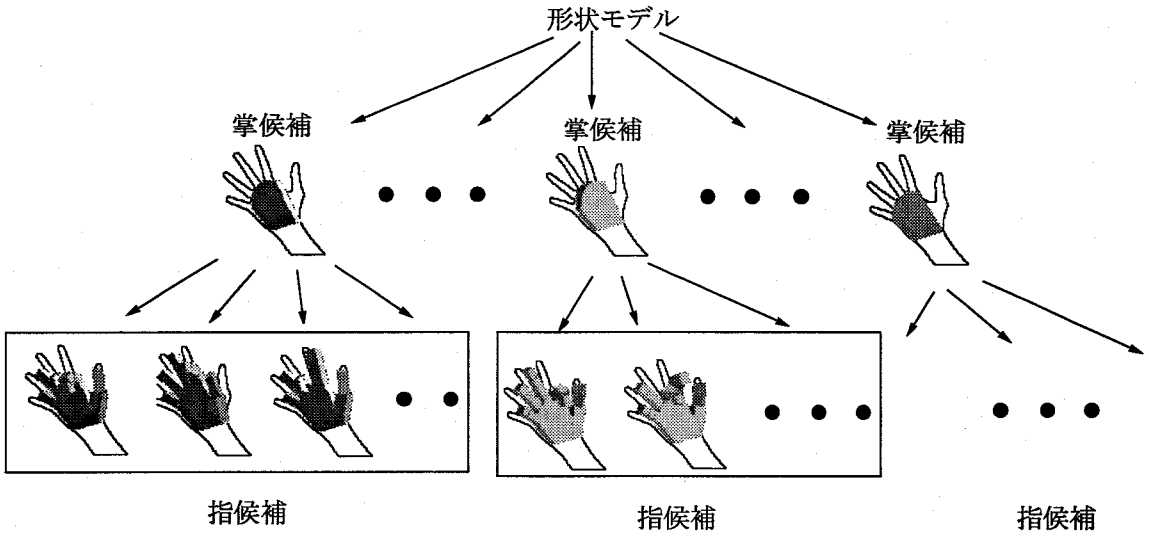



図 **2.13**: 階層的な姿勢候補の生成

### 2.5.1 階層的な姿勢候補の生成

手指は掌と指の各部に分けられる．手指の自由度は極めて高いのですべての自由度に対する全探索は現実的ではない．指は掌の先に附属しているので，まず掌を姿勢候補を生成し，入力シルエットに適合する掌候補についてのみ各指の姿勢候補を生成，評価する階層的な姿勢候補の生成を考える．[18] は深さ優先で姿勢の候補を生成したが，この場合は局所最適に陥ることがある．そこで本手法では，もしシルエットに適合する掌候補が複数えられたときは，それぞれの掌候補に対して指姿勢候補を生成する．

本来各指の自由度は 4 つあるが，親指以外の指では複雑な場合を除き，縦に曲がる 3 つの関節はほぼ同じ角度で曲がることが多い．指の姿勢はもっと少ない数のパラメータによってモデリングできるという報告 [22] もある．本研究の趣旨は，3 章以降で述べる精密な形状と姿勢の推定の初期値として利用する大まかな手指の姿勢を推定することにある．そこで今回はこれらを同一角度という拘束を用い，姿勢候補を大まかに探索した．

### 2.5.2 姿勢パラメータの適応的量子化

一般にある点の 3 次元運動をカメラで観測する場合，速度ベクトルがカメラの光軸に垂直なほど画面上での見かけの速度は大きくなる．ある領域が剛体運動した場合にカメラで観測されるその輪郭の変形や移動の度合にも同様なことがいえる．したがって，関節の単位角度変化による剛体回転によって画面上に生じる領域輪郭の変形は，その部分の三次元姿勢に依存する．姿勢パラメータを変化させながらシルエットに適合する姿勢候補を探索する場合に

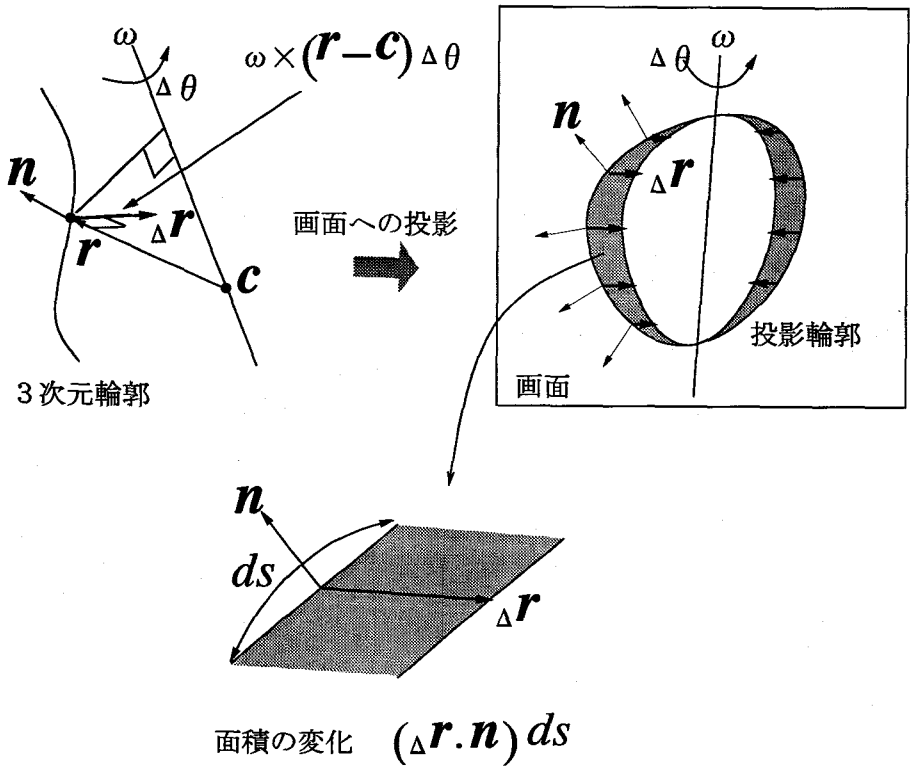



図 2.14: 関節角の変化による輪郭形状の変形度合

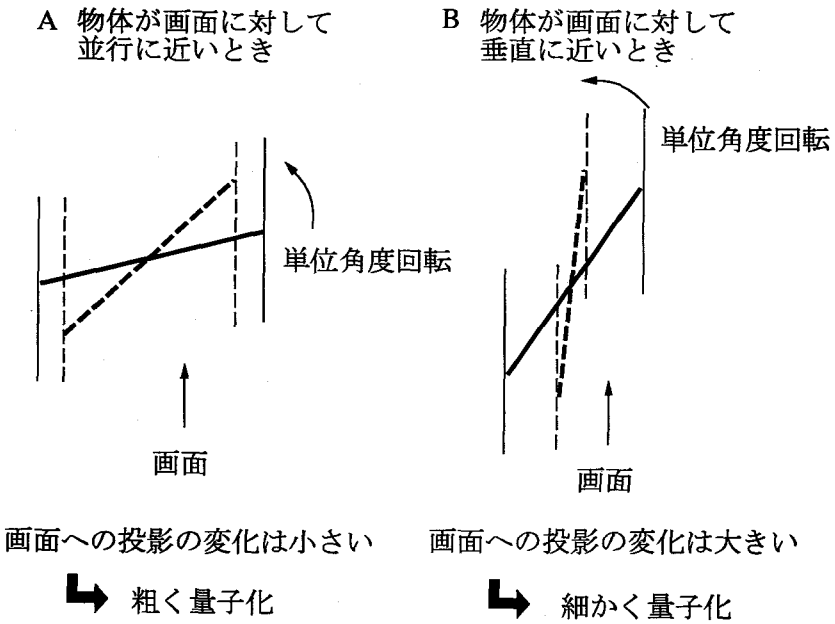



図 2.15: 姿勢パラメータの適応的量子化

パラメータを均等間隔で変化させると，投影画像が極端に変化したりほとんど変化しなかったりする場合が多い．投影画像が極端に変化してしまうところではその途中の姿勢を考慮にいれることができず，ほとんど変化のないところでは同じような候補を何回も評価する無駄が生じる．手指の形状モデルは既知なので，モデルから計算される観測輪郭の変形が均一になるように姿勢パラメータを変化させることにより，この問題を解決できる．

手指のある関節が回転した場合に，画面上で観測されるシルエットの面積の変化量を以下のように見積もる．シルエットが変化するのは，その関節より指先側にある部位全体だから，この部分の投影輪郭で囲まれた領域 $S$ について考える．$S$ は前時刻における推定姿勢を用いて求める．図 2.14に示すように，点 $\boldsymbol{c}$ を通って方向が $\boldsymbol{\omega}(|\boldsymbol{\omega}| = 1)$ で与えられる軸回りに微小角度 $\Delta\theta$ だけ回転することを考えた場合，$S$ の輪郭上の点 $\boldsymbol{r}$ の変位 $\Delta\boldsymbol{r}$ は $\boldsymbol{\omega} \times (\boldsymbol{r}-\boldsymbol{c})\Delta\theta$ で与えられる．このとき変位 $\Delta\boldsymbol{r}$ が $S$ の画面上における面積変化に寄与する分は，輪郭線素 $ds$ あたり $\Delta\boldsymbol{r} \cdot \boldsymbol{n}ds$ と近似できる (図 2.14参照)．ただし $\boldsymbol{n}$ は $S$ の投影輪郭の法線ベクトル (つまり $z$ 成分が 0) である．従って，$s$ を $S$ 上の位置パラメータ ($s \in [0,1)$) として

$$D = \oint_S |[\boldsymbol{\omega} \times \{\boldsymbol{r}(s) - \boldsymbol{c}\}] \cdot \boldsymbol{n}(s)| ds \tag{2.3}$$

とおくと，微小回転によって引き起こされる $S$ の画面上での面積の変化量 $\Delta S$ は，$D\Delta\theta$ と見積もることができる．手指の姿勢パラメータが決まれば $D$ が計算できるので，変形量 $\Delta S$ が定数 $S_c$ になるような $\Delta\theta = S_c/D$ を探索幅として用いれば，画像の変形が均一になるようなステップで探索することが期待できる．

この議論は輪郭が別の部分に隠されたり隠されていた輪郭が新たに現れることを無視しており $\Delta\theta$ が微小な範囲でしか成立しないが，ある関節角度について細かく探索するか荒く探索するかの指針を与える．なお，探索幅が荒くなりすぎるの防ぐために $\Delta\theta$ の上限値 $\Delta\theta_{max}$ をあらかじめ設定しておく．

### 2.5.3 動きに基づく事前確率順探索

本研究では動画像を用いているので，動きの滑らかさを仮定すれば，図 2.16に示すように，時刻 $k-2$ と時刻 $k-1$ の推定から線形予測によって時刻 $k$ の姿勢が予測できる．予測が適正であれば適合度の高い候補が早い段階で見つかる．そこで，本研究では可能性の高い姿勢候補を効率的を生成するために，予測にもとづく事前確率の高い順に各候補の投影画像を生成する．

予測姿勢 $\tilde{\boldsymbol{\theta}}$ の姿勢パラメータ空間における近傍を $V$ とする．$V$ は予測からのずれの範囲であり，画像のサンプリング時間と考慮すべき関節角の加速度の最大値から決められる．

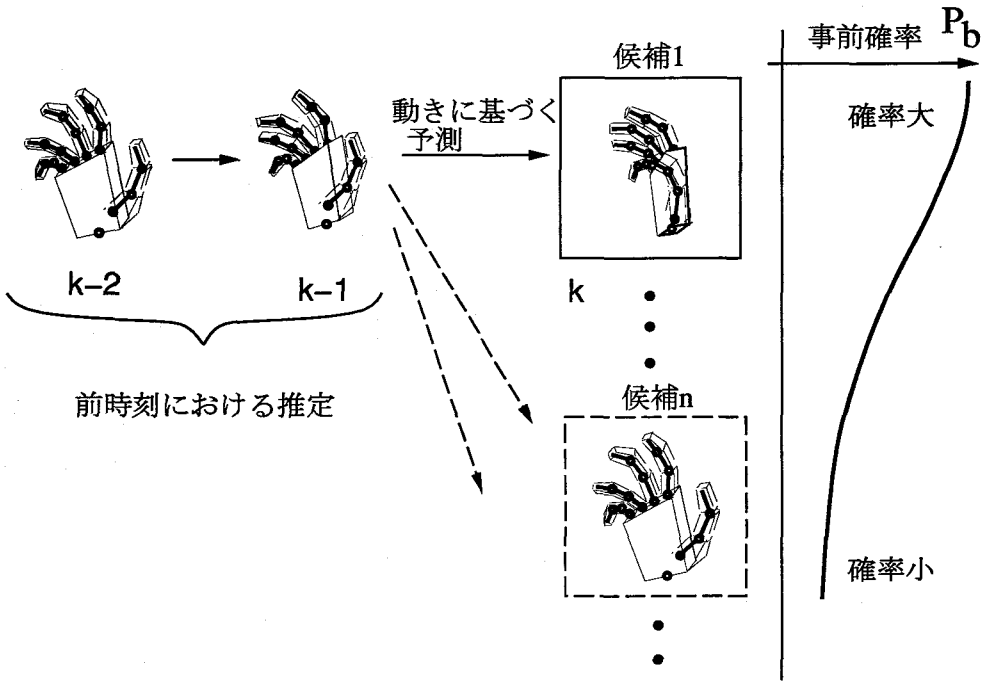



図 2.16: 動きに基づく姿勢予測

$V$ は前節で述べた各関節に対する探索幅によって適応的に量子化されている．適合度の高い候補は予測姿勢を中心に分布する，と仮定することは予測が適正であるかぎり妥当である．そこで, 姿勢候補 $\boldsymbol{\theta}_j \in V$($j$ は候補のインデックス) に対してこの分布を，平均 $\tilde{\boldsymbol{\theta}}$，分散が $\Lambda_t$ である正規分布

$$P_b(\boldsymbol{\theta}_j) = N(\tilde{\boldsymbol{\theta}}, \Lambda_t) \tag{2.4}$$

であると近似する．$\Lambda_t$ は各姿勢パラメータの可能なフレーム間変位から決められる定数である．この $P_b$ を姿勢候補 $\boldsymbol{\theta}_j \in V$ の事前確率として用いる．$V$ に含まれる各候補についてこの $P_b$ を計算することは軽い処理であるので，予測が適正であれば $P_b$ の高い候補から探索することによって，適合度の高い候補が早い段階で見つかる．探索は一定数の適合度の高い候補が見つかるか $P_b$ があるしきい値以下になった時点で終了する．

## 2.6 姿勢候補の評価

本節では前節の手法によって生成された各姿勢候補について，入力シルエットに対する適合度を評価する手法について述べる．手指は掌と指の各部に分けられる．

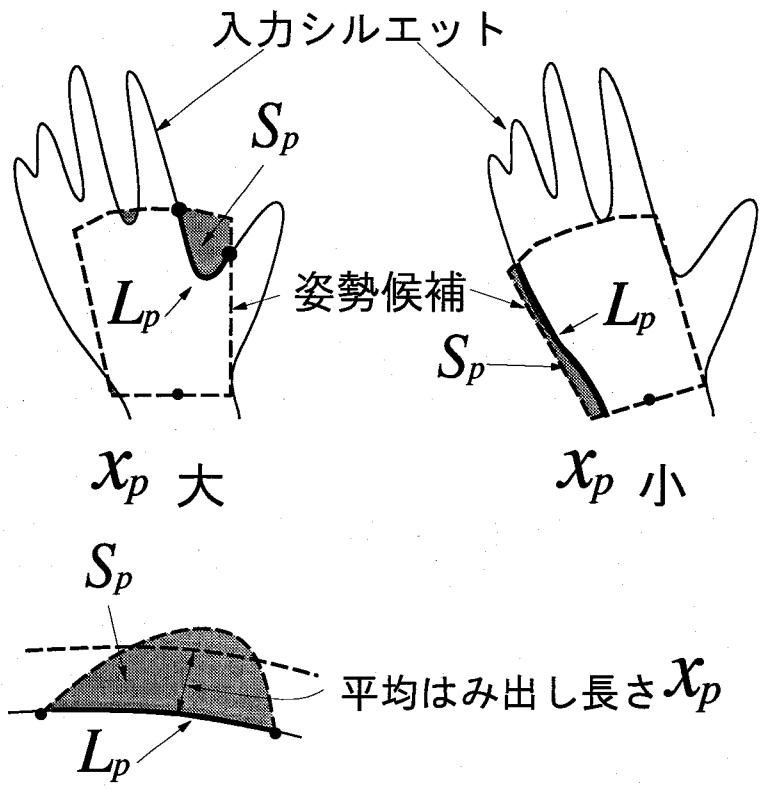


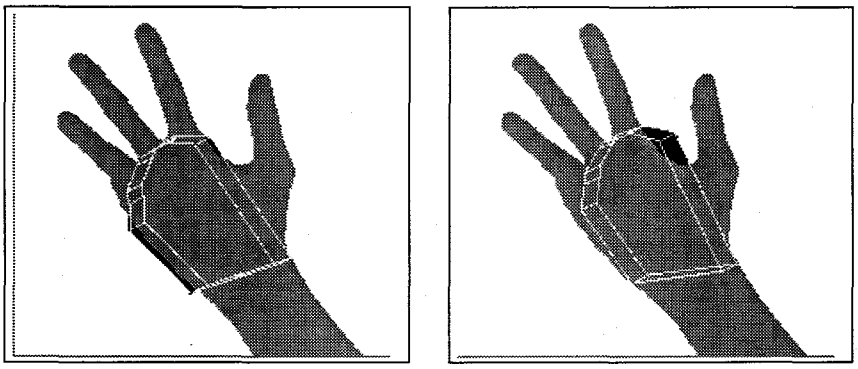

図 **2.17**: 掌姿勢候補の評価

### 2.6.1 掌姿勢候補の評価

掌のモデルは掌自体の形状と五本の指の可動範囲から構成されている．各掌姿勢候補に対して生成される投影画像は次の二つの観点で評価される．まず指であると考えられる，入力シルエットから抽出される突起状の特徴が，指の可動範囲を画面に投影した領域内に収まっているかどうかを調べる．次にこの条件を満たすものについて，掌モデルの投影画像が入力シルエットからはみ出す度合を評価する．

[16][18] は評価の基準として，掌モデルの投影画像と入力の手指領域の重なり面積の割合を用いている．しかし，もともと手指が変形をともなう非剛体であることや姿勢パラメータ空間を量子化していることからくる位置ずれを許容しなければならない．このとき人間が判断した場合明らかに間違いであるとわかる姿勢候補までも許容範囲のなかに入ってしまうという場合がある．多くの場合，次節で述べる指姿勢候補の評価の時点で間違いを除去できるが，各指姿勢候補の生成と評価は計算コストが非常に高いため，余分な候補を掌姿勢候補の評価の段階で除去することは計算量を削減する上で重要である．

本研究では図 2.17に示すように，はみ出し面積を両領域の接触長さで割った平均はみ出



(a) 候補 1　　(b) 候補 2

図 **2.18**: 掌候補の評価例

し長さ $x_p$ を次式で計算する．

$$x_p = \frac{\sum_i S_p(i)}{C_{scale} \sum_i L_p(i)} \tag{2.5}$$

ここで，$S_p(i)$ はモデル投影が入力からはみ出した $i$ 番目の領域面積であり $L_p(i)$ ははみ出した領域と入力シルエットの接触する長さである．$C_{scale}$ は形状モデルを画面に平行投影する際の拡大率であり，$x_p$ は投影時の大きさによらない正規化された量になる．$x_p$ を用いることにより，図 2.17に示すような，はみ出し面積 $S_p$ が同じ二つの場合を区別することができる．もし $x_p$ があるしきい値より小さければその候補は有望とみなし，各指の姿勢推定に入る．

図 2.18は掌部の二つの姿勢候補の適合度を評価した例である．白線で示されているのが掌モデルの投影であり，黒い領域が投影がシルエットからはみ出した部分である．モデルの投影面積に対するはみ出し面積の比は，(a) が 95.84%，(b) が 94.39% でありほぼ同じであるが，平均はみ出し長さ $x_p$ は (a) が 2.51，(b) が 10.14 となり (a) のほうが良い評価を得た．この例からシルエットのはみだし形状を評価することで，はみ出し面積をつかった場合より適切な候補を選択できたことがわかる．この適合度による候補の評価の精度は形状モデルの精度に依存する．しかし本研究では候補の大まかな姿勢が得られればよいので探索幅を大きくとっており，シルエットの差異が大きいものを上述の評価法により十分識別できる．

### 2.6.2　突起状特徴と指モデルとの対応づけ

2.3節で抽出された突起状特徴 (図 2.7) は指に対応すると考えることができる．この突起状特徴を利用して指姿勢候補を評価するためには，まず突起状特徴がどの指モデルと対応す

るのかを決定する必要がある．最も単純な対応付けは，突起状特徴の掌側端点と指モデルのMP 関節の投影位置の間の距離を基準にするものである．そこで以下の処理を行なう．

1. 5 個の指モデルと抽出された突起状特徴 (1 ～ 5 個) のすべての組合せについて突起状特徴の掌側端点と指モデルの MP 関節の投影位置の距離を計算し小さい順にソートする．

2. ソートされた順に対応を決定する．もしすでにその指モデルや突起状特徴が対応づけされていた場合は多重に対応付けを行わないようにする．

オクルージョンが生じている場合は突起状特徴の掌側の端点が MP 関節に対応しないために，対応を間違えることがある．そこで，まず上述の基準で対応を仮定し，次節で述べる指姿勢候補の適合度を評価する．その結果適合度が低く棄却される指があった場合に，その突起状特徴の対応付けを次のように修正する．

1. 特徴に対応付けされていないの指モデルがあるときは，それらのうち画面上での距離の近いものから優先して対応付けをやりなおす．

2. もし 1 であるしきい値より適合度の高い指姿勢候補が得られないときは，既に突起状特徴に対応付けされている指モデルと交換し，指姿勢候補の生成と評価を試みる．交換した指モデル双方についてしきい値以上の適合度が得られれば，その対応を新たな対応とする．得られなければ，別の対応済み指モデルと交換を試みる．

3. どの指モデルに対応づけしても適合度の高い姿勢候補が得られない特徴があった場合は，掌姿勢候補が誤っていると判断しこれを棄却する．

このような対応付けの修正を施すことで，指モデルと突起状特徴のすべての対応に関して探索と評価を行わずに正しい姿勢候補を得ることができる．

### 2.6.3 指姿勢候補の評価

前節の方法で仮定した対応づけのもとで，指姿勢候補の適合度の評価を行なう．もしオクルージョンによって抽出された特徴が五つより少ない場合，各指モデルは特徴に対応する指とそうでない指の二つに大別できる．

特徴に対応づけされた指の候補については，図 2.20に示すように，まずモデルの指先と各関節の投影位置を計算する．オクルージョンのために突起状特徴の掌側端点は必ずしも

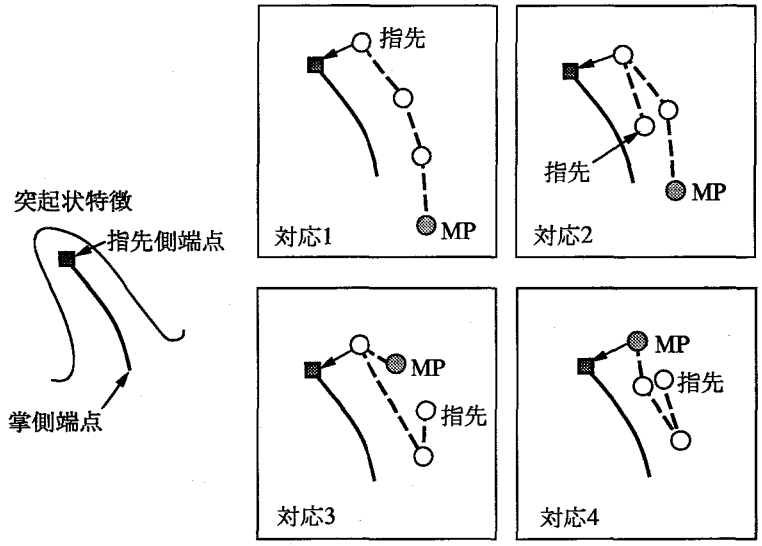



図 **2.19**: 指先，関節と突起状特徴の端点との対応の可能性

MP 関節 (図 2.4参照) に対応するとは限らないが，[23] は少なくとも指先側の端点はモデルの指先か関節に対応すると考えた (図 2.19)．そこで本研究では特徴の端点とモデルの指先及び指関節の対応すべてについて，突起状特徴の指先側端点との画面上距離 $x_{ft}$ を指候補の位置ずれとして評価する．次に，対応点が重なるように指モデルを画面上で $x_{ft}$ だけオフセットする．このとき指モデルの中心軸と突起状特徴が囲む面積 $x_{fa}$ で，両者の角度ずれを評価する．

$$x_{ft} = |\boldsymbol{P}_M - \boldsymbol{P}_I| \tag{2.6}$$

$$x_{fa} = \sum d(i)/L \tag{2.7}$$

ただし，$\boldsymbol{P}_M$ はモデルの指先もしくは各関節の投影位置，$\boldsymbol{P}_I$ は突起状特徴の指先側端点であり，$L$ は特徴と指モデルの投影の中心軸のうち短いほうの長さである．また $d(i)$ は，モデルと突起状特徴のそれぞれの中心軸をたどって道のり $i$ にある点どうしの距離である．$x_{ft}$，$x_{fa}$ がともにあるしきい値より小さいものを推定姿勢として選ぶ．投影位置と投影形状の類似性を独立に評価するので，投影位置が入力と厳密に一致しなくても，投影形状が類似している候補を選択できる．

特徴に対応しない指については上述の特徴に基づく評価ができないので，，掌の場合と同様に指モデルの投影領域を生成し入力シルエットからの平均はみ出し長さを用いて評価を行う．

以上の指姿勢候補の評価の有効性を実験例を用いて示す．図 2.21はオクルージョンが発

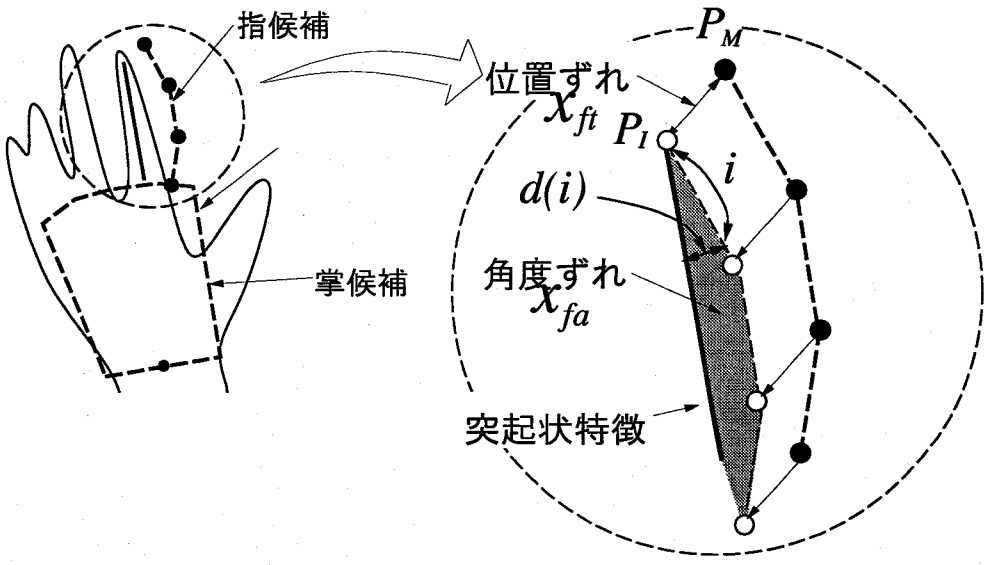


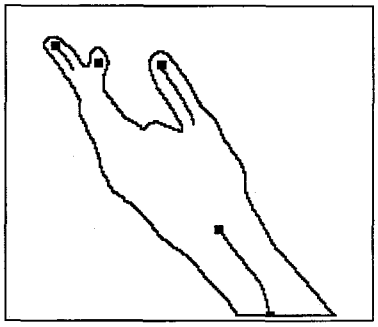

図 **2.20**: 指姿勢候補の評価

生して指の一部分しか突起状特徴として抽出されない場合に対する推定結果の例である．図 2.21(a) は突起状特徴の端点が指先に対応する場合の入力シルエットの輪郭である．輪郭内部の黒い実線は抽出された突起状特徴を示している．(b) は (a) に対する姿勢推定結果である．この例の場合，実際の手指形状とモデルの形状には差があるために，親指の位置が多少ずれている候補しか生成することができなかった．しかし本手法では広域な輪郭形状特徴と指モデルとのマッチングを行っているので，精密な形状モデルを用いなくても形状の類似性を正しく評価して，位置のずれに対してロバストに画像を照合することができた．また，(c) はオクルージョンのために突起状特徴の端点が指先ではなく指関節に対応する場合の入力シルエットであり，(d) はその姿勢推定結果である．このように，必ずしも指先が輪郭として観測されない場合でも，指関節を突起状特徴の端点に対応させた場合を考慮し適合度を評価することによって，姿勢を正しく推定することができた．

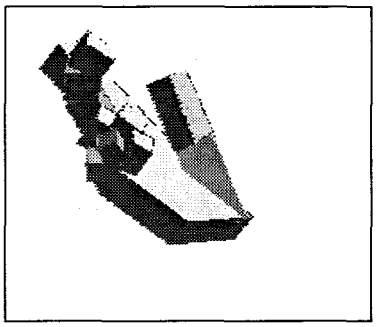
### 2.6.4 手指全体の評価

前節のようにして得られた適合度の高い姿勢候補は一般に複数ある．そこで手指全体としての適合度の評価を行ってそれらを順位付ける．前節までにおいて姿勢候補の評価に用いた手指の各部分に対する適合度 $x_n \in \{x_p,\ x_{ft},\ x_{fa}\}$ は異なった種類の観測量であるから，そのままでは統合することができない．

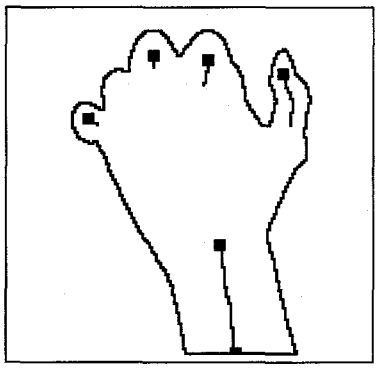
そこで，本手法ではベイズ推定に基づく確率モデルによって $x_p, x_{ft}, x_{fa}$ のもとでの手指



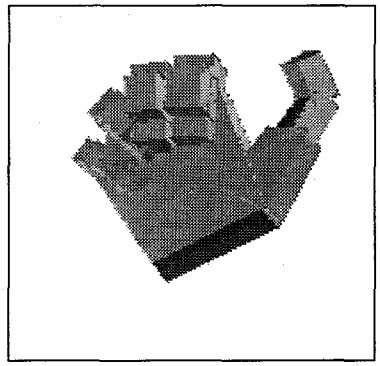
(a) 入力画像列1 (フレーム6)



(b) (a) に対する推定姿勢



(c) 入力画像列2 (フレーム7)



(d) (c) に対する推定姿勢

図 **2.21**: 手指姿勢の推定例

姿勢候補 $\boldsymbol{\theta}_j$ の条件付き確率を用いて統合を行う．

前節で得られた適合度の高い姿勢候補 $\boldsymbol{\theta}_j$ は各姿勢パラメータの探索幅 $\Delta\theta_m(m=1,\cdots,M$(手指モデルの自由度)) によって量子化された区間の代表値であるから，真の姿勢 $\boldsymbol{\theta}$ は区間 $I[\theta_{j,m}-\frac{\Delta\theta_m}{2} \leq \theta_m \leq \theta_{j,m}+\frac{\Delta\theta_m}{2}; m=1,\cdots,M]$ の間で一様分布する．またモデルが実際の手指形状と完全に一致すると仮定すると，$x_n$ の理想的な観測量 $x_n^*$ を手指モデルを用いて $\boldsymbol{\theta}$ の関数として導くことができる (図 2.22中の太線)．本論文では $\boldsymbol{\theta}_j$ の近傍にある $\boldsymbol{\theta}_i$ での $x_n^*$ をモデルから実際に合成画像を生成することにより求め，補間によって $x_n^*(\boldsymbol{\theta})$ を計算する．しかし手指は非剛体であるためにその形は変形するので，実際の観測 $x_n$ は $x_n^*$ からの誤差を生じる．$x_n$ は $x_n^*$ の周りで正規分布すると仮定して，$x_n$ の確率密度分布 $f(x_n)$ を，

$$f(x_n|\boldsymbol{\theta}) = N(x_n^*(\boldsymbol{\theta}), \sigma_n^2) \tag{2.8}$$

と表現する (図 2.22中の細線)．ここで $\sigma_n^2$ は観測の種類ごとに与えられる定数である．以上

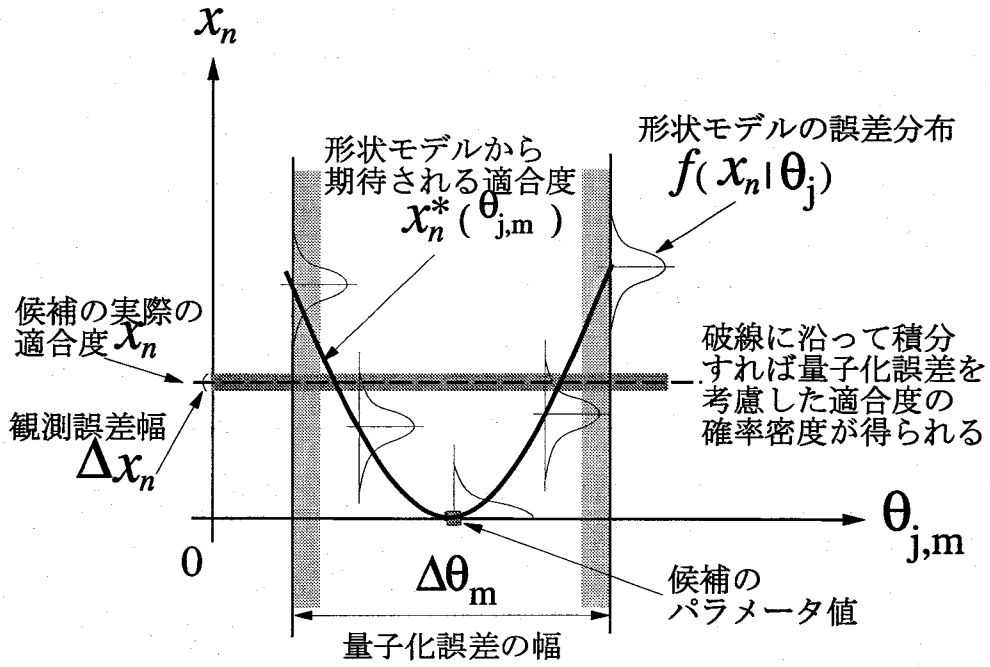



図 **2.22**: 適合度 $x_n$ の確率的評価

の議論から，離散空間における観測 $x_n$ の確率密度分布は

$$p(x_n|\boldsymbol{\theta}_j) = \frac{1}{\prod_m \Delta\theta_m} \int_I f(x_n|\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{2.9}$$

と表現される．観測誤差幅 $\Delta x_n$ が微小であれば，候補 $\boldsymbol{\theta}_j$ の適合度が区間 $[x_n - \Delta x_n/2, x_n + \Delta x_n/2]$ に入る確率は

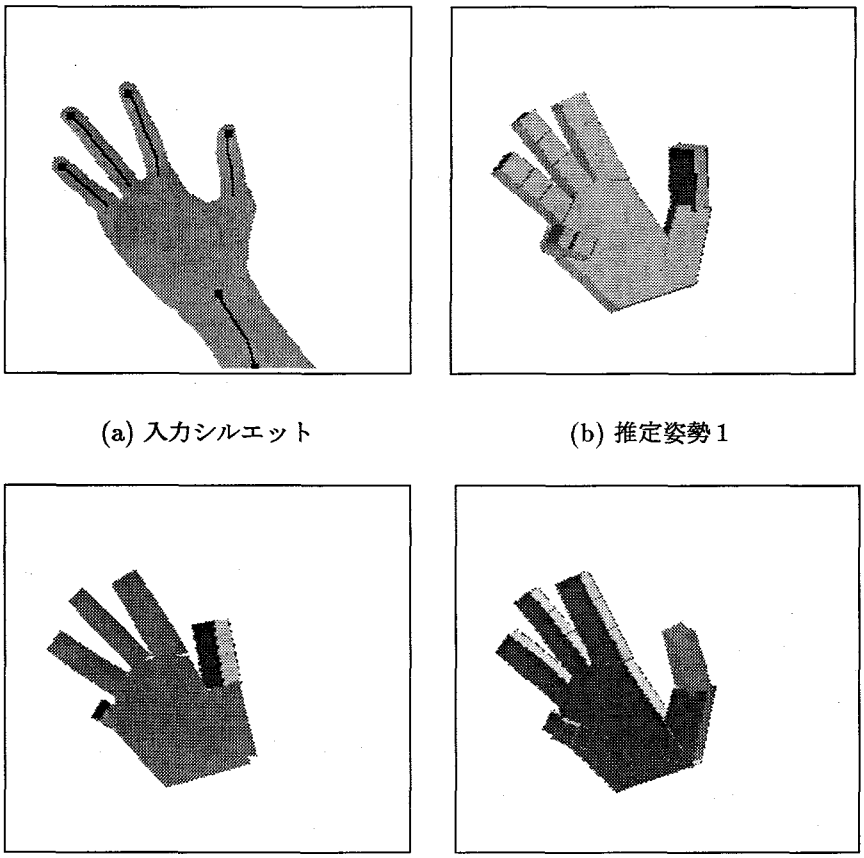
$$P(x_n|\boldsymbol{\theta}_j) = p(x_n|\boldsymbol{\theta}_j) \cdot \Delta x_n \tag{2.10}$$

とできる．

ここで，おのおのの $x_n$ は厳密に独立ではない (例えば掌と適合度と指の適合度) が，もともと適合度の高いものを選択しているのであるからずれの観測自体は十分に小さく，独立とみなしても不都合はない．そこで観測 $\boldsymbol{x} = \{x_n\}$ の下での候補 $\boldsymbol{\theta}_j$ の条件付き確率を以下のように近似する．

$$\begin{aligned} P(\boldsymbol{\theta}_j|\boldsymbol{x}) &= \frac{P_b(\boldsymbol{\theta}_j) \prod_n P(x_n|\boldsymbol{\theta}_j)}{\sum_j [P_b(\boldsymbol{\theta}_j) \prod_n P(x_n|\boldsymbol{\theta}_j)]} \\ &= \frac{P_b(\boldsymbol{\theta}_j) \prod_n p(x_n|\boldsymbol{\theta}_j)}{\sum_j [P_b(\boldsymbol{\theta}_j) \prod_n p(x_n|\boldsymbol{\theta}_j)]}. \end{aligned} \tag{2.11}$$

ここで $P_b(\boldsymbol{\theta}_j)$ は式 2.4で与えられる候補 $\boldsymbol{\theta}_j$ の事前確率である．



(a) 入力シルエット　(b) 推定姿勢 1

(c) 推定姿勢 2　(d) 推定姿勢 3

図 **2.23**: 複数の解釈が成立する場合

実際の計算においては分母の総和は手指各部の適合度の高かった候補に対してのみ行う．$P(\boldsymbol{\theta}_j|\boldsymbol{x})$ を手指候補の総体的な適合度とし，複数の推定候補が得られた場合はこの値を用いて各候補の順位付けを行う．

## 2.7 複数の姿勢候補保持とビーム探索法

生成される手指姿勢候補のうち，入力画像に対する適合度が最大の姿勢候補がその時点での最良の姿勢候補である．しかし図 2.23に示すように，入力シルエット形状の曖昧さによって複数の解釈が成立する場合や，オクルージョンやモデルと実際の手指の形状の違いからくる解釈の誤りが生じた場合には，その後の入力を考慮すると必ずしも最良とはいえないことがある．この場合，その時点での最良の候補しか考慮しないと，一時的に推定を誤っただけで次の時刻の推定ができなくなる（図 2.24(a)）．

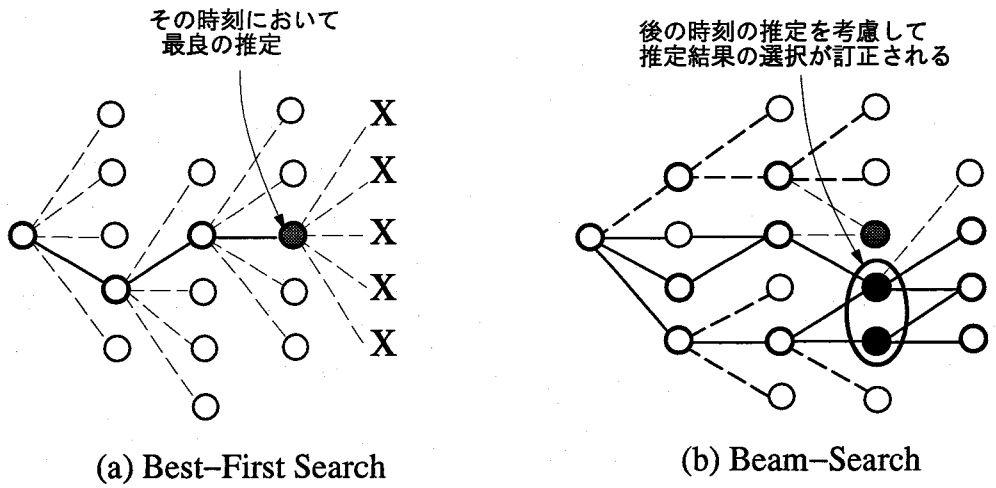



図 **2.24**: ビーム探索の利用

ところがこのような場合でも，通常捨てられてしまう二番目，三番目，…に適合度の高い候補の中から，より適した候補が見つかることが多い．そこで本手法では，時間方向に渡って広域的に最適な推定を得るために，各時刻で適合度 $P(\boldsymbol{\theta}_j|\boldsymbol{x})$ の高い順に一定個数の候補を一定時間保存しておくビーム探索法を用いる．このようにすると，あいまいなシルエットに対して手ぶりを複数通りに解釈することができる（図 2.24(b) の実線）．このときある姿勢候補がその前の時刻のどの候補から遷移したかを考えて，時間的な遷移の鎖構造を保存しておく．

さて，2.5.3節で述べたように，候補の生成は予測に基づく事前確率順に行なうが，過去における推定姿勢が複数個あれば予測も複数得られる (図 2.25)．それぞれの予測に対して動きに基づく候補の生成を行なうと，探索範囲が広くなってしまい現実的ではない．そこで2.5.3節で述べた手法を以下のように修正する．

まず，線形予測を用いて予測を行なうが，この場合

時刻 $k-2$,$k-1$ での保持されている推定の組合せの数だけ予測が生成できる．しかし，あまり大きな変位を生じる組合せについてはほとんど起きないと考えてよい．したがって，時刻 $k-1$ における推定姿勢が，$k-2$ での推定姿勢の近傍に入らないような組合せは除外しておく．生成された $i$ 番目の予測姿勢 $\widehat{\boldsymbol{\theta}}_i$ の姿勢パラメータ空間における近傍を $V_i$ とし，$V_i$ の和集合を 2.5.3節で述べた候補の探索領域 $V$ として用いる．

さて，姿勢候補 $\boldsymbol{\theta}_j \in V$ の予測姿勢 $\widehat{\boldsymbol{\theta}}_i$ に基づく事前確率 $P_t(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\theta}}_i)$ は式 2.4の定義と同様

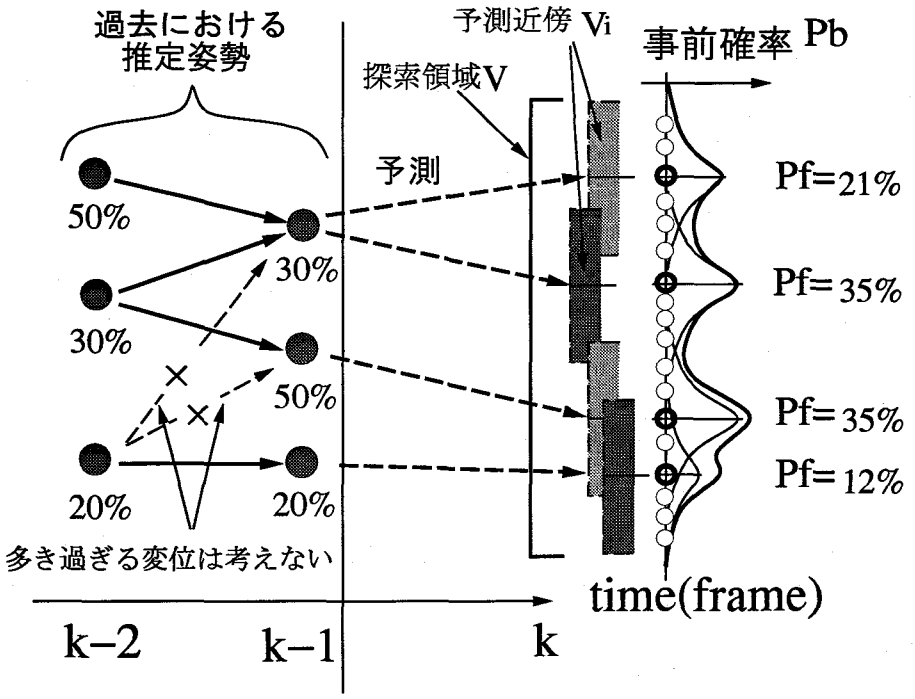



図 2.25: 予測に基づく事前確率

の議論により

$$P_t(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\theta}}_i) = \begin{cases} N(\widetilde{\boldsymbol{\theta}}_i, \Lambda_t) & \cdots \boldsymbol{\theta}_j \in V_i \\ 0 & \cdots \boldsymbol{\theta}_j \not\in V_i \end{cases} \tag{2.12}$$

と定義できる．

時刻 $k$ では時刻 $k-2, k-1$ における各推定候補とその確率が式2.11によってすでに得られているので，予測 $\widehat{\boldsymbol{\theta}}$ の事前確率 $P_f(\widehat{\boldsymbol{\theta}})$ を予測の根拠になった時刻 $k-2, k-1$ での推定候補の確率の積として定義する．すると，姿勢候補 $\boldsymbol{\theta}_j \in V$ の事前確率は，複数得られているそれぞれの予測に基づく事前確率の重ね合わせであるから，

$$P_b(\boldsymbol{\theta}_j) = \sum_i P_f(\widehat{\boldsymbol{\theta}}_i) P_t(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\theta}}_i) \tag{2.13}$$

と書ける．上式の $P_b$ を用いて式2.11を計算する．

このように複数の候補を保持する探索法と評価法を用いることで，その時点では最良と思われた推定が後の時系列を観測した結果誤りであった場合でも，追跡に失敗することなく，かつバックトラックを必要とせずに姿勢を推定しつづけることができる．また現時刻の推定結果を考慮して過去の推定姿勢を訂正することができる．

## 2.8 実験結果

本手法の有効性を確かめるために，実際の手指画像列に対して姿勢推定を試みた．

2.26,2.27,2.28に時系列に対して本手法を適用して姿勢推定を行なった例を示す．画像のサンプリング時間は0.1秒である．なお撮影した手は右手であることがわかっているとして，探索対象候補を限定している．

図2.26は手首を回転させた例である．動きはじめにおいてはほとんど入力シルエットに変化がないため，最初の段階では回転方向があいまいであり，推定結果が複数の可能性に分岐するが，やがてシルエットが変形するにつれ，腕軸まわりに回転していることがわかる．

図2.27は各指を一斉に曲げた例である．指がシルエットの突起特徴として観測できる間はその特徴を利用し，掌と重なりあって観測できなくなるとシルエットのはみ出し量を利用して姿勢推定を行っている．本手法を，仮想現実中の物体を掴んだり離すことの指示に利用できることが，この例からわかる．

図2.28では第2，第3フレームにおいて急にオクルージョンが生じ，輪郭として観測されない指が発生している．本手法では図2.21ですでに示したように，突起状の特徴と指モデルとの対応を解決しながら指の姿勢を推定する．そのため，このように唐突に画像上から消失する指があっても姿勢推定が可能である．

最後に推定の安定性に対する，複数の候補を保持する手法の効果を示す．図2.28の第4フレームにおいて，誤った姿勢推定結果が得られている．これは過去の動きから現在の姿勢を予測して探索を行うためであり，この候補がこの時点の入力に対しては最良の評価値をもつ候補として選択された．従来手法のように各時刻で唯一の候補しか保存しない場合には，このように動き予測を誤った時点で次の時刻の姿勢推定に失敗してしまう．これに対し，本手法では，図2.29に示すように，各時刻である一定以上の可能性のある姿勢推定候補を複数保持している．第4フレームで得られる姿勢推定候補は3つ存在し，第2，第3の候補は正しい推定姿勢である．したがって第2，第3の推定候補からの動きを考慮して探索することで，第5フレーム以後も推定をつづけることができた．本手法ではビーム探索法を用いているので，一時的に推定を誤った場合でもバックトラックを必要としないが，本手法では図2.30のように一定時間過去のビーム探索木を保存しているので，現時刻の推定結果から過去の推定をバックトラックして，過去の正しい推定姿勢を得ることも可能である．図2.28の場合，第4フレームでの正しい推定は後のフレームにおける推定姿勢から探索木をバックトラックすると図2.29(c)であることがわかる．

本手法の実験はSun SPARC20上で行なわれた．あらかじめサンプリングしてある画像に対して特徴抽出処理（実行時間約10秒／フレーム）を行ない，その結果を利用して姿勢

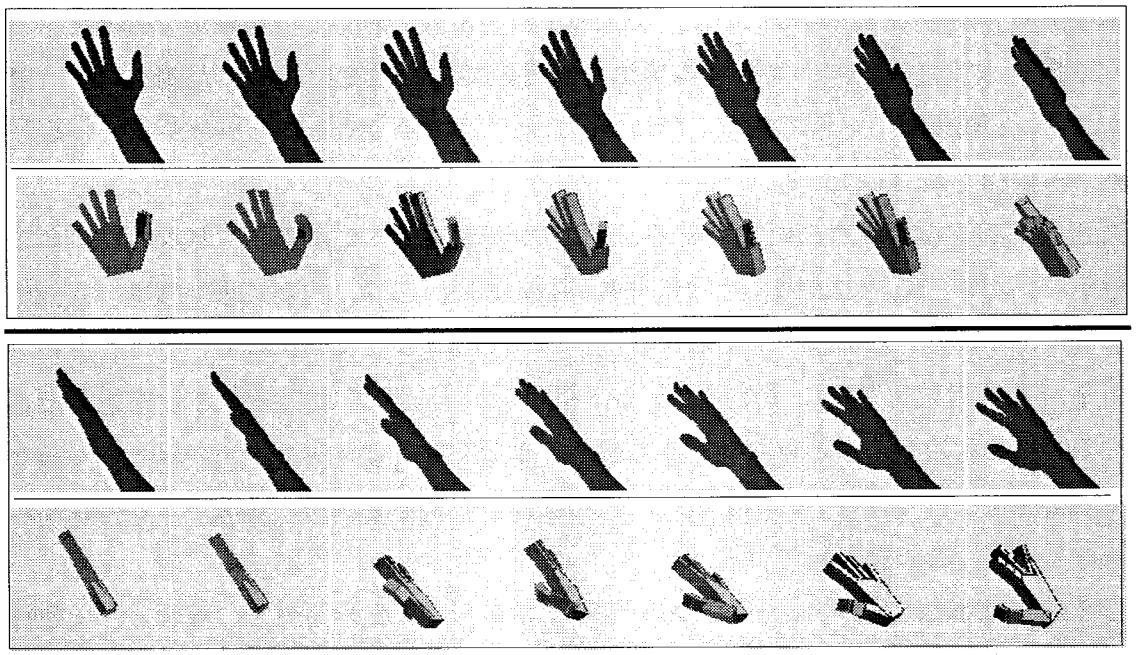


図 **2.26**: 時系列に対する実験例 (1)

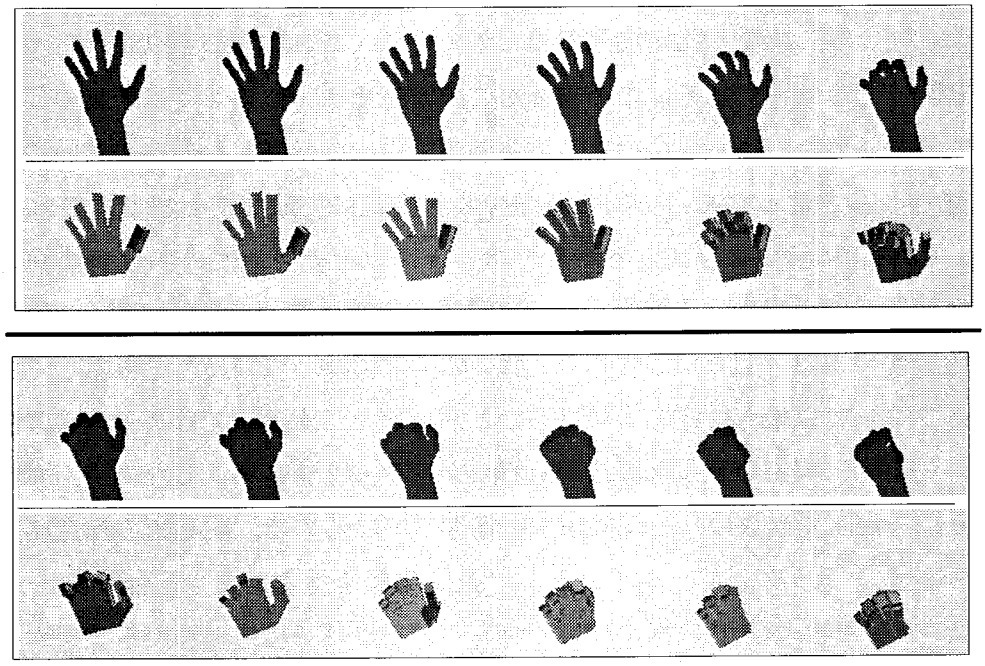


図 **2.27**: 時系列に対する実験例 (2)

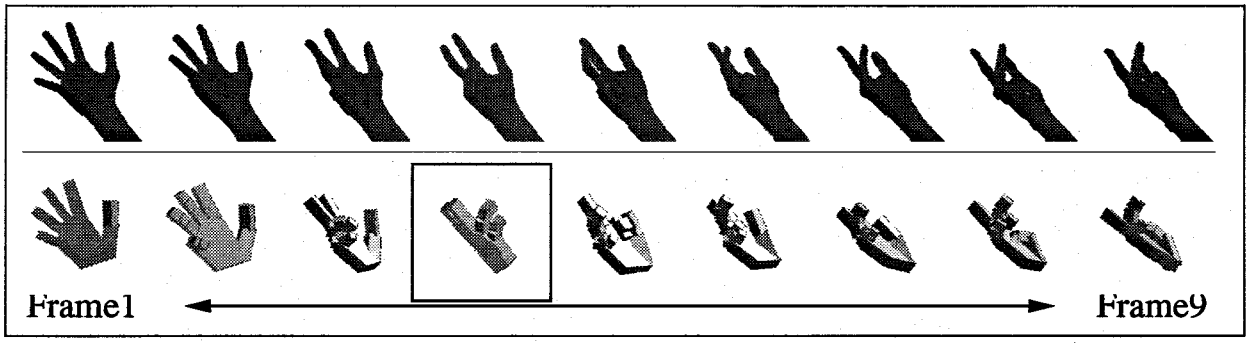
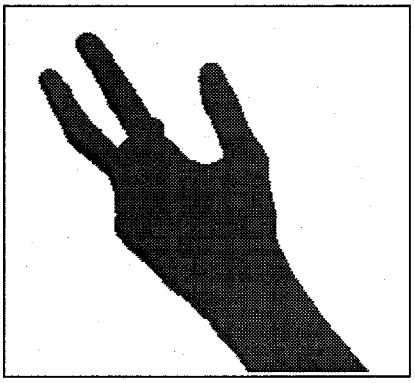
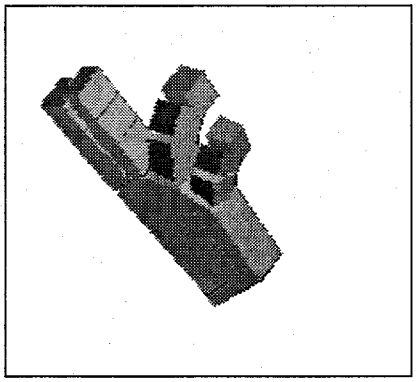


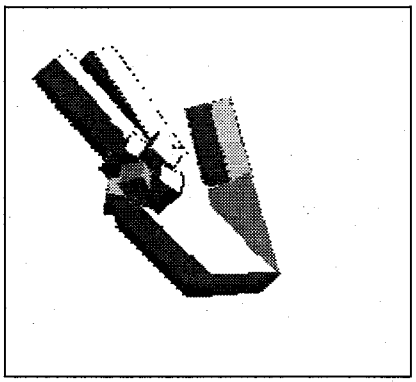

図 **2.28**: 時系列に対する実験例 (3)



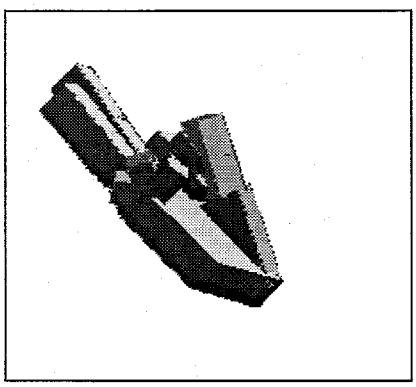
(a) 入力シルエット



(b) 第 1 の候補



(c) 第 2 の候補



(d) 第 3 の候補

図 **2.29**: 実験例 (3) 第 4 フレームにおける推定姿勢

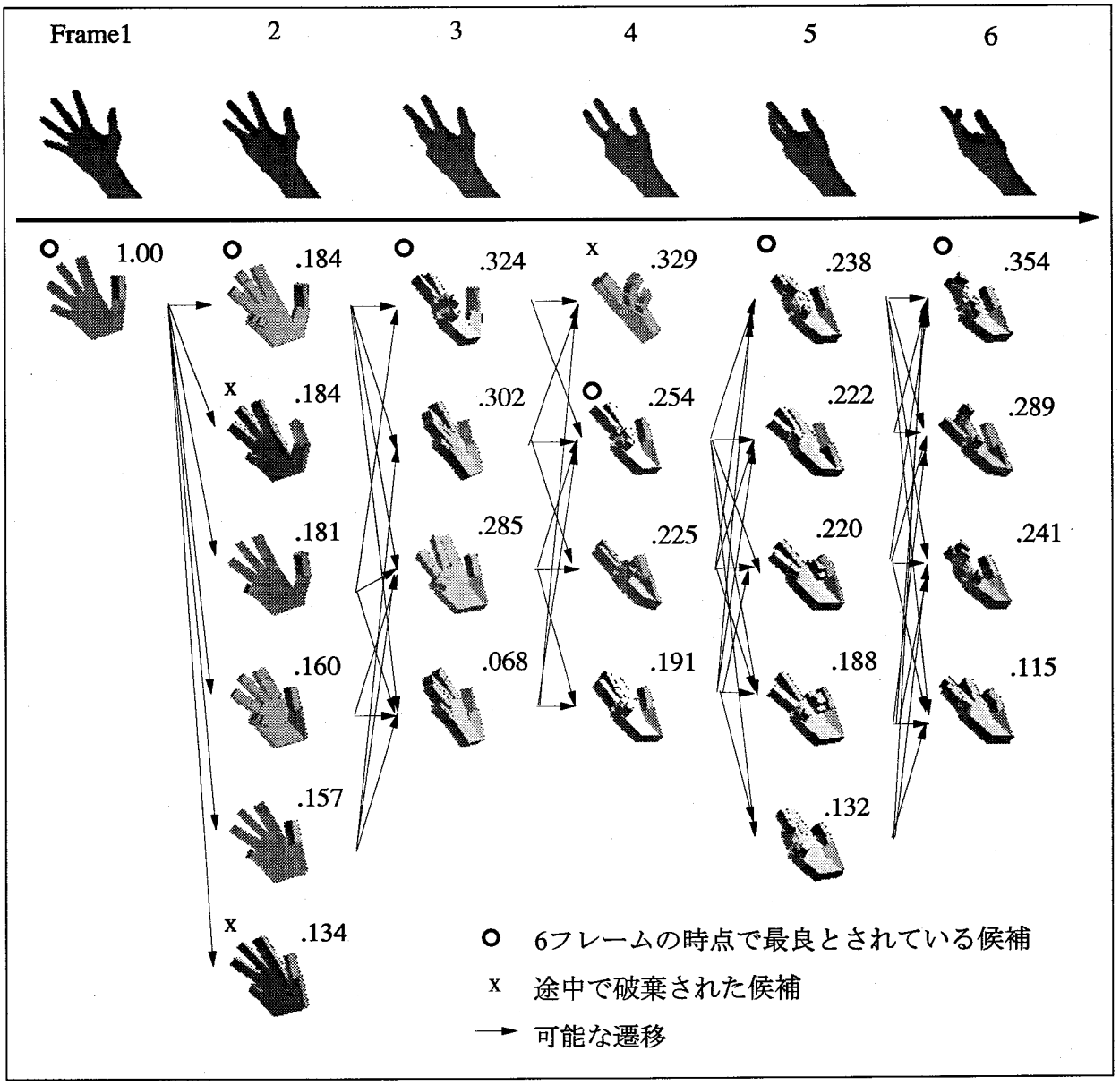



図 **2.30**: 実験例 (3) のビーム探索木

候補の探索，評価の処理を行なった．処理時間を見ると候補の探索と評価に総計で平均約20-30秒／フレームかかっている．とくに候補の評価プロセスでは，適合度を統合する時に理想的な観測量を見積もるため，実際にモデルを用いて合成画像いくつも描画し，スキャンするという多量の時間を要する処理を行なう．図2.28の実験例(3)のように，予測がはずれた場合には，非常に多くの姿勢候補を探索することになり，1フレームに3分近くかかる場合もあった．

特徴抽出処理では，画像処理専用ハードウェアの活用が望ましい．また，現状のプログラムではとくに並列計算は行っていないが，生成された各候補は並列的に評価できるので，この点を考慮すれば計算効率の向上が期待できる．

## 2.9 まとめ

本章では単眼視時系列シルエット画像を用いた手指の三次元姿勢推定法について述べた．本手法は考えられる手指の姿勢候補群を三次元形状モデルを利用して生成し，その投影が入力シルエットにもっとも適合するものを探索する手法に基づいている．広い探索範囲を効率的に探索するために，まず関節角パラメータの変動がもたらす投影画像の変形度合を考慮して探索空間を量子化した．次に動き情報による姿勢予測に基づいた事前確率を各候補に対して求め，事前確率の高い候補から順に姿勢候補を探索した．さらにシルエットの広域な形状特徴を利用して探索範囲を限定した上で，入力画像に対する候補の適合度を評価した．候補の評価においては，各部位ごとに形状の特徴を考慮して適切な適合度評価を行った．このとき，入力シルエット特徴とモデルの対応を可能性の高い順に仮定し検証することにより，対応づけの曖昧性を解決した．また手指の各部の適合度を統合するために各適合度を確率として計算した．さらに，長い時系列に渡って最適な推定を得るためにおのおのの時刻において複数の推定候補を保持しつつ探索を行った．オクルージョンのためにシルエットの解釈が曖昧な場合に対して，本章で述べた手法はロバストに推定を行えることを実験例を用いて示した．

本研究では掌の形状が板状であるという性質や指らしい形状特徴として突起状の領域に着目し，候補の探索，評価に利用しているが，シルエットが簡単に抽出できない複雑な背景の場合にはこれらは直接利用できない．シルエットが得られても，手指だけではなくより一般的な関節構造物の姿勢推定をする場合には，これらの前提は必ずしも用いることができない．したがってより一般の関節構造物体の姿勢推定に本手法を適用するためには，一般の物体の三次元形状のうちどのような形状が姿勢推定の手がかりになりうるか，またどのような規範を用いて評価するかを検討する必要がある．

また，手指に限定した場合でも，形状モデルの誤差や計算時間短縮のために探索幅を広くとるため，大まかな姿勢の推定しか得ることができない．したがってえられた大まかな姿勢を出発点にしてより正確に推定を修正することが考えられる．これについては次章で論ずる．

インターフェースとしての手指推定を考えた場合，実時間での動作は重要である．しかし現状では本手法を一台の汎用計算機上でシミュレートした段階である．本手法では特徴の抽出や処理候補の適合度を評価するときには画像をスキャンするため，時間がかかってしまう．しかし本手法は姿勢候補をつぎつぎと生成しては適合度を評価していくので処理の並列性が高い．この点をいかせばより高速な処理が可能になると考える．これらが今後の課題である．

# 第 3 章

# 関節物体の姿勢と形状の同時推定の原理

本章と３，４章では一般の関節物体について単眼視動画像から形状と姿勢の双方を同時に推定する手法について論じる．まず本章ではまず関連研究と本研究の関わりについて論じた後，本研究における形状と姿勢の推定原理について論じる．

## 3.1 関連研究と本研究の関わり

前章では手指のモデルを用いてその３次元姿勢を推定した．モデルには関節がどうつながっているかという構造の知識と，関節でつながれる各部の長さや太さ，表面の起伏といった形状の知識が含まれている．前章では両方の知識があたえられたもとで，手指の姿勢を各関節の角度として推定した．

前章で論じたシルエットマッチングの手法では，あらかじめ与えられるモデルの形状と実際の手指の形状の間の誤差のために正確に関節角度を推定できない．しかし，はじめは大雑把な形状と姿勢を与えておき，動画を次々と入力してやることによって，その形状を次第に正確に修正すれば，姿勢の推定もより正確にできる．

画像から形状と姿勢を復元する場合には，画像中の領域が物体のどこに対応するかを決定しなくてはならないが，ここでは特徴追跡をによる手法や [8]，前章で論じたシルエットマッチングの手法を用いて対応づけを解決しておく．対応が得られたもとで形状と姿勢を推定するには最小二乗法やフィルタリングの手法が良く用いられる．これらの手法に基づいて [9][24] は複数のカメラからの画像や距離データから推定を行なっており，また単眼視動画像を用いた例には [10][8] がある．しかし単眼視では奥行きがわからないので，これらの研究では物体の形状（長さや太さの比）は既知としている．

単眼視動画像から姿勢と物体の形状を同時に推定する手法として Factorization Method がある [25]．この手法を拡張した [26] は，異なる運動をする複数の物体が混在した場合でも

形状と姿勢を推定ができるが，特異値分解を利用するため，多くのビュー間で多くの観測点の対応を与える必要がある．

単眼視動画像から姿勢と形状が一意に解けるための必要条件については従来から精力的に研究，整理されてきた [27][28][29]．これらの多くは未知の一般的な物体を対象にしているので，剛体性以外の対象に関する知識は仮定されていない．このような場合，十分な観測点がなければ距離に関して曖昧性が生じてしまう（不可観測）．したがって多くの従来研究は推定に十分な観測が得られる可観測なシステムのみを対象にしてきた．

しかし我々は多くの場合良く見知った物体を観察するので，その物体の形状や構造に対する大まかな知識を推定に利用でき，曖昧性を解消できる可能性がある．このような不可観測なシステムにおける推定問題への取り組みはあまりなされていないが，[30] が関節物体の各部の長さや太さ，各関節角度のとりうる範囲を，それらに関する制約条件を用いて形状と姿勢を推定する手法を提案している．この手法では観測から得られた姿勢と形状の各パラメータの最大値と最小値を，制約伝播ネットワークを通して相互に伝播しあう．したがって，直接制約されていないパラメータに他のパラメータを通じて間接的に情報が伝播するには２以上のステップが必要になる．また収束するまでには伝播が繰り返されなければならない．

本研究では関節物体の姿勢と形状を表すパラメータ空間を考え，考えられる姿勢と形状をその空間中でのパラメータの領域で表現する手法を用いることを提案する．観測を満たすパラメータ領域，制約を満たす領域，初期パラメータ領域の３つ領域の積集合をとることで推定を行なう．この手法では観測から得られる情報と知識として与えられる制約情報をまったく同じに扱うことができ，また両者の統合の際に制約伝播のような繰り返しの必要がない．入力が単眼視の動画像であっても，さまざまな姿勢の画像を入力することにより，姿勢と形状の範囲の両方をしだいに限定していくことができる．

しかし，関節物体を画像で観測する場合，観測方程式は形状や姿勢のパラメータに関して一般に非線形であるので，扱いが困難である．当初本研究ではパラメータ領域を確率分布としてとらえ，観測を局所線形化する拡張カルマンフィルタ [31] を用いた．この方法では，異なるさまざまな観測が入力される場合には満足できる結果が得られたが，類似の観測が長い時間に渡って入力された場合は，奥行きの曖昧性と非線形性に起因する推定誤りが発生することが明らかになった (４章で詳述)．

そこで，本研究ではパラメータ領域を確率分布ではなく

1. 楕円体の内部領域

2. 平行な超平面で囲まれた領域

として近似する手法 (セットメンバシップ法 [32]) を用いる．この手法に基づく剛体物体の姿勢推定にはいくつかの従来研究 [33][34] があるが，いずれも観測が線形かつ可観測な場合に対して適用したものであった．またセットメンバシップ法は，状態パラメータの次元が高くなると指数オーダで計算量が増す欠点があるため，従来研究では少ない個数のパラメータ推定にしか利用されていなかった．これに対し，本研究では推定パラメータの数が多く観測が非線形かつ不可観測の場合についてこの手法を適用する方法を提案し，その有効性を示す．

まず観測の非線形性を補償するために，パラメータ空間の各点における非線形性の度合を各フレームでの推定の際に見積もりながら，EKF のような誤りが起きない統合を行う．また推定パラメータをモード分解することによって計算量を削減する．さらに，観測の非線形性によりとりうるパラメータ領域がパラメータ空間中で複数の領域に分かれる場合があるので，本研究では複数の楕円体を用いることで適切に取り扱う．

## 3.2 形状と姿勢に関する制約知識の利用による奥行きの曖昧性の除去

本研究では関節構造が既知である関節物体を単眼視動画像で観測した場合に，その形状と姿勢を推定することを考える．図 3.1のような関節物体の形状と姿勢はリンク長パラメータ $r$ と関節角度パラメータ $\omega,\theta$ によって記述できるとし，以後これらのパラメータを推定することを考える．なお本研究では投影は正射影を仮定する．

人体や手指のシルエット画像からは，突起上の領域とその中心軸が特徴として得られることが多いので，図 3.1に示すように線分が観測できるとする．すると単眼視画像からは奥行きが観測できないため，図 3.2に示すように，線分の３次元的傾きや長さは一意に決定できない．動画像からの複数の入力画像を用いた場合でも，リンクの３次元長さがわかっていないので，たとえば，

1. 長いリンクが画面に垂直に近い状態で小さい角度動いた場合

2. 短いリンクが画面に平行な状態で大きな角度動いた場合

の２つはやはり区別できない．

しかし，人体のような物体を考えると，形状と姿勢のパラメータは完全に独立ではない．各関節には可動範囲があり，画像中の物体の大きさもある程度わかっていることが多い．また，いくつかの関節角度はともなって変化し，各部の長さや太さといった形状パラメータには相関関係がある．関節物体の各部分を個別に考えると曖昧性が存在するが，各部分の関係

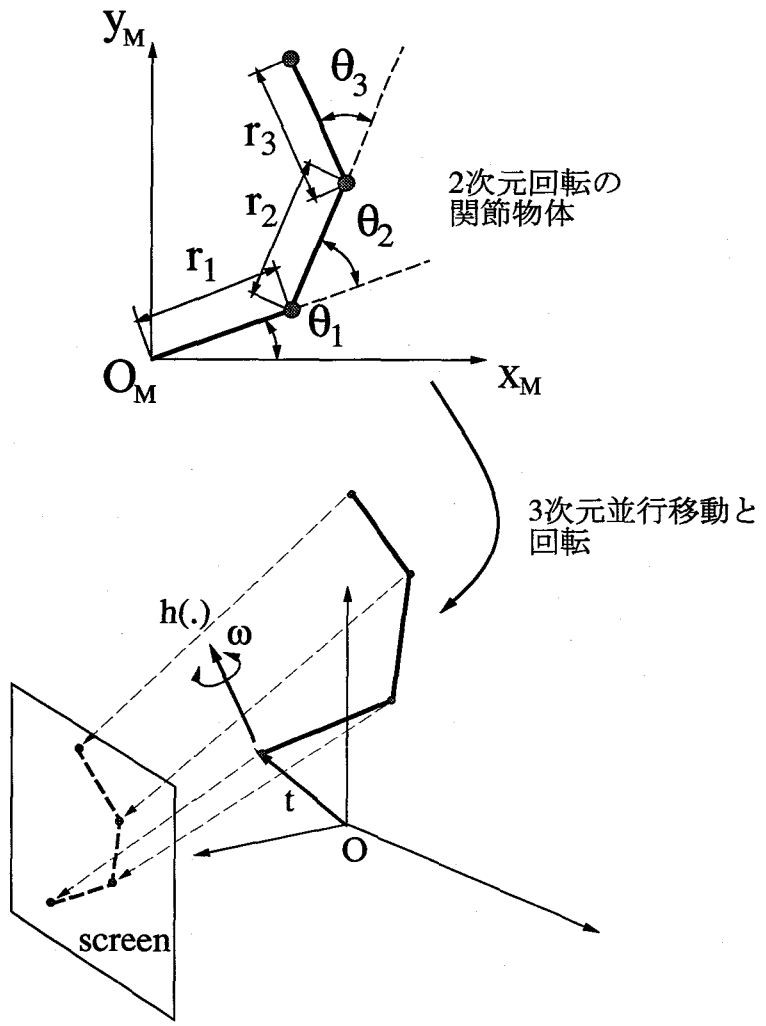



図 **3.1**: 直線リンクで構成された関節物体

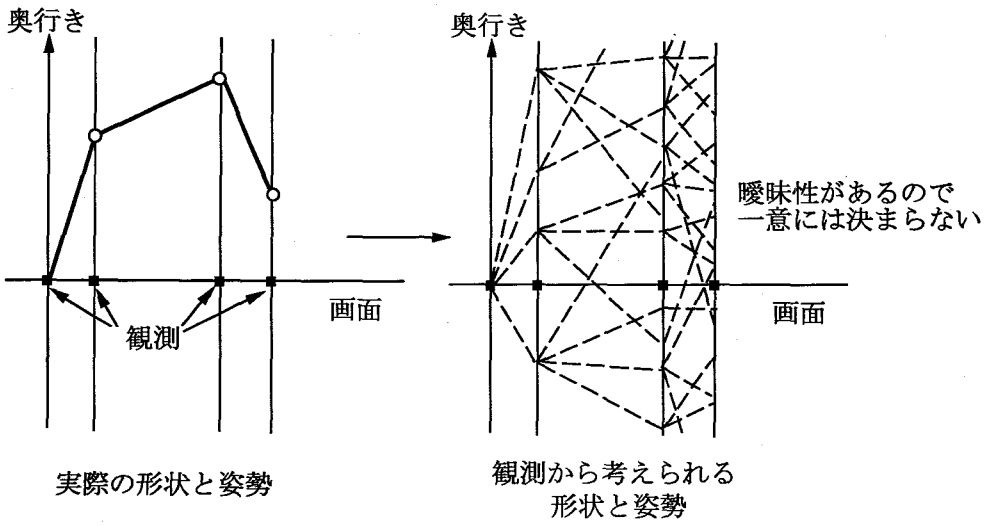



図 **3.2**: 単眼視における奥行きの曖昧性

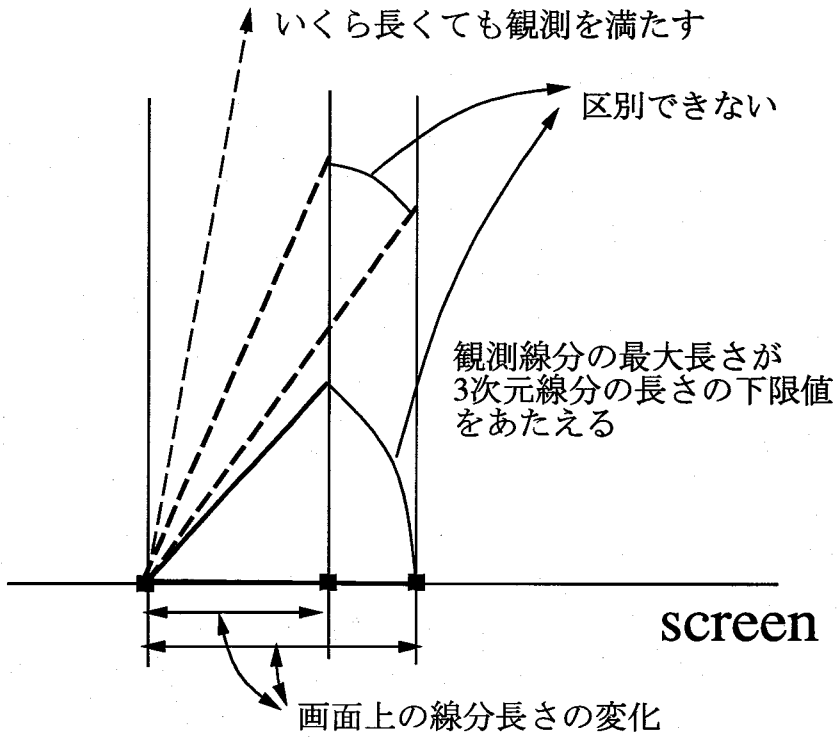



図 **3.3**: リンク長の不変性があたえる幾何的制約

を考慮すれば，それぞれの部分の観測が他の部分のパラメータを拘束することによって，可能な形状と姿勢が限定できる可能性がある．そこで，本研究ではこれらの制約知識を形状と姿勢の推定に利用することを考える．

人体のような関節物体についての制約知識は以下の４つに整理される．

1. 各関節角度は時間的に連続に変化する
2. 各リンクの長さは時間的に不変である．
3. 各時刻における各関節角度は互いに相関をもつ．
4. 各リンクの長さは互いに相関をもつ．

制約を利用した形状と姿勢パラメータ範囲の限定のされ方に関して，幾何的な考察から以下のことがわかる．

**考察 1** リンク長が不変な場合は，観測された線分の長さはリンク長の下限値を与える．

これは図 3.3に示すように，正射影では線分がその３次元長さよりも短く観測されることがないことから明らかである．ところが線分の３次元長さが観測された長さよりも長い場合は線分を画面に対して傾けた姿勢が考えられるので，長さの上限は決まらない．

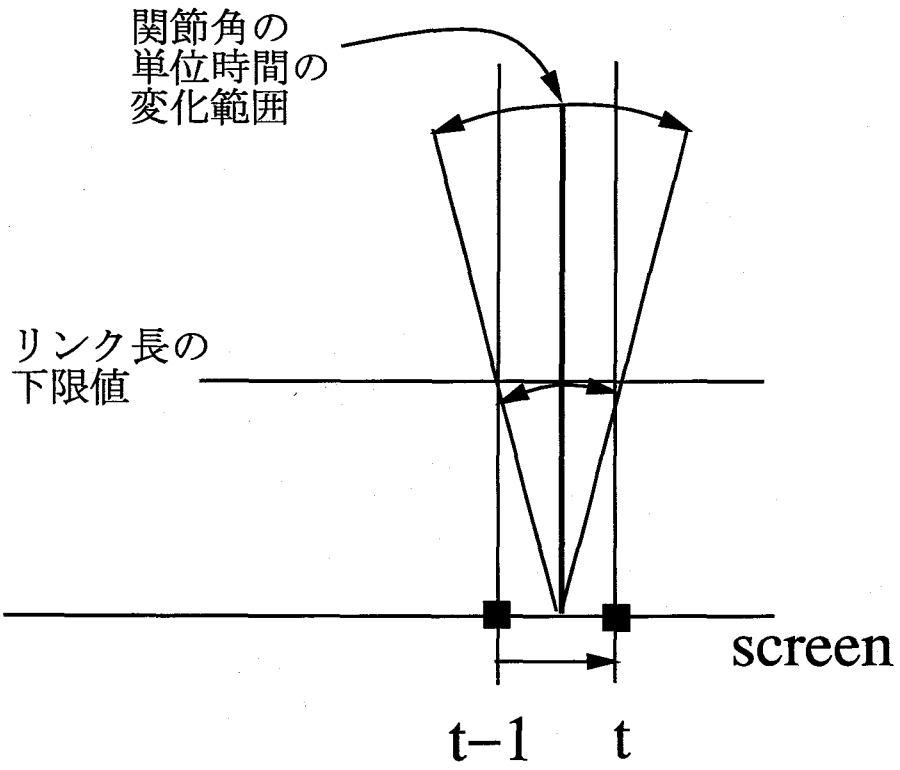



図 **3.4**: 関節角変化の上限があたえる幾何的制約 (1)

**考察 2** 単位時間あたりの関節角の変化に上限値を設定する．このとき得られた観測値の変化からリンク長の下限値が決まる．

説明図を 3.4に示す．リンクの関節角度の変化が一定ならば，画面上で可能な観測の最大変化量はリンク長が長いほど大きい．そしてその最大変化はリンクが画面に垂直になったとき生じる．したがって，観測の変化が得られるとその変化を生じうる最小のリンク長が求められる．

さらに，もしリンク長の上限値が初期パラメータ領域によってあたえられていた場合は次のことがわかる．

**考察 3** 考察 2 と同様に単位時間あたりの関節角の変化に上限値を設定する．同時にリンク長の上限値が与えられたときには，得られた観測値の変化からその時刻での関節角の範囲を限定できる．

説明図を 3.5に示す．リンク長に上限がある場合には画面に対するリンクの傾きは無制限にできない．図の破線のように，リンクが画面に対して角度が浅く平行に近いときは，観測さ

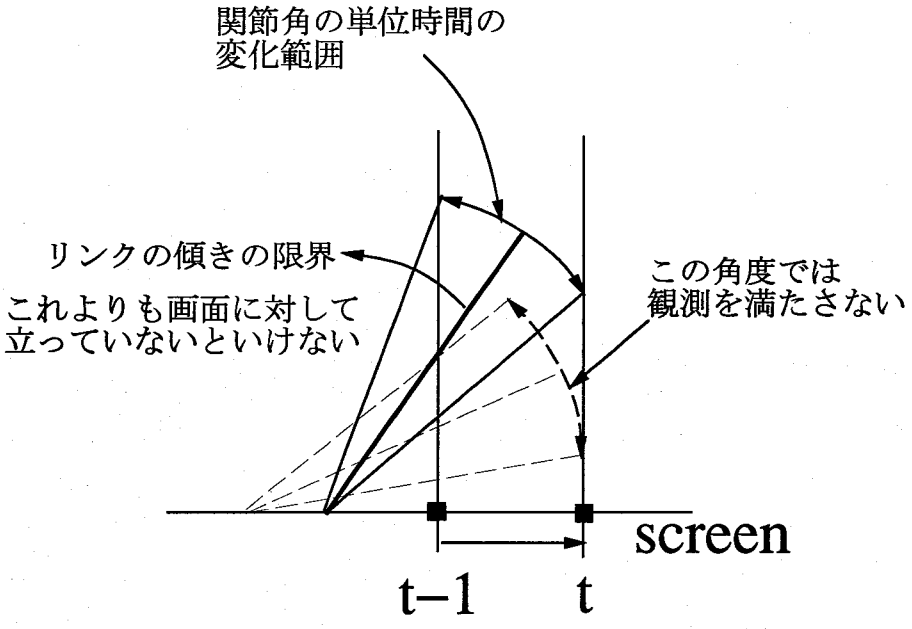



図 **3.5**: 関節角変化の上限があたえる幾何的制約 (2)

れた大きさの変化を実現できない場合がある．したがってリンクがある値以上に画面に対して傾きをもつように関節角度は制約される．

考察 1，2，3で示された幾何的制約は，すべてリンク長の下限値と画面に対するリンクの傾きの下限値を示すものであった．つまりこれらの制約のみでは「リンク長が極端に長く，リンクが画面に対して垂直に近い場合」を除去できず，初期パラメータ領域で設定されたリンク長の上限値は小さくならない．

しかし，人間の関節のように，関節角度のとりうる範囲や各関節角度どうしの相関関係を制約として導入すると以下のことがわかる．

**考察 4** 関節角の可動範囲を設定すると，状況によってはリンク長の上限値がきまる．

図 3.6のような２次元の２リンク関節物体を使って説明する．対象の関節物体が「関節角度は常に０度から 90 度の範囲にある」という制約を満たすとする．簡単のため，根元側の関節角度の範囲が図のようにわかっていたとしよう．図 3.6に黒い四角マークで示す１次元の３点が観測として得られた場合，関節角度の範囲制約を考慮すれば，末端側のリンクの状態はＡからＢまでの可能性が考えられる．ここでＢがが最大リンク長をあたえる推定である．またリンク長の下限値はＣが与える．この段階では末端側のリンクが図中上方にのびている場合と下方にのびている場合の両方が考えられる．

さて，この状況で次に図 3.7に示すような観測が得られたとする．図 3.6と同様に根元側

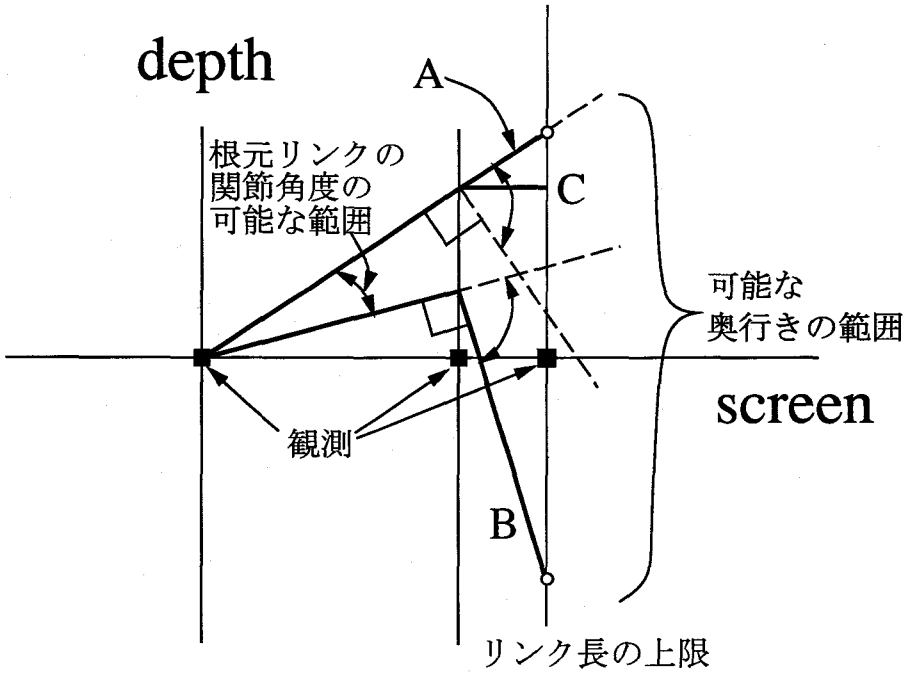



図 **3.6**: 関節角の可動範囲のあたえる幾何的制約（時刻 $t-1$ における推定１）

の関節角度の範囲が図のようにわかっていたとすると，リンク長は時間的に変化しないから，図 3.6で末端側の最大リンク長を与えたＢはＥへ写る．しかし，Ｅは関節角度の最大値 90 度を越えるのでありえない．したがってリンク長の上限値はB' が与えることになり，上限値は小さくなる．同様に，最小リンク長を与えたＡはA' に写るが，A' は長さが短過ぎて観測を満たさないから，リンク長の下限値はA' ではなくＦがあたえることになり，したがって下限値は大きくなる．

ここで再び時刻 $t-1$ での推定を考えよう．図 3.7中のＤの候補は時刻 $t-1$ では図 3.8のD' にあったことになるが，この場合関節角度は 0 度以上という制約に反する．残る可能性はD' を下に向かって対称に折り返したD" であるが，もしここで関節角度の変化の上限が与えられていると，D" → Ｄの変化は角度変化が大き過ぎるので，D" も不可能な場合がある．したがってこの場合，時刻 $t$ （図 3.7）ではリンクの存在位置はB' からＦの範囲に限定できる．

もし，複数の関節角度の間に相関関係があれば，それによって関節の可動範囲はさらに限定されるので，上記の効果はより強く生じる．また複数のリンク長に相関関係があれば，ひとつのリンク長が限定されることによって他のリンク長も限定されるから，よりパラメータの範囲は狭まる．

以上の考察から，形状と姿勢の推定にパラメータに関する制約知識を導入することで効果的に可能性を限定できることがわかる．

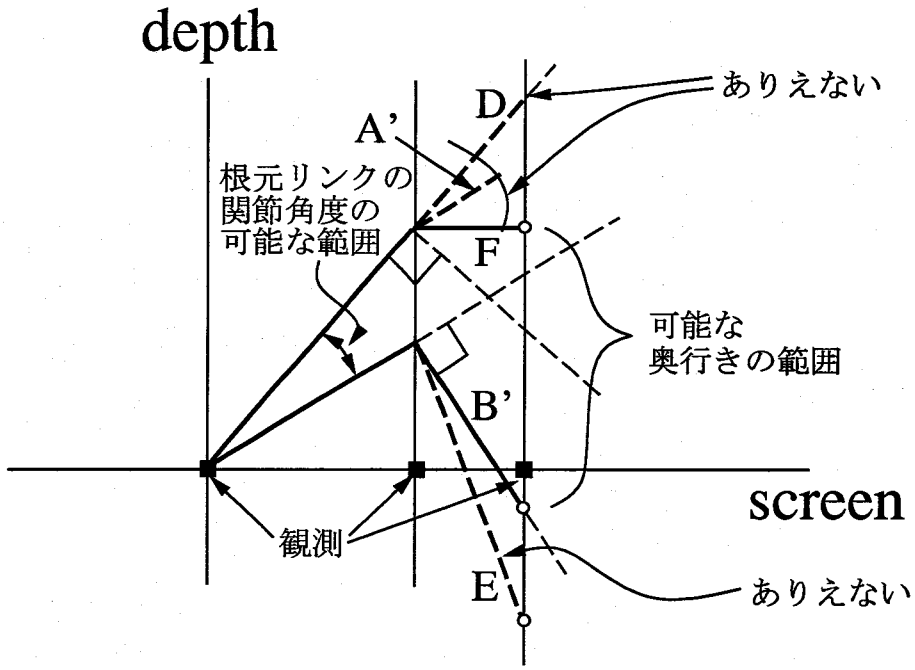



図 3.7: 関節角の可動範囲のあたえる幾何的制約（時刻 $t$ における推定）

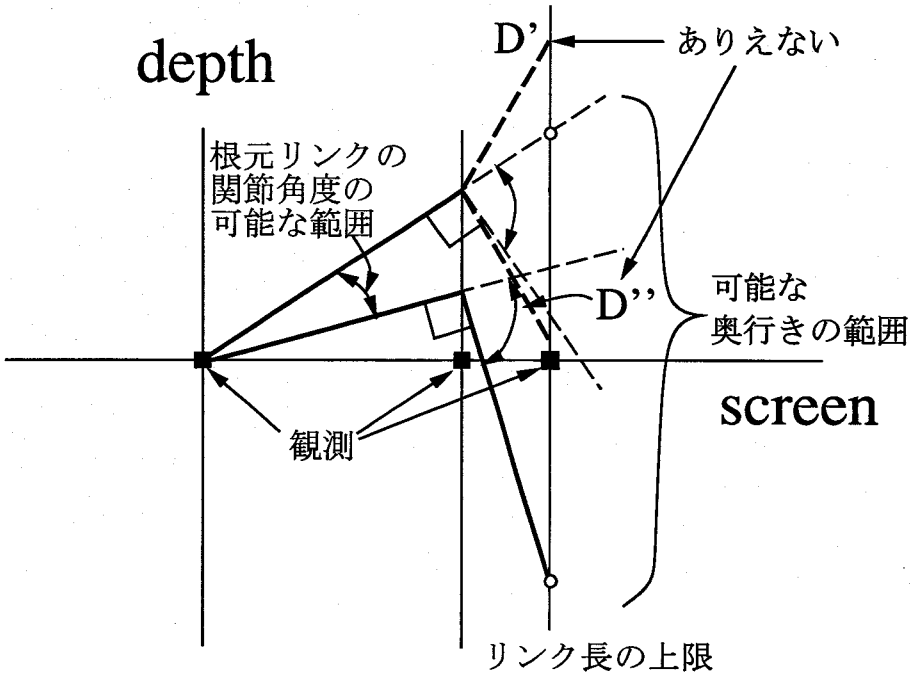



図 3.8: 関節角の可動範囲のあたえる幾何的制約（時刻 $t-1$ における推定２）

## 3.3 形状と姿勢ならびに制約知識の領域表現

前節で述べた関節物体に関する制約知識を形状と姿勢の推定に利用するためには，画像から得られる観測と制約知識とを統合するしくみが必要である．本研究では，形状と姿勢のパラメータ空間における領域表現を用いて統合する枠組を考える．

図3.1の関節物体の形状と姿勢を推定する場合について考える．物体の形状，姿勢や位置を，関節角度パラメータ $\boldsymbol{\omega}$(手首など3自由度の関節)，$\theta_i$（1自由度）とリンク長 $r_i$（ただし $i$ はリンクの番号）からなる状態ベクトル $\boldsymbol{x}$ で表すことにする．また画像からは各リンクの投影線分 $\boldsymbol{y}$ が観測されるとする．

このとき各リンクの長さと画面に対する傾きは $\boldsymbol{y}$ によって拘束を受けるので，

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) \tag{3.1}$$

と表現できる．$\boldsymbol{y}$ の次元を $m$，$\boldsymbol{x}$ の次元を $n$ とすると，一般に単眼視画像では $n > m$ なので，式3.1の解 $\boldsymbol{x}$ は $n$ 次元ユークリッド空間中の $m$ 次元多様体 $H$ を構成する (図3.9)．実際には観測はノイズによる誤差を含むから，それを考慮すると式3.1は

$$|\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{x})| \leq \Delta \boldsymbol{y} \tag{3.2}$$

と変更できる．式3.2は $H$ のまわりに厚みをもった帯状の領域 $\bar{H}$ を意味する．

さて，最初にパラメータの初期領域 $\Omega_0$ を与えると，解の領域は $H$ と $\Omega_0$ の積集合で与えられる．すると時系列にわたって，さまざまな観測 $\boldsymbol{y}_t$ から得られる領域 $\bar{H}_t$ と $\Omega_0$ との逐次的な積集合

$$\Omega_t = \Omega_{t-1} \cap \bar{H}_t \tag{3.3}$$

を求めていくことが考えられる．しかし，関節角度パラメータは時間とともに変化するので，そのダイナミクス

$$\boldsymbol{x}_t = \boldsymbol{s}(\boldsymbol{x}_{t-1}) \tag{3.4}$$

を考慮する必要がある．人体のような複雑な動きをすることのある物体について，ダイナミクスを精密にモデリングするのは困難であり，どうしても誤差が含まれる．したがってダイナミクスは，

$$|\boldsymbol{x}_t - \boldsymbol{s}(\boldsymbol{x}_{t-1})| \leq \Delta \boldsymbol{x} \tag{3.5}$$

と表現される．

するとパラメータ領域 $\Omega_{t-1}$ は次の時刻 $t$ において，

$$\tilde{\Omega}_t = \{\boldsymbol{x}_t \mid |\boldsymbol{x}_t - \boldsymbol{s}(\boldsymbol{x}_{t-1})| \leq \Delta \boldsymbol{x}, \boldsymbol{x}_{t-1} \in \Omega_{t-1}\} \tag{3.6}$$

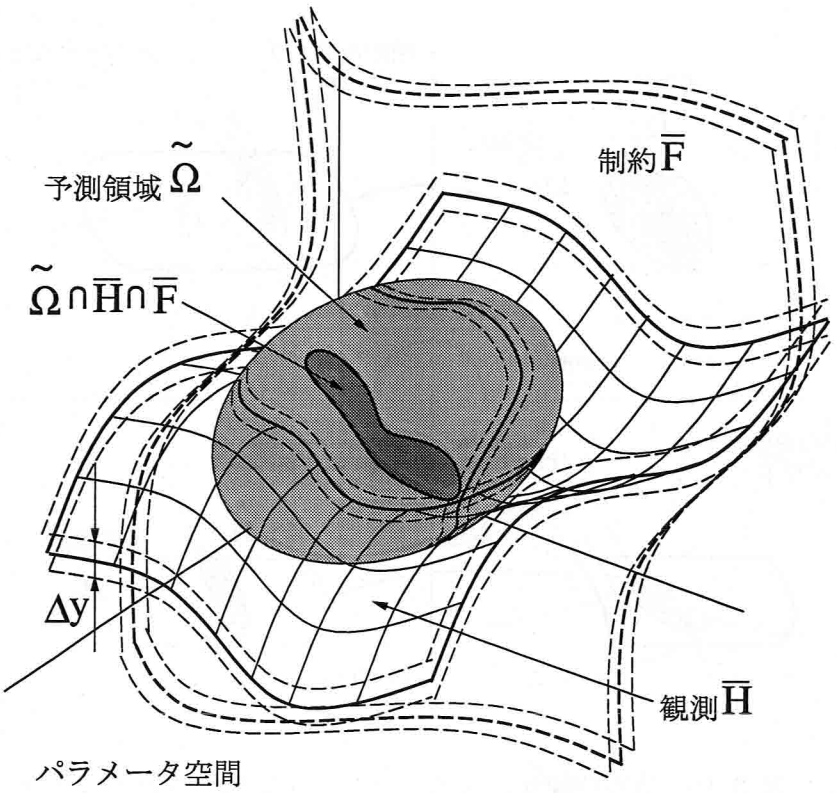



図 **3.9**: パラメータ空間における初期領域と観測による拘束の関係

なる予測領域 $\tilde{\Omega}_t$ に広がる．よって式 3.3は次のように修正される．

$$\Omega_t = \tilde{\Omega}_t \cap \bar{H}_t \tag{3.7}$$

次に，形状と姿勢のパラメータの制約として，次のような式を考える．長さ，太さなどの形状パラメータベクトルを $\boldsymbol{r}$，関節角度のパラメータベクトルを $\boldsymbol{\theta}_t$ （$t$ は時刻）として，

$$f_i(\boldsymbol{r}) = 0 \ (i = 1 \cdots p) \tag{3.8}$$

$$g_{t,j}(\boldsymbol{\theta}_t) = 0 \ (j = 1 \cdots q). \tag{3.9}$$

ただし $p$,$q$ はそれぞれの制約の個数である．上記の式は形状姿勢パラメータをパラメータ空間上の多様体 $F_t$ の上に拘束する．しかし実際の手指では上の拘束式からの誤差を生じる．その誤差を考慮して

$$|f_i(\boldsymbol{r})| \leq \Delta f_i \ (i = 1 \cdots p) \tag{3.10}$$

$$|g_{t,j}(\boldsymbol{\theta}_t)| \leq \Delta g_{t,j} \ (j = 1 \cdots q) \tag{3.11}$$

とする[1]．この式は $F_t$ のまわりに厚みをもった帯状の領域 $\bar{F}_t$ を意味する．そこで図 3.9に示

[1]場合によっては $\boldsymbol{r}$ と $\boldsymbol{\theta}$ の間の制約式を用いることもできる．

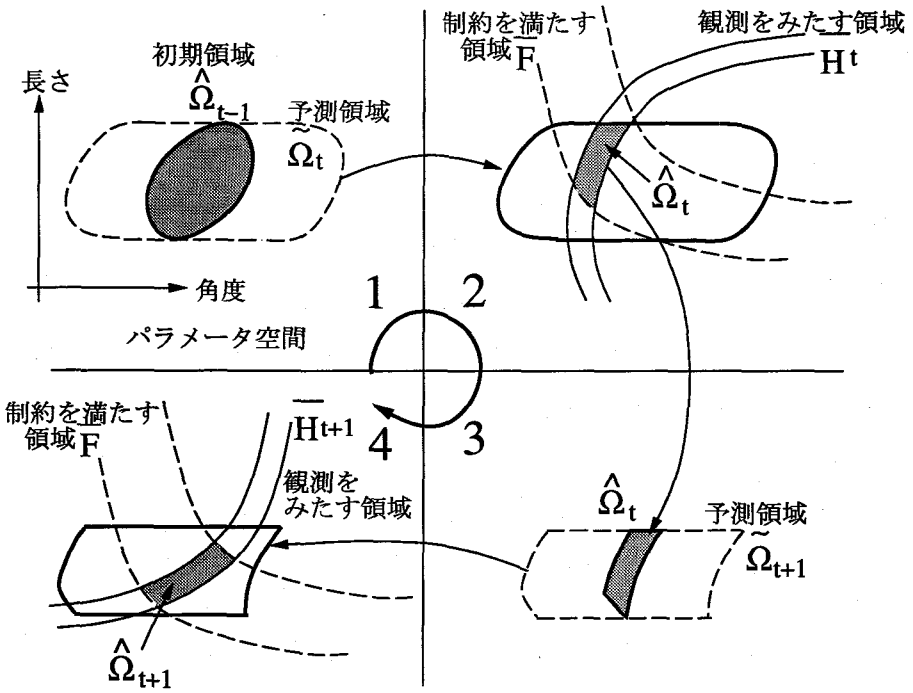



図 **3.10**: 時系列観測によるパラメータ領域の逐次的な更新

すように，予測領域 $\tilde{\Omega}_t$，観測を満たす領域 $\bar{H}$，領域 $\bar{F}_t$，の積集合をとることを考える．逐次的な表現は式 3.7を修正して

$$\Omega_t = \tilde{\Omega}_t \cap \bar{H}_t \cap \bar{F}_t \tag{3.12}$$

となる．

式 3.12を用いて姿勢と形状のパラメータ領域の更新を図 3.10に示す．初期パラメータ領域からダイナミクスを考慮した予測により可能性のある領域が広がる．この予測領域と画像から得られた観測と制約知識をともにみたす領域がその時刻の推定パラメータ領域になる．次の時刻では，この推定領域をもとにして再び予測によってパラメータ領域が広がり，制約と次の時刻の観測から領域を限定される．このように，十分さまざまな観測が入力されると，パラメータの範囲はしだいに限定されていく．

この式中の $\Omega,\bar{H},\bar{F}$ を具体的に表現するには

1. 確率分布としての表現

2. 不等式による領域としての表現

の２種類の方法がある．とくに，観測やダイナミクスの誤差をノイズ分布と考えると，確率分布としての扱い，すなわちフィルタリングを用いることが有力である．そこで次章では，カルマンフィルタを用いた式 3.12の推定手法について論じる．

# 第 4 章

# カルマンフィルタを用いた形状と姿勢の推定とその問題点

## 4.1 システムの方程式

図 3.1の関節物体について，カルマンフィルタを用いて形状と姿勢の推定を試みる．各関節は同一面内で回転し，全体が 3 次元平行移動と剛体回転をする．この関節物体の形状と姿勢の状態を各リンクの長さ $r_i$ と関節角度 $\theta_i$, 関節角速度 $\dot{\theta}_i$ ($i$ はリンクの指標），並行移動成分 $\boldsymbol{t}$, 3 次元剛体回転成分 $\boldsymbol{\omega}$ ($\boldsymbol{\omega}$ 軸まわりに $|\boldsymbol{\omega}|$ だけ回転）で表現する．

状態ベクトル $\boldsymbol{x}$ を

$$\boldsymbol{x} = (\boldsymbol{t}^T, \boldsymbol{\omega}^T, \theta_1, \cdots, \theta_m, \dot{\theta}_1, \cdots, \dot{\theta}_m, r_1, \cdots, r_n)^T \tag{4.1}$$

と定義する． $r$ は時間不変とするので $\dot{r}$ は状態に含まれない．次の時刻の状態予測に線形予測を用いれば，更新方程式は

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{u}_t \tag{4.2}$$

と表現できる． $\boldsymbol{u}_t$ は予測誤差であり，平均 0，分散 $\boldsymbol{U}$ の確率分布に従う白色雑音と仮定する． $r$ は時間不変なので $r$ に対する予測誤差の分散は 0 とする．ここで等速度予測を用いると， $\boldsymbol{A}$ は次の $(2m+n) \times (2m+n)$ の行列になる．

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{I}_n \end{bmatrix} \tag{4.3}$$

ここで $\boldsymbol{I}_m$ は $m \times m$ の単位行列である．このとき $\boldsymbol{U}$ は $\theta_i$ の加速度の範囲を仮定して決定する．

次に観測方程式を構成する．観測ベクトルを $\boldsymbol{y}_t$ とすると観測方程式は

$$\boldsymbol{y}_t = \boldsymbol{h}(\boldsymbol{x}_t) + \boldsymbol{w}_t \tag{4.4}$$

となる．$\boldsymbol{h}$ は状態ベクトルから観測ベクトルへの写像であり，カメラ座標へのモデルの変換と画面への投影を含む．$\boldsymbol{w}_t$ は観測誤差であり，平均 0，分散 $\boldsymbol{W}$ の確率分布に従う白色雑音と仮定する．

$\boldsymbol{h}$ は投影法と観測の種類によって異なる．$\boldsymbol{R}(\boldsymbol{\omega})$ を剛体回転行列 (付録 A参照) とすると，根元の 3 次元座標 $X_0$ と，根元から $i$ 番目のリンクの端点の 3 次元座標 $X_i$ は，

$$X_0 = \boldsymbol{t} \tag{4.5}$$

$$X_i = \boldsymbol{R}(\boldsymbol{\omega}) \sum_{j=1}^{i} \left\{ r_j \cos(\sum_{k=1}^{j} \theta_k), r_j \sin(\sum_{k=1}^{j} \theta_k), 0 \right\}^T + \boldsymbol{t} \tag{4.6}$$

となる．各リンクの端点と関節位置のうち $l$ 個が観測できる場合，画面での観測は

$$\boldsymbol{h}(\boldsymbol{x}) = (\boldsymbol{h}_0^T(\boldsymbol{x}), \cdots, \boldsymbol{h}_l^T(\boldsymbol{x}))^T \tag{4.7}$$

$$\boldsymbol{h}_i(\boldsymbol{x}) = \boldsymbol{L}(X_i) \tag{4.8}$$

である．ここで $\boldsymbol{L}(\cdot)$ は投影関数である．

一方，線分の端点が観測できず，直線の方程式 $m^T x + \rho = 0$ の $m$,$\rho$ のみが観測できるときは

$$\boldsymbol{h}_i(\boldsymbol{x}) = (m_i^T, \rho_i)^T \tag{4.9}$$

$$m_i = \boldsymbol{L}(X_i) - \boldsymbol{L}(X_{i-1}) \tag{4.10}$$

$$\rho_i = -m_i^T \boldsymbol{L}(X_i) \tag{4.11}$$

となる．

## 4.2 拡張カルマンフィルタ

前節で導いたの観測方程式は非線形であるのでそのままではカルマンフィルタを構成できない．そこで，観測を予測点の近傍で局所的に線形化する拡張カルマンフィルタ (Extended Kalman Filter: EKF) を適用する [31]．式 4.4は次のように線形化される．

$$\boldsymbol{y}_t \simeq \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t}\right|_{\widetilde{\boldsymbol{x}}_t} (\boldsymbol{x}_t - \widetilde{\boldsymbol{x}}_t) + \boldsymbol{h}(\widetilde{\boldsymbol{x}}_t) + \boldsymbol{w}_t \tag{4.12}$$

ここで $\widetilde{\boldsymbol{x}}_t$ は $\widehat{\boldsymbol{x}}_{t-1}$ を時刻 $t$–1 での推定の平均値としたときの時刻 $t$ の予測の平均値で

$$\widetilde{\boldsymbol{x}}_t \equiv \boldsymbol{A}\widehat{\boldsymbol{x}}_{t-1} \tag{4.13}$$

である．推定の平均値と分散の更新は次のようになる．

$$\begin{aligned} \widehat{\boldsymbol{x}}_t &= \widetilde{\boldsymbol{x}}_t + \boldsymbol{K}_t(\boldsymbol{y}_t - \boldsymbol{C}_t\widetilde{\boldsymbol{x}}_t) & (4.14) \\ \boldsymbol{P}_t &= (\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{C}_t)\boldsymbol{M}_t & (4.15) \end{aligned}$$

ここで，

$$\begin{aligned} \boldsymbol{M}_t &\equiv \boldsymbol{A}\boldsymbol{P}_{t-1}\boldsymbol{A}^T + \boldsymbol{U} & (4.16) \\ \boldsymbol{C}_t &\equiv \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t}\right|_{\widetilde{\boldsymbol{x}}_t} & (4.17) \\ \boldsymbol{K}_t &\equiv \boldsymbol{M}_t\boldsymbol{C}_t^T(\boldsymbol{C}_t\boldsymbol{M}_t\boldsymbol{C}_t^T + \boldsymbol{W})^{-1}. & (4.18) \end{aligned}$$

以上により式3.12中の $\widetilde{\Omega}_t \cap \bar{H}_t$ の部分が実装できた．次に制約知識 $\bar{F}_t$ の統合について次節で述べる．

## 4.3 制約による推定分布の修正

3.2節で述べた各関節角度と各部の形状についての制約知識を次のように線形不等式で表現する[1]．

$$\theta_{min,i} \leq \theta_i \leq \theta_{max,i}, \quad |\theta_i - \theta_j| \leq \Delta\theta_{ij} \tag{4.19}$$

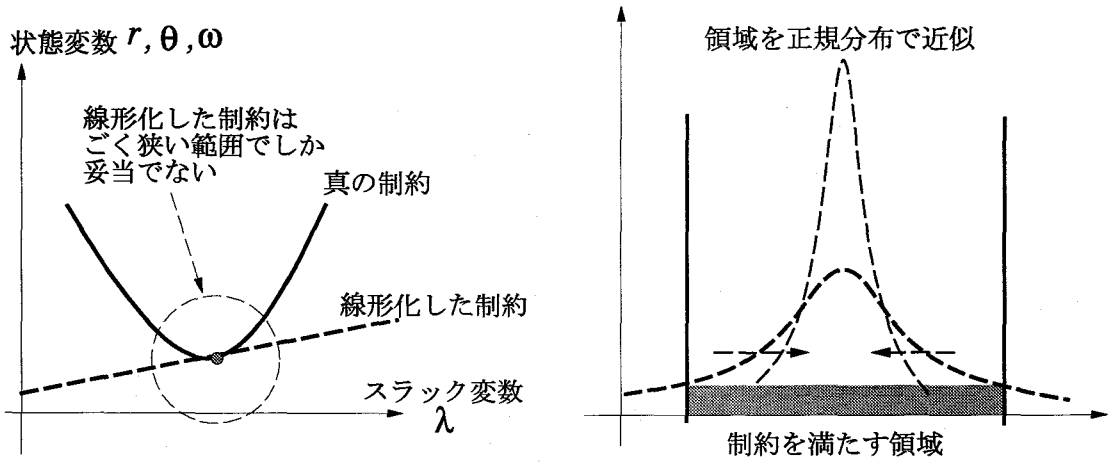
$$r_{min,i} \leq r_i \leq r_{max,i}, \quad |r_i - r_j| \leq \Delta r_{ij} \tag{4.20}$$

線形な等式による制約ならば，分散0の観測と考えることで容易にカルマンフィルタに組み込むことができるが，本研究では不等式による制約を扱わねばならない．[24] では式4.20のような不等式を，状態ベクトルに新たなスラック変数 $\lambda$ を導入して

$$g(\boldsymbol{x}) - \lambda^2 = 0 \tag{4.21}$$

の形に表現して EKF に組み込んでいる．しかしこの方法では $g(\boldsymbol{x})$ が線形でも $\lambda$ については線形化しなければならない．このように線形化された制約は図 4.1(a) に示すように，局所的なごく狭い範囲でしか妥当とはいえないので，$\Delta\theta_{ij}$ や $\Delta r_{ij}$ が大きい制約の場合は制約条件を正確に反映しない．

[1]線形ならば3つ以上のパラメータについての制約条件を用いることもできる




(a) スラック変数を用いたときの問題点　　(b) 制約を正規分布で近似したときの問題点

図 **4.1**: 従来手法における不等式制約の取り扱い方の問題点

他の手段として，図 4.1(b) に示すように，不等式の制約を平均と誤差の確率分布で置き換えてカルマンフィルタに組み込むことが考えられるが，この方法はひとつの制約を各時刻で毎回統合することを意味する．すると，予測分散よりも観測後の推定分散が小さくなるというカルマンフィルタの性質と，$r$ の予測誤差の分散が 0 であることから，仮に観測がまったく得られなくても制約だけで $r$ の分散が小さくなっていく不都合が生じる．これは制約を過剰に考慮して推定を行なっていることになる．したがってフィルタリングにおける不等式制約は観測とは別の取り扱いが必要である．

## 4.4 線形な制約による分布切りとり法

カルマンフィルタによる推定は正規分布でとらえることができる．このうち制約条件 (4.19), (4.20) をみたす領域の平均と分散を計算することにより，制約条件を統合する．

式 (4.19),(4.20) は一次式だから

$$\varphi_k^T \boldsymbol{x} \ \leq \ b_k \ (1 \leq k \leq K) \tag{4.22}$$

と一般化できる．したがって正規分布で表現される推定値の分布を平面 $\varphi_k^T \boldsymbol{x} = b_k$ で切りとって，残りの部分の平均と分散を求めればよい（図 4.2）．しかし，一度にすべての制約をみたす部分の平均と分散を計算するのは困難なので，一つずつの平面で順番に分布を切りとっていく計算法で近似する．

ここで，平均 $\boldsymbol{q}_{k-1}$，分散 $\boldsymbol{Q}_{k-1}$ の正規分布のうち，$\varphi_k^T \boldsymbol{x} \leq b_k$ の制約を満たす領域の平均

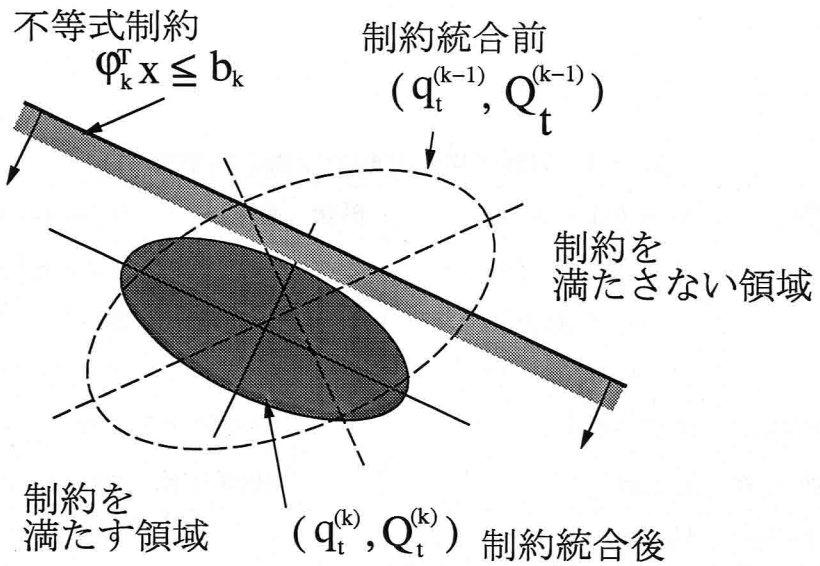



図 **4.2**: 確率分布の切りとり

$\boldsymbol{q}_k$ と分散 $\boldsymbol{Q}_k$ を計算することを考える．この問題は次の線形変換

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{T}^T(\boldsymbol{x} - \boldsymbol{q}_{k-1}) \tag{4.23}$$

によって，座標軸に平行な平面と平均０，分散 $\boldsymbol{I}$，制約不等式 $(1, 0, \cdots, 0)^T\boldsymbol{x}' \leq c_k$ の場合に帰着できる．ただし $\boldsymbol{R}, \boldsymbol{T}$ は直交行列, $\boldsymbol{W}$ は対角行列であり，

$$\boldsymbol{T}\boldsymbol{W}\boldsymbol{T}^T = \boldsymbol{Q}_{k-1} \tag{4.24}$$

$$\boldsymbol{R}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{T}^T\varphi_k = (1, 0, \cdots, 0)^T \tag{4.25}$$

$$c_k = \frac{b_k - \varphi_k^T\boldsymbol{q}_{k-1}}{\left(\varphi_k^T\boldsymbol{Q}_{k-1}\varphi_k\right)^{\frac{1}{2}}}. \tag{4.26}$$

をみたす．切りとったあとの $\boldsymbol{x}'$ の分布の平均 $\mu$ と分散 $\boldsymbol{S}$ は

$$\mu = -\sqrt{\frac{2}{\pi}}\exp(-\frac{c_k^2}{2})/(1 + \mathrm{erf}(\frac{c_k}{\sqrt{2}})) \tag{4.27}$$

$$S_{ii} = \begin{cases} 1 + c_k\mu - \mu^2 & \cdots i = j = 1 \\ 1 & \cdots i = j \neq 1 \\ 0 & \cdots \text{otherwise.} \end{cases} \tag{4.28}$$

となる[2]．これを式 (4.23) の逆変換でもとの$\boldsymbol{x}$の座標に戻せば,

$$\boldsymbol{q}_k = \boldsymbol{T}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{R}^T\mu + \boldsymbol{q}_{k-1} \tag{4.29}$$

$$\boldsymbol{Q}_k = \boldsymbol{T}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{R}^T\boldsymbol{S}\boldsymbol{R}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{T}^T. \tag{4.30}$$

[2]詳しくは付録 B参照.

表 4.1: 実験で用いた制約知識と初期値 (1)

| | | | | | |
|---|---|---|---|---|---|
| 姿勢 | 制約 | $\lvert\theta_1-\theta_2\rvert \le \pi/4$, | 形状 | 制約 | $\lvert r_1-r_2\rvert \le 40$, |
| | | $\lvert\theta_2-\theta_3\rvert \le \pi/6$, | | | $\lvert r_2-r_3\rvert \le 10$, |
| | | $0 \le \theta_1 \le 3\pi/4$, | | | $30 \le r_1 \le 90$, |
| | | $0 \le \theta_2, \theta_3 \le \pi/2$ | | | $25 \le r_2, r_3 \le 70$ |
| | 初期推定 | 真値 ±0.2rad | | 初期推定 | 真値 ±10pixels |
| | 初期分散 | $0.2\mathrm{rad}^2$ | | 初期分散 | $400\mathrm{pixels}^2$ |
| 観測 | 誤差分散 | $5\mathrm{pixels}^2$ | | | |

となる.

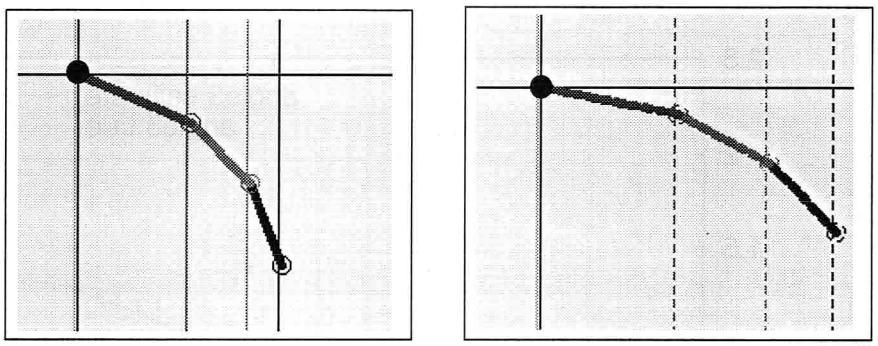
制約考慮前の推定平均を $\boldsymbol{q}_0$, 推定分散を $\boldsymbol{Q}_0$ とおいて, 上記の計算をすべての $k$ について再帰的に実行すれば, 分布切りとり後の平均 $\hat{\boldsymbol{x}}_t = \boldsymbol{q}_K$ と分散 $\boldsymbol{P}_t = \boldsymbol{Q}_K$ が得られる. これは近似計算なので, 誤差をできるだけ少なくするために, もとの平均値からのマハラノビス距離 $c_k$ が大きい平面から順に実行する.

## 4.5 実験 1

ここで本手法を適用した関節物体の形状と姿勢の推定実験を示す. 簡単のために 2 次元の 3 関節物体について計算機による推定シミュレーションを行なった. 観測として各関節と端点の位置が 1 次元で得られると仮定し, 計算機によって生成した. 観測に加えるノイズは白色ガウス分布に従うノイズを乱数により発生させた. なお, 関節物体の 1 番根元の関節は原点に固定されているとした. 初期推定は図 4.4に示すように与えた.

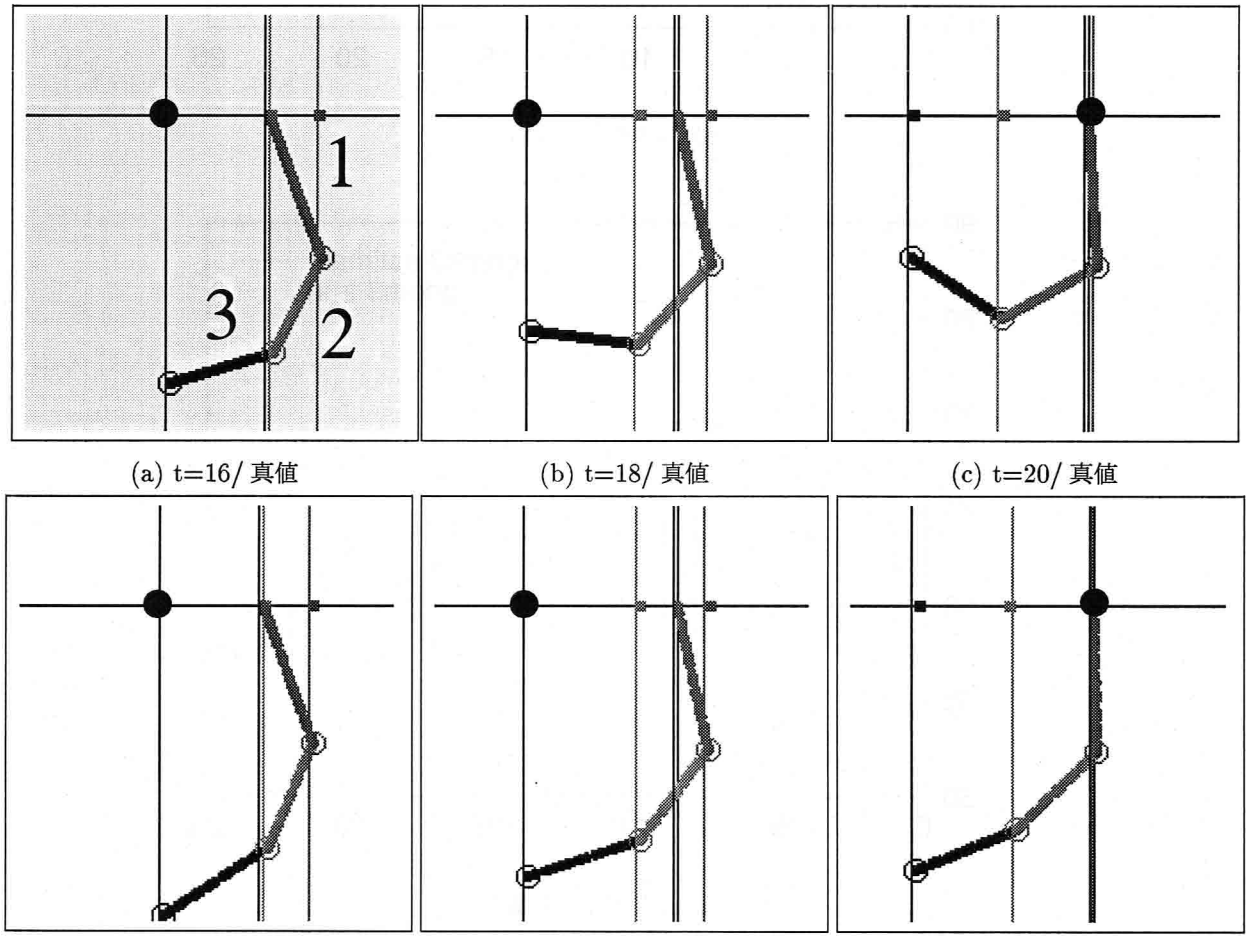
ここで物体に関する知識として表 4.1に示す制約を導入して推定を行なった結果を図 4.4に示す. 図 4.4の上段は真の姿勢であり, 下段は推定の結果である. 図 4.5は, 第 3 リンク (もっとも末端のリンク) の関節角度とリンク長の推定結果のグラフである. 破線は真値を表し, マークと縦線はそれぞれ推定の平均と分散を表している (分散は $2\sigma$ の範囲を表示).

図 4.4からわかるように, 真の姿勢は第 3 リンクが下から上に向かって回転しているにもかかわらず ((a)-(c)), 推定結果はそうなっていない. 第 20 フレーム以降では推定は真値からはずれていき, この後の推定が不可能になった. これは奥行きが観測できないことに由来する推定の多義性のためである. そこで次節ではこの問題への対処について述べる.



(a) t=1/ 真値　　(b) t=1/ 推定

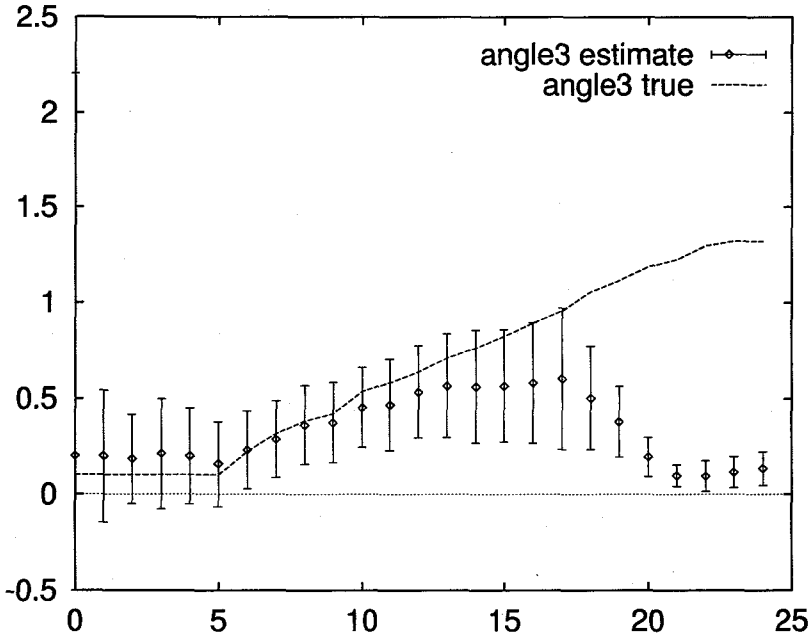
図 **4.3**: 実験１：初期推定



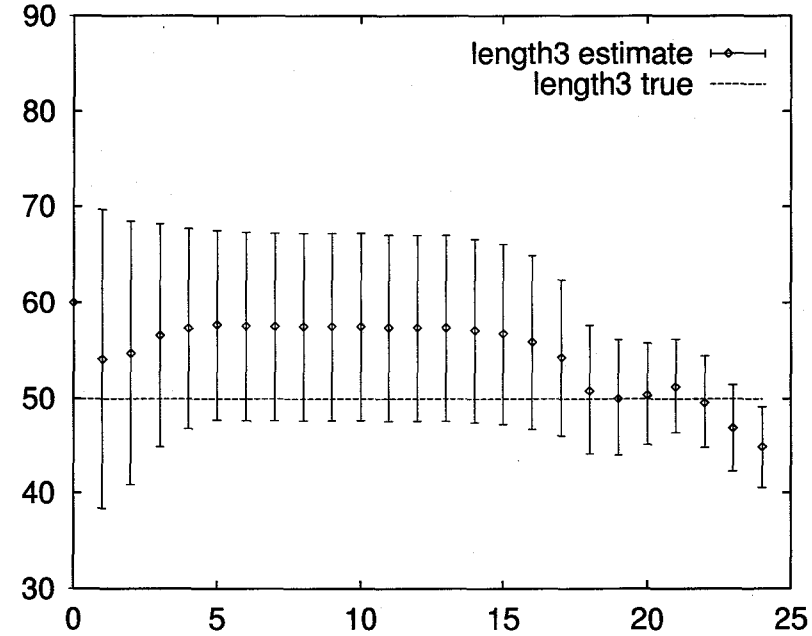

(a) t=16/ 真値　　(b) t=18/ 真値　　(c) t=20/ 真値

(d) t=16/ 推定　　(e) t=18/ 推定　　(f) t=20/ 推定

図 **4.4**: 実験１：EKF と分布切りとり法による推定




(a) $\theta_3$ の推定と分散




(b) $r_3$ の推定と分散

図 **4.5**: 実験１：第３リンクの関節角とリンク長の推定の平均と分散

観測のヤコビ行列Cが0に近い場合

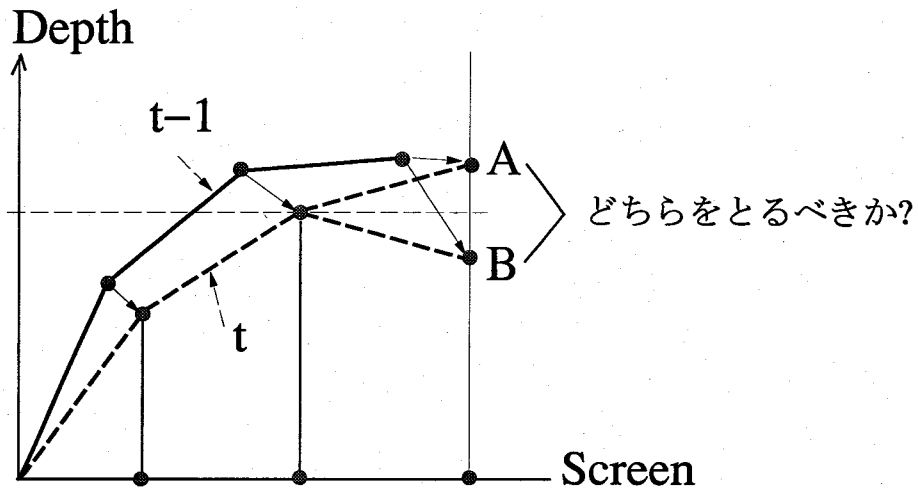



図 4.6: 複数の解釈が可能な場合

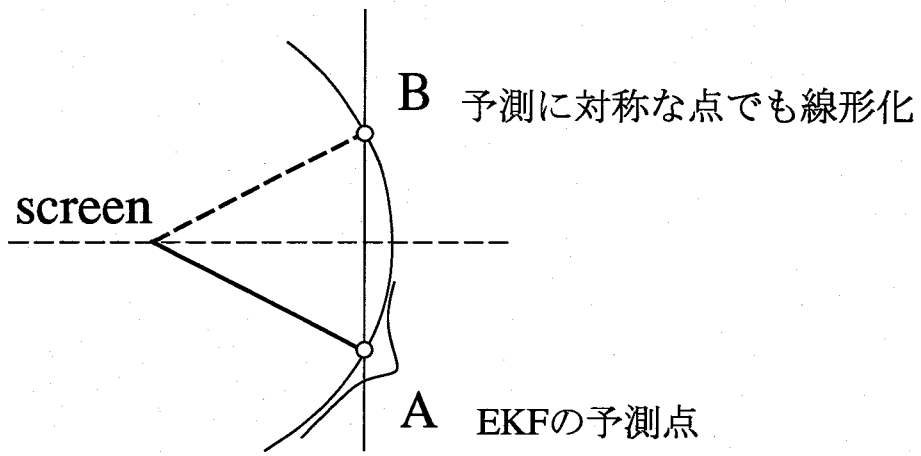



図 4.7: 観測を線形化する位置が複数ある場合

## 4.6 複数候補の生成による非線形性の補償

EKF は非線形なシステムの線形近似にすぎないので，前節の実験で示したように場合によっては非線形性の影響を無視できないことがある．

図 4.6のような場合，リンクが奥へ向かっているか，手前に向かっているかの違いで，A,B の両方に解の可能性がある．しかし，通常の EKF では A か B のどちらかの推定しか出力できない．これは観測 $\boldsymbol{h}(\boldsymbol{x}_t)$ の非線形性により尤度分布は多峰性であるのに，線形化によって単峰性として計算することに起因する．

[35] は平均や分散などの統計量を推定するのではなく，サンプリング法 [36] を利用して，多くのサンプルで多峰性の分布を表現する手法を提案している．この手法は非常に多くのサ

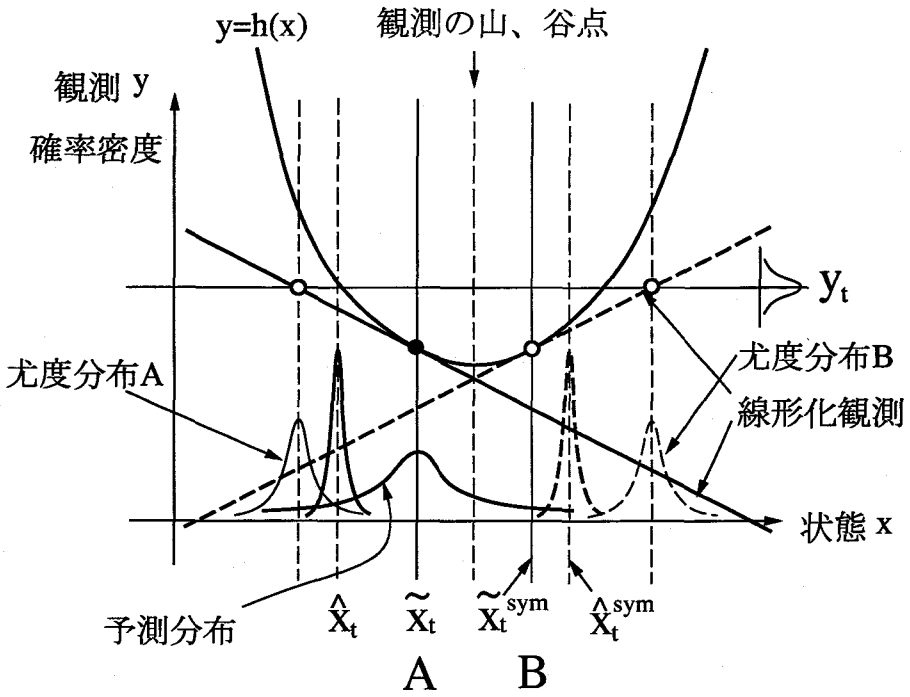



図 4.8: 複数の線形化観測の統合の様子

ンプルを必要とするので，計算量の問題から状態ベクトルの次元が大きいものには適用しにくい．

そこで，本研究では複数の EKF による推定を考える．図 4.6のリンク姿勢におけるパラメータ空間での観測の線形化の様子，ならびに分布の統合の様子を図 4.8に示す．図 4.6のような非線形性の影響が無視できない場合はローカルな角度パラメータ変化が観測をほとんど変化させない．このような場合は観測 $\boldsymbol{h}(\boldsymbol{x})$ の山もしくは谷付近でおきる[3]．その近傍では図 4.7,4.8中の A，B の 2 点で線形化が可能である．その結果，図 4.8下部のグラフに示すように複数の尤度分布が得られる．これらをそれぞれ予測分布と統合して複数の推定分布が得られる．

処理をまとめると以下のようになる．

1. $i$ 番目の関節に対し

$$\left.\frac{\partial \boldsymbol{h}_i}{\partial \boldsymbol{\Theta}_i}\right|_{\tilde{\boldsymbol{x}}_t} \simeq 0 \tag{4.31}$$

が成立する場合には，以下の処理を行う（$\boldsymbol{\Theta}_i$ は $i$ 番目のリンクを直接動かすすべての角度パラメータ）．

[3] $\boldsymbol{x}$の次元が 2 以上の場合は必ずしも極とはかぎらない．ある 1 方向の断面が $\boldsymbol{h}(\boldsymbol{x})$ のすべての成分に関して（$\boldsymbol{h}$ はベクトルであることに注意）山もしくは谷状になってさえいれば，その近傍では複数の解釈が生じる．

2. EKF の予測 $\tilde{\boldsymbol{x}}_t$ を得る.

3. $i$ 番目のリンクが画面に関して $\tilde{\boldsymbol{x}}_t$ と対称になる姿勢 $\tilde{\boldsymbol{x}}_t^{sym}$ を生成する.

4. $\tilde{\boldsymbol{x}}_t$,$\tilde{\boldsymbol{x}}_t^{sym}$ の位置で観測をそれぞれ線形化する.

5. EKF を用いて, 4 で得られたそれぞれの線形化観測を予測分布と統合する.

6. 5 で得られたそれぞれの推定に対し, 制約による分布切りとりを行う. ただし, 大きく制約を外れるものは間違いとみなして棄却する.

上の処理の結果, すべてのリンクが画面とほぼ平行な場合に, 最大 $2^n$ 個 ($n$: リンクの数) の予測が生成される可能性がある. これらの推定群をすべて保持するのはコストがかかるので, 予測分布に基づく各推定の平均値の尤度を用いて各推定を順位づけする. 尤度の高いものから一定数の推定のみを保存し, それらに対して次の時刻の推定を行なう (ビーム探索).

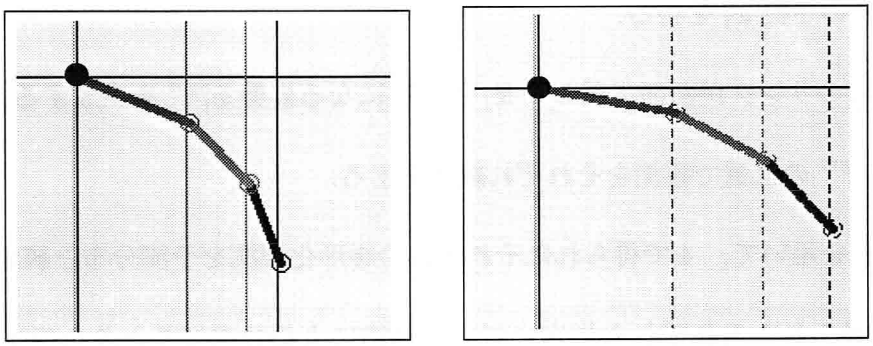
## 4.7 実験 2

実験 1 と同じ設定で複数の姿勢推定の実験を試みた. 図 4.4は初期推定である. 図 4.10上段は真の姿勢であり, 下段は推定結果である. 図 4.11から 4.14は第 1 のリンクと第 3 のリンクについて関節角度 ($\theta_1$,$\theta_3$) とリンク長 ($r_1$,$r_3$) の変化を表したグラフで, 破線は真値を表し, マークと縦線はそれぞれ推定の平均と分散を表している (分散は $2\sigma$ の範囲を表示).

図 4.11と図 4.12によると, $\theta_1$ は正確に推定されているが, $r_1$ はそれに比べてそれほど正確に推定されていない. これは, シーケンスを通じて第 1 のリンクが画面に垂直に近かったためである.

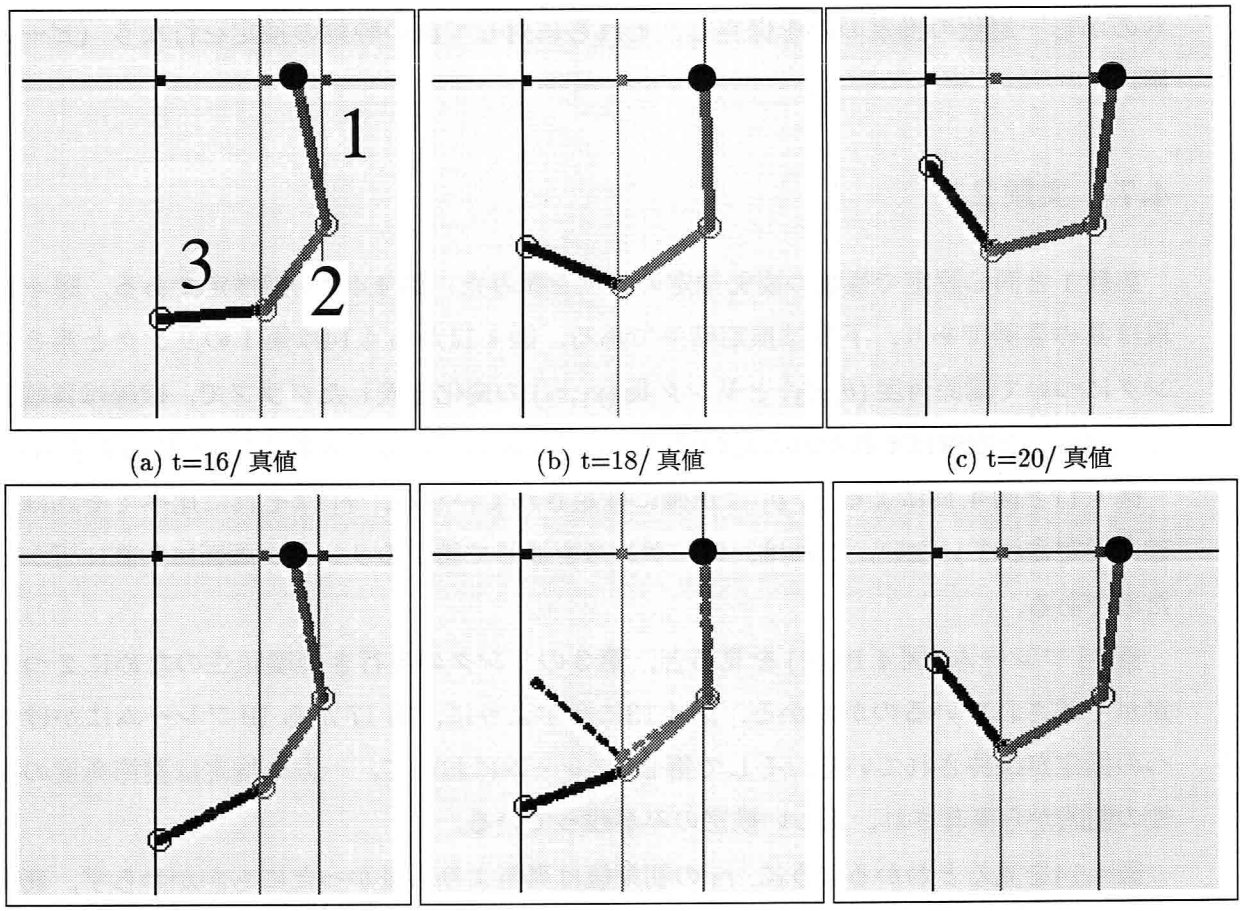
第 18 フレーム (図 4.10(e)) を見ると, 第 3 のリンクの奥行きの曖昧性のために 2 つの推定が生成されているのがわかる. 図 4.13に示すように, 第 17 から 19 フレームにかけて 2 つの推定が保持されている. そして第 20 フレームにおいて, 一方の推定は関節角度の相関性の制約から棄却され, 正しい推定のみが残っている.

図 4.14を見るとわかるように, $r_3$ の初期値は真値よりも長かったにもかかわらず, 真値に近付いている. 制約がなければ, あらゆる観測はより長いリンクによって説明できるので, この効果は制約知識を用いた成果である. 第 20 から 30 フレームにかけて第 2 のリンクは画面にほぼ平行だったので, 長さ $r_2$ が正確に推定できた. これが各リンクの長さの相関性によって $r_3$ に伝播したと考えられる.



(a) t=1/ 真値　　(b) t=1/ 推定

図 **4.9**: 実験 2 ：初期推定




(a) t=16/ 真値　　(b) t=18/ 真値　　(c) t=20/ 真値

(d) t=16/ 推定　　(e) t=18/ 推定　　(f) t=20/ 推定

図 **4.10**: 実験 2 ：複数の EKF による推定

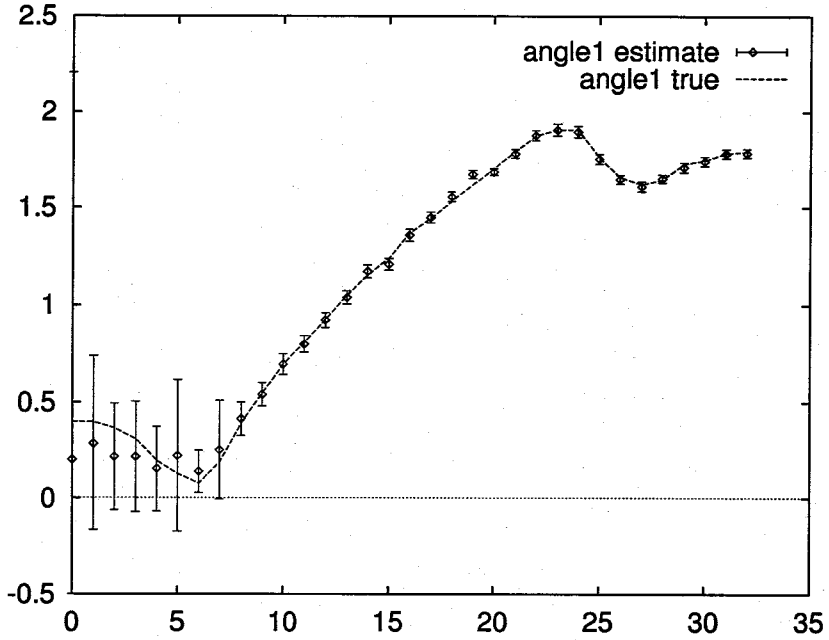



図 **4.11**: $\theta_1$ の推定と分散

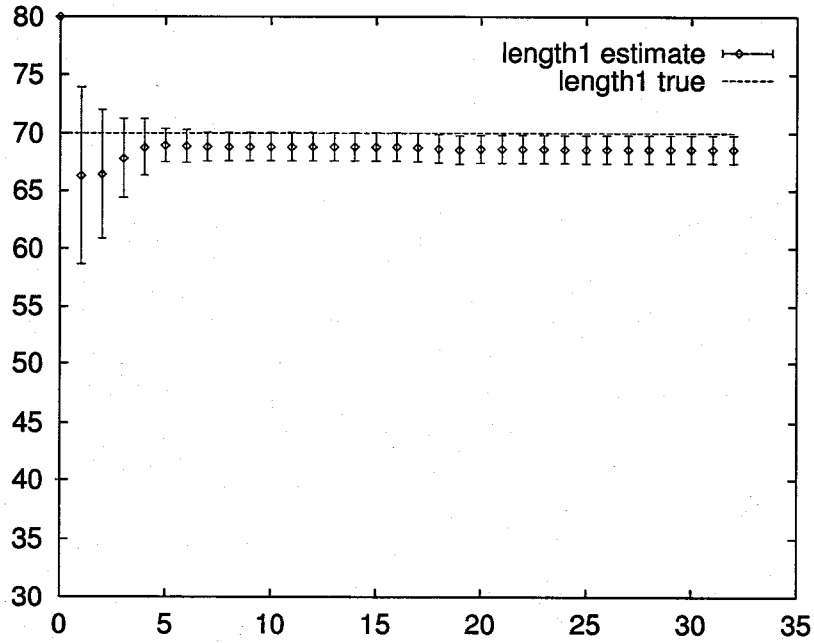



図 **4.12**: $r_1$ の推定と分散

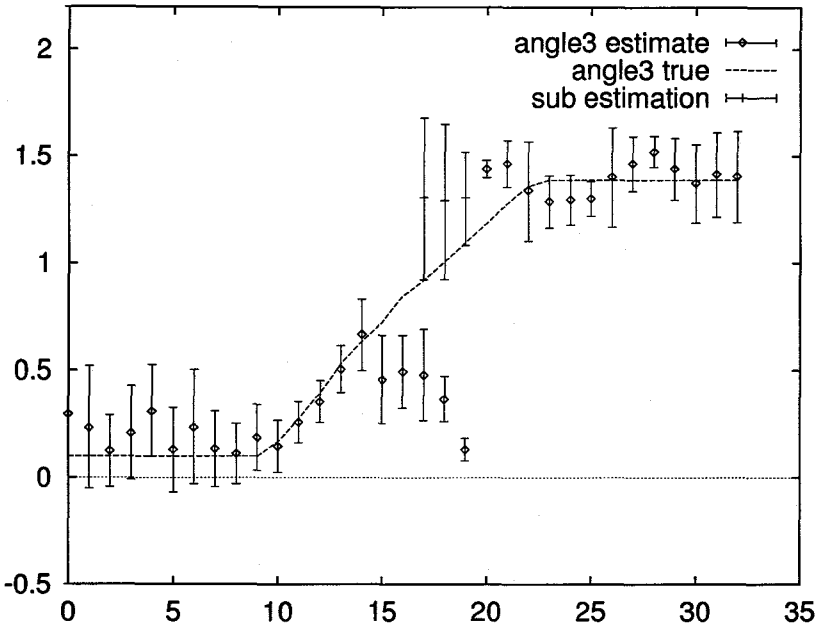



図 **4.13**: $\theta_3$ の推定と分散

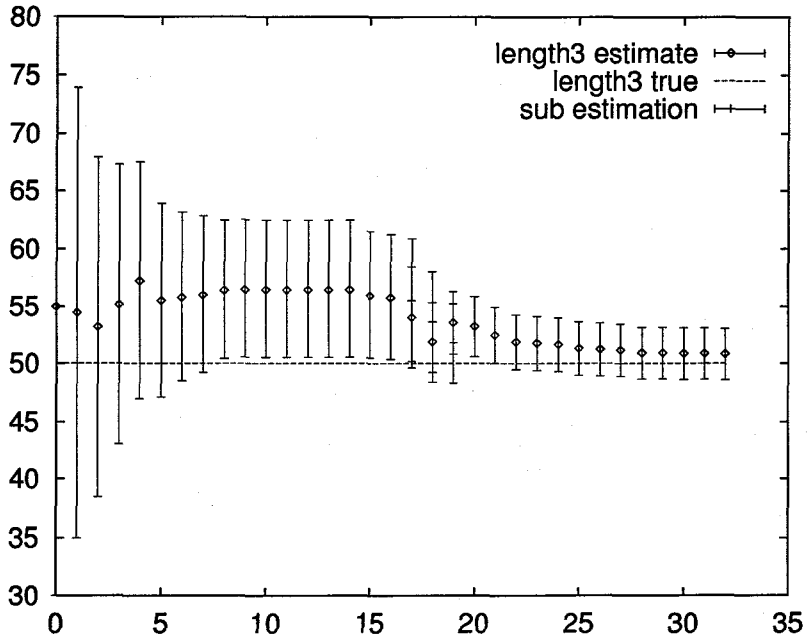
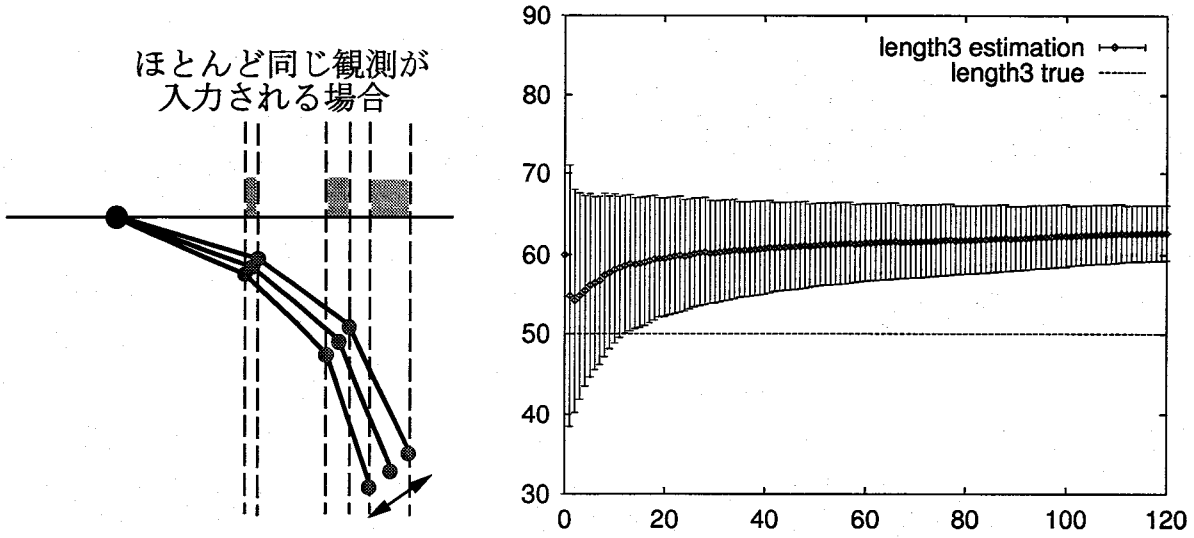



図 **4.14**: $r_3$ の推定と分散




(a) 類似した観測の連続入力　　(b) $r_3$ の推定結果

図 **4.15**: 類似の観測を長い時間に渡って入力したときの推定誤り

## 4.8 拡張カルマンフィルタの問題点

実験 2 は，入力する観測データを十分に変化させたシーケンスについての推定例であった．しかしほぼ類似する観測データが長い時間に渡って入力された場合には問題が発生する．

### 4.8.1 リンク長の推定値が誤っていても分散が小さくなる問題点

図 4.15(b) は最初のフレームから真の姿勢を動かさない状態で推定をつづけたときの，$r_3$ の推定結果のグラフである．推定分散は通常推定の確からしさを表すと考えられる．そこで分散が一定値以下になった場合は長さの確からしい推定値が既に得られたと考えて，角度のみの推定に切替えることを考える．図 4.15の実験の場合，新たな情報がまだ得られていないのだから，推定の分散は変化しないのが望ましい．しかし EKF を用いた推定では，リンク長の推定ではグラフに示すように誤った推定値のまま分散が小さくなってしまう．

これはカルマンフィルタの性質と元の観測の非線形性に由来する．線形かつ可観測でないシステムにおいて，カルマンフィルタは状態パラメータ空間中で図 4.16左図のようにふるまう．図の座標軸は各状態パラメータの軸を表す．破線で描かれた直線は観測を満たすパラメータの集合である．可観測でない線形システムでは，観測によって拘束される $\boldsymbol{x}$ の空間の傾きはいかなる $\boldsymbol{y}$ についても等しい．したがって，ある初期分布のもとで観測を統合す

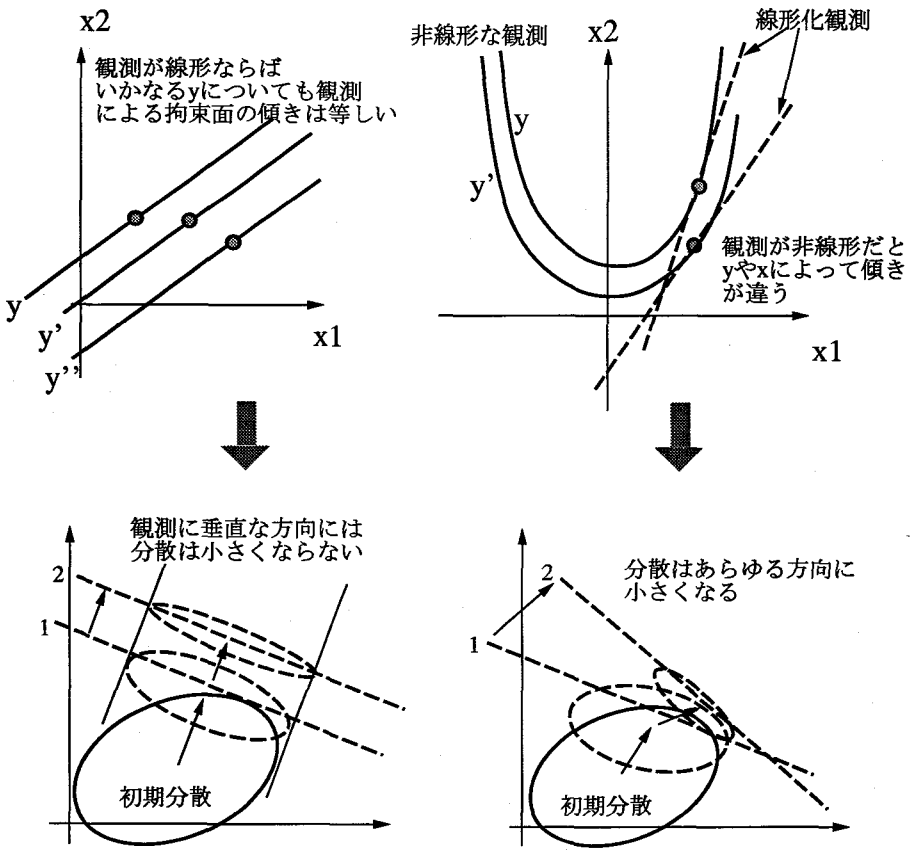



図 **4.16**: 可観測でない場合における EKF の悪い振舞い

ると，観測できるモード（図の破線に垂直な方向）の分散だけが小さくなり，観測できないモード（破線に平行な方向）の分散は変化しない[4]．したがって物体の形状と姿勢に関する制約によってのみその方向の分散は小さくなる．

一方，観測が非線形の場合， EKF によって観測を線形化するが，図 4.16の右図のようにその傾きは観測 $\boldsymbol{y}$ や線形化する場所 $\boldsymbol{x}$ によって異なる．したがって EKF では分散が小さくならない方向が存在しない．つまり EKF で近似したシステムは事実上可観測になっている．

実験では観測には乱数によって白色ガウスノイズを加えたので，真の姿勢がまったく動かなくても観測値は微小に変化する．観測値が微小に変動すると推定ならびに予測が変動するので，観測を線形化する位置が変化する．したがって状態空間中における観測の傾き，すなわち式 4.12の $\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t}$ が変化する．このため，長い時間同じような観測を与え続けると，ノイズ変動のせいで推定分散がどんどん小さくなっていく．これでは，推定値が間違ったままでも推定分散が十分小さくなる可能性があるので，いつリンク長の推定が確からしく得られた

[4]正確には観測できるモードと観測できないモードの相関が 0 になるまで，観測できないモードの分散は小さくなりうるが，それ以上は変化しない．

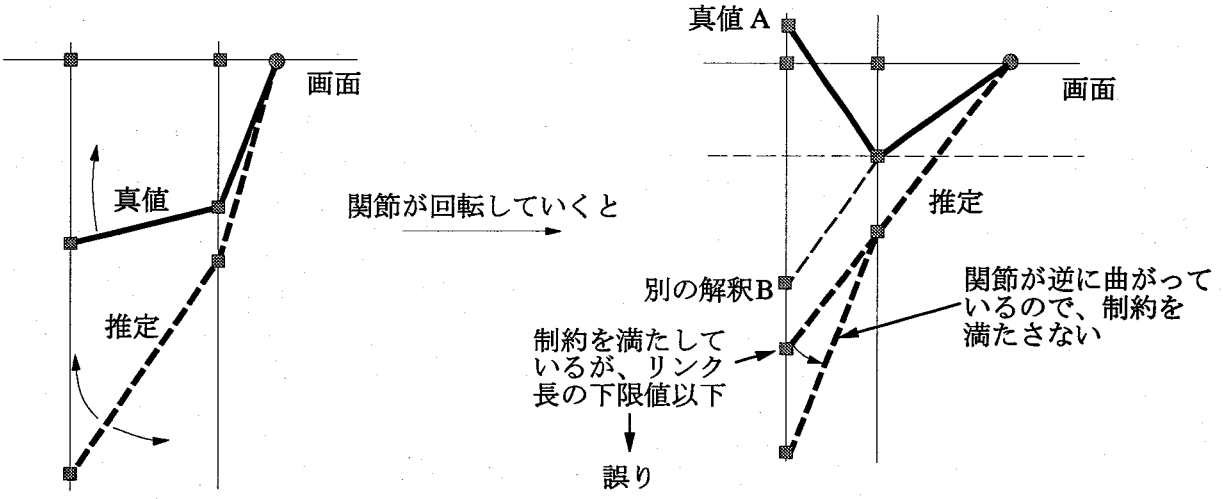



図 4.17: リンク長の推定が真値より極端に長い場合

か判断することができない.

### 4.8.2 複数の推定候補が生成できない問題点

線形なシステムにおけるカルマンフィルタでは，前節で述べたように誤った推定値のまま分散が小さくなっても，さまざまな観測を十分入力すれば推定は真値に近付いていく．しかし，非線形かつ不可観測なシステムでは必ずしもそうならない.

リンク長を真値より極端に長く推定している状態では，図 4.17に示すように真のリンクが画面と平行になっていても，推定値は画面に対して深い角度をもつ．すると実際にはAとBの2通りの解釈が生じているにもかかわらず，式 4.31の左辺が0にならないので，Bに相当する推定候補しか生成できない．すると関節がある方向にしか可動しないという制約とリンク長の下限値の制約があるときは，図 4.17のように制約を満たさない候補しか生成できない場合がある.

以上のように，単眼視動画像において関節物体の形状と姿勢を推定する場合，EKF では望ましい結果が得られないことがわかった．EKF の問題を解決するには非線形のフィルタリング [37] を行なう必要があるが，一般に非線形フィルタリングは計算が困難である．そこで，次に本研究では確率分布としての扱いではなく，領域の境界を表現することで式 3.12の実装することを考えた．この手法を次章で述べる.

# 第 5 章

# 楕円近似法を用いた形状と姿勢の推定

## 5.1 楕円近似法

本節では 3.3節で述べた更新式

$$\Omega_t = \widetilde{\Omega}_t \cap \bar{H}_t \cap \bar{F}_t \tag{3.12}$$

の $\Omega$,$\bar{H}$,$\bar{F}$ を確率分布ではなく，境界をもった領域として表現する．式 3.12中の記号'∩' は集合の積であるから，それぞれの領域が交差する領域を表現したいが，空間中のあらゆる領域を正確に表現する一般的な方法はない．従って本節では，

1. 推定範囲 $\Omega$ については多次元の楕円体
2. 観測，制約を満たす領域 $\bar{H}$,$\bar{F}$ については多次元の平行な二つの超平面にはさまれた領域

でパラメトリックに近似表現し，推定の更新式 3.12を計算する．

図 5.1に示すように，楕円体と平行な二平面の両方に挟まれた領域 $\Omega_t$ は，厳密には楕円体ではない．しかし，楕円近似法 [38][39] を用いると，$\Omega_t$ を完全に含む楕円体 $\widehat{\Omega}_t$ を以下に示す要領で容易に計算することができる．

まず $\Omega_{t-1}$ を中心 $\widehat{\boldsymbol{x}}_{t-1}$ とノルム行列 $\boldsymbol{P}_{t-1}$ を用いて

$$(\boldsymbol{x}_t - \widehat{\boldsymbol{x}}_{t-1})^T \boldsymbol{P}_{t-1}^{-1} (\boldsymbol{x}_t - \widehat{\boldsymbol{x}}_{t-1}) \leq 1 \tag{5.1}$$

と表現する．また $\bar{H}$,$\bar{F}$ をともに複数の不等式

$$|\boldsymbol{m}_i^T \boldsymbol{x}_t - \rho_i| \leq c_i \tag{5.2}$$

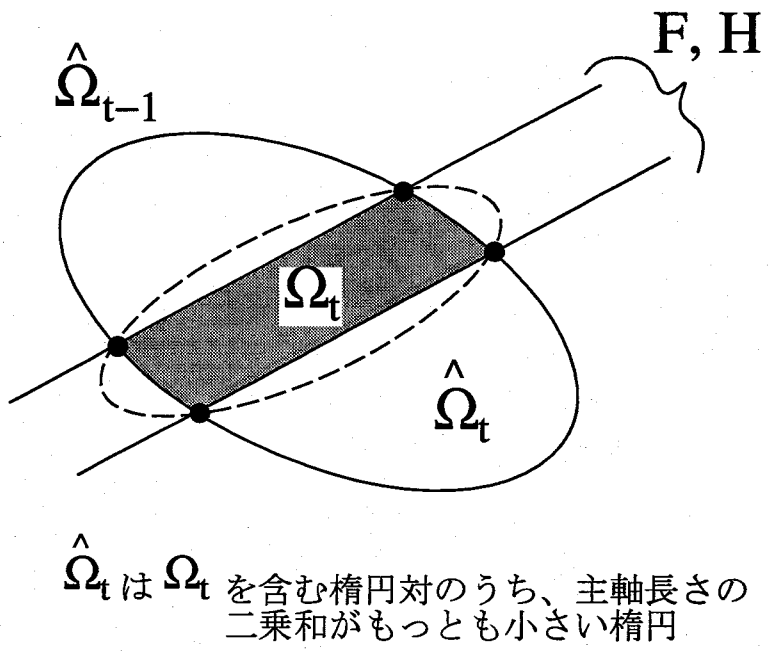



図 **5.1**: 楕円近似法

で表現する．$\boldsymbol{m}_i$ は面の法線，$\rho_i$ は平面の原点からのオフセット，$c_i$ は誤差の許容量を表す．両者に含まれる領域を完全に含む楕円体は $\lambda \geq 0$ を用いて

$$(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_{t-1})^T \boldsymbol{P}_{t-1}^{-1}(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_{t-1}) + \lambda|\boldsymbol{m}_i^T \boldsymbol{x}_t - \rho_i|^2 \leq 1 + \lambda c_i^2 \tag{5.3}$$

と表せる．

このような楕円体のうち，各主軸長さの二乗和 ($\mathrm{tr}\boldsymbol{P}_t$) が最小[1]になる $\lambda$ は，

$$\mu \equiv \mathrm{tr}\boldsymbol{P}_{t-1} \tag{5.4}$$

$$\gamma \equiv \boldsymbol{m}_i^T \boldsymbol{P}_{t-1}^2 \boldsymbol{m}_i \tag{5.5}$$

$$G \equiv \boldsymbol{m}_i^T \boldsymbol{P}_{t-1} \boldsymbol{m}_i \tag{5.6}$$

$$\varepsilon \equiv \rho_i - \boldsymbol{m}_i^T \hat{\boldsymbol{x}}_{t-1} \tag{5.7}$$

$$\beta_1 \equiv 3/G \tag{5.8}$$

$$\beta_2 \equiv [G\mu(3c_i^2 - \varepsilon^2) - \gamma\{2(c_i^2 - \varepsilon) + G\}]/[G^2 c_i^2(\mu G - \gamma)] \tag{5.9}$$

$$\beta_3 \equiv \{\mu(c_i^2 - \varepsilon^2) - \gamma\}//[G^2 c_i^2(\mu G - \gamma)] \tag{5.10}$$

としたときの三次方程式

$$\lambda^3 + \beta_1\lambda^2 + \beta_2\lambda + \beta_3 = 0 \tag{5.11}$$

[1]体積が最小にな解も求めることができるが，本研究においては不都合な面があるので採用しなかった．付録 C参照．

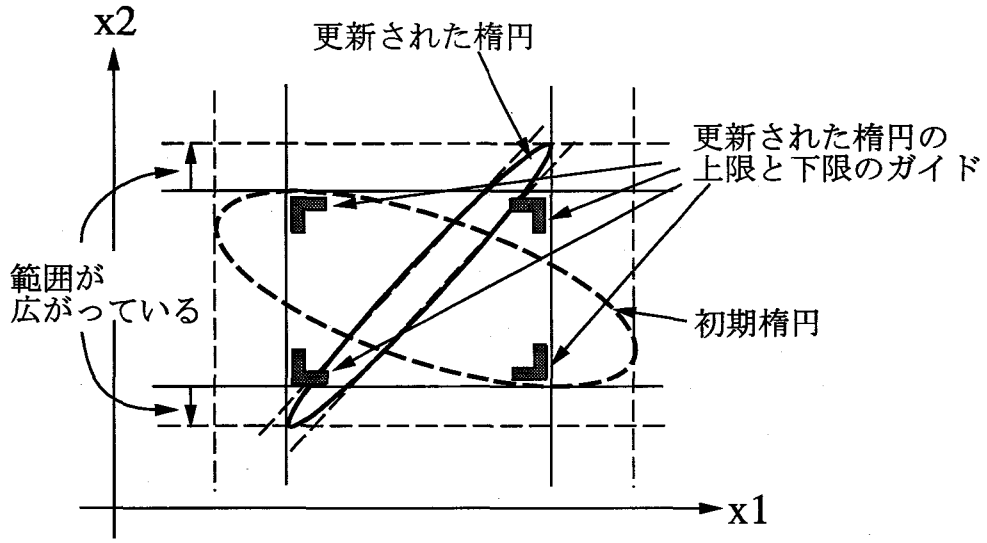



図 5.2: パラメータの上限と下限のガイド

の解である．ただし $\lambda < 0$ のときは $\lambda = 0$ とする[2]．$\widehat{\Omega}_t$ を表す $\widehat{\boldsymbol{x}}_t$ と $\boldsymbol{P}_t$ は次のように更新される．

$$\widehat{\boldsymbol{x}}_t = \widehat{\boldsymbol{x}}_{t-1} + \lambda \boldsymbol{Q} \boldsymbol{m}_i \varepsilon \tag{5.12}$$

$$\boldsymbol{P}_t = z\boldsymbol{Q} \tag{5.13}$$

ただし，

$$\boldsymbol{Q} \equiv \boldsymbol{P}_{t-1} - \lambda \frac{\boldsymbol{P}_{t-1}\boldsymbol{m}_i\boldsymbol{m}_i^T\boldsymbol{P}_{t-1}^T}{1+\lambda G} \tag{5.14}$$

$$z \equiv 1 + \lambda c_i^2 - \frac{\lambda \varepsilon^2}{1+\lambda G} \tag{5.15}$$

$$\tag{5.16}$$

である．

これをすべての $i$ について順に適用することによって $\widehat{\Omega}_t$ を更新する．

このようにして更新された $\widehat{\Omega}_t$ は真の領域 $\Omega_t$ を完全に含むことが，式5.3によって保証されるので，次の定理が成立する．

**定理 1** 式 *5.12,5.13*による領域の更新の際，式 *5.2*の $i$ をどの順に適用しても，最終的な $\widehat{\Omega}_t$ は必ず真の領域 $\Omega_t$ を完全に含む．

この楕円近似法による領域の更新では，定理1によって，領域が不当に小さくはならないことが保証されるが，近似した楕円体は余分な領域を含んでしまう．図5.2のような場合，

[2] $\lambda > 0$ の実解はたかだか1つしかない [38].

姿勢形状パラメータ $x_2$ の可能な範囲が更新後の楕円では初期楕円よりも広がってしまっている．あまり余分な領域を含んでしまうと，リンク長の推定範囲が限定できない場合がある．

そこで本研究ではこれを補正するために，パラメータ空間の各座標軸方向に楕円領域の上限値 $x_i^{max}$ と下限値 $x_i^{min}$ をガイドとして記憶しておく．これにより，楕円の内側でもガイドの外側は対象外であることが判断できる．従って $\widehat{\Omega}_t$ は，中心 $\widehat{\boldsymbol{x}}_t$，ノルム行列 $\boldsymbol{P}_t$ に加えて，上限ベクトルと下限ベクトル

$$\boldsymbol{x}_t^{max} = \{x_1^{max}, \cdots, x_n^{max}\} \tag{5.17}$$

$$\boldsymbol{x}_t^{min} = \{x_1^{min}, \cdots, x_n^{min}\} \tag{5.18}$$

によって記述される．

このガイド付き楕円近似法を使って式3.12を計算するには，観測を満たす領域 $\bar{H}$，制約を満たす領域 $\bar{F}$ をの式5.2による表現と，予測過程における $\Omega_t$ から $\widetilde{\Omega}_{t+1}$ への更新が必要である．本研究では線形な制約式4.19,4.20のみを考えるので $\bar{F}$ は式5.2で表現できる．そこで，次節と次々節では，非線形な観測の式5.2による表現と，予測過程の表現について順に述べる．

## 5.2 観測の非線形性の見積もりと補償

3.2節では，誤差を含んだ時刻 $t$ での観測 $\boldsymbol{y}_t$ を満たすパラメータ領域 $\bar{H}$ の表現として

$$|\boldsymbol{y}_t - \boldsymbol{h}(\boldsymbol{x}_t)| \leq \Delta \boldsymbol{y} \tag{3.2}$$

を考えた[3]．これは

$$|y_{t,i} - h_i(\boldsymbol{x}_t)| \leq \Delta y_i \ \cdots 1 \leq i \leq m \tag{5.19}$$

の組合せであるので，定理1から，それぞれの $i$ に関して順に楕円近似法による更新を適用すればよい．しかし $\boldsymbol{h}$ は非線形であるので，式3.2の領域は式5.2で表現できない．

そこでEKFと同様に予測楕円の中心点で $h_i$ を局所的に線形化して

$$\left| y_{t,i} - \frac{\partial h_i}{\partial \boldsymbol{x}_t}(\boldsymbol{x}_t - \widetilde{\boldsymbol{x}_t}) - h_i(\widetilde{\boldsymbol{x}_t}) \right| \leq \Delta y_i \tag{5.20}$$

と表現したいが，次のような問題がある．

[3]観測$\boldsymbol{y}$の誤差の各成分の間に相関がある場合は式3.2は余分な領域を多く含んでしまう．これに関する議論は5.5節と付録Eを参照．

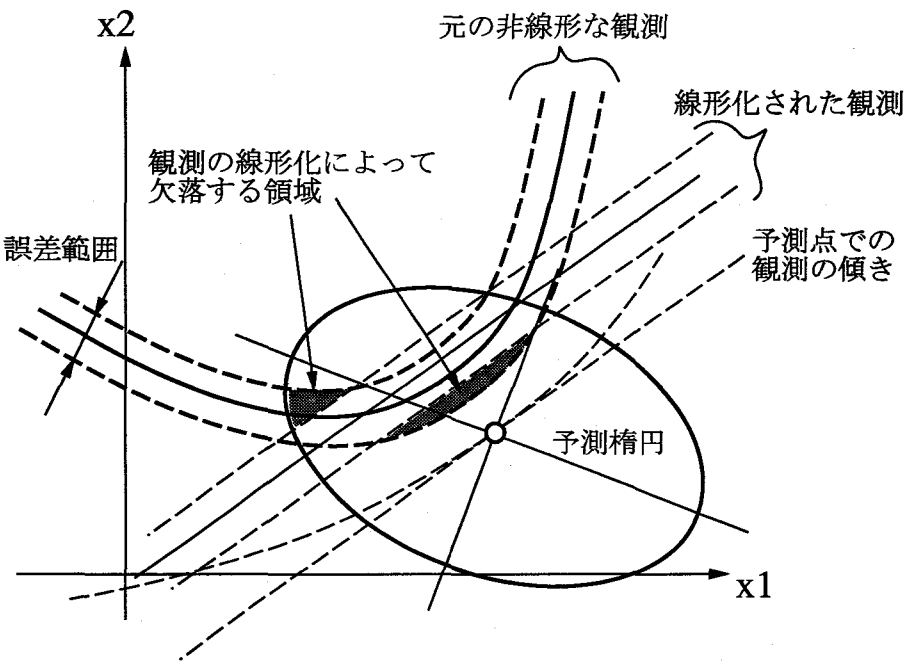



図 **5.3**: 観測を線形化した場合の領域欠落

図 5.3に示すように，元の非線形な観測式を満たす領域と線形化された観測を満たす領域が異なることである．図の網かけの領域は初期楕円内でかつ観測を満たす領域であるにもかかわらず，線形化の影響で欠落してしまう．

これを防ぐため本研究では図 5.4に示すように，非線形性を考慮して，線形化した観測の誤差範囲を広く見積もる．もっとも誤差が大きくなるのは，図の濃い網かけの領域のうち，線形化した観測による拘束平面

$$y_{t,i} - \frac{\partial h_i}{\partial \boldsymbol{x}_t}(\boldsymbol{x}_t - \widetilde{\boldsymbol{x}_t}) - h_i(\widetilde{\boldsymbol{x}_t}) = 0 \tag{5.21}$$

までの距離が最大のところである．そこで，その点を通るよう式 5.20を次のように変更する．

$$\left| b_{t,i} - \frac{\partial h_i}{\partial \boldsymbol{x}_t}\boldsymbol{x}_t \right| \leq d_{t,i} \tag{5.22}$$

式 5.22の $b_{t,i}$,$d_{t,i}$ を計算するために，予測楕円の近傍で元の非線形な観測を満たす領域から点をサンプリングし，線形化した観測による拘束平面（式 5.21）までの符号付距離

$$e_i = y_{t,i} - \frac{\partial h_i}{\partial \boldsymbol{x}_t}(\boldsymbol{x}_t - \widetilde{\boldsymbol{x}_t}) - h_i(\widetilde{\boldsymbol{x}_t}) \tag{5.23}$$

の最大値 $e_i^{max}$ と最小値 $e_i^{min}$ を探す．これから

$$b_{t,i} = (e_i^{max} + e_i^{min})/2 \tag{5.24}$$

$$d_{t,i} = (e_i^{max} - e_i^{min})/2 \tag{5.25}$$

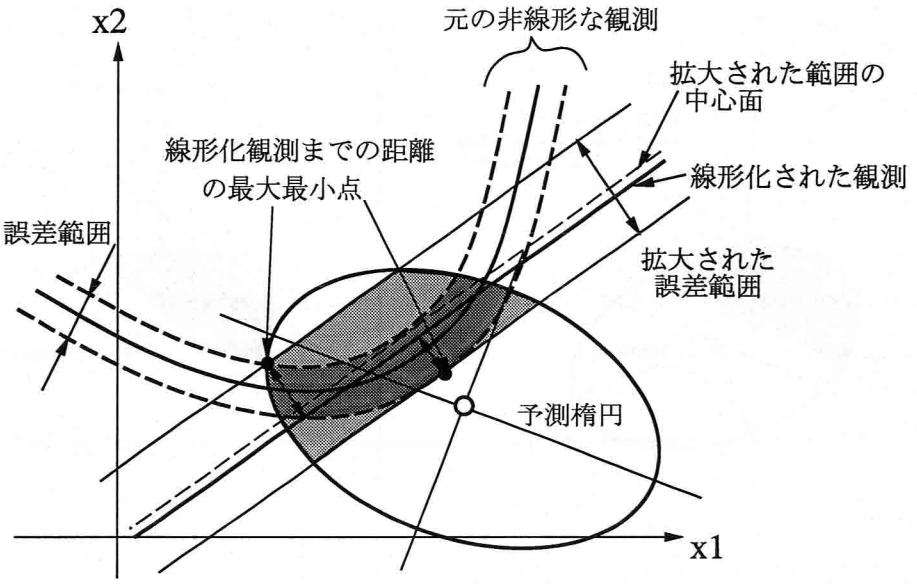



図 5.4: 観測の非線形性の見積もり

として計算できる．この処理は $\boldsymbol{x}$ の次元が高くなると計算量が非常に大きくなるが，これに対する対処は5.5節で述べる．

さて，このようにして観測の誤差を修正した観測による制約（式5.22）を用い，楕円近似法によってパラメータ領域を更新する．更新された楕円領域は定理1により必要な領域をすべて含むようになる．しかし同時に余分な領域も含んでいる．より精度よく推定するためにはできるだけ余分な領域を含まないような楕円近似が望ましい．

ところが，図5.5に示すように，予測楕円の中心点ので線形化を行うと，余分な領域が多く楕円に含まれてしまうことがある．これは予測楕円の中心点は観測による制約を満たしていないので，そこで線形化してしまうと拘束平面の傾き $\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t}$ が正確でないからである．そこで本研究ではIterated EKF[37] に基づく繰り返し計算によって観測制約の適切な線形化位置を求める（付録D参照）．

## 5.3 楕円近似による予測

本研究で対象にする関節物体は動物体であるから，ある時刻 $t$ での推定

$$\widehat{\boldsymbol{x}}_t, \boldsymbol{P}_t, \boldsymbol{x}_t^{min}, \boldsymbol{x}_t^{max} \tag{5.26}$$

から時刻 $t+1$ での状態を推定するには，運動による状態の変化を考慮する必要がある．本節では予測過程によるパラメータ領域の更新について述べる．ダイナミクスはカルマンフィ

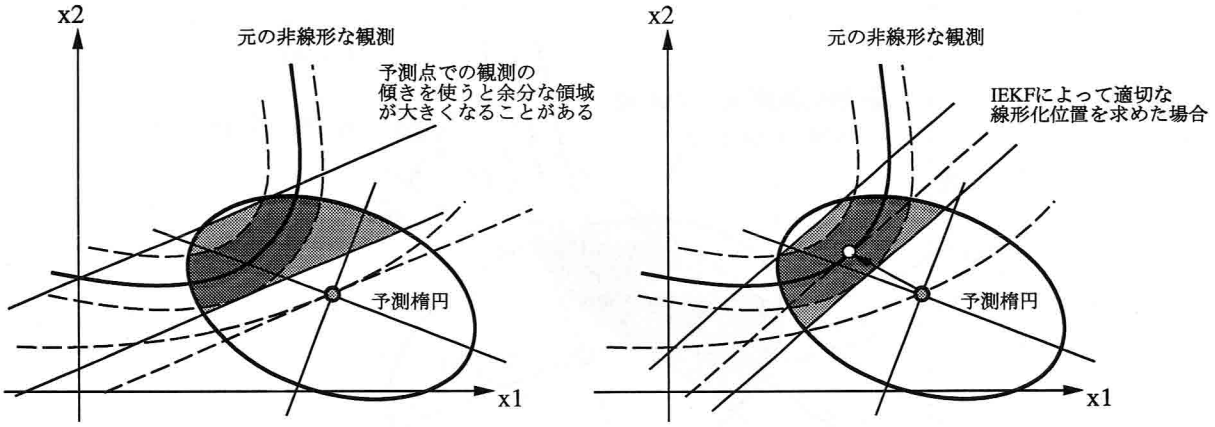



図 **5.5**: 線形化する位置による楕円近似の精度の違い

ルタの場合 4.2と同様に線形式

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{u}_t \tag{5.27}$$

であたえられるとする．$\boldsymbol{A}$ は，前時刻の状態を予測とするときは単位行列，等速度予測を用いるときは式 4.3になる．

カルマンフィルタでは状態ベクトルと予測誤差が無相関であるという仮定から，予測過程での分散の更新を，推定分散と予測誤差の分散の和で表現した．一方，楕円近似法において前時刻の推定から予測される領域は，予測誤差のとりうる領域を前時刻の推定パラメータ領域で掃引した図形になる．しかし前時刻の推定領域と予測誤差の領域がともに楕円だったとしても，予測される領域は正確には楕円にはならない．

そこで予測される領域を楕円で近似することにする．もし予測誤差が等方的ならば，前時刻での推定楕円の主軸方向に分解して考えることができるので，予測誤差のとりうる領域を球

$$\boldsymbol{u}_t^T \boldsymbol{u}_t \leq \Delta u^2 \tag{5.28}$$

と仮定しよう[4]．すると予測されるパラメータ領域は，前時刻の推定楕円の主軸の長さを予測誤差球の半径分長くした楕円（図 5.6）

$$(\boldsymbol{x}_{t+1} - \widetilde{\boldsymbol{x}}_{t+1})^T \boldsymbol{M}_{t+1} (\boldsymbol{x}_{t+1} - \widetilde{\boldsymbol{x}}_{t+1}) \leq 1 \tag{5.29}$$

$$\widetilde{\boldsymbol{x}}_{t+1} = \boldsymbol{A}\widehat{\boldsymbol{x}}_t \tag{5.30}$$

[4]予測誤差が等方でないときは誤差範囲を楕円で近似する．この場合適当な線形変換によって等方の場合に帰着できる．

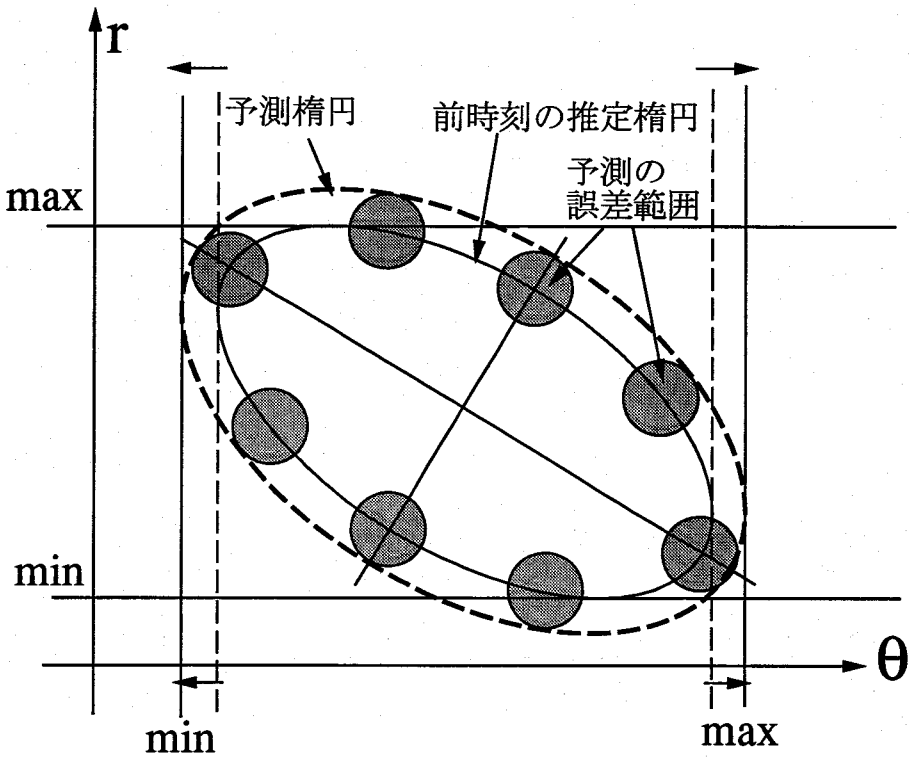



図 5.6: 予測される領域

$$\boldsymbol{M}_{t+1} = \boldsymbol{P}_t + \Delta u \boldsymbol{I} \tag{5.31}$$

によって近似できる．この近似楕円領域が本来の予測される領域を含むように予測誤差 $\Delta u$ を大きめにとるようにする．

予測楕円が得られたら，図 5.6のように上限値ベクトルと下限値ベクトルを上で得られた予測楕円の端を指すように変更する．しかしリンク長は時間不変なので予測誤差は 0 である．つまり各リンク長の上限値と下限値は実際には変化しないから，上述の等方的予測誤差の仮定は成立しない．そこで，

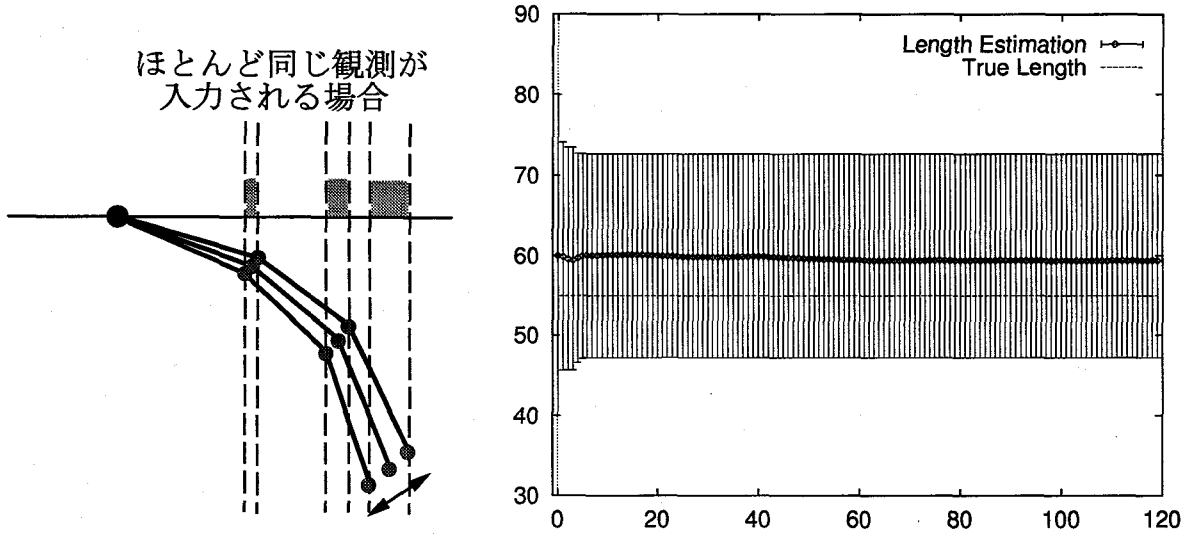
$$x_{t+1,i}^{min} = \begin{cases} x_{t,i}^{min} - \Delta u & \cdots \quad i: \text{角度} \\ x_{t,i}^{min} & \cdots \quad i: \text{リンク長} \end{cases} \tag{5.32}$$

$$x_{t+1,i}^{max} = \begin{cases} x_{t,i}^{max} + \Delta u & \cdots \quad i: \text{角度} \\ x_{t,i}^{max} & \cdots \quad i: \text{リンク長} \end{cases} \tag{5.33}$$

とガイドを更新したうえで，以下の平行な 2 平面で囲まれた領域

$$|(x_{t,i}^{max} + x_{t,i}^{min})/2 - x_{t+1,i}| \leq (x_{t,i}^{max} - x_{t,i}^{min})/2 \quad \cdots i: \text{リンク長} \tag{5.34}$$

と予測楕円を 5.1節で述べた計算方法で統合することにより修正しておく．




(a) 類似した観測の連続入力

(b) $r_3$ の推定結果

図 5.7: 類似の観測を長い時間に渡って入力したときの楕円近似法による推定結果

## 5.4 実験1

ここで，観測の非線形性を考慮した楕円近似法による推定法によって，4.8節で指摘した拡張カルマンフィルタによる推定法の問題点を解決できることを，実験によって示す．

図 5.7は，4.8節での実験と同じく，類似の観測が長い時間に渡って入力された場合の，楕円近似法によるリンク長の推定結果のグラフである．拡張カルマンフィルタでは，推定分散が誤って小さくなってしまった (図 4.15)．しかし楕円近似法では，推定開始直後は新しい情報が入力されるので推定範囲が小さくなるが，その後は類似の観測しか入力されないので，推定範囲は完全に保存されているのがわかる．

## 5.5 モード分解による計算量の減少

再び図 3.1の関節物体について考える．また以下では線形な画像投影写像を考え，写像行列を $\boldsymbol{L}$ とする．リンクの端点と関節位置の観測方程式 $\boldsymbol{h}(\boldsymbol{x}) = \{h_i(\boldsymbol{x})\}$ は

$$h_0(\boldsymbol{x}) = \boldsymbol{L}\boldsymbol{t} \tag{5.35}$$

$$h_i(\boldsymbol{x}) = \boldsymbol{L}\left[\boldsymbol{R}(\boldsymbol{\omega})\sum_{j=1}^{i}\left\{r_j\cos(\sum_{k=1}^{j}\theta_k), r_j\sin(\sum_{k=1}^{j}\theta_k), 0\right\}^T + \boldsymbol{t}\right] \quad \cdots\ 1 \le i \le 3 \tag{5.36}$$

$$\boldsymbol{x} = (\boldsymbol{\omega}, \boldsymbol{t}, \theta_1, \theta_2, \theta_3, r_1, r_2, r_3)^T \tag{5.37}$$

と表現される．なお $h_0$ は関節物体のもっとも根元の位置の観測である．$\omega$ は3次元，$\boldsymbol{L}$ として正射影を考えれば $\boldsymbol{t}$ は2次元だから，$\boldsymbol{x}$の次元は11次元になる．5.2節で述べた観測の非線形を補償するプロセスでは予測楕円の近傍で点のサンプリングが必要になる．しかし11次元の空間では，各軸方向に5点ずつのサンプリングでも $5^{11} \sim 5 \times 10^7$ 個の点について式5.23を計算することになり，計算量が膨大になってしまう問題点がある．

そこで次のような変数変換を考える．

$$\theta_i^* = \sum_{j=1}^{i} \theta_j \tag{5.38}$$

$$y_i^* = y_i - y_{i-1} \quad \cdots \quad i \neq 0 \quad (h \text{ も同様}) \tag{5.39}$$

$$y_0^* = y_0 \quad (h \text{ も同様}) \tag{5.40}$$

上式は線形な変換なので行列 $\boldsymbol{D}$,$\boldsymbol{E}$ を用いて以下のように書ける．

$$\boldsymbol{x}^* = \boldsymbol{D}\boldsymbol{x} \tag{5.41}$$

$$\boldsymbol{y}^* = \boldsymbol{E}\boldsymbol{y} \tag{5.42}$$

$$\boldsymbol{h}^* = \boldsymbol{E}\boldsymbol{h} \tag{5.43}$$

この時予測楕円の式5.31は

$$(\boldsymbol{x}_t^* - \boldsymbol{D}\tilde{\boldsymbol{x}}_t)^T(\boldsymbol{D}\boldsymbol{M}_t\boldsymbol{D}^T)^{-1}(\boldsymbol{x}_t^* - \boldsymbol{D}\tilde{\boldsymbol{x}}_t) \leq 1 \tag{5.44}$$

に写り，式5.36は以下のようになる．

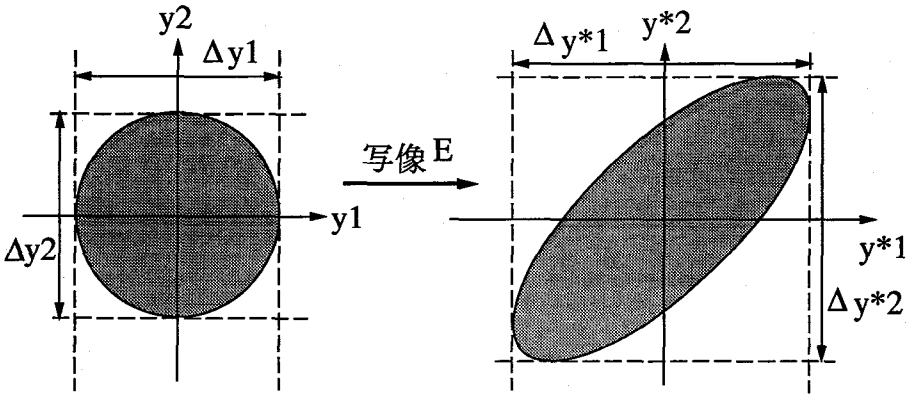
$$h_0^*(\boldsymbol{x}) = \boldsymbol{L}\boldsymbol{t} \tag{5.45}$$

$$h_i^*(\boldsymbol{x}) = \boldsymbol{L}\boldsymbol{R}(\boldsymbol{\omega})\left\{r_i\cos(\theta_i^*), r_i\sin(\theta_i^*), 0\right\}^T \quad \cdots \quad 1 \leq i \leq 3 \tag{5.46}$$

上式を見ると，$h_0$ には $\boldsymbol{t}$，$h_i$ には $(\boldsymbol{\omega}, r_i, \theta_i^*)$ しか含まれていない．すなわちモードに分解されている．したがって，式5.43によって変換された空間において楕円近似法を適用すれば，それぞれの $h_i^*$ について式5.23を計算する際にはたかだか5次元でサンプリングすればよい．よって，$h_i^*$ を順番に予測楕円に統合していく．

このとき，観測誤差の各成分 $\Delta y_i$ が無相関であっても，変換された観測誤差 $\Delta y_i^*$ は相関を持ってしまう．すると $h_i$ を予測楕円に対して順番に統合した場合，観測の満たす領域は図5.8に破線で示す矩形領域となり，余分な領域を多く含んでしまう．余分な領域を少なくするには $\Delta y_i^*$ の領域の主軸方向に $\boldsymbol{y}$,$\boldsymbol{h}$ を変数変換すればよいが，そうすると今度は複数のモード $(\boldsymbol{\omega}, r_i, \theta_i^*)$ を同時に扱わないといけないので計算量が増加してしまう[5]．

[5]複数のモードを同時に扱う手法については付録Eを参照．




無相関　　相関あり

近似精度がよい　　近似精度がよくない

図 5.8: モード分解による観測の近似精度の悪化

モード分解した観測 $h_i^*$ を順番に適用した場合は，精度は落ちるが少なくとも必要な領域（網かけ）はすべて含んでいるので，以下本研究では上記のモード分解した観測式 $h_i^*$ を用いることにする．

式5.45は線形なのでそのまま楕円近似法に適用できる．一方，式5.46は非線形なので，5.2節で述べた非線形性の補償を適用する．$h_i^*(\boldsymbol{x}) = y_i^*$ の観測が得られたとき，式5.46は $r_i$ について陽に解ける．

$$r_i = y_i^* / \left[ \boldsymbol{LR}(\boldsymbol{\omega}) \{\cos(\theta_i^*), \sin(\theta_i^*), 0\}^T \right] \tag{5.47}$$

よって図5.9に示すように，$r_i$ 以外の４つの変数についてサンプリングを行えば，$r_i$ は式5.47で計算できる．各軸方向に５点ずつのサンプリングならば，$5^4 \times 2 = 1250$ 個のサンプリングになる．

観測の非線形性の補償の処理手順は以下のようになる．

1. $\boldsymbol{\omega}, \theta_i^*$ を予測楕円の中から等間隔にサンプリングする．

2. $y_i^* - \Delta y_i^*$ と $y_i^* + \Delta y_i^*$ の場合について式5.47から $r_i$ を計算する．

3. サンプリングされた点が予測楕円ならびに上限ベクトル $\boldsymbol{x}_t^{max}$, 下限ベクトル $\boldsymbol{x}_t^{min}$ で指定される矩形内に入るかどうかをチェックする．

4. すべてのサンプリング点の中から式5.23の最大値と最小値を求め，式5.22の $b_{t,i}$ を $d_{t,i}$ を計算する．

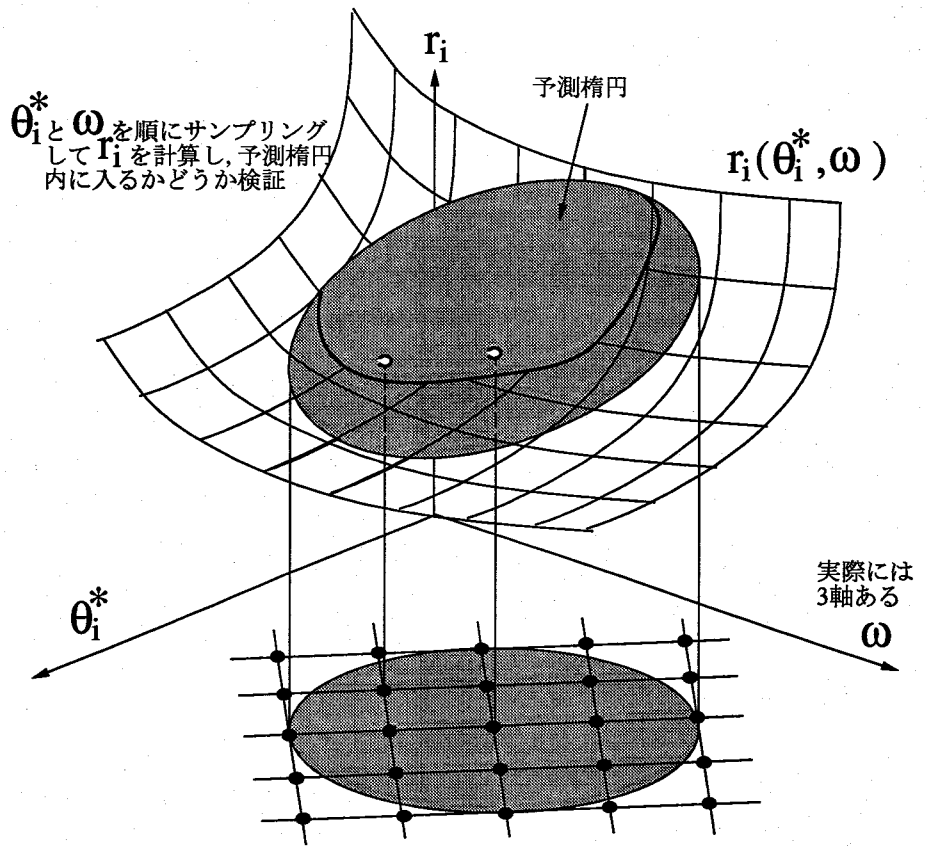



図 5.9: サンプリングによる観測の非線形性の補償

サンプリングが粗過ぎる場合は，最初のサンプリングで楕円内に入った点の近傍をさらに細かく階層的にサンプリングして，$b_{t,i}$，$d_{t,i}$ の精度を上げる．

以上のようにして観測の非線形性の補正に必要な計算量を減らす．

## 5.6 複数候補の生成

カルマンフィルタを用いた実装で見たように (4.6節)，単眼視では奥行きに曖昧性があるので，場合によっては一時的に解釈が複数に分かれることがある．本研究の楕円近似法では観測の非線形性を補償して可能なパラメータ領域を更新するので，複数の解釈を一つの楕円体でまとめて扱うことが出来る．しかし図 5.10に示すように，楕円近似したパラメータ領域には明らかに余分な領域が含まれてしまう．このような状態は，可能なパラメータ領域が複数の連続領域に分離したときに起きる．そこで本研究では，このような場合に複数の楕円をつかって近似する．

前節で述べたサンプリングの結果，予測楕円内に入った点を $(\omega, \theta_i^*)$ 空間でラベリングす

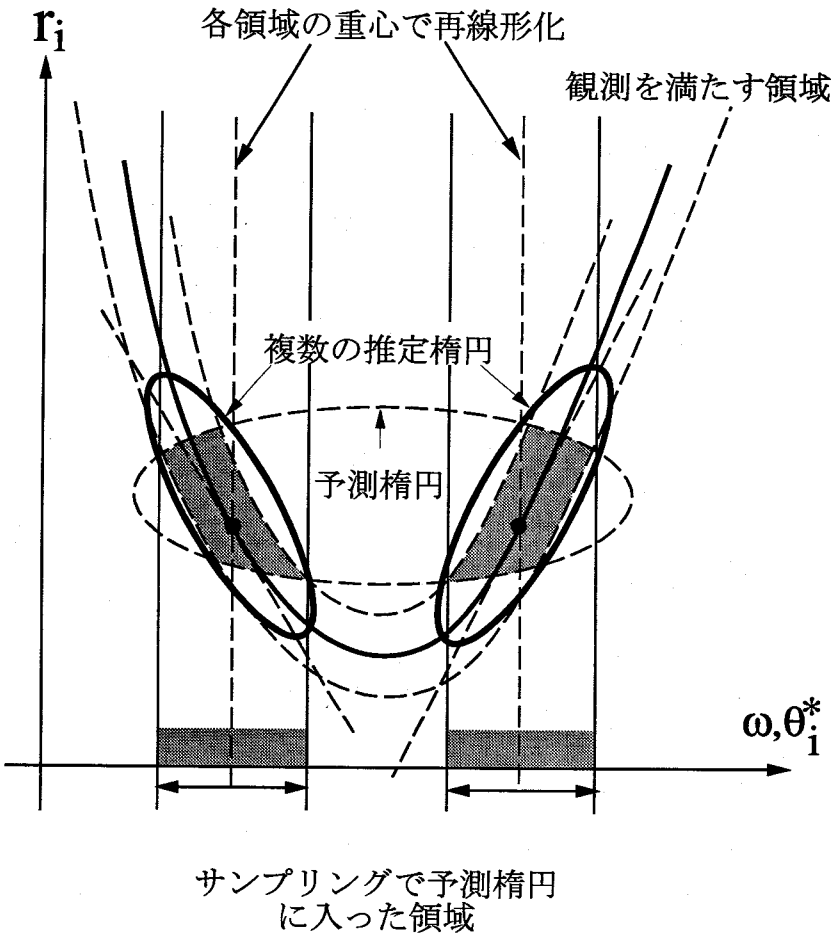



図 5.10: 複数候補の生成

る．このラベリングは多次元のラベリングになる[6]．もし得られた連結領域が２つ以上になれば，各連結領域に対して以下の分割処理を行なう．

1. 連結領域の重心をもとめ，その点で観測を線形化し直す．
2. 連結領域の中を再びサンプリングし，１で得られた線形化観測の式を用いて式5.23を計算，式5.22を作り直す．
3. 楕円近似法を用いて推定パラメータ領域を更新する．

上記の処理によって，複数の解釈が可能なときには推定パラメータ領域は複数に分割される．一度分割されても再び同じ領域になることがあるので，次の統合処理も行なう．

1. 二つの推定領域のうち，一方の楕円と $\boldsymbol{x}_t^{min}, \boldsymbol{x}_t^{max}$ による矩形領域が，他方に完全に含まれるときは含まれるほうを消去する．

[6] ２次元での８近傍は，４次元では80近傍に相当する．

表 5.1: 実験で用いた制約知識と初期値 (2)

| | | |
|---|---|---|
| 姿勢 | 制約 | $\lvert\theta_2 - \theta_3\rvert \leq \pi/6$ |
| | | $0 \leq \theta_2, \theta 3 \leq \pi/2$ |
| | 初期範囲 | 真値 ±0.1rad |
| 形状 | 制約 | $\lvert r_1 - r_2\rvert \leq 15$ |
| | | $\lvert r_2 - r_3\rvert \leq 15$ |
| | | $60 \leq r_1 \leq 80$ |
| | | $50 \leq r_2 \leq 70$ |
| | | $40 \leq r_3 \leq 60$ |
| | 初期範囲 | 真値 ±20pixels |
| 観測 | 最大誤差 | 3pixels |

2. 二つ以上の推定領域が存在する場合，予測楕円と観測を満たす領域の積集合が空集合になったときはこれを誤った推定とみなして消去する．

3. 一つしか推定領域がないときに，予測楕円と観測を満たす領域の積集合が空集合になったときは観測が誤りであったとみなして，前時刻での推定領域のまま維持する．

以上の処理の結果，複数の推定領域が残った場合は，次の時刻での予測や観測の統合をそれぞれの推定領域に対して並列に実行する．

## 5.7 実験 2

最後に楕円近似法による関節物体の形状と姿勢の推定例をコンピュータシミュレーションで示す．３章での実験と同じく，簡単のため２次元の３関節物体について実験した．物体の１番根元の関節は原点に固定されているとした．観測として各関節と端点の位置が１次元で得られると仮定し，実験では真の物体関節を制約を満たしながらさまざまに回転させ，各関節の投影位置に観測ノイズを加えて生成した．観測ノイズは白色ガウス分布に従うノイズを乱数により発生させた．推定に利用した物体に関する制約知識ならびに観測誤差の大きさを表 5.1に示す．

### 5.7.1 楕円近似法による形状と姿勢の推定例

まず，楕円近似法によって関節物体の形状と姿勢推定を行なった例を画像とグラフで示す．図 5.11は各時刻における関節物体の真の形状と姿勢 ((a),(b),(c),(g),(h),(i)) ならびに推

定された形状と姿勢 ((d),(e),(f),(j),(k),(l)) を示したものある．２次元の物体を考えているので，図の横線は画面をあらわし，縦方向は奥行き方向である．縦線は観測された関節位置の座標（１次元）に対応している．図 5.11の (a) と (d) に示すように，初期推定におけるリンク長と関節角度は真値と異なっているにもかかわらず，観測を繰り返していくと次第に推定が正確に推定できる．

図 5.11の推定例について，各パラメータの推定範囲（上限値と下限値）をグラフに示したのが図 5.12， 5.13， 5.14， 5.15， 5.16， 5.17である．

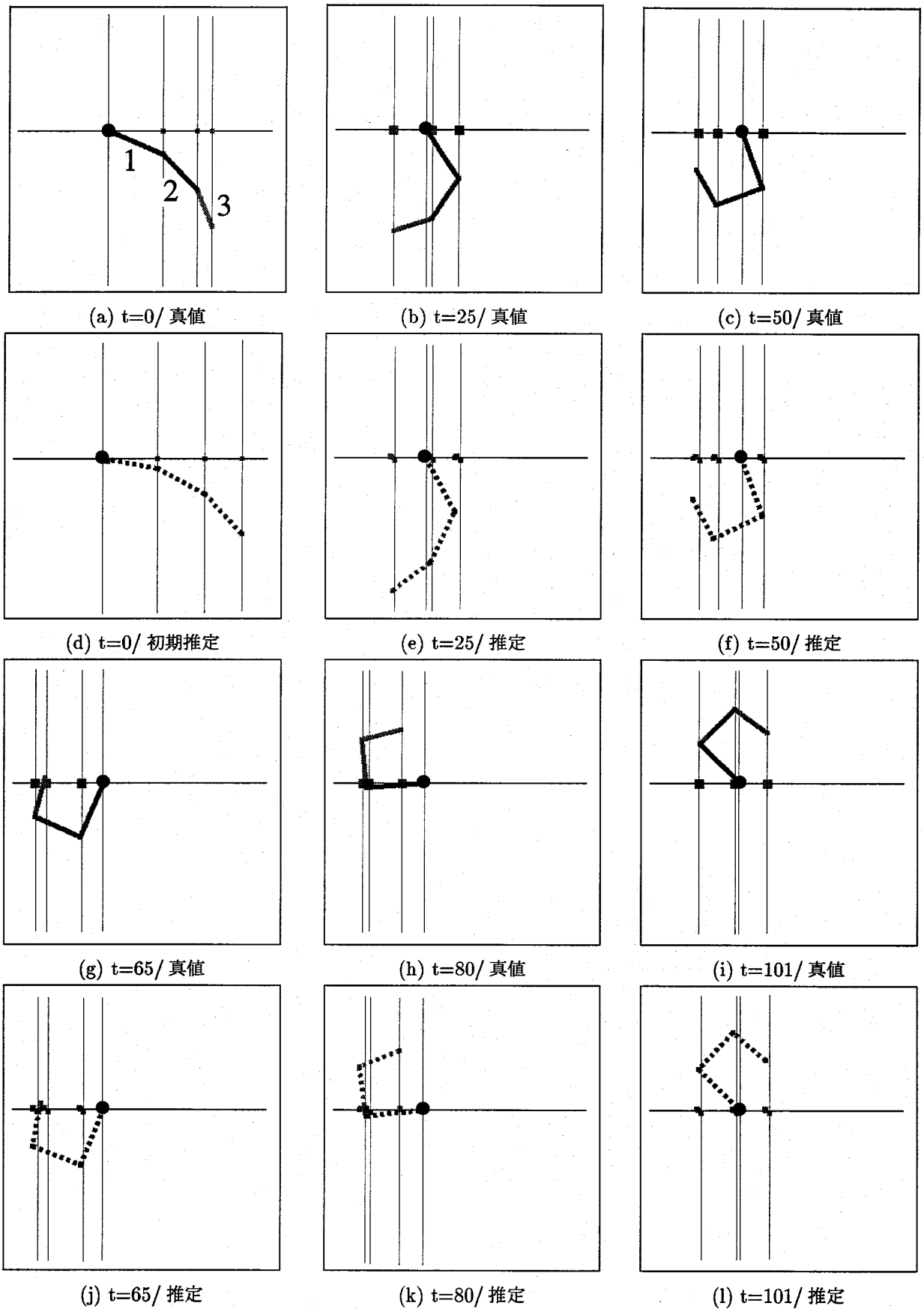
図 5.12， 5.14， 5.16の縦軸は式 5.38で与えれるモード分解された関節角度 $\theta^*$ [rad] である．各推定値が実線で示す真値と比べて正しく推定されていることがわかる．また，$\theta^*$ が $n\pi(n = 0, \pm 1, \cdots)$ に近いところでは推定された範囲が広がっている．これはリンクが画面に対して平行に近くなったため，画面に垂直なときに比べて角度の推定精度が落ちたと考えられる．

一方，図 5.13， 5.15， 5.17の縦軸は各リンク長 [pixel] である．本手法では形状と姿勢に関する制約を利用して推定を行なうので，単眼視にも関わらず可能なリンク長の範囲が限定されていくのがわかる．また，各リンクとも推定されたリンク長の範囲は ±5pixel 程度に近付いていくが，これは観測の誤差 ±3pixel よりも大きい．これは 5.2節で述べた非線形性の補償を行なう際の近似誤差のために観測の持つ情報を完全に利用できないためと考えられる．

### 5.7.2 楕円近似法における複数候補の生成例

次に，姿勢の解釈に曖昧さがある場合について，複数の候補を保持しながら推定を行なった例を図 5.18に示す．図 5.18(a),(b),(c),(g),(h) は真の形状と姿勢を示し， (d),(e),(f),(i),(j) はそれらに対応する推定結果である．

(d) では第３リンク（一番末端のリンク）が画面に対してほぼ平行であったので次のフレームではリンクが奥（画面上方）へ移動したのか，手前（画面下方）へ移動したのかが曖昧になった．そこで， (e) ではそれらに対応する２つの推定候補が生成されているのがわかる．さらに (f) では第２リンクが画面に対してほぼ平行になったので，これについても二通りの候補が生成され， (i) では全部で３つの推定候補が並列に保持されている．しかし (j) に至ると，このうち２つの候補が除去されている．これは，表 5.1に示した姿勢に関する制約のうち， $|\theta_2 - \theta_3| \leq \pi/6$ と観測を満たすパラメータ領域の積集合が空集合になったためである．この結果 (j) では正しい推定候補が得られている．




(a) t=0/ 真値　(b) t=25/ 真値　(c) t=50/ 真値

(d) t=0/ 初期推定　(e) t=25/ 推定　(f) t=50/ 推定

(g) t=65/ 真値　(h) t=80/ 真値　(i) t=101/ 真値

(j) t=65/ 推定　(k) t=80/ 推定　(l) t=101/ 推定

**図 5.11**: 楕円近似法による関節物体の形状と姿勢の推定例

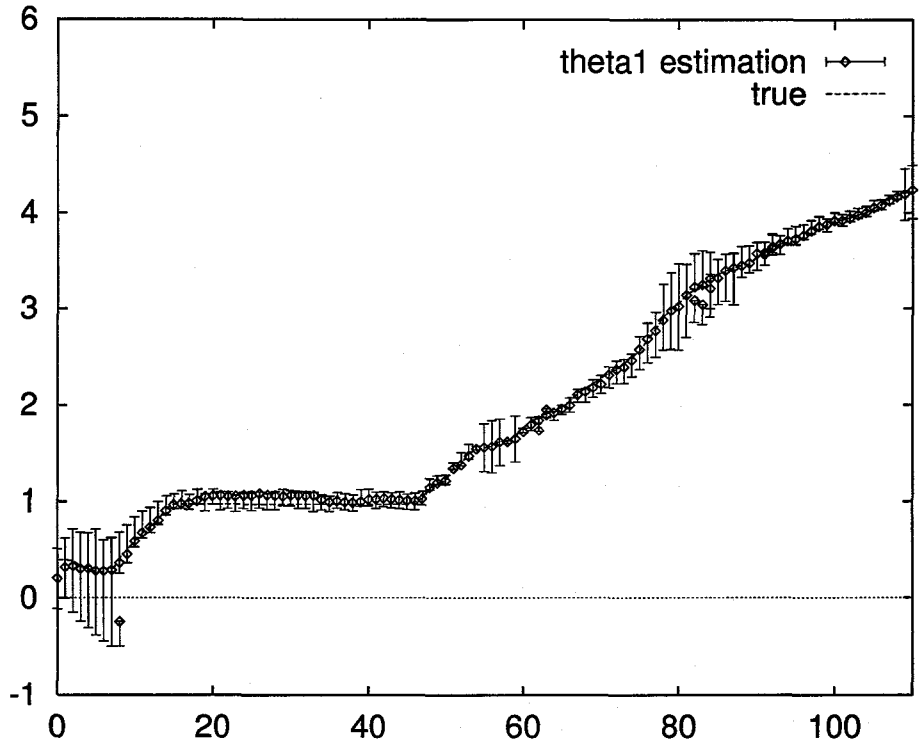



図 **5.12**: 楕円近似法による $\theta_1^*$ の推定範囲

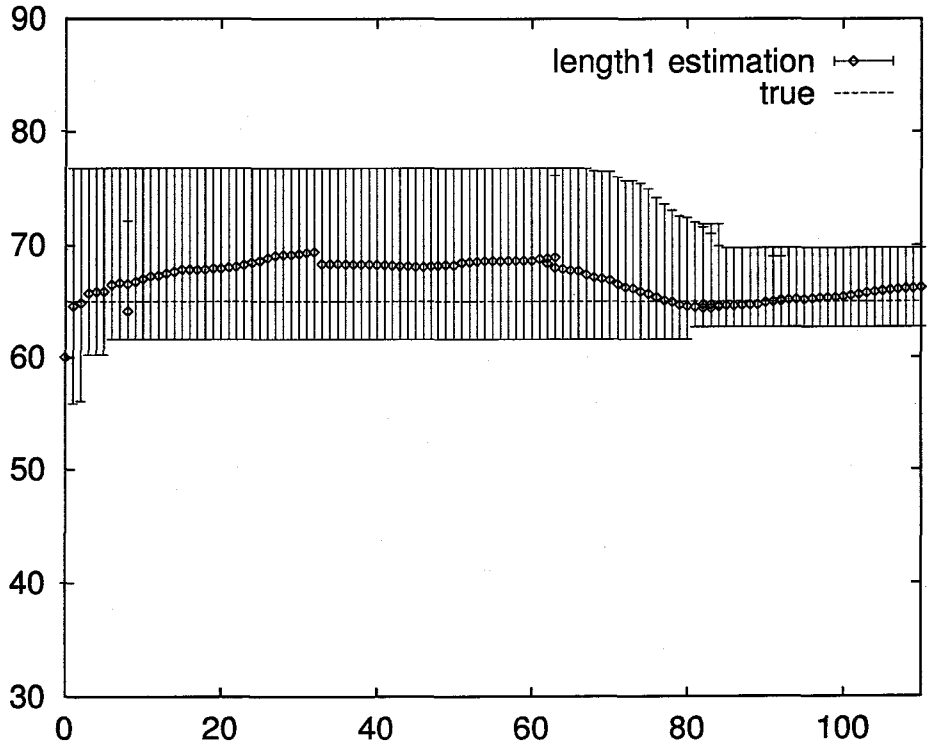



図 **5.13**: 楕円近似法による $r_1$ の推定範囲

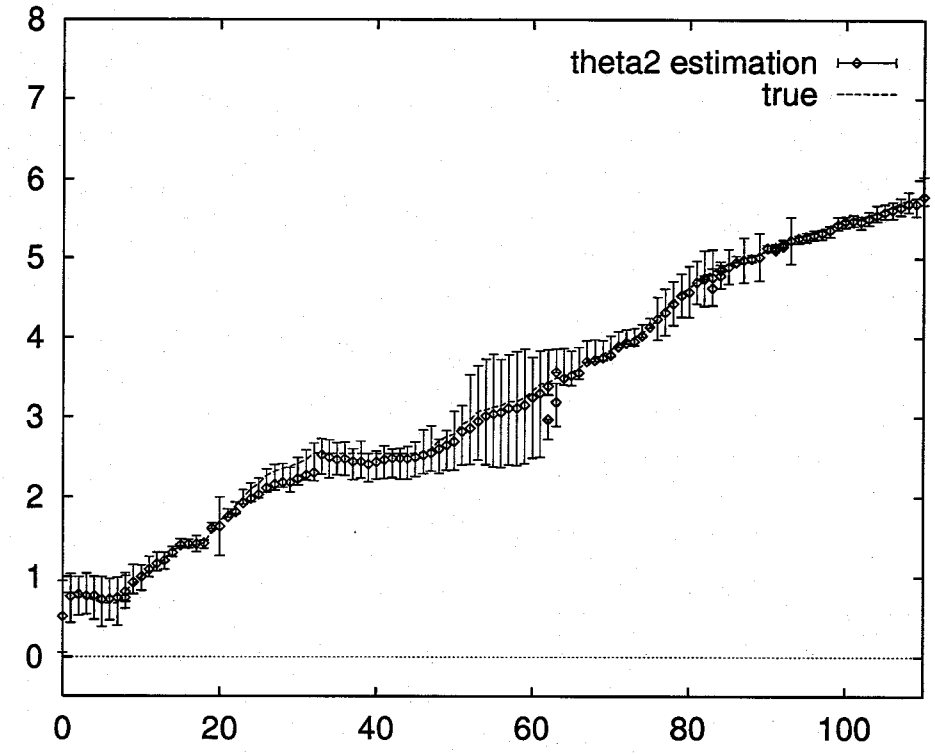



図 5.14: 楕円近似法による $\theta_2^*$ の推定範囲

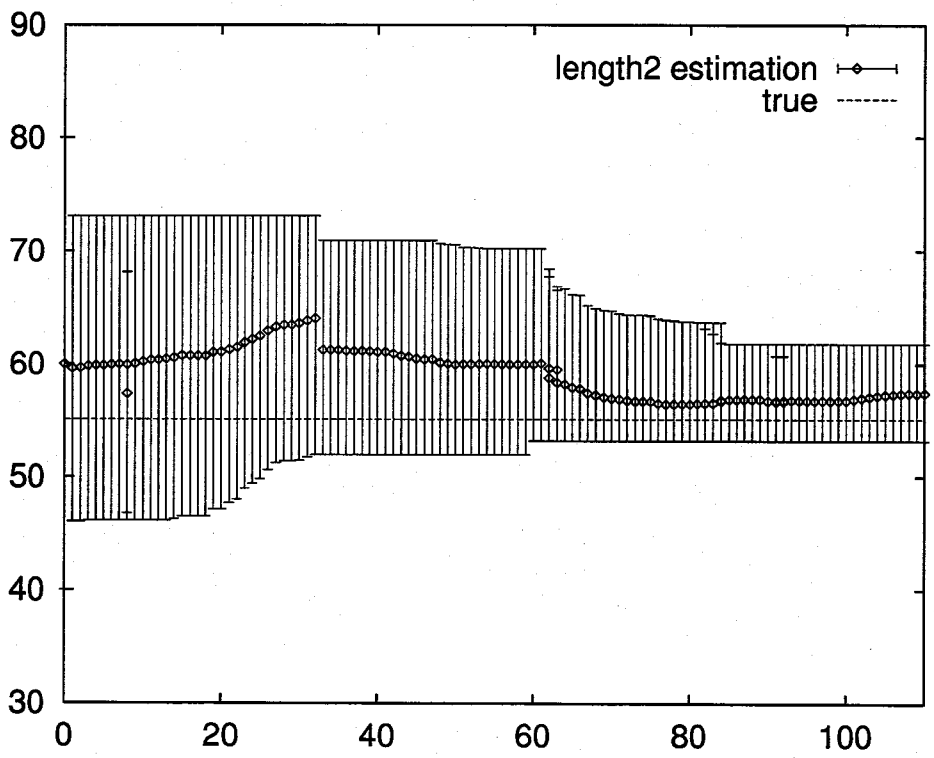



図 5.15: 楕円近似法による $r_2$ の推定範囲

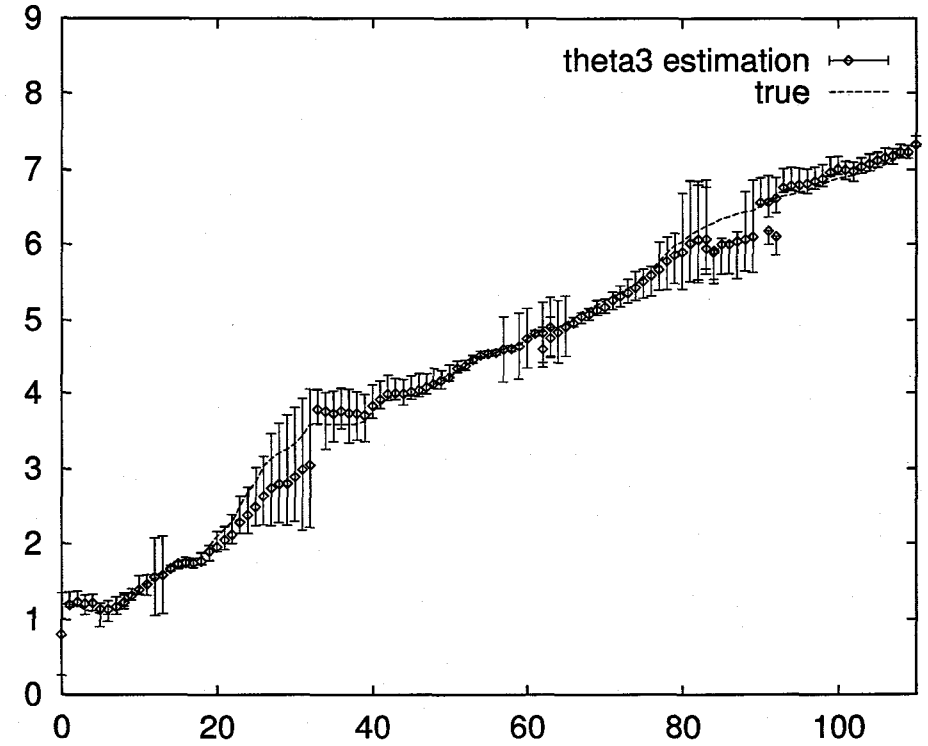



図 **5.16**: 楕円近似法による $\theta_3^*$ の推定範囲

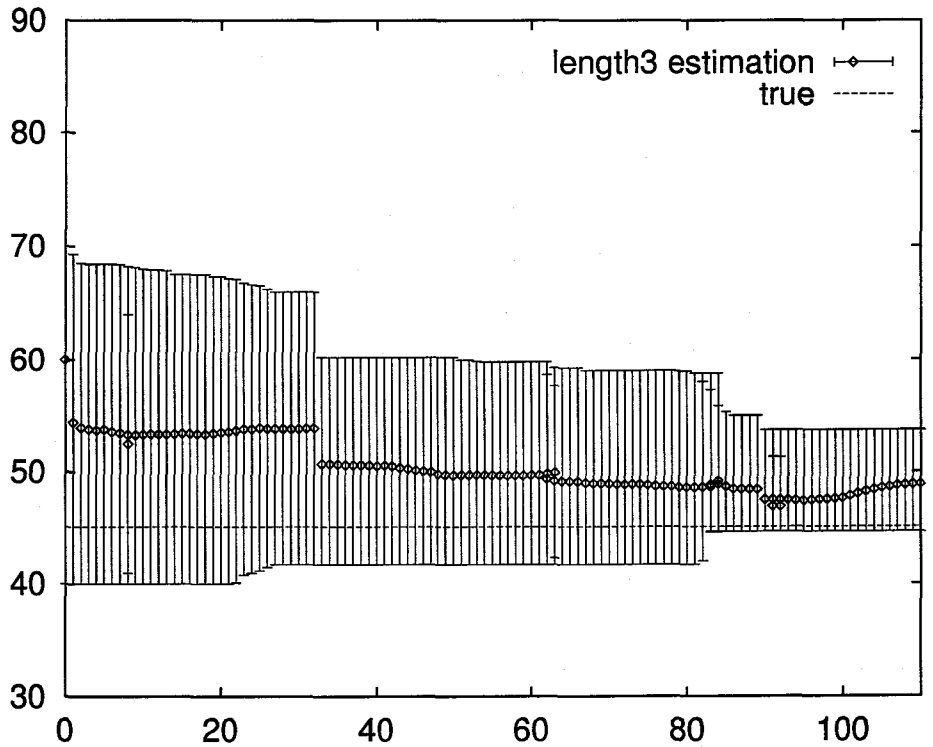
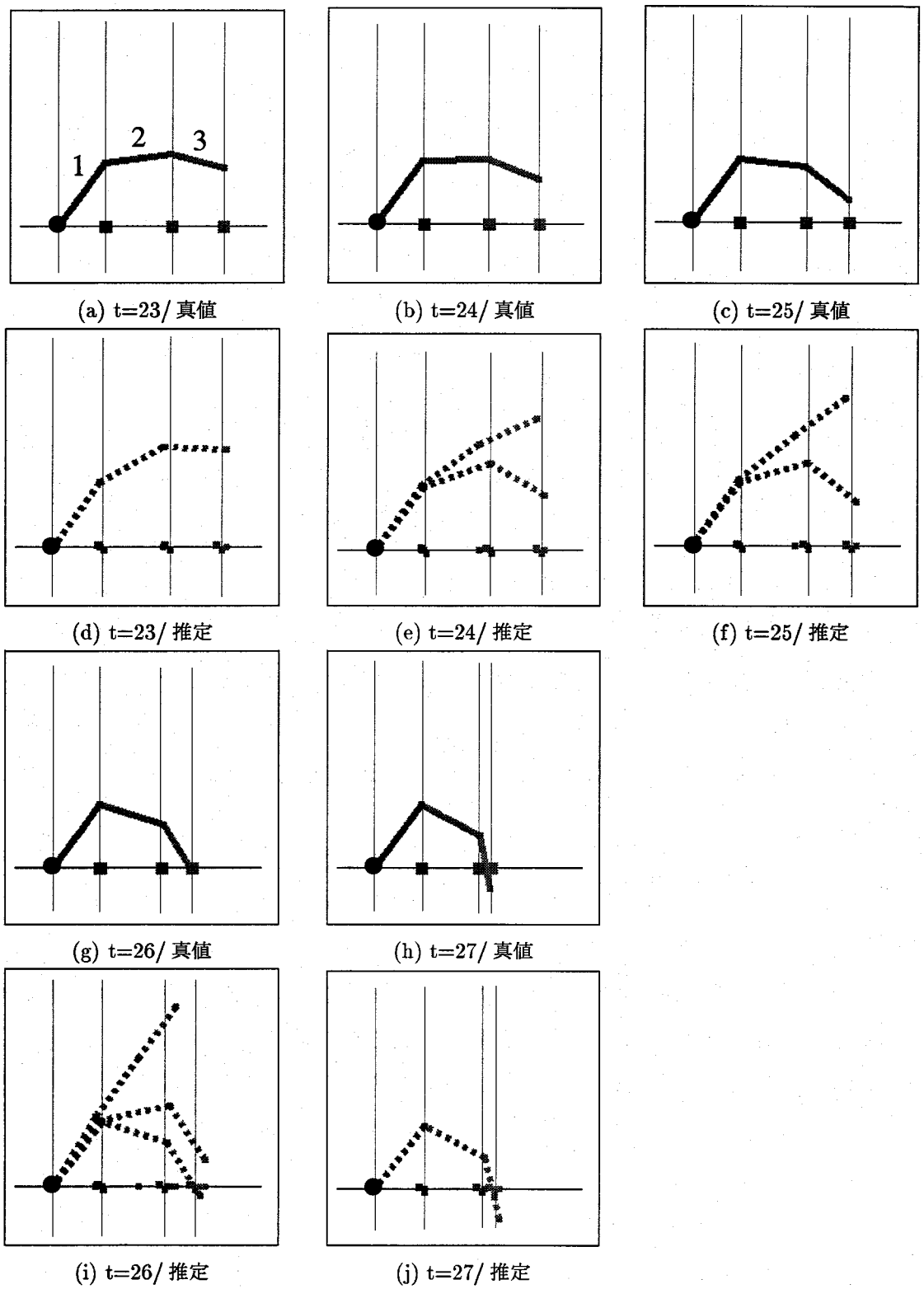



図 **5.17**: 楕円近似法による $r_3$ の推定範囲




(a) t=23/ 真値　(b) t=24/ 真値　(c) t=25/ 真値

(d) t=23/ 推定　(e) t=24/ 推定　(f) t=25/ 推定

(g) t=26/ 真値　(h) t=27/ 真値

(i) t=26/ 推定　(j) t=27/ 推定

**図 5.18**: 楕円近似法における複数候補の生成例

表 5.2: 大きさの異なる物体

| | $r_1$ | $r_2$ | $r_3$ |
|---|---|---|---|
| 物体 1 | 65 | 55 | 45 |
| 物体 2 | 75 | 65 | 55 |

表 5.3: 大きさの異なる物体に対する推定の比較実験における初期推定範囲

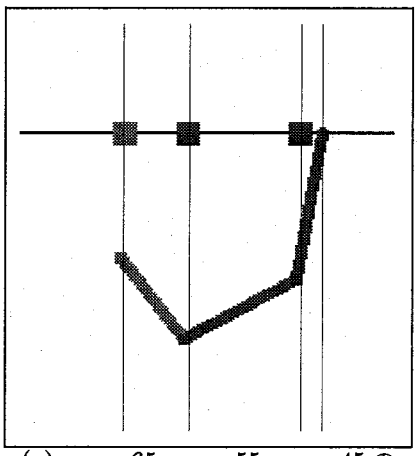
| 関節角度 | リンク長 |
|---|---|
| $0.1 \leq \theta_1 \leq 0.3$ | $60 \leq r_1 \leq 80$ |
| $0.3 \leq \theta_2 \leq 0.5$ | $50 \leq r_2 \leq 70$ |
| $0.2 \leq \theta_3 \leq 0.4$ | $40 \leq r_3 \leq 60$ |

### 5.7.3 大きさの異なる物体に対する推定の比較

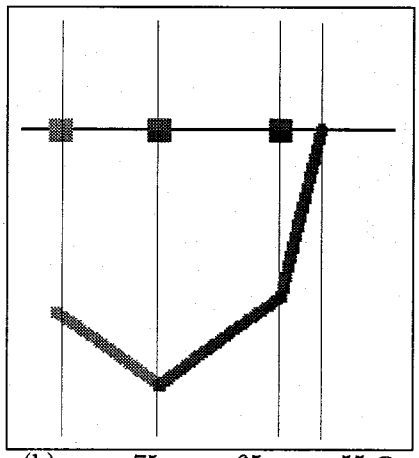
本節では，大きさが異なる２つの関節物体に対する推定を行なった場合の，結果の比較検討を行なう．実験では，表 5.2に示すリンク長が違う２つの二次元関節物体について，両者ともに表 5.1の制約を用い，初期推定範囲として表 5.3に示す範囲の範囲を与えて推定を行なった．

図 5.19はさまざまな観測を入力して上記の２つの物体に対して形状と姿勢の推定を行なった結果である．また，図 5.20, 5.21, 5.22はそれぞれの物体の，各時刻におけるリンク長 $r_1$,$r_2$,$r_3$ の推定範囲のグラフである．

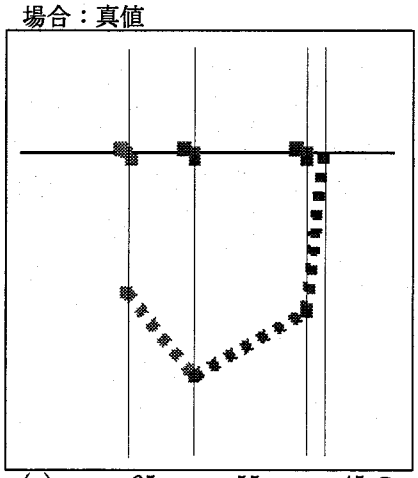
図とグラフから，同じ制約と初期値を用いたにも関わらず，両者の推定はおのおのの真値に近付いていくことがわかる．本手法は最初に与える初期推定だけではなく，入力されたさまざまな観測の情報を使って，物体の大きさを区別することができることが，この実験結果により示された．



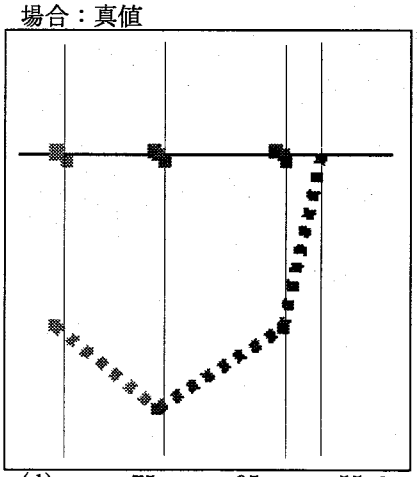
(a) $r_1 = 65, r_2 = 55, r_3 = 45$ の場合：真値



(b) $r_1 = 75, r_2 = 65, r_3 = 55$ の場合：真値

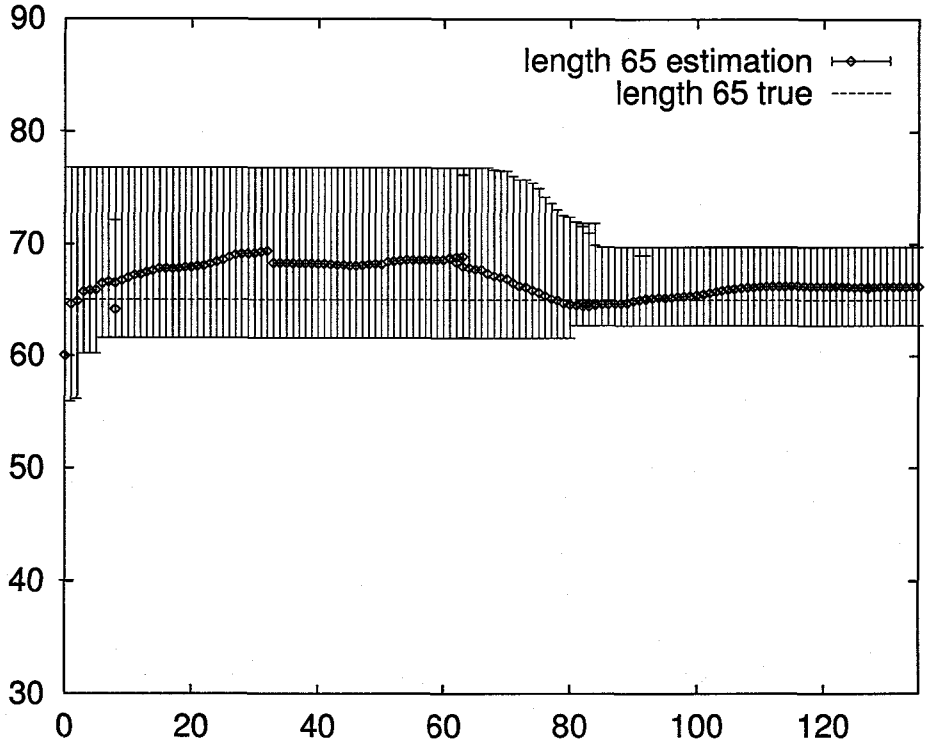


(c) $r_1 = 65, r_2 = 55, r_3 = 45$ の場合：推定

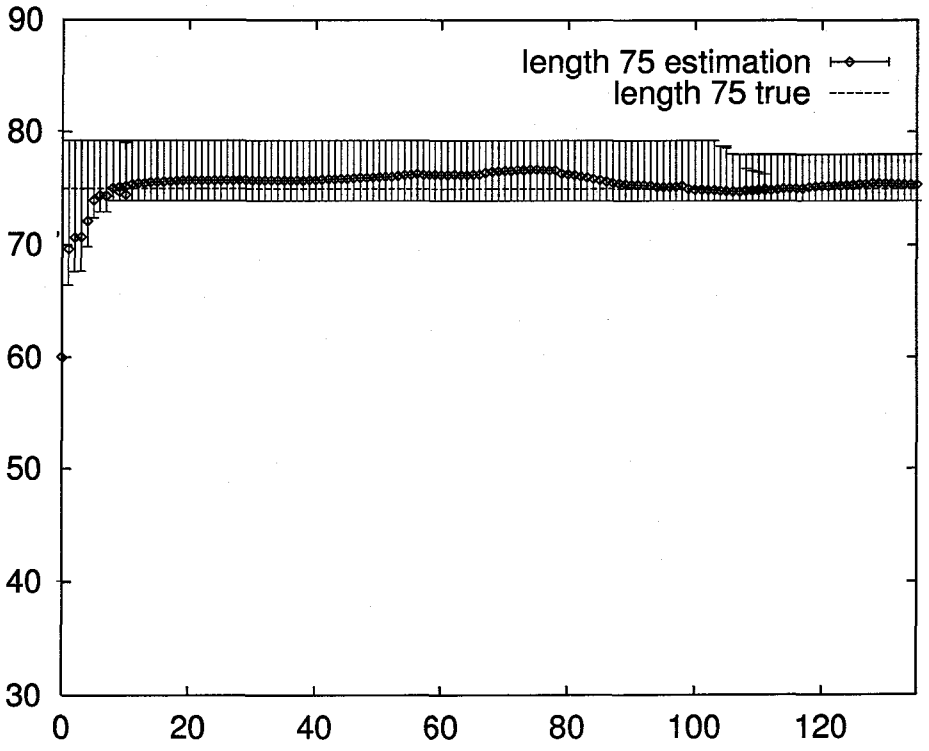


(d) $r_1 = 75, r_2 = 65, r_3 = 55$ の場合：推定

**図 5.19**: 大きさが異なる関節物体に対する推定結果の比較

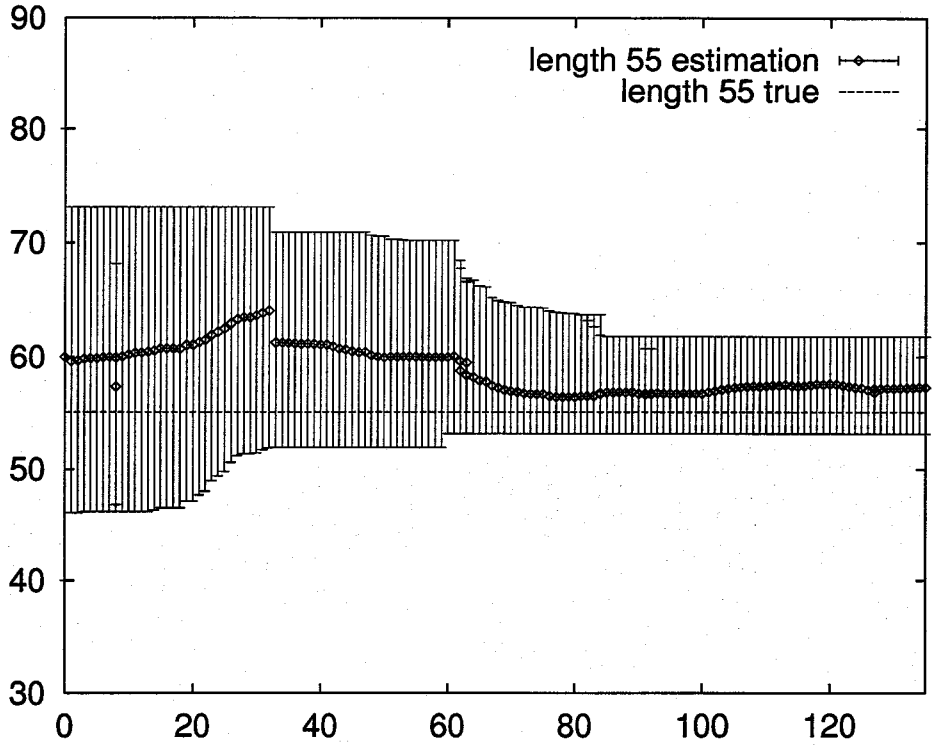



(a) $r_1 = 65$pixel の場合

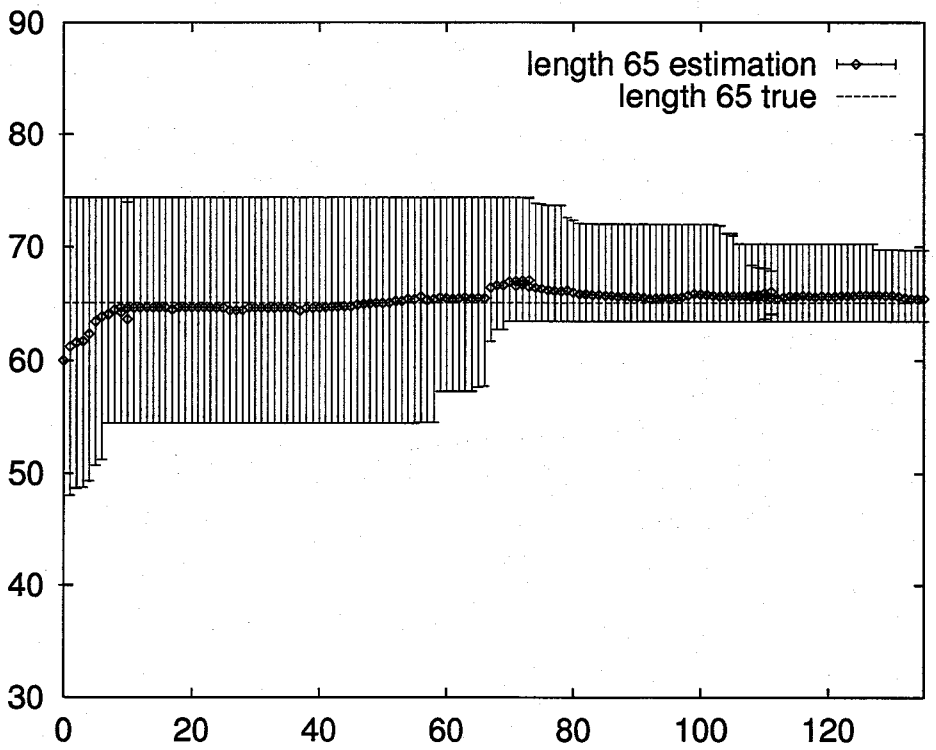



(b) $r_1 = 75$pixel の場合

図 **5.20**: 大きさが異なる物体に対する $r_1$ の推定範囲の比較

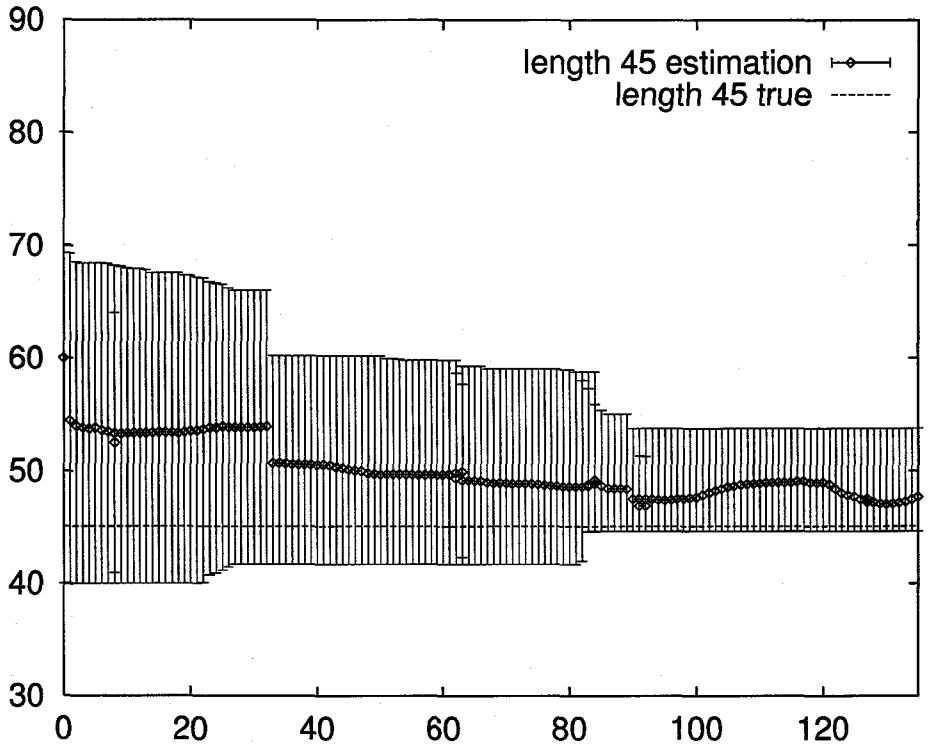



(a) $r_2 = 55$pixel の場合

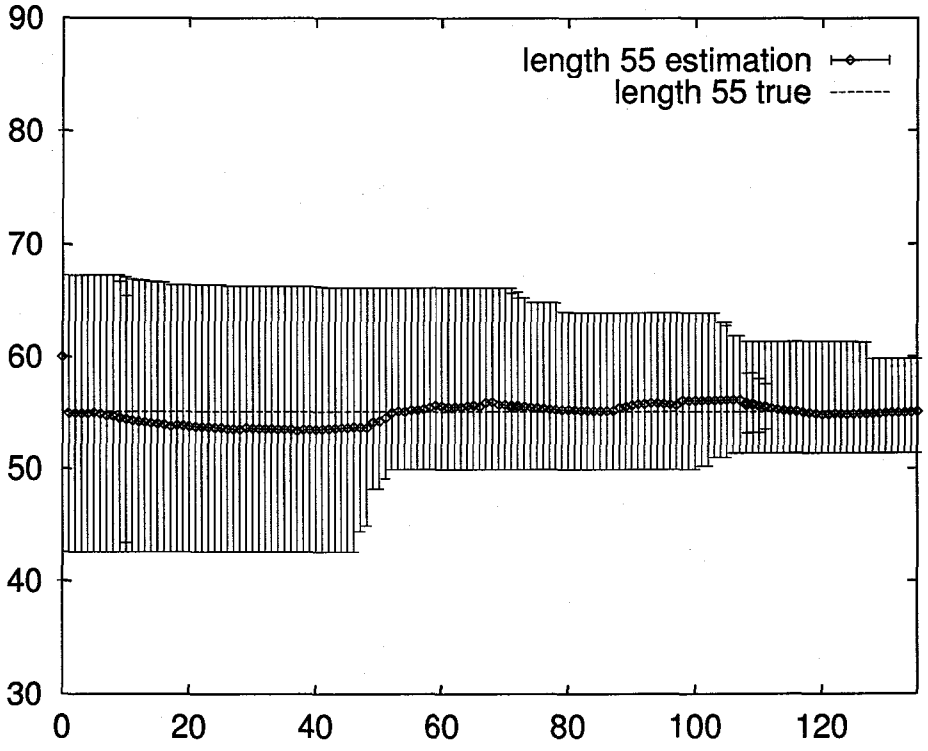



(b) $r_2 = 65$pixel の場合

**図 5.21**: 大きさが異なる物体に対する $r_2$ の推定範囲の比較




(a) $r_3 = 45$pixel の場合




(b) $r_3 = 55$pixel の場合

**図 5.22**: 大きさが異なる物体に対する $r_3$ の推定範囲の比較

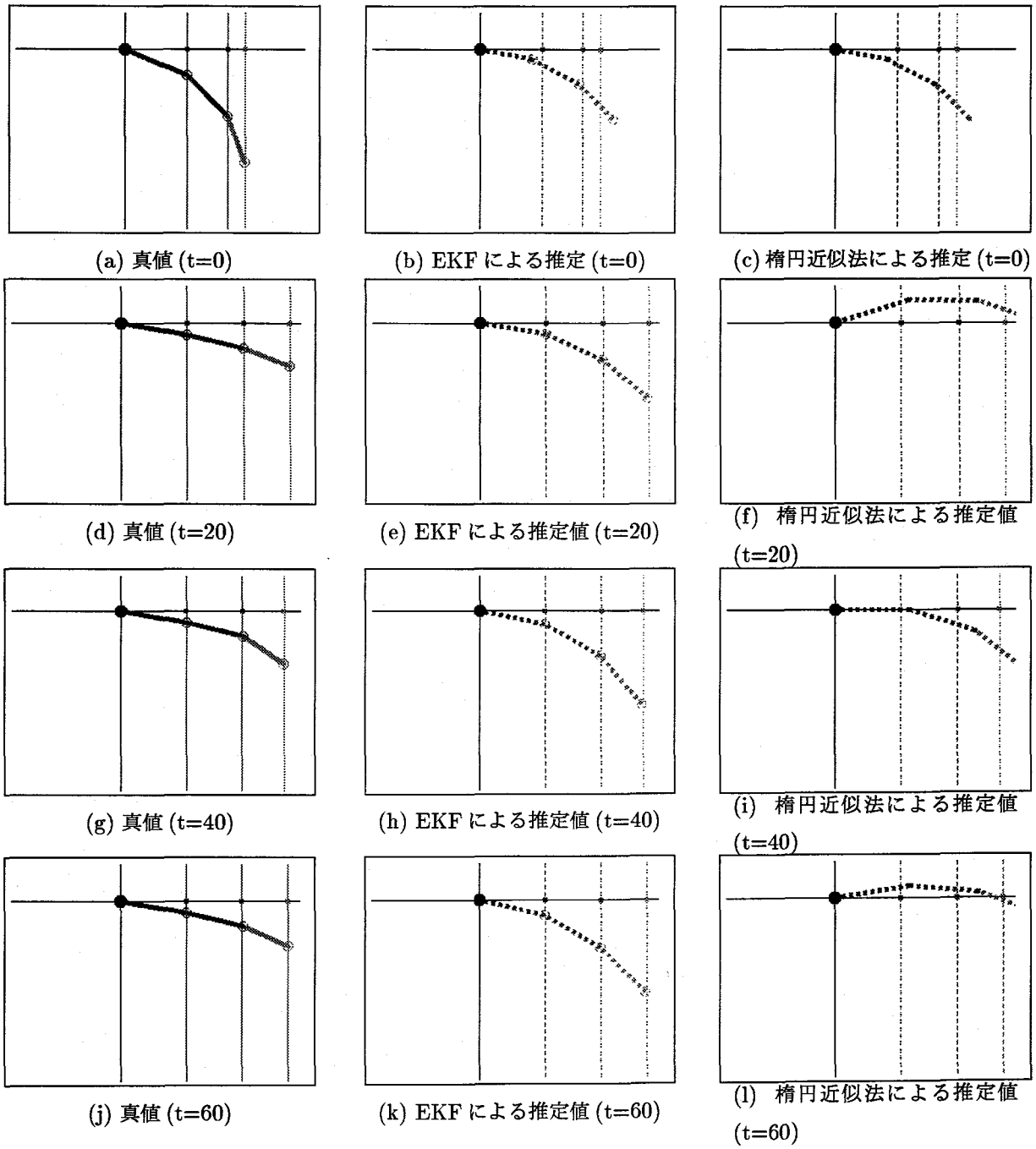
### 5.7.4 楕円近似法による EKF の問題点の改善

楕円近似法による推定法と EKF による推定法の性能を比較検討し，4.8節で述べた EKF による推定の問題点を楕円近似法によって改善できることを示す．

5.7.1節の実験と同じ制約条件と初期推定を用いて，同一の観測値を入力した場合の両者の推定結果を図 5.23,5.24に示す．図 5.25,5.26は 3 番目のリンクの角度とリンク長の推定結果のグラフである．角度は関節角度 $\theta_3$ ではなくモード変換後の角度 $\theta_3^*$ で表示してある．また，EKF による推定値のグラフは推定範囲として $2\sigma$ をつかって表示した．

実験ではまず，90 フレームあたりまでリンクを小さく振動させて，類似の観測を長時間に渡って入力したのち，関節を徐々に回転させていった．その結果，EKF による推定ではリンク長の推定値が真値よりも長い状態で分散が小さくなってしまったため，第 100 フレーム前後で真のリンクが画面に対して平行になった際にも，奥行きの曖昧性が生じたことを検知できず，画面に対して対称な推定を生成しなかった（図 5.24(h)）．その結果，後の推定が制約を満たさなくなり推定を誤ってしまった（図 5.24(k)）．

一方楕円近似法による推定では可能なリンク長の範囲を考慮しながら推定を行なうので，奥行きがわからないことによる曖昧さが発生したことを検知して複数の推定候補を生成し（図 5.24(i)），追跡に成功していることがわかる（図 5.24(l)）．



(a) 真値 (t=0)　(b) EKF による推定 (t=0)　(c) 楕円近似法による推定 (t=0)

(d) 真値 (t=20)　(e) EKF による推定値 (t=20)　(f) 楕円近似法による推定値 (t=20)

(g) 真値 (t=40)　(h) EKF による推定値 (t=40)　(i) 楕円近似法による推定値 (t=40)

(j) 真値 (t=60)　(k) EKF による推定値 (t=60)　(l) 楕円近似法による推定値 (t=60)

**図 5.23**: 楕円近似法と EKF による形状と姿勢の推定結果の比較

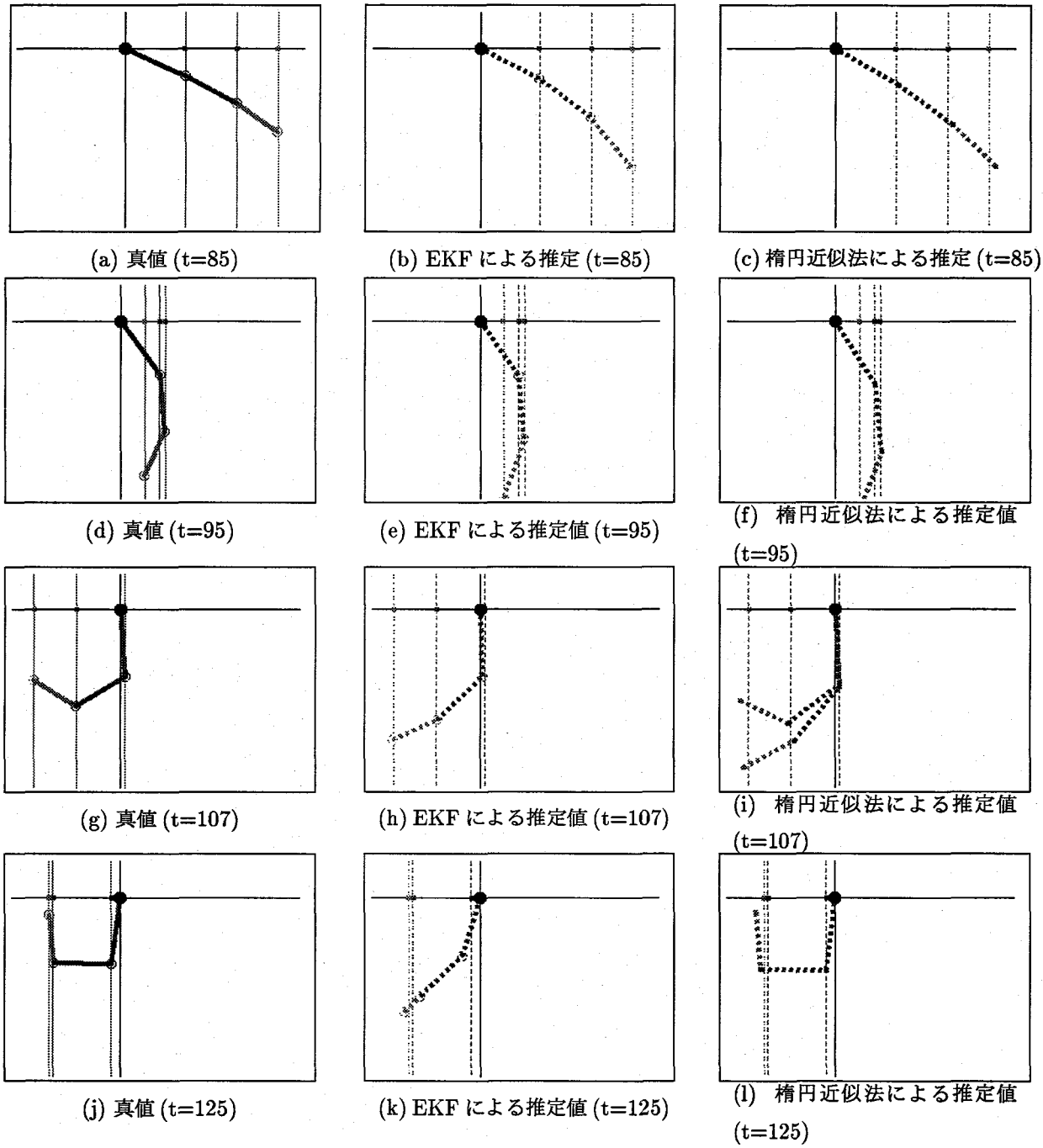
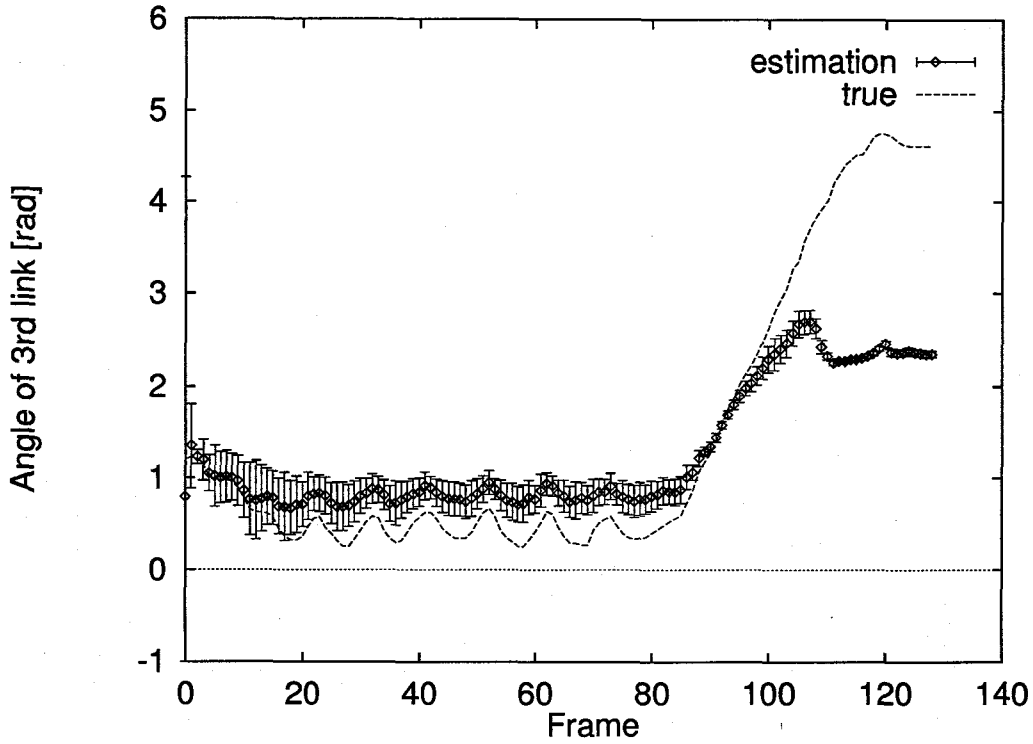
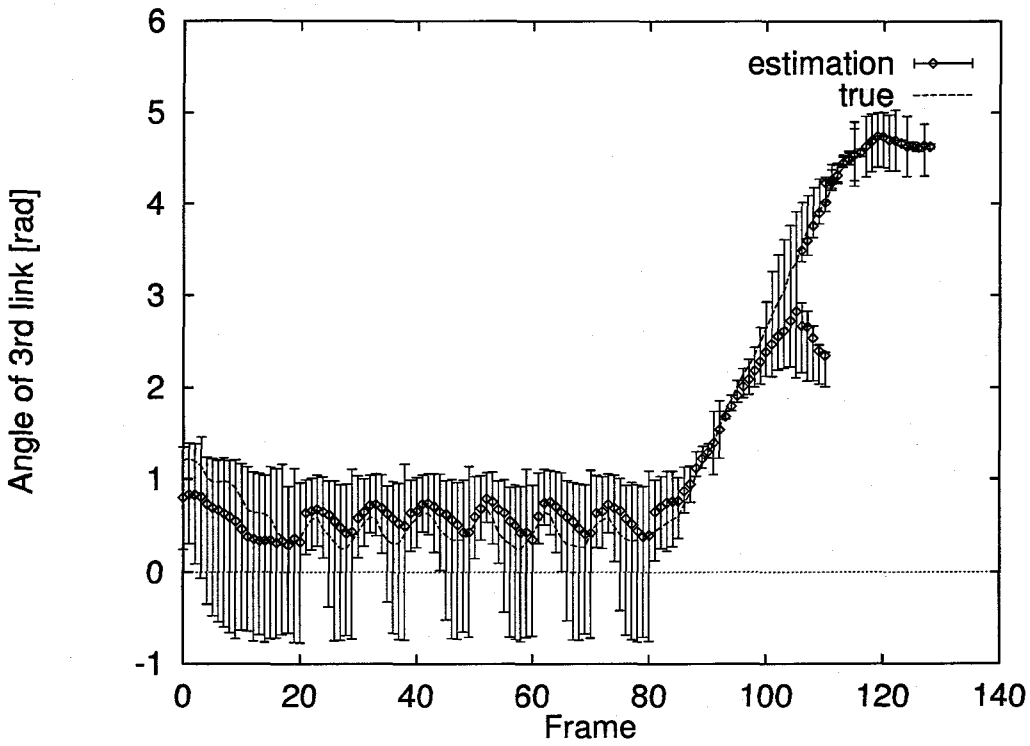



図 **5.24**: 楕円近似法と EKF による形状と姿勢の推定結果の比較（つづき）

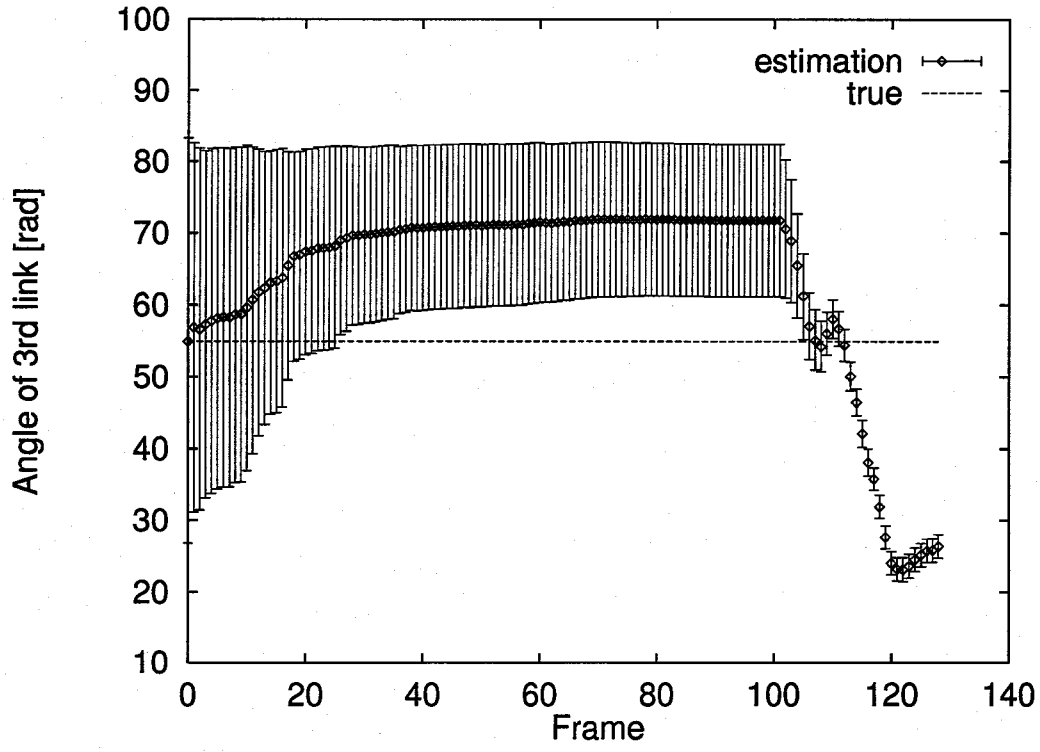



(a) EKF による推定結果

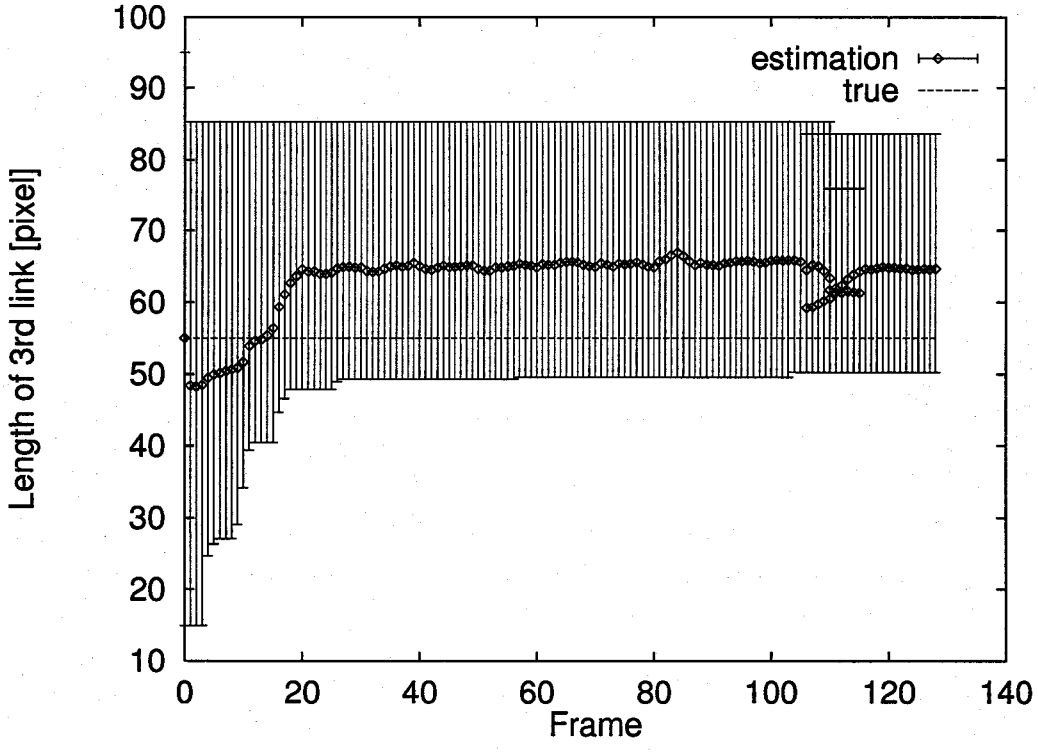



(b) 楕円近似法による推定結果

図 **5.25**: EKF と楕円近似法による $\theta_3$ の推定範囲の比較




(a) EKF による推定結果




(b) 楕円近似法による推定結果

図 **5.26**: EKF と楕円近似法による $r_3$ の推定範囲の比較

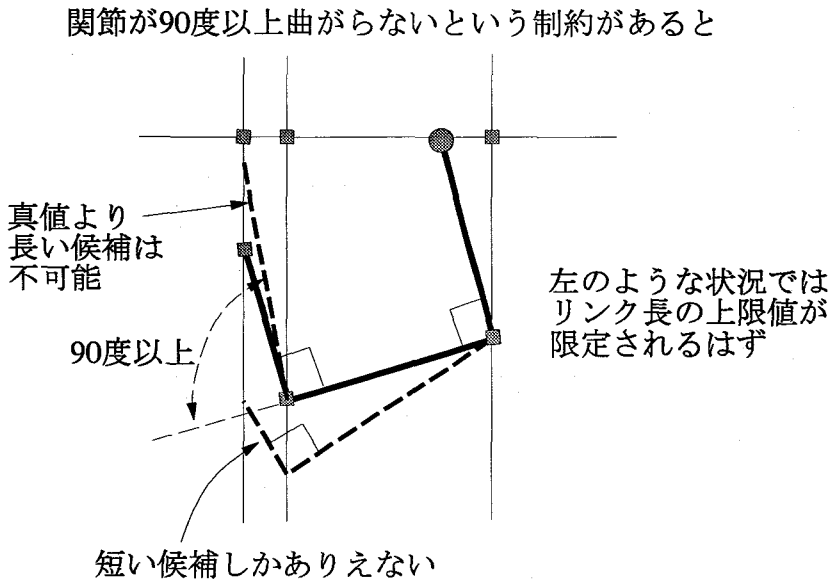



図 **5.27**: リンク長の上限値が限定される姿勢

### 5.7.5 楕円近似法による推定の精度上の問題点

図 5.26(b) のグラフを見ると，楕円近似法での推定ではリンク長の推定範囲の上限値が初期値からほとんど小さくなっていない．図 5.24(j) のような姿勢の観測を入力した場合，どのようなことがわかるかを図説したのが図 5.27である．

関節が 0 度から 90 度の範囲でしか変化しないという制約を考えると，図 5.27のように可動限界近くの姿勢を観測したときには，真値よりもリンク長が長い候補は関節角度が 90 度以上になってしまうので，リンク長の上限値は誤差を考慮しても真値付近まで小さくなるはずである．実験では図 5.24(j) のようにほぼ可動限界の姿勢を観測しているにもかかわらず，図 5.26(b) のグラフをみるとリンク長の上限値は小さくなっていない．

図 5.24(j) の時点でのパラメータの推定楕円を関節角度 $\theta_2$-$\theta_3$ 平面に投影した図が図 5.28である．図中の直線は $\theta_2$,$\theta_3$ に関する制約を示す．実線の楕円は推定パラメータ楕円であるが，各制約によってつくられるグレーの可動領域からはみ出しているのが図からわかる．

5.1節で述べた楕円近似法によるパラメータ領域の更新では，図 5.29(a) に示すように，初期楕円（予測楕円）が制約をみたす領域から 2 方向にはみ出す場合は，図中の A 点と B 点を通る楕円を求めることになるので，更新した楕円の近似精度は比較的高い．しかし，図 5.29(b) のように制約をみたす領域から 1 方向にしかはみ出さない場合は，図中の C 点を通る楕円を考えるため，更新した楕円は初期楕円からほとんど変化せず，制約からはみ出した部分が残ってしまう．関節物体の可動限界近くではこの図 5.29(b) の状態が発生しているため，実験では本来削られるべき 90 度以上の関節角に対応する長いリンク長の可能性が残ってしまったと考えられる．

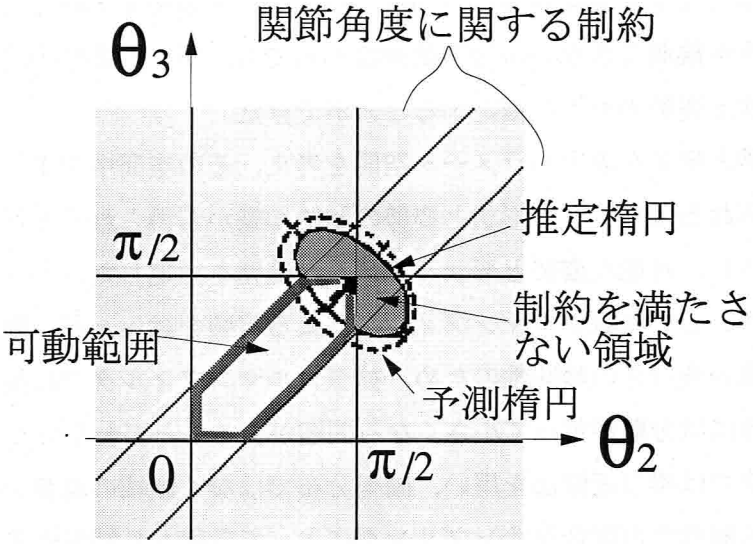
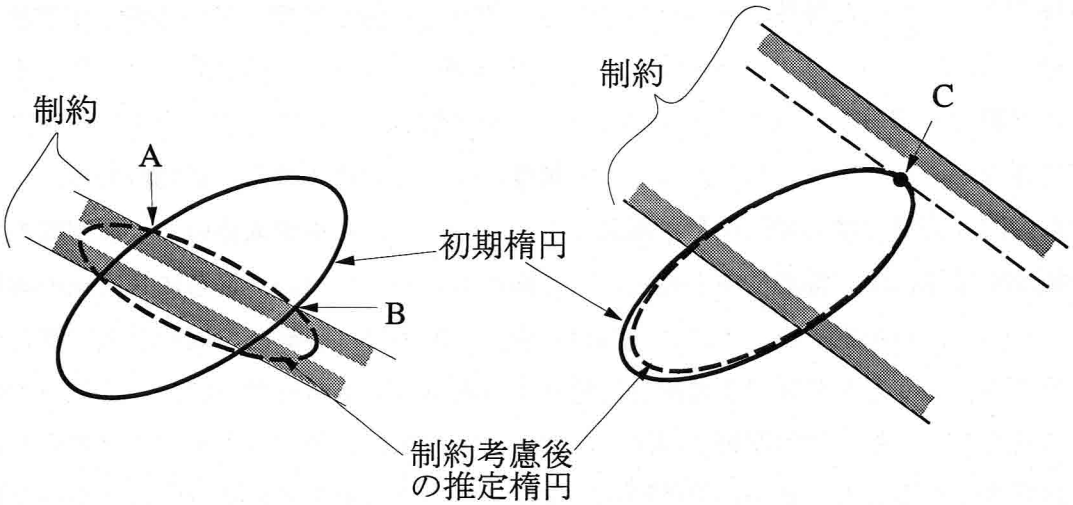



図 5.28: 可動限界の姿勢を観測した時の推定楕円




(a) 楕円が2方向にはみ出す場合

(b) 楕円が1方向にしかはみ出さない場合

図 5.29: 楕円近似法の近似精度

## 5.8 まとめ

３，４章と本章では単眼視動画像から関節物体の形状と姿勢を同時に推定する手法について述べた．奥行きが観測できない単眼視動画像からでも，形状と姿勢の制約知識を利用することにより，形状と姿勢の両方を推定することができた．

本研究では姿勢と形状を表すパラメータ空間を考え，その空間中で１）予測される領域，２）観測から得られる領域，３）形状と姿勢の制約知識から得られる領域，の３つの積集合を求めることにより，可能な姿勢と形状パラメータ範囲を推定した．パラメータ範囲を具体的に表現する手段として拡張カルマンフィルタによる実装を試みたが，観測方程式の非線形性と単眼視動画像の奥行きの曖昧性のため，拡張カルマンフィルタでは長い時間類似の観測が入力された場合には分散が誤って小さくなる問題があることがわかった．

そこで，本研究では楕円近似法を用い，確率分布ではなく領域の境界を表現する手法を提案した．観測の非線形性の度合をサンプリングによって見積もり補償することによって拡張カルマンフィルタにおける問題を解決した．このとき関節物体の高い自由度のために計算量が膨大になるので，関節物体のパラメータと観測をモードに分解することによって計算量を削減する方法を示した．また複数の楕円による領域表現を用いて，一時的な奥行きの曖昧性のために生じる複数の可能性を考慮した．この手法によって，同じ制約条件と初期推定からでも異なる形状の物体を識別することができることを，実験例で確認した．

本研究では３次元の関節物体に適用できる一般的な手法を提案しているが，本手法の性能を確認するために，まず２次元の例についての実験を示した．5.7.4節で示したように，パラメータ範囲を楕円で近似すると余分なパラメータ領域を含んでしまうため，正確に範囲を限定できないことがある．したがってより精度の高い近似法を考える必要がある．

また今回の実験では各関節位置が観測できるとしたが，手指や人体のような実際の関節物体を観測する際には，輪郭や中心軸といった観測しか得られない．さらに３次元の物体にはオクルージョンが発生する．本手法における楕円近似法では各観測の制約を順に統合するので，オクルージョンが発見できた場合は統合する観測の数が減るだけで，本手法がそのまま利用できるが，対応づけの問題を解決しなければならない．また，将来的には表面の形状といった長さや太さよりも細かい物体形状を推定していくことも考えたい．これらが今後の課題である．

# 第 6 章

# おわりに

本論文では単眼視動画像から関節物体の３次元形状と姿勢を推定する手法について論じた.

本論文では単眼視動画像から関節物体の３次元形状と姿勢を推定する問題を

1. 画像とモデルの対応づけがわからないもとでの，姿勢の大まかな推定

2. 対応づけと大まかな姿勢がわかったもとでの，形状と姿勢の同時推定

の２つの段階にわけ，それぞれについて効果的な手法を提案した.

まず前者として，単眼視時系列シルエット画像を用いて手指の三次元姿勢推定する手法について論じた．本手法では，考えられる手指の姿勢候補群を三次元形状モデルを利用して生成し，その投影が入力シルエットにもっとも適合するものを探索した．その際

1. 動きに基づく事前確率順の姿勢候補生成

2. 形状特徴とモデルの対応づけの仮定と検証

3. シルエットの形状特徴に基づく姿勢候補の適合度の評価

4. 候補の適合度の確率による統合

5. ビーム探索を用いたロバストで効率的な候補探索

を用いることによって，手指の動きが速くシルエットの変化が大きい，オクルージョンのためにシルエットの解釈が曖昧になるなど，従来手法が適用できない難しい場合に対しても，画像とモデルの対応づけを解決し，大まかな姿勢の推定ができた.

次に，対応づけと大まかな姿勢がわかったもとで，単眼視動画像から関節物体の形状と姿勢を同時に推定する手法について論じた．単眼視動画像には奥行きの曖昧さがあり，一般に

は多くの対応点がわからないと形状と姿勢は復元できない．本研究では関節物体の関節角度や各部の長さ，太さに関する制約知識を利用すれば，単眼視動画像の奥行きの曖昧性が解決でき，少ない対応点から関節物体の形状と姿勢が同時に推定できることを示した．このとき

1. 形状と姿勢の範囲，観測，制約知識のパラメータ空間における領域表現
2. 楕円近似法による観測と制約の逐次的な統合
3. サンプリングによる観測の非線形性の補償
4. 複数の楕円による奥行きの曖昧性への対処

を用いて，拡張カルマンフィルタを用いた場合に生じる問題を解決し，また同じ制約条件と初期推定からでも，異なる形状の物体を識別することができることを，実験により確認した．

本論文では両段階の手法をそれぞれ提案したが，実際に画像から関節物体の形状と姿勢を推定するためには，前者の手法により大まかな推定を得たのち，後者の手法によってより確からしい推定を行なうといった，両者を結合するシステムが必要である．とくに前者の手法は動きが速くオクルージョンが生じるなどの難しい画像から推定することが可能だが，処理が重いといった欠点がある．そこで実時間で動作するシステムには，普段は画像とモデルと対応づけをもっと簡単な処理によって行い，推定誤りが生じたときにのみ前者の手法を用いる，といったアルゴリズムを選択する機構が必要になる．これらの問題が今後の課題としてあげられる．

# 参考文献


[1] 徳永修一, 兼田雅弘, 大崎紘一. “手話の数詞の連続認識法に関する研究”. 日本機械学会論文集 (C 編) 59 巻 561 号 No.92-1205, pp. 159–165, 1993.

[2] 長谷川修, 藤木真和, 石塚満. “指サイン認識による自然感を有する仮想金魚像との実時間インタラクション”. 画像電子学会誌, 第 23 巻第 2 号, pp. 86–95, 1994.

[3] T. G. Zimmerman and J. Lanier. “A Hand Gesture Interface Device”. *in CHI+GI'87 Proceedings*, pp. 189–192, 1987.

[4] J. Davis and M. Shah. “Recognizing Hand Gestures”. *ECCV'94*., pp. 331–340, 1994.

[5] R. Chipolla, Y. Okamoto, and Y. Kuno. “Robust Structure from Motion using Motion Parallax”. In *ICCV'93*, pp. 374–382. IEEE, 1993.

[6] 渡辺賢, 岩井儀雄, 八木康史, 谷内田正彦. “カラーグローブを用いた指文字の決定木による認識”. In *MIRU'96*, pp. II–61–66, 1996.

[7] 吉野和芳, 川嶋稔夫, 青木由直. “手形状遷移に基づく手話単語推定”. In *MIRU'96*, pp. II–43–48, 1996.

[8] 岩井儀雄, 八木康史, 谷内田正彦. “単眼動画像による手の 3 次元運動復元”. 第 3 回ロボットシンポジウム, pp. 121–126, 1993.

[9] J. M. Rehg and T. Kanade. “Visual Tracking of High DOF Articulated Structures: an Application to Human Hand Tracking”. *ECCV'94*, pp. 35–46, 1994.

[10] D. Lowe. “Fitting Parameterized Three-Dimentional Models to Images”. *IEEE trans., Pattern Anal. Machine Intell.,vol.13,No.5*, pp. 441–450, 1991.

[11] 柴田勉, 杉原厚吉, 杉江昇. “正射影速度場からの関節物体の立体情報復元法”. 電子情報通信学会論文誌 vol.J68-D No.9, pp. 1689–1696, 1985.

[12] R. J. Holt, A. N. Netravali, T. Huang, and Richard J. Qian. "Determining Articulated Motion from Perspective Views: A Decomposition Approach". In *Proc. of Workshop on Motion of Non-Rigid and Ariticulated Objects '94*, pp. 126–137. IEEE, 1994.

[13] 山本正信, 川田聡, 近藤拓也, 越川和忠. "ロボットモデルに基づく人間動作の３次元動画像追跡". 電子情報通信学会論文誌 vol.J77-D-II No.8, pp. 1562–1570, 1994.

[14] 堀口和俊, 斉藤英雄, 持丸正明, 小沢慎治. "誤差を含んだ移動ベクトルからの手指運動パラメータ推定". In *PRU95-94*, pp. 92–96. 電子情報通信学会, 1994.

[15] 亀田能成, 美濃導彦, 池田克夫. "シルエットを利用した手指の三次元形状推定法". In *MIRU'92*, pp. II239–II246, 1992.

[16] Y. Kameda, M. Minoh, and K. Ikeda. "Three Dimensional Pose Estimation of an Articulated Object from its Silhouette Image". In *ACCV'93*, pp. 612–615, 1993.

[17] J. J. Kuch and T. S. Huang. "Virtual Gun: A Vision Based Human Computer Interface Using the Human Hand". In *MVA'94*, pp. 196–199, 1994.

[18] M. Mochimaru and N. Yamazaki. "The Three-dimensional Measurement of Unconstrained Motion Using a Model-matching Method". *ERGONOMICS, vol.37, No.3*, pp. 493–510, 1994.

[19] 大谷淳, 岸野文郎. "遺伝的アルゴリズムを用いたマルチ画像からの人物の姿勢検出の検討 - 上半身の関節角度の検出 -". 信学技報 PRU93-122, pp. 47–54. 電子情報通信学会, 1993.

[20] 石淵耕一, 岩崎圭介, 竹村治雄, 岸野文郎. "画像処理を用いた実時間手振り推定とヒューマンインターフェー ス". 電子情報通信学会論文誌, D-II vol.J79-DII No.7, pp. 1218–1229, 1996.

[21] B. T. Lowerre and R. D. Reddy. "The Harpy Speech Understanding System". In W. A. Lea, editor, *Trends in Speech Recognition*, pp. 340–360. PresticeHall, Englewood Cliffs, NJ, 1980.

[22] 李亜兵, 辺見一男, 森隆知, 井上和夫. "２指の動作モデルの構築". 電子情報通信学会論文誌 vol.J77-D-II No.5, pp. 1031–1034, 1994.

[23] B. Dorner. "Hand Shape Identification and Tracking for Sign Language Interpretation". In *IJCAI'93 Looking at people Workshop*, 1993.

[24] Y. Hel-Or and M. Werman. "Recognition and Localization of Articulated Objects". In *Proc. of Workshop on Motion of Non-Rigid and Ariticulated Objects '94*, pp. 116–123. IEEE, 1994.

[25] C. Tomasi and T. Kanade. "Shape and motion without depth". In *ICCV '90*, pp. 91–95. IEEE, 1990.

[26] J. Costeira and T. Kanade. "A Multi-body Factorization Method for Motion Analysis". In *Carnegy mellon Univ. Tech. Report*, pp. CMU–CS–TR–94–220, 1994.

[27] 浅田稔. ダイナミックシーンの理解. 電子情報通信学会, 1994.

[28] O. Faugeras. *Three-Dimensional Computer Vision.* The MIT Press, 1993.

[29] 尺長健. "単眼視回転推定に関する木構造基本問題の統一的解法". 電子情報通信学会論文誌, D-II vol.J76-DII No.5, pp. 976–986, 1993.

[30] J. O'Rourke and N. I. Badler. "Model-Based Image Analysis of Human Motion Using Constraint Propagation". *IEEE Trans. of Pattern Anal. and Machine Intell.,PAMI-2, No.6*, pp. 522–536, 1980.

[31] 有本卓. カルマン・フィルター. 産業図書, 1977.

[32] F. C. Schweppe. "Recursive State Estimation: Unknown but Bounded Errors and System Inputs". *IEEE Trans. on Automatic Control, vol.AC-13, No.1*, pp. 22–28, 1968.

[33] R. E. Ellis. "Geometric Uncertainties in Polyhedral Object Recognition". *IEEE Trans. on Robotics and Automation, vol.7, No.3*, pp. 361–371, 1991.

[34] D. P. Bertsekas and I. B. Rhodes. "Recursive State Estimation for a Set-Membership Description of Uncertainty". *IEEE Trans. on Automatic Control, vol.AC-16, No.2*, pp. 117–128, 1971.

[35] M. Isard and A. Blake. "Contour Tracking by Stochastic Propagation of Conditional Density". *ECCV'96.*, pp. 343–356, 1996.

[36] S. Geman and D. Geman. "Stochastic Relaxation, Gibbs Distributions, and the Baysian Restoration of Images". *IEEE Trans. of Pattern Anal. and Machine Intell., PAMI-6, No.6*, pp. 721–741, 1984.

[37] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Acadimic Press, New York, San Francisco, London, 1970.

[38] E. Fogel and Y. F. Huang. "On the Value of Information in System Identification – Bounded Noise Case". *Automatica, vol.18, No.2*, pp. 229–238, 1982.

[39] 佐野昭. “不確実さを含むシステムの同定 III”. 計測と制御 32 巻第 9 号, pp. 735–740, 1993.

[40] J.Craig 著, 三浦宏文, 下山勲訳. ロボティクス - 機構・力学・制御 -. 共立出版, 1991.

# 付録 A

# 投影行列と剛体回転行列

この章では関節物体の観測方程式を構成する，画像への投影写像行列と剛体回転行列について述べる．

## A.1 投影写像行列

手指などを画像で撮影するときはズームしていることが多く，カメラの焦点距離が長くなっている．このような場合は投影は正射影と考えてよい．よって本研究では投影法に正射影を仮定する．したがって，投影写像行列は

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{A.1}$$

となる．このとき，奥行き方向の並行移動は画像から検知できないので，並行移動のパラメータ $\boldsymbol{t}$ は画面に並行な２成分のみとする．

## A.2 ３次元の剛体回転行列

物体の空間中の向きを表現するためには３自由度の回転を表現する必要がある．それにはハミルトンの四元数 (quaternion)[28] [40] を用いると便利である．四元数 $\boldsymbol{q} = (s, l, m, n)$ は方向が回転軸を表し，大きさが回転角を表す回転ベクトル $\boldsymbol{\omega}$ をもちいて

$$s = \cos(|\boldsymbol{\omega}|/2) \tag{A.2}$$

$$\boldsymbol{v} \equiv (l, m, n)^T = \boldsymbol{\omega} \sin(|\boldsymbol{\omega}|/2) \tag{A.3}$$

で表現される (実部が $s$，虚部が $\boldsymbol{v}$ の複素数と考えて良い) ．すなわち $||\boldsymbol{q}|| = 1$ である．四元数は３自由度の剛体回転写像を４次元空間中の単位球面に写像する．剛体回転行列を $\boldsymbol{q}$

を用いて表現すると，

$$\boldsymbol{R}(\boldsymbol{q}) = \begin{bmatrix} s^2+l^2-m^2-n^2 & 2(lm-sn) & 2(ln+sm) \\ 2(lm+sn) & s^2-l^2+m^2-n^2 & 2(mn-sl) \\ 2(ln-sm) & 2(mn+sl) & s^2-l^2-m^2+n^2 \end{bmatrix} \tag{A.4}$$

になることが知られている．しかし $||\boldsymbol{q}|| = 1$ の制約は二次式なので四元数をそのままカルマンフィルタの状態量に用いることはできない．

他の手段として，ワールド座標の３つの固定軸回りに順番に回転する行列

$$\boldsymbol{R}_x(\omega_x)\boldsymbol{R}_y(\omega_y)\boldsymbol{R}_z(\omega_z) \tag{A.5}$$

を用いる方法があるが，この表現の場合，$\boldsymbol{\omega}$ が大きい姿勢では非線形性が強くなり，拡張カルマンフィルタにおける誤差が大きくなる．

そこでワールド座標系から時刻 $t-1$ における物体座標系への変換を $\boldsymbol{q}_{t-1}$ で表現しておいて，時刻 t における物体の向きを $\boldsymbol{q}_{t-1}$ と時刻 $t-1$ から $t$ の間の物体座標系における $\boldsymbol{\omega}$ 軸回りのローカルな剛体回転 $\boldsymbol{R}_t$ で表現することを考える．こうすると $t-1$,$t$ 間の運動が小さければ $\boldsymbol{R}_t$ は

$$\boldsymbol{R}_t(\boldsymbol{\omega}) = \begin{bmatrix} 1 & -\omega_z & \omega_y \\ \omega_z & 1 & -\omega_x \\ -\omega_y & \omega_x & 1 \end{bmatrix} \tag{A.6}$$

でよく近似できる ($|\boldsymbol{\omega}|$ は回転角度を意味することに注意)．これより物体座標系で $\boldsymbol{X}$ の点はワールド座標系で

$$\boldsymbol{X}_w(t) = \boldsymbol{R}(\boldsymbol{q}_{t-1})\boldsymbol{R}_t(\boldsymbol{\omega})\boldsymbol{X} \tag{A.7}$$

に写る．上式は線形なので，カルマンフィルタに用いて推定を行う．推定の結果，$\hat{\boldsymbol{\omega}}$ が得られたら，式 A.2,A.3から対応する四元数 $\boldsymbol{q}_\omega$ を求め，

$$\boldsymbol{q}_t = \boldsymbol{q}_{t-1} \times \boldsymbol{q}_\omega \tag{A.8}$$

と更新する[1]．ただし'×' 記号は四元数の積であり次のように定義される．

$$\boldsymbol{q}_1 \times \boldsymbol{q}_2 \equiv (s_1 s_2 - \boldsymbol{v}_1 \cdot \boldsymbol{v}_2, s_1\boldsymbol{v}_2 + s_2\boldsymbol{v}_1 + \boldsymbol{v}_1 \wedge \boldsymbol{v}_2) \tag{A.9}$$

ただし，'·' は内積，'∧' は３次元の外積である．

[1]回転の合成は四元数の積で表現できる．証明は略す．

次の時刻の推定を行うときは，ワールド座標に対する物体座標系の向きが変化しているので，$\boldsymbol{\omega}$の平均は0，$\boldsymbol{\omega}$の分散は

$$\boldsymbol{R}(\boldsymbol{q}_\omega)V[\boldsymbol{\omega}]\boldsymbol{R}(\boldsymbol{q}_\omega)^T \tag{A.10}$$

と座標変換しておく．

以上のようにして，3次元剛体回転を表現する．

# 付録 B

# EKF における分布切り取りによる制約の統合

平均 $\boldsymbol{q}_{k-1}$，分散 $\boldsymbol{Q}_{k-1}$ の正規分布のうち，$\varphi_k^T \boldsymbol{x} \leq b_k$ の制約を満たす領域の平均 $\boldsymbol{q}_k$ と分散 $\boldsymbol{Q}_k$ を計算する問題は

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{T}^T(\boldsymbol{x} - \boldsymbol{q}_{k-1}) \tag{B.1}$$

なる線形変換によって，座標軸に平行な平面と平均 0，分散 $\boldsymbol{I}$，制約不等式 $x_1' \leq c_k$ の場合に帰着できる．ただし $\boldsymbol{R}, \boldsymbol{T}$ は直交行列, $\boldsymbol{W}$ は対角行列であり，

$$\boldsymbol{T}\boldsymbol{W}\boldsymbol{T}^T = \boldsymbol{Q}_{k-1} \tag{B.2}$$

$$\boldsymbol{R}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{T}^T\varphi_k = (1, 0, \cdots, 0)^T \tag{B.3}$$

$$c_k = \frac{b_k - \varphi_k^T \boldsymbol{q}_{k-1}}{\left(\varphi_k^T \boldsymbol{Q}_{k-1}\varphi_k\right)^{\frac{1}{2}}}. \tag{B.4}$$

を満たす．

変換された分布は等方分布なので各成分は独立であり，確率密度分布は周辺分布の積で表現される．

$$p(\boldsymbol{x}') = \prod_i p_i(x_i') = \prod_i \frac{1}{\sqrt{2\pi}} \exp(-\frac{{x_i'}^2}{2}) \tag{B.5}$$

制約不等式は $x_1' \leq c_k$ であるから，$x_1' \leq c_k$ の範囲で積分して 1 になるように確率密度分布を正規化する．

$$\begin{aligned} \int_{-\infty}^{c_k} \cdots \int_{-\infty}^{\infty} p(\boldsymbol{x}') dx_1' \cdots dx_n' &= \int_{-\infty}^{c_k} \frac{1}{\sqrt{2\pi}} \exp(-{x_1'}^2) dx_1' \\ &= \frac{1}{2}\left\{1 + \mathrm{erf}\left(\frac{c_k}{\sqrt{2}}\right)\right\} \end{aligned} \tag{B.6}$$

ただし $\mathrm{erf}(\cdot)$ は誤差関数であり，

$$\mathrm{erf}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}) dt \tag{B.7}$$

である．この関数は解析関数ではないが，数値演算ライブラリによって短時間に値を求めることができる．よって制約を満たす部分の確率密度分布は

$$p'(\boldsymbol{x}') = \alpha \prod_i p_i(x'_i) \tag{B.8}$$

$$\alpha \equiv \frac{\sqrt{2}\exp(-{x'_i}^2/2)}{\sqrt{\pi}(1+\mathrm{erf}(c_k/2)} \tag{B.9}$$

となる．この確率密度分布を用いて制約を満たす部分の平均と分散を計算する．各 $x'_i$ は独立であるであることに注意すれば

$$\begin{aligned}
\mu_1 = E[x'_1] &= \alpha \int_{-\infty}^{c_k} x'_1 \exp(-{x'_1}^2/2) dx'_1 \\
&= -\alpha \exp(-\frac{c_k^2}{2}) && \text{(B.10)} \\
\mu_i = E[x'_i] &= 0 \ (i \neq 1) && \text{(B.11)} \\
\sigma_{11}^2 = E[x'_1 - \mu_1][x'_1 - \mu_1] &= \alpha \int_{-\infty}^{c_k} (x'_1 - \mu_1)^2 \exp(-\frac{{x'_1}^2}{2}) dx'_1 \\
&= 1 + \alpha c_k \exp(-\frac{c_k^2}{2}) - \mu_1^2 && \text{(B.12)} \\
\sigma_{ii}^2 = E[x'_i - \mu_i][x'_i - \mu_i] &= 1 \ (i \neq 1) && \text{(B.13)} \\
\sigma_{ij}^2 = E[x'_i - \mu_i][x'_j - \mu_j] &= 0 \ (i \neq j) && \text{(B.14)}
\end{aligned}$$

となるから，制約を満たす領域の平均と分散は

$$\boldsymbol{\mu} = (\mu_1, 0, \cdots, 0) \tag{B.15}$$

$$\boldsymbol{S} = \begin{bmatrix} \sigma_{11}^2 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{B.16}$$

となる．

# 付録 C

# 楕円近似法におけるトレース最小解と体積最小解

5.1節で述べた楕円近似法では,
式5.3で表現される楕円体は $\lambda \leq 0$ のとり方で幾通りも考えれるが，そのうち

1. 体積 ($|\boldsymbol{P}|$) が最も小さいもの

2. 主軸長さの二乗和 ($\mathrm{tr}\boldsymbol{P}$) が最も小さいもの

を求めることは容易である [38].

実際に両者の計算を行なった結果を表示したのが図 C.1である．この２つを比較すると，体積最小の解は楕円領域が平たくなる傾向があり，トレース最小の解は円に近くなる傾向がある．本研究においては観測誤差が予測誤差に対して比較的正確な場合を扱っているので，初期楕円に比べて並行な２平面間の距離が小さいことが多い．このような場合では図 C.1のように，体積 (det) 最小解のほうが楕円の長軸方向の誤差が大きくなりやすい．つまり推定範囲が不正確な方向に楕円が大きくなるので，最も不正確なパラメータがいつまでも推定範囲が限定されない不都合が生じることが多い．そこで本研究ではトレース最小解を用いることとした.

以下に式5.3で表現される楕円体の体積を最小にする $\lambda$ の計算方法を示す． $\boldsymbol{x}$ の次元を $n$ として

$$G \equiv \boldsymbol{m}_i^T \boldsymbol{P}_{t-1} \boldsymbol{m}_i \tag{C.1}$$

$$\varepsilon \equiv \rho_i - \boldsymbol{m}_i^T \widehat{\boldsymbol{x}}_{t-1} \tag{C.2}$$

$$\alpha_1 \equiv (n-1)c_i^2 G^2 \tag{C.3}$$

$$\alpha_2 \equiv G\{(2n-1)c_i^2 - G + \varepsilon^2\} \tag{C.4}$$

$$\alpha_3 \equiv n(c_i^2 - \varepsilon^2) - G \tag{C.5}$$

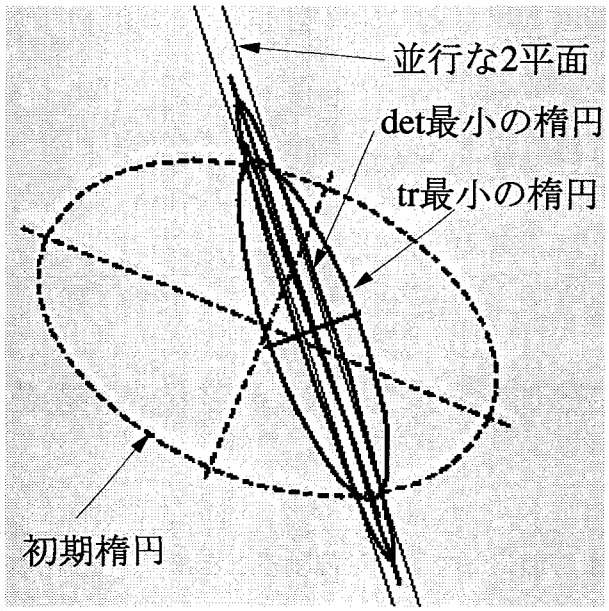



図 **C.1**: トレース最小解と体積最小解

としたときの二次方程式

$$\alpha_1\lambda^2 + \alpha_2\lambda + \alpha_3 = 0 \tag{C.6}$$

を解く．ただし $\lambda < 0$ のときは $\lambda = 0$ とする．このとき $\hat{\boldsymbol{x}}_t$ と $\boldsymbol{P}_t$ は，トレース最小解と同じく式5.12，5.13によって更新される．

# 付録 D

# 楕円近似法における適切な線形化位置の計算

観測制約による曲面を線形化して平面近似するときに，予測楕円の中心で線形化したのでは正確な傾きが求められない．そこでIterated Extended Kalman Filter[37] を応用した繰り返し計算によって適切な線形化位置を求める．

予測楕円の近傍のうち観測制約をみたす点で観測を線形化したいので，予測楕円の中心からの距離が最小になる観測制約上の点を求める．しかし，図D.1に示すようにユークリッド距離最小の点は予測楕円の外に出てしまうことがある．そこで，できるだけ予測楕円と観測制約の交差領域の中央で線形化するために，予測楕円によるマハラノビス距離を最小にする点を求める．これはちょうど観測の分散を0にしたときのIterated Extended Kalman Filterに相当する．

そこで次の計算を行なう．予測楕円と観測を満たす領域を

$$(\boldsymbol{x}_t - \tilde{\boldsymbol{x}}_t)^T \boldsymbol{M}_t^{-1} (\boldsymbol{x}_t - \tilde{\boldsymbol{x}}_t) \leq 1 \tag{D.1}$$

$$|\boldsymbol{y}_t - \boldsymbol{h}(\boldsymbol{x})| \leq \Delta \boldsymbol{y} \tag{D.2}$$

とおく．このときカルマンゲイン $\boldsymbol{K}_t$ は

$$\boldsymbol{K}_t \equiv \boldsymbol{M}_t \boldsymbol{C}_t^T (\boldsymbol{C}_t \boldsymbol{M}_t \boldsymbol{C}_t^T)^{-1}. \tag{D.3}$$

と計算できる．ただし

$$\boldsymbol{C}_t \equiv \left. \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t} \right|_{\tilde{\boldsymbol{x}}_t} \tag{D.4}$$

である．次の繰り返し計算

$$\boldsymbol{\eta}_0 = \tilde{\boldsymbol{x}}_t \tag{D.5}$$

$$\boldsymbol{\eta}_{i+1} = \tilde{\boldsymbol{x}}_t + \boldsymbol{K}_t (\boldsymbol{y}_t - \boldsymbol{h}(\boldsymbol{\eta}_i) - \boldsymbol{C}_t (\hat{\boldsymbol{x}}_t - \boldsymbol{\eta}_i) \tag{D.6}$$

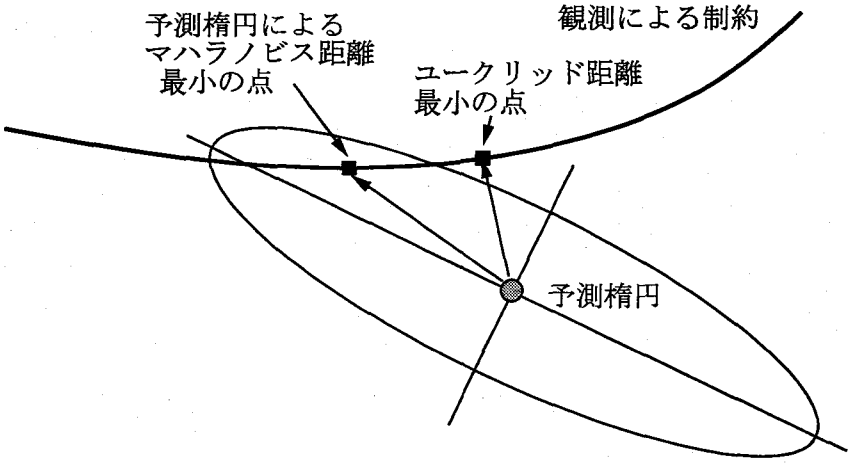



図 **D.1**: 適切な線形化位置の計算

の結果 $\boldsymbol{\eta}_i$ がしきい値より大きな変化をしなくなるか，もしくは繰り返しが一定の回数に達したら，その $\boldsymbol{\eta}_i$ で $\boldsymbol{h}$ を線形化する．

# 付録 E

# 楕円近似法における観測の相関性を考慮したモード分解

5.5節でのモード分解を適用すると，観測 $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x})$ が無相関でもモード分解された観測 $\boldsymbol{y}^*$ は各成分が相関をもってしまう．観測のモード分解行列 $\boldsymbol{E}$ は式5.39から

$$\begin{bmatrix} I & 0 & 0 \\ -I & I & 0 \\ 0 & -I & I \end{bmatrix} \tag{E.1}$$

という形をしており ($I$ は $2 \times 2$ 単位行列)， $\boldsymbol{y}$ の誤差範囲が単位球だったとしたとき $\boldsymbol{y}^*$ は

$$\boldsymbol{y}^{*T}(\boldsymbol{E}\boldsymbol{E}^T)^{-1}\boldsymbol{y}^* \leq 1 \tag{E.2}$$

という範囲に写る．ノルム行列 $\boldsymbol{E}\boldsymbol{E}^T$ の主軸は $\boldsymbol{y}^*$ のどの軸に対しても傾いており，モード変換後の空間で楕円近似法による統合を行うと図5.8に示したような精度の悪化は避けれられない．

しかし以下のようにすると，計算量は増加するが精度を改善することが可能である．

$\boldsymbol{y}^*$ のうち2つの成分 $y_i^*$ と $y_j^*$ は，式E.1の $2 \times 2$ 主小行列 $\boldsymbol{Z}_{ij}$ を用いて

$$(y_i^*, y_j^*)^T(\boldsymbol{Z}_{ij}\boldsymbol{Z}_{ij}^T)^{-1}(y_i^*, y_j^*) \leq 1 \tag{E.3}$$

を満たす．なぜなら上式は式E.2の $(y_i^*, y_j^*)$ 平面への投影だからである．式E.3の楕円領域は主軸が座標軸に対して傾いているので， $\boldsymbol{Z}_{ij}\boldsymbol{Z}_{ij}^T$ を

$$\boldsymbol{Z}_{ij}\boldsymbol{Z}_{ij}^T = \boldsymbol{T}\boldsymbol{W}\boldsymbol{T}^T \tag{E.4}$$

と直交化する行列 $\boldsymbol{T}$ を用いて

$$(y_i^+, y_j^+)^T = \boldsymbol{T}^T(y_i^*, y_j^*)^T \tag{E.5}$$

$$(h_i^+, h_j^+)^T = \boldsymbol{T}^T(h_i^*, h_j^*)^T \tag{E.6}$$

と変換する．すると変換された観測 $y_i^+$ は以下のように書ける．

$$y_i^+ = \alpha h_i^*(\boldsymbol{\omega}, \theta_i^*, r_i) + \beta h_j^*(\boldsymbol{\omega}, \theta_j^*, r_j) \tag{E.7}$$

上式は $(\boldsymbol{\omega}, \theta_i^*, \theta_j^*, r_i, r_j)$ の関数であるので，7 次元である．式 E.7は $r_i$ について陽にとけるから，残りの 6 次元について非線形性の補償のためにサンプリングを実行する．各変数に 5 つずつサンプリングすれば $5^6 = 15625$ 個のサンプリングが必要になる．これをすべての $(i, j)$ の組合せについて実行するが，

$$\boldsymbol{E}\boldsymbol{E}^T = \begin{bmatrix} 1 & 0 & -1 & \cdots & 0 \\ -0 & 2 & \ddots & \ddots & \vdots \\ -1 & \ddots & \ddots & \ddots & -1 \\ \vdots & \ddots & \ddots & 2 & 0 \\ 0 & \cdots & -1 & 0 & 2 \end{bmatrix} \tag{E.8}$$

を見ると，$(i, i \pm 2)$ の組合せ以外は相関 0 なので，実際にはこの組合せについてのみ実行すればよい．そのような組合せは

$$\boldsymbol{Z}_{ij} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \tag{E.9}$$

となる組合せである（他は $\boldsymbol{Z}_{ij} = I$ となり無相関である）．

## 謝辞

本研究は著者が大阪大学大学院工学研究科に在学中に大阪大学工学部電子制御機械工学科の白井良明教授の御指導のもと行なった研究をまとめたものです．白井教授には著者が研究のテーマを探していた当初から御助言下さり，本研究を進めていく際にも終始懇切丁寧なる御指導を頂き，また激励頂きましたことに深い敬意と心よりの感謝の念を表します．

本論文をまとめるにあたり大阪大学工学部電子制御機械工学科の浅田稔教授，大阪大学産業科学研究所の北橋忠宏教授には懇切丁寧なる御教示と御助言を頂きましたことを深く感謝致します．

日々の研究においては著者の所属する白井研究室の久野義徳助教授，三浦純助手には適切な御指導と貴重な御助言を頂きました．ここに心より感謝致します．

また白井研究室の大谷真理子事務官は本論文提出に関わる事務処理をして下さいました．白井研究室の中山收文氏には本研究の内容に関し活発に討論して頂き，多くの示唆を頂きました．この他白井研究室の学生のみなさんには本研究に関わるさまざまな議論を通じて多くの知見を与えて頂きました．ここに深く感謝致します．

最後に常に心の支えとなり励ましをくれた妻知美に心からの感謝をこめて謝辞と致します．