

視触覚を用いた能動的センサフュージョンシステムの研究

向 井 利 春

目次

序	1
第1章 視覚を用いた能動的センサフュージョンシステムの研究	1
1.1 研究の背景	2
1.2 研究の目的	3
1.3 研究の概要	4
1.4 論文の構成	5
第2章 視覚を用いた能動的センサフュージョンシステムの構築	6
2.1 はじめに	7
2.2 システムの構成	10
2.3 視覚を用いた能動的センサフュージョンシステムの構築	11
2.4 システムの評価	12
2.5 システムの応用	13
2.6 システムの将来展望	14
2.7 システムのまとめ	15
第3章 視覚を用いた能動的センサフュージョンシステムの応用	16
3.1 はじめに	17
3.2 システムの応用	18
3.3 システムの評価	19
3.4 システムの応用	20
3.5 システムの応用	21
3.6 システムの応用	22
3.7 システムの応用	23
3.8 システムの応用	24
3.9 システムの応用	25
3.10 システムの応用	26
3.11 システムの応用	27
3.12 システムの応用	28
3.13 システムの応用	29
3.14 システムの応用	30
3.15 システムの応用	31
3.16 システムの応用	32
3.17 システムの応用	33
3.18 システムの応用	34
3.19 システムの応用	35
3.20 システムの応用	36
3.21 システムの応用	37
3.22 システムの応用	38
3.23 システムの応用	39
3.24 システムの応用	40
3.25 システムの応用	41
3.26 システムの応用	42
3.27 システムの応用	43
3.28 システムの応用	44
3.29 システムの応用	45
3.30 システムの応用	46
3.31 システムの応用	47
3.32 システムの応用	48
3.33 システムの応用	49
3.34 システムの応用	50
3.35 システムの応用	51
3.36 システムの応用	52
3.37 システムの応用	53
3.38 システムの応用	54
3.39 システムの応用	55
3.40 システムの応用	56
3.41 システムの応用	57
3.42 システムの応用	58
3.43 システムの応用	59
3.44 システムの応用	60
3.45 システムの応用	61
3.46 システムの応用	62
3.47 システムの応用	63
3.48 システムの応用	64
3.49 システムの応用	65
3.50 システムの応用	66
3.51 システムの応用	67
3.52 システムの応用	68
3.53 システムの応用	69
3.54 システムの応用	70
3.55 システムの応用	71
3.56 システムの応用	72
3.57 システムの応用	73
3.58 システムの応用	74
3.59 システムの応用	75
3.60 システムの応用	76
3.61 システムの応用	77
3.62 システムの応用	78
3.63 システムの応用	79
3.64 システムの応用	80
3.65 システムの応用	81
3.66 システムの応用	82
3.67 システムの応用	83
3.68 システムの応用	84
3.69 システムの応用	85
3.70 システムの応用	86
3.71 システムの応用	87
3.72 システムの応用	88
3.73 システムの応用	89
3.74 システムの応用	90
3.75 システムの応用	91
3.76 システムの応用	92
3.77 システムの応用	93
3.78 システムの応用	94
3.79 システムの応用	95
3.80 システムの応用	96
3.81 システムの応用	97
3.82 システムの応用	98
3.83 システムの応用	99
3.84 システムの応用	100
3.85 システムの応用	101
3.86 システムの応用	102
3.87 システムの応用	103
3.88 システムの応用	104
3.89 システムの応用	105
3.90 システムの応用	106
3.91 システムの応用	107
3.92 システムの応用	108
3.93 システムの応用	109
3.94 システムの応用	110
3.95 システムの応用	111
3.96 システムの応用	112
3.97 システムの応用	113
3.98 システムの応用	114
3.99 システムの応用	115
3.100 システムの応用	116

向井 利春

目次

1	序論	1
1.1	はじめに	1
1.2	センサフュージョン	2
1.3	視触覚融合	3
1.4	アクティブセンシング	5
1.5	研究の目的	6
1.6	本論文の構成	7
2	観測ノイズがオフセットを持つ場合の融合値と分散の求め方	9
2.1	はじめに	9
2.2	モデルが曖昧さを含む場合の推定	10
2.3	特定の場合同	13
2.3.1	同一センサで複数回観測	13
2.3.2	異なるセンサで一回ずつ観測	14
2.3.3	独立な複数センサで繰り返し観測する場合	14
2.4	モデルの曖昧さの推定	15
3	センサ位置選択の一般論	19
3.1	はじめに	19
3.2	誤差を小さくするセンサ位置選択の方法	19
3.3	アクチュエータの動きの誤差を含む場合	23
3.4	連合を良く行なえるセンサ位置選択の方法	23
3.5	評価関数の設定と融合の仕方	25
3.6	計算アルゴリズム	26
3.6.1	対象が1つの場合	26
3.6.2	対象が複数の場合	27
3.7	シミュレーション—ターゲットトラッキング	28

4 触覚センサによる接触位置の求め方	31
4.1 はじめに	31
4.2 対象認識	32
4.3 形状モデルに誤差がない場合の位置計測	36
4.4 形状モデルに誤差がある場合の位置計測	41
4.5 シミュレーション	44
5 逆問題を簡単化するアクティブセンシング	47
5.1 一般論	47
5.2 接触による接触対象の求め方	48
5.3 カメラ-マニピュレータ間の座標変換を求める	53
6 実験	55
6.1 システム構成	56
6.1.1 システム全体	56
6.1.2 マニピュレータ	56
6.1.3 フォース/トルクセンサ	60
6.2 カメラ-マニピュレータ変換	61
6.3 フォース/トルクセンサによるエッジ位置の計測	62
6.4 視触覚による対象認識	65
6.5 視触覚融合による精度の良い位置計測	69
7 結論	73
謝辞	75
参考文献	76

第 1 章

序論

1.1 はじめに

知能ロボットのように、未知、または限定された知識しか得られていない環境内でタスクを行なうシステムでは、多くの、しかも種類の異なったセンサを使って、環境の情報を取り込む能力が必要不可欠である。このように複数のセンサを持ったシステムによって、単一のセンサからは得られない機能を工学的に実現しようという手法はセンサフュージョンと呼ばれ、近年盛んに研究されている [山崎, 石川, 1992] [Luo and Kay, 1989]。

センサフュージョンで使われるセンサには様々な種類があるが、空間知覚を得るためには、視覚と触覚が代表的に使われる。視覚は大域的だが粗く、触覚は局所的だが接触部分の正確な情報を得られるというように、視覚と触覚にはそれぞれ補い合う特徴があるので、組み合わせることで、より有効なセンシングが可能となる。特に、触覚はその局所性ゆえに、有効に使うためには視覚の補助が必要である場合が多い。このように視覚と触覚を使ったシステムは視触覚融合システムと呼ばれ、対象認識などで研究されている。

また、一般にセンサフュージョンシステムは、何らかのタスクを実現するというシステムの目的から、アクチュエータを備えていることが多い。例えば視触覚融合システムの多くでは、対象のハンドリングが最終目標として考えられ、マニピュレータやハンドを備えている。このようなシステムでは、アクチュエータを本来の目的に加えて、センシングを構成する一要素として使うことができる。特に、触覚センサはその局所性を補うために、アクチュエータと共に使われることが不可欠である。このようにアクチュエータをセンシングシステムの一部として利用する考え方はアクティブセンシングと呼ばれおり、視触覚融合システムのように、触覚を用いるシステムでは必然的に生じてくる考え方である。

これまでいくつかの視触覚融合システムが製作されたり、シミュレーションが行なわれたりしてきたが、それらは視触覚融合システムが必然的に持つアクティブセンシングという面を有効に利用しているとは言いがたい。わずかに、あらかじめ決められたいくつかの方向のどちらからなら触覚センサで対象に触れるかを視覚から求めて、触覚センサのガ

イドに使った研究や、対象認識で対象の確認のためにはどこを探索したら良いかを求めている研究があるが、触覚センサによってどこをどのように探索すれば効率的なセンシングができるかを自律的に求めるという問題はほとんど考えられていない。

そこで本論文では、アクティブセンシングを利用した視触覚融合の方法を提案する。その中でも特に、触覚センサで探索する位置を系統的に求める方法に重点を置く。視覚でおおまかな情報を得た後に、より正確な対象認識や対象位置の計測を行なうためには触覚センサでどのような探索を行えば良いかをセンシングの効率の面から求める。さらに、提案した方法の有効性を確認するために行なった実験について述べる。

なお、本論文で述べる手法は視触覚融合システムのために考えられたものであるが、センサフュージョンシステム一般への拡張は容易である。

1.2 センサフュージョン

人間は、俗に五感と言われるように、多数、多種類のセンサを使って実世界の情報を取り入れ、それらの情報と既に持っている知識を統合・融合することにより、実世界を理解・認識している。このような、複数センサからの情報を統合的に処理することにより、単一のセンサからは得られない機能を工学的に実現しようとする考え方がセンサフュージョン (sensor fusion) であり、自律移動車 (ALV: Autonomous Land Vehicle) [Shafer et al., 1986], 知能ロボットやヒューマンインターフェイス [松本, 亀島, 1990] などの分野で盛んに研究されている [山崎, 石川, 1992] [Luo and Kay, 1989] [Mukai et al., 1993]。

このセンサフュージョンという言葉は、字義通りに解釈すると複数のセンサの情報を融合・結合するという意味であるが、実際にはより大きな概念を表すために用いられることが多い。山崎ら [山崎, 石川, 1992] は、前者を狭義の“fusion (融合)”と呼び、「いくつかのセンサの情報から一つの知覚表象を得る過程」と位置づけている。一方、後者を「複数のセンサ情報の処理の総体を対象としており、より広い概念で扱おうとするもの」とし、カタカナの「センサフュージョン」という言葉を当てている。そして、「従来のセンシング技術は、計測にかかわるある現象において、被測定量が観測量に及ぼす因果過程に対する逆向きの求解過程を意味していたのに対し、センサフュージョンでは、これに加えて対象とする現象を選択・設計・評価・判断する過程をもセンシングの問題として取り込まなければならない」と述べている。

また、Luoら [Luo and Kay, 1989] は、前者を Multisensor fusion と呼び、“any stage in the integration process where there is an actual combination (or fusion) of different sources of sensory information into one representational format” と定義している。また、“The information to be fused may come from multiple sensory devices during a single period of time or from a single sensory device over an extended time period” とし、単一のセンサから異なる時間に得られた情報を融合する場合も Multisensor fusion の

一つであるとしている。一方、後者を Multisensor integration と呼び、“synergistic use of the information provided by multiple sensory devices to assist in the accomplishment of a task by a system” と定義している。Synergistic use という言葉は、“This synergistic effect from the multisensor integration can be achieved either by using the information from one sensor to provide cues or guide the operation of other sensors, or by actually combining or fusing the information from multiple sensors” という文からわかるように、あるセンサで別のセンサをガイドするような場合も含む、広い概念の言葉である。

本論文では「センサフュージョン」という言葉で後者を示し、特に前者を区別する場合には「狭義のセンサフュージョン」という言葉を用いることにする。

異なるセンサからの情報を融合するアルゴリズムは、おおまかに次の3種類に分類される [山崎, 石川, 1992]。一つはカルマンフィルタや最小2乗法などの信号処理・統計処理を使う方法である。次に、データベースなどの人工知能・知識工学を使う方法がある。また、対象についての知識が少ない場合には、ニューラルネットワークが使われる場合もある。本論文では、アクティブセンシングを行なうために測定値の分散を使用するので、測定値とともに分散を求められる方法として、カルマンフィルタを拡張した方法を利用する。

また、センサフュージョンで行なわれる異なるセンサからの情報の利用の仕方として、

1. 複数のセンサからの観測値を、処理系内だけで処理してより高次の情報を得る場合（狭義のセンサフュージョン）と、
2. 逐次的に観測値が得られ、先に得た情報を使って、次に行なうセンシングにより良く適するようにセンシングシステムの構成に変更を加える場合

とが考えられる。2 はアクティブセンシングに通じる考え方である。2 の代表的な適用分野として、視触覚融合での視覚による触覚センサのガイドが上げられるが、今までの研究ではこの機能を有効に生かしているとは言えない。例えば Stansfield [Stansfield, 1988] は、あらかじめ決めたいくつかの方向のどちらからなら触覚センサが対象に触れるかを視覚から求めて、触覚センサのガイドに使っているが、視覚で得られた対象のどの部分を優先的に探索するかというような問題に対しては何も答えていない。

本論文で扱う視触覚融合では、触覚センサの探索行動を視覚でガイドする系統的方法を提案するが、これは2にあたり、その後、視覚と触覚から得た情報を融合することにより対象についての精度の良い情報を得るが、これは1に相当する。

1.3 視触覚融合

視覚と触覚は空間知覚を得るための代表的なセンサである。

視覚は、非接触で対象についての大局的な情報が得られるが、そのかわり対象の3次元位置を簡便に精度良く得ることは難しい。また、対象の裏側やオクルージョンの起こっ

ている領域の情報は得られない。さらに、反射率や照明の変化により画像が変化するので、画像のみから必要とされている情報を全てを得ることは難しい。

一方、触覚センサは接触している対象についての3次元情報は得やすいが、センサの接触している部分のみの情報しか得られない。つまり、情報は非常に局所的である。そのため、対象全体の情報を得るためには、触覚センサを動かすアクティブセンシングが必要になる。これを行なうためにはセンサの動かし方を決めなければならないが、それには対象の大局的な情報が必要なので、視覚から得た情報を補助として使うことが有効である場合が多い。

また、視覚や触覚の一方でしか得られない情報もある。例えば、色情報は視覚のみからしか得られないし、対象のコンプライアンス、熱特性、弾性などの情報は触覚からしか得られない。

そこで、視覚と触覚の両方を使うシステムが有効であると考えられ、視触覚融合システムとして研究されている。Allen [Allen, 1988] は動眼機構を持った2台のCCDカメラによるステレオビジョンと、マニピュレータに装着した触覚センサからの情報を使って、コーヒーカップなどの認識を行なっている。このシステムは物体のモデルをデータベースの形で持っており、データベースとのマッチングを行なうための前処理として、視覚と触覚の融合を行なっている。また、Stansfield [Stansfield, 1988] 固定された2台のカメラとマニピュレータに取りつけた触覚センサを使って同様のことを行なっている。このシステムでは、最初に一回取った視覚情報を用いて触覚のガイドを行なっているが、それは上、前、左右のどちらからなら対象に触れられるかを判別するというように、単純なレベルである。また、Yamada et al. [Yamada et al., 1993] は視覚の大局的で粗い情報と、触覚の詳細しくて局所的な情報を統合するために、どのような形状の内部モデルを使うかについて提案シミュレーションを行なっている。また、坂根ら [坂根ら, 1994] は視覚センサから得られる情報と、ロボットハンドの手首に装着した力覚センサから得られる情報とを融合することにより、把持物体と他物体の接触位置を推定する方法を提案し、ロボットシステムで実験を行なっているが、積極的に何かを探索するというようなことはしていない。

このようにいくつかの視触覚融合システムが提案されているが、多くのシステムでは触覚は視覚で得られた対象の領域全て、または ad hoc に選んだ位置を探索している。効率の良いセンシングを行なうためには、触覚を用いてどこを探索すればよいかを決めなければならないが、これを系統的に求める方法は考えられていない。そこで本論文では、視覚情報から触覚による探索の仕方を求めることを主な目的とし、これに関して新たな提案を行なう。

なお、視触覚融合システムと関連するものとしてハンドアイシステムがある [稲葉, 井上] [坂根ら, 1986]。これはセンサとして視覚のみを用い、ハンドはアクチュエータとしてのみ用いるシステムであるので、触覚によってもセンシングを行なおうとする視触覚融合とは異なる。

1.4 アクティブセンシング

センシングシステムの一部としてアクチュエータを利用する手法は、アクティブセンシング (active sensing) として知られている [石川, 1988] [Shimojo and Ishikawa, 1993] [Aloimonos et al., 1988] [石川, 1993] [金子, 1993] [木下, 出口, 1992] [金子, 1992] [山口, 1993] [平井, 1988]。

アクティブセンシングはまだ発展途中の分野であり、その言葉の意味も明確に定まっているわけではない。エネルギーを供給してセンサが電波などを発することでセンシングを行なうレーダーのようなものをアクティブセンシングと呼ぶ場合もあるが、ここでは、アクチュエータがセンサと一体となってシステムを作っていて、センサやセンシング対象を動かすことによりセンシングを行なうものをアクティブセンシングと呼ぶ。

アクティブセンシングと類似の考え方は、心理学の分野でも、Gibson [Gibson, 1970] のアフォーダンスの概念 (対象は、それに対する行為も含んだ存在である) や Neisser [Neisser, 1976] の知覚循環 (知覚は、計測 → モデルの同定 → 予測 → 計測を繰り返す) などに見られ、これらは人間が行なうセンシングの基本動作の一つと考えられている。

触覚における弾性や摩擦のように、アクティブセンシングを行なわなければ得られない量もあり、これらのセンシングも興味深い分野ではあるが、本論文では分野を限定することによって議論を明確にするために、位置情報の取得に関するアクティブセンシングを主に扱う。このようにアクティブセンシングの一部に限定した場合でも、この方法の利点として、次のようなものが挙げられる。

1. センサを移動することにより、単一のセンサで、異なる場所に複数のセンサがあるのと同じ効果を得られる。これにより、システムのセンサ数が減り配線も減るので、コストが下がり、信頼性が上がる。
2. 上と関係するが、触覚センサのように局所的なセンサは、動かすことにより大局的な情報を得ることができる。
3. それまでに得た情報を利用することにより、次に観測を行なう観測のためのより良いセンサ位置を逐次的に決められる。
4. 立体視などを行なう場合、異なる視点からの画像の対応点のマッチングが必要である。このように対応点のマッチングが必要な場合、1つのセンサを動かして異なった位置からの画像を得れば、追跡を行なうことにより対応点が見つけやすい [Aloimonos et al., 1988]。
5. 異なる位置にある複数の固定したセンサと、その位置間を一つのセンサで移動することを比べると、単一のセンサを移動する方法は移動中にも情報が得られるため、情

報量をはるかに多い。そのため、情報にノイズが乗っているような場合には、カルマンフィルタなどを利用することにより精度の良い情報が得られる。

6. 測定の際に、測定に関係するパラメータの一部を自由に決められることにより、精度が良くなったり、得たい量を計算する式が簡単になったりする。これは後の第5章でより詳しく考察する。
7. 対象を操作する場合、必要な情報はアクチュエータ座標系での対象の位置である。そこで計測を、対象の操作に使うアクチュエータ自身で行なうことにより、センサからアクチュエータへの座標変換の誤差から自由になる。また、アクチュエータのキネマティクスの誤差が吸収できる。これにより、例えばマニピュレータのリンクの長さがモデルと実際とで少し異なっていたりしても、アクチュエータ座標系上での対象の位置にはほとんど影響がなくなり、タスクを実現する際の支障がなくなる。

このようにアクティブセンシングには多くの利点が考えられる。しかし、計測技術としてのアクティブセンシングはいくつかの研究がなされてはいるものの、まだ定義もはっきりとは定まっておらず、今後の研究が期待されている。

アクティブセンシングを実現するためには、未知の被測定対象の計測に際し、それまでに得られた情報を最大限に利用して、より良い計測ができるように次のアクチュエータの動作を決定する必要がある。これはセンシング戦略を立てることに対応する。

1.5 研究の目的

上述したように視触覚融合システムでは、触覚の局所性を避けるために触覚センサを動かすアクチュエータが必要である。また、視触覚融合システムの最終的な目的としては、対象をマニピュレータによりハンドリングすることである場合が多い。これにより、視触覚融合システムは必然的にアクチュエータを含むシステムとなる。そのようなシステムでは、アクティブセンシングを考えることが自然である。

しかし、現在までに研究されている視触覚融合システムは、アクティブセンシングという面を積極的に生かしているとはいえない。特に、効果的なセンシングを行なうために触覚センサでどこを探索するかを系統的に求めているシステムはない。

そこで本論文では、アクティブセンシングを積極的に用いた視触覚融合の方法を提案する。その中でも特に、触覚センサで探索する位置を求める方法に重点を置く。この方法では測定値の分散を、センシングに対する知識として使用する。センサ出力を融合し、それと同時にその値の分散を求める方法としてはカルマンフィルタが有名であるが、これはセンサのモデルに含まれるノイズの平均値が0であるという仮定をしている。その結果、観測を繰り返すと分散が限りなく0に近付いてしまい、本論文で用いる方法に適さない。

そこで、複数センサの情報を融合する方法としてセンサモデルのノイズの平均値にオフセットがあっても良い方法を提案し、本論文ではこれを使用する。

次に、より良い観測を行なうためのセンサ位置選択の方法の一般論を提案する。ここでは、良いセンサ位置の基準として、まず、求めたい値の推定値の分散が小さいことを採用する。さらに、観測対象が複数ある場合を考慮し、これらの対象とセンサ出力の対応がとりやすいことをもう一つの基準とする。これらの基準からセンシングを行なう際に望ましいセンサ位置を求める。

また、以上で述べた融合の方法と、センサ位置選択の方法を合わせて、視覚でおおまかに対象をとらえた後に、対象認識や対象の位置を求める際には触覚センサでどこを探索すれば良いかを求める方法を提案する。そのために、対象の3次元形状モデルを用い、この形状モデルが誤差を含んでいても使用できるような、一般的な方法を提案する。

さらに、センサをアクティブに動かすことにより測定の状況をコントロールし、精度の良い計測を行なったり、逆問題を解き易くする方法を提案する。またこの方法の具体例として、フォース / トルクセンサにより対象との接触位置を求める方法と、CCDカメラとマニピュレータ間の座標変換を求める方法が提案される。

最後に、提案した方法を CCDカメラ、マニピュレータ、フォース / トルクセンサから構成される実験システムに実装し、そのシステムを使った実験により提案した方法の有効性を確認する。なお、分布型触覚センサは扱いが大変なため、この実験では触覚情報を得るためのセンサとしてフォース / トルクセンサを使用する。

これらの研究によって、アクティブセンシングを効果的に利用した視触覚融合システムが構成できることがわかった。特に、触覚センサで探索する位置を視覚から求める方法が明らかになった。これにより、対象認識や対象の位置を求める際に、効率的なセンシングが行なえるようになった。

1.6 本論文の構成

本論文の構成を図 1.1 に示す。

本章以下では、まず第 2 章で、センサモデルの観測ノイズの平均値にオフセットがあっても良い場合に複数センサの情報を統合し、同時に融合値の分散を求める方法を提案する。次に第 3 章で、より良い観測を行なうためのセンサ位置選択の方法の一般論を提案する。そして第 4 章でこの 2 つを統合することによって、視触覚システムにおいて、視覚で得た大まかな情報を使って、触覚センサで探索する位置を求める方法を提案する。

また、第 5 章ではセンサをアクティブに動かすことにより測定の状況をコントロールし、精度の良い計測を行なったり、逆問題を解き易くする方法を提案する。この方法の具体例として、フォース / トルクセンサにより対象の位置を求める方法と、CCDカメラとマニピュレータ間の座標変換を求める方法が提案される。

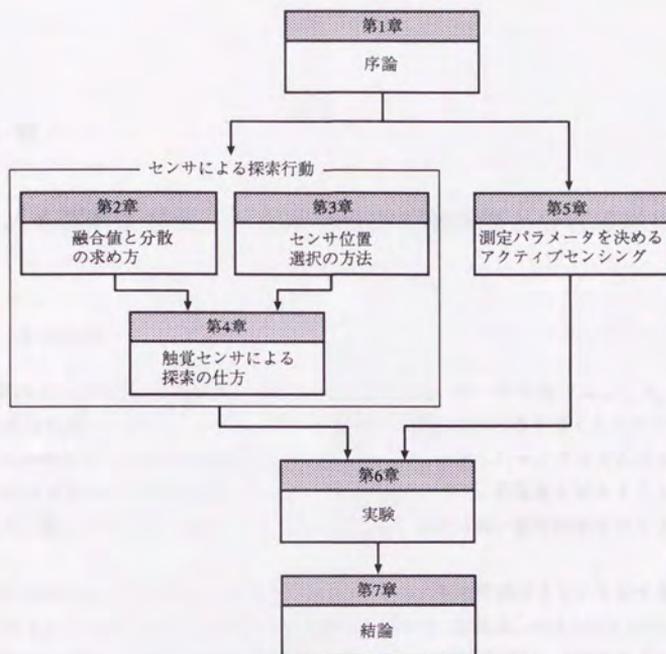


図 1.1 本論文の構成

次の第6章で、上で述べた方法を実装する視触覚融合システムについて説明し、そのシステムを使った実験について述べる。実験では、扱いの難しい分布型触覚センサは使わず、第5章で提案したフォース / トルクセンサにより触覚情報を得る方法を用いる。最終章で、全体をまとめた結論を述べる。

第 2 章

観測ノイズがオフセットを持つ場合の融合値と分散の求め方

2.1 はじめに

複数のセンサの出力を融合する際には、信頼度の高いデータを選び出したり、データがどの程度信頼できるかによって、データに対する重みの付け方を変えたりすることが良く行なわれる。データを取捨選択することによって、故障したセンサなどの全く信用できないデータをはずして精度を上げることができるし、また、計算量を減らすことができる。また、重みのつけ方を変えることによって、より精度の高い融合結果を得ることができる。

そのためには、センサ出力とともに、その出力がどの程度信頼できるかを表す量を扱う必要がある。この量はさまざまな方法で定義されており、信頼度, confidence factor, 無矛盾性 [McKendall and M.Mintz, 1988] [Luo and Kay, 1989] [浅田ら, 1990] などの名前で呼ばれている。これらの値は ad hoc に決められているものも多いが、その中で、推定値の誤差の分散を使う方法は、理論的な根拠が明確である。例えばカルマンフィルタ (Kalman filter) [片山, 1983] [Ayache, 1991] は、推定値とともにその共分散行列も推定していて、その値を推定値を更新する際に使っている。しかし、この方法には信頼度が無限大になってしまうという問題点がある。

ここで、単純な場合でカルマンフィルタによって得られる分散の問題点を説明する。ノイズが平均ゼロのガウシアンであることを仮定すると、推定値の分散はバイズの定理から以下のように計算できる。

$$1/\sigma_{\text{fused}}^2 = 1/\sigma_a^2 + 1/\sigma_b^2 \quad (2.1)$$

ここで、 σ_a, σ_b はそれぞれ、観測値 a, b の分散であり、 σ_{fused} は a と b を融合したデータの分散である。

しかしこれでは、融合を行なえば行なうほど分散が無限に小さくなってしまい、確率分布はどんどん急峻になり、推定値が無限に信用できることになってしまう。つまり、一つのセンサで何回も観測すると、いくらでも分散の小さい、つまり、いくらでも信頼性の高い

データが得られることになる。しかし、これは現実にはおかしい。実際の融合では、信頼度は最初のうちは向上するが、ある程度融合を重ねると飽和すると考えられる。

なぜこのようなことが起こるかという点、モデルが完全に正しいと仮定していることに原因がある。特にノイズの平均がゼロとする所に大きな問題がある。このために、観測を重ねれば重ねるほど、分散が小さくなる。実際にはオフセットは必ず残るはずであるが、多くの研究ではオフセットについて考えられていない。文献 [McKendall and M. Mintz, 1988] ではオフセットについて考えているが、センサ出力は $z = x + \text{noise}$ の形に限られ、また、逐次的な計算方法も考えられていない。

そこで、本章では、センサのモデルに誤差があつてオフセットが残っていると、これが確率的にしかわからないとしたときの、融合の方法を考える。同一センサからのデータを融合してもオフセットは変わらないが、異なるセンサからのデータを融合すると、オフセットについてもより信頼できる値が得られるとする。さらに、この計算を逐次的に行なえる方法を提案する。

2.2 モデルが曖昧さを含む場合の推定

センサのモデルを以下のように定義する。センサ i の時刻 t での観測を

$$z_i(t) = H_i x(t) + v_i(t) \quad (2.2)$$

とする。カルマンフィルタなどでは $v_i(t)$ を平均ゼロのガウシアンノイズと仮定するが、ここでは $v_i(t)$ を平均が m_i のガウシアンノイズと仮定する。言い換えると、ノイズを次のように2種類に分けて考える。一つはセンサをモデル化するときに生じる曖昧さ m_i であり1つのセンサについては固定されているとする。これをモデルの曖昧さと呼ぶ。もう一つは測定ごとに生じるノイズ $s_i(t)$ であり測定時ノイズとよぶ。そして $v_i(t) = m_i + s_i(t)$ を観測ノイズと呼ぶ。カルマンフィルタではこれをまとめて平均ゼロのガウシアンノイズとしている。これに従って上式を書き換えると、

$$z_i(t) = H_i x(t) + m_i + s_i(t) \quad (2.3)$$

のようになる。

測定ノイズ $s_i(t)$ は、平均ゼロ、分散 $S_i = \sigma_{S_i}^2$ のガウシアンノイズ

$$N(0, S_i) \quad (2.4)$$

であるとする。特定のセンサ i での、観測についての期待値と分散を $E_i[\cdot]$ 、 $\text{Var}_i[\cdot]$ と表すと

$$E_i[s_i(t)] = 0 \quad (2.5)$$

$$\text{Var}_i[s_i(t)] = S_i \quad (2.6)$$

である。

モデルの曖昧さ \mathbf{m}_i の確率分布は平均ゼロ、分散 $M_i = \sigma_{M_i}^2$ のガウス分布

$$N(0, M_i) \quad (2.7)$$

で与えられるとする。これは、一つひとつのセンサについてはオフセット \mathbf{m}_i が固定されているが、異なるセンサではこれが異なり、たくさんのセンサを比較するとオフセットが与えられるようなガウス分布になるということを意味する。センサのモデルを作る時の可能性で期待値と分散をとることを $\bar{E}[\cdot]$, $\bar{\text{Var}}[\cdot]$ と表すと、

$$\bar{E}[\mathbf{m}_i] = 0 \quad (2.8)$$

$$\bar{\text{Var}}[\mathbf{m}_i] = M_i \quad (2.9)$$

である。

これで、一つのセンサでいくら観測しても分散はゼロにならず、また、異なるセンサと融合すればオフセットの分散が小さくなるというような、直観と一致する結果が得られる。

この考えかたでより詳細な考察を試みる。まず、一般的な場合について述べる。

以上で述べた観測のモデルをもう一度まとめて書く。センサ N 個での観測を

$$\mathbf{z}_i(t) = H_i(t)\mathbf{x} + \mathbf{m}_i + \mathbf{s}_i(t) \quad (2.10)$$

と定義する。それぞれのセンサはモデルの曖昧さ \mathbf{m}_i ($i = 1, \dots, N$) を持つとする。また、測定時ノイズを $\mathbf{s}_i(t)$ ($i = 1, \dots, N$) とする。それぞれのノイズの平均をゼロとする。つまり

$$E[\mathbf{s}_i(t)] = \bar{E}[\mathbf{m}_i] = 0 \quad (2.11)$$

である。ノイズの独立性を

$$E[\mathbf{s}_i(t)\mathbf{s}_j^T(l)] = \delta_{ij}S_i, \quad \bar{E}[\mathbf{s}_i(t)\mathbf{m}_j^T] = 0,$$

$$E[\mathbf{s}_i(t)\mathbf{m}_j^T] = 0$$

のように定義する。

一般的な式を導くため、(2.10)の i, t による番号づけを付け変えて、 λ だけで式の index となるようにする。観測

$$\mathbf{z}(\lambda) = H(\lambda)\mathbf{x} + \mathbf{m}(\lambda) + \mathbf{s}(\lambda) \quad (2.12)$$

が $\lambda = 1, \dots, A$ で行なわれるとする。これは、とりあえずどのセンサで観測が行なわれたかは無視し、単に一回の観測と考えることを意味する。実際は $H(\lambda) = H_i(t)$, $\mathbf{m}(\lambda) = \mathbf{m}_i$, $\mathbf{s}(\lambda) = \mathbf{s}_i(t)$ のような形となる。

そのときの推定値は仮に $\mathbf{m}(\lambda)$ がわかっているとすると

$$\eta(\lambda) \equiv \mathbf{z}(\lambda) - \mathbf{m}(\lambda) = H(\lambda)\mathbf{x} + \mathbf{s}(\lambda) \quad (2.13)$$

にカルマンフィルタ (を少し変形したもの) を適用することにより

$$\hat{\mathbf{x}}(\lambda) = \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \right\}^{-1} \cdot \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{z}(\lambda) - \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{m}(\lambda) \right\} \quad (2.14)$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S(\lambda) \quad (2.15)$$

$$P_S^{-1}(\lambda) = \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \quad (2.16)$$

となる。ここで、 $\mathbf{m}(\lambda)$ が確率変数であることを考慮すると

$$\hat{\mathbf{x}}(\lambda) = \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \right\}^{-1} \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{z}(\lambda) \quad (2.17)$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S(\lambda) + P_M(\lambda) \quad (2.18)$$

$$P_S^{-1}(\lambda) = \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \quad (2.19)$$

$$P_M(\lambda) = \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \right\}^{-1} \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{m}(\lambda) \right\} \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{m}(\lambda) \right\}^T \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} H(\lambda) \right\}^{-T} \quad (2.20)$$

となる。 λ の逐次式になるようにこれを書き換える。まず、

$$\hat{\mathbf{x}}(\lambda) = P_S^{-1}(\lambda) (P_S(\lambda-1) \hat{\mathbf{x}}(\lambda-1) + H(\lambda)^T S(\lambda)^{-1} \mathbf{z}(\lambda)) \quad (2.21)$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S(\lambda) + P_M(\lambda) \quad (2.22)$$

$$P_S^{-1}(\lambda) = P_S^{-1}(\lambda-1) + H(\lambda)^T S(\lambda)^{-1} H(\lambda) \quad (2.23)$$

$$P_M(\lambda) = P_S(\lambda) L(\lambda) P_S^T(\lambda) \quad (2.24)$$

$$L(\lambda) = \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{m}(\lambda) \right\} \left\{ \sum_{\lambda} H(\lambda)^T S(\lambda)^{-1} \mathbf{m}(\lambda) \right\}^T \quad (2.25)$$

とする。ここで更に $L(\lambda)$ を逐次的に書き換える。 λ 回目の観測がセンサ i の t 回目だということ、つまり

$$H(\lambda) = H_i(t), \quad \mathbf{m}(\lambda) = \mathbf{m}_i, \quad \mathbf{s}(\lambda) = \mathbf{s}_i(t) \quad (2.26)$$

だということを

$$I(\lambda) = i, \quad T(\lambda) = t \quad (2.27)$$

と書くことにする。

$$\tilde{L}_i(\lambda) = \begin{cases} \tilde{L}_i(\lambda-1) + H(\lambda)^T S(\lambda)^{-1} = \tilde{L}_i(\lambda-1) + H_i(T(\lambda)) S_i(T(\lambda)) & \text{if } I(\lambda) = i \\ \tilde{L}_i(\lambda-1) & \text{otherwise} \end{cases} \quad (2.28)$$

と定義すると

$$L(\lambda) = \left\{ \sum_{i=1}^N \tilde{L}_i(\lambda) \mathbf{m}_i \right\} \left\{ \sum_{j=1}^N \tilde{L}_j(\lambda) \mathbf{m}_j \right\}^T \quad (2.29)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \tilde{L}_i(\lambda) M_{ij} \tilde{L}_j^T(\lambda) \quad (2.30)$$

$$= \sum_{i=1}^N \tilde{L}_i(\lambda) M_{ii} \tilde{L}_i^T(\lambda) + \sum_{i=2}^N \sum_{j=1}^{i-1} (\tilde{L}_i(\lambda) M_{ij} \tilde{L}_j^T(\lambda) + \tilde{L}_j(\lambda) M_{ij} \tilde{L}_i^T(\lambda)) \quad (2.31)$$

$$(2.32)$$

となる。

まとめると

$$\hat{\mathbf{x}}(\lambda) = P_S^{-1}(\lambda) (P_S(\lambda - 1) \hat{\mathbf{x}}(\lambda - 1) + H(\lambda)^T S(\lambda)^{-1} \mathbf{z}(\lambda)) \quad (2.33)$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S(\lambda) + P_M(\lambda) \quad (2.34)$$

$$P_S^{-1}(\lambda) = P_S^{-1}(\lambda - 1) + H(\lambda)^T S(\lambda)^{-1} H(\lambda) \quad (2.35)$$

$$P_M(\lambda) = P_S(\lambda) L(\lambda) P_S^T(\lambda) \quad (2.36)$$

$$L(\lambda) = \sum_{i=1}^N \tilde{L}_i(\lambda) M_{ii} \tilde{L}_i^T(\lambda) + \sum_{i=2}^N \sum_{j=1}^{i-1} (\tilde{L}_i(\lambda) M_{ij} \tilde{L}_j^T(\lambda) + \tilde{L}_j(\lambda) M_{ij} \tilde{L}_i^T(\lambda)) \quad (2.37)$$

$$\tilde{L}_i(\lambda) = \begin{cases} \tilde{L}_i(\lambda - 1) + H(\lambda)^T S(\lambda)^{-1} = \tilde{L}_i(\lambda - 1) + H_i(t) S_i(t) & \text{if } I(\lambda) = i \\ \tilde{L}_i(\lambda - 1) & \text{otherwise} \end{cases} \quad (2.38)$$

となり、逐次的に推定値とその分散を計算するための一般式が得られる。

2.3 特定の場合

第2.2節で上げた式をわかり易くするため、特定の状況に適用した場合、式がどのようになるかについて説明する。

2.3.1 同一センサで複数回観測

同一センサで複数回観測した場合を考える。観測を

$$\mathbf{z}(t) = H\mathbf{x} + \mathbf{m} + \mathbf{s}(t) \quad (2.39)$$

と定義する。第2.2節で述べた式は

$$\hat{\mathbf{x}}(t) = P_S^{-1}(t) (P_S(t - 1) \hat{\mathbf{x}}(t - 1) + H^T S^{-1} \mathbf{z}(t)) \quad (2.40)$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S(t) + P_M(t) \quad (2.41)$$

$$P_S^{-1}(t) = P_S^{-1}(t - 1) + H^T S^{-1} H = t H^T S^{-1} H \quad (2.42)$$

$$\begin{aligned}
 P_M(t) &= P_S(t)L(t)P_S^T(t) \\
 &= P_S(t)(tHS^{-1})M(tHS^{-1})^T P_S^T(t) \\
 &= (H^T S^{-1}H)^{-1}H^T S^{-1}MS^{-T}H(H^T S^{-1}H)^{-T}
 \end{aligned} \tag{2.43}$$

となる。

この式に従うと、 P_M は何回融合しても変化しない。よって、

$$\lim_{t \rightarrow \infty} \text{Var}[\hat{\mathbf{x}}(t)] = (H^T S^{-1}H)^{-1}H^T S^{-1}MS^{-T}H(H^T S^{-1}H)^{-T} \neq 0 \tag{2.44}$$

である。これにより、観測結果の分散がゼロにはならなくなり有限の値に保たれるので、推定値が無限に信用できることになるという問題はなくなる。

2.3.2 異なるセンサで一回ずつ観測

次にセンサ1とセンサ2でそれぞれ1回観測した結果を融合する場合を考える。観測を

$$\mathbf{z}_1 = H_1 \mathbf{x} + \mathbf{m}_1 + \mathbf{s}_1 \tag{2.45}$$

$$\mathbf{z}_2 = H_2 \mathbf{x} + \mathbf{m}_2 + \mathbf{s}_2 \tag{2.46}$$

とする。ただし、センサ1と2のモデルの曖昧さは独立、つまり $M_{12} = 0$ とする。すると

$$\hat{\mathbf{x}} = P_S^{-1}(H_1^T S_1^{-1} \mathbf{z}_1 + H_2^T S_2^{-1} \mathbf{z}_2) \tag{2.47}$$

$$\text{Var}[\hat{\mathbf{x}}] = P_S + P_M \tag{2.48}$$

$$P_S^{-1} = H_1^T S_1^{-1} H_1 + H_2^T S_2^{-1} H_2 \tag{2.49}$$

$$P_M = P_S(H_1^T S_1^{-1} M_1 S_1^{-T} H_1 + H_2^T S_2^{-1} M_2 S_2^{-T} H_2) P_S^T \tag{2.50}$$

となる。

特に、 $H = H_1 = H_2$, $S = S_1 = S_2$, $M = M_1 = M_2$ の場合

$$P_M = \frac{1}{2}(H^T S^{-1}H)^{-1}H^T S^{-1}MS^{-T}H(H^T S^{-1}H)^{-T} \tag{2.51}$$

となり、分散はセンサ1つで観測した場合の半分になる。この式は、同一の精度を持つと考えられるセンサが複数ある場合、同一のセンサで繰り返し観測するよりも、異なるセンサで観測する場合の方が精度が良くなることを表している。つまり、センサフュージョンの状況で、異なるセンサを使うことの意義を数式で表している。

2.3.3 独立な複数センサで繰り返し観測する場合

それぞれのセンサが独立、つまり $M_{ij} = 0$ ($i \neq j$) の場合を考える。この場合には第2.2節の $L(\lambda)$ を計算する式は以下のように計算できる。

$$L(\lambda) = \sum_{i=1}^N \bar{L}_i(\lambda) \tag{2.52}$$

$$\bar{L}_i(\lambda) = \tilde{L}_i(\lambda)M_{ii}\tilde{L}_i^T(\lambda) \quad (2.53)$$

とすると

$$\bar{L}_i(\lambda) = (\tilde{L}_i(\lambda-1) + H^T(\lambda)S^{-1}(\lambda))M_{ii}(\tilde{L}_i(\lambda-1) + H^T(\lambda)S^{-1}(\lambda))^T \quad (2.54)$$

$$\begin{aligned} &= \tilde{L}_i(\lambda-1) + H^T(\lambda)S^{-1}(\lambda)M_{ii}\tilde{L}_i^T(\lambda-1) + \tilde{L}_i(\lambda-1)M_{ii}S^{-T}(\lambda)H(\lambda) \\ &\quad + H^T(\lambda)S^{-1}(\lambda)M_{ii}S^{-T}(\lambda)H(\lambda) \end{aligned} \quad (2.55)$$

となり、 $\bar{L}_i(\lambda-1)$ が逐次的に計算できる。これは (2.37), (2.38) の代わりに使うことができる。なお、 $M_{ij} \neq 0$ の場合に (2.37) の第 2 項について同様の式を求めることも容易にできる。

2.4 モデルの曖昧さの推定

実際の状況では、モデルの曖昧さ M_{ij} がわかっていないことが多いと考えられる。そこで、センサが多数あって、観測値 z_i ($i = 1, \dots, N$) が同じモデルの曖昧さ M に従って得られたときに M を推定することを考える。ただし、 S_i はわかっているものとする。

例えば同一の M に従っているとわかるセンサが同一の対象を観測した場合の z_i どうしが近ければ、 M は小さいと推定できる。真値が \tilde{z} のときの複数のセンサの出力は図 2.1 のように表せる。

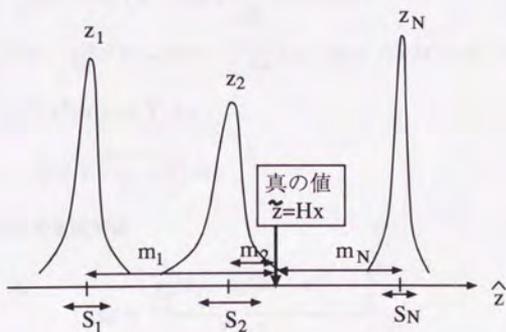


図 2.1 センサ出力の様子

今まで通り、

$$z_i = H_i x_i + m_i + s_i \quad (2.56)$$

という観測を考える。

まず簡単のため、センサ2個の場合を考える。

$$E[(z_1 - z_2)(z_1 - z_2)^T] = E[(v_1 - v_2)(v_1 - v_2)^T] \quad (2.57)$$

$$= E[(m_1 + s_1 - m_2 - s_2)(m_1 + s_1 - m_2 - s_2)^T] \quad (2.58)$$

$$= 2M + S_1 + S_2 \quad (2.59)$$

だから

$$M = \frac{(z_1 - z_2)(z_1 - z_2)^T - S_1 - S_2}{2} \quad (2.60)$$

が M の推定値となる。測定値によっては $M \leq 0$ となってしまうが、その場合には十分に小さな値を M にとる。

一般的にして、 N 個のセンサで観測を行なうとする。 z の推定値を

$$\hat{z} = \frac{1}{N} \sum_{j=1}^N \hat{z}_j \quad (2.61)$$

とする。ただし、 \hat{z}_j はセンサ j についての (複数の) 観測値から求めた推定値であり、以下の式では s_j は測定時ノイズ (を融合したもの) とする。すると、

$$E[\sum_i (\hat{z}_i - \hat{z})^2] \quad (2.62)$$

$$= \sum_i E\left[\frac{1}{N} \{(N-1)(m_i + s_i) - \sum_{j \neq i} (m_j + s_j)\}^2\right] \quad (2.63)$$

$$= \sum_i \frac{1}{N^2} [(N-1)^2(M + S_i) + \sum_{j \neq i} (M + S_j)] \quad (2.64)$$

$$= \frac{1}{N^2} [(N-1)^2 NM + (N-1)^2 \sum_i S_i + (N-1)NM + (N-1) \sum_i S_i] \quad (2.65)$$

$$= \frac{N-1}{N^2} [N^2 M + N \sum_i S_i] \quad (2.66)$$

$$= (N-1)M + \frac{N-1}{N} \sum_i S_i \quad (2.67)$$

となる。よって M の推定値は

$$M = \frac{\sum_{i=1}^N (\hat{z}_i - \hat{z})(\hat{z}_i - \hat{z})^T}{N-1} - \frac{\sum_{i=1}^N S_i}{N} \quad (2.68)$$

とできる。これにより、観測値から M を推定することができる。

別の \hat{z} の取り方をして、 z の推定値を

$$\hat{z} = \left(\sum_j S_j^{-1}\right)^{-1} \sum_j S_j^{-1} \hat{z}_j \quad (2.69)$$

とする場合について考えておく。このほうが z の推定値としては分散が考慮されている分正確であり、この式はベイズ推定から導かれる。しかしこのように定義すると、以下に

示すように M を求める式が複雑になってしまうので、 M の推定には向いていないことがわかる。

次の式では式の表示を簡略化するため、ベクトル a の 2 乗は aa^T を表すことにする。 S_i, M は対象行列であることに注意して変形すると

$$E[\sum_i (z_i - \bar{z})^2] \quad (2.70)$$

$$= E[\sum_i (z_i - (\sum_j S_j^{-1})^{-1} \sum_j S_j^{-1} z_j)^2] \quad (2.71)$$

$$= E[\sum_i (z + m_i + s_i - (\sum_j S_j^{-1})^{-1} \sum_j S_j^{-1} (z + m_j + s_j))^2] \quad (2.72)$$

$$= \sum_i E[\{(\sum_j S_j^{-1})^{-1} \{(\sum_{j \neq i} S_j^{-1})(m_i + s_i) - \sum_{j \neq i} S_j^{-1} (m_j + s_j)\}\}^2] \quad (2.73)$$

$$= \sum_i [(\sum_j S_j^{-1})^{-1} \cdot \{(\sum_{j \neq i} S_j^{-1})(M + S_i)(\sum_{j \neq i} S_j^{-1}) + \sum_{j \neq i} S_j^{-1}(M + S_j)S_j^{-1}\}(\sum_j S_j^{-1})^{-1}] \quad (2.74)$$

$$= \sum_i (\sum_j S_j^{-1})^{-1} \cdot \{\sum_{j \neq i} S_j^{-1} M \sum_{k \neq i} S_k^{-1} + \sum_{j \neq i} S_j^{-1} M S_j^{-1} + \sum_{j \neq i} S_j^{-1} S_i \sum_{k \neq i} S_k^{-1} + \sum_{j \neq i} S_j^{-1}\} (\sum_j S_j^{-1})^{-1} \quad (2.75)$$

となる。ここで

$$\sum_i \sum_{j \neq i} A_j = (N-1) \sum_i A_i \quad (2.76)$$

$$\sum_i (\sum_{j \neq i} S_j^{-1} M \sum_{k \neq i} S_k^{-1}) = \sum_j S_j^{-1} \sum_i M \sum_k S_k^{-1} - \sum_i (S_i^{-1} M S_i^{-1} + S_i^{-1} M \sum_{j \neq i} S_j^{-1} + \sum_{j \neq i} S_j^{-1} M S_i^{-1}) \quad (2.77)$$

$$= N \sum_j S_j^{-1} M \sum_k S_k^{-1} - \sum_i S_i^{-1} M S_i^{-1} - \sum_i \sum_{j \neq i} (S_i^{-1} M S_j^{-1} + S_j^{-1} M S_i^{-1}) \quad (2.78)$$

$$= (N-2) \sum_j S_j^{-1} M \sum_k S_k^{-1} + \sum_i S_i^{-1} M S_i^{-1} \quad (2.79)$$

$$\sum_i (\sum_{j \neq i} S_j^{-1} S_i \sum_{k \neq i} S_k^{-1}) = \sum_j S_j^{-1} \sum_i S_i \sum_k S_k^{-1} - \sum_i S_i^{-1} - 2(N-1) \sum_i S_i^{-1} \quad (2.80)$$

$$= \sum_j S_j^{-1} \sum_i S_i \sum_k S_k^{-1} - (2N-1) \sum_i S_i^{-1} \quad (2.81)$$

となることを使うと

$$E[\sum_i (z_i - \bar{z})^2] \quad (2.82)$$

$$= (\sum_j S_j^{-1})^{-1} \{ (N-2) \sum_j S_j^{-1} M \sum_k S_k^{-1} + \sum_i S_i^{-1} M S_i^{-1} + (N-1) \sum_i S_i^{-1} M S_i^{-1} + \sum_j S_j^{-1} \sum_i S_i \sum_k S_k^{-1} - N \sum_i S_i^{-1} \} (\sum_j S_j^{-1})^{-1} \quad (2.83)$$

$$= (N-2)M + N \left(\sum_j S_j^{-1} \right)^{-1} \sum_i S_i^{-1} M S_i^{-1} \left(\sum_j S_j^{-1} \right)^{-1} + \sum_i S_i - N \left(\sum_j S_j^{-1} \right)^{-1} \quad (2.84)$$

M の推定値を得ることを考える。

$$\sum_i (z_i - \bar{z})^2 \simeq (N-2)M + N \left(\sum_j S_j^{-1} \right)^{-1} \sum_i S_i^{-1} M S_i^{-1} \left(\sum_j S_j^{-1} \right)^{-1} + \sum_i S_i - N \left(\sum_j S_j^{-1} \right)^{-1} \quad (2.85)$$

より

$$(N-2)M + N \left(\sum_j S_j^{-1} \right)^{-1} \sum_i S_i^{-1} M S_i^{-1} \left(\sum_j S_j^{-1} \right)^{-1} = \sum_i (z_i - \bar{z})^2 - \sum_i S_i + N \left(\sum_j S_j^{-1} \right)^{-1} \quad (2.86)$$

となる。これが直接解ければ良いが、現在のところ、解法はわかっていない。 M の要素をばらして、ベクトルに並び変えて解くという方法があるが計算は大変である。よって、(2.69) のように推定値を求める方法は、 M を求めるのには向いていないことがわかる。

第3章

センサ位置選択の一般論

3.1 はじめに

本章ではアクティブセンシングの一つの形態として、センサの位置を変えることによって位置情報の計測をより効率的に行なう方法を考える [向井, 石川, 1994] [Mukai and Ishikawa, 1994]。何が良いセンサ位置かを表す基準を、ここでは次のように2つ取る。一つは、推定値の分散を小さくすることであり、第3.2節、第3.3節で扱う。もう一つは測定値の連合を行ないやすくすることであり、第3.4節で扱う。

さらに、これらの基準の統合、推定値の求め方を第3.5節で、更に計算アルゴリズムをその次の第3.6節で扱う。最後に簡単なシミュレーション結果をあげる。

なお、本章の方法の定式化はカルマンフィルタに対して行なうが、実際には推定値の分散が計算できる方法ならば何でも良い。つまり、第2章で提案した観測モデルがオフセットを持つ場合の推定方法についても、同様の議論ができる。

3.2 誤差を小さくするセンサ位置選択の方法

センサの位置を変えることにより、時間変化する対象の状態をより良い精度で観測する問題を考える。例としては、ハンドアイカメラによる動き回る対象のトラッキングがあげられる。簡単のためにまず、対象とセンサがともに1つである場合を考える。

観測対象の状態変数 \mathbf{x} の状態方程式と、観測値 \mathbf{z} が、

$$\mathbf{x}(t+1) = F(t)\mathbf{x}(t) + G(t)\mathbf{w}(t) \quad (3.1)$$

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t), t) + \mathbf{v}(t) \quad (3.2)$$

と記述されるとする。ただし、 t は時刻を表し、 F, G は行列、 $\mathbf{h}(\cdot)$ は非線形関数である。また、 $\mathbf{w}(t), \mathbf{v}(t)$ は平均が0で互いに独立なノイズであり、その共分散行列を $Q(t), R(t)$ とする。観測値 \mathbf{z} から状態変数 \mathbf{x} を推定するために、カルマンフィルタを用いる。カルマンフィルタの特徴としては、

1. 逐次的に推定ができる、

2. 1 時刻先の予測値がフィルタの計算に組み込まれている,
3. 推定値の共分散行列が得られる,
4. 全てのノイズがガウス過程なら最小分散推定値が得られ, そうでなくても線形最小分散推定値が得られる,

などが上げられる。ここでは非線形なシステムなので, 特に拡張カルマンフィルタを用いる。具体的には, 時刻 t において状態方程式 (3.1) を用いた時刻 $t+1$ の状態 \mathbf{x} およびその共分散行列 P の予測は,

$$\hat{\mathbf{x}}(t+1|t) = F(t)\hat{\mathbf{x}}(t|t) \quad (3.3)$$

$$P(t+1|t) = F(t)P(t|t)F^T(t) + G(t)Q(t)G^T(t) \quad (3.4)$$

と書け, 時刻 $t+1$ における観測後の推定として,

$$\begin{aligned} \hat{\mathbf{x}}(t+1|t+1) &= \hat{\mathbf{x}}(t+1|t) \\ &\quad + W(t+1)[\mathbf{z}(t+1) - \mathbf{h}(\hat{\mathbf{x}}(t+1|t), t)] \end{aligned} \quad (3.5)$$

$$\begin{aligned} P^{-1}(t+1|t+1) &= P^{-1}(t+1|t) \\ &\quad + \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)} \right)^T R^{-1}(t+1) \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)} \end{aligned} \quad (3.6)$$

が得られる。ただし,

$$W(t+1) = P(t+1|t+1) \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)} \right)^T R^{-1}(t+1) \quad (3.7)$$

である。また, $\hat{\mathbf{x}}(t|t)$ は $\mathbf{x}(t)$ の時刻 t における推定値であり, $\hat{\mathbf{x}}(t+1|t)$ は時刻 t までの観測値と (3.1) を使った $\mathbf{x}(t+1)$ の予測値である。 $\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)}$ は $\mathbf{h}(\mathbf{x}, t)$ の \mathbf{x} による $\hat{\mathbf{x}}(t+1|t)$ でのヤコビアンを表す。また, $P(t|t)$ は誤差の共分散行列に相当する。つまり,

$$P(t|t) = E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t))(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t))^T] \quad (3.8)$$

である。ただし $E[\cdot]$ は期待値を表す。

ここで, カルマンフィルタを使った状態推定が効率良く行なえるように, センサを移動することを考える。本方法の概念を図 3.11 で説明する。時刻 t に対象が楕円で示される領域の中にあると推定されるとする。この領域は状態方程式 (3.2) により, 時刻 $t+1$ には新たな楕円の領域に移動すると予測される。さらに, 観測の結果新たな推定値が得られ, 推定値が存在すると考えられる領域はその前に予測された楕円より小さくなる。位置 1 にあるカメラと位置 2 にあるカメラで観測を行なった場合の領域が図のようになると仮定する。そのような場合, 領域の体積を比較すると位置 2 から得られるものの方が小さいので, 位置 1 より位置 2 の方がより良い位置であるとする。

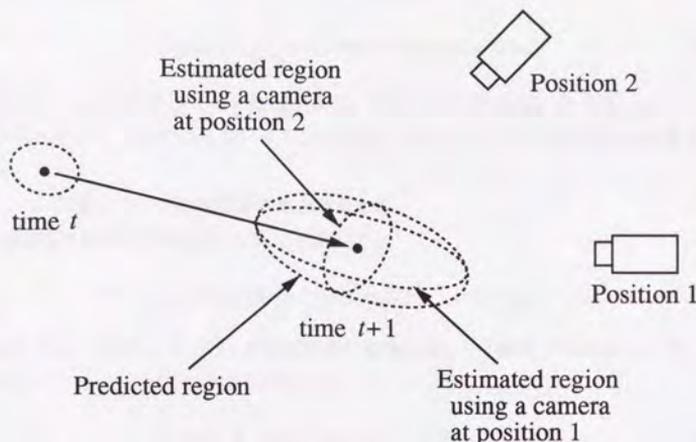


図 3.1 Target tracking

以上で説明した概念を定式化する。センサの位置を移動するというは、 $h(\cdot)$ を変化させることを意味する。そこで、 $h(\cdot)$ や P などがセンサの位置に依存することを明示したい場合には、 $h(\cdot; \mathbf{x}_s)$, $P(\cdot; \mathbf{x}_s)$ のように書く。ただし、 \mathbf{x}_s で (複数個の) センサ位置を表す。ここでは望ましいセンサ位置を $\hat{\mathbf{x}}$ の分散が小さくなる位置と考える。つまり、時刻 $t+1$ で観測を行なった時の $P(t+1|t+1; \mathbf{x}_s)$ を予測して、

$$\mathcal{R}(\mathbf{x}_s) = \det P(t+1|t+1; \mathbf{x}_s) \quad (3.9)$$

を計算し、これが小さくなるような位置 \mathbf{x}_s を選ぶ。 $P(t+1|t+1; \mathbf{x}_s)$ は (3.6) からわかるように、時刻 $t+1$ の観測値 $\mathbf{z}(t+1)$ を得なくても計算できる。また、 P は正定値なので、 $\mathcal{R}(\mathbf{x}_s)$ は常に正となる。 $\mathcal{R}(\mathbf{x}_s)$ は、 $P(t+1|t+1)$ の全ての固有値の積となるので、 $\hat{\mathbf{x}}$ の分散に相当すると考えられる。 $\mathcal{R}(\mathbf{x}_s)$ を小さくするというは、状態変化のモデルによる予測と、それぞれのセンサの観測値を融合した結果、一番良いデータが得られるセンサ位置を選択することを意味する。なお、(3.9) のように推定値の共分散行列の行列式を評価関数として用いる方法は、実験計画法の分野では D 最適 [Dodge et al., 1988] として知られていて、共分散行列の行列式は一般化分散と呼ばれている。

しかし、定義した $\mathcal{R}(\mathbf{x}_s)$ を直接計算するのは大変なので、ここでは位置 \mathbf{x}_s を離散化し、有限個の候補の中から最適な位置を選ぶことにする。また、 $\mathcal{R}(\mathbf{x}_s)$ が小さくなる \mathbf{x}_s を

求めるということは,

$$\mathcal{E}^P(\mathbf{x}_s) \equiv \frac{1}{\mathcal{R}(\mathbf{x}_s)} = \det P^{-1}(t+1|t+1; \mathbf{x}_s) \quad (3.10)$$

が大きくなる \mathbf{x}_s を求めることと等価である。このような評価関数 \mathcal{E}^P を用いると, (3.6) から得られる P^{-1} をそのまま用いることができるため, $P^{-1} \rightarrow P$ の逆行列の計算が不要となる。

ここで, 対象とセンサが複数ある場合を考える。

観測対象 i の状態方程式が (3.1) に対応して,

$$\mathbf{x}_i(t+1) = F_i(t)\mathbf{x}_i(t) + G_i(t)\mathbf{w}_i(t) \quad (i = 1, \dots, N) \quad (3.11)$$

であるとする。ただし, 添え字 i が観測対象の番号を表し, 全体で N 個あるとする。この対象のセンサ j への出力が (3.2) に対応して

$$\mathbf{z}_{j;i}(t) = \mathbf{h}_j(\mathbf{x}_i(t), t) + \mathbf{v}_{j;i}(t) \quad (j = 1, \dots, M) \quad (3.12)$$

と表されるとする。ただし, M はセンサの数で, $\mathbf{z}_{j;i}$ は得られない場合もある。また, $\mathbf{v}_{j;i}(t)$ は平均 0 で, $\mathbf{w}_l(l = 1, \dots, N)$, $\mathbf{v}_{m;i}(m = 1, \dots, M; m \neq j)$ と独立なノイズで, その共分散行列が $R_{j;i}(t)$ であるとする。

まず, 複数の対象に対しては, それぞれの対象に対して独立に

$$\mathcal{E}_i^P(\mathbf{x}_s) \equiv \frac{1}{\mathcal{R}_i(\mathbf{x}_s)} = \det P_i^{-1}(t+1|t+1; \mathbf{x}_s) \quad (3.13)$$

を求める。第 3.5 節でこれを統合する。

また, 複数のセンサはまとめて大きな次元の出力ベクトルを持つセンサが一つあると考える。つまり, (3.2) の \mathbf{z} , $\mathbf{h}(\cdot)$ やセンサ位置 \mathbf{x}_s を

$$\begin{cases} \mathbf{x}_s = [\mathbf{x}_{s1}^T, \mathbf{x}_{s2}^T, \dots, \mathbf{x}_{sM}^T]^T \\ \mathbf{h} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_M^T]^T \\ \mathbf{z} = [\mathbf{z}_1^T, \mathbf{z}_2^T, \dots, \mathbf{z}_M^T]^T \end{cases} \quad (3.14)$$

と書く。観測ノイズの共分散行列は, $\mathbf{v}_{j;i}$ の定義に従って求められ, 特に, センサのノイズには相互に相関がないという仮定がある場合には

$$R = \text{diag}(R_1, R_2, \dots, R_M) \quad (3.15)$$

のように書ける。これらの式に対して, センサ一つについて求めたセンサ位置選択の方法がそのまま適用できる。

3.3 アクチュエータの動きの誤差を含む場合

一般にセンサを動かす場合には、アクチュエータの動きの誤差によって、センサ位置にも誤差が生じる。特に、自律移動車が移動して画像などを取る場合、センサ位置の誤差は大きくなると考えられる。この場合に対応するようにセンサ選択の方法を拡張する。

センサ位置 \mathbf{x}_s に対する、センサを動かしたときのアクチュエータによる誤差を $\mathbf{v}_a(\mathbf{x}_s)$ とし、その分散を $R_a(\mathbf{x}_s)$ とする。ただし、 $\mathbf{v}_a(\mathbf{x}_s)$ は他のノイズと無相関であるとする。つまり、

$$\begin{aligned} E[\mathbf{v}_a(\mathbf{x}_{si})\mathbf{v}_a(\mathbf{x}_{sj})^T] &= \delta_{ij}R_a(\mathbf{x}_{si}), \\ E[\mathbf{v}_a(\mathbf{x}_s)\mathbf{v}(t)] &= 0, \quad E[\mathbf{v}_a(\mathbf{x}_s)\mathbf{w}(t)] = 0 \end{aligned} \quad (3.16)$$

となる。この場合、観測を表す式 (3.2) は

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{x}_s) + \mathbf{v}(t) + \frac{\partial \mathbf{h}}{\partial \mathbf{x}_s} \mathbf{v}_a(\mathbf{x}_{sj}) \quad (3.17)$$

と書き換えられる。つまり、(3.2) の観測の式に関して

$$\begin{aligned} \mathbf{v}(t) &\rightarrow \mathbf{v}(t) + \frac{\partial \mathbf{h}}{\partial \mathbf{x}_s} \mathbf{v}_a(\mathbf{x}_{sj}) \\ R(t) &\rightarrow R(t) + \frac{\partial \mathbf{h}}{\partial \mathbf{x}_s} R_a(\mathbf{x}_{si}) \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}_s} \right)^T \end{aligned} \quad (3.18)$$

のように置き換える。この置きかえををするだけで、残りの部分は以前と同じ式が使える。

3.4 連合を良く行なえるセンサ位置選択の方法

複数センサから、複数の対象に関する情報が入ってくる場合、異なるセンサ間でどのデータが同じ対象から得られたデータかの連合 (association) をする必要がある。

センサ j から時刻 t に L 個の対象物に関する出力が得られたとして、それを

$$\tilde{\mathbf{z}}_{j;k}(t) \quad (k = 1, \dots, L) \quad (3.19)$$

とする。どの対象物からの出力かを判別するためには色々な手法が考えられるが、ここでは単純に Nearest Data Association Algorithm [Blake and Yuille, 1992] を使うことにする。これは

$$\tilde{\mathbf{z}}_{j;i}(t+1|t) = \mathbf{h}_j(\hat{\mathbf{x}}_i(t+1|t)) \quad (3.20)$$

として、マハラノビス距離 (Mahalanobis generalized distance)

$$\begin{aligned} \mu_{j;ik} &\equiv (\tilde{\mathbf{z}}_{j;i}(t+1|t) - \tilde{\mathbf{z}}_{j;k}(t+1|t))^T S_{j;i}(t+1|t)^{-1} \\ &\quad \cdot (\tilde{\mathbf{z}}_{j;i}(t+1|t) - \tilde{\mathbf{z}}_{j;k}(t+1|t)) \end{aligned} \quad (3.21)$$

を最小にする $\hat{z}_{j,k}(t+1)$ を $z_{j,i}(t+1)$ とする方法である。ただし、 $S_{j,i}(t+1|t)$ は $\hat{z}_{j,i}(t+1|t)$ の共分散行列で、対象 i に対する共分散行列を P_i とし、

$$S_{j,i}(t+1|t) = \frac{\partial h_j}{\partial x} |_{x_i(t+1|t)} P_i(t+1|t) \left(\frac{\partial h_j}{\partial x} |_{x_i(t+1|t)} \right)^T \quad (3.22)$$

と計算できる。また、適当に定めた閾値より $\mu_{j,ik}$ が大きくなったら $z_{j,i}(t+1)$ に該当する出力はなしとする。このようにして求めた $z_{j,i}(t+1)$ を j について並べたものを対象 i についての観測値 $z_i(t+1)$ とすると、その共分散行列 $R_i(t+1)$ は $R_{j,i}(t+1)$ を対角上に並べたものとなる。この $z_i(t+1)$ を用いて、(3.3)~(3.7) に添え字 i をつけた対象 i についてのカルマンフィルタが構成できる。

(3.21) を使って正しく連合できるためには、異なった対象から得られるセンサデータが離れている必要がある。このことを実現するための方法の概念を図 3.2 で説明する。この図で、カメラの画像からは対象が黒いか白いかわからないとする。すると、カメラ 2 の画像上では対象が離れているので連合がうまくいく可能性が高いが、カメラ 1 では対象が近付いているため、失敗する可能性が高い。そこで、カメラ 2 の位置の方が連合にとって良い位置であるとする。ただし、実際にはスクリーン上の対象の距離はマハラノビス距離で測られる。

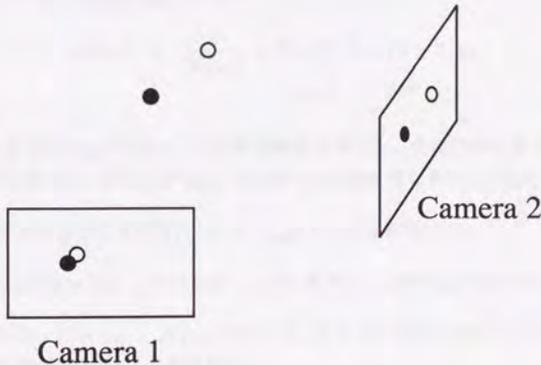


図 3.2 Target tracking of two objects

以上の概念を定式化する。

(3.21) は予測値とマハラノビス距離が一番近いデータをその対象からの出力として選ぶということを意味するので、これが有効に働くためには、任意の 2 つの対象の予測値のマハラノビス距離、

$$D_{j\hat{z}_{i\alpha}, i\beta} \equiv (\hat{z}_{j\hat{z}_{i\alpha}}(t+1|t) - \hat{z}_{j\hat{z}_{i\beta}}(t+1|t))^T (S_{j\hat{z}_{i\alpha}} + S_{j\hat{z}_{i\beta}})^{-1}$$

$$\begin{aligned} & \cdot (\hat{\mathbf{z}}_{j;i_\alpha}(t+1|t) - \hat{\mathbf{z}}_{j;i_\beta}(t+1|t)) \\ & (i_\alpha, i_\beta = 1, \dots, N; i_\alpha \neq i_\beta) \end{aligned} \quad (3.23)$$

ができるだけ大きくなるようなセンサ位置を選ぶことが必要である。ただし、対象 i_α, i_β がセンサのレンジからはずれている場合には、 $D_{j;i_\alpha,i_\beta}$ は得られないとする。ここでは、距離が近付くと急激に分離が悪くなることを考慮して、

$$\mathcal{E}^a(\mathbf{x}_s) \equiv - \sum_{j=1}^M \frac{1}{D_j} \quad (3.24)$$

を、センシングシステム全体での連合のしやすさを表す評価関数とする。ただし、

$$D_j \equiv \min_{i_\alpha, i_\beta} D_{j;i_\alpha, i_\beta} \quad (3.25)$$

である。また、センサ j に関して1つも $D_{j;i_\alpha, i_\beta}$ が得られない場合には、 D_j に充分小さな値を設定する。 $\mathcal{E}^a(\mathbf{x}_s)$ が大きいほど、対象の連合が行ないやすくなる。

3.5 評価関数の設定と融合の仕方

それぞれの対象 i の精度の良さを表す

$$\begin{aligned} \mathcal{E}_i^p(\mathbf{x}_s) & \equiv \frac{1}{\mathcal{R}_i(\mathbf{x}_s)} \equiv \det P_i^{-1}(t+1|t+1; \mathbf{x}_s) \\ & (i = 1, \dots, N) \end{aligned} \quad (3.26)$$

と連合のしやすさ $\mathcal{E}^a(\mathbf{x}_s)$ を大きくする評価関数を考える。そのためにまず、センサ、対象ともに複数ある場合に、 $\mathcal{E}_i^p(\mathbf{x}_s), \mathcal{E}^a(\mathbf{x}_s)$ を計算する手順を明らかにしておく。

1. 候補となるあるセンサ位置に対して $\hat{\mathbf{z}}_{j;i}(t+1|t)$ を計算する。
2. $\hat{\mathbf{z}}_{j;i}(t+1|t)$ ($i = 1, \dots, N; j = 1, \dots, M$) を使って、 $\mathcal{E}^a(\mathbf{x}_s)$ を求める。
3. それぞれの i ($i = 1, \dots, N$) について、 $\hat{\mathbf{z}}_{j;i}(t+1|t)$ が出力のレンジ内に入っているセンサを使って $\mathcal{E}_i^p(\mathbf{x}_s)$ を計算する。

システム全体の誤差を表す評価関数を

$$\mathcal{R}_{\text{All}}(\mathbf{x}_s) \equiv \sum_{i=1}^N \mathcal{R}_i(\mathbf{x}_s) \quad (3.27)$$

とすると、これを小さくするということは

$$\mathcal{E}_{\text{All}}^p(\mathbf{x}_s) \equiv \frac{1}{\mathcal{R}_{\text{All}}(\mathbf{x}_s)} = \frac{1}{\sum_{i=1}^N 1/\mathcal{E}_i^p(\mathbf{x}_s)} \quad (3.28)$$

を大きくすることに対応する。これと $\mathcal{E}^a(\mathbf{x}_s)$ を用いて、精度の良さと連合のしやすさをともに満たす評価関数

$$\mathcal{I}(\mathbf{x}_s) \equiv \mathcal{E}_{\text{All}}^p(\mathbf{x}_s) + \lambda \mathcal{E}^a(\mathbf{x}_s) \quad (3.29)$$

を定義する。ただし、 λ は適当な正数である。 $\mathcal{I}(\mathbf{x}_s)$ を最大にする位置を望ましいセンサ位置と定義する。 $\mathcal{E}_{\text{All}}^p$ を大きくする位置と \mathcal{E}^a を大きくする位置は必ずしも両立せず、その場合には、正しく連合できる中で、一番予測誤差の小さくなる位置を選ぶことが望ましい。 λ の値によって、正しい連合をすることと、データの融合結果の精度のどちらをどの程度重要視するかが決まる。しかし、 λ を適切に定めるためには $\mathcal{E}_{\text{All}}^p$ や \mathcal{E}^a の値を知る必要があるため、現在のところ、事前に λ を定めるための指針は得られていない。

また、分散したセンサからの情報を融合する方法は、(3.5) をそれぞれの対象について考えたものを変形して、

$$\begin{aligned} \hat{\mathbf{x}}_i(t+1|t+1) &= P_i(t+1|t+1)[P_i^{-1}(t+1|t)\hat{\mathbf{x}}_i(t+1|t) \\ &\quad + \sum_{j=1}^N \left(\frac{\partial \mathbf{h}_j}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}_i(t+1|t)} \right)^T R_{j;i}^{-1}(t+1) \{ \mathbf{z}_{j;i}(t+1) \\ &\quad - \mathbf{h}_j(\hat{\mathbf{x}}_i(t+1|t)) + \frac{\partial \mathbf{h}_j}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}_i(t+1|t)} \hat{\mathbf{x}}_i(t+1|t) \}] \end{aligned} \quad (3.30)$$

で与えられる。これは並列カルマンフィルタ (parallel Kalman filter) [Hashemipour, 1988] と同様の式である。

3.6 計算アルゴリズム

3.6.1 対象が1つの場合

まず簡単のため、対象が1つしかない場合 ($N=1$) を考える。対象が1つなので連合のしやすさ \mathcal{E}^a は考慮しなくて良い。対象が1つでもセンサが複数あって、 \mathbf{x}_s の次元が高い場合には計算が大変である。今、それぞれのセンサ j ($j=1, \dots, M$) に対応する $\mathbf{x}_{sj}, \mathbf{h}_j(\cdot), \mathbf{z}_j, R_j$ を考える (対象が1つしかないので添え字 i は省略)。センサのノイズには相互に相関がないとすると、

$$\begin{cases} \mathbf{x}_s = [\mathbf{x}_{s1}^T, \mathbf{x}_{s2}^T, \dots, \mathbf{x}_{sM}^T]^T \\ \frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \left[\frac{\partial \mathbf{h}_1}{\partial \mathbf{x}}, \frac{\partial \mathbf{h}_2}{\partial \mathbf{x}}, \dots, \frac{\partial \mathbf{h}_M}{\partial \mathbf{x}} \right]^T \\ \mathbf{z} = [\mathbf{z}_1^T, \mathbf{z}_2^T, \dots, \mathbf{z}_M^T]^T \\ R = \text{diag}(R_1, R_2, \dots, R_M) \end{cases} \quad (3.31)$$

と書ける。ここで、

$$\bar{P}_j^{-1}(t+1) \equiv \left(\frac{\partial \mathbf{h}_j}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)} \right)^T R_j^{-1}(t+1) \frac{\partial \mathbf{h}_j}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}(t+1|t)} \quad (3.32)$$

と定義すると, (3.6) は,

$$P^{-1}(t+1|t+1) = P^{-1}(t+1|t) + \sum_{j=1}^M \bar{P}_j^{-1}(t+1) \quad (3.33)$$

と書ける。とりあえずセンサを一つだけ考え,

$$P^{-1}(t+1|t+1; \mathbf{x}_{s1}) = P^{-1}(t+1|t) + \bar{P}_1^{-1}(t+1; \mathbf{x}_{s1}) \quad (3.34)$$

を使った評価関数

$$\mathcal{E}^P(\mathbf{x}_{s1}) \equiv \det P^{-1}(t+1|t+1; \mathbf{x}_{s1}) \quad (3.35)$$

からセンサ位置 \mathbf{x}_{s1} を求める。次に \mathbf{x}_{s1} が確定したとして,

$$P^{-1}(t+1|t+1; \mathbf{x}_{s1}, \mathbf{x}_{s2}) = P^{-1}(t+1|t) + \bar{P}_1^{-1}(t+1; \mathbf{x}_{s1}) + \bar{P}_2^{-1}(t+1; \mathbf{x}_{s2}) \quad (3.36)$$

を使った評価関数 $\mathcal{E}^P(\mathbf{x}_{s1}, \mathbf{x}_{s2})$ から, \mathbf{x}_{s2} を求める。これを繰り返すことによって, 近似的に良い位置 \mathbf{x}_s が求まる。これは, 貪欲算法と呼ばれる考え方である。

3.6.2 対象が複数の場合

次に, 対象が複数存在する場合を考える。この場合は連合のしやすさ \mathcal{E}^a を考慮する必要がある。実際の計算は対象が一つの場合と同様に, センサ位置を一つずつ決めていく。(3.33) を対象が複数あることを考慮して書き換えると,

$$P_i^{-1}(t+1|t+1) = P_i^{-1}(t+1|t) + \sum_{j=1}^M \bar{P}_{ij}^{-1}(t+1) \quad (3.37)$$

となる。そこで, まずセンサを一つだけを考え,

$$\mathcal{I}(\mathbf{x}_{s1}) = \left(\sum_{i=1}^N \frac{1}{\det P_i^{-1}(t+1|t+1; \mathbf{x}_{s1})} \right)^{-1} - \lambda \frac{1}{D_1(\mathbf{x}_{s1})} \quad (3.38)$$

を使って, 望ましいセンサ位置 \mathbf{x}_{s1} を求める。ただし

$$P_i^{-1}(t+1|t+1; \mathbf{x}_{s1}) = P_i^{-1}(t+1|t) + \bar{P}_{i1}^{-1}(t+1; \mathbf{x}_{s1}) \quad (3.39)$$

である。次に, \mathbf{x}_{s1} は固定して,

$$\begin{aligned} \mathcal{I}(\mathbf{x}_{s1}, \mathbf{x}_{s2}) &= \left(\sum_{i=1}^N \frac{1}{\det P_i^{-1}(t+1|t+1; \mathbf{x}_{s1}, \mathbf{x}_{s2})} \right)^{-1} \\ &\quad - \lambda \left(\frac{1}{D_1(\mathbf{x}_{s1})} + \frac{1}{D_2(\mathbf{x}_{s2})} \right) \end{aligned} \quad (3.40)$$

を使って \mathbf{x}_{s2} を求める。ただし,

$$\begin{aligned} P_i^{-1}(t+1|t+1; \mathbf{x}_{s1}, \mathbf{x}_{s2}) &= P_i^{-1}(t+1|t) \\ &\quad + \bar{P}_{i1}^{-1}(t+1; \mathbf{x}_{s1}) + \bar{P}_{i2}^{-1}(t+1; \mathbf{x}_{s2}) \end{aligned} \quad (3.41)$$

である。これを繰り返すことによって, 近似的に良い位置 \mathbf{x}_s が求まる。

3.7 シミュレーション — ターゲットトラッキング

壁で囲まれた2次元平面上を動きまわる2つの対象を、2つのマニピュレータ(自由度3)の先につけたハンドアイカメラで協調して追跡するシミュレーションを行なった(図3.3)。対象が2つなので(3.29)で定義される評価関数 I を用いた。

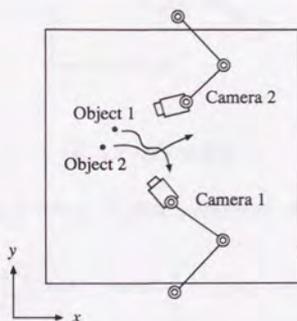


図 3.3 Hand-eye camera tracking system

対象 $i(i = 1, 2)$ の状態変数を、位置 (x_i, y_i) (壁の一边の長さが1となるように単位をとる) と速度を用いて $\mathbf{x}_i = [x_i, y_i, \dot{x}_i, \dot{y}_i]^T$ と定義し、(3.11)に対応する状態方程式を用いた。ただし $F_i(t)$ を

$$F_i = \begin{bmatrix} 1 & \Delta t & & \\ & 1 & \Delta t & \\ & & 0.98 & \\ & & & 0.98 \end{bmatrix} \quad (3.42)$$

とした。 Δt は時間を離散化した間隔である。また、対象は壁にぶつくと跳ね返るとした。(3.12)に対応する観測は、

$$\mathbf{h}_j(\mathbf{x}(t), t) + \mathbf{v}_{j;i}(t) \equiv -\frac{f\xi_{j;i}}{\eta_{j;i}} + [C_1 + C_2 \left\{ \frac{f\xi_{j;i}}{\eta_{j;i}} \right\}^2] \bar{\mathbf{v}} \quad (3.43)$$

と表した。ただし $\bar{\mathbf{v}}$ は一定の大きさの分散を持つノイズであり、 $(\xi_{j;i}, \eta_{j;i})$ はカメラ j のカメラ座標系での観測対象 i の位置座標である(図3.4)。また、 f は焦点距離、 C_1, C_2 は正の定数である。これにより、画像の中心から離れると大きなノイズが加わることを表す。なお、センサの後ろに対象がある場合には、出力は得られないとした。また、マニピュレータのそれぞれの関節は5度刻みに動き、一時刻に最大20度まで動けるとした。

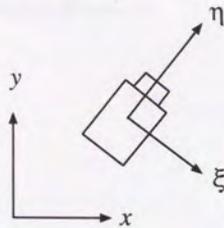


図 3.4 カメラ座標

カメラを固定した場合とアクティブに動かした場合の、実際の位置とその推定値との誤差

$$d(t) \equiv \sum_{i=1}^2 \{(x_i(t) - \hat{x}_i(t|t))^2 + (y_i(t) - \hat{y}_i(t|t))^2\}^{\frac{1}{2}} \quad (3.44)$$

の時間経過を図 3.5 に示す。

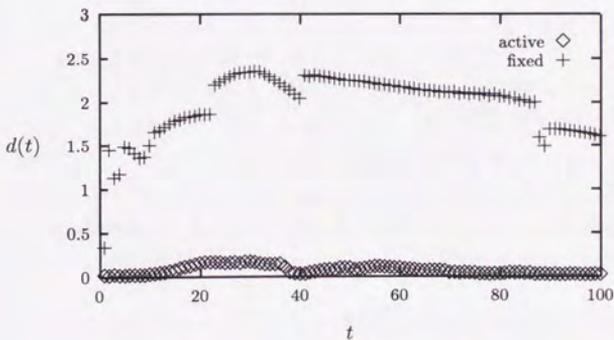


図 3.5 Results of simulation

本方法を使ってアクティブセンシングを行なうことにより、対象の動きに合わせてハンドアイカメラが望ましい位置に移動する。その結果、誤差 $d(t)$ がカメラを固定した場合に比べて小さな値に保たれていることがわかる。また、2つの対象が近距離にある時には2つのカメラとも両方の対象を視野に入れるのに対し、対象が離れている時にはそれぞれ

の対象を別々のカメラが追跡することが観察された。これにより、対象の状態に応じて適切なセンシング戦略が構成されることがわかった。

論文要約

観測カメラによる複数目標の追跡

1. はじめに

この論文では、大規模な環境の中で、複数のカメラで観測された複数の目標物の追跡問題を扱う。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。

この論文では、大規模な環境の中で、複数のカメラで観測された複数の目標物の追跡問題を扱う。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。

この論文では、大規模な環境の中で、複数のカメラで観測された複数の目標物の追跡問題を扱う。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。

この論文では、大規模な環境の中で、複数のカメラで観測された複数の目標物の追跡問題を扱う。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。目標物の位置と速度は、観測カメラからの観測データに基づいて推定される。

第 4 章

触覚センサによる接触位置の求め方

4.1 はじめに

CCD カメラは大域的な情報を得ることができるが、対象の3次元位置についての情報を得るためには、ステレオビジョンなどのそれなりの手順が必要である。また、それでも精度の良い位置情報を得ることは難しい。さらに、画像は照明によって大きな影響を受け、オクルージョンや影、光学的なひずみ、さらにはカメラのキャリブレーションの問題などが起こる。

一方、触覚センサは CCD カメラと対照的な特徴を持つ。触覚センサはセンサが接している対象の局所的な情報しか得られない。しかし、センサ自身の位置を利用することにより、センサが接している頂点やエッジなどの対象の3次元的位置を容易に得ることができる。更に、今回の論文では考慮していないが、対象のコンプライアンスや、熱特性など、視覚だけでは決して得られない情報を得ることも可能である。

触覚センサはその局所性のゆえに、大域的なセンサである視覚にガイドされることにより有効に利用することが可能となる。このように視覚と触覚を利用して、より効果的なセンシングを行なおうとする方法は視触覚融合と呼ばれて研究されてきた [Allen, 1988] [Stansfield, 1988] [Yamada et al., 1993] [坂根ら, 1994]。しかし、これらの研究では効率の良いセンシングを行なうためにはどのように触覚センサを動かせば良いかを求める方法は考えられていない。

そこで、本章では視触覚融合によって3次元対象の認識と、対象の位置・姿勢を効率良く得るための、触覚センサによる探索の仕方を求めることを目標とする [向井, 石川, 1994] [Mukai and Ishikawa, 1994]。まず第4.2節で、対象認識のために視触覚を使う方法を述べる。そして第4.3節、第4.4節で、第3章で述べたセンサ位置選択の方法を応用して、視覚から触覚センサで探索する位置を求める方法を提案する。そのために、CCD カメラでまず対象の大域的なおおまかな情報を取り、その情報を用いて、触覚センサで触る位置を求める。CCD カメラで撮った画像から対象の3次元位置を推定し、触覚センサで探索する位置を求めるために、この方法では対象の形状モデルを利用する。まず、第4.3節で対象の形

状モデルが誤差を含まない場合を説明し、次の第4.4節でこれを誤差を含む場合に拡張する。最後に第4.5節で、簡単なシミュレーションについて述べる。

4.2 対象認識

本節では複数の対象の形状モデルを持っている場合に、観測された対象がその中のどれに当たるかを認識し、それと同時に画像上の点やセグメント（モデル上の頂点やエッジに対応）がモデル上のどの部分に相当するかを求める方法を提案する。形状モデルは対象の3次元的な形状を情報として持つとし、そのモデルを構成する頂点やエッジの位置は誤差を含むとする。まず、CCDカメラの画像から頂点とエッジの情報を抽出し、形状モデルとのマッチングを行なうことにより対象認識を行なう。

このマッチングの方法として、文献[Gaston and Lozano-Perez, 1984]でGastonにより提案されている解釈木(IT: Interpretation Tree)を使う方法を参考にする。文献[Gaston and Lozano-Perez, 1984]では解釈木は触覚センサでの対象認識のために用いられているが、視覚にも適用可能であるようにこれを拡張する。マッチングを行なう際には、モデルやセンシングの誤差も考慮に入れてこれを行なう。そして、視覚ではマッチングが一意に定まらない場合に、必要な部分を触覚センサで探索することにより情報を補う。

まず、文献[Gaston and Lozano-Perez, 1984]で述べられている方法を簡単に説明する。 f 本の指を持つ手で n 個ある対象の中のどれか一つを把持したとする。それぞれの対象は O_j と表され、図4.1のように、 e_j 本のエッジを持つとする。 f 本の指はそれぞれ位置 P_i にあり、それはどれかはわからないが対象のエッジ上にあるとする。

ここで、 P_i がどの対象のどのエッジ上にあるかを判別するために、ITが使われる。図4.1は対象 O_j のITを表している。例えば P_1 levelは P_1 が e_j 本のエッジの中のどれに当たるかを表している。 P_2 level以下も同様である。異なる指が同一のエッジを触ることもあるため、それぞれのノードは e_j 本の子孫を持つ。ルートから P_f レベルにいたる経路により、一つの解釈が表される。何も条件がないと解釈は

$$\sum_{j=1}^n (e_j)^f \quad (4.1)$$

通り存在する。この中から、一つの解釈だけを選び出すことが目標である。

そのために、モデルの特徴とセンサからの入力を用いて枝刈りを行なう。例えば、2つの指の位置が、モデルから計算されるエッジの距離の最小値より小さければ、そのような解釈はあり得ない。このような拘束条件を用いて可能な解釈を減らし、最後に、モデルを幾何学的に回転、平行移動して入力データと一致するかを調べるという操作を行なう。この操作の計算量が多いので、それに至るまでに可能な限り枝刈りを行なうことにより、高速なマッチングが可能となる。

Gastonの方法を画像用に拡張する。まず画像からエッジ（画像上ではセグメントとなる）と、そのつながりのノードとなる頂点を抽出する。その際には図4.2のように、頂

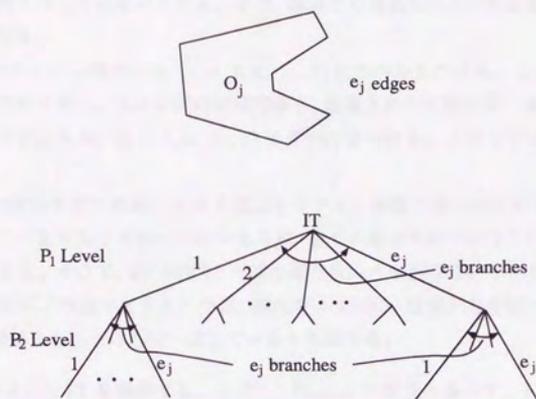


図 4.1 Interpretation tree by Gaston

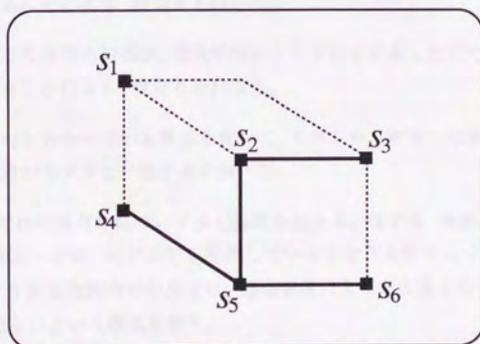


図 4.2 エッジ抽出

点やエッジは全部は抽出されなくても良いとする。この図では、実線が抽出されたエッジ、点が抽出された頂点で、点線が抽出に失敗した部分であるとする。何も無いところに誤ってエッジを検出することはないとする。また、画像上の頂点やエッジの位置は誤差を含むこともあるとする。

抽出されたエッジに適当に s_i ($i = 1, 2, \dots, I$) とラベルをつける。ここで、 I は抽出されたエッジの数である。エッジ間の接続関係も、画像上から可能な限り検出する。また、モデル上のエッジにも m_j ($j = 1, 2, \dots, J$) とラベルをつける。 J はモデルのエッジ数である。

基本的には幾何学的な移動により3頂点をモデルと画像で合わせてみて、残りの頂点とエッジが全て一致するかを調べるのであるが、全ての組合せについてこれを行なうと計算量が膨大になる。そこで、IT を作り、可能な限り枝刈りを行なう。これを考慮したマッチングの手順は以下の通りとする。なお、頂点やエッジは、位置が完全に一致していなくても、誤差が閾値より小さければ一致していると判断する。

1. 図 4.3 のような IT を製作する。ただし、Gaston の場合と違って、下のレベルでは、上のレベルで出てきた枝は使わない。
2. 画像上のエッジの接続関係から枝刈りを行なう。
3. 画像のノードから出ている分岐のなす角がモデルの分岐の角度より小さければマッチしないので、枝刈りを行なう。
4. モデル上で平行なエッジが画像上でも、透視変換によるずれを考慮した上で、平行かチェックすることにより、枝刈りを行なう。
5. 平行なエッジの長さの関係が、透視変換によるずれを考慮した上で、保たれているかチェックすることにより、枝刈りを行なう。
6. 確実に頂点だとわかっている数点を使って、モデルから画像への変換を求め、残りのエッジと頂点がモデルと一致するか調べる。

ここで、それぞれの枝刈りについて少し説明を加える。2では、画像上から接続しているとわかっているエッジが、モデルでも接続しているかどうかのチェックを行なう。3では、2直線のなす3次元空間内での角度は、透視変換によって大きくなることはあっても小さくなることはないという事実を使う。

4, 5では、もし透視変換ではなく正射影が行なわれるなら、平行なエッジは平行に投影され、長さの比は保たれるという性質を使う。透視投影では少し正射影の場合からずれるが、対象の距離がある程度以上で、対象の大きさがある範囲内ならば、ずれの上限が求まる。そこで、このずれと画像の誤差よりもずれが大きければ、候補からはずす。

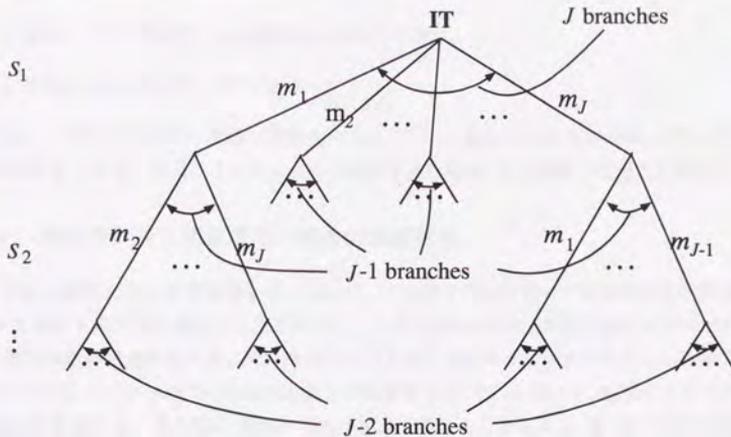


図 4.3 Interpretation tree for vision

6では、3頂点からカメラに対する対象の位置と姿勢を求める。この時の、残りの頂点とエッジの位置から、透視変換によって画像上の位置を計算し、画像から検出された頂点とエッジがこれらのどれかに一致するかを調べる。

この手続きを行なった後で、候補が複数出てきた場合、触覚センサにより、新たな情報を得る。触覚センサでは、面に対しては接触位置と方線方向の、エッジに対しては接触位置と方向ベクトルの情報が得られるとする。触覚センサは探索行動を行なうことにより、対象の位置が正確にわかっていなくても触りたい面やエッジに到達することが可能である。まず、触りやすい面について考える。次のように候補を選ぶ。

1. 上面、側面のように触覚センサで触れられて、カメラで検出されていない面で、ある程度以上の面積を持つものを選ぶ。
2. その中で、候補となるモデル間で、面の傾きまたは接触した時の位置が異なるものを選ぶ。

このようにして選ばれた面との接触を、一つ目の候補から順番に行なっていく。該当する面があったらその候補が確定となる。該当しない面が検出された場合、その情報で枝刈りを行なう。

面で選べなかった場合、エッジを触ることにより、認識を行なう。触るエッジは次の条件に当てはまるものを選ぶ。

- カメラで検出されていない。
- 触覚センサで触われ、ある程度以上の長さを持つ。
- 候補となる解釈間で一致しない。

これも、一つ目の候補から順番に接触を行なっていく。該当するエッジがあったら、その候補が確定となる。該当しないがエッジが検出された場合、その情報で枝刈りを行なう。

4.3 形状モデルに誤差がない場合の位置計測

対象の精度の良い位置情報を得るために、3次元モデルを使って視覚情報から触覚センサで探索する位置を求める方法を述べる。まずは簡単のため、対象の形状モデルが正確で、誤差がない場合を考える。センサはカメラ1台とマニピュレータの先についた触覚センサ1つで、まずカメラで対象のおおよその位置をとり、それを使って、触覚センサでどこを触るかを求める。その際に、対象の形状モデルを考えることにより、第3章の方法が使え

る。まず、回転を定義しておく。 $\phi = [\phi_x, \phi_y, \phi_z]^T$ で順番に x 軸回りに ϕ_x , y 軸回りに ϕ_y , z 軸回りに ϕ_z の回転を表す。その回転を表す行列は

$$R(\phi) = R_z(\phi_z)R_y(\phi_y)R_x(\phi_x) \quad (4.2)$$

$$R_z(\phi_z) = \begin{bmatrix} \cos \phi_z & -\sin \phi_z & 0 \\ \sin \phi_z & \cos \phi_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\phi_y) = \begin{bmatrix} \cos \phi_y & 0 & \sin \phi_y \\ 0 & 1 & 0 \\ -\sin \phi_y & 0 & \cos \phi_y \end{bmatrix}$$

$$R_x(\phi_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_x & -\sin \phi_x \\ 0 & \sin \phi_x & \cos \phi_x \end{bmatrix}$$

となる。

対象の3次元モデルを、対象の代表点Oを原点とするモデル座標系で記述される点と線分の集合で表す。対象の位置を、代表点Oの世界座標系での位置 $\mathbf{X}_O = [X_O, Y_O, Z_O]^T$ と、対象のO回りの回転 $\phi_O = [\phi_{Ox}, \phi_{Oy}, \phi_{Oz}]^T$ で表す。ただし簡単のため、 \mathbf{X}_O, ϕ_O は時

間変化しないとする。モデル座標系上の座標値 \mathbf{x}_M に対応するワールド座標系上の座標値 \mathbf{x}_W は

$$\mathbf{x}_W(\mathbf{X}_O, \phi_O; \mathbf{x}_M) = R(\phi_O)\mathbf{x}_M + \mathbf{X}_O \quad (4.3)$$

となる。

まず、カメラでの観測を定義する。簡単のため、カメラ中心は常に原点にあつて、光軸は z 軸方向を向いているとする。このときの観測のモデルとしてピンホールモデルを使う。第2章のフィルタの定式化にしたがつて、観測を

$$\mathbf{x}_c = \mathbf{h}_c(\mathbf{x}_W) + \mathbf{m}_c + \mathbf{s}_c = \begin{bmatrix} \frac{x_W f}{z_W} \\ \frac{y_W f}{z_W} \\ z_W \end{bmatrix} + \mathbf{m}_c + \mathbf{s}_c \quad (4.4)$$

と定義する。 \mathbf{x}_c が画像面上の座標で、 f は焦点距離である。また、 \mathbf{m}_c がモデルの曖昧さ、 \mathbf{s}_c が測定時ノイズを表し、それぞれの共分散行列が M_c, S_c であるとする。 \mathbf{x}_W に (4.3) を代入して、

$$\begin{aligned} \mathbf{x}_c &= \mathbf{h}_c(R(\phi_O)\mathbf{x}_M + \mathbf{X}_O) + \mathbf{m}_c + \mathbf{s}_c \\ &= \tilde{\mathbf{h}}_c(\mathbf{X}_O, \phi_O; \mathbf{x}_M) + \mathbf{m}_c + \mathbf{s}_c \end{aligned} \quad (4.5)$$

となる (図 4.4)。

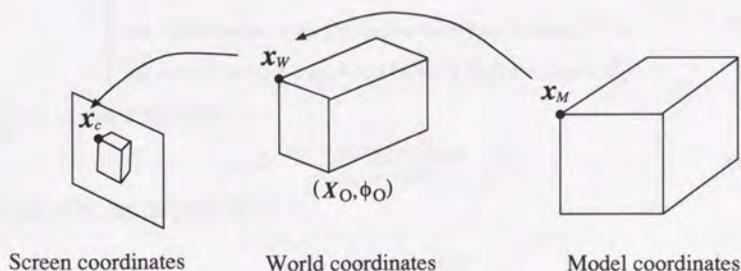


図 4.4 位置と姿勢の計算

カメラを使った計測では、画像から特徴点の検出をして、それがモデルのどの点に対応するか求め、それらを用いて状態変数 \mathbf{X}_O, ϕ_O を求めるとする。ここでは、画像上の点の座標 \mathbf{x}_c に対応するモデルの点の座標 \mathbf{x}_M は前節の方法でわかるものとし、 \mathbf{X}_O, ϕ_O を求めるために第2章のフィルタ (2.33)~(2.38) を使う。(4.5) が (2.3) に対応する観測の式であり、 \mathbf{X}_O, ϕ_O が (2.3) の状態変数 \mathbf{x} に対応する求めるべき変数である。

次に、マニピュレータの先についた触覚センサを考える。触覚センサでは、センサに接触している観測対象の世界座標系での頂点の座標やエッジのパラメータが検出できるとする。まず、頂点の観測について説明する。モデル座標系で座標値 \mathbf{x}_M を持つ頂点の世界座標が \mathbf{x}_W であるとする、それを触覚センサで検出する式をカメラと同様に

$$\mathbf{x}_T = \mathbf{x}_W(\mathbf{X}_O, \phi_O; \mathbf{x}_M) + \mathbf{m}_t + \mathbf{s}_t \quad (4.6)$$

と定義する。ただし、 $\mathbf{m}_t, \mathbf{s}_t$ はモデルの曖昧さと測定時ノイズで、これらの共分散行列が M_t, S_t であるとする。この式が第2章のフィルタの観測の式となる。 \mathbf{X}_O, ϕ_O が求めるべき変数である。

次に、触覚センサがエッジの一部に接触している場合を考える。この際には、センサ出力からエッジを表す式のパラメータが求まり、モデル上の対応するエッジも前節の方法でわかるとする。この関係から、対象の位置と方向 \mathbf{X}_O, ϕ_O を求める方法を考える。

まず、モデル座標系でのエッジ (z 軸に垂直でない) を

$$\begin{cases} x_M = a_M z_M + p_M \\ y_M = b_M z_M + q_M \end{cases} \quad (4.7)$$

と表す。このエッジの世界座標系での表現を求める。(4.7)のエッジ上の点を \mathbf{X}_O, ϕ_O で変換すると、 $R(\phi_O) = (r_{ij})$ において

$$\begin{cases} x_W = r_{11}(a_M z_M + p_M) + r_{12}(b_M z_M + q_M) + r_{13}z_M + X_O \\ y_W = r_{21}(a_M z_M + p_M) + r_{22}(b_M z_M + q_M) + r_{23}z_M + Y_O \\ z_W = r_{31}(a_M z_M + p_M) + r_{32}(b_M z_M + q_M) + r_{33}z_M + Z_O \end{cases} \quad (4.8)$$

となる。この第3式を変形した

$$z_M = \frac{z_W - r_{31}p_M - r_{32}q_M - Z_O}{r_{31}a_M + r_{32}b_M + r_{33}} \quad (4.9)$$

を(4.8)の第1式、第2式に代入して、

$$\begin{cases} x_W = a_W z_W + p_W \\ y_W = b_W z_W + q_W \end{cases} \quad (4.10)$$

となる。これが、(4.7)のエッジを \mathbf{X}_O, ϕ_O で変換した式である。ただし

$$\begin{cases} a_W = \frac{r_{11}a_M + r_{12}b_M + r_{13}}{r_{31}a_M + r_{32}b_M + r_{33}} \\ b_W = \frac{r_{21}a_M + r_{22}b_M + r_{23}}{r_{31}a_M + r_{32}b_M + r_{33}} \\ p_W = r_{11}p_M + r_{12}q_M + X_O - \frac{r_{31}p_M + r_{32}q_M + Z_O}{r_{31}a_M + r_{32}b_M + r_{33}}(r_{11}a_M + r_{12}b_M + r_{13}) \\ q_W = r_{21}p_M + r_{22}q_M + Y_O - \frac{r_{31}p_M + r_{32}q_M + Z_O}{r_{31}a_M + r_{32}b_M + r_{33}}(r_{21}a_M + r_{22}b_M + r_{23}) \end{cases} \quad (4.11)$$

である。この関係を $\mathbf{p}_W = [a_W, b_W, p_W, q_W]^T$, $\mathbf{p}_M = [a_M, b_M, p_M, q_M]^T$ として、

$$\mathbf{p}_W = \boldsymbol{\xi}(\mathbf{X}_O, \phi_O; \mathbf{p}_M) \quad (4.12)$$

と表す。

対象のエッジの世界座標系でのパラメータ値を触覚センサで検出できるとして、その値を $\mathbf{p}_T = [a_T, b_T, p_T, q_T]^T$ とする。この測定値を求める際にも、点の分布からエッジのパラメータを求めるためにカルマンフィルタを使えるが、ここでは複雑になるので省略する。観測を表す式を

$$\mathbf{p}_T = \mathbf{p}_W + \mathbf{m}_T + \mathbf{s}_T = \boldsymbol{\xi}(\mathbf{X}_O, \phi_O; \mathbf{p}_M) + \mathbf{m}_T + \mathbf{s}_T \quad (4.13)$$

とする。ただし、これまでと同様に $\mathbf{m}_T, \mathbf{s}_T$ はモデルの曖昧さと測定時ノイズであり、その共分散行列を M_T, S_T とする。また、この式のモデルの曖昧さ \mathbf{m}_T は、(4.6) のモデルの曖昧さ m_t と独立であるとする。つまり

$$\bar{E}[\mathbf{m}_t \mathbf{m}_T^T] = 0 \quad (4.14)$$

とする。ただし、 $\bar{E}[\cdot]$ はモデルを作る時の可能性で期待値をとることを表す。このことは、触覚センサで点を検出する時と、エッジを検出する時で独立なセンサを使うこと、または、同じセンサを使っても、検出の仕組みが大きく異なることを意味する。

これが、エッジを検出する際の観測の式であり、 \mathbf{X}_O, ϕ_O が求めるべき変数である。

以上の説明で、モデルを使って対象の位置と方向を求めるために第2章のフィルタが使えることがわかる。(4.5), (4.6), (4.13) が観測の式である。(4.6), (4.13) は、どの座標値 \mathbf{x}_M を持つ頂点や、パラメータ \mathbf{p}_M を持つエッジを触るかによって式が変わる。これに、第2章のセンサ位置を選択する方法を使うと、カメラで対象のおおよその位置を求めてから、次に触覚センサでどこを触ったら良いかが求められることがわかる。

具体的には、 $\mathbf{Y}_O = [\mathbf{X}_O^T, \phi_O^T]^T$, \mathbf{Y}_O の共分散行列を $P_{\mathbf{Y}_O}$ とし、まず(4.5)に第2章のフィルタを適用する。すると、

$$\hat{\mathbf{Y}}_O(t+1) = \hat{\mathbf{Y}}_O(t) + W_c(t+1)[\mathbf{x}_c(t+1) - \tilde{\mathbf{h}}_c(\hat{\mathbf{Y}}_O(t); \mathbf{x}_M(t+1))] \quad (4.15)$$

$$\text{Var}[\hat{\mathbf{Y}}_O(t+1)] \equiv P_{\mathbf{Y}_O}(t+1) = P_S(t+1) + P_M(t+1) \quad (4.16)$$

$$P_S^{-1}(t+1) = P_S^{-1}(t) + J_{hc}(t+1)^T S_c^{-1}(t+1) J_{hc}(t+1) \quad (4.17)$$

$$P_M(t+1) = P_S(t+1) \tilde{L}_c(t+1) M_c \tilde{L}_c^T(t+1) P_S^T(t+1) \quad (4.18)$$

$$\tilde{L}_c(t+1) = \tilde{L}_c(t) + J_{hc}(t+1)^T S_c^{-1} \quad (4.19)$$

となる。ただし、

$$W_c(t+1) = P_S(t+1) J_{hc}(t+1) S_c^{-1}(t+1) \quad (4.20)$$

$$J_{hc}(t+1) = \frac{\partial \tilde{\mathbf{h}}_c}{\partial \mathbf{Y}_O} \Big|_{\mathbf{Y}_O(t); \mathbf{x}_M(t+1)} \quad (4.21)$$

である。ここで、 t は計算に用いた特徴点を区別するためのインデックスであり、実際の時間には対応していない。また、(4.15) の $\mathbf{x}_M(t+1)$ は観測点 $\mathbf{x}_c(t+1)$ に対応するモデルの点のモデル座標系での座標値である。これらの式からカメラの出力 \mathbf{x}_c を使って対象のおおよその位置が求められる。計算に必要なヤコビアンは、

$$\frac{\partial \tilde{\mathbf{h}}_c}{\partial \mathbf{X}_O} = \frac{\partial \mathbf{h}_c}{\partial \mathbf{x}_W} \frac{\partial \mathbf{x}_W}{\partial \mathbf{X}_O} = \frac{\partial \mathbf{h}_c}{\partial \mathbf{x}_W} \quad (4.22)$$

$$\frac{\partial \tilde{\mathbf{h}}_c}{\partial \phi_k} = \frac{\partial \mathbf{h}_c}{\partial \mathbf{x}_W} \frac{\partial \mathbf{x}_W}{\partial \phi_k} = \frac{\partial \mathbf{h}_c}{\partial \mathbf{x}_W} \frac{\partial R(\phi_O)}{\partial \phi_k} \mathbf{x}_M \quad (4.23)$$

と計算できる。ただし、 $k = x$ or y or z である。また、

$$\frac{\partial \mathbf{h}_c}{\partial \mathbf{x}_W} = \begin{bmatrix} \frac{f}{z_W} & 0 & -\frac{x_W f}{z_W^2} \\ 0 & \frac{f}{z_W} & -\frac{y_W f}{z_W^2} \end{bmatrix} \quad (4.24)$$

である。

(4.6), (4.13) に第2章のフィルタを適用した式も同様に計算でき、CCDカメラの観測が終了した時点での t を τ とし、

$$L_{\text{camera}} \equiv \tilde{L}_c(\tau) M_c \tilde{L}_c^T(\tau) \quad (4.25)$$

とすると、 \mathbf{x}_T を観測する時のフィルタが、

$$\hat{\mathbf{Y}}_O(t+1) = \hat{\mathbf{Y}}_O(t) + W_t(t+1) [\mathbf{x}_T(t+1) - \mathbf{x}_W(\hat{\mathbf{Y}}_O(t); \mathbf{x}_M(t+1))] \quad (4.26)$$

$$\text{Var}[\hat{\mathbf{Y}}_O(t+1)] = P_{\mathbf{Y}_O}(t+1) = P_S(t+1) + P_M(t+1) \quad (4.27)$$

$$P_S^{-1}(t+1) = P_S^{-1}(t) + J_{ht}(t+1)^T S_t^{-1}(t+1) J_{ht}(t+1) \quad (4.28)$$

$$P_M(t+1) = P_S(t+1) L(t+1) P_S^T(t+1) \quad (4.29)$$

$$L(t+1) = L_{\text{camera}} + \tilde{L}_t(t+1) M_t \tilde{L}_t^T(t+1) + \tilde{L}_T(t+1) M_T \tilde{L}_T^T(t+1) \quad (4.30)$$

$$\begin{cases} \tilde{L}_t(t+1) = \tilde{L}_t(t) + J_{ht}(t+1)^T S_t^{-1} \\ \tilde{L}_T(t+1) = \tilde{L}_T(t) \end{cases} \quad (4.31)$$

となる。ただし、

$$W_t(t+1) = P_S(t+1) J_{ht}(t+1) S_t^{-1}(t+1) \quad (4.32)$$

$$J_{ht}(t+1) = \frac{\partial \mathbf{x}_W}{\partial \mathbf{Y}_O} \Big|_{\mathbf{Y}_O(t); \mathbf{x}_M(t+1)} \quad (4.33)$$

である。同様に、 \mathbf{p}_T を観測する時のフィルタが、

$$\hat{\mathbf{Y}}_O(t+1) = \hat{\mathbf{Y}}_O(t) + W_T(t+1) [\mathbf{p}_T(t+1) - \boldsymbol{\xi}(\hat{\mathbf{Y}}_O(t); \mathbf{p}_M(t+1))] \quad (4.34)$$

$$\text{Var}[\hat{\mathbf{Y}}_O(t+1)] = P_{\mathbf{Y}_O}(t+1) = P_S(t+1) + P_M(t+1) \quad (4.35)$$

$$P_S^{-1}(t+1) = P_S^{-1}(t) + J_{hT}(t+1)^T S_T^{-1}(t+1) J_{hT}(t+1) \quad (4.36)$$

$$P_M(t+1) = P_S(t+1) L(t+1) P_S^T(t+1) \quad (4.37)$$

$$L(t+1) = L_{\text{camera}} + \tilde{L}_i(t+1) M_i \tilde{L}_i^T(t+1) + \tilde{L}_T(t+1) M_T \tilde{L}_T^T(t+1) \quad (4.38)$$

$$\begin{cases} \tilde{L}_i(t+1) = \tilde{L}_i(t) \\ \tilde{L}_T(t+1) = \tilde{L}_T(t) + J_{hT}(t+1)^T S_T^{-1} \end{cases} \quad (4.39)$$

となる。ただし、

$$W_T(t+1) = P_S(t+1) J_{hT}(t+1) S_T^{-1}(t+1) \quad (4.40)$$

$$J_{hT}(t+1) = \frac{\partial \xi}{\partial \mathbf{Y}_O} \Big|_{\mathbf{Y}_O(t); \mathbf{x}_M(t+1)} \quad (4.41)$$

である。

(4.27) や (4.35) は $\mathbf{x}_M(t+1)$ や $\mathbf{p}_M(t+1)$ に依存するので、これらの式から求まる

$$\mathcal{E}^p = \det P_{\mathbf{Y}_O}^{-1}(t+1) \quad (4.42)$$

が大きくなるように、 $\mathbf{x}_M(t+1)$ や $\mathbf{p}_M(t+1)$ を求める。このことは、触覚センサで触る位置を決めることに対応する。その結果得られる触覚センサの出力を使って、(4.26)~(4.41) から、より正確な \mathbf{Y}_O を求める。これにより対象の位置 \mathbf{X}_O と方向 ϕ_O を効率良く求めることができる。

4.4 形状モデルに誤差がある場合の位置計測

実際に本方法を使う場合は、現実にある対象の形状モデルを作成することになる。しかし、正確な対象のモデルを得ることは不可能な場合が多く、モデルは誤差を含むことになる。この場合には、対象の正確な位置を求めることはできない。なぜなら、位置の計算にモデルに含まれる値を使うからである。つまり、センサ情報と形状モデルから得られる対象の位置は、常に形状モデルの誤差による影響を受ける。そこで、触覚センサで少しでも形状モデルの誤差の少ない点に触ることが考えられる。その場合に、触覚センサで探索する望ましい位置を求められるように、前節の方法を拡張する。なお、形状モデルに誤差がある場合触覚センサを使っても、剛体としての対象全体の位置や姿勢には誤差が残るが、触覚センサで触ったエッジや頂点に関しては、(触覚センサに誤差がなければ) 正確な値を得ることができる。

モデルの点の区別をつけたい場合には、点 i を \mathbf{x}_{Mi} と書くことにする。モデル座標系上の点の位置の真値を \mathbf{x}_{Mi} とし、実際に得られるモデルの点の座標が

$$\bar{\mathbf{x}}_{Mi} = \mathbf{x}_{Mi} + \mathbf{v}_{Mi} \quad (4.43)$$

であるとする。ただし \mathbf{v}_{Mi} は平均 0, 共分散行列 V_{Mi} のノイズである。また, \mathbf{v}_{Mi} は $\mathbf{v}_c, \mathbf{v}_t, \mathbf{v}_T, \mathbf{m}_c, \mathbf{m}_t$ や \mathbf{m}_T と無相関であるとする。また, モデルの異なる点でも, 無相関であるとする。つまり

$$\bar{E}[\mathbf{v}_{Mi}\mathbf{v}_{Mj}^T] = \delta_{ij}V_{Mi} \quad (4.44)$$

である。

モデル上の座標値 $\bar{\mathbf{x}}_{Mi}$ に対応するワールド座標系上の値を

$$\begin{aligned} \bar{\mathbf{x}}_W &= R(\phi_O)\bar{\mathbf{x}}_{Mi} + \mathbf{X}_O = R(\phi_O)(\mathbf{x}_{Mi} + \mathbf{v}_{Mi}) + \mathbf{X}_O \\ &= \mathbf{x}_W + R(\phi_O)\mathbf{v}_{Mi} \end{aligned} \quad (4.45)$$

と表す。CCD カメラを使った特徴点の観測を表す (4.5) は, $\bar{\mathbf{x}}_W$ でテラー展開すると

$$\begin{aligned} \mathbf{x}_c &\doteq \mathbf{h}_c(\bar{\mathbf{x}}_W) - J_{h_c}R(\phi_O)\mathbf{v}_{Mi} + \mathbf{m}_c + \mathbf{s}_c \\ &= \bar{\mathbf{h}}_c(\mathbf{Y}_O; \bar{\mathbf{x}}_{Mi}) + J_{h_c}R(\phi_O)\mathbf{v}_{Mi} + \mathbf{m}_c + \mathbf{s}_c \end{aligned} \quad (4.46)$$

となる。ただし, J_{h_c} は $\mathbf{h}_c(\mathbf{x}_W)$ の \mathbf{x}_W による $\bar{\mathbf{x}}_W$ での Jacobian を表す。ここで $J_{h_c}R(\phi_O)\mathbf{v}_{Mi}$ について考えてみる。 \mathbf{v}_{Mi} はモデルが決まった時点で固定されるので, 同じ点の観測を行なうと, 毎回 $J_{h_c}R(\phi_O)\mathbf{v}_{Mi}$ だけ観測値にオフセットを生じる。つまり, $J_{h_c}R(\phi_O)\mathbf{v}_{Mi}$ は (2.3) の観測モデルの曖昧さ \mathbf{m}_t と同じ種類の誤差である。ただし, センサごとに固定されるのではなく, 観測点ごとに固定される。 $J_{h_c}R(\phi_O)\mathbf{v}_{Mi}$ の共分散行列は

$$J_{h_c}R(\phi_O)V_{Mi}R(\phi_O)^T J_{h_c}^T \quad (4.47)$$

と計算できる。同様に, 触覚センサによる頂点の観測を表す (4.6) は

$$\mathbf{x}_t = \bar{\mathbf{x}}_W(\mathbf{Y}_O; \bar{\mathbf{x}}_{Mi}) + R(\phi_O)\mathbf{v}_{Mi} + \mathbf{m}_t + \mathbf{s}_t \quad (4.48)$$

となる。 $R(\phi_O)\mathbf{v}_{Mi}$ がモデルの曖昧さと同様の値で, その共分散行列は $R(\phi_O)V_{Mi}R(\phi_O)^T$ である。これらの式に, 前節と同様に第 3 章で述べたセンサ位置選択の方法が適用できる。ただし, 分散を計算するために第 2 章の方法を使う際に, \mathbf{v}_{Mi} が原因で生じる曖昧さはセンサごとではなく, 観測点ごとに固定値となることに注意する。

次に, エッジについて考える。まず, 誤差を含んだモデル上の点の位置から, それをたぐエッジのパラメータと, その誤差の分散の求め方を述べる。2 点の座標とその誤差の分散がわかっているとす。つまり, モデル上の点の座標

$$\bar{\mathbf{x}}_1 = (\bar{x}_1, \bar{y}_1, \bar{z}_1), \quad \bar{\mathbf{x}}_2 = (\bar{x}_2, \bar{y}_2, \bar{z}_2) \quad (4.49)$$

が, 真値 $\mathbf{x}_1, \mathbf{x}_2$ と誤差 $\mathbf{v}_{x1}, \mathbf{v}_{x2}$ を使って

$$\bar{\mathbf{x}}_1 = \mathbf{x}_1 + \mathbf{v}_{x1}, \quad \bar{\mathbf{x}}_2 = \mathbf{x}_2 + \mathbf{v}_{x2} \quad (4.50)$$

と得られるとし、 $\mathbf{v}_{x1}, \mathbf{v}_{x2}$ の共分散行列を

$$V_{x1}, \quad V_{x2} \quad (4.51)$$

とする。このとき、 \bar{x}_1, \bar{x}_2 を通る線分

$$\begin{cases} x = az + p \\ y = bz + q \end{cases} \quad (4.52)$$

のパラメータ $\bar{\mathbf{p}}_M = [a, b, p, q]^T$ を求める。

$$\begin{cases} \bar{x}_1 = a\bar{z}_1 + p \\ \bar{y}_1 = b\bar{z}_1 + q \\ \bar{x}_2 = a\bar{z}_2 + p \\ \bar{y}_2 = b\bar{z}_2 + q \end{cases} \quad (4.53)$$

が成立するから変形して、

$$\bar{\mathbf{p}}_M = \frac{1}{\bar{z}_1 - \bar{z}_2} \begin{bmatrix} \bar{x}_1 - \bar{x}_2 \\ \bar{y}_1 - \bar{y}_2 \\ \bar{x}_2\bar{z}_1 - \bar{z}_2\bar{x}_1 \\ \bar{y}_2\bar{z}_1 - \bar{z}_2\bar{y}_1 \end{bmatrix} \quad (4.54)$$

となり \mathbf{p}_M がわかる。

この分散を求める。まず、 $\bar{\mathbf{x}} = [\bar{x}_1^T, \bar{x}_2^T]^T$ として、上式を $\bar{\mathbf{x}}$ で線形化した式を

$$\mathbf{p}_M = A \delta \mathbf{x} + \bar{\mathbf{p}}_M \quad (4.55)$$

とする。ただし $\delta \mathbf{x}$ は $\bar{\mathbf{x}}$ からの変位である。また、

$$A = \frac{1}{\bar{z}_1 - \bar{z}_2} \begin{bmatrix} 1 & 0 & \frac{\bar{x}_2 - \bar{x}_1}{\bar{z}_1 - \bar{z}_2} & -1 & 0 & \frac{\bar{x}_1 - \bar{x}_2}{\bar{z}_1 - \bar{z}_2} \\ 0 & 1 & \frac{\bar{y}_2 - \bar{y}_1}{\bar{z}_1 - \bar{z}_2} & 0 & -1 & \frac{\bar{y}_1 - \bar{y}_2}{\bar{z}_1 - \bar{z}_2} \\ -\bar{z}_2 & 0 & \frac{\bar{x}_2(\bar{z}_1 - \bar{z}_2) - (\bar{x}_2\bar{z}_1 - \bar{z}_2\bar{x}_1)}{\bar{z}_1 - \bar{z}_2} & \bar{z}_1 & 0 & \frac{-\bar{x}_1(\bar{z}_1 - \bar{z}_2) + (\bar{x}_2\bar{z}_1 - \bar{z}_2\bar{x}_1)}{\bar{z}_1 - \bar{z}_2} \\ 0 & -\bar{z}_2 & \frac{\bar{y}_2(\bar{z}_1 - \bar{z}_2) - (\bar{y}_2\bar{z}_1 - \bar{z}_2\bar{y}_1)}{\bar{z}_1 - \bar{z}_2} & 0 & \bar{z}_1 & \frac{-\bar{y}_1(\bar{z}_1 - \bar{z}_2) + (\bar{y}_2\bar{z}_1 - \bar{z}_2\bar{y}_1)}{\bar{z}_1 - \bar{z}_2} \end{bmatrix} \quad (4.56)$$

である。真の \mathbf{p}_M と $\bar{\mathbf{p}}_M$ の関係を求めるために、(4.55) に

$$\delta \mathbf{x} = \mathbf{x} - \bar{\mathbf{x}} = -[\mathbf{v}_{x1}^T, \mathbf{v}_{x2}^T]^T \equiv -\mathbf{v}_x \quad (4.57)$$

を代入すると

$$\mathbf{p}_M = -A\mathbf{v}_x + \bar{\mathbf{p}}_M \quad (4.58)$$

となる。よって、 $\bar{\mathbf{p}}_M$ の誤差は $A\mathbf{v}_x$ となり、その分散は

$$V_{P_M} = AV_x A^T \quad (4.59)$$

となる。ただし、 V_x は \mathbf{v}_x の共分散行列で

$$V_x = \text{diag}(V_{s1}, V_{s2}) \quad (4.60)$$

で与えられる。(4.12) に対応する

$$\bar{\mathbf{p}}_W = \boldsymbol{\xi}(\mathbf{X}_O, \phi_O; \bar{\mathbf{p}}_M) = \boldsymbol{\xi}(\mathbf{X}_O, \phi_O; \mathbf{p}_M + A\mathbf{v}_x) \quad (4.61)$$

を使って $\bar{\mathbf{p}}_W$ を求める。このとき $\bar{\mathbf{p}}_W$ の誤差 $\mathbf{v}_{P_W} \equiv \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{p}_M} A\mathbf{v}_x$ の共分散行列は

$$V_{P_W} = \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{p}_M} V_{P_M} \left(\frac{\partial \boldsymbol{\xi}}{\partial \mathbf{p}_M} \right)^T \quad (4.62)$$

となる。よって (4.13) は

$$\mathbf{p}_T = \bar{\mathbf{p}}_W(\mathbf{X}_O, \phi_O; \bar{\mathbf{p}}_M) + \mathbf{v}_{P_W} + \mathbf{m}_T + \mathbf{s}_T \quad (4.63)$$

となる。 \mathbf{v}_{P_W} が、観測点ごとに固定値となる曖昧さである。これに第3章で述べたセンサ位置選択の方法を適用することができる。

触覚センサで接触位置を選ぶ際には、できるだけ形状モデルの曖昧さの小さいところを選ぶことになる。しかし、同じ場所ばかり触っていると、第2章で述べた同じセンサばかり使うことと対応して、分散の減り方が飽和する。そこで、適度に異なる場所も触るようになる。そのようなセンシング戦略が、第2章と第3章の方法を合わせて使うことにより、自動的に得られる。

4.5 シミュレーション

図4.5に示すコーヒークップ状の形状を持つ対象について、モデルが誤差を含む場合に視覚と触覚を用いて対象の位置と姿勢を正確に求めるシミュレーションを行なった。ただし、融合値と分散を求めるためのフィルタとしては、第2章のフィルタでなくカルマンフィルタを使った。どちらを使っても、この章で提案した3次元形状モデルを使うことの有効性は確認できる。

図中の数字はエッジに付けられたラベルを表している。簡単のため、視覚から得られた点やエッジはモデルとの対応がわかっているものとした。また、触覚センサにより任意のエッジの情報が、他のエッジと間違えることなく得られるとした。

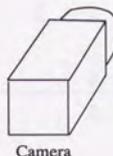
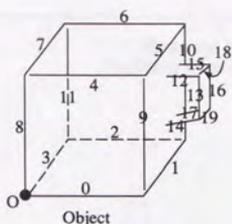


図 4.5 シミュレーション

まず、カメラで一回だけ画像を撮る。その画像を使って対象の位置を得る。しかし、画像のみでは形状モデルの誤差と CCD カメラの観測ノイズのために正確な位置は得られない。そこで、この画像を使って、第 4.4 節で提案した方法を用いて触覚センサで触るためのエッジを選ぶ。

触覚センサで触る時の誤差の時間経過を図 4.6 に示す。縦軸は誤差

$$\text{err} \equiv (\mathbf{Y}_O - \hat{\mathbf{Y}}_O(7))^2 = \sum_{i=1}^2 \{(x_i(t) - \hat{x}_i(t|t))^2 + (y_i(t) - \hat{y}_i(t|t))^2\}^{\frac{1}{2}} \quad (4.64)$$

を表す。ただし、カメラでの計測が終った時点の値で正規化してある。参考のため、エッジを 0 → 19 のように順番に触った時の誤差も示す。

表 4.1 選択されたエッジ

t	1	2	3	4	5	6	7
エッジ	3	0	1	2	3	17	3

$k = 7$ での誤差を、エッジを単に順番にラベル 0 からラベル 6 まで触った場合と較べると、本方法が $\text{err} = 0.0246$ 、順番に触った場合が $\text{err} = 0.277$ となった。これにより、本

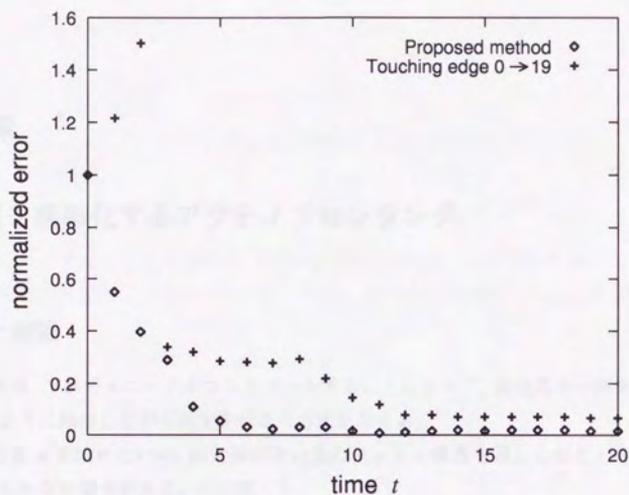


図 4.6 触覚センサを用いた時の誤差の時間経過

方法の有効性が確かめられた。

なお、参考のため、触覚センサで触るために選択されたエッジの一例を表 4.1 に示す。

第 5 章

逆問題を簡単化するアクティブセンシング

5.1 一般論

本章では、アクチュエータをコントロールすることによって、測定系の一部を測定にふさわしいように操作しながら測定を行なう方法を考える。

未知な量 \mathbf{x} を求めたいが、測定値がその他の値 \mathbf{p} との関数を通した値としてしかセンサから得られない場合がある。式で書くと

$$\mathbf{z} = \mathbf{f}(\mathbf{x}, \mathbf{p}) \quad (5.1)$$

のように表される。 \mathbf{z} がセンサ出力である。ここで \mathbf{p} を何らかの方法で測定または推定できるとする。関数 $\mathbf{f}(\cdot)$ がわかっている、その逆関数が求められれば

$$\mathbf{x} = \mathbf{f}^{-1}(\mathbf{z}, \mathbf{p}) \quad (5.2)$$

とすることにより、原理的には \mathbf{x} が求められる。

しかし、passive なセンシングでは \mathbf{p} は自由にならず、この逆演算は原理的に可能でも、式が複雑で計算が大変だったり、解が不安定になって精度が悪くなったりする [岡本, 1992]。能動的センシングでは、アクチュエータをコントロールすることにより \mathbf{p} を逆問題を解くために適した値にすることができ、計算を簡単にすることができる場合がある。

解の精度の問題を数式から考察する。 \mathbf{x}, \mathbf{p} が $\mathbf{x}_0, \mathbf{p}_0$ に近い値をとることがわかっているとすると、(5.1) は

$$\mathbf{z} = [A_x, A_p] \begin{bmatrix} \mathbf{x} - \mathbf{x}_0 \\ \mathbf{p} - \mathbf{p}_0 \end{bmatrix} + \mathbf{b} \quad (5.3)$$

のように線形化される。ただし、

$$A_x = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_0, \mathbf{p}_0}, \quad A_p = \frac{\partial \mathbf{f}}{\partial \mathbf{p}} \Big|_{\mathbf{x}_0, \mathbf{p}_0}, \quad \mathbf{b} = \mathbf{f}(\mathbf{x}_0, \mathbf{p}_0) \quad (5.4)$$

である。これを書き換えると

$$\mathbf{x} = A_x^{-1}(\mathbf{z} + A_x \mathbf{x}_0 - A_p(\mathbf{p} - \mathbf{p}_0) - \mathbf{b}) \quad (5.5)$$

となる。ただし A_x^\dagger は A_x の一般化逆行列を表す。

ここで、測定や推定を行なった値 z や p が誤差 δz , δp を持つとすると、 x の誤差 δx は

$$\delta x = A_x^\dagger \delta z - A_x^\dagger A_p \delta p \quad (5.6)$$

となる。 p を変えることによって A_x, A_p を変えることができる。そこで、(5.6), つまり

$$\delta x = \left(\frac{\partial f}{\partial x} \Big|_{x_0, p_0} \right)^\dagger \delta z - \left(\frac{\partial f}{\partial x} \Big|_{x_0, p_0} \right)^\dagger \frac{\partial f}{\partial x} \Big|_{x_0, p_0} \delta p \quad (5.7)$$

が小さくなるような A_x, A_p を取るように p を選べば, $\delta x, \delta p$ による誤差は小さくできる。 p の性質が何もわかっていないとすれば, これは A_x^\dagger の固有値の積と $A_x^\dagger A_p$ の固有値の積, つまり

$$\det A_x^\dagger, \det(A_x^\dagger A_p) \quad (5.8)$$

を小さくすることを意味する。 $\det(A_x^\dagger A_p) = \det A_x^\dagger \det A_p$ なので, これは結局

$$\det A_x^\dagger, \det A_p \quad (5.9)$$

を小さくすることになる。現実の問題では, いちいち (5.9) を計算しなくても, どのような p を選べば精度が良くなるかは直観的にわかることも多い。

p をコントロールすることによって逆問題を解きやすくしたり, 精度を上げる方法の具体例として, 第5.2節でロッドを対象に接触させることによって, 対象の位置や姿勢を求める問題を扱う。この方法は分布型触覚センサを用いることの代用として使うことができ, 第6章の実験では実際にこの方法を使用する。また, 第5.3節では, カメラ-マニピュレータ間の座標変換を求める方法を扱う。

5.2 接触による接触対象の求め方

ロボットが対象物を把持する場合, 視覚情報はもちろん大切であるが, それだけでは外界の情報としては不十分であり, 他のセンサを使って情報を補う必要がある。触覚は, 接触対象との接触位置の他に, 視覚からは得られない対象の硬さや, 摩擦係数なども得ることができ, 視覚を補うために有用なセンサである。ここでは, 接触対象の位置と姿勢の検出に話を限定して考察する。

接触点検出のための方法として概念的に最も簡単なものは, 接触するハンドやロッドなどの全体に分布型触覚センサを貼りつけるという方法であろう。しかしこの方法は,

- 曲面にセンサを貼りつけることが困難
- 信号線, 電源線が多数必要で, 信頼性に問題が出る
- コストが上がる

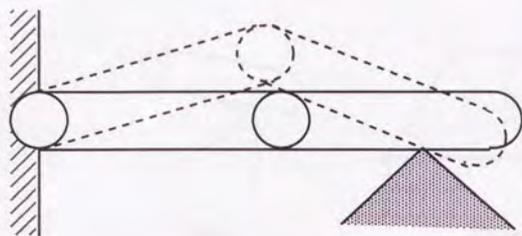


図 5.1 自己姿勢変形動作

などの理由で、問題がある。

これを解決するための方法として、いくつかの方法が提案されているが、その中に力覚センサを用いて自己姿勢変形動作から接触位置を求める方法がある [金子, 1993]。この方法は、概念的に図 5.1 のように表される。この図で、実線で表される状態から点線で表される状態への変化が自己姿勢変形動作である。この図で表される 2 つの状態から、もし対象が先のリンクに一点のみで接触しているなら、その位置は一意に定まる。

しかしこの方法は、リンクがエッジに垂直に接触していることを仮定している。つまり、図 5.1 で求めているエッジは紙面に垂直であることを仮定している。ところが、実際はエッジの位置は正確にはわかっていないため、リンクに対して斜めにあたるのが一般的である。また、アクチュエータの可動範囲の制限のために、垂直に触れないこともある。そのような場合には、この方法はもとのままでは使えない。

そこで、対象に対してリンクが 3 次元的にどのように接触するかわからない場合に、対象の位置や姿勢を求める方法を考える。対象としては、頂点、エッジ、平面を考える (図 5.2)。接触しているかどうかは、リンクの根本に付けられたフォース / トルクセンサを取りつけ、この出力をモニタすることで判別可能となる。リンクの形状は円筒形で、先端が半球状になっているものとする。

ここで、本節での考え方を、エッジへの接触で説明する。

ハンドがコンプライアンスを持つ場合を考える。何回か触るということは、複数の円筒 (色々な接触の仕方に対応) がエッジに接触することになる。これらに共通の接線が求まれば良い。

エッジの直線の方程式を

$$\boldsymbol{x} = \boldsymbol{x}_0 + t\boldsymbol{d} \quad (5.10)$$

と表す。この未知数は 4 個である。例えば、直線が yz 平面に平行でないとすれば

$$\boldsymbol{x}_0 = [0, y, z]^T, \quad \boldsymbol{d} = [1, b, c]^T \quad (5.11)$$

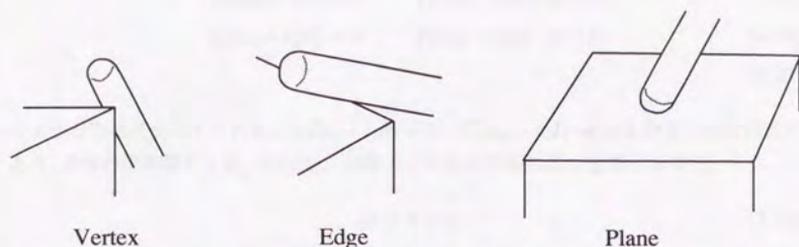


図 5.2 頂点, エッジ, 平面への接触

とできる。視覚により、大まかなエッジの位置はわかるので、どの平面に平行でないかということは設定できる。

リンクを表す円筒の方程式を

$$f_i(\mathbf{x}) = 0 \quad (5.12)$$

と表す。 i は円筒の位置のインデックスである。また、 \mathbf{x} での円筒の方線を

$$\mathbf{f}_i^n(\mathbf{x}) \quad (5.13)$$

と表す。実際には、円筒の標準形を

$$x^2 + y^2 = r^2 \quad (5.14)$$

$$0 \leq z \leq h \quad (5.15)$$

として、これを位置 \mathbf{X}_c と姿勢 ϕ_c で移動した円筒を考える。 \mathbf{X}_c, ϕ_c の表す同次変換を $A(\mathbf{X}_c, \phi_c)$ とすると、移動した円筒は

$$\tilde{\mathbf{x}}_t = A^{-1}(\mathbf{X}_c, \phi_c) \tilde{\mathbf{x}} \quad (5.16)$$

$$(x_t/w_t)^2 + (y_t/w_t)^2 = r^2 \quad (5.17)$$

$$0 \leq z_t/w_t \leq h \quad (5.18)$$

を満たす $\tilde{\mathbf{x}}$ となる。ただし、 $\tilde{\mathbf{x}}$ は \mathbf{x} の同次座標

$$\mathbf{x} = [x, y, z, 1]^T \quad (5.19)$$

とする。

リンクで複数回エッジに接触することにより

$$f_1(\mathbf{x}_0 + t_1 \mathbf{d}) = 0 \quad \mathbf{f}_1^n(\mathbf{x}_0 + t_1 \mathbf{d}) \cdot \mathbf{d} = 0 \quad (5.20)$$

$$f_2(\mathbf{x}_0 + t_2 \mathbf{d}) = 0 \quad \mathbf{f}_2^n(\mathbf{x}_0 + t_2 \mathbf{d}) \cdot \mathbf{d} = 0 \quad (5.21)$$

$$f_3(\mathbf{x}_0 + t_3 \mathbf{d}) = 0 \quad \mathbf{f}_3^n(\mathbf{x}_0 + t_3 \mathbf{d}) \cdot \mathbf{d} = 0 \quad (5.22)$$

$$\vdots \quad (5.23)$$

のような方程式がたてられる。 $f_i(\mathbf{x}_0 + t_i \mathbf{d}) = 0$, $\mathbf{f}_i^n(\mathbf{x}_0 + t_i \mathbf{d}) \cdot \mathbf{d} = 0$ が2つの方程式であり、全体の未知数が4個、それに、一回触るごとに1つ未知数 t_i が増えるので、

$$2k \geq 4 + k \quad (5.24)$$

なら解ける。つまり、最低、4回の接触で求められる。それ以上接触すれば、より正確な情報が得られる。しかし、この方法は原理的には可能であるが、式が複雑になって、結局非線形の連立方程式を数値的に解くことになり、計算が大変になる。

そこで、リンクの位置と姿勢の6自由度の中の、姿勢の3自由度をコントロールしながら触ることを考える。第5.1節でいえば、位置が \mathbf{x} に、姿勢がコントロールする値 \mathbf{p} にあたる。マニピュレータを姿勢が硬くなるようにハイブリッド制御し、同じ姿勢、違う位置で二回さわる。ハイブリッド制御とは、剛体の6自由度の中のいくつかの要素を力制御し、残りの部分を位置制御する方法である [高瀬, 1990]。これにより、平面が一つ簡単に定まる。先端の座標値の差を表すベクトルと、先端の z 軸方向のベクトルの外積がその平面の方線ベクトルとなる。また、円筒上で方線方向を考えることにより、平面上の点が定まる。詳しく書くと、ワールド座標で求めた方線を先端リンク（ロッド）の座標系に変換、そこで円筒との交点が簡単に求まる（図5.3）。これを基準座標系に戻す。同じことを違う角度で行なうと、もう一つ平面が定まる。この二つの平面の交線が求めるエッジとなる。

同じ姿勢でエッジに触るためには、姿勢を硬くするハイブリッド制御を行なう方法が考えられる。最初、エッジがあると推定される領域からある程度離れた位置にロッドを持っていき、そこから位置を少しずつエッジがあるはずの方向へ近づけていく。そして、フォース/トルクセンサにより接触を感知する。

なお、解の精度を上げるためには平面を求める際の2本の円筒の距離をできるだけ離れた方がよいことは、直観的にわかる。もちろん、求めるエッジに触れなければならないので、確実にエッジに触れる範囲内で、できるだけ離す。また、接触回数を増やして、本方法を何らかの統計的手法を使って拡張したものを使ってエッジを求めれば、誤差はもちろん減る。

同様のことを、頂点と平面について考える。まず、頂点について述べる。リンクの姿勢を同じにして2回触れば共通の線分が求まり（図5.4）、この上に求める頂点となる。そこで3回目は異なる角度で頂点に触り、線分との交点を求めれば、これが求める頂点である。

解の精度を上げるためには、共通の線分を求める時には2本の円柱が線分においてで

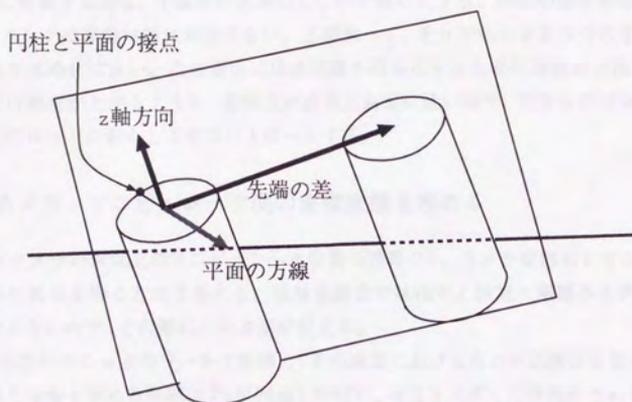


図 5.3 2つの接触から平面が1つ決まる

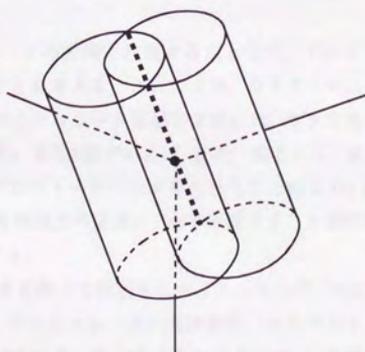


図 5.4 頂点の求め方

きるだけ直角に交わるようにすれば良い。3回目に触る時には、円柱の軸が前の2回の軸に対してできるだけ直角になるようにすれば良い。

平面に接触する時は、半球状の先端のどこかが触るとする。球状の部分が接触しさえすれば、リンクの姿勢は特に制限はない。3回触って、それから定まる3つの半球に共通な接平面を求めれば良い。この場合には逆問題を簡単にするための接触の方法はないが、できるだけ精度が上がるように、接触点が直角三角形に近い形で、できるだけお互いに離れるようにロッドの動かし方をコントロールする。

5.3 カメラ - マニピュレータ間の座標変換を求める

CCDカメラの画像上のマニピュレータ位置の情報から、カメラ座標系とマニピュレータ座標系の関係を得る方法を考える。視触覚融合では視覚と触覚の座標系を関連づけなければならぬので、その際にこの方法が使える。

3次元空間中の n 点の世界座標と、その画像における点の対応関係を使って、カメラの位置と姿勢を求める問題はPnP問題と呼ばれ、今日まで多くの研究がされてきた[出口, 1991]。これにより、3点の位置が知れていることが、この問題の解が存在するための最小の点数であること、6点一般的な位置であれば、解は一意に決まることなどが知られている。

これらの方法では、紙などに書かれた点や、立体上の点をカメラで撮影することにより、3次元内の点を扱っている。紙の場合には、一般的な位置の(平面上でない)点を得るために、何らかの方法で紙を動かす必要がある。また、カメラ - マニピュレータ間の座標変換を求めるためには、それらの点とマニピュレータ座標系との関係も何らかの方法で求めなければならない。

そこで、マニピュレータの先端を注視する点として、マニピュレータを動かすことにより、複数の点を得ることを考える。これにより、カメラ - マニピュレータ間の座標変換が一度に求められる。マニピュレータを動かす際には、カメラ画像を利用したビジュアルフィードバックを行なう。第5.1節でいえば、 p の一部として、点の画像上の位置が考えられる。さらに、ビジュアルフィードバックを行なうことにより、変換を得るための計算を非常に簡単にするような画像上の位置に、点を配置することが可能となる。この方法については、詳しくは後述する。

また、マニピュレータを使って校正するメリットとして、マニピュレータ自身を含んだ座標系で校正するため、マニピュレータの座標変換(キネマティクス)の誤差も含めて校正されるという点もあげられる。このようなシステムでは、最終的な目的はアクチュエータを使って対象の操作をすることである場合が多い。その際には、センサとアクチュエータで座標形があっていることが必要であり、分解してセンサだけ、アクチュエータだけで校正して合わせるよりも、センサからアクチュエータまで一度に座標変換を求めてしまう

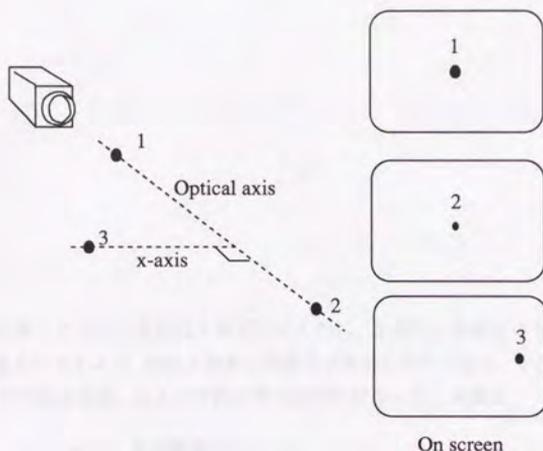


図 5.5 ビジュアルフィードバックで点を定める

方が、個々の誤差が吸収されて、より望ましいと思われる。

マニピュレータを動かして、先端の点を画像上の逆問題を解くために決められた位置に持ってくる。具体的にはまず、画像の中心に近い場合と遠い場合の2点のマニピュレータ座標の値をとる。これでカメラの光軸が定まる。次に、画像上で中心からx方向にある距離だけ離れた位置に点を動かして1点とる。この点から光軸に垂線を降ろすことによりx軸が定まる。これでカメラのマニピュレータ座標系に対する姿勢がわかる(図5.5)。あとは降ろした垂線の長さとも画像上の距離より、カメラから垂線への距離がわかる。つまり、カメラのマニピュレータ座標系での位置がわかる。これで、カメラからマニピュレータへの座標変換が求まる。

なお、マニピュレータやカメラの精度が位置に依存しないと仮定すると、解の精度を上げるためには測定する点をできるだけお互いに離せば良い。

第 6 章

実験

本論文で提案した方法の有効性を確認するために、基礎的な実験を行なった。具体的には、視触覚融合システムで、視覚と触覚の関係を自動的に求めてから、その関係を使って視触覚により対象物の認識、および対象位置の計測を行なった。実験は、

1. カメラ-マニピュレータ変換求める
2. フォース / トルクセンサによりエッジの位置を求める
3. 視触覚融合による対象認識
4. 視触覚融合による精度の良い位置計測

のように分解して行なった。

本章では、まず実験システムについての説明を行なう。次に、視触覚融合の予備実験として行なった2つの実験について述べる。1つは、第5.3節の方法に従ってカメラ-マニピュレータ変換を求める実験である。この変換により、視覚の情報と触覚の情報を関連づけることができるようになる。言い替えると、視覚による情報はカメラ座標系で求まるので、この変換を使うことにより、CCDカメラを使うだけでは不十分な情報を、触覚センサで補うことが可能となる。もう1つは、第5.2節で述べたフォース / トルクセンサを使って対象の位置や姿勢を求める実験である。以下の視触覚融合の実験では、この方法を触覚として用いる。

この2つの方法により、視触覚融合を行なう準備ができる。視触覚融合の実験としては、第4.2節で述べた対象認識と、第4.5節で述べた対象の位置および姿勢を求める実験を行なったので、これらについての説明を行なう。どちらも効率の良いセンシングを実現するために、触覚センサで探索する位置を、視覚の情報から求めている。これにより、視覚のみを用いるよりも精度の良いセンシングが実現されている。

6.1 システム構成

6.1.1 システム全体

本章では実験について述べるので、まず、実験を行なった視触覚融合システムについての説明を行なう。

このシステムは、センサフュージョンの実験を行なうことを目的として作られたものであり、マニピュレータ、それに附属するフォース/トルクセンサと CCD カメラより構成される。マニピュレータは、制御の研究よりも、他のセンサと組み合わせて全体としてセンシングシステムを構成することを目的として用いられる。CCD カメラの位置は、現在は固定されている。また、今回の実験では使っていないが、マニピュレータの関節の速度フィードバックの電流をモニタすることにより、電流によるトルクセンサを構成することも可能である。

CCD カメラ、マニピュレータ、フォース/トルクセンサ、およびその制御システムの構成を図 6.1 に示す。CCD カメラの映像は画像処理装置のメモリに入力される。このメモリは、workstation のメモリ空間上にダイレクトにマップされている。画像処理装置の能力に限界があるので、画像処理は基本的にこの workstation 上で行なう。一方、マニピュレータとフォース/トルクセンサのコントロールは PC-98 で行なう。workstation と PC-98 はイーサネットにつながれている。現在の、PC-98 でのコントロールのフィードバック周期は、約 20ms である。現在はまだ接続されていないが、将来はこのシステムにさらに CCD カメラを動かすための視線移動装置、および対象を把持することによってセンシングを行なうためのハンドを加える予定である。また、制御システムはより高速な処理の可能な DSP (Digital Signal Processor) で置きかえる予定である。

6.1.2 マニピュレータ

三菱重工工業株式会社製作の 7 軸マニピュレータを使用する。このマニピュレータの大きな特徴として、コントローラ部分の製作がユーザに任されていることが上げられる。これにより、ユーザはコントローラの部分にまで踏み込んで望みの制御系が構成できる。前述のように、このマニピュレータは他のセンサと組み合わせられて、計測システムを構成する要素として使われる。

マニピュレータの写真を図 6.2 に、模式図を図 6.3 に示す。図 6.3 に関節の回転方向、関節間の距離、およびそれぞれの関節に付随する座標軸の取り方が示されている。マニピュレータの代表的な特性を表 6.1 にまとめておく。

マニピュレータのサーボドライバユニットとしては、安川電機の SGD-01BS を使用する。これは速度制御とトルク制御が可能であるが、本実験では速度制御を使用する。速度制御のフィードバックの一段上に、PC-98 により位置制御のフィードバックループを構成

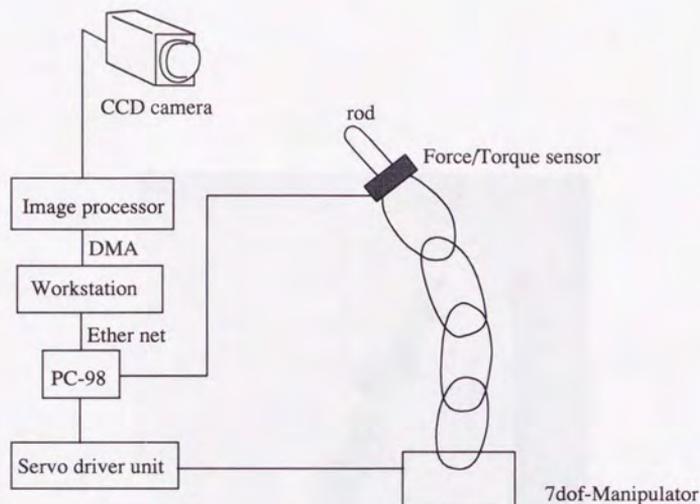


図 6.1 視触覚融合システム

表 6.1 マニピュレータの主な特性

	M1	M2	M3	M4	M5	M6	M7
動作範囲 (deg)	±225	+102 -94	±225	±114	±225	±111	ソフトリミ ットによる
最大速度 (rad/s)	0.8	0.45	2.27	2.03	3.9	3.9	5.0
減速比	480	480	206.67	208	120	120.59	80
周波数特性 (Hz)	22.0	11.7	35.7	12.2	26.2	9.0	9.8
カウンタ出力 パルス (p/r)	64 (モータ軸 1 回転当たり)						
重量 (N)	マニピュレータ本体 193 (ドライバユニット 181)						



図 6.2 マニピュレータの写真

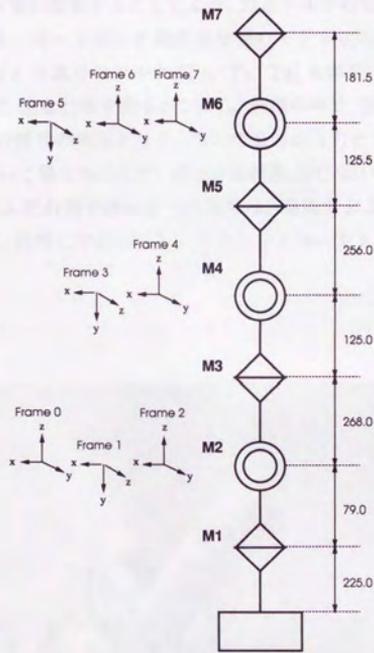


図 6.3 マニピュレータの関節

する。現在の位置制御のフィードバックの周期は20msである。このPC-98は後述するように、将来はDSPに置き換える予定である。なお、今回の実験では使用していないが、速度制御を行なうためのフィードバック系により流される電流をモニタすることにより、トルクのおおよその値を知ることが可能である。

6.1.3 フォース / トルクセンサ

マニピュレータの7軸の先にはフォース / トルクセンサを接続する。その先にロッドを接続し、その部分で対象に接触することにより、力とトルクのセンシングを行なう。

センサはビー・エル・オートテック株式会社製のモデル5/50を使用する。これは3成分の力(F_x , F_y , F_z)と3成分のトルク(T_x , T_y , T_z)を測定する6軸力覚センサである。この写真を図6.4に、主な仕様を表6.2に示す。仕様の中で、伝送速度が最大0.5ms(2000Hz)というのは、ユーザ側の信号の読みとりと、センサ出力から力とトルクを求めるための行列の計算が追いついていく場合のことで、現在は処理系(PC-98)のスピードで制限を受けている。現在のシステムでの信号読みとりの周期は、最高でおよそ240Hzである。実際に実験を行なう際には、同時にマニピュレータのコントロールも行なうため、さらに遅くなっている。



図 6.4 フォース / トルクセンサの写真

表 6.2 フォース / トルクセンサの主な特性

定格荷重		分解能			最大伝送速度 (ms)
力 (kgf)	トルク (kgf-cm)	Fx, Fy (gf)	Fz (gf)	Tx, Ty, Tz (gf-cm)	
5	50	5	15	30	0.5

6.2 カメラ - マニピュレータ変換

第 5.3 節で述べた方法に従ってカメラ - マニピュレータ変換を自動的に求める実験を行った。まず、マニピュレータの前に CCD カメラを適当においた。ただし、ビジュアルフィードバックができる程度にはカメラとマニピュレータの相対的な関係はわかっているとした。また、マニピュレータにとりつけたロッドの先端を求める点とし、これが判別できるように色をつけた。先端がカメラスクリーン上で目標位置にあるかどうかは、x 軸、y 軸ともに目標点から 1 dot 以内に収まっているかどうかで判断した。

3 点としては図 6.5 の座標系で、(256,240)、(256,240)、(100,240) を用いた。ただし、1 点目と 2 点目はカメラからの距離を変える。1 点目はカメラから約 100cm、2 点目は約 130cm、3 点目は約 115cm となった。

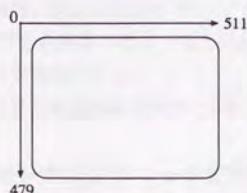


図 6.5 スクリーン座標系

この 3 点のデータからカメラ - マニピュレータの座標変換を求めた。この変換の精度を調べるために、カメラスクリーン上のいくつかの位置に、カメラからのフィードバックを使わずにロッドの先端を動かして見て、スクリーン座標系での誤差を調べた。この時の位置としては、計測に用いた点と用いなかった点の両方を測った。計測に用いた点は (256,240)、(100,240) であり、計測に用いなかった点としては、中心に関して (100,240) と対称な (412,240) と、計測に使った点と独立な方向の (256,100) を用いた。また、カメラからのロッドの先端の距離として 100cm、110cm、120cm、130cm の 4 種類を用いて実験を

行なった。この時の誤差を表 6.3 に示す。ただし誤差としては、画像上での pixel を単位とした距離を 30 回測って、その平均をとった。

表 6.3 CCD カメラ - マニピュレータ間の座標変換の誤差

	計測に用いた点		計測に用いなかった点	
	(256,240)	(100,240)	(412,240)	(256,100)
100cm の時の誤差 (とずれ)[dot]	6.7 (2.6)	9.7 (1.1)	4.2 (1.0)	17.3 (1.7)
110cm の時の誤差 (とずれ)[dot]	3.0 (2.0)	4.5 (0.81)	4.8 (1.0)	18.4 (1.3)
120cm の時の誤差 (とずれ)[dot]	2.4 (0.95)	1.7 (0.54)	8.4 (1.9)	20.1(1.1)
130cm の時の誤差 (とずれ)[dot]	2.0 (0.82)	1.0 (0.75)	9.4 (1.7)	20.3 (0.85)

これらの誤差には、マニピュレータの制御の誤差も含まれている。定常偏差は完全にはゼロにならないし、また、摩擦と制御の I 動作の影響で、ときどき位置が突然変わることがある。そこで、表 6.3 に 30 点の平均位置からの個々の点の、ずれの平均値も示す。

一般的な傾向として、計測に用いなかった点は、計測に用いた点より誤差が大きくなることがわかる。

今回は最低の数である 3 点しかとってはず、それらは x 軸上にならんでいるので、y 軸方向にずれた点 (256,100) では特に誤差が大きくなっている。カメラの内 部パラメータとしてカタログ値を使ったのでその影響が出ていると思われる。4 以上の点を用いると、座標変換の精度が向上することが期待できる。

いずれにせよ、第 5.3 節の方法で座標変換が非常に手軽に求められることがわかった。

6.3 フォース / トルクセンサによるエッジ位置の計測

第 5.2 節で述べた方法に従って、フォース / トルクセンサで接触する対象の位置と姿勢を求める実験を行なった。対象の代表としてエッジを選び、これにロッドが垂直にあたる場合と、斜めにあたる場合の計測の誤差を測った。誤差は、実際にアクチュエータを使用するときの誤差が問題になるという観点から、マニピュレータに取り付けたロッドの先を計測したエッジのそばに位置させてみて、その位置の誤差を測った。

まず、実験 1 として、ロッドがエッジに対して垂直にあたる場合の計測を行なった。4 回の接触の様子を図 6.6 に示す。この接触から第 5.2 節で述べた方法に従ってエッジの位置を求めた。その位置の誤差を調べるために、A 点の付近と B 点の付近にロッドの先を移動させ、その時の目標位置との誤差を測った。その値を表 6.4 に示す。なお、この誤差は定規を使って目視で測ったので、1mm 程度の誤差はありと考えられる。

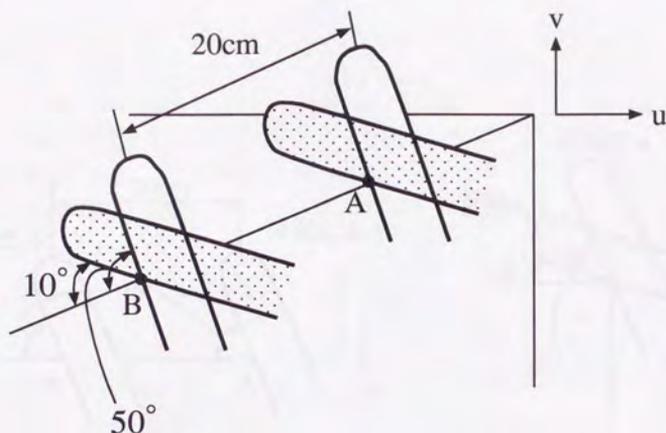


図 6.6 実験1の接触の様子

表 6.4 実験1の誤差

	u 方向	v 方向
A 点の誤差	1mm	-1mm
B 点の誤差	0mm	-2mm

次に実験2として、エッジに対してロッドが斜めにあたる場合の計測を行なった。4回の接触の様子を図6.7に示す。この図にあるように、ロッドはエッジに対して70度の角度で接触するとした。この時の、実験1と同様に求めた誤差を表6.5に示す。

マニピュレータを動かした時の誤差によって計測の誤差を測ったので、それぞれの誤差はマニピュレータの制御の誤差も含んでいる。実験1は満足できる結果だが、実験2は少し誤差が大きい。それでも、視覚のみから得られる結果よりはかなり精度が良いと思われる。

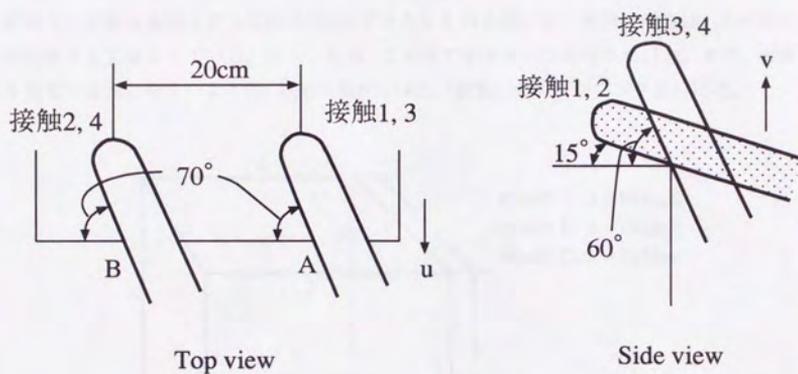


図 6.7 実験2の接触の様子

表 6.5 実験2の誤差

	u 方向	v 方向
A 点の誤差	14mm	-15mm
B 点の誤差	15mm	-15mm

6.4 視触覚による対象認識

第4.2節の方法に従って、視触覚による対象認識を行なった。この方法を使うことによって、対象の認識を行なうのと同時に、画像上の頂点やエッジが三次元形状モデルのどの部分に対応するかも同時にわかる。その結果を使って次の第6.5節で対象の位置計測を行なう。対象の形状モデルは図6.8に示すようなものを用いた。実際の対象は、モデルAに対応する形状をしている。エッジには、この図で示すような番号をつけた。また、対象を視覚で識別しやすいように、対象の頂点には色（黄色）のついたマークをつけた。

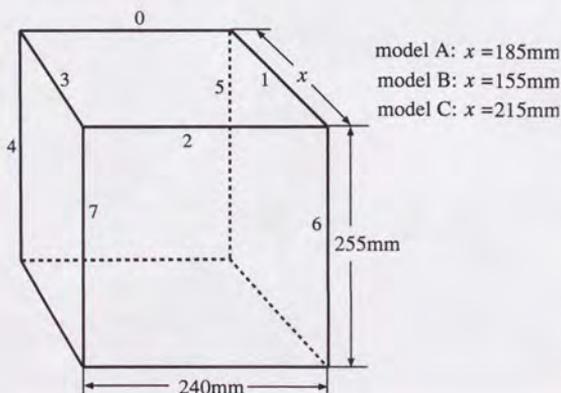


図 6.8 対象のモデル

まず、CCDカメラで対象の画像を一回取る。この画像にたいして、以下のような画像処理を行なう。最初に色情報を使って対象の頂点を検出する。次に、検出された頂点を結ぶすべての組み合わせについて、その部分にエッジがあるかどうかを調べる。これは、エッジがあると想定される部分をエッジと垂直にサーチして、その部分の輝度変化の大きさをもとに判断する。形状モデルとのマッチングのアルゴリズムを簡単にするため、頂点、およびエッジの検出では、ないものを誤ってあると認識することはないとする。その逆の、画像にある頂点やエッジを検出できないことはあっても良いとする。頂点とエッジの検出結果の一例を図6.9に示す。この図で白い部分が検出された頂点、およびエッジを表している。

頂点とエッジを検出したあとに、三次元形状モデルとのマッチングを行なう。その際に、第4.2節で述べた解釈木を用いる方法を利用する。この方法では、モデルの頂点と画像上の頂点の対応関係についてまず全ての組合せを考え、その後計算量を減らすために枝刈



図 6.9 頂点とエッジの検出結果

りを行なうが、その条件としては、第 4.2 節で述べたように様々な種類が考えられる。今回の実験では

1. 画像上のエッジの接続関係
2. モデル上で平行なエッジが画像上でも平行かどうか

を用いることによって枝刈りを行ない、その後、残った候補についてモデルから画像への変換（回転、平行移動、透視変換）を求め、モデルと画像が一致するかを調べた。モデルから画像への変換は、まず画像とモデル上の 6 点の対応関係から解析的に求め [出口, 1991]、その後、最小二乗法によってその解を反復的に改良することにより求めた。その時の画像と画像上に変換したモデルの一例を図 6.10 に示す。緑色の部分が画像上に変換された対象の形状モデルである。

図 6.8 のモデル A について、枝刈りによって候補数がどの程度減ったか、およびその時のモデルから画像への変換を求めるまでの計算時間を測ったものを表 6.6 に示す。計測は



図 6.10 画像上に変換したモデル

5回異なる画像（頂点の検出数6）で行ない、その時の平均値を表に示した。また、計算時間はSunのSPARCstation10で測ったものである。モデルから画像への変換を計算するのは非常に計算コストのかかる作業なので、枝刈りを行なうことによって大幅に計算時間が短縮されていることがわかる。なお、モデルから画像への変換を求める際の計算時間は、解析の手法にかかる時間のみを測った。実際にはその後の反復法による解の改良でさらに時間がかかるので、枝刈りを行なう場合と行なわない場合の差は、この表よりさらに大きくなると考えられる。

画像処理を行なうだけでは、画像や形状モデルの誤差のために対象を特定できない場合もある。実際に、表6.6の結果を得るために行なった画像処理では、対象に対応するモデルをAだけに限定しても、マッチングの仕方の候補数は最終的に2~4個残った。その際には第4.2節で述べたように、触覚センサーで情報を補うことにより対象の識別を行なう。そのための簡単な実験を行なった。形状が似ていて、画像のみからは誤差のためにどちらか認識できない対象のモデルを3つ用意し、そのどれに当たるかを視触覚を用いて判断し

表 6.6 枝刈りの効果 (5 回の平均)

	枝刈り後の候補数	計算時間 [s]
枝刈りなし	5040	114.5
接続関係のみで枝刈り	1118	26.7
接続関係+平行性で枝刈り	514	14.6

た。対象は図 6.8 のように、辺の長さが 30mm ずつ異なる ($x=185, 155, 215\text{mm}$) 3 種類のモデルを用意した。モデル A が実際の対象に対応するモデルである。

まず、視覚のみを用いて対象が判別できるかを調べるために、モデルと画像の正しい対応が取れた際の、モデルから画像への変換を求めて撮像面上の対応する点の誤差を 5 回条件を変えながら計測したら、表 6.7 のようになった。ここで、モデルから画像への変換は全て最小二乗法による改良を 1000 回行ったあとのものである。また、撮像面上のずれは、各対応点のずれの RMS(Root Mean Square) を用いた。なお、撮像面の大きさは $8.8\text{mm} \times 6.6\text{mm}$ である。この誤差が一番小さいモデルが、対象に対応するモデルと判断される。モデル A が実際の対象に対応するのであるが、この計測結果の平均値と標準偏差を考慮すると、視覚を使って 3 種類のモデルからモデル A を正確に判別するのは不可能なことがわかる。

表 6.7 視覚によるマッチングの誤差 [mm]

回数	モデル A	モデル B	モデル C
1	0.103	0.035	0.213
2	0.039	0.119	0.166
3	0.110	0.134	0.157
4	0.043	0.083	0.122
5	0.106	0.150	0.117
平均	0.080	0.095	0.155
標準偏差	0.032	0.034	0.035

一方、触覚から $x=185\text{mm}$ の辺の長さを条件を変えながら 5 回求めたら表 6.8 のようになった。この結果を使うことにより、モデル A が実際の対象に対応することがわかる。

この結果から、モデルで差がある部分を触覚で触ることにより、視覚からは判別できない対象でも判別が可能となることがわかる。

表 6.8 触覚による長さ 185mm の計測結果

回数	センシング結果 [mm]
1	190.8
2	178.1
3	184.3
4	178.5
5	182.6
平均	182.9
標準偏差	4.62

6.5 視触覚融合による精度の良い位置計測

第 4.5 節で行なったシミュレーションを現実の系で実験した。

まず、第 6.2 節の実験の方法でカメラ-マニピュレータ変換を求め、第 6.4 節の方法で対象とモデルの頂点やエッジの対応関係を求める。その結果を使って第 4 章の方法により、対象の位置を正確に得るために触覚センサで触るエッジを求め、実際に接触することにより対象のエッジの位置を計測する。その際に、第 6.3 節の、フォース/トルクセンサによりエッジの位置を求める方法を利用する。その結果を使って対象の位置を求め、この時の対象位置の誤差を測定した。実験の様子を図 6.11 に示す。なお、カメラ-マニピュレータ変換を求める際には正確さを増すため、第 6.2 節より 1 点増やし、4 点使って計算を行なった。

この時の誤差の推移を図 6.12 に示す。比較のために、エッジを番号の順番に（ただしマニピュレータが到達可能な位置にあるものだけ）触った時のエラーの推移を同じグラフで示す。また、この時に選択されたエッジを表 6.9 に示す。なお、エッジ 2 はマニピュレータの可動範囲外にあり触れられなかったため、候補からはずした。

この結果、接触するエッジを第 4 章の方法に従って選択することにより、計測の誤差が小さくなることがわかった。このことは、対象位置のセンシングがより効率良く行なえるようになることを意味する。なお、図で誤差が増えることもあるのは、マニピュレータで触ることによる計測に誤差があるからである。



図 6.11 実験の様子

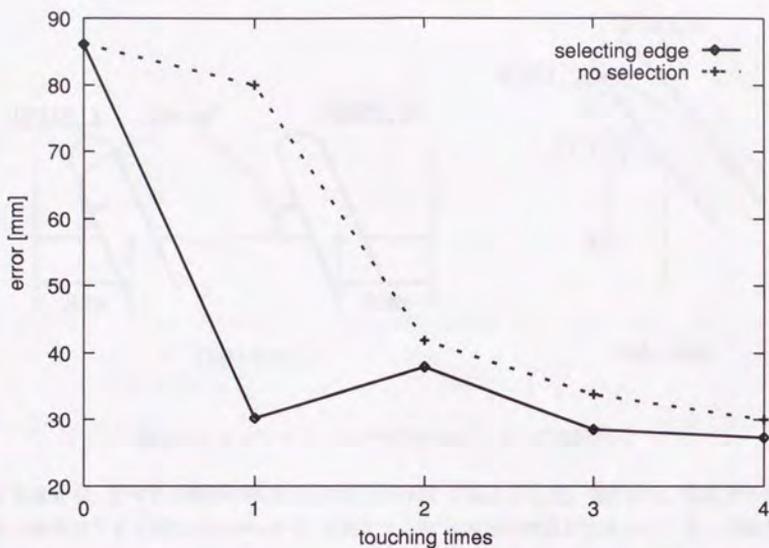


図 6.12 対象位置の測定値の誤差の推移

ロッドのマニピュレータへの接触角度（の推定値）は、接触可能なエッジを増やすために、図 6.13 のようになり甘くとした。そのため、エッジの位置計測の誤差は第 6.3 節の実験 2 よりもさらに悪くなっていると予想される。これは、可動範囲の広いマニピュレータを使えば改善される。

誤差は、マニピュレータに取りつけたロッドの先を対象の頂点があると予想される位置の付近に持っていき、目標位置との距離を測ることによって求めた。これを、2つの頂点について 2 種類の姿勢（つまり 4 回）で行ない、その平均を対象の位置の誤差とした。誤差は、マニピュレータで触る位置やマニピュレータの姿勢に依存するが、今回はこの点

表 6.9 選択されたエッジ

回数	1	2	3	4
選択した場合	3	1	5	1
順番に触る場合	0	1	3	4

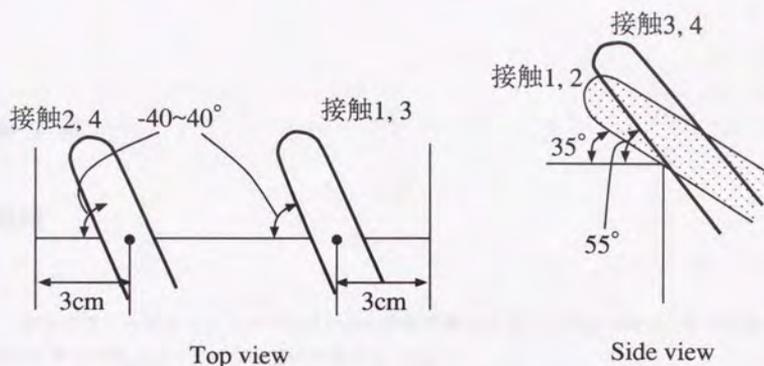


図 6.13 フォース / トルクセンサのエッジとの接触角度

は考慮せず、すべての接触で誤差は確率的には同じであるとした。実際には、特異点の付近で制御がうまく行なえなかったり、姿勢によっては関節の摩擦が大きかったりして誤差が大きくなることがあると考えられ、本来はこれも考慮するべきであるが、これは将来の課題である。なお、誤差の計測の際には、対象の位置計測の誤差の他に、制御の誤差、誤差計測（定規で測った）の誤差なども含まれていることに注意が必要である。

実験を行なう際には、次のような問題が起こった。対象の推定位置が実際と大きくずれている場合には推定位置にロッドを持っていくのでは触れない場合がある。また、マニピュレータの姿勢によってはロッド以外の部分が対象や、対象をのせてある台に接触してしまう場合もある。前者はロッドによる接触の際にもビジュアルフィードバックを使うことによって、後者はマニピュレータの冗長性を利用するようなプログラムを用意すればある程度解決可能である。しかしこのような問題は実験の本質ではないので、今回は人間が補助的にマニピュレータを操作することで解決した。

第 7 章

結論

センサフュージョンシステムのなかから視触覚融合システムを取りあげ、そのなかでのアクティブセンシングについて研究を行なった。

まず第 1 章で、視触覚融合、アクティブセンシング、センサフュージョンについて説明し、視触覚融合システムでアクティブセンシングについて研究することの意味を説明した。

第 2 章では、センサの観測モデルのノイズの平均値にオフセットがあっても良い場合に、複数センサの出力の融合値とその分散を逐次的に求める方法を提案した。次に、第 3 章で、より良い観測を行なうためのセンサ位置選択の方法の一般論を提案した。より良い観測の基準としては、推定値の誤差が小さいことと、センサ出力の連合が行ないやすいことを選んだ。そして第 4 章で、この 2 つを統合して、触覚センサで探索する位置を求める方法を提案した。その際には対象の 3 次元形状モデルを利用した。

また、第 5 章ではセンサをアクティブに動かすことにより測定の状況をコントロールし、精度の良い計測を行なったり、逆問題を解き易くする方法を提案した。この方法の具体例として、フォース / トルクセンサを使って頂点、エッジ、面などの位置を求める方法と、カメラ - マニピュレータ間の座標変換を求める方法を提案した。

次の第 6 章で、上で述べた方法を実装する視触覚融合システムについての説明、及び、そのシステムを使った基礎的な実験について述べた。この実験により、視覚だけでは判別できない対象が触覚も使うことにより認識可能となること、および能動性を積極的に使うことによって対象位置の効率の良いセンシングが可能となることがわかった。

これらの研究により、視触覚融合システムにおいてアクティブセンシングをどのように行なったら良いかが示された。特に、効率の良い計測を行なうためには、どのように触覚センサを動かせば良いかを求める方法が明らかになった。なお、このセンサ位置を求める方法は、視触覚融合などの、対象の位置を計測する一般のセンサフュージョンシステムにも適用可能である。

これから、ますますシステムの自動化は必要とされ、自律的なシステムが必要になるだろう。そのようなシステムでは、システムの抽象的な目的に応じてサブゴールを作りだし、

そのサブゴールを実現するために、自律的にセンサシステムを構成したり、センシング戦略を作り出したりすることが必要になるであろう。また、多種多様な現実に応じるためにも、センサはアクチュエータと組み合わせられ、目的に応じて位置を変えたりする必要が出てくるであろう。つまり、センサフュージョンやアクティブセンシングの研究は、これからはますます必要になってくると思われる。本研究が、そのようなシステムを構築するための一歩となれば幸いである。

謝辞

指導教官の石川正俊先生には、この論文の執筆をはじめとして、研究テーマの選び方、研究の進め方、発表の仕方など、多くの面でお世話になりました。また、私生活の面でも、人生設計に対する先生の言葉は私にとって大いに参考になるものでした。先生に対する感謝の気持ちは、とても言葉では言い表せないくらいです。

藤村貞夫先生には、研究に対する数々の助言をいただきました。また、本論文の副査を引き受けていただきました。心から感謝します。また、伊藤直さん、喜安千弥さんにも、研究に対する助言のほか、私生活でも大いに助けられました。

また、中野馨先生、館璋先生、森下巖先生にも本論文の副査を引き受けていただき、貴重な助言をいただきました。深く感謝致します。

研究室のメンバーにも大いに助けられました。彼らとの研究についての議論や助言はとても参考になりました。また、研究室のなごやかな雰囲気は、私にとってかけがえのないものでした。センサフェージョンのグループの石井君、山口さん、矢野君、並木君、中坊君とは、システム構築のための、回路の製作、ソフトのインストールなどを共同で進めました。また、研究についての良き話し相手であり、一緒に学習を行なう良き仲間でした。彼らがいなければ、この論文で述べたシステムは完成しなかったでしょう。また、田畑君は workstation の管理を通して私の仕事を大いに助けてくれました。石田君の教えてくれたけん玉は、息抜きとして大いに役にたちました。また、渡部君、今元君、大野君もイベントなどを企画してくれて、私の生活を充実させてくれました。

最後に、博士課程での生活を精神的、金銭的に暖かく見守ってくれた両親に感謝します。

参考文献

- [Allen, 1988] P.K. Allen: "Integrating vision and touch for object recognition task," Int. Journal of Robotics Research, vol.7, no.6, pp.15-33, 1988.
- [Aloimonos et al., 1988] J. Aloimonos, I. Weiss, and A. Bandyopadhyay: "Active Vision," Int. J. of Computer Vision, pp. 333-356, 1988.
- [浅田ら, 1990] 浅田, 木村, 白井: "時系列距離画像を用いた高さ地図の動的統合", 第8回日本ロボット学会学術講演会予稿集, pp.863-866, 1990.
- [Ayache, 1991] N. Ayache: "Artificial Vision for Mobile Robots," The MIT Press, 1991.
- [Blake and Yuille, 1992] A. Blake and A. Yuille: "Active Vision," MIT Press, 1992.
- [出口, 1991] 出口 光一郎: "コンピュータビジョン, グラフィックスのための射影幾何学 [IV]", 計測と制御, Vol.30, No.3, pp.241-246, 1991.
- [Dodge et al., 1988] Y. Dodge, V.V. Fedrov and H. P. Wynn eds.: "Optimal design and analysis of experiments," Elsevier science publishers, 1988.
- [Gaston and Lozano-Perez, 1984] P.C.Gaston and T.Lozano-Perez: "Tactile Recognition and Localization Using Object Models: The Case of Polyhedra on a Plane," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 3, pp.257-266, 1984.
- [Gibson, 1970] J.J. Gibson: "The Ecological Approach to Visual Perception," Houghton Mifflin, 1970.
- [Hashemipour, 1988] H.R. Hashemipour, S. Roy, and A. J. Laub: "Decentralized Structure for Parallel Kalman Filtering," IEEE Trans. Automatic Control, 33-1, pp. 88-94, 1988.
- [平井, 1988] 平井: "注視特徴の不動回転実現による把持物体-ハンド変換の推定法-基本手順のシミュレーションによる評価-", 第27回SICE予稿集, pp.383-384, 1988.

- [稲葉, 井上] 稲葉, 井上: “ロボットによる紐のハンドリング”, 日本ロボット学会誌, 3-6, pp.538-547, 1985.
- [石川, 1988] 石川 正俊: “並列処理を用いた能動的センサシステム”, 計測自動制御学会論文集, 24-8, pp. 860-866, 1988.
- [石川, 1993] 石川 正俊: “ロボットハンドとアクティブセンシング”, 日本ロボット学会誌, 11-7, pp.938-942, 1993.
- [金子, 1992] 金子 真: “Active Antenna”, 第10回日本ロボット学会予稿集, pp. 671-674, 1992.
- [金子, 1993] 金子 真: “把握までのセンシング戦略 その1: 接触点及び形状検出”, 日本ロボット学会誌, 11-7, pp.953-958, 1993.
- [片山, 1983] 片山 徹: “応用カルマンフィルタ”, 朝倉書店, 1983.
- [木下, 出口, 1992] 木下, 出口: “能動視覚による3次元形状認識”, 計測自動制御学会論文集, 28-1, pp.144-153, 1992.
- [Luo and Kay, 1989] R.C. Luo and M.G. Kay: “Multisensor Integration and Fusion in Intelligent System,” IEEE Trans. Syst., Man, Cybern., SMC-19-5, pp. 901-931, 1989.
- [松本, 亀島, 1990] 松本, 亀島: “ヒューマン・フレンドリ・ファクトリへの提唱”, 日本機械学会誌, 93, pp.831-835, 1990.
- [McKendall and M.Mintz, 1988] R.McKendall and M.Mintz: “Robust Fusion of Location Information,” Proc. IEEE Int. Conf. Robotics and Automat., pp.1239-1244, 1988.
- [Mukai et al., 1993] T. Mukai, T. Mori and M. Ishikawa: “A Sensor Fusion System Using Mapping Learning Method,” Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.391-396, 1993.
- [向井, 石川, 1994] 向井利春, 石川正俊: “複数センサによる予測誤差を用いたアクティブセンシング”, 日本ロボット学会誌, 12-5, pp.715-721, 1994.
- [Mukai and Ishikawa, 1994] T. Mukai and M. Ishikawa: “An Active Sensing Method Using Estimated Errors for Multisensor Fusion Systems,” Proceedings of the 1994 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems, pp.615-622, 1994.

- [Neisser, 1976] U. Neisser: "Cognition and Reality," Freeman, 1976.
- [岡本, 1992] 岡本良夫: "逆問題とその解き方", オーム社, 1992.
- [坂根ら, 1986] 坂根, 石井, 柿倉: "ハンドアイ行動シミュレータ: HEAVEN システムに基づく視覚センサのオクルージョン回避", 日本ロボット学会誌, 4-4, pp.382-392, 1986.
- [坂根ら, 1994] 坂根, 石川, 佐藤: "視覚と力覚のセンサフュージョンによる把持物体の接触位置推定", 日本ロボット学会誌, vol.12, no.5, pp.677-684, 1994.
- [Shafer et al., 1986] S.A. Shafer, A. Stentz, C.E. Thorpe: "An Architecture for Sensor Fusion in a Mobile Robot," Proc. IEEE Int. Conf. on Robotics and Automation, pp. 2002-2011, 1986.
- [Shimojo and Ishikawa, 1993] M. Shimojo and M. Ishikawa: "An Active Touch Sensing Method Using a Spatial Filtering Tactile Sensor," Proc. 1993 IEEE Int. Conf. on Robotics and Automation, 1, pp. 948-954, 1993.
- [Stansfield, 1988] S.A. Stansfield: "A robotic perception system utilizing passive vision and active touch," Int. J. Robotics Res., vol.7, no.6, pp.138-161, 1988.
- [高瀬, 1990] 高瀬國克: "剛体アームの力制御", ロボット工学ハンドブック (日本ロボット学会編), コロナ社, pp.249-258, 1990.
- [Yamada et al., 1993] Y. Yamada, A. Ishiguro and Y. Uchikawa: "A method of 3D object reconstruction by fusing vision with touch using internal models with global and local deformations," Proc. IEEE Int. Conf. on Robotics and Automation, pp.782-787, 1993.
- [山口, 1993] 山口 陽滋: "把握までのセンシング戦略 その2: 滑りと静摩擦係数の検出～表面粗さ情報の取得", 日本ロボット学会誌, 11-7, pp.959-965, 1993.
- [山崎, 石川, 1992] 山崎 弘郎, 石川 正俊 編著: "センサフュージョン", コロナ社, 1992.



inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM, Kodak



Kodak Gray Scale



© Kodak, 2007 TM, Kodak

A 1 2 3 4 5 6 **M** 8 9 10 11 12 13 14 15 **B** 17 18 19

