

Title	ハンドアイシステムにおける視覚系と運動系の融合
Author(s)	満田, 隆
Citation	大阪大学, 1998, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3144055
rights	
Note	

Osaka University Knowledge Archive : OUKA

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

Osaka University

1111

ハンドアイシステムにおける
視覚系と運動系の融合

満田 隆

大阪大学大学院 基礎工学研究科

1998年1月



ハンドアイシステムにおける
視覚系と運動系の融合

満田 隆

大阪大学大学院 基礎工学研究科

1998年1月

論文要旨

未知あるいは動的な環境の中で作業を行う知能ロボットにとって、視覚は必要不可欠なものである。視覚情報を用いてロボットを制御しようとするとき、一般的には視覚によって環境を観る手法（ロボットビジョン・視覚系）と環境モデルの中でロボットを動かす手法（ロボティクス・運動系）に分けて考えられる。すなわち‘見てから動く’という考え方である。しかし、視覚情報を用いてロボットを高速に動作させる場合、すなわち、‘見ながら動く’動作を行う場合には、視覚系と運動系の相互関係がロボットの‘ふるまい’に影響を与えるため、視覚系と運動系をハードウェア、ソフトウェアの両面から適切に融合することが重要である。本研究ではカメラを備えたロボットマニピュレータ（ハンドアイシステム）の制御において、カメラ配置とマニピュレータの運動学的構造が視覚情報を用いたフィードバック制御動作にどのような影響を及ぼすかを考察する。そして、人間の眼と腕に似せた構造のハンドアイシステムにおいては、視覚情報とマニピュレータ間のマッピングが単純化され、計算量の少ない単純な構成式のフィードバック制御が可能となること、さらにこのフィードバック制御はパラメータ誤差や時間遅れなどの外乱要因に対してロバストであることを示す。また、視覚フィードバック制御に用いるカメラの画像情報の特徴量検出法として、計算量の少ない大雑把なオプティカルフローの検出法を提案し、視覚系単独で用いる場合には非力な手法が、運動系と組み合わせて用いると大変優れた手法と成りうることを示す。

目次

1 序論	1
1.1 本研究の背景	1
1.1.1 視覚センサを用いたロボットの制御	1
1.1.2 ビジュアルフィードバック	2
1.1.3 ビジュアルサーボにおけるハンドアイマッピング	4
1.2 本研究の目的と本論文の概要	5
2 人間型ハンドアイシステムにおける線形近似逆運動学	7
2.1 緒言	7
2.2 人間型ハンドアイシステムのモデル	8
2.3 関節空間	9
2.4 視空間	11
2.4.1 眼球視空間	12
2.4.2 網膜視空間	13
2.4.3 首視空間	15
2.5 視空間と関節空間の線形性	16
2.6 視空間と関節空間の線形近似	18
2.7 線形近似に適したカメラ配置とマニピュレータの運動学的構造	21
2.8 結言	24
3 視空間を用いたビジュアルサーボ	27
3.1 緒言	27
3.2 ビジュアルサーボ	28
3.3 線形ビジュアルサーボ	30
3.4 漸近安定性	32
3.5 実験	33
3.5.1 実験設備	33
3.5.2 軌道特性	34
3.5.3 カメラ角度に依存しない性質	35
3.6 モデル誤差に対するロバスト性	36
3.6.1 カメラの姿勢誤差による影響	36
3.6.2 マニピュレータの関節角度誤差による影響	36
3.7 ビジュアルサーボに適したカメラ配置とマニピュレータの運動学的構造	38

3.7.1	ヤコビアン非線形性が及ぼす軌道への影響	38
3.7.2	ハンドアイシステムのモデル	39
3.7.3	ハンドアイシステムの構造とヤコビアンの関係	40
3.8	結言	42
4	ビジュアルサーボと画像特徴検出法	43
4.1	緒言	43
4.2	ビジュアルサーボと画像特徴検出法	44
4.3	粗いオプティカルフローを用いたビジュアルサーボ	46
4.3.1	はじめに	46
4.3.2	粗いオプティカルフローを用いた画像の並進量と回転量の検出	46
4.3.3	粗いオプティカルフローに関する補足説明	48
4.3.4	階層画像を用いた段階的位置決め	50
4.3.5	平面運動の検出実験	51
4.3.6	ビジュアルサーボの漸近安定性	58
4.3.7	ビジュアルサーボ実験	63
4.3.8	環境の明度変化に対するロバスト性	66
4.3.9	エッジ抽出を行った画像への適用	69
4.3.10	おわりに	71
4.4	結言	71
5	結論	73
	謝辞	77
	付録	79
A	ステレオカメラの運動学	79
B	疑似逆行列を用いたステレオビジュアルサーボ	83
B.1	ステレオ視のモデル	83
B.2	イメージヤコビアン	84
B.3	フィードバック入力の生成	85
	参考文献	87
	関連文献	93

目次

1.1	視覚センサを用いたロボット制御	1
1.2	‘見てから動く’ (look and move)	1
1.3	‘見てから動く’ ビジュアルフィードバック	2
1.4	ビジュアルサーボ	3
1.5	ハンドアイマッピング	4
2.1	人間型ハンドアイシステムのモデル	9
2.2	ステレオカメラの幾何学	9
2.3	デカルト空間に投影した関節空間	10
2.4	視空間の構成	11
2.5	眼球視空間	12
2.6	デカルト空間に投影した眼球視空間	12
2.7	網膜視空間	13
2.8	デカルト空間に投影した網膜視空間	14
2.9	デカルト空間に投影した首視空間	15
2.10	デカルト空間に投影した関節空間 (広領域)	17
2.11	デカルト空間に投影した関節空間 (狭領域)	17
2.12	視空間Aに投影した関節空間 (広領域)	17
2.13	視空間Aに投影した関節空間 (狭領域)	17
2.14	視空間Bに投影した関節空間 (広領域)	17
2.15	視空間Bに投影した関節空間 (狭領域)	17
2.16	デカルト空間を用いた線形近似結果 (広領域) と誤差分布	19
2.17	視空間Aを用いた線形近似結果 (広領域) と誤差分布	19
2.18	視空間Bを用いた線形近似結果 (広領域) と誤差分布	19
2.19	デカルト空間を用いた線形近似結果 (狭領域) と誤差分布	20
2.20	視空間Aを用いた線形近似結果 (狭領域) と誤差分布	20
2.21	視空間Bを用いた線形近似結果 (狭領域) と誤差分布	20
2.22	人間型ハンドアイシステムの構造評価	21
2.23	人間の手先の可動領域 (近似対象領域)	21
2.24	眼球位置と線形近似精度 (対象領域S)	22
2.25	眼球位置と線形近似精度 (対象領域L)	22
2.26	前腕, 上腕の長さの比と線形近似精度 (対象領域S)	23
2.27	前腕, 上腕の長さの比と線形近似精度 (対象領域L)	23
2.28	視空間Aを用いた場合の線形近似結果と誤差分布	24

2.29	視空間Bを用いた場合の線形近似結果と誤差分布	25
3.1	従来のビジュアルサーボとの比較	28
3.2	デカルト空間における吸引領域	33
3.3	関節空間における吸引領域	33
3.4	実験設備	33
3.5	線形ビジュアルサーボと従来の手法の比較	34
3.6	カメラ固定時と回転時の手先軌道	35
3.7	カメラ回転時の画像座標 X^L の軌道	35
3.8	カメラ回転時の視空間における誤差の軌道	35
3.9	カメラの姿勢誤差による影響	37
3.10	マニピュレータの関節角度誤差による影響 (C-VS の場合)	37
3.11	ステレオ型ハンドアイシステムのモデル	40
3.12	単眼型ハンドアイシステムのモデル	40
3.13	ステレオ型システムにおける時不変安定度	41
3.14	単眼型システムにおける時不変安定度	41
3.15	リンク長さの比と時不変安定度	41
3.16	カメラ位置と時不変安定度	41
4.1	カメラを用いた平面上の物体の位置決め	43
4.2	画像の並進	46
4.3	並進量の検出モデル	46
4.4	階層画像を用いた段階的な位置決め	50
4.5	使用した画像 (150×150)	51
4.6	原画像 (96×96)	52
4.7	シフト画像 (96×96)	52
4.8	粗画像 (12×12)	52
4.9	シフト粗画像 (12×12)	52
4.10	オプティカルフロー	52
4.11	並進量の推定	53
4.12	回転量の推定	54
4.13	x 方向, y 方向に平行移動させた場合の推定並進量 v_x, v_y とビジュアルサーボにおける収束安定性 dU/dt	56
4.14	x 方向に平行移動させて, 回転させた場合の推定並進量 v_x と推定回転量 v_θ とビジュアルサーボにおける収束安定性 dU/dt	57
4.15	漸近安定性の解析 (dU/dt の分布図)	60
4.16	dU/dt の分布図 (96×96)	61
4.17	dU/dt の分布図 (12×12)	62
4.18	実験システム	63
4.19	X 座標の変化	65

4.20	Y 座標の変化	65
4.21	θ の変化	65
4.22	X 座標の変化	65
4.23	Y 座標の変化	65
4.24	θ の変化	65
4.25	基準画像	67
4.26	明るい画像	67
4.27	暗い画像	67
4.28	明るい画像での推定並進量	68
4.29	明るい画像での推定回転量	68
4.30	暗い画像での推定並進量	68
4.31	暗い画像での推定回転量	68
4.32	空間 1 次微分オペレータ	69
4.33	明るい画像での推定並進量	70
4.34	明るい画像での推定回転量	70
4.35	暗い画像での推定並進量	70
4.36	暗い画像での推定回転量	70
A.1	ステレオカメラの幾何学	80
A.2	注視におけるステレオカメラの幾何学	82
B.3	ステレオカメラのモデル	84

表目次

2.1 人間型ハンドアイシステムの各パラメータ	8
4.1 ビジュアルサーボに用いる画像特徴検出法	44
4.2 各解像度で画像処理に要する時間	63
4.3 最終的な収束誤差	64
5.1 サフィックスの定義	83

第 1 章 序論

1.1 本研究の背景

1.1.1 視覚センサを用いたロボットの制御

現在、自動車工場などの多くの生産現場では、組み立て、塗装、溶接などの作業を、人間をはるかに凌ぐようなスピードと精度で実行することができるロボットマニピュレータが活躍している。これらのロボットは、あらかじめ操作者によって手先のとおる目標軌道を教示されており、実行時には教示された手先の位置と速度を達成するようにロボットの各関節角を制御することによって作業を実現している [1]。これらのロボットの多くは、内界センサの情報のみに基づいて制御を行っているために、環境の変化に対応することができない。たとえば、ロボットの形状が変化したり、ロボットの関節に取り付けられた角度センサの精度が劣化した場合には、目標とする手先の位置と実際の位置に誤差が生じる。また、操作対象である物体の位置や形状が変化した場合には、操作者は再度、ロボットに教示を行わなければならない [2]。これに対して、外界センサを用いることにより、ロボットや環境の変化に対応できる知能ロボットの研究も盛んに行われており、一部は実用化されている。中でも視覚センサは非接触で外部環境を認識するのに最も有効なセンサとして期待されている [3]。

視覚センサを用いてロボットを制御する場合 (Fig.1.1), 通常は視覚により作業環境を理解してからロボットを動作させる '見てから動く (look and move)' 方法がとられる (Fig.1.2)。たとえば、ロボットマニピュレータでコップを把持する場合、まず、視覚セン

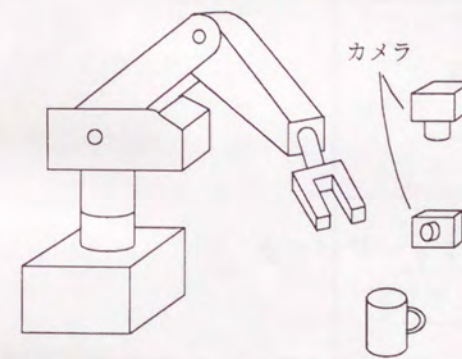


Fig. 1.1: 視覚センサを用いたロボット制御

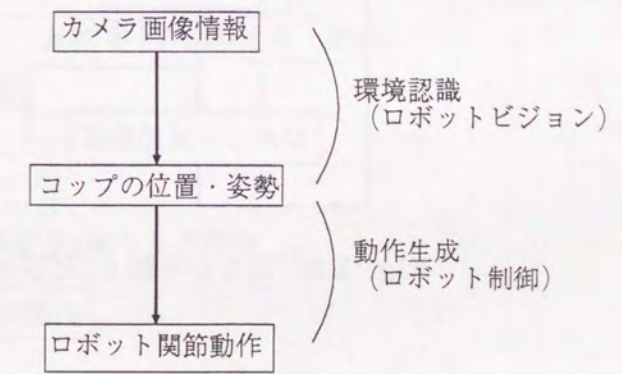


Fig. 1.2: '見てから動く' (look and move)

サでコップの位置と姿勢を計測する。そして、コップの位置に手先を伸ばす軌道を計画し、

ロボットを動作させる。この方法では、視覚センサの座標系とロボットの座標系のキャリブレーション誤差がロボットの動作結果にそのまま誤差として現れる問題がある。たとえば、視覚センサによって計測したコップの位置が正確であったとしても、ロボットの座標系とカメラの座標系がずれていれば、ロボットの手先の到達位置はコップの位置とは一致しない。この問題はカメラとロボットの正確なキャリブレーションにより解消することができるが、キャリブレーションには大変な手間と時間が必要である [4]。

1.1.2 ビジュアルフィードバック

視覚センサより得た情報を参照しながらフィードバックを行うことで、これらの誤差に対してロバストに動作を行うことができるロボットの研究も行われており、一部は実用化されている [5]。たとえば、視覚センサによってコップの位置とロボットの手先の位置を同時に計測し、その相対位置関係をもとにフィードバックを行うことによって、視覚センサの座標系とロボットの座標系がずれた場合でも正確にロボットの手先の位置決めを行うことができると考えられる。このように視覚情報を用いてフィードバックを行う制御は一般にビジュアルフィードバックと呼ばれる [6]。カメラ画像を用いたビジュアルフィードバック

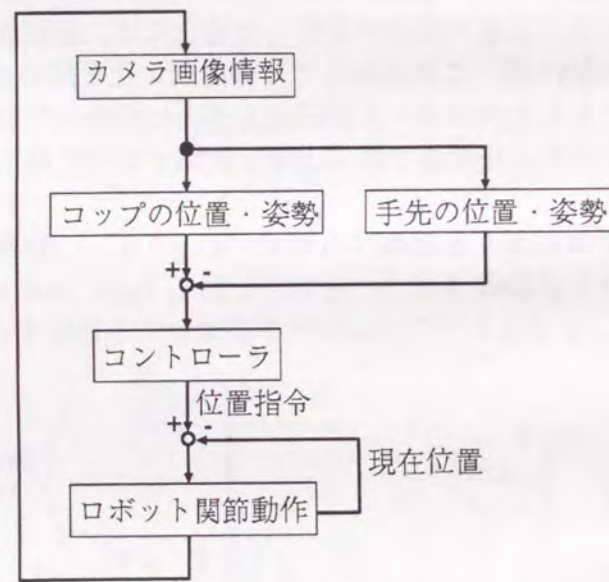


Fig. 1.3: '見てから動く' ビジュアルフィードバック

の研究はかなり古くから行われているが [7] (1973年)、カメラで得られる画像データは一般に数万から数十万個の濃淡値データであるために、画像中の物体の位置と姿勢を求めためには非常に多くの計算量が必要で、研究が始められた当時のハードウェアでは、画像を取り込んでからロボットの動作までかなりの時間を要した。このため、ビジュアルフィードバックは実用レベルに達することができず、以後はロボットビジョンの研究とロ

ット制御の研究に分化して研究が続けられた。しかし、近年はコンピュータの処理速度の飛躍的な向上によって、簡単な画像処理であればリアルタイムに行えるようになり、ビジュアルフィードバックも実用可能なレベルに達しつつある。

ビジュアルフィードバックの初期の研究では、コントローラがロボットへ手先の目標位置（またはロボットの関節角度）を指令した後、ロボットの動作が完了した時点で次の画像を取り込み、次のループへと処理を進める手法が用いられていた。すなわち、従来の'見てから動く'制御を繰り返して行う手法が取られていた (Fig.1.3)。この手法は'見てから動く'ビジュアルフィードバックと呼ばれている [8]。'見てから動く'ビジュアルフィードバックでは、視覚センサによる環境認識（視覚系）とロボット制御（運動系）を分けて考えることができるので、既に長年にわたって研究されてきたロボットビジョン [9] とロボット制御 [10] という二つの研究分野の資産を有効に活用できる利点があるが、高速な動作は望めない。これに対して近年は、コントローラがロボットへ手先の目標速度（またはロボットの関節角速度、関節トルク）を指令した後、すぐに次の画像を取り込み、次のループへと処理を進めるビジュアルサーボと呼ばれる手法が注目を浴びている (Fig.1.4)。ビジュアルサーボは高速にフィードバック動作を行うため、動く物体を扱うようなタスクにも応用可能である。しかし、ビジュアルサーボにおいては、視覚系のフィードバックループに運動系が組み込まれ、視覚系と運動系のそれぞれが並列に動作を行うために、視覚系と運動系の融合方法が制御特性に影響を与える重要な要因となる。たとえば、カメラとコンピュータの配置、画像特徴の抽出法とその選択法、視覚系のフィードバックループと運動系のフィードバックループのサンプリングタイムの違いなど、視覚系と運動系が相互に関連しあう問題は数多く存在するが、これらの研究はまだ始められたばかりである [8]。

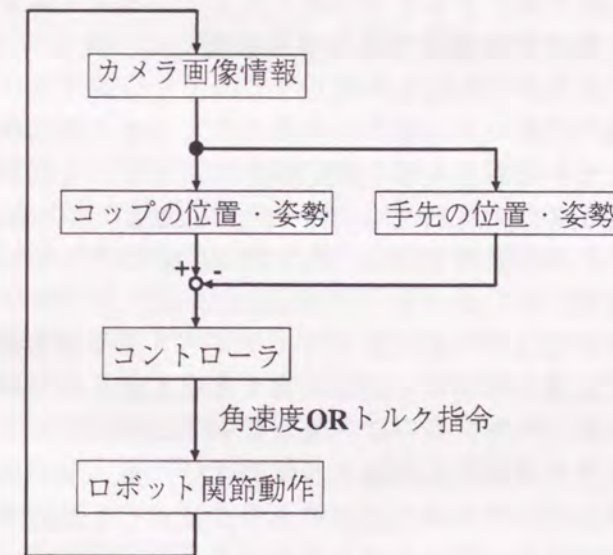


Fig. 1.4: ビジュアルサーボ

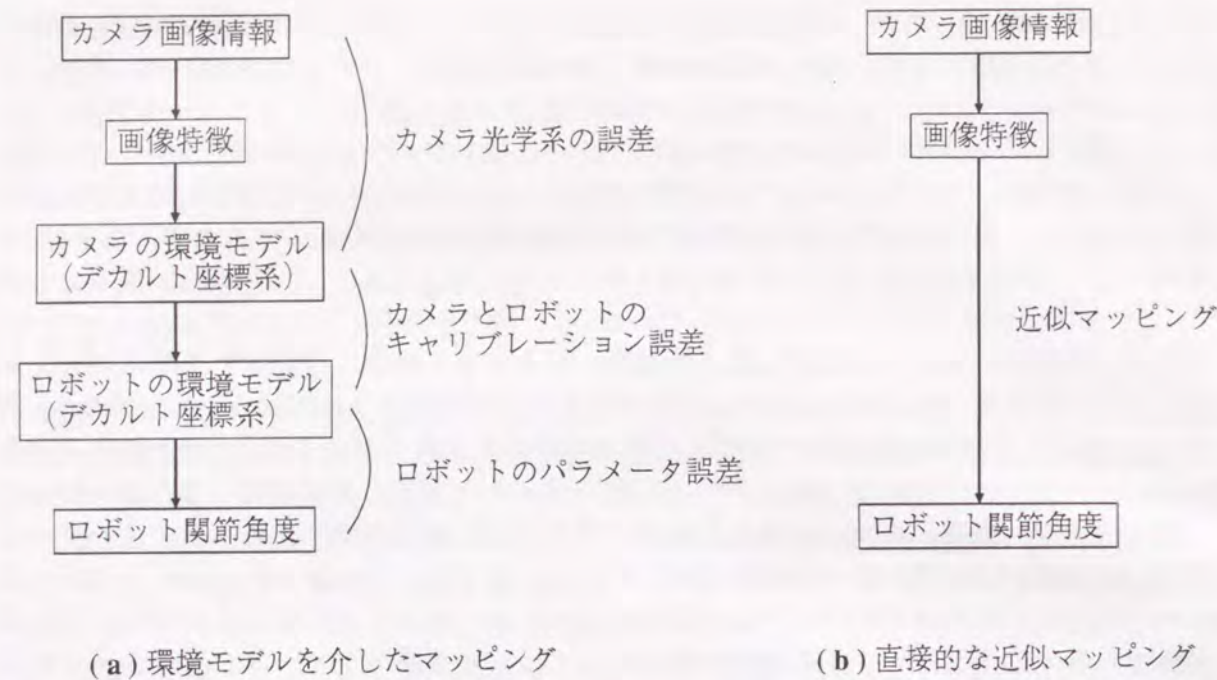


Fig. 1.5: ハンドアイマッピング

1.1.3 ビジュアルサーボにおけるハンドアイマッピング

カメラを用いてロボットを制御する場合には、システムモデルとして、カメラの画像情報とロボットの関節角度のマッピング（ハンドアイマッピング）が必要である。そして通常はデカルト座標系で表現する環境モデル（先の例では手先またはコップの位置と姿勢）を介してマッピングを行う（Fig.1.5(a)）。これまでに提案されているビジュアルサーボのほとんどの手法も、デカルト座標系を介したマッピングに基づいたものである。このマッピングは、デカルト空間における幾何学的拘束を用いた座標変換（ステレオカメラによる三角測量の原理、ロボットマニピュレータの運動学など）を基礎としたものであり、カメラ画像情報からカメラの環境モデル、カメラの環境モデルからロボットの環境モデル、そしてロボットの環境モデルからロボットの関節角度と大きく分けて3つの変換を用いる。このため、各変換においてはそれぞれ多くのパラメータが必要であり、これらのパラメータ誤差が重なり合って全体のマッピングに影響を及ぼす。これらの誤差をすべてキャリブレーションによってキャンセルするのは、大変な手間と時間が必要である [11]。

さて、カメラを用いてロボットを制御する場合には、カメラの画像情報とロボットの関節角度を結びつけるマッピングのみが必要であり、必ずしも環境モデルとしてのデカルト座標系を介させる必要はない。さらに、ビジュアルサーボを行う場合には、Fig.1.5(b)に示すように、カメラの画像情報とロボットの関節角度を結びつける近似的なマッピングを用いても、その誤差が収束漸近安定を満たす範囲内であれば収束動作は可能である。また、キャリブレーションが不正確である、あるいはパラメータが絶えず変動するような環境においてビジュアルサーボを行う場合、デカルト座標系を用いた厳密なマッピングは、単

に計算量を増やすだけでなく、マッピングの途中に含まれるパラメータの誤差が厳密なモデルのもとでの計算によって増幅される可能性がある。

以上の理由により、ビジュアルサーボにおいては近似的であっても単純なマッピングを行ったほうが、キャリブレーションの手間の簡略化、計算量の抑制、パラメータ誤差に対するロバスト性の面から有利であると考えられる。ただし、これはカメラ画像情報とロボットの関節角度の関係が単純なマッピングとして近似できるような場合に限っての議論である。非線形で常に大きく変動するような複雑なマッピングを単純なマッピングとして近似しても、近似誤差が大きく影響して、逆に不安定なシステムとなってしまう。つまり、ビジュアルサーボにおいては、カメラ画像情報とロボットの関節角度のマッピングが単純となるようなカメラとマニピュレータの構造および配置（ハードウェア的な視覚系と運動系の融合）と、時変的なパラメータをなるべく含まない単純な近似マッピング（ソフトウェア的な視覚系と運動系の融合）が重要であると考えられる。

1.2 本研究の目的と本論文の概要

本研究では、マニピュレータを用いたビジュアルサーボにおいて、カメラの配置とマニピュレータの運動学的構造が、カメラとロボットのマッピングにどのような影響を及ぼすのかを考察し、制御特性からみたハンドアイ構造の評価法ならびに最適設計法を構築することを目的とする。また近似的なマッピングを用いたビジュアルサーボが、制御特性およびパラメータ誤差に対するロバスト性にどのような影響を及ぼすかを考察し、ビジュアルサーボに適したハンドアイシステムの構造ならびに画像特徴検出法の提案を行う。

本論文では、まず2章において、従来、非線形と考えられてきたステレオ画像とロボットの関節角度のマッピングが、人間の眼と腕に似せた運動学的構造を持つハンドアイシステムにおいては、時不変な線形マッピングとして近似できることを示す。2台の回転型カメラを備えた人間型のステレオカメラシステムにおいて、カメラ姿勢およびカメラ画像からデカルト座標系へのマッピング（3次元復元）は、多くのパラメータと回転行列を含む複雑なものである。しかし、同じ人間型のステレオカメラシステムにおけるカメラ姿勢およびカメラ画像から、カメラの注視動作における輻射角と視線方向角によって定義する視空間座標系へのマッピングは、単純な時不変線形変換として近似することができることを示す。また、回転関節型ロボットマニピュレータにおいて、デカルト座標系からロボットの関節角度へのマッピング（逆運動学）は、複雑な非線形代数方程式として表される。しかし、人間の眼と腕に似せた運動学的構造を持つハンドアイシステムにおいては、視空間座標系からロボットの関節角度へのマッピングが時不変線形変換として近似できることを示す。これらの性質を用いることにより、カメラ姿勢とカメラ画像からロボットの関節角度へのマッピングが、視空間座標系を介させることで、単純な時不変線形式として近似できることを示す。さらに、時不変線形式として最も精度良く近似が行えるカメラの配置とマニピュレータの運動学的構造は、人間の眼と腕の構造と似たものであることを示す。

3章においては、ハンドアイシステムのロボットの手先をターゲットへ位置決めするビジュアルサーボを行う場合に、2章で示した時不変線形近似マッピングを適用すると、ビ

ジュアルサーボのゲインを時不変とした制御が広範囲において適用可能であり、単純な構成式で計算量が少なく、カメラとロボットのキャリブレーション誤差に対してロバストなシステムとすることができることを示す。また、線形ビジュアルサーボと名付けたこの方法は、カメラ画像情報のみに基づいて制御を行うため、カメラ姿勢の計測は不要であり、制御動作中にカメラを回転させても、マニピュレータの動作にはほとんど影響がないことを示す。このため、本手法は、カメラを回転させて画像処理の高度化を目指すアクティブビジョンにも適用しやすいと考えられる。次に、ビジュアルサーボの軌道特性とキャリブレーション誤差および時間遅れの関係について述べ、カメラ配置とマニピュレータの運動学的構造が、この関係にどのような影響を与えるかを考察する。また、カメラ画像情報とロボット関節角度のマッピングのヤコビアンの時不変性が、キャリブレーション誤差や時間遅れに対してビジュアルサーボの軌道が受ける影響を和らげる性質を示すことに着目し、ビジュアルサーボの軌道安定性の指標として‘ヤコビアンの時不変安定度’を提案する。また、‘ヤコビアンの時不変安定度’を用いてハンドアイシステムの構造を評価すると、人間の眼と腕に似せた構造は広範囲においてヤコビアンの時不変性を有しており、時不変ゲインを用いたビジュアルサーボの軌道が広範囲において安定することをシミュレーションによって示す。

次に 4 章においては、ビジュアルサーボに適した画像特徴の検出法について考察を行う。近似的なモデルに基づく簡便な画像処理を用いても、ビジュアルサーボにおいては十分な制御性能が得られる一例として、粗いオプティカルフローを用いたビジュアルサーボを提案する。通常、オプティカルフローは精度良い結果を得るためには多くの計算量が必要であり、リアルタイムのフィードバック制御に用いられることは少ない。しかし、計算量の少ない大雑把なオプティカルフローの検出方法を用いても、ビジュアルサーボと併用することにより、高精度な並進と回転の位置決めを行うことができることを示す。

最後に 5 章で、本研究の結果をまとめ今後の課題を示す。

第 2 章 人間型ハンドアイシステムにおける線形近似逆運動学

2.1 緒言

カメラを用いてロボットマニピュレータを制御する場合、通常は、まずカメラによりカメラ座標系におけるマニピュレータの手先の位置と対象物の位置を計測する。次に、それらをマニピュレータの作業座標系に変換し、逆運動学やロボットヤコビアン（マニピュレータの手先速度と各関節の角速度を関係づけるヤコビ行列）を用いることにより、マニピュレータの各関節角を制御する。一般に、カメラ座標系や作業座標系としてはデカルト座標系がよく用いられており、マニピュレータの逆運動学は複雑な非線形代数方程式として表される。また、ロボットヤコビアンは、マニピュレータの各関節の角度を係数として含んでいるために、常に正確な関節角度を必要とする。したがって、組み立てなどの作業を実行するためには、あらかじめマニピュレータの正確なキャリブレーションが必要であり、複雑な非線形の逆運動学を解かなければならない。また、カメラ座標系と作業座標系の正確なキャリブレーションも必要である [11]。

これらの問題に対処する方法として、近年、ビジュアルサーボ [8, 12, 13, 14] が注目されている。この方法は、カメラで撮られた対象物イメージの特徴をダイレクトに利用してロボットの関節角コマンドを生成する方法であり、正確なキャリブレーションを必要としない点が最大の特徴である。ただし、安定なサーボ系を実現するには、関節角コマンドの計算に必要なカメラの画像座標系とロボットの関節座標系を結ぶヤコビアンの推定が必要である。このため、従来の単眼によるビジュアルサーボでは、モデルを利用してヤコビアンを計算したり、ロボットの駆動結果に基づいてヤコビアンを推定する方法 [15] が提案されている。また、ステレオ視を用いたビジュアルサーボでは、画像情報を直接用いてヤコビアンを計算する方法 [16, 17, 18] がとられている。いずれにしても、従来のこれらの方法では、カメラから得られる画像情報をロボットの関節コマンドに反映させるために、一旦デカルト座標に変換する操作が含まれる。

一方、生理学の分野においては、人間の空間知覚と上肢運動感覚の統合に関する研究が数多くなされている [19, 20, 21]。人間の腕をロボットアームと考えると、これは人間における逆運動学と考えることができる。人間における空間知覚系としては、生体と直接関係のないデカルト座標系ではなく、眼球運動や上肢運動にみられる生体本来のもつ情報形態に基づいた座標系が想定されている [22, 23]。カメラを用いてマニピュレータを制御する場合においても、ロボットとは直接関係のないデカルト座標系上の逆運動学を用いるよりも、カメラに基づく座標系を用いてカメラからマニピュレータへの直接的なマッピングを行うほうが、座標変換の手数も減り、単純なアルゴリズムにできると考えられる。

本章では、人間の眼と腕に似せた構造を持つ人間型ハンドアイシステムにおいて、カメラ角度と両眼視差に基づく視空間座標系を作業座標系として用いると、マニピュレータの逆運動学が線形式として近似できることを示す。この結果は、次章に示す時不変ゲインを用いたビジュアルサーボ方式を可能とするものである。本章では、まず、人間と似せた視覚運動系を持つハンドアイシステムにおけるカメラとロボットアームのモデルを示し(2.2節)、視空間(2.4節)と関節空間(2.3節)について説明する。次に、視空間と関節空間の線形性(2.5節)を示したうえで、視空間から関節空間への線形近似変換、すなわち、視空間を作業座標系とした線形近似逆運動学(2.6節)を構成し、その近似精度が十分実用可能であることを示す。また、線形近似の誤差を最小とするカメラ配置とロボットアームの運動学的構造が人間と類似したものであることを示す(2.7節)。

2.2 人間型ハンドアイシステムのモデル

Fig.2.1に、人間の運動学的構造を参考として作成したハンドアイシステムのモデルを示す。2台のカメラは、光学的焦点を中心として水平方向に独立に回転し(α_L, α_R)、垂直方向に同じ角度(α_D)回転する。カメラは基線長(Baseline length = $2E$)の間隔でベースフレームに搭載されている。また、ベースフレームはカメラの中心に位置する首関節(NECK)を中心に水平方向に回転する(ζ)。それぞれのカメラにおいては理想的な透視投影変換が成り立ち、空間上の一点は焦点(Focal point)を通して結像面に投影されるとする。画像座標とカメラ姿勢の関係はFig.2.2に示す。図中、 $(X^L, Y^L), (X^R, Y^R)$ は左右のカメラ画像座標をあらわす。なお、カメラの焦点距離(Focal length = f)は、左右等しく設定する。

つぎにロボットアームは、上腕(L_u)と前腕(L_f)の2リンクと肘関節(ELBOW)と肩関節(SHOULDER)の2関節で構成し、肘関節に水平方向に1自由度(j_1)、肩関節に水平方向(j_1)と垂直方向(j_0)の2自由度を有する。Tab.2.1に本モデルの各パラメータの値を示す。これらのパラメータは成人男子の実測値をもとに設定した。前腕の長さ(L_f)は肘関節から人差指の先までの長さを用いた。一般には2リンク型ロボットアームのリンク長さは等しく設定される場合が多いが、本モデルでは上腕よりも前腕がやや長いのが特徴である。2.7節で述べるが、これらのパラメータの値は視覚情報(視空間)からロボットアームの関節空間への変換を線形近似するうえで最適な値に近い。

Tab. 2.1: 人間型ハンドアイシステムの各パラメータ

Link length	$L_u = 250, L_f = 380$
Camera position	$W = 200, K = 200, G = 0$
Focal length	$f = 3.5$
Baseline length /2	$E = 35$ [mm]

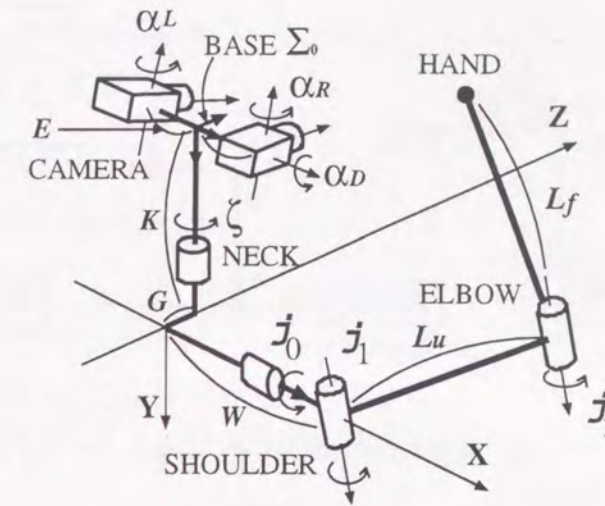


Fig. 2.1: 人間型ハンドアイシステムのモデル

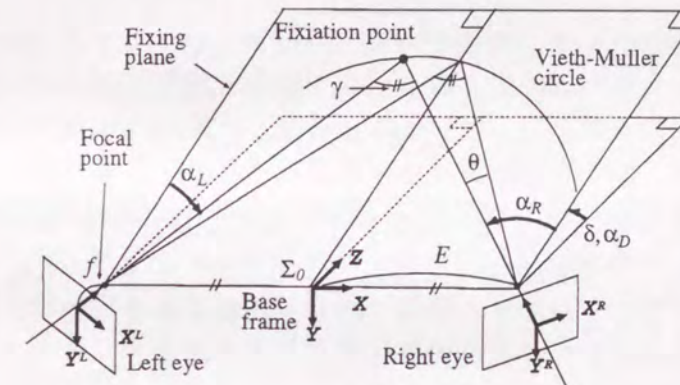


Fig. 2.2: ステレオカメラの幾何学

2.3 関節空間

Fig.2.1に定義するロボットアームの順運動学(関節空間からデカルト空間への変換)と逆運動学(デカルト空間から関節空間への変換)を次式に示す。順運動学と逆運動学は非線形式として表される。

順運動学

$$\begin{aligned}
 x &= L_u \cos(j_1) + L_f \cos(j_1 + j_2) + W \\
 y &= \bar{z} \sin(j_0) + K
 \end{aligned}$$

$$z = \tilde{z} \cos(j_0) - G$$

$$\text{ここで } \tilde{z} = L_u \sin(j_1) + L_f \sin(j_1 + j_2) \quad (2.1)$$

逆運動学

$$j_0 = \tan^{-1}(\hat{y}/\hat{z})$$

$$j_1 = \tan^{-1}(\hat{z}/\hat{x}) - \tan^{-1}(\sqrt{d}/\hat{x}^2 + \hat{z}^2 + L_u^2 - L_f^2)$$

$$j_2 = \tan^{-1}(\sqrt{d}/\hat{x}^2 + \hat{z}^2 - L_u^2 - L_f^2)$$

ただし,

$$\hat{x} = x - W$$

$$\hat{y} = y - K$$

$$\hat{z} = \sqrt{\hat{y}^2 + (z + G)^2}$$

$$d = (\hat{x}^2 + \hat{z}^2 + L_u^2 + L_f^2)^2 - 2\{(\hat{x}^2 + \hat{z}^2)^2 + L_u^4 + L_f^4\} \quad (2.2)$$

肩関節をとる水平面 ($j_0 = 0$) において, 関節空間をデカルト空間に投影した像を Fig.2.3 に示す. 図示した関節空間は $0 \leq j_1 \leq 100, 10 \leq j_2 \leq 110$ [deg] の領域であり, 人間が手先を楽に動作することができる領域である. 本図よりその形状は曲線的であり, 関節空間はデカルト空間に対して非線形性が強いことがわかる.

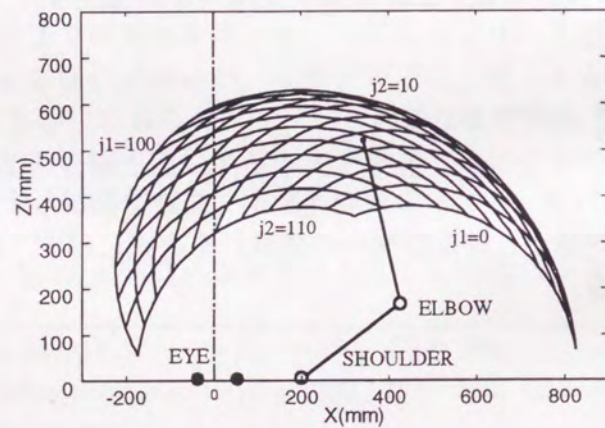


Fig. 2.3: デカルト空間に投影した関節空間

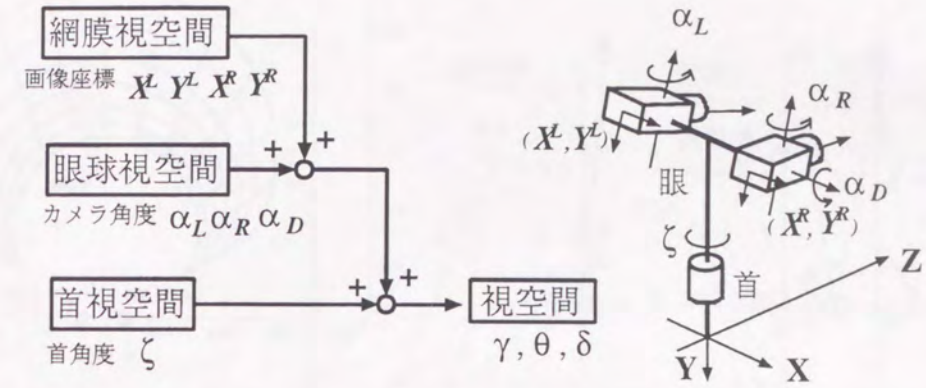


Fig. 2.4: 視空間の構成

2.4 視空間

人間型ハンドアイシステムでは, 首角度とカメラ角度とカメラの画像座標の3つの視覚情報を用いて空間上の3次元位置を計測する. これら3つのパラメータを用いて空間上の1点の3次元位置をデカルト空間として表現するためには, 多くの回転変換を含んだ煩雑な計算が必要である (付録 A 章参照). これに対して本研究では, Fig.2.4 に示すように, 首角度により定める首視空間と, カメラ角度により定める眼球視空間, 画像座標により定める網膜視空間の3つの視空間の線形和としてあらわす空間座標系 (視空間) によって空間上の位置を表現する. 視空間 $V = (\gamma, \theta, \delta)^T$, および他の3つの視空間は, (2.3) 式に示すように, 首角度とカメラ角度とカメラ画像座標の線形式として定義されるので, 単純で少ない計算量で3次元位置を得ることができる.

$$V = \begin{bmatrix} \gamma(\text{奥行き}) \\ \theta(\text{水平方向}) \\ \delta(\text{垂直方向}) \end{bmatrix} = \begin{bmatrix} 0 \\ \zeta \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha_L - \alpha_R \\ (\alpha_L + \alpha_R)/2 \\ \alpha_D \end{bmatrix} + \begin{bmatrix} (X^L - X^R)/f \\ (X^L + X^R)/2f \\ (Y^L + Y^R)/2f \end{bmatrix} \quad (2.3)$$

$$\boxed{\text{視空間}} = \boxed{\text{首視空間}} + \boxed{\text{眼球視空間}} + \boxed{\text{網膜視空間}}$$

以下, それぞれの視空間が線形和としてどのように結合されるかを説明する.

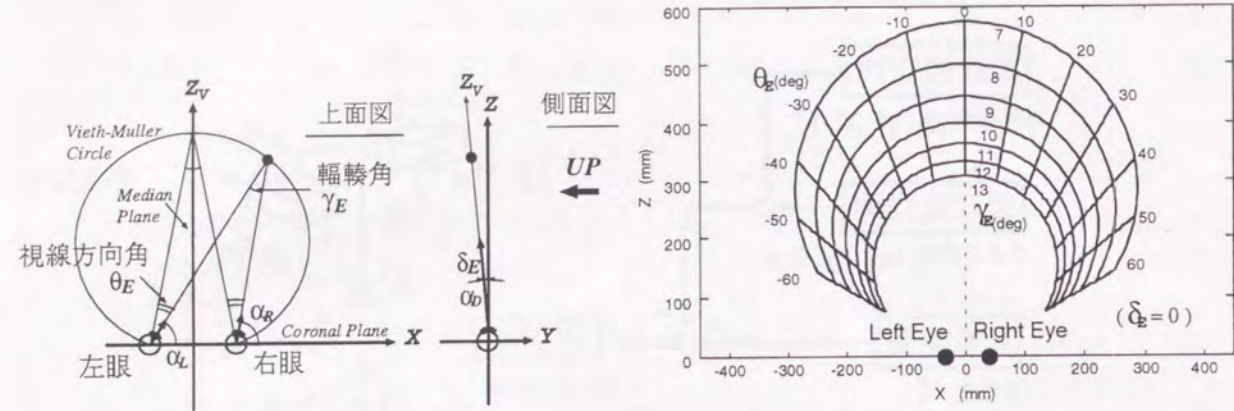


Fig. 2.5: 眼球視空間

2.4.1 眼球視空間

眼球視空間は1点を注視したときの輻輳角 γ_E と視線方向角 θ_E, δ_E を用いて定義される座標系 (Fig.2.5参照) で, 人間の視覚的な空間知覚モデルとして, 生理学, 心理学において用いられている [24]. 次式に示すように, 輻輳角と視線方向角は左右のカメラ角度の差と和より得ることができる.

$$\mathbf{V}_E = \begin{bmatrix} \gamma_E (\text{輻輳角}) \\ \theta_E (\text{水平方向角}) \\ \delta_E (\text{垂直方向角}) \end{bmatrix} = \begin{bmatrix} \alpha_L - \alpha_R \\ (\alpha_L + \alpha_R)/2 \\ \alpha_D \end{bmatrix} \quad (2.4)$$

首関節を正面方向 $\zeta = 0$ に固定するとき, 眼球視空間は次式によりデカルト空間へ変換される.

$$\begin{aligned} x &= E \sin(2\theta_E) / \sin(\gamma_E) \\ y &= \bar{z} \sin(\delta_E), \quad z = \bar{z} \cos(\delta_E) \\ \text{ここで } \bar{z} &= E \{ \cos(\gamma_E) + \cos(2\theta_E) \} / \sin(\gamma_E) \end{aligned} \quad (2.5)$$

カメラをとおりる水平面 ($\delta_E = 0$) において, 眼球視空間をデカルト空間に投影した像を Fig.2.6に示す. その形状は極座標に似ており, 眼球視空間からデカルト空間への変換は, 明らかに非線形なものであることがわかる.

Fig. 2.6: デカルト空間に投影した眼球視空間

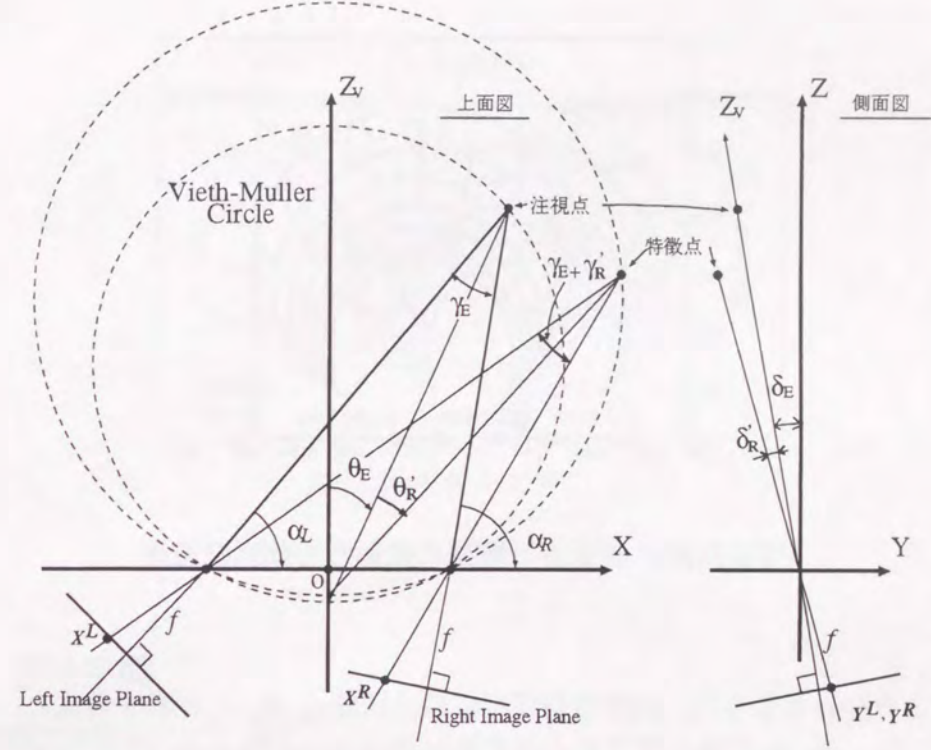
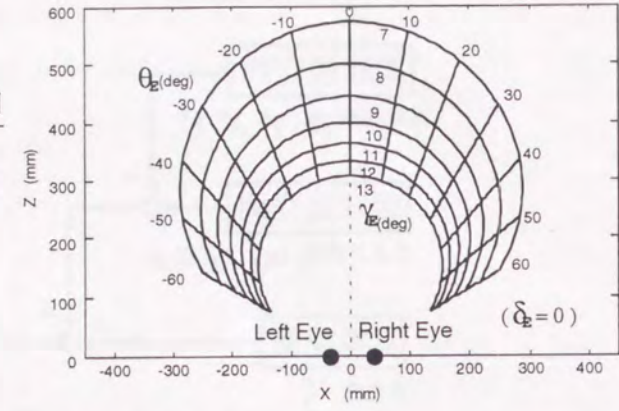


Fig. 2.7: 網膜視空間

2.4.2 網膜視空間

左右のカメラ画像に投影された特徴点の画像座標 X^L, X^R, Y^L, Y^R を用いると, 注視点に対する特徴点の3次元的な相対位置を, 次式により得ることができる.

$$\mathbf{V}_R = \begin{bmatrix} \gamma_R (\text{視差, 奥行}) \\ \theta_R (\text{水平方向角}) \\ \delta_R (\text{垂直方向角}) \end{bmatrix} = \begin{bmatrix} (X^L - X^R)/f \\ (X^L + X^R)/2f \\ (Y^L + Y^R)/2f \end{bmatrix} \quad (2.6)$$

γ_R は, 左右のカメラ画像に投影された特徴点の水平座標の差であり, カメラからの奥行きを表している. γ_R は一般に視差と呼ばれる. θ_R は, 特徴点の水平座標の平均であり, カメラからの水平方向を表している. また, δ_R は特徴点の垂直座標の平均であり, カメラからの垂直方向を表している. 本研究ではこの座標系を網膜視空間と呼ぶ. 網膜視空間の原点は注視点に一致する.

網膜視空間座標 $\gamma_R, \theta_R, \delta_R$ は, 次式に定義する近似値 $\gamma'_R, \theta'_R, \delta'_R$ を用いると, 幾何学的には Fig.2.7のように表すことができる.

$$\begin{aligned} \gamma'_R &= \tan^{-1}((X^L - X^R)/f) \simeq \gamma_R \\ \theta'_R &= \tan^{-1}((X^L + X^R)/2f) \simeq \theta_R \\ \delta'_R &= \tan^{-1}((Y^L + Y^R)/2f) \simeq \delta_R \end{aligned} \quad (2.7)$$

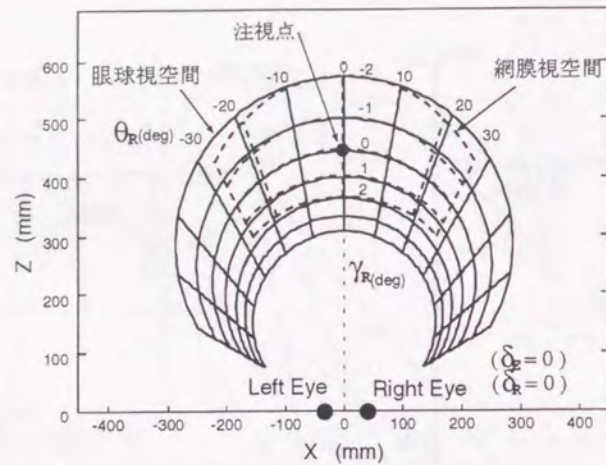


Fig. 2.8: デカルト空間に投影した網膜視空間

Fig.2.7より分かるように、網膜視空間座標の近似値 $\gamma'_R, \theta'_R, \delta'_R$ が表す角度と、注視点の眼球視空間座標 $\gamma_E, \theta_E, \delta_E$ が表す角度のそれぞれの和 $\gamma'_R + \gamma_E, \theta'_R + \theta_E, \delta'_R + \delta_E$ は、特徴点を眼球視空間座標として表すときの角度にほぼ一致している。よって、網膜視空間座標の近似値 $\gamma'_R, \theta'_R, \delta'_R$ が網膜視空間座標 $\gamma_R, \theta_R, \delta_R$ とほぼ等しいと考え、次式に示すように、カメラ角度より得る眼球視空間座標 V_E と画像座標より得る網膜視空間座標 V_R の線形和によって、特徴点の位置を眼球視空間を用いて‘大まかに’定めることができる。本座標系 V_{ER} を、本研究では眼球網膜視空間と呼ぶ。

$$V_{ER} = V_E + V_R = \begin{bmatrix} \alpha_L - \alpha_R \\ (\alpha_L + \alpha_R)/2 \\ \alpha_D \end{bmatrix} + \begin{bmatrix} (X^L - X^R)/f \\ (X^L + X^R)/2f \\ (Y^L + Y^R)/2f \end{bmatrix} \quad (2.8)$$

本座標系は‘大まかに’しか空間上の位置を定めていない点だが、通常の座標系とは異なる。すなわち、網膜視空間座標の近似値 $\gamma'_R, \theta'_R, \delta'_R$ と網膜視空間座標 $\gamma_R, \theta_R, \delta_R$ の誤差により、(2.8)式で得られる V_{ER} で定まる眼球網膜視空間と(2.5)式より定まる眼球視空間には、誤差が存在する (Fig.2.8に、図中央の黒点を注視している際の網膜視空間の形状を点線、眼球視空間の形状を実線で示す)。また、 V_E と V_R は実空間に対して冗長であるため、 V_{ER} の値が求まっても、 V_E と V_R の値がそれぞれ定まらなければ実空間の位置が一意には定まらない。このため、 V_{ER} より得られる特徴点の位置は、‘大まかに’しか定められない。しかし、デカルト座標系を用いて空間上の位置を表すためには、付録A章に示すような煩雑な非線形式を計算しなければならない。これに対して、本座標系では、‘大まか’ではあるが、非常に単純な線形式の計算だけで位置を記述できる点がポイントである。

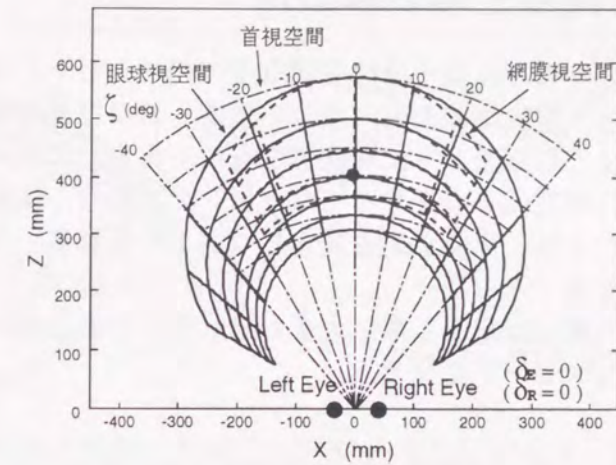


Fig. 2.9: デカルト空間に投影した首視空間

2.4.3 首視空間

首視空間の定義を次式に示す。

$$V_N = \begin{bmatrix} 0 \\ \zeta \text{ (首方向角)} \\ 0 \end{bmatrix}$$

首視空間は首の角度で眼からの方向をあらわす座標系である。Fig.2.9に首視空間と眼球視空間と網膜視空間を重ねてプロットしたものを示す。首視空間の方向角は、眼球視空間、網膜視空間の方向角とほぼ一致している。また、奥行きも注視点近傍 (図中央の黒点) においてはほぼ一致している。よって、空間上の一点がカメラに投影されたときのカメラ角度と画像座標と首角度が得られると、(2.3)式を用いることで首視空間、眼球視空間、網膜視空間の線形和として点の位置を‘大まかに’定めることができる。

2.5 視空間と関節空間の線形性

一般に視覚情報からマニピュレータの関節角度へのマッピングはデカルト空間を介在させて行われる。しかし、視覚情報からデカルト空間、およびデカルト空間から関節空間へのマッピングは前節までに見たように非線形なものである。これに対して本研究では、デカルト空間の代わりに視空間を介在させてマッピングを行う。視覚情報から視空間への変換は、前節で説明したように単純な線形式としてあらわされる。次に視空間から関節空間へのマッピングを考える。

視空間は空間的に一意には定まらない大まかな座標系であるが、ここでは次の2通りの視空間を用いてそのマッピングを考える。

- 視空間A
首は正面に固定し、物体を注視したときのカメラ角度より得られる視空間。このとき $\zeta = 0, X^L = X^R = Y^L = Y^R = 0$ となるので、(2.3)式は

$$\mathbf{V} = \begin{bmatrix} \alpha_L - \alpha_R \\ (\alpha_L + \alpha_R)/2 \\ \alpha_D \end{bmatrix} \quad (2.9)$$

となり、Fig.2.6の形状となる。すなわち首関節を正面に固定させた場合の眼球視空間と一致する。

- 視空間B
首を物体が正面となるように回転したうえで注視を行い、そのときの首角度とカメラ角度で得られる視空間。このとき、 $\alpha_L + \alpha_R = 0, X^L = X^R = Y^L = Y^R = 0$ となるので、(2.3)式は

$$\mathbf{V} = \begin{bmatrix} 0 \\ \zeta \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha_L - \alpha_R \\ 0 \\ \alpha_D \end{bmatrix} \quad (2.10)$$

となる。その形状は Fig.2.9の一点鎖線と一致する。

視空間Aはカメラの回転可動範囲により有効視野が制限されるので正面の比較的狭い領域のマッピングに有効であると考えられ、視空間Bは広範囲を認識する場合に有効であると考えられる（人間が眼球動作だけで情報探索が可能な水平方向の範囲は、頭部前方 ± 15 [deg]であり、頭部運動（首回転）を用いて注視動作が安定に行える水平方向の範囲は、頭部前方 ± 45 [deg]である [25]）。Fig.2.10に示した関節空間 $0 \leq j_1 \leq 100, 10 \leq j_2 \leq 110$ [deg] をこれら2つの視空間に投影した結果を Fig.2.12, Fig.2.14に示す。また、これよりやや狭い関節空間 $20 \leq j_1 \leq 60, 60 \leq j_2 \leq 100$ [deg] (Fig.2.11) を視空間に投影した結果を Fig.2.13, Fig.2.15に示す。これらの図より視空間上における関節空間の形状は平行四辺形に近く、関節空間と視空間のマッピングは線形的であることがわかる。つまり視空間と関節空間の変換は線形近似可能であることが予想される。

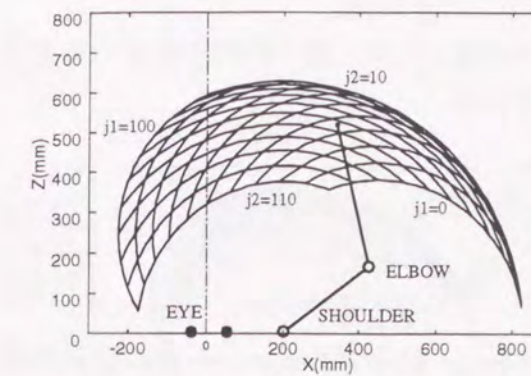


Fig. 2.10: デカルト空間に投影した関節空間（広領域）

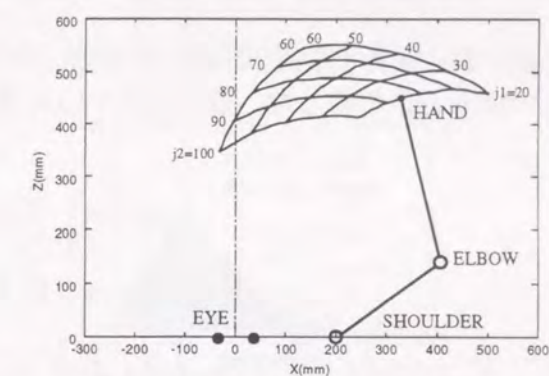


Fig. 2.11: デカルト空間に投影した関節空間（狭領域）

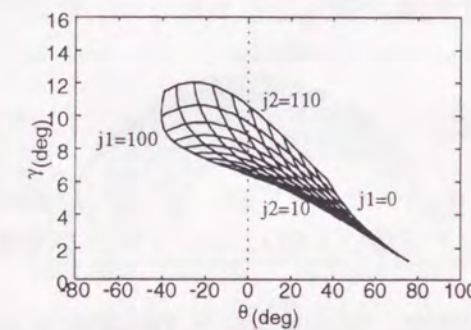


Fig. 2.12: 視空間Aに投影した関節空間（広領域）

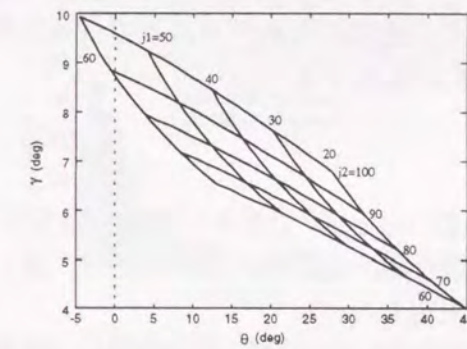


Fig. 2.13: 視空間Aに投影した関節空間（狭領域）

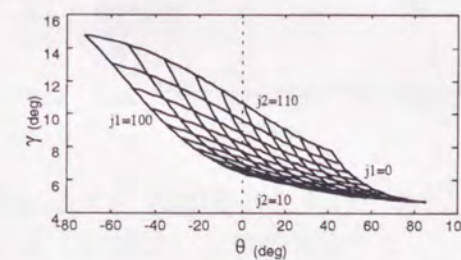


Fig. 2.14: 視空間Bに投影した関節空間（広領域）

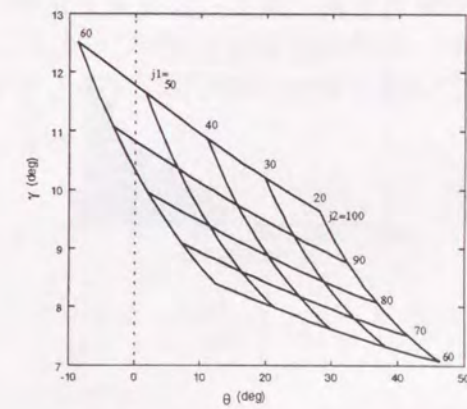


Fig. 2.15: 視空間Bに投影した関節空間（狭領域）

2.6 視空間と関節空間の線形近似

(2.11) 式に定義する近似誤差の自乗和 (SSD) を評価値として, 最小自乗法を用いて逆運動学の線形近似を行った. 近似式を (2.12) 式に定義する.

$$SSD = \sum_{j_1, j_2} (j'_1 - j_1)^T (j'_1 - j_1) \quad (2.11)$$

$$j' = R V + C \quad (2.12)$$

ここで $V = (\gamma, \theta)^T$, $j = (j_1, j_2)^T$, $j' = (j'_1, j'_2)^T$ であり, V はマニピュレータの関節角度が j のときの手先の視空間座標, j' は近似関節角度, R, C は線形近似の定数行列, 定数ベクトルである. 線形近似は, ロボットアームの運動を肩関節をとる水平面上 ($y = 200$) に制限し, 先に示した 2 種類の関節空間 (Fig. 2.10 (広領域), Fig. 2.11 (狭領域)) を用いて行った. R, C は前節と同じく (2.11) 式に示す SSD についての最小自乗法により決定した. また, 比較のためデカルト空間から関節空間への変換の線形近似も行った. 近似式を (2.13) 式に定義する.

$$j' = R_D D + C_D \quad (2.13)$$

ここで $D = (x, z)^T$ であり, D はマニピュレータの手先のデカルト座標, R_D, C_D は線形近似の定数行列, 定数ベクトルである. R_D, C_D は R, C と同様に SSD についての最小自乗法により決定した.

広領域を用いた場合の近似結果と, 誤差分布, 近似係数, SSD を Fig. 2.16 ~ Fig. 2.18 に示す. また, 狭領域を用いた場合の結果を同様に, Fig. 2.19 ~ Fig. 2.21 に示す. 近似結果の図において, 目標とする手先位置は実線で示し, 線形近似により得られた手先位置は点線で示す (広領域の近似においては, 外枠のみ太い実線で示す). 誤差分布より, 視空間 A, 視空間 B を用いると, デカルト空間に比べて精度の良い線形近似が行えることがわかる. なお, 本近似結果は 3.4 節で示すように, 3.3 節で提案するビジュアルサーボに対して十分に広い漸近安定領域を保証するものである.

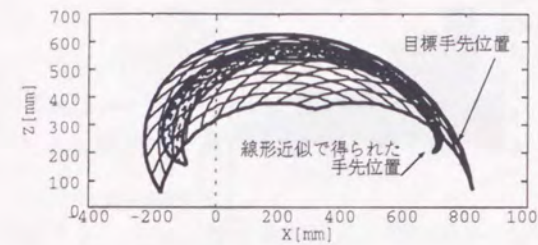
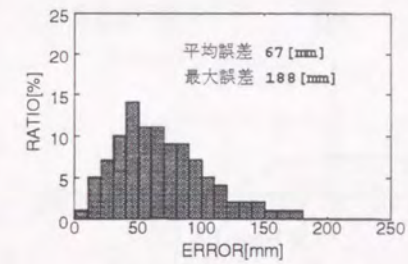


Fig. 2.16: デカルト空間を用いた線形近似結果 (広領域) と誤差分布



$$R_D = \begin{pmatrix} -0.00141 & 0.00172 \\ -0.00110 & -0.00349 \end{pmatrix}, C_D = \begin{bmatrix} 0.446 \\ 3.08 \end{bmatrix}$$

$$SSD = 26.7[\text{rad}^2]$$

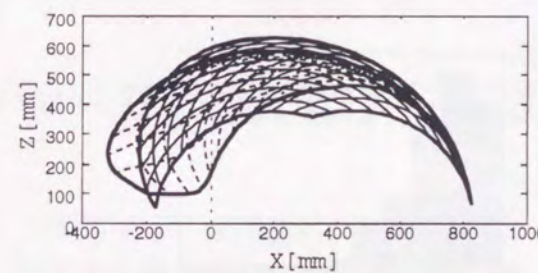
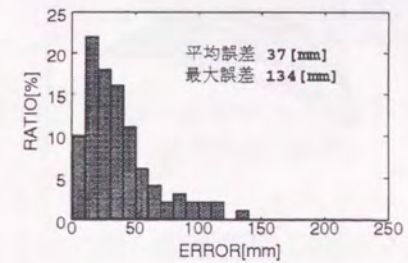


Fig. 2.17: 視空間 A を用いた線形近似結果 (広領域) と誤差分布



$$R = \begin{pmatrix} -10.3 & -1.78 \\ 21.0 & -1.18 \end{pmatrix}, C = \begin{bmatrix} 2.60 \\ -1.60 \end{bmatrix}$$

$$SSD = 14.9[\text{rad}^2]$$

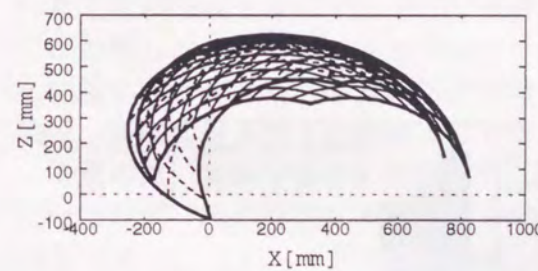
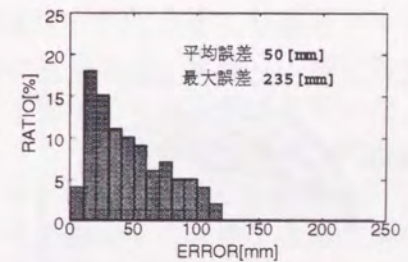


Fig. 2.18: 視空間 B を用いた線形近似結果 (広領域) と誤差分布



$$R = \begin{pmatrix} -11.5 & -1.47 \\ 18.8 & -0.608 \end{pmatrix}, C = \begin{bmatrix} 2.83 \\ -1.53 \end{bmatrix}$$

$$SSD = 13.9[\text{rad}^2]$$

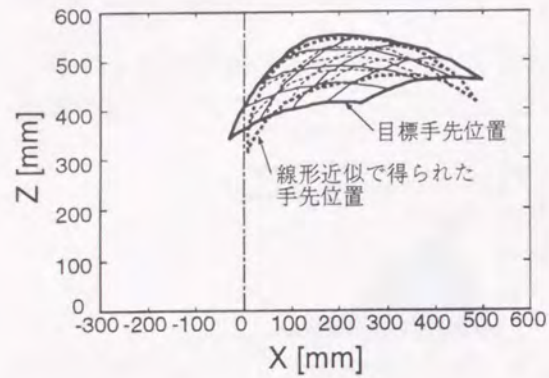


Fig. 2.19: デカルト空間を用いた線形近似結果 (狭領域) と誤差分布

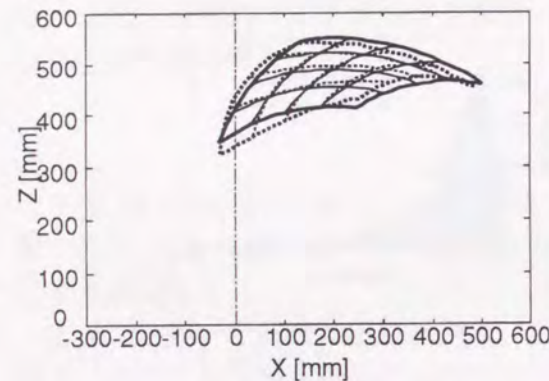
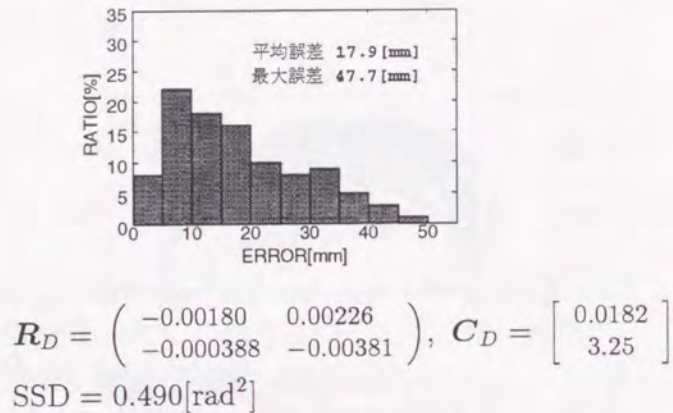


Fig. 2.20: 視空間Aを用いた線形近似結果 (狭領域) と誤差分布

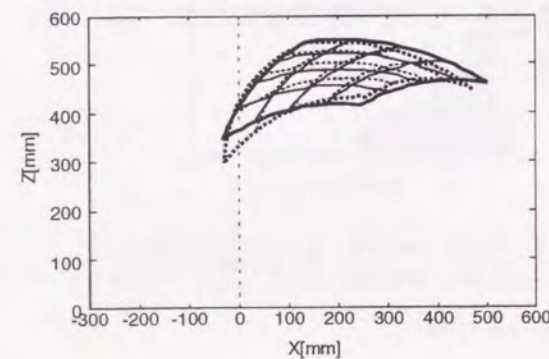
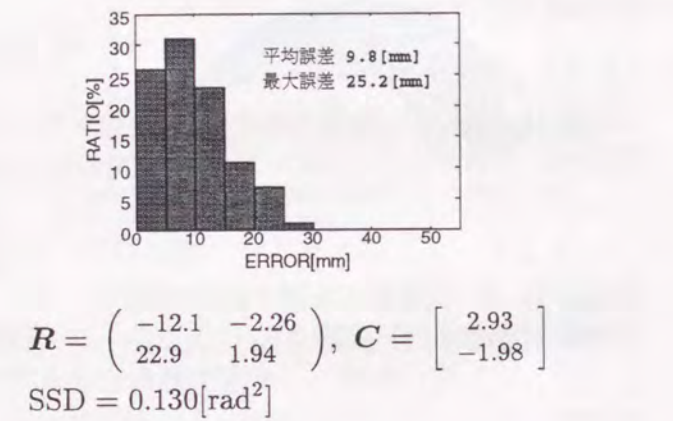
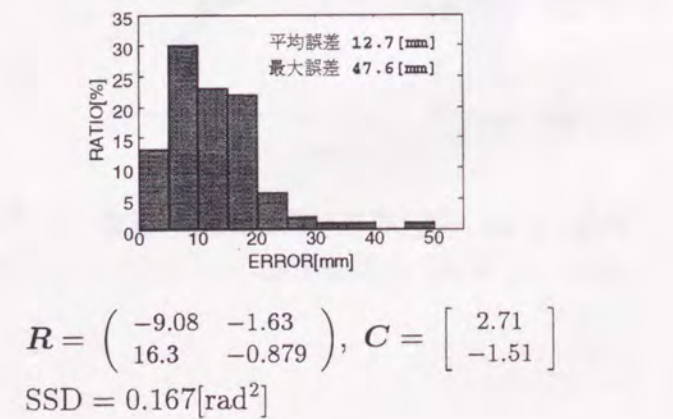


Fig. 2.21: 視空間Bを用いた線形近似結果 (狭領域) と誤差分布



2.7 線形近似に適したカメラ配置とマニピュレータの運動学的構造

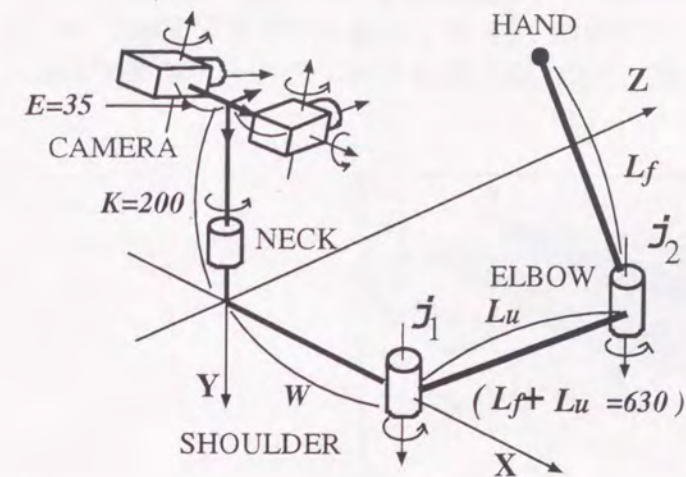


Fig. 2.22: 人間型ハンドアイシステムの構造評価

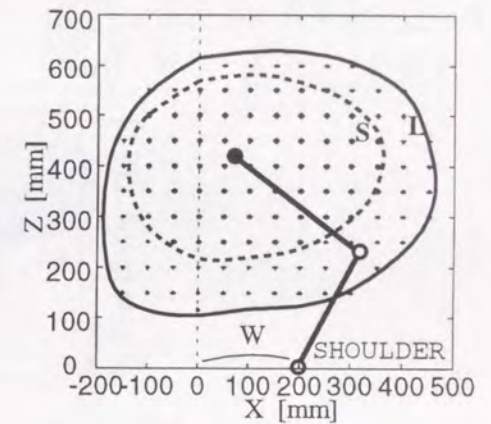


Fig. 2.23: 人間の手先の可動領域 (近似対象領域)

本節では、人間型ハンドアイシステムにおける視空間と関節空間の線形マッピングの近似誤差を最小化するカメラ配置とマニピュレータの運動学的構造の評価を行う。構造評価においては、Fig.2.22に示すように、カメラの肩関節に対する高さ ($K = 200$) とマニピュレータのリンク長さの和 ($L_u + L_f = 630$) を固定したうえで、カメラと肩関節の水平距離 (W) とマニピュレータの前腕と上腕の長さの比率 ($ratio = L_u / (L_u + L_f)$) について評価を行った。カメラの基線長は $E = 35$ に固定しているが、 E は視空間のスケールを定める定数に相当するので ((2.5) 式参照)、基線長を変化させても線形マッピングの精度には影響しない。また、本節ではマッピングの対象領域を、実際に人間が肩を通る水平面上で腕を動かすときの手先の可動領域と、楽に手先が動かせる領域を用いた。これらの領域は、実際に成人男子の被験者に手先を動かしてもらうことで計測し、Fig.2.23に示すように定めた。実線で囲まれた領域が実験により得られた手先の可動領域 (L)、点線で囲まれた領域が、腕の筋肉に不自然さを感じずに手先を動かすことができる作業のし易い領域 (S) である。これらの領域内に 50[mm] 間隔で格子状に設定した点を対象に、視空間と関節空間の線形マッピングを行った。線形マッピングの近似パラメータは、前節と同様、関節角度の近似誤差の自乗和を評価量とした最小自乗法により定め、カメラ配置とマニピュレータの運動学的構造の評価には、線形マッピングにより得られた手先位置のデカルト空間における誤差の自乗和 ($C-SSD$) を用いた。

$$C-SSD = \sum_{S \text{ or } L} (D' - D)^T (D' - D) \tag{2.14}$$

ここで、 D はマニピュレータの手先のデカルト座標、 D' は線形近似により得られたマニピュレータの手先のデカルト座標である。

まず、アームのリンク長さの比を人間と同じ値 ($ratio = 0.4$, $L_u = 250$, $L_f = 380$) に固定したうえで、カメラと肩関節の距離 W を変化させながら線形近似マッピングを行ったときの W と $C-SSD$ の関係を調べた。マッピングの対象領域を手先の動かし易い領域 (S) とした場合の結果を Fig.2.24 に、手先の可動領域 (L) とした場合の結果を Fig.2.25 に示す。図中の実線は視空間 A を用いた場合、点線は視空間 B を用いた場合の結果である。

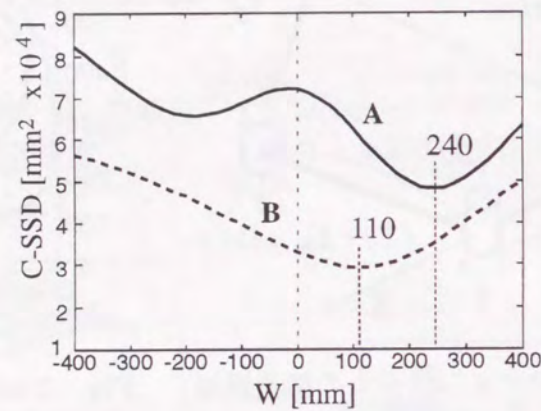


Fig. 2.24: 眼球位置と線形近似精度 (対象領域 S)

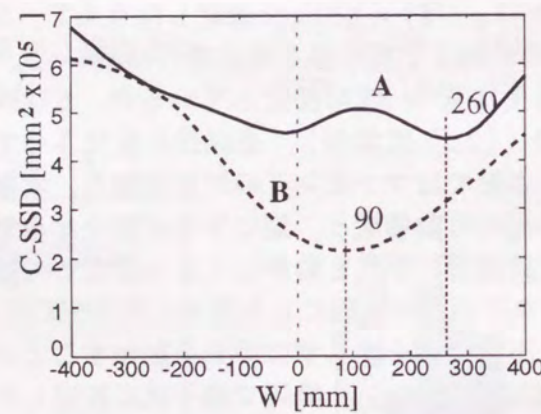


Fig. 2.25: 眼球位置と線形近似精度 (対象領域 L)

いずれの場合も、近似誤差を最も小さくする W の値は $W > 0$ であり、カメラとマニピュレータの肩関節はやや水平距離をおいて設置したほうが、線形近似の精度は良くなることわかった。

次にカメラと肩関節の水平距離 W とリンクの前腕と上腕の長さの比 $ratio$ を同時に変えながら線形近似マッピングを行ったときの W , $ratio$ と $C-SSD$ の関係を調べた。視空間 A を用いて領域 L に対してマッピングを行ったときの結果を Fig.2.26 に、視空間 A を用いて領域 S に対してマッピングを行ったときの結果を Fig.2.27 に示す。マッピングの対

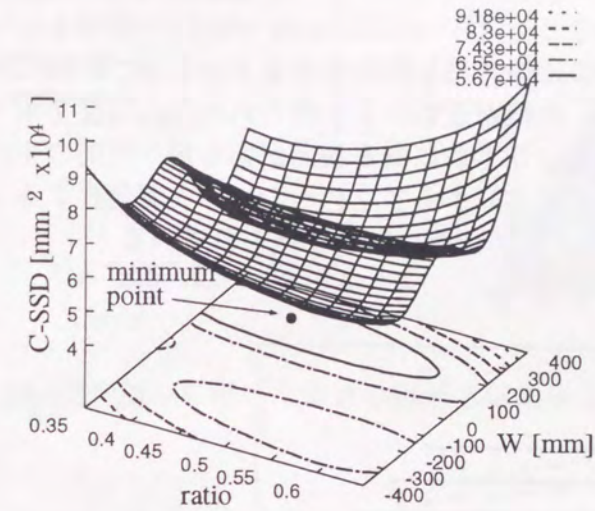


Fig. 2.26: 前腕、上腕の長さの比と線形近似精度 (対象領域 S)

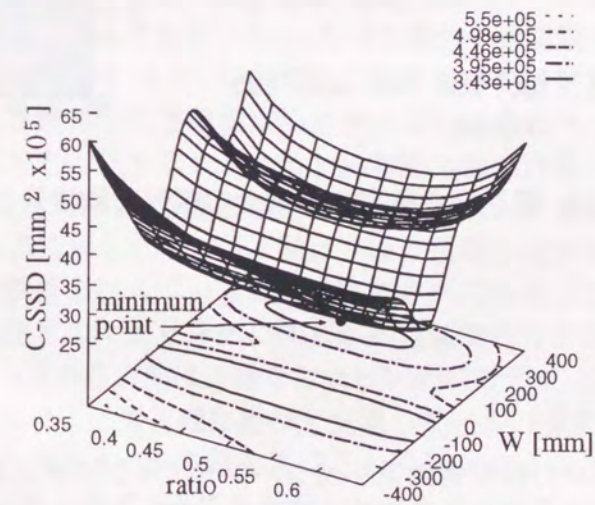


Fig. 2.27: 前腕、上腕の長さの比と線形近似精度 (対象領域 L)

象領域を領域 L としたときは、 $ratio = 0.43$, $W = 240$ で最も近似誤差が少ないことがわ

かる。この値は人間 ($ratio = 0.4, W = 200$) とほぼ一致している。また、領域 S を対象領域とした場合も、 $ratio = 0.47, W = 290$ に最小値が存在しており、この値も人間と似たものである。なお、視覚性空間として視空間 B を用いた場合には、近似誤差は腕の長さの比 $ratio$ に対してほとんど変化しなかった。以上より、視空間とマニピュレータの関節角度のマッピングは、カメラ配置とマニピュレータの運動学的構造が、人間と似たものであるときに、線形的となることがわかった。

次に構造を人間と同じとし ($W = 200, ratio = 0.4$)、領域 S について線形マッピングを行ったときの手先位置の近似結果と誤差分布を Fig. 2.28, 2.29 に示す。図中、実線は目標とする手先位置であり、点線が近似により得られた手先位置を示す。これらの図では、近似された位置が目標位置より手前にある傾向があるが、人間の手のリーチング動作の実験によっても、手の到達位置に関するアンダーシュートが発生することが確かめられており [19]、生体・生理工学的にも興味深い結果が得られている。

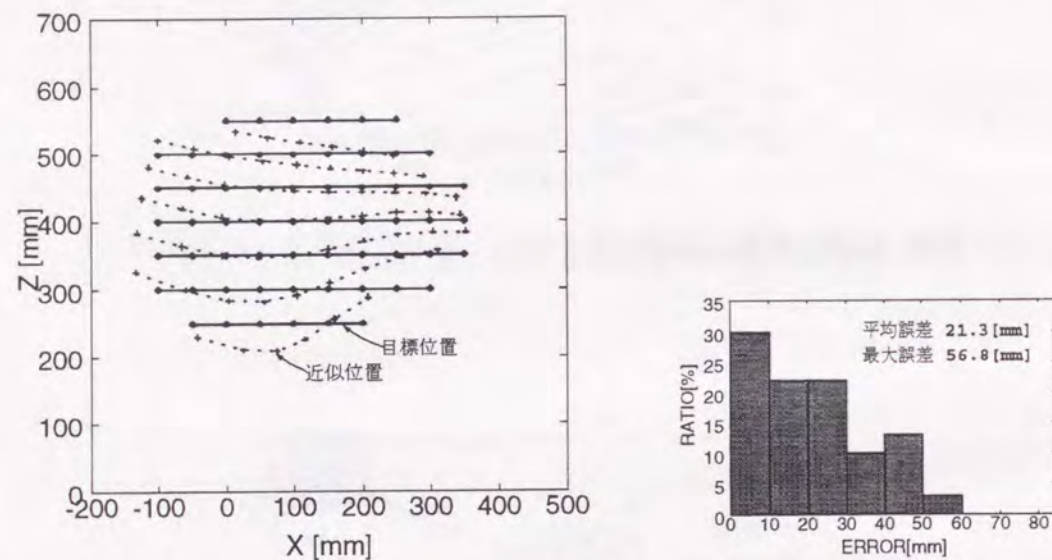


Fig. 2.28: 視空間Aを用いた場合の線形近似結果と誤差分布

2.8 結言

本節では、人間と似せた視覚運動系を持つハンドアイシステムにおいて、カメラ角度と両眼視差に基づく視空間座標系を作業座標系として用いると、ロボットアームの逆運動学が線形式として近似できることを示した。また、線形近似に適したカメラ配置とロボットアームの構造が人間と類似したものであることを示した。

通常、カメラを用いたロボットの制御においては、デカルト空間における幾何学的拘束に基づいた座標変換 (ロボットアームの運動学、ステレオカメラによる三角測量など) が

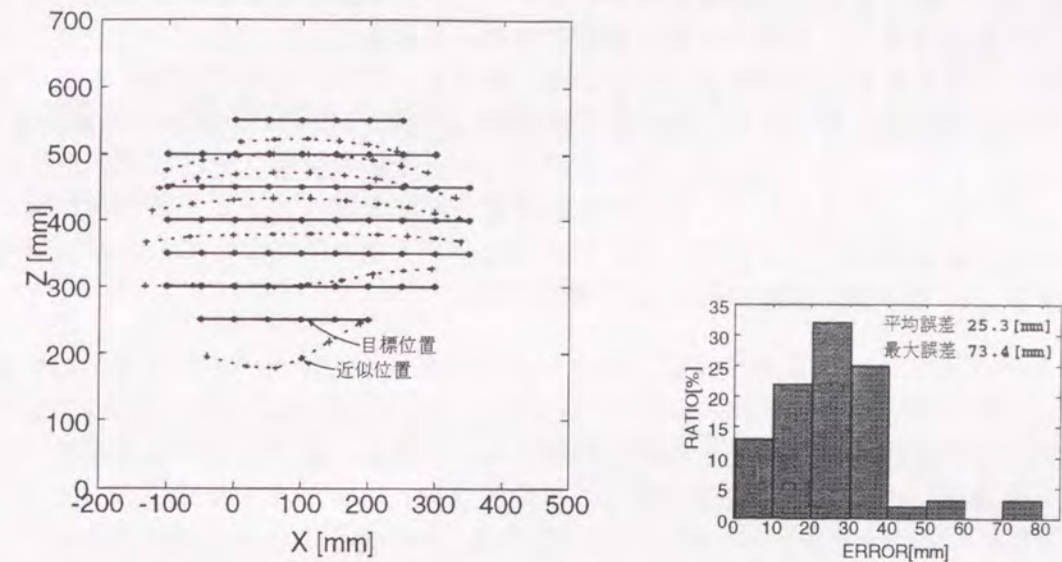


Fig. 2.29: 視空間Bを用いた場合の線形近似結果と誤差分布

基礎となっている。一部には、ニューラルネットワークを用いた直接的なマッピングを利用した研究 [26, 27] も見られるが、その数は少ない。これに対して、本研究では、変換誤差が生じることを前提として、単純な時不変線形マッピングを用いる点が従来にない新しい概念である¹。変換誤差が存在することを前提とすることは、動作の正確さを追求する産業用ロボットの研究からみると逆行する技術であるかもしれない。しかし、現在、多くのロボット研究者が研究を行っている、いわゆる‘知能ロボット’においては、‘環境、ロボットの形状、タスクの変化に対応できる’といった『柔軟性』が重要視されている。また、柔軟性を持ったロボットとしては‘人間’が参考とされる場合が多いが、果して‘人間’は産業用ロボットが行うような幾何学的な座標変換を行っているのだろうか？まず、人間のような複雑な構造をすべてモデル化することはほとんど不可能である。また、本研究で行ったような視覚情報より得た空間的な位置と手先位置のマッピングを生体・生理工学的に検証した研究 [19, 21, 29] は多くみられるが、人間はさほど正確にはマッピングを行っておらず、逆にある一定の傾向の誤差を持っていることが確認されている。この事実は、人間の手先位置の感覚と視覚的空間とのマッピングに何らかの近似的なモデルが用いられている可能性を示唆していないだろうか？人間は正確なマッピングのモデルは持っていないが、フィードバックを利用することによって正確な動作を実現している。『柔軟性』を持つ‘知能ロボット’においても、多くのパラメータの値を常に正しく同定することは非常に困難であること、また、マッピングが不正確であっても外部センサを用いたフィードバック制御により正

¹本研究では、カメラの画像座標およびカメラ角度と、マニピュレータの関節角度のマッピング (ハンドアイマッピング) を取り扱っているが、デカルト座標とマニピュレータの関節角度のマッピング (逆運動学) については、指令値 (マニピュレータの手先位置) の多項式近似によりマニピュレータの関節角度を求める手法が新井により提案されている [28]。

確な動作が可能であることを考えあわせると、必ずしも厳密なモデルに基づいたマッピングを行う必要はなく、人間と同様に近似的なモデルに基づいたマッピングとフィードバック制御とを併用することが妥当ではないかと考える。また、この近似モデルとしては、なるべく単純で直接的なマッピングであるほうが、学習に必要なパラメータも少なく済み、次節で示すように、サーボ動作におけるパラメータ変動の影響を受けにくいシステムとすることができると考えられる。本研究では視覚と関節角度のマッピングのみを取り扱っているが、力覚と関節トルクのマッピングなど他のマッピングにおいても、センサの配置を工夫することで同様の議論が行える可能性がある。

第 3 章 視空間を用いたビジュアルサーボ

3.1 緒言

本章では、カメラとマニピュレータを人間と似せたカメラ配置とマニピュレータの運動学的構造にすれば、視空間座標系におけるマニピュレータの手先位置と関節角度が、線形変換により近似結合できる性質に着目し、視空間座標系がカメラ画像上の位置と直結した座標系であることを利用することで、定数行列による線形変換という極めて単純なビジュアルサーボが構成可能であることを示す。線形ビジュアルサーボと名付けたこの制御方法は、時不変線形式に基づくためアルゴリズムも非常に単純であり計算量も非常に少ない。また、従来のビジュアルサーボではマニピュレータの関節角度をリアルタイムに検出して、常に適切なゲインを計算する必要があったが、本サーボの構成式には関節角度は含まれず、カメラ画像情報のみに基づく制御が可能である。このため、マニピュレータのキャリブレーションエラーに対してもロバストである。さらに、従来のカメラを用いた制御ではカメラ姿勢の検出が必要であったが、本サーボの構成式にはカメラ姿勢も含まれないため、制御動作中にカメラ姿勢が変化しても、サーボ動作に及ぼす影響は非常に小さい。このため、人間の眼のように回転するカメラを用いたハンドアイシステムを用いた場合でも、カメラの運動を考慮することなくマニピュレータを動作させることができ、カメラのキャリブレーションエラーに対してもロバストである。キャリブレーションにおいても、従来はカメラとマニピュレータの各パラメータを一つずつ修正する必要があり、手間がかかる上に、破損等による大きなモデルの変化にはまったく対応することができなかった。これに対し、本方法では、カメラ座標系からマニピュレータ座標系への変換を用いず、カメラ画像上の位置とマニピュレータの関節角度を直接関係づける近似線形式に基づいているため、カメラとマニピュレータを一まとめにしたグローバルなキャリブレーションが可能である。さらに本方法は、画像座標と直結した三次元座標（視空間座標系）と関節座標系との線形結合にも基づいているため、画像特徴を状態量とする特徴ベースのビジュアルサーボでありながらも、軌道の予測は容易であり、位置ベースの利点も併せ持つものである。

本章ではまず、3.2節において、これまでに提案されているビジュアルサーボについて説明を行い、本章で提案する線形ビジュアルサーボとの比較を行う。次に 3.3節において線形ビジュアルサーボの構成式を示し、その特徴を述べる。3.4節においては、本サーボ方式の漸近安定性をリアプノフの定理を用いて解析する。つぎに 3.5節において静止ターゲットへの位置決め動作の実験により、線形ビジュアルサーボの持つカメラ回転に対するロバスト性の確認を行う。3.6節においては、モデル誤差に対するロバスト性をシミュレーション実験によって解析する。さらに、3.7節においては、カメラの配置とマニピュレータの運動学的構造が及ぼすビジュアルサーボの制御特性への影響について考察しその評価を行う。

3.2 ビジュアルサーボ

本節では、これまでに提案されたビジュアルサーボについて説明を行い、本章で提案する線形ビジュアルサーボとの比較を行う。ビジュアルサーボについての一般的な説明は、1.1.2節を参照されたい。

これまでに提案されたビジュアルサーボは、大きく位置ベース法 [30, 31, 32] と特徴ベース法 [33, 34, 35] に分けることができる (Fig.3.1 (左))。

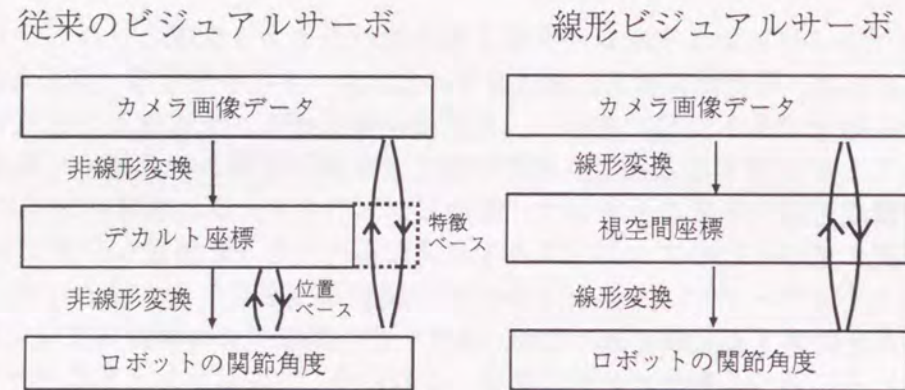


Fig. 3.1: 従来のビジュアルサーボとの比較

位置ベース法は、カメラ画像に基づいて推定した物体とマニピュレータの手先（またはカメラ）の相対的な3次元位置（デカルト座標）を状態量としてフィードバックを行うもので、これまでに数多く研究されてきたカメラ画像から3次元復元を行う研究（ロボットビジョン）や、デカルト座標系におけるロボット制御の研究（ロボティクス）の豊富な資産を活用できる利点がある。しかし、カメラ画像からデカルト座標への3次元復元を安定に行うためには多くの計算量が必要であり、実時間制御が要求されるビジュアルサーボにおいては、画像ノイズの影響を受けやすい問題がある。そこで、カルマンフィルターを用いることによってノイズの影響を少なくした手法などが提案されている [32]。位置ベース法は、デカルト座標系における物体とマニピュレータの手先（またはカメラ）の相対的な位置関係を制御入力として、ロボット関節の運動とデカルト空間における変位を関係づけるヤコビ行列（ロボットヤコビアン）を用いて、ロボットの関節への指令値を次式のように生成する。

$$\begin{aligned} \mathbf{u} &= -\lambda \mathbf{J}_R^{-1}(\boldsymbol{\theta}_R) (\mathbf{p} - \mathbf{p}_d) \\ \mathbf{J}_R &= \frac{\partial \mathbf{p}}{\partial \boldsymbol{\theta}_R} \end{aligned} \quad (3.1)$$

ただし、 \mathbf{u} はロボットの関節への指令値、 λ はゲイン、 \mathbf{J}_R はロボットヤコビアン、 $\boldsymbol{\theta}_R$ はロボットの関節角度、 \mathbf{p} はロボットの手先（またはカメラ）のデカルト座標、 \mathbf{p}_d は目標物体のデカルト座標である。デカルト座標からロボットの関節角度への変換は非線形式で表

されるので、ロボットヤコビアンは、ロボットの関節角をパラメータとして含んだ時変な行列である。このため、安定した制御を行うためには、常に正確なロボットヤコビアンを推定しなければならない。

特徴ベース法は、カメラ画像から抽出した特徴量（特徴の重心位置、面積、エッジの長さなど）を状態量としてフィードバックを行うもので、3次元位置の推定を行わないため、前者と比べて計算量が少なく、カメラのキャリブレーションエラーに対してもロバストである。特徴ベース法は、次式に示すように、ロボットの関節の運動と画像から抽出した特徴量の変化を関係づけるヤコビ行列 (\mathbf{J}_{RI}) を用いて、カメラの画像情報からロボットの関節への指令値を直接的に生成する。

$$\begin{aligned} \mathbf{u} &= -\lambda \mathbf{J}_{RI}^{-1}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_C) (\mathbf{I} - \mathbf{I}_d) \\ \mathbf{J}_{RI} &= \frac{\partial \mathbf{I}}{\partial \boldsymbol{\theta}_R} \end{aligned} \quad (3.2)$$

ここで \mathbf{I} は画像特徴量、 \mathbf{I}_d は目標とする画像特徴量、 $\boldsymbol{\theta}_C$ はカメラ姿勢をあらわす。ここで、ヤコビアン \mathbf{J}_{RI} は次式に示すように、ロボットヤコビアン \mathbf{J}_R とデカルト空間における運動と画像中の特徴量の変化を関係づけるヤコビ行列（イメージヤコビアン \mathbf{J}_I ）を合成したものと考えることができる。

$$\begin{aligned} \mathbf{J}_{RI} &= \frac{\partial \mathbf{I}}{\partial \boldsymbol{\theta}_R} \\ &= \frac{\partial \mathbf{I}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial \boldsymbol{\theta}_R} \\ &= \mathbf{J}_I \mathbf{J}_R \\ \mathbf{J}_I &= \frac{\partial \mathbf{I}}{\partial \mathbf{p}} \end{aligned} \quad (3.3)$$

このため、従来提案されている特徴ベース法は、間接的にデカルト座標系を介在したマッピングに基づいた制御を行っているものがほとんどである。たとえば、Feddemraらはロボットヤコビアンとイメージヤコビアンの逆行列を用いて次式で制御を行った [40]。

$$\mathbf{u} = -\lambda \mathbf{J}_R^{-1} \mathbf{J}_I^{-1} (\mathbf{I} - \mathbf{I}_d) \quad (3.4)$$

また、橋本らは冗長な画像特徴量に対して一般化逆行列を用いて次式で制御を行った [34]。

$$\mathbf{u} = -\lambda \mathbf{J}_R^{-1} \mathbf{J}_I^+ (\mathbf{I} - \mathbf{I}_d) \quad (3.5)$$

ここで、 \mathbf{J}_I^+ は \mathbf{J}_I の一般化逆行列である。これらの手法では、ロボットヤコビアンに加えてイメージヤコビアンも実時間で推定しなければならない。また、Espiauらは目標点におけるイメージヤコビアンを用いて次式によって制御を行った [36]。

$$\mathbf{u} = -\lambda \mathbf{J}_R^{-1} \mathbf{J}_{Id}^+ (\mathbf{I} - \mathbf{I}_d) \quad (3.6)$$

ここで \mathbf{J}_{Id}^+ は、目標点におけるイメージヤコビアン \mathbf{J}_{Id} の一般化逆行列である。この手法では、制御のはじめに目標点におけるイメージヤコビアンを推定すれば、制御中はイメー

ジャコビアンを推定する必要はない。しかし、収束が保証される領域は目標点近傍に限定されている。また、橋本らは同様に次式によって制御をおこなった [38]。

$$\mathbf{u} = -\lambda \mathbf{K} (\mathbf{J}_{Id} \mathbf{J}_{Rd})^T (\mathbf{I} - \mathbf{I}_d) \quad (3.7)$$

ここで \mathbf{J}_{Rd} は、目標点におけるロボットヤコビアンであり、 \mathbf{K} はゲイン行列である。この手法では制御のはじめに目標点におけるイメージヤコビアンとロボットヤコビアンを推定すれば、制御中はこれらのヤコビアンを一切推定する必要はない。しかし、同様に収束が保証される領域は目標点近傍に限定されている。

さて、これまでの手法において、イメージヤコビアンやロボットヤコビアンを実時間で推定しなければならない、または、収束範囲が目標点近傍に限定されるといった問題は、イメージヤコビアンとロボットヤコビアンが時変であることに起因している。つまり、カメラの画像より得た画像特徴量とデカルト座標およびロボットの関節角度の変換が非線形であるために、これらの変換のヤコビアンが時変となり、これらの問題を引き起こしている。これに対して本研究では、カメラとマニピュレータを人間と似たカメラ配置とマニピュレータの運動学的構造にすれば、視空間座標系におけるマニピュレータの手先位置と関節角度が、線形変換により近似結合できる性質を利用して、次節で示すように時不変ゲインを用いたビジュアルサーボを提案する (Fig.3.1 (右))。 (3.7) 式に示した手法は、同じく時不変ゲインを用いたビジュアルサーボであるが、本研究で提案する手法はカメラ配置とマニピュレータの運動学的構造を工夫することにより、ほぼ可動領域の全域で収束漸近安定性が保証されること、また、視空間を介在することによりカメラが回転するステレオカメラを用いても、時不変ゲインを用いて制御が行える点が特徴である。

3.3 線形ビジュアルサーボ

本節では、前章で示した人間型ハンドアイシステムにおける視空間とマニピュレータの関節角度の線形性を利用して、時不変ゲインを用いたビジュアルサーボを提案する。

Fig.2.1に示した人間型ハンドアイシステムにおいて、マニピュレータの手先とターゲットが同時に左右のカメラで観測されているとき、手先の視空間座標とターゲットの視空間座標の差より関節角速度を生成するフィードバック制御を考えると、(2.12) 式を用いて次式のフィードバック式が構成できる。

$$\mathbf{u} = -\lambda \mathbf{R} (\mathbf{V} - \mathbf{V}_d) \quad (3.8)$$

\mathbf{u} : 関節速度指令値
 \mathbf{V} : ハンドの視空間座標
 \mathbf{V}_d : 目標点の視空間座標
 λ : ゲイン

ここで、 \mathbf{R} は前章で示した最適な線形近似を得る定数行列のパラメータである。手先とターゲットを同時に観測しているときには、手先とターゲットの視空間座標 \mathbf{V}, \mathbf{V}_d に含まれる

カメラ角度と首角度は等しいので、(2.3) 式を (3.8) 式に代入すると、カメラ角度 $\alpha_L, \alpha_R, \alpha_D$ と首角度 ζ の項が消去され、カメラ画像座標のみに依存する次のサーボ式が得られる。

$$\begin{aligned} \mathbf{u} &= -\lambda \mathbf{R} \begin{bmatrix} \{(X^L - X^R) - (X_d^L - X_d^R)\}/f \\ \{(X^L + X^R) - (X_d^L + X_d^R)\}/2f \\ \{(Y^L + Y^R) - (Y_d^L + Y_d^R)\}/2f \end{bmatrix} \\ &= -\lambda \mathbf{R} \mathbf{T} (\mathbf{I} - \mathbf{I}_d) \\ \text{ここで } \mathbf{T} &= \begin{pmatrix} 1/f & -1/f & 0 & 0 \\ 1/2f & 1/2f & 0 & 0 \\ 0 & 0 & 1/2f & 1/2f \end{pmatrix} \\ \mathbf{I} &= (X^L, X^R, Y^L, Y^R)^T \end{aligned} \quad (3.9)$$

ここで、 \mathbf{I} は手先の画像座標であり、 \mathbf{I}_d はターゲットの画像座標である。以下、このビジュアルサーボを線形ビジュアルサーボと呼ぶ。一般にステレオカメラを用いたビジュアルサーボのゲインは時変であるが、人間型ハンドアイシステムにおいて視空間を用いると、このように時不変のゲインを用いたビジュアルサーボが構成できる。線形ビジュアルサーボの特徴を以下に示す。

- カメラ姿勢角、首角度、マニピュレータの関節角の計測が不要である。
 カメラ角度、首角度、マニピュレータの関節角の計測が不要であり、これらのキャリブレーションエラーに対してロバストである。また、制御入力として用いる手先とターゲットの視空間座標の差は、カメラ姿勢や首姿勢が変化してもほぼ一定値として計測されるので、サーボ中にカメラや首関節が回転しても、サーボ動作に及ぼす影響が非常に小さいシステムである。よって、アクティブビジョンを用いた画像処理システムとの整合性も良いと考えられる。
- 単純な制御則である。
 計算量が少ない。特徴ベースのビジュアルサーボでありながら、視空間座標系に基づくサーボ方式でもあるため、ハンドの軌道は、画像空間に比べて3次元的に認識の容易な視空間上で推定することができる。
- 柔軟でグローバルなキャリブレーションが可能である。
 カメラとマニピュレータのパラメータ (焦点距離、リンク長さなど) を用いない直接的なキャリブレーションが可能である。本手法におけるキャリブレーションとは、カメラ画像とマニピュレータの関節角度をマッピングする行列 \mathbf{R} を定めることを意味する。キャリブレーションを行う領域を限定すると、作業領域に適した行列 \mathbf{R} を設定することが可能であるなど、柔軟なキャリブレーションが可能である。

3.4 漸近安定性

本節では、前節で提案したビジュアルサーボの収束安定性を検証する。ここで、マニピュレータは関節速度指令値どおりに理想的に動作するものとする。 V_d を静止ターゲットとしたとき、誤差 $e_V = V - V_d$ を定義すると、誤差システムは次式で表される。

$$\dot{e}_V = M(j)\dot{j} \quad (3.10)$$

ここで、 $M(j)$ は、関節空間から視空間へのヤコビアンである。線形ビジュアルサーボにおいて、関節速度は (3.8) 式で与えられる。このとき、閉ループ系のダイナミクスは次式であらわされる。

$$\dot{e}_V = -\lambda M(j)R e_V \quad (3.11)$$

以下、線形近似が行われた領域内 (Fig.2.11 参照) の初期位置から同領域内に設定された目標点に収束することを検証する。ここで、つぎのスカラー関数を定義する。

$$U(V) = \frac{1}{2} e_V^T e_V \quad (3.12)$$

上式は、(3.11) 式のシステムにおいて連続である。上式を微分すると次式を得る。

$$\frac{d}{dt} U(V) = -\lambda e_V^T M(j)R e_V \quad (3.13)$$

つぎに Ω を次式で定義される閉領域とする。

$$U(V) < U_0 \quad (3.14)$$

ここで、 U_0 は、正の定数である。もし、閉領域 Ω に含まれる目標点 V_d 以外の全ての V について、 $\frac{d}{dt} U(V) < 0$ が成り立つなら、不変集合定理 [37] より、 Ω を初期位置とする軌道はすべて目標点に収束することが保証される。ここで、閉領域 Ω は、吸引領域と呼ばれ、数値的に定義することができる。Fig.3.2 と Fig.3.3 は、線形近似を行った領域周辺における吸引領域を示す。

ここでは、状態ベクトル V を 2次元とし、肩関節をとる水平面 ($Y = 200$) 上において解析を行った。 $(V = (\gamma, \theta))$ 、また R は、Fig.2.20 の 2×2 の行列とする。図中の線は U_0 の等高線をあらわす。斜線でおおわれた領域は、マニピュレータの可動範囲とその境界領域における傾きを考慮して求めた吸引領域を示す。点でおおわれた領域は、 $U(V)$ の時間微分値が負で、吸引領域に含まれると推定される領域を示す。本図より、吸引領域は、線形近似が行われた領域に対して十分広い領域であることがわかる。

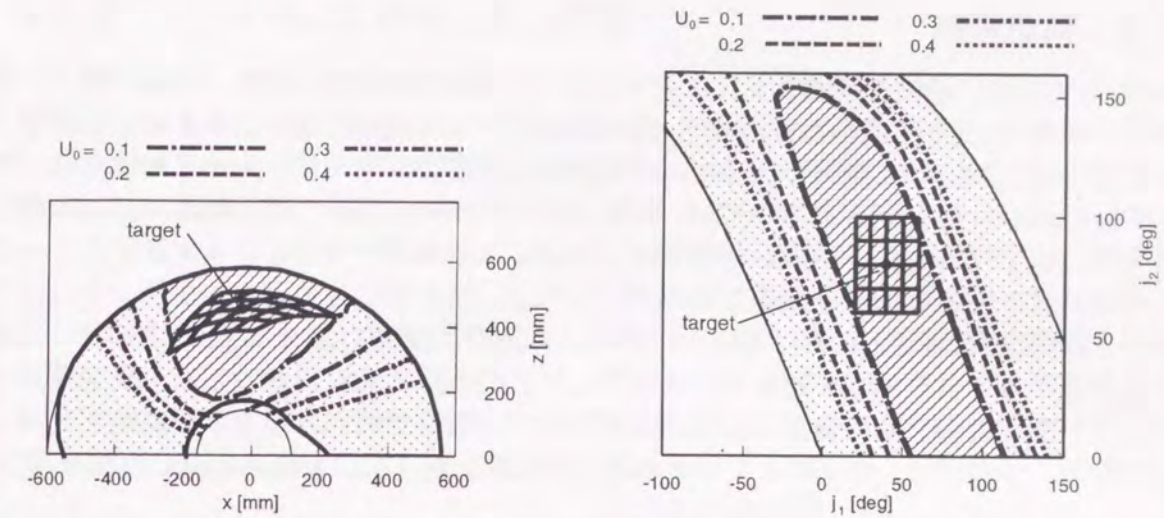


Fig. 3.2: デカルト空間における吸引領域

Fig. 3.3: 関節空間における吸引領域

3.5 実験

本節では、実験とシミュレーションにより線形ビジュアルサーボの有効性を確認する。まず、従来のステレオカメラを用いたビジュアルサーボとの比較として、疑似逆行列を用いるステレオビジュアルサーボ [35] (付録 B 章参照) との比較実験を行う。以下、線形ビジュアルサーボを L-VS、疑似逆行列を用いるビジュアルサーボを C-VS として記述する。つぎに、線形ビジュアルサーボがカメラ回転にロバストであることを実験により検証する。

3.5.1 実験設備

Fig.3.4 に実験設備の概略図を示す。ここで用いるステレオビジョンシステムのサンプリングレートは 30Hz である。実験においてシステムは大雑把に設定しており、キャリブレーションは行っていない。また、マニピュレータの動作は前節と同じく水平面上に限定して行った。

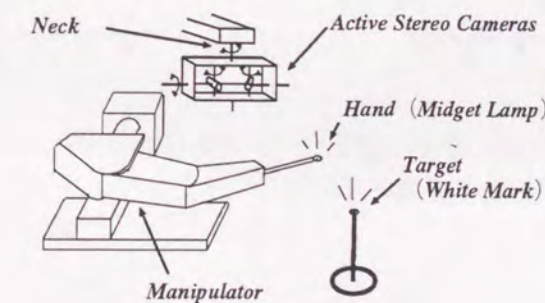


Fig. 3.4: 実験設備

3.5.2 軌道特性

L-VSとC-VSの軌道を実験とシミュレーションにより比較した。まず、初期位置と目標位置をそれぞれ $(x, y)^T = (-100, 400)^T$, $(300, 500)^T$ として実験を行い、つぎに初期位置と目標位置を逆に設定して実験を行った。また、同じ条件でシミュレーションを行った。結果を Fig. 3.5 に示す。L-VS においては、シミュレーションと実験で得た軌道は、ほぼ同じであるが、C-VS では大きく異なっている。これは、L-VS がキャリブレーションエラーに対してロバストであることを示唆している。

また、L-VS は視空間上におけるフィードバック制御であるために、視空間上において直線軌道を描こうとする傾向があるが、人間型ハンドアイシステムにおいては、視空間とマニピュレータの関節空間は線形的な関係にあるので、関節空間においても直線軌道を描く傾向がある。このため、マニピュレータの非可動領域を避けた収束軌道を描く特性を持つ。

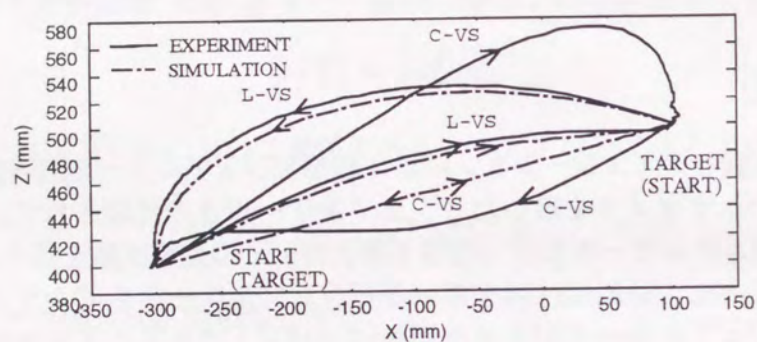


Fig. 3.5: 線形ビジュアルサーボと従来の手法の比較

3.5.3 カメラ角度に依存しない性質

L-VS はカメラ角度の計測が不要でありながら、カメラ回転の影響が非常に小さい。このことを検証するために、カメラを手で約 1Hz の周期で水平方向に約 ± 10 [deg] 回転させながら実験を行い、カメラを固定して実験を行った場合の軌道との比較を行った。Fig. 3.6 は、実験より得られたハンドの軌跡を示す。軌道はどちらもほぼ等しく、L-VS においては、サーボ中のカメラ回転が空間軌道にほとんど影響を及ぼさないことがわかる。Fig. 3.7 は、左カメラ画像上の特徴点の時間変化を示す。軌道の周期的な変動は、カメラの回転をあらわしている。Fig. 3.8 は、画像上におけるハンドと目標点の偏差の時間変化を示す。偏差は滑らかに収束しており、L-VS においては、サーボ中のカメラ回転が時間軌道に対してほとんど影響を及ぼさないことがわかる。ここで制御中にカメラの角度は一切用いていないことに注意されたい。

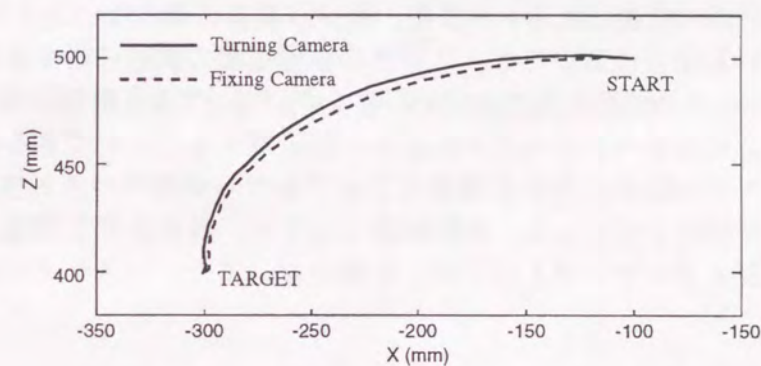


Fig. 3.6: カメラ固定時と回転時の手先軌道

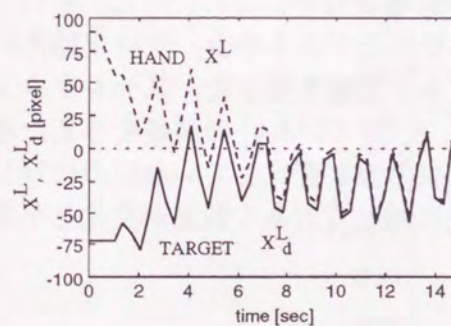
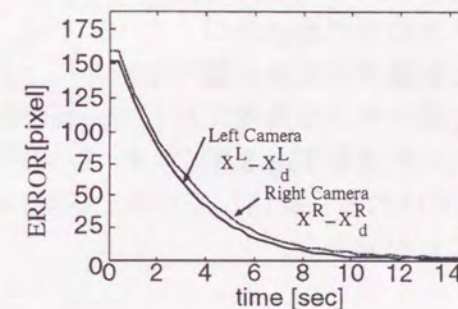
Fig. 3.7: カメラ回転時の画像座標 X^L の軌道

Fig. 3.8: カメラ回転時の視空間における誤差の軌道

3.6 モデル誤差に対するロバスト性

本章では、線形ビジュアルサーボのモデル誤差に対するロバスト性をシミュレーションにより検証する。従来のビジュアルサーボ [30, 31, 32, 33, 34, 35] では、カメラ姿勢とマニピュレータの関節角度は、正確に計測されているという前提に基づいていた。そのため、これらの計測誤差はサーボ動作に悪影響を及ぼすと考えられる。これに対し、線形ビジュアルサーボでは、構成式にこれらのパラメータが含まれないため、計測自体が必要ない。

また、従来のビジュアルサーボは、カメラは固定されたものと仮定しており、このことがカメラのキャリブレーションエラーに対するロバスト性を保証していた。これに対し、線形ビジュアルサーボはカメラ姿勢に依らないシステムであるため、カメラが回転する場合でも、同等のロバスト性が保証される。現在、研究されている次世代ビジョンシステムの多くが、回転可能なカメラシステムを用いており、実環境におけるカメラの回転を利用した画像処理の有効性が報告されている [41, 42] ことを考えると、カメラの回転角を用いないビジュアルサーボの意義は大きい。また、カメラ姿勢に依らないシステムであることは、固定カメラを用いる場合においても、カメラの設置姿勢の誤差に対するロバスト性を保証する。マニピュレータの関節角度の計測が不要な点についても、計測誤差の影響を受けないだけでなく、マニピュレータのモデル誤差を一部吸収することができるという利点がある。

以上のようなモデル誤差に対する線形ビジュアルサーボのロバスト性について、シミュレーションにより検証を行なった。比較対象としては、前章までと同様、『疑似逆行列を用いたステレオビジュアルサーボ』(C-VS)を用いた。また、システムは前節と同じモデルを用いた。

3.6.1 カメラの姿勢誤差による影響

L-VS では、カメラ姿勢が変化しても、サーボ動作には影響がない。これに対し、平行ステレオカメラのモデルに基づく C-VS では、カメラの姿勢誤差がある場合、サーボ動作にどのような影響があるかを調べるために、シミュレーション実験を行った。ステレオカメラのうち左側のカメラ角度に対して、反時計まわりに 2, 5, 8, 10 [deg] の誤差を与えた場合に得られた軌道を Fig.3.9 に示す。L-VS では、軌道がほとんど変化していないのに対して、C-VS では大きく変化している。C-VS はカメラ誤差に対して大きく軌道が変化し不安定となることが分かる。

3.6.2 マニピュレータの関節角度誤差による影響

L-VS の構成式にはマニピュレータの関節角度が含まれないため、関節角度の計測誤差による影響は受けない。C-VS において、関節 j_1 の計測値が、実際より 5, 10, 20 [deg] の誤差 (Δj) があるとしてシミュレーションを行った結果を Fig.3.10 に示す。C-VS では、マニピュレータの関節角度誤差とともに軌道が変化することがわかる。

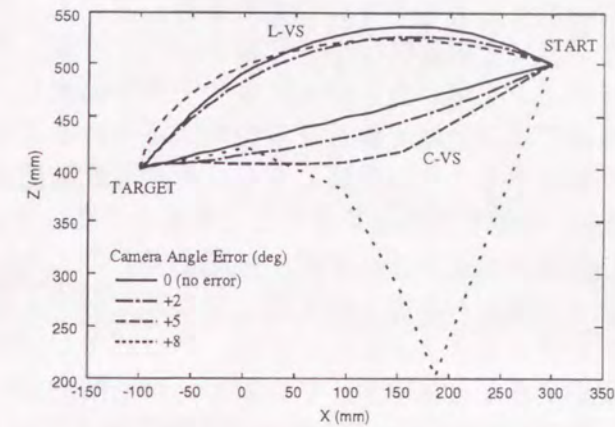


Fig. 3.9: カメラの姿勢誤差による影響

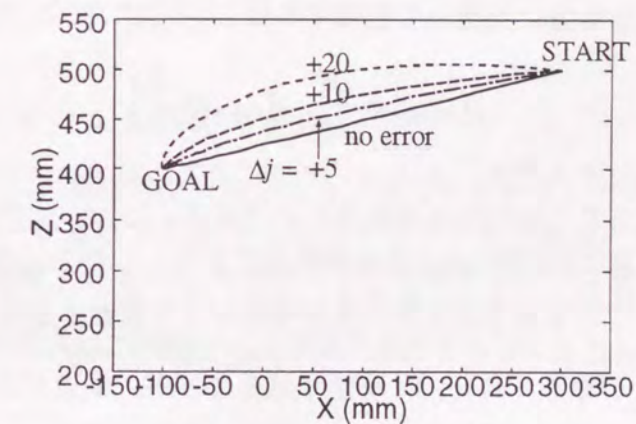


Fig. 3.10: マニピュレータの関節角度誤差による影響 (C-VS の場合)

3.7 ビジュアルサーボに適したカメラ配置とマニピュレータの運動学的構造

本節ではカメラを用いたマニピュレータの運動制御において、カメラの配置とマニピュレータの運動学的構造が及ぼす制御特性への影響について考察しその評価を行う。

前節までは、時不変ゲインを用いたビジュアルサーボを提案し、人間型ハンドアイシステムにおいては可動領域のほぼ全域で漸近安定性が保証されることを示した。しかし、漸近安定領域は、軌道が収束することは保証しても、その収束軌道が適切であることは保証していない。最悪の場合、渦巻状の軌道を描くこともありうる。そこで、本節では、ヤコビアンを時不変定数としたビジュアルサーボが、空間上でどのような収束軌道を描くかを評価し、直線的な収束軌道が得られるカメラ配置とマニピュレータ構造について考察を行なう。また、人間型ハンドアイシステムにおいては、ヤコビアンが広い範囲で空間的に一様となり、収束軌道も安定することを示す。ヤコビアンが空間的に一様であるカメラとマニピュレータの構造と配置は、他のビジュアルサーボの制御法を用いる場合においても、ヤコビアンの変動が小さいため、時間遅れなどによるヤコビアンの推定誤差が小さくなり、収束軌道が安定すると考えられる。

3.7.1 ヤコビアンの非線形性が及ぼす軌道への影響

まずビジュアルサーボの一般的な定式化を行う。画像特徴量ベクトルを ξ 、ロボットの位置姿勢ベクトルを p 、ロボットの関節角ベクトルを θ とすると、特徴量を目標値 ξ_d に収束させるためには、

$$u = -\lambda J_{robot}^{-1} J_{image}^{-1} (\xi - \xi_d) \quad (3.15)$$

に基づいて関節速度指令値 u を用いて制御すればよい[12]。ここで、

$$J_{image} = \frac{\partial \xi}{\partial p}, \quad J_{robot} = \frac{\partial p}{\partial \theta} \quad (3.16)$$

と定義する。これらはそれぞれ、イメージヤコビアン、ロボットヤコビアンと呼ばれ、ロボットの位置姿勢に応じて変化する時変な行列である。(3.15)式の制御系の設計思想に基づけば、その理想軌道は ξ 空間における直線軌道である。しかし、(3.15)式は、本来非線形である制御系を離散時間内において線形系と近似していることにより、得られる収束軌道はサンプリング周期とゲインにより大きく変動し、一般的には曲線的な軌道を描く[18]。すなわち、 $J_{robot}^{-1} J_{image}^{-1}$ が離散化時間内に時変であることが、軌道が曲がる直接的原因となっている。

一方、本研究では、次式に示す時不変ゲインを用いた制御を提案した。

$$u = -\lambda R (\xi - \xi_d) \quad (3.17)$$

ここで、 R は画像特徴ベクトルとロボットの関節角ベクトルを時不変線形式により結びつける近似定数である。(3.17)式においては、もともと R が近似値であるため、 R と $J_{robot}^{-1} J_{image}^{-1}$

の差により軌道は曲線的となる。つまり、いずれの手法においても軌道が曲線となるのは、 $J_{robot}^{-1} J_{image}^{-1}$ が時変であることに起因している。

さて、デカルト空間におけるマニピュレータの手先の誤差ベクトル $e = p - p_d$ を定義し、 $\xi - \xi_d = J_{image} e$ が満たされると仮定する。ここでマニピュレータが(3.17)式に基づいて理想的に動作し、 $\dot{\theta} = u$ が成り立つと仮定すると、デカルト空間におけるマニピュレータの手先の誤差システムは次式のようにあらわすことができる。

$$\begin{aligned} \dot{e} &= J_{robot} u \\ &= -\lambda J_{robot} R J_{image} e \end{aligned} \quad (3.18)$$

本式において、 $J_{robot} R J_{image}$ が、単位行列のスカラー倍に近い場合に、 e は直線的に収束する。そこで、(3.18)式におけるヤコビアン($M = J_{robot} R J_{image}$)に対して、次の評価量 L を定義する。

$$\begin{aligned} L(M) &= \frac{m_1^T l_1}{|m_1|} \times \frac{m_2^T l_2}{|m_2|} \times \cdots \times \frac{m_n^T l_n}{|m_n|} \\ &\times \frac{(m_1 + m_2 + \cdots + m_n)^T (l_1 + l_2 + \cdots + l_n)}{|m_1 + m_2 + \cdots + m_n| |l_1 + l_2 + \cdots + l_n|} \end{aligned}$$

ただし、

$$M = (m_1, m_2, \cdots, m_n)$$

$$l_i = (n_1, n_2, \cdots, n_n)^T \{n_i = 1, n_{other} = 0\} \quad (3.19)$$

$L(M)$ は M が単位行列に近いほど1に近づくスカラー値である。以下、本評価量 L を用いてヤコビアンの一様性(軌道の安定性)を評価する。また、評価量 L をヤコビアンの時不変安定度と呼ぶ。

3.7.2 ハンドアイシステムのモデル

ハンドアイシステムの構造評価に用いるカメラとマニピュレータのモデルをFig.3.11に示す。マニピュレータは2リンク2自由度のスカラー型である。また、カメラは基線長 $2E = 70[\text{mm}]$ 、焦点距離 $f = 3[\text{mm}]$ の理想的な透視投影モデルに基づく平行ステレオカメラを用いる。また、画像特徴量としては特徴点の視空間座標 $\xi = (\gamma, \theta)^T$ を用いる。左右の画像平面上に投影された特徴点 $(X^L, Y^L), (X^R, Y^R)$ は次式で視空間座標 ξ に変換される。

$$\xi = ((X^L - X^R)/f, (X^L + X^R)/2f)^T \quad (3.20)$$

また、比較のため、マニピュレータに対して上面にカメラを1台取り付け付けたハンドアイシステムを考える(Fig.3.12)。このとき、カメラ画像にはデカルト座標 (x, z) がそのまま投影されており、 $\xi = p$ 、すなわち J_{image} が単位行列であると考えられる。一般には、このようにカメラを配置することが多い。

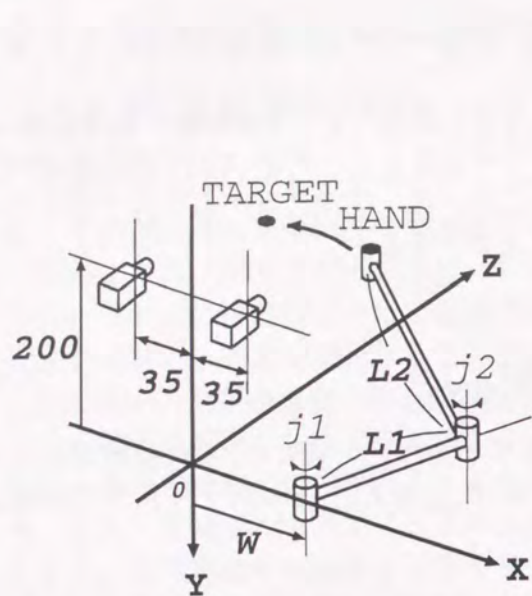


Fig. 3.11: ステレオ型ハンドアイシステムのモデル

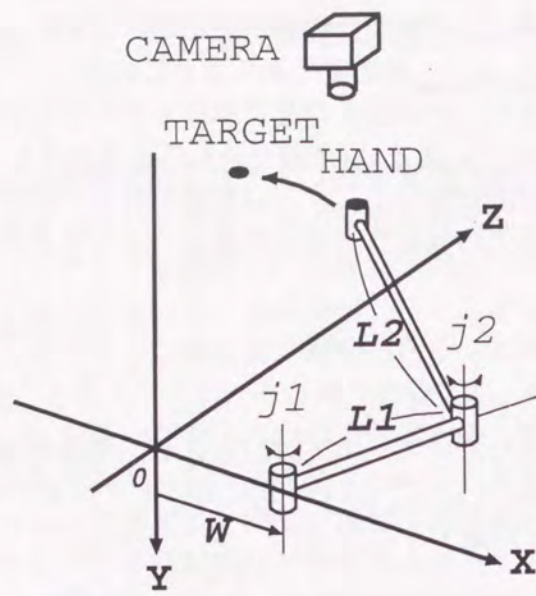


Fig. 3.12: 単眼型ハンドアイシステムのモデル

3.7.3 ハンドアイシステムの構造とヤコビアンの関係

まず, Fig.3.11に示すハンドアイシステムにおいてカメラ配置, リンク長さを $L_1 = 250, L_2 = 380, W = 200, K = 200$ と人間と似た値に設定したときの評価量 L の分布を Fig.3.13に示す. 行列 R の値は, 2.6節と同様にカメラに対して正面の領域を対象に ξ と θ の線形マッピングを行なう際の近似パラメータとして最小自乗法により定めた. また, Fig.3.12に示すカメラを上面に取り付けたハンドアイシステムを用いる場合の評価量 L の分布を Fig.3.14に示す. カメラ配置, リンク長さは同じく $L_1 = 250, L_2 = 380, W = 200$ と人間と似た値に設定した. Fig.3.13と Fig.3.14を比較すると, ステレオカメラを用いるほうが広範囲において評価量 L の値が大きいことが分かる. 図中の矢印線は, (3.17)式に基づくシミュレーションにより得られたマニピュレータの手先の軌道である. ここで, ゲインは $\lambda = 3$, サンプルングタイムは $33[\text{msec}]$ とし, ステレオカメラ, 単眼カメラ, それぞれにおいて3通りの試行を行なった. L が大きい領域においては, 軌道は直線的となることが分かる.

以下, Fig.3.11のシステムを用いる場合について, ヤコビアンが空間的に一様となるカメラ配置とマニピュレータの運動学的構造を評価量 L を用いて評価する. $L_1 + L_2 = 630$ の条件のもとで, L_1, L_2 の割合をそれぞれ変化させた場合に, 可動範囲内における $L > 0.6$ の領域の面積 $area$ を求めた. 結果を Fig.3.15に示す. 図より $L_1 = 230$ において最も広い面積が得られることがわかる. また, カメラと肩関節の水平距離 W を変化させた場合の $L > 0.6$ の領域の面積を Fig.3.16に示す. $W = -150$ と $W = 230$ に局所最小値が見られるが, $W = 230$ と作業領域をカメラに対してやや右前方とする配置のほうが, 面積 $area$ の変動も小さく安定していることがわかる. これらの結果は, L の閾値を 0.6 以外にした場合でも同様に得られる. 以上より, 人間の眼と腕の構造と配置においては, ヤコビアン

が空間的に広い範囲で一様であり, 収束軌道も安定することがわかる.

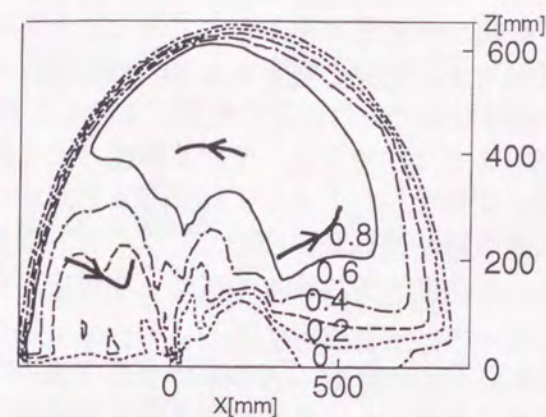


Fig. 3.13: ステレオ型システムにおける時不変安定度

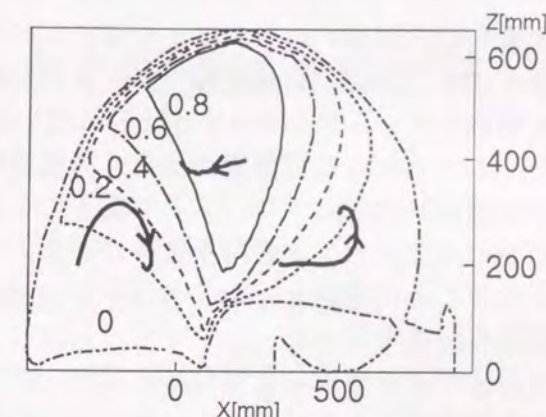


Fig. 3.14: 単眼型システムにおける時不変安定度

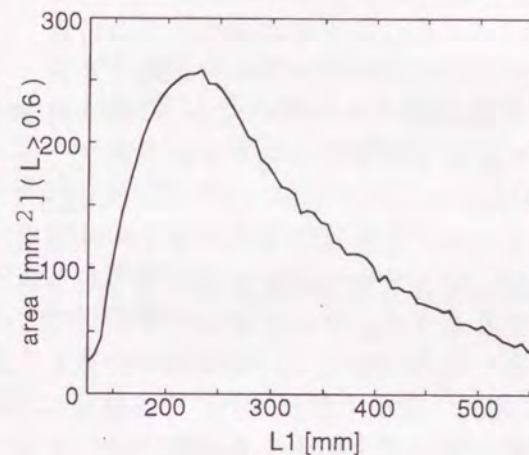


Fig. 3.15: リンク長さの比と時不変安定度

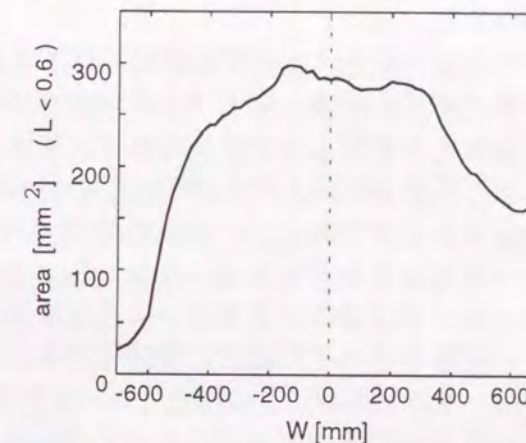


Fig. 3.16: カメラ位置と時不変安定度

3.8 結言

本章では人間型のハンドアイシステムにおいて従来、非線形と考えられてきたステレオ画像とロボットの関節角度のマッピングを線形近似できる性質を利用すると、時不変ゲインを用いたビジュアルサーボが十分に広い領域で動作可能であることを実験により示した。また、本システムは計算量が少なく単純な制御が行えるだけでなく、モデル誤差（とくにカメラ回転）に対してロバストであることを実験により示した。さらに、ビジュアルサーボにおけるヤコビアンの時不変性がパラメータ誤差や時間遅れなどに対してロバストな性質を示すことに着目し、シミュレーションによりハンドアイシステムの構造と軌道安定性の関係の評価を行った。

時不変ゲインを用いたビジュアルサーボの研究としては、他に, Espiau ら [36] や橋本ら [38] の研究がある。ビジュアルサーボに時不変ゲインを用いることは、計算量の低減やパラメータ変動に対するロバスト化など多くのメリットが存在するが、収束可能な領域の広さが問題となる。本節で示したシステムは、可動領域のほぼ全域にわたってヤコビアンが一様であるために、漸近安定性が保証されるとともに、収束軌道も安定して得られる点が特徴である。

ハンドアイシステムの構造評価に関する研究としては、他に Sharma ら [39] が、ビジュアルサーボにおけるハンドアイシステムのカメラ配置に関する評価法として 'motion perceptibility' を提案している。これは、マニピュレータの可操作性 [10] をカメラとマニピュレータの関係に拡張したもので、マニピュレータの動きがどの程度カメラで観測されるかを評価するものである。この評価法では、マニピュレータの可操作性とカメラの知覚性に分けて考えることができる。すなわち、ロボットが動作しやすい場所を作業領域として設定し、カメラは操作が見やすいような場所に設定するように、マニピュレータの構造とカメラの配置をそれぞれ独立に最適化することになる。これに対して、本研究で提案した評価法は、マニピュレータの構造とカメラ配置は共に関係しあっており、それぞれ独立には評価することができない点がこれまでの研究にない興味深い点である。本研究では、カメラとマニピュレータを独立に設置したシステムについて評価を行ったが、カメラがマニピュレータの手先に固定されたシステムについても同様に評価が行えると考えられる。

第4章 ビジュアルサーボと画像特徴検出法

4.1 緒言

本章では、ビジュアルサーボに用いる画像特徴検出法について考察を行い、ビジュアルサーボに適した画像特徴検出法の一手法として粗いオプティカルフローを用いた平面運動の検出法を提案する。本検出法は近似的なモデルに基づいた簡易な方法であるため、計測精度は良いとは言えないが、微小な移動量の検出には優れている。また、非常に少ない計算量で高速な計測が可能であるため、ビジュアルサーボと組み合わせることによって、高速かつ高精度な位置決め制御が実現できる。このように視覚系単独で用いる場合には非力な手法が、運動系と組み合わせること（ビジュアルサーボ）によって大変優れた手法と成りうることを本例は示している。

以下、本章ではまず、4.2節においてビジュアルサーボと画像特徴検出法の融合について考察を行い、ビジュアルサーボに用いる画像特徴検出法としては、環境復元の正確さよりも、目標とのずれ量が少ない計算量で高速に求まることが重要であることを述べる。次に4.3節において粗いオプティカルフローを用いた平面運動の検出法ならびに本手法を用いたビジュアルサーボを提案する。平面上の物体の位置決め実験 (Fig.4.1) によって、本手法がサブピクセル精度の位置と姿勢に関する位置決めを高速に行うことができることを示す。

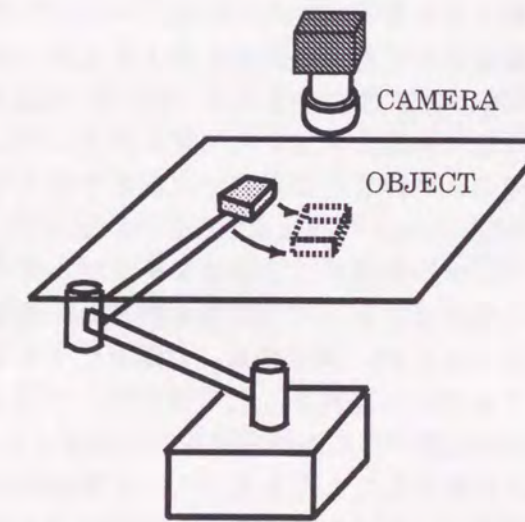

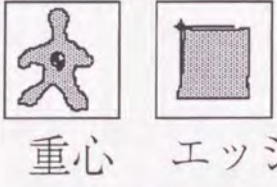



Fig. 4.1: カメラを用いた平面上の物体の位置決め

4.2 ビジュアルサーボと画像特徴検出法

	相関法 	図形的特徴  重心 エッジ	従来の オプティカルフロー 
設定の手間	○	×	○
計算量	△	○	×
計測精度	△	○	△
回転量の検出	×	○	○

Tab. 4.1: ビジュアルサーボに用いる画像特徴検出法

カメラを用いて平面上の物体の位置と姿勢を目標に合わせるタスクを高精度に行うためには、高精度な位置計測技術と物体を動かすロボットの高精度な制御技術が必要である。しかし、カメラとロボットがいくら高精度であっても、カメラとロボットの位置関係が正確に同定されなければ、対象物体の位置決め精度も向上しない。また、このようなキャリブレーションには、大変な手間が必要である[4, 11]。そこで、視覚情報をもとにフィードバック制御を行うことにより、カメラやロボットのモデル誤差に対してロバストな位置決め制御技術が近年は盛んに研究されている。これまでに提案されたビジュアルサーボに用いられている画像特徴の検出方法について Tab.4.1 にまとめる。

相関法（テンプレートマッチング法）は、目標をテンプレートとして取り込み、画像中の探索領域を走査しながら各位置において相関値を計算し、最も相関値の大きい位置を目標の位置として定める方法である[9]。本手法は、目標をテンプレートとして取り込むだけで利用できるというシステム設定の簡便さと汎用性の高さが利点である。計算量が大きい欠点があるが、最近では実時間処理が行える専用画像処理装置（トラッキングビジョン：富士通株式会社）[43] が安価に市販されたこともあって、近年は特に多く用いられている。相関法における計測単位は通常は1ピクセルである。補間処理を併用することにより、1/数～1/10ピクセル程度の精度で検出が可能とされているが[44, 45]、ハードウェアによるサポートはまだ行われていない。また、回転量の計測には、少しずつ回転させたテンプレートを多数枚用意し、これらとのマッチング結果より最適なものを選ぶ手法が通常用いられ

るが[9]、計算量が多くなる欠点がある。また、複数のテンプレートのマッチング結果より線形補間により回転量を得ようとする試みもみられる[46]。

次に図形的特徴を利用する方法は、2値化などの前処理を行って得られた画像特徴の重心位置やエッジ位置などを用いる方法である[47]。本手法は、計算量が少なく済み、環境の明度変化などに対してロバストである反面、システム設計時には環境にあわせて画像処理を調整したり、画像処理が適切に行えるように環境を設定する必要があり、システム設定に手間がかかる難点がある。しかし、回転量の検出やサブピクセル精度の計測などは容易に可能であり、画像処理のカスタマイズを行うぶんだけ自由度は大きい。現在のビジュアルサーボの研究においては、処理の高速性から最も多く用いられている手法である。

オプティカルフローとは時間的に連続した画像より得られる画像上の速度場のことである。画像中の物体を剛体と仮定することで、オプティカルフローより、物体の運動やカメラの運動を計測することができる[48]。通常、オプティカルフローは時間的に連続した画像を用いて、画像中の物体やカメラの速度を求めるのに用いられる[49]。これに対して、基準画像と現在の画像を用いてオプティカルフローを求めると、物体やカメラのずれ量を基準画像に対する位置として計測することもできるはずである。しかし、オプティカルフローは大きな運動量に対しては計測精度が悪いので、一般にこのような使用例は見られない。オプティカルフローの検出法としては、これまでに数多くの手法が開発されている[50]。しかし、精度の良いオプティカルフローを得るためには多くの計算量が必要であるために、今のところリアルタイムにオプティカルフローを利用した例としては、画像中の動物体の抽出のように精度が問題とされないタスク[30, 52]に限定されている。

さて、ビジュアルサーボを用いた位置決め制御においては、最終的な位置決め精度が重要であり、途中の軌道はさほど問題とならない場合が多い。このような場合には、ビジュアルサーボに用いる特徴量としては、高速に得られることと、微小なずれ量が検出されることが重要であり、収束動作の途中における計測精度は大きな問題とはならない。先に述べた3つの特徴量を比較すると（Tab.4.1）、計算量の少なさと計測精度の高さから、図形的特徴量を用いる方法が最も適しているように思われるが、システム設定の手間が欠点である。そこで本研究では、オプティカルフローがシステム設定の手間が少なく済み、および微小な運動に対しては少ない計算量でも高精度な検出が可能であることに着目し、計算量の少ないオプティカルフローの検出法を提案する。また、リアルタイムにオプティカルフローを検出しながら位置決め制御を行うビジュアルサーボを提案する。オプティカルフローは大きな移動量に対しては、安定した検出が困難であるが、本研究では画像解像度を変化させながら段階的に位置決めを行うことで、大きな誤差が存在する場合にも高精度な位置決めを実現する。

4.3 粗いオプティカルフローを用いたビジュアルサーボ

4.3.1 はじめに

本節では、計算量の少ない粗いオプティカルフローの検出法ならびに本手法を用いたビジュアルサーボを提案する。Fig.4.1に示すような平面上の物体の位置決め制御実験によって、本手法は少ない計算量でありながら、サブピクセル精度で位置と姿勢の位置決めが行えることを示す。さらに本手法が画像の明度変化に対してもロバストであることを示し、また、本手法を用いた位置決め制御の収束漸近安定性をリアプノフ関数を用いて考察する。

4.3.2 粗いオプティカルフローを用いた画像の並進量と回転量の検出

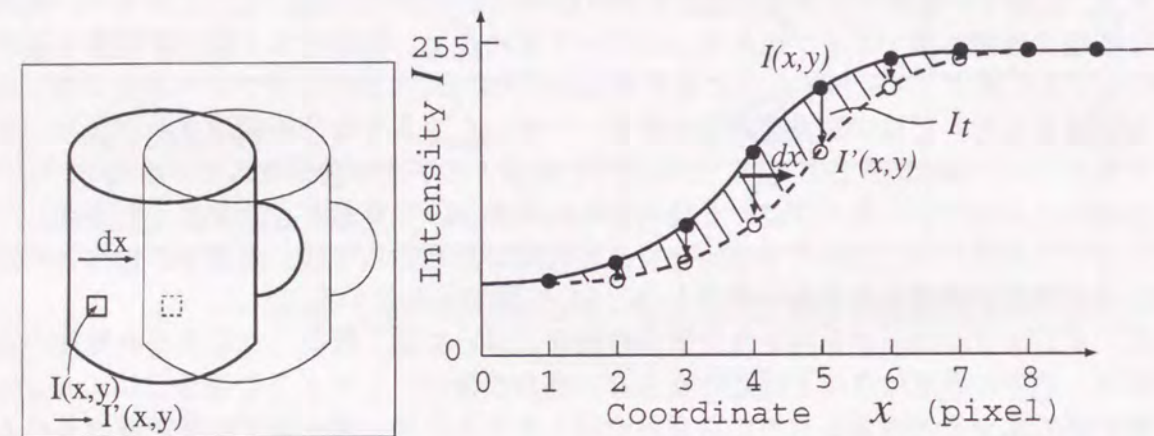


Fig. 4.2: 画像の並進

いま、画像上の点 (x, y) の濃度を $I(x, y)$ とし、画像全体（画像に写っている物体と背景、またはカメラ）が x 軸方向に dx だけ移動したとする（Fig.4.2）。また、移動後の画像上の点の濃度は $I'(x, y)$ とする。Fig.4.3は、このときの画像上の x 軸方向の一画素列の濃度変化を示している。実線は移動前の濃度、点線は移動後の濃度である。このとき、画像上の物体と背景の濃度が不変であると仮定すると、実線と点線で囲まれた斜線部の面積は移動量 dx に対して比例することが知られている [51]。画像上の x 軸方向の各画素列における斜線部の面積の総和を次式によって計算することで、画像全体の x 軸方向の並進量を近似的に推定することができる。

$$s_x = \sum_{x,y} |I_t| \quad (4.1)$$

$$I_t = I(x, y) - I'(x, y)$$

ここで s_x は推定並進量（斜線部の面積の総和）、 I_t は濃度変化量である。しかし、 s_x では移動方向が判別できない。そこで、各点における空間的な濃度勾配の符号と濃度変化量の

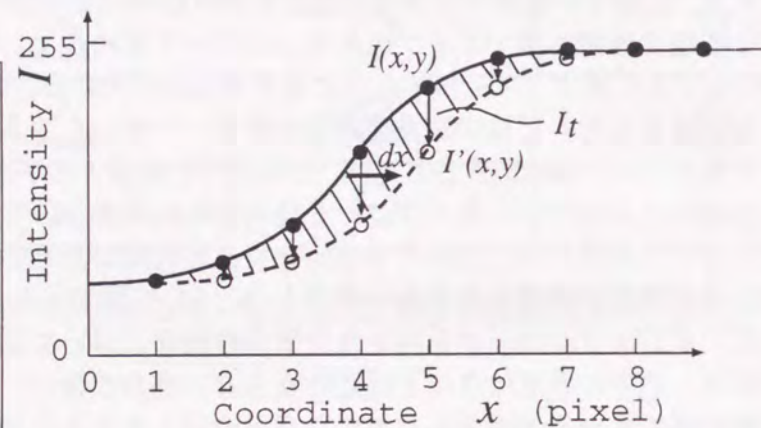


Fig. 4.3: 並進量の検出モデル

4.3. 粗いオプティカルフローを用いたビジュアルサーボ

符号から画像の移動方向が判別できる性質を利用して、次式によって面積に方向性を持たせる。

$$s'_x = \sum_{x,y} (I_t \operatorname{sign}(\frac{\partial I}{\partial x}))$$

$$\operatorname{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \quad (4.2)$$

ここで、 s'_x は方向性を持たせた面積、 $\frac{\partial I}{\partial x}$ は点 (x, y) における空間的な濃度勾配である。画像が x 軸の正の方向に移動したときは s'_x は正であり、負の方向に移動したときは s'_x は負となる。

次に画像全体が x 軸方向に dx 、 y 軸方向に dy だけ移動した場合を考える。このとき、(4.2) 式の I_t は、 dx だけでなく dy によっても影響を受けて変化するが、 dy による I_t の変化量と $\frac{\partial I}{\partial x}$ の符号が無関係であると仮定すると、画像全体の総和を計算することで (4.2) 式に対する dy の影響はキャンセルされる。このとき、 s'_x と dx の比例関係は維持されるので、 y 軸方向の移動量が存在する場合でも、(4.2) 式によって x 軸方向の移動量が推定できる。また同様に、 y 軸方向の移動量も次式によって推定できる。

$$s'_y = \sum_{x,y} (I_t \operatorname{sign}(\frac{\partial I}{\partial y})) \quad (4.3)$$

(4.2) 式、(4.3) 式は2枚の画像を平均して離散化すると、次式のようにあらわすことができる。

$$s'_x = \sum_{x,y} (\Delta_t \operatorname{sign}(\Delta_x))$$

$$s'_y = \sum_{x,y} (\Delta_t \operatorname{sign}(\Delta_y))$$

$$\Delta_t = \{(I(x-1, y-1) - I'(x-1, y-1)) + (I(x, y-1) - I'(x, y-1)) + (I(x-1, y) - I'(x-1, y)) + (I(x, y) - I'(x, y))\} / 4$$

$$\Delta_x = \{(I(x, y) - I(x-1, y)) + (I'(x, y) - I'(x-1, y))\} / 2$$

$$\Delta_y = \{(I(x, y) - I(x, y-1)) + (I'(x, y) - I'(x, y-1))\} / 2 \quad (4.4)$$

ここで、 s'_x 、 s'_y はそれぞれ s'_x 、 s'_y の離散化値、 Δ_t 、 Δ_x 、 Δ_y はそれぞれ I_t 、 $\frac{\partial I}{\partial x}$ 、 $\frac{\partial I}{\partial y}$ の離散化値である。

次に画像全体の回転量を求める。画像の回転中心が画像中心近傍にあると仮定すると、各点におけるフローベクトル（オプティカルフロー）が画像中心に対してなす回転モーメン

トの和は、回転量に対して比例する。そこでまず、オプティカルフローを検出する。オプティカルフローの検出については、これまでに数多くの手法が開発されているが、本研究では計算量の少なさを重視して、(4.4)式の S'_x, S'_y が画像上の各点におけるフローベクトルの総和を表していると考え、 Σ を外した次式によってフローベクトルを求める。

$$\begin{aligned} u_x &= \Delta_t \text{sign}(\Delta_x) \\ u_y &= \Delta_t \text{sign}(\Delta_y) \end{aligned} \quad (4.5)$$

$(u_x, u_y)^T$ は点 (x, y) におけるフローベクトルを表す。これを粗いオプティカルフローと呼ぶ。画像中心の座標を (x_c, y_c) とすると、画像全体のフローベクトルがなす回転モーメントは次式によって求まる。

$$S'_r = \sum_{x,y} (u_x (y - y_c) - u_y (x - x_c)) \quad (4.6)$$

ここで、 S'_r は回転モーメントの総和であり、画像の回転量に比例する値である。(4.5)式で求める粗いオプティカルフローは、大雑把な近似に基づくものであり、検出精度は悪いが、(4.4)式および(4.6)式で求める画像の並進量および回転量は、粗いオプティカルフローが画像全体に対して平均をされるために、4.3.5節、4.3.6節で示すようにビジュアルサーボにとっては十分な精度で検出される。(4.5)式のオプティカルフローの検出方法については、次節で再度説明を行う。

さて、(4.4)式、(4.6)式で求めた並進量および回転量は相対的なものであり単位は未知である。そこで目標画像を x 方向、 y 方向に1[pixel]シフト、または1[deg]回転して作成した画像ともとの画像を比較し、それぞれについて求めた並進量 \bar{S}'_x, \bar{S}'_y 、および回転量 \bar{S}'_r を基準値として次式のように並進量 v_x, v_y および回転量 v_θ の単位付けを行う。

$$\begin{aligned} v_x &= S'_x / \bar{S}'_x [\text{pixel}] \\ v_y &= S'_y / \bar{S}'_y [\text{pixel}] \\ v_\theta &= S'_r / \bar{S}'_r [\text{deg}] \end{aligned} \quad (4.7)$$

以上、画像の並進量と回転量の検出方法について説明した。本手法の実験結果は4.3.5節に示す。また、本手法を用いた位置決め制御の方法については、4.3.4節に示す。

4.3.3 粗いオプティカルフローに関する補足説明

本節では(4.5)式に示したオプティカルフローについて、一般的なオプティカルフローの算出方法の観点から説明を行う。いま、画像上の点 (x, y) の時刻 t における濃度を $I(x, y, t)$ とし、微小時間 dt の間に、点が dx, dy だけ移動したとする。画像上の物体の濃度が不変であると仮定すると次式が成り立つ[53]。

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (4.8)$$

右辺をテイラー展開し、2次以上の項を無視すると次式を得る。

$$I(x, y, t) = I(x, y, t) + dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t} \quad (4.9)$$

整理すると次式を得る

$$\frac{\partial I}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial I}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (4.10)$$

ここで見掛けの速度ベクトル $(\frac{dx}{dt}, \frac{dy}{dt})^T$ を $(u_x, u_y)^T$ 、空間的な濃度勾配 $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ を I_x, I_y 、時間的な濃度勾配 $\frac{\partial I}{\partial t}$ を I_t とすれば次のように表現される。

$$I_x \cdot u_x + I_y \cdot u_y + I_t = 0 \quad (4.11)$$

本式はオプティカルフローの時空間勾配条件と呼ばれる。この条件だけからフローベクトル $(u_x, u_y)^T$ を一意に決定することはできないので、通常これにさまざまな制約を付加することでフローベクトルの推定が行われる。

たとえば、あるブロック内の画素に関して、時空間勾配条件式の2乗誤差最小の条件の下で(4.11)式を解くと、次式が得られる[54]。

$$\begin{aligned} u_x &= -\frac{\sum_{x,y} I_t I_x}{\sum_{x,y} I_x^2} \\ u_y &= -\frac{\sum_{x,y} I_t I_y}{\sum_{x,y} I_y^2} \end{aligned} \quad (4.12)$$

ただし、次式が近似的に成り立つと仮定している。

$$\sum_{x,y} I_x I_y = 0 \quad (4.13)$$

本式は画素ブロック内に斜め方向の勾配が存在しないことを意味している。(4.12)式は更に次のように簡略化することができる[55]。

$$\begin{aligned} u_x &= -\frac{\sum_{x,y} I_t \text{sign}(I_x)}{\sum_{x,y} |I_x|} \\ u_y &= -\frac{\sum_{x,y} I_t \text{sign}(I_y)}{\sum_{x,y} |I_y|} \end{aligned} \quad (4.14)$$

(4.12)式、(4.14)式は画像上のブロック内の画素について計算することで、各ブロック毎の移動量を検出する方法である。これに対して本研究で提案する(4.5)式は前式を更に簡略化したうえに、各ピクセル毎に計算するようにしたものと考えることができる。このように(4.5)式に提案した粗いオプティカルフローは非常に簡略化を施しており、計算量が少ない反面、計測精度も悪い。しかし、後節で示すように、微小運動に対しては高精度な検出が可能であり、ビジュアルサーボに対しては十分に適用可能である。

4.3.4 階層画像を用いた段階的位置決め

一般にオプティカルフローは、画像の動き量が大きい場合には精度よく安定に推定することができない。そこで佐藤ら [57] は階層化された画像をもとにオプティカルフローを大まかな動きから細かな動きに向かって順次推定する方法を考案している。この方法では、まず原画像を縦横 1/2 に順次縮小した画像列を作成する。そして、それぞれの階層画像に基づいて推定された動き量はその量だけつぎの階層画像を変位させることによって順次伝播される。このように速度ベクトルを反復的に繰り返し補正することによって精度良く推定する方法である。

さて、本研究では、最終的な位置決め精度の向上が目的であるので、佐藤らと同様に階層画像を用意して、まず粗い画像で位置決め制御を行ったのちに、段階的に細かい画像で位置決め制御を行うことによって、画像の動きが大きい場合にも精度の良い位置決めが行えるようにする (Fig.4.4)。解像度を切り替えるタイミングは、各解像度において検出された並進推定量がある閾値より小さくなった時点とする。閾値の設定については 4.3.7 節で述べる。

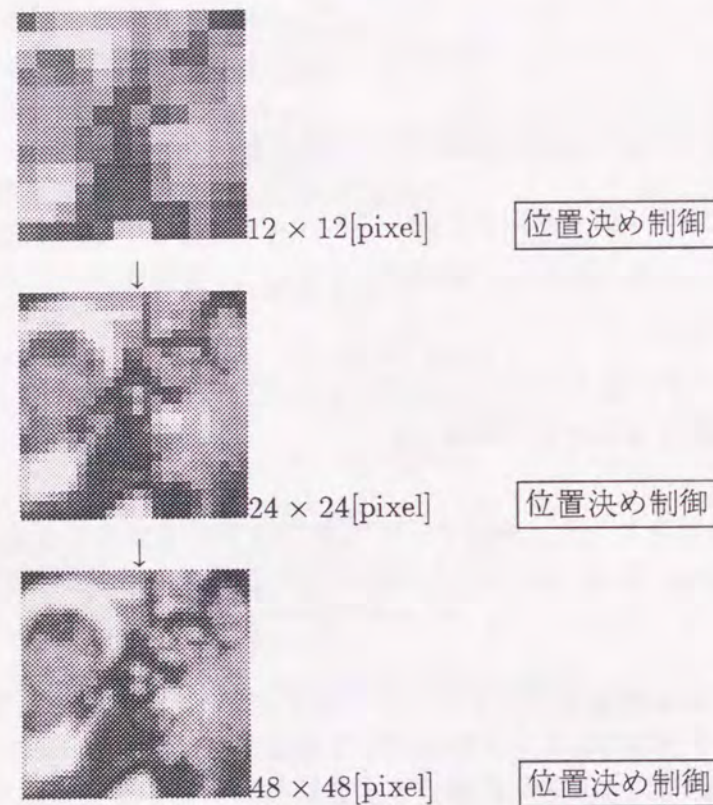


Fig. 4.4: 階層画像を用いた段階的な位置決め

4.3.5 平面運動の検出実験

本節では前述した平面運動検出法の実験結果を示す。実験には、写真を CCD カメラで撮影して取り込んだ 150×150 画素の濃淡画像 (Fig.4.5) を用いた。以下、この画像をソフト的に並進および回転させ、図中央の矩形領域 (96×96 画素) がカメラで観測されていると考えて、移動前後の矩形領域の画像を比較することで画像の並進量、回転量の推定を行う。

粗いオプティカルフローの検出例

画像を水平方向に $10[\text{pixel}]$ 並進移動し、 $10[\text{deg}]$ 回転させたときの移動前後の矩形領域の画像を、Fig.4.6, Fig.4.7 にそれぞれ示す。これらの画像より縦横 1/2 に順次縮小した階層画像を作成する。画像の階層化は、 $2 \times 2[\text{pixel}]$ の画素の濃度値の単純平均を用いて行った。Fig.4.8, Fig.4.9 はそれぞれ 12×12 画素に解像度を下げた画像である。

Fig.4.8 と Fig.4.9 について、(4.5) 式を用いてオプティカルフローを求めた結果を Fig.4.10 に示す。(4.5) 式より得られるオプティカルフローは、計算量が小さい反面、精度の悪い大雑把なものであることがわかる。このとき (4.7) 式を用いて得られた推定並進量、推定回転量は $v_x = 8.4[\text{pixel}]$, $v_y = 0.7[\text{pixel}]$, $v_\theta = 12.9[\text{deg}]$ であった。

Fig. 4.5: 使用した画像 (150×150)



Fig. 4.6: 原画像 (96×96)

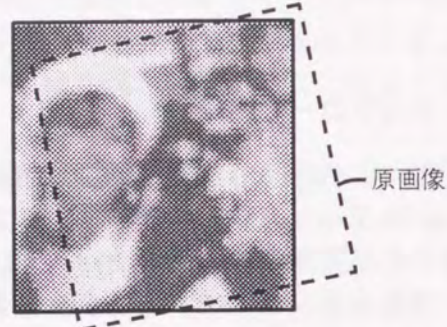


Fig. 4.7: シフト画像 (96×96)

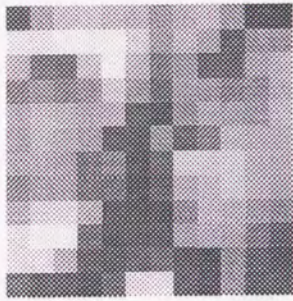


Fig. 4.8: 粗画像 (12×12)

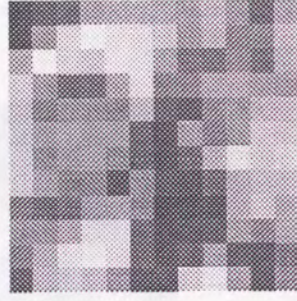


Fig. 4.9: シフト粗画像 (12×12)

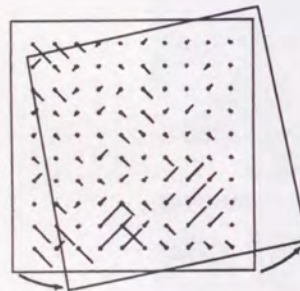


Fig. 4.10: オプティカルフロー

同画像を x 方向に $-20[\text{pixel}] \sim 20[\text{pixel}]$ の範囲で平行移動させた後に解像度の階層化を行い、それぞれの階層画像で (4.7) 式を用いて並進量を推定した結果を Fig.4.11 に示す。本図の単位は、 96×96 ピクセルの画像におけるピクセルを用いている。図より解像度が高くなるほど検出可能範囲が狭くなる様子がわかる。また、逆に解像度が低いと微小運動が検出されない様子がわかる。この結果より、解像度を段階的に高くしながら位置決めを行うことで、大きな動き量に対しても精度良く位置決めが行えることがわかる。また、1 ピクセル以内の小さな動きに対しても、これらの運動量は検出できていることからサブピクセル精度の位置決めが可能であることがわかる。一般にサブピクセルの手法を用いた位置計測においては、0.1 ピクセル程度の精度が期待されるが [58]、本手法においても実験で約 0.1 ピクセル以内の精度で誤差が検出できることが確認された。

次に同画像を画像中心を回転軸として、 $-6[\text{deg}] \sim 6[\text{deg}]$ の範囲で回転させた場合に回転量を推定した結果を Fig.4.12 に示す。この図より、回転についても解像度が高くなるほど検出可能範囲が狭くなる、また、解像度が低くなると微小回転が検出されない様子がわかる。また、回転量に関しては 6×6 ピクセルではややノイズが大きく、 12×12 ピクセルの画像が最低解像度として適当であると思われる。

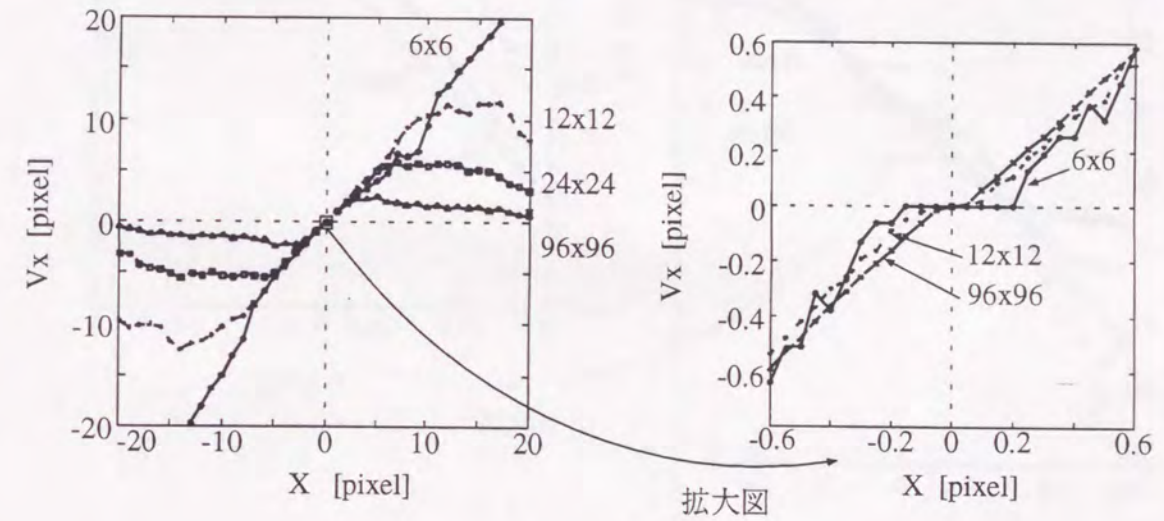


Fig. 4.11: 並進量の推定

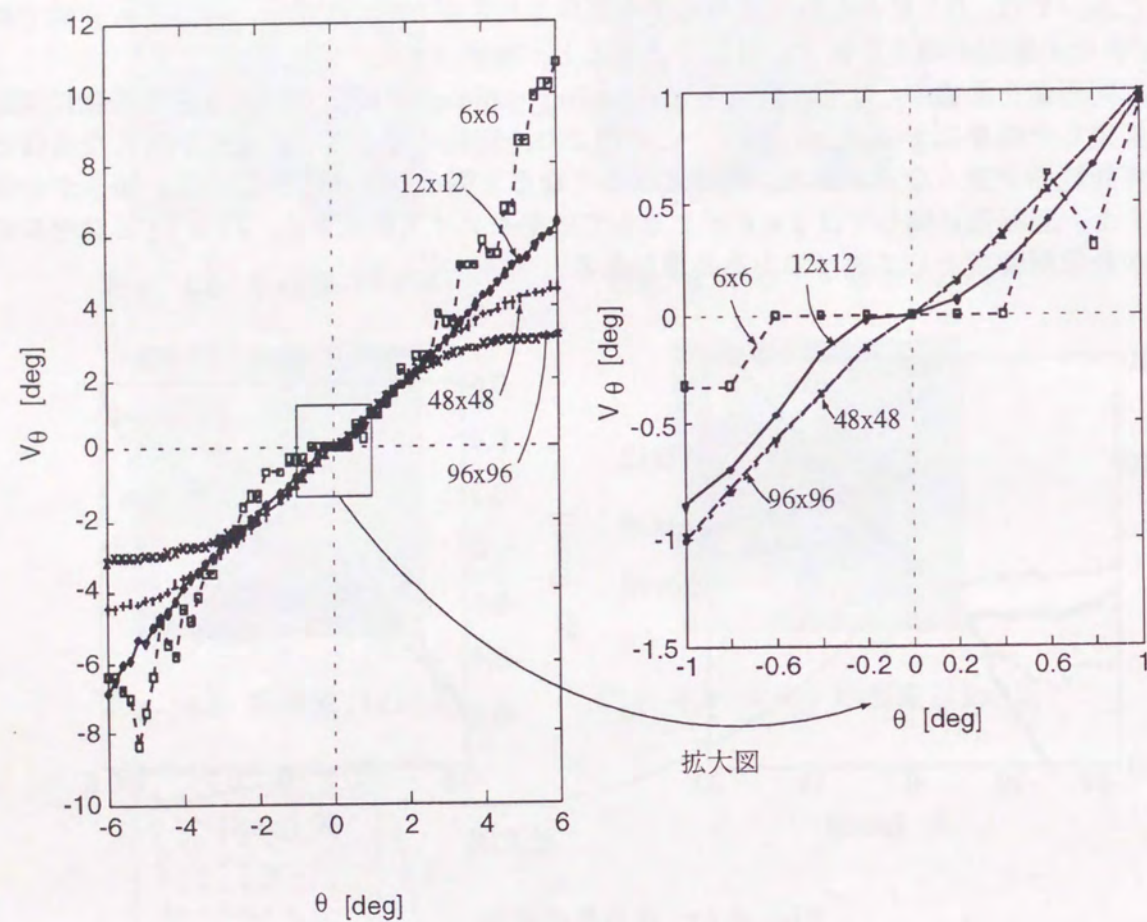


Fig. 4.12: 回転量の推定

次に同画像を x, y 両方向に同時に平行移動を行った場合の 96×96 ピクセル, 24×24 ピクセル, 6×6 ピクセル画像における推定結果を Fig.4.13 (左側 2 列) に示す. 理想的には, これらのグラフは傾き 1 の平面となるはずである. まず, 6×6 ピクセルの画像における推定結果を見ると, ほぼ平面が現われており, x, y 方向の並進が同時にある場合にも, それぞれ独立に並進量が推定できていることがわかる. 次に 24×24 ピクセル, 96×96 ピクセルの推定結果を見ると, グラフ中央部のみにおいて平面が現われており, 先に示した x 軸方向のみに並進を与えた場合の推定結果 (Fig.4.11) と同様に, 解像度が高くなるほど, 並進量が推定できる領域が狭くなることわかる.

次に x 方向の並進と画像中心周りの回転を同時に与えた場合の推定結果を Fig.4.14 (左側 2 列) に示す. 同じくグラフに平面が現われており, 並進量と回転量も独立に検出されている様子わかる.

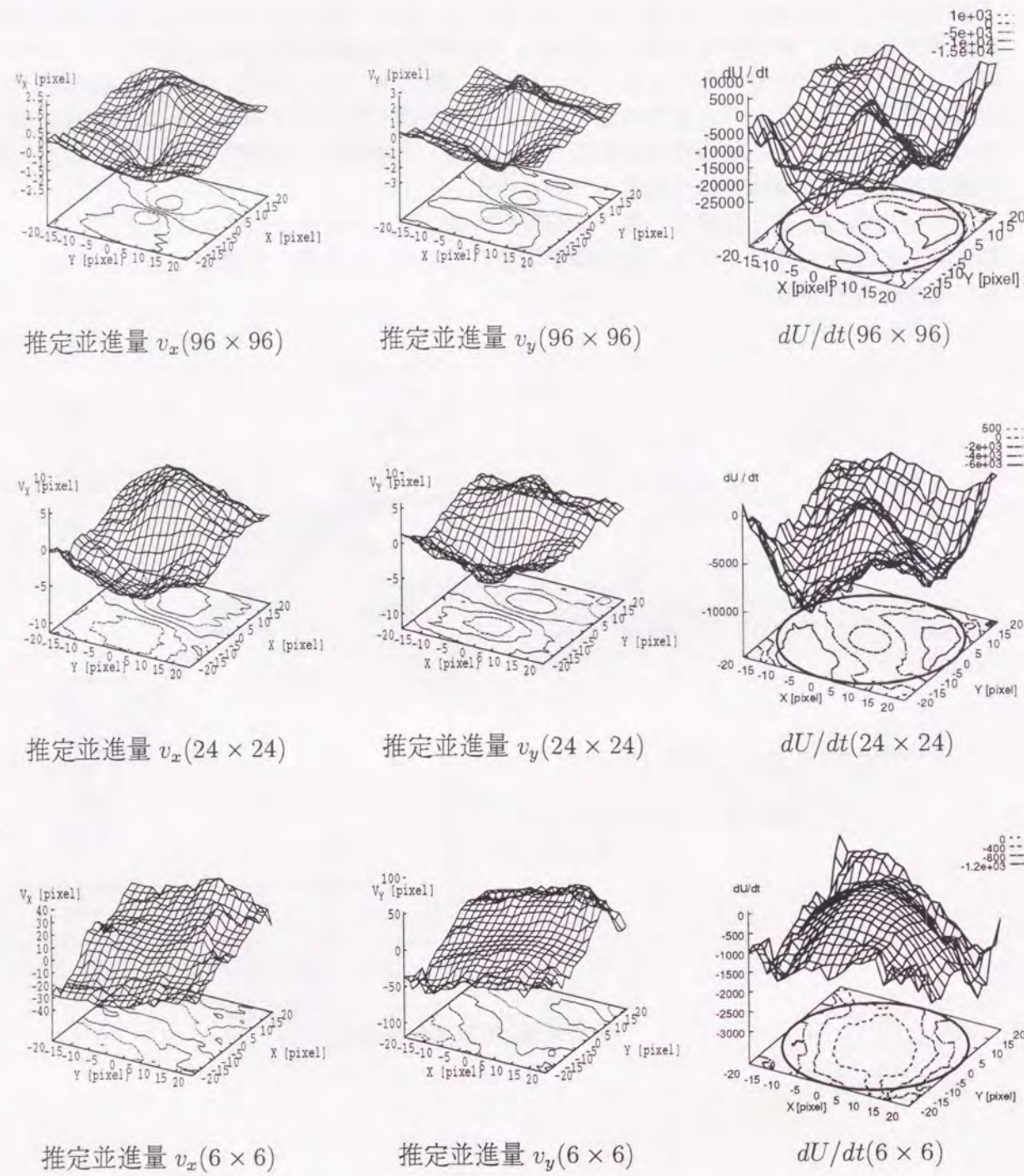


Fig. 4.13: x 方向, y 方向に平行移動させた場合の推定並進量 v_x, v_y とビジュアルサーボにおける収束安定性 dU/dt

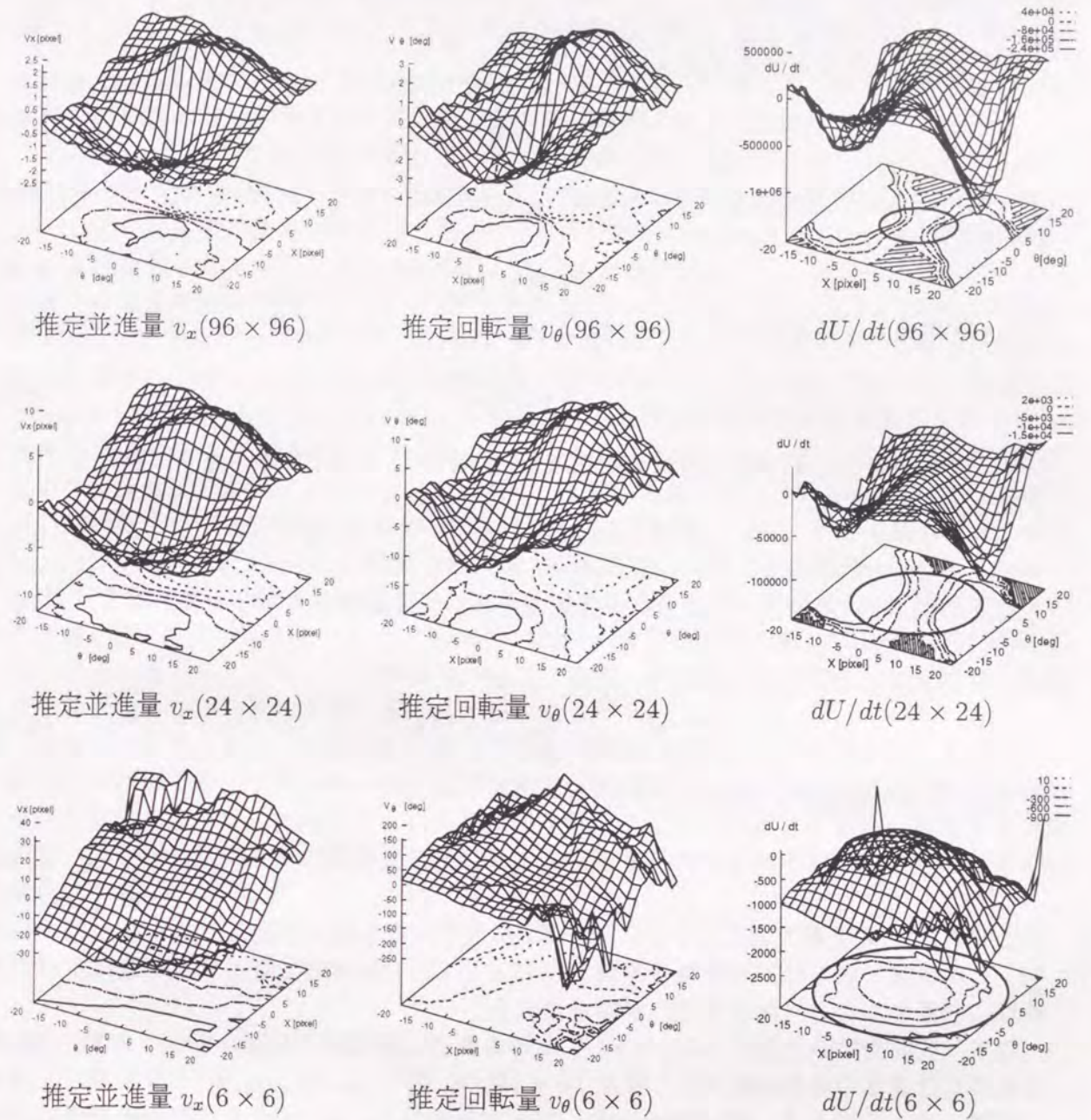


Fig. 4.14: x 方向に平行移動させて, 回転させた場合の推定並進量 v_x と推定回転量 v_θ とビジュアルサーボにおける収束安定性 dU/dt

4.3.6 ビジュアルサーボの漸近安定性

本節では、(4.7)式を用いて位置決めフィードバック制御を行う場合の漸近安定性をリアプノフ関数を用いて考察する。ここで物体が速度指令値 $\mathbf{u} = (u_x, u_y, u_\theta)^T$ とおりに理想的に動作するものとする。まず、目標位置姿勢に対する誤差ベクトル $\mathbf{e} = (V_x, V_y, V_\theta)^T$ を定義すると、誤差システムは次式で表される。

$$\dot{\mathbf{e}} = \mathbf{u} \quad (4.15)$$

(4.7)式で得られた推定誤差ベクトル $\hat{\mathbf{e}} = (v_x, v_y, v_\theta)^T$ にスカラーゲイン λ を乗じて、速度指令値を次式で与えたとする。

$$\mathbf{u} = -\lambda \hat{\mathbf{e}} \quad (4.16)$$

このとき閉ループ系のダイナミクスは、(4.16)式を(4.15)式に代入して次式で表される。

$$\dot{\hat{\mathbf{e}}} = -\lambda \hat{\mathbf{e}} \quad (4.17)$$

以下、上式が目標位置姿勢に収束する領域を検証する。ここで次のスカラー関数(リアプノフ関数)を定義する。

$$U = \frac{1}{2} \mathbf{e}^T \mathbf{e} \quad (4.18)$$

上式は(4.17)式のシステムにおいて連続である。上式を微分すると次式を得る。

$$\begin{aligned} \frac{d}{dt} U &= \mathbf{e}^T \dot{\mathbf{e}} \\ &= -\lambda \mathbf{e}^T \hat{\mathbf{e}} \end{aligned} \quad (4.19)$$

次に Ω を次式で定義する閉領域とする。

$$U < U_0 \quad (4.20)$$

ここで U_0 は正の定数である。もし、閉領域 Ω に含まれる目標点(原点)以外の \mathbf{e} について $\frac{d}{dt} U < 0$ が成り立てば、不変集合定理[37]によって、 Ω を初期位置とする軌道はすべて目標点に収束すること(漸近安定)が保証される。

以下、前節で用いた画像(Fig.4.6)における漸近安定領域の検証を行う。まず、回転を与えずに並進量のみの制御を行う場合 ($\mathbf{e} = (V_x, V_y, 0)^T$, $\hat{\mathbf{e}} = (v_x, v_y, 0)^T$ とする場合) について考える。このとき、 96×96 [pixel], 24×24 [pixel], 6×6 [pixel] の各解像度画像において制御を行う場合の $\frac{d}{dt} U$ の分布をそれぞれ、Fig.4.13の右側1列に示す。図中、実線の円は $\frac{d}{dt} U < 0$ の領域に含まれる最大半径の閉領域 Ω を示している。すなわち、実線の円の内部が漸近安定が保証される領域である。これらの図では、ほとんどの領域で $\frac{d}{dt} U < 0$ が成り立ち、 $\frac{d}{dt} U > 0$ の領域(斜線部)はほとんど含まれない。また、これらの図における Ω はすべて $U_0 = 20^2$ であり、半径 20 [pixel] 以内の誤差であれば漸近安定が保証されることがわかる。

次に、 x 軸方向の並進と回転の制御を行う場合 (y 軸方向の並進を行わない場合、 $\mathbf{e} = (V_x, 0, V_\theta)^T$, $\hat{\mathbf{e}} = (v_x, 0, v_\theta)^T$ とする場合) について検証する。このとき、 96×96 [pixel], 24×24 [pixel], 6×6 [pixel] の各解像度画像において制御を行う場合の $\frac{d}{dt} U < 0$ の分布をそれぞれ、Fig.4.14の右側1列に示す。実線の円で示した Ω の領域は、それぞれ $U_0 = 8^2$, $U_0 = 14^2$, $U_0 = 20^2$ である。すなわち、並進誤差[pixel]と回転誤差[deg]のノルムがそれぞれ $8, 14, 20$ 以内であれば漸近安定が保証される。これらの図より、解像度が高くなるにつれて斜線部の領域が大きくなり、漸近安定領域が狭くなるのがわかる。

次に x, y 方向の並進と回転を含む3次元の制御を行う場合 ($\mathbf{e} = (V_x, V_y, V_\theta)^T$, $\hat{\mathbf{e}} = (v_x, v_y, v_\theta)^T$ とする場合) について検証する。Fig.4.15の右側、およびFig.4.16に、 96×96 [pixel] の解像度画像で制御を行う場合の $\frac{d}{dt} U$ の分布を -20 [deg] \sim 20 [deg] の回転誤差毎における断面図として示す。図中の斜線は、 $\frac{d}{dt} U > 0$ の領域を示し、実線の円は $\frac{d}{dt} U < 0$ に含まれる最大半径の閉領域 ($U_0 = 15^3$) を示している ($\theta = -20, -15, 15, 20$ [deg] の図は、閉領域 Ω が含まれないため実線の円は図示していない)。これらの図より、3次元の制御においては 96×96 [pixel] の解像度で制御を行う場合、 $\mathbf{e}^T \mathbf{e} < 15^3$ の領域であれば(並進誤差[pixel]と回転誤差[deg]のノルムが15以下であれば)漸近安定が保証されることがわかる。また、一辺の長さが10.6の立方体が Ω に内接することから、 10 [pixel], 10 [deg] 以内の誤差であれば漸近安定が保証されることがわかる。また、同様に 12×12 [pixel] の解像度画像を用いて制御を行う場合の $\frac{d}{dt} U > 0$ の分布をFig.4.15の左側、およびFig.4.17に示す。 Ω は $U_0 = 17^3$ であり、 $\mathbf{e}^T \mathbf{e} < 17^3$ の領域であれば(並進誤差[pixel]と回転誤差[deg]のノルムが17以下であれば)漸近安定が保証されることがわかる。また、同様に一辺の長さが12の立方体が Ω に内接することから、 12 [pixel], 12 [deg] 以内の誤差であれば漸近安定が保証されることがわかる。このように解像度が低くなると、漸近安定領域は広くなるので、解像度を切り替えながら制御を行う有効性が分かる。

前節で示したように、並進回転誤差の推定量の精度は悪くとも、本節で示したようになり広い領域でビジュアルサーボの漸近安定性は保証される。しかし、漸近安定性が保証される領域の広さは、本節のように実際に誤差量を推定して確認する以外に推定する方法がない。画像より何らかの特徴量を検出して、収束安定領域を求める手法の開発が今後の課題である。

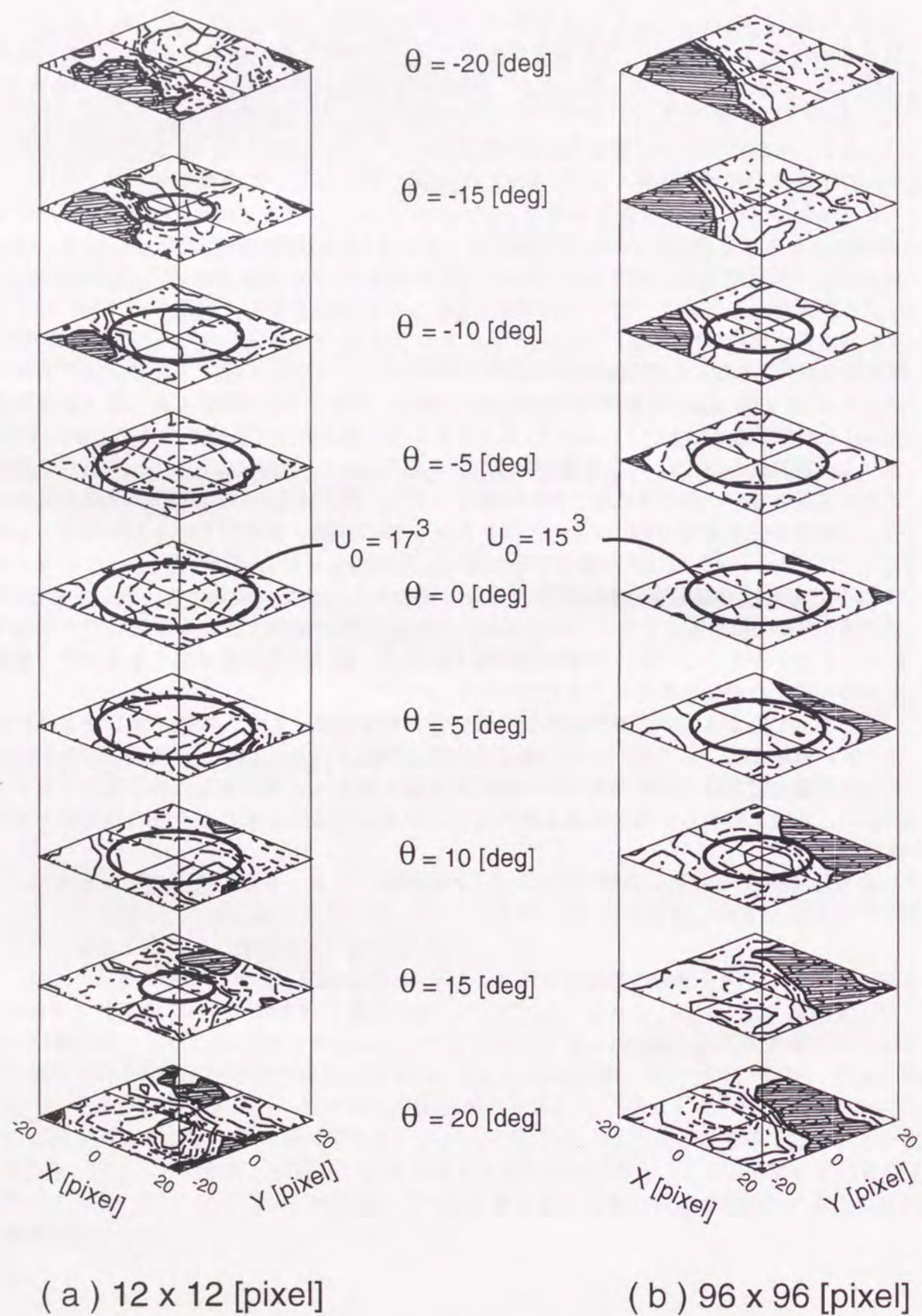


Fig. 4.15: 漸近安定性の解析 (dU/dt の分布図)

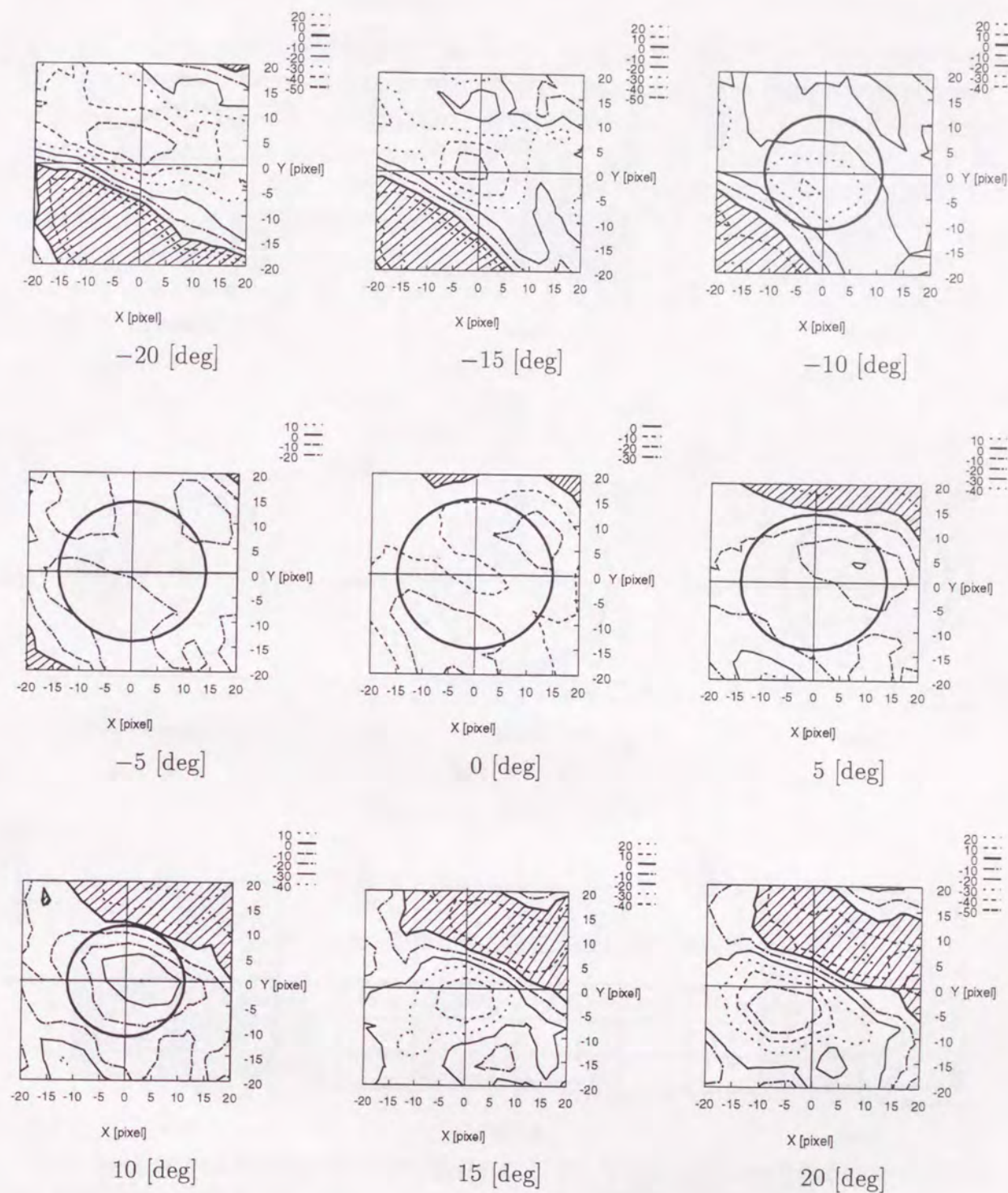


Fig. 4.16: dU/dt の分布図 (96 x 96)

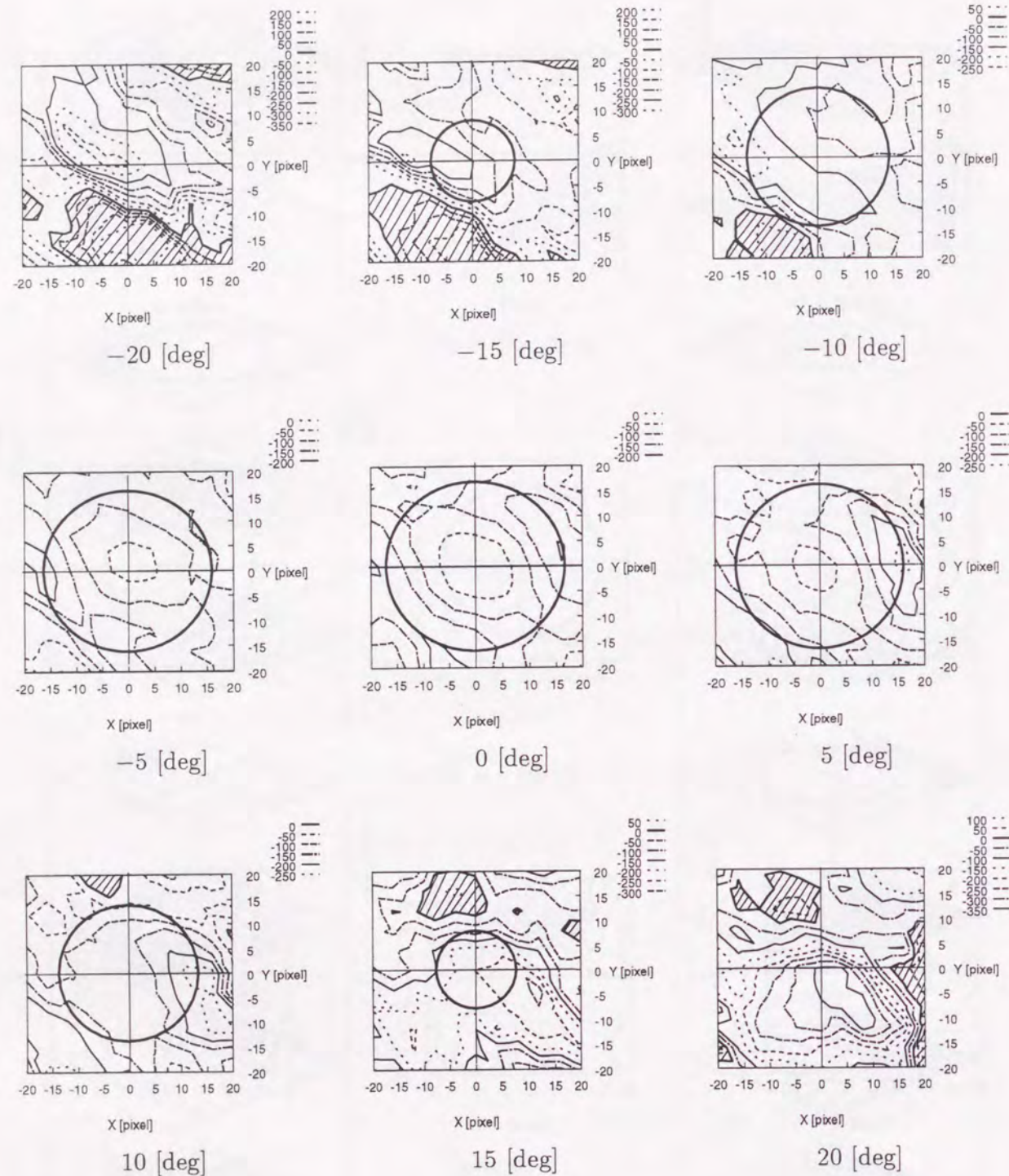


Fig. 4.17: dU/dt の分布図 (12 × 12)

4.3.7 ビジュアルサーボ実験

前述した画像の並進回転量検出法に基づいて、位置決めビジュアルサーボの実験を行った。まず、実験システムの構成について説明する。Fig.4.18に実験システムの概観を示す。マニピュレータ（ムーブマスタ RV-M2：三菱電機（株）製）の手先にFig.4.5に示す写真を固定し、500[mm]上方から CCD カメラで撮影した。画像の取り込みには TRP-IMG（（株）コンカレントシステムズ製）を用いて、写真上の 63 × 53[mm] の領域を 144 × 144[pixel] の画像データ（0～255 の濃淡値データ）として 60Hz 周期で取り込んだ。画像上の 1[pixel] は写真上で 0.44 × 0.37[mm] に相当する。画像処理とマニピュレータの制御にはトランスピュータ（インモス T 8 0 5：（株）コンカレントシステムズ製）を用いた。各処理に要した時間を Tab.4.2 に示す。

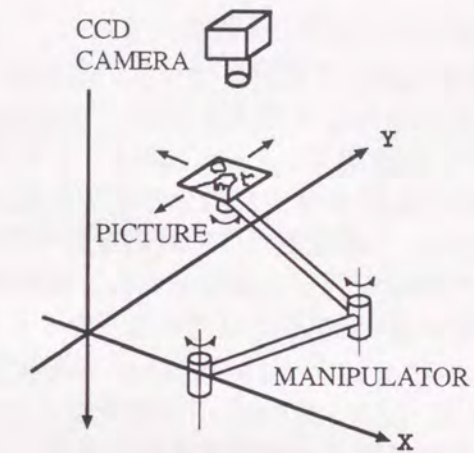


Fig. 4.18: 実験システム

Tab. 4.2: 各解像度で画像処理に要する時間

(単位 [msec])	12 × 12[pixel]	24 × 24[pixel]	48 × 48[pixel]
階層画像作成	10	7	29
並進回転量推定	3	20	95

マニピュレータは各関節を角速度制御することで、手先に固定した写真を任意の並進速度、回転速度で動作させた。なお、マニピュレータの位置決め精度は約 0.3[mm] である。

次に具体的な実験方法について説明する。まず、写真を原点 ($x = y = \theta = 0$) に固定した状態で画像を取り込み、48 × 48, 24 × 24, 12 × 12[pixel] の各階層画像を作成する。画像の階層化は、3 × 3, または 2 × 2[pixel] の画素の濃度値の単純平均を用いて行った。次に、写真の位置と姿勢の初期誤差を与えた後に画像を取り込み、(4.7) 式による並進量と回転量

の推定値にスカラーゲインを乗じて、写真の並進速度と回転速度の指令値を生成した。はじめは 12×12 [pixel] の解像度で制御を行い、推定並進量が x, y ともに ± 0.1 [pixel] 以内かつ、推定回転量が ± 2 [deg] 以内になったときに、解像度を 24×24 [pixel] に上げて制御を行い、さらに、推定並進量が x, y ともに ± 0.1 [pixel] 以内かつ、推定回転量が ± 1 [deg] 以内になったときに、解像度を 48×48 [pixel] に上げて制御を行った¹。

並進量 $(10, 10)$ [mm], 回転量 -10 [deg] を初期誤差として与えた場合の実験結果を Fig. 4.19, 4.20, 4.21 に示す。図中の縦線は解像度が切り替わったタイミングを示している。粗い画像より段階的に解像度を切り替えながら位置決めが行われている様子がわかる。 48×48 画像 (1[pixel] のサイズは 1.3×1.1 [mm]) を用いた場合の最終的な位置決め誤差は $(0.0, 0.0)$ [mm], 0.3 [deg] であり、サブピクセル精度の位置決めが実現できていることがわかる。また、 12×12 [pixel] の解像度の画像を用いた場合でも、かなりの精度で移動量を検出していることがわかる。

次に 12×12 [pixel] の解像度だけを用いる場合、または、 12×12 と 24×24 [pixel] の解像度を用いて制御を行う場合について実験を行った。本実験ではカメラと画像との距離、および写真を変更し、 200×286 [mm] の写真を 144×144 [pixel] の画像データとして取り込んだ。取り込んだ画像上の 1[pixel] は、 1.4×1.3 [mm] に相当する。初期誤差として、並進移動量 $(7, 7)$ [mm], 回転量 -7 [deg] を与えた場合の実験結果を Fig. 4.22, 4.23, 4.24 に示す。図中、実線は 12×12 [pixel] の解像度だけで制御を行った結果であり、点線は 12×12 と 24×24 [pixel] の解像度で制御を行った結果である。解像度は推定並進量が x, y ともに ± 0.1 [pixel] 以内かつ、推定回転量が ± 2 [deg] 以内になったときに切り替えた。 12×12 [pixel] の解像度のみを用いた場合と 12×12 と 24×24 [pixel] の解像度を用いた場合の最終的な位置決め精度を Tab. 4.3 に示す。 12×12 [pixel] の解像度より 24×24 [pixel] の解像度のほうが精度良く位置決めが行えていることが確認できる。また、どちらの解像度においても約 0.1 [pixel] 以下の精度で位置決めが行えていることがわかる²。

Tab. 4.3: 最終的な収束誤差

解像度 (1[pixel] のサイズ)	x [mm] ([pixel])	y [mm]([pixel])	θ [deg]
24×24 (8.4×7.8 [mm])	0.9 (0.11)	0.2 (0.026)	0.0
12×12 (16.8×15.6 [mm])	1.7 (0.10)	-0.9 (-0.06)	0.4

¹解像度を切り替えるタイミングは、現在は経験的に定めている。タイミングが早すぎる（または遅すぎる）と、解像度に対する誤差量が大きすぎて（または小さすぎて）推定誤差の精度が悪くなる可能性がある（4.3.5節参照）。また、4.3.6節に示したように、収束動作の漸近安定領域は、解像度が高くなるにつれて狭くなるために、切り替えるタイミングが早すぎると漸近安定が保証されなくなる可能性もある。そのため、切り替えるタイミングはやや遅めに設定したほうが（閾値を小さく設定したほうが）安全であるが、誤差に対する解像度が低すぎると、推定誤差量が小さくなる傾向があるために収束時間が長くなる傾向がある。画像に適した解像度切り替えのタイミングの設定は今後の課題である。

²本実験では、ロボットマニピュレータの位置決め精度が限られているために、 48×48 の画像を用いた実験を行っていないが、高精度な制御が可能なマニピュレータで、より高解像度の画像で制御を行えば、さらに高精度な位置決めが行えると予想される。

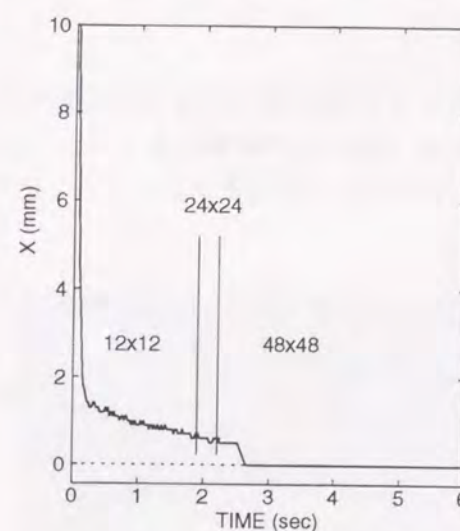


Fig. 4.19: X座標の変化

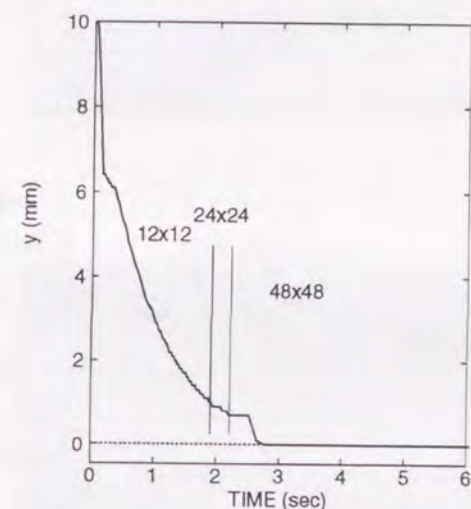


Fig. 4.20: Y座標の変化

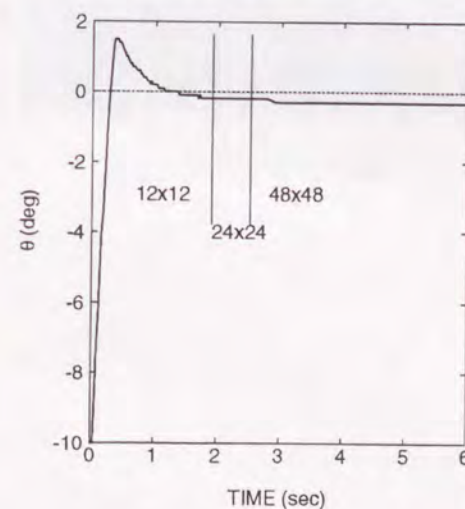
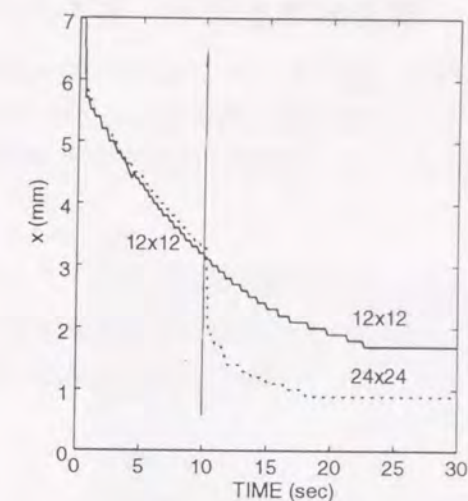
Fig. 4.21: θ の変化

Fig. 4.22: X座標の変化

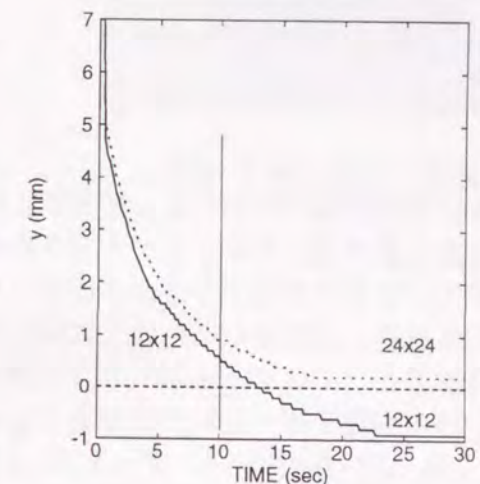
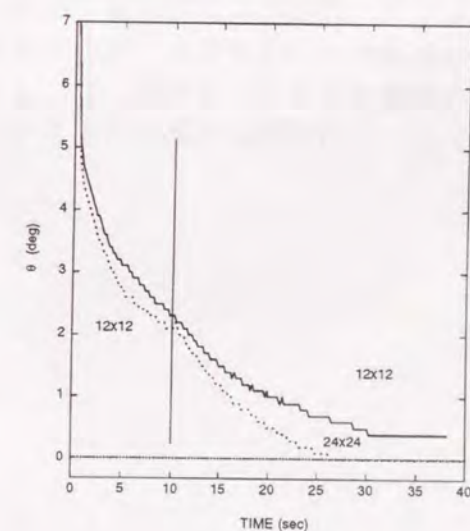


Fig. 4.23: Y座標の変化

Fig. 4.24: θ の変化

4.3.8 環境の明度変化に対するロバスト性

本節では、粗いオプティカルフローが環境の変化による画像全体の明度変化に対してもロバストに平面運動を検出できることを説明する。(4.4)式の S'_x の計算において、 $\Delta_x > 0$ の領域と $\Delta_x < 0$ の領域に分割すると次式のようにあらわすことができる。

$$S'_x = - \sum_{\Delta_x > 0} \Delta_t + \sum_{\Delta_x < 0} \Delta_t$$

ここで、環境の明度変化により目標画像に対して観測画像全体の濃度が c だけ大きくなったと仮定すると、目標画像と観測画像の濃度差 Δ_t も c だけ大きくなるので、上式は次式のようになる。

$$\tilde{S}'_x = - \sum_{\Delta_x > 0} (\Delta_t + c) + \sum_{\Delta_x < 0} (\Delta_t + c)$$

ここで、 $\sum_{\Delta_x > 0}$ と $\sum_{\Delta_x < 0}$ の領域面積が等しいと仮定すると、

$$\begin{aligned} \tilde{S}'_x &= - \sum_{\Delta_x > 0} (\Delta_t) + \sum_{\Delta_x < 0} (\Delta_t) \\ &= S'_x \end{aligned}$$

のように画像濃度の変化による増分値 c がキャンセルされて、画像明度が変化しない場合と同じ推定量が得られることが予想される。 $\sum_{\Delta_x > 0}$ と $\sum_{\Delta_x < 0}$ の領域面積は、ある程度の濃淡の変化が存在する画像であればほぼ等しいと考えられる。そこで実際に環境明度を変化させて並進量、回転量がどのように推定されるかを実験した。

目標画像としてはFig.4.25に示す画像を用いた。環境を明るく、または暗くした状態でFig.4.25の画像を取り込んだ画像をFig.4.26, Fig.4.27に示す。それぞれの画像の平均濃度は、165, 137, 106である。Fig.4.26, Fig.4.27の画像をソフト的に $-6 \sim 6$ [pixel]の範囲で並進させたときの推定並進量をFig.4.28, Fig.4.30に示す。また、 $-6 \sim 6$ [deg]の範囲で回転させたときの推定回転量をFig.4.29, Fig.4.31に示す。これらの図より画像全体の濃度が変化しても、並進および回転量はほぼ変わらずに計測されることがわかる。しかし、並進量、回転量が0の場合でも、微小ながら並進量、回転量が計測されるため、収束位置には微小な誤差が生じる。これは、 $\sum_{\Delta_x > 0}$ と $\sum_{\Delta_x < 0}$ の領域面積が微小に異なるためであると考えられる。この問題は次節で示すように空間的な微分画像を用いることで解決できる。



Fig. 4.25: 基準画像



Fig. 4.26: 明るい画像



Fig. 4.27: 暗い画像

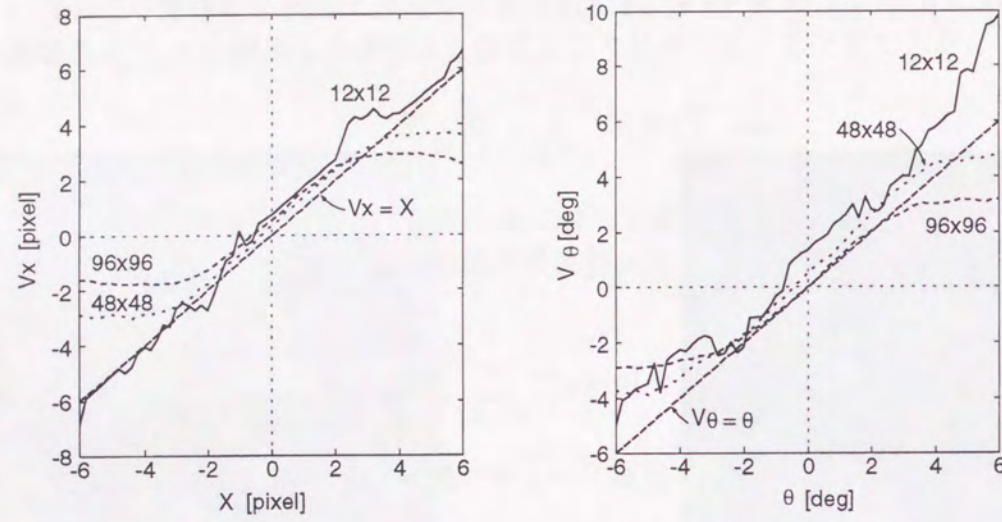


Fig. 4.28: 明るい画像での推定並進量 Fig. 4.29: 明るい画像での推定回転量

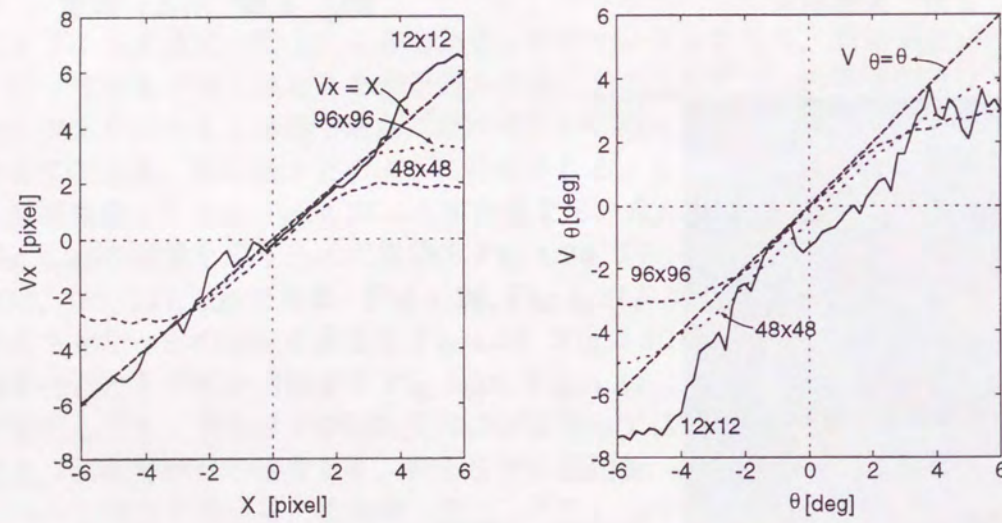


Fig. 4.30: 暗い画像での推定並進量 Fig. 4.31: 暗い画像での推定回転量

原画像を用いた場合の推定結果

4.3.9 エッジ抽出を行った画像への適用

前節まで、オプティカルフローの検出には CCD カメラより取り込んだ画像をそのまま用いていた。しかし、従来の画像処理で行われている各種のフィルターをかけた画像に対しても本手法はそのまま適用することができる。そこで、本節では環境の明度変化にロバストな画像処理として空間 1 次微分を施した画像に対して並進量、回転量の推定実験を行った。空間 1 次微分は Fig.4.32 に示すオペレータをマスク処理することによって行った。前節と同じ条件で、空間 1 次微分を施した画像に対して行った実験結果を Fig.4.33 ~ Fig.4.36 に示す。環境明度が変化しても、空間 1 次微分を施した画像ではほとんど画像濃度は変化しないので、明るい画像を用いた場合と暗い画像を用いた場合でほとんど同じ結果が得られている。また、前節の実験で検出された 0 点における並進量、回転量の検出誤差も検出されおらず、空間 1 次微分を施した画像に対する運動量の推定は環境の明度変化に対して大変ロバストであることがわかる。

0	-1	0
-1	0	1
0	1	0

Fig. 4.32: 空間 1 次微分オペレータ

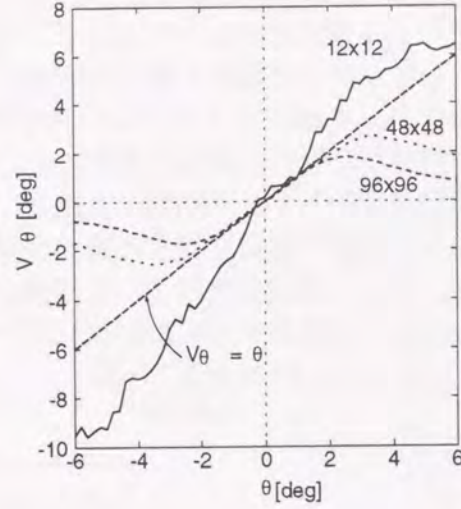
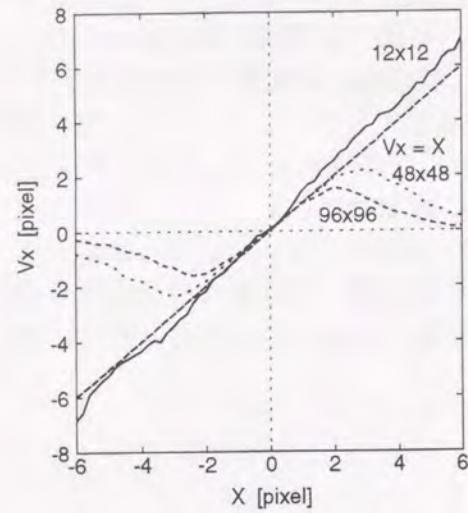


Fig. 4.33: 明るい画像での推定並進量 Fig. 4.34: 明るい画像での推定回転量

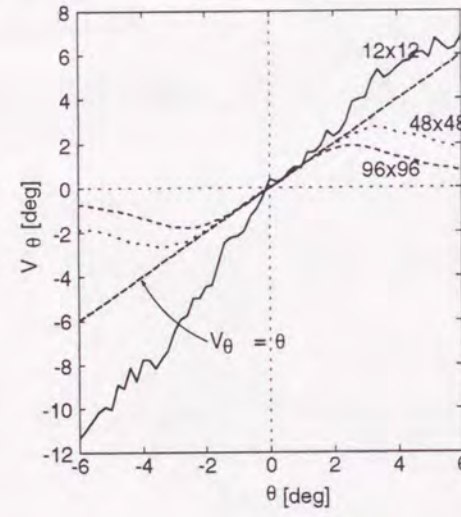
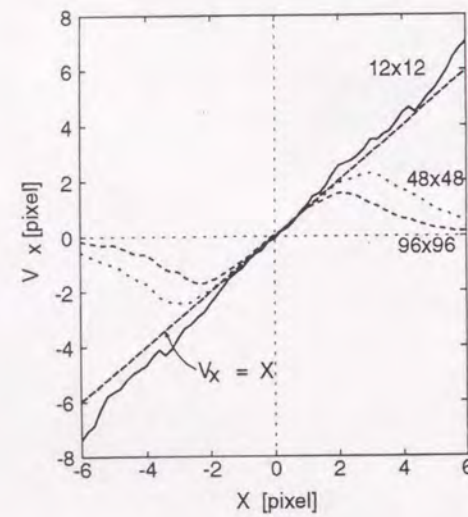


Fig. 4.35: 暗い画像での推定並進量 Fig. 4.36: 暗い画像での推定回転量

一次微分画像を用いた場合の推定結果

4.3.10 おわりに

本節では、計算量の少ないオプティカルフローの検出方法として、'粗いオプティカルフロー'を提案し、画像の並進量と回転量が、近似的なモデルに基づく簡素な画像処理を用いても、環境の明度変化に対してロバストでかつ高速に検出できることを示した。また、画像解像度を段階的に変化させながら、本手法で得た画像のずれ量に基づいてビジュアルサーボを行うことで、サブピクセル精度の位置と姿勢に関する位置決め制御が高速に行えることを示した。本手法は、実画像を直接用いて計測を行うので、画像の図形的特徴量（重心、エッジなど）を用いる方法のように、事前に画像の特徴を設定したり、対応点問題を考慮する必要がない。また、テンプレートマッチング法に比べると計算量が少ないので、特徴量の設定が難しい大きな画像（写真や印刷物など）を高精度に位置決めするようなタスクにおいて特に効果を発揮すると考えられる。

本手法の最大の問題点は、収束安定領域が画像により異なる点である。オプティカルフローを検出するためには、画像全体にある程度の濃度変化が存在しなければならない。また、この濃度変化は各階層画像（各解像度の画像）においてそれぞれに存在しなければならない。例えば、細かい模様はあるが全体的にはのっぺりとした画像（例えば格子状の模様）が画面全体にあるような画像の場合には、本手法の適用可能な誤差範囲は狭いと思われる。また、本手法を適用した場合の漸近安定領域は、目標画像をソフトまたはハード的にシフト、回転させて誤差量の検出を実際に行って確認する以外、現状では確認する手段はない。画像が与えられたときに、少ない計算量で漸近安定領域を調べる手法の開発が今後の課題である。また、本研究では、カメラに対して物体を移動させて位置決めするタスクを扱ったが、画像中のテンプレートを検出した移動量に基づいて動かすことで、画像中の動物体の追跡のようなタスクにも応用できると考えられる。

4.4 結言

本章では、ビジュアルサーボに適した画像特徴の抽出法として、粗いオプティカルフローを用いた平面運動の検出法を提案し、ビジュアルサーボに適用することで近似的なモデルに基づく小さい計算量の簡素な画像処理によっても、高精度に並進と回転の位置決めを行うことができることを示した。ビジュアルサーボにおいては、高速な画像処理が必要とされるが、一方で本章に示したように簡素な画像処理で得た画像特徴を用いても、十分な制御性が得られる特徴がある。しかし、現在のビジュアルサーボの研究の多くは、画像処理の手間を少なくするために、実験室内の整備された環境においてしか行われておらず、実際の複雑な環境においてビジュアルサーボを適用することに重点をおいた研究は少ない。出口らは、実画像を固有空間法によって次元圧縮することでカメラ運動と画像情報を結合し、ビジュアルサーボを行う手法を提案した [59]。本手法は実画像を直接的に用いたビジュアルサーボを目指したもので、概念的には大変興味深いものであるが、莫大な計算量が必要であるといった欠点がある。しかし、本章に示したようにビジュアルサーボは近似的な大雑把なモデルを用いた画像特徴量であっても十分な制御が可能であるので、出口らのようなアプローチも計算量が少なくなるように十分に簡易化ができれば有用となるであろう。

今後のビジュアルサーボの研究では、タスクに応じて適切な特徴量を如何に抽出するかといった実環境における具体的な応用が重要となると思われる。

第 5 章 結論

本研究では、ハンドアイシステムにおける視覚系と運動系の融合について以下の点を明らかにした。

1. 人間の眼と腕に似せた運動学的構造を持つハンドアイシステムにおいては、従来、非線形と考えられてきたステレオ画像とロボットの関節角度のマッピングが時不変な線形マッピングとして近似できる。
2. 1. における時不変線形マッピングの近似精度について、カメラの配置とマニピュレータの運動学的構造を評価すると、最も精度良く近似が行える構造は、人間の眼と腕の構造と似たものである。
3. 通常、ビジュアルサーボにおいては、カメラの姿勢とマニピュレータの関節角度を実時間で計測して、適切なゲイン（ヤコビアン）を計算する必要があるが、1. の性質を利用すると、これらを計測せずにゲインを時不変としても十分に広い領域で動作可能なシステムが構成できる。また、時不変ゲインを用いたビジュアルサーボは、制御式が単純で計算量が少なく、パラメータ変動（とくにカメラ回転）に対してロバストである。
4. ビジュアルサーボにおけるパラメータ変動や時間遅れに対するロバスト性は、カメラ情報とマニピュレータの関節角度のマッピングのヤコビアンの空間的な一様性を示す‘ヤコビアンの時不変安定度’で評価することができる。
5. 4. の‘ヤコビアンの時不変安定度’を用いてカメラ配置とロボットアームの運動学的構造を評価すると、最も広い領域でヤコビアンが一様である構造（パラメータ変動や時間遅れに対するロバスト性が広範囲で得られる構造）は人間の眼と腕の構造と似たものである。
6. ビジュアルサーボに用いる画像特徴の検出法としては、正確な環境復元よりも、目標とのずれ量が小さい計算量で高速に求まる手法が適切である。粗いオプティカルフローを用いた平面運動の検出方法は、近似的なモデルに基づいた大雑把な手法であるが、ビジュアルサーボサーボと組み合わせることによって平面上の物体の位置と姿勢を高速かつ高精度に位置決めすることができる。

以下、それぞれについて考察と今後の課題を記す。

1. について ハンドアイシステムの制御においては、カメラ情報よりデカルト空間へのマッピング（3次元復元）とデカルト空間よりロボットの関節角度へのマッピング（キネ

マティクス)を用いて、デカルト空間を介したマッピングを基礎とするのが一般的であるが、本研究ではハンドアイシステムのマッピングにおいては変換誤差(キャリブレーション誤差)が存在することを前提として、カメラ情報よりロボットの関節角度への直接的な時不変線形マッピングを用いた点が従来にない新しい概念である。カメラ情報からロボットの関節角度への直接的なマッピングを試みた研究は古くから存在しており、ニューラルネットワークを用いる手法や、メモリーベースのマッピングを用いる手法が見られる。しかし、これらの研究のほとんどは、制御対象が未知であったり、制御対象が時間的に変動する場合の学習方法の開発を目的としたものであるため、学習速度や学習精度といったマッピングの効率は議論されても、マッピングに適した制御対象についての議論は見られない。本研究で示した線形近似マッピングの近似精度は決して良いものではないが、マッピングが線形的である場合には、ニューラルネットワークを用いたマッピングを行う場合にも、その収束速度が速いことが報告されている。すなわち、制御に適したマッピングを行うためには、マッピングの手法の選択と同時に制御対象をいかに設計するかが重要であることを本結果は示している。

2. について 人間における視覚情報と上肢運動感覚のマッピングに関する研究は、生体・生理工学の分野で古くから盛んに行われている。しかし、本研究で示したような時不変線形マッピングは、工学的には計算量が少ない、単純であるといった利点はあっても、生体の仕組みを説明する場合には、生体の構造が非常に複雑であり、そのほとんどが非線形的な要素で構成されているなどの理由によって、このようなアプローチはほとんどない。しかし、先に述べたようにマッピングが線形的であることは、他のマッピングの手法を用いる場合でも収束速度が速いなど多くの利点を持つと考えられる。また、フィードバック制御のゲインを時不変とすることができる。生体におけるフィードバックのゲインが時変であるか時不変であるかは、マッピングの仕組みとは別に取り扱うことが可能であり、生体・生理学的にも興味のある問題である。本研究では、上肢を単純な2リンクモデルとして平面上のマッピングのみを取り扱ったが、3次元またはさらに多次元のマッピングを行う場合に同様の手法が適用できるかどうかは、今後の課題である[61]。また、本研究ではハンドアイのマッピング、すなわちカメラ情報とロボットの関節角度のマッピングを取り扱ったが、他のセンサに関するマッピング、たとえば手先の力覚と関節トルクなどのマッピングも同様に線形マッピングとして取り扱うことができるかどうか興味ある課題である。
3. について 本研究では時不変ゲインを用いたビジュアルサーボがカメラ回転に対してロバストであることを示したが、具体的にどのようにカメラを回転させればいいのかは述べていない。しかし、人間において眼球の回転は空間知覚を担う重要な要素であることが知られており、ロボットビジョンにおいてもカメラ回転を利用した画像処理(アクティブビジョン)の研究が盛んであることを考えると、カメラ回転に対してロバストであることは有用であると思われる。ビジュアルサーボに適したカメラ回転は、画像特徴の検出法とも密接に関連すると思われるので重要な課題である。

6. について オプティカルフローは従来計算量が膨大であり、リアルタイムのフィードバック制御には適さないとされていたが、検出精度を犠牲にすると少ない計算量でしかも最終的な位置決め精度が優れた特徴量の抽出法と成りうることを示した。本手法は計算量の少なさと対応点問題を解く必要がない点が利点であり、広い領域の位置決めが要求されるタスク(たとえば印刷物の位置決め、および印刷ミス抽出など)に有用であると思われる。本研究ではカメラを固定した状態で目標物体を位置決めするタスクを扱ったが、他にも、テンプレートマッチングのように画像中の目標物体の追跡などのタスクにも応用可能であると思われる。

謝辞

本研究に際して終始一貫御指導、貴重な御助言ならびに暖かい励ましの御言葉を頂きました宮崎文夫教授、和歌山大学 丸典明講師に心から御礼申し上げます。そして、論文作成にあたりましては、新井健生教授ならびに田中正夫教授から多くの御意見を賜りました。厚く御礼申し上げます。また、本研究を進めるにあたり、升谷保博講師、周桑完助手、西川敦助手には多くの御助言、御教示をいただきました。さらに、実験設備の構築においては、林清重技官に御指導をいただきました。また、三菱電気(株)長沢 雅人さんにはカメラシステムの製作のためにモーターや各種部品、資材を提供していただきました。また、富士川和延さん(現:三菱電機(株))、小荒健吾さん(本学大学院生)、宮崎陽司さん(本学部生)には研究全般にわたって御協力いただきました。また、財団法人浦上奨学会からは、博士課程就学において大きなご援助をいただきました。心から御礼申し上げます。

最後に、終始ご援助を頂きました宮崎研究室の皆様方、いつも励ましを与えてくれた友人達、そして最後まで暖かく見守って頂いた父、母、妹に心から感謝いたします。

満田 隆

1998年1月

付 録

A ステレオカメラの運動学

本章では、ステレオカメラの画像座標とデカルト空間座標の変換について説明を行う。なお本章は [60] を参考としている。

Fig.A.1に示す3つの座標系を定義する。 Σ_W は環境中に固定されたワールド座標系、 Σ_R はロボットの頭部に固定されたロボット座標系、 Σ_C はカメラプラットフォームに固定されたカメラ座標系で、全て右手系の3次元直交座標系とする。 Σ_R, Σ_C の x 軸はステレオカメラの基線と一致するものとし、基線の midpoint をこれらの座標系の原点とする。また、ステレオカメラの両視軸はカメラ座標系 Σ_C の xz 平面内で運動するものとする。ロボットの頭部は、環境に対して1軸の回転と並進の自由度を持つとし、ロボット座標系の、ワールド座標系に対する y 軸まわりの回転角を ζ 、 x 軸方向への並進距離を d とすると、ロボット座標系 Σ_R からワールド座標系 Σ_W への同時変換行列 ${}^W\vec{T}_R$ は次式となる。

$${}^W\vec{T}_R = \begin{bmatrix} \cos \zeta & 0 & \sin \zeta & d \\ 0 & 1 & 0 & 0 \\ -\sin \zeta & 0 & \cos \zeta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{A.1})$$

また、カメラ座標系の、ロボット座標系に対する x 軸まわりの回転角を α_D とすると、カメラ座標系 Σ_C からロボット座標系 Σ_R への同時変換行列 ${}^R\vec{T}_C$ は次式となる。

$${}^R\vec{T}_C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha_D & -\sin \alpha_D & 0 \\ 0 & \sin \alpha_D & \cos \zeta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{A.2})$$

まず、環境中のある点をステレオカメラの視線が注視しているとき、カメラの姿勢からその視線の交点 ${}^W\vec{P}$ の座標を求める。左右のカメラの焦点位置と回転中心は一致するものとし、カメラ座標系において、これらの回転中心をそれぞれ、 ${}^C\vec{R}(E, 0, 0)$ 、 ${}^C\vec{L}(-E, 0, 0)$ 、回転角を α_R 、 α_L とする。ここで、 $2E$ は基線長である。このとき、カメラ座標系にお

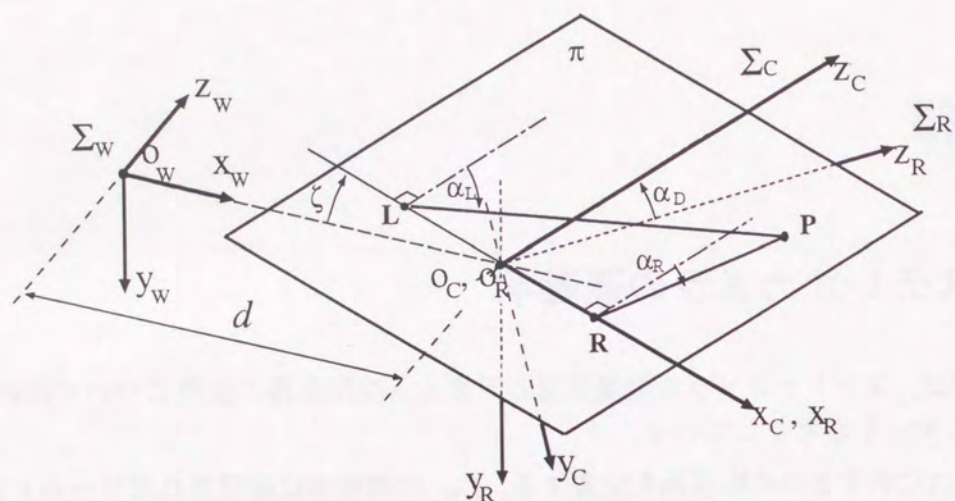


Fig. A.1: ステレオカメラの幾何学

ける注視点 ${}^C\vec{P}({}^C x_g, {}^C y_g, {}^C z_g)$ をカメラ回転角 α_R, α_L で表すと,

$${}^C\vec{P} = \begin{bmatrix} {}^C x_g \\ {}^C y_g \\ {}^C z_g \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{E(\tan \alpha_R + \tan \alpha_L)}{(\tan \alpha_R - \tan \alpha_L) - 2E} \\ \frac{\tan \alpha_R - \tan \alpha_L}{-2E} \\ \frac{\tan \alpha_R - \tan \alpha_L}{1} \end{bmatrix} \quad (\text{A.3})$$

となる.

また, ${}^W\vec{P}$ は同時変換行列 ${}^W\vec{T}_R, {}^R\vec{T}_C$ を用いると,

$${}^W\vec{P} = {}^W\vec{T}_R {}^R\vec{T}_C {}^C\vec{P} \quad (\text{A.4})$$

となる. 従って, 環境中に固定された注視点 ${}^W\vec{P}$ はロボットの関節変数 $(\alpha_R, \alpha_L, \alpha_D, \zeta, d)$ により (A.3), (A.4) 式で表すことができる.

次に, 点 ${}^W\vec{P}$ を注視するロボットの関節変数を ${}^W\vec{P}$ により表すことを考える. しかし, 空間中の 1 点を注視することを目的とする場合には, このビジョンロボットは 2 自由度冗長である為に, 逆運動学の解は無限に存在し, 拘束条件を与えなければ解くことができない. 頭部の回転角 ζ , 並進量 d は与えられるものとして $(\alpha_R, \alpha_L, \alpha_D)$ を注視点の座標

${}^R\vec{P}({}^R x_g, {}^R y_g, {}^R z_g)$ と ζ, d を用いて表すと,

$$\begin{aligned} \alpha_R &= \tan^{-1} \left(\frac{({}^R x_g - E) \cos \alpha_D}{{}^R z_g} \right) \\ \alpha_L &= \tan^{-1} \left(\frac{({}^R x_g + E) \cos \alpha_D}{{}^R z_g} \right) \\ \alpha_D &= \tan^{-1} \left(-\frac{{}^R y_g}{{}^R z_g} \right) \end{aligned} \quad (\text{A.5})$$

となる. ここで, ${}^R\vec{P}$ は

$${}^R\vec{P} = {}^W\vec{T}_R^{-1} {}^W\vec{P} = \begin{bmatrix} \cos \zeta & 0 & -\sin \zeta & -d \cos \zeta \\ 0 & 1 & 0 & 0 \\ \sin \zeta & 0 & \cos \zeta & -d \sin \zeta \\ 0 & 0 & 0 & 1 \end{bmatrix} {}^W\vec{P} \quad (\text{A.6})$$

である.

よって, 頭部の回転角 ζ および並進量 d が与えられるとき, 空間中の 1 点 ${}^W\vec{P}$ を注視するカメラの回転角 $(\alpha_R, \alpha_L, \alpha_D)$ が決まる.

また, Fig.Aに示すように, カメラ座標系 Σ_C において, ステレオカメラが点 $\vec{P} = ({}^C x_g, 0, {}^C z_g)$ を注視しているとき, 空間中のある点 $\vec{Q} = ({}^C x, {}^C y, {}^C z)$ の画像中の投影位置は,

$$\begin{aligned} X^R &= f \frac{(x - E) \cos \alpha_R - z \sin \alpha_R}{(x - E) \sin \alpha_R + z \cos \alpha_R} \\ Y^R &= f \frac{y}{(x - E) \sin \alpha_R + z \cos \alpha_R} \\ X^L &= f \frac{(x + E) \cos \alpha_L - z \sin \alpha_L}{(x + E) \sin \alpha_L + z \cos \alpha_L} \\ Y^L &= f \frac{y}{(x + E) \sin \alpha_L + z \cos \alpha_L} \end{aligned} \quad (\text{A.7})$$

となる. ここで f は焦点距離であり, $(\alpha_R, \alpha_L, \alpha_D)$ は (A.5) 式により求まる.

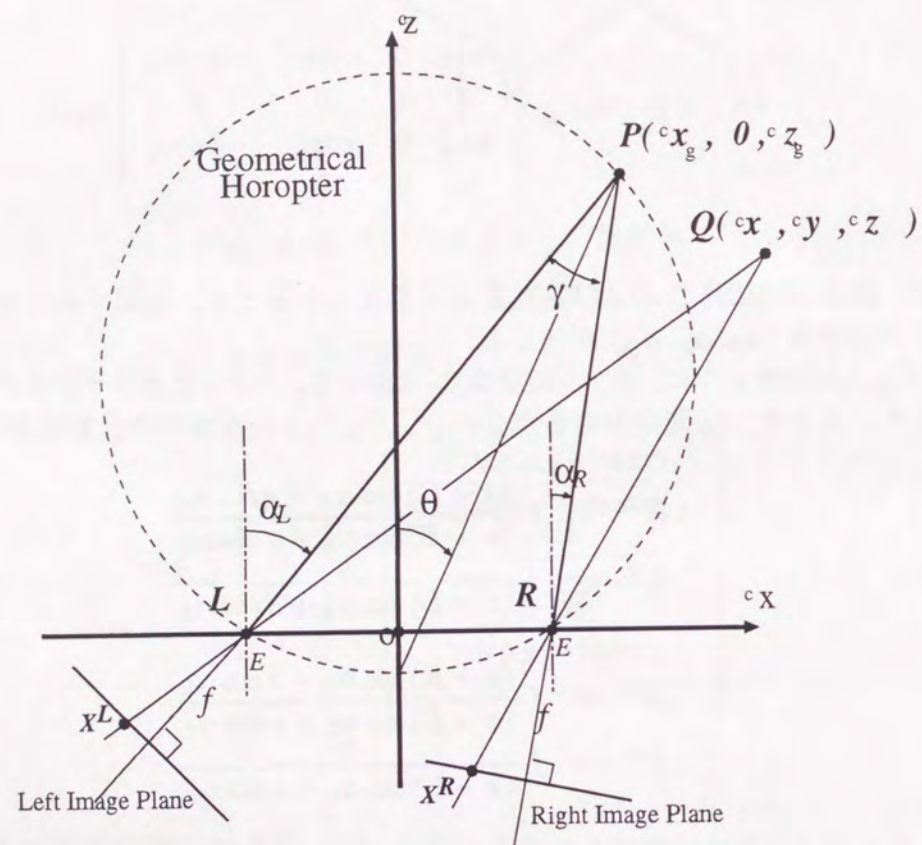


Fig. A.2: 注視におけるステレオカメラの幾何学

Tab. 5.1: サフィックスの定義

座標系	添字
ワールド座標系 : Σ_w	w (World)
カメラ座標系 : Σ_c	c (Camera)
左画像座標系 : Σ_l	l (Left)
右画像座標系 : Σ_r	r (Right)
仮想画像座標系 : Σ_I	I (Image)

B 疑似逆行列を用いたステレオビジュアルサーボ

本章では、3章の実験およびシミュレーションにおいて、比較対象として用いた『疑似逆行列を用いたステレオ視によるビジュアルサーボ』 [35] について説明する。 [35] では、カメラをマニピュレータの手先に搭載したシステムを想定しているが、本研究は手と眼が分離したシステムであるため、以下に示すような方法を用いた。

B.1 ステレオ視のモデル

ステレオ視における各座標系を定義し、ステレオ視のモデルについて説明する。また、各座標系における値は、表 5.1 の添字により表現することにする。ここで用いたステレオ視のモデルは、一般に標準配置のステレオ視モデルと呼ばれるもので、以下のような座標系を定義する。図 B.3 に示すように、2 台のカメラ中心を基線長 $2E$ だけ離し、それらの光軸の方向が基線に対して直角になるように配置し、2 台のカメラ間の中心にカメラ座標系 Σ_c の原点をとる。このとき、カメラ中心から光軸方向に焦点距離 f だけ離れた位置で光軸と直角に交わる平面が画像平面であり、光軸と画像平面の交点を左右それぞれ左画像座標系 Σ_l 、右画像座標系 Σ_r の原点とする。また、ワールド座標系 Σ_w は、適当な位置に原点がとられているものとする。各座標系の x, y, z 軸のとり方は、図 B.3 に示す通りとする。

左右の画像座標値をカメラ座標系の値を用いて表現すると次式を得る。

$$X^L = f \frac{c_x + E}{c_z} \tag{B.8}$$

$$X^R = f \frac{c_x - E}{c_z} \tag{B.9}$$

$$Y^L = f \frac{c_y}{c_z} \tag{B.10}$$

$$Y^R = f \frac{c_y}{c_z} \tag{B.11}$$

これがステレオ視のモデルとなる。

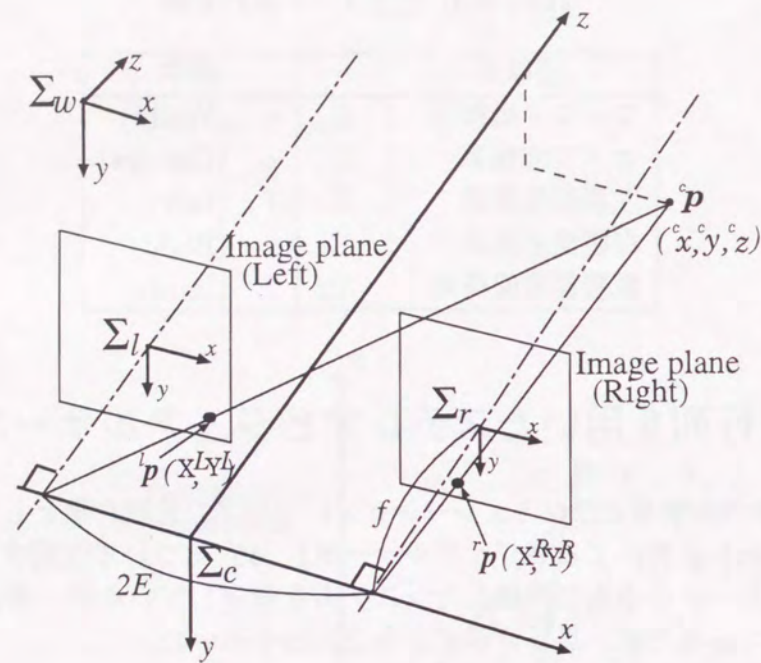


Fig. B.3: ステレオカメラのモデル

B.2 イメージャコピアン

いま、対象物体上の特徴点 p を考え、その画像中の座標を (X^L, Y^L) , (X^R, Y^R) とする。これら左右の画像座標をまとめて仮想的な画像座標系 I を定義し、画像中の特徴点の現在位置を次式のように記述する。

$${}^I p = \begin{pmatrix} X^L \\ X^R \\ Y^L \\ Y^R \end{pmatrix} \quad (\text{B.12})$$

画像中の特徴点速度 ${}^I \dot{p}$ とカメラ座標系から見た特徴点速度 ${}^c \dot{p}$ の間には、その相互関係を表すヤコビ行列 ${}^I J_c$ を用いて、次式の関係が成立する。

$${}^I \dot{p} = {}^I J_c {}^c \dot{p} \quad (\text{B.13})$$

B. 疑似逆行列を用いたステレオビジュアルサーボ

ただし、

$${}^I J_c = \begin{pmatrix} \frac{\partial X^L}{\partial c_x} & \frac{\partial X^L}{\partial c_y} & \frac{\partial X^L}{\partial c_z} \\ \frac{\partial X^R}{\partial c_x} & \frac{\partial X^R}{\partial c_y} & \frac{\partial X^R}{\partial c_z} \\ \frac{\partial Y^L}{\partial c_x} & \frac{\partial Y^L}{\partial c_y} & \frac{\partial Y^L}{\partial c_z} \\ \frac{\partial Y^R}{\partial c_x} & \frac{\partial Y^R}{\partial c_y} & \frac{\partial Y^R}{\partial c_z} \end{pmatrix} \quad (\text{B.14})$$

である。これにステレオ視のモデル (B.8)~(B.11) 式を代入すると以下ようになる。

$${}^I J_c = f \begin{pmatrix} 1 & 0 & -\frac{c_x + E}{c_z} \\ \frac{c_z}{1} & 0 & -\frac{c_x - E}{c_z} \\ 0 & 1 & -\frac{c_y}{c_z} \\ 0 & \frac{c_z}{1} & -\frac{c_y}{c_z} \end{pmatrix} \quad (\text{B.15})$$

${}^c R_w$ をワールド座標系からカメラ座標系への回転行列とすると、画像中の特徴点速度 ${}^I \dot{p}$ とワールド座標系における特徴点速度 ${}^w \dot{p}$ の間には、次式の関係が成立する。

$${}^I \dot{p} = {}^I J_c {}^c R_w {}^w \dot{p} \quad (\text{B.16})$$

${}^I J_c$ をイメージャコピアンと呼ぶ。

B.3 フィードバック入力の生成

マニピュレータの手先速度と関節速度を関係づける行列をロボットヤコビアンと呼ぶ。マニピュレータの手先に特徴点を配置すると、ワールド座標系における特徴点速度 ${}^w \dot{p}$ と関節角速度 \dot{j} の間には、ロボットヤコビアン ${}^w J_j$ を用いた次式の関係が成立する。

$${}^w \dot{p} = {}^w J_j \dot{j} \quad (\text{B.17})$$

(B.13) 式を代入すると、画像中の特徴点速度 ${}^I \dot{p}$ と関節角度 \dot{j} の間には、次式の関係が成立する。

$${}^I \dot{p} = {}^I J_c {}^c R_w {}^w J_j \dot{j} \quad (\text{B.18})$$

画像中の特徴点位置 ${}^I p$ と画像中の目標位置 ${}^I p_d$ より、マニピュレータ関節へのフィードバック入力 u は次式のように生成する。

$$u = -\lambda {}^w J_r^{-1} {}^c R_w^{-1} {}^I J_c^+ ({}^I p - {}^I p_d) \quad (\text{B.19})$$

${}^I J_c^+$ は、イメージャコピアンの疑似逆行列

$${}^I J_c^+ = ({}^I J_c^T {}^I J_c)^{-1} {}^I J_c^T \quad (\text{B.20})$$

であり、 λ はゲインである。

参考文献

- [1] 川辺, 新井: “ロボットコントローラの現状と動向”, 日本ロボット学会誌, Vol.14, No.6, pp.764-767, 1996.
- [2] 松島: “オフライン教示”, 日本ロボット学会誌, Vol.13, No.5, pp.611-614, 1995.
- [3] 谷内田: “ロボットビジョンの概要と今後の展望”, 日本ロボット学会誌, Vol.10, No.2, pp.140-145, 1992.
- [4] 高橋: “ロボットビジョンのためのカメラキャリブレーション”, 日本ロボット学会誌, Vol.10, No.2, pp.177-184, 1992.
- [5] 佐藤, 長谷川: “視覚フィードバックによるロボットの制御”, 計測と制御, Vol.26, No.2, pp.122-127, 1987.
- [6] 橋本: “視覚フィードバック制御-静から動へ”, システム/制御/情報, Vol.38, No.12, pp.659-665, 1994.
- [7] Y.Shirai, H.Inoue: “Guiding a Robot by Visual Feedback in Assembly Tasks”, Pattern Recognition, Vol.5, pp.99-108, 1973.
- [8] P.I.Corke: “Visual Control of Robot Manipulator - A REVIEW”, Workshop of the 1994 IEEE International Conference on Robotics and Automation, 1994.
- [9] 谷内田 正彦: “ロボットビジョン” 人工知能シリーズ 11, 昭晃堂, 1990.
- [10] 吉川 恒夫: “ロボット制御基礎論” コンピュータ制御機械システムシリーズ 10, コロナ社, 1988.
- [11] Roger Y.Tsai, Reimar K.Lenz: “A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration”, IEEE Transactions on Robotics and Automation, Vol.5, No.3, pp.345-358, 1996.
- [12] 橋本: “ビジュアルサーボイング”, 計測と制御, Vol.35, No.4, pp.282-285, 1996.
- [13] K.Hashimoto: “Visual Servoing”, World Scientific, 1993.
- [14] S.Hutchinson: “A Tutorial on Visual Servo Control”, IEEE Transactions on Robotics and Automation, Vol.12, No.5, 1996.

- [15] K.Hosoda,M.Asada: "Versatile Visual Servoing without Knowledge of True Jacobian" Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Vol.1, pp.186-191, 1994
- [16] 加瀬, 丸, 西川, 宮崎: "ステレオ視によるマニピュレータのビジュアルサーボ", システム制御情報学会誌, Vol. 6, No. 8, pp. 360-367, 1993.
- [17] N. Maru H. Kase S. Yamada A.Nishikawa and F. Miyazaki: "Manipulator Control by Visual Servoing with the Stereo Vision," Proceedings of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Yokohama, Japan, Vol. 3, pp. 1865-1870, 1993.
- [18] G. Hager, W. Chang, and A. Morse. "Robot Hand-Eye Coordination Based on Stereo Vision", IEEE Control Systems magazine, Vol.15, No.1, pp.30-38, 1995.
- [19] 前田・館: "視覚性到達運動における両眼視と上肢位置感覚の統合", 計測自動制御学会論文集, Vol.29, No.2, pp.201-210, 1993.
- [20] 磯部・川人・鈴木: "両眼網膜座標における手の繰り返し学習制御", 電子通信学会技術研究報告, MBE86-80, pp.17-24
- [21] 笠井: "眼と手の協調動作と空間知覚" 電子情報通信学会誌, Vol.76, No.22, pp.1156-1163, 1993.
- [22] R.Caminiti, P.B.Johnson, A.Urbano: "Making Arm Movements Within Different Parts of Space Dynamic Aspects in the Primate Motor Cortex" The journal of Neuroscience, Vol.10(7), pp.2039-2058, 1990.
- [23] Michael S.A.Graziano, Charles G.Gross: "A bitmodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields" Experimental Brain Research, Vol.97, pp.96-109, 1993.
- [24] R.K.Luneburg: "The Metric of Binocular Visual Space" Journal of the Optical Society of America, Information and Communication Engineers, Vol.40, No.10, pp.627-642, 1950.
- [25] 福田: "生体情報システム論" 産業図書, 1995.
- [26] W.Thomas Miller, Filson H.Glanz, L.Gordon Kraft: "Application of a General Learning Algorithm to the Control of Robotic Manipulators" The International Journal of Robotics Research, No.6, pp.84-98, 1987.
- [27] Michael Kuperstein: "Newral Model of Adaptive Hand-Eye Coordination for Single Posture" Science, Vol.139, pp.13080-1311, 1988.

- [28] 新井: "多関節マニピュレータ制御における演算処理手法" 計測自動制御学会論文集, Vol.18, No.1, 1982.
- [29] Martha Flanders, Stephen I.Helms Tilley, John F. Soechting: "Early stages in a sensorimotor transformation" Behavioral and brain science, Vol.15, pp.309-362, 1992.
- [30] P.K.Allen,B.Yoshimi,A.Timucenko: "Real-time visual servoing" Proceedings of the 1991 IEEE/RSJ International Conference on Robotics and Automation, pp.851-856, 1991.
- [31] A.J.Koivo and N.Houshangi: "Real-time vision feedback for servoing robotic manipulator with self-tuning controller" IEEE Trans.Systems,Man,and Cybernetics,21(1), pp.134-142, 1991
- [32] D.B. Westmore and W.J.Wilson: "Direct dynamic control of robot using an endpoint mounted camera and Kalman filter position estimation" Proceedings of the 1991 IEEE/RSJ International Conference on Robotics and Automation, pp.851-856, 1991.
- [33] L.E.Weiss,A.C.Sanderson,C.P.Newman: "Dynamic sensor-based control of robotics with visual feedback" IEEE J. Robotics and Automation, RA-3(5), pp.404-417, 1991.
- [34] K.Hashimoto et al: "Manipulator control with image-based visual servo" Proceedings of the 1991 IEEE/RSJ International Conference on Robotics and Automation, pp.2267-2272, 1991.
- [35] N.Maru, H.Kase et al: "Manipulator control by visual servoing with the stereo vision" Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.1865-1870, 1993.
- [36] B. Espiau, F. Chaumette and P. Rives: "A New Approach to Visual Servoing in Robotics" IEEE Trans. Robotics and Automation, RA-8-3, 313-326, 1992.
- [37] J.E. Slotine, Weiping Li: "Applied nonlinear control" Prentice Hall, 1991.
- [38] 橋本・海老根・木村: "ビジュアルサーボイング最適制御によるアプローチ" 日本ロボット学会誌, Vol.12, No.5, pp.773-778, 1994
- [39] R. Sharma, and S. Hutchinson: "Optimizing Hand/Eye Configuration for Visual-Servo Systems" Proceeding of the IEEE International Conference on Robotics and Automation, Vol.1, pp.172-177, 1995.
- [40] John T.Feddema, and Owen R.Mitchell: "Vision-Guided Servoing with Feature-Based Trajectory Generation" IEEE Journal of Robotics and Automation, Vol.5, No.5, pp.691-700, 1989.

- [41] 久野: “アクティブビジョン-歴史と展望-” 人工知能学会誌, Vol.10, No.4, pp.493-499, 1995.
- [42] R.Sharma: “Active Vision for Visual Servoing: A REVIEW”, Workshop of the 1994 IEEE International Conference on Robotics and Automation, 1994.
- [43] 森田, 沢崎, 内山, 佐藤: “カラートラッキングビジョン” 第14回日本ロボット学会学術講演会予稿集, Vol.1, pp.279-280, 1996.
- [44] 五味: “テンプレートマッチング法によるサブピクセル精度の追跡” 第15回日本ロボット学会学術講演会予稿集, Vol.3, pp.993-994, 1997.
- [45] 清水, 桑原: “画像間相関を利用したリアルタイム計測” 画像ラボ, Vol.5, No.1, pp.74-77, 1996.
- [46] 田中, 佐野, 大原, 奥平: “パラメトリックテンプレートとそのサブピクセルマッチングへの応用” 第2回知能メカトロニクスワークショップ講演論文集, pp.74-79, 1996.
- [47] Peter.I.Corke: “Visual Control of Robots: High-performance visual servoing” ROBOTICS AND MECHATRONICS SERIES, RESEARCH STUDIES PRESS LTD., 1996.
- [48] 飯間, 小野木, 西村: “時空間微分法によるオプティカルフローの検出可能性と3次元運動推定の一方法” 計測自動制御学会論文集, Vol.24, No.1, pp.76-82, 1988.
- [49] 中島, 安居院: “画像処理 [4]: 動画画像処理” 計測と制御, Vol.28, No.8, pp.711-718, 1989.
- [50] 三池, 古賀: “パソコンによる動画画像処理” 森北出版, 1993.
- [51] J.O.LIMB and J.A.Murphy: “Measuring the Speed of Moving Objects from Television Signals” IEEE TRANS.Commun., Vol.COM-23, No.4, pp.474-478, 1975.
- [52] P. K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman: “Real-time visual servoing” Proceeding of the IEEE International Conference on Robotics and Automation, pp.1850-1856, 1992.
- [53] Horn B.K.P and Schunk B.G.: “Determining optical flow” Artificial Intelligence, Vol.17, pp.185-203, 1981.
- [54] 吹抜: “画像信号による動画画像の移動量, 速度の測定” 信学技法, IE78-67, pp.1850-1856, 1978-10.
- [55] C.Cafforio and F.Rocca: “Methods for Measuring Small Displacements of Television Images” IEEE TRANS.Inf.Theory, Vol.IT-22, No.5, pp.573-579, 1976.

- [56] 和田, 山口: “反復勾配法による動画画像信号の動き量検出” 電子通信学会論文誌, Vol.J68-D, No.4, pp.663-670, 1985.
- [57] 佐藤, 佐々木: “動画画像における動きベクトルの階層的推定法” 電子通信学会論文誌, Vol.J69-D, No.5, pp.771-776, 1986.
- [58] 野村, 相良, 成瀬, 井手: “エッジ位置計測のサブピクセル化と誤差解析” 電子情報通信学会論文誌, Vol.J73-D-2, No.9, pp.1458-1467, 1990.
- [59] 出口: “動きと視覚の幾何学的結び付きの記述とその視覚サーボへの応用” 第1回ロボティクスシンポジウム, pp.81-86, 1996.
- [60] 田中, 丸, 宮崎: “アクティブなステレオ視覚システムによる移動物体の追跡” 大阪大学大学院基礎工学研究科修士論文, 1995.
- [61] 満田 隆, 丸 典明, 宮崎 文夫: “人間型ハンドアイシステムにおける線形ビジュアルサーボ-視空間と手先姿勢の線形マッピング-”, 第15回日本ロボット学会学術講演会予稿集, No.3, pp.999-1000, 1997.

関連文献

1. 満田 隆, 丸 典明, 富士川 和延, 升谷 保博, 宮崎 文夫: “視空間と関節空間の線形近似に基づくビジュアルサーボ”, 第12回日本ロボット学会学術講演会予稿集, no.1, pp.331-332, 1994.
2. 満田 隆: “人間型ハンドアイシステムにおける線形ビジュアルサーボ”, 大阪大学大学院基礎工学研究科修士論文, 1995.
3. 満田 隆, 丸 典明, 富士川 和延, 宮崎 文夫: “逆運動学の線形近似に基づくビジュアルサーボ”, 日本ロボット学会誌, vol.14, no.5, pp.743-750, 1996.
4. 満田 隆, 丸 典明, 富士川 和延, 宮崎 文夫: “視空間を用いた逆運動学の線形近似”, 日本ロボット学会誌, vol.14, no.8, pp.1145-1151, 1996.
5. 満田 隆, 丸 典明, 富士川 和延, 升谷 保博, 宮崎 文夫: “人間型ハンドアイシステムにおける線形ビジュアルサーボ”, 第1回ロボティクスシンポジウム予稿集, no.1, pp.75-80, 1996.
6. T.Mitsuda, N.Maru, K.Fujikawa, F.Miyazaki: “Visual Servoing based on the Use of Binocular Visual Space”, Proc. of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol.3, pp.1104-1111, 1996.
7. 満田 隆, 丸 典明, 富士川 和延, 宮崎 文夫: “視空間に基づくビジュアルサーボイング”, 計測自動制御学会論文集, vol.33, no.1, pp.35-41, 1997.
8. 満田 隆, 丸 典明, 宮崎 文夫: “定数近似ヤコビアンに基づくビジュアルサーボにおける構造と軌道安定性の解析”, 第14回日本ロボット学会学術講演会予稿集, no.3, pp.765-766, 1996.
9. T.Mitsuda, N.Maru, F.Miyazaki: “Visual Servoing Based on the Constant Linear Mapping”, Proc. of the 2nd Asian Control Conference, vol.3, pp.335-338, 1997.
10. 満田 隆, 丸 典明, 宮崎 文夫: “視空間と上肢関節空間の線形時不変マッピングに適する運動学的構造”, 第12回生体・生理工学シンポジウム論文集, pp.413-416, 1997.
11. 満田 隆, 丸 典明, 宮崎 文夫: “人間型ハンドアイシステムにおける線形ビジュアルサーボ-視空間と手先姿勢の線形マッピング-”, 第15回日本ロボット学会学術講演会予稿集, no.3, pp.999-1000, 1997.

12. 満田 隆, 丸 典明, 宮崎 文夫: “テンプレート画像の濃度差に基づいた高精度な2次元位置決め法”, 第2回知能メカトロニクスワークショップ講演論文集, pp.68-73, 1997.
13. T.Mitsuda, N.Maru, F.Miyazaki: “Binocular Visual Servoing Based on Linear Time-invariant Mapping”, Advanced Robotics,1998 (in print).
14. 満田 隆, 宮崎 陽司, 丸 典明, 宮崎 文夫: “粗いオプティカルフローを用いた高精度な位置決めビジュアルサーボ”, 日本ロボット学会誌投稿中 (1998)

