

視覚と制御

東京大学情報理工学系研究科

橋本浩一

概要: 近年の急速なコンピュータ技術の進展にともない, ビジョンセンサや画像処理システムが非常に安価に利用可能になってきた. それにともないビジョンセンサを制御システムに組み込む要求も高まってきた. しかし撮像システムの特性は高度に非線形であり, そのパラメータの同定・キャリブレーションも容易ではない. また, 各種機械システムの制御に要求されるサンプリング周波数と比較して一般の視覚装置のサンプリング周波数はかなり低い. 視覚センサを用いて制御システムを構成するには, 非線形システムの動特性を考慮して安定性の議論をおこなうことも重要な問題となってくる. 本ワークショップではそのような視覚システムの特性を明らかにした上で, その特性を生かした制御法やパラメータ誤差にロバストな制御法について議論する.

1 はじめに

人間にとって視覚・聴覚・触覚などの感覚器はとても重要である. なかでも視覚は, 外部情報の 87% を占めるといわれており, 知的情報を獲得するためのもっとも重要な感覚器であることは間違いない.

一方, さまざまなシステムの設計・運用に関しても, 視覚は重要である. 自動車運転支援・安全化・無人化, 監視システム (高速道路, 交差点, 踏み切り, ビル), 大形建設機械 (クレーン, 掘削機械) の効率化, 遠隔操縦 (ロボット, 潜水艇), 外観検査など, 多岐にわたる用途において実用化されている.

本稿においては, 機械システムにおける行動と認識に対し, 脳研究からのアナロジーに基づいてシステムのアーキテクチャを概説し, 視覚の重要性を振り返る. その後, 視覚と制御に限定して, コンピュータビジョンの基礎, ビジュアルフィードバック制御の歴史, 最近の話題について言及する.

とくに, 近年のビジュアルフィードバック制御の研究においては, 動く物体の実時間トラッキングを目標としたビジュアルサーボ法に重点がおかれている. これは作業環境を理解してからおもむくに動き出す “見てから動く” 方法とはことなり, 作業対象物とハンドの間の状態を観測しながら反射的に反応する “見ながら動く” 方法で, センサ情報処理, 軌道計画, モータ制御がすべて並列して動作するという特徴をもつ. したがって, 従来の静的な視覚フィードバック制御とその歴史, 現在の動的なビジュアルサーボ法の研究について詳しく解説する [41, 39]. なお, ビジュアルフィードバック制御に関する詳しい文献は [13, 65] にまとめられているので参照していただきたい.

2 認識と行動の情報処理

サイバネティクスはいまから約 50 年前に Wiener によって提案された「動物と機械の制御と通信に関する学問」であり, 彼は

- 動物 (人間) の行動を機械の原理との類推を用いて解明する
- 動物 (人間) に似た行動をする機械を作る

ことを目指した [122]. Wiener は神経生理学者との共同研究を通じて, 動物の随意運動の制御が機械の制御と同じフィードバックを原理としていることを発見したが, これがサイバネティクスを提案する具体的な動機となった. この問題は前者に属する研究課題であるが, 現在, 脳のモータ制御として脳生理学で活発に研究されている. また, 学習をする機械の提案は後者の研究課題の顕著な例であり, ロボティクスに限らず, 制御, 通信, 情報処理などの多くの研究分野に影響を与えた.

Wiener の視点に立ってこれからの制御工学を展望した木村による解説 [80, 79] は非常に興味深い. その中で, 木村は学習制御のアーキテクチャとして川人らのフィードバック誤差学習法 [78] を取り上げている. この学習法はシナプス可塑性の神経回路モデル [68] に基づくものであり, 生理学的に妥当性が検証されている方法である.

2.1 計算論的神経科学 [77]

人間がどのようにして視覚情報を処理し, 認識しているかについては古くから研究の対象となっている. また, 視覚情報をもとにどのような仕組みで行動が制御されているかについてもさかんに研究されている. 脳の情報処理機能を理解するために必要となる研究のレベルについて, Marr は以下のように分類している [85, 66].

計算理論 計算の目標は何か、なぜそれが適切なのか、そしてその実行可能な方略の論理は何か。

表現とアルゴリズム この計算理論はどのようにして実行することができるのか、とくに入力と出力の表現は何か、そして変換のためのアルゴリズムは何か。

ハードウェアによる実現 表現とアルゴリズムがどのようにして物理的に実現されるか。

もっとも高いレベルでは、脳における計算の目的は何かという抽象的なレベルである。人間のなめらかな運動を実現するために、脳は何を入力として、何を出力として、どのような計算をしているのかを理解する必要がある。次のレベルでは、その計算にあたっての、入力と出力（さらに、あればそれらの中間）の表現を見極めなければならない。そして入力を出力に変換するための具体的なアルゴリズムを明確にする必要がある。ここでの入出力表現およびアルゴリズムは一意的ではない。同じ結果を得るためにはいろいろなアルゴリズムが考えられる。そしてもっとも具体的なレベルでは、上述のアルゴリズムを実行するハードウェア、すなわち神経回路モデルの研究が必要となる。

このようにして提案された計算理論は心理実験、行動実験を通して検証される。たとえば、錯視によって視覚認識のモデルを検証したり、感覚器からの入力を絶って運動させることにより運動制御モデルにおけるフィードバックの役割を検証したりする。一方、神経回路モデルは生理実験などにより検証される。脳内の情報表現とアルゴリズムに関する仮説を、脳内の信号を見ることなく、脳をブラックボックスとして外部からの観察だけで行うのはかなりの困難がともなう。さらに神経回路モデルを実験的に検証するには、脳内の実際の神経回路はニューロンの電気的な活動の計測に基づくことが必要になる。小脳のパーセプトロンモデルは小脳のシナプス可塑性の仮説「小脳プルキンエ細胞の登上線維と平行線維のふたつの入力系に同時に入力が入ると、2種類のシナプス間で干渉が起こり、平行線維とプルキンエ細胞間のシナプス伝達の効率が可塑的に変化する」のもとに提案された有名なモデル [84, 1] であるが、シナプス可塑性の検証には生理学的な実験が必要となる。

2.2 空間知覚 [127]

体操の選手は空中で身体をいろいろの向きに回転させながら、鉄棒をつかんだり飛び越えたりという運動を正確に行うことができる。また、テニスの選手は俊敏な身のこなしで高速のサーブのボールをとらえ、相手コートのコーナーに正確に決めることができる。人間はこの

ように激しい運動をしているときでも、周囲の世界がどのようにになっているかを正確に認識する能力と、周囲の世界に対して正確に運動を制御する能力をもっている。

一方で、ビデオカメラをでたらめに振り回して撮像すると映像は非常に乱れ、カメラの周囲の世界がどのような構造であったかを知ることは困難となる。目や頭が動いているとき、網膜から映像を受け取る脳もまったく同じ状況に直面しているはずである。

この能力の違いは、情報処理の過程で感覚情報と運動情報の統合がなされているか否かに帰することができる。つまり、脳は網膜に写った像のみから周囲の世界を理解しているのではなく、目や頭などビジョンセンサの搭載されている本体の動きの情報と網膜像の情報を統合して周囲空間の構造を脳の中に再現していく。そうすることによってビジョンセンサに対する相対的な姿勢ではなく、ビジョンセンサの姿勢に依存しない絶対的な姿勢として周囲の世界を把握することが可能になる。

2.2.1 視覚性空間

視野内のいろいろな対象物に対して次々と視線の方向を切り替える眼球運動をサッケードという。これは単に網膜上の像と網膜中心の位置のずれを解消する眼球運動ではなく、対象物の空間的位置をまず把握し、その位置に対する眼球運動命令を生成している。

Zee らが与えたサッケードのモデルは以下の通りである [128]。脳は網膜からの視覚情報とそのときの頭を基準とした眼の方向から、頭に対する目標の方向を計算する。約 0.2 秒の潜時の後、この目標に対するサッケードを行う。サッケードの開始時には、その時点における視線の方向と目標の方向を比較し、その偏差が十分小さくなるまでサッケードの運動指令を出し続ける。

このモデルはいくつかの点で工学的にも興味のある考え方を含んでいる。第一は、脳が眼の瞬時瞬時の方向をモニタしていて、ゴールの方向と一致するかどうかを比較しながら運動命令をつくるという点、第二は、眼の運動方向という時間的情報と網膜像という空間的情報の統合をしている点である。

眼球の運動命令は、「この速度で動け」という速度指令と、それを神経回路で積分してつくった「この位置変化にしたがって動け」という位置指令から構成されていることは生理学的にもほぼ確定している [105]。時々刻々の眼球方向を表す信号は、この位置信号の一部（脳から運動機関へと向かう遠心性信号のコピー: efference copy）が使われているとする考え方と、眼筋の自己受容器 (proprioceptor) の信号が用いられているという 2 説がある。

たとえば、外眼筋受容器からの求心性神経（脳へと向かう神経: afferent neuron）を切断した実験、病気で外眼筋

の自己受容器の機能が一時停止している患者が視標の位置を手で指示する実験，外力で眼の回転に強制的に変位を加えて視標を手で指す実験などから総合的に判断すると，視覚目標の位置知覚の形成には，外眼筋の自己受容器の信号が用いられ，眼球運動の制御には遠心性コピーの信号が用いられているようである．

これらの信号の違いをまとめてみると，遠心性コピーでは

- 眼球運動命令の一部であり，自己受容器信号に比べて遅れない
- 両眼の同じ方向への動きの成分のみを表している
- 眼の頭に対する絶対的な方向ではなく，現在の眼の方向に対する相対的な角度を表す
- 遠心性コピーの信号が両目の動きを正しく表現できるためには眼筋の自己受容器信号を参照して常時学習する必要がある

一方，自己受容器信号では

- 10ms 前後の遅れがある
- 各眼の頭に対する絶対的な角度を算出できる情報を含む
- 両眼の輻輳角，したがって視覚対象物までの距離を算出できる情報を含む

といえる．

2.2.2 視覚性空間と運動性空間の関係

視覚・眼球運動系が獲得した対象物の位置情報は，視線サッケードのため頭と眼球の運動発現に用いられると同時に手の運動発現にも用いられる．手の運動系においても視覚・眼球運動系からの位置情報だけでなく，自己受容器情報に基づいて物体や身体の一部に触れる非視覚的对象物への運動を発現する．

長時間自分の手を見る機会を奪われると，視覚目標を開ループ的に(視覚フィードバックなしに，すなわち自分の手を見ずに)指先で指示する位置精度が明らかに低下する．ところがサッケードや，鼻頂に手を触れる動作のように各系内で閉じている動作はその影響を受けない．つぎに，プリズムを通して右または左にシフトして見える物体に手を伸ばした場合，手を見ることができなければ物体に触れることはできない．しかし，プリズムを通して自分の手を見る視覚体験をつむことにより新たな眼と手の位置の対応関係を学習することができる．これらの実験結果は，視覚・眼球運動系と手動作系の間で配置の対応関係が可塑的であり相互にドリフトすること，手を

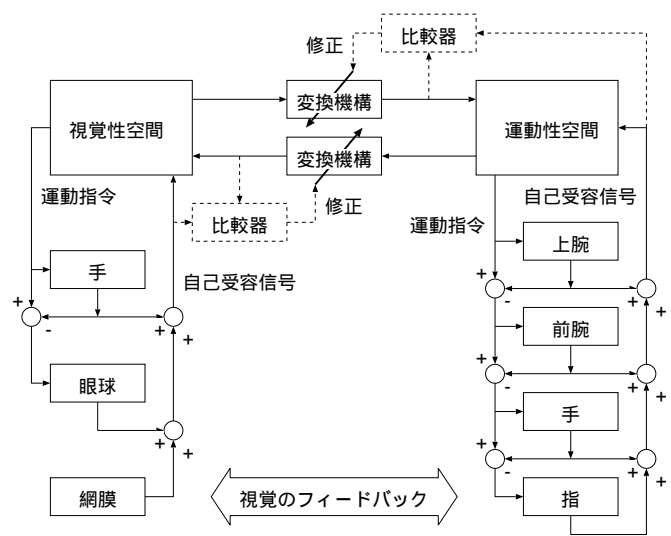


Fig. 1: Visuo-Motor Control

目視確認することによって相互の位置の対応関係が校正されることを意味する．この考えをもとに，視覚系と運動系の関係をブロック図に示す．

右の手動作系では，中枢から指先に対して出された運動命令の一部を体幹側にある上腕が実行し，残りの動きのうちさらに一部を前腕が実行する．つまり，より体幹側が実行した動きを次々と補償する形で末端の運動が実行される．腕や手の動きは自己受容器で検出され，中枢に身体イメージを報告すると同時に，より体幹側の動きを補償する信号として用いられる．左の視覚・眼球運動系についても同様である．

この両系の間で配置情報を伝達する経路が双方向に存在し，視覚目標に対する手動作や，自分の指先を注視する眼球運動の発現に供される．両系の対応関係が可塑的であったり，ドリフトしたりすることは，両系を結ぶ経路の中間にある機構の入出力関係を調整する仕組みによって表現されている．正しい入出力関係は，手を目視確認する視覚フィードバックにより保持される．

2.3 認識行動システムのアーキテクチャ

2.3.1 視覚の階層化

Marr は視覚の役割を以下の三段階に分類している [85, 66] ．

初期視覚 網膜に投影された 2 次元画像から外界の 3 次元空間の表面を推測する

中間視覚 さまざまな視覚情報を統合して $2\frac{1}{2}$ 次元スケッチをつくる

高次視覚 3 次元のモデルに基づいてパターン認識を行う

1970年以降、生理学的実験を通して脳のハードウェアとしての研究が長足の進展を遂げた。並行して、たくさんの心理実験、行動実験によって脳をブラックボックスと見立ててその機能を解明しようとする研究も進展した。そして現在では Marr の理論は厳しい理論的、実験的な批判にさらされているが、それでも Marr の研究は、視覚認識に関してはじめて理論的な体系を構築したという意味で、功績は偉大である。

以下、Marr の分類に基づいてコンピュータシステムにおける視覚認識について考えてみる。

初期視覚 第一の段階は、2次元の画像から3次元の構造を推測する問題である。たとえば、画像の中の輪郭線を抽出するという問題は人間にとっては簡単である。1960年代に人工知能の研究者たちがコンピュータを使ってこの問題に取り組んだが、コンピュータで解くのは非常に難しいことがわかった。その根本的な原因は画像が作られる過程で3次元の情報が2次元に落とされているにもかかわらず、元の世界を推測しようとしているからである。もちろん、元の世界に関する何らかの先見の情報がなければこの問題は不良設定であり解けないわけであるが、人間は意識するしないに関わらず膨大なデータベースを持っており、それと照合することによって断片的な観測からでも元の世界の情報を容易に復元できる。画家エッシャーが描いたような、実在しそうで存在し得ない物体、だまし絵、錯視図形などにだまされるのはこのためである(もちろんこの説明はあまりにも一元的で、不十分である。脳における情報処理はもっと多元的で複雑であることが実験的に検証されている [4])。積み木の世界のようなかなり制限的な仮定においても、コンピュータによる処理では撮像装置の解像度、ノイズ、影、オクルージョン(ある物体がある物体を遮蔽すること)などが障害となり、やはり難しい。

さらに Marr は、視覚系はモジュール構造をしていると主張している。明るさ、色、形、動き、テクスチャーなどを別々に処理するモジュールがあると考え。あるひとつのモジュールが種類の情報に基づいて3次元復元を行うことを初期視覚と名付けた。コンピュータによる画像処理においても、これらのさまざまな情報を統合的に取り扱うのは困難で、やはり画像から何らかの特徴量を抽出するわけであるが、この過程においてもモジュール化が行われる。

中間視覚 第二の段階は、初期視覚で計算された各モジュールの出力を全体として統合した2次元画像と3次元立体モデルの中間的な表現(2 $\frac{1}{2}$ スケッチ)である。この表現は、面の奥行きと向きの両方を、面が観察者からどれくらい離れているか、面の向きが観察者の視線と比べ

てどの方向にあるかという、観察者中心の座標系で表現するものである。色、形、動き、陰影、テクスチャーなどを別々に処理する複数の初期視覚モジュールの出してきた結果を統合し、うまく組み合わせることでこの2 $\frac{1}{2}$ スケッチを作り出す機能を、Marr は中間視覚と呼んだ。

コンピュータビジョンの分野においては、環境認識のためのデータ表現について万能の確立された方法はまだ提案されていない。積み木の世界においてはシーンの解釈に有効な表現として線画が用いられ、多くの成果を修めたが、3次元立体モデルの認識にとくに有効な表現ではない。奥行き分布図(depth map)は、画像中の各画素に対する視線方向の距離であり、距離計測装置(range finder)からの出力としては自然な形式であるが、この形式は物体が回転した場合の変換が複雑になる。また、照度差を用いたステレオから得られる表現は、各画素に対応する表面の向きを与えるが、この表現(法線分布図: needle diagram)も表面の回転に関して簡潔な表現を与えるものではない。一方、たとえば、拡張ガウス像(extended Gaussian image)は法線分布図や奥行き分布図から容易に計算でき、物体の認識および3次元空間における姿勢の決定に有利な方法として知られている [60]。そしてこの表現は視線方向に対する面の法線の角度に面の面積情報を付与したものであり、一種の2 $\frac{1}{2}$ スケッチといえなくもない。

高次視覚 第三の段階はパターン認識などの高次視覚に関するものである。Marr は一般化円錐と呼ばれる円錐と円柱を拡張したような3次元部品をいくつか組み合わせ、3次元物体のモデルを作り、それをもとにしてパターン認識、あるいは記憶を行うと考えた。

先に述べたように、2次元の画像から3次元の構造を推測しなければならないので、視覚の問題は難しい。ランダムドットステレオグラムにおける対応づけ、陰影からの構造復元において光源方向情報の欠落による複数解の存在など、脳にとっても難しい。コンピュータによる解釈では、対応づけを簡単にするために対象物の連続性の仮定をおき、光源の方向を既知としても、パターン認識は現在も大きな研究課題である。

2.3.2 感覚情報と運動情報の相互作用

生理学的構造を考慮にいれた感覚情報と運動情報との相互作用の階層モデルがDeanによって提案されている [20]。そのモデルのブロック図を Fig.2 に示す。この制御機能の表現は神経系の構造と直接対応をもつわけではないことを注意しておく。運動制御に関して、機能ごとに制御レベルが階層化されている。もっとも下の層には特定の効果器(生物物理的性質を含む)が、もっとも上の層には動作計画が配置されており、感覚器からの信号(求心

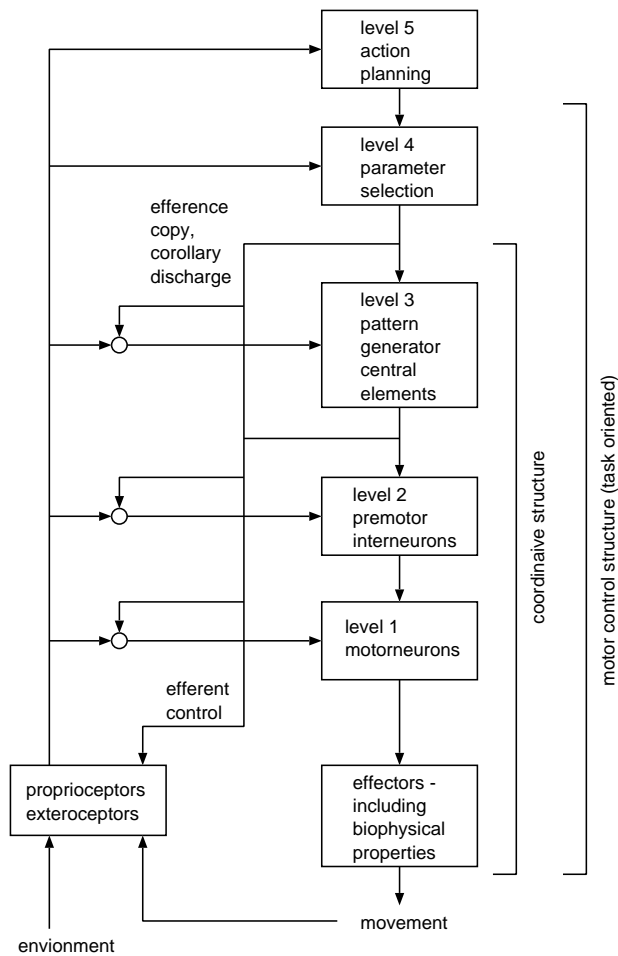


Fig. 2: Efferent/afferent interaction model

性信号: afferent information) とパターンジェネレータへのコマンドのコピー (遠心性コピー: efferent copy) がそれぞれの階層に作用する。

一番下位の構造レベルは反射を表す。感覚神経線維 (自己受容器: 空間における生体の位置と姿勢を感受する。筋、腱、関節の受容器。proprioceptor および外受容器: 生体の近傍の外界の情報を受感受する。皮膚受容器。exteroceptor) の活動が運動ニューロンに直接渡される。この結合は反射の選択から活動レベルの詳細な調整まで多岐にわたる働きをする。また、この感覚器と運動ニューロンの直接結合により、けがを避けるときなどにあらわれるすばやい反応が可能となるが。

運動前介在ニューロンは発信器 (パターンジェネレータ) と運動ニューロンの間の中間的なレベルである。このレベルはすべてのシステムに存在するわけではないが、時間的・空間的な積分効果を表す。つまり、運動前介在ニューロンは瞬間的な感覚情報による反応を引き延ばしたり、感覚情報をさまざまな用途に適した情報に変換したりするのに用いられる。これによって反応の組み合わ

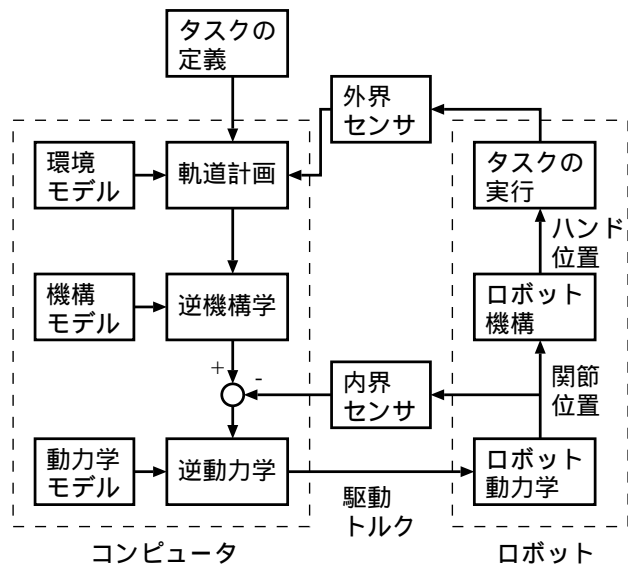


Fig. 3: Hierarchical Sensor-Based Control

せを増加させることができる。

パターンジェネレータへの感覚情報は、制御系への入力情報として筋骨格系の状態を知らせるもので、四肢に何か負荷が加わって質量が増加したときには、それに合わせたパターンを生成する。

運動系の高位のレベルにおいて、感覚情報は適切な運動制御構造を選択し、そのパラメータの決定に用いられる。この機能は外環境に対して適切な感覚情報を抽出する機能をもつ。そのためには人間の内部的表現や人間と環境との相互作用に相当する複雑な信号処理が必要になる。

3 階層的な制御

従来のロボットは内界センサの信号のみにもとづいて制御されていたので、ロボットの機構モデルや環境のモデルに未知または不正確な要素があればタスクの実行に障害が生じた。したがってロボットの自律的な制御のためにはタスクの実行状況をフィードバックするための外界センサが必要になる。通常内界センサにもとづくモータ制御機構と外界センサにもとづく軌道生成機構は階層的に構成される (Fig.3)。つまり、アームやモータの慣性などの動的なモデルにもとづいてモータの制御を行うジョイントレベル、ロボットの機構モデルにもとづいて目標軌道の変換を行うハンドレベル、ハンドと作業対象物との関係を外界センサを用いて計測し、作業のモデルにもとづき新しいハンドの軌道を生成するタスクレベルの3レベルが階層的に構成される。

実際に対象物との接触をとまなうような作業では外界

センサとして力センサが用いられる。しかし対象物の位置が不正確にしかわからない場合には、その位置を知るために非接触センサが必要になる。視覚はもっとも汎用的かつ手軽な非接触外界センサである。視覚情報から物体の位置・姿勢を求めることはタスクの実行状況を知るためにはもっとも基本的なことからである。このような研究は従来から盛んに行なわれてきたが、ほとんどの研究は複雑なシーンから対象物を認識し作業状況を理解することを目的とした静的なものであった。このような人工知能的な手法は認識・理解に時間がかかるため、対象が動くような動的な環境においての使用には適していない。

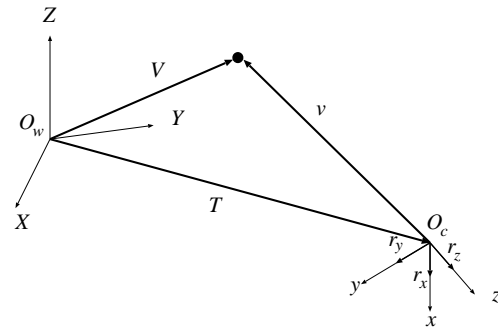


Fig. 5: World and Camera Coordinates

4 視覚の基礎

4.1 画像生成過程と画像入力

カメラでは前方の3次元の世界がレンズを通して2次元の画像へと投影され、像を結ぶ。では画像はどのように生成され、計算機に入力されるのであろうか。画像生成過程には2つの重要な問題がある。

- 3次元空間のある点の画像上での位置はどのような要因で決定するのか。
- 3次元空間のある表面の画像上での明るさはどのような要因で決定するのか。

これらの問題の答を得るには、画像投影 (image projection) と画像の放射測定 (image radiometry) の知識が必要となる [60]。放射測定に関しては文献 [60, 23] を参照していただくことにして、画像投影について説明する。

4.1.1 透視投影 [23]

対象とその画像への結像の位置関係を一番単純に理想化したモデルがピンホールカメラ (pin-hole camera) である。ピンホールカメラでは、単レンズ系のレンズ中心にピンで穴を開け、ピンホールを通して入射する光のみが画像面に到達するカメラである。Fig.4 に概略を示す。このピンホールの位置を光学中心 (optical center) といい、光学中心を通りスクリーンに垂直な直線を光軸 (optical axis) という。また、空間上の点と光学中心を結ぶ直線を視線 (line of sight) という。

Fig.4(a) では、画像面での像が逆さまになっておりまぎらわしいので、(b) のように、ピンホール前方に画像面を移動して考える。

さらに、数学的定式化のため、画像面上に2次元座標系 (x, y) 、カメラには光学中心に原点をもつ3次元座標系 (X, Y, Z) を考える。画像面では水平方向右向きに x 軸、垂直方向上向きに y 軸をとるのが自然である。カメラ座

標系 (camera coordinates) もそれに合わせて、 X, Y 軸を設定する。さらに、座標系は右手系が便利であるから、 Z 軸は光軸とは逆向きにとる。光学中心 O から画像面までの距離はレンズの焦点距離に相当する。一般的にはレンズの焦点距離を f とおくが、画像面の座標が $Z = -f$ となり表記上面倒なので、焦点距離に相当する定数 f を負と定義し、画像面の座標を $Z = f$ と定める。こうするとレンズの前方の点の座標は $Z < 0$ となることに注意する。このとき、カメラ座標系において (X, Y, Z) にある点は画像座標系において

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \frac{X}{Z} \\ f \frac{Y}{Z} \end{bmatrix} \quad (1)$$

に写像される。この変換のことを透視変換 (perspective transformation) という。透視変換は透視射影 (perspective projection) または中心射影 (central projection) と呼ばれることもある。

4.1.2 射影近似 [73]

透視変換はカメラ座標系の Z 成分 (奥行き: depth) が分母にきており、非線形である。このことにより、注视着ている対象の奥行き変化の範囲が小さい場合、その変化を用いて3次元の再構成をする際に解が不安定になることがある。逆に、対象物の大きさがカメラから物体までの距離と比較して十分小さい場合には、透視変換は線形な変換で近似できる。

線形近似は、度合いに応じて弱透視射影 (weak perspective projection) と疑似透視射影 (paraperspective projection) にわけられる。平行射影 (orthographic projection) は、弱透視射影の特殊なケースであり [3, 126]、アフィン射影 (affine projection) はこれらの線形化射影を一般化したものである [89]。詳細は参考文献に譲る。

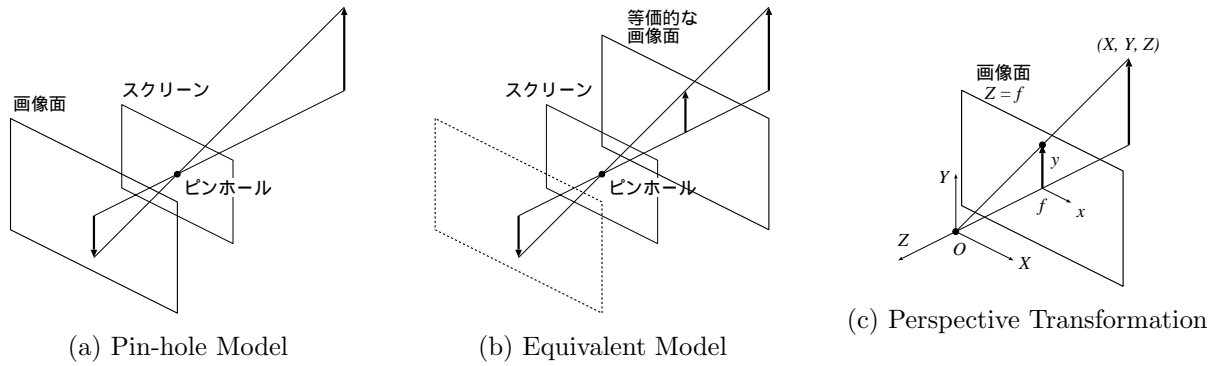


Fig. 4: Pin-hole and Perspective Models

4.2 画像とカメラ位置・姿勢の関係

カメラの外に固定された基準座標系を考える．これをワールド座標系 (world coordinates) と呼ぶ．Fig.5 にワールド座標系 $O_w - XYZ$ とカメラ座標系 $O_c - xyz$ を示す．ワールド座標系においてベクトル T で表される位置にカメラ座標系の原点 O_c があり，カメラ座標系の x, y, z 軸方向の単位ベクトルはワールド座標系においてそれぞれ r_x, r_y, r_z であるとする．

カメラ座標系においてベクトル $v = (x, y, z)^T$ で与えられる点は，ワールド座標系では

$$V = xr_x + yr_y + zr_z + T = Rv + T \quad (2)$$

となる．ここで

$$R = [r_x \ r_y \ r_z] \quad (3)$$

は回転行列 (rotation matrix) と呼ばれ， $R^T R = I$ をみたす．ただし I は 3×3 の単位行列である．また，この行列 R は，あるベクトルのカメラ座標系での表現 (たとえば (2) 式の v) をワールド座標系での表現 (Rv) に変換する座標変換行列 (coordinate transformation matrix) となっている．

4.3 同次変換

撮像過程や位置・姿勢の関係を簡潔に表現するために同次変換を用いる．たとえば透視変換 (1) では，画像平面上のベクトルとカメラ座標系におけるベクトルをそれぞれ 1 次元拡張し，

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4)$$

を考える．これらの拡張ベクトル (augmented vector) を用いると透視変換は

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (5)$$

と書ける．ただし左辺の s はスカラーで，左辺の第 3 要素が 1 となるように定めるものとする．

また，(2) の剛体変換 (rigid transformation) は v, V を拡張して

$$V = Dv, \quad D = \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \quad (6)$$

と書ける．ここで $\mathbf{0}$ は 1×3 のゼロベクトルである．(6) 式の D または D の第 4 行を取り除いた 3×4 行列

$$D' = [R \quad T] \quad (7)$$

はカメラの外部パラメータ行列 (external parameters matrix) と呼ばれる．

4.4 画像座標

コンピュータに取り込まれるときの座標と上記の画像平面上の座標はことなる．カメラ内部の撮像素子 (Charge Coupled Device: CCD) の座標 (画像座標 image coordinates) への変換が必要となる．この変換の図を Fig.6 に示す．重要な点は以下の通りである [23]．

第一に，光軸 (カメラ座標の z 軸) は画像面の中心を通るわけではない．つまり，光学系の中心は CCD 素子の中心とは必ずしも一致しない．

第二に，画像面は厳密には $x - y$ 平面と平行ではない．いいかえると，光軸は CCD 面と正確に直交しているわけではない．

第三に，画像座標 (u, v) でのスケールとカメラ座標 (x, y, z) のスケールがことなる．1 画素のサイズは，走

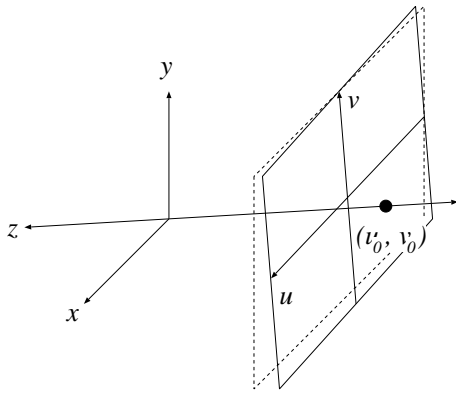


Fig. 6: Image Coordinates

査線の方向 (横方向) では A/D 変換器のクロック周期で決まり, 走査線間の方向 (縦方向) は CCD の画素サイズで決まる.

これらを考慮すると, カメラ座標と画像座標系の間ではつぎのような変換が必要である.

- 画像面上で光軸の横切る点を (u_0, v_0) とする.
- $u-v$ 平面と $x-y$ 平面のずれを $u-v$ 平面に射影して, u 軸と v 軸のなす角度 ϕ の直角からのずれに変換する.
- 縦, 横のスケール (1 画素の実質的な大きさの逆数) を k_u, k_v とする.

その結果画像面上の座標は

$$\begin{aligned} u &= fk_u \frac{x + y \cot \phi}{z} + u_0 \\ v &= fk_v \frac{y}{z \sin \phi} + v_0 \end{aligned} \quad (8)$$

となる. f は焦点距離である. $\phi = \pi/2, k_u = k_v = 1, u_0 = v_0 = 0$ であれば, 理想的なピンホールモデル (1) に一致する. さらにレンズの収差 (lens aberration) ひずみを考えることでより精度の高い補正が可能である [119] が, 一般には, よいレンズをよい条件で使う限り収差ひずみは無視できる.

4.5 カメラキャリブレーション

実際に用いるカメラについて, (6) や (8) の各パラメータを決定することをカメラキャリブレーション (camera calibration) という. キャリブレーションは, 形状が正確にわかっている物体を 3 次元ワールド座標の正確な位置におき, その像を取り込んでワールド座標系とカメラ座標系・画像座標系との対応関係を調べる. 式 (6) の回転行列と並進ベクトルを外部パラメータ (external parameters),

(8) の焦点距離, 画像中心などのパラメータを内部パラメータ (internal parameters) と呼ぶ.

ワールド座標系においてベクトル T で表される位置にカメラ座標系の原点があり, カメラ座標系の x, y, z 軸方向の単位ベクトルはワールド座標系においてそれぞれ r_x, r_y, r_z であるとする. カメラの姿勢を $R = [r_x, r_y, r_z]$ で表す. ワールド座標系で $V = (X, Y, Z)$ にある点がカメラ座標系で $v = (x, y, z)$ と表されるとすると, (6) 式より

$$v = R^T V - R^T T \quad (9)$$

を得る. v, V を拡張し, この式と (8) 式をあわせると, 画像座標とワールド座標を結び付ける同次変換は

$$w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (10)$$

となる. ここで, P は透視投影行列 (projection matrix) と呼ばれる行列で,

$$P = A [R^T \quad | \quad -R^T T] \quad (11)$$

で定義される. ただし

$$A = \begin{bmatrix} fk_u & fk_u \cot \phi & u_0 \\ 0 & fk_v / \sin \phi & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (12)$$

である. この $f, k_u, k_v, \phi, u_0, v_0$ を含む 3×3 行列 A を内部パラメータ行列 (internal parameters matrix) と呼ぶ.

P は 3×4 行列であり, その成分を

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \quad (13)$$

とおく. このとき, (10) の w を消去して

$$\begin{aligned} u &= \frac{P_{11}X + P_{12}Y + P_{13}Z + P_{14}}{P_{31}X + P_{32}Y + P_{33}Z + P_{34}} \\ v &= \frac{P_{21}X + P_{22}Y + P_{23}Z + P_{24}}{P_{31}X + P_{32}Y + P_{33}Z + P_{34}} \end{aligned} \quad (14)$$

を得る. ここで P_{11} から P_{34} までを定数倍しても (14) の関係は変わらないので, たとえば $P_{34} = 1$ と固定すると, 未知数は 11 個になる. したがって, (u, v) の組が 6 個あれば, P の成分を決定できる. また, 多数の点があれば最小 2 乗あてはめなどにより精度をあげることが期待できる (ただし, すべての点が同一平面上にあってはいけない).

4.6 ステレオ

左右に少し離れた位置に2台のカメラをおくと、それらのカメラの像は少し異なる。その違いはカメラの相対的な位置と対象表面までの距離によって決まる。したがって左右のカメラの相対的な位置がわかり、左右のカメラで見ている対象の対応がつくならば、画像の違いを用いて距離が逆算できる。この原理は三角法と呼ばれる。

この原理からわかるように、ステレオでは二つの問題がある。相対位置と対応づけである。相対位置の計測は、一種のキャリブレーション問題である。適当にセットしたカメラでステレオを構成したり、アクティブに動くカメラから得られる画像の系列から3次元形状を復元したりするときに必要になる。対応づけの問題は画像処理の一般の問題である。しかしステレオでは、対応点決定においてステレオ固有の幾何学的な拘束を利用できる [108, 75, 76]。

4.6.1 ステレオの原理

ふたつのカメラをカメラ1とカメラ2とする。空間の注目する点の世界座標での位置を (X, Y, Z) とする。この点のカメラ1, 2における画像座標を $(u_1, v_1), (u_2, v_2)$ とおく。また、それぞれのカメラの透視投影行列を P_1, P_2 とする。このとき、

$$w_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = P_i \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (15)$$

が成り立つ。したがって、カメラの透視投影行列が既知であれば、それらの画像上での対応する点の位置座標の組 $(u_1, v_1), (u_2, v_2)$ から (15) を w_1, w_2, X, Y, Z に関する方程式として解くことで注目する点の3次元位置を求めることができる。

4.6.2 基礎方程式と基礎行列 [23]

カメラ1, カメラ2の内部パラメータ行列をそれぞれ A_1, A_2 とする。Fig.7に示すように、カメラ1のカメラ座標の世界座標に対する相対的な位置・姿勢を R_1, T_1 で、カメラ2の世界座標に対する相対的な位置・姿勢を R_2, T_2 で表す。そして、カメラ1, 2の相対的な位置関係をカメラ1のカメラ座標を基準としてカメラ2のカメラ座標への回転 R と並進 t とで表す。すなわち、

$$R = R_1^T R_2, \quad t = R_1^T (T_2 - T_1) \quad (16)$$

である。

空間の位置を $M = (X, Y, Z, 1)^T$ 、その像のそれぞれのカメラの画像面上での位置を $m_1 = (u_1, v_1, 1)^T, m_2 =$

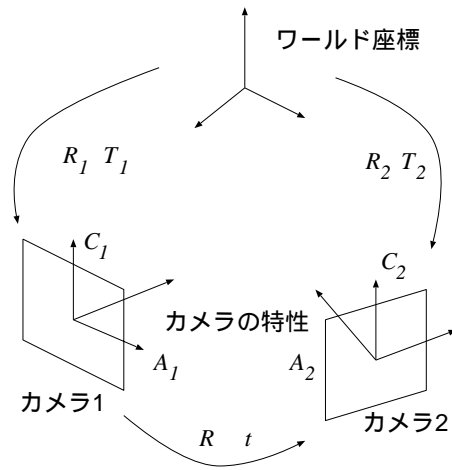


Fig. 7: Camera Coordinates in Stereo

$(u_2, v_2, 1)^T$ において、書き直すと

$$\begin{aligned} w_1 m_1 &= A_1 [R_1^T \quad | \quad -R_1^T T_1] M \\ w_2 m_2 &= A_2 [R_2^T \quad | \quad -R_2^T T_2] M \end{aligned} \quad (17)$$

となる。したがって

$$\begin{aligned} w_1 (A_1 R_1^T)^{-1} m_1 &= M - T_1 \\ w_2 (A_2 R_2^T)^{-1} m_2 &= M - T_2 \end{aligned} \quad (18)$$

を得る。この両式から M を消去すると

$$w_1 R_1 A_1^{-1} m_1 - w_2 R_2 A_2^{-1} m_2 = T_2 - T_1 \quad (19)$$

ゆえに

$$w_1 A_1^{-1} m_1 - w_2 R A_2^{-1} m_2 = t \quad (20)$$

を得る。この式より、3本のベクトル $A_1^{-1} m_1, R A_2^{-1} m_2, t$ は同一平面上にあることがわかる。このことは、物理的には、第一のカメラと注目点を結ぶベクトル、第二のカメラと注目点を結ぶベクトル、第一のカメラと第二のカメラを結ぶベクトルがひとつの平面上にあることを表している。

ここでベクトルの外積を \wedge で表すと、 $A_1^{-1} m_1$ と $t \wedge (R A_2^{-1} m_2)$ は直交する。ベクトル t に対して $t \wedge x = T x$ となる歪対象行列 T を導入すると、

$$m_1^T (A_1^{-1})^T T R A_2^{-1} m_2 = 0 \quad (21)$$

を得る。

つまり、

$$F = (A_1^{-1})^T T R A_2^{-1} \quad (22)$$

とおくと、この行列 F は2台のカメラを配置したとたんにそれらの結像特性と相対的な位置関係で決定され、そ

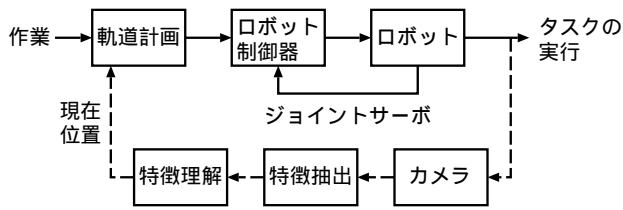


Fig. 8: Conventional Visual Feedback

して、その2台のカメラによるステレオでは、ふたつの画像上のすべての対応点 m_1, m_2 に対して

$$m_1^T F m_2 = 0 \quad (23)$$

という関係があることを示す。この行列 F を基礎行列 (fundamental matrix) といい、この方程式を基礎方程式 (fundamental equation) という。

ひとつのカメラが移動しながら画像をとる場合や、カメラは固定して対象が移動する場合は、まったく同じ内部パラメータのカメラ2台によるステレオとみなすことができ、そのときは基本行列 (essential matrix)

$$E = TR \quad (24)$$

が本質的な役割を果たす。また、上述したように、この方程式の幾何学的意味はエピ曲面、エピ曲線として解釈できる [23, 108, 97, 74]。この概念はステレオにとってきわめて重要であるが、本解説では省略する。

5 静的な視覚フィードバック制御

視覚フィードバック制御のもっとも初期の文献としては [110] があげられる。積木の世界とはいえ完全な自動化を目指した研究であったが、背景からの物体の切り出しと位置認識のために入力画像を線画に変換し、解釈するという人工知能的な手法を用いていたため、視覚フィードバックに10秒を要した。この方法 (Fig.8) では、画像入力、特徴抽出 (線画の作成)、特徴理解、軌道計画のブロックはアームが動作を開始する前に1度しか実行されない。

この研究の背景を考えると、電総研では白井が線画抽出のための非階層的な手法を提案し [109]、高瀬らがトルク制御のマニピュレータを開発している [114]。そして同時期、日立では江尻らがプリント版の欠陥検出や半導体組み立てのための自動機械を発表している [25]。また、MITではシーンアナリシスにおける線画の解釈が確立されつつあり [125]、Stanfordにおいてロボットアームの軌道計画と制御法の基礎が確立されつつあった [100]。このような流れのもとで、視覚フィードバック制御はマシン

ビジョン・マシンインテリジェンスの研究とマニピュレーションの研究に分化し、前者の応用の一分野に後者を取り込む形で続けられた。

その後、SRIにおいてボルトの挿入やコンベア上の物体のつまみあげなどへの視覚フィードバックの応用が研究され [106]、Hillらは2値化を用いた物体位置検出の簡単なアルゴリズムを提案した [58]。さらに動く部品の追跡と把握の実験を行ない、閉ループ視覚フィードバック制御における動的な問題を指摘した。

1980年代に入り、Weissは画像ベース法のビジュアルサーボ法を提案し [117]、適応制御法を用いて3自由度マニピュレータに対するシミュレーションを行なった。この方法は画像の解釈をしないので実時間制御に向いている。この研究は後に論文となって発表されるが [118]、最近の視覚サーボ法の原点ともいえる研究である。画像ベース法は特徴ベース法とも呼ばれる。

一方、有名な Andersson のピンポンロボット [5] はビジュアルサーボ法を用いずに、ボールの軌道のモデルを用いてボールを打つためのハンドの軌道を動的にプランニングするものであるが、人工知能的な視覚フィードバック手法のひとつのハイライトであろう。

6 動的な視覚フィードバック制御

Weissら [118] 以降の研究はビジョンセンサをフィードバックループの中に組み込んだ動的なビジュアルサーボ法に重点が移行する。ビジュアルサーボ系は目標値入力の方法の違いにより位置ベース法と特徴ベース法の2つに大別される。

6.1 位置ベース法

位置ベース法ではサーボ系への目標値入力は物体とハンド (またはカメラ) の相対位置・姿勢により与えられる (Fig.9)。カメラからの画像にもとづいて物体の3次元位置・姿勢を計算しロボットを制御するので、従来の静的な方法の自然な拡張になっており、ロボット制御、画像処理、画像認識などが並列に実行されることになる。位置ベース法においてはロボット制御系は通常的位置制御系であり、これまでに開発されている多様な位置制御の方法を使うことができる。そこで位置ベースのサーボ系の研究においては主にノイズに対してロバストかつ高速な位置・姿勢の検出法が研究の対象となっている。

3次元情報の獲得のために人間はいろいろな方法を複雑に組み合わせている。たとえば物体の模様、明るさ、大きさ、透明感、そして両眼によるステレオなどがあげられる。コンピュータビジョンにおいても同様の手法が開

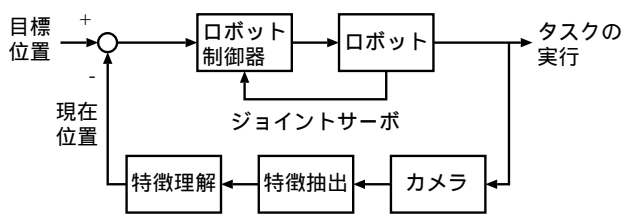


Fig. 9: Position-Based Visual Servo

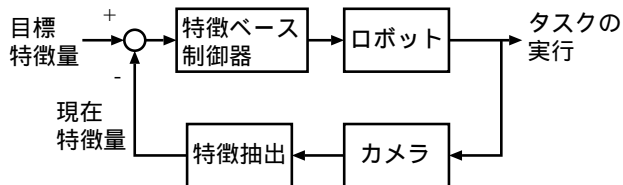


Fig. 10: Feature-Based Visual Servo by Weiss

発されているが [72], そのうちビジュアルサーボ法で主に用いられている手法は写真測量法とステレオ法である.

写真測量法は既知の大きさ・形状の物体の位置・姿勢を一枚の画像から求める方法で, 画像上の特徴点の位置を用いる. 特徴点とは物体の像の中の特徴的な点のことで, 物体の角や穴の位置, または面の中心などが用いられる. 特徴点の数が3または4の時には幾何学的な解が求まっているが, 画像ノイズの影響を受けやすく, 精度が悪い [30]. したがってコンピュータビジョンにおいては, 多くの特徴点を用いて最小2乗法によりノイズの影響を避ける. しかしビジュアルサーボ法においては実時間の要求があるため, 少ない特徴点に対してその時系列を用いて Kalman filter [123] や AR モデル [81] などを併用することにより精度の向上および予測が行なわれる.

一方, ステレオ法は2台のカメラ画像の対応を用いて三角測量により物体の位置・姿勢を求める方法である. 一般的には2枚の画像の対応づけが困難であるが, 暗い背景の中で白いボールを追跡したり [104], オプティカルフローを用いて静止した背景を消す [2] などの手法により, お手玉 [104], おもちゃの列車の追跡・把握 [2] などが実現されている.

6.2 特徴ベース法

特徴ベース法は画像上の特徴量をフィードバックして制御する (Fig.10). ここで特徴量とは画像中の領域の面積や中心位置, 線分の長さや傾きなどの, 物体の像の特徴を表す量であり, 画像の解釈を行わずに簡単な前処理で得られる量のことをいう (Fig.11). サーボ系への入力の特徴量で与えられ, 画像から抽出された特徴量との差

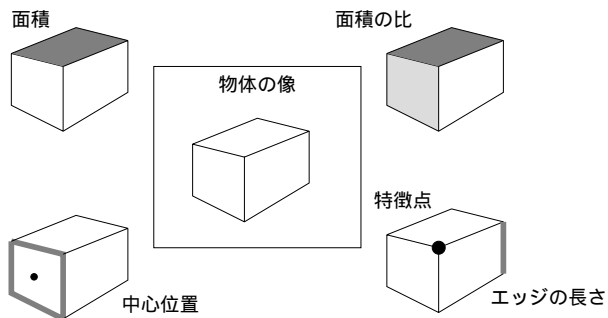


Fig. 11: Features

でロボットが駆動される. この方法の特長を以下に示す.

1. 画像の解釈を必要としないので高速
2. 物体の位置やハンドの位置の計算が不要なので, 物体のモデルやロボットの機構学的モデルの不正確さに影響されない
3. 目標値を teach-by-showing (実際に目標位置にカメラを持っていき, 画像を記憶して目標画像として用いる) により生成すれば環境のモデルを必要としない

特徴量からロボットの軌道を生成するために Jacobian of ideal inverse interpretation というヤコビ行列が用いられる. これは feature Jacobian [29], image Jacobian [50], feature sensitivity matrix [71], interaction matrix [9] などいろいろな呼ばれ方をするが, ここでは画像ヤコビアンと呼ぶことにする. 画像ヤコビアンはハンドの位置・姿勢の微小変化に対する特徴量の変化の割合として定義される. つまり特徴量ベクトルを ξ , ハンドの位置・姿勢を p とすると, 物体が動かない場合には ξ は p の関数になる.

$$\xi = \psi(p) \quad (25)$$

このとき画像ヤコビアンは

$$J_i = \frac{\partial \xi}{\partial p} \quad (26)$$

で定義される. J_i が正方で (これはハンドの持つ自由度と特徴量の数が同じことを意味する) 逆行列が存在するならば, 特徴量の偏差が減少する方向にハンドを動かせばよいわけだから,

$$\dot{p} = J_i^{-1}(\xi_d - \xi) \quad (27)$$

により特徴量を目標値 ξ_d に近づけることが可能になるであろう. ジョイントの制御にはたとえば古典的な分解速度制御 (resolved motion rate control) [121, 31, 90, 100, 113]

$$\dot{\theta} = \lambda J^{-1}(\xi_d - \xi), \quad J = J_i J_a \quad (28)$$

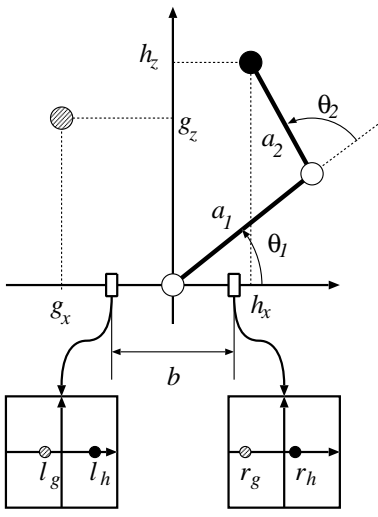


Fig. 12: 2 Link Arm with Stereo Setup

を用いて速度制御すればよい．ここで θ はロボットのジョイント角， J_a はロボット制御でよく用いられるヤコビ行列，

$$J_a = \frac{\partial p}{\partial \theta} \quad (29)$$

λ はスカラーのフィードバックゲインである．また，トルク制御の場合にはロボットのダイナミクスが

$$M\ddot{\theta} + h = \tau \quad (30)$$

で与えられるとすると (M は慣性行列， h は遠心力，コリオリ力，重力，摩擦力などの非線形な力， τ はジョイントトルク)，ビジュアルサーボ制御器は

$$\tau = MJ^{-1}(\lambda(\xi_d - \xi) - \dot{j}\dot{\theta}) + h \quad (31)$$

とすればよい．もちろん実装に関してはいろいろな問題はあるが，重要なことは“ J_i を正則とするように特徴量を選択すれば，特徴量を外界センサの出力とみなして従来のハンドレベルの制御法が適用できる”ということである．

7 位置ベース法と特徴ベース法の比較

位置ベース法と特徴ベース法の違いを明らかにするために，簡単な例題を考える．Fig.12 に 2 リンクロボットとステレオカメラの例を示す．ステレオカメラで得られた画像をもとにハンド位置をゴールに位置決めする．ロボットは水平面内を動き，カメラはその平面内に設置されている．カメラの光軸は z 軸方向，画像面は z 軸に垂直で，ベースライン長を b とする．このような設定をベースラインステレオ (base-line stereo) という．

左のカメラの出力においてハンドの像を l_h ，ゴールの像を l_g ，右のカメラにおけるハンド像を r_h ，ゴール像を r_g と定義する．2 リンクロボットの関節角を $\theta = (\theta_1, \theta_2)$ ，リンク長さを $a = (a_1, a_2)$ ，ハンド位置を $h = (h_x, h_z)$ ，ゴール位置を $g = (g_x, g_z)$ とする．(前節ではハンド位置を p と表記していたが，この節では混乱を避けるために h と表記する．)

7.1 位置ベース法

位置ベース法では，ステレオ画像からゴール位置を計算し，ロボットを制御する．ハンド位置は関節角から計算する．

7.1.1 撮像モデル

撮像モデルを $C_g : g \rightarrow \xi_g$ とする．ここで ξ_g は左右のカメラにおけるゴール画像 $\xi_g = (l_g, r_g)$ である．簡単な計算により，

$$l_g = f \frac{g_x + b/2}{g_z}, \quad r_g = f \frac{g_x - b/2}{g_z} \quad (32)$$

を得る． f は焦点距離を表す．

7.1.2 ステレオモデル

左右の画像からゴール位置を推定するには，(32) を逆算すれば良い．したがって，

$$g_x = f \frac{b(l_g + r_g)}{d_g}, \quad g_z = f \frac{b}{d_g} \quad (33)$$

を得る．ここで d は視差 (disparity) と呼ばれ， $d_g = l_g - r_g$ で定義される．人間の 3 次元位置再構成にも視差に相当する信号が重要な役割を果たすことが知られている．この過程をステレオモデルを $E_g : \xi_g \rightarrow \hat{g}$ と呼ぶことにする． \hat{g} はゴール位置 g の推定値を表し，カメラの内部パラメータ (この例では f のみ) や外部パラメータ (この例では b) の推定誤差にとまらぬ，ゴール位置の推定値にも誤差を伴うことに注意する．

7.1.3 機構学モデル

ハンド位置は関節角度から計算する．この計算を機構学モデル $K : \theta \rightarrow \hat{h}$ と呼ぶ．機構学モデルは

$$h_x = l_1 \cos \theta_1 + \cos(\theta_1 + \theta_2), \quad h_z = l_1 \sin \theta_1 + \sin(\theta_1 + \theta_2) \quad (34)$$

で表される．機構学パラメータ (この例では l_1, l_2) の誤差により手先位置の推定値に誤差が生じる．

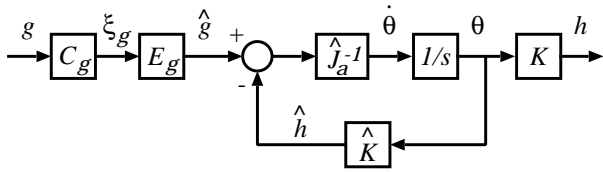


Fig. 13: Position-Base Stereo

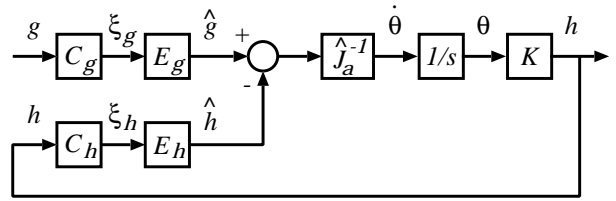


Fig. 14: Position-Base Stereo II

7.1.4 制御則

この例では，ロボットは速度制御型を仮定し，制御則はもっとも単純な分解速度制御を用いる．関節速度と手先の速度の関係は

$$J_a = \frac{\partial h}{\partial \theta} \quad (35)$$

$$= \begin{bmatrix} -l_1 \sin \theta_1 - l_2 \sin(\theta_1 + \theta_2) & -l_2 \sin(\theta_1 + \theta_2) \\ l_1 \cos \theta_1 + l_2 \cos(\theta_1 + \theta_2) & l_2 \cos(\theta_1 + \theta_2) \end{bmatrix}$$

で与えられる．したがって，制御則は

$$\dot{\theta} = \lambda \hat{J}_a^{-1} (\hat{g} - \hat{h}) \quad (36)$$

となる． J_a の要素には機構学パラメータが含まれるので，その誤差により J_a の推定にも誤差が含まれる．制御系のブロック図を Fig.13 に示す．

7.1.5 ロバスト性

パラメータ誤差に関するロバスト性について明確にしておく．

まず， E および \hat{K} に誤差がある場合，フィードバックによる修正が機能せず， h は g に収束しない．これはブロック図 (Fig.13) をみれば明らかである．

次に， E および \hat{K} が正確であるとして h は g に収束するための \hat{J}_a に関する条件を求める．リャプノフ関数の候補として

$$V = (h - g)^T (h - g) \quad (37)$$

をとる．このとき，

$$\dot{V} = -2\lambda (h - g)^T J_a \hat{J}_a^{-1} (h - g) \quad (38)$$

であるから， $\dot{V} \leq 0$ となるための条件は $J_a \hat{J}_a^{-1} > 0$ となる．この条件は「 E および \hat{K} が正確」という条件に比べるとはるかにゆるい．

7.2 位置ベース法 II

ハンドがカメラで計測できる場合，ハンド位置を関節角から計算せず，ステレオ法により算出することもできる．ハンドの撮像モデルを C_h ，ハンドの像 ξ_h から \hat{h} を

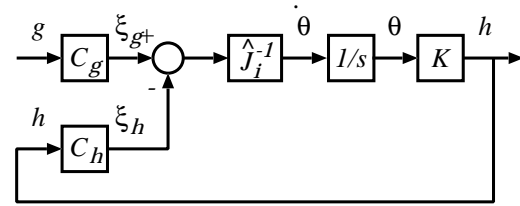


Fig. 15: Feature-Base Stereo

計算するステレオモデルを E_h とすると，この制御法のブロック図は Fig.14 となる．

この場合， \hat{K} は用いない．また， E_g と E_h において同じカメラパラメータ (内部・外部) を使うことにすれば，パラメータ誤差に対してロバストになる．たとえば， b と f が b' と f' になったとすると，この誤差が \hat{g}, \hat{h} に与える影響は，すべての要素に均一に $b'f'/bf$ 倍であり，パラメータ誤差に関係なく h を g に収束させることが可能である．また，ヤコビ行列のロバスト性に関しては前節の位置ベース法のとおり同様である．

7.3 特徴ベース法

特徴ベース法では，画像面内におけるハンドの像を直接制御する．関節速度と画面内のハンドの像 $\xi_h = (l_h, r_h)$ の時間微分の関係は

$$J_i = \frac{\partial \xi_h}{\partial \theta} = \frac{\partial \xi_h}{\partial h} \frac{\partial h}{\partial \theta} \quad (39)$$

で与えられる．ただし

$$\frac{\partial \xi_h}{\partial h} = \begin{bmatrix} f \frac{b}{2h_x} & -f \frac{h_x + b/2}{h_z^2} \\ -f \frac{b}{2h_x} & -f \frac{h_x - b/2}{h_z^2} \end{bmatrix} = \begin{bmatrix} d_h/2 & -l_h d_h / fb \\ -d_h/2 & -r_h d_h / fb \end{bmatrix} \quad (40)$$

で， $d_h = l_h - r_h$ はハンドに関する視差である．また， $\frac{\partial h}{\partial \theta}$ は (36) で定義されている．分解速度制御を採用すると，制御則は

$$\dot{\theta} = \lambda \hat{J}_i^{-1} (\xi_g - \xi_h) \quad (41)$$

となる．制御系のブロック図を Fig.13 に示す．

この制御法は機構学モデルの計算やステレオモデルの計算を含まないので，機構パラメータやカメラパラメー

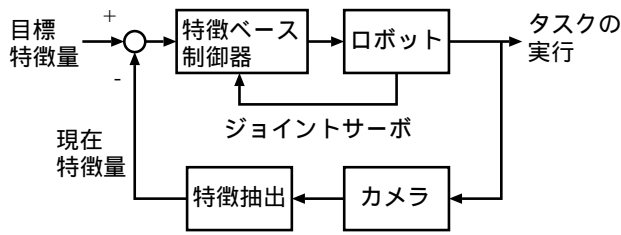


Fig. 16: Feature-Based Visual Servo

タの誤差は直接的には影響しない。しかし $\frac{\partial \xi_h}{\partial h}$ にはカメラパラメータが、 $\frac{\partial h}{\partial \theta}$ の要素には機構学パラメータが含まれるので、その誤差により J_i の推定に誤差が含まれる。ヤコビ行列の誤差に対するロバスト性は、前の節でみたように、 $J_i \hat{J}_i^{-1} > 0$ がみたされれば収束性に影響を与えない。つまり J_i に関する正定性の条件をみたせば h が g に収束することが保証される。

8 特徴ベースのビジュアルサーボ法

(28) の J はロボットの姿勢 p に依存して変動する。Weiss らは λJ^{-1} を時変なゲインとしてとらえ、各入出力間を独立とみなして 1 入力 1 出力の model-reference adaptive control を用いて J を推定しながら制御することを試みた [118]。さらに各入出力間を独立とみなすことができるための条件を導出し、シミュレーションにより有効性を検討した。しかし“画像のサンプリング周期に比べてロボットの制御に必要なサンプリング周期は数十倍短い”という極めて重要な問題が考慮されていないことが (Fig.10 でジョイントサーボが施されていないことに注意)。

Feddema らはこのサンプリング周期の問題を解決するため、画像のサンプリング周期内において連続的な特徴量軌道の生成法を提案した [29]。つまり、

$$\dot{\theta} = \lambda J_a^{-1} J_i^{-1} f(\xi_d - \xi(k), t) \quad (42)$$

にもとづいてジョイントを制御することを提案した (Fig.16)。 f はなめらかな特徴量軌道で k は時間を表すパラメータである。Feddema らはその後特徴量位置を推定しながら制御する適応制御法 [27] や特徴量の自動選択法などの研究 [28] を行なった。

Corke らは Feddema らの方法において画像フィードバックのタイミング、画像入力の遅れなどについて詳細に検討し、動特性のモデル化を行なった [15]。また、複数の特徴点に対する解析と実装における注意点を検討した [12]。その後、Kalman filter を用いて物体の運動の推定を行ないながら追従することを提案した [16]。

Papanikolopoulos らは SSD (sum of square difference) optical flow を導入し、ノイズに対してロバストな特徴量抽出の研究を行なうとともに、簡単な 3 自由度のモデルにもとづいて PI (proportional and integral), 極配置, LQG (linear quadratic Gaussian) などの制御法の比較実験を行なった [99]。その後, [98] では self-tuning regulator により深さを推定しながら制御する方法を提案した。

Nelson らはビジュアルサーボ問題をセンサ配置問題ととらえ [95], 特異点回避や画像ヤコビアン of オンライン推定法などを研究した [96]。

Espiau らは task function を用いてビジュアルサーボ系を構成した [26]。タスクを $\|e(p(k))\|$ の最小化問題として定義し、ビジュアルサーボ系においては e を

$$e(p(k)) = J_d^T (\xi(p(k)) - \xi_d) \quad (43)$$

と定義した。ここで J_d は $\xi = \xi_d$ のときの J_i である。 $\|e\|$ を減少させるため $\dot{e} = -\lambda e$ とおき

$$\dot{p}(k) = \left(\frac{\partial e}{\partial p} \right)^{-1} \dot{e} \approx -\lambda J_d^\dagger (\xi(p(k)) - \xi_d) \quad (44)$$

を用いた。ここで λ は正のスカラであり、 J_d^\dagger は J_d の一般化逆行列である。

Chaumette らは離散時間化した task function と Kalman filter を用いて物体の運動を推定しながら追跡を行なった [11]。また、点以外の特徴量に対して具体的に画像ヤコビアンを求め、シミュレーションによりその有効性を示した [10]。また、ビジュアルサーボの安定領域の局所性を指摘し [8], 大域的に安定化可能な方法を提案した [83]。

著者らは一般化逆行列を用いて関節角変位 $\Delta \theta$ を

$$\Delta \theta = -J_a^{-1} J_i^\dagger (\xi - \xi_d) \quad (45)$$

にもとづいて計算し、それを内挿して制御する方法を提案した [50]。また、位置ベース法と特徴ベース法の比較シミュレーションにより特徴ベース法のロバスト性を示し、PUMA560 を用いた実時間実験により特徴ベース法の有効性を確認した。その後、冗長な特徴量に対する LQ (linear quadratic) 最適制御 [49, 40] や非線形制御 (task level inverse dynamics) を提案した [48]。非線形法の実験結果は [52] に、LQ 法との比較は [53] に述べられている。また、物体の運動のモデルにもとづいて運動パラメータを推定する非線形適応オブザーバを提案し、非線形制御法と併用することにより追従偏差が漸的に零になることを証明した [51]。さらに、特徴量を冗長とすることで制御系の感度および一種のゲインが上昇することを示した [56, 44]。また、安定領域の問題を検討し、ポテンシャル切り替え法を提案した [57, 45]。

Jang らは特徴量の数学的な定義を与えた [71] . この定義は計算量の面で実用的ではないが, 興味深い研究である . さらに彼らは (10) と同様の制御則を用いて実時間制御を行なったが, 画像処理に時間がかかり遅れが大きい . その後, Bien らはこの特徴量にもとづいて fuzzy self-organizing controller を提案した [7] .

Ghosh らは物体の運動のモデルを与え, 物体が透視変換を経て観測される場合の運動パラメータの推定問題を研究した [36] . さらに, 固定したステレオカメラによりひとつの特徴点とロボットハンドを観測し, 非線形出力フィードバックレギュレータを用いてハンドを物体に追従させ, その偏差が漸近的に零になることを示した [82] .

Jägersand らはモデル推定法をビジュアルサーボに適用し, 電球の把握や組み立て作業などを行った [69, 70] .

出口らはビジュアルサーボにおいて並進と回転を非干渉化し, ロボットの軌道の最適性について議論した [21] . また, 能動カメラによりビジュアルサーボを行いながら同時に 3 次元形状を計測する方法について研究した [22] .

細田らは画像ヤコビアンを適応的に推定しながらビジュアルサーボを行う方法について研究した [64, 61] . また, ビジュアルサーボに基づく位置制御と力制御 [62] や脚式ロボットの制御 [63] などを行った .

石川らは 1ms で画像入力・処理が可能な高速視覚システムを開発し [92, 91], それを用いて高速に運動する物体の把握 [93] やセンサフュージョン [94] の研究を行った .

藤田らは制御理論的観点からビジュアルサーボを研究し, 受動性に基づくロバストな制御法 [86] や自動車の自動操縦の研究 [32] を行った .

以下の節では, 冗長な特徴量を用いたビジュアルサーボの特性と大域的安定化, オブザーバに基づくビジュアルサーボ, 超高速ビジョンを用いた把握について詳しく述べる .

9 冗長な特徴量に基づくビジュアルサーボ

9.1 可制御性

特徴ベースのビジュアルサーボ法の研究において, これまで可制御性と安定性はあまり議論されていない . 特徴ベースでは特徴量を状態変数にとるのが自然であるから, 可制御性問題とは “任意に与えられた ξ_d に対して $\xi - \xi_d = 0$ とするような入力が存在するかどうか?” という問題である . Weiss, Feddema, Papanikolopoulos らは, 制御入力の決定において特徴量を J_i が正則になるように選択しており, 暗黙のうちに可制御になっている . しかし Espiou, Hashimoto, Jang らの一般化逆行列を用いた方法では入力の数よりも多くの特徴量を選択することができ, その場合には可制御ではない . ここでは冗長な

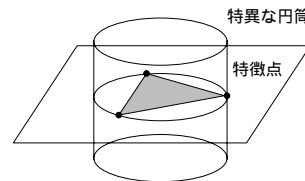


Fig. 17: Cylinder of Singularity

特徴量を用いる必要性, 可制御性, 安定性について考えたい .

たとえば 6 自由度のマニピュレータのハンドにカメラを固定して, 3 つの特徴点 (画面上の 3 点の $x-y$ 座標を特徴量と考えるとその数は 6 になる) に対して位置決めを行なう場合, 3 点を含む円筒で 3 点を含む平面に垂直なもの (Fig.17) の上にカメラがあるとき J_i は正則でなくなる [87] . これでは 3 角形の内部にカメラを位置決めすることはできない . 特徴点を 4 点にすれば J_i はカメラが 4 点を含む平面内にあるときを除き full rank になり一般化逆行列は存在するが, 特徴量の数は 8 になり冗長になる .

状態変数を特徴量偏差 $\bar{\xi} = \xi - \xi_d$ にとり, 入力をハンドの速度 $u = \dot{p}$ と考えれば状態方程式は

$$\dot{\bar{\xi}} = J_i u \quad (46)$$

となる . このときでたために ξ_d を与えては可制御でなくなる . なぜなら特徴点は剛体上にあり自由に動くことが許されないからである . そこで, 自由に動くことのできる変数 p を使い, ハンドの目標値を p_d として目標特徴量を $\xi_d = \psi(p_d)$ と定める (teach by showing) . すると, このようにして定められた任意の ξ_d に対して ($p = p_d$ とする u は必ずあるから) $\xi = \xi_d$ とする u は存在する (可制御である) .

次に安定性を考える . p_d が既知で p が計測可能なら $p \rightarrow p_d$ は容易である . 問題は p_d ではなく ξ_d が与えられ p ではなく ξ が計測されることである . たとえば制御則

$$u = -\lambda J_i^+ \bar{\xi} \quad (47)$$

を考えると偏差系のダイナミクスは

$$J_i^T (\dot{\bar{\xi}} + \lambda \bar{\xi}) = 0 \quad (48)$$

となる . 定常状態における偏差を $\bar{\xi}_\infty$ と書くと, 中間値の定理より

$$\bar{\xi}_\infty = M(p - p_d) \quad (49)$$

となる行列 M が存在し, $p \rightarrow p_d$ のとき $M \rightarrow J_d$ であるから, J_d が full rank であれば $p = p_d$ の近傍において $p \rightarrow p_d$, すなわち $\bar{\xi} \rightarrow 0$ が示される (安定である) .

Espiou らは (47) ではなく (44) にもとづいてほぼ同じ結果を導いている。また、特徴量が不足して制御に自由度が余る場合についても考察している [26]。

冗長な特徴量の場合には

1. J_d の full rank 性
2. $J_i^T M$ が正則となる p の範囲
3. 2. の範囲内に p を収束させることができるか?

などが問題になる。4 点の場合には J_i はほとんどの範囲で full rank であり p の初期値が p_d に近い場合には 2., 3. は問題ないと思われる。しかし J_i が p の関数でありその構造は与えられた作業に依存して変化するという性質上、2., 3. の問題の厳密な解析はまだなされていない。

9.2 最適制御法

特徴量を冗長にとる場合の制御則は Espiou, Jang らによって提案されているが、いずれも制御性能の調節に用いるパラメータはひとつのスカラー λ のみであり、制御系設計の面から見れば不十分である。そこで著者らは LQ 最適制御法を提案した [49]。ここでは [40] に沿って解説する。具体的には (45) を離散化時間したモデル

$$\bar{\xi}(k+1) = \bar{\xi}(k) + T J_d u(k) \quad (50)$$

を考える。 T はサンプリング周期である。冗長な場合、偏差系 (50) はこのままでは不可制御である。しかし、それは ξ のすべての成分が独立に動くことを許しているためであり、目標値を適切に設定すれば安定化可能なはずである。そこで ξ は p の関数であり、 $\xi = \psi(p)$ で生成される多様体上を動くことに注意する。そして特徴量ベクトル $\bar{\xi}$ を $\bar{\xi} = 0$ における接平面に射影する。

$$z(k) = J_d^T \bar{\xi}(k) \quad (51)$$

このとき $z \rightarrow 0$ なら $\bar{\xi} \rightarrow 0$ となることは明らかであり、 $z \rightarrow 0$ を制御目的とすればよい。次に $B = J_d^T J_d$, $v = Tu$ とおき、偏差系を

$$z(k+1) = z(k) + Bv(k) \quad (52)$$

と変換する。 J_d が full rank のとき B は正則となり (52) は可制御となる。このシステムに対して適当な評価関数を定め、それを最小にする状態フィードバック $v = -Kz$ を求めることは容易である。そしてジョイントの目標変位コマンドを $\Delta\theta = v$ として

$$\Delta\theta = -K J_d^T (\xi - \xi_d) \quad (53)$$

にもとづき目標ジョイント角度を生成し、通常ジョイント角度サーボで制御すればよい。この方法では、モデ

リング段階において J_d が計算されていれば実時間での物体位置の推定や画像ヤコビアン計算は一切必要ない。さらに評価関数の重みを調節することにより性能の改善が容易に行なえる。

さらに著者らは、制御偏差を (51) のように定義することにより撮像過程のヤコビ行列を正則にできることに注目し、ロボットのダイナミクスを厳密に線形化する視覚フィードバック則を導いた [52]。

9.3 特徴量の選択と感度

9.3.1 感度の定義

特徴量を冗長にすることでビジュアルサーボ系の特性がどのように変化するかを考察する。まず、画像ヤコビアン J_i の特異値分解を $J_i = UDV^T$ とする。このとき、

$$J_i v_k = d_k u_k \quad (k = 1, 2) \quad (54)$$

が成り立つ。 u_k, v_k は U, V の第 k 列で、 $|u_k| = 1, |v_k| = 1$ をみたす。また、 d_k は D の (k, k) 成分で、 $k < j$ なら $d_k \geq d_j$ ある。

(54) は以下のように解釈できる。 u_k 方向の単位大きさのカメラの運動は d_k 倍されて画像の v_k 方向の動きとなる。したがって、最大特異値 d_1 に対応した特異ベクトル u_1 方向のカメラの運動に対しては画像特徴量は大きく変化するが、最小特異値 d_n に対応する特異ベクトル u_n 方向のカメラの運動に対しては画像特徴量は d_n に相当する量しか変化しない。もし、最小特異値が 0 であれば、カメラが動いても画像が変化しない方向が存在することになり、ビジュアルサーボにはそのような特徴量は不適である。逆に、最小特異値がある程度大きければカメラのすべての運動方向に対して画像が適度に変化するので、そのような画像特徴量はビジュアルサーボに適している。したがって、画像ヤコビアン J_i の最小特異値をビジュアルサーボ系の感度と呼び、特徴量選択の重要な指標となる [44]。また、特徴量の数と感度の関係は [47] で詳しく解析されている。

9.3.2 2 自由度の例

ここで簡単な例題を考える。Fig.18 にカメラと 2 つの注目点を示す。この 2 つの注目点を目標位置に保つよう、ビジュアルサーボを行う。カメラの自由度は紙面内上下左右 (X, Z) の 2 自由度とする。注目点はカメラから距離 D の位置にあり、間隔 $2B$ である。このとき、2 点の像は

$$\xi = (x_1, x_2) = \left(-f \frac{B}{D}, f \frac{B}{D}\right) \quad (55)$$

である。2 自由度であるから、 y 方向 (画面内の縦方向) の座標は考えない。

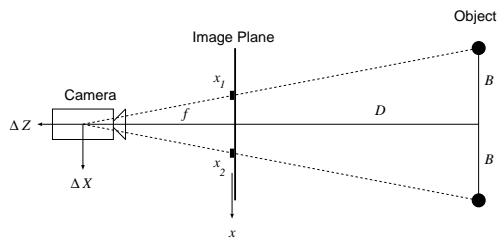


Fig. 18: 2 points

カメラの微小移動 $\Delta p = (\Delta X, \Delta Z)$ に対する特徴量 ξ の変化量は画像ヤコビ行列 J_i であるが、この例では

$$\frac{\partial \xi}{\partial p} = \begin{bmatrix} -\frac{f}{D} & \frac{fB}{D^2} \\ -\frac{f}{D} & -\frac{fB}{D^2} \end{bmatrix} \quad (56)$$

となる。数値例として $D = 1000, B = 100, f = 100$ を当てはめてみれば、特異値分解は

$$U = \begin{bmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{bmatrix}, \quad D = \begin{bmatrix} 0.141 & 0.0 \\ 0.0 & 0.014 \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (57)$$

となる。したがって、カメラの X 方向の動きに対する画像の変化 $d_1 = 0.141$ と Z 方向の動きに対する画像の変化 $d_2 = 0.014$ には 10 倍の差があることがわかる。ビジュアルサーボでは特徴量偏差に比例する運動指令値を与えることが一般的であるので、Fig.18 の構成においてはカメラの運動は X 方向に関しては剛性が高いが、 Z 方向に関しては剛性が低い。この剛性を示すためにカメラの運動方向に対応する画像特徴量の変化を楕円で表現すると Fig.19 となる。

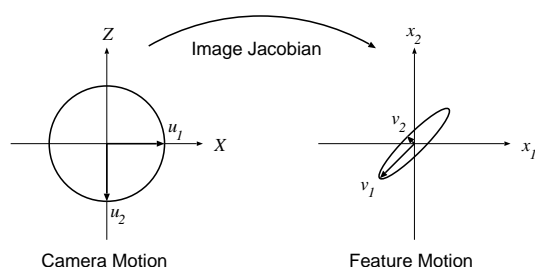


Fig. 19: Ellipses of Sensitivity

10 運動する物体の追従

10.1 これまでの研究

単なるフィードバックだけでなく、物体の運動が予測できればフィードフォワード制御により追従特性が改善

できる。とくにビジュアルサーボでは、CCD カメラのサンプリングの遅れが大きく、画像処理にも時間がかかるため、フィードフォワードは効果的である。ビジュアルサーボにおける動的な効果については Corke と Good が研究している [12, 15, 16, 17]。また、[14] において Corke はビジュアルサーボ系を時間遅れのある線形システムとしてモデル化し、詳細な解析を行った。

フィードフォワード構造をもつビジュアルサーボの研究はたくさんなされている。 $\alpha-\beta-\gamma$ フィルタに基づく運動推定 [2]、Kalman フィルタを用いた位置ベースのビジュアルサーボ [120, 123, 124]、特徴ベースで一定加速度を仮定した研究 [11]、一般化尤度比試験を突然の運動変化の検出に用いた研究 [6]、ロボットの運動特性の推定に GPC (Generalized predictive controllers) を用いた方法 [33] と GPC を形状追従に拡張した研究 [34]、AR モデルに基づく研究 [81] などがあげられる。

ここで注意が必要なのは、 $\alpha-\beta-\gamma$ フィルタや Kalman フィルタ、GPC、AR モデルなどは観測の瞬間のシステムの状態を予測するだけで、サンプリングの中間の状態を推定するには適さないということである。いいかえると、これらのフィルタでは、ビジョンのサンプリングレートで状態(たとえば、対象の位置と速度)は推定できるが、関節角のサンプリングレートでは推定が困難であることを示す。したがって、これらのフィルタを Fig.16 のような 2 つの異なるサンプリングレートをもつシステムに適用するにはかなりの変更が必要となる。

一方、もし、対象の運動のモデルが使えるならば、Luenberger 型のオブザーバがサンプリング間隔内の情報を推定するのに利用可能である。Ghosh らは対象の速度を推定するオブザーバを提案した [36, 38, 37]。Rizzi と Koditchek は対象のトラッキングのためのウィンドウ位置の予測を研究した [102, 103]。著者らは関節角サーボのサンプリングレートで対象の運動速度を推定するオブザーバを提案し、線形・非線形の制御器を提案した [54, 55, 47, 51, 46]。さらに、実時間実験により物体の運動の推定が精度良く行なわれること、そしてロボットによる追従性能がかなり改善されることを示した。

10.2 例題

Fig.20 に示す 2 リンクロボットを考える。ロボットの手先は水平面に平行な平面内を動く。カメラの光軸は Z_w 方向とする。

10.2.1 2 つのヤコビアン

ロボットの関節角を $q = (q_1, q_2)$ 、リンク 1 の長さを l_1 、リンク 2 の原点からカメラまでのベクトルを (l_{cx}, l_{cy}) と

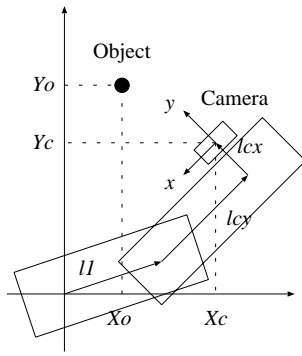


Fig. 20: Two Link Direct Drive Robot (overview)

する．このとき，カメラ位置・姿勢 s_c は

$$s_c = \begin{bmatrix} \ell_1 c_1 + \ell_{cx} c_{12} - \ell_{cy} s_{12} \\ \ell_1 s_1 + \ell_{cx} s_{12} + \ell_{cy} c_{12} \\ q_1 + q_2 \end{bmatrix} \quad (58)$$

となる．ここで $s_1 = \sin q_1$, $c_1 = \cos q_1$, $s_{12} = \sin(q_1 + q_2)$, $c_{12} = \cos(q_1 + q_2)$ とおいた．オブジェクト位置を $s_o = (X_o, Y_o)$ とする．オブジェクトまでの距離を Z ，焦点距離を f とすると，特徴量ベクトル $\xi = [x \ y]^T$ は

$$\xi = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix} = F \begin{bmatrix} \ell_{cx} + \ell_1 c_2 - X_o c_{12} - Y_o s_{12} \\ -\ell_{cy} + \ell_1 s_2 - X_o s_{12} + Y_o c_{12} \end{bmatrix}, \quad (59)$$

となる．ここで $F = \frac{f}{Z}$, $s_2 = \sin q_2$, $c_2 = \cos q_2$ とおいた．

この関係式 (59) から，ロボットが動くことによる画像の変化 $J = \frac{\partial \xi}{\partial q}$ とオブジェクトが動くことによる画像の変化 $L = \frac{\partial \xi}{\partial s_o}$ は次式のようになる．

$$J = F \begin{bmatrix} X_o s_{12} - Y_o c_{12} & -\ell_1 s_2 + X_o s_{12} - Y_o c_{12} \\ -X_o c_{12} - Y_o s_{12} & \ell_1 c_2 - X_o c_{12} - Y_o s_{12} \end{bmatrix} \\ L = F \begin{bmatrix} -c_{12} & -s_{12} \\ -s_{12} & c_{12} \end{bmatrix} \quad (60)$$

ここで，簡単な計算により

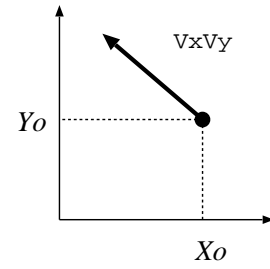
$$\det J = F^2 \ell_1 (X_o s_1 - Y_o c_1), \quad \det L = -F^2. \quad (61)$$

を得る．したがって， J は第一ジョイントと第二ジョイントを結ぶ線上にカメラが来たときに特異になることがわかる．また， L は常に正則である．

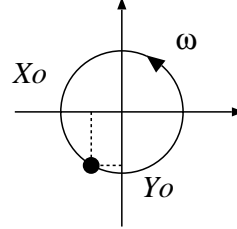
10.3 運動のモデル

オブジェクトの運動は次の自励系の出力であると仮定する．

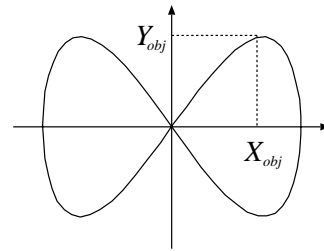
$$\dot{p} = W(p)\theta^* \quad (62)$$



(a) Linear



(b) Circular



(c) Figure 8

Fig. 21: Object Motions

ここで p はオブジェクト位置， $W(p)$ は適当な次元の行列で p の関数である．ベクトル θ^* は速度パラメータで，式 (62) をオブジェクトの運動モデルと呼ぶ．このモデルは単純であるが，かなり広い運動のクラスを表現することができる．たとえば直線，円，楕円，“8の字”などである．また，オブジェクトの運動は円運動と直線運動の組み合わせで生成できることを考えるとこのモデルは妥当である．同様のモデルは [35] で考察されており，[38] では2段階のオブザーバに拡張されている．

以下にオブジェクトの運動モデルの例をあげる．

直線運動 Fig.21(a) に示すように，オブジェクトの運動が直線で X 方向の速度が v_X ， Y 方向の速度が v_Y とすると，上記の運動モデル (62) は

$$W(p) = I, \quad \theta^* = [v_X \ v_Y]^T \quad (63)$$

となる．これはもっとも簡単な運動パターンである．

円運動 オブジェクトの運動が Fig.21(b) に示すような円運動の場合，角速度を ω とすると，オブジェクトの位

置・速度は

$$\begin{bmatrix} X_o \\ Y_o \end{bmatrix} = \begin{bmatrix} r \cos \omega t \\ r \sin \omega t \end{bmatrix}, \quad \frac{d}{dt} \begin{bmatrix} X_o \\ Y_o \end{bmatrix} = \begin{bmatrix} -r\omega \sin \omega t \\ r\omega \cos \omega t \end{bmatrix}. \quad (64)$$

で表される．ここで r は円運動の半径である．したがってオブジェクトの運動モデルは

$$W(p) = \begin{bmatrix} -Y_o \\ X_o \end{bmatrix}, \quad \theta^* = \omega. \quad (65)$$

となる．

オブジェクトの運動の中心が原点ではなく， (c_x, c_y) の場合，オブジェクトの運動は

$$\begin{bmatrix} X_o \\ Y_o \end{bmatrix} = \begin{bmatrix} r \cos \omega t + c_x \\ r \sin \omega t + c_y \end{bmatrix}. \quad (66)$$

となる．速度は (64) と同じである．中心位置が不明の場合，運動モデルを

$$W(p) = \begin{bmatrix} -Y_o & 0 & 1 \\ X_o & -1 & 0 \end{bmatrix}, \quad \theta^* = \begin{bmatrix} \omega \\ \omega c_x \\ \omega c_y \end{bmatrix} \quad (67)$$

とすることにより，中心位置も同時に推定できる．

8 の字 Fig.21(c) に示すような 8 の字運動の場合，オブジェクト位置は

$$\begin{bmatrix} X_o \\ Y_o \end{bmatrix} = \begin{bmatrix} R_1 \cos \omega t \\ R_2 \sin 2\omega t \end{bmatrix} \quad (68)$$

となる．このとき運動モデルは

$$W(p) = \begin{bmatrix} 0 & -\frac{Y_o}{X_o} \\ X_o^2 & -\frac{Y_o^2}{X_o^2} \end{bmatrix}, \quad \theta^* = \begin{bmatrix} \frac{1}{R} \omega \\ R \omega \end{bmatrix}, \quad (69)$$

で与えられる．ここで $R = \frac{R_1^2}{2R_2}$ である．

10.4 動特性を補償する制御器

ロボットの動特性を補償する制御器 (31) においては，特徴点が動くのはロボットが動くからであると考えていた．しかし，オブジェクトの運動を陽に考慮するとなれば，それも制御器に取り込むべきである．また，特徴点の速度，加速度もロボットおよび物体の運動の速度，加速度から計算でき．オブジェクトの運動が (63) で与えられることを考慮すると，特徴点の加速度は以下ようになる (計算はめんどうなので結果だけを掲載する)．

$$\ddot{\xi} = JM^{-1}(\tau - h) + \mu + N\theta^* + \Phi\kappa(\theta^*) \quad (70)$$

ここで

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix}, \quad N = \begin{bmatrix} N_1 \\ \vdots \\ N_m \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_m \end{bmatrix}, \quad (71)$$

であり，各々の要素は

$$\begin{aligned} \mu_i &= \dot{q}^T \frac{\partial^2 \xi_i}{\partial q^2} \dot{q}, \quad N_i = 2\dot{q}^T \frac{\partial^2 \xi_i}{\partial p \partial q} W, \\ \Phi_{ijk} &= \left[\frac{\partial}{\partial p} \left\{ W^T \left(\frac{\partial \xi_i}{\partial p} \right)^T \right\} W \right]_{jk}, \\ \Phi_i &= \begin{bmatrix} \Phi_{i11} & \cdots & \Phi_{i1\ell} & \Phi_{i21} & \cdots & \Phi_{i\ell\ell} \end{bmatrix}, \\ \kappa(\theta^*) &= \begin{bmatrix} \theta_1^{*2} & \cdots & \theta_1^* \theta_\ell^* & \theta_2^* \theta_1^* & \cdots & \theta_\ell^{*2} \end{bmatrix}^T \end{aligned}$$

である．ロボットの関節角が q ，オブジェクトの位置が p で， θ はオブジェクトの速度パラメータであることに注意する． J は正則なので，アクチュエータトルクを

$$\begin{aligned} \tau &= MJ^{-1}(v - \mu - N\theta^* - \Phi\kappa(\theta^*)) + h \\ v &= -K_1 \xi - K_2 \dot{\xi}, \end{aligned} \quad (72)$$

とすることで完全な逆動力学制御器が構成できる [52, 43]．

10.5 オブザーバ

上述の制御器では特徴量速度 $\dot{\xi}$ と運動パラメータ θ^* を必要とする．それらを推定するために Luenberger 型の適応オブザーバを構成する．運動パラメータの推定値を $\hat{\theta}$ ，特徴量の推定値を $\hat{\xi}$ とおくと，特徴量のダイナミクス $\dot{\xi} = J\dot{q} + L\dot{p}$ とオブジェクトの運動モデル (62) に基づいて以下のオブザーバが構成できる [88, 116]．

$$\dot{\hat{\xi}} = J\dot{q} + LW\hat{\theta} + H(\hat{\xi} - \xi), \quad \dot{\hat{\theta}} = -W^T L^T P(\hat{\xi} - \xi). \quad (73)$$

このパラメータ更新則は最小 2 乗同定のグラディエントアルゴリズム [107] に基づくものである．行列 H および P はオブザーバゲインに相当し， H は任意の安定行列で，任意の正定行列 Q に対して P は以下の式をみたすように選ばれる．

$$H^T P + PH = -Q \quad (74)$$

このとき偏差系を考えると

$$\bar{\xi} = \xi - \hat{\xi}, \quad \bar{\theta} = \theta^* - \hat{\theta}, \quad e = \begin{bmatrix} \bar{\xi} \\ \bar{\theta} \end{bmatrix}. \quad (75)$$

となり，安定性はリャプノフ関数を

$$V = e^T \tilde{P} e \quad \text{where} \quad \tilde{P} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}. \quad (76)$$

として，証明できる [42]

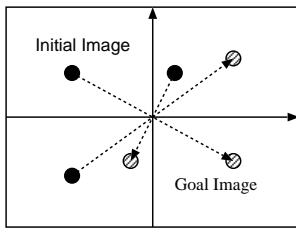


Fig. 22: Symmetric Features

10.6 オブザーバ併合制御器

オブジェクトの運動が直線だとすると制御器は

$$\begin{aligned}\ddot{\xi} &= J\ddot{\xi} + \mu + N \begin{bmatrix} v_X \\ v_Y \end{bmatrix}, \\ \mu &= \left[(\dot{q}_1 + \dot{q}_2) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} J + F\ell_1\dot{q}_1 \begin{bmatrix} 0 & c_2 \\ 0 & s_2 \end{bmatrix} \right] \dot{q}, \\ N &= 2F(\dot{q}_1 + \dot{q}_2) \begin{bmatrix} s_{12} & -c_{12} \\ -c_{12} & -s_{12} \end{bmatrix}, \\ \tau &= MJ^{-1}(-K_1\xi - K_2\dot{\xi} - \mu - N \begin{bmatrix} v_X \\ v_Y \end{bmatrix}) + h. \quad (77)\end{aligned}$$

となる．

オブジェクトの運動が円運動であれば，制御器は

$$\begin{aligned}\ddot{\xi} &= J\ddot{q} + \mu + N\omega + \Phi\omega^2, \\ \Phi &= F \begin{bmatrix} -s_{12} & c_{12} \\ c_{12} & s_{12} \end{bmatrix} \begin{bmatrix} -Y_o \\ X_o \end{bmatrix}, \\ N &= 2(\dot{q}_1 + \dot{q}_2)\Phi, \\ \tau &= MJ^{-1}(-K_1\xi - K_2\dot{\xi} - \mu - N\omega - \Phi\omega^2) + h(78)\end{aligned}$$

となる．ただし μ は (77) と同じである．

11 大域的安定化

11.1 ポテンシャルによる定式化

特徴量を冗長にとる場合，感度やゲインが増す [47]．この場合の典型的な制御則は画像ヤコビアン疑似逆行列を用いる方法 (47) である．このときの問題点は [8] に述べられているように，安定領域の局所性およびカメラ軌道の複雑さである．

たとえば，次の制御則を考える．

$$u = \lambda J^\dagger (\xi_d - \xi) \quad (79)$$

もし，Fig.22 に示すように，目標特徴量 ξ_d と現在特徴量 ξ が画像平面の原点に対して点対象である場合，特徴量

偏差ベクトル $\xi_d - \xi$ において各特徴点に対応するベクトルはすべて画像原点に向かう．すなわち，(79) により生成されるロボット制御コマンドはすべての特徴点を原点に収束させる方向にカメラを動かす．その結果，カメラはオブジェクトから遠ざかる．

また，(79) の制御則は特徴量を直線的に変化させるので，カメラの描く軌道は直感的にはわかりづらく，現場での実際の応用には不適切な場合がある．出口からはカメラの回転と並進を分解し，カメラの軌道が最適になるような制御則を提案した [21]．Corke と Hutchinson も同様な考え方によりオブジェクトの像が視野からはずれない制御則を提案した [18]．

制御則 (79) の問題点はほかにもある．特徴点を冗長にしているので，画像ヤコビアン J は縦長である．したがって $J^\dagger = (J^T J)^{-1} J^T$ であり，

$$\xi_d - \xi \in \text{Ker} J^T \quad (80)$$

のとき， $u = 0$ となる．特徴量の数を増すと $\text{Ker} J^T$ の次元も増すので，注意しなければならない．

この問題はポテンシャルを $V = (\xi_d - \xi)^T (\xi_d - \xi)$ ととったポテンシャル最小化問題と考えるとわかりやすい．制御則 (79) は重み $(J^T J)^T$ をかけた最急降下法であり，(80) をみたく ξ (正確には “ ξ を生成するようなロボットの位置，姿勢”) はそのポテンシャルの極 (極小または極大) である．したがって，特徴量を冗長にとったビジュアルサーボ問題は，ポテンシャル場 V における経路計画問題と考えることもできる．

安定領域に関連する研究として，Chaumette は目標位置に達する前にカメラの運動が停止するための条件を導出し，目標方向以外の方向にカメラが動く理由を定性的に説明している [8]．この条件はかなり一般的なものであり，Chaumette は制御則の選択に利用する方向で議論を展開している．しかし，どのような場合にどの制御則を選択すれば良いかという結論には至っていない．

Cowan と Koditschek [19] はカメラの動きを平面内に限定した場合に対して navigation function を用いた大域的安定化の方法を提案している．彼らの定式化によると，視覚サーボ問題は糸に通されたビーズ玉を目標位置に位置決めするための各ビーズの軌道プランニング問題となる．きわめて限定された対象に対するシミュレーション例しか報告されていないが，カメラの視野角まで定式化に組み込んだ興味深い論文である．

Malis, Chaumette と Boudet は，画像特徴量とカメラの姿勢パラメータの両方を制御量に組み込むことにより大域的安定化を図る方法を提案した [83]．この方法は，純粋な特徴ベース視覚サーボではないが，大域的に安定である以外に，注視対象物を確実に視野内に収めておくこ

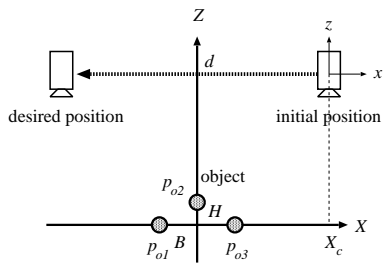


Fig. 23: Camera motion (X translation)

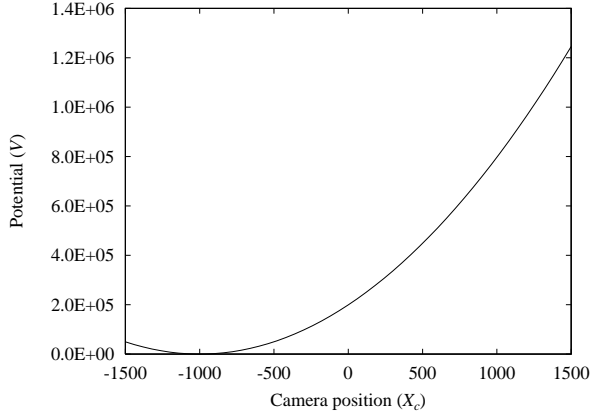


Fig. 24: Potential plot (X translation)

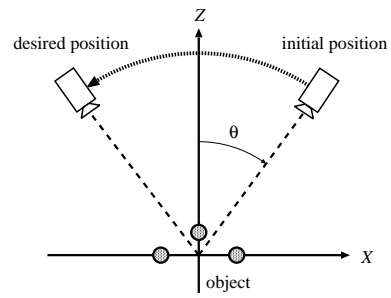


Fig. 25: Camera motion (circular)

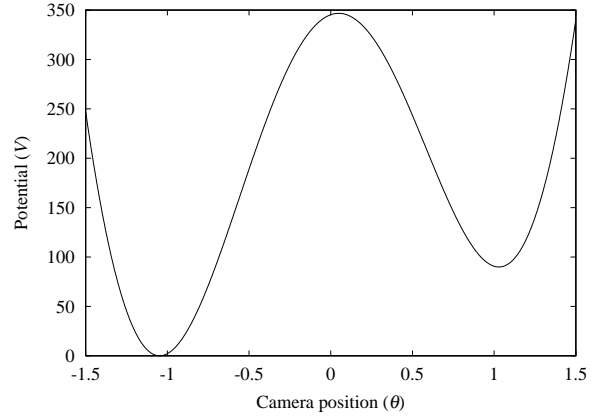


Fig. 26: Potential plot (circular)

とが可能で、カメラキャリブレーション誤差に対してもロバストであり、極めて興味深い。ただし、目標姿勢と現在姿勢の差を実時間で推定する必要があり、計算量が多く、ノイズに対しても弱いと報告されている。

著者らはポテンシャルの極小にトラップされたときに目標特徴量を変化させて人工的にポテンシャル形状を変化させ、極小を抜け出す方法“ポテンシャル切り替え法”を提案した [45]。

11.2 ポテンシャルの例

カメラの運動は作業座標系 $X-Z$ 平面に拘束されており、注視対象物は作業座標系 $X-Z$ 平面におかれた底辺 $2B$ 高さ H の 2 等辺 3 角形で、3 つの頂点を特徴点とする。

11.2.1 1 自由度直線運動

Fig.23 に示すように、カメラの運動を深さ一定の直線上に限定する。つまり、カメラ位置を $p_c = [X_c \ Y_c \ Z_c]^T$ とするとき、カメラの動きを $Y_c = 0, Z_c = d$ (一定) なる直線上の並進運動に限定し、カメラの光軸は鉛直下向に固定する。カメラ座標系の方向は作業座標系と一致させる。カメラの一般化座標を $q = X_c$ ととるとき、 ${}^c J_r = [1 \ 0 \ 0 \ 0 \ 0]$

である。 $f = -256, B = 100, H = 20, d = 1000$ とし¹、目標位置を $X_d = -1000$ に設定したときのポテンシャル $V(X_c)$ は Fig.24 に示すように $V(X_d) = 0$ なる下に凸の 2 次曲線となり、安定領域は大域的である。

11.2.2 1 自由度円運動

カメラの運動を $X-Z$ 平面内の原点を中心とする半径 d の円弧に限定し、カメラの方向は常に光軸が原点を通るように設定する。これは Y 軸を回転軸とする 1 自由度のアームの先に回転軸が画像平面の y 軸と重なるようにカメラを取りつけた場合に相当する。アームが鉛直上向きにあるときの回転角を $\theta = 0$ とする。カメラの運動の自由度は 1 で、一般化座標を $q = \theta$ とする。このとき ${}^c J_r = [d \ 0 \ 0 \ 0 \ 1 \ 0]$ となる。目標画像を $\theta_d = -\pi/3$ のときに得られる画像に設定したときのポテンシャルを Fig.26 に示す。明らかに極小点 θ_{lm} が存在しており、 $\theta > 0.05$ なる初期値に対しては θ_{lm} に収束することがわかる。また、 θ_{lm} において得られる画像を ξ_{lm} とすると、 $\xi_{lm} \neq \xi_d$ であり、極点条件 $J^T(\xi_d - \xi_{lm}) = 0$ をみたす。

¹カメラ座標系においてカメラ前方は ${}^c Z < 0$ となる。したがって $f < 0$ である。

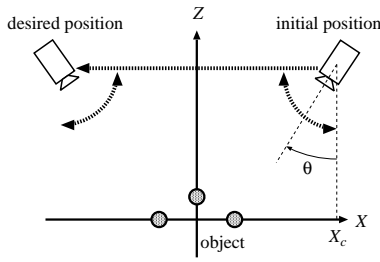


Fig. 27: Camera motion (X translation and rotation)

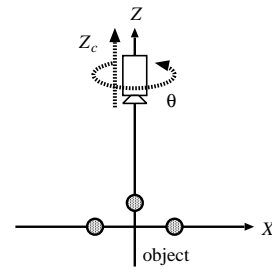


Fig. 29: Camera motion (2D rotation)

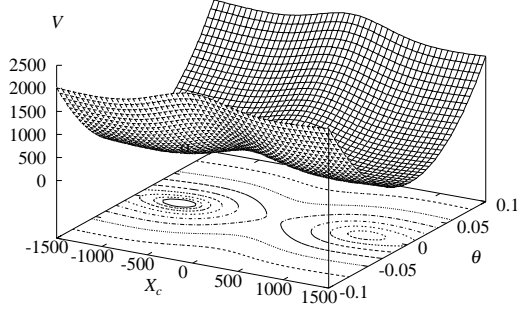


Fig. 28: Potential plot (Close up 2D X and q)

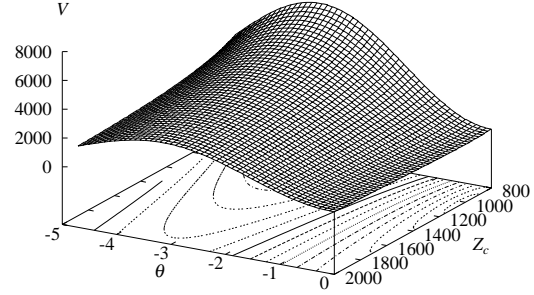


Fig. 30: Potential plot (2D rotation)

11.2.3 2自由度直線運動

カメラの位置は高さ d が一定の直線上に限定するが、光軸の方向を制御できる場合を考える (Fig.27)。つまり、カメラ位置は $Y_c = 0, Z_c = d$ の直線上に限定するが、カメラの光軸は Y 軸まわりに回転できる場合である。このときカメラの運動は X 軸方向の並進 X_c と Y 軸回りの回転 θ の2自由度であり、一般化座標を $q = [X_c \theta]^T$ とする。このとき、ヤコビ行列は

$${}^c J_r = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (81)$$

となる。目標位置を $X_d = -1000, \theta_d = \pi/4$ に設定すると、ポテンシャル面はカメラが対象物の方向を向くとき ($\theta = \arctan(X_c/d)$ のとき) 谷を持つ。ポテンシャル面の谷の様子を詳しく見るために $\theta' = \theta - \arctan(X_c/d)$ なる変換を施し、座標 $-0.1 \leq \theta' \leq 0.1$ の付近を拡大したのが Fig.28 である。ベース面に引かれた等高線からわかるように、この場合も極小点が存在し、その点において極点条件をみやす。

11.2.4 2自由度回転運動

カメラの運動を Z 軸方向の並進と Z 軸回りの回転に限定する。カメラは注視対象物の真上にあり、カメラの

光軸は常に注視対象物方向を向くが、深さの変化と光軸回りの回転運動が可能とする。カメラの自由度は2となり、カメラ高さを Z_c 、カメラの回転角を θ として、カメラの一般化座標を $q = [Z_c \theta]^T$ とする。このとき、ヤコビ行列は

$${}^c J_r = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (82)$$

となる。目標画像を $Z_c = 1000, \theta = 0$ のときに得られる画像に設定し、 $800 \leq Z_c \leq 2000, -1.5\pi \leq \theta \leq 0$ の範囲に対してポテンシャルをプロットする。 $\theta > 0$ の範囲のプロットは $\theta < 0$ の範囲のプロットと対称であるから省略する。結果を Fig.30 に示す。目標画像に対し点対称な初期値 ($q_0 = [1000 \ -\pi]^T$) を与えると、 θ の変化方向には不安定な平衡状態にあり、 Z_c の方向には単調減少であることがわかる。したがって、目標画像に点対称な初期画像を与えると、カメラは回転せずにまっすぐに後退することがわかる。これは Chaumette の観察 [8] と一致する。また、正確に点対称でなくても、それに近い初期値においては θ に関しては目標値方向に傾いているが、 Z_c の増加方向により大きな傾きを持つため、カメラは大きく後退しながら徐々に回転することがわかる。

12 階層的アーキテクチャに基づくビジュアルサーボ

感覚情報と運動情報との相互作用の階層モデル (efferent/afferent interaction model: Fig.2) に基づく階層的ロボット制御モデルについて解説する．ロボット制御システムを階層的にモジュール化することによって反射的な行動と環境に対する適応的な学習が同時に行える．この制御アーキテクチャに基づいて DSP 並列処理ネットワークが実装され，運動するオブジェクトの動的な把握とマニピュレーションを実現している [93]．

アクティブビジョンによるビジュアルトラッキング，アームの位置制御，ハンドのプリシェイピングなどの複数の異なる種類のタスクを実行するには，階層的構造が適している．そして，動的な把握はこれらのタスクの協調動作として実現される．ビジョン信号を含むすべてのセンサ信号をサンプリング周期 1ms でフィードバックする．そのために 1ms ビジョンチップとアクティブビジョンシステムを開発した [67]．

12.1 階層的並列センサフィードバックシステム

階層システムにタスクを分配するためにセンサ情報を 3 つのカテゴリに分割する．そして制御アルゴリズムを各レイヤに配置する．最後に，階層化モデルに基づく並列処理システムを構成する．

12.1.1 センサ信号の分割

サーボ Fig.31(a) は視覚によって得られるハンド位置誤差をフィードバックしてハンド位置を制御する例を示している．アクティブビジョンは内界センサと同様のもっとも低レベルのセンサとして用いられている．一般的にビジョンセンサはジョイントサーボの上位の外界センサとして用いられる [24, 65]．しかし，この例ではジョイントループは Fig.10 のように完全にビジョンにより制御されている．このような制御構造を実現するためには 1ms ビジョンチップのような高速なビジョンセンサが必須である．

目標軌道の生成 センサフィードバックはハンドの目標軌道を生成するためにも用いられる．Fig.31(b) は動いている障害物に対する回避運動の最適軌道をビジョン信号により生成する例を示している．先の例と同様に，動的に変化する環境においてはリアルタイムのセンサフィードバックが必要である．

状況に応じたタスク切り替え Fig.31(c) は 3 つのサブタスク (グラスピング，オブジェクトハンドリング，障害物回避) が環境の状況に応じて切り替えられる例を示して

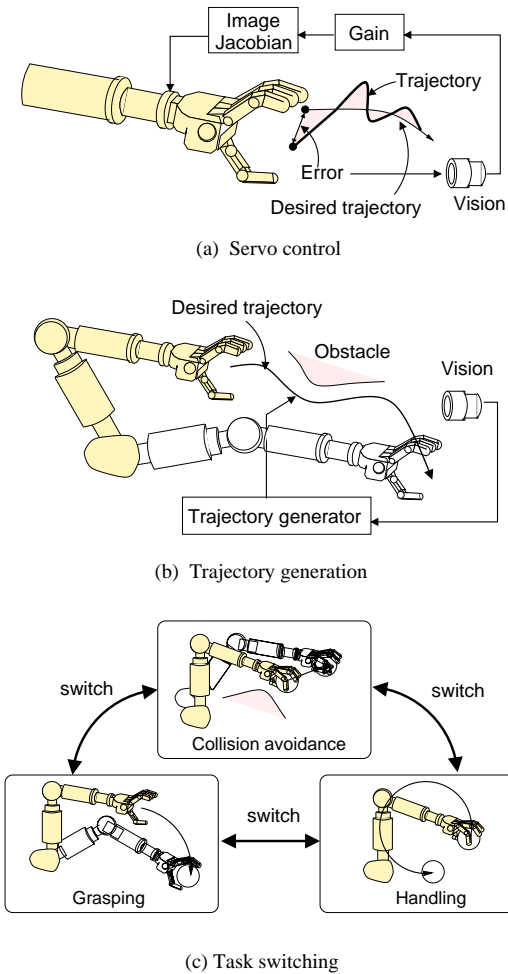


Fig. 31: Use of realtime sensory feedback

いる．最適なサブタスクを選択するために高速なセンサフィードバックが必要である．

12.2 計算モデル

システムの状態を $z \in R^{m_z}$ で表す．たとえばアームの関節角やアクティブビジョンの視線方向など，この状態の一部は制御可能である．そして全状態のなかから制御可能な状態 $\theta \in R^{m_\theta}$ を取り出す関数を f とする．つまり $\theta = f(z)$ とする．この制御可能な状態を運動パラメータ (motion parameter) と呼ぶ．また，センサ信号の表す空間 $s \in R^{m_s}$ を考える．この信号空間はシステムの全状態を推定することができるだけの十分な情報をもつと仮定する．

サーボ層 Fig.31(a) に対応するサーボ層では Fig.2 のモデルのレベル 1 と 2 を実装する．

サーボが必要なサブタスクの数を M_z とする．そして各タスクに対応する状態を $z_i(s) \in R^{m_z}$ とする．ただし $i = 1, \dots, M_z$ である．たとえば $i = 1$ はアクティブビジョンのパン角 q_p に対応し，パン角が全状態 z の第 3 成

分だとすると、 $z_1(s)$ はセンサ信号 s からパン角を選択する．すなわち $z_1 = [0, 0, q_p, 0, \dots, 0]^T$ である．もしパン角の制御に他の信号も必要ならば z_1 はそのセンサ信号も選択し、組み合わせて z の m_z 要素として出力する．

次に、全状態を現在のセンサ出力から推定されるすべてのサブタスクの線形結合として定義する．つまり

$$z = \sum_{i=1}^{M_z} S_i(s) z_i(s). \quad (83)$$

とする．係数行列 $S_i(s) \in \mathbf{R}^{m_z \times m_z}$ ($i = 1, \dots, M_z$) はセンサ出力 s に依存し、パラメータの適応学習に用いられる．すべての係数行列は半正定で、関係式 $\sum_{i=1}^{M_z} S_i(s) = I_{m_z}$ をみたす．ここで I_{m_z} は $m_z \times m_z$ の単位行列である．

すべてのアクチュエータが速度制御である場合、次のような単純な比例制御則が採用できる．

$$\dot{\theta} = \frac{\partial f}{\partial z} (z_d - z), \quad (84)$$

ここで z_d は目標状態である．制御則そのものは重要ではなく、どのような制御則でも良いが、重要なのはシステムの状態の推定方法、すなわち“複数の種類の異なるサブタスクを実現するために適切な情報を選択し、組み合わせること”である．目標状態の生成についてはこのあとに述べる．

12.2.1 行動パターンの生成

このレイヤでは、Fig.31(b) で述べたような目標軌道の生成を行う．Fig.2 のモデルのレベル3 を実装する．種類の異なるサブタスクを取り扱うため、目標軌道（たとえば視線方向、アームの動き、指の形状など）はすべてのサブタスクの軌道の線形結合で記述する．

制御パラメータに影響するすべてのサブタスクの数を M_{dz} とする．各々のサブタスクの目標軌道を $z_{di}(s) \in \mathbf{R}^{m_z}$ ($i = 1, \dots, M_{dz}$) とすると、実際の行動パターン（全状態に対する目標軌道）は

$$z_d = \sum_{i=1}^{M_{dz}} U_i(s) z_{di}(s). \quad (85)$$

で生成される．係数行列 $U_i(s) \in \mathbf{R}^{m_z \times m_z}$ ($i = 1, \dots, M_{dz}$) は半正定で、 $\sum_{i=1}^{M_{dz}} U_i(s) = I_{m_z}$ をみたす．

12.2.2 適応（タスクの選択）

第3のレイヤは Fig.31 (c) に対応する適応であり、サブタスクの選択に用いられる．タスクの選択は係数行列 $S_i(s)$ と $U_i(s)$ により実現される．

これらの基本的アルゴリズムに基づいた階層的並列処理アーキテクチャを Fig.32 に示す．すべてのフィードバツ

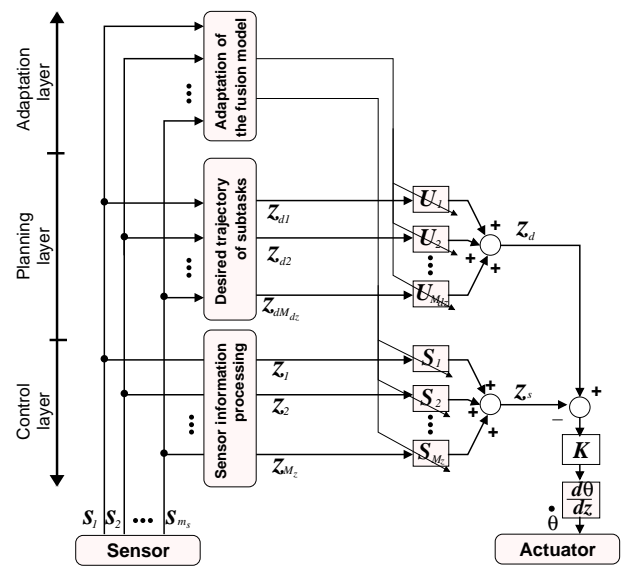


Fig. 32: Hierarchical parallel realtime sensory feedback model

グループは 1ms のサンプリング周期が保証されており、環境の変化に応じた適応とタスクの選択が実装されている．さらに、この計算方式はシステムの構造的な変化（たとえば、フィードバックループの増減）に容易に対応できる．

12.3 アルゴリズム

アルゴリズムと制御則を説明する前に、ロボットの自由度は6とし、多指ハンドとひとつのアクティブビジョンを仮定する．指のすべての関節には力センサがあり、指同士の干渉はそれで計測できる．タスクは一連の一般的マニピュレーション、たとえばオブジェクトのグラスピング、ハンドと他のオブジェクトの相互操作からなる．

12.3.1 アーム

アームの制御には次のモードを設定する．

1. Tracking: ハンドとオブジェクトの相対位置を制御することにより把握のための最適位置を維持する (Fig.33(a))
2. Reaching: アームをオブジェクトに近づける (Fig.33(b))
3. Handling: 把握したオブジェクトを適当な位置に制御する (Fig.33(c))
4. Collision avoidance: 把握したオブジェクトと他のオブジェクトの干渉を回避する (Fig.33(d))

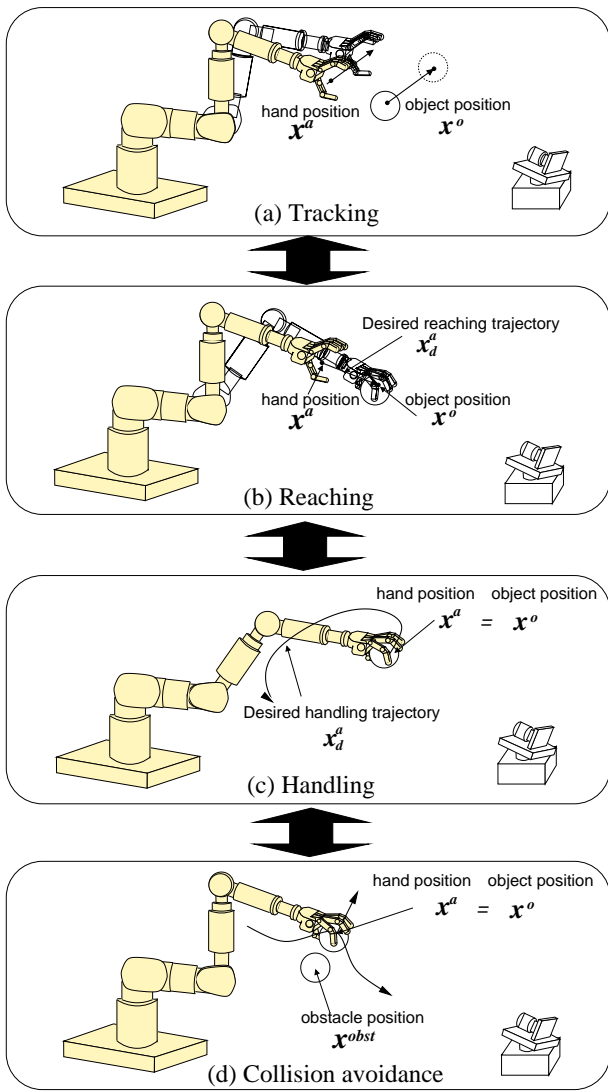


Fig. 33: Desired subtasks

12.3.2 制御

アームのジョイント角を $\theta^a \in R^6$, ハンドの位置・姿勢を $x^a \in R^6$, オブジェクトの位置・姿勢を $x^o \in R^6$, 手先・指にかかる力/トルクを $F^a \in R^6$ とする. オブジェクト位置はビジョンによって, 力/トルクはハンドと指の力センサにより計測される. 典型的な力制御則 [59] により, 目標のジョイント角速度 $v_d^a \in R^6$ は

$$v_d^a = J^{a-1} K^a (x_d - x_s) - K^{av} \dot{\theta}^a + J^{aT} K^{af} F^a \quad (86)$$

に基づいて決定される. ただし x_d と x_s は目標ハンド位置と現在のハンド位置を表す. 行列 $J^a \equiv \frac{\partial x^a}{\partial \theta^a}$ はアームのヤコビアンで行列 K^a , K^{av} , K^{af} は対角のゲインである. 式 (86) において, 第一項はビジョンセンサに基づく位置制御, 第二項は速度フィードバック, 第三項は力フィードバックの項である. (x_d と x_s はビジョンセンサによって計測されるオブジェクト位置とハンド位置によ

り決定される. 詳細は次節)

12.3.3 タスクエンコーディング

(86) 式の第一項に現れる変数は次式によって定義される.

$$\begin{aligned} x_d \equiv & (I_6 - G^a)(I_6 - G^m)(I_6 - G^s) x^o \\ & + (I_6 - G^a)(I_6 - G^m) G^s x_d^a, \\ & + (I_6 - G^a) G^m x_d^o \\ & + G^a x_d^c, \end{aligned} \quad (87)$$

$$x_s \equiv (I_6 - C^m) x^a + C^m x^o, \quad (88)$$

ここで x_d は各サブタスクの目標軌道を統合することにより生成されるハンドの目標軌道である. x_s は実際のハンド位置である. 目標軌道はオブジェクト位置 x^o , ハンドリングのための目標軌道 x_d^o , 障害物回避のための目標軌道 x_d^c , リーチングのための目標軌道 x_d^a に基づいて生成される. 行列 G^a , G^m はサブタスクを選択するのに用いられ, 行列 G^s は作業座標系をトラッキングのための部分空間とリーチングのための部分空間に分割するのに用いられる. ベクトル x_s はタスクに依存したシステムの状態を表し, 視覚情報 x^o と内部センサの情報 x^a (関節角 θ^a から計算される指先の位置など) を統合することによって計算される. 行列 C^m により制御する点を切り替える. つまり, ハンドがオブジェクトを把持しているとき制御点はオブジェクトとなり, それ以外のときには制御点は指先となる.

12.3.4 選択行列

(87) や (88) に現れた係数行列は以下のように定義される.

$$G^s = \text{diag}(1, 0, 0, 0, 1, 1), \quad (89)$$

$$G^m = \text{diag}(g_i^m), \quad i = 1, 2, \dots, 6, \quad (90)$$

$$g_i^m = \begin{cases} 1 & (1 < \bar{g}_i^m) \\ \bar{g}_i^m & (0 \leq \bar{g}_i^m \leq 1) \\ 0 & (\bar{g}_i^m < 0) \end{cases}, \quad (91)$$

$$\bar{g}_i^m = \int_{t'=0}^t \gamma_i^m \text{sgn}(\tau^h(t') - \tau_o^h) dt', \quad (92)$$

$$\tau^h = \sqrt{\sum_{i=1}^4 \tau_i^h{}^T H_i \tau_i^h}, \quad (93)$$

$$G^a = \text{diag}(g_i^a), \quad i = 1, 2, \dots, 6, \quad (94)$$

$$g_i^a = \frac{1}{1 + \exp(\gamma_i^a (l - l_o))}, \quad (95)$$

$$C^m = C^\alpha G^m, \quad (96)$$

ここで τ^h は指の関節トルク $\tau_i^h \in \mathbf{R}^3$ の加重平均のサイズ, $\tau_o^h \in \mathbf{R}$ は閾値, $l \in \mathbf{R}$ はハンドリングしているオブジェクトと障害物の距離, $l_o \in \mathbf{R}$ は回避行動を開始するかどうかの閾値である. 行列 C^α , H_i とスカラー γ_i^m , γ_i^a は適当な定数である.

12.4 アームサブタスクの統合

制御則 (88) の各々のサブタスクの意味を以下にまとめる.

12.4.1 トラッキングとリーチング

トラッキングとリーチングにおいては指の関節トルクはゼロである ($\tau^h = 0$). したがって, 選択行列は $G^m = O_6$, $G^a = O_6$ となる. したがって, 制御則 (86) は以下になる.

$$\begin{aligned} v_d^a = & J^{a^{-1}} K^a \{(I_6 - G^s)(x^o - x^a) + G^s(x_d^a - x^a)\} \\ & - K^{av} \dot{\theta}^a + J^{a^T} K^{af} F^a, \end{aligned} \quad (97)$$

ただしトラッキングの運動方向は $(I_6 - G^s)(x^o - x^a)$ はリーチングの方向 $G^s(x_d^a - x^a)$ と直交するのでトラッキングとリーチングは干渉しない.

12.4.2 ハンドリング

オブジェクトを把握したあとはハンドリングモードにはいる. オブジェクトをしっかりと把握したなら, スwitch関数 g_i^m は1になる. したがって, 選択行列は $G^m = I_6$, $G^a = O_6$ となる. このとき, 制御則 (86) は以下になる.

$$v_d^a = J^{a^{-1}} K^a (x_d^o - x_s^o) - K^{av} \dot{\theta}^a + J^{a^T} K^{af} F^a \quad (98)$$

ここで $x_s^o \equiv (I_6 - C^\alpha)x^a + C^\alpha x^o$ は視覚情報と内界センサの情報を融合することにより生成されたハンドの位置である.

12.4.3 障害物回避

オブジェクトを把握後, 障害物が視野内に現れたときに障害物回避モードに切り替わる. このサブタスクにおいては選択行列は $G^m = I_6$, $G^a = I_6$ になる. したがって, (86) の制御則は

$$v_d^a = J^{a^{-1}} K^a (x_d^c - x_s^o) - K^{av} \dot{\theta}^a + J^{a^T} K^{af} F^a \quad (99)$$

となる. 障害物を避けようとする力は $(x_d^c - x_s^o)$ という項により発生する. この項は視覚情報と内界センサの情

報を融合することにより生成される. 障害物回避モードのときの目標軌道は次式により定義される.

$$x_d^c \equiv \begin{cases} x_s^o + \begin{bmatrix} n^l(l_o - l) \\ \mathbf{0} \end{bmatrix} & \text{if } l_o - l > 0, \\ x_s^o & \text{otherwise,} \end{cases} \quad (100)$$

ただし $n^l \in \mathbf{R}^3$ は障害物からハンドに向かう方向の単位ベクトルである.

12.4.4 補足

サブタスクの統合 このアルゴリズムでは, リーチングからオブジェクトハンドリングへの移行は τ^h (把握力) が引金となる. オブジェクトハンドリングから障害物回避へはビジョンセンサによって得られる l (把握しているオブジェクトと障害物の間の距離) が引金となる.

センサフュージョン センサフュージョンの構造は τ^h (指関節トルク) に基づき変化する. したがって, ハンドで把握するまではオブジェクトに関する情報はビジョンセンサのみによって獲得される. 把握後は, ビジョンセンサと内界センサの情報を融合することによって獲得される.

一般性 本手法では, トラッキングとリーチングが直交しているのをこれらを明確に区別する仕組みは必要ない. しかし, 一般的には, 環境の状態によりトラッキングとリーチングを切り替えるための何らかの条件が必要になる.

適応 適応層における演算に関しては, (90) や (94) 以外の適切な係数を導入することにより同様の効果を発現させることが可能であろう. しかし全体システムの大域的な安定性はみとされる必要がある. この点に関しては今後の課題である.

12.5 実験

12.5.1 実験システム

実験システムは 1ms センサフュージョンシステム [94] が基盤となる. システムの構成を Fig.35 に示す. このシステムは以下の要素で構成され, すべての処理は 1ms のサンプリングタイムで実行される.

- 階層的並列処理システム: 11 個の DSP (TMS320C40, Texas Instrument Ltd.) が相互に結合されている.
- 7 軸アーム: 三菱重工 PA-10
- アクティブビジョンシステム: 2 自由度カメラ台, 高速ビジョンチップシステム SPE-256 (解像度: 16 × 16 pixel) [92]

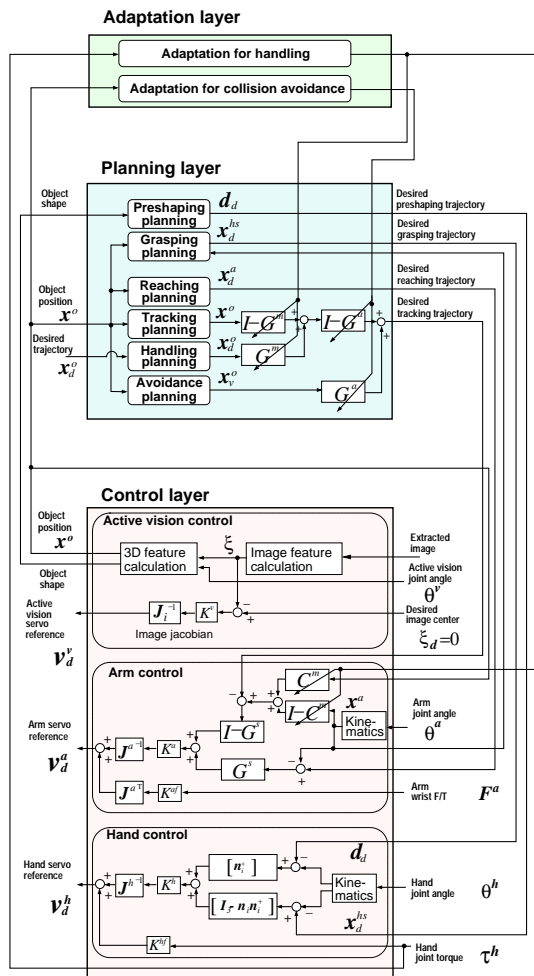


Fig. 34: Block diagram

- 4指ハンド: 14自由度. すべての関節に力センサが配置されている.

12.5.2 実験結果

ハンドリングタスクはあらかじめ定められた位置にオブジェクトを静止保持することとする. Fig.36(a) はオブジェクトとハンドの位置の時間履歴を表す. オブジェクトを把持した後では, 障害物の軌跡もプロットしてある. この図からトラッキング, グラスピング, ハンドリング, 障害物回避のサブタスクが適切に切り替えられ実行されていることがわかる.

サブタスクの切り替えのメカニズムをみるために, 適応のための係数 G^a , G^m の時間履歴を Fig.36(b) に示す. また, 把持力 τ^h の時間履歴を Fig.36(c) に示す. リーチングからハンドリングへのタスク移行は把持力 τ^h の変化に伴う係数行列 G^m の変化により引き起こされることに注意する. また, オブジェクトハンドリングから障害物回避への移行は, 障害物の接近に伴う係数行列 G^a の変化により実現される.

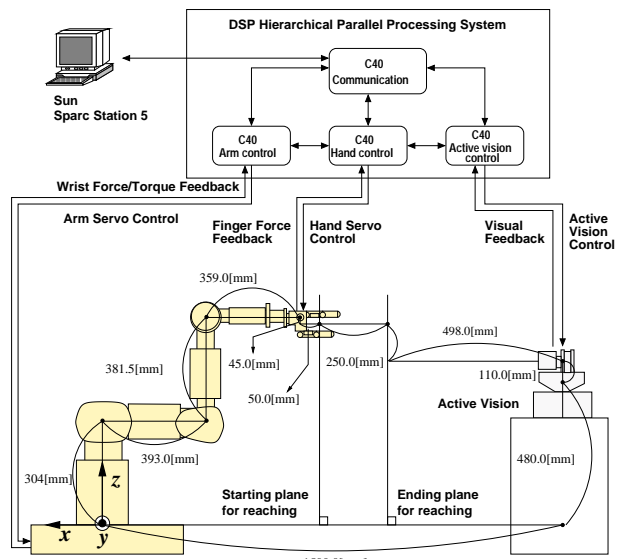


Fig. 35: Experimental system

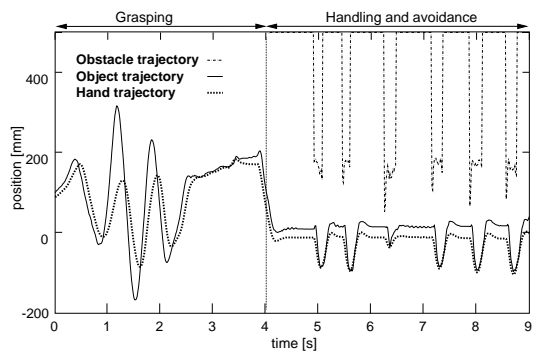
Fig.37 は, 全体プロセスを連続写真で表している. 写真は 0.3[s] ごとに撮影された. トラッキングサブタスクは 0.0[s] から 1.5[s] の間に, オブジェクトハンドリングは 1.8[s] から 3.6[s], 障害物回避は 3.9[s] から 4.5[s] に実現されている.

13 おわりに

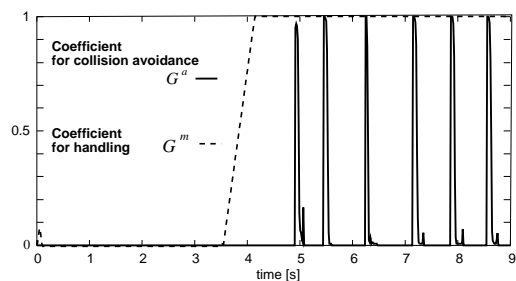
サイバネティクスから脳の機能を概説し, 視覚の基礎, ビジュアルサーボの基礎的事項を少し詳しく解説した. そして最後に脳の機能の工学的実現を目指した階層的センサフィードバックシステムの構築に関し詳しく紹介した.

“見て, 考えて, 理解してから動く” 視覚フィードバック制御の研究は人工知能の応用の一部として研究され, 歴史は古い. しかし我々が作業をするときには, 作業対象物の位置を正確に測定してから手をのぼすわけではなく, 常に手と対象物を見ながら行動する. このような“見ながら動く” ビジュアルサーボ系の研究は始められたばかりである.

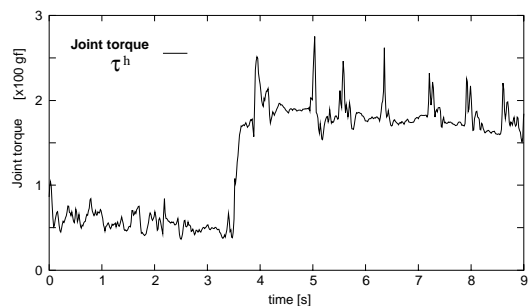
筆者は制御理論の応用として視覚フィードバックの研究を始めた. 制御理論的観点からは理想化されたビジュアルサーボ系は非線形な動的システムとみなせることがわかってきた. 非線形システムの研究は盛んに行なわれているが, ビジュアルサーボ系においては特徴量の抽出と選択, 冗長な特徴量の処理, 視覚ループとモータループのサンプリング周期の違いなど, 定式化さえできていない問題がたくさん残されている. 非線形システムの解析に微分幾何学が有効であることは誰もが認めるところである. ビジュアルサーボ系の解析にもそのようなアプ



(a) Trajectory



(b) Coefficient



(c) Joint torque

Fig. 36: Time response: task switching

ローチ [26, 87, 36] が重要である。

一方、射影幾何や不変量といった画像処理の典型的な解析手法に時間にもなう変化 (time evolution) を導入することにより、動画の解析や兵器の制御、マン・マシンインタフェースなどに応用する研究が、まさに制御理論研究者により開拓されつつある [115, 111, 112, 101]。制御理論は視覚の分野においてもやはり強力な解析手法なのである。

参考文献

- [1] J. S. Albus. A theory of cerebellar functions. *Mathematical Bioscience*, Vol. 10, pp. 25–61, 1971.
- [2] P. K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Trans. Robotics and Automation*, Vol. 9, No. 2, pp. 152–165, 1993.



Fig. 37: Experimental result: collision avoidance

- [3] J. Aloimonos. Perspective approximations. *Image and Vision Computing*, Vol. 8, No. 3, pp. 177–192, 1990.
- [4] 甘利俊一, 酒田英夫. 脳とニューラルネット. 朝倉書店, 1994.
- [5] R. L. Andersson. *A Robot Ping-Pong Player: Experiment in Real Time Intelligent Control*. MIT Press, Cambridge, Massachusetts, 1988.
- [6] F. Bensalah and F. Chaumette. Compensation of abrupt motion changes in target tracking by visual servo. In *Int. Conf. Intelligent Robots and Systems*, pp. 181–187, Pittsburgh, Pennsylvania, 1995.
- [7] Z. Bien, W. Jang, and J. Park. Characterization and use of feature-Jacobian matrix for visual servoing. In K. Hashimoto, editor, *Visual Servoing*, pp. 317–363. World Scientific, Singapore, 1993.
- [8] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In D. J. Kriegman, G. D. Hager, and A. S. Morse, editors, *The Confluence of Vision and Control*, pp. 66–78. Springer-Verlag, London, 1998.
- [9] F. Chaumette, P. Rives, and B. Espiau. Positioning of a robot with respect to an object, tracking it and estimating its velocity by visual servoing. In *IEEE Int. Conf. Robotics and Automation*, pp. 2248–2253, Sacramento, Calif., 1991.
- [10] F. Chaumette, P. Rives, and B. Espiau. Classification and realization of the different vision-based tasks. In K. Hashimoto, editor, *Visual Servoing*, pp. 199–228. World Scientific, Singapore, 1993.
- [11] F. Chaumette and A. Santos. Tracking a moving object by visual servoing. In *12th IFAC World Congress, Vol.9*, pp. 409–414, Sydney, Australia, 1993.
- [12] P. I. Corke. Video-rate robot visual servoing. In K. Hashimoto, editor, *Visual Servoing*, pp. 257–283. World Scientific, Singapore, 1993.
- [13] P. I. Corke. Visual control of robot manipulators — a review. In K. Hashimoto, editor, *Visual Servoing*, pp. 1–32. World Scientific, Singapore, 1993.
- [14] P. I. Corke. *Visual Control of Robots: high performance visual servoing*. Research Studies Press, Somerset, England, 1997.
- [15] P. I. Corke and M. C. Good. Dynamic effects in high-performance visual servoing. In *IEEE Int. Conf. Robotics and Automation*, pp. 1838–1843, Nice, France, 1992.
- [16] P. I. Corke and M. C. Good. Controller design for high speed-performance visual servoing. In *12th IFAC World Congress, Vol.9*, pp. 395–398, Sydney, Australia, 1993.
- [17] P. I. Corke and M. C. Good. Dynamic effects in visual closed-loop systems. *Trans. on Robotics and Automation*, Vol. 12, No. 5, pp. 671–683, 1996.
- [18] P. I. Corke and S. A new hybrid image-based visual servo control scheme. In *Proc. 39th IEEE Conf. on Decision and Control*, pp. 2521–2526, Sydney, Australia, 2000.
- [19] N. J. Cowan and D. E. Koditschek. Planar image based visual servoing as a navigation problem. In *IEEE Int. Conf. Robotics and Automation*, pp. 611–617, Detroit, Michigan, 1999.
- [20] J. Dean. The neuroethology of perception and action. In *Relationships Between Perception and Action*, Edited by O. Neuman and W. Prinz, Springer-Verlag, pp. 81–131, Berlin Heidelberg, 1990.
- [21] 出口光一郎, 石山壘. 画像ベースド視覚サーボにおける並進と回転の非干渉化によるロボットの最適軌道制御. 日本ロボット学会誌, Vol. 17, No. 5, pp. 720–727, 1999.
- [22] 出口光一郎. 能動カメラによる運動物体追跡と実時間3次元形状復元システム. 計測自動制御学会論文集, Vol. 35, No. 5, pp. 675–683, 1999.
- [23] 出口光一郎. ロボットビジョンの基礎. コロナ社, 2000.
- [24] K. Hashimoto ed. *Visual Servoing —Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*. World Scientific, Singapore, 1993.
- [25] M. Ejiri, et al. A process for detecting defects in complicated patterns. *Computer Graphics and Image Processing*, Vol. 2, p. 326, 1973.
- [26] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. Robotics and Automation*, Vol. 8, No. 3, pp. 313–326, 1992.
- [27] J. T. Feddema and C. S. G. Lee. Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera. *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 20, No. 5, pp. 1172–1183, 1990.
- [28] J. T. Feddema, C. S. G. Lee, and O. R. Michell. Weighted selection of image features for resolved rate visual feedback control. *IEEE Trans. Robotics and Automation*, Vol. 7, No. 1, pp. 31–47, 1991.
- [29] J. T. Feddema and O. R. Michell. Vision guided servoing with feature-based trajectory generation. *IEEE Trans. Robotics and Automation*, Vol. 5, No. 5, pp. 691–700, 1989.
- [30] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395, 1981.
- [31] K. S. Fu, R. C. Gonzalez, and C. S. G. Lee. *Robotics: Control Sensing, Vision, and Intelligence*. McGraw-Hill, New Yoak, 1987.
- [32] 藤原幸広, 小森鉄平, 藤田政之, 内田健康. 自動車の自動操舵—ビジュアルサーボによるアプローチ—. 計測と制御, Vol. 39, No. 2, pp. 130–134, 2000.
- [33] J. A. Gangloff, M. de Mathelin, and G. Abba. 6 DOF high speed dynamic visual servoing using GPC controllers. In *IEEE Int. Conf. Robotics and Automation*, pp. 2008–2013, Leuven, Belgium, 1998.
- [34] J. A. Gangloff, M. de Mathelin, and G. Abba. Visual servoing of a 6 DOF manipulaor for unknown 3D profile following. In *IEEE Int. Conf. Robotics and Automation*, pp. 3236–3242, Detroit, Michigan, 1999.
- [35] B. K. Ghosh, M. Jankovic, and Y. T. Yu. Perspective problems in system theory and its application to machine vision. *Journal of Mathematical Systems, Estimation and Control*, Vol. 4, No. 1, pp. 3–38, 1994.

- [36] B. K. Ghosh, M. Lei, and E. P. Loucks. Visionics: A new vision guided estimation of a dynamical system. In *12th IFAC World Congress, Vol.9*, pp. 415–418, Sydney, Australia, 1993.
- [37] B. K. Ghosh and E. P. Loucks. A realization theory for perspective systems with application to parameter estimation problems in machine vision. *IEEE Trans. on Automatic Control*, Vol. 41, No. 12, pp. 1706–1722, 1996.
- [38] B. K. Ghosh, E. P. Loucks, and M. Jankovic. Multistage nonlinear estimation with application to image based parameter estimation. In *13th IFAC World Congress, Vol.F*, pp. 447–452, San Francisco, 1996.
- [39] 橋本浩一. ビジュアル・サーボイング. 計測と制御, Vol. 35, No. 4, pp. 282–285, 1994.
- [40] 橋本浩一, 海老根巧, 木村英紀. 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.
- [41] 橋本浩一. 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.
- [42] 橋本浩一, 木村英紀. ビジュアルサーボイング—非線形オブザーバアプローチ—. 日本ロボット学会誌, Vol. 13, No. 7, pp. 986–993, 1995.
- [43] 橋本浩一, 井上貴博, 木村英紀. ビジュアルサーボイング—非線形制御アプローチ—. 日本ロボット学会誌, Vol. 13, No. 2, pp. 263–269, 1995.
- [44] 橋本浩一, 青木篤人, 則次俊郎. 冗長な特徴量に基づく視覚サーボ. 日本ロボット学会誌, Vol. 16, No. 3, pp. 384–390, 1998.
- [45] 橋本浩一, 田中浩平, 則次俊郎. 視覚サーボにおけるポテンシャル切り替え制御. 計測自動制御学会論文誌, Vol. 36, No. 8, pp. 660–667, 2000.
- [46] K. Hashimoto. Observer-based visual servoing. In N. Xi B. Ghosh and T. J. Tarn, editors, *Chapter 2 of Control in Robotics and Automation*, London, 1999. Academic Press.
- [47] K. Hashimoto, A. Aoki, and T. Noritsugu. Visual servoing with redundant features. In *35th IEEE Conf. on Decision and Control*, pp. 2482–2484, Kobe, Japan, 1996.
- [48] K. Hashimoto, T. Ebine, K. Sakamoto, and H. Kimura. Full 3D visual tracking with nonlinear model-based control. In *American Control Conference*, pp. 3180–3184, San Francisco, Calif., 1993.
- [49] K. Hashimoto, et al. Image-based dynamic visual servo for a hand-eye manipulator. In *MTNS-91*, pp. 609–614, Kobe, Japan, 1991.
- [50] K. Hashimoto, et al. Manipulator control with image-based visual servo. In *IEEE Int. Conf. Robotics and Automation*, pp. 2267–2272, Sacramento, Calif., 1991.
- [51] K. Hashimoto, T. Inoue, T. Ebine, and H. Kimura. Visual servoing based on adaptive identifier. In *1994 Japan-U.S.A. Symposium on Flexible automation, Vol.1*, pp. 39–42, Kobe, Japan, 1994.
- [52] K. Hashimoto and H. Kimura. Dynamic visual servoing with nonlinear model-based control. In *12th IFAC World Congress, Vol.9*, pp. 405–408, Sydney, Australia, 1993.
- [53] K. Hashimoto and H. Kimura. LQ optimal and nonlinear approaches to visual servoing. In K. Hashimoto, editor, *Visual Servoing*, pp. 165–198. World Scientific, Singapore, 1993.
- [54] K. Hashimoto and H. Kimura. Visual servoing with nonlinear observer. In *IEEE Int. Conf. Robotics and Automation*, pp. 484–489, Nagoya, Japan, 1995.
- [55] K. Hashimoto and T. Noritsugu. Observer-based control for visual servoing. In *13th IFAC World Congress, Vol.F*, pp. 453–458, San Francisco, 1996.
- [56] K. Hashimoto and T. Noritsugu. Performance and sensitivity in visual servoing. In *IEEE Int. Conf. Robotics and Automation*, pp. 2321–2326, Leuven, Belgium, 1998.
- [57] K. Hashimoto and T. Noritsugu. Potential switching control in visual servo. In *IEEE Int. Conf. Robotics and Automation*, pp. 2765–2770, San Francisco, Calif., 2000.
- [58] J. Hill and W. T. Park. Real time control of a robot with a mobile camera. In *9th Int. Symp. on Industrial Robots*, pp. 233–246, 1979.
- [59] N. Hogan. Stable execution contact tasks using impedance control. *Proc. of IEEE Conf. on Robotics and Automation*, pp. 1047–1054, 1987.
- [60] B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, Massachusetts, 1986.
- [61] 細田耕, 浅田稔. 構造やパラメータに関する先験的な知識を必要としないフィードフォワード補償器を持つ適応型ビジュアルサーボ系の構成. 日本ロボット学会誌, Vol. 14, No. 2, pp. 159–165, 1996.
- [62] 細田耕, 五十嵐克司, 浅田稔. 未知環境内で動作するロボットのビジュアルサーボ/力サーボハイブリッド制御系. 日本ロボット学会誌, Vol. 15, No. 4, pp. 642–647, 1997.
- [63] 細田耕, 竈門光彦, 浅田稔. ビジュアルサーボによる脚式ロボットの揺動の実現. 日本ロボット学会誌, Vol. 17, No. 5, pp. 750–753, 1999.
- [64] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true Jacobian. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 186–193, 1994.
- [65] S. Hutchinson, G. D. Hager, and P. I. Corke. A tutorial on visual servo control. *Trans. on Robotics and Automation*, Vol. 12, No. 5, pp. 651–670, 1996.
- [66] 乾敏郎, 安藤広志 (訳). D. Marr 著 ビジョン—視覚の計算理論と脳内表現. 産業図書, 1987.
- [67] I. Ishii, Y. Nakabo, and M. Ishikawa. Target tracking algorithm for 1ms visual feedback system using massively parallel processing vision. *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 2309–2314, 1996.
- [68] M. Ito, M. Sakurai, and P. Tongroach. Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *Journal of Physiology London*, Vol. 324, pp. 113–134, 1982.
- [69] M. Jägersand, O. Fuentles, and R. Nelson. Acquiring visual-motor models for precision manipulation with robot hands. *Proc. of 4th European Conf. on Computer Vision*, pp. 603–612, 1996.

- [70] M. Jagersand and R. Nelson. Visual space task specification, planning and control. *Proc. IEEE Int. Symp. on Computer Vision*, pp. 521–526, 1995.
- [71] W. Jang and Z. Bien. Feature-based visual servoing of an eye-in-hand robot with improved tracking performance. In *IEEE Int. Conf. Robotics and Automation*, pp. 2254–2260, Sacramento, Calif., 1991.
- [72] R. A. Jarvis. A perspective on range finding techniques for computer vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 2, pp. 122–139, 1983.
- [73] 徐剛, 辻三郎. 3次元ビジョン. 共立出版, 1998.
- [74] 徐剛. コンピュータビジョンにおけるエピポラ幾何. コンピュータビジョン—技術評論と将来展望—, pp. 80–96. 新技術コミュニケーションズ, 1998.
- [75] 金谷健一. 画像理解—3次元認識の数理—. 森北出版, 1990.
- [76] K. Kanatani. *Geometric Computation for Machine Vision*. Oxford Science Publications, 1993.
- [77] 川人光男. 脳の計算理論. 産業図書, 1996.
- [78] M. Kawato. Feedback-error-learning neural network for supervised motor learning. In *Advanced Neural Computers (R. Eckmiller ed.)*, pp. 365–372. North-Holland, Amsterdam, 1990.
- [79] 木村英紀. 「横割り」学問の復権を目指して: 21世紀の制御理論とその応用. 計測と制御, Vol. 40, No. 1, pp. 16–21, 2001.
- [80] 木村英紀. サイバネティクスとこれからの制御. システム/制御/情報, Vol. 45, No. 6, pp. 157–161, 2001.
- [81] A. J. Koivo and N. Houshangi. Real-time vision feedback for servoing robotic manipulator with self-tuning controller. *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 21, No. 1, pp. 134–142, 1991.
- [82] M. Lei and B. K. Ghosh. Visually-guided robotic motion tracking. In *Lecture held at Osaka University*, 1993.
- [83] E. Malis, F. Chaumette, and S. Boudet. 2-1/2-D visual servoing. *IEEE Trans. Robotics and Automation*, Vol. 15, No. 2, pp. 238–250, 1999.
- [84] D. Marr. A theory of cerebellar cortex. *Journal of Physiology*, Vol. 202, pp. 437–470, 1969.
- [85] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information*. Freeman, W. H. and Company, New York, 1982.
- [86] 丸山章, 藤田政之. 適応 H^∞ 制御に基づくマニピュレータのロバスト視覚 フィードバック制御. システム制御情報学会論文誌, Vol. 12, No. 10, pp. 579–585, 1999.
- [87] H. Michel and P. Rives. Singularities in the determination of the situation of a robot effector from the perspective view of 3 points. Technical Report n. 1850, INRIA, 1993.
- [88] A. P. Morgan and K. S. Narendra. On the stability of nonautonomous differential equations $\dot{x} = [a + b(t)]x$ with skew symmetric matrix $b(t)$. *SIAM J. Control*, Vol. 15, pp. 163–176, 1977.
- [89] J. L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, 1992.
- [90] R. M. Murray, Z. Li, and S. S. Sastry. *Robotic Manipulation*. CRC Press, Boca Raton, Florida, 1994.
- [91] 中坊嘉宏, 石井抱, 石川正俊. 超並列・超高速ビジョンを用いた1msターゲットトラッキングシステム. 日本ロボット学会誌, Vol. 15, No. 3, pp. 417–421, 1997.
- [92] Y. Nakabo, I. Ishii, and M. Ishikawa. High speed target tracking using 1ms visual feedback system. In *Video Proceedings of IEEE Int. Conf. Robotics and Automation*, Minneapolis, 1996.
- [93] A. Namiki, Y. Nakabo, I. Ishii, and M. Ishikawa. High speed grasping using visual and force feedback. *Proc. IEEE Int. Conf. on Robotics and Automation*, 1999.
- [94] A. Namiki, Y. Nakabo, I. Ishii, and M. Ishikawa. 1ms sensory-motor fusion system. *IEEE/ASME Trans. on Mechatronics*, Vol. 5, No. 3, pp. 244–252, 2000.
- [95] B. J. Nelson, N. P. Papanikolopoulos, and P. K. Khosla. Dynamic sensor placement using controlled active vision. In *12th IFAC World Congress, Vol. 9*, pp. 419–422, Sydney, Australia, 1993.
- [96] B. J. Nelson, N. P. Papanikolopoulos, and P. K. Khosla. Visual servoing for robotic assembly. In K. Hashimoto, editor, *Visual Servoing*, pp. 139–164. World Scientific, Singapore, 1993.
- [97] 奥富正敏. ステレオ視. コンピュータビジョン—技術評論と将来展望—, pp. 123–137. 新技術コミュニケーションズ, 1998.
- [98] N. Papanikolopoulos and P. K. Khosla. Adaptive robotic visual tracking: theory and experiments. *IEEE Trans. Automatic Control*, Vol. 38, No. 3, pp. 429–445, 1993.
- [99] N. Papanikolopoulos, P. K. Khosla, and T. Kanade. Vision and control techniques for robotic visual tracking. In *IEEE Int. Conf. Robotics and Automation*, pp. 857–864, Sacramento, Calif., 1991.
- [100] R. P. Paul. *Robot Manipulators: Mathematics, Programming, and Control*. MIT Press, Cambridge, MA, 1981.
- [101] G. Picci, editor. *Dynamical Systems, Control, Coding and Computer Vision*. Birkhauser, 1999.
- [102] A. A. Rizzi and D. E. Koditchek. A dynamical sensor for robot juggling. In K. Hashimoto, editor, *Visual Servoing*, pp. 229–256. World Scientific, Singapore, 1993.
- [103] A. A. Rizzi and D. E. Koditchek. An active visual estimator for dexterous manipulation. *Trans. on Robotics and Automation*, Vol. 12, No. 5, pp. 697–713, 1996.
- [104] A. Rizzi and D. E. Koditschek. Progress in spatial robot juggling. In *IEEE Int. Conf. Robotics and Automation*, pp. 775–780, Nice, France, 1992.
- [105] D. A. Robinson. *Oculomotor control signals*. Pergamon Press, 1975.
- [106] C. Rosen. Machine intelligence research applied to industrial automation. *Technical Report, SRI International*, 1976, 1978.
- [107] S. Sastry and M. Bodson. *Adaptive Control*. Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
- [108] 佐藤淳. コンピュータビジョン—視覚の幾何学—. コロナ社, 1999.

- [109] Y. Shirai. A context sensitive line finder for recognition of polyhedra. *Artificial Intelligence*, Vol. 4, No. 2, pp. 95–119, 1973.
- [110] Y. Shirai and H. Inoue. Guiding a robot by visual feedback in assembly tasks. *Pattern Recognition*, Vol. 5, pp. 99–108, 1973.
- [111] S. Soatto, R. Frezza, and P. Perona. Motion estimation via dynamic vision. *IEEE Trans. on Automatic Control*, Vol. 41, No. 3, pp. 393–414, 1996.
- [112] S. Soatto and P. Perona. Recursive 3-d visual motion estimation using subspace constraints. *Int. Journal of Computer Vision*, Vol. 22, No. 3, pp. 235–259, 1997.
- [113] M. W. Spong and M. Vidyasagar. *Robot Dynamics and Control*. John Wiley & Sons, New York, 1989.
- [114] K. Takase, et al. Torque control manipulators. *ETL Report*, Vol. 37, No. 3, S48.
- [115] A Tannenbaum. Pleanary lecture ii: Visual information in a feedback loop: a control/computer vision synthesis. In *39th IEEE Conf. on Decision and Control*, Sydney, 2000.
- [116] A. Teel, R. Kadiyala, P. Kokotovic, and S. Sastry. Indirect techniques for adaptive input-output linearization of non-linear systems. *Int. J. Control*, Vol. 53, No. 1, pp. 193–222, 1991.
- [117] L. E. Weiss. *Dynamic Visual Servo Control of Robots: an Adaptive Image-Based Approach*. PhD thesis, Carnegie-Mellon University, 1984.
- [118] L. E. Weiss, A. C. Sanderson, and C. P. Newman. Dynamic sensor-based control of robots with visual feedback. *IEEE J. Robotics and Automation*, Vol. RA-3, No. 5, pp. 404–417, 1987.
- [119] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 10, pp. 965–980, 1992.
- [120] D. B. Westmore and W. J. Wilson. Direct dynamic control of a robot using an end-point mounted camera and Kalman filter position estimation. In *IEEE Int. Conf. Robotics and Automation*, pp. 2376–2384, Sacramento, Calif., 1991.
- [121] D. E. Whitney. Resolved motion rate control of manipulators and human prosthetic. *IEEE Trans. Man-Machine Systems*, Vol. MMS-10, No. 2, pp. 47–53, 1969.
- [122] N. Wiener. *Cybernetics*. John Wiley and Sons, 1948.
- [123] W. J. Wilson. Visual servo control of robots using Kalman filter estimates of relative pose. In *12th IFAC World Congress, Vol.9*, pp. 399–404, Sydney, Australia, 1993.
- [124] W. J. Wilson, C. C. Williams Hulls, and G. S. Bell. Relative end-effector control using Cartesian position based visual servoing. *Trans. on Robotics and Automation*, Vol. 12, No. 5, pp. 684–696, 1996.
- [125] P. H. Winston (ed.). *The Psychology of Computer Vision*. McGraw Hill, 1975.
- [126] G. Xu and Z. Zhang. *Epipolar geometry in Stereo, Motion and Object Recognition: A Unified Approach*. Kluwer Academic, 1996.
- [127] 山崎弘郎, 石川正俊 (編著). センサフュージョン, 実世界の能動的な理解と知的再構成. コロナ社, 1992.
- [128] D. S. Zee, J. D. Cook L. M. Optican, D. A. Robinson, and W. K. Engel. Slow saccades in spinocerebellar degeneration. *Arch. Neurology*, Vol. 33, pp. 243–251, 1976.