

Title	動画像を用いた多関節物体の三次元動作認識に関する研究
Author(s)	浅野, 英城
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1112
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士

修士論文

動画像を用いた多関節物体の 三次元動作認識に関する研究

指導教官 阿部 亨 助教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

浅野英城

1998年2月13日

要旨

本稿では、画像情報を用いることで人体などの多関節物体の三次元動作を認識する手法を提案する。オプティカルフローおよび領域モデルを使用したアプローチを紹介するとともに、実験結果により本手法の有効性を示す。

目次

1	序論	1
1.1	研究の背景と目的	1
1.2	従来手法	2
1.3	認識処理の流れ	6
2	オプティカルフローの抽出	8
2.1	オプティカルフローとは	8
2.2	Lucas & Kanade 法	11
2.3	実験および抽出結果	13
2.4	まとめ	15
3	運動パラメータの推定	16
3.1	三次元運動の解釈	16
3.2	領域モデル	19
3.3	運動パラメータの算出	21
3.3.1	Zの推測法	21
3.3.2	上膊部運動パラメータの算出	24
3.3.3	下膊部運動パラメータの算出	26
3.4	まとめ	28
4	領域の姿勢推定・追跡処理	29
4.1	概略	29
4.2	探索範囲	32
4.3	比較評価	32

4.4	上膊部姿勢推定	33
4.5	下膊部姿勢推定	35
4.6	まとめ	37
5	実験と考察	38
5.1	CGによるシミュレーション	38
5.2	実画像に対する実験	43
5.2.1	対象物体	43
5.2.2	二次元方向の運動 (z 軸まわりの回転運動)	43
5.2.3	奥行き方向の運動 (x,y 軸まわりの回転運動)	47
5.3	考察とまとめ	50
6	結論	51

目 次

1.1	処理全体の流れ	7
2.1	時空間勾配の関係 (1次元)	10
2.2	速度空間	11
2.3	動画像からオプティカルフロー抽出	14
3.1	直交射影	17
3.2	座標系と速度ベクトル	17
3.3	領域モデル	19
3.4	スティックと腕の傾きの関係	20
3.5	領域内の点 (画像上)	22
3.6	円筒型の成分	22
3.7	Z の推測	24
3.8	p1 を原点とする local 座標系	25
3.9	領域の境界付近を除去	25
3.10	上膊部、下膊部を含む local 座標系	26
3.11	Z の決定 (上膊部、下膊部)	27
4.1	姿勢推定処理の流れ	30
4.2	モデルの領域推定	31
4.3	ベクトルの評価	33
4.4	モデルの移動 (上膊部)	34
4.5	モデルの移動 (下膊部)	36
5.1	CG モデル (円筒型)	39

5.2	x 軸回転に対する運動パラメータ推定	40
5.3	y 軸回転に対する運動パラメータ推定	41
5.4	z 軸回転に対する運動パラメータ推定	42
5.5	z 軸まわりの回転運動 1	44
5.6	z 軸まわりの回転運動 2	45
5.7	x,y 軸まわりの回転運動	48

第 1 章

序論

1.1 研究の背景と目的

人間や動物といった多関節物体の動作を画像情報から解析しようとする試みは、コンピュータアニメーション、スポーツ選手の運動分析や医用画像処理など多くの分野での有用性により盛んに研究が取り組まれている。画像情報のみからの運動・姿勢解析はマーキングや装着物を使用した手法に比べ人間と計算機との親和性に優れているが、認識精度の低下や処理時間の増大などの問題点がある。このため、認識の高精度化、処理の簡略化、動作・姿勢解析の高速化などの要求を満たした画像解析法の確立が望まれる。

現在のところ画像解析による物体(特に人体、手指)の動作・姿勢認識に関する研究は様々な手法が提案されているが、それらの多くは大きく二通りに分類される。一つは、画像情報から動きや領域などを表す特徴を抽出し、その情報を利用し物体の動作・姿勢を解析するボトムアップ的処理である。ここで特徴とは主にオプティカルフロー、シルエット、エッジ、細線化画像などがある。もう一つは、対象物体の特徴を考慮にいれたモデルを用い、モデルの動きを調べることで認識を行うトップダウン的処理である。モデルには主に、汎用性を重視する観点から作成されたモデル(スティックモデル、リボン)と、より高精度な解析を目指すために作成されたモデル(三次元モデル、ロボットモデル)などが提案されている。近年では上記のトップダウン、ボトムアップ処理を双方向から進めていき、特徴とモデルとを対応付けさせることによって認識処理を行う手法が主流であるといえる。

ただし、従来の手法には様々な問題点が挙げられる。まず特徴抽出に関しては、画像の

ノイズによる影響を受けるため必ずしも正確な情報が得られるとは限らない。これをそのまま使用することは誤認識につながる恐れがある。また、シルエット、エッジなどは姿勢(位置)情報しか持たないため、モデルに記憶してある全ての姿勢候補の中から対応付けを行わなければならないうえ、同様の処理を毎フレームで実施しなければならず、結果的に多くの処理時間を要する。また、運動情報を持たない特徴(細線、シルエットなど)はそれだけでは物体の動作解析は行えないので、モデル側に特別な条件を与え、それを利用し解析を行わなければならない。

モデルにもいくつかの問題点がある。モデルに汎用性を求めれば、その分特徴との対応付けが正確に行えない、また二次元的なモデルは作成は簡単であるが、複雑な動作(例えば三次元へ拡張した動作)には対応できない。つまり、認識可能な動作・姿勢が限定される。逆に高精度な認識を求めれば、精密なモデルが必要である。しかしこれではモデル作成が複雑なうえ、汎用性に欠ける。

そこで本手法はオプティカルフローと領域モデルを使用した多関節物体の動作・姿勢推定法を提案する。オプティカルフローは領域情報(同一物体の上のフローは一樣である性質)と運動情報(画素単位での動きをベクトルで表示)の両方を持つので動作認識に有効であり、さらに姿勢推定の処理を軽減させる働きも持つ。また本手法では、オプティカルフローから三次元動作を解析する手法を取り入れることにより対象物体の三次元動作・姿勢を推定する。

領域モデルは2点と幅情報を与えることにより簡易に作成できるものを用いる。これにより汎用性のあるモデルでの解析が行える。さらにこのモデルに三次元運動をあらゆるパラメータ(運動パラメータ)を持たせることで物体の三次元動作に対応させる。また、オプティカルフローというのは局所的な情報であり画像のノイズに影響されやすいので、これを大局的な情報として活用するためにモデルを使用しフローを統合する。

本手法では、人体などの多関節物体の三次元動作および姿勢の推定を行うことを目的とする。

1.2 従来手法

ここでは、画像情報による物体の動作・姿勢認識に関する従来の研究についていくつか紹介する。

画像情報による物体の動作・姿勢認識法の多くは、対象物体を参考にしたモデルを使用

し、それを画像から抽出された特徴とマッチングさせることで認識を行う手法をとる。モデルに注目すると、より対象に近い精密なモデルを使用する手法と、汎用性を重視した簡易なモデルを使用する手法とに分類される。

まずは、精密なモデルを使用する手法を紹介する。

精密なモデルを用いた姿勢推定

亀田ら [1] は、1枚のシルエット画像から、その画像中に写された関節物体の姿勢を、モデルマッチングによって推定する手法を提案している。対象となる関節物体の精密なモデルをあらかじめ構築しておき、モデルの投影とシルエット画像との被覆率によりモデルマッチングを行っている。モデルマッチングは、精密な人体モデルの接続関係や知識に基づきモデルの投影を計算しシルエット画像の輪郭を正確に照合することにより各体節の輪郭がはっきりしている場合には、正確な姿勢推定が行える。しかし、モデルの姿勢候補の中から全て照合させ推定するので、処理時間が非常にかかる。また、体節どうしが重なった場合(オクルージョン)の正確な推定が行えない。さらに被験者に対する精密なモデルを必要とするため、汎用性にかけるなどの欠点がある。

三次元モデルを用いた手指姿勢の推定

島田ら [3] は、シルエット画像から三次元形状モデルを用いて手指姿勢を解釈する手法を提案している。手指モデルを手の平部分と指部分に分け、シルエット領域の突起部分に注目して姿勢の推定を行っている。この手法ではモデルの自由度を制限したうえで、取りうる全ての姿勢候補とシルエットとのマッチングをしているが、処理時間の画期的な短縮には至っていない。

CADモデルとオプティカルフローによる人体動作推定

山本ら [16] はCADモデルによる立体モデルを使用し、それをオプティカルフローと対応付けさせ人体の動作および姿勢の推定を行っている。画像上の点および速度ベクトルをモデルの幾何学的構造と対応させ、モデル全体で最適化することで姿勢を推定している。また、運動の予測や追跡ずれの修正を行うことで高精度に認識を行っている。しかし立体モデルは物体の三次元形状に対応させやすい分、構築が複雑かつ多大な知識を要する。

次に汎用性を重視した簡易モデルを使用する手法を挙げる。

リボンモデルを使用した姿勢推定

藤富ら [7] は、人間の姿勢を 1 枚の静止画像から推定する手法を提案している。画像中の関節物体を一般化円筒として捉え、二次元的に処理するために人物をリボン表現で記述する。リボンはシルエット画像の細線化で得られる骨格情報をもとに記述され、各リボンに対して接続条件（関節特性などの拘束条件）や形状に関する知識を与えることでパーツの対応づけを行ない、この記述をもとに姿勢推定を行なっている。対象物体として認識の容易な人形を用いていることで、特徴抽出が容易な環境を作り上げているが、誤認識が多い。また、パーツ間の二次元的長さを一定としているためカメラ軸方向に曲がっている軸については認識できない。さらにパーツ同士が接近している場合（ex. 手と胴）にはパーツの幅情報が使えないので誤認識してしまうおそれがある。

エッジ画像とリボンを使用した関節物体の姿勢推定

倉掛ら [4] は、物体の異なる見えや変形を含む複数の画像を処理することで関節部位を検出し、関節物体のセグメンテーションと記述の獲得を行なう手法を提案している。一般化円筒の二次元版であるリボンをパーツの基本的表現とするため、各フレームのエッジ画像からリボンの初期記述を得て、フレーム間のリボンマッチングにより初期記述を比較して関節部位を検出している。そして検出された部位をもとに各初期記述を選択的に統合して最終的な記述を得ている。

しかしながらこの手法は、物体の関節部を抽出することに重点をおいているため、複雑な姿勢および動作の解析までは十分に行えていない。またこのリボンでは物体の三次元姿勢および動作までは認識できないなどの欠点がある。

スティックモデルによる歩行の認識

木本ら [5] は、対象を人物の歩行運動に限定し、画像上での二次元的な足の動きを三次元スティックモデルの動きパラメータで記述することを試みている。ここでは、足のモデルを簡単なものから複雑なものまで 3 段階に分け、関節のないモデルから順次複雑なモデルへと変化させることにより、足の動きパラメータを段階的に求めている。

しかしながら、被験者の歩行スタート位置や歩行方向などが指定されており認識の自由度が低い。

スティックモデルを用いた手振りの認識

クンラボンら [20] は、手振りをおこなう人物の上半身画像を対象にして、輝度情報と動き情報のほか、対象物に関する知識を利用した領域分割結果に基づくスティックモデルの生成法を提案した。この手法では、対象画像中の人物領域を切り出せると仮定し、人体を六つの部位に分割し、剛体の連結体で近似している。領域割り当て問題は MRF (マルコフ確率場) によってモデル化され、各部位および部位間の属性に関する知識による最適化問題によって得られる。

しかし、領域割り当て問題のためにモデルおよび方法に対する設定が複雑になることや、全ての割り当て候補に対し評価を行うと計算コストが莫大になるなどの欠点がある (そのため様々な条件付けがされている)。また、得られる特徴はあくまで二次元的なものなので必然的に認識できる姿勢も限定される。

オプティカルフローと領域モデルを使用した動作・姿勢推定

安藤ら [6] は、オプティカルフローを領域モデルを使って統合し、大局的な情報として活用し、人物の腕の動作を推定した。オプティカルフローは局所的な情報ではあるが、領域情報と運動情報の両方を持つので、動作認識において有効な特徴である。

しかし、この手法ではオプティカルフローから二次元的動作しか解析していない。したがって、認識可能な動作も二次元的に制限される。

以上の画像情報による認識法は、特徴抽出の難しさや動作・姿勢解析の複雑さなどの面から、高精度な認識法は未だ提案されていないのが現状である。特徴抽出に関しては、多くの手法はエッジやシルエットなどの姿勢 (位置) 情報のみを持つ特徴を使用しているが、これでは動作認識は行えないどころか姿勢推定においてもモデルの姿勢候補が多くなり、結果計算時間がかかる。

モデルに関しては、精密なモデルほど複雑な姿勢に対応しているが、汎用性が低いうえ、モデルに与える情報、知識が多くなる。また、複雑な処理を施すほど計算コストは増大してしまう。

そこで、これらの問題点を改善するために以下の特徴を持った手法が望まれる。

- 動き情報を持った特徴を利用することで、姿勢推定における探索範囲を縮小する。
- 汎用的かつ簡易なモデルを使用することで、モデルに関する知識を軽減させる。
- 局所的かつ誤差を含みやすい特徴を、部分的な領域で統合し大局的に扱う。

1.3 認識処理の流れ

本手法による物体の三次元動作・姿勢推定処理は以下の手順で行われる (図 1.1)。

1. 動画像からオプティカルフロー抽出 (図 1.1(1.))。
2. 領域モデルにより、領域内すべてのオプティカルフローを統合 (図 1.1(2.))。
オプティカルフローは画素単位での動き情報のため、局所的である。これを体節単位での運動として大局的に活用する。
3. 運動パラメータの推定 (図 1.1(3.)) 。
4. 運動パラメータを利用し、比較評価により次フレームでのモデルの姿勢を推定 (図 1.1(4.))。
5. 1. ~ 4. を全てのモデルに対して行い、人体としての姿勢を推定。(図 1.1(5.))
体節ごとの動きを今度は人物全体の動きとして統合し、これにより人体の動作・姿勢推定を目指す。

以上の処理を各フレーム毎に行うことで、対象物体 (本研究では人体の腕) の三次元動作認識および姿勢の追跡を行う。

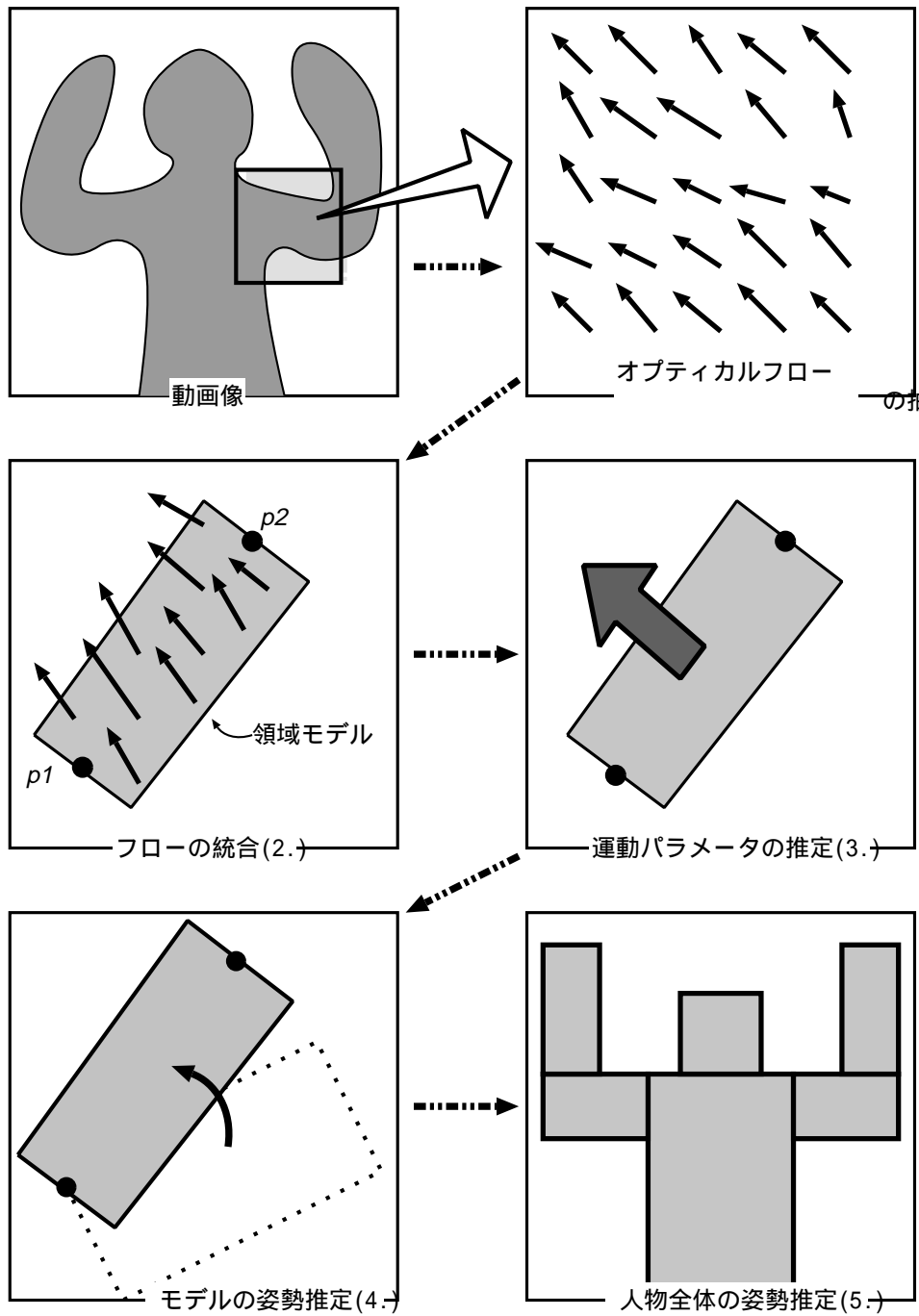


図 1.1: 処理全体の流れ

第 2 章

オプティカルフローの抽出

画像情報を用いて人体などの姿勢認識を行う場合、シルエット画像を用いる手法や人体領域を細線化する手法などが提案されているが、これらの多くは姿勢 (位置) 情報のみしか持たない。そのため、モデルに設定された全ての姿勢候補の中から最適な姿勢を検出する必要があるうえ、毎フレームで同様の処理を施さねばならず計算時間を要する。

これに対してオプティカルフローは領域情報 (同一物体上のフローは一様である性質) と運動情報 (運動ベクトル) の両方を持っているので、動作認識に非常に有効であり、かつ姿勢推定の処理を軽減させる役目も果たす。

本章ではまずオプティカルフローについて説明し、本手法で用いた Lucas & Kanade 法の紹介とともにフローの出力結果も併せて示す。

2.1 オプティカルフローとは

動画像からの速度検出に関する研究は 1970 年ころより取り組まれている。その多くは動画像中の濃淡パターンの対応付けの考えをもとに速度場を計算する手法が一般的である。こうした対応付けの考えに基づき、動画像より検出される見かけの速度場に相当するのがオプティカルフローベクトルとよばれる。従来より提案されている代表的な手法は、マッチング法とグラディエント法に大別される。

マッチング法とは一言でいうと相関による対応付けに基づきフローベクトルを抽出する方法である。連続する画像フレーム間において明度一定の仮定により、現フレームでのある小領域 (ウィンドウと呼ぶ) の明度パターンは次フレームでの対応するウィンドウの明

度パターンと (理想的には) 一致する。この考えにより探索範囲内で明度パターンの最も一致するウィンドウ領域を見つけ、得られる変位ベクトルから速度場を決定する方法である。ただしマッチング法の問題点は一般的に計算時間が多いことや、必ずしも明度パターンが一致せず曖昧さが生じやすいことなどが挙げられる。

これに対してグラディエント法は画像輝度の時空間微分によりフローベクトルを抽出する方法である。画像上で直接的な対応付けはせず、時空間における輝度勾配の拘束を用い移動ベクトルを抽出する手法である。いま、画像上のある点 (x, y) の時刻 t における輝度を $I(x, y, t)$ とし、微小時間 δt 後の移動点を $(x + \delta x, y + \delta y)$ とする。物体上の輝度は不変であると仮定すれば以下の式が成立する。

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (2.1)$$

上式の右辺を Taylor 級数展開すると、

$$I(x, y, t) = I(x, y, t) + \delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} + e \quad (2.2)$$

ここで e は $\delta x, \delta y, \delta t$ に関する 2 次以上の高次の項で微小であるとし無視する。

両辺を δt で割り、

$$\frac{\delta x}{\delta t} \frac{\partial I}{\partial x} + \frac{\delta y}{\delta t} \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (2.3)$$

δt の極限として $\delta t \rightarrow 0$ とすると、次式を得る。

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (2.4)$$

ここで、見かけの速度ベクトルを $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, 空間的な輝度勾配を $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, 時間的な輝度勾配を $I_t = \frac{\partial I}{\partial t}$ とすれば、次のように書き直せる。

$$I_x u + I_y v + I_t = 0 \quad (2.5)$$

これはフローベクトルの成分 u と v に関する拘束を表しているためオプティカルフローの拘束方程式と呼ばれる。

また、拘束方程式は以下の形式で書かれることもある。

$$\nabla I(\mathbf{p}, t) \cdot \mathbf{v} + I_t(\mathbf{p}, t) = 0 \quad (2.6)$$

上式において、 $\mathbf{p} = (x, y)$ 、 $I(\mathbf{p}, t) = I(x, y, t)$ 、 $\mathbf{v} = (u, v)^T$ 、 $I_t(\mathbf{p}, t) = \frac{\partial I(\mathbf{p}, t)}{\partial t}$ 、 $\nabla I(\mathbf{p}, t) = (I_x(\mathbf{p}, t), I_y(\mathbf{p}, t))$ である。

ここで、式 (2.6) の直接的な意味を説明する。式 (2.2) の Taylor 展開で、2 次以降の高次項を無視しているため、輝度勾配は線形で近似されている。簡単のために画像平面を 1 次元に縮退して考える (x のみの 1 次元とする) とフローと輝度勾配の関係は図 2.1 で表される。いま点 (x_0) 上にある物体の移動ベクトル (速度ベクトル u) を求めたい。そこで、 x_0 での時間的な輝度の変化量 (I_t) を調べる。ここで空間的な輝度勾配 I_x は一定なので以下の式が成立する。

$$I_x \cdot u + I_t = 0 \tag{2.7}$$

これを二次元に拡張すれば式 (2.6) が得られる。 u, v を座標軸とする二次元空間 (速度空間 (図 2.2)) を考えるとオプティカルフローの拘束方程式を満足する (u, v) の値は速度空間上で直線上に位置することになる。

本手法では上記のグラディエント法によりオプティカルフローを求めることにした。

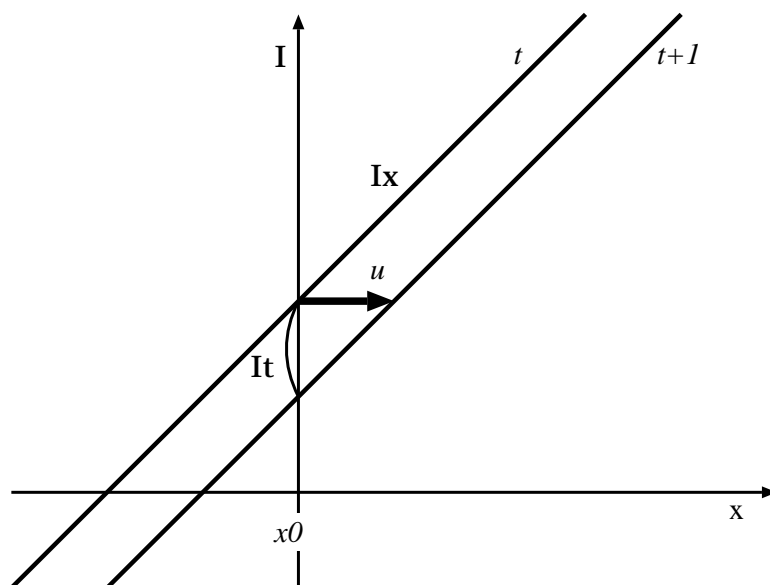


図 2.1: 時空間勾配の関係 (1 次元)

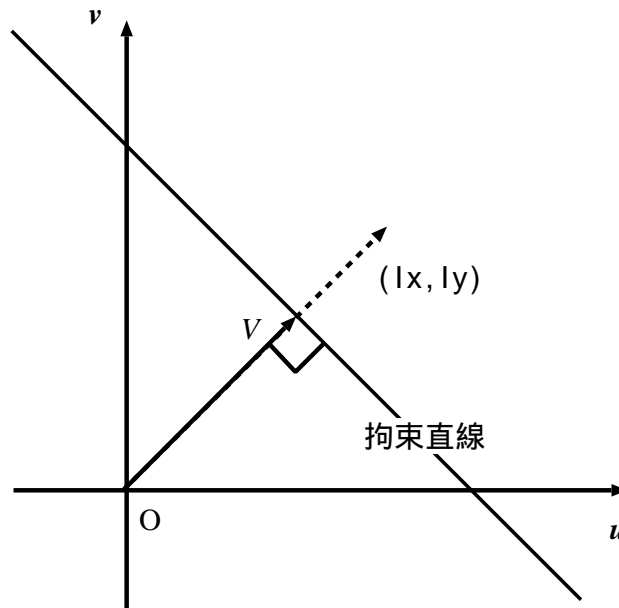


図 2.2: 速度空間

2.2 Lucas & Kanade 法

実際にフローベクトルを求めようとすれば、一般的に式 (2.6) だけでは各点におけるフローを一意に決定できない。そこでこのフローを決定するための様々な手法が提案されている。本手法ではその中でも Lucas & Kanade の提案した拘束条件を利用しフローベクトルを求めることにした。この手法を採用した理由は Barron らが行った比較研究において、最も良好な結果を得ているからである [2]。

Lucas & Kanade は、「同一物体の濃淡パターン上の局所領域では、オプティカルフローはほぼ一定と考えられる、すなわち局所領域で得られるオプティカルフローの拘束方程式は同一の解をもつ」という仮定から、オプティカルフロー V を求めた。つまり、各局所領域 Ω 中の V の拘束として、オプティカルフローの拘束方程式 (2.6) の加重二乗和

$$\sum_{p \in \Omega} W^2(p) [\nabla I(p, t) \cdot V + I_t(p, t)]^2 = E^2 \quad (2.8)$$

の最小化を考えた。ここで $W(x)$ は近傍の周囲よりも中心における拘束に重みをおくための窓関数である。式 (2.8) の解は、 $\frac{\partial E^2}{\partial u} = 0, \frac{\partial E^2}{\partial v} = 0$ より

$$A^T W^2 A \mathbf{V} = A^T W^2 \mathbf{b} \quad (2.9)$$

を解くことで得られる。

ここで単一時間 t における n 個の点 $\mathbf{p}_i \in \Omega [i = 1, \dots, n]$ に対して、

$$\begin{aligned} A &= [\nabla I(\mathbf{p}_1), \dots, \nabla I(\mathbf{p}_n)]^T, \\ W &= \text{diag}[W(\mathbf{p}_1), \dots, W(\mathbf{p}_n)], \\ \mathbf{b} &= -(I_t(\mathbf{p}_1), \dots, I_t(\mathbf{p}_n))^T \end{aligned}$$

また、

$$A^T W^2 A = \begin{bmatrix} \sum W^2(\mathbf{p}) I_x^2(\mathbf{p}) & \sum W^2(\mathbf{p}) I_x(\mathbf{p}) I_y(\mathbf{p}) \\ \sum W^2(\mathbf{p}) I_x(\mathbf{p}) I_y(\mathbf{p}) & \sum W^2(\mathbf{p}) I_y^2(\mathbf{p}) \end{bmatrix} \quad (2.10)$$

であり実際にはこの 2×2 行列 $A^T W^2 A$ が正則なときに解が求まる。

また、 $A^T W^2 A$ の固有値 λ はフローの信頼性の評価値として用いることができる。ここで固有値 λ が表しているのは空間上の輝度勾配の大きさである [17]。一般に輝度勾配の大きい所ほどフローは正確で、勾配の小さい所ほど不正確であると考えられる。例えば背景が白い空間上に白い物体が移動したとしても正確に動きが捕らえられないのと同じで、フローを正確に求めるにはある程度輝度勾配がある所でないといけない。そこで、 λ が閾値以上の u, v を採用することにする。

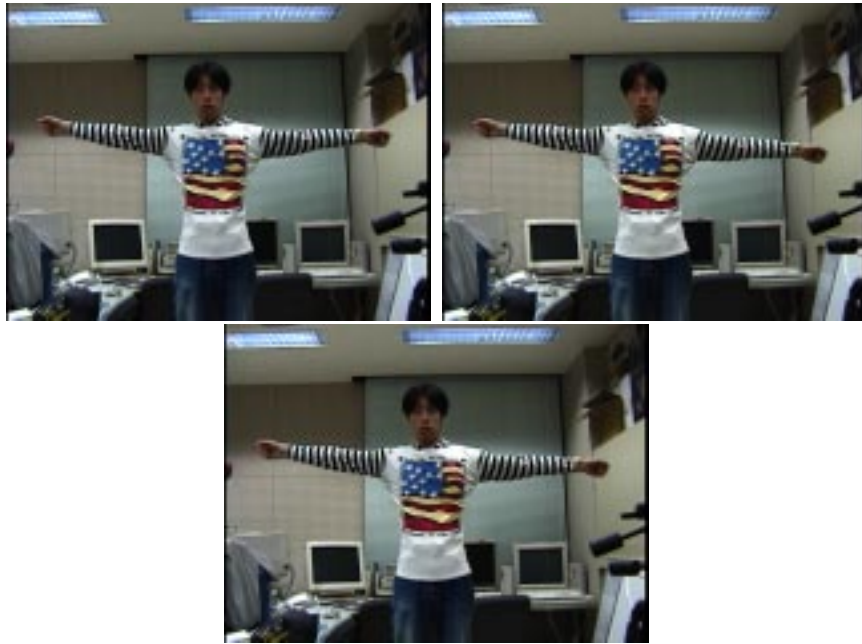
2.3 実験および抽出結果

実際に動画像からオプティカルフローを抽出した例を示す。入力画像についての設定は以下の通りである。

- デジタルビデオカメラで撮影。
- 照明は特に制限せず、一般的な室内蛍光灯のみ。
- 前処理 - ガウシアンフィルタ。
- 320×240 画素、24bit カラー画像、フレーム間隔 $1/30\text{sec}$ 。
- 重み関数 W は二次元のガウス関数を用いる。

オプティカルフロー抽出には複数の連続する画像が必要である。今回は3枚の連続フレーム画像を用い ($Frame(t-1), Frame(t), Frame(t+1)$)、フローを抽出した。これにより求めたフロー場 ($Opt(t)$) は時刻 t での画素単位での速度ベクトルを表している (図 2.3)。

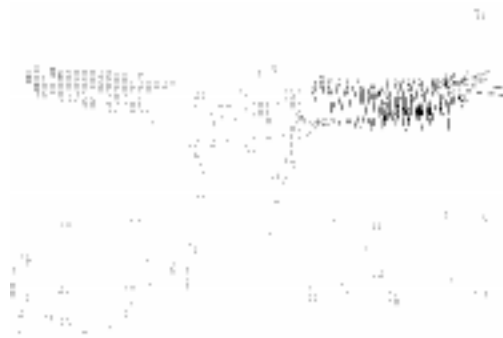
ここで、フローの現れてない箇所は信頼評価により固有値 λ が閾値以下だった点を示している。



Frame($t - 1$)[左上]

Frame(t)[中下]

Frame($t + 1$)[右上]



Opt(t)

図 2.3: 動画像からオプティカルフロー抽出

2.4 まとめ

オプティカルフローの抽出法は、大きく分けてマッチング法とグラディエント法の2つがある。前者は連続フレーム間での直接的な対応付け(相関による対応付け)を行い移動ベクトルを求めるものである。後者は時空間による輝度勾配の拘束に基づき速度場の解析を行うことで移動ベクトル(速度ベクトル)を抽出する方法である。本手法では処理時間の軽減などの理由によりグラディエント法を採用した。また、拘束方程式のみではフローベクトルは一意に決定できないので更なる拘束条件が必要である。いくつかの手法が提案されているが、比較的精度のよい Lucas&Kanade 法を本手法では取り入れている。

第 3 章

運動パラメータの推定

前章では実画像からオプティカルフローを抽出した。そこで次は求めたフローから対象物体の三次元動作を解析する処理を行う。具体的にはオプティカルフローを腕などのパーツの動きとして統合し、パーツとしての運動をパラメータとして表現しこれを算出する。

3.1 三次元運動の解釈

オプティカルフローは二次元の時系列画像から抽出されたものであるため、あくまで二次元的であり画素単位での運動情報しか持たない (x, y 軸方向それぞれのベクトル成分 u, v)。これらの二次元運動情報から物体の三次元運動および姿勢を復元するために以下の制約条件および解析の準備を適用する。[17]

- 剛体性

対象物体の剛体性を仮定することは、三次元運動の解析に数理的手法を用いるために必要な前提条件である。これにより物体像が時間的に変化するとき、その原因が物体自体の形状変化ではなく動作による見かけの変化であることが保証される。本研究では、対象物体として人体と限定しているため、この条件に充分当てはまる。

- 直交射影

通常使用されるカメラモデルとして、直交射影と中心射影がある。本手法では直交射影で撮影されたものと仮定し処理を行う (図 3.1)。これは中心射影の場合のカメ

ラと物体との距離 Z が極めて大きい場合に相当し、空間内の P の座標 (X, Y, Z) が奥行き成分 Z に無関係に画像上の点 $p(x = X, y = Y)$ に射影される。

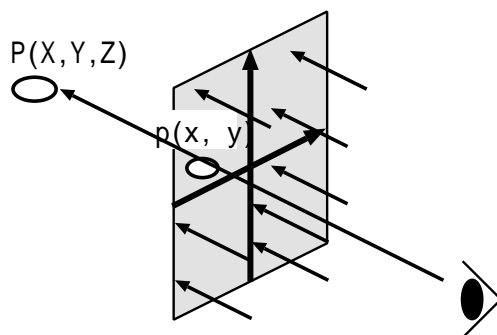


図 3.1: 直交射影

次に物体の剛体運動の一般的解釈を考える。物体が空間内を運動 (移動) するときそれは回転と並進の運動から成り立っている。すなわち物体の中心を通る任意の軸に対する回転と並進である。任意軸に対する回転、並進はそれぞれの運動成分に分解でき、それらのベクトルの合成として表現できる。図 (3.2) にその様子を示す。

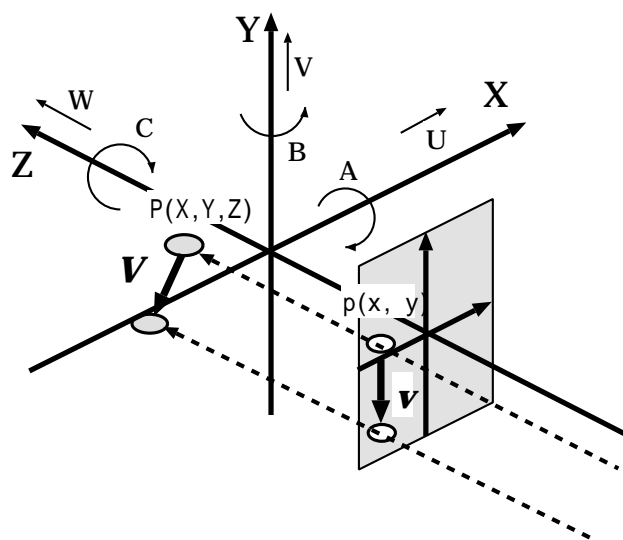


図 3.2: 座標系と速度ベクトル

ここで、各軸に対する軸回りの角速度を (A, B, C) 、軸方向の並進速度を (U, V, W) とする。これら全ての運動は同時発生的に行われている。よって順番に依存しないのでベクトル合成可能である。

ここで、三次元空間での速度場を考えてみる。物体が三次元空間内を角速度 (A, B, C) 、並進速度 (U, V, W) で運動したとする。三次元空間内の点 $P = (X, Y, Z)$ の持つ速度ベクトル $\mathbf{V} = (\dot{X}, \dot{Y}, \dot{Z})$ は次式で表される。

$$\mathbf{V} = \mathbf{T} + \boldsymbol{\Omega} \times \mathbf{P} \quad (3.1)$$

ここで

$$\mathbf{T} = (U, V, W), \quad \boldsymbol{\Omega} = (A, B, C), \quad \mathbf{P} = (X, Y, Z)$$

ただし、本研究では対象物体を人体の腕の動作に限定している。腕の動作というのは関節を軸とした回転運動のみで成り立っていると考えるので、並進運動の成分は無いものとする。よって式 (3.1) の各要素は

$$\begin{cases} \dot{X} &= -BZ + CY \\ \dot{Y} &= -CX + AZ \\ \dot{Z} &= -AY + BX \end{cases}$$

で表わされる。また先ほどの直交射影の説明より、画像上の点 (x, y) がそのまま空間内の点の X, Y 成分なので、

$$(x, y) = (X, Y) \quad (3.2)$$

また画像上のフローベクトル (u, v) は物体の三次元運動の (\dot{X}, \dot{Y}) をそのまま投影したものと考えられるので、これを式 (3.2)(3.2) と併せて、

$$\begin{cases} u &= \dot{X} &= -BZ + Cy \\ v &= \dot{Y} &= AZ - Cx \end{cases} \quad (3.3)$$

となる。これをフローベクトルと三次元運動との関係式とし、運動パラメータの推定に用いる。

3.2 領域モデル

認識対象を限定した場合、画像から抽出された特徴に加え、その認識対象に関する知識を有効に利用するよう、モデルを用いることが行われる。本手法ではオプティカルフロー情報を基に人体の各体節（上膊、下膊など）の運動パラメータと姿勢の推定を行うために領域情報をもったモデルとして円筒の二次元版であるリボンを個々の体節の基本モデルとする。以降、このモデルのことを「領域モデル」と呼ぶ。

領域モデルを使用する目的は次の理由による。オプティカルフローというのは、画素単位での移動量を表したのもであり局所的な情報である。また画像のノイズによる影響を受けやすいなどの欠点をもつ。そこで領域モデルを用いることで領域内の全ての点におけるフローベクトルを統合し、モデル全体として一貫性のある運動を抽出する。これにより局所的なフロー情報を大局的に扱うことができる。

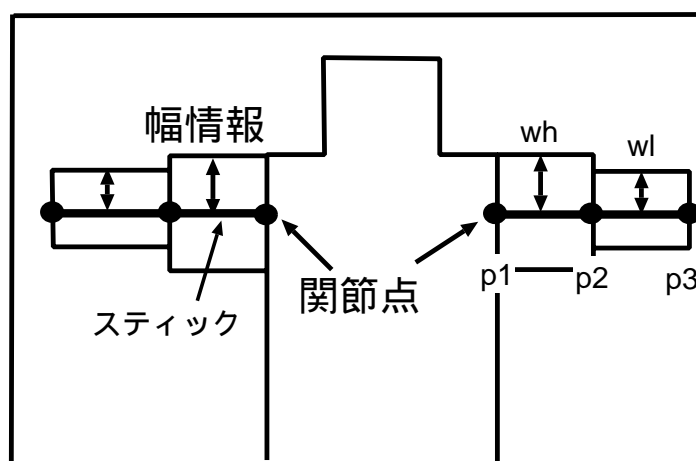


図 3.3: 領域モデル

モデルは人体の上半身を表現したものである。人体を剛体の連結体と考え、その画像面への投影である二次元的形状は長方形で近似し、関節点情報（座標）および幅情報を持たせることで体節の領域を表現する。人体領域は胴体部（頭部含む）、右上膊部、右下膊部、左上膊部、左下膊部の5つの領域に分割する。各領域は長方形のリボンで近似し、リボンの連結により人体の姿勢を表わす。リボン間は関節で接続され、木構造により接続関係を表す。肩を表す関節点を (p1)、肘を表す点を (p2)、そして手首を表す点を (p3) とする。

本研究では対象物体（人体）の動作を、胴体固定のもと腕のみの運動（回転運動）に限定

しているため、実際の認識処理に必要な領域は左右の上膊・下膊部のみである。また、画像情報を直交射影で撮影したものと仮定しているため、実際には物体の前後関係は不明である。そこで、腕の移動範囲を限定し、胴体よりも前で動くものと仮定する。これにより、領域モデルの姿勢から腕の三次元姿勢が解析できる。

次に領域の作成法について述べる。1つの領域は回転中心点 (p) と先端点 (p') の2点の位置情報と領域の幅情報 (w) を持つ。2点は具体的には人体の関節点を指す。以上3つの情報により1つの領域が作成される。ただし、初期設定ではユーザー側が画像を参考に手動で作成し、次フレーム以降は姿勢推定 (次章参考) により次フレームでの2点の座標が得られるので (幅は不変)、自動で作成される。

また、初期設定の段階でスティック (領域の中心を通る線分) の本来の長さ (L) も併せて設定する。この L は腕 (上膊または下膊) の長さに対応する。領域モデルのスティック長 (l) は可変であるが L は不変なので、この l と L を比較することにより腕がどれだけ Z 方向に傾いているか (どれだけ前へ出ているか) を表すことができる (図 3.4)。これにより腕の空間内での姿勢および運動を表現できるのである。

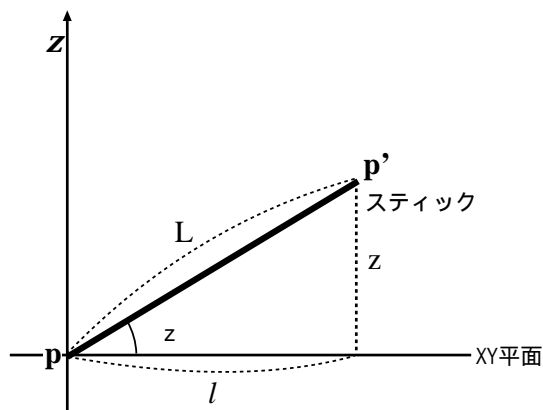


図 3.4: スティックと腕の傾きの関係

領域モデルのスティック長 (l) と腕本来の長さ (L) の関係は以下の通りである。

$$0 \leq l \leq L$$

$l = L$ のときは、腕は XY 平面上に存在し、 $l = 0$ のときは腕は XY 平面に対し垂直に存在していることを示す。

3.3 運動パラメータの算出

3.1 節では物体の三次元運動を表すパラメータについて述べた。本節では腕の運動パラメータ (A,B,C) を推定する方法について述べる。

3.3.1 Z の推測法

3.1 節、式 (3.3) でオプティカルフローと三次元運動パラメータの関係を示した。この式を利用することによりパラメータ (A,B,C) が導出できるが、空間内での Z が既知でなければ求めることはできない。そこで画像上の平面座標と領域モデルを利用することにより、Z を推測する。

いま、図 3.5 に示す領域内の点 (x, y) に対応する物体の Z 成分を求めたい。領域の回転中心点 $p1 = (p1x, p1y)$ 、先端点 $p2 = (p2x, p2y)$ 、幅 (w)、スティック長 (l) とすると、ここで得られる情報は、

- スティックの画像平面に対する傾き (α)
- スティックの空間内の XY 平面に対する傾き (θ_z)
- 点 (x, y) と $p1$ との距離 (d)
- (d) が画像平面になす角度 (θ_{xy})
- 点 (x, y) からスティックへ下ろした垂線の長さ ($h = d \cdot \sin(\theta_{xy} - \alpha)$)
- 垂線とスティックが交わる点と $p1$ を結んだ長さ ($d2 = d \cdot \cos(\theta_{xy} - \alpha)$)

次に図 3.6 に示すものはモデルを円筒型と考えたときのその円筒の切断面である。Z を決定するとき、腕の (XY 平面に対する) 傾きだけでなく腕自体のふくらみも考慮しなければならない。そこでモデルを円筒型と仮定し、その円筒のふくらみを決定する。円筒の幅は領域モデルの幅 (w) をそのまま当てはめる。これにより領域内の点 (x, y) に対応する物体のふくらみを得ることができる。

- 領域内の点 (x, y) に対応する円筒モデルのふくらみ ($Zh = \sqrt{w^2 - h^2}$)

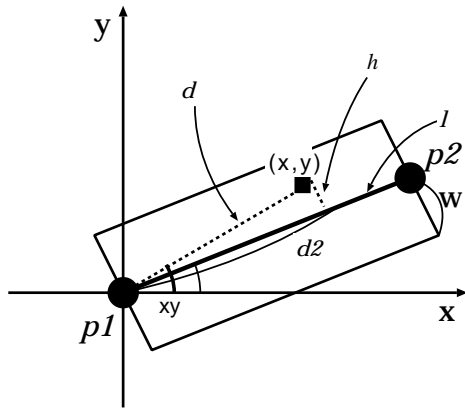


図 3.5: 領域内の点 (画像上)

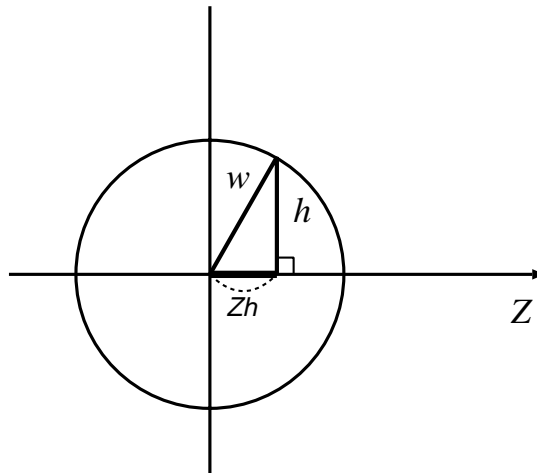


図 3.6: 円筒型の成分

図 3.7はスティックの XY 平面に対する傾きと円筒のふくらみから Z を決定する様子を示したものである。ここで、

- Z_a : スティックの傾きによる z 成分 ($Z_a = d2 \cdot \tan(\theta_z)$)
- Z_b : 円筒モデルのふくらみ成分 ($Z_b = Zh / \cos(\theta_z)$)

以上より、 Z が次式で決定される。

$$Z = Z_a + Z_b \quad (3.4)$$

ここで、上膊部の回転中心点を空間座標系の原点とする。つまり、 $p1$ において $Z = 0$ である。

ただし、ここで注意すべきことは、入力画像を直交射影と仮定しているために z 軸方向に対する動きが制限されることである。本手法では処理の簡単のために、直交射影で撮影された画像を使用するものと仮定し、処理を行っている。このため、厳密には物体の奥行き情報は獲得できない。しかし実画像は中心射影で撮影されているため、実際は焦点距離と物体の空間座標の z 成分が大きく関わってくる。

そこで以下の条件を用意する。

1. 腕は胴体に対し平行、およびそれよりも前方に動くものとする。
2. z 軸方向に対する物体の移動量は、カメラと物体との距離 Z よりも微小である。

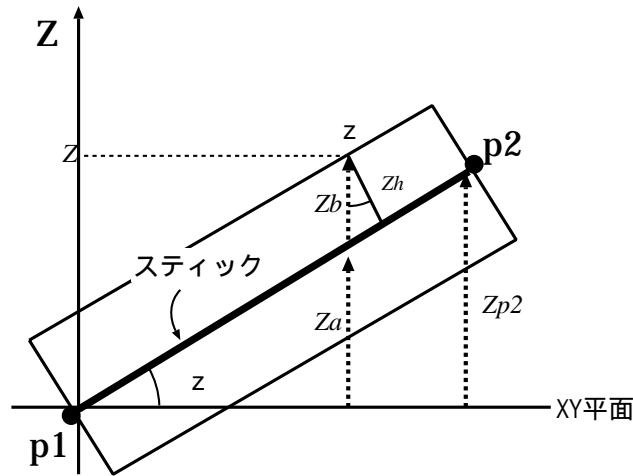


図 3.7: Z の推測

3.3.2 上膊部運動パラメータの算出

本節では、上膊部の運動パラメータ (A,B,C) 算出の過程について説明する。上膊は固定された胴体と連結されているため、その動作は他の体節による運動の影響を受けない。つまり、領域内に現れているオプティカルフローは上膊部のみの運動を表しているものとして扱うことができる。(図 3.8)

$$\begin{cases} u_{(upper)} = -BZ + Cy_{(upper)} \\ v_{(upper)} = AZ - Cx_{(upper)} \end{cases} \quad (3.5)$$

ここで、

- $(u_{(upper)}, v_{(upper)})$: 上膊部モデルの運動を表すフローベクトル
- $(x_{(upper)}, y_{(upper)})$: 上膊部の回転中心座標 ($p1 = (p1x, p1y)$) を原点とした local 座標
 $(x_{(upper)}, y_{(upper)}) = ((x - p1x), (y - p1y))$
- Z : 画像上の座標 (x, y) に対応する、円筒モデルの z 成分

ここで、(上膊部) 領域内の全ての点におけるフローベクトル (u, v) を取り出し、運動パラメータの計算を行いたいが、領域の境界付近におけるフローは輝度勾配が不連続であるため正確な値が抽出されていない場合が多い。そこで、領域の境界付近のフローは処理

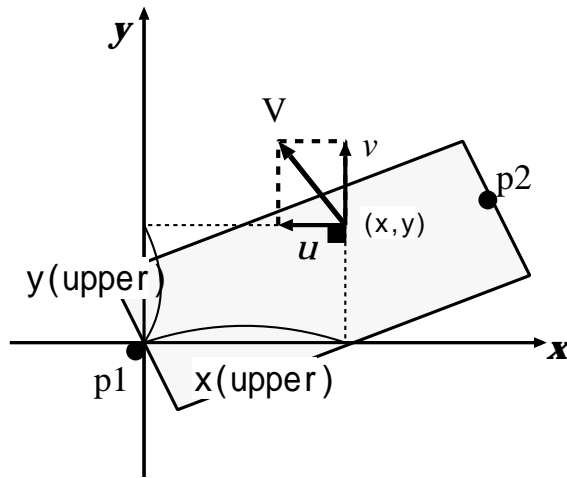


図 3.8: p1 を原点とする local 座標系

に用いないことにする。具体的には図 3.9 に示すように、領域の幅およびスティック長の 80% 以内に存在する点のフロー値のみを使用することにする。

以上により、以下の式 (3.6) に示す最小自乗問題を解くことで、上膊部の運動パラメータ (A,B,C) を導出することができる。

$$\sum_{p \in \Omega} \left\{ (u - BZ + Cy_{(upper)})^2 + (v + AZ - Cx_{(upper)})^2 \right\} \quad (3.6)$$

Ω は上膊部の境界付近を除く領域を示す。

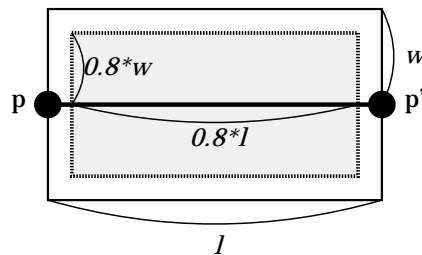


図 3.9: 領域の境界付近を除去

3.3.3 下腕部運動パラメータの算出

下腕部運動パラメータの算出も基本的には上腕部運動パラメータの算出法と同じだが、ただ注意すべき点は、下腕の運動は上腕の運動の影響を受けながら運動することである。つまり、下腕部運動を表すフローは上腕運動と下腕運動の合成ベクトルで表現される (図 3.10 参照)。よって、下腕部の運動パラメータを推定するためには上腕部による運動ベクトルを除去する必要がある。

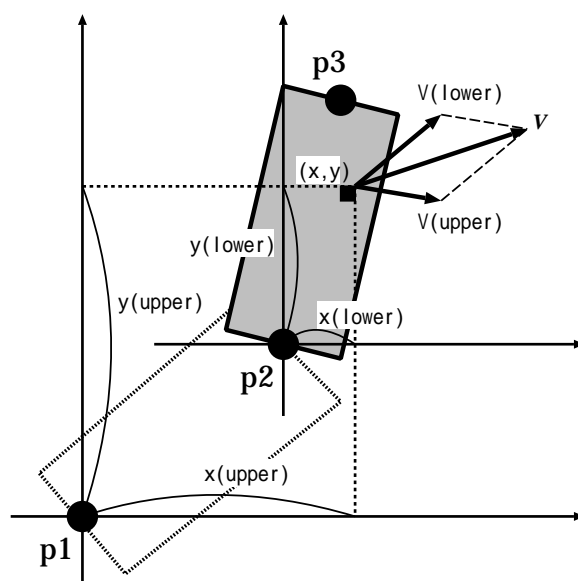


図 3.10: 上腕部、下腕部を含む local 座標系

ここで下腕部領域内の点 ($p=(x,y)$) におけるフローベクトルを V_p 、上腕部の影響による運動ベクトルを $V_p(\text{upper})$ 、下腕部の運動ベクトルを $V_p(\text{lower})$ とすると、

$$V_p(\text{lower}) = V_p - V_p(\text{upper}) \quad (3.7)$$

これを更に (u, v) 成分で記述すると、

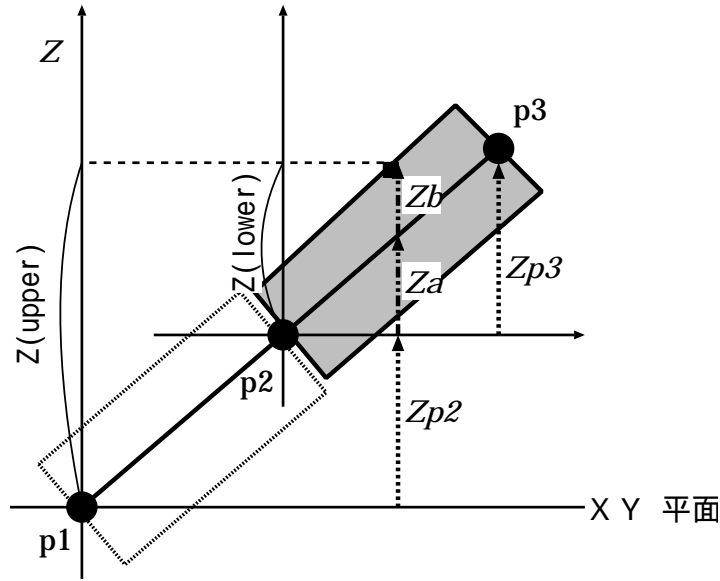


図 3.11: Z の決定 (上膊部、下膊部)

$$\begin{cases} u_{\mathbf{p}(lower)} = u_{\mathbf{p}} - (BZ_{(upper)} - Cy_{(upper)}) \\ v_{\mathbf{p}(lower)} = v_{\mathbf{p}} - (-AZ_{(upper)} + Cx_{(upper)}) \end{cases} \quad (3.8)$$

- $x_{(upper)}, y_{(upper)}$: 上膊部の回転中心座標 ($\mathbf{p1} = (p1x, p1y)$) を原点とした local 座標
- (A, B, C) : 上膊部運動パラメータ
- $Z_{(upper)}$: $\mathbf{p1}$ を原点としたときの z 成分 (図 3.11 参照)

以上により、以下の式 (3.9) に示す最小自乗問題を解くことで、下膊部の運動パラメータ (A_2, B_2, C_2) を導出することができる。

$$\sum_{\mathbf{p} \in \Omega} \left\{ (u_{\mathbf{p}(lower)} - B_2 Z_{(lower)} + C_2 y_{(lower)})^2 + (v_{\mathbf{p}(lower)} + A_2 Z_{(lower)} - C_2 x_{(lower)})^2 \right\} \quad (3.9)$$

ここで、

- $(x_{(lower)}, y_{(lower)})$: 下膊部の回転中心座標 ($\mathbf{p2} = (p2x, p2y)$) を原点とした local 座標
- Ω : 下膊部領域内 (ただし境界付近を除く)。
- $Z_{(lower)}$: $\mathbf{p2}$ を原点としたときの z 成分 (図 3.11 参照)

3.4 まとめ

本章では、オプティカルフローを利用し、物体の三次元動作を推定する手法について述べた。まず、二次元的情報であるオプティカルフローから、三次元的な運動を解析する方法を説明した。また本研究では対象物体を人体の腕と限定したので、回転運動のみによるパラメータで動作を表現できることを示した。

しかし、オプティカルフローは局所的情報かつ誤差を含みやすいので、これを直接利用するのは困難である。そこでこれらのフローを統合的、大局的に扱うため領域情報を持ったモデルを用いた。これによりモデルで作成された領域内にあるフローを全て取り出し、統合することでパーツ(上膊部、下膊部)の運動として表す。

下膊部の運動に関しては上膊部の運動の影響を受けているので、予めフローから上膊部の運動ベクトルを除去したものを使用する。

第 4 章

領域の姿勢推定・追跡処理

3 章では、オプティカルフローから物体の三次元運動を表すパラメータを推定する方法を述べた。本章ではこの運動パラメータを使用し次フレームでの領域モデルの姿勢を推定する手法を説明する。これにより対象物体の動作認識および姿勢の追跡を自動で行う。

4.1 概略

ひとつの領域に対する姿勢推定は次の手順で行う (図 4.1, 4.2)。

1. 現フレームから得られた運動パラメータをもとに移動パラメータを作成。
2. 移動パラメータに従い領域モデルを移動。
3. 移動後の領域内全ての点においてフローベクトルを復元。(これを復元フローと呼ぶ)
4. 次フレームでのオプティカルフローを計算。
5. 次フレームのオプティカルフローと復元フローとの対応を評価。
6. 2. ~ 5. の処理を探索範囲内で繰り返し、最も評価の高かったモデルの位置を次フレームでの姿勢として選択。

以上の処理は、ひとつの領域に対する姿勢推定の流れを示しているが、これを上膊部から下膊部の順で行うことにより腕としての姿勢を決定する。

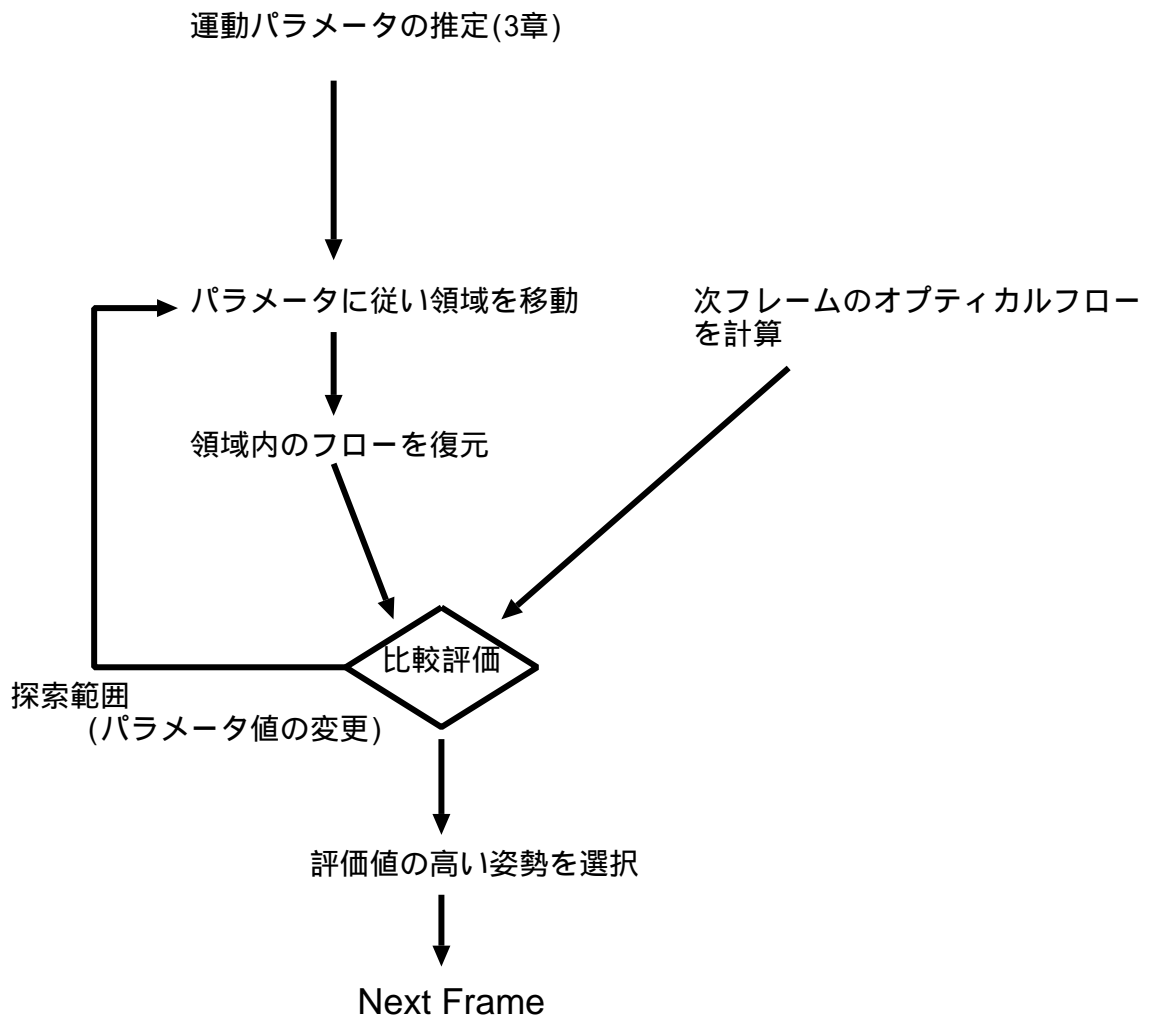


図 4.1: 姿勢推定処理の流れ

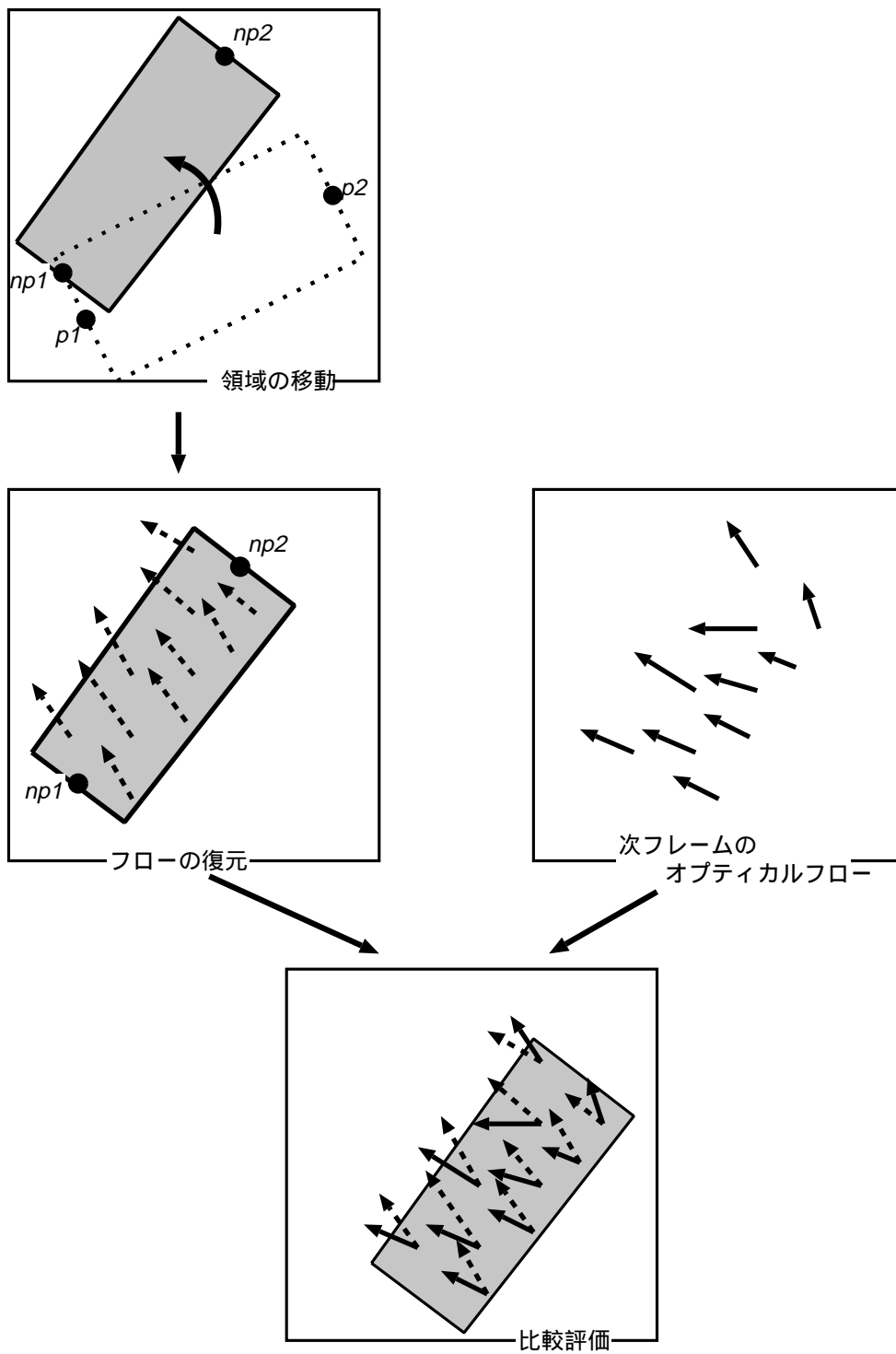


図 4.2: モデルの領域推定

4.2 探索範囲

ここで前節に出てきた探索範囲について説明する。

本手法では3章で得られた運動パラメータを基に領域モデルの位置を移動し、これを次フレームでの姿勢とする。連続するフレーム間の時間間隔が微小な場合、現フレームでの物体の運動は次フレームでも同様の性質をもつと考えられる。

しかしながら、運動パラメータは一般に誤差を含みやすいオプティカルフローを使用し推定されているので、必ずしも正確な値であるとは限らない。これをそのまま使用し領域モデルの姿勢を更新すると姿勢に狂いが生じる恐れがある。

そこで、パラメータの値を調整し姿勢の候補をいくつか作る。これを探索範囲と呼ぶ。この範囲内で評価を行うことで最適な姿勢を決定する。

4.3 比較評価

探索範囲内で最適な姿勢を決定するためには、姿勢の評価を行う必要がある。本節では姿勢推定の評価法について説明する。

探索範囲内での領域で復元したフロー場とそれに対応する次フレームでのオプティカルフロー場とを照合し評価する(図4.3)。いま、移動パラメータに従い移動したモデルがあるとする。このモデルの領域内にある点 p における復元フローを $RV_p = (Ru_p, Rv_p)$ 、次フレームでのオプティカルフローを $OV_p = (Ou_p, Ov_p)$ とすれば、領域内での2種類のフローベクトルの誤差は以下の式で求められる。

$$\sum_{p \in \Omega} \left\{ |Ou_p - Ru_p|^2 + |Ov_p - Rv_p|^2 \right\} = e^2 \quad (4.1)$$

ここで Ω とは移動先のモデルの領域である。

この計算を探索範囲内全ての領域に対して行い、最も評価が高かった(誤差が最小だった)領域が次フレームでのモデルの姿勢として決定される。

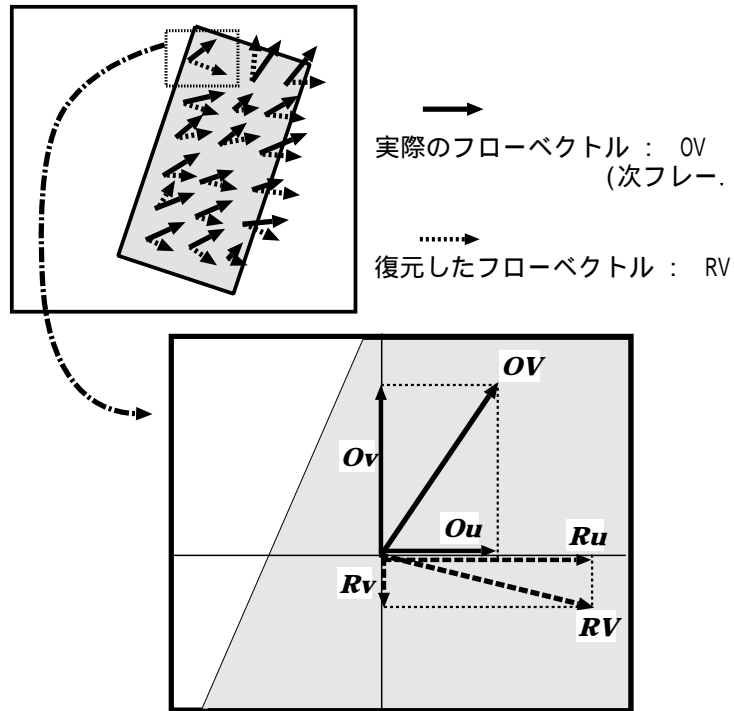


図 4.3: ベクトルの評価

4.4 上膊部姿勢推定

上膊部の運動は3章でも説明した通り、他の体節の運動に依存しない独立した運動をする。したがって、上膊部における姿勢推定も独立で行える。

3.3.2章で推定された上膊部運動パラメータを (A, B, C) とする。これを基にモデルを移動させる移動パラメータを作成する。移動パラメータ (a, b, c) は以下の範囲をとるものとする。

$$\begin{cases} 0.5A \leq a \leq 1.5A \\ 0.5B \leq b \leq 1.5B \\ 0.5C \leq c \leq 1.5C \end{cases}$$

現フレームでの領域モデルの回転中心点を $p1 = (p1x, p1y)$ 、先端点を $p2 = (p2x, p2y)$ とするとそれぞれの移動先の点 $np1, np2$ は

$$\begin{cases} np1x = p1x \\ np1y = p1y \end{cases} \quad (4.2)$$

$$\begin{cases} np2x = p2x + (bZ_{p2} - cy_{(up)}) \\ np2y = p2y + (-aZ_{p2} + cx_{(up)}) \end{cases} \quad (4.3)$$

で表される。ここで、

- $(x_{(up)}, y_{(up)})$: 点 p1 を原点とした local 座標系における点 p2 の座標成分 (図 4.4)
- Z_{p2} : 先端点 p2 の z 成分 (24ページ図 3.7参照)

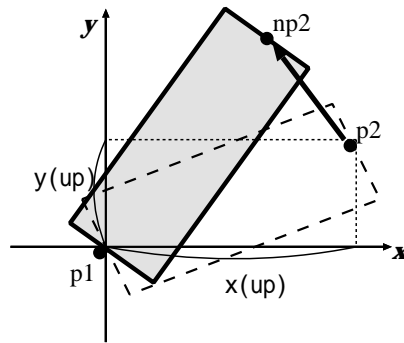


図 4.4: モデルの移動 (上膊部)

また腕は回転運動のみで動いているので、基本的には回転中心点 (p1) は移動することはない。

次に復元フローの作成について説明する。上膊部領域内の点 $p = (x, y)$ における復元フロー $RV = (Ru_p, Rv_p)$ は次式で与えられる。

$$\begin{cases} Ru_p = bZ - cy_{(upper)} \\ Rv_p = -aZ + cx_{(upper)} \end{cases} \quad (4.4)$$

ここで、

- $\{a, b, c\}$: 移動パラメータ
- Z : 点 p に対応する空間座標の z 成分 (24ページ 図 3.7参照)
- $(x_{(upper)}, y_{(upper)})$: 回転中心点 np1 を原点とする local 座標に対応する p の座標成分 (25ページ 図 3.8参照)

$$(x_{(upper)}, y_{(upper)}) = ((x - np1x), (y - np1y))$$

以上により移動した領域内全ての点において復元フローが生成される。これを使用し、式(4.1)の評価を行い、次フレームでのモデルの姿勢を決定する。

4.5 下腕部姿勢推定

次に下腕部の姿勢推定についてであるが、下腕部の運動は上腕部の運動に依存している。したがって、推定処理は下腕部運動成分に加えて上腕部の運動成分も考慮しなければならない。

ここで、3.3.3章で推定された下腕部運動パラメータを (A_2, B_2, C_2) とする。これを基にモデルを移動させる移動パラメータを作成する。下腕部移動パラメータ (a_2, b_2, c_2) は以下の範囲をとるものとする。

$$\begin{cases} 0.5A_2 \leq a_2 \leq 1.5A_2 \\ 0.5B_2 \leq b_2 \leq 1.5B_2 \\ 0.5C_2 \leq c_2 \leq 1.5C_2 \end{cases}$$

以上に加え、前節で使用した上腕部移動パラメータを (a, b, c) とする。現フレームでの領域モデルの回転中心点を $p2 = (p2x, p2y)$ 、先端点を $p3 = (p3x, p3y)$ とすると移動先の点 $np3$ は

$$\begin{cases} np3x = p3x + (b(Z_{p2} + Z_{p3}) - cy_{(up)}) + (b_2Z_{p3} - cy_{(lo)}) \\ np3y = p3y + (-a(Z_{p2} + Z_{p3}) + cx_{(up)}) + (-a_2Z_{p3} + cx_{(lo)}) \end{cases} \quad (4.5)$$

で表される。ここで、

- $(x_{(up)}, y_{(up)})$: 点 $p1$ を原点とした local 座標系における点 $p3$ の座標成分 (図 4.5)
- $(x_{(lo)}, y_{(lo)})$: 点 $p2$ を原点とした local 座標系における点 $p3$ の座標成分 (図 4.5)
- Z_{p2} : 回転中心点 $p2$ の z 成分 (27ページ図 3.11参照)
- Z_{p3} : 先端点 $p3$ の z 成分 (27ページ図 3.11参照)

ここで、下腕部回転中心点 $np2$ は上腕部先端点と結合しているので同一の点である。したがって、前節(上腕部姿勢推定)で決定された $np2$ をそのまま下腕部回転中心点とする。

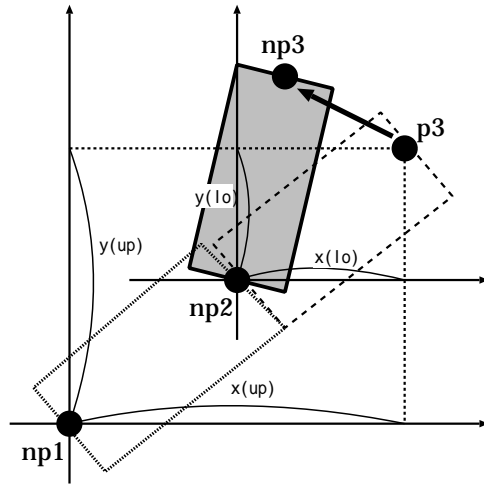


図 4.5: モデルの移動 (下腕部)

次に復元フローの作成について説明する。下腕部領域内の点 $p = (x, y)$ における復元フロー $RV = (Ru_p, Rv_p)$ は次式で与えられる。

$$\begin{cases} Ru_p = bZ_{(upper)} - cy_{(upper)} + b_2Z_{(lower)} - c_2y_{(lower)} \\ Rv_p = -aZ_{(upper)} + cx_{(upper)} - a_2Z_{(lower)} + c_2x_{(lower)} \end{cases} \quad (4.6)$$

ここで、

- $\{a, b, c\}$: 上腕部移動パラメータ
- $\{a_2, b_2, c_2\}$: 下腕部移動パラメータ
- $Z_{(upper)}$: 点 p に対応する、 $p1$ を原点としたときの z 成分 (27ページ 図 3.11参照)
- $Z_{(lower)}$: 点 p に対応する、 $p2$ を原点としたときの z 成分 (27ページ 図 3.11参照)
- $(x_{(upper)}, y_{(upper)})$: 上腕部回転中心点 $np1$ を原点とする local 座標に対応する p の座標成分 (26ページ 図 3.10参照)
- $(x_{(lower)}, y_{(lower)})$: 下腕部回転中心点 $np2$ を原点とする local 座標に対応する p の座標成分 (26ページ 図 3.10参照)

$$(x_{(upper)}, y_{(upper)}) = ((x - np1x), (y - np1y))$$

以上により移動した領域内全ての点において復元フローが生成される。これを使用し、式(4.1)の評価を行い、次フレームでのモデルの姿勢を決定する。

4.6 まとめ

本章では、運動パラメータを用いて次フレームにおける領域モデルの姿勢推定を行う手法について述べた。この手法により、対象物体(人体の腕)の姿勢推定・追跡を自動で行うことができる。

処理の手順は、

1. 運動パラメータを基に移動パラメータを作成。
2. 移動パラメータ従い、領域モデルを移動。
3. 領域内全てのフローを復元。
4. 復元フローと次フレームでのオプティカルフローとの比較評価。
5. 1.~4. までの処理を探索範囲内で繰り返し、最も評価の高いときの姿勢を選択。

というプロセスにより処理をおこなう。移動パラメータを採用する理由は運動パラメータはオプティカルフローを使用して求めるため、不正確な値を示す場合があるからである。これにより、より正確な姿勢推定が行えると考える。オプティカルフローと復元フローとの比較評価は2つのフローベクトルの最小自乗誤差を解くことで行われる。

以上の処理を上腕部から下腕部の順に行うことで、腕としての姿勢を推定する。

第 5 章

実験と考察

これまで説明した本手法を用いて、動作認識実験を行い、実験結果に対する考察を行う。今回は 2 種類の実験を行ない、それぞれの結果を示すことで本手法の有効性を検証する。

5.1 CG によるシミュレーション

ここではまず、CG によって作成されたモデルを使用し、このモデルの運動を認識する。対象となる CG モデルは円筒型をしており (図 5.1)、その運動は各軸まわりの回転運動のみから成る。モデルに与える運動情報 (ここでは角速度) は既知であるため、これと本手法で推定された運動パラメータとを比較することで、本手法における運動パラメータ推定の正確性を検証する。

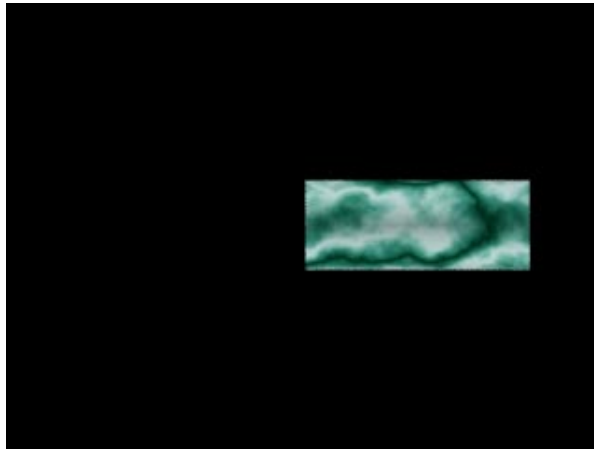


図 5.1: CG モデル (円筒型)

以下に示すのは、角軸に対する回転運動を与えたモデルの動作認識を行なった結果である。図 5.2, 図 5.3, 図 5.4はそれぞれモデルに x, y, z 軸周りの回転運動を与えた時の、運動パラメータの推定結果をグラフに示したものである。ここで、実線は対象モデルに与えた運動成分、点線は本手法におけるパラメータ推定により求めた運動パラメータを表している。また、角速度が正の値を示している時は時計周りの回転を、負の値を示している時は逆向きの回転を示す。

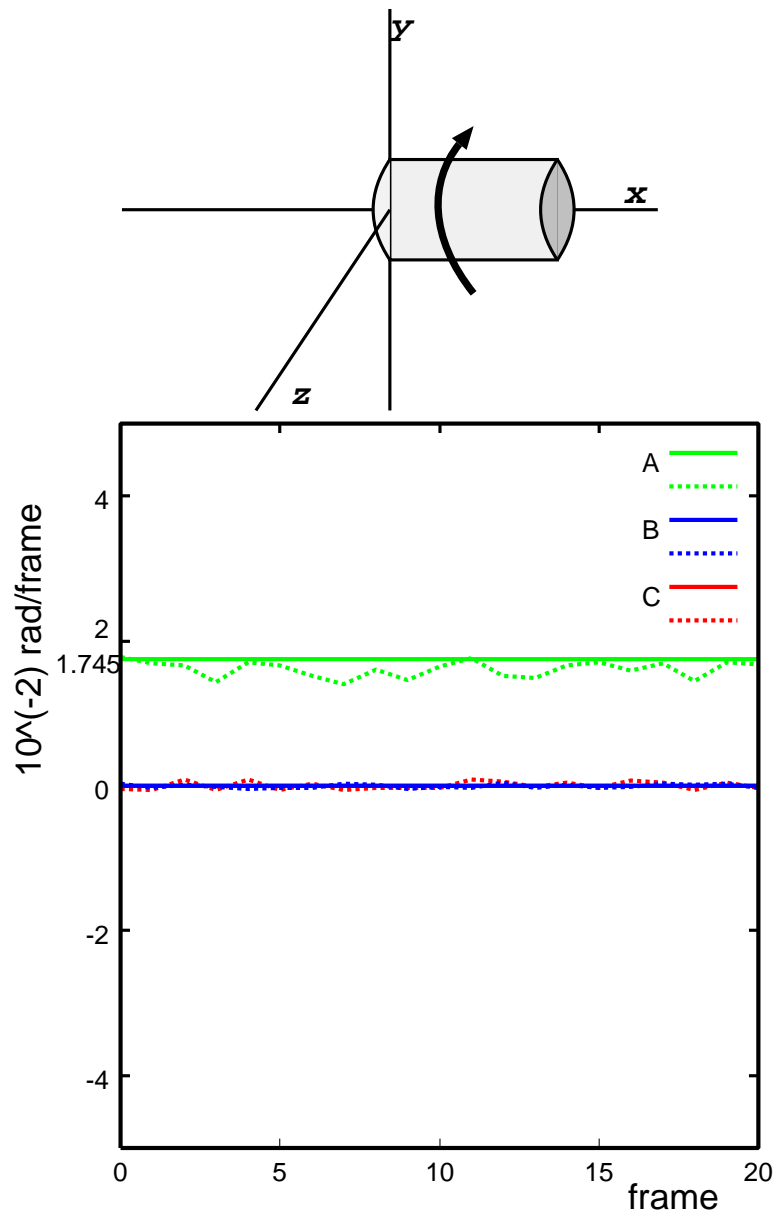


図 5.2: x 軸回転に対する運動パラメータ推定

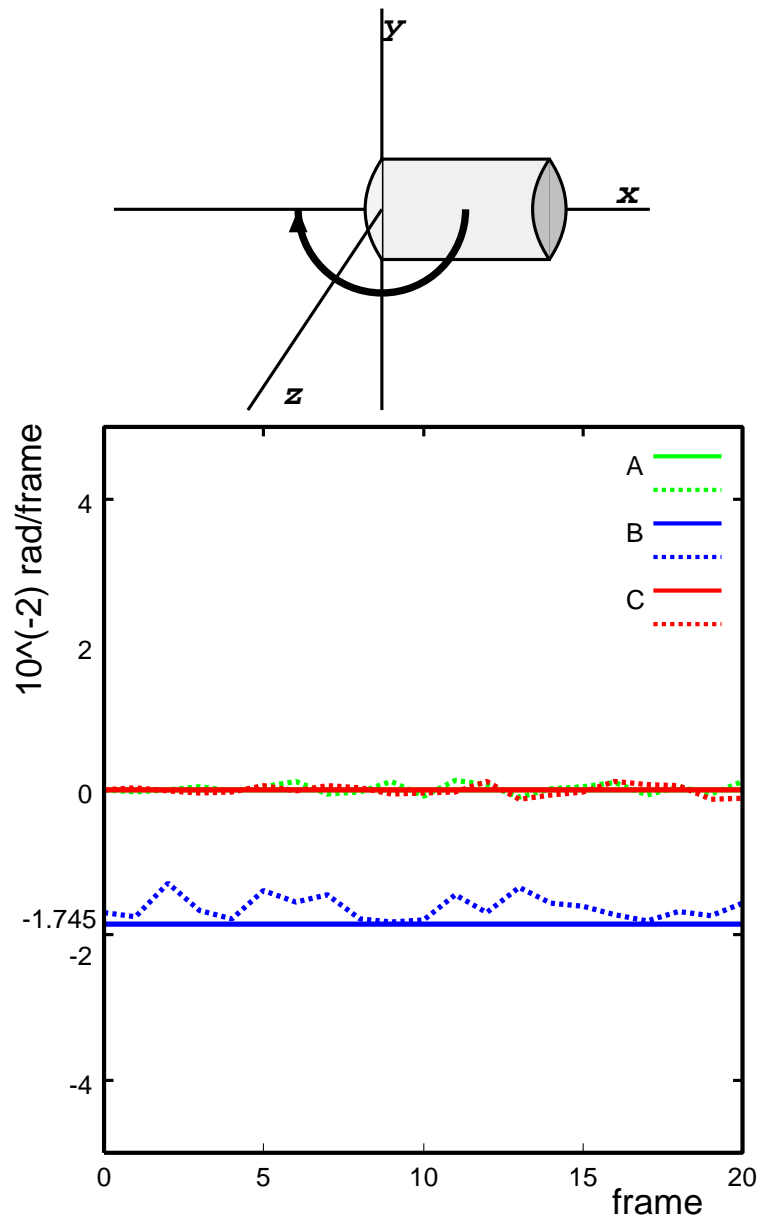


図 5.3: y 軸回転に対する運動パラメータ推定

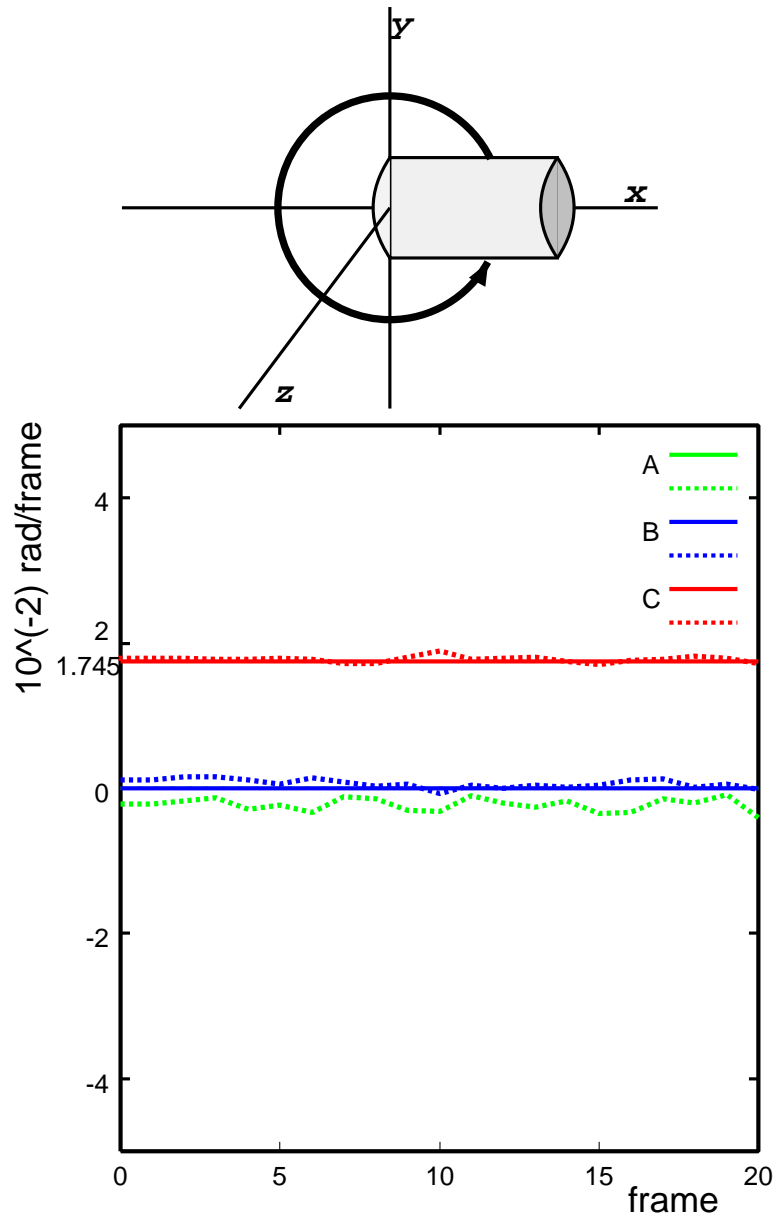


図 5.4: z 軸回転に対する運動パラメータ推定

以上の結果により、各軸まわりの回転運動に対するパラメータ推定は回転の大きさ(角速度)、および向きにほぼ正確に対応している事が示された。特に z 軸まわりの回転運動における推定精度が優れて良かった。これは対象物体に奥行き方向(カメラの光軸方向)の動きが無いために、 Z の値が一定で安定して求められたのが原因だと考える。対して x,y 軸まわりの回転運動では、 Z 成分は常に変化しており、それを画像上から簡易に求めているために正確さに欠けていたものと考えられる。別の原因としては、 x,y 軸の回転運動は画像上に投影しても見かけの動きが小さいことによる。これによりオプティカルフローが正確に抽出されず、これが直接の原因になったといえる。

今回の実験では、モデルに各軸にそれぞれ角速度 ($-5^\circ \sim 5^\circ/\text{frame}$) の範囲で回転を与えた。 x,y 軸まわりの回転に対しては角速度が変化しても、ほぼ安定した精度で推定を行っていた。 z 軸まわりの回転では $|1^\circ \sim 2^\circ|$ の範囲では高精度に推定を行っていたが、それ以上になると急激に精度が劣った。これは画素の移動量が大きくなったためにオプティカルフローが正確に抽出されなかったことが原因である。対して x,y 軸の回転運動は画像上ではそれほど大きな動きとして表れないために、フローが安定して抽出されたと考えられる。(ただし、先程述べたように動きが小さすぎて不正確になる場合もある)

5.2 実画像に対する実験

次に領域モデルの姿勢推定・追跡処理の精度を調べるために実画像を用いた実験を行なう。

5.2.1 対象物体

本手法では動作・姿勢認識の対象として、関節で連結された剛体の動作に注目する。今回の実験においては、人体の上半身をとりあげ、胴体固定のもと腕のみの動作を行うものとした。実験環境は一般的な室内で、背景は特に制限しない。ただし、被験者以外に移動物体はないものとする。被験者は画像中に一人存在し、被験者のみが運動を行う。

また、本手法では背景差分などの特別な前処理はせず、オプティカルフロー抽出に必要な処理(フィルタリング等)以外は特に施さないものとする。

5.2.2 二次元方向の運動 (z 軸まわりの回転運動)

ここでは被験者の腕の運動として、画像面に対し平行な動きを与えた。つまりこれは空間内での z 方向(カメラの光軸方向)への移動がない動作である。この一連の動作では、各体節(上膊、下膊)は関節を中心に z 軸回りの回転運動のみから成り立っていると考えられる。図 5.5, 5.6 にモデルの追跡結果を示す(追跡過程を 10 フレームおきに表示)。ここで領域モデルの枠部分を赤線、スティック部分を緑線で示す。



図 5.5: z 軸まわりの回転運動 1



図 5.6: z 軸まわりの回転運動 2

$\times 10^{-2}rad(1.745 \times 10^{-2}rad = 1^\circ)$

	A	B	C	A_2	B_2	C_2
Frame(0)	-2.810	0.097	-3.221	-4.868	-1.251	2.447
Frame(20)	-0.076	0.006	-2.686	-0.510	0.038	0.258
Frame(40)	0.160	0.051	-2.098	0.477	1.358	-2.014
Frame(60)	1.780	3.220	-1.124	-0.690	1.555	-0.815
Frame(80)	3.763	1.030	-1.363	-0.776	0.812	0.694
Frame(100)	0.845	0.208	-1.780	0.603	-0.131	0.363
Frame(120)	0.516	-0.853	-1.454	0.279	-0.202	0.035
Frame(140)	0.285	-2.250	-0.947	-0.263	-0.303	-1.206

表 5.1: 運動パラメータ

また、運動パラメータの推定結果を表 5.1 に示す。ここで上膊部運動パラメータを (A, B, C) 、下膊部運動パラメータを (A_2, B_2, C_2) で表示する。これらはそれぞれ各軸 (x, y, z) に対する回転角速度を表している。また負の数値を表している時は軸方向に対して時計まわりの回転を示し、正の時は逆方向の回転を示している。

ここで、実験結果に対する考察を述べる。まず、図 5.5, 5.6 の追跡結果から視覚的にはあるが、ほぼ正確に追跡していることがわかった。特に上膊部の運動に対しては、他の体節の運動に依存しないので独立で測定でき、結果として高精度に追跡が行えた。下膊部に対しては、上膊の運動に依存して運動しているので、解析が多少困難である。故に追跡結果は上膊部よりも精度が劣っている。しかし、腕全体の動きとして考えれば、充分良好な結果を示しているといえる。

次に表 5.1 に示した運動パラメータについて考える。この実験における被験者の腕の動作は z 軸まわりの回転運動のみから成り立っているため、パラメータ C にその値がよく現れている。ここで、 -1.00 未満の数値は許容誤差として判断しても問題ない。注目するのは、60 フレームから 80 フレームにかけて A, B の値が大きく現れているところである。この付近は腕が胴体と重なりはじめた箇所である (図 5.5 下段)。そのため、腕が z 軸方向 (前方) に移動したために A, B に現れたと考える。

以上より、 z 軸まわりの回転に対しては、良好に運動推定および追跡処理を行っている

ことが示された。

5.2.3 奥行き方向の運動 (x,y 軸まわりの回転運動)

次に被験者の腕の運動として、 Z の変化量がある動きを与えた。つまりこれは空間内での z 方向 (カメラの光軸方向) への移動を含む運動である。この一連の動作では、各体節 (上膊、下膊) は関節を中心に x,y 軸回りの回転運動から成り立っている。

図 5.7 にモデルの追跡結果を示す (追跡過程を 10 フレームおきに表示)。



図 5.7: x,y 軸まわりの回転運動

$\times 10^{-2}rad(1.745 \times 10^{-2}rad = 1^\circ)$

	A	B	C	A_2	B_2	C_2
Frame(0)	0.025	1.032	0.294	0.996	-1.201	0.170
Frame(10)	0.469	1.744	0.296	0.643	-0.413	0.199
Frame(20)	0.539	1.225	-0.288	0.037	0.184	0.717
Frame(30)	0.325	1.150	-0.320	0.053	0.316	0.547
Frame(40)	-0.224	0.513	0.255	0.363	1.606	-1.319
Frame(50)	-0.094	1.017	-0.972	0.214	0.722	0.480
Frame(60)	-0.251	0.846	-0.807	0.086	0.914	0.116

表 5.2: 運動パラメータ

また、運動パラメータの推定結果を表 5.2 に示す。

以上の結果を基に考察する。まず、対象の動作を x,y 軸まわりの回転運動に限定し、腕が z 軸方向に移動するような動きに対する追跡処理を行ない本手法の有効性を検証した。図 5.7 に示される通り、追跡結果はほぼ良好であった。この図における腕の動作は、特に y 軸まわりの回転運動のみで成り立っているが、表 5.2 に示されるように、運動パラメータは B に大きく現れている。よって、腕が z 軸方向に向かう動作を行っても充分動作推定は行えているといえる。

ただ上記の結果には示されていないが、長時間追跡処理を行う過程においては推定誤差の蓄積による追跡ずれが生じる問題が発生する。これは先ほどの z 軸まわりの回転運動にも同様のことがいえるが、特に x,y 軸まわりの回転運動に対する追跡処理の方が、追跡ずれが大きく現れた。また、運動パラメータの推定においてもモデルの微妙な位置ずれに敏感で、正確にモデルを位置付けなければ、パラメータの値に大きく誤差が生じるという結果も示された。

ここで、 x,y 軸まわりの運動が z 軸まわりの運動に対して精度が劣る理由を考える。まず 1 つは、 x,y 軸まわりの運動は画像上に投影してもあまり大きな動きとして現れないことにある。これにより、領域モデルの移動範囲が縮小されてしまい、十分な追跡が行えない。もう 1 つは、 Z 成分の推測法にある。本手法では、物体の Z 成分を画像から簡易に推

測しているため、必ずしも正確な値とは限らない。今後、この Z 成分を正確に求める方法を確立する必要があると考える。

また、第 70 フレーム以降は物体のさらなる z 軸方向への移動によりモデルの領域が完全に遮蔽してしまい追跡不可能となった。

5.3 考察とまとめ

本章では、本手法の有効性を示すため、2 種類の認識実験を行なった。まず 1 つめは、CG を用いたシミュレーション実験を行ない、運動パラメータの推定に関して注目して調べた。結果、ほぼ正確にパラメータ推定が行なわれていることが示された。

2 つめの実験は、実画像を用い、領域モデルの追跡処理に注目して実験を行なった。対象物体は人体の上半身で、胴体固定のもと腕のみの動作に限定した。腕の動きにも 2 種類用意し、1 つは Z の変化がない z 軸まわりの回転運動、もう 1 つは Z の変化がある x, y 軸まわりの回転運動を行なった。

結果は、 z 軸まわりの回転運動に対しては、視覚的にはあるが良好に追跡していた。 x, y 軸まわりの回転運動に対しては、 z 軸まわりの運動ほどではないがまず良好に追跡していた。

問題点は、画像のノイズによる影響のため運動パラメータ自体に多少の誤差が含まれることである。これにより処理を長時間行うと、推定誤差が蓄積されるため、モデルの姿勢にずれが生じる。本手法ではこの追跡ずれに対する処理を行わず、途中手動で修正しなおすことで補ったが、今後は追跡ずれに対する処理に関しても検討する必要があると考える。

また、 z 軸方向への移動によりモデルの領域が遮蔽した場合、以降の処理が不可能になる。領域が遮蔽した場合の追跡処理に関しても検討する必要がある。

第 6 章

結論

人体のような多関節物体の動作や姿勢を、画像を用いて認識、記述しようとする試みは、機械の制御、コミュニケーションの手段などの目的でさかんに行われている。しかし、人体は複雑な形状をしておりかつ非剛体な変形を伴うため、その動作・姿勢の画像からの解析には困難を要する。

そこで、本研究では人間との親和性および処理の軽減を考慮し、従来の問題点を解決するための新しい手法を提案した。

まず、同一物体上では同様の性質を持つオプティカルフローを活用し、この局所的な情報を領域を持ったモデルを用いて統合することによって大局的な信頼性のある情報とした。また、人物領域を体節ごとの領域に分割することで、各体節からの情報を得、さらにそれを統合することで人物全体の姿勢を追跡する方法を行った。

また、本手法ではオプティカルフローから三次元運動を解析し、領域単位で統合することで、人体の三次元動作および姿勢を推定した。

以上の本手法による有効性を確認するために実験を行った。実験結果は、視覚的な評価しか行っていないが、十分満足する結果が得られた。

ただし問題点も明らかになった。本手法は空間内での姿勢を推定する際、 z 成分 (奥行き成分) を簡単な推測により求めたが、値の正確性が充分でなかったことや認識可能な姿勢が限定されてしまった。今後は、より正確かつ幅広い姿勢に対応できるような手法を確立することが望まれる。また、 z 方向の運動により領域が完全に遮蔽した場合の追跡法も考慮しなければならない。さらに、長時間追跡可能にするには、推定誤差による追跡ずれを修正する方法についても考慮する必要がある。

また、今回は対象を人体の腕に限定し、回転運動のみの解析を行うことで動作および姿勢の認識を行った。今後は更に幅広い対象に対応させるためにも、並進運動を含む三次元運動の解析法を確立する必要がある。

謝辞

本研究を行うにあたり、御指導御鞭撻を戴いた北陸先端科学技術大学院大学阿部 亨 助教授に深く感謝致します。

また、さまざまな面で御教授戴きました、堀口 進 教授に心から感謝致します。
小谷 一孔 助教授には修士論文審査に関して、御指導戴き、誠に感謝致します。
木村 正行 教授には、サブテーマで熱心に御指導を戴き、厚く御礼申し上げます。
また、日頃よりお世話になった研究室の皆様に心より感謝致します。

参考文献

- [1] 亀田能成, 美濃導彦, 池田克夫, シルエット画像からの関節物体の姿勢推定法, 信学論 (D-II), vol.J79-D-II, no.1, pp. 26-35, 1993.
- [2] Barron,Fleet,Beauchemin, Performance of Optical Flow Techniques, PhD Dissertation, Dept. of Computer Science, Carnegie-Mellon Univ. ,1984.
- [3] 島田伸敬, 白井良明, 久野義徳, 確率に基づく探索と照合を用いた画像からの手指の3次元姿勢推定, 電子情報通信学会論文誌, D-II Vol. J79-D-II No. 7 pp.1210–1217, Jul 1996.
- [4] 倉掛政治, ラマカント ネヴァティア, 関節のある運動物体の記述と追跡, 信学論 (D-II), vol.J76-D-II, no.4, pp. 854-862, 1993.
- [5] 木本伊彦, 梶谷昭彦, 安田靖彦, スティックモデルに基づく単眼視動画像からの人体歩行運動の解析の一手法, 信学論 (D-II), vol.J74-D-II, no.3, pp. 376-387, 1991.
- [6] 安藤栄治, 動画像の領域分割に基づく人体の姿勢推定と追跡に関する研究, 北陸先端科学技術大学院大学修士論文, 1997.
- [7] 藤富知行, 大西昇, 杉江昇, 可変形状物体の姿勢推定についての検討, 通学技法, pp. 17-24, 1994.
- [8] 鈴木貴彦, 吉野和芳, 川嶋稔夫, 青木由直, 特異画像の解析を基にした周期的動作の記述, コンピュータビジョン, pp. 43-50, 1995.
- [9] 木村和広, 太田直哉, 金谷健一, 精密なノイズモデルによるオプティカルフローの検出, コンピュータビジョン, pp. 37-42, 1996.

- [10] Haibo Li, Pertti Roivainen, Robert Forchheimer, 3-D Motion Estimation in Model-Based Facial Image Coding, IEEE T-PAMI, pp. 545-555, 1993.
- [11] 谷内田正彦, ロボットビジョン, 東京 昭晃堂, 1990.
- [12] 嶋中徹, オプティカルフローの正確な推定に関する研究, 北陸先端科学技術大学院大学修士論文, 1996.
- [13] 戸島靖, 手指の動作認識に関する研究, 北陸先端科学技術大学院大学修士論文, 1994.
- [14] 永井克明, スティックモデルを用いた人物動画像の認識と追跡, 北陸先端科学技術大学院大学修士論文, 1995.
- [15] 尾崎弘, 谷口慶治, 画像処理 –その基礎から応用まで, 2nd ed, 東京, 共立出版, 1988.
- [16] 山本正信, 人間動作の動画像解析, 第2回 パターン計測シンポジウム, 1997.
- [17] 浅田稔, ダイナミックシーンの理解, 東京, 電子情報通信学会, 1994.
- [18] 高橋勝彦, 関進, 小島浩, 関 隆一, ジェスチャー動画像のスポットティング認識, 電子情報通信学会論文誌 (D-II), J77-D-II, No.8, pp. 1552-1561, 1994.
- [19] 大和淳司, 大谷淳, 石井健一郎, 隠れマルコフモデルを用いた動画像からの人物の行動認識, 電子情報通信学会論文誌 (D-II), J76-D-II, No.12, pp. 2556-2563, 1993.
- [20] クンラポン ユーニパン, 木下宏揚, 酒井善則, スティックモデルを用いた手振りの認識, 電子情報通信学会論文誌 (D-II), J77-D-II, No.1, pp. 51-61, 1994.
- [21] 八木伸行, 井上誠喜, 林正樹, C 言語で学ぶ実践画像処理, 東京, オーム社, 1992.
- [22] Sourabh A. Niyogi, Edward H. Adelson, Analyzing and Recognizing Walking Figures in XYT, IEEE Computer Vision and Pattern Recognition, IEEE COMPUTER SOCIETY PRESS, pp. 469-474, 1994.

研究業績

- [1] 動画像を用いた多関節物体の三次元動作認識,
浅野 英城, 阿部 亨, 堀口 進
電気関係学会北陸支部連合大会 F-51, pp360, 1997