

Title	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
Author(s)	安藤, 栄治
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1003
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士

修士論文

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

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

安藤栄治

1997年2月14日

目次

1	序論	1
1.1	研究の背景と目的	1
1.2	本論文の構成	2
2	人物動作認識	3
2.1	はじめに	3
2.2	画像処理による手法	4
2.3	まとめ	5
3	オプティカルフローの推定	7
3.1	はじめに	7
3.2	オプティカルフロー	8
3.2.1	オプティカルフローの拘束方程式	8
3.2.2	Lucas と Kanade の拘束条件	10
3.3	まとめ	11
4	人体動作の認識・追跡処理	12
4.1	はじめに	12
4.2	座標系	13
4.3	認識対象	13
4.4	処理の流れ	14
4.5	人体モデル	16
4.6	認識処理	19
4.6.1	人物領域の抽出	19

4.6.2	速度パラメータの推定	23
4.6.3	運動領域の推定・追跡処理	33
4.7	まとめ	38
5	実験と考察	39
5.1	はじめに	39
5.2	入力画像	39
5.3	体節の重なりがない場合	40
5.4	体節の重なりが生じる場合	40
5.5	実験結果の考察	41
5.6	まとめ	42
6	結論	47

目 次

3.1	移動ベクトルと速度ベクトルの違い	8
3.2	速度空間	10
4.1	対象物体と座標軸	13
4.2	オプティカルフローの計算	14
4.3	処理の流れ	15
4.4	人体モデル	18
4.5	人体モデルの姿勢	18
4.6	画像上の点のベクトル表現	20
4.7	入力画像	22
4.8	出力画像	22
4.9	モデルの初期設定	23
4.10	各関節 (回転中心における)local 座標軸	24
4.11		25
4.12	フローの分割	26
4.13	速度パラメータ	28
4.14	回転中心の誤差	29
4.15	投票分布	31
4.16	上腕部による速度ベクトル	32
4.17	モデルの領域推定	34
4.18	角誤差	36
5.1	屈伸運動 1	43
5.2	屈伸運動 2	44

5.3	回転運動	45
5.4	領域の重なりが生じた場合	46

第 1 章

序論

1.1 研究の背景と目的

人体の動きや姿勢を検出する技術は、高度なマンマシンインターフェースの実現において重要な課題である。これらの技術は、ジェスチャ認識、スポーツ選手の運動解析としての応用がある。近年これらの技術に関する研究が盛んに行われているが、それらの多くは対象にマーキングを施し特徴点を明示したり、データスーツなどのセンサーから情報を得るといった特別な条件を必要としている [1]。これに対し、人間と計算機の親和性を重視する観点から、特別な条件を必要としない画像処理のみによる研究も試みられている。亀田らは、精密な人体のモデルを用いて 1 枚のシルエット画像と人体モデルの被覆率を計算し人体の姿勢推定を行っている [2]。倉掛らは明確な人体モデルをもたず、画像中のエッジからリボンを抽出し、そのリボンの接続関係を考察することによって平面内での関節物体の動作を認識している [5]。また木本らは、人体をスティックモデルで表現し、対象のシルエット画像の骨格線を抽出することによってモデルとの対応づけを行い、姿勢の解析を行っている [6]。しかし従来から提案されている手法では、胴体や腕などの領域の重なりに対して個々の領域の検出が困難であるため、認識できる動作が制限されていたり、重なった領域の位置推定が曖昧であるなどの問題点がある。

こういった問題に対して、本研究では、複雑な動作で領域の重なりが生じる人体の上半身を取り上げ、その動作および姿勢を動画像から認識する手法について新しい方法で検討を行う。その際、人体の重なりを正確に認識するためには、重なった領域を個々の体節に正確に領域分割することが重要であり、そのためには、従来までの線情報や点情報あるい

はシルエット情報を基にした追跡ではなく、個々の体節領域ごとに追跡を行うことが必要である。そこで本研究では、領域情報を用いることにより、領域単位で追跡を行う方法を提案する。領域情報においては、重なった領域に対して同一物体上では、同じ性質をもつオプティカルフローを活用する。

従来、オプティカルフローを移動ベクトルと考え、モデルの移動を試みてきたが、オプティカルフローは速度場を中心とした解析手法のため、本研究では速度ベクトルとして解析する。オプティカルフローは、局所領域で計算されるため誤差も多く信頼性が低いが、領域情報を持った人体モデルを用いてモデル内の個々のオプティカルフロー情報を全体で一つの情報と考えることによって、局所情報を信頼性のある大局的な情報として活用する。オプティカルフロー情報および、人体モデルを有効に活用することによって、マーキングなどの特殊な準備を施さずに撮影した動画像から、重なった領域に対しても正確に人体の姿勢推定、追跡処理を行うことを目的とする。

1.2 本論文の構成

本論文の構成は以下の通りである。

第2章では、画像処理のみによる人物動作認識の代表的な手法について簡単に紹介する。

第3章では、オプティカルフローの計算方法と適用方法について述べる。

第4章では、人体の姿勢推定・追跡処理について述べる。

第5章では、実際の人物動作を撮影した動画像を用いて本手法による有効性を視覚的に検討し、その考察について述べる。

第6章では、本研究の結果をまとめるとともに、今後の課題について述べる。

第 2 章

人物動作認識

2.1 はじめに

近年、コンピュータアニメーションやジャスチャ認識などを目的として、人間の姿勢や動作を認識する研究が盛んに行われている。この問題に対するアプローチは大きく 3 つに分割することができる。

1 つは、被験者にデータグローブや磁気センサなどのデータ入力装置を取り付け、直接その位置や動きを測定するものである [7]。この方法は、認識に要する位置や角度が容易かつ正確に得られオクルージョン問題もないため非常に有効な結果が得られる。しかしながら測定機器の装着の手間や装着物による拘束のため、自然な動きができない。また、センサ類はセンシティブであるため取扱いに注意を要するといった欠点がある。

2 つめは、被験者にマーカーを取り付け、その位置を画像処理により求める方法である。この方法は、被験者にマーカーを施すだけであるので、被験者に対する拘束は無くなり、しかも比較的容易に特徴点の検出ができる。しかしながらマーカーの取り付けに時間を費し、画像中にマーカーとの同色成分が無いことが前提となるため、撮影できる環境条件は厳しいものとなる。

3 つめは、背景や被験者にはできる限り条件をつけずに、画像処理のみにより認識する手法である。この方法は、被験者にはセンサやマーカーの取り付けを必要としないため、人間との親和性を考えた場合優れた手法であるといえる。また、機器の取り付けなどが必要としないため、数多くの被験者がいる場合にもデータだけを先にとり、後で処理することも可能であり被験者に対しての時間の拘束もない。しかしながら、人間を認識する際には

剛体の認識とは異なり対象の動きや変形など抽出すべき要素が多く、それらの認識も困難であるため、必ずしも良い結果は得られない。

以下、特に画像処理を用いた認識に注目し従来手法について述べる。

2.2 画像処理による手法

ここでは、被験者に特殊な準備を施さず、画像処理のみによる従来認識手法をいくつか紹介する。

精密なモデルを用いた姿勢推定

亀田ら [2] 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。しかし重なりが生じた体節の正確な推定が行えない。人物領域の輪郭が正確に抽出されているという特殊な条件を必要としている。また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。

1枚の画像からの姿勢推定

藤富ら [8] 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体積表現である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。しかしながら、認識に用いられる画像が1枚しかないために、背景を制限し、対象に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2次元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。

リボン抽出による認識

倉掛ら [5] 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信頼性高く関節部位を検出し、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2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。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。この手法の利点は、複数フレームにまたがりリボンマッチングすることにより、各フレームで同じ位置にあるエッジ成分は背景として削除できるため、背景が既知である必要がなく、同時にフレーム間で情報を補いあうために、信頼性が増す点にある。しかしながら、手先などの閾値以下の短いパーツや、複雑に各部位が接近している場合にはリボンをうまく抽出できないといった欠点がある。

スティックモデルによる足の認識

木本ら [6] は、対象を歩行運動に限定し、画像上での2次元的な足の動きを3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。

以上のように画像処理による認識法は、装着物等による被験者に対する拘束はなくなるが、それによる特徴抽出の難しさや動きの複雑さの面からより高度な画像処理技術が要求される。しかしながら、現状では「認識可能な動作が限定されている」、「人体の重なりには対応できない」など、さまざまな問題を抱えており、良好な認識手法が確立されていない。

2.3 まとめ

これまで行われてきた人物を対象とした認識は、データグローブや磁気センサなどの接触型センサを装着するものと、認識対象に装着物をつけずに、自由な動作を認識するため

の画像処理によるものの2つに大別される。接触型センサを用いるものは、速応性や精度の面で非常に優れた認識手法であるが、被験者に対する負担が重く、汎用性に欠け、一般的ではない。

このような欠点を克服するために、近年、画像処理のみによる認識する手法の提案が盛んに行われている。この方法は、人間との親和性、汎用性に優れているが

- 抽出できる特徴点が少ないため、認識にはより高度な画像処理技術が要求される。
- 処理が複雑となるためリアルタイム性にかける。
- モデルマッチングを行うものは、多大な計算機資源と計算量を要する。

といった一般的に考えられる問題点がある。これら問題点において、従来手法では画像情報だけではなく、対象に関する知識としてモデルを用いることにより画像からの情報とモデル情報との照合を行うことで問題の解決を試みている。従来手法においては、領域の細線化による特徴点や、エッジ情報、精密なモデルの投影を用いるシルエット情報などを用いて照合しているため情報量が少なく重なった領域に対する有効な情報が得られないため、

- 認識可能な動作が制限されている。
- 人体の体節の重なりに対して認識が曖昧である。

など重大な問題が残されている。そこで本研究では、重なった領域に対しても正確な姿勢推定を行うためには領域単位での追跡およびモデルの対応づけが必要であると考え、領域情報を持った人体モデルを用いる。重なった領域に対して運動情報、領域情報が得られるオプティカルフローを用いて、領域情報、人体モデルとともに活用することによってこれらの問題点を解決することを試みる。

第 3 章

オプティカルフローの推定

人体などを対象とした画像認識を行うために、人体のシルエットなどの情報を用いる手法や人体領域を細線化して特徴点を得る手法があるが、これでは人体の重なった領域の推定が正確には行えない。そこで重なった領域において個々の領域を分割する何らかの情報が必要となる。

本研究では、重なった領域に対しても正確に姿勢推定および追跡を行うため、同一物体上では同じ性質をもつオプティカルフローを活用する。オプティカルフローは、動画像において連続するフレーム間の対応する各点を対応づけることにより各点の運動ベクトルを推定する手法である。また、フローベクトルは、局所的、平面的ではあるが運動情報（各点における運動ベクトル）、領域情報（同一性質をもつフローベクトルは同一物体上の点である）として有効である。

3.1 はじめに

動画像からの速度の検出に関する研究は 1970 年ころより行われている。多くの論文 [3] では、動画像中の濃淡パターンの対応づけの考え方をもとに速度場（画像中の各点に速度ベクトルを対応させたもの）を計算する手法が一般的である。こうした対応づけの考え方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ものが、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検出法は、マッチング法、グラディエント法などに大別される。

連続する画像フレーム間において濃淡パターンの対応づけを実行し、得られる変位ベク

トルからその地点での動きベクトルを決定する手法がマッチング法とよばれる。通常、適当なサイズのテンプレートを用い、空間相互相関などの解析法により対応づけを行う。

本研究ではグラディエント法を用いる。これは、濃淡パターンの特徴を表す画像関数(明度関数)が運動に際し不変に保たれるとの仮定より、ある点 (x, y) における動きベクトルの速度と、動画像の濃淡分布の空間勾配および時間勾配とを関係づける式(拘束方程式)をもとに速度ベクトルを検出する解析手法である。

しかしグラディエント法を用いる上で考慮しなければならないのは、本研究のように回転運動のパラメータを推定する場合、マッチング法は、直接的な対応づけにより移動ベクトルを得るものであるが、グラディエント法は、フレーム間隔が非常に短い場合を想定しており、速度場での解析が中心となっている。図 3.1 に示すように、回転運動においては移動ベクトルと速度ベクトルは異なるということである。

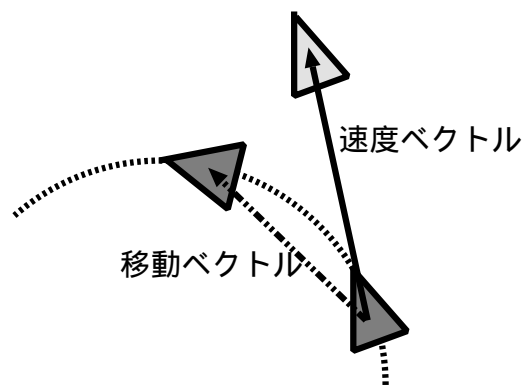


図 3.1: 移動ベクトルと速度ベクトルの違い

3.2 オプティカルフロー

3.2.1 オプティカルフローの拘束方程式

動画像中の時刻 t における点の画素位置 (x, y) の明度を $E(x, y, t)$ と表す。微小時間 δt 後の時刻 $t + \delta t$ の点 $(x + \delta x, y + \delta y)$ における明るさは不変と仮定すると次式が成り立つ。

$$E(x, y, t) = E(x + \delta x, y + \delta y, t + \delta t) \quad (3.1)$$

上式の右辺を Taylor 級数展開すると,

$$E(x, y, t) = E(x, y, t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} + e \quad (3.2)$$

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,

$$\frac{\delta x}{\delta t} \frac{\partial E}{\partial x} + \frac{\delta y}{\delta t} \frac{\partial E}{\partial y} + \frac{\partial E}{\partial t} = 0 \quad (3.3)$$

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

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \quad (3.4)$$

ここで, 見かけの速度ベクトルを $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, 空間的な明度勾配を $E_x = \frac{\partial E}{\partial x}$, $E_y = \frac{\partial E}{\partial y}$, 時間的な明度勾配を $E_t = \frac{\partial E}{\partial t}$ とすれば, 次のように書き直せる.

$$E_x u + E_y v + E_t = 0 \quad (3.5)$$

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. u, v を座標軸とする 2 次元空間 (速度空間)(図 3.2) 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拘束直線上にある.

このような局所的計測では拘束直線を決定することのみが可能である. また, 拘束方程式は以下の形式で書かれることもある.

$$\Delta I(\mathbf{x}, t) \cdot \mathbf{v} + I_t(\mathbf{x}, t) = 0 \quad (3.6)$$

ここで, $\mathbf{x} = (x, y)$, $I(\mathbf{x}, t) = E(x, y, t)$, $\mathbf{v} = (u, v)^T$, $I_t(\mathbf{x}, t) = \frac{\partial I(\mathbf{x}, t)}{\partial t}$, $\Delta I(\mathbf{x}, t) = (I_x(\mathbf{x}, t), I_y(\mathbf{x}, t))$

実際にフローベクトルを求めようとすれば, 式 (3.6) だけでは一意に決定できない. そこでもうひとつの拘束条件として Lucas, Kanade の拘束条件を導入する.

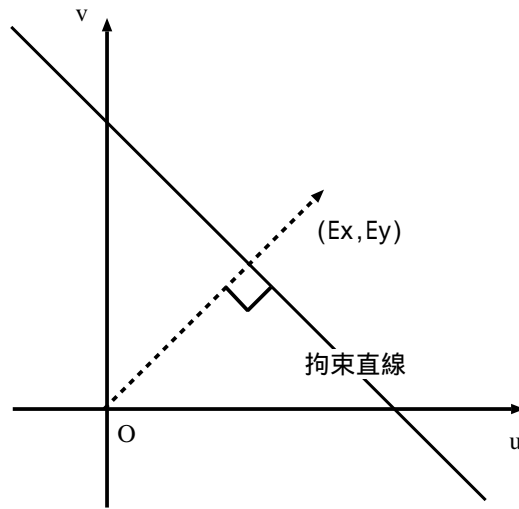


図 3.2: 速度空間

3.2.2 Lucas と Kanade の拘束条件

本研究では, Barron らが行った比較研究 [3] において最良の結果を得た Lucas と Kanade の拘束条件を用いる.

Lucas と 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, 局所的滑らかさの拘束を用いた. つまり, 各局所領域 Ω 中のオプティカルフロー速度 \mathbf{v} の拘束として, オプティカルフロー拘束方程式 (3.6) の加重二乗和

$$\sum_{\mathbf{x} \in \Omega} W^2(\mathbf{x}) [\nabla I(\mathbf{x}, t) \cdot \mathbf{v} + I_t(\mathbf{x}, t)]^2 \quad (3.7)$$

の最小化を考えた. ここで $W(\mathbf{x})$ は近傍の周囲よりも中心における拘束に重きをおくための窓関数である. 式 (3.7) の解は,

$$A^T W^2 A \mathbf{v} = A^T W^2 \mathbf{b} \quad (3.8)$$

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

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

式 (3.8) の解は 2×2 行列 $A^T W^2 A$ が正則なとき求まり,

$$\mathbf{v} = [A^T W^2 A]^{-1} A^T W^2 \mathbf{b}$$

ここで,

$$A^T W^2 A = \begin{bmatrix} \sum_{\mathbf{x} \in \Omega} W^2(\mathbf{x}) I_x^2(\mathbf{x}) & \sum_{\mathbf{x} \in \Omega} W^2(\mathbf{x}) I_x^2(\mathbf{x}) I_y^2(\mathbf{x}) \\ \sum_{\mathbf{x} \in \Omega} W^2(\mathbf{x}) I_y^2(\mathbf{x}) I_x^2(\mathbf{x}) & \sum_{\mathbf{x} \in \Omega} W^2(\mathbf{x}) I_y^2(\mathbf{x}) \end{bmatrix} \quad (3.9)$$

3.3 まとめ

オプティカルフローをの抽出法は、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、連続動画像においてフレーム間での直接的な対応づけにより移動ベクトルを抽出する手法である。後者は、時空間における勾配の拘束を用いた解析法であり移動ベクトルというより、速度場の解析が中心となっている。本研究では、グラディエント法によるオプティカルフローの抽出を行い、フローベクトルを速度ベクトルとして活用することによって、速度場を用いた人体動作の認識法について検討する。

第 4 章

人体動作の認識・追跡処理

4.1 はじめに

本章では、あらかじめ用意された背景画像と、その背景の前で動作する人物の動画像系列から、人物の各部位の 2 次元的位置とその動きを追跡するための処理過程について述べる。

本研究では、認識処理において、領域情報をもつ人体モデルを用いる。局所速度情報であるオプティカルフローは、各々 1 点だけでは、誤差も多く信頼性が低い。各体節を各々剛体と考えることで領域内のオプティカルフローは、全体として一貫した運動を示すものであると考える。このように、局所速度情報としてのオプティカルフローを大局的な体節の速度情報とすることによって正確な各体節の速度パラメータを推定する。本研究で用いるオプティカルフローは、3 章で示したように速度場を中心とした解析により速度ベクトルを推定している。それゆえ、フローベクトルを各点での移動ベクトルとしてボトムアップ的なモデルの移動は行わない。人物領域をフレームごとに正確に領域分割することにより、順次、フローベクトルをモデル情報に照らし合わせて各体節の速度パラメータを推定する。また本研究では、微小時間における解析を行うため、次フレームでの人体の運動は、現フレームでの運動とほぼ同じであると仮定する。次フレームでのモデル位置を探索範囲内で少しずつずらし、現フレームでの速度パラメータに基づき各モデル位置における速度場候補を生成する。生成した各速度場候補と実際に画像から計算された速度場（以下、明確な区別をつけるためフロー場と呼ぶ）との照合により領域単位でのモデルの対応づけを行う。モデルの対応づけは、上腕部、下腕部の順にモデル位置および領域を推定する。以下

に処理の詳細を述べる.

4.2 座標系

画像上の位置を示す座標系として以下のように設定した. 画像上での座標軸は, 対象物体の画像面への正射影を仮定することによって, 画像面に対して垂直方向を z 軸, 画像面上を x 軸, y 軸とする (図 4.1).

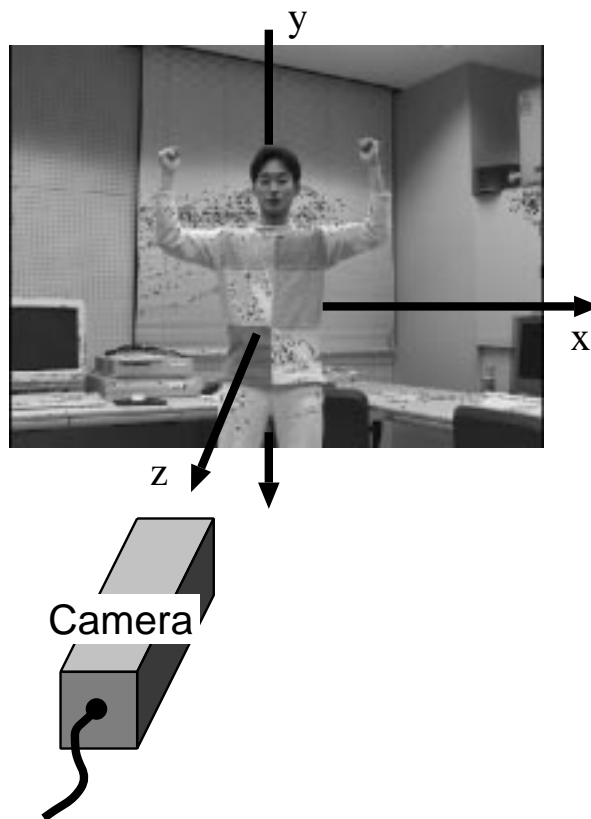


図 4.1: 対象物体と座標軸

4.3 認識対象

本研究では, 認識対象として人物の上半身の動作, 姿勢推定に限定し, 胴体および頭部に関してはビデオカメラに対して正面を向いた状態を保ち固定と仮定する. これは, 速度ベ

クトルを用いて人体の動作認識をする本手法の有効性を確認するための初期段階として、できるだけ問題の簡略化を図るのが目的である。また、人物動作および姿勢の概略を求めることを目的としており、手、指の認識は行わない。提案する手法により、領域情報を用いてフロー情報を大局的に見るために各モデルに対応する体節領域は、ある程度の大きさを持っていることとする。これは、上腕部および下腕部に対して極度に z 軸方向へ向かった姿勢に対して制限をすることである。

4.4 処理の流れ

全体の処理の概略を図 4.3 に示す。まず、動画像系列の各フレーム枚に被験者の姿勢推定を行うための前処理として、背景画像との差分により人物領域の抽出を行う。初期フレームにおいては、モデルの初期設定を行うことにより対象となる人体の関節位置、モデルの長さおよび幅情報を決定し、各体節ごとにモデル領域を求めておく。次に、初期フレームに続く時系列画像 2 枚を入力し、計 3 枚からフローベクトルを連続する画像間で現時点のフローベクトル (flow0) と 1 フレーム後のフローベクトル (flow1) として計算する (図 4.2)。各モデル領域ごとに、体節領域内の全フローベクトルを用いて投票を行い、現フレームでの体節の速度パラメータを推定する。推定された速度パラメータを用いることにより、次フレームにおけるモデルの速度場候補を生成する。次フレームにおけるフロー場上で速度場との照合により領域単位でのモデルの対応づけを行いモデル領域を更新する。続く時系列画像を入力し、再びフローベクトルを求める処理に戻る。

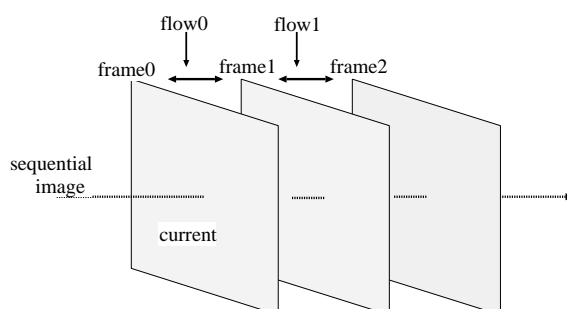


図 4.2: オプティカルフローの計算

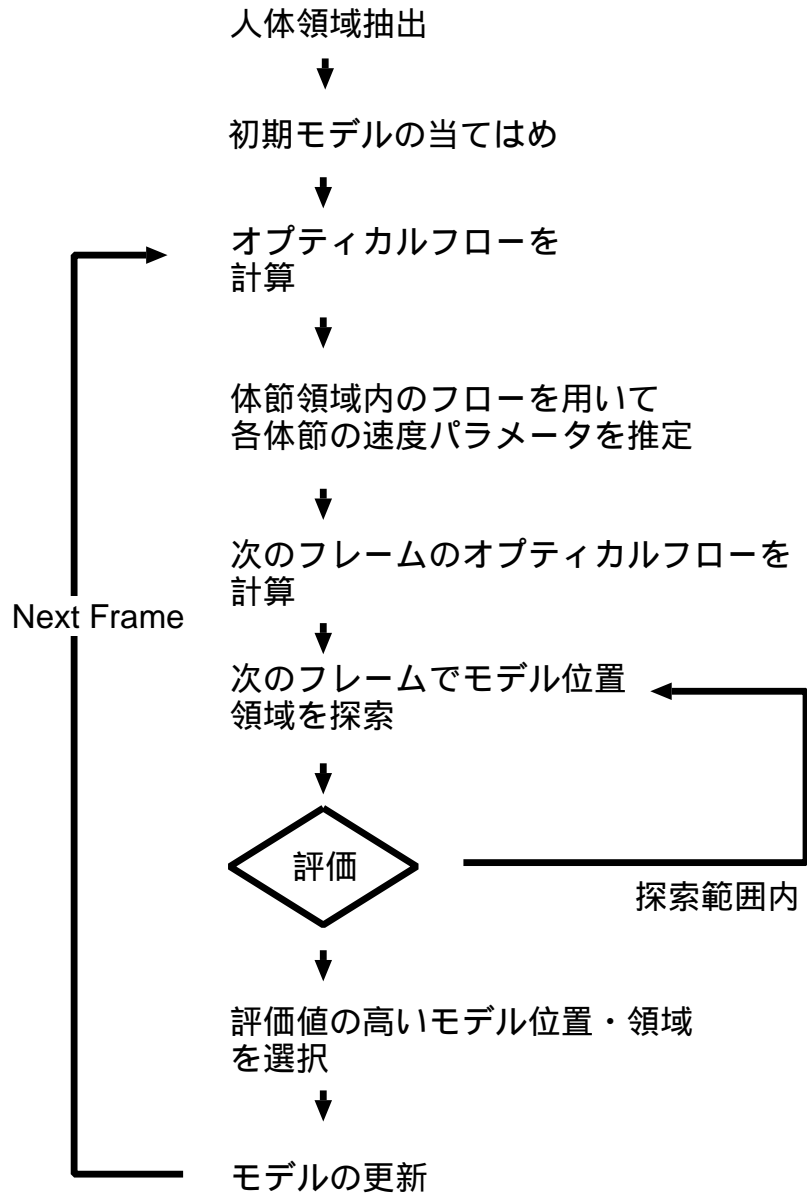


図 4.3: 処理の流れ

4.5 人体モデル

認識対象を限定した場合、画像から抽出できる特徴に加え、その認識対象に関する知識を有効に利用するために、モデルを用いることが行われる。これまで認識対象によりさまざまなモデルが提案されている。認識対象が立方体や円柱など剛体の場合には、その形状の多面体モデルが用いられ、人間の頭部のような物体の表面のモデリングにはワイヤフレームモデルが適している [10]。本研究での認識対象である人体は、複雑な形状を持ち多様な変形を行うため、そのままモデル化するのは容易ではない。そこで、個々の体節を剛体とみなすことにより、人体を剛体の連結体と考え一般化円筒や、その 2 次元版であるリボンにより表現するのが一般的であり良く用いられている。また、より簡単なモデルとして人体の特徴点 (端点・節点) 間を辺で結ぶスティックモデルなどもよく用いられる。

本研究では、オプティカルフロー情報を基に人体の各体節の速度パラメータの推定と領域の推定を行うために、領域情報をもつモデルとして一般化円筒の 2 次元版のリボンを個々の体節の基本モデルとする。本研究で用いるリボンモデルは、各体節を剛体と考え、その画像面への投影である 2 次元的な形状は、長方形で近似し、長さ情報および幅情報を持たせることによって平面的な体節の領域を表現する。また、人体は、剛体の連結体としてリボンの連結で表す。リボン間は、関節で接続され接続関係は木構造で表す。

本研究では、以下の条件を考慮にいてモデルを構築した。

1. モデルの使用目的はなにか。
2. 汎用性のあるモデルであるか。
3. どの程度の認識を行うか。

条件 1 について、本研究では時系列な動画画像から抽出される局所的な速度情報を表すオプティカルフローを用いる。フローベクトルは、一般的に誤差を含むため大局的に見る必要がある。領域を持ったモデルを用いることによって領域内の各点におけるフローベクトルはモデル全体として一貫性のある運動を示すものであると考え、局所的な情報から大局的な情報を得ることができる。人体に関する知識として多関節物体における運動の依存関係を得ることができる。また、人体モデルを用いることによって推定すべき領域の個数、および探索範囲が限定される。

条件 2 について、精密なモデルは被験者ごとのモデルを構築することが難しいが、本研究で用いるリボンモデルは、特徴点 (端点・関節点) 間を結ぶスティックの情報 (モデルの長さ) に、領域の幅情報を加えることにより比較的簡単に領域情報を持つモデルを構築することができ、汎用性に優れている。

条件 3 について、本研究では人物動作の概略を求めることを目的としており、手首や指などの細かな動きの認識は行わないためそれほど厳密なモデルを要求しない。

以上の 3 点の条件を満たすものとして図 4.4 のようなリボンモデルを提案する。また上半身領域は、胴体部および頭部は固定と考え、下腕部および上腕部の姿勢推定、認識処理を行う。このモデルは、人体を剛体の連結体と考え、人体領域は、胴体部および頭部 (T)、右上腕部 (RUA)、右下腕部 (RLA)、左上腕部 (LUA)、左下腕部 (LLA) の 5 つの領域に割り当てる。各領域は、長方形のリボンで近似する。各リボンは人体の接続関係に従い、それぞれ関節 (右肩 (sh-r), 右肘 (el-r), 左肩 (sh-l), 左肘 (el-l): 計 4 箇所) で連結することで構成される。また各リボンは、リボンの中心線としてのスティック情報 (長さ情報 (l)) に加え、幅情報 (w) を持つ。人体モデルの姿勢は、胴体位置と回転中心 (関節) の座標およびリボンの方向、モデルの長さ、幅で表現し、モデルの自由度は、各関節においてそれぞれ z 軸を中心とする回転 1 自由度のみとする (図 4.5)。

ただし、 z 軸方向への回転を考慮し、 z 軸方向への 3 次元的な移動はリボンの長さで表現する。またリボンの長さについてあらかじめ最大長、最小長の設定を対象領域に合わせて初期設定とともに手動で行う。以上、モデル情報を表 4.1 に示す。

リボンの最小長 \leq リボンの長さ \leq リボンの最大長

リボン	胴体 (T), 右上腕部 (RUA), 右下腕部 (RLA), 左上腕部 (LUA), 左下腕部 (LLA): 計 5 個
関節	右肩 (sh-r), 右肘 (el-r), 左肩 (sh-l), 左肘 (el-l): 計 4 箇所
リボン領域	長さ (l), 幅 (w) を用いて長方形で近似
人体モデルの姿勢	胴体位置, 回転中心 (関節の座標), リボンの方向 (スティックの始点と終点の座標)

表 4.1: モデル情報

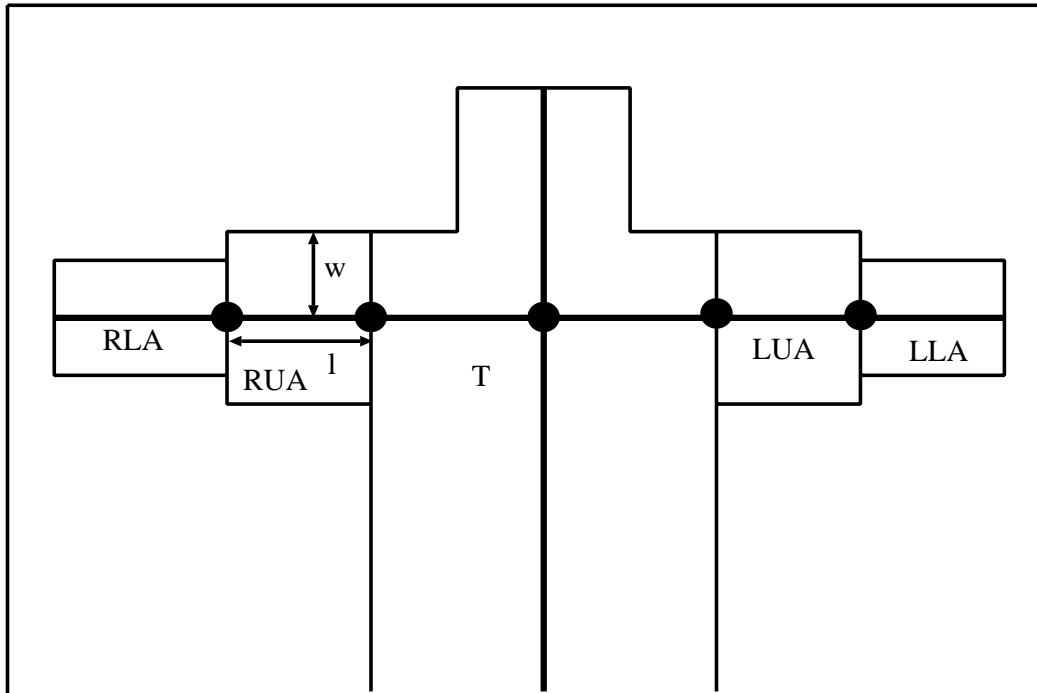


図 4.4: 人体モデル

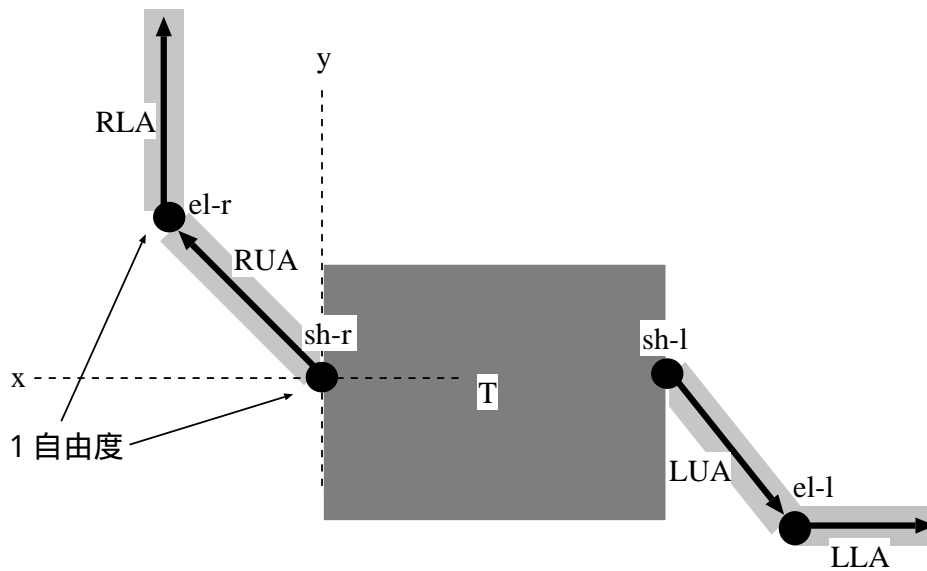


図 4.5: 人体モデルの姿勢

4.6 認識処理

4.6.1 人物領域の抽出

本研究では、オプティカルフローを動作認識における有効な情報として活用する。そこで、人物領域上でフローベクトルを正確にかつ効率的に計算することが重要となってくる。フローベクトルの計算は一般的に非常に時間がかかるので、必要な部分のみを計算することが要求される。そこで画像中の人物領域上のみを計算すれば効率的である。オプティカルフローは、グラディエント法を用いているため、動画像中の濃淡パターンが滑らかに変化していることが条件となっているため計算する画像に対して画像を一様に滑らかにする正規関数(ガウス関数)をかける。この際、切り出した人物領域画像を用いることによって背景による濃淡パターンの影響を無視することができる。そのため、シーン中の対象物体の領域を正確に切り出す必要がある。

入力画像

処理に用いる入力画像は、背景とその前で動作する被験者を単眼固定の $6mm$ デジタルカメラで撮影した動画像系列 (640×480 画素・濃淡 256 階調, 1 フレーム = $1/30sec$) である。本研究で用いる入力画像には、撮影の際に背景、照明などについて特に制限する必要はない。

前処理

画像入力において、ノイズが混入しているため画像の濃度情報を正しく求められない。そのため静止している背景の部分でも階調の差分値が 0 にならない場合がある。この影響を低減するために画像の平滑化を行い、ノイズを除去する。

前述したガウシアンフィルタ以外に、平滑化フィルタの代表的なものとして、画素の近傍画素の濃度平均値をその画素の濃度値と置き換える移動平均法と、ある画素の近傍画素の中央値をその画素の中央値とするメディアンフィルタとがある [15]。

前者は画像をぼかすことにより細かいノイズの影響を受けてしまう。後者の場合は、前者に比べ計算時間を要するが、フィルタリング後の濃度値としてノイズ成分は選択されにくいので、ノイズ除去においても、エッジの保存においても非常に優れた手法である。

本研究では、人物領域を正確に抽出したいため、エッジを保存する平滑化フィルタとして、 3×3 のメディアンフィルタを用いることとし、背景画像を含み、対象となる全ての画像に対し、平滑化を施しておく。

背景分離

従来、あらかじめ用意した背景画像と人物画像の各点における明度の差を適当な閾値で二値化し、人物領域を抽出する背景差分がよく用いられている。しかし背景差分においては、背景との明度差が少ない場合や影による明度変化の影響が無視できないなどの問題があり人物領域の抽出が困難である。そこで本研究では、背景画像と人物画像との正規化相互相関を計算する。正規化相互相関は、各点と近傍の画素を用いることによって明度パターンの類似度を計算したもので、これが高い画素を背景とする。

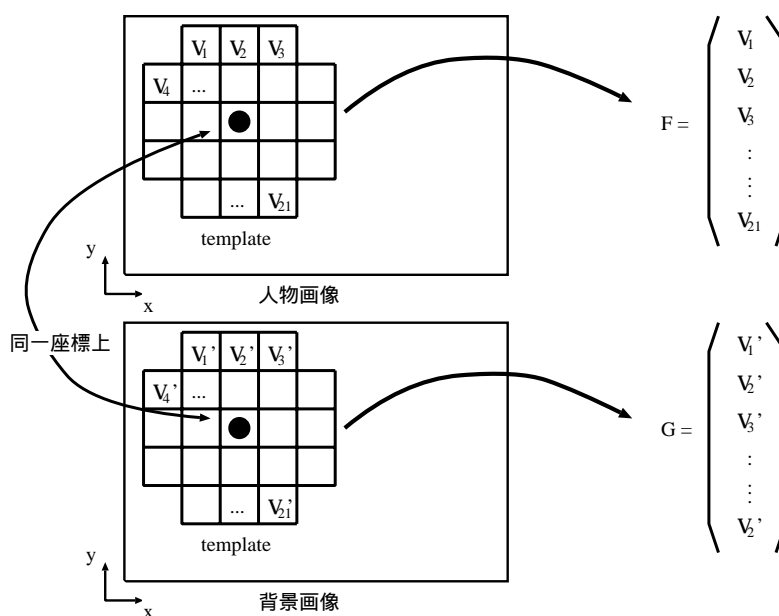


図 4.6: 画像上の点のベクトル表現

正規化相互相関法の手順を以下に示す.

1. 各点を近傍 (テンプレート内) の画素を用いてベクトルで表現する (図 4.6).
2. 背景画像と人物画像の対応する点で相関値を計算する.
3. 相関値が高い点は, 人物画像と背景画像の類似度が高い.
4. 相関値が閾値以上の点を背景とする.

$$\text{正規化相互相関} = \frac{F^T G}{\|F\| \|G\|} \quad (4.1)$$

F: 人物画像上の点のベクトル表現

G: 背景画像上の点のベクトル表現

この手法の利点は, 従来の背景差分法における影を人体領域として抽出してしまう問題を解決している. また背景と人体との明度差が少ない点も良好に抽出が行え, 特に人物領域の境界上などパターンが強く変わる部分に対して正確に人物領域は抽出できる.

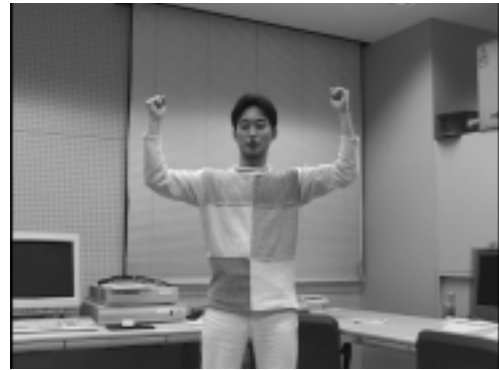
領域補正

正規化相互相関法によって背景から切り出した人物領域は, ノイズの影響や背景との明度パターンが同じであることにより欠落した部分がある. 本研究ではオプティカルフローを用いるため領域に欠落があればその部分でフロー情報が得られないという問題が生じてくる. そのためより正確に領域を得るために切り出し 2 値化した人物領域に対して領域補正処理を行う.

領域補正処理は, 画素の連結数を探查していき連結数の少ない領域はノイズとして排除する. また人物領域内では, 連結数の少ない領域は, 背景として抽出されなかった欠落領域として人物領域に加える.



(a) 背景画像



(b) 人物画像

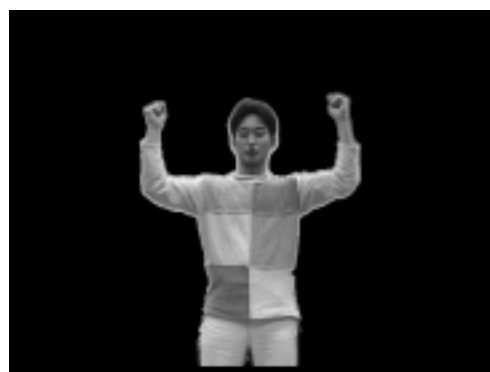
図 4.7: 入力画像



(a) 正規化相互相関



(b) 領域補正



(c) 人物領域

図 4.8: 出力画像

4.6.2 速度パラメータの推定

本節では、局所速度情報としてのオプティカルフロー情報から「リボン内のフローベクトルは、全体で一貫したリボンの運動を表すものである」という考えのもと大局的な運動情報として各リボンにおける速度パラメータの推定方法について述べる。

まず、各点でのフロー情報を大局的に活用するためにモデルの初期設定として各リボンと対応する体節を手動で設定する。このとき各リボンにおける幅情報、長さ情報（最大長、最小長）も設定しておく（図 4.9）。

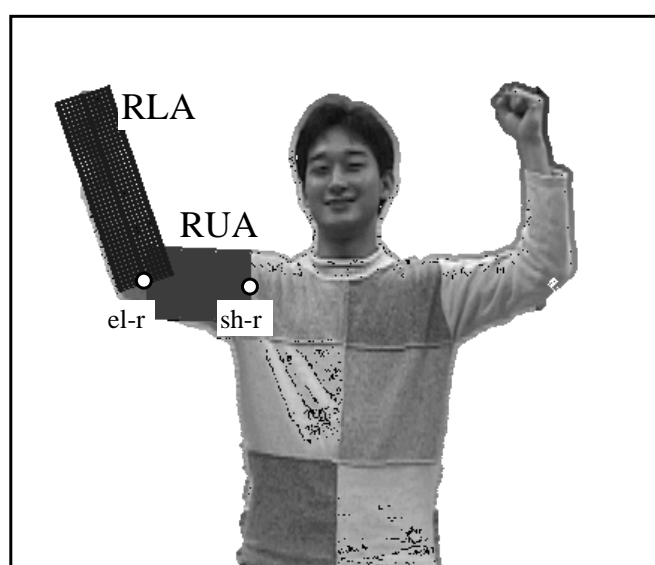


図 4.9: モデルの初期設定

次に、各リボンの速度パラメータを求めるために、抽出した人物領域上でフローベクトルを計算する。本研究では、グラディエント法を用いているため、空間的に画像を滑らかにすることを目的に人物画像に対して、 $\sigma = 3$ のガウシアンフィルタ式 (4.2) を通した後、フローベクトルを計算する。

$$g(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-m)^2}{2\sigma^2}} \quad (4.2)$$

3次元運動とオプティカルフローの解釈

初期設定を行った後、各リボンの速度パラメータを求めるために、3次元運動を投影した2次元の時系列動画像上から計算されるオプティカルフローの解釈について述べる。

各関節(回転中心)における local 座標軸は、画像座標系で z 方向を z 軸、画像座標系 $x-y$ 平面上で体節(リボン)方向、体節に垂直方向を x 軸、 y 軸とする(図 4.10)。

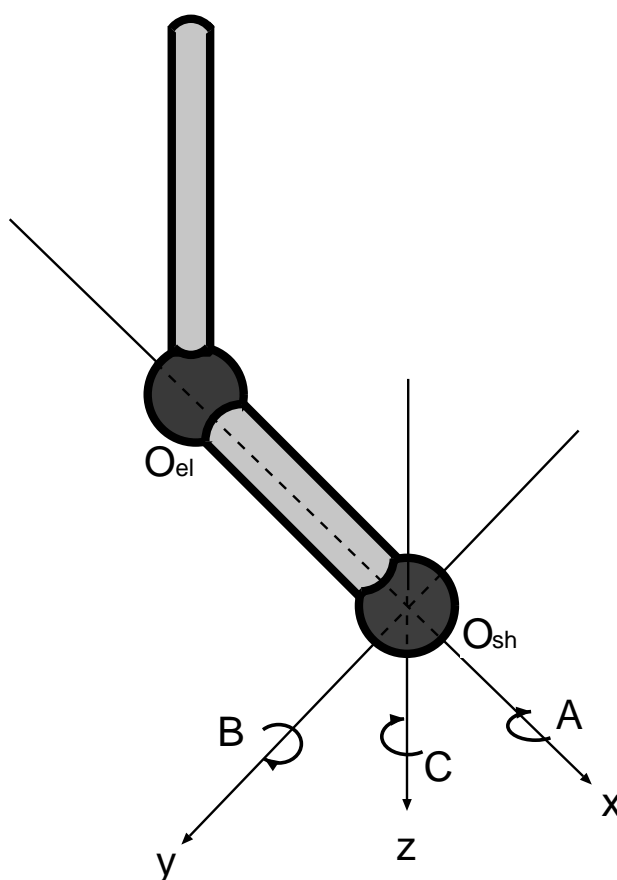


図 4.10: 各関節(回転中心における)local 座標軸

本研究では、胴体部は固定であるという条件を課しているため、人間の動作は、各関節における local 座標軸において x 軸、 y 軸、 z 軸を中心とする回転運動によって行われる。よってリボンの速度パラメータは、回転運動だけからなる場合を考えればよい。

3次元回転運動における3次元速度ベクトルの投影であるフローベクトル成分について考察する。3次元座標軸を X, Y, Z 軸とし、各軸における回転角速度を A, B, C とする。

フローベクトルの回転成分を各軸周りの成分に分けて考えると、 z 軸周りの回転によるフローベクトルは、回転角速度 C と回転軸からの距離により決まる。

図 4.11(a) にその幾何学的な意味を示す。 Z 軸周りの回転では、画像上で原点中心の同心円上にフローベクトルが存在し、その方向は、同心円の接線方向、即ち、原点からその点に向かうベクトルに垂直である。

X 軸周りの回転によるフローベクトルは、 x 軸からの距離と角速度 A による成分のみで決まる。図 4.11(b) にその幾何学的な意味を示す。 X 軸周りの回転では、 X 軸を中心軸とする円筒の垂直断面の円周上に 3 次元ベクトルが存在し、それを画像に写像すると、正射影を仮定しているため回転軸方向に垂直な速度ベクトルとなる。

同様に Y 軸周りの回転によるフロー成分も、 y 軸からの距離と角速度 B による成分のみで決まる。幾何学的な意味は X 軸周りと同様である。いずれの回転ベクトルも 3 次元位置 (X, Y, Z) に依存せず画面上の位置によって決まり、画像上での回転中心からの距離によって線形である。

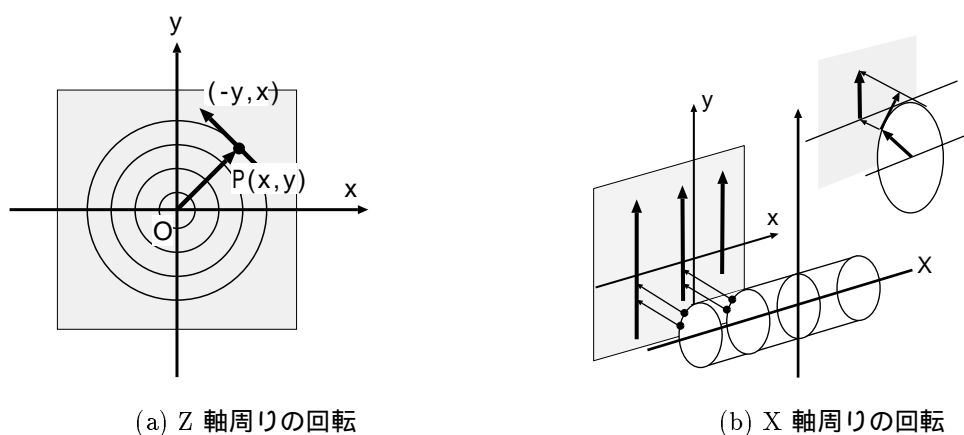


図 4.11:

以上のような、3次元運動に対応するフローベクトルの性質を用いて各リボンにおけるフローベクトルを個々の回転成分に分割し、リボンの速度パラメータを推定する。

フローベクトルの分割

本研究では, 各体節の運動は各関節を回転中心とする local 座標軸において y 軸周りの回転運動と z 軸周りの回転運動のみを行っていると考える. x 軸における回転運動において回転半径は, y 軸および z 軸の回転半径に比べ非常に短く, 運動におけるフローベクトルも微小であるとして無視する. フローベクトルは, z 軸および y 軸を中心とする回転運動における個々の速度ベクトルの合成であるとする. リボンの速度パラメータは, local 座標軸における y 軸回転成分, z 軸回転成分について推定する.

まず, 速度パラメータを推定するために実際に計算されたリボン内の各点におけるフローベクトルを各運動における速度成分に分割する (図 4.12).

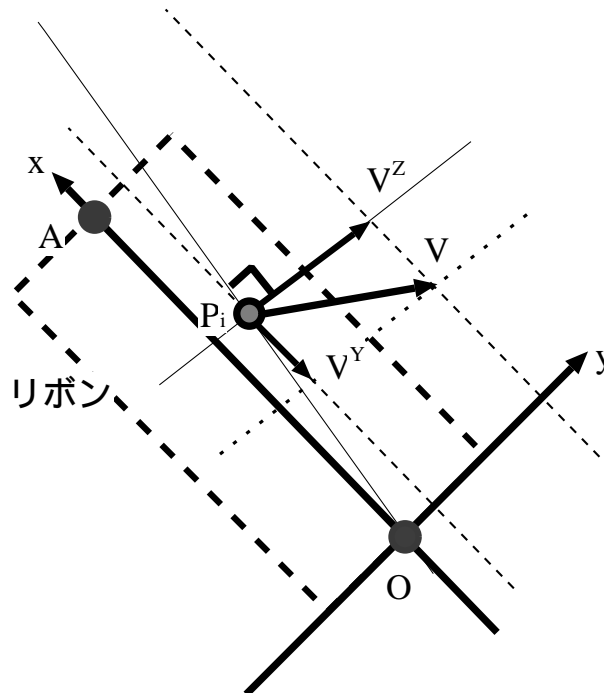


図 4.12: フローの分割

いま, リボンにおける回転中心を原点とする local 座標軸からみたフローベクトルの分割方法を述べる. この座標軸上で原点 (リボンの始点) を $O(0, 0)$, リボンの終点を $A(x_a, y_a)$, リボン内の点を $P_i(x_i, y_i)$ とする. 点 P_i におけるフローベクトル $V_i = (u_i, v_i)^T$, y 軸周りの回転運動による速度ベクトル $V_i^Y = (u_i^Y, v_i^Z)$, z 軸周りの回転運動による速度ベクトル $V_i^Z = (u_i^Z, v_i^Z)$ は次式の関係が成り立つ.

$$\mathbf{V} = \mathbf{V}_i^Y + \mathbf{V}_i^Z \quad (4.3)$$

$$= n \frac{\vec{OP}}{\|\vec{OP}\|} R(\theta_1) + m \frac{\vec{OA}}{\|\vec{OA}\|} R(\theta_2) \quad (4.4)$$

ここで,

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (4.5)$$

$$\theta_1 = \begin{cases} \frac{\pi}{2} & : \text{回転方向が時計周りのとき} \\ -\frac{\pi}{2} & : \text{それ以外} \end{cases} \quad (4.6)$$

$$\theta_2 = \begin{cases} \pi & : \text{回転方向がカメラに近づく方向のとき} \\ 0 & : \text{それ以外} \end{cases} \quad (4.7)$$

となる. n, m は, $\mathbf{V}_i^Y, \mathbf{V}_i^Z$ のノルムを表し式 (4.4) において (u_i, v_i) について解くことによって, 未知数 $2(n, m)$ で式が 2 つなので, モデル内の全ての点について $\mathbf{V}_i^Y, \mathbf{V}_i^Z$ となる速度ベクトルが得られる.

投票による速度パラメータ推定

オプティカルフローは, 局所速度情報であり誤差が多いため一点のフローベクトルからだけではリボンの運動を決定するには信頼性が低い. 各リボン内の「フローベクトルは, 全体として一貫した運動を示すものである」という考えに基づき, 各点からの局所速度情報を大局的なリボンの速度情報として取り扱う方法として, 投票により運動パラメータを決定する. 投票を行うのは, リボン内のある部分点をリボンの動きを示す代表点とするのではなく, 全ての点において最も最適な運動を推定するということである.

投票において z 軸周りの回転成分は, 回転中心に対して同心円上に距離に比例して速度ベクトルの大きさが線形であるため, 角速度 $\delta\theta$ で, y 軸周りの回転成分においては, y 軸からの距離に比例して速度ベクトルの大きさが線形であるため, 距離に対する速度ベクトルの比率 δz として全ての点において統一的に扱える (図 4.13).

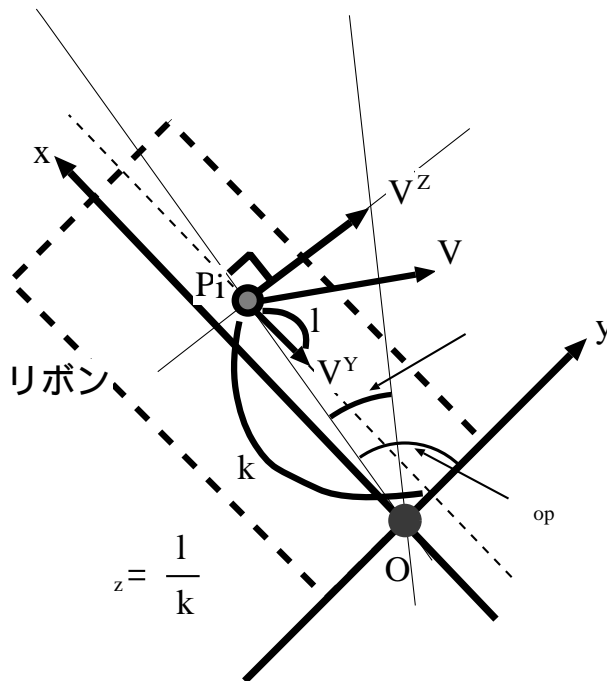


図 4.13: 速度パラメータ

点 $P_i(x_i, y_i)$ における回転成分は,

$$\delta\theta(i) = \tan^{-1}\left(\frac{\|\mathbf{V}_i^Z\|}{\|\vec{OP}_i\|}\right) \quad (4.8)$$

および,

$$\delta z(i) = \frac{\|\mathbf{V}_i^Y\|}{\|\vec{OP}_i\| \sin(\theta_{OP_i})} \quad (4.9)$$

となる。 θ_{OP_i} は \vec{OP}_i と y 軸との角度である。

2つのパラメータについてそれぞれ投票を行う。投票空間において対象とした腕の長さを考慮すると、回転中心からの距離 100 画素の点において 1 画素程度の動きを解析できる分解能として $\delta\theta$, δz とともに小数第 2 位までとし回転角の単位はラジアンを用いる。

投票において、図 4.14 のように回転中心に誤差があった場合、回転中心に近い点においては $\delta\theta$, δz に大きな誤差が出る。また回転中心付近においては、衣服による領域の変形お

よび移動が観測されるため、投票において信頼性が高いのは回転中心から遠い点と考え、距離によって投票値に重みづけする。

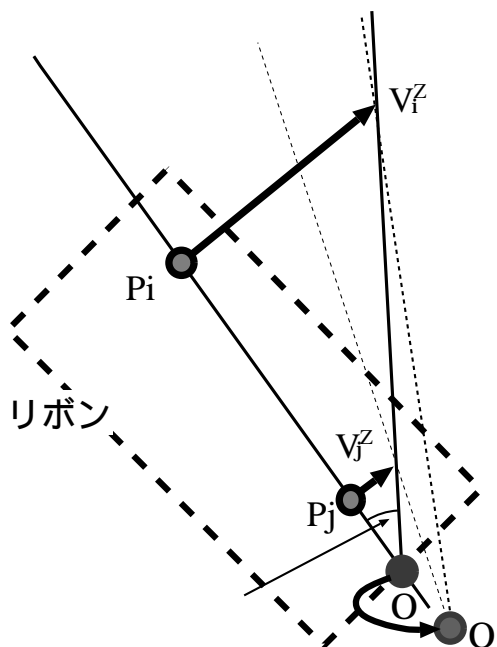


図 4.14: 回転中心の誤差

リボン内のある点 $P_i(x_i, y_i)$ における重みを W_i^Z, W_i^Y とし、 l_i^Z, l_i^Y を以下のように定義する。

$$l_i^Z = \frac{\|\vec{OP}_i\|}{length_{(parts)}} \quad (4.10)$$

$$l_i^Y = \frac{\|\vec{OP}_i\| \sin(\theta_{OP_i})}{length_{(parts)}} \quad (4.11)$$

ここで $length_{(parts)}$ は各リボンの長さで、本研究では重みを表 4.2 のようにする。

投票は、リボン内 (D) の全ての点 $P_i(x_i, y_i), i \in D$ において行う。また、微小時間での解析を行っているため投票範囲は、 z 軸周りの回転において $(-1.00 < \delta\theta < 1.00)$ 、 y 軸周りの回転において長さを越える変化はないとして $(-1.00 < \delta z < 1.00)$ とする。

投票分布において上腕部における最大投票値を得たパラメータを $\delta\theta_{max}, \delta z_{max}$ とすると、

l_i^Z	$l_i \leq 0.2$	$0.2 < l_i \leq 0.4$	$0.4 < l_i \leq 0.6$	$0.6 < l_i \leq 0.8$	$0.8 < l_i$
\mathbf{W}_i^Z	1	2	3	4	5
l_i^Y	$l_i \leq 0.2$	$0.2 < l_i \leq 0.4$	$0.4 < l_i \leq 0.6$	$0.6 < l_i \leq 0.8$	$0.8 < l_i$
\mathbf{W}_i^Y	1	2	3	4	5

表 4.2: 距離による重み

$$f(\delta\theta) = \begin{cases} 1 & \delta\theta - \Delta\theta < \delta\theta(i) \leq \delta\theta + \Delta\theta (\Delta\theta = 0.005) \\ 0 & \text{それ以外のとき} \end{cases} \quad (4.12)$$

とすると, z 軸周りの回転成分の投票分布関数は,

$$F(\delta\theta) = \sum_{i \in \mathbf{D}} (f(\delta\theta) \cdot \mathbf{W}_i^Z) \quad (4.13)$$

また,

$$g(\delta z) = \begin{cases} 1 & \delta z - \Delta z < \delta z(i) \leq \delta z + \Delta z (\Delta z = 0.005) \\ 0 & \text{それ以外のとき} \end{cases} \quad (4.14)$$

とすると, y 軸周りの回転成分の投票分布関数は,

$$G(\delta z) = \sum_{i \in \mathbf{D}} (f(\delta z) \cdot \mathbf{W}_i^Y) \quad (4.15)$$

よって,

$$\delta\theta_{(max)} = \max_{\delta\theta} F(\delta\theta) \quad (4.16)$$

$$\delta z_{(max)} = \max_{\delta z} G(\delta z) \quad (4.17)$$

となる (図 4.15).

ここまでは, 他の体節による運動に依存しないリボンにおける速度パラメータ, つまり上腕部における速度パラメータの推定方法である. 下腕部における運動は, 上腕部の影響を受ける. 下腕部の回転軸を中心とするリボンの速度パラメータを推定するためには上腕部による速度ベクトルを排除する必要がある.

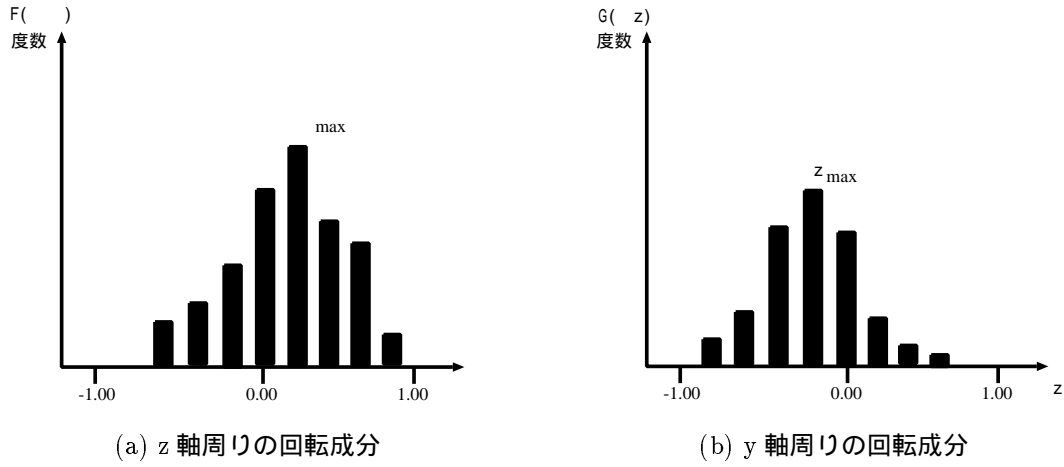


図 4.15: 投票分布

いま, 推定された上腕部における速度パラメータを $\delta\theta_{(upper)}$, $\delta z_{(upper)}$, 下腕部の回転軸を中心とする運動の速度パラメータを $\delta\theta_{(lower)}$, $\delta z_{(lower)}$ として下腕部に対応するリボンの速度パラメータの推定法を述べる.

上腕部の回転中心を原点とする local 座標軸からみて, 下腕部のリボン内の点 $P_j(x_j, y_j) \in D_{lower}$, 上腕部の回転中心を O , 下腕部の回転中心を A とすると, 点 P_j における上腕部回転運動による速度ベクトルは, z 軸周りの回転運動においては,

$$\mathbf{V}_{\delta\theta_{(upper)}} = \frac{O\vec{P}_j}{\|O\vec{P}_j\|} R(\theta_1) \cdot \|O\vec{P}_j\| \cdot \|\tan(\delta\theta_{(upper)})\| \quad (4.18)$$

$$= O\vec{P}_j R(\theta_1) \cdot \|\tan(\delta\theta_{(upper)})\| \quad (4.19)$$

$$\theta_1 = \begin{cases} \frac{\pi}{2} & : \delta\theta_{(upper)} \text{ が正のとき} \\ -\frac{\pi}{2} & : \delta\theta_{(upper)} \text{ が負のとき} \end{cases}$$

$$\mathbf{V}_{\delta z_{(upper)}} = \frac{O\vec{A}}{\|O\vec{A}\|} R(\theta_2) \cdot \|O\vec{P}_j\| \sin(\theta_{OP_j}) \cdot \|\delta z_{(upper)}\| \quad (4.20)$$

$$\theta_2 = \begin{cases} 0 & : \delta z_{(upper)} \text{ が正のとき} \\ \pi & : \delta z_{(upper)} \text{ が負のとき} \end{cases}$$

となる. ここで θ_{OP_j} は, $O\vec{P}_j$ と y 軸の角度である.

点 P_j におけるフローベクトルを V_j , 上腕部の影響による速度ベクトルを $V_{j(upper)}$, 下腕部の運動のみにおける速度ベクトルを $V_{j(lower)}$, とすると,

$$V_{j(lower)} = V_j - V_{j(upper)} \quad (4.21)$$

$$= V_j - V_{\delta\theta(upper)} - V_{\delta z(upper)} \quad (4.22)$$

となる (図 4.16).

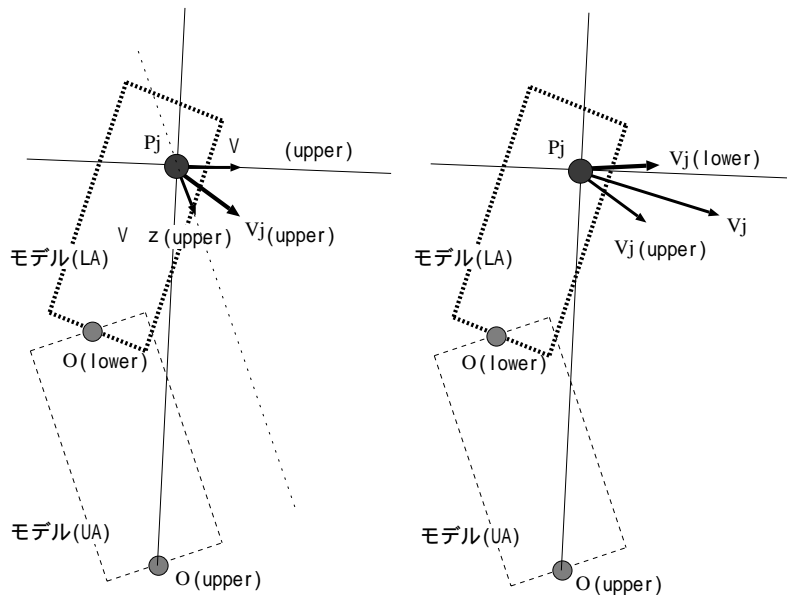


図 4.16: 上腕部による速度ベクトル

このように上腕部による速度成分を下腕部の全点におけるフローベクトルに対して排除する. $V_{j(lower)}$ を下腕部の回転運動によるフローベクトルとしてリボンの速度パラメータ $\delta\theta_{(lower)}$, $\delta z_{(lower)}$ を上腕部と同様に推定する. また, 各ベクトルは上腕部回転中心とする local 座標軸上で計算しているため座標軸の変更にもないベクトルを変換する必要がある.

4.6.3 運動領域の推定・追跡処理

運動領域の推定, 追跡を行うために本研究では領域単位でのモデルの対応づけを行う. 領域単位での対応づけの情報として速度場を用いる. 前節で推定した速度パラメータにおいて, 微小時間における解析のため各体節の運動は前フレームと同じような性質をもつ, つまり各リボンにおける速度パラメータは, 前フレームとほぼ同様のパラメータであるといえる.

運動領域の推定は, 次のような手順で処理を行う.

1. 現フレームのフローベクトルから各リボンにおける速度パラメータを推定する.
2. リボン位置を拘束に従い移動させる.
3. 移動後のリボン内の全点おいての現フレームの速度パラメータを用いて速度ベクトルを計算. これを速度場候補とする.
4. 探索範囲内で (2), (3) の処理を繰り返す.
5. 次フレームにおける実際のオプティカルフローを計算する. フロー場を抽出.
6. フロー場と速度場候補の対応を評価し評価の高いものを次フレームにおけるリボン位置として選択する.
7. 全てのリボン (上腕部, 下腕部の順に行う) に対して次フレームにおけるリボン位置を決定することで人体モデルの姿勢を決定する.

以上の操作により領域の認識, 追跡処理を行う. 図 4.17に概略を示す. 以下処理の詳細について述べる.

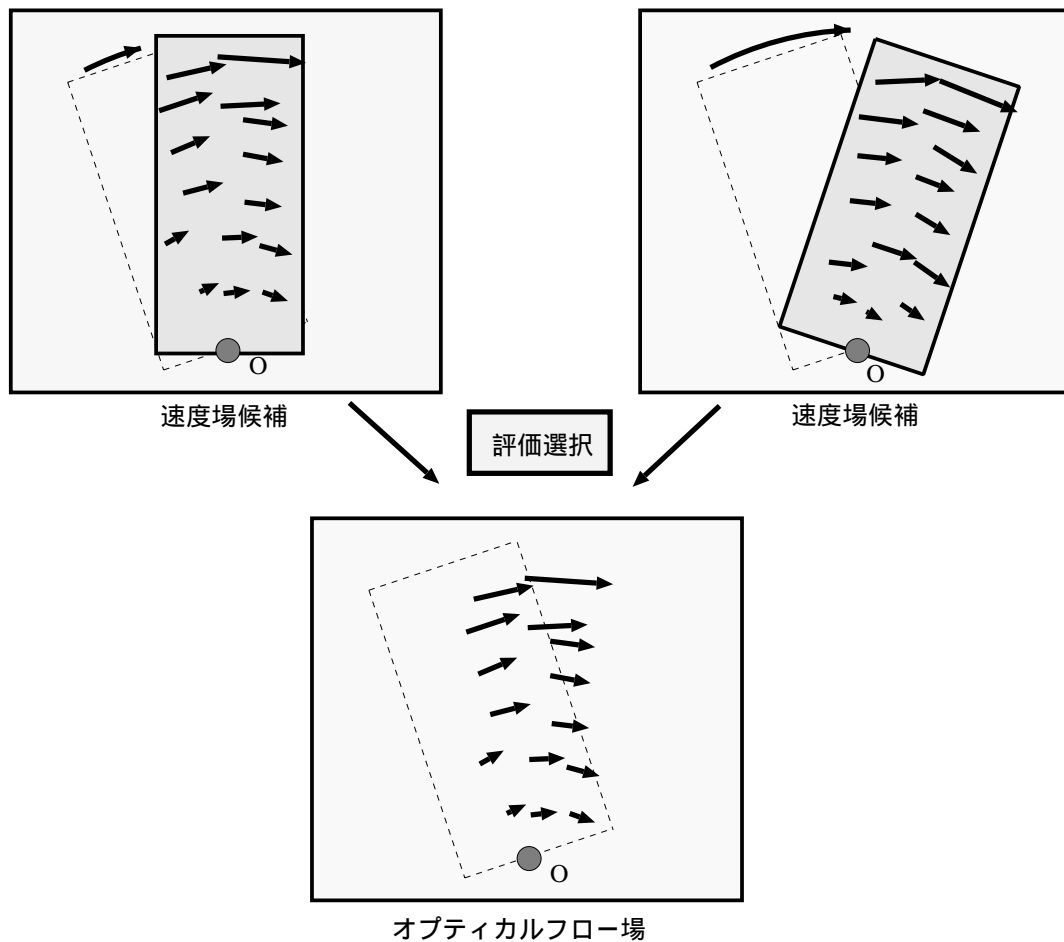


図 4.17: モデルの領域推定

速度場候補の作成

各リボンにおいて推定した速度パラメータを用いて、リボン移動後の各位置における速度場候補を作成するために、各点における速度ベクトルを計算する。速度ベクトルの決定は、4.5.2 で述べたように、local 座標において回転中心、リボンの方向、リボン内の各点の座標によって一意に決定できる。各点における速度ベクトルは、式 (4.19)、式 (4.20)。

上腕部の場合、他の体節の運動に依存しないため独立に計算でき、式 (4.19)、式 (4.20) において各リボンにおける回転中心、計算点座標 (回転中心を原点とする local 座標)、およびリボンの速度パラメータを代入することによって、各点における速度ベクトルを決定する。

下腕部の場合、上腕部の運動に依存するため、下腕部速度パラメータのみによって計算

された速度ベクトルに接続関係にある上腕部の運動による速度ベクトルを合成する。この関係は、式 (4.21)、式 (4.22) で表される。全節では、上腕部による回転成分を排除したが、ここでは上腕部、および下腕部速度パラメータによる速度ベクトルの合成ベクトルをを計算するために用いる。

速度場候補は、全リボンにつき探索範囲内で人体モデルの可動範囲内で少しずつ動かした各位置においてそれぞれ作成する。速度場は、リボン内の各点における速度ベクトルで構成する。

モデルの評価

本研究では、領域単位でのモデルとの対応づけを行うことによって、姿勢推定、追跡を行う。領域の対応づけにおいては、オプティカルフロー情報を用いてフロー場と、領域情報である各リボンにおいて作成した速度場候補との評価により行う。

評価基準として以下の条件を考慮する。

1. 領域内の各点においてフローベクトルと速度ベクトル方向の誤差。
2. 領域内の各点においてフローベクトルと速度ベクトル長の誤差。
3. フロー場と速度場候補の領域の重なりの誤差。

以上の誤差が最小になるようなフロー場をもつ領域 (リボン位置) がモデルと領域の正しい対応づけであると考えられる。

ここで、各位置でのリボン内のある点 $P_i(x_i, y_i)$ における候補となるの速度ベクトルを $\mathbf{V}_j = (u_j, v_j)^T$ 、実際のフローベクトルを $\mathbf{V}_i = (u_i, v_i)^T$ とする。この二つのベクトルの差異をこの点における評価とする。また、ベクトル長に関する差異も考慮に入れるため、2次元ベクトル $\mathbf{V}_j, \mathbf{V}_i$ を3次元ベクトル $\mathbf{V}'_j = (u_j, v_j, w)^T$ 、 $\mathbf{V}'_i = (u_i, v_i, w)^T$ として取り扱う。このとき、2つのベクトルの差異をはかる測度として、角測度を

$$\Psi_{p_i}(\mathbf{V}'_j, \mathbf{V}'_i) = \cos^{-1}\left(\frac{\mathbf{V}'_j}{\|\mathbf{V}'_j\|} \cdot \frac{\mathbf{V}'_i}{\|\mathbf{V}'_i\|}\right) \quad (4.23)$$

とする。式 (4.23) を用いてリボン内の対応する同座標軸上のフローベクトルと速度ベクトルの角誤差を計算するしこの点における評価とする (図 4.18)。ここで、 i, j は、それぞれ同じ座標軸における点で j はベクトルを区別するため便宜上使っている。ここで、 w は、

評価基準 (1),(2) における重みづけとしての役割を果たす. つまり w の値を大きくすると角度による誤差は小さくなる.

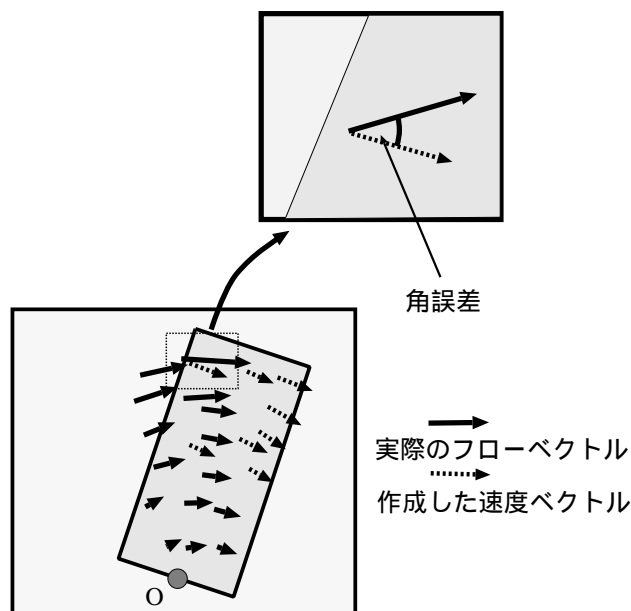


図 4.18: 角誤差

式 (4.23) は, 評価基準 (1),(2) に関する評価を行える. ここで, 評価基準 (3) に関しては, 以下のように評価する.

リボンの評価には, リボン内の全点における角誤差の平均を用い, 領域全体として誤差の少ない速度場候補つまりリボン位置を決定する. 角誤差の平均を各リボン位置におけるモデルの評価とする上で, 速度パラメータの決定において回転軸からの距離によって重みづけを行っていることを考慮にいれる. 推定した速度パラメータは, 回転中心から遠い点におけるフローベクトルを重視しているためリボン全体の誤差の計算においても回転中心からの距離によって重みつき評価を行う.

しかし, 投票における重みづけでは, z 軸中心の回転運動による成分と y 軸中心の回転運動における成分の重みづけは, 同じ点においても回転軸からの距離が違うため正確には異なる. ここでは, 各成分とも遠くの点を重視する重みづけをしているため各軸からの距離に関する重みづけは, ほぼ同じと考える.

リボン内のある点 $P_i(x_i, y_i)$ における重み W_i を, 表 (4.3) のように設定する. リボンの長さを $length_{(parts)}$ とすると, リボンの長さにおける P_i の回転中心からの距離の比率は,

次式で表せる.

$$l_i = \frac{\|\vec{OP}_i\|}{length_{(parts)}} \quad (4.24)$$

l_i	$l_i \leq 0.2$	$0.2 < l_i \leq 0.4$	$0.4 < l_i \leq 0.6$	$0.6 < l_i \leq 0.8$	$0.8 < l_i$
\mathbf{W}_i	1	2	3	4	5

表 4.3: 評価における重み

重みづけした リボンの角誤差は,

$$\Psi_{(parts)} = \frac{\sum_{i \in \mathbf{D}} \mathbf{W}_i \Psi_{P_i}}{\sum_{i \in \mathbf{D}} \mathbf{W}_i} \quad (4.25)$$

となる. 評価値は, 角誤差で表し誤差が少ないほど評価が高いとする.

モデルの修正

人体動作の追跡を行うために, 人体の動作にもなってモデルの修正を行う必要がある. 本研究では, 微小時間における解析を行っているため, 各リボンにおいて現フレームでの速度パラメータは, 次のフレームでの速度パラメータとほぼ同じであるという考えのもと, 人体モデルを用いてリボン位置を探索する. 人体モデルの拘束および推定した速度パラメータを用いることによって, リボン位置の探索範囲を限定することは, 計算量においても誤認識を避けるという点においても有効である.

提案するモデルは, 各関節において z 軸周りの回転運動 1 自由度である. y 軸周りの回転運動に対しては, モデルの長さにより 2 次元的な姿勢を決定する.

上腕部においては, 次の条件に従い探索している.

1. 長さによる評価は, 肘の位置において下腕部との領域がはっきりしないため z 軸周りの回転のみの探索を行う.
2. 速度パラメータにより探索方向を決定する.

3. 速度パラメータは、角速度でありモデルの探索範囲においてある程度の指標を与えるものとして活用する。しかしあくまでもフロー場における領域単位での対応づけを行うため、そのまま移動ベクトルとは考えない。

以上のような条件により探索範囲を決定し実際には回転角 $\Delta\theta = 0.01rad$ ごとにモデルを回転させ各リボン位置において領域単位でのモデルの対応づけを行い評価値の高い(平均誤差が最小値)対応づけを選択する。

下腕部においては次のような条件に従い探索を行っている。

1. 上腕部との依存関係により回転軸が移動するため、肘における回転各を固定とし、上腕部の回転軸中心にモデル位置を修正する。
2. 上腕部における長さの探索を行い、下腕部の評価値により決定する。この場合、肘による回転も同時に行いながら探索する。
3. 長さの探索においても速度パラメータにより範囲を限定する。

以上の探索によりリボン位置を決定した後、下腕部における長さの評価を行う。全てのリボン位置・領域を決定したのち、これを次フレームのモデルとしてモデルの修正を行う。

4.7 まとめ

本章では、動画像から、オプティカルフロー情報、領域情報、人体モデルを用いた領域単位での追跡、対応づけによる人体の動作認識、および各体節の位置、領域推定法について具体的な処理方法を述べた。

以上の処理をまとめると、

1. 背景からの人物領域の分離。
2. 初期設定により個々の体節(リボン)へ領域分割。
3. フロー情報によりリボンの速度パラメータを推定。
4. 次フレームのフロー場上で領域単位の各体節の追跡、姿勢の推定

という処理により新しいアプローチを示した。

第 5 章

実験と考察

5.1 はじめに

第 4 章で述べた本手法の有効性を検証するために 2 種類の動作認識実験を行った。一つは体節の重なりのない動作で、もう一つは体節の重なりが生じる動作について実験を行っている。

まず入力画像に対して、あらかじめ背景差分処理を行った。正規化相互相関における閾値は、手動で決定している。次に、初期フレームにおいて、モデルの初期設定を手動で行う。このとき人体の知識として幅情報、長さ情報も設定する。以下、本手法による実験結果を示す。

5.2 入力画像

実験に用いる人物の動きは単眼固定の $6mm$ デジタルカメラにより撮影する。撮影された動画像を 1 フレーム (1 フレーム = $1/30sec$) ごとにファイルにおとし、それを入力画像系列として実験を行う。ビデオカメラの出力画像サイズは、 640×480 画素、RGB 各 256 階調のものであり、本手法の実験には、この画像を 256 階調の濃淡画像に変換したものをを用いる。撮影は、室内で一人の被験者に対して行い、画像中では被験者以外に移動物体は無いものとする。また対象は、上半身のみを撮影する。照明は特に制限せず一般的な室内蛍光灯のみとしている。処理の前に必要となる背景画像は人物が動作を行う前に同じ環境で撮影しておく。

5.3 体節の重なりがない場合

実験には、図 5.1, 図 5.2は、屈伸運動を追跡した結果で、図 5.3 は回転運動を追跡した結果である。これらの画像は、体節の重なりがない動作を行う被験者を撮影したものを入力画像として用いた。認識結果は、全てのフレームのモデルの中心線となるスティックを背景分離を行った人物領域画像に重ね合わせることで示す。本手法による結果は、全てのフレームにおいて視覚的ではあるが正確に追跡・姿勢推定が行えたと考えられる。

5.4 体節の重なりが生じる場合

実験には、図 5.4 に示す体節の重なりが生じる動作を行う被験者を撮影したものを入力画像として用いた。認識結果は、全てのフレームのモデルの中心線となるスティックを背景分離した人物領域画像に重ね合わせることで示す。

本手法による結果は、全てのフレームにおいて視覚的ではあるが被験者のおおよその動作、姿勢を追跡できていることが解る。重なりが生じない場合に比べて、比較的精度は落ちるが、重なりがなくなった時点で正確にモデルの修正が行えている。

5.5 実験結果の考察

以上の結果に基づき本手法の有効性について考察する。背景の分離はあらかじめ用意された背景画像と人物画像との正規化相互相関を計算することにより、全ての画像に対して領域の欠落や不必要な領域もなく良好な結果が得られているといえる。

領域の重なりのない場合の認識結果は、視覚的ではあるが非常に良好に行えているといえる。屈伸運動においては、急激に運動方向が変わるが、微小時間における解析を行っているため、速度の変化も微小であり良好な結果が得られている。しかし、本手法では前フレームでの速度情報を用いてフロー場と作成した速度場候補の対応づけを行っているため、もっと早い動作では、ずれが生じると考えられる。急激な速度、方向の変化がおきた場合、フロー場と速度場候補の誤差が大きくなり、いったん対応が外れてしまえばフレームごとに速度パラメータを推定しているため、致命的であると考えられる。

回転運動の場合においても良好な結果が得られていると考えられる。この場合、速度がほぼ等速回転運動を行っているため、速度パラメータは安定していて、良好な結果が得られている。

領域の重なりが生じる場合の結果は、重なりのない場合に比べると精度は落ちるが、2次元的な姿勢の概略は得られており良好な結果といえる。この場合、精度が落ちている理由として、人物領域上でガウシアン関数により平滑化を行っているため、他の領域における輝度の影響やフロー場が全体にぼやけている、などが考えられる。また、4章で述べたように、 x 軸の回転に対する速度ベクトル成分を考慮にいれてないことが影響を及ぼしていると考えられる。

5.6 まとめ

本章では, 第4章で述べた本手法の有効性を検証するために2種類の動作認識実験を行い, その結果についての検討を行った.

結果の検討を以下にまとめる.

- 正規化相互相関法により, 正確に人物領域の抽出が行えた.
- 重なりの生じない場合においては, 十分満足できる結果が得られた.
- 屈伸運動においては, 速度変化がゆっくりであるため, 領域の対応づけに問題はなかった.
- 重なりの生じる場合は, 平滑化の影響によりずれが生じたが, 視覚的だが大体満足できる結果が得られた.
- 胴体固定など制限が多いが, 本手法の有効性が確認できた.

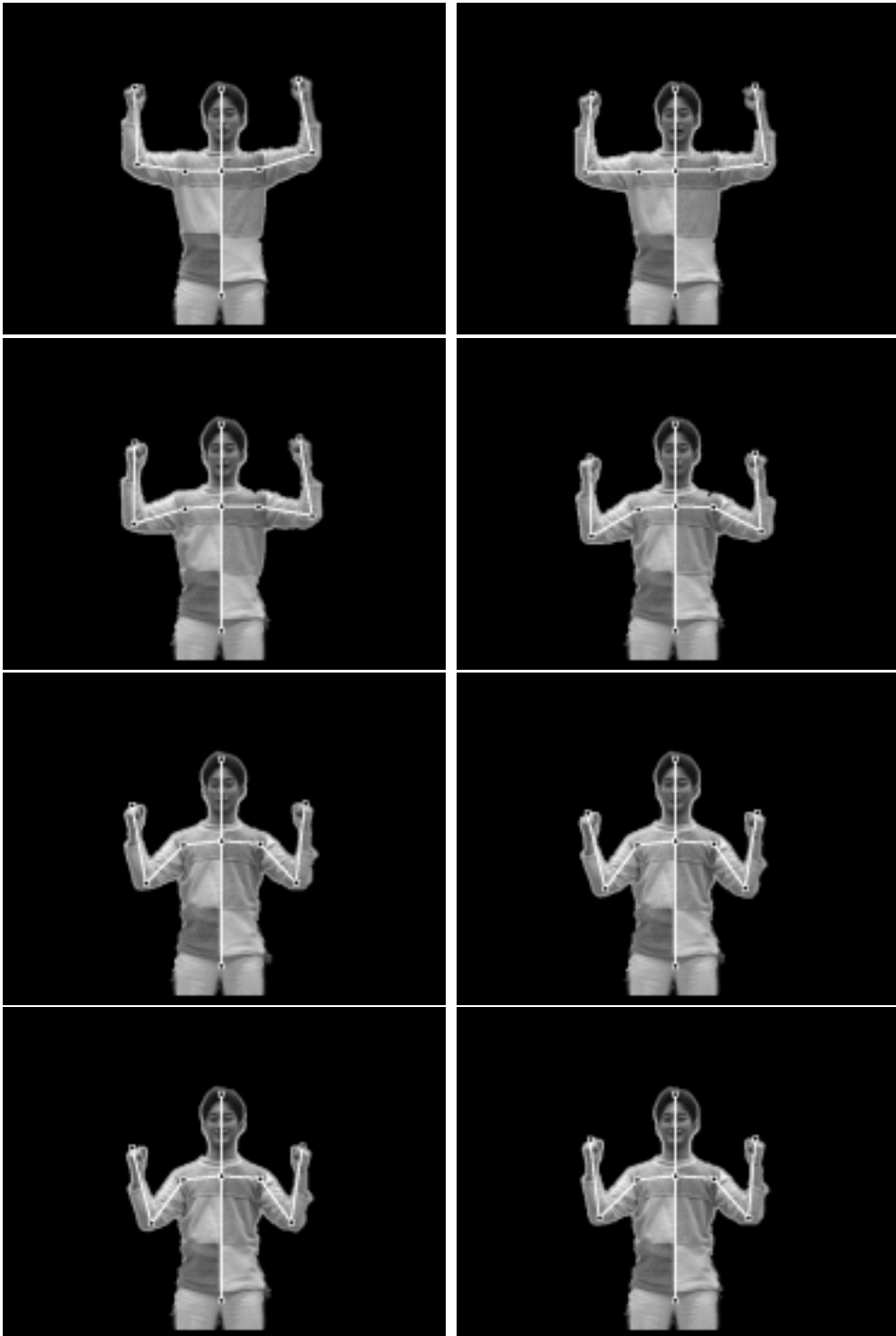


図 5.1: 屈伸運動 1

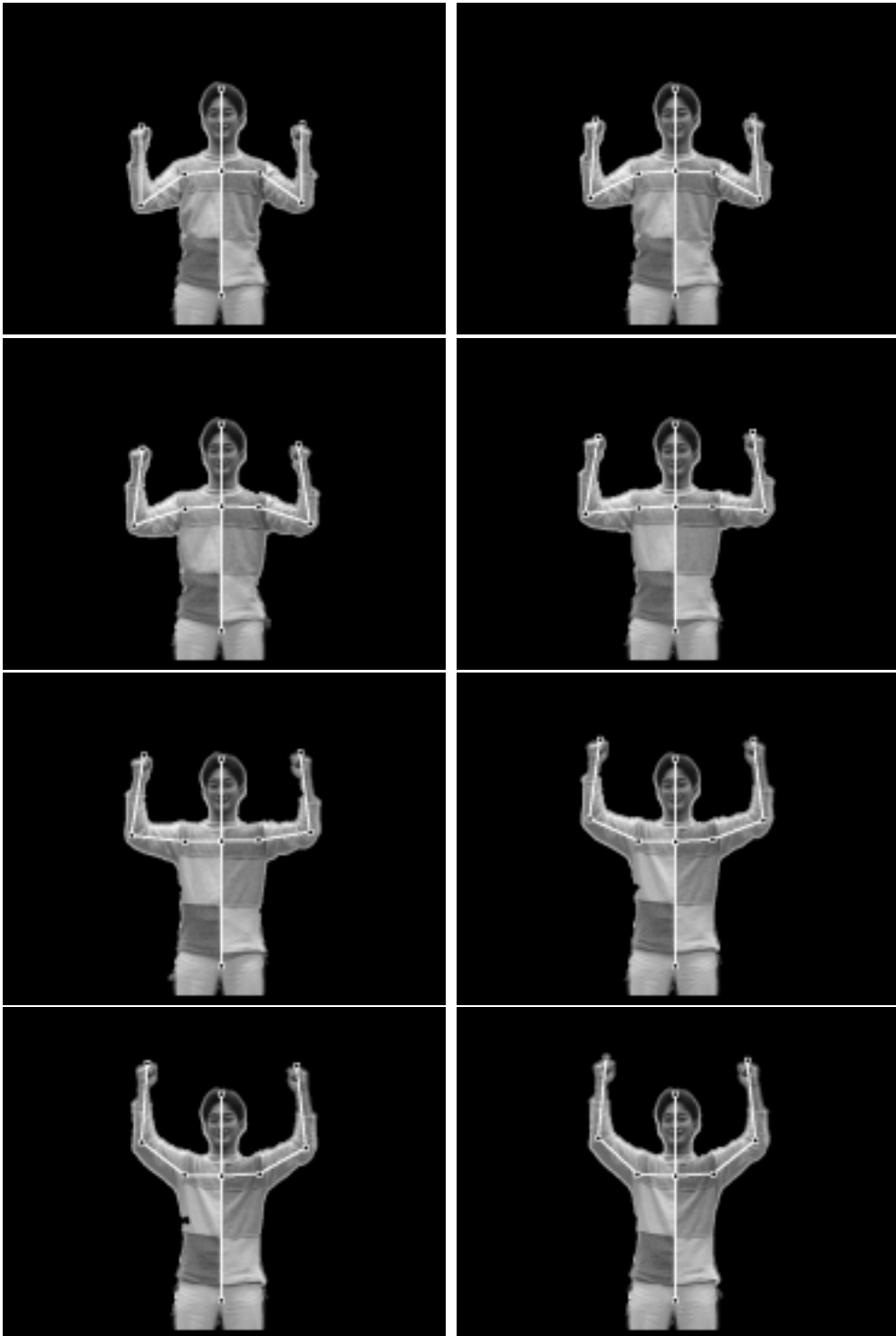


図 5.2: 屈伸運動 2

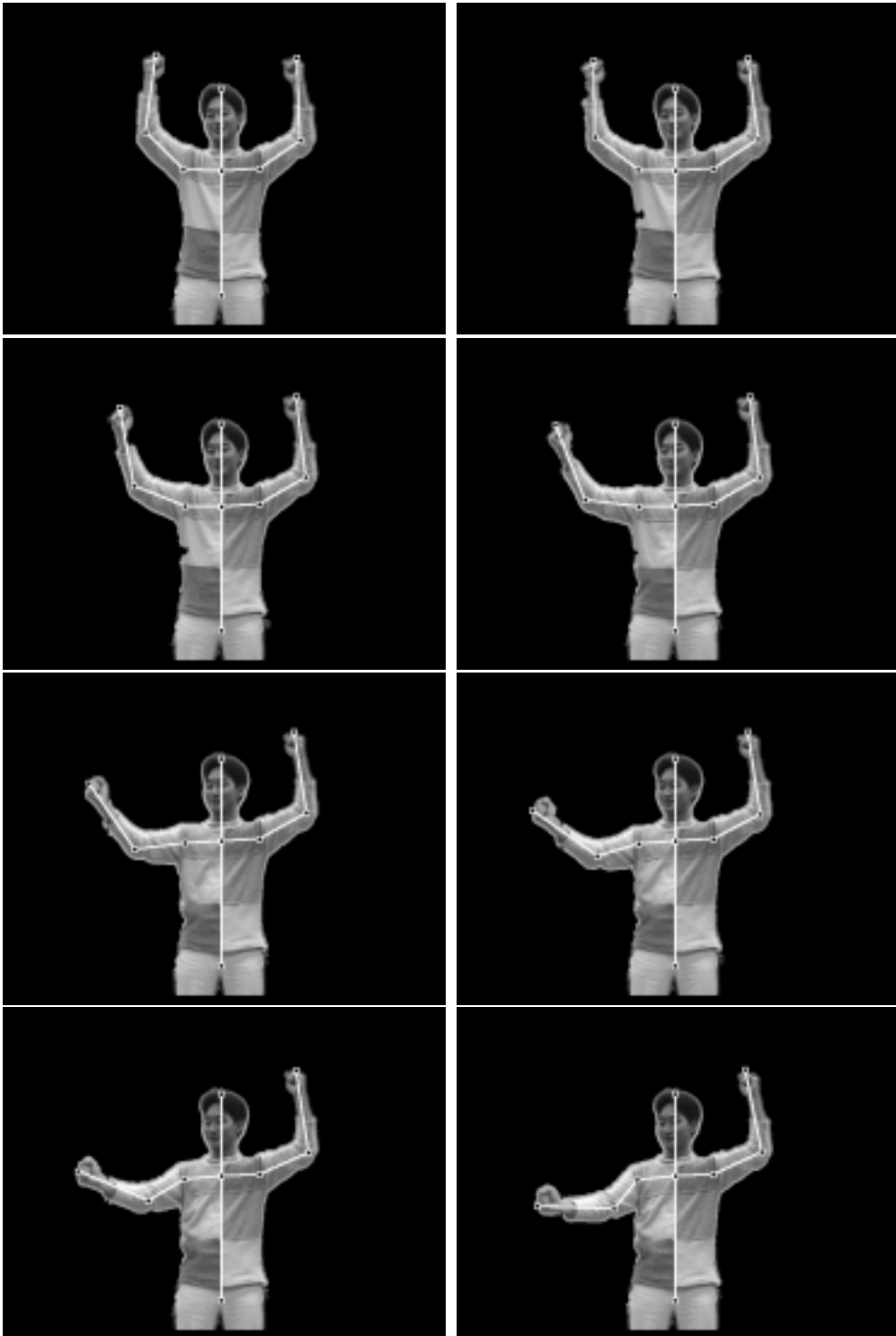


図 5.3: 回転運動

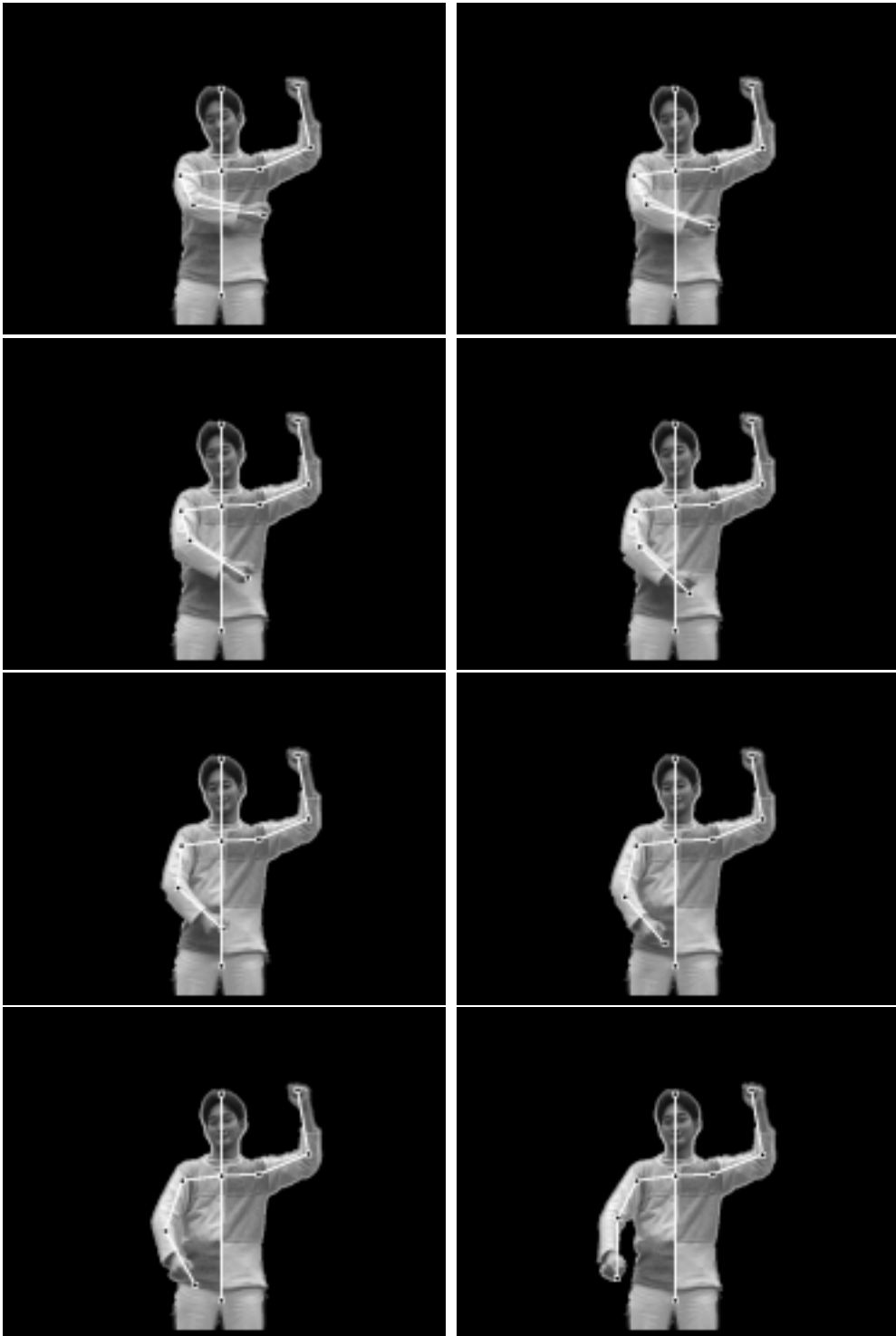


図 5.4: 領域の重なりが生じた場合

第 6 章

結論

人体のような多関節柔軟物体の姿勢や動作を、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、人体は大変複雑な形状をしておりかつ非剛体な変形を伴うため、その動作の画像からの認識には困難がある。そのため、対象物の一部にマーキングをほどこして特徴点を明示したり、特殊な装着物を取りつけてセンサーなどから効率的に情報を得るといった方法を用いることが多い。

しかし、このような方法は被験者に課する負担が多く、適当とはいえない。そこで本研究では、マーキングなど特殊な装置および環境を一切施さない対象物体の姿勢や動きを、画像処理のみから認識することを前提とした。

一方、マーキングなどを施さない方法として、人体動作を対象としてその動画像からの姿勢推定を、シルエット化した画像を細線化し、特徴点情報と簡単な形状モデルとのマッチングをとることにより行うという方法がある。しかし従来手法においては、認識動作に制限があるものや、人体の個々の体節の重なりにおいて姿勢推定が曖昧であるなどの問題点があり、適当な認識手法が確立されていない。

そこで、本研究では人間との親和性を考慮し、画像処理のみにおける人体の姿勢推定法について複雑な動作および体節の重なりが生じる上半身の動作を用いて、従来の問題点である領域の重なりに対する解決法として新しい手法を提案した。

人物領域を個々の領域に分割することで、各体節からの情報を得る領域単位での追跡方法を行った。また、重なった領域における情報として、同一物体上では同じような性質を持つオプティカルフローを活用し、局所情報としての誤差に対しては、領域を持ったモデ

ルを用いることによって大局的な信頼性のある情報とした。

また、従来手法においてはオプティカルフロー法を、移動ベクトルとして考えボトムアップ的にモデルの移動を試みてきた。しかし本手法では、3次元運動におけるオプティカルフローの性質を解析し、またフロー場を用いて領域単位でのモデルの対応づけを行うという正確な方法を行った。

以上、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と言える。

今後の展望としては、本手法は、重なった領域における認識手法としての初期段階として多くの制限を課したが、本手法の有効性は確認された。よって今後これら制約を外していくことで、より一般的で、汎用性のある認識が期待できる。そのためには、 x 軸 (腕方向) 回転も考慮にいれより緻密な解析を行う必要がある。

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。

横田 治夫 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、ひごろよりお世話になった研究室の皆様にも心より感謝致します。

参考文献

- [1] 高橋友一, 岸野文郎, 手振り認識方法とその応用, 信学論 (D-II), vol.J73-D-II, no.12, pp. 1985-1992, 1990.
- [2] 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.
- [3] Barron,Fleet,Beauchemin, Performance of Optical Flow Techniques, PhD Dissertation, Dept. of Computer Science, Carnegie-Mellon Univ. ,1984.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 萩原芳彦, 手形状入力装置を用いた連続した指文字動作の認識に関する研究, 北陸先端科学技術大学院大学修士論文, 1996.
- [8] 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.
- [9] 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.

- [10] 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.
- [11] Haibo Li, Pertti Roivainen, Robert Forchheimer, 3-D Motion Estimation in Model-Based Facial Image Coding, IEEE T-PAMI, pp. 545-555, 1993.
- [12] 谷内田正彦, ロボットビジョン, 東京 昭晃堂, 1990.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [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] クンラボン 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.