

# 画像中の物体および人物領域の 抽出手法に関する研究

平成12年度

名古屋大学大学院工学研究科情報工学専攻

博士課程後期 学位論文

玉木 徹



# 目次

第1章 序論	1
1.1 はじめに	1
1.2 視覚における抽出処理過程	2
1.3 対象物体抽出の応用分野	5
1.4 従来研究とその問題点	6
1.4.1 従来研究における抽出処理	6
1.4.2 従来手法の問題点	7
1.5 本研究の目的とその特徴	8
1.6 本論文の概要	9
第2章 情景理解のための計算機による画像情報処理	11
2.1 領域分割の形式的表現	11
2.2 しきい値処理による二値化	12
2.3 特徴空間におけるクラスタリング	13
2.3.1 $K$ -平均法	13
2.3.2 初期クラスタとクラスタ数の決定	14
2.4 分割併合法	14
2.5 領域分割とパラメータ推定の同時処理	16
2.5.1 EMアルゴリズム	17
2.5.2 混合分布に対するEMアルゴリズム	17
2.6 動的輪郭手法	18
2.6.1 原理	18
2.6.2 変分法による解法	19
2.6.3 差分方程式による実装	20
2.7 移動領域検出	22
2.8 色空間と色差	23
第3章 対象物体に依存しない静止画像からの物体抽出	27
3.1 はじめに	27
3.2 過分割の二つの要因	28
3.3 領域間の情報と併合のための特徴量	29
3.3.1 配置情報	30
3.3.2 隣接情報：境界線の共有部分	31

3.3.3	隣接情報：接続している二つの境界線の非共有部分	33
3.3.4	面積比情報	35
3.4	併合の方略	36
3.4.1	併合に用いる特徴量の順序	36
3.4.2	しきい値の設定 / 更新	37
3.4.3	画像の性質による例外	37
3.5	領域間の特徴量に基づいた併合法	38
3.5.1	初期分割による部分領域生成	38
3.5.2	併合アルゴリズム	38
3.5.3	領域抽出	39
3.6	実画像を用いた実験	40
3.6.1	領域分割過程	40
3.6.2	他の領域分割手法との比較	41
3.6.3	その他の領域抽出結果	41
3.7	特徴量の有効性についての考察	42
3.8	従来手法との比較	52
3.9	おわりに	56
<b>第4章</b>	<b>人物形状の知識を利用した画像系列からの人物領域の抽出</b>	<b>57</b>
4.1	はじめに	57
4.2	差分による移動領域の抽出	58
4.3	エッジの曲線部抽出	59
4.3.1	交点のないエッジの抽出	59
4.3.2	$k$ -曲率による曲線エッジの抽出	59
4.3.3	曲線と直線の分類	62
4.4	曲線エッジの円近似による頭部候補領域決定	63
4.5	人物を覆う閉包の作成	63
4.6	詳細な境界の決定	64
4.7	実験と結果	65
4.8	おわりに	69
<b>第5章</b>	<b>回転モデルに基づいた動画からの人物の腕領域の抽出と運動パラメータ推定</b>	<b>71</b>
5.1	はじめに	71
5.2	Curlに基づく回転領域抽出	73
5.2.1	回転ベクトル場のCurlの性質	73
5.2.2	Curlの例	73
5.2.3	Curlに基づいた領域分割	75
5.2.4	透視投影下のCurl	76
5.3	運動パラメータの推定	76
5.4	EMアルゴリズムを用いた領域分割と推定	77
5.4.1	反復アルゴリズム	78
5.4.2	非回転運動の除去	80
5.4.3	回転面(軸)の決定	80

5.5 実画像を用いた実験結果 . . . . .	81
5.5.1 マニピュレータを用いた推定精度評価 . . . . .	81
5.5.2 実画像系列からの腕領域抽出 . . . . .	88
5.5.3 状況を変化させた場合の実験結果 . . . . .	93
5.5.4 対象を変化させた場合の実験結果 . . . . .	93
5.6 おわりに . . . . .	96
<b>第6章 結論</b>	<b>99</b>
謝辞	101
付録 A 滑らかさの特徴量 $V_s$ の特性	103
付録 B Curlによる領域分割のシミュレーション実験	107
付録 C フローの curl に与える透視投影の影響	111
付録 D 式(5.16) の導出	117
参考文献	119



# 目 次

1.1	さまざまな情景とそこに存在する対象人物（物体）	2
1.2	ルビンの「顔と杯」	3
1.3	ゲシュタルト崩壊の症例の一つ：アフリカ地図	4
2.1	$L^*a^*b^*$ 表色系	24
3.1	部分領域、それを併合した領域、物体に対応する領域	29
3.2	二つの領域の配置	30
3.3	領域が隣接している場合の境界線の共有部分と非共有部分	31
3.4	疑似輪郭	32
3.5	境界線の共有部分	33
3.6	境界線の非共有部分の方向の変化	33
3.7	本の例	34
3.8	偶発的に連続的となる例	34
3.9	T接点での境界線のなす角度	35
3.10	面積比の違い	35
3.11	画像の辺における連続的な接続	38
3.12	重み関数	40
3.13	実画像（男性）を用いた実験結果	43
3.14	実画像（女性）を用いた実験結果	44
3.15	実画像を用いた実験結果	45
3.15	（続き）実画像を用いた実験結果	46
3.15	（続き）実画像を用いた実験結果	47
3.16	併合に用いられた特徴量によって色分けされた領域の境界線	48
3.17	二種類の理想分割との差	49
3.18	図3.14(a)の初期分割での領域数を197とした場合の初期分割、併合結果、理想分割	50
3.19	図3.15(d)の初期分割での領域数を124とした初期分割、併合結果、理想分割	50
3.20	図3.18(b)の併合過程における理想分割との差	51
3.21	図3.19(b)の併合過程における理想分割との差	51
3.22	大津のしきい値による二値化	53
3.23	輝度分布を評価し併合する分割併合法による領域分割	54
3.24	動的輪郭モデルSnakeの収束過程	55
4.1	8近傍隣接画素	60

4.2	点 $s$ における $k$ -曲率 $\varphi$ ( $k = 6$ のとき)	60
4.3	点列 $\{s_i\}$ に対する $\varphi - s$ 曲線	61
4.4	曲線エッジを円弧と仮定した時の扇型	62
4.5	人物像の楕円モデル	64
4.6	snake の収束	65
4.7	詳細な人物領域の決定	65
4.8	人物抽出結果	66
4.9	人物抽出例	67
4.10	人物抽出が失敗した画像	68
4.11	各フレーム毎の抽出結果	69
4.12	抽出失敗数で整理した画像系列としての抽出結果	69
5.1	スティックモデル	74
5.2	マニピュレータを用いた実験	82
5.3	傾き角 $\phi = 0, 10, 20, 30$ [deg] の時の領域分割の結果	83
5.4	傾き角 $\phi = 0, 10, 20, 30$ [deg] の時の領域抽出結果	84
5.5	各領域の $\omega$ と $\xi$ の変化の様子	85
5.6	$\alpha$ - $\gamma$ 空間で表されたマニピュレータの運動パラメータの推定過程	86
5.7	マニピュレータの回転中心の推定結果	87
5.8	実画像に対する実験結果	89
5.9	30~38フレームにおける腕抽出結果	90
5.10	回転中心の推定値の軌跡	91
5.11	角速度と傾き角の推定値の軌跡	92
5.12	二本の腕が運動している場合	94
5.13	背後に移動物体が存在する場合	95
5.14	肘を回転中心とした場合	96
5.15	画像内での領域の大きさを変えた場合	97
A.1	ノイズを付加された濃淡画像のグラデーション	104
A.2	輝度勾配と滑らかさの特徴量のグラフ	105
B.1	curl の値による領域分割	108
B.2	回転面が傾いた場合の curl の値による領域分割	109
C.1	特異点の軌跡	113
C.2	透視投影下の curl : $Z = 20, 80$	114
C.3	透視投影下の curl : $Z = 140, 200$	115
D.1	$x$ 軸方向から見た座標系	118



# 表 目 次

3.1	しきい値の初期値と更新の上限値、更新比率 . . . . .	42
3.2	図 3.20 において各特徴量が併合に使われた回数と併合一回あたりの差の減少量 . . .	52
3.3	図 3.21 において各特徴量が併合に使われた回数と併合一回あたりの差の減少量 . . .	52
5.1	パラメータの符号による運動の分類 . . . . .	81
5.2	回転面の決定 . . . . .	81
5.3	異なる $\phi$ に対する推定結果 . . . . .	88



# 第1章 序論

## 1.1 はじめに

図1.1のような写真を見たとき、複雑なシーンにもかかわらず、人間は容易に被写体が何であるのか、またそれが誰であるのかを推測することができる。たとえ写真に写っている人物を知らないとしても、その状況においてその人物がどのような行動を起こしているかを理解することができる。また、舞台となっている場所を知らなかったとしても、そこが木々に囲まれたどこかのグラウンドであることは想像できるであろう。このような情景理解は、一見すると取るに足らない当たり前のことのように思えてしまうかも知れないが、実際は我々の生活の中できわめて重要な役割を果たしている。例えば、今いる場所からどこかに移動する場合、または何か物をつかんだり、それを誰かに渡したりする時には、自分を取り巻く環境を目で見て正しく認識する必要がある。情景理解は、適切な行動をとるために必要不可欠なものであると言える。

コンピュータビジョンや画像認識は、計算機にカメラを搭載し、人間の視覚に備わっている様々な機能を計算機によって実現しようとする研究分野である。人間の持つ優れた視覚機能を計算機によって実装することができれば、その有用性は計り知れないものとなる。その中でも情景理解は特に有用性が高いものである。しかし、実際にそれを実現するのは簡単なことではない。なぜなら、人間が容易に行っている情景の理解という処理は、計算機で行うためにはいくつかの段階に分けて構築しなければならないが、その一つ一つの段階でさえも、計算機で実現するには困難であるからだ。

例えば、その中の一つに対象を抽出するという処理がある。画像中のどの位置に人物がいるのか、その人物はどのような姿勢をとっているのか、腕の領域と背景の領域の境界線はどこなのか。これらのことは、情景理解を行う上で必要な段階であるが、人間はいともたやすく行っている事であるにもかかわらず、計算機によって対象を安定して抽出することは難しい。

情景理解を困難なものにしている原因の一つに、この物体抽出の難しさが挙げられる。本論文は、この物体抽出の手法について論ずるものである。本研究のアプローチは工学的なものであるが、人間の視覚における処理を知るのは示唆に富むことであるし、また実際視覚の仕組みを取り入れた研究も多い。そこでまず、これまで人間が物を見る仕組みに関



図 1.1: さまざまな情景とそこに存在する対象人物（物体）。

しての研究が行われてきた視覚心理学において示唆されている抽出処理過程を概観し、次いで抽出処理の応用面について述べることにする。

## 1.2 視覚における抽出処理過程

実際に人間が情景を認識する過程で、どのような処理を行なっているのかというはっきりとした心理学的知見はいまだ得られていない[下條 95]。したがって、情景理解を行う際に、対象物体をどのように抽出しているのか、また抽出した対象に対してどのような処理を行っているのかは定かではない。しかし、まったく手がかりが得られていないというわけではなく、断片的ではあるけれども興味あるいくつかの事柄が、視覚心理学において示されている。認知心理学では、視覚情報処理過程を初期、中期、高次の3つのレベルに分けているが、まず最初に、初期の視覚情報処理に目を向けてみよう。

視覚の初期の段階における処理は、全体的並列処理である前注意的処理（前注意過程, preattentive process）と、逐次処理である注意的処理（集中的注意過程, attentive process）に分けられるとしている[乾 95a, 川人 94, 乾 90]。目の前に視覚刺激が提示された時、まず前者が視野内の広い範囲に渡る情報処理を素早く行なう。それに続いて、注意を向けた限られた範囲での処理を、後者が時間をかけて行なう。前注意的処理により即座に処理される視覚刺激として挙げられているのは、円や四角などの形の違い、白や黒、赤や緑などの明るさや色の違い、垂直線と斜め線の違い、平行線と非平行線の違いなどがある。これらの中には、個々の視覚刺激を順番に探索していくというよりは、即座にすべてが目飛び込んでくるような感覚を与えるものがあり、そのような現象はポップアウトと呼ばれている。この初

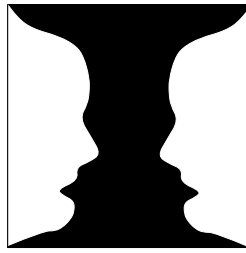


図 1.2: ルビンの「顔と杯」

期視覚における二つの処理過程を説明するものとして、特徴統合理論 [Treisman80, 川人 94] がある。それによると、これらの処理されるべき個々の視覚刺激（平行線や非平行線など）の存在がそれぞれ別の特徴マップに書き込まれ、すべての視覚刺激の位置の情報はマスターマップに書き込まれる。このマスターマップを参照することで、複数の特徴マップの統合が可能になるという。このように、初期視覚においては視覚刺激の抽出やその処理過程が示されている。

線分や色の違いといった低次の視覚刺激ではなく、視野内の領域を抽出する処理は、視覚の中期過程において見出されている。人間がある物体を対象として見た時、その視野内では注目しているその対象物体が中心的な存在である。たとえそれ以外にさまざまな物が存在していたとしても、それらは対象物体に対する背景でしかなく、重要性は格段に低くなる。このような状況は絵画などを見た時にも同様に起こり得るものであり、注目している対象を図 (figure) と呼び、それ以外に存在する背景となるものを地 (ground) と呼ぶ。この現象は認知心理学では図地分離 (figure-ground separation) と呼ばれ、認識すべき対象を背景から分離する働きは、パターン知覚の前提になる重要な処理であるとされている。図 1.2 に示す有名なルビンの「顔と杯」の図形でも分かるように、図と地のどちらを見るかによって図形の持つ意味が変わってくる。図は明確な形を持ち前面に現れて見えるのに対し、地は明確な形を持たずに図の背後まで広がって見える。このような図と地の知覚に対し、どのような形状や特徴を持った物が図や地に見えやすいのか、という視覚刺激要因が視覚心理学の研究で明らかにされてきた。これによると、明るい領域 / 面積の小さい領域 / 閉じた領域 / 規則的な形をした領域 / 空間周波数の高い領域 / 時間周波数の低い領域 / 斜め方向の空間周波数成分を持った領域 / などが、そうでない領域に比べて図になりやすいとされている [川人 94, 乾 95a, 乾 95b]。しかし現在のところ、図地分離の心理学的な知見は、図 1.2 のように実験対象がその実験に特化したもののみ得られている。そのため、与えられた二つ（またはいくつか）の視覚刺激のうち、どちらがより図として見えやすく優勢になるのかを予測することはできても、実際のシーン内のどのような場所が図となりまた地となるのかを、予測することはできない。

高次の視覚情報処理の過程においても、どのようにして対象となる物体を抽出しているのかということははっきりと示されていない。しかし、シーン内の一部分を取り出して処理を

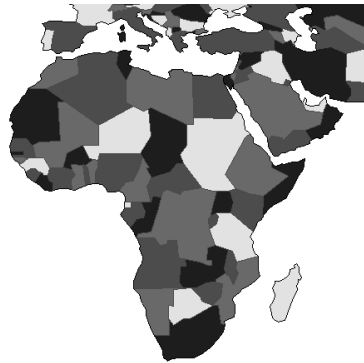


図 1.3: ゲシュタルト崩壊の症例の一つ：アフリカ地図

行なっていることを表す現象として、臨床心理学の分野で知られている視覚失認症 (visual agnosia) [乾93] が挙げられる。これは視覚の認知障害であり、対象が見えているのにも関わらず、それが何であるのかが分からないという症状を指す。このとき、その他の認知機構には別段異常はなく、知能の低下も認められないため、痴呆とは完全に区別される。このような状況は、病気や外傷などによって脳の一部が損傷し、脳内での認知機能が阻害されることによって起こるとされている。視覚失認にはその症例によっていくつかの分類がなされており、その中の一つに、同時失認 (simultanagnosia) と呼ばれるものがある [乾93]。これは、ある状況を絵画や写真などで提示された場合、その細部は正しく認知することができるが、全体としての意味は把握することができないという症状である。例えば、提示された絵画を見た場合、人がいることや怒っている表情をしていること、窓が割れていることなどは理解できるが、全体としてその絵が何を意味しているのか (ボールを投げて窓を割ってしまい、怒られている) ということを認識することができない。

また、ゲシュタルト崩壊 (Gestalzerfall) [乾90] と呼ばれて報告されているものは、全体としての認知を保っていることができず、すぐに部分の認識しかできなくなってしまうという症例である。例えばトラックを見た時に、最初はそれをトラックとして認識することができる。しかしそのまま注視していると、すぐにモーターや車台や運転室など、それぞれの部分しか認識することができなくなってしまう、一旦他方を見てから再び対象を見なければ、もう一度全体としてまとまって見えることはない。色分けされた世界地図のアフリカ (図1.3) をみても、アフリカと認識できるのは最初だけで、すぐにただの色の集まりにしか見えなくなるという。

これらの例は、全体と部分の位置関係というものが認知過程の中で扱われており [乾90]、その際に全体の中における部分というものが取り出されて処理されていることを示唆している。ここで部分という言葉が指すものが、すなわち情景認識における対象物体であると言いきれるものではない。しかし、部分として示されているものが、点や線のような低次の視覚刺激にまで落ちることなく、一つの意味あるまとまりの最小単位として認識されている点は興味深い。

## 1.3 対象物体抽出の応用分野

前節で見たように、人間の視覚がどのように処理を行っているのかは明確には分からない。そのため、人間のするように理解することを目指すのは難しいかもしれない。しかし計算機による情景理解は、必ずしもそれと一致していなくとも、システムとして用いることができる。そのためにも、画像の中の対象を抽出する領域抽出や、画像をいくつかの対象に切り分ける領域分割は、情景を解析する上で必要不可欠な処理である。また、これまでは画像認識や情景理解の一部としての側面が大きかったが、近年は以下のような様々な方面から応用が期待されており、重要な役割を担う処理となっている。

- 領域ベース符号化のための対象物抽出・領域分割

より低ビットレートの画像符号化を行なうために、動画像符号化の国際標準としてMPEG4[渡邊98, 甲藤98]が制定された。この中で、画像をオブジェクト単位の領域に分割し、それぞれの領域で符号化を行なう、オブジェクトプレーンと任意形状符号化の概念が取り入れられ、さまざまな符号化手法が研究されている[齊藤00, 鹿喰97]。また対象物体の領域とそれ以外の領域を選択的に符号化するような手法[奥宮96]も提案されている。しかし、MPEG4の仕様やこれらの研究では、画像中の対象物体はあらかじめ切り出されていることが前提となっており、画像内の物体を抽出することができなければ効率的に符号化することは困難である。

また動画像符号化においては、三次元空間での物体という単位ではなく、画像上でまとまった動きを持つ領域単位での記述が重要な役割を持っていることが言われている[栄藤97, 栄藤98]。そのために、異なる運動をする領域に動画像を分割する研究が盛んに行なわれている。

- 画像データベースのインデックス、検索キー作成の自動化

通常のデータベースでは検索キーとして単語や文章が用いられるが、画像データベースでは蓄積してある膨大な量の画像に対して、十分対応できるインデックスをあらかじめ作成しておかなければならない。画像に写されている物体を抽出し認識できれば、効率的にインデックスを作成することができる。また、特定の対象物体が写ったシーンを検索したいというような場合には、検索キーとしてテキストではなく画像や映像パターンそのものを検索キーとすることが必要になってくる[Forsyth97]。

- 動作認識処理のための人物領域の抽出

ロボットへの指示や操作などを身振りで行なったり、計算機とのインターフェイスとしてジェスチャというノンバーバルな表現を用いるマルチモーダルなシステムが研究されている[岡99, 末永95]。現在のジェスチャーや動作の認識の研究では抽出され

た動きが何を意味するのかを認識することに主眼がおかれており、抽出処理を簡単にするために撮影環境をセッティングしなければならない。そこで、認識の前処理として画像中の人物を切り出すことができれば、さまざまな環境で撮影された画像に対して認識が行なえることになる。

- セキュリティ監視システム

近年、セキュリティの監視システムにおいて、カメラによる監視システムの実現が期待されている [増田 95, 増田 97, 佐藤 98, 山口 99]。赤外線センサや振動センサを用いるような従来のシステムでは、侵入者が入ってきたことを検知するだけである。それだけではなく、何が起こったのか、誰が侵入してきたのか、などを知ることは、セキュリティの重要な要素である。このとき、移動物体を追跡するようなカメラシステム [Takeuchi94] は有用な手段である。それだけでなく、捜査時の資料とするために、カメラに写った人物をズームアップして大きく拡大して撮影し、対象人物の人相や体型を精細に捉えることが望まれている [吉野 99]。

## 1.4 従来研究とその問題点

### 1.4.1 従来研究における抽出処理

物体抽出の研究は従来から多く行われているが、コンピュータビジョンや画像認識の応用の一部として行われることが多く、一般的な手法の開発よりも利用目的に応じた個別の手法の開発の方が盛んである。

まず動画像の中から移動する物体を抽出する手法として広く用いられているものに、クロマキー手法がある [中嶋 98a, 中嶋 97]。撮影環境を整えることができる場合に、背景を一律な色 (例えば黒一色など) にして、その色以外を抽出する手法である。背景を青一色にするいわゆるブルーバック手法以外にも、最近は特殊な反射特性を持つ布を背景にして、カメラに取り付ける照明用 LED を変化させる手法がテレビ番組制作現場などで実用化されている [川本 00]。

屋外など撮影環境を整えることができない場合に用いられる手法に背景差分がある。これは、移動物体が存在しない時に撮影された背景 (参照) 画像と現在の画像との差分を取る方法であり、処理が単純で扱いやすい。ただし屋外では、天候や時刻の違いによって照明が大きく変動することがあり、参照画像と現在の画像との差分だけでは、移動物体のみを抽出することが困難になるという問題点がある。そのため、参照画像を動的に更新していく手法 [松下 98, 波部 98] が提案されている。背景差分は、照明条件の変化しにくい屋内で用いる場合には扱いやすく、ステレオと組み合わせた手法も提案されている [石橋 91]。

もう一つの移動物体を抽出する手法として、連続するフレームでの差分を取る方法がある。このフレーム間差分の手法は、参照画像のように日光などの照明に大きく影響される



ことはなく、広く用いられている。ただし、前後二枚の画像だけを見て安定に検出することは難しく、多数のフレームを観察し画素値の変化を検出する手法も多く用いられている [Haritaoglu98, Khan00]。

静止画像から物体を抽出する手法は、動画像からの場合と違い、動きの情報を使うことができない。そのため、人間が大まかな輪郭を与えて抽出する手法 [淡96] や、誘目度を評価して観測者が注目する物体を抽出する手法 [田中98a, 田中98b] が提案されている。形の決まった工業部品を抽出したり [谷内田90b, 谷内田90a]、写真の中から人間の顔の位置を検出したり [Rowley95] するなど、個々の特定の物体の抽出については研究され、産業にも応用されている。

### 1.4.2 従来手法の問題点

前述した研究では、従来からそれぞれの処理の一部として物体を抽出しているが、それ以降の処理に重点が置かれることが多い。そのため、抽出処理が容易に行えるように、主に以下のような前提条件の下で行なわれている。

- 対象となる被写体の画像内での位置や、大まかな輪郭が与えられている
- 対象となる被写体を抽出しやすくするために
  - 単純な背景に限定している
  - シーン内に他の物体が存在しないことを仮定している

これらの仮定は、環境に対する適切な条件を設け、手法の適用範囲を明確にし、また予測しやすくするためには必要なことである。

しかしその一方で、多くの応用分野において実用的な場面で使用するためには、それらは取り払わなければならない課題でもある。容易に変えることのできない環境で処理が行われることも考えられるし、また実際にどのような環境でシステムが使用されるのかをすべて予測することもできない。さらに、実環境では時々刻々と環境そのものが変化する場合もある。例えば、カメラから得た画像情報を基に自律的に行動するロボット [増田00, 人間98] として研究が進められているものに、自動走行車両やサービスロボット [Waldherr00, Thrun00, Bischoff00, 松井99] などがある。それらは保持しているマップの中を動きまわるが、その既知の環境の中であっても、物が動かされて障害物となったり、共に協力したり支援すべき多数の人間たちが、徒歩だけでなく時に車や自転車などで絶えず動き回っているため、ロボットをとりまく環境は変わっていく。また画像認識を用いたインターフェイスが広く一般に用いられるようになると、現在実用化されているような研究所への入退室用だけでなく、駅構内や街角、店内の隅や屋上などにも認識用にカメラが設置されるかもしれない。

したがって、撮影環境を限定せず、多数の移動物体が存在する状況で人物を認識する必要も出てくるであろうし、位置や輪郭を与えることなく一般的な対象物を抽出することも重要である。

## 1.5 本研究の目的とその特徴

本論文では、画像から対象物体や対象人物を抽出するという処理を、未知の複雑な背景において他の移動物体が存在するという、より実的な状況において用いようとする際に問題となる点を明らかにし、それを克服する手法を提案する。具体的には以下の通りである。

第一に、不特定の物体を静止画像から抽出する方法について考察する。写真のような静止画の被写体となる対象は、人物から花や風景などの静物まで幅広い。それゆえ、様々な物体を扱うには、物体についての固有の知識を用いずに抽出できることが望ましい。そのための一つの方法として、画像の領域分割を用いるアプローチが考えられるが、従来の領域分割では、一つの物体に対して一つの領域を抽出することは困難である。まず初めに、その困難さの原因である二つの要因 — 光学的要因と物理的要因 — について述べる（第3章参照）。この要因によって、本来は一つの領域になるべき物体が多数の領域に分割されてしまう。従来の研究では光学的要因を考慮する手法はあったが、物理的要因は考慮されなかった。本研究の特徴は、物体に固有でない一般的な情報を用いて、光学的要因だけでなく、物理的要因も排除することで、物体を一つの領域として抽出することである。

第二に、対象を人間とした場合に、より実的な状況において、人物の形状モデルを用いた人物領域の抽出について検討する。様々な目的のために人物を被写体にすることは工学的な応用分野で多く行われているため、他の移動物体が存在する状況においても人物を抽出することが必要である。従来の移動物体の検出手法の多くは、シーン内の動物体の数を一つに限定している。また複数動物体の研究でも、全て同じ種類の物体、例えば全て人物で、それぞれの人物を抽出しているに留まっている。しかし実環境では、人物以外に車や動物などが存在することもあり、複数の異なる種類の移動物体から人物以外を排除することが必要である。本研究の特徴は、人や車などの複数の移動物体が存在するような動画像からでも、人物の簡単な形状モデルを用いることで人物の領域を抽出することである。

第三に、対象を人間の腕とした場合に、より実的な状況において、腕の運動モデルを用いた人物の腕領域を抽出する手法を提案する。ジェスチャなどでは情報を伝えるために腕の動作が頻繁に用いられるため、人物の腕の抽出は、動作認識において重要な課題である。しかし従来の動作認識手法では、未知の複雑な背景において、他の移動物体が存在するという、より実的な状況での抽出は考慮されなかった。本研究の特徴は、そのような条件の下であっても、腕の運動モデルを用いることで、回転する人物の腕の領域を抽出することである。また、それと同時に、認識の際に有用な情報となる腕の運動パラメータを推定し、それを基に他の物体を排除することができる。

## 1.6 本論文の概要

本論文は6章から構成されている。

2章では、物体抽出処理と密接な関係にある領域分割の手法についてまとめる。

3章では、物体についての固有の知識を用いずに、静止画像から物体を抽出する方法について述べる。静止画に対する従来の領域分割では、一つの物体に対して一つの領域を抽出することは困難である。まず初めに、その困難さの原因である二つの要因 — 光学的要因と物理的要因 — について述べ、物体に固有ではない一般的な情報を用いて物理的要因及び光学的要因を排除し、物体を一つの領域として抽出する領域分割手法を提案する。この手法は分割併合法に基づいており、まず画像を一つの特徴量で表される部分領域に分割する。そして、領域内の物体の領域が物理的要因によって分割されたのかどうかを判定するために、二つの領域間の配置・隣接・面積比の情報を導入し、それに基づいて併合を行うための四つの特徴量を提案する。それらの特徴量は、一つの領域の内部についてのものではなく、二つの領域間についての特徴量である。これによって、物体表面が別々の色を持っているというような、従来の領域分割では扱えないような場合であっても、一つの物体の領域を得ることができる。

4章では、人や車などの複数の移動物体が存在する動画像から、簡単な人物の形状モデルを用いて人物全身の領域を抽出する方法について述べる。この章で提案する手法は、まずフレーム間差分を用いて移動する人物領域候補の抽出を行う。次に、背景に人物以外の動物体がある場合でも人物領域を得るために、エッジから人物の頭部とみなせる部分を特定する。そしてその頭部位置を元に、抽出される人物像についての楕円モデルを作成し、動的輪郭手法により人物領域に収束させる。そして最後に色による領域分割を行い、詳細な人物の輪郭を抽出する。

5章では、複数の運動物体が存在する動画像から、回転運動モデルを用いて人物の腕の領域を抽出する方法について述べる。この章では、未知の複雑な背景において、他の移動物体が存在するという条件の下で、回転する人物の腕の領域抽出と、その運動パラメータの推定を行う手法を提案する。まず動画像からオプティカルフローを求め、その方向によって大まかに分割を行う。そして画像中の複数の運動物体を扱うためにEM (Expectation Maximization) アルゴリズムを導入し、反復計算によって異なる運動物体の領域に分割し、同時にその運動パラメータを求める。最後に求められたパラメータから回転物体であるかどうかを判定する条件を示し、他の物体を排除する。

最後に6章で本論文をまとめるとともに、今後の課題について検討する。



## 第2章 情景理解のための 計算機による画像情報処理

計算機によって画像から対象物体の領域を切り出す手法は、従来からさまざまな方法が開発されてきた。それらは大別して、対象物体の輪郭線のエッジを線として抽出するエッジ抽出と、画像を対象物体に対応した領域に分割する領域分割に分けられる[江91]が、この章では、物体抽出と密接な関係にある領域分割手法について整理する。

領域分割は画像認識において基礎的かつ重要な処理であり、これまでにさまざまな手法が提案されている[Gonzalez77, Kanade80, Ballard82, Nev86, Shirai87, Gonzalez92, Pal93, Horn93, Skarbek94, Jain95]。領域分割とは、画像を幾つかの重なりのない領域に分けることである。ここで領域とは、一般的には色やテクスチャなどが同じ画素の集合を指す。領域分割の目的は、画像を赤い領域や緑の領域に、または壁や床などの領域に分割することであるので、後者のように、一つの物体に対して一つの領域を形成することは、容易ではないが非常に有用である。実際、多くの画像解析や認識の手法は、画像中の一つの対象ごとに適用されるものであるため、後続の処理を行なう前に別々の対象の領域に分割しておくことが必要になる[Horn93]。

### 2.1 領域分割の形式的表現

領域分割を形式的に表すことはさまざまに行なわれているが、述語論理 (predicate logic) を用いることが多い[Gurari82, Vernon91, Gonzalez92, Pal93, Jain95]。 $R_i$  を画像中の  $i$  番目の領域とし、 $P$  を predicate (たとえば色の分布が一定以内である、という命題) とすると、 $P(R_i) = TRUE$  であれば、ある色の画素の集合が領域  $R_i$  となる。つまり、その predicate を満たすかどうかを領域の基準とするものである。

領域分割処理によって生成された領域  $R_1, \dots, R_N$  が満たすべき条件としては、例えば次のように定式化できる[Gonzalez92]。

- (a)  $\bigcup_{i=1}^N R_i = R$  (すべての領域の和は画像全体である)
- (b)  $R_i$  は連結した領域 (画像上で分裂していない)

- (c)  $R_i \cap R_j = \phi$  ( $i \neq j$ ) (領域同士の重なりはない)
- (d)  $P(R_i) = TRUE$  (領域は predicate を満たす)
- (e)  $P(R_i \cup R_j) = FALSE$  ( $i \neq j$ ) (異なる領域の和は predicate を満たさない)

これらは、形式的に領域分割を取り扱うという点においては重要な意味を持つが、この定義から実際の領域分割処理が導けるわけではない。

## 2.2 しきい値処理による二値化

画像中に二つの領域、例えば白と黒の領域しかない場合には、輝度値をしきい値で二値化することによって、白と黒の領域に分割することができる。しきい値を輝度値のヒストグラムから決定する手法として、分散比を最大にする設定法 [谷口 96] (大津のしきい値 [大津 80]) がある。

ヒストグラムをあるしきい値  $T$  で二つのクラス  $C_1, C_2$  に分割することを考える。このそれぞれのクラスの平均を  $\mu_1, \mu_2$ 、分散を  $\sigma_1^2, \sigma_2^2$  とする。またそれぞれのクラスの生起確率 (クラスに属する画素数を全画素数で割った正規化画素数) を  $\omega_1, \omega_2$  とする。

ここで  $\mu$  を全画素についての平均とすると、次の式によって与えられる  $\sigma_W^2(T), \sigma_B^2(T)$  をクラス内分散 (within-class variance)、クラス間分散 (between-class variance) という [石井 98]。

$$\sigma_W^2(T) = \omega_1 \sigma_1^2 + \omega_2 \sigma_2^2 \quad (2.1)$$

$$\sigma_B^2(T) = \omega_1 (\mu_1 - \mu)^2 + \omega_2 (\mu_2 - \mu)^2 \quad (2.2)$$

クラス内分散  $\sigma_W^2(T)$  はそれぞれのクラスの分散の重みつき和であり、それぞれのクラスがそれだけ密集しているかを表す。クラス間分散  $\sigma_B^2(T)$  は、全体の平均から、それぞれのクラスの平均がどれだけ離れているかを表し、二つのクラスタ間の広がりを示す。

これらを用いて、クラス分離度

$$\eta(T) = \frac{\sigma_B^2(T)}{\sigma_W^2(T)} \quad (2.3)$$

を定義することができる。この分離度が大きくなるほど、二つのクラスタがはっきり分かれていることになるので、 $\eta(T)$  を最大にするしきい値

$$T^* = \operatorname{argmax}_T \eta(T) \quad (2.4)$$

を、ヒストグラムの区間内から決定する。この手法を用いた二値化処理の例は 3.8 節を参照のこと。

この手法を多数のクラスタについての分離度に拡張した手法も提案されている [福井 97]。

## 2.3 特徴空間におけるクラスタリング

濃淡画像におけるしきい値処理を拡張してカラー画像のそれぞれの色や色相などのヒストグラムをしきい値処理する方法もある [Shirai87, 谷内田 90b] が、複数の特徴量を一度に扱う手法として、クラスタリングがある。これは、各画素の特徴量をサンプルとする特徴量空間において、サンプルの分布が密になっている部分（クラスタ）を検出し、各サンプルをいずれかのクラスタに分類するものである。クラスタリングのアルゴリズムにはさまざまな手法がある [画像 97] が、ここでは  $K$ -平均法について述べる。4.6節と 3.5.1節で行う（初期）領域分割には、この  $K$ -平均法を用いている。

### 2.3.1 $K$ -平均法

$K$ -平均法 [安居院 92] は、クラスタ数が  $K$  個に決定しているときのクラスタリングアルゴリズムである。クラスタリングの対象となる与えられたサンプルの集合を  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  とする。ここで  $\mathbf{x}_i \in \mathcal{R}^n$ 、つまり各サンプルは  $n$  次元実ベクトルである。

1.  $K$  個のクラスタ  $\{C_{j(1)}\}$  ( $j = 1, \dots, K$ ) について、 $X$  中の任意のサンプルを重ならないように選んで、初期クラスタのクラスタ中心  $\mathbf{c}_{j(1)}$  とする。

$$\mathbf{c}_{j(1)} \in X, \quad \mathbf{c}_{j(1)} \neq \mathbf{c}_{i(1)} \quad (i \neq j) \quad (2.5)$$

2. 以下の処理を、 $t = 1, 2, \dots$  に対して反復する。

- (a) クラスタ中心  $\mathbf{c}_{1(t)}, \mathbf{c}_{2(t)}, \dots, \mathbf{c}_{K(t)}$  を用いて、各サンプル  $\mathbf{x}_i$  を、それに最も近いクラスタ中心をもつクラスタに所属させる。すなわち、サンプル  $\mathbf{x}_i$  とクラスタ  $C_{j(t)}$  との距離を  $d(\mathbf{x}_i, C_{j(t)})$  とすると、

$$J = \operatorname{argmin}_{1 \leq j \leq K} d(\mathbf{x}_i, C_{j(t)}) \quad (2.6)$$

$$\mathbf{x}_i \in C_{J(t+1)} \quad (2.7)$$

ここで距離をユークリッド距離にとれば、

$$d(\mathbf{x}_i, C_{j(t)}) = \|\mathbf{x}_i - \mathbf{c}_{j(t)}\| \quad (2.8)$$

である。

- (b) 新しいクラスタ中心  $\mathbf{c}_{j(t+1)}$  ( $j = 1, 2, \dots, K$ ) を、(a) で  $C_{j(t+1)}$  に所属したサンプルの平均値に定める。

$$\mathbf{c}_{j(t+1)} = \frac{1}{|C_{j(t+1)}|} \sum_{\mathbf{x}_i \in C_{j(t+1)}} \mathbf{x}_i \quad (2.9)$$

ここで  $|C_{j(t+1)}|$  は  $C_{j(t+1)}$  に属するサンプルの個数である。

(c) クラスタ中心が一つも変化しない、すなわち、

$$\forall j; \mathbf{c}_{j(t+1)} = \mathbf{c}_{j(t)} \quad (2.10)$$

ならば終了。このとき

$$\mathbf{X} = \bigcup_{j=1}^K C_{j(t)} \quad (2.11)$$

$$C_{j(t)} \cap C_{i(t)} = \phi \quad (i \neq j) \quad (2.12)$$

である  $\{C_{j(t)}\}$  を最終的なクラスタリングとする。

そうでないときは  $t \leftarrow t+1$  としてステップ2.(a)へ。

### 2.3.2 初期クラスタとクラスタ数の決定

$K$ -平均法はクラスタ数が  $K$  個に固定されているが、どのようにクラスタリングされるかは、クラスタ数と初期クラスタによって変化する [中谷 93, 境田 98]。そのため、ここではクラスタ内分散の平均が最小になるように  $K$  の値を決定する [原 95] 手法について述べる。

$$A(K) = \log V_{mean} + \frac{2}{K} \quad (2.13)$$

$$V_{mean} = \frac{1}{K} \sum_j V_j \quad (2.14)$$

$$V_j = \frac{1}{|C_j|} \sum_{\mathbf{x}_i \in C_j} \|\mathbf{c}_j - \mathbf{x}_i\|^2 \quad (2.15)$$

ここで  $V_{mean}$  は各クラスタの分散  $V_j$  の、全クラスタについての平均を表している。そして  $A(K)$  を最小にする  $K$  である

$$K_{min} = \underset{K}{\operatorname{argmin}} A(K) \quad (2.16)$$

を最適クラスタ数とする。

まず適切なクラスタ数と初期クラスタを求めるために、予備クラスタリングを行なう。1/10に縮小した画像に対してランダムに初期クラスタを設定し、 $K$ -平均法によりクラスタリングを行ない、クラスタ中心の移動がなくなるまで処理を反復する。この予備クラスタリングを  $K = 1, 2, \dots$  について行ない、式(2.16)によって  $K_{min}$  を求める。

このクラスタ数  $K_{min}$  と、その予備クラスタリング終了時に形成されたクラスタを初期クラスタとして、同様に  $K$ -平均法によるクラスタリングを行ない、領域分割画像を得る。

## 2.4 分割併合法

特徴空間におけるサンプルの分布により分割するのではなく、各画素の特徴とその近傍の画素の特徴により同じ領域に属するかどうかを判定して、特徴が似ていれば（通常はし



きい値で判定し)それらを併合して、ひとまわり大きい領域を生成する。これを繰り返す、領域を成長させていく手法が領域成長(拡張)法(region growing)である。

領域成長法は種(seed)と呼ばれる一画素の領域から出発するが、他の手法によって多数の領域を生成しておき、それらの特徴量が似ているかどうかを判定し、併合していく手法もある。これは領域併合法と呼ばれ、前述の領域分割法と合わせて分割併合法(split-and-merge)として多数の手法が提案されている。3章で提案する手法もこの分割併合法の一種である。

併合のために用いられる特徴量には領域内の画素の輝度や色の平均値[Ballard82, Horn93]が一般的であるが、領域内の輝度の分布を直接用いて平均分散から検定を行なう手法[Jain95]や、分布と境界線を符号化したときの最小符号長(MDL, Minimum Description Length)[Kanungo94, Zhu96]なども提案されている。また幾何的な併合条件として、領域の境界での明度差は大きくはつきりしているのか[Castleman96]、さらには境界の全周囲長に対する共通境界の長さの比[Brice70, Nev86, Horn93]といったものもある。

ここでは輝度分布に基づく検定による併合基準[Jain95]について述べる。この手法は、領域内の輝度の分布に正規分布を仮定して、二つの隣接した領域の輝度分布の尤度と、それらを併合した一つの領域の輝度分布の尤度の比がしきい値を下回る場合に併合を行う。ここで、隣接する二つの領域を  $R_1, R_2$  とし、それぞれの画素数を  $m_1, m_2$  とする。画素は  $R_1, R_2$  に属する順に1から  $m_1 + m_2$  まで番号づけされていて、 $i$  番目の画素の画素値を  $I_i$ 、その画素値を取り得る確率を  $P(I_i)$  とすると、その画素が属する領域の画素値が正規分布  $N(\mu, \sigma^2)$  に従うならば、

$$P(I_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(I_i - \mu)^2}{2\sigma^2}\right\} \quad (2.17)$$

である。

仮説(Hypothesis)  $H_0$  を、この二つの領域が同じ領域であるとする。その領域の輝度の分布を  $N(\mu_0, \sigma_0^2)$  とすると、 $H_0$  の下で画素値が  $I_1, \dots, I_{m_1+m_2}$  を取り得る同時確率は以下のようになる。

$$P(I_1, \dots, I_{m_1+m_2} | H_0) = \prod_i^{m_1+m_2} P(I_i | H_0) \quad (2.18)$$

$$= \frac{1}{(\sqrt{2\pi}\sigma_0)^{m_1+m_2}} \exp\left(-\frac{m_1 + m_2}{2}\right) \quad (2.19)$$

ここで、各画素でとる画素値は独立であると仮定する。また  $\mu_0, \sigma_0^2$  はすべての点について求めた輝度値の平均と分散である。

また対立仮説  $H_1$  を、この二つの領域は別々の領域であるとする。二つの領域の輝度の分布を  $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$  とすると、画素値の独立性を仮定し、同時確率は以下のようになる。

$$P(I_1, \dots, I_{m_1}, I_{m_1+1}, \dots, I_{m_1+m_2} | H_1) \quad (2.20)$$

$$\begin{aligned}
&= \prod_{i=1}^{m_1} P(I_i|H_1) \prod_{i=m_1+1}^{m_2} P(I_i|H_1) \\
&= \frac{1}{(\sqrt{2\pi}\sigma_1)^{m_1}} \exp\left(-\frac{m_1}{2}\right) \frac{1}{(\sqrt{2\pi}\sigma_2)^{m_2}} \exp\left(-\frac{m_2}{2}\right) \quad (2.21)
\end{aligned}$$

ここで  $\mu_1, \sigma_1^2$  は領域  $R_1$  内の点について、 $\mu_2, \sigma_2^2$  は領域  $R_2$  内の点について求めた輝度値の平均と分散である。

この二つの仮説の尤度比  $L$  は次のようになる。

$$\begin{aligned}
L &= \frac{P(I_1, \dots, I_{m_1+m_2}|H_1)}{P(I_1, \dots, I_{m_1+m_2}|H_0)} \\
&= \frac{\sigma_0^{m_1+m_2}}{\sigma_1^{m_1} \sigma_2^{m_2}} \quad (2.22)
\end{aligned}$$

$$\log L = (m_1 + m_2)\sigma_0 - m_1\sigma_1 - m_2\sigma_2 \quad (2.23)$$

この  $L$  (つまり  $\log L$ ) が大きいほど  $H_0$  が、逆に小さいほど  $H_1$  が採用されやすくなる。したがって、対数尤度比  $\log L$  がしきい値を下回った場合に仮説  $H_0$  を採用し、その二つの領域を併合する。この手法を用いた領域併合の例は3.8節を参照のこと。

## 2.5 領域分割とパラメータ推定の同時処理

領域分割として、異なるパラメータをもつ領域毎に分割するという方法もある。パラメータとしては、任意曲線を直線で分割する際にはその傾きであったり、画像中の異なる移動物体を抽出するにはその運動パラメータであったりする。この方法は、与えられたデータを、パラメータを共有するデータの集合に分割するが、特徴量そのものの差を用いて分割するクラスタリングでは取り扱えない場合に用いられることが多い。

例えば、異なる移動物体に分ける場合を考えると、同じ物体上に存在する点の運動は同じ運動パラメータから生成されている。しかし、それらの点の画像上での動き（オプティカルフロー）は異なっているので、動きの方向や大きさを特徴量としてクラスタリングを行なうと異なる領域に分割されてしまう。そこで、ある範囲の領域から計算される運動パラメータによって異なる領域かどうかを判定することになる。しかし、ここで問題となるのは、そのパラメータを推定することと領域を分割することとが、相互に依存していることである。すなわち、パラメータ推定には分割された領域が必要となるが、領域分割には推定されたパラメータが要求されるのである。

この問題を克服するためにはいくつかの手段が提案されているが、ここではEM (Expectation-Maximization) アルゴリズム [Dempster77, 栗田98] について述べる。EM アルゴリズムは反復法的一种で、初期値として大まかな領域分割を行い、それらの領域に対してパラメータを推定し、再び領域分割を行うという反復処理をすることで相互依存の問題を解決するものである。5章で提案する手法はこのEM アルゴリズムを用いている。

### 2.5.1 EM アルゴリズム

EM アルゴリズムは、観測できない変数（隠れ変数）を含む確率分布に対する最尤推定を行なう方法である [大羽 99]。ここで、サンプルを生成する確率変数が多数ある場合には、各サンプルがどの分布から生成されたのか、またはどの分布に属するのかを観測できない変数とみなすことができる。

観測される変数を  $x$ 、観測できない変数を  $z$  とする。EM アルゴリズムは、あるパラメータ  $\theta$  についての分布  $p$  の尤度を最大にするために、反復  $t = 0$  での初期値  $\theta^t$  から始め、以下の二つのステップを反復する [栗田 98]。

- E step

$x, z$  の対数尤度  $\log p(x, z; \theta)$  の条件つき期待値

$$Q(\theta|\theta^t) = E[\log p(x, z; \theta)|x, \theta^t] \quad (2.24)$$

を計算する。

- M step

$\theta^t$  が与えられている時に、条件つき期待値を最大にするパラメータを

$$\theta^{t+1} = \operatorname{argmax}_{\theta} Q(\theta|\theta^t) \quad (2.25)$$

として求める。

この二つのステップを反復することで、尤度を単調増加させることができる。ただし局所最適解に陥ることが多いため、次に述べるような混合モデルの場合にモデルを分割併合して尤度を増加させる方法 [上田 99] などが提案されている。

### 2.5.2 混合分布に対する EM アルゴリズム

複数の対象がそれぞれ確率分布で表されており、複数の対象全体としては、それらが重なりあった確率密度関数で表されているような場合、これを混合分布と呼ぶ [石井 98]。ここでは、ある一つの対象  $i$  の確率分布が平均  $\mu_i$ 、分散  $\sigma_i^2$  の正規分布  $\phi(x; \mu_i, \sigma_i^2)$  であり、すべての対象についての全体の確率分布がその重みつき線形和

$$p(x; \theta) = \sum_i \xi_i \phi(x; \mu_i, \sigma_i^2) \quad (2.26)$$

で表される正規混合分布を考える。ここで  $\theta = \{\mu_i, \sigma_i^2\}$  である。

これに対して EM アルゴリズムを適用する場合には、反復するステップを以下のように書くことができる [栗田 98]。

(1) 各サンプル  $x_j$  に対する、 $z$  の条件つき確率を計算する。

$$w_{zj}^{(t)} \equiv p(z|x_j; \theta^t) = \frac{\xi_z^{(t)} \phi(x_j; \mu_z^{(t)}, \sigma_z^{2(t)})}{\sum_i \xi_i^{(t)} \phi(x_j; \mu_i^{(t)}, \sigma_i^{2(t)})} \quad (2.27)$$

(2) 重み確率を計算する。

$$\xi_i^{(t+1)} = \frac{1}{N} \sum_j w_{ij}^{(t)} \quad (2.28)$$

(3) 各パラメータを計算する。

$$\mu_i^{(t+1)} = \frac{1}{N \xi_i^{(t+1)}} \sum_j w_{ij}^{(t)} x_j \quad (2.29)$$

$$\sigma_i^{2(t+1)} = \frac{1}{N \xi_i^{(t+1)}} \sum_j w_{ij}^{(t)} x_j^2 - (\mu_i^{(t+1)})^2 \quad (2.30)$$

(3) の計算を見ればわかるとおり、パラメータは各サンプルの重み確率で重み付けされた平均と分散になっている。正規混合分布の場合の EM アルゴリズムは、 $K$ -平均法においてユークリッド距離ではなくマハラノビス距離を用い、クラスタの平均だけでなく分散も計算する場合に相当する。

## 2.6 動的輪郭手法

画像を領域に分割するのではなく、輪郭線自体を移動させて領域抽出を行なう動的輪郭モデル (Active Contour Model) も提案されている。これは、輪郭線抽出や物体抽出問題をエネルギー汎関数の最小化問題としてモデル化する手法である [Sakaue99, 天野 98]。

動的輪郭モデルとしては、画像中の一次元の輪郭線である Snake [Kass98] が有名であるが、それ以外にも、メッシュ状の Active Net [坂上 91] や Balloon などが提案されている。以下では、3章で提案する手法で用いている、一次元動的輪郭モデルの Snake について述べる。

### 2.6.1 原理

パラメータ  $s$  で表される二次元平面上の任意の閉曲線である Snake を  $\mathbf{v}(s) = (x(s), y(s))$  ( $s \in [0, 1]$ ) とする。そのときの Snake エネルギー  $E_{snake}$  は

$$E_{snake} = \int_0^1 (E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s))) ds \quad (2.31)$$

と定義される。

$E_{int}$  は内部 (スプライン) エネルギーを表す項で、一次微分 (1次スプライン) と二次微分 (2次スプライン) の和として、以下の式で計算される。

$$E_{int} = \frac{1}{2} (\alpha(s) \| \mathbf{v}_s(s) \|^2 + \beta(s) \| \mathbf{v}_{ss}(s) \|^2) \quad (2.32)$$

$$= \frac{1}{2} (\alpha(s)(x_s^2 + y_s^2) + \beta(s)(x_{ss}^2 + y_{ss}^2)) \quad (2.33)$$

$$\mathbf{v}_s = \frac{\partial \mathbf{v}}{\partial s} = (x_s, y_s)^T \quad (2.34)$$

$$x_s = \frac{\partial x}{\partial s}, \quad x_{ss} = \frac{\partial^2 x}{\partial s^2}, \quad y_s = \frac{\partial y}{\partial s}, \quad y_{ss} = \frac{\partial^2 y}{\partial s^2} \quad (2.35)$$

ここで  $\alpha(s), \beta(s)$  は重みである。

$E_{image}$  は画像のエネルギーを表す項である。これは画像上の情報を用いて Snake を誘導するものである。この項に様々な関数を定義することで、ある程度物体についての知識を採り入れた Snake を作るができる。たとえばある色に近いほど値の大きくなるような関数を作れば、Snake はその色に収束することになる。以下は、ある輝度値  $I_0$  に収束する例と、エッジ  $\nabla I$  が急峻な個所に収束させる例である。

$$E_{image}(x, y) = |I(x, y) - I_0| \quad (2.36)$$

$$E_{image}(x, y) = -|\nabla I(x, y)|^2 \quad (2.37)$$

$E_{con}$  は外部からの制御力を表す項である。これはタスクによって様々な形を取り、輪郭に凹凸のある場合にも適切な輪郭線が得られるように、圧力項 [坂口 90] や面積項 [坂口 91] などが提案されている。

## 2.6.2 変分法による解法

変分法とは、積分形式で与えられる汎関数の極値を求める方法を与えるものである [篠崎 91]。汎関数の極値 (停留値) を与える関数を停留関数と呼び、その関数によって表される曲線を停留曲線と呼ぶ。

最も単純な形式の汎関数  $J(y(x))$  を次式に示す。

$$J(y(x)) = \int_{x_0}^{x_1} F(x, y, y') dx \quad (2.38)$$

ここで  $y$  は  $x$  についての関数  $y(x)$  であり、 $y' = \frac{dy}{dx}$  である。

この汎関数の極値を与えるための条件であるオイラー・ラグランジュの微分方程式は

$$F_y - \frac{d}{dx} F_{y'} = 0 \quad (2.39)$$

である。これを  $y$  について解くことで、停留曲線を得ることができる。

一般に  $n$  階導関数含む汎関数

$$J(y(x)) = \int_{x_0}^{x_1} F(x, y, y', \dots, y^{(n)}) dx \quad (2.40)$$

のオイラー・ラグランジュ方程式は次のようになる。

$$F_y - \frac{d}{dx}F_{y'} + \frac{d^2}{dx^2}F_{y''} + \cdots + (-1)^n \frac{d^n}{dx^n}F_{y^{(n)}} = 0 \quad (2.41)$$

動的輪郭手法の目的は式(2.31)のエネルギー汎関数を最小にすることである。この汎関数は独立変数  $s$  の関数  $x(s), y(s)$  とその導関数  $x_s, y_s, x_{ss}, y_{ss}$  を含むので、そのオイラー・ラグランジュ方程式は以下のようなになる。

$$F_x - \frac{d}{ds}F_{x_s} + \frac{d^2}{ds^2}F_{x_{ss}} = 0 \quad (2.42)$$

$$F_y - \frac{d}{ds}F_{y_s} + \frac{d^2}{ds^2}F_{y_{ss}} = 0 \quad (2.43)$$

ここで  $F$  は式(2.31)の右辺積分内を表す。すなわち、

$$F(s, x, y, x_s, y_s, x_{ss}, y_{ss}) = E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s)) \quad (2.44)$$

$$= \frac{1}{2}(\alpha(x_s^2 + y_s^2) + \beta(x_{ss}^2 + y_{ss}^2)) + E_{image}(x, y) + E_{con}(x, y) \quad (2.45)$$

導関数はそれぞれ

$$F_x = \frac{\partial E_{image}}{\partial x} + \frac{\partial E_{con}}{\partial x} \quad (2.46)$$

$$F_y = \frac{\partial E_{image}}{\partial y} + \frac{\partial E_{con}}{\partial y} \quad (2.47)$$

$$\frac{d}{ds}F_{x_s} = \frac{d}{ds}(\alpha x_s) = \alpha x_{ss} \quad (2.48)$$

$$\frac{d}{ds}F_{y_s} = \frac{d}{ds}(\alpha y_s) = \alpha y_{ss} \quad (2.49)$$

$$\frac{d^2}{ds^2}F_{x_{ss}} = \frac{d^2}{ds^2}(\beta x_{ss}) = \beta x_{ssss} \quad (2.50)$$

$$\frac{d^2}{ds^2}F_{y_{ss}} = \frac{d^2}{ds^2}(\beta y_{ss}) = \beta y_{ssss} \quad (2.51)$$

となる。したがって、解くべきオイラー・ラグランジュ方程式である式(2.42)、式(2.43)の具体的な形式は次のようになる。

$$\frac{\partial E_{image}}{\partial x} + \frac{\partial E_{con}}{\partial x} - \alpha x_{ss} + \beta x_{ssss} = 0 \quad (2.52)$$

$$\frac{\partial E_{image}}{\partial y} + \frac{\partial E_{con}}{\partial y} - \alpha y_{ss} + \beta y_{ssss} = 0 \quad (2.53)$$

### 2.6.3 差分方程式による実装

Snake を計算機上で実現するためには、微分方程式を差分方程式に変換する必要がある。以下では、Snake を構成する  $n$  個の点列を  $\mathbf{v}_i(s) = (x_i(s), y_i(s))$  とし、 $x$  に対しては以下

の差分式で近似する [坂上 91] が、 $y$  に対しても同様である。なお本来  $\alpha, \beta$  は  $s$  の関数であるが、ここでは定数として扱う。また  $E_{con} = 0$  とする。

$$x_s = x_{i+1} - x_i \quad (2.54)$$

$$x_{ss} = x_{i+1} - 2x_i + x_{i-1} \quad (2.55)$$

$$x_{sss} = x_{i+1} - 3x_i + 3x_{i-1} - x_{i-2} \quad (2.56)$$

$$x_{ssss} = x_{i+2} - 4x_{i+1} + 6x_i - 4x_{i-1} + x_{i-2} \quad (2.57)$$

$$\frac{\partial E_{image}}{\partial x} = E_{image}(x_i, y_i) - E_{image}(x_i + 1, y_i) \quad (2.58)$$

Snake は閉曲線であるので、 $x_i$  の添字  $i$  が範囲を超えた場合、すなわち  $i \geq n$  もしくは  $i < 0$  となった場合には、 $n$  で割った余り ( $i \bmod n$ ) を用いる。

これを用いると、式 (2.52) に対応する  $x_i$  に関する差分方程式は次のようになる。

$$\beta x_{i+2} + \Delta x_{i+1} + \Gamma x_i + \Delta x_{i-1} + \beta x_{i-2} = -(E_{image}(x_i, y_i) - E_{image}(x_i + 1, y_i)) \quad (2.59)$$

ここで  $\Gamma = 2\beta + 6\alpha, \Delta = -(\alpha + 4\beta)$  とする。

この式をすべての  $x_i, y_i$  に対して立て、連立方程式を解く。以下  $x$  についてのみ解法を述べるが、 $y$  についても同様に行なう。 $x$  に関する連立方程式は次のようになる。

$$\begin{bmatrix} \Gamma & \Delta & \beta & 0 & \cdots & \beta & \Delta \\ \Delta & \Gamma & \Delta & \beta & \cdots & 0 & \beta \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Delta & \beta & 0 & 0 & \cdots & \Delta & \Gamma \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = - \begin{bmatrix} E_{image}(x_1, y_1) - E_{image}(x_1 + 1, y_1) \\ E_{image}(x_2, y_1) - E_{image}(x_2 + 1, y_1) \\ \vdots \\ E_{image}(x_n, y_n) - E_{image}(x_n + 1, y_n) \end{bmatrix} \quad (2.60)$$

式 (2.60) 左辺の係数行列を  $A$ 、変数ベクトルを  $\mathbf{x}$ 、右辺を  $\mathbf{b}(\mathbf{x})$  と置いて、連立方程式を

$$A\mathbf{x} = \mathbf{b}(\mathbf{x}) \quad (2.61)$$

と書く。この式は連立一次方程式であるが、右辺が  $\mathbf{x}$  を含んでいるため、QR 分解や特異値分解などで一般化逆行列 [小柳 91, p.914] を計算して解くことは困難である。そこで、この連立方程式を線形反復法である SOR 法 (successive over relaxation method) [名取 90] を用いて解く。

まず、行列  $A$  を対角行列  $D$ 、左下三角行列  $L$ 、右上三角行列  $U$  の和に分ける。

$$A = L + D + U \quad (2.62)$$

これから第  $m$  近似解  $\mathbf{x}^{(m)}$  を、次の反復式によって計算する。

$$\boldsymbol{\xi}^{(m+1)} = D^{-1}(\mathbf{b}(\mathbf{x}^{(m)}) - (L + U)\mathbf{x}^{(m)}) \quad (2.63)$$

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \omega(\boldsymbol{\xi}^{(m+1)} - \mathbf{x}^{(m)}) \quad (2.64)$$

ここで  $\omega$  は、SOR法において加速因子として用いられるもので、通常は  $1 < \omega < 2$  の値を取る。この反復計算によって、画像エネルギーで定義された物体の輪郭に Snake が収束していく。

## 2.7 移動領域検出

移動する物体の領域を抽出する方法には、1.4.1節で述べた差分法のほか、オプティカルフローを用いる方法がある。オプティカルフローとは、画像上での動きベクトル場であり[三池93]、異なるフレーム間での対応付けにより求められる。オプティカルフローを計算する手法として、マッチング法や勾配法などがあるが、これらは、物体が移動したりカメラが動いたりして画像上での物体の位置が変わったとしても、その輝度分布が不変であることを仮定している。

ここでは勾配法について述べる。前述の仮定、すなわち、時刻  $t$  の画像上の点  $(x, y)$  が  $\delta t$  時刻で  $(\delta x, \delta y)$  だけ動いた時、その点の輝度  $I$  が不変であることから、

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (2.65)$$

が成立する。右辺を  $(x, y, t)$  のまわりでテーラー展開すると、

$$I = I + I_x \delta x + I_y \delta y + I_t \delta t + O(\delta x, \delta y, \delta t) \quad (2.66)$$

ここで  $O()$  は二次以上の項であり、これを無視すると

$$I_x \delta x + I_y \delta y + I_t \delta t = 0 \quad (2.67)$$

$\delta t$  で両辺を割ると、

$$I_x \frac{\delta x}{\delta t} + I_y \frac{\delta y}{\delta t} + I_t = 0 \quad (2.68)$$

$(u, v) = (\delta x/\delta t, \delta y/\delta t)$  として

$$I_x u + I_y v + I_t = 0 \quad (2.69)$$

が得られる。この式をオプティカルフロー  $(u, v)$  の基本式、拘束式などと呼ぶ。このとき  $u, v$  は  $x, y$  の関数で、各点の速度ベクトルに対応する。

この基本式だけでは速度ベクトルを一意に決定することができないため、さらに条件が必要となる。速度ベクトルが空間的に滑らかに変化するという条件[Horn93]を用いる空間的大域最適化法[三池93]の場合、基本式の誤差の総和と空間変化の総和を最小にする目的関数  $E$  は次のようになる。

$$E = \iint \left\{ (I_x u + I_y v + I_t)^2 + \lambda (u_x^2 + u_y^2 + v_x^2 + v_y^2) \right\} dx dy \quad (2.70)$$



この目的関数を最小にする関数  $u, v$  を求めればよい。ここで  $\lambda$  は係数である。変分法により関数  $u, v$  を求めるためのオイラーラグランジュ方程式は以下のようになる。

$$\nabla^2 u = \lambda(I_x u + I_y v + I_t)I_x \quad (2.71)$$

$$\nabla^2 v = \lambda(I_x u + I_y v + I_t)I_y \quad (2.72)$$

また別の条件も考えられる。ある小領域  $S$  の内部でオプティカルフローが一定であるという条件を用いる空間的局所最適化法 [三池 93] では、その局所領域の各点から求まる式を最小二乗法を用いて求める。

$$E = \iint_S (I_x u + I_y v + I_t)^2 dx dy \quad (2.73)$$

ここでは  $u, v$  は関数ではなく、求めるべき定数である。

最小二乗法の計算では、式 (2.69) における誤差分布が正規分布であることを暗に仮定しているが、Ohta [Ohta91] は誤差項  $I_r$  を次のように明示的に導入した。

$$I_t = -I_x u - I_y v + I_r \quad (2.74)$$

空間微分関数  $I_x, I_y$  と誤差関数  $I_r$  の線形和が時間微分  $I_t$  と等しいというモデルであり、 $I_r$  は  $I_x$  と  $I_y$  に直交する場合に最小となる。すなわち、

$$I_r \cdot I_x = 0 \quad (2.75)$$

$$I_r \cdot I_y = 0 \quad (2.76)$$

ここで  $\cdot$  は次式で定義される関数の内積である。

$$f \cdot g = \iint_S f(x, y)g(x, y) dx dy \quad (2.77)$$

式 (2.75) と式 (2.76) が成立する時、式 (2.74) より

$$- \begin{pmatrix} I_t \cdot I_x \\ I_t \cdot I_y \end{pmatrix} = \begin{pmatrix} I_x \cdot I_x & I_x \cdot I_y \\ I_x \cdot I_y & I_y \cdot I_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (2.78)$$

となる。各内積は定義より計算できるので、上式を  $(u, v)$  について解けば、小領域  $S$  のオプティカルフローが計算される。この手法を実装した C 言語によるプログラムが公開されている [Ohta96]。この手法により求めたオプティカルフローは 5.5 節を参照のこと。

## 2.8 色空間と色差

人間が物を見る時に用いる重要な特徴として物体の色がある。人間の網膜の中には色刺激に反応する神経細胞である錐体があり、その神経信号が脳内の外側膝状体を經由して、

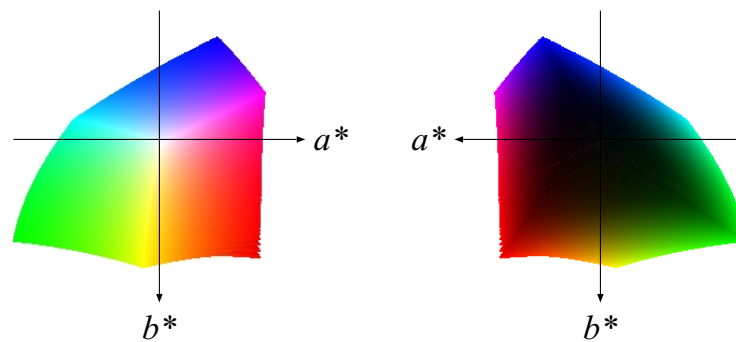


図 2.1:  $L^*a^*b^*$  表色系の  $a^*b^*$  平面への投影。左は  $L^*$  軸を上から見た場合、右は  $L^*$  軸を下から見た場合。

大脳後頭部の1次視覚野へと送られ、処理されている [色彩研 93a, 乾 90, 川人 94]。視覚による色の弁別能力は高く、色によっては波長が数 nm 異なるだけでもそれを知覚することができる [色彩研 93a]。したがって、客観的に同じ色かどうかを判断する場合には、その色の差を正確に測定する必要がある。二つの色の差を定量的に評価する方法として、CIE (国際照明委員会 Commission Internationale de l'Eclairage) が 1976 年に  $CIE L^*a^*b^*$  均等色空間とその色差式を勧告した [色彩研 93b] (図 2.1 参照)。他のいくつかの表色系と同じように、輝度 (luminance) を  $L^*$  軸に、色度 (chromaticity) を  $a^*b^*$  軸に割り当て、両者を分離している。この表色系は、二つの色の距離が一定であればどの色領域においてもほぼ一定の知覚的な差を与えるように定めたものであり、日本工業規格 (JIS Z 8729) として、色見本と同じ製品を再現して製造するような工業にも広く用いられている。

領域分割では色を特徴量とした手法が数多く提案されており [Skarbek94]、様々な色空間での領域分割の比較 [高橋 99] も行われているが、 $L^*a^*b^*$  表色系は領域分割において広く用いられている。

$RGB$  表色系から  $L^*a^*b^*$  表色系への変換は、次のようにして行うことができる。まず、 $RGB$  表色系を  $XYZ$  表色系に変換する。実際に用いられている  $RGB$  にはさまざまな実装があるが [Prrat91, Ford98]、NTSC 信号の  $RGB$  からの変換の場合は以下の式を用いる [Ford98]。

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (2.79)$$

ここで  $R, G, B \in [0, 1]$  である。

$L^*a^*b^*$  表色系は  $XYZ$  表色系を用いて次のように定義される [Prrat91, 色彩研 93b, Ford98, Poynton97]。

$$L^* = \begin{cases} 116\sqrt[3]{\frac{Y}{Y_n}} - 16 & \left(\frac{Y}{Y_n} > 0.008856\right) \\ 903.3\frac{Y}{Y_n} & \left(\frac{Y}{Y_n} < 0.008856\right) \end{cases} \quad (2.80)$$

$$a^* = 500 \left( f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right) \quad (2.81)$$

$$b^* = 200 \left( f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right) \quad (2.82)$$

$$f(t) = \begin{cases} \sqrt[3]{t} & (t > 0.008856) \\ 7.787t + \frac{16}{116} & (t < 0.008856) \end{cases} \quad (2.83)$$

ここで  $X_n, Y_n, Z_n$  は標準の光の下の完全拡散面の三刺激値である。 $(R, G, B) = (1, 1, 1)$  に対して、 $(X_n, Y_n, Z_n) = (0.981, 1.000, 1.182)$  となる。相対輝度  $\frac{Y}{Y_n}$  が 0 から 1 に正規化されているとき、 $L^*$  は 0 から 100 の値をとる。

二つの色の  $L^*a^*b^*$  を  $L_1^*, a_1^*, b_1^*, L_2^*, a_2^*, b_2^*$  とすると、その色差  $\Delta$  は  $L^*a^*b^*$  色空間におけるユークリッド距離である次式で計算される。

$$\Delta = (L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2 \quad (2.84)$$



# 第3章 対象物体に依存しない 静止画像からの物体抽出

## 3.1 はじめに

本論文の目的は、より一般的な状況で撮影された画像から対象を抽出することである。まずこの章では、静止画として撮影された画像から対象物体を抽出する手法について述べる。写真など静止画像の被写体となるものは様々なので、ある決まった物体を抽出するという手法ではなく、撮影された未知の対象を抽出するための処理を行わなければならない。そしてその処理では、特定の物体についての固有の知識を用いることはできないが、このような場合には領域分割の手法が有用である。

2章で述べたように、領域分割処理は画像内の一つの物体に一つの領域を対応させる結果を出力することが理想的である。したがって、一つの物体を複数の領域に分割してしまう状況は好ましくない。この状況は過分割と呼ばれるが、できるだけ過分割を避ける処理を行なうことが望ましい。しかし一般的な領域分割処理においては、画像内の領域とはもつと狭義に捉えられており、輝度や色、テクスチャなどの特徴量が類似した画素の集合として定義されている(2.1節参照)。このような領域の定義を用いる処理では、過分割を避けることは難しい。なぜなら、二つの領域の輝度や色が異なっていたとしても、それらが一つの物体の領域である場合があるからである。そのため、複数の領域の情報を基に、それらが物体の領域を形成しているかどうかを判定することが必要となる。

過分割を抑制するために、領域を併合することで対処する研究が従来からある。 $K$ -平均法のクラスタ数[中谷93]あるいは初期クラスタの位置[境田98]を変化させるとクラスタリングの結果が変わることを利用し、生成された境界線の位置が変化するのかどうかを判定基準として領域を併合する手法も提案されているが、それらは領域の境界線がエッジに一致することを狙ったものである。また、輪郭形状、コントラスト、色分布に基づく同一領域度に基づいて領域併合を行うという手法[志治00]も提案されている。しかし、併合基準が一つしかないために、形状による二つの領域が同一領域度が大きかったとしても、コントラストが大きかったり色が異なっている場合は併合される可能性が低くなり、この手法では前述の過分割に対処することはできない。

この章では、物体に固有ではない一般的な情報を用いて、二つの領域の輝度や色が異なっ

ているような場合でも、物体を一つの領域として抽出する領域分割手法を提案する。まず過分割の原因となる物理的要因及び光学的要因について述べる。そして、一つの物体の領域が物理的要因によって分割されたのかどうかを判定するために、二つの領域間の配置・隣接・面積比の情報を導入し、それに基づいて併合を行うための四つの特徴量を提案する。それらの特徴量は、一つの領域の内部についてのものではなく、二つの領域間についての特徴量である。これによって、物体表面が様々な反射率を持っているというような、類似した特徴量を持つ画素の集合という従来の領域の定義では扱えないような場合であっても、物体の領域を得ることができる。

以下3.2節では物理的要因及び光学的要因について述べ、3.3節では二つの要因を排除するための領域間の特徴量について述べる。領域併合の戦略と領域分割アルゴリズムを3.4節と3.5節で、実画像を用いた実験結果と考察をそれぞれ3.6節と3.7節で示す。

## 3.2 過分割の二つの要因

領域分割処理の過分割には、大きく分けて以下の二つの理由が挙げられる。

まず一つ目は、同一の反射率を持つ同じ物体の表面でありながら、陰影、ハイライトや相互反射などさまざまな影響で照明条件が変化するため、場所によって輝度や色が変わってしまうという理由がある。ここではこれを光学的要因と呼ぶ<sup>1</sup>。この光学的要因のために、一様な反射率を持つ物体表面上では一様な色や輝度を持つという仮定が成立しなくなり、一つの物体表面上に複数の領域が形成されることになる[浜本96]。そのため、反射モデルを用いてこれらの影響を排除する手法[富永98]や色分布上でハイライトを分割する手法[堀田98]などがとられている。

二つ目は、物体表面が実際に複数の異なる反射率を持っているという理由である。これを物理的要因と呼ぶことにする。例えば、サッカーボールは白と黒のパーツで構成されていたり、また本の表紙などはカラフルな色で印刷されている。このような場合、同じ物体の表面であっても多数の領域に分割されることは容易に想像できる。

一般的に、実世界では、一つの特徴量しか持たないという物体はほとんどないと言える。そのため通常の領域分割では、一つの領域は一つの物体には対応せず、それぞれのパーツに対応する部分領域に分割するところまでしか行われていない。この問題は、いくつかの特徴量をまとめた特徴ベクトルを用いても同じである。一つの部分領域に対して計算される特徴量を用いるだけでは、特徴量が大きく異なれば複数の領域に分割してしまう。つまり、異なる特徴量を持つ部分領域同士を組み合わせたものが一つの物体に対応する領域を形成するかどうかを考慮しなければ、物理的要因による過分割は解決できない。

---

<sup>1</sup>光も物理学的なものではあるが、ここでは言葉の定義上、照明条件に対して光学的、物体の構成の性質に対して物理的という単語を用いている。

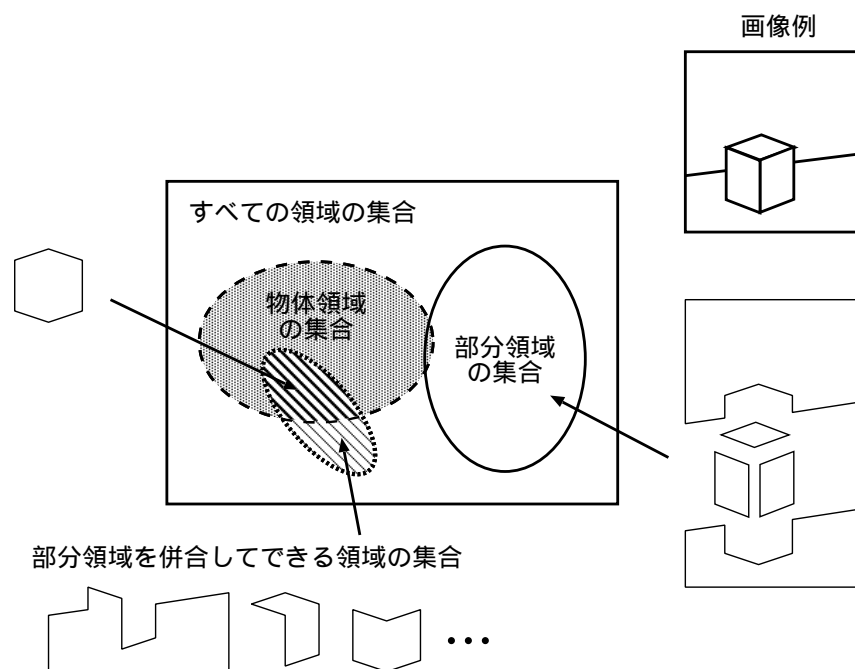


図 3.1: すべての領域の集合における、部分領域、それを併合した領域、物体に対応する領域

### 3.3 領域間の情報と併合のための特徴量

従来研究では、前述したように、二つの領域の輝度や色が異なっていたとしても、それらが一つの物体の領域である場合があるということは考慮されていない。そのため、複数の領域の情報を基に、それらが抽出すべき物体の領域を形成しているかどうかを判定することが必要である。以下では、通常領域分割で得られる、類似した特徴量を持つ画素の集合である領域（2.1節の定義）を部分領域と呼ぶことにする。もし一つの物体の領域が、物理的要因によって二つの部分領域に分割されているのであれば、それらの部分領域を併合することで一つの物体の領域が得られると考えられる。

ここで用いる一つの物体の領域とは、色や輝度の違いに関わらず、三次元物体の画像平面への投影像を指す。例えば、人間が写っている画像であれば、顔や目鼻、髪、服などの部分領域が存在するが、それらは人間の一部分であり、人間の投影像、すなわち物体領域を形成している。ただし、図3.1に示すように、部分領域を併合してできる領域には、物体に対応する領域ではなく、実世界の物体とは関係のない意味のない領域も存在する（破線外側の斜線部）。また、物体固有の知識を用いなければ物体に対応する領域が得られないこともあり、図3.1の破線内部の斜線部分は網掛部分とは一致しない。物体の領域（破線内部）かどうかを定義するには、実際の物体というものを定義しなければならないが、それは困難であるので、破線の内部と外部の区別は明確ではない。したがって、ここでは破線内部の領域のうち、物理的要因によって分割されているものを見つけ、併合によってそ

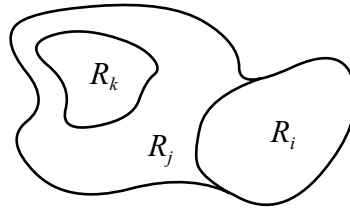


図 3.2: 二つの領域の配置。  $R_j$  と  $R_k$  は包含、  $R_i$  と  $R_j$  は隣接、  $R_i$  と  $R_k$  は非隣接の関係にある。

の領域を得ることを目指す。

そこで、領域が物理的要因によって分割されたのかどうかを判断するために、

- 二つの領域がどのような配置にあるのか（配置情報）
- どのように二つの領域が隣接しているのか（隣接情報）
- 二つの領域の相対的な大きさはどれぐらいなのか（面積比情報）

という三つの領域間の情報を用いる。以下では、それらの情報から判断できる物理的要因および光学的要因について述べ、それらを定量的に測るために導入する特徴量について説明する。この特徴量に基づいて、隣接する二つの領域を併合すべきかどうかを決定するアルゴリズムは、次節で示す。

以下では、画像中の領域を  $R_i$  ( $i = 1, \dots$ ) と書くことにする。

### 3.3.1 配置情報

配置情報とは、二つの領域が画像中でどのように配置されているか、その位置関係を表す。二つの領域の配置は、包含、隣接、非隣接の三つに分けられる（図 3.2 参照）。非隣接の場合には、配置情報からだけでは同じ物体の領域に属するかどうかは不明である。しかし、隣接している領域であれば同じ物体の領域に属する可能性がある。

包含は隣接関係の特殊な場合である。「人物の顔の領域には目や口の領域が包含されている」というような例にも見られるように、それらの部分領域は、物体を構成するパーツに対応している。したがって、包含の関係にある領域は物理的要因によって生成されたと考えられ、それらを併合することで物体の領域が得られると考えられる。

画像上で包含となる別の可能性として、二つの物体が異なる奥行きで前後に配置されている場合が考えられる。これには物体に穴があいている場合も含まれるが、穴のあいている物体の場合には、その穴ではなく物体の方が重要であり、また一般的には、視野内でより大きな物体の方が小さな物体よりも注目されるのが自然であるので、これらの場合も併合してよいと考えられる。



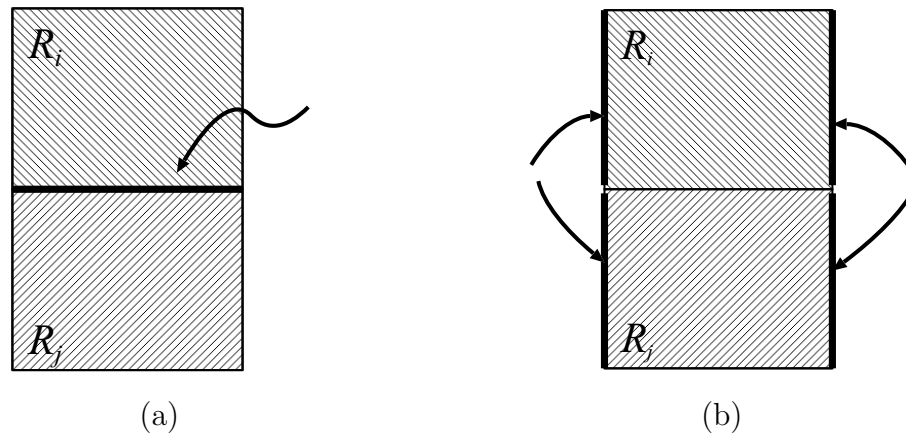


図 3.3: 領域  $R_i$  と  $R_j$  が隣接している場合の境界線の (a) 共有部分と (b) 非共有部分

包含されているかどうかは、隣接する領域の数によって決定できるので、以下の包含関係を表す特徴量（包含の特徴量）を導入する。

$$V_{inclusion}(i, j) = \begin{cases} 1 & Q(i) = \{j\} \\ 0 & \text{その他} \end{cases} \quad (3.1)$$

ここで  $Q(i)$  は、領域  $R_i$  に隣接する領域の番号の集合であり、 $R_j$  が  $R_i$  に隣接するとき、 $j \in Q(i)$  である。この特徴量は、 $R_i$  に隣接している領域が  $R_j$  ただ一つしか存在しない場合、つまり領域  $R_j$  に包含されている場合に1となる。

### 3.3.2 隣接情報：境界線の共有部分

二つの領域が隣接している場合に、どのように隣接しているか（隣接情報）は、領域の境界線を用いて表される。領域が隣接している場合に、注目すべき境界線の部分は二ヶ所ある。一つはT接点で接続している境界線のうち両者が共有している部分であり、もう一つは共有されていない（非共有）部分である（図3.3 参照）。

共有されている境界線に注目して、二つの領域が同じ物体の領域に属するかどうか判断するためには、その境界線が滑らかかが重要となる（図3.5 参照）。もし境界線が滑らかで直線に近ければ、その二つの領域の間には明確な特徴量の変化（エッジ）が存在することになる。この場合は、二つの領域が同じ領域に属しているとするならば、特徴量が明らかに違うため、それらは物理的要因によって分けられているはずである。しかし、物理的要因によって分けられているのかがどうかは領域の形状から判断しなければならないが、共有されている境界線からはそれを判断することはできない。

しかし、もし境界線が滑らかではなくぎざぎざしたものであれば、それは主として光学的要因によって生じる擬似輪郭である場合がある。一般に擬似輪郭とは、図3.4に示すように、画像の階調を落としたとき、実際の輪郭は存在しないのににもかかわらず、緩やかな

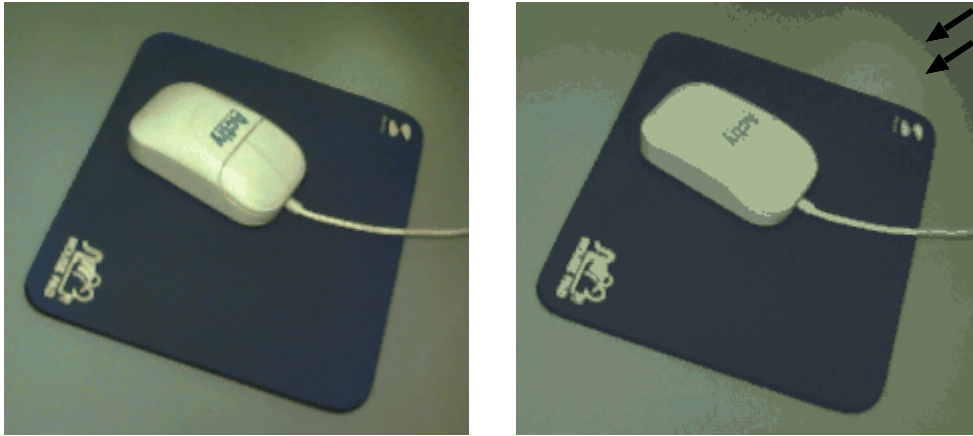


図 3.4: 階調を落した時に現れる疑似輪郭 (左) 原画像 (右) 階調を落した画像。矢印は顕著な疑似輪郭を指す。

輝度勾配があった場所に現れる、画像上の階調の急激な変化を指す。領域分割においては、特徴量が緩やかに変化している場所が、わずかに異なる特徴量を持つ二つの領域に分割された場合に、それらの領域の間の境界線が疑似輪郭となる。この境界付近では、ノイズの影響でどちらの領域に属するかが画素毎に頻繁に変わってしまう。そのため、二つの領域の境界線は滑らかにはならず、ぎざぎざした線になる。

木などのように物体の輪郭線自体が滑らかではない物体も存在するが、実際の輪郭線の滑らかさはその物体を見る距離に依存する。遠くから木全体を観察するとき、詳細な輪郭を見るには解像度が足りず、大まかな輪郭が知覚され、滑らかに見える。逆に近い距離から観察すると、木の輪郭ではなく個々の葉の輪郭が見えてくるので、やはり滑らかな輪郭が見えることになる。しかし、実際に物体の形状自体がぎざぎざしている鋸の歯のような境界も存在するため、以下では物体はそのような境界を持たないとする。

したがって、共有される境界線が滑らかではない場合、その境界線は物体の外輪郭ではなく、疑似輪郭であるとみなすことにする。よって、これらを併合することで、物体の領域が得られると考える。これらの領域は、物理的要因ではなく光学的要因によって生成されたものではあるが、境界線の情報を用いて判断することができる。

共有されている境界線の滑らかさを評価するために、以下の特徴量 (滑らかさの特徴量) を導入する。

$$V_{smoothness}(i, j) = \frac{1}{L_{ij}} \int_{\Omega} \frac{\|\mathbf{v}_i(s) - \mathbf{v}_i(s + \ell)\|}{\ell} ds \quad (3.2)$$

$$\Omega = \{s \mid \mathbf{v}_i(s) \in \partial R_i \cap \partial R_j \text{ かつ } \mathbf{v}_i(s + \ell) \in \partial R_i \cap \partial R_j\} \quad (3.3)$$

ここで  $\partial R_i$  は領域  $R_i$  の境界線を表す。 $\mathbf{v}_i(s)$  は  $R_i$  の境界線上の点を表し、 $\mathbf{v}_i(s) \in \partial R_i$  である。 $L_{ij}$  は  $R_i$  と  $R_j$  が共有する境界線の長さを表し、 $L_{ij} = \int_{\Omega} ds$  である。 $\ell$  はある定数であり、滑らかさを定義する単位長である。この特徴量は、境界線上の道のりで  $\ell$  だけ

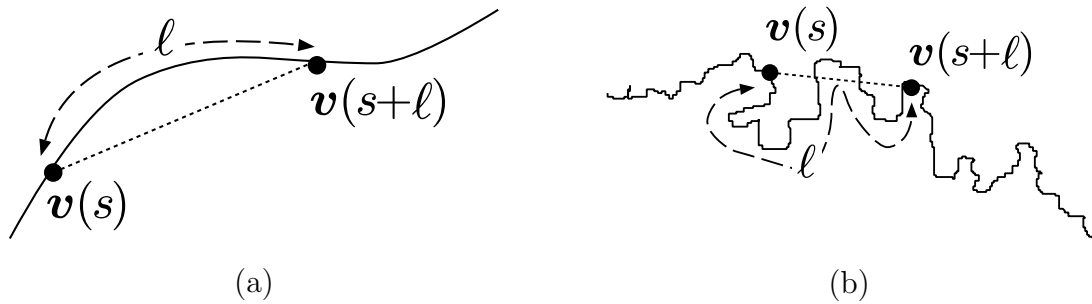


図 3.5: 境界線の共有部分 (a) 滑らかな場合は  $V_{smoothness}$  が大きい。(b) ギザギザした場合は  $V_{smoothness}$  が小さい。点線は二点間の直線距離、破線は道のり  $l$  を表す。

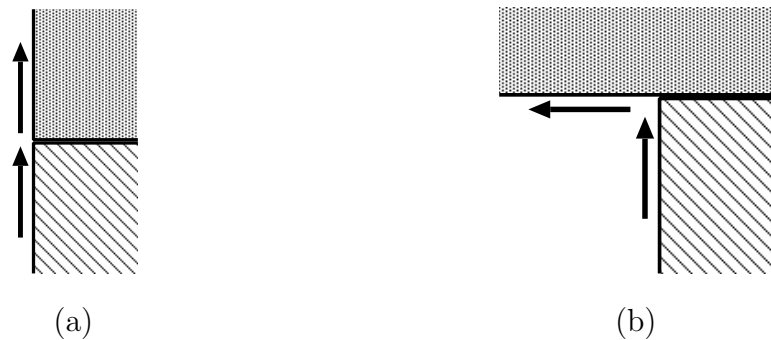


図 3.6: 境界線の非共有部分の方向 (矢印) の変化 (a) 変化は緩やかで、境界線は連続的に接続。(b) 変化は著しく、境界線の接続は連続的ではない。

離れた二点間の直線距離と道のりの比を表し (図 3.5 参照) 境界線が直線のとときに 1 となり、滑らかでなくなるほど 0 に近くなる (付録 A 参照)。

### 3.3.3 隣接情報：接続している二つの境界線の非共有部分

次に二つの領域の境界線の非共有部分に注目すると、非共有部分が接続している接点における境界線の方向の変化が著しいのか緩やかなのかという情報が重要になる (図 3.6 参照)。この変化によって、物理的要因で生じる二つの部分領域が、物体の外輪郭で隣接しているという状況なのかどうかを判断することができる。

非共有部分が接続している接点には、共有部分も接続しているので、境界線は T 接点を形成している。この T 接点において境界線の二つの部分の方向が同じ (または近い) とき、接続が連続的であるとする (図 3.6(a))。逆に方向が急激に変化しているときは、連続的ではないとする (図 3.6(b))。

一つの隣接領域に対して T 接点は二つ (もしくは偶数個) あるが、もしそれらすべての T 接点で非共有部分の接続が連続的であるならば、その二つの領域は物理的要因によって

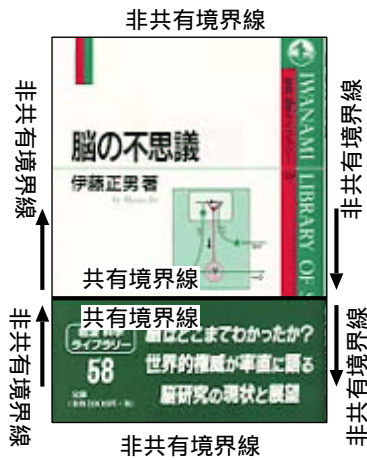


図 3.7: 本の例

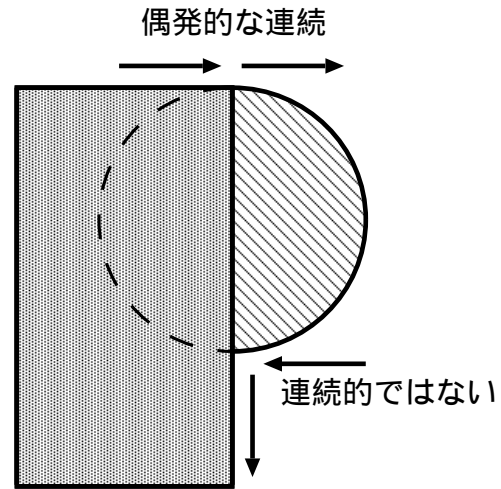


図 3.8: 偶発的に連続的となる例

分割されたと考えることができる。例えば、図3.7のように本の表紙が上下別の色に印刷されていて、上部の白色の領域と下部の緑色の領域の二つの領域が生成されるような場合を考えよう。緑の領域の境界線は、本の縁にあるT接点を挟んで、白の領域と境界線を共有する部分と、それ以外の非共有部分とに分けられる。このT接点において二つの非共有部分の方向の変化は、図3.7の矢印が示すように緩やかで、連続的に接続している。

もちろん図3.8のように、T接点において連続的に接続する境界線が存在しても、それが物体の外輪廓ではない可能性もある。例えば、物体の角などに対応して境界線が急激に折れ曲がったちょうどその場所に、背景にある領域の境界線が似た角度で存在するような場合である。しかし、そのような偶発的な場合が、複数のT接点すべてにおいて起こる可能性は非常に低い。

したがって、二つ（または偶数個）のT接点において非共有部分が連続的に変化している領域は、物理的要因によって物体の領域が分割されて形成された部分領域であるといえる。そして、これらの部分領域を併合することで物体の領域が得られると考えられる。

接続が連続的かどうかを評価するために、T接点における非共有部分のなす角度によって定義される次の特徴量（連続性の特徴量）を導入する。

$$V_{continuity}(i, j) = \frac{1}{k_n(i, j)} \sum_{k=1}^{k_n(i, j)} (\pi - \theta_k(i, j)) \quad (3.4)$$

$k_n(i, j)$  は領域  $R_i$  と  $R_j$  の境界線によって形成されたT接点の個数、 $\theta_k(i, j)$  は  $k$  番目のT接点で境界線の非共有部分がなす角度で、図3.9に示すように、 $\theta_k(i, j) \leq \frac{\pi}{2}$  である。つまりこの特徴量はT接点での境界線のなす角度の平均値であり、接続が連続的であるほど（ $\theta_k$  は  $\pi$  に近付くので）この特徴量は小さくなり、直線的に接続していれば0となる。

またこの連続性は、ゲシュタルト心理学でよく知られている郡化の要因（体制化の法則）

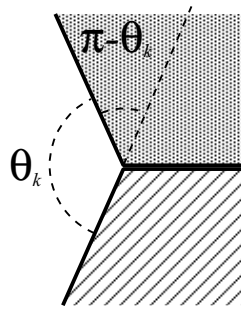


図 3.9: T接点での境界線のなす角度  $\theta_k$

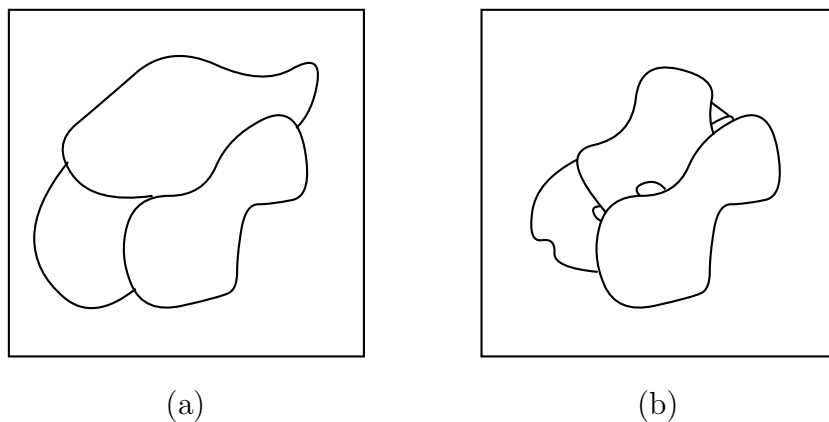


図 3.10: 面積比の違い (a) どの領域も同じ程度の大きさ (b) 比較的小さい領域が存在する

の一つである、よい連続の要因に対応している [大山 00, 感覚 94]。

### 3.3.4 面積比情報

面積比情報とは、二つの領域の大きさがどれくらい違っているのかを示す。隣接する二つの領域が、両方ともに画像に対して比較的大きいならば、それらの領域はその画像において重要な意味を持つと考えられる (図 3.10(a))。しかし、両方とも数画素程度の小さい領域である場合や、一方に比べて他方が非常に小さい場合、それらはノイズか、あるいはより大きな物体の一部に対応する小さな領域であると考えられる (図 3.10(b))。

ただし、一画素程度の大きさの物体が多数あるような状況も考えられる。そのような例としては、草原や布地などの統計的テクスチャがあるが、これらは別の処理 (例えばテクスチャによる領域分割など) によって、あらかじめ一つにまとめられるべきである。

先の三つの領域間の情報とは異なり、この小さな領域をどの隣接領域に併合させるかということについては、一意には決まらずいくつかの可能性がある。最も面積が大きい隣接領域に併合してしまうと、その領域の境界線は小さな領域が併合された分だけ滑らかでは

なくなってしまうかもしれず、滑らかさの特徴量に影響を与えてしまう。もし小さな領域がノイズによるものであれば、併合すべき領域の特徴に近いと思われるので、大きい領域にではなく領域の色などの特徴量が似た領域に併合されることが望ましいと思われる。したがって、面積比の情報で併合する場合は、色などが似た領域へ併合する。

面積比の特徴量は、二つの領域の面積比で定義する。

$$V_{area}(i, j) = \frac{S(R_i)}{S(R_j)} \quad (3.5)$$

ここで  $S(R_i)$  は領域  $R_i$  の面積とする。 $V_{area}(i, j)$  は、領域  $R_i$  とそれに隣接する領域  $R_j$  との面積比を表し、隣接領域のほうが大きければ1より小さな値を取り、そうでなければ大きな値を取る。

### 3.4 併合の方略

前節では領域間の特徴量を導入した。本手法ではそれら領域間の情報に基づいて領域の併合を行う。したがって、始めに画像を部分領域に分割し、それから併合操作を行うことになる。

ある二つの領域の間で、前節で導入した特徴量を計算する。もし特徴量が設定されたあるしきい値に満たない場合、その二つの領域を併合する。そして新しい領域を加えて再び特徴量を計算し、併合を繰り返す。どの二つの領域についても、計算された特徴量がしきい値以上の場合、処理を終了する。

以上が本手法における領域併合基本的な戦略であるが、いくつか補足をする。

#### 3.4.1 併合に用いる特徴量の順序

二つの領域を併合するか否かを判断する際に、四つの特徴量のうちでどれを先に用いるかの優先順位を決めておかなければならない。滑らかさと連続性の特徴量は境界線の情報から計算されるので、小さな領域の短い境界線に対して計算してもあまり信頼できる結果は得られない。よって、それらより先に包含と面積比の特徴量を用いて併合を行ない、小さな領域をできるだけ排除しておくことになる。

したがって、用いる特徴量の順序は以下ようになる。まず包含されている領域を先に排除するために、包含の特徴量を用いる。この特徴量を用いて併合を行っても、他の領域の境界線の情報には影響を与えない。次に、画像中の細かい領域を排除するために、面積比の特徴量を用いる。これらの特徴量を用いて併合を繰り返すと、比較的小さな領域が少なくなり、どの領域も同程度に大きくなると考えられるので、境界線の情報を用いることができる。その後、滑らかさの特徴量と、連続性の特徴量を用いる。

それぞれの特徴量を用いる段階では、画像中のすべての隣接する二つの領域についてその特徴量を計算し、しきい値以下で最も小さい特徴量をもつ二つの領域を併合する。最も小さい特徴量がしきい値を上回る場合は処理を終了する。

### 3.4.2 しきい値の設定 / 更新

併合するか否かを判断するしきい値の設定の方針や、それが画像の大きさに依存するかどうかは、それぞれの特徴量によって異なる。これはある程度経験的に決めなければならないが、どのように決定するのかという方針は必要である。

包含の特徴量は0か1なので、しきい値を必要としない。

滑らかさの特徴量は、局所的な輪郭線の情報を用いるため、画像の大きさには依存しないが、パラメータ  $l$  に依存する。予備実験の結果、 $l = 20$  の時、擬似輪郭ではおよそ0.5~0.65以下の値を取ることがわかったのでその値をしきい値とすることができる。この予備実験の詳細は付録Aを参照のこと。

連続性の特徴量はT接点の角度なので、画像の大きさには寄らない。この角度が0°であればもっとも連続的であるが、実際にはそのようなことは少ないので、5°程度か、大きくとも20°をしきい値とする。

面積比の特徴量は生成される領域の面積、すなわち画像の大きさに依存する。しかし、画像の大きさの1/30程度より小さい領域は意味のあるものではないと考えた場合には、しきい値は最高で1/30とすることができる。

以上がしきい値の設定の方針であるが、画像によって適切なしきい値は異なるので、それらのしきい値を初めに固定してしまうと、しきい値以下の特徴量を持つ領域同士が一斉に併合されることになる。それを防ぐために、まずしきい値を小さな値に設定しておき、併合が進むにつれて徐々に大きな(ただしある値以下の)値に更新していくというように、動的にしきい値を変更する(詳細は3.5.2節 Step.5 参照)。

### 3.4.3 画像の性質による例外

3.3節で述べた領域間の情報は一般的な場合についてのものであり、画像中の領域にそのまま適用するにはいくつかの問題がある。

一つ目は、包含についての問題である。背景の中に一つだけ物体が存在している場合には、背景の領域が物体の領域を包含しているので、これらは併合されてしまい、物体を領域として抽出することはできない。したがって、画像の中の領域が二つだけになった場合には、それらは図と地の二つになったと考えられるので、併合しないことにする。

もう一つの問題は、本来は境界ではない画像の辺において連続的な接続ができてしまうことである(図3.11参照)。画像は四つの直線に囲まれているので、画像の辺に接してい

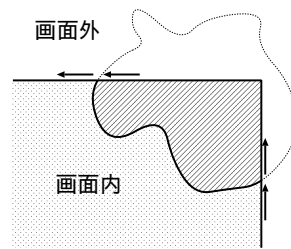


図 3.11: 画像の辺における連続的な接続

る領域の境界線上の T 接点では、必ず直線的な接続になってしまう。よって、画像の辺に接する T 接点では、連続性の特徴量（つまり式(3.4)）は計算しないこととする。

## 3.5 領域間の特徴量に基づいた併合法

前節で述べた領域併合の方略に基づいて、本手法で提案する領域分割手法の具体的な処理の手順 — 初期分割、併合アルゴリズム、領域抽出 — を示す。

### 3.5.1 初期分割による部分領域生成

初期領域分割は従来用いられているクラスタリング手法 [原95] を使用する。2.8節で述べた  $L^*a^*b^*$  表色系に変換し、各画素の色 ( $L^*, a^*, b^*$ ) と位置  $(x, y)$  にそれぞれ重みをつけた5次元ベクトルについて、2.3節で述べた  $K$ -平均法を用いてクラスタリングを行う。

この初期分割によって、類似した色の画素の集合である  $N$  個の部分領域が生成される。それらを  $R_i$  ( $i = 1, \dots, N$ ) とする。

### 3.5.2 併合アルゴリズム

初期分割によって生成された領域に対して、3.4節で述べた戦略によって領域の併合を行う。ここでは本手法で提案する併合のアルゴリズムを示す。以下では、 $Th_a, Th_s, Th_c$  はそれぞれ面積比、滑らかさ、連続性の特徴量を用いて併合するかどうかを決定するしきい値であり、 $h$  はしきい値の更新回数<sup>2</sup>、 $M$  は各時点で残っている領域数である。

Step 0.

$$h = 0, M = N \text{ とする。}$$

Step 1.

<sup>2</sup>これは併合処理を行った回数とは異なる。併合回数は  $N - M$  で表される。



領域数が2 ( $M = 2$ ) であれば、処理を終了する。

そうでなければ、もし  $V_{inclusion}(i, j) = 1$  なる領域  $R_i, R_j$  が存在するなら、 $R_i$  と  $R_j$  を併合する。  $M \leftarrow M - 1$  として Step 1. へ。

存在しなければ、Step 2. へ。

#### Step 2.

$V_a = \min_i \min_{j \in Q(i)} V_{area}(i, j)$  とする。

もし  $V_a \geq Th_a^{(h)}$  ならば、Step 3. へ。

そうでなければ、 $V_a$  を与える  $R_i$  と、 $R_i$  の隣接領域のうち最も色差が小さい領域と併合する。  $M \leftarrow M - 1$  として Step 1. へ。

#### Step 3.

$V_s = \min_i \min_{j \in Q(i)} V_{smoothness}(i, j)$  とする。

もし  $V_s \geq Th_s^{(h)}$  ならば、Step 4. へ。

そうでなければ、 $V_s$  を与える  $R_i$  と  $R_j$  を併合する。  $M \leftarrow M - 1$  として Step 1. へ。

#### Step 4.

$V_c = \min_i \min_{j \in Q(i)} V_{continuity}(i, j)$  とする。

もし  $V_c \geq Th_c^{(h)}$  ならば、Step 5. へ。

そうでなければ、 $V_c$  を与える  $R_i$  と  $R_j$  を併合する。  $M \leftarrow M - 1$  として Step 1. へ。

#### Step 5.

しきい値  $Th_f^{(h)} (f = a, s, c)$  に更新比率  $r_f (> 1)$  を乗じて値を更新する。更新によってしきい値の値は増えることになるが、設定された上限  $Th_f^E$  を超えたしきい値は、以後その上限の値を用い、更新を行わない。つまり、各特徴量  $f = a, s, c$  について

$$Th_f^{(h+1)} = \begin{cases} r_f Th_f^{(h)}, & r_f Th_f^{(h)} \leq Th_f^E \\ Th_f^E, & \text{otherwise} \end{cases} \quad (3.6)$$

そして、もしすべてのしきい値が上限に達した場合は処理を終了する。

そうでなければ  $h \leftarrow h + 1$  として Step 1. へ。

### 3.5.3 領域抽出

上記の併合アルゴリズムが終了した時点では、いくつかの(最低でも二つの)領域  $R_{i'} (i' = 1, \dots, M; 2 \leq M \leq N)$  が残っていることになる。

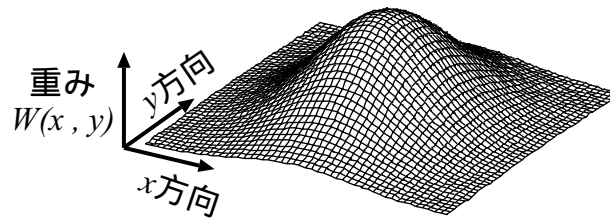


図 3.12: 重み関数

一つの物体の領域を抽出するために、画像の中心付近で大きな値を取る次のような重み関数  $W(x, y)$  を定義する (図 3.12 参照)。

$$W(x, y) = G\left((x, y); \left(\frac{w}{2}, \frac{h}{2}\right), \frac{\min(w, h)}{4}\right) \quad (3.7)$$

ここで  $(x, y)$  は各画素の座標、 $w, h$  は画像のそれぞれ幅と高さ、 $G(x, y)$  は二次元のガウス関数で、その平均は  $(\frac{w}{2}, \frac{h}{2})$ 、共分散行列は  $\begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ 、 $\sigma = \min(w, h)/4$  である。

各領域内で重みの和を計算し、最も大きい和を持つ領域  $R_I$  を抽出する。つまり、

$$I = \operatorname{argmax}_{1 \leq i' \leq M} \sum_{(x, y) \in R_{i'}} W(x, y) \quad (3.8)$$

これによって、画像全体に広がっている背景領域ではなく、中心付近にある物体の領域を抽出できる。ここでは物体を一つだけ取り出すと仮定しているが、もし二つ以上の物体が存在していることがわかっているならば、この重み和の大きい方から順に領域を抽出する。

## 3.6 実画像を用いた実験

本手法を計算機上に実装し、実画像に対して実験を行った。表 3.1 に、実験の際に用いたしきい値の初期値、上限、更新比率を示すが、初期値と上限は 3.4 節の議論に基づいて、比率は経験的に決定した。また滑らかさの特徴量のパラメータ  $\ell$  は 20 とした。初期領域分割のクラスタリングにおける 5 次元ベクトルの重みは、画素の色  $(L^*, a^*, b^*)$  に 1.0、位置  $(x, y)$  に 0.01 とした。以下の実験では意図的に物体を配置した画像も用いているが、特に記述がない場合は画像内に抽出すべき物体は一つであると仮定している。

実装は C++ を用いて PC (200MHz CPU) で行った。計算時間は画像のサイズやテクスチャにも依存するが、およそ、初期分割に 20 秒から 1 分ほど、併合に 3,4 分であった。

### 3.6.1 領域分割過程

図 3.13 に人物が写っている画像に対しての結果を示す。(a) が原画像 (432×301)、(b) がその初期領域分割結果である。領域分割では 59 の部分領域が生成されており、その境界

線を黒線で表示してある。(c)~(k) は領域併合過程の様子であり、各画像の下の数字はその時点で残っている領域の数を表す。背景の擬似輪郭は(f)までにほぼ取り除かれており、物体と背景の間に生じた小領域も併合が進むにつれて取り除かれている((c)(d)参照)。(k)に示すように、最終的に画像は人物と背景に対応する二つの領域に分割され、抽出処理によって背景ではなく物体の領域(図3.13(1)における黒い領域)が抽出された。この最終結果では、目や口や鼻のような顔のパーツも、人物という物体の領域に属しており、顔や服のハイライトや影の部分も一つの領域に併合されている。

### 3.6.2 他の領域分割手法との比較

別の画像を用いた実験結果を図3.14に示す。(a)が原画像(116×261)、(b)が初期領域分割結果である。最終的な併合結果(j)では、上半身と下半身は一つの物体に属するべきであるが、二つの領域に分けられている。これは、腕が胴体に垂直に横切っているために、連続的に接続しなくなっているためである。

図3.14(1)は文献[Zhu96]の領域分割手法による結果で、個々の色の異なるパーツ、つまり、腕、服、髪の毛などがそれぞれ一つずつの領域として分割されている。これは、人物像という一つの領域として抽出されるべき物体が、3.2節で述べた「一つの物体表面が複数の異なる反射率(色)をもつ」という物理的要因によって、複数の領域に分割されてしまっている状況であると言える。

本手法の結果(図3.14(k))では、髪の毛や顔、服、腕などの領域が併合されて一つの領域として、また服の色が異なっているために分割されていた領域が人物の下半身の領域として一つに併合されていることがわかる。人物の形状などの対象物体に関する知識を用いずに、服や肌の色が大きく異なるという物理的要因を排除したことで、上半身と下半身をそれぞれ一つの領域とすることができたことから、物理的要因を排除して物体の領域を抽出するという観点から見れば、本手法は有効であるといえる。

### 3.6.3 その他の領域抽出結果

その他の実験結果を図3.15に示す。どの画像についても、物体の領域として抽出されると思われる領域が大体抽出されている。

(a)(b)(c)では、さまざまな複雑な背景に対しても人物の領域がほぼ特定されているが、人物領域の一部が背景と併合されてしまっている。これは、(b)では人物の頭髮と背景の写真がどちらも暗いため、また(c)では背景の壁にあるタイルの溝の領域と人物の頭髮の領域とがともに暗い色であるため、初期領域分割時に背景と頭髮の領域が一部ですでにつながってしまっていることが原因である。そのため、特徴量のしきい値などのパラメータを調整しても良好な結果は得られないが、初期分割を改善すれば抽出結果を向上させることができると思われる。また(b)では、初期分割時に人物の服のテクスチャによって服の領域

表 3.1: しきい値の初期値と更新の上限値、更新比率。  $Th_s$  は  $\ell=20$  の時の値。

$Th_f$	$Th_f^{(0)}$	$Th_f^E$	$r_f$
$Th_a$	1/1200	1/30	1/0.9
$Th_s$	0.5	0.65	1.01
$Th_c$	5	20	1.01

が多数の細かい領域に分割されているので、3.3.4節で述べたように、本手法で用いたような色による初期分割ではなく、テクスチャによる統計的な手法による初期分割が適していると思われる。

(d) は最も簡単な画像であるが、亀の黒い目と緑の胴体が一つに併合されており、この結果は本手法の意図するところである。ただし、影の部分が併合されてしまっている。

(i) と (f) では抽出された領域には穴がある ((f) の方の穴は閉じてはいないが) ため、抽出領域数をそれぞれ3と2とした。二番目と三番目の領域はそれぞれ濃灰色と薄灰色で示してある。このような穴が存在するのは、用いた包含の定義によるものである。包含を定義する際に (i) では二つ以上の領域を同時に含んでいる、(f) では完全に囲まれていなくとも包含されている割合が大きいかどうか、ということを考慮することで避けられると思われる。

(g) の結果では本は一冊だけ抽出されているが、3.5.3節で述べたように仮定を変えることで、両方とも抽出することができる。

(j) と (k) の違いはホチキスがあるかないかだけであるが、結果は大きく違っている。この理由としては以下の二つが考えられる。実際の連続性の特徴量はT接点付近でのみ角度を求めている (図 3.9) ので、T接点から少し離れたところで境界線が曲がっている場合などは考慮していない。滑らかさの特徴量は、境界線が  $\ell$  より短ければ計算できず、境界線が長くなければ安定して計算できないので、物体内で細くなっている領域の短い境界線に対しては擬似輪郭かどうかの判定が正確ではない。以上のような実装上の問題点の解決は今後の課題である。

### 3.7 特徴量の有効性についての考察

一般的に、領域分割結果の評価を行うことは容易ではない。しかし、実験結果において良好な結果を得るために、どの特徴量がどの程度寄与したのかを分析することは、本手法の適用範囲を考察するために必要なことである。

まずどの境界線がどの特徴量によって併合されたのかを示すために、用いられた特徴量

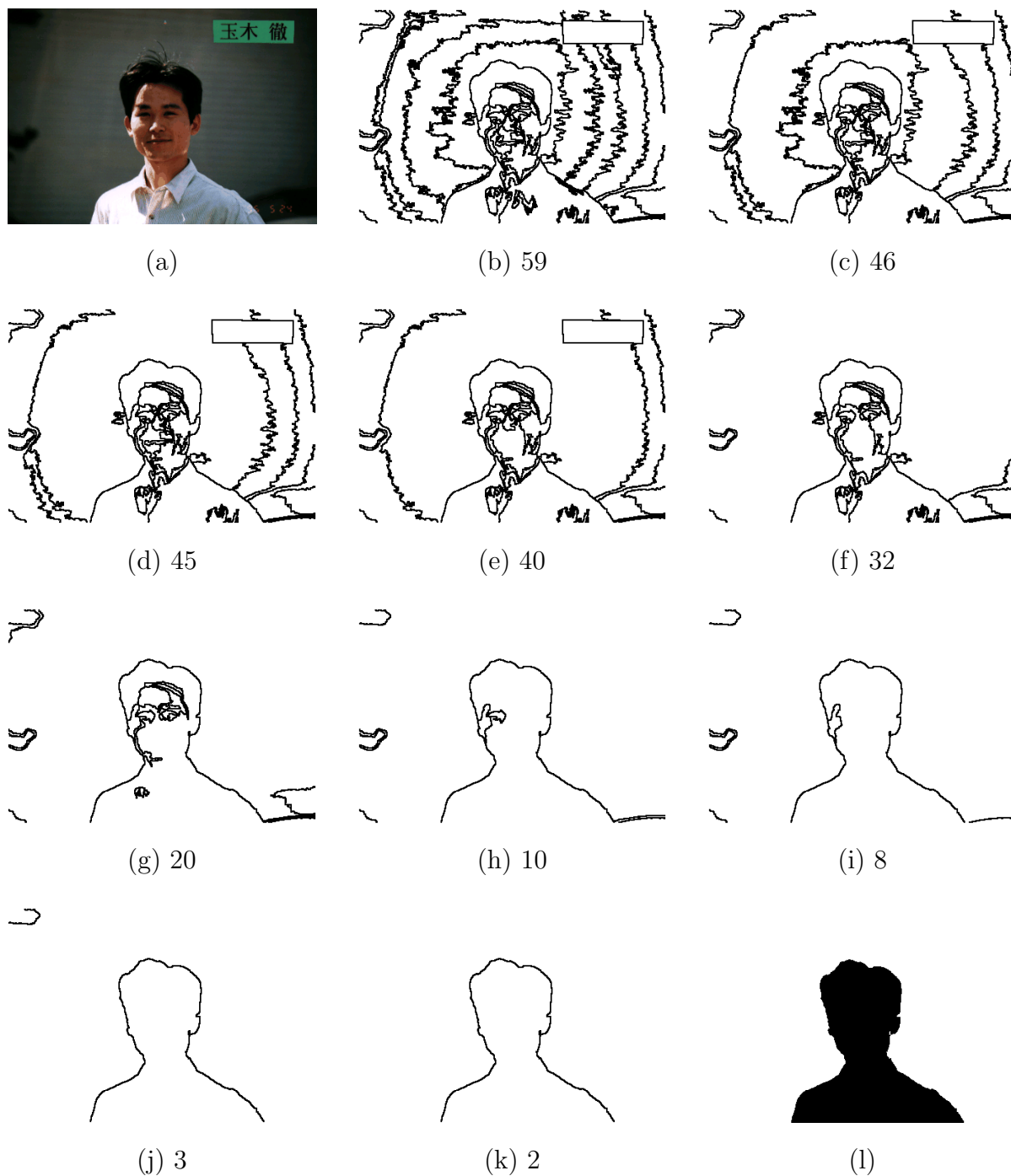


図 3.13: 実画像を用いた実験結果 (a) 原画像 (b) 初期領域分割 (c)~(k) 併合処理の様子。数字は領域数。(l) 抽出結果

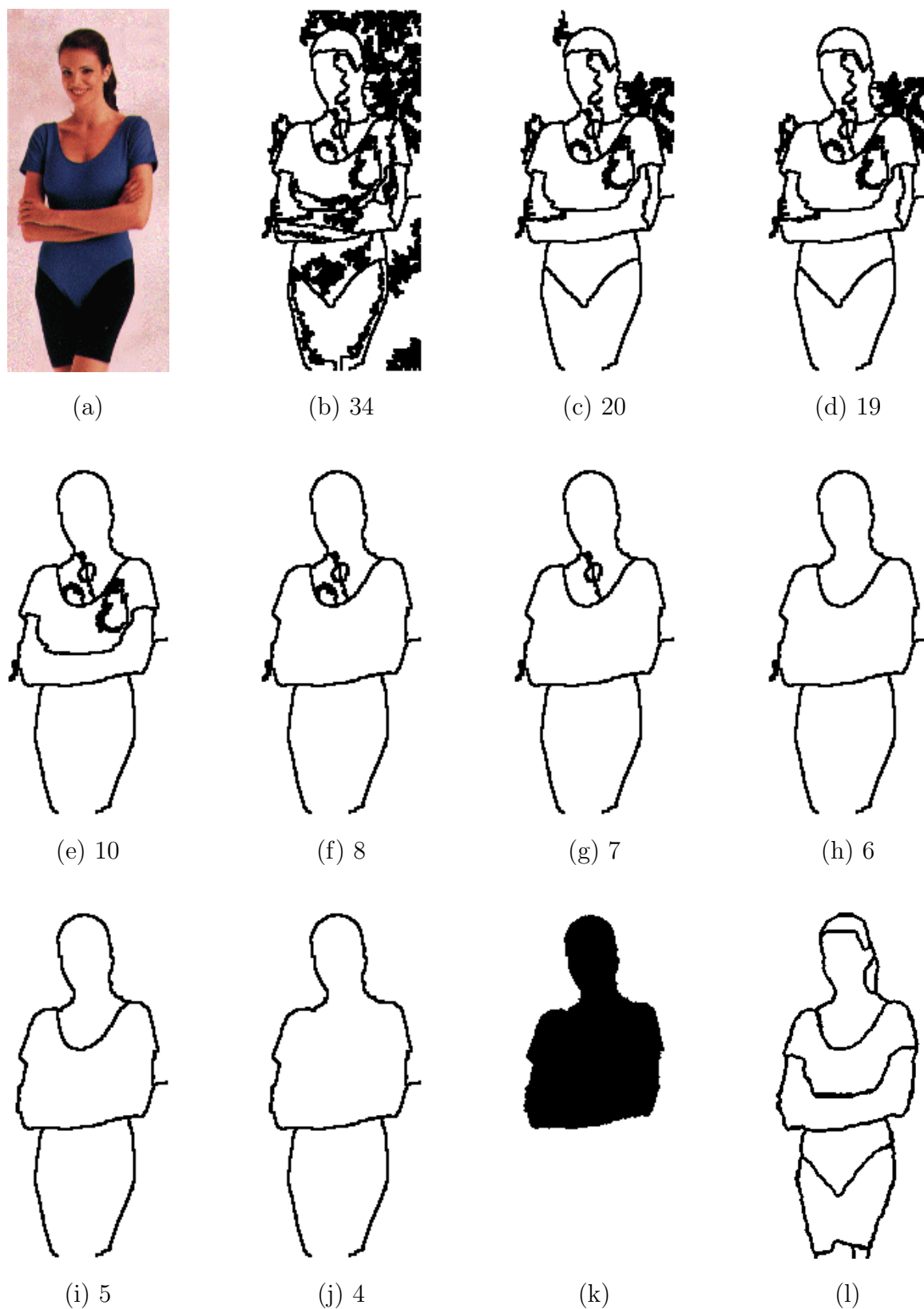
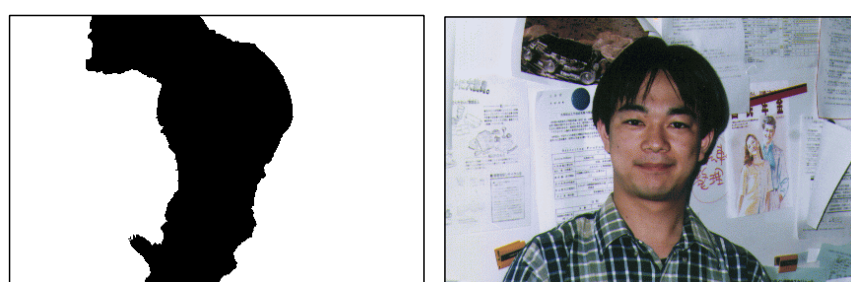


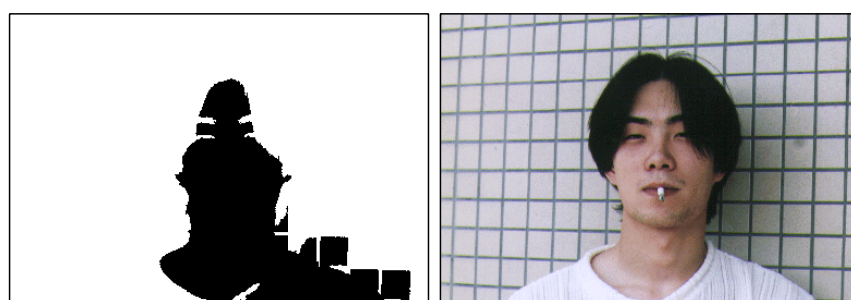
図 3.14: 実画像を用いた実験結果 (a)原画像 (© 1996 IEEE) (b)初期領域分割 (c)~(j)併合処理の様子。数字は領域数。(k)抽出結果 (l)文献[Zhu96]の結果 (© 1996 IEEE)



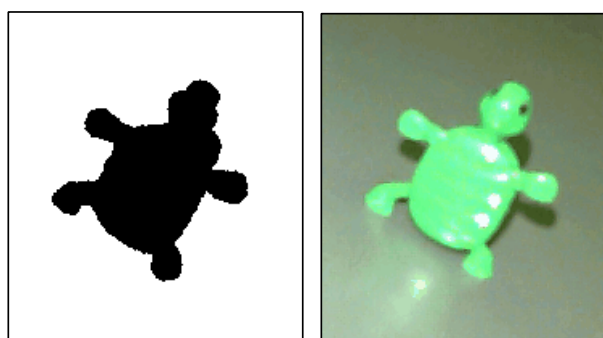
(a)



(b)

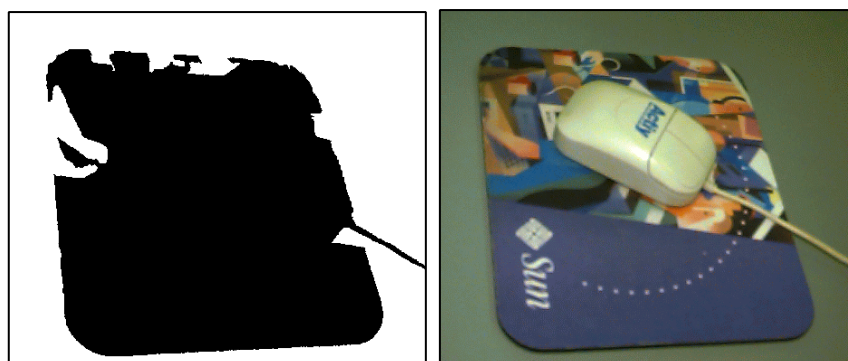


(c)

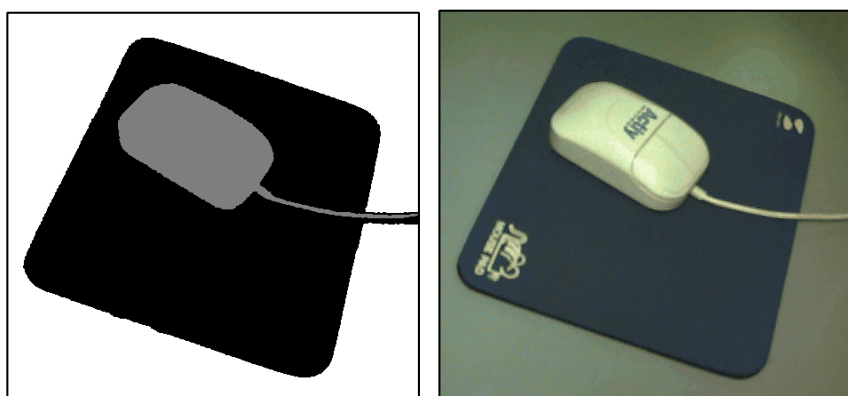


(d)

図 3.15: 実画像を用いた実験結果 (右) 各画像の原画像 (左) 抽出結果



(e)



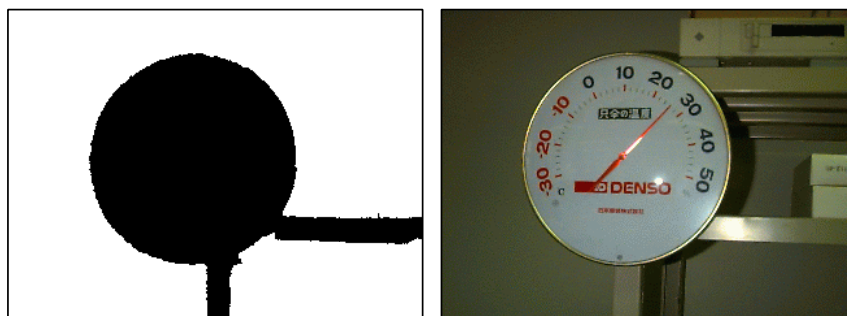
(f)



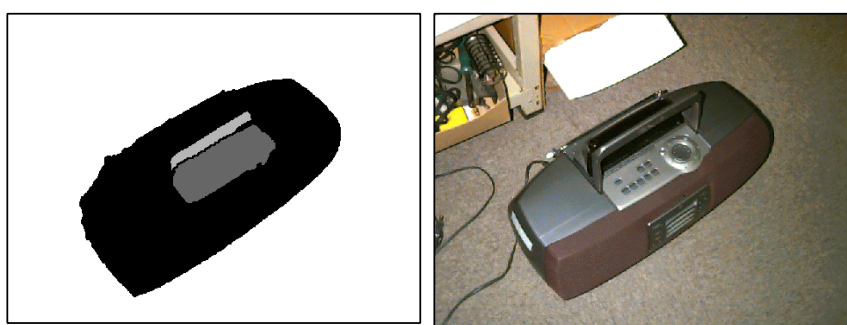
(g)

図 3.15: ( 続き ) 実画像を用いた実験結果 ( 右 ) 各画像の原画像 ( 左 ) 抽出結果

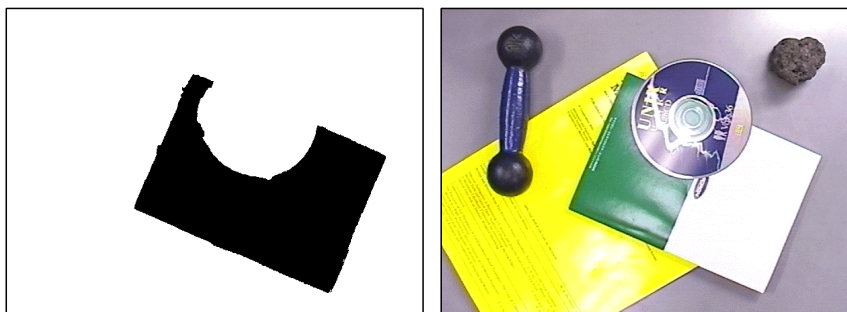




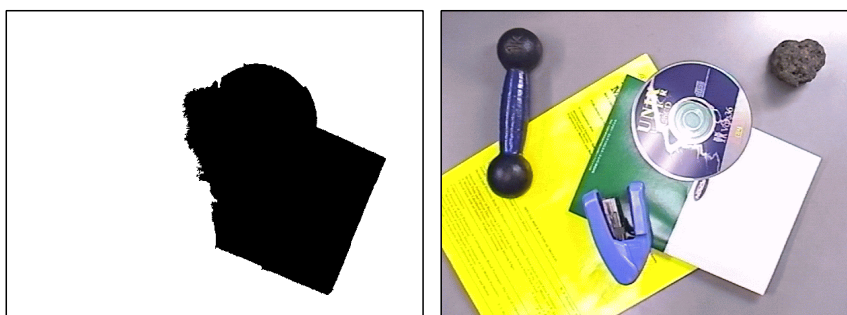
(h)



(i)



(j)



(k)

図 3.15: ( 続き ) 実画像を用いた実験結果 ( 右 ) 各画像の原画像 ( 左 ) 抽出結果

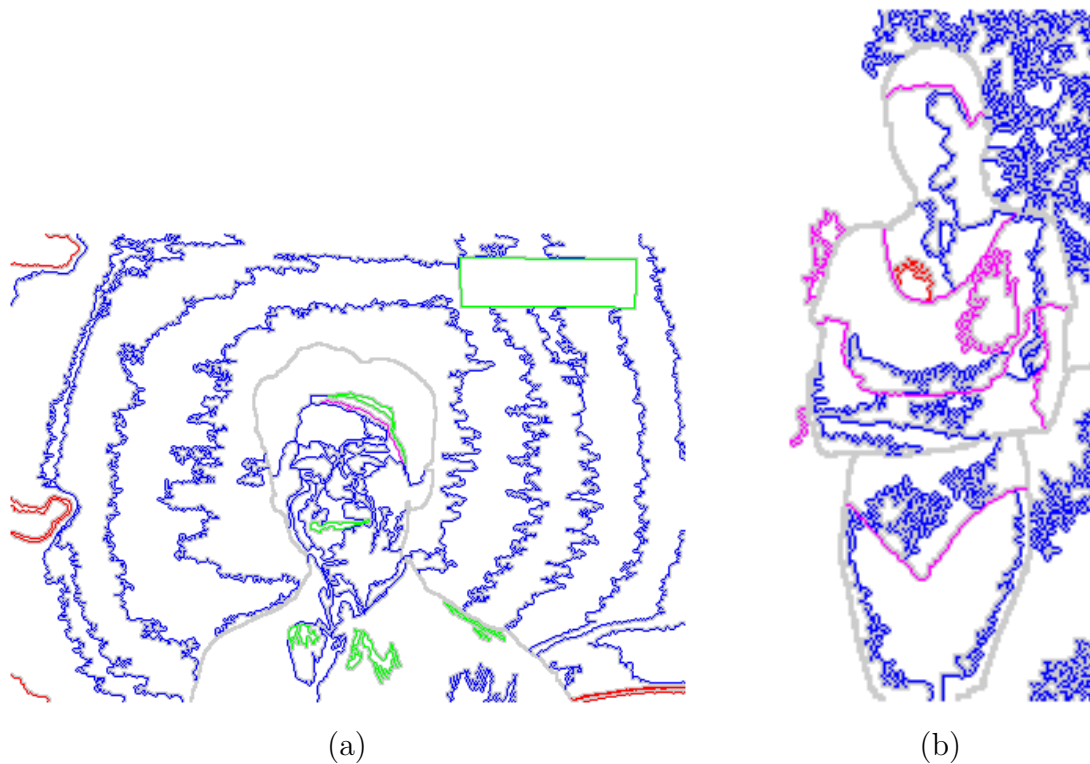


図 3.16: 併合に用いられた特徴量によって色分けされた (a) 図 3.13 と (b) 図 3.14 の境界線。  
 緑: 包含、青: 滑らかさ、紫: 連続性、赤: 面積比

毎に色分けした境界線を図 3.16 に示す。この図では、例えば、青い境界線で接した二つの領域は滑らかさの特徴量によって併合されたことを示す。図 3.16(a) をみると、背景の緩やかな輝度勾配によって生じた擬似輪郭が青く色分けされていることから、それらは滑らかさの特徴量によって排除されたことがわかる。これは図 3.16(b) でも同様にみることができ、さらに、髪と顔、胸部と服、脚と腰の領域が連続性の特徴量（紫）によって併合されていることがわかる。一方で、例えば図 3.16(a) の髪と顔の領域が滑らかさの特徴量（青）で併合されているように、予期しない併合が行われた場合もあった。しかし全体的には、これらの画像では、滑らかさと連続性の特徴量が意図したような個所で領域併合の特徴量として働いている。

このことを定量的に示すために、あらかじめ与えた理想的な分割との差（理想分割に近づけるための処理コスト [Vincken97, Koster97]）を定義し、領域併合処理を行う毎に、現在の分割結果と理想分割との差を計算する、という評価を行った。つまり、ある特徴量を用いた併合の際にこの差が大きく減少すれば、その特徴量が有効であったといえることになる。

理想分割は、初期領域分割によって生成された領域のうち、抽出されるべき物体の領域であると人間が判断した領域を塗りつぶすという方法で生成した。このようにすることで、併合処理を行っても、併合結果の境界線が理想分割の境界線とずれることがなくなる。

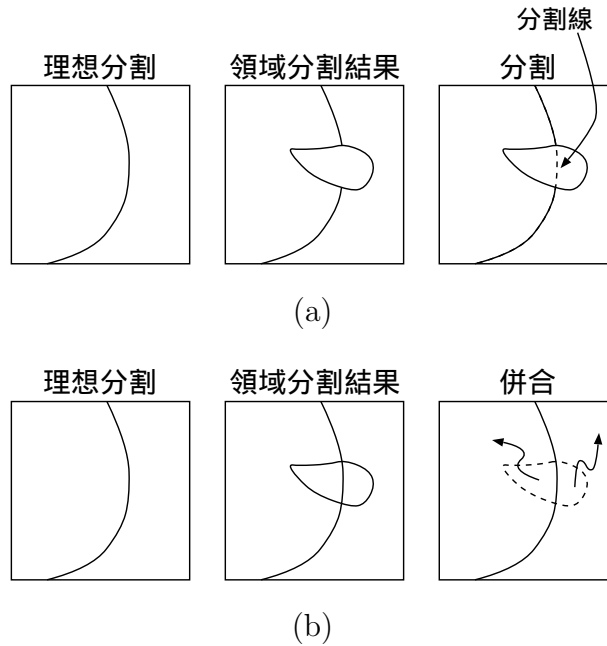


図 3.17: 二種類の理想分割との差 (a) 理想分割での二つ以上の領域にまたがっている場合 (b) 現在の分割結果で二つの領域に分かれている場合

理想分割との差は以下のように定義した。

- 現在の分割結果でのある領域が理想分割での二つ以上の領域にまたがっている時 (図 3.17(a))、その領域を理想分割の境界に沿って分割すれば差が減るので、その分割線の長さ  $C$  の二乗に比例した値 (ここでは  $C^2/4\pi$  とした<sup>3</sup>) を差とする。
- 理想分割でのある領域が現在の分割結果で二つの領域に分かれている時 (図 3.17(a))、その二つの領域を併合すれば差が減るので、その二つのうち小さな方の領域の面積を差とする。

以上の差を用いて考察を行なうが、図 3.14の結果は領域数が少なく、実際には包含と面積比の特徴量が用いられなかったため、以下では図 3.14の結果のかわりに、初期分割のクラスタ数を変え領域数を増やした初期分割 (図 3.18(a)) の併合結果 (図 3.18(b)) を用いることにする。

図 3.18の分割結果 (b) がどのように理想分割 (c) に近づいているのかを表したグラフを図 3.20に示す。同時にこのグラフには、どの特徴量がいつ用いられたかの累積回数を表示しており、どの特徴量が用いられた時にどれだけ差が減っているかがわかる。その使用回数と、併合一回あたりの差の減少量を表 3.2に示す。

<sup>3</sup>ここでは、領域が円のときに二種類の差が対等に扱われるように決定した。すなわち、周囲長  $2\pi r$  を分割線  $C$ 、面積を  $\pi r^2$  として、 $\pi r^2 = \frac{4\pi^2 r^2}{4\pi} = \frac{(2\pi r)^2}{4\pi} = \frac{C^2}{4\pi}$  である。実際には係数を変えてもそれほど影響がない。

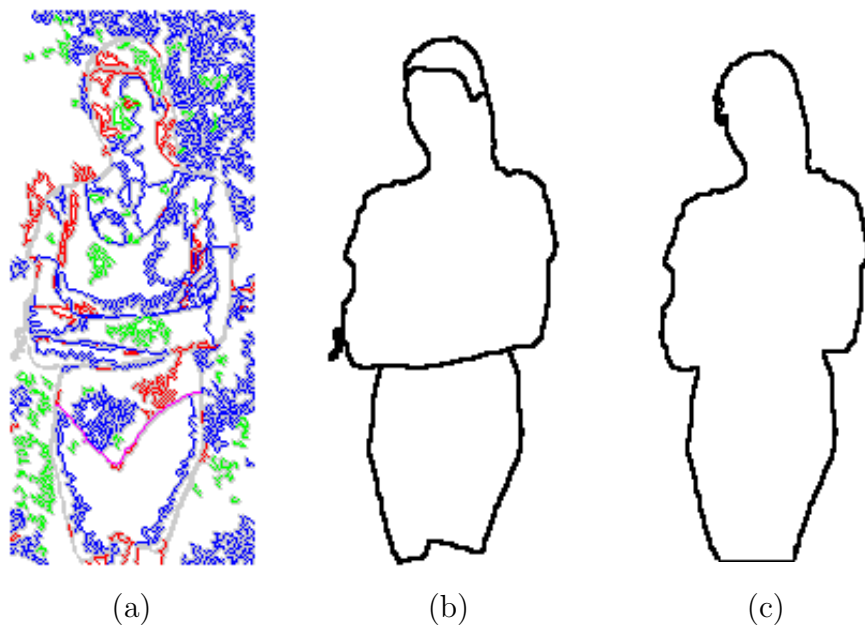


図 3.18: 図 3.14(a) の初期分割での領域数を 197 とした (a) 初期分割。図 3.16 と同様に色分けしてある。(b) 併合結果。(c) 理想分割。

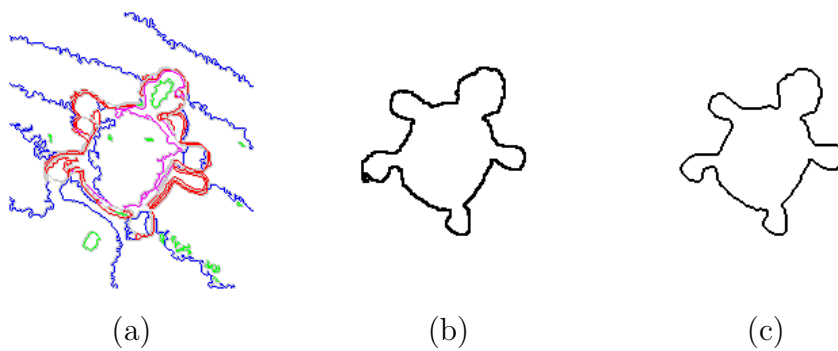


図 3.19: 図 3.15(d) の初期分割での領域数を 122 とした (a) 初期分割。図 3.16 と同様に色分けしてある。(b) 併合結果 (c) 理想分割

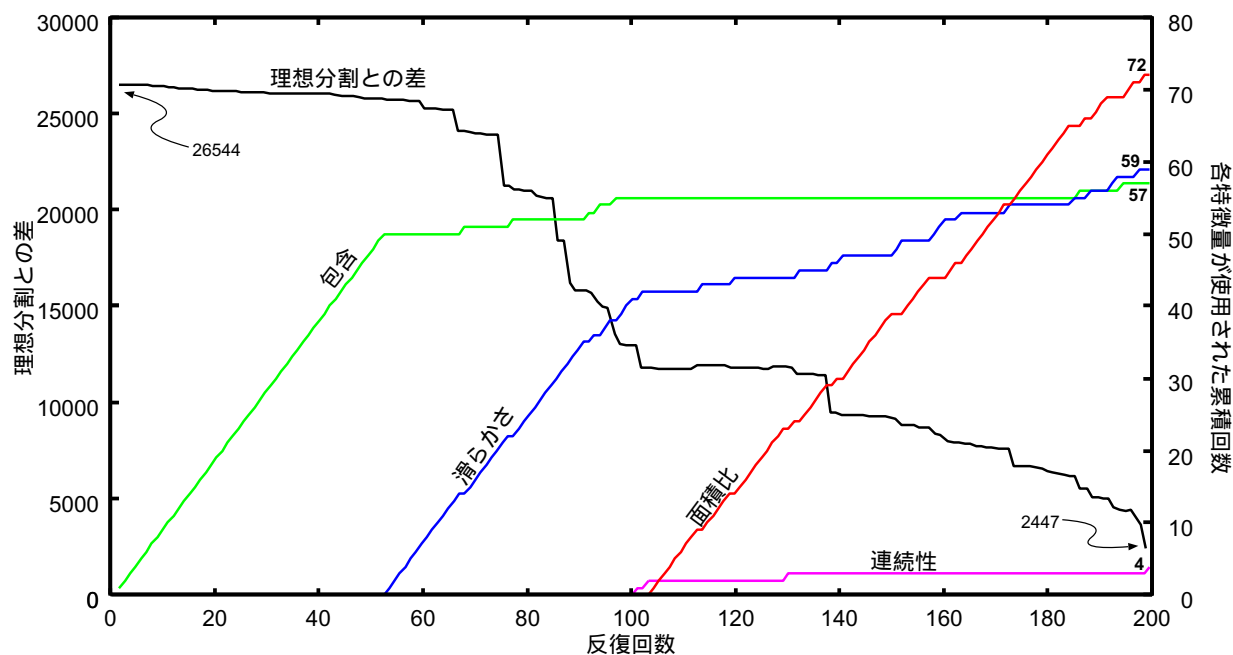


図 3.20: 図 3.18(b) の併合過程における理想分割との差。横軸は併合回数、縦軸（左）は理想分割との差、縦軸（右）は特徴量が用いられた累積回数。

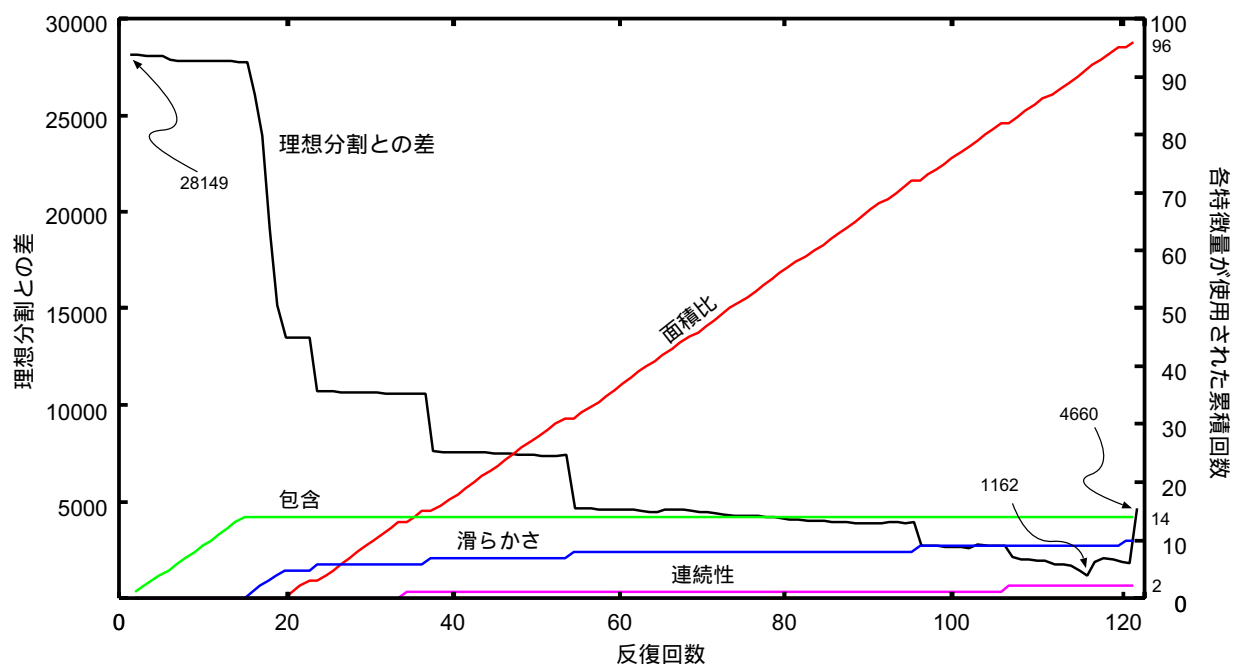


図 3.21: 図 3.19(b) の併合過程における理想分割との差。横軸は併合回数、縦軸（左）は理想分割との差、縦軸（右）は特徴量が用いられた累積回数。

表 3.2: 図 3.20において各特徴量が併合に使われた回数と併合一回あたりの差の減少量

	包含	面積比	滑らかさ	連続性	合計
使用回数	57	72	59	4	192
差減少量	18.5	24.9	319	607	

表 3.3: 図 3.21において各特徴量が併合に使われた回数と併合一回あたりの差の減少量

	包含	面積比	滑らかさ	連続性	合計
使用回数	14	96	10	2	122
差減少量	31	-16.4	2405.5	286	

表 3.2をみると、滑らかさの特徴量は差を大きく減らしており、結果を理想的な領域分割に近づけている。連続性の特徴量は、包含や面積比よりも比較的高いレベルの特徴であり数回しか使われていないが、滑らかさと同じように大きく差を減少させている。大きく差を減らしているこの二つは結果を良好にするために有効な特徴量であるといえる。ただし、表 3.2では連続性の方が滑らかさよりもより差を減少させて有効である。

一方、面積比の特徴量はそれほど差を減少させないため、この評価からはあまり有効ではないといえる。しかし、併合の際に最も多く用いられる特徴量であるので、小さな領域を減らすことに有効であると思われる。

他の画像について行なった結果を表 3.3に示す。これは、図 3.15(d)の初期分割を、図 3.18(a)と同様に領域数を増やし、その併合結果である図 3.19(b)がどのように理想分割図 3.19(c)に近づいているのかを表すグラフである図 3.21を用いて得られた。滑らかさと連続性の特徴量が差を減らすのに有効であることと、面積比の特徴量の使用される回数が多いことは、表 3.2の結果と同じである。

他の画像についても行なった結果同様の傾向が見られたが、滑らかさと連続性の特徴量のうち、どちらの特徴量がより有効かは画像により異なった。

### 3.8 従来手法との比較

図 3.14の実験において文献 [Zhu96]の手法との結果を比較したが、そのような洗練された手法ではなく、もっとシンプルで従来からある代表的な領域分割手法に比べて、本手法の



図 3.22: 大津のしきい値による二値化

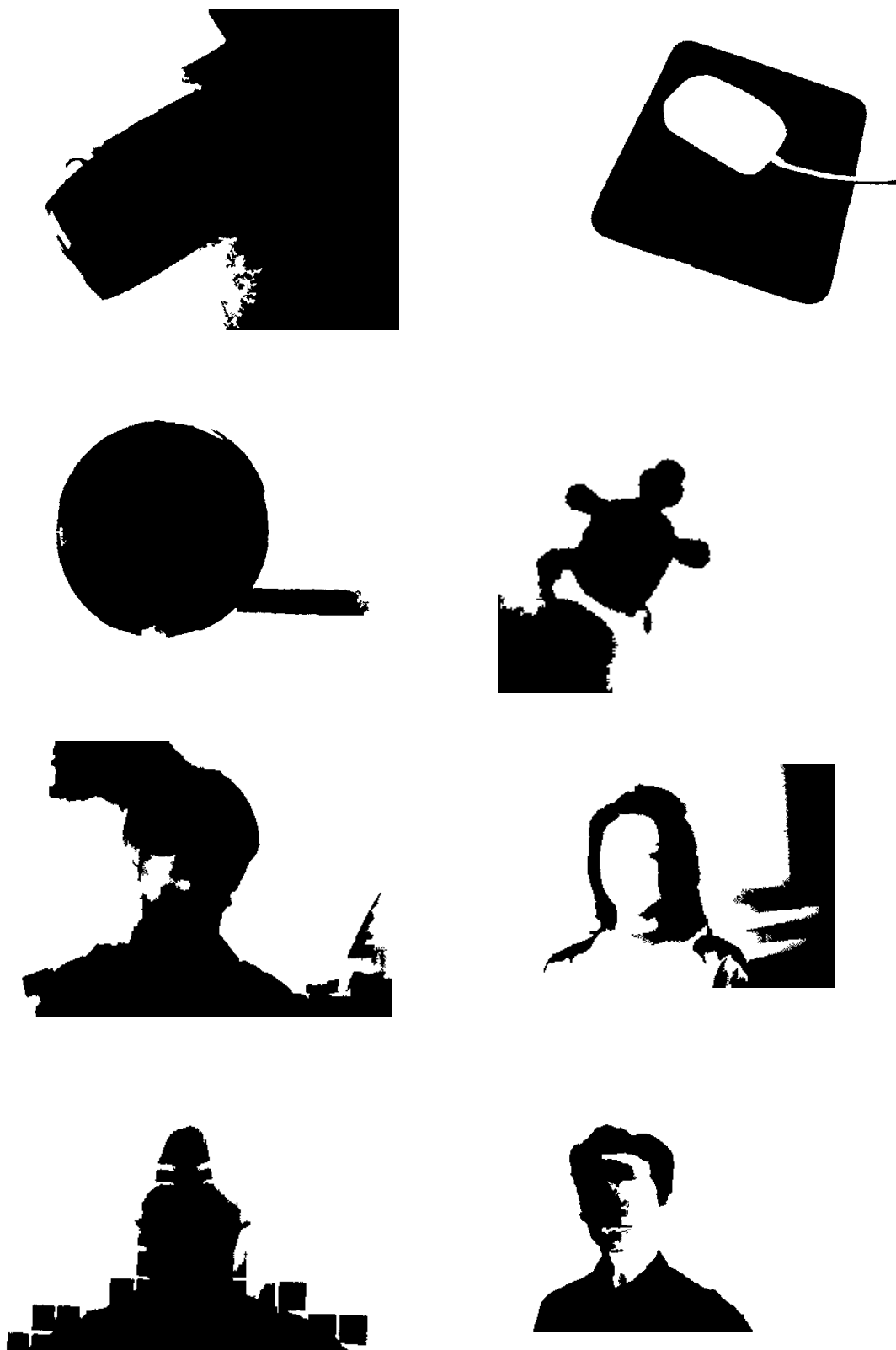


図 3.23: 輝度分布を評価し併合する分割併合法による領域分割



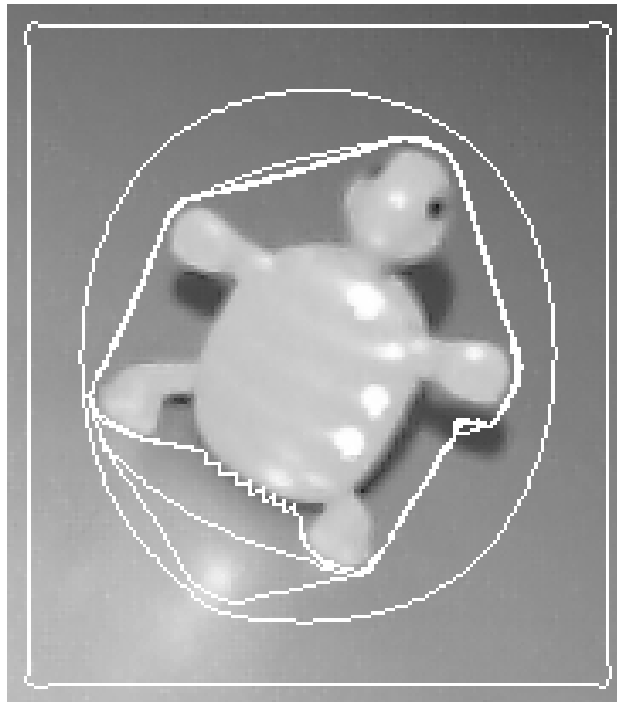


図 3.24: 動的輪郭モデル Snake の収束過程。最も外側の白線が初期輪郭。

性能がどの程度のものであるかを評価することも重要である。先の実験に用いた画像のうちいくつかに対して、2章で述べたしきい値による二値化(2.2節)と、輝度分布を評価し併合する分割併合法(2.4節)を適用した結果を示す。

図3.22にしきい値により輝度値を二値化した結果を示す。もともと二値化処理は領域分割に適した方法ではないにもかかわらず、物体と背景のコントラストが強い画像に対しては、物体がはっきりと分離されている。しかし、画像全体に渡って緩やかに輝度が変化しているような場合では、背景を一つのしきい値で分離することは困難であるのがわかる。

次に図3.23に、輝度分布を評価する分割併合法を用いて領域分割を行った結果を示す。画像の分割には先の実験と同じ初期分割を用いた。併合処理は、すべての二つの隣接する領域に対して尤度比を計算し、最も小さい値を取る二つの領域を併合するという操作を、領域数が2になるまで繰り返した。

この手法も、物体と背景のコントラストが強い画像では、背景と物体の領域に分割することができる。しかし、適切な結果が得られるための併合のしきい値は画像によって大きく異なり、すべての画像に対して同じしきい値を適用することは困難である。したがって、しきい値を大きくして併合を促進させると、背景の領域同士が併合される前に、背景と物体の領域と併合されてしまうことは避けられない。

これらに比べて本手法はどの画像に対してもしきい値の範囲や更新比率を変えることなく処理することができる。また併合のための特徴量は物理的要因を考慮しているため、コントラストに影響されずに物体の領域を抽出することができる。

また、図3.15(d)は背景が単純で物体が一つしか存在しないため、動的輪郭モデルであるSnake (2.6節参照)によるアプローチが考えられる。この画像にSnakeを図3.15(d)に適用した結果を図3.24に示す。初期輪郭が最も外側に配置され、画像のエネルギーとしてエッジを用いており、Snakeが物体のエッジに収束している過程を示している。図3.24の結果が示すように、Snakeはその一次/二次微分もエネルギーとして小さくするため、輪郭に凹凸のある物体の場合はその包絡が抽出される。圧力項や面積項を導入したとしても、輝度差の小さいエッジ部分が抽出されなくなることもあり、パラメータの調整が難しい。したがって、正確な輪郭を抽出するという事は困難であると思われる。

### 3.9 おわりに

この章では、物体に固有ではない一般的な情報である配置・隣接・面積比情報を用いて、物理的要因及び光学的要因を排除し、物体を一つの領域として抽出する領域分割手法を提案した。本手法は四つの特徴量、つまり包含、滑らかさ、連続性、面積比の特徴量を用い、一つの物体の領域が物理的要因によって分割されたのかどうかを判定する。これらの特徴量は、一つの領域の内部についてのものではなく、二つの領域間についての特徴量である。そのため、これにより物体表面が様々な反射率を持っているというような場合でも、物理的要因を排除し物体の領域を得ることができる。そして併合のためのアルゴリズムを示し、実画像に対して実験を行った結果、物体の領域として抽出されと思われる領域をほぼ抽出することができた。

本手法で用いた領域間の情報は経験的なものであり、導入した特徴量が最適でそれら以外には存在しない、ということはない。一般的な情報を用いて物理的・光学的要因を排除するという事を念頭に、他の特徴量をどのように取り入れていくのかは今後の課題である。特徴量を統合する場合には、本手法のように直列に判断していく方法と、統計的性質を考慮し並列に判断していく方法が考えられるが、それらをうまく組み合わせることが必要になるだろう。また、本手法の併合結果はその順序に依存するが、どの特徴量で併合されたかを保存することはできる。その情報を解析し、階層的な木構造 [Salembier00] などで保持された領域分割結果を用いて、物体の認識 [Arita95] や探索などへの発展が考えられる。

## 第4章 人物形状の知識を利用した 画像系列からの人物領域の抽出

### 4.1 はじめに

カメラから得られる映像は、写真のような静止画像だけでなく、ビデオカメラで撮影される動画像もあり、これもまた有用な情報源である。計算機でこのような動画像を処理したいという要求は高まっており、近年のCCDカメラの低価格化とPCの普及も手伝って、幅広い分野で応用が期待されている。写真の被写体となる対象は人物から花や風景などの静物まで幅広いが、ビデオに撮られる興味の対象の多くは、当然のことながら動くものである。工学的な応用分野に限って言えば、様々な目的のために人物を被写体にすることが多い。ここでもやはり、多様な状況で撮影された画像から対象を抽出することが必要である。前章では静止画に対する一般的な対象物体の抽出手法について述べたが、この章と次章においては、対象として人物が撮影されている動画像を扱う。

まずこの章では、カメラによるセキュリティ監視システムや計測などに多く用いられている、移動する人物の全身像を抽出する手法について述べる。カメラで撮影された物体が何であるかを判断し、侵入者であれば警告するようなシステムを実現するとき、動画像から移動物体を検出し、その中から人物の領域を自動的に抽出する処理は重要な役割を果たす。

動画像を用いる場合、画像での動きの情報を用いて移動する物体が存在するかどうかを知ることが基本となる。1.4.1節でも述べたように、これまでも画像中の動物体を検出する研究は数多くなされている。その中でも背景参照画像を用いる手法は、処理が単純であるが、天候や時刻の違いによって参照画像と現在の画像とで照明が大きく変動すると、移動物体のみを抽出することが困難になるという問題点がある。そのために参照画像を動的に更新していく手法[松下98, 波部98]もあるが、背景に写る車なども移動物体と見なされてしまい、特定の対象以外を背景であると認識することは容易ではない。また連続するフレーム間の差分を用いる手法[安居院92]は参照画像を用いるときのように照明に大きく影響されることはないが、移動速度の遅い物体は抽出されにくく、人物以外に移動する物体が背景にある場合には全てを抽出してしまうため、差分のみでは人物抽出にそのまま適用することはできない。差分を用いずに二次元画像との時間方向の三次元から人物の特徴的

な歩行パターンから人物像を検出する手法 [Niyogi94] もあるが、抽出には多数のフレームが必要となり、フレーム間差分のような迅速な抽出が困難である。またこれらの研究の多くは、シーン内の動物体の数を一つに限定し、人物像の解析・動作認識のために領域を抽出している [Ohya94, Guo94]。また、複数動物体の研究もあるが、全て同じ種類の物体、例えば全て人物で、それぞれの人物を抽出している [緒方98, Haritaoglu98]。しかし、監視システムの対象とする実環境では、人物以外に車や動物などが存在することもあり、複数の異なる種類の移動物体から人物以外を排除することが必要である。

そこでこの章では、人、車など、複数の移動物体が存在するときに、人物領域を抽出する方法を提案する。移動する人物の抽出のためにフレーム間差分を用い、背景に人物以外の動物体がある場合でも人物領域を得るために、抽出される人物像とその頭部についての簡単なモデルを用いる。

人物抽出のアルゴリズムの大まかな流れを示す。まずフレーム間差分を求めて、注目フレーム中の動きのある領域を抽出する。次に画像から求めたエッジを曲線、直線に分類し、曲線を円の一部分として表す。そして人物の頭部を「動いている」「曲線部に囲まれている」とすることにより、差分と曲線エッジから画像中の人物頭部に相当する部分を検出する。そこで求めた位置に人間の簡易的な輪郭モデルおき、Snake を用いて収束させて人物の大まかな外形を表す閉包を得る。最後に人物領域の正確な輪郭を得るために領域分割を行ない、閉包内部の分割されている領域を併合し、最終的な人物の領域を抽出する。

以下この章では、4.2節で差分による動領域抽出について、4.3節と4.4節では人物頭部を表すエッジによる円の検出について述べる。また、簡易人物モデルについては4.5節で、領域分割による詳細な領域抽出については4.6節で説明し、4.7節で本手法を実画像に用いた実験とその結果について述べる。

## 4.2 差分による移動領域の抽出

移動している人物の領域を特定するために、フレーム間差分による動領域抽出を行なう。まず入力であるカラー画像を、次式を用いて輝度画像に変換する。

$$I_t(x, y) = 0.299R_t(x, y) + 0.587G_t(x, y) + 0.114B_t(x, y) \quad (4.1)$$

ここで  $I_t(x, y)$ ,  $R_t(x, y)$ ,  $G_t(x, y)$ ,  $B_t(x, y)$  はそれぞれ、時刻  $t$  の画像  $I_t$  の座標  $(x, y)$  にある画素の輝度値と  $R, G, B$  の値である。

次に同一画素のフレーム間のノイズを減らすために、各フレームを Gaussian フィルタを用いて平滑化する。そして画像系列の注目する時刻  $t$  の画像  $I_t$  とその前後の時刻  $t+1, t-1$  の画像  $I_{t-1}, I_{t+1}$  の3枚から、まず2枚ずつの差分画像  $I'_{-t}, I'_{+t}$  を求める。この2枚の差分画像に対して、しきい値以上を1、それ以下を0とし、二値画像  $I''_{-t}, f''_{+t}$  を求める。つぎにこれらの論理積をとり、注目する時刻  $t$  の画像  $I_t$  に対応するフレーム間差分画像  $I'_t$  を求

める。

この二値画像  $I'_t$  の画素値1の領域が、画像上で動いている物体の領域を表す。ここで、面積が一定以下の小領域はノイズによって生じたとみなせるため、収縮処理により削除する。また同じ理由で、領域に穴がある場合は拡散処理によって穴を埋める。こうして求めた移動物体領域の重心座標  $G(g_x, g_y)$  を計算し、この座標を移動領域の中心位置とする。

## 4.3 エッジの曲線部抽出

### 4.3.1 交点のないエッジの抽出

画像のエッジを抽出するために、輝度画像に対して Sobel フィルタを用い微分画像を計算する。この微分画像を二値化してエッジを抽出する。二値化された微分画像を Hilditch の細線化アルゴリズム [安居院 92] を用いて細線化処理を施し、エッジを線として抽出する。

しかし細線化で得られたエッジは一般にひげと呼ばれる分岐した線が多い。そこで、以降の処理を行なうために、分岐の交点を検出し、交点を除去して分岐のないセグメント(線分)に分ける。

ここでは連結数 (conective number) [鳥脇 88, 安居院 92] を用いて、交点を検出する。連結数とは、ある図形 (画素値1) の画素が接する背景 (画素値0) の画素の数を指す。4近傍の連結数  $N_c^4$  の値は0から4で、1は端点、3以上は交点を表し、次の式で計算される。

$$N_c^4 = \sum_{k \in \{1,3,5,7\}} B(P_k) \{1 - B(P_{k+1})B(P_{k+2})\} \quad (4.2)$$

ここで  $P_k (k = 0, 1, \dots, 8)$  は注目画素  $P_0$  の8近傍画素 (図 4.1) を、 $B(P_k)$  はその画素値 (0または1) を表す。

エッジ上の各画素に対してをこの連結数を求め、連結数が3以上である画素を交点とみなし、その画素をエッジ線から削除する。

### 4.3.2 $k$ -曲率による曲線エッジの抽出

一つのエッジセグメントは直線と曲線から成り立っている。ここでは  $k$ -曲率 [Shirai87] を用いて、エッジセグメントを直線と曲線に分割する。ある点での  $k$ -曲率とは、図 4.2 に示すように、その点とその点の  $k$  個前の点とを結んだ直線と、 $k$  個後の点とを結んだ直線のなす角で、点  $s$  での曲率  $\varphi(s)$  の近似である。

点列  $\{s\}$  に対して  $k$ -曲率  $\varphi(s)$  をプロットしたグラフを  $\varphi-s$  曲線といい、 $\varphi-s$  曲線の水平線が元の画像での曲率一定の曲線に (ただし  $\varphi(s) = 0$  の水平線は元の画像の直線に) 対応する。図 4.3 に示すように、画像上での曲線と、曲線と直線との接点は、 $\varphi-s$  曲線において大きな値 (すなわち曲率が大きい) を持つ区間に対応する。この区間は、最も大き

$P_4$	$P_3$	$P_2$
$P_5$	$P_0$	$P_1$
$P_6$	$P_7$	$P_8$

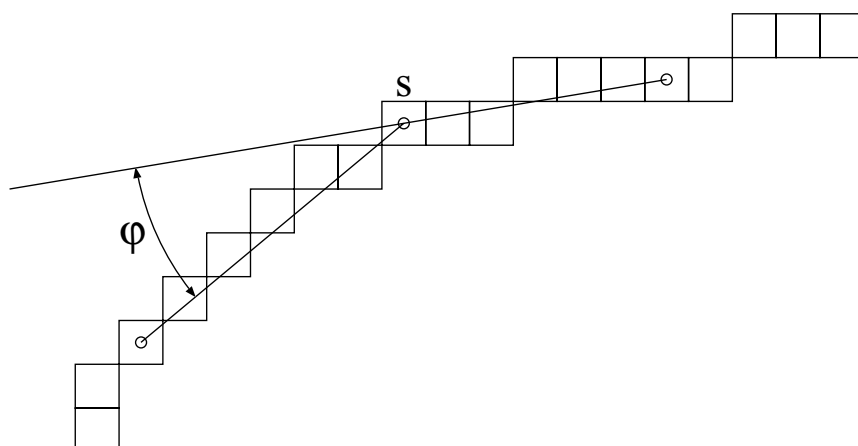


図 4.1: 8近傍隣接画素

図 4.2: 点  $s$  における  $k$ -曲率  $\varphi$  ( $k = 6$  のとき)

い曲率がしきい値以上で、かつその区間から外れた所で曲率が急激に小さくなるという特徴を持つ [Shirai87]。具体的には、その区間  $I_i = [L_i, R_i]$  は、区間内の曲率の最大値  $\varphi_m$  がしきい値以上の場合に次のように定義される。

$$\varphi_m = \max_{s \in I_i} \{\varphi(s)\} > Th \quad (4.3)$$

$$|\varphi(s)| \begin{cases} > h(\varphi_m), & \forall s \in I_i \\ < h(\varphi_m), & s = L_i - 1, R_i - 1 \end{cases} \quad (4.4)$$

ここで  $h(x)$  は  $x$  の線形関数

$$h(x) = ax, \quad 0 < a < 1 \quad (4.5)$$

である。

上の条件を満たす  $\varphi - s$  曲線の区間は、画像上では曲線（直線）と接点の両方に対応するため、接点の区間を除去して曲線（直線）の区間のみを得るアルゴリズムを以下に示す。このアルゴリズムは、画像上のセグメントの各画素の座標を連結数を基に追跡して一次元配列に格納するステップ（1.~3.）と、得られた配列の先頭から曲率の大きい区間を見つけしていくステップ（4.(a)~(d)）から構成される。

まず、得られている多数のエッジセグメントに対してラベリングを施す。そして各セグメント  $j = 1, \dots$  に対して、以下の処理を行なう。

[曲線または直線の区間を求めるアルゴリズム]

#### 1. 端点の検出

セグメント  $j$  に属し、かつ連結数が1（端点）である画素を見つけ、 $s_1$  とする。見つからない場合、連結数が2である点を  $s_1$  とする。

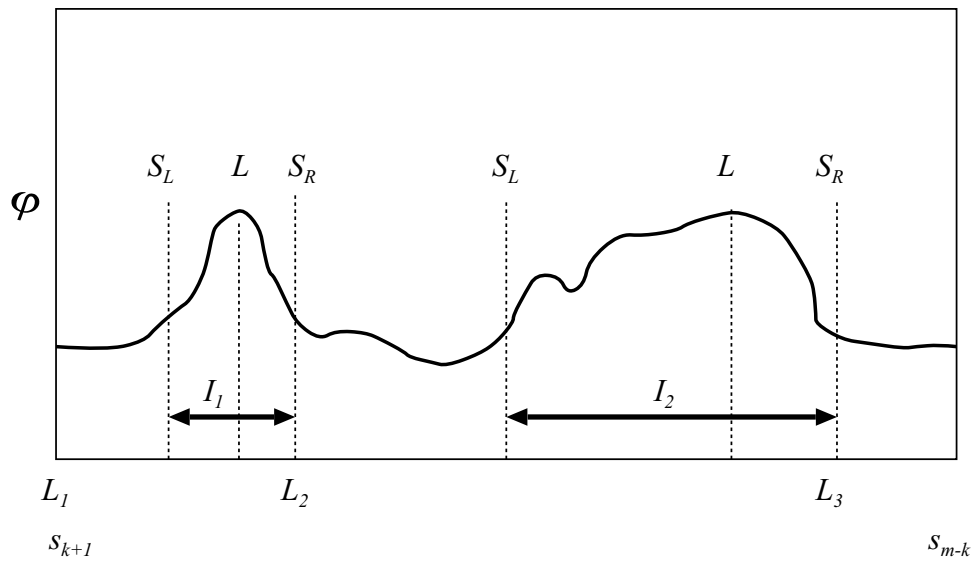


図 4.3: 点列  $\{s_{k+1}, \dots, s_{m-k}\}$  に対する  $\varphi - s$  曲線。区間  $I_1$  は曲率が大いいが短いので、画像上で曲線と直線の接点に対応する。区間  $I_2$  は長く、区間にわたって曲率が大いなので、画像上では曲線に相当する。

## 2. セグメントの追跡

$s_1$  に連結し、かつ連結数が 2 である画素を追跡する。追跡してきた点以外の連結数が 1 になる点  $s_m$  まで、追跡した画素の座標を配列  $s = \{s_1, s_2, \dots, s_m\}$  に記憶する。

## 3. $k$ -曲率の計算

セグメント  $j$  を構成する配列の  $s_{k+1}, \dots, s_{m-k}$  の各点において、 $\varphi(s)$  を計算する。

$i = 1$  として、最初の区間探索の始まり  $L_1$  を  $L_1 = s_{k+1}$  とおく。

## 4. 第 $i$ 番目の区間 $I_i$ の検出 (図 4.3 参照)

### (a) 曲率の大きい個所 $L$ の検出

$\max_{s \in [L_i, L]} \varphi(s) - \min_{s \in [L_i, L]} \varphi(s) > Th$  なる  $L \in [L_i, s_{m-k}]$  を見つける。

### (b) 区間の始まり $s_L$ の検出

$\varphi(s_L) > h(\varphi(L))$  かつ  $\varphi(s_L - 1) < h(\varphi(L))$  なる  $s_L \in [L_i, L]$  を見つける。

### (c) 区間の終わり $s_R$ の検出

$\varphi(s_R) > h(\varphi(L))$  かつ  $\varphi(s_R + 1) < h(\varphi(L))$  なる  $s_R \in [L, s_{m-k}]$  を見つける。

### (d) 区間の長さによる接点の除去

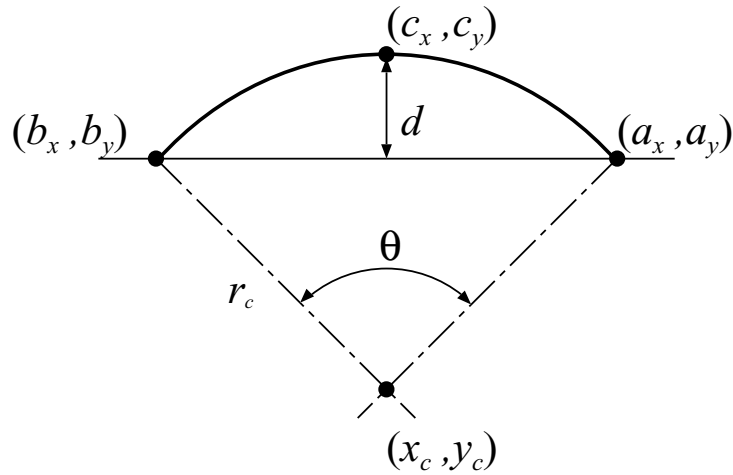


図 4.4: 曲線エッジを円弧と仮定した時の扇型

$s_R - s_L > th$  ならば区間  $[s_L, s_R]$  を曲線または直線の区間とする。そうでなければ接点であるとしてこの区間を除去する。

#### 5. 終了条件

$s_R = s_{m-k}$  ならば終了。

そうでなければ次の区間探索の始まり  $L_{i+1}$  を  $L_{i+1} = R_i + 1$  とおき、 $i \leftarrow i + 1$  として、ステップ 4 へ。

### 4.3.3 曲線と直線の分類

次に、前項で得られた各区間  $I_i$  に対応するエッジを、曲線と直線とに分類する。ある区間に対応するエッジを円弧であると仮定すると、図 4.4 における弧の角度  $\theta$  と高さ  $d$  について以下の式が導かれる [Shirai87]。

$$\theta = \varphi_a \frac{m-1}{k} \quad (4.6)$$

$$d = \frac{m}{\theta} \left(1 - \cos \frac{\theta}{2}\right) \quad (4.7)$$

ここで、 $m$  は区間の長さ、 $k$  は  $k$ -曲率を計算する長さ、 $\varphi_a$  は  $s_{k+1}, \dots, s_{m-k}$  についての  $k$ -曲率の平均値であり、

$$\varphi_a = \frac{1}{m-2k} \sum_{i=k+1}^{m-k} \varphi(s_i) \quad (4.8)$$

上式に基づいて以下のように各区間  $I_i$  を曲線と直線に分類できる [Shirai87]。

- もし  $m < 2k$  ならば曲率が計算できないため不定



- $m \geq 2k$  の時、 $d > d_u$  ならば曲線  
(しきい値  $d_u$  は明らかな曲線のみを区別できるために十分大きくする)
- $d \leq d_u$  の時、 $\theta < \theta_t$  ならば直線 ( $d$  が小さいので曲率も小さい)
- $\theta \geq \theta_t$  の時、 $m < m_t$  かつ  $d < d_l$  ならば直線  
( $m_t, d_l$  はしきい値、ただし  $d_l < d_u$  とする)
- そうでない時、曲線

#### 4.4 曲線エッジの円近似による頭部候補領域決定

エッジの各区間  $I_i$  が直線と曲線に分類されたら、曲線エッジを円の一部と仮定し、そのエッジを円弧で近似する。それぞれの曲線エッジの始点、終点とその中点をそれぞれ  $(a_x, a_y), (b_x, b_y), (c_x, c_y)$  とする (図4.4)。この三点を通る円の中心座標  $(x_c, y_c)$  と半径  $r_c$  は次の式で与えられる。

$$(x_c, y_c) = \begin{cases} (C, B), & b_y = c_y \\ (A, D), & a_y = b_y \\ \left( \frac{-\beta C + D + \alpha A - B}{\alpha - \beta}, \frac{\alpha\beta(A - C) + \alpha D - \beta B}{\alpha - \beta} \right), & otherwise \end{cases} \quad (4.9)$$

$$r_c = \sqrt{(x_c - a_x)^2 + (y_c - a_y)^2} \quad (4.10)$$

$$\alpha = \frac{a_x - b_x}{b_y - a_y} \quad \beta = \frac{b_x - c_x}{c_y - b_y}$$

$$A = \frac{a_x + b_x}{2} \quad B = \frac{a_y + b_y}{2} \quad C = \frac{b_x + c_x}{2} \quad D = \frac{b_y + c_y}{2}$$

上式により計算された円を、人物の頭部候補位置のひとつとする。

一方、画像中で人物は移動しているため、頭部は移動する領域として抽出されているはずである。つまり、頭部候補 (円) として検出した曲線エッジのうち、差分で抽出した移動物体領域と共通部分がないものを排除する。

そして絞りこんだ頭部候補位置である円のうち、中心座標が最も高い位置にある円の中心座標を  $(x_h, y_h)$  とし、これを頭部位置を表す円とする。

#### 4.5 人物を覆う閉包の作成

人物の頭部位置と移動領域を基に、人物の大まかな領域を抽出する。頭部位置を表す円の  $x$  座標  $x_h$  と、差分画像の重心の  $y$  座標  $g_y$  を用いて、以下の式で表される楕円形に沿っ

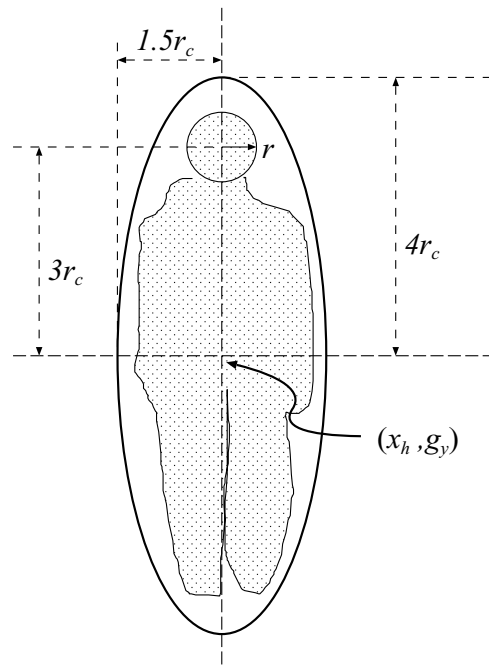


図 4.5: 人物像の楕円モデル

て簡略な人物輪郭モデルを配置し (図 4.5)、これを収束させて人物を包絡する閉包を作る。

$$\frac{(x - x_h)^2}{2.25r_c^2} + \frac{(y - g_y)^2}{16r_c^2} = 1 \quad (4.11)$$

このモデルの収束には Snake を用いる (2.6節参照)。

図 4.6 に示すように Snake を円領域に収束させるために、領域の外であれば収縮するベクトル  $\mathbf{n}$ 、内であれば膨張するベクトル  $-\mathbf{n}$  をとるように  $E_{image}$  を定める [原 95]。

$$\mathbf{n} = \left( -\frac{y_{i+1} - y_i}{\|\mathbf{v}_{i+1} - \mathbf{v}_i\|}, \frac{x_{i+1} - x_i}{\|\mathbf{v}_{i+1} - \mathbf{v}_i\|} \right) \quad (4.12)$$

ここで Snake の点列  $\mathbf{v}_i = (x_i, y_i)$  が画像上で右回りであるとする。またここでは  $E_{con} = 0$  とする。

2.6節で述べた反復計算によって解を計算し、Snake の各点の移動の和がしきい値以下になった時点で反復計算を終了する。

## 4.6 詳細な境界の決定

Snake によって抽出した閉包は大まかな人物領域なので、より詳細な輪郭線を抽出するために、色による領域分割を行ない、閉包を基にして人物領域を決定する。入力であるカラー画像を  $L^*a^*b^*$  表色系 (2.8節参照) に変換し、各画素の色 ( $L^*, a^*, b^*$ ) と位置  $(x, y)$  に

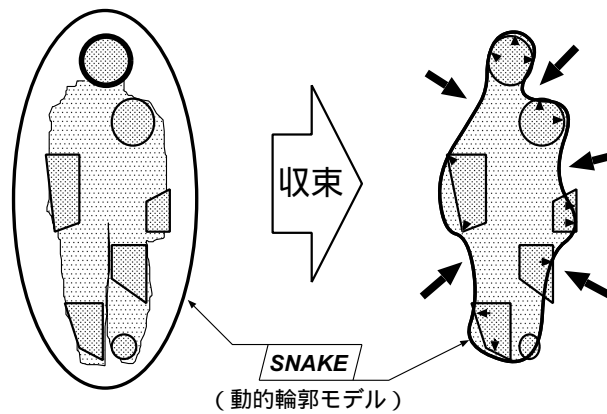


図 4.6: snake の収束

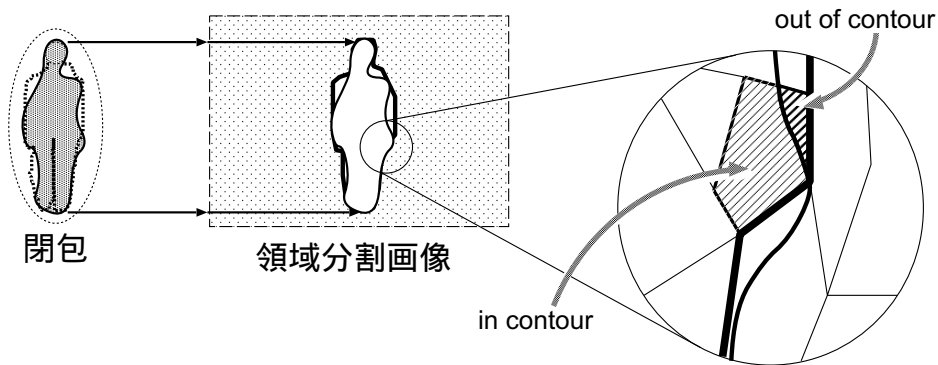


図 4.7: 詳細な人物領域の決定

それぞれ重みをつけた5次元ベクトルについて、2.3節で述べた  $K$ -平均法を用いてクラスタリングを行う。この初期分割によって、類似した色の画素の集合である  $N$  個の部分領域が生成される。それらを  $R_i$  ( $i = 1, \dots, N$ ) とする。

次に、領域  $R$  の面積を  $S(R)$ 、閉包内に含まれている面積を  $S_{in}(R)$  として、図4.7に示すように、閉包の中に含まれている割合が一定以上の領域

$$\frac{S_{in}(R_i)}{S(R_i)} > Th \quad (i = 1, \dots, N_R) \quad (4.13)$$

を最終的な人物領域とし、それらを併合する。

## 4.7 実験と結果

本手法の有効性を調べるために、実画像を対象に実験を行なった。屋外において固定したカメラ (SONY CCD-TR11) で画像系列を撮影し、ワークステーション (SGI O2) を



図 4.8: 人物抽出結果。白線で囲まれている部分が人物領域。

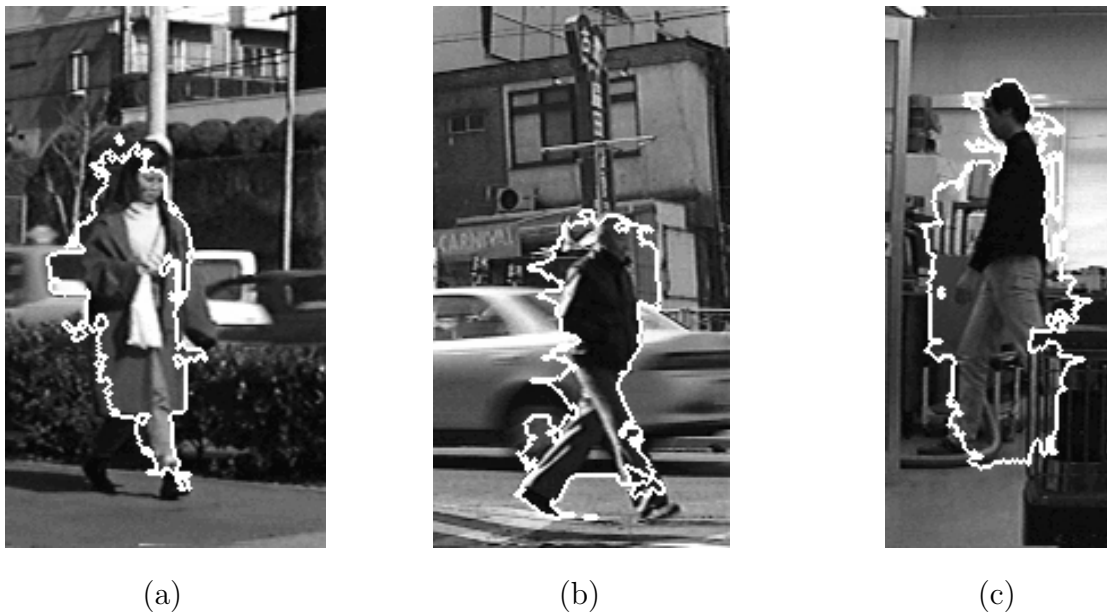


図 4.9: 人物抽出例

用いて大きさ  $620 \times 460$  画素、RGB 各 256 色のフルカラー画像で取り込んだ。まず各画像の領域分割と差分領域の計算を行ない、それらの結果と原画像を用いて他の処理を行なった。各フレームに対して全ての処理時間の合計は、PC (200MHzCPU) 上の C 言語による実装で約 1 分半程度である。

屋外で自動車など人物以外の動物体がある画像系列に本手法を適用した結果を図 4.8 に示す。カメラは舗道から道路に向けて固定しており、左から右へ人物が歩いている背景には走行する車が存在している。図 4.8 の各画像は左上から右下へ、0.1 秒間隔で撮影された時系列順に並べてある。それぞれの画像において、抽出された人物領域の境界線 (白線) を元の画像に重ねて表示している。

図 4.8 の結果からわかるように、人物は画像中のどの位置にあっても抽出されている。また、人物の背景に存在する車などの移動物体の一部が人物領域に含まれているものの、人物の存在位置が大きく変わってしまっていることはない。抽出された人物領域の輪郭が滑らかでない部分があるが、これは領域分割時にグラデーションなどの影響によって疑似輪郭が生成されているためである。

そのほかの画像系列での抽出例を図 4.9 に示す。図 4.9(c) は、他に移動物体がない屋内のシーンである。図 4.9(b) では、背景にぶれるほどの速度で走る車が存在しているが、抽出に成功している。

図 4.8 に示す 15 枚の画像系列では人物が抽出されているが、このように連続した 34 の画像系列に対して同様の処理を行なった。図 4.11 は、人物抽出に成功したフレーム数の割合を示す。この計算に用いた画像系列は、それぞれ 7 枚の画像からなり、その内差分の計算

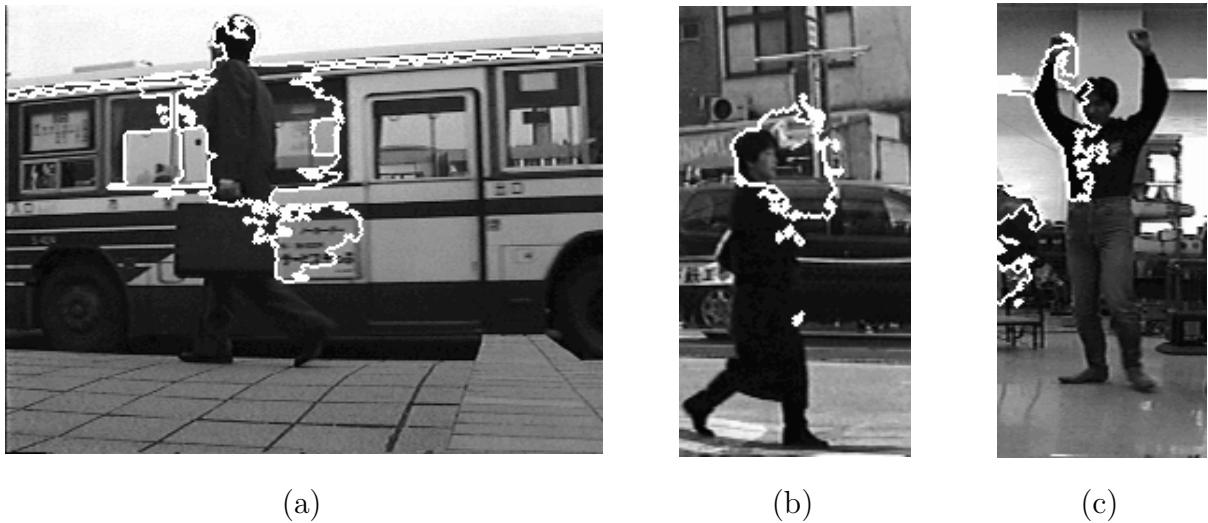


図 4.10: 人物抽出が失敗した画像

できる5枚について人物を抽出した結果、全170枚の画像に対して、123枚(約72%)において人物領域が得られた。また図4.12は、抽出失敗の数で整理した系列数を示し、一つの系列に含まれる5枚の画像全てについて人物を抽出できた系列数は、全体の約42%の14系列であった。ここでは、実際の人物領域を含んで抽出された場合(例えば図4.9(c))と、人物の大部分を抽出した場合(例えば図4.9(a))を成功とし、人物領域をほとんど抽出できなかったり、人物以外の領域を多く含んだ場合には失敗とした。

人物領域が抽出できなかった原因としては、背景でバスが移動している図4.10(a)のように、人物の移動に対して背景の移動物体の移動量が多い場合、閉包の収束がうまく行われなかったことが挙げられる。また抽出失敗の傾向として、曲線エッジの曲率の計算で実際よりも小さい円に近似されてしまい、楕円が人物の上部しか覆わず、全身の抽出に失敗している場合(図4.10(b))や、図4.10(c)のように、手など他のものを頭部と認識してしまった場合などが多かった。

図4.8の14枚目の画像のように、歩行動作によっては領域に含まれない部分が出てくることがある。これは楕円モデルから手足の先がはみ出し、Snakeがうまく収束しなかったためである。しかし、図4.8の15枚目のように、手足が領域分割において一つの領域を形成すれば、この領域を式(4.13)で人物領域と判定することができる。

図4.8、図4.10には昼間に撮影した画像のみを用いている。しかし、特定の色を特徴として用いているわけではないので、昼夜の違いによって明るさが変化しても、人物と背景が判別できれば、領域分割を行った結果を用いることができる。しかし、領域分割では色差の違いによって色空間クラスタを生成しているため、人物も含めて画像全体が暗くなってしまう場合には、期待するような領域分割は行われぬ。そのため、夜間の撮像においては、照明が必要となるが、本手法が適用できると思われる。

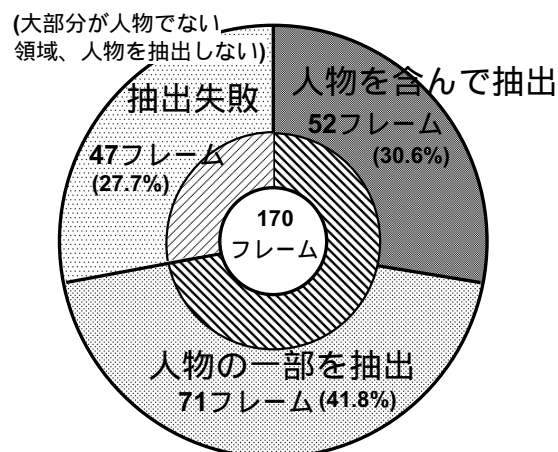


図 4.11: 各フレーム毎の抽出結果  
1シーン7フレームの中で、差分が計算できる中5フレーム分を計算した。

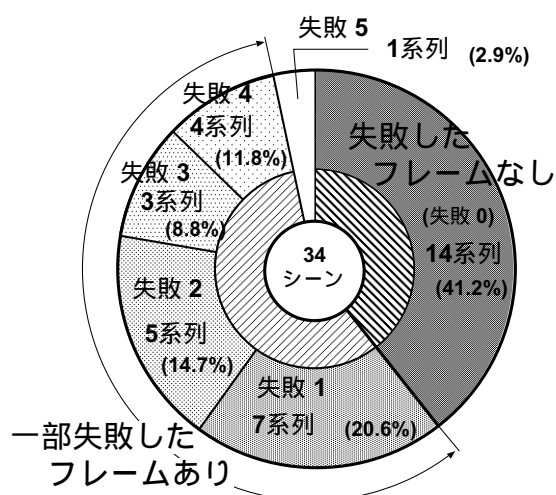


図 4.12: 抽出失敗数で整理した画像系列としての抽出結果

## 4.8 おわりに

この章では、頭部位置を抽出するアルゴリズムを示し、それとフレーム間差分を用いて画像中の人物領域を抽出する手法を提案した。また本手法を計算機上に実装し、実画像を用いて抽出実験を行なった。本手法によって、前後三枚のフレームのみで自動的に画像内の人物の領域を抽出することが可能になった。

本手法はフレーム間差分を用いているため、移動速度の遅い人物像の抽出が困難である。また、人物の頭より上方に手などが存在する場合に頭部位置を誤認するため、人物の姿勢に影響を受けるといった問題がある。

今後の課題として、これらの問題点の解決法を検討すると共に、現在では差分を計算することにのみ用いている前後のフレームの情報を、エッジや領域分割時にも用いて抽出を向上させることなどが挙げられる。また、抽出した領域のより正確な輪郭を得るために、姿勢推定などに用いられる精度の高い人物モデル[佐藤97, Gavril96]を画像中の人物像と一致させることなどが考えられ、動作認識の前処理としての応用などを検討している。





# 第5章 回転モデルに基づいた 動画像からの人物の腕領域の 抽出と運動パラメータ推定

## 5.1 はじめに

ビデオで人物を撮影することは、前章の冒頭で示したような監視目的だけとは限らず、様々なところで行われている。例えば、現在はキーボードやマウスなどの入力機器を介して行っている計算機とのインターフェイスの代わりに、カメラを設置してジェスチャで指示を出すことが研究されている [岡 99]。また、一過性であるスポーツ選手の動作や、失われていく伝統芸能の舞踊など [わらび座 96]、実世界で行う運動をビデオに撮り、計算機上でCGを用いて三次元の動作を再現したり、後世に伝えるために保存するといった、バーチャルリアリティの研究もなされている。

腕や手足がどのように動いているのかを知るためには、あらかじめ腕や脚を抽出しておく必要がある。そのため、抽出処理は人物動作の認識システムにおいて重要な一部分となっている。この章では、前章に引き続きビデオカメラなどから得られる動画像を扱うが、前章で述べた手法が画像から得られる情報を重視したボトムアップ的な処理だったのに対し、この章で述べるものは認識されるべき動作をモデル化したトップダウン的なものであり、認識処理を意識したものである。

現在までに実用化されている動作認識 [中嶋 98b] に用いられている手法としては、対象人物にマーカー [Goncalves98, 木本 91] やセンサ [戸川 97, 岩田 97] を取り付けて認識をするものが多い。これらは人物の動きを正確に測定できる反面、システムが設置される場所、対象者の動作やその行動範囲を制限しなければならないといった問題点がある。一方で、カメラからの画像に基づいた認識システムも研究され、この方式には、カメラが複数台であっても比較的容易に設置できること、また人物に何も取り付ける必要がなく、行動が制限されないという利点がある。

画像に基づいた人物動作認識の研究には、人体モデルを用いるものとそうでないものがある。人体モデルを用いない方法 [Yamato92, 大和 94, 高橋 94, Nishimura98] は、どんな動作をしているのかということ进行分类することはできるが、動作を再現することはできない。

そこで、実際の体の各部の動きを再現するために、パラメータ表現された人体モデルが多く用いられており、この時に、対象となる人物の領域を抽出することが必要となる。これらのモデルは、通常、関節で結合されたリンク [Morris98, Guo94, 石井 93] や、2D の矩形 [Ju96, Haritaoglu98]、3D の円筒や多面体 [Gavrila96, Ohya94, 大谷 94, 山本 96, 亀田 96, Minoh98] などで構成されている。こういった人体モデルを用いた認識手法には、大きく分けて以下の二つがある。

一つは、ある姿勢をとったモデルを画像平面に投影し、その投影像と、あらかじめ抽出しておいた対象人物の領域との違いを評価し照合する手法 [Gavrila96, Ohya94, Guo94, Minoh98] である。これにより、抽出された領域にシルエットが最も似ているモデルのパラメータが得られ、対象者の姿勢が認識できる。この時、人物領域の抽出には、背景差分や動物体検出などの手法が用いられる。その多くは、背景差分を簡単にするために背景が既知か少なくとも一様な色であることを仮定しており、また背景に他の移動物体が存在しないことが前提となっている。また照合過程において膨大な探索を必要とすることも問題である。

もう一つは、人体の運動学的モデルを用いて人物の運動を追跡する手法 [Morris98, Ju96, 山本 96] である。これらの手法は、動きの情報を用いているため、背景の複雑さには影響されないという利点がある。しかし実画像から得られるオプティカルフローは、エッジや動きの不連続な部分では信頼できるフローが得られず [六日市 99]、それを避けるための過度な平滑化はパラメータの推定に影響を与える。そのため、画素値の時空間勾配を直接用いる手法 [山本 96] や、ステレオによって動領域を分割する [岡田 99] などの手法がとられているが、計算コストが高いことや複数のカメラを必要とすることが問題となる。また多くの場合、追跡する初期フレームや蓄積誤差が大きくなったフレームで [太田 98]、その時刻での画像内の人物の位置と姿勢を指定しなければならず、モデルを人物に当てはめるための対話的な操作には労力を必要とする。

以上のように、未知の複雑な背景において他の移動物体が存在する、より実際的な状況で、人物を抽出することが必要になってくる。そのためにまず最初に、ジェスチャなどで情報を多く伝える、人物の腕を抽出することを考える [Gavrila99]。

この章では、未知の複雑な背景において他の移動物体が存在する、より実際的な状況において、ジェスチャなどで情報を多く伝える、人物の腕を抽出する手法について述べる。本手法は人物の腕の運動を回転モデルを用いて表し、その運動パラメータを同時に推定する。提案手法はオプティカルフローの curl に基づいているので、背景に依存せずに回転している腕領域を抽出することができ、背景の移動物体も除去できるという特長がある。また複数の運動を扱うために EM アルゴリズムを導入し、実画像から得られたノイズの多いオプティカルフローから良好な推定を行うことができる。

以下 5.2 節ではオプティカルフローに基づいた領域抽出について説明し、5.3 節において運動パラメータの推定手法について、5.4 節では EM アルゴリズムを用いた領域分割と推定のアルゴリズムについて述べ、実画像を用いた評価と実験結果を 5.5 節で示す。

## 5.2 Curlに基づく回転領域抽出

本手法の領域抽出は、ベクトル場の curl に基づいている。この節では、回転物体によって生成されたオプティカルフローの curl の性質について説明する。そして単純な回転と一般的な回転についての curl の例を示し、それによって腕領域の抽出が行えることを示す。

### 5.2.1 回転ベクトル場の Curl の性質

三次元空間の点  $\mathbf{x} = (x, y, z)^T$  の速度を  $\dot{\mathbf{p}}(\mathbf{x}) = (u, v, w)^T$  とすると、この速度ベクトルはベクトル場を形成する。このベクトル場の curl<sup>1</sup> は以下の式で定義される。

$$\nabla \times \dot{\mathbf{p}}(\mathbf{x}) \equiv \left( \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)^T \quad (5.1)$$

回転によって生成されたベクトル場において、この curl は次のような性質を持つ。

$$\nabla \times \{\dot{\mathbf{p}}(\mathbf{x} - \mathbf{c}) + \mathbf{t}\} = \nabla \times \dot{\mathbf{p}}(\mathbf{x}) \quad (5.2)$$

ここで  $\mathbf{c}$  は任意のベクトル、 $\mathbf{t}$  は位置に依存しない定数ベクトル場である。二次元画像平面内のオプティカルフローは、 $z = w = 0$  としたベクトル場と考えることができる。回転物体によって生成されたオプティカルフローについて考えると、 $\mathbf{c}$  は回転中心であり、 $\mathbf{t}$  は（正射影下では）物体の並進運動であると見なすことができる。したがって、画像内のどこにあっても curl の値は変わらず、また物体が並進運動していても影響がない。 $\mathbf{t}$  はまた、カメラや背景の運動物体によって生じる並進運動とみなすこともできるので、curl の性質によってそれらの影響を排除することができる。

### 5.2.2 Curl の例

まず初めに、画像平面に平行な面上で回転している図 5.1(a) のような単純な場合を考える（以下では正射影を仮定している）。点  $\mathbf{p}$  を、回転中心である原点から  $r$  ( $0 \leq r \leq r_1$ ) だけ離れ、角速度  $\omega$  で回転し、角度  $\theta (= \omega t)$  の位置にある腕上の点とする。このとき、 $\mathbf{p}$  の速度  $\dot{\mathbf{p}}$ （画像上でのオプティカルフロー）とその curl  $\nabla \times \dot{\mathbf{p}}$  は次のように表される。

$$\begin{aligned} \mathbf{p} &= (x, y, 0)^T \\ &= (r \cos \theta, r \sin \theta, 0)^T \end{aligned} \quad (5.3)$$

$$\begin{aligned} \dot{\mathbf{p}} &= (-\omega r \sin \theta, \omega r \cos \theta, 0)^T \\ &= (-\omega y, \omega x, 0)^T \end{aligned} \quad (5.4)$$

$$\nabla \times \dot{\mathbf{p}} = (0, 0, 2\omega)^T \quad (5.5)$$

<sup>1</sup>回転や rot という語が使われることが多い[小西 78]が、実際に物体が回転運動をすることを指す場合との混乱を避けるために、ここでは curl という語を用いることにする。

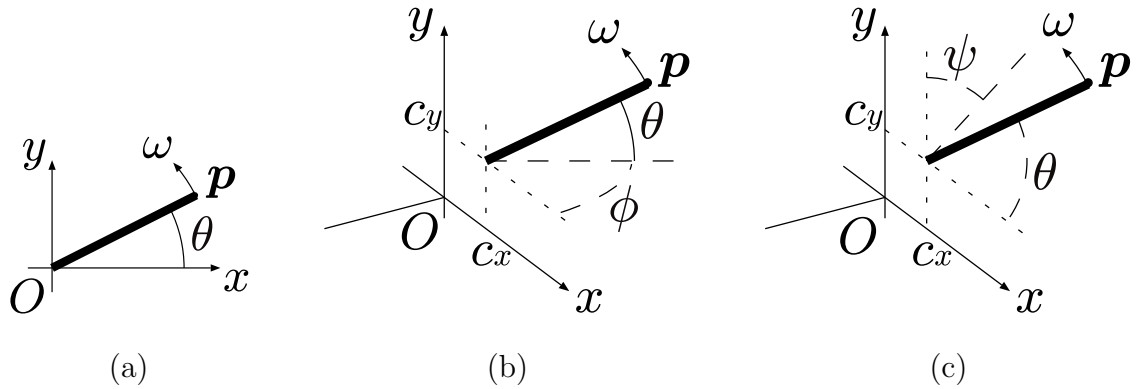


図 5.1: スティックモデル (a) 画像平面と平行な面上での回転 (b) 回転面が画像平面 ( $O-xy$ ) に対して  $y$  軸回りに傾いている場合 (c)  $x$  軸回りに傾いている場合

式 (5.5) より、curl の  $z$  成分は  $2\omega$  であり、 $\theta$  や  $r$ 、 $x, y$  にも依存しない。すなわち、回転中心付近の小さいフローと腕の先端の大きなフローとが同じ特徴量を持つことになる。

一般に回転中心は原点ではなく、画像中のどこにあるのかは未知であるので、次に腕が点  $c = (c_x, c_y, c_z)^T$  を中心に回転している場合について考える。このときの点  $p$  の速度  $\dot{p}$  と curl は以下ようになる。

$$\dot{p} = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = \begin{pmatrix} -\omega(y - c_y) \\ \omega(x - c_x) \\ c_z \end{pmatrix} \quad (5.6)$$

$$\begin{aligned} \nabla \times \dot{p} &= \left( 0, 0, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)^T \\ &= (0, 0, 2\omega)^T \end{aligned} \quad (5.7)$$

この場合も、式 (5.7) の  $z$  成分は式 (5.5) と同じであり、同様に  $\omega$  のみに依存すると言える。

以上で、腕が画像平面に平行な面上で回転している場合について述べたが、三次元運動への拡張として、以下では画像平面に対して傾いた面上での回転を考える。図 5.1(b) に示すように、腕の回転面は、画像平面 ( $O-xy$ ) に対して  $y$  軸回りに角  $\phi$  ( $\neq \frac{\pi}{2}$ ) だけ傾いているとすると、このときの点  $p$  とその速度  $\dot{p}$  は次ようになる。

$$\begin{aligned} p &= (x, y, z)^T \\ &= (r \cos \phi \cos \theta, r \sin \theta, r \sin \phi \cos \theta)^T \end{aligned} \quad (5.8)$$

$$\begin{aligned} \dot{p} &= (-\omega r \cos \phi \sin \theta, \omega r \cos \theta, -\omega r \sin \phi \sin \theta)^T \\ &= \left( -\omega \cos \phi y, \frac{\omega x}{\cos \phi}, -\omega \sin \phi y \right)^T \end{aligned} \quad (5.9)$$

回転中心を  $\mathbf{c} = (c_x, c_y, c_z)^T$  とすれば、 $\dot{\mathbf{p}}$  は

$$\dot{\mathbf{p}} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\omega \cos \phi (y - c_y) \\ \frac{\omega(x - c_x)}{\cos \phi} \\ -\omega \sin \phi (y - c_y) \end{pmatrix} \quad (5.10)$$

となり、その curl の  $z$  成分は次のようになる。

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{1 + \cos^2 \phi}{\cos \phi} \quad (5.11)$$

ここで、(フレームレート程度の) 微小時間では  $\phi$  は一定であると仮定しているので、式(5.11) はやはり  $\omega$  にのみ依存しているとみなせる。実際、任意の微小な三次元の回転運動は、ある回転軸まわりの回転と見なすことができ [Craig91]、その回転軸は回転面の法線に一致する。

次に図5.1(c)のように、腕の回転面が  $x$  軸回りに角  $\psi$  ( $\neq \frac{\pi}{2}$ ) だけ傾いているとすると、次のように式(5.10)、式(5.11) とよく似た結果が得られる。

$$\begin{aligned} \mathbf{p} &= (x, y, z)^T \\ &= (r \cos \theta, r \cos \psi \sin \theta, r \sin \psi \sin \theta)^T \end{aligned} \quad (5.12)$$

$$\begin{aligned} \dot{\mathbf{p}} &= (-\omega r \sin \theta, \omega r \cos \psi \cos \theta, \omega r \sin \psi \cos \theta)^T \\ &= \left( \frac{-\omega y}{\cos \psi}, \omega \cos \psi x, \omega \sin \psi x \right)^T \end{aligned} \quad (5.13)$$

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{-\omega(y - c_y)}{\cos \psi} \\ \omega \cos \psi (x - c_x) \\ \omega \sin \psi (x - c_x) \end{pmatrix} \quad (5.14)$$

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{1 + \cos^2 \psi}{\cos \psi} \quad (5.15)$$

一般には、回転面は  $\phi \neq 0$  かつ  $\psi \neq 0$  である。この場合、速度はやや複雑になるが、curl は単純な式になる (導出は付録Dを参照)。

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \omega \frac{\cos^2 \phi + \sin^2 \psi + \sin^2 \phi \cos^2 \psi}{\cos \phi \sin \psi} \quad (5.16)$$

5.3節以降の議論では主に  $\phi$  の場合について考えるが、 $\psi$  についても同様であり、5.4.3節で両者の関係について述べる。一般的な場合の議論は5.6節で述べる。

### 5.2.3 Curlに基づいた領域分割

式(5.11)、式(5.15)、式(5.16)で示したように、 $\phi$  又は  $\psi$  が一定という条件で curl の  $z$  成分は  $\omega$  にのみ依存するが、それは速度  $\dot{\mathbf{p}}$  の  $(u, v)$  成分であるオプティカルフローから計

算することができる。この  $z$  成分は、ある回転している腕の内部では同じ値をとり、それ以外の領域とは異なるので、これを用いて、腕の領域を抽出することができる（具体的な領域分割の方法は付録 B を参照）。画像中で複数の腕が回転していて、かつ接している場合でも、それらの角速度が異なる値である限り、別々の物体として分離することができる。

### 5.2.4 透視投影下の Curl

5.2節の議論は、すべて正射影を仮定している。これは簡単化のためでもあるが、実際、物体の奥行きがカメラと物体の距離に比べて非常に小さいことを仮定してもよい場合が多いからである。ただし、透視投影下での curl の  $z$  成分は、回転面が画像平面と平行である限り、正射影下のものと同じである（付録 C 参照）。

## 5.3 運動パラメータの推定

5.2.3節で、三次元空間内の回転する腕の領域を抽出する方法について述べた。ここでは、その腕の運動パラメータ、つまり回転中心  $(c_x, c_y)$ 、角速度  $\omega$ 、傾き角  $\phi$ 、腕の角度  $\theta$  の推定方法について述べる。

実画像から得られるオプティカルフローはノイズを含んでいるため、抽出された領域内のフローを用いて、最小二乗法によりパラメータを推定する。そのためには、領域内の各点  $\mathbf{p}_j = (x_j, y_j)^T$  とそのフロー  $\dot{\mathbf{p}}_j = (u_j, v_j)^T$  について式 (5.10) を連立して解けばよい。

式 (5.10) は非線形であるので、このまま連立方程式を解くことは難しいため、以下のよう書きなおす。

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} \alpha y_j + \beta \\ \gamma x_j + \delta \end{pmatrix} \quad (5.17)$$

ここで係数はそれぞれ、

$$\alpha = -\omega \cos \phi, \quad \beta = \omega c_y \cos \phi, \quad \gamma = \frac{\omega}{\cos \phi}, \quad \delta = \frac{-\omega c_x}{\cos \phi} \quad (5.18)$$

である。これを行列形式に変形する。

$$\dot{\mathbf{p}}_j = A_j \mathbf{q} \quad (5.19)$$

$$A_j = \begin{pmatrix} y_j & 1 & 0 & 0 \\ 0 & 0 & x_j & 1 \end{pmatrix} \quad (5.20)$$

$$\mathbf{q} = (\alpha, \beta, \gamma, \delta)^T \quad (5.21)$$

すなわち、式 (5.19) のようにオプティカルフローはパラメータ  $\alpha, \beta, \gamma, \delta$  の線形式で表すこ

とができる。よって、式(5.19)を以下のように連立する。

$$\begin{pmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix} \mathbf{q} \quad (5.22)$$

この優決定の連立一次方程式を、QR分解を用いて最小二乗推定問題として解く [小柳91, p.914]。これにより得られた  $\mathbf{q}$  から、運動パラメータは以下のように求められる。

$$c_x = \frac{-\delta}{\gamma}, \quad c_y = \frac{-\beta}{\alpha}, \quad |\omega| = \sqrt{-\alpha\gamma}, \quad \cos \phi = \sqrt{\frac{-\alpha}{\gamma}} \quad (5.23)$$

ここで  $\omega$  が解を持つためには、回転運動において  $-\alpha\gamma > 0$  が満たされなければならない。 $\phi$  についても、 $0 < \frac{-\alpha}{\gamma} \leq 1$  が成立する必要がある。これらの制約と、 $\omega$  の符号の決定については5.4.2節で述べる。

以上に述べた手法によって運動パラメータを推定できるが、式(5.22)は腕の角度  $\theta$  の情報を含まないため、この推定では  $\theta$  を求めることはできない。したがって、抽出された腕の領域の主軸を  $\theta$  の推定値として用いることにする。

## 5.4 EMアルゴリズムを用いた領域分割と推定

5.3節では、腕領域内のオプティカルフローを用いた、腕の運動パラメータの推定手法について述べた。そして、その領域はあらかじめオプティカルフローのcurlの  $z$  成分を用いて抽出(5.2.3節)してあることが前提であった。つまり、ある領域が得られれば、パラメータ推定を行うことができる。

一方、領域分割のためには精度のよいcurlの計算が必要となるが、curlは微分演算なのでノイズに敏感であり、実画像において各点で正確に計算することは困難である。そこで、式(5.17)においてオプティカルフローはパラメータ表現されているので、微分演算  $\left(\frac{\partial v}{\partial x}, \frac{\partial u}{\partial y}\right)$  の代わりに、推定されたオプティカルフローの勾配  $(\gamma, \alpha)$  を用いることができる。

ここで問題となるのは、推定と領域分割の相互依存である。つまり、パラメータ推定は領域分割を必要とするが、領域分割には推定されたパラメータが要求される。この問題を克服するために、本手法ではEMアルゴリズム [栗田98, Dempster77] を用いる。2.5節で述べたように、EMアルゴリズムは反復法の一つであり、初期値として大まかな領域分割を行えばよく、それらの領域に対してパラメータを推定し、再び領域分割を行うという反復処理をすることで相互依存の問題を解決することができる。

また2.5.2節で述べたように、EMアルゴリズムは隠れ変数を含む正規混合分布の最尤推定に適用できる [大羽99]。つまり、画像中の複数物体の領域を分割し、かつそれぞれのパラメータを推定することができるということである。複数の運動物体はそれぞれの運動パラメータを持っていて、ある物体の領域内の点は、その物体のパラメータで表されるフロー

を中心にして分布しているはずである。しかし、各点がどの物体に属しているのかは未知である。ここで、各点がどの物体（またはそのパラメータ）に属しているかを表す、陽に現れない変数が隠れ変数となる。この隠れ変数をもとにクラスタリングを行うことで、それぞれの運動物体に分割することができる。

ここでは、式(5.19)における、求められたフロー  $\dot{p}_j$  とパラメータ表現された  $A_j \mathbf{q}$  との誤差の分布は、二次元の正規分布に従うと仮定する。したがって、パラメータ  $\mathbf{q} = (\alpha, \beta, \gamma, \delta)^T$  が与えられた時に、点  $\mathbf{p}_j$  でのフローが  $\dot{p}_j$  である条件付確率は以下ようになる。

$$P(\dot{p}_j | \mathbf{p}_j, \mathbf{q}, \Sigma) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp \left\{ \frac{-1}{2} (\dot{p}_j - A_j \mathbf{q})^T \Sigma^{-1} (\dot{p}_j - A_j \mathbf{q}) \right\} \quad (5.24)$$

ここで  $\Sigma$  は誤差  $\dot{p}_j - A_j \mathbf{q}$  の共分散行列であるが、 $u (= \alpha y + \beta)$  と  $v (= \gamma x + \delta)$  は別々に推定されるため、それぞれの誤差は互いに独立であると仮定すると、上式は次のように二つの確率密度関数の積で表すことができる。

$$\begin{aligned} P(\dot{p}_j | \mathbf{p}_j, \mathbf{q}, \sigma_x^2, \sigma_y^2) \\ = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp \left\{ \frac{-(u_j - \alpha y_j - \beta)^2}{2\sigma_x^2} \right\} \cdot \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left\{ \frac{-(v_j - \gamma x_j - \delta)^2}{2\sigma_y^2} \right\} \end{aligned} \quad (5.25)$$

複数の異なる運動物体が存在する場合における点  $\mathbf{p}_j$  でのフローが  $\dot{p}_j$  である確率を、各運動物体  $i$  のパラメータを  $\mathbf{q}_i$ 、重みを  $\xi_i$  として、以下の混合分布で表す。

$$P(\dot{p}_j | \mathbf{p}_j, \mathbf{q}_1, \mathbf{q}_2, \dots, \Sigma_1, \Sigma_2, \dots) = \sum_i \xi_i P(\dot{p}_j | \mathbf{p}_j, \mathbf{q}_i, \Sigma_i) \quad (5.26)$$

### 5.4.1 反復アルゴリズム

式(5.25)で定義された確率密度関数を、2.5.2節で示した混合分布に対するEMアルゴリズムに適用し、腕領域の抽出とそのパラメータを推定するアルゴリズムを以下に示す。

1. 各点  $\mathbf{p}_j = (x_j, y_j)^T$  ( $j = 1, \dots, N$ ) についてオプティカルフロー  $\dot{p}_j = (u_j, v_j)^T$  を計算する。
2. 初期値としての大まかな領域分割としてクラスタリングを行ない、クラスタ  $R_i$  ( $i = 1, \dots$ ) を得る。
3. 重み  $w_{ij}$  は、パラメータ  $\mathbf{q}_i = (\alpha_i, \beta_i, \gamma_i, \delta_i)$  で表されるクラスタ  $R_i$  に点  $\mathbf{p}_j$  が属する確率を表すとする。この重みの初期値を以下のように決める。

$$w_{ij} = \begin{cases} 1 & (\mathbf{p}_j \in R_i) \\ 0 & (\mathbf{p}_j \notin R_i) \end{cases} \quad (5.27)$$



4.  $i = 1, \dots$  について、 $w_{ij}$  を正規化する。

$$\xi_i = \frac{1}{N} \sum_j w_{ij} \text{ として、 } w_{ij} \leftarrow \frac{w_{ij}}{\xi_i} \quad (5.28)$$

5. 各クラス  $R_i$  について、以下の式で表される重みつき二乗誤差を最小にするパラメータ  $\mathbf{q}_i$  を推定する。

$$\sum_j w_{ij} \|\dot{\mathbf{p}}_j - A_j \mathbf{q}_i\|^2 = \sum_j \|(\sqrt{w_{ij}} \dot{\mathbf{p}}_j) - (\sqrt{w_{ij}} A_j) \mathbf{q}_i\|^2 \quad (5.29)$$

すなわち、次の連立一次方程式をQR分解を用いて解き、 $\mathbf{q}_i$  を求める。

$$\begin{pmatrix} \sqrt{w_{i1}} \dot{\mathbf{p}}_1 \\ \sqrt{w_{i2}} \dot{\mathbf{p}}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{w_{i1}} A_1 \\ \sqrt{w_{i2}} A_2 \\ \vdots \end{pmatrix} \mathbf{q}_i \quad (5.30)$$

6. 各クラス  $R_i$  の重みつき分散  $\sigma_{x_i}^2, \sigma_{y_i}^2$  を計算する。

$$\sigma_{x_i}^2 = \frac{1}{N} \sum_j w_{ij} (u_j - \alpha_i y - \beta_i)^2 \quad (5.31)$$

$$\sigma_{y_i}^2 = \frac{1}{N} \sum_j w_{ij} (v_j - \gamma_i y - \delta_i)^2 \quad (5.32)$$

7. 式(5.30)~(5.32)で得られた  $\mathbf{q}_i, \sigma_{x_i}^2, \sigma_{y_i}^2$  を用いて、次式の事後確率を計算し、重み  $w_{ij}$  を更新する。

$$w_{ij} = \frac{\xi_i P(\dot{\mathbf{p}}_j | \mathbf{p}_j, \mathbf{q}_i, \sigma_{x_i}^2, \sigma_{y_i}^2)}{\sum_k \xi_k P(\dot{\mathbf{p}}_j | \mathbf{p}_j, \mathbf{q}_k, \sigma_{x_k}^2, \sigma_{y_k}^2)} \quad (5.33)$$

8. もし

$$\sum_i \|\mathbf{q}_i - \mathbf{q}_i^{old}\|^2 > \varepsilon \quad (5.34)$$

ならば、つまり前回の推定値  $\mathbf{q}_i^{old}$  と今回の推定値  $\mathbf{q}_i$  の差がしきい値  $\varepsilon$  よりも大きければ、 $\mathbf{q}_i^{old} \leftarrow \mathbf{q}_i$  としてステップ 4. へ。

そうでなければ、反復処理を停止する<sup>2</sup>。ステップ 9. へ。

9. 領域分割を行う。反復処理を停止した時点で、点  $\mathbf{p}_j$  は、各クラスに対してそれぞれ重み  $w_{ij}$  をもっている。その中で最大の重みを求め、点  $\mathbf{p}_j$  をそのクラス  $R_{i^*}$  ( $i^* = \operatorname{argmax}_i w_{ij}$ ) に分類する。そして画像上で同じクラスに属する画素を連結し、それぞれの領域に分割する。

<sup>2</sup>EMアルゴリズムの終了条件は本来ならば尤度を用いるべきであるが、非常に収束が遅い[栗田98]ので、今回はある程度まで推定値が収束した時点で処理を打ち切っている。

### 5.4.2 非回転運動の除去

前節で述べたアルゴリズムによって、各クラスタのパラメータ  $q_i$  が求まるので、式 (5.23) によって  $c_x, c_y, \omega, \phi$  が計算できる。5.3節で述べたように、 $\omega$  は実数であるので、式 (5.23) から  $-\alpha\gamma > 0$  が満たされなければならない。しかし、 $\alpha$  と  $\gamma$  はその領域が回転運動しているかどうかにかかわらず別々に推定されるため、条件が満たされない領域も出てくる。そのような領域は、並進や拡大、縮小などの非回転運動に対応するので、これらの運動を分類し、非回転運動の領域を排除することを考える。

表 5.1 に示すように、 $-\alpha\gamma$  の符号 (つまり  $\alpha$  と  $\gamma$  の符号) によって運動を分類することができる。 $\alpha$  と  $\gamma$  が異なる符号を持つとき ( $-\alpha\gamma$  が正のとき)、パラメータに対応する運動は回転運動であり、その回転の方向は  $\alpha (= -\omega \cos \phi)$  からわかる。 $\alpha$  と  $\gamma$  が共に 0 のとき、その領域のオプティカルフローは一定である。すなわち、その領域は並進運動をしているか、または静止している。最後に  $-\alpha\gamma \leq 0$  の場合であるが、 $\alpha$  と  $\gamma$  がともに正 (またはどちらか一方が 0 で他方が正) であるときは拡大運動であり、ともに負 (またはどちらか一方が 0 で他方が負) のときは縮小運動である。

### 5.4.3 回転面 (軸) の決定

前節では回転運動の領域を選択する方法を示した。次に、その回転運動が 5.2.2 節で述べた二つの三次元運動、つまり図 5.1(b) のように回転面が  $y$  軸回りに角  $\phi$  だけ傾いている場合と、 $x$  軸回りに角  $\psi$  だけ傾いている場合 (図 5.1(c)) のうち、どちらが適切なのかを決定し、選択しなければならない。

5.3節で述べたように、式 (5.23) より、 $\phi$  について満たされなければならない条件がある。回転運動については  $-\alpha\gamma > 0$  なので、 $0 \leq \frac{-\alpha}{\gamma}$  は満たされる。さらに  $|\alpha| \leq |\gamma|$  であれば、 $\frac{-\alpha}{\gamma} \leq 1$  となり条件を満たすので、この場合は  $y$  軸回りに  $\phi$  だけ傾いている場合と考えてよい。ここで式 (5.14) と式 (5.17) より、 $\cos \psi = \sqrt{\frac{-\gamma}{\alpha}}$  なので、 $x$  軸回りに  $\psi$  だけ傾いている場合、 $\phi$  とは逆に、 $|\alpha| \geq |\gamma|$  が満たすべき条件となる。

以上をまとめた結果を表 5.2 に示すが、これを用いることで、どちらの三次元運動を選択するかを決定することができる。すなわち、もし  $|\alpha| \leq |\gamma|$  ならば、回転面は画像平面に対して  $\phi$  だけ傾いており、回転面の法線は  $(\sin \phi, 0, \cos \phi) = \left( \sqrt{1 - \frac{-\alpha}{\gamma}}, 0, \sqrt{\frac{-\alpha}{\gamma}} \right)$  となる。そうでなければ、 $(0, \sin \psi, \cos \psi)$  である。

表 5.1: パラメータの符号による運動の分類

条件	パラメータの符号	対応する運動
$-\alpha\gamma > 0$	$\alpha > 0 \quad \gamma < 0$	回転 ( $\omega < 0$ )
	$\alpha < 0 \quad \gamma > 0$	回転 ( $\omega > 0$ )
$-\alpha\gamma = 0$	$\alpha = 0 \quad \gamma = 0$	並進/静止
$-\alpha\gamma \leq 0$	$\alpha \geq 0 \quad \gamma \geq 0$	拡大
	$\alpha \leq 0 \quad \gamma \leq 0$	縮小

表 5.2: 回転面の決定

条件	軸	パラメータ
$ \alpha  \leq  \gamma $	$y$ 軸	$\cos \phi = \sqrt{\frac{-\alpha}{\gamma}}, \quad \psi = 0$
$ \alpha  \geq  \gamma $	$x$ 軸	$\phi = 0, \quad \cos \psi = \sqrt{\frac{-\gamma}{\alpha}}$

## 5.5 実画像を用いた実験結果

本手法を計算機上に実装し、実画像を用いて実験を行った。アルゴリズムのステップ1におけるオプティカルフローの計算には、2.7節最後で述べたプログラムを用いた。ステップ2における初期領域分割には、フローの方向のヒストグラムを24分割し、検出したピークの数だけクラスタを生成するという単純なクラスタリング手法を用いた。

### 5.5.1 マニピュレータを用いた推定精度評価

提案手法の推定精度を評価するため、角度を正確に制御できるマニピュレータ（川崎重工 JS005-A、6軸、1軸のみの回転では  $\pm 0.001[\text{deg}]$  の誤差）を用いて実験を行った。図 5.2(a) に示すように、長さ約 58[cm] の円筒にテクスチャを張り（オプティカルフローを計算しやすくするため）、マニピュレータの第6軸の回転軸に垂直に取り付けた。第5軸の角度が  $0[\text{deg}]$  の時に、カメラとマニピュレータは正面で向き合うように設置してあり（カメラとの距離は約 440[cm]）、このときに第6軸の角度  $\theta$  を操作することで、円筒はカメラと並行な平面で回転する。第5軸の角度は、円筒の回転する回転平面の、画像平面に対する傾き

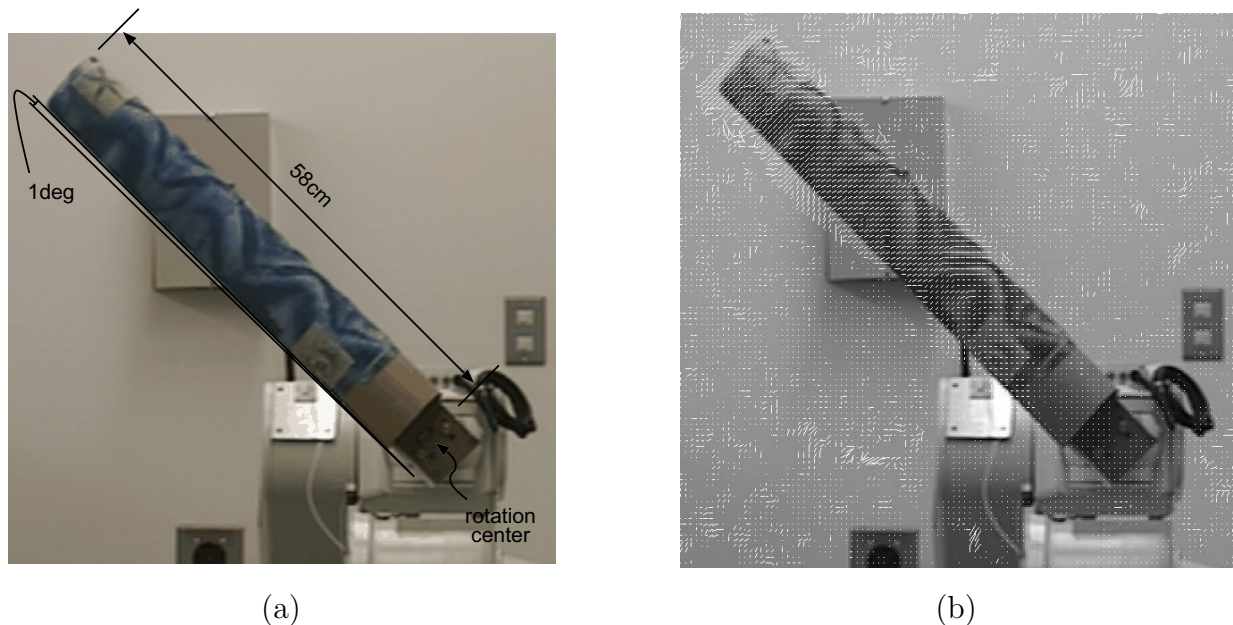
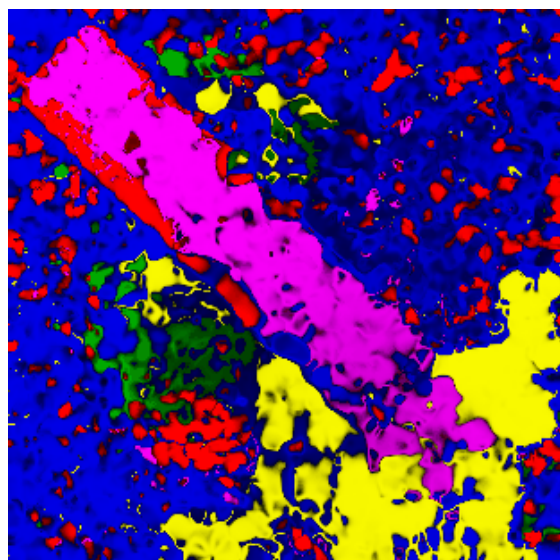
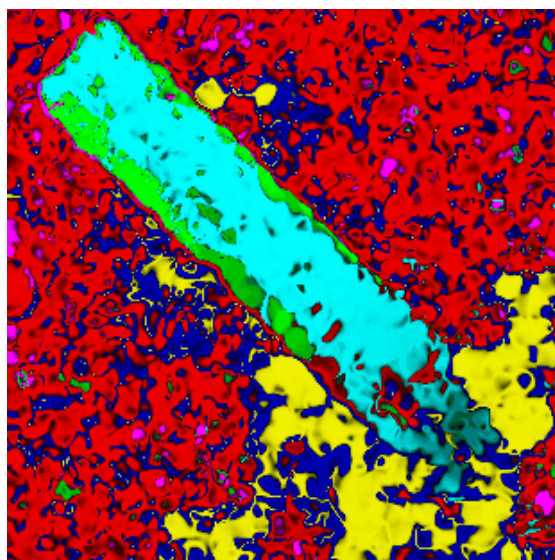
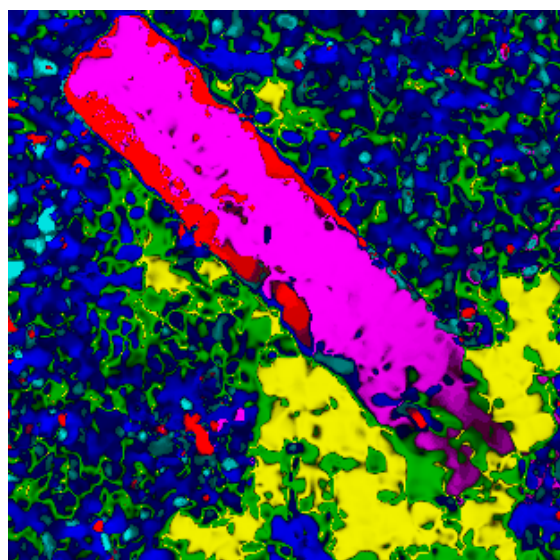
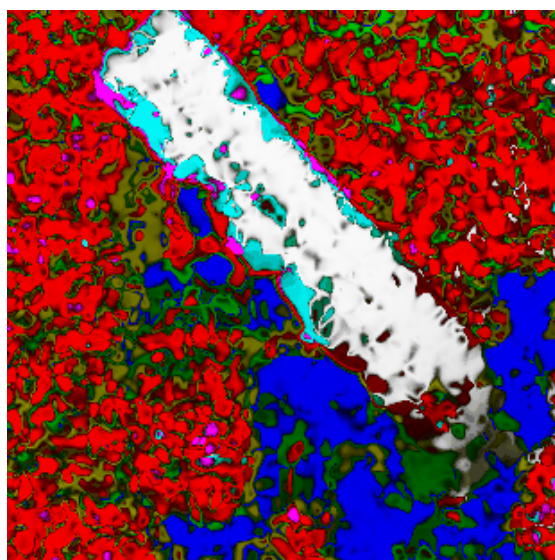


図 5.2: マニピュレータを用いた実験 (a) 原画像 (b) オプティカルフロー

角  $\phi$  となり、 $\phi = 0, 10, 20, 30[\text{deg}]$  について実験画像を撮影した。それぞれの傾き角において、 $\theta$  を固定した画像を 1 フレーム目として撮影し、その後  $\theta$  を  $1[\text{deg}]$  ずらして（円筒の先端での移動距離は約  $1[\text{cm}]$  となる）、つまり角速度を  $\omega = 1 [\text{deg}/\text{frame}]$  に設定して、2 フレーム目として画像を撮影した。ここではノイズを抑制するために、同じ角度設定の画像を 10 枚ずつ撮影し、画素値を平均した平均画像をそれぞれ作成した。図 5.2(b) に、二枚の平均画像から計算されたオプティカルフローを示す。

それぞれの傾き角において領域分割を行なった結果を図 5.3 に示す。この図では異なるクラス毎に色分けし、輝度は重みに比例するように表示しているため、暗ければ暗いほど、クラスに属する確率が小さいことを表している。この領域分割によって回転しているアームの部分が良好に分割されていることが分かる。領域分割によって生成された各領域において、推定された運動パラメータから角速度を計算し、最も角速度の大きい領域を抽出した結果を図 5.4 に示す。背景の領域は回転運動をしてはいないが、パラメータから角速度を計算することはできる。しかしその値は回転しているアーム部分の領域の角速度に比べると極めて小さく、容易に判別することができる。

それを示すのが図 5.5(a) である。これは  $\phi = 0$  のときの各領域の推定された角速度  $\omega_i$  を、アルゴリズムの各反復毎にプロットしたグラフである。色分けは図 5.3 と同じであり、マニピュレータの領域は紫色の線に対応する。マニピュレータの角速度（紫）は十数回の反復で 0.9 付近に収束しているが、それ以外の背景領域の角速度は 0 に近く、十分判別できる。また図 5.5(b) は、 $\phi = 0$  のときの各領域の重み確率  $\xi_i$  の変化の様子をプロットしたグラフである。赤と青の領域はいずれも背景であり、同じようなパラメータであるために競

(a)  $\phi = 0$ (b)  $\phi = 10$ (c)  $\phi = 20$ (d)  $\phi = 30$ 図 5.3: 傾き角  $\phi = 0, 10, 20, 30$  [deg] の時の領域分割の結果

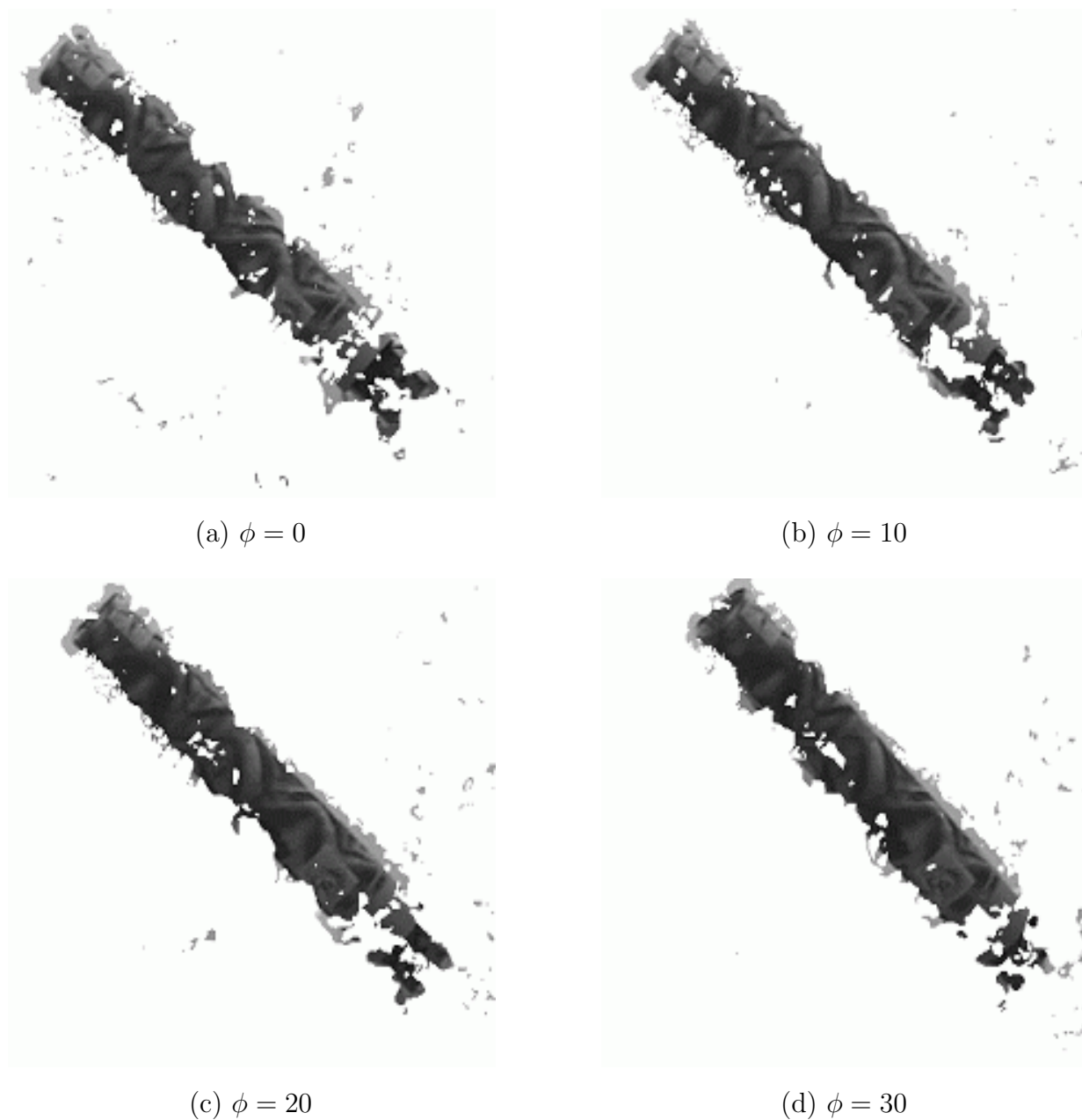
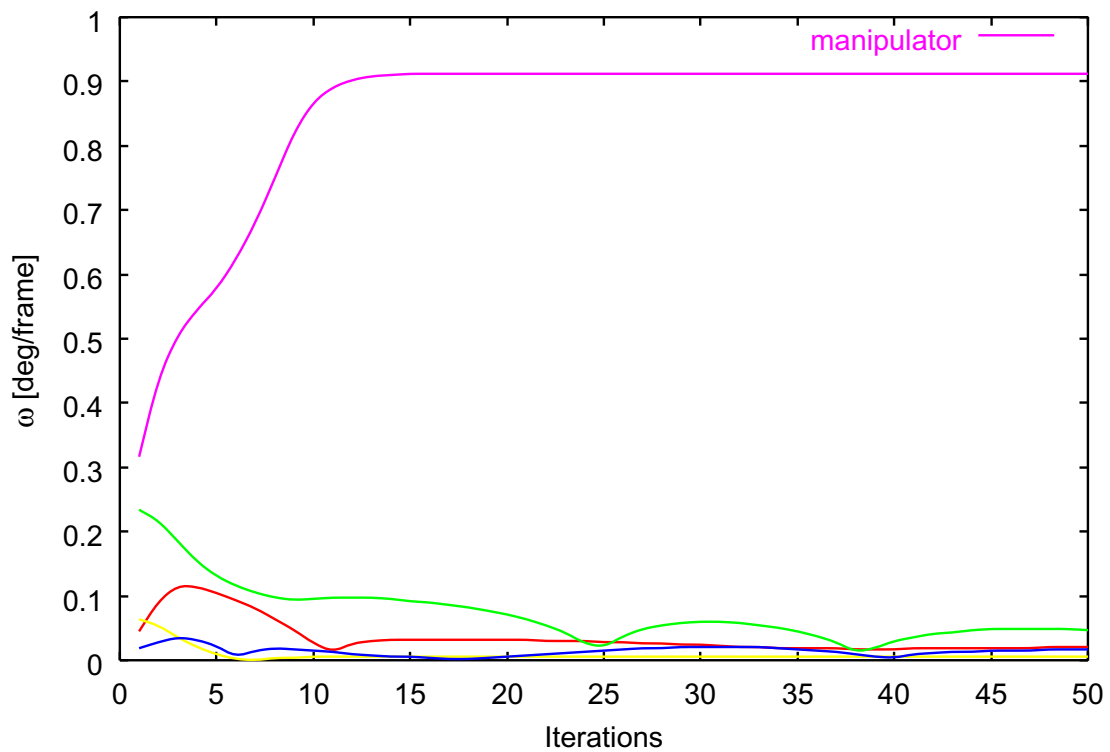
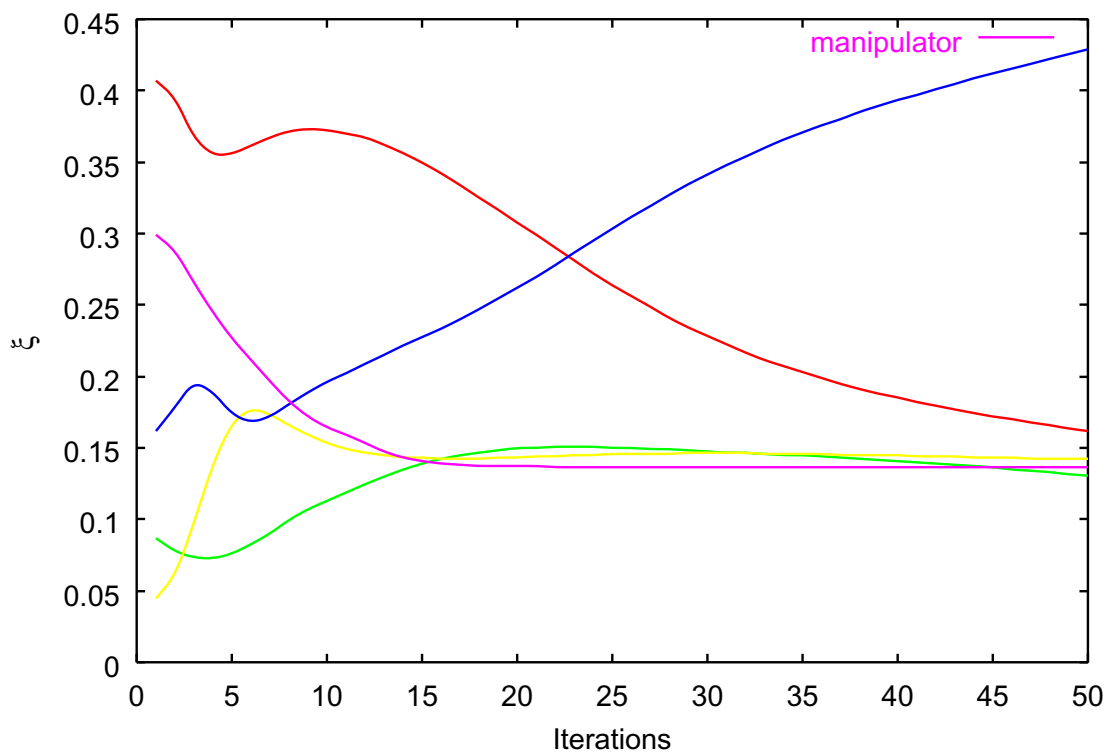


図 5.4: 傾き角  $\phi = 0, 10, 20, 30$  [deg] の時の領域抽出結果



(a)



(a)

図 5.5: (a)  $\phi = 0$  のときの各領域の角速度  $\omega$  推定過程。色分けは図 5.3(a) と同じで、マニピュレータの領域の角速度は紫色の線で示してある。(b)  $\phi = 0$  のときの各領域の重み確率  $\xi$  の変化の様子。

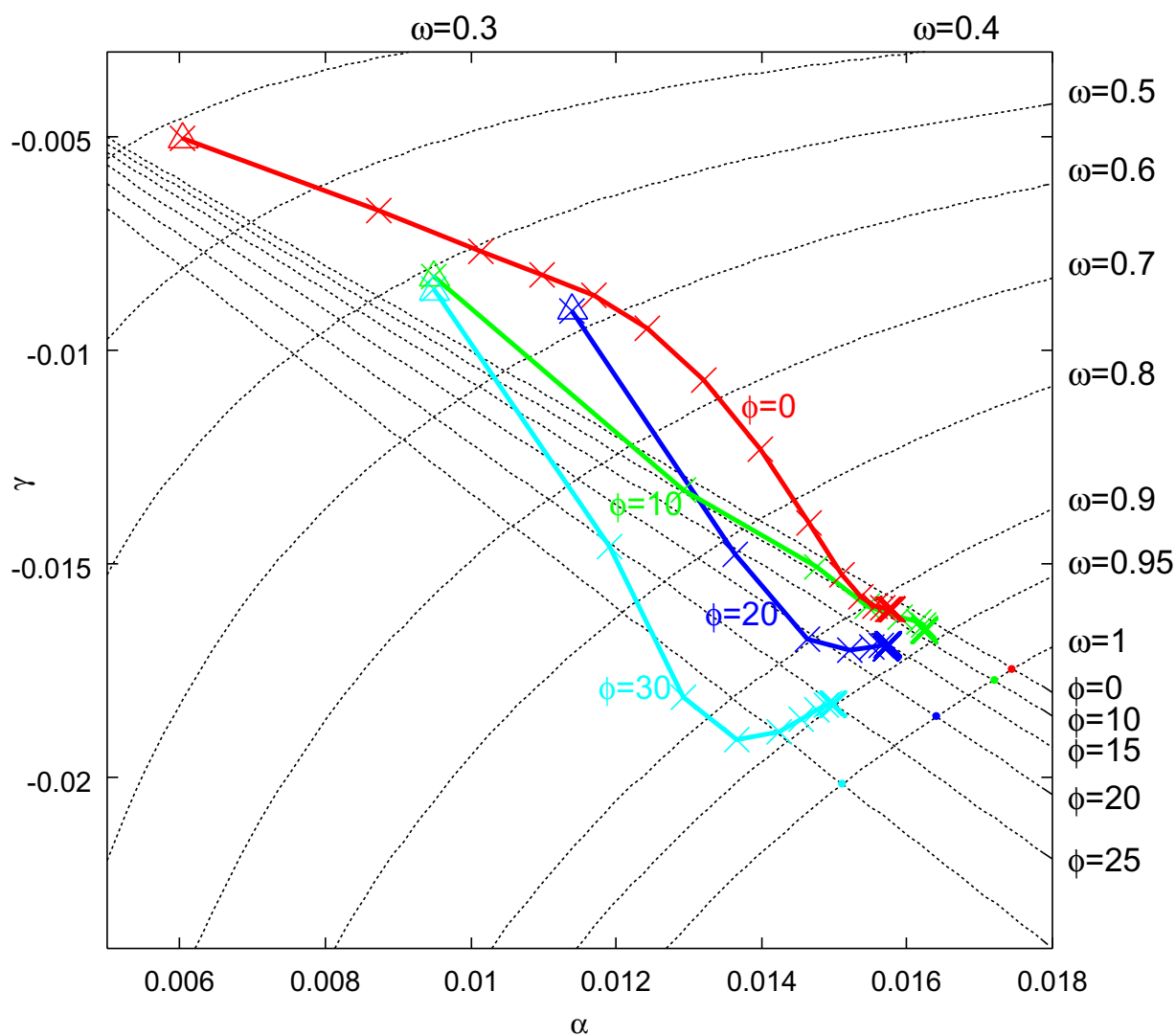


図 5.6:  $\alpha$ - $\gamma$  空間で表されたマニピュレータの運動パラメータの推定過程。傾き角  $\phi$  によって色分けしてある。初期推定値を  $\triangle$  で、各反復毎に推定された値を  $\times$  で、理想値を  $\bullet$  で表してある。



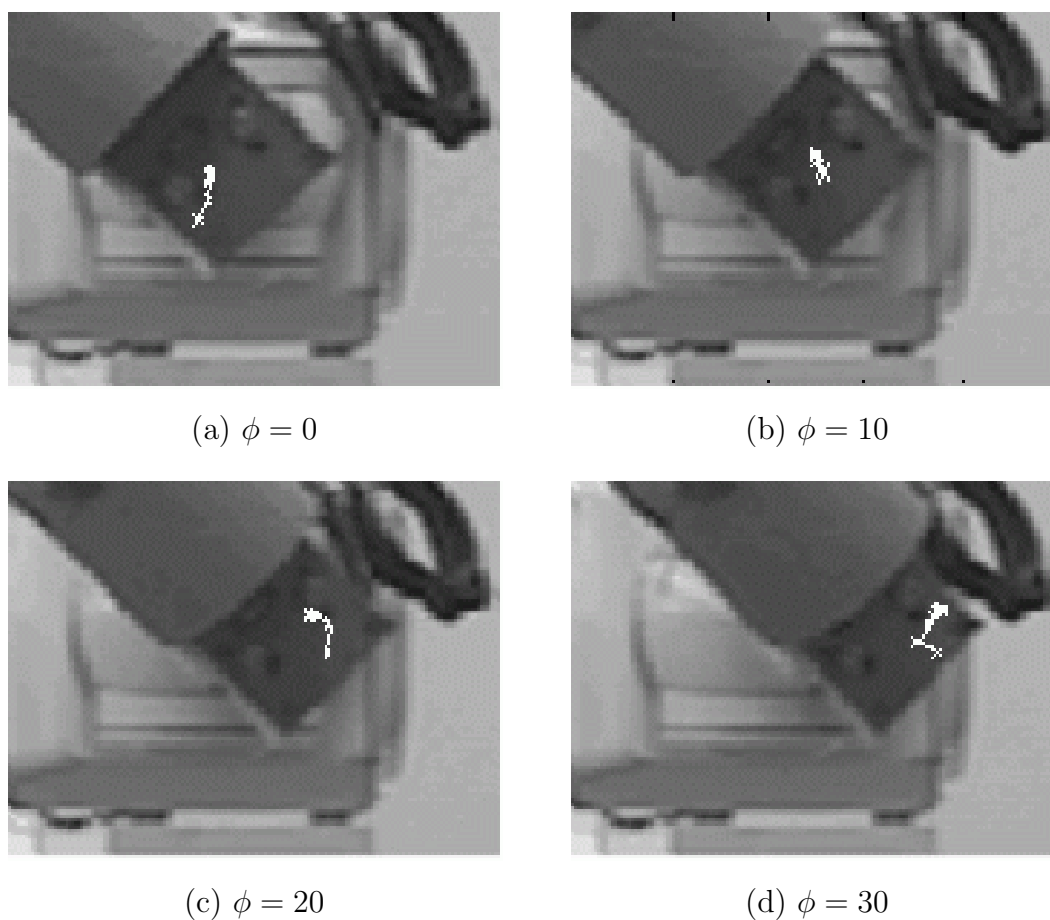


図 5.7: 傾き角  $\phi = 0, 10, 20, 30[\text{deg}]$  の時のマニピュレータの回転中心の推定結果。二つのボルトの midpoint が真の回転中心である。

表 5.3: 異なる  $\phi$  に対する推定結果

$\phi$	0	10	20	30	[deg]
$\phi$ の推定値	7.18	7.97	15.8	25.2	[deg]
$\omega$ の推定値	0.913	0.939	0.936	0.948	[deg/frame]
$\cos \phi$	1	0.9848	0.9397	0.8660	
$\cos \phi$ の推定値	0.9923	0.9903	0.9624	0.9048	

合して収束は遅いが、その他の領域についてはやはり十数回の反復で収束している。

角速度と傾き角の推定過程を  $\alpha$ - $\gamma$  座標でプロットしたグラフを図 5.6 に、最終的な推定値を表 5.3 に示す。 $\omega$  ( $=1$ ) は 10% 程度の誤差で約 0.9~0.95 [deg/frame] と推定された。この誤差の原因には、角速度が小さくなるにつれて最小二乗法による推定精度は悪くなること、また回転中心付近で生じるフローが一画素よりも小さくなるため、うまく計算されないことなどがあげられる。一方  $\phi$  の推定値は比較的誤差が大きい ( $\cos \phi$  は誤差が小さいが) ので、こちらの精度向上は今後の課題である。

この時に推定された回転中心の結果を図 5.7 に示す。この図は、反復計算で回転中心座標の推定値 (白線) が収束していく様子を、回転中心付近を拡大した画像 (100×80) に重ねてプロットした。この図から分かるように、どの傾き角での画像においてもほぼ真の値に近いところに収束していることが分かる。

### 5.5.2 実画像系列からの腕領域抽出

次に、実際に動作を行っている腕の画像を用いて実験を行った。一般的な屋内環境を想定して、背景には壁などを避けて多数の機材が写るようにしたほかは手を加えず、人物の服装に関しても何も操作をしなかった。動作はカメラから約 3m 離れて行い、毎秒 10 フレームで撮影した。インターレースの影響を排除するために、撮影された画像を間引いて 1/4 のサイズ (340 × 223) に縮小した。

実験で撮影した画像系列の、33 フレーム目の画像と、そのオプティカルフローをそれぞれ図 5.8(a)(b) に示す。この画像系列では、正面を向いた人物が右腕を水平位置から下に振り下ろすという運動を行っており、その際意図的に肘を曲げず肩を動かさないようにした。この運動は途中で停止させることなく、また画像面に対してほぼ平行に行っている。毎秒 10 枚という撮影条件の関係上、動きが速すぎるとフレーム間での対応が取れなくなるため、水平位置から下にきて止まるまでに約 3 秒をかけた。

アルゴリズムの反復によって領域分割結果が変化する様子を見るために、アルゴリズム

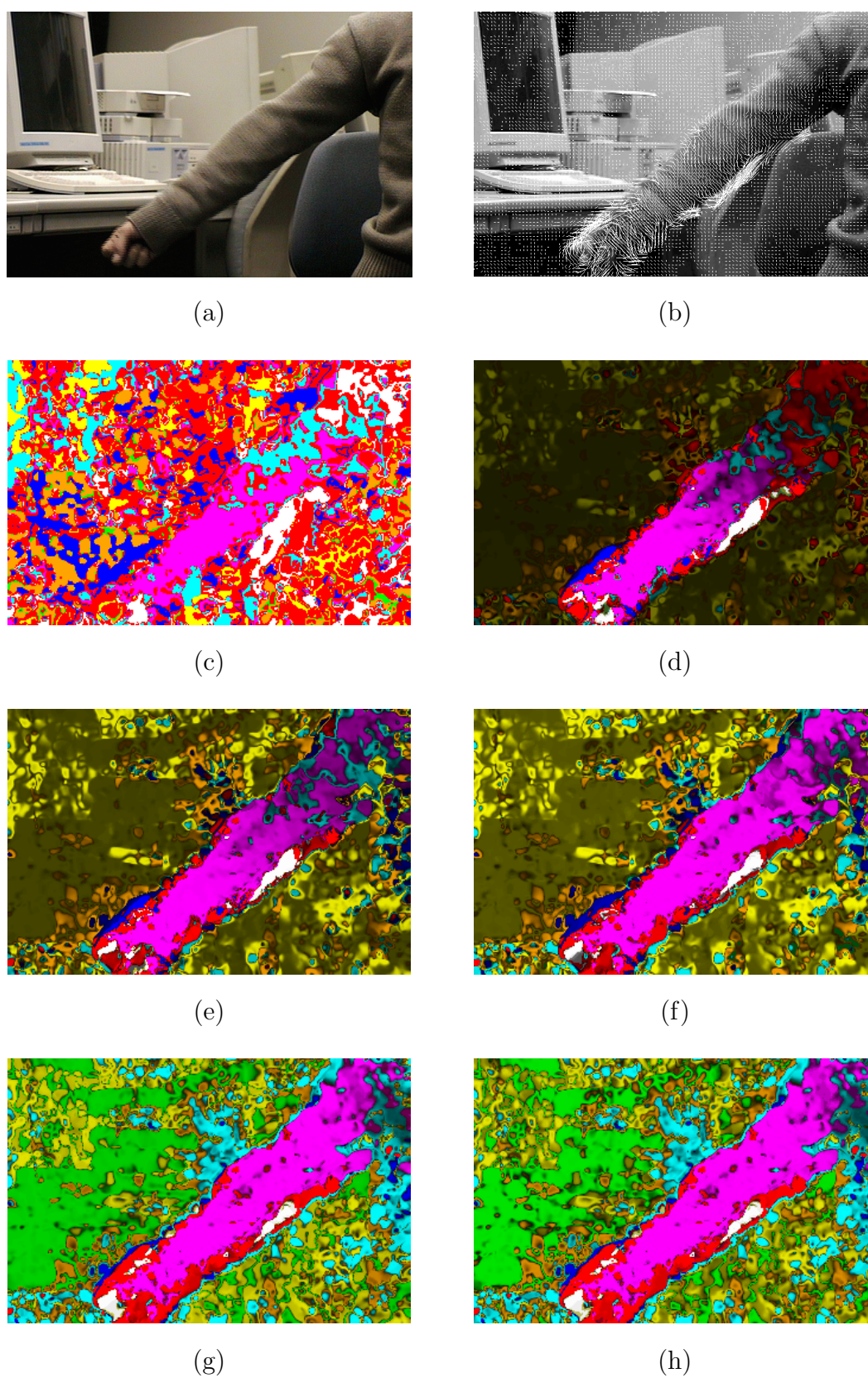


図 5.8: 実画像に対する実験結果 (a) 第33フレーム原画像 (b) オプティカルフロー (c) 初期分割 (d) 反復1回目の領域分割結果 (e) 反復2回目の領域分割結果 (f) 反復3回目の領域分割結果 (g) 反復10回目の領域分割結果 (h) 反復20回目の領域分割結果

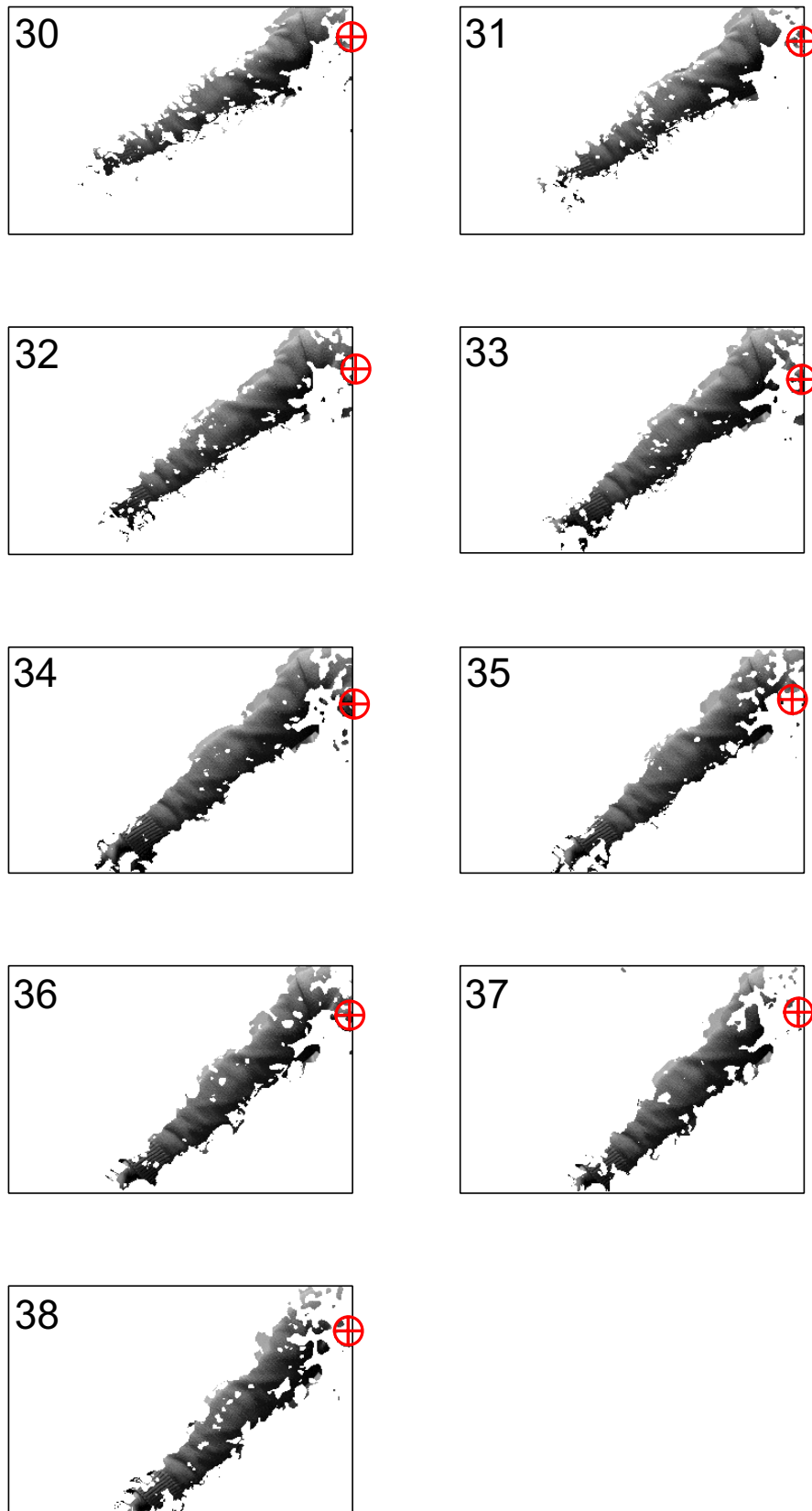


図 5.9: 30~38 フレームにおける腕抽出結果。記号 ⊕ は推定された回転中心位置を示す。

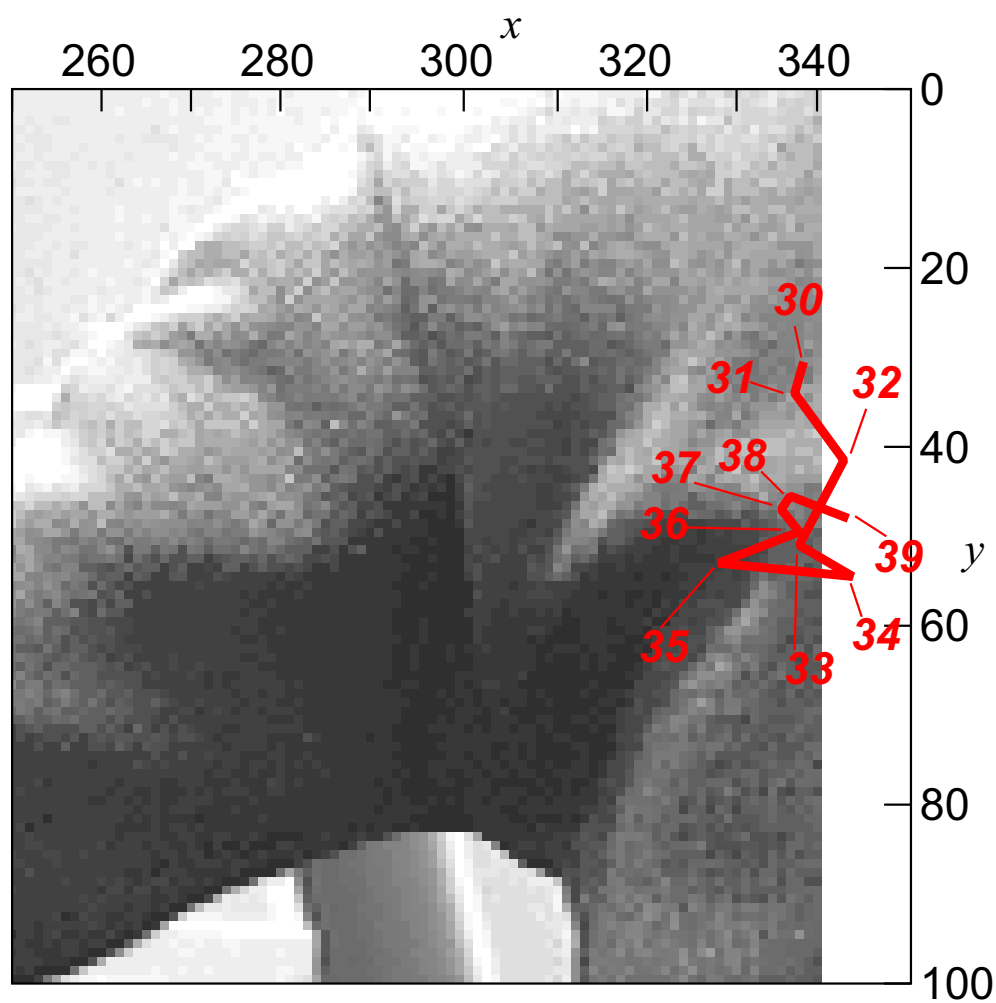


図 5.10: 各フレームで推定された回転中心の推定値の軌跡。赤い数字はフレーム番号。原画像（第30フレーム）の肩のあたりに重ねて表示している。

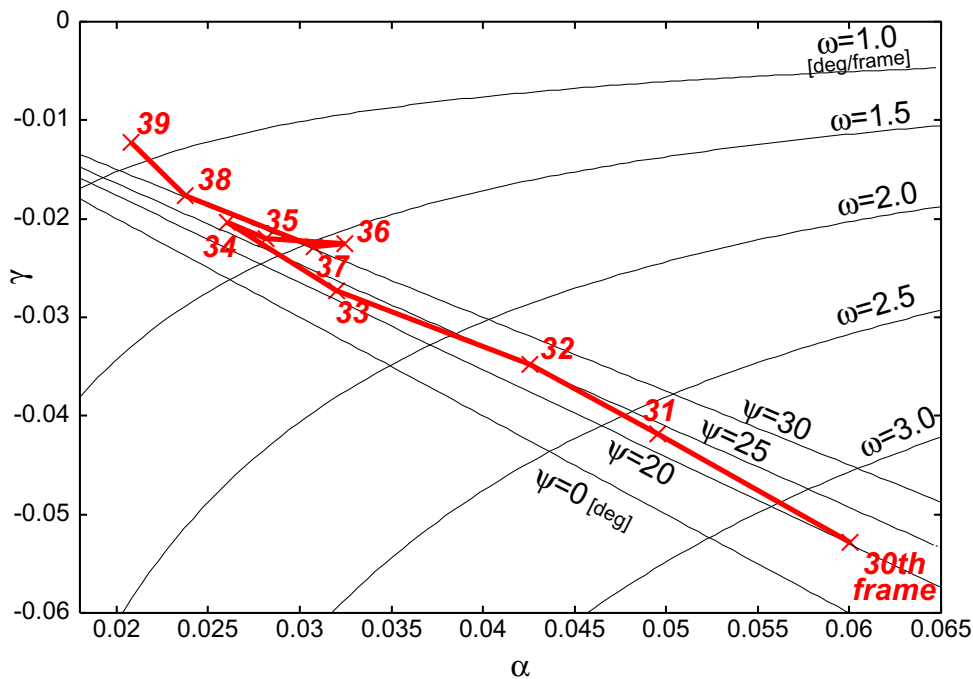


図 5.11: 角速度と傾き角の推定値の軌跡 (太線)。細線は  $\alpha\gamma = -\omega^2$  と  $\alpha = -\cos^2 \phi\gamma$  の曲線。

中の式 (5.33) によって重みが更新されるたびに、ステップ 9 によって領域分割を行った (図 5.8(d)~(h))。図 5.3 と同様に、この図では異なるクラスタ毎に色分けし、輝度は重みに比例するように表示しているため、暗ければ暗いほど、クラスタに属する確率が小さいことを表している。

初期分割 (図 5.8(c)) では、上腕と下腕が別々の領域に分かれてしまっているが、数回の反復によって、肩から手首まで腕全体が一つの領域に覆われていることがわかる。領域分割は最初の数回の反復で大きく変化し、反復が 10 回を超えるとほとんど変化しなくなり (図 5.8(g),(h))、パラメータ推定の方もほぼ収束する。

反復 20 回目 (図 5.8(h)) では、クラスタ数が物体の数よりも多いので、背景が複数のクラスタに分割されてしまっている。そのため、背景がすべて黒くなっている反復 1 回目 (図 5.8(d)) のほうがよい分割結果であるように見えるかもしれない。しかし、前述したように、クラスタに属する確率が低いほど暗く表示しているため、1 回目の結果の背景部分の暗い画素は、背景に (そしてどのクラスタにも) 属する確率が低く、信頼できない結果であることを示している。反復が進むにつれて、それぞれのクラスタに属する確率が高くなる (したがって明るい表示に) になっていくので、過分割にはいるが反復 20 回目の結果の方が良い結果であると言える。このように過分割とならないために AIC などを用いてクラスタ数を決定する [栗田 98] ことについては今後の課題である。

画像中で最も大きい角速度を持つ領域が腕の領域であると仮定して、同じ画像系列の

30~38 フレームにおいて腕領域を抽出した結果を図 5.9 に示す。この図が示すように、実画像から得られるノイズが含まれたオプティカルフローを用いて、本手法は安定して腕領域を抽出できることがわかる。この図ではそれぞれのフレームで  $\oplus$  によって推定された回転中心を表示しているが、それらを一つの画像に重ねて表示した軌跡を図 5.10 に示す。回転中心は次第に下がっていき、系列の後半ではやや肩の下辺りで留まっているのがわかる。他の推定されたパラメータの軌跡を図 5.11 に示す。このグラフは、パラメータ  $\omega$  と  $\cos\phi$  が式 (5.23) において  $\alpha$  と  $\gamma$  で表されているので、それらを  $\alpha$ - $\gamma$  空間にプロットしたものである。腕は振り下ろされるにしたがって遅く ( $\omega$  が小さく) なっているが、 $\psi$  はそれほど大きく変化していないことがわかる。

### 5.5.3 状況を変化させた場合の実験結果

他の画像を用いた実験結果を図 5.12 と図 5.13 に示す。

複数の腕が同時に運動している場合の実験 (図 5.12) は、左側の人物が左腕を振り下ろし、右側の人物は右腕を振り上げている、という画像系列を用いた。実験環境は、背景が異なるほかは図 5.8 の実験と同じである。ここでは二つの運動が存在するので、最大角速度を持つ領域を抽出すると片方の人物の腕しか抽出することができない (図 5.12(g))。そこで、表 5.1 の回転運動であるという制約条件を満たし、かつ角速度がしきい値 (ここでは 3 deg/s とした) 以上である領域を抽出することで、それぞれの腕を抽出することができた (図 5.12(h))。

図 5.13 は、背景に移動物体が存在する場合の実験である。図 5.8 の実験と同様にある人物が左腕を振り下ろしており、その背後約 1m のところを、別の人物が右から左に移動しているという画像系列を使用した。後ろの人物が背後を通過するのに合わせて手前の人物が腕を動かしており、両者の領域は接しているが、歩いている人物の領域は回転運動を含まないため、5.4.2 節で述べた条件によって除去され、腕の領域だけが抽出された (図 5.13(g))。

これらの実験で、本手法の目的である「背景に依存せず、他の移動物体が存在する状況」での抽出が行えることが示された。

### 5.5.4 対象を変化させた場合の実験結果

前項では複数の物体が存在するなど状況を変えて実験を行ったが、ここでは肘を回転中心とした場合や領域の大きさを変えた場合など、対象を変化させて実験を行った結果を示す。

回転中心を肩ではなく、肘にした場合の結果を図 5.14 に示す。この場合は腕全体であった対象が下腕になり、回転モデルで用いているリンクの長さが変化するだけであるので、肩を回転中心とした場合と何ら変わることなく、領域を抽出することができる。

次に、対象となる腕の大きさを変化させて抽出した結果を図 5.15 に示す。カメラのズームを変え画面の中での腕の大きさを変化させるたびに同じ動作を撮影し、オプティカルフ

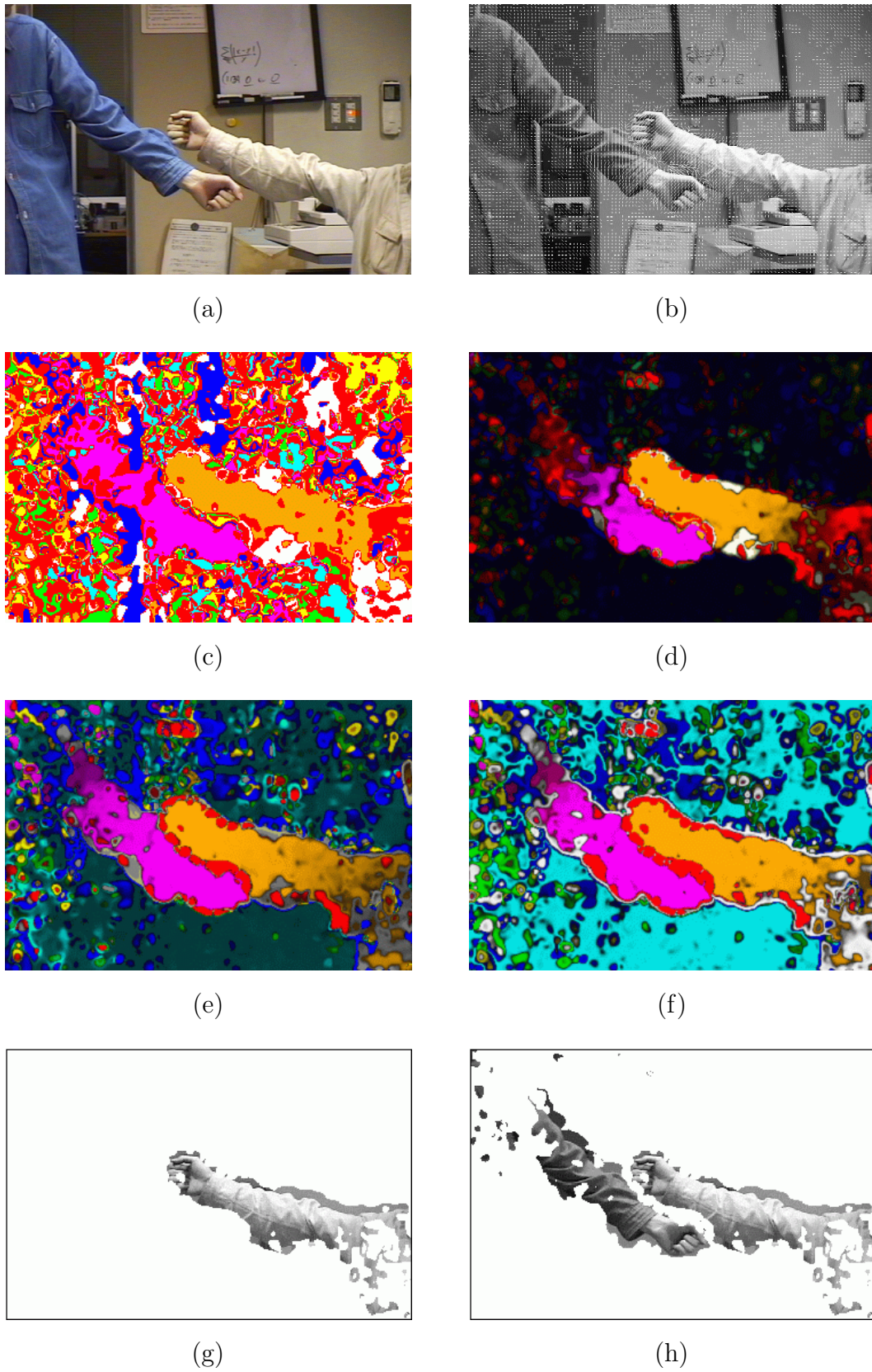


図 5.12: 二本の腕が運動している場合 (a) 原画像 (b) オプティカルフロー (c) 初期分割 (d) 反復1回目 (e) 反復5回目 (f) 反復10回目 (g) 最大角速度を持つ領域の抽出 (h) しきい値以上の角速度を持つ領域の抽出



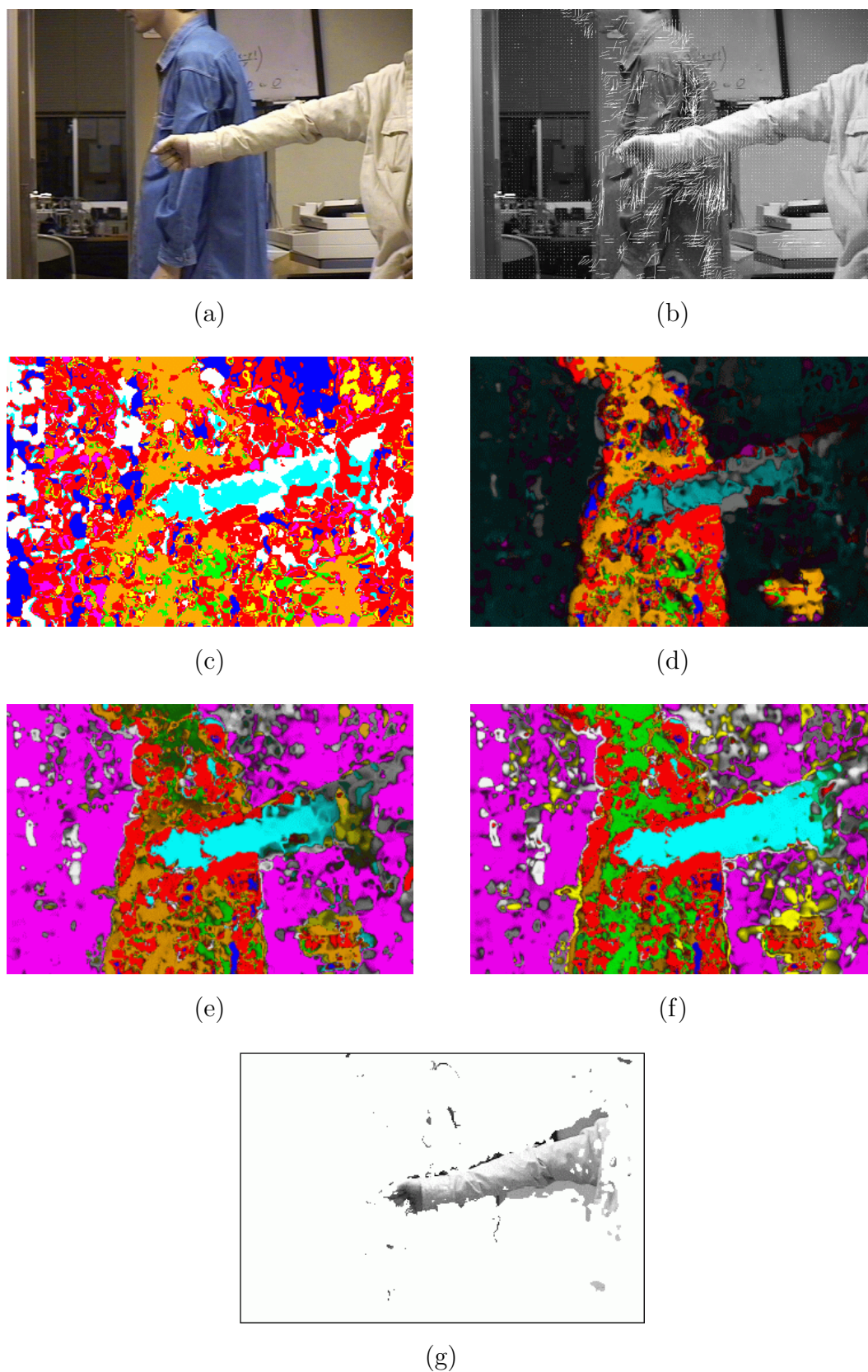


図 5.13: 背後に移動物体が存在する場合 (a) 原画像 (b) オプティカルフロー (c) 初期分割 (d) 反復1回目 (e) 反復10回目 (f) 反復20回目 (g) 最大角速度を持つ領域の抽出

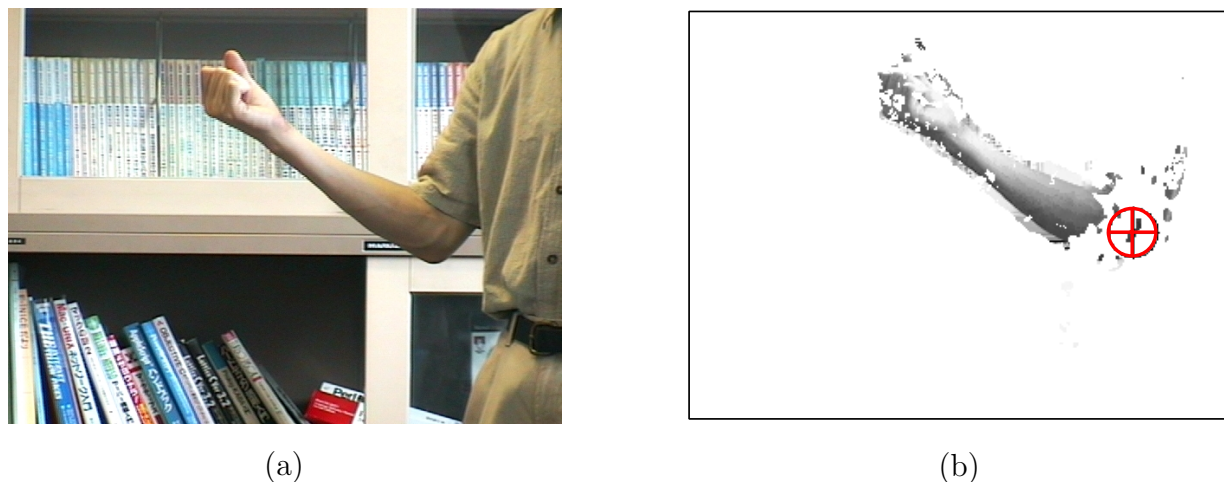


図 5.14: 肘を回転中心とした場合 (a) 原画像 (b) 抽出結果

ローを計算した。それぞれの画像系列で腕の大きさが異なるが、比較のために、腕の位置がほぼ同じ場所にあるフレームでの抽出結果を並べた。

大きく写っている場合（上の方の画像）は問題ないが、画像中に写る腕が小さくなるにつれ（下の方の画像）上腕部の抽出が困難になった。これは画像内での大きさが小さくなるにつれて計算されるオプティカルフローの精度が悪くなっていくためである。また、これ以上腕が小さく写る画像では、速度の速い腕先端部のオプティカルフローさえもノイズに埋もれてしまい、抽出することができなかった。

したがって、オプティカルフローの計算手法に依存してしまうとはいえ、対象人物の腕がある程度大きく写っていることが必要となる。

## 5.6 おわりに

この章では、オプティカルフローの curl に基づいた、回転する腕の領域抽出と、その運動パラメータを推定する手法を提案した。提案手法は、オプティカルフローの勾配による領域分割と運動パラメータの最小二乗法による推定からなり、EM アルゴリズムを用いて推定と領域分割を同時に行う。そして推定されたパラメータに基づいて、画像内に存在する並進運動と回転運動を区別することができ、背景で移動する物体を除去することができる。提案手法を実装し実画像に対して実験を行い、パラメータ推定には多少の誤差を含むが、実画像から得られたノイズを含むオプティカルフローからでも良好に腕の領域を抽出することができることを確認した。また複数の腕が同時に動作している場合や、背後に移動物体が存在する場合についても実験を行い、実環境における本手法の有効性を示した。

本手法はオプティカルフローが正確に計算されることが必要であるので、対象が画像中で小さくなりすぎるとそのフローがノイズに埋もれてしまい、抽出が困難になる。また本

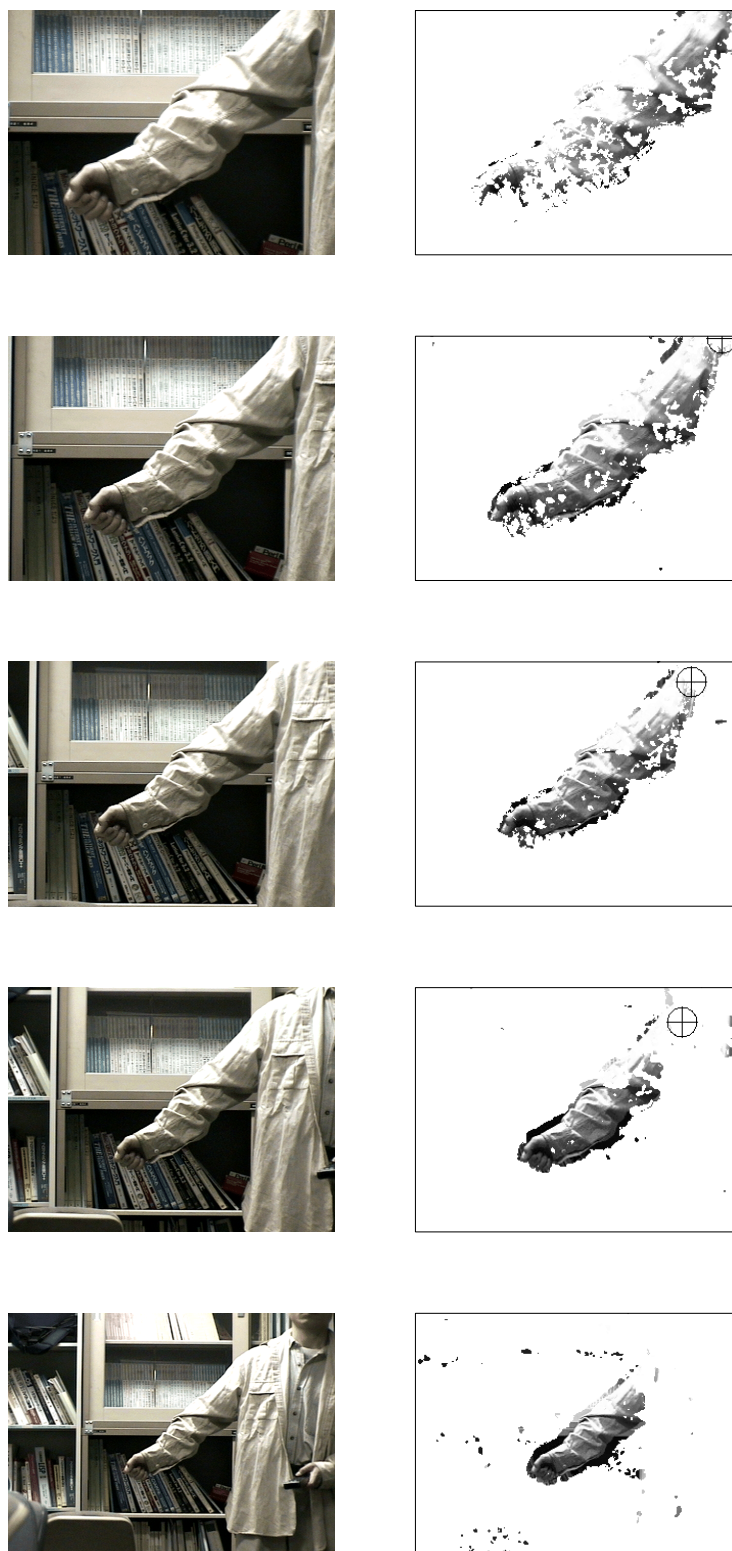


図 5.15: 画像内での領域の大きさを変えた場合の抽出結果。上の方が大きく、下の方が小さい領域。

手法は正射影を仮定しているが、これは透視投影の近似であり（付録C参照）、ノイズのある実画像のフローではこの近似はノイズに埋もれてしまい、それほど影響はないと思われる。

本手法は  $\phi$  と  $\psi$  の一方が0である場合のみを扱っているが、一般的な三次元運動 ( $\phi \neq 0$ ,  $\psi \neq 0$ ) については、式(5.17)の代わりに、 $u$  と  $v$  を  $\alpha x + \beta y + \gamma$  という平面を用いて近似することで、5.3節の手法が適用できるように拡張することができる。また、本手法はカメラが移動したときや、肘などの関節がある場合を考慮していない。これらはいずれも、式(5.2)で述べた定数ベクトル  $t$  が存在する場合であり、運動モデルが異なるため、本手法をそのまま適用すると回転中心など推定されるべき運動パラメータに影響を与える[朴99]。これについては、 $t$  の時間的变化に（線形であるなどの）仮定を設けることで、パラメータを補正することができ、カメラの移動や関節物体を扱うことができると考えている。

## 第6章 結論

本論文では、画像から対象物体や対象人物を抽出するという処理をより一般的な環境において用いようとする際に問題となる点を明らかにし、静止画像からの物体抽出と動画からの人物抽出において、それらの問題点を克服する手法について述べた。

1章では本研究の背景と人間の視覚において行なわれている抽出処理過程を概観し、応用分野からみたその有用性について述べ、従来研究と比較した本研究の特徴を明らかにし、本研究の位置づけを行なった。

2章では、物体抽出処理と密接な関係にある領域分割の、計算機で実装される代表的な処理についてまとめた。

3章では、物体についての固有の知識を用いずに、過分割の原因となる物理的および光学的要因を排除し、静止画像から物体を抽出する方法について述べた。過分割されている領域を併合する基準として、対象物体に依存しない領域間の情報に基づいた四つの特徴量（包含、滑らかさ、連続性、面積比）を導入し、分割併合法に基づく併合アルゴリズムを提案した。提案手法を用いた実画像に対する実験を行ない、物体の領域をほぼ抽出することができた。そしてどの特徴量が効果的に働いているかの考察をし、また他の手法との比較を行なった。

4章では、人や車などの複数の移動物体が存在する動画から、簡単な人物の形状モデルを用いて人物全身の領域を抽出する方法について述べた。フレーム間差分とエッジから抽出された円形部から人物の頭部位置候補を決定した。それを元に作成された楕円モデルを動的輪郭手法により人物領域に収束させて、領域分割によって人物の詳細な輪郭を抽出した。提案手法を用いて実画像系列による実験を行ない、34画像系列全170枚の画像中の123枚（約72%）において人物領域が得られ、14系列の全てのフレームで人物を抽出することができた。

5章では、複数の運動物体が存在する動画から、オプティカルフローを計算し、回転運動モデルを用いて人物の腕の領域の抽出とその運動パラメータの推定を行なう方法について述べた。動画から求められたオプティカルフローを、その方向によって大まかに分割を行い、EMアルゴリズムを用いて異なる運動物体の領域に分割し、同時にその運動パラメータを求めた。推定されたパラメータに基づいて非回転運動領域を除去する条件を示した。実画像による実験で、複数の腕が運動している場合や、背後に移動する物体が存在す

る場合でも人物の腕が抽出できることを示した。

今後の課題としては以下の事柄が挙げられる。

静止画像からの物体抽出においては、本研究で提案した特徴量だけが唯一ではない。しかし、領域分割の研究は「どうやって分割するか」という手法だけが先行して発展しており、「何を分割するか」という特徴量や基準の考察や検討が置き去りにされている。さらなる併合の基準やクラスタリングの特徴量を採り入れるためには、シーンの理解という点に立ち戻って、どのような特徴や基準を用いるべきかを考える必要がある。

動画からの物体抽出については、モデルをより積極的に用いること、また動きだけでなくさまざまな画素の情報を有効に利用することが重要である。今回用いたような簡単な楕円モデルではなく、より細かく人物を表すロボットモデルなどを用いることで、人物の抽出と同時に動きを解析する手法に発展させることが考えられる。また運動モデルを用いる場合は、より複雑で多用な運動に対処できるようにさまざまなモデルを作成しておくことが考えられる。

また、静止画像に対する抽出処理と動画に対する抽出処理を統合して、より頑健な処理方法を構築することができる。動画に対する処理は、静止画に対してはまったく効力をなさない。これはすなわち、移動していた物体が停止してしまったならば対処できなくなることを意味する。一方静止画の処理は動画の各時刻のフレームに対して処理をすることはできるが、動きという重要な手がかりを無視している。これらの統合により、様々な対象や状況に対して適用できるようになると考えられる。

本研究が将来におけるロボットの視覚情報処理の発展に貢献することを期待して、本論文のまとめとしたい。

# 謝辞

本研究を進めるにあたり、多くの適切な御指導と御助言を頂きました名古屋大学情報メディア教育センター基礎系 大西昇 教授（前名古屋大学工学研究科情報工学専攻教授）に深く感謝致します。大西先生には、学部生の頃から通算して六年間、指導教官として本研究を進める機会を与えて頂きました。また、明確な研究指導と、熱心かつ的確な御教示、時宜を得た激励、そして絶えまない援助を頂いたことにより、本論文をまとめることができました。大西先生の御指導、御鞭撻は、今後の研究活動の支えでもあります。ここに、心から感謝の意を表します。

また本論文の作成にあたり、多忙な中時間を割いて本論文を丁寧に読んで頂き、貴重な御意見と激励を頂いた名古屋大学工学研究科情報工学専攻 鳥脇純一郎 教授、名古屋大学工学研究科計算理工学専攻 末永康仁 教授に深く感謝致します。

本研究を進めるにあたり、多くの方々から有益な助言と討論を頂いたことに深く感謝致します。愛知県立大学情報科学部 山村毅 助教授（前名古屋大学工学研究科情報工学専攻講師）には、視覚心理学における知見を数多く教示して頂き、また長時間に渡る熱心な議論を幾度となく交わして頂きました。取分け本文第3章の研究は山村先生なくして進めることはできませんでした。名城大学理工学部情報科学科 田中敏光 助教授（前名古屋大学大型計算機センター助教授）には、毎回の研究会での討論において数々の助言をいただきました。また、名古屋大学情報メディア教育センター基礎系 工藤博章 助教授（前名古屋大学工学研究科情報工学専攻講師）、松本哲也 助手、竹内義則 助手（前名古屋大学工学研究科情報工学専攻助手）、並びに大西研究室の皆様には、熱心な討論と有益な意見を頂きました。さらに、さまざまな心遣いを頂くとともに、迅速かつ丁寧な事務を行なって頂いた名古屋大学工学部情報工学科事務室の皆様、並びに大西研究室の秘書の方々に感謝致します。

本研究の一部は、平成11年度から12年度にかけて、理科学研究所 バイオミメティックコントロール研究センター（名古屋市守山区）にて、ジュニア・リサーチ・アソシエイト（JRA）として携わったものです。機器を使用させて頂きました理科学研究所バイオミメティックコントロール研究センター 故 伊藤正美 前センター長、佐田登志夫 センター長に深く感謝致します。理研という環境で研究をさせて頂いたことは、掛替えのない財産であると思っております。また、生体ミメティックセンサ研究チームの皆様には、熱心な討論と数多くの有益な意見を頂きました。指導教官でもある 大西昇 チームリーダー、コンピュー

タビジョンの先輩研究者として数々の指摘を頂きました 向井利春 元研究員、数学的素養の必要性を教えて頂いた Ali MANSOUR 研究員をはじめとする研究員とJRAの皆様、並びにセンサチームの秘書として色々心配りを頂いた 近藤亜紀 様に感謝致します。また、生体ミメティックコントロール研究チーム 小田島正 研究員には、研究者として、また先輩として、大学時代から様々な形で支えて頂き、その博識と慧眼に啓発されることは数多くありました。本当にありがとうございました。

本研究の成果を国際学会にて発表するための渡航費用は、財団法人電子通信普及財団海外渡航費用援助平成10年12月期、財団法人西秋奨学会1999年度海外渡航旅費助成、社団法人電子情報通信学会東海支部平成11年度後期学生研究発表助成、の助成によるものです。

最後に、長期間に渡る教育を受けることを快く支え、後押しをしてくれた両親と家族に心から感謝します。

2001年1月



## 付録 A 滑らかさの特徴量 $V_S$ の特性

本論文3章において、滑らかさの特徴量を式(3.2)で定義した。この特徴量は境界線が直線のとときに1となり、滑らかでなくなるほど0に近くなるが、他の特徴量と異なり、直感的に数値を把握できない。そこで、実際にどの程度の値を取るのかをシミュレーションで実験した。

図A.1に示すように、画像中央を中心とする輝度勾配を指定した濃淡画像のグラデーション(400×200)を生成し、その後輝度値にガウスノイズを付加した。輝度が[0,255]の範囲を超えた場合は0または255にした。次に、輝度0と255を中心とするクラスタリングを行った。生成された領域に対して包含と面積比の特徴量を用い(アルゴリズム Step1,2)、大きな二つの領域が形成されるまで併合を行った。

ガウスノイズの分散  $\sigma = 100$  の場合について、輝度の傾きを0.75から50まで変えて生成したグラデーションと、領域の境界線を図A.1に示す。傾きが小さいグラデーションはエッジには見ないが、10を超えるあたりから明確な境界線を認識できる。

輝度勾配と滑らかさの特徴量の関係を表すグラフを図A.2に示す。勾配が大きくなるにつれて特徴量の値は大きく(滑らかに)、またノイズが大きくなるにつれて特徴量は小さく(ぎざぎざに)なっているのがわかる。しかし、ノイズが大きい場合であっても、勾配が小さいうちは特徴量は0.5程度だが、勾配が大きくなると0.65を超える。

このことから、滑らかさの特徴量のしきい値  $Th_s$  を0.5~0.65に設定することで、擬似輪郭と物体の輪郭線とを区別することができる。

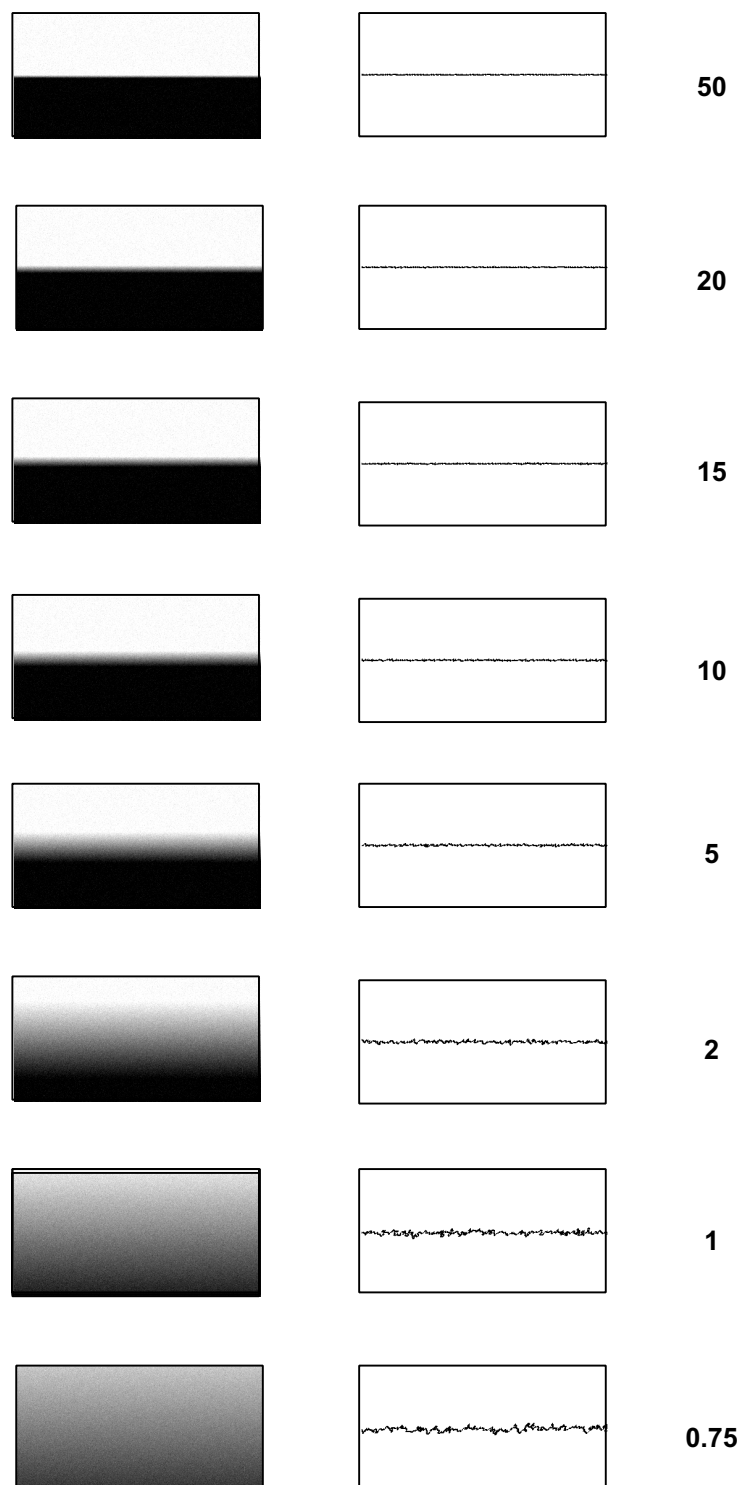


図 A.1: ノイズ ( $\sigma = 100$ ) を付加された濃淡画像 ( $400 \times 200$ ) のグラデーション (左列) とそのときの境界線 (右列)。数字は輝度勾配を示す。

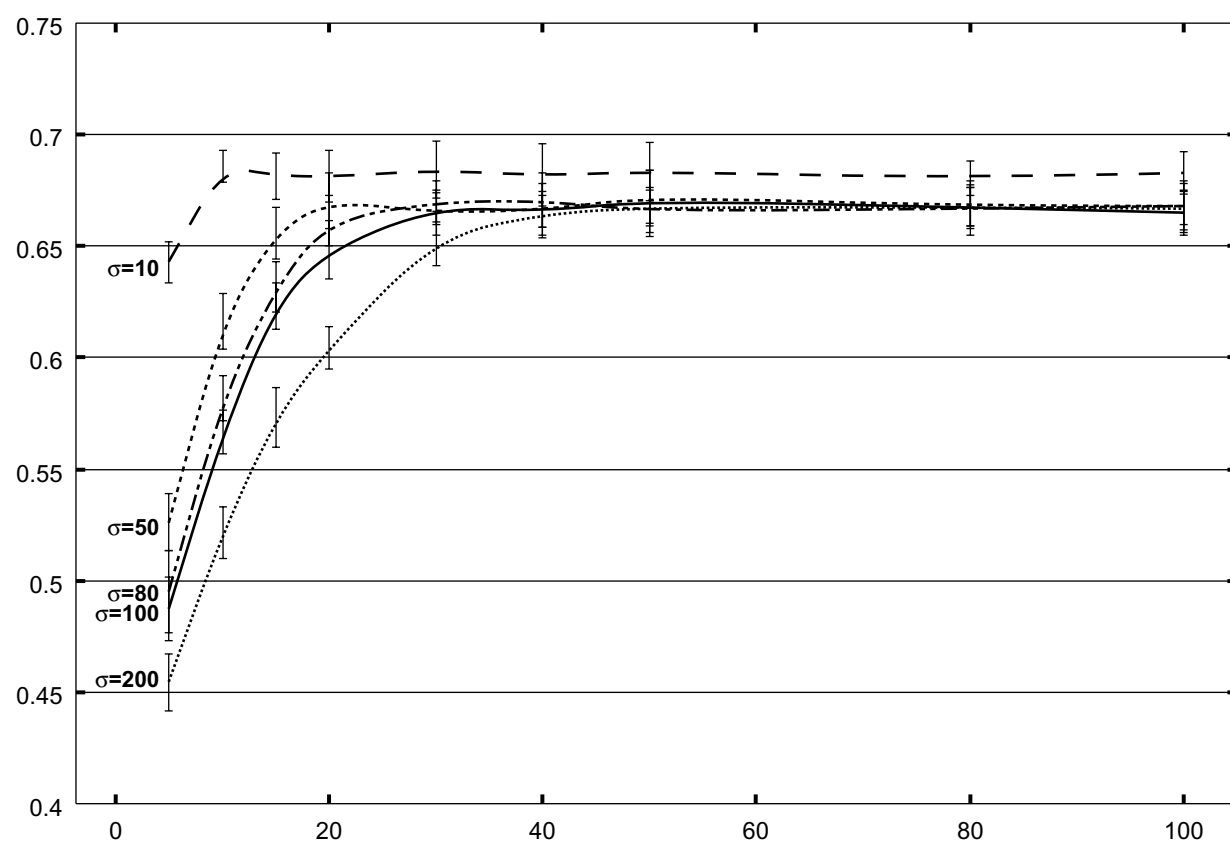


図 A.2: 輝度勾配 (横軸) と滑らかさの特徴量 (縦軸) のグラフ。バーは20回計算した標準偏差を示す。付加したノイズ ( $\sigma = 10, 50, 80, 100, 200$ ) のそれぞれに対してプロットした。



## 付録 B Curl による領域分割の シミュレーション実験

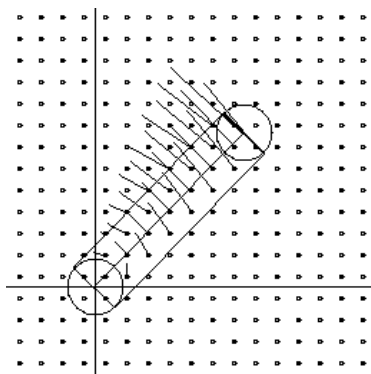
本論文 5.2.3 章において、curl の値による領域分割について述べた。ここでは、人工的に生成したオプティカルフローについて、その curl を計算し、領域分割を行なったシミュレーション結果について述べる。

生成したオプティカルフロー（速度ベクトル  $\dot{p}$ ）を図 B.1(a) に示す。この図の大きさは  $256 \times 256$  で、15 ドット毎に円を起点にしてフローを描画している。用いたパラメータは、回転中心の座標 (50,50)、角速度  $\omega = 0.4[\text{rad/s}]$ 、腕の長さ  $r = 150[\text{dot}]$ 、腕の幅  $20[\text{dot}]$ 、腕の位置  $\theta = 0.8[\text{rad}]$  である。

図 B.1(b) に、フローの curl ( $\nabla \times \dot{p}$ ) を計算した結果を示す。curl の値が大きいほど黒く、小さいほど白く表示している。ここでは微分計算には単純に画素値の差分を用いた。腕の部分に相当する灰色の部分の値は  $0.79 \sim 0.80 \quad 2\omega$  であり、図 B.1(c) に示すように回転しているフローの領域が抽出できることがわかる。

図 B.1(b) はノイズのない場合であったが、フローにノイズを付加して同様の実験を行なった。フローの  $x$  成分と  $y$  成分のそれぞれに  $\pm 5$  の一様乱数を加え、ガウスフィルター ( $\sigma = 2$ ) で平滑化した後、curl を計算した結果を図 B.1(d) に示す。これはノイズの影響で図 B.1(b) よりも明瞭ではないが、平均輝度による領域分割を行なうことで回転物体の領域が分離できる（図 B.1(e)）。

また回転面が傾いている場合 ( $\phi \neq 0$ ) でも、図 B.2 に示すように、同様に領域分割によって抽出することができる。



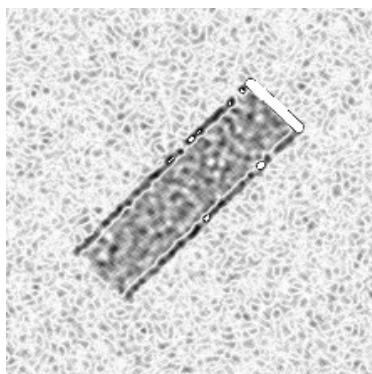
(a)



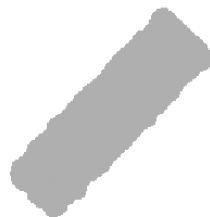
(b)



(c)



(d)



(e)

図 B.1:  $\phi = 0$  の場合の curl の値による領域分割 (a) 生成したオプティカルフロー (b) フローの curl (c) curl の  $z$  成分によって分割した結果 (d) ノイズを付加したフローの curl (e) ノイズを付加したフローの curl の  $z$  成分によって分割した結果

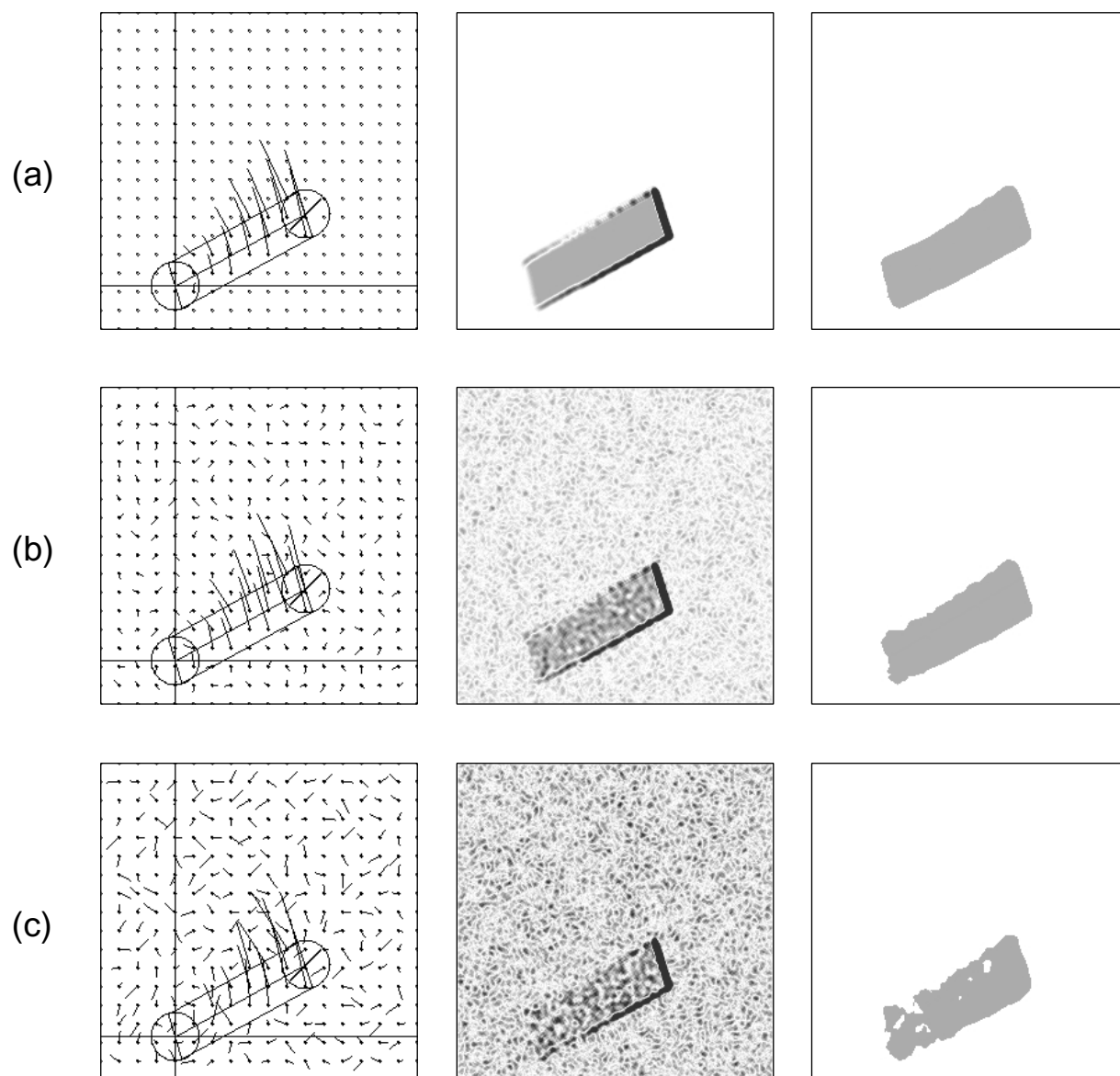


図 B.2: 回転面が傾いた場合の curl の値による領域分割 (256×256)。左下を原点に  $(c_x, c_y) = (60, 35)$ ,  $\omega = 0.5$ ,  $\phi = 0.7$ ,  $\theta = 0.3$  (a) ノイズなし (b)  $\pm 5$  の一様ノイズ付加 (c)  $\pm 10$  の一様ノイズ付加。左列は生成したオプティカルフロー、中列はオプティカルフローの curl の  $z$  成分、右列は領域抽出結果である。





## 付録 C フローの curl に与える 透視投影の影響

本論文 5.2.3 節では、回転モデルに基づいたオプティカルフローの curl による領域分割について述べたが、それらはすべて正射影を仮定していた。しかし、回転面が画像平面と平行であれば、議論の結果は正射影と同じになる。

$H(\cdot)$  を透視投影、 $f$  を焦点距離、 $Z$  を回転する面までの奥行きとして、正射影と同様に透視投影の場合を考える。角速度  $\omega$  で回転し、角度  $\theta (= \omega t)$  の位置にある腕上の点を  $p$  とする。このとき、 $p$  の速度  $\dot{p}$  (画像上でのオプティカルフロー) とその curl  $\nabla \times \dot{p}$  は次のように表される。

$$\begin{aligned} p &= H((r \cos \theta, r \sin \theta, Z)) \\ &= \left( r \cos \theta \frac{f}{Z}, r \sin \theta \frac{f}{Z}, f \right) \end{aligned} \quad (\text{C.1})$$

$$\begin{aligned} \dot{p} &= \left( -\omega r \sin \theta \frac{f}{Z}, \omega r \cos \theta \frac{f}{Z}, f \right) \\ &= (-\omega y, \omega x, f) \end{aligned} \quad (\text{C.2})$$

$$\nabla \times \dot{p} = (0, 0, 2\omega) \quad (\text{C.3})$$

つまり、curl の  $z$  成分が角速度  $\omega$  の二倍になるという正射影と同じ結果になる。これは、スケールが変化しても画像上での角速度は変わらないという理由による。

回転面が画像平面と傾いている時には、傾きの大きさ  $\phi$  と、その物体までの距離  $Z$  が関係してくる。回転面が角度  $\phi$  だけ傾いていた場合 (図 5.1(b))

$$p = H((r \cos \phi \cos \theta, r \sin \theta, r \sin \phi \cos \theta)) \quad (\text{C.4})$$

$$= \left( \frac{xf}{z+Z}, \frac{yf}{z+Z}, f \right) \quad (\text{C.5})$$

$$= \left( \frac{r \cos \phi \cos \theta f}{r \sin \phi \cos \theta + Z}, \frac{r \sin \theta f}{r \sin \phi \cos \theta + Z}, f \right) \quad (\text{C.6})$$

$$\dot{p} = \left( \frac{-frZ\omega \cos \phi \sin \theta}{(r \sin \phi \cos \theta + Z)^2}, \frac{fr\omega(Z \cos \theta + r \sin \phi)}{(r \sin \phi \cos \theta + Z)^2}, f \right) \quad (\text{C.7})$$

$$\nabla \times \dot{\boldsymbol{p}} = \left( \begin{array}{c} 0, 0, \frac{\omega \left( \frac{Z^2 \cos \phi (Z \cos \theta + r(1 + \sin^2 \theta) \sin \phi)}{Z \cos \theta + r \sin \phi} + \sec \phi (Z^2 - rZ \cos \theta \sin \phi - 2r^2 \sin^2 \phi) \right)}{Z(Z + r \cos \theta \sin \phi)} \end{array} \right) \quad (\text{C.8})$$

となり、curlの $z$ 成分は一定値にはならない。しかし $\phi = 0$ を代入すると $z$ 成分は $2\omega$ となり、傾いていないときの結果には一致する。また、焦点距離 $f$ には依存しない。

一定値にはならないが、値の変化はどの程度なのかを調べた。 $\phi = 0.2[\text{rad}]$  (約 $10^\circ$ )、 $\omega = 0.4[\text{rad/s}]$ として、奥行き $Z$ と腕の角度 $\theta$ によって、curlの $z$ 成分がどの程度の値をとるのかを計算した結果を図C.2と図C.3に示す。

円盤(渦巻き)のように見えるのがcurlの $z$ 成分の値である。回転中心から30cmの距離までを $-\pi \leq \theta \leq \pi$ の範囲で計算した。つまり、 $Z = 200$ の図C.3(b)ならば、奥行き200cmの距離から見た長さ30cmの回転物体のcurlが、場所 $(r, \theta)$ によって変化する様子を示している。

理想的な傾きのない場合( $\phi = 0$ )は、値が $0.8 (= 2\omega)$ の円盤になる。しかし傾きがある場合は、奥行きが小さい時(図C.2(a))に $z$ 成分の値は大きく歪んだ円盤ようになる。奥行きが大きくなるにしたがって透視投影の影響は小さくなり、 $\theta$ によらず $0.8$ に近づいていく(図C.2(b))。たしかに $Z \rightarrow \infty$ のときの極限を計算すると $\omega \frac{1 + \cos^2 \phi}{\cos \phi}$ となり、正射影の結果に一致する。

いずれの場合も $\theta = \pm \frac{\pi}{2}$ 付近で特異点がある。これは、式(C.8)の分子第一項の分母である $Z \cos \theta + r \sin \phi$ が0になるためである。これを0とおいて以下のように変形すると、

$$Z \cos \theta + r \sin \phi = 0 \quad (\text{C.9})$$

$$Z \frac{x}{\sqrt{x^2 + y^2}} + \sqrt{x^2 + y^2} \sin \phi = 0 \quad (\text{C.10})$$

$$Zx + (x^2 + y^2) \sin \phi = 0 \quad (\text{C.11})$$

$$Ax + x^2 + y^2 = 0, \quad A = \frac{Z}{\sin \phi} \quad (\text{C.12})$$

$$\left(x + \frac{A}{2}\right)^2 + y^2 = \frac{A^2}{4} \quad (\text{C.13})$$

となり、原点を通る円になることがわかる。その軌跡を $Z$ を変化させて描いたのが図C.1である。この図からも分かるように、 $Z$ が大きくなるにしたがって、軌跡の円の半径は大きくなり、特異点が存在する場所は $y$ 軸上に近づいていく。しかしそれ以外の場所では、奥行き $Z$ が腕の長さ $r$ に比べて十分大きければ、一定値に近似してもよいと思われる。

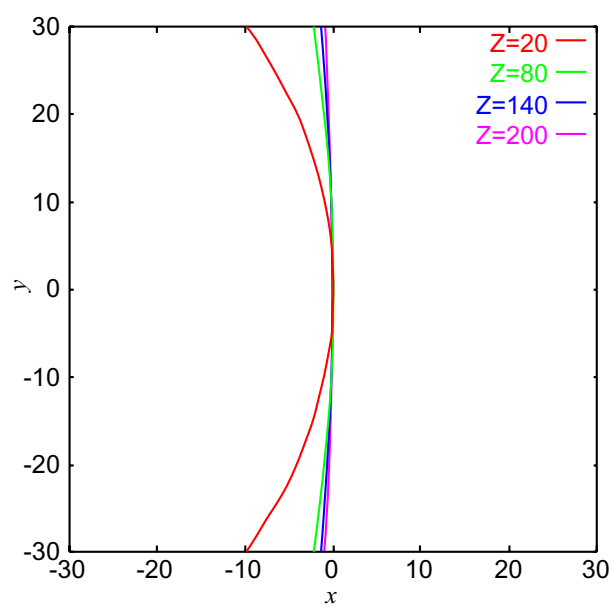


図 C.1: 特異点の軌跡

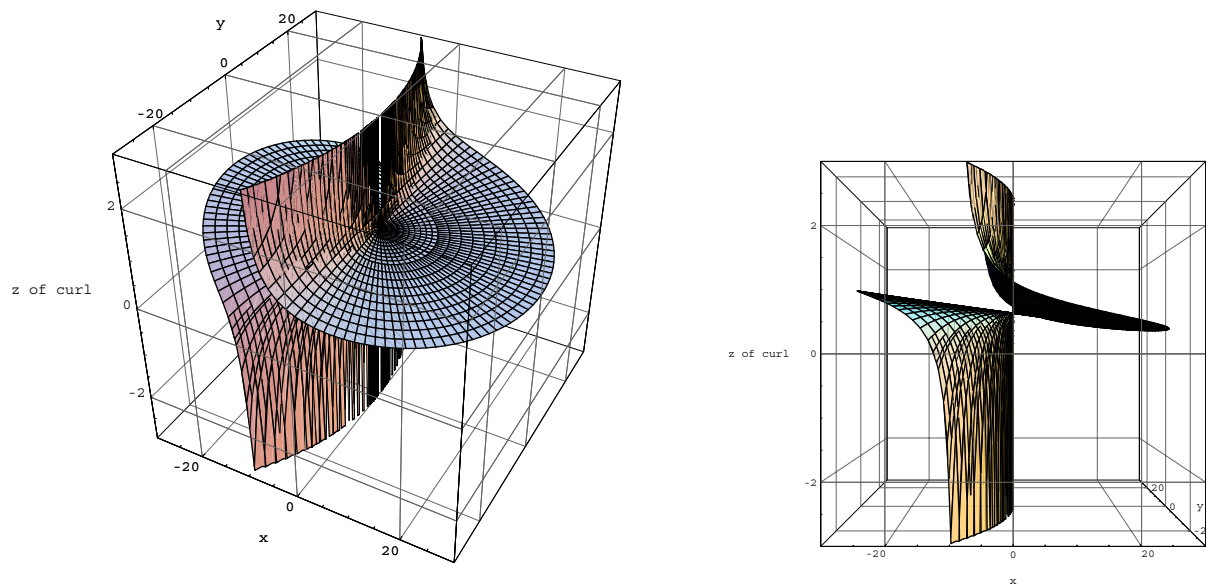
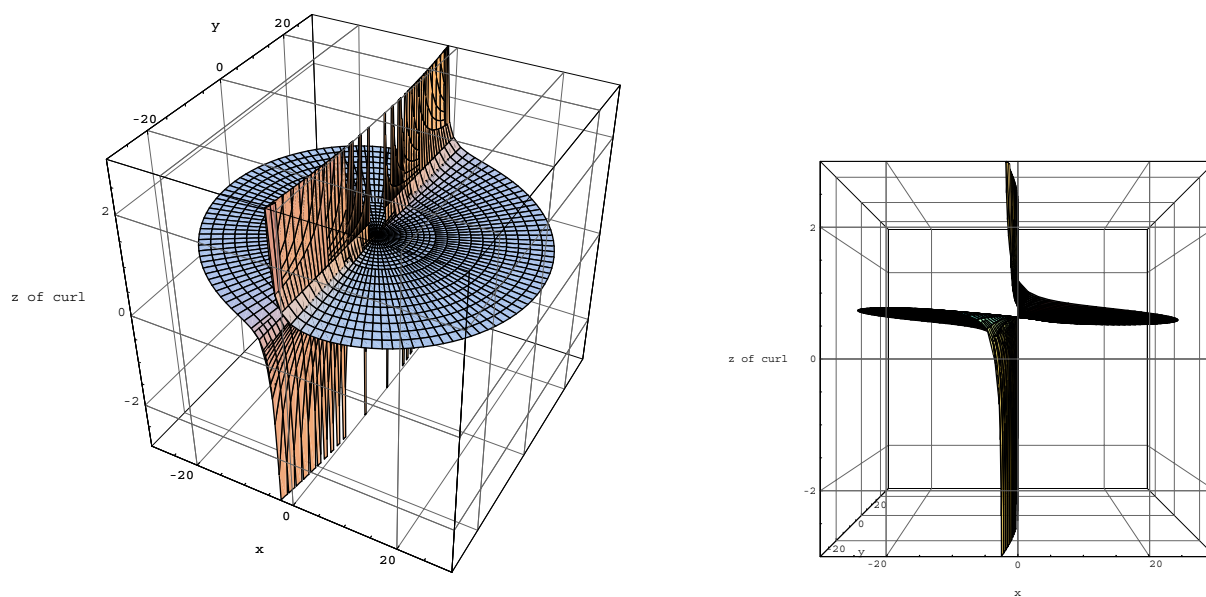
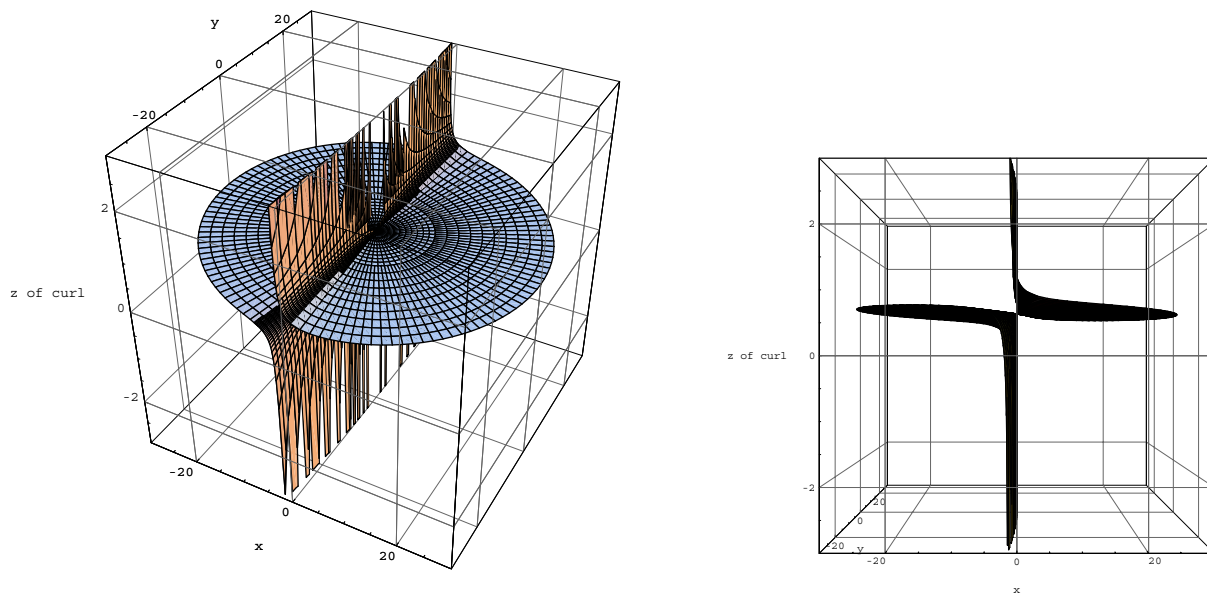
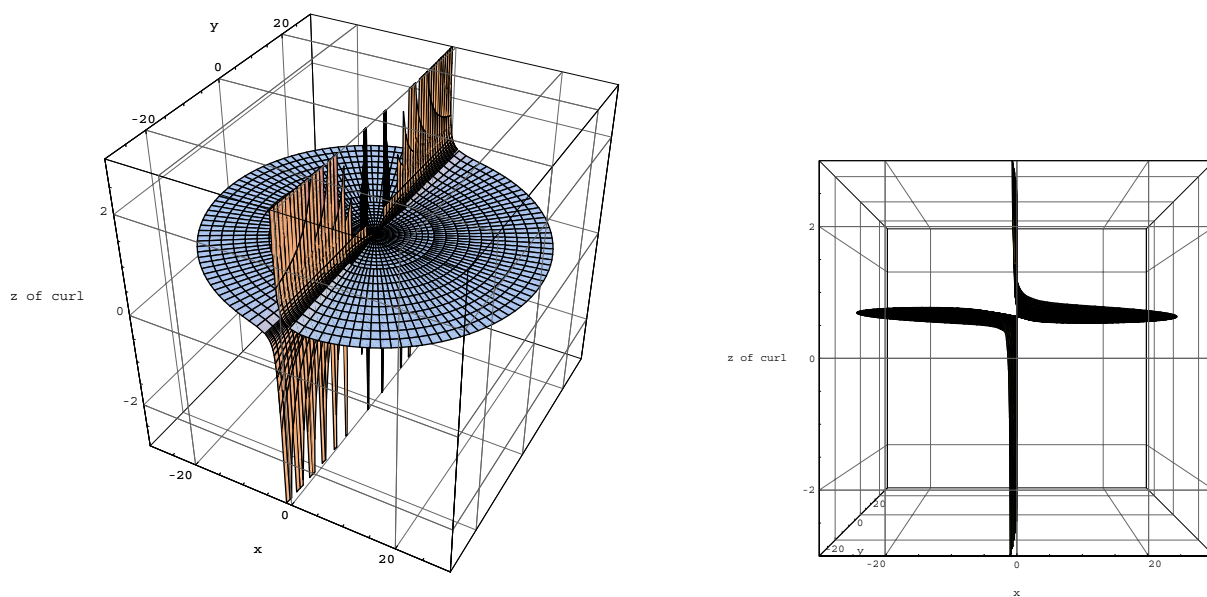
(a)  $Z = 20$ (b)  $Z = 80$ 

図 C.2: 透視投影下で、 $\phi = 0.2$ 、 $\omega = 0.4$  の場合の curl を計算した結果。それぞれ、俯瞰図 (左) と、それを水平方向から見た図 (右)



(a)  $Z = 140$



(b)  $Z = 200$

図 C.3: 透視投影下で、 $\phi = 0.2$ 、 $\omega = 0.4$  の場合の curl を計算した結果。それぞれ、俯瞰図 (左) と、それを水平方向から見た図 (右)



## 付録 D 式 (5.16) の導出

5.2.2節において、運動の回転面が  $\phi \neq 0$  かつ  $\psi \neq 0$  である一般な場合の、運動のフローの curl を式 (5.16) に示した。この式の導出を以下に示す。

まず図 5.1(b) と同様に、回転面を  $y$  軸回りに  $\phi$  だけ回転させ、次に  $x$  軸回りに  $\psi$  だけ回転させる。回転の順序を変えると以下の導出と式も変わるが、同様の手順で導くことができる。

このときの点  $p$  とその速度  $\dot{p}$  は次ようになる。

$$\begin{aligned} \mathbf{p} &= (x, y, z)^T \\ &= (r \cos \phi \cos \theta, r \sin \theta \cos(\frac{\pi}{2} - \psi) + r \cos \theta \sin \phi \sin(\frac{\pi}{2} - \psi), \\ &\quad r \cos \theta \sin \phi \cos(\frac{\pi}{2} - \psi))^T \end{aligned} \quad (\text{D.1})$$

$$= (r \cos \phi \cos \theta, r \sin \theta \sin \psi + r \cos \theta \sin \phi \cos \psi, r \cos \theta \sin \phi \sin \psi)^T \quad (\text{D.2})$$

$$\begin{aligned} \dot{\mathbf{p}} &= (u, v, w)^T \\ u &= -r\omega \cos \phi \sin \theta \\ &= -\omega \cos \phi \frac{y - \frac{x}{\cos \phi} \sin \phi \cos \psi}{\sin \psi} \\ &= -y\omega \cos \phi \csc \psi + x\omega \cot \psi \sin \phi \end{aligned} \quad (\text{D.3})$$

$$\begin{aligned} v &= r\omega(\cos \theta \sin \psi - \cos \psi \sin \phi \sin \theta) \\ &= \omega \left( \frac{x \sin \psi}{\cos \psi} - \cos \psi \sin \phi \frac{y - \frac{x}{\cos \phi} \sin \phi \cos \psi}{\sin \psi} \right) \\ &= -y\omega \cos \psi \sin \phi + x\omega(\sec \phi \sin \psi + \cos \psi \cot \psi \sin \phi \tan \phi) \end{aligned} \quad (\text{D.4})$$

$w$  は省略する。

上式から curl の  $z$  成分は以下のように計算できる。

$$\frac{\partial v}{\partial x} = \omega(\sec \phi \sin \psi + \cos \psi \cot \psi \sin \phi \tan \phi) \quad (\text{D.5})$$

$$\frac{\partial u}{\partial y} = -\omega \cos \phi \csc \psi \quad (\text{D.6})$$

$$\begin{aligned} \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} &= -\omega \cos \phi \csc \psi - \omega(\sec \phi \sin \psi + \cos \psi \cot \psi \sin \phi \tan \phi) \\ &= \omega \frac{\cos^2 \phi + \sin^2 \psi + \sin^2 \phi \cos^2 \psi}{\cos \phi \sin \psi} \end{aligned} \quad (\text{D.7})$$

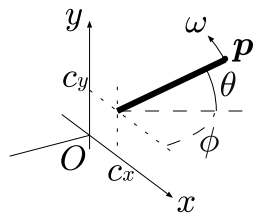
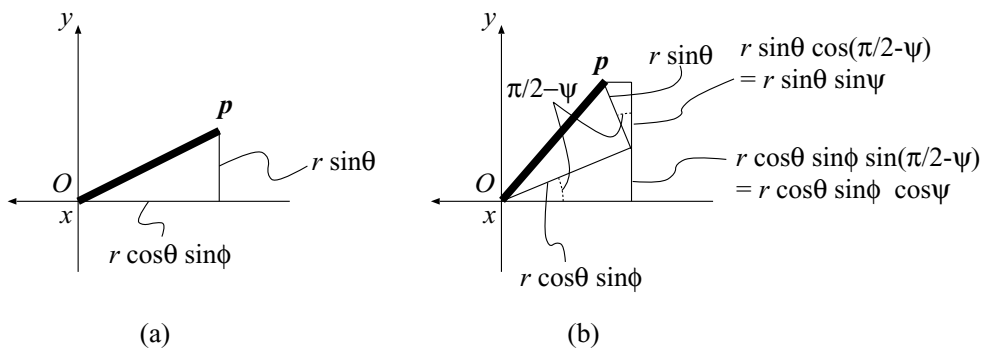


図 5.1(b) (再掲)

図 D.1: 図 5.1(b) を  $x$  軸方向から見た座標系

(a)  $\phi \neq 0, \psi = 0$  の場合 (図 5.1(b)) (b)  $\phi \neq 0, \psi \neq 0$  の場合



## 参考文献

- [Arita95] Daisaku Arita, Naoyuki Tsuruta, Rin-ichiro Taniguchi, and Makoto Amamiya. An Object Recognition System Based on a Model Generated from Image Examples. In *Proc. of ACCV'95*, Vol. 3, pp. 161–165, 1995.
- [Ballard82] Dana H. Ballard and Christopher M. Brown. *Computer Vision*. Prentice-Hall, Inc., 1982.
- [Bischoff00] R. Bischoff. Towards the development of ‘plug-and-play’ personal robots. In *Proc. of 1st IEEE-RAS Intl. Conf. on Humanoid Robots*, 2000.
- [Brice70] Claude R. Brice and Claude L. Fennema. Scene analysis using regions. *Artificial Intelligence*, Vol. 1, pp. 205–226, 1970.
- [Castleman96] Kenneth R. Castleman. *Digital Image Processing*. Prentice-Hall, Inc., 1996.
- [Craig91] John J. Craig. *ロボティクス — 機構・力学・制御 —*. 共立出版, 1991.
- [Dempster77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM Algorithm. In *J. Roy. Statist. Soc. B*, Vol. 39, pp. 1–38, 1977.
- [Ford98] Adrian Ford and Alan Roberts. Colour space conversions. <http://www.inforamp.net/~poynton/PDFs/coloureq.pdf>, 8 1998.
- [Forsyth97] David Forsyth, Jitendra Malik, and Robert Wilensky. Searching for digital pictures. *Scientific American*, pp. 88–93, Jun 1997. (邦訳)「絵の特徴から選び出す画像検索法」, 谷千田正彦 岩井儀雄 訳, 日経サイエンス 1997年9月号, pp.86–93.
- [Gavrila96] D. M. Gavrila and L. S. Davis. 3-D model-based tracking of humans in action: a multi-view approach. In *Proc. of ICPR'96*, pp. 73–80, 1996.
- [Gavrila99] D. M. Gavrila. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, Vol. 73, No. 1, pp. 82–98, 1999.
- [Goncalves98] L. Goncalves, E. D. Bernardo, and P. Perona. Reach out and touch space (Motion learning). In *Proc. of FG'98*, pp. 234–239, 1998.
- [Gonzalez77] Rafael C. Gonzalez and Paul Wintz. *Digital Image Processing*. Addison-Wesley, 1977.

- [Gonzalez92] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley, 1992.
- [Guo94] Yan GUO, Gang XU, and Saburo TSUJI. Understanding human motion patterns. In *Proc. of ICPR'94*, pp. 325–329, 1994.
- [Gurari82] Eitan M. Gurari and Harry Wechsler. On the difficulties involved in the segmentation of pictures. *IEEE Trans. on PAMI*, Vol. 4, No. 3, pp. 304–306, 1982.
- [Haritaoglu98] Ismail Haritaoglu, David Harwood, and Larry S. Davis.  $W^4$ : Who? When? Where? What? A real time system for detecting and tracking people. In *Proc. of FG'98*, pp. 222–227, 1998.
- [Horn93] B. K. P. Horn. *ロボットビジョン*. 朝倉書店, 1993.
- [Jain95] Ramesh Jain, Rangachar Kasturi, and Brian G. Schunck. *MACHINE VISION*. McGraw-Hill, New York, 1995.
- [Ju96] S. X. Ju, M. J. Black, and Y. Yacoob. Cardboard people: A parameterized model of articulated image motion. In *Proc. of FG'96*, pp. 38–44, 1996.
- [Kanade80] Takeo Kanade. Survey region segmentation: Signal vs semantics. *Computer Graphics and Image Processing*, Vol. 13, pp. 279–297, 1980.
- [Kanungo94] Tapas Kanungo, Byron Dom, Wayne Niblack, and David Steele. A fast algorithm for MDL-Based multi-band image segmentation. In *Proc. of IEEE CVPR'94*, pp. 609–616, 1994.
- [Kass98] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Intl. J. Comp. Vis.*, Vol. 1, No. 3, pp. 321–331, 1998.
- [Khan00] Sohaib Khan and Mubarak Shah. Tracking people in presense of occlusion. In *Proc. of ACCV2000*, pp. 1132–1137, 2000.
- [Koster97] A. S. E. Koster, K. L. Vincken, and M. A. Viergever. Heuristic linking models in multi-scale image segmentation. *Computer Vison and Image Understanding*, Vol. 65, No. 3, pp. 382–402, 1997.
- [Minoh98] M. Minoh. 3D Model centered framework for CV and VR. In *Proc. of ACCV'98*, pp. 332–339, 1998.
- [Morris98] D. D. Morris and J. M. Rehg. Singularity analysis for articulated object tracking. In *Proc. of CVPR'98*, pp. 289–296, 1998.
- [Nev86] R. Nevatia. Image Segmentation. In Tzay Y. Young and King-Sun Fu, editors, *Handbook of Pattern Recognition and image Processing*, chapter 9. Academic Press, 1986.

- [Nishimura98] T. Nishimura, T. Mukai, and R. Oka. Non-monotonic continuous dynamic programming for spotting recognition of hesitated gestures from time-varying images. In *Proc. of ACCV'98*, pp. 734–739, 1998.
- [Niyogi94] Sourabh A. Niyogi and Edward H. Adelson. Analyzing and recognizing walking figures in xyt. In *Proc. of CVPR'94*, pp. 469–474, 1994.
- [Ohta91] N. Ohta. Image movement detection with reliability indices. *IEICE Transactions*, Vol. E74, No. 10, pp. 3379–3388, 1991.
- [Ohta96] N. Ohta. <http://www.ail.cs.gunma-u.ac.jp/Labo/Program/Flow-R00.tar.gz>, 1996.
- [Ohya94] J. Ohya and F. Kishino. Human posture estimation from multiple images using genetic algorithm. In *Proc. of ICPR'94*, pp. 750–753, 1994.
- [Pal93] Nikhil R. Pal and Sankar K. Pal. A review on image segmentation techniques. *Pattern Recognition*, Vol. 26, No. 9, pp. 1277–1294, 1993.
- [Poynton97] Charles Poynton. Frequently asked questions about color (Color FAQ). <http://www.inforamp.net/~poynton/ColorFAQ.html>, 1997.
- [Pratt91] William K. Pratt. *Digital Image Processing*. John Wiley & Sons, Inc., 2nd edition, 1991.
- [Rowley95] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. Human face detection in visual scenes, July 1995. CMU-CS-95-158.
- [Sakaue99] Katsuhiko Sakaue, Akira Amano, and Naokazu Yokoyama. Optimization approaches in computer vision and image processing. *IEICE Trans. Inf. & Syst.*, Vol. E82-D, pp. 534–547, 1999.
- [Salembier00] Philippe Salembier and Luis Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. on Image Proc.*, Vol. 9, pp. 561–576, 2000.
- [Shirai87] Yoshiaki Shirai. *Three-Dimensional Computer Vision*. Springer-Verlag, 1987.
- [Skarbek94] Wladyslaw Skarbek and Andreas Koschan. Colour image segmentation — A survey — . Technical Report 94-32, Institute of Technical Informatics, Technical University of Berlin, 1994.
- [Takeuchi94] Yoshinori Takeuchi, Zeng-Fu Wang, Noboru Ohnishi, and Noboru Sugie. Real time visual tracking system mimicking saccadic movements. In *Proc. of ACCV'95*, pp. I131–I135, December 1994.
- [Tamaki99] Toru Tamaki, Tsuyoshi Yamamura, and Noboru Ohnishi. Image segmentation and object extraction based on geometric features of regions. In *SPIE Vol. 3653 Visual Communications and Image Processing '99*, Vol. 2, pp. 937–945, 1999.

- [Tamaki00a] Toru Tamaki, Tsuyoshi Yamamura, and Noboru Ohnishi. An automatic camera calibration method with image registration technique. In *Proc. of SCI2000 ; The 4th World Multiconference on Systemics, Cybernetics and Informatics*, Vol. 5, pp. 317–322. IIS; International Institute of Informatics and Systemics, 2000.
- [Tamaki00b] Toru Tamaki, Tsuyoshi Yamamura, and Noboru Ohnishi. Extraction of human limb regions and parameter estimation based on curl of optical flow. In *Proc. of ACCV2000 ; the Fourth Asian Conference on Computer Vision*, Vol. 2, pp. 1008–1013, 2000.
- [Thrun00] S. Thrun, M. Beets, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *Intl. J. Robotics Res.*, Vol. 19, pp. 972–999, 2000.
- [Treisman80] A. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, Vol. 12, pp. 97–136, 1980.
- [Vernon91] David Vernon. *Machine Vision*. Prentice Hall Intl., 1991.
- [Vincken97] Koen L. Vincken, André S. E. Koster, and Max A. Viergever. Probabilistic multiscale image segmentation. *IEEE Trans. on PAMI*, Vol. 19, No. 2, pp. 109–120, 1997.
- [Waldherr00] Stefan Waldherr, Roseli Romero, and Sebastian Thrun. A gesture based interface for human-robot interaction. *Autonomous Robots*, Vol. 9, No. 2, pp. 151–173, 2000.
- [Yamato92] J. Yamato, J. Ohta, and K. Ishii. Recognizing human action in time-sequential images using hidden markov model. In *Proc. of CVPR'92*, pp. 379–385, 1992.
- [Zhu96] Song Chun Zhu and Alan Yuille. Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Trans. on PAMI*, Vol. 18, No. 9, pp. 884–900, 1996.
- [安居院 92] 安居院猛, 長尾智晴. 画像の処理と認識. 昭晃堂, 1992.
- [天野 98] 天野晃. コンピュータビジョンにおける最適化手法. 松山隆司, 久野義徳, 井宮淳 (編), コンピュータビジョン: 技術評論と将来展望, 第 11 章. 新技術コミュニケーションズ, 1998.
- [淡 96] 淡誠一郎, 鄭小江, 北橋忠宏. 対話型対象物抽出アルゴリズムの高速化. 信学論 D-II, Vol. J79-D-II, No. 11, pp. 1984–1987, 1996.
- [石井 93] 石井浩史, 望月研二, 岸野文郎. 人物像合成のためのステレオ画像からの動作認識法. 信学論 D-II, Vol. J76-D-II, No. 8, pp. 1805–1812, 1993.
- [石井 98] 石井健一朗, 上田修功, 前田英作, 村瀬洋. わかりやすいパターン認識. オーム社, 1998.

- [石橋 91] 石橋聡, 小林幸雄, 岸野文郎. 背景参照画像と両眼視差を用いた任意背景中の人物像抽出. テレビジョン学会誌, Vol. 45, No. 10, pp. 1270–1276, 1991.
- [乾 90] 乾敏郎. 視覚情報処理の基礎. サイエンス社, 1990.
- [乾 93] 乾敏郎 (編). 認知と学習. 丸善, 1993.
- [乾 95a] 乾敏郎 (編). 知覚と運動. 認知心理学 1. 東京大学出版会, 1995.
- [乾 95b] 乾敏郎 (編). 認知心理学, 第1巻. 東京大学出版会, 1995. 運動と知覚.
- [岩田 97] 岩田洋夫. バーチャルリアリティへの応用. 計測と制御, Vol. 36, No. 9, pp. 639–642, 1997.
- [上田 99] 上田修功, 中野良平. 混合モデルのための併合分割操作付き EM アルゴリズム. 信学論 D-II, Vol. J82-D-II, No. 5, pp. 930–940, 1999.
- [栄藤 97] 栄藤稔. 動画像符号化におけるコンピュータビジョン. 信学論 D-II, Vol. J80-D-II, No. 2, pp. 399–406, 1997.
- [栄藤 98] 栄藤稔. 画像圧縮: 画像内容の理解による符号化効率の向上. 日本ロボット学会誌, Vol. 16, No. 6, pp. 754–758, 1998.
- [太田 98] 大田佳人, 山際貴志, 山本正信. キーフレーム拘束を利用した単眼動画像からの人間動作の追跡. 信学論 D-II, Vol. J81-D-II, No. 9, pp. 2008–2018, 1998.
- [大津 80] 大津展之. 判別および最小2乗基準に基づく自動しきい値選定法. 信学論, Vol. J63-D, No. 4, pp. 349–356, 1980.
- [大山 00] 大山正. 視覚心理学への招待. サイエンス社, 2000.
- [大谷 94] 大谷淳, 岸野文郎. 遺伝的アルゴリズムを用いたマルチ画像からの人物の姿勢検出の検討—上半身の間接角度検出—. 信学技報, Vol. PRU93-12, pp. 47–54, 1994.
- [大羽 99] 大羽成征, 石井信. 正規混合分布の推定によるクラスタリングの種々手法の比較. 信学技報, Vol. NC99-12, pp. 109–116, 1999.
- [岡田 99] 岡田隆三, 白井良明, 三浦純, 久野義徳. オプティカルフローと距離情報の統合による3次元運動する人間の追跡. 信学論 D-II, Vol. J82-D-II, No. 8, pp. 1252–1261, 1999.
- [岡 99] 岡隆一, 西村拓一, 向井理郎. しぐさで伝える. 電子情報通信学会誌, Vol. 82, No. 4, pp. 332–339, 1999.
- [緒方 98] 緒方敏博, 園田頼信. 複数対象物の運動軌跡の抽出. 画像の認識・理解シンポジウム (MIRU'98) 論文集, 第I巻, pp. I-431–I-436, 1998.
- [奥宮 96] 奥宮啓司, 金子正秀, 原島博. 動画像の抽象化索引映像の作成とその応用. テレビジョン学会誌, Vol. 50, No. 10, pp. 1558–1566, 1996.

- [甲藤 98] 甲藤二郎. MPEG4 最終版解説とデモンストレーション : MPEG-4 SYSTEM の仕様. 信学技報, Vol. PRUM97-178, pp. 45-52, 1998.
- [亀田 96] 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, 1996.
- [川人 94] 川人光男, 行場次朗, 藤田一郎, 乾敏郎, 力丸裕. 視覚と聴覚. 岩波講座 認知科学 3. 岩波書店, 1994.
- [川本 00] 川本哲也. LED リングと特殊反射材を用いたバーチャルスタジオ開発. 2000 年度電気関係学会東海支部連合大会講演論文集, p. 263, 2000.
- [感覚 94] 大山正, 今井省吾, 和気典二 (編). 新編 感覚・知覚心理学ハンドブック. 誠信書房, 1994.
- [画像 97] 画像処理標準テキストブック編集委員会 (編). イメージプロセッシング <画像処理標準テキストブック>. 財団法人画像情報教育振興会, 1997.
- [木本 91] 木本伊彦, 梶尾昭彦, 安田靖彦. スティックモデルに基づく単眼動画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.
- [栗田 98] 栗田多喜夫, 赤穂昭太郎. 画像理解における統計的手法. 松山隆司, 久野義徳, 井宮淳 (編), コンピュータビジョン : 技術評論と将来展望, 第 13 章, pp. 195-206. 新技術コミュニケーションズ, 1998.
- [江 91] 江浩, 鈴木秀智, 鳥脇純一郎. 領域情報とエッジ情報を併用したセグメンテーションの一手法. 信学論 D-II, Vol. J74-D-II, pp. 1651-1660, 1991.
- [小柳 91] 小柳義夫. 最小 2 乗法と特異値分解. 現代数理科学辞典, 第 XVI[6]1-2 章, pp. 913-917. 大阪書籍, 1991.
- [小西 78] 小西栄一, 深見哲造, 遠藤静男. 線形代数・ベクトル解析. 培風館, 1978.
- [齊藤 00] 齊藤隆弘. 映像符号化の分類・要素技術. テレビジョン学会誌, Vol. 54, No. 5, pp. 629-635, 2000.
- [境田 98] 境田慎一, 鹿喰善明, 田中豊, 湯山一郎. K 平均アルゴリズムの初期値依存性を利用した統合処理による画像の領域分割. 信学論 D-II, Vol. J81-D-II, No. 2, pp. 311-322, 1998.
- [坂上 91] 坂上勝彦, 山本和彦. 動的な網のモデル Active Net とその領域抽出への応用. テレビジョン学会誌, Vol. 45, No. 10, pp. 1155-1163, 1991.
- [坂口 90] 坂口嘉之, 美濃導彦, 池田克夫. SNAKE パラメータの設定についての検討. 信学技報, Vol. PRU90-21, pp. 43-49, 1990.
- [坂口 91] 坂口俊文, 大山公一. 面積項をもつスネーク. 信学春季全大, Vol. D-555, , 1991.
- [佐藤 97] 佐藤明知, 川田聡, 大崎喜彦, 山本正信. 多視点動画からの人間動作の追跡と再構成. 信学論 D-II, Vol. J80-D-II, No. 6, pp. 1582-1589, 1997.

- [佐藤 98] 佐藤宏介. 画像応用セキュリティシステム. システム/制御/情報, Vol. 42, No. 3, pp. 150–155, 1998.
- [鹿喰 97] 鹿喰善明. 領域サポート DCT を用いた任意形状テクスチャ符合化. 信学論 D-II, Vol. J80-D-II, No. 7, pp. 1810–1817, 1997.
- [色彩研 93a] (財)日本色彩研究所(編). 色の見えるしくみ. 色彩ワンポイント 1. 日本規格協会, 1993.
- [色彩研 93b] (財)日本色彩研究所(編). 色彩管理の基礎. 色彩ワンポイント 2. 日本規格協会, 1993.
- [志治 00] 志治亜矢子, 浜田望. Watershed アルゴリズムと輪郭情報を用いたカラー画像の領域分割. 信学論 D-II, Vol. J83-D-II, No. 2, pp. 593–600, 2000.
- [篠崎 91] 篠崎寿夫, 松森徳衛, 吉田正廣. 現代工学のための変分学入門. 現代工学社, 1991.
- [下條 95] 下條信輔. 視覚の冒険. 産業図書, 1995.
- [末永 95] 末永康仁. 人物像を読む — ヒューマンインターフェースのためのコンピュータビジョン — . 電子情報通信学会誌, Vol. 78, No. 8, pp. 800–804, 1995.
- [高橋 94] 高橋勝彦, 関進, 小島浩, 岡隆一. ジェスチャー動画像のスポットティング認識. 信学論 D-II, Vol. J77-D-II, No. 8, pp. 1552–1561, 1994.
- [高橋 99] 高橋圭子, 阿部圭一. ISODATA クラスタリング法を用いたカラー画像の領域分割. 信学論 D-II, Vol. J82-D-II, No. 4, pp. 751–762, 1999.
- [田中 98a] 田中昭二. 見ているところを見る. *ATR Journal*, Vol. 33, pp. 2–3, Autumn 1998.
- [田中 98b] 田中昭二, 井上正之, 井上誠喜, 中津良平. 誘目性に寄与する物理的特徴量を基にした画像注目領域の抽出. 映像情報メディア学会誌, Vol. 52, pp. 881–890, 1998.
- [谷口 96] 谷口慶治. 画像処理工学 — 基礎編 — . 共立出版, 1996.
- [玉木 97] 玉木徹, 山村毅, 大西昇. 画像からの人物領域の抽出. 平成 9 年度 電気関係学会東海支部連合大会 講演論文集, p. 405, 1997.
- [玉木 98a] 玉木徹, 山村毅, 大西昇. 画像中の人物像の検出について. 第 11 回若手セミナー「インテリジェント計測システム」若手研究者研究発表会講演資料, pp. 41–42. 電気学会東海支部, 1998.
- [玉木 98b] 玉木徹, 山村毅, 大西昇. 画像中の物体領域の抽出. 信学技報, Vol. PRMU97-216, pp. 63–70, 1998.
- [玉木 98c] 玉木徹, 山村毅, 大西昇. 人物モデル当てはめによる人物の位置決め. 情報処理学会第 57 回 (平成 10 年後期) 全国大会講演論文集, 第 2 巻, pp. 142–143, 1998.
- [玉木 99a] 玉木徹, 山村毅, 大西昇. オプティカルフローを用いた腕の領域抽出とパラメータ推定. 平成 11 年度 電気関係学会東海支部連合大会 講演論文集, p. 227, 1999.

- [玉木 99b] 玉木徹, 山村毅, 大西昇. フローのcurlに基づいた回転する人物の腕の抽出とパラメータ推定. 信学技報, Vol. PRMU99-35, pp. 41–48, 1999.
- [玉木 99c] 玉木徹, 山村毅, 大西昇. 画像系列からの人物領域の抽出. 電気学会論文誌 C分冊, Vol. 119-C, pp. 37–43, 1999.
- [玉木 00a] 玉木徹, 山村毅, 大西昇. オプティカルフローを用いた複雑背景下における人物の腕領域の抽出とパラメータ推定. 電気学会論文誌 C分冊, Vol. 120-C, pp. 1801–1808, 2000.
- [玉木 00b] 玉木徹, 山村毅, 大西昇. 画像レジストレーションに基づいた画像歪み補正. 2000年電子情報通信学会 情報・システムソサイエティ大会講演論文集, p. 235, 2000.
- [玉木 00c] 玉木徹, 山村毅, 大西昇. 非線形最適化による画像の歪曲収差補正. 2000年映像情報メディア学会冬期大会講演予稿集, p. 58, 2000.
- [玉木 00d] 玉木徹, 山村毅, 大西昇. 非線形最適化を用いた人物の腕領域抽出. 平成12年度電気関係学会東海支部連合大会 講演論文集, p. 265, 2000.
- [戸川 97] 戸川達男. 運動計測技術の展望. 計測と制御, Vol. 36, No. 9, pp. 605–608, 1997.
- [富永 98] 富永昌治. コンピュータビジョンにおけるカラー情報の表現と解析. 松山隆司, 久野義徳, 井宮淳 (編), コンピュータビジョン: 技術評論と将来展望, 第5章. 新技術コミュニケーションズ, 1998.
- [鳥脇 88] 鳥脇純一郎. 画像理解のためのデジタル画像処理 [II]. 昭晃堂, 1988.
- [中嶋 97] 中嶋正之. デジタル映像処理の基礎と応用 [I] — デジタル映像処理入門 — . 電子情報通信学会誌, Vol. 80, No. 12, pp. 1288–1293, 1997.
- [中嶋 98a] 中嶋正之. デジタル映像処理の基礎と応用 [II] — 二次元デジタル映像処理入門 — . 電子情報通信学会誌, Vol. 81, No. 1, pp. 68–75, 1998.
- [中嶋 98b] 中嶋正之. デジタル映像処理の基礎と応用 [IV] — コンピュータアニメーション入門 — . 電子情報通信学会誌, Vol. 81, No. 3, pp. 282–288, 1998.
- [中谷 93] 中谷広正, 大崎賢治, 阿部圭一. 複数の領域分割に基づく対象物境界線検出. 信学論 D-II, Vol. J76-D-II, No. 4, pp. 914–916, 1993.
- [名取 90] 名取亮. 数値解析とその応用. コロナ社, 1990.
- [波部 98] 波部斉, 大矢崇, 松山隆司. 動的環境における頑健な背景差分の実現法. 画像の認識・理解シンポジウム (MIRU'98) 論文集, 第I巻, pp. I-467–I-472, 1998.
- [人間 98] 人間共存型ロボット研究専門委員会. 人間共存型ロボットシステムにおける技術課題. 日本ロボット学会誌, Vol. 16, No. 3, pp. 288–294, 1998.
- [浜本 96] 浜本隆之, 半谷精一郎, 宮内一洋. 照明の影響を受けたカラー画像の領域分割. 信学論 D-II, Vol. J79-D-II, pp. 351–357, 1996.



- [原 95] 原義明, 大西昇, 杉江昇. カラー画像からの物体抽出について. 信学技報, Vol. IE94-150, PRU94-150, pp. 35-42, 1995.
- [朴 99] 朴基守, 増田基, 伊東敏夫, 金田悠紀夫. アフィン変換導入によるオプティカルフロー推定の一手法. 電気学会論文誌 C 分冊, Vol. 119-C, No. 12, pp. 1468-1475, 1999.
- [福井 97] 福井和広. 領域間の分離度に基づく物体輪郭抽出. 信学論 D-II, Vol. J80-D-II, No. 6, pp. 1406-1414, 1997.
- [堀田 98] 堀田裕弘, 宮本剛, 村井忠邦. 遺伝的アルゴリズムを用いたカラー画像の領域分割. 信学論 D-II, Vol. J81-D-II, No. 11, pp. 2685-2687, 1998.
- [増田 95] 増田功. セキュリティ監視システム. テレビジョン学会誌, Vol. 49, No. 3, pp. 261-267, 1995.
- [増田 97] 増田功. セキュリティのための人物像の自動認識. テレビジョン学会誌, Vol. 51, No. 8, pp. 1154-1158, 1997.
- [増田 00] 増田良介. センシング技術. 日本ロボット学会誌, Vol. 18, No. 7, pp. 909-912, 2000.
- [松井 99] 松井俊浩. オフィスロボット Jijo-2. 工業技術, Vol. 40, No. 4, 1999.
- [松下 98] 松下賢也, 竹中昭武, 加藤邦人, 山本和彦. 非整備環境におけるハト追跡システムの構築について. 電気学会論文誌 C 分冊, Vol. 118-C, No. 3, pp. 391-397, 1998.
- [三池 93] 三池秀敏, 古賀和利 (編). パソコンによる動画像処理. 北森出版, 1993.
- [六日市 99] 六日市正弘, 浜田望. エッジにおける速度拘束式とエッジ連結を考慮したオプティカルフロー推定. 信学論 D-II, Vol. J82-D-II, No. 3, pp. 407-414, 1999.
- [谷内田 90a] 谷内田正彦 (編). コンピュータビジョン. 丸善, 1990.
- [谷内田 90b] 谷内田正彦. ロボットビジョン. 昭晃堂, 1990.
- [山口 99] 山口順一. セキュリティ画像処理の現状と課題. 画像ラボ, Vol. 10, No. 3, pp. 37-40, 1999.
- [大和 94] 大和敦司, 倉掛正治, 伴野明, 石井健一郎. カテゴリー別 VQ を用いた HMM による動作認識法. 信学論 D-II, Vol. J77-D-II, No. 7, pp. 1333-1318, 1994.
- [山本 96] 山本正信, 川田聡, 近藤拓也, 川越和忠. ロボットモデルに基づく人物動作の 3 次元動画像追跡. 信学論 D-II, Vol. J79-D-II, No. 1, pp. 71-83, 1996.
- [吉野 99] 吉野峰生. 犯罪捜査における顔画像の個人識別. 信学技報, Vol. PRMU99-28, pp. 67-74, 1999.
- [渡邊 98] 渡邊敏明. MPEG4 最終版解説とデモンストレーション — MPEG-4 VIDEO の仕様 — . 信学技報, Vol. PRUM97-177, pp. 37-44, 1998.

- [わらび座96] (株)わらび座. 民族芸能の3次元デジタル舞踊符に関するCD-ROM制作. 通商産業省 マルチメディアコンテンツ支援事業成果集1, 1996. 財団法人マルチメディアコンテンツ振興協会 平成8年度補正予算マルチメディアコンテンツ制作支援事業 素材開発 34.

---

2001年1月23日 初版第1刷

著者 : 玉木 徹

tamaki@ohnishi.nuie.nagoya-u.ac.jp

印刷 : 〒464-8603 名古屋市千種区不老町

名古屋大学大学院工学研究科情報工学専攻 第七講座 大西研究室

Tel : 052-789-5149/5150/3811

Fax : 052-789-3814

製本 : 名古屋大学生協印刷部

---