

Title	場景變動を考慮した移動物体の追跡に関する研究
Author(s)	副島, 義貴
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1163
Rights	
Description	Supervisor:阿部 亨, 情報科学研究科, 修士

修士論文

場景変動を考慮した 移動物体の追跡に関する研究

指導教官 阿部亨 助教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

副島義貴

1998年2月15日

要旨

動的輪郭モデル (Snakes) を用いた移動物体の追跡に関する手法が幾つか提案されているが、従来提案されている手法では、多くの場合背景が変化しない、また視点の移動が無いといった撮影環境が制約された状態で追跡を行っている。しかしながら実際に動画像を扱う場合に必ずしもその範疇に収まらない動画像を対象にした追跡を行う場合がでてくる。そのような場合、フレーム間差分、optical flow、背景差分などに依存した画像エネルギーを用いる従来の手法では、対処するのが非常に困難である。本研究では、より一般的な状況（視点移動・背景変化等）での移動物体の追跡を実現するために、差分情報に依存しない画像領域の色を基にした画像エネルギーを Snakes に導入する。それにより目的の色領域の追跡を行う。それと同時に複数物体の追跡をより正確に行う為に、Snakes の輪郭線の分離・融合機能と時系列画像全体を通して Snakes の状態を管理するために履歴の概念を導入する。

目次

1	序論	1
1.1	研究の背景	1
1.2	研究の目的	2
1.3	本論文の構成	3
2	動的輪郭モデルによる移動物体の抽出・追跡法	5
2.1	はじめに	5
2.2	領域分割を用いた手法	5
2.3	輪郭抽出を用いた手法	6
2.3.1	エッジ抽出法	6
2.3.2	動的輪郭モデル	6
2.4	従来の動的輪郭モデルを用いた手法の問題点	13
2.5	おわりに	14
3	動的輪郭モデル Snakes	16
3.1	Snakes の概要	16
3.2	拡張された Snakes	18
3.2.1	面積項を導入した Snakes	18
3.2.2	分離・融合機能を導入した Snakes(C-Snakes)	20
3.3	Snakes の移動物体追跡への応用	26
3.3.1	移動物体の追跡処理	26
3.4	場景変動を考慮した移動物体の抽出・追跡法	29
3.4.1	場景変動への対処	29

3.4.2	新しい画像エネルギーを導入した Snakes	31
3.5	まとめ	35
4	 场景変動を考慮した移動物体追跡の手法	37
4.1	提案手法の流れ	37
4.2	移動物体の追跡処理	38
4.2.1	輪郭抽出・追跡	38
4.2.2	追跡領域の設定	41
4.2.3	閾値 の取扱い	41
4.3	時系列画像全体の対象物体の軌跡管理	41
4.3.1	従来の軌跡管理の問題点	42
4.3.2	対象物体の軌跡管理のしくみ	42
5	 評価実験	45
5.1	実験内容の説明	45
5.1.1	実験内容	45
5.1.2	実験に用いる画像	46
5.2	色情報によって抽出した結果	46
5.3	指定領域の設定と抽出結果との関係	46
5.4	軌跡管理を導入の有効性	48
5.5	场景変動時 (カメラが動くような画像) における追跡結果	48
6	 結論	56
	謝辞	58

第 1 章

序論

1.1 研究の背景

動画像を対象にした画像処理において、移動物体の抽出・追跡は基本的かつ非常に困難な問題の一つである。汎用的な移動物体の抽出・追跡の技術の確立は、認識のみならず画像符号化など画像処理全般にわたって重要な課題であり、またその技術は多くの分野に応用可能だと考えられる。

画像情報から移動対象の抽出・追跡の手法としていくつか提案されているが、主に次の2つの手法に分けられる。一つは領域分割による手法、もう一つは輪郭線抽出による手法である。

領域分割による手法では、画素の特徴量に基づく分類(色、輝度、位置などのクラスタリング)により領域を抽出する。領域分割が正確に行える場合にはこの手法は有効であるが、分割した各領域が必ずしも追跡したい対象とは対応しておらず、対応させるためには複雑な後処理が必要であるという点で問題がある。

一方、輪郭線抽出を用いた手法の一つとして、エッジ抽出を行う手法がある。これは検出されたエッジ点を局所的に接続して目的の輪郭線を得るものである。しかしながら実画像のようにノイズが多い状況下では、このような局所的なエッジ点の接続から正確な閉領域を得ることは非常に困難であり、なにかしらのヒューリスティックな情報に基づく接続処理を行わないと不可能である。

このような問題に対処するものとして、M.Kass ら [1] によって動的輪郭モデル (Snakes) が提案されている。Snakes は初期輪郭として常に閉じた領域が与えられるため、エッジ

抽出では困難だったノイズの入った画像中でも閉領域を抽出することが可能である。特に連続的な画像においては、前のフレームでの処理結果を次のフレームでの処理の初期値として、用いることにより移動物体の追跡を行うことが可能である。このように Snakes は移動物体の抽出・追跡を同時に行うことが可能で、特に動画像のような時間的に連続する画像を処理する場合に有効である。

また剛体で構成されるような物体を抽出・追跡する場合、輪郭線の形状に制約を設け動きをモデル化する手法がよく用いられる。しかし複雑な運動をするような物体では、そのような手法を用いることは困難である。複雑な運動をする物体の抽出・追跡に対しても、輪郭線の形状に制約を受けない Snakes は有効な手法である。

1.2 研究の目的

動画像処理に有効な Snakes を用いた移動物体の抽出・追跡を行う手法がいくつか提案されているが、実際に動画像から移動物体の抽出・追跡を行う上での課題として次の2点が挙げられる。

- オクルージョンへの対処
 - 位相の変化
 - 対応付け
- 場景変化への対処

まずオクルージョンへの対処であるが、これに対して2つの問題がある。一つは Snakes 自身が単一の輪郭線をモデル化しているので、追跡途中で追跡対象同士の重なりにより起こるオクルージョンに対処できない。もう一つは複数の Snakes を扱う場合、オクルージョン発生後の対象物体の対応付けが困難という問題である。最初の問題は、Snakes の輪郭線の形状が変化する際に発生するねじれ(自己交差)現象を検出して輪郭線を切断したり、また複数の輪郭線が重なる際には、Convex hull などの幾何学的手法で統合することで、オクルージョン発生時の輪郭線の位相変化を行うことが可能になった [2]。もう一方の問題としてオクルージョン発生前後の対象物体の対応付けがある。これは輪郭線を統合、分離することで Snakes だけでは前後のフレーム間での対応付けが困難になるというもので、これに対しては領域情報(色、輝度等)や動きの情報を用いて対応付けを行うと

いう方法がある [4] . しかし従来行われてる手法ではオクルージョンの発生前後でしか対応付けを行っておらず , 頻繁にオクルージョンが起こった際の対象物体の対応付けが正確に行えない場合がある .

またもう一方の課題である場景変化への対応であるが , ここで述べる場景変化とは , 背景の状態が変化したり , カメラの視点が変動したりする状態を指す . 従来手法の Snakes を用いた手法では Snakes の画像エネルギーとして , 背景画像との差分や前後フレーム間の差分 , またオプティカルフロー等を用いている . これらの画像エネルギーを用いた場合は , 撮影条件に ” 場景変化がない ” という制約がつくことになる . しかし実際の追跡を行う際には , そのような制約は無い方が好ましい . この場合これに対処する手法として次の 2 つが考えられる . 一つは画像中の背景構造を理解して , アフィン変換などを用いて背景位置合わせを行って画像を補正する [5] というものである . しかしこの場合も背景の動き等に制約が付いてしまうといった問題がある . そこで背景の動きに制約されず場景変動が起こっても追跡が可能なもう一つの手法が Snakes に色情報のエネルギーを導入して追跡を行うものである .

本研究で提案する手法は , まず時系列画像全体を通して対象物体の管理を行い , オクルージョン発生時の対応づけを正確に行い , 場景変化に対しては画像中の対象とする領域の色情報を Snakes の新しい画像エネルギーとして導入することで抽出・追跡を行う . これにより動画像中の移動物体を追跡する際に撮影条件の制約を緩め , 動画像中の移動物体をより汎用的に抽出・追跡することを目指すものである .

1.3 本論文の構成

本論文は次のような構成になっている .

第 2 章では , 従来研究されている研究について述べる . 従来手法として , 領域分割を用いた手法 , 輪郭線を用いた手法について述べた後 , 本研究で用いる動的輪郭モデルを使った従来手法について検討する .

第 3 章では , Kass によって提案された , 動的輪郭モデル (Snakes) について簡単に説明した後 , 凹領域抽出や輪郭線の位相変化に対応した , 拡張された Snakes について説明する . 最後にそれら Snakes を用いた , 移動物体の追跡への応用について説明を行う .

第 4 章では , 本研究で提案する手法について述べる . 場景変化に対処するために新しい画像エネルギーを導入して移動物体の抽出・追跡を行う方法について述べ , 次に対象物体

の時系列全体での管理について述べる．

第 5 章では，提案した手法をもとに場景変化の起こる環境での追跡結果を示す．

第 6 章では，本研究のまとめとして結論と問題点について検討し，今後の課題について述べる．

第 2 章

動的輪郭モデルによる移動物体の抽出・追跡法

2.1 はじめに

本章では画像から目的の対象領域を抽出・追跡する際に従来行われてる手法として動的輪郭モデルを用いた手法を中心に検討を行い，従来提案されている手法の問題点について述べる．

2.2 領域分割を用いた手法

領域分割法は主に画素の特徴量に基づく分類(色，輝度，位置などのクラスタリング)を基本とし，またノイズに対しても比較的強いという特徴をもつ．近年、領域分割を用いた研究としては，顔画像からの顔部品の抽出のための顔領域の推定などに用いられている例 [6][7] や，また複数の特徴量から領域分割を行い移動物体の 2 次元的な動きを推定するのに用いられている例 [8] などがある．また領域分割を用いた対象領域の追跡手法 [9][10][11] では，顔領域の推定や，動き推定などといった領域分割による処理を時系列画像毎に行った後に連続する前後のフレームで対応する対象同士の対応づけをして追跡を行う．

しかしながらこのような領域分割を用いた手法の場合，分割された領域が各フレーム間画像で一致するとは限らず，前後フレーム間で追跡対象を一致させるために，別の処理を行わなければならないという問題がある．このような処理は一般的に複雑な作業となる．

実際の画像中から複数の物体を追跡する際には，更に複雑な処理が必要になる可能性がある．

2.3 輪郭抽出を用いた手法

2.3.1 エッジ抽出法

画像から目的の物体を抽出することは輪郭線を抽出することと同じ意味を持つことから，輪郭線抽出を用いた物体の抽出・追跡が行われている．しかし全体像の情報なしにエッジ点を抽出して局所的に接続するというエッジ抽出法のボトムアップな処理では，目的とする物体の閉じた輪郭線を得ることは困難な場合が多い．とくに画像中にノイズが含まれるような場合は，ヒューリスティックな情報なしに目的の輪郭線をえることは，ほとんど不可能である．輪郭線を用いた物体の抽出・追跡の場合に最も問題になるのがこの点であり，いかに安定して輪郭線を抽出するかが課題になる．

2.3.2 動的輪郭モデル

輪郭線抽出を直接的に行う方法として Kass らによって動的輪郭モデル (Snakes) が提案されている．Snakes は最初から輪郭線をモデル化した関数を定義して，それを目的の画像に対してフィッティングすることで抽出を行うものである．モデル化した関数のエネルギー項により画像上でその輪郭線を動的に移動させ目的の領域を抽出することから動的輪郭モデルと呼ばれる．またモデル化した関数の収束解を求めるため繰り返し計算を行い，その過程を画像上では輪郭線が蛇のように動くことから Snakes とも呼ばれている．M.Kass が提案した Snakes は滑らかさの制約とエッジ点に近づくようにモデル化している．また関数の形を決めず，関数として自由度を大きくしてどのような形の関数でも表現できるものになっている．

Snakes の特徴としては次の 2 点が挙げられる．

- 閉領域を囲む輪郭線を安定して抽出可能
- 移動物体の抽出・追跡を同時並行に行うことが可能

第一の特徴として，初期輪郭として閉曲線がはじめから与えられることから同じ輪郭線抽出法であるエッジ抽出とは異なり，閉領域を安定して抽出することが可能である．

第二の特徴としては連続的な画像においては，前のフレームでの処理結果を次のフレームでの処理の初期値として，用いるとすることができるので動画像中の物体の抽出・追跡には特に有効である．

Snakes を用いた物体の抽出・追跡の手法として次のようなものが提案されている．

Sampled Active Countour Model

M.Kass らによって提案された Snakes は，ノイズに捕獲されやすく，エネルギー関数の重みづけをするパラメータの決定が複雑になり，収束解を求めるために関数エネルギーの最小化問題を解く必要があった．エネルギーの最小化問題を解くことは，非常に計算コストがかかり処理に時間を要することから，橋本ら [12] は，新しい動的輪郭モデルとして処理が単純で離散的な Sampled Active Contour Model(S-ACM) を提案した．S-ACM は独立した輪郭点 p_i で構成されている．各輪郭点は，S-ACM を収縮させる圧力と輪郭点間に働く引力により独立に移動し，結果として S-ACM の形状が変化する．

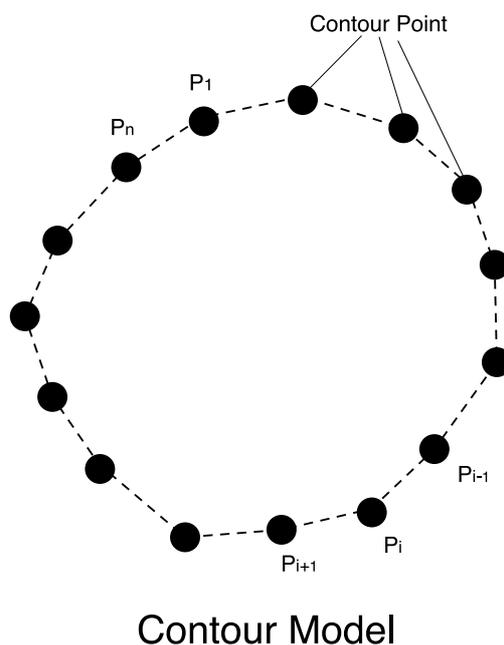


図 2.1: 輪郭モデル (S-ACM)

輪郭点は次の式 (2.1) で定義された力により移動する．

$$F^{(i)} = F_{p^{(i)}} + F_{a^{(i)}} + F_{r^{(i)}} \quad (2.1)$$

ここで, $F_{(i)}$ は輪郭線上の p_i にかかる力の総和である. $F_{p(i)}$ は輪郭線上の p_i にかかる周りからの圧力, $F_{a(i)}$ は輪郭線上の p_i に隣接する p_{i-1}, p_{i+1} との間で働く引力,

$F_{r(i)}$ は輪郭線上の p_i が画像の濃度値から決定されるエッジによりうける抗力である.

S-ACM では輪郭点 p_i が隣り合った 2 点 p_{i-1}, p_{i+1} 以外に影響されないため, 高速に輪郭線を修正できる.

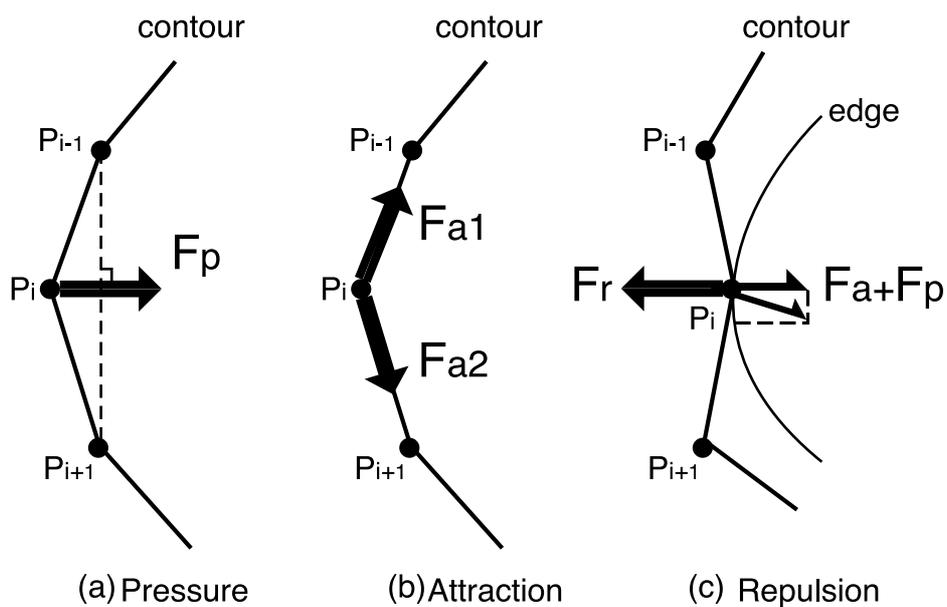


図 2.2: S-ACM で輪郭線上のサンプリング点 p_i にかかる力の種類: (a) 輪郭線上の p_i にかかる圧力 $F_{p(i)}$; (b) 輪郭線上の p_i にかかる引力 $F_{a(i)}$; (c) 輪郭線上の p_i が対象のエッジに近づくことにより生じる抗力 $F_{r(i)}$.

S-ACM の特徴をまとめると以下のようになる .

1. ノイズに強い.

輪郭線上の p_i がノイズで停止すると隣接する 2 点 p_{i-1} , p_{i+1} の引力により引き寄せられ, ノイズを飛び越えることができる.

2. 処理時間が短い.

輪郭線上の p_i の移動は, 隣接する 2 点 p_{i-1} , p_{i+1} 以外に影響されないため処理時間が短くてすむ.

3. 並列化が可能である.

輪郭線上の p_i は輪郭線の形状とは独立して移動する (隣接する 2 点 p_{i-1} , p_{i+1} 以外とは独立である) ため力の計算を並列化することができる.

4. 凹型の輪郭線抽出が困難である.

対象の凹型の場所へ p_i が入り込もうとすると力と隣接する 2 点 p_{i-1} , p_{i+1} の引力が相殺する .

Active Net

坂上ら [13] は領域を抽出する際に 1 次元的な輪郭上の情報だけでは不十分と考え, 領域情報を有効に活用するという観点から, 1 次元的な輪郭モデルを 2 次元的な面モデルへ拡張し, エネルギー最小化原理に基づいた動的な網のモデル Active Net を提案した .

Active Net の網モデルは, 次の式 (2.2) の 2 次元の点列パラメータで表現する .

$$v(p, q) = (x(p, q), y(p, q)) \quad (2.2)$$

Active Net の各格子点は, 4 近傍の点 $v(p, q - l)$, $v(p - k, q)$, $v(p + k, q)$, $v(p, q + l)$ と継れて網を形成する .

Active Net ではこの網に対して 3 種類のエネルギーを定義する . 網自身の内部歪みエネルギーを E_{int} とする, また網と画像との適合性エネルギーを E_{image} とし, 外部からの

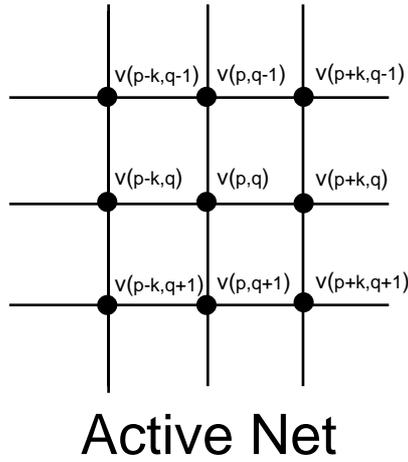


図 2.3: 動的な網モデル (Active Net)

強制力に対応するエネルギーを E_{con} とする．網のエネルギーはこの 3 つのエネルギー関数の線形結合で以下の式 (2.3) で記述される．

$$E_{net} = \int_0^1 \int_0^1 \{E_{int}(\mathbf{v}(p, q)) + E_{image}(\mathbf{v}(p, q)) + E_{con}(\mathbf{v}(p, q))\} dpdq \quad (2.3)$$

Active Net は，初期網の位置を初期値として，式 (2.3) を反復演算して求まる近似解に対応して動き，最終的に平衡状態になって解が収束した状態（すなわち領域にフィットイングした網の形状）で停止する．

Active Net の特徴をまとめると以下のようなになる．

1. 領域の内部の特徴を考慮した対象抽出が可能である．

Active Net は，面的な網のモデルを採用したことから，画像のテクスチャー情報などの領域情報をエネルギー関数として定義することが可能になり，領域の内部の特徴を考慮した領域抽出が可能になった．

Active Tubes

古川ら [14] は，連続画像から非剛体物体の運動を解析するために，Active Tubes と呼ばれるモデルを提案した．Active Tubes は，連続画像の各フレーム上の Snakes を時間軸に沿って積み重ね，隣接フレーム上の対応点を接続したモデルであり，Snakes と同様のエネルギー最小化の手法を用いて物体の運動を解析する．

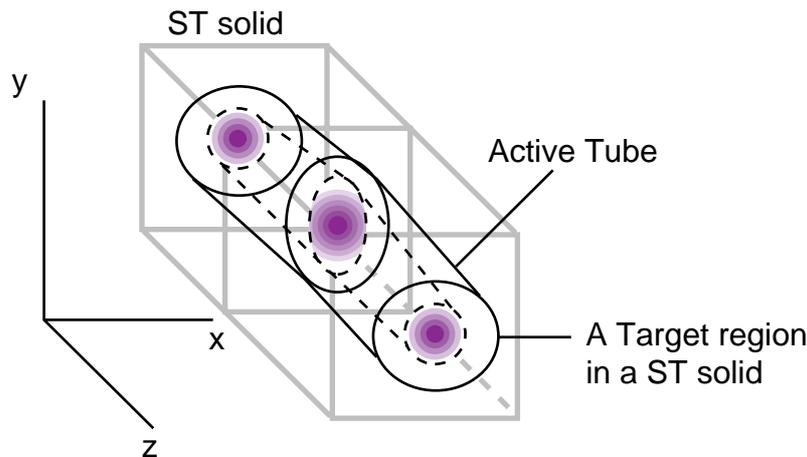


図 2.4: Active Tubes

Active Tubes は，時空間画像中の動物体に対応する領域を抽出するためのモデルである．このモデルは図 (2.4) に示すように時空間内で管状の形状をしており，時空間画像中の柱状領域に引き寄せられる性質をもつ．

Active Tubes は，抽出したい領域の境界に張りつくように変形，収束する．収束した後にその内部領域を取りだすことで動物体の抽出が実現できる．

Active Tubes は，その表面上の点の位置ベクトル $v(s,t)$ の集合で表現される．ここで s は空間領域のパラメータを， t は時間領域のパラメータを表す．また Active Tubes のエネルギー関数は以下の式 (2.4) で定義される．

$$E_{tube} = \int_T \int_0^1 \{E_{int}(v(s,t)) + E_{image}(v(s,t))\} ds dt \quad (2.4)$$

ここで E_{int} は内部エネルギーと呼ばれ， $v(s,t)$ 周辺の Active Tubes の形状によって定義され，Active Tubes の形状を滑らかにし，管状の形状を維持するように働く．一方 E_{ext} は外部エネルギーと呼ばれ，時空間画像中の各点で定義される一種のポテンシャルエネルギー

ギーである。Active Tubes を目的の動物体に収束させるために、外部エネルギーは対応する領域の表面の極小値を持つように定義される。

Active Tubes の特徴をまとめると以下のようなになる。

1. オクルージョンが途中で発生しても抽出領域の復元が可能である。

Active Tubes は、そのエネルギー関数の中に Snakes などにはない時間軸方向への滑らかさや連続性の項を持つことから、あるフレームで収束にその前後のフレームからの拘束力を働かせることができ、またその拘束力をエネルギー関数の重みづけで調整することが可能である。これを利用して追跡途中のあるフレームで対象領域にオクルージョンが発生した場合にその前後のフレームからの拘束力を用いることで、抽出領域の復元が可能になる。

2. 初期位置の抽出が困難である。

Active Tubes は、すべてのフレームにおいて初期位置を準備していなければならない、2次元の動的輪郭モデルを利用する場合に比べて、初期位置の抽出がより困難である。

C-Snakes

今まで述べてきた、従来手法としての動的輪郭モデルの2次元、3次元への拡張は単一領域の抽出・追跡を前提にしたものであったため分離・統合といった輪郭線の位相変化への対応は考慮されておらず、画像上での重なりにより発生するオクルージョンへの対処や、複数の対象を追跡することは不可能であった。そこで提案されたモデルが M.Kass らの Snakes をもとに位相変化に対応できるように拡張した Cross Snakes (C-Snakes) である。

C-Snakes のエネルギー関数は E_{int} からなる形状に依存した内部エネルギーと、 E_{image} からなる画像エネルギーから構成されており、両エネルギー関数の線形結合からなる。また C-Snakes は内部エネルギーに領域内の面積に相当するエネルギー項を持ち（この面積項については次章で詳しく述べる）このエネルギーの作用により S-ACM などでは不可能だった、凹領域を抽出することが可能になった。

しかしながら輪郭線を収縮する過程で図 (2.5) のような Snakes の輪郭線同士が自己交差する現象が起こる場合がある。この現象が起こった場合、この自己交差によってできた

領域は負の面積エネルギーを持つことからこのままの状態では変形を続けると自己交差によってできた領域が増大していくことになる。またこのような領域は対象領域を含まないので、自己交差によってできる領域の生成を行わないことが望ましい。

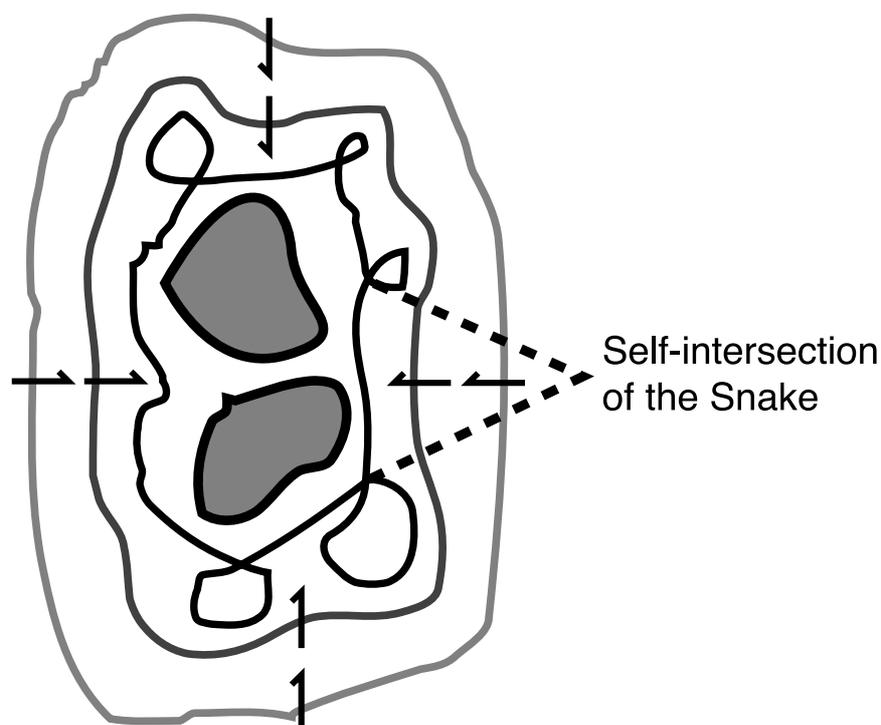


図 2.5: Snakes の自己交差現象

C-Snakes ではこの現象を利用して、輪郭線の自己交差を検出し一つの輪郭から複数の輪郭への分離を可能にしている。これにより複数の対象を抽出・追跡することが可能になった。C-Snakes を用いた抽出・追跡手法として C.Vieren ら [3] や荒木ら [2], 金田ら [4] が提案している。いずれの手法も複数物体への対応が可能で、分離だけでなく幾何学的な処理により統合まで可能なように拡張されている。

2.4 従来の動的輪郭モデルを用いた手法の問題点

しかしながらこれらの提案された手法は、画像エネルギーとして背景との差分や連続画像におけるフレームの前後での差分、オプティカルフロー等を用いている。この場合背景変化のない撮影環境でないと抽出・追跡を行うことが困難である。また分離、融合が可能

になり対象物体が複数になったことからこれら複数の対象を管理する必要がでてきた。しかしながら金田の提案している手法では、融合直前の情報のみから対象の対応づけを行っていたことから正確な追跡が困難になる場合があった。

C-Snakes の特徴をまとめると以下のようなになる。

1. 凹領域の抽出が可能である。

C-Snakes は、面積をエネルギーにもつ面積項を導入し、また輪郭線上のサンプリング点すべてに対して毎回エネルギー計算を行うことにより S-ACM では不可能だった対象物体の凹領域に対しても入り込むことが可能である。

2. 初期輪郭の設定をする必要がない。

C-Snakes は、分離が可能なことから、前述した S-ACM や Active Net, Active Tubes のように対象領域近傍に初期輪郭線 (初期網) を設定する必要がない。

3. 複数の対象を抽出・追跡することが可能である。

上述したように C-Snakes は、輪郭線分離が可能なことから複数の物体を同時に抽出・追跡することが可能である。また対象同士が重なりあって起こるオクルージョンに対しても輪郭線を統合できるように拡張することで対処可能である。

2.5 おわりに

本章では従来の手法について動的輪郭モデルを中心に述べてきたが、動画像中から移動物体の抽出・追跡を行う。

まずオクルージョンへの対処であるが、領域分割を用いた手法では対象同士のフレーム間での対応が保証されていないことから、この問題に対処するのは困難である。また対象同士のフレーム間での対応がある程度保証されている動的輪郭モデルを用いた場合は、輪郭線分離、統合といった位相変化に対応したモデルでないと不可能である。位相変化を考慮にいれた場合 C-Snakes がこれに対処可能である。しかしながらオクルージョンが発生

して輪郭線が統合され、また分離するような過程では対象同士のフレーム間での対応が保証されなくなることから、オクルージョン発生時に対応づけを正確に行うための処理が必要で、従来の手法では問題がある。

次に場景変化への対処であるが、C-Snakes を用いた抽出・追跡手法としていくつか挙げられていたがいずれの場合も画像エネルギーとして場景変化がない撮影条件に依存した画像エネルギーを用いていることから、これらの手法をそのまま用いるのは困難である。そこで場景変化のない撮影条件に依存しない画像エネルギーを考慮する必要がある。領域情報を積極的に活用するのに向いているのは Active Net のような面モデルであるが、オクルージョンへの対処という点で、Active Net は位相変化に対応できていない。

そこで本研究では C-Snakes のに従来のような背景との差分や連続画像におけるフレームの前後での差分、オプティカルフロー等を用いたものでなく、領域の色情報をエネルギーにする新たな画像エネルギーを導入して抽出・追跡を行う。

第 3 章

動的輪郭モデル Snakes

ここでは Snakes 全般の説明を行うとともに、本研究で使用している拡張された Snakes を各機能ごとにまとめて説明を行う。

3.1 Snakes の概要

M.Kass らは輪郭抽出問題をプロファイルフィッティングとして扱い Snakes を提案した。Snakes 画像中に設定する閉曲線 $\mathbf{v}(s) = (x(s), y(s))$ 上での、内部エネルギー E_{int} 、画像エネルギー E_{image} 、外部エネルギー E_{con} の線形結合からなるエネルギー関数 E_{Snakes} を定義する。Snakes を用いて輪郭抽出を行うことは、エネルギー関数 E_{Snakes} の最小化問題を解くことと等価である。Snakes は、初期値として初期輪郭を与え、エネルギー関数 E_{Snakes} を最小化するように輪郭線の形状を修正していくことで目的の領域の抽出を行う。エネルギー関数 E_{Snakes} は、以下の式 (3.1) で表される。

$$E_{snakes} = \int_0^1 \{E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s))\} ds \quad (3.1)$$

次に各エネルギー関数 E_{int} 、 E_{image} 、 E_{con} の説明を行う。

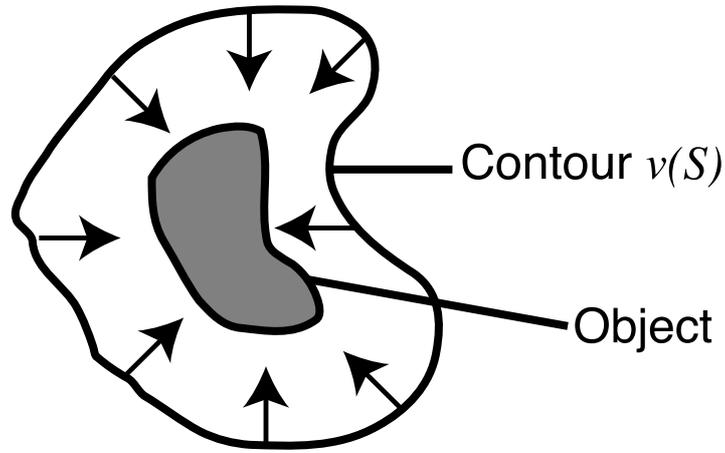


図 3.1: Snakes の輪郭線の挙動

内部エネルギー

内部エネルギー E_{int} は, Snakes の輪郭の形状に依存しており輪郭線の形状を滑らかにしながら収縮させる. エネルギーとして以下の式 (3.2) 定義される.

$$E_{int}(\mathbf{v}(s)) = E_{spline}(\mathbf{v}(s)) = \frac{1}{2} \{ \alpha \|\mathbf{v}_s(s)\|^2 + \beta \|\mathbf{v}_{ss}(s)\|^2 \} \quad (3.2)$$

ここで, α, β は E_{int} の重み係数を表す. また, $\mathbf{v}_s(s) = d\mathbf{v}(s)/ds$, $\mathbf{v}_{ss}(s) = d\mathbf{v}_s(s)/ds$ である. E_{int} を減少させることは, $\mathbf{v}(s)$ を収縮させかつ滑らかな曲線となるように変形させることである.

画像エネルギー

E_{image} は画像エネルギーと呼ばれ, $\mathbf{v}(s)$ が存在する位置の画像の状態により定まるエネルギーである. 式 (3.3) で定義される画像濃度の 1 次微分 (画像のエッジ) が用いられることが多い.

$$E_{image}(\mathbf{v}(s)) = -\frac{1}{2} w_{image} \|\nabla I(\mathbf{v}(s))\|^2 \quad (3.3)$$

ここで, w_{image} は E_{image} の重み係数を表し, $\mathbf{v}(s)$ の位置の画像濃度値を $I(\mathbf{v}(s))$ とする. E_{image} を減少させることは, $\mathbf{v}(s)$ がエッジ点に引き寄せられるようにすることである.

外部エネルギー

E_{con} は外部エネルギーと呼ばれ、外部から与える制約であり、 $v(s)$ の形状を強制的に変形させたいときに利用する。

3.2 拡張された Snakes

M.Kass の提案した Snakes をもとに目的に応じた拡張がなされてきた。ここでは移動物体の抽出・追跡という観点から本研究で用いている拡張された Snakes について述べる。

3.2.1 面積項を導入した Snakes

Snakes の輪郭内の面積に相当するエネルギーとして坂口等が提案した面積項 [15] がある。このエネルギーを内部エネルギー E_{int} もしくは、外部エネルギー E_{con} として用いることにより、面積を収縮させて凹形状の領域が抽出可能になる。面積項は以下の式 (4.3) で定義される。

$$E_{area}(v(s)) = \frac{1}{2} w_{area} \{x v_y(s) - v_x(s) y\} \quad (3.4)$$

ここで、 w_{area} は面積項の重み係数を表す。

次に面積項 E_{area} の算出原理を図 3.2.1 を用いて説明する。 E_{area} を以下のように 2 つの要素に分けて考える。

- $E_{area1} = x v_y(s)$
- $E_{area2} = v_x(s) y$

いま図 3.2.1 のように $v(s)$ が形成されているとする。

- $v(s)$ 上の s_i, s_{i+1} の線分の範囲で E_{area} を計算した場合

$v(s)$ 上の s_i, s_{i+1} の線分の範囲で E_{area} を計算して求まる面積は図 3.2.1(a) の領域となる。

- E_{area1} だけを計算した場合

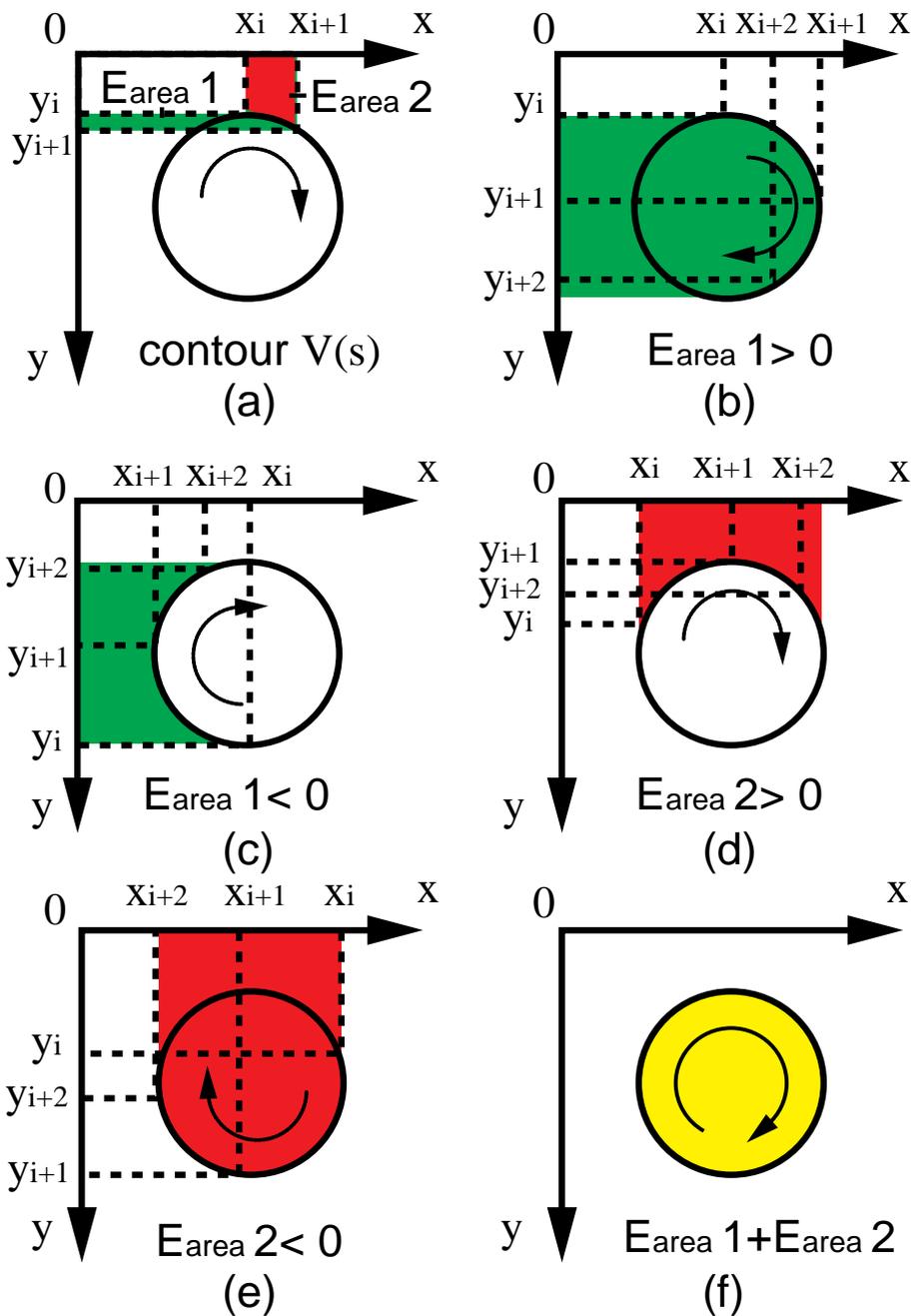


図 3.2: 面積項 E_{area} の算出原理:

(a) $v(s)$ 上の s_i, s_{i+1} の線分の範囲で E_{area} を計算して求まる面積 (E_{area1} と E_{area2}); (b) $v(s)$ が $v_y(s) > 0$ となる線分の範囲で E_{area1} を計算して求まる面積; (c) $v(s)$ が $v_y(s) < 0$ となる線分の範囲で E_{area1} を計算して求まる面積; (d) $v(s)$ が $v_x(s) > 0$ となる線分の範囲で E_{area2} を計算して求まる面積; (e) $v(s)$ が $v_x(s) < 0$ となる線分の範囲で E_{area2} を計算して求まる面積; (f) E_{area} を計算して求まる面積 .

E_{area1} だけを計算して求まる面積の値は $v(s)$ が $v_y(s) > 0$ となる線分の範囲では正となり (図 3.2.1(b)), $v_y(s) < 0$ となる線分の範囲では負となる (図 3.2.1(c)). これらを輪郭にそって求めた総和は Snakes の輪郭内の面積になることがわかる .

- E_{area2} だけを計算した場合

E_{area1} と同様に Snakes の領域の面積が求まる .

- E_{area} を計算した場合

E_{area} を計算して求まる面積は E_{area1} と E_{area2} の総和を半分にしたものとなる (E_{area1} , E_{area2} で求まる面積が重なるため面積を半分にする必要がある図 3.2.1(f)).

3.2.2 分離・融合機能を導入した Snakes(C-Snakes)

本節では荒木 [2] によって提案された位相変化へ対応した Snakes(C-Snakes) について説明を行う . 位相変化に対応することは , 移動物体の抽出・追跡において必須の機能といえる . また位相変化に対応することで複数物体に対しても抽出・追跡が可能になる . ここでは交差の検出 , 分離の仕組み , 融合の仕組みについて順に説明を行う .

輪郭線の交差検出

前節で述べた面積項を Snakes に導入することで，面積が収縮する際に輪郭線同士がお互いに交わる自己交差という現象が起こる．C-Snakes はこの自己交差を検出して輪郭線を分離する．交差検出の流れについて以下に述べる．

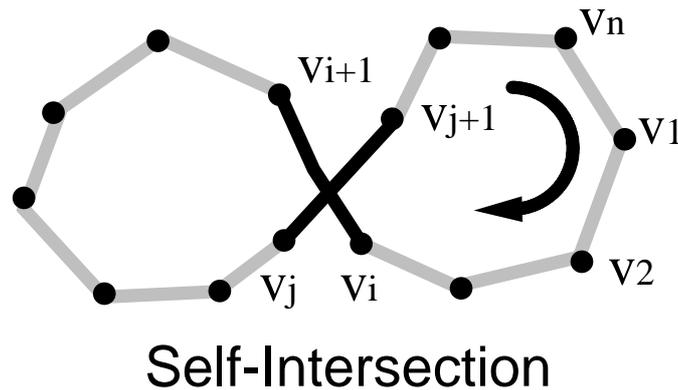


図 3.3: Snakes の交差

交差検出は Snakes の輪郭線上の隣接する離散点間の線分とそれ以外の隣接する離散点間の線分で行われる．(例えば線分 $V_1 V_2$ に対して線分 $V_2 V_3, V_3 V_4, \dots, V_{n-1} V_n$ のように) この際，線分が交差しているかどうかを式 3.5 を用いて判定する．

$$p(\mathbf{v}_{i+1} - \mathbf{v}_i) + \mathbf{v}_i = q(\mathbf{v}_{j+1} - \mathbf{v}_j) + \mathbf{v}_j \quad (3.5)$$

図 3.3 のように線分 $V_i V_{i+1}$ と線分 $V_j V_{j+1}$ が交差している場合には式 (3.5) を満たす実数 $p(0 \leq p \leq 1)$, $q(0 \leq q \leq 1)$ が存在する．これにより Snakes の交差判定を行う．C-Snakes では交差判定により輪郭線の分離・統合を行う．

輪郭線の分離

前節で述べた交差判定をもとにここでは，C-Snakes の分離の流れを説明する．

1. 交差の判定

初期輪郭から輪郭線を修正する過程で自己交差が起こる．C-Snakes は輪郭線を修正すると同時にその時点で交差が起こっていないか交差判定を行う．

2. 輪郭線の切断

式 (3.5) による交差判定を行い交差が発生する場合には、交差している線分の始点と終点を交互に結ぶ。図 (3.8) の場合 2 箇所交差が起こり、それにより 3 つの輪郭線に分かれる。

3. 面積項による判定

最後に面積項による判定を行う。前の手続きで 3 つの輪郭線に分かれたが、この 3 つうち本当に対象物体を含んでる輪郭線だけを真の輪郭線としてそうでない輪郭線を削除する。時計回りの方向の面積項の値を正と定義すると図 (3.8) では、3 つの輪郭線の E_{area} を計算した場合、輪郭線内に何も含んでない E_{area} が負の値を取るこれにより E_{area} が正の輪郭線だけを真の輪郭線として残し、負の輪郭線を削除して最終的に 2 つの輪郭線になる。

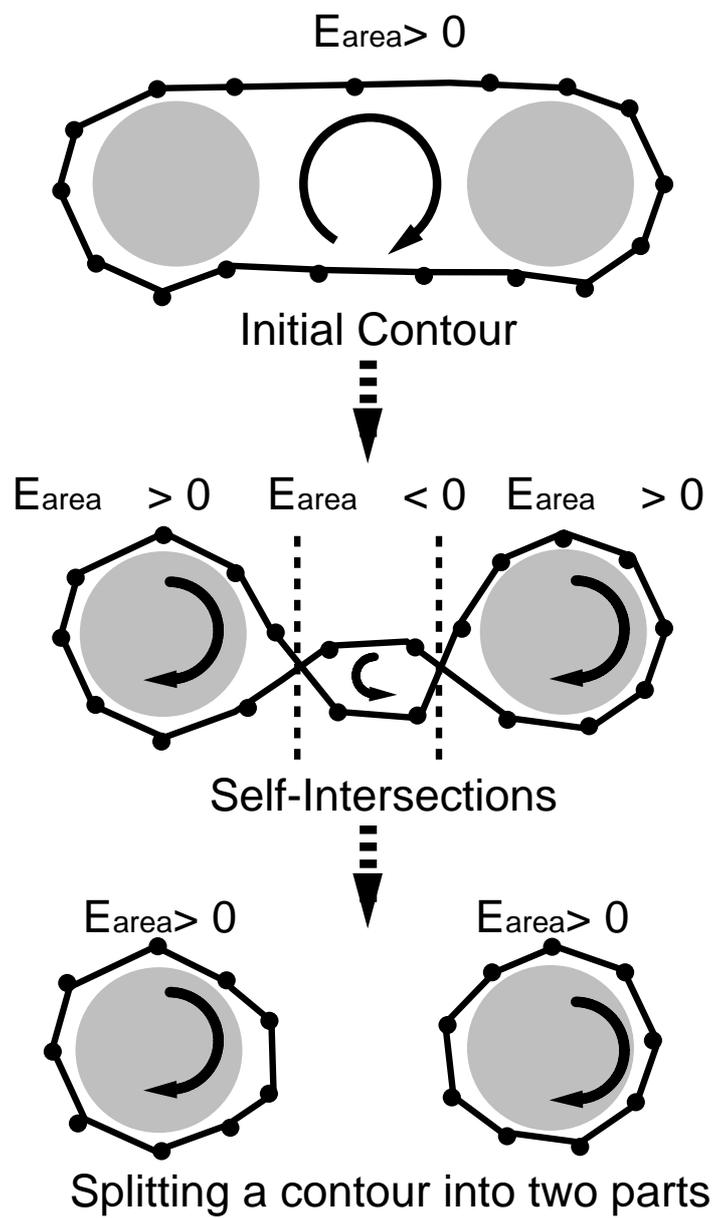


図 3.4: Snakes の分離の仕組み

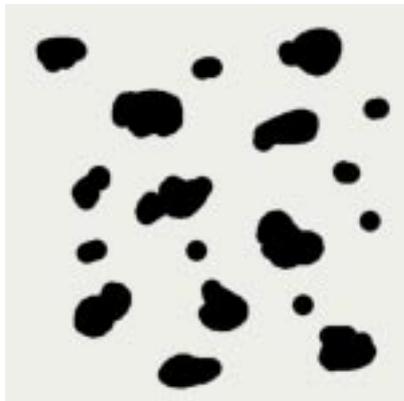


図 3.5: 入力画像

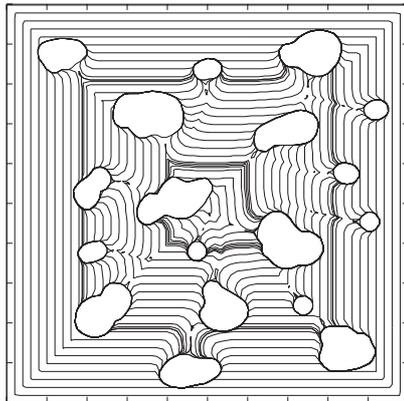


図 3.6: C-Snakes 実行

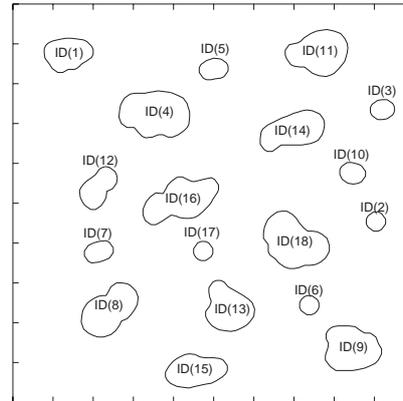


図 3.7: 輪郭抽出結果

図 (3.5) , (3.6) , (5.12) にシミュレーション画像を使った C-Snakes の実行結果を示す . シミュレーション画像は , 512 × 512 画素の 256 階調グレースケール画像を使用し , 初期輪郭は画像外枠に設定した . 画像中には 18 個の斑点がありそれを C-Snakes を実行して抽出した . 前述した C-Snakes の分離の手続きにより 18 個全ての輪郭線が抽出できていることがわかる .

輪郭線の統合

最後に位相変化に対応するためのもう一方の機能として輪郭線の統合について説明する . 輪郭線の統合は 2 つの物体が画像中で重なり合う際に起こる . 以下の手順で輪郭線の統合を行う .

1. 交差の判定

C-Snakes は輪郭線の自己交差を検出するようになっていたが , さらにここでは外部の他の輪郭線との交差検出を行うように C-Snakes を拡張する . 判定には分離の時に用いたものと同じ式 (3.5) を用いて判定する . 一方の隣接した離散点間の線分に対して交差しているもう一方の輪郭線の離散点間の線分すべてとの交差判定を行う .

2. 輪郭線の統合

輪郭線の統合については幾何学的な処理としていくつかのアルゴリズムが提案されているが、本研究ではもっとも基本的な手法である凸包問題 (convex hull) で処理を行った。ここでの手順としては、式 (3.5) による交差判定を行い交差が発生する場合には、新しく形成される統合された輪郭線が、両方の輪郭線上の全ての離散点を含む最小凸多角形になるように処理を行う。ここでの交差判定は交差の有無だけを判定する。輪郭線が凹形状をしている場合、輪郭線は複雑に交差を起こす場合があるがこの手法ではそのような複雑な交差に対しても単純な処理で済む。

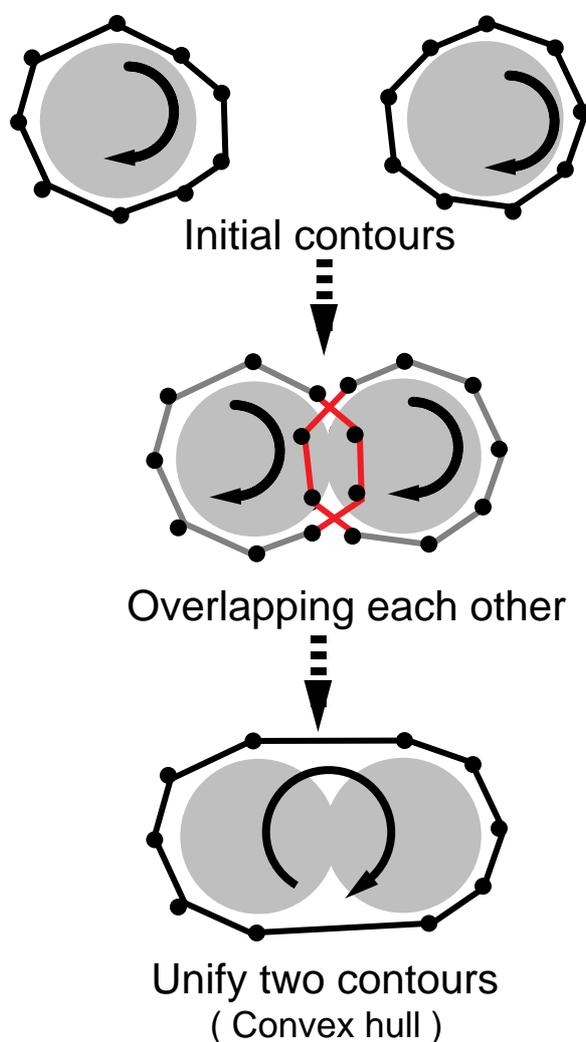


図 3.8: Snakes の統合の仕組み

3.3 Snakes の移動物体追跡への応用

Snakes が特に動画処理に有効であることは序論で述べたが、実際に動画中の移動物体の追跡への適用についてはここでは説明を行う。

3.3.1 移動物体の追跡処理

ここでは C-Snakes を用いた追跡処理と追跡を行ったあとの対象物体の軌跡決定処理について述べる。

追跡の流れ

1. 初期輪郭線の設定 (初期フレーム画像)

初期フレーム画像に対してすべての対象物体を囲むように初期輪郭を設定する (図 3.9(a))。例えば、画像フレーム枠の近傍に設定する。

2. 初期輪郭線の修整

対象物体の抽出を行うため輪郭線の修整を行う。

3. 輪郭線の追跡 (第 2 フレーム画像以降)

- (a) 初期輪郭線の設定

前フレーム画像において抽出された輪郭線 (図 3.9(c)) を少量膨張させ、これを現フレーム画像に対する初期輪郭として設定する (図 3.9(d))。

- (b) 輪郭線の統合処理

抽出された輪郭線を膨張させ次フレーム画像の初期輪郭として与えるために、輪郭同士の間が重なりが起こる場合がある、その場合、重なった輪郭線に対して輪郭線同士を統合する。(図 3.9(e))。

- (c) 繰り返し処理

(b) で得られた輪郭線に対し、追跡手順の (a)、(b) と同様にして、輪郭線を抽出する。

4. 領域内属性を用いた軌跡の決定

対象物体の軌跡を決定するために、移動する物体が重なり合う前の状態の領域の属性と物体が分離した後の状態の領域の属性とのマッチング処理を行うことで、軌跡の決定を行う。(図 3.9(f)).

使用する Snakes のエネルギー

ここでは追跡に用いた各エネルギー関数について簡単に説明する。まず内部エネルギーとしては、 E_{spline} 、と面積項 E_{area} 、輪郭線上の隣接する格子点間の距離を一定に保つための距離平均化項 E_{dist} を使用し、画像エネルギーとしては、エッジポテンシャル項 E_{edge} と輪郭線が対象領域内に入り込まないように制御する画像濃度項 E_{intens} を用いる。

$$\mathbf{E}_{snakes} = \mathbf{E}_{spline} + \mathbf{E}_{area} + \mathbf{E}_{dist} + \mathbf{E}_{edge} + \mathbf{E}_{intens} \quad (3.6)$$

$$E_{spline} = \frac{1}{2} \sum_{i=1}^n \{ \alpha \|\mathbf{v}'_i\|^2 + \beta \|\mathbf{v}''_i\|^2 \} \quad (3.7)$$

$$E_{area} = \frac{1}{2} \sum_{i=1}^n \{ w_{area} (x_i \mathbf{v}_y - \mathbf{v}_x y_i) \} \quad (3.8)$$

$$E_{dist} = \frac{1}{2} \sum_{i=1}^n \{ w_{dist} (d_{ave} - \|\mathbf{v}'_i\|)^2 \} \quad (3.9)$$

$$E_{edge} = -\frac{1}{2} \sum_{i=1}^n \{ w_{edge} \|\nabla I(\mathbf{v}_i)\|^2 \} \quad (3.10)$$

$$E_{intens} = \sum_{i=1}^n \{ w_{intens} I(\mathbf{v}_i) \} \quad (3.11)$$

移動物体の軌跡決定

軌跡決定処理は、輪郭追跡処理によって得られた輪郭線を基に動画像中で移動する対象物体の軌跡を決定する。この処理を行うことで、移動する対象物体が重なった場合でも対象物体の軌跡を決定することができる。

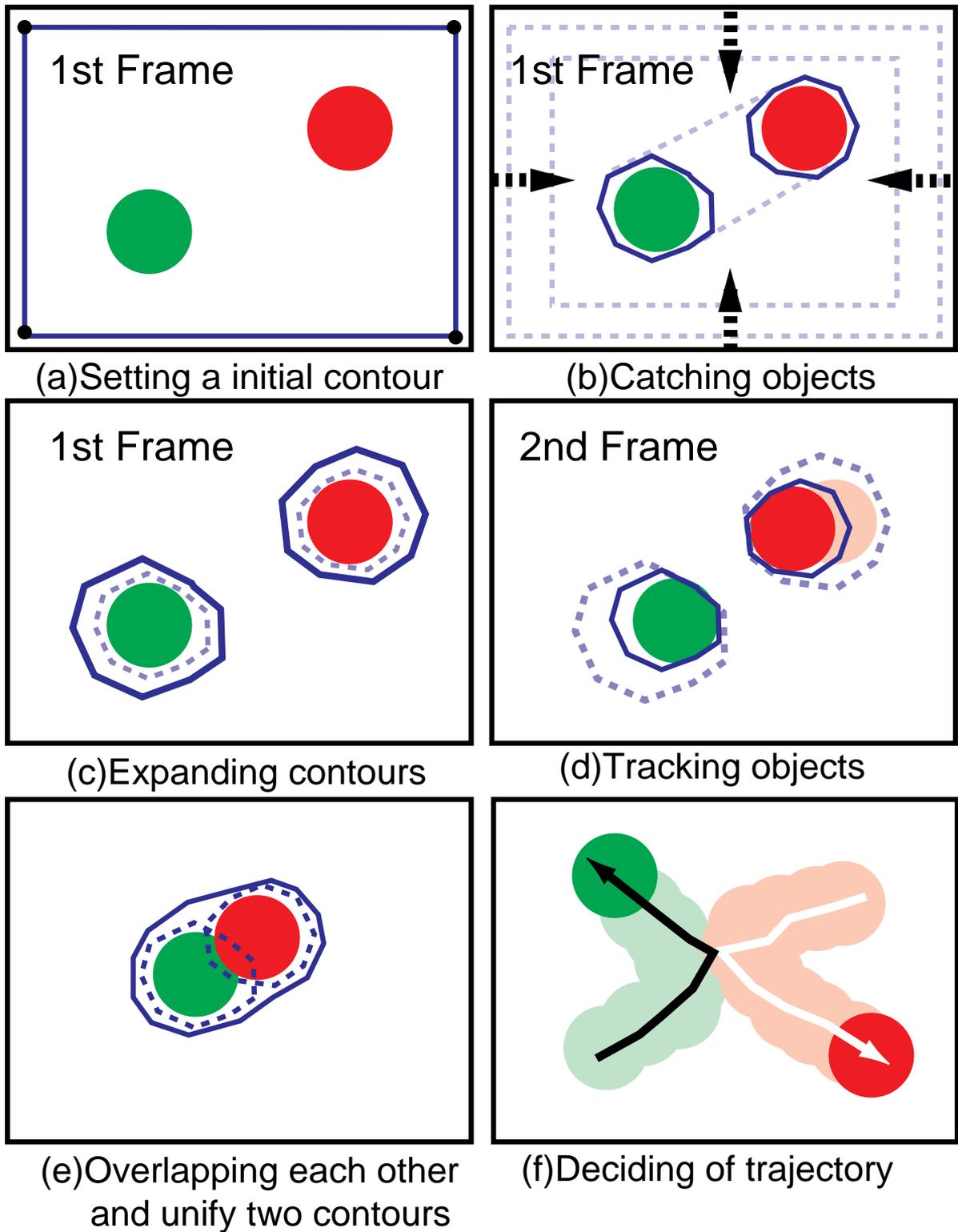


図 3.9: Snakes による追跡の流れ

対象物体の軌跡を決定するためには、輪郭線が統合される前のフレームと統合された後のフレーム間で輪郭線を対応づける処理が必要となる。対応づけには領域の色情報を用い、単純類似度法 (式 4.7) によって輪郭線の対応づけを行う [16]。

$$S_s [a, b] = \frac{(a, b)^2}{\|a\|^2 \|b\|^2} \quad (3.12)$$

ここで、式 (4.7) 中の特徴ベクトルを a, b として、本手法ではそれぞれ RGB 画像の平均輝度値を用いている。

次に背景差分に依存する画像エネルギーを用いた移動物体の追跡実験の結果 (図 3.10) を示す。

3.4 場景変動を考慮した移動物体の抽出・追跡法

ここまでは拡張された C-Snakes によって位相変化への対処とそれによって複数移動物体の抽出・追跡が可能になったことについて述べた。ここでは序論で述べたもう一方の課題である場景変動への対処について検討する。

3.4.1 場景変動への対処

動画像を処理する上で動画像中で目的の対象が移動しながら、背景の一部が動いたり、またカメラの動きから画面全体が移動するようなシーンを目にすることがある。このような動画像から目的の対象を追う方法として背景の位置をあわせがある。画像の背景に注目点を設定しそれを基準にして背景を補正する [5]。しかしこの場合背景の全ての動きをアフィン変換で仮定しているため誤検出がでる場合がある。

以上が背景位置合わせ的なアプローチであったが、本研究ではこのような背景位置合わせ的な手法でなく、Snakes に用いる画像エネルギーを変えることで場景変動時の移動物体の抽出・追跡を実現する。

Snakes は動画像処理には非常に有効な手段であることはこれまで述べてきたが、従来の Snakes を用いた追跡では背景画像との差分や連続する画像の前後フレームとの差分などを画像エネルギーとして用いてきたので、場景変動に対しての考慮がなされていなかった。そこで本研究では従来手法では軌跡の決定などにしか用いてこなかった領域内の情報で、とくに最も一般的な色情報を Snakes の画像エネルギーとして導入することで場景変動するような状況での移動物体の抽出・追跡を試みる。

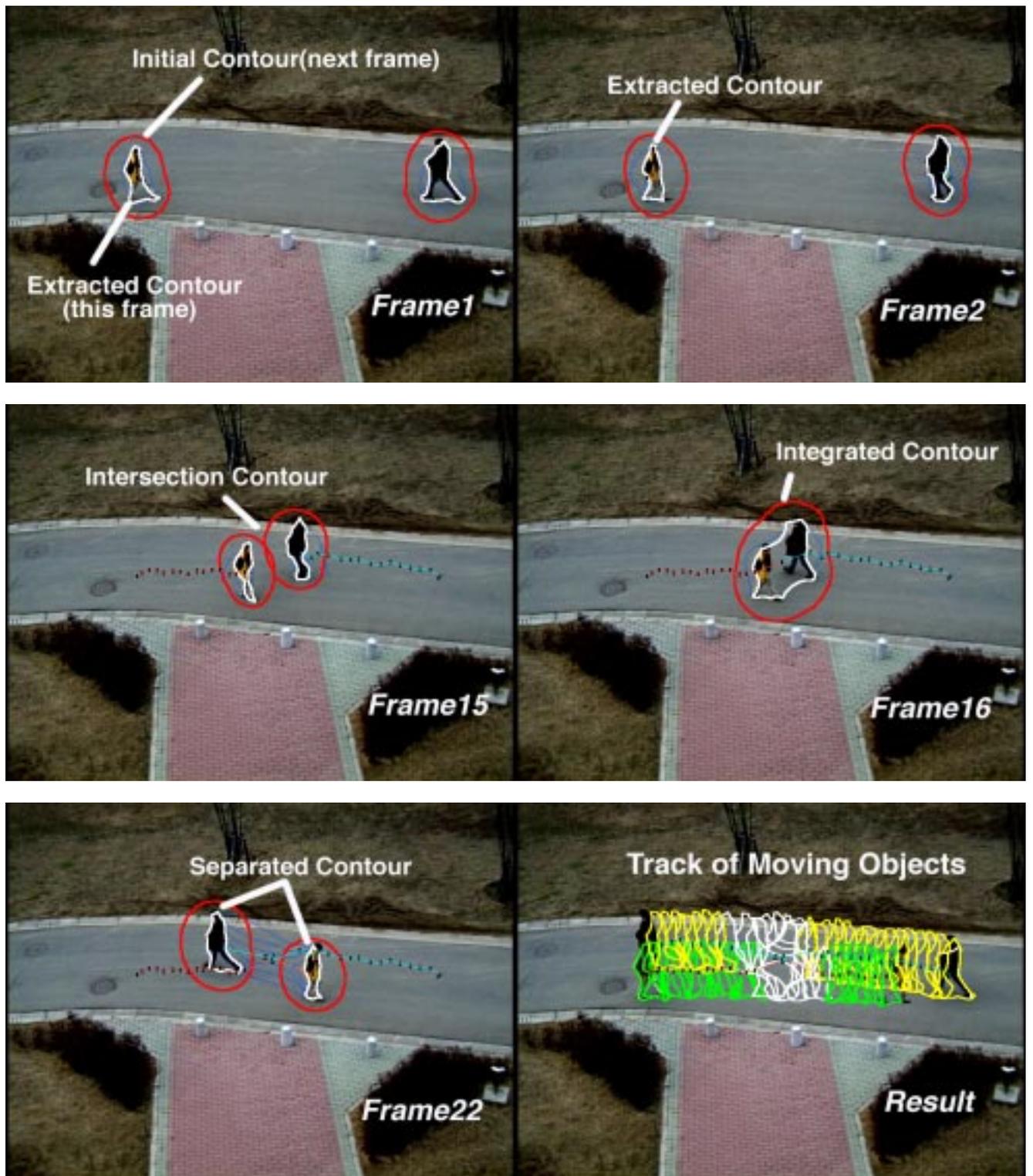


図 3.10: 背景差分に依存する画像エネルギーを用いた追跡結果: 撮影領域内に侵入する移動物体を自動で追跡. 白線は収束した輪郭, 赤線は収束した輪郭線をもとに膨張させた次フレームでの初期輪郭を示す.

3.4.2 新しい画像エネルギーを導入した Snakes

色情報を Snakes に導入する場合に考慮する点として次の 2 点がある .

- 画像エネルギーに入力するデータとしてできるだけ単純なものを用いる .
- ヒューリスティックな閾値設定を行わない .

Snakes に限らず動画処理は実時間での処理を目標にしており , カメラからの入力画像に対して最も一般的に用いられている色情報として RGB 表色系を本手法では用いる . またエネルギーとして用いる場合には , 自動化という点から特定のヒューリスティックな閾値を設定することは望ましくないと考えられる . 以上の観点から色情報の画像エネルギーへの導入について , J.lvins ら [17] の提案した手法と本研究で用いる K . P . Ngoi が提案している手法について検討を行う .

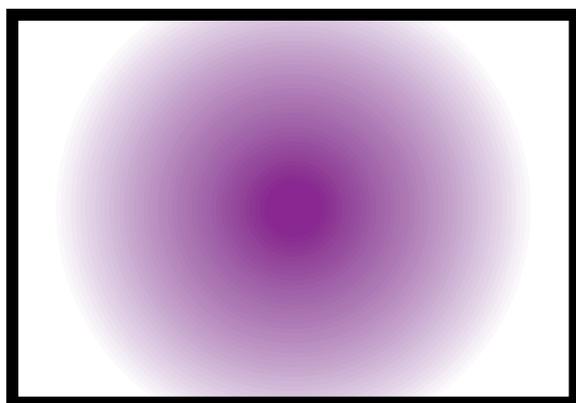
Multi-dimensional binary pressure

J.lvins ら [17] の提案した手法は , 抽出したい領域に対してその中に含まれる同じ色の領域 (seed region) 内の画素の色情報 (r, g, b) から入力画像全体の点の画素の色情報と seed region と色情報の Mahalanobis 距離 χ^2 を求め統計的な閾値 th を設定することでその閾値を越えない範囲で seed region の輪郭を Snakes を用いて成長させる . 図 (3.4.2) にその手順を説明する .

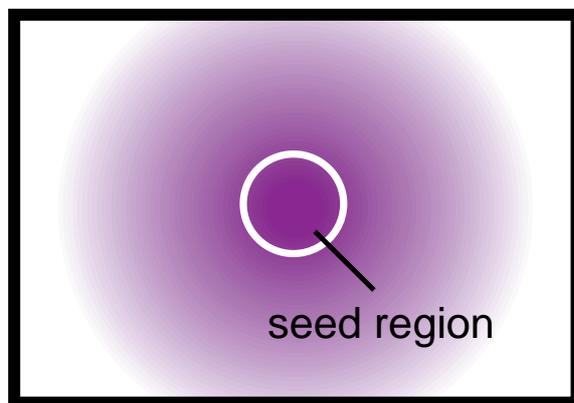
Color image energy

前述した手法ではヒューリスティックな閾値を用いていたが , 自動化という点でヒューリスティックな閾値を用いることは好ましくない . また実画像において特に自然の風景などで抽出する際の配慮がなく , 単純に基準になる領域とその他の領域との (r, g, b) 特徴量の距離で判断していた .

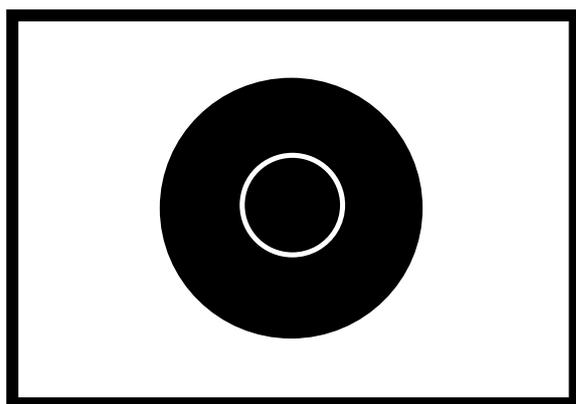
しかしながら太田 [19] らの研究によると従来領域分割などにもちいられてきた色情報も画像の構造によって , とくに野外のシーンでは大変影響を受けやすいことが報告されている . そこで K . P . Ngoi は , そのような影響を押えるためにカラー画像の明度 $E_{intensity}$ 式 (4.6) とコントラスト $E_{contrast}$ 式 (4.5) からなる 2 次元特徴空間を用いて Snakes の新しい画像エネルギー E_{image} 式 (3.13) として導入した .



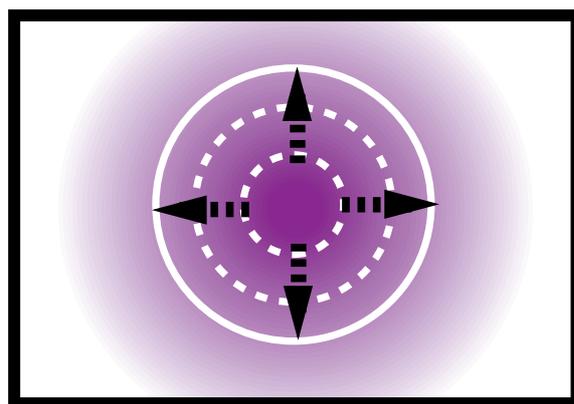
(a)Input image



(b)set seed region



(c)binary pressure image



(d)expanding contour

図 3.11: Multi-dimensional binary pressure:

(a) 入力画像; (b) 入力画像に対して抽出したい色と同じ領域 (seed region) を設定する . ; (c) seed region の領域内の色情報から seed region の平均画素値から全領域の画素値との Mahalanobis 距離 χ^2 を求めて、ある閾値で 2 値化し binary pressure を得る . ; (d) c の処理を行って求めた binary pressure をもとに seed region から Snakes で輪郭を膨張させ、目的の領域を得る . ;

$$E_{image} = \gamma E_{contrast} + (1 - \gamma) E_{intensity} \quad (3.13)$$

$$E_{contrast} = \sum_{i=0}^m \ln^2 \left\{ \frac{1+r_i}{1+r_0} \right\} + \ln^2 \left\{ \frac{1+g_i}{1+g_0} \right\} + \ln^2 \left\{ \frac{1+b_i}{1+b_0} \right\} \quad (3.14)$$

$$E_{intensity} = \sum_{i=0}^m \ln \{ 1 + |(r_i - r_0) + (g_i - g_0) + (b_i - b_0)| \} \quad (3.15)$$

次に式 (3.13) の γ について説明を行う。この γ は、 $E_{intensity}$ と $E_{contrast}$ との間のバランスを決定するもので、この値を最適に設定することによって目的の対象物体を Snakes がより正確に抽出することが可能になる。

本研究ではこの γ の値を自動決定するために Fisher の線形識別を用いる。2次元特徴空間 $E_{intensity} - E_{contrast}$ では対象物体 (object) と背景 (background) はできるだけ離れて分布し、object 同士、background 同士はなるべく固まった分布をとることが望ましい。それを実現するために Fisher の線形識別を用いて最適な座標変換を次の式 (3.16) で求める。

$$y = w^t x \quad (3.16)$$

これによりパターン x をパターン y に変換する。この w を重みベクトルと呼び、この場合重みベクトルは次のような式 (3.17) で与えられる。

$$w = \begin{pmatrix} w_x \\ w_y \end{pmatrix} = (S_o + S_b)^{-1} (m_o - m_b) \quad (3.17)$$

ここで m_o は object の平均、 S_o は分散を表し、 m_b は background の平均、 S_b は分散を表している。

最後に を次の式 (3.18) で求める .

$$\gamma = 2/\pi(\tan^{-1} |w_y/w_x|) \quad (3.18)$$

この γ は , 座標変換したあとの特徴空間がもともとの特徴空間に対してどれだけ傾いているかを示しており , E_{image} における $E_{intensity}$ と $E_{contorast}$ のバランスの最適な値を表す . $0 < \gamma < \pi/4$ で intensity が強調され , $\pi/4 < \gamma < \pi/2$ で contrast が強調される .

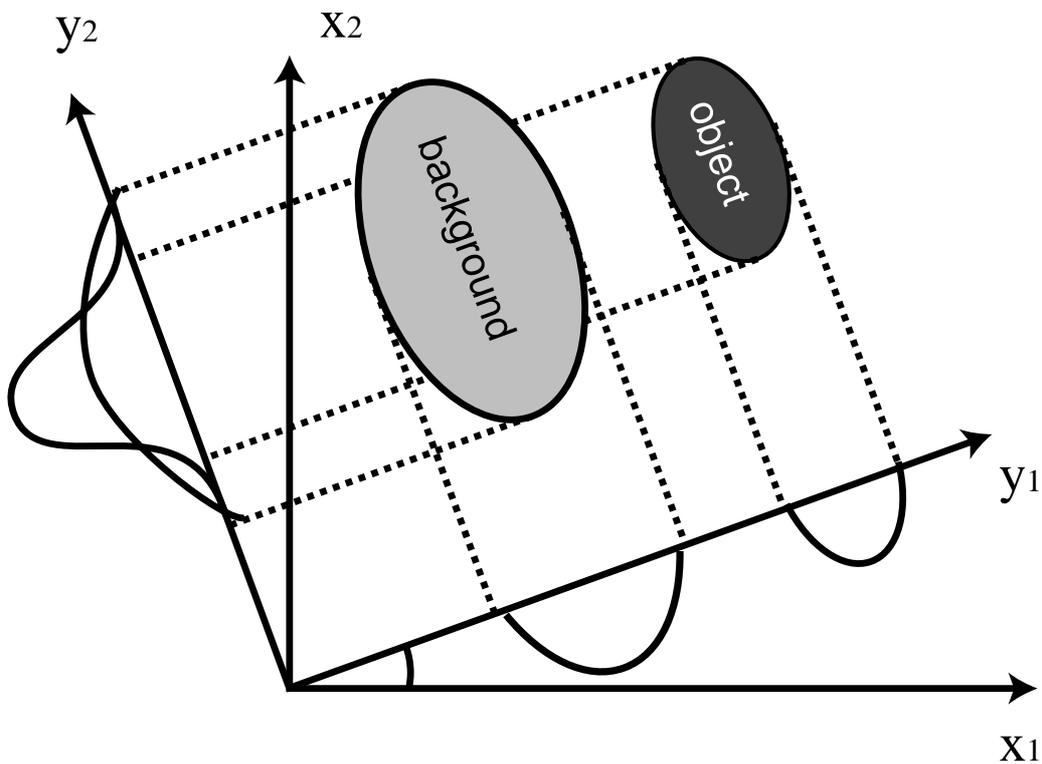


図 3.12: Fisher の線形識別による座標変換:

Fisher の線形識別を用いて $E_{intensity} - E_{contarst}$ 2次元特徴空間内の object クラスと background クラスを分離するのに最適な座標系に変換する .

このようにして最適な γ を決定することで画像エネルギー E_{image} を求め , Snakes を実

行して目的の領域を抽出する．図 (3.4.2) にこの画像エネルギーを用いた対象物体の手順を示す．

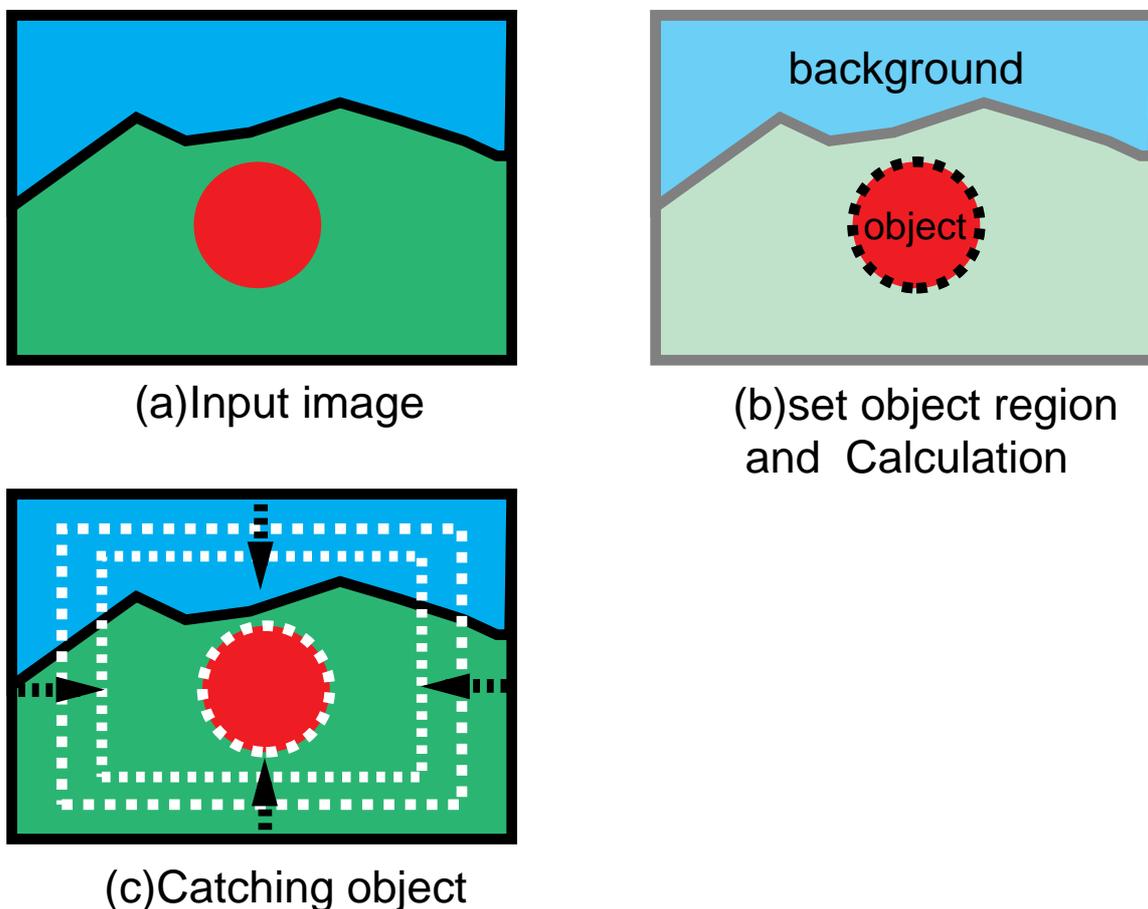


図 3.13: 新しい画像エネルギーを用いた追跡法:

(a) 入力画像; (b) 入力画像に対して抽出したい領域 (object) を設定する . object と background の領域の色情報より最適な λ の値を決定する; (c) 最適な λ の値を用いて E_{image} を求め Snakes を実行し目的の領域を得る .

3.5 まとめ

本章では , まず従来の Snakes を位相変化に対応させた C-snake について述べた . これにより複数物体の抽出が可能になった . 次にこの C-Snakes を用いた移動物体の追跡への

応用について述べた．最後に場景変動に対処する為の新しい画像エネルギーについて検討を行った．

第 4 章

場景変動を考慮した移動物体追跡の手法

本章では位相変化と色情報に対応した Snakes を用いて場景変動時の移動物体追跡を行う手法を提案する。

4.1 提案手法の流れ

場景変動を考慮した本手法での追跡の手順について図 (4.1) をもとに説明を行う。

- 手順 1: 追跡物体の指定

入力画像に対して追跡したい物体の領域を指定する。一つの場合は一つの領域を、複数の物体の場合はそれに対応した物体の領域を指定する。図 (4.1(b))

- 手順 2: α の計算

指定した領域の色情報をもとに $E_{intensity}$ と $E_{contrast}$ の最適なバランスを決める α の値を求める。この場合は一つの指定領域に対して1つの α を求める。複数の場合はそれぞれの領域に対して α を求める。図 (4.1(b))

- 手順 3: E_{image} の計算と Snakes による追跡

手順 2 より求めた α の値をもとに Snakes の E_{image} を計算する。それをもとに Snakes を実行する。オクルージョンが発生しない場合には C-Snakes と同様の手法で追跡を行う。図 (4.1(c))

- 手順 4: の計算 (2)

現フレームでの対象物体の抽出が完了したら, その抽出領域の色情報をもとに次のフレームで用いる の値を計算する. 手順 2 同様に一つの追跡領域に対して一つの を求める. 図 (4.1(d))(4.1(e))

- 手順 5: オクルージョン発生時の対応

オクルージョンが発生した場合は, C-Snakes での処理同様に輪郭線を統合するが, 场景変化することを前提にしているので統合後に追跡に問題が生じないように工夫する必要がある. (それに関しては後で述べる). 図 (4.1(f))

- 手順 6: 軌跡決定処理

本手法では対象物体を時系列画像全体を通して管理する. 各フレームでの追跡終了後に各対象の位置の記録と Snakes 自身が何を捉えてるのかの情報またその領域内の情報 (色情報) を管理してそれぞれの対象の軌跡を正確に決定する. 図 (4.1(g))

4.2 移動物体の追跡処理

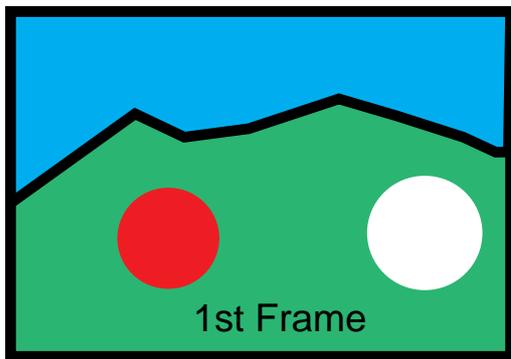
4.2.1 輪郭抽出・追跡

本手法で用いる Snakes のエネルギー関数

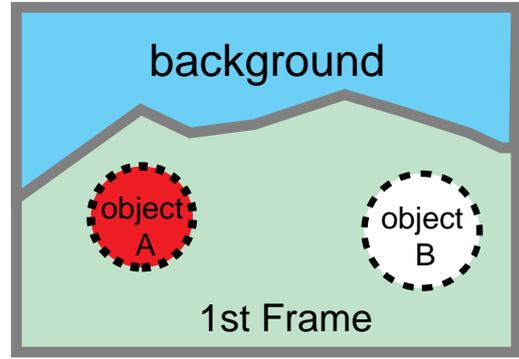
本手法では场景変動に対処する為に位相変化への対応と色情報を画像エネルギーとして扱う次式 (4.1) の Snakes を用いる.

$$E_{\text{snakes}} = E_{\text{spline}} + E_{\text{area}} + E_{\text{dist}} + \gamma E_{\text{contrast}} + (1 - \gamma) E_{\text{intensity}} \quad (4.1)$$

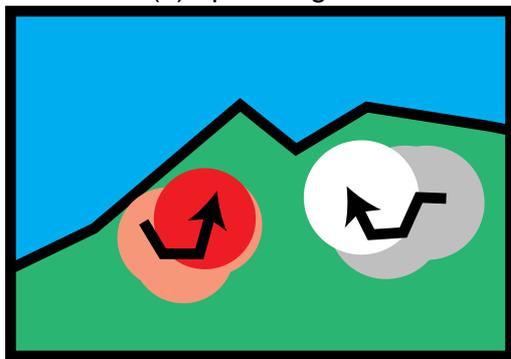
導入するエネルギー関数は内部エネルギー E_{int} ではスプライン項 (式 4.2), 面積項, (式 4.3), 格子点間平均化項 (式 4.4), 画像エネルギー E_{image} ではコントラスト項 (式 4.5) と明度項 (式 4.6) を用いる. なお は第 3 章で述べた処理により計算する.



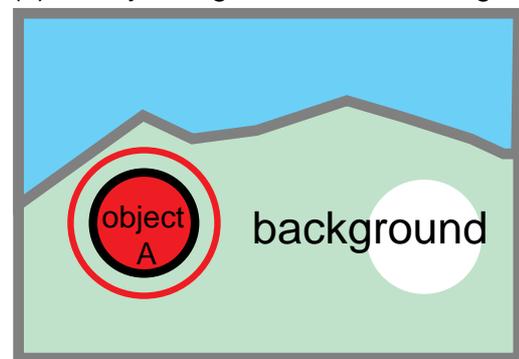
(a)Input image



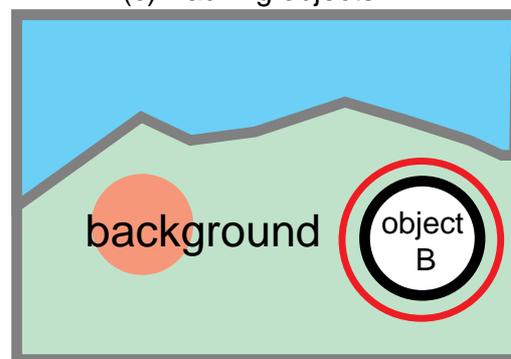
(b)set object regions and calculating



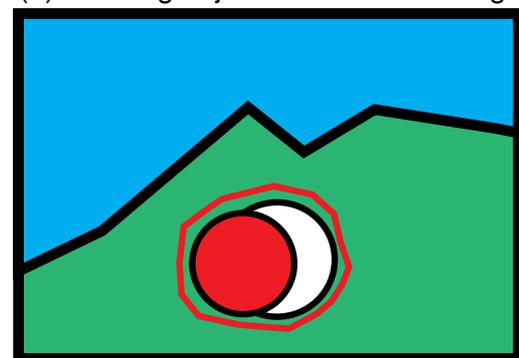
(c)Tracking objects



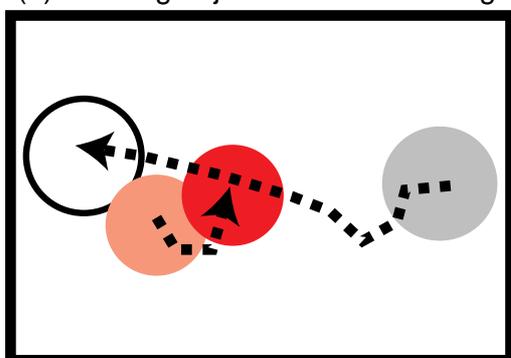
(d)Catching object A and Calculating



(e)Catching object Band Calculating



(f)Overlapping each other and unify two contours



(g)Deciding of trajectory

図 4.1: 本手法での追跡の流れ

$$E_{spline} = \frac{1}{2} \sum_{i=1}^n \{ \alpha \| \mathbf{v}'_i \|^2 + \beta \| \mathbf{v}''_i \|^2 \} \quad (4.2)$$

$$E_{area} = \frac{1}{2} \sum_{i=1}^n \{ w_{area} (x_i \mathbf{v}_y - \mathbf{v}_x y_i) \} \quad (4.3)$$

$$E_{dist} = \frac{1}{2} \sum_{i=1}^n \{ w_{dist} (d_{ave} - \| \mathbf{v}'_i \|^2) \} \quad (4.4)$$

$$E_{contrast} = \sum_{i=0}^m \ln^2 \{ (1 + r_i) / (1 + r_0) \} + \ln^2 \{ (1 + g_i) / (1 + g_0) \} + \ln^2 \{ (1 + b_i) / (1 + b_0) \} \quad (4.5)$$

$$E_{intensity} = \sum_{i=0}^m \ln \{ 1 + |(r_i - r_0) + (g_i - g_0) + (b_i - b_0)| \} \quad (4.6)$$

輪郭線上の離散点の扱い

Snakes は離散点によって構成されている . Snakes の離散点間の距離と一つの離散点移動する量は Snakes での抽出結果に影響を与える . 本研究での離散点の扱いは次のようにした .

- 離散点間の距離

離散点間の距離は E_{dist} によって隣接する離散点間の距離を均等に保つように働く . しかし輪郭線が膨張したり収縮することで離散点間の距離が狭くなったり広くなったりする . これを抑えるためにある隣接する離散点間の長さがある値より長ければ離散点間の中間に新しい離散点を生成し , 逆に短くなる場合は離散点を消滅させる . 本研究では離散点の間隔が 2 画素から 10 画素の間になるように調整した

- 離散点の移動

1 回の計算での離散点の移動距離を大きくとれば Snakes は高速に収束できるが , 抽出精度は粗くなる . 一方距離を小さくとれば抽出精度は良くなるが , 収束に時間がかかる . 速度と精度のトレードオフがある . 本研究では 1 回での移動距離を 2 画素または 3 画素に設定して計算させた .

- 計算のアルゴリズム

本研究では Snakes の計算を greedy algorithm[20] を用いて行った .

4.2.2 追跡領域の設定

提案した手法は初期フレームで人間が対象物体を指定することで、その対象を追跡する。その場合最初にその物体を指定する際にどのような領域で指定するかをきめる必要がある。本研究では次に挙げる 2 点から指定する領域を決定した。

- 領域指定する際に複雑な作業を必要としない。

最も理想的な指定領域は、対象物体を正確に切り出すことであるが、対象が複雑な図形の場合は指定が困難になる。よって指定する領域はなるべく単純な図形が好ましい。

- 対象物体以外の領域はなるべく含まない。

指定する領域はなるべく単純な図形が好ましいが、それと同時に指定した領域内には対象物体以外の領域は含まない方がよい。

本手法では対象物体を指定する際にその対象物体の内側にある最大の四角形の領域を指定領域とした。第 5 章で指定領域と Snakes による抽出結果との関係を調べる実験を行った。

4.2.3 閾値 の取扱い

提案した手法で用いる Snakes の画像エネルギーは $E_{intensity}$ と $E_{contrast}$ のバランス で決定される。しかしながら は抽出した領域の色情報から計算するので、Snakes 実行後でなければ決めることができない。本来第 n フレームを処理するのに必要な の値は第 n フレーム終了後でなければ求めることができないため 1 フレーム遅延することになる。そこで本研究では時間的分解能が十分な場合は、前後フレーム間でそれほど の値は変化しないという仮定で第 $n + 1$ フレームでの処理に第 n フレームで求めた を用いた。

4.3 時系列画像全体の対象物体の軌跡管理

本研究で用いる軌跡管理の方法について説明する。

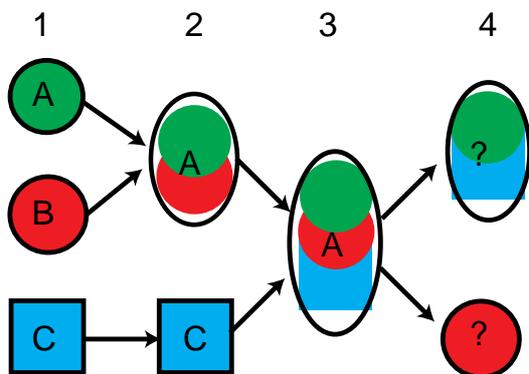


図 4.2: 従来手法 (直前でのみ管理)

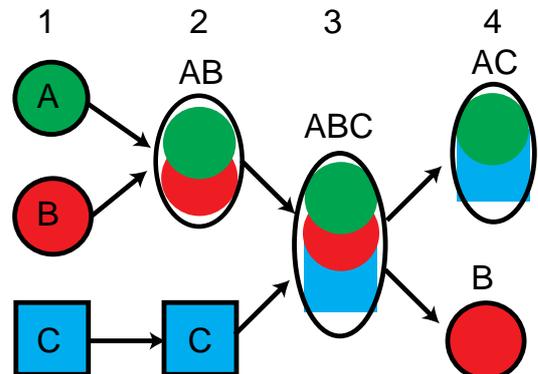


図 4.3: 提案手法 (全体を通して管理)

4.3.1 従来の軌跡管理の問題点

Snakes では 3 次元的な現象を 2 次元的な画像をもとに取り扱うため追跡において複数物体が見かけ上重なってオクルージョンが発生する場合がある。Snakes だけで追跡を行った場合にオクルージョン発生前と発生後の対象物体の対応づけが困難になる場合がある。従来提案されている手法ではオクルージョン発生直前のフレームでの領域情報 (色情報) だけを見て対応づけを行っていた為、オクルージョンが頻繁に発生するような状況では対応づけがうまく行かない場合が多々発生した。例えば図 (4.2) では直前の情報からのみ対応づけを行っている為、第 3 フレームで 3 つ同時に重なっている状態から第 4 フレームで分離する際に第 2 フレーム以前の対象物体の情報が無い為対応づけが行えない。そこで本研究では直前の情報だけで対応づけを行うのではなく、時系列画像全体で対象物体を管理する。図 (4.3) では輪郭線内にとらえてる対象物体を各フレーム毎に管理しているため第 2 フレーム以前の対象物体の統合情報も使える。それにより従来の手法でうまく行かなかった 3 つ同時に重なった時の対応づけが可能になる。

4.3.2 対象物体の軌跡管理のしくみ

追跡時に管理する情報は次の 3 つである。この 3 つの情報を時系列画像全体で管理することで対象物体の軌跡を正確に決定する。

- 各輪郭線内の対象物体の識別情報

対象物体を囲んでいる Snakes の輪郭線内にどの対象物体を含んでいるかの情報 .

- 各対象物体の領域情報 (色情報)

対象物体を最初に抽出した時の領域内の色情報 (平均画素値) .

- 各対象物体の各フレームでの位置情報

各対象物体が時系列画像中の各フレームでどの位置に存在していたかの情報 .

これらの情報を使った対応つけの手順について説明する . 領域内の属性を利用した軌跡の決定は以下の手順で行う .

- 手順 1

初期フレームで抽出した輪郭線に対してその領域の色情報を記録する . 対象物体の各フレームでの位置情報と一緒に記録する . 対象物体とは別に追跡対象を内包する輪郭線に対しても , それぞれどの対象を含むかを記録する . ただし初期フレームでは 1 つの輪郭線は 1 つの対象物体しか含まないものとする . 図 (4.3) 第 1 フレーム

- 手順 2

追跡途中で , 対象物体同士が重なり合っって輪郭線の統合が発生する場合は , その輪郭線同士がお互いにどの追跡対象の識別情報をもとに新しく統合された輪郭線に対してその情報を記録する . 例えば図 (4.3) 第 2 フレームでは A と B の対象物体が重なるがその際に新しくできた輪郭線に対して A と B が含まれてるという情報を対応つけて記録する .

- 手順 3

統合したあとに分離が発生した場合には , 輪郭線がどの対象を含んでいるかという識別情報を用いてその中に含まれる各対象物体同士の輪郭内の属性を使って単純類似度法 (4.7) を用いて決定する . 図 (4.3) 第 4 フレーム

- 手順 4

各フレームに対して手順 1 から 3 のいずれかを実行する.

なお単純類似度法は以下の式 (4.7) で表現される . 本研究では, 式 (4.7) 中の特徴ベクトルを \mathbf{a} , \mathbf{b} として , それぞれ対象物体の領域の RGB の平均輝度値を用いている .

$$S_s [\mathbf{a} , \mathbf{b}] = \frac{(\mathbf{a} , \mathbf{b})^2}{\|\mathbf{a}\|^2 \|\mathbf{b}\|^2} \quad (4.7)$$

第 5 章

評価実験

ここでは提案する場景変動を考慮した移動物体の追跡法の有効性を調べるために実験を行った結果を示す。実験には人工的に生成したシミュレーション画像とビデオカメラより撮影した実画像を用いた。

5.1 実験内容の説明

5.1.1 実験内容

1. 新しい画像エネルギーの有効性

新しい画像エネルギーが抽出の際にどの程度有効か示す。また指定領域の設定でどの程度抽出結果が変化するかを示す。

2. 軌跡管理導入の効果

提案した軌跡管理を用いることで対象が同時に 3 つ重なる場合でも対象の正しい軌跡を決定できることを示す。

3. 場景変動時の本手法での追跡結果

本研究で提案した手法をもとに場景変動時の追跡結果について示す。

5.1.2 実験に用いる画像

実験で用いた動画は、320x240 画素のカラー画像であり、DV カメラから取り込んで処理を行った。

5.2 色情報によって抽出した結果

ここでは本手法で用いた新しい画像エネルギーを導入した Snakes を用いてシミュレーション画像および実画像に適用した結果を示す。図 (5.1) では人工的に作ったシミュレーション画像を生成し、それに対して Snakes を用いて抽出した。このような抽出領域に色ムラのない画像に対してはうまく抽出することができている。次に比較的同じ色で構成されている物体 (5.2) の実画像を用いて抽出を行った。この場合も瓶の先のほうが削れているが良好な抽出が行われている。次に医療画像 (図 (5.3)) を用いて病変部分の抽出実験を行った。病変部分があまり一様ではないので実際の病変部分の位置からずれているが、その部分を含むようには抽出できている。このように実画像の場合でも比較的一様な対象に対してはうまく抽出が行われている。

5.3 指定領域の設定と抽出結果との関係

本手法では最初のフレームで追跡すべき物体を指定しその指定した領域の色情報を用いて対象を抽出する。しかしながら人が領域を指定する際にはなるべく複雑な操作を行わないで領域を指定することが望ましい。そこで対象物体の領域指定と Snakes の抽出結果の関係を調べるために実画像を用いて実験を行った結果を示す。図 (5.4) 図 (5.7) 図 (5.10) は、それぞれ対象物体を含む四角形を指定、対象物体を正確に指定、対象の一部を指定の3つで実験を行った。また図 (5.5) 図 (5.8) 図 (5.11) は、それぞれその領域を指定した場合に得る Snakes の画像エネルギー E_{image} を示したもので、色が黒い程エネルギーは小さいことを意味する。図 (5.4) のような領域を指定した場合には、Snakes での抽出がうまく行われていない (図 (5.6))。これは対象領域以外の背景も指定領域に含んでしまっているため、Snakes の画像エネルギーが誤った領域も含んで計算を行った為うまく抽出がされていない。図 (5.7) のように対象領域を正確に指定した場合は、Snakes による抽出もうまく行われている (図 (??))。また図 (5.10) のように領域を指定した場合も、正確に

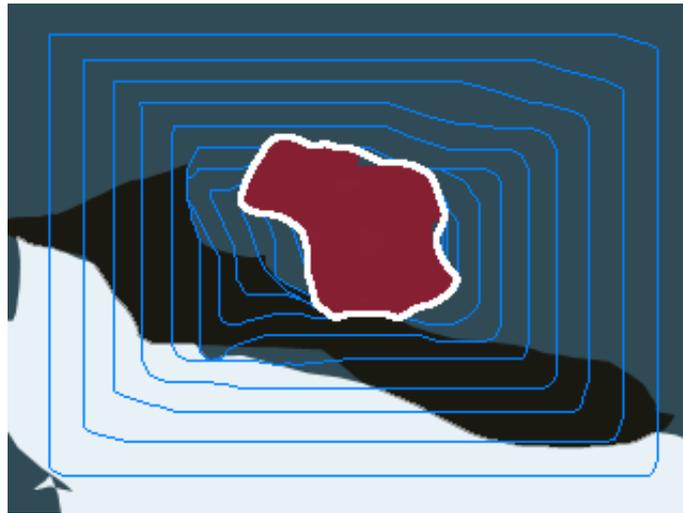


図 5.1: シミュレーション画像

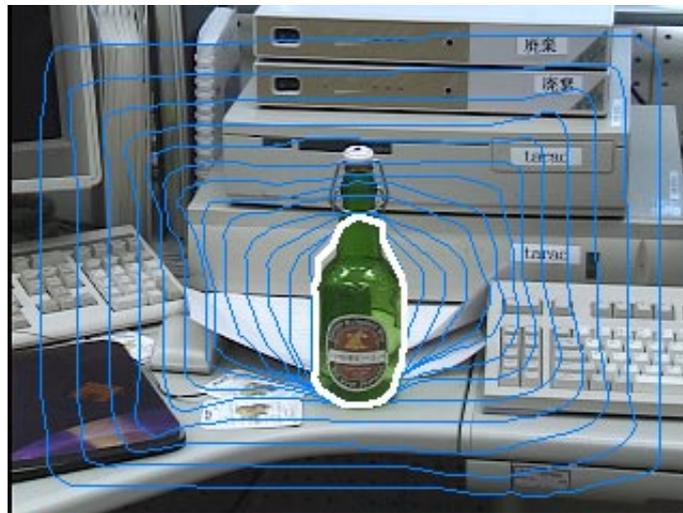


図 5.2: 比較的一様な色の物体 (ボトル)

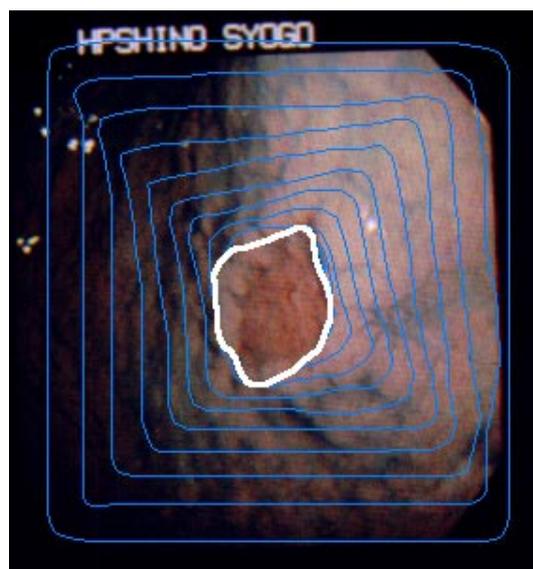


図 5.3: 臓器の病変部分の抽出

領域を指定した場合と同様に Snakes による抽出がうまく行えている (図 (??)) . 領域を指定する際には図 (5.4) のような領域の指定の仕方が、もしくは図 (5.10) のような指定の仕方をした場合が操作としては最も簡単であるが、これらの実験より対象の一部を指定するやりかたでも良好な抽出が行えることがわかる .

5.4 軌跡管理を導入の有効性

ここでは本研究で提案した軌跡管理を実際にシミュレーション画像を用いて実験した (図 5.13) 各物体は図 5.13 のように動く . まず第 1 フレームで、それぞれの対象を Snakes がとらえている . 次に第 11 フレームで、赤と緑の円が接近して輪郭線が統合されている . 第 13 フレームでは 3 物体とも重なってに 1 つの輪郭線に統合されている . 第 21 フレームで 3 つ同時に重なっていた物体が 2 つに分離する . この際従来の手法では第 11 フレーム以前の情報 (赤と緑が統合しているという情報) が使えないことから第 21 フレームでの対応つけができなかったが、提案手法では最終的に 3 つに分離する第 25 フレームまで正確に対象物体の軌跡を決定していることがわかる .

5.5 场景変動時 (カメラが動くような画像) における追跡結果

ここでは本研究での最終目標である场景変動が起こる状態での移動物体の追跡実験を行う . まず最初にシミュレーション画像を用いて、背景の雲模様が各フレームでランダムに変わる状態での追跡を図 (5.17) のように行った . まず第 1 フレームで Definition Region のように領域を指定した . そこで指定した領域の色情報をもとに追跡を開始した . 抽出結果を見てみると第 1 フレームで目的の対象がうまく抽出されていることがわかる . その後、背景の変化する各フレームでもうまく抽出されており、その後第 7, 第 10, 第 14 フレームでも同様にうまく抽出できている . このように一様な色領域をもつシミュレーション画像を用いた場合は比較的良好に追跡が行えた .

次に実画像を用いた実験を行った . カメラが横に連続的に動く中で対象物体も同時に動くようなシーン (図 (5.15)) を用いて実験した . ここでカメラの動きは右から左に動いてゆき、その中に写っているチンパンジーも右から左に動いていく . まず初期フレームで (図 (5.15)) の Definition Region のように対象の領域を指定した . 初期フレームでは対象物体の抽出が比較的うまく行えている . また第 6 フレームと第 12 フレームでは輪郭線が対象



圖 5.4: 指定領域

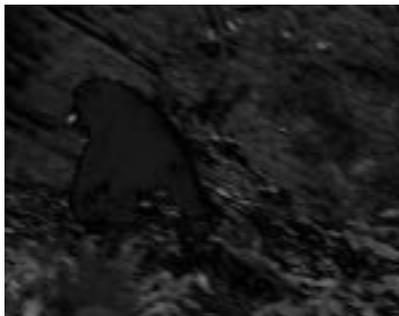


圖 5.5: E_{image}



圖 5.6: 輪廓抽出結果



圖 5.7: 指定領域



圖 5.8: E_{image}



圖 5.9: 輪廓抽出結果



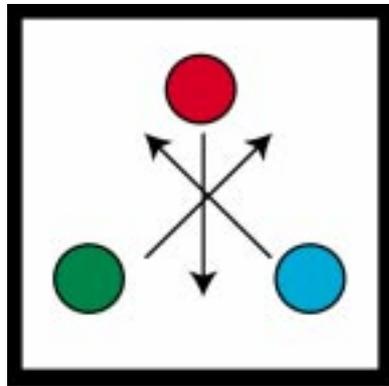
圖 5.10: 指定領域



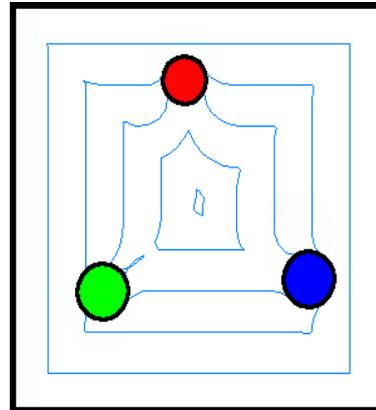
圖 5.11: E_{image}



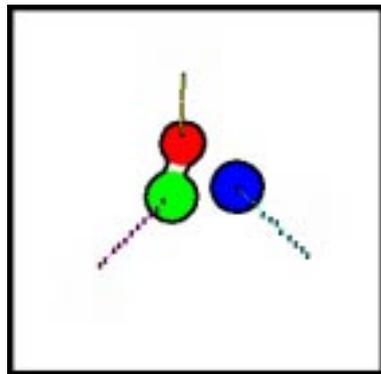
圖 5.12: 輪廓抽出結果



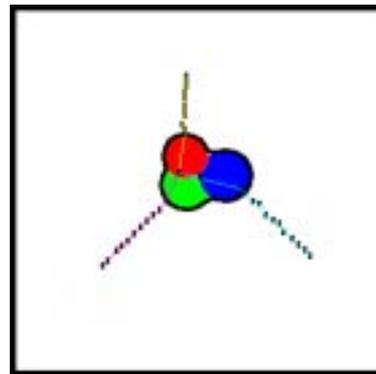
Direction of Moving Object



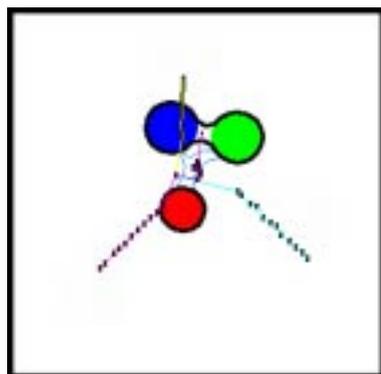
FRAME1



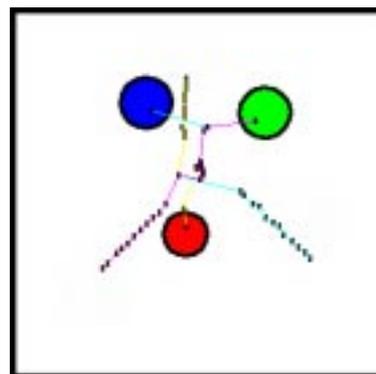
FRAME11



FRAME13



FRAME21



FRAME25

図 5.13: 提案手法での軌跡決定の結果

の方に入り込めずにいるが、これはチンパンジーの色とチンパンジーの影の色がほぼ同じであるために Snakes がそれにとらえられた為であると考えられる。また本手法では 1 フレーム前の E_{image} を用いて画像エネルギー E_{image} を求めているが、この場合は追跡を行う上で問題にはなっていない。

また図 (5.16) ではスノーボードに乗って移動している人間を赤い線で描いたカメラワークで追跡しながら撮影している。チンパンジーの時と同じように Definition Region を初期フレームで設定する。この場合追跡する人物の上着の一部を指定した。まず第 1 フレームでは上着の部分全体が比較的良好に抽出できている。このまま第 11 フレームまで良好に追跡が行えていることがこの結果からわかる。以上のように対象の一部を指定してその領域の色エネルギーを Snakes で用いることで比較的良好な追跡が行えた。

最後に図 (5.17) では歩行する人物を背景が変動する状態で追跡する実験を行った。歩行者の服の一部を指定してそれをもとに追跡を行った。スノーボードの時と同様に比較的良好に追跡が行えていることがわかる。

なおこの処理を行うの pentiumII プロセッサ 300Mhz を載せたコンピュータで初期フレームでの抽出までにおよそ 6 秒、それ以降のフレームでの処理には、平均で約 3 秒を要した。

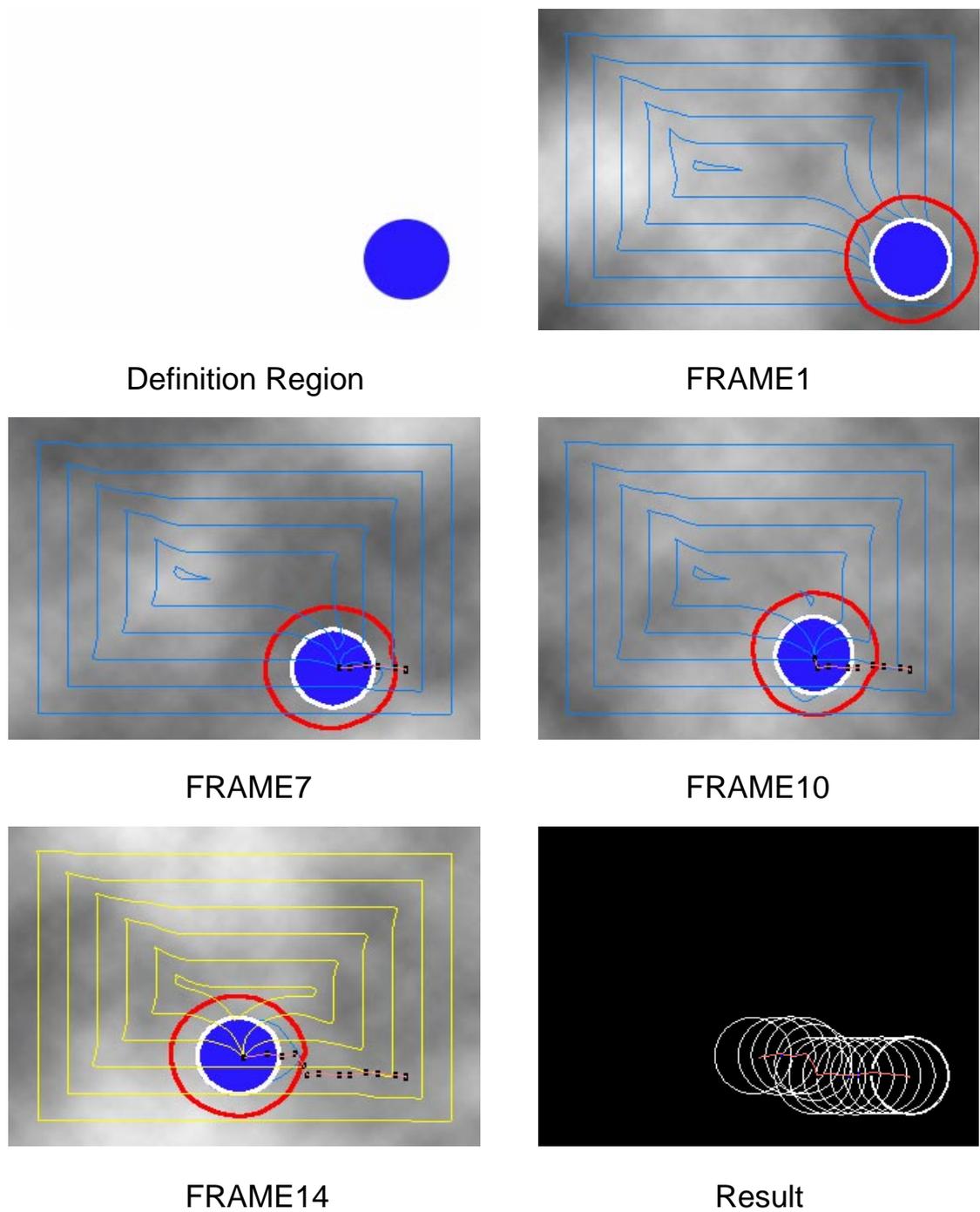


図 5.14: シミュレーション画像を用いた実験結果 (simulation): 背景の雲模様がランダムに変わる中で一様に青い円を動かして追跡。白線は収束した輪郭線, 赤線は収束した輪郭線を膨張させた次フレームでの初期輪郭を示す。



Camera Work



Definition Region



Result 1FRAME *E image*



Result 6FRAME *E image*



Result 12FRAME *E image*



図 5.15: 実画像を用いた実験結果 (chimpanzee): 白線は収束した輪郭線, 赤線は前フレームで収束した輪郭線を膨張させた現フレームでの初期輪郭を示す



Camera Work

Definition Region



Result

1FRAME



E image



Result

7FRAME



E image



Result

11FRAME



E image

図 5.16: 実画像を用いた実験結果 (Snowbord):白線は収束した輪郭線, 赤線は前フレームで収束した輪郭線を膨張させた現フレームでの初期輪郭を示す



図 5.17: 場景変動時の歩行者の追跡実験:歩行者の服の一部を指定して視点が変動していく状態で追跡を行った。白線は収束した輪郭線,赤線は前フレームで収束した輪郭線を膨張させた現フレームでの初期輪郭を示す

第 6 章

結論

本論文では従来手法では対応が困難である場景変動中の移動物体の追跡に対処する手法を提案した．また時系列画像全体で対象物体の軌跡を管理することで，頻繁にオクルージョンがおこる状況での軌跡の決定が困難という問題に対処する軌跡管理の手法を提案した．

以下に本研究で提案した手法をまとめる．

- 場景変動時の移動物体の追跡法

従来の Snakes を用いた移動物体の追跡方法では，背景との差分画像や連続する前後フレームとの差分画像などに依存する画像エネルギーを用いていたため，カメラの移動や背景が変化するような状況では移動物体の追跡を行うことが困難であった．しかしながら K . P . Ngoi らが提案した画像エネルギー項を Snakes に新たに導入することで，対象の色情報を用いて，場景変動がおこるような状況でも動画像中からの移動物体の追跡が可能になった．

- 頻繁にオクルージョンが発生する場合の軌跡管理

従来手法ではオクルージョン発生の前後の情報のみで軌跡の決定を行っていたため，頻繁にオクルージョンが起こる状態では軌跡の決定が困難になる場合があった．提案手法では各対象の統合状態を追跡全体を通して管理することで軌跡決定を正確に行えるようになった．

しかしながら本研究で提案した移動物体追跡法にもいくつかの課題がある。

- 背景に追跡対象と同じ色が入って来た場合

本研究で提案した手法は，対象物体の色に依存した追跡を行う為に追跡対象の初期輪郭の中に背景として同じような色の物体が入って来た場合，その領域も誤って抽出してしまう．この場合はテクスチャー等の他の差分情報に依存しない特徴を E_{image} に導入することで対処できると考えられる．

- 追跡している対象同士が同じ色の場合

追跡対象が複数の場合にオクルージョンが発生することがあるが，その場合に対象同士が同じ色の場合に，再びお互いが離れた時にオクルージョン発生前の対象を決定することができない．これに対処する方法の一つは本研究で提案した軌跡決定の手法を拡張しいくつかの Snakes を群として管理することで対処が可能と考えられる．例えば上着とズボンのようにお互いの領域が離れないようなもの追跡対象に対して複数設定させることでオクルージョンで重なる相手の色と群の一つの Snakes の領域が同じでも群の残りの Snakes 領域の色の違いで軌跡を決定することが可能であると考えられる．

- 自動化との兼ね合い

従来 of 场景変動を考慮してない状況での追跡は領域内の移動物体を差分情報により自動で検出することができていた．しかし本研究で提案した手法では，差分に基づくエネルギーを用いていない為追跡開始時に追跡対象を手動で指定する必要がある．自動化という兼ね合いから考えるとこの手法では改善の余地がある．

謝辞

本研究を進めるにあたり、終始暖かく御指導を頂いた北陸先端科学技術大学院大学 マルチメディア統合システム講座 阿部 亨 助教授に厚く御礼申し上げます。

また、さまざまな面で御教授頂いた北陸先端科学技術大学院大学 マルチメディア統合システム講座 堀口 進 教授に深く感謝致します。

本研究に際して、審査、御助言を頂いた 北陸先端科学技術大学院大学 像情報処理学講座 小谷 一孔 助教授に深く感謝致します。

またサブテーマで熱心に御指導を頂いた北陸先端科学技術大学院大学 ソフトウェア基礎講座 片山 卓也 教授に深く感謝致します。

また、日頃よりお世話になった当研究室の皆様にも心より感謝致します。

参考文献

- [1] M.Kass, A.Witkin, D.Terzopoulos, Snakes : Active Contour Model, International Journal of Computer Vision, pp.321-331, 1988.
- [2] 荒木 昭一, 横矢 直和, 岩佐 英彦, 竹村 治雄, 交差判定により分裂する動的輪郭モデルとその応用, 信学技報, pp.1-8, 1995.7.
- [3] C.Vieren, F.Cabestaing, J.Postaire, Catching moving objects with Snakes for motion tracking, Pattern recognition Letters 16, pp.678-685, 1995.
- [4] 金田 丘, 動的輪郭モデルを用いた多方向に移動する複数物体の追跡に関する研究, 北陸先端科学技術大学院大学 修士論文1997.
- [5] 寺久保 敏, 横矢 直和, 岩佐 英彦, 竹村 治雄, 2 段階アフィン変換を利用した移動カメラ画像からの動物体の抽出, 信学技報, pp.31-pp.36,1995.12.
- [6] 平山 泰崇, 中村 納 表情変化にロバストな新しい個人識別方式 信学技報, pp.25-pp.32,1997.11.
- [7] 徐 剛, 杉本 岳夫, カラー画像を用いた実時間顔検出追跡システム 信学技報, pp.9-pp.16,1997.11.
- [8] 栄藤 稔, 白井 良明, 色, 位置, 輝度こう配に基づく領域分割による 2 次元動き推定 電子情報通信学会論文誌, Vol.J76-D-II No.11 pp.2324-pp.2332,1993.11.
- [9] M.Tabb, N.Ahuja, A Multiscale Region-Based Approach to Image matching, Trans. on Patt. Anal. Mach. Intell., pp.415-419, 1994.

- [10] D.Koller, J.Weber, J.Malik, Robust Multiple Car Tracking with Occlusion reasoning, In Proc. Third European Conference on Computer Vision, pp.189-196, Stockholm Sweden, May 2-6, 1994.
- [11] D.Koller, J.Weber, J.Malik, Machine Vision Based Surveillance System for California Roads PATH project MOU-83 Final Report, Computer Science Division University of California.
- [12] 橋本 晶寛, 木下 宏揚, 酒井 善則, Sampled Active Contour Model による輪郭抽出法, 電子情報通信学会論文誌 D-II, Vol j77-D-II, No.11, pp.2171-2178, 1994.11.
- [13] 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.
- [14] 古川 亮, 今井 正和, 鳥野 武, 時空間画像を利用した弾性輪郭モデルとその収束方法 電子情報通信学会論文誌 D-II, Vol j79-D-II, No.6, pp.1054-1063, 1996.6.
- [15] 阪口 俊文, 大山 公一, 面積項を持つスネーク, 電子情報通信学会 春季大会講演論文集, D-555, 1991.
- [16] 飯島 泰蔵, パターン認識理論, 基礎情報工学シリーズ 6, 森北出版, pp.111-129, 1989.
- [17] Jim Ivins, John Porrill Active region models for segmenting textures and colours, Image and Vision Computing, Volume 13 Number 5, pp.431-438, 1995.
- [18] K.P.Ngoi, J.C.Jia, A new colour image energy for active contours in natural scenes Pattern Recognition Letters 17, pp.1271-1277, 1996.
- [19] Ohta, Y Knowledge-based Interpretation of Outdoor Natural Color Scenes, Pitman Publishing Inc., Cambridge, MA.
- [20] DONNA J.Williams, MUBARAK SAHA, A Fast Algorithm for Active Contours and Curvature Estimation, CVGIP:IMAGE UNDERSTANDING, Vol.55, No.1, January, pp.14-26, 1992.

研究業績

- 1 追跡機能を統合した動的輪郭モデルによる移動物体の追跡
副島義貴, 阿部 亨, 堀口 進
電気関係学会北陸支部連合大会, F-50, pp359, Nov.1996.