

〔公 開〕

TR-C-0108

S n a k e s による

複雑な輪郭の追跡

畑 幸一
Kouichi HATA

大谷 淳
Jun OHYA

岸野 文郎
Fumio KISHINO

1 9 9 5 3 . 1 2

A T R 通信システム研究所

Snakesによる複雑な輪郭の追跡

畑 幸一 大谷 淳 岸野 文郎

1995年3月12日

第I部

Survey of Snakes: Active Contour Models

要旨

現在コンピュータビジョンの分野で重要な技術の一つとなっている Active Contour Models とその関連技術について、Kass らの Snakes を基準にまとめている。Kass らの Snakes で、主に問題となっている初期値、収束性、パラメータ、外部拘束力、ポテンシャル画像について異なるアプローチがなされたものや Snakes を拡張したもの、またコンピュータビジョンの問題に適用されたものについて述べている。

1 はじめに

コンピュータビジョンや画像処理の分野では、3次元実世界におけるカメラからの物体までの距離の違いや物体表面の不連続、物体表面のテクスチャなどによりあらわれる画像上の明るさの不連続な部分をエッジと呼んでいる。それらエッジの一部分によって、物体の形状の輪郭が構成されている。物体の形状の輪郭は、我々がそれらから、3次元形状に関する情報を得ることができるように、画像中の2次元的存在であるにも関わらず、3次元形状を顕著に表している。[1]。そのような観点からも、画像中のパターン認識や3次元計測において、対象物体の輪郭の抽出は重要な処理となっており、対象物体の正確な輪郭が得ることができれば、対象物体の形状獲得に有効だけでなく、画像合成、医療用画像解析等に役立つ情報を提供することができる[2][3]。

コンピュータビジョンの分野では、エッジ抽出、ステ

レオマッチングや動き追跡等の低レベルな視覚機能を実現するために、低レベルの情報をそのままボトムアップしていくアプローチが多くとられているが、そのようなアプローチでは、低レベルの情報の間違いが後のより高レベルな処理に危機的な間違いを引き起こしてしまう[4]。そこで、より高レベルな処理に低レベルな情報を引き渡す前に、高レベルな処理に有効な記述に換えてやる必要がある。この変換を最適化問題に帰着させたモデルが、Kass らによって提案された、Snakes:Active Contour Models[4]である。

一般に Active Contour Models は、パラメータで表された変形可能なモデルと画像そのものや画像から得られる特徴との照合度を何らかの評価関数で表し、パラメータを変化させてその評価関数を最小化させることでモデルと画像をマッチングさせる手法の、特殊化されたものの一つであると考えられている。Active Contour Models は、モデルとして長さや曲がりによる内部エネルギーを持つスプラインを持ち、外部からの拘束力や、特徴点へ引き付ける画像からの力によって影響を受け、評価関数であるエネルギー関数が最小化される。ユーザは外部拘束力や画像からの力をさまざまに定義することにより、その振舞いを制御することができ、一部にユーザの意図を含んだエッジ抽出、ライン抽出、主観的な輪郭の抽出、追跡、ステレオマッチングなどに利用される。

この報告書では、Kass らによる Active Contour Models を軸に、それに関連する文献をそれと際違った違いに基いてまとめる。Kass らの Snakes は、Tezorpoulos らの Deformable Models の一部を特殊化したものであるが、それらも Kass らの拡張として Kass らを基準にまとめてある。表1に、取り上げた文献とその文献の Kass らによるものとの違った特徴を示す。他の文献で多くその内容を研究され、Kass らと異なった提案がされている部分を七つあげ、○印によって各文献がどの部分にお

年代	著者	初期値	パラ メータ	収束性	ポテン シャル	外部力	拡張	適用	ターゲット
1986年	Terzopoulos ら [5]						○		
1987年	Terzopoulos ら [6]						○		
1987年	Terzopoulos ら [7]		○				○		パイパイとポテトの絵
1988年	Terzopoulos ら [8]						○		指、ピカソの絵、ポテトの絵
1988年	Zucker ら [9]	○			○				指紋
1988年	Kass ら [4]	—	—	—	—	—	—	—	唇の輪郭、パイパイ
1989年	Yuille ら [10]						○		目、唇
1989年	Staib ら [11]						○		脳 (NMR)、超音波心臓
1990年	Amini ら [12]			○					コーラ缶、葉、マウス
1990年	Williams ら [13]			○					ボトル、コップ
1990年	Berger ら [14]	○							手、スタンドライト
1990年	David ら [15]	○	○		○				指紋
1990年	Menet ら [16]			○	○				ビルディングトップ
1990年	Cohen ら [17]	○		○	○	○			超音波心臓
1990年	Berger ら [18]	○							angiographics
1990年	Fua ら [19]	○	○						道路、ビル
1990年	坂口ら [20]		○						ボトル、複数の顔
1991年	Cohen ら [21]	○		○	○	○			超音波心臓
1991年	Samadani [22]		○						CG、オーロラ
1991年	上田ら [23]				○	○			駐車場の車、本立て
1992年	Basclé ら [24]	○					○		室内シーン、テーブル
1992年	Leymarie ら [25]							○	CG のスケルトン抽出
1992年	Couvignou ら [26]							○	本、傘
1992年	栄藤ら [2]				○	○			人物の輪郭
1992年	藤村ら [27]	○							なめくじ
1992年	福井ら [28]			○	○				顔
1993年	Leymaire ら [29]	○	○	○	○				白血球
1993年	Ferrie ら [30]							○	人形
1993年	Cohen ら [31]	○		○	○	○	○		超音波心臓
1993年	Couvignou ら [32]							○	—
1993年	Blake ら [33]						○		手
1993年	Xu ら [34]	○				○			積木、プリンタ
1993年	Basclé ら [35]						○	○	室内シーン、長靴、テーブル
1993年	天野ら [36]					○			歩行者、MRI心臓
1994年	Neuenschwander ら [37]	○							道路の航空写真、CG
1994年	Henricsson ら [38]	○							ビーナス像、船、とうがらし
1994年	Curwen ら [39]						○	○	花瓶、箱
1994年	Chakraborty ら [40]						○		超音波心臓
1994年	Lai ら [41]						○		手書き文字、鍵
1994年	Basclé ら [42]						○		車、腕、顔
1994年	Cipolla ら [43]							○	手
1994年	福井ら [44]				○	○			顔
1994年	橋本ら [3]				○	○			ボトル、ぬいぐるみ
1994年	梅山ら [45]						○		スタンプ、ハンマー、CG

表 1: 各文献の特徴

いて異っているかを示している。まず,2章では Kass らによる Active Contour Models をある程度詳細に述べる。3章, 4章,5章,6章,7章では, それぞれ初期値, パラメータ, 収束性, ポテンシャル画像, 外部力について述べている文献をまとめている。8章では, Kass らとはアプローチがかなり違っているが近い技術としていくつかの技術を挙げている。9章では Active Contour Models を使った研究について触れている。

2 Kass らによる Active Contour Models

Active Contour Models[4] は, スプライン曲線の長さや曲がりによって定義された内部エネルギー, ユーザや他のメカニズムによって定義される外部拘束力, エッジやラインといった画像から得られる特徴に引き付ける画像力によってその全体のエネルギーが定義される。それがスプライン曲線と画像の特徴との適合度の評価関数となっており, 最小化されることによってスプライン曲線を目的の特徴に適合させる。

2.1 スネークのエネルギー

空間的なパラメータ s を用いて, スプライン曲線の位置 $\mathbf{v}(s)$ を, x 座標と y 座標で表す。

$$\mathbf{v}(s) = (x(s), y(s)) \quad (1)$$

これより, スネークのエネルギーは以下のように定義できる。

$$\begin{aligned} E_{snake}^* &= \int_0^1 E_{snake}(\mathbf{v}(s)) ds \quad (2) \\ &= \int_0^1 E_{int}(\mathbf{v}(s)) ds + \int_0^1 E_{image}(\mathbf{v}(s)) ds \\ &\quad + \int_0^1 E_{ext}(\mathbf{v}(s)) ds \quad (3) \end{aligned}$$

ここで, E_{int} はスプライン曲線の長さ, 曲がりなどによる内部エネルギーで, E_{ext} は外部拘束力によるエネルギー, E_{image} は画像からの力によるエネルギーである。つまり, 各エネルギーすべての和が全体のエネルギーとなっている。

2.2 内部エネルギー

内部エネルギーは, スプライン曲線の一次微分の項と二次微分の項を用いて以下のように表せる。

$$E_{int}(\mathbf{v}(s)) = \frac{\alpha(s)|\mathbf{v}_s(s)|^2 + \beta(s)|\mathbf{v}_{ss}(s)|^2}{2} \quad (4)$$

ここで, $\alpha(s)$ と $\beta(s)$ はそれぞれの項の重みを制御する係数である。一次の項によって ACM は膜のようなものが常に縮もうとする動きをする。また, 二次の項によって薄い板のようなものが常にまっすぐになろうと振舞う。もちろん, 重み係数 $\alpha(s)$ と $\beta(s)$ によってそれぞれの振舞いの相対的な強弱を調節することができ, $\beta(s)$ をある点で 0 にすれば, その点で輪郭のコーナなどの二次の不連続にも対応できる。これらのパラメータについて, Samadani[22] や坂口ら [20] が研究を行なっているがそれらについては 4章で述べる。

2.3 外部拘束力によるエネルギー

外部拘束力は, Kass らによって二つの種類が提案されている。一つは ACM の一点と引き付けたい一点とをバネのように結ぶもので, ACM 上の一点を \mathbf{x}_1 , 引き付けたい点を \mathbf{x}_2 とすると,

$$E_{spr} = -k|\mathbf{x}_1 - \mathbf{x}_2|^2 \quad (5)$$

と表す。もう一つは ACM をはねつけたい部分に, エネルギーの山 (Kass らは火山と呼んだ) を置いて ACM を遠ざけるもので, 山の位置を \mathbf{x}_{vol} とすると,

$$E_{vol} = -\frac{k}{|\mathbf{x} - \mathbf{x}_{vol}|^2} \quad (6)$$

のように定義される。これは, 画像からの力によるエネルギーとして扱われることもある。この外部拘束力をさまざまに定義することによって, スネークの振舞いを制御できる。Kass らはこの外部拘束力をユーザとのインタラクションのために用いたが, 画像から引き出せる情報や物理的拘束条件を用いて, 人間とのインタラクションを減らす試みもされている。これらの力について数多くされている提案については 7章で述べる。

2.4 画像からの力によるエネルギー

Kass らは, 画像の特徴に ACM を引き付けるために, 以下のような三種類の画像からの力によるエネルギーを定

義した。

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term} \quad (7)$$

まず、画像の特徴として一番簡単なものは画像の輝度 $I(x, y)$ である。

$$E_{line} = I(x, y) \quad (8)$$

重み関数である w_{line} の符合をかえることで、ACM を明るい部分に引き寄せるようにしたり、暗い部分に引き寄せるようにしたりできる。第二に、次のようなエネルギーを考えると、ACM を画像中のエッジに引き寄せることができる。

$$E_{edge} = -|\nabla I(x, y)|^2 \quad (9)$$

さらに、ローカルミニマムに落ちいるのを避けるために、画像にガウシアンフィルタをかけたものを用いることが多い。

$$E_{edge} = -(G_\sigma * \nabla^2 I)^2 \quad (10)$$

ガウシアン分散 σ を大きくすれば大まかなエッジにかなりはなれた ACM を引き付けることができ、小さくすれば近くの ACM を細かいエッジに引き付けることができる。第三の特徴として、Kassらはラインセグメントの端点やコーナといったものをあげている。画像をスムージングするためガウシアンをかけた $C(x, y) = G_\sigma(x, y) * I(x, y)$ をもちい、そのエネルギーを以下のように定義している。ここで $C_x = \frac{\partial C}{\partial x}$ である。

$$E_{term} = \frac{C_{yy}C_x^2 - 2C_{xy}C_xC_y + C_{xy}C_y^2}{(C_x^2 + C_y^2)^{\frac{3}{2}}} \quad (11)$$

これらのうち、一部もしくは全部のエネルギーを用いユーザは ACM を目的の画像の特徴へ引き付けることができる。これらについても、数々の研究がなされ異なる手法が提案されている。それらについては6章で述べる。

2.5 ダイナミクス

Kassらは、画像からのエネルギーと外部拘束力からのエネルギーを合わせて $E_{ext} = E_{image} + E_{con}$ としたものを最小化するため、

$$\min E_{ACM} = \min \left(\int E_{int}(\mathbf{v}(s))ds + \int E_{ext}(\mathbf{v}(s))ds \right) \quad (12)$$

オイラーラグランジュ方程式を用いた。 $\alpha(s) = \alpha, \beta(s) = \beta$ とそれぞれ定数にして、以下のオイラーの方程式を導きだす。

$$\alpha x_{,ss} + \beta x_{,ssss} + \frac{\partial E_{ext}}{\partial x} = 0 \quad (13)$$

$$\alpha y_{,ss} + \beta y_{,ssss} + \frac{\partial E_{ext}}{\partial y} = 0 \quad (14)$$

Kassらは、オイラーの方程式を差分近似法を用いて解を求めた。 $\mathbf{v}(s)$ を $\mathbf{v}_i = (x_i, y_i) = (x(ih), y(ih))$ と、ステップ幅 h で離散化し、内部エネルギーを以下のように表した。

$$E_{int}(i) = \frac{\alpha_i |\mathbf{v}_i - \mathbf{v}_{i-1}|^2}{2h^2} + \frac{\beta_i |\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}|^2}{2h^4} \quad (15)$$

これにより、オイラーの方程式は以下のようになる。

$$\begin{aligned} & \alpha_i(\mathbf{v}_i - \mathbf{v}_{i-1}) - \alpha_{i+1}(\mathbf{v}_{i+1} - \mathbf{v}_i) \\ & + \beta_{i-1}(\mathbf{v}_{i-2} - 2\mathbf{v}_{i-1} + \mathbf{v}_i) \\ & - 2\beta_i(\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}) \\ & + \beta_{i+1}(\mathbf{v}_i - 2\mathbf{v}_{i+1} + \mathbf{v}_{i+2}) \\ & + (f_x(i), f_y(i)) = 0 \end{aligned} \quad (16)$$

これらは、以下のようにマトリックス的に書ける。

$$Ax + f_x(x, y) = 0 \quad (17)$$

$$Ay + f_y(x, y) = 0 \quad (18)$$

これを反復法を用いて解くと、ステップ幅を γ として、

$$Ax_i + f_x(x_{i-1}, y_{i-1}) = -\gamma(x_i - x_{i-1}) \quad (19)$$

$$Ay_i + f_y(x_{i-1}, y_{i-1}) = -\gamma(y_i - y_{i-1}) \quad (20)$$

$$x_i = (A + \gamma I)^{-1}(x_{i-1} - f_x(x_{i-1}, y_{i-1})) \quad (21)$$

$$y_i = (A + \gamma I)^{-1}(y_{i-1} - f_y(x_{i-1}, y_{i-1})) \quad (22)$$

行列 $A + \gamma I$ は、対角バンド行列なので LU 分解を用いて、逆行列が計算でき、それにより解が与えられる。しかし、ステップ幅や逆行列によっては解が安定しないことがある。これに対する他のアプローチとしては、Cohenら [17][21][31] の FEM を用いたものや Aminiら [12] の DP を用いたものがあるが、これらについては5章で述べる。

2.6 初期値

Kass らの方法では, スネークが望まないローカルミニマムに陥らないため, 目的の特徴に対するかなり正確なスネークの初期位置を仮定しなければならない. この仮定をなくす, もしくは緩めるため数々の提案があるが, これらについては3章で述べる.

3 初期値について

Kass らの方法では, スネークが望まないローカルミニマムに陥らないため, 望ましい特徴に対するかなり正確なスネークの初期位置を仮定し, 設定してやらなければならない. この仮定はスネークが追跡等に用いられる場合は妥当と考えられるが, 複雑な物体の輪郭の抽出を目標とする場合にはかなり困難な仮定である. 画像からの力をガウシアンによって広域に及ぼすようにし, 遠くのスネークを引き付けても良いが, 細かな特徴が失われてしまう. これらに対するアプローチを述べる.

3.1 Zucker らのアプローチ

Zucker ら [9] と David ら [15] はタンジェントフィールドからのエッジ抽出を行なっている. 画像の局所的な記述であるタンジェントフィールドは, 曲線の位置, 方向, 曲率の情報を持つセグメントである. David はそれらからポテンシャル画像を作り曲線を抽出するのに, 初期値としてもまたタンジェントフィールドのセグメントを用い, それらの重なりで曲線を抽出した.

3.2 Leymarie らのアプローチ

Leymarie ら [29] と藤村ら [27] は, 異なるスケールの画像を用いて初期値収束させる, 同じようなアプローチをした. まず, 大きな σ の Gaussian フィルタによってぼかされた画像からポテンシャル画像を作り, 遠くのかなり大まかに与えられた初期値を収束させる. その収束結果を用いより小さい σ でぼかした画像から作ったポテンシャル画像により収束させる. これを何段階か繰り返し最終的に小さな σ でぼかした画像で収束させる. このことにより, かなり離れた初期値を目的の特徴に引き寄せ, さらに細かな特徴までカバーした輪郭が得られる. この手法では, 初期値に対する仮定が少し緩められている.

3.3 Cohen らのアプローチ

Cohen ら [17][21][31] や Xu ら [34] も初期値について述べている.

Cohen らは, Active Contour Models に風船のように膨らむ力を加えた (7章で述べる). このようなモデルを用いれば, Active Contour Models の初期値は目的の輪郭の内部にあれば, どのようなものでも構わないといったほどの緩い仮定になった.

Xu らもパラメータの正負により, 膨らむもしくは縮むモデルを用いた (7章で述べる). この手法では, 初期値は輪郭の内部にあるものか, もしくは輪郭を外部から囲むものであれば, ほぼどんなものでも構わないといった仮定となる.

3.4 Berger らのアプローチ

Berger ら [14][18] や Henricsson ら [38], Neuenschwander ら [37] は自動的な Active Contour Models を目指した.

Berger らは Active Contour Models をそのタンジェント方向に伸ばしていくモデルを用い, "growing snake" と呼んだ. Berger らによれば, まず適当に置かれた初期値から Snakes を収束させ, 輝度勾配の Snakes と垂直な方向の変化を表す以下の式によって, その Snakes が目的とする輪郭に収束しているかを評価した.

$$\frac{\partial |\nabla I(\mathbf{v}(s))|}{\partial n(\mathbf{v}(s))} \quad (23)$$

これらがある閾値より小さい部分を目的の輪郭に収束している部分として抽出し, 以下のアルゴリズムを用いて Snakes を伸ばしていった.

アルゴリズム:

- 目的とする輪郭の近傍に置かれた初期スネーク S_0 から始める.
- Snakes アルゴリズムにより収束させ, 評価式により目的の輪郭に収束した部分の曲線 C_0 を抽出する.
- 曲線が伸ばせるなら以下を繰り返す.
 - 曲線 C_i をタンジェント方向に伸ばしたものを新たなスネーク S_i とする.
 - Snakes アルゴリズムにより C_i を収束させ C_{i+1} を得る.
 - C_{i+1} から評価式により輪郭に収束した部分 C_{i+1} を得る.

これによると、初期値は目的とする輪郭に最低一点で交わってればよく、その仮定は一段と緩められている。

Neuenschwander らは両端点と両端の傾きだけが正確な大まかな初期値から、正確な初期値を得るための方法を提案している。式 17, 18において (x_0, y_0) と (x_n, y_n) における位置と傾きをきめて、以下のように書き直した。

$$A'x' + f'_x(x, y) = 0 \quad (24)$$

$$A'y' + f'_y(x, y) = 0 \quad (25)$$

ここで、 x' は $(x_1, x_2, \dots, x_{n-1})$ であり (y' も同様)、 A' は $(n-1) \times (n-1)$ 行列である。これを用いて、まず両端点の位置と傾きがわかっている状態で、画像のポテンシャルエネルギーを付加することなくスネークを収束させる。次に両端点から少し内側の部分まで画像のポテンシャルエネルギーを付加し、それから内側には付加しない状態でスネークを収束させる。画像のポテンシャルエネルギーを付加させる部分を内側に伸ばしていき、最終的に輪郭にフィットしたスネークを得る。この手法では、目的とする輪郭に両端の位置と傾きだけがあっている初期値を与えれば良い。

Henricsson らは、画像からのポテンシャルエネルギー、画像のローカルオリエンテーション、キーポイント、エッジ上の一点から Active Contour Models を自動的に当てはめる方法について述べた。そのアルゴリズムは Berger らの "growing snakes" と似ており、エッジ上の一点からローカルオリエンテーションを用いて Snakes の伸ばす方向を決め、キーポイントにおいてその伸長を止めた。これよれば、初期値はエッジ上の一点を指定するだけで良い。

4 パラメータについて

Active Contour Models の内部エネルギーにある Kass らによって $\alpha(s), \beta(s)$ と表された、張力と剛体性を決定するパラメータについての検討がされている。David ら [15], Terzopoulos ら [7], Leymarie ら [29] は、それぞれのパラメータを望ましい Active Contour Models の長さ、曲がりと実際の長さ、曲がりとの関数として表している。Samadani [22] は、Active Contour Models の粘度を表す γ [5] や β をアダプティブに変化させる手法を述べている。坂口ら [20] は α, β が Active Contour Models

に与える影響について調べた。

4.1 Terzopoulos らのアプローチ

Terzopoulos ら [7], David ら [15], Leymarie ら [29] は、 $\alpha(s)$ を Active Contour Models の望ましい長さ $L(s)$ と実際の長さ $h(s)$ との関数で、 $\beta(s)$ を望ましい曲がり $C(s)$ と実際の曲がり $k(s)$ の関数で以下のようにした。

$$\alpha(s) = h(s) - L(s) \quad (26)$$

$$\beta(s) = k(s) - C(s) \quad (27)$$

これにより、Active Contour Models の長さが望ましいものより長くなると α は正になり縮むようになり、短くなると負となり伸びるようになる。

さらに、David らや Leymarie らはさらに、 $\alpha(s)$ として、次式を用いた。

$$\alpha(s) = \tau(h(s) - L(s)) \quad (28)$$

$\tau()$ は頭打ち関数で、ある値以上 (以下) になるとそこで頭打ちになり大きく (小さく) なり過ぎないようにしている。 $\beta(s)$ においては、あまり重大でないとして経験的な定数にしている。さらに Leymarie らにおいては、これをインプリメントする際に Active Contour Models を構成する点を同じ間隔に再サンプリングする方法をとった。

4.2 Samadani のアプローチ

Samadani は適応的に張力、剛体性、ダンピング (8章で述べる) を表すパラメータ α, β, γ を変化させる方法を述べている [22]。Samadani は、ダンピングパラメータの現在の値 γ で 1 ステップとったものと 2γ で 2 ステップとったものとの結果を比べ適応させる方法を述べている。 γ による 1 ステップは近似的に以下のように表せる。

$$v_n^s = v_{n-1} - \frac{1}{\gamma} f(v_{n-1}) \quad (29)$$

また、 2γ による 2 ステップの 1 ステップ目は近似的に以下のように表せ、

$$v_m^t = v_{n-1} - \frac{1}{2\gamma} f(v_{n-1}) \quad (30)$$

2 ステップめは以下のように表せる。

$$v_n^t = v_{n-1} - \frac{1}{2\gamma} f(v_{n-1}) - \frac{1}{2\gamma} f(v_m^t) \quad (31)$$

ここで、制御したいのは以下の式である。

$$\begin{aligned} \|\delta v\| &= \|v_n^t - v_n^s\| \\ &= \frac{1}{2\gamma} \|f(v_{n-1}) - f(v_m^t)\| \end{aligned} \quad (32)$$

これより、以下のような適応式を定義した。

$$\gamma_n = \{(1 - \rho_\gamma)(\|\delta v\| - \delta_\gamma)\} \gamma_{n-1} \quad (33)$$

ここで新しく、 ρ_γ , δ_γ といったパラメータが出てくるが、これらはそれぞれ Active Contour Models の動きの残差と適応レートを示しており、それほど微妙な調整を必要としない。Samadami は、 α や β の適応アルゴリズムも述べている。

4.3 坂口らのアプローチ

坂口らは円にランダムノイズを加えた Active Contour Models を初期値としそれを収束させて、 α, β の Active Contour Models に対する影響を調べた [20]。それらの評価は、Active Contour Models の重心と構成する各点との距離を用いた。それらによると、

- α, β とも、三次以下の周波数成分に影響は少ない。
- α は、三次以上七次以下の周波数成分に影響を与える。
- β は、七次以上の周波数成分に影響を与える。

そのことから、 α は β に比べ低域の周波数成分を抑制するので小さな値を設定し、 β に大きな値を与え対雑音性を向上させると良いと述べられている。

5 収束安定性について

kass らによる差分近似法では、方程式の収束の安定性に問題があった。これらに対して、Cohen らは有限要素法を用い、Amini らは動的計画法を用いた。Williams らは動的計画法の計算量が多くなることに対して工夫を行った。

5.1 Cohen らのアプローチ

Cohen ら [17][31] は、各要素の点の位置だけしか用いない差分近似法に対して、その間の情報をも統合した形

になる有限要素法を用いた。これにより、同じ N 個のセグメントに分割した Active Contour Models における解く方程式の量は増えるものの、差分近似法に比べてセグメント数がかかり少なくて済み、結果計算量は増える。また、ノード間の情報も含んでいるので、収束性、正確さともに差分近似法より優れたものが得られると述べている。

5.2 Amini らのアプローチ

動的計画法を用いた Amini ら [12] の方法は収束性に優れている。以下の漸化式によって v_{i-1} は、 v_i と v_{i+1} の関数として表せる。

$$\begin{aligned} S_i(v_{i+1}, v_i) &= \min_{v_{i-1}} S_{i-1}(v_i, v_{i-1}) + \alpha(|v_i - v_{i-1}|)^2 \\ &\quad + \beta|v_{i+1} - 2v_i + v_{i-1}|^2 + E_{ext}(v_i) \end{aligned} \quad (34)$$

それらを DP テーブルに記憶しておき一番最後から再びトレースして最適解を求める。 v_i の候補が m , $i = 1, \dots, n$ とすると、計算量は $O(nm^2)$ となり、候補数が増えると計算量は爆発的に増大する。これらは、開輪郭について定式化されており、閉輪郭については上田が述べている。

5.3 Williams らのアプローチ

Williams ら [13] は、Amini らの動的計画法による求解を、単純化しその計算量を $O(nm^2)$ から $O(nm)$ に減らした方法を提案した。欲張り法と言われるこの手法は以下のようなアルゴリズムからなっている。ACM の要素の点の数を n 、各要素の近傍の数を m とすると、

- 要素の 1 から n まで。
 - 各近傍 1 から m について。
 - $E_j = \alpha_i E_{cont,j} + \beta_i E_{curv,j} + \gamma_i E_{image,j}$ の内最小の E_j を与える j を得る。
- 要素を j に移動。

以上を移動した点の数が少なくなるまで繰り返す。ここで、 $E_{cont,j}, E_{curv,j}, E_{image,j}$ はそれぞれ、 j における長さ、曲がり、画像によるエネルギーである。

6 画像からの力について

画像の特徴点からの力によるエネルギーを生成するため、画像にフィルタリングなどの処理を施した画像を作り、その画像はポテンシャル画像と呼ばれる。Kassらは輪郭抽出のポテンシャル画像として $E_{image} = -|\nabla I(x, y)|^2$ をもちい、それはまた一般的に用いられているが、画像からの力によるエネルギーに関して、さまざまな工夫が行なわれている。

6.1 Davidらのアプローチ

Davidら [15] は、初期値としても用いたタンジェントフィールドを加工し、ポテンシャル画像を生成した。タンジェントフィールドと平行な方向と垂直な方向に Gaussina フィルタをかけ、フィルタをかけられたタンジェントフィールドの重ね合わせをポテンシャル画像とした。 (x_i, y_i) にある傾き θ_i のセグメントに対して、以下のようにポテンシャルを定義した。

$$\begin{aligned} G_i(x, y) &= -(K_E e^{-(f_i(x, y) - x_i)^2 / \sigma_E^2})(K_B e^{-(g_i(x, y) - y_i)^2 / \sigma_B^2}) \\ &= -K_E K_B e^{-(f_i(x, y) - x_i)^2 / \sigma_E^2 - (g_i(x, y) - y_i)^2 / \sigma_B^2} \quad (35) \end{aligned}$$

ここで、 $f_i(x, y), g_i(x, y)$ はそれぞれ傾き $\theta_i, \theta_i + \frac{\pi}{2}$ に回転させるものである。

$$\begin{aligned} f_i(x, y) &= x_i + (x - x_i) \cos \theta_i + (y - y_i) \sin \theta_i \quad (36) \end{aligned}$$

$$\begin{aligned} g_i(x, y) &= x_i - (x - x_i) \sin \theta_i + (y - y_i) \cos \theta_i \quad (37) \end{aligned}$$

$$U_{TF}(x, y) = \sum_{i=1}^n G_i(x, y) \quad (38)$$

最終的に、タンジェントフィールドのあるところにポテンシャルの谷を作り、そこに Active Contour Models を収束させてエッジ抽出を行なった。

6.2 Cohenらのアプローチ

Cohenら [17][21][31] は、ポテンシャル画像からの力を正規化してもちいた。

$$Ax + f_x(x, y) = 0 \quad (39)$$

$$Ay + f_y(x, y) = 0 \quad (40)$$

において、 $f_x(x, y), f_y(x, y)$ は各点での画像からの力を表し、Kassらの方法ではポテンシャル画像 $P(x, y)$ の勾配をそのまま用いて、

$$F(x, y) = (f_x(x, y), f_y(x, y)) = -\nabla P \quad (41)$$

としたが、これでは $F(x, y)$ が大きい所では近くの望ましい極値を飛び越してしまい戻ってこれず、安定な解が得られない。そこで Cohen は、それを以下のように正規化し用いた。

$$F(x, y) = k \frac{-\nabla P}{\|\nabla P\|} \quad (42)$$

また、もとの $P(x, y)$ には、Canny によるエッジ抽出法を用いその二値化画像にガウシアンフィルタをかけたものを用いた。さらに、その力に ACM の法線方向に働く外部拘束力を加え、風船のように膨らむ ACM を提案した (7章参照)。

6.3 上田らのアプローチ

上田ら [23] はポテンシャル画像として、微分演算子を用いて画像の輝度勾配を計算し、適当な閾値で二値化したものに距離変換を施したものを提案している。 $d(x, y)$ を距離変換後の画素の値として、Active Contour Models の画像からのエネルギーを

$$E_{image}(v_i) = \mu_3 d(v_i) \quad (43)$$

としている。

7 外部拘束力について

Active Contour Models の持つ大きな特徴の一つである対話性の源である外部拘束力には多岐に渡る研究がなされている。ここでは、縮んでいく ACM に外部から力をかけて異なる振舞いをさせたもの、画像からの力を用いず、画像に対する別の記述から得られたものを外部拘束力として利用したものについて述べる。

7.1 Cohenらのアプローチ

Cohenら [17][31] は ACM の法線方向に働く外部拘束力を加え、風船のように膨らむ ACM を提案した。

$$F = k_1 n(s) - k \frac{\nabla P}{\|\nabla P\|} \quad (44)$$

以上のように6章でのべた,正規化された画像からの力とともに,ACMの方線方向 $\mathbf{n}(s)$ に力をかけ小さい初期ACMを膨らませ,目的の輪郭まで膨らんで収束するようにしている.

Xuら[34]は内部エネルギーにより縮んでゆく力をそのまま打ち消す外部力を加えさらにACMの動きを制御するエネルギー項を加えることによって膨らむACM,縮むACMを実現した.

$$E_{snakes} = \int \left\{ \frac{1}{2}(\alpha|v_s(s)|^2 + \beta|v_{ss}(s)|^2 + w_{press}E_{press} + w_{control}E_{control} + w_{edge}|\nabla I(x,y)|^2) \right\} ds \quad (45)$$

ACMは $w_{control}$ の正負により,膨らむ・縮むといった異なる振舞いをする.

7.2 上田らのアプローチ

上田ら[23]はACMに弾性を持たせたモデルを提案している.元となるACMを \mathbf{v}_i^0 として,以下のように弾性的なエネルギーを定義している.

$$E_{elastic}(\mathbf{v}_i) = \frac{1}{2}(\mu_1(|\mathbf{v}_{i+1} - \mathbf{v}_i| - |\mathbf{v}_{i+1}^0 - \mathbf{v}_i^0|) + \mu_2(\text{ang}(\mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}) - \text{ang}(\mathbf{v}_i^0, \mathbf{v}_{i+1}^0, \mathbf{v}_{i+2}^0))) \quad (46)$$

このような,エネルギーを加えることでACM \mathbf{v}_i は元のACM \mathbf{v}_i^0 と同じ形を保とうとし,ACM \mathbf{v}_i にある種の硬さを持たせたことになる.

天野ら[36]は同様に,モデルとするACMと同じエネルギー状態になった時が最良のACMであるとの観点から,モデルとするACMの内部エネルギー $E_{int}^0(\mathbf{v}(s))$ と画像エネルギー $E_{image}^0(\mathbf{v}(s))$ を用いて以下のようにエネルギーを決めた.このエネルギーの最小化により,モデルとするACMに似たACMが抽出できる.

$$E_{msnake} = \int \left\{ |E_{int}(\mathbf{v}(s)) - E_{int}^0(\mathbf{v}(s))| + w_{image}|E_{image}(\mathbf{v}(s)) - E_{image}^0(\mathbf{v}(s))| \right\} ds \quad (47)$$

7.3 栄藤らのアプローチ

栄藤ら[2]は画像の輝度勾配に基づいた画像エネルギーを用いず,画像の混合密度記述から,画素の領域へ属する尤度を推定し,それを外部エネルギーとしてもちいてACMを制御した.ISODATAと呼ばれる,k平均法と評価尺度としてAICを用いたクラスの分割・統合からなるクラスタリング法によって,領域内画素の分布モデルのパラメータとクラス数を推定し,領域の混合密度記述を導いた.その混合密度記述より画素の尤度が得られ,それを用いて,以下のようなエネルギーを用いた.

$$-E_{region}(\mathbf{v}_i) = \begin{cases} \rho_{in}(\mathbf{v}_i) & \text{新制御点が入点} \\ \rho_{out}(\mathbf{v}_i) & \text{新制御点が外点} \\ 0 & \text{上記以外} \end{cases} \quad (48)$$

$$\rho_{in}(\mathbf{v}_i) = -\rho_{out}(\mathbf{v}_i) = l(\mathbf{y}|\mathbf{R}^{out}, \theta^{out}) - l(\mathbf{y}|\mathbf{R}^{in}, \theta^{in}) \quad (49)$$

ここで, $l(\mathbf{y}|\mathbf{R}^{out}, \theta^{out})$ と $l(\mathbf{y}|\mathbf{R}^{in}, \theta^{in})$ はそれぞれ背景領域に対する対数尤度,対象領域に対する対数尤度である.さらに,尤度の勾配を表す尤度のエッジエネルギーも考慮している.

7.4 福井らのアプローチ

福井らも,輝度勾配に基づいた画像からの力を用いず,領域の分離度を定義しそれを用いた外部エネルギーにより,特徴へACMを収束させた.

$$E_{snakes} = \int (E_{int}(\mathbf{v}(s)) + \frac{1}{\eta_b(\mathbf{v}(s))}) ds \quad (50)$$

ここで, $\eta_b(\mathbf{v}(s))$ は分離度であり,以下のように表している.

$$\eta_b = \frac{\sigma_b^2}{\sigma_T^2} \quad (51)$$

$$\sigma_b^2 = n_1(\bar{P}_1 - \bar{P}_m)^2 + n_2(\bar{P}_2 - \bar{P}_m)^2 \quad (52)$$

$$\sigma_T^2 = \sum_{i=0}^{N-1} (P_i - \bar{P}_m)^2 \quad (53)$$

ここで, N は探索領域内の全画素数, n_1, n_2 は探索領域1,2内の画像数, σ_T は全領域の分散, $P_i, \bar{P}_1, \bar{P}_2, \bar{P}_m$ はそれぞれ位置 i の,領域1の平均の,領域2の平均の,全体領域の平均の輝度を表している.この分離度を用いることにより,不明瞭なエッジを持ち,輝度勾配が検出しにくい物体の輪郭を検出している.

8 拡張

Kass らの Snakes と同じように、エネルギー関数の最小化の考えをベースにしたものや、輪郭追跡・抽出の関連技術など Active Contour Models と関係の深い技術をいくつか挙げる。

8.1 Terzopoulos らの拡張

Terzopoulos ら [7] は一般的な deformable model のエネルギー関数について述べている。パラメータの数を p 、変形の自由度を d とすると、パラメータ空間上の一点は $\mathbf{x} = (x_1, \dots, x_p)$ と表せ、モデルの $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \dots, v_d(\mathbf{x}))$ にマッピングされる q 次の deformable model のエネルギー関数は以下のように表せる。

$$E_q(\mathbf{v}) = \sum_{m=1}^q \sum_{|j|=m} \frac{m!}{j_1! \dots j_p!} \int w_j(\mathbf{x}) \left| \frac{\partial^m \mathbf{v}(\mathbf{x})}{\partial x_1^{j_1} \dots \partial x_p^{j_p}} \right|^2 dx + \int P[\mathbf{v}(\mathbf{x})] dx \quad (54)$$

ここで、 $j = (j_1, \dots, j_p)$ はパラメータに対するインデックスで、 $|j| = j_1 + \dots + j_p$ である。Kass らによる Active Contour Models はこの一般化されたエネルギー関数の ($p = 1, d = 2, q = 2$) のものであると述べられている。

8.2 Yuille らの拡張

Yuille ら [10] は、顔の特徴をパラメータ化したテンプレートを用意し、エネルギー関数を輝度のピーク、輝度そのもの、エッジを用いて定義し最急降下法をもちいて最小化して、目や口などの抽出を行なった。

8.3 Cohen らの拡張

Cohen ら [31] は三次元のサーフェースモデルについて述べている。パラメータ s, r により面 $v(s, r)$ を

$$\mathbf{v}(s, r) = (v_1(s, r), v_2(s, r), v_3(s, r)) \quad (55)$$

と表し、

$$E(\mathbf{v}) = \int w_{10} \left\| \frac{\partial \mathbf{v}}{\partial s} \right\|^2 + w_{01} \left\| \frac{\partial \mathbf{v}}{\partial r} \right\|^2 + 2w_{11} \left\| \frac{\partial^2 \mathbf{v}}{\partial s \partial r} \right\|^2 + w_{20} \left\| \frac{\partial^2 \mathbf{v}}{\partial s^2} \right\|^2$$

$$+ w_{02} \left\| \frac{\partial^2 \mathbf{v}}{\partial r^2} \right\|^2 + P(\mathbf{v}(s, r)) ds dr \quad (56)$$

このエネルギー関数を最小化することにより、面モデルを最適化する。

8.4 Staib らの拡張

Staib ら [11] や Chakraborty ら [40] は以下のようにフーリエ領域でパラメータ化したスプライン曲線を用いている。

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos kt \\ \sin kt \end{bmatrix} \quad (57)$$

これらのパラメータ \mathbf{p} を画像の輝度勾配に最適化するために最尤推定によって推定する。

$$\max_{\mathbf{p}} P(\nabla I | \mathbf{p}) P(\mathbf{p}) \quad (58)$$

以上の最大化をバウエル法によって行ないパラメータの推定値を得る。

8.5 Blake らの拡張

Blake ら [33] はカルマンフィルタを用いた統計的な deformable model について述べている (Kalman Snake と呼ばれている)。輪郭を (\mathbf{X}, \mathbf{Y}) で表し、そのダイナミクスを時間的に連続にガウシアンノイズの力が加わるものと仮定すると、

$$\frac{d}{dt} \begin{pmatrix} \mathbf{X} \\ \dot{\mathbf{X}} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{X}} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \quad (59)$$

と書ける (\mathbf{Y} も同様)。この仮定の元で、輪郭 $(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$ を推定するためにカルマンフィルタを用いることができる (K はカルマンゲイン)。

$$\frac{d}{dt} \begin{pmatrix} \hat{\mathbf{X}} \\ \dot{\hat{\mathbf{X}}} \end{pmatrix} = \begin{pmatrix} \dot{\hat{\mathbf{X}}} \\ 0 \end{pmatrix} + K(\mathbf{X}_f - \hat{\mathbf{X}}) \quad (60)$$

さらに、輪郭の位置の分散を、輪郭のアフィン変換によるサブスペースとの距離の分散と見ることでアフィン変換に不変なフィルタとしている。

8.6 Lai らの拡張

Lai ら [41] は輪郭の各要素点をその近傍の線形結合で表す、Shape Matrix を用いた deformable model を提案

している。Shape Matrix を A , 輪郭を U とすると,

$$AU^T = 0 \quad (61)$$

といった, Shape equation を定義する.

$$A = \begin{bmatrix} 1 & -\beta_1 & -\alpha_1 & 0 & \cdots & 0 \\ -\alpha_2 & 1 & -\beta_2 & 0 & 0 & \cdots \\ 0 & -\alpha_3 & 1 & -\beta_3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -\alpha_{n-1} & 1 & -\beta_{n-1} \\ 0 & 0 & \cdots & -\beta_n & -\alpha_n & 1 \end{bmatrix} \quad (62)$$

これらより, 内部エネルギー, 外部エネルギーを以下のようにしている.

$$E_{int}(u_i) = \frac{\|u_i - \alpha_i u_{i\alpha} - \beta_i u_{i\beta}\|^2}{l(U)} \quad (63)$$

$$E_{ext}(u_i, g) = 1 - h_1^T f(u_i + g) \quad (64)$$

ここで, $l(U)$ は輪郭の長さ, g, u_i は輪郭の重心と各要素の点までの変位, $h_1^T f(u_i + g)$ は, 例えば, 0 から 1 の値をとるノイズ入りの輝度テンプレートである. Lai らは, DP を利用して

$$\{U, g\} = \arg \min_{U, g} \sum_{i=1}^n \left\{ \frac{E_{int}(u_i)}{\sigma_i^2} + \frac{E_{ext}(u_i, g)}{\sigma_g^2} \right\} \quad (65)$$

を求めている. これらは, Kass の Snakes のモデルベースとなったものになっている.

8.7 梅山の拡張

梅山は複素自己回帰モデルによって, 点列 (x_j, y_j) を

$$z_i = \sum_{k=1}^m a_k z_{j-k} + \epsilon_j \quad (66)$$

とモデル化した. ϵ は複素ノイズ. これによる内部エネルギーを以下のように表し, 最急降下法で極小値を求めた.

$$E_{int} = \frac{1}{N} \sum_{j=0}^{N-1} \|z_j - \sum_{k=1}^m a_k z_{j-k}\|^2 \quad (67)$$

8.8 Bascle らの拡張

Bascle らは, アフィン変換や剛体の動きを行列で表したモーションモデルと Active Contour Models を組み

合わせて追跡を行ない [42]. また, 領域分割から動き推定されたモーションと Active Contour Models を組み合わせた追跡も行なっている [42].

8.9 福井らの拡張

福井ら [28] は人間の顔の輪郭を追跡するのに, 局所的な遮蔽や隠れていた輪郭の出現による追跡の失敗の伝播を防ぐため Snakes を小さく分裂したマルチスネークと呼ぶものを用いた. さらに, それら小さなスネーク同士をバネ・ダンパでつなぎ全体的な動きにロバストにしている.

9 適用

Couvignou ら [26][32] は, hand-eye ロボットで剛体を追跡する際の制御に, active deformable models を用いている. また, Curwen ら [39] は, ロボットの空間の探索やバスプランニングに deformable contour を用いての追跡を利用している. 人間の手を deformable template で追跡し, 指さしたものを認識するマンマシンインターフェースに利用されているものもある [43]. その他にも, ステレオマッチング [24][35] に用いられているものなどもある.

10 おわりに

Snakes: Active Contour Models とその関連技術について Kass らとのアプローチの違いに観点を置いて述べた. 画像から得られる最低のレベルの情報をわずかな制約によって, 一段高いレベルの情報へ変換できるという点で Active Contour Model は広いコンピュータビジョンの分野で多岐に渡って利用される基盤技術となり得る. 今後, Active Contour Models について Deformable Models の元になるモデルの画像からの構築や, 望まない局所的極小値の回避方法についての発展が待たれる. これらは全て, Kass らによってユーザとのインタラクションとして解決されていたものを何らかの拘束条件で計算機によって代行することによって実現できると考えられる. 実際, Kass ら以降の Active Contour Models はユーザとのインタラクションを減らすためになされたと考えられるものが多い. Active Contour Models によ

て与えられたわずかな拘束条件をさらに発展させていき、一段高度な情報をさらに高度な情報へ押し上げていくことがこれから Active Contour Models に必要とされているのではないだろうか。

第 II 部

Snakes による人物像追跡

11 人物像の輪郭

11.1 人物像の輪郭の重要性

人物の動きや姿勢を自動的かつ非接触な方式で検出する技術は、知的画像通信や高度なマンマシンインターフェースとして、広い応用範囲を持っている。それゆえ、画像通信システム等の分野でそのような技術の重要性が増しており、コンピュータビジョンの分野でも人物像の認識・姿勢推定が精力的に取り組まれている。

人物の姿勢推定に関するものの多くは人物像のシルエットを利用しており、胴体に重なった腕など部品と部品の重なりが検出できず、重なった部品の位置、姿勢の推定は非常に困難である。それを解決するため複数のカメラを用いる方法 [46] も提案されているが、人物像の輪郭の位置や動きといった情報が得られ利用できれば、姿勢推定の信頼性も向上すると考えられる。

11.2 人物像の輪郭の特性

人物の姿勢推定に人物像の輪郭情報は有効に働くと考えられるが、それは画像中から抽出するには困難な特性を持っている。第一に、人物が関節を持った物体であるため、その輪郭は非常に複雑に変形する。それゆえ、部分的に剛体性は仮定できるものの、関節位置が検出できなければその仮定さえ利用することができない。第二に、その輪郭は頻繁に出現・遮蔽を繰り返す。たとえば、複雑な変形に耐え、追跡に成功したとしても、その輪郭はすぐに他の部分に遮蔽され、また全く新しい輪郭が他の部分から出現する。追跡には全く不向きなこれら二つの特性を持つため、人物像の輪郭の追跡は困難なものとなっている。

12 Active Contour Models による人物像輪郭の追跡

12.1 Active Contour Models の特性

Active Contour Models は、それ自体の滑らかさと画像への適合度をそれぞれエネルギーとして表し、それら

の和を最小化するため、特徴点のあるところでは特徴点にそって滑らかに輪郭を形成し、特徴点の無いところでも滑らかな輪郭を形成することができる。また、局所的に乱れた特徴点を入力しても、全体でエネルギーを最小化するため、平滑化することができる。ユーザは初期輪郭を与えると、特徴点が乱れたものであっても、滑らかな変形や動きを追跡できる。

12.2 Active Contour Models の問題点

初期輪郭から滑らかな変形・動きを追跡するのに有効と考えられる Active Contour Models の特性も、複雑な変形や動きをする追跡には適していない。その理由の一つは初期輪郭を与えなければならないことである。複雑な変形や動きによっておこる輪郭の遮蔽や出現に対応するためには、初期輪郭を計算機自身が計算することができなければならない。また、輪郭の遮蔽や出現による誤った特徴点の追跡を行なってしまい、さらに平滑化の作用のためその影響が Active Contour Models 全体に伝播してしまう。

12.3 従来のアプローチ

従来、Active Contour Models による複雑物体の輪郭追跡は前述の問題のため、あまり行なわれていない。しかし、それら個別の問題には取り組まれている。初期輪郭の問題には多くの試みがなされている。しかし、それらも完全に計算機による初期輪郭の設定をしたものではなく、単に人間の初期輪郭設定をかなり正確なものでもなくとも大まかなものでよいとしたものにとどまっている [14][17][38]。また、全体に局所的な影響が伝播することを防ごうとしたものには、スネークを細かなスネークに分けて適用するマルチスネーク [28] があるが、それらをつなぐバネを仮定しているため完全に遮蔽や出現の影響を防ぐものではなかった。

13 Snakes を用いた複雑形状輪郭の自動抽出法

そこで、人間による初期設定を行わず輪郭に自動的に Snakes を当てはめる方法を提案する。これによれば、非常に複雑な輪郭の追跡を行なうことができると考

えられる。

13.1 提案するアルゴリズムの概要

アルゴリズムは以下の図1のように5つのステップよりなっている。生成ステップではエッジピクセルよりACMの素を抽出し、同じ輪郭線にあると考えられるものは結合ステップで一つのACMに結合する。ACMを当てはめられた輪郭はそのACMによって追跡される。遮蔽や追跡失敗によって輪郭から外れたと考えられるACMは分裂や消滅ステップによって削除し、出現や追跡の失敗した輪郭には生成ステップによってACMが当てはめられる。

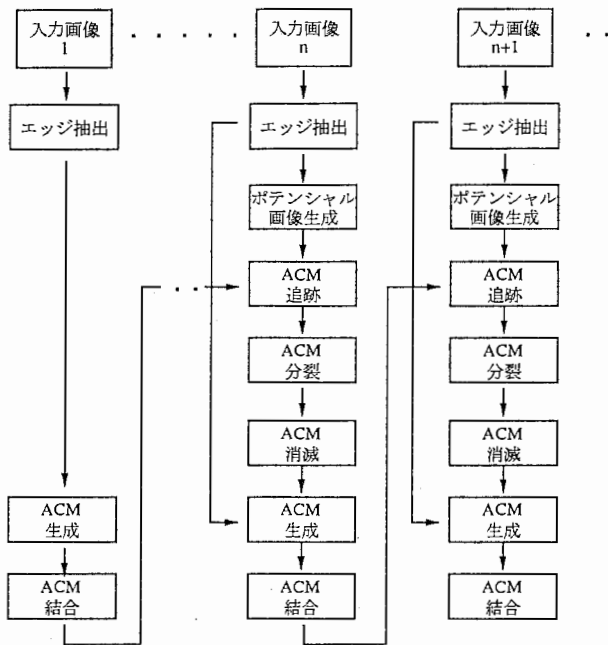


図1: アルゴリズムの概要

13.2 生成ステップ

生成ステップでは、エッジピクセルよりACMの素(最小形のACM)を抽出する。まず、画像にエッジオペレータを施す、今回はソーベルと非極大値抑制を用いたがエッジピクセルを抽出できればラプラシアンガウシアンフィルタやCannyのエッジ抽出法やその他の手法でも構わない、抽出された画像は以下の図2のようになるが、結合ステップによってACMの両端の傾きが条件になることを考慮し、それにロバストになるよう隣合う

エッジピクセルを間引きした図3のようなエッジ画像を用いる。

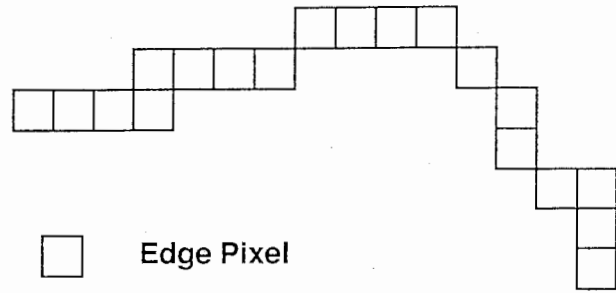


図2: エッジピクセル

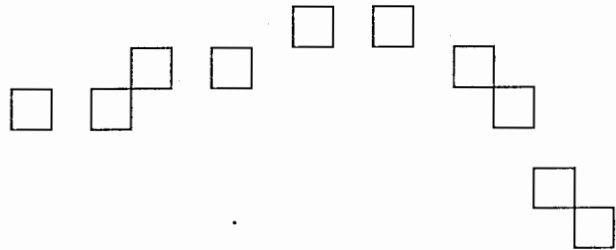


図3: 間引きしたエッジピクセル

また、追跡されている輪郭に新しくACMを当てはめる必要はないので、そのエッジピクセルを消す。つまり、処理 $n = 1, \dots, n$ のうち $n \geq 2$ に関しては、前の追跡の結果のACMデータを用い、以下の条件を満たすエッジピクセルは消去する。ACMの各要素を $V_i = (x_i, y_i) (i = 0, \dots, n)$ 、エッジピクセルを $P_i = (px_i, py_i) (i = 0, \dots, m)$ とすると、

$$\frac{|(y_{i-1} - y_i)px_j - (x_{i-1} - x_i)py_j - y_{i-1}x_i + x_{i-1}y_i|}{\sqrt{(y_i - y_{i-1})^2 + (x_i - x_{i-1})^2}} \leq \text{threshold} \quad (68)$$

$$(x_i - x_{i-1})(px_j - x_{i-1}) + (y_i - y_{i-1})(py_j - y_{i-1}) \geq 0 \quad (69)$$

$$(x_i - x_{i-1})(px_j - x_i) + (y_i - y_{i-1})(py_j - y_i) \leq 0 \quad (70)$$

今回は閾値として5 pixelをもちいた。

こうして得られたエッジピクセルから以下の条件を満たす2点を結ぶ直線を最小形のACMとする(図4)。

エッジピクセルを $P_i (i = 0, \dots, n)$ とすると、

$$P_k \cdot P_l^T \leq \text{threshold} \quad (71)$$

$$l = \underset{i}{\operatorname{arg\,min}}(P_k \cdot P_i^T) \quad (72)$$

閾値として2.5 pixelを用いた.

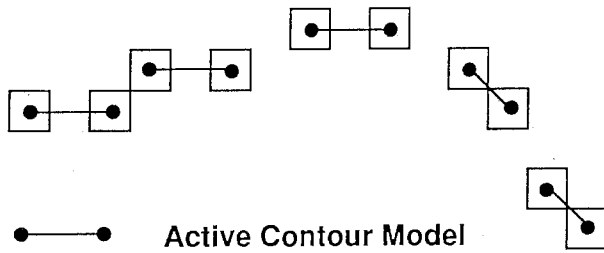


図4: ACMの最小形

このようにして得たACMの最小形と、画像 $n \geq 2$ に関しては前のACMによる追跡データとを用いて、結合ステップへ進む。

13.3 結合ステップ

結合ステップは、追跡しているACMデータとACMの最小形とから、同じ輪郭上にあると考えられるACMデータを結合させる。ここでは輪郭の連続性を用いて、結合させるか判定を行なう。まず、輪郭はつながっていると考えられるので、両端の端点の距離を判定する。ACMデータから任意に選んだ二つのACM, A, BについてA, Bを

$$A: \quad (ax_i, ay_i) \quad i = 1, \dots, n \quad (73)$$

$$B: \quad (bx_i, by_i) \quad i = 1, \dots, m \quad (74)$$

とすると、

$$(ax_n - bx_1)^2 + (ay_n - by_1)^2 \leq \text{threshold} \quad (75)$$

を条件の一つ目とする。また、輪郭は滑らかにつながっていると仮定すると、ACMの端点の傾きの差が一定値以下でなければならない。そこで、

$$\left| \arctan \frac{ay_{n-1} - ay_n}{ax_{n-1} - ax_n} - \arctan \frac{by_1 - by_2}{bx_1 - bx_2} \right| \leq \text{threshold} \quad (76)$$

を二つ目の条件とする。さらにACMが同じ輪郭上であれば、端点の両側の輝度差が似たものであると考えられる。そこで、

$$\left| \frac{\partial I}{\partial n(ax_n, ay_n)} - \frac{\partial I}{\partial n(bx_1, by_1)} \right| \leq \text{threshold} \quad (77)$$

ここで、 $\mathbf{n}(ax_n, ay_n)$ はACMにおける (ax_n, ay_n) の法線ベクトルであり、 I は輝度値である。今回は閾値にそれぞれ3.0, 0.25, 10.0を用いた。こうして得られた、ACMデータによって次フレームの追跡を行なう。

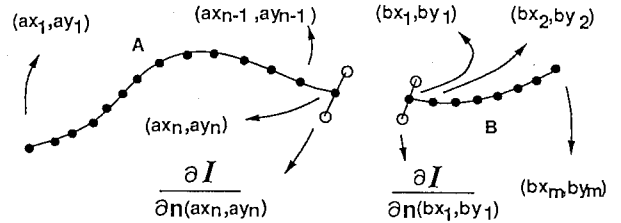


図5: ACMの結合

13.4 追跡ステップ

ここでは、生成・結合によって得られたACMデータによって追跡を行なう。追跡の手法は上田らによって提案された弾性モデルと、Aminiらによる動的計画法を用いた。以下に上田らによる弾性モデルのエネルギー関数を示す。

$$E_{snake} = \sum_{i=1}^n (E_{elastic} + E_{field}) \quad (78)$$

$$E_{elastic}(\mathbf{v}_i) = \frac{1}{2} (\mu_1 (|\mathbf{v}_{i+1} - \mathbf{v}_i| - |\mathbf{v}_{i+1}^0 - \mathbf{v}_i^0|)^2 + \mu_2 (\text{ang}(\mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}) - \text{ang}(\mathbf{v}_i^0, \mathbf{v}_{i+1}^0, \mathbf{v}_{i+2}^0))^2) \quad (79)$$

ここで、 \mathbf{v}_i^0 は初期輪郭であり、 $\text{ang}(\mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{v}_{i+2})$ は、 $\mathbf{v}_i \mathbf{v}_{i+1}$ と $\mathbf{v}_{i+1} \mathbf{v}_{i+2}$ のなす角である。また、画像のポテンシャルエネルギーにはACMの生成の際に使ったエッジ画像を、距離変換したものを用いた。

$$E_{field}(x, y) = d(S(x, y)) \quad (80)$$

$d()$ は距離変換画像の値。 $S(x, y)$ はエッジ画像を表す。今回は、 $\mu_1 = 1.0, \mu_2 = 1.0$ を用いた。

13.5 分裂ステップ

輪郭の一部の遮蔽により、ことなる特徴点に引き付けられたACMを分裂させる。ACMの両側の輝度差を判定の基準とする。ACM上に一点をとり、その点で二分

したそれぞれの ACM で両側の輝度差, C, C' を計算する.

$$C = \frac{1}{k} \sum_{i=1}^k C_i \quad (81)$$

$$C' = \frac{1}{n-k} \sum_{i=k+1}^n C_i \quad (82)$$

それらの差を最大にする点を分裂点 d とし, その分裂点で二つの ACM に分割する.

$$d = \arg \max_k |C - C'| \quad (83)$$

今回は最大となる輝度差が 2.0 を越えた時のみ分裂を行

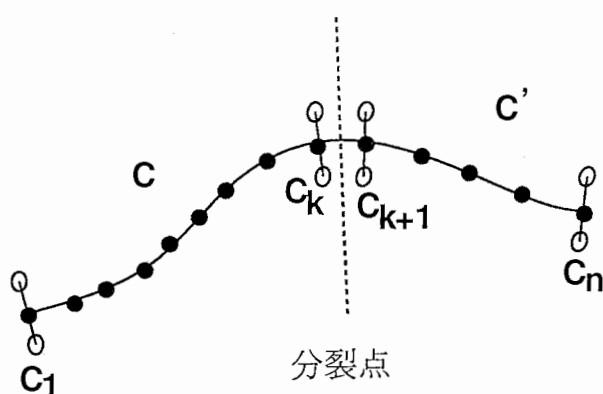


図 6: ACM の分裂

なった.

13.6 消滅ステップ

ACM の両側の輝度差を計算し, それによって ACM が輪郭にのっているか評価する.

$$C = \frac{1}{n} \sum_{i=1}^n C_i \quad (84)$$

輝度差が小さいものは, ACM が輪郭から外れたと見なし消去する.

$$C \leq \text{threshold} \quad (85)$$

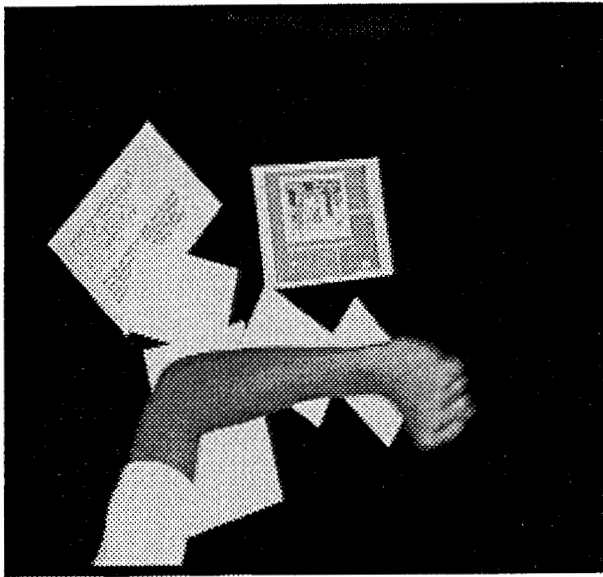
閾値は 1.0 とした.

14 結果と考察

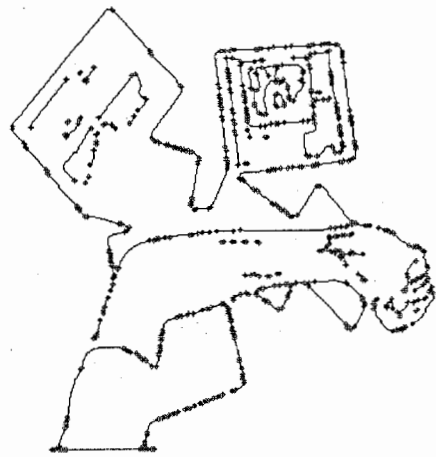
本手法を, 机の上に置かれた紙の上を通過する腕の画像に適用した結果を次ページに示す. それぞれ, 1, 6,

11, 16, 21, 26 フレーム目の画像とそれに当てはめられた ACM である. 点から始まり点まで引かれた線分が一つの ACM になっている. 滑らかにつながり, 両側の輝度差が一定である輪郭に一つの ACM が当てはめられていることがわかる. また, 紙の輪郭と腕の輪郭とが複雑に遮蔽・出現を起こしているが, 輪郭の隠れたものからは ACM が外され, 新しく出てきた輪郭には新しく ACM が当てはめられていることがわかる.

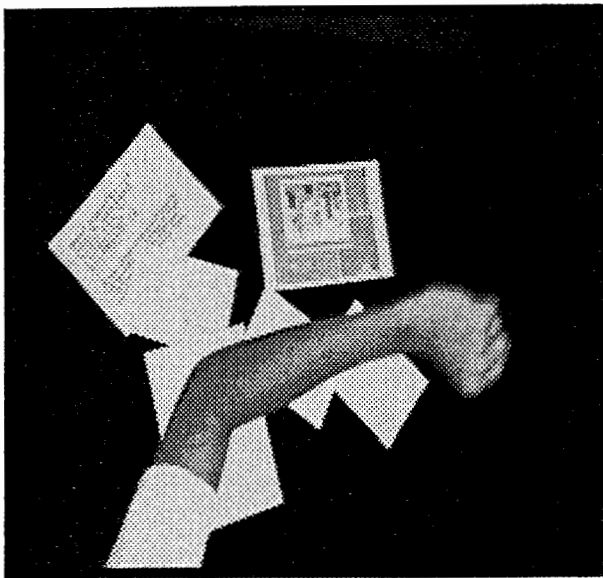
本手法では, 滑らかにつながっており両側の輝度差が一定の輪郭セグメントに一つの ACM をあてはめることによる輪郭の自動抽出と追跡を目的とした. しかし, 腕の輪郭や紙の輪郭を自動的に抽出するには不十分であり, 今後, 当てはめられた ACM の動きの情報やモデルの情報を利用することによって, それらが達成できると考えられる. また, 結合条件には閾値を基にした条件を用いたが, 閾値は画像に大きく依存し閾値の調節も問題となる. したがって, 閾値とは別の結合の評価関数を定義し, それを用いた結合条件を設定することも一つの問題である.



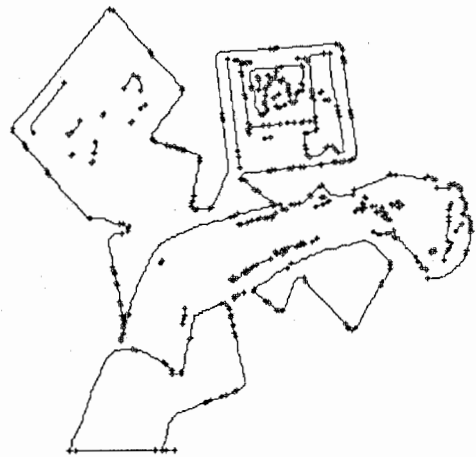
1 フレーム目



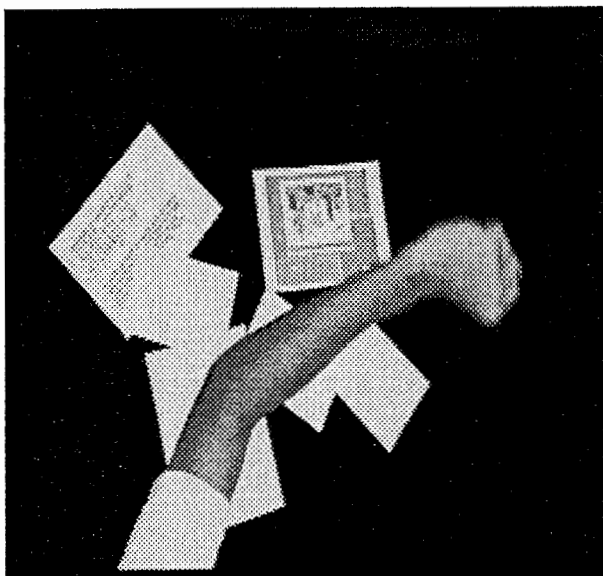
抽出結果



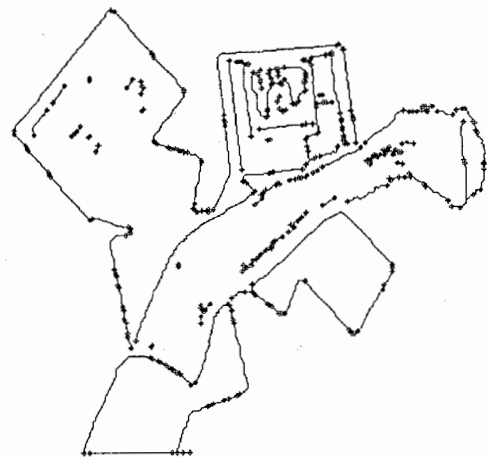
6 フレーム目



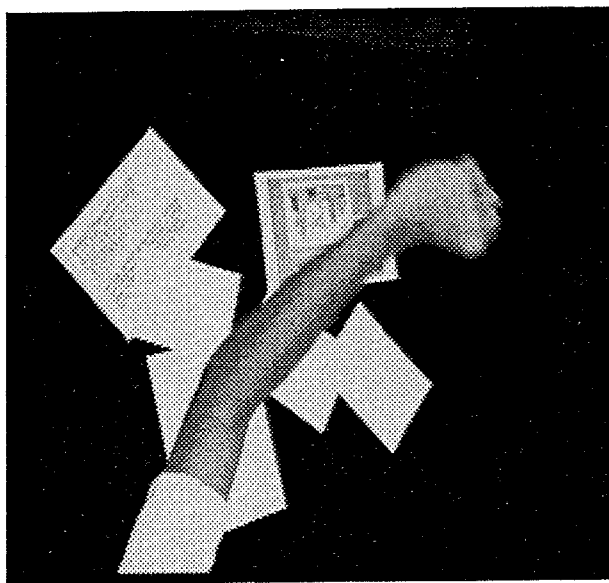
抽出結果



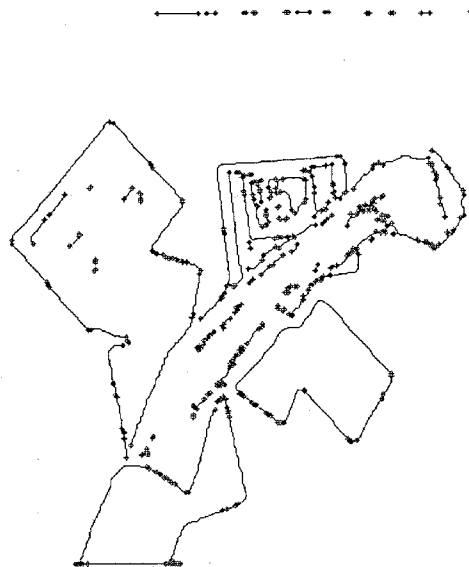
11 フレーム目



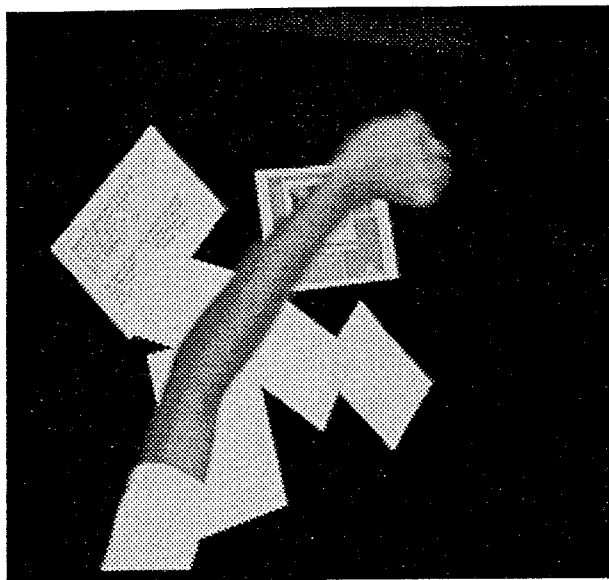
抽出結果



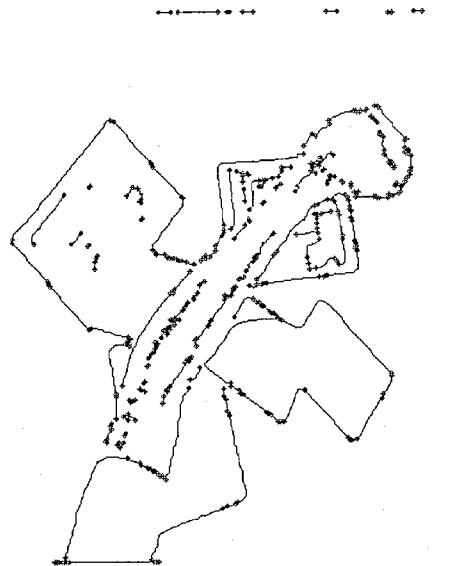
16 フレーム目



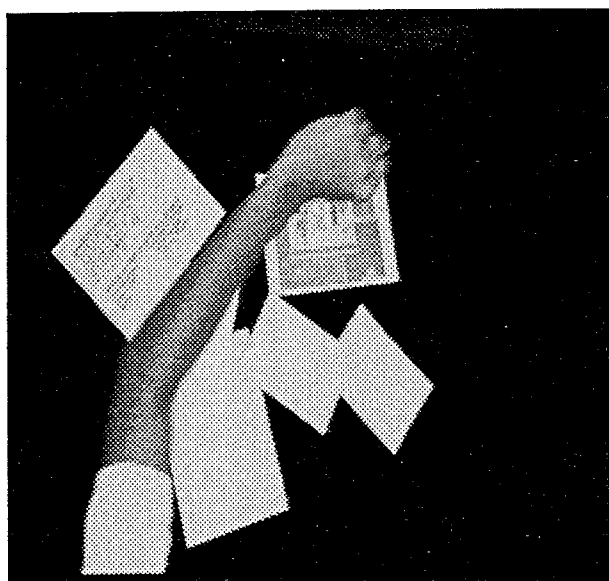
抽出結果



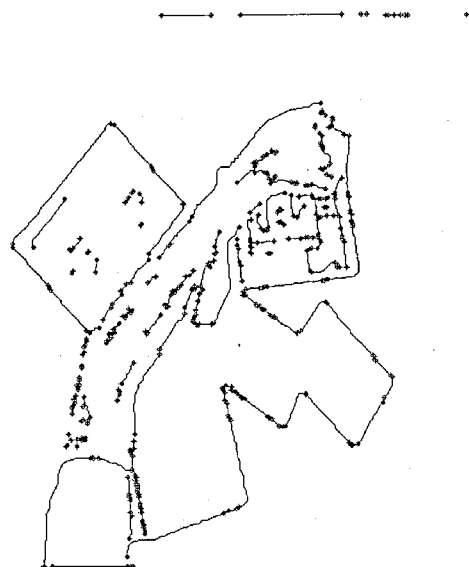
21 フレーム目



抽出結果



26 フレーム目



抽出結果

参考文献

- [1] David Marr. *VISION A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H.Freeman and Company, 1982.
- [2] 柴藤稔, 白井良明, 浅田稔. クラスタリングから得られる領域の記述に基づく動的な輪郭抽出. 電子情報通信学会論文誌, Vol. J75-DII, No. 7, pp. 1111-1119, July 1992.
- [3] 橋本昌寛, 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, Vol. J77-DII, No. 11, pp. 2171-2178, November 1994.
- [4] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321-331, September 1988.
- [5] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Regularization of inverse visual problems involving discontinuities. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 413-422, 1986.
- [6] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Symmetry-seeking models and 3d object recognition. *International Journal of Computer Vision*, Vol. 3, No. 1, pp. 211-221, 1987.
- [7] Demetri Terzopoulos. On matching deformable models to images. In *Topical meeting on machine vision, Technical Digest Series*, Vol. 12, pp. 160-167, 1987.
- [8] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artificial Intelligence*, Vol. 36, pp. 91-123, 1988.
- [9] Steven W.Zucker and Chantal David. The organization of curve detection: Coarse tangent fields and fine spline covering. In *Proc. International Conference on Computer Vision*, pp. 568-577, 1988.
- [10] Alan L.Yuille, David S.Cohen, and Peter W.Hallinan. Feature extraction from faces using deformable templates. In *Proc. Computer Vision and Pattern Recognition*, pp. 104-109, San Diego, CA, 1989.
- [11] Lawrence H.Staib and James S.Duncan. Parametrically deformable contour models. In *Proc. Computer Vision and Pattern Recognition*, pp. 98-103, San Diego, CA, 1989.
- [12] Amir A. Amini, Terry E. Weymouth, and Ramesh C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 9, pp. 321-331, September 1990.
- [13] Donna J.Williams and Mubarak Shah. A fast algorithm for active contours. In *Proc. International Conference on Computer Vision*, pp. 592-595, Osaka, Japan, 1990.
- [14] Marie-Odile Berger and Roger Mohr. Towards autonomy in active contour models. In *Proc. International Conference on Pattern Recognition*, pp. 847-851, ??, 1990.
- [15] Chantal David and Steven W.Zucker. Potentials, valleys, and dynamic global coverings. *International Journal of Computer Vision*, Vol. 5, No. 3, pp. 219-238, 1990.
- [16] Sylvie Menet, Philippe Saint-Marc, and Gérard Medioni. B-snakes: implementation and application. In *Proc. DARPA Image Understanding Workshop*, pp. 720-726, ??, 1990.
- [17] Laurent D.Cohen and Isaac Cohen. A finite element method applied to new active contour models and 3d reconstruction from cross sections. In

- Proc. International Conference on Computer Vision*, pp. 587-591, Osaka, Japan, December 1990.
- [18] Marie-Odile Berger. Snake growing. In *Proc. 1st European Conference on Computer Vision*, pp. 570-572, ??, 1990.
- [19] Pascal Fua and G.Leclerc. Model driven edge detection. *Machine Vision and Applications*, Vol. 3, No. 1, pp. 45-56, 1990.
- [20] 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. 電子情報通信学会技術報告, Vol. PRU90-21, No. ?, pp. 43-44, 1990.
- [21] Laurent D.Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, Vol. 53, No. 2, pp. 211-218, March 1991.
- [22] Ramin Samadami. Adaptive snakes: control of damping and material parameters. In *Proc. SPIE. Conf. Geometric Methods in Computer Vision Vol.1570*, pp. 202-213, ??, 1991.
- [23] 上田修巧, 間瀬, 末永. 弾性輪郭モデルとエネルギー最小化原理による輪郭追跡手法. 電子情報通信学会論文誌, Vol. J75-D-II, No. 1, pp. 111-120, 1991.
- [24] B.Basclé and R.Deriche. Features extraction using parametric snakes. In *Proc. International Conference on Pattern Recognition*, pp. 659-662, ??, 1992.
- [25] Frédéric Leymarie and Martin D.Levine. Simulating the grassfire transform using an active contour model. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 1, pp. 56-75, January 1992.
- [26] Philippe A.Couvignou, Nikolaos P.Papanikolopoulos, and Pradeep K.Khosla. Hand-eye robotic visual servoing around moving objects using active deformable models. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1855-1862, Raleigh, NC, July 1992.
- [27] 藤村恒太, 横矢直和, 山本和彦. 多重スケール画像を用いた非剛体物体の動的輪郭追跡. 情報処理学会研究会資料, Vol. CV78-4, No. ?, pp. 25-32, May 1992.
- [28] 福井和広, 久野義徳. マルチスケールによる動物体の輪郭追跡. 情報処理学会研究会資料, Vol. CV80-24, No. ?, pp. 177-184, November 1992.
- [29] Frédéric Leymarie and Martin D.Levine. Tracking deformable objects in the plane using an active contour model. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, No. 6, pp. 617-634, June 1993.
- [30] Frank P.Ferrie, Jean Lagarde, and Peter Whaite. Darboux frames, snakes, and super-quadratics: Geometry from the bottom up. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, No. 8, pp. 771-784, August 1993.
- [31] Laurent D.Cohen and Isaac Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, No. 11, pp. 1131-1147, November 1993.
- [32] Philippe A.Couvignou, Nikolaos P.Papanikolopoulos, and Pradeep K.Khosla. On the use of snakes for 3-d robotic visual tracking. In *Proc. Computer Vision and Pattern Recognition*, pp. 750-751, New York City, NY, June 1993.
- [33] Andrew Blake, Rupert Curwen, and Andrew Zisserman. Affine-invariant contour tracking with automatic control of spatiotemporal scale. In *Proc. International Conference on Computer Vision*, pp. 66-75, Berlin, Germany, 1993.
- [34] Gang Xu, Eigo Segawa, and Saburo Tsuji. A robust active contour model with insensitive parameters. In *Proc. International Conference on Computer Vision*, pp. 562-566, Berlin, Germany, 1993.

- [35] B.Basclé and Deriche R. Stereo matching, reconstruction and refinement of 3d curves using deformable contours. In *Proc. International Conference on Computer Vision*, pp. 421-430, Berlin, Germany, 1993.
- [36] 天野晃, 坂口嘉之, 美濃導彦, 池田克夫. サンプル輪郭モデルを利用した snakes. 電子情報通信学会論文誌, Vol. J76-D-II, No. 6, pp. 1168-1176, June 1993.
- [37] W.Neuenschwander, P.Fua, G.Székely, and O.Kübler. Initializing snakes. In *Proc. Computer Vision and Pattern Recognition*, pp. 658-663, Seattle, Washington, June 1994.
- [38] Olof Henricsson and Walter Neuenschwander. Controlling growing snakes by using key-points. In *Proc. International Conference on Pattern Recognition*, pp. 68-73, ??, 1994.
- [39] R.Curwen, A.Blake, and A.Zisserman. Real-time visual tracking for surveillance and path planning. In *Proc. European Conference on Computer Vision*, pp. 879-883, ??, 1992.
- [40] Amit Chakraborty, Lawrence H.Staib, and James S.Duncan. Deformable boundary finding influenced by region homogeneity. In *Proc. Computer Vision and Pattern Recognition*, pp. 624-627, Seattle, Washington, June 1994.
- [41] Kok F.Lai and Roland T.Chin. Deformable contours: Modeling and extraction. In *Proc. Computer Vision and Pattern Recognition*, pp. 624-627, Seattle, Washington, June 1994.
- [42] Benedicte Basclé, Patrick Bouthemy, Rachid Deriche, and Francois Meyer. Tracking complex primitives in an image sequence. In *Proc. International Conference on Pattern Recognition*, pp. ??-??, ??, 1994.
- [43] Roberto Cipolla, Paul A.Hadfield, and Nicholas J.Hollinghurst. Uncalibrated stereo vision with pointing for a man-machine interface. In *IAPR Workshop on Machine Vision Applications*, pp. 163-166, Kawasaki, Japan, December 1994.
- [44] 福井和広. 領域間の分離度に基づく物体輪郭抽出. 情報処理学会研究会資料, Vol. CV88-2, No. ?, pp. 9-15, March 1994.
- [45] 梅山伸二. 複素自己回帰モデルを用いた輪郭抽出. 電子情報通信学会技術報告, Vol. PRU94-65, pp. 9-16, November 1994.
- [46] 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. 電子情報通信学会技術報告, Vol. PRU93-122, No. ?, pp. 47-54, 1994.