

対象追跡 – フレーム間の類似度に着目した手法から動きのモデルに着目した手法まで –

加藤 丈和[†] 深尾 隆則[‡] 羽下 哲司^{††}

[†]和歌山大学 情報通信システム学科

[‡]神戸大学 機械工学科

^{††}三菱電機 先端技術総合研究所

E-mail : [†]t.kato@ieee.org, [‡]fukao@mech.kobe-u.ac.jp, ^{††}Haga.Tetsuji@wrc.melco.co.jp

あらまし

本稿では、対象追跡に関連する研究について、主に対象の探索方法に注目し、動きのモデルに着目した時系列フィルタリングに基づく手法と、フレーム間の類似度に着目した勾配法に基づく手法を中心に最近の動向を紹介する。

キーワード 対象追跡, 時系列フィルタリング, 勾配法

Target Tracking: From Similarity Based Approaches to Dynamics Based Approaches

Takekazu Kato[†], Takanori Fukao[‡], Tetsuji Haga^{††}

[†]Department of Computer and Communication Science, Wakayama University

[‡]Department of Mechanical Engineering, Kobe University

^{††}Advanced Technology R&D Center, Mitsubishi Electric Corporation

E-mail : [†]t.kato@ieee.org, [‡]fukao@mech.kobe-u.ac.jp, ^{††}Haga.Tetsuji@melco.co.jp

Abstract This survey article introduces recent researches for target tracking. We focus on the target tracking between frames and mainly introduce approaches based on stochastic filtering and gradient method.

Keywords target tracking, stochastic filtering, gradient method

1 はじめに

対象追跡は、与えられた時系列画像から、指定した対象が画像上でどのように移動するかを推定する問題であり、時系列画像の解析の中でも中心的な話題の一つである。対象追跡では、その目的は様々であるが、実時間処理が求められる場合が多く、一枚一枚の画像を高速に処理する必要がある。

1970年代には、航空写真の解析で追跡技術が検討されている[1]。しかし、この頃の追跡技術はハードウェア的な制約から、オフライン処理を前提とするものであった。その後1980年代から一度対象追跡の研究が下火になっていくが、リアルタイム処理が可能なハードウェア条件が整い始めた1980年代後半から1990年代前半には再度さかんに研究

されるようになってきた。

1990年代以降に提案された手法は、リアルタイム処理を前提とした手法が中心となり、時系列での追跡の安定化や計算コスト削減のために、動きのモデル化により対象の動き予測を行う手法や、フレーム間の類似度の勾配から対象の位置、姿勢の変化を推定する手法が多く提案されている。これらの流れは、その後1990年代後半から2000年代にかけて様々な発展を遂げている。

2000年代に入ると、人物監視などの用途から環境変動をともなうシーンでの追跡についての研究がさかんにになり、パターン認識技術に基づく手法や、クラスタリングに基づく手法など、新たな動きが見え始めている。

本稿では、特に1990年代以降のリアルタイム処

理を前提とした追跡手法について紹介する。個々の追跡技術の特徴づける要素は、対象のモデル表現、追跡に用いる特徴量、類似度の定義など様々な要素が考えられるが、ここでは対象を時系列でどう対応付けるかという観点から、追跡技術の紹介を行う。このような観点から追跡技術を分類すると、多くの追跡手法はフレーム間の類似度に着目し、その勾配から対象の移動先を推定する手法と、対象の動きをモデル化し、対象の移動先を予測しながら追跡する手法に分類することができる。以降の章では、これらの二つのアプローチについて重点的に紹介する。

2章では Lucas-Kanade の手法を利用した特徴点追跡を中心とした類似度の勾配に着目した手法を紹介し、3章では動きのモデルに着目した手法として、時系列フィルタリングに基づく手法を紹介する。また、4章では追跡に用いる特徴量という観点から、上記の分類に収まらない技術を中心に、最近の動向について紹介する。

2 Point-Feature-based Tracking

連続する画像系列の情報に基づく Ego-Motion や Structure from Motion アルゴリズムを始めとする様々な手法は、高い精度を得るために、ある点を出来るだけ正確に追従することを要求する。しかし、一点のみの情報を用いて追従することは困難であるため、ある点周りの領域 (Point Feature) を含めて追従していくことになる。このような Point Feature に基づいた追跡をこの章では考える。ここで、オクルージョンなどのために領域内のすべての点を一致させることは出来ず、領域の中心点を出来るだけ正確に追従することが目標となる。

この目標を達成するための方法は、大きく二つに分けられる。一つは、元の位置周辺に領域を直接移動させて全探索し、何らかの指標を最小化 (または最大化) するところを対応点とするブロックマッチングアルゴリズムである。この指標としては SSD (Sum of Squared Differences), SAD (Sum of Absolute Differences), CC (Cross Correlation) などがある。もう一つは、SSD などの指標の最小化を輝度勾配の情報に基いて探索、求解する方法である。この始まりは Lucas-Kanade のアルゴリズムであり、現在の代表的な手法である。以下では基本的な方法をまず示し、その発展形や関連する事項について述べる。

2.1 ブロックマッチングアルゴリズム

最も基本となるのは、SSD を評価規範として用いるものであり、二つの画像 I, J に関して、 I の考慮する画像領域 \mathcal{W} を周辺にスライドさせ、次の

2乗誤差和を最小化する変位量 d を画像 J にて発見する方法である。

$$\epsilon = \int_{\mathcal{W}} [J(\mathbf{x} + \mathbf{d}) - I(\mathbf{x})]^2 dx \quad (1)$$

SSD は外れ値 (Outlier) が存在するときに、大きな影響を受けるため、単純に以下のような絶対値の誤差和を用いる方法も考えられている。

$$\epsilon = \int_{\mathcal{W}} |J(\mathbf{x} + \mathbf{d}) - I(\mathbf{x})| dx \quad (2)$$

さらに、SSD から導かれる指標として以下の相互相関を用い、最大化する変位量 d を求める方法も高速な実装が可能なものとして用いられている。

$$\gamma = \int_{\mathcal{W}} [I(\mathbf{x}) - \bar{I}][J(\mathbf{x} + \mathbf{d}) - \bar{J}] dx \quad (3)$$

ここで、 \bar{I} は領域 \mathcal{W} における I の平均、 \bar{J} は J の平均を表す。

また、特に照明変化のある時などに、次の正規相互相関 (Normalized Cross Correlation) を指標として用いる方法 [2] も一般的である。

$$\gamma = \frac{\int_{\mathcal{W}} [I(\mathbf{x}) - \bar{I}][J(\mathbf{x}) - \bar{J}] dx}{\sqrt{\int_{\mathcal{W}} [I(\mathbf{x}) - \bar{I}]^2 dx \int_{\mathcal{W}_1} [J(\mathbf{x}) - \bar{J}]^2 dx}} \quad (4)$$

これらを基とした方法は多数提案されており、それ以外にも他のノルムや相関を用いる方法も提案されている [3, 4]。

さらに、原理的には、アフィン変形 $A\mathbf{x} + \mathbf{d}$ が生じる場合などにも拡張可能であるが、このような直接移動・変形させて全探索する方法に要する計算量は、現在の計算機の実力では現実的でない。

以上のような領域探索法は、一画素単位の探索となり、あまり精度が良くないという問題点もあり、現在のところは、次節で述べる Lucas-Kanade のアルゴリズムに基づくものが現実的である。

2.2 Lucas-Kanade のアルゴリズム

様々な発展や変形がなされているが、基本となる Lucas-Kanade のアルゴリズム [5] を概観する。次の重み付き 2 乗誤差和を最小化するような d, A を発見することにより、追従を行う。

$$\epsilon = \int_{\mathcal{W}} [J(A\mathbf{x} + \mathbf{d}) - I(\mathbf{x})]^2 w(\mathbf{x}) dx \quad (5)$$

ここでは、アフィン変形 $A\mathbf{x} + \mathbf{d}$ を考え、 $w(\mathbf{x})$ はガウス関数などにより表される重み関数である。

説明の簡単化のため、特徴点の周辺領域が平行移動、回転、対称的な拡大・縮小だけを伴うと仮定

する．このとき， A は次のように表される．

$$\begin{aligned} A &= \begin{bmatrix} s_1 & 0 \\ 0 & s_1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \\ &\approx \mathbf{1} + \begin{bmatrix} s_0 & -\theta \\ \theta & s_0 \end{bmatrix} \\ &\equiv \mathbf{1} + D \end{aligned} \quad (6)$$

ここで， $\mathbf{1}$ は単位行列， $s_1 = 1 + s_0$ とおいている．このとき， $s_0 \ll 1$ ， $\theta \ll 1$ となる．また，2 次以上の項はここでは無視している．

変位 $D\mathbf{x} + \mathbf{d}$ が \mathbf{x} より十分に小さい時，輝度関数 J を \mathbf{x} の回りで近似線形化すると，

$$\begin{aligned} J(A\mathbf{x} + \mathbf{d}) &\approx J(\mathbf{x} + D\mathbf{x} + \mathbf{d}) \\ &\approx J(\mathbf{x}) + \frac{\partial J}{\partial \mathbf{x}}(D\mathbf{x} + \mathbf{d}) \end{aligned} \quad (7)$$

となり，式 (5) は次のように書ける．

$$\epsilon \approx \int_{\mathcal{W}} \left[J(\mathbf{x}) - I(\mathbf{x}) + \frac{\partial J}{\partial \mathbf{x}}(D\mathbf{x} + \mathbf{d}) \right]^2 w(\mathbf{x}) d\mathbf{x} \quad (8)$$

そこで，式 (8) を最小化 (極小化) する $\mathbf{z} = [\theta, s_0, d_x, d_y]^T$ を求めるのであるが，これは式 (8) を $\mathbf{z} = [\theta, s_0, d_x, d_y]^T$ で偏微分するとゼロとなる場合ということであるから，次の関係式が得られる．

$$G\mathbf{z} = \mathbf{e} \quad (9)$$

ここで，

$$\begin{aligned} G &= \int_{\mathcal{W}} \begin{bmatrix} G_a & G_b \\ G_b^T & G_c \end{bmatrix} w(\mathbf{x}) d\mathbf{x} \\ \mathbf{e} &= \int_{\mathcal{W}} [I(\mathbf{x}) - J(\mathbf{x})] \begin{bmatrix} -g_x y + g_y x \\ g_x x + g_y y \\ g_x \\ g_y \end{bmatrix} w(\mathbf{x}) d\mathbf{x} \end{aligned}$$

$$\begin{aligned} G_a &= \begin{bmatrix} (-g_x y + g_y x)^2 & (-g_x y + g_y x)(g_x x + g_y y) \\ (-g_x y + g_y x)(g_x x + g_y y) & (g_x x + g_y y)^2 \end{bmatrix} \\ G_b &= \begin{bmatrix} (-g_x y + g_y x)g_x & (-g_x y + g_y x)g_y \\ (g_x x + g_y y)g_x & (g_x x + g_y y)g_y \end{bmatrix} \\ G_c &= \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \end{aligned}$$

ただし， $g_x = \frac{\partial J}{\partial x}$ ， $g_y = \frac{\partial J}{\partial y}$ とする．

式 (9) を Newton 形式で繰り返し解くことにより，文献 [5, 6, 7] に示されているように式 (5) は最小化されていく．

この手法では，2 フレーム間の変化が十分に小さいと仮定したが，時間の経過を考えると大きくなるため，最初のフレームをテンプレートとして

用い，現在のフレームまでの追従結果を積算して変換する Warping が必要となる．この Warping の仕方には，順手法と逆手法があり，逆手法が可能である場合はテンプレートの G を求める計算が一度で済むという点や精度の面から，これを用いることが望ましい [8, 9]．また，テンプレートを固定せずに，直前のフレームをテンプレートとすることで，この Warping は必要ではなくなる．これがいわゆる KLT トラッカであり [6]，点の平行移動成分のみを求めるだけで良い場合には非常に有効である．ただし，平行移動以外の変形が大きい時には追従に失敗することも多く，さらに一度追従に失敗すると，回復することは出来ないため，ほぼ平行移動のみである場合に用いるか，変形の影響があまり現れないように，小さな領域の追従を考える必要がある．

この手法は，より一般的なアフィン変形の場合でも，先に述べたブロックマッチングアルゴリズムとは異なり，計算量は現実的なものである．ただし，変形が大きい時の追従性能は局所解に陥るなど完全ではないために，追従を二段階に分離して求めるアルゴリズム [10, 9] が提案されている．また，Lucas-Kanade のアルゴリズムに基づく手法の詳細な解析も行われている [8, 11, 12, 13]．

さらに，照明変化やオクルージョンに対応するためには，次節で述べるようなロバスト手法が必要となる．

2.3 ロバスト手法

今までに述べてきた手法は，照明変化やオクルージョンなどに対してあまりロバストでないことが多く，これを解決するためのロバスト手法が考えられている．基本的な一つの手法として，指標とする関数を SSD ではなく，外れ値の影響を低減する以下のようなものを用いることである [14, 15, 16]．

$$\epsilon = \int_{\mathcal{W}} \rho(J(\mathbf{x} + \mathbf{d}) - I(\mathbf{x}), \sigma) d\mathbf{x} \quad (10)$$

ここで， ρ は以下のような誤差 $e = J - I$ が大きくなると飽和するような関数である．

$$\rho(e, \sigma) = \frac{e^2}{e^2 + \sigma^2} \quad (11)$$

ただし， σ は閾値に相当するもので，これを適切に設定することで，大きな外れ値の影響を抑えることが可能となる．

Lucas-Kanade のアルゴリズムのように，輝度勾配に基いた非線形最適化手法を用いる場合にも，このようなロバスト化関数を利用する方法が提案されている [15, 17, 18]．また，式 (5) の重み w の重み付けや，テンプレートを分割して部分的に除去するなどにより，変動の影響を抑える方法も提案されている [19, 20]．

2.4 追従に適した特徴点の発見法

追従すべき領域を選ぶとき、領域内で輝度の差が少ないときなどに、開口問題 (Aperture Problem) が生じる。特に、平行移動のみならず、より一般的なアフィン変形を考える場合には、開口問題は起こりやすい。そこで、指定された領域内で、追従に適した特徴点を自動的に発見する方法は重要である [6, 7, 21, 9]。それぞれのアルゴリズムに対して、十分な輝度変化が含まれているかが基本となり、Lucas-Kanade のアルゴリズムに関しては、式 (9) の G の固有値が十分大きいということである [6]。

2.5 テンプレートの更新

テンプレートを固定している場合、変形が線形範囲に収まらなくなると、それ以上の追従が困難になる。そこで、テンプレートの更新を行う必要があるが、それに関する手法が考案されている [22, 23]。ただし、追従の成功・失敗の判断は難しく、どのタイミングで更新を行うかというのは非常に難しい問題である。

2.6 その他の発展

今までは、点の追従ということを考えてきたが、Lucas-Kanade のアルゴリズムは、点の追従のみならず、様々に発展している。その基本となる概念は、文献 [24] に詳しいが、ある関係や拘束条件を利用した何らかのパラメータ表現によるモデル化を行い、そのパラメータを求めていくというようなものである。例えば、coarse-to-fine アプローチによる階層的なモデル化 [24, 25] や、3次元風景の一部を構成する複数画像の統合である Image Mosaicing [26]、画像の固有空間とこの線形結合を利用して追従を行う Eigen Tracking [16]、従来ブロックマッチングが用いられていたオプティカルフロー [27]、Active Blob による非剛体物の追従 [28]、確率的なモデルである Active Appearance Model [29, 30] を用いた追従などである。

また、ブロックマッチングアルゴリズムの発展としては、リアルタイム性を考慮したアフィン変形の場合への対処法 [31] などが提案されている。

3 時系列フィルタリングに基づく追跡手法

3.1 問題の定式化

時系列フィルタリングは、もともと制御理論の分野で研究されてきた理論で、過去に観測された

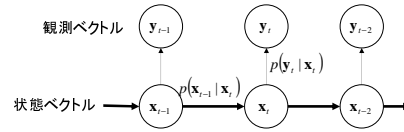


図 1: 状態遷移と観測ベクトル

信号系列から、未来の値を予測したり、ノイズを含む観測値から対象の状態を推定する問題などが取り扱われてきた。

時系列フィルタリングに基づく追跡技術では、対象をパラメトリックなモデルで表現し、画像そのものや画像から得られる特徴量を観測ベクトルとして、対象の位置、姿勢、変形などの状態を推定する。

一般的な時系列フィルタリングでは、時刻 $t, t+1$ の対象の状態ベクトル x_t, x_{t+1} と、観測ベクトル y_t の関係を次式のように定式化する。

$$x_{t+1} = f_t(x_t) + G_t w_t \quad (12)$$

$$y_t = h_t(x_t) + v \quad (13)$$

ここで、 $f_t(x_t)$ は、時刻 t と時刻 $t+1$ との状態変化を表わす状態遷移関数、 $h_t(x_t)$ は、ある状態のときに得られる観測値の関係を表わす観測関数、 w_t, v_t はそれぞれ、システムノイズ、観測ノイズである。

ここで、時刻 t までの観測値の集合を $Y_t = \{y_1, \dots, y_t\}$ としたとき、追跡問題は x_t の事後確率分布 $p(x_t | Y_t)$ や、 x_t の最小分散推定量 \hat{x}_t を推定する問題として定式化できる。 \hat{x}_t は x_t の条件付き期待値

$$\hat{x}_t = \mathcal{E}(x_t | Y_t) = \int x_t p(x_t | Y_t) dx_t \quad (14)$$

として与えられる。

観測値 Y_t から状態 x_t の事後確率 $p(x_t | Y_t)$ を直接推定することは困難であるため、時系列フィルタリングでは、ベイズの定理と x_t のマルコフ性から逐次的な推定問題に置き換えて推定する (図 1)。事後確率 $p(x_t | Y_t)$ は、ベイズの定理により次式のように事前確率と尤度の積に置き換えられる。

$$\begin{aligned} p(x_t | Y_t) &= p(x_t | y_t, Y_{t-1}) = \frac{p(x_t, y_t, Y_{t-1})}{p(y_t, Y_{t-1})} \\ &= \frac{p(t | x_t, Y_{t-1}) p(x_t, Y_{t-1})}{p(y_t, Y_{t-1})} \\ &= \frac{p(y_t | x_t) p(x_t | Y_{t-1})}{p(y_t | Y_{t-1})} \end{aligned} \quad (15)$$

ここで、 $p(y_t | x_t)$ は、ある状態 x_t のときに、観測値 y_t を得る確率 (尤度) であり、式 (13) の観測モデルによって与えられる。また、分母の $p(y_t | Y_{t-1})$ は、状態 x_t とは無関係な項であり、実際の問題で

は $\int p(\mathbf{x}_t|\mathbf{Y}_t)d\mathbf{x}_t = 1$ となるような正規化定数 k_t と考えることができる。

$p(\mathbf{x}_t|\mathbf{Y}_{t-1})$ は、時刻 t における \mathbf{x}_t の事前確率であり、 \mathbf{x}_t のマルコフ性により、次式のように与えられる。

$$\begin{aligned} p(\mathbf{x}_t|\mathbf{Y}_{t-1}) &= \int p(\mathbf{x}_t, \mathbf{x}_{t-1}|\mathbf{Y}_{t-1})d\mathbf{x}_{t-1} \\ &= \int p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{Y}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{Y}_{t-1})d\mathbf{x}_{t-1} \quad (16) \\ &= \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{Y}_{t-1})d\mathbf{x}_{t-1} \end{aligned}$$

ここで、 $p(\mathbf{x}_{t-1}|\mathbf{Y}_{t-1})$ は時刻 $t-1$ における事後確率であり、 $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ は時刻 $t-1$ と t の間の状態推移確率であり、式 (12) によって与えられる。

時系列フィルタリングは、式 (15), (16) によって与えられる尤度 $p(\mathbf{y}_t|\mathbf{x}_t)$ と状態遷移確率 $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ を用いて各時刻の事後確率分布 $p(\mathbf{x}_t|\mathbf{Y}_t)$ を求める問題とすることができる。

時系列フィルタリングは制御理論や統計学の分野では古くから研究されているが、近年動画からの対象追跡問題にこれらの手法を応用した研究が報告されている。時系列フィルタリングでは、システムプロセス、観測プロセスがとに線形である場合には、カルマンフィルタと呼ばれる手法で、線形に事後分布を推定することができる。

追跡問題では、状態ベクトル \mathbf{x}_t は物体の位置や姿勢、速度、加速度などのパラメータであり、システムプロセス $f_t(\mathbf{x}_t)$ は物体の動作モデルを表わしており、多くの場合線形に記述できる。

一方、観測モデル $h_t(\mathbf{x}_t)$ は、物体のパラメータから、観測される画像や画像特徴の生成プロセスを表わしており、観測値として何を用いるかによって異なるが一般には非線形となる場合が多い。

以後、観測プロセスも線形と仮定し、カルマンフィルタを用いて追跡する手法と、観測プロセスが線形でない場合にも対応したモンテカルロフィルタを用いた手法を紹介する。

3.2 カルマンフィルタ

カルマンフィルタは、1960年に Kalman [32] が提案した状態の予測モデルの一つであり、過去の状態から現在の状態を予測するとともに、予測された状態と実測値との誤差を用いて予測値の補正を行う方法である。カルマンフィルタでは、システムプロセス $f_t(\mathbf{x}_t)$ と観測モデル $h_t(\mathbf{x}_t)$ が線形であり、システムノイズ \mathbf{w}_t 、観測ノイズ \mathbf{v}_t が白色のガウシアンであることを仮定している。

$f_t(\mathbf{x}_t)$, $h_t(\mathbf{x}_t)$ が線形な場合、式 (15), (16) は状態遷移行列 \mathbf{F}_t と観測行列 \mathbf{H}_t を用いて次式のよう

に書ける。

$$\mathbf{x}_{t+1} = \mathbf{F}_t\mathbf{x}_t + \mathbf{G}_t\mathbf{w}_t \quad (17)$$

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t \quad (18)$$

このとき、事前分布 $p(\mathbf{x}_t|\mathbf{Y}_{t-1})$ 、事後分布 $p(\mathbf{x}_t|\mathbf{Y}_t)$ はガウス分布となり、時刻 t における推定値 $\hat{\mathbf{x}}_{t/t} = E\{\mathbf{x}_t|\mathbf{Y}_t\}$ は次のように線形に計算できる。

$$\hat{\mathbf{x}}_{t/t} = \hat{\mathbf{x}}_{t/t-1} + \mathbf{K}_t[\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t/t-1}] \quad (19)$$

$$\hat{\mathbf{x}}_{t/t-1} = \mathbf{F}_{t-1}\hat{\mathbf{x}}_{t-1/t-1} \quad (20)$$

ここで、 \mathbf{K}_t はカルマンゲインと呼ばれ、次式で与えられる。

$$\mathbf{K}_t = \mathbf{P}_{t/t-1}\mathbf{H}_t^T[\mathbf{H}_t\mathbf{P}_{t/t-1}\mathbf{H}_t^T + \mathbf{R}_t]^{-1} \quad (21)$$

$$\mathbf{P}_{t+1/t} = \mathbf{F}_t\mathbf{P}_{t/t}\mathbf{F}_t^T + \mathbf{G}_t\mathbf{Q}_t\mathbf{G}_t^T \quad (22)$$

$$\mathbf{P}_{t/t} = \mathbf{P}_{t/t-1} - \mathbf{K}_t\mathbf{H}_t\mathbf{P}_{t/t-1} \quad (23)$$

ただし、 \mathbf{Q}_t , \mathbf{R}_t はそれぞれ、 \mathbf{w}_t , \mathbf{v}_t の共分散行列である。

カルマンフィルタは、単純な線形関数で予測が可能であり、計算量も少なく安定という利点を持つ一方で、観測関数が線形であり、事前分布、事後分布としては単峰性のガウシアンのみという強い制約がある。カルマンフィルタを追跡に利用する方法としては、過去の追跡結果から次のフレームの追跡結果 $\hat{\mathbf{x}}_{t/t-1}$ を予測し、その近傍から対象の位置を探索する方法が多数提案されている [33, 34, 35, 36]。これらの手法では、状態ベクトルを対象の位置と速度 (と加速度) を用いて $\mathbf{x}_t = [x_t, y_t, \dot{x}_t, \dot{y}_t]^T$ と表現し、状態遷移行列 F を次式のように与えて対象の動きの予測を行っている。

$$F = \begin{bmatrix} 1 & 0 & \delta t & 0 \\ 0 & 1 & 0 & \delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (24)$$

また、観測行列 H を次式のように与え、位置のみが観測されるとしている。

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (25)$$

文献 [35, 37] は、アフィン変換によるオブジェクトの変形を状態として扱う方法を提案し、文献 [38, 39] では、3次元の位置、姿勢と速度、角速度を状態として、3次元の位置姿勢の追跡を行っている。また、Blackら [40] は Homography を使って多数のカメラの座標を同一の座標に変換し、複数カメラを用いた追跡を行っている。また、過去の結果から未来の状態を推定できるという特性を用いて、オクルージョンに対応させる研究も報告されている [41, 42]。

また、カルマンフィルタの $f_t(x_t)$, $h_t(x_t)$ が線形であるという制限を緩め、これらの関数が滑かであるという制約のもとで、推定値近傍で線形に近似し、カルマンフィルタを適用する、拡張カルマンフィルタを用いた追跡も報告されている [43, 34, 39] .

3.3 モンテカルロフィルタリング

線形システムを扱うカルマンフィルタや、推定値近傍で線形近似を行う拡張カルマンフィルタに対して、式 (16) の積分をランダムサンプリングによるモンテカルロ積分によって計算することで、非線形のシステムを直接扱う方法が、1980 年代から統計学の分野ではさかんに研究されているが、1990 年代後半からこの方法を追跡に適用する研究が報告されている .

これら手法では、事後分布 $p(x_t|Y_t)$ 、事前分布 $p(x_t|Y_{t-1})$ を、ランダムに生成された仮説と、その仮説に対する尤度の組で離散的に表現する . 以下、Isard ら [44] の Condensation を例に、モンテカルロフィルタを用いた追跡の具体的な手順を説明する .

時刻 $t-1$ における状態 x_{t-1} のサンプル集合を $\{s_{t-1}^{(1)}, \dots, s_{t-1}^{(N)}\}$ 、それぞれのサンプルに対する尤度を $\{\pi_{t-1}^{(1)}, \dots, \pi_{t-1}^{(N)}\}$ とする . このとき次の時刻 t におけるサンプル集合を以下の手順で生成する .

まず、時刻 $t-1$ におけるサンプル $s_{t-1}^{(j)}$ を尤度 $\pi_{t-1}^{(j)}$ の比に従ってランダムに選択し、これを $s_t^{\prime(n)}$ と置く . 次に、推移確率 $p(x_t|x_{t-1} = s_t^{\prime(n)})$ に従って $s_t^{\prime(n)}$ をランダムに変化させ、時刻 t におけるサンプル $s_t^{(n)}$ とする . これを N 回くりかえすことによって、 N 個のサンプル集合を生成する . さらに次の時刻 $t+1$ におけるサンプルは、時刻 t のサンプル集合 $s_t^{(n)}$ と尤度 $\pi_t^{(n)} = p(y_t|x_t = s_t^{(n)})$ から決定する .

ある時刻 t における追跡結果は、次式のように尤度による重み付き平均によって期待値として求める .

$$\mathcal{E}(x_t) = \frac{\sum_{n=1}^N \pi_t^n s_t^{(n)}}{\sum_{n=1}^N \pi_t^n} \quad (26)$$

このようなランダムサンプリングによる状態分布の推定を行うことによって、ノンパラメトリックな状態分布を効率良く推定できるだけでなく、追跡結果の候補をサンプルとして多数保持することで、追跡結果に多様性をもたせることができ、類似の特徴を持つ物体が接近して、一時的に追跡に失敗しても復旧することが可能となる .

モンテカルロフィルタは、観測モデルとして線形でない任意のモデルを適用できる、対象の確率分布として非ガウス性、ノンパラメトリックな分布を表現できるなどの特徴から、近年、対象追跡に適用する手法が数多く提案されている . これらの手法は、Sequential Monte Carlo, Particle Filter

[45, 46, 47, 48, 49, 50, 51], Condensation[44, 52, 53, 54, 55, 56, 57] とも呼ばれている .

3.4 複数対象への拡張

モンテカルロフィルタリングを応用した追跡手法では、対象のパラメータをノンパラメトリックな分布として表現できる反面、複数対象を扱う場合には、単一対象に対する分布か、複数対象に対する分布かを区別することが困難であるという問題点が生じる . この問題を解決し、複数対象に対応させる手法が提案されている [58, 59, 60, 61] . Isard ら [58] は、対象の個数を状態変数に含め、対象の数が違う状態の尤度推定法を提案し、個数が未知の複数対象追跡を実現している . また、Tweed と Calway [61] は、複雑なシーンで 30 羽までの鳥の群の追跡を行っている .

3.5 計算量の削減

モンテカルロフィルタリングの欠点は、詳細かつ安定な追跡を実現するためには、多数のサンプルを生成する必要があり、速度が低下してしまう点である . このような欠点を解決するために、効率的なサンプリングの手法が提案されている . Isard ら [53] は、あらかじめ肌色検出によって求めた顔や手らしい領域に重点的にサンプリングを行う Importance sampling を提案している . また、MacCormick ら [62] は、密度の粗いサンプリングと詳細なサンプリングの 2 回のサンプリングによって、効率の良く推定する Partitioned Sampling と呼ばれる手法を提案している .

4 追跡に用いる特徴量による分類

4.1 重心の位置情報を用いる方法

追跡対象が剛体で、追跡の基準点 (例えば重心) の位置情報を安定して求めることができる場合、フレーム間で基準点を対応付けることにより追跡を行うことができる . しかし非常にシンプルな反面、他の対象への誤った対応付けが生じる可能性もある . そこでこのような誤対応を避けるために、通常は重心の位置、速度、加速度からなる 3 次の方程式で、平均ゼロの加速度項をノイズとして、位置情報が可観測なカルマンフィルタを用い、次フレームでの位置情報を推定しながら追跡を行うことが一般的である . 例えば、航空機に取り付けた自由移動カメラから得られた、背景の揺れのある映像中で移動対象を追跡するシステムなどでは $\alpha - \beta - \gamma$ トラッカーとも呼ばれ、古くから良く知られている [63] .

近年では、追跡領域のマッチングによる検証プロセスと、追跡領域の統合、分割プロセスをカルマンフィルタによる追跡のループの中に組み込んで、領域追跡のロバスト性を向上させた手法 [64] や、部分的なオクルージョンにも対応できるように、フィルタの係数に補正值 (定数) を掛けてモデルを拡張したもの [65]、追跡の枠組みに形状マッチングのプロセスも組み込んだ手法 [66, 67] などが数多く提案されている。基準点の選び方にも工夫が凝らされ、追跡対象の骨格を抽出し、その位置情報をフィルタの入力とすることにより先行車両の安定な追跡を実現した手法 [68] や、対象の重心を中心としてエッジコーナー検出器により抽出した複数点のエッジ位置情報をフィルタの入力とすることにより精度を向上させたもの [69, 70] などがある。この他に、複数対象の追跡を取り扱ったもの [42] や、セルフオクルージョンに対応するために、追跡対称をパッチと呼ばれる複数の小領域の集合として表現し、ロバスト M 推定アルゴリズムによって、各パッチからなる領域の大局的な動きを求め、線形カルマンフィルタによって各パッチのアフィン動き推定を行い、次フレームでの領域を予測する手法を用いた顔や手の追跡、航空映像からの移動物体の追跡において実現されている [71, 72, 73]。

例えば対象の重心のように少ない情報量であっても、カルマンフィルタのような時系列のモデルを用いることにより、安定な追跡を行なうことが出来る。しかしモデルに合わない動き、急加速、急カーブ、急停止には対応することが困難であるため、このようなフィルタは追跡対象の動きの性質を良く見極めて用いる必要がある。

4.2 アピアランスモデルを用いる方法

追跡対象の重心位置だけでなく、その姿勢をも追跡する必要がある場合、対象のコーナーエッジなどの特徴点から構成される 3 次元のモデルを用い、モデルマッチングによる追跡を行う方法が提案されている [74, 75]。

同様の手法を人の頭部の追跡に応用し、目、鼻、口、耳などを特徴点とする頭部モデルの追跡を行った手法も提案されている [76] が、さらに頭部特徴をよりよく抽出するために、Gabor Wavelet によって目、鼻、口、耳などの特徴点を抽出し、顔の 3 次元モデルの追跡を行う方法も提案されている [77]。これらの手法は追跡対象がほぼ剛体で、convex な形状を有し、常に十分多くの特徴点が得られる場合は有効な手法であると言える。

対象が自転運動を行うなど極端に大きなセルフオクルージョンが生じず、そして運動が 2 次元で近似できると仮定できる場合には、物体を 2 次元のモデルで表現し、演算コストを抑えて追跡を行う手法が提案されている [78, 79]。同様の条件下では、コーナーエッジモデルではなく、古典的なテンプレートマッチングの手法に、テンプレートの変

形、明るさの補正といった拡張を行い、追跡を行う手法が提案されている [80, 81]。

あらかじめ様々な方向から対象を撮影したビューベースのアピアランスモデルが用意できる場合、それらを用いてテンプレートを切り替えながら追跡を行うことができる。例えばこのような方法で、人の顔の位置と姿勢の追跡を行う手法が提案されている [82]。

4.3 動き (フロー) 情報を用いる方法

対象のアピアランスが追跡の途中で大きく変わる場合、その全てをあらかじめモデルとして準備することは容易ではない。しかし対象の移動速度がおよそ一定で、急変しない場合には動きをベースとしたセグメンテーションが可能である。

例えば、対象のエッジマップを時間的に伝播して正確なセグメンテーションを行い、領域輪郭の射影の累積分布により中心位置を正確に出すことにより追跡を行なう手法が提案されている [83]。さらに、単一の動きモデルでは不整合が生じる領域でのみ境界の再計算を行って安定な追跡を実現した手法 [84]、モフォロジカルなモーションフィルタを用いてオブジェクト形状を整形しながら追跡する手法 [85]、動きのパラメータをそのベイズ則に従う信頼度に基づいて次のフレームで予測しながら追跡をおこなうもの [86] などが提案されている。

これらの手法は、サイズが極端に小さな対象の追跡は不得意である。

4.4 色、エッジなどの情報を用いる方法

追跡対象の形状やアスペクトの変形が大きな場合でも、その色ヒストグラムやエッジ方向のヒストグラムなどが安定して得られ、背景領域との間で明瞭な差がある場合、これらの特徴を用いて追跡を行なうことができる。

例えば、色を用いた手法では、画像中の各画素の色 (R,G,B) と位置 (i,j) からなる特徴空間でクラスタリングを行い、追跡を行なう手法 [87]、色ヒストグラムの違いによって背景から移動物体を切り出し、重心位置と幅と高さを更新しながら追跡する手法 [88]、chromatic plane に変換したカラー情報と、面積、隣接四角形の充填率からなる特徴量のマッチングにより追跡を行う手法 [89]、 Y, r, θ で構成される色空間でオンライン学習を行い、重心位置を補正しながら追跡する手法 [90]、HSI カラー系で B スプラインカーブを用いたクラスタリングを行いながら追跡を行なう手法 [91]、輝度変化に対してオンラインでモデルを修正しながら色ヒストグラムマッチングを行う手法 [92] など、様々な色空間の特徴を用いた方法が提案されている。

一方、エッジ特徴を用いたものでは、RGB 空間でそれぞれ求めた x 方向 y 方向の勾配を用いてエッ

ジの方向を求め、10度毎に36分割されたエッジヒストグラムを特徴量として追跡を行うものが提案されている [93]。

いずれも、サイズが小さな対象の追跡は、色やエッジのヒストグラムが統計的に十分な量だけ得られないため、不得意であると言える。

4.5 パラメトリックなモデルを用いる方法

追跡対象の輝度分布や色ヒストグラム分布がドラスティックに変化しない、あるいは背景の同等の分布が急変しないという条件下では、追跡対称のこれらの特徴量の分布にパラメトリックなモデルを当てはめ、そのモデルパラメータをEMアルゴリズムによってフレーム間で推定することにより、セグメンテーションを行ないながら、追跡を行なうことができる。モデルとしては、重心と広がりをパラメータとする2次元正規分布を重み付け加算した、混合正規分布が用いられることが多い。

例えば、入力画像をHSI色空間に変換し、二次元のH、S座標系で混合正規分布によるクラスタリングを行い、EMアルゴリズムに従って追跡を行なう手法 [94] や、色ヒストグラムを特徴とする手法 [95, 96] などが提案されている。

さらに発展させて、オブジェクトの形状や動きのダイナミクスといった事前知識を取り込んだ手法 [97] や、オブジェクト毎にレイヤーを当てはめて動きの推定と追跡を行う際に、レイヤーの二次元の動き、画素毎のレイヤーへのセグメンテーション、アピアランスの拘束条件を用いて、カメラ自身が移動する航空画像から複数の移動物体の追跡を行う手法 [98, 99] が提案されている。

また、色情報だけでなく、重心位置およびGabor Wavelet 関数で表現されるテクスチャの情報をパラメトリック推定する手法 [100] や、有限の混合隠れマルコフモデル (HMMs) をEMアルゴリズムを用いて、各シーケンスをk-means法に従って確率密度の重み付けを行い、複数のHMMに割り当てた手法 [101]、オプティカルフロー、カルマンフィルタによる動きのモデル、そして重み付きのテンプレートマッチングを特徴とするアピアランスモデルをパラメトリック推定する追跡手法 [102] などが提案されている。

4.6 輪郭情報を用いる方法

照明条件の変化などによって、追跡対称や背景の輝度や色の分布は追跡途中で変化する場合があります。しかし、そのような場合でも、対象が画像中で十分に大きな面積を占めている場合は、その輪郭形状を比較的安定に求めることが出来る。対象が完全な剛体でない場合でも、その変動が極端に大きくなければ、Condensation, あるいはParticle

Filtering の手法を用いて、追跡対象の輪郭形状をパラメータ空間における確信度の分布で表し、確信度分布の伝播と再構成を繰り返す事により、輪郭形状を推定しながら移動する対象を追跡することができる。

例えば、手や人の頭部など曲線で囲まれた領域の追跡に適用した例では、カルマンフィルタよりも精度良く追跡することができるとされている [103]、またこの手法の発展させて、出力を平滑化したもの [58]、さらに importance sampling と呼ばれる統計的な手法を用いて高精度と高速処理を両立させた手法 [104] が提案されている。

4.7 認識処理の結果を用いる方法

追跡の枠組みの中に、対象そのものを画像中から抽出する、学習器の概念を取り入れたものも近年提案されている。

例えば物体追跡を抽出された特徴のフレーム間での一致度に基づくアフィン変換式を求めることであると見え、不感帯のマージンを徐々に狭めながらサポートベクターのくり返し計算を行うことにより、誤った特徴量のマッチング、すなわちアウトライヤは自動的に抽出されて除外される。この手法では理論的にはアウトライヤが全データの50%以上に達する場合でも、それらをすばやく取り除くことが出来るとされている [105]。

また、初期フレームでは全てのサポートベクターを用いてSVMのスコアを計算し、追跡過程においては、小さな可変の部分サポートベクターのみを用いてスコアの更新を行う手法が提案されている。このとき、全てのサポートベクターを同じサイズの部分サポートベクターに分解してサイクリックに用いる方法と、参考テストパターンとの内積が前のフレームで最も大きく変化したサポートベクターを選択する方法とを組合せた手法を用いることにより、毎フレームSVMのスコアを再計算するよりも少ない処理演算量で安定な対象追跡を実現することができ、車載カメラからの先行車のリアルタイム追跡を実現している [106]。

4.8 Mean Shift Vector を用いる方法

近年、追跡対象の、次フレームの予測位置における色ヒストグラム(確率密度分布)を用いて、勾配法に類似した「場のポテンシャル」を定義して、相関を最大にする位置ずれ量 (Mean Shift Vector) を山登り法によって求めて追跡を行う手法が良く用いられている。2章で述べた追跡対称のモデルの類似度探索でも、また3章で述べた動きのモデルとも違う別の概念から成り立つ追跡手法であるため、追跡対象の極端に大きな形状変形に強く、またフレーム間でのダイナミクスのモデルの伝播が無いため、(カーネルサイズによって定義される) 探

索範囲を越えない限り、極端な急加速、急減速、急停止、急カーブにも対応して追跡が可能である。

この追跡手法の枠組みである Mean-Shift という概念自体は既に 1975 年の Fukunaga と Hostetler の論文 [107] で登場しているが、近年、Comaniciu らによって数学的な解析が行われ、非剛体の追跡に有効であることが示された [108]。この手法はさらに発展し、色ヒストグラムの特徴量を用いてモデル化する際に、等方性の空間フィルタカーネルを重みとして畳み込むことによって、空間的な平滑化を行い、勾配法 (山登り法) による局所解を得易くした手法が提案されている [109]。

同等のアルゴリズムを用いた、CAMSHIFT と呼ばれる追跡手法では、ウィンドウサイズや色ヒストグラムを追跡過程においてフレームごとに動的に変化させることにより、顔領域の実時間トラッキングを実現し、3D ゲームのインターフェースを実現している [110, 111]。

さらに、Mean shift 追跡における重みに負の値が使えるように拡張して、スケールやオフセットに依存しないようにアルゴリズムを改修して、通常の Mean shift で問題となっていたカーネルサイズの設定をオンライン推定しながら追跡を行うことを可能とした手法が提案されている [112]。

また、さらなる追跡のロバスト化をめざして複数の手法と組合せた手法が提案されている。例えば、対象の向きや照明条件が変動する環境下での対象追跡を実現するために、SVM を用いた対象認識と、Mean shift を用いた追跡、そしてカルマンフィルタをベースとする赤外照明を組み合わせたシステムで、目のトラッキングを実現した手法 [105] や、人のグローバルな動きに着目するのではなく、人体の各パーツのアピアランスモデルに着目し複数人物の追跡を行う手法が提案され、人数カウント、個別追跡、一時的なオクルージョンからの復帰、大まかな体の姿勢推定が可能であるとされている [113]。

Mean shift 追跡は、演算処理量が少なく、高速に実行できる追跡アルゴリズムであるが、追跡対象の色ヒストグラムが急激に変動した場合、偽の Mean Shift Vector が求められ、追跡に失敗する場合がある。そこで、追跡対象と次フレームの色ヒストグラム類似性だけでなく、(統計的に生成した) 背景画像と次フレームの色ヒストグラムの非類似性をも評価関数に組み込んだ、haybrid mean shift が提案されている [114, 115]。この方法を用いると、追跡対象の色ヒストグラムが背景のそれと同一にならない限り、追跡を維持することが可能となる。

また、モノクロ画像や、夜間での使用、あるいは照明条件によって有効な色ヒストグラムが得られない場合は、原理的に Mean Shift Vector を正しく求めることが出来ない。しかし、この点を改善し、赤外線画像における対象追跡の輝度値と Gabor フィルタの応答値を特徴量とするテクスチャ情報を組合せ、Mean Shift Vector を求める手法も提案されている [116]。

5 まとめ

本稿では、1990 年代以降のリアルタイム処理を目指した対象追跡の研究について、フレーム間の類似度の勾配に基づく手法、時系列フィルタリングに基づく手法を重点的に紹介した。また、追跡に基づく特徴量という観点から、様々な追跡技術を分類し紹介した。

本稿では、対象をフレーム間で対応付ける方法に着目して、その枠組の解説を含めて関係論文を重点的に紹介した。そのため、追跡分野を網羅的に紹介するという観点では不足する点も多いと思われるが、その点に関してはご容赦いただきたい。また、不足する点を著者らまで御指摘頂ければ幸いです。

参考文献

- [1] Lesse, J., Novak, C. and Taylor, V.: An Automated Technique for Obtaining Cloud Motion from Geosynchronous Setellite Data Usin Cross-Correlation, *Journal of Appl. Meteorology*, Vol. 10, pp. 118–132 (1971).
- [2] Rosenfeld, A. and Kak, A. C.: *Digital Picture Processing*, Vol. 1, Academic Press, 2nd edition (1982).
- [3] Crowley, J. and Martin, J.: Experimental comparison of correlation techniques, in *International Conference on Intelligent Autonomous Systems* (1995).
- [4] Fitch, A., Kadyrov, A., Christmas, W. and Kittler, J.: Orientation correlation, in *British Machine Vision Conference*, pp. 133–142 (2002).
- [5] Lucas, B. and Kanade, T.: An iterative image registration technique with an application to stereo vision, in *IJCAI* (1981).
- [6] Tomasi, C. and Kanade, T.: Detection and tracking of point features, Technical Report CMU-CS-91-132, Carnegie Mellon University (1991).
- [7] Shi, J. and Tomasi, C.: Good features to track, in *IEEE CVPR* (1994).
- [8] Baker, S. and Matthews, I.: Lucas-Kanade 20 years on: A unifying framework: Part 1, Technical Report CMU-RI-TR-02-16, Carnegie Mellon University (2002).
- [9] 深尾, 金出: 2 段階特徴点追従アルゴリズム, 第 141 回コンピュータビジョンとイメージメディア研究発表会 (2003).

- [10] Hager, G. and Toyama, K.: X Vision: A portable substrate for real-time vision applications, *Computer Vision and Image Understanding*, Vol. 69, No. 1, pp. 23–37 (1998).
- [11] Baker, S., Gross, R., Ishikawa, T. and Matthews, I.: Lucas-Kanade 20 years on: A unifying framework: Part 2, Technical Report CMU-RI-TR-03-01, Carnegie Mellon University (2003).
- [12] Baker, S., Gross, R. and Matthews, I.: Lucas-Kanade 20 years on: A unifying framework: Part 3, Technical Report CMU-RI-TR-03-35, Carnegie Mellon University (2003).
- [13] Baker, S., Gross, R. and Matthews, I.: Lucas-Kanade 20 years on: A unifying framework: Part 4, Technical Report CMU-RI-TR-04-14, Carnegie Mellon University (2004).
- [14] Black, M. and Anandan, P.: The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, *Computer Vision and Image Understanding*, Vol. 63, No. 1, pp. 75–104 (1996).
- [15] Hager, G. and Belhumeur, P.: Efficient region tracking with parametric models of geometry and illumination, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, No. 10, pp. 1025–1039 (1998).
- [16] Black, M. and Jepson, A.: EigenTracking: Robust matching and tracking of articulated objects using a view-based representation, in *ECCV '96* (1996).
- [17] Jurie, F. and Dhome, M.: Real time robust template matching, in *British Machine Vision Conference*, pp. 123–132 (2002).
- [18] Lai, S.: Robust image matching under partial occlusion and spatially varying illumination change, *Computer Vision and Image Understanding*, Vol. 78, pp. 84–98 (2000).
- [19] Ishikawa, T., Matthews, I. and Baker, S.: Efficient image alignment with outlier rejection, Technical Report CMU-RI-TR-02-27, Carnegie Mellon University (2002).
- [20] 深尾, 金出: オクルージョンに対してロバストな2段階特徴点追従アルゴリズム, 第141回コンピュータビジョンとイメージメディア研究発表会 (2003).
- [21] Tommasini, T., Fusiello, A., Trucco, E. and Roberto, V.: Making good features track better, in *IEEE CVPR* (1998).
- [22] Seelen, U. C. and Bajcsy, R.: Adaptive correlation tracking of targets with changing scale, Technical Report MS-CIS-96-22/GRASP LAB 405, University of Pennsylvania (1996).
- [23] Matthews, I., Ishikawa, T. and Baker, S.: The template update problem, *IEEE PAMI*, Vol. 26, No. 6, pp. 810–815 (2004).
- [24] Bergen, J., Anandan, P., Hanna, K. and Hingorani, R.: Hierarchical model-based motion estimation, in *2nd European Conference on Computer Vision*, pp. 237–252 (1992).
- [25] Szeliski, R. and Coughlan, J.: Hierarchical spline-based image registration, in *CVPR '94* (1994).
- [26] Szeliski, R.: Image mosaicing for tele-reality applications, in *WACV '94* (1994).
- [27] Bainbridge-Smith, A. and Lane, R.: Determining optical-flow using a differential method, *IVC*, pp. 11–22 (1997).
- [28] Sclaroff, S. and Isidoro, J.: Active blobs, in *ICCV '98* (1998).
- [29] Cootes, T., Edwards, G. and Taylor, C.: Active appearance model, in *ECCV '98* (1998).
- [30] Baker, S., Matthews, I. and Schneider, J.: Image coding with active appearance models, Technical Report CMU-RI-TR-03-13, Carnegie Mellon University (2003).
- [31] 張, 清水, 奥富: 画像のレジストレーションにおける同時推定法の高速化手法, 第141回コンピュータビジョンとイメージメディア研究発表会, pp. 51–58 (2005).
- [32] Kalman, R.: A new approach to linear filtering and prediction problems, *Trans. ASME – Journal of Basic Engineering*, pp. 35–45 (1960).
- [33] Bar-Shalom, Y. and Fortmann, T. E.: *Tracking and Data Association*, Academic Press (1988).
- [34] Wang, H., Goh, W. L., Chua, C. S. and Sim, C. T.: Real-time object tracking, in *21st International conference on industrial electronics*, pp. 1366–1370 (1995).

- [35] Reid, I. D. and Murray, D. W.: Active tracking of foveated feature clusters using affine structure, *International Journal of Computer Vision*, Vol. 18, pp. 41–60 (1996).
- [36] Ziliani, F. and Moscheni, F.: Kalman filtering motion prediction for recursive spatio-temporal segmentation and object tracking, in *Workshop on Image Analysis for Multimedia Interactive Services* (1997).
- [37] Manku, G. S., Jain, P., Aggarwal, A., Kumar, L. and Banerjee, S.: Object Tracking using Affine Structure for Point Correspondences, in *Computer Vision and Pattern Recognition (CVPR'97)*, pp. 704–708 (1997).
- [38] Jung, S. K. and Wohn, K. Y.: 3-D tracking and motion estimation using hierarchical Kalman filter, in *Vision, Image and Signal Processing*, pp. 293–298 (1997).
- [39] Jebara, T., Russell, K. and Pentland, A.: Mixtures of eigenfeatures for real-time structure from texture, in *6th International Conference on Computer Vision*, pp. 128–135 (1998).
- [40] Black, J., Ellis, T. and Rosin, P.: Multi View Image Surveillance and Tracking, in *Workshop on Motion and Video Computing*, pp. 169–174 (2002).
- [41] Nguyen, H. T., Worring, M. and Boomgaard, R.: Occlusion robust adaptive template tracking, in *International Conference on Computer Vision (ICCV'01)*, pp. 678–682 (2001).
- [42] Loutas, E., Diamantaras, K. and Pitas, I.: Occlusion Resistant Object Tracking, in *IEEE Symposium on Signal Processing Information Technology (ISSPIT)*, pp. 65–67 (2001).
- [43] Harris, C.: Tracking with rigid models, in Blake, A. and Yuille, A. eds., *Active Vision*, MIT Press, pp. 59–74 (1992).
- [44] Isard, M. and Blake, A.: Visual tracking by stochastic propagation of conditional density, in *4th European Conference Computer Vision (ECCV1996)*, pp. 343–356 (1996).
- [45] Gordon, N. and Salmond, D.: Bayesian state estimation for tracking and guidance using the bootstrap filter, *Journal of guidance, control and dynamics*, Vol. 18, No. 6, pp. 1434–1443 (1995).
- [46] Carpenter, J., Clifford, P. and Fearnhead, P.: An improved particle filter for non-linear problems, in *Technical report, Dept of Statistics, University of Oxford* (1997).
- [47] Pitt, M. and Shepherd, N.: Filtering via simulation and auxiliary particle filters, in *Technical report, Nuffield College, University of Oxford* (1997).
- [48] Hue, C. and Cadre, J.-P. L.: A particle filter to track multiple objects – Multi-Object Tracking, in *WOMOT2001*, p. 61 (2001).
- [49] Torma, P. and Szepesvari, C.: Efficient object tracking in video sequences by means of LS-N-IPS, in *ISPA2001*, pp. 277–282 (2001).
- [50] Rui, Y. and Chen, Y.: Better proposal distributions: object tracking using unscented particle filter, in *CVPR2001*, pp. 786–793 (2001).
- [51] Bolio, M., Hong, S. and Djurio, P. M.: Performance and Complexity Analysis of Adaptive Particle Filtering for Tracking Applications, in *Signals Systems and Computers*, pp. 853–857 (2002).
- [52] Isard, M. and Blake, A.: CONDENSATION – Conditional Density Propagation for Visual Tracking, *International Journal of Computer Vision*, Vol. 29, No. 1, pp. 5–29 (1998).
- [53] Isard, M. and Blake, A.: ICondensation: Unifying low-level and high-level tracking in a stochastic framework, in *5th European Conference Computer Vision (ECCV1998)*, pp. 893–908 (1998).
- [54] Isard, M. and Blake, A.: A smoothing filter for CONDENSATION, in *5th European Conference on Computer Vision*, Vol. 1, pp. 767–781 (1998).
- [55] Chang, T., Gong, S. and Ong, E.: Tracking Multiple People Under Occlusion Using Multiple Cameras, in *British Machine Vision Conference*, Vol. 2, pp. 566–575 (2000).
- [56] Kang, H., Kim, D. and Bang, S. Y.: Real-time Multiple People Tracking Using Competitive Condensation, in *IPIP 2002*, pp. 325–328 (2002).

- [57] Kang, H., Kim, D. and Bang, S. Y.: Real-time Multiple People Tracking Using Competitive Condensation, in *ICPR2002*, pp. 413–416 (2002).
- [58] Isard, M. and MacCormick, J.: BraMBLE: A bayesian multiple blob tracker, in *International Conference on Computer Vision (ICCV2001)*, Vol. 2, pp. 34–41 (2001).
- [59] MacCormick, J. and Blake, A.: A probabilistic exclusion principle for tracking multiple objects, *International Journal of Computer Vision*, Vol. 39, No. 1, pp. 57–71 (2000).
- [60] Tao, H., Sawhney, H. and Kumar, R.: A sampling algorithm for tracking multiple objects, in *IEEE Workshop on Vision Algorithms*, pp. 53–68 (1999).
- [61] Tweed, D. and Calway, A.: Tracking Many Objects Using Subordinated CONDENSATION, in *Proceedings of the British Machine Vision Conference*, pp. 283–292 (2002).
- [62] MacCormick, J. and Isard, M.: Partitioned sampling, articulated objects, and interface-quality hand tracking, in *European Conference on Computer Vision*, pp. 3–19 (2000).
- [63] Boyce, J. F. and Toulson, D. L.: Moving object tracking using camera motion corrected Kalman filtering, in *ICIP*, pp. 81–84 (1992).
- [64] Vigus, S. A., Bull, D. and Canagarajah, N.: Video object tracking using region split and merge and a kalman filter tracking algorithm, in *ICIP2001*, pp. 650–653 (2001).
- [65] Niu, W., Jiao, L., Han, D. and Wang, Y.-F.: Real-time multi-person tracking in video surveillance, in *ICICS-PCM2003* (2003).
- [66] Marcenaro, L., Ferrari, M., Marchesotti, L. and Regazzoni, C. S.: Multiple object tracking under heavy occlusions by using kalman filters based on shape matching, in *ICIP 2002*, pp. 341–344 (2002).
- [67] Cao, G., Jiang, J. and Chen, J.: An improved object tracking algorithm based on image correlation, in *ISIE 2003*, pp. 598–601 (2003).
- [68] Foresti, G. L.: Object recognition and tracking for remote video surveillance, in *IEEE 1999*, p. 1045 (1999).
- [69] Wang, H., Goh, W. L., Chua, C. S. and Sim, C. T.: Real-time object tracking, in *IEEE 1995*, pp. 1366–1371 (1995).
- [70] Manku, G. S., Jain, P., Aggarwal, A., Kumar, L. and Banerjee, S.: Object tracking using affine structure for point correspondences, in *CVPR1997*, p. 704 (1997).
- [71] Huang, Y., Huang, T. and Niemann, H.: Segmentation-based object tracking using image warping and kalman filtering, in *ICIP 2002*, pp. 601–604 (2002).
- [72] Karlsson, H. and Jonas, N.: Robust and efficient tracking in image sequences using a kalman filter and an affine motion model, in *RSJ 2002*, pp. 98–104 (2002).
- [73] Gao, J.: Self-occlusion immune video tracking of objects in cluttered environments, in *AVSS 2003*, p. 79 (2003).
- [74] Jurie, F.: Model-based object tracking in cluttered scenes with occlusions, in *IROS 1997*, pp. 886–892 (1997).
- [75] Ayromlou, M., Vincze, M. and Ponweiser, W.: Probabilistic matching of image- to model-features for real-time object tracking, in *ICPR 2002*, pp. 692–695 (2002).
- [76] Gao, J. and Kak, A.: Multi-frame based motion estimation for semantic object tracking in the presence of occlusion, in *ICIP 2002*, pp. 881–884 (2002).
- [77] He, C., Zheng, Y. F. and Ahalt, S. C.: Object tracking using the gabor wavelet transform and the golden section algorithm, pp. 528–538 (2002).
- [78] Matsuzawa, Y. and Kumazawa, I.: Coarse-to-fine object tracking using a shape representation network with continuous parameters determining shape details, in *IJCNN 2000*, pp. 273–278 (2000).
- [79] Oberti, F., Calcagno, S., Zara, M. and Regazzoni, C. S.: Robust tracking of humans and vehicles in cluttered scenes with occlusions, in *ICIP 2002*, Vol. III, pp. 629–632 (2002).
- [80] Jurie, F. and Dhome, M.: A simple and efficient template matching algorithm, in *ICCV 2001*, pp. 554–549 (2001).

- [81] Chalana, T., Haynor, V. and Kim, Y.: Video object tracking with a sequential hierarchy of template deformations, pp. 1171–1182 (2001).
- [82] L. P. Morency, A. R. and Darrell, T.: Adaptive view-based appearance models, Vol. I, pp. 803–810 (2003).
- [83] Lee, K. W., Ryu, S. W., Lee, S. J. and Park, K. T.: Motion based object tracking with mobile camera, pp. 256–258 (1998).
- [84] Lim, J., Cho, H. K. and Ra, J. B.: An improved video object tracking algorithm based on motion re-estimation, in *ICIP 2000*, pp. 339–342 (2000).
- [85] Shi, L., Zhang, Z. and An, P.: Automatic segmentation of video object plane based on object tracking and matching, pp. 510–513 (2001).
- [86] Kumar, A. and Gupta, S.: A novel probabilistic approach for real time motion segmentation and tracking, in *ISSPA*, pp. 136–139 (2001).
- [87] Heisele, B., Kressel, U. and Ritter, W.: Tracking non-rigid, moving objects based on color cluster flow, in *Computer Vision and Pattern Recognition*, pp. 257–260 (1997).
- [88] Grove, T. D., Baker, K. D. and Tan, T. N.: Colour based object tracking, in *Pattern Recognition*, pp. 1442–1444 (1998).
- [89] Bakowski, A. and Jones, G. A.: Video surveillance tracking using colour region adjacency graphs, pp. 794–798 (1999).
- [90] Nakamura, T. and Ogasawara, T.: On-line visual learning method for color image segmentation and object tracking, in *IROS1999*, pp. 222–228 (1999).
- [91] Lee, Y.-B., Youg, B.-J. and Lee, S.-W.: A real-time color-based object tracking robust to irregular illumination variations, in *ICRA*, pp. 1659–1664 (2001).
- [92] Chen, T. M., Luo, R. C. and Hsiao, T. H.: Visual tracking using adaptive color histogram model, in *IECON '99*, pp. 1336–1341 (1999).
- [93] Mason, M. and Duric, Z.: Using histograms to detect and track objects in color video, in *AIPR 2001*, pp. 154–159 (2001).
- [94] McKenna, S. J., Raja, Y. and Gong, S.: Tracking colour objects using adaptive mixture models, pp. 225–231 (1999).
- [95] Withagen, P. J., Schutte, K. and Groen, F.: Likelihood-based object detection and object tracking using color histograms and EM, in *ICIP 2002*, pp. 589–592 (2002).
- [96] Withagen, P. J., Schutte, K. and Groen, F.: Likelihood-based object detection and object tracking using color histograms and EM, in *ICIP2002*, pp. 589–592 (2002).
- [97] Chen, Y. and Huang, T. S.: Hierarchical MRF model for model-based multi-object tracking, in *ICIP 2001*, pp. 385–388 (2001).
- [98] Tao, H., Sawhney, H. S. and Kumar, R.: Dynamic layer representation with applications to tracking, in *CVPR 2000*, p. 2134.
- [99] Tao, H., Sawhney, H. S. and Kumar, R.: Object tracking with bayesian estimation of dynamic layer representations, in *IEEE2002*, pp. 75–89 (2002).
- [100] Rares, A. and Reinders, M. J. T.: Object tracking by adaptive modeling, in *IEEE 2000*, pp. 74–77 (2000).
- [101] Alon, J., Sclaroff, S., Kollios, G. and Pavlovic, V.: Discovering clusters in motion time-series data, in *CVPR 2003* (2003).
- [102] Han, T. X., Huang, T. S. and Rui, Y.: Tracking by online appearance learning and dynamic model, in *ICME 2004* (2004).
- [103] Blake, A. and Isard, M.: Conditional density propagation and application to visual tracking, in *IJCV98* (1998).
- [104] Isard, M. and Blake, A.: Unifying low-level and high-level tracking in a stochastic framework, in *ECCV'98*, pp. 893–908 (1998).
- [105] Zhu, W., Wang, S., Lin, R.-S. and Levinson, S.: Tracking of object with SVM regression, in *CVPR 2001*, Vol. II, pp. 240–245 (2001).
- [106] Avidan, S.: Subset selection for efficient SVM tracking, in *Computer Vision and Pattern Recognition*, Vol. I, pp. 85–92 (2003).

- [107] Fukunaga, K. and Hostetler, L.: The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition, *IEEE Trans. Information Theory*, Vol. 21, pp. 32–40 (1975).
- [108] Comaniciu, D., Ramesh, V. and Meer, P.: Real-time tracking of non-rigid objects using mean shift, in *CVPR2000*, pp. 142–149 (2000).
- [109] Comaniciu, D., Ramesh, V. and Meer, P.: Kernel-based object tracking, pp. 564–577 (2003).
- [110] Bradski, G. R.: Computer vision face tracking for use in a perceptual user interface (1998).
- [111] Boyle, M.: The effects of capture conditions on the camshift face tracker (2001).
- [112] Collins, R. T.: Mean-shift blob tracking through scale space, in *CVPR 2003*, p. 234 (2003).
- [113] Ramanan, D.: Finding and tracking people from the bottom up, in *CVPR2003* (2003).
- [114] Porikli, F. and Tuzel, O.: Object tracking in low-frame-rate video, in *SPIE- Image and Video Communications and Processing Conference*, Vol. III (2005).
- [115] Zhao, T. and Nevatia, R.: Tracking multiple humans in crowded environment, in *CVPR 2004*, pp. 406–413 (2004).
- [116] Yilmaz, A., Shafique, K., Lobo, N., Li, X., Olson, T. and Shah, M. A.: Target-tracking in flir imagery using mean-shift and global motion compensation, in *CVBVS2001* (2001).