

## ■2 群 (画像・音・言語) - 2 編 (パターン認識とビジョン)

# 4 章 動画解析

(執筆者：藤吉弘亘) [2010 年 9 月受領]

### ■ 概要 ■

動画像の理解には、動画中の人や物体の動きを求める手法や対象物の形状を推定する動画解析が必要不可欠である。この動きを求める手法としてオプティカルフローがある。オプティカルフローは、画像上のある点の動きを表したベクトルであり、動画像解析のほか、形状復元、領域抽出、対象追跡などの前処理として用いられる。オプティカルフローでは、点に対する動きであったが、それを領域や対象物の動きに拡張したものがトラッキングとなる。トラッキングでは、オプティカルフローと同様に前処理として用いられるだけでなく、特定の対象物のトラッキングを行うことができるため、監視の分野に応用されている。

これらのオプティカルフローやトラッキング処理により、対象物の動画中の位置や動きを算出することができるようになり、フレーム間における同一対象物の対応が必要な 3 次元運動復元に応用することができる。そのため、追跡した対象物の 3 次元形状を復元することができる。

### 【本章の構成】

本章ではオプティカルフロー (4-1 節) として勾配法、勾配法の改良、Lucas-Kanade 法について述べ、, トラッキング (4-2 節) としてテンプレートマッチング、アクティブ探索による物体追跡、Mean Shift、パーティクルフィルタについて述べ、3 次元運動復元 (4-3 節) として射影行列からの運動復元、2 画像からの運動復元、動画像からの運動復元、バンドル調整による中心射影画像からの運動復元について述べる。

## ■ 2 群 - 2 編 - 4 章

## 4-1 オプティカルフロー

(執筆者: 大塚和弘) [2010 年 9 月 受領]

## 4-1-1 オプティカルフローとは

オプティカルフロー (Optical Flow) とは、カメラと外界との相対的運動によって生ずる画像上での見かけの動きのことであり、時刻  $t$  の画像と微小時間後  $t + \delta t$  の画像との間での点対応に相当する。オプティカルフロー (または、単にフロー) という場合、画像上のある 1 点での速度ベクトルを指す場合もあれば、画像上における密な速度ベクトル場を指す場合もある。オプティカルフローの計算は、動画像解析、形状復元、領域抽出、対象追跡など様々なタスクの前処理として重要な役割をもつ。なお、類似の用語である運動場 (Motion Field) は、3 次元空間中に存在する対象表面上の 3 次元速度ベクトルを画像平面に射影して得られる速度場を指すものであり、オプティカルフローとは異なる。一方、オプティカルフローは、画像から解釈可能な見かけの動きであり、被写体の物理的な運動とは必ずしも一致しない。また、観測画像からオプティカルフローが一意に定められないという性質もある。これまで数多くオプティカルフロー計算法が提案されている<sup>1)</sup>。本節では、代表的な方法として、勾配法<sup>2)</sup>とその改良<sup>3)</sup>、及び、Lucas-Kanade 法<sup>4)</sup>を紹介する。

## (1) 勾配法

勾配法では、画像輝度の空間的・時間的な勾配を利用してオプティカルフローを計算する方法である。勾配法では、対象の動きによって画像間での対応点の輝度は変わらないことを仮定する。この仮定は輝度の不変性と呼ばれる。いま、時刻  $t$  における画像上の位置  $(x, y)$  の輝度を  $I(x, y, t)$  と記す。この点  $(x, y)$  が微小時間後  $t + \delta t$  において  $(\delta x, \delta y)$  だけ変位したとすると、画像輝度の不変性は、

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

と書ける。ここで画像間での変位量が小さいことを仮定して、画像  $I(x, y, t)$  についてテイラー展開を行うと次式が得られる。

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t + H.O.T.$$

ただし、 $H.O.T.$  は高次の項を表す。ここで高次の項が無視できるほど小さいと仮定することで、オプティカルフロー拘束式が次式のように導かれる。

$$\begin{aligned} \frac{\partial I}{\partial x} \frac{\delta x}{\delta t} + \frac{\partial I}{\partial y} \frac{\delta y}{\delta t} + \frac{\partial I}{\partial t} \frac{\delta t}{\delta t} &= 0, \quad \text{or} \\ \frac{\partial I}{\partial x} u_x + \frac{\partial I}{\partial y} u_y + \frac{\partial I}{\partial t} &= 0. \end{aligned} \quad (4.1)$$

ここで、 $u_x, u_y$  は、オプティカルフローの  $x$  成分、 $y$  成分を表し、

$$u_x = \frac{\delta x}{\delta t}, \quad u_y = \frac{\delta y}{\delta t}$$

である。ここで式 (4.1) 中の偏導関数を

$$I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$$

と表すと、式 (4.1) のオプティカルフロー拘束式は次式のように書き直すことができる。

$$I_x u_x + I_y u_y = -I_t \quad \text{or} \quad \nabla I \mathbf{u} = -I_t. \quad (4.2)$$

ただし、 $\nabla I = [I_x \ I_y]$  である。また、 $\mathbf{u} = [u_x \ u_y]^T$  であり、これは画素  $(x, y)$  におけるオプティカルフローまたは速度ベクトルと呼ばれる。このオプティカルフロー拘束式には未知数が 2 個含まれ、一つのオプティカルフロー拘束式からはフローを一意に決定することができない。この問題は開口問題 (aperture problem) として知られている。開口問題は、局所的な輝度分布が完全なフローを決定するために十分な構造をもたないことに起因する。直感的には、図 4.1 のように直線的なエッジが移動する様子を覗き窓を通して観測する場合、その観測情報のみからは真の動きが決定できないこととして説明される。図 4.1 では、図中矢印で示したように動きについて多様な解釈が可能である。この場合、エッジの法線方向の動き成分 (ノーマルフローと呼ばれる) のみ得られ、エッジの接線方向の動き成分は一意に決定できない。この動きの曖昧性は、図 4.2 の速度空間上においてオプティカルフロー拘束式に対応する直線として示唆される。

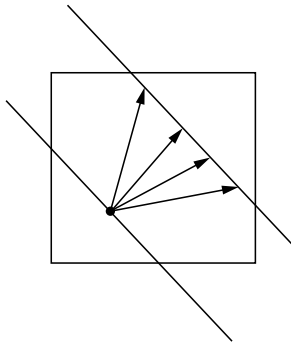


図 4.1 開口問題

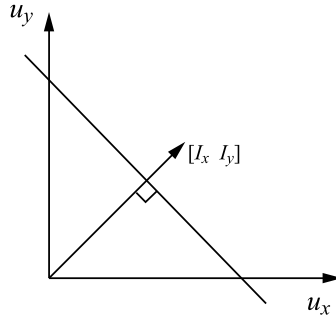


図 4.2 オプティカルフロー拘束式の速度空間上での解釈

勾配法では、この問題を回避し、解を一意に定めるため、輝度不変性より導かれるオプティカルフロー拘束式に加えて、空間的に速度ベクトルは滑らかに変化するという拘束条件を導入する。この二つの拘束条件それぞれに対応する評価関数を定義し、それらを足し合わせた評価関数の最適化問題として、オプティカルフロー推定の問題が定式化される。この評価関数は、エネルギー関数とも呼ばれる。この評価関数は一般に次式のように表される。

$$E(\mathbf{u}) = E_{\text{data}}(\mathbf{u}) + \alpha \cdot E_{\text{smooth}}(\mathbf{u}). \quad (4.3)$$

右辺第一項は、データ項と呼ばれ、輝度の不変性の制約条件を表し、第二項は、平滑化項と呼ばれ、速度場の滑らかさの制約条件を表す。ここで、 $\alpha$  は二つの制約の強さを調整するパラメータである。このように解が一意に定まらない場合に、付加的な拘束条件を導入することで安定解を求める手法は、一般に正則化と呼ばれる。ここで、オプティカルフロー評価関数のデータ項は、

$$E_{\text{data}}(\mathbf{u}) = \int_{\Omega} \left\| I_x u_x + I_y u_y + I_t \right\|^2 dx \quad (4.4)$$

のように定義される。ここで  $\Omega$  は、画像の全領域を表し、 $\mathbf{x} = [x \ y]^T$  は各画素を表す。また、 $\|\cdot\|^2$  は  $L2$  ノルムを表す。一方、平滑化項は、

$$E_{\text{smooth}}(\mathbf{u}) = \int_{\Omega} \left[ |\nabla u_x|^2 + |\nabla u_y|^2 \right] dx = \int_{\Omega} \left\{ \left( \frac{\partial u_x}{\partial x} \right)^2 + \left( \frac{\partial u_x}{\partial y} \right)^2 + \left( \frac{\partial u_y}{\partial x} \right)^2 + \left( \frac{\partial u_y}{\partial y} \right)^2 \right\} dx \quad (4.5)$$

のように定義される。式 (4.3) の評価関数を最小化するためには、次式のようなオイラー・ラグランジェ方程式を用い、反復により解を求める。

$$u_x^{k+1} = \bar{u}_x^k - \frac{I_x [I_x \bar{u}_x^k + I_y \bar{u}_y^k + I_t]}{\alpha^2 + I_x^2 + I_y^2},$$

$$u_y^{k+1} = \bar{u}_y^k - \frac{I_y [I_x \bar{u}_x^k + I_y \bar{u}_y^k + I_t]}{\alpha^2 + I_x^2 + I_y^2}.$$

ただし、 $k$  は反復回を表す。 $u_x^0, u_y^0$  は、初期のフローの推定値を表し、通常ゼロとされる。 $\bar{u}_x^k, \bar{u}_y^k$  は、それぞれ  $u_x, u_y$  について近傍の平均値である。

## (2) 勾配法の改良

Horn と Schunck によるオリジナルの勾配法<sup>2)</sup>は、その輝度不変性の仮定により照明変動に対して敏感であるという欠点がある。近年、勾配法の改良として、様々な不変性の制約が提案され、これら制約を輝度の不変性と併用することの有効性が示されている<sup>3)</sup>。その制約の例としては、輝度勾配の不変性、ヘッセ行列の不変性、ラプラシアンの不変性、勾配のノルムの不変性、ヘッセ行列のノルムの不変性、ヘッセ行列の行列式の不変性などがあげられる。これら様々な不変性に関する評価関数の重み付き和として、評価関数のデータ項が

$$E_{\text{data}}(\mathbf{u}) = \int_{\Omega} \sum_{i=1}^N \gamma_i D_i(I, \mathbf{u}, \mathbf{x}) dx \quad (4.6)$$

のように定義される。ここで  $D_i$  は、各々の仮定に対応した評価関数であり、 $\gamma_i$  はその重みである。 $N$  は不変性の制約の個数を指す。なお、式 (4.6) において、 $N = 1$  として、 $D_1$  を式 (4.4) の  $E_{\text{data}}$  とした場合、オリジナルの勾配法に相当する。

また、外れ値に対する頑健性を向上させるために、上式の評価関数についてある種のロバスト関数を導入して、評価関数を次のように定義することもできる。

$$E_{\text{data}}(\mathbf{u}) = \int_{\Omega} \Psi \left( \sum_{i=1}^N \gamma_i D_i(I, \mathbf{u}, \mathbf{x}) \right) d\mathbf{x}.$$

ただし、 $\Psi(\cdot)$  はロバスト性向上のための非二次関数である。

また、平滑化項に対しても同様のロバスト化を施した評価関数が提案されている。

$$E_{\text{smooth}}(\mathbf{u}) = \int_{\Omega} \Psi(|\nabla u_x|^2 + |\nabla u_y|^2) d\mathbf{x}. \quad (4\cdot7)$$

式(4・5)の古典的な滑らかさの制約条件では、物体と背景の境界部分などに生ずる速度場の不連続部分において適切なフローが得られないが、式(4・7)では、速度場の不連続部分を外れ値として許容するような平滑化項となっている。

以下、輝度の不変性の制約  $D_1$  以外に、これまで提案されている制約をいくつか紹介する。

#### (a) 勾配の不変性

これは輝度の空間的な勾配が動きに対して不変であることの仮定である。輝度の勾配は、輝度値のシフトに不変であり、輝度のスケールリングは、勾配ベクトルの長さのみに影響を与える。勾配の不変性に対応した評価関数は、

$$D_2(I, \mathbf{u}, \mathbf{x}) := \|\nabla I(\mathbf{x} + \mathbf{u}) - \nabla I(\mathbf{x})\|^2 = 0$$

のように表すことができる。

#### (b) ヘッセ行列の不変性

ヘッセ行列の不変性では、輝度の二階偏導関数を要素とするヘッセ行列の不変性を仮定する。この仮定は、並進及び拡大・縮小の動きに適しているが、運動の方向が変化する回転運動には適さない。ヘッセ行列の不変性に対応した評価関数は、

$$D_3(I, \mathbf{u}, \mathbf{x}) := \|\mathcal{H}_2 I(\mathbf{x} + \mathbf{u}) - \mathcal{H}_2 I(\mathbf{x})\|^2 = 0$$

のように表すことができる。ただし、 $\mathcal{H}_2$  は、ヘッセ行列を表す。

#### (c) ラプラシアンの不変性

動き方向の変化に対して不変な特徴として、ラプラシアンが利用可能である。これはヘッセ行列のトレースに相当する。ラプラシアンの不変性に対応する評価関数は、

$$D_4(I, \mathbf{u}, \mathbf{x}) := \|\Delta I(\mathbf{x} + \mathbf{u}) - \Delta I(\mathbf{x})\|^2 = 0,$$

$$\text{where } \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

のように表すことができる。ただし、 $\Delta$  はラプラシアンである。

#### (d) 勾配のノルムの不変性

輝度の空間勾配は、方向とノルムの成分に分解できる。回転により方向成分は変化するが、ノルム成分は変化しない。したがって、回転運動に対して、輝度勾配のノルムについて不変性を仮定した制約条件が有用である。この制約に対応する評価関数は、

$$D_5(I, \mathbf{u}, \mathbf{x}) := (\|\nabla I(\mathbf{x} + \mathbf{u})\| - \|\nabla I(\mathbf{x})\|)^2 = 0$$

のように表すことができる。

局所解への落ち込みを緩和するアプローチとして、疎密戦略 (coarse-to-fine strategy) がよく用いられている。疎密戦略では、まず、逐次的なダウンサンプリングにより画像ピラミッドを得る。その後、ピラミッドの最上部、つまり、最も粗い階層からオプティカルフローの計算を始め、その結果を、次に一つ下の階層 (より細かい) におけるオプティカルフロー計算の初期値として用いる。この処理を最下層の画像、つまり、入力画像の解像度の階層まで繰り返す。また、大きな変位の推定を可能にするアプローチとして、疎密変形法 (coarse-to-fine warping techniques) が知られている。これは、画像ピラミッドの各階層において、一つ上の粗い階層で計算されたフローを用いて、画像を変形させ、次の階層でのオプティカルフロー計算の入力画像とするアプローチである。

### (3) Lucas-Kanade 法

Lucas-Kanade 法<sup>4)</sup>は、局所近傍領域内においてオプティカルフローが一様であることを仮定して、その領域内の各画素について輝度不変性のオプティカルフロー拘束条件を課して、フローの最小二乗解を求める方法である。いま、フロー推定の対象画素を中心とした  $m \times m$  画素の近傍領域を考える。この領域中の各画素について、式 (4・2) のオプティカルフロー拘束式が成り立つと仮定すると次の連立方程式を得ることができる。

$$\begin{cases} I_{x_1}u_x + I_{y_1}u_y = -I_{t_1} \\ I_{x_2}u_x + I_{y_2}u_y = -I_{t_2} \\ \vdots \\ I_{x_n}u_x + I_{y_n}u_y = -I_{t_n} \end{cases} \quad (4\cdot8)$$

ただし、 $n$  は、近傍領域の画素数  $n = m \times m$  であり、各画素に  $1, \dots, n$  の番号を付与した。式 (4・8) は、次のように書き換えられる。

$$\mathbf{A}\mathbf{u} = -\mathbf{b}, \quad (4\cdot9)$$

$$\text{where } \mathbf{A} = \begin{bmatrix} I_{x_1} & I_{y_1} \\ I_{x_2} & I_{y_2} \\ \vdots & \vdots \\ I_{x_n} & I_{y_n} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \mathbf{b} = \begin{bmatrix} I_{t_1} \\ I_{t_2} \\ \vdots \\ I_{t_n} \end{bmatrix}.$$

フロー  $\mathbf{u}$  は、評価関数  $E = \|\mathbf{A}\mathbf{u} + \mathbf{b}\|^2$  の最小二乗解として、

$$\begin{aligned} \mathbf{A}^T \mathbf{A} \mathbf{u} &= \mathbf{A}^T (-\mathbf{b}), \\ \mathbf{u} &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T (-\mathbf{b}) \end{aligned}$$

のように計算される。

#### ■参考文献

- 1) J.L. Barron, D.J. Fleet, and S. Beauchemin, "Performance of optical flow techniques," International Journal of Computer Vision, vol.12, no.1, pp.43-77, 1994.

- 2) B. Horn and B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol.17, pp.185-203, 1981.
- 3) A. Bruh, T. Brox, S. Didas, and J. Weickert, "Highly accurate optic flow computation with theoretically justified warping," *International Journal of Computer Vision*, vol.67, no.2, pp.141-158, 2006.
- 4) B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," In *Proc. Seventh International Conference on Artificial Intelligence*, pp.674-479, 1981.

## ■ 2 群 画像・音・言語 - 2 編 パターン認識とビジョン - 4 章 動画解析

### 4-2 トラッキング

(執筆著者: 藤吉弘互) [2010 年 9 月 受領]

トラッキングとは、与えられた動画画像から、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る。トラッキングは、通常、図 4・3(a) のように、被写体である対象の初期位置  $x_0$  が、次フレームにおいてどのように変化したかという移動量  $\Delta x$  を順方向に探索することである。この処理を毎フレーム繰り返すことで、現在のフレームにおける対象物体の位置を知ることができる。本節では、このような処理を逐次処理によるトラッキングと呼ぶことにする。逐次処理によるトラッキングは、各フレームごとの処理が高速であれば、リアルタイムに被写体の追跡処理が可能となる。そのため、このような逐次処理によるトラッキングは、オートフォーカスや自動カメラワークといった撮影時の処理に利用することができる。

一方、図 4・3(b) のように、動画を時空間としてとらえ、対象領域をボリュームとして抽出する追跡アプローチも考えられる<sup>2)3)</sup>。本節では、このような処理を一括処理によるトラッキングと呼ぶことにする。一括処理によるトラッキングでは、あるフレームで追跡対象位置  $x_0$  を指定し、追跡対象が動画画像中でどのように移動したかを、動画画像全体を用いて順方向と逆方向から最適化することで移動量  $\Delta x_0, \dots, \Delta x_n$  を求める。そのため、逐次処理よりもノイズやオクルージョンに対して頑健に追跡することができる。一括処理によるトラッキングは複数の時系列画像からなるフレームバッファに対して行うため、画像合成や動画編集などの撮影後の処理として利用される。本節では、撮影時に利用可能な逐次処理によるトラッキングとして、領域ベースのトラッキングについて述べる。

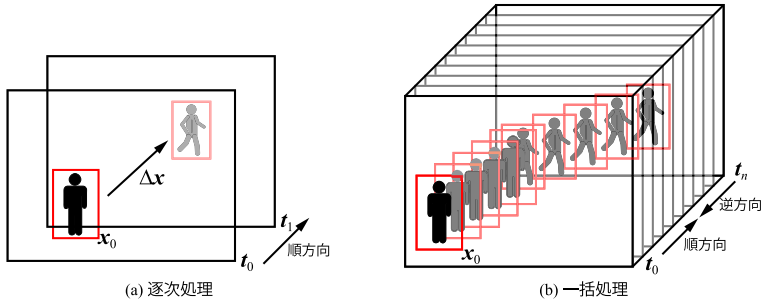


図 4・3 トラッキングの種類

#### 4-2-1 テンプレートマッチング (ブロックマッチング)

被写体の追跡をするには、追跡対象領域を指定する必要がある。ここでは、ユーザが指定した矩形領域から得られるカラーヒストグラムを特徴量として追跡する手法について考える。領域ベースの追跡には図 4・4 に示すように、全探索による手法と局所探索による手法がある。

全探索の追跡手法として代表的なものは、テンプレートマッチングである。テンプレートマッチングは探索領域内をラスタスキャンしてマッチング位置を求める。そのため、回転や



スケール変化などを考慮した場合、非常に計算量が多くなるという問題がある。この問題を解決するアプローチとして、アクティブ探索法<sup>4)</sup>が提案されている。

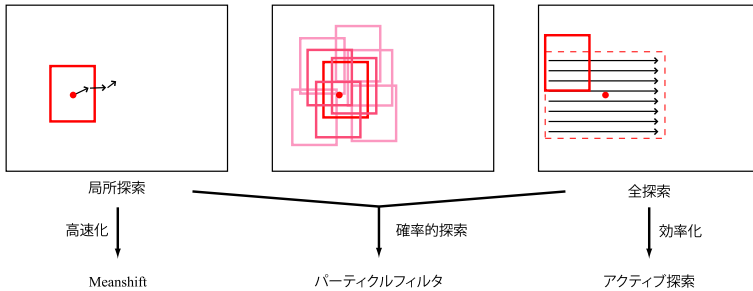


図 4・4 領域ベースの追跡手法

一方、探索領域内の局所探索を効率的かつ高速に行う手法として、Mean Shift<sup>1)</sup>が提案されている。また、全探索と局所探索の中間的な手法として、確率的に物体位置を推定するパーティクルフィルタ<sup>5)</sup>が提案されている。

これらの追跡手法では、特徴量にカラーヒストグラムを用いることが一般的である。カラーヒストグラムは歩行中の人のように形状が変化しても、その色分布は変化しないため、追跡に用いると非剛体や物体の見えの変化に対して頑健で高速な追跡が可能となる。以下に、探索領域全体を効率的に探索するアクティブ探索法、近傍領域を効率的に探索する Mean Shift と、確率的に次の物体位置を推定するパーティクルフィルタについて述べる。

#### 4-2-2 アクティブ探索による物体追跡

テンプレートマッチングは全探索してマッチング位置を求めため、計算量が多くなる。効率的に全探索する手法として、村瀬らによりアクティブ探索<sup>4)</sup>が提案されている。アクティブ探索法は、物体の形状変化に安定な色ヒストグラムを特徴として利用し、入力画像中のある位置の類似値からその近傍の類似値の上限値を計算する。上限値が探索値より小さければ、その領域での探索が省略できるため、照会回数を大幅に減らすことが可能となる。

入力画像中から矩形領域により抽出された領域の正規化ヒストグラムを  $p$ 、参照画像の正規化ヒストグラムを  $q$  とすると、ヒストグラムの類似値  $S_{p,q}$  は、

$$S_{p,q} = \sum_{u=1}^I \min(p_u, q_u) \tag{4・10}$$

によって与えられる。ここで、 $u$  はヒストグラムのビンで、 $p_u, q_u$  はビン  $u$  での頻度である。 $S_{p,q}$  は正規化されているため、0 から 1 の値をとる。

一般に、入力画像中のある位置での部分領域と参照画像との類似値は、その位置の近傍での類似値に類似している。図 4・5 の A の領域と B の領域とのヒストグラムの違いは、A と B の重なりあった部分の部分ヒストグラムは同じであるため、図の差分部分だけであるという関係を利用することが可能である。

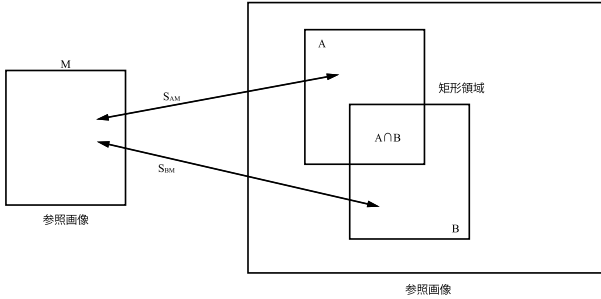


図 4・5 参照画像と二つの局所領域との関係

図 4・5 に示すように、 $A, B$  を画像中の任意の領域、 $M$  を参照画像とする。  $S_{AM}$  を  $A$  と  $M$  の類似値、  $S_{BM}$  を  $B$  と  $M$  の類似値とする。  $A \cap B$  を  $A$  と  $B$  の共通領域、  $A - B$  を  $A$  の領域から  $B$  の領域除いた領域、  $|A|$  を  $A$  の画素数とする。  $|A| \leq |B|$  の場合には、

$$S_{AM} = \frac{\min(S_{BM}|B|, |B \cap A|) + |A - B|}{|A|} \quad (4 \cdot 11)$$

という関係が成り立つ。 右辺は  $S_{BM}$  が計算されたときの  $S_{AM}$  の上限値である。 この関係を利用すれば、ある位置での類似値が計算されればその近傍の類似値の上限値が分かることになるため、もし探索している領域の類似値よりも上限値が小さければ、この領域での類似値の計算をする必要はない。

図 4・6 にアクティブ探索により矩形領域で指定したテンプレートと入力画像の照合を行った箇所を示す。 対象領域では密な探索を行い、それ以外の領域では疎に探索していることが分かる。

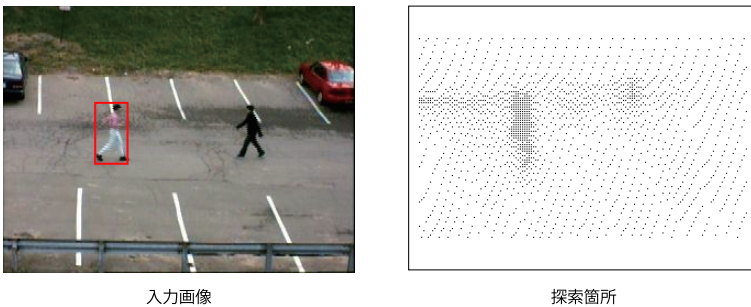


図 4・6 入力画像と照合した箇所

### 4-2-3 Mean Shift

Mean Shift によるトラッキング<sup>1)</sup>は、Comaniciu らにより提案された手法である。 Mean

Shift では、追跡対象のカラーヒストグラム間の類似度関数から求められる各ピクセルの重みの分布に対して探索を行うため、追跡対象の形状は部分的なオクルージョン問題に対して頑健である。この手法では、次フレームの状態  $\mathbf{y}$  を変数とする Bhattacharyya 係数の関数を以下のように定義する。

$$B(p(\mathbf{y}), q) = \sum_{u \in U} \sqrt{p_u(\mathbf{y})q_u} \tag{4.12}$$

式 (4.12) を一次微分の項までテイラー展開すると式 (4.13) が得られる。

$$B(p(\mathbf{y}), q) \approx \frac{1}{2} \sum_{u \in U} \sqrt{p_u(\mathbf{y}_0)q_u} + \frac{1}{2} \sum_{u \in U} p_u(\mathbf{y}) \sqrt{\frac{q_u}{p_u(\mathbf{y}_0)}} \tag{4.13}$$

式 (4.13) より、第 2 項を最大化させることで、Bhattacharyya 係数が最大となることが分かる。そこで、Mean Shift vector により Bhattacharyya 係数が大きくなる方向を求め、逐次更新することで Bhattacharyya 係数を最大化させる。Mean Shift vector による最大化は以下の式により求める。

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i w_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)} \tag{4.14}$$

$$w_i = \sum_{u=1}^m \delta[b(\mathbf{x}_i) - u] \sqrt{\frac{q_u}{p_u(\mathbf{y}_j)}} \tag{4.15}$$

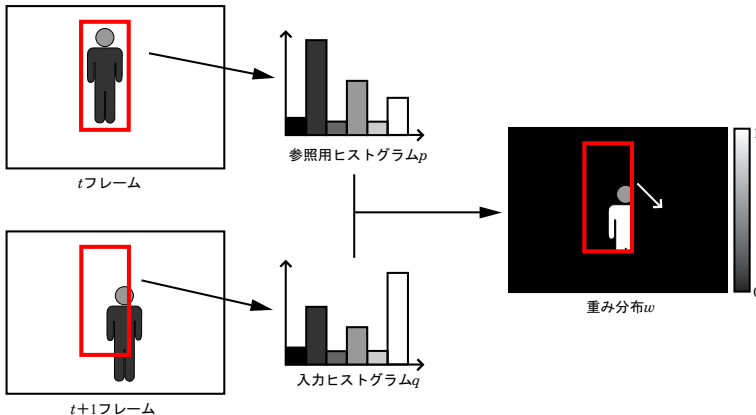


図 4.7 Mean Shift によるトラッキング

ここで、 $h$  は探索範囲を決めるカーネル関数  $g(x)$  のバンド幅、 $\mathbf{x}_i$  は各ピクセル、 $b(\mathbf{x}_i)$  は入力したピクセルのピンを返す関数である。式 (4.15) の重み  $w_i$  は、 $t$  フレームで領域内に多

く含まれる色が  $t+1$  フレームで減少したとき、重みは高くなる。そのため、図 4・7 に示すように、 $t$  フレームで頻度が高い色を  $t+1$  フレームでも高くなるように注目領域を移動させるベクトルの重みが高くなる。Mean Shift vector は、これらのベクトルの重み付き平均により求められ、これを逐次更新し、Bhattacharyya 係数が極大値となる位置を求めることで追跡することができる。

#### 4-2-4 パーティクルフィルタ

パーティクルフィルタ<sup>5)</sup>は、Condensation とも呼ばれ、追跡する対象を状態量と重み（尤度）をもつ多数のパーティクルにより状態空間内全体の確率分布を近似することで、ノイズや環境の変動に対して頑健な追跡を行うことができる。

時刻  $t$  における事後確率密度  $p(\mathbf{x}_t | \mathbf{z}_t)$  を、状態  $\mathbf{x}$  の仮説とその重みの  $N$  個の組からなるパーティクル群

$$s_t^{(i)} = \{\mathbf{x}_t^{(i)}, \pi_t^{(i)}\} (i = 1, \dots, N) \tag{4-16}$$

により近似することで、対象物体を追跡する。ここで時刻  $t$  における  $i$  番目の追跡対象の状態量  $\mathbf{x}_t^{(i)}$ 、 $\pi_t^{(i)}$  はその尤度である重みを示している。時刻  $t$  において画像から観測値  $\mathbf{z}_t$  が得られると、追跡対象の状態  $\mathbf{x}_t$  を確率変数とする確率密度は、事後確率密度  $p(\mathbf{x}_t | \mathbf{z}_t)$  として表される。この  $p(\mathbf{x}_t | \mathbf{z}_t)$  はベイズの法則を用いて以下のように表すことができる。

$$p(\mathbf{x}_t | \mathbf{z}_t) = \alpha p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{t-1}) \tag{4-17}$$

$\alpha$  は正規化のための定数である。

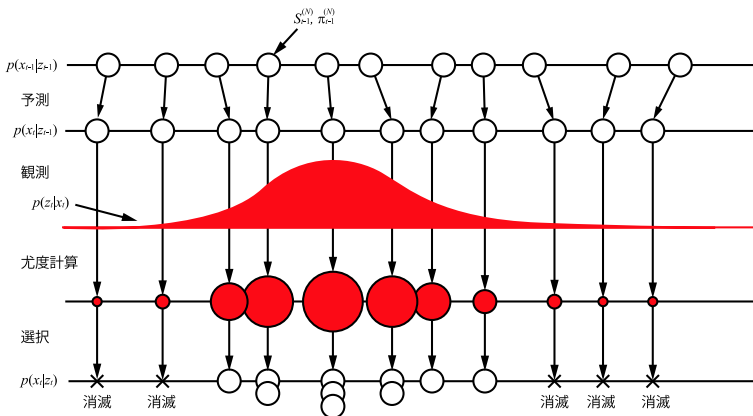


図 4・8 パーティクルフィルタのアルゴリズム

パーティクルフィルタによる移動体の追跡アルゴリズムを以下に示す（図 4・8 参照）。

**Step1. 初期化** 対象物体を指定し、 $N$  個のパーティクルを生成

**Step2. 予測** パーティクルを状態遷移モデルに基づき移動

**Step3. 観測** 各パーティクルについて矩形を作成，矩形内の色ヒストグラムを取得

**Step4. 尤度計算** 各パーティクルより取得した色ヒストグラムと前状態の対象物体より取得した参照ヒストグラムを比較し，パーティクルの尤度  $L_t^{(i)}$  を算出

**Step5. 対象推定** 各パーティクルの尤度を重みとした重み付き平均を求め，対象物体を推定

**Step6. 選択** 尤度が高い順番で尤度の高さに比例する確率で  $N$  個のパーティクルを選択

**Step7. 次状態推定**  $t + 1$  として Step2 の予測を実行

この処理を各フレームで行うことにより，対象物体を追跡することが可能となる。

#### ■参考文献

- 1) D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," In *Proceedings of the IEEE Computer Science Conference on Computer Vision and Pattern Recognition(CVPR-00)*, vol.2, pp.142-149, 2000.
- 2) J. Sun, W. Zhang, X. Tang, and H. Shum, "Bi-directional tracking using trajectory segment analysis," In *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV' 05)*, vol.1, pp.717-724, 2005.
- 3) Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," *Int. J. Comput. Vision*, vol.70, no.2, pp.109-131, 2006.
- 4) 村瀬 洋, V. Vinod, "局所色情報を用いた高速物体探索 : アクティブ探索法," *信学論 (D-II)*, vol.81, no.9, pp.2035-2042, 1998.
- 5) P. Bransnett, L. Mihaylova, D. Bull, and N. Canagarajah, "Sequential Monte Carlo tracking by fusing multiple cues in video sequences," *Image Vision Comput.*, vol.25, no.8, pp.1217-1227, 2007.

■ 2 群 画像・音・言語-2 編 パターン認識とビジョン-4 章 動画解析

4-3 3次元運動復元

(執筆者: 加藤 文和) [2010 年 9 月 受領]

3次元運動復元とは、図 4・9 に示すように、与えられた動画像から、対象物体やカメラ自身の 3次元位置・姿勢の相対的な変化を推定する問題である。3次元の位置・姿勢は、基準とする座標系(ワールド座標や、1フレーム目のカメラ座標、対象座標など)に対する3次元の並進ベクトル  $\mathbf{t} = [t_x \ t_y \ t_z]^T$  と 3次元の回転行列  $\mathbf{R}$  の 6 自由度で表現できる。

いま、3次元空間中で静止している  $N$  個の点を  $\mathbf{M}_i = [X_i \ Y_i \ Z_i]^T (i = 1, 2, \dots, N)$  とし、この点のある  $f$  フレームに観測したときの画像上の座標を  $\mathbf{m}_{i,f} = [u_{i,f} \ v_{i,f}]^T$  とする。このときのカメラの位置、姿勢を  $\mathbf{t}_f = [t_{x,f} \ t_{y,f} \ t_{z,f}]^T, \mathbf{R}_f$  とし、カメラの内部パラメータ行列を  $\mathbf{A}$  とすると、 $\mathbf{m}_{i,f}$  と  $\mathbf{M}_i$  の関係は次式のように表される。

$$\mathbf{s}\tilde{\mathbf{m}}_{i,f} = \mathbf{A}[\mathbf{R}_f \ \mathbf{t}_f]\tilde{\mathbf{M}}_i \tag{4・18}$$

ただし、 $\mathbf{m}_{i,f}, \tilde{\mathbf{M}}_i$  は、それぞれ  $\mathbf{m}_{i,f}, \mathbf{M}_i$  の同時座標表現である。このとき、3次元運動復元問題は、3次元空間中の静止点を観測した座標系列  $\mathbf{m}_{i,f}$  から、カメラの並進ベクトル  $\mathbf{t}_f$  と回転ベクトル  $\mathbf{R}_f$  を推定する問題となる。ただし、カメラ内部パラメータ行列  $\mathbf{A}$  はあらかじめキャリブレーションにより求めているものとする。

なお、ここでは、3次元空間中の点の座標は静止し、カメラが運動する場合について述べるが、静止するカメラで剛体物体を観測した場合も同様の問題であるといえる(図 4・10)。

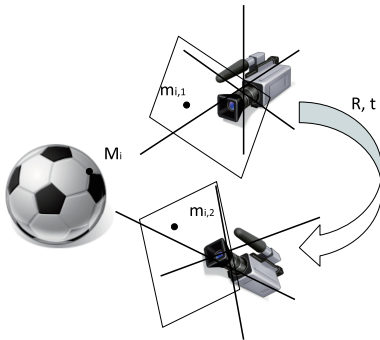


図 4・9 3次元運動復元(カメラ運動)

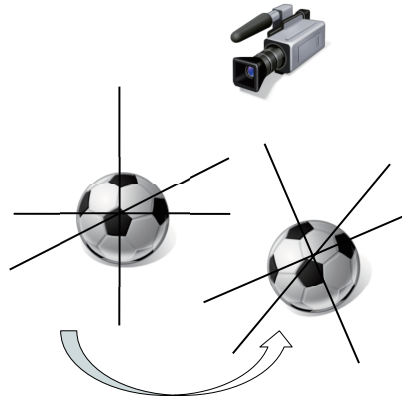


図 4・10 3次元運動復元(剛体物体の運動)

4-3-1 射影行列からの運動復元

対象の 3次元形状、すなわち 3次元座標  $\mathbf{M}_i = [X_i \ Y_i \ Z_i] (i = 1, 2, \dots, N)$  が既知であれば、3次元座標  $\mathbf{M}_i$  と対応する画像座標  $\mathbf{m}_{i,f}$  の対応づけから各フレーム射影行列  $P_f$  を推定することができる。

$$\mathbf{s}\tilde{\mathbf{m}}_{i,f} = \mathbf{P}_f\tilde{\mathbf{M}}_i \quad (4\cdot 19)$$

各  $\mathbf{P}_f = \mathbf{A}[\mathbf{R}_f \mathbf{t}_f]$  の形に分解することで、並進ベクトルと回転ベクトルを計算できる。式 (4・19) の行列を展開すると次のように書き表すことができる。

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & 1 \end{bmatrix} = k \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_X \\ R_{21} & R_{22} & R_{23} & t_Y \\ R_{31} & R_{32} & R_{33} & t_Z \end{bmatrix} \quad (4\cdot 20)$$

ここで、 $\mathbf{A}$  が上三角行列であり、 $\mathbf{R}$  が正規直行列であることから、 $\mathbf{P}$  の  $3 \times 3$  部分を QR 分解により上三角行列と直行列に分解することにより  $\mathbf{A}$  と  $\mathbf{R}$  を計算できる。

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (4\cdot 21)$$

また、並進ベクトル  $\mathbf{t}$  も求めた  $\mathbf{A}$  より次のように求められる。

$$\begin{bmatrix} t_X \\ t_Y \\ t_Z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}^{-1} \begin{bmatrix} p_{14} \\ p_{24} \\ 1 \end{bmatrix} \quad (4\cdot 22)$$

### 4-3-2 2 画像からの運動復元

対象の 3 次元形状が未知であり、2 視点で撮影した画像間の対応点のみが与えられている場合、これらの対応点間には次式で表すエピポーラ方程式が成り立つことが知られている。

$$\tilde{\mathbf{m}}_1^T \mathbf{F} \tilde{\mathbf{m}}_2 = 0 \quad (4\cdot 23)$$

この方程式の  $\mathbf{F}$  は基礎行列と呼ばれ、8 点以上の対応点から最小自乗法によって推定することができる<sup>1)</sup>。2 視点のカメラの内部パラメータ  $\mathbf{A}_1, \mathbf{A}_2$  が既知である場合、次式のようにエッセンシャル行列  $\mathbf{E}$  に変換できる。

$$\mathbf{E} = \mathbf{A}_1^T \mathbf{F} \mathbf{A}_2 \quad (4\cdot 24)$$

ここで、 $\mathbf{m}_1$  から内部パラメータの影響を取り除いた正規化座標系を  $\mathbf{x}_1$  としたとき、 $\tilde{\mathbf{m}}_1 = \mathbf{A}_1 \tilde{\mathbf{x}}_1 = \mathbf{A}_1 [\mathbf{R}_1 \mathbf{t}_1] \tilde{\mathbf{M}}_1$  となり、式 (4・23) は次式のように書き直せる。

$$\tilde{\mathbf{x}}_1^T \mathbf{E} \tilde{\mathbf{x}}_2 = 0 \quad (4\cdot 25)$$

このとき、 $\mathbf{E} = [\mathbf{t}']_{\times} \mathbf{R}'$  となる。ただし、 $\mathbf{R}' = \mathbf{R}_1^{-1} \mathbf{R}_2$  であり 2 視点間の相対的な回転行列を表し、 $\mathbf{t}' = \mathbf{R}_1^{-1}(\mathbf{t}_2 - \mathbf{t}_1)$  は 2 視点間の平行移動を表す。また、 $[\mathbf{t}']_{\times}$  はベクトル  $\mathbf{t}'$  による外積を表す歪対象行列である。

つまり、 $\mathbf{E}$  を  $\mathbf{t}'$  と  $\mathbf{R}'$  に分解することができれば、2 視点間の相対的なカメラ運動を復元

することができる<sup>2)</sup>.

ここで、 $\mathbf{t}'$  は、2 視点間の並進ベクトル、つまり言い換えれば 1 個目の視点のカメラから見た、2 視点目のカメラの投影中心であるので、すべての  $\mathbf{x}$  について次式が成り立つ。

$$\mathbf{t}'^T \mathbf{E} \mathbf{x} = 0 \quad (4\cdot26)$$

すなわち、 $\mathbf{E}^T \mathbf{t}' = \mathbf{0}$  を満たす  $\mathbf{t}'$  を求めればよい。ここで、エビポーラ方程式のスケールパラメータは任意なので、 $s\mathbf{t} = \mathbf{t}'$ 、 $|\mathbf{t}| = 1$  と置き、 $\|\mathbf{E}^T \mathbf{t}\|^2$  の最小自乗解を求めると次式の最小化問題となる。

$$\mathbf{t}'^T \mathbf{E} \mathbf{E}^T \mathbf{t} + \lambda(1 - \mathbf{t}^T \mathbf{t}) \quad (4\cdot27)$$

ただし、 $\lambda$  はラグランジュ乗数である。これを  $\mathbf{t}$  で微分して 0 とすると、 $2\mathbf{E} \mathbf{E}^T \mathbf{t} - 2\lambda \mathbf{t} = \mathbf{0}$  となり、 $\mathbf{E} \mathbf{E}^T$  の固有方程式となるので、 $\mathbf{t}$  は  $\mathbf{E} \mathbf{E}^T$  の最小固有値に対応する固有ベクトルとして求められる。また、 $\mathbf{E} = [\mathbf{t}']_{\times} \mathbf{R}'$  より、 $\mathbf{E} \mathbf{E}^T = [\mathbf{t}']_{\times} \mathbf{R}' \mathbf{R}'^T [\mathbf{t}']_{\times}^T = [\mathbf{t}']_{\times} [\mathbf{t}']_{\times}^T = \mathbf{t}'^T \mathbf{t}' \mathbf{I} - \mathbf{t}' \mathbf{t}'^T$  となり、この式に  $s\mathbf{t} = \mathbf{t}'$  を代入して  $\mathbf{t}'$  の長さ  $s$  を求めることができる。

また、回転行列  $\mathbf{R}'$  については、 $\mathbf{E} = [\mathbf{t}']_{\times} \mathbf{R}'$  より上記で求めた  $\mathbf{t}'$  を用いて次式のように与えられる

$$\mathbf{R}' = \underset{\mathbf{R}}{\operatorname{argmin}} \|\mathbf{E} - [\mathbf{t}']_{\times} \mathbf{R}\|^2 \quad (4\cdot28)$$

### 4-3-3 動画像からの運動復元

$N$  個の点について、 $F$  フレームの画像で対応付けが与えられているとき、点の 3 次元位置とカメラの 3 次元運動を線形に分解し、同時に復元する因子分解法と呼ばれる手法が提案されている<sup>3)</sup>。ここでは、カメラの投影モデルとして弱中心射影を仮定する。3 次元空間中の点を  $\mathbf{M}_i = [X_i \ Y_i \ Z_i]^T$ 、それを  $f$  フレーム目の視点位置で観測した画像座標上の点を  $\mathbf{m}_{i,f} = [u_{i,f} \ v_{i,f}]$  とすると弱中心射影モデルでは次式が成り立つ。

$$\begin{bmatrix} u_{i,f} \\ v_{i,f} \end{bmatrix} = s_f \begin{bmatrix} \mathbf{r}_{i,f}^{1T} \\ \mathbf{r}_{i,f}^{2T} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \quad (4\cdot29)$$

ただし、 $s_f$  は  $f$  フレーム目の画像のスケール、 $\mathbf{r}_f^{1T}, \mathbf{r}_f^{2T}$  は  $f$  フレーム目のカメラの回転行列  $\mathbf{R}_f$  の 1 番目と 2 番目の行ベクトルである。ここで、1 フレーム目のカメラのスケールと回転行列を基準とするとし、 $s_1 = 1, \mathbf{r}_f^1 = [1 \ 0 \ 0]^T, \mathbf{r}_f^2 = [0 \ 1 \ 0]^T$  とする。このとき、 $N$  個の点について、 $F$  フレームすべてで観測点が得られたとき、これらすべてを行列で表現すると次式のように書ける。



$$\begin{pmatrix} u_{1,1} & \cdots & u_{i,1} & \cdots & u_{N,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{1,f} & \cdots & u_{i,f} & \cdots & u_{N,f} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{1,F} & \cdots & u_{i,F} & \cdots & u_{N,F} \\ v_{1,F} & \cdots & v_{i,F} & \cdots & v_{N,F} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{1,f} & \cdots & v_{i,f} & \cdots & v_{N,f} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{1,F} & \cdots & v_{i,F} & \cdots & v_{N,F} \end{pmatrix} = \begin{pmatrix} s_1 \mathbf{r}_1^{1T} \\ \vdots \\ s_f \mathbf{r}_f^{1T} \\ \vdots \\ s_f \mathbf{r}_F^{1T} \\ s_1 \mathbf{r}_1^{2T} \\ \vdots \\ s_f \mathbf{r}_f^{2T} \\ \vdots \\ s_f \mathbf{r}_F^{2T} \end{pmatrix} \begin{bmatrix} X_1 & \cdots & X_i & \cdots & X_F \\ Y_1 & \cdots & Y_i & \cdots & Y_F \\ Z_1 & \cdots & Z_i & \cdots & Z_F \end{bmatrix} \quad (4\cdot30)$$

ここで、この三つの行列をそれぞれ  $\mathbf{D}, \mathbf{M}, \mathbf{S}$  と置くと、この式は  $\mathbf{D} = \mathbf{MS}$  と書ける。つまり、 $\mathbf{M}$  はフレームごとのカメラの位置・姿勢を表す行列であり、 $\mathbf{S}$  は 3 次元座標を表す行列であるので、観測行列  $\mathbf{D}$  が与えられたときに、これを  $\mathbf{M}$  と  $\mathbf{S}$  に分解することができれば、対象の 3 次元形状を復元するとともにカメラの 3 次元動作を復元できることになる。

$\mathbf{M}$  は  $3 \times 2F$  であり、 $\mathbf{S}$  は  $N \times 3$  であるので、これらの行列積である観測行列  $\mathbf{D}$  のランクは 3 である。ここで、 $\mathbf{D}$  を特異値分解すると次式のようになる。

$$\mathbf{D} = \mathbf{V}\Sigma\mathbf{U}^T \quad (4\cdot31)$$

ここで  $\Sigma$  は、特異値を対角行列にもつ対角行列であり、 $\mathbf{V}$  と  $\mathbf{U}$  はそれぞれ左直行列と右直行列である。 $\mathbf{D}$  のランクが 3 なので特異値は三つあり残りは 0 となるはずであるが、観測行列  $\mathbf{D}$  に含まれる誤差のために一般的には 4 個目以降の特異値も値をもつ。そこで、4 個目以降の特異値を削除し、それに対応する特異ベクトルも削除して、 $\mathbf{D}' = \mathbf{V}'\Sigma'\mathbf{U}'^T$  とする。

このとき、 $\mathbf{M}' = \mathbf{V}'(\Sigma')^{\frac{1}{2}}$ 、 $\mathbf{S}' = (\Sigma')^{\frac{1}{2}}\mathbf{U}'^T$  と置くと、 $\mathbf{D}' = \mathbf{M}'\mathbf{S}'$  のように  $3 \times 2F$  行列  $\mathbf{M}'$  と  $N \times 3$  行列  $\mathbf{S}'$  に分解できる。ただし、この分解は任意の  $3 \times 3$  の正則行列  $\mathbf{Q}$  に対して以下のような不定性をもっている。

$$\mathbf{D}' = (\mathbf{M}'\mathbf{Q})(\mathbf{Q}^{-1}\mathbf{S}') \quad (4\cdot32)$$

ここで、 $\mathbf{M}$  について回転行列が正規直交であることから  $\mathbf{r}_f^{1T}\mathbf{r}_f^1 = 1, \mathbf{r}_f^{2T}\mathbf{r}_f^2 = 1, \mathbf{r}_f^{1T}\mathbf{r}_f^2 = 0$  という拘束条件を与えることができ、この拘束条件を用いて  $\mathbf{M} = \mathbf{M}'\mathbf{Q}$  となる  $\mathbf{Q}$  を求めることができる。

#### 4-3-4 バンドル調整による中心射影画像からの運動復元

前々節で紹介した 2 画像からの運動復元では、2 視点ごとの運動を復元するため、これを多視点の画像系列へ適用すると各画像間の誤差が蓄積して最終的には大きな誤差となってしまう。また、前節で紹介した因子分解法による運動復元では、多数の画像系列から運動を復元できるが、線形に解くためにカメラモデルとして弱中心投影を用いているため、運動視差

が大きくなるような状況では、誤差が大きくなるという問題がある。

そこで、中心投影モデルのもとで、多数の画像系列から対象の 3 次元形状と各フレームの 3 次元位置・姿勢を、非線形の最小自乗法によって直接推定するバンドル調整という手法が提案されている<sup>4)</sup>。これは、各画像で観測されている座標と、推定した 3 次元座標とカメラ運動パラメータから推定される画像座標との間の誤差を最小にするように、3 次元座標とカメラ運動を推定する手法である。

$f$  フレーム目の投影行列を  $\mathbf{P}_f = \mathbf{A}_f[\mathbf{R}_f \mathbf{t}_f]$  とし、 $i$  番目の点の 3 次元座標を  $\mathbf{M}_i = [X_i Y_i Z_i]^T$  とすると、画像上の座標  $\mathbf{m}_{i,f} = [u_{i,f} v_{i,f}]^T$  は次のように与えられる。

$$s\tilde{\mathbf{m}}_{i,f} = P_f \tilde{\mathbf{M}}_i \quad (4 \cdot 33)$$

ここで、 $\mathbf{P}_f = [p_f^1 p_f^2 p_f^3]^T$  とおくとこの式は次のように展開できる。

$$\begin{bmatrix} su_{i,f} \\ sv_{i,f} \\ s \end{bmatrix} = \begin{bmatrix} p_f^{1T} \tilde{\mathbf{M}}_i \\ p_f^{2T} \tilde{\mathbf{M}}_i \\ p_f^{3T} \tilde{\mathbf{M}}_i \end{bmatrix} \quad (4 \cdot 34)$$

これは、同時座標を削除すると次式のように表される。

$$\begin{bmatrix} u_{i,f} \\ v_{i,f} \end{bmatrix} = \begin{bmatrix} \frac{p_f^{1T} \tilde{\mathbf{M}}_i}{p_f^{3T} \tilde{\mathbf{M}}_i} \\ \frac{p_f^{2T} \tilde{\mathbf{M}}_i}{p_f^{3T} \tilde{\mathbf{M}}_i} \end{bmatrix} \quad (4 \cdot 35)$$

よって、両辺の平均自乗誤差は次式によって与えられる。

$$\mathcal{E} = \frac{1}{NF} \sum_{f=1}^F \sum_{i=1}^N \left[ \left( u_{i,f} - \frac{p_f^{1T} \tilde{\mathbf{M}}_i}{p_f^{3T} \tilde{\mathbf{M}}_i} \right)^2 + \left( v_{i,f} - \frac{p_f^{2T} \tilde{\mathbf{M}}_i}{p_f^{3T} \tilde{\mathbf{M}}_i} \right)^2 \right] \quad (4 \cdot 36)$$

この誤差  $\mathcal{E}$  を最小とする 3 次元座標  $\mathbf{M}_i$  とカメラ位置・姿勢  $\mathbf{P}_f$  として求められる。

$$(\tilde{\mathbf{M}}_1, \dots, \tilde{\mathbf{M}}_N, \hat{\mathbf{p}}_1^1, \hat{\mathbf{p}}_1^2, \hat{\mathbf{p}}_1^3, \dots, \hat{\mathbf{p}}_F^1, \hat{\mathbf{p}}_F^2, \hat{\mathbf{p}}_F^3) = \underset{(\mathbf{M}_1, \dots, \mathbf{M}_N, \mathbf{p}_1^1, \mathbf{p}_1^2, \mathbf{p}_1^3, \dots, \mathbf{p}_F^1, \mathbf{p}_F^2, \mathbf{p}_F^3)}{\operatorname{argmin}} \mathcal{E} \quad (4 \cdot 37)$$

この最適化問題は非線形のものであり、適切な初期値からマルカート法などの繰り返し計算によって計算される。初期値としては前々節で述べた基礎行列の分解によって得た画像間の並進・回転を合成したものや、前節で述べた弱中心投影モデルのもとでの因子分解法の解を用いることができる。

#### ■参考文献

- 1) G. Xu and Z. Zhang, "Epipolar Geometry in Stereo, Motion and Object Recognition," A Unified Approach. Kluwer Academic Publishers, 1996.
- 2) O.D. Faugeras, "Three-Dimensional Computer Vision," A Geometric Viewpoint. MIT Press, Cambridge, MA, 1993.
- 3) C. Tomasi and T. Kanade, "Shape and motion from image streams under orthograph : a factorization method," International Journal of Computer Vision, vol.9., no.2, pp.137-154, 1992.
- 4) R. Szeliski and S.B. Kang, "Recovering 3d shape and motion from image streams using nonlinear least squares," Journal of Visual Communication and Image Representation, vol.5, no.1, pp.10-28, 1994.