

# 物体数が未知の運動軌跡からの運動物体分離

酒匂祐太郎<sup>1,a)</sup> 菅谷保之<sup>1,b)</sup>

**概要:** 本稿では, ビデオ画像中に含まれる剛体運動をする物体数が未知な場合の特徴点軌跡から複数の運動を分離する方法を提案する. 提案手法では, 特徴点の運動軌跡を二つのアフィン空間に分離することを繰り返すことで運動物体数が未知な特徴点軌跡の分離を実現する. 特徴点軌跡を分割するかどうかの指標として分離前と分離後のアフィン空間に対する尤度を用いる. また, 既存研究でも指摘されている同一剛体運動をする特徴点の過分割を防ぐために, 分割の際に使用した尤度による統合を行うことで, 過分割に対応する. 複数の運動物体が存在するビデオ画像を用いた実験によって本手法の有効性を確認した.

## Multibody motion segmentation for an arbitrary number of independent motions

YUTARO SAKO<sup>1,a)</sup> YASUYUKI SUGAYA<sup>1,b)</sup>

**Abstract:** We propose a new method for segmenting feature point trajectories tracked through a video sequence without assuming a number of independent motions. Our method realizes motion segmentation of feature point trajectories by repeatedly dividing the trajectories into two affine spaces in a situation that we do not know the number of independently moving objects. We judge that input trajectories should be divided by comparing likelihoods computed from those trajectories of before/after segmentation. We also consider integration of the resulting segmented trajectories for avoiding too much segmentations. By using real video images, we confirmed the efficiency of our proposed method.

### 1. はじめに

ビデオ画像上の複数の運動物体を同一の剛体運動をする物体に分離することはコンピュータビジョンの重要なタスクのひとつである. 運動物体分離が実現できれば, 防犯カメラ映像からの不審者検出やビデオカメラの画像上の複数の物体の3次元復元することが可能になるなど, さまざまな応用に利用することが期待される.

ビデオ画像から特徴点の追跡を行い, その特徴点の運動軌跡から運動分離を行う研究はこれまで多くなされてきた. Costeira[1]らはアフィンカメラを仮定するとビデオ画像上の特徴点の軌跡が高次元空間の1点と同一視でき, 同一の運動をする特徴点の軌跡は共通の部分空間に含まれることから運動物体の分離を行っている. 菅谷ら[8]は運動物体の運動の仕方によって特徴点軌跡が含まれる空間が変化することを考慮して, 複数の運動モデルを組み合わせる多段

階最適化による分離を提案した. Vidalら[12], [13]は多項式の当てはめによって高次元空間の点集合を複数の部分空間に分類する一般主成分分析(GPCA)を提案した. 田中ら[11]は菅谷ら[8]の多段階最適化とVidalら[12], [13]の一般主成分分析を組み合わせた手法を提案している.

これらの既存研究の多くは, 運動物体数が既知であるという条件で運動分離を実現している. 運動物体数を推定する手法には金谷ら[5]の研究があるが, この手法は軌跡全体が拘束されるアフィン空間の次元数を推定して, その次元数をもとに物体数を推定するものである. しかし, 運動の仕方によってはその軌跡が含まれる空間の次元が退化する問題や, 同一剛体運動をする特徴点の軌跡であっても異なる運動と判定されてしまう問題があることが指摘されている.

そこで本研究では, 特徴点軌跡を分離するかどうかを軌跡の分離前と分離後の尤度を用いて判定して, 運動軌跡を2分割することを繰り返すことで, 運動物体数が未知の場合でも特徴点軌跡の分離を行う手法を提案する. また, 分離結果に対して分離の際に用いた尤度による統合を行うことで, 同一物体上の特徴点の過分割の問題に対応する.

<sup>1</sup> 豊橋技術科学大学情報・知能工学系  
Department of Computer Science and Engineering, Toyohashi University of Technology

a) sako@iim.cs.tut.ac.jp

b) sugaya@iim.cs.tut.ac.jp

## 2. アフィン空間拘束条件

まず、本研究で基礎となる特徴点軌跡のアフィン空間拘束条件について説明する [6].

運動物体を移動するビデオカメラで撮影したビデオ画像をフレームごとに切り出し、切り出した第 1 フレーム上にある特徴点の位置を求め、その特徴点を残りのフレームで追跡する。ここで、 $N$  個の特徴点  $\{p_\alpha\}$  を  $M$  フレームの画像に渡って追跡したときの第  $\kappa$  画像における  $\alpha$  番目の特徴点  $p_\alpha$  の画像座標を  $(x_{\kappa\alpha}, y_{\kappa\alpha})$ ,  $\kappa = 1, \dots, M, \alpha = 1, \dots, N$  と表す。この運動軌跡を次の  $2M$  次元ベクトルで表し、軌跡ベクトルと呼ぶ。

$$p_\alpha = (x_{1\alpha}, y_{1\alpha}, x_{2\alpha}, y_{2\alpha}, \dots, x_{M\alpha}, y_{M\alpha})^\top \quad (1)$$

これにより、各特徴点の軌跡を  $2M$  次元空間の 1 点とみなすことができる。本稿ではこれまでの我々の研究と同様に、カメラの光軸を  $Z$  軸とする  $XYZ$  カメラ座標系を定義し、これに相対的に物体が運動すると考える。物体に物体座標系を任意に固定し、特徴点  $p_\alpha$  の物体座標系に関する座標を  $(a_\alpha, b_\alpha, c_\alpha)$  とする。第  $\kappa$  フレームでの物体座標系の原点を  $t_\kappa$  とし、各座標軸の基底ベクトルをカメラ座標系で表したものを  $\{i_\kappa, j_\kappa, k_\kappa\}$  とすると、特徴点  $p_\alpha$  の第  $\kappa$  フレームにおける 3 次元位置  $r_{\kappa\alpha}$  はカメラ座標系では式 (2) で表せる。

$$r_{\kappa\alpha} = t_\kappa + a_\alpha i_\kappa + b_\alpha j_\kappa + c_\alpha k_\kappa \quad (2)$$

平行投影や弱透視投影、擬似透視投影を一般化したアフィンカメラ [3] を仮定すると、3 次元点  $r_{\kappa\alpha}$  の画像面への投影は式 (3) のようになる。

$$\begin{pmatrix} x_{\kappa\alpha} \\ y_{\kappa\alpha} \end{pmatrix} = A_\kappa r_{\kappa\alpha} + b_\kappa \quad (3)$$

ここで、 $A_\kappa, b_\kappa$  はそれぞれ第  $\kappa$  フレームでのカメラの位置や内部パラメータによって定まる  $2 \times 3$  行列および 2 次元ベクトルである。式 (2) を式 (3) に代入すると式 (4) のようになる。

$$\begin{pmatrix} x_{\kappa\alpha} \\ y_{\kappa\alpha} \end{pmatrix} = \tilde{m}_{0\kappa} + a_\alpha \tilde{m}_{1\kappa} + b_\alpha \tilde{m}_{2\kappa} + c_\alpha \tilde{m}_{3\kappa} \quad (4)$$

ここで  $\tilde{m}_{i\kappa}$ ,  $i = 0, 1, 2, 3$  は  $A_\kappa, b_\kappa, t_\kappa, i_\kappa, j_\kappa, k_\kappa$  から計算される 2 次元ベクトルである。そして、式 (4) を  $\kappa = 1, \dots, M$  に渡って縦に並べると式 (5) のように書ける。

$$p_\alpha = m_0 + a_\alpha m_1 + b_\alpha m_2 + c_\alpha m_3 \quad (5)$$

ただし、 $m_i, i = 0, 1, 2, 3$  は  $\tilde{m}_{i\kappa}$  をフレーム  $\kappa = 1, \dots, M$  に渡って縦に並べた  $2M$  次元ベクトルである。

式 (5) は特徴点の軌跡が  $m_0, m_1, m_2, m_3$  の張る 4 次元部分空間に含まれることを意味する。さらに、式 (5) 内の

$m_0$  の係数は全ての  $\alpha$  に共通に 1 であることから、特徴点の軌跡は 4 次元部分空間内の 3 次元アフィン空間に含まれる。これをアフィン空間拘束条件と呼ぶ。

## 3. アフィン空間拘束条件による分離とその問題

アフィン空間拘束条件を用いて特徴点軌跡を互いに異なるアフィン空間に分類することで、ビデオ画像中で異なる剛体運動をする物体を分離することが可能となる。しかし、運動分離の対象となるビデオ画像では、背景や物体が画像内で 2 次元的な運動をすることが多く、特徴点軌跡が含まれるアフィン空間の次元が退化することがある。

物体の回転が光軸回りのみの場合は、物体座標系の基底ベクトル  $k_\kappa$  を  $Z$  方向にとれば、アフィンカメラのもとでは画像面に投影されないで、式 (5) の  $m_3$  を含む項がなくなり、軌跡ベクトルが  $m_0$  を通り、 $\{m_1, m_2\}$  の張る 2 次元アフィン空間に含まれる。さらに、物体も背景も回転しなければ、式 (2) の  $i_\kappa, j_\kappa$  をそれぞれ  $X$  方向、 $Y$  方向の基底  $i, j$  に固定してよい。これは物体と背景に共通であるから、式 (5) の  $m_1, m_2$  も物体と背景に共通になり、それぞれの 2 次元アフィン空間は平行な 2 次元アフィン空間になる。このとき、それぞれの特徴点軌跡が含まれるアフィン空間の次元が退化するだけではなく、これらの二つの平行な 2 次元アフィン空間を含む 3 次元アフィン空間が存在することになり、特徴点軌跡を異なる 3 次元アフィン空間に分離することではこのような軌跡を正しく分離することが困難である。

ビデオ画像中の運動物体数が既知であれば、その軌跡が含まれるアフィン空間の次元を推定してから、その軌跡に最適なアフィン空間への当てはめにより分離を行ったり、菅谷ら [8]、田中ら [11] のように当てはめるアフィン空間の次元を多段階に変化させることにより分離を実現できる。しかし、運動物体数が未知の場合は、個々の軌跡が含まれる最適なアフィン空間の次元と運動物体数を同時に推定するのが困難である。

## 4. 運動物体数が未知の場合の分離手法

アフィン空間拘束条件により、剛体運動をする物体の軌跡ベクトルは 2 次元または、3 次元アフィン空間に拘束される。前節で述べたように運動軌跡の含まれるアフィン空間の次元はその物体の運動の仕方によって異なり、また、運動物体数に応じて軌跡全体が含まれるアフィン空間の次元も変化する。

しかし、特徴点軌跡をその軌跡が含まれるアフィン空間の次元以下の二つのアフィン空間に分割することを繰り返せば、最終的には異なる 2 運動物体を分離することが可能だと考えられる。具体的には、まず分離前の特徴点軌跡が含まれるアフィン空間の次元を幾何学的 MDL [4], [5] を用い

て推定し、その次元が3次元以上であればその軌跡をEMアルゴリズムを用いて二つのアフィン空間に分離する。次にその分離が妥当かどうかを分離前のアフィン空間に対する尤度と分離後のアフィン空間に対する尤度を比較して判断する。これを分離後の特徴点軌跡が含まれるアフィン空間の次元が2次元になるか、尤度の比較によって軌跡が分離されなくなるまで繰り返す。提案手法のアルゴリズムを以下に示す。

- (1) 幾何学的MDLによって、特徴点軌跡が含まれるアフィン空間の次元  $d$  の推定をする。
- (2)  $d \leq 2$  なら終了する。
- (3) 特徴点軌跡に二つのアフィン空間を当てはめて分離する。
  - (a) 特徴点軌跡を3次元ベクトルに変換する。
  - (b) Taubin法による2平面当てはめによって軌跡を分離する。
  - (c) 特徴点軌跡を  $d$  次元ベクトルに変換する。
  - (d) 2平面当てはめによる分離結果を初期値として、 $d$  次元EMアルゴリズムによって、2つのアフィン空間に分離する。
- (4) 分離後の軌跡数がどちらか一方でも3点以下になった場合には分離前の状態で終了する。
- (5) 各特徴点軌跡ごとに分離前のアフィン空間に対する尤度と分離後のアフィン空間に対する尤度を計算し、それぞれの平均値  $P$  と  $P'$  を計算する。
- (6) 計算した尤度の平均値を対数尤度として比較して、以下の不等式を満たす場合は分離結果を確定する。そうでなければ、この特徴点軌跡の分離を行わずに終了する。

$$[\log_{10} P'] - [\log_{10} P] > 0 \quad (6)$$

ここで、 $[\cdot]$  は床関数を表す。

上記の処理をすべての特徴点軌跡が分離されなくなるまで繰り返す。以下の節からアルゴリズムの各処理の詳細について説明する。

## 5. 幾何学的MDLによるモデル選択を用いたアフィン空間の次元推定

特徴点軌跡が含まれるアフィン空間の次元を調べるには、軌跡のモーメント行列の固有値を計算し、正の固有値の数を調べればよい。しかし、データに誤差が含まれている場合には本来ゼロになるべき固有値も厳密にはゼロにならず、アフィン空間の次元の推定が困難である。本研究では、当てはめ残差に加えてモデルの自由度を考慮した指標である幾何学的MDL[5]を利用して特徴点軌跡が含まれるアフィン空間の次元を推定する。

$2M$ 次元ベクトルで表された特徴点軌跡に対して、モーメント行列

$$M = \sum_{\alpha=1}^N (\mathbf{p}_{\alpha} - \mathbf{p}_C)(\mathbf{p}_{\alpha} - \mathbf{p}_C)^{\top}, \mathbf{p}_C = \frac{1}{N} \sum_{\alpha} \mathbf{p}_{\alpha} \quad (7)$$

の固有値を計算し、 $\lambda_1 \geq \dots \geq \lambda_{2M}$  とする。このとき、特徴点軌跡を  $r$  次元アフィン空間に当てはめたときの残差  $J_r$  は次のように計算できる。

$$J_r = \sum_{i=r+1}^{2M} \lambda_i \quad (8)$$

また、 $2M$ 次元空間の  $r$ 次元アフィン空間の自由度は  $(r+1)(2M-r)$  であるから、 $N$ 個のデータ点に対する幾何学的MDLは次式で表せる。

$$\text{G-MDL}(r) = J_r + (rN + (r+1)(2M-r)) \epsilon^2 \log\left(\frac{\epsilon}{L}\right)^2 \quad (9)$$

ここで、 $\epsilon$  は特徴点座標に加わる誤差の標準偏差であり、ノイズレベルと呼ぶ、また  $L$  はデータ空間の大きさを表す基準長である。

式(9)で定義される幾何学的MDLの値を  $r=2, \dots, 2M$  に対して計算し、次式を満たす次元  $\hat{r}$  を特徴点軌跡が含まれるアフィン空間の次元とする。

$$\hat{r} = \arg \min_r \text{G-MDL}(r) \quad (10)$$

## 6. EMアルゴリズムを用いた軌跡の分離

### 6.1 軌跡の次元変換

$2M$ 次元空間の軌跡ベクトルを  $d$ 次元ベクトルに変換するには、軌跡ベクトルの主成分分析を行い、第  $d$ 主成分までの基底を用いて軌跡ベクトルを表現すればよい。その手順は以下のとおりである。

- (1) 軌跡ベクトル  $\mathbf{p}_{\alpha}, \alpha=1, \dots, N$  の重心  $\mathbf{p}_C$  からの変位  $\mathbf{p}'_{\alpha}$  を計算する。

$$\mathbf{p}'_{\alpha} = \mathbf{p}_{\alpha} - \mathbf{p}_C, \quad \mathbf{p}_C = \frac{1}{N} \sum_{\alpha=1}^N \mathbf{p}_{\alpha} \quad (11)$$

- (2) 次の  $2M \times N$  行列を特異値分解する。

$$W = (\mathbf{p}'_1, \dots, \mathbf{p}'_N) = U \text{diag}(\sigma_1, \dots, \sigma_r) V^{\top} \quad (12)$$

ここで、 $r = \min(2M, N)$  であり、 $U$  は  $r$ 本の正規直交系の列ベクトルとして持つ  $2M \times r$  行列、 $V$  は  $r$ 本の正規直交系の列ベクトルとして持つ  $N \times r$  行列である。また、 $\sigma_1 \geq \dots \geq \sigma_r (\geq 0)$  は特異値であり、 $\text{diag}(\sigma_1, \dots, \sigma_r)$  はそれらを対角成分として持つ対角行列である。

- (3) 行列  $U$  の第  $i$ 列を  $\mathbf{u}_i$  として、次の  $d$ 次元ベクトル  $\mathbf{r}_{\alpha}, \alpha=1, \dots, N$  を計算する。

$$\mathbf{r}_{\alpha} = \left( (\mathbf{p}'_{\alpha}, \mathbf{u}_1), \dots, (\mathbf{p}'_{\alpha}, \mathbf{u}_d) \right)^{\top} \quad (13)$$

## 6.2 2平面当てはめ

3次元点に対して、2平面を当てはめる方法を説明する。3次元空間の平面  $Ax + By + Cz + D = 0$  はベクトル  $n, x$  を

$$n = (A, B, C, D)^\top, \quad x = (x, y, z, 1)^\top \quad (14)$$

と置けば  $(n, x) = 0$  と書ける。ここで  $n$  は何倍しても同じ平面を表すので、定数倍の不定性がある。3次元空間に2平面  $(n_1, x) = 0, (n_2, x) = 0$  があるとき、次のように一つの式にまとめることができる。

$$(n_1, x)(n_2, x) = (x, n_1 n_2^\top x) = (x, Qx) = 0 \quad (15)$$

ただし、対称行列  $Q$  を次のようにおいた。

$$Q = \frac{n_1 n_2^\top + n_2 n_1^\top}{2} \quad (16)$$

行列  $Q$  はランク2であり、固有値0を2重解として持つ。そして、残りの二つの固有値は異符号である。固有値を大きい順に  $\lambda_1 > \lambda_2 = \lambda_3 (= 0) > -\lambda_4$  とし、対応する単位固有ベクトルを  $u_1, u_2, u_3, u_4$  とすると、 $Q$  は次のようにスベクトル分解される。

$$\begin{aligned} Q &= \lambda_1 u_1 u_1^\top - \lambda_4 u_4 u_4^\top \\ &= \left( \sqrt{\frac{\lambda_1}{2}} u_1 + \sqrt{\frac{\lambda_4}{2}} u_4 \right) \left( \sqrt{\frac{\lambda_1}{2}} u_1 - \sqrt{\frac{\lambda_4}{2}} u_4 \right)^\top \\ &\quad + \left( \sqrt{\frac{\lambda_1}{2}} u_1 - \sqrt{\frac{\lambda_4}{2}} u_4 \right) \left( \sqrt{\frac{\lambda_1}{2}} u_1 + \sqrt{\frac{\lambda_4}{2}} u_4 \right)^\top \end{aligned} \quad (17)$$

式(16)と見比べ、 $Q$ に定数倍の不定性があることを考慮すると、 $n_1, n_2$ は次のように定まる。

$$n_1 = \sqrt{\lambda_1} u_1 + \sqrt{\lambda_4} u_4 \quad n_2 = \sqrt{\lambda_1} u_1 - \sqrt{\lambda_4} u_4 \quad (18)$$

したがって、誤差により完全に2平面にはない  $N$  点を式(14)のベクトル  $x$  の形で  $x_1, \dots, x_N$  と表し、

$$(x_\alpha, Qx_\alpha) \approx 0, \quad \alpha = 1, \dots, N \quad (19)$$

となる  $Q$  を求めれば、式(18)によって2平面を指定する  $n_1, n_2$  が定まる。

3次元空間点  $(x, y, z)$  から平面  $Ax + By + Cz + D = 0$  までの距離  $d$  は次のように書ける。

$$d = \frac{|Ax + By + Cz + D|}{\sqrt{A^2 + B^2 + C^2}} \quad (20)$$

これを計算して、各データ点を近い方の平面に分類することにより、点集合を2クラスに分類する。本研究では Taubin 法を用いて、3次元点に変換した特徴点軌跡を2平面に分類し、EMアルゴリズムによる軌跡分割の初期値とする。

## 6.3 EMアルゴリズムを用いた2分割手法

$d$ 次元EMアルゴリズムを用いて軌跡ベクトルを2つのアフィン空間に分離する方法について述べる。これは菅谷ら[8]の多段階最適化の中で用いられた手法と同一である。 $n$ 次元ベクトルに変換された  $N$ 個の軌跡  $\{r_\alpha\}$  に二つの  $d$ 次元アフィン空間を当てはめて、2クラスに分類するEMアルゴリズムは次のようになる。

(1) 初期分類を用いて、各軌跡  $r_\alpha$  の各クラス  $k$  に対する重み  $W_\alpha^{(k)}$  を次のように計算する。

$$W_\alpha^{(k)} = \begin{cases} 1 & \text{軌跡 } r_\alpha \text{ がクラス } k \text{ に属する} \\ 0 & \text{それ以外} \end{cases} \quad (21)$$

(2) 各クラス  $k = 1, 2$  について以下の計算を行う。

(a) クラス  $k = 1, 2$  の重み  $w^{(k)}$  を次のように計算する。

$$w^{(k)} = \frac{1}{N} \sum_{\alpha=1}^N W_\alpha^{(k)} \quad (22)$$

(b)  $w^{(k)} \leq d/N$  なら計算を中断して終了し、その時点での分類結果を出力する。

(c) クラス  $k$  の重心  $r_C^{(k)}$  を次のように計算する。

$$r_C^{(k)} = \frac{\sum_{\alpha=1}^N W_\alpha^{(k)} r_\alpha}{\sum_{\alpha=1}^N W_\alpha^{(k)}} \quad (23)$$

(d) クラス  $k$  のモーメント行列

$$M^{(k)} = \frac{\sum_{\alpha=1}^N W_\alpha^{(k)} (r_\alpha - r_C^{(k)})(r_\alpha - r_C^{(k)})^\top}{\sum_{\alpha=1}^N W_\alpha^{(k)}} \quad (24)$$

の  $n$  個の固有値  $\lambda_1^{(k)} \geq \dots \geq \lambda_n^{(k)}$  に対応する単位固有ベクトル  $u_1^{(k)}, \dots, u_n^{(k)}$  を計算する。

(e) クラス  $k$  への射影行列  $P^{(k)}$  とその外側方向への射影行列  $P_\perp^{(k)}$  を次のように計算する。

$$P^{(k)} = \sum_{i=1}^d u_i^{(k)} u_i^{(k)\top}, \quad P_\perp^{(k)} = I - P^{(k)} \quad (25)$$

(f) 二乗ノイズレベル  $\sigma^2$  を次のように推定する。

$$\sigma^2 = \min \left[ \frac{N \cdot \text{tr}(w^{(1)} P_\perp^{(1)} M^{(1)} P_\perp^{(1)} + w^{(2)} P_\perp^{(2)} M^{(2)} P_\perp^{(2)})}{(n-d)(N-d-1)}, \sigma_{\min}^2 \right] \quad (26)$$

ここで、 $\sigma_{\min}^2$  は微小定数であり、誤差のない、あるいは誤差の極めて小さいデータに対して計算が破綻することを防止するためのものである\*1。また、 $\text{tr}$  は行列のトレースを表す。

(3) クラス  $k = 1, 2$  の共分散行列  $V^{(k)}$  を次のように計算

\*1 本研究では、 $\sigma_{\min}^2$  を 1.0 画素とした。

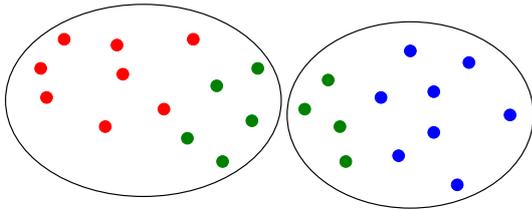


図 1: 過分割が起こるような分割の例. 赤, 緑, 青の点がそれぞれ異なる運動物体を表す. 2 分割を行うとこのように同一物体に属する点が 2 分割されてしまうことがある.

する.

$$\mathbf{V}^{(k)} = \mathbf{P}^{(k)} \mathbf{M}^{(k)} \mathbf{P}^{(k)} + \sigma^2 \mathbf{P}_{\perp}^{(k)} \quad (27)$$

(4) 各軌跡  $r_{\alpha}$  について次の計算を行う.

(a) 軌跡  $r_{\alpha}$  の各アフィン空間に対する尤度  $P(\alpha|k)$ ,  $k = 1, 2$  を次のように計算する.

$$P(\alpha|k) = \frac{e^{-(r_{\alpha} - r_C, \mathbf{V}^{(k)-1}(r_{\alpha} - r_C))/2}}{\sqrt{\det \mathbf{V}^{(k)}}} \quad (28)$$

(b) 軌跡  $r_{\alpha}$  の重み  $W_{\alpha}^{(k)}$ ,  $k = 1, 2$  を次のように更新する.

$$W_{\alpha}^{(k)} = \frac{w^{(k)} P(\alpha|k)}{w^{(1)} P(\alpha|1) + w^{(2)} P(\alpha|2)} \quad (29)$$

(5) ステップ 2 に戻って,  $\{W_{\alpha}^{(k)}\}$  が収束するまで繰り返す.

(6) 収束したら, 各軌跡  $r_{\alpha}$  を  $\{W_{\alpha}^{(k)}\}$  が大きいクラス  $k$  に分類する.

## 7. アフィン空間に対する尤度比較による分離判定

提案手法で特徴点軌跡を 2 分割する際, 過剰な分割を防ぐためにその分離が妥当であるかを判定する必要がある. アフィン空間への当てはまり具合を判定するには当てはめ残差を用いて幾何学的モデル選択を行う方法も考えられる. しかし, 我々の研究で EM アルゴリズムによる軌跡の分離の際に計算した軌跡ごとのアフィン空間に対する尤度を用いる方が正しい判定結果を得ることが多かったため, 尤度を用いた分離判定を行うことにした.

さらに式 (28) で計算される尤度は非常に小さな値を取ることが実験によってわかっており, 尤度自身を直接比較しても安定した判定を行うことが困難である. そこで本研究では, 分離前の尤度の平均値  $P$  と分離後の尤度の平均値  $P'$  に対して, 常用対数をとった値を比較して式 (6) を満たす場合に分離結果が妥当であると判定する, すなわち, 軌跡を分離する.

ただし, 分離後の軌跡の数が 3 以下である場合には正しい分離が行われていないと判断して軌跡は分離しないことにする. これは, 特徴点軌跡の数が 3 である場合, その軌跡に対して 2 次元アフィン空間が誤差なく当てはめってしまうため, そのアフィン空間が適切に物体の運動を反映したものであるかわからないためである.

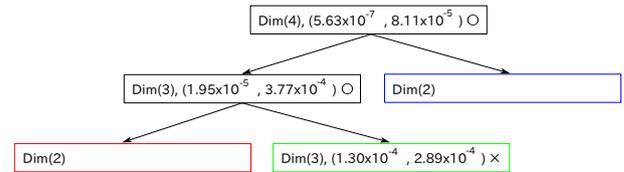


図 2: 図 3 のデータに対する分離の流れ. 図中の矩形内は, Dim(a) は対象となる軌跡が含まれるアフィン空間の次元の推定結果, (b, c) はそれぞれ軌跡の分離前後のアフィン空間に対する平均尤度, o 印は式 (6) を満たし軌跡を分離する, x 印は式 (6) を満たさず軌跡を分離しないことを表す. 図中の青, 赤, 緑枠は図 3(d) の最終分離結果の特徴点の色に対応する.

## 8. 分離した軌跡の統合判定

提案手法によって特徴点軌跡を分割していくと, 図 1 に示す例のような三つ以上の運動物体が存在する場合の分割では, 緑色で示した軌跡が分割によって異なるグループに分割されてしまうことがある. そのため, 最終の分離結果から本来同一物体上の軌跡であったものを統合する仕組みが必要となる.

本研究では, この統合の仕組みとして分離判定に用いた基準を同様に利用して分離された軌跡の統合判定を行うことにした. このとき, 直前の分離処理によって分離された軌跡同士は分離判定によって分離が妥当であると判定されているため, 統合判定は行わない. また, 二つの軌跡グループを統合するときの尤度は, まず, それらが含まれるアフィン空間の次元を 5 節に示した方法で推定し, その次元のアフィン空間に軌跡を当てはめたときの尤度を式 (28) により計算する. そして, 統合前の尤度平均値を  $P$ , 統合後の尤度平均値を  $P'$  として, 次式を満たす場合に軌跡を統合する.

$$\lfloor \log_{10} P' \rfloor - \lfloor \log_{10} P \rfloor > 0 \quad (30)$$

## 9. 実画像実験

提案手法の有効性を確認するために, 運動物体が三つ存在するビデオ画像を対象として特徴点軌跡の分離実験を行った. 実験データは Johns Hopkins 大学の Hopkins155 データベース [7] から菅谷らの方法 [9] でアウトライアを除去した T-Hopkins155 データベース [10] のデータを使用した.

### 9.1 分離過程の確認

図 3 のデータを用いて提案手法によってどのように特徴点軌跡が分離されるかを確認した. 図 3(a) は 30 フレームのビデオ画像から 5 枚を抜き出した画像に追跡した特徴点を描画したものである. このデータに対する分離の流れを図 2 を参照しながら説明する. まず, 入力軌跡が含まれるアフィン空間の次元を推定すると 4 次元であったため, 4 次元 EM アルゴリズムを用いて軌跡を分割した結果が図



(a) 入力画像と追跡した特徴点 (フレーム数 30, 特徴点数 533 点, 抜き出した画像はそれぞれ 1, 8, 15, 22, 30 フレーム目)



(b) 1 段階目の分離結果



(c) 2 段階目の分離結果



(d) 最終結果

図 3: 分離過程の確認

3(b) である. 分離前のアフィン空間への当てはめ尤度は  $5.63 \times 10^{-7}$ , 分離後の尤度は  $8.11 \times 10^{-5}$  であったため, 式 (6) の分離条件を満たし, 図 3(b) の分離結果が確定した (図 2 上段). 次に図 3(b) に示す 2 分割した軌跡に対して, それぞれのアフィン空間の次元を推定すると, 緑点 (□) で示した軌跡は 3 次元, 青点 (□) で示した軌跡は 2 次元と推定された. これにより, 青点 (□) の軌跡の分割はこれで終了する (図 2 中段の右). 図 3(b) の緑点 (□) の軌跡に対して 3 次元 EM アルゴリズムによって分離を行った結果が図 3(c) である. この分離結果についても分離前後のアフィン空間への当てはめ尤度を比較して, 分離前の尤度が  $1.95 \times 10^{-5}$ , 分離後の尤度が  $3.77 \times 10^{-4}$  であったため, 式 (6) を満たし, 図 3(c) の分離結果が確定した (図 2 中段左). さらに分離したこれらの軌跡に対するアフィン空間を推定すると, 図 3(c) の緑点 (□) で示した軌跡は 3 次元, 赤点 (□) で示した軌跡は 2 次元と推定された. 緑点 (□) で示した軌跡を 3 次元 EM アルゴリズムにより分離したが, 分離前の尤度が  $1.30 \times 10^{-4}$ , 分離後の尤度が  $2.89 \times 10^{-4}$  であったため, 式 (6) を満たさず, 分離を行わず (図 2 下段右), これによりすべての分離処理が終了した.

また, このデータでは, 8 節で説明した統合処理を施しても統合される軌跡は存在しなかった. 分離後の最終結果を図 3(d) に示す. 提案手法により三つの運動物体に正しく分離されていることを確認した.

## 9.2 統合処理の確認

図 5 のデータを用いて分離後の特徴点軌跡の統合処理について確認する. 図 5(a) は 30 フレームのビデオ画像から 5 枚を抜き出した画像に追跡した特徴点を描画したものである. まず, 先の実験と同様に入力特徴点軌跡を提案手法を

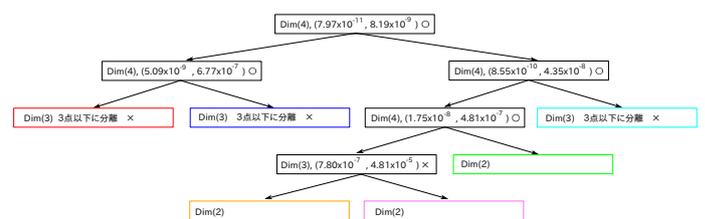


図 4: 図 5 のデータに対する分離の流れ.

用いて分離した. 分離の流れを図 4 に, 分離結果を図 5(b) から (f) に示す. 図 5(f) の分離の最終結果を確認すると, 水色点 (□) とピンク点 (□) で誤分割があり, 青点 (□), 緑点 (□) とオレンジ点 (□), ピンク点 (□) は過分割であることがわかる.

これらの分割した特徴点軌跡に対して統合判定を行った結果を表 1 に示す. 8 節で説明したように, 分割した特徴点軌跡はすぐ直前の分離前の状態に戻ることはないので, この分離結果では, 赤色点 (□) と青点 (□) の軌跡, オレンジ点 (□) とピンク点 (□) の軌跡の統合判定は行わなかった. 表 1 の結果より, 青点 (□) と緑点 (□) とオレンジ点 (□) は統合された. この統合判定に基づく軌跡の統合結果を図 5(g) に示す. 分離誤りは残っているが, 過分割された軌跡は正しく統合される結果を得ることができ, アフィン空間の尤度を用いた統合判定は正しく動作していることが確認できた.

## 9.3 その他のデータに対する分離結果

T-Hopkins155 データベースに含まれる運動物体が三つのデータに対する提案手法による分離結果を図 6 に示す. また, 前述の二つのデータも含めた分離結果の正解率を表 2 に示す. 図 6(a) では背景上の特徴点が異なる運動として二つの異なる運動として分離されてしまい, 統合判定を



(a) 入力画像と追跡した特徴点 (フレーム数 30, 特徴点数 217 点, 抜き出した画像はそれぞれ 1, 8, 22, 30 フレーム目)



(b) 第 1 段階目の分離結果



(c) 第 2 段階目の分離結果 1



(d) 第 2 段階目の分離結果 2



(e) 第 3 段階目の分離結果



(e) 第 4 段階目の分離結果



(f) 分離最終結果



(g) 統合処理を施した結果

図 5: 分離した軌跡の統合実験

表 1: 図 5(f) の分離結果に対する統合判定

|           | 青点 (□) | 水色点 (□)  | 緑点 (□)   | オレンジ点 (□)  | ピンク点 (□)  |
|-----------|--------|--|--|--|---|
| 赤色 (□)    | —      | $4.89 \times 10^{-7}$<br>$8.40 \times 10^{-10}$<br>× | $8.76 \times 10^{-7}$<br>$5.73 \times 10^{-9}$<br>×  | $3.76 \times 10^{-8}$<br>$4.65 \times 10^{-9}$<br>×  | $8.81 \times 10^{-7}$<br>$6.15 \times 10^{-9}$<br>× |
| 青色 (□)    |        | $1.05 \times 10^{-7}$<br>$1.73 \times 10^{-9}$<br>×  | $1.03 \times 10^{-10}$<br>$6.03 \times 10^{-9}$<br>○ | $3.12 \times 10^{-10}$<br>$5.52 \times 10^{-9}$<br>○ | $2.93 \times 10^{-7}$<br>$1.68 \times 10^{-8}$<br>× |
| 水色 (□)    |        |  | $1.36 \times 10^{-6}$<br>$2.24 \times 10^{-9}$<br>×  | $7.48 \times 10^{-7}$<br>$3.02 \times 10^{-9}$<br>×  | $7.07 \times 10^{-8}$<br>$3.57 \times 10^{-9}$<br>× |
| 緑色 (□)    |        |  |  | $1.52 \times 10^{-29}$<br>$6.58 \times 10^{-9}$<br>○ | $4.26 \times 10^{-8}$<br>$9.15 \times 10^{-9}$<br>× |
| オレンジ色 (□) |        |  |  |  | —   |

行っても統合されなかった。しかし、分離結果には誤りはなかった。図 6(b), (d) には若干分離誤りがあったが良好な分離結果を得ることができた。図 6(f) は画像中の左右の車がそれぞれ異なる運動をしているため、全体として三つの運動に分離されるべきであるが、二つの運動にしか分離できなかった。これは、分離の際に過分割された二つの車上の軌跡が今回用いた統合処理によって一つの運動軌跡に統合されてしまったためである。

## 10. まとめ

本稿では、ビデオ画像中に含まれる剛体運動をする物体数が未知な場合の特徴点軌跡から複数の運動を分離する方法を提案した。提案手法では、軌跡が拘束されるアフィン空間を EM アルゴリズムによって 2 分割することを繰り返して運動物体数が未知の特徴点軌跡を分離することを実現した。まず、分離対象となる運動軌跡が含まれるアフィン



(a) フレーム数 15, 特徴点数 112



(b) フレーム数 20, 特徴点数 210



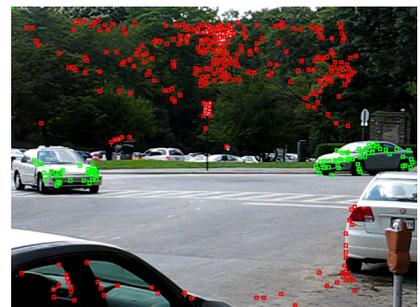
(c) フレーム数 20, 特徴点数 538



(d) フレーム数 61, 特徴点数 212



(e) フレーム数 31, 特徴点数 266



(f) フレーム数 37, 特徴点数 392

図 6: その他のデータに対する分離結果

表 2: 分離結果の正解率

|                | 正しい分離数 / 全特徴点数 | 正解率     |
|----------------|----------------|---------|
| データ 1 (図 3)    | 533 / 533      | 100 %   |
| データ 2 (図 5)    | 213 / 217      | 98.16 % |
| データ 3 (図 6(a)) | 112 / 112      | 100 %   |
| データ 4 (図 6(b)) | 209 / 210      | 99.52 % |
| データ 5 (図 6(c)) | 538 / 538      | 100 %   |
| データ 6 (図 6(d)) | 210 / 212      | 99.06 % |
| データ 7 (図 6(e)) | 266 / 266      | 100 %   |
| データ 8 (図 6(f)) | 344 / 392      | 87.76 % |

空間の次元を幾何学的 MDL を用いて推定し、それに基づいて軌跡を二つの異なるアフィン空間に EM アルゴリズムを用いて分離する。そして、分離前と分離後のアフィン空間への当てはまり具合を尤度で測り、分離を確定させるかどうかを判定する。この処理を分離が行われなくなるまで繰り返すことで、運動物体数をパラメータとすることなく特徴点軌跡の分離を可能にした。また、分離後の結果に対して分離に用いた尤度による統合判定を行うことで、分離の過程で過分割された軌跡を統合する試みも行った。統合処理についてはまだ検討段階であり、今後さらに検討していく予定である。

謝辞: 本研究の一部は日本学術振興会科学研究費 (基盤研究 (C) No. 26330192) の助成による。

#### 参考文献

[1] J. P. Costeira and T. Kanade, A multibody factorization method for independently moving objects, *Int. J.*

*Computer Vision*, **29**-3, pp. 159–179, Sept. 1998.

- [2] E. Elhamifar, R. Vidal, Sparse Subspace Clustering, *CVPR* 2009, pp. 2790–2797, June 2009.
- [3] 金出武雄, コンラッド・ポールマン, 森田俊彦, 因子分解法による物体形状とカメラ運動の復元, 電子情報通信学会論文誌 D-II, **J74-D-II-8**, pp. 1497–1505, 1993.
- [4] 金谷健一, 統計的推測と幾何学的当てはめにおけるモデル選択, 情報処理学会研究報告, 2000-CVIM-122-1, pp. 1–8, Mar. 2000.
- [5] 金谷健一, 黒澤典義, 松永力, モデル選択によるランク推定と複数運動の分離, 情報処理学会研究報告, 2002-CVIM-126-3, pp. 17–24, Mar. 2001.
- [6] 黒澤典義, 金谷健一, アフィン空間分離法による運動物体の分離, 情報処理学会研究報告, 2001-CVIM-123-3, pp. 25–32, Mar. 2001.
- [7] R. Tron and R. Vidal, A benchmark for the comparison of 3-D motion segmentation algorithms, *Proc. IEEE Conf. Comput. Vision Patt. Recog.* June 2007.
- [8] 菅谷保之, 金谷健一, 複数運動分離の教師なし学習による多段階最適化, 情報処理学会研究報告, 2003-CVIM-138-25, pp. 185–192, May 2003.
- [9] 菅谷保之, 松下裕一, 金谷健一, 複数運動ビデオデータベース Hopkins155 の誤追跡除去, 情報処理学会研究報告, 2013-CVIM-187-16, pp. 1–8, May 2013.
- [10] T-Hopkins データベース (<http://www.iim.cs.tut.ac.jp/T-Hopkins/>)
- [11] 田中慎也, 田中勇桂, 原裕貴, 菅谷保之, 金谷健一, 一般主成分分析による複数運動分離の多段階最適化, 情報処理学会研究報告, 2009-CVIM-168-8, 2009.
- [12] R. Vidal, Y. Ma and S. Sastry, Generalized principal component analysis (GPCA), *IEEE Trans. Patt. Anal. Mach. Intell.*, **27**-12, pp. 1945–1959, 2005.
- [13] R. Vidal, R. Tron and R. Hartley, Multiframe motion segmentation with missing data using PowerFactorization and GPCA, *Int. J. Comput. Vision*, **79**-1 pp. 85–105, 2008.