

## 一般主成分分析による複数運動分離の多段階最適化

田中 慎也<sup>†2</sup> 田中 勇 桂<sup>†1</sup> 原 裕 貴<sup>†1</sup>  
菅 谷 保 之<sup>†2</sup> 金 谷 健 一<sup>†1</sup>

ビデオ画像上の複数の運動を分離する新しい方法を提案する。これは菅谷・金谷(2003)の多段階最適化法を改良したものである。菅谷・金谷は初期分類を Costeira・Kanade (1998) の作用行列と幾何学的 AIC による探索によって求めたのに対して、本論文では Vidal ら (2005) の GPCA の考え方に基づいた Taubin 法による 3 次元空間の 2 平面の当てはめを用いる。これにより解が探索なしに直接に求まるだけでなく、それ自身でよい精度の分離が得られ、それを段階的な EM アルゴリズムで最適化する。そして、シミュレーションおよび実ビデオデータを用いて提案方法の有効性を実証し、その理由を視覚的に表示する。

### Multi-stage Optimization of Multi-body Motion Segmentation Using Generalized Principal Component Analysis

SHINYA TANAKA,<sup>†2</sup> YUUKI TANAKA,<sup>†1</sup>  
HIROTAKA HARA,<sup>†1</sup> YASUYUKI SUGAYA<sup>†2</sup>  
and KENICHI KANATANI <sup>†1</sup>

We propose a new method for segmenting multi-body motions in a video stream. Our method improves the multi-stage optimization of Sugaya and Kanatani (2003). While they computed initial segmentation by search based on the shape interaction matrix of Costeira and Kanade (1998) and model selection by the geometric AIC, we fit two planes in 3-D by the Taubin method based on the GPCA of Vidal et al. (2005). As a result, we can obtain without doing any search an initial segmentation which is by itself a very good approximation. It is then progressively optimized by the EM algorithm. Using simulated and real video data, we demonstrate the effectiveness of our method. We also visually display the underlying geometric structure.

### 1. ま え が き

ビデオ画像上を移動する特徴点の軌跡から独立な運動を分離する研究が最近、新たな関心を集めている。この問題の発端は Costeira ら<sup>1)</sup> であり、アフィンカメラを仮定すると、ビデオ画像上の軌跡が高次元空間の 1 点と同一視でき、同一の運動をする点は共通の部分空間に含まれるので(後述)、高次元空間の点を異なる部分集合に分離すればよい。Costeira ら<sup>1)</sup> は因子分解法の計算で定義される「作用行列」の要素の非零判定によって軌跡を複数の部分空間に分類した。このようなアフィンカメラの下での軌跡の部分空間への分類が以後の主流となった。Gear<sup>3)</sup> は行列の標準形を用いてグラフのマッチングに帰着させ、市村ら<sup>6)</sup> は行列の QR 分解を用い、市村<sup>5)</sup> は大津の判別規準を適用した。井上ら<sup>7)</sup> はファジークラスタリングを用いた。黒澤ら<sup>14),15)</sup> は幾何学的 AIC によるモデル選択と投票によるロバスト推定を組合せた。Wu ら<sup>26)</sup> は部分空間の直交分解を適用した。Gruber ら<sup>4)</sup> らは因子分解法に EM アルゴリズムを用いた。菅谷ら<sup>19)</sup> は複数のモデルを組合せる多段階最適化を提案した。Vidal ら<sup>24),25)</sup> 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)」を提唱した。Fan ら<sup>2)</sup> や Yan ら<sup>27)</sup> は高次元空間の点を投票によって部分空間に分類する方法を提案した。Schindler ら<sup>17)</sup> や Rao ら<sup>16)</sup> は最小記述長 (MDL) 原理によるモデル選択を用いている。

本論文では菅谷ら<sup>19)</sup> の多段階最適化法と Vidal ら<sup>24),25)</sup> の一般主成分分析 (GPCA) を組み合わせた新しい分離手法を提案する。菅谷ら<sup>19)</sup> の方法は非常に高い性能があることが Tron ら<sup>23)</sup> によって報告されているが、これは EM アルゴリズムを用いるものである。EM アルゴリズムはよい初期値を必要とする<sup>11)</sup>。菅谷ら<sup>19)</sup> は初期値として Costeira ら<sup>1)</sup> の作用行列と幾何学的 AIC<sup>9)</sup> を組合せた探索を用いた。本論文では独立な運動が 2 個の場合に、初期分類に Vidal ら<sup>24),25)</sup> の GPCA に基づく退化 2 次曲面の当てはめを用い、これに段階的に EM アルゴリズムを適用することにより、高い精度の分離が実現されることを示す。そしてこれを筆者らの画像例、および Johns Hopkins 大学の Hopkins155 画像データ<sup>23)</sup> を用いて実証する。同時に、データの幾何学的な構造を可視化して、なぜよい分離ができるかを視覚的に表示する。

<sup>†1</sup> 岡山大学大学院自然科学研究科

Department of Computer Science, Okayama University

<sup>†2</sup> 豊橋技術科学大学情報工学系

Department of Information and Computer Sciences, Toyohashi University of Technology

## 2. アフィンカメラモデル

$N$  個の特徴点  $\{p_\alpha\}$  を  $M$  枚の画像に渡って追跡し、第  $\kappa$  画像における  $\alpha$  番目の特徴点  $p_\alpha$  の画像座標を  $(x_{\kappa\alpha}, y_{\kappa\alpha})$ ,  $\kappa = 1, \dots, M, \alpha = 1, \dots, N$  とする．そしてその運動履歴を次の  $2M$  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 点と同一視できる．本論文ではカメラの光軸を  $Z$  軸とする  $XYZ$  カメラ座標系をとり、これに相対的にシーンが運動すると解釈する．シーン中に物体座標系を任意に固定し、特徴点  $p_\alpha$  のそ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．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次式となる．

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

平行投影や弱透視投影や疑似透視投影を抽象化した「アフィンカメラ」<sup>8)</sup> は、3 次元点  $r_{\kappa\alpha}$  が次のように画像上に投影されると仮定するものである．

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

ここに  $A_\kappa, b_\kappa$  は第  $\kappa$  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) を代入すると、式 (3) は次のように書ける．

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

ただし、 $\tilde{m}_{0\kappa}, \tilde{m}_{1\kappa}, \tilde{m}_{2\kappa}, \tilde{m}_{3\kappa}$  は第  $\kappa$  フレームでのカメラの位置や内部パラメータで決まる 2 次元ベクトルである．これをフレーム  $\kappa = 1, \dots, M$  に渡って式 (1) のように縦に並べると、式 (1) の軌跡ベクトル  $p_\alpha$  が次のように書ける．

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

ここに、 $m_i, i = 0, 1, 2, 3$  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．

## 3. 軌跡の拘束条件

式 (5) は、同一の剛体運動をする特徴点  $p_\alpha$  の軌跡が  $2M$  次元空間中の  $\{m_0, m_1, m_2, m_3\}$  の張る「4 次元部分空間」に含まれることを意味する．したがって、観測した特徴点

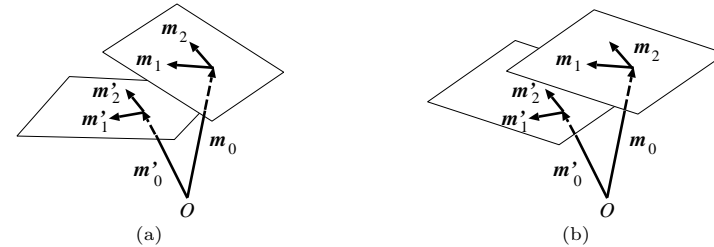


図 1 (a) 平面運動では物体と背景の軌跡ベクトルはそれぞれ 2 次元アフィン空間に含まれる．(b) 物体も背景も回転せずに単に並進すると、二つの 2 次元アフィン空間互いに平行になる．

を異なる剛体運動に分離するには、それらの軌跡ベクトル  $\{p_\alpha\}$  を互いに異なる 4 次元部分空間に分類すればよい．しかし、式 (5) において  $m_0$  の係数はすべての  $\alpha$  に共通に 1 である．このため  $p_\alpha$  は  $\{m_0, m_1, m_2, m_3\}$  の張る 4 次元部分空間内の「3 次元アフィン空間」に含まれる．したがって、特徴点の運動を分離するには軌跡ベクトル  $\{p_\alpha\}$  を互いに異なる 3 次元アフィン空間に分類してもよい．

ところが、通常のシーンでは物体と背景が画像内で 2 次元な運動をし、回転は  $Z$  軸（光軸）回りのみのことが多い．以下これを「平面運動」と呼ぶ（奥行き方向の並進があってもよい．これはアフィンカメラでは観測できないので、並進は  $XY$  面内であるとみなせる）．このとき、物体座標系の基底ベクトル  $k_\kappa$  を  $Z$  軸方向に取ればアフィンカメラのもとでは画像面に投影されないから、式 (5) の  $m_3$  が 0 となり、背景も物体も軌跡ベクトルが  $m_0, m_1, m_2$  の張る「2 次元アフィン空間」に含まれる（図 1(a)）．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、背景に共通であるから、式 (5) の  $m_1, m_2$  も物体、背景に共通になり、それぞれの 3 次元アフィン空間は互いに「平行な 2 次元アフィン空間」になる（図 1(b)）．

平面運動によって 3 次元アフィン空間が 2 次元に退化すると Costeira ら<sup>1)</sup> の作用行列に基づく方法では正しい分離を行なうことができない．さらに二つの 2 次元アフィン空間が平行になると、それらが一つ 3 次元アフィン空間に含まれ、3 次元アフィン空間による分離ができない．菅谷ら<sup>19)</sup> の多段階最適化はこれに対処するものであり、平行なアフィン空間の当てはめから始めて、得られた解により一般のモデルを段階的に EM アルゴリズムで当てはめるものである．その初期分類は Costeira ら<sup>1)</sup> の作用行列と幾何学的 AIC<sup>9)</sup> による探索に基く黒澤ら<sup>15)</sup> のアフィン空間分離法を用いた．本論文で Vidal ら<sup>24)</sup> の GPCA に基づ

く方法により、探索を行うことなく初期分類を計算し、段階的に EM アルゴリズムを適用する。そして、これらの計算を高次元空間を主成分分析によって圧縮した最小次元の空間において行う。

#### 4. 軌跡の圧縮

式 (1) の軌跡ベクトルが  $2M$  次元空間の二つの 3 次元アフィン空間上に分離されているとすると、その二つの 3 次元アフィン空間を含む 7 次元アフィン空間が存在する。したがって、軌跡の分類はその 7 次元アフィン空間内で考えればよい (データに誤差があるとその 7 次元アフィン空間から多少はみ出すが、その 7 次元アフィン空間に直交する誤差成分は分類には影響しない)。その 7 次元アフィン空間を原点を通るように平行移動し、7 個の基底ベクトルをとってデータをそれらの線形結合で表せば、データを 7 次元空間の点と同一視できる。同様に、軌跡ベクトルが  $2M$  次元空間中の二つの 2 次元アフィン空間上にあるときは、その二つの 2 次元アフィン空間を含む 5 次元アフィン空間が存在するので、データは 5 次元空間の点と同一視できる。さらに  $2M$  次元空間中の二つの 2 次元アフィン空間が平行であれば、その二つの 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ので、データは 3 次元空間の点と同一視できる。

$2M$  次元空間の軌跡ベクトルを  $d$  次元空間の点と同一視するには次の計算を行えばよい。

- (1) 軌跡ベクトル  $\{p_\alpha\}$  の重心  $p_C$  とそれからの差  $\tilde{p}_\alpha$  を計算する。

$$p_C = \frac{1}{N} \sum_{\alpha=1}^N p_\alpha, \quad \tilde{p}_\alpha = p_\alpha - p_C \quad (6)$$

- (2) 次の  $2M \times N$  行列を特異値分解する。

$$\begin{pmatrix} \tilde{p}_1, \dots, \tilde{p}_N \end{pmatrix} = U \text{diag}(\sigma_1, \dots, \sigma_r) V^\top \quad (7)$$

ただし  $r = \min(2M, N)$  であり、 $U$  は  $r$  本の正規直交系の列をもつ  $2M \times r$  行列、 $V$  は  $r$  本の正規直交系の列をもつ  $N \times r$  行列である。そして  $\sigma_1 \geq \dots \geq \sigma_r (\geq 0)$  が特異値であり、 $\text{diag}(\sigma_1, \dots, \sigma_r)$  はそれらを対角要素とする対角行列である。

- (3) 行列  $U$  の第  $i$  列を  $u_i$  として、次の  $d$  次元ベクトル  $r_\alpha, \alpha = 1, \dots, N$  を計算する。

$$r_\alpha = \begin{pmatrix} (\tilde{p}_\alpha, u_1), \dots, (\tilde{p}_\alpha, u_d) \end{pmatrix}^\top \quad (8)$$

#### 5. 初期分類

本論文で提案する初期分類法は、軌跡ベクトルを 3 次元空間の点と同一視して、それらに 2 平面 (= 2 次元アフィン空間) を当てはめるものである。物体と背景が共に並進していれば、すべての点は平行な 2 平面上にある。データに誤差があったり、回転成分があればこれは成り立たないが、誤差が小さく、運動が並進に近い場合は、ほぼ平行な 2 平面がほぼ当てはまると期待される。3 次元空間の平面  $Ax + By + Cz + D = 0$  はベクトル  $n, x$

と置けば  $(n, x) = 0$  と書ける。以下、ベクトル  $a, b$  の内積を  $(a, b)$  と書く。ただし、 $n$  は何倍しても同じ平面を表すので、 $n$  には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、これらを次のように一つの式にまとめることができる。

$$n = (A, B, C, D)^\top, \quad x = (x, y, z, 1)^\top \quad (9)$$

ただし、対称行列  $Q$  を次のように置いた (2 次形式の係数行列は対称成分、すなわちその転置と足して 2 で割ったもののみが意味がある<sup>11)</sup>)。

$$(n_1, x)(n_2, x) = (x, n_1 n_2^\top x) = (x, Qx) = 0 \quad (10)$$

ただし、対称行列  $Q$  を次のように置いた (2 次形式の係数行列は対称成分、すなわちその転置と足して 2 で割ったもののみが意味がある<sup>11)</sup>)。

$$Q = \frac{n_1 n_2^\top + n_2 n_1^\top}{2} \quad (11)$$

この式からわかるように、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<sup>10)</sup>。固有値を大きい順に  $\lambda_1 > 0 > -\lambda_2$  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、 $Q$  は次のようにスペクトル分解される<sup>11)</sup>。

$$Q = \lambda_1 u_1 u_1^\top - \lambda_2 u_4 u_4^\top = \left( \sqrt{\frac{\lambda_1}{2}} u_1 + \sqrt{\frac{\lambda_2}{2}} u_4 \right) \left( \sqrt{\frac{\lambda_1}{2}} u_1 - \sqrt{\frac{\lambda_2}{2}} u_4 \right)^\top + \left( \sqrt{\frac{\lambda_1}{2}} u_1 - \sqrt{\frac{\lambda_2}{2}} u_4 \right) \left( \sqrt{\frac{\lambda_1}{2}} u_1 + \sqrt{\frac{\lambda_2}{2}} u_4 \right)^\top \quad (12)$$

式 (11) と見比べ、 $n_1, n_2$  (したがって  $Q$ ) に定数倍の不定性があることを考慮すると  $n_1, n_2$  が次のように定まる。

$$n_1 = \sqrt{\lambda_1} u_1 + \sqrt{\lambda_2} u_4, \quad n_2 = \sqrt{\lambda_1} u_1 - \sqrt{\lambda_2} u_4 \quad (13)$$

したがって、誤差のために完全には 2 平面上にはない  $N$  点を式 (9) のベクトル  $x$  形で  $x_1, \dots, x_N$  と表し、

$$(x_\alpha, Qx_\alpha) \approx 0, \quad \alpha = 1, \dots, N \quad (14)$$

となる  $Q$  を求めれば (解法は次節)、式 (13) によって 2 平面を指定する  $n_1, n_2$  が定まる。3 次元空間の点  $(x, y, z)$  から平面  $Ax + By + Cz + D = 0$  までの距離  $d$  は次のように書ける<sup>10)</sup>。

$$d = \frac{|Ax + By + Cz + D|}{\sqrt{A^2 + B^2 + C^2}} \quad (15)$$

これを計算して、各データ点を近いほうの平面に割りつけることによって、点集合を 2 クラス

に分類できる．これを以後の EM アルゴリズムの初期値とする．Vidal ら<sup>24),25)</sup> の GPCA は複数の同次式（定数項のないもの）で表される（原点を通る）部分空間をまとめて一つの次数の高い多項式で表し，点データにそのような多項式を当てはめることによって部分空間に分類するものである．ここではその考え方を 2 平面（= 二つのアフィン空間）の当てはめに転用したものである．

## 6. Taubin 法による当てはめ

式 (14) となる  $Q$  の計算法を考える． $x$  を式 (9) のように定義すると， $Q$  が一般の対称行列のとき，式  $(x, Qx) = 0$  は 3 次元空間の 2 次曲面（楕円面，双曲面，放物面およびそれらの退化）を表す<sup>10)</sup>．2 平面は退化した 2 次曲面であり<sup>10)</sup>，式 (14) となる  $Q$  の計算は，3 次元空間の点集合に 2 次曲面を当てはめる問題とみなせる．これは数学的には平面上の点列に 2 次曲線（楕円，双曲線，放物線およびそれらの退化）を当てはめる問題と同じであり，同じ計算手法が使える<sup>13),20)</sup>．9 次元ベクトル  $z_\alpha, u$  を

$$\begin{aligned} z_\alpha &= (x_\alpha^2, y_\alpha^2, z_\alpha^2, 2y_\alpha z_\alpha, 2z_\alpha x_\alpha, 2x_\alpha y_\alpha, 2x_\alpha, 2y_\alpha, 2z_\alpha)^\top \\ v &= (Q_{11}, Q_{22}, Q_{33}, Q_{23}, Q_{31}, Q_{12}, Q_{41}, Q_{42}, Q_{43})^\top \end{aligned} \quad (16)$$

と置くと，式 (14) は次のように書ける．

$$(z_\alpha, v) + Q_{44} \approx 0, \quad \alpha = 1, \dots, N \quad (17)$$

このような  $v, Q_{44}$  を反復なしに計算する方法として知られているのが次の Taubin 法であり，単純な最小二乗法より精度がよいことが知られている（理論的には最尤推定のほうが精度が高いが<sup>12),13),20)</sup>，最尤推定は反復を要し，誤差が非常に大きいとき収束しないことがある<sup>12),13),20)</sup>．

まず  $x_\alpha, y_\alpha, z_\alpha$  に独立に期待値 0，標準偏差  $\sigma$  の正規分布に従う誤差  $\Delta x_\alpha, \Delta y_\alpha, \Delta z_\alpha$  が加わるときの  $z_\alpha$  の誤差を  $\Delta z_\alpha$  とする．式 (16) より  $\Delta x_\alpha, \Delta y_\alpha, \Delta z_\alpha$  の 2 次以上の項を無視すると次のようになる．

$$\Delta z_\alpha = (2x_\alpha \Delta x_\alpha, 2y_\alpha \Delta y_\alpha, 2z_\alpha \Delta z_\alpha, 2\Delta y_\alpha z_\alpha + 2y_\alpha \Delta z_\alpha, \dots, 2\Delta z_\alpha)^\top \quad (18)$$

次に  $z_\alpha$  の共分散行列  $V[z_\alpha] = E[\Delta z_\alpha \Delta z_\alpha^\top]$  を計算する．関係  $E[\Delta x_\alpha] = E[\Delta y_\alpha] = E[\Delta z_\alpha] = 0$ ， $E[\Delta y_\alpha \Delta z_\alpha] = E[\Delta z_\alpha \Delta x_\alpha] = E[\Delta x_\alpha \Delta y_\alpha] = 0$ ， $E[\Delta x_\alpha^2] = E[\Delta y_\alpha^2] = E[\Delta z_\alpha^2] = \sigma^2$  を代入すると次のようになる．

$$V[z_\alpha] = \sigma^2 V_0[z_\alpha] \quad (19)$$

$$V_0[z_\alpha] = \begin{pmatrix} x_\alpha^2 & 0 & 0 & 0 & z_\alpha x_\alpha & x_\alpha y_\alpha & x_\alpha & 0 & 0 \\ * & y_\alpha^2 & 0 & y_\alpha z_\alpha & 0 & x_\alpha y_\alpha & 0 & y_\alpha & 0 \\ * & * & z_\alpha^2 & y_\alpha z_\alpha & z_\alpha x_\alpha & 0 & 0 & 0 & z_\alpha \\ * & * & * & y_\alpha^2 + z_\alpha^2 & x_\alpha y_\alpha & z_\alpha x_\alpha & 0 & z_\alpha & y_\alpha \\ * & * & * & * & z_\alpha^2 + x_\alpha^2 & y_\alpha z_\alpha & z_\alpha & 0 & x_\alpha \\ * & * & * & * & * & x_\alpha^2 + y_\alpha^2 & y_\alpha & x_\alpha & 0 \\ * & * & * & * & * & * & 1 & 0 & 0 \\ * & * & * & * & * & * & * & 1 & 0 \\ * & * & * & * & * & * & * & * & 1 \end{pmatrix} \quad (20)$$

ただし，\* は対称位置の要素をコピーをすることを表す．Taubin 法は次の  $J_{TB}$  を最小化するものである<sup>12),13),20)</sup>．

$$J_{TB} = \frac{\sum_{\alpha=1}^N \left( (z_\alpha, v) + Q_{44} \right)^2}{\sum_{\alpha=1}^N (v, V_0[z_\alpha] v)} \quad (21)$$

これを最小化する  $v, Q_{44}$  は次のように得られる<sup>13),20)</sup>．

(1)  $\{z_\alpha\}$  の平均  $\bar{z}$ ，および各  $z_\alpha$  の平均からのずれ  $\tilde{z}_\alpha$  を次のように計算する．

$$\bar{z} = \frac{1}{N} \sum_{\alpha=1}^N z_\alpha, \quad \tilde{z}_\alpha = z_\alpha - \bar{z} \quad (22)$$

(2) 次の行列  $9 \times 9$  行列  $M_{TB}, N_{TB}$  を計算する．

$$M_{TB} = \sum_{\alpha=1}^N \tilde{z}_\alpha \tilde{z}_\alpha^\top, \quad N_{TB} = \sum_{\alpha=1}^N V_0[z_\alpha] \quad (23)$$

(3) 一般固有値問題

$$M_{TB} v = \lambda N_{TB} v \quad (24)$$

の最小一般固有値  $\lambda$  に対する単位一般固有ベクトル  $v$  を計算する．

(4)  $Q_{44}$  を次のように計算する．

$$Q_{44} = -(\bar{z}, v) \quad (25)$$

## 7. 多段階最適化

本論文の提案方法は，以上のようにして得られる初期分類から出発して，次の順に EM アルゴリズムによってアフィン空間の当てはめを行うものである．

(1) 3 次元空間の平行な 2 平面の当てはめ

(2) 5 次元空間の二つの 2 次元アフィン空間の当てはめ

(3) 7 次元空間の二つの 3 次元アフィン空間の当てはめ

物体も背景も並進のみなら第 1 段階で最適解が得られ，その解は第 2, 3 段階でも最適解

であるから何の変更も行われぬ。物体と背景が回転を伴う平面的運動なら第2段階で最適解が得られ、その解は第3段階でも最適解であるから何の変更も行われぬ。物体と背景が共に3次元的な運動をしていれば第3段階で最適解が得られる。このように、退化した運動は一般の運動の特殊な場合であることから、特殊な運動から順に判定を行えば、どのような運動でも正しく判定できるというのが菅谷ら<sup>19)</sup>の多段階最適化の考え方である。

$n$ 次元に圧縮したデータ  $r_\alpha, \alpha = 1, \dots, N$  に二つの  $d$ 次元アフィン空間 ( $n \geq 2d+1$ ) を当てはめて、2クラスに分類するEMアルゴリズム<sup>11)</sup>は次のようになる。

(1) 初期分類を用いて  $r_\alpha$  の各クラス  $k = 1, 2$  への所属を表わす重み  $W_\alpha^{(k)}$  を次のように定義する。

$$W_\alpha^{(k)} = \begin{cases} 1 & \text{点 } r_\alpha \text{ がクラス } k \text{ に属するとき} \\ 0 & \text{それ以外} \end{cases} \quad (26)$$

(2) 各クラス  $k = 1, 2$  について次の計算を行なう。

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

$$w^{(k)} = \frac{1}{N} \sum_{\alpha=1}^N W_\alpha^{(k)} \quad (27)$$

(b)  $w^{(k)} \leq d/N$  なら計算を中断する(点数が少なすぎて  $d$ 次元アフィン空間が張れない場合)。

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

$$r_C^{(k)} = \frac{\sum_{\alpha=1}^N W_\alpha^{(k)} r_\alpha}{\sum_{\alpha=1}^N W_\alpha^{(k)}} \quad (28)$$

(d) クラス  $k$  のモーメント行列

$$M^{(k)} = \frac{\sum_{\alpha=1}^N W_\alpha^{(k)} (r_\alpha - r_C^{(k)}) (r_\alpha - r_C^{(k)})^\top}{\sum_{\alpha=1}^N W_\alpha^{(k)}} \quad (29)$$

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。

$$P^{(k)} = \sum_{i=1}^d u_i^{(k)} u_i^{(k)\top}, \quad P_\perp^{(k)} = I - P^{(k)} \quad (30)$$

(3) 二乗ノイズレベル  $\sigma^2$  を次のように推定する。

$$\hat{\sigma}^2 = \min \left[ \frac{N}{(n-d)(N-d-1)} \text{tr}(w^{(1)} P_\perp^{(1)} M^{(1)} P_\perp^{(1)} + w^{(2)} P_\perp^{(2)} M^{(2)} P_\perp^{(2)}), \sigma_{\min}^2 \right] \quad (31)$$

ただし、 $\sigma_{\min}$  は微小定数であり(誤差のない、あるいは極めて小さいデータで計算が破綻するのを防止する微小定数であり、1.0(画素)程度に設定する)、tr は行列のトレースである。

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

$$V^{(k)} = P^{(k)} M^{(k)} P^{(k)} + \hat{\sigma}^2 P_\perp^{(k)} \quad (32)$$

(5) 各点  $r_\alpha, \alpha = 1, \dots, N$  について次の計算をする。

(a) 点  $r_\alpha$  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。

$$P(\alpha|k) = \frac{e^{-(r_\alpha - r_C^{(k)}, V^{(k)-1}(r_\alpha - r_C^{(k)})) / 2}}{\sqrt{\det V^{(k)}}} \quad (33)$$

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

$$W_\alpha^{(k)} = \frac{w^{(k)} P(\alpha|k)}{w^{(1)} P(\alpha|1) + w^{(2)} P(\alpha|2)} \quad (34)$$

(6) ステップ2に戻って、 $\{W_\alpha^{(k)}\}$  が収束するまで反復する。

(7) 収束したら(または中断したら)各点  $r_\alpha$  を  $W_\alpha^{(k)}, k = 1, 2$  が大きいクラス  $k$  に分類する。

上記の手順で  $n = 5, d = 2$  とすれば多段階最適化の第2段階となり、 $n = 7, d = 3$  とすれば第3段階となる。しかし、第1段階は2平面が平行であるという制約が必要となる。そのためには  $n = 3, d = 2$  とし、ステップ2(d)で得られる  $M^{(k)}$  を合併したモーメント行列

$$M = w^{(1)} M^{(1)} + w^{(2)} M^{(2)} \quad (35)$$

の  $n$ 個の固有値  $\lambda_1 \geq \dots \geq \lambda_n$  に対応する単位固有ベクトル  $u_1, \dots, u_n$  を計算し、各アフィン空間への射影行列  $P^{(k)}$  とその法線方向への射影行列  $P_\perp^{(k)}$  を共通に  $P^{(1)} = P^{(2)} = P, P_\perp^{(1)} = P_\perp^{(2)} = P_\perp$  とする。

$$P = \sum_{i=1}^d u_i u_i^\top, \quad P_\perp = I - P \quad (36)$$

ステップ3のノイズレベル  $\sigma^2$  の推定は次のように行う。

$$\hat{\sigma}^2 = \min \left[ \frac{N}{(n-d)(N-d-2)} \text{tr}(P_\perp M P_\perp), \sigma_{\min}^2 \right] \quad (37)$$

それ以外は同じである。ただし、一つ問題がある。シミュレーションにおいて、誤差が全くなく、厳密なアフィンカメラモデルにおける完全な平面運動データを生成すると、7次元空間において、軌跡ベクトルの各クラスの点の分布が厳密に2次元的になり、その共分散行列のランクが2となって、正規分布に基づく尤度  $P(\alpha|k)$  が定義できない(正規分布の共分散行列は正値対称行列であり、式(32)の  $V^{(k)}$  はランク  $n$ 、式(33)の分母の  $\det V^{(k)}$  は正

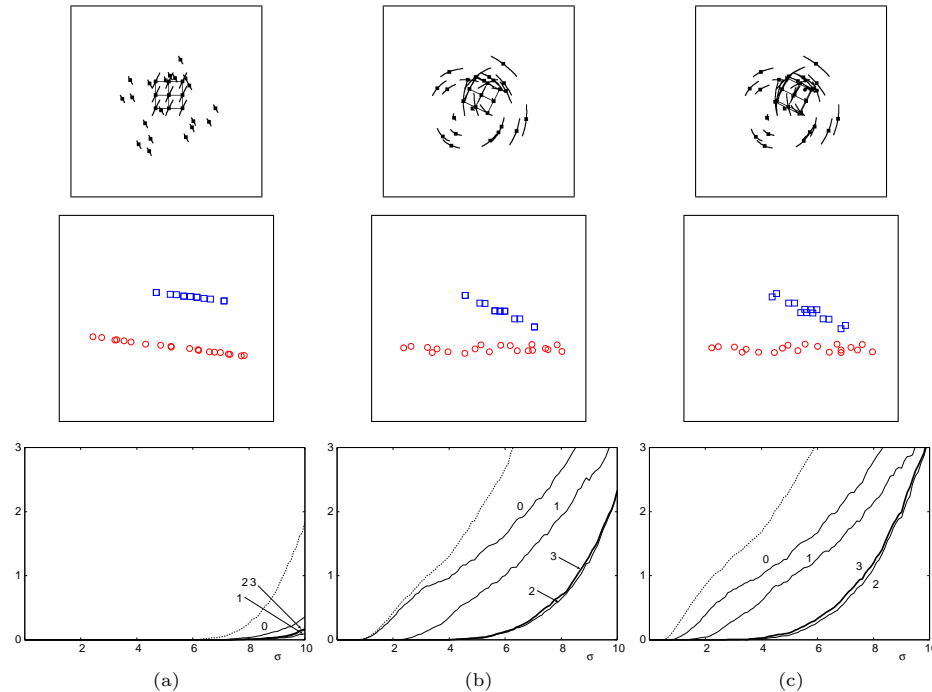


図 2 上: 背景点 (20 個) と物体点 (14 個) の運動 (a) 並進運動, (b) 平面運動, (c) 一般の 3 次元運動. 下: 多段階最適化の効果 (誤差を変えた 5000 回の試行の平均誤り率). 横軸は誤差の標準偏差  $\sigma$ , 縦軸は平均の分離誤り率. 0) Taubin 法による初期分類. 1) 3 次元空間の平面の当てはめ. 2) 5 次元空間のアフィン空間の当てはめ. 3) 7 次元空間のアフィン空間の当てはめ. 点線は最小二乗法による初期分類.

でなければならない). そこで提案システムでは, このような, あるいはそれに非常に近い状況が生じているかどうかを幾何学的 AIC<sup>(9)</sup> によって判定して, 判定されれば 7 次元空間の 3 次元アフィン空間を 2 次元アフィン空間に置き換えている (詳細省略).

## 8. 実験

### 8.1 シミュレーション実験

図 2 の上欄はシーン中で 20 個の背景点と 14 個の物体点がそれぞれ独立に移動している  $512 \times 512$  画素を想定したシミュレーション画像である. 図中の太い曲線は 10 フレームに

渡る移動軌跡を第 5 フレーム上に記入したものである. 見やすくするために物体点は線分で結んでいる. 図 2(a) では物体も背景も回転せず並進している. 図 2(b) は回転を含む平面運動である. 図 2(c) は物体も背景も一般の 3 次元運動をしている.

中欄はそれぞれの例の軌跡ベクトルを 3 次元に圧縮して, 3 次元空間内に表示したものである. 運動 (a) は回転がないので, 軌跡ベクトルは平行な 2 平面に含まれている. 運動 (b), (c) は本来は平行な 2 平面に含まれないにもかかわらず, おおよそ平行な 2 平面に近い分布をしている. このことから Taubin 法による初期分類それ自体が既に高い精度を達成すると期待される.

下段は上段のそれぞれの例に対して, 各点の位置の  $x, y$  座標に独立に期待値 0, 標準偏差  $\sigma$  (画素) の正規分布に従う誤差を加えて背景点と物体点を分離し, 横軸の各  $\sigma$  に対して誤差を変えて 5000 回試行した平均の分離誤り率 (誤って分類された点の割合) を各段階ごとに縦軸にプロットしたものである. 点線は Taubin 法の代わりに最小二乗法 (Vidal ら<sup>24),25</sup>) の GPCA は最小二乗法に基づいている) で計算した初期分類の誤り率である. これから分かるように, 運動 (a) では初期分類がすでに十分正しく, 第 1 段階でほぼ正しい分類が得られている. そして, 運動 (b) では第 2 段階で, 運動 (c) では第 3 段階で, ほぼ正しい分類が得られている. また, 初期分類に用いる Taubin 法は最小二乗法よりも精度が高いことも分かる.

### 8.2 実ビデオ画像実験

図 3 の上段は Johns Hopkins 大学で作成された画像 Hopkins155 データベース<sup>\*1,23</sup>) から取り出した 6 例である. 下段はその軌跡の 3 次元表示である. 表 1 は提案方法<sup>\*2</sup>) の段階ごとの正解率であり, 比較として菅谷らの方法 (プログラムは注 \*2 のサイトに公開), 注 \*1 のサイトにプログラムが公開されている Vidal らの方法<sup>24</sup>), RANSAC による方法, Yan らの方法<sup>27</sup>) の結果を示すこれから分かるように, 提案手法ではどの例に対しても比較的早い段階で高い正解率が得られ, 最終的にすべて正解率が 100%となっている. それに対して他の方法では必ずしも 100%正しい分離がなされていない.

従来の研究では特定の画像例やデータベースに対しての平均正解率を計算して, 他の方法との優位性を主張しているものが多い. しかし, 正解率はどのような画像例やデータベースを用いるかに大きく依存する. 特定データによる評価しかできないのは, 方法が抽象的, 数

\*1 <http://www.vision.jhu.edu/data/hopkins155>

\*2 <http://www.iim.ics.tut.ac.jp/~sugaya/public-e.html>

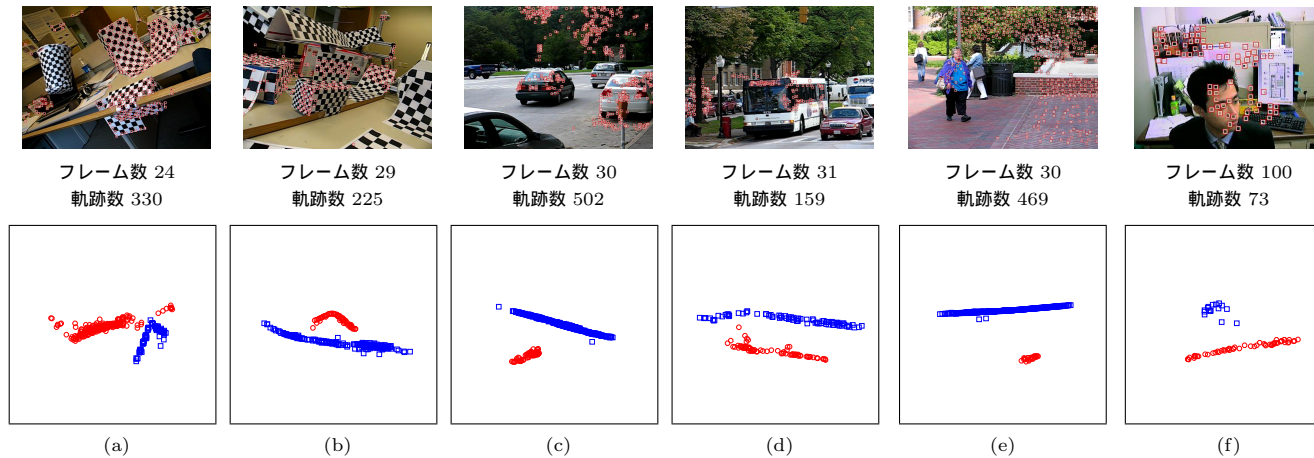


図3 Hopkins155 データベースのビデオ画像の特徴点(上)と軌跡の3次元表示(下)。

表1 図3の画像例に対する提案手法の各段階の正解率(%)と菅谷ら, Vidalら, RANSAC, Yanらの結果。

	(a)	(b)	(c)	(d)	(e)	(f)
初期分類	88.8	99.1	98.0	100.0	100.0	98.6
第1段階	99.7	99.6	100.0	100.0	100.0	100.0
第2段階	98.8	99.6	100.0	100.0	100.0	100.0
第3段階	100.0	100.0	100.0	100.0	100.0	100.0
菅谷ら	99.7	99.6	100.0	100.0	100.0	100.0
Vidalら	88.2	99.6	99.2	99.4	100.0	100.0
RANSAC	91.8	99.6	96.6	97.5	100.0	100.0
Yanら	98.5	98.2	97.4	94.3	99.8	80.8

学的な議論に基いているため、どのような状況で正しく分離でき、どのような状況では正しく分離されにくいのが考察できないためである。それに対して、本論文では運動のタイプとその退化などの幾何学的構造の解析に立脚しているため、そのような考察が可能となる。図3の下段から分かるように、複雑な3次元運動をしているように見えても、3次元表示では軌跡がほぼ平行な2平面上に載ることが多く、提案方法の高い性能はこの事実立脚している。従来この事実には十分注意が払われていなかった。

## 9. まとめ

本論文ではビデオ画像上の複数の運動を分離する新しい方法を提案した。これは著者らの以前の方法(菅谷ら<sup>19)</sup>)を改良したものである。最大の改良点は、菅谷ら<sup>19)</sup>では初期分類をCosteiraら<sup>1)</sup>の作用行列と幾何学的AIC<sup>9)</sup>による探索によって求たのに対して、本論文ではVidalら<sup>24),25)</sup>のGPCAの考え方に基づいたTaubin法による3次元空間の2平面的当てはめを用いたことである。これにより解が探索なしに直接に求まるだけでなく、それ自身でよい精度の分離が得られ、これを段階的なEMアルゴリズムによって最適化した。さらに菅谷ら<sup>19)</sup>では8次元空間において分離を行っていたが、本論文では段階ごとに3次元, 5次元, 7次元と次元を上げた。また菅谷ら<sup>19)</sup>では誤差のないシミュレーションデータで計算が破綻することがあったが、本論文ではそれを判定して破綻を防いでいる。そして、

シミュレーションおよび実ビデオデータを用いて提案方法の有効性を実証した。さらに軌跡データを3次元表示によって、提案方法の幾何学的な構造を視覚的に示した。

謝辞: 本研究に関して有益な議論をして頂いた米国 Johns Hopkins 大学の René Vidal 博士の感謝する。本研究の一部は文部科学省科学研究費基盤研究 C (No. 21500172) 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, 159-179, Sept. 1998.
- 2) Z. Fan, J. Zhou and Y. Wu, Multibody grouping by inference of multiple subspace from high-dimensional data using oriented-frames, *IEEE Trans Patt. Anal. Mach. Intell.*, **28**-1 (2006-1), 91-105.
- 3) C. W. Gear, Multibody grouping from motion images, *Int. J. Comput. Vision*, **29**-2, 133-150, Aug./Sept. 1998.
- 4) A. Gruber and Y. Weiss, Multibody factorization with uncertainty and missing data using the EM algorithm *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, Vol.1, pp. 769-775, June-July 2004, Washington, DC, U.S.A.
- 5) 市村直幸, 形状空間への直交射影行列と判別基準を用いた複数運動の分割, 情報処理学会研究報告, 2000-CVIM-120-3, 17-24, Jan. 2000.
- 6) 市村直幸, 富田文明, 形状行列からの特徴選択に基づく動きの分割, 電子情報通信学会論文誌 D-II, **J81-D-II-12**, 2757-2766, Dec. 1998.



- 7) 井上光平, 浦浜喜一, クラスタリングによる動画像中の複数物体の分離, 電子情報通信学会技術研究報告, PRMU2000-45, 29-36, July 2000.
- 8) 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**, 1497-1505, Aug. 1993.
- 9) K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, the Netherlands, 1996; Reprinted, Dover, New York, NY, U.S.A., 2005.
- 10) 金谷健一, 「形状CADと図形の数学」, 共立出版, 1998.
- 11) 金谷健一, 「これなら分かる最適化数学—基礎原理から計算手法まで—」, 共立出版, 2005.
- 12) K. Kanatani, Statistical optimization for geometric fitting: Theoretical accuracy analysis and high order error analysis, *Int. J. Comput. Vision*, **80-2** (2008-11), 167-188.
- 13) K. Kanatani and Y. Sugaya, Performance evaluation of iterative geometric fitting algorithms, *Comp. Stat. Data Anal.*, **52-2** (2007-10), 1208-1222.
- 14) 黒澤 典義, 金谷 健一, 部分空間分離法とモデル選択による運動物体の分離, 情報処理学会研究報告, 2000-CVIM-124-4, 25-32, Nov. 2000.
- 15) 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-125-3, 25-32, Mar. 2001.
- 16) S. R. Rao, R. Tron, R. Viadl and Y. Ma, Motion segmentation via robust subspace separation in the presence of outlying, incomplete, or corrupted trajectories, *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, June 2008, Anchorage, AK, U.S.A.
- 17) K. Schindler, D. Suter and H. Wang, A model-selection framework for multibody structure-and-motion of image sequences, *Int. J. Comput. Vision*, **79-2** (2008-8), pp. 159-177.
- 18) 菅谷保之, 金谷健一, 部分空間分離法による特徴点追跡のアウトライア除去, 情報処理学会研究報告, 2002-CVIM-133-24, 177-184, May 2002.
- 19) 菅谷 保之, 金谷 健一, 複数運動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, 185-192, May 2003.
- 20) 菅谷保之, 金谷健一, 画像の三次元理解のための最適化計算 [II] —楕円の当てはめ—, 電子情報通信学会誌, **92-4** (2009-4), 301-306.
- 21) G. Taubin, Estimation of planer curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **13-11** (1991-11), 1115-1138.
- 22) C. Tomasi and T. Kanade, *Detection and Tracking of Point Features*, CMU Tech. Rep. CMU-CS-91-132, Apr. 1991; <http://vision.stanford.edu/~birc/klt/>.
- 23) R. Tron and R. Vidal, A benchmark for the comparson of 3-D motion segmentation algorithms, *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, June 2007, Minneapolis, MN, U.S.A.
- 24) R. Vidal, Y. Ma and S. Sastry, Generalized principal component analysis (GPCA), *IEEE Trans. Patt. Anal. Mach. Intell.*, **27-12** (2005-12), 1945-1959.
- 25) R. Vidal, R. Tron and R. Hartley, Multiframe motion segmentation with missing data using PowerFactorization and GPCA, *Int. J. Comput. Vision*, **79-1** (2008-8), 85-105.
- 26) Y. Wu, Z. Zhang, T. S. Huang and J. Y. Lin, Multibody grouping via orthogonal subspace decomposition, sequences under affine projection, *Proc. IEEE Conf. Computer Vision Pattern Recog.*, Vol.2, pp.695-701, Kauai, Hawaii, U.S.A., Dec. 2001.
- 27) J. Yan and M. Pollefeys, A general framework for motion segmentation: Independent, articulate, rigid, non-rigid, degenerate and nondegenerate, *Proc. Euro. Conf. Comput. Vision.*, Vol. 4, pp. 94-104, Graz, Austria, May 2006.