モデル選択によるランク推定と複数運動の分離

金谷 健一 * 　黒澤 典義 * 　松永 　力 †

* 群馬大学工学部情報工学科 　†(株) 朋栄 放送システム開発部

複数画像間の対応点を独立に移動する物体に分割する Costeira-Kanade 法や Gear の方法、およびそれらを拡張した「部分空間分離法」や「アフィン空間分離法」では、物体の個数の推定が最も困難である。これは数学的には誤差のある行列のランクの推定に帰着する。本論文ではこれに「モデル選択」によって経験的なしきい値なしに判定できることを示し、その規準として「幾何学的 AIC」、「幾何学的 MDL」、および「大津・市村の判別規準」を実験的に比較する。最後に、得られた分割の事後評価や解の選択の基準のための F 検定およびモデル選択を定式化する。

キーワード: 運動物体分離、幾何学的モデル選択、領域分割、ランク推定、幾何学的 AIC、幾何学的 MDL

Rank Estimation by Model Selection for Multibody Motion Segmentation

Kenichi Kanatani*, Noriyoshi Kurosawa* and Chikara Matsunaga†

*Department of Computer Science, Gunma University, Kiryu, Gunma 376-8515 Japan
†Broadcast Division, FOR-A Co. Ltd., 2-3-3 Ohsaku, Sakura, Chiba 285-0802 Japan

In the *subspace separation* method or the *affine space separation* method that generalize the Costeira-Kanade method and Gear's method for multibody motion segmentation, the most difficult task is to estimate the number of objects. Mathematically, it reduces to estimating the rank of a matrix in the presence of noise. We show that we can estimate the rank by *model selection* without introducing any thresholds that need to be adjusted empirically. We experimentally compare the *geometric AIC*, the *geometric MDL*, and the *Otsu-Ichimura criterion*. Finally, we formulate the F-test and the model selection for a posteriori evaluation of the separation and selection of the solution.

Key words: motion segmentation, geometric model selection, image segmentation, rank estimation, geometric AIC, geometric MDL

謝辞: 本研究の一部は文部科学省科学研究費基盤研究 C (2) (No. 11680377) によった。

*376-8515 桐生市天神町 1–5–1 群馬大学工学部情報工学科
Tel: (0277)30-1841, Fax: (0277)30-1801, E-mail: kanatani@cs.gunma-u.ac.jp, kurosawa@ail.cs.gunma-u.ac.jp
†285-0802 千葉県佐倉市大作 2–3–3 (株) 朋栄 放送システム開発部
Tel: (043)498-1239, Fax: (043)498-2223, E-mail: matsunaga@for-a.co.jp

1. 序論

画像中の物体を背景から切り出す「領域分割」はコンピュータビジョンの重要な問題である。Costeira と Kanade [1] は動画像から 3 次元復元する Tomasi-Kanade の「因子分解法」[18] をヒントにして、複数画像間の対応点を独立に移動する物体に分割する手法を提案し、種々の応用や改良が研究された [2, 4, 14, 15]。これらはアフィンカメラのもとではシーン中を剛体運動する点の画像上の運動履歴がある 4 次元部分空間に拘束されるという事実に基づく。

市村 [3] は Costeira-Kanade 法に大津の判別規準 [16] を適用した。黒澤・金谷 [12, 13] はモデル選択やロバスト推定などの手法を組み合わせて「部分空間分離法」および「アフィン空間分離法」と呼ぶ手法にまとめた。これが Costeira-Kanade 法や市村の方法と異なるのは、分離をデータから導かれるある行列の要素の零・非零に帰着させるのではなく、誤差の入る元々のデータ空間を解析している点である。

この解析のために文献 [12, 13] では移動物体の個数は既知としているが、物体の個数の推定は非常に難しい問題であり、文献 [1, 2, 15] でもこれが分離の手続きよりも困難であると指摘している。特に Gear [2] は誤差の確率モデルに基づく統計解析を行い、結論として各点がそれぞれ独立な物体と判定される傾向にあると述べている。これは当然である。なぜなら独立な運動はまとまった運動の一種であるから、データの説明を目標とする限り、ゆるい条件ほどデータによくあてはまるからである。

これを解決するには、まとまった運動が望ましいという何らかの規準が必要である。もう一つの問題は、従来は判定のしきい値を経験的に調節していることであり [1, 2, 15]、一般性がない。これを解決するのが「モデル選択」である。

数学的には物体の個数の推定は行列のランクの推定に帰着する。時系列データの解析で類似の問題が現れ、データ長に関して漸近解析した赤池の AIC や Rissanen の MDL が適用されているが [19]、本論文ではノイズレベルに関する摂動解析 [10, 11] による幾何学的 AIC と幾何学的 MDL を適用する。さらにこれらを大津・市村の判別規準とも比較する。

最後に、得られた分離の正しさの事後評価のための F 検定およびモデル選択を定式化する。これは、部分空間分離法の解とアフィン空間分離法の解が異なる場合の解の選択の基準にも利用できる。

2. 運動物体の分離

静止したカメラ座標系を世界座標系と同一視し、XY 面を画像面、Z 軸をカメラの光軸とみなす。ある剛体に属する特徴点 p_α の第 κ 画像中の画像座標を $(x_{\kappa\alpha}, y_{\kappa\alpha})$ とし、これを $\kappa = 1, ..., M$ に渡って縦に並べた $2M$ 次元ベクトル $\boldsymbol{p}_\alpha$ を次のように定義する。

$$\boldsymbol{p}_\alpha = \begin{pmatrix} x_{1\alpha} & y_{1\alpha} & x_{2\alpha} & y_{2\alpha} & \cdots & y_{M\alpha} \end{pmatrix}^\top \tag{1}$$

物体に任意に物体座標系を固定し、特徴点 p_α の物体座標系に関する座標を $(a_\alpha, b_\alpha, c_\alpha)$ とする。時刻 κ での物体座標系の原点の位置と各座標基底ベクトルをそれぞれ $\boldsymbol{t}_\kappa$、$\{\boldsymbol{i}_\kappa, \boldsymbol{j}_\kappa, \boldsymbol{k}_\kappa\}$ とすると、特徴点 p_α の時刻 κ における位置 $\boldsymbol{r}_{\kappa\alpha}$ は次のように書ける。

$$\boldsymbol{r}_{\kappa\alpha} = \boldsymbol{t}_\kappa + a_\alpha \boldsymbol{i}_\kappa + b_\alpha \boldsymbol{j}_\kappa + c_\alpha \boldsymbol{k}_\kappa \tag{2}$$

平行投影を仮定し、ベクトル $\boldsymbol{t}_\kappa$, $\boldsymbol{i}_\kappa$, $\boldsymbol{j}_\kappa$, $\boldsymbol{k}_\kappa$ の投影 (Z 座標を取り除いた 2 次元ベクトル) を $\kappa = 1, ..., M$ に渡って縦に並べた $2M$ 次元ベクトルをそれぞれ $\boldsymbol{m}_0$, $\boldsymbol{m}_1$, $\boldsymbol{m}_2$, $\boldsymbol{m}_3$ とすると、式 (1) の $2M$ 次元ベクトル $\boldsymbol{p}_\alpha$ が次のように表せる。

$$\boldsymbol{p}_\alpha = \boldsymbol{m}_0 + a_\alpha \boldsymbol{m}_1 + b_\alpha \boldsymbol{m}_2 + c_\alpha \boldsymbol{m}_3 \tag{3}$$

これから、各特徴点 p_α の履歴が $2M$ 次元空間 $\mathcal{R}^{2M}$ の 1 点となり、N 個の点 $\{\boldsymbol{p}_\alpha\}$ がベクトル $\{\boldsymbol{m}_0, \boldsymbol{m}_1, \boldsymbol{m}_2, \boldsymbol{m}_3\}$ の張る 4 次元部分空間 $\mathcal{L}$ に含まれることがわかる。特に $\boldsymbol{m}_0$ の係数はすべての α に共通に 1 であるから、$\{\boldsymbol{p}_\alpha\}$ はさらに $\mathcal{L}$ 内のある 3 次元アフィン空間 $\mathcal{A}$ に拘束されている。これは平行投影のみならず、弱透視変換や疑似透視変換を含む一般のアフィンカメラ [5] でも同様である [13]。

したがって、もし m 個の物体が独立に運動していれば、データ $\{\boldsymbol{p}_\alpha\}$ は $2M$ 次元空間 $\mathcal{R}^{2M}$ の $4m$ 次元部分空間および $4m - 1$ 次元アフィン空間に含まれる。すなわち物体の個数はデータ $\{\boldsymbol{p}_\alpha\}$ のランク (= それらの張る部分空間の次元) およびアフィンランク (= それらを通るアフィン空間の次元) の推定に帰着する。このことから m 個の物体を分離するには $2m$ 枚以上の画像が必要であることもわかる。

物体が画像面内で 2 次元剛体運動をする場合は式 (3) で $\boldsymbol{m}_3$ が恒等的に $\boldsymbol{0}$ となるから、データ $\{\boldsymbol{p}_\alpha\}$ のランクは $3m$、アフィンランクは $3m - 1$ となり、m 個の物体を分離するには $1.5m$ 枚以上の画像が必要である。

3. ランクの推定

誤差がないとき、N 本の n 次元ベクトル $\{\boldsymbol{p}_\alpha\}$ のランク r を計算する方法に 3 通りある。

1. $n \times n$ モーメント行列

$$\boldsymbol{M} = \sum_{\alpha=1}^{N} \boldsymbol{p}_\alpha \boldsymbol{p}_\alpha^\top \tag{4}$$

の固有値を $\lambda_1 \geq \cdots \geq \lambda_n$、対応する固有ベクトルの正規直交系を $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_n\}$ とする。r は $\boldsymbol{M}$ の正の固有値の個数、すなわち $\boldsymbol{M}$ のランクに等しい。

2. $N \times N$ 計量行列 $\boldsymbol{G} = (G_{\alpha\beta})$

$$G_{\alpha\beta} = (\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta) \tag{5}$$

の固有値を $\lambda_1 \geq \cdots \geq \lambda_N$、対応する固有ベクトルの正規直交系を $\{\boldsymbol{v}_1, ..., \boldsymbol{v}_N\}$ とする。r は $\boldsymbol{G}$ の正の固有値の個数、すなわち $\boldsymbol{G}$ のランクに等しい。

3. $n \times N$ 観測行列 $\boldsymbol{W}$ を

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{p}_1 & \cdots & \boldsymbol{p}_N \end{pmatrix} \tag{6}$$

とし、$n > N$ のときその特異値分解を

$$\boldsymbol{W} = \boldsymbol{U}_{n\times N}\mathrm{diag}(\sigma_1, \sigma_2, ..., \sigma_N)\boldsymbol{V}_{N\times N}^\top \tag{7}$$

とする ($\mathrm{diag}(\cdots)$ は $\cdots$ を対角要素とする対角行列)。$\boldsymbol{U}_{n\times N}$, $\boldsymbol{V}_{N\times N}$ はそれぞれ列が直交する $n \times N$, $N \times N$ 行列である。$N \geq n$ のときは $\boldsymbol{W}^\top$ を

$$\boldsymbol{W}^\top = \boldsymbol{V}_{N\times n}\mathrm{diag}(\sigma_1, \sigma_2, ..., \sigma_n)\boldsymbol{U}_{n\times n}^\top \tag{8}$$

と特異値分解する。$\boldsymbol{V}_{N\times n}$, $\boldsymbol{U}_{n\times n}$ はそれぞれ列が直交する $N \times n$, $n \times n$ 行列である。いずれの場合も r は正の特異値の個数、すなわち $\boldsymbol{W}$, $\boldsymbol{W}^\top$ のランクに等しい。

関係 $\boldsymbol{M} = \boldsymbol{W}\boldsymbol{W}^\top$, $\boldsymbol{G} = \boldsymbol{W}^\top\boldsymbol{W}$ から次のことがわかる。

- モーメント行列 $\boldsymbol{M}$ と計量行列 $\boldsymbol{G}$ のランクは等しい。それらの固有値 $\{\lambda_i\}$ は零の個数を除いて一致し、正の固有値を r 個もつ。
- 観測行列 $\boldsymbol{W}$ とその転置 $\boldsymbol{W}^\top$ のランクは等しい。それらの特異値 $\{\sigma_i\}$ は零の個数を除いて一致し、正の特異値を r 個もつ。
- 行列 $\boldsymbol{M}$, $\boldsymbol{G}$ の非零の固有値は対応する特異値の二乗である。$\lambda_i = \sigma_i^2$, $i = 1, ..., r$.
- モーメント行列 $\boldsymbol{M}$ の非零の固有値に対応する固有ベクトル $\{\boldsymbol{u}_i\}$, $i = 1, ..., r$ は、$n > N$ のときは式 (7) 中の行列 $\boldsymbol{U}_{n\times N}$ の最初の r 本の列に等しく、$N \geq n$ のときは式 (8) 中の行列 $\boldsymbol{U}_{n\times n}$ の最初の r 本の列に等しい。
- 計量行列 $\boldsymbol{G}$ の非零の固有値に対応する固有ベクトル $\{\boldsymbol{v}_i\}$, $i = 1, ..., r$ は、$n > N$ のときは式 (7) 中の行列 $\boldsymbol{V}_{N\times N}$ の最初の r 本の列に等しく、$N \geq n$ のときは式 (8) 中の行列 $\boldsymbol{V}_{N\times n}$ の最初の r 本の列に等しい。

このように種々のランクの計算法があるが、実際の計算では効率および安定性の観点から式 (7), (8) の特異値分解がよいとされている。ただし、これは計算の便宜であり、幾何学的構造と計算の便宜とを混同すべきでない[1]。数学的な構造はモーメント行列 $\boldsymbol{M}$ や計量行列 $\boldsymbol{G}$ に反映されており、特異値分解からは幾何学的構造を正しく理解するのが難しい。

4. アフィンランクの推定

誤差がないとき、n 次元空間中の N 個の点 $\{\boldsymbol{p}_\alpha\}$ のアフィンランク r を計算するには、各点の重心

$$\boldsymbol{p}_C = \frac{1}{N}\sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha \tag{9}$$

からの変位

$$\boldsymbol{p}'_\alpha = \boldsymbol{p}_\alpha - \boldsymbol{p}_C \tag{10}$$

を計算し、$\{\boldsymbol{p}'_\alpha\}$ のランクを計算すればよい。これは前節に述べた計算に帰着する。

5. ランクのモデル選択

誤差のあるときにベクトルの張る部分空間の次元を推定する素朴な考え方は、次元を仮定して部分空間を最適に当てはめて、その残差が最も小さいものを選ぶことである。しかし、高い次元の部分空間を当てはめるほど残差が小さくなり、十分高い次元の部分空間、特に全空間を当てはめたとき残差が 0 になる。

これは次元の高い部分空間ほど自由度、すなわち調節できるパラメータ数が多いためである。このため残差と自由度のバランスを考える必要がある。これを考慮する代表的な規準に「幾何学的 AIC」と「幾何学的 MDL」がある [7, 8]。

n 次元空間の r 次元部分空間の自由度は $r(n-r)$ であるから[2]、N 個のデータ点に対する幾何学的 **AIC** は次のようになる [7, 8]。

$$\text{G-AIC} = \hat{J}_r + 2\Big(rN + r(n-r)\Big)\epsilon^2 \tag{11}$$

ただし ϵ はノイズレベルであり、ここでは n 次元空間の各データ点の各座標値に誤差 0、標準偏差 ϵ の正規分布に従う誤差が独立に入るとする。また $\hat{J}_r$ はデータ $\{\boldsymbol{p}_\alpha\}$ に r 次元部分空間を各点からの距離二乗和が最小となるように当てはめた残差である。これは次のように表せる [12, 13]。

$$\hat{J}_r = \sum_{i=r+1}^{\nu}\lambda_i \tag{12}$$

---

[1] Tomasi-Kanade の "因子分解法"[18] の原理は、アフィンカメラのもとで剛体運動する点の画像上の運動履歴が 3 次元アフィン空間に拘束されることである。単一物体のとき、その重心を原点にとるとデータは 3 次元部分空間に拘束され、これに 3 次元部分空間を当てはめて 3 次元復元ができる [9]。しかし、命名の不適切さから特異値分解がその原理であると錯覚されている。

[2] r 個の点の座標を指定して定まるが、それらを r 次元部分空間内で任意に移動させてもよいから、実質的なパラメータ数は $nr - r^2$ である。

ただし $\{\lambda_i\}$ は式 (4) の行列 $\boldsymbol{M}$ および式 (5) の行列 $\boldsymbol{G}$ の正の固有値であり、$\nu = \min(n, N)$ である。一方、情報理論的符号長から導かれた幾何学的 **MDL** は次のようになる [10, 11]。

$$\text{G-MDL} = \hat{J}_r - \Big(rN + r(n-r)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{13}$$

ただし L はデータ空間の大きさを表す基準長である。これを厳密に定めるには符号長の高次の展開が必要となるが、近似的にはデータがほぼ 1 のオーダーになるように定めればよい。ϵ が十分小さい場合は L を無視することもできる。

$r = 1, 2, ...$ に対して式 (11) または式 (13) を評価して、これが最小になる値をランク r とする。このとき必要になるノイズレベル ϵ の推定のために、ランク r はある値 $r_{\max}$ 以下であることが既知であるとする。データ $\{\boldsymbol{p}_\alpha\}$ に $r_{\max}$ 次元部分空間を最適に当てはめたときの残差を $\hat{J}_{r_{\max}}$ とすれば、ノイズレベル ϵ が次のように推定できる [12, 13]。

$$\hat{\epsilon}^2 = \frac{\hat{J}_{r_{\max}}}{(n - r_{\max})(N - r_{\max})} \tag{14}$$

これは $\hat{J}_{r_{\max}}/\epsilon^2$ が自由度 $(n - r_{\max})(N - r_{\max})$ の χ^2 分布に従うことから得られる[3]。

幾何学的 AIC も幾何学的 MDL も恣意的なしきい値を必要としない点が特徴的であるが、同じくしきい値を用いない判定として、市村 [3] は大津の判別規準 [16] を Costeira-Kanade 法に適用した。これをランクの推定に転用することもできる。観測行列 $\boldsymbol{W}$ の特異値を $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_\nu$ とするとき、問題はこれを d 個の非零のグループと $\nu - d$ 個の零のグループとに分割することである。大津・市村の規準は次の量 (仮に OIC と記す) を最大にする r を採用することである (その他さまざまな同値な変形ができる [16])。

$$\text{OIC} = \frac{r(\nu - r)(\mu_1 - \mu_2)^2}{\sum_{i=1}^{r}(\sigma_i - \mu_1)^2 + \sum_{i=r+1}^{\nu}(\sigma_i - \mu_2)^2} \tag{15}$$

ただし次のように置いた。

$$\mu_1 = \frac{1}{r}\sum_{i=1}^{r}\sigma_i, \qquad \mu_2 = \frac{1}{\nu - r}\sum_{i=r+1}^{\nu}\sigma_i \tag{16}$$

6. アフィンランクのモデル選択

n 次元空間の r 次元アフィン空間の自由度は $(r+1)(n-r)$ であるから[4]、N 個のデータ点に対する幾何学的 AIC は次のようになる [7, 8]。

$$\text{G-AIC} = \hat{J}'_r + 2\Big(rN + (r+1)(n-r)\Big)\epsilon^2 \tag{17}$$

幾何学的 MDL は次のようになる。

$$\text{G-MDL} = \hat{J}'_r - \Big(rN + (r+1)(n-r)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{18}$$

そして $r = 1, 2, ...$ に対して式 (17) または式 (18) を評価して、これが最小になる値をアフィンランク r とすればよい。ノイズレベル ϵ の推定のために、アフィンランク r はある値 $r_{\max}$ 以下であることが既知であるとする。データ $\{\boldsymbol{p}_\alpha\}$ に $r_{\max}$ 次元アフィン空間を当てはめたときの残差を $\hat{J}'_{r_{\max}}$ とすれば、ノイズレベル ϵ が次のように推定できる [12, 13]。

$$\hat{\epsilon}^2 = \frac{\hat{J}'_{r_{\max}}}{(n - r_{\max})(N - r_{\max} - 1)} \tag{19}$$

これは $\hat{J}'_{r_{\max}}/\epsilon^2$ が自由度 $(n - r_{\max})(N - r_{\max} - 1)$ の χ^2 分布に従うことから得られる[5]。

OIC は式 (6) において $\boldsymbol{p}_\alpha$ を式 (10) の $\boldsymbol{p}'_\alpha$ に置き換えた観測行列 $\boldsymbol{W}'$ の特異値 $\{\sigma'_i\}$ に対して式 (15) を適用すればよい。

7. 数値実験

区間 $[-1, 1]$ 上の一様乱数を独立に 200 個発生させ、これを要素とする 10×20 行列 $\boldsymbol{B}$ を次のように特異値分解した。

$$\boldsymbol{B} = \boldsymbol{V}\mathrm{diag}(\sigma_1, ..., \sigma_{10})\boldsymbol{U}^\top \tag{20}$$

これから次の行列 $\boldsymbol{A}$ を定義した。

$$\boldsymbol{A} = \boldsymbol{V}\mathrm{diag}(\sigma_1, ..., \sigma_5, \gamma\sigma_5, 0, ..., 0)\boldsymbol{U}^\top \tag{21}$$

計算すると σ_1, ..., σ_5 はそれぞれ 3.81, 3.58, 3.09, 2.98, 2.75 であった。行列 $\boldsymbol{A}$ のランクは $\gamma = 0$ のとき 5、$0 < \gamma < 1$ のとき 6 である。

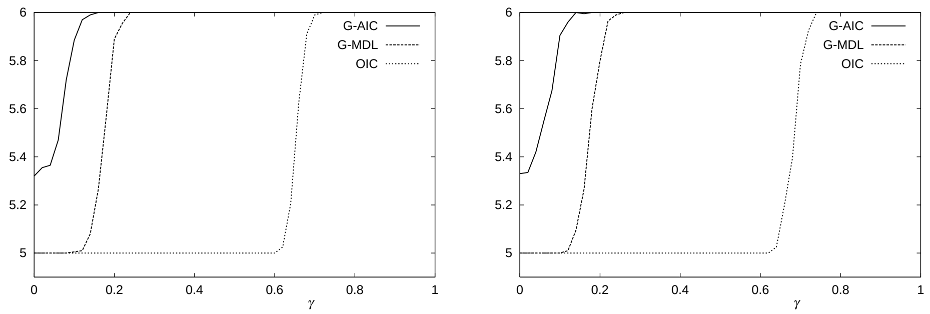
図 1(a) は横軸に γ をとり、$\boldsymbol{A}$ の各要素に独立に期待値 0、標準偏差 0.05 の誤差を独立に加えて、$r_{\max} = 6$ として 200 回試行し、推定したランクの平均値をプロットしたものである。基準長 L は 1 とした。

幾何学的 AIC は $\gamma = 0$ (真のランクが 5) のときでもデータに誤差があると一定の割合でランク 6 と判定し、$\gamma > 0$ (真のランクが 6) のときは γ が非常に小さくてもランク 6 と判定することが多い。これに対して幾何学的 MDL は $\gamma = 0$ (真のランクが 5) のときは正しくランク 5 と判定するが、$\gamma > 0$ (真のランクが 6) のときでも γ がある程度の値までランク 5 と判定することが多い。

[3] 誤差の自由度 (= モデルの余次元) は $n - r_{\max}$ であり、モデルの自由度は $r_{\max}(n - r_{\max})$ であるから、$\hat{J}_{r_{\max}}/\epsilon^2$ の自由度は $(n - r_{\max})N - r_{\max}(n - r_{\max})$ となる [7]。

[4] $r+1$ 個の点の座標を指定して定まるが、それらを r 次元アフィン空間内で任意に移動させてもよいから、実質的なパラメータ数は $n(r+1) - r(r+1)$ である。

[5] 誤差の自由度 (= モデルの余次元) は $n - r_{\max}$ であり、モデルの自由度は $(r_{\max}+1)(n - r_{\max})$ であるから、$\hat{J}'_{r_{\max}}/\epsilon^2$ の自由度は $(n - r_{\max})N - (r_{\max}+1)(n - r_{\max})$ となる [7]。



(a) (a)

図 1: 推定したランク (a) とアフィンランク (b) の平均値。実線: 幾何学的 AIC、破線: 幾何学的 MDL、点線: OIC。

一方、 OIC は本来は 0 でない小さい特異値を零のグループに分類する傾向が強いので、相当に大きい γ に対してもほとんどの場合にランク 5 と判定する。

式 (21) の行列 $\boldsymbol{A}$ のアフィンランクは $\gamma = 0$ のとき 5、 $0 < \gamma < 1$ のとき 6 である。図 1(b) は図 1(a) と同様に誤差のあるときのアフィランクの平均値である。これについても図 1(a) と同じ傾向が見られる。

このように、幾何学的 AIC は誤差に比較的忠実にランクをやや大きめに推定し、幾何学的 MLD は誤差を無視してもランクをやや少なめに推定する。このような幾何学的 AIC と対照的な挙動は他の応用でも見られる [11]。

ただし、 $r_{\max}$ を大きくとると式 (14) の分子の残差 $\hat{J}_{r_{\max}}$ が非常に小さくなり、数値計算において有効数字が減少するため $\hat{\epsilon}$ が小さめに推定される (極端な場合は 0 となる)。したがって $r_{\max}$ は可能な限り小さくとる必要がある。

8. 分離手法

文献 [12, 13] では「部分空間分離法」および「アフィン空間分離法」を 2 物体に適用しているが、これを任意個数の物体に拡張すれば、 N 個の n 次元ベクトル $\{\boldsymbol{p}_\alpha\}$ を m 個の d 次元部分空間あるいは $d-1$ 次元アフィン空間に分類することができる (3 次元運動の場合は $d = 4$、 2 次元運動の場合は $d = 3$)。

物体の個数 m の推定はランクまたはアフィンランクの推定に帰着するが、ランクは d の倍数、アフィンランクはそれから 1 を引いたものであるので、そのような場合を互いに比較すればよい。ただしノイズレベルの推定のために、個数の最大値 $m_{\max}$ を指定する。指定がなければデフォルト値 $m_{\max} = \lceil \min(n, N)/d \rceil - 1$ を用いる ($\lceil \cdot \rceil$ は天井関数)。

物体の個数 m が推定されれば文献 [12, 13] に従って、次元補正を行いながら作用行列と幾何学的 AIC を併用した類似度の大きい順に併合し、 m 個のクラスにまとまった時点で終了する。ただし、併合の過程で d 個以下の要素をもつものあれば、それらを優先して併合する。

このようにして初期分類が得られた後に、ロバスト推定によって再分類を行う [12, 13]。まず各クラスからそれぞれノルムの大きい半分を選ぶ (ただし d 個以上)。それらに d 次元部分空間 (または $d-1$ 次元アフィン空間) を最適に当てはめる。次に各クラスについて、残りのクラスに当てはめた部分空間 (またはアフィン空間) の内の最も近いものまでの距離が大きい半分の点を改めて選ぶ (ただし d 個以上)。そして、それらに d 次元部分空間 (または $d-1$ 次元アフィン空間) を最適に当てはめ、すべての点をそれらの内の最も近いものに帰属させる。最後に各クラスに最小メジアン法 [17] によって再度それらに改めて d 次元部分空間 (または $d-1$ 次元アフィン空間) を最適に当てはめ、すべての点を再度最も近いものに帰属させる。

9. シミュレーション実験

図 2は 512×512 画素を想定した 2 次元平面上を独立に運動する 20 個の背景点とそれぞれ 9 個の点からなる 2 個の物体の 10 枚の動画像から 1 枚おきに抜き出したものである。見やすくするために運動物体をワイヤーフレーム表示している。各画像上の特徴点の各々の座標に独立に期待値 0、標準偏差 ϵ の正規乱数誤差を加えたものをデータとし、物体数を推定した。

図 3(a) は部分空間分離法において推定した物体数の平均値を示す。横軸は ϵ であり、 $m_{\max} = 4$ として 500 回独立に試行した。実線が幾何学的 AIC、破線が幾何学的 MDL、点線が OIC である。基準長 L は 600 とした。幾何学的 AIC は個数を平均的にやや多めに推定しているが、安定している。一方、幾何学的 MDL は誤差が非常に小さいときは正しい個数を推定しているが、誤差が増えると実際より少なく推定している。それに対して OIC は常に物体数を 1 と推定している。図 3(b) はアフィン空間分離法の場合であり、幾何学的 AIC と OIC は同様な傾向があるが、幾何学的 MDL は正しい結果となっている。

なお図 3(a), (b) の左端の不規則な挙動は、誤差が小さすぎて、残差の計算やノイズレベルの推定に有効数字が不足するためである。

10. 部分空間分離法の仮説検定

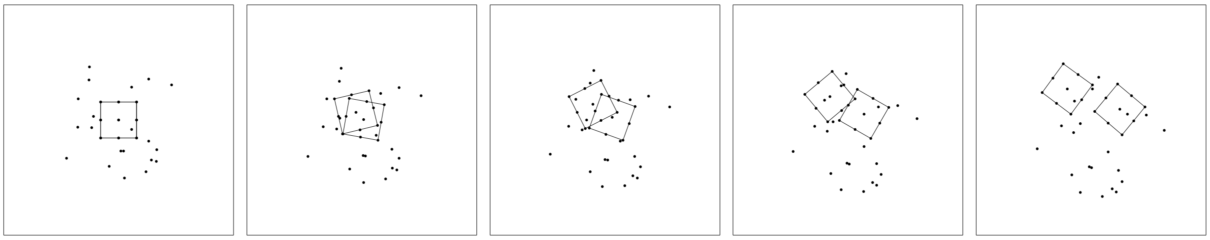
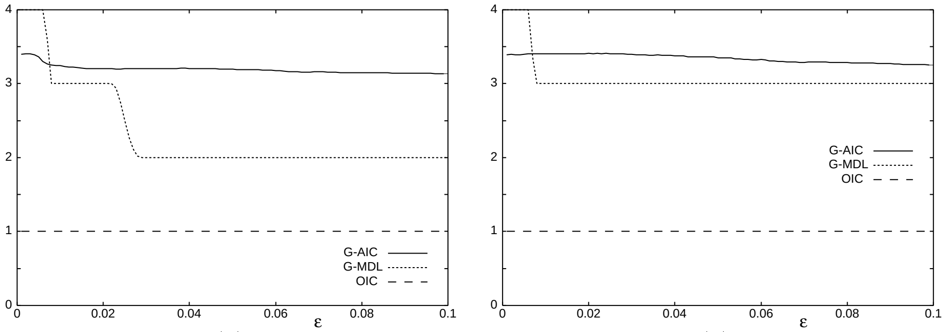


図 2: 2 次元平面上を運動する点。



(a) (a)

図 3: 推定した物体数の平均値。(a) 部分空間分離法。(b) アフィン空間分離法。

実画像例ではシミュレーションと異なり、真の解は (人間が目で見て指定しなければ) 不明であるから、シミュレーションで高性能であることが保証されても、個々の結果が正しいという保証はない。そこで結果の信頼性の事後評価を考える。

N 個のデータ点が m 個のグループに N_i 個ずつ分離されたとする。各グループに d 次元部分空間 $\mathcal{L}_i$ を当てはめた残差を $\hat{J}_i$ とすると、分離が正しければ $\hat{J}_i/\epsilon^2$ は χ^2 分布する。部分空間 $\mathcal{L}_i$ の余次元は $n-d$ であり、自由度が $d(n-d)$ であるから自由度は次のようになる [7]。

$$\phi_i = (n-d)N_i - d(n-d) = (n-d)(N_i - d) \tag{22}$$

N 個のデータ点全部に md 次元部分空間 $\mathcal{L}_t$ を当てはめた残差を $\hat{J}_t$ とすると、分離の正しさによらず $\hat{J}_t/\epsilon^2$ も χ^2 分布し、その自由度は次のようになる。

$$\begin{aligned}\phi_t &= \sum_{i=1}^{m}(n-md)N_i - md(n-md) \\ &= (n-md)(N-md)\end{aligned} \tag{23}$$

$\hat{J}_t$ は各点の部分空間 $\mathcal{L}_t$ までの距離の二乗和であり、$\sum_{i=1}^{m}\hat{J}_i$ は各点のそれが属す部分空間 $\mathcal{L}_i$ までの距離の二乗和である。これは各点から $\mathcal{L}_{\rm total}$ に下ろした垂線の長さと、その垂線の足から $\mathcal{L}_i$ に下ろした垂線の長さの二乗和を合計したものである (図 4)。ゆえに、部分空間 $\mathcal{L}_{\rm total}$ 内での各点のそれが属す部分空間 $\mathcal{L}_i$ までの距離の二乗和は $\sum_{i=1}^{m}\hat{J}_i - \hat{J}_t$ である。$(\sum_{i=1}^{m}\hat{J}_i - \hat{J}_t)/\epsilon^2$ も χ^2 分布するが、各 $\hat{J}_i$ に寄与する誤差は互いに独立であり、かつ $\hat{J}_t$ に寄与する誤差は $\hat{J}_1$, ..., $\hat{J}_m$ に寄与する誤差の一部であるから、コクランの定理によりその自由度は次のようになる [6, 7]。

$$\sum_{i=1}^{m}\phi_i - \phi_t = (m-1)d(N-md) \tag{24}$$

さらに $(\sum_{i=1}^{m}\hat{J}_i - \hat{J}_t)/\epsilon^2$ に寄与する誤差成分と $\hat{J}_t$ に寄与する誤差成分とは直交するから互いに独立である。ゆえに分離が正しければ

$$F_s = \frac{(\sum_{i=1}^{m}\hat{J}_i - \hat{J}_t)/(m-1)d(N-md)}{\hat{J}_t/(n-md)(N-md)} \tag{25}$$

は自由度 $(m-1)d(N-md)$, $(n-md)(N-md)$ の F 分布に従う。したがって、

$$F_s > F_\alpha^{(s)} \tag{26}$$

なら分離は有意水準 α で棄却される。ただし $F_\alpha^{(s)}$ は自由度 $(m-1)d(N-md)$, $(n-md)(N-md)$ の F 分布の上側 α% 点である。

11. アフィン空間分離法の仮説検定

アフィン空間分離法でも同様である。各グループに $d-1$ 次元アフィン空間 $\mathcal{A}_i$ を当てはめた残差を $\hat{J}'_i$ とすると、分離が正しければ $\hat{J}'_i/\epsilon^2$ は χ^2 分布する。アフィン空間 $\mathcal{A}_i$ の余次元は $n-d+1$ であり、自由度が $d(n-d+1)$ であるから自由度は次のようになる。

$$\phi'_i = (n-d+1)N_i - d(n-d+1) = (n-d+1)(N_i-d) \tag{27}$$

N 個のデータ点全部に $md-1$ 次元アフィン空間 $\mathcal{A}_t$ を当てはめた残差を $\hat{J}'_t$ とすると、分離の正しさによ

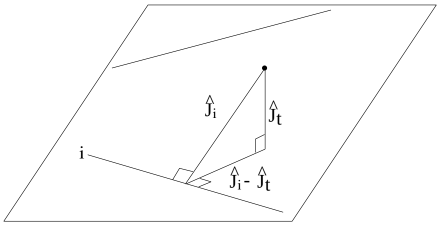


図 4: 部分空間あてはめの残差。

らず $\hat{J}'_t/\epsilon^2$ も χ^2 分布し、その自由度は次のようになる。

$$\phi'_t = \sum_{i=1}^{m}(n-md+1)N_i - md(n-md+1) = (n-md+1)(N-md) \tag{28}$$

そして $(\sum_{i=1}^m \hat{J}'_i - \hat{J}'_t)/\epsilon^2$ も χ^2 分布し、自由度は次のようになる。

$$\sum_{i=1}^{m}\phi'_i - \phi'_t = (m-1)d(N-md) \tag{29}$$

ゆえに分離が正しければ

$$F_a = \frac{(\sum_{i=1}^m \hat{J}'_i - \hat{J}'_t)/(m-1)d(N-md)}{\hat{J}'_t/(n-md+1)(N-md)} \tag{30}$$

は自由度 $(m-1)d(N-md)$, $(n-md+1)(N-md)$ の F 分布に従う。したがって、

$$F_a > F_\alpha^{(a)} \tag{31}$$

なら分離は有意水準 α で棄却される。ただし $F_\alpha^{(a)}$ は自由度 $(m-1)d(N-md)$, $(n-md+1)(N-md)$ の F 分布の上側 α% 点である。

12. 部分空間分離法のモデル選択

F 検定は有意水準に左右される。幾何学的 AIC を用いて有意水準に依存しないモデル選択を行うと次のようになる。各グループおよび全体の幾何学的 AIC および幾何学的 MDL は各々次のようになる。

$$\text{G-AIC}_i = \hat{J}_i + 2\Big(dN_i + d(n-d)\Big)\epsilon^2$$
$$\text{G-MDL}_i = \hat{J}_i - \Big(dN_i + d(n-d)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{32}$$

$$\text{G-AIC}_t = \hat{J}_t + 2\Big(mdN_i + md(n-md)\Big)\epsilon^2$$
$$\text{G-MDL}_t = \hat{J}_t - \Big(mdN_i + md(n-md)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{33}$$

二乗ノイズレベル ϵ^2 は全体の当てはめは分離の正しさによらないから、次のように推定される。

$$\hat{\epsilon}^2 = \frac{\hat{J}_t}{(n-md)(N-md)} \tag{34}$$

これから分離が適切でないとされる判定条件 $\text{G-AIC}_t < \sum_{i=1}^m \text{G-AIC}_i$, $\text{G-MDL}_t < \sum_{i=1}^m \text{G-MDL}_i$ はそれぞれ次のようになる。

$$F_s > 2, \qquad F_s > -\log\Big(\frac{\epsilon}{L}\Big)^2 \tag{35}$$

ここに F_s は式 (25) の F 統計量である。通常 $-\log(\epsilon/L)^2$ は 2 より大きいので、幾何学的 AIC は (常に成立する) 一般モデルを選び易く、幾何学的 MDL は (それを特定化した) 特殊モデルを選びやすい [11]。

13. アフィン空間分離法のモデル選択

アフィン空間分離法では各グループおよび全体の幾何学的 AIC、幾何学的 MDL は各々次のようになる。

$$\text{G-AIC}_i = \hat{J}'_i + 2\Big((d-1)N_i + d(n-d+1)\Big)\epsilon^2$$
$$\text{G-MDL}_i = \hat{J}'_i - \Big((d-1)N_i + d(n-d+1)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{36}$$

$$\text{G-AIC}_t = \hat{J}'_t + 2\Big((md-1)N_i + md(n-md+1)\Big)\epsilon^2$$
$$\text{G-MDL}_t = \hat{J}'_t - \Big((md-1)N_i + md(n-md+1)\Big)\epsilon^2 \log\Big(\frac{\epsilon}{L}\Big)^2 \tag{37}$$

二乗ノイズレベル ϵ^2 は全体の当てはめは分離の正しさによらないから、次のように推定される。

$$\hat{\epsilon}^2 = \frac{\hat{J}'_t}{(n-md+1)(N-md)} \tag{38}$$

これから分離が適切でないとされる判定条件 $\text{G-AIC}_t < \sum_{i=1}^m \text{G-AIC}_i$, $\text{G-MDL}_t < \sum_{i=1}^m \text{G-MDL}_i$ はそれぞれ次のようになる。

$$F_a > 2, \qquad F_a > -\log\Big(\frac{\epsilon}{L}\Big)^2 \tag{39}$$

ここに F_a は式 (30) の F 統計量である。

14. 実画像実験

図 5の上段は、静止した 3 物体と運動する 1 物体を移動カメラから連続して撮影し、5 枚をサンプルしたものである (512 × 768 画素)。下段に示した特徴点を手で選んで $m_{\max} = 2$ として物体数を推定すると[6]、部分空間 / アフィン空間分離法による幾何学的 AIC/MDL でどの場合も $m = 2$ となったが、OIC では常に $m = 1$ であった。そして $m = 2$ として部分空間 / アフィン空間分離法で分離すると、運動する 1 物体が正しく取り出された。

[6] $m_{\max} = 3$ とするためには 7 枚以上の画像が必要である

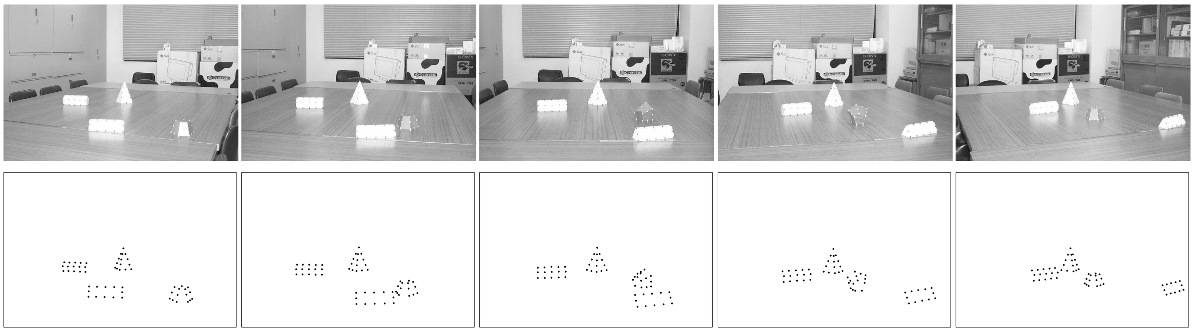


図 5: 背景と独立に移動する物体の実画像 (上) とその特徴点 (下)

これら 5 枚の画像の中間の画像を追加した 10 枚で $m_{\max} = 3, 4$ として実験すると、幾何学的 AIC/MDL ではどの場合も $m = 3$ となり、OIC ではやはり $m = 1$ となった。$m = 3$ として分離すると、運動する 1 物体が正しく取り出され、残りの 3 物体のうちの一つが独立に運動するとみなされた。この分離も誤りとはいえない。

図 5の分離の結果の信頼性を評価すると、式 (25), (30) の F 統計量 F_s, F_a はそれぞれ 0.893, 1.483 となった。それぞれの上側 5% 点は 1.346, 1.293 であるから、有意水準 5% で部分空間分離法の分離は棄却されないが、アフィン空間分離法の分離は棄却される。しかし幾何学的 AIC では共に選択される。一方、基準長 L を 600 とすると $-\log(\epsilon/L)^2$ は部分空間分離法およびアフィン空間分離法でそれぞれ 13.5, 13.1 となるので幾何学的 MDL でも選択され、先に指摘したように、幾何学的 AIC よりも緩い評価となる。

前報 [13] に示したように、透視効果の影響の度合いによって部分空間分離法とアフィン空間分離法の解の信頼性が異なる。図 5のような透視効果の強い画像ではアフィン空間分離法の信頼性が低下し、判定のしきい値の差にそれが現れている。これを利用すると、部分空間分離法とアフィン空間分離法を両方適用して異なる解が得られたとき、より信頼性の高い解を選択することができる。

15. まとめ

本論文では著者らが先に提案した部分空間 / アフィン空間分離法における物体の個数を推定する方法を示した。これは誤差のある行列のランクの推定に帰着する。これをモデル選択によって経験的に設定すべきしきい値なしに判定できることを示し、幾何学的 AIC、幾何学的 MDL、および OIC (大津・市村の判別規準) を比較した。幾何学的 AIC は物体数をやや多めに推定するが、それによる分離は誤りとはいえない。幾何学的 MDL も似たような結果を与える。しかし OIC は物体数を過小評価するので、この目的には適していない。

最後に、得られた分割の事後評価や解の選択の基準のための F 検定およびモデル選択を定式化した。

参考文献

[1] J. P. Costeira and T. Kanade, A multibody factorization method for independently moving objects, *Int. J. Comput. Vision*, **29**-3 (1998), 159–179.

[2] C. W. Gear, Multibody grouping from motion images, *Int. J. Comput. Vision*, **29**-2 (1998), 133–150.

[3] 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, 情報処理学会コンピュータビジョンとイメージメディア研究会, 2000-CVIM-120-3 (2000-1), 17–24.

[4] 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 (1998), 2757–2766.

[5] 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, **J76-D-II**-8 (1993), 1497–1505.

[6] 金谷健一，「空間データの数理 — 3次元コンピューティングに向けて —」, 朝倉書店, 1995.

[7] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, 1996.

[8] 金谷健一, 情報量基準による幾何学的モデル選択, 情報処理学会論文誌, **37**-6 (1996), 1073 – 1080.

[9] 金谷健一，因子分解を用いない因子分解法：平行投影から透視変換へ，電子情報通信学会パタン認識とメディア理解研究会, PRMU98-26 (1998-6), 1 – 8.

[10] 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，情報処理学会コンピュータビジョンとイメージメディア研究会, 2000-CVIM-122-1 (2000), 1–8.

[11] 金谷健一，松永力，幾何学的ＭＤＬとそのメディア応用，情報処理学会コンピュータビジョンとイメージメディア研究会, 2000-CVIM-122-2 (2000), 9–16.

[12] 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，情報処理学会コンピュータビジョンとイメージメディア研究会, 2000-CVIM-124-4 (2000-11), 25–32.

[13] 黒澤典義，金谷健一，アフィン空間分離法とモデル選択による運動物体の分離，情報処理学会コンピュータビジョンとイメージメディア研究会, 2001-CVIM-125-3 (2001-1), 25–32.

[14] 牧淳人，渡邉睦，C. Wiles, Geotensity 拘束による３次元形状獲得, 電子情報通信学会論文誌 D-II, **J83-D-II**-8 (2000), 1714–1752.

[15] 長崎健，川嶋稔夫，青木由直，因子分解法に基づく運動画像列解析による多関節物体の構造推定, 電子情報通信学会論文誌 D-II, **J81-D-II**-3 (1998), 483–492.

[16] 大津展之，判別および最小２乗規準に基づく自動しきい値選定法, 電子通信学会論文誌 D, **J63-D**-4 (1980), 349–356.

[17] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 1987.

[18] C. Tomasi and T. Kanade, Shape and motion from image streams under orthography—A factorization method, *Int. J. Comput. Vision*, **9**-2 (1992), 137–154.

[19] M. Wax and T. Kailath, Detection of signals by information theoretic criteria, *IEEE Trans. Acoust. Speech, Signal Process.*, **33**-2 (1985), 387–392.