アフィン空間分離法による運動物体の分離*

黒澤 典義† 金谷 健一†
群馬大学工学部情報工学科

前報では、アフィンカメラのもとでは剛体運動が履歴空間の部分空間に対応することを利用して、複数画像間の対応点を独立に移動する物体に分割する Costeira-Kanade 法に次元補正法、幾何学的 AIC によるモデル選択、最小メジアン法等を適用して性能を向上させた。本報では、剛体運動はさらにその部分空間内のアフィン空間に対応することを利用してアルゴリズムを強化する。そしてシミュレーション実験を行い、性能がさらに向上することを示すとともに、透視効果の影響を評価する。

キーワード: 運動物体分離、アフィン空間分離法、幾何学的モデル選択、領域分割、幾何学的 AIC、ロバスト推定

Motion Segmentation by Affine Space Separation

Noriyoshi Kurosawa† and Kenichi Kanatani†
Department of Computer Science, Gunma University, Kiryu, Gunma 376-8515 Japan

In our previous report, we improved the Costeira-Kanade method for motion segmentation by incorporating such techniques as dimension correction, model selection by the geometric AIC, and least-median fitting. In this report, we strengthen the algorithm by exploiting the fact that a rigid motion in the scene corresponds to an affine space with a smaller dimension. Doing numerical simulation, we demonstrate that the accuracy improves further. We also study the effect of perspective distortion.

Key words: motion segmentation, affine space separation, geometric model selection, image segmentation, geometric AIC, robust estimation

1. 序論

前報 [8] では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 [1] を Tomasi-Kanade [14] の「因子分解法」から切り離し、線形代数の定理として「部分空間分離法」を定式化した。そして、種々の技法を導入してロバストなアルゴリズムを構成した。

部分空間分離法は、アフィンカメラのもとではシーン中を剛体運動する各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く [8]。しかし実際にはさらに強く、その 4 次元部分空間内の 3 次元アフィン空間 (平面運動の場合は 2 次元アフィン空間) に拘束されている。この事実はよく知られていたが、これまではこの拘束は活用されていなかった。

これは、Costeira-Kanade 法が「作用行列」と呼ぶ行列の要素の、誤差がないときの零・非零に基くものであり、それを支える線形代数の範囲内ではアフィン拘束条件を活用する手段が存在しないからである。この結果、Costeira-Kanade [1] や Gear [2] の定式化やその各種の拡張や変形 [3, 4, 10] にもアフィン拘束条件は利用されなかった。

部分空間分離法の出発点は Costeira-Kanade 法と等価であるから、この部分にはアフィン拘束条件は利用できない。しかし、前報ではそれをより効果的にするために次元補正法、幾何学的 AIC によるモデル選択等の手法、最小メジアン法によるロバスト当てはめなどの手法を導入した。これらは作用行列ではなく、"元々のデータ空間" の解析に基いているから、この部分にアフィン拘束条件が適用できる。本報ではこれによってアルゴリズムの性能が一層向上することを実験的に検証するとともに、透視効果の影響を評価する。前報と同様に本手法も、経験的に調節すべきパラメータを何も含んでいない。

部分空間法は "部分空間に拘束された点を分離する" という弱い条件に基く手法であるため、運動物体の分離だけでなく、複数光源の分離 [9] などの数学的に同じ構造を持つ問題にも適用できる。このため非常に一般性があるが、アフィン拘束は運動物体分離に限って成立する性質であるから、本報の手法はこれに特化した手法である。

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

*謝辞: 本研究の一部は文部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)30-1844, Fax: (0277)30-1801
E-mail: kurosawa@ail.cs.gunma-u.ac.jp, kanatani@cs.gunma-u.ac.jp

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。

$$\boldsymbol{p}_\alpha = \left(x_{1\alpha}\ y_{1\alpha}\ x_{2\alpha}\ y_{2\alpha}\ \cdots\ y_{M\alpha} \right)^\top \tag{1}$$

物体に任意に物体座標系を固定し、α 番目の特徴点 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、すなわち 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。しかし $\boldsymbol{m}_0$ 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 (付録 A)。

以上より 3 次元運動をする複数物体を分離するには $2M$ 次元空間 $\mathcal{R}^{2M}$ の点集合を互いに異なる **3** 次元アフィン空間に分割すればよい。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 **0** となるから、$2M$ 次元空間 $\mathcal{R}^{2M}$ の点集合を互いに異なる **2** 次元アフィン空間に分割すればよい。

3. 部分空間分離法

部分空間分離法をまとめると次のようになる。n 次元空間 $\mathcal{R}^n$ の N 点 $\boldsymbol{p}_\alpha$, $\alpha = 1, ..., N$ が r 次元部分空間を張るとする。$N \times N$ 計量行列 $\boldsymbol{G} = (G_{\alpha\beta})$ を次のように定義する ($(\boldsymbol{a}, \boldsymbol{b})$ はベクトル $\boldsymbol{a}, \boldsymbol{b}$ の内積)。

$$G_{\alpha\beta} = (\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta) \tag{4}$$

この固有値を $\lambda_1 \geq \cdots \geq \lambda_N$、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。そして $N \times N$ 作用行列 $\boldsymbol{Q} = (Q_{\alpha\beta})$ を次のように定義する。

$$\boldsymbol{Q} = \sum_{i=1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^\top \tag{5}$$

添え字集合 $\mathcal{I} = \{1, ..., N\}$ を次のように m 個の部分集合に分割する。

$$\mathcal{I}_1 \cup \cdots \cup \mathcal{I}_m = \mathcal{I}, \quad \mathcal{I}_1 \cap \cdots \cap \mathcal{I}_m = \emptyset \tag{6}$$

そして第 i 集合 $\{\boldsymbol{p}_\alpha\}$, $\alpha \in \mathcal{I}_i$ の張る部分空間を $\mathcal{L}_i$ とする。これら m 個の部分空間 $\mathcal{L}_i$, $i = 1, ..., m$ が互いに線形独立であれば次の定理が成り立つ [8]。

【定 理 1】 点 $\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta$ が異なる部分空間に属せば $\boldsymbol{Q}$ の $(\alpha\beta)$ 要素は 0 である。

$$Q_{\alpha\beta} = 0, \ \ \alpha \in \mathcal{I}_i, \ \ \beta \in \mathcal{I}_j, \ \ i \neq j \tag{7}$$

4. 分離手法

前報 [8] と同様に、簡単のため 2 個の物体が独立に運動する場合を考えるが、3 個以上の場合も容易に拡張できる。

定理 1 より $|Q_{\alpha\beta}|$ が大きい点 $\boldsymbol{p}_\alpha$, $\boldsymbol{p}_\beta$ を同一部分空間に併合し、行列 $\boldsymbol{Q}$ の第 α, β 行および第 α, β 列が連続するように行と列を入れ換えていけば $\boldsymbol{Q}$ が最終的に近似的なブロック対角行列になるというのが Costeira-Kanade 法 [1] の考え方である。形式的に書けば、第 i 部分空間 $\mathcal{L}_i$ と第 j 部分空間 $\mathcal{L}_j$ の類似度を

$$s_{ij} = \begin{cases} 0 \ \ \mathcal{L}_i = \emptyset \text{ または } \mathcal{L}_j = \emptyset \\ \max_{\mathbf{p}_\alpha \in \mathcal{L}_i, \mathbf{p}_\beta \in \mathcal{L}_j} |Q_{\alpha\beta}| \ \ \text{その他} \end{cases} \tag{8}$$

と定義し、これが大きい二つの部分空間 (ブロック) を一つに併合する操作を続けることになる。

しかしデータに誤差があると一つの特徴点の誤差が $\boldsymbol{Q}$ の全部の要素に影響を与え、各 $Q_{\alpha\beta}$ はデータ誤差の複雑な非線形関数となる。また本来零でない $Q_{\alpha\beta}$ の値については何の情報も存在しない。このため誤差解析が困難で、最適な規準を得ることが難しい [2]。市村 [3] はこれを避けるために $\boldsymbol{Q}$ の各行に大津の判別基準 [11] を用いたが、誤差の複雑な相関のために必ずしもよい結果を与えない [8]。

前報 [8] では誤差に対処する統計的手法を作用行列 $\boldsymbol{Q}$ ではなく元々のデータ空間を考え、次元補正法、幾何学的 AIC によるモデル選択等の手法、最小メジアン法によるロバスト当てはめなどの手法を導入した。本報ではこれらをアフィン拘束条件に変形する。

5. 次元補正法

以下、部分空間の次元を d とする。同一の運動をする点はその中の $d-1$ 次元アフィン空間に拘束さ

れている (各運動成分は d 個以上の点を含むとする)。2 次元運動の場合は $d = 3$ であり、3 次元運動の分離の場合は $d = 4$ である。

前報 [8] の方法をアフィン拘束に適用するには、d 個以上を併合した場合にはそれらに最適に $d-1$ 次元アフィン空間を当てはめ (付録 B, C)、各々の要素をそのアフィン空間に射影した点に置きかえればよい。ただし、併合した要素を次のステップでさらに別の要素と併合する場合は、改めて原データをその併合したアフィン空間に射影する。

この計算のため、$d-1$ 次元アフィン空間に併合する点集合を $\{\boldsymbol{p}_\alpha\}$, $\alpha = 1, ..., N$ とし、重心

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

を計算し、その周りの $n \times n$ モーメント行列

$$\boldsymbol{M}' = \sum_{\alpha=1}^{N} (\boldsymbol{p}_\alpha - \boldsymbol{p}_C)(\boldsymbol{p}_\alpha - \boldsymbol{p}_C)^\top \tag{10}$$

の大きい順に並べた固有値 $\lambda'_1 \geq \cdots \geq \lambda'_\nu$ と対応する単位固有ベクトル $\{\boldsymbol{u}'_1, ..., \boldsymbol{u}'_\nu\}$ を計算する ($\nu = \max(n, N)$)。そして次の射影行列を計算する。

$$\boldsymbol{P}_{d-1} = \sum_{i=1}^{d-1} \boldsymbol{u}'_i \boldsymbol{u}'^\top_i \tag{11}$$

併合後の位置 $\{\boldsymbol{p}'_\alpha\}$ は次のように与えられる。

$$\tilde{\boldsymbol{p}}_\alpha = \boldsymbol{p}_C + \boldsymbol{P}_{d-1}(\boldsymbol{p}_\alpha - \boldsymbol{p}_C) \tag{12}$$

このときの残差 (データと当てはめたアフィン空間との距離の二乗和) は次のようになる。

$$\hat{J}' = \sum_{i=d}^{\nu} \lambda'_i \tag{13}$$

6. 幾何学的モデル選択

前報 [8] と同様に、各点 $\boldsymbol{p}_\alpha$ の各要素に期待値 0、標準偏差 ϵ の正規分布に従う独立な誤差が加わると仮定する。併合の候補となる二つのアフィン空間に対して、別々のアフィン空間を当てはめた場合と一つのアフィン空間を当てはめた場合に対して残差を比較し、幾何学的 **AIC**[6, 7] によってどちらが統計的に妥当か判定する。

いま $\mathcal{A}_i$, $\mathcal{A}_j$ を併合の候補となるアフィン空間とし、含まれる点数をそれぞれ N_i, N_j とする。それぞれの残差 $\hat{J}'_i$, $\hat{J}'_j$ は次元補正を行う過程で式 (13) で計算される。これらを一つのアフィン空間に合併した場合の残差を $\hat{J}'_{i\oplus j}$ とする。n 次元空間の $d-1$ 次元アフィン空間は $d(n-d+1)$ 個のパラメータで指定されるから[1]、一つのアフィン空間を当てはめる場合の幾何学的 AIC は次のようになる。

$$\text{G-AIC}_{i\oplus j} = \hat{J}'_{i\oplus j} + 2\Big((d-1)(N_i+N_j) + d(n-d+1)\Big)\epsilon^2 \tag{14}$$

二つのアフィン空間を別々に当てはめる場合は自由度はそれぞれのアフィン空間の和となるから、幾何学的 AIC は次のようになる。

$$\text{G-AIC}_{i,j} = \hat{J}'_i + \hat{J}'_j + 2\Big((d-1)(N_i+N_j) + 2d(n-d+1)\Big)\epsilon^2 \tag{15}$$

そして $\text{G-AIC}_{i\oplus j} < \text{G-AIC}_{i,j}$ なら併合することが統計的に妥当であると考えられる。これと $|Q_{\alpha\beta}|$ の持つ情報を加味して、第 i アフィン空間 $\mathcal{A}_i$ と第 j アフィン空間 $\mathcal{A}_j$ の類似度を

$$s_{ij} = \begin{cases} 0 & \mathcal{A}_i = \emptyset \text{ または } \mathcal{A}_j = \emptyset \\ \dfrac{\text{G-AIC}_{i,j}}{\text{G-AIC}_{i\oplus j}} \max_{\mathbf{p}_\alpha \in \mathcal{A}_i, \mathbf{p}_\beta \in \mathcal{A}_j} |Q_{\alpha\beta}| & \text{その他} \end{cases} \tag{16}$$

と定義し、これが大きいアフィン空間から次々に併合し、二つのグループになったところで終了する。

仮定より各アフィン空間は d 個以上の点を含まなければならないが、場合によっては最終的に d 個未満の点のグループと残りと要素とに分離されることがある。これを防ぐために、併合の過程で d 個未満の要素のものがある限り、それらを優先して併合する。

7. ノイズレベルの推定

幾何学的 AIC の計算には誤差の標準偏差 ϵ が必要である。その推定には、誤差がなければデータ $\{\boldsymbol{p}_\alpha\}$ は n 次元空間の $r-1$ $(= 2d-1)$ 次元アフィン空間にあることを利用する[2]。$\{\boldsymbol{p}_\alpha\}$ に $r-1$ 次元アフィン空間を最適に当てはめた残差を $\hat{J}'_{\text{total}}$ とすると、次の形の ϵ^2 の不偏推定量が得られる。

$$\hat{\epsilon}^2 = \frac{\hat{J}'_{\text{total}}}{(n-r+1)(N-r)} \tag{17}$$

残差 $\hat{J}'_{\text{total}}$ は全データの重心を $\boldsymbol{p}_C$ とし、全体のモー

[1] n 次元アフィン空間中に d 個の点を指定すればよい。各点は $d-1$ 次元アフィン空間に沿って任意に動かしてよいから、実質的なパラメータ数は $nd - d(d-1)$ である。

[2] n $(\geq r+s-1)$ 次元の空間中の一般の位置にある r 次元アフィン空間と s 次元アフィン空間を含む最低次元のアフィン空間の次元は $r+s-1$ である。

メント行列

$$\boldsymbol{M}'_{\rm total} = \sum_{\alpha=1}^{N} (\boldsymbol{p}_\alpha - \boldsymbol{p}_C)(\boldsymbol{p}_\alpha - \boldsymbol{p}_C)^\top \qquad (18)$$

の固有値を大きい順に $\lambda'_1 \geq \cdots \geq \lambda'_n$ とするとき、次のように計算される (付録 B, C)。

$$\hat{J}'_{\rm total} = \sum_{k=r}^{n} \lambda'_k \qquad (19)$$

8. 最終的な再分類

以上の手順でアフィン空間を併合していくと、途中で誤って分類された点は最後まで残る。そこで最終的に得られたアフィン空間 $\mathcal{A}_1$, $\mathcal{A}_2$ に属するデータからアウトライアを除去し、再分類する。

原点に近い点は誤分類されやすいと考えられるので、まず $\mathcal{A}_1$, $\mathcal{A}_2$ 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)。それらにアフィン空間 $\mathcal{A}'_1$, $\mathcal{A}'_2$ を最適に当てはめ、$\mathcal{A}_1$ 内で $\mathcal{A}'_2$ からの距離が大きい半分 (ただし d 個以上)、および $\mathcal{A}_2$ 内で $\mathcal{A}'_1$ からの距離が大きい半分 (ただし d 個以上) を選び、それぞれに再度アフィン空間 $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2$ を当てはめる。そして各点 $\boldsymbol{p}_\alpha$ を $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2$ のうちの距離の小さいほうに割り当てる。得られたそれぞれのクラスにさらに最小メジアン法 [13] によってアフィン空間を当てはめ (付録 D)、各点を距離の小さいほうに割り当てる。

点 $\boldsymbol{p}$ のアフィン空間 $\mathcal{A}$ からの距離 $D(\boldsymbol{p}, \mathcal{A})$ は、式 (11) の重心 $\boldsymbol{p}_C$ と式 (9) の射影行列 $\boldsymbol{P}_{d-1}$ とを用いて次のように表せる ($\boldsymbol{I}$ は単位行列)。

$$D(\boldsymbol{p}, \mathcal{A}) = \sqrt{(\boldsymbol{p} - \boldsymbol{p}_C, (\boldsymbol{I} - \boldsymbol{P}_{d-1})(\boldsymbol{p} - \boldsymbol{p}_C))} \qquad (20)$$

9. 精度の限界および計算の高速化

前報と同様に、一様等方の正規分布に従う誤差モデルのもとでの理想的な分類は "最尤推定"、すなわち各点 $\boldsymbol{p}_\alpha$ を "真の" アフィン空間 $\bar{\mathcal{A}}_1$, $\bar{\mathcal{A}}_2$ のうちの距離の小さいほうに割り当てることである。もちろん $\bar{\mathcal{A}}_1$, $\bar{\mathcal{A}}_2$ は未知であるから、実データに適用することはできない。しかし、解が既知のシミュレーションでは、これによる分類の誤り率を一つの限界として比較することにより、緒手法の性能の評価ができる。

また前報に述べたのように、固有値問題は特異値分解によって計算が効率化される。ただし、これは計算の便宜であり、背景となる幾何学的な構造とは区別しなければならない。

10. シミュレーション実験

図 1 は 2 次元平面上を独立に運動する 20 個の背景点と 9 個の物体点の動画像である。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、物体と背景の分離を試みた。

図 2 に横軸を ϵ とし、 縦軸に単純な Costeira-Kanade 法 [1]、市村の方法 [3]、部分空間分離法、本報のアフィン空間分離法の各々に対して、独立な 500 回の試行の平均誤り率をプロットした。図 3 は 3 次元空間を独立に運動する 20 個の背景点と 14 個の物体点の物体点の動画像であり、図 4 は図 2 と同様に評価した結果である。これらの結果からアフィン空間分離法は部分空間分離法に比べて著しく性能がよいことがわかる。

なお図 2, 4 とも 9 節に述べた限界を計算すると、ここに示した誤差の範囲ではすべて 0 となった。

11. 透視効果の影響

部分空間分離法もアフィン空間分離法もアフィンカメラの仮定に基いているが、実際の画像は透視変換によるものであり、その食い違いは誤差に繰り込まれる。しかし部分空間分離法とアフィン空間分離法ではアフィンカメラの仮定への依存度が異なるため、モデル化誤差も異なる。

モデル化誤差を期待値 0 の一様等方正規分布の標準偏差 ϵ に換算したものを実効ノイズレベルと呼ぶことする。部分空間分離法では、データ点のそれが属すべき部分空間からの距離の二乗和を 2 物体についてそれぞれ $\hat{J}_1$, $\hat{J}_2$ とすると、次式から評価できる。

$$\epsilon_s^2 = \frac{\hat{J}_1 + \hat{J}_2}{(n-d)(N-d)} \qquad (21)$$

アフィン空間分離法ではデータ点のそれが属すべきアフィン空間からの距離の二乗和をそれぞれ $\hat{J}'_1$, $\hat{J}'_2$ とすると、次式となる。

$$\epsilon_a^2 = \frac{\hat{J}'_1 + \hat{J}'_2}{(n-d+1)(N-d)} \qquad (22)$$

これらを図 3 の 3 次元運動を画角 α (フレームの両端を見込む角度) の透視変換に直したものについて計算した。図 5(a) に画角を $\alpha = 0^\circ$ (平行投影), 40°, 80° の透視投影としたときの図 3 の 3 枚目の画像を表示している。計算した実効ノイズレベルは図 5(b) のようになる。このようにアフィン空間分離法のほうが

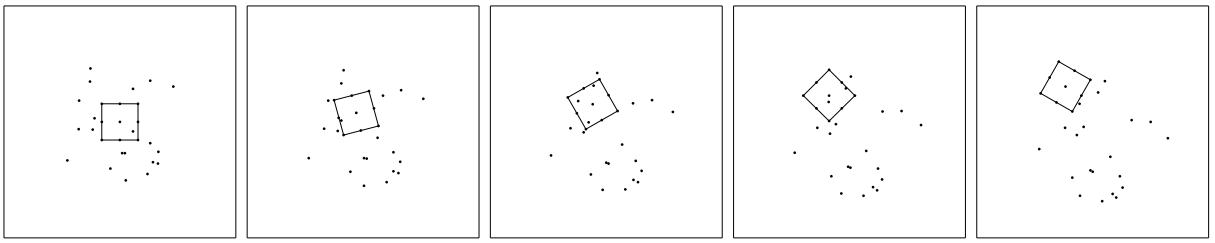


図 1:　2 次元平面上を運動する点。

アフィンカメラの仮定により強く依存しているので実効ノイズレベルが大きくなる。

図 6 は画角 $\alpha = 0°, 40°, 80°$ に対してランダム誤差を加えて図 4 と同様に評価したものである。実効ノイズレベルが等しければ部分空間分離法よりアフィン空間分離法が優れるが、アフィン空間分離法のほうが実効ノイズレベルが大きいので、総合的な評価はそのバランスに依存する。実際、図 6 から誤差がごく小さいときにアフィン空間分離法と部分空間分離法の性能が逆転する部分のあることがわかる。しかしランダム誤差が大きくなると、アフィン空間分離法のほうがランダム誤差を打ち消す力が強いので再び優位となる。

一方、実際の場面によく現れる 2 次元運動ではモデル化誤差が存在しないのでアフィン空間分離法が大きな威力を発揮する。

12. 実画像実験

図 7 の上段は背景と独立に移動する物体の実画像列である。下段に示した特徴点を手で (正確に) 選んで 3 次元運動の分離を行なうと、単純な Costeira-Kanade 法と市村の方法では誤りが生じたが、部分空間分離法およびアフィン空間分離法では共に正しく分離された。式 (21), (22) で評価した実効ノイズレベルは $\epsilon_s = 0.64$ (画素)、$\epsilon_a = 0.94$ (画素) であり、予想通りアフィン空間分離法のほうが大きい。各特徴点の画像座標に期待値 0、標準偏差 1 画素の正規乱数の誤差を加えて 500 回試行したところ、分離の平均正解率は部分空間分離法では 100%、アフィン空間分離法では 90.9% であった。標準偏差 2 画素の場合はそれぞれ 100%, 83.5% となった。この例のように透視効果が強く、かつ特徴点の誤差が小さい場合は部分空間分離法とアフィン空間分離法の性能が逆転する。

12. まとめ

本報では、シーン中を剛体運動する点の画像上の運動がその履歴の空間内のあるアフィン空間に拘束されていることを利用し、複数画像間の対応点を独立に移動する物体に分割する部分空間分離法を強化

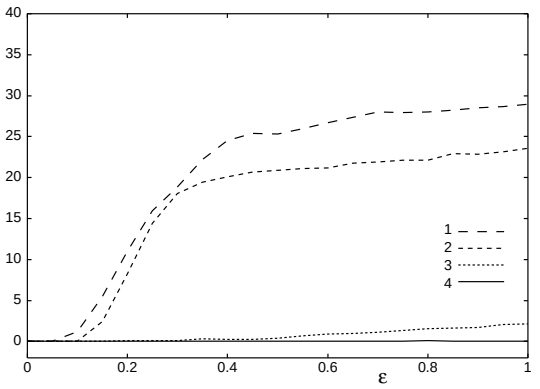


図 2:　2 次元運動の分離の誤り率。1. Costeira-Kanade 法、2. 市村の方法、3. 部分空間分離法、4. アフィン空間分離法。

した。これは、Costeira-Kanade 法にアフィン拘束条件に対する次元補正法、幾何学的 AIC によるモデル選択、および最小メジアン法によるロバスト推定を適用して高精度のアルゴリズムを構成したものである。その性能をシミュレーションや実画像に対して確認するとともに、透視効果の影響を評価した。結論として、透視効果が少ない場合は極めて高性能であるが、透視効果が強く、かつ誤差が小さいときには部分空間分離法に比べて不利になることがわかった。

前報と同様に本手法は経験的に調節すべきパラメータを何も含んでいない。本論文では物体の個数は既知としたが、次報でその自動判定法について述べる。

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

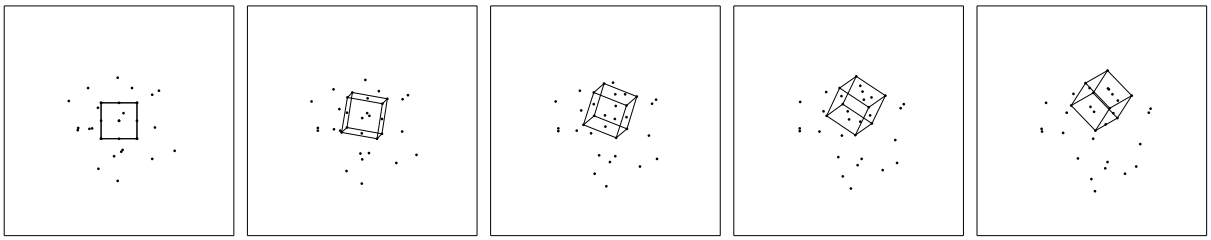


図 3:　3 次元空間中を運動する点。

[7] 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.

[8] 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.

[9] 牧淳人, 渡邊睦, 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.

[10] 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.

[11] 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.

[12] 佐藤淳,「コンピュータビジョン—視覚の幾何学—」, コロナ社, 1999.

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

[14] 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.

[15] 徐剛, 辻三郎,「3 次元ビジョン」, 共立出版, 1998.

付録 A: アフィンカメラ

XYZ カメラ座標系を世界座標系と同一視し、Z 軸を光軸とみなす。空間の点 (X, Y, Z) が画像座標 (x, y) の点に投影されるとき、平行投影とは $x = X$, $y = Y$ を仮定するものである。これは次のように書ける。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{\Pi} \begin{pmatrix} X \\ Y \\ Z \\ f \end{pmatrix} \tag{23}$$

ただし f はカメラの焦点距離であり、$\mathbf{\Pi}$ は次のように定義した投影行列である [12, 15]。

$$\mathbf{\Pi} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{24}$$

物体に固定した座標系の原点を $\boldsymbol{t}$、その正規直交基底を $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ とすると、物体座標 (a, b, c) の点は世界座標系では $\boldsymbol{r} = \boldsymbol{t} + a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}$ の位置にある。これが投影される点の画像座標 (x, y) は次のようになる。

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \mathbf{\Pi} \begin{pmatrix} \boldsymbol{r} \\ f \end{pmatrix} \\ &= \mathbf{\Pi} \left(\begin{pmatrix} \boldsymbol{t} \\ f \end{pmatrix} + a \begin{pmatrix} \boldsymbol{i} \\ 0 \end{pmatrix} + b \begin{pmatrix} \boldsymbol{j} \\ 0 \end{pmatrix} + c \begin{pmatrix} \boldsymbol{k} \\ 0 \end{pmatrix} \right) \\ &= \tilde{\boldsymbol{t}} + a\tilde{\boldsymbol{i}} + b\tilde{\boldsymbol{j}} + c\tilde{\boldsymbol{k}} \end{aligned} \tag{25}$$

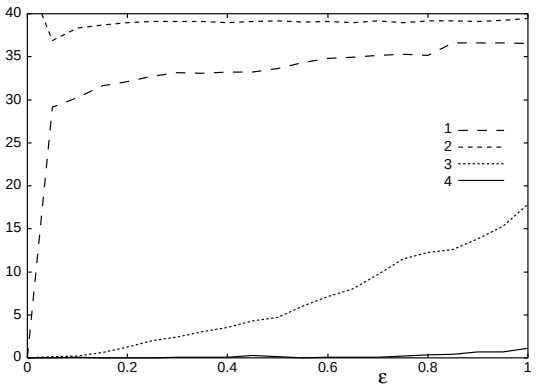


図 4:　3 次元運動の分離の誤り率。1. Costeira-Kanade 法、2. 市村の方法、3. 部分空間分離法、4. アフィン空間分離法。

ただし次のように置いた。

$$\begin{aligned} \tilde{\boldsymbol{t}} &= \mathbf{\Pi} \begin{pmatrix} \boldsymbol{t} \\ f \end{pmatrix}, \quad \tilde{\boldsymbol{i}} = \mathbf{\Pi} \begin{pmatrix} \boldsymbol{i} \\ 0 \end{pmatrix} \\ \tilde{\boldsymbol{j}} &= \mathbf{\Pi} \begin{pmatrix} \boldsymbol{j} \\ 0 \end{pmatrix}, \quad \tilde{\boldsymbol{k}} = \mathbf{\Pi} \begin{pmatrix} \boldsymbol{k} \\ 0 \end{pmatrix} \end{aligned} \tag{26}$$

弱透視変換とは空間の点 (X, Y, Z) が次のような関係で定まる画像座標 (x, y) の点に投影されるとみなすものである [5]。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_3} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{27}$$

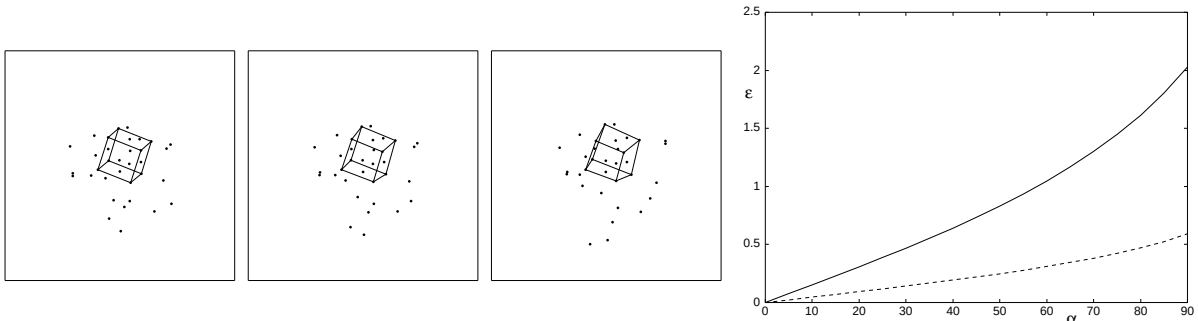
ここに t_3 は $\boldsymbol{t}$ の第 3 成分である。式 (27) は次のような投影行列 $\mathbf{\Pi}$ を定義すれば式 (23) のように書ける。

$$\mathbf{\Pi} = \frac{f}{t_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{28}$$

疑似透視変換とは空間の点 (X, Y, Z) が次のような関係で定まる画像座標 (x, y) の点に投影されるとみなすものである [5]。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_3} \left(\begin{pmatrix} X \\ Y \end{pmatrix} + \left(1 - \frac{Z}{t_3}\right) \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \right) \tag{29}$$

ここに t_1, t_2, t_3 はそれぞれ $\boldsymbol{t}$ の第 1, 第 2, 第 3 成分である。式 (29) は次のような投影行列 $\mathbf{\Pi}$ を定義す



(a) (b)

図 5: (a) 画角を $\alpha = 0^\circ$ (平行投影), 40°, 80° の透視投影としたときの図 3 の中央の画像。(b) 画角 α の透視変換に対する実効ノイズレベル。実線がアフィン空間分離法、破線が部分空間分離法。

れば式 (23) のように書ける。

$$\mathbf{\Pi} = \frac{f}{t_3} \begin{pmatrix} 1 & 0 & -t_1/t_3 & t_1/f \\ 0 & 1 & -t_2/t_3 & t_2/f \end{pmatrix} \tag{30}$$

これを一般化し、空間の点 (X, Y, Z) が式 (23) の形で画像座標 (x, y) の点に投影されるような個々の点に依存しない射影行列 $\mathbf{\Pi}$ が存在するとき、その投影モデルをアフィンカメラと呼ぶ [12, 15]。ただし $\mathbf{\Pi}$ はフレームごとに、また対象物体ごとに異なってもよい。アフィンモデルでは常に式 (25) が成り立つ。カメラの撮像は本来は透視変換としてモデル化するべきであるが、多くの実際的な場面ではアフィンカメラモデルがよい近似であることが知られている [12, 15]。

付録 **B:** 部分空間の当てはめ

n 次元空間の N 点 $\{\boldsymbol{p}_\alpha\}$ に d 次元部分空間 $\mathcal{L}$ を当てはめるとする。$\mathcal{L}$ 内に正規直交基底 $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_d\}$ をとり、その直交補空間 $\mathcal{L}^\perp$ の正規直交基底を $\{\boldsymbol{u}_{d+1}, ..., \boldsymbol{u}_n\}$ とする。点 $\boldsymbol{p}_\alpha$ から $\mathcal{L}$ までの距離は $\sum_{j=d+1}^{n} (\boldsymbol{u}_j, \boldsymbol{p}_\alpha)^2$ となる。誤差が独立、等方、一様な正規分布であれば、最尤推定は次の関数を最小化することを意味する。

$$J = \sum_{\alpha=1}^{N} \sum_{j=d+1}^{n} (\boldsymbol{u}_j, \boldsymbol{p}_\alpha)^2 = \sum_{j=d+1}^{n} (\boldsymbol{u}_j, \sum_{\alpha=1}^{N} \boldsymbol{p}_\alpha \boldsymbol{p}_\alpha^\top \boldsymbol{u}_j)$$
$$= \sum_{j=d+1}^{n} (\boldsymbol{u}_j, \boldsymbol{M}\boldsymbol{u}_j) \tag{31}$$

ただしモーメント行列を次のように置いた。

$$\boldsymbol{M} = \sum_{\alpha=1}^{N} \boldsymbol{p}_\alpha \boldsymbol{p}_\alpha^\top \tag{32}$$

基底 $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_n\}$ が J を最小化するように選ばれているなら、その任意の変分に対して J の第 1 変分が 0 にならなければならない。特に $\mathcal{L}^\perp$ のある $\boldsymbol{u}_j$, $j > d$ と $\mathcal{L}$ のある $\boldsymbol{u}_i$, $i \leq d$ の張る面内でこれらを微小角度 $\delta\theta$ だけ回転することを考える。このとき第 1 変分

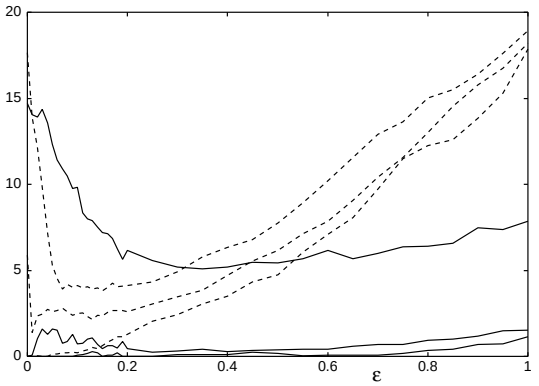


図 6: 透視変換の場合の分離の誤り率。実線がアフィン空間分離法、破線が部分空間分離法。下から画角が 0°, 40°, 80°。

は $\delta\boldsymbol{u}_j = \delta\theta\boldsymbol{u}_i$, $\delta\boldsymbol{u}_i = -\delta\theta\boldsymbol{u}_j$ であるから、J の第 1 変分は次のようになる。

$$\delta J = 2(\delta\boldsymbol{u}_j, \boldsymbol{M}\boldsymbol{u}_j) = 2(\boldsymbol{u}_i, \boldsymbol{M}\boldsymbol{u}_j)\delta\theta \tag{33}$$

これが任意の $\delta\theta$, $j > d$, $i \leq d$ で 0 でなければならないから、$i = 1, ..., d$, $j = d+1, ..., n$ に対して次の関係が成り立つ。

$$(\boldsymbol{u}_i, \boldsymbol{M}\boldsymbol{u}_j) = 0, \tag{34}$$

これは半正値対称行列 $\boldsymbol{M}$ が基底 $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_n\}$ に関して $d \times d$ および $(n-d) \times (n-d)$ の区画にブロック対角化されることを意味する。したがって $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_d\}$ および $\{\boldsymbol{u}_{d+1}, ..., \boldsymbol{u}_n\}$ をそれぞれ適当に回転すると、当てはめた部分空間 $\mathcal{L}$ を変えることなく行列 $\boldsymbol{M}$ を対角化することができる。その結果得られる基底 $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_n\}$ は $\boldsymbol{M}$ の固有ベクトルであり、対応する固有値 $\lambda_1, ..., \lambda_n$ (≥ 0) がその対角要素となる。このとき式 (31) から、残差 J は次のように表せる。

$$J = \sum_{j=d+1}^{n} \lambda_j \tag{35}$$

明らかに $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_d\}$ が $\boldsymbol{M}$ の大きい d 個の固有値に対応する固有ベクトル、残りが $\{\boldsymbol{u}_{d+1}, ..., \boldsymbol{u}_n\}$ のときこれが最小となる。

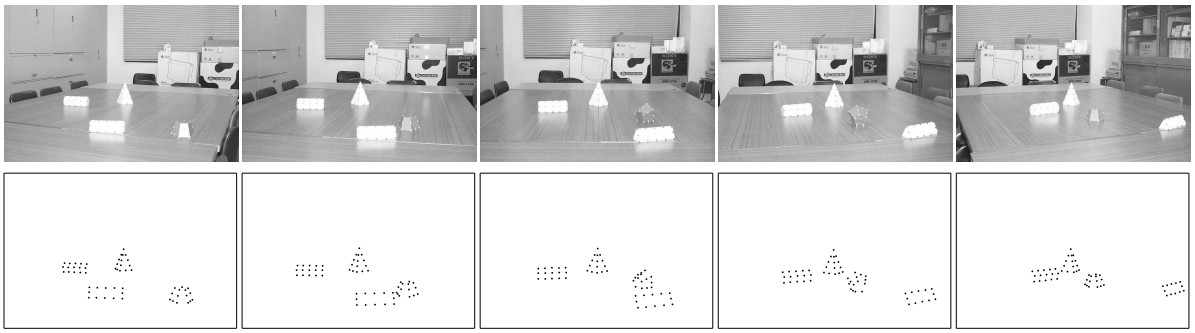


図 7: 背景と独立に移動する物体の実画像 (上) とその特徴点 (下)

付録 C: アフィン空間の当てはめ

n 次元空間の N 点 $\{\boldsymbol{p}_\alpha\}$ に $d-1$ 次元アフィン空間 $\mathcal{A}$ を当てはめるとする。$\mathcal{A}$ 上の 1 点を $\boldsymbol{p}_C$ をとし、$\mathcal{A}$ 内に正規直交系 $\{\boldsymbol{u}_1, ..., \boldsymbol{u}_d\}$ をとり、それに直交する正規直交系を $\{\boldsymbol{u}_{d+1}, ..., \boldsymbol{u}_n\}$ とする。点 $\boldsymbol{p}_\alpha$ から $\mathcal{A}$ までの距離は $\sum_{j=d+1}^{n}(\boldsymbol{u}_j, \boldsymbol{p}_\alpha - \boldsymbol{p}_C)^2$ となる。誤差が独立、等方、一様な正規分布であれば、最尤推定は次の関数を最小化することを意味する。

$$J = \sum_{\alpha=1}^{N}\sum_{j=d+1}^{n}(\boldsymbol{u}_j, \boldsymbol{p}_\alpha - \boldsymbol{p}_C)^2 \tag{36}$$

これを $\boldsymbol{p}_C$ で微分すると次のようになる。

$$\begin{aligned}\nabla_{\boldsymbol{p}_C} J &= -2\sum_{\alpha=1}^{N}\sum_{j=d+1}^{n}(\boldsymbol{u}_j, \boldsymbol{p}_\alpha - \boldsymbol{p}_C)\boldsymbol{u}_j \\ &= -2\boldsymbol{P}^{n-d+1}\Big(\sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha - N\boldsymbol{p}_C\Big)\end{aligned} \tag{37}$$

ただし次のように置いた。

$$\boldsymbol{P}^{n-d+1} = \sum_{j=d+1}^{n}\boldsymbol{u}_j\boldsymbol{u}_j^\top \tag{38}$$

これは $\mathcal{A}$ に直交する $n-d+1$ 次元方向への射影行列である (式 (11) の $\boldsymbol{P}_{d-1}$ を用いれば $\boldsymbol{P}^{n-d+1} = \boldsymbol{I} - \boldsymbol{P}_{d-1}$ と書ける)。式 (37) を $\boldsymbol{0}$ と置いて変形すると次の関係を得る。

$$\boldsymbol{P}^{n-d+1}\Big(\boldsymbol{p}_C - \frac{1}{N}\sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha\Big) = \boldsymbol{0} \tag{39}$$

これは $\boldsymbol{p}_C - \sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha/N$ が $\mathcal{A}$ と同じ方向にあること、すなわち $\boldsymbol{p}_C$ は重心 $\sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha/N$ を $\mathcal{A}$ に沿って任意に移動したものであることがわかる。特に $\boldsymbol{p}_C$ を重心 $\sum_{\alpha=1}^{N}\boldsymbol{p}_\alpha/N$ そのものにとってよい。

座標系の原点を重心 $\boldsymbol{p}_C$ に平行移動すれば、アフィン空間 $\mathcal{A}$ の当てはめは部分空間の当てはめに帰着するから、付録 B のように定まる (ただし次元は d ではなく $d-1$ となる)。以上より式 (9)–(13) および式 (20) が得られる。

付録 D: 最小メジアン法

n 次元ベクトル $\{\boldsymbol{p}_\alpha\}$, $\alpha = 1, ..., N$, に d 次元部分空間を最小メジアン法で当てはめるには次のようにする [13]。

1. $S_m = \infty$ (実際には十分大きい数) と置く。
2. $\{\boldsymbol{p}_\alpha\}$ からランダムに d 個を選ぶ。
3. それらに d 次元部分空間 $\mathcal{L}$ を当てはめる。
4. データ $\{\boldsymbol{p}_\alpha\}$ から当てはめた $\mathcal{L}$ までの距離に二乗のメジアンを計算する。

$$S = \mathrm{med}_{\alpha=1}^{N} D(\boldsymbol{p}_\alpha, \mathcal{L})^2 \tag{40}$$

5. $S < S_m$ なら $S_m \leftarrow S$ と更新する。
6. ステップ 2 に戻り、S_m が十分収束するまでこれを反復する。

n 次元ベクトル $\{\boldsymbol{p}_\alpha\}$, $\alpha = 1, ..., N$, に $d-1$ 次元アフィン空間を最小メジアン法で当てはめるには次のようにする [13]。

1. $S_m = \infty$ (実際には十分大きい数) と置く。
2. $\{\boldsymbol{p}_\alpha\}$ からランダムに d 個を選ぶ。
3. それらに $d-1$ 次元アフィン空間 $\mathcal{A}$ を当てはめる。
4. データ $\{\boldsymbol{p}_\alpha\}$ から当てはめた $\mathcal{A}$ までの距離に二乗のメジアンを計算する。

$$S = \mathrm{med}_{\alpha=1}^{N} D(\boldsymbol{p}_\alpha, \mathcal{A})^2 \tag{41}$$

5. $S < S_m$ なら $S_m \leftarrow S$ と更新する。
6. ステップ 2 に戻り、S_m が十分収束するまでこれを反復する。