信頼性評価を備えた最適な射影変換の計算プログラム

清水 慶行 [†]　太田 直哉 [†]　金谷 健一 [†]

[†] 群馬大学工学部情報工学科

本論文では平面上に与えられた二つの点集合間の射影変換を最適に計算するとともに計算の信頼性評価を与えるプログラムパッケージを記述する。その原理は最尤推定を "くりこみ法" と呼ばれる手法で計算するものである。まず入出力と信頼性評価の表現法、およびアルゴリズムの概要を述べ、プログラムの仕様を説明し、最後に実行例を示し、プログラムの特徴や注意事項を述べる。本プログラムは極めて効率的、高精度、かつロバストであり、コンピュータビジョンの様々な応用に広く利用されることが期待される。

キーワード: 射影変換, プログラム, ソフトウェアライブラリ, 最適当てはめ, くりこみ法, 信頼性評価

Optimal Homography Computation Program with a Reliability Measure

Yoshiyuki Shimizu[†], Naoya Ohta[†] and Kenichi Kanatani[†]

[†]Department of Computer Science
Gunma University, Kiryu, Gunma 376-8515 Japan

This paper describes a program package for optimally computing the homography (projective transformation) between two given sets of points on a plane and evaluating the reliability of the resulting mapping. The basic principle is computing maximum likelihood estimation by a technique called *renormalization*. We first describe the representation of input/output and reliability evaluation. Then, we summarize the algorithm and describe the program specification. Finally, we show execution examples and comment on characteristics of this program and related topics. This program is very efficient, accurate, and robust and is expected to be used widely for many computer vision applications.

Key words: homography, program, software library, optimal fitting, renormalization, reliability evaluation

謝辞: 本研究の一部は文部省科学研究費基盤研究 C (2) (No. 09680352) 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)30-1842, Fax: (0277)30-1801,
E-mail: {shimizu|ohta|kanatani}@ail.cs.gunma-u.ac.jp

1．序論

射影変換とは異なる位置から撮影した平面物体の画像間の関係として現われる画像面の変換であり、直線は直線に写像され、共点性 (複数の直線が一点で交わる)、共線性 (複数の点が同一直線上にある) が保存される [3, 9]。そして任意の一般の位置にある 4 点 (どの 3 点も共線ではない) を任意の一般の位置にある 4 点に写す射影変換がただ一つ存在し、任意の一般の位置にある 4 直線 (どの 3 直線も共点ではない) を任意の一般の位置にある 4 直線に写す射影変換がただ一つ存在する。これらの点や直線は "無限遠点" や "無限遠直線" であってもよい。

2 画像間の射影変換が定まれば、撮影位置が未知でもカメラ座標系に関するその平面物体の 3 次元位置が定数倍を除いて計算できる[1][3, 10, 12, 15, 16, 18]。この性質を利用して未校正カメラによる 2 画像から地面や建物のような平面部分をもつ物体を 3 次元復元するシステムも研究されている [17]。

物体が平面ではなくても、地面や床のような平面上に建物、樹木、人物、車両等があるとき、地面や床のマーカー等の対応から画像間の射影変換を計算すれば、一方の画像に計算した写像を施して他方の画像に重ならない部分がそれ以外の物体であると判定できる。これを利用して道路上の障害物を識別したり [13]、その重ならない程度から 3 次元形状を復元したりする試みがある [2]。また平面パタンを設置してステレオ視システムを校正する場合も 2 画像間の射影変換の計算が基本となる [11]。スリット光投影システムの校正においても射影変換の計算が重要な役割を果たす [1]。

このように 2 画像の対応から射影変換を計算することはコンピュータビジョンの最も基本となる演算の一つである。従来は射影変換の式を適当に変形してから最小二乗法を適用するのが普通であった [3, 11, 13]。これはそれにより計算が固有値問題に帰着して計算が簡単になるからである。しかし、このような方法では精密な解が得られず、また解に統計的な偏差が存在することが指摘されている [10, 15]。これは、問題を簡単化する変形によって最小化すべき量がもはや画像上の誤差の大きさではなく、その非線形な関数となるからである。そこでは成分間の比のみが意味をもつ "同次座標" があたかも普通の座標のように取り扱われ、元来の座標値と人為的に導入した定数とが同等に扱われている。

射影変換は "射影幾何学" の基礎をなすものであり、射影幾何学は同次座標によって代数的に非常に美しい形で記述できる [9]。しかし、これは誤差がない場合のことである。誤差は画像座標 ("非同次座標") に入り、その挙動はユークリッド幾何学に従うので、実データを用いる計算を射影幾何学の枠内のみで完全に定式化することはできない。

これに対して金谷 [7, 8] は各種の幾何学的推定問題を抽象化し、データ点数を固定して微小誤差に対する第 1 近似による解析を行い、最尤推定に基く一般的な最適推定の理論を提唱した。この理論を射影変換の計算に適用すると、最適な射影変換が得られるだけでなく、その信頼性が定量的に評価できる。その基本形は文献 [10, 15] に示されている。

本論文の目的は、この射影変換の計算を文献 [10, 15] に従ってプログラムパッケージとして提供することにある。まず入出力と信頼性評価の表現法、およびアルゴリズムの概要を述べ、プログラムの仕様を説明し、最後に実行例を示し、このプログラムの特徴や注意事項を述べる。

2．入出力の表現法

射影変換とは次の方程式で表される写像のことである[2] [3, 6, 9]。

---

[1]数学的には最大 8 個の解が存在するが、物体が両方のカメラの前方にあるという条件から高々 2 個、多くの場合 1 個の解に絞られる [4, 8]。

[2]数学的には式 (2) の行列 $\mathbf{H}$ の行列式が 0 でないと約束する。このときこの写像は射影平面 (無限遠点まで含めた xy 面) の全単射 (1 対 1 かつ上への写像) となり、その全体は群 (**2** 次元射影変換群) を作る [9]。しかし、計算上では行列式が (まれに)0 となる可能性を排除していない。

$$x' = \frac{Ax + By + C}{Px + Qy + R}, \qquad y' = \frac{Dx + Ey + F}{Px + Qy + R} \tag{1}$$

行列 $\boldsymbol{H}$ とベクトル $\boldsymbol{x}$, $\boldsymbol{x}'$ を

$$\boldsymbol{H} = \begin{pmatrix} A & B & C/f \\ D & E & F/f \\ P & Q & R/f \end{pmatrix}, \qquad \boldsymbol{x} = \begin{pmatrix} x/f \\ y/f \\ 1 \end{pmatrix}, \qquad \boldsymbol{x}' = \begin{pmatrix} x'/f \\ y'/f \\ 1 \end{pmatrix} \tag{2}$$

と定義すると式 (1) は次のように書ける。

$$\boldsymbol{x}' = Z[\boldsymbol{H}\boldsymbol{x}] \tag{3}$$

ただし $Z[\,\cdot\,]$ は z 成分を 1 とするスケールの正規化を表す[3]。また f は x/f, y/f を 1 のオーダにするためのスケール因子である。本プログラムでは平面上に与えられた点集合 $\{(x_\alpha, y_\alpha)\}_{\alpha=1}^N$, $\{(x'_\alpha, y'_\alpha)\}_{\alpha=1}^N$ の各点をベクトル $\boldsymbol{x}_\alpha = (x_\alpha/f, y_\alpha/f, 1)^\top$, $\boldsymbol{x}'_\alpha = (x'_\alpha/f, y'_\alpha/f, 1)^\top$ で表す ($\top$ は転置)。定数 f は x_α/f, y_α/f, x'_α/f, y'_α/f, $\alpha = 1, \ldots, N$ がほぼ ± 1 の範囲内に収まるように選ぶ[4]。出力は射影変換を指定する行列 $\hat{\boldsymbol{H}}$ である。これには定数倍の不定性があるので $\|\hat{\boldsymbol{H}}\| = 1$ と正規化する (行列 $\boldsymbol{H}$ $= (H_{ij})$ のノルムは $\|\boldsymbol{H}\| = \sqrt{\sum_{i,j=1}^3 H_{ij}{}^2}$ と定義する)。

3. 信頼性評価の表現法

各点 (x_α, y_α), (x'_α, y'_α) には不確定さがあるする。その共分散行列をそれぞれ $\boldsymbol{\Sigma}_\alpha$, $\boldsymbol{\Sigma}'_\alpha$ とすると、ベクトル表示 $\boldsymbol{x}_\alpha$, $\boldsymbol{x}'_\alpha$ の共分散行列が次のように定義される ($\boldsymbol{0}$ は零ベクトル)。

$$V[\boldsymbol{x}_\alpha] = \begin{pmatrix} \boldsymbol{\Sigma}_\alpha & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{pmatrix}, \qquad V[\boldsymbol{x}'_\alpha] = \begin{pmatrix} \boldsymbol{\Sigma}'_\alpha & \boldsymbol{0} \\ \boldsymbol{0}^\top & 0 \end{pmatrix}, \tag{4}$$

以下では各点の共分散行列 $V[\boldsymbol{x}_\alpha]$, $V[\boldsymbol{x}'_\alpha]$, が定数倍を除いて与えられるとする。すなわち未知の定数 ϵ (ノイズレベルと呼ぶ) が存在して、

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \qquad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha] \tag{5}$$

と書けるとし、$V_0[\boldsymbol{x}_\alpha]$, $V_0[\boldsymbol{x}'_\alpha]$ (正規化共分散行列と呼ぶ) を入力として指定する。これはデータのノイズ依存性を指定するものであり、ノイズに特に方向性がない場合はデフォルト値として $V_0[\boldsymbol{x}_\alpha] = V_0[\boldsymbol{x}'_\alpha] = \mathrm{diag}(1,1,0)$ とする ($\mathrm{diag}(\cdots)$ は $\cdots$ を対角要素とする対角行列)。本プログラムでは点 p をその位置ベクトル $\boldsymbol{x}$ と正規化共分散行列 $V_0[\boldsymbol{x}]$ とをもつオブジェクトとみなし、そのクラスを画像位置と呼ぶ。

このような不確定性をもつデータから計算した射影変換も不確定性を持つ。これを評価するために、パラメータ空間で最も誤差が生じやすい方向に正負の標準偏差だけ偏った点の表す二つの射影変換の行列 $\boldsymbol{H}^{(+)}$, $\boldsymbol{H}^{(-)}$ を計算する。これを標準偏位と呼ぶ [8]。

4. プログラムの概要

[3] $Z[\mathbf{Hx}]$ の第 3 成分が 0 のとき x は無限遠点となる。しかし、計算ではデータとして得られる点列のみを考え、無限遠点は考慮しない。

[4] データがカメラ実画像画像から得られ、座標の単位が画素の場合には f として焦点距離 (単位は画素)、あるいはその近似値を用いればよい

入力: 画像位置の列 $\{p_\alpha\}_{\alpha=1}^N$, $\{p'_\alpha\}_{\alpha=1}^N$ $(N \geq 4)$。
出力: 行列[5] $\hat{\boldsymbol{H}}$、$\boldsymbol{H}^{(+)}$, $\boldsymbol{H}^{(-)}$。
アルゴリズム:

1. $c = 0$, $\boldsymbol{W}_\alpha = \boldsymbol{I}$, $\alpha = 1, ..., N$ と置く。
2. 次のようにテンソル $\mathcal{M}$ を計算する。

$$\mathcal{M} = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3} W_\alpha^{(kl)}(\boldsymbol{e}^{(k)} \times \boldsymbol{x}'_\alpha \boldsymbol{x}_\alpha^\top) \otimes (\boldsymbol{e}^{(l)} \times \boldsymbol{x}'_\alpha \boldsymbol{x}_\alpha^\top) \tag{6}$$

3. 次のようにテンソル $\mathcal{N} = (N_{ijkl})$ を計算する。

$$N_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{m,n,p,q=1}^{3} \varepsilon_{imp}\varepsilon_{knq} W_\alpha^{(mn)}\Big(V_0[\boldsymbol{x}_\alpha]_{jl} x'_{\alpha(p)} x'_{\alpha(q)} + V_0[\boldsymbol{x}'_\alpha]_{pq} x_{\alpha(j)} x_{\alpha(l)}\Big) \tag{7}$$

4. 次のテンソルの 9 個の固有値 $\lambda_1 \geq \cdots \geq \lambda_9$ と対応するノルム 1 の固有行列の正規直交系 $\{\boldsymbol{H}_1, ..., \boldsymbol{H}_9\}$ を計算する。

$$\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N} \tag{8}$$

5. $\lambda_9 \approx 0$ であれば λ_8, $\boldsymbol{H}_8$, $\boldsymbol{H}_9$, c を返す。
6. そうでなければ c, $\boldsymbol{W}_\alpha$ を次のように更新してステップ 2 へ戻る。

$$c \leftarrow c + \frac{\lambda_9}{(\boldsymbol{H}_9; \mathcal{N}\boldsymbol{H}_9)} \tag{9}$$

$$\boldsymbol{W}_\alpha \leftarrow \Big(\boldsymbol{x}'_\alpha \times \boldsymbol{H}_9 V_0[\boldsymbol{x}_\alpha]\boldsymbol{H}_9^\top \times \boldsymbol{x}'_\alpha + (\boldsymbol{H}_9\boldsymbol{x}_\alpha) \times V_0[\boldsymbol{x}'_\alpha] \times (\boldsymbol{H}_9\boldsymbol{x}_\alpha)\Big)_2^- \tag{10}$$

7. 二乗ノイズレベルを次のように推定する。

$$\hat{\epsilon}^2 = \frac{c}{1 - 4/N} \tag{11}$$

8. 次の計算を行う。

$$\boldsymbol{H}^{(+)} = N[\hat{\boldsymbol{H}} + \sqrt{\frac{\hat{\epsilon}^2}{\lambda_8 N}}\boldsymbol{H}_8], \qquad \boldsymbol{H}^{(-)} = N[\hat{\boldsymbol{H}} - \sqrt{\frac{\hat{\epsilon}^2}{\lambda_8 N}}\boldsymbol{H}_8] \tag{12}$$

9. $\hat{\boldsymbol{H}}$, $\boldsymbol{H}^{(+)}$, $\boldsymbol{H}^{(-)}$ を返す。

5. 記号の用語の説明

前節の記述に用いた記号と用語を説明する。$\boldsymbol{I}$ は単位行列である。式 (6) に現われる $x_{\alpha(i)}$, $x'_{\alpha(i)}$ はそれぞれベクトル $\boldsymbol{x}_\alpha$, $\boldsymbol{x}'_\alpha$ の第 i 成分である。$W_\alpha^{(kl)}$ は行列 $\boldsymbol{W}_\alpha$ の (kl) 要素である。演算 $\otimes$ はテ

[5] 零でない値を得るためには $N > 5$ が必要である。

ンソル積を表し、行列 $\boldsymbol{A} = (A_{ij})$, $\boldsymbol{B} = (B_{ij})$, テンソル $\mathcal{C} = (C_{ijkl})$ に対して $\mathcal{C} = \boldsymbol{A} \otimes \boldsymbol{B}$ は $C_{ijkl} = A_{ij}B_{kl}$ と定義する。また $\boldsymbol{e}^{(1)}$, $\boldsymbol{e}^{(2)}$, $\boldsymbol{e}^{(3)}$ は次の単位ベクトルである。

$$\boldsymbol{e}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{13}$$

式 (7) の $V_0[\boldsymbol{x}_\alpha]_{ij}$, $V_0[\boldsymbol{x}'_\alpha]_{ij}$ はそれぞれ正規化共分散行列行列 $V_0[\boldsymbol{x}_\alpha]$, $V_0[\boldsymbol{x}'_\alpha]$ の (ij) 要素である。記号 ε_{ijk} はエディントンのイプシロンであり、(ijk) が (123) の偶置換か奇置換かそれ以外かに応じて $1, -1, 0$ をとると約束する。

テンソル $\hat{\mathcal{M}} = (\hat{M}_{ijkl})$ と行列 $\boldsymbol{H} = (H_{ij})$ の積 $\hat{\mathcal{M}}\boldsymbol{H}$ は (ij) 要素が $\sum_{k,l=1}^{3} \hat{M}_{ijkl}H_{kl}$ の行列であると定義する。このとき、行列 $\boldsymbol{H}$ がテンソル $\hat{\mathcal{M}}$ の固有値 λ の固有行列であるとは $\hat{\mathcal{M}}\boldsymbol{H} = \lambda\boldsymbol{H}$ となることである。この計算には行列 $\boldsymbol{H} = (H_{ij})$ とテンソル $\hat{\mathcal{M}} = (\hat{M}_{ijkl})$ を次のような 9 次元ベクトル $\boldsymbol{h}$ と 9 次元行列 $\hat{\boldsymbol{M}}$ とみなす。

$$\boldsymbol{h} = \begin{pmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ H_{33} \end{pmatrix}, \quad \hat{\boldsymbol{M}} = \begin{pmatrix} \hat{M}_{1111} & \hat{M}_{1112} & \hat{M}_{1113} & \hat{M}_{1121} & \hat{M}_{1122} & \hat{M}_{1123} & \hat{M}_{1131} & \hat{M}_{1132} & \hat{M}_{1133} \\ \hat{M}_{1211} & \hat{M}_{1212} & \hat{M}_{1213} & \hat{M}_{1221} & \hat{M}_{1222} & \hat{M}_{1223} & \hat{M}_{1231} & \hat{M}_{1232} & \hat{M}_{1233} \\ \hat{M}_{1311} & \hat{M}_{1312} & \hat{M}_{1313} & \hat{M}_{1321} & \hat{M}_{1322} & \hat{M}_{1323} & \hat{M}_{1331} & \hat{M}_{1332} & \hat{M}_{1333} \\ \hat{M}_{2111} & \hat{M}_{2112} & \hat{M}_{2113} & \hat{M}_{2121} & \hat{M}_{2122} & \hat{M}_{2123} & \hat{M}_{2131} & \hat{M}_{2132} & \hat{M}_{2133} \\ \hat{M}_{2211} & \hat{M}_{2212} & \hat{M}_{2213} & \hat{M}_{2221} & \hat{M}_{2222} & \hat{M}_{2223} & \hat{M}_{2231} & \hat{M}_{2232} & \hat{M}_{2233} \\ \hat{M}_{2311} & \hat{M}_{2312} & \hat{M}_{2313} & \hat{M}_{2321} & \hat{M}_{2322} & \hat{M}_{2323} & \hat{M}_{2331} & \hat{M}_{2332} & \hat{M}_{2333} \\ \hat{M}_{3111} & \hat{M}_{3112} & \hat{M}_{3113} & \hat{M}_{3121} & \hat{M}_{3122} & \hat{M}_{3123} & \hat{M}_{3131} & \hat{M}_{3132} & \hat{M}_{3133} \\ \hat{M}_{3211} & \hat{M}_{3212} & \hat{M}_{3213} & \hat{M}_{3221} & \hat{M}_{3222} & \hat{M}_{3223} & \hat{M}_{3231} & \hat{M}_{3232} & \hat{M}_{3233} \\ \hat{M}_{3311} & \hat{M}_{3312} & \hat{M}_{3313} & \hat{M}_{3321} & \hat{M}_{3322} & \hat{M}_{3323} & \hat{M}_{3331} & \hat{M}_{3332} & \hat{M}_{3333} \end{pmatrix} \tag{14}$$

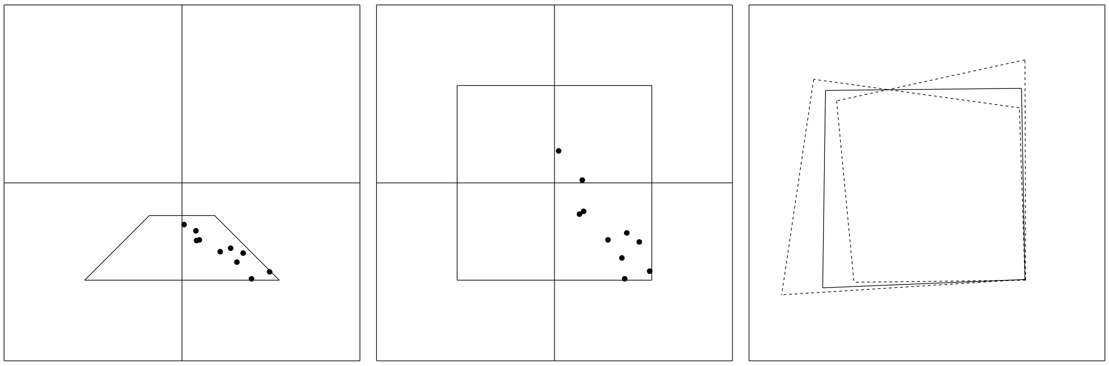
このとき $\boldsymbol{H}$ が $\hat{\mathcal{M}}$ の固有値 λ の固有行列であること ($\hat{\mathcal{M}}\boldsymbol{H} = \lambda\boldsymbol{H}$) と、$\boldsymbol{h}$ が $\hat{\boldsymbol{M}}$ の固有値 λ の固有ベクトルであること ($\hat{\boldsymbol{M}}\boldsymbol{h} = \lambda\boldsymbol{h}$) とが等価になる。また、行列 $\boldsymbol{H}$ のノルム $\|\boldsymbol{H}\|$ は置き換えたベクトル $\boldsymbol{h}$ のノルム $\|\boldsymbol{h}\|$ に一致する。

式 (9) の右辺の分母に現われる演算 $(\,\cdot\,;\,\cdot\,)$ は行列の内積であり、$\boldsymbol{A} = (A_{ij})$, $\boldsymbol{B} = (B_{ij})$ に対して $(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i,j=1}^{3} A_{ij}B_{ij}$ と定義する。式 (10) の右辺に現われるベクトル $\boldsymbol{a} = (a_i)$ と行列 $\boldsymbol{A} = (A_{ij})$ の $\boldsymbol{a} \times \boldsymbol{A} \times \boldsymbol{a}$ の形の積は (ij) 要素が $\sum_{i,j=1}^{3} \varepsilon_{ikl}\varepsilon_{jmn}a_k a_m A_{ln}$ の対称行列であると定義する。演算 $(\,\cdot\,)_2^-$ はランクを 2 に拘束した一般逆行列を表す (8 節参照)。式 (12) に現われる演算 $N[\,\cdot\,]$ はノルムを 1 にする正規化作用素である ($N[\boldsymbol{A}] = \boldsymbol{A}/\|\boldsymbol{A}\|$)。

6. プログラムの仕様

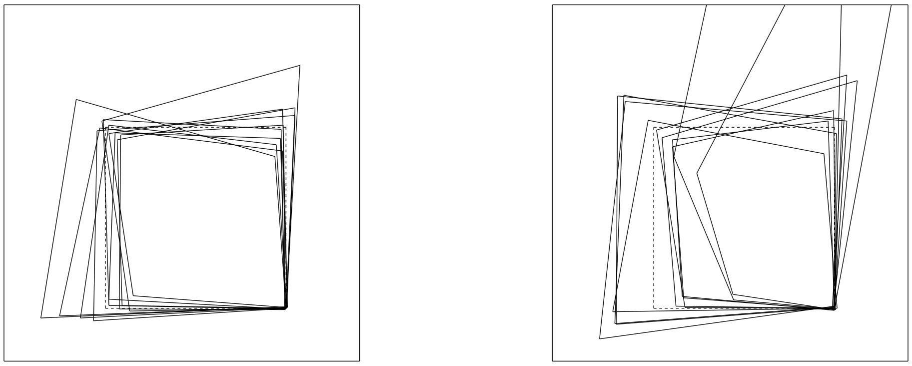
4 節の計算に用いた技法はくりこみ法と呼ばれ (8 節参照)、第 1 近似において最適であることが証明される [7, 8]。すなわち、高次の項を除いて他のどんな方法によってもよりよい結果を得ることは不可能である。このプログラムは C++ により書かれており、付属ライブラリと一体化されている。その構成は次のようになっている。

1. ベクトル・行列計算部
2. テンソル計算部
3. 幾何学的統計計算部



(a) (b) (c)

図 1: (a) 台形領域 (透視画像)。 (b) 正方形領域 (真の形状)。 (c) 境界線の写像とその標準偏位。



(a) (b)

図 2: 誤差を変えて写像した境界線。 (a) 最適推定。 (b) 最小二乗法。

1. では種々のベクトル・行列演算を定義している。基本的な線形計算 (逆行列、行列式、連立 1 次方程式、固有値・固有ベクトル、等) には文献 [14] のプログラムを使用している。 2. は 1. を種々のテンソル演算に拡張したものである。 3. は共分散行列をもった幾何学的対象とその信頼性計算を定義するものであり、 4 節のプログラムもここに収められている。そして、これらを用いて、標準入力から点列データを読み込み、画像位置オブジェクトに変換し (正規化共分散行列はデフォルト値を用いる)、行列 $\hat{\boldsymbol{H}}$, $\boldsymbol{H}^{(+)}$, $\boldsymbol{H}^{(-)}$ を計算し、標準出力に出力するプログラムを別途提供している。これらは `http://www.ail.cs.gunma-u.ac.jp/Labo/research.html` に置かれている。

7. 実行例

図 1(a) は 4 点 $(100, -100)$, $(100/3, -100/3)$, $(-100/3, -100/3)$, $(-100, -100)$ を頂点とする台形領域であり、その内部の右半分にランダムに 10 点をとる。図 1(b) はその各点に射影変換

$$x' = -\frac{100x}{y}, \quad y' = -\frac{200(y+50)}{y} \tag{15}$$

を施したものである。 3 次元的には図 1(a) は、図 1(b) の正方形領域 (“物体モデル”) をカメラの光軸に平行に置いてやや上方から撮影した “画像” であると解釈してもよい。

図 1(a), (b) の各点の座標値に独立に独立に期待値 0、標準偏差 1 の乱数誤差を加えたものから射影変換の行列 $\hat{\boldsymbol{H}}$ とその標準偏位 $\boldsymbol{H}^{(+)}$, $\boldsymbol{H}^{(-)}$ を計算する。図 1(c) は台形領域の境界線をそれぞれに対応する射影変換で写像したものを重ねて図示したものである。実線は $\hat{\boldsymbol{H}}$ に、破線はその標準偏位

$\boldsymbol{H}^{(+)},\ \boldsymbol{H}^{(-)}$ にそれぞれ対応している。図 2(a) は乱数誤差をいろいろ変えて得られる境界線 10 例を重ねて表示したものである (破線は真の形状)。図 1(c) の標準偏位が誤差によるずれの可能性を定量的によく表現していることがわかる。図 2(b) は比較のために最小二乗法で計算した結果を示している。最小二乗法とは式 (3) が

$$\boldsymbol{x}' \times \boldsymbol{H}\boldsymbol{x} = \boldsymbol{0} \tag{16}$$

と等価であることから

$$\sum_{\alpha=1}^{N} \|\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha\|^2 \to \min \tag{17}$$

によって $\boldsymbol{H}$ を定めるものである。左辺は $\boldsymbol{H}$ の 2 次形式であるから、これを最小化するには最小固有値に対する固有行列を計算すればよい。この計算の単純さのためにこの最小二乗法は広く用いられている [3, 11, 13]。しかし、その代償として精度が劣ることが図 2(b) からも確認できる。その理由の一つはベクトル $\boldsymbol{x},\ \boldsymbol{x}'$ の成分が射影幾何学における同次座標であって、それらの第 3 成分は定義により常に 1 であるにもかかわらず、最小二乗法ではデータも定数も同等に扱っているからである。

この例の本プログラムによる当てはめの実行時間は UNIX ワークステーション (Sun Ultra-1) 上で平均 1.58 秒 (CPU time) であり、同じ問題に最小二乗法を用いた場合の約 41.2 倍である。図 1(a), (b) では変形が大きい上、データが領域内で偏在しているため最適推定と最小二乗法との差が大きいが、変形が少なくデータの分布が領域内で一様になるにつれて差は少なくなる。そのような場合は最小二乗法を用いれば精度を落とさずに計算が効率化される。これは実時間処理には有利である。しかし、そのためには、考えている状況であらかじめ最適推定と最小二乗法の両方を比較して結果に差がないことを確認しなければならない。このように最適推定は実行時間では不利でもそのようなベンチマークとしての意義が大きい。

注意すべきことは、二つの方法で計算した $\hat{\boldsymbol{H}}_1,\ \hat{\boldsymbol{H}}_2$ に対して単に $\|\hat{\boldsymbol{H}}_1 - \hat{\boldsymbol{H}}_2\|$ が小さいからといって安心できないことである。なぜなら射影変換は部分的に極端に拡大縮小が起こり得るので、行列表示は似ていても画像上の対応する変換は非常に異なることが多いからある。実際、図 2の比較においても行列 $\boldsymbol{H}$ 自体は差が非常に小さい。

8. 関連する事項

解法の原理は文献 [10, 15] に述べられているが、基本的には最尤推定である [7]。これは形式的には行列 $\boldsymbol{H}$ を含む重み行列 $\boldsymbol{W}_\alpha(\boldsymbol{H})$(式 (10) 参照) をもつ最小二乗法

$$\sum_{\alpha=1}^{N} (\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha, \boldsymbol{W}_\alpha(\boldsymbol{H})(\boldsymbol{x}'_\alpha \times \boldsymbol{H}\boldsymbol{x}_\alpha)) \to \min \tag{18}$$

とみなせる [8]。ただし $(\boldsymbol{a}, \boldsymbol{b})$ はベクトル $\boldsymbol{a},\ \boldsymbol{b}$ の内積を表す。左辺は定数倍と定数項を除いて、行列 $\boldsymbol{H}$ で定まる射影変換のもとにデータ $\{\boldsymbol{x}_\alpha,\ \boldsymbol{x}'_\alpha\}$ が観測される尤度の対数の符号を換えたものに等しい。特に誤差が等方かつ一様の場合はデータのあるべき位置からの距離の二乗和に等しい。

式 (18) より、 $\boldsymbol{W}_\alpha(\boldsymbol{H})$ が定数行列であれば解 $\boldsymbol{H}$ はテンソル $\mathcal{M}$ (式 (6) 参照) の最小固有値に対する固有行列である。しかし、これを $\boldsymbol{W}_\alpha(\boldsymbol{H})$ と残りの部分を交互に反復計算したのでは解に統計的な偏差が生じる [6, 8]。一方、誤差がないとき $\mathcal{M}$ の最小固有値が 0 であることから、毎回の反復で $\mathcal{M}$ の最小固有値を 0 に近付ける補正を行えば精度の理論限界 [7] を達成する精度の解が得られる [8]。これがくりこみ法と呼ばれる手法である [5, 6]。

式 (9) にランクを 2 に拘束した一般逆行列を用いているが、これを計算するには固有値と固有ベクトルを計算して対角行列に対角化し、最小固有値を 0 に置き換えて、残りの固有値をそれぞれ逆数に置き換えてもとの形に戻せばよい [6, 8]。この演算が必要となる理由は、式 (3)、あるいはこれ等価な式(16)

の 3 個の成分方程式のうち **2** 個のみが独立であることにある。これを反映して行列 $\boldsymbol{W}_\alpha$ は誤差がなければランクが 2 である。しかし、データを代入した行列は一般にランクが 3 であり、最小固有値は厳密には 0 ではない。一般逆行列を計算する際にランクを 2 に強制するのはデータから計算した値で $\boldsymbol{W}_\alpha$ の真の値を近似しようとするためである [8]。

文献 [10, 15] では固有値の摂動を 2 次の項まで展開して補正する "2 次のくりこみ法" が示されている。本プログラムではデフォルトとして 1 次の項までの展開を用いる "1 次のくりこみ法" を採用しているが (式 (9) 参照)、オプションとして 2 次のくりこみ法も指定できる。 2 次のくりこみ法を用いると収束までの反復回数が減少する (1 次で平均 4, 5 回が 2 次では平均 2, 3 回)。しかし一回の反復の計算量が増えるので、全体の計算時間はやや増加する。また最終的な精度も有効数字の範囲内でほぼ差がない。

このくりこみ法は極めてロバストであり、相当大きい誤差を与えても速やかに収束する。

9. まとめ

本論文では平面上の二つの点集合から射影変換を最適に計算するとともに計算の信頼性評価を与えるプログラムパッケージを記述した。その原理は最尤推定をくりこみ法によって計算するものであり、文献 [10, 15] に示されている。まず入出力と信頼性評価の表現法、およびアルゴリズムの概要を述べ、プログラムの仕様を説明し、実行例を示し、プログラムの特徴や注意事項を述べた。本プログラムは極めて効率的、高精度、かつロバストであり、コンピュータビジョンの様々な応用に広く利用されることが期待される。

参考文献

[1] D. Q. Huynh, Calibration of a structured light system: A projective approach, *Proc. IEEE Conf. Comput. Vision Patt. Recog.*, June 1997, Puerto Rico, pp. 225–230.

[2] M. Irani and P. Anandan, Parallax geometry of pairs of points for 3D scene analysis, *Proc. 4th European Conf. Comput. Vision*, April 1996, Cambridge, U.K., Vol. 1, pp. 17–30.

[3] 金谷健一, 「画像理解 — 3次元認識の数理 —」, 森北出版, 1990.

[4] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, 1993.

[5] 金谷健一, コンピュータビジョンのためのくりこみ法, 情報処理学会論文誌, **35**-2, 201–209 (1994).

[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] 金谷健一, "当てはめ問題の最適推定と精度の理論限界," 情報処理学会論文誌, **36**-8, 1865–1873 (1995).

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

[9] 金谷健一, 「形状ＣＡＤと図形の数学」, 共立出版, 1998.

[10] K. Kanatani and S. Takeda, 3-D motion analysis of a planar surface by renormalization, *IEICE Trans. Inf. & Syst.* **E78-D**-8 (1995), 1074–1079.

[11] 蚊野浩, 金出武雄, 任意のカメラ配置におけるステレオ視とステレオカメラ校正, 電子情報通信学会論文誌 D-II, **J79-D-II**-11 (1996), 1810–1818.

[12] H. C. Longuet-Higgins, The reconstruction of a plane surface from two perspective projections, *Proc. Roy. Soc. Lond.*, **B227** (1986), 399–410.

[13] 野口卓, 奥富正敏, ステレオ画像からの道路平面に対する射影変換行列の導出, 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, 97-CVIM-108-4 (1997), 23–30.

[14] 奥村晴彦, 「C 言語による最新アルゴリズム辞典」, 技術評論社, 1991.

[15] 武田佐知男, 金谷健一, くりこみ法による平面の3次元運動解析, 情報処理学会コンピュータビジョン研究会, 94-CV-90-6 (1994), 41–48.

[16] R. Y. Tsai and T. S. Huang, Estimating three-dimensional motion parameters of a rigid planar patch, *IEEE Trans. Acous. Speech Sig. Proc.*, **29**-6 (1981), 1147–1152.

[17] T. Viéville, C. Zeller and L. Robert, Using collineations to compute motion and structure in an uncalibrated image sequences, *Int. J. Comput. Vision*, **20**-3 (1996), 213–242.

[18] J. Weng, N. Ahuja and T. S. Huang, Motion and structure from point correspondences with error estimation: Planar surfaces, *IEEE Trans. Patt. Anal. Mach. Intell.*, **39**-12 (1991), 2691–2717.