基礎行列の最適計算とその信頼性評価

三島 等 * 金谷 健一 *

* 群馬大学工学部情報工学科

必ずしも等方一様でない独立な正規分布に従う誤差のもとで 2 組の対応点から基礎行列を最適に計算する手法を述べる。まずこの誤差モデルのもとでの精度の理論限界を導き、次にこれを達成するアルゴリズムを記述する。これはまずランク拘束を考慮せずにくりこみ法により基礎行列を計算し、次にこれがランク拘束を満たすように補正するものである。そして本手法の精度が理論限界を実際に達成していることを実験的に検証する。したがって、本手法は厳密に最適であり、もはや改良の余地はない。また、本アルゴリズムにより最適解が得られるだけでなく、その信頼性も同時に評価される。シミュレーションおよび実画像実験によりエピ極点の信頼性を評価する例を示す。

キーワード: 基礎行列、くりこみ法、エピ極点、信頼性評価、焦点距離、精度の理論限界

Optimal Computation of the Fundamental Matrix and Its Reliability Evaluation

Hitoshi Mishima* **and Kenichi Kanatani***

*Department of Computer Science
Gunma University, Kiryu, Gunma 376-8515 Japan

This paper presents an optimal algorithm for computing the fundamental matrix from two sets of corresponding points in the presence of independent Gaussian noise not necessarily isotropic or homogeneous. We first derive a theoretical accuracy bound and then present an algorithm that attains it. This algorithm first applies a technique called renormalization without considering the rank constraint and then corrects the solution to impose the constraint. We demonstrate by experiments that our algorithm indeed attains the accuracy bound. Hence, our algorithm is optimal in the strict sense: no further improvement is possible. Our algorithm produces not only an optimal estimate but also evaluates its reliability. We show simulated and real-image examples of evaluating the reliability of the computed epipoles.

Key words: fundamental matrix, renormalization, epipole, reliability evaluation, theoretical accuracy bound, renormalization

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)30-1841, Fax: (0277)30-1801
E-mail: mishima@ail.cs.gunma-u.ac.jp, kanatani@cs.gunma-u.ac.jp

1. 序論

2 画像間の「基礎行列」を計算することは画像から3次元復元を行うあらゆる問題の核心である [2, 16, 19, 20]。まず、基礎行列から「エピ極線」と「エピ極点」が計算できるから、一方の画像中の各点の他方の画像中で対応する位置がその点のエピ極線上に制限される [7, 8, 2, 11]。この性質はステレオ画像からの3次元復元の最も基礎となる事実である。また、基礎行列からその撮影カメラの知識と合わせてカメラの相対的な位置や物体の3次元構造を計算することもできる [3, 14, 15]。このためこれまで基礎行列を計算するさまざまなな方法が提案された [1, 4, 5, 13, 17, 18, 21, 22]。これらはすべて、誤差がなければ 0 となるある関数 (目的関数) を最小化するものであり、この最小化が「最適化」と呼ばれてきた。

しかし、そもそも真の解が求まらないのはデータに誤差があるためであるから、精度のよい解を得るにはその誤差の性質を明らかにしなければならない。誤差のモデルが与えられれば統計的解析によって「その誤差モデルのもとでの最適解」が定まる。ところがこれまでの研究では誤差モデルにほとんど触れず、新奇な目的関数を思い付くままに列挙し、シミュレーションおよび実画像実験で比較しているのみである [1, 4, 5, 13, 21, 22]。

この理由はあるゆる場で有効な「万能アルゴリズム」が唯一存在すると暗黙に仮定されているためと思われる。その要因の一つは特徴点の誤対応に起因する「アウトライア」である。そして、その統計的モデル化が困難であるという理由で誤差のモデル化自体が排除されてきた。その結果、アウトライアのない場合に精密なアルゴリズムはアウトライアがあれば有効でないという理由で排除され、精度は悪いがアウトライアに影響されにくいアルゴリズムが重視されてきた [17, 18]。しかし、アウトライアの検出とアウトライアのない場合の高精度アルゴリズムとは別問題である。アウトライア検出に有効な手法によってアウトライアを除去した後でアウトライアのない場合の高精度手法を適用すべきである。

また、従来の論文ではどれも種々の手法の実験結果を互いに比較しているのみである。しかし、手法の最適性は精度の理論限界と比較して初めて実証される。これを達成していなければ改良の余地があるし、達成していればもはや改良の余地がない。したがって理論限界を導き、これと比較することが不可欠である。

本論文では「各特徴点の誤差が微小かつランダムであり互いに独立で期待値 0 の正規分布に従う」というモデルのもとでの基礎行列の最適計算アルゴリズムを述べる。ただし誤差分布は最も一般的なものとし、「一様性」(場所によらない) も「等方性」(方向によらない) も仮定しない。まず、このモデルのもとでの精度の理論限界を示す。次にこれを達成するアルゴリズムを記述し、実際に限界を達成していることを実験的に示す。したがって、本アルゴリズム厳密な意味で「最適」であり、仮定した誤差モデルのもとではもはや改良の余地はない。

本アルゴリズムは単に最適解を計算するのみならず、計算した解の信頼性も同時に評価する。したがって、計算した基礎行列から3次元復元を行えば復元した3次元形状の信頼性も評価することができる。

カメラがあらかじめキャリブレーションされ、焦点距離が既知の透視変換と仮定できれば、基礎行列は「基本行列」と呼ばれ、より複雑な拘束条件 (「分解可能条件」[6, 8]) が課される。この場合の最適推定は既になされ、それによる3次元復元の信頼性も評価されている [10, 11]。本論文のアルゴリズムはこれから分解可能条件を取り除いたものである。

2. 基礎行列

任意の画像座標系により表された N 組の対応点 $\{(x_\alpha, y_\alpha)\}$, $\{(x'_\alpha, y'_\alpha)\}$, $\alpha = 1, ..., N$(単位は画素) を次のベクトルで表す。

$$\boldsymbol{x}_\alpha = \begin{pmatrix} x_\alpha/f_0 \\ y_\alpha/f_0 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}'_\alpha = \begin{pmatrix} x'_\alpha/f_0 \\ y'_\alpha/f_0 \\ 1 \end{pmatrix} \tag{1}$$

f_0 はスケール因子 (単位は画素) であり、x_α/f_0, y_α/f_0, x'_α/f_0, y'_α/f_0 が $O(1)$ となるように選ぶ [1]。このとき、ある行列式 0 の行列 $\boldsymbol{F}$ が存在して

$$(\boldsymbol{x}_\alpha, \boldsymbol{F}\boldsymbol{x}'_\alpha) = 0 \tag{2}$$

となっているとき、二つの点集合は**エピ極線拘束条件**を満たすといい、式 (2) を**エピ極線方程式**と呼ぶ。そして $\boldsymbol{F}$ を**基礎行列**と呼ぶ。ただし、本論文ではベクトル $\boldsymbol{a}$, $\boldsymbol{b}$ の内積を $(\boldsymbol{a}, \boldsymbol{b})$ と表す。式 (2) は対応点がシーン中の静止した特徴点を 2 台のカメラで観測したものであるための必要十分条件である [2, 16, 19, 20]。式 (2) より $\boldsymbol{F}$ はスケールの不定性があるので $\|\boldsymbol{F}\| =$

[1]例えば画像サイズを用れればよい。カメラの焦点距離 (単位は画素) の近似値が既知なら、それを用いてもよい。

1と正規化する。ただし、行列 $\boldsymbol{A} = (A_{ij})$ のノルムを $\|\boldsymbol{A}\| = \sqrt{\sum_{i,j=1}^{3} A_{ij}^2}$ と定義する。

3. 誤差モデル

ベクトル $\boldsymbol{x}_\alpha$, $\boldsymbol{x}'_\alpha$ の真の値を $\bar{\boldsymbol{x}}_\alpha$, $\bar{\boldsymbol{x}}'_\alpha$ とし、次のように書く。

$$\boldsymbol{x}_\alpha = \bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha, \qquad \boldsymbol{x}'_\alpha = \bar{\boldsymbol{x}}'_\alpha + \Delta\boldsymbol{x}'_\alpha \tag{3}$$

$\Delta\boldsymbol{x}_\alpha$, $\Delta\boldsymbol{x}'_\alpha$ を期待値0の正規分布に従う独立な確率変数とみなし、その共分散行列をそれぞれ $V[\boldsymbol{x}_\alpha]$, $V[\boldsymbol{x}'_\alpha]$ とする。ただし、これらの絶対的な大きさを知る必要はなく、その相対的な傾向のみ既知とする。そして

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \qquad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha] \tag{4}$$

と書き、未知の定数 ϵ をノイズレベルと呼ぶ。$V_0[\boldsymbol{x}_\alpha]$, $V_0[\boldsymbol{x}'_\alpha]$ は誤差の場所や方向への相対的な依存を表し、**正規化共分散行列**と呼ぶ。ベクトル $\boldsymbol{x}_\alpha$, $\boldsymbol{x}'_\alpha$ は第3成分が1であるため $V_0[\boldsymbol{x}_\alpha]$, $V_0[\boldsymbol{x}'_\alpha]$ は第3行第3列が0の特異行列である。誤差の出方に特に傾向がない場合はデフォルト値 $\mathrm{diag}(1,1,0)$ を用いる。ただし $\mathrm{diag}(\cdots)$ は $\cdots$ を対角要素とする対角行列を表す。

問題は $\{\bar{\boldsymbol{x}}_\alpha, \bar{\boldsymbol{x}}'_\alpha\}$ がエピ極線方程式 (2) を満たすとき、誤差のある $\{\boldsymbol{x}_\alpha, \boldsymbol{x}'_\alpha\}$ から基礎行列 $\boldsymbol{F}$ を推定することである。

4. 精度の理論限界

基礎行列 $\boldsymbol{F}$ のある推定値を $\hat{\boldsymbol{F}}$ とし、その真の値を $\bar{\boldsymbol{F}}$ とする。この**共分散テンソル** $\mathcal{V}[\hat{\boldsymbol{F}}]$ を次のように定義する。

$$\mathcal{V}[\hat{\boldsymbol{F}}] = E[\mathcal{P}\Big((\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}}) \otimes (\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}})\Big)\mathcal{P}^\top] \tag{5}$$

右辺の $E[\,\cdot\,]$ は期待値を表す。演算 $\otimes$ はテンソル積であり、行列 $\boldsymbol{A} = (A_{ij})$, $\boldsymbol{B} = (B_{ij})$ に対して $\boldsymbol{A} \otimes \boldsymbol{B}$ は $(ijkl)$ 要素が $A_{ij}B_{kl}$ のテンソルである。テンソル $\mathcal{P} = (P_{ijkl})$, $\mathcal{T} = (T_{ijkl})$ に対して $\mathcal{P}\mathcal{T}\mathcal{P}^\top$ は $(ijkl)$ 要素が $\sum_{m,n,p,q=1}^{3} P_{ijmn}P_{klpq}T_{mnpq}$ のテンソルである。式 (5) 中の $\mathcal{P} = (P_{ijkl})$ は次のように定義された射影テンソルである。

$$P_{ijkl} = \delta_{ik}\delta_{jl} - \bar{F}_{ij}\bar{F}_{kl} \tag{6}$$

ただし δ_{ij} はクロネッカデルタであり、$i = j$ のとき1、そうでないとき0をとる。**モーメントテンソル** $\bar{\mathcal{M}} = (\bar{M}_{ijkl})$ を次のように定義する。

$$\bar{\mathcal{M}} = \frac{1}{N}\sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{x}}_\alpha \otimes \bar{\boldsymbol{x}}'_\alpha \otimes \bar{\boldsymbol{x}}_\alpha \otimes \bar{\boldsymbol{x}}'_\alpha, \tag{7}$$

$$\bar{W}_\alpha = \frac{1}{(\bar{\boldsymbol{x}}'_\alpha, \bar{\boldsymbol{F}}^\top V_0[\boldsymbol{x}_\alpha]\bar{\boldsymbol{F}}\bar{\boldsymbol{x}}'_\alpha) + (\bar{\boldsymbol{x}}_\alpha, \bar{\boldsymbol{F}}V_0[\boldsymbol{x}'_\alpha]\bar{\boldsymbol{F}}^\top\bar{\boldsymbol{x}}_\alpha)} \tag{8}$$

このとき統計的最適化理論 [11] から共分散テンソル $\mathcal{V}[\hat{\boldsymbol{F}}]$ の下界が次のように得られる。

$$\mathcal{V}[\hat{\boldsymbol{F}}] \succ \frac{\epsilon^2}{N}\Big(\mathcal{P}^{\mathcal{S}}\bar{\mathcal{M}}\mathcal{P}^{\mathcal{S}\top}\Big)^-_7 \tag{9}$$

ただし、テンソル $\mathcal{T}$, $\mathcal{S}$ に対して $\mathcal{T} \succ \mathcal{S}$ は $\mathcal{T} - \mathcal{S}$ が半正値テンソル (固有値がすべて非負のテンソル) であることを意味する。$\mathcal{P}^{\mathcal{S}} = (P^{\mathcal{S}}_{ijkl})$ は次のように定義された射影テンソルである。

$$P^{\mathcal{S}}_{ijkl} = \delta_{ik}\delta_{jl} - \frac{\bar{F}^\dagger_{ji}\bar{F}^\dagger_{lk}}{\|\bar{\boldsymbol{F}}^\dagger\|^2} \tag{10}$$

$\bar{\boldsymbol{F}}^\dagger$ は行列 $\bar{\boldsymbol{F}}$ の余因子行列である。テンソル $\mathcal{T}$ と行列 $\boldsymbol{A}$ に対して $\mathcal{T}\boldsymbol{A} = \lambda\boldsymbol{A}$ となるとき、行列 $\boldsymbol{A}$ はテンソル $\mathcal{T}$ の固有値 λ の**固有行列**であるという。テンソル $\mathcal{T} = (T_{ijkl})$ と行列 $\boldsymbol{A} = (A_{ij})$ の積 $\mathcal{T}\boldsymbol{A}$ は (ij) 要素が $\sum_{k,l=1}^{3} T_{ijkl}A_{kl}$ の行列である。テンソル $\mathcal{T}$ の固有値と固有行列は $\mathcal{T}$ を 9×9 行列と同一視して固有値を計算し、得られる9次元固有ベクトルを 3×3 行列と同一視すればよい [11]。$(\cdot)^-_r$ はランクを r に拘束した**一般逆テンソル**である。テンソル $\mathcal{T}$ の固有値を $\lambda_1 \geq \cdots \geq \lambda_9$、対応するノルム1の固有行列を $\boldsymbol{U}_1$, ..., $\boldsymbol{U}_9$ とする。$\lambda_r > 0$ となる r に対して $\mathcal{T}^-_r$ は次のように計算される。

$$\mathcal{T}^-_r = \sum_{i=1}^{r} \frac{\boldsymbol{U}_i \otimes \boldsymbol{U}_i}{\lambda_i} \tag{11}$$

推定値 $\hat{\boldsymbol{F}}$ の**平方平均二乗誤差** (root-mean-square error) を次のように定義する。

$$\mathrm{rms}[\hat{\boldsymbol{F}}] = \sqrt{E[\|\mathcal{P}(\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}})\|^2]} \tag{12}$$

射影テンソル $\mathcal{P}$ の定義より、取り得る値の範囲は $0 \leq \mathrm{rms}[\hat{\boldsymbol{F}}] \leq 1$ である。式 (9) から次の理論的下界が得られる。

$$\mathrm{rms}[\hat{\boldsymbol{F}}] \geq \sqrt{\mathrm{tr}\mathcal{V}[\hat{\boldsymbol{F}}]} \tag{13}$$

ただし、テンソル $\mathcal{T} = (T_{ijkl})$ の**トレース** $\mathrm{tr}\mathcal{T}$ を次のように定義する。

$$\mathrm{tr}\mathcal{T} = \sum_{k,l=1}^{3} T_{klkl} \tag{14}$$

注: $\boldsymbol{F}$ はその 9 個の要素を座標とする 9 次元空間 $\mathcal{R}^9$ の 1 点とみなせる。この空間の中で $\det \boldsymbol{F} = 0$ となる点の集合は 8 次元代数多様体 $\mathcal{S}$ となる。そして、$\mathcal{S}$ 上の点 $\bar{\boldsymbol{F}}$ における接空間 $T_{\bar{\mathbf{F}}}(\mathcal{S})$ への射影テンソル $\mathcal{P}^{\mathcal{S}} = (P^{\mathcal{S}}_{ijkl})$ が式 (10) で与えられる。一方、正規化 $\|\boldsymbol{F}\| = 1$ より $\boldsymbol{F}$ はその空間中の原点を中心とし、半径 1 の 8 次元球面 S^8 上にある。式 (6) は S^8 上の点 $\bar{\boldsymbol{F}}$ における接空間 $T_{\bar{\mathbf{F}}}(S^8)$ への射影テンソルに他ならない。したがって $\boldsymbol{F}$ は S^8 と $\mathcal{S}$ の交わりに拘束されている。そして、$T_{\bar{\mathbf{F}}}(\mathcal{S})$ は $\bar{\boldsymbol{F}}$ において $\bar{\boldsymbol{F}}^{\dagger\top}$ と直交する線形空間であり、$T_{\bar{\mathbf{F}}}(S^8)$ は $\bar{\boldsymbol{F}}$ において $\bar{\boldsymbol{F}}$ と直交する線形空間である。ただし行列 $\boldsymbol{A} = (A_{ij})$, $\boldsymbol{B} = (B_{ij})$ の内積を $(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i=1}^{3} A_{ij}B_{ij}$ と定義する。さらに $\bar{\boldsymbol{F}}$ と $\bar{\boldsymbol{F}}^{\dagger\top}$ が直交することが次のように示せる。

$$(\bar{\boldsymbol{F}}^{\dagger\top}; \bar{\boldsymbol{F}}) = \mathrm{tr}(\bar{\boldsymbol{F}}^{\dagger}\bar{\boldsymbol{F}}) = (\det \bar{\boldsymbol{F}})\mathrm{tr}\boldsymbol{I} = 0. \qquad (15)$$

この結果、射影テンソル $\mathcal{P}^{\mathcal{S}}$ は接空間 $T_{\bar{\mathbf{F}}}(S^8)$ をそれ自身の中に射影する。定義よりモーメントテンソル $\bar{\mathcal{M}}$ の定義域は接空間 $T_{\bar{\mathbf{F}}}(S^8)$ であるが式 (9) の右辺の定義域は $T_{\bar{\mathbf{F}}}(\mathcal{S}) \cap T_{\bar{\mathbf{F}}}(S^8)$ である。これは $\bar{\boldsymbol{F}}$ における 7 次元空間であり、その零空間は $\bar{\boldsymbol{F}}$ と $\bar{\boldsymbol{F}}^{\dagger\top}$ の張る 2 次元空間である。

5. アルゴリズムの概要

以下に述べるアルゴリズムは2段階から成る。第1段階では $\det \boldsymbol{F} = 0$ の制約を考えずにくりこみ法 [9, 10, 11] と呼ぶ手法を用いて $\boldsymbol{F}$ を最適に計算する。これは最小二乗法による解を計算してはそれに含まれる統計的偏差を逐次的に除去するものである。第2段階では第1段階で得られる $\boldsymbol{F}$ の共分散テンソル $\mathcal{V}[\hat{\boldsymbol{F}}]$ に基づいて $\det \boldsymbol{F} = 0$ となるように最適に補正する。これは $\mathcal{V}[\hat{\boldsymbol{F}}]$ に関するマハラノビス距離を最小化するものである。

5.1 くりこみ法

1. $c = 0$, $W_\alpha = 1$, $\alpha = 1, ..., N$, $J = \infty$ と置く。% 実際の計算では ∞ は十分大きい数とする (例: 10^{10})。
2. 次のようにテンソル $\mathcal{M} = (M_{ijkl}), \mathcal{N} = (N_{ijkl})$ を計算する。

$$M_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha x_{\alpha(i)} x'_{\alpha(j)} x_{\alpha(k)} x'_{\alpha(l)},$$

$$N_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha (V_0[\boldsymbol{x}_\alpha]_{ik} x'_{\alpha(j)} x'_{\alpha(l)} + V_0[\boldsymbol{x}'_\alpha]_{jl} x_{\alpha(i)} x_{\alpha(k)}) \qquad (16)$$

% $x_{\alpha(i)}$, $x'_{\alpha(i)}$ はそれぞれベクトル $\boldsymbol{x}_\alpha$ の第 i 成分であり、$V_0[\boldsymbol{x}_\alpha]_{ij}$, $V_0[\boldsymbol{x}'_\alpha]_{ij}$ はそれぞれ $V_0[\boldsymbol{x}_\alpha]$, $V_0[\boldsymbol{x}'_\alpha]$ の (ij) 要素である。

3. テンソル $\mathcal{M}$ 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 $\{\boldsymbol{F}_1$, ..., $\boldsymbol{F}_9\}$ を計算する。
4. 次の計算を行う。

- c を次のように更新する。

$$c \leftarrow c + \frac{\lambda_9}{(\boldsymbol{F}_9; \mathcal{N}\boldsymbol{F}_9)} \qquad (17)$$

- W_α, $\alpha = 1, ..., N$ を次のように計算する。

$$W_\alpha = 1\Big/\Big((\boldsymbol{x}'_\alpha, \boldsymbol{F}_9^\top V_0[\boldsymbol{x}_\alpha]\boldsymbol{F}_9\boldsymbol{x}'_\alpha) + (\boldsymbol{x}_\alpha, \boldsymbol{F}_9 V_0[\boldsymbol{x}'_\alpha]\boldsymbol{F}_9^\top \boldsymbol{x}_\alpha)\Big) \qquad (18)$$

- 式 (16) によってテンソル $\mathcal{M}$, $\mathcal{N}$ を計算する。
- 次の計算を行う。

$$J' \leftarrow J, \quad J \leftarrow (\boldsymbol{F}_9, \mathcal{M}\boldsymbol{F}_9) \qquad (19)$$

- $J' < J$ であれば $J \leftarrow J'$ とする。
- そうでなければ次のテンソルの 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 $\{\boldsymbol{F}_1, ..., \boldsymbol{F}_9\}$ を計算する。

$$\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N} \qquad (20)$$

5. これを $J' \leq J$ または $|\lambda_9| \approx 0$ となるまで反復する。% これによって収束が保証される。

5.2 最適補正

1. くりこみ法で得られた行列 $\boldsymbol{F}_9$ を $\boldsymbol{F}$ とする。
2. 二乗ノイズレベルの推定値 $\hat{\epsilon}^2$ を次のように計算する。

$$\hat{\epsilon}^2 = \frac{J}{1 - 8/N} \qquad (21)$$

3. $\boldsymbol{F}$ の正規化共分散テンソルを次のように計算する。

$$\mathcal{V}_0[\boldsymbol{F}] = \frac{1}{N}\sum_{i=1}^{8} \frac{\boldsymbol{F}_i \otimes \boldsymbol{F}_i}{\lambda_i} \qquad (22)$$

4. $\det \boldsymbol{F} \approx 0$ となるまで次の計算を反復する。

- 行列 $\boldsymbol{F}$ を次のように更新する。

$$\boldsymbol{F} \leftarrow N[\boldsymbol{F} - \frac{(\det \boldsymbol{F})\mathcal{V}_0[\boldsymbol{F}]\boldsymbol{F}^{\dagger\top}}{(\boldsymbol{F}^{\dagger\top}; \mathcal{V}_0[\boldsymbol{F}]\boldsymbol{F}^{\dagger\top})}] \tag{23}$$

% $N[\cdot]$ はノルムを1とする正規化作用素である。

- 射影テンソル $\mathcal{P} = (P_{ijkl})$ を次のように定義する。

$$P_{ijkl} = \delta_{ik}\delta_{jl} - F_{ij}F_{kl} \tag{24}$$

- 正規化共分散テンソル $\mathcal{V}_0[\boldsymbol{F}]$ を次のように更新する。

$$\mathcal{V}_0[\boldsymbol{F}]_{ijkl} \leftarrow \sum_{m,n,p,q=1}^{3} P_{ijmn}P_{klpq}\mathcal{V}_0[\boldsymbol{F}]_{mnpq} \tag{25}$$

% $\boldsymbol{F}$ の誤差は $\boldsymbol{F}$ に直交するので、$\boldsymbol{F}$ が変われば誤差の定義域も変わる。上式は誤差の挙動を新しい接空間 $T_{\mathbf{F}}(S^8)$ に射影するものである。

6. エピ極点とエピ極線

$\boldsymbol{F}^\top$ の固有値 0 に対する単位固有ベクトルを $\boldsymbol{u}$ とする。第1画像中の**エピ極点**とは次のように表される点である。

$$\boldsymbol{x}_e = Z[\boldsymbol{u}] \tag{26}$$

ただし $Z[\cdot]$ は第3成分を1とする正規化作用素である。このエピ極点は第2画像を撮影したカメラのレンズ中心の画像上の位置である。ただし焦点距離が与えられない限り、その物理的な方向を知ることはできない。

第1画像中の**エピ極線**とは次のように定義される直線である。

$$(\boldsymbol{n}_\alpha, \boldsymbol{x}) = 0, \quad \boldsymbol{n}_\alpha = N[\boldsymbol{F}\boldsymbol{x}'_\alpha] \tag{27}$$

$\boldsymbol{n}$ はこの直線の**Nベクトル**と呼ばれる [7, 8]。すべてのエピ極線はエピ極点 $\boldsymbol{x}_e$ で交わる。

第2画像についても同様である。$\boldsymbol{F}$ の固有値 0 に対する単位固有ベクトルを $\boldsymbol{v}$ とすると、エピ極点が次のように定義される。

$$\boldsymbol{x}'_e = Z[\boldsymbol{v}] \tag{28}$$

これは第1画像を撮影したカメラのレンズ中心の画像上の位置であるが、やはり焦点距離が与えられない限り、その物理的な方向を知ることはできない。エピ極線は次のように定義される。

$$(\boldsymbol{n}'_\alpha, \boldsymbol{x}') = 0, \quad \boldsymbol{n}'_\alpha = N[\boldsymbol{F}^\top\boldsymbol{x}_\alpha] \tag{29}$$

すべてのエピ極線はエピ極点 $\boldsymbol{x}'_e$ で交わる。

データに誤差があるとき、エピ極点 $\boldsymbol{x}_e$ の共分散行列は固有値問題の**摂動定理** [11] を用いて次のように導ける。

$$V[\boldsymbol{x}_e] = \frac{\boldsymbol{Q}_{\bar{\mathbf{x}}_e}V[\boldsymbol{u}]\boldsymbol{Q}_{\bar{\mathbf{x}}_e}^\top}{(\boldsymbol{k}, \bar{\boldsymbol{u}})^2} \tag{30}$$

バーは真の値を意味し、射影行列 $\boldsymbol{Q}$ を次のように定義する。

$$\boldsymbol{Q}_{\mathbf{x}} = \boldsymbol{I} - \boldsymbol{x}\boldsymbol{k}^\top \tag{31}$$

式 (30) 中のベクトル $\boldsymbol{u}$ の共分散行列 $V[\boldsymbol{u}]$ は次のように与えられる。

$$V[\boldsymbol{u}] = \epsilon^2(\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^\top)_2^- \boldsymbol{G}(\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^\top)_2^- \tag{32}$$

ただし、行列 $\boldsymbol{G} = (G_{ij})$ を次のように置く。

$$G_{ij} = \sum_{j,k=1}^{3} E_{ijkl}\bar{u}_k\bar{u}_l \tag{33}$$

テンソル $\mathcal{E} = (E_{ijkl})$ は次のように定義される。

$$E_{ijkl} = \sum_{m,n=1}^{3}\Big(\mathcal{V}_0[\boldsymbol{F}]_{imjn}\bar{F}_{km}\bar{F}_{ln} + \mathcal{V}_0[\boldsymbol{F}]_{imln}\bar{F}_{km}\bar{F}_{jn} + \mathcal{V}_0[\boldsymbol{F}]_{kmjn}\bar{F}_{im}\bar{F}_{ln} + \mathcal{V}_0[\boldsymbol{F}]_{kmln}\bar{F}_{im}\bar{F}_{jn}\Big) \tag{34}$$

同様に、エピ極点 $\boldsymbol{x}'_e$ の共分散行列は次のように書ける。

$$V[\boldsymbol{x}'_e] = \frac{\boldsymbol{Q}_{\bar{\mathbf{x}}'_e}V[\boldsymbol{v}]\boldsymbol{Q}_{\bar{\mathbf{x}}'_e}^\top}{(\boldsymbol{k}, \bar{\boldsymbol{v}})^2}, \tag{35}$$

$$V[\boldsymbol{v}] = \epsilon^2(\bar{\boldsymbol{F}}^\top\bar{\boldsymbol{F}})_2^- \boldsymbol{G}'(\bar{\boldsymbol{F}}^\top\bar{\boldsymbol{F}})_2^- \tag{36}$$

ただし、行列 $\boldsymbol{G}' = (G'_{ij})$ を次のように置く。

$$G'_{ij} = \sum_{j,k=1}^{3} E'_{ijkl}\bar{u}_k\bar{u}_l, \tag{37}$$

$$E'_{ijkl} = \sum_{m,n=1}^{3}\Big(\mathcal{V}_0[\boldsymbol{F}]_{minj}\bar{F}_{mk}\bar{F}_{nl} + \mathcal{V}_0[\boldsymbol{F}]_{minl}\bar{F}_{mk}\bar{F}_{nj} + \mathcal{V}_0[\boldsymbol{F}]_{mknj}\bar{F}_{mi}\bar{F}_{nl} + \mathcal{V}_0[\boldsymbol{F}]_{nknl}\bar{F}_{mi}\bar{F}_{nj}\Big) \tag{38}$$

7. シミュレーション実験

図1は移動するカメラで焦点距離を変えながら撮影した格子パタンからなる環境モデルのシミュレーショ

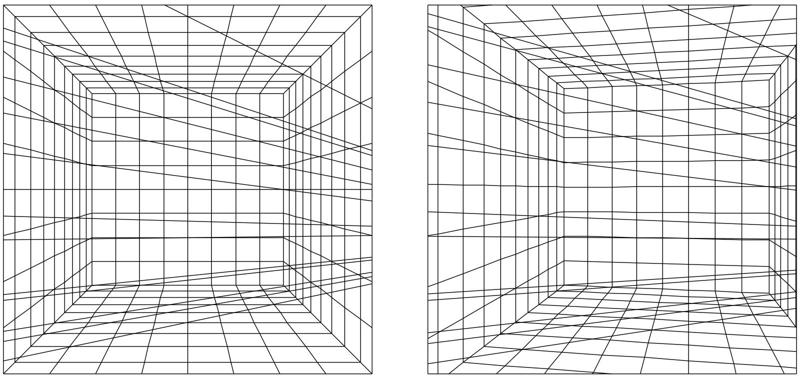

図 1: 3次元環境のシミュレーション画像とそのエピ極線。

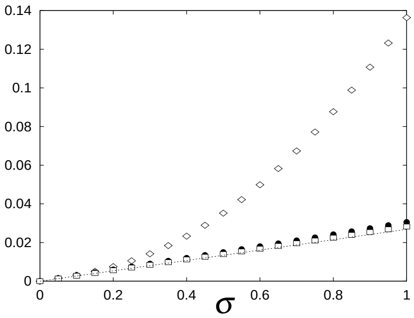

図 2: 計算した基礎行列の平方平均二乗誤差。□: 本手法、●: 最適補正を行わない場合、◇: 最小二乗法。破線は理論的下界。

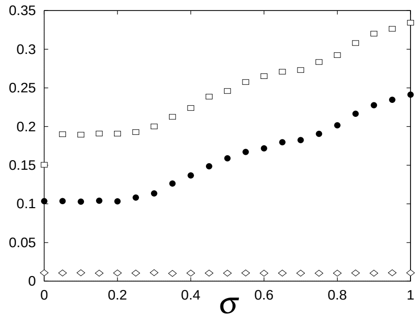

図 3: 平均実行時間 (秒)。□: 本手法、●: 最適補正を行わない場合、◇: 最小二乗法。

ン画像である (512×512 画素)。いくつかのエピ極線も重ねてある。この二つの画像中の格子点の x, y 座標に独立に期待値 0、標準偏差 σ(画素) の正規乱数を独立に加え、それを対応点として基礎行列を計算した。5.1 節のくりこみ法は収束が保証されており、通常は 3, 4 回の反復で収束する。

図 2は横軸を σ とし、各 σ に対して異なる誤差を加えて 100 回計算して、平方平均二乗誤差

$$\sqrt{\frac{1}{100}\sum_{a=1}^{100}\|\mathcal{P}(\hat{\boldsymbol{F}}^{a}-\bar{\boldsymbol{F}})\|^2} \tag{39}$$

をプロットしたものである ($\hat{\boldsymbol{F}}^{a}$ は a 回目の推定値、$\bar{\boldsymbol{F}}$ は真の値、$\mathcal{P}$ は式 (6) の射影テンソル)。□ が本手法によるものであり、破線は式 (13) の理論的下界である。● は 5.1 節のくりこみ法のみを行い、5.2 節の最適補正を行わない場合である。また ◇ は**最小二乗法 (代数的距離最小法)**[4, 5] によるものである。これは式 (2) から直接に

$$\frac{1}{N}\sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha, \boldsymbol{F}\boldsymbol{x}'_\alpha)^2 \to \min \tag{40}$$

とするものであり、5.1 節のくりこみ法のステップ 4 の反復を行わないものに相当している。

図からわかるように、本手法は精度の理論限界に到達しており、もはや改良の余地がない。これは理論解析からも示せるが [11]、実験的にも確認された。図 3は図 2の計算のワークステーション Sun Ultra-30 (SunOS 5.6) 上での平均実行時間である。当然ながら、本手法は最小二乗法に比べて時間がかかるが、この程度の計算時間で精度を限界まで高めることができるともいえる。

図 4(a) は図 1の第1画像のエピ極点付近を拡大したものであり、これに $\sigma = 0.5$(画素) として誤差をいろいろに変えて計算したエピ極点を 100 点プロットしている。楕円は真のエピ極点 $\bar{\boldsymbol{x}}_e$ を中心とし、式 (30) の共分散行列 $V[\boldsymbol{x}_e]$ を表すものである (各方向の標準偏差を示す)。再び本手法が精度の理論限界に到達していることがわかる。図 4(b) は最小二乗法による結果である。この場合は計算した基礎行列 $\boldsymbol{F}$ の行列式が必ずしも 0 でないから厳密に $\boldsymbol{F}^\top\boldsymbol{u} = \boldsymbol{0}$ となるベクトル $\boldsymbol{u}$ が存在しない。そこで $\|\boldsymbol{F}^\top\boldsymbol{u}\|^2 \to \min$ となる単位ベクトル $\boldsymbol{u}$ を計算する ($\boldsymbol{F}\boldsymbol{F}^\top$ の最小固有値に対応する単位固有ベクトルに等しい)。図からエピ極点の

分散が大きいのみらならず、統計的偏差が存在することがわかる。

図 5(a),(b) は第2画像について同様に計算した結果である。やはり最小二乗法については $\|\boldsymbol{F}\boldsymbol{v}\|^2 \to \min$ となる単位ベクトル $\boldsymbol{v}$ ($\boldsymbol{F}^\top\boldsymbol{F}$ の最小固有値に対応する単位固有ベクトル) を計算する。この場合も本手法が最適であり、最小二乗解には統計的偏差が存在することがわかる。

図 6は横軸を σ とし、各 σ に対して異なる誤差を加えて 100 回計算して、エピ極点 $\boldsymbol{x}_e$ の平方平均二乗誤差

$$\sqrt{\frac{1}{100}\sum_{a=1}^{100}\|\hat{\boldsymbol{x}}_e^a - \bar{\boldsymbol{x}}_e\|^2} \tag{41}$$

をプロットしたものである ($\hat{\boldsymbol{x}}_e^a$ は a 回目の推定値)。□ が本手法、◇ が最小二乗法によるものであり、破線は理論的下界 $\sqrt{\mathrm{tr}V[\boldsymbol{x}_e]}$ である。図 7はエピ極点 $\boldsymbol{x}'_e$ に対して同様に計算した結果である。

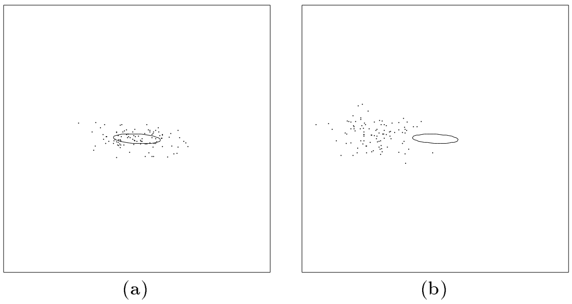

図 4: 図 1 の第1画像のエピ極点の信頼性。(a) 本手法、(b) 最小二乗法。

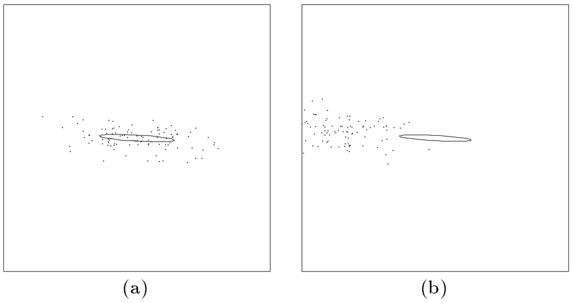

図 5: 図 1 の第2画像のエピ極点の信頼性。(a) 本手法、(b) 最小二乗法。

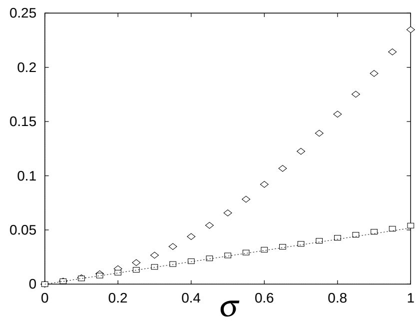

図 6: 第1画像のエピ極点の平方平均二乗誤差。□: 本手法、◇: 最小二乗法。破線は理論的下界。

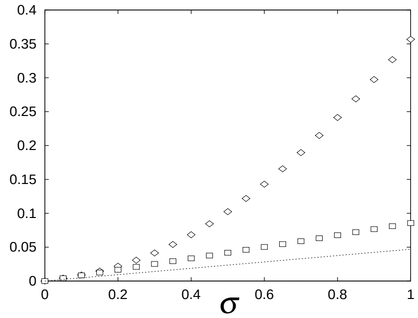

図 7: 第2画像のエピ極点の平方平均二乗誤差。□: 本手法、◇: 最小二乗法。破線は理論的下界。

8. 実画像実験

図 8は移動するカメラで焦点距離を変えながら撮影した室内シーンの実画像である (768×512 画素)。これから図中にマークした特徴点を選んで基礎行列を計算し、それから得られたエピ極線を図示している。エピ極点の信頼性 (式 (30), (35) にデータと推定値を代入して計算した $\sqrt{\mathrm{tr}V[\boldsymbol{x}_e]}$, $\sqrt{\mathrm{tr}V[\boldsymbol{x}'_e]}$) は運動前後でそれぞれ 3.44 画素、1.74 画素である。このように、本手法では最適解が得られるだけでなく、計算した解の信頼性の評価が可能となる。これを用いると、推定した基礎行列から復元した3次元形状の信頼性も計算できる [10, 11]。

9. 公開プログラム

本論文に述べたアルゴリズムは C++ により公開プログラムが作成されている [2]。これは最適に計算した基礎行列 $\hat{\boldsymbol{F}}$ だけでなく、その標準偏位 $\boldsymbol{F}^{(+)}$, $\boldsymbol{F}^{(-)}$ も出力する。標準偏位はパラメータ空間で $\hat{\boldsymbol{F}}$ から最も誤差が生じやすい両方向に標準偏差だけずれた値を示すものである。これは式 (9) の右辺にデータと推定値 $\hat{\boldsymbol{F}}$ を代入したテンソルの最大固有値を $\lambda_{\max}$、対応するノルム 1 の固有行列を $\boldsymbol{U}_{\max}$ とするとき、次のように定義される。

$$\boldsymbol{F}^{(+)} = N[\hat{\boldsymbol{F}} + \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}]$$

[2] http://www.ail.cs.gunma-u.ac.jp/Labo/research.html

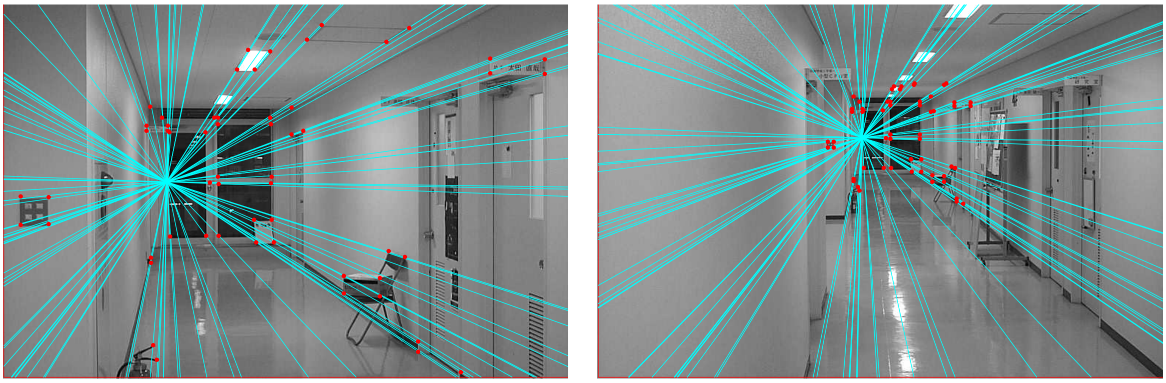

図 8: 実画像と計算したエピ極線。

$$\boldsymbol{F}^{(-)} = N[\hat{\boldsymbol{F}} - \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}] \qquad (42)$$

例えば $\boldsymbol{F}^{(+)}$ と $\boldsymbol{F}^{(+)}$ が有効数字 3 桁で一致していれば、解 $\hat{\boldsymbol{F}}$ にほぼ有効数字 3 桁の精度があることが保証される。このように、解の信頼性が定量的に評価されるのが特徴である。

よく知られているように、特徴点が「退化」と呼ばれる特殊な配置 [3] にあれば、原理的に基礎行列が一意的に定まらない。そこで、計算の過程で式 (42) 中の $\lambda_{\max}$ が 1 程度になると予測されると退化と判定し、警告メッセージを出して計算を中止する。

10. まとめ

本論文では必ずしも等方一様でない独立な正規分布に従う誤差のもとで 2 組の対応点から基礎行列を最適に計算する手法を述べた。まず精度の理論限界を導き、次にこれを達成するアルゴリズムを記述した。これはまずランク拘束を考慮せずにくりこみ法により基礎行列を計算し、次にランク拘束を満たすように補正するものである。そして本手法の精度が理論的限界を実際に達成していることを実験的に検証した。したがって、本手法は厳密に最適であり、もはや改良の余地はない。また、本アルゴリズムにより最適解が得られるだけでなく、その信頼性も同時に評価される。シミュレーションおよび実画像実験によりエピ極点の信頼性を評価する例を示した。

参考文献

[1] M. Bober, N. Geogis and J. Kittler, On accurate and robust estimation of fundamental matrix, *Comput. Vision Image Understanding*, **72**-1 (1998), 39–53.

[2] O. D. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA, U.S.A., 1993.

[3] 並進が零に近い場合や物体が平面または「臨界曲面」と呼ばれる特殊な2次曲面の場合に生じる [8]。

[3] R. I. Hartley, Estimation of relative camera position for uncalibrated cameras, *Proc. 2nd Euro. Conf. Comput. Vision*, May 1992, Santa Margherita Ligure, Italy, pp. 579–587.

[4] R. I. Hartley, In defense of the eight-point algorithm, *IEEE Trans. Patt. Anal. Mach. Intell.*, **97**-6 (1997), 580–593.

[5] R. I. Hartley, Minimizing algebraic error, *Phil. Trans. R. Soc. Lond.*, A, **3569** (1998), 1175–1192.

[6] T. S. Huang and O. D. Faugeras, Some properties of the *E* matrix in two-view motion estimation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **11**-12 (1989), 1310–1312.

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

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

[9] 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 (1994), 201–209.

[10] K. Kanatani, Renormalization for motion analysis: Statistically optimal algorithm, *IEICE Trans. Inf. & Syst.*, **E77-D**-11 (1994), 1233–1239.

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

[12] 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.

[13] Q.-T. Luong and O. D. Faugeras, The fundamental matrix: Theory, algorithm, and stability analysis, *Int. J. Comput. Vision*, **17**-3 (1996), 43–7.

[14] Q.-T. Luong and O. D. Faugeras, Self-calibration of a moving camera from point correspondences and fundamental matrices, *Int. J. Comput. Vision*, **23**-3 (1997), 261–289.

[15] G. N. Newsam, D. Q. Huynh, M. J. Brooks and H.-P. Pan, Recovering unknown focal lengths in self-calibration: An essentially linear algorithm and degenerate configurations, *Int. Arch. Photogram. Remote Sensing*, **31**-B3-III, July 1996, Vienna, Austria, pp. 575–580.

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

[17] P. H. S. Torr and A. Zissermann, Performance characterization of fundmental matrix estimation under image degradation, *Mach. Vision Appl.*, **9** (1997), 321–333.

[18] P. H. S. Torr and Z. Zisserman, Robust detection of degenerate configurations while estimating the fundamental matrix, *Comput. Vision Image Understanding*, **71**-3 (1998), 312–333.

[19] 徐剛, 辻三郎, 「3次元ビジョン」, 共立出版, 1998.

[20] G. Xu and Z. Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*, Kluwer Academic, Dordrecht, The Netherlands, 1996.

[21] Z. Zhang, Determining the epipolar geometry and its uncertainty: A review, *Int. J. Comput. Vision*, **27**-2 (1998), 161–195.

[22] Z. Zhang, On the optimization criteria used in two-view motion analysis, *IEEE Trans. Patt. Anal. Mach. Intell.*, **20**-7 (1998), 717–729.