

くりこみ法による平面の3次元運動解析

武田 佐知男 金谷 健一

群馬大学工学部情報工学科

カメラを移動させながら平面物体を写した2枚の画像上の対応点から、平面の方程式とカメラの3次元運動とを計算する。2枚の画像は射影変換で結ばれるので、射影変換の行列が決まれば、平面の方程式とカメラの運動パラメータとが解析的に計算できる。しかし、変換行列を対応点から通常の最小二乗法で計算すると、解に統計的な偏差が生じる。本論文ではまずこれを指摘する。そして一般的な画像の誤差モデルを導入し、統計的に最適な解法を定式化する。次に数値探索によらずに誤差を適応的に推定しながら解を求める「くりこみ法」と呼ぶ数値解法を構成する。最後にシミュレーションによって誤差が除去されていることを確認する。

キーワード: 3次元運動解析, 平面の運動, 射影変換, くりこみ法, 統計的偏差

3-D Motion Analysis of a Planar Surface by Renormalization

Sachio Takeda and Kenichi Kanatani

Department of Computer Science
Gunma University, Kiryu, Gunma, 376, Japan

Given two images of a planar surface taken by a camera in motion, we reconstruct the 3-D shape of the surface and the 3-D motion of the camera. Since the two images are interrelated by a projective transformation, the surface equation and the camera motion parameters are analytically computed if the matrix of the projective transformation is determined. However, if the transformation matrix is computed by the usual least-squares method, statistical bias is introduced to the solution. We introduce a general model of image noise and derive a statistically optimal scheme for estimating the transformation matrix. Then, we present a numerical scheme, called *renormalization*, for computing the solution by adaptively estimating image noise without doing numerical search. Finally, we conduct a numerical simulation to confirm that statistical bias is indeed removed.

Key words: 3-D motion analysis, planar surface motion, projective transformation, renormalization, statistical bias

謝辞 本研究において、プログラミングに関するいろいろな協力を頂いた群馬高等専門学校の金沢靖助手および群馬大学大学院生の川島徹氏に感謝します。

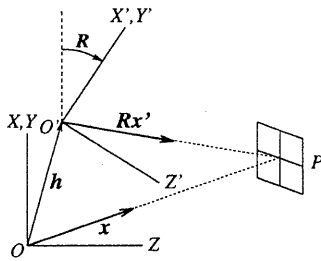


図 1: カメラ運動

1. 序論

空間中を運動する物体の 2 画像, あるいは静止した物体に対してカメラを移動させながら撮影した 2 画像の特徴点の対応から物体の形状や (物体またはカメラの) 3 次元運動を計算することは, コンピュータビジョンの最も基本的な問題であり, これまでに多くの研究が発表されている [1, 2, 4, 7, 8, 11]. しかし, すべての特徴点が臨界曲面と呼ばれる特殊な曲面上にあれば, その 3 次元解釈には必然的に多義性があり, 解が不定となることが知られている [4, 10]. 臨界曲面は一般には二つのカメラのレンズの中心を通る一葉双曲面であり, 実の線織面である. しかしその縮退として, 任意の平面も臨界曲面になっており, すべての特徴点が平面上にあるときは別の解法が必要になる [2, 4, 9, 11].

空間中の特徴点がすべて同一平面上にあれば, その運動を画像面上に投影したものは画像面に射影変換を引き起こす [2, 4]. 射影変換の行列は, 特徴点の対応から最小二乗法によってロバストに計算できるように思われるが [2, 4], 本論文ではまずそのような最小二乗法によって計算した解には統計的偏差が存在することを指摘する. そして, 画像の誤差モデル [3] を導入し, 毎回の反復によって誤差の挙動を推定しながら偏差を取り除くくりこみ法 [5] と呼ぶ方法を用いて解を最適に計算する. 最後に乱数を用いたシミュレーションによってこの方法の有効性を示す.

2. 平面の 3 次元運動解釈

カメラのレンズの中心を原点 O とし, その光軸を Z 軸とする XYZ 座標系をとり, 平面 $Z = f$ を画像面とする. 座標原点 O を視点, 定数 f を焦点距離と呼ぶ. 以下, カメラが物体に対して運動する場合を考える. カメラの運動を視点の周りの回転 R と並進 h とで指定し, $\{h, R\}$ を運動パラメータと呼ぶ (図 1).

空間中の物体の特徴点を $\{P_\alpha\}, \alpha = 1, \dots, N$ とし, 各点のカメラ運動前後の画像座標を $(\bar{x}_\alpha, \bar{y}_\alpha), (\bar{x}'_\alpha, \bar{y}'_\alpha)$

とする. 3 次元ベクトル $\bar{x}_\alpha, \bar{x}'_\alpha$ を

$$\bar{x}_\alpha = \begin{pmatrix} \bar{x}_\alpha/f \\ \bar{y}_\alpha/f \\ 1 \end{pmatrix}, \quad \bar{x}'_\alpha = \begin{pmatrix} \bar{x}'_\alpha/f \\ \bar{y}'_\alpha/f \\ 1 \end{pmatrix} \quad (1)$$

と定義する.

空間中の特徴点 $\{P_\alpha\}, \alpha = 1, \dots, N$ が同一平面上にあれば, ある正則行列 A が存在して, 式 (1) のベクトル $\bar{x}_\alpha, \bar{x}'_\alpha$ は

$$\bar{x}'_\alpha \propto A \bar{x}_\alpha \quad (2)$$

の関係を満たす [2, 4]. これは画像上で射影変換が引き起こされることを意味している [1, 10]. 射影変換を定義する行列 A が求まれば, これから平面の方程式や運動パラメータが解析的に計算できる [2, 4, 9, 11]. しかし, 実際に観測できるのは真の画像座標 $(\bar{x}_\alpha, \bar{y}_\alpha), (\bar{x}'_\alpha, \bar{y}'_\alpha)$ ではなく, それらに誤差が入った $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha)$ である. これらを式 (1) のように 3 次元ベクトル x_α, x'_α で表せば, 平面の 3 次元運動解析は次の問題に帰着する.

【問題】 $\alpha = 1, \dots, N$ に対して

$$\bar{x}'_\alpha \times A \bar{x}_\alpha = 0 \quad (3)$$

が成立する行列 A をデータ $\{x_\alpha\}, \{x'_\alpha\}$ から推定せよ.

空間中の平面を考える. その (視点 O から見て外向きの) 単位法線ベクトルを n , 視点 O からの距離を d とすると, 平面の方程式は $(n, r) = d$ と表せる (本論文ではベクトル a, b の内積を (a, b) と書く). このとき式 (3) の射影変換の行列が次のように表せる (付録参照).

$$A = R^T (hn^T - dI) \quad (4)$$

式 (3) から行列 A を定めようとするとき定数倍の不定性が残ることがわかる. これは式 (4) において h と d が定数倍の自由度をもつことを意味する. このことは, 近くの小さい物体の小さい運動と遠くの大きい物体の大きい運動とが画像のみからは区別できないことを表している. 以下, 定数倍の不定性を除くために $\|A\| = 1$ と拘束する. ただし, 行列 $A = (A_{ij})$ のノルムを $\|A\| = \sqrt{\sum_{i=1}^3 A_{ij}^2}$ と定義する.

3. 統計的に最適な推定

観測したデータ x_α, x'_α は厳密な値ではなく誤差を含んでいるため, 式 (3) はデータ x_α, x'_α に対しては成立しない. そこで, A の最適な推定を行なう方法を考える.

まず, 行列 $X_\alpha^{(k)}, k=1, 2, 3$ を

$$X_\alpha^{(k)} = e^{(k)} \times x'_\alpha x_\alpha^\top \quad (5)$$

と定義する. ただし, $e^{(1)} = (1, 0, 0)^\top, e^{(2)} = (0, 1, 0)^\top, e^{(3)} = (0, 0, 1)^\top$ である. 式(5)で x_α, x'_α をそれぞれ $\bar{x}_\alpha, \bar{x}'_\alpha$ に置き換えたものを $\bar{X}_\alpha^{(k)}$ とすれば, 式(3)は次のように表せる.

$$(\bar{X}_\alpha^{(k)}; A) = 0, \quad k=1, 2, 3 \quad (6)$$

データ x_α, x'_α を次のように表す.

$$x_\alpha = \bar{x}_\alpha + \Delta x_\alpha, \quad x'_\alpha = \bar{x}'_\alpha + \Delta x'_\alpha \quad (7)$$

そして x_α, x'_α の共分散行列をそれぞれ

$$V[x_\alpha] = E[\Delta x_\alpha \Delta x_\alpha^\top], \quad V[x'_\alpha] = E[\Delta x'_\alpha \Delta x'_\alpha^\top] \quad (8)$$

とする ($E[\cdot]$ は期待値を表す). 共分散行列 $V[x_\alpha], V[x'_\alpha]$ は特徴点ごとに異なっているがよい. しかし, 一般に誤差の生じる傾向 (例えば等方性) はわかっても, その絶対量を正確に予測するのは困難である. そこで, 誤差の平均的な強度を表す量 ϵ を用いて, 共分散行列 $V[x_\alpha], V[x'_\alpha]$ を

$$V[x_\alpha] = \epsilon^2 V_0[x_\alpha], \quad V[x'_\alpha] = \epsilon^2 V_0[x'_\alpha] \quad (9)$$

と表す. そして, ϵ は未知であるが, $V_0[x_\alpha], V_0[x'_\alpha]$ は既知であるとする. ϵ をノイズレベル, $V_0[x_\alpha], V_0[x'_\alpha]$ を正規化共分散行列と呼ぶ.

誤差は小さいと仮定し, 独立な誤差の積を正規分布で近似すると, 行列 A の最適な推定値は次のように定義した関数 $J[A]$ の最小化によって求まる [6].

$$\begin{aligned} J[A] &= \sum_{\alpha=1}^N \sum_{k,l=1}^3 W_\alpha^{(kl)}(A)(X_\alpha^{(k)}; A)(X_\alpha^{(l)}; A) \\ &= \sum_{\alpha=1}^N (x'_\alpha \times Ax_\alpha, W_\alpha(A)x'_\alpha \times Ax_\alpha) \quad (10) \end{aligned}$$

ただし $W_\alpha^{(kl)}(A)$ は, 次のように定義した行列 $W_\alpha(A)$ の (kl) 要素である.

$$\begin{aligned} W_\alpha(A) &= (x'_\alpha \times AV_0[x_\alpha]A^\top \times x'_\alpha \\ &\quad + (Ax_\alpha) \times V_0[x'_\alpha] \times (Ax_\alpha) \\ &\quad + [AV_0[x_\alpha]A^\top \times V_0[x'_\alpha]])_2 \quad (11) \end{aligned}$$

ここで記号 $(\cdot)_2$ はランク 2 の一般逆行列, すなわちスベクトル分解して大きい固有値を 2 個残して, 残りの固有値を 0 に置き換えて計算した一般逆行列である [6]. また, 3 次元ベクトル $a = (a_i), b = (b_i)$ と 3 次

元行列 $A = (A_{ij}), B = (B_{ij})$ に対して $a \times A \times b$ と $[A \times B]$ をそれぞれ, その (ij) 要素が次のように表される行列と定義する.

$$(a \times A \times b)_{ij} = \sum_{k,l,m,n=1}^3 \epsilon_{ikl} \epsilon_{jmn} a_k b_m A_{ln} \quad (12)$$

$$[A \times B]_{ij} = \sum_{k,l,m,n=1}^3 \epsilon_{ikl} \epsilon_{jmn} A_{km} B_{ln} \quad (13)$$

ただし ϵ_{ijk} はエディンガトンのイブシロンである ((ijk) が (123) の偶順列のとき 1, 奇順列のとき -1, その他のとき 0 をとる).

式(10) は非線形最適化問題である. この解を数値探索を用いなくて計算する手法を構成する.

4. 最小二乗法の統計的偏差

式(10)の中に現れる行列 $W_\alpha(A) = (W_\alpha^{(kl)}(A))$ を定数行列 $W_\alpha = (W_\alpha^{(kl)})$ に置き換えると, 式(10)の最小化は次の最小二乗法の形に書くことができる.

$$\bar{J}[A] = (A; MA) \rightarrow \min \quad (14)$$

ただし, モーメントテンソル M を次のように定義する.

$$M = \frac{1}{N} \sum_{\alpha=1}^N \sum_{k,l=1}^3 W_\alpha^{(kl)} X_\alpha^{(k)} \otimes X_\alpha^{(l)} \quad (15)$$

そして, テンソル $M = (M_{ijkl})$ と行列 $A = (A_{ij})$ との積 MA を, その (ij) 要素が $\sum_{k,l=1}^3 M_{ijkl} A_{kl}$ の行列と定義し, 行列 $A = (A_{ij})$ と $B = (B_{ij})$ の内積 $(A; B)$ を $\sum_{i,j=1}^3 A_{ij} B_{ij}$ と定義する.

条件 $\|A\| = 1$ のもとでの最小化 (14) の解はテンソル M の最小固有値に対するノルムが 1 の固有行列として与えられる. ただし, テンソル M と行列 A に対して $MA = \lambda A$ が成立するとき, λ と A をそれぞれテンソル M の固有値, 固有行列と呼ぶ [4]. テンソル M の固有値と固有行列の計算は 3×3 行列 A を 9 次元ベクトルで表し, テンソル M を 9×9 行列で表して, 固有値を計算し, それに対する 9 次元固有ベクトルを再び 3×3 行列 A に変換すればよい [4].

そこで, 行列 A の適当な初期推定値 A_0 を式(11)に代入したものを $W_\alpha = W_\alpha(A_0)$ とし, 上の手順で計算した行列 A の値を A_1 とする. これを再び式(11)に代入して $W_\alpha = W_\alpha(A_1)$ とし, また同じ手順で行列 A の値 A_2 を計算し, 以下これを収束するまで反復すれば最適解が求まるように思える. しかし, このような方法では解に統計的な偏差が生じる. これは, 式(15)で, $X_\alpha^{(k)}, X_\alpha^{(l)}$ をそれぞれ $\bar{X}_\alpha^{(k)}, \bar{X}_\alpha^{(l)}$ に

置き換えたものを \bar{M} と書くと, $E[M]$ は一般には \bar{M} とは等しくないからである. 実際, $X_\alpha^{(k)} = \bar{X}_\alpha^{(k)} + \Delta X_\alpha^{(k)}$ と置くと, $E[\Delta X_\alpha^{(k)}] = O$ であるから, テンソル M の期待値は次のようになる.

$$\begin{aligned} E[M] &= \frac{1}{N} \sum_{\alpha=1}^N \sum_{k,l=1}^3 W_\alpha^{(kl)} E[\bar{X}_\alpha^{(k)} \otimes \bar{X}_\alpha^{(l)}] \\ &\quad + \frac{1}{N} \sum_{\alpha=1}^N \sum_{k,l=1}^3 W_\alpha^{(kl)} E[\Delta X_\alpha^{(k)} \otimes \Delta X_\alpha^{(l)}] \\ &= \bar{M} + O(\epsilon^2) \end{aligned} \quad (16)$$

したがって, どんな重み $W_\alpha^{(kl)}$ を用いても (例え真の値でも), 固有値問題の摂動定理 [2, 4] により M の固有行列の期待値には $O(\epsilon^2)$ の偏差が生じる.

5. 不偏推定とくりこみ法

次のようなテンソル $\mathcal{N}^{(1)} = (N_{ijkl}^{(1)})$ と $\mathcal{N}^{(2)} = (N_{ijkl}^{(2)})$ を定義する.

$$\begin{aligned} N_{ijkl}^{(1)} &= \frac{1}{N} \sum_{\alpha=1}^N \sum_{m,n,p,q=1}^3 \epsilon_{imp} \epsilon_{knq} W_\alpha^{(mn)} \\ &\quad \left(V_0[x_\alpha]_{jl} x'_{\alpha(p)} x'_{\alpha(q)} + V_0[x'_\alpha]_{pq} x_{\alpha(j)} x_{\alpha(l)} \right) \end{aligned} \quad (17)$$

$$\begin{aligned} N_{ijkl}^{(2)} &= \\ &\quad \frac{1}{N} \sum_{\alpha=1}^N \sum_{m,n,p,q=1}^3 \epsilon_{imp} \epsilon_{knq} W_\alpha^{(mn)} V_0[x_\alpha]_{jl} V_0[x'_\alpha]_{pq} \end{aligned} \quad (18)$$

そして, 不偏共分散テンソル \hat{M} を

$$\hat{M} = \bar{M} - \epsilon^2 \mathcal{N}^{(1)} + \epsilon^4 \mathcal{N}^{(2)} \quad (19)$$

と定義すれば,

$$E[\hat{M}] = \bar{M} \quad (20)$$

となることがわかる. したがって, この不偏共分散テンソルを用いた次の最適化の解には統計的な偏差が生じない.

$$\hat{J}[A] = (A; \hat{M}A) \rightarrow \min \quad (21)$$

しかし, \hat{M} は未知のノイズレベル ϵ を含んでいる. これを実際の画像から推定するのは困難である. そこで, これを推定しながら反復によって偏差を取り除き, 同時に重み W_α も更新する次の手法を用いる. これを (2次の) くりこみ法と呼ぶ [5].

1. $c = 0, W_\alpha = I, \alpha = 1, \dots, N$ と置く.

2. テンソル $M, \mathcal{N}^{(1)}, \mathcal{N}^{(2)}$ を式 (15), (17), (18) によって計算する.

3. テンソル

$$\hat{M} = M - c\mathcal{N}^{(1)} + c^2\mathcal{N}^{(2)} \quad (22)$$

の最小固有値に対する単位固有行列 A を計算し, その最小固有値を λ とする.

4. $\lambda \approx 0$ であれば A, c, \hat{M} を返して終了する. そうでなければ定数 c と重み W_α を次のように更新する.

$$\begin{aligned} D &= \left((A; \mathcal{N}^{(1)}A) - 2c(A; \mathcal{N}^{(2)}A) \right)^2 \\ &\quad - 4\lambda(A; \mathcal{N}^{(2)}A) \end{aligned} \quad (23)$$

$D \geq 0$ のとき

$$c \leftarrow c + \frac{(A; \mathcal{N}^{(1)}A) - 2c(A; \mathcal{N}^{(2)}A) - \sqrt{D}}{2(A; \mathcal{N}^{(2)}A)} \quad (24)$$

$D < 0$ のとき

$$c \leftarrow c + \frac{\lambda}{(A; \mathcal{N}^{(1)}A)} \quad (25)$$

$$\begin{aligned} W_\alpha &\leftarrow \left(x'_\alpha \times A P_k A^T \times x'_\alpha \right. \\ &\quad \left. + (A x_\alpha) \times P_k \times (A x_\alpha) \right. \\ &\quad \left. + c[A P_k A^T \times P_k] \right)_2 \end{aligned} \quad (26)$$

5. ステップ 2 へ戻る.

6. 運動パラメータと平面のパラメータの計算

変換行列 A から運動パラメータ $\{R, h\}$ と平面のパラメータ $\{n, d\}$ は次のように解析的に計算できる [2, 4, 9, 11]. ただし 1 節で指摘したように, 並進 h と平面までの距離 d のスケールが不定であるから, $h \neq 0$ を仮定し, $\|h\| = 1$ となるスケールを用いる ($h = 0$ かどうかの判定は容易である [2, 4]). ここで, カメラの運動は平面の同じ側で起こる場合 (場合 1) と平面の反対側に回り込む場合 (場合 2) を区別する.

1. 行列 A を $\det A = 1$ となるようにスケールを変える.

2. 行列 AA^T の固有値を $\lambda_1 \geq \lambda_2 \geq \lambda_3 (> 0)$ とし, 対応する固有ベクトルのつくる正規直交系を $\{u_1, u_2, u_3\}$ とする. そして $\sigma_i = \sqrt{\lambda_i} (> 0)$, $i = 1, 2, 3$ と置く (σ_i は行列 A の特異値である [2]).

3. 平面のパラメータ $\{n, d\}$ を次のように計算する.

(場合 1)

$$n = N \left[\sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sqrt{\sigma_2^2 - \sigma_3^2} u_3 \right] \quad (27)$$

$$d = \frac{\sigma_2}{\sigma_1 - \sigma_3} \quad (28)$$

(場合 2)

$$n = N \left[\sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sqrt{\sigma_2^2 - \sigma_3^2} u_3 \right] \quad (29)$$

$$d = \frac{\sigma_2}{\sigma_1 + \sigma_3} \quad (30)$$

ただし $N[\cdot]$ は単位ベクトルへの正規化を表す.

4. 運動パラメータ $\{h, R\}$ を次のように計算する.

(場合 1)

$$h = N \left[-\sigma_3 \sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sigma_1 \sqrt{\sigma_2^2 - \sigma_3^2} u_3 \right] \quad (31)$$

$$R = \frac{1}{\sigma_2} \left(I + \frac{\sigma_2^3}{d} n h^\top \right) A^\top \quad (32)$$

(場合 2)

$$h = N \left[\sigma_3 \sqrt{\sigma_1^2 - \sigma_2^2} u_1 \pm \sigma_1 \sqrt{\sigma_2^2 - \sigma_3^2} u_3 \right] \quad (33)$$

$$R = \frac{1}{\sigma_2} \left(-I + \frac{\sigma_2^3}{d} n h^\top \right) A^\top \quad (34)$$

5. 得られた 2 つの解のそれぞれについて, $(n, k) < 0$ であればベクトル n, h の符号を同時に反転させる.

以上から 8 組の解が定まる. 明らかに $\{h, n\}$ が解であれば $\{-h, -n\}$ も解である. この 2 つの解に対する平面は視点に対して互いに鏡像の位置にある. そして, それぞれに対して運動後にそれをカメラの前後からながめている 2 つの解釈が求まるので, 本質的に異なる解は 2 組ある [4]. 平面がカメラの背後にあっても, それが画像面に平行でない限り見えることに注意 (例えばカメラを斜め上に向けても, 視野が十分広ければ画像の一部に地平線が写る). α 番目の特徴点の空間中の位置の XY 平面および $X'Y'$ 平面からの距離 (カメラの前方が正) はそれぞれ次のようになる (付録参照).

$$Z_\alpha = \frac{d}{(n, x_\alpha)}, \quad Z'_\alpha = \frac{d - (n, h)}{(n, R x'_\alpha)} \quad (35)$$

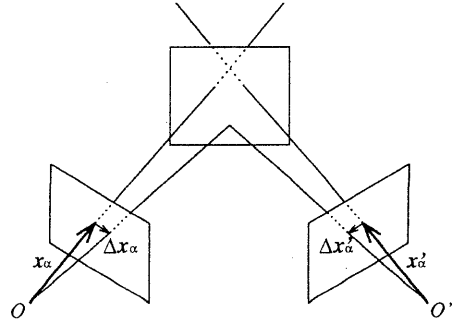


図 2: 特徴点の補正

したがって, すべての特徴点がカメラの運動前後ともカメラの前方にある条件は次のように書ける.

$$Z_\alpha > 0, \quad Z'_\alpha > 0, \quad \alpha = 1, \dots, N \quad (36)$$

これから解が高々 2 組に限定され, 多くの場合は解が一意的に定まる. 解が 2 組生じる場合には何らかの補助的情報によって, 2 組のうち真の解に近い解が選べると仮定する.

7. 特徴点の補正

α 番目の特徴点の 3 次元位置は第 1 のカメラ座標系の原点 O を始点とし, ベクトル x_α の定義する視線方向の, 式 (35) で与えられる距離 Z_α のところにある. また, 第 2 のカメラ座標系の原点 O' を始点とし, ベクトル x'_α の定義する視線方向の, 式 (35) で与えられる距離 Z'_α のところにある. しかし, x_α, x'_α が誤差を含んでいれば, これら二つの視線が交わり, かつその交点が平面 $(n, r) = d$ の上にあるとは限らない. それはデータ x_α, x'_α が誤差のために厳密に式 (3) を満たすとは限らないからである. そこで式 (3) が厳密に満たされるように最適に補正する (図 2). x_α, x'_α をそれぞれ $\Delta x_\alpha, \Delta x'_\alpha$ だけ補正したものが, 式 (3) を満たす条件は

$$(x'_\alpha + \Delta x'_\alpha) \times A(x_\alpha + \Delta x_\alpha) = 0 \quad (37)$$

である. 展開して 1 次近似をとると次のようになる.

$$\Delta x'_\alpha \times A x_\alpha + x'_\alpha \times A \Delta x_\alpha = -x'_\alpha \times A x_\alpha \quad (38)$$

最適な補正は式 (38) の制約のもとでのマハラノビス距離の最小化

$$J = (\Delta x_\alpha, V[x_\alpha]^{-1} \Delta x_\alpha)$$

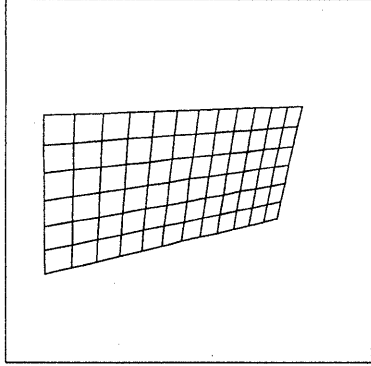


図 3: カメラ運動前の画像

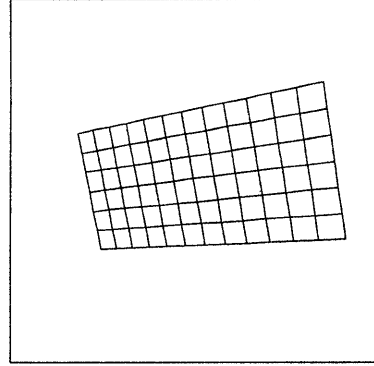


図 4: カメラ運動後の画像

$$+(\Delta x'_\alpha, V[x'_\alpha]^{-1} \Delta x'_\alpha) \rightarrow \min \quad (39)$$

によって求まる。この解 $\Delta x_\alpha, \Delta x'_\alpha$ は次のように定まる [6]。

$$\Delta x_\alpha = -V_0[x_\alpha](x'_\alpha \times A)^T W_{0\alpha}(x'_\alpha \times Ax_\alpha) \quad (40)$$

$$\Delta x'_\alpha = -V_0[x'_\alpha]((Ax_\alpha) \times W_{0\alpha})(x'_\alpha \times Ax_\alpha) \quad (41)$$

$$W_{0\alpha} = (((Ax_\alpha) \times I)V_0[x_\alpha]((Ax_\alpha) \times I)^T + (x'_\alpha \times A)V_0[x'_\alpha](x'_\alpha \times A)^T)^{-1} \quad (42)$$

ただし、式 (38) は式 (37) の 1 次近似であるから、1 回の計算で式 (37) が満たされるとは限らない。そこで補正 $x_\alpha \leftarrow x_\alpha + \Delta x_\alpha, x'_\alpha \leftarrow x'_\alpha + \Delta x'_\alpha$ を反復する。このように補正すれば、これから計算した 3 次元位置は厳密に求めた平面上に存在する。

補正した $\hat{x}_\alpha, \hat{x}'_\alpha$ が式 (3) を満たせば、その 3 次元位置は次のように与えられる。

$$r_\alpha = \frac{d\hat{x}_\alpha}{(n, \hat{x}_\alpha)} \quad (43)$$

8. 数値実験

空間中に仮想的に 7×13 の平面格子ボタン作成し、平面の反対側に回り込まないようにカメラを移動させ、2 枚の画像を得る。ただし、各画像とも 512×512 (画素) に想定し、焦点距離を $f = 600$ (画素) とする。

運動前後の各画像の格子点の x, y 座標に独立に標準偏差 σ 画素の正規乱数を誤差として加える。このと

き正規化共分散行列とノイズレベルは次のように書ける。

$$\epsilon = \frac{\sigma}{f} \quad (44)$$

$$V_0[x_\alpha] = V_0[x'_\alpha] = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 1 \end{pmatrix} \quad (45)$$

ただし計算では ϵ は未知数として扱う。そして、運動パラメータと平面のパラメータをそれぞれ計算し、平面、並進、回転の誤差を以下のような誤差ベクトルによって評価する。

- 真の平面を $(\bar{n}, r) = \bar{d}$ とするとき、平面の誤差ベクトル Δu を次のように定義する。

$$\Delta u = P_{\bar{n}}(n - \bar{n}) + \frac{\|\bar{h}\| |d - \bar{d}|}{\bar{d}} \bar{n} \quad (46)$$

ただし、 $P_{\bar{n}} = I - \bar{n}\bar{n}^T$ である。

- 真の並進を \bar{h} とするとき、並進の誤差ベクトル Δh を次のように定義する。

$$\Delta h = P_{N[\bar{h}]}(h - N[\bar{h}]) \quad (47)$$

ただし、 $P_{N[\bar{h}]} = I - N[\bar{h}]N[\bar{h}]^T$ である。

- 真の回転を \bar{R} とするとき、

$$\bar{R} = R\bar{R}^T \quad (48)$$

と置き、相対回転行列 \bar{R} の回転軸を \bar{l} (単位ベクトル)、回転角を $\Delta\bar{\Omega}$ とする。回転の誤差ベクトル $\Delta\bar{\Omega}$ を次のように定義する。

$$\Delta\bar{\Omega} = \Delta\bar{\Omega}\bar{l} \quad (49)$$

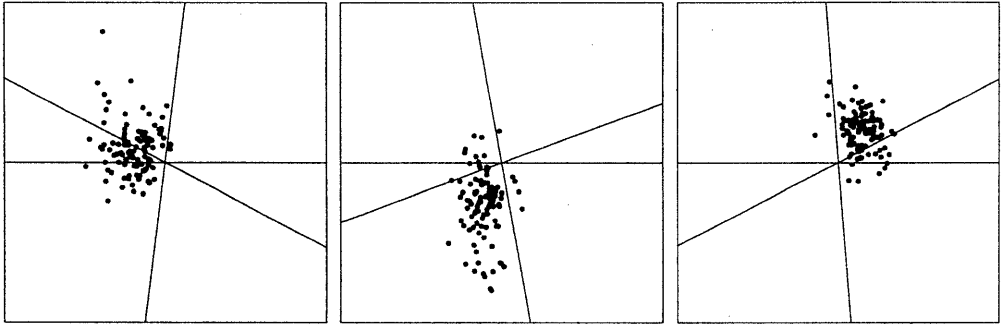


図 5: 最小二乗法 (左: 平面の誤差, 中: 並進の誤差, 右: 回転の誤差)

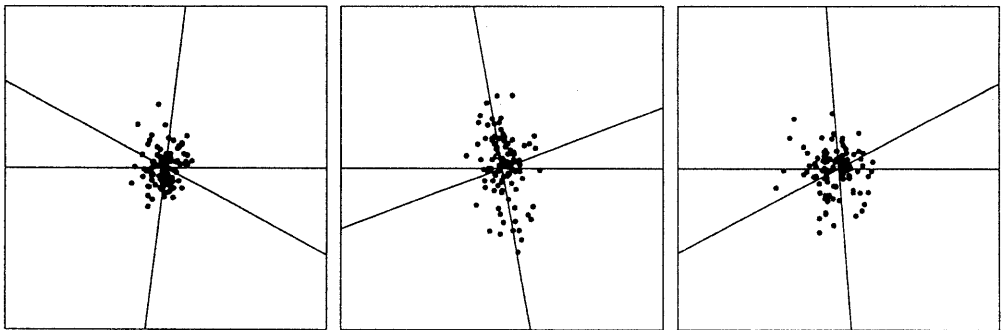


図 6: 本手法 (左: 平面の誤差, 中: 並進の誤差, 右: 回転の誤差)

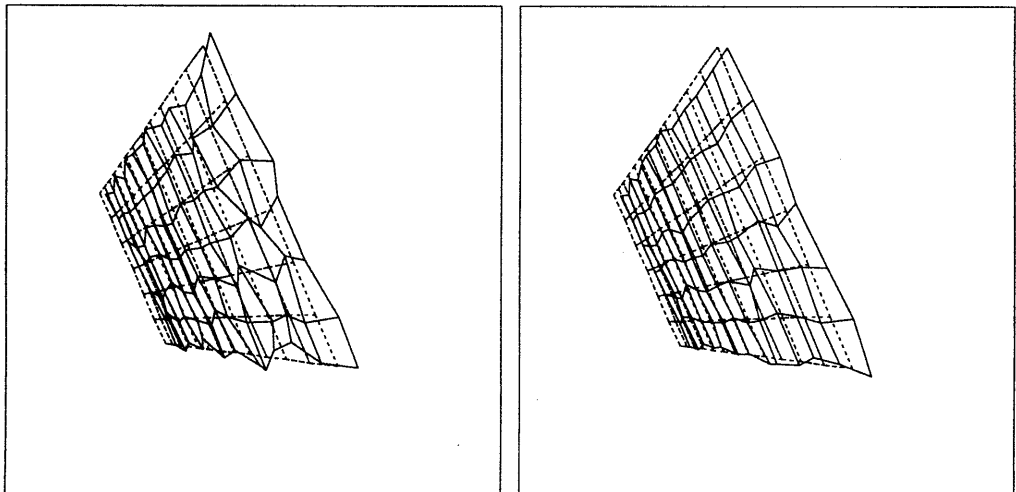


図 7: 格子パタンの 3次元復元 (左: 特徴点の補正を行わない, 右: 補正を行なう)

そして、異なる独立な誤差に対して100回のシミュレーションを行ない、各誤差ベクトルを計算し、それらをそれぞれ3次元空間の原点の周りにプロットする。最適な最小二乗法(最適な重みを変換行列 A の真の値 \bar{A} から計算する)を用いたものと、本手法によるものとの比較を行なった。 $\sigma = 5$ 画素に対する結果を図5と6に示す。最小二乗法による方法は原点から少し離れたところに誤差ベクトルの中心があるのに対し、本手法では原点を中心とした分布となっている。これより、最小二乗法では取り除けなかった偏差が本手法では取り除かれていることがわかる。

誤差の与え方によって復元の結果には違いが見られるが、図7に標準的と思われるものを一例として、特徴点の補正を行なわずに式(43)によって3次元復元したもの(左)と補正してから3次元復元したもの(右)を示す。これは側面からながめたものである。また、真の位置を点線で表示する。明らかに補正を行なった方が復元の精度が向上している。

9. まとめ

本論文では、対象とする物体が平面であるとき、2枚の画像からその3次元運動を計算するアルゴリズムを検討し、誤差解析に基づいて最良と思われるアルゴリズムを示した。これは画像面の射影変換行列の計算において、反復によって誤差の挙動を推定しながら偏差を取り除く「くりこみ法」と呼ぶ方法を用い、平面の復元において画像上の特徴点を統計的に最適に補正するものである。最後に、乱数を用いたシミュレーションを行なってその有効性を確認した。

参考文献

- [1] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA, U.S.A., 1993.
- [2] 金谷健一, 「画像理解/3次元認識の数理」 森北出版, 1990.
- [3] 金谷健一, 画像の3次元解釈の統計的信頼性, 情報処理学会論文誌, Vol. 34, No. 10, pp. 2062-2070, 1993.
- [4] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford Univ. 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, Vol. 35, No. 2, 201-209, 1993
- [6] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Lecture, Note, Department of Computer Science, Gunma University, April 1994.

- [7] 金谷健一, くりこみ法による3次元解析: 統計的に最適なアルゴリズム, 情報処理学会研究報告, 94-CV-88, 1994-03.
- [8] K. Kanatani, "Renormalization for motion analysis: Statistically optimal algorithm," *IEICE Trans. Infor. Sys.*, Vol. E77-D, No. 11, 1994
- [9] H. C. Longuet-Higgins, "The reconstruction of a planar surface from two perspective projections," *Proc. Roy. Soc. Land.*, Ser. B, Vol. 227, pp. 399-410, 1986.
- [10] S. Maybank, *Theory of Reconstruction from Image Motion*, Springer, Berlin, 1993.
- [11] J. Weng, T. S. Huang and N. Ahuja, *Motion and Structure from Image Sequences*, Springer, Berlin, F.R.G., 1993.

付録.

第1, 第2のカメラ座標系の原点をそれぞれ O, O' とすると, $\vec{OP} = O\vec{O}' + O'\vec{P}$ である。ベクトル $O'\vec{P}$ が第2カメラ座標系に関して r' であるとする, 第2のカメラ座標系は第1のカメラ座標系に対して R だけ回転しているから, 第1のカメラ座標系に関しては Rr' である。 $r' = Z'x'$ と表すと, $\vec{OP} = Zx$, $O\vec{O}' = h$ であるから, 次の式を得る。

$$Zx = h + Z'Rx' \quad (50)$$

一方, 空間中の点 $r = Zx$ が平面 $(n, r) = d$ の上にあれば, 式 $(n, Zx) = d$ が満たされるから, Z が次のように定まる。

$$Z = \frac{d}{(n, x)} \quad (51)$$

また, 式(50)より $r = h + Z'Rx'$ であるから, これを $(n, r) = d$ に代入すると Z' が次のように定まる。

$$Z' = \frac{d - (n, h)}{(n, Rx')} \quad (52)$$

式(50)と式(51)と組み合わせると次の関係が得られる。

$$x' = \frac{Z}{Z'} R^T \left(x - \frac{h}{Z} \right) \quad (53)$$

$$= \frac{Z}{Z'} R^T \left(x - \frac{(n, x)h}{d} \right) \quad (54)$$

$$= -\frac{Z}{Z'd} R^T (hn^T - dI)x \quad (55)$$

したがって, 行列 A を式(4)のように定義すれば $x' \propto Ax$ と書ける。