

2 画像からの 3 次元復元の最新アルゴリズム

山田 健人^{†1} 金澤 靖^{†1}
金谷 健一^{†2} 菅谷 保之^{†3}

2 画像の対応点からシーンの 3 次元形状を復元する基本原理はよく知られているが、本論文では著者らの統計的最適化の研究の最新の成果を取り入れてこれを再構成する。この結果、本システムは理論的に従来の方法より精度が向上するだけでなく、プログラムの構成がより明確になり、本システムをさまざまな応用に組み込むのが非常に容易になる。また、本システムではカメラの移動が注視運動になる退化の場合も考慮している。実際の撮影ではほとんどの場合が注視運動であることを考えると、これは非常に実用的である。最後に実画像を用いて本システムの性能を実証する。

Latest Algorithm for 3-D Reconstruction from Two Views

KENTO YAMADA,^{†1} YASUSHI KANAZAWA,^{†1}
KENICHI KANATANI^{†2} and YASUYUKI SUGAYA^{†3}

This paper presents a new algorithm for reconstructing the 3-D shape of the scene from point correspondences over two views. The basic principle is well known, but we incorporate into it the latest results of the authors' studies of statistical optimization techniques. As a result, not only the accuracy increases but also the organization of the program becomes simpler, allowing this system to be used in a wider range of practical application. Also, our system take into consideration the degenerate situation where the camera is in a fixating configuration. This makes our system very practical, since this is the situation most frequently encountered in practice. Finally, we demonstrate the performance of our system using real images.

^{†1} 豊橋技術科学大学知識情報工学系

Department of Knowledge-based Information Engineering, Toyohashi University of Technology

^{†2} 岡山大学大学院自然科学研究科

Department of Computer Science, Okayama University

^{†3} 豊橋技術科学大学情報工学系

Department of Information and Computer Sciences, Toyohashi University of Technology

1. まえがき

2 画像の対応点からその 3 次元位置とカメラの内部および外部パラメータと計算する問題は 1980, 90 年代に精力的に研究され、現在では動画像解析 (structure from motion) の典型例として確立された技術である^{5),7)}。これはロボットビジョンや仮想現実生成などのさまざまな応用に使われ、その原理はビデオ画像列からの 3 次元復元に発展している⁵⁾。しかし、原理は解明されていても、部品となる基本計算は、例えば基礎行列の計算一つをとっても、精度や計算効率を向上をさせるさまざまな技法が追及されている。このような背景を考慮して、本論文では 2 画像からの 3 次元復元のアルゴリズム全体を著者らの統計的最適化の研究の最新の成果を取り入れて再構成する。基本的な構成は金谷・三島⁹⁾の通りであるが、各要素を技術的な進歩を考慮したものに置き換える。具体的には次の部分である。

- (1) 金谷・三島⁹⁾では対応点からの基礎行列の計算にくりこみ法を用い、それをランク拘束を満たすように最適補正を施している。本論文では、これをランク拘束を満たすように基礎行列を最適に計算する拡張 FNS 法¹⁰⁾を組み込んだ厳密な最尤推定計算法¹¹⁾に置き換える。
- (2) 基礎行列の計算は反復法であるために初期値が必要である。金谷・三島⁹⁾のくりこみ法の初期値計算には最小二乗法が使われているが、本論文では Taubin 法¹⁸⁾を用いる。これは精度が非常に高いことが知られ¹⁸⁾、以降の最適化の収束が非常に速くなる。
- (3) 金谷・三島⁹⁾では計算した基礎行列から 2 台のカメラの焦点距離を計算するのに Bougnoux の公式¹⁾を用いている。この計算はカメラが注視運動のとき破綻する(後述)。しかし、実用上は注視運動が最も普通である。本論文では高橋ら¹⁹⁾に従い、注視運動にも対応できる計算法を組み込む。
- (4) 金谷・三島⁹⁾では計算した基礎行列から対応点を補正し、3 次元位置をカメラ配置の幾何学的な関係から計算している。しかし、本論文では計算したカメラの運動パラメータから投影行列を計算し、最適な三角測量の方法¹³⁾に基づいて 3 次元位置を計算する。

2. アルゴリズムの構成

本論文のアルゴリズムの入出力は次の通りである。

入力: 対応する特徴点の座標の組 $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha), \alpha = 1, \dots, N (\geq 8)$.

出力: 各々の3次元位置 $(X_\alpha, Y_\alpha, Z_\alpha)$, $\alpha = 1, \dots, N$.

ただし, 次の前提を置く.

- カメラの光軸点 (光軸と画像面との交点) は既知とし, それを $(0, 0)$ とする xy 画像座標系を定義する. そして, 上方を x 軸, 右方を y 軸とする
- 画像に幾何学的な歪はない, あるいは既に補正されているとし, 画素の配列は縦方向と横方向が直交し, アスペクト比が1であるとみなす.
- 2台のカメラの焦点距離は等しい, あるいは同じことであるが, ズームやフォーカスを変えずに1台のカメラを移動して撮影する*1.

多数の画像間の対応点が得られていれば, それらから光軸点 (さらには幾何学的歪) を推定する理論があるが⁵⁾, 2画像のみの場合は上記の仮定がなければ3次元復元はできない. 現在の市販のデジタルカメラでは画像面の中央を光軸点とし, 幾何学的な歪はないと仮定しても問題ないであろう. そうでない場合は事前の補正が必要となる. 画像の座標軸を上方に x 軸, 右方に y 軸とするのは, 光軸の奥行方向を z 軸と想定するためである. 普通に右方に x 軸, 上方に y 軸をとると xyz 座標系が左手系となり, 後述のカメラの回転行列の行列式が -1 となるので, これを防ぐものである*2. 本論文のアルゴリズムの流れは図1のようになる.

3. 基礎行列の計算

最初のステップは対応点の組 $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha)$ から基礎行列 $F = (F_{ij})$ を計算することである. 以下の計算ではこれを次の9次元ベクトル u と同一視する.

$$u = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^\top \quad (1)$$

計算の流れは図2のようになる.

3.1 初期値の計算

図2は反復計算であり, 初期値が必要である. 金谷・三島⁹⁾のくりこみ法では, 初期値は最小二乗法解を用いていたが, 本論文では次の Taubin 法¹⁸⁾を用いる. これは最小二乗法に比べて非常に高精度であることが知られ¹⁸⁾, 以降の最適化の収束が非常に速くなる. 以下, $\mathcal{N}[\cdot]$ で単位ベクトルへの正規化を表す ($\mathcal{N}[a] = a/\|a\|$).

*1 実際には本アルゴリズムは焦点距離を変化させても有効である. その場合にはカメラの移動が注視運動であってはならない^{8),19)}.

*2 そのためにはもちろん右方に x 軸, “下方” に y 軸をとってもよい.

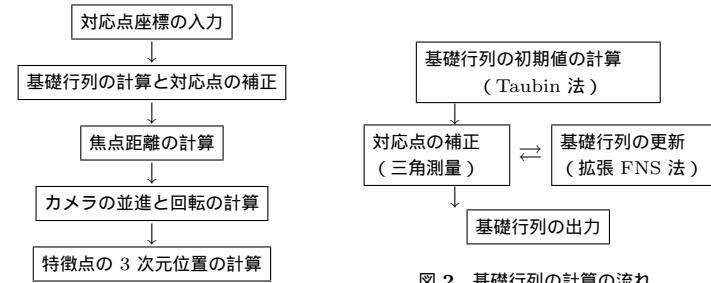


図2 基礎行列の計算の流れ.

図1 本論文のアルゴリズムの流れ.

- (1) 対応点の組 $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha)$ を合わせて次の8次元ベクトル z_α で表す*3.

$$z_\alpha = (x_\alpha x'_\alpha, x_\alpha y'_\alpha, f_0 x_\alpha, y_\alpha x'_\alpha, y_\alpha y'_\alpha, f_0 y_\alpha, f_0 x'_\alpha, f_0 y'_\alpha)^\top \quad (2)$$

- (2) 次の 8×8 行列 $V_0[z_\alpha]$ を計算する.

$$V_0[z_\alpha] = \begin{pmatrix} x_\alpha^2 + x'^2_\alpha & x'_\alpha y'_\alpha & f_0 x'_\alpha & x_\alpha y_\alpha & 0 & 0 & f_0 x_\alpha & 0 \\ x'_\alpha y'_\alpha & x_\alpha^2 + y'^2_\alpha & f_0 y'_\alpha & 0 & x_\alpha y_\alpha & 0 & 0 & f_0 x_\alpha \\ f_0 x'_\alpha & f_0 y'_\alpha & f_0^2 & 0 & 0 & 0 & 0 & 0 \\ x_\alpha y_\alpha & 0 & 0 & y_\alpha^2 + x'^2_\alpha & x'_\alpha y'_\alpha & f_0 x'_\alpha & f_0 y_\alpha & 0 \\ 0 & x_\alpha y_\alpha & 0 & x'_\alpha y'_\alpha & & & & \\ y_\alpha^2 + y'^2_\alpha & f_0 y'_\alpha & 0 & f_0 y_\alpha & & & & \\ 0 & 0 & 0 & f_0 x'_\alpha & f_0 y'_\alpha & f_0^2 & 0 & 0 \\ f_0 x_\alpha & 0 & 0 & f_0 y_\alpha & 0 & 0 & f_0^2 & 0 \\ 0 & f_0 x_\alpha & 0 & 0 & f_0 y_\alpha & 0 & 0 & f_0^2 \end{pmatrix} \quad (3)$$

- (3) 次の8次元ベクトル $\bar{z}, \tilde{z}_\alpha$ を計算する.

$$\bar{z} = \frac{1}{N} \sum_{\alpha=1}^N z_\alpha, \quad \tilde{z}_\alpha = z_\alpha - \bar{z} \quad (4)$$

- (4) 次の 8×8 行列 M_{TB}, L_{TB} を計算する.

$$M_{TB} = \sum_{\alpha=1}^N \tilde{z}_\alpha \tilde{z}_\alpha^\top, \quad L_{TB} = \sum_{\alpha=1}^N V_0[z_\alpha] \quad (5)$$

- (5) 次の一般固有値問題の最小一般固有値 λ に対する一般固有ベクトル v を計算する.

$$M_{TB} v = \lambda L_{TB} v \quad (6)$$

*3 f_0 を用いるのは計算中のベクトルの成分や行列の要素のオーダーをそろえて数値計算を安定化させるためであり²⁾, 最終結果には影響しない. 本システムでは $f_0 = 600$ (画素) とした

(6) 次の9次元単位ベクトル \mathbf{u} を返す.

$$\mathbf{u} = \mathcal{N}\left[\begin{pmatrix} \mathbf{v} \\ -(v, \mathbf{z})/f_0^2 \end{pmatrix}\right] \quad (7)$$

3.2 対応点の補正

このように計算した基礎行列を表すベクトル \mathbf{u} を用いて, 以下の反復によって各特徴点 (x_α, y_α) , (x'_α, y'_α) を $(\hat{x}_\alpha, \hat{y}_\alpha)$, $(\hat{x}'_\alpha, \hat{y}'_\alpha)$ に補正する. これは「ステレオ画像の三角測量」と呼ばれる処理¹³⁾に相当する. 同時にその過程で基礎行列を最適化する. 以下の処理は文献12)に基く.

(1) 再投影誤差を $E^0 = \infty$ (十分大きい数) とし, 次のように置く.

$$\hat{x}_\alpha = x_\alpha, \quad \hat{y}_\alpha = y_\alpha, \quad \hat{x}'_\alpha = x'_\alpha, \quad \hat{y}'_\alpha = y'_\alpha, \quad \tilde{x}_\alpha = \tilde{y}_\alpha = \tilde{x}'_\alpha = \tilde{y}'_\alpha = 0. \quad (8)$$

(2) 次の9次元ベクトル ξ_α と 9×9 行列 $V_0[\xi_\alpha]$ を計算する.

$$\xi_\alpha = \begin{pmatrix} \hat{x}_\alpha \hat{x}'_\alpha + \hat{x}'_\alpha \tilde{x}_\alpha + \hat{x}_\alpha \tilde{x}'_\alpha \\ \hat{x}_\alpha \hat{y}'_\alpha + \hat{y}'_\alpha \tilde{x}_\alpha + \hat{x}_\alpha \tilde{y}'_\alpha \\ f_0(\hat{x}_\alpha + \tilde{x}_\alpha) \\ \hat{y}_\alpha \hat{x}'_\alpha + \hat{x}'_\alpha \tilde{y}_\alpha + \hat{y}_\alpha \tilde{x}'_\alpha \\ \hat{y}_\alpha \hat{y}'_\alpha + \hat{y}'_\alpha \tilde{y}_\alpha + \hat{y}_\alpha \tilde{y}'_\alpha \\ f_0(\hat{y}_\alpha + \tilde{y}_\alpha) \\ f_0(\hat{x}'_\alpha + \tilde{x}'_\alpha) \\ f_0(\hat{y}'_\alpha + \tilde{y}'_\alpha) \\ f_0^2 \end{pmatrix} \quad (9)$$

$$V_0[\xi_\alpha] = \begin{pmatrix} \hat{x}_\alpha^2 + \hat{x}'_\alpha^2 & \hat{x}'_\alpha \hat{y}'_\alpha & f_0 \hat{x}'_\alpha & \hat{x}_\alpha \hat{y}_\alpha & 0 & 0 & f_0 \hat{x}_\alpha & 0 & 0 \\ \hat{x}'_\alpha \hat{y}'_\alpha & \hat{x}_\alpha^2 + \hat{y}'_\alpha^2 & f_0 \hat{y}'_\alpha & 0 & \hat{x}_\alpha \hat{y}_\alpha & 0 & 0 & f_0 \hat{x}_\alpha & 0 \\ f_0 \hat{x}'_\alpha & f_0 \hat{y}'_\alpha & f_0^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hat{x}_\alpha \hat{y}_\alpha & 0 & 0 & \hat{y}_\alpha^2 + \hat{x}'_\alpha^2 & \hat{x}'_\alpha \hat{y}'_\alpha & f_0 \hat{x}'_\alpha & f_0 \hat{y}_\alpha & 0 & 0 \\ 0 & \hat{x}_\alpha \hat{y}_\alpha & 0 & \hat{x}'_\alpha \hat{y}'_\alpha & \hat{y}_\alpha^2 + \hat{y}'_\alpha^2 & f_0 \hat{y}'_\alpha & 0 & f_0 \hat{y}_\alpha & 0 \\ 0 & 0 & 0 & f_0 \hat{x}'_\alpha & f_0 \hat{y}'_\alpha & f_0^2 & 0 & 0 & 0 \\ f_0 \hat{x}_\alpha & 0 & 0 & f_0 \hat{y}_\alpha & 0 & 0 & f_0^2 & 0 & 0 \\ 0 & f_0 \hat{x}_\alpha & 0 & 0 & f_0 \hat{y}_\alpha & 0 & 0 & f_0^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (10)$$

(3) サブプログラム EFNS (3.3 節) を呼んで, \mathbf{u} を更新する.

(4) $\tilde{x}_\alpha, \tilde{y}_\alpha, \tilde{x}'_\alpha, \tilde{y}'_\alpha$ を次のように更新する.

$$\begin{pmatrix} \tilde{x}_\alpha \\ \tilde{y}_\alpha \end{pmatrix} \leftarrow \frac{(\mathbf{u}, \xi_\alpha)}{(\mathbf{u}, V_0[\xi_\alpha] \mathbf{u})} \begin{pmatrix} u_1 & u_2 & u_3 \\ u_4 & u_5 & u_6 \end{pmatrix} \begin{pmatrix} \hat{x}'_\alpha \\ \hat{y}'_\alpha \\ f_0 \end{pmatrix},$$

$$\begin{pmatrix} \tilde{x}'_\alpha \\ \tilde{y}'_\alpha \end{pmatrix} \leftarrow \frac{(\mathbf{u}, \xi_\alpha)}{(\mathbf{u}, V_0[\xi_\alpha] \mathbf{u})} \begin{pmatrix} u_1 & u_4 & u_7 \\ u_2 & u_5 & u_8 \end{pmatrix} \begin{pmatrix} \hat{x}_\alpha \\ \hat{y}_\alpha \\ f_0 \end{pmatrix} \quad (11)$$

(5) $\hat{x}_\alpha, \hat{y}_\alpha, \hat{x}'_\alpha, \hat{y}'_\alpha$ を次のように更新する.

$$\hat{x}_\alpha \leftarrow x_\alpha - \tilde{x}_\alpha, \quad \hat{y}_\alpha \leftarrow y_\alpha - \tilde{y}_\alpha, \quad \hat{x}'_\alpha \leftarrow x'_\alpha - \tilde{x}'_\alpha, \quad \hat{y}'_\alpha \leftarrow y'_\alpha - \tilde{y}'_\alpha \quad (12)$$

(6) 再投影誤差 E を次のように計算する^{*1}.

$$E = \sqrt{\frac{1}{N-7} \sum_{\alpha=1}^N (\tilde{x}_\alpha^2 + \tilde{y}_\alpha^2 + \tilde{x}'_\alpha^2 + \tilde{y}'_\alpha^2)} \quad (13)$$

(7) $|E - E^0| \approx 0$ であれば (本システムでは $|E - E^0| < 10^{-4}$ とした), 次のように定めた基礎行列 \mathbf{F} を返して終了する.

$$\mathbf{F} = \begin{pmatrix} u_1 & u_2 & u_3 \\ u_4 & u_5 & u_6 \\ u_7 & u_8 & u_9 \end{pmatrix} \quad (14)$$

そうでなければ $E^0 \leftarrow E$ としてステップ2に戻る

3.3 基礎行列の更新

上述のアルゴリズムのステップ4で呼ぶ EFNS は「拡張 EFNS 法」と呼ぶ次の計算である¹⁰⁾.

(1) 次の 9×9 行列 \mathbf{M}, \mathbf{L} を計算する.

$$\mathbf{M} = \sum_{\alpha=1}^N \frac{\xi_\alpha \xi_\alpha^\top}{(\mathbf{u}, V_0[\xi_\alpha] \mathbf{u})}, \quad \mathbf{L} = \sum_{\alpha=1}^N \frac{(\mathbf{u}, \xi_\alpha)^2 V_0[\xi_\alpha]}{(\mathbf{u}, V_0[\xi_\alpha] \mathbf{u})^2}. \quad (15)$$

(2) 次の9次元ベクトル \mathbf{u}^\dagger と 9×9 行列 $\mathbf{P}_{\mathbf{u}^\dagger}$ を計算する.

$$\mathbf{u}^\dagger = \mathcal{N}\left[\begin{pmatrix} u_5 u_9 - u_8 u_6 \\ u_6 u_7 - u_9 u_4 \\ u_4 u_8 - u_7 u_5 \\ u_8 u_3 - u_2 u_9 \\ u_9 u_1 - u_3 u_7 \\ u_7 u_2 - u_1 u_8 \\ u_2 u_6 - u_5 u_3 \\ u_3 u_4 - u_6 u_1 \\ u_1 u_5 - u_4 u_2 \end{pmatrix}\right], \quad \mathbf{P}_{\mathbf{u}^\dagger} = \mathbf{I} - \mathbf{u}^\dagger \mathbf{u}^{\dagger\top}. \quad (16)$$

(3) 次の 9×9 行列 \mathbf{X}, \mathbf{Y} を計算する.

$$\mathbf{X} = \mathbf{M} - \mathbf{L}, \quad \mathbf{Y} = \mathbf{P}_{\mathbf{u}^\dagger} \mathbf{X} \mathbf{P}_{\mathbf{u}^\dagger}. \quad (17)$$

*1 このように定義すると, 最終的にこれが各点の各座標の誤差の標準偏差の推定量となる⁷⁾.

(4) 行列 Y の小さい二つの固有値^{*1}に対応する 9 次元単位固有ベクトル v_1, v_2 を計算する .

(5) 次の 9 次元ベクトル \hat{u} を計算する .

$$\hat{u} = (u, v_1)v_1 + (u, v_2)v_2 \quad (18)$$

(6) 次の 9 次元ベクトル u' を計算する .

$$u' = \mathcal{N}[P_{u'} \hat{u}]. \quad (19)$$

(7) 符号を除いて $u' \approx u$ であれば (本システムでは符号をそろえて $\|u' - u\| < 10^{-6}$ とした) , u' を u の更新値として返して終了 . そうでなければ $u \leftarrow \mathcal{N}[u + u']$ としてステップ 1 に戻る .

このようにステレオ画像の最適な三角測量による特徴点の補正と EFNS 法による基礎行列の更新の組み合わせによって , 厳密な最尤推定解が計算できる¹²⁾ .

4. 焦点距離の計算

次に , 計算した基礎行列 F から 2 台のカメラの焦点距離 f, f' (単位は画素) を計算する . これには二通りの方法がある¹⁹⁾ . 一つは , f, f' を (実際には同じ値であるが) 計算上は異なる値として扱う方法である . もう一つは初めから $f = f'$ として計算する方法である . 前者を「自由焦点距離の方法」, 後者を「固定焦点距離の方法」と呼ぶ . さらに自由焦点距離の方法で計算される焦点距離の重み付き平均をとることも考えられる . これを「平均焦点距離の方法」と呼ぶ .

4.1 自由焦点距離の方法

対称行列 $FF^T, F^T F$ の最小固有値に対する 3 次元単位固有ベクトルをそれぞれ e, e' とし^{*2} , 次のように f, f' を計算する . 以下 , 大域的に $k = (0, 0, 1)^T$ と定義する .

$$\xi = \frac{\|Fk\|^2 - (k, FF^T Fk)\|e' \times k\|^2 / (k, Fk)}{\|e' \times k\|^2 \|F^T k\|^2 - (k, Fk)^2},$$

$$\eta = \frac{\|F^T k\|^2 - (k, FF^T Fk)\|e \times k\|^2 / (k, Fk)}{\|e \times k\|^2 \|Fk\|^2 - (k, Fk)^2} \quad (20)$$

$$f = \frac{f_0}{\sqrt{1 + \xi}}, \quad f' = \frac{f_0}{\sqrt{1 + \eta}} \quad (21)$$

*1 文献 10) では絶対値が小さい二つの固有値としていたが , 後の実験により , 小さい固有値を選ぶほうが誤差の大きいデータに対して収束性能が向上することが確認された .

*2 e, e' は「エビ極点」を表す . すなわち $(f_0 e_1 / e_3, f_0 e_2 / e_3)$ が第 1 画像上の第 2 カメラの視点の投影位置であり , $(f_0 e'_1 / e'_3, f_0 e'_2 / e'_3)$ が第 2 画像上の第 1 カメラの視点の投影位置である .

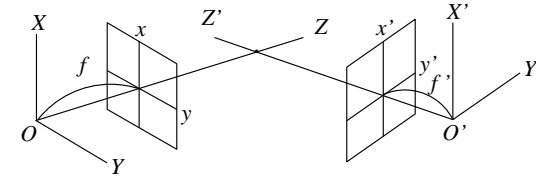


図 3 注視の配置 . 2 画像を撮影するカメラの光軸が交わる .

式 (21) の根号の中身が 0 または負になればエラーメッセージを出して終了する . 式 (20) は Bougnoux¹⁾ が示した式を書き直した式である (導出は文献 8) 参照) . 式 (20) から直ちに分かるように , $(F_{33} =) (k, Fk) = 0$ のときに計算が破綻する . k は視点から画像の原点を指すベクトルであるから , $(k, Fk) = 0$ は第 1 画像の原点と第 2 画像の原点が対応するという「エビ極線方程式」である . これはカメラの配置が「注視」(シーン中の同一点が 2 画像のそれぞれの中心に写るようなカメラを移動) であること意味する (図 3) . これはよく知られた事実であるが^{8), 19)} , 実際の撮影では注視の配置が最も普通である . このため , あらかじめ $(k, Fk) \approx 0$ かどうかを判定して^{*3} , 注視とみなされればエラーメッセージを出して計算を中止する .

4.2 平均焦点距離の方法

自由焦点距離の方法で解が得られた場合でも , 2 台のカメラの焦点距離が同一の撮影に対して計算上は一般に異なる値が得られる . それらを一つの値に等号するには次のような ξ と η の重み付き平均をとるのが合理的である (導出省略) .

$$\tilde{\xi} = \tilde{\eta} = \frac{(H_{11} + H_{12})\xi + (H_{22} + H_{12})\eta}{H_{11} + 2H_{12} + H_{22}}. \quad (22)$$

ただし , H_{ij} を次のように定義する .

$$H_{11} = 2(k, Fk)^4 \eta^2 + 4(k, Fk)^2 \|F^T k\|^2 \eta + 2\|F^T k\|^4 - \left((k, Fk)^2 \eta + \|F^T k\|^2 \right)^2,$$

$$H_{22} = 2(k, Fk)^4 \xi^2 + 4(k, Fk)^2 \|Fk\|^2 \xi + 2\|Fk\|^4 - \left((k, Fk)^2 \xi + \|Fk\|^2 \right)^2,$$

$$H_{12} = 4(k, Fk)^4 \xi \eta + 4(k, Fk)^2 \left(\|F^T k\|^2 \xi + \|Fk\|^2 \eta \right) + 4(k, Fk)(k, FF^T Fk)$$

*3 本システムでは $|(k, Fk)| < 0.1 \min(\|Fk\|, \|F^T k\|) / f_0$ を判定条件としている .

$$\begin{aligned}
 & -\left((\mathbf{k}, \mathbf{F}\mathbf{k})^2\xi + \|\mathbf{F}\mathbf{k}\|^2\right)\left((\mathbf{k}, \mathbf{F}\mathbf{k})^2\eta + \|\mathbf{F}^\top\mathbf{k}\|^2\right) \\
 & -(\mathbf{k}, \mathbf{F}\mathbf{k})^2\left((\mathbf{k}, \mathbf{F}\mathbf{k})^2\xi\eta + \|\mathbf{F}^\top\mathbf{k}\|^2\xi + \|\mathbf{F}\mathbf{k}\|^2\eta + \|\mathbf{F}\|^2\right)
 \end{aligned} \quad (23)$$

そして式 (22) の ξ, η を式 (21) に代入したものを \tilde{f}, \tilde{f}' とする.

4.3 固定焦点距離の方法

$(\mathbf{k}, \mathbf{F}\mathbf{k}) = 0$ の注視の場合でも固定焦点距離の方法を用いれば焦点距離 $f = f'$ が計算できることが知られている^{8),19)}. これは次のように行う.

(1) 次の $a_1 \sim a_5$ を計算する.

$$\begin{aligned}
 a_1 &= \frac{1}{2}(\mathbf{k}, \mathbf{F}\mathbf{k})^4 \\
 a_2 &= (\mathbf{k}, \mathbf{F}\mathbf{k})^2(\|\mathbf{F}^\top\mathbf{k}\|^2 + \|\mathbf{F}\mathbf{k}\|^2) \\
 a_3 &= \frac{1}{2}(\|\mathbf{F}^\top\mathbf{k}\|^2 - \|\mathbf{F}\mathbf{k}\|^2)^2 + (\mathbf{k}, \mathbf{F}\mathbf{k})(4(\mathbf{k}, \mathbf{F}\mathbf{F}^\top\mathbf{F}\mathbf{k}) - (\mathbf{k}, \mathbf{F}\mathbf{k})\|\mathbf{F}\|^2) \\
 a_4 &= 2(\|\mathbf{F}\mathbf{F}^\top\mathbf{k}\|^2 + \|\mathbf{F}^\top\mathbf{F}\mathbf{k}\|^2) - (\|\mathbf{F}^\top\mathbf{k}\|^2 + \|\mathbf{F}\mathbf{k}\|^2)\|\mathbf{F}\|^2 \\
 a_5 &= \|\mathbf{F}\mathbf{F}^\top\|^2 - \frac{1}{2}\|\mathbf{F}\|^4
 \end{aligned} \quad (24)$$

(2) 関数 $K(\xi)$ を次のように置く.

$$K(\xi) = a_1\xi^4 + a_2\xi^3 + a_3\xi^2 + a_4\xi + a_5 \quad (25)$$

(3) $(\mathbf{k}, \mathbf{F}\mathbf{k}) \approx 0$ であれば ξ を次のように置く.

$$\xi = -\frac{a_4}{2a_3} \quad (26)$$

(4) そうでなければ, 3 次多項式

$$K'(\xi) = 4a_1\xi^3 + 3a_2\xi^2 + 2a_3\xi + a_4 = 0 \quad (27)$$

の実数解を計算する.

(5) 式 (27) が唯一の実数解をもつときはそれを ξ とする. 3 個の実数解 $\xi_3 \leq \xi_2 \leq \xi_1$ をもつときは次のように置く.

$$\xi = \begin{cases} \xi_1 & \xi_3 \leq -1 \text{ または } K(\xi_3) < 0 \text{ または } K(\xi_1) \leq K(\xi_3) \\ \xi_3 & 0 \leq K(\xi_3) < K(\xi_1) \\ -1 & \text{それ以外} \end{cases} \quad (28)$$

(6) 次のように f, f' を計算する.

$$f = f' = \frac{f_0}{\sqrt{1+\xi}} \quad (29)$$

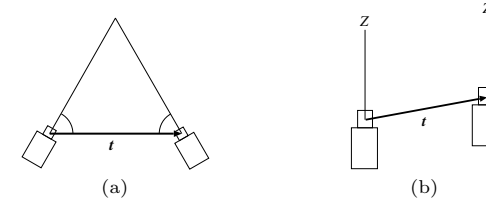


図 4 (a) 対称なカメラ配置. (b) カメラの平行移動.

ただし, 根号の中身が 0 または負になればエラーメッセージを出して終了する.

4.4 問題点

固定小数点の方法を用いれば一般には注視であっても $f = f'$ が計算できるが, カメラの配置が対称 (2 台のカメラの視点と注視点が二等辺三角形を作る. 図 4(a)) あるいは平行移動 (カメラの光軸が平行. 図 4(b)) の場合は焦点距離が計算できないことが知られている⁸⁾. これらはステレオ視の標準的なカメラ配置である. もちろん, あらかじめカメラ校正を行って焦点距離が既知であれば 3 次元復元ができる. しかし, 焦点距離が未知で「自己校正」(撮影した画像から計算する) を行いたい場合はこのようなカメラ配置は避けなければならない. 可能なら注視の配置も避ける. しかし, 実際の撮影ではカメラが厳密にそのような破綻する配置になくても, それに近いと, しばしば式 (21), (29) 中の根号の中身が負になる「虚数焦点距離」が生じる^{*1}. 自由焦点距離の方法, 平均焦点距離の方法, 固定焦点距離の方法の一つから解が得られればそれを使えばよいが, すべてが破綻する場合はあらかじめ経験的に定めたデフォルト値 f, f' を使うしかない. 一方, 複数の解が得られた場合はどれかを選択しなければならない. 高橋ら¹⁹⁾ は経験的な判定条件を示しているが, しきい値の設定が難しい.

より実際的なのは, 複数の解を保存し, それぞれに対して以下の計算で 3 次元復元を行うことであろう. 最終的にどれを選ぶかは, 例えば復元形状を調べて, 直角となるべき物体の頂点が正しく直角になっているかで判定することが考えられる. そのような判断が難しい場合は, 復元形状の再投影誤差が小さいほうを選ぶことが考えられる (後述).

*1 それ以外に, 光軸点が仮定した位置からある程度離れると虚数焦点距離が生じることが知られている³⁾.

5. 運動パラメータの計算

基礎行列 F と焦点距離 f, f' から 2 台のカメラの相対的な並進 t と回転 R を計算する。 $\{t, R\}$ を「運動パラメータ」と呼ぶ。ただし、並進 t は定数倍を除いてしか定まらないので、 $\|t\| = 1$ となる解を求める。この計算は 1980 年代からよく知られている^(6),7)。

- (1) 次の基本行列 E を計算する。

$$E = \text{diag}(1, 1, \frac{f_0}{f}) F \text{diag}(1, 1, \frac{f_0}{f'}) \quad (30)$$

- (2) 対称行列 EE^T の最小固有値に対する単位固有ベクトルを t とする。
(3) 対応点の組 $(x_\alpha, y_\alpha), (x'_\alpha, y'_\alpha)$ を次のベクトルで表す。

$$x_\alpha = \begin{pmatrix} x_\alpha/f \\ y_\alpha/f \\ 1 \end{pmatrix}, \quad x'_\alpha = \begin{pmatrix} x'_\alpha/f' \\ y'_\alpha/f' \\ 1 \end{pmatrix} \quad (31)$$

- (4) 次の条件が成り立てば t の符号を換える。

$$\sum_{\alpha=1}^N |t, x_\alpha, Ex'_\alpha| < 0 \quad (32)$$

ただし、 $|a, b, c|$ はベクトル a, b, c のスカラ三重積である。

- (5) 行列 $-t \times E$ を次のように特異値分解する。

$$-t \times E = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V^T \quad (33)$$

ただし、ベクトル t と行列 E の積 $t \times E$ は t と E の各列とのベクトル積を列とする行列である。

- (6) 回転行列 R を次のように計算する。

$$R = U \text{diag}(1, 1, \det(UV^T)) V^T \quad (34)$$

並進 t はステップ 2 で固有ベクトルとして計算されるが、固有ベクトルには符号が不定である。ステップ 4 の判定はこれを E と適合させるためである^(6),7)。

6. 3次元位置の計算

6.1 三角測量

運動パラメータ $\{t, R\}$ が定まれば、各特徴点の 3 次元位置は各カメラの視点から出発して画像上のその点を通る視線の交点として得られる。これが「三角測量」と呼ばれる処理である。しかし、画像上の対応点に誤差があれば視線が交わるとは限らない。古くは二つの視

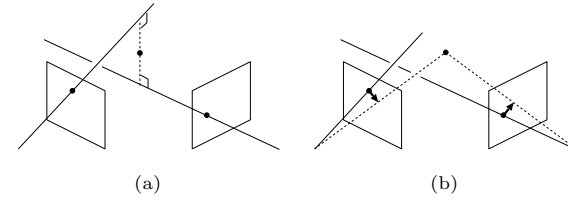


図 5 三角測量。(a) 中点法。(b) 最適な方法。

線を最短に結ぶ線分の中点を計算することが行われていたが(図 5(a))、統計的には最適な方法は、各対応点を視線が交わる位置に「最適」に移動することである(図 5(b))。「最適」とは移動の距離の二乗和(「再投影誤差」)を最小にするという意味であり、正規分布に従う誤差を仮定すると最尤推定となる^(4),7)。このための計算が効率的な金谷ら⁽¹³⁾の方法は次のようになる。まず計算した運動パラメータ $\{t, R\}$ から基礎行列 F を次のように再計算する。

$$F = \text{diag}(1, 1, \frac{f}{f_0})(t \times R) \text{diag}(1, 1, \frac{f'}{f_0}) \quad (35)$$

そして、これを式 (1) の 9 次元ベクトルによって表す。以下の手順は 3.2 節の計算と同一である。ただし、ステップ 3 の u の更新は行わない。そしてステップ 7 では基礎行列 F を返すのではなく、補正位置 $(\hat{x}_\alpha, \hat{y}_\alpha), (\hat{x}'_\alpha, \hat{y}'_\alpha), \alpha = 1, \dots, N$ を返す。この計算は運動パラメータ $\{t, R\}$ の計算において自由焦点距離の方法で得られる f, f' をそのまま用いれば、得られる $(\hat{x}_\alpha, \hat{y}_\alpha), (\hat{x}'_\alpha, \hat{y}'_\alpha)$ と再投影誤差 E は 3.2 節の結果と同一になる。しかし、平均焦点距離の方法の $\tilde{f} = \tilde{f}'$ 、あるいは固定焦点距離の方法の解 $f = f'$ を用いると 3.2 節の値とは必ずしも同一にはならず、再投影誤差 E も一般に増加する。これは自由焦点距離の方法が再投影誤差 E を最小にする F をそのまま分解するためである。以上のことを利用して解の選択ができる。すなわち、平均焦点距離 $\tilde{f} = \tilde{f}'$ と固定焦点距離 $f = f'$ の両方が得られているとき、再投影誤差 E が小さい方を採用することができる*1。

6.2 3次元空間への逆投影

前節のようにして特徴点の補正位置 $(\hat{x}_\alpha, \hat{y}_\alpha), (\hat{x}'_\alpha, \hat{y}'_\alpha)$ が得られれば、その 3 次元位置 $(X_\alpha, Y_\alpha, Z_\alpha)$ は次のように計算される。

- (1) 2 台のカメラの 3×4 投影行列 P, P' を次のように計算する。

$$P = \text{diag}(1, 1, \frac{f_0}{f}) \begin{pmatrix} I & 0 \end{pmatrix}, \quad P' = \text{diag}(1, 1, \frac{f_0}{f'}) \begin{pmatrix} R^T & -R^T t \end{pmatrix} \quad (36)$$

*1 平均焦点距離と固定焦点距離を比較する。自由焦点距離 $f \neq f'$ は常に E が最小なので比較できない。

- (2) 次の連立 1 次方程式を解いて, 3 次元位置 $(X_\alpha, Y_\alpha, Z_\alpha)$, $\alpha = 1, \dots, N$ を計算する (P_{ij}, P'_{ij} はそれぞれ P, P' の (ij) 要素).

$$\begin{pmatrix} \hat{x}_\alpha P_{31} - f_0 P_{11} & \hat{x}_\alpha P_{32} - f_0 P_{12} & \hat{x}_\alpha P_{33} - f_0 P_{13} \\ \hat{y}_\alpha P_{31} - f_0 P_{21} & \hat{y}_\alpha P_{32} - f_0 P_{22} & \hat{y}_\alpha P_{33} - f_0 P_{23} \\ \hat{x}'_\alpha P'_{31} - f_0 P'_{11} & \hat{x}'_\alpha P'_{32} - f_0 P'_{12} & \hat{x}'_\alpha P'_{33} - f_0 P'_{13} \\ \hat{y}'_\alpha P'_{31} - f_0 P'_{21} & \hat{y}'_\alpha P'_{32} - f_0 P'_{22} & \hat{y}'_\alpha P'_{33} - f_0 P'_{23} \end{pmatrix} \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix} = - \begin{pmatrix} \hat{x}_\alpha P_{34} - f_0 P_{14} \\ \hat{y}_\alpha P_{34} - f_0 P_{24} \\ \hat{x}'_\alpha P'_{34} - f_0 P'_{14} \\ \hat{y}'_\alpha P'_{34} - f_0 P'_{24} \end{pmatrix} \quad (37)$$

- (3) $\sum_{\alpha=1}^N \text{sgn}(Z_\alpha) < 0$ なら ($\text{sgn}(x)$ は符号関数であり, $x > 0, x = 0, x < 0$ に応じてそれぞれ 1, 0, -1 を返す), t とすべての $(X_\alpha, Y_\alpha, Z_\alpha)$ の符号を換える.

式 (36) は第 1 カメラのレンズ中心を原点とし, その光軸を Z 軸とする XYZ 座標系を仮定するものであり, 復元された点の座標 $(X_\alpha, Y_\alpha, Z_\alpha)$ はこの座標系に関する記述となる. ステップ 3 は鏡像解 (カメラの後方にある反転した形状) を除去するものである. 式 (32) の判定で t の符号を E に適合させているが, 基本行列 E は式 (30) により基礎行列 F から定まる. しかし, 基礎行列 F は符号が不定であり, この符号が正しくないと鏡像解が生じる. そこで全特徴点が第 1 カメラの前方にあるように t と 3 次元位置 $(X_\alpha, Y_\alpha, Z_\alpha)$ の符号を調節する. このとき $\sum_{\alpha=1}^N Z_\alpha < 0$ でなく, 符号関数 sgn を用いるのは, 正しい復元でも十分遠方の点の視線がほぼ平行になり, 計算誤差によって $-\infty$ に近いところに交点が計算され, $Z_\alpha \approx -\infty$ となって $\sum_{\alpha=1}^N$ がそれに引きずられるのを防ぐためである.

復元した形状にはスケールの不定性がある. これはカメラに近い小さい物体に対してカメラを少し移動しても, カメラから遠い大きい物体に対してカメラを大きく移動しても, 同じ画像が撮影されるためである^{6),7)}. 上記のアルゴリズムでは $\|t\| = 1$ となる解を計算している. もしカメラの移動距離 d が既知なら d を, あるいはある二つの特徴点の空間的位置, 例えば r_1, r_2 の実間隔 d_{12} が既知なら $d_{12}/\|r_2 - r_1\|$ を t とすべての $X_\alpha, Y_\alpha, Z_\alpha$ に掛ければ実寸の復元となる. 第 1 カメラを基準にするのではなく, シーンに固定した $\bar{X}\bar{Y}\bar{Z}$ 世界座標を基準にするには次のように変換すればよい. 第 1 カメラが世界座標の周りに回転行列 \bar{R} だけ回転し, ベクトル \bar{t} だけ並進した位置にあるとすると, 各特徴点のこの世界座標系に関する 3 次元座標 $(\bar{X}_\alpha, \bar{Y}_\alpha, \bar{Z}_\alpha)$ は次のようになる.

$$\begin{pmatrix} \bar{X}_\alpha \\ \bar{Y}_\alpha \\ \bar{Z}_\alpha \end{pmatrix} = \bar{t} + \bar{R} \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix} \quad (38)$$

7. 実 験

図 6 の 2 画像から特徴点を抽出し, 金谷・三島⁹⁾ の方法で 3 次元復元を行った. 特徴点

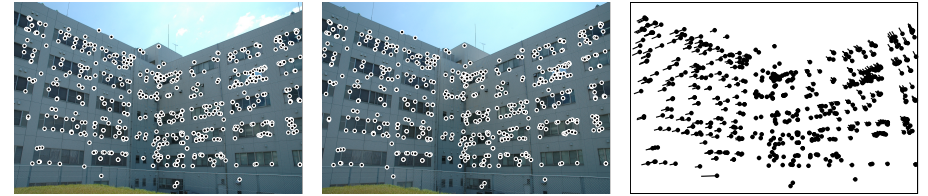


図 6 実画像シーンから抽出した対応点とそのフロー.



図 7 図 6 の 3 次元復元結果を上から眺め, Google Map の画像に重ねたもの. (a) 金谷・三島⁹⁾ の方法. (b) 本システムの平均焦点距離の方法. (c) 本システムの固定焦点距離の方法.

表 1 各方法による焦点距離 (画素) と再投影誤差 (画素). 焦点距離のカメラ校正値は 1156.0 画素.

	金谷・三島 ⁹⁾	平均焦点距離	固定焦点距離
焦点距離	847.8/850.3	1133.5	1148.2
再投影誤差	0.407	0.510	0.521

の抽出と対応付けに SIFT¹⁵⁾ を使い, RANSAC により誤対応を除去した (文献 14) の適用例を参照). 対応付けた特徴点を画像中にマークしている. 左はその対応のフロー (画像フレームを重ねて対応する特徴点を線で結んだもの) である. 図 7(a) はそれらの復元した 3 次元位置を上方から見たものを Google Map の画像上に重ねたものである. 図 7(b), (c) はそれぞれ本システムの平均焦点距離および固定焦点距離による復元である. 表 1 はカメラ校正による焦点距離 (OpenCV で提供されている方法を用いた), およびそれぞれの方法で得られた値と再投影誤差を示す (単位は画素).

図 6 はほぼ注視画像であり, 虚数焦点距離が生じやすいが, この例ではどの場合も実数焦点距離が得られている. 図 7(a) では自由焦点距離の方法で得られた f, f' をそのまま使っている. しかし, 値が不正確なために, 直交すべき角度がやや鋭角になっている. 図 7(b),

(c) の復元形状はほとんど差がないが、再投影誤差は図 7(b) のほうが小さく、焦点距離は図 7(c) のほうがより正しい値になっている。

8. まとめ

本システムと金谷・三島⁹⁾を比較すると、金谷・三島⁹⁾では基礎行列の計算にくりこみ法と最適補正を用いているところを本論文では厳密な最尤推定の計算^{11),12)}に置き換えている。このため理論的には精度が向上する。また、反復の初期値計算に Taubin 法を用いているので、最適化の収束が極めて速くなる。さらに金谷・三島⁹⁾はテンソル式などを用いたかなり複雑な記述であるのに対して、本システムアルゴリズムが明解であり、3次元復元を組み込んだシステムを実装する場合に大きなメリットとなる^{*1}。

金谷・三島⁹⁾ではカメラが注視運動する場合を考慮していなかったが、ユーザーが無意識に2画像を撮影すると、カメラの移動が注視運動になるのが普通である。本システムでは自由焦点距離の方法と固定焦点距離の方法を併用しているので、計算の破綻が少なくなる。ただし、どうしても計算が破綻するカメラ配置が存在し(カメラの対称な配置やカメラの平行移動など)、その場合に計算自体は破綻しなくても精度が著しく低下する。これは2画像しか使わないことの理論的な制約である。ユーザーはこのことを理解して、そのようなカメラ配置を避けるか、避けられない場合は事前にカメラ校正を行う必要がある。

なお、本論文では記述していないが、計算した3次元特徴点を結んで多面体を定義し、各面に原画像からテクスチャマッピングして表示する場合、その物体自身が多面体であれば、表示した多面体とその多面体物体との辺が適合しない(例えば特徴点を結んだ線分が物体の辺と交差する)ことが生じる。そのような場合には特徴点間の結び方を変えて、物体形状に適合するように修正する必要がある。これを自動的に行う方法が文献 16), 17) に記述されている。

謝辞: 本研究の一部は文部科学省科学研究費基盤研究 C (No. 21500167, 21500172) の助成による。

参考文献

- 1) S. Bougnoux, From projective to Euclidean space under any practical situation, a criticism of self calibration, *Proc 6th Int. Conf. Comput. Vision*, January 1998, Bombay, India, pp. 790–796.
- 2) R. I. Hartley, In defense of the eight-point algorithm, *IEEE Trans. Patt. Anal. Mach. Intell.*, **19**-6 (1997-6), 580–593.

- 3) R. Hartley and C. Silpa-Anan, Reconstruction from two views using approximate calibration, *Proc. 5th Asian Conf. Comput. Vision*, January 2002, Melbourne, Australia, Vol.1, pp.338–343.
- 4) R. I. Hartley and P. Sturm, Triangulation, *Comput. Vision Image Understand.*, **68**-2 (1997-11), 146–157.
- 5) R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- 6) 金谷健一, 「画像理解—3次元認識の数理—」, 森北出版, 1990.
- 7) K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice* Elsevier, Amsterdam, the Netherlands, 1996; reprinted, Dover, York, NY, U.S.A., 2005.
- 8) 金谷健一, 松永 力, 基礎行列の分解: 焦点距離の直接的表現, 情報処理学会研究報告, 2000-CVIM-120-7 (2000-1), 49–56.
- 9) 金谷健一, 三島等, 未校正カメラによる2画像からの3次元復元とその信頼性評価, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, **42**-SIG 6 (2001-6), 1–8.
- 10) 金谷健一, 菅谷保之, 制約付きパラメータ推定のための拡張 FNS 法, 情報処理学会研究報告, 2007-CVIM-158-4, (2007-3), 25–32.
- 11) 金谷健一, 菅谷保之, 高ノイズレベルにおける基礎行列の最尤推定, 情報処理学会研究報告, 2007-CVIM-160-9, (2007-9), 49–56.
- 12) 金谷健一, 菅谷保之, 幾何学的当てはめの厳密な最尤推定の統一的計算法, 情報処理学会研究報告, 2008-CVIM-164-1 (2008-9), 17–24.
- 13) 金谷健一, 菅谷保之, 新妻弘崇, 2画像からの三角測量: Hartley vs. 最適補正, 情報処理学会研究報告, 2008-CVIM-162-54 (2008-3), 335–342.
- 14) 金澤 靖, 金谷健一, 大域的な整合性を保証するロバストな画像の対応づけ, 情報処理学会論文誌: CVIM, **44**-SIG 17 (2003-12), 70–77.
- 15) D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vision*, **60**-2 (2004-11), 91–110.
- 16) 村田 正和, 中辻 敦忠, 菅谷 保之, 金谷 健一, 画像からの形状復元に適合した三角網の生成, 情報処理学会研究報告, 2004-CVIM-145-2 (2004-9), 9–16.
- 17) 中辻 敦忠, 菅谷 保之, 金谷 健一, ビデオ画像からの形状復元のための三角網の最適化, 情報処理学会研究報告, 2005-CVIM-149-2 (2005-5), pp. 7–14.
- 18) 菅谷保之, 金谷健一, 基礎行列の高精度計算法とその性能比較, 情報処理学会研究報告, 2006-CVIM-153-32 (2006-3), 207–214.
- 19) 高橋 茂雄, 中辻 敦忠, 金谷健一, 未校正2画像からの3次元復元のための焦点距離計算の安定化, 情報処理学会研究報告, 2003-CVIM-141-12 (2003-11), 79–86.

*1 以下にプログラムを公開している。 <http://www.img.tutkie.tut.ac.jp/programs.html>