

## 3 基礎行列の分解による 3 画像からの 3 次元形状復元

山崎 大輔<sup>1,a)</sup> 金澤 靖<sup>1,b)</sup> 菅谷 保之<sup>1,c)</sup> 金谷 健一<sup>2,d)</sup>

概要：本論文では 3 画像からの 3 次元復元の計算法を示す。これは 3 画像のみが得られている場合のほか、複数画像からのバンドル調整の初期値としても用いることができる。本論文の方法は 3 画像間の対応を求める必要はなく、2 画像ごとの対応から 3 個の基礎行列が計算されればよい。3 個の基礎行列は 3 次元復元のための過剰な情報を含んでいるので、その冗長性を利用して、3 台のカメラの位置や向きや焦点距離を最適に計算する。そして、3 画像で対応している点はその対応から、2 画像のみ対応している点はその対応から 3 次元位置が計算し、シミュレーションにより 2 画像からの復元よりも精度が向上することを示す。さらに実画像例として内視鏡画像を用いる医用応用を示し、3 枚の画像を用いると 2 画像よりも広い範囲の復元ができること、および 2 画像からの復元の計算が破綻する場合も 3 画像から最適な復元ができることを示す。

### 3-D Shape Reconstruction from Three Views by Decomposing Three Fundamental Matrices

DAISUKE YAMASAKI<sup>1,a)</sup> YASUSHI KANAZAWA<sup>1,b)</sup> YASUYUKI SUGAYA<sup>1,c)</sup> KENICHI KANATANI<sup>2,d)</sup>

**Abstract:** This paper presents how we can reconstruct a 3-D shape from three views. The proposed method can be used not only for given three images but also for computing an initial value of bundle adjustment for multiple images. We need not obtain correspondences among the three images; all we need is three fundamental matrices computed from pair-wise correspondences. Since three fundamental matrices have redundant information for 3-D reconstruction, we exploit that redundancy to optimize the positions and orientations of the three cameras and their focal lengths. The 3-D positions of individual points are computed from three views if observed in all of them and from two views if observed only in them. We do simulation to show that the reconstruction accuracy is higher for three views than two views. We also show a medical application using endoscopic images of intestine tracts. We observe that a wider range can be reconstructed using three images and that optimal 3-D reconstruction is possible even when reconstruction from pairs of images fails.

#### 1. まえがき

画像からの 3 次元形状復元は、以前は一部の研究者が自ら開発した高度な手法を用いてデモを行っていたが、誰でも使える形に整備されたソフトウェアが公開され、現在では多くの人によってさまざまな応用に利用されている。現在の主流は「バンドル調整」と呼ばれる方法である [1], [9]。これは複数の画像間の対応点から、各点の 3 次元座標と各画像を撮影したカメラの位置や向きや内部パラメータをすべて未知数として、高次元パラメータ空間を探索するもの

である。その原理は、そのような 3 次元位置をそのようなカメラで撮影したら得られるであろう画像上の位置と観測位置のずれ（「再投影誤差」）を最小にすることである。

これを行う公開ソフトウェアとして Lourakis ら [8] による SBA、およびそれを対応点抽出と組み合わせた Snavely ら [10], [11] の bundler がよく知られている。これらは反復解法であり、探索の初期値はよく対応のとれた 2 画像から計算するのが普通である。これは、2 画像の対応から比較的簡単な計算で対応点の 3 次元位置やカメラパラメータが計算できるからであり、特徴点抽出の誤差を考慮した精度の高い方法が研究されている [12]。

しかし、2 画像からの復元にはさまざまな制約がある。最大の問題は 2 画像を撮影するカメラが「注視の位置」（光軸が交わる位置）にあってはならないことである。しかし、実際の場面では関心のある対象を二カ所から撮影すると、それを注視するように撮影するのが最も自然である。また、注視を避けても、2 画像から得られる情報は必要最

<sup>1</sup> 豊橋技術科学大学情報・知能工学系 Department of Computer Science and Engineering, Toyohashi University of Technology, Japan

<sup>2</sup> 岡山大学大学院自然科学研究科 Department of Computer Science, Okayama University, Japan

a) d-yamasaki@img.cs.tut.ac.jp

b) kanazawa@cs.tut.ac.jp

c) sugaya@iim.cs.tut.ac.jp

d) kanatani@suri.cs.okayama-u.ac.jp

小限であり、未知数と方程式の数が一致する。これは解が解析的に得られるという利点をもたらすが、データに誤差があるとしばしば解が存在しない。これに対する対処法がいろいろ考察されているが [5], 2 画像を用いる限り、この問題は解決できない。

そこで本論文では 3 画像からの形状復元を考える。ただし 3 画像間の対応を求める必要はなく、2 画像ごとの対応から 3 個の基礎行列が計算されればよい。その 3 個の基礎行列から 3 台のカメラの位置や向きや焦点距離を計算する。そして、3 画像で対応している点はその対応から、2 画像のみ対応している点はその対応から 3 次元位置を計算する。

2 画像からの復元と異なり、3 個の基礎行列が得られれば条件が過剰になり、未知数よりも方程式の数が多くなる。そこでこの冗長性を利用し、すべての式がほぼ満たされるように最適化を行う。そしてシミュレーションにより 2 画像からの復元よりも精度がよいことを示す。最後に実画像例として内視鏡画像から 3 次元復元を行う医用応用を示す。

## 2. 基礎行列の計算

2 画像からの復元では光軸点が画像フレームの中心にあり、アスペクト比は 1 で画像の歪みはない、あるいは事前にそのように校正されていると仮定する必要がある [12]。これは実際の応用では大きな制約ではない。3 画像の場合はこれを弱めることが可能であるが、本論文では冗長性による最適化を行うので、2 画像の場合と同じ条件を仮定する。そして、画像のフレーム中心を原点とし、 $x$  軸を上方に、 $y$  軸を右方にとる。これは光軸方向を  $z$  軸とみなしたときに  $xyz$  座標系を右手系にするためであり、カメラ回転の計算に必要となる ( $x$  軸を右方に、 $y$  軸を下方にとってもよい)。

同一シーンを 3 台のカメラで (あるいは 1 台のカメラを移動して) 撮影し、得られる画像をそれぞれ第 0, 1, 2 画像と呼ぶ。第 0 画像の点  $(x, y)$  が第 1 画像の点  $(x', y')$  に対応しているとき、これらを次の 3 次元ベクトルで表す。

$$\mathbf{x} = \begin{pmatrix} x/f_0 \\ y/f_0 \\ 1 \end{pmatrix}, \quad \mathbf{x}' = \begin{pmatrix} x'/f_0 \\ y'/f_0 \\ 1 \end{pmatrix} \quad (1)$$

$f_0$  はオーダをそろえるスケール定数であるが、これをカメラの焦点距離の近似値にとり、「デフォルト焦点距離」と呼ぶ (実験では 600 画素とした)。このときエビ極線方程式は次のように書ける。

$$(\mathbf{x}, \mathbf{F}_{01}\mathbf{x}') = 0 \quad (2)$$

本論文ではベクトル  $a, b$  の内積を  $(a, b)$  と書く。 $\mathbf{F}_{01}$  が「基礎行列」と呼ばれ、2 画像間に 8 組以上の対応点を与えられれば計算できる (理論的には 7 組でも可能であるが一意的とは限らない) [2]。本論文では基礎行列の計算に、対応点抽出の誤差を考慮した現時点で最も優れた方法と考え

られる「拡張 FNS 法」に基づく方法 [6] を用いる。同様にして、第 0, 1 画像間の基礎行列  $\mathbf{F}_{02}$ 、第 1, 2 画像間の基礎行列  $\mathbf{F}_{12}$  を計算する。

第 0 カメラの座標系 (レンズ中心が原点、光軸が  $Z$  軸) を世界座標系とみなす。そして第 1, 2 カメラのレンズ中心を  $t_1, t_2$ 、第 0 カメラに相対的な回転を  $\mathbf{R}_1, \mathbf{R}_2$  とする。3 台のカメラの焦点距離をそれぞれ  $f, f', f''$  とすると、計算に誤差がなければ基礎行列  $\mathbf{F}_{01}, \mathbf{F}_{02}, \mathbf{F}_{12}$  は次のように表される [2]。

$$\begin{aligned} \mathbf{F}_{01} &\simeq \text{diag}(1, 1, \frac{f}{f_0})\mathbf{t}_1 \times \mathbf{R}_1 \text{diag}(1, 1, \frac{f'}{f_0}), \\ \mathbf{F}_{02} &\simeq \text{diag}(1, 1, \frac{f}{f_0})\mathbf{t}_2 \times \mathbf{R}_2 \text{diag}(1, 1, \frac{f''}{f_0}), \\ \mathbf{F}_{12} &\simeq \text{diag}(1, 1, \frac{f'}{f_0})(\mathbf{R}_1^\top(\mathbf{t}_2 - \mathbf{t}_1)) \times (\mathbf{R}_1^\top \mathbf{R}_2) \\ &\quad \text{diag}(1, 1, \frac{f''}{f_0}) \end{aligned} \quad (3)$$

記号  $\simeq$  は定数倍を除いて等しいことを表し、 $\text{diag}(a, b, c)$  は  $a, b, c$  を対角要素とする対角行列である。またベクトル  $v$  と行列  $A$  に対して  $v \times A$  は  $v$  と  $A$  の各列のベクトル積を列とする行列である。課題は式 (3) を用いて、与えられた基礎行列  $\mathbf{F}_{01}, \mathbf{F}_{02}, \mathbf{F}_{12}$  から  $f, f', f'', t_1, t_2, \mathbf{R}_1, \mathbf{R}_2$  を計算することである。

## 3. 焦点距離の計算

### 3.1 一般の場合

焦点距離  $f, f', f''$  を直接に計算する代わりに次の  $x, y, z$  を計算する。

$$x = \left(\frac{f_0}{f}\right)^2 - 1, \quad y = \left(\frac{f_0}{f'}\right)^2 - 1, \quad z = \left(\frac{f_0}{f''}\right)^2 - 1 \quad (4)$$

$\mathbf{k} = (0, 0, 1)^\top$  と置くと、 $x, y$  は 4 次多項式

$$\begin{aligned} K_{01}(x, y) &= (\mathbf{k}, \mathbf{F}_{01}\mathbf{k})^4 x^2 y^2 + 2(\mathbf{k}, \mathbf{F}_{01}\mathbf{k})^2 \|\mathbf{F}_{01}^\top \mathbf{k}\|^2 x^2 y \\ &\quad + 2(\mathbf{k}, \mathbf{F}_{01}\mathbf{k})^2 \|\mathbf{F}_{01}\mathbf{k}\|^2 x y^2 + \|\mathbf{F}_{01}^\top \mathbf{k}\|^4 x^2 + \|\mathbf{F}_{01}\mathbf{k}\|^4 y^2 \\ &\quad + 4(\mathbf{k}, \mathbf{F}_{01}\mathbf{k})(\mathbf{k}, \mathbf{F}_{01}\mathbf{F}_{01}^\top \mathbf{F}_{01}\mathbf{k}) x y + 2\|\mathbf{F}_{01}\mathbf{F}_{01}^\top \mathbf{k}\|^2 x \\ &\quad + 2\|\mathbf{F}_{01}^\top \mathbf{F}_{01}\mathbf{k}\|^2 y + \|\mathbf{F}_{01}\mathbf{F}_{01}^\top\|^2 - \frac{1}{2} \left( (\mathbf{k}, \mathbf{F}_{01}\mathbf{k})^2 x y \right. \\ &\quad \left. + \|\mathbf{F}_{01}^\top \mathbf{k}\|^2 x + \|\mathbf{F}_{01}\mathbf{k}\|^2 y + \|\mathbf{F}_{01}\|^2 \right)^2 \end{aligned} \quad (5)$$

を最小にする値であり、その最小値は 0 であることが知られている [5]。同様に  $x, z$  の 4 次多項式  $K_{02}(x, z)$ 、 $y, z$  の 4 次多項式  $K_{03}(y, z)$  を定義すると、 $x, z$  は  $K_{02}(x, z)$  の最小値 0 をとり、 $y, z$  は  $K_{03}(y, z)$  の最小値 0 をとる。したがって、 $K_{01}(x, y)$  から  $x, y$  が、 $K_{02}(x, z)$  から  $x, z$  が、 $K_{03}(y, z)$  から  $y, z$  が定まり、しかも解は解析的な式 (「Bougnoux の式」) によって定まる [5]。しかし、基礎行列に誤差があると一般に解が互いに矛盾する。そこで

$$F(x, y, z) = K_{01}(x, y) + K_{02}(x, z) + K_{03}(y, z) \quad (6)$$

を最小にする  $x, y, z$  を解として採用する。最小化には

ニュートン法を用い、初期値は  $x = y = z = 0$  とする ( $f = f' = f'' = f_0$  に相当).  $x, y, z$  が求まれば、式 (4) より  $f, f', f''$  は次のようになる.

$$f = \frac{f_0}{\sqrt{1+x}}, \quad f' = \frac{f_0}{\sqrt{1+y}}, \quad f'' = \frac{f_0}{\sqrt{1+z}} \quad (7)$$

このとき  $x, y, z$  のどれかが  $-1$  以下になると計算が破綻する. これが「虚数焦点距離の問題」であり、2 画像の場合に解析的な解を用いるとよく生じる. その一つの原因は、解析的な解の導出に最小値が 0 となる条件が使われていることにある (実際には 0 とは限らない). 上の最適化ではこの条件を用いていないので虚数焦点距離の問題が緩和されると期待される.

### 3.2 焦点距離が等しい場合

前節の方法は 3 台のカメラの焦点距離をそれぞれ別々の変数とみなすものであり、文献 [5], [12] では「自由焦点距離の方法」と呼んでいる. しかし、実際の場合ではカメラのズームを変えず撮影することが多い. 3 枚の画像の撮影の焦点距離が等しいことが分かっている場合は  $K_{01}(x) = K_{01}(x, x)$ ,  $K_{02}(x) = K_{02}(x, x)$ ,  $K_{12}(x) = K_{12}(x, x)$  と置いて

$$F(x) = F_{01}(x) + F_{02}(x) + F_{12}(x) \quad (8)$$

をニュートン法で最小化し、求めた  $x$  から式 (7) によって  $f (= f' = f'')$  を定める. これを「固定焦点距離の方法」と呼ぶ [5], [12]. あるいは、まず自由焦点距離の方法で  $x, y, z$  を計算してから平均してもよい. このとき、単純平均ではなく、式 (6) の増加量が最小になるように  $x, y, z$  を一定値  $\bar{x}$  に補正するのが合理的である. 式 (6) を最小値を与える点の近傍で 2 次関数で近似すると解は次のようになる.

$$\begin{aligned} \bar{x} = & \left( (F_{xx} + F_{xy} + F_{yz})x + (F_{yx} + F_{yy} + F_{yz})y \right. \\ & \left. + (F_{zx} + F_{zy} + F_{zz})z \right) / \left( F_{xx} + F_{yy} + F_{zz} \right. \\ & \left. + 2(F_{xy} + F_{yz} + F_{zx}) \right) \end{aligned} \quad (9)$$

ただし、 $F_{xx}, F_{xy}, \dots$  は  $F(x, y, z)$  の 2 階微分の求めた  $(x, y, z)$  での値である. この  $\bar{x}$  から (7) によって  $f (= f' = f'')$  を定める. これを「平均焦点距離の方法」と呼ぶ [12].

3 台のカメラの焦点距離が等しい場合は自由焦点距離、固定焦点距離、平均焦点距離の三つの方法が可能であり、どれがよいとは言えないが、データによっては虚数焦点距離が生じるので、複数の方法があれば実数値が選べる可能性が増える.

### 3.3 注視の配置

次のことが知られている [5]. 第 0, 1 画像が注視の配置のとき式 (5) の  $K_{01}(x, y)$  は最小値が退化して  $xy$  面上の曲線となるので、解が定まらない. このとき、固定焦点距離を仮定すれば、その曲線と直線  $x = y$  の交点として解が定

まる. しかし、カメラ配置が“二等辺三角形”(注視点から 2 台のカメラのレンズ中心まで等距離)のときその曲線が直線  $x = y$  に接するので解が定まらない. 第 0, 2 画像, 第 1, 2 画像についても同様である.

本論文のように式 (8) を最小化すると、ある 2 画像が注視の配置でも、注視でない 2 画像があれば解が一意的に定まる. しかし、3 画像が同時に注視の配置であれば、式 (8) の  $F(x, y, z)$  の最小値においてヘッセ行列の行列式が 0 となるので解が定まらない. どれも二等辺三角形の配置でなければ、理論的には固定焦点距離を仮定して解が定まるが、実際の場合では注視の配置がほぼ二等辺三角形になることが多く、厳密には二等辺三角形の配置でなくても解が非常に不安定になる.

以上のことから、2 画像が注視の配置となって 2 画像のみからは復元できない場合でも、3 枚目の画像があれば、すべてが同時に注視でなければ解が求まることが分かる.

## 4. 並進の計算

2 画像間の基礎行列から 2 台のカメラ間の並進が計算できる [12]. したがって、三つの基礎行列  $F_{01}, F_{02}, F_{12}$  からそれぞれのカメラ間の並進が計算できる. しかし、基礎行列から計算される並進はスケールが不定であり、通常は単位ベクトルに正規化される. 画像を用いる限り、絶対的なスケールは復元できない. ただし、3 個の並進の相対的なスケールは「三つの並進ベクトルの作る三角形が閉じる」という条件から定めることができる. しかし、基礎行列に誤差があると、それぞれから計算した並進ベクトルが厳密には同一平面上にはなく、閉じた三角形を作ることができない. そこで、三つの並進ベクトルの三角形が閉じるように 3 個の基礎行列から最適に並進を計算する. しかし、そのためにはカメラの回転既知でなければならず、並進のみを独立に計算することができない. そこで、本論文では並進と回転を交互に求める反復を行う.

まず、計算した焦点距離  $f, f', f''$  を用いて次のように「基本行列」 $E_{01}, E_{02}, E_{12}$  を計算する.

$$\begin{aligned} E_{01} &= \text{diag}(1, 1, \frac{f_0}{f}) F_{01} \text{diag}(1, 1, \frac{f_0}{f'}) \\ E_{02} &= \text{diag}(1, 1, \frac{f_0}{f'}) F_{02} \text{diag}(1, 1, \frac{f_0}{f''}) \\ E_{12} &= \text{diag}(1, 1, \frac{f_0}{f''}) F_{12} \text{diag}(1, 1, \frac{f_0}{f'}) \end{aligned} \quad (10)$$

これは誤差がなければ式 (3) より次の関係を満たす.

$$E_{01} \simeq t_1 \times R_1, \quad E_{02} \simeq t_2 \times R_2, \quad E_{12} \simeq t_{12} \times R_1^\top R_2 \quad (11)$$

ただし

$$t_{12} = R_1^\top (t_2 - t_1) \quad (12)$$

は第 1 カメラから見た第 2 カメラのレンズ中心であり、こ

のように表せることが“三角形が閉じる”条件である。しかし、この条件を使うには回転  $R_1$  を知らなければならない。ところが回転を計算するには並進を知る必要がある。そこで次のような交互操作を行う。まず式 (11) から次の関係が分かる。

$$E_{01}^\top t_1 = 0, \quad E_{02}^\top t_2 = 0, \quad E_{12}^\top t_{12} = 0 \quad (13)$$

これは基礎行列が正しく計算されている場合である。そこで次の最小化によって  $t_1, t_2, t_{12}$  を計算する。

$$\begin{aligned} \|E_{01}^\top t_1\|^2 &= (t_1, (E_{01} E_{01}^\top) t_1) \rightarrow \min, \\ \|E_{02}^\top t_2\|^2 &= (t_2, (E_{02} E_{02}^\top) t_2) \rightarrow \min, \\ \|E_{12}^\top t_{12}\|^2 &= (t_{12}, (E_{12} E_{12}^\top) t_{12}) \rightarrow \min \end{aligned} \quad (14)$$

解はそれぞれ  $E_{01} E_{01}^\top, E_{02} E_{02}^\top, E_{12} E_{12}^\top$  の最小固有値に対する単位固有ベクトルで与えられる。この段階では  $t_1, t_2, t_{12}$  のスケールは不定であり、すべて仮に単位ベクトルとして計算する。さらに  $t_1, t_2, t_{12}$  の符号も不定である。そこで次の不等式が成り立つように符号を選ぶ [12]。

$$\begin{aligned} \sum_{\alpha} |t_1, x_{\alpha}, E_{01} x'_{\alpha}| > 0, \quad \sum_{\alpha} |t_2, x_{\alpha}, E_{02} x''_{\alpha}| > 0, \\ \sum_{\alpha} |t_{12}, x'_{\alpha} E_{12} x''_{\alpha}| > 0 \end{aligned} \quad (15)$$

ただし、 $x_{\alpha}, x'_{\alpha}, x''_{\alpha}$  は対応点の座標を式 (1) のようにベクトルで表したものである。添字は対応点の番号であり、各 2 画像間に対応する点について総和する。式 (15) は  $E_{01}, E_{02}, E_{12}$  の符号が正しいと仮定したときに、ほとんどの対応点がどのカメラに関しても“同じ側”にあるという条件である（式 (2) のエビ極線方程式は点がカメラの後にあっても成り立つ）。

このようにして  $t_1, t_2, t_{12}$  の方向が定めれば、回転  $R_1, R_2$  を定めることができる（次節）。回転  $R_1$  が定めれば、これを式 (12) に代入して、式 (14) ではなく、次式を最小にするように  $t_1, t_2$  を再計算する。

$$\|E_{01}^\top t_1\|^2 + \|E_{02}^\top t_2\|^2 + \|E_{12}^\top t_{12}\|^2 = \left( \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, G \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \right) \quad (16)$$

ただし  $6 \times 6$  行列  $G$  を次のように置いた。

$$G = \begin{pmatrix} E_{01} E_{01}^\top + R_1 E_{12} E_{12}^\top R_1^\top & -R_1 E_{12} E_{12}^\top R_1^\top \\ -R_1 E_{12} E_{12}^\top R_1^\top & E_{02} E_{02}^\top + R_1 E_{12} E_{12}^\top R_1^\top \end{pmatrix} \quad (17)$$

式 (16) を最小にする解は  $G$  の最小固有値に対する単位固有ベクトル  $\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$  である。これは  $\|t_1\|^2 + \|t_2\|^2 = 1$  とするスケールで計算されている。符号は  $t_1, t_2$  か  $-t_1, -t_2$  のうち、再計算前の  $t_1, t_2$  の向きのものを選ぶ。

このように  $t_1, t_2$  を再計算したら、式 (12) によって  $t_{12}$

を再計算する。この再計算した  $t_1, t_2, t_{12}$  から再び  $R_1, R_2$  を計算する（次節）。そして得られた  $R_1$  を用いて式 (17) の  $G$  の最小固有値に対する単位固有ベクトルを計算して、 $t_1, t_2$  を更新する。これを収束するまで反復する。

## 5. 回転の計算

並進  $t_1, t_2, t_{12}$  の方向が与えられたとき、式 (11) を満たす  $R_1, R_2$  を計算する。そのために次式を最小にする  $R_1, R_2$  を求める。

$$\|E_{01} - t_1 \times R_1\|^2 + \|E_{02} - t_2 \times R_2\|^2 + \|E_{12} - t_{12} \times R_1^\top R_2\|^2 \quad (18)$$

この最小化は

$$\begin{aligned} K_{01} &= -t_1 \times E_{01}, \quad K_{02} = -t_2 \times E_{02}, \\ K_{12} &= -t_{12} \times E_{12} \end{aligned} \quad (19)$$

と置くと、次式の最大化に等価である [4]。

$$J = \text{tr}[K_{01}^\top R_1] + \text{tr}[K_{02}^\top R_2] + \text{tr}[K_{12}^\top R_1^\top R_2] \quad (20)$$

これを最大化するには次の定理を用いる [4]。

定理： $K = V \Lambda U^\top$  と特異値分解したとき、 $\text{tr}[K^\top R]$  を最大化する回転  $R$  は  $R = V \text{diag}(1, 1, \det(VU^\top)) U^\top$  で与えられる。

まず  $R_1$  を  $\text{tr}[K_{01}^\top R_1]$  を最大化するものを選ぶ。式 (20) は次のように書き直せる。

$$J = \text{tr}[K_{01}^\top R_1] + \text{tr}[(K_{02} + R_1 K_{12})^\top R_2] \quad (21)$$

そこで、定めた  $R_1$  から  $\text{tr}[(K_{02} + R_1 K_{12})^\top R_2]$  を最大化する  $R_2$  を計算する。式 (20) は次のように書き直せる。

$$J = \text{tr}[K_{02}^\top R_2] + \text{tr}[(K_{01} + R_2 K_{12}^\top)^\top R_1] \quad (22)$$

そこで、定めた  $R_2$  から  $\text{tr}[(K_{01} + R_2 K_{12}^\top)^\top R_1]$  を最大化する  $R_1$  を計算する。 $R_1$  が定まると再び  $\text{tr}[(K_{02} + K_{12} R_1)^\top R_2]$  を最大化する  $R_2$  を計算する。反復ごとに  $J$  が増加するので、収束するまでこれを続ける。

この計算を実行するには問題が生じる。それは基礎行列  $F_{01}, F_{02}, F_{12}$  が式 (3) の記号  $\simeq$  から分かるように、符号の不定性があるため式 (10) で定めた基本行列  $E_{01}, E_{02}, E_{12}$  に符号の不定性があることである。絶対的な符号はそもそも不定であるが、3 個の基本行列の符号間に矛盾があってはならない。式 (15) で定まる並進は  $E_{01}, E_{02}, E_{12}$  の符号が正しいことを仮定しているため、基本行列の符号を換えると対応する並進の符号も換える必要がある。そこで仮に  $E_{01}$  の符号が正しいと仮定して、 $E_{02}, E_{12}$  およびそれから計算される  $t_2$  や  $K_{12}$  の符号を適合させる。これは、誤差がなければ  $E_{12}^\top R_1^\top (t_2 - t_1) = 0, E_{12} \simeq t_{12} \times R_1^\top R_2$  であることから、誤差があることを考慮して次の規則を適用する。

- $\|E_{12}^T R_1^T (t_2 - t_1)\| > \|E_{12}^T R_1^T (t_2 + t_1)\|$  であれば  $t_2$ ,  $E_{02}$  の符号を換える .
- $\|E_{12} - t_{12} \times R_1^T R_2\| > \|E_{12} + t_{12} \times R_1^T R_2\|$  であれば  $K_{12}$  の符号を換える .

## 6. 特徴点位置の補正

このようにして並進  $t_1, t_2$  と回転  $R_1, R_2$  が定まったら, 次のように基本行列を再計算する .

$$\begin{aligned} E_{01} &= t_1 \times R_1, & E_{02} &= t_2 \times R_2, \\ E_{12} &= \left( R_1^T (t_2 - t_1) \right) \times R_1^T R_2 \end{aligned} \quad (23)$$

そして, 各画像上の特徴点位置がこれらの基本行列が定めるエッジ極線条件を満たすように最適に補正する . この処理は 2 画像の場合は文献 [6], [7], [12] に示されているが, 3 画像に拡張すると次のようになる . まず, 3 画像上の対応点  $(x, y)$ ,  $(x', y')$ ,  $(x'', y'')$  を式 (1) のようにベクトル  $x, x', x''$  で表す . ただし, 式 (1) 中の  $f_0$  をそれぞれ計算した焦点距離  $f, f', f''$  に置き換える . そして次の計算を行う .

(1)  $E_0 = \infty$  (十分大きい数),  $\hat{x} = x, \hat{x}' = x', \hat{x}'' = x'', \tilde{x} = \tilde{x}' = \tilde{x}'' = 0$  と置く .

(2) 次の連立 1 次方程式を解いて  $\lambda, \lambda', \lambda''$  を定める .

$$\begin{pmatrix} (P_k E_{01} \hat{x}', E_{01} \hat{x}') + (P_k E_{01}^T \hat{x}, E_{01}^T \hat{x}) \\ (P_k E_{01} \hat{x}', E_{02} \hat{x}'') \\ (P_k E_{01}^T \hat{x}, E_{12} \hat{x}'') \\ (P_k E_{02} \hat{x}'', E_{01} \hat{x}') \\ (P_k E_{02} \hat{x}'', E_{02} \hat{x}'') + (P_k E_{02}^T \hat{x}, E_{02}^T \hat{x}) \\ (P_k E_{02}^T \hat{x}, E_{12} \hat{x}') \\ (P_k E_{12} \hat{x}'', E_{01}^T \hat{x}) \\ (P_k E_{12}^T \hat{x}', E_{02}^T \hat{x}) \\ (P_k E_{12} \hat{x}'', E_{12} \hat{x}'') + (P_k E_{12}^T \hat{x}', E_{12}^T \hat{x}') \end{pmatrix} \begin{pmatrix} \lambda \\ \lambda' \\ \lambda'' \end{pmatrix} = \begin{pmatrix} (\hat{x}, E_{01} \hat{x}') + (\tilde{x}, E_{01} \hat{x}') + (\tilde{x}', E_{01}^T \hat{x}) \\ (\hat{x}, E_{02} \hat{x}'') + (\tilde{x}, E_{02} \hat{x}'') + (\tilde{x}'', E_{02}^T \hat{x}) \\ (\hat{x}', E_{12} \hat{x}'') + (\tilde{x}', E_{12} \hat{x}'') + (\tilde{x}'', E_{12}^T \hat{x}') \end{pmatrix} \quad (24)$$

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

$$\begin{aligned} \tilde{x} &\leftarrow \lambda P_k E_{01} \hat{x}' + \lambda' P_k E_{02} \hat{x}'', \\ \tilde{x}' &\leftarrow \lambda P_k E_{01}^T \hat{x} + \lambda'' P_k E_{12} \hat{x}'', \\ \tilde{x}'' &\leftarrow \lambda' P_k E_{02}^T \hat{x} + \lambda'' P_k E_{12}^T \hat{x}' \end{aligned} \quad (25)$$

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

$$\hat{x} \leftarrow x - \tilde{x}, \quad \hat{x}' \leftarrow x' - \tilde{x}', \quad \hat{x}'' \leftarrow x'' - \tilde{x}'' \quad (26)$$

(5) 再投影誤差  $E = \|\tilde{x}\|^2 + \|\tilde{x}'\|^2 + \|\tilde{x}''\|^2$  を計算し,  $E \approx E_0$  なら  $\hat{x}, \hat{x}', \hat{x}''$  を返して終了する . そうでなければ  $E_0 \leftarrow E$  としてステップ 2 に戻る .

得られた  $\hat{x}, \hat{x}', \hat{x}''$  のそれぞれの第 1, 2 成分を  $f, f', f''$  倍

したものが補正された位置  $(\hat{x}, \hat{y}), (\hat{x}', \hat{y}'), (\hat{x}'', \hat{y}'')$  となる . 2 画像間のみ対応している点に対しては文献 [6], [12] のように計算すればよい .

## 7. 3次元位置の計算

計算した焦点距離  $f, f', f''$ , 並進  $t_1, t_2$ , 回転  $R_1, R_2$  から次の投影行列  $P, P', P''$  を計算する .

$$\begin{aligned} P &= \text{diag}(1, 1, \frac{f_0}{f}) \begin{pmatrix} I & 0 \end{pmatrix}, \\ P' &= \text{diag}(1, 1, \frac{f_0}{f'}) \begin{pmatrix} R_1^T & -R_1^T t_1 \end{pmatrix}, \\ P'' &= \text{diag}(1, 1, \frac{f_0}{f''}) \begin{pmatrix} R_2^T & -R_2^T t_2 \end{pmatrix} \end{aligned} \quad (27)$$

3次元位置  $(X_\alpha, Y_\alpha, Z_\alpha)$  と各画像に投影された位置  $(\hat{x}_\alpha, \hat{y}_\alpha), (\hat{x}'_\alpha, \hat{y}'_\alpha), (\hat{x}''_\alpha, \hat{y}''_\alpha)$  は次の関係を満たす . これは前節の補正により厳密に成立する .

$$\begin{pmatrix} x_\alpha/f_0 \\ y_\alpha/f_0 \\ 1 \end{pmatrix} \simeq P \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix}, \quad \begin{pmatrix} x'_\alpha/f_0 \\ y'_\alpha/f_0 \\ 1 \end{pmatrix} \simeq P' \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix}, \quad \begin{pmatrix} x''_\alpha/f_0 \\ y''_\alpha/f_0 \\ 1 \end{pmatrix} \simeq P'' \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix} \quad (28)$$

これから  $(X_\alpha, Y_\alpha, Z_\alpha)$  が次の連立 1 次方程式を解いて定まる .

$$\begin{pmatrix} \hat{x}_\alpha P_{31} - f_0 P_{11} & \hat{x}_\alpha P_{32} - f_0 P_{12} & \hat{x}_\alpha P_{33} - f_0 P_{13} \\ \hat{y}_\alpha P_{31} - f_0 P_{21} & \hat{y}_\alpha P_{32} - f_0 P_{22} & \hat{y}_\alpha P_{33} - f_0 P_{23} \\ \hat{x}'_\alpha P'_{31} - f_0 P'_{11} & \hat{x}'_\alpha P'_{32} - f_0 P'_{12} & \hat{x}'_\alpha P'_{33} - f_0 P'_{13} \\ \hat{y}'_\alpha P'_{31} - f_0 P'_{21} & \hat{y}'_\alpha P'_{32} - f_0 P'_{22} & \hat{y}'_\alpha P'_{33} - f_0 P'_{23} \\ \hat{x}''_\alpha P''_{31} - f_0 P''_{11} & \hat{x}''_\alpha P''_{32} - f_0 P''_{12} & \hat{x}''_\alpha P''_{33} - f_0 P''_{13} \\ \hat{y}''_\alpha P''_{31} - f_0 P''_{21} & \hat{y}''_\alpha P''_{32} - f_0 P''_{22} & \hat{y}''_\alpha P''_{33} - f_0 P''_{23} \end{pmatrix} \begin{pmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{pmatrix} = - \begin{pmatrix} \hat{x}_\alpha P_{34} - f_0 P_{14} \\ \hat{y}_\alpha P_{34} - f_0 P_{24} \\ \hat{x}'_\alpha P'_{34} - f_0 P'_{14} \\ \hat{y}'_\alpha P'_{34} - f_0 P'_{24} \\ \hat{x}''_\alpha P''_{34} - f_0 P''_{14} \\ \hat{y}''_\alpha P''_{34} - f_0 P''_{24} \end{pmatrix} \quad (29)$$

式が未知数より多いので, 適当な 3 行を除いて解いても最小二乗法で解いてもよい . 解は同じである . 2 画像のみ対応している点是对応のない画像に対する行を除いて解けばよい . 最後に

$$\sum_{\alpha=1}^N \text{sgn}(Z_\alpha) < 0 \quad (30)$$

であれば, すべての 3次元位置  $(X_\alpha, Y_\alpha, Z_\alpha)$  の符号を反転する .  $\text{sgn}(x)$  は  $x > 0, x < 0, x = 0$  に応じて  $1, -1, 0$  を返す符号関数である . これは基本的行列  $E_{01}$  の符号が正しいとして計算を進めたからである .  $E_{01}$  の符号を反転する

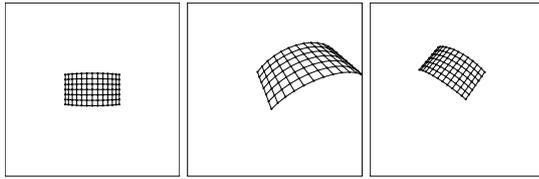


図 1 曲面格子のシミュレーション画像．左から順に第 0, 1, 2 画像．第 0 画像と第 2 画像のカメラは注視に近い位置にある．

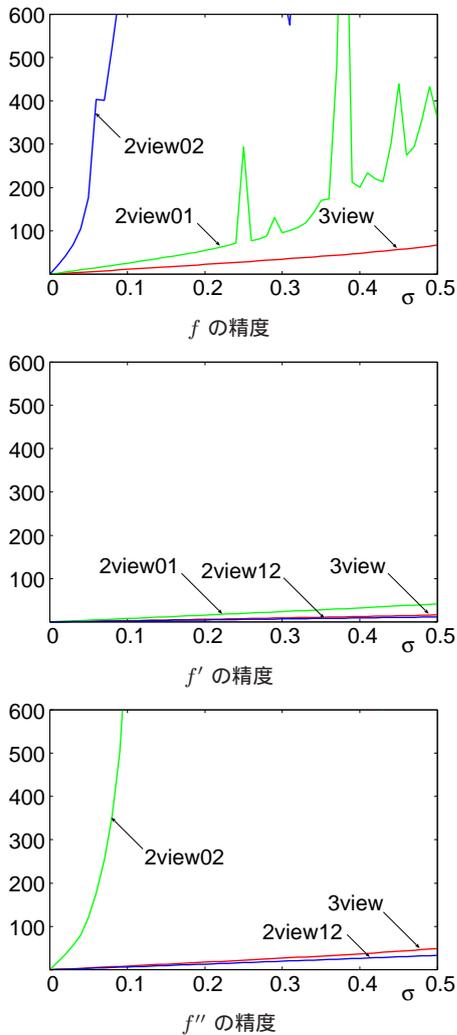


図 2 画像の誤差  $\sigma$  に対する焦点距離の精度．2view01 は第 0, 1 画像から計算した値（他も同様）．3view は 3 画像から計算した値．

と復元される 3 次元形状も反転し，互いに鏡像の関係となる．そこで，カメラの前方にあるほうの形状を選ぶ．これは 2 画像の場合 [12] と同様である．

## 8. シミュレーション実験

図 1 は曲面格子を 3 方向から眺めたシミュレーション画像である．画像サイズは  $800 \times 800$  画素，焦点距離は  $f = f' = f'' = 600$  画素を想定している．各格子点の  $x, y$  座標に期待値 0，標準偏差  $\sigma$  画素の正規乱数を加えて，格子点に対応点として 3 次元復元を行った．

図 2 は計算した焦点距離  $f, f', f''$  の誤差を 2 画像のみ

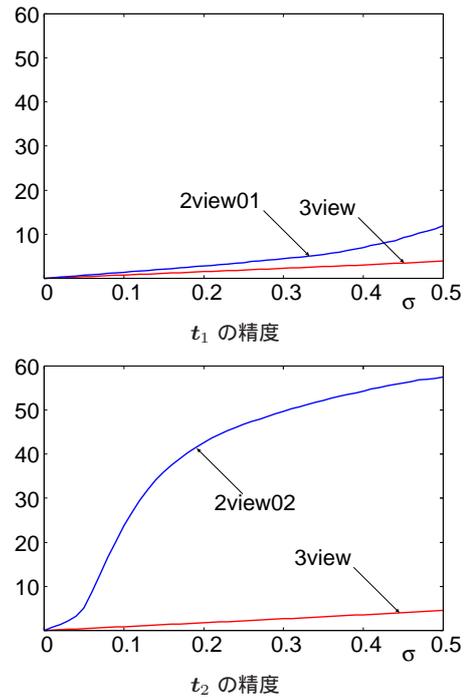


図 3 画像の誤差  $\sigma$  に対する並進の精度（単位は度）．2view01 は第 0, 1 画から像計算した値（他も同様）．3view は 3 画像から計算した値．

から計算した場合と 3 画像から計算した場合を  $\sigma$  についてプロットしたものである．計算には自由焦点距離の方法を用いた．計算精度として， $f$  の真値を  $\bar{f}$  とするときの誤差を  $\Delta f = f - \bar{f}$  とし，その RMS（平方平均二乗）誤差

$$E_f = \sqrt{\frac{1}{K} \sum_{a=1}^K \Delta f_a^2} \quad (31)$$

を評価した．添字  $a$  は  $a$  回目の試行の値を表す，独立な誤差を用い， $K = 10000$  回試行した．ただし，虚数焦点距離が生じた場合は  $f_a = 0$  とみなした．図 1 の第 0, 2 画像のカメラは注視に近い位置にある．このため，第 0, 2 画像から計算した焦点距離  $f, f''$  に大きな誤差が生じている．またプロットが激しく変化しているのは虚数焦点距離が発生していることを示している．しかし，注視の配置を含んでいても 3 枚の画像を用いると，図から分かるように，どの焦点距離も高い精度で計算できることが分かる．このノイズの範囲では 3 枚の画像では虚数焦点距離は発生しなかった．

図 3 は同様にして並進の精度を比較したものである．ただし，並進は絶対的なスケールが不定なので，計算値  $t$  と真値  $\bar{t}$  の成す角度  $\Delta\theta = \cos^{-1}(t, \bar{t}) / \|t\| \cdot \|\bar{t}\|$ （単位は度）の RMS 誤差

$$E_t = \sqrt{\frac{1}{K} \sum_{a=1}^K \Delta\theta_a^2} \quad (32)$$

を評価し，これを 2 画像から計算した場合と 3 画像を用いた場合を比較した．虚数焦点距離が生じて並進が計算でき

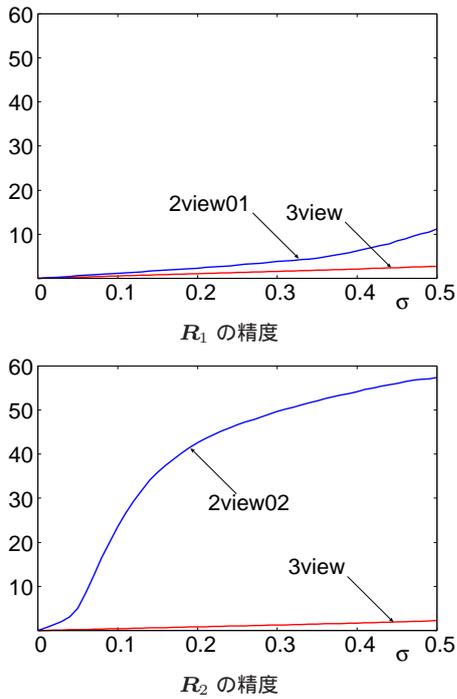


図 4 画像の誤差  $\sigma$  に対する回転の精度 (単位は度). 2view01 は第 0, 1 画像から計算した値 (他も同様). 3view は 3 画像から計算した値.

ない場合は  $\Delta\theta = 90^\circ$  とした. 図 4 は回転の精度を比較である. 計算値を  $R$ , 真値を  $\bar{R}$  とするとき, その相対回転  $R\bar{R}^T$  の回転角を  $\Delta\Omega$  (単位は度) の RMS 誤差

$$E_R = \sqrt{\frac{1}{K} \sum_{a=1}^K \Delta\Omega_a^2} \quad (33)$$

を評価し, これを 2 画像から計算した場合と 3 画像を用いた場合を比較した. 虚数焦点距離が生じて回転が計算できない場合は  $\Delta\Omega = 90^\circ$  とした.

図 3, 図 4 から分かるように, 第 0, 2 画像のカメラが注視に近い位置にあって計算した焦点距離  $f, f''$  の精度が低いため, 並進と回転の精度が低い. しかし, 注視の配置を含んでいても 3 枚の画像を用いると, 図から分かるように, 並進も回転も高い精度で計算できることが分かる.

## 9. 実画像実験

図 5 は腸管モデルの内視鏡実画像である. これはカメラが腸管に沿って後退しながら撮影した連続する 3 フレームである. このようにカメラを前進あるいは後退させながら撮影すると, 注視の配置に近いので虚数焦点距離が生じやすく, 3 次元復元が非常に困難であることが知られている. この 3 画像に対して平井ら [3] の方法で特徴点の抽出と対応付けを行った. これから 3 次元復元した特徴点位置を正面および側面から見た図を図 6(a) に示す. 図 6(b), (c), (d) は比較のために第 0, 1 フレーム, 第 0, 2 フレーム, 第 1, 2 フレームのみから復元した形状である. それぞれその 2 画像で対応している点のみの 3 次元位置を表示している.



図 5 腸管モデルの内視鏡実画像. 腸管に沿って後退しながら撮影した連続する 3 フレーム.

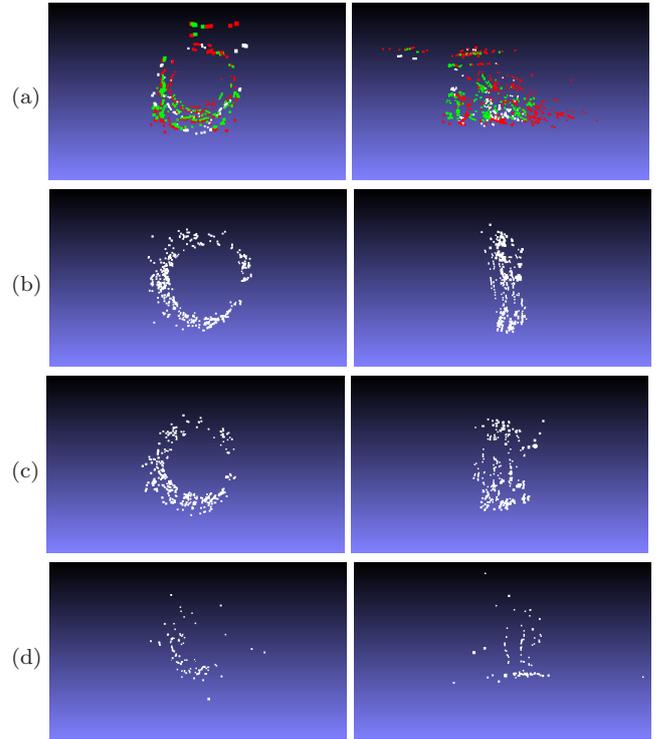


図 6 図 5 から復元した特徴点を正面および側面図. (a) 3 フレームを用いた場合. (b) 第 0, 1 フレームからの復元 (c) 第 0, 2 フレームからの復元 (d) 第 1, 2 フレームからの復元

実画像であるため (a), (b), (c), (d) のどれが最も正しいかは不明であるが, (a) はより多くの特徴点を含んでいるので, 奥行き方向により長い範囲が復元されている. 復元に誤差がなければ (b), (c), (d) の復元位置のスケールを調節し (それぞれの復元はスケールが不定であることに注意), すべてを合成すると (a) の位置に重なるはずである. しかし, 誤差があると一般には重ならない. この意味で, 本論文の 3 画像からの復元は, 3 組の 2 画像からの復元のスケールを自動的に調節し, 誤差を平均化して最適に合成したものと解釈できる.

図 7 は別の 3 フレームであり, 図 8(a) が復元結果である. 図 8(b) は第 1, 2 フレームから復元した結果である. この場合は第 0, 1 フレーム, および第 0, 2 フレームからは虚数焦点距離が生じて復元ができない. しかし, このように 2 画像では復元ができない場合も 3 枚目の画像があれば, すべての 2 画像間の対応点の 3 次元位置が計算できる.



図 7 腸管モデルの内視鏡実画像の別の 3 フレーム .

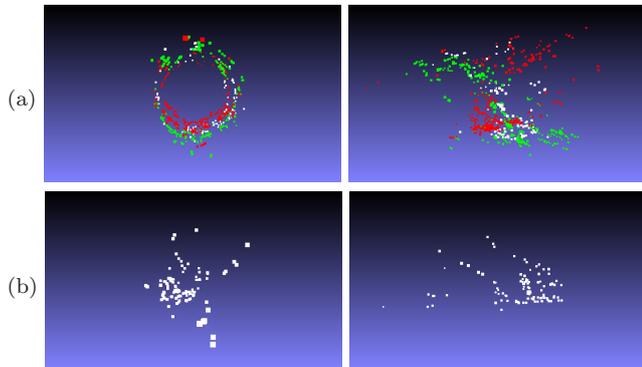


図 8 図 7 から復元した特徴点を正面および側面図 . (a) 3 フレームを用いた場合 . (b) 第 1, 2 フレームからの復元 . 第 0, 1 フレーム, および第 0, 2 フレームからは計算が破綻して復元できない .

## 10. まとめ

本論文では 3 画像からの 3 次元復元の計算法を示した . これは 3 画像のみが得られている場合のほか , 複数画像からのバンドル調整の初期値の計算に用いることができる . 本論文の方法は 3 画像間の対応を求める必要はなく , 2 画像ごとの対応から 3 個の基礎行列が計算されればよい . 3 個の基礎行列は 3 次元復元のための過剰な情報を含んでいるので , 本論文ではその冗長性を利用して , 3 台のカメラの位置や向きや焦点距離を最適に計算する方法を示した . 3 画像で対応している点はその対応から , 2 画像のみ対応している点はその対応から 3 次元位置が計算できる . そして , シミュレーションにより , 注視の位置の 2 画像を含んでも高い精度で復元できることを示した . また実画像例として内視鏡画像を用いる医用応用を示し , 3 枚の画像を用いるとより広い範囲の復元ができること , および 2 画像からの復元の計算が破綻する場合も 3 画像から最適な復元ができることを示した .

謝辞: 本研究の一部は日本学術振興会科学研究費 ( 挑戦的萌芽研究 24650086 , 基盤研究 C 24500202 , 若手研究 B 23700202 ) , 平成 24 年度物質・デバイス領域共同研究課題 ( 2012242 ) の助成によった .

## 参考文献

- [1] 岩元祐輝, 菅谷保之, 金谷 健一, 3 次元復元のためのバンドル調整の実装とその評価, 情報処理学会研究報告, 2011-CVIM-175-19 (2011-1), 1-8.
- [2] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, U.K., 2004.
- [3] 平井克広, 金澤靖, 佐川立昌, 八木康史, 腸管の 3 次元復元のための内視鏡画像間の対応づけ, *Medical Imaging Technology*, **29**-1 (2011-1), 36-46.
- [4] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, U.K., 1993.
- [5] K. Kanatani, A. Nakatsuji and Y. Sugaya, Stabilizing the focal length computation for 3-D reconstruction from two uncalibrated views, *Int. J. Comput. Vis.*, **66**-2 (2006-2), 109-122.
- [6] K. Kanatani and Y. Sugaya, Compact fundamental matrix computation, *IPJS Trans. Comput. Vis. Appl.*, **2** (2010-3), 59-70.
- [7] K. Kanatani, Y. Sugaya, and H. Niitsuma, Triangulation from two views revisited: Hartley-Sturm vs. optimal correction, *Proc. 19th British Machine Vis. Conf.*, September 2008, Leeds, U.K., pp. 173-182.
- [8] M. I. A. Lourakis and A. A. Argyros, SBA: A software package for generic sparse bundle adjustment, *ACM Trans. Math. Software*, **36**-1 (2009-3), 2:1-30.
- [9] 岡谷貴之, バンドルアジャストメント, 情報処理学会研究報告, 2009-CVIM-167-37 (2009-6), 1-16.
- [10] N. Snavely, S. Seitz and R. Szeliski, Photo tourism: Exploring photo collections in 3d, *ACM Trans. Graphics*, **25**08 (1995), 835-846.
- [11] N. Snavely, S. Seitz and R. Szeliski, Modeling the world from internet photo collections, *Int. J. Comput. Vis.*, **80**-22 (2008-11), 189-210.
- [12] 山田健人, 金澤靖, 金谷 健一, 菅谷保之, 2 画像からの 3 次元復元の最新アルゴリズム, 情報処理学会研究報告, 2009-CVIM-168-15 (2009-8), 1-8.