

アフィンエピポーラ幾何の因子分解法による解釈

蔵田 武志[†] 藤木 淳[†] 坂上 勝彦[†]

本稿では、一般アフィン射影 (GAP) モデルの因子分解法の観点から捉えたアフィンエピポーラ幾何についての簡潔な解釈を示す。GAP モデルの因子分解法では、3 画像以上あればユークリッド空間における復元解が一意に得られ、2 画像の場合も 1 自由度を陽に含む復元解が得られる。この 1 自由度は、GAP モデルの 2 画像の位置関係をオイラー角表示したときの第 2 要素と一致する。この性質を用いてアフィンエピポーラ方程式を具体的なパラメータを含む形で求めることができる。実験では、本手法の合成データを用いた評価と、応用の一例として実時系列画像からの密な 3 次元形状復元を行う。

Affine Epipolar Geometry via Factorization Method

TAKESHI KURATA,[†] JUN FUJIKI[†] and KATSUHIKO SAKAUE[†]

We present a simple interpretation of the affine epipolar geometry in terms of a factorization method for the generalized affine projection (GAP) model. The factorization method for the GAP model gives the unique Euclidean estimate from three or more images, and gives the estimate which explicitly includes one degree of freedom from two images. The remaining one degree of freedom is same as the second element of the Euler angle between the two images. Using the property the affine epipolar geometry can be obtained with concrete parameters. We show some experiments using synthetic data and perform the dense shape reconstruction of an object using a real image sequence as one of applications of this method.

1. はじめに

時系列画像からカメラと対象物の相対運動および対象物の 3 次元形状を同時に復元する structure-from-motion 問題は、コンピュータビジョンやその周辺領域において、基本的かつ重要な問題である。

この問題を解くための条件について、まず Ullman が、正射影モデルを仮定した場合にユークリッド空間における解を一意^{*}に求めるには同一平面上にない 4 点が異なる 3 画像で観測されていることが必要であることを最初に示した¹⁷⁾。また Keonderink らは、異なる 2 画像があればある 1 軸回転の自由度を残して運動と形状の復元が可能であることを示した³⁾。この 1 軸回転の自由度は、3 次元空間の回転を $R_z(\theta_1)R_y(\theta_2)R_z(\theta_3)$ とオイラー角表示した場合、2 画像平面間の回転を表すオイラー角の第 2 要素 $R_y(\theta_2)$ に相当する。さらに、Shapiro らはこの 2 画像間の関係のオイラー角表示による解釈を、アフィンエピポーラ幾何の観点からより体系的に示した^{11),19)}。また杉本らはこれを応用

し、まず 3 画像の各 2 画像間のアフィンエピポーラ幾何を weak perspective 射影モデルで復元し、各 2 画像間のオイラー角の第 2 要素を 3 画像間の関係から求める手法を提案した^{14),18)}。

以上の手法は、主に 2 画像または 3 画像を対象としている。一方、多数枚の画像を同時に均等に扱うことにより数値計算的に安定で、簡潔に定式化できるという利点を持つ手法として因子分解法が提案されている。これは、Tomasi らにより正射影モデルにおいて提案され¹⁵⁾、次に Poelman らにより weak perspective (scaled orthographic) 射影モデル、paraperspective 射影モデル⁷⁾の各アフィン射影モデルに拡張された^{8)☆☆}。しかし、前述のとおり、因子分解法の研究目的がこれまで主に多数枚の画像を扱うことであつたため、2 画像間の関係の解釈についてはあまり言及されていない。そこで本稿では、因子分解法の計量拘束で用いる各画像から得られる部分計量に着目することにより、オイラー角表示による 2 画像間の関係

^{*} 本稿では鏡映解の議論はしない。

^{☆☆} 最近では、点特徴から線特徴^{6),10)}に、単体から複数物体¹⁾に、アフィン射影モデルから透視射影モデル^{13),16)}に拡張する研究もなされている。

を一般アフィン射影 (GAP) モデルの因子分解法²⁾の観点から簡潔に解釈し、オイラー角の第 2 要素の 1 自由度を陽に含む形での因子分解法の定式化を行う。さらにその過程で得られるパラメータを用いてアフィンエピポーラ方程式を記述し⁴⁾、因子分解法とアフィンエピポーラ幾何との関係についての、合成データを用いた評価実験、考察を行う。

また、本手法の応用の 1 つとして、実時系列画像から密な 3 次元形状を復元する。因子分解法によって得られる 3 次元形状は各画像で対応付けられた特徴点に関してだけである。この特徴点を増やすことで密な 3 次元形状を復元することは不可能ではないが、運動の未知な多数枚の画像間で精度良く対応付けをするためには、特徴点が 2 次元的なテクスチャ (輝度勾配) を十分に持つ必要がある¹²⁾。よって特徴点を増やせるかどうかは画像の内容に依存する。また、特徴点の対応付けは計算コストの高い 2 次元探索を必要とする。そこで本稿ではまず、十分な画像特徴を持つ特徴点のみを追跡処理により対応付けし、因子分解法により全画像の相対運動 (計量) を復元する。次にその運動情報をもとに本手法によって適当な 2 画像を仮想ステレオ画像対に変換 (rectification) し、画像特徴の 1 次元探索であるステレオマッチングにより密な 3 次元形状を復元する。これらの処理により、各画像の運動復元には全画像を均等に扱える因子分解法を利点を利用して、また、計算コストの低い 1 次元探索による密な 3 次元形状復元が可能となる。

2. 一般アフィン射影 (GAP) モデルの因子分解法

まず、GAP モデルの因子分解法について概説する²⁾。なお、本稿では簡単のために、光軸と画像平面は直交し、それらの交点は既知とし、アスペクト比は 1:1 とする。また、焦点距離は l とする。

2.1 GAP モデルによる因子分解法の記述

第 f 画像上の第 p 特徴点のカメラ座標を $\mathbf{x}_{fp} = (x_{fp}, y_{fp})^T$ とし、それらの重心 $\mathbf{x}_f = (x_f, y_f)^T$ からの各特徴点の相対座標を、 $\mathbf{x}'_{fp} = \mathbf{x}_{fp} - \mathbf{x}_f = (x'_{fp}, y'_{fp})^T$ とする ($f = 1, \dots, F; p = 1, \dots, P$)。また、シーン中の第 p 特徴点のワールド座標を 3 次元ベクトル \mathbf{s}_p で表す。それらの重心を \mathbf{s} とし、 \mathbf{s} からの各特徴点の相対座標を \mathbf{s}'_p とする。GAP モデルでは \mathbf{s} の像は \mathbf{x}_f と一致するため、 \mathbf{s}'_p と \mathbf{x}'_{fp} の関係は、

$$\mathbf{x}'_{fp} = A_f C_f \mathbf{s}'_p \quad (2 \times 3) (3 \times 3)$$

と表すことができる。ここで、 $C_f = (\mathbf{i}_f, \mathbf{j}_f, \mathbf{k}_f)^T$ は画像面 Π_f 上の正規直交基底 $\{\mathbf{i}_f, \mathbf{j}_f\}$ とカメラの光軸方向の単位ベクトル \mathbf{k}_f からなるカメラ座標系の基底行列、 A_f はアフィンカメラモデルを表すアフィン射影行列である。修正計測行列 W' を、

$$W' = \begin{pmatrix} W'_1 \\ \vdots \\ W'_F \end{pmatrix}, \quad W'_f = (\mathbf{x}'_{f1}, \dots, \mathbf{x}'_{fP})$$

と定義する。ここで、 $S' = (\mathbf{s}'_1, \dots, \mathbf{s}'_P)$ とおき、 $M_f = (\mathbf{m}_f, \mathbf{n}_f)^T$ を $M_f = A_f C_f$ によって定めると、

$$W' = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix} S', \quad W'_f = M_f S'$$

と表せるので、以下のように W' をカメラ運動を表す行列 $M (= (M_1^T, \dots, M_F^T)^T)$ と特徴点の 3 次元情報を表す行列 S' の積に因子分解できることが分かる。

$$W' = \underset{(2F \times 3)}{M} \underset{(3 \times P)}{S'} \quad (1)$$

2.2 計量拘束

W' からユークリッド空間における動きと形状を復元するために用いることのできる条件は、カメラ座標系の基底が正規直交基底である、つまり、 $C_f C_f^T = I_3$ ということなので、 M_f について、

$$M_f M_f^T = A_f C_f C_f^T A_f^T = A_f A_f^T \quad (2)$$

という条件が成り立つ。しかし、1 度の操作でこの条件を満たすように W' を分解するのは難しいため、通常、2 つのステップに分けて分解する。

まず第 1 ステップでは、 W' が rank3 であることを利用し、特異値分解などを用いて、

$$W' = \underset{(2F \times P)}{\hat{M}} \underset{(2F \times 3)}{\hat{S}'} \underset{(3 \times P)}{\hat{S}'}$$

のように分解する。なお、 \hat{M} が rank3 であるためには、異なる射影方向 ($\mathbf{m}_f \times \mathbf{n}_f$) を持つ画像が 2 画像あればよく、 \hat{S}' が rank3 であるためには、同一平面上にない最低 4 点の \mathbf{s}'_p があればよい²⁾。

ここで、任意の 3 次の正則行列を A とすると、 $W' = \hat{M} A A^{-1} \hat{S}'$ のような分解も可能であるため、この時点では解は一意に決定できていない。ただし、この分解ですでに線形変換の自由度を持つ運動と形状は復元できていることになる。

第 2 ステップでは、条件 (2) を満たすように A を求め、ユークリッド空間における復元解を一意に決定する。ここで、 $\hat{M} = (\hat{M}_1^T, \dots, \hat{M}_F^T)^T$ 、 $\hat{M}_f = (\hat{\mathbf{m}}_f, \hat{\mathbf{n}}_f)^T$ とし、 $Q = A A^T$ とすると式 (2) から、

$$\hat{M}_f Q \hat{M}_f^T = A_f A_f^T \quad (3)$$

という拘束条件を得る（各アフィンカメラモデルの具体的なパラメータは付録を参照）．これを満たす計量行列 Q は、本質的に異なる 3 枚以上の画像があれば決定できる²⁾．

3. 因子分解法によるオイラー角表示を用いた 2 画像間の関係の解釈

3.1 1 画像から得られる条件

正射影と weak perspective 射影モデルの場合には、 $m_f \times n_f$ と $k_f (= i_f \times j_f)$ の方向が一致するが、paraperspective 射影モデルでは、画像中心と x_f とのずれ（位置効果）により、 $m_f \times n_f$ 方向が変化し、 k_f 方向と必ずしも一致しない．そこで、対象物体と各画像間関係を理解しやすくするために、 $\{i_f, j_f\}$ で張られる画像面 Π_f の代わりに、 $\{m_f, n_f\}$ で張られる仮想画像面 π_f を考える．

$M_f (= (m_f, n_f)^T)$ を特異値分解した結果を、

$$M_f = R_f \Sigma_f V_f^T, \quad (4)$$

(2×3) (2×2) (2×2) (2×3)

$R_f^T R_f = V_f V_f^T = I_2$, $\Sigma_f = \text{diag}\{p_f, q_f\}$ とし、 $V_f = (v_f, w_f)^T$, $\hat{V}_f = (\hat{v}_f, \hat{w}_f)^T$ を、

$$V_f = \Sigma_f^{-1} R_f^T M_f \quad (5)$$

$$\hat{V}_f = \Sigma_f^{-1} R_f^T \hat{M}_f \quad (6)$$

とおく．式 (2), (4) より、

$$M_f M_f^T = A_f A_f^T = R_f \Sigma_f^2 R_f^T$$

となるので、式 (3), (6) を用いて以下のように、 \hat{V}_f における Q の拘束条件を得ることができる．

$$\begin{aligned} \hat{V}_f Q \hat{V}_f^T &= \Sigma_f^{-1} R_f^T \hat{M}_f Q \hat{M}_f^T R_f \Sigma_f^{-1} \\ &= \Sigma_f^{-1} R_f^T A_f A_f^T R_f \Sigma_f^{-1} \\ &= \Sigma_f^{-1} R_f^T R_f \Sigma_f^2 R_f^T R_f \Sigma_f^{-1} \\ &= I_2 \end{aligned} \quad (7)$$

また、式 (5) より、式 (1) は以下のように書き換えられる．

$$W_f' = M_f S' = R_f \Sigma_f V_f S' \quad (8)$$

$V_f V_f^T = I_2$ なので、この式は、「 S' はどのアフィンカメラを仮定しても正規直交基底となる $\{v_f, w_f\}$ により仮想画像面 π_f に正射影され、 $R_f \Sigma_f$ により π_f 上において各アフィンカメラモデル固有の変換がなされる。」と解釈できる．

3.2 2 画像の因子分解法：オイラー角の第 2 要素

図 1 は、ある 2 画像 π_1, π_2 における、 Q が未知な場合（上段）、および Q が既知な場合、つまりユークリッド空間での解が得られた場合（下段）の仮想画像

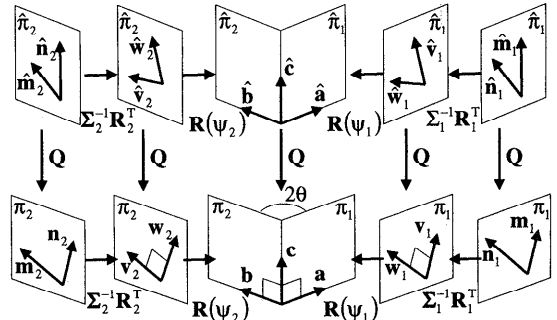


図 1 仮想画像面とその基底
Fig. 1 Virtual image planes and the bases.

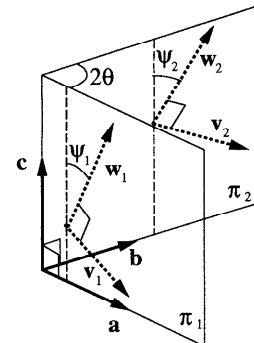


図 2 $R(\psi_1)$ と $R(\psi_2)$
Fig. 2 $R(\psi_1)$ and $R(\psi_2)$.

面とその基底を示している*．この 2 画像 π_1, π_2 から得られる部分計量を考えるために、まず式 (7) を以下のようにベクトルを用いて書き下す．

$$|\hat{v}_f|_Q = |\hat{w}_f|_Q = 1, \langle \hat{v}_f, \hat{w}_f \rangle_Q = 0 \quad (f = 1, 2)$$

図 1 中央および図 2 に示されるように、 π_1 と π_2 の交線方向のベクトルをそれぞれの画像面の基底の 1 つとすることができるので、これら 6 つの条件は互いに独立ではない．そこで、これらと同値となる互いに独立な 5 つの条件に書き直す．

$$\begin{aligned} |\hat{a}|_Q = |\hat{b}|_Q = |\hat{c}|_Q = 1, \\ \langle \hat{b}, \hat{c} \rangle_Q = \langle \hat{c}, \hat{a} \rangle_Q = 0 \end{aligned} \quad (9)$$

ここで、 $\hat{c} \in \pi_1 \cap \pi_2$ であり、 \hat{a}, \hat{b} はそれぞれ上の条件を満たす $\hat{\pi}_1$ と $\hat{\pi}_2$ 上のベクトルとする（図 1）．

Q は 6 自由度あるが、拘束条件 (9) は 5 つであり、 $\langle \hat{a}, \hat{b} \rangle_Q$ についての 1 自由度を決定できない．ここで、 $\langle \hat{a}, \hat{b} \rangle_Q^2 \leq |\hat{a}|_Q^2 |\hat{b}|_Q^2 = 1$ であるので、 $\langle \hat{a}, \hat{b} \rangle_Q = \cos 2\theta$ とおくことができる．これにより、 π_1 と π_2 のなす角であるオイラー角の第 2 要素を 2θ で表したことにな

* ここで扱う 2 画像を仮に第 1 画像、第 2 画像としているが説明の一般性は失わない．

る. ここで, $\hat{P}^T = (\hat{a}, \hat{b}, \hat{c})$ とおくと,

$$\hat{P}Q\hat{P}^T = \begin{pmatrix} 1 & \cos 2\theta & 0 \\ \cos 2\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

であるから,

$$Q = AA^T = \hat{P}^{-1} \begin{pmatrix} 1 & \cos 2\theta & 0 \\ \cos 2\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} (\hat{P}^T)^{-1}$$

となり, $\theta - \frac{\pi}{4} = \varphi$ とおくと,

$$A = \hat{P}^{-1}G(\theta),$$

$$G(\theta) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となる. このとき,

$$M(\theta) = \hat{M}\hat{P}^{-1}G(\theta), \quad S'(\theta) = G(\theta)^{-1}\hat{P}\hat{S}'$$

が成立し, 2 画像の因子分解をオイラー角の第 2 要素を陽に含む形

$$W' = M(\theta)S'(\theta) \tag{10}$$

で表すことができることが分かる.

3.3 オイラー角の第 1 要素, 第 3 要素の復元

\hat{a} , \hat{b} , \hat{c} の計算に計量 Q は必要ない. まず, \hat{c} については,

$$\hat{c} = \alpha_1 \hat{v}_1 + \beta_1 \hat{w}_1 = \alpha_2 \hat{v}_2 + \beta_2 \hat{w}_2$$

とおけるので, $\hat{v}_f \times \hat{w}_f = \hat{r}_f$ とすると,

$$\alpha_1^2 + \beta_1^2 = \alpha_2^2 + \beta_2^2 = 1,$$

$$\alpha_1(\hat{r}_2^T \hat{v}_1) + \beta_1(\hat{r}_2^T \hat{w}_1) = 0,$$

$$\alpha_2(\hat{r}_1^T \hat{v}_2) + \beta_2(\hat{r}_1^T \hat{w}_2) = 0$$

を解けばよい. よって, $\alpha_f, \beta_f (f = 1, 2)$ は以下のようになる.

$$\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = \frac{\pm 1}{\sqrt{(\hat{r}_2^T \hat{w}_1)^2 + (\hat{r}_2^T \hat{v}_1)^2}} \begin{pmatrix} -\hat{r}_2^T \hat{w}_1 \\ \hat{r}_2^T \hat{v}_1 \end{pmatrix}$$

$$\begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = \frac{\pm 1}{\sqrt{(\hat{r}_1^T \hat{w}_2)^2 + (\hat{r}_1^T \hat{v}_2)^2}} \begin{pmatrix} -\hat{r}_1^T \hat{w}_2 \\ \hat{r}_1^T \hat{v}_2 \end{pmatrix} \tag{11}$$

なお, α_f, β_f の符号は \hat{c} の向きが同じになるように選ぶ. \hat{a} , \hat{b} については,

$$\hat{a} = \beta_1 \hat{v}_1 - \alpha_1 \hat{w}_1, \quad \hat{b} = \beta_2 \hat{v}_2 - \alpha_2 \hat{w}_2$$

とおけるが, これらも \hat{c} と同様に解くことができる. まとめると,

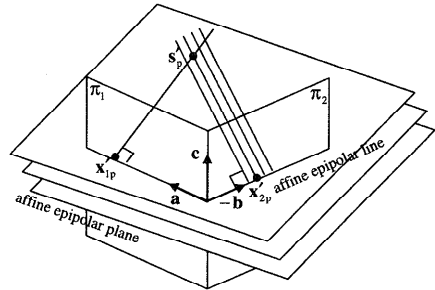


図 3 アフィンエピポーラ拘束
Fig. 3 Affine Epipolar Constraint.

$$(\hat{a}, \hat{c})^T = R(\psi_1)\hat{V}_1, \quad (\hat{b}, \hat{c})^T = R(\psi_2)\hat{V}_2$$

$$R(\psi_f) = \begin{pmatrix} \beta_f & -\alpha_f \\ \alpha_f & \beta_f \end{pmatrix} \tag{12}$$

となる. これらから,

$$(a, c)^T = (\hat{a}, \hat{c})^T A = R(\psi_1)\hat{V}_1 A = R(\psi_1)V_1$$

$$(b, c)^T = (\hat{b}, \hat{c})^T A = R(\psi_2)\hat{V}_2 A = R(\psi_2)V_2 \tag{13}$$

と書けるため, π_1 上で $\{v_1, w_1\}$ を $\{a, c\}$ に重ねる回転 $R(\psi_1)$ と, π_2 上で $\{v_2, w_2\}$ を $\{b, c\}$ に重ねる回転 $R(\psi_2)$, つまりオイラー角の第 1, 第 3 要素を得ることができる (図 2). なお, paraperspective 射影モデルでは, 位置効果による歪みを修正するために R_f^T, Σ_f^{-1} による回転を含む変換が施されるため, $R(\psi_1), R(\psi_2)$ は, それらが修正された後の画像間の回転成分となる.

3.4 アフィンエピポーラ方程式の導出

式 (5), (13) から以下の変換式が得られる.

$$(a, c)^T = R(\psi_1)\Sigma_1^{-1}R_1^T M_1$$

$$(b, c)^T = R(\psi_2)\Sigma_2^{-1}R_2^T M_2 \tag{14}$$

s'_p を π_1, π_2 へ正射影した点はそれぞれ, $(a, c)^T s'_p, (b, c)^T s'_p$ と書ける. ここで, $x'_{fp} = M_f s'_p$ であるから, 式 (14) から以下の関係を導き出せる.

$$(a, c)^T s'_p = R(\psi_1)\Sigma_1^{-1}R_1^T x'_{1p}$$

$$(b, c)^T s'_p = R(\psi_2)\Sigma_2^{-1}R_2^T x'_{2p}$$

図 3 に示すように, π_1, π_2 上において $c^T s'_p$ の値が等しいということがエピポーラ拘束であることから, エピポーラ方程式 $f_1 x'_{1p} + f_2 y'_{1p} = f_3 x'_{2p} + f_4 y'_{2p}$ を以下のように具体的なパラメータを含む形で記述できることが分かる.

$$(0, 1)R(\psi_1)\Sigma_1^{-1}R_1^T x'_{1p} = (0, 1)R(\psi_2)\Sigma_2^{-1}R_2^T x'_{2p} \tag{15}$$

ここで、エピポーラ方程式の復元を最終目的とした場合の処理手順をまとめて示しておく。

- [1] 全画像を用いて W' を \hat{M} と \hat{S}' に分解する。
- [2] エピポーラ方程式を求めたい 2 画像の \hat{M}_f を用いて、付録より R_f, Σ_f を求める。
- [3] R_f^T, Σ_f^{-1} を用いて、式 (6) より $\{\hat{v}_f, \hat{w}_f\}$ を求める。
- [4] $\{\hat{v}_f, \hat{w}_f\}$ を用いて、式 (11), (12) より、 $R(\psi_f)$ を求める。
- [5] 式 (15) よりエピポーラ方程式が得られる。

まず、GAP モデルの場合、[1] ですでに線形変換の自由度を持つ動きと運動と形状が分離でき、3 画像以上が入力された場合は、計量拘束により運動と形状を一意に求めることができる。入力が 2 画像しかない場合も、[2], [3], [4] より $R_f, \Sigma_f^{-1}, R(\psi_f)$ を求めることで、式 (10) に示したオイラー角の第 2 要素を含む形で復元が可能である。また、その $R_f, \Sigma_f^{-1}, R(\psi_f)$ により、位置効果による歪み、スケール調整、エピポーラ線の水平化が可能であり、後述の応用実験で示すステレオマッチングなどに有効な画像の変換 (rectification) が可能である。このように、因子分解法の場合、運動や形状、画像変換のパラメータは、エピポーラ方程式を導出する過程ですでに計算できていることが分かる。

4. 合成データによる評価実験

合成データを用いて、提案手法によるアフィンエピポーラ方程式の復元についての評価実験を行った。

20 フレームからなる合成データを以下のように生成した。

- 特徴点 $\{s_p\}_{p=1}^{40}$ は一辺が 200 mm の立方体内にランダムに一様分布
- カメラはピンホールカメラを仮定し、焦点距離は 35 mm カメラ相当で 84 mm
- 画像は 640×480 (pixel) で各画素は一辺 $8 \mu\text{m}$ の正方形
- オイラー角の各要素は毎フレーム 2 度ずつ回転
- 奥行きは $z_1 = 2000 \text{ mm}$, $z_{20} = 1600 \text{ mm}$, 画像中心からのずれは $u_1 = 115 \text{ pixel}$, $u_{20} = 101 \text{ pixel}$ となるように滑らかに移動

図 4 は、第 1 フレームと第 20 フレームの特徴点の見え方を示している。対象物体が、左下から右上に、また徐々に大きく写されていることが分かる。

復元誤差は、特徴点とそれに対応するエピポーラ線との距離を 2 画像それぞれで計算したものの平均とし、評価には異なる観測誤差を与えて計算した復元誤

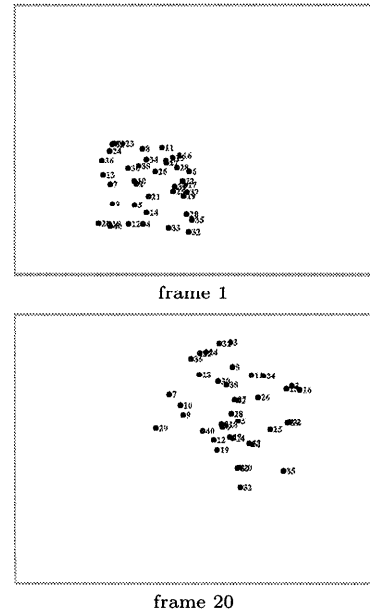


図 4 シミュレーションの入力画像例
Fig. 4 Examples of input images for simulation.

差の 100 回分の平均値を用いた。観測誤差は、平均 0, 標準偏差 (σ) 1 画素および 2 画素のガウシアンノイズ、および観測誤差を加えなかったものの 3 通りを設定した。

エピポーラ方程式の復元には weak perspective 射影モデルを用いた。また、次節のような応用を仮定し、全部で 40 点ある特徴点を、

データ A: 20 点 エピポーラ方程式の復元と誤差評価用。観測誤差を付加。

データ B: 20 点 誤差評価用。観測誤差なし。

の 2 グループに分け、観測誤差はデータ A のみに加えられた。これにより、データ A の誤差評価では、それ自身で復元したエピポーラ方程式へのあてはまり具合を評価でき、データ B では、新たに特徴点 (画像特徴) を対応付けるために用いられるエピポーラ拘束の適切さを評価できる。エピポーラ方程式の復元対象としては、連続する 2 画像計 19 対を用いた。

これらはすべて、3.4 節 [1] の W' の分解による \hat{M} の計算にエピポーラ方程式を復元する 2 画像のみを用いた場合と、全 20 画像を用いた (\hat{M} の最小自乗解を求めた) 場合について評価し、参考のために従来法との比較も行った。なお、従来法は、文献 19) の手法を実装した Zhang によるフリーウェアの AffineF²⁰⁾ を使用した。

図 5 に (a) 観測誤差なし, (b) $\sigma = 1$, (c) $\sigma = 2$ の結果を示す。グラフの “2frame”, “All” は W' の分解

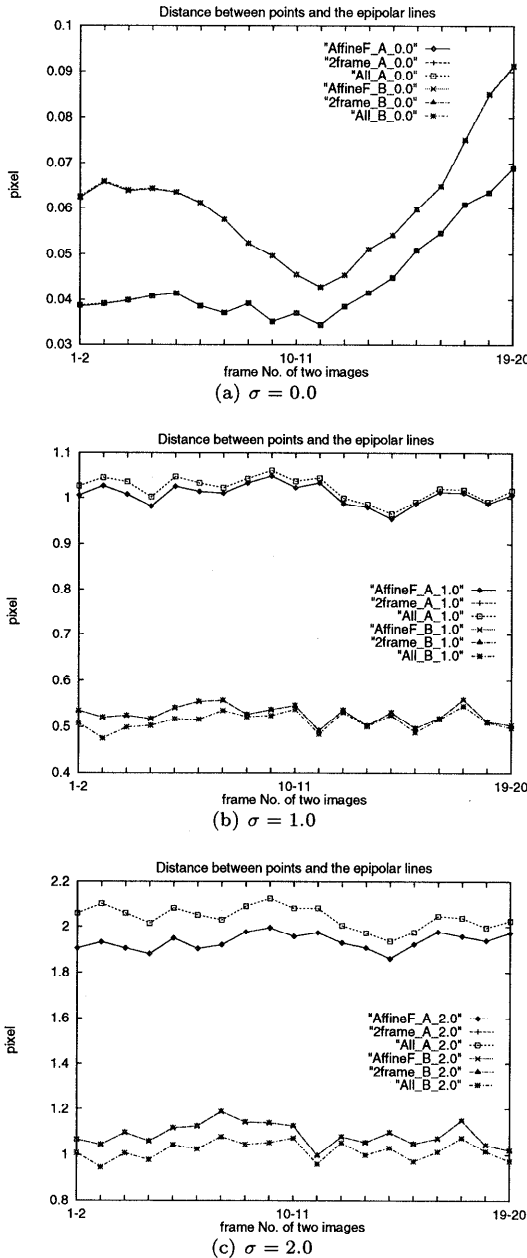


図5 特徴点とエピポラ線との距離

Fig. 5 Distance between points and the epipolar lines.

にそれぞれ、2画像のみ、全20画像を用いたことを表し、“AffineF”は従来法を示す。また、“_A_”、“_B_”はそれぞれデータA、データBを示す。

(a)では、すべての結果において実用上十分な精度が得られた。観測誤差を含まないため、データAの方がデータBよりも誤差は小さく、また、2画像、全画像どちらの \hat{M} を用いてもほとんど変化はなかった。

(b), (c)では、(a)とは逆にデータBの方が誤差が

小さくなった。その差は、(b)で約0.5pixel, (c)では約1pixelであった。この結果からは、エピポラ方程式の復元に用いた特徴点に多少観測誤差が含まれていても、新たに探索される画像特徴はエピポラ拘束を満たし、対応付けに成功することが期待できる。

また、データAでは、2画像から計算された \hat{M} を用いる方が、全画像を用いるよりも誤差が小さいが、データBでは、その逆に全画像を用いる方が2画像のみよりも誤差が小さく、その差は(b)で最大約0.05pixel, (c)では最大約0.12pixelであった。

さらに、従来法との比較であるが、図5中でほとんど重なっていることから分かるように、“2frame”と“AffineF”との結果にはほとんど差がなかった。これは、従来法も2フレームの情報のみを用いているからであると考えられる。

これらの結果からは、全画像を用いて \hat{M} (ならびに \hat{S}) を計算することで、各画像に誤差を分散させ3次的に矛盾の少ないエピポラ方程式が得られることが期待できる。これは、全画像を均等に扱え、数値計算的に安定であるという因子分解法の特徴を利用しているといえる。なお、このように全画像を同時に用いることで、各画像で得られる運動や形状の座標系を統一できていることも利点である。

以上の結果は paraperspective 射影モデルで復元した場合も同じであった。ここで、3次元運動および形状の復元について考察する。復元には、データAを用いた。また、観測誤差は $\sigma = 1$ とした。

運動復元誤差は、

$$e_i(f) = \cos^{-1} \| (i_f^{\text{estimated}})^T i_f^{\text{true}} \|$$

により評価した。 j_f , k_f についても同様に評価したが、図6に示すように大きな差が生じた。対象物体は、第12画像において最も画像中心に近く撮像されており、第1、第20画像に近いほど画像中心からずれて撮像されている。weak perspective 射影モデルでは、位置効果を表現できないため、第1、第20画像に近いほど運動復元誤差も大きくなっている。一方、paraperspective 射影モデルでは、どの画像においても1度以内に誤差が収まっている。形状に関しては歪み方の評価が難しいがこの実験では有意な差は得られなかった(両モデルとも約4%の誤差)。ただし、運動復元誤差を考慮にいと、ユークリッド空間における復元解を得るためには weak perspective 射影モデルでは不十分なことが確認された。

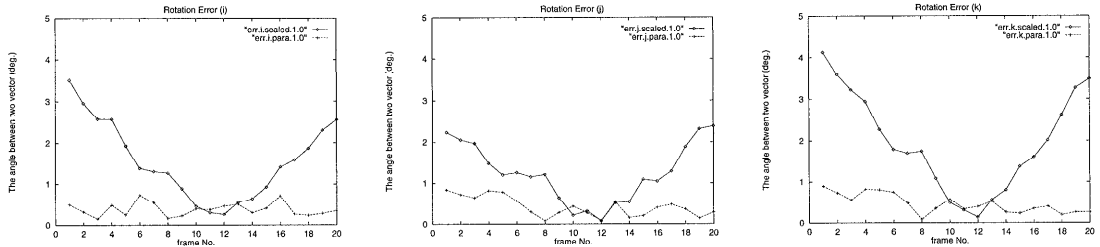


図6 回転運動の復元誤差 (左: i_f , 中央: j_f , 右: k_f)

Fig. 6 Rotation error.



図7 入力時系列画像 (左: frame 1, 中央: frame 50, 右: frame 99)

Fig. 7 Input image sequence.

5. 実時系列画像からの密な3次元形状復元

次に応用として, 提案手法を実時系列画像からの密な3次元形状復元に適用した. まず, 計算コストの高い2次元対応探索を2次元的なテクスチャを含む小方形特徴(特徴点)のみで行い, それらの対応付けから各画像の運動を因子分解法により復元した. 次にその結果を用いて得られるエピポーラ拘束による1次元探索を行い, 密な3次元形状を復元した.

実験に用いられた時系列画像(図7)は, 対象物体のぬいぐるみを手に持って動かし, カメラも適当にパン・チルトさせながら撮影されたものである. 特徴点の自動抽出・追跡は日立のIP5000を用いた.

LMedS 推定を用いた因子分解法⁵⁾により対象物体に属していると判断された23点の全99画像間の対応付けの結果を用いて計量を拘束し, ユークリッド空間における対象物体とカメラとの相対運動を復元した. また, 密な画像特徴の対応付けのために, 第64, 65, 66画像の3画像を第1, 2, 3眼とする仮想3眼ステレオマッチングを行った. 以下にその処理の概要を示す.

- [1] $R(\psi_f)$, Σ_f^{-1} , R_f^T を用いたエピポーラ線の水平化, スケール合わせにより第1, 2眼を変換し, ステレオマッチング用の画像対を生成(図8)
- [2] 3×3 ブロックの正規化相関を用いた1次元ブ

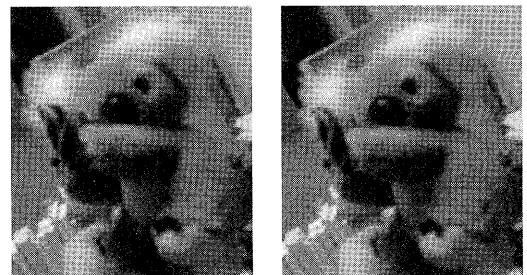


図8 ステレオマッチングのために変換された画像対
Fig. 8 Rectified images for stereo matching.

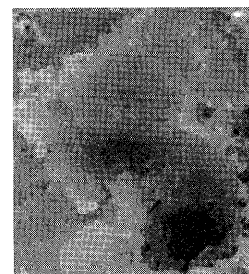


図9 視差画像
Fig. 9 Disparity map.

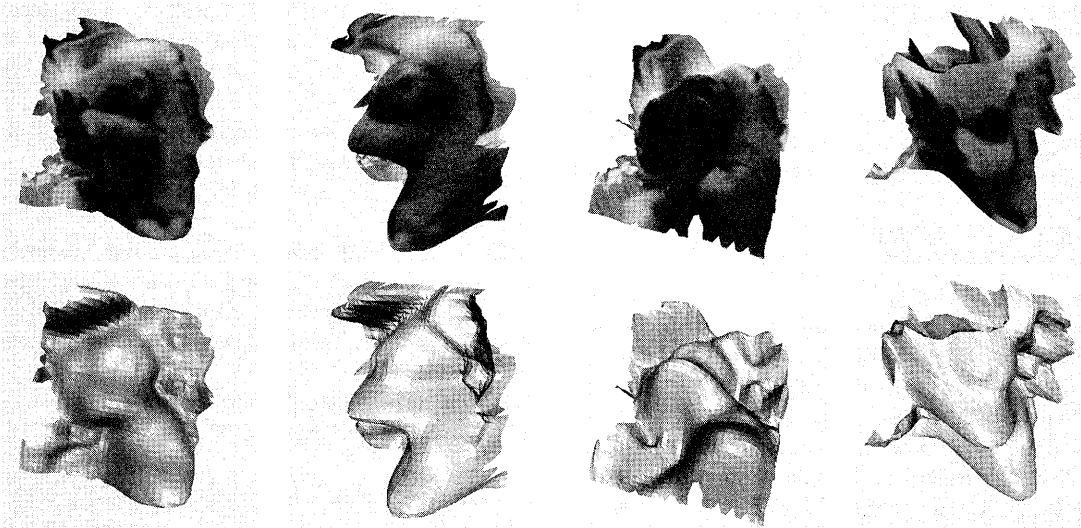


図 10 復元された対象物体の密な 3 次元形状
Fig.10 Reconstructed dense shape of the toy.

ロックマッチングにより視差画像を生成し (図 9), 全 99 画像で求められた計量を用いてユークリッド空間における各点の 3 次元座標を復元

- [3] [1]と同様の変換を第 2, 3 眼に施し, 復元された 3 次元座標を用いて第 2 眼と第 3 眼の各点の対応をとり, 3×3 ブロックの正規化相関値により [2]の結果を評価
- [4] エピポーラ線方向への輝度勾配がない点や [3]により正しく復元されていると評価されなかった点の 3 次元座標を周辺の点を用いて内挿
- [5] 各点を頂点とするメッシュからなるポリゴンモデルを生成

カメラの内部パラメータが未知であるので, 3 次元復元には weak perspective 射影モデルを用いた. 以上の処理により復元された対象物体のポリゴンモデルをいくつかの視点から表示した結果を図 10 に示す. この結果は, 少数の特徴点を追跡し対応付けるだけで, 因子分解法によりユークリッド空間における全画像の運動が復元され, その結果を利用して仮想ステレオ画像対 (エピポーラ拘束) が得られるため, 密な 3 次元形状が復元できることを示している. ただし, 射影モデルの近似誤差だけではなく, ステレオマッチングやポリゴンモデルの生成に単純な手法を用いた影響が結果に表れているため, 異なる画像特徴や多眼を用いた精度の高いステレオマッチング手法や, 適応的なポリゴンモデル生成手法を導入する必要がある. また, ここでは一方向からの密な 3 次元形状しか復元していない. 複数方向からの密な復元形状の統合についても検

討が必要である.

6. む す び

本稿では, 一般アフィン射影モデルによる因子分解法の観点から, まず, 2 画像の因子分解法について述べ, 次に, アフィンエピポーラ幾何の解釈について簡潔に示した. また, 具体的なパラメータを含んだ形でのエピポーラ方程式の定式化を行った.

シミュレーションでは, エピポーラ方程式の復元誤差を評価し, ある程度の観測誤差を含む特徴点の対応付けからエピポーラ拘束を導いても, 精度良く新たな画像特徴を探索することが可能であることを示した. 実画像による応用実験では, 実際にエピポーラ拘束を用いてステレオマッチングを行い, 密な 3 次元形状を復元した.

今後の課題としては, 実画像で paraperspective 射影モデルを用いるための, 焦点距離の獲得についてや, それを含まない形での計量拘束についての検討があげられる⁹⁾. また, GAP モデルでは, 透視変換による奥行き方向の歪みを近似できない反面, 簡単な線形計算による復元が可能であるため, 実時間性の重視される応用への適用についての検討も行いたい.

謝辞 本研究はリアルワールドコンピューティング (RWC) プログラムの一環として行われた. 本研究の機会を与えて下さった知能情報部長大津展之博士, 情報科学部長中島秀之博士, 情報ダイナミクスラボリーダー梅山伸二博士に深く感謝いたします. また, 様々なご議論をいただきました, 適応ビジョンラボ, 情報

ダイナミクスラボの皆様へ深く感謝いたします。

参考文献

- 1) Costeira, J. and Kanade, T.: A Multi-body Factorization Method for Independently Motion Objects, *IJCV*, Vol.29, No.3, pp.159-179 (1998).
- 2) 藤木 淳, 蔵田武志: 一般アフィン射影モデルの因子分解法の数理, 信学技法, PRMU97-142, pp.101-108 (1997).
- 3) Koenderink, J. J. and van Doorn, A. J.: Affine Structure from Motion, *J. Optic. Soc. Am.*, Vol.8, No.2, pp.377-395 (1991).
- 4) Kurata, T., Fujiki, J. and Sakaue, K.: Affine Epipolar Geometry via Factorization Method, *Proc. ICPR98*, Vol.1, pp.862-866 (1998).
- 5) 蔵田武志, 藤木 淳, 坂上勝彦: ロバスト推定を用いた一般アフィン射影モデルの因子分解法, 信学技法, PRMU98-194, pp.17-24 (1999).
- 6) Morris, D.D. and Kanade, T.: A Unified Factorization Algorithm for Points, Line Segments and Planes with Uncertainty Models, *Proc. ICCV98*, pp.696-702 (1998).
- 7) Ohta, Y., Maenobu, K. and Sakai, T.: Obtaining surface orientation from texels under perspective projection, *Proc. 7th IJCAI*, pp.746-751 (1981).
- 8) Poelman, C.J. and Kanade, T.: A Paraperspective Factorization Method for Shape and Motion Recovery, *PAMI*, Vol.13, No.3, pp.206-218 (1997).
- 9) Quan, L.: Self-Calibration of an Affine Camera from Multiple Views, Vol.19, No.1, pp.93-105 (1996).
- 10) Quan, L. and Kanade, T.: A Factorization Method for Affine Structure from Line Correspondences, *Proc. CVPR96*, pp.803-808 (1996).
- 11) Shapiro, L.S., Zisserman, A. and Brady, M.: 3D Motion Recovery via Affine Epipolar Geometry, *IJCV*, Vol.16, No.2, pp.147-182 (1995).
- 12) Shi, J. and Tomasi, C.: Good Features to Track, *Proc. CVPR94*, pp.593-600 (1994).
- 13) Sturm, P. and Triggs, B.: A Factorization Based Algorithm for Multi-Image Projective Structure and Motion, *Proc. ECCV96*, pp.709-720 (1996).
- 14) 杉本典子, 徐 剛: 弱中心射影画像からオイラー角を利用したモーション復元の線形アルゴリズム, 信学論, Vol.J81-D-II, No.4, pp.681-688 (1998).
- 15) Tomasi, C. and Kanade, T.: Shape and Motion from Image Streams under Orthography, *IJCV*, Vol.9, No.2, pp.137-154 (1992).
- 16) Ueshiba, T. and Tomita, F.: A Factorization

Method for Projective and Euclidean Reconstruction from Multiple Perspective Views via Iterative Depth Estimation, *Proc. 5th ECCV*, Vol.1, pp.296-310 (1998).

- 17) Ullman, S.: *The Interpretation of Visual Motion*, MIT Press, Cambridge, MA (1979).
- 18) 徐 剛, 辻 三朗: 3次元ビジョン, 共立出版 (1998).
- 19) Xu, G. and Zhang, Z.: Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach, *Computational Imaging and Vision*, Vol.6, Kluwer Academic Publishers (1996).
- 20) Zhang, Z.: <http://www-sop.inria.fr/robotvis/personnel/zhang/> (1997).

付録 各アフィン射影モデルの具体的なパラメータ

正射影モデル

$$A_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad R_f = I_2, \quad \Sigma_f = I_2$$

weak perspective (scaled orthographic) 射影モデル

$$A_f = \frac{l}{z_f} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad R_f = I_2$$

$$\Sigma_f = \frac{l}{z_f} I_2, \quad z_f = \frac{l}{|m_f|} = \frac{l}{|n_f|}$$

ここで, z_f は全特徴点の平均奥行きである。

paraperspective 射影モデル

$$A_f = \frac{1}{z_f} \begin{pmatrix} l & 0 & -x_f \\ 0 & l & -y_f \end{pmatrix}$$

$$R_f = \frac{1}{u_f} \begin{pmatrix} x_f & -y_f \\ y_f & x_f \end{pmatrix}$$

$$\Sigma_f = \frac{1}{z_f} \begin{pmatrix} \sqrt{l^2 + u_f^2} & 0 \\ 0 & l \end{pmatrix}$$

$$z_f = \frac{\sqrt{l^2 + x_f^2}}{|m_f|} = \frac{\sqrt{l^2 + y_f^2}}{|n_f|} = \sqrt{\frac{x_f \cdot y_f}{\langle m_f, n_f \rangle}}$$

ここで $u_f = \sqrt{x_f^2 + y_f^2}$ である。

なお, どの射影モデルにおいても, $\{z_f\}_{f=1}^F$ の値は比の形でしか求まらない。2画像について考えた場合, z_1 と z_2 の比は,

$$R_f^T \hat{M}_f Q \hat{M}_f^T R_f = \Sigma_f^2$$

$$c = \alpha_1 \hat{m}_1 + \beta_1 \hat{n}_1 = \alpha_2 \hat{m}_2 + \beta_2 \hat{n}_2$$

という関係から求めることができる。たとえば, weak perspective 射影モデルの場合は,

$$z_2 = \sqrt{\frac{\alpha_2^2 + \beta_2^2}{\alpha_1^2 + \beta_1^2}} z_1,$$

となる. $\alpha_f, \beta_f (f = 1, 2)$ は以下の式を解くことにより得られるため, Q は必要ない.

$$\alpha_1(\hat{r}_2^T \hat{m}_1) + \beta_1(\hat{r}_2^T \hat{n}_1) = 0$$

$$\alpha_2(\hat{r}_1^T \hat{m}_2) + \beta_2(\hat{r}_1^T \hat{n}_2) = 0$$

$$\beta_1(\hat{m}_{12}^T \hat{n}_1) = \beta_2(\hat{m}_{12}^T \hat{n}_2)$$

ここで, $\hat{r}_1 = \hat{m}_1 \times \hat{n}_1$, $\hat{r}_2 = \hat{m}_2 \times \hat{n}_2$, $\hat{m}_{12} = \hat{m}_1 \times \hat{m}_2$ である.

(平成 11 年 1 月 6 日受付)

(平成 11 年 6 月 3 日採録)



滝田 武志 (正会員)

1993 年筑波大学第三学群情報学類卒業. 1996 年同大学院工学研究科修士課程修了. 同年電子技術総合研究所入所. コンピュータビジョンの研究に従事.



藤木 淳

1993 年東京大工学部計数工学科卒業. 1995 年同大学院工学系研究科修士課程修了. 同年電子技術総合研究所入所. コンピュータビジョンの研究に従事. 電子通信情報学会員.



坂上 勝彦 (正会員)

1976 年東京大工学部電子工学科卒業. 1981 年同大学院工学系研究科博士課程修了. 同年電子技術総合研究所入所. 画像処理, コンピュータビジョン等の研究に従事. 工学博士. 1979 年度電子情報通信学会学術奨励賞受賞. 1985 年本会論文賞受賞. IEEE, 電子情報通信学会, 電気学会, 映像情報メディア学会各会員.