

焦点距離が未知のステレオカメラによる三次元復元

植芝 俊夫 富田 文明

電子技術総合研究所

つくば市梅園 1-1-4

E-mail: {ueshiba | tomita}@etl.go.jp

内部パラメータが未知のステレオカメラにおけるエピポラ拘束を特徴づける基礎行列には、7つの自由度がある。本稿では、カメラの内部パラメータのうち焦点距離のみが未知という条件のもとで、基礎行列を左右のカメラの焦点距離(1+1自由度)およびカメラ間の相対的位置(2自由度)・姿勢(3自由度)に分解する方法を提案し、その閉形式解を示す。これによってユークリッド空間におけるシーンの復元が可能になる。さらに、解が一意に決定できない退化運動の構造を明らかにし、この場合にも射影的復元よりも豊かな幾何学的情報が得られることを示す。特に同一平面上の4点のアフィン座標に関する不変式が存在することを指摘する。最後に、実画像を用いた実験によって理論の正当性を検証する。

Metric 3D Reconstruction from Stereo Cameras with Unknown Focal Lengths

Toshio Ueshiba Fumiaki Tomita

Electrotechnical Laboratory

Tsukuba, Ibaraki, 305-8568 Japan

E-mail: {ueshiba | tomita}@etl.go.jp

This fundamental matrix, which encodes the complete epipolar geometry between two perspective cameras with unknown intrinsic parameters, has seven degrees of freedom. This paper presents a new simple algorithm for decomposing the fundamental matrix into two focal lengths and camera motion parameters under an assumption that all the intrinsic parameters other than focal lengths are available. We derive a closed-form solution for focal lengths and characterize the degenerate motions under which no unique solution is possible. Even in these cases, it will be shown that we can elicit richer geometrical information than projective reconstruction. Especially, an invariant expression for affine coordinates of four co-planer points is defined under gaze motion.

1 はじめに

2台のカメラによって得られた画像間の点対応にはエビポーラ拘束が成り立ち、特にカメラ内部パラメータが未知の場合には、エビポーラ拘束は基礎行列 (fundamental matrix) と呼ばれる 3×3 行列によって特徴づけられる [8]。しかし、対応点の画像座標と基礎行列のみから復元した3次元情報には射影変換の不定性があり、ユークリッド空間におけるシーンの構造を一意に決定できないことが知られている [3, 5]。これは、内部パラメータが未知のステレオカメラにおけるエビポーラ拘束を特徴付ける基礎行列には7自由度しかなく、ステレオカメラの自由度数15(左右カメラの内部パラメータ: $5+5=10$ およびカメラ間の並進方向:2と回転:3)より少ないためである。しかし、左右のカメラの内部パラメータのうち2自由度のみが未知であるならば、それら2つのパラメータと並進・回転を基礎行列から取り出して、ユークリッド空間における3次元復元が可能になるものと考えられる。

Ponceらは、左右のカメラの内部パラメータが同一との仮定のもとで焦点距離とアスペクト比を基礎行列より取り出した [11]。本稿では、左右カメラの内部パラメータのうち焦点距離のみが未知という条件のもとで、基礎行列を焦点距離(1+1自由度)およびカメラ間の相対的位置(2自由度)・姿勢(3自由度)に分解する方法を提案し、それから導かれる性質について論じる。この問題はHartley [4]によって最初に扱われたが、具体的な解の形は与えられていなかった。その後、Newsamら [10] や Bougnoux [1] によって閉形式解が示されたが、その導出は複雑であった。

我々は、より簡便に閉形式解を導出する方法を示し、解が一意に決まらない退化運動がどのような場合に生じるかを明らかにする。さらに、これまで知られていなかった新しい性質として、退化運動のもとでも射影復元よりも豊かなシーンの幾何学的情報が得られることを指摘し、その場合に平面上の4点を用いてアフィン座標に関する不変式が定義できることを示す。

2 ステレオカメラと基礎行列

三次元空間中のある点 $\mathbf{x} \in \mathcal{R}^3$ を2台の中心投影カメラで観測した像を、斉次座標を用いてそれぞれ \mathbf{u}, \mathbf{u}' なる3次元ベクトルで表わす。すると、左右のカメラの投影式は未知のスカラー λ, λ' を用いて

$$\begin{cases} \lambda \mathbf{u} &= \mathbf{K} \mathbf{R}^T (\mathbf{x} - \mathbf{c}) \\ \lambda' \mathbf{u}' &= \mathbf{K}' \mathbf{R}'^T (\mathbf{x} - \mathbf{c}') \end{cases} \quad (1)$$

となる。ここで、3次元ベクトル \mathbf{c} と 3×3 直交行列 \mathbf{R} はそれぞれ左カメラの位置と姿勢を表す。同様に、 \mathbf{c}'

と \mathbf{R}' は右カメラの位置と姿勢を表わす。また、 \mathbf{K} と \mathbf{K}' は2台のカメラの内部パラメータを表わし、一般にこれらは 3×3 上半三角行列である。しかし、画像の横軸と縦軸が直交し、アスペクト比(画素の横と縦のサイズの比)および画像中心(焦点から画像面に降ろした垂線の足)が既知ならば、横軸と縦軸のスケールが等しく原点が画像中心になるような画像座標系をとることによって

$$\mathbf{K} = \text{diag}(k, k, 1), \quad \mathbf{K}' = \text{diag}(k', k', 1) \quad (2)$$

のような対角行列になる。ただし、 k, k' はそれぞれ左右のカメラの焦点距離(画素単位で表した焦点と画像面の距離)である。

(1) から \mathbf{x} を消去することによって $\lambda \mathbf{R} \mathbf{K}^{-1} \mathbf{u} - \lambda' \mathbf{R}' \mathbf{K}'^{-1} \mathbf{u}' = \mathbf{c}' - \mathbf{c}$ を得る。これは、3つのベクトル $\mathbf{R}' \mathbf{K}'^{-1} \mathbf{u}'$, $\mathbf{R} \mathbf{K}^{-1} \mathbf{u}$ および $\mathbf{t} \equiv \mathbf{c}' - \mathbf{c}$ が線形従属であることを示しているから、これらのスカラー3重積は0になる。すなわち、

$$\mathbf{u}'^T \underbrace{\mathbf{K}'^{-T} \mathbf{R}'^T}_{\mathbf{F}} [\mathbf{t}]_x \mathbf{R} \mathbf{K}^{-1} \mathbf{u} = 0 \quad (3)$$

ただし、 $\mathbf{K}'^{-T} \equiv (\mathbf{K}'^{-1})^T$ である。また、 $\mathbf{t} = [t_1, t_2, t_3]^T$ に対して反対称行列 $[\mathbf{t}]_x$ を

$$[\mathbf{t}]_x \equiv \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$$

と定義した。(3)をエビポーラ拘束と呼び、 \mathbf{F} を基礎行列 (fundamental matrix) という [8]。基礎行列は、左右の対応点の(画素単位で指定された)座標値の間に成立するエビポーラ拘束を特徴づける行列であり、2台のカメラの内部パラメータおよびそれらの相対的位置関係によって決まる。0でない \mathbf{t} に対しては $\text{rank} [\mathbf{t}]_x = 2$ なので \mathbf{F} の rank も2である。さらに(3)を満たす \mathbf{F} には定数倍の不定性がある。したがって、 \mathbf{F} は 3×3 行列であるがその自由度は7である。

さらに、

$$\mathbf{e} \simeq \mathbf{K} \mathbf{R}^T \mathbf{t}, \quad \mathbf{e}' \simeq \mathbf{K}' \mathbf{R}'^T \mathbf{t} \quad (4)$$

を満たす3次元ベクトル \mathbf{e}, \mathbf{e}' で表される画像上の点をそれぞれ左エビポール、右エビポールと呼ぶ*。これらは一方のカメラ焦点のもう一方のカメラへの投影像であり、 \mathbf{F} が与えられればそれぞれ $\mathbf{F} \mathbf{e} = \mathbf{0}$ および $\mathbf{F}' \mathbf{e}' = \mathbf{0}$ の解として(定数倍の不定性を除いて)容易に求めることができる。

*記号 \simeq は定数倍の違いを除いて等しいことを表す。

3 基礎行列からの焦点距離の抽出

本節では、基礎行列 F が与えられたときに、それから左右のカメラの焦点距離を取り出す方法を述べる。

3.1 基礎行列の変換

R と R' についてそれぞれ $R = [r_1, r_2, r_3]$, $R' = [r'_1, r'_2, r'_3]$ の如く列を陽に書けば、(2), (4) より左右のエピポール $e = [e_1, e_2, e_3]^T$, $e' = [e'_1, e'_2, e'_3]^T$ は

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \simeq \begin{bmatrix} k r_1^T t \\ k r_2^T t \\ r_3^T t \end{bmatrix}, \quad \begin{bmatrix} e'_1 \\ e'_2 \\ e'_3 \end{bmatrix} \simeq \begin{bmatrix} k' r'_1{}^T t \\ k' r'_2{}^T t \\ r'_3{}^T t \end{bmatrix}$$

となる。いま、 $e_1^2 + e_2^2 \neq 0$ かつ $e'_1{}^2 + e'_2{}^2 \neq 0$ と仮定すると、

$$\begin{aligned} T &\equiv \frac{1}{\sqrt{e_1^2 + e_2^2}} \begin{bmatrix} e_1 & -e_2 & 0 \\ e_2 & e_1 & 0 \\ 0 & 0 & \sqrt{e_1^2 + e_2^2} \end{bmatrix} \\ &= \frac{1}{\mu} \begin{bmatrix} r_1^T t & -r_2^T t & 0 \\ r_2^T t & r_1^T t & 0 \\ 0 & 0 & \mu \end{bmatrix} \end{aligned} \quad (5)$$

なる T (ただし $\mu \equiv \sqrt{(r_1^T t)^2 + (r_2^T t)^2}$) と同様に定義した T' (および $\mu' \equiv \sqrt{(r'_1{}^T t)^2 + (r'_2{}^T t)^2}$) を用いて、左右のカメラの画像座標をそれぞれ $\hat{u} = T^{-1}u$, $\hat{u}' = T'^{-1}u'$ と変換する。これは、左右のカメラの画像面においてエピポールが横軸に乗るよう、座標系をカメラ光軸のまわりに回転することを意味する (図1参照)。

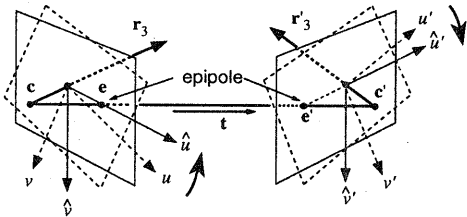


図1: 画像座標系の回転

この新しい座標系のもとでは、左右のエピポール $\hat{e} = [\hat{e}_1, \hat{e}_2, \hat{e}_3]^T$, $\hat{e}' = [\hat{e}'_1, \hat{e}'_2, \hat{e}'_3]^T$ はそれぞれ

$$\begin{bmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{bmatrix} \simeq \begin{bmatrix} k\mu \\ 0 \\ r_3^T t \end{bmatrix}, \quad \begin{bmatrix} \hat{e}'_1 \\ \hat{e}'_2 \\ \hat{e}'_3 \end{bmatrix} \simeq \begin{bmatrix} k'\mu' \\ 0 \\ r'_3{}^T t \end{bmatrix} \quad (6)$$

となる。また、基礎行列 F は $\hat{u}'^T \hat{F} \hat{u} = 0$ を満たす 3×3 行列 \hat{F} に変換される。(2), (3) および (5) を用いれば、

この \hat{F} は

$$\begin{aligned} \hat{F} &= T'^T F T \\ &\simeq \text{diag}(r_3^T t, 1, k'\mu') \\ &\begin{bmatrix} (r_3 \times r'_3)^T t & -(r'_3 \times t)^T (r_3 \times t) & -(r_3 \times r'_3)^T t \\ (r'_3 \times t)^T (r_3 \times t) & t^T t (r_3 \times r'_3)^T t & -(r'_3 \times t)^T (r_3 \times t) \\ -(r_3 \times r'_3)^T t & (r'_3 \times t)^T (r_3 \times t) & (r_3 \times r'_3)^T t \end{bmatrix} \\ &\text{diag}(r_3^T t, 1, k\mu) \end{aligned} \quad (7)$$

となる (付録A参照)。

3.2 焦点距離の抽出

\hat{F} の (i, j) 成分を f_{ij} と書くと、 $\mu^2 + (r_3^T t)^2 = t^T t$ であることに注意すれば (6) と (7) より

$$\frac{1}{k^2} + \frac{\hat{e}_3^2}{\hat{e}_1^2} = \frac{1}{k^2} + \frac{(r_3^T t)^2}{k^2 \mu^2} = \frac{t^T t}{k^2 \mu^2} = -\frac{f_{22} f_{32}}{f_{23} f_{33}}$$

を得る。同様にして $\frac{1}{k'^2} + \frac{\hat{e}'_3^2}{\hat{e}'_1^2} = -\frac{f'_{22} f'_{32}}{f'_{23} f'_{33}}$ を得る。よって、これらより k と k' を

$$\begin{cases} k = \left(-\frac{f_{22} f_{32}}{f_{23} f_{33}} - \frac{\hat{e}_3^2}{\hat{e}_1^2} \right)^{-\frac{1}{2}} \\ k' = \left(-\frac{f'_{22} f'_{32}}{f'_{23} f'_{33}} - \frac{\hat{e}'_3^2}{\hat{e}'_1^2} \right)^{-\frac{1}{2}} \end{cases} \quad (8)$$

のように求めることができる。

3.3 カメラ運動とシーンのユークリッド復元

k と k' が得られれば、(3) より基本行列 $E \equiv K^T F K$ を求め、特異値分解によってこれをカメラ運動の並進と回転に分離できることはよく知られている [7]。さらに、この情報を用いてユークリッド空間におけるシーンの構造を復元することが可能である。

4 退化運動

§3の焦点距離の計算では、 $e_1^2 + e_2^2 = 0$ または $e'_1{}^2 + e'_2{}^2 = 0$ となると、画像座標の変換行列 T または T' が定義できなくなる。また、 f_{23}, f_{32}, f_{33} のいずれかが0になると (8) の計算ができない ($e_1^2 + e_2^2 \neq 0$, $e'_1{}^2 + e'_2{}^2 \neq 0$ ならば、 \hat{e}_1, \hat{e}'_1 が0になることはない)。本節では、カメラがどのように運動するとこれらの破綻が生じてシーンのユークリッド復元が不可能になるのかを解析する。

並進の方向がいずれかの光軸方向と一致する場合

$e_1^2 + e_2^2 = 0$ または $e'_1{}^2 + e'_2{}^2 = 0$ となるのは、 $t \simeq r_3$ または $t \simeq r'_3$ の場合、すなわち並進ベクトルが左または右カメラの光軸上にある場合である。

$t \simeq r_3$ の場合を考えよう。(3) より

$$F \simeq \text{diag}(1, 1, k') \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix} [r_3]_x [r_1 \ r_2 \ r_3] \text{diag}(1, 1, k)$$

$$= \begin{bmatrix} \mathbf{r}_1^T \mathbf{r}_2 & -\mathbf{r}_1^T \mathbf{r}_1 & 0 \\ \mathbf{r}_2^T \mathbf{r}_2 & -\mathbf{r}_2^T \mathbf{r}_1 & 0 \\ k' \mathbf{r}_3^T \mathbf{r}_2 & -k' \mathbf{r}_3^T \mathbf{r}_1 & 0 \end{bmatrix}$$

すなわち、 \mathbf{F} は k を含まないのでその値を決定することはできない。一方、 k' については、 \mathbf{F} の (i, j) 成分を f_{ij} とすれば

$$f_{11}^2 + f_{21}^2 + \frac{1}{k^2} f_{31}^2 = f_{12}^2 + f_{22}^2 + \frac{1}{k'^2} f_{32}^2 \quad (9)$$

が成り立つので、 $f_{31}^2 \neq f_{32}^2$ ならば k' を決定できる。 $f_{31}^2 = f_{32}^2$ となるのは $(\mathbf{r}_3^T \mathbf{r}_2)^2 = (\mathbf{r}_3^T \mathbf{r}_1)^2$ である場合、すなわち右カメラの光軸が左カメラの横軸・縦軸と同一の角度を成す場合に相当する。

2つのカメラの光軸が直交する2平面上にある場合

$\hat{f}_{23} = \hat{f}_{32} = 0$ となるのは、 $(\mathbf{r}_3^T \mathbf{t})^T (\mathbf{r}_3^T \mathbf{t}) = 0$ のとき、すなわち左右のカメラの光軸が直交する2平面上にあり、それらの交線が並進ベクトルに一致する場合(図2参照)である。このとき $\hat{\mathbf{F}}$ は

$$\hat{\mathbf{F}} \simeq \begin{bmatrix} \mathbf{r}_3^T \mathbf{t} \mathbf{r}_3^T \mathbf{t} & 0 & -k \mu \mathbf{r}_3^T \mathbf{t} \\ 0 & \mathbf{t}^T \mathbf{t} & 0 \\ -k' \mu' \mathbf{r}_3^T \mathbf{t} & 0 & k' \mu' k \mu \end{bmatrix}$$

となり、 $\mu^2 + (\mathbf{r}_3^T \mathbf{t})^2 = \mu'^2 + (\mathbf{r}_3^T \mathbf{t})^2 = \mathbf{t}^T \mathbf{t}$ に注意すれば

$$\hat{f}_{11}^2 + \frac{1}{k^2} \hat{f}_{13}^2 + \frac{1}{k'^2} \hat{f}_{31}^2 + \frac{1}{k^2 k'^2} \hat{f}_{33}^2 = \hat{f}_{22}^2 \quad (10)$$

が成り立つことがわかる。

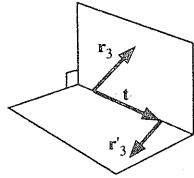


図2: 2つのカメラの光軸が直交する2平面上にある場合

2つのカメラの光軸が同一平面上にある場合

$\hat{f}_{33} = 0$ となるのは $(\mathbf{r}_3 \times \mathbf{r}_3')^T \mathbf{t} = 0$ のとき、すなわち左右のカメラの光軸および並進ベクトルが同一平面上に存在する場合である。このとき $\hat{\mathbf{F}}$ は

$$\hat{\mathbf{F}} \simeq \begin{bmatrix} 0 & -\mathbf{r}_3^T \mathbf{t} & 0 \\ \mathbf{r}_3^T \mathbf{t} & 0 & -k \mu \\ 0 & k' \mu' & 0 \end{bmatrix} \quad (11)$$

となり、 $\mu^2 + (\mathbf{r}_3^T \mathbf{t})^2 = \mu'^2 + (\mathbf{r}_3^T \mathbf{t})^2$ に注意すれば、 k と k' の間には

$$\hat{f}_{21}^2 + \frac{1}{k^2} \hat{f}_{23}^2 = \hat{f}_{12}^2 + \frac{1}{k'^2} \hat{f}_{32}^2 \quad (12)$$

が成り立つことがわかる。

もしも左右のカメラの焦点距離が等しいことがわかっているならば、 $k = k'$ として上式を解いてこれを求めることが可能である。ただし、この場合も $\hat{f}_{23}^2 = \hat{f}_{32}^2$ すなわち $\mu^2 = \mu'^2$ の時は解が求まらない。これは、 $(\mathbf{r}_3^T \mathbf{t})^2 = (\mathbf{r}_3'^T \mathbf{t})^2$ を意味し、左光軸 \mathbf{r}_3 と並進ベクトル \mathbf{t} のなす角が右光軸 \mathbf{r}_3' と \mathbf{t} のなす角に等しいことを示す。具体的には左右の光軸の交点と左右のカメラの焦点が二等辺三角形を成すか、左右の光軸が平行であるかのいずれかに相当する。並進成分に垂直な軸まわりの回転のみを許すステレオカメラについて、Brooksら[2]も同様の指摘を行っている。

5 注視運動による不変式

§4で扱った退化運動のうち、第一のタイプはエピポールすなわち視差を生じない点が画像の中心付近にあってその近辺では高精度の三次元復元が期待できず、本来のステレオの目的にはそぐわない運動である。また、第二のタイプは左右のカメラで共通の視野を確保しなければならない実際の状況下ではほとんど起こり得ない。これに対し、第三のタイプはある固定点を注視しながら移動するという極めて自然な運動であり、最も頻繁に生じる。本節ではこのような注視運動からどのような幾何学的情報を引き出すことができるかを論じる。

5.1 無限遠直線に対する拘束

左右のカメラの焦点をそれぞれ C, C' とし、それらの光軸の交点を V とする。 C, C', V によって定義される平面を Ψ とし、運動平面と呼ぶ。

まず、 C から直線 VC' におろした垂線の足を B' とする。同様に B を定義する。さらに、 C を通り直線 BC' と平行な直線と直線 VC' との交点を D' とする。同様に D を定義する(図3参照)。

次に、 $e \equiv -\hat{f}_{21}/\hat{f}_{23}$, $e' \equiv -\hat{f}_{12}/\hat{f}_{32}$, $\nu \equiv -\hat{f}_{32}/\hat{f}_{23}$, $a \equiv 1/k$, $a' \equiv 1/k'$ とおく。すると(12)は

$$e^2 + a^2 = \nu^2 (e'^2 + a'^2) \quad (13)$$

となる。また、このとき $\hat{\mathbf{e}} \simeq [1, 0, e]^T$, $\hat{\mathbf{e}}' \simeq [1, 0, e']^T$ であるから、並進ベクトル \mathbf{t} を左右のカメラ座標系から見たものはそれぞれ $\mathbf{K}^{-1} \hat{\mathbf{e}} \simeq [a, 0, e]^T$, $\mathbf{K}'^{-1} \hat{\mathbf{e}}' \simeq [a', 0, e']^T$ となる。これより B と C の間の(符号付き)距離を \overline{BC} などと書けば $\overline{BC}/\overline{BC'} = e/a$, $\overline{B'C'}/\overline{B'C} = e'/a'$ である。よって、 $\overline{BC} = e$ となるように全体のス

ケールを決めれば, (13)は $\overline{CC'}^2 = e^2 + a^2 = \nu^2(e^2 + a^2)$ であることを表す。したがって, $\overline{B'C'} = \nu e'$, $\overline{B'C} = \nu a'$ である。さらに, $\overline{VB} = b$, $\overline{VB'} = \nu b'$ とおけば, $\triangle VBC'$ と $\triangle VB'B$ が相似だから $\nu b'/b = (b+e)/\nu(b'+e)$, すなわち

$$\frac{b(b+e)}{b'(b'+e)} = \nu^2 \quad (14)$$

である。

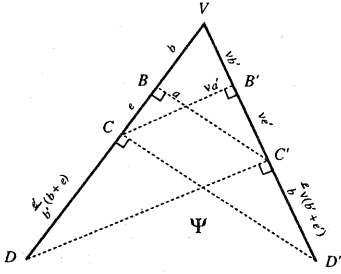


図3: 一次元中心投影カメラによるステレオ

一方, 平面 Ψ を 2 次元射影空間と見て, V, C, D, C', D' の位置をそれぞれ $\bar{v}, \bar{c}, \bar{d}, \bar{c}', \bar{d}'$ なる 3 次元ベクトルで表す。同様に, Ψ を 2 次元アフィン空間と見たときの C, D, C', D' の位置をそれぞれ $\bar{c}_A, \bar{d}_A, \bar{c}'_A, \bar{d}'_A$ なる 2 次元ベクトルで表す。このとき, 2 つの空間が 3×3 行列 \bar{H} で表される射影変換によって

$$\begin{bmatrix} \bar{c}_A \\ 1 \end{bmatrix} \simeq \bar{H} \bar{c}, \quad \begin{bmatrix} \bar{d}_A \\ 1 \end{bmatrix} \simeq \bar{H} \bar{d}, \dots$$

などと結ばれるものとしよう。

まず, V は直線 BC と直線 $B'C'$ の交点だから

$$\begin{aligned} \bar{v} &\simeq (\bar{c} \times \bar{d}) \times (\bar{c}' \times \bar{d}') \\ &= -(\bar{c}' \times \bar{d}')^T \bar{d} \bar{c} + (\bar{c}' \times \bar{d}')^T \bar{c} \bar{d} \end{aligned}$$

が成り立つ。これに \bar{H} を適用して,

$$\bar{H} \bar{v} \simeq -(\bar{c}' \times \bar{d}')^T \bar{d} \bar{h}^T \bar{c} \begin{bmatrix} \bar{c}_A \\ 1 \end{bmatrix} + (\bar{c}' \times \bar{d}')^T \bar{c} \bar{h}^T \bar{d} \begin{bmatrix} \bar{d}_A \\ 1 \end{bmatrix}$$

を得る。ただし \bar{h}^T は \bar{H} の第 3 行であり, 平面 Ψ 上の無限遠直線を表す。上式右辺に現れるベクトルの係数は C, D に関する V のアフィン座標であるから, V が線分 CD を $1:1+e'/b'$ に外分することを考慮すると,

$$1 + \frac{e'}{b'} = \frac{(\bar{c}' \times \bar{d}')^T \bar{d} \bar{h}^T \bar{c}}{(\bar{c}' \times \bar{d}')^T \bar{c} \bar{h}^T \bar{d}} \quad \text{すなわち}$$

$$b' = e' \frac{(\bar{c}' \times \bar{d}')^T \bar{c} \bar{h}^T \bar{d}}{\bar{h}^T (\bar{c} \bar{d}^T - \bar{d} \bar{c}^T) (\bar{c}' \times \bar{d}')} = -e' \frac{(\bar{c}' \times \bar{d}')^T \bar{c}^T \bar{h}^T \bar{d}}{\bar{h}^T \{(\bar{c} \times \bar{d}) \times (\bar{c}' \times \bar{d}')\}}$$

であることがわかる。同様に C', D' に関する V のアフィン座標を考えることによって

$$b = e \frac{(\bar{c} \times \bar{d})^T \bar{c}' \bar{h}^T \bar{d}'}{\bar{h}^T \{(\bar{c} \times \bar{d}) \times (\bar{c}' \times \bar{d}')\}}$$

が成り立つから, これら 2 つの式を (14) に代入して

$$\frac{(\bar{c} \times \bar{d})^T \bar{c}' (\bar{c} \times \bar{d})^T \bar{d}' \bar{h}^T \bar{c}' \bar{h}^T \bar{d}'}{(\bar{c}' \times \bar{d}')^T \bar{c} (\bar{c}' \times \bar{d}')^T \bar{d}' \bar{h}^T \bar{c} \bar{h}^T \bar{d}} = \frac{e'^2}{e^2} \nu^2 \quad (15)$$

を得る。この式は平面 Ψ を記述する射影座標系の取り方によらない。

射影復元を行えば射影空間におけるカメラの C, C' の位置を知ることができる。また, 光軸を含めたカメラの 3 軸も射影復元によって求まるので, 左右の軸同士の交点として定義される D, D' の位置も復元することができる。すなわち, $\bar{c}, \bar{c}', \bar{d}, \bar{d}'$ は基礎行列 F のみから求めることができる。無限遠平面または直線はシーンのアフィン構造を反映しており, 一般の射影復元ではその位置は全く任意であるが, (15) より, 注視運動のもとでは無限遠直線は基礎行列から計算されるあるコニック上に拘束されることがわかる。

5.2 平面上の 4 点より計算される不変式

運動平面 Ψ に垂直でない任意の平面 Π 上にある 4 点 X, P, Q, R を考え, これらのアフィン空間での位置をそれぞれ $\bar{x}_A, \bar{p}_A, \bar{q}_A, \bar{r}_A$ なる 2 次元ベクトルで表す。また, これら 4 点の平面 Ψ への直交投影を $\bar{X}, \bar{P}, \bar{Q}, \bar{R}$ とし (図 4 参照), そのアフィン空間での位置を $\bar{x}_A, \bar{p}_A, \bar{q}_A, \bar{r}_A$ なる 2 次元ベクトルで表す。 P, Q, R に関する X のアフィン座標を $[p_A, q_A, r_A]$ とすれば, $\bar{x}_A = p_A \bar{p}_A + q_A \bar{q}_A + r_A \bar{r}_A$ である。平面から平面への直交投影はアフィン変換であるから, 投影後も

$$\bar{x}_A = p_A \bar{p}_A + q_A \bar{q}_A + r_A \bar{r}_A \quad (16)$$

が成り立つ。

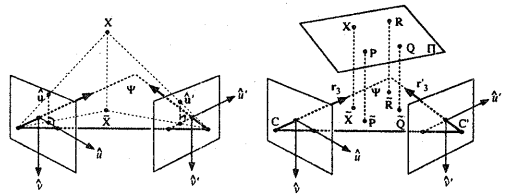


図4: 平面上の 4 点による不変式: 一般の点の運動平面への投影法 (左), 平面上の 4 点の運動平面への投影 (右)

一方, $\bar{X}, \bar{P}, \bar{Q}, \bar{R}$ の射影空間での位置を $\bar{x}, \bar{p}, \bar{q}, \bar{r}$ なる 3 次元ベクトルで表すと, これら 4 点の共面性よ

り, あるスカラ p, q, r が存在して

$$\bar{x} \simeq p\bar{p} + q\bar{q} + r\bar{r} = \begin{bmatrix} \bar{p} & \bar{q} & \bar{r} \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} \quad (17)$$

が成り立つ. すると,

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} \simeq \begin{bmatrix} (\bar{q} \times \bar{r})^T \\ (\bar{r} \times \bar{p})^T \\ (\bar{p} \times \bar{q})^T \end{bmatrix} \bar{x} = \begin{bmatrix} (\bar{q} \times \bar{r})^T \bar{x} \\ (\bar{r} \times \bar{p})^T \bar{x} \\ (\bar{p} \times \bar{q})^T \bar{x} \end{bmatrix}$$

一方, (17) に \bar{H} を適用してアフィン座標系に変換すれば

$$\bar{H}\bar{x} \simeq p\bar{h}^T \bar{p} \begin{bmatrix} \bar{p}_A \\ 1 \end{bmatrix} + q\bar{h}^T \bar{q} \begin{bmatrix} \bar{q}_A \\ 1 \end{bmatrix} + r\bar{h}^T \bar{r} \begin{bmatrix} \bar{r}_A \\ 1 \end{bmatrix}$$

であるから, (16) と比較して

$$\begin{bmatrix} p_A \\ q_A \\ r_A \end{bmatrix} \simeq \begin{bmatrix} p\bar{h}^T \bar{p} \\ q\bar{h}^T \bar{q} \\ r\bar{h}^T \bar{r} \end{bmatrix} = \begin{bmatrix} p\bar{p}^T \\ q\bar{q}^T \\ r\bar{r}^T \end{bmatrix} \bar{h}$$

が成り立つ. これより, 無限遠直線 \bar{h} を

$$\bar{h} \simeq \begin{bmatrix} \frac{\bar{q} \times \bar{r}}{p} & \frac{\bar{r} \times \bar{p}}{q} & \frac{\bar{p} \times \bar{q}}{r} \end{bmatrix} \begin{bmatrix} p_A \\ q_A \\ r_A \end{bmatrix} \\ \simeq \begin{bmatrix} \frac{\bar{q} \times \bar{r}}{(\bar{q} \times \bar{r})^T \bar{x}} & \frac{\bar{r} \times \bar{p}}{(\bar{r} \times \bar{p})^T \bar{x}} & \frac{\bar{p} \times \bar{q}}{(\bar{p} \times \bar{q})^T \bar{x}} \end{bmatrix} \begin{bmatrix} p_A \\ q_A \\ r_A \end{bmatrix} \quad (18)$$

のように P, Q, R に関する X のアフィン座標 p_A, q_A, r_A で表現することができる. (18) を (15) に代入すれば, アフィン座標に関する不変式 (焦点距離にかかわらず常に一定値をとる) を得る.

この不変式を計算するには, Π 上の点の Ψ への投影についてその射影座標が必要である. これには, まず左右の画像における Ψ 上の点の像をそれぞれ \hat{u} 軸と \hat{v} 軸に投影し, 次にこれらに対応点として射影復元を行えば良い (図 4 (左) 参照).

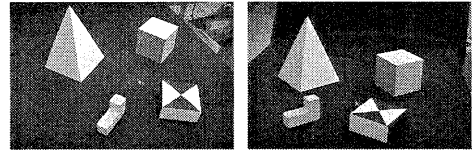
この不変式は, 平面上の 4 点から成るパターンに関してモデルとシーンのマッチングに利用できる. すなわち, あらかじめモデルパターンからアフィン座標を計算しておけば, これをシーン中にある平面上の 4 点の射影復元から得た不変式に代入して値を調べることに, マッチングの検証ができる[†]. 一般には, 射影復元しか得られていなければ, 射影変換に対する不変量である複比を求めるには同一平面上の 5 点 (または同一直線上の 4 点) が必要である [9] が, 我々の場合は無限遠直線に対する拘束を利用して 4 点で判定できる.

[†]ただし, 不変式の値が正しいことは対応が正しいための必要条件にしか過ぎないので, すべての誤対応を除去できるとは限らない.

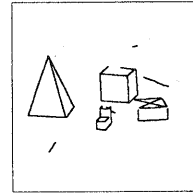
6 実験結果

6.1 焦点距離とシーンのユークリッド復元

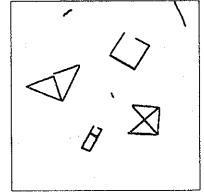
焦点距離が未知のカメラで撮像した二枚の画像からシーンのユークリッド構造を復元する実験を 2 種類行った. いずれの場合も, 角検出器を適用して得た特徴点のマッチングを画像間でとり, これらの対応点の組から基礎行列を求めた. 基礎行列の計算には, くりこみ法の後にランク 2 への補正を行う 2 段階最適アルゴリズム [14] を簡易化したものを用いた. 画像中心は与えられた画像の中央にあるものとし, 基礎行列を分解して焦点距離とカメラ運動パラメータを計算した後, 我々のグループで開発したセグメントベースステレオビジョンシステム [13, 12] を用いてシーン全体の三次元復元を行った.



(a) Stereo images with variable focal lengths



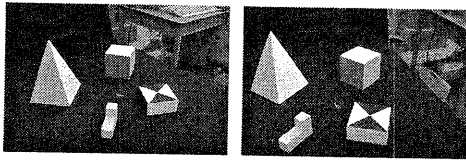
(b) Upper-side view



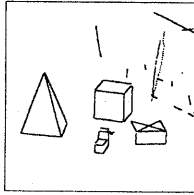
(c) Top view

図 5: 可変焦点距離・非退化運動のもとでのユークリッド復元

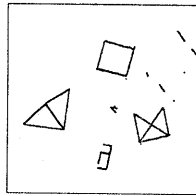
図 5 のシーンは, 上方から短焦点距離で第一画像を撮った後, 右後方に下がりズームして (退化運動を避けるため) 低い位置から第二画像を撮像したものである. 同図 (b),(c) は, 直角であるべき箇所がそのように復元されていることを示しており, シーンのユークリッド復元が成功していることがわかる. 推定された焦点距離はそれぞれ 756 pixel, 1343 pixel であった. 参照パターンを用いる事前のオフラインキャリブレーションによれば前者の焦点距離は 834 pixel であったから, 9.3% 程度の誤差がある. また, 退化運動への近さの指標であるアオリ角 (並進ベクトルと左光軸によって張られる平面と右光軸が成す角) は, 3.9 deg であった. Kahl らの数値実験 [6] によれば, 三次元復元に比べて焦点距離の正確な推定はより困難であり, 特にアオリ角が 5



(a) Stereo images with constant focal lengths



(b) Upper-side view



(c) Top view

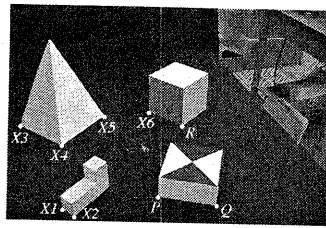
図6: 同一焦点距離・注視運動のもとでのユークリッド復元

deg 以下になると急激に精度が落ちることが報告されている。

図6は、焦点距離を一定に保ちながら注視運動を行って得たシーンに対する実験結果である。まず左上方から第一画像を撮った後、よりシーンに近付いて右上方から第二画像を撮像した。2つの光軸が同一平面上に乗るよう、両画像とも右奥立方体の左下角を画面の中央付近でとらえるようにカメラを動かした。推定された焦点距離は715 pixelであり、オフラインキャリブレーションの結果(834 pixel)とは14%の誤差を生じた。アオリ角の推定値は0.04 degであった。

6.2 不変式の検証

§5で述べた平面上の4点のアフィン座標に関する不変式が、実際に(15)で予想される値になるかを検証した。まず、図7の3点 P, Q, R をアフィン基底とし、点 $X_i (i = 1, \dots, 6)$ のこれらに関するアフィン座標を求めた。次に、注視運動を行いながら2枚の画像を撮り、 P, Q, R, X_i の運動平面 Ψ への投影 $\tilde{P}, \tilde{Q}, \tilde{R}, \tilde{X}_i$ の射影復元を行った。この結果を用いて各 \tilde{X}_i について(15)と(18)よりアフィン座標が満たすべき不変式を作り、それに X_1, \dots, X_6 のアフィン座標値を代入してどのような値をとるか調べた。その結果を表1に示す。表の対角線上のエントリが正しい対応に相当し、ほぼ基礎行列から予想される値をとっていることがわかる。厳密に一致しないのは、射影復元そのものに含まれる誤差の影響に加え、基礎行列が厳密に(11)の形を満たさないためと考えられる。



P : (398, 287)
 Q : (487, 240)
 R : (503, 494)
 X_1 : (228, 295)
 X_2 : (246, 273)
 X_3 : (130, 584)
 X_4 : (228, 479)
 X_5 : (330, 578)
 X_6 : (445, 573)
 (mm)

図7: 不変式の検証に用いた点とその位置

表1: 点 X_i とアフィン座標の組み合わせに対する不変式の値(予想値: 1.29868)

Point used with $\tilde{P}, \tilde{Q}, \tilde{R}$	Substituted affine coordinates					
	X_1	X_2	X_3	X_4	X_5	X_6
\tilde{X}_1	1.42	-0.211	2.05	1.83	1.66	1.49
\tilde{X}_2	2.09	1.27	2.38	2.09	1.82	1.56
\tilde{X}_3	0.898	0.879	1.38	0.737	1.18	1.20
\tilde{X}_4	0.967	0.935	3.32	1.30	1.06	1.17
\tilde{X}_5	0.723	0.686	1.10	0.908	1.29	0.996
\tilde{X}_6	0.561	0.529	0.828	0.721	0.858	1.41

7 おわりに

ステレオビジョンにおいて左右のカメラの内部パラメータのうち焦点距離のみが未知である場合に、基礎行列を左右の焦点距離およびカメラ間の並進・回転に分解する簡便な方法を提案し、その閉形式解を示した。この問題はこれまでもしばしば取り上げられ閉形式解の存在も示されていたが、本手法ではエビポールが横軸に乗るよう左右の画像座標系を光軸まわりに回転することにより、見通しよく解が導出できる。また、解が一意に決まらない退化運動の解析も非常に簡単になった。さらに退化運動の一つである注視運動のもとでは、運動平面上の無限遠直線があるコニック上に拘束され、この性質を用いて同一平面上の4点に対してそのアフィン座標に関する不変式が定義できることを示した。

実験の結果、2つのカメラの光軸が共面になる退化運動を十分に避ければ(アオリ角4-5 deg以上)、カメラの焦点距離が異なってもシーンのユークリッド復元が可能であることがわかった。また、退化運動の場合でも、注視点までの距離がカメラ間で十分に異なれば、焦点距離一定の条件のもとでユークリッド復元が可能である。

しかし、常に退化運動を避けようとするカメラ運動が非常に限定されてしまう。特に注視運動を避けてアオリ角を十分に確保しようとする、2台のカメラの視野の共通部分が小さくなるため対応点を局所的にしかとることができず、基礎行列を正確に推定するこ

とが困難になる。したがって、現実の状況下では退化運動のもとで不変となる幾何学的情報の利用が不可欠である。

提案した共面な4点のアフィン座標に関する不変式はこの目的のために構成したものである。実験の結果、これがほぼ予測どおりの値をとることが示された。しかし、そのノイズに対する振る舞いや精度についてはまだ明らかでないため、今後の解析と実験によって解明してゆきたい。

謝辞: 有益なコメントをいただくと共に多数の文献を御教示くださった群馬大学の金谷健一教授に心から感謝いたします。

A 式(7)の導出

$K^{-1} \simeq \text{diag}(1, 1, k)$ に注意すれば、まず

$$\begin{aligned} RK^{-1}T & \simeq \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1^T \mathbf{t} & -\mathbf{r}_2^T \mathbf{t} & 0 \\ \mathbf{r}_2^T \mathbf{t} & \mathbf{r}_1^T \mathbf{t} & 0 \\ 0 & 0 & k\mu \end{bmatrix} \\ & = \begin{bmatrix} (\mathbf{r}_1 \mathbf{r}_1^T + \mathbf{r}_2 \mathbf{r}_2^T) \mathbf{t} & -\mathbf{r}_1 \mathbf{r}_2^T \mathbf{t} + \mathbf{r}_2 \mathbf{r}_1^T \mathbf{t} & k\mu \mathbf{r}_3 \\ (\mathbf{I}_3 - \mathbf{r}_3 \mathbf{r}_3^T) \mathbf{t} & (\mathbf{r}_1 \times \mathbf{r}_2) \times \mathbf{t} & k\mu \mathbf{r}_3 \\ (\mathbf{r}_3 \times \mathbf{t}) \times \mathbf{r}_3 & \mathbf{r}_3 \times \mathbf{t} & k\mu \mathbf{r}_3 \end{bmatrix} \end{aligned}$$

であるから、

$$\begin{aligned} \hat{\mathbf{F}} & = \mathbf{T}^T \mathbf{K}'^{-T} \mathbf{R}'^T \begin{bmatrix} \mathbf{t} \end{bmatrix} \times RK^{-1}T \\ & \simeq \begin{bmatrix} \{(\mathbf{r}'_3 \times \mathbf{t}) \times \mathbf{r}'_3\}^T \\ (\mathbf{r}'_3 \times \mathbf{t})^T \\ k' \mu' \mathbf{r}'_3{}^T \end{bmatrix} \\ & \quad \begin{bmatrix} \mathbf{t} \times \{(\mathbf{r}_3 \times \mathbf{t}) \times \mathbf{r}_3\} & \mathbf{t} \times (\mathbf{r}_3 \times \mathbf{t}) & k\mu (\mathbf{t} \times \mathbf{r}_3) \end{bmatrix} \\ & = \begin{bmatrix} \{(\mathbf{r}'_3 \times \mathbf{t}) \times \mathbf{r}'_3\}^T \\ (\mathbf{r}'_3 \times \mathbf{t})^T \\ k' \mu' \mathbf{r}'_3{}^T \end{bmatrix} \\ & \quad \begin{bmatrix} (\mathbf{r}_3 \times \mathbf{t}) \mathbf{r}_3^T \mathbf{t} & \mathbf{t} \times (\mathbf{r}_3 \times \mathbf{t}) & -k\mu (\mathbf{r}_3 \times \mathbf{t}) \end{bmatrix} \\ & = \begin{bmatrix} \mathbf{r}'_3{}^T \{(\mathbf{r}_3 \times \mathbf{t}) \times (\mathbf{r}'_3 \times \mathbf{t})\} \mathbf{r}'_3{}^T \mathbf{t} & -\mathbf{r}'_3{}^T \mathbf{t} (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \\ (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \mathbf{r}'_3{}^T \mathbf{t} & \mathbf{t}^T \{(\mathbf{r}_3 \times \mathbf{t}) \times (\mathbf{r}'_3 \times \mathbf{t})\} \\ -k' \mu' (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \mathbf{r}'_3{}^T \mathbf{t} & k' \mu' (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \\ & -k\mu \mathbf{r}'_3{}^T \{(\mathbf{r}_3 \times \mathbf{t}) \times (\mathbf{r}'_3 \times \mathbf{t})\} \\ & -k\mu (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \\ & k' \mu' k\mu (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \end{bmatrix} \\ & = \begin{bmatrix} \mathbf{r}'_3{}^T \mathbf{t} (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \mathbf{r}'_3{}^T \mathbf{t} & -\mathbf{r}'_3{}^T \mathbf{t} (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \\ (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \mathbf{r}'_3{}^T \mathbf{t} & \mathbf{t}^T (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \\ -k' \mu' (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \mathbf{r}'_3{}^T \mathbf{t} & k' \mu' (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} -k\mu \mathbf{r}'_3{}^T \mathbf{t} (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \\ -k\mu (\mathbf{r}'_3 \times \mathbf{t})^T (\mathbf{r}_3 \times \mathbf{t}) \\ k' \mu' k\mu (\mathbf{r}_3 \times \mathbf{r}'_3)^T \mathbf{t} \end{bmatrix}$$

となって(7)を得る。ただし、以上の導出で三次元ベクトル $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ に関する公式 $\mathbf{a}^T (\mathbf{b} \times \mathbf{c}) = \mathbf{b}^T (\mathbf{c} \times \mathbf{a}) = \mathbf{c}^T (\mathbf{a} \times \mathbf{b}), (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -(\mathbf{b}^T \mathbf{c}) \mathbf{a} + (\mathbf{c}^T \mathbf{a}) \mathbf{b}, (\mathbf{a} \times \mathbf{b})^T (\mathbf{c} \times \mathbf{d}) = \mathbf{a}^T \mathbf{c} \mathbf{b}^T \mathbf{d} - \mathbf{a}^T \mathbf{d} \mathbf{b}^T \mathbf{c}$ を用いた。 ■

参考文献

- [1] S. Bougnoux. From Projective to Euclidean Space under any practical situation, a criticism of self-calibration. In *Proc. 6th International Conference on Computer Vision*, pp. 790-796, 1998.
- [2] M. J. Brooks, L. de Agapito, D. Q. Huynh, and L. Baumela. Towards robust metric reconstruction via a dynamic uncalibrated stereo head. *Image and Vision Computing*, Vol. 16, No. 14, pp. 989-1002, 1998.
- [3] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. 2nd European Conference on Computer Vision*, pp. 563-578, 1992.
- [4] R. I. Hartley. Estimation of Relative Camera Positions for Uncalibrated Cameras. In *Proc. 2nd European Conference on Computer Vision*, pp. 579-587, 1992.
- [5] R. Hartley, R. Gupta, and T. Chang. Stereo from Uncalibrated Cameras. In *Proc. Computer Vision and Pattern Recognition*, pp. 761-764, 1992.
- [6] F. Kahl and B. Triggs. Critical Motions in Euclidean Structure from Motion. In *Proc. Computer Vision and Pattern Recognition*, Vol. 2, pp. 366-372, 1999.
- [7] K. Kanatani. *Geometric Computation for Machine Vision*. Oxford University Press, 1993.
- [8] Q.-T. Luong and O. D. Faugeras. The Fundamental Matrix: Theory, Algorithms, and Stability Analysis. *International Journal of Computer Vision*, Vol. 17, No. 1, pp. 43-75, 1996.
- [9] J. L. Mundy and A. Zisserman, editors. *Geometric Invariance in Computer Vision*. MIT Press, 1992.
- [10] G. N. Newsam, D. Q. Huynh, M. J. Brooks, and H.-P. Pan. Recovering Unknown Focal Lengths in Self-Calibration: An Essentially Linear Algorithm and Degenerate Configurations. *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, No. B3-III, pp. 575-580, 1996.
- [11] J. Ponce, D. H. Marimont, and T. A. Cass. Analytical Methods for Uncalibrated Stereo and Motion Reconstruction. In *Proc. 4th European Conference on Computer Vision*, pp. 463-470, 1994.
- [12] T. Ueshiba, Y. Kawai, Y. Ishiyama, Y. Sumi, and F. Tomita. An Efficient Matching Algorithm for Segment-Based Stereo Vision Using Dynamic Programming Technique. In *Proc. IAPR Workshop on Machine Vision Applications*, pp. 61-64, 1998.
- [13] 河井良浩, 植芝俊夫, 石山豊, 角保志, 富田文明. セグメントベースステレオにおける連結性に基づく対応評価. *情処論*, Vol. 40, No. 8, pp. 3219-3229, 1999.
- [14] 三島等, 金谷健一. 基礎行列の最適計算とその信頼性評価. *情処研報*, No. CVIM118-10, pp. 67-74, September 1999.