

## 未校正因子分解法：カメラモデルを指定しないユークリッド復元

金谷 健一 浅原 清太郎 菅谷 保之 ハノ・アッカーマン

岡山大学大学院自然科学研究科

Tomasi・Kanade の因子分解法によって 3 次元復元を行うにはカメラモデル（平行，弱，疑似透視）を指定する必要がある．本論文ではそのような特定のカメラモデルを指定する必要のない方法を提案する．まず因子分解法の原理を，撮像がアフィンカメラであるという以外には何も仮定しない形で記述し，カメラモデルに依らない計量条件を導く．次に，撮像が透視投影を近似するための最小限の要請を置くと 2 個の不定関数を含むカメラモデルが得られることを示す．そして，その関数値を入力画像から自己校正によって定めれば自動的に適切なモデルが選ばれることを実験的に検証する．

### Uncalibrated Factorization: Euclidean Reconstruction without Specifying a Camera Model

Kenichi Kanatani, Seitaro Asahara, Yasuyuki Sugaya, Hanno Ackermann

Department of Computer Science, Okayama University, Okayama 700-8530 Japan

In order to reconstruct 3-D Euclidean shape by the Tomasi-Kanade factorization, one needs to specify a camera model such as orthographic, weak perspective, and paraperspective. We present a new method that does not require any such specific camera models. We first state the principle of 3-D reconstruction in the most general form without assuming anything about the camera except that it is affine and derive a camera-model-free metric constraint. We then prove that a minimal requirement for the affine imaging geometry to mimic perspective projection leads to a camera model that has two free functions. We experimentally confirm that if we optimally determine their values from input images by self-calibration, an appropriate camera model is automatically selected.

#### 1. まえがき

ビデオ画像上で特徴点を追跡して得られる軌跡からシーンの 3 次元形状を復元する手法として Tomasi・Kanade [18] の因子分解法 (factorization) がよく知られている．これはカメラの撮像をアフィン変換で近似してシーンの 3 次元形状を計算するものであり [9]，反復や探索なしに線形計算のみで実行できる点が最大の魅力である．復元結果に高精度を要求しなければ十分実用的であり，また，透視投影に基づく厳密な復元の反復計算 [4] の初期値として用いられる．

因子分解法ではカメラモデルを特定しなければ，復元形状は真の形状にアフィン変換を加えたもの（アフィン復元）である [4]．正しい形状（ユークリッド復元<sup>1</sup>）を得るためにはカメラモデルを特定しなければならない．従来はこれに平行投影 (orthographic)，弱透視 (weak perspective)，および疑似透視 (paraperspective) のカメラモデルが用いられていた [6, 14]．しかし，常にこの順に復元精度が高いわけではな

い [3]．実際の応用でどれを選ぶかは，すべてを試みて結果が最もよいものを採用するしかない．このことから次の疑問が生じる．

- これら以外のカメラモデルは存在しないのか．特に，疑似透視よりもよく透視投影を近似するカメラモデルは存在しないのか．
- これらは別々のモデルであり，一つが他を特別な場合として含むという包含関係がない．すべてを包括し，パラメータの取り方によって個々のモデルが得られるような一般モデルは存在しないのか．
- 最も適切なカメラモデルが自動的に選ばれ，ユーザが指定しなくてもよい方法は存在しないのか．

本論文は初めてこれに答えるものである．これまでこの問題が研究されなかったのは，研究者の関心が因子分解法の透視投影への拡張 [1, 2, 12, 15, 19, 21, 22, 23] に向かったことであると思われる．

以下ではまず，アフィンカメラであるという以外には何も仮定せずに，3 次元復元の原理を述べ，これがユークリッド復元となるための計量条件 (metric condition) の特定のカメラモデルに依らない形を導

<sup>†</sup>700-8530 岡山市津島中 3-1-1, (086)251-8173 {kanatani, asahara, sugaya, hano}@suri.it.okayama-u.ac.jp

<sup>1</sup>絶対的なスケールは不定であり，正しくは“相似復元”と呼ぶべきであるが，この呼び方が定着している．

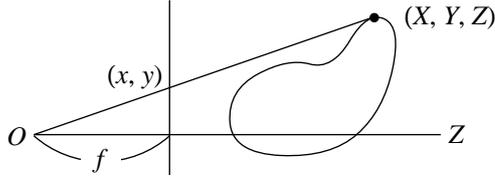


図 1: 透視投影 .

く . これは射影復元をユークリッド復元に高める絶対 2 級束面拘束条件 [4] に対応するものである .

次に , アフィンカメラが透視投影を近似するための最小限の要請を置くと , 2 個の不定関数を含むカメラモデルが一意的に定まることを示す . その関数を特定すると平行 , 弱透視 , 疑似透視投影が得られる .

本論文ではこれを特定せず , それらの値を入力画像から自己校正 (self-calibration) によって最適に定める . そして , これによって適切なモデルが自動的に選ばれることを実験的に検証する . しかも , 計算手順も計算量も基本的には従来と同じである .

## 2. アフィンカメラモデル

視点を原点  $O$  , カメラの光軸を  $Z$  軸とする  $XYZ$  座標系を考える . 透視投影 (perspective projection) とは空間の点  $(X, Y, Z)$  から次の画像座標  $(x, y)$  への投影である (図 1) .

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z} \quad (1)$$

定数  $f$  を焦点距離 (focal length) と呼ぶ . 式 (1) は通常のカメラのよい記述になっている . アフィンカメラ (affine camera) とはこれを次の 1 次関数で近似するものである .

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13} \\ \Pi_{21} & \Pi_{22} & \Pi_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} \quad (2)$$

$2 \times 3$  行列  $\Pi = (\Pi_{ij})$  と 2 次元ベクトル  $\pi = (\pi_i)$  をそれぞれ投影行列 , 投影ベクトルと呼ぶ . それらの要素は内部パラメータ (カメラに固有な定数) および外部パラメータ (シーンのカメラに相対的な位置や向きの) の関数である . したがって , カメラに相対的にシーンが運動していれば , 内部パラメータが一定でも  $\Pi, \pi$  は時間と共に変化する . ただし , シーンの個々の点の位置には依存しない .

代表的なアフィンカメラモデルは , シーン中に基準点  $(t_x, t_y, t_z)$  をとり , 透視投影 (1) を次のように近似するものである .

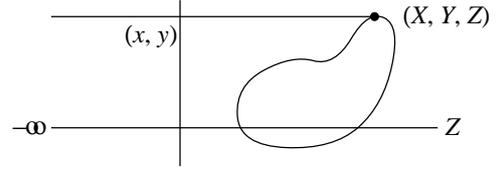


図 2: 平行投影 .

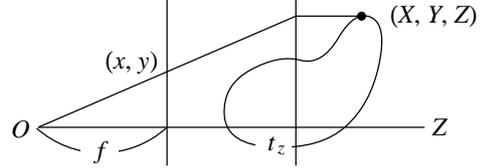


図 3: 弱透視投影 .

平行投影 (図 2)

$$\Pi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \pi = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3)$$

弱透視投影 (図 3)

$$\Pi = \begin{pmatrix} f/t_z & 0 & 0 \\ 0 & f/t_z & 0 \end{pmatrix}, \quad \pi = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (4)$$

疑似透視投影 (図 4)

$$\Pi = \begin{pmatrix} f/t_z & 0 & -ft_x/t_z^2 \\ 0 & f/t_z & -ft_y/t_z^2 \end{pmatrix}, \quad \pi = \begin{pmatrix} ft_x/t_z \\ ft_y/t_z \end{pmatrix} \quad (5)$$

図 2~4 から , それぞれが別々の幾何学的構造であり , 疑似透視は弱透視を , 弱透視は平行投影をその特殊な場合として含むというような階層関係が存在しないことがわかる .

## 3. アフィン空間拘束条件

本論文では静止した  $XYZ$  カメラ座標系に相対的にシーンが移動するとし<sup>2</sup> , シーンに固定した座標系を考える . このシーン座標系の時刻  $\kappa$  の原点の位置を  $t_\kappa$  , 座標基底ベクトルの作る正規直交基底を  $\{i_\kappa, j_\kappa, k_\kappa\}$  とする . ベクトル  $t_\kappa$  はシーン座標系のカメラ座標系からの「並進」を表す .

シーン中に  $N$  個の特徴点  $p_1, p_2, \dots, p_N$  をとると , 点  $p_\alpha$  の時刻  $\kappa$  における位置  $r_\alpha$  は次のように表せる .

$$r_{\kappa\alpha} = t_\kappa + a_\alpha i_\kappa + b_\alpha j_\kappa + c_\alpha k_\kappa \quad (6)$$

<sup>2</sup>伝統的な因子分解法 [6, 13, 14, 18] では静止したシーンに対してカメラが移動するとみなすが , 数学的にはどちらでも等価である . 以下の議論では本論文の定式化のほうが便利である .

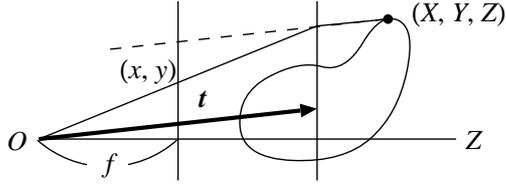


図 4: 疑似透視投影 .

シーン座標系に関する特徴点  $p_\alpha$  のシーン座標  $(a_\alpha, b_\alpha, c_\alpha)$  は時刻  $\kappa$  には依存しない .

カメラモデル (2) より特徴点  $p_\alpha$  の時刻  $\kappa$  における画像座標  $(x_{\kappa\alpha}, y_{\kappa\alpha})$  は次のように書ける .

$$\begin{pmatrix} x_{\kappa\alpha} \\ y_{\kappa\alpha} \end{pmatrix} = \tilde{t}_\kappa + a_\alpha \tilde{i}_\kappa + b_\alpha \tilde{j}_\kappa + c_\alpha \tilde{k}_\kappa \quad (7)$$

ただし  $\tilde{t}_\kappa, \tilde{i}_\kappa, \tilde{j}_\kappa, \tilde{k}_\kappa$  はそれぞれ次の 2 次元ベクトルである .

$$\begin{aligned} \tilde{t}_\kappa &= \Pi_\kappa t_\kappa + \pi_\kappa, & \tilde{i}_\kappa &= \Pi_\kappa i_\kappa \\ \tilde{j}_\kappa &= \Pi_\kappa j_\kappa, & \tilde{k}_\kappa &= \Pi_\kappa k_\kappa \end{aligned} \quad (8)$$

ここに  $\Pi_\kappa, \pi_\kappa$  はそれぞれ時刻  $\kappa$  の投影行列, 投影ベクトルである .

ビデオ画像上で特徴点  $\{p_\alpha\}$  を  $M$  フレームに渡って追跡すると, 各点  $p_\alpha$  の運動の履歴は  $2M$  次元空間の 1 点として表せる .  $2M$  個の座標をまとめたベクトル

$$p_\alpha = \begin{pmatrix} x_{1\alpha} & y_{1\alpha} & x_{2\alpha} & y_{2\alpha} & \dots & x_{M\alpha} & y_{M\alpha} \end{pmatrix}^\top \quad (9)$$

を点  $p_\alpha$  の軌跡ベクトルと呼ぶ . これは式 (7) より次のように表せる .

$$p_\alpha = m_0 + a_\alpha m_1 + b_\alpha m_2 + c_\alpha m_3 \quad (10)$$

ただし,  $m_0, m_1, m_2, m_3$  はそれぞれ次の  $2M$  次元ベクトルである .

$$\begin{pmatrix} \tilde{t}_1 \\ \tilde{t}_2 \\ \vdots \\ \tilde{t}_M \end{pmatrix}, \begin{pmatrix} \tilde{i}_1 \\ \tilde{i}_2 \\ \vdots \\ \tilde{i}_M \end{pmatrix}, \begin{pmatrix} \tilde{j}_1 \\ \tilde{j}_2 \\ \vdots \\ \tilde{j}_M \end{pmatrix}, \begin{pmatrix} \tilde{k}_1 \\ \tilde{k}_2 \\ \vdots \\ \tilde{k}_M \end{pmatrix} \quad (11)$$

これから, すべての特徴点の運動履歴を表すベクトルは  $2M$  次元空間  $\mathcal{R}^{2M}$  の  $m_0$  を通り,  $m_1, m_2, m_3$  の張る 3 次元アフィン空間に含まれることがわかる . これをアフィン空間拘束条件と呼ぶ .

この事実は因子分解法による 3 次元復元だけでなく, 複数運動の分離などの多くのビデオ画像処理の基本原則であり [5, 10, 11, 16, 17, 20], 特定のカメラモデルによらない .

#### 4. アフィン基底の変換

シーン座標系の原点は任意に取れるから,  $N$  個の特徴点  $\{p_\alpha\}$  の重心に取る . すると  $\sum_{\alpha=1}^N a_\alpha = \sum_{\alpha=1}^N b_\alpha = \sum_{\alpha=1}^N c_\alpha = 0$  であるから, 式 (10) より

$$\frac{1}{N} \sum_{\alpha=1}^N p_\alpha = m_0 \quad (12)$$

となり,  $m_0$  は  $2M$  次元空間  $\mathcal{R}^{2M}$  の軌跡ベクトル  $\{p_\alpha\}$  の重心に一致する .

式 (11) より,  $m_0$  の第  $(2(\kappa-1)+1)$ , 第  $(2(\kappa-1)+2)$  成分を取り出したものが  $\tilde{t}_\kappa$  であるから, すべての  $\tilde{t}_\kappa = (\tilde{t}_{x\kappa}, \tilde{t}_{y\kappa})^\top$ ,  $\kappa = 1, \dots, M$  が定まる .

重心を差し引いたベクトルを

$$p'_\alpha = p_\alpha - m_0 \quad (13)$$

とすると, 式 (10) は次のように表せる<sup>3</sup> .

$$p'_\alpha = a_\alpha m_1 + b_\alpha m_2 + c_\alpha m_3 \quad (14)$$

すなわち  $p'_\alpha$  は  $m_1, m_2, m_3$  の張る 3 次元部分空間  $\mathcal{L}$  に含まれる . したがってモーメント行列

$$C = \sum_{\alpha=1}^N p'_\alpha p'^\top_\alpha \quad (15)$$

のランクは 3, すなわち  $C$  の非零の固有値は 3 個あり, その固有ベクトルの正規直交系  $\{u_1, u_2, u_3\}$  が部分空間  $\mathcal{L}$  の基底となる .  $m_1, m_2, m_3$  はそれらの 1 次結合として次のように表される .

$$m_i = \sum_{j=1}^3 A_{ji} u_j \quad (16)$$

$A_{ij}$  を  $(ij)$  要素とする  $3 \times 3$  行列を  $A$  とし,  $m_1, m_2, m_3$  を列ベクトルとする  $2M \times 3$  行列および  $u_1, u_2, u_3$  を列ベクトルとする  $2M \times 3$  行列を

$$M = \begin{pmatrix} m_1 & m_2 & m_3 \end{pmatrix}, \quad U = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \quad (17)$$

とすると式 (16) は次のように書ける<sup>4</sup> .

$$M = UA \quad (18)$$

<sup>3</sup>伝統的な因子分解法 [6, 13, 14, 18] では  $p'_1, \dots, p'_N$  を列とする観測行列 (または計測行列) を  $W = (p'_1 \dots p'_N)$ , 形状行列を  $S = \begin{pmatrix} a_1 & \dots & a_N \\ b_1 & \dots & b_N \\ c_1 & \dots & c_N \end{pmatrix}$  とし, 運動行列  $M$  を式 (17) のように定め, 式 (14) を  $W = MS$  と表している . しかし, 以下の解析には本論文の定式化のほうが便利である .

<sup>4</sup>伝統的な因子分解法 [6, 13, 14, 18] では観測行列  $W$  を  $W = U\Lambda V^\top$  と特異値分解し,  $M = UA$ ,  $S = A^{-1}\Lambda V^\top$  としているが, 以下の解析には本論文の定式化のほうが便利である .

## 5. 一般アフィンカメラの計量条件

行列  $A = (A_{ij})$  を定める原理は、式 (11) のベクトル  $m_1, m_2, m_3$  が正規直交系  $\{i_\kappa, j_\kappa, k_\kappa\}$  のアフィンカメラによる投影像として式 (8) ように定まるという事実である。式 (8) から次の関係を得る。

$$\begin{pmatrix} \tilde{i}_\kappa & \tilde{j}_\kappa & \tilde{k}_\kappa \end{pmatrix} = \Pi_\kappa \begin{pmatrix} i_\kappa & j_\kappa & k_\kappa \end{pmatrix} = \Pi_\kappa R_\kappa \quad (19)$$

ただし、正規直交基底  $\{i_\kappa, j_\kappa, k_\kappa\}$  を列ベクトルとする行列  $R_\kappa$  を次のように定義した。

$$R_\kappa = \begin{pmatrix} i_\kappa & j_\kappa & k_\kappa \end{pmatrix} \quad (20)$$

これはシーン座標系のカメラ座標系からの「回転」を表す。

式 (17) の行列  $M$  の転置  $M^\top$  の第  $(2(\kappa - 1) + a)$  列を  $m_{\kappa(a)}^\dagger$  ( $\kappa = 1, \dots, M, a = 1, 2$ ) とすると、式 (19) の転置は次のように書ける。

$$R_\kappa^\top \Pi_\kappa^\top = \begin{pmatrix} m_{\kappa(1)}^\dagger & m_{\kappa(2)}^\dagger \end{pmatrix} \quad (21)$$

式 (18) より  $M^\top = A^\top U^\top$  であるから、 $M^\top$  に対応して式 (17) の行列  $U$  の転置  $U^\top$  の第  $(2(\kappa - 1) + a)$  列を  $u_{\kappa(a)}^\dagger$  とすると次の関係を得る。

$$m_{\kappa(a)}^\dagger = A^\top u_{\kappa(a)}^\dagger \quad (22)$$

これを用いると、式 (21) は次のように書き直せる。

$$R_\kappa^\top \Pi_\kappa^\top = A^\top \begin{pmatrix} u_{\kappa(1)}^\dagger & u_{\kappa(2)}^\dagger \end{pmatrix} \quad (23)$$

$u_{\kappa(1)}^\dagger, u_{\kappa(2)}^\dagger$  を列とする  $3 \times 2$  行列を  $U_\kappa^\dagger$  と置く。

$$U_\kappa^\dagger = \begin{pmatrix} u_{\kappa(1)}^\dagger & u_{\kappa(2)}^\dagger \end{pmatrix} \quad (24)$$

式 (23) より次の関係を得る。

$$U_\kappa^{\dagger\top} A A^\top U_\kappa^\dagger = \Pi_\kappa R_\kappa R_\kappa^\top \Pi_\kappa^\top \quad (25)$$

$R_\kappa$  は直交行列であるから、次式を得る。

$$U_\kappa^{\dagger\top} T U_\kappa^\dagger = \Pi_\kappa \Pi_\kappa^\top \quad (26)$$

ただし、計量行列  $T$  を次のように定義した。

$$T = A A^\top \quad (27)$$

式 (26) を計量条件と呼ぶ。両辺の要素を取り出すと次のようになる。

$$\left( u_{\kappa(1)}^\dagger, T u_{\kappa(1)}^\dagger \right) = \sum_{i=1}^3 \Pi_{1i\kappa}^2,$$

$$\begin{aligned} \left( u_{\kappa(2)}^\dagger, T u_{\kappa(2)}^\dagger \right) &= \sum_{i=1}^3 \Pi_{2i\kappa}^2 \\ \left( u_{\kappa(1)}^\dagger, T u_{\kappa(2)}^\dagger \right) &= \sum_{i=1}^3 \Pi_{1i\kappa} \Pi_{2i\kappa} \end{aligned} \quad (28)$$

これに式 (3), (4), (5) の投影行列の値をそれぞれ代入することにより、よく知られた平行、弱透視、疑似透視の計量条件 [6, 13, 9, 14, 18] が得られる。上式はすべてのアフィンカメラモデルを包括する計量条件の最も一般的な形である。

式 (16) の代わりに  $m_i = u_i, i = 1, 2, 3$  としても 3次元形状が復元できるが、得られる形状は真の形状にアフィン変換が加わったものである（アフィン復元）。式 (16) は正しい復元（ユークリッド復元）となるように基底  $\{u_1, u_2, u_3\}$  を変換するものであり、その変換行列  $A$  を定めるものが計量条件 (26) である。この意味で、これは透視カメラによる射影復元をユークリッド復元にする基底の変換行列を定める絶対 2 級束面（または絶対双対 2 次曲面）拘束条件 (absolute dual quadric constraint)[4] に対応する。

## 6. アフィンカメラによる透視投影の近似

次に、式 (2) のアフィンカメラが式 (1) の透視投影を近似するための最小限の条件を課して、カメラモデルのより具体的な表現を求める。まず次の条件を課す。

条件 1 . シーン座標系の原点  $(t_x, t_y, t_z)$  を通り、 $XY$  面に平行な平面は透視カメラと同じ投影像を生じる。

これは、シーンがカメラの画像面に平行な平面であるときは、それが透視カメラと同じように写るといことであり、自然な要請である。平面  $Z = t_z$  上の点を  $(X, Y, t_z)$  とすると、式 (1) より条件 1 は次のように書ける。

$$\begin{pmatrix} fX/t_z \\ fY/t_z \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + t_z \begin{pmatrix} \Pi_{13} \\ \Pi_{23} \end{pmatrix} + \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} \quad (29)$$

これが任意の  $(X, Y)$  で恒等的に成立しなければならぬから、次の関係を得る。

$$\begin{aligned} \Pi_{11} = \Pi_{22} = \frac{f}{t_z}, \quad \Pi_{12} = \Pi_{21} = 0, \\ t_z \Pi_{13} + \pi_1 = 0, \quad t_z \Pi_{23} + \pi_2 = 0 \end{aligned} \quad (30)$$

これからカメラモデル (2) が次のように書き直せる。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_z} \begin{pmatrix} X \\ Y \end{pmatrix} - (t_z - Z) \begin{pmatrix} \Pi_{13} \\ \Pi_{23} \end{pmatrix} \quad (31)$$

このとき，式中の  $f$  は定数である必要はなく，カメラの外部パラメータの関数であってもよい．式 (31) をさらに具体化するために次の条件を課す．

条件 2 . カメラの撮像は  $Z$  軸の周りに軸対称である．

条件 3 . 投影行列  $\Pi$  はカメラの内部パラメータおよびシーン座標系の原点  $(t_x, t_y, t_z)$  の関数である．

条件 2 は，シーンをカメラの光軸の周りにある角度  $\theta$  だけ回転させると，その投影像も画像原点の周りに同じ角度  $\theta$  だけ回転するという自然な要請である．条件 3 は撮像がシーン座標系の向きによらないということであり，これも自然な要請である．なぜなら，シーン座標系は任意の向きにとれるので，観測画像がそのような任意に選べる量に影響されるのはおかしいからである．

光軸の周りの回転角を  $\theta$  とし，回転行列を

$$\mathcal{R}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (32)$$

と置くと，条件 2 は次式となる．

$$\mathcal{R}(\theta) \begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_z} \mathcal{R}(\theta) \begin{pmatrix} X \\ Y \end{pmatrix} - (t_z - Z) \begin{pmatrix} \Pi'_{13} \\ \Pi'_{23} \end{pmatrix} \quad (33)$$

ここに  $\Pi'_{13}$ ,  $\Pi'_{23}$  は条件 3 より，関数  $\Pi_{13}$ ,  $\Pi_{23}$  の引数中の  $t_x, t_y$  をそれぞれ  $t_x \cos \theta - t_y \sin \theta$ ,  $t_x \sin \theta + t_y \cos \theta$  に置き換えたものである．

式 (31) の両辺に  $\mathcal{R}(\theta)$  を掛けると次のようになる．

$$\mathcal{R}(\theta) \begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_z} \mathcal{R}(\theta) \begin{pmatrix} X \\ Y \end{pmatrix} - (t_z - Z) \mathcal{R}(\theta) \begin{pmatrix} \Pi_{13} \\ \Pi_{23} \end{pmatrix} \quad (34)$$

式 (33), (34) を比較すると，任意の  $\theta$  に対して

$$\begin{pmatrix} \Pi'_{13} \\ \Pi'_{23} \end{pmatrix} = \mathcal{R}(\theta) \begin{pmatrix} \Pi_{13} \\ \Pi_{23} \end{pmatrix} \quad (35)$$

が恒等的に成立しなければならない．これから不変量の理論 [7] でよく知られているように，

$$\begin{pmatrix} \Pi_{13} \\ \Pi_{23} \end{pmatrix} = c \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (36)$$

でなければならない．ここに  $c$  は内部パラメータと  $t_x^2 + t_y^2, t_z$  の任意の関数である．以上より次式を得る．

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{t_z} \begin{pmatrix} X \\ Y \end{pmatrix} - c(t_z - Z) \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (37)$$

したがって，任意に選べる関数は  $f/t_z$  と  $c$  の 2 個のみである．これらを組み合わせると

$$\zeta = \frac{t_z}{f}, \quad \beta = -\frac{ct_z}{f} \quad (38)$$

と置けば，カメラモデル (37) は次のように書ける．

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\zeta} \left( \begin{pmatrix} X \\ Y \end{pmatrix} + \beta(t_z - Z) \begin{pmatrix} t_x \\ t_y \end{pmatrix} \right) \quad (39)$$

投影行列  $\Pi$  と投影ベクトル  $\pi$  は次のようになる．

$$\Pi = \begin{pmatrix} 1/\zeta & 0 & -\beta t_x/\zeta \\ 0 & 1/\zeta & -\beta t_y/\zeta \end{pmatrix}, \quad \pi = \begin{pmatrix} \beta t_x t_z/\zeta \\ \beta t_y t_z/\zeta \end{pmatrix} \quad (40)$$

ここに， $\zeta, \beta$  は内部パラメータおよび  $t_x^2 + t_y^2, t_z$  の任意の関数である．

- 関数  $\zeta, \beta$  を次のように定義すると，式 (39) は式 (5) の疑似透視となる．

$$\zeta = \frac{t_z}{f}, \quad \beta = \frac{1}{t_z} \quad (41)$$

- 関数  $\zeta, \beta$  を次のように定義すると，式 (39) は式 (4) の弱透視となる．

$$\zeta = \frac{t_z}{f}, \quad \beta = 0 \quad (42)$$

- 関数  $\zeta, \beta$  を次のように定義すると，式 (39) は式 (3) の平行投影となる．

$$\zeta = 1, \quad \beta = 0 \quad (43)$$

このように，式 (39) は関数  $\zeta, \beta$  の定義の仕方によって平行，弱透視，疑似透視のすべてを包含する．そして，前述のように，条件 1~3 を満たすそのようなモデルは式 (39) 以外には存在しない．この意味で式 (39) はアフィンカメラモデルの唯一の一般形である．

以下では  $\zeta, \beta$  を特定の関数として定義するのではなく，時間と共に変化するパラメータとみなし，入力画像から自己校正によって最適に定める．

## 7. 3次元復元の計算

カメラモデルが与えられれば 3次元復元ができる．計算は次のステップからなる．これは平行，弱透視，疑似透視を含むどのカメラモデルについても共通である（従来のカメラモデルに対する計算の詳細は文献 [9] 参照）．

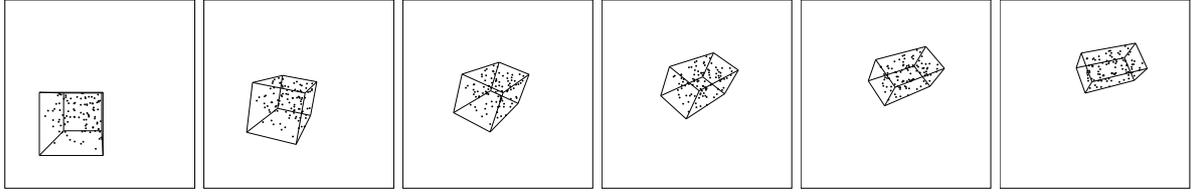


図 5: シミュレーション動画像 .

1. 軌跡ベクトル  $\{p_\alpha\}$  に 3 次元アフィン空間を当てはめる . すなわち , 式 (12) によって重心  $m_0$  を計算し , 式 (15) のモーメント行列  $C$  の大きい 3 個の固有値に対する単位固有ベクトル  $\{u_1, u_2, u_3\}$  を計算する<sup>5</sup> .
2. 計量条件 (28) から計量行列  $T$  と各フレームの投影行列  $\Pi_\kappa$  を計算する .
3. 投影行列  $\Pi_\kappa$  からカメラモデルに従って各フレームの並進  $t_\kappa$  を計算する .
4. 式 (21) の関係から各フレームの回転  $R_\kappa$  を計算する .
5. 計量行列  $T$  を式 (27) のように分解して変換行列  $A$  を計算し , これを用いて式 (16) によってベクトル  $m_1, m_2, m_3$  を計算する .
6. 各特徴点の形状ベクトル  $s_\alpha = (a_\alpha, b_\beta, c_\beta)^\top$  を式 (13) により

$$J = \sum_{\alpha=1}^N \|p'_\alpha - a_\alpha m_1 - b_\alpha m_2 - c_\alpha m_3\|^2 \quad (44)$$

を最小にする最小二乗法で定める . 式 (17) の行列  $M$  を用いれば , 解は次のように与えられる .

$$s_\alpha = M^{-1} p'_\alpha \quad (45)$$

ここに  $M^{-1}$  は  $M$  の一般逆行列である .

ただし , 解には次の不定性が残る .

1. すべての並進  $\{t_\kappa\}$  と形状ベクトル  $\{s_\alpha\}$  に共通の定数を掛けても解である .
2. すべての回転  $\{R_\kappa\}$  に共通の回転行列を掛けても解であり , 形状ベクトル  $\{s_\alpha\}$  もそれに従った変換を受ける .
3. 解  $\{s_\alpha\}, \{R_\kappa\}$  に対して鏡像解  $\{-s_\alpha\}, \{R'_\kappa\}$  が存在する .
4. 重心の奥行き  $t_z$  の絶対位置は定まらない .

1. は遠くの大きな運動と近くの小さな運動が区別できないという , 画像からの 3 次元復元の根本的な不

<sup>5</sup>伝統的な因子分解法 [6, 13, 14, 18] で観測行列  $W$  を  $W = UAV^\top$  と特異値分解することに相当する .

定性である . 2. はシーン座標系の向きは任意に取れることを意味している . 3. はアフィンカメラに特有な不定性であり , アフィンカメラモデルのもとでは除去できない .

4. はカメラモデル (39) に特有な不定性である . 式 (39) からわかるように , 各点の投影はその  $Z$  座標と重心の奥行き  $t_z$  との「差」によって定まるからである . 従来カメラモデルのように  $\zeta, \beta$  を焦点距離  $f$  と奥行き  $t_z$  の特定の関数とすれば , それを解いて  $f$  と  $t_z$  が定まる . しかし , ここでは関数形を特定せず ,  $\zeta$  と  $\beta$  自身を自己校正しているから ,  $f$  と  $t_z$  は定まらない .

## 8. 実験

### 8.1 設定と評価法

3 次元空間の直方体領域にランダムに 100 点を取り , これに並進と回転を加え , 透視カメラで観測した 11 枚のシミュレーションを作成した . 図 5 は見やすいように立方体領域の枠を示し , 1 枚おきに抜き出した 6 枚の画像である . 画像サイズは  $600 \times 600$  画素を想定している . 図 5 は焦点距離  $f = 600$  (画素) の場合であるが , この  $f$  をいろいろに変えてシミュレーション画像列を作る . そして , それぞれの  $f$  に対する透視画像から平行 , 弱透視 , 疑似透視 , および提案モデルを用いて 3 次元復元を行なった .

復元した 3 次元形状は絶対的なスケールと奥行きが不定であるから , 復元した特徴点の 3 次元位置の重心を座標原点に平行移動し , 原点から各特徴点までの平方平均二乗距離が 1 になるようにスケールを調節した . そして , 鏡像解の不定性を除去するために , 真の位置が最も手前にある特徴点と最も背後にある特徴点の復元後の前後関係から正しい形状を判定した .

このように解を正規化した後 , 各特徴点の真の位置と復元位置の間の平方平均二乗距離が最小になるように復元形状を重心の周りに回転し<sup>6</sup> , その最小値

<sup>6</sup>これは特異値分解により容易に計算される [8] .

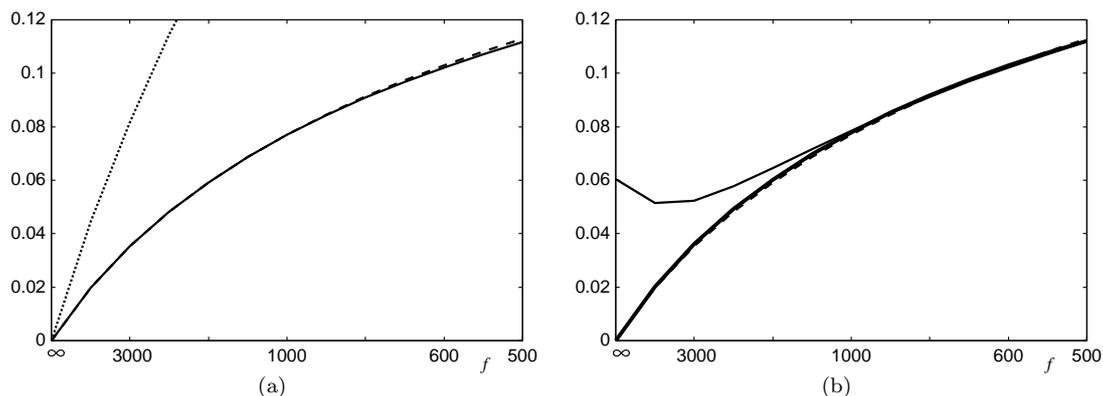


図 6: 焦点距離と各モデルの復元精度の関係。横軸は撮像の焦点距離  $f$  を  $1/f$  に比例する目盛りで表示したもの。(a) 校正済みカメラの場合。点線: 平行, 破線: 弱透視, 実線: 疑似透視。(b) 未校正カメラの場合。破線: 弱透視, 細い実線: 疑似透視, 太い実線: 提案モデル。

を評価した。

## 8.2 校正済みカメラの場合

まず, 既存の平行, 弱透視, 疑似透視の 3 モデルを比較した。このとき, 弱透視と疑似透視の式 (4), (5) に焦点距離  $f$  が含まれている。そこで, カメラは校正済みであると仮定し, 投影像を生成した透視投影の焦点距離の真値  $f$  を用いた (以下, これを「校正済み疑似透視」と呼ぶ)。

図 6(a) は横軸に  $f$  を  $1/f$  に比例する目盛りで表記したものである。 $f = \infty$  が平行投影に相当する。点線が平行, 破線が弱透視, 実線が疑似透視を用いたものである。これから分かるように, 平行投影は  $f$  が非常に大きい場合 (望遠レンズ) 以外は極めて精度が悪い。 $f$  が小さい透視 (広角レンズ) を平行投影で近似するのが無理であるのは直観的にも明らかである。それに対して弱透視は  $f$  が小さくなると近似の程度が低下するものの, その変化は穏やかであり,  $f$  の広い範囲に対してよい近似を与えている。疑似透視も同様であり, 精度がさらに向上している。

## 8.3 未校正カメラの場合

図 6(b) は弱透視, 疑似透視と提案モデルを比較したものであり, 破線が弱透視, 細い実線が疑似透視, 太い実線が提案モデルである。今度はカメラが未校正であると仮定し, 弱透視と疑似透視の  $f$  にはデフォルト値  $f_0 = 600$  (画素) を用いた。しかし, 弱透視の  $f$  は復元形状全体のスケールのみに影響するので, スケールを正規化すれば結果は  $f$  に依存しない。提案モデルはパラメータを入力画像から自己校正するので  $f$  は不要である。

それに対してデフォルト値  $f_0$  を用いる疑似透視

(以下, これを「未校正疑似透視」と呼ぶ) は図 6(b) から分かるように,  $f \approx f_0$  の場合以外は極めて誤差が大きい。これは疑似透視が  $f$  の定める基準点への視線を基準にして弱透視を補正するので (図 4), 誤った  $f$  を用いると不自然な歪みが生じるからである。一方, 提案モデルは自己校正を行うので, 校正済み疑似透視とほとんど同じ結果となる。

## 8.4 考察

以上の実験結果から, 次の結論が得られる。

- 平行投影モデルは撮像が平行投影に近い場合以外は極めて精度が悪い。
- 弱透視モデルは撮像の焦点距離の広い範囲に渡って安定した近似を与える。
- 疑似透視モデルは校正済みなら弱透視モデルよりもよい精度を与える。しかし, 未校正の場合は仮定する焦点距離に大きく依存し, 真の焦点距離と仮定した焦点距離の差が大きくなるにつれて精度が低下する。
- 提案方法は撮像の焦点距離に関らず適切なカメラモデルが自動的に選ばれ, 校正済み疑似透視モデルとほぼ同じ精度を与える。

平行投影に近い画像に疑似透視モデルを用いると精度が悪いことは Deguchi ら [3] が指摘しているが, 従来の文献 [6, 13, 14] では疑似透視の焦点距離への依存性はそれほど強調されていなかった。それに対して, 提案方法は自己校正を行うので真の焦点距離を知る必要がない。

今日, 透視投影に基く精密な 3 次元復元手法 [4] や反復による因子分解法の透視投影への拡張 [1, 2, 12, 15, 19, 21, 22, 23] が研究されているが, 元来の因子分解法の魅力はカメラモデルを詳しく知ることなく,

反復も行わず，簡単な線形計算で容易に3次元復元ができるという点にある．この点で，疑似透視にカメラ校正が必要であるというのはこの精神に反し，提案方法のほうがこの精神に合致している．

## 9. まとめ

本論文では，因子分解法による3次元復元の原理となるアフィン空間拘束条件を，撮像がアフィンカメラであるという以外には何も仮定せずに記述し，これがユークリッド復元となるための計量条件の特定のカメラモデルによらない一般形を導いた．これは射影復元をユークリッド復元に高める「絶対2級束面拘束条件」に対応するものであり，従来のカメラモデルによる3次元復元をすべて包括している．

次に，アフィンカメラが透視投影を近似するための最小限の要請を置くと，2個の不定関数を含むカメラモデルが一意的に定まることを示した．その関数を特定すると平行，弱透視，疑似透視が得られる．本論文ではそれを特定せず，時間とともに変化するパラメータとみなし，入力画像から自己校正によって定めた．これにより，適切なカメラモデルが自動的に選ばれることを実験的に検証した．しかも，計算手順も計算量も基本的に従来モデルと同一である．

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. 17500112) の助成による．

## 参考文献

- [1] S. Christy and R. Horaud, Euclidean shape and motion from multiple perspective views by affine iterations, *IEEE Trans. Patt. Anal. Machine Intell.*, **18-11** (1996-11), 1098–1104.
- [2] K. Deguchi, Factorization method for structure from perspective multi-view images, *IEICE Trans. Inf. & Syst.*, **E81-D-11** (1998-11), 1281–1289.
- [3] K. Deguchi, T. Sasano, H. Arai, and H. Yoshikawa, 3-D shape reconstruction from endoscope image sequences by the factorization method, *IEICE Trans. Inf. & Syst.*, **E79-D-9** (1996-9), 1329–1336.
- [4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- [5] 讓田賢治, 坪内貴之, 菅谷保之, 金谷健一, 移動ビデオカメラ画像からの運動物体の抽出, 情報処理学会研究報告, 2004-CVIM-143-6 (2004-3), pp. 41–48.
- [6] 金出武雄, コンラッド・ポールマン, 森田俊彦, 因子分解法による物体形状とカメラ運動の復元, 電子情報通信学会論文誌 D-II, **J74-D-II-8** (1993-8), 1497–1505.
- [7] K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer-Verlag, Berlin, Germany, 1990.
- [8] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, U.K., 1993.
- [9] 金谷健一, 菅谷保之, 因子分解法の完全レシピ, 電子情報通信学会技術報告, PRMU2003-118 (2003-10), pp. 19–24.
- [10] 黒澤典義, 金谷健一, 部分空間分離法とモデル選択による運動物体の分離, 情報処理学会研究報告, 2000-CVIM-124-4 (2000-11), pp. 25–32.
- [11] 黒澤典義, 金谷健一, アフィン空間分離法による運動物体の分離, 情報処理学会研究報告, 2001-CVIM-125-3 (2001-3), 25–32.
- [12] 宮川勲, 石川裕治, 若林佳織, 有川知彦, 空撮映像からの透視投影型因子分解法による空間連鎖復元, 電子情報通信学会論文誌 D-II, **J87-D-II-4** (2004-4), 942–957.
- [13] T. Morita and T. Kanade, A sequential factorization method for recovering shape and motion from image sequence, *IEEE Trans. Patt. Anal. Mach. Intell.*, **19-8** (1997-8), 858–867.
- [14] C. J. Poelman and T. Kanade, A paraperspective factorization method for shape and motion recovery, *IEEE Trans. Patt. Anal. Mach. Intell.*, **19-3** (1997-3), 206–218.
- [15] P. Sturm and B. Triggs, A factorization based algorithm for multi-image projective structure and motion, *Proc. 4th Euro. Conf. Comput. Vision*, April 1996, Cambridge, U.K., Vol. 2, pp. 709–720.
- [16] 菅谷保之, 金谷健一, 部分空間分離法による特徴点追跡のアウトライア除去, 情報処理学会研究報告, 2002-CVIM-133-24 (2003-5), pp. 177–184.
- [17] 菅谷保之, 金谷健一, 複数運動の教師なし学習による多段階最適化情報処理学会研究報告, 2003-CVIM-138-25 (2003-5), pp. 185–192.
- [18] C. Tomasi and T. Kanade, Shape and motion from image streams under orthography—A factorization method, *Int. J. Comput. Vision*, **9-2** (1992-10), 137–154.
- [19] Factorization methods for projective structure from motion, *Proc. IEEE Conf. Comput. Vision, Patt. Recog.*, Jun 1996, San Francisco, CA, U.S.A., pp. 845–851.
- [20] 坪内貴之, 菅谷保之, 金谷健一, 3次元アフィン復元のための途切れた特徴点追跡の延長情報処理学会研究報告, 2003-CVIM-137-17 (2003-3), 133–140.
- [21] T. Ueshiba and F. Tomita, A factorization method for perspective and Euclidean reconstruction from multiple perspective views via iterative depth estimation, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, Vol. 1, pp. 296–310.
- [22] 植芝俊夫, 富田文明, 奥行きパラメータの逐次推定による多視点透視投影画像のための因子分解法, 電子情報通信学会論文誌 D-II, **J81-D-II-8** (1998-8), 1718–1726.
- [23] 浮田宗伯, 尺長健, 透視投影因子分解法による広範囲画像系列からの3次元物体モデル生成, 電子情報通信学会技術報告 PRMU97-276 (1998-3), 81–88.