

射影因子分解法における奥行き推定が 不要となる条件とそれに基づく三次元復元

村上 洋平^{†1} 伊藤 義道^{†1} 馬場 口 登^{†1}

本報告では、射影因子分解法において、射影奥行き推定が不要となる必要十分条件を導き、その条件に基づく三次元形状復元手法を提案する。撮影の際にカメラの運動をその条件を満たすように制限することにより、射影奥行き推定に伴う膨大な計算を回避することができ、それによって三次元形状復元に要する計算時間を大幅に短縮できるものと期待できる。また、提案手法を用いた場合の計算時間、観測誤差の復元精度への影響、カメラの運動が条件から外れた場合の復元精度への影響を、シミュレーションを通して検証する。

Depth-Estimation-Free Condition in Projective Factorization and Its Application to Metric 3D Reconstruction

YOHEI MURAKAMI,^{†1} YOSHIMICHI ITO^{†1}
and NOBORU BABAGUCHI^{†1}

In this report, we derive a necessary and sufficient condition for depth-estimation-free projective factorization. Based on the condition, we propose a method for metric 3D reconstruction, which is expected that the computational time for 3D reconstruction is significantly reduced compared to existing methods. Through simulation, we evaluate the computational time of the proposed method. We also evaluate the reconstruction accuracy against several types of the perturbations of camera motion as well as measurement noise.

^{†1} 大阪大学大学院工学研究科
Graduate School of Engineering, Osaka University

1. はじめに

射影因子分解法は、実際のカメラをモデル化した透視投影モデルに基づく因子分解法として Strum ら¹⁾ が提案して以来、非常に多くの研究成果が報告されている。それらの研究において中心的な役割を果たすのが射影奥行きの計算法であり、現在ではエビポーラ幾何に基づく手法¹⁾⁻³⁾ や、反復計算に基づく手法⁴⁾⁻¹¹⁾ など、実に様々な有用な手法が提案されている。しかしながら、射影奥行きは全てのフレームにおける全ての特徴点に対して推定しなければならず、三次元形状復元に要する計算時間のほとんどがその計算に費やされる。

一方、陽には述べられていないものの、結果的に射影奥行きを推定せずに射影因子分解法を実行する手法も存在する^{12),13)}。また、文献 14) においては、観測データが特定の条件を満たした場合、射影奥行き推定が不要となることに言及している。しかしながら、そこではいくつかの例を通して断片的に述べているに過ぎず、射影奥行き推定が不要となるための必要十分条件やカメラ運動に関する制約などを、系統的に論じている研究はない。この条件を明らかにすることができれば、その条件を満たすカメラ運動をさせることによって、射影奥行き推定に伴う膨大な計算をすることなく三次元形状復元が可能となる。

そこで本報告では、射影因子分解法において射影奥行き推定が不要となるための必要十分条件を明らかにする。また、その条件を満たすカメラ運動及び点の配置を明らかにし、その条件下における三次元形状復元手法を提案する。さらに、シミュレーションによって射影復元に要する計算時間を、現在提案されている射影復元手法の中で最も高速とされている文献 11) の手法と比較するとともに、観測誤差、及びカメラ運動における奥行き推定不要条件から外れる変動が、復元結果の精度に与える影響を評価する。

2. 射影奥行き推定が不要となるための条件

本節では、射影因子分解法において射影奥行き推定が不要となるための必要十分条件を示す。始めに、カメラモデルについてまとめた後、必要となる概念を定義する。次に、射影奥行き推定が不要となるための必要十分条件を導き、その条件を満たすカメラ運動及び点の配置を明らかにする。

2.1 準備

三次元空間中の N 個の特徴点を、 F 枚のフレームからなる動画像によって観測することを考える。また、第 i フレームを撮影したカメラの投影行列 \mathbf{P}_i ($i = 1, \dots, F$) を

$$\mathbf{P}_i = \mathbf{K} \mathbf{P}_f \mathbf{M}_w^i \quad (1)$$

とする。ただし、

$$\mathbf{K} = \begin{bmatrix} f k_u & f k_s & u_0 \\ 0 & f k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{P}_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{M}_w^i = \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (2)$$

$$\mathbf{R}_i = \begin{bmatrix} r_{11}^i & r_{12}^i & r_{13}^i \\ r_{21}^i & r_{22}^i & r_{23}^i \\ r_{31}^i & r_{32}^i & r_{33}^i \end{bmatrix}, \quad \mathbf{t}_i = \begin{bmatrix} t_1^i \\ t_2^i \\ t_3^i \end{bmatrix} = -\mathbf{R}_i \mathbf{T}_i, \quad \mathbf{T}_i = \begin{bmatrix} t_x^i \\ t_y^i \\ t_z^i \end{bmatrix} \quad (3)$$

であり、 \mathbf{R}_i と \mathbf{T}_i は、それぞれ世界座標系から第 i フレームを撮影したカメラの座標系に変換する回転行列、世界座標系の原点から第 i フレームを撮影したカメラの座標系の原点への並進ベクトルである。

三次元空間中の第 j 特徴点 \mathbf{X}_j ($j = 1, \dots, N$)、および \mathbf{X}_j の第 i フレームへの投影像 \mathbf{x}_{ij} ($i = 1, \dots, F; j = 1, \dots, N$) の斉次座標による表現を、それぞれ

$$\mathbf{X}_j = \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix}, \quad \mathbf{x}_{ij} = \begin{bmatrix} u_{ij} \\ v_{ij} \\ 1 \end{bmatrix} \quad (4)$$

とする。このとき、特徴点 \mathbf{X}_j の第 i フレームへの投影像 \mathbf{x}_{ij} は、次式によって表される。

$$\lambda_{ij} \mathbf{x}_{ij} = \mathbf{P}_i \mathbf{X}_j. \quad (5)$$

ただし、 λ_{ij} は射影奥行きである。式 (5) を $i = 1 \dots F; j = 1 \dots N$ に関して並べると次式を得る。

$$\mathbf{W} = \mathbf{M} \mathbf{S} \quad (6)$$

ただし、 $\mathbf{W} \in \mathbf{R}^{3F \times N}$ 、 $\mathbf{M} \in \mathbf{R}^{3F \times 4}$ 、 $\mathbf{S} \in \mathbf{R}^{4 \times N}$ はそれぞれ次式によって表される。

$$\mathbf{W} = \begin{bmatrix} \lambda_{11} \mathbf{x}_{11} & \lambda_{12} \mathbf{x}_{12} & \dots & \lambda_{1N} \mathbf{x}_{1N} \\ \lambda_{21} \mathbf{x}_{21} & \lambda_{22} \mathbf{x}_{22} & \dots & \lambda_{2N} \mathbf{x}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{F1} \mathbf{x}_{F1} & \lambda_{F2} \mathbf{x}_{F2} & \dots & \lambda_{FN} \mathbf{x}_{FN} \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_F \end{bmatrix}, \quad \mathbf{S} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_N]. \quad (7)$$

\mathbf{W} 、 \mathbf{M} 、 \mathbf{S} を、それぞれ観測行列、運動行列、形状行列と呼ぶ。

以下、射影奥行き推定が不要となるための条件を考察する際に必要な概念を導入する。

定義 1 式 (7) によって与えられる観測行列 \mathbf{W} における射影奥行き λ_{ij} ($i = 1, \dots, F; j = 1, \dots, N$) から、次式によって構成される行列 $\mathbf{\Lambda}$ を射影奥行き行列、あるいは単に、奥行

き行列と呼ぶ。

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1N} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{F1} & \lambda_{F2} & \dots & \lambda_{FN} \end{bmatrix}. \quad (8)$$

□

式 (5) の代わりに、非零の定数 α_i, β_j を用いて $\tilde{\mathbf{P}}_i = \alpha_i \mathbf{P}_i$ 、 $\tilde{\mathbf{X}}_j = \beta_j \mathbf{X}_j$ と置き換えて得られる $\alpha_i \beta_j \lambda_{ij} \mathbf{x}_{ij} = \tilde{\mathbf{P}}_i \tilde{\mathbf{X}}_j$ を用いたとしても物理的な意味は変わらないため、式 (7) の観測行列において、 λ_{ij} を $\alpha_i \beta_j \lambda_{ij}$ に置き換えても問題はないことはよく知られている。このことを踏まえて、以下に射影奥行き推定不要の概念を定義する。

定義 2 奥行き行列において、各行各列に対して非零の定数倍を施すことによって、 $\lambda_{ij} = 1$ ($i = 1, \dots, F; j = 1, \dots, N$) とできる場合、射影奥行き推定は不要である、あるいは単に、奥行き推定は不要であると呼ぶ。 □

本報告では、議論を簡単にするため、次の仮定を置く。この仮定を置いても一般性は失われないことに注意しておく。

仮定 1 第 1 フレームを撮影するカメラの座標系を世界座標系とする。 □

この仮定により、 $\mathbf{R}_1 = \mathbf{I}$ 、 $\mathbf{t}_1 = \mathbf{T}_1 = \mathbf{0}$ となる。ただし、 \mathbf{I} は単位行列を表す。

2.2 奥行き推定が不要となるための必要十分条件の導出

以上の準備の下、以下では奥行き推定が不要であるための必要十分条件を導き、その条件を満たすカメラ運動及び点の配置との関係を示す。まず、次の定理を示す。

定理 1 奥行き推定不要であるための必要十分条件は、奥行き行列 $\mathbf{\Lambda}$ が次式によって表現できることである。

$$\mathbf{\Lambda} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_F \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_N \end{bmatrix}, \quad (9)$$

$$\alpha_i \neq 0 (i = 1, \dots, F), \quad \beta_j \neq 0 (j = 1, \dots, N). \quad (10)$$

□

証明 定義 2 より、 $\mathbf{\Lambda}$ が奥行き推定不要であることは、次式を満たす行列 \mathbf{A} 、 \mathbf{B} が存在する

ことと等価である．

$$\mathbf{A}\mathbf{\Lambda}\mathbf{B} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}. \quad (11)$$

ただし， \mathbf{A} と \mathbf{B} は，それぞれ $\mathbf{R}^{F \times F}$ ， $\mathbf{R}^{N \times N}$ に属する正則な対角行列である．

まず，十分性を示す． $\mathbf{\Lambda}$ が式 (9) 及び式 (10) の条件を満たしているとき，行列 \mathbf{A} ， \mathbf{B} として以下の式によって表される行列を選ぶ．

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\alpha_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\alpha_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\alpha_F} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \frac{1}{\beta_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\beta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\beta_N} \end{bmatrix}. \quad (12)$$

式 (9) の左から行列 \mathbf{A} を，右から行列 \mathbf{B} を掛けると式 (11) が得られる．従って，十分性が証明された．

次に，必要性を示す．式 (11) の行列 \mathbf{A} ， \mathbf{B} はともに正則行列であるため， $\text{rank}(\mathbf{A}\mathbf{\Lambda}\mathbf{B}) = \text{rank}(\mathbf{\Lambda})$ が成り立つ．ただし， $\text{rank}(\mathbf{A})$ は行列 \mathbf{A} のランクである．式 (11) の右辺はランクが 1 であるため， $\mathbf{\Lambda}$ もランクが 1 でなければならない．従って， $\mathbf{\Lambda}$ は式 (9) のように分解できなければならない．式 (9) において，ある i に対して $\alpha_i = 0$ が成り立つとすると，奥行き行列の第 i 行は全て要素が 0 となり，正則な対角行列 \mathbf{A} ， \mathbf{B} をどのように選んでも式 (11) を成立させることは不可能である．従って， $\alpha_i \neq 0$ ($i = 1, \dots, F$) でなければならない． β_j の場合も同様に考えると $\beta_j \neq 0$ ($j = 1, \dots, N$) が導かれる．従って，必要性が証明された．以上より定理 1 が証明された． \square

次に，三次元空間中の任意の点の配置を想定したユークリッド復元を考えた場合の，奥行き推定が不要となるためのカメラ運動に関する定理を与える．

定理 2 三次元空間中の任意の点の配置を想定した場合，以下の 3 条件は等価である．

- (i) 奥行き推定は不要である．
- (ii) 第 i フレーム ($i = 1, \dots, F$) を撮影するカメラの回転行列 \mathbf{R}_i と並進ベクトル \mathbf{T}_i は

次式で与えられる構造を持つ．

$$\mathbf{R}_i = \begin{bmatrix} r_{11}^i & r_{12}^i & 0 \\ r_{21}^i & r_{22}^i & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{T}_i = \begin{bmatrix} t_x^i \\ t_y^i \\ 0 \end{bmatrix}. \quad (13)$$

- (iii) 全てのカメラ座標系の原点が同一平面上にあり，かつ全てのカメラの光軸がその平面に垂直な同一方向を向く．

\square

証明 条件 (ii) と条件 (iii) が等価であることは明らかである．以下では，条件 (i) と条件 (ii) の等価性を示す．

始めに，条件 (i) が成り立つならば条件 (ii) が成り立つことを証明する．式 (1) ~ 式 (4) を用いると，式 (5) の第 3 行は次式のように書ける．

$$\lambda_{ij} = r_{31}^i X_j + r_{32}^i Y_j + r_{33}^i Z_j + t_3^i = \begin{bmatrix} r_{31}^i & r_{32}^i & r_{33}^i & t_3^i \end{bmatrix} \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix}. \quad (14)$$

式 (14) を $i = 1, \dots, F; j = 1, \dots, N$ に関して並べると次式を得る．

$$\mathbf{\Lambda} = \mathbf{M}_{3s} \mathbf{X}_s. \quad (15)$$

ただし， \mathbf{M}_{3s} ， \mathbf{X}_s はそれぞれ以下に示す行列である．

$$\mathbf{M}_{3s} = \begin{bmatrix} r_{31}^1 & r_{32}^1 & r_{33}^1 & t_3^1 \\ r_{31}^2 & r_{32}^2 & r_{33}^2 & t_3^2 \\ \vdots & \vdots & \vdots & \vdots \\ r_{31}^F & r_{32}^F & r_{33}^F & t_3^F \end{bmatrix}, \quad \mathbf{X}_s = \begin{bmatrix} X_1 & X_2 & \dots & X_N \\ Y_1 & Y_2 & \dots & Y_N \\ Z_1 & Z_2 & \dots & Z_N \\ 1 & 1 & \dots & 1 \end{bmatrix}. \quad (16)$$

定理 1 より，奥行き推定不要である場合， $\mathbf{\Lambda}$ について式 (9) が成り立つため， $\mathbf{\Lambda}$ のランクは 1 である．また，三次元空間中の点の配置は任意であることから， N が十分大きい場合 ($N \geq 4$ の場合)， \mathbf{X}_{3s} は行フルランクの行列である．その場合， $\text{rank}(\mathbf{M}_{3s}) = \text{rank}(\mathbf{M}_{3s} \mathbf{X}_s) = \text{rank}(\mathbf{\Lambda})$ が成り立つため¹⁵⁾， \mathbf{M}_{3s} のランクは 1 でなければならない．

仮定 1 より $\mathbf{R}_1 = \mathbf{I}$ ， $\mathbf{t}_1 = \mathbf{0}$ となるため， $[r_{31}^1 \ r_{32}^1 \ r_{33}^1 \ t_3^1] = [0 \ 0 \ 1 \ 0]$ である．このことと， \mathbf{M}_{3s} のランクが 1 であり $[r_{31}^i \ r_{32}^i \ r_{33}^i \ t_3^i]$ ($i = 2, \dots, F$) が $[r_{31}^1 \ r_{32}^1 \ r_{33}^1 \ t_3^1]$ の定数倍でなければならないこと，及び $[r_{31}^i \ r_{32}^i \ r_{33}^i]$ が回転行列 \mathbf{R}_i を構成する行ベクトルであり，そのノルムが 1 であることから，次式が成り立つ．

$$\begin{bmatrix} r_{31}^i & r_{32}^i & r_{33}^i & t_3^i \end{bmatrix} = \begin{bmatrix} 0 & 0 & \pm 1 & 0 \end{bmatrix}. \quad (i = 2, \dots, F) \quad (17)$$

この式から、カメラの回転は光軸方向 (Z 軸方向) の周りの回転に限定されること、及びカメラの並進ベクトルは光軸方向の成分を持たないこと ($t_3^i = t_z^i = 0$) が分かる。ここで、 $r_{33}^i = -1$ とすると、第 1 フレームを撮影したカメラと第 i フレームを撮影したカメラが互いに逆方向を向くことになり、全ての点が全てのフレームに撮影されることに矛盾する。従って、 $r_{33}^i = 1$ でなければならない。従って、 M_{3s} の要素について次式に示す条件が成り立つ。

$$\begin{bmatrix} r_{31}^i & r_{32}^i & r_{33}^i & t_3^i \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \quad (i = 1, \dots, F). \quad (18)$$

式 (18) は、カメラが条件 (ii) を満たすことを意味する。

次に、条件 (ii) が成り立つならば条件 (i) が成り立つことを証明する。式 (18) を式 (14) に代入すると、次式を得る。

$$\lambda_{ij} = Z_j. \quad (19)$$

また、点 X_j はカメラ座標系の X - Y 平面上にないので、 $Z_j \neq 0$ である。これらのことは、 λ_{ij} が定理 1 の条件を満たすことを意味しており ($\alpha_i = 1, \beta_j = Z_j$ と選ぶ)、これによって、条件 (ii) が成り立つならば、条件 (i) が成り立つことが示された。以上より、定理 2 が証明された。□

定理 2 の条件 (iii) が満たされるならば奥行き推定が不要となることは、文献 14) においても例として示されている。定理 2 は、三次元空間中の任意の点配置におけるユークリッド復元を考えた場合、その条件が奥行き推定不要であるための必要条件にもなっていることを示している。

定理 2 は、三次元空間中の任意の点配置に対して奥行き推定が不要となるための必要十分条件を与えている。以下では、点が平面上、あるいは直線上に制限された場合に対して、奥行き推定が不要となるための条件を考える。平面上の点の集合も直線上の点の集合も、行フルランクの行列 Π を適切に選ぶことによって、 $\Pi X = 0$ を満たす X の集合、即ち Π の零化空間として表される。制限された点の集合をこのように表すことによって、奥行き推定が不要となるための必要十分条件を得ることができる。

定理 3 $\Pi \in \mathbf{R}^{c \times 4}$ を行フルランクとし、 $\begin{bmatrix} \mathbf{e}_3^\top \\ \Pi \end{bmatrix}$ も行フルランクであるとする。ただし、 $\mathbf{e}_3^\top = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ とする。また、 $\mathbf{p}_3^{i\top} = \begin{bmatrix} r_{31}^i & r_{32}^i & r_{33}^i & t_3^i \end{bmatrix}$ とおく。このとき、 Π の零化空間における任意の点の配置に対して、奥行き推定が不要となるための必要十分条件

は、以下の条件を満たす $a_i \neq 0, \mathbf{b}_i \in \mathbf{R}^c$ ($i = 2, \dots, F$) が存在することである。

$$\mathbf{p}_3^{i\top} = a_i \mathbf{e}_3^\top + \mathbf{b}_i^\top \Pi, \quad \mathbf{p}_3^{i\top} \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} \mathbf{p}_3^i = 1. \quad (20)$$

ただし、 \mathbf{R}^c は全ての要素が実数からなる c 次元の列ベクトルの集合を表す。□

定理 3 において $\Pi X = 0$ という制約がない場合、式 (20) の条件は式 (18) に一致する。従って、定理 3 は定理 2 の拡張とみなすことができる。定理 3 の証明は紙面の都合上省略する。この証明は文献 16) に示している。

注 1 定理 3 において、 $\begin{bmatrix} \mathbf{e}_3^\top \\ \Pi \end{bmatrix}$ の行フルランク性を仮定しているのは次の理由による。 Π の行フルランク性から、 $\begin{bmatrix} \mathbf{e}_3^\top \\ \Pi \end{bmatrix}$ が行フルランクでないならば、 \mathbf{e}_3^\top は Π の行ベクトルの線形結合で表される。従って、 $\Pi X = 0$ を満たす全ての点 X は、 $\mathbf{e}_3^\top X = 0$ 、すなわち $Z = 0$ を満たす。このことは $\Pi X = 0$ を満たす点 X は、平面 $Z = 0$ の上に限定されることを意味する。仮定 1 よりこの平面は第 1 フレームを撮影するカメラの座標系の X - Y 平面上に一致する。このときそのカメラからはその平面上の点が見えないため、 $\begin{bmatrix} \mathbf{e}_3^\top \\ \Pi \end{bmatrix}$ は行フルランクでなければならない。□

定理 3 の条件を満たす特別な場合として、 $\Pi = \begin{bmatrix} 0 & 0 & 1 & -10 \end{bmatrix}$ の場合を考える。このとき、式 (20) と奥行きが正となることから $a_i + \mathbf{b}_i^\top = 1$ が必要であり、 $r_{31}^i = 0, r_{32}^i = 0, r_{33}^i = 1, t_3^i = -t_z^i = 10(1 - a_i)$ となる。従って、 $a_i \neq 0$ であれば式 (20) の条件が満たされる。これは、カメラの光軸が、点の存在する平面に対して常に垂直であれば、どのような姿勢であっても奥行き推定が不要であることを示している。この例は、文献 14) にも掲載されている。従って、定理 3 は、文献 14) の結果を特殊な場合として含む、より一般的な結果であるといえる。

3. 射影奥行き推定が不要となる条件化における三次元形状復元

本節では、定理 2 の条件を強めた次の仮定の下で、射影復元からユークリッド復元を求める計算法を与える。

仮定 2 全てのカメラは定理 2 の条件 (iii) を満たしており、さらに全てのカメラの姿勢は同じである。□

仮定 2 に基づくカメラ運動は、前節において導いた条件を満たすカメラ運動の一つであ

り, $\mathbf{R}_i = \mathbf{I}$ ($i = 1, \dots, F$) であることを意味する。しかし, 一般にカメラの回転成分が存在しない, あるいは1軸周りのみ存在する場合は自己校正, 即ち, カメラの内部パラメータ \mathbf{K} と外部パラメータ $\mathbf{R}_i, \mathbf{t}_i$ を同時に求めることは不可能であることが知られている¹⁷⁾。従って, カメラ運動に定理2の条件 (iii), あるいは仮定2の制限がある場合, それらを同時に求めることはできない。そこで, 次の仮定を置く。

仮定3 カメラの内部パラメータは既知である。 □

以下,

$$\mathbf{W} = \mathbf{M}_s \mathbf{S}_s \quad (21)$$

の分解が与えられたとして,

$$\mathbf{M} = \mathbf{M}_s \mathbf{H}, \quad \mathbf{S} = \mathbf{H}^{-1} \mathbf{S}_s, \quad (22)$$

がユークリッド復元となるような \mathbf{H} を求める方法を述べる。ただし, 適当な前処理により, \mathbf{M}_s と \mathbf{X}_s は次の条件を満たしているものとする。

$$\mathbf{M}_s = \begin{bmatrix} \mathbf{P}_s^1 \\ \mathbf{P}_s^2 \\ \vdots \\ \mathbf{P}_s^F \end{bmatrix}, \quad \mathbf{S}_s = \begin{bmatrix} \mathbf{X}_s^1 & \mathbf{X}_s^2 & \dots & \mathbf{X}_s^N \end{bmatrix}, \quad \mathbf{P}_s^1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}. \quad (23)$$

ただし, $\mathbf{P}_s^i \in \mathbf{R}^{3 \times 4}$ ($i = 1, \dots, F$), $\mathbf{X}_s^j \in \mathbf{R}^{4 \times 1}$ ($j = 1, \dots, N$) である。ユークリッド復元後の第 i カメラの投影行列を \mathbf{P}_i , 第 j 点を \mathbf{X}_j とすると, 式 (22), 及び式 (23) より, $\mathbf{P}_s^i, \mathbf{X}_s^j$ との間に次式に示す関係が成り立つ。

$$\mathbf{P}_i = \mathbf{P}_s^i \mathbf{H}, \quad (24)$$

$$\mathbf{X}_j = \mathbf{H}^{-1} \mathbf{X}_s^j. \quad (25)$$

仮定1より, $\mathbf{R}_1 = \mathbf{I}$, $\mathbf{t}_1 = \mathbf{0}$ となる。これらと式 (2) を式 (1) に代入すると, 次式を得る。

$$\mathbf{P}_1 = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}. \quad (26)$$

式 (23) の第3式, 及び式 (26) を式 (24) に代入すると, 次式を得る。

$$\begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \mathbf{H} = \mathbf{H}_{3 \times 4}. \quad (27)$$

ただし, $\mathbf{H}_{3 \times 4}$ は \mathbf{H} の上3行を取りだした $\mathbf{R}^{3 \times 4}$ に属する行列である。従って, \mathbf{H} につい

て次式が成り立つ。

$$\mathbf{H} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{v}^\top & k \end{bmatrix}. \quad (28)$$

ただし, $\mathbf{v} \in \mathbf{R}^3$, k は非零の定数である。 \mathbf{H} は正則行列であるため $k \neq 0$ であり, k はスケールのみに関わる変数であるため, $k = 1$ として一般性を失わない。従って, \mathbf{v} を求めることと \mathbf{H} を求めることは等価である。

ここで仮定2より, $\mathbf{R}_i = \mathbf{I}$ ($i = 2, \dots, F$) が成り立つ。これらと式 (2) 及び式 (3) を式 (1) に代入すると, 次式を得る。

$$\mathbf{P}_i = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{I}\mathbf{T}_i \\ \mathbf{0}^\top & 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{I} & -\mathbf{T}_i \end{bmatrix}. \quad (29)$$

ここで, \mathbf{P}_s^i を次のように表記する。

$$\mathbf{P}_s^i = \begin{bmatrix} \mathbf{A}_s^i & \mathbf{B}_s^i \end{bmatrix} \quad (30)$$

ただし, $\mathbf{A}_s^i, \mathbf{B}_s^i$ はそれぞれ $\mathbf{A}_s^i \in \mathbf{R}^{3 \times 3}, \mathbf{B}_s^i \in \mathbf{R}^3$ である。式 (28), 式 (29) 及び式 (30) を式 (24) に代入すると, 次式を得る。

$$\mathbf{K} \begin{bmatrix} \mathbf{I} & -\mathbf{T}_i \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s^i & \mathbf{B}_s^i \end{bmatrix} \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{v}^\top & 1 \end{bmatrix}. \quad (31)$$

式 (31) より, 未知の行列 \mathbf{v} について次式が成り立つ。

$$(\mathbf{I} - \mathbf{A}_s^i) \mathbf{K} = \mathbf{B}_s^i \mathbf{v}^\top \quad (i = 1, \dots, F). \quad (32)$$

本手法では, 式 (32) を厳密に満たす \mathbf{v} を考える代わりに, 次式で与えられる J_1 を最小化する最小二乗解 \mathbf{v} を考える。

$$J_1 = \sum_{i=1}^F \left\| (\mathbf{I} - \mathbf{A}_s^i) \mathbf{K} - \mathbf{B}_s^i \mathbf{v}^\top \right\|^2. \quad (33)$$

ただし, $\|\mathbf{A}\|$ は行列 \mathbf{A} のフロベニウスノルムを表す。この解は次式によって与えられる。

$$\mathbf{v} = \left(\sum_{i=1}^F \mathbf{K}^\top (\mathbf{I} - \mathbf{A}_s^i)^\top \mathbf{B}_s^i \right) \left(\sum_{i=1}^F \mathbf{B}_s^{i\top} \mathbf{B}_s^i \right)^{-1}. \quad (34)$$

式 (34) によって算出された \mathbf{v} を式 (28) に代入することによって, 正則行列 \mathbf{H} が求まる。従って, \mathbf{M}_s 及び \mathbf{S}_s は, 式 (22) によって, ユークリッド空間における運動行列 \mathbf{M} , 形状行列 \mathbf{S} へと変換され, 射影不定性は取り除かれる。

4. シミュレーションによる検討

4.1 射影復元に要する計算時間比較

ノイズの影響を全く受けずに理想的な投影をするという仮定の下において、シミュレーションによって、既存手法と提案手法の射影復元に要する時間を比較した。なお、 X - Y - Z 空間における点の配置パターンは立方体領域内部の点、円柱側面上の点、球面上の点の3種類であり、それぞれ100回の試行の計算時間の平均を取ることにより比較した。また、既存手法として、現在提案されている射影復元の中で最も高速とされている文献11)の手法を用いた。実験に用いた計算機のCPUはInter(R)Core(TM): 7-2600CPU 3.40GBであり、メモリサイズは3.49GBである。

図1に3種類のシミュレーションデータを示す。図1の上段のデータは、重心の座標が(50, 50, 150)、一辺の長さが100である立方体内に100個の点をランダムに配置したものである。このデータのことを以降ではboxと呼ぶ。図1の中段のデータは、底面の円が X - Z 平面上にあり、中心が(50, 0, 150)、半径が50、高さが100である円柱を、底面の円の中心角が 90° となるよう X - Z 平面と垂直に切り出し、その側面上に100個の点を規則的に配置したものである。このデータのことを以降ではcylinderと呼ぶ。図1の下段のデータは中心が(50, 50, 150)、半径が50である球を、大円の中心角が 90° となるよう X - Z 平面と垂直に切り出し、その面上に100個の点を規則的に配置したものである。このデータのことを以降ではsphereと呼ぶ。

また、いずれのデータも、焦点距離 $f = 600$ 、画像中心 $(u_0, v_0) = (240, 160)$ であるカメラを、 $(0, 0, 0)$ において光軸が Z 軸方向に向くよう配置し、 $(0, 0, 0)$ から $(100, 0, 0)$ まで、回転運動をすることなく、 X 方向に1ずつ並進運動させながら計101枚の画像を撮影した状況を想定した。

射影復元に要した時間を表1に示す。既存手法では反復計算を用いて射影奥行きを算出するが、提案手法では射影奥行きを推定しないため、計算時間を大幅に短縮できることが分かる。

	既存手法	提案手法
box	3.532	0.014
cylinder	2.391	0.015
sphere	2.718	0.014

4.2 観測誤差の影響

実環境を撮影した動画像を入力として射影因子分解法を行うためには、動画像内において

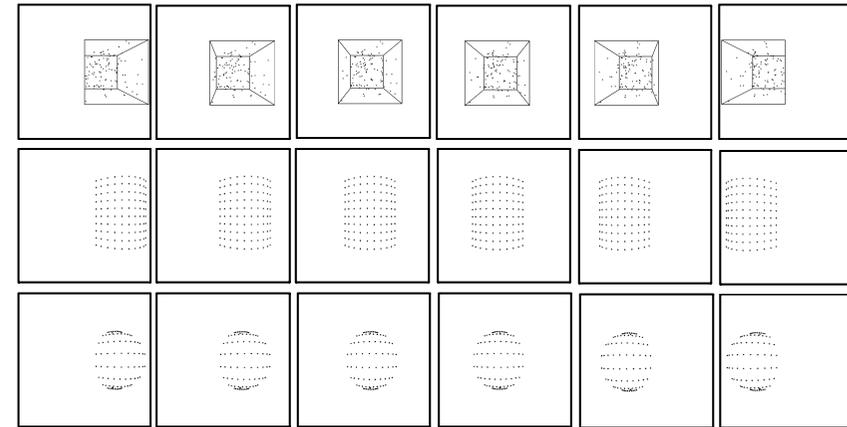


図1 シミュレーションデータ

特徴点の投影像を正確に追跡する必要がある。しかし、実際の追跡結果には観測誤差が生じる。このことを想定し、ノイズを加えたデータを入力として、提案手法による復元誤差をシミュレーションによって算出した。

4.1節と同様のカメラ運動を想定し、box, cylinder, sphereの3種類のデータの各投影像に対して、平均 μ 、標準偏差 σ ピクセルの正規分布に従うランダムな値をノイズとして与えた。そして、ノイズの加えられた各投影像を入力とし、提案手法によって形状復元した。

一般に、射影因子分解法による復元結果には鏡像解の不定性、及びスケールの不定性がある。従って、文献8)と同様の方法によって両不定性を取り除いた。これを取り除くために、世界座標系によって表現される、復元された特徴点の座標を第1フレームを撮影したカメラの座標系によって表現される座標に変換した。そして、各特徴点について、変換された座標の Z の値が正であるか負であるかを判定し、負である特徴点の方が多い場合は座標値に -1 を掛けることにより鏡像解を取り除いた。

一方、スケールの不定性を取り除くためには、まずシミュレーションデータ及び復元結果の重心を計算し、重心が原点となるよう各々を平行移動した。次に、原点から各特徴点までの平均距離が1になるよう正規化し、両データのスケールを揃えた。そして、正規化されたシミュレーションデータの各特徴点と、正規化された復元結果の各特徴点との平均距離が最小になる回転行列を算出し、算出された回転行列によって正規化された復元結果の位置を変換した。そして、正規化されたシミュレーションデータの各特徴点と算出された回転行列に

よって変換された復元結果の各特徴点の平均距離を算出した。

この試行を各 σ ごとに 100 回繰り返して、その平均をノイズの標準偏差が σ であるときの復元誤差とした。 σ を 1 から 10 まで 1 ずつ変化させ、復元結果がノイズから受ける影響を評価した。結果を図 2 に示す。

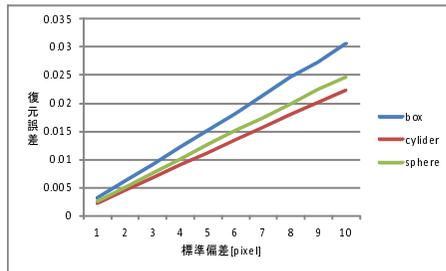


図 2 観測誤差による復元誤差

図 2 より、提案手法による復元結果は観測誤差に対して安定であることが分かる。

4.3 奥行き推定不要条件から外れる変動の影響

第 2 節において示したように、射影奥行き推定不要となる条件を満たすカメラ運動は、定理 2 の仮定 (iii) を満たす運動である。4.1 節及び 4.2 節におけるシミュレーションでは、全てのカメラ座標系の原点が X - Y 平面上にあり、かつ全ての光軸が Z 軸方向を向いていたため、この条件を満たしていた。このような条件を厳密に満足させることは、現実的には非常に困難であると考えられる。そこで、このような条件からカメラの姿勢や動きが外れて変動した場合の影響を、シミュレーションによって検証する。

奥行き推定不要条件から外れる変動として、 Z 軸方向の並進運動、 X 軸周りの回転運動、及び Y 軸周りの回転運動を考え、これらのカメラ運動によって復元結果が受ける影響を評価した。

Z 軸方向の並進運動

4.1 節と同様のカメラ運動に対し、平均 0、標準偏差 σ の正規分布に従うランダムな値を、 Z 軸方向の並進運動として各カメラの運動に付与した。そして、box, cylinder, sphere の 3 種類のデータを投影し、提案手法によって形状復元した。この試行を各 σ ごとに 100 回繰り返して、その平均をノイズの標準偏差が σ であるときの復元誤差とした。 σ を 2 から 20 まで 2 ずつ変化させ、復元結果が Z 軸方向の並進運動から受ける影響を評価した。結果を図

3 に示す。図 3 より、 Z 軸方向の並進運動に起因する復元誤差は観測誤差に起因する復元

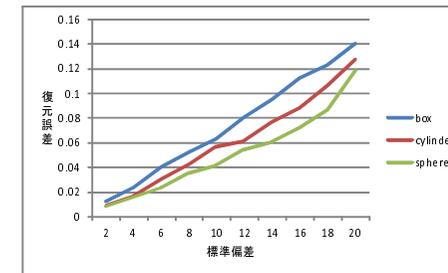


図 3 Z 軸方向の並進運動による復元誤差

誤差よりも大きく、変動の標準偏差に対して線形に変化する傾向が見られる。

X 軸周りの回転運動

4.1 節と同様のカメラ運動に対し、平均 0、標準偏差 σ 度の正規分布に従うランダムな値を、 X 軸周りの回転角度として各カメラの運動に付与した。そして、box, cylinder, sphere の 3 種類のデータを投影し、提案手法によって形状復元した。この試行を各 σ ごとに 100 回繰り返して、その平均をノイズの標準偏差が σ であるときの復元誤差とした。 σ を 2 から 20 まで 2 ずつ変化させ、復元結果が X 軸周りの回転運動から受ける影響を評価した。結果を図 4 に示す。

Y 軸周りの回転運動

X 軸周りの回転運動と同様のランダムな値を Y 軸周りの回転角度として与え、同様の評価によって、復元結果が Y 軸周りの回転運動から受ける影響を評価した。結果を図 5 に示す。

図 4 及び図 5 から、回転運動に起因する復元誤差は観測誤差に起因する誤差よりも大きく、さらに角度の変動に対して線形に変化するのではなく、急激に悪化する傾向が見られた。この結果より、カメラ運動が光軸を除く軸周りの回転成分を持っているときは、復元誤差が大きいことが分かる。

5. おわりに

本論文では、射影因子分解法において射影奥行き推定が不要となる必要十分条件を導出した。その結果、三次元空間中における任意の点の配置に対して射影奥行き推定が不要となる

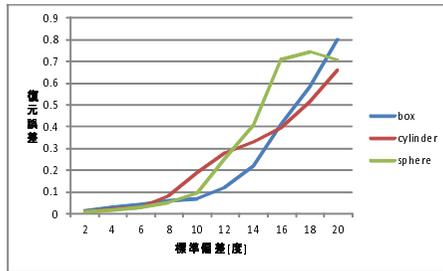


図4 X 軸周りの回転運動による復元誤差

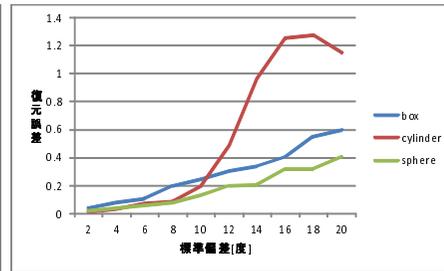


図5 Y 軸周りの回転運動による復元誤差

カメラ運動は、全てのカメラ座標系の原点が同一平面上にあり、かつ全てのカメラの光軸がその平面に垂直な同一方向を向くことであることを示した。また、点の配置が平面上や直線上に限定された場合にも、射影奥行き推定が不要となる必要十分条件を導いた。この結果は、Hartley の例¹⁴⁾ の特殊な場合を含む、より一般的な結果である。さらに、射影奥行き推定不要となるカメラ運動のうち、回転運動及び Z 軸方向の運動成分が存在しない場合は、ユークリッド復元が、最小二乗法によって代数的に算出可能であることを示した。

また、既存手法¹¹⁾ と射影復元に要する時間を比較し、提案手法では大幅に短縮できることを示した。さらに、シミュレーションによって、提案手法の復元結果は観測誤差に対しては安定であるものの、光軸を除く軸周りの回転には不安定であることを示した。

今後の課題として、既存の射影因子分解法との復元精度の比較が挙げられる。提案手法では、射影奥行きを推定しないことから、観測誤差が大きくなる場合、もしくは奥行き推定不要条件からのカメラの姿勢や運動の変動が大きくなる場合は、復元精度が既存の射影因子分解法の復元結果と比較して下がることが予想される。両手法の性能を比較することにより、提案手法の適用範囲を明らかにすることが今後の課題である。

参 考 文 献

- 1) Strum, P. and Triggs, W.: A Factorization Based Algorithm for Multi-image Projective Structure and Motion, *In Proceedings of the 4th European Conference on Computer Vision*, Vol.2, pp.709–720, (1996).
- 2) 出口光一郎：多視点からの透視投影像の因子分解による 3 次元形状復元，計測自動制御学会論文誌，Vol.34, No.10, pp.1321–1328, (1998).
- 3) 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，パターン認識・メディア理解研究会，Vol.97,

No.596, pp.81–88, (1998).

- 4) Ueshiba, T. and Tomota, F.: A Factorization Method for Perspective and Euclidean Reconstruction from Multiple Perspective Views Via Iterative Depth Estimation,” *In Proceedings of the 5th European Conference on Computer Vision*, Vol.1, pp.296–310, (1998).
- 5) Heyden, A. et al.: An Iterative Factorization Method for Projective Structure and Motion From Image Sequences,” *Image and Vision Computing*, Vol.17, No.13, pp. 981–991, (1999).
- 6) Mahamud, S. and Hebert, M.: Iterative Projective Reconstruction From Multiple Views, *In Proceedings of the 14th IEEE Conference on Computer Vision and Pattern Recognition*, pp.430–437, (2000).
- 7) Han, M. and Kanade, T.: Multiple Motion Scene Reconstruction From Uncalibrated Views, *In Proceedings of the 8th IEEE International Conference on Computer Vision*, pp.163–170, (2001).
- 8) 金谷健一，森昭延，菅谷保之：自己校正法の最新レシピ，情報処理学会研究報告，コンピュータビジョンとイメージメディア研究会，2006-CVIM-153-31, pp.199–206, (2006).
- 9) 森昭延，金谷健一，菅谷保之：最新の自己校正法の性能評価，情報処理学会研究報告，コンピュータビジョンとイメージメディア研究会，2006-CVIM-154-37, pp.347–354, (2006).
- 10) Ackermann, H. and Kanatani, K.: Fast Projective Reconstruction: Toward Ultimate Efficiency, *IPSJ Transactions on Computer Vision and Image Media*, Vol.49, No.SIG 6, pp.68–78, (2008).
- 11) ハノ・アッカーマン，新妻弘崇，金谷健一：自己校正法のための射影復元の計算量削減，情報処理学会研究報告，コンピュータビジョンとイメージメディア研究会，2007-CVIM-160-11, pp.63–70, (2007).
- 12) 宮川勲，石川裕治，若林佳織，荒川賢一：車両運動投影モデルに基づく全方位画像系列からの市街地空間の 3 次元形状復元，情報処理学会論文誌，コンピュータビジョンとイメージメディア，Vol.45, No.SIG 13 (CVIM10), pp.34–53, (2004).
- 13) 宮川勲，若林佳織，荒川賢一：魚眼投影モデルに基づく画像系列からの平面運動と 3 次元形状復元，電子情報通信学会論文誌，Vol.J87-D-II, No.5, pp.1120–1132, (2004).
- 14) Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press (2004).
- 15) Harville, D.A.: *Matrix Algebra From a Statistician's Perspective*, Springer, (2008).
- 16) 村上洋平：射影奥行き推定が不要となる条件下における三次元形状復元，修士論文，大阪大学大学院工学研究科 (2012).
- 17) Strum, P.: Critical Motion Sequences for Monocular Self-calibration and Uncalibrated Euclidean Reconstruction, *In Proceedings of the 11th IEEE Conference on Computer Vision and Pattern Recognition*, pp.1100–1105, (1997).