

未校正マルチカメラの相互投影を用いた弱校正と物体形状の復元

杉 村 良 彦[†] 佐 藤 淳[†]

マルチカメラ間の位置や姿勢の情報は対象物形状を復元したり、任意視点画像を生成したりするうえで非常に重要である。カメラが2台存在する場合にはこれらのカメラ間の関係はエピポーラ幾何で表される。一般にエピポーラ幾何を安定に計算することは難しいことが知られているが、近年お互いのカメラを相互に投影しあうことでエピポーラ幾何を安定に計算することができる相互投影法が提案された。そこで本論文では、相互投影法を拡張し一般の N 台のカメラ間の関係を相互投影を用いて計算する方法を提案する。相互投影法を用いることで各カメラ間のエピポーラ幾何が安定に求まるとともに、ある1台のカメラを基準にすべてのカメラが弱校正可能であることを示す。また、提案法により対象物形状復元を行い、その有効性を示す。

Weak Calibration of Multiple Cameras and Reconstruction of 3D Objects from Mutual Projection of Cameras

YOSHIHIKO SUGIMURA[†] and JUN SATO[†]

The relative position and orientation of multiple cameras are very important for reconstructing 3D objects and for generating arbitrary views of the scene. The relationship between these multiple cameras can be described by the epipolar geometry. Recently, it has been shown that the epipolar geometry can be computed accurately and efficiently by using the mutual projection of cameras. In this paper, we propose a method for calibrating a large amount of multiple cameras efficiently by extending the mutual projection method for two cameras. We also show that the proposed method can be applied for reconstructing 3D objects accurately.

1. はじめに

近年コンピュータビジョンにおいては、複数の視点から得られたカメラ画像から対象物の3次元形状の復元を行ったり、ロボットを目的地へ誘導したりする視覚誘導などの研究が行われている。さらに、複数のカメラ画像から物体の形状を復元し、仮想空間や現実空間へ融合する複合現実感などの研究も進められている。このように、複数のカメラ画像を用いて3次元形状を復元したり視覚誘導したりするには、複数のカメラの位置や姿勢などのカメラ情報を取得することが必要である。このような複数のカメラ間の相対的な位置や姿勢などの情報はエピポーラ幾何によって表される。そしてエピポーラ幾何は、2枚の画像間では7自由度の fundamental 行列、3枚の画像間では18自由度の trifocal tensor、4枚の画像間では29自由度の quadrifocal tensor で記述され、画像を2枚から4枚まで増やすことによりカメラ画像間においてより多く

の拘束条件が得られることが知られている⁴⁾。

これまでに fundamental 行列や trifocal tensor は画像上の複数の投影点から線形解法により求める方法が提案されている^{1),3)}。しかし、このような計算法は画像ノイズや投影条件に大きく影響を受けるため、安定に計算するのは非常に難しいことが知られている。これに対し著者らは2つのカメラを相互に投影しあうことにより、それらのカメラの投影像として画像中から直接エピポールの情報を取得し、この情報を用いることにより2つのカメラ間の fundamental 行列を安定に計算する手法⁸⁾や、3つのカメラ間の trifocal tensor を安定に計算する手法⁵⁾、そして4つのカメラ間の quadrifocal tensor を計算する手法⁷⁾を提案してきた。

本論文では2つのカメラ間の相互投影によるエピポーラ幾何計算法⁸⁾を拡張し、 N 台のカメラが相互に投影されている場合に、これら N 台のカメラ間のエピポーラ幾何を安定に求め形状復元する手法を提案する。

[†] 名古屋工業大学

Nagoya Institute of Technology

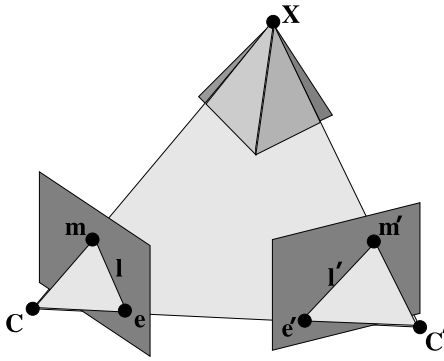


図 1 エピポーラ幾何

Fig. 1 Epipolar geometry.

2. エピポーラ幾何

よく知られているように、2つのカメラの相対的な位置や姿勢などの関係はエピポーラ幾何によって記述することができる。エピポーラ幾何では、それぞれのカメラはエピポール e, e' として互いの画像面に投影される。今、図1のように空間中の点 X がそれぞれのカメラに投影されているとし、その投影点を m, m' とする。これらにはエピポーラ方程式と呼ばれる次式で示す関係式が成り立つ。

$$\tilde{\mathbf{m}}'^T \mathbf{F} \tilde{\mathbf{m}} = 0 \quad (1)$$

ここで $(\sim)$ は斉次座標であることを表す。このとき、fundamental 行列 $\mathbf{F}$ は 3×3 の自由度 7 の行列である。

$\mathbf{F}$ 行列はそれぞれ自由度 2 のエピポール e, e' と自由度 3 のエピポーラ・ホモグラフィー $\mathbf{H}$ によって次のように記述することができる⁶⁾。

$$\mathbf{F} = \begin{bmatrix} \mathbf{RH} & -\mathbf{RHe} \\ -\mathbf{e}'^T \mathbf{RH} & \mathbf{e}'^T \mathbf{RHe} \end{bmatrix} \quad (2)$$

ここで $\mathbf{R}$ は

$$\mathbf{R} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3)$$

である。エピポール e, e' はカメラの並進運動を表し、エピポーラ・ホモグラフィー $\mathbf{H}$ は一方の画像中のエピポーラ線をもう一方の画像中のエピポーラ線に変換する 1 次元射影変換である。

このような $\mathbf{F}$ 行列を計算する方法として 8 点法が広く用いられているが、これは次式の線形方程式を解くことにより求めるものである³⁾。

$$\mathbf{Df} = 0 \quad (4)$$

ここで $\mathbf{F}$ 行列の i 行 j 列の要素を f_{ij} 、各投影点を $\mathbf{m}_i = [x_i, y_i]^T$ 、 $\mathbf{m}'_i = [x'_i, y'_i]^T$ とすると、

$$\begin{aligned} \mathbf{f} &= [f_{11}, \dots, f_{33}]^T \\ \mathbf{D} &= [\mathbf{Z}_1^T, \dots, \mathbf{Z}_8^T]^T \\ \mathbf{Z}_i &= [x_i x'_i, y_i x'_i, x'_i, x_i y'_i, y_i y'_i, y'_i, x_i, y_i, 1] \end{aligned}$$

である。このとき、画像座標 m, m' を正規化することにより、線形計算を安定化させるのが一般的である³⁾。

しかし、このような 8 点法は、対応点が少なかったり、対応点が 3 次元空間中で平面的であったりするなど縮退に近い場合には、画像ノイズに対して非常に敏感であることが知られている。

そこで、本論文ではカメラを相互に投影することにより画像中からエピポールを直接求めることを考える。この場合には非常に精度良くエピポールを求めることが可能である。したがって、 $\mathbf{F}$ 行列の要素の一部をあらかじめ精度良く求めることができるので、より少ない投影点から安定に $\mathbf{F}$ 行列全体を計算することが可能になる⁸⁾。本論文では、この相互投影によるエピポーラ幾何計算法を一般の N 台のカメラの場合に拡張する。

3. 相互投影を用いたマルチカメラの校正

3.1 2 台のカメラの相互投影法

まず 2 台のカメラ間のエピポーラ幾何を相互投影法により求める方法⁸⁾ について簡単に説明する。

$\mathbf{F}$ 行列とエピポール e, e' には次の関係がある。

$$\mathbf{F}\tilde{\mathbf{e}} = 0 \quad (5)$$

$$\mathbf{F}^T \tilde{\mathbf{e}}' = 0 \quad (6)$$

2 台のカメラが相互に投影されている場合、カメラの像としてエピポール e, e' を画像中から直接求めることができる。このようにして得られた e, e' より次に示す $\mathbf{f}$ に関する線形拘束が得られる。

$$\mathbf{M}_1 \mathbf{f} = 0 \quad (7)$$

ここで $\mathbf{M}_1$ は次に示すとおりである。

$$\mathbf{M}_1 = \begin{bmatrix} e_u & e_v & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e_u & e_v & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & e_u & e_v & 1 \\ e'_u & 0 & 0 & e'_v & 0 & 0 & 1 & 0 & 0 \\ 0 & e'_u & 0 & 0 & e'_v & 0 & 0 & 1 & 0 \\ 0 & 0 & e'_u & 0 & 0 & e'_v & 0 & 0 & 1 \end{bmatrix}$$

式 (7) には 6 つ式があるが、このうち独立なものは 5 つのみである。したがって、一般の対応点 3 点によ

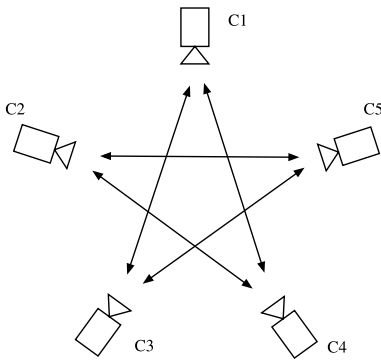


図 2 カメラ配置

Fig. 2 Configuration of cameras.

る拘束式と組み合わせることで、次式より f が求まる。

$$Mf = 0 \quad (8)$$

ここで M は次に示す 9×9 行列である。

$$M = \begin{bmatrix} M_1^T & Z_1^T & Z_2^T & Z_3^T \end{bmatrix}^T$$

以上のように 2 つのカメラが相互投影されている場合には最低 3 点の対応点より F 行列が求まる。この相互投影によるエピポーラ幾何計算法では、対応点が少なくすむのみでなく、求まる F 行列が、従来の 8 点法に比べて格段に安定化することが分かっている⁸⁾。

3.2 マルチカメラの相互投影

そこで本論文では 3.1 節で示した 2 台のカメラ間の相互投影法を拡張して、 N 台のカメラを相互に投影しあうことで、これら N 台すべてのカメラ間の校正を少ない対応点から安定に行う手法を提案する。

一般にカメラの台数が増えるにつれ、これらのカメラが相互に投影される確率が高くなる。しかしカメラの画角は限られているため、すべてのカメラが相互に投影されることはない。そこで本研究では、各カメラに他の 2 台のカメラが投影されており、このような相互投影がチェーン状につながっている状態を考える。たとえば図 2 に示すように 5 台のカメラ C_1, C_2, C_3, C_4, C_5 が存在する場合には、 C_1 C_3 間、 C_3 C_5 間、 C_5 C_2 間、 C_2 C_4 間、 C_4 C_1 間でそれぞれ相互投影が起こっているとすると、以下では、 N 台の各カメラに他の 2 台のカメラが相互投影されており、この相互投影がチェーン状につながっている、すなわち、どのカメラへも相互投影のつながりをたどることにより行き着くことのできるような連結グラフとなっているという条件のもとで、すべてのカメラ間のエピポーラ幾何の計算を行い、全カメラを弱校正する方法を提案する。なお、このような相互投影の連結グラフは閉じている必要はない。

3.3 相互投影によるマルチカメラの校正

簡単のため、まず 5 台のカメラが相互に投影されている場合について説明する。図 2 のように $C_1, \dots, C_5$ の 5 台のカメラが配置されているとする。このとき、 C_1 には C_3 および C_4 が投影されているとする。同様にその他のカメラにおいても、そのカメラに隣り合わない 2 つのカメラが投影されているとする。

すると 3.1 節で述べた相互投影法により、カメラの投影像を利用しながらカメラ C_1 とカメラ C_3 の間のエピポーラ幾何が計算できる。同様にして C_3 と C_5 間、 C_5 と C_2 間、 C_2 と C_4 間、 C_4 と C_1 間においても、相互投影法によりエピポーラ幾何を計算することができる。以上のことはそのまま N 台のカメラに拡張して考えることができる。すなわち、 N 台のカメラが存在する状態において、どのカメラにおいてもその他の 2 台のカメラが投影されており、かつこの相互投影が連結グラフ状となっている場合、以上に述べた方法を繰り返していくことにより、すべてのカメラ間の F 行列を求めることができる。

2 台のカメラ C_i, C_j 間の fundamental 行列 F_{ij} が求まると、 i 番目のカメラ C_i のカメラ座標を基準とした i 番目と j 番目のカメラのカメラ行列 P_{ii}, P_{ij} がそれぞれ以下のように求まる²⁾。

$$P_{ii} = \begin{bmatrix} I & 0 \end{bmatrix} \quad (9)$$

$$P_{ij} = \begin{bmatrix} M_{ij} & e_{ji} \end{bmatrix} \quad (10)$$

ここで、 P_{ij} はカメラ C_i を基準としたカメラ C_j のカメラ行列であり、 M_{ij} は次に示す 3×3 行列である。

$$M_{ij} = [e_{ji}]_{\times} F_{ij} \quad (11)$$

また e_{ji} はカメラ i に関するカメラ j 中のエピポールを表し、 $[\cdot]_{\times}$ は 3 次元ベクトルの 3 つの要素よりなるベクトル積を表す歪対称行列である⁶⁾。これより、カメラ C_1 とカメラ C_3 間の F_{13} 行列を用いて、カメラ C_1 を基準としたカメラ C_1 のカメラ行列 P_{11} とカメラ C_3 のカメラ行列 P_{13} をそれぞれ次のように求めることができる。

$$P_{11} = \begin{bmatrix} I & 0 \end{bmatrix} \quad (12)$$

$$P_{13} = \begin{bmatrix} M_{13} & e_{31} \end{bmatrix} \quad (13)$$

同様にして、カメラ C_3 とカメラ C_5 間の F_{35} 行列を用いて、カメラ C_3 を基準としたカメラ C_3 のカメラ行列 P_{33} とカメラ C_5 のカメラ行列 P_{35} がそれぞれ次のように計算できる。

$$\mathbf{P}_{33} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \quad (14)$$

$$\mathbf{P}_{35} = \begin{bmatrix} \mathbf{M}_{35} & \mathbf{e}_{53} \end{bmatrix} \quad (15)$$

同様に他の 2 台のカメラ間でもカメラ行列を求めることができる。

しかし、このようにして求めたカメラ行列は一方のカメラを基準としたもう一方のカメラのカメラ行列であり、全体として 1 つの基準座標に統一されていない。そこで、次にカメラ C_1 を基準としたカメラ行列に統一する方法を示す。上記の方法でカメラ C_3 のカメラ行列は 2 通り求まっている。すなわち、カメラ C_1 を基準とした $\mathbf{P}_{13}$ とカメラ C_3 を基準とした $\mathbf{P}_{33}$ である。そこで以下のようにこれらのカメラ行列を変換する射影変換行列 $\mathbf{H}_{13}$ を考える。

$$\mathbf{P}_{13} \sim \mathbf{P}_{33}\mathbf{H}_{13} \quad (16)$$

ここで $(\sim)$ は定数倍の不定性を除いて等しいことを表す。このような射影変換行列 $\mathbf{H}_{13}$ が求まれば、カメラ C_1 を基準としたカメラ C_5 のカメラ行列 $\mathbf{P}_{15}$ は $\mathbf{P}_{35}$ より以下のように求めることができる。

$$\mathbf{P}_{15} \sim \mathbf{P}_{35}\mathbf{H}_{13} \quad (17)$$

そこで $\mathbf{H}_{13}$ の計算法について考える。まず、 $\mathbf{P}_{13}$ および $\mathbf{P}_{33}$ が得られているのでこれらを用いて式 (16) より $\mathbf{H}_{13}$ に関する拘束が得られる。カメラ行列 $\mathbf{P}_{ij}$ は 11 自由度なので式 (16) は $\mathbf{H}_{13}$ に関して 11 の拘束を与える。一方式 (12), (13) よりカメラ行列が得られると、これらより画像に投影された点をカメラ C_1 を基準に射影復元することができる。この復元結果を $\mathbf{X}_1$ とする。さらにこの点が C_5 にも投影されている場合には、式 (14) と式 (15) より求めたカメラ行列よりカメラ C_3 を基準にこの点を射影復元することができる。この結果を $\mathbf{X}_3$ とすると、 $\mathbf{X}_1$ と $\mathbf{X}_3$ の間には次の関係がある。

$$\tilde{\mathbf{X}}_3 \sim \mathbf{H}_{13}\tilde{\mathbf{X}}_1 \quad (18)$$

ここで、式 (18) の両辺に $\mathbf{P}_{33}$ をかけて式 (18) の上 3 行のみ取り出して考えることにする。

$$\mathbf{P}_{33}\tilde{\mathbf{X}}_3 \sim \mathbf{P}_{33}\mathbf{H}_{13}\tilde{\mathbf{X}}_1 \quad (19)$$

復元結果 $\mathbf{X}_1$ を $\mathbf{P}_{13}$ で投影したものと $\mathbf{X}_3$ を $\mathbf{P}_{33}$ で投影したものはともにカメラ C_3 の投影像を表しているため、次式が成り立つ。

$$\mathbf{P}_{13}\tilde{\mathbf{X}}_1 \sim \mathbf{P}_{33}\tilde{\mathbf{X}}_3 \quad (20)$$

式 (20) を式 (19) に代入することにより次式を導くことができる。

$$\mathbf{P}_{13}\tilde{\mathbf{X}}_1 \sim \mathbf{P}_{33}\mathbf{H}_{13}\tilde{\mathbf{X}}_1 \quad (21)$$

式 (16) と式 (21) より、式 (16) が与えられた場合、式 (19) は $\mathbf{H}_{13}$ に関して何も新しい情報を与えないことが分かる。したがって、式 (16) と式 (18) を同時に用いる場合には、式 (18) からは $\mathbf{H}_{13}$ に関して独立な拘束が 1 つのみ得られる。よって、空間中に 4 つの対応点があれば式 (18) より $\mathbf{H}_{13}$ に関して 4 つの拘束が得られる。これを式 (16) による拘束と結合させることにより 15 の拘束が得られ、15 自由度の $\mathbf{H}_{13}$ を計算することができる。ただし、このとき用いる 4 点は同一平面上にあってはならない。

次に、 C_5 を基準とした C_2 のカメラ行列 $\mathbf{P}_{52}$ を C_1 を基準としたカメラ行列 $\mathbf{P}_{12}$ に変換する行列 $\mathbf{H}_{15}$ を計算する方法を示す。 C_3 を基準とした C_5 のカメラ行列 $\mathbf{P}_{35}$ と C_5 を基準とした C_5 のカメラ行列 $\mathbf{P}_{55}$ の間には次のような関係がある。

$$\mathbf{P}_{35}\mathbf{H}_{13} \sim \mathbf{P}_{55}\mathbf{H}_{15} \quad (22)$$

$\mathbf{H}_{13}$ は先に得られているため、式 (22) より $\mathbf{H}_{15}$ に関して 11 の拘束が得られる。また、この射影変換行列 $\mathbf{H}_{15}$ と射影復元結果には以下の関係が成り立つので、1 つの点がカメラ C_1, C_2, C_3, C_5 に共通に映っていれば、 $\mathbf{H}_{15}$ に関して 1 つの拘束が得られる。

$$\tilde{\mathbf{X}}_5 \sim \mathbf{H}_{15}\tilde{\mathbf{X}}_1 \quad (23)$$

ここで $\mathbf{X}_5$ は C_5, C_2 による復元結果である。したがって 4 点が共通に投影されていれば、式 (22), (23) より $\mathbf{H}_{15}$ に関して 15 の拘束が得られ、 $\mathbf{H}_{15}$ を計算することができる。このようにして得られた $\mathbf{H}_{15}$ よりカメラ C_1 を基準としたカメラ C_2 のカメラ行列 $\mathbf{P}_{12}$ は以下のように計算できる。

$$\mathbf{P}_{12} \sim \mathbf{P}_{52}\mathbf{H}_{15} \quad (24)$$

以下、同様の処理を繰り返すことにより、カメラ C_1 を基準とした各カメラのカメラ行列 $\mathbf{P}_{11}, \mathbf{P}_{12}, \mathbf{P}_{13}, \mathbf{P}_{14}, \mathbf{P}_{15}$ を計算することができる。以上より、 N 台のカメラ間において相互投影の連鎖が存在する場合には、最低 4 点の対応点より N 台すべてのカメラが弱校正可能であることが分かる。このようにしてすべてのカメラがある基準座標をもとに弱校正できれば、通常の復元法により、シーン中の物体の 3 次元形状を復元することができる。

3.4 相互投影の必要条件

以上のとおり、各カメラに他の 2 台のカメラが相互投影されており、このような相互投影が連結グラフをなしている場合には、相互投影情報を用いてすべてのカメラを完全に校正することができる。このようなマ

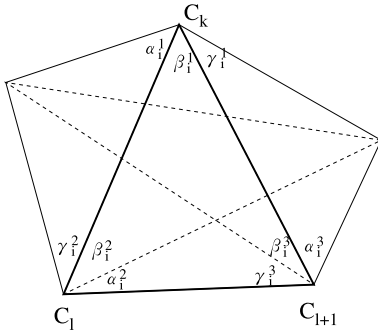


図 3 視野角

Fig. 3 Field of view.

ルチカメラを用いる場合には、一般に対象物を囲むようにカメラを配置し、対象物の全周形状を復元する場合が多い。そこで本節ではこのような N 台のカメラが対象物を囲むように内向きに配置され、かつそれぞれのカメラに向かい合う 2 台のカメラが投影されるような相互投影の連結グラフが構成されるための必要条件について考える。

一般の N 角形の頂点にカメラを配置し、カメラ $C_k (k = 1, \dots, N)$ が向かい合う 2 つのカメラ $C_l, C_{l+1} (l = 1, \dots, N)$ を映す場合を考える。代表的な例として、5 角形の場合を図 3 に示す。図 3 のようにカメラ C_k, C_l, C_{l+1} の角に関して $\alpha_i^j, \beta_i^j, \gamma_i^j$ ($j = 1, \dots, 3, i = 1, \dots, N$) を考える。このとき、カメラ C_k, C_l, C_{l+1} からなる三角形の内角の和は π であることから以下の式が成り立つ。

$$\beta_i^1 + \beta_i^2 + \alpha_i^2 + \beta_i^3 + \gamma_i^3 = \pi \quad (25)$$

N 個の頂点すべてに関してこのような三角形を考えてその内角の総和 S をとると以下のとおり $N\pi$ となる。

$$S = \sum_{i=1}^N (\beta_i^1 + \beta_i^2 + \alpha_i^2 + \beta_i^3 + \gamma_i^3) = N\pi \quad (26)$$

この式は以下のように分解できる。

$$S = \sum_{i=1}^N (\alpha_i^2 + \beta_i^1 + \gamma_i^3) + \sum_{i=1}^N \beta_i^2 + \sum_{i=1}^N \beta_i^3 \quad (27)$$

ここで、 $\sum_{i=1}^N \beta_i^1 = \sum_{i=1}^N \beta_i^2 = \sum_{i=1}^N \beta_i^3$ であるので、これらを $\sum_{i=1}^N \beta_i$ と表記することにする。また、 α, γ についても同様に $\sum_{i=1}^N \alpha_i, \sum_{i=1}^N \gamma_i$ と表記する。さらに、 $\sum_{i=1}^N (\alpha_i + \beta_i + \gamma_i)$ は一般の N 角形の内角の和であることから、その総和が $(N-2)\pi$ となることを用いると、 S は次式のように表せる。

$$\begin{aligned} S &= \sum_{i=1}^N (\alpha_i + \beta_i + \gamma_i) + 2 \sum_{i=1}^N \beta_i \\ &= (N-2)\pi + 2 \sum_{i=1}^N \beta_i \end{aligned} \quad (28)$$

内角の総和 S は $N\pi$ であるので、式 (28) より以下の式を導き出すことができる。

$$\sum_{i=1}^N \beta_i = \pi \quad (29)$$

これより、 N 角形の頂点にカメラを配置し、あるカメラが向かい合う 2 つのカメラを映す場合、 N 個のカメラの視野角の和はつねに π となることが分かる。以上より、カメラを正 N 角形に配置したとき、カメラの視野角は π/N となり、このとき N 台のカメラのうちで最も小さな視野角を持つカメラの視野角が最大となる。なぜなら、視野角の総和はつねに π であるため、あるカメラの視野角が π/N より小さくなった場合には、その他のカメラのうちで少なくとも 1 つの視野角が π/N 以上になるからである。

なお、これは相互投影の連結グラフを作るための必要条件であり十分条件ではないことに注意が必要である。

4. 実験

本章では 3 章で述べた理論を用いてシミュレーション実験によるカメラ位置の計算、形状復元の安定性評価、および実画像による形状復元を行い、本論文で提案した手法の有効性を示す。なお、本手法ではエピボール座標の取得を手動で行っている。手動で取得することによる誤差はせいぜい数ピクセルであり、従来法による計算でエピボールを求めた場合の誤差と比較すると非常に小さいものであるといえる。

4.1 相互投影によるカメラ視点の計算および物体形状の復元

本実験では図 4 のように 5 台のカメラ $C_i (i = 1, \dots, 5)$ が配置されており 5 台のカメラのほぼ中央部分に対象物が存在する。それぞれのカメラには相対する方向の 2 台のカメラと対象物が投影されている。

これら 5 台のカメラで得られた画像を図 5 (a) ~ (e) に示す。それぞれの画像中で黒丸は相対する 2 台のカメラの像を表し、白丸は提案法で用いる 4 つの基底点の像を表す。

まず、画像ノイズを印加した場合に提案法によるカメラ視点の計算および物体形状の復元を行い、相互投影を用いた場合と従来の 8 点法³⁾を用いた場合との比

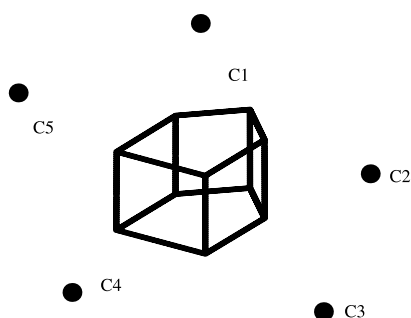


図 4 カメラと対象物の配置

Fig. 4 Experimental configuration of cameras and an object.

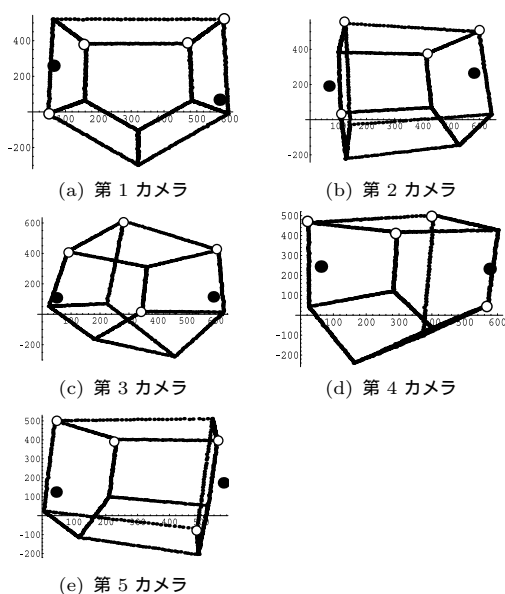


図 5 カメラ画像

Fig. 5 Camera images.

較を行った．なお，本論文では 8 点法においても相互投影法においても論文 3) に基づいて画像座標の正規化を行ったうえで F 行列計算を行った．図 5 の各投影点に標準偏差 1 pixel の画像ノイズを印加してそれぞれの手法でカメラ視点の計算を 100 回行った．このようにして得られたカメラ視点の計算結果を図 6 に示す．

カメラ視点，物体形状の復元結果には射影変換の不定性が残っているが，この結果に対して物体上の 5 点に正しい基底座標を与えてユークリッド復元に変換し，計算結果の安定性を評価した．このときに基準とした 5 つの射影基底点を図 6 に白丸で示す．図 6 中の楕円体はカメラ視点の復元結果に対する 3σ の不確定領域を表している．図 6 (a) は従来法を，また (b) は提案法を用いて計算した結果である．従来の 8 点法では

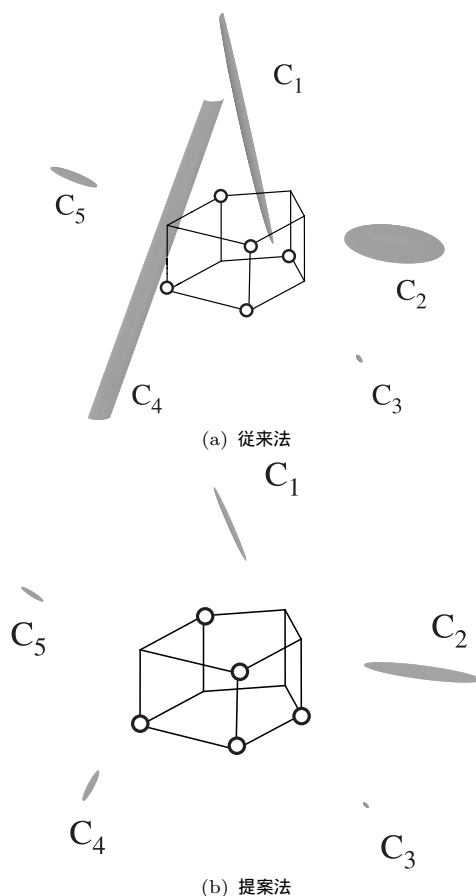


図 6 カメラ視点の計算結果

Fig. 6 Results of the computation of camera viewpoints.

表 1 カメラ視点の二乗誤差

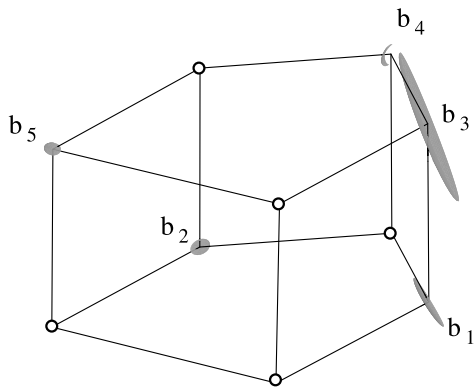
Table 1 Square errors of extracted viewpoints.

	C ₁	C ₂	C ₃	C ₄	C ₅
従来法	34.300	7.8146	0.0748	1759.7	0.8661
提案法	2.6490	3.9906	0.0261	0.7013	0.1983

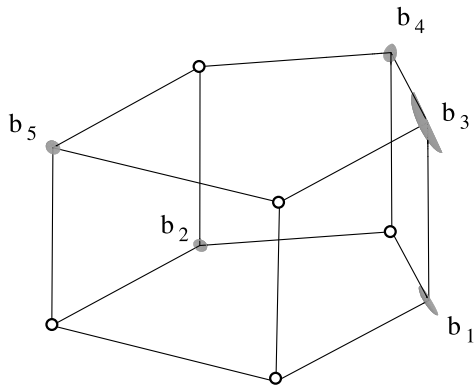
求めたカメラ視点の不確定領域が大きく，安定に復元することが困難であることが分かる．これに対して，提案法では非常に安定にカメラ視点が見出されていることが分かる．また，計算した各カメラ視点の正しい位置に対する二乗誤差を表 1 に示す．表 1 より，従来法に比べて提案法の誤差が全体的に非常に小さく，優れていることが確認できる．なお，図 6 に示す不確定領域のカメラ間でのばらつきに関しては，その理由を今後解析していく必要がある．

次に図 5 の各投影点に標準偏差 2 pixel の画像ノイズを印加してそれぞれの手法で物体形状の復元を 100 回行った．物体形状の復元結果を図 7 に示す．

先と同様に図中の白丸の 5 点に正しい射影基底座標



(a) 従来法



(b) 提案法

図 7 物体形状の復元結果

Fig. 7 Results of the reconstruction of objects.

表 2 物体形状の二乗誤差
Table 2 Square errors of object shapes.

	b_1	b_2	b_3	b_4	b_5
従来法	0.2906	0.0362	2.1527	0.0423	0.0684
提案法	0.0938	0.0235	0.3657	0.0339	0.0499

を与えて評価を行った．図中の楕円体は対象物上の各点の 3σ の不確定領域を表す．(a) は従来法による復元を，また (b) は提案法による復元を表す．対象物上に基底点をとっているためそれほど大きな差が見られないが，やはり従来法よりも提案法の方がより安定な結果が得られていることが分かる．また，復元した物体上の特徴点の正しい位置に対する二乗誤差を表 2 に示す．表 2 より，数値的にも従来法に比べて提案法の方が誤差が小さいことが確認できる．

4.2 実画像によるカメラ視点の計算および射影復元
次に，実画像を用いたカメラ視点の計算と射影復元について示す．本実験では図 8 に示すように 5 台のカメラを配置し，これらのカメラから得られる画像を用いて，従来法と提案法を用いてカメラ視点計算と射影復元を行い，それぞれの精度を比較した．図 9 に

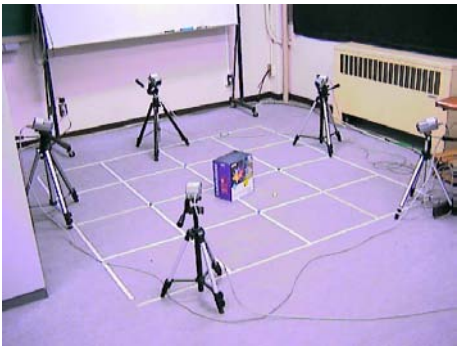
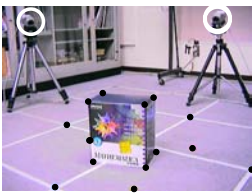


図 8 実画像実験におけるカメラ配置

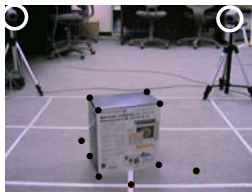
Fig. 8 Experimental configuration of cameras and an object.



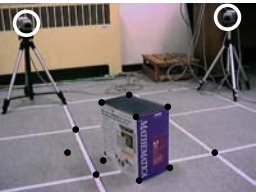
(a) 第 1 カメラ



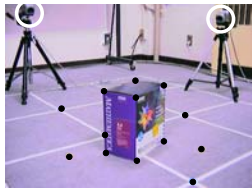
(b) 第 2 カメラ



(c) 第 3 カメラ



(d) 第 4 カメラ



(e) 第 5 カメラ

図 9 カメラ画像 (実画像)

Fig. 9 Camera images.

5 台のカメラより得られた画像を示す．これらの画像では，対象物（画像中央の箱）とともに相対する 2 台のカメラ（画像中，左上と右上）が相互投影されていることが分かる．また，図中の黒丸は対応点を表す．従来法では画像中の 8 組の対応点から，提案法では画像中の 3 組の対応点と 1 組のエピポールから fundamental 行列を計算し復元を行った．なお，提案法で用いたエピポールは画像中のカメラ像の中心を手動で指定することにより得た．このようにしてカメラ視点の計算を行った結果を図 10 に示す．図中の白丸はカメラ視点の復元結果を表し，黒丸は対応点を表す．

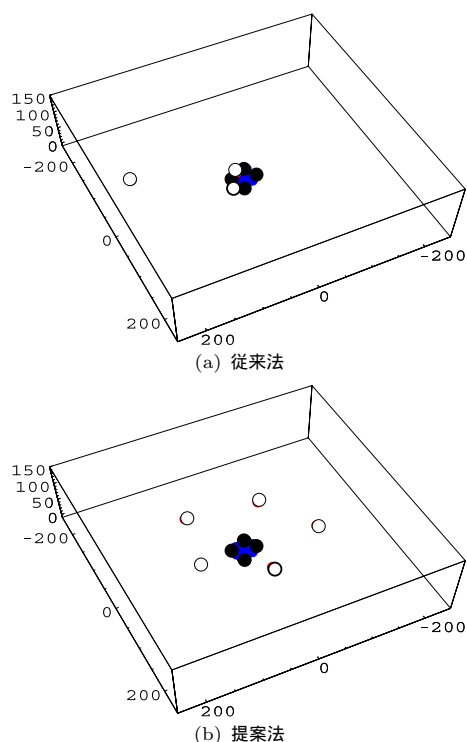


図 10 カメラ視点の計算結果 (実画像)

Fig. 10 Results of the computation of camera viewpoints.

図 10 より従来法ではカメラ視点の計算結果が非常に不安定であるのに対し、提案法ではほぼ実際のカメラ配置どおりに復元されていることが分かる。また、復元したカメラ視点の正しい位置に対する二乗誤差の平均は、従来法では 29653.9 cm^2 であるのに対し提案法では 341.63 cm^2 であり、提案法により大幅な精度向上が実現できていることが分かる。

次に対象物の形状復元を行った結果を図 11 に示す。図 11 (a) は従来法による復元結果、(b) は提案法による復元結果を表し、また (c) はほぼ同一視点から見た実際の対象物の形状を示す。なお、図 11 では白丸で示す点を射影基底とし、これらの点を基準にその他の点の復元安定性を表示している。図 11 より従来法の復元では形が大きくゆがみ、安定に形状復元できていないのに対し、提案法による復元では非常に安定に形状復元できていることが分かる。また、実際の物体形状と比較しても違和感なく復元できていることが分かる。また、復元した特徴点の正しい位置に対する二乗誤差の平均は従来法では 60.16 cm^2 であるのに対し、提案法では 1.14 cm^2 であり、やはり提案法の方が優れていることが分かる。



(a) 従来法



(b) 提案法



(c) 実画像

図 11 実画像における物体形状の復元結果

Fig. 11 Results of the reconstruction of objects.

5. ま と め

本論文ではカメラの位置や姿勢の情報が分からないマルチカメラを相互投影法を用いることで完全に校正し、3次元物体を復元する手法を提案した。この手法ではマルチカメラの相互投影より計算した2台のカメラ間のエピポーラ幾何と、それぞれのカメラ対での射影復元の結果より、すべてのカメラを射影的な不定性を除いて校正できることを示した。特に提案法ではすべてのカメラ間で最低4点が共通に投影されていればすべてのカメラを弱校正できることを示した。実際に提案法を用いてカメラの視点および物体の形状を復元

し、従来法と比較することにより提案法の有効性を示した。

今後の課題としては、1つの画像中に投影されるカメラの台数が増えた場合に対応して、trifocal tensor や quadrifocal tensor を用いた手法に拡張していくことが考えられる。また連結グラフが閉じている場合において、bundle adjustment を用いてより復元精度を向上させる方法への拡張も考えられる。

参 考 文 献

- 1) Longuet-Higgins, H.C.: A Computer Algorithm for Reconstructing a Scene from Two Projections, *Nature*, Vol.293, pp.133-135 (1981).
- 2) Luong, Q.T. and Vieville, T.: Canonic Representations for the Geometries of Multiple Projective Views, *Proc. 3rd European Conference on Computer Vision*, Vol.1, pp.589-599 (1994).
- 3) Hartley, R.I.: In Defense of the Eight-point Algorithm, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.19, No.6, pp.580-593 (1997).
- 4) Hartley, R.I. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press (2000).
- 5) 杉村健之, 佐藤 淳: カメラの相互投影による trifocal tensor の計算と形状復元の安定化, 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.43, No.SIG 11(CVIM 5), pp.113-120 (2002).
- 6) 佐藤 淳: コンピュータビジョン—視覚の幾何学, コロナ社 (1999).
- 7) 佐藤 淳: カメラの相互投影を用いたマルチカメラの弱校正, 計測自動制御学会システムインテグレーション部門講演会論文集 (III), pp.83-84 (2002).

- 8) 伊藤 満, 佐藤 淳: カメラの相互射影によるエピソード幾何計算と射影復元の安定化, *Trans. Institute of Electronics, Information and Communication Engineers*, Vol.J85-D-II, No.3, pp.416-424 (2002).

(平成 16 年 3 月 29 日受付)

(平成 16 年 9 月 11 日採録)

(担当編集委員 杉本 晃宏)



杉村 良彦

平成 15 年名古屋工業大学工学部電気情報学科卒業。現在同大学大学院博士前期課程在学中。コンピュータビジョン、物体形状復元の研究に従事。



佐藤 淳 (正会員)

昭和 59 年名古屋工業大学工学部卒業。平成 8 年ケンブリッジ大学大学院博士課程修了。同年ケンブリッジ大学工学部助手。平成 10 年名古屋工業大学工学部助教授。平成 16 年より同大学大学院情報工学専攻教授。この間に、ATR 人間情報通信研究所客員研究員等。コンピュータビジョン、視覚誘導、視覚的ユーザインタフェースの研究に従事。博士 (Ph.D.)。著書に『コンピュータビジョン—視覚の幾何学』(コロナ社)等。BMVC'94 最優秀科学論文賞, BMVC'97 最優秀科学論文賞, 第 5 回画像センシングシンポジウム論文賞等受賞。British Machine Vision Association 各会員。