社団法人 情報処理学会　研究報告
IPSJ SIG Technical Report

2005－CVIM－148　(3)
2005／3／3

エピポーラ幾何の進展

佐藤 淳[†]　杉本晃宏[‡]　木下敬介[††]

† 名古屋工業大学 情報工学専攻, 〒466-8555 名古屋市昭和区御器所町
‡ 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋
†‡ATR 人間情報科学研究所, 〒619-0288 京都府相楽郡精華町
junsato@nitech.ac.jp, sugimoto@nii.ac.jp, kino@atr.jp

あらまし：２画像間の対応付け問題として始まったエピポーラ幾何の研究は、その後複数カメラ間の幾何学的性質を解明する研究へと展開し、今や多視点画像間の幾何学的関係の全てを記述する大きな理論へと発展した。このような多視点幾何は、複数のカメラが存在する場合だけでなく、対称性を持つシーンの投影やミラー幾何、さらにはプロジェクタカメラシステムなど多くの視覚現象を記述することのできる非常に汎用性の高い理論である。さらに近年では、非剛体運動や非単焦点カメラなど、より一般的な対象とカメラを扱う多視点幾何へと拡張されつつある。本稿では、このように大きな発展を遂げた多視点幾何について、最近の研究動向をまとめる。

キーワード：エピポーラ幾何, 多視点幾何, multifocal tensor, multilinear 拘束

Recent Progress in Epipolar Geometry

Jun Sato[†]　Akihiro Sugimoto[‡]　Keisuke Kinoshita[†‡]

† Nagoya Institute of Technology,Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
‡ National Institute of Informatics, Chiyoda-ku, Tokyo 101-8430, Japan
†‡ ATR Human Information Science Laboratories, Soraku-gu, Kyoto 619-0288, Japan
junsato@nitech.ac.jp, sugimoto@nii.ac.jp, kino@atr.jp

Abstract: The research on the epipolar geometry was initiated for solving correspondence problems, and has been grown to the research on the general theory of multiple view geometry. The multiple view geometry can describe not only multiple camera views, but also the projection of symmetric scene, mirror geometry and projector-camera systems. Furthermore, it has recently been extended for describing more general visual phenomena, such as non-rigid motions and non-centric projections. In this paper, we investigate the recent progress in the research on multiple view geometry.

Keywords : epipolar geometry, multiple view geometry, multifocal tensor, multilinear constraints

1 はじめに

コンピュータビジョンでは、古くから複数のカメラ画像を用いることが考えられてきた。これは、カメラ画像というものがそもそも３次元の世界を投影して得られる２次元の情報であり、元の３次元の世界に対して１次元分の情報が不足していることに起因する。２次元の世界を対象とするパターン認識に対し、３次元の世界を対象とするコンピュータビジョンでは、複数のカメラ画像を用いることの意義は非常に大きい。

２台のカメラ間の幾何は古くよりエピポーラ幾何と呼ばれ、ステレオ視における対応点探索問題を中心に利用されて来た。その後、３眼視や多眼視の有効性が示され、さらにカメラの低コスト化や計算機の高速化により大量のカメラ画像をリアルタイムに処理できる可能性が広がったことから、より多くのカメラが用いられる時代となった。

このように多数のカメラの使用が一般化すると共に、それまで２台のカメラの幾何に限定されていたエピポーラ幾何を、３台以上のカメラを扱う幾何へと拡張する研究が９０年代半ばから急速に進展した。このようにして３台のカメラを扱う３視点幾何 (three-view geometry)、４台のカメラを扱う４視点幾何 (four-view geometry)、そして一般のN台のカメラを扱う N 視点幾何 (N-view geometry) が詳しく解析されその性質が明らかにされた。これらを総称して多視点幾何 (multiple view geometry) と呼ぶ。

多視点幾何は、複数の単焦点射影カメラの幾何学的性質を記述するのみでなく、単焦点射影カメラに変換可能な全方位カメラやプロジェクタなどの中心放射型投光器などに対して共通に適用することのできる非常に汎用性の高い幾何である。さらに近年では、非剛体運動に関する多視点幾何や非単焦点カメラに関する多視点幾何など、より一般的な対象とカメラとの組み合わせのもとでの多視点幾何が構築されつつある。本稿では、このように大きな進展を見せた多視点幾何に関する最近の研究動向をまとめる。

尚、本稿で述べる内容のうちの基礎論に関しては、次のような最近の解説書が特に参考になるであろう [9, 70, 2, 4, 28, 7, 40]。

2 多視点幾何の基礎

多視点画像に関する幾何は当初は幾何学的な考察のもとに研究が進められて来た。幾何学的な解析は直観的に理解しやすい半面、一般の多視点幾何には拡張しずらい。このようなことから多視点幾何の問題は次第に以下に示すように代数的に考えられるようになった。

今、空間中に N 台の射影カメラが存在し、空間中の点 $\mathbf{X}=[X^1,X^2,X^3,X^4]^\top$ がこれら N 台のカメラに投影されているとする。このとき、i 番目のカメラのカメラ行列を $\mathbf{P}_i$ とし、i 番目のカメラにおける $\mathbf{X}$ の投影像を $\mathbf{x}_i=[x_i^1,x_i^2,x_i^3]^\top$ とすると N 台のカメラへの投影は以下のように表すことができる。

$$\lambda_i \mathbf{x}_i = \mathbf{P}_i \mathbf{X} \quad (i=1,\cdots,N) \tag{1}$$

式 (1) において、形状に関するパラメータである $\mathbf{X}$ および λ_i とそれ以外とを分離して整理すると式 (1) は以下のように記述し直すことができる。

$$\begin{bmatrix} \mathbf{P}_1 & \mathbf{x}_1 & 0 & 0 & \cdots & 0 \\ \mathbf{P}_2 & 0 & \mathbf{x}_2 & 0 & \cdots & 0 \\ \mathbf{P}_3 & 0 & 0 & \mathbf{x}_3 & \cdots & 0 \\ \vdots & & & & & \vdots \\ \mathbf{P}_N & 0 & 0 & 0 & \cdots & \mathbf{x}_N \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ -\lambda_1 \\ -\lambda_2 \\ -\lambda_3 \\ \vdots \\ -\lambda_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2}$$

ここで式 (2) を以下のように記号で書き表すことにする。

$$\mathbf{M}\mathbf{Y} = \mathbf{0} \tag{3}$$

行列 $\mathbf{M}$ は $(3N)\times(N+4)$ 行列であるが、(3) がゼロベクトルでない解 $\mathbf{Y}$ を持つことから、行列 $\mathbf{M}$ の階数は常に以下の条件を満たす。

$$\operatorname{rank}\mathbf{M} < N+4 \tag{4}$$

すなわち、行列 $\mathbf{M}$ から $(N+4)\times(N+4)$ の部分行列 $\mathbf{M}'$ を取り出すと、次に示す通り、どの部分行列 $\mathbf{M}'$ もその行列式は 0 となる。

$$\det \mathbf{M}' = 0 \tag{5}$$

この時、式 (5) で表される拘束が N 視点幾何の拘束であり、幾何学的には、複数のカメラにおいて視

点と投影像とを結ぶ直線を伸ばした時に、これらの直線が３次元空間中において互いに１点で交わることを表している。これらは、２視点幾何の場合にはエピポーラ拘束 (epipolar constraints) あるいは bilinear 拘束 (bilinear constraints)、３視点幾何の場合には trilinear 拘束 (trilinear constraints)、４視点幾何の場合には quadrilinear 拘束 (quadrilinear constraints) と呼ばれ、またこれらを総称して multilinear 拘束 (multilinear constraints) と呼ぶ。

N 視点幾何において、行列 M の $3N$ 行から行列 M' の $N+4$ 行を選ぶ選び方は一般に一通りではないため複数の multilinear 拘束の式が得られる。しかし N 視点幾何に関する有効な拘束を得るためには、N 個のカメラ行列 $\mathbf{P}_i (i = 1, \cdot\cdot, N)$ のいずれからも２行以上ずつ選ばなければならない (あるカメラ行列から１行のみ選ぶとそのカメラの要素は単純にスケール項となってしまい意味をなさない)。すなわち、N 視点幾何において有効な $(N+4) \times (N+4)$ 行列 M' を得るためには、以下に示す条件を満たす必要がある。

$$2N \leq N + 4 \tag{6}$$

以上より、$N \geq 5$ では、有効な multilinear 拘束が得られないことがわかる。従って、multilinear 拘束は４視点幾何までしか存在しない。

一般に N 個の射影カメラが存在する場合、これらの画像が持つ自由度は $11N-15$ しか存在しない。これは、N 台の射影カメラがカメラ行列 $\mathbf{P}$ で表される通りそれぞれ 11 自由度を持つが、これら N 台のカメラが同一の射影空間 (15 自由度) 中に存在するという拘束条件があるためである。従って、２視点幾何は７自由度、３視点幾何は 18 自由度、４視点幾何は 29 自由度を持つことがわかる。

このような多視点幾何の代数的解析をもとに N 視点幾何が以下の節に示すように再構築された。代数的解析では、その幾何学的意味が見えずらいが、一方で、このような解析により多視点幾何が統一的に扱えるようになり、さらに多視点幾何の様々な代数的性質が明らかにされた。

3　２視点幾何

２台のカメラが存在するとき、これらのカメラのカメラ行列を $\mathbf{P}, \mathbf{P}'$ とし、ある３次元点 $\mathbf{X}$ のこれらのカメラにおける投影像を $\mathbf{x}, \mathbf{x}'$ とすると、先の議論によれば、式 (5) より以下に示す通り 6×6 行列の行列式が 0 となる。

$$\det \begin{bmatrix} P^1 & x^1 & 0 \\ P^2 & x^2 & 0 \\ P^3 & x^3 & 0 \\ P'^1 & 0 & x'^1 \\ P'^2 & 0 & x'^2 \\ P'^3 & 0 & x'^3 \end{bmatrix} = 0 \tag{7}$$

ここで、P^i はカメラ行列 $\mathbf{P}$ の第 i 行目を表し、x^i は $\mathbf{x}$ の第 i 要素を表す。式 (7) を展開して整理し直すと、以下に示すよく知られたエピポーラ方程式が得られる。

$$x^i x'^j \mathcal{F}_{ji} = 0 \tag{8}$$

但し、ここではエピポーラ方程式をテンソル表記を用いて表しており、$\mathcal{F}_{ji}$ はいわゆる $\mathbf{F}$ 行列の j 行 i 列の要素を表す。このような $\mathbf{F}$ 行列のテンソル表記 $\mathcal{F}_{ji}$ を bifocal tensor と呼ぶようになった。式 (7) を式 (8) と表したことから、$\mathcal{F}_{ji}$ と２つのカメラ行列, $\mathbf{P}, \mathbf{P}'$ との間には以下の関係があることがわかる。

$$\mathcal{F}_{ji} = \epsilon_{ipa}\epsilon_{jqb} \det \begin{bmatrix} P^p \\ P^a \\ P'^q \\ P'^b \end{bmatrix} \tag{9}$$

ここで、ϵ_{ijk} は $\{i, j, k\}$ から $\{1, 2, 3\}$ への置換が偶置換であれば 1、奇置換であれば -1、それ以外であれば 0 の値を取るテンソルである。bifocal tensor はカメラ行列が与えられている場合には式 (9) から、また画像中の対応点が得られている場合には式 (8) の拘束を用いて求めることができる（詳しい計算法に関しては９節参照)。

２視点幾何の自由度は７であることから、幾何学的な条件を満たす bifocal tensor の自由度は７であることがわかる。bifocal tensor は 3×3 で９個の要素を持つが、定数倍の不定性を持つことから、bifocal tensor の各要素間には $8-7=1$ 個の拘束が存在することがわかる。この拘束は以下に示す通り、bifocal tensor が full rank ではないということである。

$$\det \mathcal{F}_{ji} = 0 \tag{10}$$

4　3視点幾何

次に3台のカメラが存在する場合について考える。これら3台のカメラのカメラ行列をそれぞれ $\mathbf{P}, \mathbf{P}', \mathbf{P}''$ とし、またある3次元点 $\mathbf{X}$ のこれらのカメラへの投影像を $\mathbf{x}, \mathbf{x}', \mathbf{x}''$ とすると、式 (5) より以下の式が得られる。

$$\det \begin{bmatrix} P^1 & x^1 & 0 & 0 \\ P^2 & x^2 & 0 & 0 \\ P^3 & x^3 & 0 & 0 \\ P'^1 & 0 & x'^1 & 0 \\ P'^2 & 0 & x'^2 & 0 \\ P''^1 & 0 & 0 & x''^1 \\ P''^2 & 0 & 0 & x''^2 \end{bmatrix} = 0 \tag{11}$$

式 (11) では、$\mathbf{P}', \mathbf{P}''$ に関しては、第1, 第2行を取り出して用いているが、同様に第2, 第3行や、第3, 第1行を用いることもできるので、$\mathbf{P}'$ および $\mathbf{P}''$ からの行の取り方にはそれぞれ3通りの場合が存在する。従って式 (11) のような行列式の条件は合計9通り考えることができる。これらの式を展開して整理し直すと、以下に示す trilinear 拘束が得られる。

$$x^i x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr} = 0_{uv} \tag{12}$$

このとき、$\mathcal{T}_i^{jk}$ は3階のテンソル $(3 \times 3 \times 3)$ であり、trifocal tensor と呼ばれる。式 (11) と式 (12) より、trifocal tensor はカメラ行列を用いて以下のように表せることがわかる。

$$\mathcal{T}_i^{jk} = \epsilon_{ipa} \det \begin{bmatrix} P^p \\ P^a \\ P'^j \\ P''^k \end{bmatrix} \tag{13}$$

式 (13) より明らかなように、$\mathcal{T}_i^{jk}$ はカメラ行列のみからなるテンソルであり、射影空間中における3台のカメラ間の幾何学的な関係を完全に表現している。すなわち、3つのカメラのカメラ行列が与えられれば $\mathcal{T}_i^{jk}$ が計算でき、逆に $\mathcal{T}_i^{jk}$ が与えられれば3つのカメラのカメラ行列が射影的に一意に決まる。

Trifocal tensor はカメラ行列が与えられている場合には式 (13) から、また画像中の対応点が得られている場合には式 (12) の拘束を用いて求めることができる。(詳しい計算法に関しては9節参照)。式 (12) に示す trilinear 拘束には9つの式があるが、Hartley[18] はこの内で線形独立なものは4つのみであることを示した。この性質は trifocal tensor を対応点より求める場合に重要となる。

3台のカメラ間の幾何は、当初は2台のカメラのエピポーラ幾何を組み合わせて考えられた。太田ら [1] は2台のカメラ間のエピポーラ幾何を組み合わせることにより、3台のカメラ間での対応点拘束が直線ではなく一点にできることを示した。その後、しばらくの間は2台のカメラのエピポーラ幾何をベースに多眼視が研究された [11, 12]。3台のカメラ間ではエピポーラ幾何を複数用いる以上の拘束が得られることを初めて示したのは Spetsakis ら [55] であると思われる。Spetsakis らは校正済みカメラによる復元に trilinear 拘束の性質を初めて用いた。この trilinear 拘束は後に Shashua ら [51, 52] によって深く研究され trifocal tensor による3視点幾何の表現が示された。これに対し Hartley[18] は射影幾何をベースに3視点幾何を解析し直し、未校正な射影カメラに基づく trilinear 拘束を示した。さらに Hartley は、式 (13) において点 $\mathbf{x}$ とこの点を通る直線 $\mathbf{l}$ との間の関係を用いることにより、3視点における点と直線との対応関係や直線同士の対応関係が、以下に示すように一つの trifocal tensor によって記述できることを明らかにした。

$$x^i x'^j l''_r \epsilon_{jqu} \mathcal{T}_i^{qr} = 0_u \tag{14}$$

$$x^i l'_q l''_r \mathcal{T}_i^{qr} = 0 \tag{15}$$

$$l_p l'_q l''_r \epsilon^{pit} \mathcal{T}_i^{qr} = 0^t \tag{16}$$

ここで、l_i は直線 $\mathbf{l}$ の斉次座標の第 i 要素を表す。

trifocal tensor は $3 \times 3 \times 3$ で 27 個の要素を持つが、3視点幾何には 18 自由度しか存在しないことから、幾何学的条件を満たす trifocal tensor は 18 自由度しか持たない。すなわち、定数倍の不定性を除いて考えると、trifocal tensor の各要素間には $26 - 18 = 8$ 個の拘束が存在する。Faugeras ら [10] は、Grassman-Cayley 代数を用いることにより、これら8個の拘束を明らかにした。また、Canterakis[6] は異なる方法でやはりこれら 8 個の拘束を示している。

一方、3視点幾何は trifocal tensor で表す代わりに、3つの bifocal tensor を用いて表すこともできる。このような3視点幾何の表現に関しては6節において詳しく述べる。

5　4視点幾何

次に4台のカメラが存在する場合について考える。4台のカメラのカメラ行列をそれぞれ $\mathbf{P}$, $\mathbf{P}'$, $\mathbf{P}''$, $\mathbf{P}'''$ とし、ある3次元点 $\mathbf{X}$ のこれらのカメラへの投影像を $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}''$, $\mathbf{x}'''$ とすると、式 (5) よりこれらの間には以下の関係が得られる。

$$\det \begin{bmatrix} P^1 & x^1 & 0 & 0 & 0 \\ P^2 & x^2 & 0 & 0 & 0 \\ P'^1 & 0 & x'^1 & 0 & 0 \\ P'^2 & 0 & x'^2 & 0 & 0 \\ P''^1 & 0 & 0 & x''^1 & 0 \\ P''^2 & 0 & 0 & x''^2 & 0 \\ P'''^1 & 0 & 0 & 0 & x'''^1 \\ P'''^2 & 0 & 0 & 0 & x'''^2 \end{bmatrix} = 0 \qquad (17)$$

式 (17) では、それぞれのカメラ行列の第1, 第2行に関する式を取り出して用いているが、第2, 第3行、又は第3, 第1行を用いることもできるので、式 (17) に示すような拘束式は合計 $3^4 = 81$ 通り考えることができる。これらの式を展開して整理すると、以下にしめす quadrilinear 拘束が得られる。

$$x^i x'^j x''^k x'''^l \epsilon_{ipa}\epsilon_{jqb}\epsilon_{krc}\epsilon_{lsd} Q^{pqrs} = 0_{abcd} \qquad (18)$$

このとき、Q^{pqrs} は4階のテンソル $(3\times3\times3\times3)$ であり、quadrifocal tensor と呼ばれるようになった [66, 8]。式 (17) および式 (18) より、Quadrifocal tensor はカメラ行列を用いて以下のように表されることがわかる。

$$Q^{pqrs} = \det \begin{bmatrix} P^p \\ P'^q \\ P''^r \\ P'''^s \end{bmatrix} \qquad (19)$$

式 (19) より明らかなように quadrifocal tensor はカメラ行列のみからなるテンソルであり、4つのカメラ間の射影幾何学的な関係を完全に表現している。すなわち、4つのカメラのカメラ行列が与えられれば quadrifocal tensor が計算可能であり、逆に quadrifocal tensor が与えられれば、4つのカメラのカメラ行列を射影的に一意に決定することができる。式 (18) に示す quadrilinear 拘束には 81 個の式が有るが、この内で線形独立なものは 16 個のみである。さらに複数の対応点から式 (18) の拘束を得ると、これらの間には従属関係が生ずるため、N 個の対応点から得られる線形独立な式は $16N - {}_NC_2$ 個であることが示された [20]

3視点の場合と同様に、式 (18) において点 $\mathbf{x}$ とこの点を通る直線 $\mathbf{l}$ との間の関係を用いれば、以下に示すような点と直線に関する Quadrilinear 拘束や直線のみに関する quadrilinear 拘束を得ることができる。

$$\begin{aligned} x^i x'^j x''^k l'''_s \epsilon_{ipa}\epsilon_{jqb}\epsilon_{krc} Q^{pqrs} &= 0_{abc} && (20) \\ x^i x'^j l''_r l'''_s \epsilon_{ipa}\epsilon_{jqb} Q^{pqrs} &= 0_{ab} && (21) \\ x^i l'_q l''_r l'''_s \epsilon_{ipa} Q^{pqrs} &= 0_a && (22) \\ l_p l'_q l''_r l'''_s Q^{pqrs} &= 0 && (23) \end{aligned}$$

ただし、4直線に関する拘束 (23) は、画像中の4つの直線が3次元空間中の同一直線 L の投影像でなくとも、L 上の同一点 X が画像中の4直線上に投影されていさえすれば成り立つことに注意が必要である。

quadrifocal tensor は $3\times3\times3\times3$ であり 81 個の要素を持つが、4視点幾何には 29 自由度しか存在しないことから、幾何学的な条件を満たす quadrifocal tensor は 29 自由度しか持たない。従って、定数倍の不定性を除くと、quadrifocal tensor の各要素間には $80 - 29 = 51$ 個の拘束が存在する。Shashua ら [53] は homography tensor を導入することにより、これら 51 個の拘束を示した。

4視点幾何は quadrifocal tensor で表現する代わりに、2つの trifocal tensor や、5つの bifocal tensor で表現することもできる。これらに関しては6節において述べる。

6　多視点幾何の間の関係

以上の通り、2視点幾何、3視点幾何、4視点幾何はそれぞれ bifocal tensor, trifocal tensor, quadrifocal tensor によって表される。一方で3台のカメラ C_1, C_2, C_3 が存在する場合には、これらの関係を、C_1 と C_2 の間の2視点幾何、C_2 と C_3 の間の2視点幾何、C_3 と C_1 の間の2視点幾何、という具合に3つの2視点幾何に分けて考えることもできる。すなわち、3台のカメラ間の trifocal tensor と、これら3台を2台ずつ組みにした場合の bifocal tensor との間には、何らかの関係が存在する。Heyden[31] は、これらの multifocal tensor の各要素が、N 次元射影空間中の部分空間がなす Grassman 多様体の Grassman 座標 (Plücker 座標) となっていることに

着目し、Grassman 座標間に成り立つ Plücker の関係式を用いることにより、このようなテンソル間の関係を統一的に明らかにした。このようにして、以下に示すように C_1-C_2 間の bifocal tensor ${}_{12}\mathcal{F}_{ji}$ と C_1-C_3 間の bifocal tensor ${}_{13}\mathcal{F}_{ji}$ がこれら 3 視点間の trifocal tensor $\mathcal{T}_i^{jk}$ のそれぞれ左零空間と右零空間として得られることが明らかになった。

$$\mathcal{T}_i^{j1}\,{}_{12}\mathcal{F}_{1j} = 0_i,\ \mathcal{T}_i^{j2}\,{}_{12}\mathcal{F}_{2j} = 0_i,\ \mathcal{T}_i^{j3}\,{}_{12}\mathcal{F}_{3j} = 0_i \quad (24)$$

$$\mathcal{T}_i^{1k}\,{}_{13}\mathcal{F}_{1k} = 0_i,\ \mathcal{T}_i^{2k}\,{}_{13}\mathcal{F}_{2k} = 0_i,\ \mathcal{T}_i^{3k}\,{}_{13}\mathcal{F}_{3k} = 0_i \quad (25)$$

さらに文献 [31, 33] では、このような関係を用いて、trifocal tensor から 3 つの bifocal tensor を求める方法が示されている。

ここで、trifocal tensor の自由度は 18 であるのに対して、3 つの bifocal tensor が持つ自由度は $3 \times 7 = 21$ であることから、これら 3 つの bifocal tensor は独立ではなく、これらの間には $21 - 18 = 3$ 個の拘束が存在することがわかる。これら 3 つの拘束は以下のように表すことができる。

$$\mathbf{e}_{13}^i\mathbf{e}_{23}^j\,{}_{12}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{21}^i\mathbf{e}_{31}^j\,{}_{23}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{32}^i\mathbf{e}_{12}^j\,{}_{31}\mathcal{F}_{ji} = 0 \quad (26)$$

ここで、$\mathbf{e}_{pq}^i$ は第 q 視点を第 p 視点に投影して得られるエピポール $\mathbf{e}_{pq}$ の第 i 要素を表す。幾何学的にはこれらの式は、ある視点が残りの画像に投影された場合のエピポーラ拘束を表している (例えば第 1 式は、第 3 カメラを第 1 カメラと第 2 カメラに投影した像 $\mathbf{e}_{13}$, $\mathbf{e}_{23}$ が式 (8) のエピポーラ拘束を満たすことを表している)。従って、3 つの bifocal tensor が 3 視点幾何を正しく表現するためには、式 (26) に示す 3 つの拘束を満たさなければならない。一方、trifocal tensor から bifocal tensor を求める場合には、これら 3 つの拘束を満たす bifocal tensor が求まることになる。

同様に、trifocal tensor と quadrifocal tensor 間にも次に示す関係がある [31]。

$$\mathcal{Q}^{ijkl}\,{}_{123}\mathcal{T}_i^{jk}=0^{jkl},\ \mathcal{Q}^{ijkl}\,{}_{124}\mathcal{T}_i^{jl}=0^{jkl},\ \mathcal{Q}^{ijkl}\,{}_{134}\mathcal{T}_i^{kl}=0^{jkl}$$

$$\mathcal{Q}^{ijkl}\,{}_{213}\mathcal{T}_j^{ik}=0^{ikl},\ \mathcal{Q}^{ijkl}\,{}_{214}\mathcal{T}_j^{il}=0^{ikl},\ \mathcal{Q}^{ijkl}\,{}_{234}\mathcal{T}_j^{kl}=0^{ikl}$$

$$\mathcal{Q}^{ijkl}\,{}_{312}\mathcal{T}_k^{ij}=0^{ijl},\ \mathcal{Q}^{ijkl}\,{}_{314}\mathcal{T}_k^{il}=0^{ijl},\ \mathcal{Q}^{ijkl}\,{}_{324}\mathcal{T}_k^{jl}=0^{ijl}$$

$$\mathcal{Q}^{ijkl}\,{}_{412}\mathcal{T}_l^{ij}=0^{ijk},\ \mathcal{Q}^{ijkl}\,{}_{413}\mathcal{T}_l^{ik}=0^{ijk},\ \mathcal{Q}^{ijkl}\,{}_{423}\mathcal{T}_l^{jk}=0^{ijk}$$

ここで、${}_{pqr}\mathcal{T}_i^{jk}$ は $\{p, q, r\}$ の 3 視点間の trifocal tensor を表す。文献 [33] では、このような trifocal tensor と quadrifocal tensor との間の関係を用いて 4 視点間の quadrifocal tensor からこれら 4 視点中の全ての 3 視点間の trifocal tensor を計算する方法が示されている。

4 視点画像の幾何を表現するには、quadrifocal tensor 1 つで表現する代わりに、次に示すように、trifocal tensor 2 つで表現したり、bifocal tensor 5 つで表現することも可能である。今、4 台のカメラ C_i $(i = 1, \cdots 4)$ が存在するとすると、これら 4 視点の関係は C_1, C_2, C_3 間の trifocal tensor ${}_{123}\mathcal{T}_i^{jk}$ と C_2, C_3, C_4 間の trifocal tensor ${}_{234}\mathcal{T}_i^{jk}$ によって表すことができる。ここで自由度を考えてみると、2 つの trifocal tensor が持つ自由度は合計 $18 = 36$ であるのに対して 1 つの quadrifocal tensor が持つ自由度は 29 であることから、2 つの trifocal tensor による表現にはさらに 7 つの拘束が存在するはずである。これら 7 つの拘束は、${}_{123}\mathcal{T}_i^{jk}$ から得られる C_2, C_3 間の bifocal tensor ${}_{23}\mathcal{F}_{ji}$ と ${}_{234}\mathcal{T}_i^{jk}$ から得られる C_2, C_3 間の bifocal tensor ${}_{23}\mathcal{F}'_{ji}$ とが等しいことであると考えることができる (bifocal tensor の自由度は 7 であることに注意)。すなわち、2 つの trifocal tensor を用いて 4 視点幾何を表現する場合には、このような拘束を満たす trifocal tensor でなければならない。

一方でこれら 4 台のカメラ間の幾何は 5 つの bifocal tensor によって表現することもできる。例えば、${}_{12}\mathcal{F}_{ji}$, ${}_{13}\mathcal{F}_{ji}$, ${}_{23}\mathcal{F}_{ji}$, ${}_{14}\mathcal{F}_{ji}$, ${}_{24}\mathcal{F}_{ji}$ は 4 台のカメラ間の幾何を表す。ここで再び自由度を考えてみると、これら 5 つの bifocal tensor が持つ自由度は合計 $7 \times 5 = 35$ であり quadrifocal tensor の自由度は 29 であることから、これら 5 つの bifocal tensor 間には $35 - 29 = 6$ 個の拘束が存在するはずである。これら 6 つの拘束は以下のように表すことができる。

$$\mathbf{e}_{13}^i\mathbf{e}_{23}^j\,{}_{12}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{21}^i\mathbf{e}_{31}^j\,{}_{23}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{12}^i\mathbf{e}_{32}^j\,{}_{13}\mathcal{F}_{ji} = 0$$

$$\mathbf{e}_{14}^i\mathbf{e}_{24}^j\,{}_{12}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{12}^i\mathbf{e}_{42}^j\,{}_{14}\mathcal{F}_{ji} = 0 \quad \mathbf{e}_{21}^i\mathbf{e}_{41}^j\,{}_{24}\mathcal{F}_{ji} = 0 \quad (27)$$

これらは先と同様に、視点の投影像の対応関係を表すエピポーラ拘束に他ならない。従って、5 つの bifocal tensor によって 4 視点幾何を表現する場合には、これらの bifocal tensor が式 (27) の 6 つの拘束を満たさなければならない。

7 多視点幾何の縮退

3 次元空間中の点や視点が特殊な配置にある場合には、画像中の対応点から多視点幾何が一意に決

まらない状態、すなわち multifocal tensor が計算できない状態が生ずる。このような点と視点の配置を critical configuration と呼ぶ。critical configuration に近い状態にある場合には、計算した multifocal tensor が画像ノイズの影響を非常に大きく受けて計算が不安定になる。従って multifocal tensor の計算においては、このような critical configuration に近い状態が生じないよう十分に注意する必要がある。

2視点幾何における critical configuration は、Maybank[41] によって解析された。この結果、対応に用いる3次元点と2つの視点が全て一つの線織2次曲面上に存在する状態が critical configuration であることが明らかにされた。この一例は、全ての3次元点が平面上に存在する場合であり、このような場合には bifocal tensor は一意に求まらない (bifocal tensor を計算するためには平面上にない点が少なくとも2点必要である)。またこの状態に近い場合には、計算可能であるが結果は非常に不安定になる。

また、Maybank[42] はこの定理の双対として、6点と N 視点 ($N \geq 3$) とが線織2次曲面上にある場合には、やはり critical configuration となることを示した。

一方3視点幾何における critical configuration の一般形は Hartley[26] によって解析が行われた。彼は Maybank の結果を3視点に拡張し、critical configuration となる3つの場合を示した。

その後 N 視点幾何における critical configuration の一般形が Kahl ら [37, 36, 27] によって示された。Kahl らは一般の N 視点幾何における critical configuration が、3つの線織2次曲面の交叉の上に N 個の視点と全ての点が乗ることであることを明らかにし、さらにその幾つかの例を示した。

8 テンソルの最小表現

以上の通り、N 視点画像の幾何を表すテンソルは 3^N 個の要素よりなるが、これらのテンソルの自由度は $11N-15$ しか存在しない。すなわち、3^N 個の要素によるテンソル表現は、それぞれのテンソルの最小表現 (minimal parametrization) にはなっていない。このことは、非線形最適化を用いてこれらのテンソルを推定する場合に特に問題となる。すなわち、3^N 個のパラメータをそのまま最適化すると、得られる trifocal tensor は幾何学的な拘束を満たさない。そこで、これらのテンソルを最小のパラメータ数で表現する方法が研究された。

8.1 Bifocal Tensor の最小表現

Bifocal tensor は2自由度のエピポール2つと3自由度のエピポーラホモグラフィーによって表すことができる [4, 28, 7]。これは bifocal tensor の一つの最小表現であり、また幾何学的にも意味のある表現である。

一方、Bartoli ら [5] は bifocal tensor の特異値分解が2つの正規直交行列と rank 2 の対角行列に分解できることを用いて、2つの3自由度の正規直交行列と定数倍の不定性を除いた1自由度の対角行列によって最小表現する方法を示し、これを非線形最適化による bifocal tensor 計算のコスト低減に用いた。

8.2 Trifocal Tensor の最小表現

Trifocal tensor の最小表現は bifocal tensor に比べて遥かに難しくなる。このような最小表現の最初の試みはカメラ行列を用いて行われた。Hartley[18] は幾何学的に整合性の取れる trifocal tensor の表現法として、カメラ行列を用いた表現を示した。これは第1カメラを基準に3台のカメラのカメラ行列を $\mathbf{P} = [\mathbf{I}, \mathbf{0}]$, $\mathbf{P}' = A_i^j$, $\mathbf{P}'' = B_i^j$ と置き、trifocal tensor を $\mathcal{T}_i^{jk} = A_i^j B_4^k - A_4^j B_i^k$ と表現するものである。しかしこの表現には $11 \times 2 = 22$ 個の独立なパラメータが存在するため、最小表現とはなっていない。

これに対して、Torr ら [65] は3視点画像の幾何が3画像中の対応点6点で最小表現可能であることを示した。各画像において6点中の4点が射影画像平面上の基底点であると考えると、残りの2点はこれら4つの基底点で表現できる。すなわち、この従属な2点の座標が視点間の拘束を表していることになる。3つの画像中では合計 $2 \times 3 = 6$ 個の従属な画像点が存在するので、これらが trilinear 拘束を表すことになる。これら6点の斉次座標はそれぞれ3自由度であり合計18自由度を持つことから、画像中の対応点6点が trifocal tensor の最小表現を与えることがわかる。

また、Papadopoulo ら [46] は trifocal tensor の各要素間の拘束を用いることにより、やはり最小表現が可能であることを示している。Torr らの方法

では実際には3通りの表現が出て来るのに対して、Papadopoulo らの方法では一意に最小表現することが可能である。

8.3 Quadrifocal Tensor の最小表現

Heyden[30, 34, 32] は quadrifocal tensor をより少ないパラメータで表現する方法を示し、これを reduced quadrifocal tensor と呼んだ。これは、各画像中の対応点の座標を基準座標に変換すると、変換後の対応点座標に関する quadrifocal tensor の表現が簡単化されるというものである。実際にこのような変換を行うと、quadrifocal tensor の 81 個の要素の内で 0 でない要素は 36 個のみとなる。幾何学的条件を満たす quadrifocal tensor の自由度は 29 なので、定数倍の不定性を考えれば、reduced quadrifocal tensor の各要素間にはさらに $35-29=6$ 個の拘束が存在することがわかる。9 節に示すように、Heyden と Hartley はこのような表現を用いた quadrifocal tensor の計算法をそれぞれ示している。

また Shashua ら [53] は、homography tensor を導入することにより quadrifocal tensor を表現する方法を提案している。

9 Multifocal Tensor の計算法

画像から multifocal tensor を計算する方法としては、解析的に計算可能な線形解法と、繰り返し計算を必要とする非線形解法が考えられる。以下では、bifocal tensor, trifocal tensor, quadrifocal tensor それぞれに関してこれらの計算法を見て行く。

9.1 Bifocal Tensor の計算法

9.1.1 線形解法

bifocal tensor を最低 8 点の対応点から計算する 8 点法 (8 point algorithm)[39] は、bifocal tensor を線形に計算する方法として古くから知られている。これは式 (8) を展開整理して得られる以下の線形方程式を解く方法である。

$$\mathbf{Mf}=\mathbf{0} \tag{28}$$

ここで、$\mathbf{M}$ は N 個の対応点の座標よりなる $N\times 9$ 行列であり、また $\mathbf{f}=[\mathcal{F}_{11},\cdots,\mathcal{F}_{33}]^\top$ である。

このとき 8 点から計算された bifocal tensor はそのままでは式 (10) を満たさず幾何学的条件を満足していない。そこで、この条件を満たすよう bifocal tensor を線形計算後に修正する方法が提案された [67]。これは線形計算で求めた bifocal tensor を SVD で分解し、対角行列の rank を 2 としたのち再合成するものである。

8 点法は簡便な方法ではあるが、その計算不安定性は当初から問題となっていた。8 点法の不安定性の大きな要因の一つは、画像座標を斉次座標 $[x_1,x_2,x_3]^\top$ で考えた場合に、通常のカメラでは、x_1 と x_2 に対して x_3 が極端に小さな値となるため、行列計算時に計算不安定性を生む点にある。Hartley[19, 21] はこの点に着目し、画像座標の 3 つの成分がほぼ同じ大きさとなるよう座標値を正規化したのち bifocal tensor を計算する正規化 8 点法 (normalized 8 point algorithm) を提案した。この方法は bifocal tensor を簡便に求める方法として、また非線形解法の良い初期値を与える方法として現在では広く用いられている。

線形解法である 8 点法は上記の正規化を行うことにより、かなり安定性の改善を行うことができる。しかし 8 点法では、そもそも bifocal tensor が 7 自由度しか持たないのに、これを 8 自由度とみなして計算を行うため、条件によっては画像ノイズの影響を大きく受けて計算が不安定になる場合がある。一方、伊藤ら [35] はカメラ同士が相互に投影されているなどエピポールが既知の条件下では、線形解法により幾何学的条件を満たす bifocal tensor が直接求まることを示した。

また bifocal tensor は、式 (8) と式 (10) を組み合わせることにより、7 組の対応点から解析的に解が得られることが知られている [28]。この場合には、3 次方程式を解くことになるので、解が 1 つあるいは 3 個得られる。

9.1.2 非線形解法

非線形解法では何らかのコストを最小化するよう、最小パラメタライゼーションで表現した bifocal tensor や rank 2 を満たすパラメータを繰り返し計算する。この時用いるコストとしては、式 (28) にもとづく代数的距離 $||\mathbf{Mf}||$ を用いる方法と幾何学的距離を用いる方法が提案されている。

現在のところ、bifocal tensor 計算において最も良い精度が得られる方法は、計測した画像座標 $\mathbf{x}$, $\mathbf{x}'$ と推定した bifocal tensor から計算した画像座標 $\hat{\mathbf{x}}$, $\hat{\mathbf{x}}'$ との間の幾何学的距離 $\sum_i d(\mathbf{x}_i,\hat{\mathbf{x}}_i)^2+d(\mathbf{x}'_i,\hat{\mathbf{x}}'_i)^2$ をコストとして最適化を行う最尤推定であると考えられており、文献 [28] では Gold Standard 法として紹介されている。この方法では、$\hat{\mathbf{x}}_i$ と $\hat{\mathbf{x}}'_i$ を計

算するために、これらに対応する3次元空間中の点 $\mathbf{X}_i$ を同時に推定する必要があり、対応点が N 点存在する場合には $7+3N$ 個のパラメータを同時推定する必要がある。この方法では理論限界の精度が得られる半面、多くのパラメータの同時推定を行うので、計算コストがかかると同時に、局所解に陥らないよう十分に良い初期値を与える必要がある。

これに対して、より少ないパラメータの推定により、Gold Standard法に近い精度を得る方法も考案されている。これは、3次元座標 $\mathbf{X}_i$ とは無関係なコストを用いることにより実現できる。このようなコストとしては、一方の画像の対応点を推定したbifocal tensorを用いてもう一方の画像のエピポーラ線に変換したときの対応点とエピポーラ線との幾何学的距離が考えられる。文献 [64, 71] では $\sum_i \frac{(\mathbf{x}_i'^\top \mathbf{F}\mathbf{x}_i)^2}{(\mathbf{F}\mathbf{x}_i)_1^2+(\mathbf{F}\mathbf{x}_i)_2^2+(\mathbf{F}^\top\mathbf{x}_i')_1^2+(\mathbf{F}^\top\mathbf{x}_i')_2^2}$ や $\sum_i d(\mathbf{x}_i', \mathbf{F}\mathbf{x}_i)^2 + d(\mathbf{x}_i, \mathbf{F}^\top\mathbf{x}_i')^2$ をコストとして用いる方法が比較されている。Gold Standard法では多くのパラメータの推定が必要であるのに対して、これらエピポーラ幾何にもとづく幾何学的距離を用いる方法はbifocal tensorの7パラメータの推定のみで済むため計算コストが遥かに低い。

一方、Hartley[23] は、代数的距離を用いることにより、さらに推定するパラメータ数を削減できることを示した。この方法では、エピポールが与えられると、幾何学的条件を満たすbifocal tensorの残りの要素が線形に求まることを利用して、bifocal tensorの7パラメータを求める問題をエピポール1つを推定する問題に置き換えた。一般に代数的距離は幾何学的な距離を必ずしも反映していないため、代数的距離を最小化しても幾何学的に最適な解とはならない。しかし、代数的距離を用いると、bifocal tensorの7パラメータの推定問題を3パラメータの推定問題に置き換えることができ、また推定結果も幾何学的距離を最小化したものに対して遜色がないことから、有用な方法であることが報告されている。

一方、金谷ら [38] は、ノイズモデルを仮定した場合のbifocal tensor計算の理論限界精度を示し、さらにこの理論限界精度を達成するbifocal tensor計算法を提案した。

また、実際に2画像からbifocal tensorを計算する場合には、誤対応などにより座標値に大きな誤差を含む対応点を計算中に入れない工夫が必要である。このためにRANSAC[13] を用いて幾何学的な整合性が取れる点のみを用いてbifocal tensorの計算を行うロバスト法 [63, 64] などが提案されている。

9.2 Trifocal Tensorの計算法

9.2.1 線形解法

Trifocal tensorの計算法としても線形解法と非線形解法を考えることができる。Bifocal tensorの線形解法を拡張したtrifocal tensorの線形解法はHartley[18, 19, 22] によって提案された。これは、$\mathbf{t} = [\mathcal{T}_1^{11}, \cdots, \mathcal{T}_3^{33}]^\top$ と置いた場合に、式 (28) と同様に $\mathbf{Mt} = \mathbf{0}$ なる線形方程式を解くものである。ここで $\mathbf{M}$ は3視点中の点や直線の座標よりなる行列である。$\mathbf{t}$ に関する独立な拘束は、式 (12) からは4つ、式 (14) からは2つ、式 (15) からは1つ、式 (16) からは2つ得られることから、p 個の点-点-点対応と q 個の点-点-直線対応と r 個の点-直線-直線対応と s 個の直線-直線-直線対応が存在する場合には $4p+2q+r+2s \geq 26$ であればtrifocal tensorを線形に求めることができる。しかし線形解法では本来18自由度しか持たないtrifocal tensorを26自由度の問題として解くため、得られるtrifocal tensorは幾何学的な拘束を満たさない。これに対して杉村ら [3] は、bifocalの場合と同様にカメラの相互投影などによりエピポールが既知である場合には、線形解法により幾何学的条件を満たすtrifocal tensorが計算可能であることを示した。

9.2.2 非線形解法

Trifocal tensorを非線形最適化により計算する場合においても、最小化するコストにはいくつかの選択肢が考えられる。

最も良い結果が得られる方法はやはり計測した画像座標と推定した画像座標との間の幾何学的距離 $\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}_i', \hat{\mathbf{x}}_i')^2 + d(\mathbf{x}_i'', \hat{\mathbf{x}}_i'')^2$ をコストとする方法 (Gold Standard法) である [28, 7]。N 個の対応点が存在する場合、この方法では $18+3N$ 個のパラメータを同時推定することになる。ただし、trifocal tensorの場合には、bifocal tensorとは異なり、点対応だけでなく点と直線との対応や直線同士の対応を計算に用いることができるため、このような対応を用いる場合には、これらの幾何学的距離の定義が必要となる。特に直線同士の距離には明確な定義がないが、計測された直線の両端点から推定された直線への距離などが用いられている [7]。

一方、Hartley[25] は bifocal tensor の場合と同様に代数的距離をコストとする方法を提案した。この方法では、3視点画像中の6個のエピポールの内の2個を既知と仮定すると、幾何学的条件を満たす trifocal tensor が対応点から線形に求まる性質を利用して、trifocal tensor を推定する問題を2個のエピポールを推定する問題に置き換えている。この結果、6パラメータ (2個のエピポールの斉次座標) の最適化により trifocal tensor が求まり、また得られる結果も Gold Standard 法に対して遜色ないことが報告されている。

9.3 Quadrifocal Tensor の計算法

Quadrifocal tensor は 29 自由度の幾何を表すのに 81 個ものパラメータを使用しており、過度に冗長な表現となっていることから、これを直接計算する方法に関しては余り研究されて来なかった。

Heyden は quadrifocal tensor の要素の数を削減した表現を提案しこれを reduced quadrifocal tensor と呼んだ [30, 34]。さらに Heyden はこの reduced quadrifocal tensor 表現を用いて quadrifocal tensor を計算する方法を示した。

また、Hartley[24] はこの reduced quadrifocal tensor の表現を用い、さらに１２個のエピポールの内の3個が既知であると仮定するとこの表現の各要素が線形計算可能であることを用いて、代数的距離を最小化することで quadrifocal tensor を計算する非線形解法を提案した。この方法では、代数的距離をコストとしているが、それでもほぼ理論限界に近い結果が得られることを報告している。

一方、佐藤 [49] は、4視点画像における 12 個のエピポールと quadrifocal tensor の関係を解析し、N 個のエピポールが既知である場合の quadrifocal tensor の線形計算法とこの時の必要対応点数を明らかにした。

10 アフィンカメラの多視点幾何

アフィンカメラは射影カメラの特殊な場合であることから、射影カメラにおける多視点幾何の性質は全てアフィンカメラにおいても成り立つ。一方アフィンカメラは射影カメラに対して制限が加えられていることから、アフィンカメラの多視点幾何では射影カメラには無い新たな性質が生ずる。

アフィンカメラではカメラ行列 $\mathbf{P}$ の第3行が $[0,0,0,1]$ で表現される。従ってこの場合には1つのカメラ行列が持つ自由度は 8 である。このようなアフィンカメラが2台存在する場合、2つのカメラ行列は $2\times 8=16$ 自由度を持つが、これら2つのカメラには同一の3次元アフィン空間中に存在するという拘束が有る。3次元アフィン変換の自由度は 12 なので、アフィンカメラに関する2視点画像幾何には $2\times 8-12=4$ 自由度しかないことがわかる。すなわち、affine bifocal tensor (affine fundamental matrix) は 3×3 行列であるが、その自由度は 4 である。実際に第3行目が $[0,0,0,1]$ であるようなカメラ行列 $\mathbf{P},\mathbf{P}'$ を用いて式 (9) より $\mathcal{F}_{ji}$ を計算すると、$\mathcal{F}_{11}=\mathcal{F}_{12}=\mathcal{F}_{21}=\mathcal{F}_{22}=0$ となり、さらに定数倍の不定性が存在することから、$\mathcal{F}_{ji}$ の独立な要素の数は 4 個であることがわかる [72]。すなわち、affine bifocal tensor はそれ自身が最小表現となっており、各要素間には拘束が存在しない。

同様にアフィンカメラの3視点幾何では、trifocal tensor の自由度は $3\times 8-12=12$ であることがわかる [10, 43, 62]。affine trifocal tensor では 27 個の要素のうちの 11 個が 0 となることから、tensor の各要素間には $26-11-12=3$ 個の拘束が存在する [62]。

一方アフィンカメラの4視点幾何では、quadrifocal tensor の自由度は $4\times 8-12=20$ である。affine quadrifocal tensor では 81 個の要素のうちの 34 個が 0 であることから、tensor の各要素間には 26 個の拘束が存在する [61, 29]。

11 全方位カメラ，非単焦点カメラの多視点幾何

以上はアフィンカメラを含む射影カメラに基づく多視点幾何であるが、近年では全方位カメラの需要が増したことなどから、ミラーなどを用いた特殊な単焦点カメラに関する多視点幾何が研究されるようになった [59, 58, 48, 14, 15, 16]。

さらに最近ではカメラへの入射光が単一点で交差しない非単焦点カメラに関する多視点幾何が研究されつつある。当初は、特殊な非単焦点カメラに関する多視点幾何であったが [47, 54]、近年では以下に示すように、従来の単焦点カメラをも包含する非単焦点カメラの一般理論が明らかになりつつある。

Grossberg ら [17, 60] は、非単焦点カメラの各画素に入射する光線 (3次元空間中の直線) 同士が交わってできる曲面 (caustic 曲面) を非単焦点カメラ

の視点と考えることにより、非単焦点カメラをモデル化する方法を提案した。

Seitz[50] は、2台の非単焦点カメラの光線対が1点で交わるための条件を解析し、光線対が乗る曲面 (すなわちエピポーラ曲面) の種類が、平面、双曲面、双曲放物面のいずれかとなることを示した。また3台以上のカメラが存在する場合には、全ての2台のカメラ対が同一のエピポーラ曲面を持たなければならない。この条件から、Seitz は3台以上の非単焦点カメラが存在する場合にもエピポーラ曲面はやはり平面、双曲面、双曲放物面のいずれかしか取りえないことを示した。このような非単焦点カメラのエピポーラ曲面の性質は、同時期に Pajdla[44, 45] によっても示されている。

Sturm ら [56] は、Grossberg らと同様のカメラモデルをベースに、カメラに入射する光線を各画素毎に求める方法を示した。この方法では既知の座標を持つ対象物に未知の運動を与え、画像上の1点に対応する運動前後の3次元空間中の点が同一直線上に存在する拘束条件を用いて運動を求め光線を計算する。このとき、3次元運動前後の点が同一直線上に存在するという条件は、一種の multilinear 拘束であることを示した。

一方、Wexler ら [68] は、パラメトリックモデルでは表せない非単焦点カメラ間のエピポーラ幾何を、学習により獲得する方法を示した。2台の非単焦点カメラが存在する場合、これらのカメラ間の画像対応を表すエピポーラ線は曲線となるが、Wexler らは多くの点の対応から画像中での対応関係の存在確率を計算し、エピポーラ曲線を学習する方法を示した。

12 高次元空間における多視点幾何

以上の通り、点と直線に関する多視点幾何は近年大きな進展を見せたが、一方でこれらの画像特徴以外に関する多視点幾何の研究も進められている。杉本 [57] は、円錐曲線の5つの独立なパラメータが5次元射影空間中の点であると考えることにより、円錐曲線の投影を5次元射影変換として捉え、円錐曲線の投影像間に成り立つ bilinear 拘束と trilinear 拘束を示した。

一方、高次元空間における多視点幾何は、非剛体運動に関する多視点画像間の拘束を表し得ることも示されている。Wolf ら [69] は一般の N 次元空間から2次元画像への投影に関する多視点幾何の性質に関して解析を行い、各点が独立に等速直線運動を行う場合など、いくつかの非剛体運動パターンに関する幾何学的拘束がこのような投影のもとでの多視点幾何によって記述できることを示している。

13 まとめ

2画像間の対応付け問題として始まったエピポーラ幾何の研究は、射影幾何学的あるいは代数幾何学的アプローチの導入により90年代半ばから急速に進展し、今日までに多視点幾何に関する多くの性質を明らかにした。さらに最近では、剛体を対象とした単焦点カメラの理論に限られていたものが、非剛体や非単焦点カメラを扱う理論へとさらに拡張されつつある。このような研究の躍進ぶりからすると、近年完成の域に達した剛体単焦点カメラの多視点幾何理論が、古典的多視点幾何と呼ばれる日もそう遠くないかも知れない。

参考文献

[1] 大田, 池田. 3眼視ステレオについて. 情報処理学会第29回全国大会論文集, Vol. 2M-4, pages 1059–1060, 1984.

[2] 徐剛, 辻三郎. 3次元ビジョン. 共立出版, 1998.

[3] 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, 43(SIG11(CVIM5)):113–120, 2002.

[4] 佐藤淳. コンピュータビジョン—視覚の幾何学—. コロナ社, 1999.

[5] A. Bartoli and P. Sturm. Nonlinear estimation of the fundamental matrix with minimal parameters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(3):426–432, 2004.

[6] N. Canterakis. A minimal set of constraints for the trifocal tensor. In *Proc. 6th European Conference on Computer Vision*, pages 84–99, 2000.

[7] O.D. Faugeras and Q.T. Luong. *The Geometry of Multiple Images*. MIT Press, 2001.

[8] O.D. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences beterrn N images. In *Proc. 5th International Conference on Computer Vision*, pages 951–956, 1995.

[9] O.D. Faugeras and T. Papadopoulo. A theory of the motion fields of curves. *International Journal of Computer Vision*, 10(2):125–156, 1993.

[10] O.D. Faugeras and T. Papadopoulo. Grassmann-cayley algebra for modeling systems of cameras

and the algebraic equations of the manifold of trifocal tensors. *Philosophical Transaction of the Royal Society of London*, A-356:1123–1152, 1998.

[11] O.D. Faugeras and L. Robert. What can two images tell us about a third one? In *Technical Report, INRIA-2018*, 1993.

[12] O.D. Faugeras and L. Robert. What can two images tell us about a third one? In *Proc. 3rd European Conference on Computer Vision*, Vol. 1, pages 485–492, 1994.

[13] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Graphics and Image Processing*, 24(6):381–395, 1981.

[14] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 43:223–243, 2001.

[15] C. Geyer and K. Daniilidis. Properties of the catadioptric fundamental matrix. In *Proc. 7th European Conference on Computer Vision*, 2002.

[16] C. Geyer and K. Daniilidis. Mirrors in motion: Epipolar geometry and motion estimation. In *Proc. 9th International Conference on Computer Vision*, pages 766–773, 2003.

[17] M.D. Grossberg and S.K. Nayar. A general imaging model and a method for finding its parameters. In *Proc. 8th International Conference on Computer Vision*, Vol. 2, pages 108–115, 2001.

[18] R.I. Hartley. Lines and points in three views – an integrated approach. In *Proc. Image Understanding Workshop*, pages 1009–1016, 1994.

[19] R.I. Hartley. A linear method for reconstruction from lines and points. In *Proc. 5th International Conference on Computer Vision*, pages 882–887, Cambridge, Massachusetts, 1995.

[20] R.I. Hartley. Multilinear relationship between coordinates of corresponding image points and lines. In *Proc. International Workshop on Computer Vision and Applied Geometry*, 1995.

[21] R.I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(6):580–593, 1997.

[22] R.I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125–140, 1997.

[23] R.I. Hartley. Minimising algebraic distance. In *Proc. Image Understanding Workshop*, pages 631–637, 1997.

[24] R.I. Hartley. Computation of the quadrifocal tensor. In *Proc. 5th European Conference on Computer Vision*, pages 20–35, 1998.

[25] R.I. Hartley. Minimizing algebraic error in geometric estimation problems. In *Proc. 6th International Conference on Computer Vision*, pages 469–476, 1998.

[26] R.I. Hartley. Ambiguous configurations for 3-view projective reconstruction. In *Proc. 6th European Conference on Computer Vision*, pages 922–935, 2000.

[27] R.I. Hartley and F. Kahl. A critical configuration for reconstruction from rectilinear motion. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 511–517, 2003.

[28] R.I. Hartley and A. Zisserman. *Multiple View Geometry*. Cambridge University Press, 2000.

[29] E. Hayman, T. Thorhallsson, and D.W. Murray. Zoon-invariant tracking using points and lines in affine views an application of the affine multifocal tensors. In *Proc. 7th International Conference on Computer Vision*, pages 269–276, 1999.

[30] A. Heyden. *Geometry and algebra of Multiple Projective Transformations*. PhD thesis, Lund University, 1995.

[31] A. Heyden. A common framework for multiple view tensors. In *Proc. 5th European Conference on Computer Vision*, Vol. 1, pages 3–19, 1998.

[32] A. Heyden. Reduced multilinear constraints: theory and experiments. *International Journal of Computer Vision*, 30(1):5–26, 1998.

[33] A. Heyden. Tensorial properties of multiple view constraints. *Mathematical Methods in the Applied Sciences*, 23:169–202, 2000.

[34] A. Heyden and K. Astrom. Simplifications of multilinear forms for sequences of images. *Image and Vision Computing*, 15:749–757, 1997.

[35] M. Ito, T. Sugimura, and J Sato. Recovering structures and motions from mutual projection of cameras. In *Proc. International Conference on Pattern Recognition*, Vol. 3, pages 676–679, 2002.

[36] F. Kahl. *Geometry and Critical Configurations of Multiple Views*. PhD thesis, Lund University, 2001.

[37] F. Kahl, R. Hartley, and K. Astrom. Critical configurations for n-view projective reconstruction. In *Proc. Conference on Computer Vision and Pattern Recognition*, 2001.

[38] K. Kanatani. Optimal fundamental matrix computation: algorithm and reliability analysis. In

第6回画像センシングシンポジウム予稿集, pages 291–296, 2000.

[39] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.

[40] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry. *An invitation to 3-D vision: from images to geometric models*. Springer, 2004.

[41] S.J. Maybank. *Theory of Reconstruction from Image Motion*. Springer-Verlag, 1993.

[42] S.J. Maybank and A. Shashua. Ambiguity in reconstruction from images of six points. In *Proc. 6th International Conference on Computer Vision*, pages 703–708, 1998.

[43] P.R.S. Mendonca and R. Cipolla. Analysis and computation of an affine trifocal tensor. In *British Machine Vision Conference*, pages 125–133, 1998.

[44] T. Pajdla. Epipolar geometry of some non-classical cameras. In *Proc. Computer Vision Winter Workshop*, pages 223–233, 2001.

[45] T. Pajdla. Stereo with oblique cameras. *International Journal of Computer Vision*, 47:161–170, 2002.

[46] T. Papadopoulo and O. Faugeras. A new characterization of the trifocal tensor. In *Proc. 5th European Conference on Computer Vision*, pages 109–123, 1998.

[47] S. Peleg and M. Ben-Ezra. Stereo panorama with a single camera. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 338–343, 1997.

[48] R. Pless. Discrete and differential two-view constraints for general imaging systems. In *Proc. 3rd Workshop on Omnidirectional Vision*, pages 53–59, 2002.

[49] J. Sato. Recovering multiple view relations from mutual projections of multiple cameras. In *Proc. Asian Conference on Computer Vision*, Vol. 1, pages 7–12, 2004.

[50] S.M. Seitz. The space of all stereo images. In *Proc. 8th International Conference on Computer Vision*, Vol. 1, pages 26–33, 2001.

[51] A. Shashua. Trilinearity in visual recognition by alignment. In *Proc. 4th European Conference on Computer Vision*, Vol. 1, pages 479–484, 1994.

[52] A. Shashua and M. Werman. Trilinearity of three perspective views and its associated tensor. In *Proc. 5th International Conference on Computer Vision*, pages 920–925, 1995.

[53] A. Shashua and L. Wolf. On the structure and properties of the quadrifocal tensor. In *Proc. 6th European Conference on Computer Vision*, Vol. 1, pages 710–724, 2000.

[54] H. Shum, A. Kalai, and S.K. Seitz. Omnivergent stereo. In *Proc. 7th International Conference on Computer Vision*, Vol. 1, pages 22–29, 1999.

[55] M. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4:171–183, 1990.

[56] P. Sturm and S. Ramalingam. A generic concept for camera calibration. In *Proc. 8th European Conference on Computer Vision*, 2004.

[57] A. Sugimoto. Multilinear relationship between the coordinates of corresponding image conics. In *Proc. International Conference on Pattern Recognition*, Vol. 1, pages 550–554, 2000.

[58] T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49(1):23–37, 2002.

[59] T. Svoboda, T. Pajdla, and V. Hlavac. Epipolar geometry of panoramic cameras. In *Proc. 5th European Conference on Computer Vision*, pages 218–231, 1998.

[60] R. Swaminathan, M.D. Grossberg, and S.K. Nayar. Caustics of catadioptric cameras. In *Proc. 8th International Conference on Computer Vision*, Vol. 2, pages 2–9, 2001.

[61] T. Thorhallsson. *Object Symmetry in Multiple Affine Views*. PhD thesis, University of Oxford, 1999.

[62] T. Thorhallsson and D.W Murray. The tensors of three affine views. In *Proc. Conference on Computer Vision and Pattern Recognition*, Vol. 1, pages 450–456, 1999.

[63] P. Torr. Motion segmentation and outlier detection. PhD thesis, University of Oxford, 1995.

[64] P. Torr and D.W. Murray. The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24(3):271–300, 1997.

[65] P. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*, 15:591–605, 1997.

[66] B. Triggs. Matching constraints and the joint image. In *Proc. 5th International Conference on Computer Vision*, pages 338–343, 1995.

[67] R.Y. Tsai and T.S. Huang. Uniqueness and estimation of three-dimensional motion parameters of a rigid objects with curved surfaces. *IEEE*

Trans. on Pattern Analysis and Machine Intelligence, 6(1):13–26, 1984.

[68] Y. Wexler, A.W. Fitzgibbon, and A. Zisserman. Learning epipolar geometry from image sequences. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 209–216, 2003.

[69] L. Wolf and A. Shashua. On projection matrices $P^k \rightarrow P^2, k = 3, \cdots, 6$, and their applications in computer vision. In *Proc. 8th International Conference on Computer Vision*, Vol. 1, pages 412–419, 2001.

[70] G. Xu and Z. Zhang. *Epipolar geometry in stereo, motion and object recognition - A unified approach.* Kluwer Academic Publishers, 1996.

[71] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161–195, 1998.

[72] A. Zisserman. Notes on geometric invariance in vision. Lecture note (BMVC92 tutorial), 1992.