

カメラの幾何学と再構成問題

Camera geometry and reconstruction problems

東京大学大学院数理科学研究科 三浦 真人

Makoto Miura

Graduate School of Mathematical Sciences, The University of Tokyo

1 序文

コンピュータビジョンにおける3次元形状の再構成問題とは、カメラによって撮影された写真から3次元空間の中にある被写体の形状を復元する手続きにまつわる問題のことである。現在、再構成問題に関連するアルゴリズムは広い応用を持つ技術となっており、例を挙げれば、現実の物体に対する3Dモデルを作ることや、実写の映像の中にCGを合成することなどに用いられている。本稿では、再構成問題の幾何学的な構造について議論する。そのための道具として、代数幾何、とくにトリーク幾何の言葉を用いるが、これらについての知識は前提としない。一般に、コンピュータビジョンにおけるさまざまな問題に代数幾何的な手法を用いて答え、理論的な深化をもたらすことを目指す代数ビジョンという分野がある。本稿は、この分野への入門的な解説を行い、これに関連した伊藤敦氏、植田一石氏と筆者の共同研究 [6] の結果を非専門家に向けて紹介する。

2 ピンホールカメラ模型

ピンホールカメラとは、レンズの代わりに針穴（ピンホール）を用いる原始的なカメラの仕組みをいう。被写体から放射・散乱された光線のうち針穴を通過したものだけが背面のスクリーンに衝突することで、スクリーン上に逆さまの像が映し出される。この単純な光学現象は紀元前から知られており、ルネサンス期の西洋画家たちはこの現象を応用したさまざまな装置を写実画の制作に役立てていたようだ。まずは、このピンホールカメラの数理模型をつくることから始めよう。先に結論を述べれば、コンピュータビジョンで扱うピンホールカメラ模型とは、射影空間における射影に他ならない。

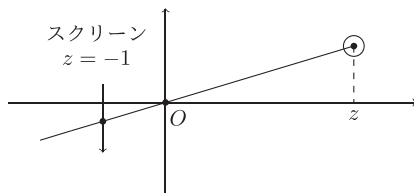


図 2.1 ピンホールカメラ. 一つ次元を落として描いたピンホールカメラの概略図。針穴 O を通過した光線のみが後ろのスクリーンに衝突し像を結ぶ。たとえば、 O を原点とした座標でスクリーンが $z = -1$ に置かれているとき、奥行き z にある被写体は $1/z$ 倍の大きさに映って見える。

はじめに、有理関数を用いて

$$\varphi^\circ : \mathbf{k}^3 \dashrightarrow \mathbf{k}^2, \quad (x, y, z)^T \mapsto \left(\frac{x}{z}, \frac{y}{z} \right)^T \quad (2.1)$$

と記述する模型を考えるのが自然であろう。本稿を通じて $\mathbf{k} = \mathbb{R}$ または \mathbb{C} とする。 z は奥行きを表して、スクリーンは図 2.1 のように $z = -1$ に置かれている。(2.1) で破線矢印を用いたのは、写像の定義域が実際には \mathbf{k}^3 全体ではなく、 \mathbf{k}^3 のザリスキー開集合^{*1}であることを明示するためであり、代数幾何における有理写像の標準的な記法である。今の場合、 φ° は $z \neq 0$ で定義されている。この φ° では写らないスクリーンと平行な平面 $z = 0$ はコンピュータビジョンではカメラの主平面 (principal plane) と呼ばれている。さて、この模型ではあくまでも点の行き先だけを見ていて、ピクセルの明るさや色 (すなわち光線の物理的特性) についての情報は無視していることに注意しよう^{*2}。一方、点の行き先について見てみると、(2.1) は現実のカメラよりも多くの点を写している。実際、スクリーンの後ろ側も写し、物陰に隠れた点も透過して写る。この事情が、再構成問題においてはカイラリティ (cheirality) という手がかりとして用いられ、CG シミュレーションにおいては隠線処理 (hidden-line removal) が必要なることに反映される。どちらの場合でも、まずは (2.1) が基本にあり、本来写らないはずの点の処理は追加で行われることになる。

次に、(2.1) が射影空間における射影に自然に延長されることを見る。たとえば、 $y = 1$ と表される地面のように、スクリーンと平行でなく原点も通らない \mathbf{k}^3 内の平面をピンホールカメラで写すことを考えてみる。すると、写真 \mathbf{k}^2 にはこの平面のどの点にも対応しない消失点の集合 (地平線) が浮かび上がる。これを \mathbf{k}^3 の「無限遠」の一部が映っていると捉え、それを含むように (2.1) を延長したものが、次の有理写像である。

$$\varphi : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2, \quad [x : y : z : w] \mapsto [x : y : z]. \quad (2.2)$$

ここで射影空間 $\mathbb{P}^n = \mathbb{P}_{\mathbf{k}}^n = \mathbb{P}(\mathbf{k}^{n+1})$ とは \mathbf{k}^{n+1} の 1 次元部分空間からなる多様体である。 \mathbb{P}^n 上の点を \mathbf{k}^{n+1} のゼロでないベクトル $\mathbf{x} = (x_0, \dots, x_n)^T$ を用いた斉次座標で表す：

$$[\mathbf{x}] = [x_0 : \dots : x_n] = [\lambda \mathbf{x}] \quad (\lambda \neq 0). \quad (2.3)$$

射影空間の間の有理写像 (rational map) とは k 次同次多項式 $P_0(\mathbf{x}), \dots, P_m(\mathbf{x})$ を用いて、

$$\mathbb{P}^n \dashrightarrow \mathbb{P}^m, \quad [\mathbf{x}] \mapsto [P_0(\mathbf{x}) : \dots : P_m(\mathbf{x})] \quad (2.4)$$

と記述される \mathbb{P}^n のザリスキー開集合で定義された写像のことをいう。正確には、低次元の代数的集合を足し引きするような定義域の些細な違いは無視するように同値類をとる^{*3}。実際に (2.2) で導入した有理写像 φ が (2.1) の φ° の延長になっていることを見るには、アフィン空間 \mathbf{k}^n を \mathbb{P}^n において $x_n \neq 0$ で定義されるザリスキー開集合と見ればよい：

$$\mathbf{k}^3 \subset \mathbb{P}^3 \dashrightarrow \mathbb{P}^2 \supset \mathbf{k}^2, \quad (x, y, z)^T \mapsto [x : y : z : 1] \mapsto [x : y : z] \mapsto \left(\frac{x}{z}, \frac{y}{z} \right)^T. \quad (2.5)$$

こうして見ると、有理写像 φ では \mathbf{k}^3 の「無限遠」(すなわち $w = 0$ の点) が写るようになったばかりでなく、先ほどは写らなかった主平面 $z = 0$ 上の点もスクリーン側の「無限遠」に写るようになったことが分かる。こ

^{*1} いくつかの多項式の共通零点で定まるような集合 (代数的集合) を閉集合と定めた位相をザリスキー位相という。射影空間やその部分多様体のザリスキー位相についても同様。

^{*2} たとえば、明るさや色の情報を用いる CT 画像再構成法なども広くは 3 次元形状の再構成問題に含まれる。しかし、ピンホールカメラ模型を扱う場合は、被写体が (たとえばその特徴点からつくられる) 点の集合 (point cloud) と見なせる状況を想定している。

^{*3} つまり空でないザリスキー開集合で一致するような有理写像は同値と見なす。これは、射影空間の間の有理写像を制限して閉部分多様体の間の有理写像を定義するときも同様である。今の場合は単に多項式 $P_0(\mathbf{x}), \dots, P_m(\mathbf{x})$ が共通因子を持たないと約束してもよい。

のように、射影空間 \mathbb{P}^3 では平面（スクリーン）と直線（光線）が必ず交わるため、例外的な状況を意識することなく幾何学を議論できるメリットがある。ただし、 φ もカメラの中心 $O = [0 : 0 : 0 : 1]$ では定義されていない。カメラはカメラ自身を写すことができないのだ。さて、一般に線形空間の間の全射線形写像 $s : V \rightarrow W$ ($\dim V > \dim W$) に対して定まる射影空間の間の有理写像

$$\bar{s} : \mathbb{P}(V) \dashrightarrow \mathbb{P}(W), \quad [\mathbf{x}] \mapsto [s(\mathbf{x})] \quad (2.6)$$

を射影空間における射影（*projection*）といい、 \bar{s} が定義されない点の集合 $\mathbb{P}(\ker s) \subset \mathbb{P}(V)$ を射影の中心（*center*）という。もちろん線形写像 s と λs ($\lambda \neq 0$) は同じ射影 \bar{s} を定める。この定義から、ピンホールカメラを記述する (2.2) の有理写像 φ も、行列

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2.7)$$

の定める射影となっている。こうして、ピンホールカメラ模型が射影空間における射影に他ならないことが分かった。行列 P 、あるいはそのスカラー倍 λP ($\lambda \neq 0$) を、ピンホールカメラ模型 φ を記述するカメラ行列（*camera matrix*）という*4。

3 再構成問題の幾何学

1872 年、クライン（Felix Klein）は幾何学の発展に大きな影響を与えることになるエルランゲン・プログラム [9] を発表した。その中で強調された幾何学についての考え方は次のとおりである：

空間 X とその上に働く（ X を X 自身の上に移す）変換群 G との組が与えられたとき、 X 内の図形の性質のうち、 G のどの変換によっても変わらないもの、すなわち G に関する不変量を研究することが、 (X, G) についての幾何学である。

このように捉えることで、複数の幾何学が変換群同士の関係によって結びつく。典型的な例は、カメラの幾何学にも登場する、射影幾何、アフィン幾何、ユークリッド幾何という三つの幾何学の関係である。3 次元形状（あるいは一般に n 次元形状）の再構成問題は、射影幾何、アフィン幾何、ユークリッド幾何という幾何学の階層的な関係を反映し、それぞれ射影再構成、アフィン再構成、計量再構成と呼ばれるステップごとの再構成問題へと分けられることになる。

射影幾何、アフィン幾何、ユークリッド幾何という三つの幾何学は、射影空間の中に特定の「基本図形」を付加していくことにより、つくることができる。具体的には、射影空間 \mathbb{P}^n に何も付加しない射影幾何から始めて、超平面 H を付加することでアフィン幾何を、超平面 H 上の 2 次超曲面 Q を付加することでユークリッド幾何をつくる。以下では表 3.1 を基に、それぞれの幾何学について見ていこう。

射影幾何（*projective geometry*）とは、射影空間 \mathbb{P}^n とその自己同型群 $\text{Aut } \mathbb{P}^n$ の組についての幾何学である。一般に、体 \mathbf{k} 上の射影空間 \mathbb{P}^n の自己同型群は射影線形群 $PGL(n+1, \mathbf{k})$ と同型である：

$$\text{Aut } \mathbb{P}^n \simeq PGL(n+1, \mathbf{k}) = \{\bar{s} : \mathbb{P}^n \rightarrow \mathbb{P}^n \mid s \in GL(n+1, \mathbf{k})\}. \quad (3.1)$$

*4 一般に、射影空間における射影 $\mathbb{P}^n \dashrightarrow \mathbb{P}^m$ をピンホールカメラ模型のアナロジーとして捉える。「写真」 \mathbb{P}^m の上の射影変換は正則行列の QR 分解（あるいは射影変換群の岩澤分解）により、カメラの内部パラメータを表す上三角行列 K （ $m=2$ の場合は、焦点距離、アスペクト比、スキュー比、主点の位置という 5 つの自由度がある）とカメラの向きを表す直交行列 R に分解される。射影変換と相殺しない残りの自由度は \mathbb{P}^n における「カメラの中心」の配置の自由度（すなわちグラスマン多様体 $Gr(n-m, n+1)$ の次元）に対応する。

表 3.1 幾何学の階層性. 基礎的な空間情報は変換群の選び方に応じて階層をなす。

	空間 X	変換群 G	「基本図形」	不変量 (例)
射影幾何	\mathbb{P}^n	$\text{Aut } \mathbb{P}^n$	なし	交わり方、接し方、線形性
アフィン幾何	\mathbf{k}^n	$\text{Aff } \mathbf{k}^n$	超平面 H	平行性、体積の比、重心
ユークリッド幾何	\mathbf{k}^n	$\text{Sim } \mathbf{k}^n$	H 上の 2 次超曲面 Q	長さの比、角度
		$\text{Euc } \mathbf{k}^n$	(長さの単位)	長さ、体積

射影空間 \mathbb{P}^n 上の各変換 σ を射影変換といい、 $\text{Aut } \mathbb{P}^n$ は射影変換群とも呼ばれる。射影空間に含まれる図形が持つ性質のうち、どの射影変換でも変わらないものとしては、図形同士の交わり方、接し方、部分多様体の次数 (とくに線形性) などがある。これらはカメラの幾何学において最も基本的な不変量である*5。

アフィン幾何 (affine geometry) とは、 n 次元空間 \mathbf{k}^n とその上に働くアフィン変換群 $\text{Aff } \mathbf{k}^n$ の組についての幾何学である。アフィン変換とは、 \mathbf{k}^n の線形変換と平行移動の組み合わせで書かれる次のような変換のことである：

$$\mathbf{k}^n \rightarrow \mathbf{k}^n, \quad \mathbf{x} \mapsto A\mathbf{x} + \mathbf{t}, \quad (3.2)$$

ここで $A \in GL(n, \mathbf{k})$, $\mathbf{t} \in \mathbf{k}^n$ とした。 n 次元空間 \mathbf{k}^n 内の図形が持つ性質のうち、どのアフィン変換でも変わらないものとしては、平行性、体積の比、重心などがある。 \mathbf{k}^n のアフィン幾何は超平面 H を付加した射影空間 (\mathbb{P}^n, H) に対する射影幾何と等価になる。これは、たとえば具体的に $x_n = 0$ で定まる超平面 H をとって条件を書いてみると明らかである：

$$\text{Aut}(\mathbb{P}^n, H) := \{g \in \text{Aut } \mathbb{P}^n \mid g \cdot H \subset H\} \quad (3.3)$$

$$= \left\{ \begin{pmatrix} A & \mathbf{t} \\ O & 1 \end{pmatrix} : \mathbb{P}^n \rightarrow \mathbb{P}^n \mid A \in GL(n, \mathbf{k}), \mathbf{t} \in \mathbf{k}^n \right\} \quad (3.4)$$

$$\simeq \text{Aff } \mathbf{k}^n \simeq GL(n, \mathbf{k}) \times \mathbf{k}^n. \quad (3.5)$$

この群の $\mathbb{P}^n \setminus H \simeq \mathbf{k}^n$ への作用を考えればよい。このように \mathbf{k}^n のアフィン幾何を超平面 H を付加した射影空間 (\mathbb{P}^n, H) の射影幾何として扱うとき、 H は無限遠超平面 (*hyperplane at infinity*) と呼ばれる。

ユークリッド幾何 (Euclidean geometry) とは、 n 次元空間 \mathbf{k}^n と相似変換群 $\text{Sim } \mathbf{k}^n$ との組についての幾何学である*6。相似変換とは、 \mathbf{k}^n の回転・鏡映、平行移動と拡大・縮小を組み合わせで記述される次のような変換のことである：

$$\mathbf{k}^n \rightarrow \mathbf{k}^n, \quad \mathbf{x} \mapsto \lambda R\mathbf{x} + \mathbf{t}, \quad (3.6)$$

ここで $R \in O(n, \mathbf{k})$, $\mathbf{t} \in \mathbf{k}^n$, $\lambda \neq 0 \in \mathbf{k}$ とした。 n 次元空間 \mathbf{k}^n 内の図形が持つ性質のうち、相似変換で変わらないものとしては長さの比、角度などの計量的性質がある。アフィン幾何の場合と同様に、 \mathbf{k}^n のユークリッド幾何は超平面 H 上の滑らかな 2 次超曲面 Q を付加した射影空間 (\mathbb{P}^n, Q) に対する射影幾何と等価にな

*5 代数幾何においては、これらの射影不変量はむしろ最も具体的な部類の不変量という扱いになる。入れ物である射影空間を忘れ、それぞれの多様体の自己同型群についての幾何学を考えれば、ベッチ数やホッジ数のような不変量 (双正則不変量) が主役になる。さらに低次元の違いを忘れる双有理変換群についての幾何学を考えるなら、次元、小平次元、不正則数、多重種数、標準環といった最も基本的な部類の不変量 (双有理不変量) を扱うことになる (たとえば [11])。

*6 ユークリッドの合同変換群 $\text{Euc } \mathbf{k}^n \simeq O(n, \mathbf{k}) \times \mathbf{k}^n$ を採用する標準的な定義とは異なるが、再構成問題においては相似変換を考える方が自然である。これは事前情報なしに写真だけから物体の絶対的なスケールを知ることが原理的にできないためだ。部屋の写真だけが与えられても、それは小人の部屋かもしれないし巨人の部屋かもしれない。

る*7*8。実際、たとえば具体的に $x_n = 0$ で定まる超平面 $H \simeq \mathbb{P}^{n-1}$ の上で $x_0^2 + \cdots + x_{n-1}^2 = 0$ で定まる 2 次超曲面 Q をとると、この対応は容易に確認できる：

$$\text{Aut}(\mathbb{P}^n, Q) := \{g \in \text{Aut } \mathbb{P}^n \mid g \cdot Q \subset Q\} \quad (3.7)$$

$$= \left\{ \begin{pmatrix} \lambda R & \mathbf{t} \\ O & 1 \end{pmatrix} : \mathbb{P}^n \rightarrow \mathbb{P}^n \mid R \in O(n, \mathbf{k}), \mathbf{t} \in \mathbf{k}^n, \lambda \neq 0 \right\} \quad (3.8)$$

$$\simeq \text{Sim } \mathbf{k}^n \simeq \mathbf{k}^* O(n, \mathbf{k}) \times \mathbf{k}^n. \quad (3.9)$$

ここで $\mathbf{k}^* = \mathbf{k} \setminus \{0\}$ は \mathbf{k} の乗法群である。この Q という図形は、 \mathbb{P}^n において任意の球面

$$(x_0 - a_0 x_n)^2 + \cdots + (x_{n-1} - a_{n-1} x_n)^2 - (r x_n)^2 = 0 \quad (a_i, r \in \mathbf{k}), \quad (3.10)$$

が (H 上で) 交わる場所として特徴づけられる。クラインは、エルランゲン・プログラムの冒頭で、まさにここで説明した主張を紹介している：

計量的性質は幾何学においてはもはや空間における物自体の性質ではなくて、基本図形すなわち無限遠の球円との関連として現れるものだ。(〔9, 序文〕)

クラインのいう無限遠の球円 (unendlich fernen Kugelkreise) が Q のことであり、今日のコンピュータビジョンでは $n = 2$ で無限遠の円点 (circular points)、 $n = 3$ で絶対円錐曲線 (absolute conic, AC) などと呼ばれている。一般の次元でも絶対 2 次超曲面 (absolute quadric) と呼ぶことにしよう。計量的性質を扱うユークリッド幾何⁹が、この特定の図形を備えた射影幾何によって記述できるという事実は、再構成問題において鍵となる。



図 3.1 3次元形状の再構成の手続き。左から順に射影再構成、アフィン再構成、計量再構成によって得られる 3次元形状の例である。それぞれ射影変換、アフィン変換、相似変換で動かした 3次元形状のデータを φ で平面に射影して描いた。

3次元形状 (あるいは一般に n 次元形状) を再構成する手続きには、射影再構成、アフィン再構成、計量再構成と呼ばれる三つの基本的なステップがある。図 3.1 において各段階の再構成で得られる物体の 3次元形状を例示した。射影再構成とは、写真と手がかりから、射影変換による違いを除いて \mathbb{P}^n 内の図形の形状を復元する手続きである。典型的には、複数の写真の間の点対応 (point correspondence) が手がかりとなる。たとえば、二つのカメラ $\varphi, \varphi' : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2$ があるとき、点対応 $x \leftrightarrow x'$ というのは、3次元空間の点 $\tilde{x} \in \mathbb{P}^3$ が存在して $x = \varphi(\tilde{x})$ および $x' = \varphi'(\tilde{x})$ となるような写真上の点の組をいう。この場合、点対応 $x \leftrightarrow x'$ は有理写像

$$(\varphi, \varphi') : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^2, \quad \tilde{x} \mapsto (x, x') \quad (3.11)$$

*7 $\mathbf{k} = \mathbb{R}$ の場合は定符号の Q をとる。変換群 $\text{Aut } \mathbb{P}_{\mathbb{R}}^n \simeq \{g \in \text{Aut } \mathbb{P}_{\mathbb{R}}^n \mid g \cdot \mathbb{P}_{\mathbb{R}}^n \subset \mathbb{P}_{\mathbb{R}}^n\}$ は $\mathbb{P}_{\mathbb{C}}^n$ にも作用しているため、幾何学を変えずに空間だけを上げられる。したがって、集合としては $Q \subset \mathbb{P}_{\mathbb{C}}^n$ をとったと見なしてもよい。とくに (3.7) はそのように読む。

*8 ここでは H と Q を付加した射影空間 (\mathbb{P}^n, H, Q) に対する射影幾何と書いてもよいが、 Q から H は一意的に定まるため省略した。何となれば、 Q が異なる 2 つの超平面にのるとすると次元から線形な多様体になってしまう矛盾するからである。

の像に含まれる射影空間の直積 $\mathbb{P}^2 \times \mathbb{P}^2$ の点だとも見なせる（次章で詳しく見る）。現実の写真を扱うときは、点対応を検出するためのさまざまな事前処理が行われる*9。アフィン再構成、計量再構成とは、それぞれアフィン変換、相似変換による違いを除いて \mathbf{k}^n 内の図形の形状を復元する手続きをいう。射影再構成まで終えた状態で、アフィン再構成、計量再構成を行うには、それぞれ無限遠超平面 H 、絶対2次超曲面 $Q \subset H$ の位置を特定すればよい。そのための手がかりとして、主にカメラの種類や配置、シーンに対する事前の知識などがある。たとえば、異なる2枚の写真が平行移動した同一のカメラによって撮影されたものであれば、（適切な「写真」 \mathbb{P}^m の同型を介して）視差のない点対応 $x \leftrightarrow x$ が無限遠超平面 H から来ているものと分かる。おおざっぱには、これと射影再構成を組み合わせると n 次元のアフィン再構成が可能だ*10。3次元形状の計量再構成は、実用上とりわけ重要である。カメラの中心 O が無限遠超平面 H にのっていないときは、絶対円錐曲線の像（image of absolute conic, IAC） $\varphi(Q)$ が写真 \mathbb{P}^2 上でも円錐曲線になっており、重要な手がかりとなる。この $\varphi(Q)$ の写真 \mathbb{P}^2 における位置を知ることは、カメラのキャリブレーションといわれる。計量再構成を終えた状態からさらに必要に応じて被写体のより精密な物理的特性を再構成するための事後処理が行われる場合もある*11。もう一つ、 $\mathbf{k} = \mathbb{C}$ での再構成から $\mathbf{k} = \mathbb{R}$ での再構成が従うことをコメントしておく。これは注釈*7で述べたように、 $\mathbf{k} = \mathbb{C}$ での射影幾何、アフィン幾何、ユークリッド幾何が、 $\mathbf{k} = \mathbb{R}$ の場合のそれぞれの幾何学の自然な拡張になっているためである。仮に、 $\mathbb{P}_{\mathbb{R}}^n$ や \mathbb{R}^n の中にある被写体を $\mathbf{k} = \mathbb{C}$ で再構成して得られる n 次元形状があるとすると、この形状は必ずそれぞれの変換群で $\mathbb{P}_{\mathbb{R}}^n$ や \mathbb{R}^n の中へ移動することができて、これが $\mathbf{k} = \mathbb{R}$ での再構成となっている*12。

ここまで登場した概念を射影空間のトーリック幾何を用いて図 3.2 にまとめておこう。射影空間のトーリック幾何とは、射影空間 \mathbb{P}^n と代数的トーラス $T = (\mathbf{k}^*)^{n+1} / \mathbf{k}^* \simeq (\mathbf{k}^*)^n$ の組についての幾何学である。ここで \mathbb{P}^n の斉次座標の全ての成分がゼロでないようなザリスキー開集合を T だと見なす。代数的トーラス T の射影空間 \mathbb{P}^n への作用は次のような射影変換で与えられる：

$$\mathbb{P}^n \rightarrow \mathbb{P}^n, \quad [x_0 : \cdots : x_n] \mapsto [t_0 x_0 : \cdots : t_n x_n], \quad (3.12)$$

ここで $[t_0 : \cdots : t_n] \in T \subset \mathbb{P}^n$ とした。射影空間 \mathbb{P}^n に含まれる図形が持つ性質のうち、 T の作用で変わらないものとして、 T の軌道との交わり方がある。そこで軌道の様子を見ておきたいが、具体的に作用 (3.12) を眺めれば、単にどの座標成分がゼロであるか、という情報によって軌道が分類できることが分かる。これを踏まえ、射影空間 \mathbb{P}^3 および \mathbb{P}^2 のトーリック幾何を、それぞれの次元の単体を用いて模式的に表したものが図 3.2 である。改めて、(2.1) のピンホールカメラ模型 φ について図を見ながら振り返っておこう。

*9 それぞれの画像に対する特徴点の検出・記述、画像間での局所的な類似度の計算などである。このような事前処理においてはピンホールカメラ模型で無視してきたピクセルの色や明るさなどの情報も手がかりとして用いられるのが普通である。点対応の検出は不良設定 (ill-posed) な問題となることもしばしばであり、最適な点対応の候補を検出するために、相互相関やその他のコスト関数を考えたり、深層学習の手法を用いたりする。（たとえば [13]）。

*10 伊藤敦氏、植田一石氏との議論に基づく。

*11 たとえば、現実の写真からリアリスティックな 3D モデルを作成するような場合などである。この事後処理においてもピクセルの色や明るさなどの情報が手がかりとして用いられる。すなわち、計量再構成した高密度な点の集合を基に 3D モデルを作成し、テクスチャマッピングやシェーディングによって色や明るさの情報を反映したりする。

*12 逆は成り立たないことに注意。つまり、 $\mathbf{k} = \mathbb{R}$ で再構成ができる場合でも、 $\mathbf{k} = \mathbb{C}$ では、正しい形状の他に、変換群で $\mathbb{P}_{\mathbb{R}}^n$ や \mathbb{R}^n の中へ移すことのできない別の独立した形状の候補が得られてしまう、ということがあり得る。

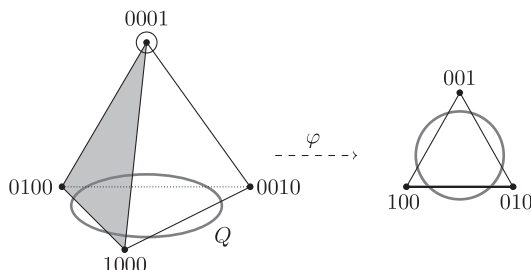


図 3.2 ピンホールカメラ模型の模式図. 左の四面体が \mathbb{P}^3 、右の三角形が φ の写した写真 \mathbb{P}^2 を表している。斉次座標について $0001 := [0 : 0 : 0 : 1] = O$ のような略記を用いた。左の \mathbb{P}^3 で、丸で囲った頂点がピンホールカメラ φ の中心 O であり、 O から下へ伸びている 3 つの辺が \mathbf{k}^3 の x 軸、 y 軸、 z 軸である。底面が $w = 0$ で定義される無限遠平面 H であって、その上に $x^2 + y^2 + z^2 = 0$ で定義される絶対円錐曲線 Q がのっている。網かけした面は中心 O を通るスクリーンと平行な平面（すなわち $z = 0$ で定義される主平面）で、最初の模型 (2.1) ではカメラに写らなかった部分である。右の \mathbb{P}^2 では、三角形の底辺が $z = 0$ で定義される無限遠直線 l であり、主平面の像となっている。上の頂点がコンピュータビジョンで主点 (principal point) と呼ばれるスクリーンの原点である。無限遠平面 $H \subset \mathbb{P}^3$ は \mathbb{P}^2 全体へ同型に写されていて、写真には絶対円錐曲線の像 (image of absolute conic, IAC) が写っている。この曲線 $\varphi(Q) \subset \mathbb{P}^2$ の位置を知ることをカメラのキャリブレーションというのであった。今の場合、 $\varphi(Q)$ は円 (虚円) なので、無限遠直線 l との交わりが、ちょうど無限遠の 2 つの円点に一致している。

4 カメラとブローアップ

ピンホールカメラ模型のグラフは、射影空間をカメラの中心に沿ってブローアップしたものになる。ブローアップとは、大ざっぱに言えば、中心の点を低い次元の射影空間に置き換える操作のことである。これは、ちょうど \mathbf{k}^n の無限遠に超平面 $H \simeq \mathbb{P}^{n-1}$ を付け加えてカメラの模型を拡張したことと同様の操作を、中心の点に無限に近いところでも行うことに相当する。まずは、このピンホールカメラ模型とブローアップの関係を、トーリック幾何を用いて図に描き、実際に観察してみよう。本章以降では、断らない限り $\mathbf{k} = \mathbb{C}$ とする。

トーリック多様体 (toric variety) とは、単項式のみでつくられる射影空間の間の有理写像

$$f : \mathbb{P}^n \dashrightarrow \mathbb{P}^N, \quad [x_0 : \cdots : x_n] \mapsto [x^{a_0} : \cdots : x^{a_N}] \quad (4.1)$$

の像のことをいう^{*13}。ここで多重指数 $\mathbf{x}^{a_j} := x_0^{a_{0j}} \cdots x_n^{a_{nj}}$ ($j = 0, \dots, N$) を用いた。必要なら射影空間における射影と合成して指数ベクトル $\mathbf{a}_0, \dots, \mathbf{a}_N \in \mathbb{Z}^{n+1}$ は全て異なるものとしてよい。一般に、指数ベクトルの集合 $\mathcal{A} = \{\mathbf{a}_0, \dots, \mathbf{a}_N\} \subset \mathbb{Z}^{n+1}$ から (4.1) により定義されるトーリック多様体を $\mathbb{P}_{\mathcal{A}}$ と書こう。それから、本稿では有理写像の像というとき、写像としての通常の像のザリスキー位相に関する閉包を取ったものを指すことにする。つまり、トーリック多様体が \mathbb{P}^N の閉部分多様体となるように定義している。有理写像のグラフというときも同じで、閉包を取ったものを指す。したがって、ピンホールカメラ模型 (2.2) のグラフ Γ_{φ} とは、有理写像

$$\tilde{\varphi} := (\text{id}_{\mathbb{P}^3} \times \varphi) : \mathbb{P}^3 \dashrightarrow \mathbb{P}^3 \times \mathbb{P}^2, \quad [x : y : z : w] \mapsto ([x : y : z : w], [x : y : z]) \quad (4.2)$$

^{*13} [12], [2] による標準的なトーリック多様体の定義とは異なるので注意。本稿では射影的でないトーリック多様体は扱わないため、(偏極付きの) 射影トーリック多様体のみを含む形での定義を採用した。ここでの定義は [7] にある (正規とは限らない) 射影トーリック多様体の定義と実質的に同じものである。ちなみに、標準的な定義というのは次のとおりである：トーリック多様体とは正規多様体 X であって代数的トーラス $T \simeq (\mathbf{k}^*)^d$ を稠密な開部分集合として含み、 T の X への作用が T 自身への自然な作用の延長として定まっているものをいう。

の像、すなわち $\mathbb{P}^3 \times \mathbb{P}^2$ のザリスキー位相に関する閉包を取ったものである。一般に、射影空間の直積 $\mathbb{P}^n \times \mathbb{P}^m$ はセグレの埋め込み (Segre embedding) $\mathbb{P}^n \times \mathbb{P}^m \rightarrow \mathbb{P}^{(n+1)(m+1)-1}$,

$$([x_0 : \cdots : x_n], [y_0 : \cdots : y_m]) \mapsto [x_0 y_0 : \cdots : x_i y_j : \cdots : x_n y_m] \quad (4.3)$$

により、大きな次元の射影空間の閉部分多様体と見なせることに注意する。とくに Γ_φ は (4.2) をセグレの埋め込み $\mathbb{P}^3 \times \mathbb{P}^2 \rightarrow \mathbb{P}^{11}$ と合成することで \mathbb{P}^{11} に埋め込まれたトーリック多様体と見なせることが分かる^{*14}。後ほど、この Γ_φ に戻ってこよう。

さて、トーリック幾何とは、トーリック多様体 \mathbb{P}_A と代数的トーラス $T \simeq (\mathbf{k}^*)^n$ の組についての幾何学である。 T の作用は (3.12) から誘導される次のような \mathbb{P}^N の射影変換をトーリック多様体 \mathbb{P}_A の上に制限したものである：

$$\mathbb{P}^N \rightarrow \mathbb{P}^N, \quad [y_0 : \cdots : y_N] \mapsto [t^{a_0} y_0 : \cdots : t^{a_N} y_N], \quad (4.4)$$

ここで $[t_0 : \cdots : t_n] \in T \subset \mathbb{P}^n$ とした。この作用は連続なので、たしかに T は \mathbb{P}_A を \mathbb{P}_A 自身の上に移している。射影空間の場合に見たように、トーリック幾何においては、トーリック多様体の中の図形が T の軌道とどのように交わるのかが不変量になるのであった。軌道の配置についてのデータは指数ベクトルの集合 $\mathcal{A} \subset \mathbb{Z}^{n+1}$ から読み取ることができる。有理写像を定義する単項式は全て同じ次数であるため、 \mathcal{A} の \mathbb{Q}^{n+1} における凸包 $\Delta = \text{Conv } \mathcal{A}$ は n 次元以下の整凸多面体となる。以下では \mathcal{A} が Δ に含まれる格子点全ての集合 $\Delta \cap \mathbb{Z}^{n+1}$ に一致している状況のみが登場する。この場合のトーリック多様体を \mathbb{P}_Δ と書き、多面体 Δ を \mathbb{P}_Δ のモーメント多面体 (moment polytope) という。トーリック幾何では、モーメント多面体 Δ の図をもって模式的にトーリック多様体 \mathbb{P}_Δ の図を描いたと見なすことができる。これは Δ の面構造がまさに \mathbb{P}_Δ における T の軌道の配置を表しているためだ。

ピンホールカメラ模型のグラフ Γ_φ の場合、指数ベクトルの集合 \mathcal{A} はセグレの埋め込みを用いて容易に計算できて、(重複を除き) 次のようになる：

$$\mathcal{A} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\} \subset \mathbb{Z}^4. \quad (4.5)$$

格子 \mathbb{Z}^4 のユニモジュラーな変換は T の生成系の取り換えに対応するだけなので幾何学を変えない。今の場合、モーメント多面体 $\Delta = \text{Conv } \mathcal{A} \subset \mathbb{Q}^4$ は 3 次元の多面体になる。そこで、(もしそちらの方が分かりやすいなら) 格子点を増やさないように \mathbb{Q}^3 に射影するかユニモジュラーな基底を選ぶなどすることで、モーメント多面体 Δ の図、すなわちピンホールカメラ模型のグラフ Γ_φ の模式図を描くことができる。このようにして描いたものが図 4.1(a) である。 Δ の頂点、辺、面がそれぞれ 0, 1, 2 次元の軌道 (あるいはその閉包として得られる部分トーリック多様体) と対応している。射影空間の直積 $\mathbb{P}^3 \times \mathbb{P}^2$ における第 1、第 2 成分への射影を $\Gamma_\varphi \subset \mathbb{P}^3 \times \mathbb{P}^2$ に制限し、それぞれ $p: \Gamma_\varphi \rightarrow \mathbb{P}^3$, $q: \Gamma_\varphi \rightarrow \mathbb{P}^2$ と書くことにしよう。この $p: \Gamma_\varphi \rightarrow \mathbb{P}^3$ がブローアップといわれる。実際、図 4.1(a) を見ると、 \mathbb{P}^3 において三角錐の頂点にあったカメラの中心 O が、 Γ_φ では網かけした三角形の表す超曲面 $E \simeq \mathbb{P}^2$ に置き換わっていることが見て取れる。一般のブローアップにおいて、カメラの中心 $O \in \mathbb{P}^3$ のように射影空間に置き換えられる点の集合をブローアップの中心 (center)、 $E \subset \Gamma_\varphi$ のように中心の逆像として得られる超曲面はブローアップの例外因子 (exceptional divisor) という。さて、ブローアップにもう少し近づいて観察したものが図 4.1(b) である。カメラの中心 O を通る直線 (光線) $\ell \subset \mathbf{k}^3$ がグラフ Γ_φ の方では $(O, \ell) \in E$ を通る直線に移っている。したがって、たとえば、 \mathbf{k}^3 の x 軸、

^{*14} (4.5) で計算するように、指数ベクトルの重複を除けば、実際には \mathbb{P}^8 に埋め込まれていることが分かる。

y 軸、 z 軸が Γ_φ に行くとき異なる三つの固定点で例外因子 E と交わることをどちらの図からも確認できる。ピンホールカメラ模型 φ ではカメラの中心 O が写真に写らなかったのだが、代わりにれっきとした写像である $q: \Gamma_\varphi \rightarrow \mathbb{P}^2$ を考えることで、全ての点が写真 \mathbb{P}^2 に写るようになった。この写像 $q: \Gamma_\varphi \rightarrow \mathbb{P}^2$ は \mathbb{P}^1 束である。例外因子 E の q による像は \mathbb{P}^2 全体を覆い、ファイバーにはスクリーンの各点に衝突する直線（光線）が並んでいる。

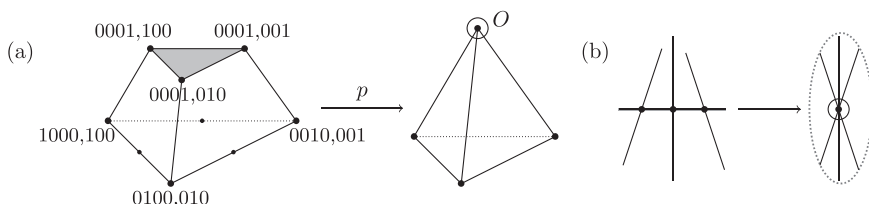


図 4.1 ピンホールカメラ模型とブローアップ. (a) ピンホールカメラ模型のグラフの模式図。左はピンホールカメラ φ のグラフ $\Gamma_\varphi \subset \mathbb{P}^3 \times \mathbb{P}^2$ をそのモーメント多面体を用いて表している。射影空間の直積 $\mathbb{P}^3 \times \mathbb{P}^2$ の斉次座標について $0001, 001 := ([0:0:0:1], [0:0:1])$ のような略記を用いた。右は \mathbb{P}^3 を再掲。矢印が表す第 1 成分への射影 $p: \Gamma_\varphi \rightarrow \mathbb{P}^3$ は中心 O でのブローアップになっており、網かけした面が例外因子 E である。(b) ブローアップの局所的な描像。ブローアップ $p: \Gamma_\varphi \rightarrow \mathbb{P}^3$ の中心 O 近傍での様子をつつ次元を下げて表している。右の \mathbb{P}^3 ではカメラの中心 O に集中していた直線（光線）が、左のグラフ Γ_φ 側では水平な線分で表された例外因子 E に沿ってばらけている。

5 射影再構成定理

二つの異なるカメラによって撮影された写真の組に対して 7 対の独立な点対応を与えれば、他の手がかりを使うことなく射影再構成ができる。これがコンピュータビジョンにおいて最も重要な二つの視点（両眼）の場合の、射影再構成定理である。7 対というのは、任意の点対応 $[\mathbf{x}] \leftrightarrow [\mathbf{x}']$ がいわゆる基礎行列 (fundamental matrix) F を用いた方程式

$$\mathbf{x}^T F \mathbf{x}' = 0 \quad (5.1)$$

を満たしており、 F が 7 個の独立なパラメータを持つことに起因している^{*15}。基礎行列 F が分かれば、カメラ行列 P, P' が簡単な公式を用いて計算でき、カメラ行列からは三角測量 (triangulation) によって 3 次元形状の復元ができる (たとえば [13])。以上の標準的な三つの手続きを幾何学の言葉に翻訳しよう。まず、点対応は有理写像 (3.11) の像 X に含まれるのであった。この像 $X \subset \mathbb{P}^2 \times \mathbb{P}^2$ はマルチビュー多様体 (multiview variety) といわれる^{*16}。つまり、(5.1) は $\mathbb{P}^2 \times \mathbb{P}^2$ におけるマルチビュー多様体 X の定義式であるから、最初のステップである基礎行列 F を決定するという手続きは、マルチビュー多様体 $X \subset \mathbb{P}^2 \times \mathbb{P}^2$ を決定することに他ならない。次に、 F からカメラ行列を求める部分は、 X から φ, φ' を (射影変換を除いて) 再構成することに相当する。最後の三角測量は (3.11) が与える双有理写像 $\mathbb{P}^3 \dashrightarrow X$ の逆を計算しているにすぎない。さ

^{*15} 行列 F と λF ($\lambda \neq 0$) が同じ条件を定めること、階数が落ちるという条件 $\det F = 0$ がある ($\ker F$ が後述のエビポール e' にあたる)。実用上は、非線形な行列式を直接解くことは避けて、8 対以上の点対応から基礎行列 F を決定する (正規化) 8 点法 (eight-point algorithm) がよく用いられている ([3], [4] など参照)。

^{*16} 二つの視点の場合に限らず、同様に定義したものをマルチビュー多様体という。ジョイントイメージ多様体 (joint-image variety) ともいわれる [8]。

て、この三つの手続きの中で、幾何学的に最も非自明なのは、アルゴリズム的には容易な、 X から φ, φ' を復元する部分である。いったいなぜ、マルチビュー多様体 X からカメラ φ, φ' を復元できるのであろうか？

二つの視点の場合、マルチビュー多様体はトーリック多様体として記述できる。実際、 \mathbb{P}^3 と \mathbb{P}^2 の射影変換で、異なる二つのカメラ φ, φ' を次のような標準形に変換できる^{*17}：

$$\varphi = (\varphi, \varphi') : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^2, \quad [x : y : z : w] \mapsto ([x : y : z], [x : y : w]). \quad (5.2)$$

この φ の像 (すなわちマルチビュー多様体) を $X_\varphi \subset \mathbb{P}^2 \times \mathbb{P}^2$ 、グラフを $\Gamma_\varphi \subset \mathbb{P}^3 \times \mathbb{P}^2 \times \mathbb{P}^2$ と書こう。前章で (4.5) を計算したときと全く同じように、セグレの埋め込みを経由して指数ベクトルの集合 \mathcal{A} を求めると、それぞれ (重複を除き) 次の行列の列ベクトルのなす集合となる。

$$X_\varphi : \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad (5.3)$$

$$\Gamma_\varphi : \begin{pmatrix} 3 & 2 & 2 & 1 & 1 & 1 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 & 0 & 1 & 0 & 3 & 2 & 1 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 & 0 & 0 & 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}. \quad (5.4)$$

これを基にそれぞれに対応するモーメント多面体 Δ の図、すなわちマルチビュー多様体 X_φ や φ のグラフ Γ_φ の模式図を描いたものが図 5.1 である。射影空間の直積 $\mathbb{P}^3 \times \mathbb{P}^2 \times \mathbb{P}^2$ の第 1、第 2、第 3 成分への射影を $\Gamma_\varphi \subset \mathbb{P}^3 \times \mathbb{P}^2 \times \mathbb{P}^2$ に制限し、これをそれぞれ $p : \Gamma_\varphi \rightarrow \mathbb{P}^3$, $q, q' : \Gamma_\varphi \rightarrow \mathbb{P}^2$ と書くことにする (前章の p, q は忘れてよい)。また、第 2, 3 成分への射影を

$$\mathbf{q} = (q, q') : \Gamma_\varphi \rightarrow X_\varphi \subset \mathbb{P}^2 \times \mathbb{P}^2 \quad (5.5)$$

とまとめて書く。今回の $p : \Gamma_\varphi \rightarrow \mathbb{P}^3$ は、二つのカメラの中心 O, O' を例外因子 E, E' に置き換えるブローアップになっている。この例外因子を $\mathbf{q} : \Gamma_\varphi \rightarrow X_\varphi$ で写した像が超曲面 $D, D' \subset X_\varphi$ である。 D, D' は点対応としては本来存在しない $\mathbb{P}^2 \times \mathbb{P}^2$ の点を含むのだが、閉包を取ったことでマルチビュー多様体 X_φ に追加された部分だ。元をたどればカメラの中心 O, O' に由来するものなので、本稿では D, D' をマルチビュー多様体 X_φ におけるカメラの因子 (camera divisor) と呼ぶことにしよう。二つのカメラの因子 D, D' は 1 点 $e = (e, e') \in X_\varphi$ で交わっている。この点はコンピュータビジョンではベースラインと呼ばれる直線 OO' 上の点からくる点対応 $e \leftrightarrow e'$ だ。写真上の点 e, e' は、それぞれのカメラの中心 O, O' をもう一方のカメラで捉えた像であり、コンピュータビジョンではエビポール (epipole) と呼ばれる。ベースライン OO' 上の点は、どちらのカメラで見ても常にエビポールに写るので、点対応だけでは分離できない。このため、ベースライン全体がマルチビュー多様体 X_φ ではエビポール対 $e = (e, e')$ の 1 点に移されるというわけだ。図 5.1 を見ると、エビポール対 e からは 4 本の軸が出ているが、これは e の近傍が、 x 軸、 y 軸、 z 軸という独立な 3 軸のみがある \mathbf{k}^3 の原点とは異なる構造を持つことを意味している。実際、この点は局所的に、

$$\{xy - zw = 0\} \subset \mathbf{k}^4 \quad (5.6)$$

と書かれる 3 次元図形の原点と同じ構造を持つ特異点 (singular point) だ。このマルチビュー多様体の特異点 $e \in X_\varphi$ は、代数幾何においては通常二重点 (ordinary double point) あるいはコニフォルド (conifold) などと呼ばれている。

^{*17} \mathbb{P}^2 の射影変換でカメラの中心以外の自由度は消せるのであった。カメラの中心が O と $O' = [0 : 0 : 1 : 0]$ になるように \mathbb{P}^3 の射影変換をするので、無限遠平面 H や絶対 2 次超曲面 Q の定義式は今までとは違うものになる。

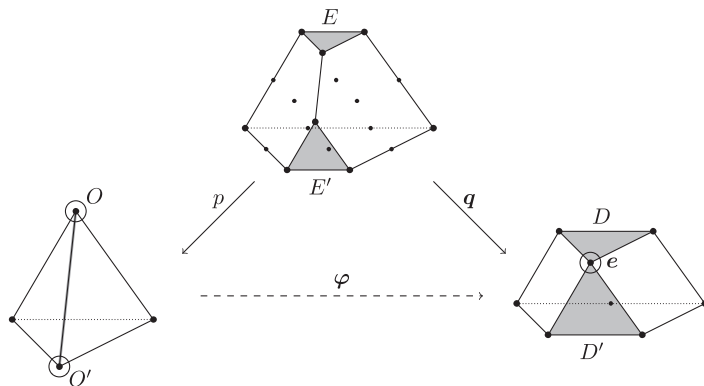


図 5.1 マルチビュー多様体の模式図。左の四面体が \mathbb{P}^3 、中央の多面体が φ のグラフ Γ_φ 、右の多面体が φ の像 (すなわちマルチビュー多様体) X_φ を表している。左の \mathbb{P}^3 では、二つの頂点 O, O' がそれぞれのカメラの中心であり、これを結ぶ直線 OO' がベースラインである。中央のグラフ Γ_φ は \mathbb{P}^3 上の 2 点 O, O' でのブローアップであって、網かけした二つの面が例外因子 E, E' である。右のマルチビュー多様体 X_φ へ行くと、ベースラインが潰れて 1 点になっている。この点 $e = (e, e') \in X_\varphi$ がエビポール対であり、通常二重点という特異点になっている。網かけした面が例外因子の像であるカメラの因子 D, D' で、エビポール対 e のみで交わっている。

二つの視点の場合には、マルチビュー多様体 $X = X_\varphi$ からカメラ $\varphi = (\varphi, \varphi')$ が復元できる様子を、トーリック幾何を使って概念的に理解することができる。まず、与えられたマルチビュー多様体 $X \subset \mathbb{P}^2 \times \mathbb{P}^2$ から、局所的な構造が異なるただ一つの特異点 $\hat{e} = (\hat{e}, \hat{e}')$ と、これを通る二つの因子 $\hat{D} = \mathbb{P}^2 \times \{\hat{e}'\}$, $\hat{D}' = \{\hat{e}\} \times \mathbb{P}^2$ を見つけることができる。このように $X \subset \mathbb{P}^2 \times \mathbb{P}^2$ だけから復元したものにはハット記号をつけて表しておこう。もちろん、今の場合はそれぞれ本当のエビポール対 e やカメラ因子 D, D' と一致する。エビポール対 $\hat{e} \in X$ は通常二重点なので、(少なくとも局所的には) ちょうど二つの小特異点解消 (small resolution) を持つことが知られている。通常二重点の小特異点解消とは、特異点を超曲面よりも低い次元の射影空間 (今の場合は \mathbb{P}^1) と置き換えるようなブローアップを行うことによって、滑らかにする操作をいう。今回は大域的にも小特異点解消が取れて、 Γ_φ の候補である $\hat{\Gamma}$ と $\hat{\Gamma}^+$ を X への双有理射 \hat{q}, \hat{q}^+ とともに得ることができる。これらもトーリック多様体なので図 5.2 のように模式図を描いてみよう。この二つの候補のうち、ピンホールカメラ模型のグラフとして許されるのは $\hat{\Gamma}$ の方だけである。実際、たとえば、カメラの因子 \hat{D} の固有変換 (proper transform)、すなわち $\hat{D} \setminus \{\hat{e}\}$ の逆像の閉包を見てみると、 $\hat{\Gamma}^+$ の方では射影平面 \mathbb{P}^2 ではなくて (ピカール数などの不変量が異なる) ヒルツェブルフ曲面 Σ_1 になっている。この固有変換はもともと例外因子 $E \simeq \mathbb{P}^2$ に一致するはずだったので、 $\hat{\Gamma}^+$ では矛盾してしまう。というわけで、 $\hat{\Gamma}$ まで復元することができた。最後に、二つの例外因子を潰す $\hat{p}: \hat{\Gamma} \rightarrow \mathbb{P}^3$ も同型を除いて一通りに定まるので、結局、射影変換を除いてカメラの復元 $\hat{\varphi} = \hat{q} \circ \hat{p}^{-1}$ は一意的だと分かる。したがって、 $\hat{\varphi}$ は射影変換を除いて φ と一致するというわけだ。

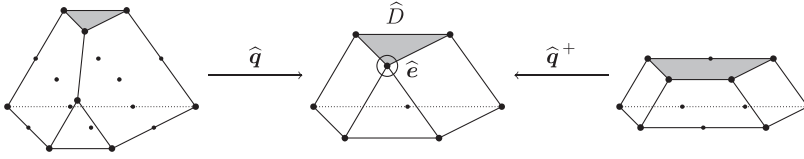


図 5.2 マルチビュー多様体の二通りの特異点解消の模式図. 左の多面体が $\widehat{\Gamma}$, 右の多面体が $\widehat{\Gamma}^+$ を表す. 中央が与えられたマルチビュー多様体 X だ. X における通常二重点 \widehat{e} を、 \widehat{q} では縦向き線分の表す \mathbb{P}^1 に、 \widehat{q}^+ では横向き線分の表す \mathbb{P}^1 にそれぞれ置き換えるブローアップをしている. マルチビュー多様体 X において特定したカメラの因子 \widehat{D} の固有変換を左右の図でも網かけをした面で表している. これが $\widehat{\Gamma}$ の方では三角形の表す \mathbb{P}^2 になるが、 $\widehat{\Gamma}^+$ の方では台形のヒルツェブルフ曲面 Σ_1 になっている.

この章で見てきた二つの視点の射影再構成定理を、多視点・一般次元の場合に拡張した次のような定理が成り立つ*18. 以下、二つの視点の場合に定めた記号を一般の場合にも拡張して用いる.

定理 5.1. 体を $\mathbf{k} = \mathbb{R}$ または \mathbb{C} 、次元を $n + 1 \leq \sum_j m_j$ とする. このとき、中心が一般的な配置にあるような r 個のカメラ $\varphi_1, \dots, \varphi_r$ に対して有理写像

$$\varphi = (\varphi_1, \dots, \varphi_r) : \mathbb{P}^n \dashrightarrow \mathbb{P}^m := \prod_{j=1}^r \mathbb{P}^{m_j} \tag{5.7}$$

の像として定まるマルチビュー多様体 $X = X_\varphi \subset \mathbb{P}^m$ が与えられているものとする. このとき、 $X \subset \mathbb{P}^m$ だけから復元されるカメラ $\widehat{\varphi}$ は、

- (a) $\mathbf{m} \neq (1^{n+1}) = (1, \dots, 1)$ の場合、射影変換を除いて φ に一致する.
- (b) $\mathbf{m} = (1^{n+1})$ の場合、射影変換を除いて二通りある φ または φ^V のどちらかに一致する.

定理 5.1 における射影再構成のおおざっぱなアイデアは次のようなものである. まず、前提である次元の条件 $n + 1 \leq \sum_j m_j$ は n 次元のマルチビュー多様体 X が \mathbb{P}^m の真の部分多様体となり、意味のある条件を与えるために必要な条件である. ピンホールカメラ模型は「写真」 \mathbb{P}^{m_i} 上の射影変換を除けば、カメラの中心 (すなわち射影空間における射影の中心) の位置によって定まるものなので、与えられた点対応からカメラの中心を見つけたいと考えるのが自然だろう. しかし、ピンホールカメラ模型ではカメラ自身を写すことができないので、点対応をいくら眺めていても、それぞれのカメラの中心を見つけることはできない (他の写真には写っているかもしれないが、点対応から対応するものを見つけれない). そこで、マルチビュー多様体 X の中で代わりのものを探すことにする. それが、カメラの因子である. カメラの因子を一つ見つければ、実はそこから芋づる式に射影再構成ができてしまう (具体的に $\widehat{\varphi}^{-1}$ が書ける). したがって、カメラの因子をどうやって見つけるのか、というのが問題となる. カメラの因子とは、カメラの中心をブローアップして得られる Γ_φ の例外因子から来ているものであった. 例外因子は $q_i : \Gamma_\varphi \rightarrow \mathbb{P}^{m_i}$ で「写真」全体を覆うように写っていたが、他のカメラからは単にカメラの中心としてしか見えていない. これがカメラの因子を特徴づける. すなわち、カメラの因子は、たとえば、 $D_1 = \mathbb{P}^{m_1} \times Y$ のような形をしていなくてはならない. ここで Y は残りのカメラからみた φ_1 の中心の点対応がなす図形である. $\mathbf{m} \neq (1^{n+1})$ の場合、 Y がどのように見えるべきかも考慮し

*18 マルチビュー多様体からの射影構成について述べた定理 5.1 は [6, Theorem 1.1] の一部であるが、[5, Theorem 5.1] で以前より知られていたグラスマンテンソルからの射影構成定理と等価になることが現在では分かっている. グラスマンテンソルとは、たとえば三つの視点の場合のトリフォーカルテンソル (trifocal tensor) のような、コンピュータビジョンで使われるマルチビューテンソルの一般化で、写真の間の部分空間対応から決まる量である.

ながら次元を勘定すれば、このような特徴を備えるマルチビュー多様体 X 上の因子（超曲面）がただ一つだけ存在し、他にないと言える。こうして、射影変換を除いて一意なカメラの復元 $\hat{\varphi}$ が得られるのである。

最後に、定理 5.1 において例外的なケースにあたる $m = (1^{n+1})$ の例を見ておこう。通常のピンホールカメラ模型のアナロジーで、次元を一つだけ下げた射影平面における射影 $\mathbb{P}^2 \dashrightarrow \mathbb{P}^1$ は、ラインカメラ（line camera）とも呼ばれている [3, §6.4.2]。三つの異なるラインカメラが与えられたとき、先と同じように \mathbb{P}^2 と \mathbb{P}^1 の射影変換で次のような標準形に変換できる：

$$\varphi = (\varphi_1, \varphi_2, \varphi_3) : \mathbb{P}^2 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1, \quad [x : y : z] \mapsto ([x : y], [y : z], [z : x]). \quad (5.8)$$

この有理写像 φ の像として得られるマルチビュー多様体 X_φ もトーリック多様体なので、セグレの埋め込みと合成して指数ベクトルの集合 \mathcal{A} を計算すると、

$$\begin{pmatrix} 1 & 2 & 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 & 1 \\ 2 & 1 & 0 & 0 & 1 & 2 & 1 \end{pmatrix} \quad (5.9)$$

という行列の列ベクトルの集合となる。モーメント多面体 Δ は図 5.3 の中央のような六角形となり、これは \mathbb{P}^2 の 3 点ブローアップで得られるデルペッツォ曲面（del Pezzo surface）だ。今回は、 $\mathbf{q} = (q_1, q_2, q_3) : \Gamma_\varphi \rightarrow X_\varphi$ がグラフとマルチビュー多様体との同型を与えている。今までのようにマルチビュー多様体 X_φ から射影再構成をしようとする、カメラの因子の選び方が一通りでないことに気が付く。実際、たとえば $z = 0$ と $z = \infty$ はどちらも $\mathbb{P}^1 \times Y$ という形をした因子を定めている（ Y は 1 点）。このどちらをカメラの因子と見なすかで、 φ もしくは

$$\varphi^\vee : \mathbb{P}^2 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1, \quad [x' : y' : z'] \mapsto ([y' : x'], [z' : y'], [x' : z']) \quad (5.10)$$

という二通りの再構成が得られるというわけだ。二つの再構成は標準クレモナ変換といわれる射影平面 \mathbb{P}^2 の双有理変換

$$\varphi^{\vee-1} \circ \varphi : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2, \quad [x : y : z] \mapsto [x' : y' : z'] = [yz : xz : xy] = \left[\frac{1}{x} : \frac{1}{y} : \frac{1}{z} \right] \quad (5.11)$$

で関係している。一般の $m = (1^{n+1})$ の場合には、 $\varphi^{\vee-1} \circ \varphi$ の与える射影空間 \mathbb{P}^n の双有理変換はベネロニ写像（Veneroni maps）と呼ばれる n 次の双有理変換になる [1]。



図 5.3 三つのラインカメラに対する二通りの射影再構成。左右の上下逆さまな三角形はどちらも \mathbb{P}^2 を、中央の六角形はマルチビュー多様体 X_φ 、あるいはそれと同型な φ のグラフ Γ_φ を表す。左の $p : \Gamma_\varphi \rightarrow \mathbb{P}^2$ は、カメラ φ の中心を太い線分の表す三つの \mathbb{P}^1 に置き換えるブローアップである。右の $p^\vee : \Gamma_\varphi \rightarrow \mathbb{P}^2$ もカメラ φ^\vee の中心を細い線分の表す三つの \mathbb{P}^1 に置き換えるブローアップである。

参考文献

- [1] M. DUMNICKI, L. FARNIK, B. HARBOURNE, T. SZEMBERG, AND H. TUTAJ-GASIŃSKA, *Veneroni maps*, in *Combinatorial Structures in Algebra and Geometry*, D. I. Stamate and T. Szemberg, eds., Cham, 2020, Springer International Publishing, pp. 31–42.

- [2] W. FULTON, *Introduction to toric varieties*, vol. 131 of Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 1993. The William H. Roever Lectures in Geometry.
- [3] R. HARTLEY AND A. ZISSERMAN, *Multiple view geometry in computer vision*, Cambridge University Press, Cambridge, second ed., 2003. With a foreword by Olivier Faugeras.
- [4] R. I. HARTLEY, *In defense of the eight-point algorithm*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997), pp. 580–593.
- [5] R. I. HARTLEY AND F. SCHAFFALITZKY, *Reconstruction from projections using Grassmann tensors*, International Journal of Computer Vision, 83 (2009), pp. 274–293.
- [6] A. ITO, M. MIURA, AND K. UEDA, *Projective reconstruction in algebraic vision*, Canad. Math. Bull., 63 (2020), pp. 592–609.
- [7] B. STURMFELS, *Gröbner bases and convex polytopes*, vol. 8 of University Lecture Series, American Mathematical Society, Providence, RI, 1996.
- [8] M. TRAGER, *Cameras, Shapes, and Contours: Geometric Models in Computer Vision*, theses, Ecole Normale Supérieure de Paris - ENS Paris, July 2018.
- [9] D. HILBERT, *Grundlagen der Geometrie*, Seventh edition, Teubner, 1930; F. KLEIN, *Vergleichende Betrachtungen über neuere geometrische Forschungen*. 1872. (ヒルベルト, クライン, 寺阪英孝・大西正男 (訳・解説), ヒルベルト 幾何学の基礎 クライン エルランゲン・プログラム. 現代数学の系譜 7, 共立出版, 1970.)
- [10] 川又雄二郎, 射影空間の幾何学. 講座 数学の考え方 11, 朝倉書店, 2001. 214 pp.
- [11] 川又雄二郎, 代数多様体論. 共立講座 21 世紀の数学 19, 共立出版, 1997. ii+223 pp.
- [12] 小田忠雄, 凸体と代数幾何学. 紀伊國屋数学叢書 24, 1985. iv+229 pp. (T. ODA, *Convex bodies and algebraic geometry*, vol. 15 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], Springer-Verlag, Berlin, 1988. An introduction to the theory of toric varieties, Translated from the Japanese. viii+212 pp.)
- [13] J. E. SOLEM, *Programming Computer Vision with Python*, (2012), 194 pp. (O’Reilly Media, Inc.) (Jan Erik Solem, 相川愛三 (訳), 実践 コンピュータビジョン. オライリー・ジャパン, 2013. xx+266pp.)