X

認知と学習

2. パターン情報の数理

2-1 視覚情報の数理

本節では人間の視覚による 3 次元空間の知覚を計算機で実現しようとする**コンピュータビジョン** (computer vision) の代表的手段である，画像からの 3 次元形状の計算法について説明する．

2-1-1 カメラモデル

画像 (image) を 3 次元シーンから 2 次元平面への写像とみなし，この写像を**カメラモデル** (camera model) とよぶ．最もよく用いられるのは，3 次元シーン中に**視点** (viewpoint) とよぶ点 O_c と**画像面** (image plane) とよぶ平面 Π_c を仮定し，シーン中の点 P が直線 O_cP と平面 Π_c との交点 p に写像されるとみなすものである (図 X 2.1)．これは理想的なピンホールカメラのモデルであり，**透視投影** (perspective projection) とよばれる．

視点 O_c を原点とし，視点 O_c から画像面 Π_c への垂線（これを**光軸** (optical axis) とよぶ）を Z_c 軸とする $X_cY_cZ_c$ 座標系をとる．そして，画像面 Π_c 上に，光軸との交点 (これを**光軸点** (principal point) とよぶ) を原点 o とし，X_c 軸，Y_c 軸にそれぞれ平行な x 軸，y 軸をもつ xy 画像座標系をとる (図 X 2.1)．このとき，シーンの点 (X_c, Y_c, Z_c) は次の点 (x, y) に投影される．

$$x = f_c\frac{X_c}{Z_c}, \qquad y = f_c\frac{Y_c}{Z_c} \tag{2.1}$$

ただし，f_c は視点 O_c から画像面 Π_c までの距離であり，**焦点距離** (focal length) とよぶ．

実際のデジタルカメラの画像面は**画素** (pixel) に対応する受光素子の配列である．その配列は複雑で（特にカラーカメラの場合）メーカーによって異なるが，概念的には縦横に平行に一定間隔で並んでいるとみなせる．しかし，縦横の方向が直交しているとは限らず，間隔も縦横で同じとは限らない．このとき左上隅の画素を $(u, v) = (0, 0)$ とし，右方向に $u = 1, 2, \ldots$ と数え，下方向に $v = 1, 2, \ldots$ と数える．この (u, v) をその画素の中心の位置とみなし，画素間はそれを補間して連続な座標系とみなしたものを**画素座標系** (pixel coordinate system) とする (図 X 2.2)．

受光素子の横の並びに x 軸をとり，それに直交する y 軸を下向きにとれば，uv 画素座標系と xy 画像座標系は次の関係にある (図 X 2.2)．

$$u = x/\alpha + (y/\alpha)\tan\theta + u_0, \quad v = y/\beta + v_0 \tag{2.2}$$

(u_0, v_0) が光軸点の画素座標であり，α, β はそれぞれ受光素子の u, v 方向の間隔である．その u 方向と v 方向の成す角を $\pi/2 + \theta$ とし，θ を**歪み角** (skew angle) とよぶ．式 (2.1), (2.2) を合わせると次のように書ける．

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \simeq \boldsymbol{K} \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} \tag{2.3}$$

以下，$\simeq$ は両辺が 0 でない定数倍の関係にあることを表す．そして次のようにおいた．

$$\boldsymbol{K} = \begin{pmatrix} f\gamma & f\gamma\tan\theta & u_0 \\ & f & v_0 \\ & & 1 \end{pmatrix} \tag{2.4}$$

ただし，$f = f_c/\beta$ であり，受光素子の v 方向の間隔を 1 とする単位で測った焦点距離である．また $\gamma = \beta/\alpha$ であり，**アスペクト比** (aspect ratio) とよぶ

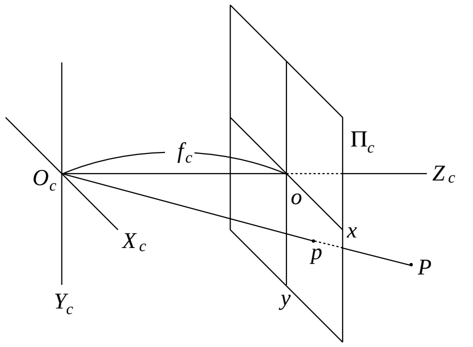

図 1-1 透視投影

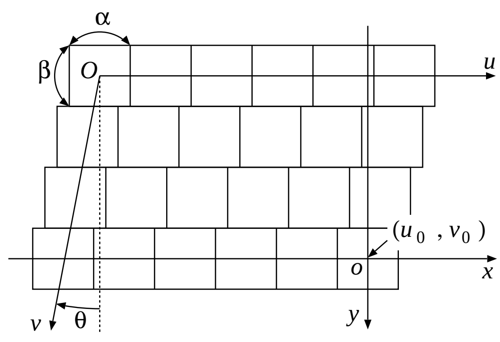

図 1-2 画素座標系

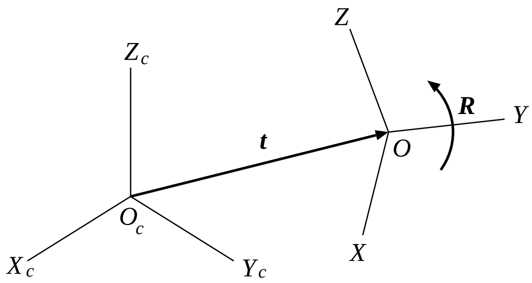

図 1-3　カメラ座標系とワールド座標系

(通常は $\gamma \approx 1$, $\theta \approx 0$ であり，(u_0, v_0) は画像面の中央)．これら f, γ, θ, u_0, v_0 をカメラの**内部パラメータ** (intrinsic parameters) とよび，$\boldsymbol{K}$ をカメラの**内部パラメータ行列** (intrinsic parameter matirx) とよぶ．

$X_cY_cZ_c$ 座標系はカメラの視点と光軸に固定した座標系であり，**カメラ座標系** (camera coordinate system) とよぶ．これとは別にシーンに固定した XYZ 座標系を考え，**ワールド座標系** (world coordinate system) とよぶ．ワールド座標系はカメラ座標系を $\boldsymbol{t}$ だけ並進し，$\boldsymbol{R}$ だけ回転した位置にあるとすると (図 X 2.3)，次の関係が成り立つ．

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \boldsymbol{t} \tag{2.5}$$

$\{\boldsymbol{R}, \boldsymbol{t}\}$ をカメラの**運動パラメータ** (motion parameters), または**外部パラメータ** (extrinsic parameters) とよぶ．

式 (2.3), (2.5) より，ワールド座標系の点 (X, Y, Z) とその投影像の画素座標 (u, v) は次式で結ばれる．

$$\boldsymbol{u} \simeq \boldsymbol{P}\boldsymbol{X} \tag{2.6}$$

ただし，$\boldsymbol{u} = (u, v, 1)^\top$, $\boldsymbol{X} = (X, Y, Z, 1)^\top$ とおいた．これらはそれぞれ画像上の点 (u, v) とシーンの点 (X, Y, Z) の**同次座標** (homogeneous coordinates) を表すベクトルである．$\boldsymbol{P}$ は次の 3×4 行列であり，**投影行列** (projection matrix)，または**カメラ行列** (camera matrix) とよぶ．

$$\boldsymbol{P} = \boldsymbol{K} \begin{pmatrix} \boldsymbol{R} & \boldsymbol{t} \end{pmatrix} \tag{2.7}$$

以下，$\boldsymbol{u}$ の表す点 (u, v) や $\boldsymbol{X}$ の表す点 (X, Y, Z) を単に "点 $\boldsymbol{u}$"，"点 $\boldsymbol{X}$" とよぶ．

式 (2.6) は同次座標の関係式であり，3 次元空間 $\mathcal{R}^3$ に無限遠点を付け加えた 3 次元射影空間 $\mathcal{P}^3$ から，2 次元空間 $\mathcal{R}^2$ に無限遠点を付け加えた 2 次元射影空間 $\mathcal{P}^2$ への投影とみなせる．$\mathcal{P}^3$ の点は同次座標 $X^1 : X^2 : X^3 : X^4$ によって指定される．$X^4 = 0$ となる点の集合 Π_∞ を**無限遠平面** (ideal plane) とよぶ．この上の $(X^1)^2 + (X^2)^2 + (X^3)^2 = 0$ で表される虚円 Ω_∞ を**絶対 2 次** (または**円錐**) **曲線** (absolute conic) とよぶ．式 (2.6) から，Ω_∞ 上の点の投影像 $\boldsymbol{u}$ が運動パラメータ $\{\boldsymbol{R}, \boldsymbol{t}\}$ によらずに次式を満たすことがわかる．

$$\boldsymbol{u}^\top \boldsymbol{\omega} \boldsymbol{u} = 0, \qquad \boldsymbol{\omega} = (\boldsymbol{K}^{-1})^\top \boldsymbol{K}^{-1} \tag{2.8}$$

これも画像面上の虚円であり，絶対 2 次曲線の投影像と解釈できる．3 次元位置が既知の図形を撮影し，その投影像からカメラの内部パラメータや運動パラメータを逆算することを**カメラ校正** (camera calibration) という．多くのカメラ校正の手続きに絶対 2 次曲線を用いた射影幾何学的な解釈を与えることができる [1]．

2-1-2　エピ極線幾何学

シーン中の点を複数の位置から撮影したとき，それらの各画像上の投影像の関係を記述するのが**エピ極線幾何学** (epipolar geometry) である [1]．シーン中の点 $\boldsymbol{X}$ を M 台のカメラで撮影し，各画像上の投影位置が $\boldsymbol{u}_\kappa$, $\kappa = 1, ..., M$ であるとする．各カメラの投影行列を $\boldsymbol{P}_\kappa$ とすると，それぞれに対して式 (2.6) が成立する．比例定数 λ_κ(これを**射影的奥行き** (projective depth) とよぶ) を用いて等号 $=$ で表すと次のようになる．

$$\lambda_\kappa \boldsymbol{u}_\kappa = \boldsymbol{P}_\kappa \boldsymbol{X} \tag{2.9}$$

$\kappa = 1, ..., M$ の式をまとめると，次のように書ける．

$$\begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{u}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{P}_2 & \boldsymbol{0} & \boldsymbol{u}_2 & \cdots & \boldsymbol{0} \\ \vdots & \boldsymbol{0} & \boldsymbol{0} & \ddots & \vdots \\ \boldsymbol{P}_M & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{u}_M \end{pmatrix} \begin{pmatrix} \boldsymbol{X} \\ -\lambda_1 \\ \vdots \\ -\lambda_M \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.10}$$

これを満たす $\boldsymbol{X}$ $(\neq \boldsymbol{0})$ と射影的奥行き λ_κ $(\neq 0)$, $\kappa = 1, ..., M$ が存在するから，左辺の $3M \times (M+4)$ 行列はランクが最大 $M+3$ である．ゆえに任意の $(M+4) \times (M+4)$ 小行列式は 0 となる．これから M $(= 2, 3, 4)$ 画像間の拘束条件が導かれる．

$M = 2$ (2 画像) のときは左辺の行列は 6×6 となり，その行列式が 0 という拘束条件が得られる．これを書き直すと次のようになる．

$$\boldsymbol{u}_1^\top \boldsymbol{F} \boldsymbol{u}_2 = 0 \tag{2.11}$$

ただし，$\boldsymbol{F}$ は (ij) 要素が次式の 3×3 行列である．

$$F_{ij} = \sum_{k,l,m,n=1}^{3} \epsilon_{ikl} \epsilon_{jmn} \det \boldsymbol{P}_{1122}^{klmn} \tag{2.12}$$

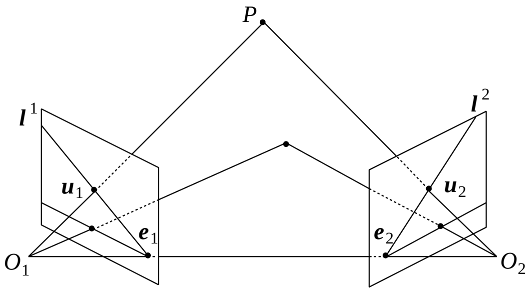

図 1-4　エピ極線とエピ極点

$\boldsymbol{P}^{klmn}_{1122}$ は $\boldsymbol{P}_1$ の k 行, $\boldsymbol{P}_1$ の l 行, $\boldsymbol{P}_2$ の m 行, $\boldsymbol{P}_2$ の n 行を順に 1 行, ..., 4 行とする 4×4 行列であり，ϵ_{ijk} は (ijk) の符号である（(ijk) が (123) の偶順列なら 1，奇順列なら -1，　それ以外は 0)．定義より $\boldsymbol{F}$ はランク 2 であり，**基礎行列** (fundamental matrix) とよばれる．

式 (2.11) は幾何学的には, 第 1 画像の視点 O_1 と点 $\boldsymbol{u}_1$ を結ぶ直線 (これを点 $\boldsymbol{u}_1$ の**視線** (line of sight) とよぶ)，および第 2 画像の視点 O_2 と点 $\boldsymbol{u}_2$ を結ぶ直線 (点 $\boldsymbol{u}_2$ の視線) が同一平面にあること，すなわち 2 点の視線が (無限遠点を含めて)1 点で交わることを表す (図 X 2.4)．

画像上で式 $\boldsymbol{l}^\top \boldsymbol{u} = 0$ を満たす点 $\boldsymbol{u}$ の集合は直線であり，$\boldsymbol{l}$ の成分がこの直線の同次座標 (非零の定数倍は同一視する) である．これを単に “直線 $\boldsymbol{l}$” とよぶ．式 (2.11) より，第 1 画像の点 $\boldsymbol{u}_1$ は直線 $\boldsymbol{l}^1 = \boldsymbol{F}\boldsymbol{u}_2$ の上にある (図 X 2.4)．この直線を “第 2 画像の点 $\boldsymbol{u}_2$ の定める**エピ極線** (epipolar line)” とよぶ．式 (2.11) はまた, 第 2 画像の点 $\boldsymbol{u}_2$ が直線 $\boldsymbol{l}^2 = \boldsymbol{F}^\top \boldsymbol{u}_1$ の上にあることも意味している (図 X 2.4)．この直線を “第 1 画像の点 $\boldsymbol{u}_1$ の定めるエピ極線” とよぶ.

要するに, 式 (2.11) は “対応する点 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ が互いに他方の定めるエピ極線上になければならない” という関係を表している．これを**エピ極線拘束** (epipolar constraint) とよぶ．したがって，2 画像間の基礎行列 $\boldsymbol{F}$ を計算すれば, 一方の画像の任意の点 $\boldsymbol{u}_1$ の他画像上の対応点 $\boldsymbol{u}_2$ を求めるには点 $\boldsymbol{u}_1$ の定めるエピ極線上を探索すればよい (図 X 2.4)．これは画像間の対応を求める画像処理の基本原理である．

基礎行列 $\boldsymbol{F}$ はランクが 2 であるから，$\boldsymbol{F}^\top \boldsymbol{e}_1 = \boldsymbol{0}$, $\boldsymbol{F}\boldsymbol{e}_2 = \boldsymbol{0}$ となる $\boldsymbol{0}$ でないベクトル $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ が存在する．これらをその成分を同次座標とする点と同一視すれば，$\boldsymbol{e}_1$ は “第 2 画像の視点 O_2 の第 1 画像上の投影像” であり，$\boldsymbol{e}_2$ は “第 1 画像の視点 O_1 の第 2 画像上の投影像” になっている (図 X 2.4)．これらをそれぞれの画像の**エピ極点** (epipole) とよぶ．

式 (2.11) より第 2 画像の任意の点の定めるエピ極線は第 1 画像のエピ極点 $\boldsymbol{e}_1$ を通り，第 1 画像の任意の点の定めるエピ極線は第 2 画像のエピ極点 $\boldsymbol{e}_2$ を通ることがわかる．すなわち，他画像のすべての点のエピ極線はエピ極点を通る放射状の**線束** (pencil of lines) をなしている (図 X 2.4)．これは，“画像上の点 $\boldsymbol{u}$ と 2 画像の視点 O_1, O_2 を結ぶ直線 (これを**基線** (baseline) とよぶ) を含む平面 (これを点 $\boldsymbol{u}$ の**エピ極面** (epipolar plane) とよぶ) と他方の画像面の交線が $\boldsymbol{u}$ の定めるエピ極線である” という関係に対応している (図 X 2.4)．

$M = 3$ (3 画像) のときは式 (2.10) から次の**三重線形拘束** (trilinear constraint) を得る．

$$\sum_{i,j,k,l,m=1}^{3} \epsilon_{jlp}\epsilon_{kmq} T_i^{jk} u_1^i u_2^l u_3^m = 0 \tag{2.13}$$

ただし，u_κ^i は $\boldsymbol{u}_\kappa$ の i 成分であり，T_i^{jk} は次のように定義する**三重焦点テンソル** (trifocal tensor) である．

$$T_i^{jk} = \sum_{l,m=1}^{3} \epsilon_{ilm} \det \boldsymbol{P}^{lmjk}_{1123} \tag{2.14}$$

第 2 画像の点 $\boldsymbol{u}_2$ の通る任意の直線 $\boldsymbol{l}^2$ と第 2 画像の視点 O_2 の張る平面 Π_2 (直線 $\boldsymbol{l}^2$ の**逆投影** (back projection) とよぶ)，および第 3 画像の点 $\boldsymbol{u}_3$ の通る任意の直線 $\boldsymbol{l}^3$ と第 3 画像の視点 O_3 の張る平面 Π_3 (直線 $\boldsymbol{l}^3$ の逆投影) を考えると，式 (2.13) は “第 1 画像の点 $\boldsymbol{u}_1$ の視線が平面 Π_2, Π_3 の交線と交わる” という関係を表す (図 X 2.5)．

$M = 4$ (4 画像) のときは式 (2.10) から次の**四重線形拘束** (quadrilinear constraint) を得る．

$$\sum_{i,j,k,l,m,n,p,q=1}^{3} \epsilon_{ima}\epsilon_{jnb}\epsilon_{kpc}\epsilon_{lqd} Q^{ijkl} u_1^m u_2^n u_3^p u_4^q = 0 \tag{2.15}$$

ただし，Q^{ijkl} は次のように定義する**四重焦点テンソル** (quadrifocal tensor) である．

$$Q^{ijkl} = \det \boldsymbol{P}^{ijkl}_{1234} \tag{2.16}$$

式 (2.15) は幾何学的には，“第 1 ～ 4 画像の点 $\boldsymbol{u}_1$ ～ $\boldsymbol{u}_4$ を通る任意の直線 $\boldsymbol{l}^1$ ～ $\boldsymbol{l}^4$ の逆投影 Π_1 ～ Π_4 が 1 点で交わる” という関係を表す (図 X 2.6)．

2-1-3　画像からの 3 次元復元

シーン中の N 点 $\boldsymbol{X}_\alpha$ を M 台のカメラで撮影して得られる第 κ 画像上の点を $\boldsymbol{u}_{\kappa\alpha}$ とする．各カメラの投影行列を $\boldsymbol{P}_\kappa$ とすると，式 (2.6) より次式が成り立つ．

$$\boldsymbol{u}_{\kappa\alpha} \simeq \boldsymbol{P}_\kappa \boldsymbol{X}_\alpha \tag{2.17}$$

投影像 $\boldsymbol{u}_{\kappa\alpha}$, $\kappa = 1, ..., M$, $\alpha = 1, ..., N$ が与え

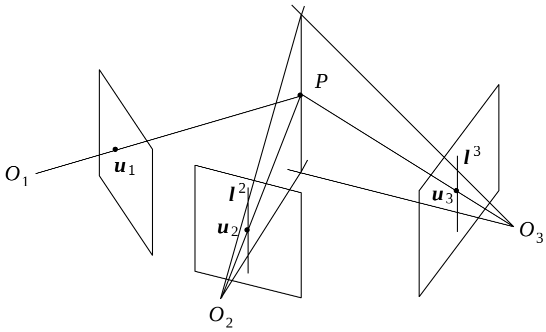

図 1-5　三重線形拘束

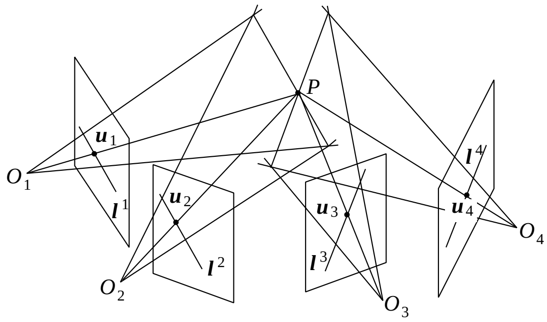

図 1-6　四重線形拘束

られたときに $\boldsymbol{X}_\alpha$, $\alpha=1, ..., N$ を計算する問題を画像からの**3 次元復元** (3-D reconstruction または structure from motion) という．これは次の 3 種類に分類できる．

(i)　各カメラの投影行列 $\boldsymbol{P}$ が既知.

(ii)　各カメラの内部パラメータ行列 $\boldsymbol{K}$ が既知 (運動パラメータ $\{\boldsymbol{R}, \boldsymbol{t}\}$ は未知).

(iii)　各カメラの投影行列 $\boldsymbol{P}$ が未知.

(i) の場合は式 (2.17) から点 $\boldsymbol{X}_\alpha$ の 3 次元座標 $(X_\alpha, Y_\alpha, Z_\alpha)$ が 1 自由度 (視線方向の**奥行き** (depth)) を除いて定まる．したがって，2 枚以上の画像があればすべての点の 3 次元位置が求まる．このようにして 3 次元復元を行うことを (**多眼**) **ステレオ視** ((multi-camera) stereo vision) とよぶ．

(ii) のとき, カメラは**校正済み** (calibrated) であるという．このときは画像間の対応点に式 (2.11) のエピ極線拘束 (あるいは式 (2.13) の三重線形拘束，式 (2.15) の四重線形拘束) を当てはめ, 異なる画像間の基礎行列 $\boldsymbol{F}$ (あるいは三重焦点テンソル T_i^{jk}，四重焦点テンソル Q^{ijkl}) を求めれば，定義式 (2.12)(あるいは式 (2.14), (2.16)) を解いて画像間の運動パラメータ $\{\boldsymbol{R}, \boldsymbol{t}\}$ が求まる（普通は基礎行列 $\boldsymbol{F}$ のみが用いられる)．したがって，問題は (i) のステレオ視に帰着する．

(iii) のときカメラは**未校正** (uncalibrated) であるといい，これから 3 次元復元を行うことを**自己校正**

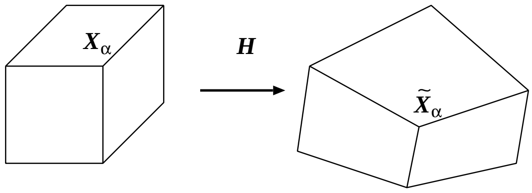

図 1-7　射影変換

(self-calibration) とよぶ．このときは式 (2.17) で $\{\boldsymbol{P}_\kappa\}$ と $\{\boldsymbol{X}_\alpha\}$ のすべてが未知数となる．しかし, カメラモデルに何の制約もなければ解は不定である．なぜなら，もし $\{\boldsymbol{X}_\alpha\}$, $\{\boldsymbol{P}_\kappa\}$ が解であるとすると, 任意の 4×4 正則行列 $\boldsymbol{H}$ に対して

$$\tilde{\boldsymbol{X}}_\alpha \simeq \boldsymbol{H}\boldsymbol{X}_\alpha,\quad \tilde{\boldsymbol{P}}_\kappa \simeq \boldsymbol{P}_\kappa\boldsymbol{H}^{-1} \qquad (2.18)$$

とおくと，$\{\tilde{\boldsymbol{X}}_\alpha\}$, $\{\tilde{\boldsymbol{P}}_\kappa\}$ も式 (2.17) を満たすからである．

式 (2.18) の第 1 式は 3 次元射影空間 $\mathcal{P}^3$ 全体に行列 $\boldsymbol{H}$ で定まる**射影変換** (projective transformation または homography) を施しているとみなせる (図 X 2.7)．このように変換した $\{\tilde{\boldsymbol{X}}_\alpha\}$ も，同一直線上にあるべき点は同一直線上に，同一平面上にあるべき点は同一平面上にあり，点や直線や平面の接続関係 (“上にある”，“通る”，“交わる” など) も正しい．しかし，角度や比は一般に異なっている．このような射影変換の不定性のある復元を**射影復元** (projective reconstruction) とよぶ．

式 (2.18) は次のようにも書ける．

$$\boldsymbol{X}_\alpha \simeq \boldsymbol{H}^{-1}\tilde{\boldsymbol{X}}_\alpha,\quad \boldsymbol{P}_\kappa \simeq \tilde{\boldsymbol{P}}_\kappa\boldsymbol{H} \qquad (2.19)$$

N 点 $\{\boldsymbol{X}_\alpha\}$ の内の一般の位置にある 5 点の 3 次元位置が既知なら，対応する 5 点 $\{\tilde{\boldsymbol{X}}_\alpha\}$ がその 5 点 $\{\boldsymbol{X}_\alpha\}$ に一致するような式 (2.18), (2.19) の射影変換 $\boldsymbol{H}$ が一意的に定まる．これを用いて正しい復元 (これを**ユークリッド復元** (Euclidean reconstruction) とよぶ) $\{\boldsymbol{X}_\alpha\}$ と正しい投影行列 $\{\boldsymbol{P}_\kappa\}$ が定まる．

そのような 5 点が知られていないときにユークリッド復元に変換する代表的な方法は**段階的復元** (stratified reconstruction) である [1]．これは，カメラモデルに何らかの制約を仮定し，射影復元から求まる $\{\tilde{\boldsymbol{P}}_\kappa\}$ を式 (2.19) の第 2 式によって変換した $\{\boldsymbol{P}_\kappa\}$ がその制約を満たすように射影変換 $\boldsymbol{H}$ を定めるものである．

式 (2.7) から関係 $\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{I}$ (単位行列) を用いて回転行列 $\boldsymbol{R}$ を消去すると，各画像について次式が成り立つ（$\mathrm{diag}(a, b, c, ...)$ は a, b, c, ... をこの順に対角要素とする対角行列)．

$$\boldsymbol{P}_\kappa \mathrm{diag}(1,1,1,0)\boldsymbol{P}_\kappa^\top = \boldsymbol{\omega}_\kappa^* \qquad (2.20)$$

ただし，3×3 行列 $\boldsymbol{\omega}_\kappa^*$ を次のようにおいた．

$$\boldsymbol{\omega}_\kappa^* \equiv \boldsymbol{K}_\kappa\boldsymbol{K}_\kappa^\top \qquad (2.21)$$

式 (2.20) に式 (2.19) の第 2 式の $\boldsymbol{P}_\kappa$ を代入すると次のように書ける．

$$\tilde{\boldsymbol{P}}_\kappa \boldsymbol{\Omega}_\infty^* \tilde{\boldsymbol{P}}_\kappa^\top \simeq \boldsymbol{\omega}_\kappa^* \tag{2.22}$$

ただし，4×4 行列 $\boldsymbol{\Omega}_\infty^*$ を次のようにおいた．

$$\boldsymbol{\Omega}_\infty^* \equiv \boldsymbol{H}\mathrm{diag}(1,1,1,0)\boldsymbol{H}^\top \tag{2.23}$$

カメラの内部パラメータ行列 $\boldsymbol{K}_\kappa$ が未知でも内部パラメータに関する何らかの制約があれば (2.21) から $\boldsymbol{\omega}_\kappa^*$ の要素の間の関係 (特定の要素が 0 である，特定に二つの要素が等しいなど) が得られる．これを用いて式 (2.22) から $\boldsymbol{\Omega}_\infty^*$ の要素に関する方程式が得られ ($\tilde{\boldsymbol{P}}_\kappa$ は仮定により既知である)，画像数 M が多く，そのような方程式が十分多数得られれば，それらを解いて $\boldsymbol{\Omega}_\infty^*$ が定まる．

内部パラメータに関する仮定としては次のものが代表的である [1]．

- 各カメラは同じ内部パラメータをもつ．
- 各カメラの光軸点は既知．
- 各カメラは歪み角 θ が 0.
- 各カメラはアスペクト比 γ が 1.

例えば内部パラメータが同じであれば (一つのカメラをそのまま移動しながら撮影すれば)，内部パラメータ行列 $\boldsymbol{K}$ が共通であるから $\boldsymbol{\omega}_1^* = ... = \boldsymbol{\omega}_M^* = \boldsymbol{\omega}^*$ $(\equiv \boldsymbol{K}\boldsymbol{K}^\top)$ であり，式 (2.22) から $\boldsymbol{\Omega}_\infty^*$ に対する $5(M-1)$ 個の拘束式が得られる．一方，光軸点が既知であれば，そこを原点とする画像座標を用いることにより式 (2.4) の $\boldsymbol{K}$ は (13), (23) 要素が 0 となるので，各 $\boldsymbol{\omega}_\kappa^* = \boldsymbol{K}_\kappa\boldsymbol{K}_\kappa^\top$ も (13), (23) 要素が 0 となり，式 (2.22) から $\boldsymbol{\Omega}_\infty^*$ に対する $2M$ 個の拘束式が得られる．さらに $\theta = 0$ であれば (12) 要素も 0 となるので，拘束式は $3M$ 個となる．$\gamma = 1$ であれば $\boldsymbol{\omega}_\kappa^*$ の (11) 要素と (22) 要素が等しくなるので，拘束式はさらに M 個増える．

これらの拘束式から $\boldsymbol{\Omega}_\infty^*$ が定まれば，式 (2.22) から $\boldsymbol{\omega}_\kappa^*$ が定まり，式 (2.23) から射影変換の行列 $\boldsymbol{H}$ が定まる（一意的ではない）．また式 (2.21) を解いて各 $\boldsymbol{K}_\kappa$ が定まる．

式 (2.8) の第 2 式と式 (2.21) を比較すれば，$\boldsymbol{\omega}_\kappa^*$ は第 κ 画像の絶対 2 次曲線 Ω_∞ の像を表す $\boldsymbol{\omega}_\kappa$ の逆行列 ($\boldsymbol{\omega}_\kappa^* = \boldsymbol{\omega}_\kappa^{-1}$) であることがわかる．$\boldsymbol{l}$ を直線を表す変数とすれば，$\boldsymbol{l}^\top\boldsymbol{\omega}_\kappa^*\boldsymbol{l} = 0$ を満たす直線の集合は式 (2.8) の第 1 式の表す虚円の接線の集合 (包絡線) である．射影幾何学の用語ではこれを 2 次曲線 $\boldsymbol{u}^\top\boldsymbol{\omega}_\kappa\boldsymbol{u} = 0$ に**双対** (dual) な **2 次線束** (dual conic) とよぶ．

式 (2.22) は $\boldsymbol{\omega}_\kappa^*$ の表す 2 次線束が $\boldsymbol{\Omega}_\infty^*$ の表す **2 次面束** (dual quadric) の投影像であることを表している．$\boldsymbol{\Omega}_\infty^*$ の表す 2 次面束とは，平面の同次座標を成分とするベクトルを $\boldsymbol{\pi}$ とするとき，$\boldsymbol{\pi}^\top\boldsymbol{\Omega}_\infty^*\boldsymbol{\pi} = 0$ を満たす平面の集合のことである．これは退化した 2 次曲面とみなした絶対 2 次曲線 Ω_∞ の接平面の集合 (包絡面) であり，**絶対 2 次面束** (または**双対絶対 2 次曲面**)(dual absolute quadric) とよぶ．このような射影幾何学的な解釈から，式 (2.22) は**絶対 2 次面束** (または**双対絶対 2 次曲面**) **拘束** (dual absolute quadric constraint) とよばれる [1), 6)]．

2-1-4 因子分解法による射影復元

前項の方法を適用するにはまず射影復元を計算する必要がある．そのためによく用いられるのは次の**因子分解法** (factorization) である．式 (2.17) を式 (2.9) のように射影的奥行きを用いて等号 = で表すと次のようになる．

$$\lambda_{\kappa\alpha}\boldsymbol{u}_{\kappa\alpha} = \boldsymbol{P}_\kappa\boldsymbol{X}_\alpha \tag{2.24}$$

$\lambda_{1\alpha}\boldsymbol{u}_{1\alpha}$, $\lambda_{2\alpha}\boldsymbol{u}_{2\alpha}$, ..., $\lambda_{M\alpha}\boldsymbol{u}_{M\alpha}$ を縦に並べてできる $3M$ 次元ベクトルを $\tilde{\boldsymbol{u}}_\alpha$ とし，$\boldsymbol{P}_1$, $\boldsymbol{P}_2$, ..., $\boldsymbol{P}_M$ のそれぞれの i 列を縦に並べてできる $3M$ 次元ベクトルを $\tilde{\boldsymbol{p}}_i$ とすると，式 (2.24) は次のように書ける．

$$\tilde{\boldsymbol{u}}_\alpha = X_\alpha^1\tilde{\boldsymbol{p}}_1 + X_\alpha^2\tilde{\boldsymbol{p}}_2 + X_\alpha^3\tilde{\boldsymbol{p}}_3 + X_\alpha^4\tilde{\boldsymbol{p}}_4 \tag{2.25}$$

ただし，X_α^i は $\boldsymbol{X}_\alpha$ の i 成分である．式 (2.25) は N 本のベクトル $\{\tilde{\boldsymbol{u}}_\alpha\}$ がすべて $\tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{p}}_2, \tilde{\boldsymbol{p}}_3, \tilde{\boldsymbol{p}}_4$ の張る $\mathcal{R}^{3M}$ の 4 次元部分空間 $\mathcal{L}$ に含まれることと意味する．これを**部分空間拘束条件** (subspace constraint) とよぶ．

式 (2.24) より射影的奥行き $\lambda_{\kappa\alpha}$ と同次座標 $\boldsymbol{X}_\alpha$ に任意の非零の定数 c_α を掛けてもよい．このとき各 $\tilde{\boldsymbol{u}}_\alpha$ が c_α 倍される．この不定性を除くために $\tilde{\boldsymbol{u}}_\alpha$ が単位ベクトルになるように正規化する．

以上より次の反復解法が得られる [3)]．まず射影的奥行き $\{\lambda_{\kappa\alpha}\}$ の初期値 $\{\lambda_{\kappa\alpha}^{(0)}\}$ を仮定し（1 としてよい），得られるベクトル $\{\tilde{\boldsymbol{u}}_\alpha\}$ に 4 次元空間 $\mathcal{L}$ を最小二乗法によって当てはめる（これに特異値分解を用いることから "因子分解法" という名称が生じた）．次に，各 $\tilde{\boldsymbol{u}}_\alpha$ から $\mathcal{L}$ までの二乗距離 J_α が最小になる $\{\lambda_{\kappa\alpha}\}$ を計算する．二乗距離 J_α は $\{\lambda_{\kappa\alpha}\}$ の 2 次形式になるので，これを正規化条件

$$\|\tilde{\boldsymbol{u}}_\alpha\|^2 = \sum_{\kappa=1}^{M}\|\boldsymbol{u}_{\kappa\alpha}\|^2\lambda_{\kappa\alpha}^2 = 1 \tag{2.26}$$

のもとで最小にする解は一般固有値問題を解いて定まる．その解を $\{\lambda_{\kappa\alpha}^{(1)}\}$ として，収束するまで上記の計算を反復する．射影的奥行き $\{\lambda_{\kappa\alpha}\}$ が定まれば，当てはめた 4 次元空間 $\mathcal{L}$ の任意の基底を $\{\tilde{\boldsymbol{p}}_i\}$ とし，各 $\tilde{\boldsymbol{u}}_\alpha$ を式 (2.25) の形に最小二乗展開して $\boldsymbol{X}_\alpha$ を定める．

2-1-5 アフィンカメラ

カメラの透視投影は，同次座標を使えば式 (2.6) のように見かけ上線形な関係式に書けるが，式 (2.1) から分かるように本質的に非線形な関係であり，解析を難しくしている．これを次のように線形な関係式で近似したものを**アフィンカメラ** (affine camera) とよぶ．

$$\begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{\Pi} \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} + \boldsymbol{\pi} \tag{2.27}$$

ただし，$\mathbf{\Pi}$ は 2×3 行列，$\boldsymbol{\pi}$ は 2 次元ベクトルであり，(X_c, Y_c, Z_c) は撮影している点のカメラ座標系に関する位置である．撮影対象がワールド座標系の原点の近傍に局在し，$\|\boldsymbol{t}\|$ が撮影対象の大きさに比較してある程度大きければ式 (2.27) のアフィンカメラモデルがよい精度で成り立つ．

式 (2.27) 中の $\mathbf{\Pi}$, $\boldsymbol{\pi}$ は運動パラメータ $\{\boldsymbol{R}, \boldsymbol{t}\}$ の関数であるが，式 (2.1) の透視投影を近似するように次の仮定をおく．

(i) 平面 $Z_c = t_z$ 上の点は式 (2.1) による透視投影と同じ像を生じる．

(ii) カメラの撮像は Z_c 軸の周りに等方である．

(iii) $\mathbf{\Pi}$, $\boldsymbol{\pi}$ は $\boldsymbol{R}$ に依存しない．

条件 (i) は撮影対象がワールド座標の原点を通り，X_cY_c 面に平行な平面であれば，透視投影と同じ像を生じるという要請であり，撮影対象がワールド座標の原点の近傍に局在しているという仮定に相当する．条件 (ii) は撮影対象を光軸の周りに角度 ϕ だけ回転しても，カメラを光軸の周りに角度 $-\phi$ だけ回転しても同じ像が得られるという仮定である．条件 (iii) は，ワールド座標系の向きは任意であり，撮影の関係がそのような任意性に影響されないという要請である．条件 (i), (ii), (iii) のもとでは式 (2.27) は次の形でなければならないことが証明される [2]．

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\zeta}\left(\begin{pmatrix} X_c \\ Y_c \end{pmatrix} + \beta(t_z - Z_c)\begin{pmatrix} t_x \\ t_y \end{pmatrix}\right) \tag{2.28}$$

ただし，t_x, t_y, t_z は $\boldsymbol{t}$ の各成分であり，ζ, β は $\sqrt{t_x^2 + t_y^2}$ および t_z の任意の関数である．ζ は対象の投影像の大きさを定め，β は対象の点が平面 $Z_c = t_z$ から離れたときの影響を定めるパラメータである．代表的な選び方は次のものである [4),5)]．

- **平行投影** (**直交射影**，**正射影**ともいう)(orthographic projection)

$$\zeta = 1, \qquad \beta = 0 \tag{2.29}$$

- **弱透視投影** (weak perspective projection) または**縮小平行投影** (scaled orthographic projection)

$$\zeta = \frac{t_z}{f_c}, \qquad \beta = 0 \tag{2.30}$$

- **疑似透視投影** (paraperspective projection)

$$\zeta = \frac{t_z}{f_c}, \qquad \beta = \frac{1}{t_z} \tag{2.31}$$

2-1-6 アフィンカメラによる 3 次元復元

シーン中の点を同次座標によるベクトル $\boldsymbol{X}_\alpha$(第 4 成分は 1) で表し，その第 κ 画像上の投影像を同次座標によるベクトル $\boldsymbol{u}_{\kappa\alpha}$(第 3 成分は 1) で表すと，式 (2.5), (2.27) から次式を得る．

$$\boldsymbol{u}_{\kappa\alpha} = \begin{pmatrix} \mathbf{\Pi}_\kappa \boldsymbol{R}_\kappa & \mathbf{\Pi}_\kappa \boldsymbol{t}_\kappa + \boldsymbol{\pi}_\kappa \\ 0\ 0\ 0 & 1 \end{pmatrix} \boldsymbol{X}_\alpha \tag{2.32}$$

ただし，$\mathbf{\Pi}_\kappa$, $\boldsymbol{\pi}_\kappa$ は第 κ カメラにおける式 (2.27) の $\mathbf{\Pi}$, $\boldsymbol{\pi}$ の値であり，$\{\boldsymbol{R}_\kappa, \boldsymbol{t}_\kappa\}$ は第 κ カメラの運動パラメータである．

式 (2.32) より，アフィンカメラとは式 (2.24) において，(i) 投影行列 $\boldsymbol{P}_\kappa$ の第 3 行が (0 0 0 1)，(ii) 射影的奥行き $\lambda_{\kappa\alpha}$ が 1，という場合に相当する．そして式 (2.25) に対応して次の関係を得る．

$$\tilde{\boldsymbol{u}}_\alpha = X_\alpha \tilde{\boldsymbol{p}}_1 + Y_\alpha \tilde{\boldsymbol{p}}_2 + Z_\alpha \tilde{\boldsymbol{p}}_3 + \tilde{\boldsymbol{p}}_4 \tag{2.33}$$

ここに $\tilde{\boldsymbol{u}}_\alpha$ は $\boldsymbol{u}_{1\alpha}$, $\boldsymbol{u}_{2\alpha}$, ..., $\boldsymbol{u}_{M\alpha}$ を縦に並べたベクトルであり，$\tilde{\boldsymbol{p}}_i$ は式 (2.32) の右辺の行列の i 列を $\kappa = 1, ..., M$ に対して縦に並べたベクトルである．各 $\boldsymbol{u}_{\kappa\alpha}$ の第 3 成分の 1 に対する部分は意味がないので除去すると，$\tilde{\boldsymbol{u}}_\alpha$ と各 $\tilde{\boldsymbol{p}}_i$ はすべて $2M$ 次元ベクトルとなる．

式 (2.33) はベクトル $\{\tilde{\boldsymbol{u}}_\alpha\}$ が $\mathcal{R}^{2M}$ の $\tilde{\boldsymbol{p}}_4$ を通り，$\tilde{\boldsymbol{p}}_1$, $\tilde{\boldsymbol{p}}_2$, $\tilde{\boldsymbol{p}}_3$ の張る 3 次元アフィン空間に拘束されることを意味する．これを**アフィン空間拘束条件** (affine space constraint) とよぶ．

これから 3 次元復元を行う方法も**因子分解法** (factorization method) とよばれる [4),5)]．まず $\{\tilde{\boldsymbol{u}}_\alpha\}$ の重心を $\tilde{\boldsymbol{p}}_4$ とし，$\{\tilde{\boldsymbol{u}}_\alpha\}$ に $\tilde{\boldsymbol{p}}_4$ を通る 3 次元アフィン空間 $\mathcal{A}$ を最小二乗法で当てはめ（これに特異値分解が用いたことから “因子分解法” という名称が生じた），それを張る正規直交基底を $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ とする．これを $\tilde{\boldsymbol{p}}_1$, $\tilde{\boldsymbol{p}}_2$, $\tilde{\boldsymbol{p}}_3$ と同一視して各 $\tilde{\boldsymbol{u}}_\alpha$ を式 (2.33) の最小二乗展開を行うことによって $(X_\alpha, Y_\alpha, Z_\alpha)$ が求まる．

しかし，これは真の形状にあるアフィン変換を施したもの (これを**アフィン復元** (affine reconstruction) とよぶ) である．これをユークリッド復元に高めるために，次のような基底の変換を行う．

$$\tilde{\boldsymbol{p}}_i = \sum_{j=1}^{3} A_{ji}\boldsymbol{q}_j \tag{2.34}$$

変換行列 $\boldsymbol{A} = (A_{ij})$ を定めるには，各 $\tilde{\boldsymbol{p}}_i$ の成分が二つごとに，ワールド座標基底を式 (2.32) のアフィンカメラで撮影したものであるという条件を用いる．これを**計量条件** (metric condition) とよぶ[4),5)]．これは

$$\boldsymbol{T} = \boldsymbol{A}\boldsymbol{A}^\top \tag{2.35}$$

とおき，式 (2.32) による投影関係から関係 $\boldsymbol{R}_\kappa\boldsymbol{R}_\kappa^\top = \boldsymbol{I}$ を用いて $\boldsymbol{R}_\kappa$ を消去すると次のように書ける．

$$\boldsymbol{Q}_\kappa^{\dagger\top}\boldsymbol{T}\boldsymbol{Q}_\kappa^\dagger = \boldsymbol{\Pi}_\kappa\boldsymbol{\Pi}_\kappa^\top \tag{2.36}$$

ただし，$\boldsymbol{Q}_\kappa^\dagger$ は次の 3×2 行列である．

$$\boldsymbol{Q}_\kappa^\dagger = \begin{pmatrix} \boldsymbol{q}_{\kappa(1)}^\dagger & \boldsymbol{q}_{\kappa(2)}^\dagger \end{pmatrix} \tag{2.37}$$

ここに，$\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ を列とする $2M \times 3$ 行列を $\boldsymbol{Q}$ とするとき，3 次元ベクトル $\boldsymbol{q}_{\kappa(1)}^\dagger$, $\boldsymbol{q}_{\kappa(2)}^\dagger$ はそれぞれ $\boldsymbol{Q}^\top$ の $(2\kappa-1)$ 列，2κ 列である．

以下，射影復元をユークリッド復元の高める場合の $\boldsymbol{\omega}^*$ の計算と同様に，仮定するカメラモデルの拘束を利用して式 (2.36) から $\boldsymbol{T}$ に関する方程式を導き，それを解いて $\boldsymbol{T}$ を定める．これを式 (2.35) の形に分解して変換行列 $\boldsymbol{A}$ が定まる．ただし，得られる 3 次元形状にはスケールと鏡像の不定性が残る．

この方法は反復なしに線形計算のみで解が求まるため，運動 1¡物体の認識や識別のような復元形状に高精度を要求しない応用によく用いられる．また，前節の方法と組合せれば，ユークリッド復元のための式 (2.22) に用いる $\tilde{\boldsymbol{P}}_\kappa$ の真の解 $\boldsymbol{P}_\kappa$ により近い値を計算することもできる．

［金谷　健一］

参 考 文 献

[2.1]
1) R. Hartley - A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Univ. Cambridge, U.K., 2000.
2) K. Kanatani - Y. Sugaya - H. Ackermann, Uncalibrated Factorization Using a Variable Symmetric Affine Camera, *IEICE Trans. Inf. & Syst.*, **E90-D**-5, 851–858, 2007.
3) S. Mahamud - M. Hebert, Iterative projective reconstruction from multiple views, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, **2**, 430–437, Hilton Head Island, SC, U.S.A, 2000.
4) C. J. Poelman - T. Kanade, A Paraperspective Factorization Method for Shape and Motion Recovery, *IEEE Trans. Patt. Anal. Mach. Intell.*, **19**-3, 206–218, 1997.
5) C. Tomasi - T. Kanade, Shape and Motion from Image Streams under Orthography—A Factorization Method, *Int. J. Comput. Vision*, **9**-2, 137–154, 1992.
6) B. Triggs, Autocalibration and the Absolute Quadric, *Proc. IEEE Conf. Compt. Vision Patt. Recog.*, 609–614, San Juan, Puerto Rico, 1997.