

## 2群(画像・音・言語) - 2編(パターン認識とビジョン)

## 3章 3次元幾何解析

(執筆者: 佐藤 淳)[2010年9月受領]

**概要**

画像情報をもとに3次元情報を獲得することを、3次元復元と呼ぶが、これは、コンピュータビジョンにおける最も重要な基本問題の一つとして考えられ、今日までに多くの研究がなされている。3次元復元には、幾何情報に基づく復元法と、輝度情報に基づく復元法とがある。本章では、特に、幾何情報に基づく3次元復元法と、これに関連する内容に関して解説する。

3次元空間中の物体は、カメラによって2次元画像に投影されるが、このときのカメラにおける投影は、レンズ中心を集光点とする透視投影によってモデル化することができる。透視投影は非線形の投影モデルであるが、1990年代よりコンピュータビジョンに射影幾何が導入され、これを線形に扱うことが可能になった。以降、幾何情報に基づく3次元復元の研究分野は急速に発展し、今日に至るまでに数多くの知見が得られている。

幾何情報に基づく3次元復元法の基本は、2台のカメラから得られた画像情報より3次元復元する両眼ステレオ法である。一方、今日では、コンピュータの処理能力が向上したことなどから、多くのカメラを用いることで、より高精度かつより広範囲な3次元復元が行われるようになってきた。このような多くのカメラに基づく復元法は多眼ステレオなどと呼ばれる。

両眼ステレオや多眼ステレオでは、これらのカメラを如何にしてキャリブレーション(校正)するかが重要な課題となる。このため、複数のカメラ間において成り立つ幾何関係である多視点幾何やカメラキャリブレーション法に関して、数多くの研究が集中的に行われ、2000年までにこれらに関する基本理論がほぼ完成の域に達した。今日では、この多視点幾何を、更に一般的なコンピュータビジョン問題へと拡張する研究や、全方位カメラなどのより一般的なカメラにおけるキャリブレーション法などの研究が進められている。

一方、3次元点のデータが得られると、次に、これらの点群データからいかにして意味のある形状情報を獲得するかが重要となる。このために、3次元点群をメッシュ情報へと変換したり、3次元点群に対してパラメトリックモデルの当てはめなどが行われる。また、一般には一度に3次元物体全体の3次元情報を得ることは難しく部分的な形状情報しか得られないことから、これらをいかに統合し全体形状を生成するかが重要な課題となる。

本章では、これらの3次元幾何解析を行ううえで重要な基本理論や技術に関して解説を行う。

**【本章の構成】**

3-1節では、3次元幾何解析を行ううえで必要な数学的基礎について解説する。3-2節では、様々なカメラのキャリブレーション法について説明する。3-3節では、2眼ステレオや多眼ステレオ、EPI解析による3次元復元法に関して解説する。3-4節では、レンジセンサから得られるレンジデータの獲得法や処理法に関して解説する。

## 2群 - 2編 - 3章

## 3-1 数学的基礎

(執筆者：佐藤 淳)[2010年9月受領]

本節では、画像に基づく3次元幾何解析において必要となる数学的基礎について述べる。

## 3-1-1 射影

まずはじめに、3次元空間中の点から2次元画像への投影について述べるが、このとき用いる座標の表記には、次に示す斉次座標(同次座標)(homogeneous coordinates)<sup>1)</sup>を用いる。

2次元平面上の点の斉次座標とは、この点の2次元座標  $[x, y]^T$  との間に  $x = x^1/x^3, y = x^2/x^3$  なる関係が成り立つような3次元ベクトル  $[x^1, x^2, x^3]^T$  である。同様に、3次元空間中の点の斉次座標とは、この点の3次元座標  $[X, Y, Z]^T$  との間に  $X = X^1/X^4, Y = X^2/X^4, Z = X^3/X^4$  なる関係が成り立つような4次元ベクトル  $[X^1, X^2, X^3, X^4]^T$  である。

今、3次元空間中の点  $[X, Y, Z]^T$  が2次元画像上の点  $[x, y]^T$  へと透視投影によって投影されたとする。すると、この投影は、3次元空間中の点の斉次座標  $\mathbf{X} = [X^1, X^2, X^3, X^4]^T$  と2次元画像上の点の斉次座標  $\mathbf{x} = [x^1, x^2, x^3]^T$  を用いて以下のように表すことができる。

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (3.1)$$

ここで  $\mathbf{P}$  は  $3 \times 4$  のカメラ行列であり、ワールド座標系におけるカメラの位置ベクトル  $\mathbf{C} = [X_c, Y_c, Z_c]^T$  と姿勢を表す  $3 \times 3$  の回転行列  $\mathbf{R}_c$  を用いて以下のように表せる。

$$\mathbf{P} = \mathbf{A} \begin{bmatrix} \mathbf{R}_c^{-1} & -\mathbf{R}_c^{-1}\mathbf{C} \end{bmatrix} \quad (3.2)$$

行列  $\mathbf{A}$  は、カメラ画像の水平方向及び垂直方向のスケールファクタ  $a_u, a_v$ 、画像中心  $u_0, v_0$ 、及び、せん断係数  $s$  よりなる次に示す  $3 \times 3$  行列である。

$$\mathbf{A} = \begin{bmatrix} a_u & s & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.3)$$

行列  $\mathbf{A}$  の各要素はカメラの内部パラメータ (intrinsic parameter) と呼ばれる。

ワールド座標系において位置  $\mathbf{C}$  と姿勢  $\mathbf{R}_c$  をもつカメラに点  $\mathbf{X}$  を投影することは、カメラの視点を基準に考えれば、物体を  $\mathbf{R}_c^{-1}$  だけ回転し、その後  $-\mathbf{R}_c^{-1}\mathbf{C}$  だけ並進して投影することに等しい。したがって、式(3.2)のカメラ行列  $\mathbf{P}$  は以下のように表すこともできる。

$$\mathbf{P} = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \quad (3.4)$$

ただし、 $\mathbf{R} = \mathbf{R}_c^{-1}$  であり、 $\mathbf{T} = -\mathbf{R}_c^{-1}\mathbf{C}$  である。以下ではこのような  $\mathbf{R}$  と  $\mathbf{T}$  を、それぞれカメラの回転と並進と呼ぶことにする。カメラの回転  $\mathbf{R}$  と並進  $\mathbf{T}$  は、カメラの外部パラメータ (extrinsic parameter) と呼ばれる。

## 3-1-2 エピポーラ幾何とF行列

次に、2台のカメラが存在する場合において、これらのカメラ間において成り立つ幾何学

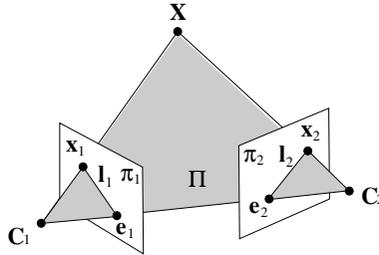


図 3・1 2 視点幾何

的な関係に関して述べる .

今, 図 3・1 に示すように, カメラが 2 台存在し, これらのカメラに 3 次元空間中の点  $X$  が, それぞれ  $x_1, x_2$  として投影されたとする . このとき, 点  $X$  と 2 台のカメラの視点  $C_1, C_2$  の 3 点を通る平面  $\Pi$  をエビポーラ面 (epipolar plane) と呼ぶ . また, 二つの視点  $C_1, C_2$  を結ぶ直線と 2 台のカメラの画像面  $\pi_1, \pi_2$  との交点  $e_1, e_2$  をエビポール (epipole) と呼ぶ . エビポールは, 一方のカメラ視点をもう一方のカメラに投影した投影像であると解釈することができる . 一方, エビポーラ面  $\Pi$  と画像面  $\pi_1, \pi_2$  との交差は直線となるが, これらの直線  $l_1, l_2$  をエビポーラ線 (epipolar line) と呼ぶ . 二つの画像中の対応点  $x_1, x_2$  はこれらのエビポーラ線上に存在する .

第 1 カメラのカメラ座標系を基準に考えれば, これら 2 台のカメラのカメラ行列  $P_1, P_2$  は次のように表せる .

$$P_1 = A_1 \begin{bmatrix} I & 0 \end{bmatrix} \quad P_2 = A_2 \begin{bmatrix} R & T \end{bmatrix} \quad (3\cdot5)$$

$R$  と  $T$  は第 1 カメラに対する第 2 カメラの相対的な回転と並進である . また,  $A_1, A_2$  は, それぞれのカメラの内部パラメータである .

3 次元空間中の点  $X$  がこれら 2 台のカメラに  $x_1, x_2$  として投影されたとする . すると, この投影は以下のように表すことができる .

$$x_1 = P_1 X \quad x_2 = P_2 X \quad (3\cdot6)$$

このとき, これらの投影像の間には次に示す関係が存在する . これをエビポーラ方程式 (epipolar equation), あるいは 2 重線形拘束 (bilinear constraints) と呼ぶ<sup>1, 2, 3)</sup> .

$$x_2^T F x_1 = 0 \quad (3\cdot7)$$

式 (3・7) の  $3 \times 3$  行列  $F$  は基礎行列 (fundamental matrix) と呼ばれ, 次に示すように, 二つのカメラ間の相対的な回転  $R$ , 並進  $T$  と, 内部パラメータ  $A_1, A_2$  よりなる .

$$F = A_2^{-T} [T]_{\times} R A_1^{-1} \quad (3\cdot8)$$

ただし,  $M^{-T}$  は行列  $M$  の転置の逆行列 (あるいは逆行列の転置) を表す . また,  $[V]_{\times}$  は, 3 次元ベクトル  $V = [V_x, V_y, V_z]^T$  の三つの要素よりなる以下に示す歪対称行列である .

$$[\mathbf{V}]_x = \begin{bmatrix} 0 & -V_z & V_y \\ V_z & 0 & -V_x \\ -V_y & V_x & 0 \end{bmatrix} \quad (3\cdot9)$$

基礎行列  $\mathbf{F}$  は  $3 \times 3$  行列であり 9 要素よりなるが、定数倍の不定性をもつこと、及びランクが 2 であり  $\det \mathbf{F} = 0$  であることから、その自由度は 7 である。

$\mathbf{F}$  は、式 (3・7) を用いることで、二つの画像中の 8 組以上の対応点組から線形計算によって求めることができる<sup>2)</sup>。また、非線形解法を用いれば、画像中の 7 組以上の対応点組から計算することもできる<sup>3)</sup>。

$\mathbf{F}$  が得られると、以下の式により、一方の画像中の点  $\mathbf{x}_1$  (または  $\mathbf{x}_2$ ) より、もう一方の画像中のエピポーラ線  $\mathbf{l}_2$  (または  $\mathbf{l}_1$ ) を求めたり、エピポール  $\mathbf{e}_1, \mathbf{e}_2$  を求めることができる。

$$\mathbf{l}_2 = \mathbf{F}\mathbf{x}_1 \quad \mathbf{l}_1 = \mathbf{F}^T\mathbf{x}_2 \quad \mathbf{F}\mathbf{e}_1 = 0 \quad \mathbf{F}^T\mathbf{e}_2 = 0 \quad (3\cdot10)$$

図 3・1 に示すように、エピポーラ線は一方の画像中の点に対するもう一方の画像中の対応点の候補の集合である。このため、エピポーラ線を用いることで、2 次元画像におけるステレオ対応点探索を、1 次元探索によって行うことが可能となる (3-3 節参照)。

また、 $\mathbf{F}$  からは、次式に示すように、第 1 カメラを基準とした射影的な不定性をもつ 2 台のカメラのカメラ行列を得ることができる。

$$\mathbf{P}_1 = [\mathbf{I} \quad \mathbf{0}] \quad \mathbf{P}_2 = [[\mathbf{e}_2]_x \mathbf{F} \quad \mathbf{e}_2] \quad (3\cdot11)$$

したがって、 $\mathbf{F}$  を求めることは、2 台のカメラを射影的な不定性を残してキャリブレーションすることに等しい。

一方、内部パラメータが既知である場合には、 $\mathbf{F}$  から次に示す行列  $\mathbf{E}$  を得ることができる。

$$\mathbf{E} = \mathbf{A}_2^T \mathbf{F} \mathbf{A}_1 = [\mathbf{T}]_x \mathbf{R} \quad (3\cdot12)$$

この行列  $\mathbf{E}$  を基本行列 (essential matrix) と呼ぶ。式 (3・12) より明らかなように、 $\mathbf{E}$  は回転  $\mathbf{R}$  と並進  $\mathbf{T}$  よりなる  $3 \times 3$  行列である。したがって、 $\mathbf{E}$  を分解することにより、2 台のカメラ間の相対的な回転  $\mathbf{R}$  と並進  $\mathbf{T}$  を得ることができる。この分解は、特異値分解などにより行うことができる<sup>2)</sup>。

### 3-1-3 多視点幾何と多焦点テンソル

式 (3・7) のエピポーラ方程式は、テンソル表記を用いれば以下のように表すことができる<sup>2,4)</sup>。

$$x_1^i x_2^j F_{ji} = 0 \quad (3\cdot13)$$

ここで、 $F_{ji}$  は基礎行列  $\mathbf{F}$  の  $j$  行  $i$  列の要素を表し、 $x_1^i$  と  $x_2^j$  はそれぞれ  $\mathbf{x}_1$  の第  $i$  要素と  $\mathbf{x}_2$  の第  $j$  要素を表す。また、式 (3・13) では、テンソル表記におけるアインシュタインの規約に基づき、総和記号を省略して表記してある。このようにテンソル表記したとき、 $F_{ji}$  を 2 焦点テンソル (bifocal tensor) と呼ぶ。

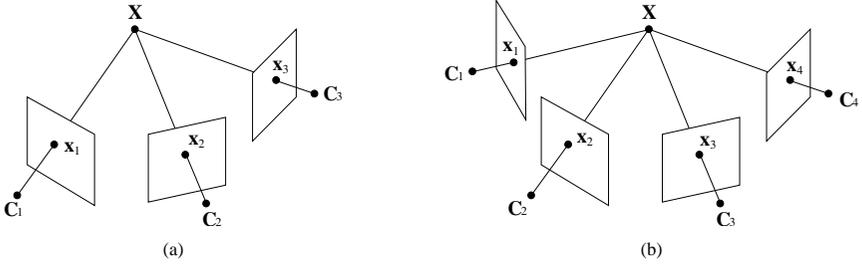


図 3・2 3 視点幾何と 4 視点幾何

次に，図 3・2(a) に示すように 3 台のカメラが存在する場合について考える．今，3 次元空間中の点  $X$  が，3 台のカメラの画像に  $x_i = [x_i^1, x_i^2, x_i^3]^T$  ( $i = 1, \dots, 3$ ) として投影されているとする．このとき，これらの対応点の関係は，テンソル表記を用いて以下のように表すことができる<sup>2, 4)</sup>．

$$x_1^j x_2^k x_3^l \epsilon_{rja} \epsilon_{skb} \mathcal{T}_i^{rs} = 0_{ab} \tag{3・14}$$

これを 3 重線形拘束 (trilinear constraints) と呼ぶ．ここで  $\epsilon_{ijk}$  は， $\{i, j, k\}$  の数値の並びが， $\{1, 2, 3\}$  に対して偶置換であれば 1 を，奇置換であれば  $-1$  を，それ以外であれば 0 の値を取る 3 階のテンソルである．式 (3・14) の  $\mathcal{T}_i^{rs}$  は，3 台のカメラ間の幾何学的な関係を表す 3 階のテンソルであり，3 焦点テンソル (trifocal tensor) と呼ばれる．3 焦点テンソル  $\mathcal{T}_i^{rs}$  は 27 個の要素をもつが，その幾何学的な自由度は 18 である． $\mathcal{T}_i^{rs}$  からは，3 台のカメラのカメラ行列を射影変換の不定性を残して求めることができる<sup>2)</sup>．したがって，3 焦点テンソルを求めることは，3 台のカメラを射影変換の不定性を残してキャリブレーションすることに等しい．

また，図 3・2 (b) に示すように，4 台のカメラが存在する場合には，これら 4 台のカメラの投影像  $x_i = [x_i^1, x_i^2, x_i^3]^T$  ( $i = 1, \dots, 4$ ) の間において以下の関係式が成り立つ<sup>2, 4)</sup>．

$$x_1^j x_2^k x_3^l x_4^m \epsilon_{pia} \epsilon_{qjb} \epsilon_{rkc} \epsilon_{sld} Q^{pqrs} = 0_{abcd} \tag{3・15}$$

これを 4 重線形拘束 (quadrilinear constraints) と呼ぶ．式 (3・15) の  $Q^{pqrs}$  は，4 台のカメラ間の幾何学的な関係を表す 4 階のテンソルであり，4 焦点テンソル (quadrifocal tensor) と呼ばれる．4 焦点テンソル  $Q^{pqrs}$  は 81 個の要素をもつが，その幾何学的な自由度は 29 である． $Q^{pqrs}$  からは，4 台のカメラのカメラ行列が射影変換の不定性を残して求まる．したがって，4 焦点テンソルを求めることは，4 台のカメラを射影変換の不定性を残してキャリブレーションすることに等しい．

$F_{ji}, \mathcal{T}_i^{rs}, Q^{pqrs}$  は，まとめて多焦点テンソル (multifocal tensor) と呼ばれる．これらのテンソルは，それぞれ，式 (3・13)，式 (3・14)，式 (3・15) を用いることで，画像中の対応点より線形解法によって求めることができる．高精度に求めたい場合には，テンソルを最小パラメータ (2 焦点テンソルは 7 個，3 焦点テンソルは 18 個，4 焦点テンソルは 29 個) で表現し，非線形解法によって求める<sup>2, 3)</sup>．また，RANSAC<sup>5)</sup>などのロバスト法と組み合わせることで，誤対応などの外れ値を除去しつつ多焦点テンソルを求めることもできる．

## 2群 - 2編 - 3章

## 3-2 カメラキャリブレーション

(執筆者: 佐藤 淳)[2010年9月受領]

カメラの内部パラメータや外部パラメータを求めることをカメラのキャリブレーション (camera calibration), あるいは校正と呼ぶ。カメラのキャリブレーションには, 大きさや形状が既知である参照物体を用いる方法と, 一般の未知の物体からキャリブレーションする方法とがある。未知の物体からキャリブレーションする方法は, セルフキャリブレーション (self calibration, auto calibration) あるいは自校正と呼ばれる。

## 3-2-1 既知の物体を用いたキャリブレーション

大きさや形状が既知な参照物体がある場合には, 参照物体上にワールド座標系をすることで, 3次元点の座標  $X$  と, これを投影した画像点の座標  $x$  の組が得られる。一方, 式(3・1)より, カメラ行列  $P$  と 3次元点  $X$  と画像点  $x$  との間には次式の関係があることが分かる。

$$[x]_{\times} P X = 0 \quad (3 \cdot 16)$$

$X$  と  $x$  が既知であることから, 式(3・16)よりカメラ行列  $P$  を求めることができる。式(3・16)は3式よりなるがこの中で独立な式は2式のみである。行列  $P$  の自由度は11であることから, 6組以上の3次元点  $X$  と画像点  $x$  の組があれば, カメラ行列  $P$  が求まる。

カメラ行列  $P$  は, 式(3・4)に示すように, 内部パラメータ  $A$  と外部パラメータ  $R, T$  よりなることから, ひとたび  $P$  が求まれば, これを分解することで, カメラの内部パラメータと外部パラメータを求めることができる。この分解には, コレスキー分解<sup>2)</sup>や  $R$  の直交性に基づく分解法などが用いられる<sup>1)</sup>。

また, 一般の3次元点の代わりに, 平面上において座標が既知な点を用意し, この平面の位置姿勢を変えて撮影した画像よりカメラをキャリブレーションする実用的な方法も提案されている。この方法は Zhang の方法<sup>6)</sup>と呼ばれ, 今日では標準的に用いられている。

## 3-2-2 セルフキャリブレーション

セルフキャリブレーションは, 既知な参照物体を用いずに, 一般の3次元シーンを撮影した画像情報のみから, カメラをキャリブレーションする方法である。キャリブレーションに使用する画像中の点の3次元座標を知る必要がない点が大きな特徴である。

画像情報のみが与えられている場合には, 3-1節で述べたように, 射影変換の不定性を残したカメラ行列を得ることができる。セルフキャリブレーションは, 何らかの制約を加えることにより, この射影変換の不定性を取り除き, 射影歪みのないカメラ行列を求めるものである。

一般に, 3次元射影変換は15自由度をもつ。キャリブレーションにおける射影歪みの除去とは, 射影的な不定性からスケール付ユークリッド変換の不定性にまで不定性を取り除くことである。このとき, スケール付ユークリッド変換は, 回転3自由度, 並進3自由度, スケール1自由度の合計7自由度をもつ。したがって  $15 - 7 = 8$  個以上の拘束が画像より得られれば, 射影変換の不定性からスケール付ユークリッド変換の不定性へと変換でき, カメラをキャリブレーションすることができる。

今,  $n$  台のカメラがあり, それぞれのカメラにおいて既知の内部パラメータが  $N_k$  個あり, 未知ではあるがカメラ間で同一な内部パラメータが  $N_f$  個あったとする. すると, 既知の内部パラメータからは  $n \times N_k$  個の拘束が得られ, 未知ではあるがカメラ間で同一な内部パラメータからは  $(n-1) \times N_f$  個の拘束が得られる. したがって, 以下の条件を満たすとき, 射影変換に関して 8 個以上の拘束式が得られ, カメラの内部パラメータと外部パラメータを求めることができる<sup>7)</sup>.

$$n \times N_k + (n-1) \times N_f \geq 8 \quad (3 \cdot 17)$$

例えば,  $n$  台のカメラにおいて, 内部パラメータの内のせん断係数  $s$  が 0 であると仮定でき, それ以外のパラメータは未知であり, かつカメラ間で異なる場合には,  $N_k = 1, N_f = 0$  である. したがって,  $n \geq 8$ , すなわち 8 台以上のカメラがあれば, これらのカメラの射影的な不定性を取り除くことができ, キャリブレーションすることができる. また, 各カメラの内部パラメータがすべて未知ではあるがカメラ間で同一である場合には,  $N_k = 0, N_f = 5$  であり,  $5(n-1) \geq 8$ , すなわち 3 台以上のカメラがあれば, これらのカメラをキャリブレーションすることができる.

セルフキャリブレーションを最も高精度に行う方法は, カメラの内部パラメータ  $\mathbf{A}$ , 外部パラメータ  $\mathbf{R}, \mathbf{T}$  と画像点を復元した 3 次元点  $\mathbf{X}$  とを, 再投影誤差を最小化するように同時推定するものであり, これをバンドルアジャストメント (bundle adjustment) と呼ぶ<sup>8)</sup>.  $N_c$  個のカメラに  $N_p$  個の 3 次元点が投影されているとき, バンドルアジャストメントでは, 以下により,  $\mathbf{A}_i, \mathbf{R}_i, \mathbf{T}_i$  ( $i = 1, \dots, N_c$ ) と 3 次元点  $\mathbf{X}_j$  ( $j = 1, \dots, N_p$ ) とを同時推定する<sup>9)</sup>.

$$\min_{\mathbf{A}_i, \mathbf{R}_i, \mathbf{T}_i, \mathbf{X}_j} \sum_{i=1}^{N_c} \sum_{j=1}^{N_p} d(\mathbf{x}_{ij}, \hat{\mathbf{x}}_{ij})^2 \quad (3 \cdot 18)$$

ここで,  $\mathbf{x}_{ij}$  は  $i$  番目のカメラにおける  $j$  番目の 3 次元点の実際の投影像であり,  $\hat{\mathbf{x}}_{ij}$  は, 推定した  $i$  番目のカメラパラメータによって推定した  $j$  番目の 3 次元点を再投影した投影像である. また,  $d(\mathbf{x}, \mathbf{y})$  は, 画像中における 2 点  $\mathbf{x}, \mathbf{y}$  の間の距離を表す.

バンドルアジャストメントは最高精度の結果が得られるが, 多くのパラメータの同時推定を行うことから, 計算コストが大きく, また良い初期値が必要である. そこで, 線形解法により初期値を計算し, これをもとにバンドルアジャストメントで高精度化するのが一般的である. 線形解法としては, 未知の内部パラメータの数を制限することで線形化する方法<sup>7)</sup>などが知られている.

### 3-2-3 レンズ歪みのキャリブレーション

以上のキャリブレーション法では, 画像中の歪みが内部パラメータ行列  $\mathbf{A}$  で記述できることを仮定している. 一方, 広角カメラなどを使用した場合には, 空間中の直線が画像上において曲線に投影されるレンズ歪み (lens distortion) が無視できないほど大きくなる. このような場合には, レンズ歪みを事前に取り除いたり, 内部パラメータと同時にレンズ歪みを推定することにより除去する<sup>6, 11)</sup>.

レンズ歪み  $\Delta \mathbf{x}$  は, 通常, 歪み中心  $\mathbf{x}_0$  から放射状に生じるラジアル歪み (radial distortion)

として考え、これを次式のように多項式により近似的に表現する<sup>6, 10)</sup>。

$$\Delta \mathbf{x} = \bar{\mathbf{x}}(k_1 r^2 + k_2 r^4 + \cdots + k_m r^{2m}) \quad (3 \cdot 19)$$

ここで、 $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{x}_0$ 、 $r = \|\bar{\mathbf{x}}\|$  であり、また次数  $m$  は一般に  $m = 2$  程度で十分とされている。

レンズ歪みの推定は、歪み中心  $\mathbf{x}_0$ 、及びラジアル歪み係数  $k_i$  ( $i = 1, \dots, m$ ) を 3 次元座標が既知の点や 2 次元座標が既知の平面パターンを用いて求めることで行う。この計算は、座標が既知の点に関して、実際の投影像と得られた  $\mathbf{x}_0$ 、 $k_i$  を用いて投影した投影像との差を最小化するよう  $\mathbf{x}_0$ 、 $k_i$  を推定することにより行う。

## 2群 - 2編 - 3章

## 3-3 ステレオ

(執筆者：佐藤 智和)[2010年9月受領]

ステレオ視覚 (stereo vision) とは、人間の目のように異なる位置で獲得された 2 枚またはそれ以上の枚数の画像群 (ステレオ画像群) より得られる立体的な視覚情報を表す。また、ステレオ画像群に含まれる視差情報を抽出し、三角測量の原理によりシーン中に存在する物体の 3 次元位置を計測することをステレオ画像計測と呼ぶ。一般に、ステレオ画像計測では前節で紹介されたカメラキャリブレーションによって、複数台のカメラの位置関係やカメラの内部パラメータがあらかじめ校正されていることを前提としている。2 枚のステレオ画像に基づく 3 次元計測法は、人間が両眼によってシーンの奥行きを知覚する手段と同様であることから、両眼ステレオまたは 2 眼ステレオと呼ばれる。これに対して 3 台以上のカメラを用いる方法を多眼ステレオと呼ぶ。

## 3-3-1 両眼ステレオ

カメラの光軸が互いに平行な 2 台のカメラが、図 3-3 に示すように一定の距離  $b$  の間隔を置いて左右に設置され、各カメラの画像面が、それぞれのカメラの投影中心から  $f$  離れた 3 次元空間中の同一平面上にあるとする。このようなカメラ配置によるステレオ画像計測は、平行ステレオまたは標準ステレオと呼ばれ、比較的簡単な式によって注目点の 3 次元位置を推定でき、またエピポーラ線が画像上で水平となるため、後述する対応点探索における探索範囲の決定が容易であるという特長をもつ。なお、ステレオ撮像系における 2 台のカメラの投影中心間の線分を基線 (baseline)、その線分の長さ  $b$  を基線長と呼ぶ。

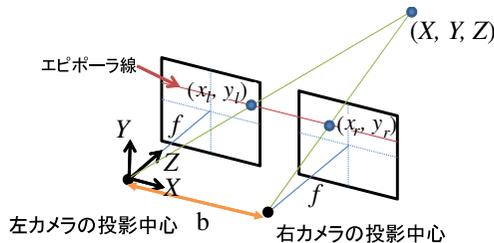


図 3-3 平行ステレオによる撮像系

ここでは、左に設置されたカメラのカメラ座標系における注目物体の座標を  $(X, Y, Z)$  とし、カメラの画像面上への注目物体の投影位置を左右のカメラの画像座標で表したものをそれぞれ  $(x_l, y_l)$ ,  $(x_r, y_r)$  とする。ただし、各カメラの画像座標の原点はカメラの光軸と画像面の交点とし、また画像面の  $x$  軸は左右のカメラの画像中心を結び直線と平行、 $y$  軸は  $x$  軸と垂直であるとする。このような画像座標系を設定した場合において、平行ステレオでは常に  $y_l = y_r$  が成り立ち、同一の 3 次元座標を表す対応点は画像中の同一の高さに存在する。このとき、左に設置されたカメラ座標系における注目点の 3 次元座標  $(X, Y, Z)$  は、以下のように求まる。

$$X = \frac{x_l b}{x_l - x_r}, Y = \frac{y_l b}{x_l - x_r}, Z = \frac{f b}{x_l - x_r} \quad (3.20)$$

ここで、 $Z$  を画素  $(x_l, y_l)$  に対する奥行き距離 (depth),  $d = x_l - x_r$  を視差 (disparity) と呼ぶ。なお、平行ステレオでない場合においても、各入力画像を平行ステレオと同等のカメラ配置で撮影した画像に変換するレクティフィケーションと呼ばれる手法を用いることで、式 (3・20) をそのまま用いることができる<sup>12)</sup>。式 (3・20) から分かるように、カメラの設定が固定されている条件下で  $(x_l, y_l)$  に対応する奥行き距離を推定する問題は、 $(x_l, y_l)$  に対する対応点の画像座標をエッジ線上で探索することによって視差  $d$  を決定する問題に帰着する。

このような問題は、ステレオ視覚における対応点探索問題と呼ばれ、対応点を決定するための手法として、注目画素ごとに対応点を決定する area-based と呼ばれる手法と、画像中の輝度エッジなどの特徴となる箇所に対して対応点を探索する feature-based と呼ばれる手法に大別できる。一般に、前者はすべての画素に対して奥行き距離を算出するため、密な奥行き情報を得ることができる。これに対して、後者は信頼度の高い奥行きを推定する代わりに、得られる奥行き情報が疎なものとなる。いずれの手法においても、対応点が必ずエッジ線上に存在するという制約を用いることで、1 次元探索によって視差を決定することができる。

視差  $d$  を決定するための対応点探索法として基礎的な方法は、SSD (Sum of Squared Differences), SAD (Sum of Absolute Distances), NCC (Normalized Cross Correlation) などの画像パターンの類似度評価指標によるブロックマッチングを用いる手法である。例えば平行ステレオを対象とした場合において、注目画素を中心とするサイズ  $(2m + 1) \times (2m + 1)$  のウィンドウを用いた SSD は以下のように定義される。

$$SSD(d) = \sum_{i=-m}^m \sum_{j=-m}^m |I_l(x_l + i, y_l + j) - I_r(x_l + d + i, y_l + j)|^2 \quad (3 \cdot 21)$$

ここで、 $I_l, I_r$  はそれぞれ左右のカメラの画像の輝度値を表しており、SSD は視差  $d$  の関数である。左カメラの画像上の各画素  $(x_l, y_l)$  に対して、SSD を最小とする視差  $d$  を順次決定していくことで、画像上のすべての画素に対応する奥行き距離を推定することができる。ただし、このような単純な方法による対応点探索には、単色の領域や多数の類似したテクスチャパターンが存在する領域において、誤対応が生じやすいという問題がある。このような問題に対して、対応点探索をエネルギー最小化問題としてとらえ、隣接する画素間において視差や奥行き距離がなめらかに変化することを仮定して、推定のロバスト性を高めるなどの工夫が用いられることが多い<sup>13, 14)</sup>。

### 3-3-2 多眼ステレオ

異なる地点に設置された 3 台以上のカメラを用いた場合においても、両眼ステレオと同様に、対応点を決定することで注目点に対応する 3 次元位置を推定できる。このような多眼ステレオ法の一つとして、複数のカメラから得られる画像群を統合的に扱い、それらの画像上において対応点を一意に決定することで注目点の 3 次元位置を計測するマルチベースラインステレオ法<sup>15)</sup>と呼ばれる手法が用いられる。マルチベースラインステレオ法では、多数の画像を同時に用いることで、両眼ステレオにおいて問題となっていた画像間の対応点探索におけるあいまいさを低減し、奥行き距離の推定におけるロバスト性を向上させる。

ここでは、図 3・4 に示すように、シーン中に存在する  $n$  台のカメラにより撮影された画像を用いるものとし、 $p$  番目 ( $1 \leq p \leq n$ ) のカメラ (以下、カメラ  $p$ ) の画像面上における画素

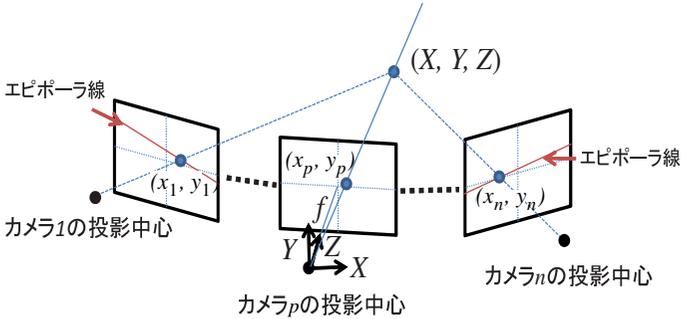


図 3・4 多眼ステレオによる撮像系

$(x_p, y_p)$  の 3 次元位置  $(X, Y, Z)$  を推定するものとする. 以下では, カメラ  $p$  の焦点距離を  $f$  とし, レンズ歪みなどは補正済みであるものとする. このとき, 画素  $(x_p, y_p)$  の奥行き距離を  $z$  とおけば, これに対応する 3 次元位置は, カメラ  $p$  のカメラ座標系において  $(x_p z / f, y_p z / f, z)$  と表せる. この 3 次元座標を  $k$  番目 ( $1 \leq k \leq n$ ) のカメラの画像面に投影した画像座標を  $(x_k, y_k)$  とする. 図 3・4 に示すように, 注目画素  $(x_p, y_p)$  に対応するカメラ  $k$  の画像上の位置  $(x_k, y_k)$  は, 両眼ステレオの場合と同様にエピポーラ線上に拘束される. マルチベースラインステレオ法では, 注目画素  $(x_p, y_p)$  とカメラ  $k$  の画像上の画素  $(x_k, y_k)$  周辺のパターンを用いて SSD を算出し, これらの総和 SSSD (Sum of SSD) を最小化することで  $(x_p, y_p)$  に対応する奥行き距離  $z$  を決定する.

$$SSSD(z) = \sum_{k=1}^n \sum_{i=-m}^m \sum_{j=-m}^m |I_p(x_p + i, y_p + j) - I_k(x_k + i, y_k + j)|^2 \quad (3 \cdot 22)$$

図 3・5 に示すように, 単一の画像対から得られる SSD を用いる場合には, 類似パターンなどの存在によって対応点の決定にあいまいさが生ずる場合が多いが, 異なる基線長からなる複数の画像対から得られる SSD を総和した SSSD を用いることで, 対応点探索におけるあいまいさを低減することができる.

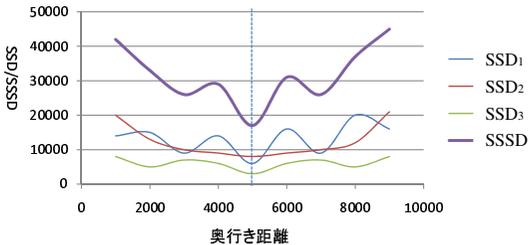


図 3・5 マルチベースラインステレオにおける類似度評価指標 SSSD

撮影対象となるシーンに動物体が存在しない場合には, 一台のカメラでシーンを移動撮影した動画像に対して多眼ステレオ法を適用することができる. この場合には, 4 章で紹介さ

れる動画画像解析による 3 次元復元手法などを用いて、動画像の各フレームに対するカメラパラメータをあらかじめ推定しておく必要がある。図 3・6 は、手持ちのカメラによって移動撮影した動画像から、マルチベースラインステレオ法を用いて各画素の奥行き距離を推定した例である<sup>16)</sup>。同図中右下の画像は、基準となる画像上の各画素に対して推定された奥行き距離を輝度値に変換したものであり、このような画像を距離画像または奥行き画像と呼ぶ。図 3・6 の例では、撮影対象である建物上に周期的なテクスチャパターンが存在するが、多数の画像を同時に用いることで、誤対応の発生が抑止されている。

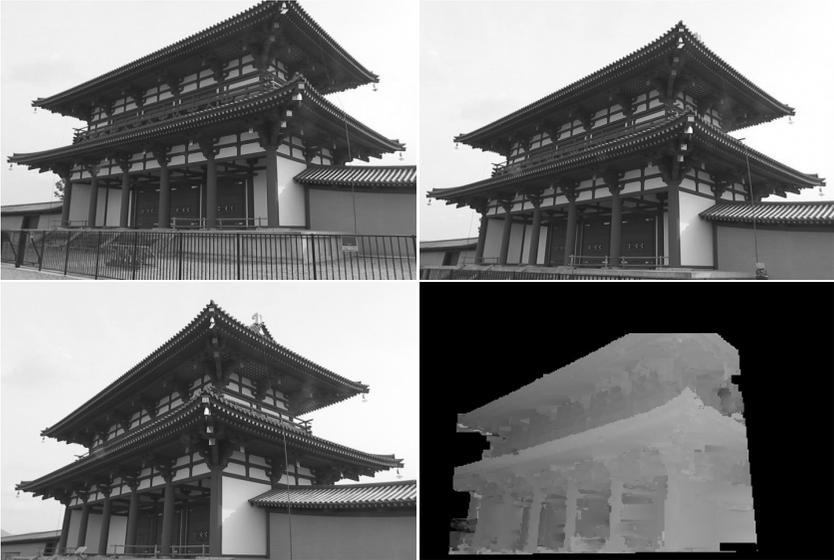


図 3・6 マルチベースラインステレオ法によって推定された距離画像と入力画像の一部

他方、比較的小物体の 3 次元計測を行う場合においては、その物体を周囲の全方向から撮影する場合が多く、このような状況下では物体上の同一の点であっても、その周辺のパターンは姿勢変化によって大きく変化してしまう。したがって、このような目的においては、SSSD を用いて対応点を決定することが難しい場合がある。これに対して、SSSD の代わりに 3 次元座標  $(X, Y, Z)$  の各画像面上への投影位置の輝度値の分散を評価尺度として用いる手法が利用される。このような輝度値の分散による尺度をフォトコンシステンシと呼ぶ<sup>17)</sup>。フォトコンシステンシを用いる方法は画像パターンを用いないため、カメラの移動による画像の変形や回転に影響されないという特長があるが、一方で対応点決定に用いる情報量が少ないため、誤対応が生じやすいという問題がある。したがって、フォトコンシステンシを評価尺度として用いる場合には、両眼ステレオの場合と同様に周囲の画素に対して推定される奥行き距離との連続性を考慮するか、または、シルエット法などのほかの 3 次元計測手法によって十分に良い初期形状が得られていることを前提として、対応点探索の範囲を限定する必要がある<sup>18)</sup>。

### 3-3-3 EPI

一定の制約条件の下で撮影された動画像を用いて EPI (Epipolar Plane Image) と呼ばれる画像を作成すれば、ステレオ画像計測における対応点探索の問題を比較的簡単に解決することができる<sup>19)</sup>。ここでは、カメラが等速直線運動をするものとし、平行ステレオの場合と同様に、カメラの光軸方向がカメラの移動方向と直交し、画像面上の横軸がカメラの移動方向と平行であるとすると、図 3・7(a) は、このような条件下で撮影された動画像の一部である。この動画像を用いて、図 3・7(b) に示すような、画像の横・縦軸 ( $x, y$ ) と時間軸  $t$  の 3 次元からなる時空間画像を作成し、その画像上で同一の高さに存在する画素群を抽出すれば、これらはすべて単一のエピポーラ面上に存在する。このような、エピポーラ面による時空間画像の断面像を、EPI と呼ぶ。図 3・7(c) は、同図 (b) の波線に対応する高さのエピポーラ面で、時空間画像を切断した EPI を表す。

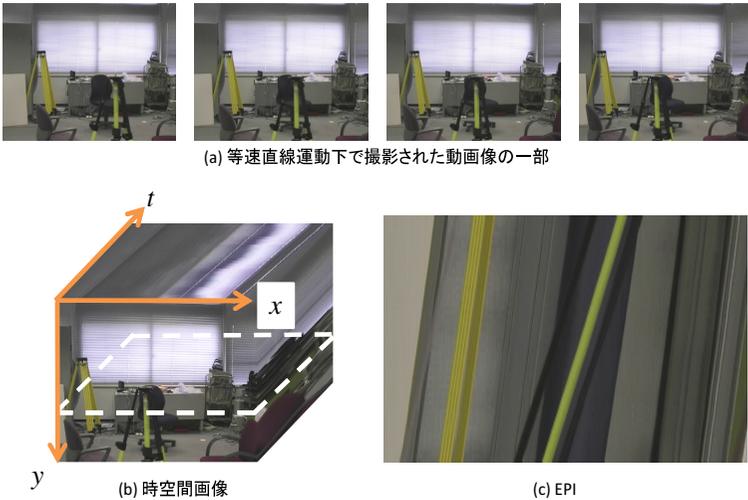


図 3・7 動画像から作成した時空間画像と EPI の例

等速直線運動によって撮影された EPI においては、シーン中に存在する物体上の点は、単一の直線として画像上に表れる。また、その直線の傾きは、カメラから物体上の点までの距離に応じて変化する。したがって、まず EPI に対して Hough 変換などを用いることによって、輝度エッジの直線を検出することで、比較的簡単に複数画像間における対応点を決定することができる。また、直線の傾きによって、その直線上に存在する物体とカメラの奥行き距離を算出することができる。

## 2群 - 2編 - 3章

## 3-4 レンジデータ

(執筆者: 増田 健)[2010年9月受領]

## 3-4-1 レンジデータの獲得

レンジデータ (range data) は, 物体の3次元形状計測結果として得られた, センサから物体表面までの距離の情報である. 3次元形状の表面は2次元多様体であり, 計測の際に画像センサや2次元走査を利用することが多いため, 一般に2次元配列の各要素  $(x, y)$  に3次元座標値  $(X(x, y), Y(x, y), Z(x, y))$  が対応する距離画像 (range image) として得られる<sup>20)</sup>.

レンジデータを計測する装置はレンジセンサ (range sensor) またはレンジファインダ (range finder) と呼ばれる. 計測方法は, 受動的な画像センサのみ使用した受動的手法と, 投射した光を利用する能動的手法に分類できる. 二眼から多眼のステレオや合焦点 (またはボケ) を利用した手法は代表的な受動的手法である.

能動的手法については, 投射した光の利用方法について様々な方式が提案されている<sup>21, 20, 22)</sup>. 二眼ステレオを構成する二つの画像センサの一方を投光装置に置き替え, どの方向に発した光が画像上で観測されたか対応づけができれば, ステレオと同様に三角測量の原理に基づき対象までの距離が求められる. 投光の方法としては, 例えばレーザビームを2次元走査することによって得られるが, 計測時間を短縮するために, スリット光を走査する光切断法や, 少ないフレーム数で方向を求めるために様々な空間コード化法が提案されている<sup>23, 22)</sup>. 主要な計測誤差の要因はステレオと同様に, 投光機と画像センサの解像度が有限であることに起因する量子化と, 想定外の反射光や復号誤りによる誤対応である. また, 画像センサと投光装置の位置が異なる必要があることから, 手前の物体の周囲に遮られて投光されず, 受光センサから見て, 計測できない箇所 (オクルージョン) が生ずるのも同様である.

光飛行時間 (TOF: Time-of-Flight) 法は, 投射した光が対象物の表面で反射されてセンサに戻ってくる時間から距離を測定する. 計測方法としてはレーザのパルスを利用する方法, 正弦波や擬似乱数などの時間変調をかけて位相を計測する方法がある. 距離画像を得るには, 点計測センサを2次元走査するか, シャッター時間を高精度に制御できる特殊な画像センサも用いられる. 距離の精度は時間の計測精度で決まる. 更に微細な光の波長程度の凹凸を計測するためには光の干渉が用いられる. フォトメトリックステレオ (photometric stereo) 法では, 複数の位置の光源を切り替えてその陰影から法線方向を推定することができ, 得られた法線方向を積分することにより奥行き値を得ることができる.

光学的手法による非接触式の3次元形状計測は, 接触式の座標測定器 (CMM: Coordinate Measuring Machine) と比較すると, 対象物の毀損が少なく高速測定が可能であるという特徴がある. また, 表面の色などの光学的特性も同時に計測できる. しかし, 対象物体の表面上に投射して拡散反射した光を画像センサで計測することを想定しているため, 鏡面反射, 相互反射, (半)透明, 黒色など想定外の反射特性をもつ表面は正しく計測できない場合がある.

### 3-4-2 レンジデータの処理

#### (1) 低レベル処理

##### (a) メッシュへの変換

距離画像は、点群 (point cloud) が 2 次元配列で構造化された形式になっており、この構造を利用することにより、点群から容易にメッシュに変換することができる。隣接要素を接続することで四辺形メッシュ (quadrilateral mesh) を構成ことができ、更に四辺形の二つの対角線のうち短いほうを接続することで三角メッシュ (triangle mesh) へ変換する処理は一般によく行われる。

##### (b) 法線の推定

レンジデータの法線は 1 次の微分量として求められる最も基本的な局所特徴量である。距離画像の各計測点における法線は、2 次元配列の構造を利用して、輝度値画像における Sobel フィルタのような微分フィルタによって得られた微係数  $(X_x, Y_x, Z_x)$  と  $(X_y, Y_y, Z_y)$  の外積を正規化して求める手法が最も簡便である。三角メッシュに変換されている場合、各頂点での法線は、その頂点を含む三角形の法線の (重みつき) 平均で求める手法も一般的である。レンジデータのエッジについては、距離値が不連続なジャンプエッジと、距離値は連続だが法線方向が不連続なルーフエッジに分類できる<sup>24)</sup>。

##### (c) 誤計測とノイズ除去

レンジデータにおいては様々な計測誤差が含まれている。計測誤差は、センサの原理、構成、対象物体の物性、対象の距離や表面の傾きなど様々な要素に依存する。レンジデータにおいても、輝度値画像での処理と同様に座標値に平滑化フィルタをかけることで計測誤差を除去することはよく行われるが、結果として形状が変わってしまうことにもなる。隣接計測点との距離や、三角メッシュに変換したときの三角形の細長さに基づいて、疑わしい計測点をあらかじめ除去してしまう処理もよく行われる。

#### (2) セグメンテーション

レンジデータを単純な形状の組合せで記述できれば、記述に必要な記憶容量を削減できるとともに、形状認識や位置合せなどの高度な処理に利用することができる。手法としては、エッジに基づく分割に基づくものと、局所的な曲面当てはめに基づくもの、両者を組み合わせるものに大別できる。分節する形状単位としては、平面・球面・2 次曲面などのパラメトリックな曲面や、同一曲率符号領域などのノンパラメトリックな領域がある<sup>25)</sup>。少ない変数で多様な曲面が記述できる superquadrics もセグメンテーションの形状要素としてしばしば利用されるが、非線形関数の当てはめが必要であり、異なった変数で似た形状が得られる不定性がある<sup>26)</sup>。

#### (3) 不変特徴量

一方のレンジデータのどの部分が他方のレンジデータのどの部分に対応しているかを見つけることは、レンジデータ中の物体を同定したり、レンジデータどうしを位置合せして比較したり統合したりする際に重要である。位置や姿勢に不変で局所的な形状記述が一般的に用いられる。

3 次元形状の局所不変特徴量として最も基本的なのは、微分幾何学で定義される曲率であ

る．単純に距離画像を二変数のパラメトリック曲面とみなして数値部分で2次微分を用いて曲率を求めても，安定な結果を得ることができない．曲率を安定に求めるには，ある程度の広さの局所領域に対して頑強な曲面の当てはめを行う必要がある．3次元曲面の曲率は二つの値の組で表されるが(平均曲率とガウス曲率，最大曲率と最小曲率など)，平均曲率とガウス曲率の符号(正・零・負)の組合せによって8種類に分類でき<sup>24, 27)</sup>，その領域の隣接関係なども特徴量として利用することができる．

不変特徴量の局所性と安定性は相反する要請である．実際のレンジデータから安定に，また，曲率より豊富な情報を有する局所特徴量を求めるには，ある程度の広さの領域を参照する必要がある．一方で，参照する領域が広すぎると，オクルージョンによるレンジデータの欠損の影響を受けやすくなる．中心点での法線と接平面により局所的な直交座標を張り，中心点の近傍の局所形状を符号化する手法が多く提案されている<sup>28, 29)</sup>．接平面上の座標値の方向には不定性があるため，最大値や主曲率方向などによる方向の正規化や回転不変量の利用によって不変量を得る．

よく参照される手法として spin image<sup>30)</sup>を簡単に紹介する．レンジデータ中のある点を中心とした近傍について，接平面からの距離と法線からの距離を直交軸とする2次元平面を法線の回りに回転させて，近傍中の計測点の個数を集積して作成した2次元ヒストグラムである．類似画像検索と同様に，主成分分析で次元圧縮し，最近傍探索することにより，点どうしの対応関係を求めることができる．誤った対応関係は，RANSACなどの手法で検出し除去する．

#### (4) 位置合せ

レンジデータ中の点の座標は，そのデータの座標系で記述されている．レンジデータとして獲得された形状データをほかの形状データと比較したり，合成したりするためには，それぞれのデータの座標系を一致させなくてはならない．座標系の位置と姿勢を形状データから推定する作業は位置合せ(registration, alignment, pose / motion estimation)と呼ばれている<sup>29)</sup>．対象物体が変形しない場合，3次元特殊ユークリッド群(special Euclidean group, SE(3))に属する回転行列と並進ベクトルで表される6自由度の剛体運動変換を推定することになる．

位置合せしたい3次元形状データ間で，形状データ中のどの点か他方のデータ中のどの点に対応するか既知な場合，最小二乗解は数学的に得ることができる<sup>31)</sup>が，一般には未知である場合が多い．対応関係は手作業で与えることもできるが，前述した局所不変特徴量やセグメンテーションされた部分形状を用いて自動的に対応づけすることもできる．このような位置合せは，粗い位置合せ(coarse registration)と呼ばれる．

粗い位置合せやロボットのエンコーダーなどにより，大まかな位置合せが既知である場合，それを初期値として，位置合せを変数とする誤差関数を反復演算により最小化するとより精度のよい位置合せが得られる．このような位置合せは，細かい位置合せ(fine registration)と呼ばれる．代表的な手法はIterative Closest Point(ICP)アルゴリズム<sup>32)</sup>である．既知の位置合せに基づいて変換したデータ形状中の各点からモデル形状中で最も近い点を求めることで対応づけを行う操作と，対応づけに基づいて剛体運動の変換係数を求める操作を交互に繰り返す．対応関係にある点間距離の二乗平均はいずれの操作でも減少するため，極小に必ず収束するが，大局的極小である保障はない．また，データ形状がモデル形状の一部であ

ることが仮定されており、計測誤差を含み部分形状であるレンジデータどうしの位置合せにはそのままでは適用できない<sup>33)</sup>。細かい位置合せは基本的に誤差最小化に帰着できるため、Levenberg-Marquardt 法<sup>34)</sup>、Genetic Algorithm<sup>35)</sup>、拡張 Kalman フィルタ<sup>36)</sup>など汎用の非線形最小化アルゴリズムも目的に応じて用いられる。

### (5) 形状モデリング

レンジデータを獲得する重要な応用の一つは、物体認識・物理シミュレーション・文化財のデジタル化などの応用分野において利用できるような、対象物体の形状モデルを作成することである。視点から見て物体の裏側になっている部分や、手前の部分に隠された部分があるため、一般にレンジデータでは対象物体の表面全部を一度に計測することはできない。様々な視点から観測されたレンジデータを位置合せする処理と、全体の形状を統合し一貫した形状記述を得る処理が必要になる。二つのレンジデータを位置合せする手法を逐次的に適用して全体を位置合せしようとしても、位置合せ誤差が蓄積するため、複数のレンジデータを同時に位置合せする手法が必要になる<sup>37, 38)</sup>。統合形状を得るには、ポリゴンメッシュを縫合する面的手法<sup>39)</sup>と、各レンジデータの符号付距離場を平均し、その零交差面を marching cube 法により抽出するような体積的手法<sup>40)</sup>があるが、後者の方が一般によく使用されている。

実際の計測の現場では更に様々な技術が必要になる。対象物表面を効率よく計測するためには、既に計測した部分から次に計測すべき部分を計画する手法 (view planning) が必要になる<sup>41)</sup>。センサの設置位置は物体そのものによって制限され、また、物体の反射特性によっても計測できる領域に限られるため、対象物体の全表面を計測することは実際には困難なことが多く、実用上もっともらしく未計測部を補完する手法 (hole filling) も必要になる。センサをいかに正確に校正していても、機械的誤差・光学的歪み・温度変化・振動により計測されたレンジデータが歪むことがあるので、その補正も必要である。形状だけでなく表面色などの輝度情報も必要な場合には、それらの情報との位置合せや統合も必要になる。

#### 参考文献

- 1) 佐藤 淳, “コンピュータビジョン 視覚の幾何学,” コロナ社, 1999.
- 2) R. Hartley and A. Zisserman, “Multiple View Geometry,” Cambridge University Press, 2000.
- 3) O.D. Faugeras and Q.T. Luong, “The Geometry of Multiple Image,” MIT Press, 2001.
- 4) 佐藤 淳, “テンソルと多視点幾,” in コンピュータビジョン 最先端ガイド 1, 八木康史, 斉藤英雄 編, pp.103-138, アドコムメディア, 2008.
- 5) 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, vol.24, no.6, pp.381-395, 1981.
- 6) Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.11, pp.1330-1334, 2000.
- 7) M. Pollefeys, R. Koch and L. Van Gool, “Self-Calibration and Metric Reconstruction In spite of Varying and Unknown Intrinsic Camera Parameter,” International Journal of Computer Vision, vol.32, no.1, pp.7-25, 1999.
- 8) B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon, “Bundle Adjustment – A Modern Synthesis,” in Vision Algorithms: Theory and Practice, ed. B. Triggs, A. Zisserman and R. Szeliski, pp.298-375, Springer Verlag, 2000.

- 9) A. Heyden and K. Astrom, "Flexible Calibration: Minimal Cases for Auto-Calibration," Proc. International Conference on Computer Vision, vol.1, pp.350-355, 1999.
- 10) G.P. Stein, "Lens Distortion Calibration using Point Correspondences," Proc. Computer Vision and Pattern Recognition, pp.602-608, 1997.
- 11) 植芝俊夫, 富田文明, "平面パターンを用いた複数カメラシステムのキャリブレーション," 情報処理学会論文誌, コンピュータビジョンとイメージメディア, vol.44, no.SIG17, pp.89-99, 2003.
- 12) 奥富他, "デジタル画像処理," CG-ARTS 協会, 2004.
- 13) D. Marr and T. Poggio, "Cooperative Computation of Stereo Disparity," Science, vol.194, No.4262, pp.283-287, 1976.
- 14) N. Yokoya, "Surface Reconstruction Directly from Binocular Stereo Images by Multiscale-multistage Regularization," Proc. Int. Conf. on Pattern Recognition, vol.1, pp.642-646, 1992.
- 15) M. Okutomi and T. Kanade, "A Multiple-baseline Stereo," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.15, no.4, pp.353-36, 1993.
- 16) T. Sato, M. Kanbara, N. Yokoya and H. Takemura, "Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo Using a Hand-held Video Camera," Int. J. of Computer Vision, vol.47, no.1-3, pp.119-129, 2002.
- 17) K.N. Kutulakos and S.M. Seitz, "Theory of Shape by Space Carving," Int. J. of Computer Vision, vol.38, no.3, pp.199-218, 2000.
- 18) Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.1-8, 2007.
- 19) R.C. Bolles, H.H. Baker and D.H. Marimont, "Epipolar-plane Image Analysis: An Approach to Determining Structure from Motion," Int. J. of Computer Vision, vol.1, no.1, pp.7-55, 1987.
- 20) 井口征士, 佐藤宏介, "三次元画像計測," 昭晃堂, 1990.
- 21) P.J. Besl, "Active Optical Range Imaging Sensors," Machine Vision and Applications, vol.1, pp.127-152, 1988.
- 22) F. Blais, "Review of 20 Years of Range Sensor Development," Journal of Electronic Imaging, vol.13, no.1, pp.231-240, 2004.
- 23) J. Salvi, J. Pagès, and J. Batlle, "Pattern Codification Strategies in Structured Light Systems," Pattern Recognition, vol.37, pp.827-849, 2004.
- 24) P.J. Besl and R.C. Jain, "Invariant Surface Characteristics for 3D Object Recognition in Range Images," Computer Vision, Graphics, and Image Processing, vol.33, pp.33-80, 1986.
- 25) A. Hoover, G. Jean-Baptiste, X. Jiang, P.J. Flynn, H. Bunke, D.B. Goldgof, K. Bowyer, D.W. Eggert, A. Fitzgibbon and R.B. Fisher, "An Experimental Comparison of Range Image Segmentation Algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.18, no.7, pp.673-689, 1996.
- 26) D. Terzopoulos and D. Metaxas, "Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.13, no.7, pp.703-714, 1991.
- 27) 横矢直和, マーチン D. レビン, "微分幾何学特徴に基づく距離画像分割のためのハイブリッド手法," 情報処理学会論文誌, vol.30, no.8, pp.944-953, 1989.
- 28) F. Stein and G.G. Medioni, "Structural Indexing: Efficient 3-D Object Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.14, no.2, pp.125-145, 1992.
- 29) J. Salvi, C. Matabosch, D. Fofi and J. Forest, "A Review of Recent Range Image Registration Methods with Accuracy Evaluation," Image and Vision Computing, vol.25, no.5, pp.578-596, 2007.
- 30) A.E. Johnson and M. Hebert, "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.21, no.5, pp.433-449, 1999.

- 31) D.W. Eggert, A. Lorusso and R.B. Fisher, "Estimating 3-D Rigid Body Transformations: A Comparison of Four Major Algorithms," *Machine Vision and Applications*, vol.9, pp.272-290, 1997.
- 32) P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.14, no.2, pp.239-256, 1992.
- 33) S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," *Proc. 3rd International Conference on 3-D Digital Imaging and Modeling*, pp.145-152, 2001.
- 34) A.W. Fitzgibbon, "Robust Registration of 2D and 3D Point Sets," *Image and Vision Computing*, vol.21, pp.1145-1153, 2003.
- 35) L. Silva, O.R.P. Bellon and K.L. Boyer, "Precision Range Image Registration using a Robust Surface Interpenetration Measure and Enhanced Genetic Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.27, pp.777-792, 2005.
- 36) J. Feldmar and N. Ayache, "Rigid, Affine and Locally Affine Registration of Free-form Surfaces," *International Journal of Computer Vision*, vol.18, no.2, pp.99-119, 1996.
- 37) P.J. Neugebauer, "Reconstruction of Real-world Objects via Simultaneous Registration and Robust Combination of Multiple Range Images," *International Journal of Shape Modeling*, vol.3, no.1&2, pp.71-90, 1997.
- 38) 増田 健, "符号付距離場の一致による複数距離画像からの形状モデル生成," *情報処理学会論文誌：コンピュータビジョンとイメージメディア*, vol.44, no.SIG5(CVIM6), pp.30-40, 2003.
- 39) G.Turk and M. Levoy, "Zippered Polygon Meshes from Range Images," *Proc. 21st Annual Conference on Computer Graphics and Interactive Techniques ACM SIGGRAPH '94*, pp.311-318, 1994.
- 40) B. Curless and M. Levoy, "A Volumetric Method for Building Complex Models from Range Images," *Proc. 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp.303-312, 1996.
- 41) W.R. Scott, G. Roth, and J.F. Rivest, "View Planning for Automated Three-dimensional Object Reconstruction and Inspection," *ACM Computing Surveys*, vol.35, no.1, pp.64-96, 2003.