

## Application of 3D scene reconstruction using Multi-Viewpoint Images based on Epipolar Geometry

遠藤 昭紀

デザイン学部技術造形学科

Akinori ENDO

Department of Art and Science

Faculty of Design

コンピュータ・ビジョン (CV) 分野における最も基本的な原理として、複数の異なる視点から、デジタルカメラで対象の 3 次元シーンを撮った複数の画像間にある幾何学的関係を統一的に表現するエピポーラ幾何学が存在する。このエピポーラ方程式を数値的に計算する応用ソフトウェアを利用して、多視点画像から元の 3 次元シーンの情報を実験的に復元し、その結果を検証しそれがどのように応用できるかを考察する。

Epipolar Geometry is well known as the most fundamental principles in Computer Vision, which gives the unified description of the geometrical relation between multi-viewpoint images of 3D scene. I have reconstructed the original 3D scene from given multi-viewpoint images by using the application software which can calculate Epipolar Equation numerically and investigated the possible applications.

### 1. はじめに

最近の 10 年間、映像デザイン分野におけるコンピュータグラフィクスを応用した新しい映像メディア作品の出現には、隔世の感がある。この背景には、半導体設計製造技術の急速な進歩により  $\mu$  プロセッサ、グラフィクス、メモリー等の VLSI が高性能・低価格で提供されるようになり、それらがパーソナルコンピュータ (PC) に搭載され、個人でも妥当な価格で強力な 3D グラフィクス機能を持つ PC が入手できるようになってきたことがある。また、デジタルカメラ、デジタルビデオカメラの発展にも著しいものがある。

今やコンピュータグラフィクス (CG) は、企業においても大学においても、機械設計、VLSI 回路設計等各種の CAD、及び 3 次元 CG などのアプリケーションソフトウェアが多く導入され、製品開発、作品制作、研究、教育等の分野に利用されて多大の貢献をしている。

この CG の学際分野に、コンピュータビジョン (CV) という分野がある。この CV では、3 次元シーンの光学像が 2 次元平面に写像されたものとして画像データを捉え、2 次元画像から 3 次元シーンという逆写像を行い、元の 3 次元シーンの情報を復元することを処理の目的としている。具体的な研究テーマとしては、画像中の陰影やテクスチャー、輪郭形状、見かけの運動情報からの 3 次元形状・運動情報の復元、ステレオ視による奥行き計測、3 次元運動の形状モデルを利用した物体認識などが取り上げられ、またロボットビジョンへの応用も研究されるようになり、産業界においてもコンピュータビジョン (CV) の地位が定着してきた。

最近の研究動向として、2 次元画像から 3

次元シーンの写像を行う CV と 3 次元シーンから 2 次元画像の写像を行う CG という相補的な情報変換プロセスを統合化した視覚情報処理が行われるようになってきた。具体的な例として、次のようなことがある。

- 画像から得られた 3 次元シーン記述を基に、撮影時とは異なる視点からシーンを見たときの画像を生成する。
- 画像から得られた 3 次元シーン記述とプログラムで生成したシーン記述を組合せ、実写画像と CG 画像を合成した画像を生成する。
- 画像に記録された色情報を、CV アルゴリズムによって物体色成分と光源色成分に分離し、物体色のみを変化させた画像を生成する。

これらは、コンピュータグラフィクスで従来行われてきた、対象世界すべてを詳細にモデリングする方法では膨大な手間がかかるため、実写画像を処理して「見え方」を直接的に生成しようという手法である。そのためには、対象の形状、表面の反射特性等の情報を何らかの形で実画像から抽出する必要があり、そこでコンピュータビジョン (CV) の技術が必要とされる。[4]、[8]、[9]、[10]、[11]

本調査研究は、複数の異なる視点から、デジタルカメラで同一の 3 次元シーンを撮った複数の画像間にある幾何学的関係を記述するエピポーラ幾何学 (エピポーラ方程式) を利用して、元の 3 次元世界の情報を復元し、それがどのように応用できるかを考察し、また新しい応用分野を模索する事を目的として行う。

### 2. カメラモデルと投影

#### 2.1 ピンホール・カメラと透視投影

視覚を情報処理として考える場合には、必

ずしも実際のカメラの物理的な投影をそのまま考える必要はない。カメラの幾何学的な性質のみを考えるのであれば、理想的なカメラを考え、レンズの中心の一点のみに注目し、撮像素子に至る光はすべてこの一点を通過してくるものと考えの方が扱いやすい。このようなカメラのことをピンホール・カメラ (pinhole camera) と呼んでいる。

図 2.1 にピンホールカメラを示す。画像平面 I から距離  $f$  のところに I に平行な面 F を置き、その上の点 C にピンホールをあける。物体から来る光はピンホール、即ち点 C を通り、画像平面に像を結ぶ。物体の点とピンホール、及び画像平面上の点は一直線上にある。このような投影を、透視投影 (perspective projection) と呼び、点 C をレンズ中心 (focal point)、または焦点 (focus)、面 F を焦点面 (focal plane)、レンズ中心から画像平面までの距離  $f$  を焦点距離 (focal length) という。点 C を通り画像平面に垂直な線を光軸 (optical axis) といい、それと画像平面との交点  $c$  を光軸点 (principal point) という。勿論、光軸は焦点面にも直交する。このようなモデルは、CCD カメラ (デジタルカメラ) をかなり正確に記述している。

まず、透視投影の方程式を記述する座標系を定義しよう。画像座標系は、光軸点  $c$  を原点とし、 $x$  軸と  $y$  軸はそれぞれカメラの素子の配置軸に合わせる。光軸を第 3 の軸としたときに、右手システムとなるように順番を決める。さらに 3 次元空間の座標系は、焦点 C

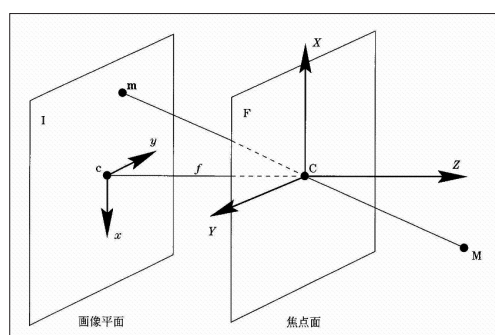


図 2.1 ピンホールカメラモデル  
ピンホールカメラは透視投影を行う。3次元座標の原点はカメラのレンズ中心にあり、画像面はレンズ中心の後ろにある。

を原点に、光軸を Z 軸とし、X 軸と Y 軸は  $x$  軸と  $y$  軸にそれぞれ平行に置くが、方向は反対にする。このような  $(C, X, Y, Z)$  座標系をカメラ座標系と呼ぶ。

この座標系を用いると、空間中の点の 3 次元座標とその 2 次元画像との間に、次の関係が成り立つ。

$$\begin{aligned} x &= f \frac{X}{Z} \\ y &= f \frac{Y}{Z} \end{aligned} \quad (2.1)$$

図 2.2 に示すように、焦点面の後ろにある画像面を、焦点の前に持ってきて、反転させても上記の関係が成り立つ。コンピュータビジョンでは、この図の方が多く用いられている。ここでは、画像上の点  $m(x, y)$  は同時にカメラ座標系において 3 次元座標  $(x, y, f)$  を持つことになる。

## 2.2 投影行列と外部変数

非線形の式 (2.1) は式 (2.2) のように線形表現できる。

ここで、 $(X, Y, Z)$  は 3 次元空間の点の座標、 $(x, y)$  はその画像への投影の座標である。 $\cong$  はスケールを除いて等しいことを意味する。

与えられた座標ベクトル  $\mathbf{x} = (x, y, \dots)^T$  に対して、要素 1 を加えたもの  $(x, y, \dots, 1)^T$  を  $\mathbf{x}$  の同次座標 (homogeneous coordinates) または拡張ベクトル (augmented vector) といひ、 $\tilde{\mathbf{x}}$  で表す。この同次座標を用いると、

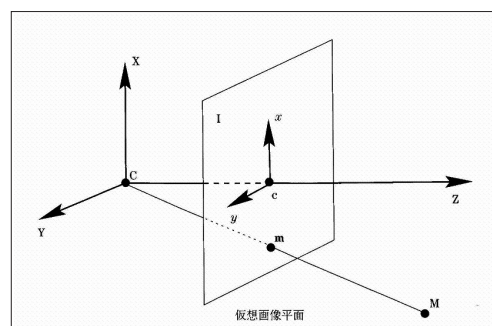


図 2.2 仮定の画像平面をレンズの中心の前に置いたピンホールカメラモデル。一般にはこのモデルが良く用いられる。

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.2)$$

上の式 (2.2) は以下のように変形できる。

$$\tilde{m} = P\tilde{M} \quad (2.3)$$

ここで、

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.4)$$

は透視投影行列という。

今までは、3次元空間内の点をカメラ座標系において記述してきたが、他の世界座標系 (world coordinate system) における記述を用いることも多い。その場合は、まず世界座標系からカメラ座標系へ変換し、その次に式 (2.2) を適用すればよい。

図 2.3 に示すように、カメラ座標系をまず回転させ、次にその姿勢を保ったまま平行移動させて、世界座標を得る。回転は  $3 \times 3$  の行列  $\mathbf{R}$ 、並進は  $3 \times 1$  のベクトル  $\mathbf{t}$  で表現される。 $\mathbf{R}$  は下記の性質を持つ。

$$\mathbf{R} \mathbf{R}^T = \mathbf{R}^T \mathbf{R} = \mathbf{I} \quad (2.5)$$

$$\det(\mathbf{R}) = 1 \quad (2.6)$$

ここで、 $\mathbf{I}$  は単位行列である。これは 6 つの独立した拘束条件となり、回転行列は 3 つ

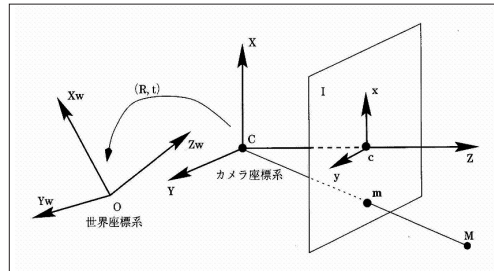


図 2.3 世界座標系とカメラ外部変数

世界座標系に記述される 3 次元点がカメラ内部変数を通じてカメラ座標系に変換され、画像平面へ射影される。

の自由度しか持たない。

ある点のカメラ座標系における座標  $\mathbf{M}_c$  とそのワールド座標系における座標  $\mathbf{M}_w$  との関係は、

$$\mathbf{M}_c = \mathbf{R} \mathbf{M}_w + \mathbf{t} \quad (2.7)$$

と表される。

### 2.3 正規化カメラ

ここでは、カメラモデルの内部パラメータについて考える。我々が用いるカメラモデルを決定するには、先ず次のことを考える必要がある。① 光軸点がどこにあるかは不明である。② 画像の両座標軸のスケールは異なるかもしれない (画素が正方形でない)。③ 実際の画像の両座標軸は必ずしも直角ではない。

ここで正規化カメラ (normalized camera) の概念を導入する。正規化カメラの画像平面は焦点から単位長のところ、すなわち、 $f = 1$  にある。正規化カメラの画像を正規化画像 (normalized image) といい、その座標  $[x, y]^T$  を正規化画像座標という。 $[x, y]^T$  と  $[X, Y, Z]^T$  との関係は次式によって決まる。

$$\begin{aligned} x &= \frac{X}{Z} \\ y &= \frac{Y}{Z} \end{aligned} \quad (2.8)$$

どのカメラを用いても、デジタル画像座標から正規化画像座標に変換できるので、それぞれのカメラの特性を考慮せず、正規化カメラを用いて CV の問題を考えることが可能になる。

最近のデジタルカメラは、製造技術の進歩により、画像の両座標軸の成す角は直角 ( $\theta = \pi/2$ )、CCD 素子も正方形と見なせるようになってきた。従って、未知の内部変数は 3 個となり、光軸点 ( $u_0, v_0$ ) と焦点距離 ( $f$ ) となる。また、光軸点を画像中心としてもそれ程問題とならないので、その場合焦点距離のみが未知である。高精度の計測が目的でない限り、この簡略モデルで十分である。

### 2.4 透視投影近似の方法

透視投影は非線形関数であり、この非線形

性によって多くの問題が難解となる。さらに遠近範囲が小さい場合は、透視投影式を用いた解が不安定になる。しかし、一定の条件を満たせば、透視投影は線形近似できる。一般的な線形近似と同じように、この線形近似が許される条件は、(2.1) 式の分母の奥行き  $Z$  があまり変化しないこと、即ち対象物体の大きさが、カメラから物体までの距離と比較して十分小さいことである。

ここで線形近似は、階数に応じて弱透視投影 (weak perspective projection)、及び平行透視投影 (paraperspective projection) がある。また、平行投影 (orthographic projection) は、弱透視投影の特殊な場合である。アフィン射影 (affine projection) はこれらの線形化射影の一般化したもので、広く用いられている。[1]、[2]、[6]

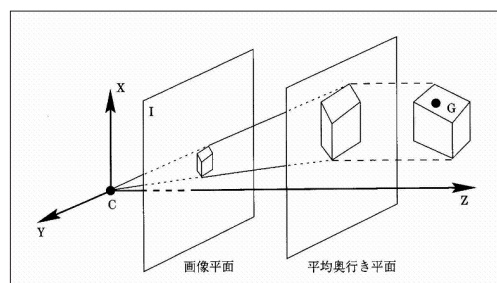


図 2.4 弱透視投影の幾何 物体重心を通る平面を物体の前に置いた。物体のカメラまでの距離によって像は大きくもなり小さくもなる。

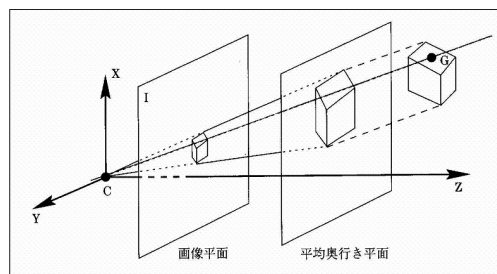


図 2.5 平行透視投影の幾何 物体重心を通る平面を物体の前に置いた。弱透視投影よりもさらに透視投影に近い。カメラまでの距離も、光軸までの距離も像に反映される。

### 3. エピポーラ幾何学

3次元の世界を2次元画像に投影すると、画像において1次元分の情報が不足するため、3次元世界に関する完全な記述が得られない。このことを不定性 (ambiguity) が残るという。従って、一つのカメラから得られた画像のみからでは、3次元物体の形状を一意に決定することはできない。しかし、二つの異なる視点で得られた画像からはこのような不定性を取り除き、形状を決定することができる。このとき、相対的なカメラの位置や姿勢などの情報が必要となる。

位置や姿勢の情報を得る方法としては、予め動きが判っている移動装置にカメラを設置して動かす方法などがあるが、一方、カメラを動かしたときの画像の変化自身から視点位置や姿勢の変化を計算することができれば、余計な装置などを使う必要がなく、より優れているといえる。

カメラを移動させたとき、あるいは複数のカメラがあるとき、複数の視点における相対的なカメラの位置や姿勢の情報は、エピポーラ幾何と呼ばれる画像特有の幾何によって記述できる。

#### 3.1 エピポーラ幾何

図 3.1 に示すように、3次元空間のある点  $X$  が視点  $C$  の画像に投影されているとする。  $X$  と  $C$  を結ぶ直線  $L$  が画像面  $\pi$  と交差してできる点  $x$  が3次元空間中の点  $X$  の画像面  $\pi$  への投影像である。同様に視点  $C'$  を考えると、  $X$  と  $C'$  とを結ぶ直線  $L'$  が第2の画像面  $\pi'$  と交差してできる点  $x'$  が空間中の点  $X$  の第2画像面  $\pi'$  への投影像である。

このとき、  $C, C', X$  の3点は空間中に一つの平面  $\Sigma$  を定義する。この平面  $\Sigma$  のことをエピポーラ平面 (epipolar plane) と呼ぶ。二つの画像点  $x, x'$  もこのエピポーラ平面上にあることは明らかである。エピポーラ平面  $\Sigma$  と画像面  $\pi$  とが交差してできる直線  $l$  はエピポーラ線 (epipolar line) と呼ばれる。同様にエピポーラ平面  $\Sigma$  が視点  $C'$  の画像面  $\pi'$  と交差してできる直線  $l'$  は第2画像面におけるエピポーラ線である。

いま第1の視点  $C$  で画像点  $x$  が観測され

たとき、これに対応する空間中の点は直線L上のどこかに存在するから、直線Lはxに対応する点の集合であると考えられる。図3.1から明らかなように、この直線Lを第2の画像面 $\pi'$ に投影したものがエピポーラ線 $l'$ である。即ち、第2の画像面上のエピポーラ線 $l'$ は第1の画像面上の点xに対応する第2の画像面上の対応点の集合である。同様に、第1の画像面 $\pi$ 上のエピポーラ線 $l$ は、第2の画像面上の点 $x'$ に対応する第1の画像面上の対応点の集合である。このように、エピポーラ線は画像面上の点同士の対応関係を表している。

また、視点Cと視点C'を結ぶ直線が二つの画像面 $\pi, \pi'$ と交差してできる点 $e, e'$ をエピポール(epipole)と呼ぶ。画像面 $\pi$ 上のエピポール $e$ は視点C'を視点Cに投影した像と考えることができる。同様に、画像面 $\pi'$ 上のエピポール $e'$ は視点Cを視点C'に投影した像と見なすことができる。このように、エピポールは視点の対応関係を表しているのである。今までは3次元空間中の一点のみにつ

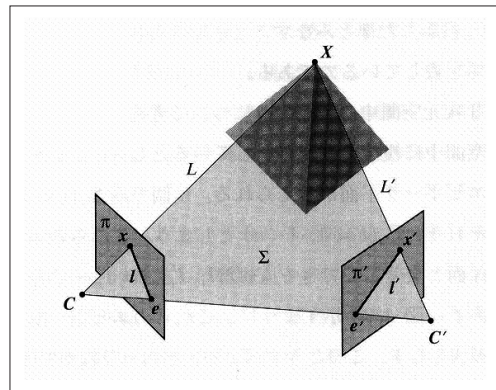


図 3.1 エピポーラ幾何

いて考えてきたが、実際には図3.2に示すように複数の点が存在しており、それぞれの点についてエピポーラ平面が考えられ、エピポーラ線が存在する。これらのエピポーラ線はすべて対応するエピポール点で交差し、この点を中心とした扇形の形状をしている。このように3次元空間の対象物を複数のカメラに投影した場合には、複数の画像間に特有な幾何が現われる。このような幾何のことをエ

ピポーラ幾何(epipolar geometry)と呼ぶ。対象とする3次元構造を復元するには、複数のカメラ間の相対的な位置関係や姿勢の情報が必要となる。このような相対的な位置や姿勢の情報を与えるのがエピポーラ幾何である。

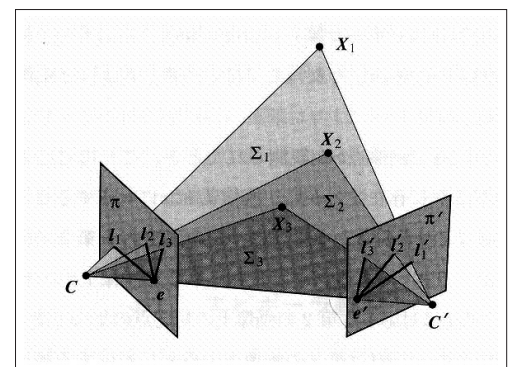


図 3.2 エピポーラ幾何

### 3.2 エピポーラ方程式

つぎに、エピポーラ方程式を導こう。(ここでは詳細な数式の導出、証明については、参考文献に記載した参考著書を参照し、結果のみを記すこととする。)[3]、[6]、[7]

二つの視点におけるそれぞれの画像面上の点 $x, x'$ に関しては、次の関係があることが判っている。

$$\lambda' \tilde{x}' = \lambda R \tilde{x} + \eta T \quad (3.1)$$

ここで、 $\tilde{x} = [x_1, x_2, x_3]^T$ は画像上の点 $x = [x, y]^T$ に対する同次座標であり、 $[x, y, 1]^T \sim [x_1, x_2, x_3]^T$ なる同値関係を持つ。同様に、 $\tilde{X} = [X_1, X_2, X_3, X_4]^T$ は3次元空間の点 $X = [X, Y, Z]^T$ の同次座標であり、 $[X, Y, Z, 1]^T \sim [X_1, X_2, X_3, X_4]^T$ なる同値関係を持つ。

また、 $T$ は並進を表す $3 \times 1$ のベクトル、 $R$ は回転を表す $3 \times 3$ の行列である。

式(3.1)は、図3.3に示すように、三つのベクトル $\tilde{x}, R\tilde{x}, T$ が同一平面上にあることを示している。同一平面上に三つのベクトルがある場合、そのうちの任意の二つのベクトルの外積と残る一つのベクトルとの内積は0となる。したがって、次の式が成り立つことがわかる。

$$\tilde{x}' \cdot (T \times R \tilde{x}) = 0 \quad (3.2)$$

これを行列演算の形で表すと次のようになる。

$$\tilde{x}'^T [T]_{\times} R \tilde{x} = 0 \quad (3.3)$$

ここで、 $[T]_{\times}$ は、 $T$ と任意のベクトル $V$ との外積がつぎのように行列演算によって表せるような $3 \times 3$ の歪対称行列である。

$$[T]_{\times} V = T \times V \quad (3.4)$$

$$[T]_{\times} = \begin{bmatrix} 0 & -T_3 & T_2 \\ T_3 & 0 & -T_1 \\ -T_2 & T_1 & 0 \end{bmatrix} \quad (3.5)$$

いま、行列 $E$ を、

$$E = [T]_{\times} R \quad (3.6)$$

と置くことにすると、式(3.3)は次のように表すことができる。

$$\tilde{x}'^T E \tilde{x} = 0 \quad (3.7)$$

これが二つの画像上の対応する点 $x$ と $x'$ との間の関係を表す式である。式(3.7)はエピポーラ方程式(epipolar equation)と呼ばれ、このときの $3 \times 3$ 行列 $E$ を基本行列、あるいは、essential行列(essential matrix)と呼ぶ。異なる視点から得られた二つの画像において、対応する点同士は必ず式(3.7)を満たす。これをエピポーラ拘束(epipolar constraints)と呼ぶ。

なお、画像のデジタル画像座標 $m$ と $m'$ が与えられたとき、式(3.7)は、

$$\tilde{x}'^T E \tilde{x} = \tilde{m}'^T F \tilde{m} = 0 \quad (3.8)$$

となる。ここで、

$$F = A'^T E A^{-1} = A'^T [t]_{\times} R A^{-1} \quad (3.9)$$

を基礎行列(fundamental matrix)という。基礎行列はカメラの内部変数と外部変数の双方

を含んでいる。同様に、 $E$ を $F$ で表現することもできる。

$$E = A'^T F A \quad (3.10)$$

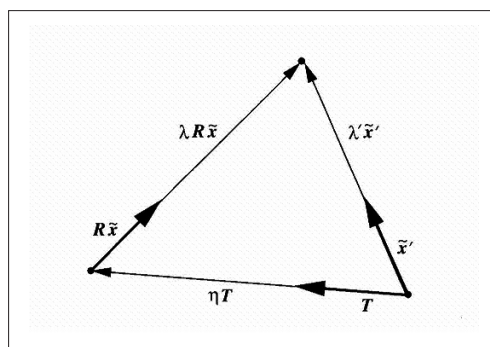


図 3.3 エピポーラ拘束

式(3.6)から明らかのように、基本行列は二つのカメラ間の回転及び並進に関する情報からなっている。この行列は $3 \times 3$ であり9つの要素があるが、実際には回転3自由度と並進2自由度の合わせて5自由度しかない。並進が3自由度ではなく2自由度であるのは、基本行列を定数倍しても式(3.7)が表す内容には変化がなく定数倍の不定性があるため、並進に関しては大きさが一意に定まらないからである。

基本行列は歪対称行列 $T$ と直交行列 $R$ の積であるから、 $E$ が求めれば、これを $T$ と $R$ に分解することができ、カメラの6自由度の運動のうちで、並進の大きさを除く回転3自由度と並進2自由度を求めることができる。

#### 4. 応用ソフト試用による結果と評価

前に説明したように、対象物の3次元形状を復元するには、エピポーラ幾何に基き、複数のカメラ間の相対的な位置関係や姿勢の情報を算出する必要がある。これを数値的に計算する商用ソフトウェアが、唯一、立命館大学情報理工学部メディア情報学科徐剛教授が起業して開発したソフトウェア商品「3D-MODEシリーズ」である。本調査研究は、このソフトウェアを利用して実施した。3D-MODEシリーズの内容は豊富であり、下記の機能を備えている。

- 3DM Modeler Professional  
3次元モデリング及び計測
  - 3DM Sticher パノラマ合成
  - 3DM Calibrator カメラ校正
  - 3DM Viewer Java Appletビューワ
- ここでは、3次元モデリングと3DMビューワの機能を利用して実験を行った。これらの操作方法については、「3D-MODE ユーザーズガイド」に任せるとして、その中でエピポラ幾何に関わる重要な事項についてのみ抜粋して記載することとする。[5]

#### 4.1 操作方法の全体の流れ

「3DM シリーズ」ソフトウェアの利用手順の全体の流れは図 4.1 で表される。

#### 4.2 写真撮影の方法

撮影には一般的なデジタルカメラを使用するが、レンズの歪の大きいカメラは避け、レンズの歪の少ないカメラでできるだけ高い解像度で撮影することが望ましい。

3DM Modeler では物体の同じ部分が共通に見えていることを利用して3次元モデルを作成するので、撮影の際には次の点に注意す

る。

- 物体の全体が写真に収まるようにする。
- 写真の中で物体ができるだけ大きく写るようにする。
- 3次元モデルを作成したい部分が複数の写真に写るようにする。

図 4.3 に表されるように、十分な特徴点がある2枚の写真の間で共通に見えているように撮



図 4.2 対象物体の共通に見えている部分

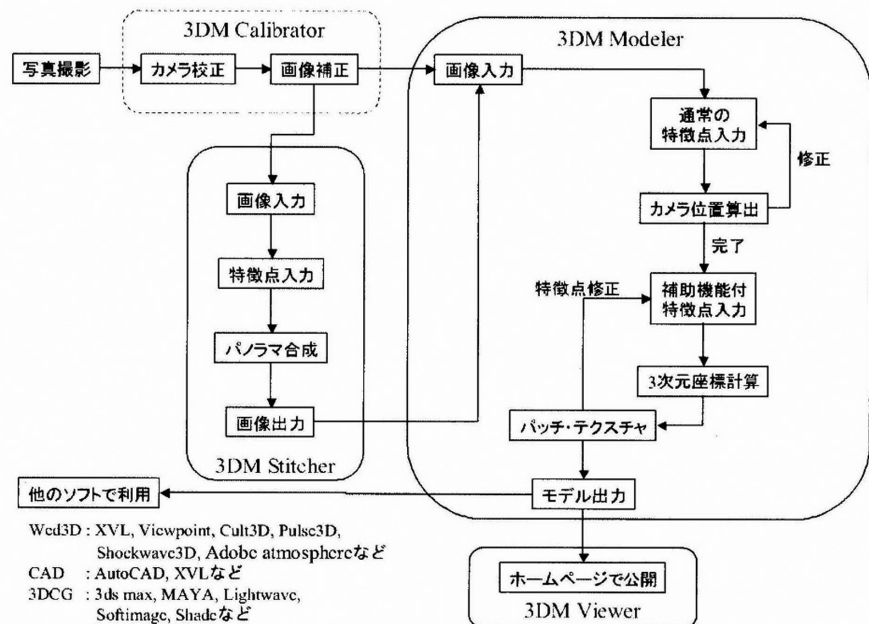


図 4.1 3DM Modeler 利用の流れ

影することが必須である。特徴点を十分な数だけ取得することができないと、カメラの位置姿勢の算出ができない。

3DM Modeler に入力する写真の枚数は、2枚以上という以外に特に制限はない。例えば、対象がディスプレイ・モニターの画像の

場合、正面の部分の3次元モデルだけが必要な場合（図 4.4 に表される）と、全周囲の全方位3次元モデルが必要な場合（図 4.5 に表される）とでは撮影する位置と枚数が異なる。

### 4.3 応用実験の結果

ここでは 3DM Modeler ソフトウェアを利用した 3次元モデリングの例として、対象物として、レーザービームプリンター（LBP-1820）に対して行った実験結果について記載する。

#### 4.3.1 レーザービームプリンターの全方位モデリング

実験として、図 4.5 のようにデジタル写真 4 枚から全方位の 3次元モデルを作成することを行った。

「図 4.6: 特徴点の表示」は、4 枚のデジタル画像を入力した後、76 個の特徴点を入力した状態の操作画面を表しており、特徴点は、対象物面上に赤い×マークで示されている。

ここで入力された画像には、その入力画像（写真）を撮影したカメラの情報（位置・姿勢）が記録されていないので、入力された特徴点の情報を利用してカメラの位置・姿勢の情報を計算する。これにより、入力された特徴点の 3次元座標を計算できるようになる。

なお、カメラの位置・姿勢の情報が計算できるためには、隣り合う画像間で特徴点の対応が理論上最低 8 点は必要であるが、実際には 2 枚の間で 10 から 12 点程度の対応点があれば、計算は安定する。従ってどの画像も、少なくとも他の 1 枚の画像との間で 10 ~ 12 点程度の対応点が入力されるようにする。

この段階で、カメラの位置・姿勢の計算を実行すると、各特徴点の 3次元空間上の座標が計算される。「図 4.7: 投影点の表示」に示されるように、特徴点が表示される入力画像上に、青い×マークが描画される。この点は、特徴点の 3次元空間上の座標から、計算されたカメラの位置・姿勢の情報を使って入力画像上に投影した点になる。特徴点と投影点のマークがほぼ同じ位置に重なっていれば誤差は少ないということであり、大きく外れていれば誤差は大きいということになる。

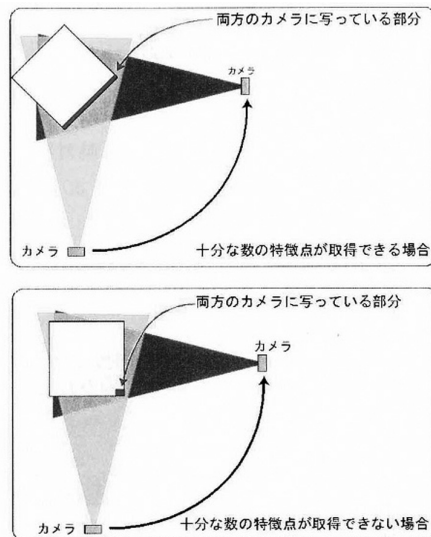


図 4.3 特徴点の数が十分な場合と不十分な場合

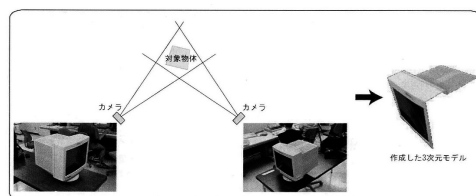


図 4.4 写真 2 枚から正面部 3次元モデルを作成した場合

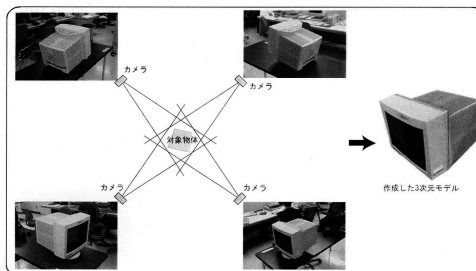


図 4.5 写真 4 枚から全方位 3次元モデルを作成した場合



画面左側にある特徴点情報リストに表示されている番号の左にある顔マークの色により、特徴点が正しく入力されているかどうかを次のように確認することができる。

- 灰色： 3次元座標未計算 (0 個)
- 緑色： 3次元座標計算済み、誤差は小 (59 個)
- 黄色： 3次元座標計算済み、誤差は中 (14 個)
- 赤色： 3次元座標計算済み、誤差は大 (3 個)

このようにして作成した 3 次元モデルは、3D Viewer (ブラウザで表示するための Java Applet ビューワ) を使用して全方位の視点から見るができる。「図 4.8: 任意の視点からの見え方」は、代表として、対象物の前後左右面の少し斜め上の 4 つの視点、及び前面と後面の 2 つの視点から見たモデルを表示したものである。本来は、アニメーションにて復元モデルを動的に見るのが最も効果的だが、本報告書では残念ながらそれができない。

当ソフトウェアがこのようなモデリングの数値計算を実行するアルゴリズムの概略は次のようである。

多視点画像から全方位モデルを直接に復元する場合には、ある評価関数を定義し、その最小化により最適な回転、並進、スケーリングなど 3 次元形状と視点間運動を求める。次に、画像上の特徴点その逆射影とのユークリッド距離の自乗和からなる評価関数を定義する。評価関数は非線形なので、最小化は非線形関数最小化のアルゴリズムを適用して行っている。

#### 4.4 モデリング結果の評価

レーザービームプリンターのモデルを撮影した 4 枚の画像を入力とし、3DM Modeler ソフトウェアを利用して全方位モデルを復元してみた結果、次のような結論を得た。

- (1) 3DM Viewer による任意視点からの表示から判るように、4 枚の画像入力 (76 個の特徴点) では、復元された形状にはやや不安定性 (歪み) が見られ

る。また、テクスチャーの貼り付けについては、入力画像から生成したものであるため、任意視点から見たときには正確さに限界がある。しかし、4 枚の画像入力だけで大まかなモデリングができれば、それを他の CG ソフトに入力して完成させるという省力化の利用価値がある。

- (2) 対象モデルの形状の複雑さにより、必要な入力画像の枚数は異なり、当然ながら凹凸の多い形状であればそれだけ特徴点の個数を多く取らなければならない。
- (3) 対象モデルの曲面、曲線に対して特徴点を入力する場合には、精度を上げるため曲線に沿って多くの特徴点個数が必要である上に、それらの対応点を二つの画像間で特定することが難しい。結局のところ、試行錯誤の方法に依るしかない。
- (4) 曲面の多い少ない、凹凸の多い少ない等の要因により、必要な入力画像枚数が異なり、枚数が多い場合には作業時間が大幅に長くなるので、必要とされるモデル形状復元の精度をよく見極めた上で 3DM Modeler の利用方法を判断することが必要である。
- (5) 当該ソフトウェアの応用操作性については、特徴点を入力するのは、すべて人の視覚による手作業によるものであり、その作業時間は 4 枚の画像入力の場合には大体 8 時間を要する。勿論、対象物体の形状の複雑さによっては、必要な入力画像枚数、及び作業時間は異なってくる。例えば、曲線、曲面の多い形状については 8 枚以上の画像が必要であるが、曲面上では対応点を特定することが難しく、さらに作業時間も大幅に増え、例えば数日間要する。



図 4.6 特徴点の表示

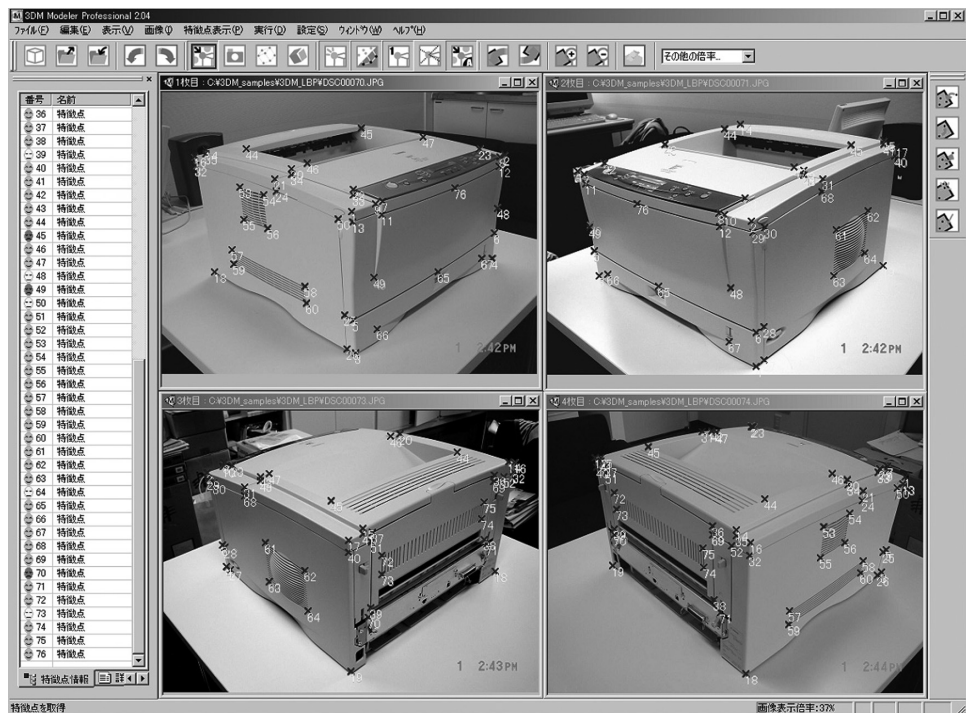


図 4.7 投影点の表示

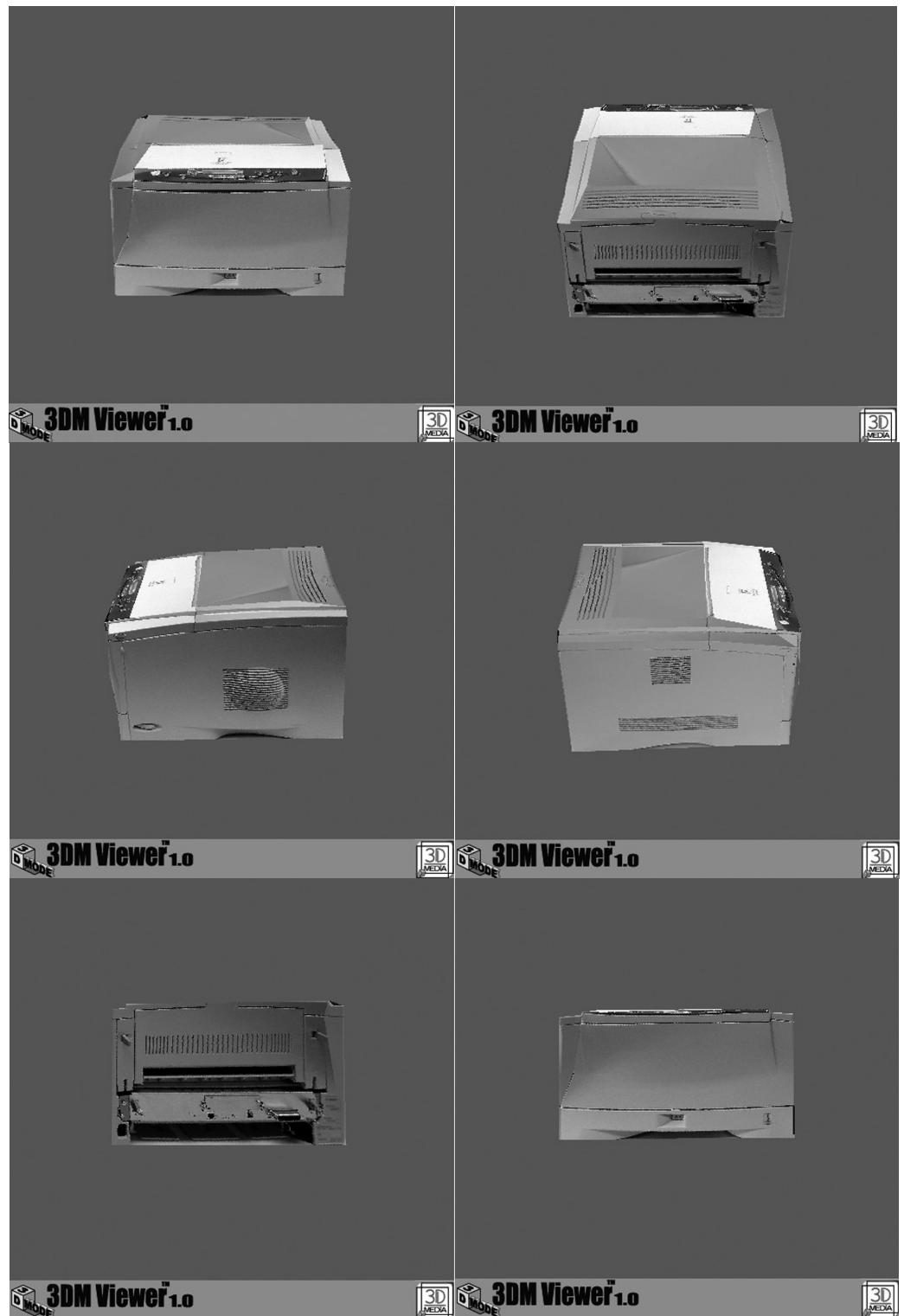


図 4.8 任意の視点からの見え方

## 5. おわりに

本調査研究は、コンピュータビジョン (CV) における基本的な問題である、複数の 2次元画像から 3次元シーンの形状を記述する「エピポーラ幾何学」におけるエピポーラ方程式を数値的に計算する応用ソフトウェアを利用して、その応用的効果を調査したものである。

得られた成果及び結論は、次のようなものである。

モデリングのための特徴点を入力するのは、すべて人の視覚による手作業によるものであり、複雑な形状の対象物に対しては長時間の作業を必要とする。従って、何らかの入力の自動化が望まれる。

この方式の最大の利点は、任意視点におけるデジタルカメラの画像入力で 3次元シーンの形状復元ができることなので、街中の建築物、近づくことができない対象物、交通事故現場などに対しては実用的な効果が期待できるとされる。また、3次元計測の機能も備わっているため、この機能も効果的に利用できるとされる。

最後に今後の方針として、下記のことを検討している。

(1) 今回の調査研究では、十分な応用実験例を試すことができなかったため、これらを発展継続していく。特に、対象としては、建築モデル、及び街中の建築物を選び、景観シミュレーションとして実用の程度を確認する。

(2) 本ソフトウェアは、主な CG, CAD 応用ソフトウェアに対する変換出力を備えているため、それらに入力して相乗効果的な応用分野を模索する。

(3) 今回利用した 3DM Modeler ソフトウェアが、最近グレードアップされ、2直線入力による交点の特徴点設定、及び限定的な特徴点の自動対応付けなどの機能が追加された。従って、再度当ソフトウェアの新バージョンを利用して本調査研究を発展継続し、その成果を続報として報告する予定である。

## 参考文献：

- [1] 徐剛・辻三郎 (1998) 3次元ビジョン、共立出版
- [2] 徐剛 (2001) 写真から作る 3次元 CG : イメージベースド・モデリング&レンダリング、近代科学社
- [3] 佐藤淳 (1999) コンピュータビジョン—視覚の幾何学、コロナ社
- [4] 松山、久野、井宮編集 (1998) コンピュータビジョン—技術評論と将来展望、新技術コミュニケーションズ
- [5] 3D-MODE ユーザーズガイド (2001) 株式会社三次元メディア
- [6] G. Xu and Z. Zhang (1996) "Epipolar Geometry in Stereo, Motion and Object Recognition": A Unified Approach, Kluwer Academic Publishers
- [7] Richard Hartley, Andrew Zisserman (2003) "Multiple View Geometry in computer vision", Cambridge Univ. Press, 2nd edition
- [8] SIGGRAPH 2001 Course Notes 2: Marc Pollefeys: "Obtaining 3D Models With a Hand-Held Camera"
- [9] SIGGRAPH 2002 Course Notes 44: (Organizer Radek Grzeszczuk) Marc Pollefeys, Leonard McMillan: "Image-Based Modeling" R.Koch, M.Pollefeys: "Calibration of Hand-held Camera Sequences for Plenoptic Modeling", M.Pollefeys, L.VanGool: "Visual Modeling with a Hand-held Camera" W.Matusik, C.Buehler: "Imaged-Based Visual Hulls" W.Matusik, H.Pfister: "Image-Based 3D Photography using Opacity Hulls"
- [10] SIGGRAPH 2003 Course Notes 15: Organized by Marc Pollefeys: "3D Models from Photos and Videos"
- [11] SIGGRAPH 2003 Course Notes 42 : Lecturer Yi Ma, Stefano Soatto, Jana Kosecka: "Multiple-View Geometry for Image-Based Modeling"