

## 半透視射影下での3次元物体認識

杉本 晃宏

室田 一雄

(株)ATR 視聴覚機構研究所

京都大学 数理解析研究所

あらまし 投影像が半透視射影によって得られるとする。このとき任意の物体について、その物体の任意の投影像は、ある条件を満たす3枚の投影像の1次結合として表わされることを示す。これにより、条件を満たす3枚の投影像を蓄積してそれを物体のモデルとすれば、未知像が与えられたとき、それが蓄積した3枚の投影像の1次結合によって表わされるか否かを判定するだけで、未知像が同じ物体から得られたものであるかどうかを判定できる。したがって、物体の3次元情報を復元して物体のモデルを構成する従来の手続きは不要となる。

和文キーワード 物体認識 半透視射影 一次結合 変換群の表現

## Object Recognition under Paraperspective Projection

Akihiro SUGIMOTO<sup>†</sup> Kazuo MUROTA<sup>‡</sup>

<sup>†</sup>ATR Auditory and Visual Perception Research Laboratories,  
Soraku-gun, Kyoto 619-02, Japan  
e-mail: sugimoto@hip.atr.co.jp

<sup>‡</sup>Research Institute for Mathematical Sciences,  
Kyoto University, Kyoto 606, Japan

Abstract This paper studies a problem of determining, for a given paraperspective image, whether it is obtained from an object to be recognized or not. Suppose that a number of 2-D paraperspective images are already stored for a given 3-D object and that the correspondence among the feature points is known. It is found that a 2-D image is considered to be an image of the object if the 2-D image can be expressed as a certain combination of the stored 2-D images. Two classes of admissible transformations for a 3-D object are investigated: Rigid transformations and affine transformations. In both cases, it is shown that three 2-D images are sufficient to describe any other image of the same 3-D object. An algorithm for object recognition based on this property is also presented.

英文 key words object recognition paraperspective projection linear combination representation of transformations

## 1 はじめに

認識すべき3次元物体をひとつ考え、複数の位置からこの物体の投影像を得て、蓄積しているとする。未知像が与えられたとき、既に蓄積している投影像を用いて、これが同じ物体から得られたものであるか否かを判断する問題(物体認識問題)を考える。この問題に対する従来の典型的な手法[2]は、三つのプロセスに分解される。蓄積している複数の投影像から物体の物理的な意味での3次元情報を復元するプロセスと、未知像を得る視線方向を決めるパラメータを決定するプロセス、さらにその決定に基づいて未知像が同じ物体から得られたものであるか否かを判定するプロセスである。第1のプロセスは未知像が与えられる前に実行でき、後者の二つのプロセスは未知像が与えられる度に実行される。

Ullman-Basri [4], [6] は、投影像が正射影によって得られた場合、3次元空間上で剛体運動する物体の任意の投影像は、適当な3枚の投影像の“線形結合”によって表せることを示した。また、物体がアフィン運動する場合の投影像は、2枚の投影像の“線形結合”によって表せることを示した。したがって、3枚の投影像をモデル像として蓄積し、各投影像上の特徴点(4個以上と仮定)間の対応が既知であるとすれば、未知像が与えられたとき、それがモデル像の“線形結合”で表せるか否かを判定するという簡単な手続きで物体認識ができる。これを発展させて、Sugimoto-Murota [5] は透視射影によって投影像が得られた場合に対しても、適当な5枚の投影像のある種の結合によって任意の投影像を表せることを示した。しかしこの場合、各投影像はモデル像の非線形結合として表されるため、未知像がモデル像の非線形結合として表されるか否かの判定は容易でない。

一方、テクスチャ解析の分野では、透視射影の下での解析が困難なので、Ohta-Maenobu-Sakai [3] は透視射影の近似として半透視射影を提案し、Aloimonos [1] がそれに基づく解析を行っている。これは、正射影と透視射影との中間に位置するモデルであって、視点と対象との距離が物体の大きさに比べて十分大きい場合には、透視射影のよい近似になっている。

本稿では、半透視射影によって投影像(第2節参照)が得られた場合に対して、物体認識問題の解析を行い、任意の投影像は3枚のモデル像に関する1次式として表されることを示す。したがって、半透視射影の下では、未知像がモデル像の1次式で表せるかどうかを判定するだけで物体認識ができる。つまり、透視射影を半透視射影で近似すれば、容易に物体を認識することができる。以下、第2節で半透視射影の説明を行い、その意味を明らかにする。第3節と第4節で、本稿で取り扱う問題と取り扱う投影像を定式化する。第5節では、任意の投影像は少数の蓄積像に関する1次式によって表現できることを示し、具体的な認識の手順を第6節で示す。なお、複数の位置から物体を見ることは、固定された視点に対して物体が運動していることと等価なので、本稿では物体が運動していると捉え、その運動としては剛体運動とアフィン運動を取り扱う。

## 2 半透視射影

### 2.1 半透視射影とは

半透視射影(Ohta-Maenobu-Sakai [3] による)とは、大域的には透視射影の性質を保ち、局所的に正射影の考え方を取り入れた投影法のモデルである。焦点距離  $f$  のレンズの中心(視点)が原点  $O$  に一致しているとする。このとき、3次元空間内の点  $p$  (その座標を  $\mathbf{x}^p$  と記す<sup>1)</sup>) の、半透視射影による像  $\tilde{\pi}^p$  は次のように定められる。基準とする点<sup>2</sup>  $G$  の座標を  $\mathbf{x}^G = (x^G, y^G, z^G)^T$  とすると<sup>3)</sup>(図1参照)。

1. 視点と  $\mathbf{x}^G$  を結ぶ直線の向きの正射影によって  $\mathbf{x}^p$  を平面  $z = z^G$  に投影した点の座標を  $\tilde{\mathbf{x}}^p (\in \mathbf{R}^3)$  で表す<sup>4)</sup>。
2. 透視射影によって  $\tilde{\mathbf{x}}^p$  を投影面(平面  $z = f$ ) 上に投影する。具体的には、投影面とベクトル  $\tilde{\mathbf{x}}^p$  が交わる点の座標が  $((\tilde{\pi}^p)^T, f)^T$  となる。ここで、 $\tilde{\pi}^p \in \mathbf{R}^2$  である。

$\mathbf{x}^p = (x^p, y^p, z^p)^T$  のとき、 $\tilde{\pi}^p$  は次のようになる:

$$\tilde{\pi}^p = (\tilde{\pi}_1^p, \tilde{\pi}_2^p)^T \quad (2.1)$$

$$= f \begin{pmatrix} \frac{1}{z^G} & 0 & -\frac{x^G}{(z^G)^2} \\ 0 & \frac{1}{z^G} & -\frac{y^G}{(z^G)^2} \end{pmatrix} \mathbf{x}^p + f \begin{pmatrix} \frac{x^G}{z^G} \\ \frac{y^G}{z^G} \end{pmatrix}. \quad (2.2)$$

ここで、 $\mathbf{x}^G$  は特徴点の重心の座標であり、

$$\mathbf{x}^G = \frac{1}{P} \sum_{p=1}^P \mathbf{x}^p. \quad (2.3)$$

一般に3次元空間上の物体は、2次元平面にゆがんで投影される。半透視射影では、投影による像のゆがみの原因を二つの成分に分解して近似している。すなわち、ステップ1では、対象が投影面に対して傾いていることにより生

<sup>1)</sup>ベクトルはすべて縦ベクトルとする。

<sup>2)</sup>物体の重心を基準点とするのが一般的であるが、本稿では特徴点の重心を採用する(式(2.3)参照)。

<sup>3)</sup> $\mathbf{x}^T$  は、 $x$  の転置を表す。

<sup>4)</sup> $\mathbf{R}$  は実数全体を表す。

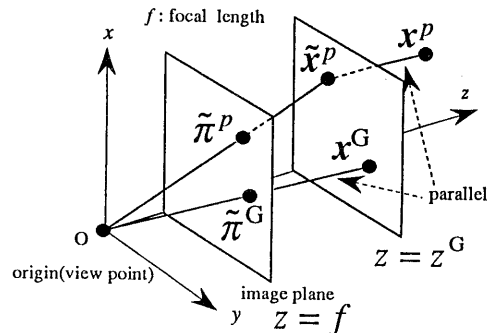


図1: 半透視射影の原理

じるゆがみと、対象の画像上の位置に起因したゆがみの一部を近似し、ステップ2では、視点と対象との距離によって生じるゆがみ、および対象の画像上の位置に起因したゆがみの残りの部分を近似している。なお、平面  $z = z^G$  上の点に関しては、半透視射影の下でも厳密に透視射影が施されている。

注意 2.1 式(2.2)で  $f \rightarrow \infty$  としても、正射影像に一致しないようにみえるが、座標系を  $z$  軸方向に  $-f$  だけ平行移動してから  $f \rightarrow \infty$  とすれば、正射影像に一致する。本稿で視点と原点が一致した座標系を採用するのは、回転運動を行列表現しやすくするためである。□

注意 2.2 透視射影の下では、3次元空間内の  $\mathbf{x}^p = (x^p, y^p, z^p)^T$  の像  $\boldsymbol{\pi}^p = (\pi_1^p, \pi_2^p)^T$  は、

$$\pi_1^p = \frac{x^p}{z^p} f, \quad \pi_2^p = \frac{y^p}{z^p} f \quad (2.4)$$

によって与えられ、これは線形な関係でない。しかし、半透視射影においては、式(2.2)のような線形な関係となり、取り扱い易い。□

## 2.2 半透視射影の意味

前節で説明した手続きによって実現される半透視射影は、どのような意味で透視射影の近似であるかを考察し、半透視射影像は透視射影像を基準点のまわりで Taylor 展開して1次近似したものであることを明らかにする。前節で導入した記号に従って、 $\mathbf{x}^p$  にある点  $p$  が、半透視射影によって  $\tilde{\boldsymbol{\pi}}^p$  に、透視射影によって  $\boldsymbol{\pi}^p$  に投影されているとする。半透視射影における基準点の座標を  $\mathbf{x}^G$  として、 $\delta \mathbf{x}^p = (\delta x^p, \delta y^p, \delta z^p)^T$  を

$$\begin{pmatrix} \delta x^p \\ \delta y^p \\ \delta z^p \end{pmatrix} := \begin{pmatrix} x^p - x^G \\ y^p - y^G \\ z^p - z^G \end{pmatrix} \quad (2.5)$$

と定義する。このとき、式(2.4)より、

$$\pi_1^p = \frac{x^G + \delta x^p}{z^G + \delta z^p} f, \quad \pi_2^p = \frac{y^G + \delta y^p}{z^G + \delta z^p} f \quad (2.6)$$

であるが、

$$|x^G| \gg |\delta x^p|, \quad |y^G| \gg |\delta y^p|, \quad |z^G| \gg |\delta z^p| \quad (2.7)$$

として、式(2.6)を  $\mathbf{x}^G$  のまわりで Taylor 展開して1次の項までかくと、

$$\pi_1^p = f \left[ \frac{x^G}{z^G} + \frac{1}{z^G} \delta x^p - \frac{x^G}{(z^G)^2} \delta z^p + \dots \right], \quad (2.8)$$

$$\pi_2^p = f \left[ \frac{y^G}{z^G} + \frac{1}{z^G} \delta y^p - \frac{y^G}{(z^G)^2} \delta z^p + \dots \right] \quad (2.9)$$

となる。一方、式(2.2)より、

$$\tilde{\pi}_1^p = f \left[ \frac{x^G + \delta x^p}{z^G} - \frac{x^G}{(z^G)^2} \delta z^p \right], \quad (2.10)$$

$$\tilde{\pi}_2^p = f \left[ \frac{y^G + \delta y^p}{z^G} - \frac{y^G}{(z^G)^2} \delta z^p \right] \quad (2.11)$$

であるから、 $\tilde{\boldsymbol{\pi}}^p$  は  $\boldsymbol{\pi}^p$  の1次近似であることがわかる。従って、視点と物体との距離に対して物体の大きさが十分小さい(式(2.7)参照)場合に半透視射影は透視射影の近似として有効である。

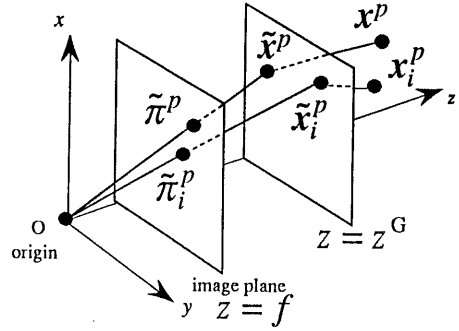


図2: 点  $p$  に変換  $i$  を施す前後の関係

## 3 問題の定式化

本稿で考察する問題を定式化する。まず、投影像に関して、次のことを仮定する:

- 物体の投影像は半透視射影によって得られる。
- 認識すべき物体の投影像に対して、特徴点が正確に抽出されている。
- 物体を構成する点の集合と、投影像上の特徴点の集合は、常に1対1に対応している。しかもこれらの対応関係は、物体の運動に依存しない。
- 投影像上の特徴点の集合は、投影像に依存せず一定で、各投影像間に対応関係は既知である。

さて、基準位置  $\mathbf{x}^p$  にある点  $p$  が、運動  $i$  をおこなって  $\mathbf{x}_i^p$  に移り、 $\tilde{\boldsymbol{\pi}}_i^p$  に像を結ぶとする(図2参照)。点  $p$  のアフィン運動は、次式で表される:

$$\mathbf{x}_i^p = R_i \mathbf{x}^p + \mathbf{t}_i, \quad (3.1)$$

ただし、

$$R_i \in \text{GL}(3), \quad \mathbf{t}_i \in \mathbb{R}^3. \quad (3.2)$$

ここで  $\text{GL}(3)$  は3次元実一般線形変換群を表す。点  $p$  の運動が剛体運動である場合には、式(3.2)の代わりに、

$$R_i \in \text{SO}(3), \quad \mathbf{t}_i \in \mathbb{R}^3 \quad (3.3)$$

となる。なお、 $\text{SO}(3)$  は3次回転群を表す。式(2.3)、(3.1)から

$$\mathbf{x}_i^G = R_i \mathbf{x}^G + \mathbf{t}_i \quad (3.4)$$

が成り立つ。さらに、

$$U_i := f \begin{pmatrix} 1/z_i^G & 0 & -x_i^G/(z_i^G)^2 \\ 0 & 1/z_i^G & -y_i^G/(z_i^G)^2 \end{pmatrix}, \quad (3.5)$$

$$V_i := f \begin{pmatrix} x_i^G/z_i^G \\ y_i^G/z_i^G \end{pmatrix} \quad (3.6)$$

とおけば、式(2.2)から

$$\tilde{\boldsymbol{\pi}}_i^p = U_i \boldsymbol{\pi}_i^p + V_i \quad (3.7)$$

と表せる。

認識すべき物体の運動  $i (i \in \{1, 2, \dots, I\})$  後に得られた投影像を  $\{\hat{\pi}_i^p\}_{p=1}^P$  で、また未知像を  $\{\hat{\pi}_*^p\}_{p=1}^P$  で表すことにする。我々の目標は、未知像  $\{\hat{\pi}_*^p\}_{p=1}^P$  が、認識すべき物体の、ある運動後に得られたものであるか否かを判定することである。まず、物体に対して許される運動(許容運動)の族を定める。物体に対して許される運動が異なれば、未知像が同じ物体から得られたものであるか否かの判定も異なるからである。本稿では、許容運動の族として、アフィン運動  $\mathcal{A}_a$  と剛体運動  $\mathcal{A}_r$  の二つを考える。それぞれは群をなし、

$$\mathcal{A}_a = \{(R, \mathbf{t}) \mid R \in \text{GL}(3), \mathbf{t} \in \mathbf{R}^3\}, \quad (3.8)$$

$$\mathcal{A}_r = \{(R, \mathbf{t}) \mid R \in \text{SO}(3), \mathbf{t} \in \mathbf{R}^3\} \quad (3.9)$$

である。  $\text{SO}(3) \subset \text{GL}(3)$  であるから、  $\mathcal{A}_r$  の要素は  $\mathcal{A}_a$  の要素のうち、ある付帯条件を満たすものであると見なすことができる。すなわち、式(3.9)を

$$\mathcal{A}_r = \{(R, \mathbf{t}) \in \mathcal{A}_a \mid R \in \text{SO}(3)\} \quad (3.10)$$

とかくことができる。以後、  $\mathcal{A}_r$  を、  $\mathcal{A}_a$  に付帯条件がついたものとして扱うことにする。また、  $(R_i, \mathbf{t}_i) \in \mathcal{A}$  を単に  $i \in \mathcal{A}$  と書く。なお、許容運動群を  $\mathcal{A}$  として、

$$\tilde{I}^p := \{\hat{\pi}_i^p \mid \hat{\pi}_i^p = U_i \mathbf{x}_i^p + V_i, \exists i \in \mathcal{A}\} \quad (3.11)$$

と定義する。

我々が解こうとしている問題は次のようになる。

**問題 3.1** 許容運動群  $\mathcal{A}$  が定められているとする。  $\{\hat{\pi}_*^p\}_{p=1}^P$  が与えられたとき、

$$\hat{\pi}_*^p \in \tilde{I}^p, \quad p \in \{1, 2, \dots, P\} \quad (3.12)$$

であるか否かを  $\{\hat{\pi}_i^p\}_{p=1}^P (i \in \{1, 2, \dots, I\})$  を用いて判定する手順を求めよ。 □

問題 3.1 を解くために運動群の表現を定めて議論を進める。求めるべき判定手順は、運動群の表現に依存するからである。

本稿(第5節)では、三つの表現を取り扱うが、いずれも  $\mathcal{A}$  の要素を  $\mathcal{A}$  の要素に関する 1 次式で表現する。これは投影像を構成する点の間の関係を 1 次式で表すためである。

## 4 投影像の取り扱い

### 4.1 投影面上の座標系の変更

式(2.2),(2.3)から、

$$\hat{\pi}_i^G = \frac{1}{P} \sum_{p=1}^P \hat{\pi}_i^p \quad (4.1)$$

となるが、これは  $P$  個の  $\hat{\pi}_i^p (p = 1, 2, \dots, P)$  から容易に計算できる。そこで、  $\hat{\pi}_i^p$  を  $\hat{\pi}_i^G$  からの増分で表現した

$$\rho_i^p := \hat{\pi}_i^p - \hat{\pi}_i^G \quad (4.2)$$

を定義する。また  $\tilde{I}^p$  に対応して、

$$\tilde{I}^p := \{\rho_i^p \mid \rho_i^p = \hat{\pi}_i^p - \hat{\pi}_i^G, \hat{\pi}_i^p \in \tilde{I}^p\} \quad (4.3)$$

とする。  $\rho_i^p$  は  $\hat{\pi}_i^p$  から容易に計算できるので、以下では  $\hat{\pi}_i^p$  の代わりに  $\rho_i^p$  を用いて議論する。つまり、  $\{\rho_i^p\}_{p=1}^P (i = 1, 2, \dots, I)$ 、  $\{\rho_*^p\}_{p=1}^P$  をそれぞれ蓄積像、未知像とみなすことにする。

ところで、式(4.2)を書き直すと、

$$\rho_i^p = \frac{f}{(z_i^G)^2} Q [\mathbf{x}_i^G] \delta \mathbf{x}_i^p \quad (4.4)$$

となる。ただし、

$$Q := \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad (4.5)$$

であり、一般に  $\mathbf{x} = (x_1, x_2, x_3)^T$  に対して、

$$[\mathbf{x}] := \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix} \quad (4.6)$$

と定義する。

$[\mathbf{x}]$  は次の性質を満たす。

**補題 4.1**  $\mathbf{x}, \mathbf{y} \in \mathbf{R}^3$  とし、  $R \in \text{GL}(3)$  とするとき<sup>5</sup>、

$$[\mathbf{x} + \mathbf{y}] = [\mathbf{x}] + [\mathbf{y}], \quad (4.7)$$

$$R^T [\mathbf{R}\mathbf{x}] R = (\det R) [\mathbf{x}]. \quad \square \quad (4.8)$$

### 4.2 蓄積像と未知像

蓄積像  $\{\rho_i^p\}_{p=1}^P (i \in \{1, 2, \dots, I\})$  に対して、

$$\mathbf{x}_i^p = R_i \mathbf{x}^p + \mathbf{t}_i, \quad p \in \{1, 2, \dots, P\} \quad (4.9)$$

が成立しているとする。これより、

$$\delta \mathbf{x}_i^p = R_i \delta \mathbf{x}^p, \quad (4.10)$$

$$\mathbf{x}_i^G = R_i \mathbf{x}^G + \mathbf{t}_i \quad (4.11)$$

となるので、補題 4.1 を用いれば、

$$[\mathbf{x}_i^G] \delta \mathbf{x}_i^p = (\det R_i) R_i^{-T} [\mathbf{x}^G + R_i^{-1} \mathbf{t}_i] \delta \mathbf{x}^p \quad (4.12)$$

が得られる<sup>6</sup>。これを式(4.4)に代入して、

$$\rho_i^p = f (\tilde{R}_i [\mathbf{x}^G] + \tilde{T}_i) \delta \mathbf{x}^p \quad (4.13)$$

となる。ただし、

$$\tilde{R}_i := \frac{\det R_i}{(z_i^G)^2} Q R_i^{-T}, \quad (4.14)$$

$$\tilde{T}_i := \frac{\det R_i}{(z_i^G)^2} Q R_i^{-T} [R_i^{-1} \mathbf{t}_i] \quad (4.15)$$

はともに  $2 \times 3$  の行列である。蓄積像は、式(4.13)によって特徴づけられる。

<sup>5</sup>  $\det R$  は正方形行列  $R$  の行列式を表す。

<sup>6</sup>  $R^{-T}$  は、  $(R^T)^{-1}$  を表す。

注意 4.1  $R_i \in \text{GL}(3)$  のときは,  $\hat{R}_i, \hat{T}_i$  はともに任意の行列となり得る.  $R_i \in \text{SO}(3)$  のときは,

$$\det R_i = 1, \quad R_i^{-T} = R_i \quad (4.16)$$

だから,  $\hat{R}_i$  の第 1 行ベクトルは  $R_i$  の第 2 行ベクトルを定数  $c$  倍したものととなり,  $\hat{R}_i$  の第 2 行ベクトルは  $R_i$  の第 1 行ベクトルを  $-c$  倍したものととなる. つまり,  $\hat{R}_i$  の二つの行ベクトルは直交し, ノルムは等しい.  $\square$

一方, 未知像  $\{\rho_*^p\}_{p=1}^P$  に対して,

$$\mathbf{x}_*^p = R_* \mathbf{x}^p + \mathbf{t}_*, \quad p \in \{1, 2, \dots, P\} \quad (4.17)$$

が成立しているとする, と同様に

$$\hat{R}_* := \frac{\det R_*}{(z^G)^2} Q R_*^{-T}, \quad (4.18)$$

$$\hat{T}_* := \frac{\det R_*}{(z^G)^2} Q R_*^{-T} [R_*^{-1} \mathbf{t}_*] \quad (4.19)$$

とおいて,

$$\rho_*^p = f(\hat{R}_* [\mathbf{x}^G] + \hat{T}_*) \delta \mathbf{x}^p. \quad (4.20)$$

## 5 投影像の表現

運動群の表現を定め, 式 (4.20) で表された未知像を式 (4.13) で表された蓄積像で表現することを考える. 運動群の 3 通りの表現を考え, それに基づいて議論を進める. 式 (4.13), (4.20) から,  $\hat{R}$  と  $\hat{T}$  の表現に注目すればよいことがわかる. そこで,  $\hat{R}$  と  $\hat{T}$  を横に並べて  $2 \times 6$  の行列を定義する:

$$\hat{C} := (\hat{R} | \hat{T}), \quad \hat{C}_i := (\hat{R}_i | \hat{T}_i), \quad \hat{C}_* := (\hat{R}_* | \hat{T}_*). \quad (5.1)$$

### 5.1 線形結合による投影像の表現

式 (5.1) の  $\hat{C}_*$  は  $2 \times 6$  の行列であるが, その全要素を並べれば  $12$  次元空間  $\mathbf{R}^{12}$  のベクトルとみなすことができる. したがって,  $\{\hat{C}_i\}_{i=1}^I$  が  $\mathbf{R}^{12}$  と同型な空間を張れば, 任意の  $\hat{C}_*$  に対して

$$\hat{C}_* = \sum_{i=1}^I \lambda_i \hat{C}_i \quad (5.2)$$

を満たす  $\{\lambda_i\}_{i=1}^I$  が存在する. 式 (5.2) を  $\hat{R}, \hat{T}$  を用いて書き直すと,

$$\hat{R}_* = \sum_{i=1}^I \lambda_i \hat{R}_i, \quad \hat{T}_* = \sum_{i=1}^I \lambda_i \hat{T}_i \quad (5.3)$$

となり, これは  $\hat{R}_*, \hat{T}_*$  の表現を与えている. これを式 (4.20) に代入すれば,

$$\rho_*^p = f \left\{ \sum_{i=1}^I \lambda_i \hat{R}_i [\mathbf{x}^G] + \sum_{i=1}^I \lambda_i \hat{T}_i \right\} \delta \mathbf{x}^p \quad (5.4)$$

$$= \sum_{i=1}^I \lambda_i f(\hat{R}_i [\mathbf{x}^G] + \hat{T}_i) \delta \mathbf{x}^p \quad (5.5)$$

$$= \sum_{i=1}^I \lambda_i \rho_i^p \quad (5.6)$$

となる.

定理 5.1 許容運動群を  $\mathcal{A}_n$  (アフィン運動群) とし,  $\{\hat{C}_i\}_{i=1}^{12}$  は線形独立であるとする. このとき,  $\forall \rho_*^p \in \hat{I}^P$  に対して

$$\rho_*^p = \sum_{i=1}^{12} \lambda_i \rho_i^p \quad (5.7)$$

満たす  $p$  に依らない  $12$  個の実数  $\lambda_i$  ( $i = 1, 2, \dots, 12$ ) が存在する.  $\square$

とくに許容運動群が  $\mathcal{A}_r$  (剛体運動群) である場合には,  $\hat{R}_*$  の第 1 行ベクトルと第 2 行ベクトルとは直交し, 両者のノルムが等しい (注意 4.1 参照) ので,

$$\hat{R}_i = \begin{pmatrix} (\hat{r}_1^i)^T \\ (\hat{r}_2^i)^T \end{pmatrix} \quad (5.8)$$

とかけば,  $\lambda_i$  には次の付帯条件が付随する<sup>7</sup>:

$$\left\| \sum_{i=1}^{12} \lambda_i \hat{r}_1^i \right\| = \left\| \sum_{i=1}^{12} \lambda_i \hat{r}_2^i \right\|, \quad (5.9)$$

$$\left( \sum_{i=1}^{12} \lambda_i \hat{r}_1^i \right) \cdot \left( \sum_{i=1}^{12} \lambda_i \hat{r}_2^i \right) = 0. \quad (5.10)$$

したがって, 次の定理が得られる.

定理 5.2 許容運動群が  $\mathcal{A}_r$  である場合には, 定理 5.1 における実数  $\{\lambda_i\}_{i=1}^{12}$  は式 (5.9), (5.10) を満たす.  $\square$

注意 5.1 定理 5.1 より, 同じ物体から得た投影像上の特徴点は式 (5.7) を満たすことが保証されるが, 式 (5.7) を満たす特徴点を持つ投影像は必ず同じ物体から得られたものであるという保証はない. しかし, 特徴点の数が十分多ければ異なる物体から得られた投影像上の特徴点が全て式 (5.7) を満たす確率は十分小さいと考えられる. 次の二つの節に述べる定理に対してもこれと同じことがいえる.  $\square$

### 5.2 1 次式による投影像の表現 I

$\hat{C}$  の各行ベクトルは 6 次元ベクトルであるので,

$$\hat{C}_i = \begin{pmatrix} (\hat{c}_1^i)^T \\ (\hat{c}_2^i)^T \end{pmatrix} \quad (5.11)$$

と書いたとき,  $\{\hat{c}_1^i\}_{i=1}^I$  が  $\mathbf{R}^6$  と同型な空間を張り, かつ,  $\{\hat{c}_2^i\}_{i=1}^I$  が  $\mathbf{R}^6$  と同型な空間を張れば,

$$\hat{C}_* = \sum_{i=1}^I D_i \hat{C}_i \quad (5.12)$$

を満たす  $D_i := \begin{pmatrix} \mu_i & 0 \\ 0 & \nu_i \end{pmatrix}$  ( $i = 1, 2, \dots, I$ ) が存在する.  $\hat{R}, \hat{T}$  を用いて表すと,

$$\hat{R}_* = \sum_{i=1}^I D_i \hat{R}_i, \quad \hat{T}_* = \sum_{i=1}^I D_i \hat{T}_i \quad (5.13)$$

となり, これも  $\hat{R}_*, \hat{T}_*$  の表現を与えている. 従って,

$$\rho_*^p = f \left\{ \sum_{i=1}^I D_i \hat{R}_i [\mathbf{x}^G] + \sum_{i=1}^I D_i \hat{T}_i \right\} \delta \mathbf{x}^p \quad (5.14)$$

$$= \sum_{i=1}^I D_i \rho_i^p \quad (5.15)$$

となる.

<sup>7</sup> $\|\mathbf{x}\|$  は, ベクトル  $\mathbf{x}$  のユークリッドノルムを表す.

定理 5.3 許容運動群を  $\mathcal{A}_6$  とし,  $\{\tilde{c}_1^i\}_{i=1}^6, \{\tilde{c}_2^i\}_{i=1}^6$  がそれぞれ線形独立であるとする. このとき,  $\forall \rho_*^p \in \tilde{H}^p$  に対して

$$\rho_*^p = \sum_{i=1}^6 \begin{pmatrix} \mu_i & 0 \\ 0 & \nu_i \end{pmatrix} \rho_i^p \quad (5.16)$$

を満たす  $p$  に依らない 12 個の実数  $\mu_i, \nu_i (i = 1, 2, \dots, 6)$  が存在する.  $\square$

定理 5.4 許容運動群が  $\mathcal{A}_6$  である場合には, 定理 5.3 における実数  $\{\mu_i, \nu_i\}_{i=1}^6$  は次式を満たす:

$$\left\| \sum_{i=1}^6 \mu_i \tilde{r}_1^i \right\| = \left\| \sum_{i=1}^6 \nu_i \tilde{r}_2^i \right\|, \quad (5.17)$$

$$\left( \sum_{i=1}^6 \mu_i \tilde{r}_1^i \right) \cdot \left( \sum_{i=1}^6 \nu_i \tilde{r}_2^i \right) = 0. \quad \square \quad (5.18)$$

式 (5.3) による表現の代わりに式 (5.13) による表現を用いれば, 任意の投影像を適当な 6 枚の投影像の 1 次式によって表現できることがわかった.

注意 5.2 本節で扱った投影像の表現は, Ullman-Basri [6] が用いた表現に倣ったものである.  $\square$

### 5.3 1 次式による投影像の表現 II

6 個の 6 次元ベクトル  $\{\tilde{c}_1^i, \tilde{c}_2^i\}_{i=1}^3$  が  $\mathbf{R}^6$  と同型な空間を張れば,

$$\tilde{C}_* = \sum_{i=1}^3 \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \tilde{C}_i \quad (5.19)$$

を満たす 12 個の実数  $a_i, b_i, c_i, d_i (i = 1, 2, 3)$  が存在する.

$$M_i := \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \quad (5.20)$$

とおいて, 式 (5.19) を  $\tilde{R}_i, \tilde{T}_i$  でかくと次のようになり, やはり  $\tilde{R}_*, \tilde{T}_*$  の表現を与えている:

$$\tilde{R}_* = \sum_{i=1}^3 M_i \tilde{R}_i, \quad \tilde{T}_* = \sum_{i=1}^3 M_i \tilde{T}_i. \quad (5.21)$$

したがって,

$$\rho_*^p = f \left\{ \sum_{i=1}^3 (M_i \tilde{R}_i) [\mathbf{x}^G] + \sum_{i=1}^3 (M_i \tilde{T}_i) \right\} \delta \mathbf{x}^p \quad (5.22)$$

$$= \sum_{i=1}^3 M_i \rho_i^p \quad (5.23)$$

となり, つぎの定理を得る.

定理 5.5 許容運動群を  $\mathcal{A}_3$  とし,  $\{\tilde{c}_1^i, \tilde{c}_2^i\}_{i=1}^3$  は線形独立であるとする. このとき,  $\forall \rho_*^p \in \tilde{H}^p$  に対して

$$\rho_*^p = \sum_{i=1}^3 \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \rho_i^p \quad (5.24)$$

を満たす  $p$  に依らない 12 個の実数  $a_i, b_i, c_i, d_i (i = 1, 2, 3)$  が存在する.  $\square$

定理 5.6 許容運動群が  $\mathcal{A}_3$  である場合には, 定理 5.5 における実数  $a_i, b_i, c_i, d_i (i = 1, 2, 3)$  は次式を満たす:

$$\left\| \sum_{i=1}^3 (a_i \tilde{r}_1^i + b_i \tilde{r}_2^i) \right\| = \left\| \sum_{i=1}^3 (c_i \tilde{r}_1^i + d_i \tilde{r}_2^i) \right\|, \quad (5.25)$$

$$\sum_{i=1}^3 (a_i \tilde{r}_1^i + b_i \tilde{r}_2^i) \cdot \sum_{i=1}^3 (c_i \tilde{r}_1^i + d_i \tilde{r}_2^i) = 0. \quad \square \quad (5.26)$$

以上によって, 許容運動群がアフィン運動群・剛体運動群いずれの場合でも, 式 (5.21) による表現を用いれば任意の投影像を適当な 3 枚の投影像の 1 次式によって表現できることが示された.

## 6 アルゴリズムと計算機実験

### 6.1 アルゴリズム

5.3 節で述べた表現を用いたとき, 蓄積像の数が増えることも少なくないので, これに対するアルゴリズムを呈示する (他の表現についても同様である).

5.3 節の議論から各  $p (p = 1, 2, \dots, P)$  に対して,  $\rho_*^p$  は  $\{\rho_i^p\}_{i=1}^3$  の 1 次式で表されることがわかったので, 実際未知像  $\{\tilde{\pi}_*^p\}_{p=1}^P$  が与えられたときは,  $\{\rho_*^p\}_{p=1}^P$  を計算し, 式 (5.24) を  $\{a_i, b_i, c_i, d_i\}_{i=1}^3$  に関する過剰決定系の線形方程式と見て最小 2 乗法を適用し, 残差が (殆ど) 0 になるか否かを判定すればよい.  $p \in \{1, 2, \dots, P\}$  に対して,

$$f_p := \left\| \sum_{i=1}^3 \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \rho_i^p - \rho_*^p \right\|^2, \quad (6.1)$$

また

$$g_1 := \left( \left\| \sum_{i=1}^3 (a_i \tilde{r}_1^i + b_i \tilde{r}_2^i) \right\| - \left\| \sum_{i=1}^3 (c_i \tilde{r}_1^i + d_i \tilde{r}_2^i) \right\| \right)^2, \quad (6.2)$$

$$g_2 := \left( \sum_{i=1}^3 (a_i \tilde{r}_1^i + b_i \tilde{r}_2^i) \cdot \sum_{i=1}^3 (c_i \tilde{r}_1^i + d_i \tilde{r}_2^i) \right)^2 \quad (6.3)$$

と定義し, 剛体運動を考える場合には

$$h := \sum_{p=1}^P f_p + g_1 + g_2, \quad (6.4)$$

アフィン運動を考える場合には

$$h := \sum_{p=1}^P f_p \quad (6.5)$$

と定義する.

以下の手続きを施すことによって未知像が同じ物体から得られたものであるか否かを判定できる.

#### アルゴリズム

1.  $\tilde{\pi}_*^G$  を計算する (式 (4.1)).
2. 各  $p (p \in \{1, 2, \dots, P\})$  に対して,  $\rho_*^p$  を計算する (式 (4.2)).

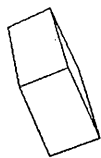


図 3: 認識すべき平行六面体

3. 評価関数  $h(\{a_i, b_i, c_i, d_i\}_{i=1}^3)$  を (殆ど) 0 にするような  $\{a_i, b_i, c_i, d_i\}_{i=1}^3$  が存在するかどうかを判定する:

- 存在する  $\implies$  同じ物体.
- 存在しない  $\implies$  異なる物体. □

注意 6.1 理論的には “ $h = 0$ ” によって Step 3 の判定をおこなえばよいが, 実際には計算による丸め誤差が存在するため, “ $h \approx 0$ ” によって判定すればよい. □

## 6.2 計算機実験

前節のアルゴリズムによって計算機実験を行った結果を示す. 投影像は計算機によって半透視射影に基づいて作成した. なお, 許容運動をアフィン運動とし, 焦点距離  $f$  の値を 1.0 とした.

認識すべき物体を平行六面体 (図 3) とした. この平行六面体は基準となる位置で 8 個の頂点, (5.00, 6.50, 10.00), (5.80, 5.96, 9.73), (6.16, 7.04, 9.37), (5.36, 7.58, 9.64), (5.45, 6.05, 11.35), (6.25, 5.51, 11.08), (6.61, 6.59, 10.72), (5.81, 7.13, 10.99) によって構成されている. 平行六面体

表 1: 蓄積像を与える運動

	rotation		translation ( $x, y, z$ )
	axis	degree	
(a)	$x$	$10^\circ$	(-0.50, 2.00, 0.00)
(b)	$x$	$15^\circ$	(2.50, 4.00, -1.00)
(c)	$y$	$-10^\circ$	(-4.50, 2.50, 0.50)

の観測できる 7 個の頂点を特徴点とした. 図 4 に示す 3 枚投影像を蓄積像とし, これを用いて未知像の判定を行った. 図 4 の各投影像は表 1 の運動<sup>8</sup>後に得られたものである.

<sup>8</sup>運動はすべて, 回転につづいて並進がおこるとする.

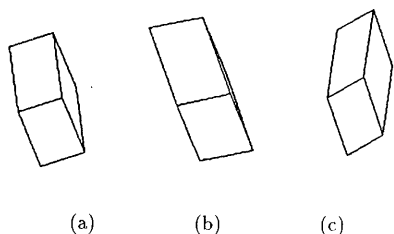


図 4: 図 3 の平行六面体に対する蓄積像

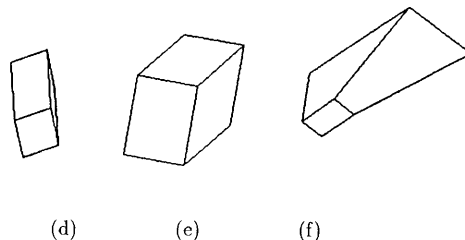


図 5: 未知像

る. 図 5 に示す三つの未知像に対して前節で呈示したアルゴリズムを実行した. その結果を表 2 (左側) に示す. 投影像の座標の値は  $10^{-2}$  程度の大きさであるので, 判定  $h \approx 0$  のしきい値を  $1.0 \times 10^{-5}$  に設定した. これにより, (d), (e) は同じ物体の投影像であり, (f) は同じ物体の投影像ではないと判定できる.

表 2: 評価関数  $h$  の最小値

	$h$ (半透視射影)	$h$ (透視射影)
(d)	$5.63 \times 10^{-14}$	$2.86 \times 10^{-7}$
(e)	$5.99 \times 10^{-13}$	$4.75 \times 10^{-6}$
(f)	$1.14 \times 10^{-2}$	$2.32 \times 10^{-3}$

事実, (d) は, 図 3 の平行六面体を  $x$  軸周りに  $30^\circ$  回転した後  $y$  軸周りに  $-30^\circ$  回転し, さらに (1.00, 4.50, -0.50) だけ並進して得られた投影像であり, (e) は, 図 3 の平行六面体を  $x$  軸,  $y$  軸,  $z$  軸の順にそれぞれ  $5^\circ$ ,  $20^\circ$ ,  $30^\circ$  回転した後, (-3.00, -5.00, 0.00) だけ並進して得られた投影像である. これに対して (f) は, ある四角錐台の投影像である.

同じ条件下で, 透視射影によって投影像が得られた場合の結果を表 2 (右側) に示す. 図 6, 図 7 はそれぞれ, 半透視射影の場合の図 4, および図 5 に対応する. 半透視射影の場合と同じしきい値によって, (d), (e) は同じ物体の投影像であり, (f) は同じ物体の投影像ではないと判定できる. このように, 前節で呈示したアルゴリズムによって, 未知像が同じ物体から得られたものであるか否かを正しく判定できた.

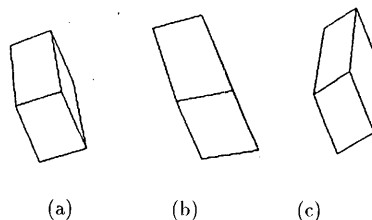


図 6: 図 3 の平行六面体に対する蓄積像 (透視射影)

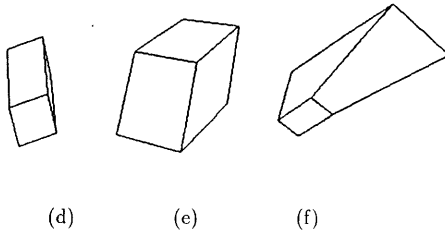


図 7: 未知像 (透視射影)

## 7 結論

半透視射影によって投影像が得られた場合について、少数の投影像をモデルとするにより、任意の物体を認識できることを示した。各投影像上の特徴点間の対応が既知であり誤差がないとすれば、物体が3次元空間内を剛体運動する場合、アフィン運動する場合、いずれの場合にも適当な3枚の投影像があれば、それらの1次式によって同じ物体の任意の投影像を表現できる。したがって、物体認識問題は、未知像をモデル像の1次式で表せるか否かを判定する問題に帰着される。これにより、蓄積像に対する前処理は不要となり、最小2乗法を用いれば容易に物体認識が行えることになる。

今後は、本稿で提案したアルゴリズムを適用したときの丸め誤差に関する検討、および、半透視射影による近似誤差に関する検討を行う予定である。

謝辞 本研究の機会を与えてくださった(株)ATR 視聴覚機構研究所 淀川英司社長に深く感謝いたします。また、研究を遂行するにあたって多大なる助言をいただいた、同研究所 上野圭一視覚研究室長ならびに、(株)ATR 人間情報通信研究所 川戸慎二郎氏に感謝いたします。

## 参考文献

- [1] J. Aloimonos: Shape from Texture, *Biological Cybernetics*, 58, 5, 345-360(1989).
- [2] P. J. Besl and R. C. Jain: Three-Dimensional Object Recognition, *ACM Computing Surveys*, 1, 17, 75-145(1985).
- [3] Y. Ohta, K. Maenobu and T. Sakai: Obtaining Surface Orientation from Texels under Perspective Projection, *Proc. of the 7th IJCAI*, 746-751, 1981.
- [4] T. Poggio: *3D Object Recognition: On a Result of Basri and Ullman*, IRST Technical Report, 9005-03, Trento, Italy, 1990.
- [5] A. Sugimoto and K. Murota: 3D Object Recognition by Combination of Perspective Images, *Proc. of SPIE'93*, to appear.
- [6] S. Ullman and R. Basri: Recognition by Linear Combinations of Models, *IEEE Trans. on Pattern Anal-*