論　文

3次元復元のための座標回転不変量の構成

正員　金谷　健一[†]

Coordinate Rotation Invariance of Image Characteristics for 3D Shape and Motion Recovery

Ken-ichi KANATANI[†], *Member*

あらまし　物体の投影像からその形状や運動を決定するには，画面上で計測した画像の特徴量と物体の形状や運動を指定するパラメータとを結びつける方程式を導いて，それを解けばよいが，多くは非線形方程式となるので解析的には解きにくい．本論文では，画像上には本来固有の座標系はなく，どの座標系も同等であることに着目して，画像の特徴量を組み合わせて座標系の回転に不変な性質をもつ量（不変量）を構成する．こうすれば画像の特徴量の幾何学的意味が明らかになると共に，解析的な解も自然に得られることも多い．このことをオプティカルフローの解析とテクスチャによる曲面の復元を例にとって示す．

1. まえがき

画像理解，画像認識，コンピュータビジョンなどと呼ばれる分野の目的のひとつは物体の3次元形状や運動を2次元投影画像から復元することである．これは一般的には次のように定式化できる*．

物体にモデルを仮定して，その形状や運動を少数のパラメータ $a_1, \cdots, a_n$ で指定できるとする．たとえば平面であればその上の一点の座標と法線方向とで指定できる．運動はある基準点における並進速度とその回りの回転速度とで表せる．このように物体の形状や運動を指定するパラメータを「物体パラメータ」と呼ぼう．

一方，画像上で計測したデータが $c_1, \cdots, c_m$ であるとする．これを画像の「特徴量」と呼ぼう．これは物体表面の反射強度のように画像の濃淡レベルから直接に得られることもあるし，画像処理を行った後の輪郭線やテクスチャやオプティカルフローを特徴づける量のこともある．よく画像の特徴量として何に着目するかで"shape from ○○○"と呼ばれている（○○○としては shading，texture，motion 等）．

物体にモデルをあてはめ，物体パラメータの値を仮定し，投影のカメラモデルが与えられているとすると，どのような画像の特徴量が得られるかが理論的に計算できる．すなわち特徴量 $c_1, \cdots, c_m$ が物体パラメータ $a_1, \cdots, a_n$ の関数として与えられる．

$$c_i = F_i(a_1, \cdots, a_n), \quad i=1, \cdots, m \qquad (1)$$

従って，特徴量 $c_1, \cdots, c_m$ の計測値を用い，式(1)を未知数 $a_1, \cdots, a_n$ に関する連立方程式とみなして解けば物体パラメータが求まる．その意味で(1)を「復元方程式」と呼ぶ．

多くの場合，このような復元方程式は非線形連立方程式となり，解析的に解くのが難しい．以下ではこのような場合に解析解を求めるための，万能ではないが極めて有力な方法を示す．その方針は次の通りである．「任意の」非線形方程式を一般的に解く方法は存在しないにしても，(1)は考えている問題の幾何学的性質を反映した何らかの「構造」をもっている．以下で着目するのは，画像面上には本来固有な座標系はなく，どの座標系も同等であるという事実である．そこで，画像の特徴量を組み合わせて座標系の回転に不変な性質をもつ量（不変量）を構成する．そして復元方程式をそのような不変量で表せば，解析解が自然に得られることが多い．また，不変量は何らかの幾何学的に意味のある性質に対応しているので（「Weylのテーゼ」：後述），画像の特徴量や復元方程式の幾何学的意味が大変理解

† 群馬大学工学部情報工学科，桐生市
Department of Computer Science, Gunma University, Kiryu-shi, 376 Japan

* これが唯一の定式化ではない．金谷(4)参照．ここに示したものをそこでは「2Dの立場」と呼んでいる．

しやすくなる．このことをオプティカルフローの解析とテクスチャによる曲面の復元を例にとって示す．

2. 座標系の回転と既約表現

画像面上にとったある座標系に関して特徴量 $c_1, \cdots, c_m$ を計測したとする．ただし座標原点は画像の中心，すなわちカメラ光軸に対応する点に一致しているとする．このとき x 軸あるいは y 軸の方向は特定の意味をもっておらず，どの方向にとっても同等である．そこで xy 座標系を原点の回りに角度 θ (反時計回りを正とする) だけ回転した $x'y'$ 座標系をとって，前と同じ特徴量を計測したところ，値が $c_1', \cdots, c_m'$ であったとする．この新しい値が古い値 $c_1, \cdots, c_m$ の線形結合で表せる場合を考えよう．すなわち

$$\begin{bmatrix} c_1' \\ \vdots \\ c_m' \end{bmatrix} = \begin{bmatrix} & & \\ & T(\theta) & \\ & & \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \tag{2}$$

あるいは略記して $\boldsymbol{c}' = T(\theta)\boldsymbol{c}$ と表せるとする．

明らかに係数行列 $T(\theta)$ は 2 次元回転群 SO(2) の表現，すなわち SO(2) から行列の積に関する群への準同型である．実際，$x'y'$ 座標系を更に角度 θ' だけ回転した $x''y''$ 座標系を考えると $\boldsymbol{c}'' = T(\theta')\boldsymbol{c}'$ の関係を得る．従って $\boldsymbol{c}'' = T(\theta')T(\theta)\boldsymbol{c}$ であるが，$x''y''$ 座標系は xy 座標系を角度 $\theta'+\theta$ だけ回転したものにほかならないから $\boldsymbol{c}'' = T(\theta'+\theta)\boldsymbol{c}$，すなわち $T(\theta')T(\theta) = T(\theta'+\theta)$ が得られる．

画像の個々の特徴量は個々の計測法に従って定まるもので，それぞれが何らかの本質的な意味をもっているとは限らない．たとえば特徴量 c_1，c_2 のかわりに $C_1 = c_1 + c_2$，$C_2 = c_1 - c_2$ を用いてもよい．特徴 $\{c_1, c_2\}$ と特徴 $\{C_1, C_2\}$ は互いに等価であり，同一の特性を記述している．そこで特徴量 $c_1, \cdots, c_m$ に線形結合をほどこして新しい特徴量 $C_1, \cdots, C_m$ をつくったとする†．この新しい特徴量に対して角度 θ の座標系の回転をほどこすと次のようになったとしよう．

$$\begin{bmatrix} C_1' \\ \vdots \\ C_l' \\ C_{l+1}' \\ \vdots \\ C_m' \end{bmatrix} = \left[\begin{array}{c|c} * & O \\ \hline O & * \end{array}\right] \begin{bmatrix} C_1 \\ \vdots \\ C_l \\ C_{l+1} \\ \vdots \\ C_m \end{bmatrix} \tag{3}$$

これは特徴 $\{C_1, \cdots, C_l\}$ と特徴 $\{C_{l+1}, \cdots, C_m\}$ とが互いに無関係に変換していることを示している．これを，「特徴 $\{C_1, \cdots, C_l\}$ と特徴 $\{C_{l+1}, \cdots, C_m\}$ とは画像の別々の性質を記述している」と解釈してよいであろう．数学的にいえば，これは 2 次元回転群 SO(2) の表現 $T(\theta)$ を二つの表現の直和に「簡約」したことにほかならない．従って簡約できる表現，すなわち「可約表現」を作る特徴量は画像の別々の性質を同時に記述しているといえる．

更に同様のことを $\{C_1, \cdots, C_l\}$ と $\{C_{l+1}, \cdots, C_m\}$ とに行って次々に簡約していくと最終的な形

$$\begin{bmatrix} C_1' \\ \vdots \\ C_m' \end{bmatrix} = \begin{bmatrix} * & & \\ & \boxed{*} & \\ & & \ddots \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix} \tag{4}$$

が得られたとする．各小区画はこれ以上簡約できない表現，すなわち「既約表現」を作っている．このように既約表現の直和に簡約できるような表現は「完全可約」であると呼ばれている．

既約表現を作っている特徴量は変換すれば必ず入り混じって分離できないのであるから，これは本質的に「ひとつの」性質を記述するものと解釈してよいであろう．それに対して既約でない場合は複数の異なった性質を同時に記述しているとみなせる．このように考えれば「ある性質をそれを構成している個々の性質に分離する」という直観的であいまいな概念を，表現の既約分解という数学的によく知られた概念に対応させることが可能となる．これは Weyl[14] が行った議論である．彼は，およそある観測量が一つの物理的性質であるといえるためには，それがその現象の本質を変えないような変換の群の既約表現に対応していなければならないと述べて，量子力学に群論を導入した[15]．以下ではこのような考え方を「Weyl のテーゼ」と呼ぶことにしよう．

3. 不変量とウェイト

よく知られているように，2 次元回転群 SO(2) の場合はすべての表現は完全可約であり，すべての既約表現は 1 次元である．従って SO(2) の表現を作るような特徴量 $c_1, \cdots, c_m$ が与えられれば，適当な線形結合を作ることによって，それぞれが独立に変換するような新しい特徴量 $C_1, \cdots, C_m$，$C_i' = T_i(\theta)C_i$，$i=1, \cdots, m$，を必ず作ることができる．各 $T_i(\theta)$ は 1 次元の表現であるから $T_i(\theta')T_i(\theta) = T_i(\theta'+\theta)$ を満たさな

† はじめの特徴量は実数であると仮定するが，線形結合の係数として複素数も許すことにする．従って $C_1, \cdots, C_m$ は複素数であるかもしれない．また両者は等価である，すなわち変換の係数行列は正則であるとする．

ければならないし，周期 2π の周期関数であって $T_i(0)=1$，$T_i(\theta+2\pi)=T_i(\theta)$ でなければならない．これから直ちに，ある整数 n があって $T_i(\theta)=e^{-in\theta}$ となっていなければならないことがわかる（i は虚数単位）．この整数 n をこの特徴量の「ウェイト」と呼ぼう†．そして，ウェイト0の特徴量，すなわち値を変えないものを「絶対不変量」，0でないウェイト n をもつものを「ウェイト n の相対不変量」と呼び，両方を合わせて単に「不変量」と呼ぼう．

以上のことをまとめれば次のようになる．式(2)のような変換を受ける特徴量 $c_1, \cdots, c_m$ に対して適当な線形結合をほどこして，新しい特徴量 $C_1, \cdots, C_m$ を

$$\begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix} = \begin{bmatrix} & & \\ & P & \\ & & \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \tag{5}$$

のように作って（P は正則行列），新しい特徴量 $C_1, \cdots, C_m$ の変換が次のように分解するようにすることができる．

$$\begin{bmatrix} C_1' \\ \vdots \\ C_m' \end{bmatrix} = \begin{bmatrix} & & \\ & P & \\ & & \end{bmatrix} \begin{bmatrix} & & \\ & T(\theta) & \\ & & \end{bmatrix} \begin{bmatrix} & & \\ & P^{-1} & \\ & & \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix}$$

$$= \begin{bmatrix} e^{-in_1\theta} & & \\ & \ddots & \\ & & e^{-in_m\theta} \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix} \tag{6}$$

すなわち，定数行列 P によって式(2)の係数行列 $T(\theta)$ を θ の値にかかわらず同時に対角化することができる．これは，2次元回転群 SO(2) はコンパクトであるから完全可約であり，また可換群であり，Schur の補題より既約表現はすべて1次元になるからである[(1),(8)]．各不変量は既約表現に対応しているから Weyl のテーゼにより，それぞれが何か特定の幾何学的な性質を表している††．

4. スカラ，ベクトル，テンソル

座標を回転しても変化しない特徴量

$$c' = c \tag{7}$$

を「スカラ」という．式(7)は自明な表現，すなわち「恒等表現」を定義している．スカラはそれ自身で絶対不変量である．

† $e^{-in\theta}$ ではなく $e^{in\theta}$ とおいた n をウェイトと呼ぶこともある．対象物の回転を考えるときはそのほうが便利である．しかし，ここでは座標系のほうを回転させるので $e^{-in\theta}$ とするほうが便利である．

†† 同じウェイトの不変量が二つ以上ある場合には，それらの線形結合をとる自由度が残って，分解は一意的でない（縮退）．このときは別の群を作用させて分解を決める（量子力学でいう摂動法）などが考えられる．

特徴量 a，b が角度 θ の座標回転に対して

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{8}$$

のように変換するとき，a，b は「ベクトル」であるという．式(8)は SO(2) の忠実な表現，「ベクトル表現」を定義している．この表現は既約ではない．なぜなら線形結合 $a+ib$，$a-ib$ をつくると

$$\begin{bmatrix} a'+ib' \\ a'-ib' \end{bmatrix} = \begin{bmatrix} e^{-i\theta} & \\ & e^{i\theta} \end{bmatrix} \begin{bmatrix} a+ib \\ a-ib \end{bmatrix} \tag{9}$$

となるからである†．従って $z=a+ib$，$z^*=a-ib$（$*$ は複素共役）はそれぞれウェイト1，-1 の相対不変量である．

$$z' = e^{-i\theta}z, \quad z^{*\prime} = e^{i\theta}z^* \tag{10}$$

特徴量 A，B，C，D が角度 θ の座標回転に対して

$$\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{11}$$

のように変換するとき，A，B，C，D は（2階の）「テンソル」であるという．式(11)は A，B，C，D から A'，B'，C'，D' への次のような線形写像を定義している．

$$\begin{bmatrix} A' \\ B' \\ C' \\ D' \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta & \cos\theta\sin\theta & \sin^2\theta \\ -\cos\theta\sin\theta & \cos^2\theta & -\sin^2\theta & \cos\theta\sin\theta \\ -\cos\theta\sin\theta & -\sin^2\theta & \cos^2\theta & \cos\theta\sin\theta \\ \sin^2\theta & -\cos\theta\sin\theta & -\cos\theta\sin\theta & \cos^2\theta \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} \tag{12}$$

これは SO(2) の（2階の）「テンソル表現」を定義している．この表現も既約ではない．まず A，B，C，D の作る行列は次のように「対称部分」と「反対称部分」とに一意的に分解できる．

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & (B+C)/2 \\ (B+C)/2 & D \end{bmatrix} + \begin{bmatrix} 0 & -(C-B)/2 \\ (C-B)/2 & 0 \end{bmatrix} \tag{13}$$

この分解が「不変」なこと，すなわち左辺を式(10)で変換したものの対称部分と反対称部分はそれぞれ右辺の対称部分と反対称部分を別々に変換したものと一致す

† 実数の範囲では既約である．しかし複素数まで考えなければ Schur の補題が成立しない．可換群の既約表現が1次元であるのは Schur の補題の直接の帰結である[(1),(8)]．

ることがすぐ確かめられる(何次元のテンソルでも同じ). 式(13)の反対称部分には独立な要素が一つしかないので $C-B$ は絶対不変量でなければならない.

対称部分は更に次のように「スカラ部分」(単位行列の定数倍)と「偏差部分」(トレース0の対称行列)とに一意的に分解できる.

$$\begin{bmatrix} A & (B+C)/2 \\ (B+C)/2 & D \end{bmatrix} = \frac{A+D}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} (A-D)/2 & (B+C)/2 \\ (B+C)/2 & -(A-D)/2 \end{bmatrix} \tag{14}$$

やはりこの分解が不変であること, すなわち左辺を変換したもののスカラ部分と偏差部分はそれぞれ右辺のスカラ部分と偏差部分を別々に変換したものと一致することがすぐ確かめられる(何次元のテンソルでも同じ). スカラ部分には独立な要素が一つしかないので $A+D$ は絶対不変量でなければならない. 偏差部分には独立な要素が二つあるが, それから $(A-D)+i(B+C)$, $(A-D)-i(B+C)$ を作るとそれぞれウェイト 2, -2 の相対不変量であることが確かめられる†.

以上より式(12)は次のように書き直せる††.

$$\begin{bmatrix} A'+B' \\ B'-C' \\ (A'-D')+i(B'+C') \\ (A'-D')-i(B'+C') \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & e^{-2i\theta} & \\ & & & e^{-2i\theta} \end{bmatrix} \begin{bmatrix} A+B \\ B-C \\ (A-D)+i(B+C) \\ (A-D)-i(B+C) \end{bmatrix} \tag{15}$$

すなわち

$$\left.\begin{aligned} &T=A+D,\ R=B-C \\ &S=(A-D)+i(B+C) \end{aligned}\right\} \tag{16}$$

とおけば, T, R は絶対不変量, S はウェイト2の相対不変量であり, 角度 θ の座標回転に対して

$$T'=T,\ R'=R,\ S'=e^{-2i\theta}S \tag{17}$$

と変換する. これらは Weyl のテーゼにより, ある特定の幾何学的意味をもっている(次章参照).

† もとの特徴量が実数であれば相対不変量はウェイトが正, 負のものが必ず対になって現れる.

†† 一般の n 次元回転群 SO(n) の r 階テンソル表現でも, このようにテンソルの(指標の)対称性と(指標間の)トレースとのみに基づいて組織的に既約表現に簡約できる(Weyl の定理)[(1),(14)].

5. オプティカルフローの解析

空間中に xy 座標系を固定し, xy 面を画像面とみなす. z 軸上の負の側に距離 f だけ隔った点 $(0,\ 0,\ -f)$ を視点として, 空間中の物体を xy 面に「中心投影」する(図1). 実際のカメラでも視点をレンズの中心, f を焦点距離とする中心投影とみなせる†. 空間中の点 $(X,\ Y,\ Z)$ はその点と視点とを結ぶ直線と xy 面との交点に投影されるから, その点の xy 座標を $(x,\ y)$ とすると

$$x=fX/(f+Z),\quad y=fY/(f+Z) \tag{18}$$

となる. $f\to\infty$ とした極限が「平行投影」である.

空間中で平面 $z=px+qy+r$ が剛体運動している場合を考えよう. p, q は面の「グラジエント」, r は z 軸に沿った面までの距離(「絶対距離」)を表す. 剛体運動はある基準点における速度 (a, b, c)(「並進速度」)とその回りの「回転速度」$(\omega_1, \omega_2, \omega_3)$(軸方向 $(\omega_1, \omega_2, \omega_3)$ の右ねじ回りの角速度 $\sqrt{\omega_1{}^2+\omega_2{}^2+\omega_3{}^2}$ (rad/sec) の回転)とによって指定される. 以下, 基準点として z 軸と面の交点 $(0,\ 0,\ r)$ をとる. 従って, この場合の物体パラメータは p, q, r, a, b, c, ω_1, ω_2, ω_3 の9個

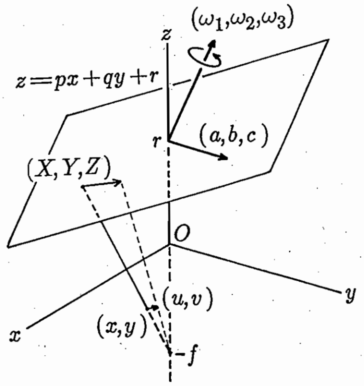

図1 平面 $z=px+qy+r$ の剛体運動は基準点 $(0,\ 0,\ r)$ における並進速度 $(a,\ b,\ c)$ とその回りの回転速度 $(\omega_1,\ \omega_2,\ \omega_3)$ とによって指定される. これに $(0,\ 0,\ -f)$ を視点とする中心投影を行うと画像面上にオプティカルフローが生じる.

Fig. 1 A plane having equation $z=px+qy+r$ is moving with translation velocity $(a,\ b,\ c)$ at $(0,\ 0,\ r)$ and rotation velocity $(\omega_1,\ \omega_2,\ \omega_3)$ around it. An optical flow is induced on the xy-plane by perspective projection, $(0,\ 0,\ -f)$ being the viewpoint.

† 厳密には「レンズの公式」$1/a+1/b=1/f$ (a: レンズ中心から物体までの距離, b: レンズ中心からフィルム面までの距離, f: 焦点距離)等を用いて補正しなければならないが, $a\gg f$ なら $b\approx f$ とみなせる.

である．

このような空間中の運動に対して，画像面上に観測される像の運動 $\dot{x}=u(x,\ y)$, $\dot{y}=v(x,\ y)$ は次のように表せる．

$$\left.\begin{aligned} u(x,\ y)&=u_0+Ax+By+(Ex+Fy)x \\ v(x,\ y)&=v_0+Cx+Dy+(Ex+Ey)y \end{aligned}\right\} \quad (19)$$

そして係数 u_0, v_0, A, B, C, D は次のように与えられる．

$$\left.\begin{aligned} u_0&=\frac{fa}{f+r},\quad v_0=\frac{fb}{f+r} \\ A&=p\omega_2-\frac{pa+c}{f+r} \\ B&=q\omega_2-\omega_3-\frac{qa}{f+r} \\ C&=-p\omega_1+\omega_3-\frac{pb}{f+r} \\ D&=-q\omega_1-\frac{qb+c}{f+r} \\ E&=\frac{1}{f}\left(\omega_2+\frac{pc}{f+r}\right) \\ F&=\frac{1}{f}\left(-\omega_1+\frac{qc}{f+r}\right) \end{aligned}\right\} \quad (20)$$

たとえば画像面上で観測した速度に式(19)の形の式を最小2乗法であてはめるなどして，係数 u_0, v_0, A, B, C, D, E, F を推定したとしよう†．すると今の場合，これが画像の特徴量となり，式(20)が物体パラメータ p, q, r, a, b, c, ω_1, ω_2, ω_3 を定めるべき復元方程式となる．この方程式の解は Longuet-Higgins[10] や Subbarao and Waxman[12] によっても与えられているが，以下の解析は Kanatani[6] に従う．

† 金谷[3] はオプティカルフローを求めることなしに，画像の「特徴」のみを測定して，これらを算出する方法を示している．

まず式(19)に座標回転をほどこすことによって，u_0, v_0 はベクトル，A, B, C, D はテンソル，E, F はベクトルであることがわかる[8]．従って次のように不変量が構成できる．

$$\left.\begin{aligned} &U_0=u_0+iv_0 \\ &T=A+D,\quad R=C-B \\ &S=(A-D)+i(B+C),\quad K=E+iF \end{aligned}\right\} \quad (21)$$

T, R は絶対不変量，U_0, K はウェイト1の相対不変量，S はウェイト2の相対不変量である．Weyl のテーゼによって，これらはそれぞれ特定の幾何学的意味をもっているはずである．対応するフローを図示してみればわかるように，U_0 は「並進」T は「発散」，R は「回転」，S は「せん断」，K は「扇状変形」を表している（図2）．

一方，物体パラメータに対しても同様に座標回転をほどこすことにより，p, q と a, b と ω_1, ω_2 はベクトル，r, c, ω_3 はスカラであることがわかる[8]．従って次のようなウェイト1の相対不変量が構成できる．

$$P=p+iq,\quad V=a+ib,\quad W=\omega_1+i\omega_2 \quad (22)$$

以上の不変量を用いると復元方程式(20)は次のように書き直せる．

$$\left.\begin{aligned} &U_0=\frac{f}{f+r}V \\ &PW'^{*}=(2\omega_3-R)-i(2c'+T) \\ &PW'=iS \\ &c'P-iW'=L \end{aligned}\right\} \quad (23)$$

ただし $c'=c/(f+r)$, $W'=W-iU_0/f$, $L=fK-U_0/f$ とおいた．このように不変量を用いると式の形が大変簡単になる．

まず $c\neq0$ であることを確かめたとしよう[6]．すると次の結果を得る（$c=0$ のときの扱いは Kanatani[6] 参照）．

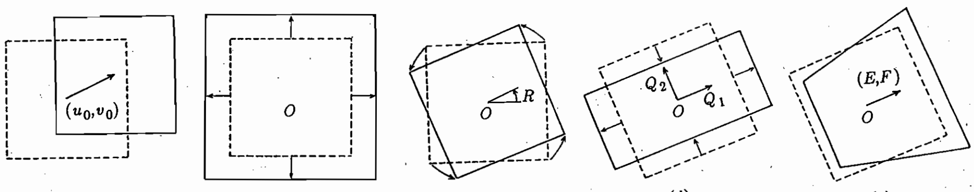

図2 (a)方向 $(u_0,\ v_0)$ の並進．(b)発散 T．(c)回転 R．(d) Q_1, Q_2 を最大引張り方向，最大圧縮方向とするせん断．(e)方向 $(E,\ F)$ の扇状変形．

Fig. 2 (a) Translation by $(u_0,\ v_0)$. (b)Divergence by T. (c)Rotation by R. (d)Shearing with Q_1 and Q_2 as the axes of maximum extension and maximum compression. (e)Fanning along $(E,\ F)$.

[定理1] 3次方程式

$$X^3+TX^2+\frac{1}{4}(T^2-|S|^2-|L|^2)X+\frac{1}{8}(\mathrm{Re}[L^2S]-T|L|^2)=0 \qquad (24)$$

は3実根をもつ．そのうちの中間の値をもつものを c' とすると復元方程式(23)の解は次のように与えられる．

$$\left.\begin{aligned}&V=\frac{f+r}{f}V_0,\quad c=(f+r)c'\\&P=\frac{1}{2c'}(L\pm\sqrt{L^2-4c'S})\\&W=\frac{i}{2}(L\mp\sqrt{L^2-4c'S})+\frac{i}{f}U_0\\&\omega_3=\frac{1}{2}R\pm\mathrm{Im}[L^*\sqrt{L^2-4c'S}]\end{aligned}\right\} \qquad (25)$$

ただしRe，Imはそれぞれ実部，虚部を表し，＊は複素共役を表す．以上より(i)絶対距離 r は不定，(ii) $a/(f+r)$，$b/(f+r)$，$c/(f+r)$ は一意的に定まり，(iii)グラジエント p，q と回転速度 ω_1，ω_2，ω_3 は2通りの解が存在する． □

定義より L はウェイト1の相対不変量である．不変量の積は各々のウェイトの和をウェイトとする不変量であり，複素共役をとるとウェイトは符号を変える．式(23)～(25)のような不変量で表した式では，右辺と左辺は同じウェイトであり，必ず同じウェイトの不変量同士しか足したり引いたりできないので式をチェックするのに大変便利である．またこれを利用して解の形をあらかじめ予測することができる場合も多い．これも不変量を導入する大きな利点の一つである．

このほかにもKanatani[6]は解の幾何学的解釈や面と面との接合の条件などの応用や数値例を示している．

式(20)で $f\to\infty$ の極限をとると，次のような平行投影の場合の復元方程式が得られる．

$$\left.\begin{aligned}&u_0=a,\quad v_0=b\\&A=p\omega_2,\quad B=q\omega_2-\omega_3\\&C=-p\omega_1+\omega_3,\quad D=-q\omega_1\end{aligned}\right\} \qquad (26)$$

最初の2式が a，b を与える．残りは不変量を用いて次のように表せる．

$$PW^*=2\omega_3-(R+iT),\quad PW=iS \qquad (27)$$

解は次のように与えられる[5]．

[定理2]

$$\left.\begin{aligned}&\omega_3=\frac{1}{2}(R\pm\sqrt{SS^*-T^2})\\&W=k\exp i\left(\frac{\Pi}{4}+\frac{1}{2}\arg(S)-\frac{1}{2}\arg(2\omega_3-(R+iT))\right)\\&P=\frac{S}{k}\exp i\left(\frac{\Pi}{4}-\frac{1}{2}\arg(S)+\frac{1}{2}\arg(2\omega_3-(R+iT))\right)\end{aligned}\right\} \qquad (28)$$

ここに k は不定の形状因子である．各 k に対して2組の解が存在する． □

これについてもKanatani[5]は面の接合条件や数値例を与えている．同時に，本質的に同一の問題を不変量を用いないで解いたSugihara and Sugie[13]との比較も論じている．彼らの方法では解は解析的な形では与えられず，また物理的に不可能な解を完全に除くことができない．

復元方程式(20)で $O(1/f^2)$ の項のみを除いて，$O(1/f)$ の項はそのままにすると E，F の式が

$$E=\omega_2/f,\quad F=-\omega_1/f \qquad (29)$$

となる．Kanatani[6]はこれを「疑似平行投影」と呼んでいる．このときの解は次のようになる[6]．

[定理3]

$$\left.\begin{aligned}&V=\frac{f+r}{f}V_0\\&P=S/L,\quad W=ifK\\&\omega_3=\frac{1}{2}(R+\mathrm{Im}[Se^{-2i\alpha}])\\&c=\frac{f+r}{2}(T-\mathrm{Re}[Se^{-2i\alpha}])\end{aligned}\right\} \qquad (30)$$

ただし，$\alpha\equiv\arg(L)$ である．これから(i)絶対距離 r は不定，(ii) $a/(f+r)$，$b/(f+r)$，$c/(f+r)$ は一意的に定まり，(iii)グラジエント p，q，回転速度 ω_1，ω_2，ω_3 は一意的に定まる．従って「にせの解」は存在しない． □

この解の幾何学的意味はKanatani[6]が論じている．

これまでは平面の運動を考えたが，一般の曲面の場合はオプティカルフローが式(19)のかわりに次のようになる(Subbarao[11])．

$$\left.\begin{aligned}&u(x,y)=u_0+Ax+By+Ex^2+2Fxy+Gy^2+\cdots\\&v(x,y)=v_0+Cx+Dy+Kx^2+2Lxy+My^2+\cdots\end{aligned}\right\} \qquad (31)$$

ただし…は x，y の高次の項を表す．この場合も不変量が次のように構成できる[8]．

$$\left.\begin{aligned}&U_0=u_0+iv_0,\quad T=A+D\\&R=C-B,\quad S=(A-D)+i(B+C)\\&H=(E+2L-G)+i(M+2F-K)\\&I=(E-2L+3G)+i(M-2F+3K)\end{aligned}\right\} \qquad (32)$$

$J=(E-2L-G)-i(M-2F-K)$

T, R は絶対不変量，U_0, H, I はウェイト1の相対不変量，S はウェイト2の相対不変量，J はウェイト3の相対不変量である．2次曲面の場合に復元方程式を不変量で表し，解を解析的な形で求めることができる(Subbarao[11])．

6. テクスチャによる曲面の復元

テクスチャをもつ曲面 $z=z(x, y)$ が z 軸に沿って xy 平面に平行投影される場合を考えよう(図3)．曲面上ではテクスチャが「一様」に分布しているとする[†]．そして投影された画像のテクチャの「密度」$\Gamma(x, y)$ が測定できるとしよう[††]．図3より，真のテクスチャ密度を ρ とすれば画像上の各点でのテクスチャ密度が次のように与えられる．

$$\Gamma(x, y)=\rho\sqrt{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2} \qquad (33)$$

曲面として2次曲面

$$z=r+px+qy+ax^2+2bxy+cy^2 \qquad (34)$$

を考えよう．するとこの場合の物体パラメータは p, q, r, a, b, c, ρ となる．式(34)を式(33)へ代入すると

$$\Gamma(x, y)=A_0\sqrt{1+A_1x+A_2y+A_3x^2+2A_4xy+A_5y^2} \qquad (35)$$

となる．各係数は次のように与えられる．

$$\left.\begin{aligned} &A_0=\rho\sqrt{1+p^2+q^2},\quad A_1=\frac{4(ap+bq)}{1+p^2+q^2}\\ &A_2=\frac{4(bp+cq)}{1+p^2+q^2},\quad A_3=\frac{4(a^2+b^2)}{1+p^2+q^2}\\ &A_4=\frac{4b(a+c)}{1+p^2+q^2},\quad A_5=\frac{4(b^2+c^2)}{1+p^2+q^2}\end{aligned}\right\} \qquad (36)$$

たとえば式(35)の形の式を，観測したテクスチャ密度にあてはめるなどして，係数 A_0～A_5 を推定したとしよう(Kanatani and Chou[9] はより巧妙な方法を用いている)．すると A_0～A_5 がこの場合の画像の特徴量であり，式(35)が物体パラメータ p, q, r, a, b, c, ρ を定めるべき復元方程式となる．

式(34)に座標回転をほどこすことによって，A_0 はスカラ，A_1, A_2 はベクトル，A_3, A_4, A_4, A_5 はテンソ

† 「一様性」を厳密に定義する必要があるが，ここでは省略する．詳細は Kanatani and Chou[9] 参照．

†† これも離散的テクスチャの「密度」とは何かを厳密に定義し，それを測定する手続きを与えなければならないが，ここでは省略する．Kanatani and Chou[9] ではこれを超関数として定義し，それに基づいた測定手続きを与えている．

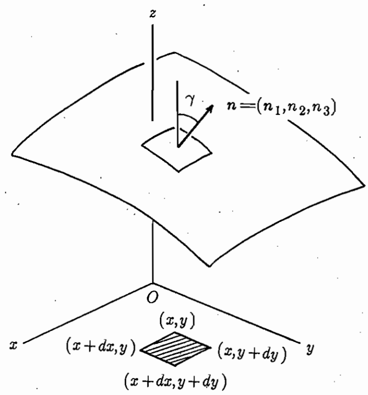

図3 z 軸に沿う平行投影では画像面上の4点 (x, y), $(x+dx, y)$, $(x+dx, y+dy)$, $(x, y+dy)$ の作る微小正方形に対応する曲面上の部分の面積は $dxdy/\cos\gamma$ である．ただし γ は曲面の法線方向 $\boldsymbol{n}$ と z 軸とのなす角である．

Fig. 3 Under orthographic projection along the z-axis, the area of the region on the surface which corresponds to the infinitesimal square on the xy-plane defined by four points (x, y), $(x+dx, y)$, $(x+dx, y+dy)$, $(x, y+dy)$ is $dxdy/\cos\gamma$, where γ is the angle made by the unit normal $\boldsymbol{n}$ to the surface and the z-axis.

ルであることがわかる[8]．従って次のような不変量が構成できる．

$$\left.\begin{aligned} &V=\frac{A_1+iA_2}{4},\quad T=\frac{A_3+A_5}{8}\\ &S=\frac{A_3-A_5}{8}+i\frac{A_4}{4}\end{aligned}\right\} \qquad (37)$$

A_0 自身と T は絶対不変量，V はウェイト1の相対不変量，S はウェイト2の相対不変量である．

また同様に式(34)に座標回転をほどこすことによって，r はスカラ，p, q はベクトル，a, b, b, c はテンソルであることがわかる．そこで次のような不変量が構成できる．

$$\left.\begin{aligned} &k=\sqrt{1+p^2+q^2},\quad v=\frac{p+iq}{k}\\ &t=\frac{a+c}{2k},\quad s=\frac{a-c}{2k}+i\frac{b}{k}\end{aligned}\right\} \qquad (38)$$

ρ 自身と k, t は絶対不変量，v はウェイト1の相対不変量，s はウェイト2の相対不変量である．これらを用いると復元方程式(36)は次のように書き直せる．

$$\left.\begin{aligned} &\rho k=A_0,\quad tv+sv^*=V\\ &t^2+ss^*=T,\quad ts=S/2\end{aligned}\right\} \qquad (39)$$

$t\neq0$, $t^2-ss^*\neq0$ の場合，解は次のようになる[8]．

[定理 4]

$$
\left.
\begin{array}{l}
t=\pm\sqrt{\dfrac{T\pm\sqrt{T^2-SS^*}}{2}},\quad s=\dfrac{S}{2t} \\
v=\dfrac{tV-sV^*}{t^3-ss^*},\quad k=\dfrac{1}{\sqrt{1-vv^*}} \\
\rho=A_0/k
\end{array}
\right\} \qquad (40)
$$

もとの物体パラメータで表せば次のようになる.

$$
\left.
\begin{array}{l}
\rho=A_0/k,\quad p=k\mathrm{Re}[v] \\
q=k\mathrm{Im}[v],\quad a=k(t+\mathrm{Re}[s]) \\
b=k\mathrm{Im}[s],\quad c=k(t-\mathrm{Re}[s])
\end{array}
\right\} \qquad (41)
$$

□

式(40)の第1式からもわかるように，解は4組存在する．平行投影なので xy 面に平行な「鏡」に関する鏡像の関係にあるものは区別できない．従って，鏡像の関係を除けば2組の解が存在する．これはテクスチャ密度の変化の割合が面の傾き角(「スラント」)のみによって定まり，傾きの方向(「ティルト」)にはよらないことに帰因している．

$t=0$ となるのは二つの主曲率が等しい絶対値をもつときであり，このとき面が楕円的(ガウス曲率が正)か双曲的(ガウス曲率が負)かを区別できないことがわかる[8]．また $t^2-ss^*=0$ となるのは面が放物的(ガウス曲率が0)のときであり，漸近方向(「峰」の方向)は定まるが，その方向のこう配が不定となることもわかる[8]．

7. むすび

本論文ではコンピュータビジョンで重要な3次元復元問題に対して，画像面上の座標系回転に対応する2次元回転群 SO(2)の既約表現を与える不変量を用いることを提案した．これによって復元方程式の形が簡単になり，解析解が導きやすくなるだけでなく，方程式や変数の幾何学的意味が明確になり(Weylのテーゼ)，式変形の見通しが大変やさしくなる．このことをオプティカルフローの解析とテクスチャによる曲面の復元を例にとって説明した．

3次元復元問題に対してWeylのテーゼを用いようとするとき，「現象の本質を変えないような変換の群」としては何も画像面上の座標変換だけとは限らない．たとえばカメラをレンズの中心の回りに回転したとき，同一物体については観測する「射線」(ray)が同一であるから，本質的には同一の情報を観測している．従って，カメラの回転に対応する3次元回転群 SO(3)の既約表現を与える不変量を構成することができる．SO(3)は可換群ではないので既約表現は一般には1次元にならないが，観測した特徴量の線形結合をとるだけでなく，非線形の代数式によって1次元の不変量にまで帰着させることができる．(詳細は文献(2)，(7)，(8)参照．)

このような考察はさまざまな問題に応用することができ，コンピュータビジョンの研究に欠くことのできないものと思われる．本論文がその一つの基礎となることを期待している．

謝辞 本研究は著者が米国Maryland大学に招聘(へい)されて行った研究の一部であり，討論頂いたAzriel Rosenfeld, Larry S. Davis, John Aloimonos, Murali Subbarao, Tsai-Chia Chouの諸氏に感謝する．また東京大学工学部計数工学科の甘利俊一教授，杉原厚吉助教授からも有益なコメントを頂いた．

文　献

(1) M. Hammermesh: "Group Theory and its Application to Physical Problems", Addison-Wesley, Reading, MA (1962).

(2) 金谷健一："オプティカルフローの変換則と不変分解"，信学技報，**PRL85**-3 (1985-05).

(3) 金谷健一："対応点を用いない物体の運動認識の理論"，情処学論，**27**，3，pp. 373-382 (1986-03).

(4) 金谷健一："3Dユークリッドか2D非ユークリッドか―画像理解の方法論―"，信学論(D)，**J70-D**，5，pp. 1035-1038 (昭62-05).

(5) K. Kanatani: "Structure and motion from optical flow under orthographic projection", Comput. Vision Graphics Image Process., **35**, pp. 181-199 (1986).

(6) K. Kanatani: "Structure and motion from optical flow under perspective projection", Comput. Vision Graphics Image Process, **37** (1987).

(7) K. Kanatani: "Camera rotation invariance of image characteristics", Comput. Vision Graphics Image Process, **38** (1987).

(8) K. Kanatani: "Group Theoretical Methods on Image Understanding", Springer, Berlin (1987).

(9) K. Kanatani and T.-C. Chou: "Shape from texture: General principle", 第17回画像工学コンファレンス論文集, pp. 129-132 (昭61-12).

(10) H.C. Longuet-Higgins: "The visual ambiguity of a moving plane", Proc. Roy. Soc. Lond., **B-223**, pp. 165-175 (1984).

(11) M.Subbarao: "Interpretation of Image Motion Fields: A Computational Study", Technical Report CAR-TR-221, Center for Automation Research, University of Maryland (Sept. 1986).

(12) M. Subbarao and A.M. Waxman: "Closed form solution to image flow equauions for planar surfaces in motion", Comput. Vision Graphics Image Process., **36**, pp. 208-228 (1986).

(13) K. Sugihara and N. Sugie: "Recovery of rigid structure from orthographically projected optical flow", Comput. Vision Graphics Image Process., **27**, pp. 309-320 (1984).

(14) H. Weyl : "The Classical Groups, Their Invariants and Representations", 2nd ed., Princeton Univ. Press, Princeton, NJ (1946).
(15) H. Weyl : "The Theory of Groups and Quantum Mechanics", Dover, New York (1950).

(昭和61年11月17日受付)

金谷　健一

昭47東大・工・計数(数理工学)卒. 昭44～45米国Case Western Reserve大学留学. 昭54東大大学院博士課程(計数工学)了. 工学博士. 同年群大・工・情報工学助手. 昭58同助教授. 昭60～61米国Maryland大学招聘研究員. 研究分野は応用力学, 粉体工学, 土質力学を経て, 現在はコンピュータビジョン, 画像理解, 画像認識.