

## 球関数による姿勢表現と姿勢推定

玉木 徹<sup>†</sup> 天野 敏之<sup>††</sup> 金田 和文<sup>†</sup><sup>†</sup> 広島大学大学院工学研究科情報工学専攻 〒739-8527 広島県東広島市鏡山 1-4-1<sup>††</sup> 奈良先端科学技術大学院大学情報科学研究科 〒630-0101 奈良県生駒市高山町 8916 番地の 5E-mail: <sup>†</sup>{tamaki,kin}@hiroshima-u.ac.jp, <sup>††</sup>amano@is.naist.jp

あらまし 本論文では、球関数を用いた新しい姿勢表現を提案する。これは見えに基づく姿勢推定による 3 自由度回転の姿勢を表現するためのものである。まず、従来行われている 1 自由度の姿勢表現方法がフーリエ基底で表せることを示し、それを球面調和関数により 2 自由度に拡張する。そして、3 自由度の姿勢を、 $SO(3)$  上の連続関数を球関数を用いて表す。この姿勢の球関数表現と単位四元数・回転行列による表現との変換も与える。この球関数表現を用いた姿勢推定の結果を示し、提案方法の有効性を議論する。

キーワード 姿勢推定, 見えに基づく推定, 球関数, 球面調和関数, 3 自由度,  $SO(3)$ , 姿勢表現, 四元数, 回転行列

## Representing pose by spherical functions for pose estimation

Toru TAMAKI<sup>†</sup>, Toshiyuki AMANO<sup>††</sup>, and Kazufumi KANEDA<sup>†</sup><sup>†</sup> Department of Information Engineering, Graduate School of Engineering, Hiroshima University

1-4-1 Kagamiyama, Higashi-hiroshima, Hiroshima, 739-8527 Japan

<sup>††</sup> Graduate School of Information Science, Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: <sup>†</sup>{tamaki,kin}@hiroshima-u.ac.jp, <sup>††</sup>amano@is.naist.jp

**Abstract** We propose a novel representation of pose in 3 degrees-of-freedom (DOF) for view-based pose estimation. First we show that a conventional representation of pose in 1 DOF is a Fourier basis, and extend the observation to 2 DOF with spherical harmonics. Then we represent 3 DOF pose with spherical functions that are continuous on  $SO(3)$ , and give transformations from the spherical functions representation to a quaternion and a rotation matrix.

**Key words** pose estimation, view-based, appearance-based, spherical functions, spherical harmonics, 3 degrees-of-freedom,  $SO(3)$ , pose representation, quaternions, rotation matrix

## 1. はじめに

姿勢とは何であろうか。3 次元幾何を扱うコンピュータビジョンでは、基準となる世界座標系からカメラや物体の局所座標系への座標系変換を回転行列  $R$  で表し、これをカメラや物体の姿勢と呼ぶ。人体やロボットなどの関節を持つ物体 (多関節物体) では、その関節角を総称して姿勢と呼んでいる。このように、画像に写る物体の状態を表すパラメータを指して姿勢 (pose) と呼ぶことが多い。この姿勢パラメータを求める姿勢推定は、様々な応用を持つ重要な問題である。

本論文が扱う主題は、見えに基づく (view-based, appearance-based) 姿勢推定手法である。この手法は、与えられた画像 (見え, view, appearance) と姿勢パラメータの関係を学習する、一種の回帰問題である。まず、画像に写る物体の姿勢パラメータ

$p_j$  と、その姿勢のときに撮影された画像  $x_j$  を対にしたものを学習セット  $\{p_j, x_j\}_{j=1,2,\dots}$  として、その間の関係  $p_j = f(x_j)$  を学習しておく。そして学習された関係  $f$  を用いて、新たに入力された未知の画像  $x$  に対応する姿勢  $p$  を推定する (つまり  $p = f(x)$ )。

本論文では、3 自由度の回転で表現される単一物体の姿勢を表す、新しい表現方法を提案する。推定手法 (つまり  $f$ ) に関する研究は多数存在するが、これまで姿勢  $p$  に関してはまったく注意が払われていなかった。しかし後述するように、見えに基づく姿勢推定には、他の応用にはない特有の問題が存在するために、3 自由度の回転に対しては適切な姿勢パラメータを用いなければならない。このことを考慮していない従来研究はすべて、関係  $f$  が線形・非線形のどちらであっても、姿勢の表現方法に本質的な問題を持っている。本論文ではこのことを指摘

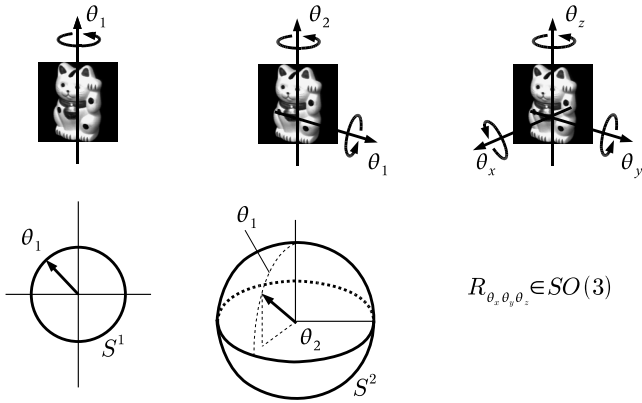


図1 1,2,3 自由度の回転と対応する形状

し、続いて必要とされる姿勢パラメータについて議論し、提案する球関数 (spherical functions) を用いた姿勢表現手法が適していることを示す。

### 1.1 物体の回転姿勢の自由度と対応する幾何形状

ここではまず、本論文で扱う回転と自由度についてまとめておく。物体の回転は、その回転に係わる回転軸の数によって、図1に示すように3種類の自由度に分けられる。

1自由度の場合、物体はある軸に対して角度  $\theta_1 \in [0, 2\pi)$  だけ回転する。基準姿勢 ( $\theta_1 = 0$ ) を決めておくと、物体の姿勢は、1自由度の姿勢は単位円  $S^1$  (注1) 上の点の偏角  $\theta_1$  と一対一の関係にある。

2自由度の場合、物体の姿勢は2つの軸に関する回転角  $\theta_1, \theta_2$  に依存する。これは緯度・経度や方位角・仰角を考えればわかりやすいだろう。基準姿勢 ( $\theta_1 = \theta_2 = 0$ ) を決めておくと、2自由度の姿勢は単位球面  $S^2$  (注2) 上の点と一対一の関係にある (注3)。

3自由度の場合、3つの直交する軸 (例えば  $x, y, z$  軸) に対してそれぞれ角度  $\theta_x, \theta_y, \theta_z$  だけ回転する。この軸周りの回転角度はオイラー角や固定角と呼ばれる [1] が、姿勢を一意に決めることはできない [2]。3自由度の姿勢は  $3 \times 3$  回転行列  $R$  によって一意に与えられる (単位行列  $I$  を基準姿勢とする)。この回転行列の集合を  $SO(3)$  (注4) といい、3自由度の姿勢は  $SO(3)$  と一対一の関係にある。

### 1.2 姿勢の表現方法 $p$ 姿勢の推定方法 $f$

従来の見えに基づく姿勢推定の研究は1自由度の回転を扱うものが多い。それらのほとんどは、画像  $x$  と姿勢  $p$  の関係  $p = f(x)$  を表す関数  $f$  に (つまり推定手法に) 関連するもので、線形関数 (行列) を用いるもの [3] ~ [7] や、カーネルを利用した非線形関数を用いるもの [8] ~ [13]、関数として与えずに手続きを定めるもの [14], [15] などがある。いくつかの研究では、その手法をそのまま3自由度の姿勢推定に適用しているだ

けであり、姿勢の表現方法について特段の考察を行っているものはない。

しかし、線形・非線形のどちらであっても、従来手法で用いられている3自由度の姿勢  $p$  の表現方法には問題がある。本研究は、線形手法を用いて見えに基づく姿勢推定が用いられるべき姿勢  $p$  についての議論に焦点を当てる。「どんな手法で推定するのか」と「どんな姿勢表現を用いるのか」を分けて議論することで、さまざまな推定手法に対して本研究を拡張することができるからである。

### 1.3 本論文の構成

本論文では、見えに基づく姿勢推定に必要な姿勢とはなにかについて議論し、そのための新しい姿勢表現方法を提案する。まず2節において、従来の1自由度の姿勢表現と、姿勢表現が持つべき性質について議論する。次に3節で、球関数を用いた3自由度の姿勢の表現方法を提案する。4節でその姿勢表現を用いた姿勢推定実験結果を示す。

## 2. 1自由度の姿勢表現

### 2.1 従来の1自由度の姿勢表現

まず1自由度の見えに基づく姿勢推定の場合に一般的に行われているやり方 (例えば [3], [4], [9], [11] ~ [13] 等) を考える。学習時において、1軸回転する物体の画像  $x_j \in \mathbb{R}^N$  とそのときの回転角  $\theta_{1j} \in [0, 2\pi)$  が与えられる (図1 (左) 参照)。姿勢表現としてはベクトル  $p_j = (\cos \theta_{1j}, \sin \theta_{1j})^T \in \mathbb{R}^2$  をとり、

$$(p_1, p_2, \dots, p_n) = F(x_1, x_2, \dots, x_n) \quad (1)$$

を満たす行列  $F \in \mathbb{R}^{2 \times N}$  を求める。

推定時には、与えられた画像  $x$  の姿勢  $Fx = \hat{p} = (\hat{p}_1, \hat{p}_2)^T$  を得て、 $\hat{\theta}_1 = \tan^{-1} \frac{\hat{p}_2}{\hat{p}_1}$  とする。推定された姿勢  $\hat{p}$  自体は単位円上にはないかもしれないが、 $\tan^{-1}$  による角度計算には問題になることはない (図2参照)。これを  $\cos, \sin$  表現と呼ぶことにする。

これから、姿勢表現についての以下の性質が見えてくる。

### 2.2 姿勢表現の性質：連続性

見えに基づく姿勢推定の場合、1自由度の姿勢に、角度  $\theta_1$  そのものを用いることはほとんどない。その理由は、画像は連続的に変化するのに、角度は  $2\pi$  において不連続となるからである。

これを理解するには、上記の見えに基づく方法がどのように姿勢を推定しているのかを考える必要がある。固有顔に代表されるように、ある特定の種類に属する画像は、与えられた学習画像ベクトルの固有ベクトルの線形和によってよく近似することができる、と仮定する。固有ベクトルが張る固有空間は学習画像が張る部分空間に含まれるので、固有ベクトルは学習画像の線形和で表現できる。したがって、ある画像  $x$  は学習画像  $x_j$  の線形和

$$x \cong b_1 x_1 + b_2 x_2 + \dots + b_n x_n \quad (2)$$

で近似することができる、とすることができる。一方、学習画像  $x_j$  の姿勢  $p_j$  は学習された線形写像  $F$  によって

(注1):  $S^1 = \{p \in \mathbb{R}^2 \mid \|p\| = 1\}$  を1次元球面という。

(注2):  $S^2 = \{p \in \mathbb{R}^3 \mid \|p\| = 1\}$  を2次元球面という。

(注3): ただし2軸で回転しても、姿勢は3自由度であるので、図1 (上段中央) は2自由度の説明として正しくない。この図は分かりやすく対比する説明のために用いている。

(注4):  $O(3) = \{R \in \mathbb{R}^{3 \times 3} \mid RR^T = R^T R = I\}$  を3次元直交群、 $SO(3) = \{R \in O(3) \mid \det(R) = 1\}$  を3次元特殊直交群という。

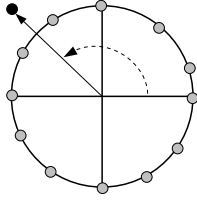


図 2 1 自由度の角度推定。灰色の点は単位円  $S^1$  上の学習角度を、黒丸は推定値  $\hat{p}$  を表す。 $\hat{p}$  が  $S^1$  上になくとも角度は計算できる。

$$p_j \cong Fx_j \quad (3)$$

となるべきである。すると、ある画像  $x$  に対応する姿勢  $p$  は、以下のように表現されなければならない。

$$p = Fx = b_1 Fx_1 + b_2 Fx_2 + \cdots + b_n Fx_n \quad (4)$$

$$\cong b_1 p_1 + b_2 p_2 + \cdots + b_n p_n \quad (5)$$

つまり、推定される姿勢  $p$  は学習された姿勢  $p_j$  の線形和で表現されることになる。

姿勢表現が連続でない場合、推定された姿勢は信頼できるものではない。もし角度  $\theta_1$  そのものを用いた場合に、例えば  $0, 5, \dots, 300, 350[\text{deg}]$  の姿勢を学習したとする。このとき  $355[\text{deg}]$  の画像の姿勢の推定値は、おそらく  $350[\text{deg}]$  と  $0[\text{deg}]$  の重み付き和である  $175[\text{deg}]$  であり、これはまったく真値とかけ離れている。Melzer ら [11] が示しているように、この問題は kernel を用いた非線形推定であっても同様である。

姿勢が連続であり、その見え方も連続的に変わるのにもかかわらず、姿勢表現が不連続点（特異点）を持つ場合、この問題はその特異点付近において必ず発生する。したがって、 $p$  の各要素は（姿勢に関して）少なくとも  $C^1$  連続でなければならない。

### 2.3 姿勢表現の性質：一対一

次の性質は、ある画像  $x$  に対しては、一つの姿勢  $p$  が一対一に対応しなければならない、というものである。ただし、同じ模様が現れる物体の場合は、見え方が同じになってしまいどのような姿勢表現を考えても一対一にならない。これはそもそも姿勢推定が不可能であるため、除外する。

一対一でない表現は、学習にも推定にも問題が生じる。例えば、ある姿勢の見え方が  $x$  である場合に、姿勢の表現の仕方が  $p_1$  と  $p_2$  の 2 通りあるとする。しかし、 $Fx = p_1$  かつ  $Fx = p_2$  となるような線形写像  $F$  は一般的に存在しない。そのため、 $x$  に対して  $p_1, p_2$  の両方を学習することは不可能である。また推定姿勢は、前述のとおり学習姿勢の線形和で表されるため、2 通りあるうちのどちらか一方を出力する、という操作が入る余地はない。

$\cos, \sin$  表現は、単位円上の位置と姿勢が一対一に対応している。しかし、角度  $\theta_1$  そのものを用いた場合には、 $\theta_1$  と  $\theta_1 + 2n\pi$  ( $n \in \mathbb{Z}$ ) が同じ姿勢を表すため、一対一ではない。

### 2.4 1 自由度の姿勢表現の基底関数による解釈

以上のことをまとめると、「1 自由度の姿勢は、単位円  $S^1$  上で定義される周期的ないくつかの連続関数  $C^1(S^1)$ （上記の場合は  $\cos \theta_1$  と  $\sin \theta_1$ ）によって一意に表現される」といえる。

$C^\infty$  連続な周期関数  $C^\infty(S^1)$  は、複素フーリエ基底

$$Y_{\ell_1} = e^{-\ell_1 \theta_1 i}, \ell_1 \in \mathbb{Z}, i = \sqrt{-1} \quad (6)$$

を用いてフーリエ級数展開できることはよく知られており、したがって複素フーリエ基底は姿勢表現の性質を満たしている、と言える。

従来の 1 自由度の姿勢表現  $(\cos \theta_{1j}, \sin \theta_{1j})^T$  は、このうち最も低周波数の（したがってフーリエ級数展開の意味でもっとも単純な） $Y_{\pm 1} = e^{\pm \theta_1 i}$  の実部と虚部を用いている、とみなすことができる。

## 3. 2 自由度の姿勢表現

### 3.1 球面調和関数による 2 自由度の姿勢表現

それでは同様に考えて、2 自由度の姿勢表現（図 1（中央））のために、2 次元球面  $S^2$  上の周期的連続関数  $C^\infty(S^2)$  を展開する低周波数の基底を用いることを考えてみる。

球面上の連続関数を展開する関数系は、球面調和関数 (spherical harmonics) としてよく知られている [16] ~ [18] (注5)。以下表記の簡単のため引数  $(\theta_1, \theta_2)$  を省略すると、 $\theta_1 \in [0, \pi)$ ,  $\theta_2 \in [0, 2\pi)$  として、

$$Y_{\ell_1, \ell_2} = \sqrt{b_{\ell_1, \ell_2}} P_{\ell_1}^{\ell_2}(\cos \theta_1) e^{-\ell_2 \theta_2 i}, \ell_1, \ell_2 \in \mathbb{Z} \quad (7)$$

ここで  $\ell_1 \geq |\ell_2|$  であり、 $P_{\ell_1}^{\ell_2}$  は Legendre 陪関数、 $b_{\ell_1, \ell_2}$  は係数である（付録参照）。

ここで、最も低周波 ( $\ell_1 = 1$ ) に対応する関数(注6)

$$Y_{1,0} = \sqrt{3} \cos \theta_1 \quad (8)$$

$$Y_{1, \pm 1} = -\sqrt{\frac{3}{2}} \sin \theta_1 e^{\mp \theta_2 i} \quad (9)$$

を姿勢表現に用いると、これらから逆に  $\theta_2$  と  $\theta_1$  を求めることができる(注7)。

$$\theta_2 = \tan^{-1} \left( \frac{-(Y_{1,-1} - Y_{1,1})/i}{-(Y_{1,-1} + Y_{1,1})} \right) \quad (10)$$

$$\theta_1 = \tan^{-1} \left( \frac{(Y_{1,-1} - Y_{1,1})/(-\sqrt{2} i \sin \theta_2)}{Y_{1,0}} \right) \quad (11)$$

### 3.2 連続関数系を用いる利点

上記の姿勢表現  $Y = (Y_{1,0}, Y_{1,1}, Y_{1,-1})^T \in C^\infty(S^2)^3$  に用いている球面調和関数は連続関数系である。このため、姿勢表現としての連続性が保証されている。

一方、従来の 2 自由度の姿勢推定では、姿勢表現に角度  $p = (\theta_1, \theta_2)^T$  が用いられていた。しかし、仰角  $\theta_1$  は極において連続ではないため、姿勢表現には適さない。仰角が連続でない部分は実験において無視されており、この問題が表に現れなかった。

(注5): spherical harmonics を spherical functions と呼ぶ場合もある（例えば [19]）ため、混同しやすい。

(注6): 一般的に用いられているものと係数が異なるのは、単に内積の定義に用いる測度の違いであり、重要ではない。

(注7): この計算は実際には  $\text{atan}(y, x)$  などの関数を用いる。このとき、実数になることを保証したり符号の整合をとるために、 $\tan^{-1}$  内の分子を  $y$ 、分母を  $x$  としてそれぞれ計算する必要がある。

このように、連続関数系を用いることで姿勢表現の連続性が保証される。次節では、この考え方を3自由度の姿勢に拡張する。

#### 4. 球関数を用いた3自由度の姿勢表現

この節において、3自由度の姿勢を「 $SO(3)$ 上で定義される連続関数  $C^\infty(SO(3))$ によって一意に表現する」ことを提案する。これは姿勢表現に必須である連続性、一対一を満たしている。

##### 4.1 従来の3自由度の姿勢表現の問題点

従来の3自由度の姿勢表現は、これらの条件を満たしていない。

オイラー角や固定角など[1]の角度表現は、特異点を持つことが知られている(つまり連続ではない)。また異なる角度の組み合わせでも同じ姿勢を表すことがあり、一対一ではない。

angle-axis 表現[20]は、3次元ベクトルの方向と長さで回転軸と回転角度を表すが、方向は連続表現であっても回転角度は連続ではない。

単位四元数は連続であるが、ある単位四元数  $q$  と  $-q$  は同じ姿勢を表すため、一対一ではない(指数写像[21]も同じ欠点を持つ)。ロボットの連続的な姿勢制御[22]やCG[2],[23],[24]における姿勢の補間[25],[26]、3次元データの位置合わせ[27]などでは、一対一ではなくても連続性があればよい。単位四元数は広く用いられている。しかし、前述のように姿勢推定には一対一という条件は必須であり、単位四元数は適切ではない[6]。これは、姿勢推定という問題に特有の性質である。

$3 \times 3$  回転行列は唯一条件を満たす。ただし行列同士の和は回転としての意味を持たないため、制御や補間には使われない。また姿勢推定においては、学習時の姿勢が直交行列で与えられたとしても、姿勢推定には行列の直交性を保つ仕組みがないため、推定された姿勢は直交行列になるとは限らない。そのため、後処理として最も近い直交行列を再推定しなければならない(付録参照)。

##### 4.2 球関数による姿勢表現

本論文で提案する姿勢表現は、前節で示した単位円  $S^1$  上や球面  $S^2$  上の連続関数系の延長線上にある。一般の  $n$  次元球面  $S^n$  上の  $C^\infty$  級連続な直交関数系  $Y_{\ell_1, \dots, \ell_n}$  は、球関数[28]と呼ばれる(一般的な定義は付録を参照)。  $n=1$  の球関数は複素フーリエ基底  $Y_{\ell_1}$  に、  $n=2$  の球関数は球面調和関数  $Y_{\ell_1, \ell_2}$  に一致する。

この球関数を用いて、3自由度の姿勢を表現することを考えるのが、本題である。それでは  $S^3$  上の球関数を見てみよう。

$$Y_{\ell_1, \ell_2, \ell_3} = \sqrt{b_{\ell_1, \ell_2, \ell_3}} C_{\ell_1}^{1, \ell_2}(\cos \theta_1) P_{\ell_2}^{\ell_3}(\cos \theta_2) e^{-\ell_3 \theta_3 i} \quad (12)$$

ここで  $C_{\ell_1}^{1, \ell_2}$  は Gegenbauer 陪関数である(付録参照)。

しかし3次元球面  $S^3$  は、 $SO(3)$  とは一対一の関係ではない(二重被覆[20, p.41][29, p.324]である)。そのため、 $S^3$  上の球関数を  $SO(3)$  の姿勢表現に用いることは適切ではないと思わ

れる。

前述した考え方に基くならば、3自由度の姿勢表現のためには、 $SO(3)$  上の連続関数系を用いる必要がある。そしてその関数系は、 $S^3$  上の球関数を  $Y_{\ell_1, \ell_2, \ell_3}$  とすると、 $Y_{2\ell_1, \ell_2, \ell_3}$  で与えられる[28]。この最も低周波 ( $2\ell_1 = 2$ ) に対応する関数は以下の9つである。

$$Y_{2,0,0} = 1 + 2 \cos 2\theta_1 \quad (13)$$

$$Y_{2,1,0} = \sqrt{24} \sin \theta_1 \cos \theta_1 \cos \theta_2 \quad (14)$$

$$Y_{2,1,\pm 1} = -2\sqrt{3} \sin \theta_1 \cos \theta_1 \sin \theta_2 e^{\mp \theta_3 i} \quad (15)$$

$$Y_{2,2,0} = \sqrt{2} \sin^2 \theta_1 (3 \cos^2 \theta_2 - 1) \quad (16)$$

$$Y_{2,2,\pm 1} = -2\sqrt{3} \sin^2 \theta_1 \sin \theta_2 \cos \theta_2 e^{\mp \theta_3 i} \quad (17)$$

$$Y_{2,2,\pm 2} = \sqrt{3} \sin^2 \theta_1 \sin^2 \theta_2 e^{\mp 2\theta_3 i} \quad (18)$$

これらを、使いやすいように実数化して係数を落とし、以下のように番号付けする。

$$Y_0 = 4 \cos^2 \theta_1 - 1 \quad (19)$$

$$Y_1 = \cos \theta_1 \sin \theta_1 \cos \theta_2 \quad (20)$$

$$Y_2 = \cos \theta_1 \sin \theta_1 \sin \theta_2 \cos \theta_3 \quad (21)$$

$$Y_3 = \cos \theta_1 \sin \theta_1 \sin \theta_2 \sin \theta_3 \quad (22)$$

$$Y_4 = \sin^2 \theta_1 (3 \cos^2 \theta_2 - 1) \quad (23)$$

$$Y_5 = \sin^2 \theta_1 \cos \theta_2 \sin \theta_2 \cos \theta_3 \quad (24)$$

$$Y_6 = \sin^2 \theta_1 \cos \theta_2 \sin \theta_2 \sin \theta_3 \quad (25)$$

$$Y_7 = \sin^2 \theta_1 \sin^2 \theta_2 (\cos^2 \theta_3 - \sin^2 \theta_3) \quad (26)$$

$$Y_8 = \sin^2 \theta_1 \sin^2 \theta_2 \cos \theta_3 \sin \theta_3 \quad (27)$$

この9変数からなる  $\mathbf{Y} = (Y_0, Y_1, \dots, Y_8)^T \in C^\infty(SO(3))^9$  を、姿勢の球関数表現と呼ぶことにする。

2自由度までと同様に、 $\theta_1, \theta_2, \theta_3$  を計算してもよいが、これらの角度は、直感的な(オイラー角などの)回転角とは異なるものである。そのため、まず  $\mathbf{Y}$  と単位四元数  $q$  との対応を与え、次に回転行列  $R$  へと変換する。

##### 4.3 四元数・回転行列と球関数表現との関係

$S^3$  上の点の座標  $\mathbf{q} = (q_1, q_2, q_3, q_4)^T$  は

$$\mathbf{q} = (\cos \theta_1, \sin \theta_1 \cos \theta_2, \sin \theta_1 \sin \theta_2 \cos \theta_3, \sin \theta_1 \sin \theta_2 \sin \theta_3)^T \quad (28)$$

と表される。点  $\mathbf{q}$  を用いると、球関数表現は

$$\mathbf{Y} = (4q_1^2 - 1, q_1 q_2, q_1 q_3, q_1 q_4, 3q_2^2 - 1 + q_1^2, q_2 q_3, q_2 q_4, q_3^2 - q_4^2, q_3 q_4)^T \quad (29)$$

と表すことができる。また符号を除いた逆変換は、

$$(q_1^2, q_2^2, q_3^2, q_4^2) = \left( \frac{Y_0+1}{4}, \frac{Y_0-4Y_4-3}{-12}, \frac{Y_0+2Y_4-6Y_7-3}{-12}, \frac{Y_0+2Y_4+6Y_7-3}{-12} \right) \quad (30)$$

である。このとき、 $q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1$  が保証されており、かつ  $\mathbf{Y}$  が  $\mathbf{q}$  と  $-\mathbf{q}$  に対応している。

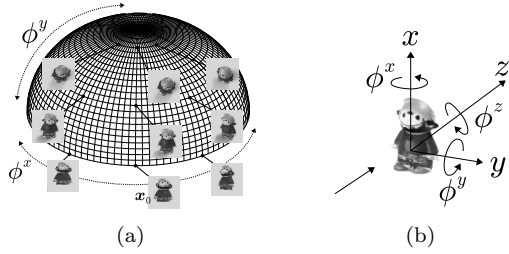


図3 実験に用いた Dwarf 画像 [30]。(a) 撮影カメラ位置。(b) Dwarf 画像の姿勢と回転軸の対応関係

ここで  $S^3$  上の点に四元数演算を定義して、 $q$  を単位四元数  $q$  とみなすことにする。すると、この単位四元数が表す回転に対して、 $\theta_2, \theta_3$  は回転軸の方向を表す角度に、 $\theta_1$  はその軸周りの回転角度に対応する。

単位四元数  $q$  に対応する回転行列  $R_q$  はよく知られている [23], [29]。

$$R_q = \begin{pmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2q_2q_3 - 2q_1q_4 & 2q_2q_4 + 2q_1q_3 \\ 2q_2q_3 + 2q_1q_4 & q_1^2 - q_2^2 + q_3^2 - q_4^2 & 2q_3q_4 - 2q_1q_2 \\ 2q_2q_4 - 2q_1q_3 & 2q_3q_4 + 2q_1q_2 & q_1^2 - q_2^2 - q_3^2 + q_4^2 \end{pmatrix} \quad (31)$$

これを用いると、球関数表現から回転行列への変換は次式で与えられる。

$$R_Y = \begin{pmatrix} \frac{Y_0 + 2Y_4}{3} & 2Y_5 - 2Y_3 & 2Y_6 + 2Y_2 \\ 2Y_5 + 2Y_3 & \frac{Y_0 - Y_4 + 3Y_7}{3} & 2Y_8 - 2Y_1 \\ 2Y_6 - 2Y_2 & 2Y_8 + 2Y_1 & \frac{Y_0 - Y_4 - 3Y_7}{3} \end{pmatrix} \quad (32)$$

ただし、推定された球関数表現  $\hat{Y}$  を用いる場合、一般に行列  $R_{\hat{Y}}$  は直交行列にならない。これは回転行列  $R$  を姿勢表現に用いた場合と同じ問題点である。

## 5. 姿勢推定実験

本研究で提案する球関数表現と他の姿勢表現方法（四元数、回転行列）を比較する姿勢推定実験を行った。

### 5.1 実験環境

実験には G. Peters [30] による “Dwarf” 人形の画像群を用いた。これらの画像は、物体を中央に固定し、それを取り囲むように半球格子状にカメラを配置して撮影したものである。図 3(a) に、カメラ配置位置を示す半球と、各カメラ位置で撮影された画像のいくつかを示す。方位角  $\phi^x = 0 \sim 360$ [deg] (100 ステップ)、仰角  $\phi^y = 0 \sim 90$ [deg] (25 ステップ) の刻み幅は 3.6[deg] で、合計 2500 枚の画像からなる。図中の  $x_0$  と示してある基準画像を撮影した位置が  $\phi^x = \phi^y = 0$ [deg] のカメラ位置である。

### 5.2 姿勢の与え方

図 3(b) に、画像と姿勢の回転軸との対応を示す。今回の実験では、人形を固定してカメラ位置を半球状に与える代わりに、カメラを固定して人形を  $x, y, z$  軸回りの順にそれぞれ  $\phi^x, \phi^y, \phi^z$  だけ回転したものとみなす。ここで  $x$  軸は物体の鉛直方向、 $y$

軸はカメラから見て右方向、 $z$  軸はカメラの光軸（奥行き）方向とする。 $\phi^x$  が方位角に対応し、 $\phi^y$  が仰角に、 $\phi^z$  は光軸回りの回転角に対応する。

回転軸  $\mathbf{n} = (n_1, n_2, n_3)^T$  回りに  $\theta$  回転する 3 次元回転を表す単位四元数は、次式で与えられる [2]。

$$q = \cos(\theta/2) + \sin(\theta/2)(n_1i + n_2j + n_3k), \quad \|\mathbf{n}\| = 1 \quad (33)$$

そこで、方位角  $\phi_j^x$ 、仰角  $\phi_j^y$  のカメラ位置から撮影された画像をさらに画面内で  $\phi_j^z$  だけ回転した  $x_j$  の姿勢を表す四元数  $q_j$  を、以下のように与える。

$$q_j^x = \cos(\phi_j^x/2) + \sin(\phi_j^x/2)(i + 0j + 0k) \quad (34)$$

$$q_j^y = \cos(\phi_j^y/2) + \sin(\phi_j^y/2)(0i + j + 0k) \quad (35)$$

$$q_j^z = \cos(\phi_j^z/2) + \sin(\phi_j^z/2)(0i + 0j + k) \quad (36)$$

$$q_j = q_j^z q_j^y q_j^x \quad (37)$$

### 5.3 学習・推定に用いた画像

学習には、方位角  $\phi^x = 0, 36, 72, \dots, 324$  [deg] の 10 ステップ、仰角  $\phi^y = 0, 36, 72$  [deg] の 3 ステップ、合計 30 枚の画像を学習した。また  $\phi^z = 0$ [deg] としたが、実際には頂点において光軸周りの回転と同等の効果が起きている。そのため、実質的に 3 自由度の姿勢を学習していることになる。

推定には、方位角仰角ともに 5 ステップ毎の画像を用いた。つまり、方位角については  $0, 3.6, 7.2, \dots$  [deg] の 100 ステップ、仰角については  $0, 3.6, 7.2, \dots$  [deg] の 25 ステップで、合計 2500 枚のうち学習した 30 枚を除く 2470 枚の画像の姿勢を推定した。

### 5.4 姿勢の推定方法

推定方法は、線形回帰 [3] を用いた。つまり、学習画像  $x_j$  に対応する姿勢を  $Y_j$  とすると、

$$Y = (Y_1, Y_2, \dots, Y_n), \quad X = (x_1, x_2, \dots, x_n) \quad (38)$$

として、 $Y = FX$  を満たす行列  $F$  を、 $X$  の最小ノルム型一般化逆行列  $X^+$  を用いて  $F = YX^+ = Y(X^T X)^{-1} X^T$  とした。

画像  $x$  の推定値  $\hat{Y}$  は、 $\hat{Y} = Fx$  として求めた。

### 5.5 比較する表現方法

画像  $x_j$  の姿勢表現方法の違いを検討するために、以下の 3 種類の姿勢表現を用いて推定と評価を行った。

- 球関数表現  $Y_j \in \mathbb{R}^9$  を学習。推定値  $\hat{Y}$  を回転行列  $R_{\hat{Y}}$  へ変換。

- 単位四元数  $q_j$  のベクトル  $q_j \in \mathbb{R}^4$  を学習。推定値  $\hat{q}$  を  $\bar{q} = \frac{\hat{q}}{\|\hat{q}\|}$  で正規化し、回転行列  $R_{\bar{q}}$  に変換。

- 回転行列  $R_j$  の要素を並べたベクトル  $R_j \in \mathbb{R}^9$  を学習。推定値  $\hat{R}$  を行列  $\hat{R}$  に戻し、Frobenius ノルムの意味で最も近い直交行列  $\bar{R}$  に変換（付録参照）。

これらを統一して評価するために、球関数と四元数の結果は回転行列へ変換した。

### 5.6 評価方法

評価は、Frobenius ノルムの平均二乗誤差 (RMSE) を用いた。つまり、真の姿勢に対応する回転行列  $R^{\text{true}}$  と、推定され



2007. <http://www.sciencedirect.com/science/article/B6V09-4K9C6GM-2/2/a1930d130b2c0f9fec6546579d50d11>.
- [13] Sanja Fidler, Danijel Skočaj, and Aleš Leonardis. Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling. *IEEE Trans. PAMI*, Vol. 28, No. 3, pp. 337–350, 2006. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2006.46>.
- [14] Hiroshi Murase and Shree K. Nayar. Visual learning and recognition of 3-D objects from appearance. *Intl. J. of Computer Vision*, Vol. 14, No. 1, pp. 5–24, 1995. <http://dx.doi.org/10.1007/BF01421486>.
- [15] 村瀬洋, シュリーナイヤー. 2次元照合による3次元物体認識パラメトリック固有空間法. 信学論, Vol. J77-D2, No. 11, pp. 2179–2187, 1994. <http://search.ieice.org/bin/summary.php?id=j77-d2.11.2179&category=D&year=1994&lang=J&abst=>.
- [16] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. Illumination from shadows. *IEEE Trans. PAMI*, Vol. 25, No. 3, pp. 290–300, 2003. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2003.1182093>.
- [17] Ping-Man Lam, Chi-Sing Leung, and Tien-Tsin Wong. Noise-resistant fitting for spherical harmonics. *IEEE Trans. Visualization and Computer Graphics*, Vol. 12, No. 2, pp. 254–265, 2006. <http://doi.ieeecomputersociety.org/10.1109/TVCG.2006.34>.
- [18] 岡部孝弘, 佐藤いまり, 佐藤洋一, 池内克史. キャストシャドウを用いた光源推定法: 球面調和関数展開に基づく解析. *CVIM*, Vol. 2002, No. 34, pp. 201–208, 2002. <http://fw8.bookpark.ne.jp/cm/ipsj/search.asp?flag=6&keyword=IPSJ-CVIM02133027&mode=PDF>.
- [19] Reiner Lenz. *Group Theoretical Methods in Image Processing*, Vol. 413 of *Lecture Notes Computer Science*. Springer Verlag, 1990. <http://staffwww.itn.liu.se/~reile/LNCS413/index.htm>.
- [20] Yi Ma, Stefano Soatto, Jana Košecká, and S. Shankar Sastry. *An Invitation To 3-D Vision*. Springer, 2004. <http://vision.ucla.edu/MASKS/>.
- [21] F. Sebastian Grassia. Practical parameterization of rotations using the exponential map. *Journal of Graphics Tools*, Vol. 3, No. 3, pp. 29–48, 1998. <http://jgt.akpeters.com/papers/Grassia98/>.
- [22] 内山勝, 中村仁彦. ロボットモーション. 岩波講座ロボット学. 岩波書店, 2004.
- [23] Andrew J. Hanson. Visualizing quaternions. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Courses*. ACM, 2005. <http://doi.acm.org/10.1145/1198555.1198701>.
- [24] John C. Hart, George K. Francis, and Louis H. Kauffman. Visualizing quaternion rotation. *ACM Trans. Graph.*, Vol. 13, No. 3, pp. 256–276, 1994. <http://doi.acm.org/10.1145/195784.197480>.
- [25] F. C. Park and Bahram Ravani. Smooth invariant interpolation of rotations. *ACM Trans. Graph.*, Vol. 16, No. 3, pp. 277–295, 1997. <http://doi.acm.org/10.1145/256157.256160>.
- [26] Ken Shoemake. Animating rotation with quaternion curves. *SIGGRAPH'85*, Vol. 19, No. 3, pp. 245–254, 1985. <http://doi.acm.org/10.1145/325165.325242>.
- [27] 阪野貴彦, 池内克史. 移動型レンジセンサによる歪み補正ICPアルゴリズム. *MIRU2006*, pp. 1225–1230, 2006. <http://www.cvl.iis.u-tokyo.ac.jp/papers/all/822.pdf>.
- [28] 竹内勝. 現代の球関数. 岩波書店, 1975.
- [29] 森田茂之. 微分形式の幾何学. 岩波書店, 2005.
- [30] Gabriele Peters, Barbara Zitova, and Christoph von der Malsburg. How to measure the pose robustness of object views. *Image and Vision Computing*, Vol. 20, No. 4, pp. 249–256, 2002. <http://ls7-www.cs.uni-dortmund.de/~peters/pages/research/modeladaptsys/modeladaptsys.vbarov.html>.
- [31] 玉木徹, 天野敏之, 金田和文. 球関数による3自由度回転物体画像系列の展開と画像補間・姿勢推定への応用. *MIRU2008*, 2008. in print.
- [32] 森口繁一, 宇田川金圭久, 一松信. 数学公式 III —特殊関数—. 岩波書店, 1960.
- [33] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. U.S. Dept. of Commerce, 10 edition, 1972. <http://www.math.sfu.ca/~cbm/aands/>.
- [34] Art Matheny and Dmitry B. Goldgof. The use of three- and four-dimensional surface harmonics for rigid and non-rigid shape recovery and representation. *IEEE Trans. PAMI*, Vol. 17, No. 10, pp. 967–981, 1995. <http://doi.ieeecomputersociety.org/10.1109/34.464561>.
- [35] Zhengyou Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research, 1998. [http://research.microsoft.com/research/pubs/view.aspx?tr\\_id=212](http://research.microsoft.com/research/pubs/view.aspx?tr_id=212).

## 付 録

### 1. Gegenbauer 陪関数

Gegenbauer の多項式  $C_n^\nu$  は超幾何関数を用いて定義される [32] が、ここでは陽に書き表されるもの [33] を示す。  $n \neq 0$  かつ  $n > -\frac{1}{2}$  の場合、

$$C_n^\nu(x) = \frac{1}{\Gamma(\nu)} \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^m \Gamma(\nu + n - m)}{m! (n - 2m)!} (2x)^{n-2m} \quad (\text{A.1})$$

ここで  $\Gamma$  はガンマ関数であり、  $\Gamma(n+1) = n!$  ( $n \in \mathbb{N}$ ) である。

Gegenbauer の陪関数 (同伴関数)  $C_n^{\nu,m}$  [28] は

$$C_n^{\nu,m}(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} C_n^\nu(x) \quad (\text{A.2})$$

で定義される。Legendre 陪関数  $P_n^m$  との間には

$$C_n^{\frac{1}{2},m}(x) = P_n^m(x) = P_n^{-m}(x) \quad (\text{A.3})$$

という関係がある。

### 2. 一般の球関数の定義

$0 \leq \theta_1, \dots, \theta_{n-1} \leq \pi, 0 \leq \theta_n < 2\pi$  のとき、  $\ell_1, \dots, \ell_n \in \mathbb{Z}$ ,  $\ell_1 \geq \ell_2 \geq \dots \geq \ell_{n-1} \geq |\ell_n|$  に対して、

$$Y_{\ell_1, \dots, \ell_n}(\theta_1, \dots, \theta_n) = \sqrt{b_{\ell_1, \dots, \ell_n}} \left\{ \prod_{k=1}^{n-1} C_{\ell_k}^{\nu_k, \ell_{k+1}}(\cos \theta_k) \right\} e^{-\ell_n \theta_n} \quad (\text{A.4})$$

とすると、  $\{Y_{\ell_1, \dots, \ell_n}\}$  は (適当な測度の元で)  $S^n$  の  $L_2$  空間  $L_2(S^n)$  の完全正規直交系となる [28]。ここで

$$\nu_k = \frac{1}{2}(n-k), \quad 1 \leq k \leq n-1 \quad (\text{A.5})$$

$$b_{\ell_1} = 1 \quad (\text{A.6})$$

$$b_{\ell_1, \dots, \ell_n} = \frac{2^{n-1} (n-2)! (n-3)! \cdots 1!}{(n-1)!}$$

$$\left\{ \prod_{k=1}^{n-2} \frac{(\ell_k - \ell_{k+1})! (\ell_k + \nu_k)}{(2\nu_k - 1 + \ell_k + \ell_{k+1})!} \right\} \cdot \frac{(\ell_{n-1} - |\ell_n|)! (2\ell_{n-1} + 1)}{2(\ell_{n-1} + |\ell_n|)!} \quad (\text{A.7})$$

### 3. $SO(3)$ の連続直交関数系

3次元実射影空間  $P_3\mathbb{R}$  は、3次元球面  $S^3$  上の点  $q \in S^3$  と  $-q \in S^3$  を同一視して得られる多様体である [28], [29]。そこで  $C^\infty(P_3\mathbb{R}) \subset C^\infty(S^3)$  とみなすと、 $\{Y_{\ell_1, \ell_2, \ell_3}; \ell_1 \in 2\mathbb{Z}, \ell_2, \ell_3 \in \mathbb{Z}, \ell_1 \geq \ell_2 \geq |\ell_3|\}$  は  $L_2(P_3\mathbb{R})$  の完全正規直交系である [28]。

一方、 $SO(3)$  と  $P_3\mathbb{R}$  は同じものとみなしてよい (微分同相である [29])。そこで、本研究では上記を  $L_2(SO(3))$  の直交系であるとみなす。このため、 $0 \leq \theta_1 \leq \frac{\pi}{2}$  となる。

$SO(3)$  上の球関数は、 $\ell_1 \in \mathbb{Z}$  とすると、

$$\begin{aligned} & Y_{2\ell_1, \ell_2, \ell_3}(\theta_1, \dots, \theta_n) \\ &= \sqrt{b_{2\ell_1, \ell_2, \ell_3}} C_{2\ell_1}^{1, \ell_2}(\cos \theta_1) C_{\ell_2}^{\frac{1}{2}, \ell_3}(\cos \theta_2) e^{-\ell_3 \theta_3 i} \quad (\text{A.8}) \end{aligned}$$

ここで  $C_n^{\frac{1}{2}, m}(x) = P_n^m(x)$  とすると、式 (12) となる。

Matheny と Goldgof [34, App. A.4] は、第2種 Chebyshev 多項式  $U_n$  を用いて、4次元 (3D 空間+時間) の球面調和関数を導出している。これは上式の  $C_{2\ell_1}^{1, \ell_2}(x)$  を、 $C_n^1(x) = U_n(x)$  という関係 [33, p.779] を用いて  $\frac{d^{\ell_2}}{dx^{\ell_2}} U_{2\ell_1}(x)$  で置き換えると、同一のものである。ただし、5次元以上に拡張できないこと、第2種 Chebyshev 多項式に有用な微分公式がないことが問題点である。その点では、前述の Gegenbauer 陪関数を用いた一般的な定式化のほうが好ましい [31]。

Lenz [19, p.66] は Jacobi 多項式を用いて  $S^4$  と  $SO(3)$  の球面調和関数を導出している。これも Jacobi 多項式と Gegenbauer 多項式、Chebyshev 多項式の関係 [32, p.95] を用いれば、同一のものに変換できる。

### 4. 行列の直交化

ある行列  $A$  を Frobenius ノルム  $\|\cdot\|_F$  の意味で最も近い直交行列  $R = (r_1^T, r_2^T, r_3^T)^T$  を求めるには、ラグランジュ未定係数法を用いて以下の目的関数を最小化する。

$$\begin{aligned} & \|A - R\|_F - \lambda_1(r_1^T r_1 - 1) - \lambda_2(r_2^T r_2 - 1) \\ & - \lambda_3(r_3^T r_3 - 1) - \lambda_4 r_1^T r_2 - \lambda_5 r_1^T r_3 - \lambda_6 r_2^T r_3 \quad (\text{A.9}) \end{aligned}$$

この解は、 $A$  の特異値分解を  $A = V\Sigma U^T$  としたとき  $R = VU^T$  である (例えば [35, App. C] 等を参照)。