

見えに基づく姿勢推定のための 複素部分空間と四元数部分空間の構築について

玉 木 徹[†] 天 野 敏 之^{††}
金 田 和 文[†] 市 原 由 美 子[†]

本稿では、3次元で1軸回転する物体の画像列を表現する複素部分空間と、3軸回転する物体の画像列を表現する四元数部分空間を構成し、それを用いた見えに基づく画像からの物体の姿勢推定手法を提案する。非線形回帰や多様体学習のような手法と異なり、本手法は線形部分空間を用いたアプローチを取り、実際の画像を用いて手法の有効性を検証する。

On the complex and quaternion subspaces for view-based pose estimation

TORU TAMAKI,[†] TOSHIYUKI AMANO,^{††} KAZUFUMI KANEDA[†]
and YUMIKO ICHIHARA[†]

This paper proposes a new method for construction of a complex subspace for rotation about a single axis and a quaternion subspace for 3DOF rotation of a 3D object, then demonstrates with the proposed method view-based pose estimation.

1. はじめに

画像に写る物体の姿勢パラメータを推定する問題はコンピュータビジョンの重要な問題の一つであり、これまでに多くの研究がなされている。その多くは剛体の既知形状モデルに対する運動や投影の3次元幾何を用い最適化問題を解く、モデルに基づく (model-based) 手法¹⁾である。それに対して、物体の形状や3次元空間での投影といった幾何学的な撮像過程を仮定せず、物体の見た目である画像そのものを学習する見えに基づく (appearance-based, view-based) 手法が数多く提案されている。

見えに基づく姿勢推定の手法は、一般に画像と姿勢推定パラメータとの関係を学習する。3次元空間中の物体がある軸回りに回転する場合、学習においては軸回りの回転角度 θ_j とその角度のときに撮影された画像 \mathbf{x}_j を対にしたものを学習セット $\{\theta_j, \mathbf{x}_j\}_{j=0,1,\dots}$ とし、テスト時には未知の画像 \mathbf{x} に対応する姿勢 θ を推定する。これらを実現する研究には画像中の特徴点を利用する手法^{2)~4)} や多様体学習を用いる手法^{5),6)}

なども提案されている。しかし、画像 \mathbf{x} をそのままベクトルとみなす線形部分空間に基づく認識手法は以前から研究されており、依然として重要である^{7),8)}。これらの研究には、固有空間を用いる手法とパラメータとの関係を学習する手法がある。

固有空間に基づく姿勢推定手法には、画像ベクトルが低次元固有空間で形成する多様体をスプラインで表現するパラメトリック固有空間法^{9),10)}があるが、線形代数的な推定を行うことはできない。一般に多様体だけでなく、固有ベクトル自体も画像によって変化するが、2次元画像を回転した特殊な状況における基底ベクトルは DCT 基底になることが知られており^{11)~15)}、3次元物体の1軸回転画像の基底を DCT で近似して姿勢推定を行う手法^{13),15)} や3軸への適用¹⁶⁾も提案されている。

一方、入力画像と出力パラメータを関連付ける線形写像を推定する手法が多数提案されている。この関連付けには、二つの与えられた学習データ同士の相関を最大化する CCA (canonical correlation analysis) やその kernel 版^{6),8),17)}、線形回帰¹⁸⁾ やその変形^{19)~21)}、サポートベクトル回帰^{22)~24)}などが用いられている。

これらに対して筆者らは、1軸回転する物体を撮影した画像列は物体が一回りすると同じ画像に戻ってくることに着目し、 n 枚の画像列に働く n 位の巡回群を用いるという、従来とは全く異なるアプローチを提案し、学習画像間の補間^{25),26)}と姿勢推定²⁷⁾に適用

[†] 広島大学大学院工学研究科情報工学専攻
Department of Information Engineering, Graduate
School of Engineering, Hiroshima University
^{††} 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology

した。この手法は、画像列中のある画像から次の画像への遷移は、部分空間における複数の2次元回転で表現する。そのため、未知画像の姿勢推定は、基準となる画像からの部分空間における回転量の推定に置き換えることができる²⁷⁾。しかしこの手法には、一回りすると元に戻るという巡回群による定式化では3自由度で回転する物体の画像列を扱えないという問題点があった。

そこで本研究では、巡回群による定式化のうち「未知画像の姿勢推定が、基準となる画像からの部分空間における回転量の推定に置き換えることができる」という部分にのみ着目し、これを実現する部分空間を構成する手法を提案する。

部分空間における回転の表現として、1軸回転の場合は偏角で角度を表せる複素数²⁵⁾を、3軸回転の場合は実部が回転量を、三つの虚部が回転軸を表す四元数²⁸⁾を用いる。1軸回転の角度 θ は 2π で0に戻るため、通常は $\sin \theta, \cos \theta$ の2パラメータを用いる^{8),17)~19)}が、ここでは1つの複素数を用いる。また3軸回転の角度表現にはさまざまなものがあるが、姿勢推定には固定角が用いられることが多い^{9),10),22),23)}。しかしこれらはそれぞれ 2π で不連続であり、 \sin, \cos を用いて適切なパラメータ化をしたとしても、ある姿勢で角度が連続にならないジンバルロックの問題が残ってしまう。そこで本研究では、3軸回転の姿勢を表現するために単位四元数を用いる。複素数も四元数も体であり、1軸回転の手法を自然に3軸回転に拡張することができる。

提案手法は、まず学習画像同士の回転を複素数または四元数で表現する部分空間を劣決定の連立方程式として定式化し、零ベクトルを求めることで複素数または四元数部分空間を構築する。姿勢推定では、未知画像をその部分空間に投影して基準画像からの回転を表す複素数または四元数を求める。本稿では、提案手法を2軸回転する物体の実画像系列に適用した実験を示し、本手法の特性などを議論する。

2. 複素対角化を用いた1軸回転の姿勢推定

ここではまず、巡回群行列を用いた1軸回転の姿勢推定の手法^{25)~27)}を概説する。

2.1 画像の遷移を表す行列 G の分解

ある物体を3次元の1軸で等間隔に回転させて撮影した n 枚の画像を $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jN})^T \in \mathbb{R}^N$ 、その画像を撮影した角度を $\theta_j = 2j\pi/n$ とする。ここで $j = 0, 1, \dots, n-1$ である。さらに実際の状況として $N > n$ 、つまり画素数は画像数より大きいことを仮定する。

まず画像 \mathbf{x}_j を \mathbf{x}_{j+1} へと遷移させるような、物体の1軸回転の1ステップ分に相当する以下のような作

用 G を考える。

$\mathbf{x}_{j+1 \bmod n} = G\mathbf{x}_j$, $\mathbf{x}_j = G^j\mathbf{x}_0$, $G^n = I_n$ (1)
ここで I_n は $n \times n$ 単位行列である。この G を、画像列に働く n 位の巡回群の生成元の表現行列と呼ぶ。この群の作用は、以下のように画像の線形変換で書くことができる。

$[\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_{n-1} \ \mathbf{x}_0] = G[\mathbf{x}_0 \ \mathbf{x}_1 \ \dots \ \mathbf{x}_{n-2} \ \mathbf{x}_{n-1}]$ (2)
上式を $X_1 = GX_0$ とおいて、 G を求める。

まず式(2)は劣決定の連立方程式であるので、 X_0 の最小ノルム型一般化逆行列²⁹⁾ X_0^+ を用いて

$$G = X_1 X_0^+ = X_1 (X_0^T X_0)^{-1} X_0^T \quad (3)$$

$$= X_1 V \Sigma^{-1} E^T \quad (4)$$

が得られる。ここで X_0 の特異値分解を $X_0 = E \Sigma V^T$ とする。

また $n \times n$ の列置換行列 M を用いて、 X_1 は次のように書ける。

$$X_1 = X_0 \begin{pmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \\ & & & & 1 & 0 \end{pmatrix} = X_0 M \quad (5)$$

ここで M の対角化²⁷⁾を、複素対角行列 D とユニタリ行列 W を用いて $M = W^H D W$ とする。

最終的に、 G の分解は以下のような行列の積で表現できる。

$$G = U_2 D U_1 \quad (6)$$

$$U_1 = W V \Sigma^{-1} E^T, \quad U_2 = E \Sigma V^T W^H \quad (7)$$

$$U_1 U_2 = I_n \quad (8)$$

ここで、1の原始 n 乗根を $\zeta_n = e^{\frac{2\pi i}{n}}$ とすると

$$D = \text{diag}(1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1}) \quad (9)$$

である。

この G を分解する複素行列 U_1, U_2 は、画像の姿勢を表す複素部分空間への投影と考えることができる。

つまり $\mathbf{x}'_j = U_1 \mathbf{x}_j \in \mathbb{C}^n$ とおくと

$$\mathbf{x}_{j+\alpha \bmod n} = G^\alpha \mathbf{x}_j = U_2 D^\alpha U_1 \mathbf{x}_j \quad (10)$$

$$\mathbf{x}'_{j+\alpha \bmod n} = D^\alpha \mathbf{x}'_j \quad (11)$$

が成り立ち、 D を何乗するかで画像の遷移を決定することができる。この遷移は学習済みの $\alpha = 0, 1, \dots, n-1$ の場合は厳密に成り立つため、それ以外の未学習画像に対しても近似的に成り立つことを仮定する。すると未知の画像 \mathbf{x} に対して

$$\mathbf{x}' = D^\alpha \mathbf{x}'_0 \quad (12)$$

を満たす α を、基準画像 \mathbf{x}_0 に対する姿勢の推定値とすることができる。

2.2 部分空間における姿勢推定

この部分空間での画像遷移を利用して、新たな画像 \mathbf{x} の姿勢パラメータ θ を推定する方法²⁷⁾を二つ示す。

^{*} 式(2)が劣決定であるため。

^{**} これらはブロック対角化を用いた手法²⁷⁾と本質的に同じである。

2.2.1 距離に基づく姿勢推定

一つ目は、以下のように部分空間におけるユークリッド距離が最小になる指数 α を姿勢パラメータの推定値 θ とする方法である。

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \|\mathbf{x}' - D^{\alpha} \mathbf{x}'_0\|^2, \quad \theta = \frac{2\pi}{n} \hat{\alpha} \quad (13)$$

2.2.2 角度に基づく姿勢推定

もう一つは、対角行列 D の一つの行に対応する 1 次元複素部分空間を用いる方法である。

$\mathbf{x}' = (x'_0, x'_1, \dots)^T, \mathbf{x}'_0 = (x'_{00}, x'_{01}, \dots)^T$ として式 (12) を要素ごとにみると、 D^{α} は対角行列なので、

$$(x'_0, x'_1, x'_2, \dots)^T = (x'_{00}, \zeta_n^{\alpha} x'_{01}, \zeta_n^{2\alpha} x'_{02}, \dots)^T \quad (14)$$

となる。つまり \mathbf{x}' は \mathbf{x}'_0 の各要素に絶対値 1 の複素数 ζ_n^{α} をかけたものである。特に x'_1 は x'_{01} の偏角を $\frac{2\pi}{n}\alpha$ 倍したものであり、これは x'_1 と x'_{01} がなす角に相当する。

そこで U_1 の ζ_n に対応する行を

$$\mathbf{u}^T = (0, 1, 0, \dots, 0) \quad U_1 \in \mathbb{C}^n \quad (15)$$

と抜き出すと、

$$x'_1 = \zeta_n^{\alpha} x'_{01} \quad (16)$$

は

$$\mathbf{u}^T \mathbf{x} = \zeta_n^{\alpha} \mathbf{u}^T \mathbf{x}_0 \quad (17)$$

と書き直せる。したがって、他の \mathbf{x}' と \mathbf{x}'_0 の要素の関係を見ずに、複素平面における x'' と x'_0 のなす角

$$\theta = \angle \zeta_n^{\alpha} = \angle \frac{\mathbf{u}^T \mathbf{x}}{\mathbf{u}^T \mathbf{x}_0} \quad (18)$$

だけを見ることで、姿勢を推定することができる。

3. 1 軸回転のための複素数部分空間の構築

前節で述べた 1 軸回転の姿勢推定手法を 3 自由度の回転の推定に拡張する。1 軸回転の場合は 2.1 節のように画像の遷移を一つの巡回群 G で表現することができるが、3 軸回転の場合は式 (2) に対応する表現が存在しない。そこで本稿では、行列の対角化という視点からではなく姿勢を表現する部分空間の構築という観点から、1 軸回転の姿勢推定手法を定式化しなおし、それを 3 軸回転に拡張する。

2.2.2 節で述べたように、複素行列 U_1 のある複素行ベクトル \mathbf{u}^T を用いると、ある画像 \mathbf{x} と基準画像 \mathbf{x}_0 の \mathbf{u}^T による投影 $x'_1, x'_{01} \in \mathbb{C}$ は、ある複素数 ζ_n^{α} を用いて $x'_1 = \zeta_n^{\alpha} x'_{01}$ と表すことができた。これは、推定したい画像をこの部分空間 \mathbf{u}^T へ投影すると、基準画像からの姿勢が複素数の積によって表現できることを示している。

ここでは、逆にそのような \mathbf{u}^T が存在するという仮定から出発する。つまり、ある $\mathbf{u}^T \in \mathbb{C}^n$ が存在して、 \mathbf{x}_j, λ_j が与えられたとき、以下の式を満たす。

$$\mathbf{u}^T \mathbf{x}_j = \lambda_j \mathbf{u}^T \mathbf{x}_0, \quad j = 0, 1, \dots, n-1 \quad (19)$$

ここで $\mathbf{u}^T \mathbf{x}_j, \mathbf{u}^T \mathbf{x}_0 \in \mathbb{C}$ であり、また $\lambda_j = \zeta_n^j$ は画像 \mathbf{x}_j の姿勢 $\angle \lambda_j = 2j\pi/n$ を表す複素数である。

$N > n$ を仮定すると、式 (19) は劣決定の連立方程式であり、以下の二乗誤差を 0 にする \mathbf{u} がかならず存在する。

$$\min_{\mathbf{u}} \sum_{j=0}^{n-1} \|\mathbf{u}^T \mathbf{x}_j - \lambda_j \mathbf{u}^T \mathbf{x}_0\|^2 \quad (20)$$

ここで上記の連立方程式を以下のように変形する。

$$\begin{pmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_{n-1}^T \end{pmatrix} \mathbf{u} = \begin{pmatrix} \lambda_1 \mathbf{x}_0^T \\ \lambda_2 \mathbf{x}_0^T \\ \vdots \\ \lambda_{n-1} \mathbf{x}_0^T \end{pmatrix} \mathbf{u} \quad (21)$$

$$\begin{pmatrix} \mathbf{x}_1^T - \lambda_1 \mathbf{x}_0^T \\ \mathbf{x}_2^T - \lambda_2 \mathbf{x}_0^T \\ \vdots \\ \mathbf{x}_{n-1}^T - \lambda_{n-1} \mathbf{x}_0^T \end{pmatrix} \mathbf{u} = \mathbf{0} \quad (22)$$

上式を $A\mathbf{u} = \mathbf{0}$ とおくと、ベクトル \mathbf{u} は行列 A の右零ベクトルである。これは $\operatorname{rank}(A) = n < N$ のとき必ず存在する。

式 (22) の幾何的な解釈は次のような興味深いものである。元の式 (19) では、画像 \mathbf{x} を \mathbf{u}^T が表す 1 次元複素部分空間に投影すると、基準画像 \mathbf{x}_0 に対する姿勢が複素数 λ の偏角で表現されていた。しかし式 (22) では、 \mathbf{u}^T に投影する前に基準画像に姿勢を表す複素数をかけ、推定したい画像との差をとると、そのベクトルが部分空間 \mathbf{u}^T と直交する。

この \mathbf{u} を用いて、新たな画像 \mathbf{x} にやはり以下の関係が成り立つと仮定する。

$$\mathbf{u}^T \mathbf{x} = \lambda \mathbf{u}^T \mathbf{x}_0 \quad (23)$$

そして式 (18) と同様に、次式で姿勢 θ を推定する。

$$\theta = \angle \lambda = \angle (\mathbf{u}^T \mathbf{x}) (\mathbf{u}^T \mathbf{x}_0)^{-1} = \angle \frac{\mathbf{u}^T \mathbf{x}}{\mathbf{u}^T \mathbf{x}_0} \quad (24)$$

4. 3 軸回転のための四元数部分空間の構築

前述の考えを 3 軸回転の場合に拡張する。つまり、ある部分空間 \mathbf{u} が存在し、推定したい画像をそこへ投影すると、基準画像からの姿勢がある数の積によって表現できることを仮定する。1 軸回転の場合は、対角化による定式化からの類推に基づいており、また回転角が複素数の偏角という 1 パラメータで表現できるという理由から、姿勢は複素数 \mathbb{C} の積によって表現していた。ここでは 3 軸回転を考えるため、姿勢の表現は複素数ではなく四元数 \mathbb{H} の積によって表すことにする²⁸⁾。以下行列はすべて四元数行列である。

4.1 四元数部分空間による定式化

前節と同様に、ある $\mathbf{u} \in \mathbb{H}^n$ が存在して、 \mathbf{x}_j, p_j が与えられたとき、以下の式を満たすという仮定から出発する。

$$\mathbf{x}_j^T \mathbf{u} = p_j \mathbf{x}_0^T \mathbf{u}, \quad j = 1, 2, \dots, n-1 \quad (25)$$

つまり、ある画像 x_j の部分空間への投影 $x_j^T \mathbf{u} \in \mathbb{H}$ は、基準画像 x_0 の投影 $x_0^T \mathbf{u} \in \mathbb{H}$ に四元数 p_j を左からかけたもの*に等しく、その p_j が画像 x_j の姿勢を表しているとする。ここで、 p_j は 3 次元回転を表すため、単位四元数 $|p_j| = 1$ とする。

$N > n$ を仮定すると、式 (25) は劣決定の連立方程式であり、以下の二乗誤差を 0 にする \mathbf{u} がかならず存在する。

$$\min_{\mathbf{u}} \sum_{j=1}^{n-1} \|x_j^T \mathbf{u} - p_j x_0^T \mathbf{u}\|^2 \quad (26)$$

ここで上記の連立方程式を以下のように変形する。

$$\begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_{n-1}^T \end{pmatrix} \mathbf{u} = \begin{pmatrix} p_1 x_0^T \\ p_2 x_0^T \\ \vdots \\ p_{n-1} x_0^T \end{pmatrix} \mathbf{u} \quad (27)$$

$$\begin{pmatrix} x_1^T - p_1 x_0^T \\ x_2^T - p_2 x_0^T \\ \vdots \\ x_{n-1}^T - p_{n-1} x_0^T \end{pmatrix} \mathbf{u} = \mathbf{0} \quad (28)$$

上式を $A\mathbf{u} = \mathbf{0}$ とおくと、ベクトル \mathbf{u} は行列 A の右零ベクトルである。これは $\text{rank}(A) = n < N$ のとき必ず存在する。

4.2 右零ベクトルの求め方

ではこの右零ベクトル \mathbf{u} を求めよう。行列 A は、各行が張る部分空間を表すとみなせる。 \mathbf{u} はその部分空間 A に直交すればよい。したがって、適当なベクトル \mathbf{u}_0 が与えられたとき、それを部分空間 A に投影したベクトルを \mathbf{u}'_0 とすると、 $\mathbf{u} = \mathbf{u}_0 - \mathbf{u}'_0$ は明らかに求めたい右零ベクトルの一つである。

\mathbf{u}_0 の A への投影 \mathbf{u}'_0 は次式で求める。

$$\mathbf{u}'_0 = A^H (AA^H)^{-1} A \mathbf{u}_0 \quad (29)$$

ここで H は共役四元数転置を取ることを意味する。したがって、

$$\mathbf{u} = (I - A^H (AA^H)^{-1} A) \mathbf{u}_0 \quad (30)$$

である。

4.3 3 軸回転の姿勢推定

新たな画像 \mathbf{x} の姿勢を以下のように推定する。その画像の姿勢を表す四元数を p とすると、

$$\mathbf{x}^T \mathbf{u} = p x_0^T \mathbf{u} \quad (31)$$

が成り立つとする。よって

$$p = (\mathbf{x}^T \mathbf{u}) (\mathbf{x}_0^T \mathbf{u})^{-1} \quad (32)$$

$$= (\mathbf{x}^T \mathbf{u}) \frac{\overline{\mathbf{x}_0^T \mathbf{u}}}{|\mathbf{x}_0^T \mathbf{u}|} = \frac{(\mathbf{x}^T \mathbf{u}) (\mathbf{u}^H \mathbf{x}_0)}{|\mathbf{x}_0^T \mathbf{u}|} \quad (33)$$

となる。ここで \bar{p} は p の共役四元数を意味する。この

* 四元数は積について非可換であるので、 p_j と \mathbf{u} の積の順序に注意。

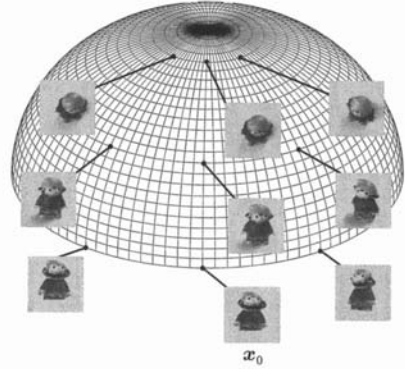


図 1 実験に用いた Dwarf 画像²⁾とその撮影カメラ位置

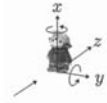


図 2 Dwarf 画像の姿勢と回転軸の対応関係

四元数 p は、(基準画像 x_0 からの) 回転の推定値を表す。

ただし、基準画像 x_0 の四元数が単位四元数になることを保障するために、以下のように正規化する。

$$p = \frac{(\mathbf{x}^T \mathbf{u}) (\mathbf{u}^H \mathbf{x}_0)}{|\mathbf{x}_0^T \mathbf{u}|^2} \quad (34)$$

5. 実 験

実験には Gabriele²⁾ 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 1 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。方位角と仰角の刻み幅はそれぞれ 3.6[deg] であり、方位角については 0 ~ 360[deg] の 100 ステップ、仰角については 0 ~ 90[deg] の 25 ステップで、合計 2500 枚の画像からなる。図 1 中の x_0 と示してある基準画像を撮影した位置が、方位角仰角ともに 0[deg] のカメラ位置である。

図 2 に、画像と姿勢の回転軸との対応を示す。今回の実験では、カメラを固定して、物体を固定軸周りに回転したものとみなして、姿勢パラメータである四元数を与えた。つまり、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 軸回りに回転した後に y 軸回りに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) 方向とし、 x 軸周りの回転が方位角に対応し、 y 軸周りの回転が仰角に対応する。したがって、方位角 θ_j 、仰角 ϕ_j の

カメラ位置から撮影された画像 x_j の姿勢を表す四元数 p_j を、以下のように与えた。

$$p_j^x = \cos(\theta_j/2) + \sin(\theta_j/2)(i + 0j + 0k) \quad (35)$$

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

$$p_j = p_j^y p_j^x \quad (37)$$

ここで $\cos(\theta/2) + \sin(\theta/2)(ai + bj + ck)$ は、回転軸 (a, b, c) 回りに角度 θ だけ回る回転を表す四元数である。この変換から分かるように、この実験画像では物体の2軸回転しか扱っていないが、四元数 p_j を扱う点では3軸回転と同等である。右零ベクトルを求めるための u_0 として、すべての値が $1 + i + j + k$ であるものを用いた。

学習には、方位角仰角ともに10ステップ毎の画像を用いた。つまり、方位角については $0, 36, 72, \dots, 324$ [deg] の10ステップ、仰角については $0, 36, 72$ [deg] の3ステップで、合計30枚の画像を学習した。

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 $0, 18, 36, 54, 72$ [deg] の5ステップで、合計500枚の画像の姿勢を推定した。推定された四元数 \hat{p} は単位四元数ではない場合があるため、 $|\hat{p}| = 1$ となるように正規化した。真の四元数 p と \hat{p} は、差を表す四元数

$$p_d = \cos(\theta_d/2) + \sin(\theta_d/2)(ai + bj + ck) \quad (38)$$

を用いて $p = p_d \hat{p}$ と書けるので、 $p_d = \hat{p} p^{-1}$ の実部の \cos^{-1} をとり θ_d を推定誤差とした。図3に姿勢推定結果を示す。学習した画像については、推定誤差は0になっている ($\phi = 0, 36, 72$ の線の、横軸 $\theta = 0, 36, 72, \dots$ は、推定誤差が0)。しかし、学習していない画像は推定誤差がかなり大きくなっており、推定精度を向上させることが今後の課題である。

学習画像に近い画像ではどの程度推定できるのかを見るために、基準画像 x_0 に一樣ノイズを加えて推定した結果を図4に示す。ノイズがない場合には誤差が0であるが、ノイズが増えるにつれて線形に推定誤差が増えている。ロバストな推定の実現も今後検討する必要がある。

6. おわりに

本稿では、3次元で1軸回転する物体の画像列を表現する複素部分空間と、3軸回転する物体の画像列を表現する四元数部分空間を構成し、それを用いた見えに基づく画像からの物体の姿勢推定手法を提案した。学習画像同士の回転を複素数または四元数で表現する部分空間を劣決定の連立方程式として定式化し、零ベクトルを求めることで複素数または四元数部分空間を構築した。姿勢推定では、未知画像をその部分空間に投影して基準画像からの回転を表す複素数または四元数を求めた。しかし、学習した画像以外の推定誤差は大きく、今後の改良の余地がある。これは右零ベクトル

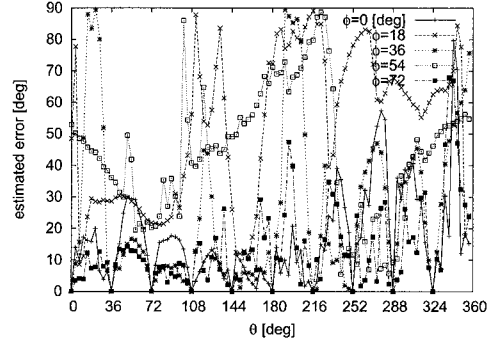


図3 推定結果。横軸は推定に用いた画像の x 軸周りの角度 (方位角) θ を、それぞれの線は y 軸回りの角度 (仰角) ϕ を示し、縦軸は推定誤差 θ_d を表している。

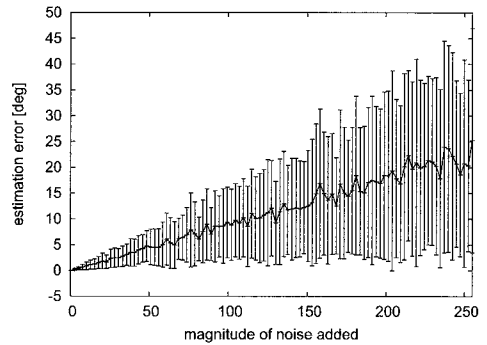


図4 基準画像 x_0 の各画素 x に一樣ノイズを加え $x \rightarrow x + d, d \in [-D, D]$ とした場合の推定誤差 θ_d (10回の平均と標準偏差)。縦軸は推定誤差を、横軸は各画素値 ($x \in [0, 255]$) に加えた一樣乱数の大きさ D を示す。

ル u を求めるときに何の制約もなく任意の u_0 を用いていることが原因の一つであると思われる。今後は提案手法に必要な制約条件を検討する。また学習画像を増やして学習データの粗密が推定結果に与える影響を調査したり、実際に3軸回転する物体画像を用いて実験することなども今後行う予定である。

参考文献

- 1) Lowe, D. G.: Fitting Parameterized Three-Dimensional Models to Images, *IEEE Trans. PAMI*, Vol.13, No.5, pp.441-450 (1991). online.
- 2) Peters, G., Zitova, B. and vonder Malsburg, C.: How to Measure the Pose Robustness of Object Views, *Image and Vision Computing*, Vol.20, No.4, pp.249-256 (2002). online.
- 3) Peters, G. and vonder Malsburg, C.: View Reconstruction by Linear Combination of Sam-

- ple Views, *British Machine Vision Conference*, Vol.1, pp.223–232 (2001). online.
- 4) Peters, G.: Efficient Pose Estimation Using View-Based Object Representations, *Machine Vision and Applications*, Vol.16, No.1, pp.59–63 (2004). online.
 - 5) Gong, H., Pan, C., Yang, Q., Lu, H. and Ma, S.: A Semi-Supervised Framework for Mapping Data to the Intrinsic Manifold, *ICCV2005*, Vol.1, pp.98–105 (2005). online.
 - 6) Sun, T. and Chen, S.: Locality preserving CCA with applications to data visualization and pose estimation, *Image and Vision Computing*, Vol.25, No.5, pp.531–543 (2007). online.
 - 7) Chen, P. and Suter, D.: An Analysis of Linear Subspace Approaches for Computer Vision and Pattern Recognition, *Intl. J. of Computer Vision*, Vol.68, No.1, pp.83–106 (2006). online.
 - 8) Fidler, S., Skočaj, D. and Leonardis, A.: Combining Reconstructive and Discriminative Subspace Methods for Robust Classification and Regression by Subsampling, *IEEE Trans. PAMI*, Vol.28, No.3, pp.337–350 (2006). online.
 - 9) 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一, *信学論*, Vol.J77-D2, No.11, pp.2179–2187 (1994). online.
 - 10) Murase, H. and Nayar, S.K.: Visual learning and recognition of 3-D objects from appearance, *Intl. J. of Computer Vision*, Vol.14, No.1, pp.5–24 (1995). online.
 - 11) Uenohara, M. and Kanade, T.: Optimal Approximation of Uniformly Rotated Images: Relationship Between Karhunen-Loeve Expansion and Discrete Cosine Transform, *IEEE Trans. IP*, Vol.7, No.1, pp.116–119 (1998). online.
 - 12) Park, R.-H.: Comments on "Optimal Approximation of Uniformly Rotated Images: Relationship Between Karhunen-Loève Expansion and Discrete Cosine Transform", *IEEE Trans. IP*, Vol.11, No.3, pp.332–334 (2002). online.
 - 13) Chang, C.-Y., Maciejewski, A. and Balakrishnan, V.: Fast Eigenspace Decomposition of Correlated Images, *IEEE Trans. IP*, Vol. 9, No.9, pp.1937–1949 (2000). online.
 - 14) Jorgan, M., Žagar, E. and Leonardis, A.: Karhunen-Loève Expansion of a Set of Rotated Templates, *IEEE Trans. IP*, Vol.12, No.7, pp. 817–825 (2003). online.
 - 15) Sengel, M. and Bischof, H.: Efficient representation of in-plane rotation within a PCA framework, *Image and Vision Computing*, Vol.23, pp. 1051–1059 (2005). online.
 - 16) Saitwal, K., Maciejewski, A.A. and Roberts, R. G.: Eigendecomposition of Correlated Images Characterized by Three Parameters, *2006 IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.203–207 (2006). online.
 - 17) Melzer, T., Reiter, M. and Bischof, H.: Appearance models based on kernel canonical correlation analysis, *Pattern Recognition*, Vol.36, pp.1961–1971 (2003). online.
 - 18) Okatani, T. and Deguchi, K.: Yet Another Appearance-Based Method for Pose Estimation Based on a Linear Model, *IAPR Workshop on Machine Vision Applications 2000*, pp.258–261 (2000).
 - 19) 天野敏之, 玉木 徹: Estimation-by-Completion: 3 次元物体の線形姿勢推定手法, *MIRU2007*, pp. 460–465 (2006). online.
 - 20) 天野敏之, 玉木 徹: EbC 画像対を用いた 3 次元物体の高速線形姿勢推定, *信学論*, Vol.J90-D, p.in print (2007).
 - 21) 玉木 徹, 天野敏之: マルチポート固有空間法, 部分空間法研究会 2006, pp.7–15 (2006). online.
 - 22) 安藤慎吾, 草地良規, 鈴木 章, 荒川賢一: サポートベクトル回帰を用いた三次元物体の姿勢推定法, *信学論*, Vol.J89-D, No.8, pp.1840–1847 (2006). online.
 - 23) Ando, S., Kusachi, Y., Suzuki, A. and Arakawa, K.: Appearance based pose estimation of 3D object using support vector regression, *ICIP2005*, Vol. 1, pp. I-341–344 (2005). online.
 - 24) 安藤慎吾, 鈴木 章, 安野貴之: 部分遮蔽と背景変化にロバストな SVR ベースの 3 次元物体姿勢推定, *PRMU*, Vol.PRMU2007-40, pp.75–80 (2007). online.
 - 25) 玉木 徹, 天野敏之, 金田和文: 巡回群行列を用いた回転物体画像の複素部分空間への投影とその応用, *MIRU2007*, pp.988–993 (2007). online.
 - 26) 玉木 徹, 天野敏之, 金田和文: 巡回群行列のブロック対角化を用いた回転物体画像の部分空間への投影とその応用, *MIRU2007*, pp.982–987 (2007). online.
 - 27) 玉木 徹, 天野敏之, 金田和文: 画像列の遷移行列のブロック対角化による部分空間を用いた物体の姿勢推定, *MIRU2007*, pp.958–963 (2007). online.
 - 28) Glassner, A. S.(ed.): *Graphics Gems*, Academic Press (1990).
 - 29) Menke, W.: 離散インバース理論 — 逆問題とデータ解析, 古今書院 (1997). 柳谷 俊, 塚田 和彦 訳.