

# $R^n$ は $R$ より高精度に $R$ を推定できるか？

— 1 回の計測で同時に得られる  $R, R^2, \dots, R^n$  から

回転行列  $R$  を推定する手法の提案 —

玉木 徹<sup>†</sup> Bisser Raytchev<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 〒630-0192 奈良県生駒市高山町 8916 番地の 5

E-mail: †{tamaki,bisser,kin}@hiroshima-u.ac.jp, ††amano@is.naist.jp

あらまし 本論文では、 $3 \times 3$  回転行列  $R$  をより高精度に推定するために、高次の回転行列  $R^2, R^3, \dots, R^n$  を用いる手法を提案する。まず、電磁波測距に基づいて  $2 \times 2$  回転行列  $R$  の推定を  $R^2, R^3, \dots$  などを用いて行う手法について述べる。そして、その手法を  $3 \times 3$  回転行列の角度推定のために、次のように定式化する。つまり、もしノイズを含む観測行列  $R, R^2, \dots, R^n$  が与えられた場合、それらから適切に  $R$  を推定する。提案手法では、まず与えられた観測行列を直交化により回転行列に変換する。次に固有値分解により回転軸と回転量を求める。最後に回転量の不定性を除去する。数値実験と、物体の姿勢推定実験により、 $R^n$  を用いることで  $R$  単独よりも高精度に回転行列を推定できることを実証する。

キーワード 回転行列, 高周波数を利用した計測, 電磁波測距, 見えに基づく姿勢推定

## Can $R^n$ estimate a rotation matrix $R$ more accurately than $R$ ?

— A method for estimating a rotation matrix  $R$

by using  $R, R^2, R^3, \dots$  obtained by an one-shot measurement —

Toru TAMAKI<sup>†</sup>, Bisser RAYTCHEV<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: †{tamaki,bisser,kin}@hiroshima-u.ac.jp, ††amano@is.naist.jp

**Abstract** In this paper, we show that a more accurate estimation of a  $3 \times 3$  rotation matrix  $R$  can be achieved by appropriately decomposing higher-order rotation matrices:  $R^2, R^3$ , and so on. First we discuss an angle estimation of a  $2 \times 2$  rotation matrix inspired by the Electronic Distance Measurement. Then we reformulate the problem for a  $3 \times 3$  rotation matrix: if noise-contaminated measurement matrices  $R, R^2, \dots, R^n$  are given, find an appropriate rotation matrix  $R$ . In the proposed method, the given measurement matrices are first transformed to rotation matrices by using the polar decomposition. Then the rotation angles are obtained by using an eigen decomposition of the rotation matrices. Finally, the ambiguity of the obtained rotation angle is removed. Numerical simulations and pose estimation experiments show that the use of  $R^n$  results in more accurate estimates than when  $R$  itself is used.

**Key words** Rotation matrix, Measurement with higher frequency, Electronic distance measurement, View-based pose estimation

### 1. はじめに

回転の推定はコンピュータビジョンにおいて重要な問

題の一つである。コンピュータビジョンの主要な問題の多くが、回転行列  $R$  と並進ベクトル  $t$  で表わされる 3 次

元姿勢の推定を必要としている。たとえば、エピポラ幾何 [1], [2] で推定される  $F$  行列や  $E$  行列は  $R$  と  $t$  をその中に含んでいる。バンドル調整や因子分解法 [3] などの多視点幾何 [1], [2] は、最適なカメラ姿勢を推定する。ICP 法 [4] などのレジストレーション問題は二つの点群データの間の変換を求め、カメラ校正法はカメラ座標系と世界座標系との間の変換を求めている。姿勢推定手法 (モデルベース手法 [5]、見えに基づく手法 [6] ~ [8]、局所特徴量を用いる手法 [9], [10]) は対象の姿勢を推定する。このように、回転の推定は多くの重要な問題にかかわるものであり、本論文では回転行列  $R$  を推定する手法について議論する。

上記のすべての手法は、回転行列  $R$  そのものを推定している。言い換えれば、その回転を特徴づける (フーリエ解析の意味で) 最も低周波な成分を用いている、と言える。近年の  $SO(3)$  上のフーリエ解析 [11] ~ [14] では、 $R$  の高周波表現も研究されている。そこで、もし  $R$  のものに加えて  $R^2, R^3, \dots$  のような高次の (つまり高周波の) 情報が得られれば、回転の推定がさらに高精度に行えることが期待できる。

本論文における高周波成分の利用は、電磁波測距 (Electronic Distance Measurement, EDM) [15], [16] にもヒントを得ている。電磁波測距は周波数の異なるいくつかの電磁波を用いており (詳細は 2 節で述べる)、単一周波数の電磁波を一つ用いる場合よりも、高い精度で距離を測定することができる。後述するように、電磁波測距は  $2 \times 2$  回転行列の角度推定とみなすこともできる。そこで、高周波成分や高次の回転行列を利用すれば、 $2 \times 2$  や  $3 \times 3$  回転行列の推定の精度を向上させられると予想できる。

本論文では、 $R^2, R^3, \dots$  などを用いて  $R$  を推定する手法を提案する。本手法は、観測行列  $R, R^2, \dots, R^n$  が与えられ、それらはノイズを含むものとする。これらの観測行列から、適切に回転行列  $R$  を決定する。このような問題設定は、基礎行列推定のための 8 点アルゴリズム [2] に似ている。8 点アルゴリズムでは、連立方程式の解として行列が与えられ、基礎行列であるという制約条件付きの最適化問題による定式化により、最適な基礎行列が決定される。提案手法の場合、観測行列が与えられ (これはタスクに依存する) 適切な回転行列が決定される。

ここで重要なのは、本研究は 1 回しかない計測機会において、ある回転行列の複数の観測  $R, R^2, \dots, R^n$  が得られる場合に、非常に有効である点である。通常のコンピュータビジョンの問題 [17] でよくあるように、ノイズを含む観測データを複数回観測すれば、統計的手法によって推定精度を向上させることができる。しかし、ノイズを含む観測が 1 回しかない場合には、このような手法は適用することができない。本研究が対象としているのはこのような場合、つまり、回転行列  $R$  の観測データを繰り返し何度も計測することはできないが、一度の計

測において  $R$  の異なる観測データを同時に計測することはできる場合である。

本手法を適用することができる応用として、見えに基づく姿勢推定手法 [6] ~ [8] がある。コンピュータビジョンの幾何学に基づく通常の推定手法には、今のところ上記のような本手法が適用できる応用場面はない。一方、見えに基づく物体の姿勢推定手法は、学習セットのパラメータを学習し、推定に用いるというものである。パラメータを変更すれば、様々なものを学習できるため、 $R$  だけでなく  $R^2, R^3, \dots$  も同時に学習することができる。さらに、それぞれの姿勢パラメータの推定を別々に行うことができるため、一回の計測において  $R, R^2, R^3, \dots$  の推定を同時に行うことができる。そのため、本手法は特に学習ベースの見えに基づく姿勢推定に重要である。本論文では、実験によって本手法が有効であることを示す。

本論文の構成は次のとおりである。まず 2 節において、電磁波測距にヒントを得た  $2 \times 2$  回転行列の角度推定について述べる。そして 3 節では、その手法を  $3 \times 3$  回転行列の角度推定に定式化しなおす。まず与えられた観測行列を直交化により回転行列に変換する。次に固有値分解により回転軸と回転量を求める。最後に回転量の不定性を除去する。4 節では数値シミュレーション実験と、画像中の物体の 3 自由度姿勢推定を行う実験により、 $R^n$  を用いることで  $R$  単独よりも高精度に回転行列を推定できることを実証する。

## 2. 1 自由度の角度推定

この節では、 $R^2$  を使うことで  $R$  だけを使う場合よりも高精度に推定できることを示す。そのために、まずここでは  $2 \times 2$  回転行列  $R$  の 1 自由度の角度推定について述べる。

ここで、 $R$  は角度  $\theta$  の回転を表すとすると、 $R$  を  $n$  乗した  $R^n$  は  $n\theta$  の回転を表す。つまり  $R^n$  は  $R$  よりも高周波成分を含んでいるといえる。

高周波成分の利用は角度の推定をより高精度にする。これは電磁波測距として知られている。そのため、まず電磁波測距について説明し、そのあと回転行列  $R$  の推定について述べる。

### 2.1 電磁波測距

電磁波測距は測量の分野 [16], [18] で用いられている距離測定手法の一つであり、波長の異なるいくつかの可視光や電磁波を用いる (図 1 参照)。電磁波測距は、送信した電磁波と受信した電磁波の位相差を計測し、送信側と受信側の距離を測定する。

ここで  $Dist$  を計測する距離、 $\lambda_1$  を 1 番目の計測に用いる電磁波の波長、 $\theta_1$  を計測した位相差とする。 $\lambda_1 > Dist$  を仮定すると、次式が成り立つ。

$$Dist = \frac{\theta_1}{2\pi} \lambda_1, \quad 0 \leq \theta_1 < 2\pi \quad (1)$$

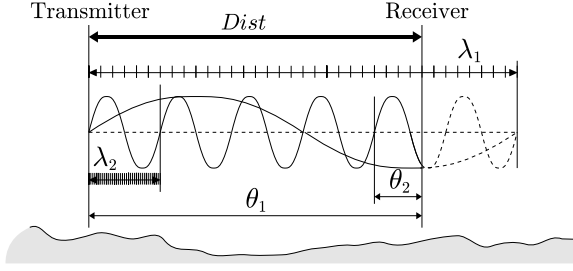


図1 電磁波測距の例。距離  $Dist$  を測るために、波長  $\lambda_1$  と  $\lambda_2$  の電磁波を用いる。

この計測の精度は波長に依存する。 $\theta_1$  を  $2\pi/1000$  の精度で計測した場合、 $Dist$  の計測精度を  $\lambda_1/1000$  にすることができる。

次に、波長  $\lambda_2 (< \lambda_1)$  を用いて2番目に計測された位相差を  $\theta_2$  とすると、次式が成り立つ。

$$Dist = \left( \frac{\theta_2}{2\pi} + k_2 \right) \lambda_2, \quad 0 \leq \theta_2 < 2\pi \quad (2)$$

ここで  $k_2$  は波長  $\lambda_2$  の波の数である。

$\lambda_2$  を  $\lambda_1$  の10分の1に設定すれば、2番目の計測を1番目の計測よりも10倍精度を高くすることができる。これは、2番目の計測精度が  $\lambda_2/1000$  だからである。もちろん  $k_2$  は未知であるが、最初の計測結果から決定することができる。

このような原理を用いた計測機器は電磁波測距儀と呼ばれ、Total Station や Geodimeter、Tellurometer [19] などの商品名で実用化されている。ただし実際の機器では、二つの周波数が少しだけ異なる電磁波のうなり (beating) を用いて、波長が大きく異なる複数の波を作り出している。

## 2.2 2 × 2 回転行列の推定

次に、上述の電磁波測距の議論を  $2 \times 2$  回転行列の推定にあてはめて考えてみる。

$2 \times 2$  回転行列  $R$  を  $R^2$  ( $R$  の2乗) を用いて推定することは、電磁波測距の特殊な場合とみなすことができる。計測対象は、距離ではなく  $R$  の回転角  $\theta$  である。つまり、 $R$  に対応する1番目の計測  $\cos \theta_1, \sin \theta_1$  と、 $R^2$  に対応する2番目の計測  $\cos \theta_2, \sin \theta_2$  を用いて、1自由度の角度  $\theta$  を推定する。 $\theta_1$  は  $\theta$  の (粗い) 推定値であり、 $\theta_2$  は  $2\theta$  の推定値である ( $R^2$  は  $R$  の回転の2倍であるため)。

$\theta_1$  と  $\theta_2$  を用いた  $\theta$  の推定を、以下のように定式化する。 $\theta_1$  のための1番目の波長は、 $0 \leq \theta < 2\pi$  なので、 $\lambda_1 = 2\pi$  とする。 $\theta_2$  のための2番目の波長は、 $\theta_2$  は  $\theta_1$  の2倍なので、 $\lambda_2 = \pi$  とする。すると、次式が得られる。

$$\theta = \frac{\theta_1}{2\pi} \lambda_1 = \theta_1, \quad (3)$$

$$\theta = \left( \frac{\theta_2}{2\pi} + k_2 \right) \lambda_2 = \frac{\theta_2}{2} + \pi k_2, \quad (4)$$

ここで  $k_2$  は0か1である。 $k_2$  の不定性は、 $\theta$  を2倍したことによるものである。もし  $2\theta$  が  $2\pi$  より大きい場合、 $2\theta = \theta + 2\pi$  であるため、 $\theta$  は  $\theta + \pi$  と区別することができない。

この不定性を除去するためには、 $\frac{\theta_2}{2}$  と  $\frac{\theta_2}{2} + \pi k_2$  のどちらが  $\theta_1$  に近いのかを判定する必要がある。ここでは、以下の二つのベクトルの距離を最小化する。

$$\min_{k_2=0,1} \left\| \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix} - \begin{pmatrix} \cos(\frac{\theta_2}{2} + \pi k_2) \\ \sin(\frac{\theta_2}{2} + \pi k_2) \end{pmatrix} \right\|^2. \quad (5)$$

これにより、最初の推定値である  $\theta_1$  を2番目の推定値である  $\frac{\theta_2}{2} + \pi k_2$  に置き換える。

上述の手法は  $R^n$  に一般化することができる。 $R^n$  の回転量を  $\theta_n$  とする。 $R^n$  は  $n$  回の  $\theta$  回転であるため、 $\theta_n$  は  $n\theta$  の推定値である。 $\theta_n$  は波長  $\lambda_n = 2\pi/n$  を用いて推定される。

$$\theta = \left( \frac{\theta_n}{2\pi} + k_n \right) \lambda_n = \frac{\theta_n}{n} + 2\pi \frac{k_n}{n}, \quad (6)$$

ここで  $k_n = 0, 1, \dots, n-1$  は回転の不定性を表す。 $k_n$  は以下の距離を最小化することで決定される。

$$\hat{k}_n = \operatorname{argmin}_{k_n} \left\| \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix} - \begin{pmatrix} \cos(\frac{\theta_n}{n} + 2\pi \frac{k_n}{n}) \\ \sin(\frac{\theta_n}{n} + 2\pi \frac{k_n}{n}) \end{pmatrix} \right\|^2 \quad (7)$$

最終的な角度の推定値は  $\frac{\theta_n}{n} + 2\pi \frac{\hat{k}_n}{n}$  である。

## 3. 3 × 3 回転行列の推定

この節では、 $R^n$  を用いて  $3 \times 3$  回転行列  $R$  の3自由度回転を推定する手法について述べる。これは前節で述べた1自由度回転推定の議論に基づいているが、3自由度回転に特有の問題も扱う必要がある。まず行列の直交化を行い、次に回転軸と回転量を決定し、最後に回転角度の不定性を除去する。

ここでは、 $R^n$  ( $n = 1, 2, \dots$ ) に対応する  $3 \times 3$  の観測行列が与えられるということを仮定する。また、それぞれの観測行列の  $n$  は既知であるとする。これは電磁波測距の場合に波長が既知であったのと同様である。

### 3.1 行列の直交化

まず最初に、行列  $R^n$  を観測した行列を回転行列に変換する。なぜなら、観測行列は通常単なる  $3 \times 3$  行列でしかなく、直交行列でも回転行列でもないからである。

$R_n$  を  $R^n$  の観測行列とする。つまり、 $R$  による回転を  $n$  回繰り返した行列であるべきものに、観測時のノイズが加わったものであるとする。次に  $\tilde{R}_n$  を  $R_n$  に最も近い回転行列とする。 $\tilde{R}_n$  はそれらのフロベニウスノルムを最小化する [20], [21] ことで得られる。

$$\min \|R_n - \tilde{R}_n\|_F + \mu \|\tilde{R}_n^T \tilde{R}_n - I\|_F \quad (8)$$

ここで  $\mu$  はラグランジュ乗数である。

この解は polar decomposition [20] として知られている。\$R\_n\$ の特異値分解を \$R\_n = U\_n \Sigma\_n V\_n^T\$ とすると、\$R\_n\$ の polar decomposition は次式で与えられる。

$$R_n = (U_n V_n^T)(V_n \Sigma_n V_n^T). \quad (9)$$

ここで polar part である \$U\_n V\_n^T\$ が求める解であり、式 (8) を最小化する直交行列である。

しかし、この解は回転行列である保証がないため、さらに修正する必要がある。つまり、\$U\_n V\_n^T\$ は単に直交行列であるというだけであり、行列式が \$-1\$ の場合もあるのである。その場合には、\$U\_n V\_n^T\$ を \$\tilde{R}\_n\$ として用いることはできない。そこで、以下のように解を修正する。

$$\tilde{R}_n = \begin{cases} U_n V_n^T & \text{if } |U_n||V_n| = +1, \\ U_n (H V_n)^T & \text{if } |U_n||V_n| = -1, \end{cases} \quad (10)$$

ここで \$H = \text{diag}(1, 1, -1)\$ とする。

この修正は以下のように解釈できる。polar decomposition は、\$R\_n\$ の特異値分解から \$U\_n V\_n^T\$ を抜き出すために、因子 \$V\_n^T V\_n\$ (これは単位行列 \$I\$ である) を挿入しているとみなせる。しかし、任意の正則行列 \$G\$ についての \$V\_n^T G^{-1} G V\_n\$ を因子として用いても、同様に成り立つ。そこで、polar part の行列式が \$1\$ になるように、行列 \$H\$ についての \$V\_n^T H^{-1} H V\_n\$ を因子として用いる。すると \$H^T H = I\$ と \$|H| = -1\$ から、\$|U\_n||V\_n| = -1\$ の場合には、元の polar decomposition を次のように書き直すことができる。

$$R_n = (U_n (H V_n)^T)(H V_n \Sigma_n V_n^T) \quad (11)$$

### 3.2 固有値分解

次に、\$\tilde{R}\_n\$ の回転量を決定する。回転行列の回転量と回転軸は、その回転行列を固有値分解することで得る。

回転行列 \$R\_n\$ の回転角度は、exponential map [2] により \$\cos \theta\_n = \frac{\text{tr} R\_n - 1}{2}\$ で求めることができる。しかし、本論文における定式化においては、この式では角度を一意に決定することができない。exponential map は回転角は \$0\$ から \$\pi\$ までと仮定しているが、本論文においては \$\theta\_n\$ の範囲は \$0\$ から \$2\pi\$ である。したがって、角度の \$\cos\$ だけでなく \$\sin\$ も知る必要がある。

回転角度と回転軸を求めるには、回転行列の固有値分解が必要である。ここで \$\omega\_n = (b, c, d)^T\$ を回転軸を表す 3 次元ベクトルとすると、exponential map [2] により、次式で与えられる。

$$\omega_n = (r_{32} - r_{23}, r_{13} - r_{31}, r_{21} - r_{12})^T, \quad (12)$$

ここで \$r\_{ij}\$ は \$\tilde{R}\_n\$ の \$ij\$ 要素である。

この \$\tilde{R}\_n\$ の回転軸は、\$\tilde{R}\_n\$ の固有値 \$1\$ に対応する固有ベクトルである [1], [22]。他の二つの固有値は回転角度に対応する。そこで、\$\tilde{R}\_n\$ の固有値分解は、ユニタリ行列 \$P\_n\$ と対角行列 \$D\_n\$ を用いて以下のように表わされ

る [1], [22], [23]。

$$\tilde{R}_n = P_n D_n P_n^H, \quad (13)$$

ここで、\$i = \sqrt{-1}\$ として

$$P_n = \frac{1}{\|\omega_n\|} \begin{pmatrix} b & \frac{\sqrt{c^2+d^2}}{2} & \frac{\sqrt{c^2+d^2}}{2} \\ c & \frac{-bc+id\|\omega_n\|}{\sqrt{2(c^2+d^2)}} & \frac{-bc-id\|\omega_n\|}{\sqrt{2(c^2+d^2)}} \\ d & \frac{-bd-ic\|\omega_n\|}{\sqrt{2(c^2+d^2)}} & \frac{-bd+ic\|\omega_n\|}{\sqrt{2(c^2+d^2)}} \end{pmatrix} \quad (14)$$

$$D_n = \text{diag}(1, e^{i\theta_n}, e^{-i\theta_n}) \quad (15)$$

\$P\_n\$ の列が固有ベクトルであり、回転軸は第 1 列に相当する。\$D\_n\$ の対角成分が固有値であり、最後の二つの固有値 \$e^{\pm i\theta\_n}\$ が回転角度に対応する。したがって、\$P\_n^H \tilde{R}\_n P\_n = D\_n\$ の対角成分を取り出せば \$\theta\_n\$ を求めることができる。

\$P\_n\$ はすべての \$R^n\$ に共通である。なぜなら回転角度だけが倍になったとしても回転軸は不変だからである。これは数式でも表わされる。\$R = P D P^H\$ を \$R\$ の固有値分解とすると、\$P^H P = I\$ なので、\$R^n = P D P^H \dots P D P^H = P D^n P^H\$ となる。

しかし、実際には \$P\_n\$ は \$\tilde{R}\_n\$ のすべてで同じではない。なぜなら観測行列にはノイズが含まれており、回転軸が同一である保証はないからである。\$P\_n\$ の最適な選択方法は今後の課題である。ここでは、\$\omega\_1, \dots, \omega\_n\$ の median をとり<sup>(注1)</sup>、ノルム \$1\$ に正規化した軸 \$\omega\_{\text{med}}\$ を用いて計算した \$P\_{\text{med}}\$ を用いることにする。ここで、

$$\omega_{\text{med}} = \underset{i=1, \dots, n}{\text{med}} \omega_i \quad (16)$$

は、各要素の median をベクトルの median と定義している。

### 3.3 ブロック対角化

上述の固有値分解を、さらに実行列を用いたブロック対角化に変換する [24], [25]。固有値分解は複素行列を用いており、複素数を扱える (Matlab などの) ソフトウェアがあるとしても、実用上は実行列を用いる方が好まれる場合が多い。

ここで、次のようなユニタリ行列 \$S\$ を定義する。

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{pmatrix}, \quad S^H S = S S^H = I. \quad (17)$$

すると固有値分解を次のように書き換えることができる。

$$\tilde{R}_n = P_n D_n P_n^H = P_n S S^H D_n S S^H P_n^H, \quad (18)$$

$$= P'_n D'_n P_n'^T, \quad (19)$$

ここで \$P'\_n = P\_n S\$、\$D'\_n = S^H D\_n S\$ である。\$P'\_n\$ は実直交

(注1): 平均よりも median の方が精度がよかったため。

行列であり、 $P'_n P'_n{}^T = P'_n{}^T P'_n = I$  が成り立つ。さらに、 $D'_n$  は以下のような実ブロック行列である。

$$D'_n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_n & \sin \theta_n \\ 0 & -\sin \theta_n & \cos \theta_n \end{pmatrix}. \quad (20)$$

$\theta_n$  は  $D'_n = P'_n{}^T \tilde{R}_n P'_n$  の二つの要素から求めることができる。この計算は実行列のみで行うことができ、実装が容易である。

### 3.4 角度の不定性の除去

最後に、回転角度の不定性を除去する。これは  $2 \times 2$  場合と同様である。

$R_1$  と  $R_n$  からそれぞれ角度の推定値  $\theta_1$  と  $\theta_n$  が得られたとする。 $\theta_n$  の不定性は以下の距離を最小にすることで決定される。

$$\hat{k}_n = \operatorname{argmin}_{\substack{k_n=0,1, \\ \dots, n-1}} \left\| \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix} - \begin{pmatrix} \cos(\frac{\theta_n}{n} + 2\pi \frac{k_n}{n}) \\ \sin(\frac{\theta_n}{n} + 2\pi \frac{k_n}{n}) \end{pmatrix} \right\|^2 \quad (21)$$

最終的な角度の推定値は  $\hat{\theta} = \frac{\theta_n}{n} + 2\pi \frac{\hat{k}_n}{n}$  である。

$R_1$  から  $R_n$  までの  $n$  個の観測行列が得られる場合には、上式において  $\theta_1$  ではなく  $\theta_{n-1}$  を用いることにする。なぜなら、 $\theta_1$  の誤差が  $\frac{2\pi}{n}$  より大きい場合には (ノイズが大きい時に生じる)  $k_n$  が正しく求まらないからである。 $\theta_1$  から順に角度の推定値を更新すると、 $\theta_{n-1}$  の推定誤差は  $\frac{2\pi}{n-1}$  以下であるため、 $k_n$  は正しく求めることができる。

最終的な回転行列の推定値は、以下で与えられる。

$$\hat{R} = P'_{\text{med}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \hat{\theta} & \sin \hat{\theta} \\ 0 & -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} P'_{\text{med}}{}^T \quad (22)$$

## 4. 評価実験

提案手法の有効性を示すために、二つの数値シミュレーション実験の結果を示す。一つ目の数値実験は、ある回転行列を真値として、それぞれの観測行列  $R_n$  ( $n = 1, 2, \dots$ ) には独立に同じタイプのノイズを加える。二つ目の数値実験は、 $R_1$  にノイズを加え、それを  $n$  乗した  $(R_1)^n$  を観測行列  $R_n$  とする。

### 4.1 数値実験 1：独立にノイズを加えた場合

一つ目の数値シミュレーション実験では、以下の手順で観測行列を生成した。

- 真の回転行列  $R$  を生成する。
- $R$  を  $n$  乗して  $R^n$  を計算する。
- $R^n$  に一樣乱数を加えて観測行列  $R_n$  を生成する。

真の回転行列  $R$  の生成には球面乱数 [26] を用いた。球面乱数により、ノルム 1 の 4 次元ベクトルを生成する。それを単位四元数とみなし、回転行列に変換する。実験

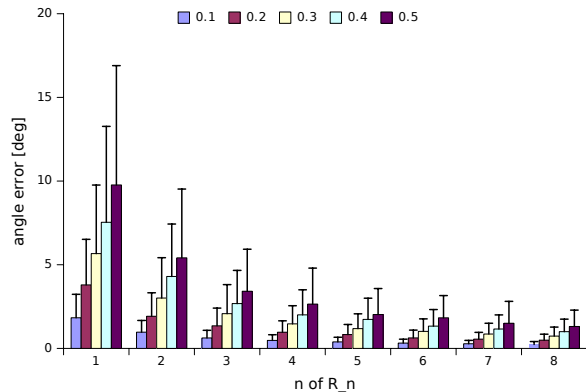


図 2 数値実験 1。一樣乱数  $\pm 0.1 \sim \pm 0.5$  に対する角度推定値の平均絶対誤差。

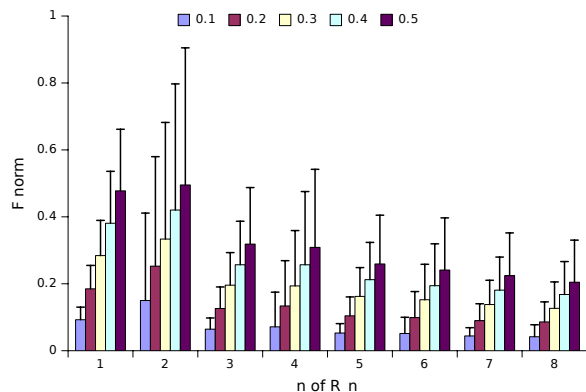


図 3 数値実験 1。一樣乱数  $\pm 0.1 \sim \pm 0.5$  に対する回転行列の推定値と真値のフロベニウスノルムで測った誤差。

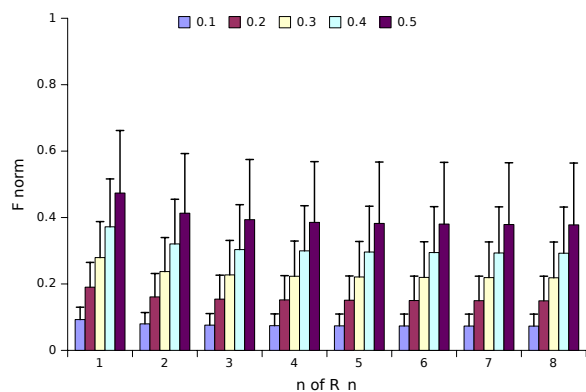


図 4 数値実験 1。一樣乱数  $\pm 0.1 \sim \pm 0.5$  に対する回転行列の推定値と真値のフロベニウスノルムで測った誤差。ただし  $P_{\text{med}}$  の代わりに  $P_1$  を用いた。

には、1000 個の回転行列を用意した。

観測行列を生成するために、要素が一樣乱数からなる行列を毎回生成し、真の回転行列に加えた。つまり、 $R_1$  と  $R_2$  に加える乱数行列は、同じ幅を持つ一樣乱数から生成されているが、別のものである。一樣乱数の幅は  $\pm 0.1$  から  $\pm 0.5$  とした。

評価には、回転角度の誤差 (真値と推定値との絶対誤差) と、回転行列としての誤差 (真値と推定値との差のフロベニウスノルム) を用いる。ここで、角度誤差の計算には回転軸の方向を考慮している。回転角度と回転軸

の真値と推定値をそれぞれ  $\theta, \omega$  と  $\hat{\theta}, \hat{\omega}$  とすると、3次元回転の性質から、 $\theta, \omega$  が表す回転は  $(2\pi - \theta), -\omega$  が表す回転と等しい。そのため次式のように、 $\omega$  と  $\hat{\omega}$  のなす角が90度以下であればそのまま角度誤差を計算し、そうでなければ推定角度を  $2\pi$  から引いて角度誤差を求めた。

$$\text{angle error} = \begin{cases} |\theta - \hat{\theta}|, & \omega^T \hat{\omega} < \frac{\pi}{2} \\ |\theta - (2\pi - \hat{\theta})|, & \text{otherwise} \end{cases} \quad (23)$$

図2に回転角度の推定結果を示す。縦軸は角度の推定誤差 [deg] を表し、横軸は推定に用いた観測行列  $R_n$  の  $n$  を表す。 $n = 1$  は  $R_1$  のみを用いた場合、 $n > 1$  は  $R_1$  と  $R_2, \dots, R_n$  を推定に用いた場合の結果である。それぞれの角度誤差は1000個の回転行列に対する推定値の平均であり、誤差棒は標準偏差を意味する。加えた一様乱数の幅 ( $\pm 0.1, 0.2, 0.3, 0.4, 0.5$ ) ごとに推定結果を示してある。

この結果から、回転角度の推定には提案手法が有効であることがわかる。加えた乱数の幅によらず、 $R_2, R_3, \dots$  を用いた方が  $R_1$  だけを用いた場合よりも角度の推定誤差が小さくなっている。 $R_2$  の推定誤差は  $R_1$  の推定誤差の7割程度にまで抑えられている。この傾向は加えた乱数の幅が大きくても顕著である。そのため、提案手法はノイズが大きい場合にも有効であると言える。

図3に、推定された回転行列と真の回転行列との誤差をフロベニウスノルムで表わしたものを示す。回転行列には角度の誤差に加えて回転軸の誤差が加わるため、この評価は前述の回転角度の評価とは異なる。 $R_2$  を用いた結果は  $R_1$  のみの場合よりも悪くなっているが、 $R_3$  まで用いると推定誤差が減ることが分かる。 $R_2$  の結果が悪い理由は、対角化行列の計算に回転軸の median を取った  $P_{\text{med}}$  を用いているためである。2つのデータの median は意味のあるものではない。 $R_2$  の結果を向上させるために、 $P_{\text{med}}$  の代わりに  $P_1$  を用いた結果を図4に示す。こうすることで、 $R_2$  の結果は  $R_1$  よりも向上する。しかし  $R_3$  以上に対しては精度に大きな変化は見られない。このように対角化行列  $P$  の取り方によって結果は変わるため、最適な  $P$  の決定方法が今後の課題である。

以上の結果は提案手法の有効性を十分に示している。この実験ではそれぞれの観測行列に別々のノイズが同程度に加わると仮定している。もし応用場面においてそのような仮定を満たせば、提案手法が非常に有効であるといえる。

#### 4.2 数値実験2：加えたノイズが独立でない場合

二つ目の数値シミュレーション実験では、以下の手順で観測行列を生成した。

- 真の回転行列  $R$  を生成する。
- $R$  に一様乱数を加えて  $R_1$  を生成する。
- $R_1$  を  $n$  乗した  $(R_1)^n$  を観測行列  $R_n$  とする。

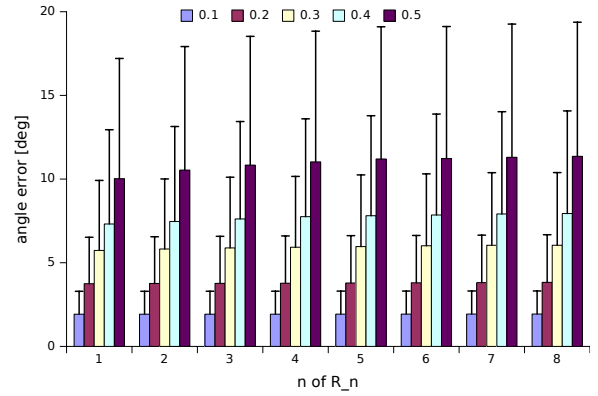


図5 数値実験2。一様乱数  $\pm 0.1 \sim \pm 0.5$  に対する角度推定値の平均絶対誤差。

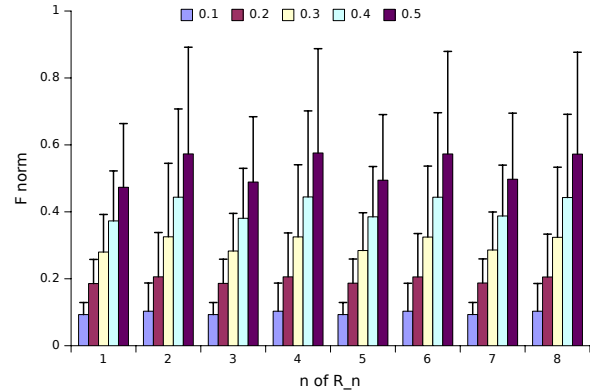


図6 数値実験2。一様乱数  $\pm 0.1 \sim \pm 0.5$  に対する回転行列の推定値と真値のフロベニウスノルムで測った誤差。

通常の応用場面では  $R_1$  しか得られないため、この実験の仮定は、一つ目の実験の仮定よりも現実的であろうと思われる。

図5に角度推定結果を、図6にフロベニウスノルムで評価した結果を示す。これらの図のフォーマットは一つ目の実験結果と同じである。

この図から、この実験においては  $R_2$  を用いても  $R_1$  より精度が向上していないことが分かる。これは、 $R_2$  が  $R_1$  を2乗して生成されているため、 $R$  に関して追加的な情報を持っていないためである。提案手法は  $R_1, R_2, \dots$  が別々に得られることを仮定しているため、ノイズを含む  $R_1$  しか得られない場合には、それ以上精度を向上させることはできないことがわかる。

#### 4.3 見えに基づく姿勢推定への応用

ここでは、本手法を見えに基づく3自由度の姿勢推定へ適用した実験結果を示す。画像全体の見えに基づく手法[6]~[8]は、モデルに基づく手法[5]や局所特徴量を用いる手法[9],[10]とは異なり、画像とパラメータとの関係を学習するものである。

3自由度の姿勢パラメータとしては、固定角やオイラー角、四元数などがあるが、ここでは回転行列  $R$  そのものを姿勢パラメータとする。そしてこの実験では、 $R, R^2, \dots, R^8$  を姿勢パラメータとして推定し、その推定



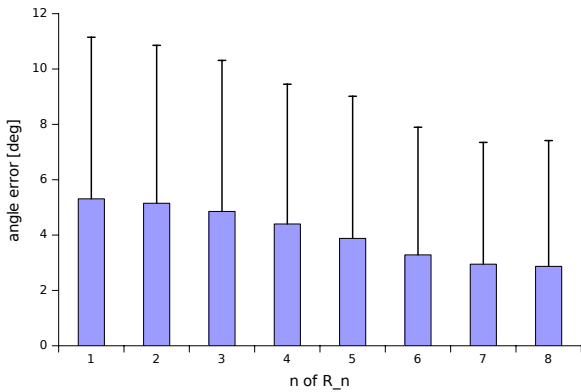


図7 見えに基づく姿勢推定に適用した結果。角度推定値の平均絶対誤差。

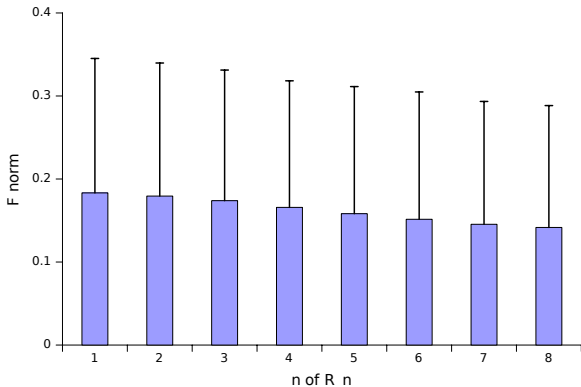


図8 見えに基づく姿勢推定に適用した結果。回転行列の推定値と真値のフロベニウスノルムで測った誤差。

値  $R_1, \dots, R_8$  を本手法の推定に用いる。

推定手法には、正準相関分析 [8] や非線形回帰 [27] などがあるが、ここではもっとも単純な線形回帰 [28] を用いる。つまり、姿勢を表す回転行列  $R$  の学習画像ベクトルを  $x_R$  として、

$$R^n = F_n x_R, \quad n = 1, \dots, 8 \quad (24)$$

を満たすように行列  $F_i$  を求めておく。ここで  $R^n$  は、 $R^n$  の9個の要素を並べたベクトルである。

未知の画像  $x$  を推定するときには、

$$R_n = F_n x, \quad n = 1, \dots, 8 \quad (25)$$

として得られた9次元ベクトル  $R_n$  の要素を、 $3 \times 3$  の行列  $R_n$  に並べ直し、本手法での観測行列とする。

実験には、3次元物体モデルを用いて作成し、学習用に100枚、推定用に1000枚のCG画像 ( $128 \times 128$ 、グレイスケール) を使用した。姿勢は、学習用と推定用ともに、正面から30度以内にランダムになるようにした。まず球面乱数により生成したノルム1の3次元ベクトルを回転軸とし、 $0 \sim 30[\text{deg}]$  の範囲で一様乱数により生成した値を回転角度とした。

図7に角度推定結果を、図8にフロベニウスノルムで評価した結果を示す。 $R_1$ のみを用いた場合では、角度誤差は  $5.31 \pm 5.84[\text{deg}]$  であったが、 $R_8$ までを用い

た場合では  $2.87 \pm 4.55[\text{deg}]$  にまで減少している。またフロベニウスノルムでの誤差も、 $R_1$ のみを用いた場合は  $0.183 \pm 0.162$  であったが、 $R_8$ までを用いた場合では  $0.142 \pm 0.146$  にまで減少している。 $n$ が大きくなるにつれて推定誤差が減少しており、この結果は本手法の有効性を示している。

この実験では線形回帰による推定結果を用いたが、本手法は線形回帰以外の手法にも適用できる。非線形手法を用いた姿勢手法を、この実験と同様に本手法に適用すれば、姿勢推定精度をさらに向上させることができると思われる。

## 5. おわりに

本論文では、回転行列  $R$  をそのべき乗の行列  $R^n$  ( $n > 1$ ) を用いて推定する手法を提案した。一様乱数を加えた回転行列を用いて数値シミュレーション実験を行い、その推定誤差を角度とフロベニウスノルムを用いて評価した。実験結果から、もし観測行列がそれぞれ独立にノイズを加えられた場合には、推定誤差は減少し、提案手法は非常に有効であることが示された。しかしノイズが独立でない場合には、精度の向上は見られなかった。見えに基づく姿勢推定に適用した結果、推定誤差が減少することが確認され、本手法が有効であることが示された。 $R$ がノイズを含むデータから推定される応用問題において、観測行列が持つノイズの独立性という仮定が満たされるようにするにはどのように問題を定式化すればよいかということについては、個々の応用問題において今後示す必要がある。

## 文 献

- [1] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004. <http://www.cambridge.org/aus/catalogue/catalogue.asp?isbn=0521540518>.
- [2] Yi Ma, Stefano Soatto, Jana Košecká, and S. Shankar Sastry. *An Invitation To 3-D Vision*. Springer, 2004. <http://vision.ucla.edu/MASKS/>.
- [3] Carlo Tomasi and Takeo Kanade. Shape and motion from image streams under orthography: a factorization method. *Intl. J. of Computer Vision*, Vol. 9, No. 2, pp. 137–154, 1992. <http://www.springerlink.com/content/q3546r136332418r>.
- [4] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Trans. PAMI*, Vol. 14, No. 2, pp. 239–256, 1992. <http://doi.ieeecomputersociety.org/10.1109/34.121791>.
- [5] D. G. Lowe. Fitting parameterized three-dimensional models to images. *IEEE Trans. PAMI*, Vol. 13, No. 5, pp. 441–450, 1991. <http://doi.ieeecomputersociety.org/10.1109/34.134043>.
- [6] 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>.
- [7] 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\\_vba\\_rov.html](http://ls7-www.cs.uni-dortmund.de/~peters/pages/research/modeladaptsys/modeladaptsys_vba_rov.html).
- [8] Thomas Melzer, Michael Reiter, and Horst Bischof. Appearance models based on kernel canonical correlation analysis. *Pattern Recognition*, Vol. 36, pp. 1961–1971, 2003. [http://dx.doi.org/10.1016/S0031-3203\(03\)00058-X](http://dx.doi.org/10.1016/S0031-3203(03)00058-X).
- [9] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Intl. J. of Computer Vision*, Vol. 60, No. 2, pp. 91–110, 2004. <http://www.springerlink.com/content/h4102691327px768>.
- [10] Fred Rothganger, Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. 3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints. *Proc. of CVPR2003*, Vol. 2, pp. 272–277, 2003. [http://www-cvr.ai.uiuc.edu/ponce\\_grp/publication/paper/cvpr03a.ps.gz](http://www-cvr.ai.uiuc.edu/ponce_grp/publication/paper/cvpr03a.ps.gz).
- [11] Ameesh Makadia and Kostas Daniilidis. Rotation recovery from spherical images without correspondences. *IEEE Trans. PAMI*, Vol. 28, No. 7, pp. 1170–1175, 2006.
- [12] Ameesh Makadia and Kostas Daniilidis. Direct 3d-rotation estimation from spherical images via a generalized shift theorem. *Proc. of CVPR2003*, Vol. 2, p. 217, 2003. <http://doi.ieeecomputersociety.org/10.1109/CVPR.2003.1211473>.
- [13] Peter J. Kostelec and Daniel N. Rockmore. Ffts on the rotation group. *Journal of Fourier Analysis and Applications*, Vol. 14, No. 2, 2008. <http://www.springerlink.com/content/r6312816xr890vr5>.
- [14] David K. Maslen and Daniel M. Rockmore. Generalized FFTs – a survey of some recent results. *Groups and Computation II, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pp. 183–238, 1997. <http://www.cs.dartmouth.edu/~rockmore/dimacs-0.pdf>.
- [15] 須田教明. 電磁波測距儀. 森北出版, 1976.
- [16] Russell Charles Brinker and Roy Minnick, editors. *The Surveying Handbook*. Chapman & Hall, 2nd edition, 1995. <http://books.google.com/books?id=2gB7w9X1NJAC>.
- [17] Zhengyou Zhang. Parameter estimation techniques: a tutorial with application to conic fitting. *Image and Vision Computing*, Vol. 15, No. 1, pp. 59–76, 1997.
- [18] 岡田清, 森忠次, 岡本厚, 小泉俊雄, 細野武庸. 測量学. 東京電機大学出版局, 2008. <http://books.google.co.jp/books?id=jk7IVyjs2LoC>.
- [19] The Alberta Land Surveyors' Association. Equipment. online, accessed 2009/2/26. <http://www.landsurveyinghistory.ab.ca/Equipment.htm>.
- [20] Gene Howard Golub and Charles F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 3rd, 1996. <http://books.google.com/books?id=m10a7wPX60YC>.
- [21] 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).
- [22] MathPages. Rotation matrices. online, accessed 2009/2/26. <http://www.mathpages.com/home/kmath593/kmath593.htm>.
- [23] Eric W. Weisstein. Rotation matrix. From MathWorld – A Wolfram Web Resource, accessed 2009/2/26. <http://mathworld.wolfram.com/RotationMatrix.html>.
- [24] Ichiro Satake. *Linear Algebra*. Marcel Dekker Inc., 1975.
- [25] 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. <http://ir.lib.hiroshima-u.ac.jp/00019809>.
- [26] GSL – GNU Scientific Library, 2008. <http://www.gnu.org/software/gsl/>.
- [27] Shingo Ando, Yoshinori Kusachi, Akira Suzuki, and Kenichi Arakawa. Appearance based pose estimation of 3D object using support vector regression. *ICIP2005*, Vol. 1, pp. 1–341–344, 2005. [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1529757&isnumber=32660](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1529757&isnumber=32660).
- [28] Takayuki Okatani and Koichiro Deguchi. Yet another appearance-based method for pose estimation based on a linear model. *IAPR Workshop on Machine Vision Applications 2000*, pp. 258–261, 2000. <http://b2.cvl.iis.u-tokyo.ac.jp/mva/proceedings/CommemorativeDVD/2000/papers/2000258.pdf>.