

線形的な 3 自由度姿勢推定のための姿勢表現の性質の実験的評価

原田 健吾[†] 玉木 徹[†] Bisser Raytchev[†] 金田 和文[†] 天野 敏之^{††}

[†] 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

^{††} 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: †{harada,tamaki,b.raytchev,kin}@eml.hiroshima-u.ac.jp, ††amano@is.naist.jp

あらまし 本論文では、3 自由度における姿勢表現の性質を実験的に評価した結果を報告する。姿勢推定に線形回帰を用いる手法を前提とする。まず、線形的な見えに基づく姿勢推定手法の仕組みを説明し、そのために必要な姿勢表現の性質を説明する。次に、4 つの姿勢表現 (回転行列、ZYX-オイラー角、指数マップ、単位四元数) が性質を満たしているかどうかを考察し、回転行列のみ性質を満たすことを説明する。そして実際に、4 つの姿勢表現をそれぞれ用いて姿勢推定を行い、推定誤差を比較する。その結果から、回転行列を用いた場合が最も推定誤差が小さいことを実験的に示す。

キーワード 姿勢推定、見えに基づく手法、姿勢表現、回転行列、ZYX-オイラー角、指数マップ、単位四元数、線形回帰、3 自由度

Experimental study on Properties of Pose representations for 3DOF linear Pose Estimations

Kengo HARADA[†], Toru TMAKI[†], Bisser RAYTCHEV[†], Kazufumi KANEDA[†], and Toshiyuki AMANO^{††}

[†] Department of Information Engineering, Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-hiroshima, Hiroshima, 739-8527 Japan

^{††} Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: †{harada,tamaki,b.raytchev,kin}@eml.hiroshima-u.ac.jp, ††amano@is.naist.jp

Abstract We report a result of an experimental study on properties of pose representations for 3DOF linear pose estimations. We use linear regression as a pose estimation method. First, we explain a method of linear pose estimation and two properties of pose representations. Next, we use four pose representations (rotation matrix, ZYX Euler angle, exponential map, and quaternion), and consider whether each pose representation satisfy the properties. We show that only rotation matrix satisfies the properties, and verify experimentally that rotation matrix has the smallest error.

Key words pose estimation, view-based, appearance-based, pose representation, rotation matrix, exponential map, ZYX Euler angle, quaternion, linear regression, 3 degrees-of-freedom

1. はじめに

本論文では、3 自由度における姿勢表現の性質を実験的に評価した結果を報告する。

姿勢推定の手法のひとつとして、見えに基づく (view-based, appearance-based) 手法がある。これは、姿勢が既知である画像とその姿勢表現の対応関係を学習し、その学習をもとに姿勢推定を行う。

3 自由度における姿勢推定では、姿勢表現の選択には注意が必要である。玉木ら [6] は、3 自由度姿勢推定において、姿勢表現に必要な性質は一对一と連続性であると、論理的な考察を行っている。そして姿勢表現はこの 2 つの性質を満たさなければならないと述べている。しかし、姿勢表現が姿勢推定にどの程度影響を与えるかについては、調査していない。

そこで本論文では、それぞれの姿勢表現が姿勢推定の

推定値に与える影響を実験的に検証する。姿勢表現には回転行列、ZYX-オイラー角 [3]、指数マップ [1]、単位四元数 [4] を用いる。それぞれの姿勢表現を用いて姿勢推定を行い、推定誤差を比較する。

2. 見えに基づく手法と姿勢表現の性質

2.1 線形的な見えに基づく手法

見えに基づく手法は、画像と姿勢表現の対応関係をあらかじめ学習しておき、その学習から推定を行う。本論文では推定手法として、線形回帰 [9] を用いる。学習画像ベクトルを x_j 、対応する姿勢表現 (姿勢パラメータ) ベクトルを p_j とおく。学習過程では、次式を満たす線形写像 F を求める。

$$p_j = Fx_j \quad (j = 1, 2, \dots, n) \quad (1)$$

この F が画像と姿勢表現の対応関係を表す。推定過程では、姿勢が未知である画像 x に対して、次式より姿勢表現 p を求める。

$$p \cong Fx \quad (2)$$

このような線形的な手法 [8] [9] [10] [11] では、1つの画像を学習画像の線形和で表現することができる。つまり、ある画像 x は学習画像 x_j の線形和

$$x \cong \sum_j b_j x_j \quad (3)$$

で近似することができる。また、学習画像 x_j とその姿勢表現 p_j は式 (1) の関係が成り立つので、式 (3) を用いると次のように書くことができる。

$$p = Fx \cong \sum_j b_j Fx_j = \sum_j b_j p_j \quad (4)$$

つまり、推定される姿勢表現 p も、学習された姿勢表現 p_j の線形和で表現される。

2.2 姿勢表現の性質

上で示した式を用いるために必要な姿勢表現の性質が、一対一と連続性である [6]。

一対一とは、1つの画像に対して1つの姿勢表現が対応するということである。例えば、1つの画像 x に2つの姿勢表現 p_1, p_2 が対応すると仮定する。このとき式 (2) から $p_1 = Fx, p_2 = Fx$ が得られるが、一般的にこのような線形写像 F は存在しない。また式 (4) より、推定姿勢は学習された姿勢表現の線形和であるから、2通りある姿勢表現のうち、どちらか一方を出力するという操作を行うことはできない。故に姿勢表現には一対一の性質が必要である。

連続性とは、画像の見えとともに姿勢パラメータが連続的に変化し、不連続となる特異点を持たないというこ

表 1 使用する姿勢パラメータ

姿勢表現	姿勢パラメータ
回転行列	$[r_{11}, r_{12}, \dots, r_{32}, r_{33}]^T$
ZYX-オイラー角	$[\theta_x, \theta_y, \theta_z]^T, (-\pi \leq \theta_{x,y,z} < \pi)$
指数マップ	$\omega = [\omega_1, \omega_2, \omega_3]^T, (0 \leq \ \omega\ \leq \pi)$
単位四元数	$q = [q_0, q_1, q_2, q_3]^T$

とである。姿勢は学習された姿勢表現の線形和であるから、不連続となる特異点付近の画像を推定するとき、問題が生じる。例えば、1軸周りの回転で回転量の範囲を $[-\pi, \pi]$ [rad] とした場合を考える。 $-\pi$ [rad] から回転し始め π [rad] まで回転すると、姿勢パラメータはまた $-\pi$ [rad] に戻る。これが特異点である。そして、姿勢パラメータが $\pm\pi$ [rad] の画像を推定するとき、誤って 0 [rad] に近い値が推定される可能性が高い。故に姿勢表現には連続性の性質が必要である。

3. 姿勢表現と一対一・連続性

この節では、本論文で使用する4つの姿勢表現について、上で述べた一対一と連続性の観点から考察する。本実験で用いる姿勢表現を表1に示す。4つの姿勢表現のうち、一対一と連続性の性質を満たすのは回転行列のみである。

3.1 回転行列

回転行列は、一対一の性質を満たす。また、連続性の性質を満たす。

3.2 ZYX-オイラー角

ZYX-オイラー角は、一対一の性質を満たさない。その一因として、ジンバルロック [2] の問題がある。X軸周りの回転量を θ_x 、Y軸周りの回転量を θ_y 、Z軸周りの回転量を θ_z とする。ZYX-オイラー角では、 $\theta_y = \pm\pi/2$ [rad] のとき、ジンバルロック状態である。このとき、 $\theta_x + \theta_z$ の値が等しい姿勢表現は同じ回転を表す。例えば、 $(\theta_x, \theta_y, \theta_z) = (\pi/6, \pi/2, \pi/6)$ と、 $(\theta_x, \theta_y, \theta_z) = (0, \pi/2, \pi/3)$ は同じ回転となる。故に、ZYX-オイラー角は一対一の性質を満たさない。

また、ZYX-オイラー角は連続性の性質を満たさない。回転は、1回りするとまた同じ姿勢となる。本実験では、回転量の範囲を $[-\pi, \pi]$ とするので、 π [rad] から $-\pi$ [rad] に変わるとき、不連続となる。故に、ZYX-オイラー角は連続性の性質を満たさない。

3.3 指数マップ

指数マップ表現は、一対一の性質を満たさない。本実験では指数マップの絶対値を回転量とおく。回転量の範囲を $[0, \pi]$ [rad] とすると、ある軸周りに π [rad] の回転と、その軸方向を反転させた π [rad] の回転は同じ回転を表す (図1)。故に、指数マップ表現は一対一の性質を満たさない。

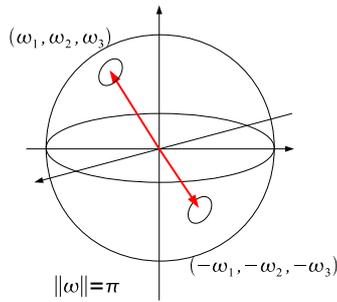


図 1 指数マップの範囲：指数マップは半径 $\|\omega\| = \pi[\text{rad}]$ の球内部に存在する。 $\|\omega\| = \pi[\text{rad}]$ の球面上では、原点对称な点は同じ回転姿勢を表す。

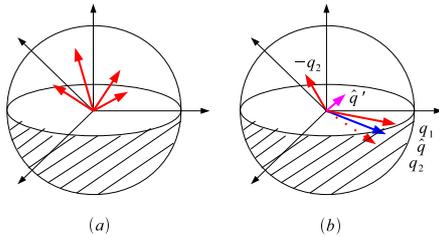


図 2 単位四元数の 1 つの要素の符号を正に揃えた場合：この時、画像と姿勢パラメータは一対一となる (a) が、半球の縁付近が不連続となる (b)。

また指数マップは連続性の性質を満たさない。ある軸周りの回転が、 $0[\text{rad}]$ から $\pi[\text{rad}]$ まで回転して、さらに回転するとき、軸方向を反転させる必要がある。このとき不連続となる。故に、指数マップは連続性の性質を満たさない。

3.4 単位四元数

単位四元数は、一対一の性質を満たさない。ある単位四元数 q とその原点对称な $-q$ は、同じ回転を表すので明らかである。

また、単位四元数は連続性の性質を満たさない。一対一の性質を満たすためには、単位四元数の 1 つの要素の符号を揃えればよい。しかしこの場合、連続性が満たされなくなる。図 2 に示すように、半球の縁付近で不連続となる。故に、単位四元数は連続性の性質を満たさない。

4. 姿勢表現の比較実験

この実験では、姿勢表現に回転行列、ZYX-オイラー角、指数マップ、単位四元数を用い、それぞれ姿勢推定を行う。そして、用いる姿勢表現によってどの程度、推定に影響が出るかを評価する。

まず、実験に使用する画像の作成方法について説明する。次に、評価方法を説明する。そして、実験方法と結果について述べる。

4.1 使用する画像の作成方法

この実験では、3D モデル [12] を用いて作成した CG 画像を使用する。3次元直交座標系の原点に 3D モデル



図 3 ランダムな姿勢の画像例



図 4 不連続となる姿勢の画像例

を置き、回転行列により回転させる。この回転行列が表す姿勢を推定する。

学習する姿勢は 2 種類作成する。1 つ目は、3 自由度で完全にランダムな姿勢、2 つ目は、ZYX-オイラー角と指数マップでは不連続となる姿勢である。

ランダムな姿勢の画像は、単位四元数を利用する。まず、ノルム 1 の 4 次元ベクトルを球面乱数 [13] により作成し、これを単位四元数の要素とする。そして単位四元数から回転行列に変換し、学習画像を作成する。画像例を図 3 に示す。

ZYX-オイラー角と指数マップでは不連続となる姿勢の画像は、次の手順で作成する。まず、 $\theta_z = \pi[\text{rad}]$ の回転を作る。その次に、球面乱数で作成した 3 次元ベクトルを回転軸とし、回転量を $[0, \pi/6][\text{rad}]$ の範囲で一様乱数とした回転を作る。この 2 つの回転を合成した回転行列を用いて画像を作成する。画像例を図 4 に示す。

4.2 評価方法

次の手順で推定値の誤差を評価する。

(1) 線形回帰による推定手法より推定された姿勢パラメータを正規化する。ZYX-オイラー角と指数マップは正規化する必要はない。単位四元数は、ノルムを 1 にする。回転行列は、特異値分解を用いて直交行列に変換する [5]。

(2) 正規化した姿勢表現の推定値を、回転行列に変換する。

(3) 推定値 \tilde{R} と真値 R_{tu} との誤差 R_e を次式で定義する。

$$R_e = R_{tu}\tilde{R}^{-1} \quad (5)$$

(4) 得られた誤差の回転量を次式 [1] で求める。

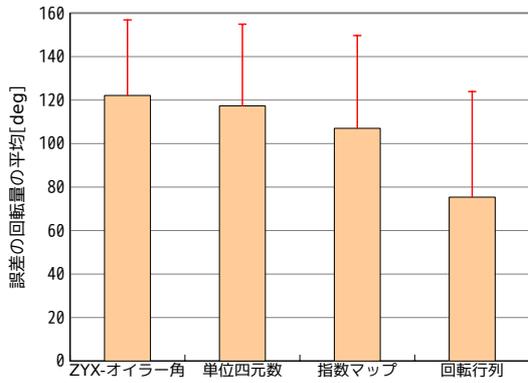


図 5 ランダムな姿勢の画像を用いた比較実験結果

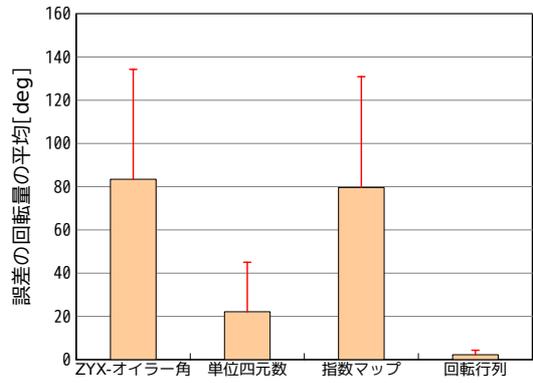


図 7 不連続となる画像を用いた結果

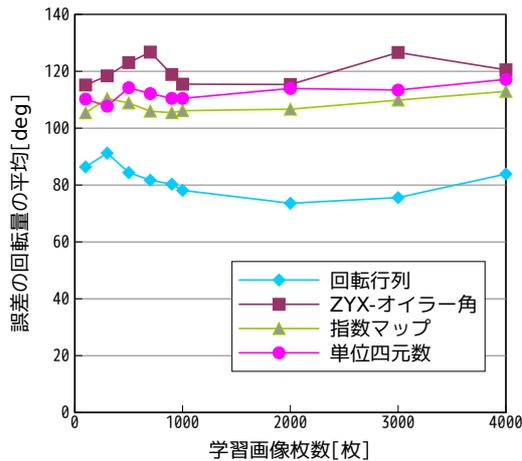


図 6 ランダムな画像を用いた場合の学習画像枚数変化による誤差の変化

$$\theta_e = \cos^{-1} \left(\frac{\text{trace}(R_e) - 1}{2} \right) \quad (6)$$

(5) 複数の推定結果を用いて、誤差の回転量の平均 $E[\theta_e]$ を求める。

4.3 ランダムな姿勢の画像を用いた比較実験

ここでは、ランダムな姿勢の画像を用いた姿勢推定結果を比較する。学習画像枚数は 2500 枚、推定画像枚数は 100 枚とした。

結果を図 5 に示す。グラフの縦軸は誤差の回転量の平均を示す。図 5 から、姿勢表現に必要な性質を唯一満たしている回転行列を用いた場合が、最も誤差が小さいことがわかる。しかし推定手法に線形回帰を用いているので、誤差が大きくなってしまっている。他の手法を用いて誤差を減らすのは今後の課題である。

次に、学習画像枚数を 100 枚から 4000 枚に変化させ、同様に誤差の回転量の平均を求めた。推定画像枚数は 100 枚とした。結果を図 6 に示す。グラフの縦軸は誤差の回転量の平均、横軸は学習画像枚数を示す。図 6 から、学習画像枚数が変化しても、回転行列を用いた場合が最も誤差が小さいことがわかる。



図 8 誤推定の例

4.4 不連続となる画像を用いた比較実験

ここでは、ZYX-オイラー角と指数マップ表現では不連続となる姿勢の画像を用いて、先ほどと同様の実験を行う。

まず、学習画像に 2500 枚、推定画像に 100 枚の画像を用いた結果を図 7 に示す。グラフの縦軸は誤差の回転量の平均を示す。この結果から、先ほどの実験同様に、回転行列を用いた場合が最も誤差が小さいことがわかる。ZYX-オイラー角と指数マップを用いた場合は、不連続となる画像が多くあったため、誤差が大きくなった。

ZYX-オイラー角と指数マップを用いた場合で誤差が大きくなった原因を考える。この実験で使用した画像の姿勢表現は、ZYX-オイラー角では $\theta_z = \pm\pi[\text{rad}]$ 、指数マップでは回転軸方向が原点对称で回転量(絶対値)が $\pi[\text{rad}]$ となるものが多い。故に推定値は、ZYX-オイラー角では $\theta_z = 0[\text{rad}]$ 、指数マップでは回転量が $0[\text{rad}]$ に近い値となるものが多いと考えられる。このように推定されたと考えられる例を図 8 に示す。

次に、学習画像枚数を 100 枚から 4000 枚に変化させ、同様に誤差の回転量の平均を求めた実験を行った。推定画像枚数は 100 枚とした。結果を図 9 に示す。グラフの縦軸は誤差の回転量の平均、横軸は学習画像枚数を示す。結果から、学習画像枚数が変化しても回転行列を用いた場合が最も誤差が小さいことがわかる。また、ZYX-

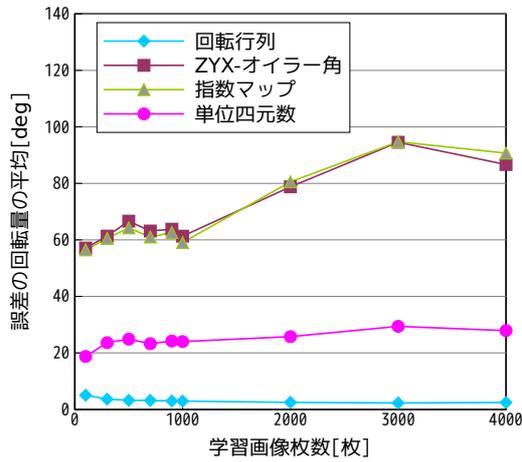


図 9 不連続となる画像を用いた場合の学習画像枚数変化による誤差の変化

オイラー角と指数マップは誤差がやはり大きいことがわかる。

これらの比較実験から、推定に用いる姿勢表現が一对一と連続性の性質を満たしていなければ、推定誤差が大きくなるといえる。

5. おわりに

本論文では、3自由度における姿勢表現の性質を実験的に評価した結果を報告した。姿勢表現として回転行列、ZYX-オイラー角、指数マップ、単位四元数を用いた姿勢推定を行い、それぞれの推定誤差を比較した。まず、見えに基づく姿勢推定手法の仕組みを説明し、そこから見えてくる姿勢表現に必要な性質である一对一と連続性について説明した。次に、4つの姿勢表現が一对一と連続性の性質を満たしているかどうかを考察した。そして、ランダムな姿勢の画像と、ZYX-オイラー角と指数マップでは不連続となる姿勢の画像を作成し、4つの姿勢表現を用いて実際に姿勢推定を行った。推定値と真値との誤差を求め、4つの姿勢表現を比較した。この実験から、用いる姿勢表現が一对一と連続性の性質を満たしていない場合には、推定誤差が大きくなることがわかった。また、性質を満たす回転行列は最も推定誤差が小さかった。このことから、姿勢表現に必要な性質が推定値に与える影響はかなり大きい。そのため、推定精度を上げるには、姿勢表現は一对一と連続性の性質を満たす必要があるといえる。

また、今回の実験では、学習画像枚数の変化による推定誤差の変化も調査した。その結果、学習画像枚数が100枚から2000枚までは、学習画像が増加するにつれて推定誤差は減少した。しかし、2000枚からは枚数が増加するにつれて、推定誤差が増加した。このような結果は論文 [7] でも述べられている。この原因を検討することも今後の課題である。

- [1] Yi Ma, Stefano Soatto, Jana Kosecka, and S. Shankar Sastry, An Invitation to 3-D Vision From Images to Geometric Models, Springer, 2006.
- [2] Frank Nielsen, Visual Computing: Geometry, Graphics, and Vision, CHARLES RIVER MEDIA, 2005.
- [3] John J. Craig, ロボティクス-機構・力学・制御-, 共立出版, 1991.
- [4] 徐剛, 辻三郎, 3次元ビジョン, 共立出版, 1998.
- [5] Zhengyou Zhang, A flexible new technique for camera calibration, Technical Report MSN-TR-98-71, Microsoft Research, 1998.
- [6] 玉木徹, 天野敏之, 金田和文, 球関数による姿勢表現と姿勢推定, MIRU2008, pp. 1134-1141, 2008.
- [7] 天野敏之, 玉木徹, 稠密なサンプル画像を用いた3次元物体の線形姿勢推定, MIRU2008, pp. 1128-1133, 2008.
- [8] 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, No. 8, pp. 2060-2069, 2007.
- [9] 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.
- [10] Thomas Melsner, Michael Reiter, and Horst Bischof, Appearance models based on kernel canonical correlation analysis, Pattern Recognition, Vol. 36, pp. 1961-1971, 2003.
- [11] 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.
- [12] Marsupi, VW New Beetle Polizei, Downloaded from DMI at <http://dmi.chez-alice.fr/modelsrescue1.html> (2009/03/18 accessed).
- [13] GSL-GNU Scientific Library, <http://www.gnu.org/software/gsl/>, 2008.