PA 8 CGを用いたカメラ校正法

玉木 徹　　山本 正信
新潟大学工学部

1 まえがき

カメラの位置や向きを推定するカメラキャリブレーション（校正）は、コンピュータビジョンの重要な課題の一つである。従来のキャリブレーションの研究は、点と点の対応からパラメータを推定するために、その対応点を与える作業が必要になっていた [1]。

本研究では、画像のレジストレーションに基づいて、対応点を与える必要のないカメラ校正手法を提案する。本手法は、CG により既知のキャリブレーションのための物体の画像を作成し、実際に物体を撮影した画像との輝度差を最小にするパラメータを非線形最小化によって推定する。

2 CG と実画像との座標変換

まず、形状が既知であるキャリブレーション用の物体の画像 I_1 を CG により生成する。$\boldsymbol{p}_1 = (x_1, y_1)^T$ を I_1 内の点とすると、CG のカメラ座標系での三次元座標 $\boldsymbol{P}_1 = (X_1, Y_1, Z_1)^T$ は、世界座標系での座標 $\boldsymbol{P}_w$ から回転 R と並進 $\boldsymbol{T}$ により変換されて、

$$\boldsymbol{p}_1 = \left(f_1 \frac{X_1}{Z_1}, f_1 \frac{Y_1}{Z_1} \right)^T$$
$$\boldsymbol{P}_1 = R\boldsymbol{P}_w + \boldsymbol{T}$$

であるとする。ここで

$$R = R_z(a)R_y(b)R_x(c) = \begin{pmatrix} \cos a & -\sin a & 0 \\ \sin a & \cos a & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos b & 0 & \sin b \\ 0 & 1 & 0 \\ -\sin b & 0 & \cos b \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos c & -\sin c \\ 0 & \sin c & \cos c \end{pmatrix}$$

$$\boldsymbol{T} = (t_x, t_y, t_z)^T$$

CG を生成する際には、既知の形状に対して変換 $R, \boldsymbol{T}$ を行ない、CG 画像 I_1 を得る。また、I_1 上の各点 $\boldsymbol{p}_1$ に対しては、CG であるので奥行き Z_1 が求まり（詳しくは後述する）、$\boldsymbol{P}_1$ は既知となる。

次に、実際にカメラで撮影した画像を I_2 とし、$\boldsymbol{p}_2 = (x_2, y_2)^T$ を I_2 内の点とすると、撮影したカメラの座標系での三次元座標 $\boldsymbol{P}_2 = (X_2, Y_2, Z_2)^T$ は、元の世界座標系での座標 $\boldsymbol{P}_w$ から $Q, \boldsymbol{S}$ により変換されて、

$$\boldsymbol{p}_2 = \left(f \frac{X_2}{Z_2}, f \frac{Y_2}{Z_2} \right)^T$$
$$\boldsymbol{P}_2 = R(Q\boldsymbol{P}_w + \boldsymbol{S}) + \boldsymbol{T}$$
$$Q = Q_z(\alpha)Q_y(\beta)Q_x(\gamma)$$
$$\boldsymbol{S} = (s_x, s_y, s_z)^T$$

となる（図 1）。ここで求めるものは、$\boldsymbol{P}_w$ から $\boldsymbol{P}_2$ への変換パラメータである $Q, \boldsymbol{S}$ である。それらを求めることで、$\boldsymbol{P}_w$ から $\boldsymbol{P}_2$ への回転 RQ と並進 $R\boldsymbol{S} + \boldsymbol{T}$ が計算できる。回転行列から各軸回りの回転角度は、容易に求めることができる [4]。

$\boldsymbol{P}_w$ から $\boldsymbol{P}_2$ への回転と並進をパラメータとして直接求めることも考えられるが、推定時の初期値が

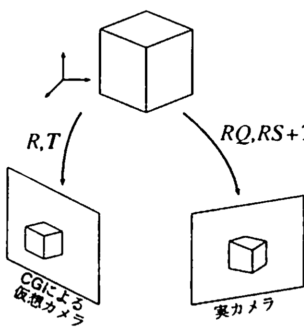


図 1: 世界座標系を介した CG 画像と実画像との関係

問題となる。たとえ値の近い R を回転の初期値として用いることができても、R との小さなずれが並進成分に大きく影響を与えてしまい、T は並進の初期値として使うことができなくなる。そのため、本手法では初期値を 0 として与えられる世界座標系での変換をパラメータとしている。

3 最適化による推定

最小化したいものは、画像 I_1 内の点 $\boldsymbol{p}_1$ の輝度 $I_1(\boldsymbol{p}_1)$ と画像 I_2 内の点 $\boldsymbol{p}_2$ の輝度 $I_2(\boldsymbol{p}_2)$ の差の二乗和である。目的関数を F、パラメータを $\boldsymbol{\theta} = (\alpha, \beta, \gamma, s_x, s_y, s_z)^T$ とすると、

$$\min_{\boldsymbol{\theta}} F, \quad F = \sum_i {r_i}^2, \quad r_i = I_1(\boldsymbol{p}_{1i}) - I_2(\boldsymbol{p}_{2i})$$

である。F は、画像 I_1 中の物体の領域内に点についてのみ和をとる。

$\boldsymbol{\theta}$ を推定するために、Gauss-Newton 法 [3] を用いる。反復によって推定値は $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \delta\boldsymbol{\theta}$ と修正され、その修正量 $\delta\boldsymbol{\theta}$ は次の連立一次方程式の解として得られる [2]。

$$\sum_i \sum_l^6 \frac{\partial r_i^u}{\partial \theta_k^u} \frac{\partial r_i^u}{\partial \theta_l^u} \delta\theta_l^u = -\sum_i r_i^u \frac{\partial r_i^u}{\partial \theta_k^u}, \quad k = 1, \ldots, 6$$

ここで $\boldsymbol{\theta} \equiv (\theta_1, \ldots, \theta_6)^T$ としている。それぞれの θ_k による偏微分は次のようになる。

$$\frac{\partial r}{\partial \theta_k} = -\frac{\partial \boldsymbol{p}_2}{\partial \theta_k} \frac{\partial I_2}{\partial \boldsymbol{p}_2} = -\left(\frac{\partial x_2}{\partial \theta_k}, \frac{\partial y_2}{\partial \theta_k} \right)^T \nabla I_2(\boldsymbol{p}_2)$$
$$\frac{\partial x_2}{\partial \theta_k} = \frac{f}{Z_2} \left(\frac{\partial X_2}{\partial \theta_k} - \frac{X_2}{Z_2} \frac{\partial Z_2}{\partial \theta_k} \right)$$
$$\frac{\partial y_2}{\partial \theta_k} = \frac{f}{Z_2} \left(\frac{\partial Y_2}{\partial \theta_k} - \frac{Y_2}{Z_2} \frac{\partial Z_2}{\partial \theta_k} \right)$$

X_2, Y_2, Z_2 の偏微分は以下のヤコビアンの要素である。

$$\frac{\partial \boldsymbol{P}_2}{\partial \alpha} = R\frac{\partial Q_z(\alpha)}{\partial \alpha}Q_y(\beta)Q_x(\gamma)R^{-1}(\boldsymbol{P}_1 - \boldsymbol{T})$$
$$\frac{\partial \boldsymbol{P}_2}{\partial \beta} = RQ_z(\alpha)\frac{\partial Q_y(\beta)}{\partial \beta}Q_x(\gamma)R^{-1}(\boldsymbol{P}_1 - \boldsymbol{T})$$
$$\frac{\partial \boldsymbol{P}_2}{\partial \gamma} = RQ_z(\alpha)Q_y(\beta)\frac{\partial Q_x(\gamma)}{\partial \gamma}R^{-1}(\boldsymbol{P}_1 - \boldsymbol{T})$$
$$\frac{\partial \boldsymbol{P}_2}{\partial \boldsymbol{S}} = R$$
$$\frac{\partial Q_z(\alpha)}{\partial \alpha} = \begin{pmatrix} -\sin\alpha & -\cos\alpha & 0 \\ \cos\alpha & -\sin\alpha & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\frac{\partial Q_y(\beta)}{\partial \beta} = \begin{pmatrix} -\sin\beta & 0 & \cos\beta \\ 0 & 0 & 0 \\ -\cos\beta & 0 & -\sin\beta \end{pmatrix}$$
$$\frac{\partial Q_x(\gamma)}{\partial \gamma} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin\gamma & -\cos\gamma \\ 0 & \cos\gamma & -\sin\gamma \end{pmatrix}$$

4 奥行きの算出

前頁で奥行き Z_1 は求まると述べたが、ここでは実際の算出方法について述べる。

CG を生成する際に、隠線消去などに z バッファが用いられている。これは CG 画像内の各点での物体の奥行きを保持しているが、物体の前後判定が可能であればよいので、その値は実際のカメラ座標系の Z 座標とは異なる。

z バッファの値を z_b、実際の奥行きを Z とすると、以下の変換式が成り立つ [5]。

$$Z = \frac{f_z n_z}{\frac{z_b}{s}(f_z - n_z) - f_z}$$

ここで f_z, n_z はそれぞれ後ろと手前の奥行きのクリップ面までの視点からの距離、s は z バッファの最大値[1]である。つまり、z バッファは $Z : [n_z, f_z] \rightarrow z_b : [0, s]$ という変換で生成されている。

この変換は線形ではなく、z バッファの精度はほとんどが手前のクリップ面付近に集中しているが、f_z と n_z を適切に設定すれば精度の問題はない。

5 実画像による実験

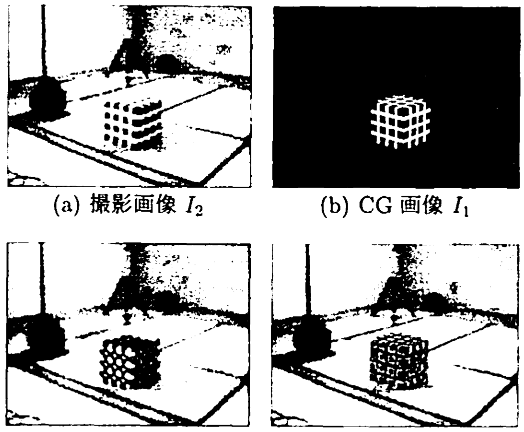
実画像を用いて本手法による実験を行なった。キャリブレーションのために使用した物体は、白黒の模様をつけた立方体（各辺 30cm）である。これを実際にカメラで撮影した画像を図 2(a) に、CG により生成した画像を図 2(b) に示す。図 2(b) は、CG 内でのカメラの位置、姿勢を GUI で制御できるアプリケーションを作成し、撮影した画像に近くなるようにした。これらの画像上で物体は正確に一致する必要はなく、図 2(c) に示すようにおおよそ重なっていればよい。この時に用いた R, T, f を表 1 に示す。

最適化のためのパラメータの初期値は 0 に近い値を乱数で与えた。この実験で推定されたパラメータの値を表 1 に示す。また、この推定値で画像 I_1 を変換して、I_2 と重ねて表示した画像を図 2(d) に示す。初期値（図 2(c)）では物体がずれて重なっていたが、推定値では物体は画像上で正確に一致しており、推定が精度良く行なわれたことを示す。

[1]例えば、符合無し 24 ビット z バッファであれば $s = 2^{24} - 1$ である。

$a = 0$	$\alpha = 1.6944$
$b = -40$	$\beta = -3.9676$
$c = 15$	$\gamma = 0.79463$
$t_x = 5$	$s_x = 34.791$
$t_y = -37$	$s_y = -4.2182$
$t_z = -285$	$s_z = -27.491$
$f = 512.476$	

表 1: $\boldsymbol{R}, \boldsymbol{T}$ の値と $Q, \boldsymbol{S}$ の推定値 [deg]



(a) 撮影画像 I_2　(b) CG 画像 I_1
(c) 初期値での差　(d) 推定値での差

図 2: 実験画像

6 あとがき

本研究では、CG 画像とのレジストレーションに基づいた、カメラ校正手法を提案し、実画像を用いた実験の結果を示した。本手法では焦点距離 f は求めていないため、他の手法で推定する必要がある。

本手法は微小な差のある二枚の画像中の物体の変換パラメータを推定しているため、フレーム間隔の小さい動画像における、既知の物体の運動の追跡 [6] に応用ができると考えている。

参考文献

[1] 浅田 尚紀,「カメラキャリブレーション」, 松山 久野 井宮 編, コンピュータビジョン, 第 3 章, 新技術コミュニケーションズ, 1998.

[2] H. S. Sawhney, S. Ayer, "Compact representations of videos through dominant and multiple motion estimation," *T-PAMI*, Vol. 18, No. 8, pp. 814–830, 1996.

[3] G. A. F. Seber, C. J. Wild, *Nonlinear Regression*, Wiley, 1989.

[4] J. J. Craig, ロボティクス, 三浦 下山 訳, 共立出版, 1991

[5] "OpenGL Developer FAQ and TroubleShooting Guide", http://www.opengl.org/developers/faqs/technical/depthbuffer.htm

[6] 山本, 川田, 近藤, 越川,「ロボットモデルに基づく人間動作の 3 次元動画像追跡」, 信学論 D-II, Vol.J79-D-II, No.1, pp.71–83, 1996