

## バーチャルスタジオにおけるカメラキャリブレーション高精度化の考察 A Study on Accurate Camera Calibration Method for Virtual Studio

大久保 英彦† 山内 結子† ミツ峰 秀樹† 深谷 崇史† 井上 誠喜†  
Hidehiko Okubo Yuko Yamanouchi Hideki Mitsumine Takashi Fukaya Seiki Inoue

### 1. はじめに

カメラワークを伴う実写とCGの合成を可能にするバーチャルスタジオの技術はいまや多くの番組制作で利用されている。CGを描画するために必要な、実際の撮影カメラの位置・姿勢などのパラメータを求める最も一般的な手法は、回転量を検出できる雲台を利用してそこから得られるパン・チルト量を元に、実カメラの相対的な姿勢変化をデータとして利用するものである。その場合、絶対的なカメラ位置がそれらパン・チルト量と正確に関連づけられていることが合成時の精度にかかわってくる重要な要素となる。

またこれまでの研究から、撮影カメラのレンズ歪をリアルタイムにシミュレートさせることで「実写の位置により正確に一致した」CG生成の可能性を見出している[1]。

本研究においては、回転量を検出する雲台を利用した、一般的なバーチャルスタジオ用カメラシステムのためのカメラキャリブレーション手法を考案した。さらに、前述のCG生成とあわせることで、幾何的に正確な実写・CG合成を実現したので報告する。

### 2. 三脚・雲台にマウントしたカメラのモデル化

実写の位置に正確にCGを一致させるためには、実際のカメラレンズをパラメータ化した「カメラパラメータ」が必要である。それは幾何学的な意味により、レンズの位置・姿勢を表す外部パラメータと、焦点距離や光軸中心位置・レンズ歪などの内部パラメータに分類される。

一般的にカメラパラメータを求めるプロセスは、カメラキャリブレーションと呼ばれ、様々な手法が提案されている。その中には、動画像解析のみからキャリブレーションを行うものも存在するが、処理時間や頑健性などの面で課題がある。そこで本研究では、図1に示すような固定した三脚にマウントしたカメラ雲台を用い、適当なカメラ位置・姿勢の初期状態においては画像解析を利用し、十分に正確な雲台の相対回転量を利用することでリアルタイムにカメラ位置・姿勢を得ることとした。その初期状態において、市松模様などの既知のパターンを撮影した画像を解析しキャリブレーションをおこなった。実際に、市松模様を利用する有効なキャリブレーションツールはいくつか存在する[2]。しかし、多くの場合はカメラの位置・姿勢に拘束条件のない状態で各種パラメータを求めており、実際にそういったツールを利用して、求められたカメラ位置と上述の雲台から得られるパン・チルトデータとの整合を直接とることはできない。

そこで、三脚と雲台にマウントされた状態のカメラ全体を表すカメラモデルを考え、それによって実現されるカメラの位置・姿勢を決定するパラメータを外部パラメータとするようなカメラモデルを想定した。

一般的にカメラの位置・姿勢をあらわす外部パラメータ  $\theta_{ext}$  は、各座標軸方向の平行移動と回転を表す6つのパラメータで規定される。

$$\theta_{ext} = [t_x, t_y, t_z, \omega, \varphi, \kappa]^T \quad (1)$$

これは、剛体変換をあらわす回転  $R$  と平行移動  $t$  からなる行列  $M$  :

$$M = \begin{bmatrix} R & t \\ \mathbf{0}_3 & 1 \end{bmatrix} \quad (2)$$

で表されることと同等である。今、空間に任意の世界座標を考える。雲台のパン・チルト回転軸は直交し一点で交差していると仮定し、この点を雲台中心と呼ぶとする。この雲台は世界座標の位置・姿勢  $M_p$  で設置されているとする。さらに、その雲台中心からオフセットされたカメラレンズ位置・姿勢を  $M_c$  とし、パン・チルト量を

$M_{PanTilt}$  で表すとする、最終的なカメラ位置・姿勢  $M$  は、

$$M = M_c M_{PanTilt} M_p \quad (3)$$

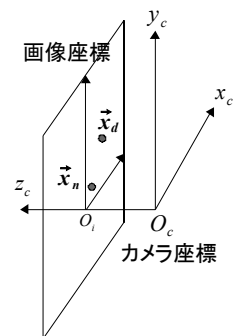
で表すことができる。この場合、カメラの位置・姿勢を決定する未知のパラメータは  $M_p$  ,  $M_c$  に対応する12個のパラメータとなる。

なお、内部パラメータ  $\theta_{int}$  としては以下の図2に示すHeikkiläのモデル[3] (アスペクト比、焦点距離、画像中心座標、レンズ歪係数の8パラメータ)を採用する。

$$\theta_{int} = [s, f, u_0, v_0, kc_1, kc_2, kc_3, kc_4]^T \quad (4)$$



図1. 三脚・雲台上的カメラ



$$\vec{x}_d = [u_d, v_d]^T \quad \rightarrow \text{歪を含んだ座標系}$$

$$\vec{x}_n = [u_n, v_n]^T \quad \rightarrow \text{無歪座標系}$$

歪  $\rightarrow$  無歪座標系への変換

$$\bar{u}_d = u_d - u_0 \quad \bar{v}_d = v_d - v_0 \quad r_d^2 = \bar{u}_d^2 + \bar{v}_d^2$$

$$\vec{x}_n = \vec{x}_d + \begin{bmatrix} \bar{u}_d (kc_1 r_d^2 + kc_2 r_d^4) \\ \bar{v}_d (kc_1 r_d^2 + kc_2 r_d^4) \end{bmatrix} + \begin{bmatrix} 2kc_3 \bar{u}_d \bar{v}_d + kc_4 (r_d^2 + 2\bar{u}_d^2) \\ kc_3 (r_d^2 + 2\bar{v}_d^2) + 2kc_4 \bar{u}_d \bar{v}_d \end{bmatrix}$$

図2. レンズ歪係数の関係式

† 日本放送協会 放送技術研究所 (人間・情報)

### 3. キャリブレーションアルゴリズム実装

前述のとおり、ここで提案するキャリブレーション手法は、既知の校正パターンを撮影した多視点(複数の)画像を解析することで各パラメータを推定するものに分類される。各パラメータは最終的にバンドル調整(格子交点の予測位置と観測位置の幾何学的距離の2乗和の最小化)をおこなうことにより、高い精度での校正が可能である。

通常の場合は、カメラの位置・姿勢に拘束条件がないために、 $K$ 枚の画像を利用する場合には、それぞれの画像に対応したカメラ位置・姿勢を未知パラメータとして扱うことになる。すなわち、評価すべきパラメータベクトルを  $\theta$  とすると、

$$\theta^T = [\theta_{\text{int}}^T, \theta_{\text{ext}}(1)^T, \theta_{\text{ext}}(2)^T, \dots, \theta_{\text{ext}}(K)^T] \quad (5)$$

となる。ここで、 $\theta_{\text{ext}}(k)$  は、 $k$ 枚目の画像に対応した撮影カメラ位置・姿勢をあらわす。

今回提案しているカメラモデルにおいては、式(3)の  $M_p$ 、 $M_c$  は、雲台中心位置を固定し、カメラの雲台マウント状態が変わらない場合は、撮影画像に関して変化しない上に、 $M_{\text{PanTilt}}$  はデータ上既知となるため、評価するパラメータベクトルはシンプルに、

$$\theta^T = [\theta_{\text{int}}^T, \theta_{\text{ext}}^p(1)^T, \theta_{\text{ext}}^c(1)^T] \quad (6)$$

となる。ここで、 $\theta_{\text{ext}}^p$ 、 $\theta_{\text{ext}}^c$  は  $M_p$ 、 $M_c$  に対応する外部パラメータベクトルをあらわす。

続いて、パラメータを最適化する際の目的関数を

$J(\theta)$  とし、画像座標系における校正パターンの特徴点の観測位置を  $e_0(n, k)$ 、カメラモデルを経て計算により推定した特徴点の位置を  $e_d(n, k)$  とすると、

$$J(\theta) \text{ は、} \quad J(\theta) = y^T(\theta) V^{-1} y(\theta) \quad (7)$$

であらわされる。なお、

$$y^T(\theta) = [(e_0(1,1) - e_d(1,1))^T, (e_0(2,1) - e_d(2,1))^T, \dots, (e_0(N,K) - e_d(N,K))^T] \quad (8)$$

であり、 $n$  は特徴点の個数 ( $n=1, \dots, N$ )、 $k$  は画像の枚数 ( $k=1, \dots, K$ ) をあらわす。さらに  $V$  は観測誤差の共分散行列をあらわす。

最適化としては、この  $J(\theta)$  を最小にする  $\theta$  を求めればよいことになる。

$$\hat{\theta} = \arg \min_{\theta} J(\theta) \quad (9)$$

## 4. 実験

### 4.1 実験手法

上述の提案キャリブレーションアルゴリズムの評価を行うべく、以下の方法で実験をすすめた。

#### ①市松模様の交点の測定：

Jean-Yves Bouguet 氏により公開されているカメラキャリブレーション Matlab スクリプト[4]の特徴点抽出部分のみを利用して、市松模様の交点の観測座標を抽出した。市松模様は  $8 \times 10$  周期で、三脚・雲台上に設置したカメラから6ポーズの姿勢に関してデータを取得した。

#### ②従来のキャリブレーション手法による評価：

①で得られた特徴点データから、カメラ位置・姿勢に関しての拘束条件をあたえずにキャリブレーションを実行

し、再投影誤差による評価を行った。

#### ③提案キャリブレーション手法による評価：

①で得られた特徴点データと、対応するパン・チルトデータを使用して、提案カメラモデルによるキャリブレーションを実行し、再投影誤差による評価を行った。

#### ④CG との合成：

レンズ歪をシミュレートするCG描画装置にキャリブレーション結果を適用しCGを生成、実写と合成を行った。

### 4.2 実験結果

上述②の結果として、すべての特徴点サンプルの平均再投影誤差が0.262ピクセル、③の結果としては、0.336ピクセルで、実質的なキャリブレーション精度という意味ではバーチャルスタジオにとって十分であり、かつ②と③も同程度の結果を得ることができた。図3に、③の結果として実写画像とCGの合成を行った例を示す。



図3. 市松模様のCG合成結果

## 5. 考察とまとめ

提案カメラモデルによるキャリブレーション手法の最大の利点は、回転雲台を利用したカメラシステムに対して、そのパン・チルトデータに整合したキャリブレーションが行える点である。

実験結果の再投影誤差の数値比較だけを行うと、従来のキャリブレーション手法と提案手法は大きな差がないように見える。しかし、雲台上にカメラがマウントされているために、外部パラメータの回転部分のデータが合成精度に対して極めて大きく影響する。拘束条件をあたえない通常のカメラモデルでキャリブレーションを行った場合は、パン・チルト角データに整合させるための何らかの補間処理などが必要になり、精度面と運用面で不利になる。

また、今回想定している三脚固定のカメラシステムだけでなくペダスタルを利用した場合でも、その移動情報を任意の位置でリセットさせることが可能になるため、カメラの位置精度と運用性の向上が見込まれる。

今後はズームレンズに対応すると同時に、実際の制作スタジオ・CG制作現場での利用を図っていく。

## 参考文献

- [1] 大久保他：「CGと実写のリアルタイム合成時におけるレンズ歪の考察」, FIT2006, p.267-268, 2006.
- [2] Intel, "Open Source Computer Vision Library", <http://www.intel.com/technology/computing/opencv/index.htm>
- [3] Heikkilä, J, "Geometric Camera Calibration Using Circular Control Points", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, pp. 1066-1077, Oct 2000.
- [4] Jean-Yves Bouguet, "Camera Calib. Toolbox for Matlab", [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html)