

全方位型マルチカメラシステムによるパノラマ動画像の生成

池田 聖[†] 佐藤 智和[†] 横矢 直和[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科

〒630-0101 奈良県生駒市高山町 8916-5

E-mail: †{sei-i,tomoka-s,yokoya}@is.aist-nara.ac.jp

あらまし 近年、全方位型カメラの開発や計算機の高機能化に伴って、遠隔地の画像を計算機に描画させ自由な見回しを可能とするテレプレゼンスシステムが実用化されつつある。このようなテレプレゼンスシステムでは解像度の高いパノラマ動画像を用いることで、利用者に対して高い臨場感を与えることが出来るが、現在までに多く用いられてきた単一のカメラを用いた全方位型カメラシステムでは画像の解像度が低いという問題がある。本研究では全方位型マルチカメラシステムの幾何学的及び光学的なキャリブレーションを行い、その結果を利用して高解像度なパノラマ動画像の作成を行う。実験では、作成したパノラマ動画像を用いて違和感のないテレプレゼンスが実現できることを確認し、定量的な評価を行なう。

キーワード 全方位型マルチカメラシステム, パノラマ動画像, カメラキャリブレーション

Generating a Panorama Movie Using an Omnidirectional Multi-camera System

Sei IKEDA[†], Tomokazu SATO[†], and Naokazu YOKOYA[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

Takayama 8916-5, Ikoma, Nara, 630-0101 Japan

E-mail: †{sei-i,tomoka-s,yokoya}@is.aist-nara.ac.jp

Abstract Recently, telepresence systems that allow us to experience a remote place virtually is widely investigated. In such systems, the problem is how to obtain panoramic images of remote scene. An omnidirectional multi-camera system is more useful to make panoramic images than those with a single camera because the multi-camera system can easily capture high-resolution images. However, the exact camera calibration is necessary to virtualize the real world accurately. In this report, we describe a method of generating a high-resolution panorama movie by calibrating an omnidirectional multi-camera system geometrically and photometrically. In experiments, a spherical panoramic video has been successfully computed to show the validity of the proposed calibration method.

Key words Omnidirectional Multi-camera System, Panorama Movie, Camera Calibration

1. はじめに

全方位型センサを用いることで遠隔地の様子をユーザに提示し、自由な見回しを可能とするテレプレゼンスシステムは、エンターテイメントや教育などの様々な分野で応用が期待されている。このようなテレプレゼンスを実現する方式として、曲面ミラー [1], [2] や広角レンズ [3], [4] を用いる方式は、広く普及している通常のカメラにアタッチメントを取りつけることで比較的容易に実現できるため、これまで広く用いられてきた。しかし、これらの方式では全方位の画像を 1 台のカメラで撮影するために解像度の不足や、視線方向による解像度の不均等が問題

となる。

これに対して複数のカメラを外向きに配置することで全方向を撮影できる全方位型マルチカメラシステムは、システムが複雑となるが高い解像度の画像が取得でき、また視線方向による解像度の変化が少ないという特徴を持つ。このような全方位型マルチカメラシステムから得られる画像をつなぎ合わせて違和感のないパノラマ動画像を生成するためには、高い精度でのカメラ間のキャリブレーションが必要である。全方位型マルチカメラシステムに対するキャリブレーション手法はいくつか提案されているが、多くの手法は没入型の大型スクリーンを用い、スクリーンに影された繰り返しパターンを撮影することでカメラ

間のキャリブレーションを行なっている [5], [6]. しかし, このような手法では各カメラに対して奥行き方向に広くマーカを配置することが出来ず, 十分なキャリブレーション精度を得ることは難しい. 実験結果からも画像間のつなぎ目の不連続が確認でき, また画像間のつなぎ合わせの精度に関する議論もされていない.

そこで我々は三次元レーザ測量機器トータルステーションとキャリブレーションボードを用いることで, 多数のマーカを空間的に配置し, 高精度にカメラキャリブレーションを行う手法を提案する. 本研究では全方位型マルチカメラシステムとして, 上方を含む 6 台のカメラによって全方位の撮影が可能で Point Grey Research 社製の Ladybug (図 1) を用いることで, 実際に全方位の高解像度画像を動画として撮影し, パノラマ画像の生成を行う. Ladybug は図 1(左) に示すような水平方向に 5 個, 上方向に 1 個の CCD カメラを配置した小型カメラヘッドと図 1(右) に示すような HDD アレイによる蓄積ユニット (30 分) を用いることで, 全方向の約 75% 以上を動画 (15fps) として撮影できるシステムである.

以下 2 章では, Ladybug の幾何学的及び光学的キャリブレーションについて述べる. 3 章ではキャリブレーション結果を用いた全天球パノラマ画像の生成方法について述べ, カメラ間の視差の影響について述べる. 4 章では, 実際に Ladybug のキャリブレーションを行い, パノラマ画像を作成する. また, 評価実験として, テレプレゼンスシステムを試作することで, 生成したパノラマ動画が違和感なくテレプレゼンスに利用可能であることを示す. さらに, パノラマ画像のつながり具合について誤差を定義し定量的な評価を行う.

2. 全方位型マルチカメラシステム Ladybug のキャリブレーション

本章では, パノラマ画像の合成に必要な Ladybug の各カメラの位置・姿勢の推定と, 入力画像の歪み補正を行う幾何学的キャリブレーション, 及び明度低下現象とカメラ間の相対的な色調を補正する光学的キャリブレーションについて述べる.

2.1 幾何学的キャリブレーション

幾何学的キャリブレーションでは各カメラの幾何学的なカメラパラメータである内部パラメータ (焦点距離, レンズ歪み, 歪み中心, アスペクト比) と外部パラメータ (カメラの位置, 姿勢) の推定を行なう. 外部パラメータに関してはカメラ間の位置及び姿勢の関係を求めるために全てのカメラについて統一した座標系での推定を行う.

これらを正確に推定するためにはシステムの全方向に三次元位置関係が既知のマーカを多数配置する必要がある. まず, マーカの配置方法について述べる. 本手法では, マーカとして図 2 に示すようなキャリブレーションボードに印刷された格子模様の交点を用いる. マーカの三次元位置は, キャリブレーションボードの 4 隅の格子点をトータルステーションで計測し線形補間して求める. キャリブレーションボードは図 2 に示すように, 各カメラに対して奥行き方向に移動させ, その都度三次元位置の計測と, キャリブレーションボードの撮影を行なう. これによ



図 1 Ladybug のカメラヘッド (左) と蓄積ユニット (右)

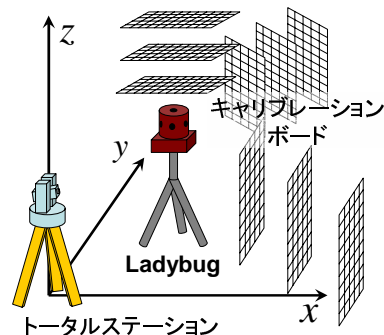


図 2 キャリブレーションボードの配置

り, 容易に多数のマーカを空間的に配置することが可能となる. 全てのカメラでマーカボードを撮影し終わるまで, トータルステーションにより張られる世界座標系を統一することで, システムの全方向にマーカを同時に配置するのと同等のデータを得ることが可能である. また, 画像上での格子点の位置は, 直線の交点を算出することによりサブピクセルの精度で検出し, キャリブレーションの精度向上を図る.

次に, 各パラメータの推定方法について述べる. 内部パラメータはあらかじめ各カメラについて Tsai [7] の手法を用いて推定しておく. ただし, 歪み係数については 2 次係数 k_2 まで考慮し精度向上を図る. 各カメラの外部パラメータは線形演算と非線形最適化により画像上の投影誤差を最小化することで求める [9]. 本手法では各カメラの外部パラメータを世界座標系におけるカメラ $c (c = 1, 2, \dots, 6)$ の位置 T_c と姿勢 R_c を用いて, 世界座標系からカメラ座標系への変換行列 M_c として表わす.

$$M_c = \begin{bmatrix} R_c & T_c \\ 0 & 1 \end{bmatrix} \quad (1)$$

M_c を求めるために, マーカ $m (m = 1, 2, \dots)$ の三次元位置 x_m とカメラ $c (c = 1, 2, \dots, 6)$ の入力画像上での位置 u_m とを対応づけ, 出口の手法 [8] を用いて世界座標系からカメラ座標系への変換行列 M'_c を計算する. しかし M'_c の自由度は 12 となり R_c が正規直交基底を成さない. そこで, 真のカメラ位置が推定された光軸上に存在すると仮定して, 線形演算により 6 自由度に補正された変換行列 M_c を得る. 最後に, マーカの三次元位置を画像上に投影した座標 v_m とマーカの入力画像上での位置 u_m との二乗誤差

$$E_c = \sum_m (u_m - v_m)^2 \quad (2)$$

を勾配法によって最小化することで M_c と内部パラメータを更新する。

2.2 光学的キャリブレーション

光学的なキャリブレーションでは、明度低下現象の補正とカメラ間の色調補正を行う。一般的な全方位型マルチカメラシステムでは、それぞれカメラで広角レンズを用いて広い視野の画像を取得するため、画像の周辺部分が暗くなる明度低下現象が生じる。明度低下現象として、広角レンズ特有の $\cos^4 \theta$ 現象 [10] と望遠レンズに特有の口径蝕現象 [11] が知られているが、広角レンズを用いたレンズ系では前者の影響が支配的であるため、ここでは $\cos^4 \theta$ 現象のみを考慮する。本研究ではレンズ径の実測値 l と焦点距離 f を使用し、以下の式を用いることで $\cos^4 \theta$ 現象を補正する。ただし、 I, I' は明度低下現象補正前後の画素の輝度値、 θ はカメラの光軸に対する入射光の角度である。

$$I' = \frac{\pi l^2 \cos^4 \theta}{f^2} I \quad (3)$$

次に、カメラ間の色調の補正方法について述べる。一般に撮像面上の放射照度 J は物体表面の放射輝度 L に対して線形的な関係 $J = aL + b$ を持つとされており、各カメラごとに a, b が異なるためカメラ間での色調の違いが発生する。本手法では、照明条件を固定した同じシーンを撮影することで各カメラ c の変換係数 a_c, b_c を推定する。変換係数は、カメラ c の画像と基準カメラ c' の画像の輝度 i に関する正規化ヒストグラム $h_c(i), h_{c'}(i)$ を一致させるように決定する。すなわち、評価関数:

$$e(a_c, b_c) = \sum_i \left\{ h_{c'}(i) - \frac{1}{a_c} h_c \left(\frac{i - b_c}{a_c} \right) \right\}^2 \quad (4)$$

を定義し、これを最小化することで各カメラの a_c, b_c を推定しカメラ間の色調を補正する。

3. パノラマ画像の作成

ここでは前述した手法によって得られるキャリブレーション結果を基にし入力画像から球面透視投影画像としてパノラマ画像を作成する方法について述べ、加えてカメラ間の視差（投影中心のずれ）の影響について考察する。

3.1 パノラマ画像の作成方法

一般的に、カメラを放射状に配置するマルチカメラシステムでは、各カメラの投影中心は一致せず、視差が発生するために厳密な一点透視投影に基づく継ぎ目のないパノラマ画像を作成することは不可能である。しかし、撮影対象までの距離がカメラの投影中心間の距離に比べて十分大きい場合は、視差の影響が相対的に小さくなるため、各カメラの投影中心が一致していると見なすことができる。そのため撮影対象は十分遠方に存在するものと仮定し、投影面である球面 S の半径を十分に大きく設定する。球面の中心は全てのカメラの投影中心の重心位置 G に設定する。このような仮定の下で、各画像上の全て画素を球面 S に投影することでパノラマ画像を生成する。ただし、球面 S 上の点 s の画素 $I_S(s)$ に対応する入力画像が複数存在する場合、複

数の画像が滑らかにつながるように画像間のブレンドを行う。カメラ c の画像面上への点 s の投影位置 u_c とその点に最も近い画像の端辺上の点までの距離を γ_c とすると、点 s の輝度値 $I_S(s)$ は、点 s からの投影が可能なカメラの集合を $C(s)$ として以下の式で与える。

$$I_S(s) = \frac{\sum_{c \in C(s)} \gamma_c I_c(u_c)}{\sum_{c \in C(s)} \gamma_c} \quad (5)$$

3.2 視差の影響

以下では、パノラマ画像を作成する際に設定する球面 S を無限遠に設定した場合、視差の影響がパノラマ画像上で 1 ピクセル以内に収まる対象物までの距離について述べる。

図 3 に示すように撮影対象 x が 2 つのカメラ c, c' 上の点 $u_c, u_{c'}$ に撮影されたとき上述の手法では、球面 S 上の $s_c, s_{c'}$ に投影され、2 重にずれた画像が生成される。球面 S の円周長が N ピクセルで構成されているとすると、 $\angle s_c G s_{c'} < \frac{2\pi}{N}$ であれば、視差の影響は 1 ピクセル以内に収まっていると考えられる。無限遠に設定された球面を考えたとき $\angle s_c G s_{c'} \rightarrow \angle s_c x s_{c'}$ となるので $\angle T_c x T_{c'} < \frac{2\pi}{N}$ という条件を満たす範囲に撮影対象 x が存在すればよい。

2 台のカメラの投影中心間の距離が d とし、カメラから撮影対象 x までの距離が等しい場合に限定すると、視差の影響が 1 ピクセル以下となるカメラ間のベースラインから撮影対象までの距離 λ は以下の式で表現できる。

$$\lambda > \frac{d}{2 \tan \frac{\pi}{N}} \quad (6)$$

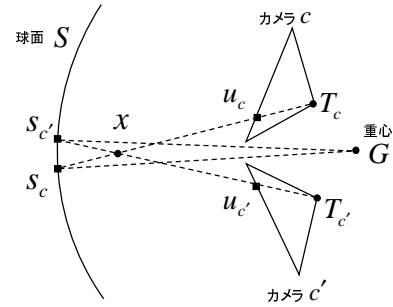


図 3 視差の影響

4. 実験

上述した手法で実際に Ladybug のキャリブレーションを行い、パノラマ動画画像を作成した。以下では、個々のキャリブレーション結果の確認、パノラマ動画画像の作成、テレプレゼンスシステムの試作、定量的な評価実験について述べる。

4.1 Ladybug のキャリブレーション

Ladybug の幾何学的なキャリブレーションでは、Ladybug を三脚に固定し、キャリブレーションボードをカメラの奥行き方向に約 50cm 間隔で移動させ、各カメラごとに 3 回撮影した。水平方向と上方向の各カメラにはキャリブレーションボードの格子点のうちそれぞれ 187 個、170 個用い、合計 561 個、510 個のマーカの三次元位置と画像上の位置を取得した。マーカの三次

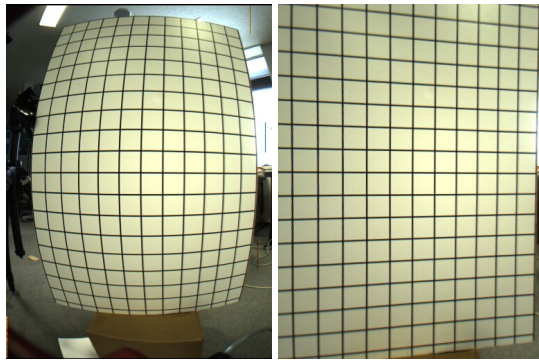


図 4 レンズ歪み補正前(左)と補正後(右)



図 5 明度低下現象補正前(左)と補正後(右)



図 6 色調補正前(左)と補正後(右)

元位置はキャリブレーションボードの 4 隅の格子点を LEICA 社製のトータルステーション TCR1105 で計測することにより算出した。

まず、図 4 に推定した内部パラメータを用いた入力画像の歪み補正結果を示す。同図より格子模様が直線に戻っていることから、おおむね正しく内部パラメータが推定されていることが確認できる。次に、図 5 に明度低下現象の補正結果を示す。同図左のように補正前の画像では、周辺部分の輝度値が全体的に低下しているのに対して、同図右のように補正後の画像ではそれらがある程度補正されていることが分かる。また、図 6 は色調補正前後でパノラマ画像を作成し、画像のつながり目部分を比較したものである。同図から色調補正前に比べて補正後のパノラマ画像の方が画像の境界が目立たなくなっていることから、おおむね正しく色調が補正されていることが確認できる。



図 7 Ladybug による屋外の撮影

4.2 パノラマ画像の作成

前節で述べたキャリブレーション結果を用いて 3.1 節の手法によってパノラマ動画像を作成した。入力動画像は図 7 の様に Ladybug を自動車の屋上に固定し屋外を走行しながら撮影することにより得た。図 8 に各カメラから得られた入力画像(解像度: 768×1024)の例を示す。また、これらの入力画像から作成したパノラマ画像(解像度: 3840×1920)を図 9 に示す。パノラマ動画像は、極座標を用いて画像を平面に展開しており図 9 の下部の黒い部分は、入力画像の存在しない部分である。図 9 から、パノラマ動画像での位置ずれや入力画像間の境界は目立ず、幾何学的にも光学的にもおおむね正しくキャリブレーションが行なえていることが分かる。

また、図 10 に示すブレンドなしのパノラマ画像を前節を同様の手法で作成し、画像のつながり目に関して比較を行った。同図と図 9 の比較によりブレンドなしのパノラマ画像では、カメラに接近した物体や、太陽光が直射した場合の撮影素子の輝度値の飽和によって画像間のつながり目が容易に確認できるが、ブレンドありのものは太陽光が直射した場合でも画像間のつながり合わせは連続的であり、非常に接近した物体を除いて殆んど画像間のつながり目を知覚することは出来なかった。

4.3 テレプレゼンスシステムの試作

実際に全方位の見回しが可能なテレプレゼンスシステムを試作することで、生成したパノラマ画像が違和感なくテレプレゼンスに利用可能であることを確認する。

試作したテレプレゼンスシステムは表 1 に示す様に球面ディスプレイ、コントローラ、PC の 3 つの装置により構成されている。上述のパノラマ動画像を JPEG 形式で PC に蓄積し、図 11 のようにユーザがコントローラを用いて指定した方向のみを切り出して球面ディスプレイに投影することで全方向の見回しを可能としている。本システムでは Ladybug の撮影時のフレームレートである 15fps を越える約 20fps での描画が可能であった。なお、球面ディスプレイの解像度は 1024×768 であり、パノラマ画像は 2048×1024 の解像度に縮小した。

テレプレゼンスシステムは解像度、視野角、対話性の面において非常に臨場感の高いシステムが出来たが、球面ディスプレイの解像度の制限のため入力画像の解像度を十分に生かしたも



図 8 入力画像 (上方向 (右下) と水平方向 (その他))



図 9 全天球動画の 1 フレーム (ブレンドあり)



図 10 全天球動画の 1 フレーム (ブレンドなし)

のではない。より解像度の高いディスプレイを利用することで、さらに臨場感の高いシステムの構築が可能である。

4.4 定量的な評価実験

異なるカメラにより得られる画像中の対応点が、パノラマ画像を生成した際にどの程度ずれて球面に投影されるかを定量的に評価した。

図 12 に示すように Ladybug では隣り合うカメラにより得られる画像には共通領域が存在する。この領域内に円形マーカを写し、画像上でのマーカの重心位置をサブピクセルの精度で算出する。本実験ではこのマーカの重心位置を $u_c, u_{c'}$ としたとき、3.2 節で述べた図 3 での角 $\angle s_c G s_{c'}$ の角度を誤差値 ϵ として評価した。ただし、4.1 節のキャリブレーション結果から視差の影響が 1 ピクセル以下となる対象物体までの距離は約 24m となるため、マーカはシステムから約 30m 離して配置し、各カメラにつき 100 点以上計測した。

表 2 に各カメラ間の画像のつなぎ合わせにおける誤差 ϵ の最大値と平均値を示す。番号が 6 のものが上向きのカメラを示し、それ以外は水平方向のカメラを示している。実験から誤差 ϵ は平均 0.0063rad であった。これは入力画像面上では約 3 ピクセルに相当する。しかし、0.0498rad のような大きな誤差値を持つ部分も存在するが、その様な大きな誤差値を持つ部分はパノラマ画像上に殆んど現れない入力画像の端の部分に分布する傾向にあることが分かった。

5. ま と め

本研究では、全方位型マルチカメラシステム Ladybug に対して、カメラ間の幾何学的及び光学的なキャリブレーションを行ない、その結果を基にして全天球パノラマ動画像を作成する手法を提案した。幾何学的キャリブレーションではキャリブレーションボードとトータルステーションを用いて、仮想的にシステムの全方向に多数のマーカを配置することで精度の向上を図った。また、実際に全方向の見回しが可能なテレプレゼンスシステムを試作し、本研究で作成したパノラマ画像が違和感なくテレプレゼンスに利用可能であることを確認した。さらに、パノラ

表 1 テレプレゼンスシステムの機器構成

球面ディスプレイ	elumens VisionStation
コントローラ	Microsoft SideWinder Game Pad Pro
PC	CPU: Intel Pentium4 1.7GHz, RAM: 1GB
グラフィクスカード	Nvidia Geforce4

表 2 誤差 ϵ [rad]

カメラ番号	1-2	2-3	3-4	4-5	5-1
サンプル数	227	258	229	216	305
最大	0.0214	0.0378	0.0258	0.0227	0.0198
平均	0.0072	0.0057	0.0054	0.0048	0.0041

カメラ番号	6-1	6-2	6-3	6-4	6-5
サンプル数	176	131	127	123	154
最大	0.0178	0.0498	0.0208	0.0107	0.0112
平均	0.0058	0.0144	0.0095	0.0063	0.0052



図 11 テレプレゼンスシステムの外観



図 12 隣りあうカメラの画像の共通領域

マ画像生成における画像のつなぎ合わせ精度を定量的に評価した結果、入力画像面上で平均約 3 ピクセル程度のずれであることが分かった。今後は、広角レンズに適した歪みモデルについて検討を行い、さらに精度の高いパノラマ画像の生成を試みる。

文 献

- [1] 山澤, 尾上, 横矢, 竹村, “全方位画像からの視線追従型実時間画像生成によるテレプレゼンス”, 電子情報通信学会論文誌 (D-II), Vol. J81-D-II, No. 5, pp. 880-887, 1998.
- [2] M. Fiala, A. Basu, “Line Segment Extraction in Panoramic Images”, Winter School of Computer Graphics 2000, Vol 10, No. 1, pp. 179-186, 2002.
- [3] K. Miyamoto, “Fish Eye Lens”, Journal of Optical Society of America, 54(8), pp. 1060-1061, 2000.
- [4] R. A. Hicks, R. Bajcsy, “Catadioptric Sensors that Approximate Wide-angle Perspective Projections”, Proc. of CVPR2000, Vol. 1, pp. 545-551, 2000.
- [5] 戎野, 山澤, 竹村, 横矢, “円筒パターンを用いた全周ステレオ画像センサのキャリブレーション手法”, 日本 VR 学会第 4 回大会論文集, pp. 211-212, 1999.
- [6] 棚橋, 佐藤, 王, 丹羽, 山本, “全方向ステレオシステム (SOS) のキャリブレーション手法”, 映像情報メディア学会誌, Vol. 56, No. 4, pp. 603-610, 2002.
- [7] R. Y. Tsai, “A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses”, IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, pp. 323-344, 1987.
- [8] 出口, “射影幾何学による PnP カメラ補正問題の統一的解法”, 情報シンポジウム, Vol. 90, pp. 41-50, 1990.
- [9] T. Sato, M. Kanbara, N. Yokoya, H. Takemura, “Dense 3D Reconstruction of an Outdoor Scene by Hundreds-Baseline Stereo Using a Hand-Held Video Camera”, Journal of Computer Vision, Vol. 47, Nos. 1-3, 2002.
- [10] B. K. P. Horn, Robot Vision, Mit Press, 1986.
- [11] 天野, 浅田, 馬場, “ズームレンズの光学的歪みの解析と補正 — 可変円筒モデルを用いた口径蝕現象の解析と画像周辺明度低下の補正 —”, 画像の認識・理解シンポジウム (MIRU'96) 論文集, Vol. 1996, No. 5, pp. I-307-312, 1996.