

全方位型マルチカメラシステムによる全天球動画の生成と テレプレゼンスへの応用

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

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全天球動画を用いることで、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得られる動画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全天球動画の生成を行う。また、生成した全天球動画の応用例として没入型ディスプレイを用いたテレプレゼンスシステムを構築する。

Generation of a Spherical Movie Using an Omnidirectional Multi-camera System and Its Application to Telepresence

Sei IKEDA Tomokazu SATO Naokazu YOKOYA
Graduate School of Information Science, Nara Institute of Science and Technology

Abstract: Telepresence systems using an omnidirectional image sensor enable us to experience remote sites with rich sensation. An omnidirectional multi-camera system is more useful to acquire high-resolution omnidirectional images of outdoor scenes than a single camera-based catadioptric system. However, exact calibration of the multi-camera system is necessary to generate a spherical movie. In this report, we describe a spherical movie generation method based on geometric and photometric calibration of the omnidirectional multi-camera system. A prototype of telepresence system is also shown as an application of generated spherical movies. The proposed high-resolution telepresence system has been proven to enable us to experience remote sites with rich sensation.

1 はじめに

全方位ビデオ映像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、エンターテイン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に使用する全方位画像センサとして、曲面ミラー [1, 2] や広角レンズ [3] 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す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全方位動画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、これら手法は没入型の大型スクリーンを用い、スクリーンに投影されたマーカ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 [4, 5]。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困難であるため、キャリブレーション結果が画像上の量子化誤差やマーカ位置の検出誤差に敏感となる [6]。

本稿では、全方位型マルチカメラシステムに対して仮想的に配置された多数のマーカによる幾何学的キャリブレーションおよび光度低下現象とカメラ間の色調を考慮した光学的キャリブレーションを行うことで、高解像度な全天球動画を生成する手法を提案する。提案手法では、キャリブレーション結果を用いて各カメラの画像を仮想的な球面に投影するこ

とで視覚的に違和感の少ない全天球動画画像を生成する。以下2章では、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画像の生成手法について述べ、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を行い、全天球画像を生成する。また、生成した全天球動画画像の応用例としてテレプレゼンスシステムを試作する。

2 全方位型マルチカメラシステムのキャリブレーション

本研究では全方位型マルチカメラシステムとしてとして、図1に示すPoint Grey Research社製のLadybugを用いる。Ladybugは図1(左)に示すように水平方向に5個、上方向に1個のCCDカメラが配置された小型カメラヘッドと図1(右)に示すHDDアレイによる蓄積ユニットからなり、全天球の約75%以上を15fpsの動画画像として約30分間撮影できるシステムである。

全方位型マルチカメラシステムのキャリブレーションは、幾何学的キャリブレーション、光学的キャリブレーションの順に行う。幾何学的キャリブレーションでは、まず入力画像の歪み補正などに用いる内部パラメータの推定を行い、次に各カメラの外部パラメータとしてカメラの位置および姿勢を推定する。光学的キャリブレーションでは、まず明度低下現象の補正に使用するパラメータの推定を行い、次にカメラ間の色調補正に使用するパラメータの推定を行う。

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クを多数配置する必要がある。まず、マークの配置手法について述べる。本手法では、マークとして図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。マークの三次元位置は、キャリブレーション



図1: Ladybugのカメラヘッド(左)と蓄積ユニット(右)

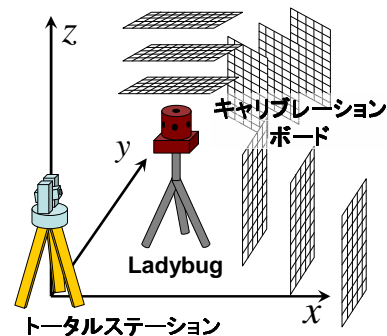


図2: キャリブレーションボードの配置

ボードの3隅の格子点をトータルステーションで計測し線形補間によって求める。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う。これにより、容易に多数のマークを空間的に配置することが可能となる。全てのカメラでキャリブレーション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ク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。

次に、各パラメータの推定手法について述べる。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 κ_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マークの画像上への投影誤差の和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 として以下のように表わす。

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

M_c を求めるために、マーク $m(m=1, 2, \dots)$ の三次

元位置 \mathbf{x}_m とカメラ $c(c = 1, 2, \dots, 6)$ の入力画像上での位置 \mathbf{u}_m とを対応づけ、線形最小自乗法を用いて世界座標系からカメラ座標系への変換行列 \mathbf{M}'_c を計算する [8]. 次に、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、線型演算により \mathbf{M}'_c を 6 自由度に補正し変換行列 \mathbf{M}_c を得る. 最後に、以下に示すマーカの三次元位置を画像上に投影した座標 \mathbf{v}_m とマーカの入力画像上での位置 \mathbf{u}_m の距離の自乗和 E_c を勾配法によって最小化することで \mathbf{M}_c を更新する.

$$E_c = \sum_m (\mathbf{u}_m - \mathbf{v}_m)^2 \quad (2)$$

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、それぞれの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、一般に、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手法について述べる. 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、異なるカメラ c と c' の撮像面上における放射輝度 I_c と $I_{c'}$ の間にも、 $I_{c'} = a_c I_c + b_c$ の関係があると考えられる. 本手法では、上式で示す色調のモデルを仮定し、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_0 の画像の輝度 i に関する正規化ヒストグラム $h_c(i), h_{c_0}(i)$ を一致させるように決定する. すなわち、評価関数:

$$e(a_c, b_c) = \sum_i \left\{ h_{c_0}(i) - \frac{1}{a_c} h_c \left(\frac{i - b_c}{a_c} \right) \right\}^2 \quad (4)$$

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 全天球画像の生成

ここでは前述の手法によって得られるキャリブレーション結果をもとに、入力画像から球面透視投影画像として全天球画像を生成する手法について述べ、加えてカメラ間の視差の影響について考察する.

3.1 全天球画像の生成

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全天球画像を作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. そのため本研究では、撮影対象は十分遠方に存在するものと仮定し、投影面である仮想的な球面 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、各画像を球面 S に投影することで全天球画像を生成する. ただし、球面 S 上の点 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処理を行う. 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边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、全天球画像を生成する際に設定する球面 S を無限遠に設定した場合、視差の影響が全天球画像上で 1 画素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 画素で構成されるとすると、 $\angle s_c G s_{c'} < \frac{2\pi}{N}$ であれば、視差の影響は 1 画素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は $\angle s_c G s_{c'} \rightarrow \angle s_c x s_{c'}$ 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を 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

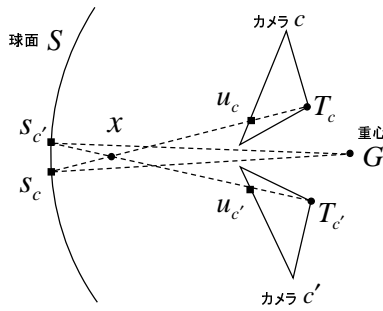


図 3: 視差の影響

1画素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.

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

Ladybug の場合, カメラの投影中心間の距離は約 40mm であり, 球面へ投影された画像の円周長が 3840 画素で構成されていると仮定すると, 視差の影響が 1画素以内に収まる距離 λ は, 約 24m である.

4 実験

2章で述べた手法を用いて実際に Ladybug のキャリブレーションを行い, 全天球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, 全天球動画画像の生成, 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 個を用い, 合計 3315 個のマーカの三次元位置と画像上の位置を取得した. マーカの三次元位置はキャリブレーションボードの 3 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. 同図から補正前の画像では, 周辺部分の輝度値が全体的に低下しているが, 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

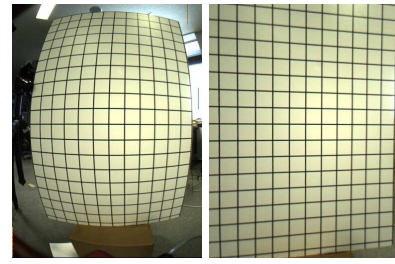


図 4: レンズ歪み補正前 (左) と補正後 (右)



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)

前後で全天球画像を生成し, 画像のつなぎ目部分を比較したものである. なお, 両画像ともブレンド処理は行っていない. 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.

4.2 全天球画像の生成

前節で述べたキャリブレーション結果を用いて 3.1 節の手法によって全天球動画画像を生成した. 入力動画画像は Ladybug を自動車のルーフ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 7 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全天球画像 (解像度: 3840×1920) を図 8 に示す. 全天球動画画像は, 極座標を用いて画像を平面に展開して



図 7: 入力画像 (上方向 (右下) と水平方向 (その他))



図 8: 全天球動画の 1 フレーム (ブレンドあり)

おり図8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8から、全天球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えていることが分かる。

4.3 テレプレゼンスシステムの試作

生成した全天球動画像の応用例として、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した。試作したテレプレゼンスシステムは表1に示すよう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。全天球動画像をJPEG形式でPCに蓄積し、図9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利用者はインタラクティブに視線方向を操作できる。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が1024×768であるため、全天球画像は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、視野角の面において十分に臨場感の高いシステムであることが確認されたが、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。

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い、その結果をもとにして全天球パノラマ動画像を生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、本研究で生成した全天球動画像を用いることで臨場感の高いテレプレゼンスが可能であることを確認した。今後

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, メモリ: 1GB
グラフィクスカード	NVIDIA Geforce4



図 9: テレプレゼンスシステムの外観

は、全天球動画像の応用として全方位の見回しだけでなく自由に視点移動可能なウォークスルーシステムの構築を試みる。

参考文献

- [1] 山澤, 八木, 谷内田, “移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案”, 電子情報通信学会論文誌, Vol. J81-D-II, No. 5, pp. 880-887, 1998.
- [2] Y. Yagi, Y. Nishizawa and M. Yachida, “Estimating Location and Avoiding Collision Against Unknown Obstacle for the Mobile Robot Using Omnidirectional Image Sensor COPIS”, Proc. Int. Workshop on Intelligent Robots and Systems(IROS-91), Vol. 2, pp. 909-914, 1991.
- [3] Movegeye Inc., “Movegeye”, 1999. <http://www.movegeye.co.jp/>
- [4] 戎野, 山澤, 竹村, 横矢, “全方位ステレオ画像センサによる実環境の仮想環境への取り込み”, 電子情報通信学会 技術研究報告, MVE99-82, 1999.
- [5] 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.
- [6] 佐藤, “複数の動画像を用いたカメラパラメータ推定に基づく屋外環境の三次元モデル化に関する研究”, 奈良先端科学技術大学院大学 博士論文, NAIST-IS-DT0161018, 2003.
- [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 Jour. 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”, 情報処理学会 コンピュータビジョン'90 ビジョンと環境理解シンポジウム論文集, 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”, Int. Jour. of Computer Vision, Vol. 47, No. 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部明度低下の補正-”, 電子情報通信学会論文誌, Vol. J80-D-II, No. 6, pp.159-166, 1997.