

NAIST-IS-MT0151005

修士論文

全方位型マルチカメラシステムのキャリブレーション に基づく高解像度パノラマ動画の生成

池田 聖

2003年2月7日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

池田 聖

審査委員： 横矢 直和 教授
小笠原 司 教授
山澤 一誠 助教授

全方位型マルチカメラシステムのキャリブレーション に基づく高解像度パノラマ動画の生成*

池田 聖

内容梗概

画像センサを用いた屋外環境の仮想化は, テレプレゼンスやウォークスルーなど様々な分野で応用が可能である. 現在までに, 実環境の画像から仮想化を試みる研究は数多く報告されており, このような研究においては全方位画像センサは実環境の映像を効率よく取得することができるという点で注目されている. しかし, 従来広く用いられてきた単一のカメラを用いた全方位画像センサでは画像の解像度の不足が問題とされてきた. これに対して, 複数のカメラを外向きに配置することにより高解像度に全方向の画像を獲得することができる全方位型マルチカメラシステムが開発されている. このようなシステムを用いて違和感の少ない仮想空間を構築するためには複数カメラ間の高精度なキャリブレーションが必要である.

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手法を提案し, その結果を利用して高解像度な全天球パノラマ動画を作成する. 実験では, まずパノラマ画像を作成し, キャリブレーション結果の確認を行う. 次に, パノラマ画像において各画像間のつながり具合を示す誤差を定義し, 定量的な評価実験を行う. また, 作成した高解像度パノラマ動画の応用例として, 様々な映像提示デバイスを用いた臨場感の高いテレプレゼンスシステムを構築する.

* 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0151005, 2003年2月7日.

キーワード

全方位型マルチカメラシステム, カメラキャリブレーション, パノラマ画像生成,
テレプレゼンス

Generating a Panoramic Movie by Calibrating an Omnidirectional Multi-camera System*

Sei IKEDA

Abstract

In the field of mixed reality, virtualization of real outdoor scenes is widely investigated for a number of applications such as telepresence and walk-through systems. An omnidirectional multi-camera system is more useful to virtualize outdoor scenes than a monocular camera system, because the multi-camera system can easily capture high-resolution omnidirectional images. However, the exact calibration of the system is necessary to virtualize the real world accurately.

In this paper, we describe a method for generating a high-resolution panoramic 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. Finally, we show some prototypes of telepresence system using various immersive displays.

Keywords:

Omnidirectional Multi-Camera System, Camera Calibration, Panoramic Image Generation, Telepresence

* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0151005, February 7, 2003.

目次

1. はじめに	1
2. 関連研究と本研究の方針	3
2.1 全方位画像センサと取得画像の特徴	3
2.1.1 単眼カメラを用いる全方位画像センサ	5
2.1.2 多眼カメラを用いる全方位画像センサ	6
2.1.3 全方位画像センサのテレプレゼンスへの利用	7
2.2 全方位型マルチカメラシステムのキャリブレーションに関する従来手法	8
2.3 本研究の方針	10
3. 全方位型マルチカメラシステムのキャリブレーション	12
3.1 全方位型マルチカメラシステム Ladybug	12
3.2 幾何学的キャリブレーション	15
3.2.1 カメラパラメータの定義	15
3.2.2 マーカの配置と検出	17
3.2.3 カメラパラメータの推定	19
3.3 光学的キャリブレーション	21
3.3.1 明度低下現象の補正	21
3.3.2 カメラ間の相対的な色調補正	21
3.4 実験結果	23
4. 全天球パノラマ動画の作成	34
4.1 パノラマ画像の作成方法	34
4.2 視差の影響	35
4.3 実験結果	37
4.3.1 全天球パノラマ動画の生成	37
4.3.2 定量的な評価実験	38

5. テレプレゼンスシステムの試作	44
5.1 ヘッドマウントディスプレイを用いたシステム	45
5.2 球面ディスプレイを用いたシステム	45
5.3 円筒ディスプレイを用いたシステム	46
6. まとめ	50
謝辞	52
参考文献	53

目 次

1	全方位画像の例 (撮影: NHK エンタープライズ 21)	4
2	全方位型マルチカメラシステムの例	9
3	Ladybug のカメラヘッド (左) と蓄積ユニット (右)	13
4	入力画像 (上方向 (右下) と水平方向 (その他))	14
5	Ladybug による画像の取り込み	14
6	世界座標系とカメラ座標系の関係	18
7	キャリブレーションボード	19
8	キャリブレーションボードの配置	19
9	マーカの検出	20
10	$\cos^4 \theta$ 現象のレンズモデル	22
11	キャリブレーションに使用したマーカの画像例	25
12	キャリブレーションボード上の格子点の検出例	26
13	レンズ歪み補正結果	27
14	カメラの位置および姿勢の推定結果	28
15	明度低下現象の補正結果	29
16	色調補正前後のヒストグラム (R 成分)	30
17	色調補正前後のヒストグラム (G 成分)	31
18	色調補正前後のヒストグラム (B 成分)	32
19	色調補正に用いた入力画像	33
20	パノラマ画像の生成手順	35
21	投影球面からパノラマ画像への展開	36
22	視差の影響	36
23	自動車による屋外の撮影	38
24	入力画像 (奈良 鴻池周辺: 上方向 (右下) と水平方向 (その他))	39
25	全天球動画画像の 1 フレーム (ブレンド処理あり)	40
26	全天球動画画像の 1 フレーム (ブレンド処理なし)	40
27	ブレンド処理によるスムージングの効果	41
28	隣り合うカメラ画像の共通領域	41

29	画像のつなぎ合わせに関する誤差 ϵ の分布	43
30	ヘッドマウントディスプレイを用いたテレプレゼンスシステムの 外観	46
31	球面ディスプレイを用いたテレプレゼンスシステムの外観	47
32	円筒ディスプレイを用いたテレプレゼンスシステムの入力画像 (飛 騨高山)	49
33	円筒ディスプレイを用いたテレプレゼンスシステムの外観	49

表 目 次

1	全方位画像センサの特徴	8
2	Ladybug の仕様	13
3	本論分で用いるカメラの内部・外部パラメータ一覧	18
4	推定された内部パラメータ	27
5	各カメラの再投影誤差 [pixel]	29
6	画像のつなぎ合わせに関する誤差 ϵ [rad]	42
7	ヘッドマウントディスプレイを用いたテレプレゼンスシステムの 機器構成	45
8	球面ディスプレイを用いたテレプレゼンスシステムの機器構成	47
9	円筒ディスプレイを用いたテレプレゼンスシステムの機器構成	49

1. はじめに

都市空間などの屋外環境の仮想化に関する技術は、近年の計算機技術の発達とともに注目され盛んに研究が行われている [1, 2, 3, 4]. これらの技術によって仮想化された屋外環境を、利用者の要求に応じてインタラクティブに映像として提示することで、仮想化された空間を擬似的に体験することが可能となる. 特に近年、仮想空間のインタラクティブな提示手法として、ビデオカメラを用いて遠隔地の情景を画像として蓄積し、利用者の要求に応じて見たい場所・時間・方向の映像を提示する、テレプレゼンス技術に関する研究が盛んである [5, 6, 7, 8, 9]. このようなテレプレゼンス技術は、遠隔地の仮想体験のほかに、機械の遠隔操縦や観光案内システムなどのように幅広い分野への応用が可能である.

このようなテレプレゼンスを実現するために、従来から、アクティブカメラを用いた手法 [10, 11, 12, 13] が提案されてきた. アクティブカメラを用いた手法 [10, 11, 12] では、遠隔地の情景をネットワークを介してリアルタイムに伝送し、ユーザの要求に従ってカメラの向きを制御することで、自由な方向の見回しを可能としている. しかし、ネットワークの遅延やカメラの回転駆動による遅延のため、要求する方向が提示されるまでに時間がかかるという問題があった. このような問題に対して、全方位画像センサと呼ばれるような水平方向 360 度の視野を撮影することが可能な画像センサの利用が注目されている. 全方位画像センサでは、得られる全方位画像のうち利用者が要求する部分だけを切り出して提示することで、見回しによる遅延が非常に小さくなり、臨場感の高いテレプレゼンスが実現できる [14].

しかし、従来から広く用いられてきた単一のカメラを用いた全方位画像センサでは画像の解像度の不足が問題とされてきた. これに対して、近年、複数のカメラを用いて高解像度に全方向の画像を獲得することができる全方位型マルチカメラシステムが開発されている. 全方位型マルチカメラシステムは高解像度であるという点に加えて、カメラの配置方法にもよるが、一般に撮影できる視野角が広く、また、解像度があまり方向に依存しないといった特長が挙げられる. しかし、このようなシステムを用いて臨場感の高いテレプレゼンスを実現するためのパノラマ画像を得るためには、複数のカメラ間の位置・姿勢および各カメラの光学的な特

性を計測する高精度なキャリブレーションが必要である。

本論文では、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、および、その結果を用いた高解像度な全天球パノラマ動画像の生成手法を提案する。実験では、提案手法を用いた全方位型マルチカメラシステムのキャリブレーション、およびパノラマ動画像の生成を行い、それらの結果を確認する。また、パノラマ画像において各画像間のつながり具合を示す誤差を定義し、定量的な評価実験を行う。さらに、作成した高解像度パノラマ動画像の応用例として、様々な映像提示デバイスを用いた臨場感の高いテレプレゼンスシステムを構築する。

以下、2章では、まず関連研究と本研究の方針について述べる。3章では、全方位型マルチカメラシステムとしてLadybug[15]の幾何学的及び光学的キャリブレーションについて詳述し、実験ではキャリブレーションの結果を示す。4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動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考察を行う。実験では、実際に屋外環境を全方位型マルチカメラシステムで撮影し、それぞれのカメラの動画像を入力として高解像度なパノラマ動画像を生成し、結果を提示する。また、パノラマ画像のつながり具合についての誤差を定義し定量的な評価を行う。5章では生成した高解像度なパノラマ動画像の応用例としてテレプレゼンスシステムの試作結果を示し、本研究で生成したパノラマ動画像が臨場感の高いテレプレゼンスに有効であることを示す。最後に、6章で本論文をまとめる。

2. 関連研究と本研究の方針

遠隔地や過去の映像を利用者の要求に応じて提示することで利用者にあたかもその場所にいるかのような感覚を与える技術はテレプレゼンスと呼ばれる。従来提案されている全方位画像センサを利用したテレプレゼンス [5, 14] では、利用者が要求する方向に応じて、図1のような全方位画像から提示するディスプレイの画角にあわせて一部分を切り出し、ディスプレイを介して平面透視投影画像を提示している。このような方法をとることで、過去に取得された映像や遠隔地からネットワークを介して伝送されたリアルタイム映像において、利用者は見回しに関する遅延を殆ど感じることなく、仮想化された環境を自由に見回すことができる。

このようなテレプレゼンスの用途への応用が可能な全方位画像センサは、これまで様々な方式のものが提案されてきた。本章では、まず全方位画像の取得手段として提案されている代表的な全方位画像センサおよびそれらのセンサにより得られる全方位画像の特徴について概観する。従来提案されている全方位画像センサのなかでも、複数のカメラを用いる全方位型マルチカメラシステムでは、パノラマ画像を取得するために正確なカメラ間のキャリブレーションが必要となるため、次に、全方位型マルチカメラシステムのキャリブレーションの従来手法について述べる。最後に、本研究の位置付けと方針について述べる。

2.1 全方位画像センサと取得画像の特徴

全方位画像センサはその構造から、1つのカメラとミラーやレンズを組み合わせることで全方向をカバーするものと、複数のカメラで全方向をカバーするものの2つに大別することができる。一方、全方位画像センサにより取得した全方位画像を臨場感の高いテレプレゼンスに利用する場合、使用する全方位画像には以下のような特徴が要求される。

解像度が高い: 全方位画像センサを用いたテレプレゼンスでは、全方位画像の一部を切り出して利用者に提示するため、入力する全方位画像には高解像度のものが要求される。解像度が低い場合には、実環境の物体のエッジや文字などの細かい部分の表現が困難となり、利用者は違和感を感じる。



図 1 全方位画像の例 (撮影: NHK エンタープライズ 21)

解像度が一様である: 全方位画像センサから得られる画像の解像度が方向によって異なる場合, 提示画像の解像度が利用者の視線変化によって変化するため, 違和感を感じる.

視野角が広い: 全方位画像の視野角が狭いと, 画像を提示できる視線方向が限定されるため, 違和感を感じる. より自由な見回しを行うためには, 広い視野角の全方位画像が必要である.

一点透視投影画像である: 一点透視投影画像でない場合, パノラマ画像上の各画素の投影中心が一致しないため視差が生じ, 撮影対象が暈けたり二重に提示されて違和感が発生する.

以上のような観点から, それぞれのセンサの特徴について説明し, 臨場感の高いテレプレゼンスに適した画像センサについて述べる.

2.1.1 単眼カメラを用いる全方位画像センサ

広く普及している通常のカメラにレンズやミラーなどのアタッチメントを取りつける方式の全方位画像センサは、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。単眼のカメラにより全方位画像の取得を実現しているものは、魚眼レンズを使用するものと曲面ミラーを用いるものとに分けられる。

魚眼レンズを用いる手法: [16, 17, 18] 地面に対して鉛直上方向に向けた通常のビデオカメラに超広角なレンズを取りつけることでの水平方向 360 度の画像を取得することが可能な手法である。得られる画像は完全な一点透視投影ではなく、また水平面以下まで撮影できるものは少ない。レンズの特性上、側方領域の解像度は低く、レンズ形状に応じたキャリブレーションが必要である。

Movingeye 社 [18] では、魚眼レンズを取り付けたビデオカメラを天井に下向きにして設置することで、複数の利用者が一度に異なる方向を監視することのできるシステムを開発している。このシステムでは、魚眼レンズカメラから得られた全方位画像の一部分を切り出し、実時間でレンズ歪みの補正を行い、要求される異なる方向の画像を同時に複数の利用者に提示している。しかし、側方領域の解像度が低く、撮影できる視野角は全天球の半分程度であるため、監視カメラのような用途には適しているが、全方位の見回しを行うテレプレゼンスのような用途には適さない。

曲面ミラーを用いる手法: [19, 20, 21] 曲面ミラーを用いる手法では、地面に対して鉛直上方向きに向けた単眼のビデオカメラに凸型の曲面ミラーを取り付けて、ミラーに映りこむ鏡像を撮影することで全方位の画像を取得する。ミラーの形状として球面 [19]、円錐 [20]、双曲面 [21]、放物面 [22] などが提案されているが、中でも双曲面および放物面ミラーを用いた手法は完全な一点透視投影画像の取得が可能という特徴がある。ただし、原理的に鏡面の影やカメラ自身の影になる領域の撮影は出来ず、また視線方向によって解像度が大きく異なるという問題がある。

山澤ら [14, 6] は双曲面ミラーを用い遠隔地において撮影された全方位画像

を実時間で平面透視投影画像に変換し複数の利用者に提示するテレプレゼンシステムを提案している。利用者はヘッドマウントディスプレイを介し、姿勢センサから得られるそれぞれの視線方向に応じた方向の映像を見ることができる。しかし、一台のビデオカメラで全方位を撮影しているため解像度の不足が目立つことや上方向の映像がないため鉛直方向の視野角が狭く感じるといった問題がある。

これらの方式では、常にシームレスな全方位画像を取得できる反面、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。曲面ミラーを用いる手法の欠点である解像度の不足を補うために双曲面ミラーと高解像度カメラを組み合わせる手法 [5] が提案されているが、解像度が方向によって大きく異なるという問題や上方向の視野が確保されていないという問題は解決されていない。

2.1.2 多眼カメラを用いる全方位画像センサ

複数のカメラにより全方向の視野を一度に撮影し、高解像度の画像を獲得することの出来るカメラシステムを全方位型マルチカメラシステムと呼ぶ。全方位型マルチカメラシステムには、角錐ミラーを用いて水平方向360度を撮影可能なものと、複数のカメラを外向きに配置して全天球の撮影が可能なものとに分けられる。

角錐ミラーを用いる手法: [23, 24] 鉛直方向に複数のカメラが配置されており、角錐ミラーに写る鏡像を撮影することにより水平方向360度の画像を一度に取得可能である。ミラーを用いて仮想的に各カメラの投影中心を一致させることで完全な一点透視投影画像を得ることができ、解像度は一様である。鉛直方向の視野が狭く、上下方向の撮影は出来ない。

島村ら [25] は、六角錐ミラーを用いてステレオ視の可能な全方位ステレオ画像センサ [23] からの高解像度なパノラマ画像を利用し、インタラクション可能な仮想環境の構築を行った。遠景については取得したパノラマ画像から動物体を排除したものを提示し、近景にはステレオ視により復元した三次元モデルを使用して建築物など表現する。それらの映像を円筒面ディスプレ

イ CYLINDRA[25, 26] 提示することで、水平方向の自由な見回しや、近景に付加された仮想物体の移動操作が可能であった。しかし、複数のカメラと角錐ミラーを用いているため装置が複雑となり、正確なキャリブレーションが困難であった。

カメラを外向きに配置する手法: [15, 27] 複数のカメラが外向き放射状に配置され全方向の視野の画像を獲得することの出来るカメラシステムである。一般に、全てのカメラの投影中心は一致しておらず、光学的特性は正確な一点透視投影ではない。特にセンサに近接する対象はパノラマ画像上で幾何学的に不連続となる。システムの構造によるが、解像度は方向にあまり依存せずほぼ一様であり、視野角については一般的には制限はない。

Point Grey Research 社では、全方位型マルチカメラシステム Ladybug[15] を用いて秒間 5 フレーム以下での画像のつなぎ合わせを行い利用者に高解像度な平面透視投影画像を提示するテレプレゼンスを実現している。しかし、キャリブレーション手法の詳細は公開されておらず、利用目的によってはキャリブレーションが問題となる。

これらの方式では、高解像度かつ広い視野の画像を取得でき、得られる画像の解像度があまり方向に依存しないなどの特長がある。複数のカメラにより得られる複数の画像をどのように統合するかが問題となり、カメラ間のキャリブレーションが必要である。

2.1.3 全方位画像センサのテレプレゼンスへの利用

前述した全方位画像センサの特徴をテレプレゼンスへの利用という観点でまとめる。表 1 は、臨場感の高いテレプレゼンスの実現という観点から各画像センサの特徴について、解像度の高さ、解像度の一様性、視野角、光学特性が一点透視投影であるかを表している。ただし、ここでは曲面ミラーを用いる手法として、一点透視投影という光学的性質を満たす双曲面ミラーについて示している。

同表に示すように全ての特徴を満たす画像センサは存在しない。しかし、カメラを外向きに配置する方式の全方位型マルチカメラシステムは、撮影対象がカメ

表 1 全方位画像センサの特徴

	解像度		視野角	一点透視投影
	高さ	一様性		
魚眼レンズを用いる方法	×	×	×	×
双曲面ミラーを用いる方法	×	×	×	
双曲面ミラーとHDカメラを用いる方法		×	×	
角錐ミラーを用いる方法			×	
カメラを外向きに配置する方法				×

ラ間の距離に対して十分遠方に存在する場合には、近似的に一点透視投影とみなすことができる。よって、屋外環境のように撮影対象が遠方に存在する環境においては、全方位型マルチカメラシステムが臨場感高いテレプレゼンスの実現に対して最も適しているといえる。ただし、全方位型マルチカメラシステムの各カメラから得られる画像を統合して、違和感のないパノラマ動画像を生成するためには、高い精度でのシステムのキャリブレーションが必要である。

2.2 全方位型マルチカメラシステムのキャリブレーションに関する従来手法

本節では、現在までに提案されてい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について述べ、それらの手法の問題点について述べる。

角錐ミラーを用いた全方位画像センサは、設計どおり各カメラの投影中心が一致するように各カメラが配置されていなければ、正確な一点透視投影画像を生成することが出来ない。そこで、川西ら [23] は、図 2(a) に示すような六角錐ミラーを用いた全方位画像ステレオセンサに対して、各カメラの位置姿勢が設計仕様を満たすようにするための物理的な調整手法を提案した。まず、撮像画像面上にある鏡面領域の位置および面積が等しくなるように各カメラの光軸方向の調整を行い、隣り合うカメラの画像の連続性を保つようにしてカメラの光軸方向の位置の



(a) 全方位ステレオ画像センサ [23]

(b) SOS[27]

図 2 全方位型マルチカメラシステムの例

調整を行った。しかし、このような物理的な調整により正確に設計仕様を満たすようにするのは非常に困難である。また、カメラの調整が隣接するカメラ間の相対的な調整のため誤差の蓄積が大きい。そのため生成したパノラマ画像中に不連続な部分が生じた。

これに対して、戎野ら [28] や棚橋ら [29] はシステムを囲む大型の没入型スクリーンに投影されたマーカを用いて蓄積誤差の発生しない方式でシステムのキャリブレーションを提案した。戎野らは、六角錐ミラーを用いた全方位ステレオ画像センサ [23] に対して、上述のような各カメラの投影中心の微妙なずれを容認し、カメラパラメータを推定することでパノラマ画像生成の精度の向上を図った。システムを円筒形ディスプレイ CYLINDRA[25, 26] の中央に配置し、スクリーンに投影したマーカを用いて、各カメラの光軸方向の物理的な調整および画角の推定を行った。棚橋らは、図 2(b) のような全方向ステレオシステム (SOS)[27] を没入型 6 面ディスプレイ COSMOS[30] の中央に配置し、スクリーンに投影したマーカを用いて各カメラの姿勢を示すカメラパラメータを推定した。ただし、カメラの位置に関しては、SOS の設計パラメータを用いている。

両手法ともに、カメラの位置を既知とし設計パラメータを使用しているが、一般の全方位型マルチカメラシステムを考えた場合、常に設計パラメータが使用できるとは限らない。また、長期間の使用に伴ってカメラパラメータが変化する可能性もあり、必ずしも設計パラメータを使用することで高精度なキャリブレーションが出来るわけではない。

また、プロジェクトにより大型の没入型スクリーンに投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位置と姿勢の関係を推定するキャリブレーション手法では、各カメラに対するマーカ配置の空間的な広がりが小さく、マーカをシステムの近くに配置させるといったことも出来ない。そのようなマーカの配置方法では、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非常に敏感となる [31]。

2.3 本研究の方針

前述したように、全方位画像センサを用いて屋外環境の映像を取得し臨場感の高いテレプレゼンスを実現するためには、全方位型マルチカメラシステムが適している。ただし、全方位型マルチカメラシステムで違和感のないパノラマ動画像を生成するためには、システムのキャリブレーションが重要となる。しかし、従来から提案されているカメラキャリブレーションの手法は、設計パラメータを用いるなど汎用性に欠けており、一般の全方位型マルチカメラシステムに適用することは困難であった。また、キャリブレーション時の各カメラに対してマーカが空間的に広く配置されていないなど、キャリブレーションの精度向上のために改善の余地が残されている。

そこで、本論文では、臨場感の高いテレプレゼンスの実現を想定し、全方位型マルチカメラシステムを用いることで高解像度なパノラマ動画像の生成を実現する。提案手法では、まず全方位型マルチカメラシステムの幾何学および光学的キャリブレーションを行い、次にキャリブレーション結果を基にして全天球パノラマ動画像を生成する。システムの幾何学的なキャリブレーションでは、設計パラメータを用いずカメラパラメータの推定を行う。本手法では、物体表面の三次元位置を計測可能なレーザ計測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使

用して各カメラに対して空間的に広くマーカを配置することでキャリブレーションの精度の向上を図る。加えて光学的なキャリブレーションでは各カメラの明度低下現象とカメラ間の色調を補正する。パノラマ画像の生成には、上記のキャリブレーションの結果を用いて、撮影対象が十分遠方にあることを仮定して、仮想的な投影中心を設定し画像のつなぎ合わせを行う。これにより、違和感の少ない高解像度な全天球パノラマ動画像を生成する。

3. 全方位型マルチカメラシステムのキャリブレーション

本章では、全方位型マルチカメラシステムによるパノラマ動画の生成に必要な各カメラの位置・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、およ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。

3.1 全方位型マルチカメラシステム Ladybug

まず、全方位型マルチカメラシステムとして本研究で使用する Point Grey Research 社製の Ladybug[15] について説明する。Ladybug は図 3(左) に示すような水平方向に 5 個、上方向に 1 個の CCD カメラを外向きに配置した小型カメラヘッドと図 3(右) に示すようなハードディスク 4 台からなる HDD アレイによる蓄積ユニットにより構成される。表 2 に示すように、Ladybug は図 4 のような横 768 ピクセル、縦 1024 ピクセルの高解像度な画像を 6 台のカメラから同期して取得し、上方を含む全方向の約 75% 以上を 15fps の動画として約 20 分間蓄積することが可能なシステムである。図 5 のように蓄積ユニットとカメラヘッドは光ファイバーケーブルにより接続されているため、蓄積ユニットはカメラヘッドから伝送される大量の画像情報を実時間で記録することができる。また、蓄積ユニットと PC の間は IEEE1394 ケーブルにより接続されており、6 個のカメラから同時に最大 5fps で PC へ画像を取り込むことができる。



図 3 Ladybug のカメラヘッド (左) と蓄積ユニット (右)

表 2 Ladybug の仕様

カメラ数	6 個
解像度	1024×768
視野	75%以上
フレームレート	15fps
蓄積可能時間	約 20 分



カメラ 0



カメラ 1



カメラ 2



カメラ 3



カメラ 4



カメラ 5

図 4 入力画像 (上方向 (右下) と水平方向 (その他))

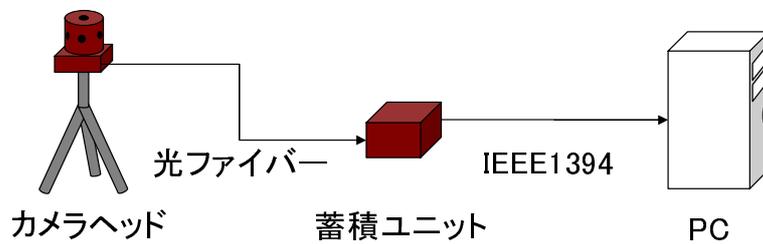


図 5 Ladybug による画像の取り込み

3.2 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.

3.2.1 カメラパラメータの定義

本論文では図 6 に示すようにカメラモデルとして一般的に用いられているピンホールカメラモデルを仮定し, カメラパラメータとして内部パラメータと外部パラメータを以下のように定義する. ただし, レンズ歪みとしては, Tsai のキャリブレーション手法 [32] で用いられている歪み中心に対して等方な非線形歪みを考慮する.

まずレンズ歪みについて述べる. 一般的なカメラのレンズを用いて撮影された画像には, レンズによる歪みが生じ画像が歪曲する. このようなレンズ歪みは Tsai の手法では, 図 6 に示すように歪んだ画像における座標 (X_d, Y_d) と歪みのない画像上での座標 (X_u, Y_u) を用いて以下のような関係式で表現される.

$$X_u = X_d(1 + \kappa_1 r^2 + \kappa_2 r^4) \quad (1)$$

$$Y_u = Y_d(1 + \kappa_1 r^2 + \kappa_2 r^4) \quad (2)$$

$$r = \sqrt{X_d^2 + Y_d^2} \quad (3)$$

一般的なカメラに使用されているレンズでは, このような歪みは小さく, 歪み係数としては κ_1 のみを使用するケースが多い. これに対して一般的な全方位型マルチカメラシステムでは, 1つのカメラで広範囲な撮影をするため, 広角なレンズが使用されていることが多く, そのために, 得られる画像では非常に歪みが大きい. 歪み係数は外部パラメータの推定精度に最も影響を与えやすい [33] ということから, 本研究では歪み係数 κ_2 についても考慮する.

次に、撮像面上に投影された画像の標本化モデルについて述べる。撮像面上に投影された光線は、 X 方向および Y 方向に d_x および d_y の間隔で配置された撮像素子により受光される。標本化される前の座標を (X_d, Y_d) と標本化後の座標 (X_f, Y_f) の間の変換は、以下の式により表現される。

$$X_d = \frac{d'_x}{s_x}(X_f - C_x), \quad Y_d = d_y(Y_f - C_y) \quad (4)$$

ただし、 C_x, C_y を歪みの中心、 s_x はアスペクト比であり、 d'_x は、 X 方向の撮像素子数 N_{cx} および走査線のサンプル数 N_{fx} を用いて補正 ($d'_x = d_x \frac{N_{cx}}{N_{fx}}$) された撮像素子の間隔である。

本論文では、各カメラの位置と姿勢は世界座標系でのカメラの位置 $T_c(t_x, t_y, t_z)$ と方向 $R_c(\alpha, \beta, \gamma)$ 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現する。

$$\begin{aligned} M_c &= \begin{bmatrix} \mathbf{R}_c & \mathbf{T}_c \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_1 c_3 + s_1 s_2 s_3 & s_1 c_2 & -c_1 s_3 + s_1 s_2 c_3 & t_x \\ -s_1 c_3 + c_1 s_2 s_3 & c_1 c_2 & s_1 s_3 + c_1 s_2 c_3 & t_y \\ c_2 s_3 & -s_2 & c_2 c_3 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (5)$$

ただし、

$$\begin{aligned} s_1 &= \sin \alpha, & s_2 &= \sin \beta, & s_3 &= \sin \gamma \\ c_1 &= \cos \alpha, & c_2 &= \cos \beta, & c_3 &= \cos \gamma \end{aligned} \quad (6)$$

また、図6に示すようにマーカの世界座標系における三次元位置を $[x_w, y_w, z_w]^T$ 、マーカのカメラ座標系における三次元位置を $[x_c, y_c, z_c]^T$ とすると、世界座標系とカメラ c のカメラ座標系の変換は以下のように記述される。

$$\begin{bmatrix} x_C \\ y_C \\ z_C \\ 1 \end{bmatrix} = M \begin{bmatrix} x_W \\ y_W \\ z_W \\ 1 \end{bmatrix} \quad (7)$$

以上をまとめると、幾何学的キャリブレーションでは、表3に示すように内部パラメータ（焦点距離 f 、歪み係数 (κ_1, κ_2) 、歪み中心 (C_x, C_y) 、アスペクト比 s_x ）と外部パラメータ（カメラ c の位置 T_c 、姿勢 R_c ）を推定しなければならない。

3.2.2 マーカの配置と検出

本項では、マーカの配置および検出手法について述べる。上述のカメラパラメータ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。本手法では、マーカとして図7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して求める。キャリブレーションボードは図8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。

次に、格子点の検出手法について述べる。画像上での格子点の位置は、以下に示す手順で格子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。

1. 大まかな格子点の位置を初期値として与える。
2. 図9に示すように、格子点位置の初期値を中心としたウィンドウを用いて、格子点を1つしか含まないようにウィンドウ内の領域を切り出す。

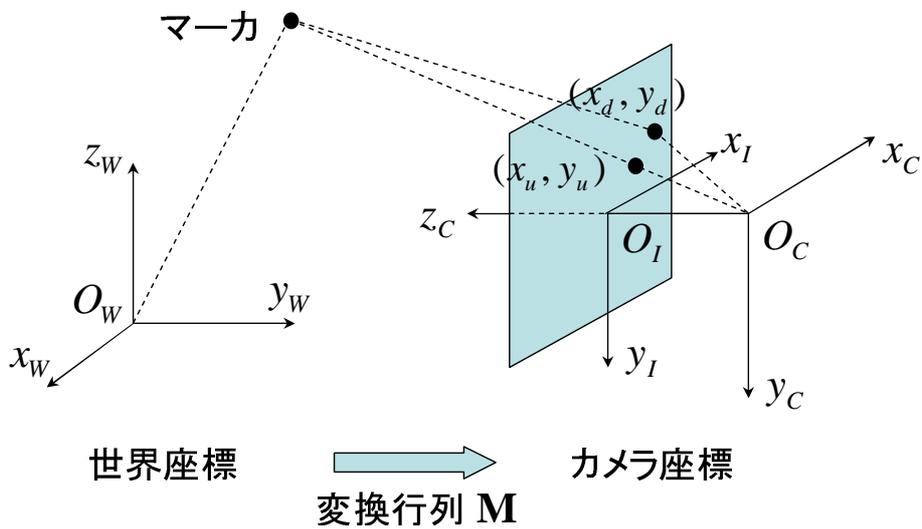


図 6 世界座標系とカメラ座標系の関係

表 3 本論分で用いるカメラの内部・外部パラメータ一覧

	パラメータ	表記
外部パラメータ	カメラの位置	$T_c(t_x, t_y, t_z)$
	カメラの姿勢	$R_c(\alpha, \beta, \gamma)$
内部パラメータ	焦点距離	f
	歪み中心	C_x, C_y
	歪み係数	κ_1, κ_2
	アスペクト比	s_x

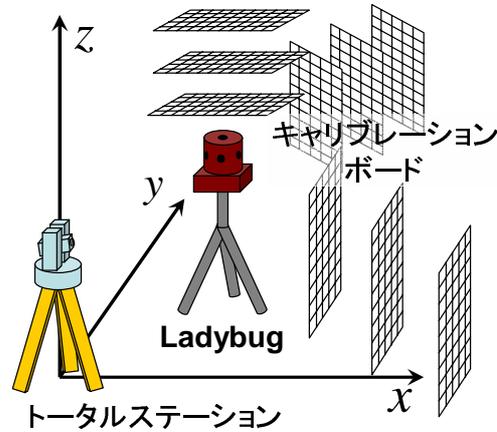
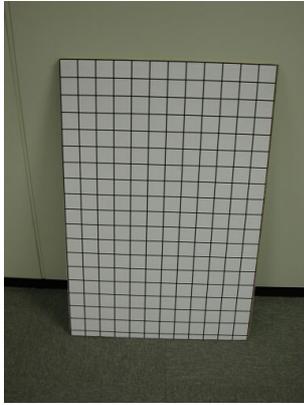


図 7 キャリブレーションボード 図 8 キャリブレーションボードの配置

3. ウィンドウ内を判別分析法 [34] を用いて 2 値化し, ウィンドウの縁を走査し, 格子線の領域を 4 つに領域に分割する.
4. 向かい合う 2 つの領域の重心をつなぐ直線の交点を求め, 格子点の位置とする.
5. 新たに求めた格子点の位置を初期値として, この手順を数回繰り返す.

始めに与える初期値が真の格子点の位置から離れている場合, 格子点の位置とウィンドウの縁が近くなり, 算出される直線の交点の精度は, ウィンドウの 2 値化の結果に敏感となる. そこで, 一端求めた格子点位置を新しい初期値として上述の手順を繰り返すことで検出精度の向上を図る.

3.2.3 カメラパラメータの推定

各カメラの内部および外部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[32] 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 [35]. 本手法では, 各カメラの外部パラメータは, 世界座

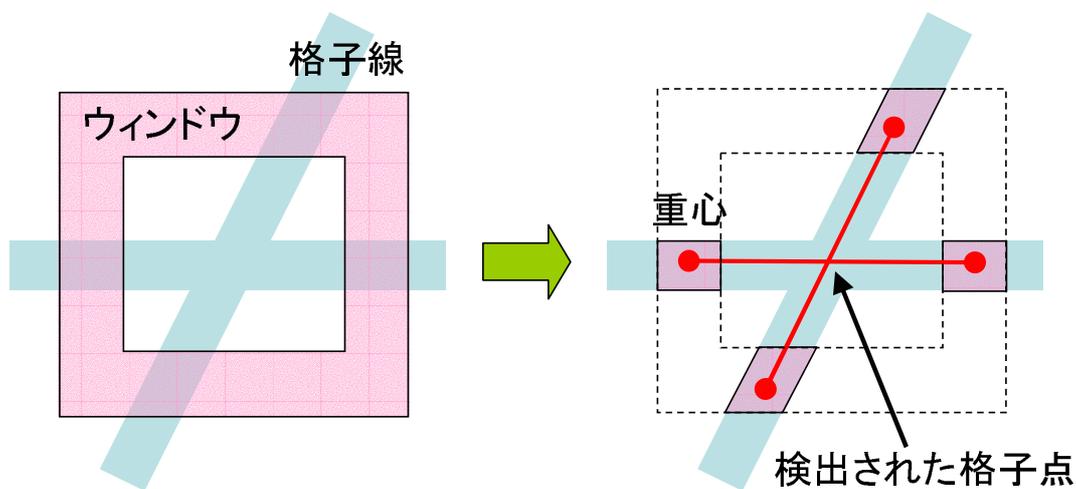


図 9 マーカの検出

標系におけるカメラ $c(c = 0, 1, \dots, 5)$ 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 として表わされる.

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 $c(c = 0, 1, \dots, 5)$ の入力画像上での位置 u_m とを対応づけ、出口の手法 [36] 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 $12(r_1, r_2, \dots, r_9, t_x, t_y, t_z)$ 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 $(\alpha, \beta, \gamma, t_x, t_y, t_z)$ 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 との差 (以降、再投影誤差) の二乗和で算出される評価関数

$$E_c = \sum_m (u_m - v_m)^2 \quad (8)$$

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.

3.3 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複数カメラ間の色調補正を行う。明度低下現象の補正では、広角レンズを考慮した $\cos^4 \theta$ 現象の補正を行い、色調補正では、ヒストグラムを用いたカメラ間の相対的な色調補正を行う。

3.3.1 明度低下現象の補正

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は、広角レンズ特有の $\cos^4 \theta$ 現象 [37] と望遠レンズに特有の口径蝕現象 [38] 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。

Horn[37] によれば、図 10 に示すように直径 l のレンズへ光軸に対して入射角 θ の角度をとる方向から入射する光の殆どがレンズの中心を通過することを仮定すれば、レンズへ入射する光の強度 I に対する撮像面上での受光強度 I' の減衰率は、レンズへ入射する光の入射角 θ のみに依存し、以下の式を満たす。

$$I' = \frac{\pi l^2 \cos^4 \theta}{f^2} I \quad (9)$$

本研究ではレンズ径の実測値 l と幾何学的キャリブレーションにおいて推定される焦点距離 f を使用し、上の式を用いることで $\cos^4 \theta$ 現象による明度低下を補正する。

3.3.2 カメラ間の相対的な色調補正

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' における撮像面上の放射照度 I_c と $I_{c'}$ の間にも線形的な関係

$$I_{c'} = a_c I_c + b_c \quad (10)$$

があると考えられる。これらの線形変換係数 a_c, b_c が各カメラごとに異なるためカメラ間での色調の違いが発生する。本研究では、色調のモデルとして上式で示

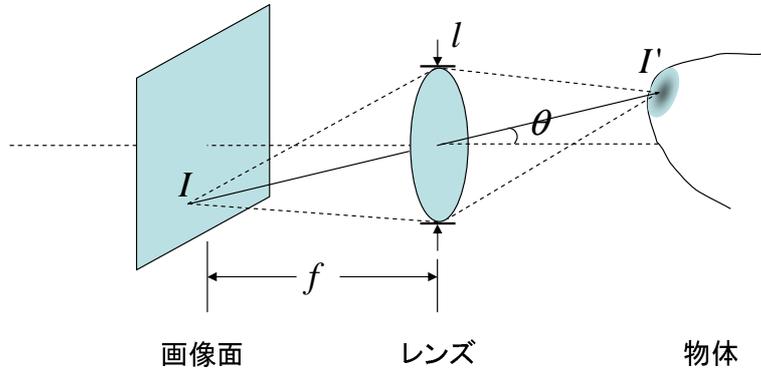


図 10 $\cos^4 \theta$ 現象のレンズモデル

されるモデルを用い, 各カメラ c の変換係数 a_c, b_c を推定するために以下のような手順をとる.

まず, 照明条件を固定した同じシーンを撮影し, RGB それぞれの輝度値についてヒストグラムを作成する. ただし, 上記の 2 つのパラメータ a_c, b_c を決定するためには, ヒストグラムに起伏がなければならない. そこで, ヒストグラムを作成するために, 目視により入力画像が類似していることを確認しながらカメラの方向を調整し, 様々な色を含む実環境の画像を取得する.

次に, カメラ c の画像と基準カメラ c_0 の画像の輝度 i に関するヒストグラム $h_c(i)$ と $h_{c_0}(i)$ が一致するようにパラメータ a_c, b_c を決定する. 上述の各 RGB 成分のヒストグラムがそれぞれ一致するように評価関数

$$e(a_c, b_c) = \sum_i \left\{ h_{c_0}(i) - \frac{1}{a_c} h_c \left(\frac{i - b_c}{a_c} \right) \right\}^2 \quad (11)$$

を定義し, これを最小化することで各カメラの a_c, b_c を推定する. なお, 輝度値は離散化されているため, ヒストグラムの値は隣接する輝度値での値で線形補間して評価関数の計算を行う. 最後に, 推定したパラメータから式 (10) に基づき, 入力画像の RGB の輝度値をそれぞれ独立に変換する.

3.4 実験結果

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回ずつ撮影した。図11にこの撮影により得られたキャリブレーションボードの画像を示す。水平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個ずつ検出し、合計561個のマーカの三次元位置と画像上の位置を取得した。上方向のカメラについては、キャリブレーションボードを固定する部分が写るため、各画像から170個ずつ検出し、合計510組のマーカの位置情報を取得した。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 XRで計測し、その他の格子点の三次元位置は線形補間により算出した。マーカの三次元位置と画像上の二次元位置との対応付けでは、マウスを使用して半自動で行うインターフェースを作成し、作業にかかる手間を削減した。図12に、検出したキャリブレーションボード上の格子点を×印で示す。

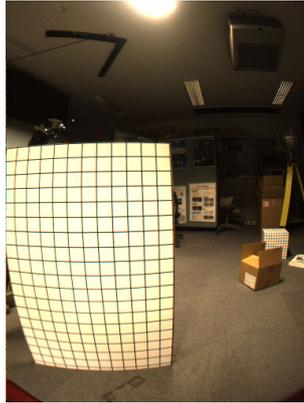
まず、表4に推定した内部パラメータを示す。カメラ0から4までが水平方向のカメラを示し、カメラ5が上向きカメラを示す。また、図13に推定した内部パラメータを用いた入力画像の歪み補正結果を示す。同図より格子模様が直線に戻っていることから、おおむね正しく内部パラメータが推定されていることが確認できる。

次に、推定した外部パラメータを用いて描画したカメラの位置と姿勢の関係を図14に示す。同図では、四角錐と直線でカメラの位置と姿勢の関係を表現している。四角錐の頂点と直線が交わる点はカメラの投影中心を、直線の方法はカメラの光軸を、四角錐の底面はカメラの光軸周りの回転を表す。また、表5に各カメラの再投影誤差の平均値、標準偏差、最大値を示す。図14に示した図形の位置および姿勢の関係がLadybugのカメラの配置に近く、表5に示した各カメラの再投影誤差の平均値が小さいことから、おおむね正しく外部パラメータが推定されていることが確認できる。

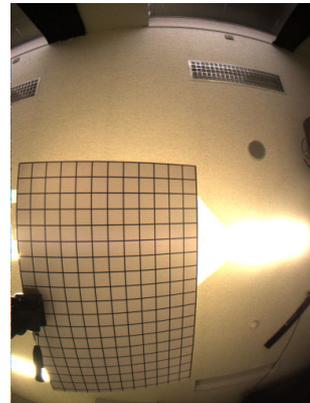
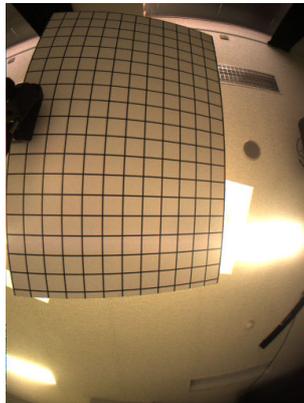
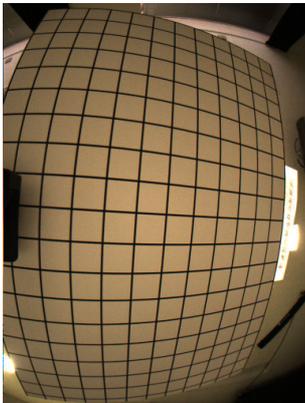
図15に明度低下現象の補正結果を示す。同図左のように補正前の画像では、周辺部分の輝度値が全体的に低下しているのに対して、同図右のように補正後の画像

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 16, 17, 18 は図 19 に示す同一シーンの色調補正前後での R, G, B 成分の正規化ヒストグラムの変化をそれぞれ示している。ただし、カメラ 0 の色調に他のカメラの色調を合わせており、補正後のヒストグラムは推定したパラメータを用いて補正前のヒストグラムを変換したものである。色調補正前では、ヒストグラムがカメラごとに大きく異なっており受光素子の感度がそれぞれ異なることが分かる。色調補正後には、それらのヒストグラムが比較的近づいていることから、カメラ間の色調がある程度補正されていることが確認できる。



水平方向のカメラ



上方向のカメラ

図 11 キャリブレーションに使用したマーカの画像例

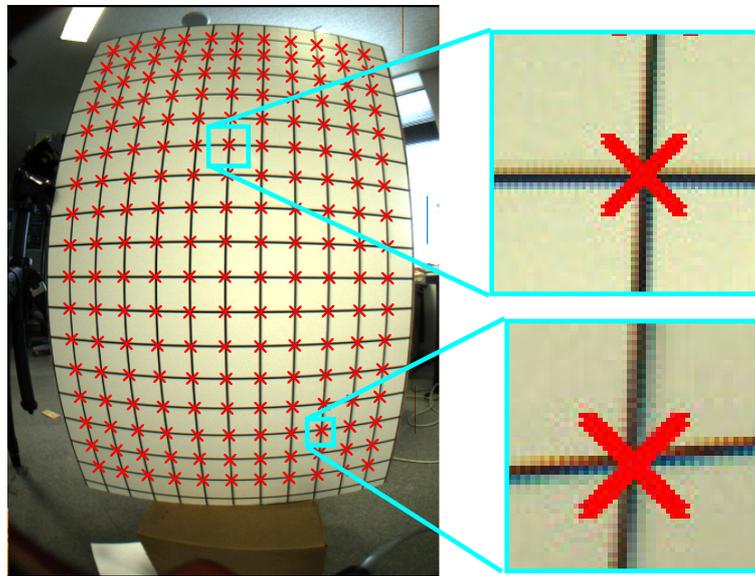
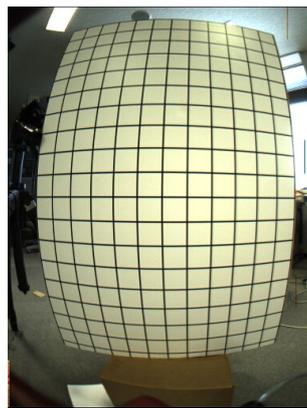


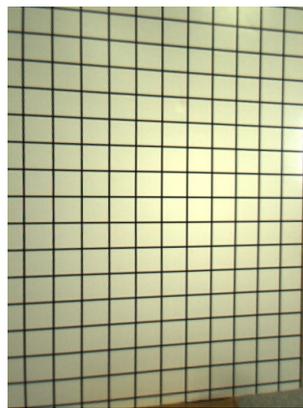
図 12 キャリブレーションボード上の格子点の検出例

表 4 推定された内部パラメータ

	カメラ 0	カメラ 1	カメラ 2
歪み係数 κ_1 [1/mm ²]	0.0326	0.0425	0.0354
歪み係数 κ_2 [1/mm ⁴]	0.0122	0.0108	0.0119
焦点距離 f [mm]	2.5517	2.5577	2.5470
アスペクト比 s_x	0.9988	0.9981	0.9972
歪み中心 C_x [pixel]	351.18	355.97	339.35
歪み中心 C_y [pixel]	507.33	492.47	477.06
	カメラ 3	カメラ 4	カメラ 5
歪み係数 κ_1 [1/mm ²]	0.0263	0.0308	0.0357
歪み係数 κ_2 [1/mm ⁴]	0.0131	0.0126	0.0109
焦点距離 f [mm]	2.5317	2.5221	2.5654
アスペクト比 s_x	0.9987	0.9989	0.9980
歪み中心 C_x [pixel]	353.87	344.04	365.02
歪み中心 C_y [pixel]	482.26	478.54	524.90



(a) 補正前



(b) 補正後

図 13 レンズ歪み補正結果

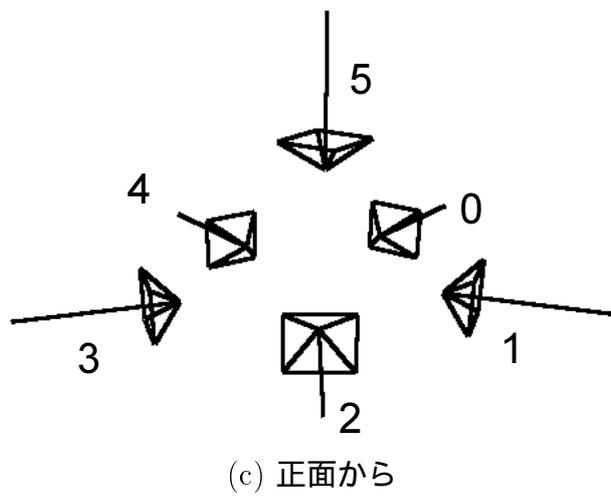
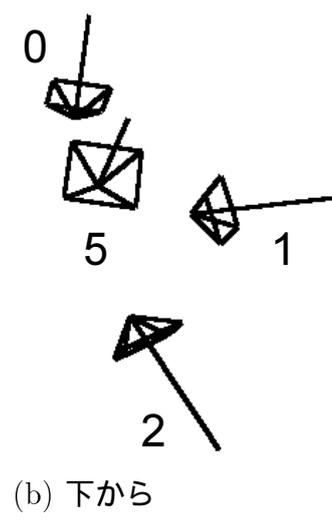
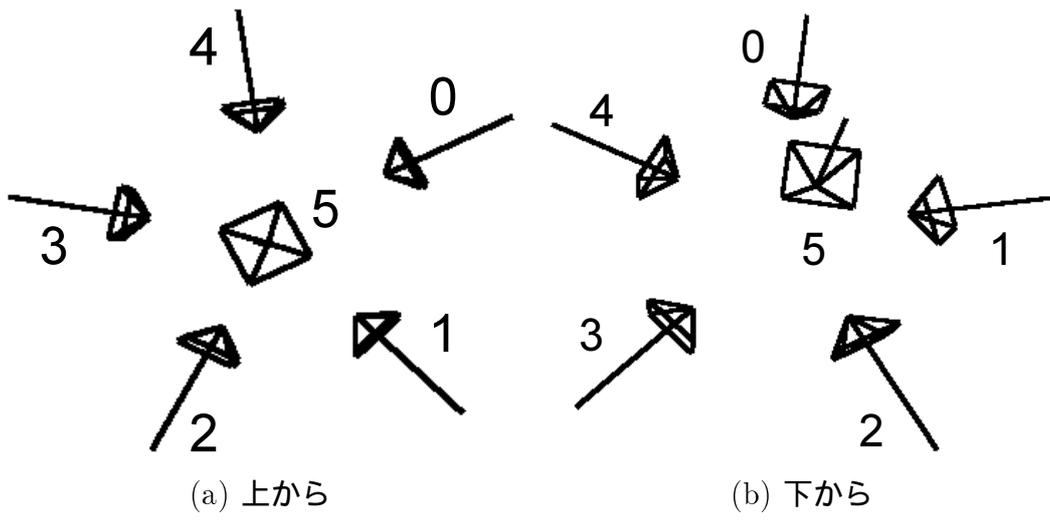


図 14 カメラの位置および姿勢の推定結果

表 5 各カメラの再投影誤差 [pixel]

	カメラ 0	カメラ 1	カメラ 2
平均	0.65	0.46	0.42
標準偏差	0.52	0.36	0.32
最大	2.66	1.94	2.39

	0	カメラ 1	カメラ 2
平均	0.56	0.55	0.42
標準偏差	0.48	0.41	0.28
最大	3.99	2.12	1.45

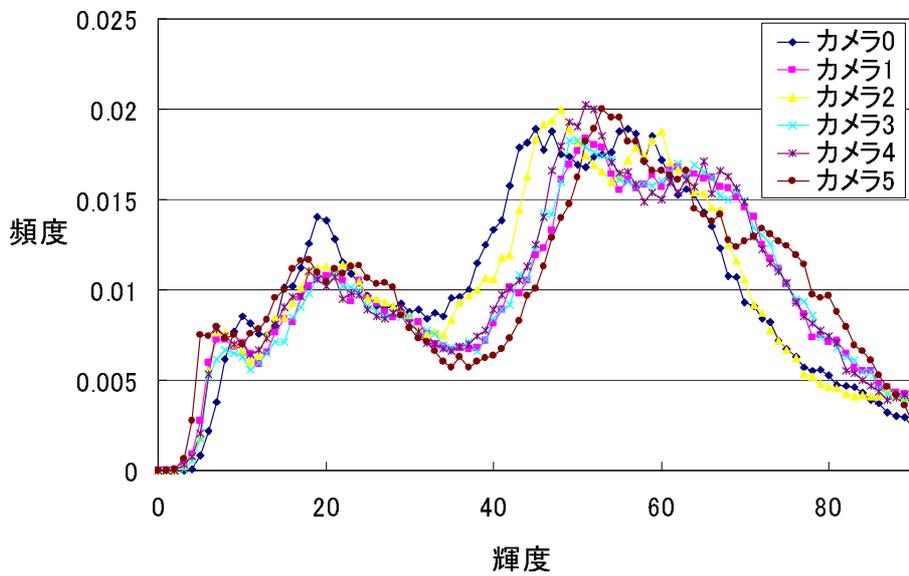


(a) 補正前

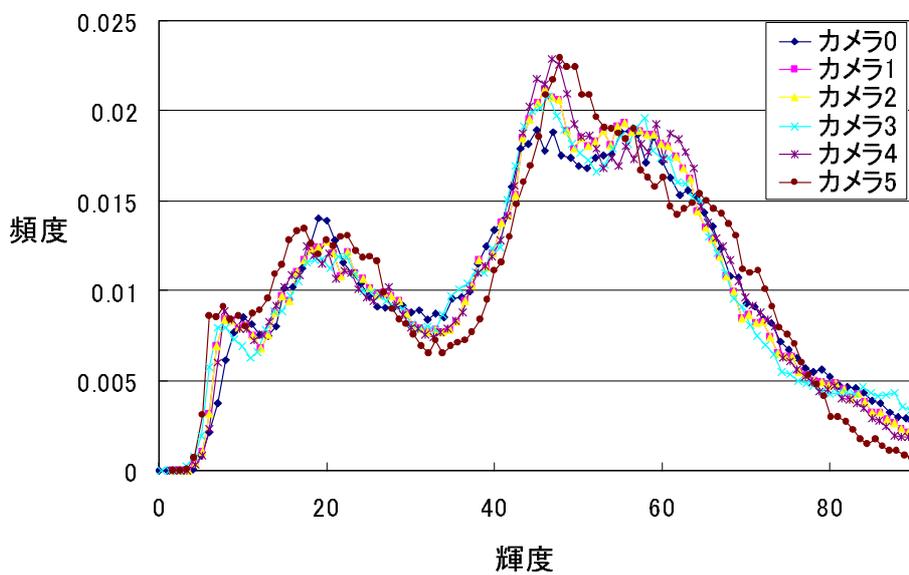


(b) 補正後

図 15 明度低下現象の補正結果

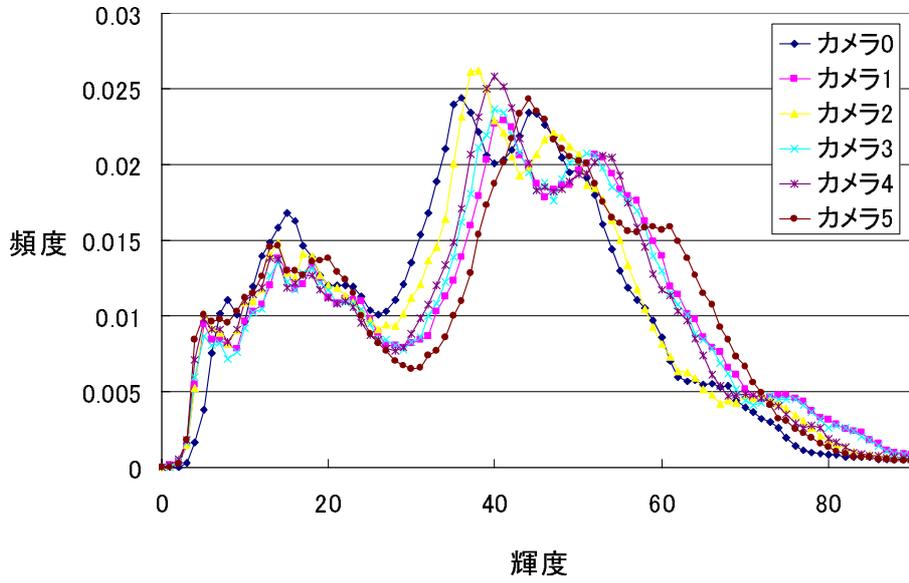


(a) 補正前

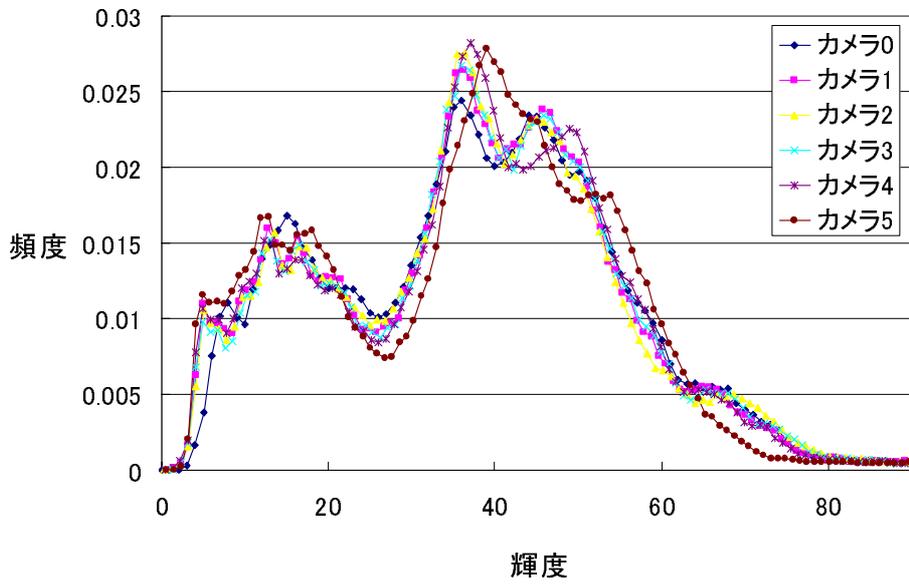


(b) 補正後

図 16 色調補正前後のヒストグラム (R 成分)

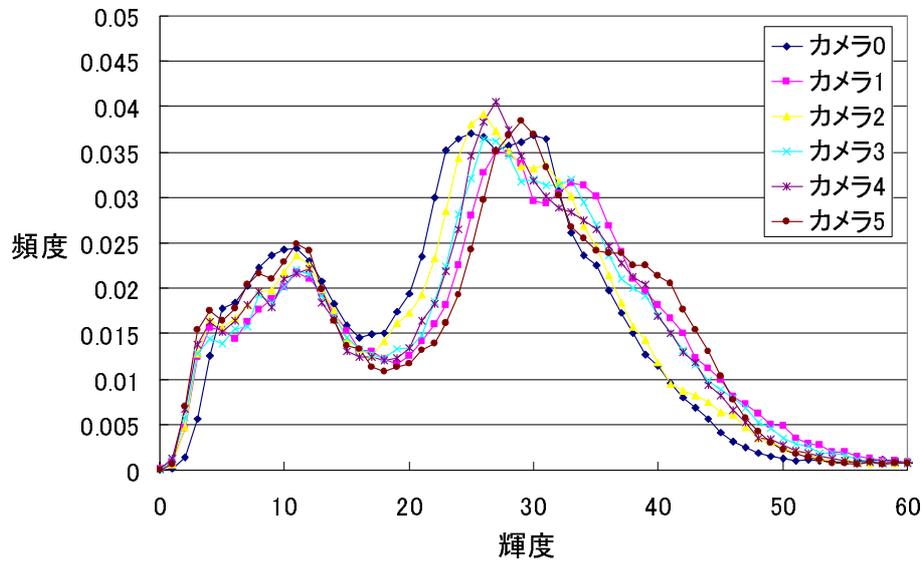


(a) 補正前

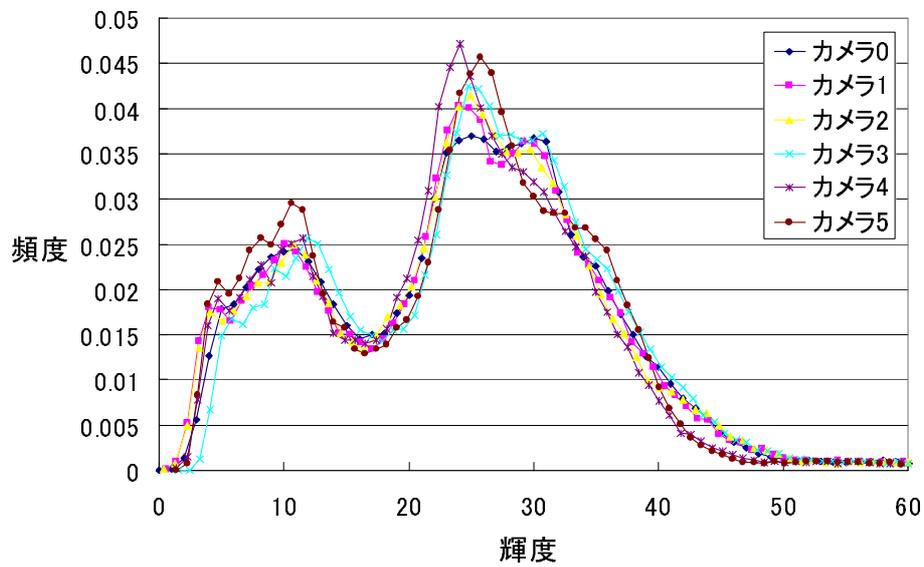


(b) 補正後

図 17 色調補正前後のヒストグラム (G 成分)



(a) 補正前



(b) 補正後

図 18 色調補正前後のヒストグラム (B 成分)



カメラ 0



カメラ 1



カメラ 2



カメラ 3



カメラ 4



カメラ 5

図 19 色調補正に用いた入力画像

4. 全天球パノラマ動画像の作成

本章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.

4.1 パノラマ画像の作成方法

本研究においては、図 20 に示すように、まず、前章で述べた入力画像の明度低下現象の補正およびカメラ間の色調補正を行い、次に仮想的な球面を設定しキャリブレーション結果を基に補正された入力画像を投影する。最後に、球面上に投影された画像を平面に展開し、パノラマ画像を生成する。以下では、仮想的な球面への画像の投影について述べる。

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。ここでは、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辺上の点までの距離を ξ_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。

$$I_S(s) = \frac{\sum_{c \in C(s)} \xi_c I_c(u_c)}{\sum_{c \in C(s)} \xi_c} \quad (12)$$

また、本研究では球面へ投影された画像は、球面上の点への方向を極座標 (θ, ϕ) で

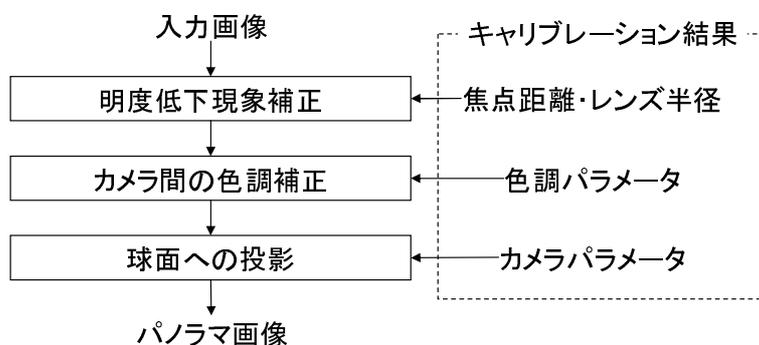


図 20 パノラマ画像の生成手順

表したとき、パノラマ画像の座標を図 21 に示すように、画像の縦が θ 、横が ϕ になるように対応付けて展開を行う。

4.2 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 22 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。

$$\lambda > \frac{d}{2 \tan \frac{\pi}{N}} \quad (13)$$

例えば、Ladybug の場合、推定された外部パラメータから隣り合う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 ピクセル

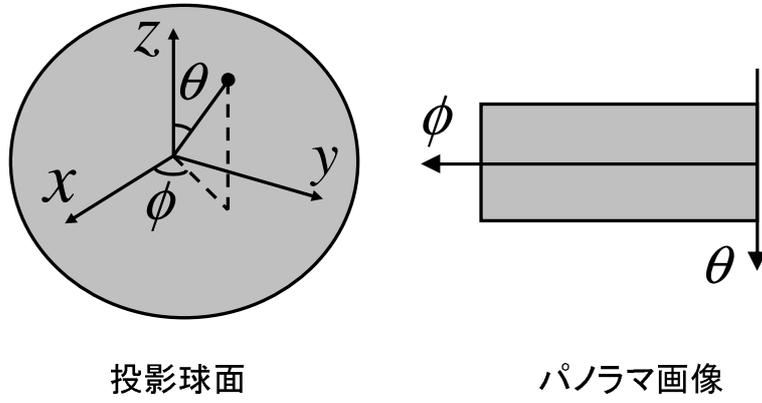


図 21 投影球面からパノラマ画像への展開

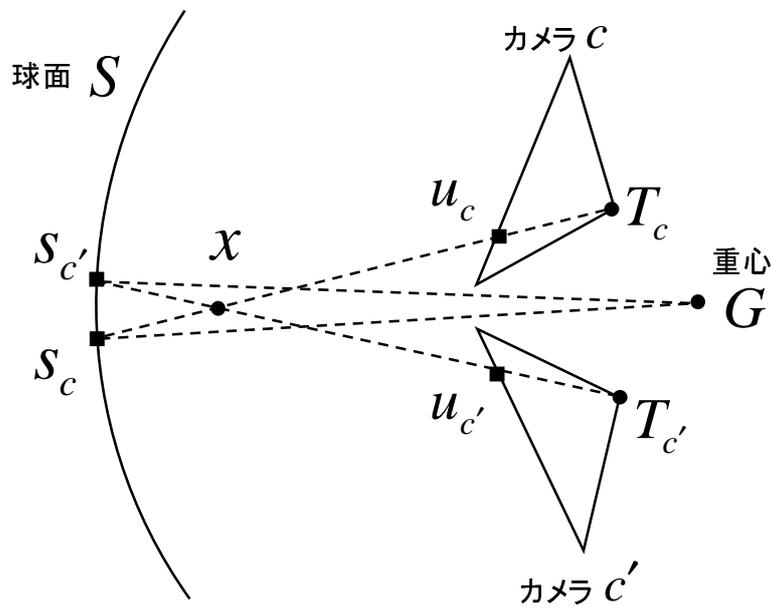


図 22 視差の影響

ルで構成されていると仮定すると、視差の影響が 1 ピクセル以内に収まる距離 λ は、約 24m である。

4.3 実験結果

4.1 節で述べた手法を用いて、実際に 6 つの動画像から全天球パノラマ動画像を作成した。以下では、パノラマ動画像の作成、定量的な評価実験について述べる。

4.3.1 全天球パノラマ動画像の生成

3 章で述べたキャリブレーション結果を用いて 4.1 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 23 に示すように Ladybug を自動車の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 24 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を図 25 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 25 の下部の黒い部分は、入力画像の存在しない部分である。ただし、本論文では、パノラマ画像の最大の水平方向の解像度は、各カメラの水平方向の解像度である 768pixel を 5 倍した 3840pixel とした。また、パノラマ画像の垂直方向の解像度は、水平方向の半分の 1920pixel とした。図 25 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 26 に示すブレンドなしのパノラマ画像を同様の手法で作成し、画像のつなぎ目に関して主観的な比較を行った。同図と図 25 の比較により、図 27 に示すような部分において、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像素子の輝度値の飽和によって画像間のつなぎ目が容易に確認できる。これに対して、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ぎ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ぎ目を知覚することは出来なかった。



図 23 自動車による屋外の撮影

4.3.2 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 28 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、4.2 節で述べた図 22 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、4.2 節で述べたように視差の影響が 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。



カメラ 0



カメラ 1



カメラ 2



カメラ 3



カメラ 4



カメラ 5

図 24 入力画像 (奈良 鴻池周辺: 上方向 (右下) と水平方向 (その他))



図 25 全天球動画像の1フレーム(ブレンド処理あり)



図 26 全天球動画像の1フレーム(ブレンド処理なし)



(a) ブレンドなし

(b) ブレンドあり

図 27 ブレンド処理によるスムージングの効果



図 28 隣り合うカメラ画像の共通領域

表6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5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のように大きな誤差値を持つ部分も存在した。

そこで、定義した誤差 ϵ の画像上における位置の分布について調べた。表6は、上述の実験において計測した全てのデータについて、誤差 ϵ と入力画像上でのマーカの歪み中心からの距離との関係を示している。入力画像の解像度は768×1024であることから、歪み中心からの距離が640pixel付近のデータは、画像の角に近い部分に撮影されたマーカによるものである。同表から、画像の端の付近では、画像のつなぎ合わせに関する誤差が急激に大きくなる傾向にあることがわかる。しかし、このような入力画像の端の部分での大きな誤差は、パノラマ画像の生成時にはブレンドの比率が小さくなり、生成した画像上における視覚的な影響は小さい。

表6 画像のつなぎ合わせに関する誤差 ϵ [rad]

カメラ番号	0-1	1-2	2-3	3-4	4-0
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	5-0	5-1	5-2	5-3	5-4
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

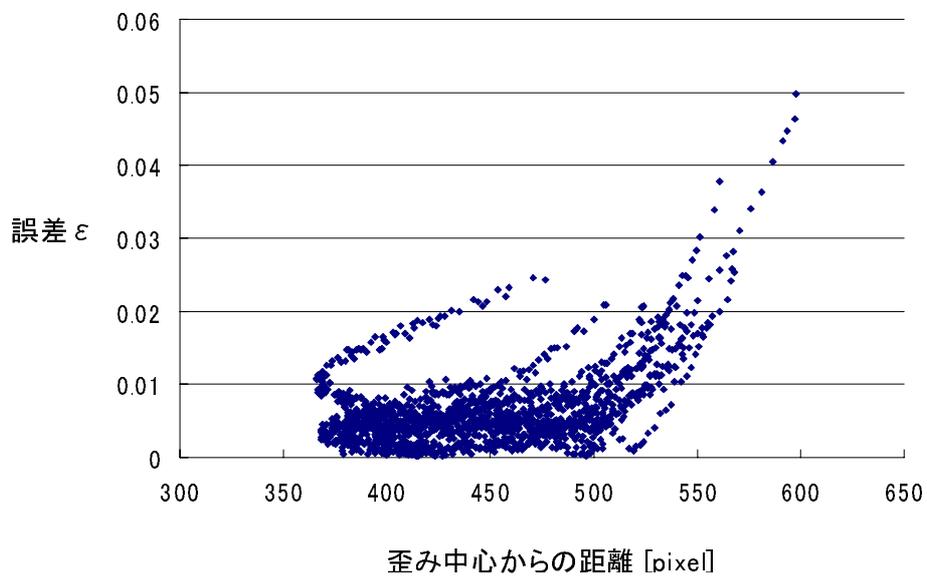


図 29 画像のつなぎ合わせに関する誤差 ϵ の分布

5. テレプレゼンスシステムの試作

生成した高解像度パノラマ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する。また、試作したシステムにより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。本研究では臨場感の高いテレプレゼンスの例として3種類の没入型のディスプレイを用いた実装を行う。以下に各ディスプレイの特徴を示す。

ヘッドマウントディスプレイ： 図30に示すように利用者の頭部に固定し両眼に映像を提示する装置。視界全てを覆いコントラストの高い映像の提示が可能であるため没入感が高い。姿勢センサと組み合わせることで頭部の方向に応じた映像を提示ができ、全方向の見回しが可能である。しかし、一般にヘッドマウントディスプレイは、後述の他の没入型ディスプレイに比べて視野が狭いという問題がある。

球面ディスプレイ： 図31に示すようにディスプレイが利用者のほぼ全ての視野を覆い臨場感の高い映像提示が可能なディスプレイである。コントローラなどと組み合わせることにより利用者が要求する方向の映像を提示することができ、全方向の見回しが可能となる。

大型円筒ディスプレイ： 図33に示すように利用者の立ち位置を中心として利用者を囲むように円筒面状のスクリーンが配置されている。スクリーンには常に全方向の映像を提示しておくことで、利用者は、見たい方向を自由に見回すことができる。水平方向に関しては広い画角で映像を提示することが可能であるが、上下方向の映像提示ができないという問題がある。

以上のように提示機器はそれぞれの長所、短所があるため目的に合わせて選択を行なう必要がある。

本章では、3種類のディスプレイを用いたテレプレゼンスシステムの試作例として、5.1節ではヘッドマウントディスプレイを用いたシステム、5.2節では球面ディスプレイを用いたシステム、5.3節では円筒面ディスプレイを用いたシステムの試作結果を示す。

5.1 ヘッドマウントディスプレイを用いたシステム

ヘッドマウントディスプレイを用いたシステムでは、表7に示すようにディスプレイの他に、姿勢センサと計算機の2つの装置により構成されている。本システムでは、生成したパノラマ画像列は算機内のハードディスク上に蓄積されており、計算機は頭部に取り付けられた姿勢センサからの頭部の姿勢情報を基に各パノラマ画像の一部分を透視投影画像に変換しディスプレイに表示している。なお、ヘッドマウントディスプレイの解像度は800×600であり、ハードディスクからの画像の読み出しを高速に行うためパノラマ画像は2048×1024の解像度に縮小しJPEG形式で蓄積されている。

本システムではLadybugの撮影時のフレームレートである15fpsでの画像の更新が可能であった。頭部の方向に応じた方向の映像が提示されるため、システムの操作性に関して殆ど違和感なく全方位を見回すことができた。また、上方向の映像があるため広い範囲で見回すことができ、視界の制限による束縛感少なかった。見回しに関する遅延は、多少知覚することができたが、これは、姿勢センサからの姿勢情報の更新に遅延があるためと考えられる。

5.2 球面ディスプレイを用いたシステム

球面ディスプレイを用いたシステムでは、表8に示すようにディスプレイの他に、コントローラと計算機の2つの装置により構成されている。本システムでは、前節のヘッドマウントディスプレイを用いたテレプレゼンスシステムと同様に、

表7 ヘッドマウントディスプレイを用いたテレプレゼンスシステムの機器構成

ディスプレイ	Olympus 社製 Eye-Trek
姿勢センサ	Polhemus 社製 Fastrack
計算機	CPU: Intel 社製 Pentium4 1.7GHz, メモリ: 1GB
グラフィクスカード	Nvidia 社製 Geforce4



図 30 ヘッドマウントディスプレイを用いたテレプレゼンスシステムの外観

生成したパノラマ画像列は計算機内のハードディスク上に蓄積されており、計算機はユーザが操作するコントローラからの方位情報を基に各パノラマ画像を球面透視投影画像に変換しディスプレイに表示している。なお、球面ディスプレイの解像度は 1024×768 であり、パノラマ画像は前節と同じ形式で蓄積されている。

本システムにおいても前節と同様に 15fps での画像の更新が可能であった。視界の殆どが映像で覆われているため、臨場感の高いテレプレゼンスが実現できた。また、見回しによる遅延は、殆ど知覚することができなかった。しかし、ゲームコントローラを用いた視線方向の操作において、利用者がその操作を間違えることがあった。より臨場感のあるテレプレゼンスを実現するためには、利用者の操作性について検討が必要である。

5.3 円筒ディスプレイを用いたシステム

円筒ディスプレイを用いたシステムでは、表 8 に示すように大型の円筒面ディスプレイ CYLINDRA と計算機の 2 つの装置により構成されている。本システムでは、表示デバイスの特性を考慮し、図 32 に示すような水平方向の全周パノラマ画像のみの形式で円筒面への透視投影画像を計算機内のハードディスク上に蓄積

表 8 球面ディスプレイを用いたテレプレゼンスシステムの機器構成

ディスプレイ	Elumens 社製 VisionStaion
コントローラ	Microsoft 社製 SideWinder Game Pad Pro
計算機	CPU: Intel 社製 Pentium4 1.7GHz, メモリ: 1GB
グラフィクスカード	Nvidia 社製 Geforce4



図 31 球面ディスプレイを用いたテレプレゼンスシステムの外観

している。なお、円筒ディスプレイの解像度は 6144×768 程度であり、パノラマ画像の解像度は 2712×512 である。パノラマ画像の解像度は、本システムにおいて 15fps での描画が可能ないように決定した。

本システムでは、他の提案システムと比較して水平方向に関して高い解像度の映像を提示することができ、鮮明な映像のテレプレゼンスが実現できた。複数の利用者が同時に同じ映像を見ることができ、かつ互いの姿を見ることができるので、仮想旅行のような応用に最も近い形でのテレプレゼンスであった。しかし、視界に現実環境が入ることが多く、他の提案システムに比べて没入感に欠けていた。

表 9 円筒ディスプレイを用いたテレプレゼンスシステムの機器構成

ディスプレイ	CYLINDRA
計算機	SGI 社製 Onyx2: CPU:MIPS R10000 250MHz × 8
グラフィクスカード	InfiniteReality2 × 2



図 32 円筒ディスプレイを用いたテレプレゼンスシステムの入力画像 (飛騨高山)



図 33 円筒ディスプレイを用いたテレプレゼンスシステムの外観

6. まとめ

本論文では、臨場感の高いテレプレゼンスの実現を目指して、全方位型マルチカメラシステム Ladybug により得られる複数のカメラ画像から高解像度な全天球パノラマ動画画像を生成する手法を提案した。提案手法では、まず全方位型マルチカメラシステム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う。幾何学的キャリブレーションでは、キャリブレーション精度を向上させるために各カメラに対してマーカを空間的に広く配置し、内部パラメータの推定と各カメラの位置及び姿勢を、設計パラメータを用いずに推定する。光学的キャリブレーションでは、各カメラの明度低下現象およびカメラ間の相対的な色調補正を行う。次に、求めたキャリブレーション結果を基に、撮影対象がシステムから遠方にあることを仮定して全天球パノラマ動画画像を作成する。これにより、高解像度かつ高視野のパノラマ動画画像を生成することができる。

実験では全方位型マルチカメラシステムとして Ladybug の幾何学的および光学的キャリブレーションを行い、キャリブレーション結果を確認した。また、実際に Ladybug を用いた撮影により得られた各カメラの画像から、提案手法を用いて高解像度なパノラマ画像の作成を行った。評価実験では、生成したパノラマ画像における画像のつなぎ合わせ精度を定量的に評価し、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。高解像度なパノラマ画像の応用例として 3 種類の特徴の異なるディスプレイを用いた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を用いることで臨場感の高いテレプレゼンスが実現可能であること確認した。

本研究により得られた全天球パノラマ動画画像や全方位型マルチカメラシステムのキャリブレーション結果は、テレプレゼンスへの応用だけでなく、従来の単眼のカメラを用いて屋外環境の画像から仮想環境の構築を行う様々な手法において、画像の取得効率の向上が期待できる。しかし、まだ課題も残されており。例えば、キャリブレーションの自動化が挙げられる。本研究においては、全方位型マルチカメラシステムをキャリブレーションする際の手軽さについては考慮しておらず、マーカの配置や計測では人手による作業が多い。通常、キャリブレーションは何度も行う必要はないが、将来的に全方位型マルチカメラシステムが普及した時は、

キャリブレーションの自動化は重要である。今後は、本研究により得られた高解像度パノラマ動画像およびそのキャリブレーション結果を用いた屋外環境の三次元形状の復元やウォークスルー環境の構築を目指したいが、同時に、キャリブレーション手法の自動化についても検討していきたい。

謝辞

まず、本研究テーマを取り上げてくださり、全過程を通して細やかな御指導を賜りました、視覚情報メディア講座 横矢直和教授に心より感謝致します。特に、研究ミーティングや発表論文の添削においては、数々の印象に残る御言葉を頂きました。また、横矢教授は、国内外において研究成果を発表する多くの機会を与えて下さりました。

本研究の遂行にあたり、御助言、御鞭撻頂いたロボティクス講座 小笠原 司教授に厚く御礼申し上げます。

そして、本研究の全過程を通して温かい御指導をして頂いた視覚情報メディア講座 山澤 一誠助教授に深く感謝申し上げます。

物心両面において常に温かい御支援を頂いた視覚情報メディア講座 神原 誠之助手に深く御礼申し上げます。特に、研究ミーティングや学外発表の練習では、多くの御指導、御助言頂きました。

研究活動の全過程を通して多くの御助言、御指導賜りました視覚情報メディア講座 佐藤 智和氏に心より感謝致します。特に、本研究のテーマの設定から本論文の執筆、その他多くの発表論文の添削、発表練習に至るまで細やかな御指導を頂きました。

日々、研究室での研究活動を支えてくださった視覚情報メディア講座 事務補佐員 北川 知代女史に心より感謝申し上げます。

最後に、研究活動において常に温かい御助言を頂いた視覚情報メディア講座の諸氏に深く感謝いたします。

参考文献

- [1] 廣瀬 通孝：“電脳都市の誕生”，PHP 研究所 (1996).
- [2] A. Lippman: “Movie-Map: An Application of the Optical Videodisc to Computer Graphics”, Proc. SIGGRAPH’80, pp. 39–46 (1993).
- [3] 宮内 信仁, 佐伯 俊彰, 福岡 久雄, 下間芳樹：“実画像テクスチャによる仮想空間表示方法を用いた3次元仮想都市の構築”, バーチャルリアリティ学会論文誌, 4, 2, pp. 367–376 (1999).
- [4] 清水 浩行, C. R. Gunadi, 山本 智幸, 相澤 清晴：“画像, レンジデータと航空測量データを利用した広域実空間のモデル化”, 画像の認識・理解シンポジウム (MIRU 2002) 講演論文集, 2, 11, pp. 177–182 (2002).
- [5] K. Yamazawa, H. Takemura and N. Yokoya: “Telepresence System with an Omnidirectional HD Camera”, Proc. 5th Asian Conf. on Computer Vision (ACCV2002), 2, pp. 535–538 (2002).
- [6] S. Morita, K. Yamazawa and N. Yokoya: “Internet Telepresence by Real-time View-dependent Image Generation with Omnidirectional Video Camera”, Proc. SPIE Electronic Imaging, 5018, pp. 51–60 (2003).
- [7] 池田 聖, 佐藤 智和, 横矢直和：“全方位型マルチカメラシステムのキャリブレーション”, 情報科学技術フォーラム (FIT) 情報技術レターズ, 1, pp. 149–150 (2002).
- [8] 池田 聖, 佐藤 智和, 横矢直和：“全方位型マルチカメラシステムのキャリブレーション全天球動画像の生成”, 電気関係学会関西支部連合大会講演論文集, G13-27, p. G329 (2002).
- [9] 池田 聖, 佐藤 智和, 横矢直和：“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”, 電子情報通信学会 技術研究報告, PRMU2002-154, pp. 49–54 (2002).

- [10] D. Coombs and C. Brown: “Real-time Smooth Pursuit Tracking for a Moving Binocular Robot”, Proc. Computer Vision and Pattern Recognition, pp. 23–29 (1992).
- [11] T. Uhlin, P. Nardlund, A. Maki and J. O. Eklundh: “Towards an Active Visual Observer”, Proc. Int. Conf. on Computer Vision, pp. 679–686 (1995).
- [12] S. Rougeaux and Y. Kuniyoshi: “Velocity and Disparity Cues for Robust Real-time Binocular Tracking”, Proc. Computer Vision and Pattern Recognition (CVPR’97), pp. 1–6 (1997).
- [13] M. Hirose, K. Yokoyama and S. Sato: “Transmission of Realistic Sensation: Development of a Virtual Dome”, Proc. Virtual Reality Annual International Symposium, pp. 125–131 (1993).
- [14] Y. Onoe, K. Yamazawa and N. Yokoya and H. Takemura: “Telepresence by real-time view-dependent image generation from omnidirectional video streams”, Computer Vision and Image Understanding, **71**, 2, pp. 154–165 (1998).
- [15] Point Grey Research, Inc.: “Ladybug Omnidirectional Camera System User Guide version 1.0” (2002).
- [16] Z. L. Cao, S. J. Oh and E. L. Hall: “Dynamic Omnidirectional Vision for Mobile Robots”, J. Robotic Systems, **3**, 1, pp. 5–17 (1986).
- [17] S. J. Oh and E. L. Hall: “Guidance of a Mobile Robot Using an Omnidirectional Vision Navigation System”, Proc. Mobile Robots IISPIE 852, pp. 288–300 (1987).
- [18] Movingeye Inc.: “Movingeye”,
<http://www.movingeye.co.jp/> (1999).
- [19] J. Hong, X. Tan, B. Pinette, R. Weiss and E. M. Riseman: “Image-based homing”, Proc. Int. Conf. on Robotics and Automation, pp. 620–625 (1991).

- [20] 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, pp. 909–914 (1991).
- [21] 山澤 一誠, 八木 康史, 谷内田正彦: “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, J79-D-II, 5, pp. 698–707 (1996).
- [22] S. K. Nayar: “Catadioptric Omnidirectional Camera”, Proc. Computer Vision and Pattern Recognition, pp. 482–488 (1997).
- [23] T. Kawanishi, K. Yamazawa, H. Takemura and N. Yokoya: “Generation of Hight-resolution Stereo Panoramic Images by Omnidirectional Imageing Sensor Using Hexagonal Pyramidal Mirrors”, Proc. 14th Int. Conf. on Pattern Recognition, 1, pp. 445–489 (1998).
- [24] ソニー株式会社: “FC2”,
<http://www.sony.co.jp/products/fourthview> (2000).
- [25] J. Shimamura, H. Takemura, N. Yokoya and K. Yamazawa: “Construction of an Immersive Mixed Environment Using an Omnidirectional Stereo Image Sensor”, Proc. IEEE Workshop on Omnidirectional Vision, pp. 62–69 (2000).
- [26] 戎野 聡一, 山澤 一誠, 竹村 治雄, 横矢直和: “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 (2000).
- [27] H. Tanahashi, K. Yamamoto, C. Wang and Y. Niwa: “Development of a Stereo Omnidirectional Imageing System (SOS)”, Proc. IEEE Int. Conf. on Industrial Electronics, Control and Instrumentation (IECON2000), pp. 289–294 (2000).

- [28] 戎野 聡一, 山澤 一誠, 竹村 治雄, 横矢直和: “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”, 日本バーチャルリアリティ学会第4回大会論文集, pp. 211–212 (1999).
- [29] 棚橋 英樹, 佐藤 淳, 王 彩華, 丹羽 義典, 山本和彦: “全方向ステレオシステム(SOS)のキャリブレーション手法”, 映像メディア学会誌, **56**, 4, pp. 603–610 (2002).
- [30] T. Yamada, M. Hirose and Y. Iida: “Development of Complete Immersive Display”, Proc. 4th Int. Conf. on Virtual Systems and Multimedia (VSMM’98), **2**, pp. 522–527 (1998).
- [31] 佐藤智和: “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).
- [32] 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, **RA-3**, 4, pp. 323–344 (1987).
- [33] J. Z. C. Lai: “On the Sensitivity of Camera Calibration”, Image and Vision Computing, **11**, 10, pp. 656–664 (1993).
- [34] 大津 展之: “判別および最小2乗基準に基づく自動しきい値選定法”, 電子情報通信学会論文誌, **4**, J63-D, pp. 349–356 (1980).
- [35] T. Sato, M. Kanbara, N. Yokoya and H. Takemura: “Dense 3D Reconstruction of an Outdoor Scene by Hundreds-Baseline Stereo Using a Hand-Held Video Camera”, Int. Journal of Computer Vision, **47**, 1-3, pp. 110–129 (2002).
- [36] 出口 光一郎: “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, **90**, 20, pp. 41–50 (1990).

- [37] B. K. P. Horn: “Robot Vision”, chapter 10, pp. 206–209, Mit Press (1986).
- [38] N. Asada, A. Amano and M. Baba: “Photometric Calibration of Zoom Lens Systems”, Proc. Int. Conf. Pattern Recognition, **A**, pp. 186–190 (1996).