

NAIST-IS-MT9651035

修士論文

六角錐ミラーを用いた全方位画像センサによる 全周ステレオパノラマ画像の作成

川西 隆仁

1998年2月13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おいて
修士(工学) 授与の要件として提出された修士論文である。

提出者： 川西 隆仁

指導教官： 横矢 直和 教授
千原 國宏 教授
竹村 治雄 助教授

六角錐ミラーを用いた全方位画像センサによる 全周ステレオパノラマ画像の作成*

川西 隆仁

内容梗概

本研究の目的は複合環境構築のための撮像装置の開発である。複合現実感のための撮像装置には、(1)1点透視投影、(2)広視野、(3)ビデオレート、(4)高分解能、(5)ステレオという5つの特性を満たすことが必要とされる。そこで本研究では六角錐ミラーと複数のカメラを用いて1点透視の光学特性を持つビデオレートで撮像可能な高分解能全方位ステレオ画像センサを試作した。

試作センサは6台のカメラにより六角錐ミラーの各面に映る像を撮像し、それらを張り合わせることで一視点からの全方位パノラマ画像を取得する。さらに、この構造を上下対称に2つ設置することによりステレオ撮像を可能にしている。本研究では、Tsaiのキャリブレーション法により、画像中の1次樽型歪みを補正し、次にそれぞれの撮像時刻の画像組に対して、全周ステレオパノラマ画像を生成した。さらに、複合現実感への応用に必要とされる全周パノラマ距離画像、任意方向の透視投影画像を得られたステレオパノラマ画像列から生成した。室内環境における実験を行い、本手法の有効性を確認し得られた全周ステレオパノラマ画像、全周パノラマ距離画像、任意の視線方向の平面透視投影画像を作成した例を示す。

キーワード

複合現実感, 全方位画像, ステレオパノラマ画像, カメラキャリブレーション, 奥行き抽出

*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-MT9651035, 1998年2月13日.

Generation of High-resolution Stereo Panoramic Images by Omnidirectional Imaging Sensor Using Hexagonal Pyramidal Mirrors*

Takahito Kawanishi

Abstract

The objective of this research is the development of the sensor for the Mixed Reality, which can take high-resolution stereo omnidirectional images at video-rate. The sensor system takes an omnidirectional view by a component constructed of six cameras and a hexagonal pyramidal mirror, and acquires stereoscopic views by symmetrically connecting two sensor components. This paper describes a method of generating stereo panoramic images by using our sensor. First, the sensor system is calibrated; that is, twelve cameras are correctly aligned with pyramidal mirrors and radial distortion of each camera image is restored by using the Tsai's method. Stereo panoramic images are then computed by registering the camera images captured at the same time. Also described is the computation of common perspective image and omnidirectional depth maps.

Keywords:

mixed reality, omnidirectional image generation, stereo panoramic image, camera calibration, depth extraction

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9651035, February 13, 1998.

目次

1. はじめに	1
2. 関連研究と本研究の方針	4
2.1 複合現実感の従来研究	4
2.2 全方位画像取得に関する従来研究	7
2.3 本研究の方針	10
3. 六角錐ミラーを用いた全方位ステレオ画像センサ	14
3.1 全方位画像センサの設計と試作	14
3.2 全方位ステレオ画像センサのキャリブレーション	18
3.2.1 樽型歪み補正	18
3.2.2 レンズ中心の位置合わせ	19
3.3 撮像システム	24
4. 全周ステレオパノラマ画像の生成	25
4.1 画像入力システム	25
4.2 全周ステレオパノラマ画像の生成	27
4.3 任意方向の透視投影画像の生成	28
4.4 全周パノラマ距離画像の生成	29
5. 実験	33
5.1 実験条件	33
5.2 実験結果	33
6. 考察	38
7. むすび	40
謝辞	41
参考文献	42

付録	I
A. Tsai のキャリブレーション法	I

図目次

2.1	全方位画像センサの視野領域	9
3.1	高分解能全方位ステレオ画像センサの設計	16
3.2	試作センサの外観	17
3.3	Tsai 法のための入力格子点	19
3.4	撮像画像 (左) と歪み除去画像 (右)	19
3.5	稜線上で重なるレンズ中心	21
3.6	鏡面領域の調整	21
3.7	レンズ方向と高さのみが一致する場合	22
3.8	レンズ中心の虚像が一致する場合	23
3.9	動画像取得システムの構成	24
4.1	画像入力システム	26
4.2	全周円筒パノラマ画像への変換	28
4.3	全周ステレオパノラマ画像におけるステレオ撮像系	30
5.2	1.5 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像	34
5.1	0 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像	35
5.3	3 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像	36
5.4	人物を注視した透視投影画像	37
A.1	Tsai のピンホールカメラモデル	II

表目次

2.1	複合現実感の分類と応用	5
2.2	全方位画像センサの比較	10

1. はじめに

センサにより取得される実際の現実環境とコンピュータにより生成された仮想環境を組み合わせることで、仮想現実感による柔軟な表現と現実環境の写実的な情報を持った、より写実的で臨場感がある仮想環境や仮想的な情報を含んだ現実環境を生成することができる。仮想環境に現実の映像を加えてより写実的で臨場感のある仮想環境を構築する技術は拡張仮想感 (Augmented Virtuality)[1] といい、現実環境に仮想物体を加えて表現力を増す技術は拡張現実感 (Augmented Reality)[2, 3] という。また拡張仮想感や拡張現実感を含む「現実環境と仮想環境を融合して利用する技術」を総称して複合現実感 (Mixed Reality)[4, 5] という。

複合現実感技術には、遠隔地の情景を仮想的にユーザの視界に提示して、ユーザにその場にいるのと同様の感覚を与えるテレプレゼンスの技術も含まれる。テレプレゼンス技術を用いることでより臨場感豊かな通信会議、さらに仮想遊園地のようなアミューズメントにも応用可能である。また、計算機上の仮想空間に現実環境のモデルを構築して、仮想的な旅行や都市計画を行うといった応用も考えられている。こういった応用のために、現実環境同等の仮想環境を CG 技術を用いて構築する場合、現実環境の全てを詳細にモデル化しなければならないという問題がある。そこで、仮想環境内に現実環境の全てをモデリングするのではなく、現実環境を画像として取り込んだり、ステレオ視のような画像処理技術を用いて現実環境のモデリングを行い、撮像した現実環境の映像を仮想環境にレンダリングすることで現実環境を仮想的に表現する研究が行われている。

複合現実感技術のために現実と同等の臨場感のある映像を仮想環境に提示するためには現実環境の多くの情報を含む情景を撮像する必要がある。すなわちビデオレートで撮像される広視野な映像の取得が求められている。このため、複合現実感のための撮像装置としては広視野画像をビデオレートで獲得することが必要となる。広視野画像の撮像装置としては、従来ロボットナビゲーションなどの分野で用いられてきた全方位画像撮像手法が広く用いられている。全周囲の画像を獲得する手法は大きく分けて、複数の撮像画像を組み合わせる全方位画像を生成するシステムと一台のカメラで全周囲の画像を一度に取り込むシステムに分類される。前者のシステムには、複数のスチルカメラにより撮像した画像を貼り合

わせ全周パノラマ画像を生成するシステムや、一台の回転カメラにより撮像された画像を利用して全周パノラマ画像を生成するシステムが提案されている。後者のシステムには、魚眼レンズやミラーを用いるシステムが提案されている。後者のシステムで使用されるミラーとしては、球面ミラー、円錐ミラー、双曲面ミラー、放物面ミラーなどの多彩なミラーが用いられている。複合現実感でこれらのセンサにより撮像される全周パノラマ画像を利用する場合、センサの持つ光学特性が重要となる。ユーザの視界に提示する画像には人間の視覚と同様に光学特性が平面透視投影である必要がある。さらに、複数の視点を持つ全方位画像から平面透視投影の画像を生成する場合、隠蔽（オクリュージョン）や違和感の問題を生じる。このため光学特性が1点透視投影である全方位画像を生成することが望ましい。複合現実感に応用されている透視投影の光学特性を持つ全方位画像取得システムには、回転カメラを用いたシステムや双曲面ミラー、放物面ミラーを用いたシステムなどがある。さらに動的な現実環境を仮想環境に構築することを考慮に入れるとビデオレートで撮像可能であることが望ましい。しかし、回転カメラを用いたシステムでは、全周囲の画像全てを同時に取得できない。また、双曲面ミラー、放物面ミラーを用いた全方位画像センサでは、1つのカメラで全周囲の環境を撮像するため、ユーザの視線方向の画像を生成した場合に解像度が低くなるという問題がある。

以上のように、複合現実感への現実環境提示のための画像センサの条件としては以下の要素が必要であると考えられる。

1. ユーザへの提示のために光学特性が1点透視投影である。
2. 環境全体を撮像するために全方位画像が取得できる。
3. 動的な環境全体を撮像するためにビデオレートで撮像される。
4. 視線方向の画像が十分な解像度がある高分解能な画像が必要である。

さらに、仮想的な物体の配置やユーザの両眼への画像提示のためには現実環境の奥行き情報が必要とされる。このため、さらに以下の条件が必要である。

5. 奥行き情報の取得のために、ステレオ撮像が可能である。

本研究では、以上の5つの条件を満たす全方位画像センサを試作した。試作したセンサは6台のカメラと六角錐ミラーから構成される撮像系を用いて同一視点からの全周囲の画像を取得することができる。さらにこの撮像系を上下対称に2つ用いてステレオ撮像も可能にしている。本研究では本試作センサの調整、システム構成を行い、本システムを用いてビデオレートで取得された画像列からそれぞれの時刻の高分解能全周ステレオパノラマ画像を生成するシステムを作成した。さらに生成した全周ステレオパノラマ画像から任意の視線方向の画像および距離画像を求め、生成されたステレオパノラマ画像と仮想環境との融合について考察した。

以下、2章では、まず複合現実感と全方位画像取得の従来手法について述べ、本研究の方針について述べる。3章では試作した六角錐ミラーを用いた全方位ステレオ画像センサの設計とキャリブレーション法について述べる。4章では、試作センサにより撮像された12個の画像列からステレオパノラマ画像を生成する手法について述べる。また、生成したステレオパノラマ画像から任意の視線方向への透視投影画像、全周パノラマ距離画像の生成法について述べる。5章では、実際に試作センサを用いて外界の映像を取り込み、撮像した時系列画像から全周ステレオパノラマ時系列画像を生成した。さらに生成した全周ステレオパノラマ画像を用いて各時刻の全周パノラマ距離画像と注視方向への透視投影画像を生成した。6章では、5章で得られた結果から、試作センサについての考察と試作センサの複合現実感へ適用に関する考察を行う。最後に7章で研究のまとめと今後の課題を述べる。

2. 関連研究と本研究の方針

複合現実感では現実の映像と仮想物体との融合により、臨場感と表現力の両方の特徴を生かした環境が生成できる。複合現実感技術のために現実環境全体の情報を取得するためには、全方位の視野を持つ画像センサが必要とされる。全方位の視野を画像センサとしては、ロボットナビゲーションの分野で様々な特性を持つ全方位画像センサが提案されている。

節 2.1 でまず、複合現実感の従来研究について述べ、節 2.2 では、従来の全方位画像センサとその特性について述べる。節 2.3 では、複合現実感のための撮像装置に必要なとされる特性について考察し、本研究における画像センサの設計方針について述べる。

2.1 複合現実感の従来研究

複合現実感技術 [4] は、センサにより撮像された実際の映像と計算機を用いて生成した仮想的な環境を融合させて双方に内在する情報を相乗的に増強しようとする技術である。人工的な要素で構成された仮想環境内での体験を扱う仮想現実感技術に比べ、複合現実感技術にはより広範な応用が期待できる反面、現実環境と仮想環境を融合させる難しさも存在する。融合の対象となるメディアには聴覚・触覚なども考えられるが、本研究では視覚的な情報のみ限定する。

複合現実感技術は、現実環境と仮想環境のどちらかに立脚して互いの融合を行うかで 2 通りのアプローチが考えられる。現実環境に存在するユーザに仮想的な情報を与える技術を拡張現実感 (Augmented Reality)[2] といい、仮想環境に没入したユーザに現実環境の映像を用いて臨場感を増す技術を拡張仮想感 (Augmented Virtuality)[1] という。どちらも両方の環境をどのようにして融合するかが技術的なポイントになる。

複合現実感における拡張現実感と拡張仮想感の分類と応用を表 2.1 に示す。左右の軸は応用例における現実の映像と仮想の映像の一般的な割合を示す。

以下に、表 2.1 中の語句の説明を示す。

- バーチャルスタジオ [9],[10]

ユーザの活動環境	利用する映像	現実映像	現実+仮想	仮想映像
	現実環境 (Augmented Reality)			←→
仮想環境 (Augmented Virtuality)				

表 2.1 複合現実感の分類と応用

電子セットとも呼ばれ、テレビ番組の製作などにおいて従来背景として用いていた大道具の代わりに、仮想の画像や実写の画像を用いるもの [9]。実際の大道具と違い、かさばらずに保存ができる。スタジオセット画像のライブラリ化により大道具製作やロケが不要となり、製作費の削減が可能である。特にサイエンス番組などでは、実際には目に見えない事象を説明する際に説得力のある演出が可能となる [10]。

- 映画・CF 製作
人目を引き付ける映像を製作するためには、あらゆるアイデアを映像化する必要がある。実写では撮れない映像を創るニーズが多い。
- 景観シミュレーション [6-8]
空間中に建築予定の建築物を表示することにより、建築前に建築後の景観を評価するのに役立つ。

- 組み立て作業支援 [11, 12]
 機械の組み立て現場などにおいて、HMD を通して実空間中に作業指示などを提示することにより、作業ミスの減少や効率化を図る [11]. 作業指示などの仮想的な映像だけでなく、小型ビデオカメラで撮像した手元の映像をインタラクティブに表示するシステムも提案されている [12].
- 遠隔作業支援 [13, 14]
 実際に人間が行けない空間（深海, 原子炉, 宇宙空間など）にロボットを送り、そこから送られてくる画像を用いて作業指示を与える. ステレオ画像中にマーカーなどを表示し、3次元位置を指定することが可能である.
- 手術シミュレーション [15]
 MRI/CT などから得られた臓器の3次元画像を、手術中の患者のビデオ画像上に重畳してHMDに表示することにより、患部の3次元位置を確認しながら手術を行なう. 手術の仮想的なトレーニングをすることにも役立つ.
- ゲーム・アミューズメント
 HMD を用いた、様々な没入型アトラクションに応用できる. さらに、映像や音声といった視覚、聴覚以外の感覚を利用したアトラクションも様々な試みられている.
- テレプレゼンス [16, 17]
 ヘッドマウントディスプレイを通じて遠隔地の映像をユーザに提示して今その場所にいるかのような感覚を与える技術であり、臨場感のある遠隔会議やアミューズメントなどへの利用が期待されている.
- コミュニティウェア
 一人のユーザ自身が仮想環境に没入するだけでなく、複数のユーザが仮想環境を共有し互いに相手を視認してコミュニケーションできるようにする技術である.

2.2 全方位画像取得に関する従来研究

複合現実感を実現するためには、部屋全体、建物全体、街全体などの情景情報を計算機に入力する必要がある。このため画像センサには、広い視野角が必要とされる。広い視野を持つ画像センサとしては従来、全方位画像センサがロボットのナビゲーションなどの分野でこれまで提案されてきた [19–23]。以下に代表的な全方位画像センサの例をあげる。

- カメラを回転する方法 [19, 20]

通常のカメラ 1 台を回転させることで、全方位画像 (パノラマ画像) を得る方法である。この方法は、後に述べる全方位画像センサに比べて、全方位で高い解像度の画像を取得することができるが、全方位の 3 次元環境を撮影するのに時間がかかるために、全方位の 3 次元情報をリアルタイムで同時に獲得するなどの拡張には不向きである。また、視点を中心にしてカメラを回転させる方法では 1 点透視投影の光学特性をもつ側方視野の円筒パノラマ画像を得ることができる。回転カメラを上下に 2 台用いてステレオ撮像を行うことができる。

- 球面ミラーを用いる方法 [21]

図 2.1(a) から分かるように、球面ミラーを鉛直下向きに取り付け、鉛直上向きに取り付けたカメラで球面ミラーに写った環境を下から撮る方法である。この方法は、下方半球を中心に側方領域まで同時に撮影することができる。しかし、画像中にカメラ自身の写り込みや床面を多く含むので、環境を提示する上で重要である側方領域の情報は少ないという欠点がある。

- 円錐ミラーを用いる方法 [22]

図 2.1(b) から分かるように、鉛直下向きの円錐ミラーと鉛直上向きに取り付けたカメラで構成されている。この方法は、側方領域の情報を多く含んだ画像が得られる。足元の情報を得るにはミラーとレンズの関係 (光学系) を十分に考慮する必要がある。

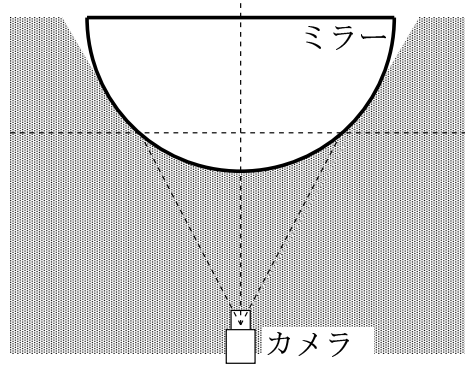
- 双曲面ミラーを用いる方法 [23]

図 2.1(c) から分かるように，鉛直下向きの双曲面ミラーと鉛直上向きに取り付けたカメラで構成される．この方法は，側方領域の情報を多く含み，さらに，足元の情報も含んだ画像が得られる．

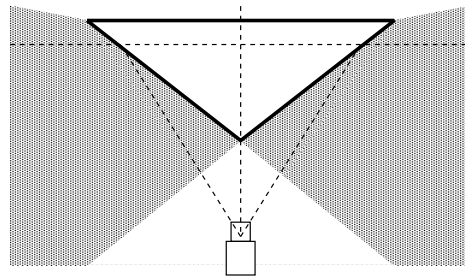
- 放物面ミラーを用いる方法 [24]

双曲面ミラーを用いる方法と同様に，鉛直下向きの放物面ミラーと鉛直上向きに取り付けられたカメラで構成される．カメラに一般の透視投影カメラでなく，正射影カメラを用いる．撮像される画像の特性は双曲面ミラーを用いた場合と同じであり，透視投影の光学特性を持つ．キャリブレーションは双曲面ミラーを用いる場合よりも容易である．

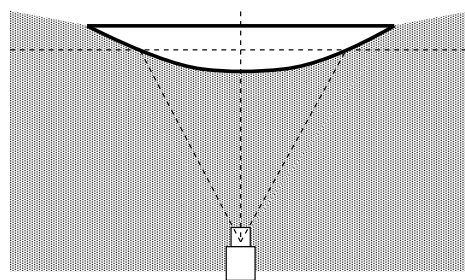
カメラを回転させる方法と双曲面ミラー，放物面ミラーを用いる方法以外の全方位画像センサは，光学系の特性が透視投影ではない．このために，人間への画像の提示を目的に使用するには適切ではない．これに対して，カメラを回転させる方法は光学特性が透視投影であり，高解像度の全周パノラマ画像を撮像できるので静的な現実環境を提示する [16] には適している．しかし環境全体を撮影するのに時間がかかるために，動的な環境における全方位の映像をビデオレートで同時に獲得するなどの拡張には不向きである．一方，双曲面ミラー，放物面ミラーを用いる全方位画像センサは，光学特性が透視投影であり，全方位の映像を一度に撮像できるので，リアルタイムで撮影した映像を利用する応用 [17] に適している．しかし，全方位の映像を一台のカメラで取得するため，分解能に制限があるという問題がある．



(a) 球面ミラー



(b) 円錐ミラー



(c) 双曲面ミラー

▨ : 視野領域

図 2.1 全方位画像センサの視野領域

2.3 本研究の方針

複合現実感の実現のためには、動的な現実環境を広い視野と高い分解能で撮像できる画像装置が必要である。

このような画像装置としては以下の特性が要求される。

1. 光学特性が 1 点透視投影
2. 視野角の範囲が全周囲
3. ビデオレート
4. 高分解能

また現実環境の距離情報の獲得のために、

5. ステレオ撮像

も必要である。

以上の特徴について従来の全方位画像センサについて考察すると、表 2.2 に示すように全ての条件を満たす画像センサは存在しなかった。

名前	1 点透視投影	ビデオレート	高分解能	ステレオ
スチルカメラ画像 [27]	△	×	○	×
回転カメラ [19]	○	×	○	○
魚眼レンズ [18]	△	○	×	△
球面ミラー [21]	×	○	×	△
円錐ミラー [22]	×	○	×	○
双曲面ミラー [23]	○	○	×	○
放物面ミラー [24]	○	○	×	○

表 2.2 全方位画像センサの比較

従来センサのうち、多くの条件を満たす回転カメラを用いた全方位画像センサと双曲面ミラー、放物面ミラーを用いた画像センサについて考察すると、回転カメラを用いた全方位画像センサは、同時に全周囲の画像を撮像できない点で問題があり、双曲面ミラー、放物面ミラーによる全方位画像センサは全周囲の画像を1つのカメラでしか撮像することができず、分解能に制限がある点で問題があった。

複合現実感の実現のためには、動的な実環境の撮影を考慮に入れると全周囲の画像をビデオレートで撮像することが必要である。また複合現実感における画像提示では取得した全周画像のうち、ユーザの視野に含まれる領域を切り出して提示するのが一般的である。このため、ユーザに提示する領域に十分な解像度を持たせられるほどに高い解像度を持った全周画像が必要である。

解像度の高い画像を得るためには、次の2つの方法がある。

- 画像センサの分解能を上げる。
- 複数の画像センサを用いる。

前者はカメラ自身の分解能の制限を緩和する方法である。1台のカメラを用いて全周囲の画像を撮像する双曲面ミラー、放物面ミラーを用いた手法などがもつ分解能の制限を改善することができる。撮像部分に高分解能なカメラ、たとえば、ハイビジョンカメラ等を用いて分解能を向上させる。しかし、ユーザに高品質の映像を提示する場合にハイビジョンの精度で表示する必要がある場合もある。こういった場合には、切り出す一部の領域がハイビジョンの分解能を持つ必要がある。そういった場合にはさらに大きな分解能を持つセンサを開発する必要がある。

一方、後者はより高分解能な撮像を行うために、撮像する全周囲の環境をいくつかの領域に分け、それぞれの領域を別々の複数のセンサを用いて撮像することで全体として高分解能な全周囲の撮像を行う方法である。より多くのセンサを用いることで高解像度の全周画像を取得できる。回転カメラによる手法がこれに含まれる。ただし、1台のカメラを回転させて全周囲の画像を取得するため、一度に全周囲の画像を取得することができない。ビデオレートで全周の画像を取得するためには、ビデオレートの範囲内で全周の画像を取得できるよう高速な回転、撮像を行う必要がある。しかし実際には安定的に高速な回転が可能な撮像装置を製作することは非常に難しい。

センサの分解能を上げる方法とセンサの数を増やす方法を比較する。

ユーザに提示する映像の画質が向上する場合のことを想定すると高分解能な画像センサを用いる場合には、ユーザに提示する映像の画質が向上するにつれて新たな高分解能カメラを開発する必要がある、負担が大きい。一方、複数のビデオカメラを用いる場合には、ユーザに提示する映像の画質が向上すれば、一般のビデオカメラの進歩に合わせてカメラを交換するか、使用するビデオカメラの台数を増やせばよく、設計の変更が比較的容易である。

そこで本研究では複数の画像センサを用いて全周囲の映像を獲得する方法を採用した。しかし複数のカメラを用いて全周囲の映像を撮像する場合に現実の1地点に複数のカメラを配置することはできないという問題がある。そのため、複数の画像から1点透視投影の全周パノラマ画像を生成する場合、視点を固定してカメラを回転させる方法が従来用いられてきた。しかし、ビデオレートで環境全体の情報を取得するためにはそれぞれの方向を同時に複数のビデオカメラで撮像する必要がある。

そこで本研究では、複数のカメラの使用と同一視点からの撮像という両方の条件をピラミッド型ミラーによる虚像を用いることで実現し、ピラミッド型ミラーを用いた高分解能全方位画像センサを提案、試作した。

ピラミッド型ミラーの側面の数は以下のようにして決定した。ピラミッド型ミラーとして多くの側面を持つミラーを選択する利点として以下の2つが上げられる。

- 全周を撮像する分解能が向上する。
- カメラの撮像視野を小さくすることが可能でカメラ収差を小さくすることができる。

一方、次の3つの点で不利である。

- カメラ配置、パラメータ調整が煩雑になる
- ミラーの稜線の精度が低いとその部分の情報が欠ける。
- センサやミラーのサイズが大きくなる。

よって、互いに矛盾するこれらの要求を考慮してピラミッド型ミラーの形状と CCD カメラの数を決定する必要がある。一方、製作上の問題としてミラーの精度を上げると大きなミラーの製作は難しくなり、小型 CCD カメラを使用しても CCD カメラの数を 6 個以上増やすことは現在の技術では難しい。そのため、実際に設計したピラミッド型ミラーの形状は正六角錐にした。これならば画像の歪みは少なく、CCD カメラの分解能も十分であると考ええる。

六角錐ミラーを用いた画像センサを利用することで回転カメラを用いた場合のように、同一視点からの複数の画像を用いて高解像度の全周パノラマ画像を作成できる。また、6 台のセンサを用いることで同時に全ての方向の画像をビデオレートで撮像できる。

さらに六角錐ミラーと 6 台のカメラからなる撮像系を上下対称に 2 つ用いることでステレオ撮像を可能にしている。

3章で試作センサの詳細な設計について述べる [26].

3. 六角錐ミラーを用いた全方位ステレオ画像センサ

本章では、本研究で設計、試作した複合環境構築のための全方位ステレオ画像センサについて述べる。複合現実感のためには動的な広域環境を高分解能で撮像できるセンサが必要である。本研究では同一視点からの全周囲の映像を複数のカメラを用いて撮像することにより実現した。複数のカメラの視点の重ね合わせは六角錐ミラーによる虚像を利用することで可能になる。以下では試作したセンサの設計、キャリブレーション法、撮像システムについて述べる。

3.1 全方位画像センサの設計と試作

複合環境構築のための画像センサには、以下の特徴が必要とされる。

1. 1点透視投影の光学特性を持つ画像を取得できる。
2. 全周囲の映像が取得できる。
3. ビデオレートで撮像できる。
4. 高分解能な撮像が行える。
5. ステレオ撮像が可能である。

以上の特徴を満たすためには同一視点からの全周画像を複数のカメラを用いてビデオレートで撮像することが必要である。

本研究では六角錐ミラーを用いることでのカメラを用いて透視投影の光学特性を持つ高分解能な全周の環境情報をビデオレートで取得することができる。また上下に2つの全周画像撮像系を配置することでステレオ撮像を可能にしている。

図 3.1 に試作センサの構成を図 3.2 に試作センサの外観を示す。

試作センサは、全方位画像を取得するための構造を上部と下部に2つ持ち、上下方向に視差のついた全方位のステレオ画像を取得できる。

全方位画像を取得するための各構造を、上部の構造を例に説明すると、鉛直上向きに六角錐ミラーとその上に鉛直下向きに配置した6台の CCD カメラから構成される。虚像上でカメラのレンズ中心が1点に重なる位置にカメラを配置する。

虚像上でカメラの視点を一致させることで、試作センサでは、高解像度の1点透視投影の光学特性を持つ全方位画像を得ることができる。6台のカメラで全周の映像を取得するため、6台のカメラにより撮像される画像は6つの異なる光軸方向を持つ平面透視投影画像となる。それぞれのカメラはミラーの虚像上で視点を共有し、それぞれが水平面上で六角錐ミラーの側面を中心を通るように視点を中心に60度ずつ回転した光軸方向を持つ。

試作センサにより撮像できる全方位画像は、同じ視点を持ち、水平面上で60度ずつ回転した光軸を持つ6枚の平面透視投影画像として得られる。



図 3.1 高分解能全方位ステレオ画像センサの設計



図 3.2 試作センサの外観

3.2 全方位ステレオ画像センサのキャリブレーション

試作センサにより、同一視点からの全方位画像を得るためには、ミラーに対して、正しい位置に各カメラを配置する必要がある。さらに、本画像センサでは6台のカメラから全方位の画像を取得するために広角(水平画角 60 度以上)のレンズを必要とする。このため、撮像される画像中に広角レンズに特有の大きな樽型歪みが生じ、あらかじめこのような樽型歪みを除去するためのパラメータを各カメラで求める必要がある。

樽型歪みの除去とカメラの正しい配置のために、Tsai のキャリブレーション法を用いて、まず、個々のカメラの樽型歪み中心、樽型歪み係数を求め、つぎに、各カメラのレンズ中心を正六角形の各頂点に配置し、配置された平面に光軸が垂直になるようにカメラを配置する。

3.2.1 樽型歪み補正

カメラ特性を求める手法は数多く提案されている [30, 31]。ここでは、樽型歪み補正のために、ピンホールカメラモデルにレンズ収差を加えた Tsai のカメラモデル [32] を用いる。

Tsai のキャリブレーション法では、CCD エLEMENT の大きさなど 6 種類のパラメータをカメラの仕様から決定した後、既知の 3 次元座標とその点のその CCD カメラへの投影点の座標の対応からカメラの内部パラメータ 5 つと外部パラメータ 6 つの計 11 パラメータを非線形最適化問題を解くことにより求める。ここで、内部パラメータは焦点距離、樽型歪み中心、樽型歪み係数、 x 軸方向のスケールファクタであり、外部パラメータは既知の 3 次元座標軸により表現されるカメラの焦点座標、光軸方向のそれぞれ 3 パラメータである。Tsai 法の詳細は付録 A で述べる。

図 3.3 に示すような、平面上に縦 34mm、横 28mm の間隔で描いた縦 16 点、横 11 点の 3 次元座標値を与えて、Tsai のキャリブレーション法による内部パラメータ推定を行う。撮像した格子点の位置は、格子点を含む領域を切り出し、その領域を 2 値化した画像の黒点の重心とした。

Tsai のキャリブレーション法によるパラメータ推定では、最小 2 乗法を用いる

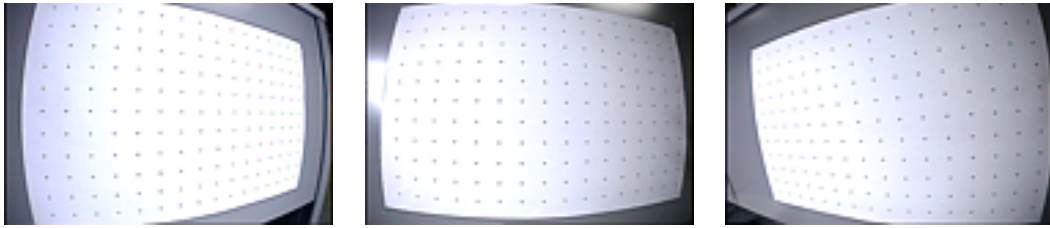


図 3.3 Tsai 法のための入力格子点

ため、格子点の座標に含まれる誤差によっては、推定値に大きな誤差を生じることがある。そのため、図 3.3 のように、3 種類の異なる視線方向からの画像から測定し、求めた焦点距離が実際のカメラの焦点距離から大きく異なるデータを除去し、残りのデータの平均を樽型歪みパラメータとした。

試作カメラから撮像された歪みを含む画像と Tsai のキャリブレーション法により樽型歪みを除去した画像を図 3.4 に示す。



図 3.4 撮像画像 (左) と歪み除去画像 (右)

3.2.2 レンズ中心の位置合わせ

本節では、本試作センサのレンズ中心をミラーの虚像上で 1 点に重ねるための位置合わせ手順について述べる。各カメラ配置の調整を以下の手順で行う。

1. 各ミラーの稜線上にうつるレンズ中心の位置を合わせる。

2. 撮像画像上のミラー領域を一致させる。

3. ミラー領域内の重複領域を除去する。

(1)において、レンズの虚像をミラーを通して、直接見てカメラの視点の位置合わせを行う。見えるレンズは外側のレンズなのでセンサのレンズ中心とは異なる。そこで(2)(3)のような撮像される映像を用いたより詳細な調整を行う。(2)においては、撮像画像におけるミラー位置の見え方に基づいて各カメラと鏡面間の相対的な位置関係とカメラの光軸方向の調整を行う。(3)においては、ミラーに映る現実環境の視野に基づいてレンズ中心の虚像の重ね合わせを行う。

鏡面上のレンズ中心の鏡像の重ね合わせ

本画像センサはミラーによる虚像を用いて同一視点からの全方位画像を取得する。片側6台のカメラの虚像を一致させるために、図3.5のように目視で隣り合う2つのカメラレンズの鏡像を六角錐ミラーの稜線上で重ね合わせる。重ね合わせのためには各カメラの方向や高さの調整を行う。ただし目視で見えるレンズ中心は最も対物側のレンズであり、実際のレンズ中心とは必ずしも一致せず、調整の精度は粗い。そのため以下に述べる撮像画像を基にしたさらに精密な調整を行う必要がある。しかし、カメラの撮像領域を用いる方法よりも容易に調整できるので、あらかじめ大まかな位置、方向を定めるために行われる。

撮像画像上の鏡面領域の調整

次に、カメラの撮像画像を用いて、より精密に各カメラの方向を調整し、高さを揃える。六角錐ミラーにおける各鏡面は合同なので、各カメラの撮像する鏡面領域が等しくなるように調整すれば、鏡面と各カメラの位置関係が等しくなる。また、鏡面の撮像領域を撮像領域全体の中央にとれば、各カメラの光軸方向を揃えることができる。そこで、各カメラに対して図3.6のように撮像画像上の鏡面領域が一定となるように調整する。この調整では、図3.7に示すように全てのカメラの光軸は虚像上の同一点で交わり、それぞれの鏡面の中心を通るようになる。またレンズ中心と虚像上の交点の距離は全てのカメラで等しくなる。以上のように各カメラとミラーの相対的な

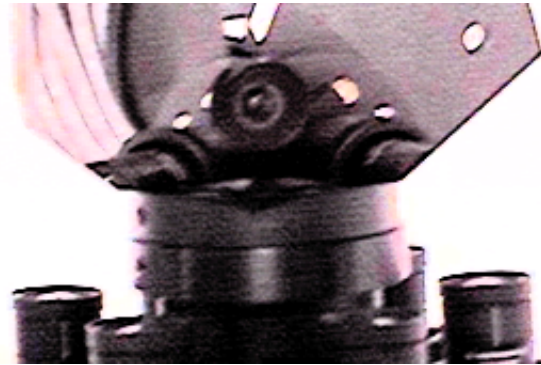


図 3.5 稜線上で重なるレンズ中心

位置を一定にすることができる。しかし、視点の虚像を 1 点に重ねることはできない。視点の虚像を 1 点に重ねるために以下に述べる重複部分の除去による調整を行う。



図 3.6 鏡面領域の調整

鏡面領域内の重複部分の除去

最後に、先に高さを揃えた全カメラのレンズ中心を光軸の交点上に移す。図 3.8 で示されるようにカメラ中心が 1 つに重なる場合、各カメラにおける視野の重複部分は全くないものの連続した全周囲の映像を取得することができる。そこで、カメラの撮像領域上の鏡面領域の視野を重複なく連続するように、カメラ 6 台全体の上下移動を行うことによってレンズ中心の重ね合

わせを行う。カメラとミラーの相対的な位置と光軸方向の調整がうまく行われておれば、全体の上下だけでレンズ中心の位置を一致させることができる。一致しない場合には、さらに、すべてのカメラのレンズ中心が正しく重なるように個々のカメラの方向の微調整を行う。

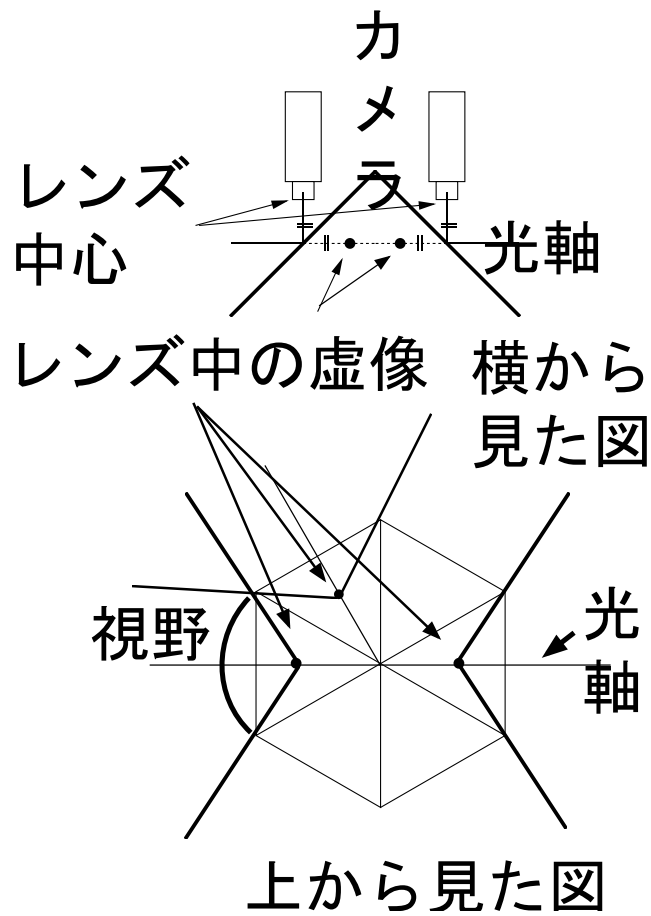


図 3.7 レンズ方向と高さのみが一致する場合

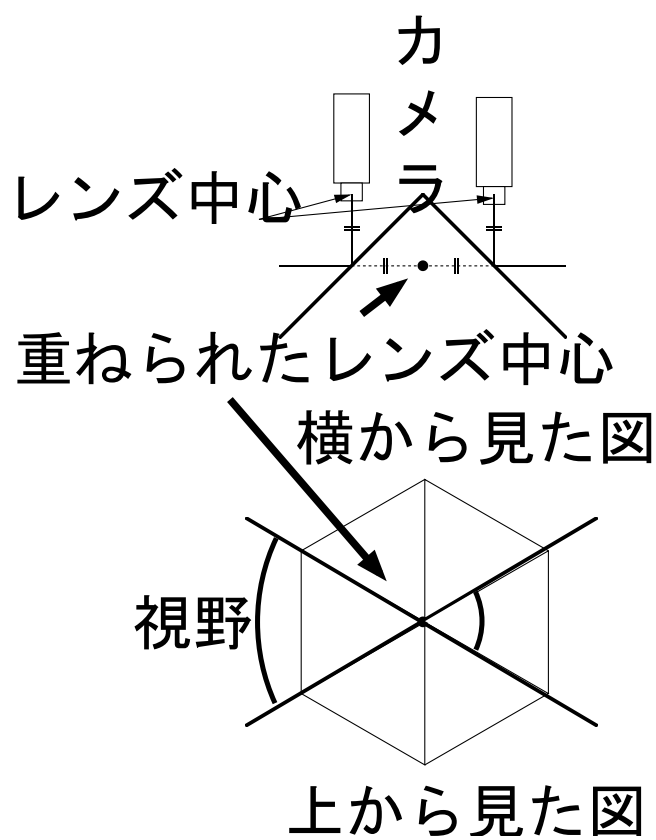


図 3.8 レンズ中心の虚像が一致する場合

3.3 撮像システム

試作センサにより得られる全周ステレオ画像情報は 12 台のカメラ毎に別々の平面透視投影画像として記録される。同時刻における全周ステレオ画像を得るためには、試作センサ上のカメラによる撮像は同期をとって行われる必要がある。そこで試作センサに用いる撮像システムを図 3.9 のように構成した。本撮像システムでは各カメラが外部同期により同時刻の画像を取得できるように、シンクジェネレータが生成する同期信号を各カメラに分配する。また、同時に同期信号をタイムコードジェネレータに与え、同期のタイミングに合わせたタイムコードを生成し、各ビデオデッキに生成したタイムコードを供給する。ビデオデッキは接続したカメラの映像をタイムコードとともにビデオテープに記録する。

以上のシステムにより、全方位の画像が時刻情報とともに、12 個のビデオテープにビデオレートで記録することができる。

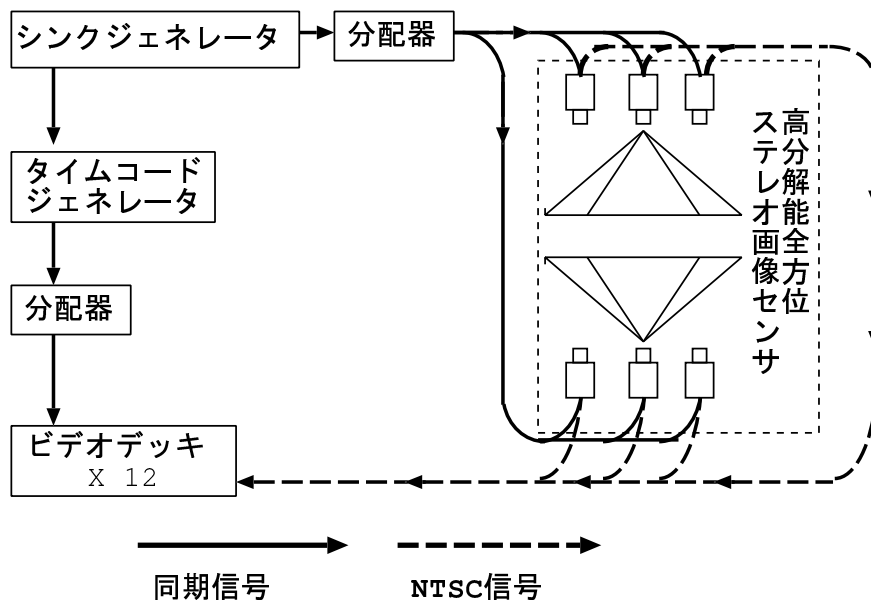


図 3.9 動画像取得システムの構成

4. 全周ステレオパノラマ画像の生成

試作画像センサで撮像された画像は個々のカメラのビデオ画像として、まず保存される。全周パノラマ画像を生成するためには、これらの動画像から各ビデオフレーム毎に、同時刻の画像を抜き出し、その抜き出した画像を貼り合わせて全周パノラマ画像を生成する必要がある。まず、複数のビデオ画像から、各フレームの静止画像を抽出するためのシステムの構成について示し、次に、抽出された12枚の静止画像から全周ステレオパノラマ画像を生成する方法を示す。次に、生成した全周ステレオパノラマ画像から、任意の視線方向の透視投影画像および全周パノラマ距離画像を生成する方法を述べる。

4.1 画像入力システム

試作センサでは、全周のステレオ画像情報を12枚の透視投影画像として記録する。ある時刻の全周ステレオパノラマ画像を生成するためには、撮像した12台のカメラのビデオ画像から同時刻の画像を取り出す計算機に取り込む必要がある。

ビデオテープから全ての動画像を計算機に入力した後、時刻情報に基づいて必要な画像を選び出す場合、計算機側に高分解能な12枚の動画を静止画の集合として保存する必要がある、大きな記憶領域を必要とする。そこで全周ステレオパノラマ画像を生成するのに必要な画像を適宜、ビデオテープから入力することを考える。しかしこの場合も12台のコントロール可能なビデオデッキと入力デバイスを用意するか、1台のビデオデッキを用いてビデオテープの出し入れを繰り返す必要があるという問題がある。そのため、全てのビデオ画像を一時的に別の記憶媒体に保存し、その記憶媒体からそれぞれの時刻の画像を取り出すことが望ましい。そこで本研究では、図4.1のようなタイムコード読み取り機能付きビデオプレーヤー、レーザービデオディスクレコーダー、ワークステーションから構成したシステムを構成した。レーザービデオディスクレコーダーを用いることで、ハードディスクなどよりも、保存した画像に対して高速にランダムアクセスが可能になる。

カメラ情報やタイムコード情報を含む画像データをレーザービデオディスクレ

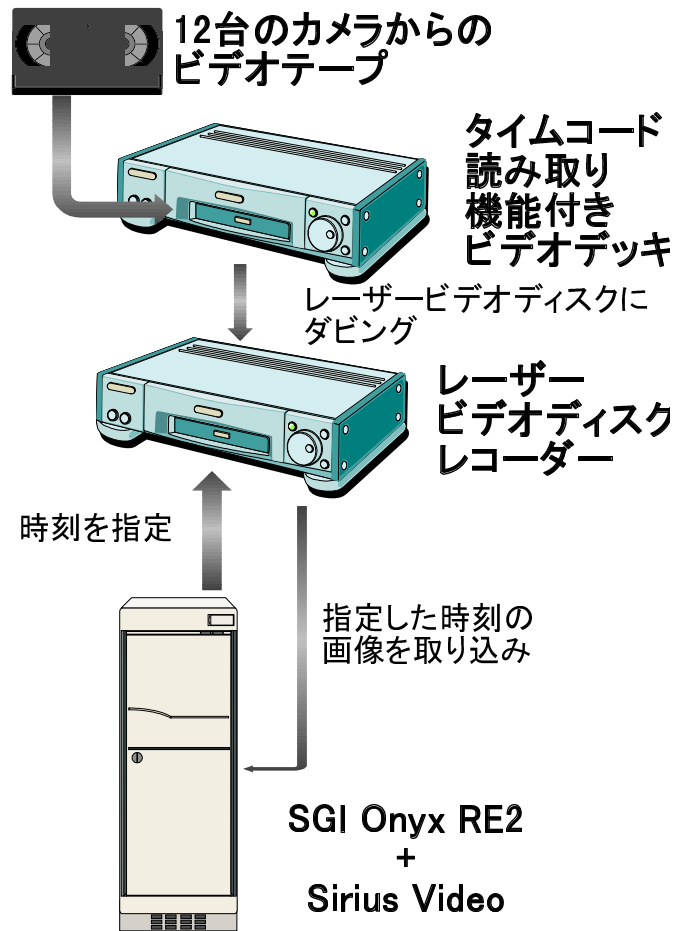


図 4.1 画像入力システム

コーダーにうつす手順は以下のようなになる。

1. レーザービデオディスクの領域を 12 台のカメラの映像保存用に割り当てる。
2. 各ビデオテープに保存された全方位画像のコピーする領域の先頭のタイムコード T_s とコピーする時間 T を決める。このとき保存範囲がレーザービデオディスクの映像保存用に割り当てた領域の範囲を超えないようにする。
3. 取り込むカメラのビデオをビデオデッキに入れる。
4. カメラのビデオテープの再生開始タイムコードを T_s に合わせる。

5. T の間, ビデオデッキとレーザービデオディスクレコーダーとの間で同期録画を行う.
6. 12 台のカメラのそれぞれの画像に対して 3, 4, 5 を繰り返す.

上のような手続きにより, 12 台のカメラの画像を 1 枚のレーザービデオディスクに保存することができる.

全周ステレオパノラマ時系列画像の作成にあたっては, 時刻毎のの画像をレーザービデオレコーダーから計算機のメモリに取り込み, 4.2節で述べる手法を用いて全周パノラマ画像を生成し, ハードディスクに保存する.

4.2 全周ステレオパノラマ画像の生成

試作センサにより同時刻に撮像された上下それぞれ 6 枚の画像は図 4.2 の左上図のような六角柱の側面に投影された全周画像を構成する. 角柱の側面への投影のため, 角柱の辺上で透視投影の投影方向が不連続になる. このため直線が折れ線に見えるといった問題がある. そこで図 4.2 のように六角柱側面へ投影された全周画像から円筒面への投影である全周円筒パノラマ画像を作成する. 円筒面への投影の場合, 透視投影の投影方向は滑らかに変化し, 直線は折れ線ではなく滑らかな曲線に投影される. 全周円筒パノラマ画像を生成することによりパン方向の視線変化に対して違和感のない全周パノラマ画像を生成することができる. 本試作センサによる上下 6 枚ずつの撮像画像で構成される六角柱側面へ投影された全周パノラマ画像から円筒パノラマ画像への変換には全周パノラマ画像上の (θ, h) と撮像画像上の点 (x, y) の対応を求める必要がある. この対応を求めるための式は図 4.2 より, 式 1, 2 のようになる.

$$x = c_x + R \cdot \tan \theta \quad (1)$$

$$y = c_y + \frac{h \cdot R}{r \cdot \cos \theta} \quad (2)$$

ここで

R : カメラの焦点距離 (pixel)

r : 生成する全周パノラマ画像の半径 (pixel)

c_x, c_y : 撮像画像中心

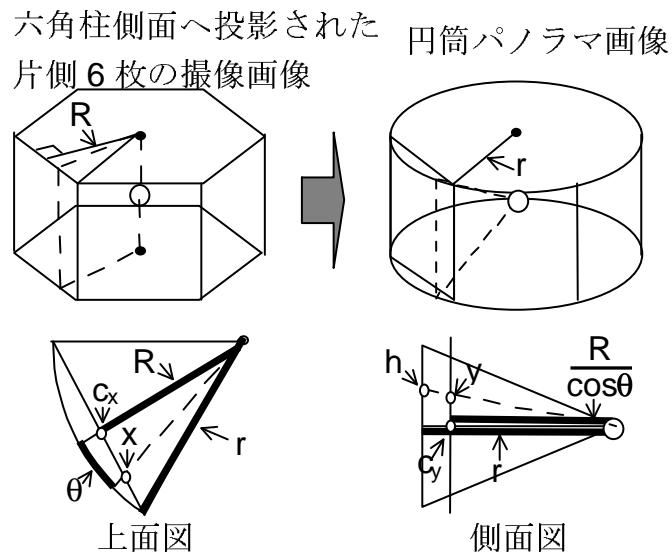


図 4.2 全周円筒パノラマ画像への変換

また、それぞれのカメラの撮像画像の中心位置 (c_x, c_y) 及び焦点距離 R はあらかじめ設定しておくパラメータであり、隣り合う画像が連続するように、また上下の画像でエピポーラ線が一致するように、全てのカメラについてヒューリスティックに求めた。

以上の変換式により、上下各々 6 枚の透視投影画像から円筒面に投影した全周ステレオパノラマ画像を生成する。

4.3 任意方向の透視投影画像の生成

全周円筒パノラマ画像から、先に述べた変換の逆変換を用いて、その画像中の任意の方向への透視投影画像を生成することができる。この変換は式 3 で表される。

$$\theta = v + \arctan \frac{(x - c_x)}{R} \quad (3)$$

$$h = \frac{(y - c_y) r \cdot \cos \theta}{R} \quad (4)$$

ここで v は左右の視線方向である. これにより生成した全周パノラマ画像中でリアルタイムでの自由な見回しが可能な蓄積型のテレプレゼンス [17] を行うことができる.

4.4 全周パノラマ距離画像の生成

本試作センサでは, 上下の全周パノラマ画像は図?? のような上下の全周パノラマ画像の円筒面座標系とワールド円筒座標系により構成される縦視差を持ったステレオ撮像系を構成する.

カメラの焦点距離を f , カメラの基線長を b とする.

ワールド円筒座標上の点 $P(\Theta, H, D)$ が上下の円筒面座標系上の座標 $p(\theta_u, h_u)$, $p_d(\theta_d, h_d)$ にそれぞれ現れたとするとエピポーラ拘束により

$$\theta_u = \theta_d \quad (5)$$

が成立し, 点 $P(\Theta, H, D)$ の座標はそれぞれ次のように表すことができる.

$$\Theta = \theta_u = \theta_d \quad (6)$$

$$H = \frac{b \cdot (h_u + h_d)}{2(h_d - h_u)} \quad (7)$$

$$D = \frac{b \cdot f}{h_d - h_u} \quad (8)$$

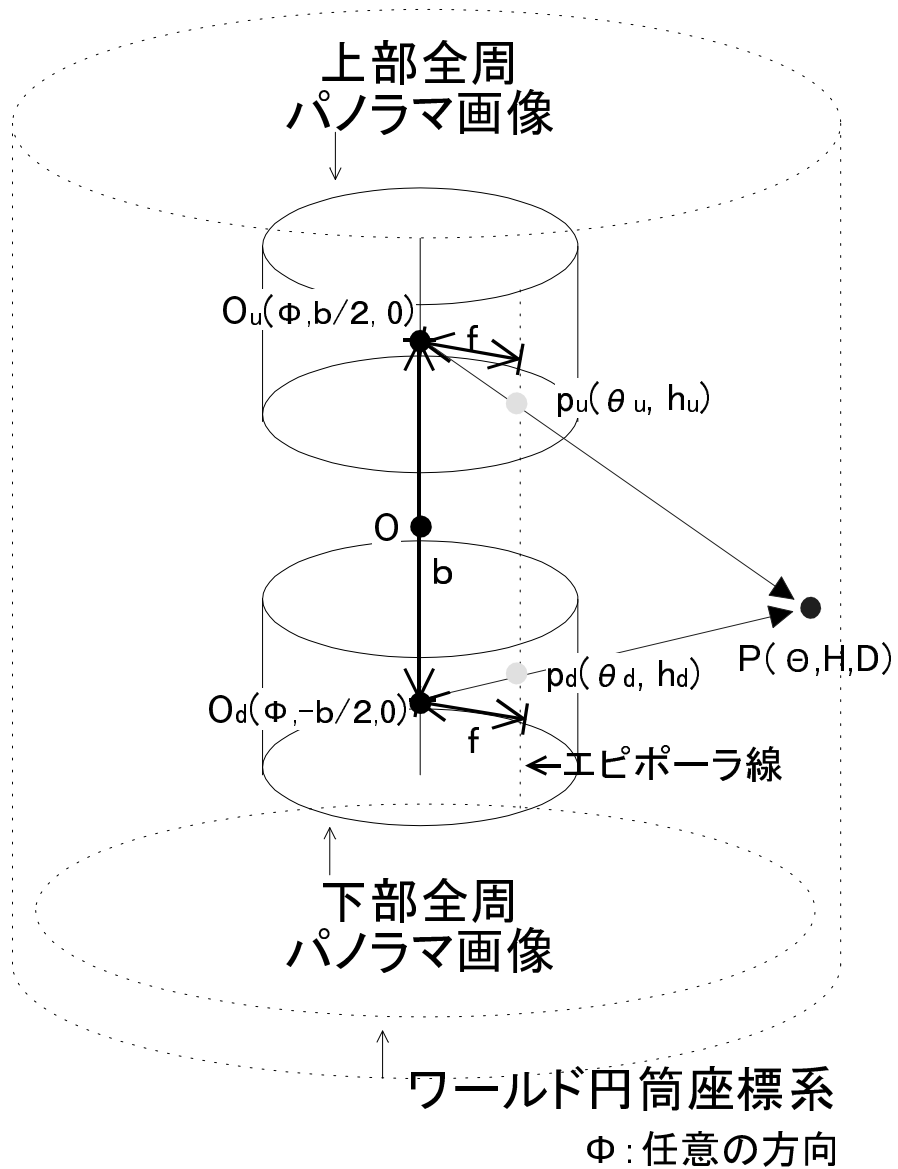


図 4.3 全周ステレオパノラマ画像におけるステレオ撮像系

式 8 より, b, f が固定であれば奥行きは視差 $d \equiv h_d - h_u$ のみに依存することがわかる. すなわち, 視差推定は奥行き推定と等価である.

視差推定では, 上下の画像間の対応点をどのようにして正しく求めるかが重要である.

本研究では、下方の全周パノラマ画像上の注目点が上方の全周パノラマ画像上のどの部分に対応するかをテンプレートマッチング法 (相関法) により求める。

以下では下方の全周パノラマ画像を基準画像として、上方の全周パノラマ画像上で対応点探索を行う方法について述べる。

対応点探索手法

式 5 に示すように、上下の全周パノラマ画像間で縦方向のエピポーラ拘束が成立する。そこで対応点探索を行う際には探索範囲がエピポーラライン上に限定できる。ある点の視差はこの探索範囲内で画像の輝度値を特徴量としてテンプレートマッチング法 (相関法) を行うことにより、対応点を発見し、視差を求めることができる。対応点探索を全周パノラマ画像の全点で行い、全方位画像間の視差マップを求めることができる。視差マップから全方位の奥行き情報を求めることができる。

以下に本研究で用いたテンプレートマッチング法の評価基準と探索領域について述べる。以下では、基準画像 I_r 上の注目点 (参照点) におけるテンプレートを参照テンプレートと呼び、探索画像上の探索点におけるテンプレートを探索テンプレートと呼ぶことにする。

マッチングの評価基準

テンプレートマッチング法の評価基準には、上下の色合いの違いを考慮して、参照テンプレート、探索テンプレート内の各輝度値を表したベクトルの正規化相互相関を用いた。テンプレートの大きさを (M, N) 、参照点 $I_r(X, Y)$ における輝度値を $P_r(X, Y)$ 、探索点 $I_s(x, y)$ における輝度値を $P_s(x, y)$ とすると、マッチングの評価値となる正規化相互相関 $Corr$ は以下のように表せる。

$$Corr = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} P_r \left(X - \left[\frac{M}{2} \right] + i, Y - \left[\frac{N}{2} \right] + j \right) \cdot P_s \left(x - \left[\frac{M}{2} \right] + i, y - \left[\frac{N}{2} \right] + j \right)}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(P_r \left(X - \left[\frac{M}{2} \right] + i, Y - \left[\frac{N}{2} \right] + j \right) \right)^2 \cdot \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(P_s \left(x - \left[\frac{M}{2} \right] + i, y - \left[\frac{N}{2} \right] + j \right) \right)^2}} \quad (9)$$

正規化相互相関は、相関を求める2つのベクトルの余弦の値を表し、 -1 から 1 の範囲をとる。この値が 1 の時、最も相関があり、 0 の時全く相関がない。 -1 の場合には逆の相関を持つ。

対応点探索を行う際には参照テンプレートと探索テンプレートの正規化相互相関 $Corr$ の値が大きいほど類似性の評価が良いことを示している。つまり正規化相互相関 $Corr$ の値が大きいほど下方の全周パノラマ画像上での参照テンプレート領域の画像と上方の全周パノラマ画像上の探索テンプレート領域の画像パターンが類似することがわかる。下方の全周パノラマ画像上の注目画素に対する上側の画像の対応画素を探索することにより、視差 d を求めることができる。

探索領域

探索領域はエピポーラ制約により x 座標が限定される。ここでは、全周パノラマ画像生成の際の位置ずれを考慮して、探索領域の x 方向の範囲は左右 10 画素程度広く設定し、 1 画素おきに探索を行った。また、上方の全周パノラマ画像上の対応点は上の座標に対応点があることはあり得ない。そこで探索領域の y 方向の範囲は参照画像の y 座標から底辺までに設定した。

5. 実験

4章までに本研究で用いる試作センサ及び撮像システムについて述べた。実験において本試作センサにより撮像された画像列及び生成した全方位ステレオパノラマ画像を示し、次に複合現実感のための応用例のために生成される任意方向の透視投影画像と距離画像を示す。

実験の目的は、静的な環境で実際にステレオパノラマ画像列を生成し、生成する際の合成画像の連結性を評価し、違和感を調べることである。以下に実験条件について述べ、撮像画像、全周ステレオパノラマ画像、任意方向の透視投影画像、全周パノラマ距離画像を示す。

5.1 実験条件

試作センサを室内に固定して、センサの周囲を物体が4人の人物が歩く環境を撮像した。センサの上方の視点は地面から約160cm、下方の視点は約140cmになるように設置した。約10秒間撮像した。試作センサ上の各カメラは24bitカラー、640x486画素の画像をビデオレートで撮像する。

5.2 実験結果

本実験では、24bitカラーの640x486画素の撮像画像12枚から、24bitカラーの3000x350画素の全周ステレオパノラマ画像を生成した。次に、生成された全周ステレオパノラマ時系列画像のうち、0, 1.5, 3秒時の撮像画像と全周ステレオパノラマ画像と全周パノラマ距離画像を図5.1~5.3に示す。(a)に撮像した画像、(b)に全周ステレオパノラマ画像、(c)に生成した全周パノラマ距離画像を示す。生成した全周ステレオパノラマ画像は現状では各画像間の色調や接続部分にやや違和感が残る。全周パノラマ距離画像は白いほど近く黒いほど遠い。全周パノラマ距離画像には誤差が多く残っているが人間領域と背景領域の奥行き程度は求まっている。最後に上部の全周パノラマ画像から生成した3方向の平面透視投影画像を図5.4に示す。それぞれの画像は一番センサに近い人物を注視する方向の画像を生成した。画像のサイズは500x350画素である。画像の接続部分での違和感が目

立つものの全周ステレオパノラマ画像から平面への投影は違和感なく実現されている。



(a) 12台のカメラによる撮像画像

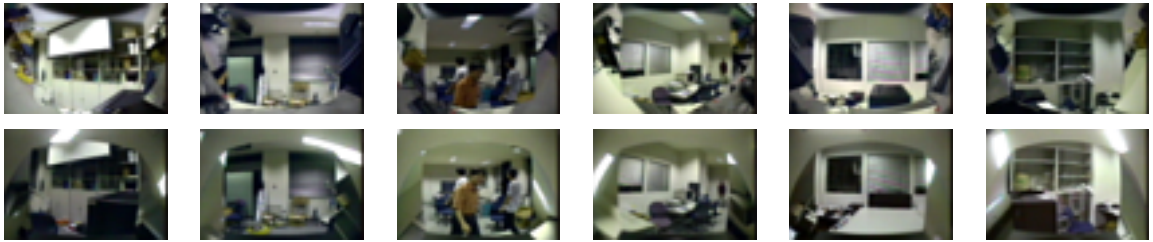


(b) 全周ステレオパノラマ画像



(c) 全周パノラマ距離画像

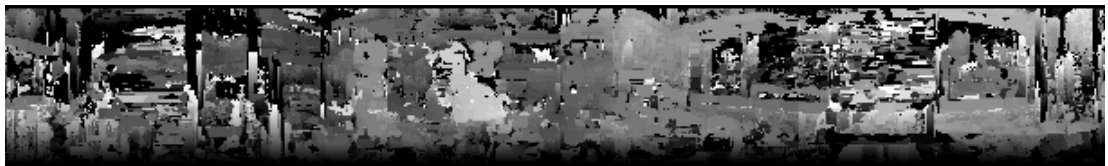
図 5.2 1.5 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像



(a) 12 台のカメラによる撮像画像

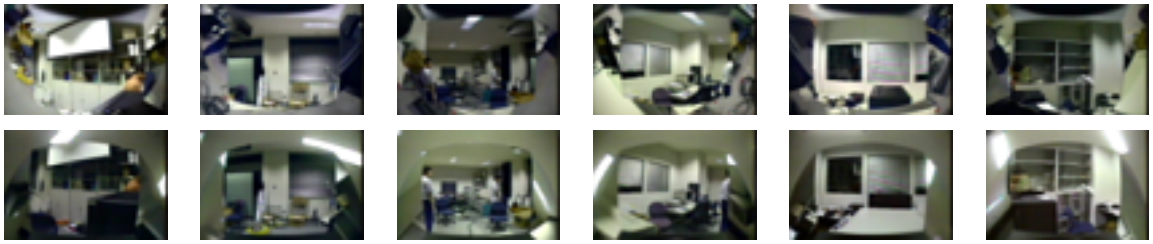


(b) 全周ステレオパノラマ画像



(c) 全周パノラマ距離画像

図 5.1 0 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像



(a) 12 台のカメラによる撮像画像



(b) 全周ステレオパノラマ画像



(c) 全周パノラマ距離画像

図 5.3 3 秒時の撮像画像, 全周ステレオパノラマ画像, 全周パノラマ距離画像



(a) 0 秒時



(b) 1.5 秒時



(c) 3 秒時

図 5.4 人物を注視した透視投影画像

6. 考察

節5で示されるように試作センサにより高解像度な全周ステレオパノラマ動画像を得ることができる。高解像度の全周ステレオパノラマ画像には、テレプレゼンスやコミュニティウェアのようなヘッドマウンテッドディスプレイを通じた遠隔地情報の提示を行う場合以下に述べる大きな利点がある。

- 視線方向の一部を切り出しても十分な分解能を持つので、見回しの許されるテレプレゼンスでも、臨場感のある映像を提供することができる。

一方、景観シミュレーションなどのように現実環境のモデルリングを行う場合においては次のような利点がある。

- 車両上に設置した画像センサを用いて移動しながら現実環境を撮像する場合、計測すべき場所を様々な位置、様々な角度から高分解能でステレオ計測することが可能である。このため現実物体、現実環境の比較的詳細なモデルを構築することが可能になると考えられる。作成したモデルに撮像した映像をレンダリングする場合には、豊富な視点、視線方向を持つ映像をあらかじめ取得できるので適した画像を選ぶことができる。

さらに環境全体の映像が取得できるので撮像中の映像以外の環境情報、例えば光源情報等を求めることができる。得られた位置情報、形状情報、光源情報は複合環境内での臨場感のある仮想物体の配置、描画に必要である。

こういった様々な複合環境の応用に試作センサから得られた全周ステレオパノラマ画像を用いるためにはまだ解決されるべき問題がある。

- カメラ間のパラメータの違い
同一の機種のカメラを用いた場合でも内部パラメータが異なる。また設置されるカメラの位置で絞りの量などは光源の位置を考慮して適切に調整される必要がある。調整がうまくいかないと色調などの光学的な要素に大きな違和感を生じる。

- 詳細な位置調整

生成した画像間に不連続部分が生じている。これはカメラの配置が正確でないからである。

不連続や違和感の取り除くためには、一般には視野の重複領域が用いられる。しかし本試作センサの場合、設計上重複領域が有ってはいけない。このため補正が非常に困難になっている。

今後、これらの違和感を解決するのが試作センサの重要な課題である。

歪みや色合いの違いなどの補正には、以下の2つの方法がアプローチが考えられる。

- 完全な同一視点, 配置を実現する。
- 視点のずれをある程度是認し, 生じた重複領域の情報を利用して, 得られた画像をソフトウェア的に補正して違和感の少ない画像を生成する。

前者の方法は完全な同一視点からの全周パノラマ画像を生成できるがその実現方法が難しい。一方, 重複領域のある複数の画像からパノラマ画像を生成する方法は, すでに提案されており, パノラマ画像生成のための市販のソフトウェアが比較的安価に, いくつか販売されている。

重複領域が小さいほど, 視点が一致した画像が得られるが, 複数の画像間の補正が難しい。重複領域が大きいほど画像間の補正は容易になるが, 視点のずれは大きくなり画像内の矛盾が大きくなる。このトレードオフのなかで最も良いバランスのよいカメラ配置を行い, 必要な補正を行う必要がある。

7. むすび

本研究では実環境の映像化を目的として設計、試作した全方位ステレオ画像センサのキャリブレーション法と撮像画像からの全周ステレオパノラマ画像の生成法について述べ、実際に生成された画像を示した。また生成した全周ステレオパノラマ画像を用いて全周パノラマ距離画像や任意の視線方向の透視投影画像を生成した。

試作センサは、6角錐ミラーと6台のカメラからなる撮像系を持ち、複数のカメラを用いて取得することで、高分解能かつビデオレートで全周囲の環境の透視投影画像を撮像することができる。試作センサではこの撮像系を上下対称に2つ配置することによって縦視差を持つステレオ画像を獲得することができる。

本研究では、試作センサにおけるカメラの歪み補正やカメラ配置を行い、実験として実際に実環境の映像を取得した。取得した映像から、全周ステレオパノラマ時系列画像を生成し、試作センサにより、実際に高分解能の全周ステレオパノラマ画像が撮像できることを確認した。さらに任意の視線方向の平面透視投影画像や全周距離画像を生成した。任意の視線方向の平面透視投影画像、全周距離画像はテレプレゼンスや仮想物体との融合といった複合現実感への応用のために利用可能である。

現状での問題点は、撮像される色調が一様でないことや手動による調整のためキャリブレーションの精度に限界があることが挙げられる。ソフトウェア的な調整を行うためには得られる画像上に重複領域があることが望ましいが、1視点の制約を満たさなくなる。このトレードオフを考慮に入れ、最適なカメラ位置を実現する必要がある。また新たに、一視点を実現した状態での、完全にカメラパラメータが同一の画像生成のためのアルゴリズムを考案することが望まれる。

また、実際に仮想物体を合成するためにはステレオ視の精度の改良が必要である。精度の改善のために従来提案されている様々な手法を比較、検討し、より良い手法を採用あるいは提案する必要がある。

今後は、上に述べた課題を解決し、臨場感のある仮想物体を描画するために全周パノラマ画像から光源情報の獲得やユーザの両眼への画像提示のために距離画像から横視差画像の生成を行う必要がある。

謝辞

本研究を進めるにあたり，終始暖かくご指導を頂いた ソフトウェア基礎講座 横矢 直和教授 に厚く御礼申し上げます．また，本学での研究におけるあらゆる面で，多大なご助言，ご協力を賜りましたことを深く感謝致します．

副指導教官としてご助言を頂いた 像情報処理講座 千原 國宏教授 ，ならびに ソフトウェア基礎講座 竹村 治雄助教授 に深く感謝致します．

さらに，本研究へのご助言や研究環境の整備にご協力を頂いた ソフトウェア基礎講座 山澤 一誠助手 ，ならびに ソフトウェア基礎講座 岩佐 英彦助手 に深く感謝致します．

ソフトウェア基礎講座の諸氏には，本研究を進めるにあたり多大なる協力を頂きました．ここに，感謝の意を表します．

また，日々の研究室の活動を支えて下さった，ソフトウェア基礎講座 事務補佐員 福永 博美女史 ならびに 村上 和代女史 に感謝致します．

参考文献

- [1] P. J. Metzger: “Adding Reality to The Virtual”, *Proc. of VRAIS '93*, pp.7–13 (1993).
- [2] A. Ronald: “A Survey of Augmented Reality, ” *Presence*, Vol.6, No.4, pp.355–385 (1997).
- [3] D. Aliga: “Virtual Objects in the Real World, ” *Communication of the ACM*, Vol.40, No.3, pp.49–54 (1997).
- [4] 太田: “複合現実感技術への期待と課題”, 第1回 VR 学会 複合現実感技術研究会 '97.
- [5] 井上, 太田: “複合現実感のための線形的手法による仮想物体像合成”, 信学技報, PRMU(1997-10).
- [6] 玉田, 亀井, 瀬尾: “ビデオ空間と CG 空間の融合による高解像度 3 次元仮想都市空間管理環境の構築”, 日本 VR 学会論文集, Vol.1, pp.167–168 (1996).
- [7] 笹田: “建築設計・都市計画におけるコンピュータグラフィックス”, 情報処理学会誌, Vol.29, No.10, pp.1177–1181 (1988).
- [8] 曾山, 安田, 横井, 鳥脇: “景観シミュレーションのための簡易情景モデル構築と動画生成の一手法”, テレビジョン学会誌, Vol.47, No.12, pp.1647–1655 (1993).
- [9] 佐々木: “テレビにおけるバーチャルスタジオの可能性”, テレビジョン学会誌, Vol.48, No.11, pp.1383–1385 (1994).
- [10] 福井: “サイエンス番組における CG と実写を融合する可視化技術”, 電気・情報関連学会連合大会講演論文集, pp.157–166 (1995).
- [11] S. Feiner, B. MacIntyre, M. Haupt, E. Solomon: “Windows on the World: 2D Windows for 3D Augmented Reality”, *Proc. UIST'93 ACM Symposium on User Interface Software and Technology*, pp.145–155 (1993).

- [12] 仲村, 大隈, 竹村, 横矢: “AR環境における情報の生成・管理手法の一考察”, 情報処理学会ヒューマンインタフェース研究会報告 98-HI-76, pp.31-36 (1998).
- [13] D. Drascic : “Stereoscopic Vision and Augmented Reality”, *Scientific Computing and Automation*, pp.31-34 (1993).
- [14] P. Milgram, S. Zhai, D. Drascic : “Applications of Augmented Reality for Human-Robot Communication”, *Proc. of IROS'93 International Conference on Intelligent Robots and Systems*, pp.1467-1-476 (1993).
- [15] F. Betting, J. Feldmar, N. Ayache, F. Devernay: “A New Framework for Fusing Stereo Images with Volumetric Medical Images”, *Computer Vision. Virtual Reality and Robotics in Medicine '95 Proc.*, pp.30-39 (1995).
- [16] 広瀬, 佐藤: “人工現実感技術を用いた臨場感の合成とその伝達”, 第9回ヒューマンインタフェースシンポジウム, pp.111-116 (1993).
- [17] 尾上, 山澤, 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”, 信学技報, PRMU97-20 (1997-05).
- [18] Z. L. Cao, S. J. Oh and E. L. Hall: “Dynamic omnidirectional vision for mobile robots”, *J. Robotic Systems*, Vol.3, No.1, pp. 5-17 (1986).
- [19] K. B. Sarachik: “Characterizing an indoor environment with a mobile robot and uncalibrated stereo”, *Proc. Int. Conf. Robotics and AutomationIEEE*, pp.984-989 (1989).
- [20] H. Ishiguro, M. Yamamoto and S. Tsuji: “Omni-directional Stereo for Making Global Map”, *Proc. IEEE 3rd Int. Conf. on Computer Vision*, pp.540-547 (1990).
- [21] J. Hong, X. Tan, B. Pinette, R. Weiss and E. M. Riseman: “Image-based homing”, *Proc. Int. Conf. Robotics and Automation*, pp. 620-625 (1991).

- [22] 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*, Osaka, pp. 909–914 (1991).
- [23] 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センサ HyperOmni Vision の提案”, 信学論 (D-II), Vol.J79-D-II, No. 5, pp.698–707 (1996).
- [24] S. K. Nayar: “Catadioptric omnidirectional camera camera”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp.482–488 (1997).
- [25] 井口, 佐藤: 三次元画像計測, 昭晃堂 (1990).
- [26] 山澤, 岩佐, 竹村, 横矢: “ピラミッド型ミラーを用いた高分解能全方位ステレオ画像センサ”, 1997 信学春季全大, D-12-146.
- [27] S. E. Chen: “QuickTime VR - An Image-Based Approach to Virtual Environment Navigation”, *Proc. SIGGRAPH 95*, pp.29–38 (1995).
- [28] 川西, 山澤, 岩佐, 竹村, 横矢: “六角錐ミラーを用いた全方位画像センサにおけるレンズ中心位置合わせ”, 1997 信学秋季全大, D-12-87.
- [29] 川西, 山澤, 岩佐, 竹村, 横矢: “六角錐ミラーを用いた全方位ステレオ画像センサによる高解像度ステレオパノラマ動画の作成”, 信学技報, PRMU97-118 (1997-10)
- [30] 浅田: “CVCV-WG 特別報告 (XI):カメラモデルとキャリブレーション手法”, 情報処理学会 CV 研究会報告, 96-CV-102-9, pp.65–74 (1996).
- [31] 出口: “カメラキャリブレーション手法の最近の動向”, 情報処理学会 CV 研究報告, 93-CV-82-1, pp.1–8 (1993).
- [32] Roger 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, August (1987).

付録

A. Tsai のキャリブレーション法

Tsaiの方法は、ロボットビジョンにおけるハンドアイ較正(hand-eye calibration)問題を扱ったものであり、ロボットの手の座標と目の座標の関係から、非線形最適化問題を解くことによって外部パラメータおよび内部パラメータを同時に求める方法を提案している。以下では、文献 [32] に述べられている方法の概要を紹介する。なお、Tsai のキャリブレーションプログラムは VISION-LIST-ARCHIVE で公開され ftp で入手可能¹ である。

カメラモデルと座標の関係 図 A.1 に示すようにワールド座標系 $O_w-x_w-y_w-z_w$ における点 P の座標を (x_w, y_w, z_w) と表し、カメラ座標系における $O-x-y-z$ における点 P の座標系を (x, y, z) と表す。ただし O はレンズ中心 (ピンホール) を表し、 z 軸はレンズの光軸に一致するように定める。

次に、 x - y 平面に並行で z 座標が f の位置に画像中心が O_i 、座標軸が X - Y の画像平面を考える。理想的なピンホールカメラの場合には、点 P の座標 (x, y, z) は画像平面上では (X_u, Y_u) と表されるが、レンズの幾何学的な歪により実際には画像平面上では (X_d, Y_d) に位置に対応するものとする。そして (X_d, Y_d) のデジタル画像上での離散化された座標を (X_f, Y_f) と表す。ただし、デジタル画像上での原点座標を (C_x, C_y) 、 X 方向 Y 方向の 1 画素の間隔をそれぞれ dx, dy とする。また、 X 方向 Y 方向の画素数をそれぞれ N_{cx}, N_{fy} とする。

点 P のワールド座標 (x_w, y_w, z_w) から、その点のデジタル画像上での座標 (X_f, Y_f) を求める手順は以下のようになる。

Step 1 (x_w, y_w, z_w) から (x, y, z) への変換

回転行列 R と並行移動ベクトル T を用いて表す。

¹<ftp://ftp.osaku-u.ac.jp/vision-list/shareware/code/calibration>

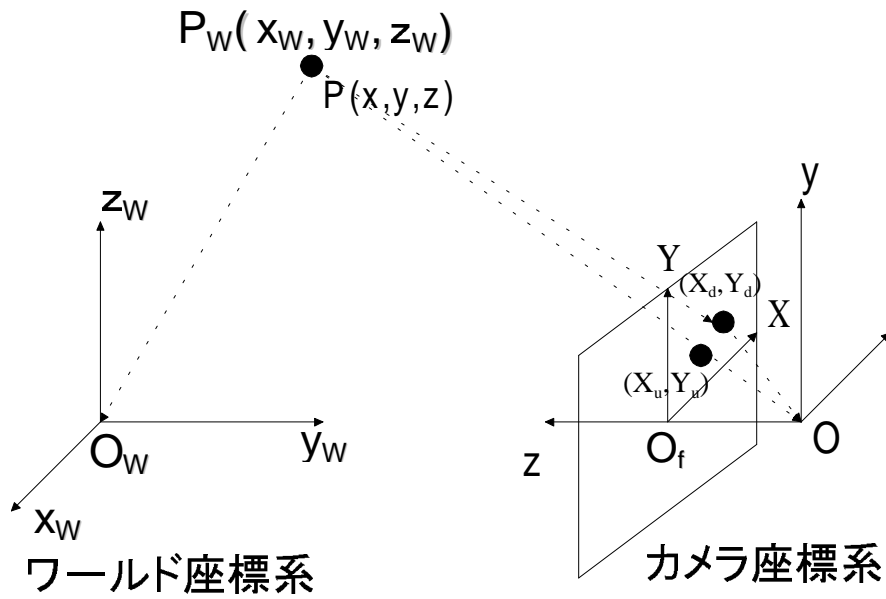


図 A.1 Tsai のピンホールカメラモデル

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} + T \quad (10)$$

ただし,

$$R = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix} \quad (11)$$

$$T = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} \quad (12)$$

Step2 透視変換による (x, y, z) から (X_u, Y_u) への変換

有効焦点距離 f を用いて表す.

$$X_n = f \frac{x}{z} \quad (13)$$

$$Y_n = f \frac{y}{z} \quad (14)$$

Step3 (X_u, Y_u) から (X_d, Y_d) への変換

$$X_d + D_x = X_n \quad (15)$$

$$Y_d + D_y = Y_n \quad (16)$$

ただし, κ_1, κ_2 をレンズ歪係数として

$$D_x = X_d (\kappa_1 r^2 + \kappa_2 r^4) \quad (17)$$

$$D_y = Y_d (\kappa_1 r^2 + \kappa_2 r^4) \quad (18)$$

$$r = \sqrt{X_d^2 + Y_d^2} \quad (19)$$

Step4 (X_d, Y_d) から (X_f, Y_f) への変換 :

$$X_f = s_x d'_x{}^{-1} X_d + C_x \quad (20)$$

$$Y_f = d'_y{}^{-1} Y_d + C_y \quad (21)$$

ただし, s_x はスケール係数を表し

$$d'_x = d_x \frac{N_{cx}}{N_{fx}} \quad (22)$$

以上の関係を整理すると

$$s_x^{-1} d'_x X + s_x^{-1} d'_x X (\kappa_1 r^2 + \kappa_2 r^4) = f \frac{x}{z} \quad (23)$$

$$d'_y Y + d'_y Y (\kappa_1 r^2 + \kappa_2 r^4) = f \frac{y}{z} \quad (24)$$

$$(25)$$

ただし,

$$X = X_f - C_x \quad (26)$$

$$Y = Y_f - C_y \quad (27)$$

$$r = \sqrt{(s_x^{-1}d'_x X)^2 + (d'_y Y)^2} \quad (28)$$

したがって

$$s_x^{-1}d'_x X + s_x^{-1}d'_x X (\kappa_1 r^2 + \kappa_2 r^4) = f \frac{r_1 x_w + r_2 y_w + r_3 z_w + T_x}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (29)$$

$$d'_y Y + d'_y Y (\kappa_1 r^2 + \kappa_2 r^4) = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (30)$$

$$(31)$$

が得られる.

カメラキャリブレーション手順

カメラキャリブレーションとは, 多数の点のワールド座標 (x_w, y_w, z_w) と画像座標 (X, Y) の組が与えられたときに, 外部パラメータとして回転行列 R (未知数 3 個) と平行移動ベクトル T (未知数 3 個), 内部パラメータとして有効焦点距離 f , レンズ歪係数 κ_1, κ_2 , スケール係数 s_x , 画像原点 (C_x, C_y) の全部で 12 個のパラメータを求める問題になる. 以下では, 平面上の多数の点を用いたカメラキャリブレーションの手順について述べる.

1. N 個のキャリブレーション点を撮影した画像から点 P_i の画像座標 (X_{fi}, Y_{fi}) を求める.
2. カメラおよび A/D 変換器の仕様から $N_{cx}, N_{fx}, d'_x, d'_y, d_y$ を決定する.
3. 画像の中心画素を原点 (C_x, C_y) とする.
4. N 点の (X_{di}, Y_{di}) を求める.

$$X_{di} = s_x^{-1}d'_x{}^{-1} (X_{fi} - C_x) \quad (32)$$

$$Y_{di} = d'_y{}^{-1} (Y_{fi} - C_y) \quad (33)$$

5. (x_{wi}, y_{wi}, z_{wi}) と (X_{di}, Y_{di}) の組から $T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x, T_y^{-1}r_4, T_y^{-1}r_5$ を未知数とする線形方程式を解く.

$$(Y_{di}x_{wi}Y_{di}y_{wi} - X_{di}y_{wi})L = X_{di} \quad (34)$$

ただし

$$L = (T_y^{-1}r_1 T_y^{-1}r_2 T_y^{-1}T_x T_y^{-1}r_4 T_y^{-1}r_5)^T \quad (35)$$

6. $T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x T_y^{-1}r_4, T_y^{-1}r_5$ から $r_1, r_2, \dots, r_9, T_x T_y$ を求める.
まず,

$$C = \begin{pmatrix} r'_1 & r'_2 \\ r'_3 & r'_4 \end{pmatrix} = \begin{pmatrix} T_y^{-1}r_1 & T_y^{-1}r_2 \\ T_y^{-1}r_4 & T_y^{-1}r_5 \end{pmatrix} \quad (36)$$

とする. もし, r'_1, r'_2, r'_4, r'_5 がすべて 0 でないなら

$$T_y = \frac{S_r - [S_r^2 - 4(r'_1 r'_5' - r'_4 r'_2')]^{\frac{1}{2}}}{2(r'_1 r'_5' - r'_4 r'_2')^2} \quad (37)$$

を求める. ただし $S_r = r'_1{}^2 + r'_2{}^2 + r'_4{}^2 + r'_5{}^2$ とする.

7. 画像原点 (Cx, Cy) から十分離れた点 (X_{fi}, Y_{fi}) とその対応する点の世界座標 (x_{wi}, y_{wi}, z_{wi}) から T_y の符号を決定する. まず, T_y の符号を正として

$$r_1 = (T_y^{-1}r_1) T_y \quad (38)$$

$$r_2 = (T_y^{-1}r_2) T_y \quad (39)$$

$$r_4 = (T_y^{-1}r_4) T_y \quad (40)$$

$$r_5 = (T_y^{-1}r_5) T_y \quad (41)$$

$$(42)$$

$$T_x = (T_y^{-1}T_x) T_y \quad (43)$$

$$x = r_1 x_w + r_2 y_w + T_x \quad (44)$$

$$y = r_4 x_w + r_5 y_w + T_y \quad (45)$$

を求め, x と X の符号が等しくかつ y と Y の符号が等しいなら T_y の符号を正, それ以外では T_y の符号を負とする.

8. 回転行列 R を決定する. まず,

$$r_1 = (T_y^{-1}r_1) T_y \quad (46)$$

$$r_2 = (T_y^{-1}r_2) T_y \quad (47)$$

$$r_4 = (T_y^{-1}r_4) T_y \quad (48)$$

$$r_5 = (T_y^{-1}r_5) T_y \quad (49)$$

$$T_x = (T_y^{-1}T_x) T_y \quad (50)$$

$$R = \begin{pmatrix} r_1 & r_2 & \sqrt{1 - r_1^2 - r_2^2} \\ r_4 & r_5 & s\sqrt{1 - r_4^2 - r_5^2} \\ r_7 & r_8 & r_9 \end{pmatrix} \quad (51)$$

ここで, $s = -\text{sgn}(r_1r_4 + r_2r_5)$ であり, r_7, r_8, r_9 は正規直交行列の性質を用いて求める. $\text{sgn}(\cdot)$ は符号を表す.

9. まず, $\kappa_1 = \kappa_2 = 0$ とする. N 点のキャリブレーション点について f と Tz を未知数として

$$(y_i - d_y Y_i) \begin{pmatrix} f \\ Tz \end{pmatrix} = w_i d_y Y_i \quad (52)$$

を解く. ただし,

$$y_i = r_4 x_{wi} + r_5 y_{wi} + T_y \quad (53)$$

$$w_i = r_7 x_{wi} + r_8 y_{wi} \quad (54)$$

これによって f と Tz の初期値を得る.

10. f と Tz の初期値を使って,

$$d_y Y + d_y Y (\kappa_1 r^2 + \kappa_2 r^4) = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z} \quad (55)$$

を非線形最適化問題として $f, T_z, \kappa_1, \kappa_2$ を求める.

ピンホールカメラモデルを用いたカメラキャリブレーションで有効焦点距離 f と呼んでいるものは、レンズ固有のパラメータではなくモデル上でのレンズと撮像面との距離を表すパラメータである. この f と 1 画素の大きさパラメータ d_x, d_y を与えることによって、モデル上での画像と実際の画像の対応付けが可能となる. すなわち、 f と d_x, d_y を独立に決定することはできない. Tsai の方法では、カメラの仕様から 1 画素の大きさパラメータを既知として与え、 f を求めている.