

博士論文

全方位型マルチカメラシステムを用いた高臨場感
テレプレゼンスシステムの構築手法に関する研究

池田 聖

2006年3月24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
博士(工学) 授与の要件として提出した博士論文である。

池田 聖

審査委員：

横矢 直和 教授 (主指導教員)

千原 國宏 教授 (副指導教員)

山澤 一誠 助教授 (副指導教員)

全方位型マルチカメラシステムを用いた高臨場感 テレプレゼンスシステムの構築手法に関する研究*

池田 聖

内容梗概

実画像提示により遠隔地の臨場感を再現するテレプレゼンスは、エンターテインメントや教育、医療などへの応用が期待されている。テレプレゼンスでは、要求される視点位置や視線方向に応じた実画像の実時間提示により実環境における視界を再現することでユーザに臨場感を与える。テレプレゼンスシステムは利用形態によって、ユーザの仮想視点を能動的に移動する方式と受動的に移動する方式に大別されるが、両方式共に、ユーザに高い臨場感を与えることが課題とされている。本研究では、受動的・能動的視点移動型の両方式のテレプレゼンスシステムにおいて広域実環境を仮想化し、高い臨場感を再現するための画像生成・提示手法の開発を目的とする。ユーザに高い臨場感を与えるために、高解像度、広視野角な全方位型マルチカメラシステムを用いて生成した全天球動画像を用い、時系列的な連続性を維持した画像提示を基本方針とする。特に、能動的視点移動型のテレプレゼンスシステムの構築では、視点移動量の入力インタフェースやそれに対応した画像提示手法の選定についても考慮する必要がある。そこで本研究では、能動的視点移動型のテレプレゼンスシステムに対しては、上記の両方式に共通する手法に加えて、歩行動作インタフェースを用いた画像提示システムを用い、利用時の歩行動作による頭部位置の変動および歩行動作と相関の無い撮影時の全方位型マルチカメラシステムの移動速度や姿勢の変動を考慮した画像提示により臨場感の向上を図る。まず 1 章では、テレプレゼンスの技術的課題と従来研究を概

* 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 博士論文, NAIST-IS-DD0361003, 2006 年 3 月 24 日.

観し，本研究の目的と意義を明らかにする．2章では，高精度な全天球動画の生成および全方位型マルチカメラシステムの移動速度・姿勢の推定に必要となる全方位型マルチカメラシステムの内部パラメータの推定法について述べる．3章では，撮影中の全方位型マルチカメラシステムの位置・姿勢を表す外部パラメータの推定法について述べる．外部パラメータの推定では，全方位型マルチカメラシステムが広範囲に移動する場合においても，推定結果に蓄積誤差が生じるのを避けるため GPS の測位情報と画像中の特徴点追跡手法を併用する．4章では，これらの全方位型マルチカメラシステムの内部・外部パラメータに基づく画像提示法および歩行動作インタフェースを用いた画像提示システムについて述べる．最後に5章で，全体を総括し，今後の展望について述べる．

キーワード

テレプレゼンス，実画像提示，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，臨場感

Construction of Telepresence Systems Using an Omnidirectional Multi-camera System*

Sei Ikeda

Abstract

Telepresence system using real images provides us with a rich sense of presence in a remote site. The sense of presence is created by reproducing a field of view according to the change of position and direction of user's view. Telepresence system can be classified into two types according to whether the movement of user's virtual view position is active or passive. For both types, it is required to provide a user with a rich sense of presence without increasing human cost for generating image contents. The purpose of this study is to develop image generation/presentation methods for passive and active telepresence systems to provide a high-quality sense of presence. To reproduce a rich sense of presence, high-resolution and omnidirectional videos acquired with an omnidirectional multi-camera system are presented, and an image presentation which maintains the temporal continuity of the videos are employed. The active type especially needs to consider some other issues including user interface to decide the movement of view point and image presentation methods suitable for the user interface. For an active telepresence system, an image presentation system using locomotion interface is proposed. In this system, the image is presented considering the variation in head position caused by user's locomotion and unintended movement of omnidirectional multi-camera system in image acquisition. In this dissertation,

* Doctoral Dissertation, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0361003, March 24, 2006.

Chapter 1 gives a perspective of the study in the area of telepresence. Chapter 2 describes a method for estimating intrinsic camera parameters required for generation of spherical video and estimation of speed and pose of omnidirectional multi-camera system. Chapter 3 describes a method for estimating extrinsic parameters which represent the position and pose of omnidirectional multi-camera system. To avoid accumulative errors in parameter estimation, GPS position of information is combined with a feature tracking-based method. Chapter 4 describes an immersive display system using a locomotion interface and an image generation method based on estimated parameters. Chapter 5 gives conclusions.

Keywords:

Telepresence, Image-based Rendering, Omnidirectional Multi-camera System, Camera Calibration, Sense of Presence

目次

1. 序論	1
1.1 テレプレゼンス技術の概観	1
1.2 画像提示方式の分類	4
1.3 テレプレゼンスシステムの技術的課題	5
1.4 全方位画像センサと取得画像の特徴	8
1.4.1 単眼カメラを用いる全方位画像センサ	8
1.4.2 多眼カメラを用いる全方位画像センサ	9
1.4.3 全方位画像センサのテレプレゼンスへの利用	10
1.5 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研究	12
1.5.1 カメラ内部パラメータ推定法	12
1.5.2 カメラ外部パラメータ推定法	13
1.6 本研究の位置づけと方針	14
1.7 本論文の構成	16
2. 全方位型マルチカメラシステムのカメラ内部パラメータ推定	17
2.1 緒言	17
2.2 全方位型マルチカメラシステム Ladybug	17
2.3 カメラ内部パラメータの推定	18
2.3.1 幾何学的キャリブレーション	18
2.3.2 光学的キャリブレーション	27
2.4 全天球画像の生成	29
2.5 視差の影響	30
2.6 実験	32
2.6.1 幾何学的キャリブレーション	32
2.6.2 光学的キャリブレーション	37
2.6.3 全天球動画画像の生成	43
2.6.4 全天球画像の生成精度評価	46

2.6.5	受動的視点移動型テレプレゼンスシステムの試作	47
2.7	結言	49
3.	全方位型マルチカメラシステムのカメラ外部パラメータの推定	50
3.1	緒言	50
3.2	カメラと GPS 受信機の幾何学的モデル	51
3.3	GPS 測位値を用いた外部パラメータ推定	53
3.3.1	特徴点と GPS による誤差の定義とその最適化	53
3.3.2	各処理の詳細	56
3.4	実験	59
3.4.1	シミュレーションによる定量評価	60
3.4.2	実環境下での実験	62
3.5	結言	69
4.	歩行装置を用いた能動的視点移動型テレプレゼンスシステム	72
4.1	緒言	72
4.2	システムの設計方針	73
4.3	システムの実装	75
4.3.1	歩行装置を用いた画像提示システム	75
4.3.2	カメラパラメータに基づく広視野角動画コンテンツの生成	79
4.4	実験	81
4.4.1	画像提示結果の定性的評価	83
4.4.2	主観評価による有効性の検証	86
4.5	結言	88
5.	結論	90
5.1	本論文のまとめ	90
5.2	今後の展望	91
	謝辞	93

目 次

1	頭部に取り付けられた CRT モニタ [CB61]	2
2	全方位画像の例 (撮影: NHK エンタープライズ 21)	9
3	全方位型マルチカメラシステムの例	11
4	Ladybug のカメラヘッド (左) と蓄積ユニット (右)	18
5	Ladybug による入力画像 (上方向 (右下) と水平方向 (その他))	19
6	Ladybug による画像の取り込み	19
7	世界座標系とカメラ座標系の関係	21
8	キャリブレーションボード	25
9	キャリブレーションボードの配置	25
10	マーカの検出	25
11	$\cos^4 \theta$ 現象のレンズモデル	28
12	全天球画像の生成手順	30
13	投影球面から全天球画像への展開	31
14	視差の影響	31
15	キャリブレーションに使用したマーカの画像例	34
16	キャリブレーションボード上の格子点の検出例	35
17	レンズ歪み補正結果	35
18	カメラの位置および姿勢の推定結果	36
19	明度低下現象の補正結果	37
20	色調補正の効果	38
21	色調補正前後のヒストグラム (R 成分)	39
22	色調補正前後のヒストグラム (G 成分)	40
23	色調補正前後のヒストグラム (B 成分)	41
24	色調補正に用いた入力画像	42
25	Ladybug による屋外の撮影	43
26	入力画像 (奈良 あやめ池遊園地: 上方向 (右下) と水平方向 (その他))	44
27	全天球動画の 1 フレーム	45
28	隣り合うカメラの画像	46

29	円形マーカ	46
30	テレプレゼンスシステムの外観	48
31	座標系と誤差の定義	52
32	提案手法の処理手順	54
33	重み変動の効果	55
34	重み係数 ω の決定	56
35	最適化の範囲	58
36	シミュレーション環境	60
37	カメラ位置の推定誤差	63
38	カメラ光軸方向の推定誤差	63
39	全方位型マルチカメラシステムと GPS 受信機	64
40	カメラ外部パラメータの推定結果	65
41	推定された外部パラメータを用いたマッチムーブの結果	66
42	パラメータ推定の誤差によって仮想物体が静止しないシーン	67
43	GPS 測位値に含まれる標準偏差	69
44	分割最適化における各シーケンスの分割範囲	70
45	分割最適化の精度	70
46	画像提示システムの外観	73
47	歩行装置を用いた画像提示システム	76
48	歩行装置の座標系	77
49	頭部のレンジデータと検出結果	78
50	足部のレンジデータと検出結果	78
51	取得した動画像の 1 フレーム	81
52	蓄積されている動画像の 1 フレーム	82
53	画像提示結果	84
54	頭部位置を考慮した画像提示結果	85
55	複数の経路が交差する環境の例	91

表 目 次

1	受動的・能動的視点移動型テレプレゼンスシステムの比較	5
2	全方位画像センサの特徴	11
3	Ladybug の仕様	20
4	全方位型マルチカメラシステムの内部パラメータ一覧	23
5	推定された内部パラメータ	33
6	各カメラの再投影誤差 [pixel]	36
7	画像生成に関する誤差 $\epsilon \times 10^{-2}$ [rad]	47
8	テレプレゼンスシステムの機器構成	48
9	観測データに付加する誤差	61
10	外部パラメータの初期値に付加する誤差	61
11	評価結果	88

1. 序論

インタラクティブな映像提示により遠隔地の疑似体験を提供するテレプレゼンスシステムは、教育や医療、娯楽などの様々な分野で応用が期待されている。しかし、遠隔地における人間の感覚を完全に再現することは難しく、現在も盛んにテレプレゼンスシステムに関する研究が続けられている。本章では、まず、従来のテレプレゼンス技術を外観し、臨場感の高いテレプレゼンスシステムを実現するために、全方位動画像が利用されてきたことを説明する。次に、テレプレゼンス技術に関する幾つかの技術的課題を取り上げ、本研究の位置づけと方針を明らかにする。

1.1 テレプレゼンス技術の概観

テレプレゼンス (telepresence) という言葉は、1980年頃、MIT (Massachusetts Institute of Technology) の Marvin Minsky が初めて用いたと言われている。Minsky は、科学技術や空想に関する雑誌 OMNI の記事 [Min80] の中で、宇宙探査における遠隔作業技術の理想形として「自分の手のような感覚を持ち、手のように動く遠隔操縦システム」を提唱し、これをテレプレゼンス技術と呼んだ。後に、LaueIらや Sheridan らは、この言葉をさらに一般的に解釈し、テレプレゼンス技術は「空間的もしくは時間的に離れた場所を身体的に体験する」技術であるとした [LF92]。また、遠隔地のロボットの操縦を目的とした遠隔操縦 (teleoperation) [米田 00, GADS04]、遠隔地にいる人とのコミュニケーションを目的とした遠隔会議 (teleconference) [Jou02, J. 04]、遠隔地の遠隔監視 (surveillance) [PGA⁺00] など様々な分野を示す言葉として単にテレプレゼンスという単語が使われる場合がある。そこで本論文では、人間が視覚刺激から最も多くの情報を得ていることを踏まえ「空間的もしくは時間的に離れたある地点の周囲環境からの視覚刺激を再現し、ユーザに遠隔地にいる感覚を与えるシステム」の総称をテレプレゼンスシステムと呼ぶこととする。

上記の意味でのテレプレゼンスシステムの最も古い形の1つは、1958年に Philco 社の Comeau らが開発した遠隔監視システム HEADSIGHT [CB61] である。この

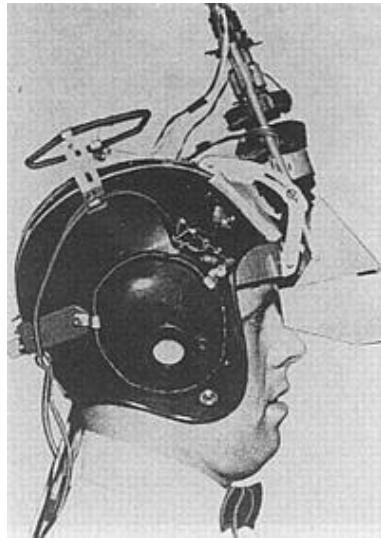


図 1 頭部に取り付けられた CRT モニタ [CB61]

システムでは、ユーザは、図 1 に示すように頭部に固定された CRT モニタにより、遠隔地のアクティブカメラからリアルタイムに伝送された映像を見ることが出来る。頭部の回転に応じてカメラが回転制御されるため、ユーザが遠隔地での首振り動作時に見える視界と同様の映像を見ることができ、その場所に居るという感覚を得ることが出来る。しかし、このシステムでは、遠隔地までの距離が大きくなると映像伝送の遅延が無視できなくなり、大きな違和感が生じる問題があった。

視線変化の再現

視線変化における遅延の問題に対して、全方位画像センサと呼ばれる水平方向 360 度の視野を撮影することが可能な画像センサが多数提案されている [ZSE86, SE87, Mov99, JXB⁺91, YYM02, 山澤 02, YKNH98, SKN03, TKHN98, FC200, Poi02, HKCY00]。全方位画像センサは、360 度の視野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、伝送されてくる動画像に遅延が生じてても、見回しによる遅延が殆ど発生しないため、臨場感の高いテレプレゼンスが実現できる [山澤 98, 森田 05]。このような画像センサを移動ロボットなどに搭載する場

合においても，遅延の少ない視線変化は再現できるが，視点移動に関しては遅延の発生が避けられない．

視点移動の再現

遅延の少ない視点移動を再現した最も古いテレプレゼンスシステムの1つは，LipmanによるAspen Movie-Map [A. 80]である．このシステムでは，予め複数の経路上を移動しながら動画像を撮影し，レーザディスクに蓄積された画像フレームの中からユーザの要求する経路の動画像がインタラクティブに提示される．ChenによるQuickTime VR [Che95]では，2次元的に分布する複数地点での全方位画像を蓄積しておき，ユーザが指定する位置および方向の画像を提示することにより，ユーザの視点移動と視線変化が再現される．ただし，QuickTime VRでは，全方位画像を生成するために各地点における異なる方向の画像を複数取得し，これらの画像の重複領域を利用した画像生成法が用いられていたため，地点間の移動時に提示される画像が不連続になり違和感が発生する．また，TaylorによるVideoPlus [C. 02]では，全方位動画像を利用することで，Aspen Movie-Map [A. 80]のような複数の経路上の連続的な視点移動とQuickTime VR [Che95]のような自由な視線変化の両方を再現できる．

これら3つのシステム [A. 80, Che95, C. 02]では，提示する各画像にその画像の取得位置が割り当てられ，異なる時刻に取得された画像が，あたかもユーザが要求した時刻に取得した画像であるかのように提示される．そのため，画像内に移動物体が映る場合などの動的な環境は正しく表現されないが，殆ど遅延が発生することなく視点移動を再現することができる．

上記の3つのシステムよりも，さらに自由な視点移動を再現する技術として，新視点画像を利用するシステムについて述べる．KanadeらのVirtualized Reality [TPP97]では，環境中に固定された複数のカメラからの画像から新しい視点位置からの画像を生成するシステムを提案している．これに類似するものとして，移動するカメラの映像から新しい視点の画像を生成し提示するシステム [THKM00, DTF⁺01] や，ステレオカメラにより新視点画像を生成するシステム [JHNK00] などがある．これらのシステムで生成される新視点画像をユーザに提示することで，

自由度の高い視点移動が可能となり，ユーザ視点を経路上に固定されることはなくなり，違和感を軽減できると考えられる．しかしながら，システムで用いられる新視点画像生成手法は，環境の形状や色が単純かもしくはこれらが十分制御されていることが前提となっており，自然風景のような複雑な環境では，画像劣化が顕著になるためテレプレゼンスシステムでの利用には適していない．以降，本論文では，この方式のテレプレゼンスシステムについては，考えないものとする．

以上のように，臨場感の高いテレプレゼンスシステムを実現するために，実画像の提示により遅延なく視線変更や視点移動を再現することが主な課題の1つとして取り上げられてきた．視線変更や視点移動を再現するテレプレゼンスシステムに利用される全方位画像センサは，これらの課題の解決に重要な役割を果たすと考えられる．以降，全方位画像センサの利用を前提にテレプレゼンスシステムに関するの従来研究および技術的課題についての議論を進める．

1.2 画像提示方式の分類

前節で述べたように，テレプレゼンスシステムは，それぞれ異なる特徴を有する．特に，全方位動画像を用いたテレプレゼンスシステムは，視点移動の形態により以下の2種類に大別できる．

受動的視点移動型テレプレゼンスシステム [山澤 98, 森田 05]

ユーザは自ら視点移動を要求せずに遠隔環境中を静止，または移動する感覚が得られる．バスや電車，ジェットコースターなど自ら制御しない乗り物に乗っている感覚を得るアプリケーションに適する．フレーム間の時間関係は保存されるため静的な環境と動的な環境の両方を再現することができる．

能動的視点移動型テレプレゼンスシステム [C. 02, MAS⁺04]

ユーザは自ら視点移動を要求し，遠隔環境中を自由な速度で移動する感覚が得られる．自ら制御している乗り物に乗っている感覚や歩行移動する感覚を得るアプリケーションに適する．フレーム間の時間関係は保存されないため動的な環境は再現できない．

両方式のテレプレゼンスシステムの特徴を表1にまとめる．ただし，ComeauらのHEADSIGHTについては，両方式に分類されないものとし比較する．表中の動作の制約では，印は，遅延なく視線変化または視点移動できることを示し，×印は，遅延が生じることを示している．環境の制約では，印はそれぞれの環境を再現できることを示し，×印は再現できないことを示す．以上のように，受動的および能動的視点移動型テレプレゼンスシステムは，アプリケーションに応じて使い分ける必要がある．

1.3 テレプレゼンスシステムの技術的課題

人間は実空間から，感覚器を介して何らかの情報を受け取り，眼前の空間を現実空間であると認識している．特に，視覚情報は全体の約80%を占めると言われており，最も重要である．視覚情報は初期過程において，色，テクスチャ，運動，位置など複数の属性に分離され，並列に処理されると言われている [TG00]．

表1 受動的・能動的視点移動型テレプレゼンスシステムの比較

	動作の制約		環境の制約	
	視線変化	視点移動	静的環境	動的環境
HEADSIGHT [CB61]	×	×		
受動的視点移動型		×		
能動的視点移動型				×

これらの属性をより忠実に再現することで，臨場感の高いテレプレゼンスシステムが実現されると考えられる．

ユーザの運動に関係しない視覚属性

ユーザの運動に関係のない色やテクスチャの属性は，撮像装置や提示装置に関係している．このことから臨場感の高いテレプレゼンスシステムに使用される撮像装置や提示装置には，以下の条件が要求される．

- (i) 解像度が高い： 視点を中心とした角度分解能が高いこと．解像度が低い場合には，現実空間のテクスチャなどの細かい部分が再現されない．
- (ii) 解像度が一様である： 視線方向により，解像度が変化しないこと．見ている方向により提示画像の解像度が異なる場合は，現実空間のテクスチャなどの細かい部分が再現されない部分が目立つ．
- (iii) 視野角が広い： 視線を固定したときに提示画像が視野を覆う範囲が広いこと．視野角が狭い場合，画像の提示領域とは異なるディスプレイのフレーム部分などが視界に入り，現実空間の色やテクスチャの再現されない部分が生じる．
- (iv) 一点透視投影画像である： 一点透視投影画像でない場合，画像上の全ての画素の投影中心が一致しないため視差が生じ，撮影対象がぼけたり多重に提示され，現実空間の色やテクスチャと異なるパターンが生じる．

以上の条件をより多く満たす画像センサを利用して画像を取得し，没入型ディスプレイに提示するためには，画像センサのカメラ内部パラメータが必要となる．

ユーザの運動に関係する視覚属性

運動や位置に関する属性は，環境と自己の間の位置・姿勢関係を示すものであり，画像の生成法や提示法に関係している．このことから臨場感の高いテレプレゼンスシステムに使用される，画像の生成法や提示法には，以下の条件が要求される．

- (v) 正しい自己位置感覚の再現： ユーザが知覚するオプティカルフローが要求した視点移動を正しく再現する．撮影時にカメラの姿勢変動が生じる場合，ユーザが要求した視点移動量と異なるオプティカルフローが知覚される．
- (vi) 正しい自己姿勢感覚の再現： ユーザが知覚するオプティカルフローが要求した視線変更を正しく再現する．撮影時にカメラの移動速度が変化する場合，ユーザには要求した視線変更量と異なるオプティカルフローが知覚される．

受動的視点移動型テレプレゼンスシステムにおいては，これらの自己位置・姿勢の再現は，単純に動画像を取得時と同じフレームレートで提示することで達成されるため，特に問題にならないことが多い．これに対して，能動的視点移動型テレプレゼンスシステムにおいては，ユーザの要求する移動速度に応じて提示する動画像のフレームレートを制御する必要があり，撮影時の画像センサの移動速度や姿勢の変動の周期はこのフレームレートに依存し，不自然な映像となる．運動と位置の属性に関する条件をより多く満たすには，画像取得時の画像センサの位置や方向を正確に把握し，これを除去した上でユーザの要求する位置や姿勢の画像を提示しなければならない．

以上の条件を満たすテレプレゼンスシステムを構築するためには，以下に示す2つの技術的な課題を解決しなければならない．

- 画像センサの内部パラメータ推定
- 撮影時の画像センサの位置・姿勢の推定

以下，1.4節では，全方位画像センサの特徴について述べ，全方位型マルチカメラシステムと呼ばれる画像センサがテレプレゼンスシステムの構築に最も適していることを明らかにする．また，1.5節では，全方位型マルチカメラシステムのためのキャリブレーションに関する従来研究について説明する．最後に，1.6節で本研究の方針について述べる．

1.4 全方位画像センサと取得画像の特徴

テレプレゼンスシステムに利用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。本節では、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し、前節で述べた臨場感の高いテレプレゼンスシステムに必要な条件を最も満たす画像センサを明らかにする。

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。

1.4.1 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： [ZSE86, SE87, Mov99] 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度をカバーする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。

曲面ミラーを用いる手法： [JXB⁺91, YYM02, 山澤02] 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 [JXB⁺91]、円錐 [YYM02]、双曲面 [山澤02]、放物面 [S. 97] 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



図 2 全方位画像の例 (撮影: NHK エンタープライズ 21)

がある。ただし、図 2 に示すように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。

これらの方式では、常に幾何学的および光学的な不連続が生じない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 [KHN02] 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。

1.4.2 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、図 3 に示すように、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: [TKHN98, FC200] 複数のカメラが鉛直上向きに配置されており、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 度の画像を高解像に取得可能である。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る。鉛直方向の視野が狭く、上下方向の撮影は出来ない。

カメラを外向きに配置する手法: [Poi02, HKCY00] 複数のカメラが放射状外向き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では生成画像上に幾何学的な不連続が生じることがある。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。

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。

1.4.3 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 2 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



(a) 全方位ステレオ画像センサ [TKHN98] (b) SOS [HKCY00]

図 3 全方位型マルチカメラシステムの例

表 2 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の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から違和感のない全方位動画像を生成し、テレプレゼンスシステムに利用するためには、高い精度でのカメラキャリブレーションが必要である。

1.5 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 従来研究

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は、撮影時に変化しないカメラ内部パラメータの推定と、撮影時のカメラの位置と姿勢を表すカメラ外部パラメータの推定がある。

1.5.1 カメラ内部パラメータ推定法

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内部パラメータの推定手法について述べ、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 [TKHN98] は、図 3(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。そのため生成した生成画像中に不連続な部分が生じた。

これに対して、戎野ら [戎野 99] や棚橋ら [棚橋 02] 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 [TKHN98] 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 [JHNK00, 戎野 00] 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 3(b) のような全方向ステレオシステム

(SOS)[HKCY00] を没入型 6 面ディスプレイ COSMOS[TMY98] 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、プロジェクタ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 [佐藤 03]。

1.5.2 カメラ外部パラメータ推定法

従来から、広範囲に移動するカメラの外部パラメータの推定手法が多数提案されている [FZ98, PKV⁺00, Dav03, SKYT02, CMC03, VLF04, GF03]。しかし、動画像のみから外部パラメータを推定する手法 [FZ98, PKV⁺00] では、外部パラメータの推定結果に蓄積誤差が発生するため、広範囲に移動する場合に著しく推定精度が低下する問題がある。また、推定される外部パラメータは、フレーム間での相対的な位置・姿勢関係を表すものであり、スケールについては不定である。そのため、複数回に分けて撮影した動画像の外部パラメータの統合や、閉曲線上の経路で撮影した動画像の外部パラメータ推定のように、同一シーンが連続しないシーケンス上に現れる場合、推定されるパラメータ間の不整合が問題となる。

これらの問題は、従来から、動画像と現実環境に関する事前知識の使用 [Dav03, SKYT02, CMC03, VLF04] や、動画像を用いずに複数のセンサから得られる値を直接外部パラメータとすること [GF03] により解決されてきた。環境の事前知識を用いる手法では、三次元的な位置関係が既知のランドマーク群 [Dav03, SKYT02] や、

三次元 CAD モデル [CMC03, VLF04] が用いられている。文献 [Dav03, SKYT02] の手法では、三次元位置が既知のランドマークと三次元位置が未知の画像特徴点 (以後、特徴点) をフレーム毎に追跡し、その再投影誤差を最小化することで外部パラメータを推定する。文献 [CMC03, VLF04] の手法では、現実環境の三次元 CAD モデルの輪郭と入力画像上のエッジを対応付けることで外部パラメータを推定する。これらの手法では、ランドマークや三次元 CAD モデルの現実環境での三次元位置情報を用いることで蓄積誤差の問題を解決している。しかし、ランドマークや三次元 CAD モデルに含まれる三次元位置情報を屋外で広範囲に取得することは困難であり、また、基準とする三次元情報の獲得のためには、人手による作業が必要であるという問題がある。

一方、GPS、磁力計、2軸角度センサーの値から算出される値を、直接動画像のカメラ外部パラメータとする手法 [GF03] では、RTK-GPS を用いることによって、誤差の蓄積しない数 *cm* 精度の位置情報 (以後、GPS 測位値) を獲得できる。しかし、センサの組み合わせによってカメラの外部パラメータを推定する手法では、カメラとセンサ類の位置関係が必要であるが、これを正確に計測することは難しい。また、一般的な GPS では、位置情報の取得レートが動画像より低い 1Hz であり、カメラの運動に含まれる高周波成分を復元することは難しい。

1.6 本研究の位置づけと方針

本論文では、屋外などの複雑な環境下での広範囲な移動感覚を再現する、臨場感の高いテレプレゼンスシステムの実現を目的としている。前節までの議論より、以下の 2 つを本研究の基本方針とする。

- (1) 全方位型マルチカメラシステムによる高解像度な全方位動画像の取得
- (2) 時系列的連続性および画素間の連結性を維持した動画像の提示

この方針のもとで、全方位型マルチカメラシステムの内部パラメータおよび外部パラメータの推定手法を提案し、これらのパラメータに基づき生成した画像を用いることで、臨場感の高い受動的および能動的視点移動型テレプレゼンスシステムを構築する。

全方位型マルチカメラシステムのカメラ内部パラメータの推定

受動的および能動的視点移動型の両方式のテレプレゼンスシステムの構築において必要である，全方位型マルチカメラシステムのカメラ内部パラメータ推定法を提案し，推定した内部パラメータを基に全方位動画像を生成する．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．全方位画像の生成には，上記のキャリブレーションの結果を用いて，全方位型マルチカメラシステムの特性を考慮した画像のつなぎ合わせ（スティッチング）を行う．これにより，幾何学的および光学的にシームレスで高解像度な全方位動画像を生成を目指す．

全方位型マルチカメラシステムのカメラ外部パラメータの推定

臨場感の高い能動的視点移動型テレプレゼンスシステムの構築に必要な，全方位型マルチカメラシステムの外部パラメータの推定手法を提案し，実験により精度検証を行う．提案手法では，動画像中の特徴点の追跡と GPS 測位値を用いた最適化に基づきカメラ外部パラメータを推定する．特徴点の追跡では，ランドマーク・特徴点追跡による手法 [SKYT02] と同様に，特徴点の仮対応づけから推定される暫定的な外部パラメータを用いて特徴点の画像上での探索範囲を限定し，誤追跡を防ぐ．GPS 測位値を用いた最適化では，動画像から推定される外部パラメータに基づく特徴点の再投影誤差および外部パラメータから推定される GPS 受信機の位置の推定誤差を同時に最小化することで誤差の蓄積を防ぐ．

受動的・能動的テレプレゼンスシステムの構築

受動的視点移動型テレプレゼンスシステムの構築では，上述の手法により推定されたカメラ内部パラメータを用いて全方位画像を生成し，自由に視線変更可能

なテレプレゼンスシステムを試作する．従来のシステムよりも，高い解像度および広い視野角の画像を提示するテレプレゼンスシステムを実現する．

能動的視点移動型のテレプレゼンスシステムの構築に対しては，歩行装置を用いたシステムを試作する．このシステムでは，受動的視点移動型テレプレゼンスシステムにおいて使用されるカメラ内部パラメータに加えて，カメラ外部パラメータを用いることで，ユーザの環境に対する位置および姿勢を再現する．すなわち，利用時の歩行動作による頭部位置の変動および歩行動作と相関の無い撮影時のカメラシステムの移動速度や姿勢の変動を考慮した画像提示により臨場感の向上を図る．

1.7 本論文の構成

本論文の構成は以下の通りである．2章では，全方位型マルチカメラシステムのカメラ内部パラメータの推定法について述べ，推定されたパラメータに基づく全天球画像の生成手法について述べる．また，生成した全天球画像を利用した受動的視点移動型テレプレゼンスシステムについて述べる．3章では，全方位型マルチカメラシステムのカメラ外部パラメータの推定法について述べる．GPS測位と画像上の自然特徴点追跡により広範囲を移動する全方位型マルチカメラシステムのカメラ外部パラメータの推定が可能であることを示す．4章では，歩行装置を用いた能動的視点移動型テレプレゼンスシステムについて述べる．推定したカメラ外部パラメータを利用した全天球画像の提示により臨場感の向上を確認する．5章では，本論文のまとめと今後の課題を述べる．

2. 全方位型マルチカメラシステムのカメラ内部パラメータ推定

2.1 緒言

本章では, 全方位型マルチカメラシステムの内部パラメータの推定について述べる. 全方位型マルチカメラシステムの内部パラメータとは, カメラシステムの内部構造を表す変数であり, 撮影中に変化するカメラシステムの位置・姿勢に対して, 撮影時に変化しないものを意味する. 複数のカメラにより構成されるカメラブロックの位置・姿勢関係, 個々のカメラ入力画像中の歪みや焦点距離を表すパラメータを推定する幾何学的キャリブレーションと, カメラ間の相対的な色調を補正し, 個々のカメラの明度低下現象の程度を表すパラメータを推定する光学的キャリブレーションがある. これらのカメラ内部パラメータは, 提示画像である全天球画像の生成や第3章で述べる外部パラメータの推定に使用されるため, 精度良く推定されなければならない. 本研究では, これらのパラメータは撮影中は変化しないものと仮定し, 全てのパラメータを予め専用パターン等を用いて推定しておくことで, 利用時に取得する画像の内容に依存せずに一定の提示画像の生成や外部パラメータの推定に利用することを考える.

以下本章では, 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 Ladybug を例に挙げ, その特徴や取得画像等を示す. 次に, 提案するカメラ内部パラメータの推定手法について説明し, 実験により有効性を示す. また, 受動的視点移動型テレプレゼンスシステムの試作についてのべ, 生成した全天球動画像を用いることにより臨場感の高いテレプレゼンスが実現できることを示す. 最後に本章のまとめと課題について述べる.

2.2 全方位型マルチカメラシステム Ladybug

全方位型マルチカメラシステムの例として本研究で使用する Point Grey Research 社製の Ladybug [Poi02] について説明する. Ladybug は図4(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



図 4 Ladybug のカメラヘッド (左) と蓄積ユニット (右)

ラヘッドと図 4(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 3 に示すように、Ladybug は図 5 のような横 768 ピクセル、縦 1,024 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 の動画像として約 20 分間蓄積することが可能なシステムである。図 6 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、オンラインで PC に記録する場合は、6 個のカメラから同時に最大 5fps で画像を取り込むことができる。

2.3 カメラ内部パラメータの推定

2.3.1 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、全方位型マルチカメラシステムを構成する各カメラに対して一般的な単眼カメラの内部パラメータおよび外部パラメータパ

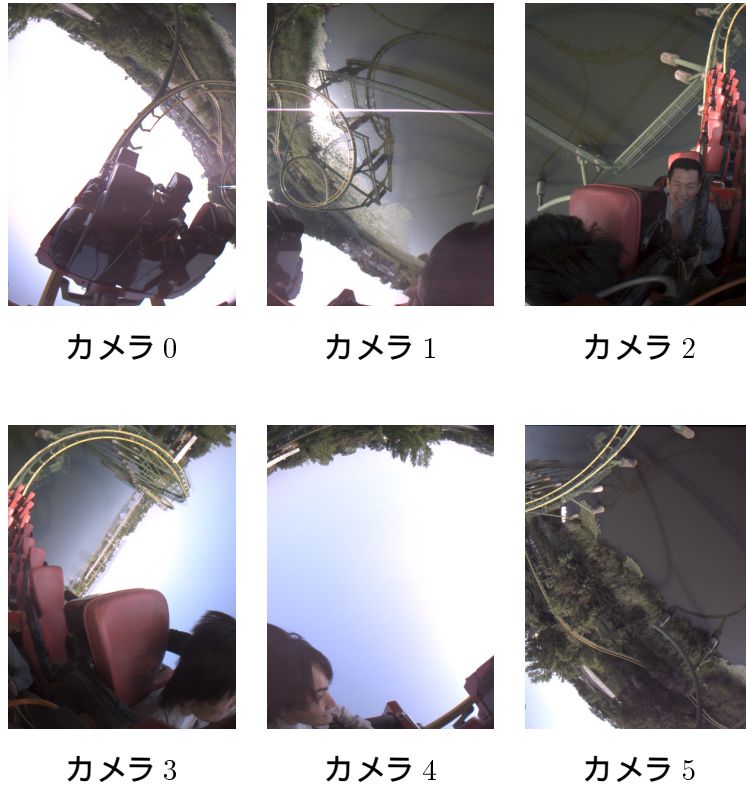


図 5 Ladybug による入力画像 (上方向 (右下) と水平方向 (その他))

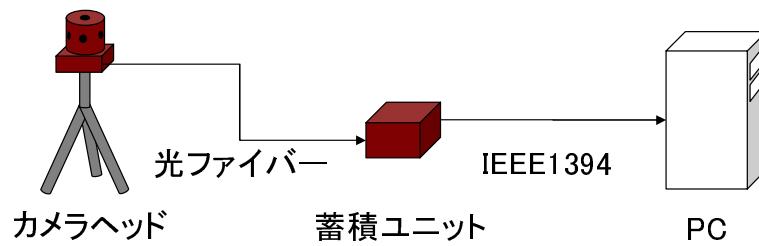


図 6 Ladybug による画像の取り込み

ラメータの推定を行う。一般的なカメラのキャリブレーションでは、内部パラメータとしては焦点距離、レンズ歪み、歪み中心、アスペクト比が推定され、外部パラメータとしてはカメラの位置および姿勢が推定される。ただし、本論文では、これらのパラメータについては、全方位型マルチカメラシステムの内部パラメータおよび外部パラメータに対して、「単眼カメラの内部パラメータ」および「単眼カメラの外部パラメータ」と呼び区別する。

幾何学的パラメータの定義

本論文では図 7 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、単眼カメラのカメラパラメータを以下のように定義する。ただし、レンズ歪みモデルおよび標本化モデルは、Tsai のキャリブレーション手法 [R. 87] で用いられているものを考え、カメラの投影モデルとしては、一般的な一点透視投影モデルを考える。

まずレンズ歪みモデルについて述べる。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 7 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。

$$x_u = x_d(1 + \kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \dots) \quad (1)$$

$$y_u = x_d(1 + \kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \dots) \quad (2)$$

表 3 Ladybug の仕様

カメラ数	6 個
解像度	1,024×768 / カメラ
視野	75%以上
フレームレート	15fps
蓄積可能時間	約 20 分

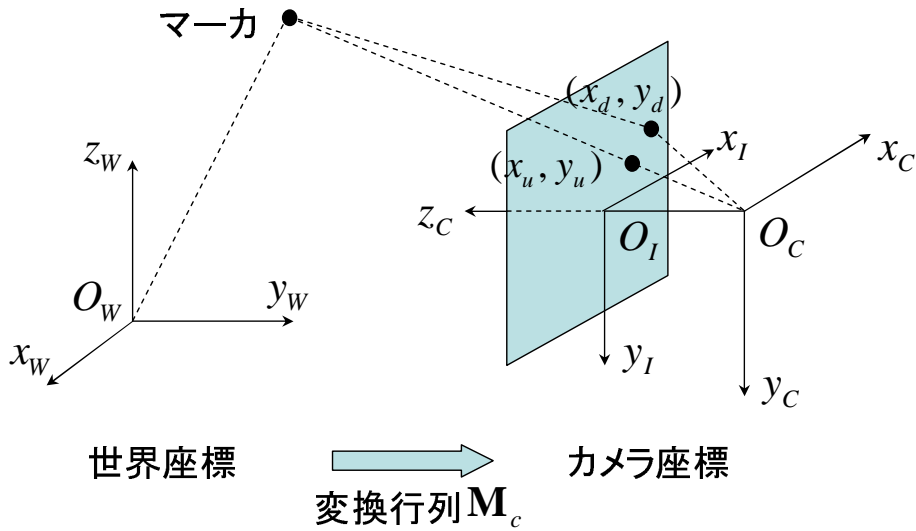


図 7 世界座標系とカメラ座標系の関係

$$r = \sqrt{x_d^2 + x_u^2} \quad (3)$$

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 のみを使用することが多い。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。歪み係数はカメラの位置・姿勢の推定精度に最も影響を与えやすい [J. 93] パラメータであることから、本研究では歪み係数 κ_2 および κ_3 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準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。標準化される前の座標を (x_d, y_d) と標準化後の座標 (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 方向の撮像の素子数 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。

最後に，カメラの投影モデルについて述べる．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 $\mathbf{T}_c(t_x, t_y, t_z)$ 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．

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

ただし，

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

また，図7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．

$$\begin{bmatrix} x_C \\ y_C \\ z_C \\ 1 \end{bmatrix} = M_c \begin{bmatrix} x_W \\ y_W \\ z_W \\ 1 \end{bmatrix} \tag{7}$$

以上をまとめると，幾何学的キャリブレーションでは，表4に示すように単眼カメラの内部パラメータに相当する焦点距離 f ，歪み係数 $(\kappa_1, \kappa_2, \kappa_3)$ ，歪み中心 (c_x, c_y) ，アスペクト比 s_x と単眼カメラの外部パラメータに相当するカメラ c の位置 \mathbf{T}_c ，姿勢 R_c を推定しなければならない．

表 4 全方位型マルチカメラシステムの内部パラメータ一覧

	パラメータ	表記
単眼カメラの 外部パラメータ	カメラの位置	$\mathbf{T}_c(t_x, t_y, t_z)$
	カメラの姿勢	$R_c(\alpha, \beta, \gamma)$
単眼カメラの 内部パラメータ	焦点距離	f
	歪み中心	c_x, c_y
	歪み係数	$\kappa_1, \kappa_2, \kappa_3$
	アスペクト比	s_x

マーカの配置と検出

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、カメラ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8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9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、カメラ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マウスクリック等により与える。
2. 図10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。
3. ウィンドウ内を判別分析法 [大津 80] 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。

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。

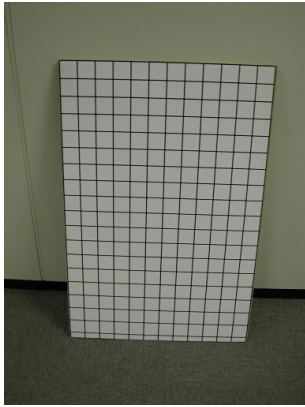


図 8 キャリブレーションボード

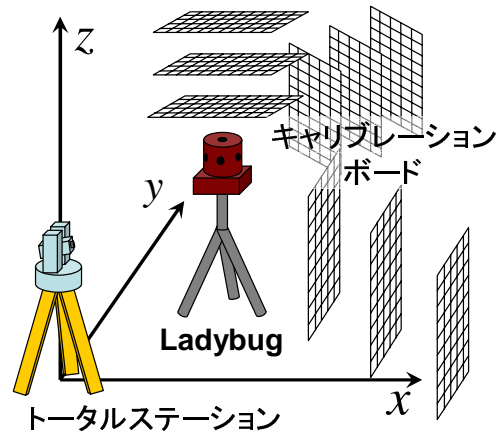


図 9 キャリブレーションボードの配置

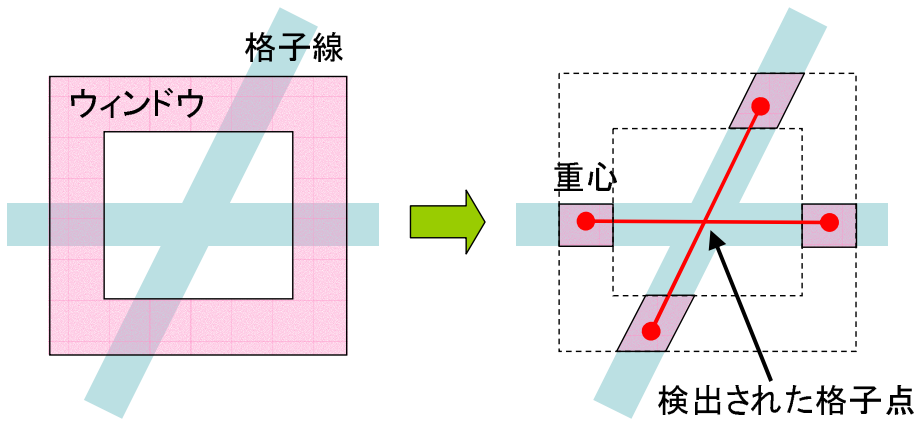


図 10 マーカの検出

幾何学的パラメータの推定

各カメラの幾何学的パラメータの推定方法について述べる. 単眼カメラの内部パラメータはあらかじめ各カメラについて Tsai [R. 87] 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 3 次係数 κ_3 まで考慮し精度向上を図る.

単眼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 [TMNH02]. 本手法では, 単眼カメラの外部パラメータは, 世界座標系におけるカメラ $c(c = 0, 1, \dots, 5)$ の位置 \mathbf{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 \mathbf{x}_m とカメラ $c(c = 0, 1, \dots, 5)$ の入力画像上での位置 \mathbf{u}_m とを対応づけ, 線形解法 [出口 90] 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 \mathbf{v}_m とマーカの入力画像上での位置 \mathbf{u}_m との差 (以降, 再投影誤差) の二乗和で算出される以下の評価関数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.

$$E_c = \sum_m |\mathbf{u}_m - \mathbf{v}_m|^2 \quad (8)$$

2.3.2 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。

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、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 [B. 86] と望遠レンズに特有の口径蝕現象 [NAM96] 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 [B. 86] によれば、図 11 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放射照度 I は、レンズへ入射する光の入射角 θ に依存して減衰し、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。

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 I は物体表面の放射輝度 L に対して線形的な関係 $I = aL + b$ 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 $I_{c'}$ の間にも以下に示す線形的な関係があると考えられる。

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

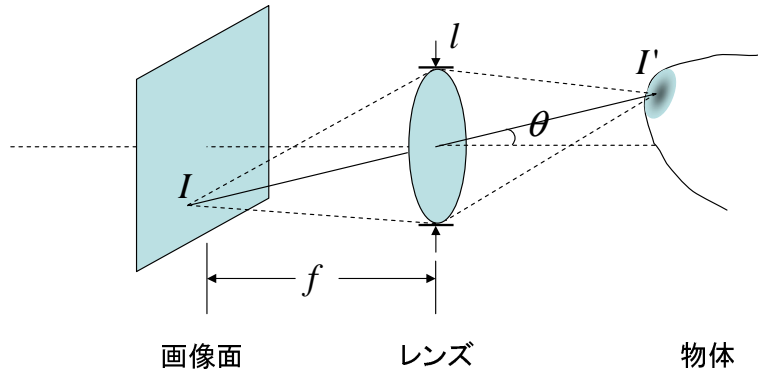


図 11 $\cos^4 \theta$ 現象のレンズモデル

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。

1. 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。
2. 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。

3. 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。

2.4 全天球画像の生成

前述の手法によって推定された全方位型マルチカメラシステムの内部パラメータを基に，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る．図 12 に示すように，まず，前節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正距円筒図法により平面に展開し，全天球画像を生成する．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．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 に投影することで全天球画像を生成する．

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) で表したとき，全天球画像の座標を図 13 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ことができる．

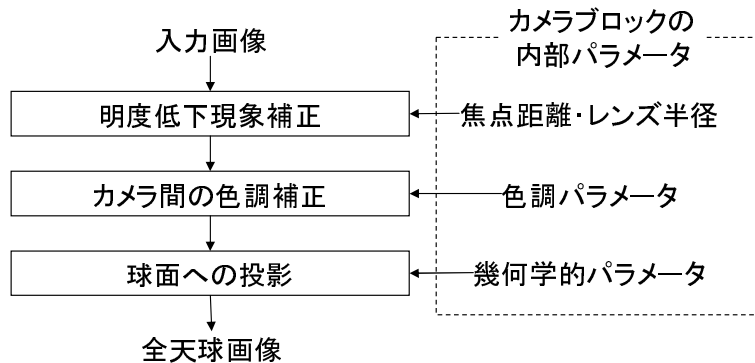


図 12 全天球画像の生成手順

2.5 視差の影響

以下では、全天球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、視差の影響が全天球画像上で 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 14 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 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、隣り合うカメラの投影中心間の距離を 40mm、球面へ投影された画像の円周長が 3,340 ピクセ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 20m である。

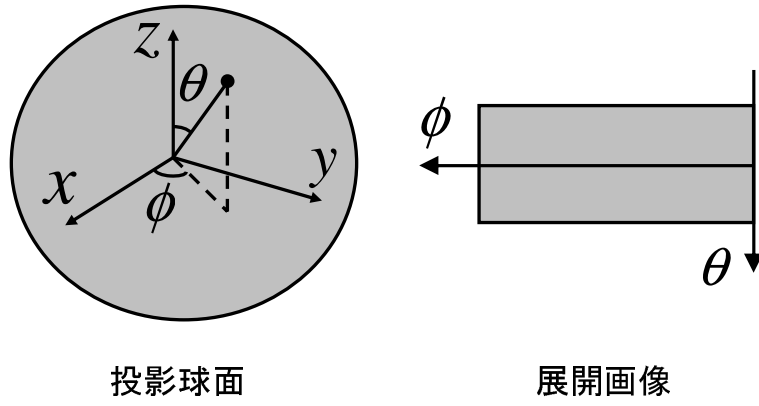


図 13 投影球面から全天球画像への展開

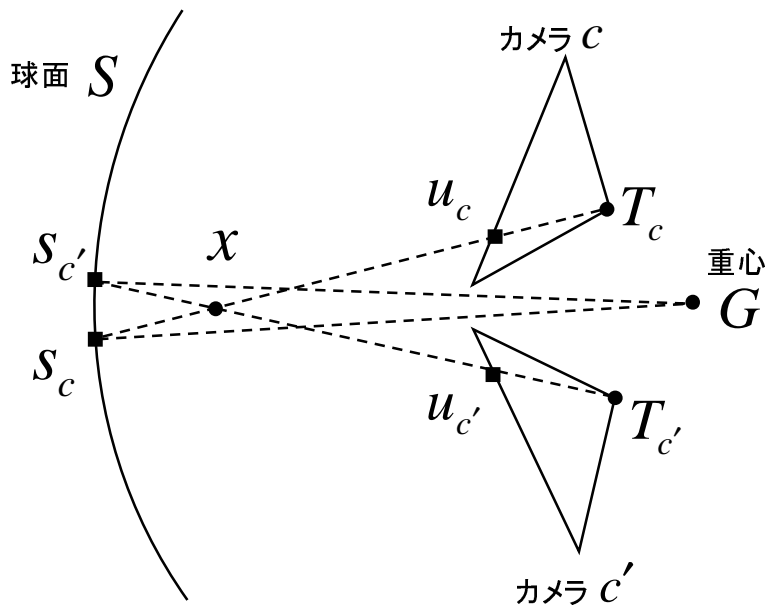


図 14 視差の影響

2.6 実験

2.6.1 幾何学的キャリブレーション

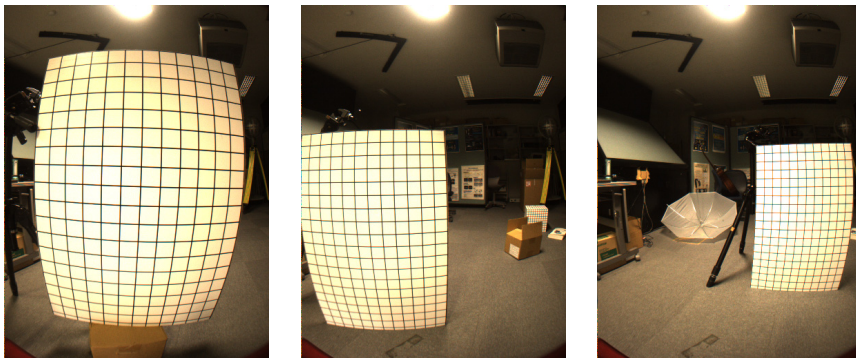
Ladybugの幾何学的なキャリブレーションでは、Ladybugを三脚に固定し、キャリブレーションボードをカメラの奥行き方向に約50cm間隔で移動させ、各カメラごとに3回ずつ撮影した。図15にこの撮影により得られたキャリブレーションボードの画像を示す。水平方向の各カメラにはキャリブレーションボードの格子点のうちそれぞれ187個ずつ検出し、合計561個のマーカの三次元位置と画像上の位置を取得した。上方向のカメラについては、キャリブレーションボードを固定する部分が写るため、各画像から170個ずつ検出し、合計510組のマーカの位置情報を取得した。マーカの三次元位置はキャリブレーションボードの3隅の格子点を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入力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16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5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17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。

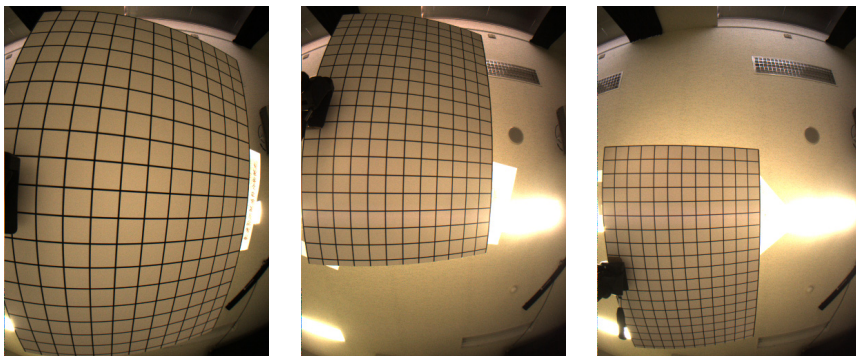
次に、推定した外部パラメータを用いて描画したカメラの位置と姿勢の関係を図18に示す。同図では、四角錐と直線でカメラの位置と姿勢の関係を表現している。四角錐の頂点と直線が交わる点はカメラの投影中心を、直線の方法はカメラの光軸を、四角錐の底面はカメラの光軸周りの回転を表す。また、表6に各カメラの再投影誤差の平均値、標準偏差、最大値を示す。図18に示した図形の位置および姿勢の関係がLadybugのカメラの配置に近く、表6に示した各カメラの再投影誤差の平均値が小さいことから、おおむね正しく外部パラメータが推定されていることが確認できる。

表 5 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]	0.0820	0.1162	0.1186
歪み係数 κ_2 [1/mm ⁴]	-0.0062	-0.0187	-0.0199
歪み係数 κ_3 [1/mm ⁶]	0.0020	0.0034	0.0035
焦点距離 f [mm]	2.61	2.65	2.65
アスペクト比 s_x	0.99	1.00	1.00
歪み中心 c_x [pixel]	356.2	352.1	341.9
歪み中心 c_y [pixel]	508.8	496.7	484.9
	カメラ 3	カメラ 4	カメラ 5
歪み係数 κ_1 [1/mm ²]	0.1028	0.0610	0.0770
歪み係数 κ_2 [1/mm ⁴]	-0.0145	-0.0018	-0.0061
歪み係数 κ_3 [1/mm ⁶]	0.0028	0.0016	0.0020
焦点距離 f [mm]	2.67	2.59	2.62
アスペクト比 s_x	0.99	0.99	0.99
歪み中心 c_x [pixel]	350.2	348.0	369.1
歪み中心 c_y [pixel]	482.8	516.5	535.0



水平方向のカメラで取得した画像



上方向のカメラで取得した画像

図 15 キャリブレーションに使用したマーカの画像例

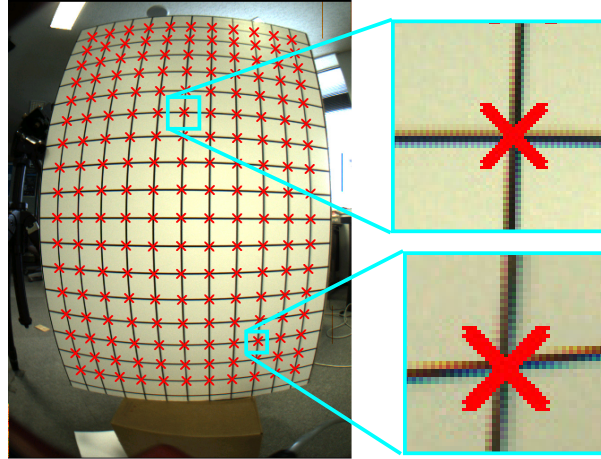
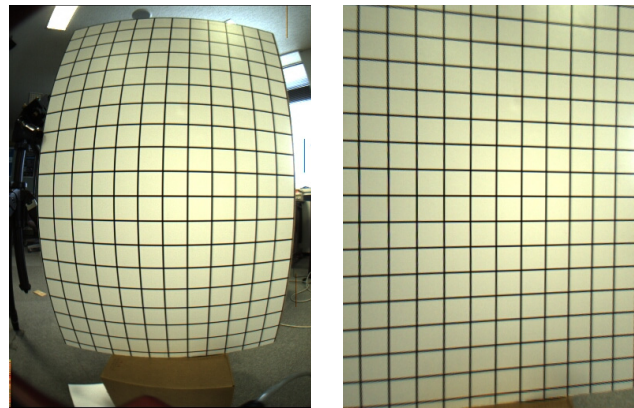


図 16 キャリブレーションボード上の格子点の検出例



(a) レンズ歪み補正前 (b) レンズ歪み補正後

図 17 レンズ歪み補正結果

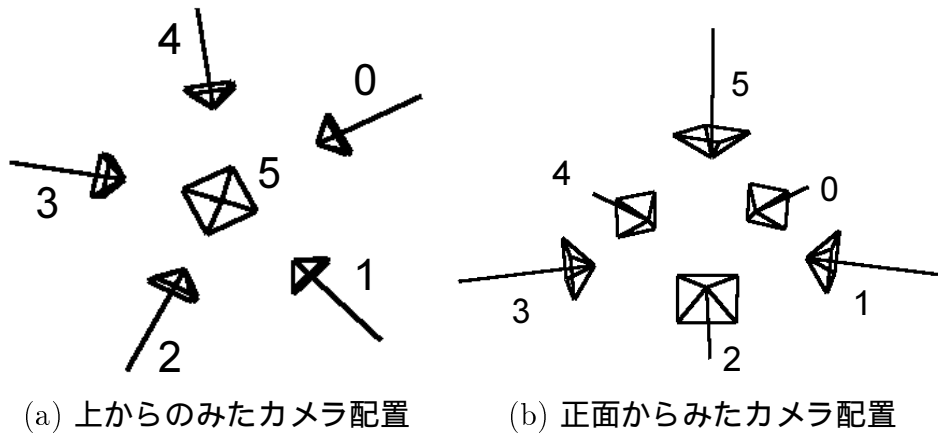


図 18 カメラの位置および姿勢の推定結果

表 6 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
	カメラ 3	カメラ 4	カメラ 5
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) 補正前の入力画像

(b) 補正後の入力画像

図 19 明度低下現象の補正結果

2.6.2 光学的キャリブレーション

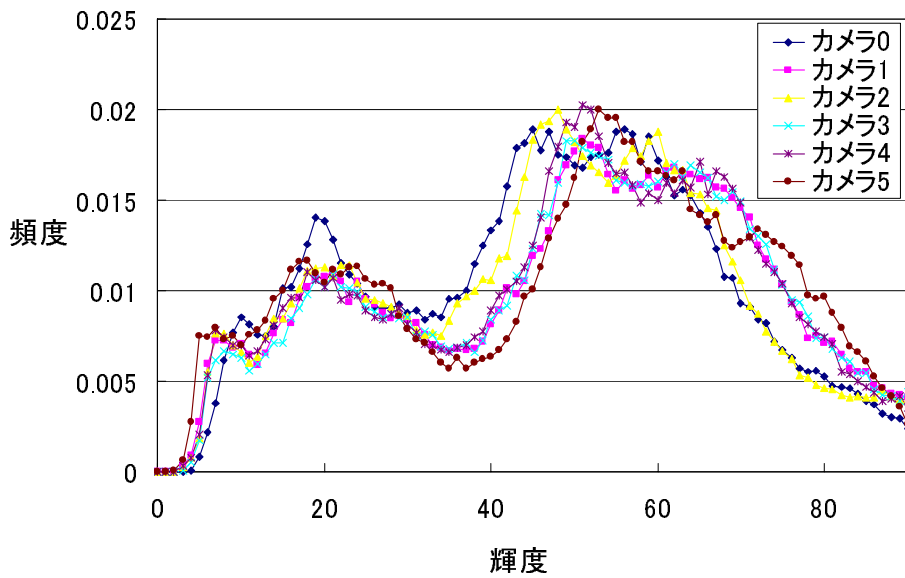
図 19 に明度低下現象の補正結果を示す。同図左のように補正前の画像では、周辺部分の輝度値が全体的に低下しているのに対して、同図右の様に補正後の画像ではそれらが補正されていることが分かる。

図 20 は、色調補正前後でブレンド処理なしに全天球画像を生成し、異なるカメラによる画像の継ぎ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生成画像では、画像の境界が目立たなくなっていることが確認できる。また、図 21, 22, 23 は図 24 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、カメラ間の色調が補正されていることが確認できる。

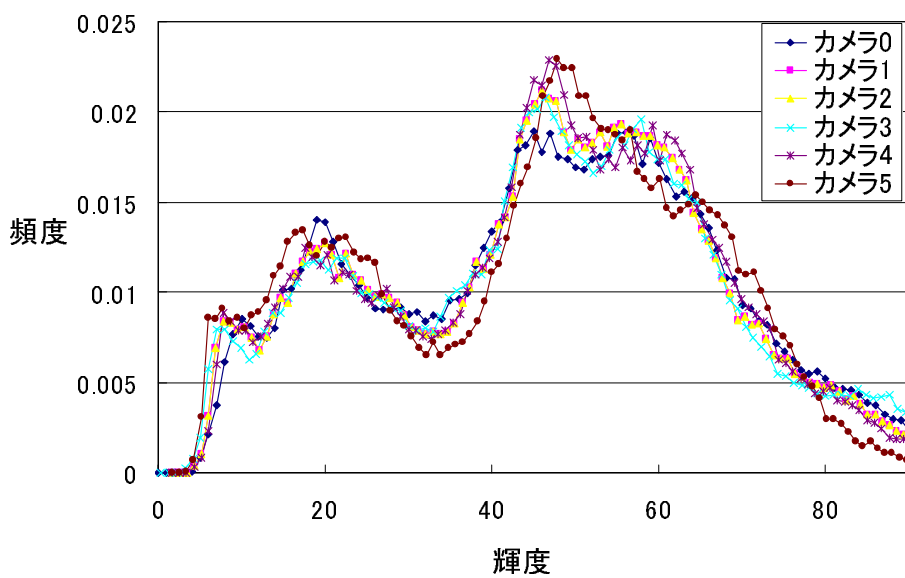


(a) 色調補正なしの生成画像 (b) 色調補した生成画像

図 20 色調補正の効果

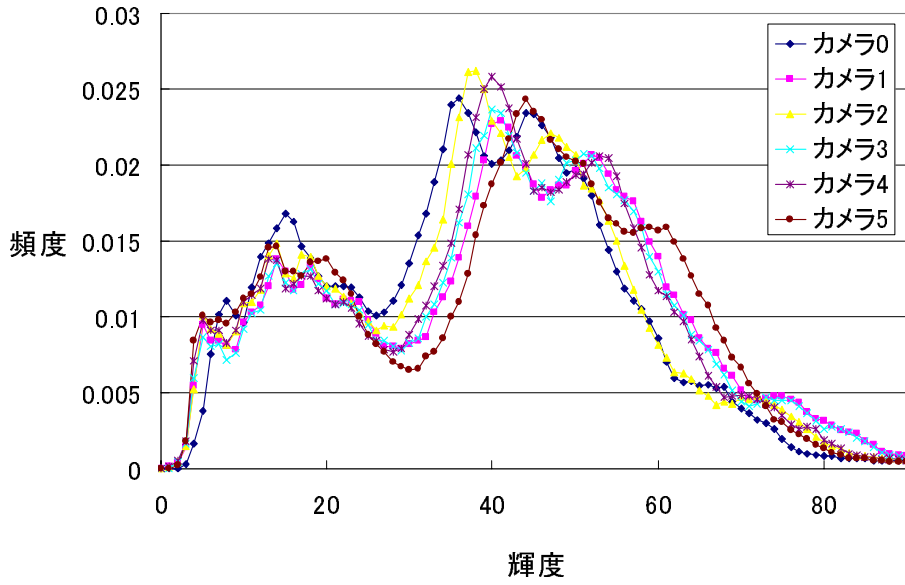


(a) 色調補正前のヒストグラム

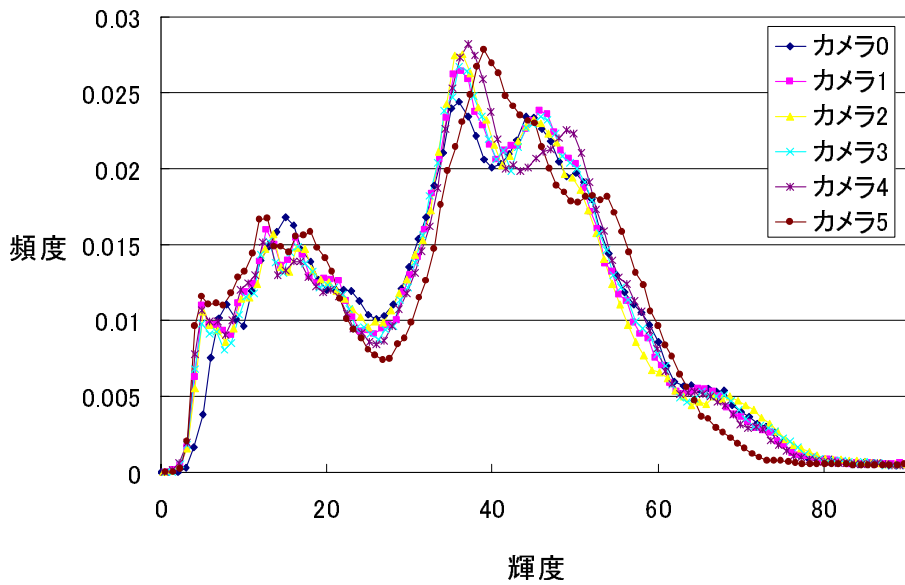


(b) 色調補正後のヒストグラム

図 21 色調補正前後のヒストグラム (R 成分)

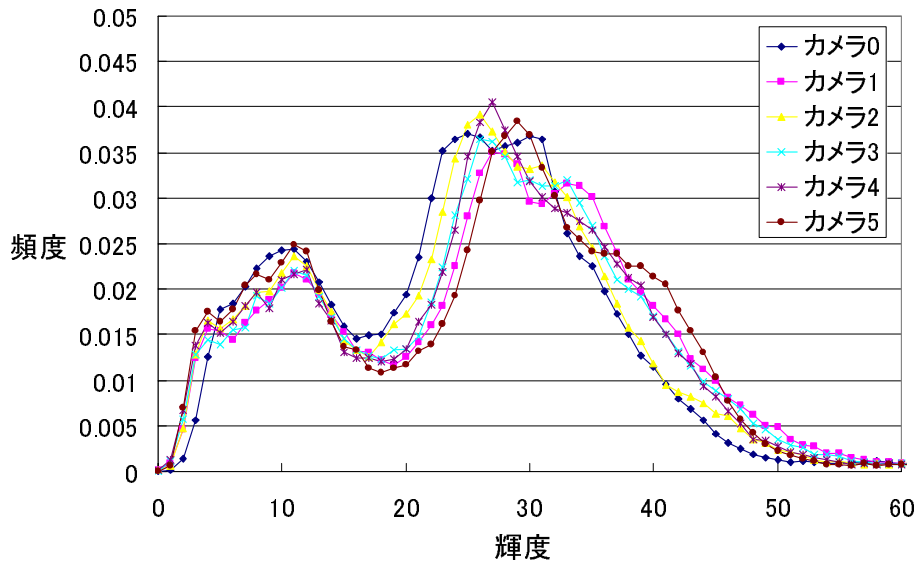


(a) 補正前のヒストグラム

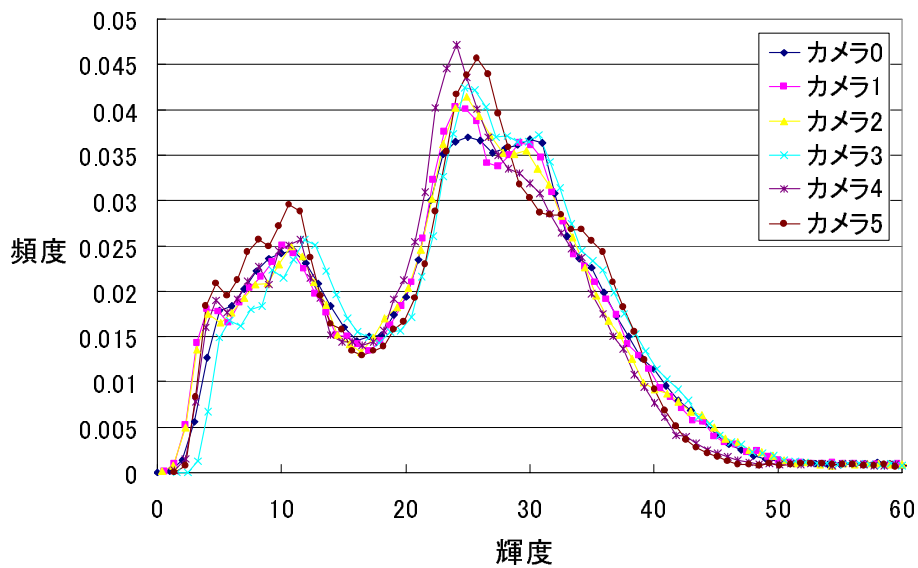


(b) 補正後のヒストグラム

図 22 色調補正前後のヒストグラム (G 成分)



(a) 補正前のヒストグラム



(b) 補正後のヒストグラム

図 23 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 24 色調補正に用いた入力画像

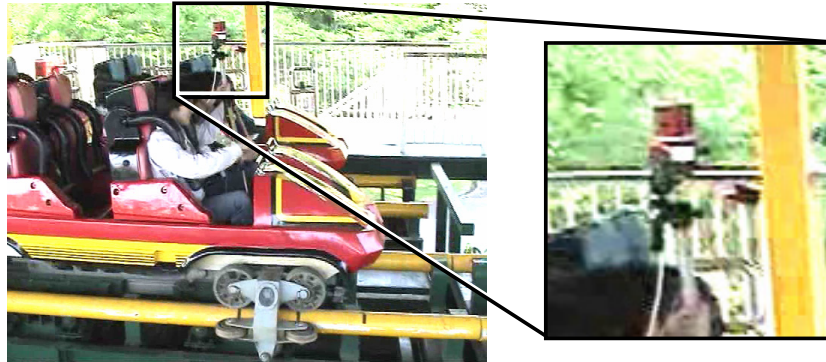


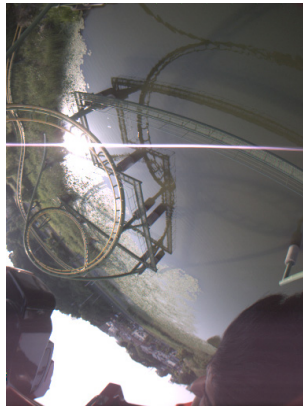
図 25 Ladybug による屋外の撮影

2.6.3 全天球動画像の生成

推定した内部パラメータを用いて全天球動画像を作成した。入力動画像は、図 25 に示すように Ladybug をジェットコースターの上に固定し走行中の景色を撮影することにより得た。図 26 に各カメラから得られた入力画像 (解像度: $768 \times 1,024$) 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を図 27 に示す。全天球動画像は、極座標を用いて画像を平面に展開しており図 27 の下部の黒い部分は、入力画像の存在しない部分である。なお、本論文では、全天球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 $3,340\text{pixel}$ とした。また、全天球画像の垂直方向の解像度は、水平方向の半分の $1,670\text{pixel}$ とした。図 27 から、全天球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。



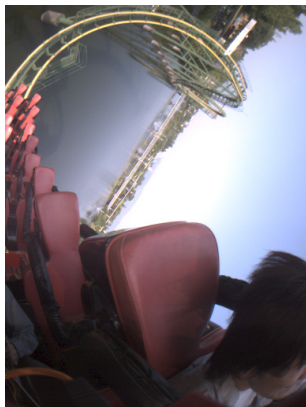
カメラ 0



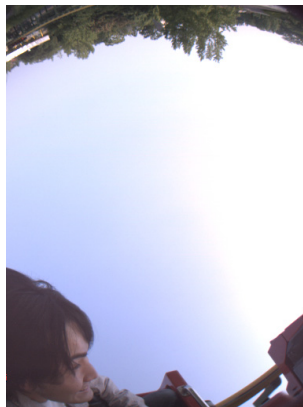
カメラ 1



カメラ 2



カメラ 3



カメラ 4

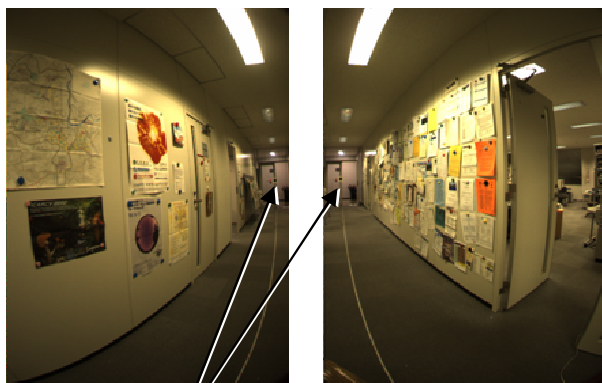


カメラ 5

図 26 入力画像 (奈良 あやめ池遊園地：上方向 (右下) と水平方向 (その他))



図 27 全天球動画像の1フレーム



円形マーカ

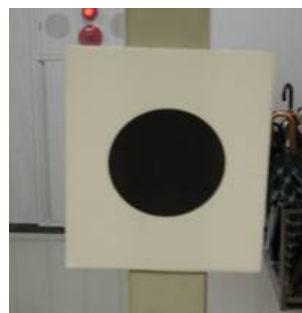


図 29 円形マーカ

図 28 隣り合うカメラの画像

2.6.4 全天球画像の生成精度評価

異なるカメラにより得られる画像中の対応点が、全天球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。この領域内に図 29 に示す円形マーカを写し、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 $u_c, u_{c'}$ とおき、2.5 節で述べた図 14 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、2.5 節で述べたように視差の影響が 1 画素以下となる対象物体までの距離は約 20m 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 7 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。実験から誤差 ϵ は平均 0.0061rad であった。これは入力画像面上では約 3 画素に相当し、表 6 に示すように幾何学的キャリブレーションにおける平均再投影誤差と同程度であることから、全天球画像生成における誤差要因は各カメラの位置および姿勢の推定誤差によるものが大きいと考えられる。全天球動画像を表示し、3 画素程度の誤差が生じている部分を目視した結果、生じているはずのずれは認知できなかった。

表 7 画像生成に関する誤差 $\epsilon \times 10^{-2}$ [rad]

カメラ番号対	0-1	1-2	2-3	3-4	4-0
マーカ数	233	288	233	244	238
最大	1.04	0.89	1.24	1.99	1.24
平均	0.62	0.51	0.65	1.36	0.37

カメラ番号対	5-0	5-1	5-2	5-3	5-4
マーカ数	185	170	169	220	194
最大	0.78	0.51	1.10	1.06	0.74
平均	0.33	0.28	0.81	0.79	0.20

2.6.5 受動的視点移動型テレプレゼンスシステムの試作

高解像度な全天球動画像を用いた、図 30 に示す全天球の見回しが可能な受動的視点移動型テレプレゼンスシステムを試作した。試作システムは表 8 に示すように、全天球視野の半分程度を一度に表示可能な球面ディスプレイ、コントローラ、PC の 3 つの装置により構成されている。球面ディスプレイの解像度が $1,024 \times 768$ であるため、全天球動画像は、 $2,048 \times 1,024$ の解像度に縮小し、JPEG 形式で PC に蓄積してある。図 30 に示すように、利用者はコントローラを用いてインタラクティブに全方向の見回しや再生スピードの変更、表示シーン切り替えを行うことができる。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。

試作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臨場感の高いシステムとなったが、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。



図 30 テレプレゼンスシステムの外観

表 8 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

2.7 結言

本章では、全天球型マルチカメラシステムの内部パラメータ推定として、幾何学的および光学的なキャリブレーションを行い、その結果を基にして全天球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、システムの全方向に多数のマーカを空間的に広く配置することで、設計パラメータを使用せずにカメラパラメータの推定を行った。本論文では Ladybug という特定のカメラシステムへの適用について述べたが、提案したカメラ内部パラメータの推定手法は、放射状にカメラを配置した全天球型マルチカメラシステム全般に適用可能である。

また、全天球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、提案したカメラ内部パラメータ推定手法では、入力画像面上で平均約 3 画素のずれが生じる精度でキャリブレーションを行えることが分かった。さらに、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画像が受動的視点移動型テレプレゼンスに利用可能であることを確認した。

3. 全方位型マルチカメラシステムのカメラ外部パラメータの推定

3.1 緒言

誤差が蓄積することなく，精度の高い外部パラメータを推定するためには，事前知識の獲得やセンサ間のキャリブレーションが必要という問題があった．本章では，屋外環境下を広範囲に移動する全方位型マルチカメラシステムで取得した動画像からのカメラ外部パラメータの推定手法を提案する．提案手法では，動画像中の特徴点の追跡と GPS 測位値を用いた最適化に基づきカメラ外部パラメータを推定する．特徴点の追跡では，ランドマーク・特徴点追跡によるカメラ外部パラメータの推定手法 [佐藤 05a] と同様に，特徴点の仮対応づけから推定される暫定的な外部パラメータを用いて特徴点の画像上での探索範囲を限定し，誤追跡を防ぐ．GPS 測位値を用いた最適化では，動画像から推定される外部パラメータに基づく特徴点の再投影誤差および外部パラメータから推定される GPS 受信機の位置の推定誤差からなる評価関数を最小化することで誤差の蓄積を防ぐ．本手法では，現実環境に関する事前知識や GPS 受信機とカメラの間の正確なキャリブレーションを必要とせず，GPS の測位座標系における外部パラメータを推定できる特長がある．

ただし，本手法では，動画像全体にわたりカメラに対する GPS 受信機の位置は固定されているものとし，容易に計測できる両センサ間の距離については既知，カメラ座標系における GPS 受信機の方位については推定するものとする．また，内部パラメータが推定されているものとし，本章で扱う入力画像は，歪み補正，周辺減光補正，カメラ間の色調補正がされているものとする．動画像と GPS 測位値の同期についても得られているものとする．

以下，3.2 節では，カメラと GPS 受信機の幾何学的関係のモデル化を行い，3.3 節では，提案手法の流れおよび各処理について詳述する．3.4 節では提案手法の有効性を確認するためにシミュレーションによる実験および実環境のデータを用いた実験を示し，最後に 3.5 節で本章をまとめる．

3.2 カメラと GPS 受信機の幾何学的モデル

本節では、まず、全方位型マルチカメラシステムのカメラ外部パラメータに関するカメラモデルについて述べ、次に従来の特徴点のみによるカメラパラメータ推定で一般的に用いられる再投影誤差について説明する。さらに、カメラと GPS 受信機の幾何学的関係をモデル化し、GPS 受信機に関する誤差について説明する。

全方位型マルチカメラシステムの外部パラメータ

本論文では、全方位型マルチカメラシステムの各カメラ間の位置および姿勢関係は、常に固定されていることを仮定しているため、全てのカメラの位置および姿勢は、代表するカメラ $c = 0$ の位置および姿勢により表現することができる。前章で述べたように、全方位型マルチカメラシステムの内部パラメータ推定により、各カメラ c のキャリブレーション時の世界座標系からカメラ座標系への変換行列 M_c が得られる。これを用いると、新たに撮影する環境の第 i フレームにおける世界座標系から各カメラ c のカメラ座標系への変換行列 N_{ic} は、以下のように表現できる。

$$N_{ic} = M_c(M_c)^{-1}N_{i0} \quad (c = 0, 1, 2, \dots) \quad (14)$$

$$= \begin{bmatrix} R_{ic} & t_{ic} \\ 0 & 1 \end{bmatrix} \quad (15)$$

ただし、 R_{ic} は、第 i フレームにおける世界座標系から各カメラ c の座標系への回転を表し、 t_{ic} は、並進を表す。このように考えることで、一般的な単眼のカメラと同様な式により全方位型マルチカメラシステムのカメラ外部パラメータ ($R_i = R_{i0}$, $t_i = t_{i0}$) を表現することができる。

特徴点の再投影誤差

再投影誤差は、特徴点追跡に基づく外部パラメータ推定手法で一般的に用いられる誤差関数である [FZ98, PKV⁺00, SKYT02]。再投影誤差の二乗和を最小化

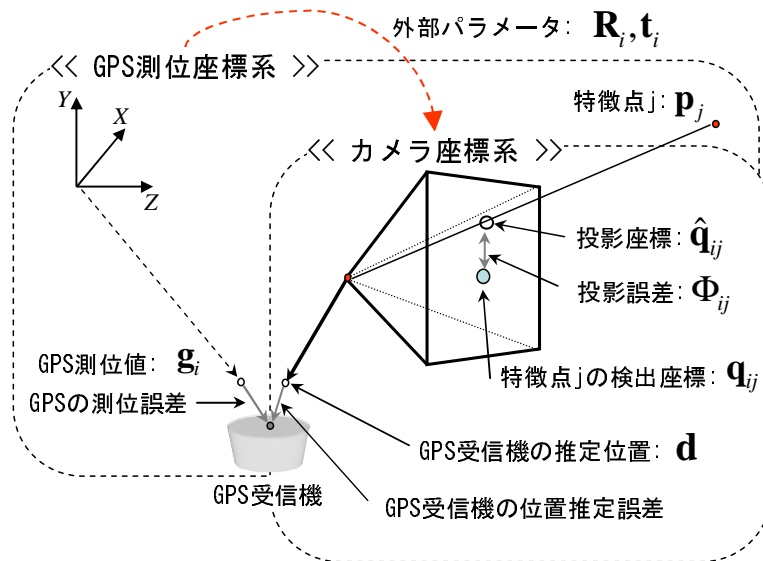


図 31 座標系と誤差の定義

することでカメラパラメータを推定する手法はバンドル調整法と呼ばれている [TPMF00]. 特徴点 j の第 i フレーム上での再投影誤差 Φ_{ij} は, 推定された三次元位置の画像上への投影座標 $\hat{\mathbf{q}}_{ij}$ と画像上で検出される座標 \mathbf{q}_{ij} を用いて, 以下のように表される.

$$\Phi_{ij} = |\mathbf{q}_{ij} - \hat{\mathbf{q}}_{ij}| \quad (j \in \mathcal{S}_i) \quad (16)$$

ただし, \mathcal{S}_i は第 i フレームの画像上で検出される特徴点の集合を表す.

GPS 受信機とカメラに関する位置誤差

GPS の測位誤差には, 真の GPS 受信機の位置を中心に発生する期待値 0 の局所的な測位誤差と平均観測位置が真値から外れる大域的な誤差の 2 種類が考えられる. 本論文では, 前者の局所的な測位誤差のみを考慮する.

一般に, GPS 測位値と外部パラメータの推定結果に誤差が含まれない場合, 第 i フレームの外部パラメータ $\mathbf{R}_i, \mathbf{t}_i$ とそれに対応する GPS 測位値 \mathbf{g}_i および, カメ

ラ座標系での GPS 受信機の位置 d には，以下の関係が成り立つ．

$$R_i g_i + t_i = d \quad (i \in \mathcal{F}) \quad (17)$$

ただし， \mathcal{F} は，GPS の計測時刻に対応する動画像のフレームの集合を表す．ここで，図 31 に示すように，GPS 測位値 g_i および外部パラメータ R_i, t_i に誤差があると，式 (17) は成り立たない．そこで本研究では，外部パラメータ R_i, t_i から推定される GPS 受信機の位置と GPS 測位値との差を GPS 受信機とカメラに関する位置誤差関数 Ψ_i とし，以下のように表す．

$$\Psi_i = |R_i g_i + t_i - d| \quad (18)$$

3.3 GPS 測位値を用いた外部パラメータ推定

提案手法では，計算量の問題から従来手法による一括処理が難しいような広範囲にカメラが移動する環境を想定し，動画像と GPS 測位値から，特徴点の三次元位置，カメラ座標系における GPS 受信機の方位，カメラ外部パラメータを推定する．本手法は，主に特徴点追跡と外部パラメータの最適化処理からなり，両処理に GPS の位置情報を反映させることで，誤差の蓄積を防止する．

本手法の流れを図 32 に示す．まずフレーム毎の処理として，(A) 特徴点の追跡，(B) カメラ外部パラメータの初期値推定を行い，次に一定フレーム間隔 k で (C) GPS 測位値を用いた最適化を行う．処理 (C) では，蓄積誤差の影響によって後続のフレームで特徴点の対応付けの精度が低下することを防ぐために狭区間での最適化を行う．処理 (D) では，処理 (C) よりも広い区間で多くの GPS 測位値を用いて外部パラメータの最適化を行う．以下では，まず処理 (C)，(D) で共通に用いる GPS 受信機に関する誤差と再投影誤差の両方を考慮した誤差関数を定義し，その最適化手法について述べる．次に処理 (A) ~ (D) についてそれぞれ詳述する．

3.3.1 特徴点と GPS による誤差の定義とその最適化

ここでは，式 (16) で示した画像上の再投影誤差 Φ_{ij} および式 (18) の GPS 受信機とカメラに関する位置誤差 Ψ_i を用いて外部パラメータ推定に用いる最適化の

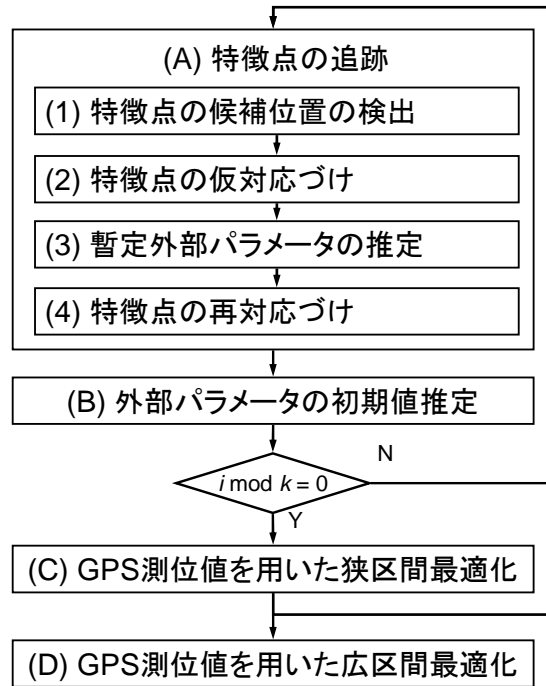


図 32 提案手法の処理手順

誤差関数 E を以下のように定義する .

$$E = \frac{\omega}{|\mathcal{F}|} \sum_{i \in \mathcal{F}} \Psi_i^2 + \frac{1}{\sum_i |\mathcal{S}_i|} \sum_i \mu_i \sum_{j \in \mathcal{S}_i} w_j \Phi_{ij}^2 \quad (19)$$

ただし, w_j は特徴点追跡処理 (A) で得た特徴点 j の信頼度であり, 全てのフレームでの再投影誤差 $\{\Phi_{0j}, \Phi_{1j}, \dots\}$ の分散の逆数とする. μ_i は各フレームに対する重み係数を表す. ω は Ψ_i および Φ_{ij} に対する重み係数であり, 予め, 特徴点の検出位置精度および GPS の測位精度を考慮し, 後述するシミュレーションにより, 真値に対する誤差が最小になるような ω を算出しておく. 一般に, シミュレーションと実利用時の環境では, GPS の測位回数 $|\mathcal{F}_i|$ と特徴点数 $\sum_i |\mathcal{S}_i|$ が異なるので, 本研究では, これらの値に依らず重み係数 ω を一定値に決めるために, $|\mathcal{F}_i|$ および $\sum_i |\mathcal{S}_i|$ により右辺の各項を正規化する.

式 (19) の誤差関数 E を用いた最適化では, 外部パラメータ R_i, t_i , 特徴点の三次元位置 p_j , およびカメラ座標系における GPS 受信機の方位をパラメータとし

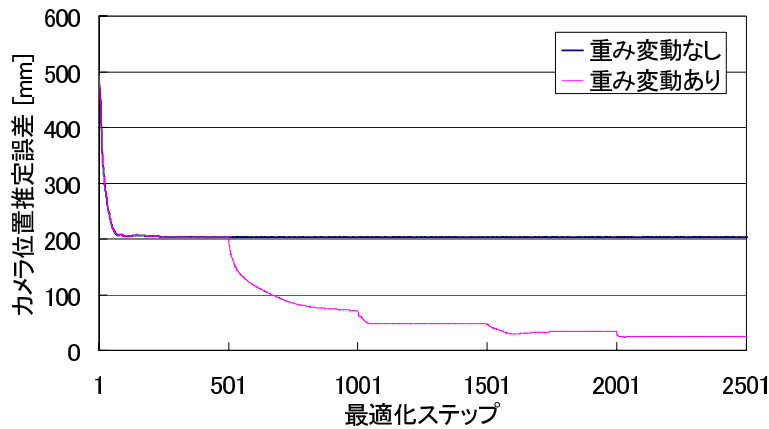


図 33 重み変動の効果

て，誤差関数 E を勾配法によって最小化する．ただし，この誤差関数 E には，局所解が多数存在するため，単純な勾配法では，大域最小解を得ることが難しい．そこで，最適化の各フレームの重み μ_i を変化させることで，局所解を回避する．具体的には，GPS 測位値があるフレームの重みを大きく設定することおよび，すべてのフレームの重みを等価に設定することを，最適化処理が収束するステップごとに交互に繰り返す．

この局所解回避策は経験的に得たものであり，効果については予めシミュレーションによる予備実験により確認している．図 33 は，真値に誤差を加えることで人為的に作成した初期値を提案手法により最適化した際のカメラ位置の推定誤差の変化を表す．このシミュレーションは，3.4.1 節で示すシミュレーションと同様の条件で行い，重み μ_i は，500 ステップ毎に変化させている．同図から，重み変動を行わない場合は，カメラ位置の平均推定誤差が 200mm 程度のときに局所解に陥っているが，重み変動を行うことにより，局所解が回避されていることが分かる．

また，重み係数 ω を決定するために，予め重み係数 ω が変化したときのカメラ位置の推定誤差の変化を予備実験により調べておく．予備実験では，3.4.1 節で示すシミュレーションと同程度のスケールの環境を使用し，カメラの移動経路につ

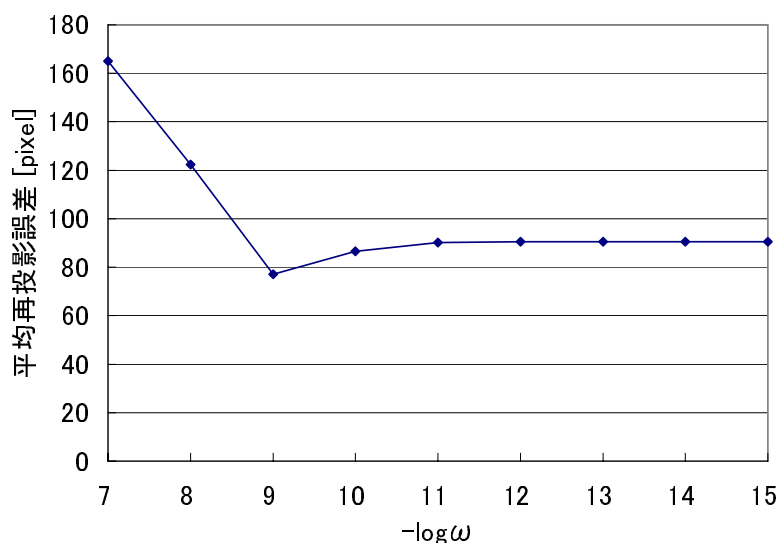


図 34 重み係数 ω の決定

いては、異なるものを設定した。また、特徴点の検出精度や GPS の測位精度は、環境に大きく依存しないと考え、後述する 3.4.1 節のシミュレーションと同じとする。我々が行った予備実験では、図 34 に示すように、 $\omega = 10^{-9}$ のときに最も真値に近いカメラ位置が得られた。この予備実験で使用した仮想環境と類似する環境で提案手法を用いる場合は、重み係数を $\omega = 10^{-9}$ に設定する。

3.3.2 各処理の詳細

本節では、先に定義した特徴点と GPS による位置誤差 E を用いたカメラパラメータの推定手順について、図 32 に示したフローチャートの各ステップ (A) ~ (D) を順に説明する。ステップ (A), (B) は毎フレーム、(C) は一定フレーム毎に実行され、(D) は最終的な最適化処理として 1 回のみ実行される。

(A) 特徴点の追跡

特徴点を入力画像上で検出し、前フレームと現フレームでの対応関係を求める。特徴点には、回転・拡大縮小等の画像の変形に対して頑健に同一の位置が検出さ

れるエッジの角や交点を用いる．さらに，特徴点の仮対応づけの結果から統計的手法によって誤対応を排除し，それにより求まる暫定的な外部パラメータを用いて，特徴点の対応関係を修正する．これにより，特徴点の周辺類似パターンとの誤対応を防止する．具体的な処理手順としては，以下のとおりである．

1. Harris オペレータ [HS88] で求められる特徴量が極大値となる位置を特徴点の候補位置とする．
2. 前フレーム上の各特徴点の候補位置の近傍領域を現フレーム上でテンプレートマッチングによって探索し，現フレームにおける仮の追跡位置を決定する．本ステップにおける特徴点の追跡は，それぞれの単眼カメラの画像内でのみ行われる．
3. 特徴点の仮の追跡位置から，LMeds 基準を用いたロバスト推定 [栗田 00] によって，暫定外部パラメータを推定する．
4. 暫定外部パラメータを用いて，前フレームにおけるすべての特徴点の三次元位置から現フレーム上での出現位置を予測し，ステップ 2 よりも小さな探索窓を用いて対応づけを行い，追跡位置を修正する．ただし，ステップ 2 の追跡処理において画像からフレームアウトし追跡が失敗した特徴点についても，他のカメラ画像上に特徴点の三次元位置を投影することにより，出現位置を予測し，カメラ間の追跡処理を行う．

(B) 外部パラメータの初期値推定

処理 (A) で求めた現フレーム (第 i フレーム) 上の特徴点に関する再投影誤差の二乗和を最小化することで特徴点の三次元位置，および現フレームの外部パラメータである回転 R_i および並進 t_i を推定する．再投影誤差の二乗和は，以下のように表される．

$$\sum_j w_j \Phi_{ij}^2 \quad (20)$$

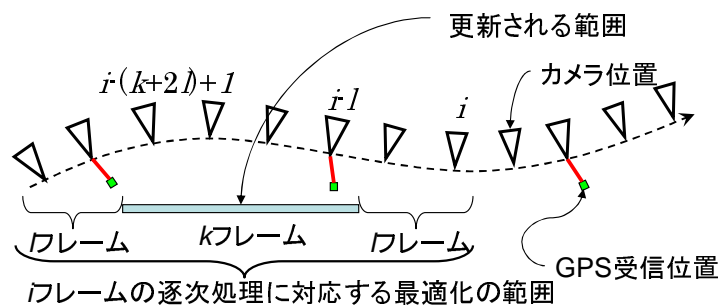


図 35 最適化の範囲

この初期値は、全ての単眼カメラに写る自然特徴点を統合的に扱い、線形解法と非線形最適化の組み合わせ [佐藤 05b] により推定される。

(C) GPS 測位値を用いた狭区間最適化

処理 (C) では、処理 (A)、処理 (B) で得られるカメラ外部パラメータを初期値として、前節で定義した E を最小化することで、GPS 測位値を反映させた外部パラメータを算出する。さらに、この処理で算出されたパラメータを処理 (A) にフィードバックすることで、特徴点对応付けの精度向上を図る。

具体的には、図 35 に示すように、処理 (A) および (B) で処理の対象となる第 i フレームに対して、第 $i - (k + 2l) + 1$ フレームから第 i フレームの特徴点追跡結果および GPS 測位値を用いて狭区間での誤差関数 E の最小化を行い、第 $i - (k + l) + 1$ フレームから第 $i - l$ フレームの k フレーム分の外部パラメータのみを更新する。これにより、更新する区間に対して、前後のフレームにおける特徴点の追跡結果および、GPS 測位値を最適化に反映させる。さらに、一度に複数のフレーム (k フレーム) を更新し、 k フレーム間隔で本処理 (C) を行うことで計算量の増大を軽減する。ただし、多数の追跡済みの特徴点が、推定結果を更新する区間 k フレーム内に写るように、 l は数十フレームに設定する。また、特徴点追跡が外部パラメータの蓄積誤差により破綻する前に GPS 測位値を反映させる必要があるため、 k は数フレームに設定する。

(D) GPS 測位値を用いた広範囲最適化

処理 (D) では，処理 (A) ~ (C) を繰り返すことで得られるカメラ外部パラメータと特徴点の三次元位置を初期値として，処理 (C) よりも広い区間で，外部パラメータの最適化を行う．これにより，特徴点追跡と GPS 測位の両情報に含まれる局所的な誤差の平均化を図る．ただし，本研究では，処理 (C) において間欠的に絶対指標を用いた最適化処理を行うため，時間的に十分離れた 2 フレームに関する外部パラメータや特徴点の三次元位置などの情報は，互いに影響しないと考えられる．そこで処理 (C) での最適化範囲 $k + 2l$ よりも十分に長い複数の $k' + 2l'$ フレームのシーケンスに動画を分割し，各シーケンスを独立に最適化する．ただし，処理 (C) と同様に，最適化区間内に外部パラメータを更新する区間 k' とその前後に更新しない区間 l' を設け，区間 l' を十分大きくとることで，各シーケンスの最適化における独立性を確保する．ただし，十分な数の GPS 測位値を用いて測位誤差を平均化するために， l' は数百フレームに設定する．また，分割されたシーケンス間の独立性を確保するために， k' についても数百フレームに設定する．

3.4 実験

実験では，まず，式 (19) で定義した誤差関数 E の最小化によって最適化処理 (D) で推定される外部パラメータの精度を定量的に評価するために，実環境から作成したデータを用いたシミュレーション実験の結果を示す．次に，提案手法における特徴点検出位置および GPS 測位値の誤差モデルや特徴点对応付けの精度を含め，提案手法の有効性を確認するために，実環境で取得した動画と GPS 測位値から外部パラメータを推定する．また，推定された外部パラメータを用いて動画上に仮想物体を重畳表示するマッチムーブを行うことで，外部パラメータが正しく推定されていることを確認する．最後に，動画全体の一括最適化によって推定される外部パラメータと，動画を複数のシーケンスへ分割して最適化し，最後に統合して得られる外部パラメータの比較を行うことで，分割による最適化の有効性を確認する．

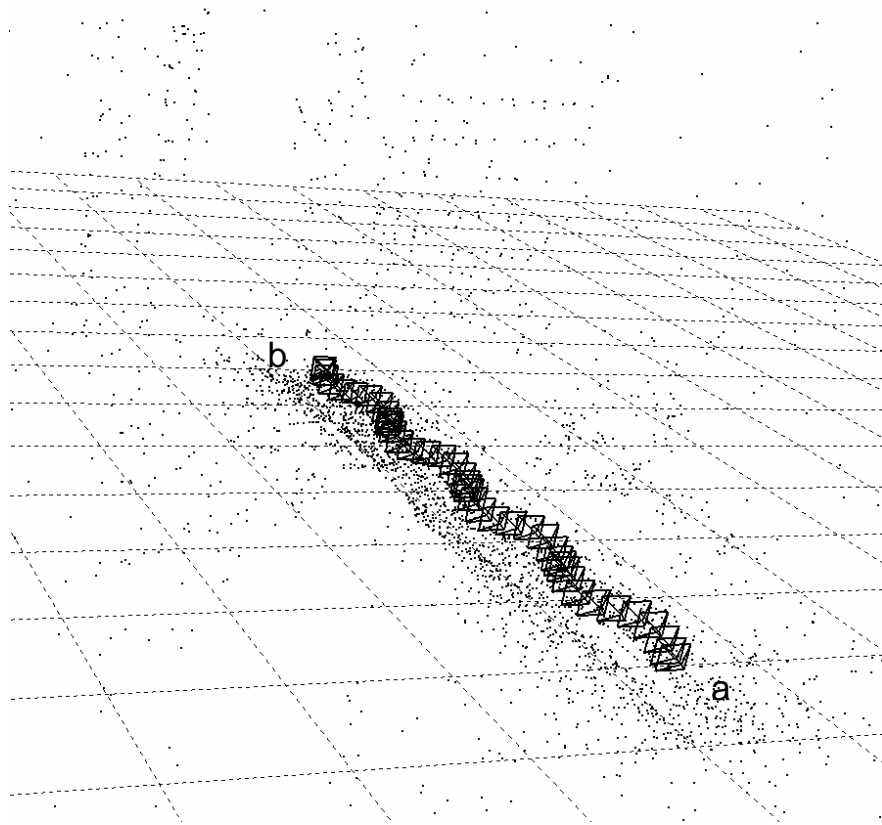


図 36 シミュレーション環境

3.4.1 シミュレーションによる定量評価

シミュレーションによる評価では，提案手法の処理 (D) によって推定される最終的な外部パラメータの推定精度を定量的に評価する．また，従来手法であるランドマーク・特徴点追跡による全方位型マルチカメラシステムの外部パラメータ推定手法 [佐藤 05a] と精度を比較することにより，提案手法の有効性を確認する．

本シミュレーションで用いる仮想環境を図 36 に示す．ここで，図中の錘台は 10 フレーム毎の代表カメラの姿勢を表し，点群は仮想環境中に配置した特徴点の三次元座標を表す．また，グリッドの一边は実環境での 10m に相当する．仮想環境中の特徴点には，事前に屋外環境を撮影した全方位動画像を入力として佐藤ら

の手法 [佐藤 05a] を用いて三次元復元を行い，これによって推定された特徴点の三次元位置を本シミュレーションの真値として用いた．本実験では，このカメラの移動経路とは異なる経路で仮想カメラを図 36 に示すように a 地点から b 地点へ移動させ，仮想的に 900 フレームの特徴点検出画像を取得した．ただし，仮想カメラの内部パラメータには，前章で述べた手法により推定した Ladybug のパラメータを用いた．カメラ座標系における GPS 受信機の位置は，後述する実環境下での実験と同じ (60,-150,250)(単位 mm) と設定し，GPS 測位値には，上記 GPS 受信機の位置に誤差を付加し，これを 1Hz に相当するフレーム間隔として 15 フレーム毎に与えた．特徴点の画像上での検出座標 q_{ij} および，GPS 測位値 g_i に付加した誤差を表 9 に示す．また，外部パラメータ R_i, t_i の初期値は，表 10 に示す誤差を真値に付加して作成した．これらの特徴点の検出誤差，カメラの位置および光軸の向きの誤差については，佐藤らの屋外環境での実験結果 [佐藤 05a] に基づき決定した．3.3.1 節で定義した式 (19) の誤差関数 E 中の重み係数 ω は，前節で説明したシミュレーションにより得た値 10^{-9} とした．また，式 (19) 中の各フレームの重み係数 μ_i については，GPS 測位値が得られていないフレームで 1.0 とし，GPS 測位値が得られているフレームでは，2.0 と 1.0 を交互に用いた．初期フレームにおける外部パラメータの初期値は，手動で大まかな姿勢を与え，位置に関しては GPS 測位値を用いた．また，提案手法の精度評価のために，ランドマーク・特徴点追跡による手法 [佐藤 05a] によって上述と同様の条件で外部パラメータを推定した．ただし，ランドマークは特徴点の信頼度 w_{ij} を十分大きくし

表 9 観測データに付加する誤差

特徴点の検出座標	標準偏差 1.6 画素のガウスノイズ + 量子化誤差
GPS の計測誤差	標準偏差 30mm のガウスノイズ

表 10 外部パラメータの初期値に付加する誤差

カメラ位置	標準偏差 500mm のガウスノイズ
カメラ光軸の向き	標準偏差 0.020rad のガウスノイズ

たものとして与え，全ての画像上での位置を与えた．また，ランドマークを与えるフレームの数や設定方法により外部パラメータの推定精度が変化する可能性があるため，全てのフレームで 15 個以上のランドマークが観測されるように設定した．以上の条件で，特徴点の三次元位置，外部パラメータ，カメラ座標系における GPS 受信機の方位を最適化した．

図 37 および図 38 に初期値として与えた真値からの誤差，提案手法における誤差，ランドマーク・特徴点追跡による手法における誤差を示す．ただし，図 37 は，外部パラメータの位置に関する誤差，図 38 については，光軸の向きに関する誤差を示す．実験の結果，ランドマークを用いる手法では，カメラ位置に関しては平均誤差 47.5mm，カメラ光軸の向きに関しては平均誤差 0.0019rad の精度で推定された．これに対して提案手法では，カメラ位置に関しては平均誤差 30.7mm，カメラ光軸の向きに関しては平均誤差 0.0023rad の精度で推定された．

提案手法では，人手による計測が必要な事前知識を用いることなく，ランドマーク・特徴点追跡による手法と同程度の精度が得られてた．推定精度の差の原因は，ランドマークを絶対指標とする手法と GPS を指標とする手法で，カメラの位置・姿勢の拘束のされ方が異なるためであると考えられる．本実験環境のようにランドマークが主に遠方に分布する場合，画像上でのランドマークの位置がカメラの並進の変化よりも回転の変化に敏感に反応することから，ランドマークを用いる場合は，姿勢に関する拘束が強く，結果として推定精度が高くなると考えられる．これに対して，GPS を用いる場合には，GPS 受信機の観測位置とカメラの投影中心間の距離が強く拘束されるため，カメラ位置に関する推定精度が従来手法よりも若干高まった反面，カメラ姿勢に関する精度は劣っていると考えられる．

3.4.2 実環境下での実験

本節では，提案手法における特徴点検出位置および GPS 測位値の誤差モデルや特徴点对応付けの精度を含めて，提案手法の有効性を確認するために，実環境で取得した動画像と GPS 測位値から外部パラメータを推定する．また，推定された外部パラメータを用いて動画像上に仮想物体を重畳表示するマッチムーブを行うことで，外部パラメータが正しく推定されていることを確認する．

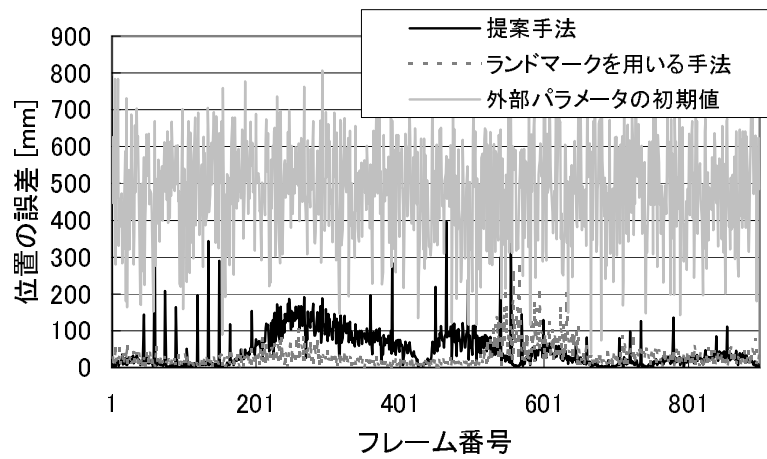


図 37 カメラ位置の推定誤差

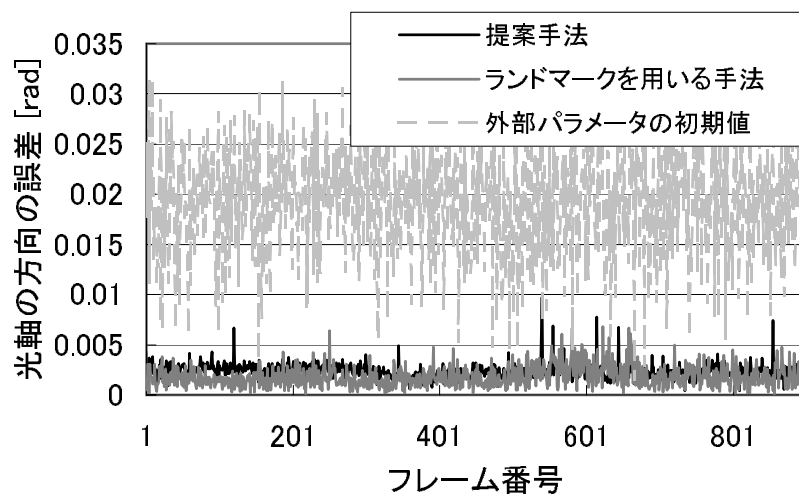


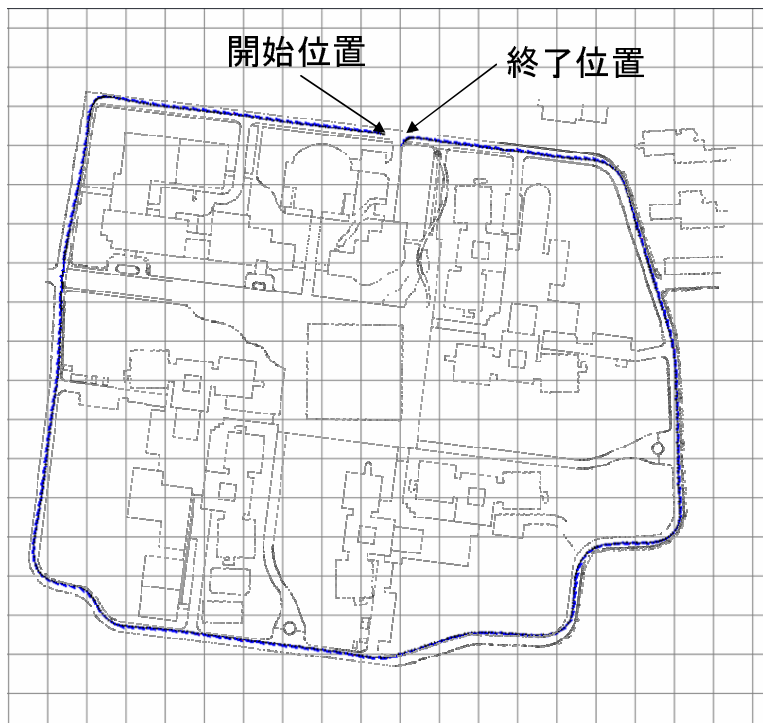
図 38 カメラ光軸方向の推定誤差



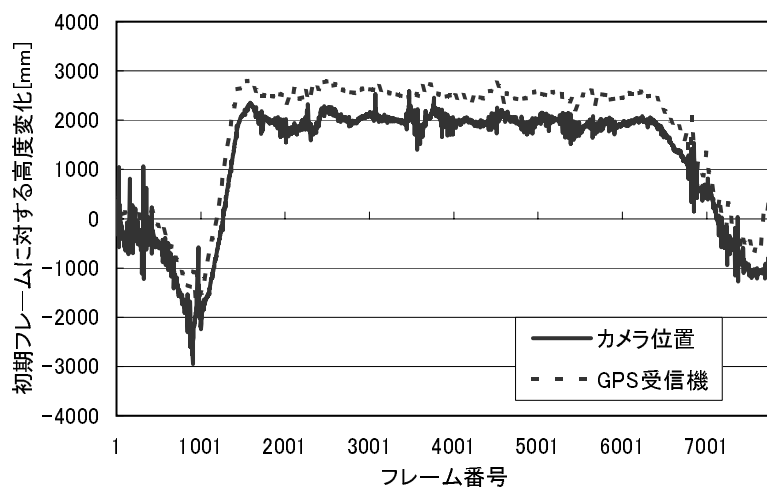
図 39 全方位型マルチカメラシステムと GPS 受信機

実験では，Ladybug と GPS 受信機 (Nikon LogPakII，水平測位精度 $\pm 3.0\text{cm}$ ，垂直測位精度 $\pm 4.0\text{cm}$) を図 39 に示すように電動スクータ (Segway LOC Segway) に固定して取り付けました．このような撮影機器を用いて，約 1.0km の距離を平均時速約 7.6km の速度で走行し，得られた 7800 枚の画像および，RTK 法により 1 秒間隔で計測された GPS 測位値を入力として用いた．ただし，動画像と GPS 測位値は手動で開始時刻を同期させ，カメラと GPS 受信機間の距離は，手動で計測した 300mm を既知な情報として設定した．初期フレームにおける外部パラメータは，前節と同様に，姿勢に関しては大まかな初期値を与え，位置に関しては GPS 測位値を用いた．最適化処理 (C) においては，更新対象フレーム $k = 5$ および重複フレーム $l = 22$ に設定した．

まず，GPS 測位値を用いた最適化で得られた外部パラメータを初期値として動画像全体で最適化した．全体最適化によって推定された特徴点の三次元位置と外部パラメータを図 40 に示す．図中 (a) の曲線は推定されたカメラの移動経路を表し，点群は外部パラメータ推定に用いられた特徴点を表す．同図から分かるよ



(a) 水平移動軌跡



(b) 高度変化

図 40 カメラ外部パラメータの推定結果



第 590 フレーム



第 610 フレーム



第 630 フレーム



第 650 フレーム



第 670 フレーム



第 690 フレーム

図 41 推定された外部パラメータを用いたマッチムーブの結果

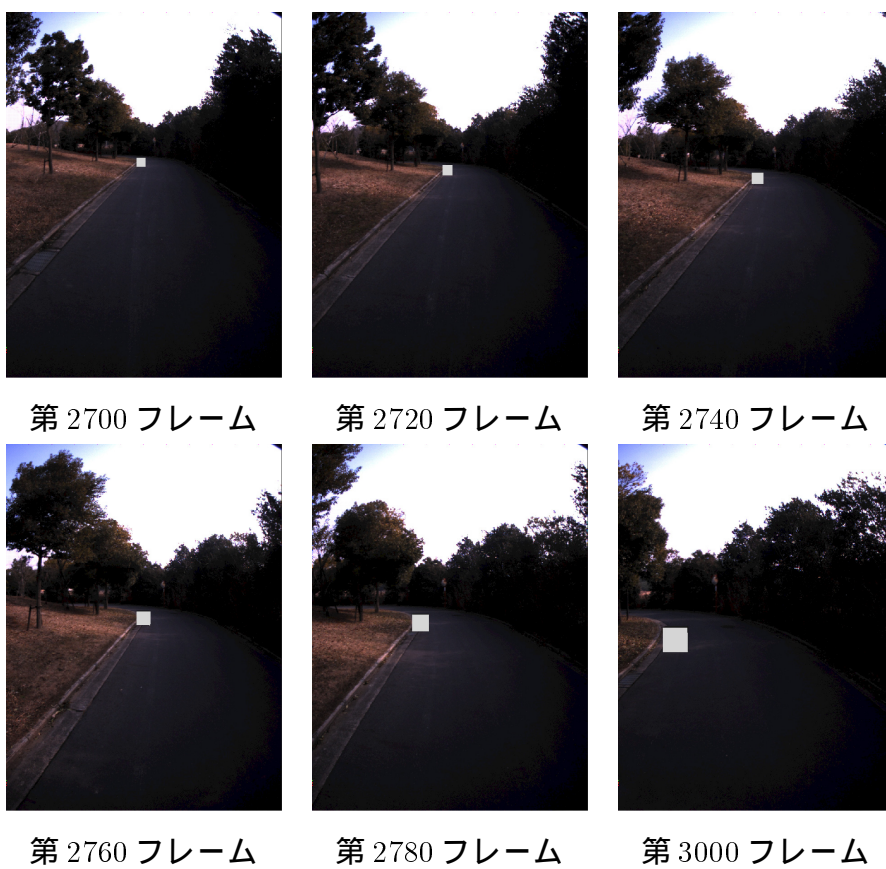


図 42 パラメータ推定の誤差によって仮想物体が静止しないシーン

うに，カメラパスは滑らかに推定されている．また，同図 (b) は，推定されたカメラ位置および GPS 測位値の高度変化を表す．同図から，カメラ位置が GPS 受信機の高度変化に沿うように推定されていることが分かる．

次に，推定された外部パラメータを用いて仮想物体を代表カメラの入力動画像に合成したマッチムーブの結果を図 41 に示す．仮想物体は，GPS による計測位置および GPS 受信機設置時の道路面からの高さから計算される道路面上に配置した．図 41 に示すシーンでは，仮想物体が現実環境中でほぼ一定の位置に合成されていることから，外部パラメータに大きな推定誤差が含まれないことを確認した．しかし，図 42 に示すように，第 2700 フレームから第 3000 フレームでは，仮想物体が一定の位置には静止せず，フレームが進むにつれ場所が移動した．これは，GPS の測位において環境中の遮蔽物等の影響により RTK 法の解が得られなかったことが考えられる．実際に，図 43 に示すように，GPS 受信機から出力される信頼度を表す標準偏差が第 2700 フレームから第 3000 フレームにおいて増加している．本研究では，遮蔽物などの影響を受ける場合については考慮していないが，このような大域的な誤差に対応するためには，GPS 測位値の信頼度等のセンサ情報を考慮したに誤差関数を導入する必要がある．

最後に，処理 (D) における分割最適化の効果を検証するために，図 44 に示すように，重複フレーム l' を 200 フレーム，更新フレーム k' を 1950 フレームとし，動画像を 4 つのシーケンスに分割し，各シーケンスの最適化後に統合した．次に，動画像を一括で最適化して得られる外部パラメータを真値とし，分割最適化によって得られた外部パラメータの誤差を検証した．図 45 に，分割最適化によって得られた外部パラメータの誤差を示す．本実験における平均誤差は，6mm であった．この値は，3.4.1 節で示した実験において示した，カメラ位置の推定誤差 32.9mm に対して比較的小さく，分割最適化により一括での最適化に近い推定結果が得られていることが分かる．計算時間は，Pentium4 3GHz, メモリ 2GB の PC を用いた場合，一括最適化が約 14 時間であったのに対して，分割最適化では各シーケンスがそれぞれ約 3 時間であった．以上により，計算機に搭載されるメモリの量の制限など，一括で動画像全体を最適化できないような場合においても，最適化の際に外部パラメータを更新しない区間 l' のフレーム数を十分大きく設定

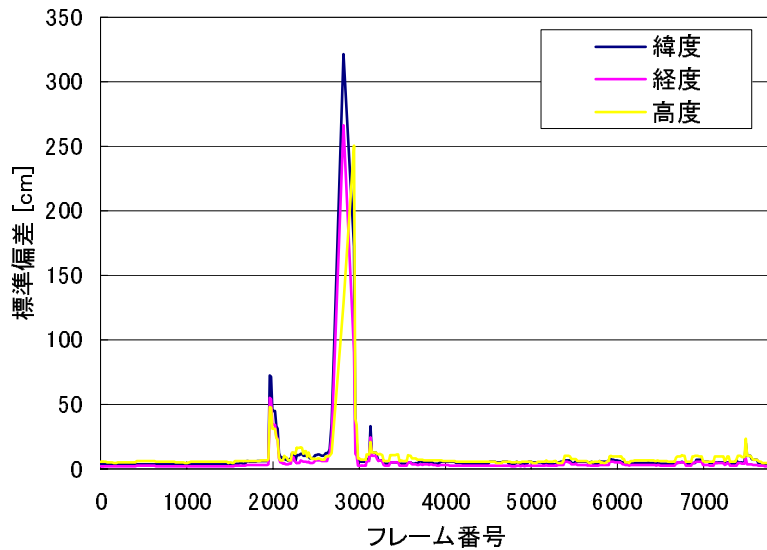


図 43 GPS 測位値に含まれる標準偏差

することで、一括で最適化する場合に近い結果が得られることが分かった。

3.5 結言

動画像と GPS 測位値から全方位型マルチカメラシステムの外部パラメータを推定する手法を提案した。提案手法では、画像上で定義される特徴点の再投影誤差と GPS 測位値によって定義される位置誤差を同時に最小化することで、外部パラメータを各フレームに対して推定することができる。

シミュレーション実験および実環境データを用いた実験により、提案手法は、ランドマーク・特徴点追跡による手法に近い精度で外部パラメータを推定可能であることを確認した。提案手法により、現実環境に関する事前知識の取得に要する人的コストの大幅な削減が可能である。しかし、GPS の電波をさえぎる遮蔽物に囲まれるような状況下では、外部パラメータの推定精度の低下が見られた。上記のような環境中に遮蔽物が多数存在する場合に対応するために、GPS 測位値の信頼度等のセンサ情報を考慮した誤差関数の導入を検討する。また、異なるシー

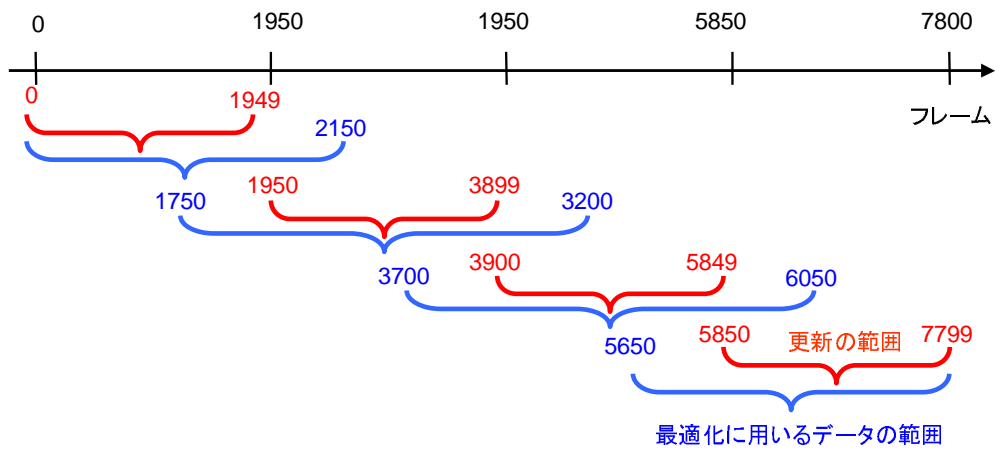


図 44 分割最適化における各シーケンスの分割範囲

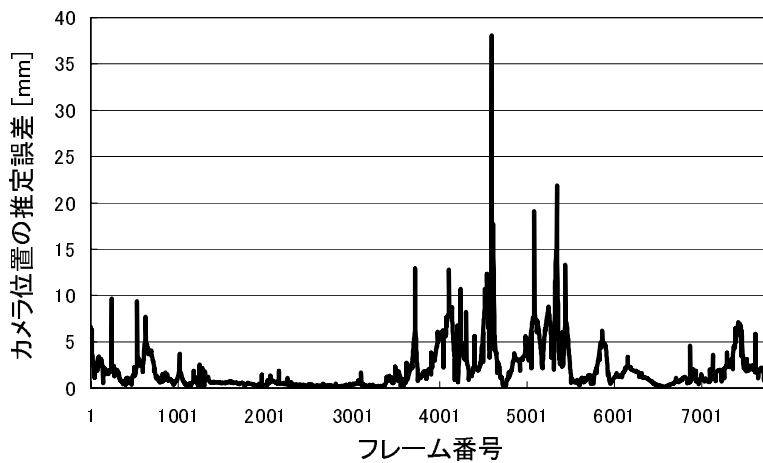


図 45 分割最適化の精度

ケースにおいて同一シーンが撮影される場合に再出現する特徴点を対応づけすることで、さらなる精度向上を図ることができると考えられる。

4. 歩行装置を用いた能動的視点移動型テレプレゼンスシステム

4.1 緒言

本章では，広域実環境を仮想化し，1つの経路上を移動する感覚を再現する能動的視点移動型テレプレゼンスシステムについて述べる．システムの設計においては，提示画像の画質を重視し，動画像の時系列的な連続性および画素間の連結性を維持した提示方式を採用する．この方式に適用可能な次に述べる4つ手法(A)～(D)を用いることで，ユーザに与える臨場感の向上を図る．提案システムでは，高い臨場感を再現するための方策として，以下の4つの特徴を持つ．

- (A) 広視野・高解像画像の提示 全方位型マルチカメラシステムにより取得し，生成した高解像度，広視野角な画像を提示する．
- (B) 提示画像のスタビライジング 歩行動作による視界変化の再現性の向上のために，歩行動作と相関の無い画像取得時における全方位型マルチカメラシステムの移動速度・姿勢の変動の除去する．
- (C) 歩行動作インタフェースの利用 歩行装置を用いた画像提示システムにより遠隔地での歩行動作中の視点位置・視線方向の変化による視界の変化の再現する．
- (D) ヘッドトラッキング 画像提示システム内での歩行動作中による頭部位置の変動を考慮した画像提示する．

本章では，まず，4.2節で臨場感向上の方策に関する考察より決定したテレプレゼンスシステムの設計方針について述べ，4.3節でシステムの実装法について説明する．4.4章では，試作システムの有効性を検証する実験を示し，最後に4.5節で本章をまとめる．



図 46 画像提示システムの外観

4.2 システムの設計方針

1章 1.1 節で述べた臨場感の高いテレプレゼンスシステムという観点から，システムの詳細な設計方針を決定する．以下では，まず画像提示法に関する方針を決定し，次に決定した画像提示法に対応する画像生成手法に関する方針を決定する．

画像提示法に関する方針

従来から，多くの視覚属性を再現する実画像提示システムが多数提案されているが，画像提示システムとしての没入型ディスプレイと視点移動インタフェースの組み合わせは，視覚属性の大部分を再現することができるため，広く研究されている [DTF⁺01, MAS⁺04, 島村 01]．代表的な没入型ディスプレイには，ヘッドマウントディスプレイ (HMD) や CAVE 型のディスプレイ [CNSD⁺92] が挙げられるが，CAVE 型のディスプレイは，高い解像度と広い視野角を確保できるため，テレプレゼンスシステムには適していると考えられる．このような CAVE 型のシステムでは，ユーザの頭部がディスプレイのスイートスポット上にある場合，ディスプレイの能力として全ての視覚属性を再現することが可能である．動きに関する視覚属性をより忠実に再現するために，システム内でのユーザの頭部位

置を検出し，頭部位置に応じた画像を実時間で描画することが一般的に行われている．視点移動インタフェースには，ゲームコントローラや全方位型の歩行装置 [KVJ03, 岩下 04]，乗り物のシミュレータ [CCEE98] などが用いられている．これらのインタフェースはアプリケーションにより使い分けられるが，一般的に，歩行装置や乗り物シミュレータなどのインタフェースは，視点移動操作が直接ユーザの身体的記憶と一致するという特長をもつため，運動属性の再現において，ゲームコントローラなどのインタフェースよりも違和感を与えにくいと考えられる．以上のことから，臨場感に寄与する主な視覚属性の全てを再現することが可能な，画像提示に CAVE 型のスクリーンと歩行装置等の移動インタフェースを組み合わせることとする．

画像生成法に関する方針

本研究では，全方位型マルチカメラシステムにより得られる画像を，時系列的連続性および画素間の連結性を維持し動画像を提示することが基本方針である．全方位型マルチカメラシステムにより取得した画像から提示画像を生成するために，2章で述べた手法により内部パラメータを推定し，利用する．画像提示におけるユーザの視点移動範囲が曲線経路上に限定されるようなアプリケーションでは，この基本方針に従うテレプレゼンスシステムにより運動や位置の属性を再現することができる．しかしながら，1章で述べたように，撮影時にカメラの姿勢や移動速度に変動がある場合，これらの属性が正しく再現されない．これらの属性を再現するために，3章で述べた全方位型マルチカメラシステムのカメラ外部パラメータ推定手法を利用する．具体的には，推定されたカメラの姿勢情報を用いて，カメラの姿勢変動を補正 [MAS⁺04] し，推定されたカメラの位置情報を用いてカメラの移動速度の変動を補正 [C. 02] する．

4.3 システムの実装

4.3.1 歩行装置を用いた画像提示システム

本研究で使用する画像提示システムは，図 47 に示すように，(a) 歩行動作インタフェース，(b) 描画用クライアント，(c) 没入型ディスプレイにより構成される．歩行動作インタフェースは，利用者の歩行動作を検出し，歩行装置を制御することで歩行運動によるユーザの移動を相殺する．描画用クライアント (CPU: Intel Pentium4 1.8GHz, Graphics Card: Geforce4 Ti4600) は，歩行動作インタフェースから受け取った情報を基にユーザの視点に応じた画像を没入型ディスプレイに描画する．以下に各部分の詳細を述べる．

(a) 歩行動作検出インタフェース

図 47(a) に示すように，歩行動作用インタフェースは歩行装置 (限界速度約 1.6m/sec)，レーザレンジファインダ (Sick LMS200) 2 台，サーバ (CPU: Intel Pentium4 2.4GHz) により構成される．サーバでは，レーザレンジファインダにより得られるレンジデータから図 48 に示すユーザの両足の重心位置 f_y ，頭部位置 h_y が算出され，これらに基づき歩行装置の制御 (図 47 中 (1)) および描画用クライアントの制御 (図 47 中 (2)) が行われる．以下では，まず，歩行動作の検出について述べ，次に歩行装置の制御法について述べる．

歩行動作は，図 48 に示すように，ユーザが立つ位置の前方および後方に設置されたレーザレンジファインダにより検出される．前方のものにより両足首，後方のものにより頭部が検出される．レンジデータからのそれぞれの足部および頭部位置の算出では，基本的には同じアルゴリズムを用いている．まず共通部分について説明し，次に足部検出処理のみに適用されるアルゴリズムを説明する．

本研究で使用するレーザレンジファインダ LMS200 は，75Hz で 180 度の範囲 401 点の奥行きを計測できる．図 50 は足部の検出結果に，図 49 は頭部の検出結果を示している．ただし，両図中の点はレンジデータを表し，長方形は設定した検出領域，放物線は曲線当てはめにより得た 2 次関数，横線は最終的な結果とし

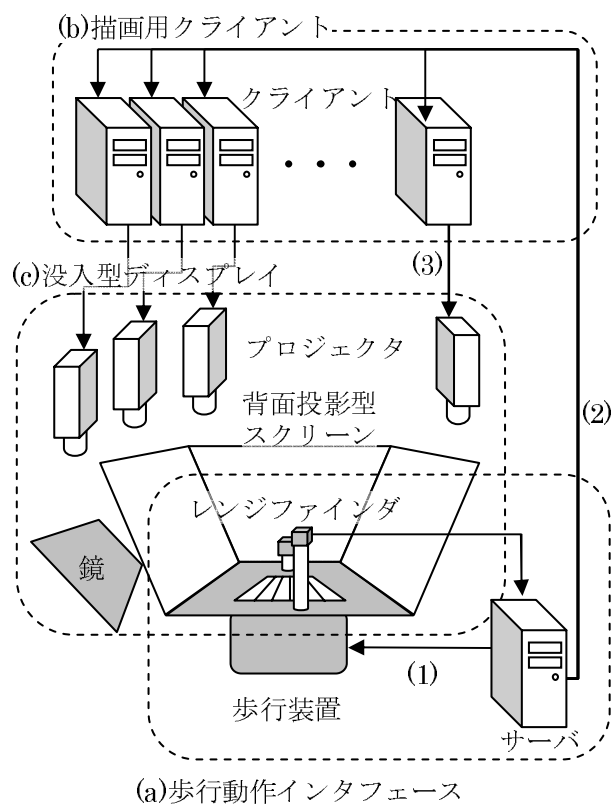


図 47 歩行装置を用いた画像提示システム.

て使用する足部もしくは頭部の検出位置を表す．LMS200 により取得したレンジデータから，以下の手順により足部および頭部の位置を算出する．

1. メディアンフィルタによりレンジデータのノイズを除去する
2. 予め設定された範囲でレンジデータをクリッピングする
3. α, β, γ を係数とする 2 次関数 $y = \alpha(x + \beta)^2 + \gamma$ の曲線当てはめる
4. 頂点座標 γ から足もしくは頭の半径分だけ差し引いたものを検出結果とする

ただし，足や頭の太さに関する係数 α は変化しないパラメータのため，システムの初期化時に取得し，2 次関数の当てはめの際は，定数として扱う．

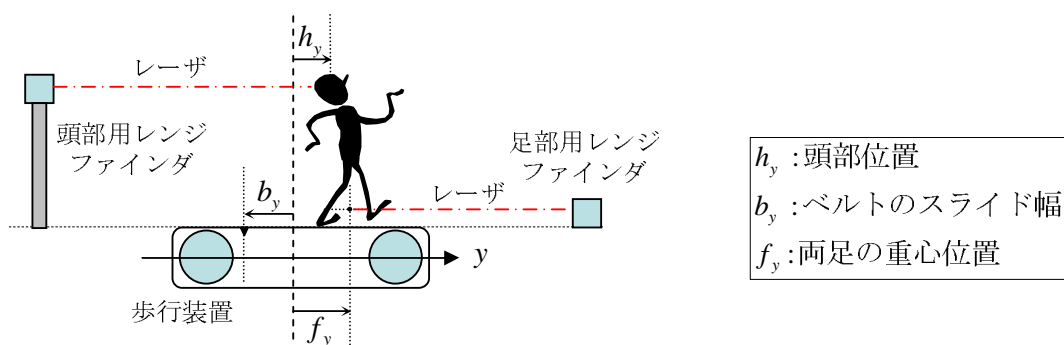


図 48 歩行装置の座標系.

特に，足部の検出においては，両足が交差する場合と両足が揃って分離できない場合についても考慮する．この場合，システム初期化の際に足の太さを計測しておき，この値から両足が揃っているのか，交差しているのか判断する．両足が交差する場合には，後ろの足の位置が算出できないので，最後に後ろの足を計測した時点での位置を現在の足の位置として使用する．

歩行装置のベルトは利用者の両足の重心位置 f_y がベルト領域の中央に位置するように制御される [Iwa99]．ベルトのスライド幅が b_y ，仮想環境内でのユーザの視点移動量 v は $v = h_y - b_y$ で表される．サーバによって算出されたユーザの視点移動量 v は，全方位画像撮影時の全方位型マルチカメラシステムの移動量に対応し，逐次，フレームインデックスに変換され，頭部位置 h_y と共に描画用クライアントに送信される．

(b) 描画用クライアント

図 47 (b) に示すように，描画用クライアントは，12 台の PC により構成される．クライアントとサーバの間は，100Mbps の LAN により接続され，ネットワークを介してクライアントが制御される．頭部位置 h_y が 0 のときの画像を JPEG 形式で圧縮したものが，予めクライアントのハードディスクに蓄積されており，ユーザの頭部位置が変化したときは，グラフィックハードウェアの機能を利用して，頭部位置に応じてワーピングされた画像が描画される．

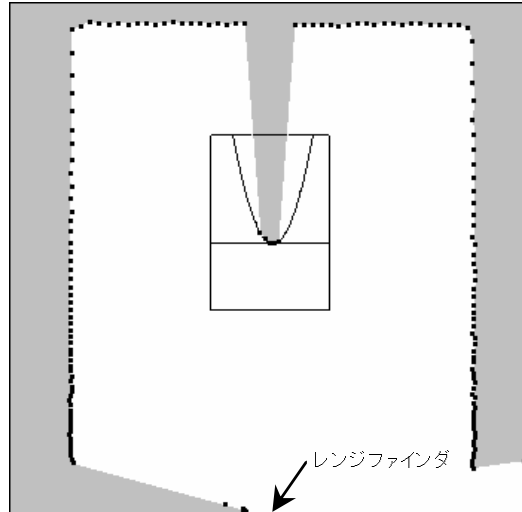


図 49 頭部のレンジデータと検出結果

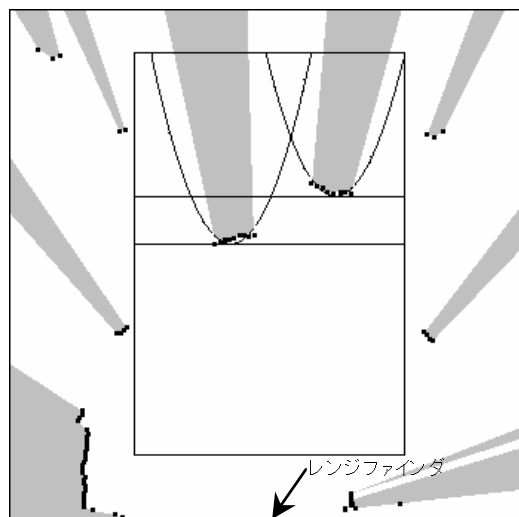


図 50 足部のレンジデータと検出結果

(c) 没入型ディスプレイ

図 47 (c) に示すように，没入型ディスプレイは高解像度な画像を提示するために，3つの傾斜型平面スクリーンと12台のプロジェクタにより構成される．4つのプロジェクタからの映像が反射鏡を介して1つのスクリーンに背面投影され，高解像度な画像がユーザに提示される．ユーザに広い視野を提供するために，スクリーンはユーザの前面および左右に配置されている．各プロジェクタの解像度は1,024 × 768 (XGA) である．ただし，複数のプロジェクタから1つのスクリーンに重複部分を持って画像が投影されるため，各スクリーンの解像度は，実質 200 万画素程度である．

4.3.2 カメラパラメータに基づく広視野角動画コンテンツの生成

提示画像は，内部・外部パラメータに基づいて再生成された動画像，およびユーザの視点移動量と動画像のフレームインデックスとの対応関係により構成される．提示画像の生成は，全方位型マルチカメラシステムによる全方位画像の取得，カメラパラメータに基づく画像生成の手順で行われる．以下，各手順について詳述する．

全方位画像の取得

画像の取得では，人間が歩いているときの視界を再現するために，全方位型マルチカメラシステムは設定された経路上を人間の目の高さに移動させ，環境を撮影する．この際，画像提示時の処理を簡単化するために，可能な限り全方位型マルチカメラシステムを一定速度で移動させる．ただし，高精度にカメラの速度を一定に保つことは難しいため，全方位型マルチカメラシステムの外部パラメータを利用して，画像提示時にユーザの歩行移動量と画像フレームを対応付ける．

提示画像の生成

第2章および第3章で述べた手法により推定されたカメラパラメータを用いて，提示画像の生成およびユーザ視点移動量と動画像のフレームインデックスとの対

応関係の算出を行う。提示画像の生成では、第2章2.4節で述べた手法にもとづき、予め光学的なパラメータを用いて周辺減光およびカメラ間の色調補正された画像を入力として、全方位型マルチカメラシステムの内部および外部パラメータを用いた画像の幾何学的な変換を行う。画像の幾何学的な変換では、ユーザに幾何学的に正しい視界を再現するために、仮想的にユーザの頭部位置 $H \equiv [h_x, h_y, h_z, 1]^T$ に応じて、スクリーンに投影する画像を生成する。このために、スクリーン上の点 $P \equiv [p_x, p_y, p_z, 1]^T$ の三次元位置をカメラパラメータを用いて画像面上に投影し、点 P に対応する画像上の位置 u を算出することで、スクリーン上の点 P の輝度値を決定する。ただし、第2章2.5節で述べたように、完全な一点透視投影画像は生成できない。そこで、画像取得環境のスケール s に合わせて、点 P を頭部位置 H を中心に s 倍に拡大した点 P' からの投影を考える。また、仮想的に頭部位置に配置する全方位型マルチカメラシステムを、推定された全方位型マルチカメラシステムの姿勢と等しくすることにより、カメラの傾きや揺れによる効果を除去する。以上の画像の幾何学的変換処理は、頭部位置 H から P' への相対位置 $P'' \equiv [p_x - h_x, p_y - h_y, p_z - h_z, 1/s]$ を用いて、以下の式で表される。

$$u = d_c(N_{c_i} R_i P'') \quad (21)$$

ただし、 $d_c(X)$ は、カメラ c のカメラ座標系での三次元位置 X から入力画像上での位置 u への変換関数である。また、屋外環境のように遠景領域が支配的な場合は、 $s \rightarrow \infty$ として、無限遠に点 P'' を設定する。 R_i は、没入型ディスプレイと撮影環境にそれぞれ固定された座標系との間の回転を表し、各フレーム i での全方位型マルチカメラシステムの移動経路に対する接線方向の画像がユーザの正面に提示されるように与える。

全方位型マルチカメラシステムの位置情報は、初期フレームからの全方位型マルチカメラシステムの移動量を算出し、ユーザの視点移動量 v とフレームインデックス i との対応関係 $i(v)$ を得るのに使用される。この対応関係 $i(v)$ は以下の手順で算出される。

1. 撮影時の Ladybug の投影中心の重心位置を各フレームごとに算出する
2. 重心位置の軌跡を平均フィルタでスムージングする



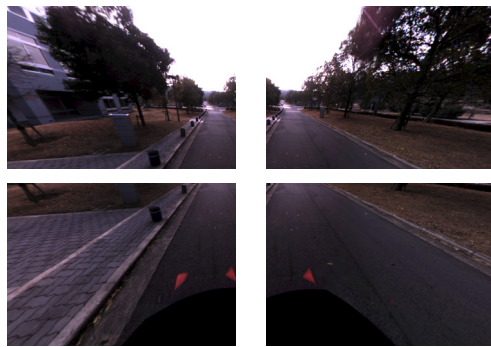
図 51 取得した動画像の 1 フレーム

3. 各フレームに対する初期位置からの Ladybug の撮影経路に沿った距離を算出する

4.4 実験

実験では、まず、提示画像の生成結果および画像提示システムの動作が設計方針を満たすことを確認するために、画像提示システムによりユーザに提示される画像を確認する。次に、主観評価実験を通して、4.1 節で述べたユーザの臨場感の向上のための方策である 4 項目 (A) ~ (D) それぞれの有効性を確認する。

両実験に用いる提示画像は、3.4.2 節の実験で取得し、4.3.2 節で述べた手法により生成したものである。4.3.2 節で述べた手法により、生成した画像を図 52 に示す。



(a) スクリーン正面用画像



(b) スクリーン左側面用画像

(c) スクリーン右側面用画像

図 52 蓄積されている動画の 1 フレーム

4.4.1 画像提示結果の定性的評価

画像提示システムによりユーザに提示される画像を確認することで、試作システムの動作を総合的に確認する。システムにより提示される画像は、ユーザの頭部位置、歩行移動により変化する。まず、頭部位置および歩行移動に依存しない部分について、提示画像が正しく生成されていることを確認し、さらに、頭部の運動依存する画像提示、歩行移動に依存する画像提示結果を順に確認する。

まず、異なるプロジェクタにより投影される画像の連続性を調べることにより、提示画像が正しく生成されていることを確認する。図 53 は、頭部位置にカメラを配置し、ディスプレイを撮影した画像である。提示画像は、直線的な物体のエッジなどが直線的に見え、異なるプロジェクタにより投影された画像間においても幾何学的な不自然さは感じられなかった。異なるスクリーン間においては、図 53 に示すように、スクリーン間の境界部分による不連続や 2 章 2.5 節で述べたカメラ間の視差による不連続が見られるが、これらの部分を除けば、提示されている画像には、幾何学的な不整合は認知できなかった。

次に、頭部位置の変動に伴う提示画像の見え方の変化および提示遅延の影響を調べた。図 54(a) は、ユーザが仮想環境内のある地点に静止しており、頭部位置が歩行装置のベルトの領域内の中央から前よりにある場合のユーザの視野の一部を示している。それに対して、図 54(b) は、図 54(a) と同一地点に静止しており、図 54(a) の場合よりも、頭部位置が後ろよりにある場合のユーザの視界に相当する画像を示している。両図に示されているスクリーン間の継ぎ目部分は、通常、頭部位置を無視して画像を提示した場合、実環境中の直線的な部分が折れ曲がって見えるなどの、幾何学的な不整合が生じやすい部分である。試作システムにおいては、提示画像に幾何学的な違和感は感じられず、提示画像に頭部位置の検出結果が反映されていることを確認した。ただし、試作システムでは、頭部位置が撮影系路上を外れる場合、正しい画像は提示されないという制限がある。試作システム上でユーザ歩行時の頭部位置を計測したところ、横方向の最大変位は約 50cm で、前後方向の約 150cm に比べて小さかった。図 54(c),(d) に示すように頭部位置が、左右に変位している場合でも提示画像の幾何学的な不整合による違和感は感じられなかった。また、足部および頭部の動きが提示画像に反映されるまでの



(a) 正面方向



(b) 右前方向



(c) 左前方向



(d) 右方向

图 53 画像提示結果



(a) 頭部位置 (h_x, h_y) が $(75\text{cm}, 0\text{cm})$ の場合 (b) 頭部位置 (h_x, h_y) が $(-75\text{cm}, 0\text{cm})$ の場合



(c) 頭部位置 (h_x, h_y) が $(0\text{cm}, 25\text{cm})$ の場合 (d) 頭部位置 (h_x, h_y) が $(0\text{cm}, -25\text{cm})$ の場合

図 54 頭部位置を考慮した画像提示結果

遅延は、振動する物体をレンジファインダにより計測し、物体の運動と提示される画像との位相差を調べることで計測した結果、17msecであった。意図的にすばやく足部および頭部を前後に動かした場合に、提示画像に遅延を感じることはあったが、歩行装置の限界速度である時速約 1.6km/h で歩行している場合にはそのような遅延は感じられなかった。

最後に、歩行装置上を歩行運動した際の提示画像の確認として、画像更新レートを調べた。その結果、最大歩行可能速度である 1.6m/sec 程度で歩行した場合、40fps で更新されていた。7.6m/sec でカメラを移動させ 15fps で画像を取得していることから、取得したフレーム全てが描画されていることが分かる。

4.4.2 主観評価による有効性の検証

構築した能動的視点移動型テレプレゼンスシステムでは、ユーザの視点移動が撮影系路上に制限されるため違和感が発生すると考えられる。この違和感により、4.1 節で述べたユーザの臨場感の向上のための方策である 4 項目 (A) 広視野・高解像画像の提示、(B) 提示画像のスタビライジング、(C) 歩行動作インタフェースの利用、(D) ヘッドトラッキングのそれぞれが無効化する可能性がある。本実験では、ユーザの臨場感を主観評価し、これらの項目の有効性を確認する。

評価方法

まず、試作システムの評価方法について説明する。テレプレゼンスシステム使用時の臨場感を絶対的な指標により定量化することは難しく、主観評価においても被験者の負担が大きい。そこで、2つのシステム X、Y を比較し、被験者はどちらのシステムを使用した時が臨場感が高かったかを「Xの方が臨場感が高い」「Yの方が臨場感が高い」「どちらとも言えない」の三択で評価する方式を採用する。被験者に説明する臨場感の定義は「映像に写っている場所にいる感覚」とした。具体的には、比較用システム A~D を設定し、それぞれを試作システムと比較することにより、ユーザの臨場感の向上のための方策である 4 項目 (A) 広視野・高解像画像の提示、(B) 提示画像のスタビライジング、(C) 歩行動作インタフェース

の利用, (D) ヘッドトラッキングの使用が, それぞれユーザの回答に影響するか調べる. 比較用システムは, 以下のように設定し, 提案システムとしては, ユーザの臨場感の向上のための方策 (A) ~ (D) を採用したものをを用いた.

システム A 方策 (A) は採用せず, 一般的なカメラの視野角および解像度を再現する意味で, 640×480 画素相当の解像度画像を前面スクリーンのみに表示する. 方策 (B) ~ (D) は採用する.

システム B 方策 (B) は採用せず, Ladybug の外部パラメータを全てのフレームで一定値に固定し, 撮影中のカメラの姿勢変動等が提示画像に表れるようにする. 方策 (A) および (C), (D) は採用する.

システム C 方策 (C) は採用せず, 歩行装置のベルトを停止させ, 市販のゲームコントローラにより歩行スピードを $0 \sim 2.0$ [km/h] の範囲で指定できるようにする. 方策 (A) および (B), (D) は採用する.

システム D 方策 (D) は採用せず, 頭部位置が没入型ディスプレイの中央に固定されているものとして画像を提示する. 方策 (A) ~ (C) は採用する.

ただし, これらの実験では, 実験順序による結果の依存性を軽減するために, 被験者は, 事前に 5 分程度各システムを体験し, 各システムの評価実験間の切り替えを 10 秒以内にした. このように事前にシステムを体験させることで, 被験者が評価実験中に学習する効果を軽減でき, 評価実験の切り替えを短時間で実施することで, 直前のシステムの使用感覚を忘却する効果を軽減できると考えられる.

評価結果

実験の目的および条件の詳しい知識を有しない 10 名を被験者として, 上述の実験を行なった. 評価結果を表 11 に示す. 比較用システム A ~ C と試作システムの比較に関しては, 有意水準 5 % のもとで統計的有意な差が確認された. これは, 多くの被験者が試作システムの特徴である, (A) 広視野・高解像画像の提示, (B) 提示画像のスタビライジング, (C) 歩行動作インタフェースの利用が, 臨場感の向上という点について有効であることを示している. ただし, システム D との

比較に関しては、「試作システムの方が臨場感が高い」と答えた被験者の数が「比較用システムの方が臨場感が高い」と答えた被験者の数と同じであり、「どちらとも言えない」と答えた被験者が一番多かった。これは、比較用システム D には採用されていない (D) ヘッドトラッキングの機能が、試作システム上での効果が低かったことを意味する。

ヘッドトラッキングの効果が低い点に対する考察を述べる。この結果の主な原因は、歩行中に被験者の頭部位置があまり動かなかったことが考えられる。実際、被験者の頭部位置は、一定速度で歩いている時は、標準偏差 10cm 程度の変動であり、図 54 で示した実験時の 20% であった。逆に、頭部位置を最大幅 150cm、標準偏差 32cm で動かす 1 人の被験者は、「試作システムの方が臨場感が高い」に投票している。このことから遠隔地での散歩を再現するシステムやトレーニング支援など、特に頭部位置を積極的に動かさないアプリケーションでは、ヘッドトラッキングの効果は小さいと考えられる。

4.5 結言

本章では、臨場感の高い能動的視点移動型テレプレゼンスシステムの構築を目的とし、全方位型マルチカメラシステムにより取得した画像を用いたシステムについて述べた。1 章 1.3 節の議論をもとにしたシステムの設計方針に従い、画像提示システムと画像生成手法を実装した。画像提示システムとして、没入型ディスプレイと歩行装置を用い、頭部位置の検出に基づく画像提示を行う。画像の生

表 11 評価結果

比較用システム	比較用システムの方が臨場感が高い	どちらとも言えない	試作システムの方が臨場感が高い
A	0	1	9
B	0	0	10
C	2	0	8
D	1	8	1

成においては，前章までに述べた全方位型マルチカメラシステムの外部パラメータを利用した画像生成を行い，カメラの姿勢変動および移動速度変動の影響を除去する．

実験では，まず画像提示システムの定性的評価を行い，システムが設計方針を満たすことを確認した．主観評価実験により，(A) 広視野・高解像画像の提示，(B) スタビライジング，(C) 歩行動作インタフェースの利用，(D) ヘッドトラッキングのそれぞれを利用する効果を確認した．その結果，少なくとも (A)～(C) の利用により，多くの被験者の臨場感が向上することを確認した．このことは，能動的視点移動型テレプレゼンスシステムの1つである歩行装置を用いたシステムにおいて，前章までに提案するカメラ内部・外部パラメータの推定技術を利用することが臨場感の向上という点に対して有効であることを示している．

5. 結論

5.1 本論文のまとめ

本研究では、遠隔地の広域実環境の臨場感を再現するテレプレゼンスシステムの構築技術として、画像生成および提示手法を提案した。テレプレゼンスシステムとしては、基本的には、要求される視点位置や視線方向に応じた実画像を提示し、実環境における視界変化を再現するものを想定し、臨場感の向上を目的として全方位型マルチカメラシステムにより取得した高解像全方位画像を用いる。こういったテレプレゼンスシステムは利用形態により、ユーザが受動的に視点移動する受動的視点移動型テレプレゼンスシステムと能動的に視点移動する能動的視点移動型テレプレゼンスシステムの2つの方式に分けることができる。提案する画像生成手法では、事前に全方位型マルチカメラシステムのカメラ内部パラメータの推定を行い、シームレスな全方位画像を生成する。特に、能動的視点型では、撮影時の全方位型マルチカメラシステムのカメラ外部パラメータを推定し、カメラの移動速度や姿勢の変動を除去した画像提示により、臨場感の向上を図る。

本論文では、実際に受動的および能動的視点移動型の両方式のテレプレゼンスシステムを構築し、上述の技術により臨場感の高いテレプレゼンスシステムが実現可能であることを示した。このことから2章および3章の技術を用いることにより、従来提案されている様々なテレプレゼンスシステムの性能の向上が期待できる。特に、受動的に視点移動する方式のテレプレゼンスシステムは、現状のシステムにおいても周辺機器等の性能不足を除けば大きな問題は残されておらず、今後の画像提示装置等の発展により、ジェットコースターの体験アトラクションや観光案内など、人物や車などの動物体を含む景観を再現するアプリケーションの実用化が十分期待できる。能動的に視点移動する方式のテレプレゼンスシステムでは、視点移動の自由度や動物体の表現など幾つかの問題は残っているが、木道や散策路などの比較的細い道上を移動し周囲の景観再現するような限定的なアプリケーションについては、実現が期待できる。



図 55 複数の経路が交差する環境の例

5.2 今後の展望

本研究に残された課題について述べる．能動的に視点移動する方式のテレプレゼンスシステムに残された課題を以下にあげる．

複数経路の統合による視点移動範囲の拡大 歩行装置を用いた能動的視点移動型テレプレゼンスシステムにおいて，視点移動範囲を拡大するために，図 55 に示すような環境の再現として，複数の交差する経路を組み合わせる方法の検討が考えられる．このようなシステムでは，経路間を移動する際に，ユーザの歩行方向や歩行移動量と提示される画像が対応しない場合，違和感を感じるという問題がある．このような違和感を最小限におさえる画像提示手法を検討する必要がある．

動物体の表現 動画像の撮影経路上の視点移動感覚を再現する能動的視点移動型テレプレゼンスシステムでは，動物体の動きが正しく表現されない．しかし，本研究で提案した提示画像の生成手法では，全方位型マルチカメラシステムのカメラ外部パラメータを用いるため，第 3 章 3.4 節の実験で示したように，動画像中に仮想物体を配置し，環境中にあたかもその物体が存在

するかのように表現することができる。この技術を用いて、動きのある仮想物体を環境中に付加することで動物体の表現が可能であると考えられる。

謝辞

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、御鞭撻頂いた像情報処理学講座 千原國宏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 智和助手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、実験、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 守屋 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。特に、本研究における実験データの取得等、数多くの場面において協力してくれた天目 隆平氏、奥村 文洋氏、福富 弘敦氏、森田 真司氏、浅井 俊弘氏、横地 裕次氏、和田 浩明氏に感謝申し上げます。

参考文献

- [A. 80] A. Lippman. Movie-Map: An application of the optical video-disc to computer graphics. In *Proc. SIGGRAPH*, pp. 32–42, 1980.
- [B. 86] B. K. P. Horn. *Robot Vision*, chapter 10, pp. 206–209. Mit Press, 1986.
- [C. 02] C. J. Taylor. VideoPlus: A method for capturing the structure and appearance of immersive environment. *IEEE Trans. Visualization and Computer Graphics*, Vol. 8, No. 2, pp. 171–182, 2002.
- [CB61] C. Comeau and J. Bryan. Headsight television system provides remote surveillance. *Electronics*, pp. 86–90, 1961.
- [CCEE98] G. U. Carraro, M. Cortes, J. T. Edmark, and J. R. Ensor. The peloton bicycling simulator. In *Proc. 3rd Symp. on Virtual Reality Modeling Language (VRML '98)*, pp. 63–70. ACM Press, 1998.
- [Che95] S. Chen. QuickTime VR: An image-based approach to virtual environment navigation. In *Proc. SIGGRAPH '95*, pp. 29–38, 1995.
- [CMC03] A. I. Comport, É. Marchand, and F. Chaumette. A real-time tracker for markerless augmented reality. In *Proc. 2nd ACM/IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR2003)*, pp. 36–45, 2003.
- [CNSD⁺92] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart. The Cave - audio visual experience automatic virtual environment. *Communication of the ACM*, Vol. 35, No. 6, pp. 64–72, 1992.
- [Dav03] Andrew J. Davison. Real-time simultaneous localisation and mapping with a single camera. *Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV2003)*, Vol. 2, pp. 1403–1410, 2003.
- [DTF⁺01] D. Kotake, T. Endo, F. Pighin, A. Katayama, H. Tamura, and M. Hirose. Cybercity Walker 2001 : Walking through and looking around

a realistic cyberspace reconstructed from the physical world. In *Proc. 2nd IEEE and ACM Int. Symp. on Mixed Reality (ISMR2001)*, pp. 205–206, 2001.

- [FC200] FC2. <http://www.sony.co.jp/products/fourthview>. 2000.
- [FZ98] A. W. Fitzgibbon and Andrew Zisserman. Automatic camera recovery for closed or open image sequences. In *Proc. 5th European Conf. on Computer Vision*, Vol. I, pp. 311 – 326, 1998.
- [GADS04] S. M. Goza, R. O. Ambrose, M. A. Diftler, and I. M. Spain. Telepresence control of the nasa/darpa robonaut on a mobility platform. In *Proc. SIGCHI Conf. on Human factors in Computing Systems (CHI '04)*, pp. 623–629. ACM Press, 2004.
- [GF03] S. Güven and S. Feiner. Authoring 3D hypermedia for wearable augmented and virtual reality. In *Proc. 7th IEEE Int. Symp. on Wearable Computers*, pp. 118–126, 2003.
- [HKCY00] H. Tanahashi, K. Yamamoto, C. Wang, and Y. Niwa. Development of a stereo omnidirectional imaging system (SOS). In *Proc. IEEE Int. Conf. on Industrial Electronics, Control and Instrumentation (IECON2000)*, pp. 289–294, 2000.
- [HS88] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conf.*, pp. 147–151, 1988.
- [Iwa99] H. Iwata. Walking about virtual environments on an infinite floor. In *Proc. IEEE Virtual Reality '99*, pp. 286–293, 1999.
- [J. 93] J. Z. C. Lai. On the sensitivity of camera calibration. *Image and Vision Computing*, Vol. 11, No. 10, pp. 656–664, 1993.

- [J. 04] J. Mulligan, X. Zabulis, N. Kelshikar, and K. Daniilidis. Stereo-based environment scanning for immersive telepresence. *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 14, No. 3, pp. 171–182, 2004.
- [JHNK00] J. Shimamura, H. Takemura, N. Yokoya, and K. Yamazawa. Construction of an immersive mixed environment using an omnidirectional stereo image sensor. In *Proc. IEEE Workshop on Omnidirectional Vision*, pp. 62–69, 2000.
- [Jou02] Norman P. Jouppi. First steps towards mutually-immersive mobile telepresence. In *Proc. 2002 ACM Conference on Computer Supported Cooperative Work (CSCW '02)*, pp. 354–363. ACM Press, 2002.
- [JXB⁺91] J. Hong, X. Tan, B. Pinette, R. Weiss, and E. M. Riseman. Image-based homing. In *Proc. Int. Conf. on Robotics and Automation*, pp. 620–625, 1991.
- [KHN02] K. Yamazawa, H. Takemura, and N. Yokoya. Telepresence system with an omnidirectional HD camera. In *Proc. 5th Asian Conf. on Computer Vision (ACCV2002)*, Vol. 2, pp. 535–538, 2002.
- [KVJ03] K. J. Fernandes, V. Raja, and J. Eyre. Cybersphere: The fully immersive spherical projection system. *Communications of the ACM*, Vol. 46, No. 9, pp. 141–146, 2003.
- [LF92] B. Lauel and S. Fisher. Be there here. *InterCommunication*, 1992.
- [MAS⁺04] M. Uyttendaele, A. Criminisi, S. B. Kang, S. Winder, R. Hartley, and R. Szeliski. High-quality image-based interactive exploration of real-world environments. *IEEE Computer Graphics and Applications*, 2004.
- [Min80] M. Minsky. Telepresence. *Omni*, Vol. 2, pp. 45–52, 1980.

- [Mov99] Movingeye. <http://www.movingeye.co.jp/>. 1999.
- [NAM96] N. Asada, A. Amano, and M. Baba. Photometric calibration of zoom lens systems. In *Proc. Int. Conf. Pattern Recognition*, Vol. A, pp. 186–190, 1996.
- [PGA⁺00] P. Peixoto, J. Goncalves, H. Antunes, J. Batista, and H. Araujo. A surveillance system integrating visual telepresence. In *Proc. 15th Int. Conf. on Pattern Recognition (ICPR'00)*, Vol. 4, 2000.
- [PKV⁺00] M. Pollefeys, R. Koch, M. Vergauwen, B. Deknuydt, and L. V. Gool. Three-dimensional scene reconstruction from images. In *Proc. SPIE*, Vol. 3958, pp. 215–226, 2000.
- [Poi02] Point Grey Research, Inc. *Ladybug Omnidirectional Camera System User Guide version 1.0*, 2002.
- [R. 87] 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.
- [S. 97] S. K. Nayar. Catadioptric omnidirectional camera. In *Proc. Computer Vision and Pattern Recognition*, pp. 482–488, 1997.
- [SE87] S. J. Oh and E. L. Hall. Guidance of a mobile robot using an omnidirectional vision navigation system. In *Proc. Mobile Robots IISPIE 852*, pp. 288–300, 1987.
- [SKN03] S. Morita, K. Yamazawa, and N. Yokoya. Internet telepresence by real-time view-dependent image generation with omnidirectional video camera. In *Proc. SPIE Electronic Imaging*, Vol. 5018, pp. 51–60, 2003.
- [SKYT02] T. Sato, M. Kanbara, N. Yokoya, and H. Takemura. Dense 3-D 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, pp. 119–129, 2002.
- [TG00] A. Triesman and G. Gelade. A feature integration theory of attention. *Cognitive Psychology*, Vol. 12, pp. 298 – 375, 2000.
- [THKM00] T. Takahashi, H. Kawasaki, K. Ikeuchi, and M. Sakauchi. Virtual driving system with real-world image. In *CD-ROM Proc. 7th World Congress on Intelligent Transportation Systems*, 2000.
- [TKHN98] T. Kawanishi, K. Yamazawa, H. Takemura, and N. Yokoya. Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors. In *Proc. 14th Int. Conf. on Pattern Recognition (ICPR '98)*, Vol. 1, pp. 445–489, 1998.
- [TMNH02] 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*, Vol. 47, No. 1-3, pp. 110–129, 2002.
- [TMY98] T. Yamada, M. Hirose, and Y. Iida. Development of complete immersive display. In *Proc. 4th Int. Conf. on Virtual Systems and Multimedia (VSMM'98)*, Vol. 2, pp. 522–527, 1998.
- [TPMF00] B. Triggs, R. Hartley P. McLauchlan, and A. Fitzgibbon. Bundle adjustment a modern synthesis. *Vision Algorithms: Theory and Practice*, pp. 298 – 375, 2000.
- [TPP97] T. Kanade, P. Rander, and P. J. Narayanan. Virtualized reality: Constructing virtual worlds from real scenes. *IEEE MultiMedia*, Vol. 4, No. 1, pp. 34–47, 1997.
- [VLF04] L. Vacchetti, V. Lepetit, and P. Fua. Combining edge and texture information for real-time accurate 3D camera tracking. In *Proc. 3rd*

IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR '04), pp. 48–57, 2004.

- [YKNH98] 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*, Vol. 71, No. 2, pp. 154–165, 1998.
- [YYM02] Y. Yagi, Y. Nishizawa, and M. Yachida. Estimating location and avoiding collision against unknown obstacle for the mobile robot using omnidirectional image sensor COPIS. In *Proc. Int. Workshop on Intelligent Robots and Systems*, pp. 909–914, 2002.
- [ZSE86] Z. L. Cao, S. J. Oh, and E. L. Hall. Dynamic omnidirectional vision for mobile robots. *Jour. Robotic Systems*, Vol. 3, No. 1, pp. 5–17, 1986.
- [岩下 04] 岩下 克, 外山 篤, 橋本 直己, 長谷川 晶一, 佐藤 誠. 足踏み動作を用いた移動インターフェースの開発. 電子情報通信学会論文誌, Vol. J87-A, No. 1, pp. 87–95, 2004.
- [戎野 99] 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.
- [戎野 00] 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.
- [大津 80] 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, Vol. 4, No. J63-D, pp. 349–356, 1980.
- [栗田 00] 栗田 多喜夫, 坂上 勝彦. ロバスト統計とその画像理解への応用. 画像の認識・理解シンポジウム (MIRU2000) 講演論文集, Vol. I, pp. 65–70, 2000.

- [佐藤 03] 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.
- [佐藤 05a] 佐藤智和, 池田聖, 横矢直和. 複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定. 電子情報通信学会論文誌 (D-II), No. 2, pp. 347-357, 2005.
- [佐藤 05b] 佐藤 智和, 池田 聖, 横矢 直和. 複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定. 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 2, pp. 347-357, 2005.
- [島村 01] 島村 潤, 山澤 一誠, 竹村 治雄, 横矢 直和. 全周パノラマステレオ画像とCGモデルの合成による複合現実環境の構築. 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol. 42, No. SIG6(CVIM2), pp. 44-53, 2001.
- [棚橋 02] 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, Vol. 56, No. 4, pp. 603-610, 2002.
- [出口 90] 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. コンピュータビジョン '90 シンポジウム 論文集, pp. 41-50, 1990.
- [森田 05] 森田 真司, 山澤 一誠, 寺沢 征彦, 横矢 直和. 全方位画像センサを用いたネットワーク対応型遠隔監視システム. 電子情報通信学会論文誌, Vol. J88-D-II, No. 5, pp. 864-875, 2005.
- [山澤 98] 山澤 一誠, 尾上 良雄, 横矢 直和, 竹村 治雄. 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, Vol. J81-D-II, No. 5, pp. 880-887, 1998.
- [山澤 02] 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, Vol. J79-D-II, No. 5, pp. 698–707, 2002.

[米田 00] 米田 美里, 山澤 一誠, 竹村 治雄, 横矢 直和. 全方位動画像からの両眼ステレオ画像生成による移動ロボットの遠隔操縦. 電子情報通信学会 技術研究報告, PRMU2000-107, 2000.

業績リスト

学術論文誌

1. 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を用いた高解像度な全天球パノラマ動画の生成とテレプレゼンスへの応用” 日本バーチャルリアリティ学会論文誌, Vol. 8, No. 4, pp. 443-450, Dec. 2003. (第2章に関連)
2. 佐藤 智和, 池田 聖, 横矢 直和: “複数動画からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定” 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 2, pp. 347-357, Feb. 2005. (第3章に関連)
3. 池谷 彰彦, 佐藤 智和, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “カメラパラメータ推定による紙面を対象とした超解像ビデオモザイク” 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 8, pp. 1490-1498, Aug. 2005.
4. 横地 裕次, 池田 聖, 佐藤 智和, 横矢 直和: “特徴点追跡とGPS測位に基づくカメラ外部パラメータの推定” 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol. 47, No. SIG5 (CVIM13), pp. 69-79, Mar. 2006. (第4章に関連)

レター

1. 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” 情報技術レターズ, Vol. 1, pp. 149-150, Sep. 2002. (第2章に関連)
2. 佐藤 智和, 池田 聖, 神原 誠之, 池谷 彰彦, 中島 昇, 横矢 直和, 山田 敬嗣: “紙面を対象としたカメラパス推定による高解像ビデオモザイク” 情報技術レターズ, Vol. 2, pp. 163-164, Sep. 2003.

3. 池谷 彰彦, 佐藤 智和, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “カメラパラメータ推定による超解像ビデオモザイク” 情報科学技術レターズ, Vol. 3, pp. 165-168, Sep. 2004.
4. 佐藤 智和, 池田 聖, 横矢 直和: “自然特徴点を利用したマルチベースラインステレオによる全方位奥行き画像推定” 情報科学技術レターズ, Vol. 3, pp. 161-164, Sep. 2004.

国際会議

1. S. Ikeda, T. Sato, and N. Yokoya: “A calibration method for an omnidirectional multi-camera system” Proc. SPIE Electronic Imaging, Vol. 5006, pp. 499-507, Jan. 2003. (第 2 章に関連)
2. S. Ikeda, T. Sato, and N. Yokoya: “High-resolution panoramic movie generation from video streams acquired by an omnidirectional multi-camera system” Proc. IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent System (MFI2003), pp. 155-160, July 2003. (第 2 章に関連)
3. S. Ikeda, T. Sato, and N. Yokoya: “Panoramic movie generation using an omnidirectional multi-camera system for telepresence” Proc. 13th Scandinavian Conf. on Image Analysis(SCIA2003), pp. 1074-1081, July 2003. (第 2 章に関連)
4. S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: “Telepresence system using high-resolution omnidirectional movies and a reactive display” Proc. IEEE and ACM Int. Symp. on Mixed Augmented Reality(ISMAR 03), pp. 292-293, Oct. 2003. (第 4 章に関連)
5. T. Sato, S. Ikeda, M. Kanbara, A. Iketani, N. Nakajima, N. Yokoya, and K. Yamada: “High-resolution video mosaicing for documents and photos by estimating camera motion” Proc. SPIE Electronic Imaging, Vol. 5299, pp. 246-253, Jan. 2004.

6. S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: “Immersive telepresence system using high-resolution omnidirectional movies and a locomotion Interface” Proc. SPIE Electronic Imaging, Vol. 5291, pp. 472-479, Jan. 2004. (第4章に関連)
7. T. Sato, S. Ikeda, and N. Yokoya: “Extrinsic camera parameter recovery of a moving omni-directional multi-camera system” Proc. Asian Conf. on Computer Vision (ACCV2004), Vol. I, pp. 306-311, Jan. 2004. (第3章に関連)
8. T. Sato, S. Ikeda, and N. Yokoya: “Extrinsic camera parameter recovery from multiple image sequences captured by an omni-directional multi-camera system” Proc. European Conf. on Computer Vision (ECCV2004), Vol. 2, pp. 326-340, May 2004. (第3章に関連)
9. S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: “An immersive telepresence system with a locomotion interface using high-resolution omnidirectional movies” Proc. 17th IAPR Int. Conf. on Pattern Recognition (ICPR2004), Vol. IV, pp. 396-399, Aug. 2004. (第4章に関連)
10. A. Iketani, T. Sato, S. Ikeda, M. Kanbara, N. Nakajima, and N. Yokoya: “Super-resolved video mosaicing for documents by extrinsic camera parameter estimation” Proc. Int. Conf. on Computer Vision and Graphics, Sep. 2004.
11. K. Yamazawa, T. Ishikawa, T. Sato, S. Ikeda, Y. Nakamura, K. Fujikawa, H. Sunahara, and N. Yokoya: “Web-based telepresence system using omni-directional video streams” Proc. 5th Pacific Rim Conf. on Multimedia, Vol. 3, pp. 45-52, Dec. 2004.
12. S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: “Immersive telepresence system using high-resolution omnidirectional video with locomotion inter-

- face” Proc. 14th Int. Conf. on Artificial Reality and Telexistence (ICAT 2004), pp. 188-193, Dec. 2004. (第4章に関連)
13. T. Ishikawa, K. Yamazawa, T. Sato, S. Ikeda, Y. Nakamura, K. Fujikawa, H. Sunahara, and N. Yokoya: “Networked telepresence system using web browsers and omnidirectional video streams” Proc. SPIE Electronic Imaging, Vol. 5664, pp. 380-387, Jan. 2005.
 14. S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: “Immersive telepresence system with a locomotion interface using high-resolution omnidirectional videos” Proc. IAPR Conf. on Machine Vision Applications (MVA2005), pp. 602-605, May 2005. (第4章に関連)
 15. A. Iketani, T. Sato, S. Ikeda, M. Kanbara, N. Nakajima, and N. Yokoya: “Video mosaicing for curved surface by 3D reconstruction using feature points” CD-ROM Proc. Int. Conf. on Computer Vision (ICCV2005), Demonstrations, Oct. 2005.
 16. Y. Yokochi, S. Ikeda, T. Sato, and N. Yokoya: “Extrinsic camera parameter estimation based-on feature tracking and GPS data” Proc. Asian Conf. on Computer Vision (ACCV2006), Vol. 1, pp. 369-378, Jan. 2006. (第3章に関連)
 17. A. Iketani, T. Sato, S. Ikeda, M. Kanbara, N. Nakajima, and N. Yokoya: “Super-resolved video mosaicing for documents based on extrinsic camera parameter estimation” Proc. Asian Conf. on Computer Vision (ACCV2006), Vol. 2, pp. 101-110, Jan. 2006.

研究会

1. 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の生成” 電子情報通信学会 技術研究報告, PRMU2002-154, Dec. 2002. (第2章に関連)

2. 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全天球動画の生成とテレプレゼンスへの応用” 日本バーチャルリアリティ学会 サイバースペースと仮想都市研究会 研究報告, Vol. 8, No. 2, pp. 31-36, May 2003. (第2章に関連)
3. 佐藤 智和, 池田 聖, 横矢 直和: “マルチカメラシステムによる全方位動画からのカメラ外部パラメータの復元” 情報処理学会 研究報告, CVIM141-13, Nov. 2003. (第3章に関連)
4. 池谷 彰彦, 中島 昇, 佐藤 智和, 池田 聖, 神原 誠之, 横矢 直和, 山田 敬嗣: “紙面を対象としたカメラパス推定によるビデオモザイクおよび超解像” 電子情報通信学会 技術研究報告, PRMU2003-223, Feb. 2004.
5. 池谷 彰彦, 中島 昇, 佐藤 智和, 池田 聖, 神原 誠之, 横矢 直和, 山田 敬嗣: “カメラパラメータ推定による紙面を対象とした超解像ビデオモザイク” 画像の認識・理解シンポジウム (MIRU2004) 講演論文集, Vol. I, pp. 505-510, July 2004.
6. 石川 智也, 山澤 一誠, 佐藤 智和, 池田 聖, 中村 豊, 藤川 和利, 砂原 秀樹, 横矢 直和: “Web ブラウザと全方位動画をを用いたテレプレゼンスシステム 車載全方位カメラや全方位型マルチカメラシステムを用いた実装 ” 画像の認識・理解シンポジウム (MIRU2004) 講演論文集, Vol. I, pp. 127-132, July 2004.
7. 佐藤 智和, 池谷 彰彦, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “カメラ外部パラメータ推定による平面を対象とした超解像ビデオモザイク” 第9回パターン計測シンポジウム講演論文集, pp. 13-20, Nov. 2004.
8. 横地 裕次, 池田 聖, 佐藤 智和, 横矢 直和: “動画とGPSによるハイブリッドなカメラ位置・姿勢推定法” 情報処理学会 研究報告 CVIM147-12, Jan. 2005. (第3章に関連)
9. 佐藤 智和, 池田 聖, 横矢 直和: “画像特徴点の数え上げに基づくマルチベースラインステレオ法” 情報処理学会 研究報告 CVIM148-20, March 2005.

10. 横地 裕次, 池田 聖, 佐藤 智和, 横矢 直和: “動画像と GPS による位置情報を用いたカメラ外部パラメータの推定” 画像の認識・理解シンポジウム (MIRU2005) 講演論文集, pp. 650-657, July 2005. (第 3 章に関連)
11. 佐藤 智和, 池谷 彰彦, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “湾曲した紙面を対象とした形状推定に基づくビデオモザイク” 画像の認識・理解シンポジウム (MIRU2005) 講演論文集, pp. 1638-1639, July 2005.
12. 佐藤 智和, 池谷 彰彦, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “動画像からの三次元復元に基づく湾曲した紙面のビデオモザイク” 電子情報通信学会 技術研究報告, PRMU, Feb. 2006. (掲載予定)

全国大会

1. 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による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, No. G13-27, Nov. 2002. (第 2 章に関連)
2. 池田 聖, 佐藤 智和, 神原 誠之, 横矢 直和: “高解像度全天球動画像と歩行装置を用いた没入型テレプレゼンスシステム” 情報科学技術フォーラム (FIT) 一般講演論文集, Vol. 3, No. K-098, Sep. 2003. (第 4 章に関連)
3. 池谷 彰彦, 中島 昇, 佐藤 智和, 池田 聖, 神原 誠之, 横矢 直和, 山田 敬嗣: “自由カメラ運動下でのビデオモザイクおよび超解像” 情報科学技術フォーラム (FIT) 一般講演論文集, Vol. 3, No. I-028, Sep. 2003.
4. 石川 智也, 山澤 一誠, 池田 聖, 横矢 直和: “Web ブラウザと全方位動画像を用いたテレプレゼンスシステム” 2004 年電子情報通信学会総合大会講演論文集, No. D-12-160, March 2004.
5. 池田 聖, 佐藤 智和, 横矢 直和: “フレーム間差分を利用した全天球動画像からの新視点画像合成” 2004 年電子情報通信学会総合大会講演論文集, No. D-12-140, March 2004.

6. 横地 裕次, 池田 聖, 佐藤 智和, 横矢 直和: “動画像と GPS 測位データからのカメラ運動の推定” 情報科学技術フォーラム (FIT) 一般講演論文集, Vol. 3, pp. 463-464, Sep. 2004. (第 3 章に関連)
7. 池田 聖, 佐藤 智和, 横矢 直和: “全天球動画像を用いた屋外環境の新視点画像生成” 情報科学技術フォーラム (FIT) 一般講演論文集, Vol. 3, pp. 129-131, Sep. 2004.
8. 佐藤 智和, 池谷 彰彦, 池田 聖, 神原 誠之, 中島 昇, 横矢 直和: “特徴点追跡による三次元復元に基づく湾曲した紙面のビデオモザイク” 2005 年電子情報通信学会総合大会講演論文集, No. D-12-12, March 2005.
9. 和田 浩明, 池田 聖, 佐藤 智和, 横矢 直和: “複数経路上を歩行可能なテレブレゼンスシステムにおけるユーザ支援手法” 日本バーチャルリアリティ学会第 10 回記念大会論文集, pp. 481-482, Sep. 2005.

解説

1. 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による高解像度な全天球パノラマ動画像の生成” 画像ラボ, Vol. 16, No. 2, pp. 5-9, Feb. 2005. (第 2 章に関連)