

博士論文

シーン構造記述に基づく
3次元形状モデルの生成・表現
に関する研究

平成15年7月

龍谷大学大学院 理工学研究科
電子情報学専攻
T03R503 正司 哲朗

博士論文

シーン構造記述に基づく
3次元形状モデルの生成・表現
に関する研究

平成15年7月

龍谷大学大学院 理工学研究科

電子情報学専攻

T03R503 正司 哲朗

シーン構造記述に基づく 3次元形状モデルの生成・表現に関する研究

要 旨

近年、実世界の環境を仮想世界に入力し、それを応用する研究が進んでいる。この研究は、産業、考古学、教育や娯楽などの分野に適用されている。しかしながら、実世界の環境は、複雑な構造情報をもつため、一般的に取得することは難しい。本論文では、実シーンの構造情報が明確ではない対象に対して、構造情報の取得と記述を目的とする。このため、シーンの構造情報の取得する際の手がかりとなる抽象モデルを導入する。抽象モデルとは、「実シーン(または対象物体)から抽象化された構造情報が記述されているもの」と定義する。例えば、抽象モデルを図式化したものには、スケッチ、鳥瞰図、立面図、平面図などがある。本論文では、抽象モデルを導入することにより、複雑な実シーンに対する構造情報の取得と記述ができ、3次元形状モデルを生成することが可能であることを示した。さらに、実シーンの構造情報を表現するために、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムを構築した。

屋内シーンを対象とした実験においては、抽象モデルを図式化した平面図と、実シーンから得られる構造情報をもとに、室内シーンの構造情報の取得と記述方法を提案した。平面図からでは、部屋の構造情報を取得することが可能であるが、室内にある対象物体の構造情報を得ることができない。そこで、ステレオ法を用いて室内にある複数の対象物体の構造情報を取得する。そして、平面図から得られた構造情報と実シーンから得られた構造情報を統合することにより、室内全体の3次元形状モデルを生成した。

屋外シーンを対象とした実験においては、抽象モデルを図式化した手書き立面図と実画像をもとに、石窟や仏塔の構造情報の取得と記述方法を提案した。ここでは、立面図の各面を1つの領域として扱い、視体積交差法を導入することにより、3次元形状モデルを生成した。また、手書き立面図から生成した3次元形状モデルは必ずしも実世界の対象とは一致しないという問題点があった。この問題を解決するために、立面図から生成した3次元形状モデルと実画像を用いて、対話的な処理により、3次元形状モデルを補正した。

拡張現実感システムの構築においては、プロジェクタと書籍画像、ブックスクリーンを利用した書籍画像表示インタフェースシステムを構築した。本システムは、キャリブレーション処理、幾何補正処理、ページ認識処理、画像表示処理の4つの処理からなる。キャリブレーション処理では、プロジェクタ座標系、画像座標系、世界座標系の3つの座標系のキャリブレーションを行った。幾何補正処理では、投影面(ブックスクリーン)が非平面であるため、投影する画像を補正する必要がある。そのため、マルチベースラインステレオ法を用いて、ブックスクリーン表面の3次元座標を求め、投影する画像の補正に必要な構造情報を取得した。ページ認識処理では、ブックスクリーンのページ番号を認識するためのバーコードを利用し、CCDカメラを用いて、ページ認識を行った。画像表示処理では、ページ認識結果からページに対応する画像を選択し、構造情報をもとに補正画像を生成した。そして、補正画像をブックスクリーンに投影した。さらに試作システムを用いて実験を行うことにより、本システムの有効性を示した。

A Study on Generation and Representation of 3-D Geometric Model Based on the Description of Scene Structure

Abstract

Recently, a technology that reconstructs real environment in virtual environment is studied in virtual reality(VR). This technology has been applied for field of industry, education, archaeology or entertainment. However, it is generally difficult to acquire structure information of scene from real world, because it has complicated structure information of scene. In this thesis, we studied on the fundamental techniques for acquiring and describing structure information of scene from real world. We propose a technique to acquire structure information from real world using "Abstract Model". In this thesis, "Abstract model" was defined as description of abstracted structure information of scene. For examples, bird's-eye view, a plane view or orthographical views are diagrammed "Abstract Model". Then, we acquire structure information of scene from real world by using a plane or orthographical views. Furthermore, in order to represent acquired structure information, we developed augmented reality system based on the correlation of human,real world and virtual world.

In experiment of indoor scene, we proposed a method to integrate the structure information acquired from a plane view and stereo method. First, structure information of objects in the room was acquired from real images based on stereo method. Next, structure information of the wall was acquired from the plane view. Then, 3-D geometric model was generated by structure information which is integrated these different structure information. As compared with multibaseline stereo method, the accuracy of the acquired structure information improved.

In experiment of outdoor scene, we proposed a method to generate 3-D geometric model from orthographical views. there are many methods for generating 3-D geometric model from orthographical views of machine parts with 2-D CAD, and vertices and edges information of these views are used. However, we are not able to generate 3-D geometric model by applying these method, because we use the orthographical views that are rough sketches of ancient structure and ruin. Our approach consists of the

following steps: First, we treated the orthographical views (front view, top view, side view) as faces. Next, we generated 3-D geometric model from these views based on the volumetric intersections method. Then, we projected the generated 3-D geometric model onto the 2-D plane, and corrected the model by comparing real images with the 2-D plane.

In construction of augmented reality system, we developed the book interface by using techniques of augmented reality. This system consists of the following processes: In the process of calibration, it was necessary to calibrate the projector coordinate systems, the image coordinate systems and the world coordinate systems. In the process of structure correction, as the projection surface (the book screen) was a non-plane, we acquired 3-D structure information on the surface of this book screen based on multibaseline stereo method. Then, in order to correct projected images, 2-D structure information were calculated from calibration parameter and acquired 3-D structure information on the surface of this book screen. In the process of page recognition, page number of the book screen was recognized by using bar code and CCD camera. In the process of image display, a book image was selected according to the recognized page number, and corrected image was generated by using the 2-D structure information. And, the corrected image was projected onto the book screen. Then, we represented book images by using this system, and had confirmed its effectiveness.

目次

1	序論	1
1.1	研究背景	1
1.2	研究目的	3
1.3	論文の構成	5
2	実シーンからの構造情報の取得・記述・表現	7
2.1	はじめに	7
2.2	シーン構造情報の取得	8
2.2.1	画像特徴を用いたシーン構造情報の取得	8
2.2.2	構造化光を用いたシーン構造情報の取得	10
2.3	シーン構造情報の記述と表現	11
2.3.1	構造情報の記述	12
2.3.2	構造情報の生成と表現	15
2.4	抽象モデルを用いたシーン構造情報の取得・記述・表現	19
2.4.1	抽象モデル	20
2.4.2	構造情報の取得と記述	22
2.4.3	3次元形状モデルの生成と表現	26
2.5	拡張現実感システムのためのシーン構造情報の表現	26
2.5.1	複合現実感	28
2.5.2	ヒューマンインタフェース	29
2.5.3	シーン構造情報の表現	30
2.6	まとめ	31

3	構造情報と抽象モデルを用いた3次元形状モデルの生成と記述	33
3.1	はじめに	33
3.2	実シーンからの構造情報の取得	35
3.2.1	3眼式マルチベースラインシステムの特徴	35
3.2.2	全周のパノラマ画像の取得	37
3.3	平面図の利用	38
3.3.1	平面図から構造情報の取得	38
3.3.2	平面図におけるカメラ位置の決定	41
3.4	構造情報の統合	42
3.5	3次元形状モデルの生成と構造情報の記述	44
3.6	実験結果と考察	45
3.7	まとめ	48
4	画像情報と抽象モデルを用いた3次元形状モデルの生成と記述	49
4.1	はじめに	49
4.2	概要	51
4.3	対話的処理による3次元形状モデルの生成と記述	53
4.3.1	前処理	53
4.3.2	視体積交差法	53
4.3.3	立面図からの3次元形状モデルの生成	54
4.3.4	対話処理による3次元形状モデルの補正と記述	58
4.3.5	実験結果と考察	62
4.4	実画像からの3次元形状モデルの生成と記述	64
4.4.1	実画像から正投影図の生成	65
4.4.2	実験結果と考察	69
4.5	まとめ	74
5	シーン構造記述に基づく拡張現実感システムの構築	75
5.1	はじめに	75

5.2	概要	76
5.3	書籍表示インタフェースの構築	77
5.3.1	本システムの条件設定と処理概要	77
5.3.2	キャリブレーション処理とブックスクリーン表面の3次元座標の取得	78
5.3.3	幾何補正処理	80
5.3.4	ページ認識処理	83
5.3.5	画像表示処理	84
5.4	実装と実験結果	85
5.4.1	システム環境と実装	85
5.4.2	ページ認識結果	87
5.4.3	幾何補正結果	89
5.4.4	画像表示結果	90
5.5	まとめ	91
6	結論	93
6.1	まとめ	93
6.2	今後の課題	95
	謝辞	97
	参考文献	99
	関連論文および口頭発表	107

目 次

1.1 実世界，仮想世界，利用者の相互関係	4
2.1 構造情報の記述法	13
2.2 シーングラフ	14
2.3 原画像とパノラマ画像生成結果	17
2.4 投影法によるパノラマ画像の生成	17
2.5 光線空間	18
2.6 抽象モデルの図式化例	21
2.7 抽象モデル	21
2.8 実シーンにおける画像特徴と抽象モデルの特徴	22
2.9 抽象モデルを用いた構造情報の取得と記述	23
2.10 ステレオ画像と平面図を用いた構造情報の取得と記述	24
2.11 実画像と立面図を用いた構造情報の記述	25
2.12 西本願寺の書院	27
2.13 ベゼクリク第9号窟	27
2.14 Mixed Reality の定義	28
2.15 実世界，仮想世界，利用者の相互関係	30
3.1 処理の流れ	34
3.2 3眼式ステレオビジョンシステム	35
3.3 パノラマ画像	37
3.4 実シーンから取得された視差画像	38

3.5	平面図と3次元形状モデル	39
3.6	標準ステレオモデル	40
3.7	平面図からの構造情報の取得	40
3.8	平面図から取得された視差画像	41
3.9	キャリブレーションパターン	41
3.10	構造情報の統合	43
3.11	図式化された抽象モデル(平面図)の生成	45
3.12	本手法における推定値と実測値(line 45)	46
3.13	3次元形状モデルの生成と抽象モデルの図式化結果	47
4.1	トユク1号窟の倒壊例(2001年11月撮影)	51
4.2	処理の流れ	52
4.3	立面図	54
4.4	立面図の分割	55
4.5	視体積交差法	55
4.6	3次元形状モデルの生成結果	55
4.7	キジル石窟の3次元モデルの生成結果	57
4.8	対話的な補正処理によるテクスチャマッピング	59
4.9	キジル8号窟内の前室部分の壁画(平凡社1983)	63
4.10	キジル8号窟の3次元形状モデルの生成と抽象モデルの図式化結果	64
4.11	弱透視投影	66
4.12	中心軸の抽出	67
4.13	モール仏塔	70
4.14	モール仏塔の3次元形状モデル生成と抽象モデルの図式化結果	71
4.15	西本願寺御影堂内陣の須弥壇と厨子の3次元形状モデルの生成と抽象モデルの図式化結果	72
5.1	処理の流れ	77
5.2	キャリブレーションパターン	79
5.3	本システムにおける画像座標系と世界座標系の関係	79

5.4	画像の幾何補正	81
5.5	ブックスクリーン	82
5.6	ページ認識に用いたバーコード	84
5.7	システム構成	85
5.8	写字台文庫の一例 (竹取物語抄)	86
5.9	時間に対するページ認識結果と幾何補正のタイミング	88
5.10	画像の幾何補正結果	91
5.11	本システムによる写字台文庫の表示結果	92

表 目 次

3.1	3眼カメラの仕様	35
3.2	インタフェースの仕様	36
3.3	奥行分解能	38
3.4	マルチベースラインステレオ法と本手法の精度	45
4.1	キジル8号窟の精度比較	63
4.2	中心軸の直線パラメータ	69
4.3	剛体変換・スケール変換のパラメータ	73
5.1	液晶プロジェクタ	86
5.2	ページ認識・画像表示用PC	87
5.3	構造情報の取得用PC	88
5.4	距離計測分解能	89

第1章

序論

1.1 研究背景

近年、実世界の環境を仮想世界に入力し、それを応用する研究が進んでいる。この研究は産業（建築、コンピュータ支援設計）、考古学（遺跡・遺構などの電子記録化）や教育（電子図書館、自然のシミュレーション）、娯楽（インタラクティブな3Dゲーム、映画）などの分野に適用されている。

実世界の環境を仮想世界に入力するとき問題となるのが、実世界の複雑な環境にどう対処するかということである。また、仮想世界に入力された環境を利用するとき問題となるのが、利用者の必要な情報を、仮想世界からどのようにして抽出し、また、抽出された情報を実世界に、どのようにして表現するかということである。これらの問題に対処するには、実世界、仮想世界、利用者の相互関係を整理する必要がある。

本論文では、コンピュータビジョン (Computer Vision:CV)、コンピュータグラフィックス (ComputerGraphics:CG)、複合現実感 (Mixed Reality:MR)、ヒューマンインタフェース (Human Interface: HI) の研究分野において、実世界、仮想世界、利用者の相互関係を明確することを目的とする。

CVでは、実世界の環境を仮想世界に入力するとき、実世界の3次元シーンの光学像を、2次元平面に写像されたもの(2次元画像)として扱っている。CVの目的の1つは、2次元画像から3次元シーンを取得し、仮想世界に3次元シーンを復元することである。CVにおいて、Marrが視覚の計算理論¹⁾を提案して以後、この理論に基づいて、

「見え方の科学 [Shape from X (X: shading, texture, contour, motion)]」に関する計算アルゴリズムが盛んに行われていた。しかしながら、2次元画像から3次元シーンを復元する処理は不良設定問題である。この不良設定問題を解くためには、光学モデルや統計処理を用いる際に、様々な仮定や前提条件が必要であり、複雑な実世界に対しては適用することが困難であった。

一方、CGの分野では、仮想世界に入力された3次元シーンをどのようにして表示するかという問題を扱っており、3次元シーンを2次元平面に写像する手法の研究が行われている。その一例として、フォトリアリスティックレンダリング (Photorealistic Rendering)²⁾³⁾が挙げられる。これは物理法則(光学等)に基づいて光と影を生成する技術であり、写実性の高い画像生成を目的としている。

また、2次元画像から3次元シーンを復元するCVと、3次元シーンから2次元画像への写像を行うCGを統合するような研究も進められている。CVとCGの統合は、一般的に次のようにして行われる。まず、実世界のシーンを撮影した2次元画像をCVアルゴリズムによって解析し、3次元シーンを仮想世界に復元する。次に、復元された3次元シーンに対して編集・加工(視点変化や時間的な変化など)を行う。それから、編集・加工された3次元シーンをCGアルゴリズムを用いて、新たな2次元画像を生成する。CGとCVの統合化研究として、Virtualized Reality⁴⁾⁵⁾がある。Virtualized Reality⁴⁾⁵⁾は、実世界から入力された時系列画像をもとにして、仮想世界に3次元シーンを復元し、視点の変化に対応した画像(視点画像)を生成したり、ドーム内に設置された多数のカメラを利用して撮影された画像から、あらゆる点のボクセルデータを計算し、時間的な変化を伴った視点画像を生成するものである。ここから得られる生成画像そのものは仮想ではなく、実世界と同等の情報をもつものである。

実世界と仮想世界の相互関係は、Paul Milgramによって提案された複合現実感(MixedReality)⁶⁾の中で明示されている。これに対し、実世界を電子的に増強・拡張するAugmented Reality(AR)の方が、既によく知られた概念であった。ARに対置する概念として、計算機内に構築される仮想世界を実世界の情報で強化することをAugmented Virtuality(AV)と呼び、ARとAVを統合し包含する概念を複合現実感と定義している。複合現実感の概念を用いてシステムを実現するために、CVとCGの基礎技術が数多く導入されている。

利用者と実世界の相互関係は、次のように考えることができる。利用者は、振舞いや行為によって、なんらかの作用を実世界に与えている。また、実世界は、視覚的あるいは触覚的な刺激などを利用者を与えている。

利用者と仮想世界の相互関係は、利用者と仮想世界の“橋渡し”の立場からインタフェースについて考察する。現在では、インタフェースの入力部分に関しては、キーボード、マウス、タブレットなどに加え、タッチパネルやデータグローブなどの入力デバイスが一般的によく用いられている。これらは、計算機側の視点からみた入力方法である。一方、人間側の視点から見た場合、ヒューマンインタフェースと呼ばれている。人間側の視点から見た入力方法については、文字・音声といった言語的な情報の入力や、人間の動作や表情といった非言語的(ノンバーバル)な情報を利用する研究が盛んに行われている。

以上のことから実世界、仮想世界、利用者の相互関係は、図 1.1 に示すようにまとめることができる。

1.2 研究目的

本論文では、図 1.1 の関係に従い、実世界の環境を仮想世界に入力する方法に関しての提案を行い、さらに実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムを構築する。

まず、実世界の環境を仮想世界に入力する方法について提案する。本論文では、実世界の環境を実シーンの構造情報として捉え、実シーンから構造情報を取得・記述することにより、仮想世界に3次元形状モデルを生成し、表現することを目的とする。

実シーンの構造取得については、実シーンの構造情報が顕著に現れている場合には、実シーンを撮影した画像から得られる画像特徴(特徴点やエッジ情報、テクスチャ情報)から構造情報を取得し、仮想世界に3次元形状モデルを入力することができる。しかしながら、実シーンの構造情報が顕著に現れていない場合には、実シーンを撮影した画像から得られる特徴が少なく、画像情報だけでは、仮想世界に3次元形状モデルを入力することが難しくなる。この場合には、シーンに対して、構造化パターン光などを照射し、シーンの構造情報を明確にさせる必要がある。もしくはシーンの構造情報

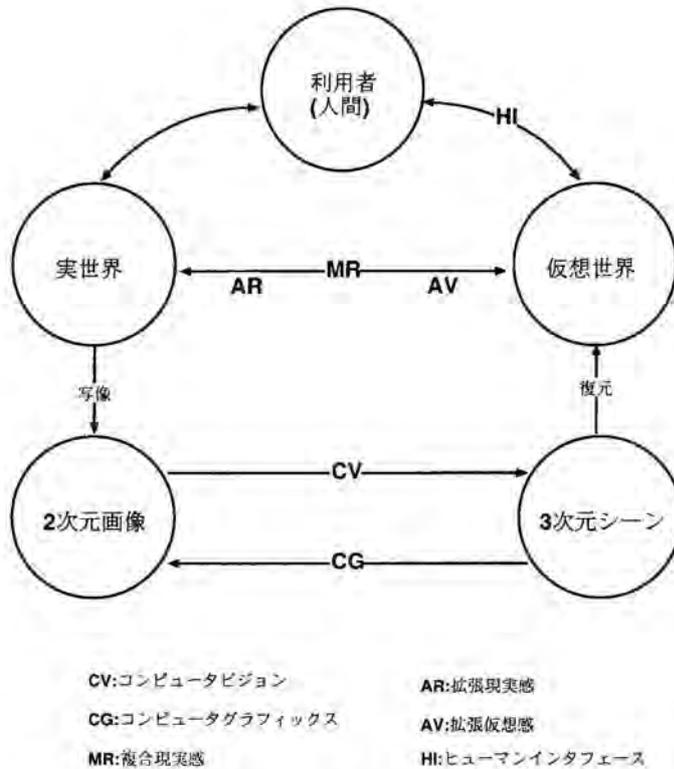


図 1.1 実世界、仮想世界、利用者の相互関係

を取得する際の手がかりとなる仮定、事前知識、モデルなどを導入する必要がある。

一方、実シーンの構造記述については、従来の体積記述法や領域記述法では複雑なシーンの構造情報に対応することが難しい。また、シーングラフのような論理構造記述では、シーンの認識や編集が可能となるが、物体認識などの高次処理が必要となる。また従来、実シーンの構造情報の取得と記述には、それぞれ独立した手法が用いられており、統合的に扱われることが少なかった。

そこで、本論文では、シーンの構造情報が明確ではない対象に対して、構造情報の取得と記述を行うことを目的とする。このため、シーンの構造情報を取得する際の手がかりとなる抽象モデルを導入する。本手法における抽象モデルとは、「実シーン(または対象物体)から抽象化された構造情報が記述されているもの」と定義し、対象に関し

での構造情報(相対的な位置関係、形状情報)が保たれているものとする。例えば、抽象モデルを図式化したものに、スケッチ、鳥瞰図、立面図、平面図などがある。本手法では、抽象モデルを導入することにより、

- 複雑な実シーンに対して、構造情報の取得と記述が可能となる。
- 構造情報の取得と記述が統合的に扱える。

などの特徴が挙げられる。実験では、屋内・屋外を対象に抽象モデルを図式化したもの(平面図・立面図)を用いて、実シーンの構造記述と3次元形状モデルの生成が可能であることを示す。

さらに、実世界、仮想世界、利用者の関係に基づいた拡張現実感システムの構築においては、利用者が仮想世界の情報を操作することにより、仮想世界から必要な情報を実世界に抽出することが可能であることを示す。利用者が仮想世界の情報を操作する方法には、すでに述べたように、

- マウスやキーボードなどの入力デバイスを使う方法
- 「利用者(人間)」の動作や表情などを情報として入力する方法

があった。ここでは、人間側の視点から見た入力方法を採用し、利用者が、本のページをめくることによって、仮想世界の情報を操作することを目的している。実験では、書籍画像、プロジェクタ、投影スクリーンとして和紙で作られた本を用いた拡張現実感システムを構築しシステムの有効性を示す。

1.3 論文の構成

本論文の構成は以下の通りである。

第2章は、実シーンから構造情報を取得・記述・表現するための一般的な手法について述べ、抽象モデルを用いた構造情報の取得と記述方法を提案する。また、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムの構築方法について述べる。

第3章では、屋内シーンから構造情報を取得・記述する方法について述べる。ここでは、構造情報を取得する際の手がかりとなる抽象モデルを図式化した平面図を用い

る。平面図では、部屋全体の形状を捉えることが可能であるが、室内にある机や計算機など物体の構造情報を得ることができない。そこで、ステレオ法を用いて、室内にある複数の物体に対しての構造情報を取得する。ただし、ステレオ法では、構造情報の取得が難しい壁に関しては平面図から得られる構造情報を用いる。次に、これらの構造情報を統合することによって室内全体の構造情報を取得し、安定した3次元形状モデルの生成が可能であることを示す。さらに、生成された3次元形状モデルを用いて抽象モデルを更新することが可能であることを示す。

第4章では、屋外シーンに対する構造情報を取得・記述する方法について述べる。ここでは、構造情報を取得する際の手がかりとなる抽象モデルを図式化した手書き立面図を用いる。次に、手書き立面図の各面を1つの領域として扱い、視体積交差法を用いて手書き立面図から3次元形状モデルを生成する。しかしながら生成した3次元形状モデルは、実世界の対象と一致するとは限らないという問題点がある。この問題に対して、ここでは手書き立面図から生成した3次元形状モデルを2次元画像に投影し、実画像と比較しながら、対話的処理を利用して、3次元形状モデルを補正する方法について述べる。さらに、補正した3次元形状モデルを用いて抽象モデルを更新することが可能であることを示す。

第5章では、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムの構築方法について述べる。このシステムは、キャリブレーション処理、幾何補正処理、ページ認識処理、画像表示処理からなる。キャリブレーション処理ではプロジェクタ座標系、画像座標系、世界座標系のキャリブレーションを行う。幾何補正処理では投影面(ブックスクリーン)が非平面であるため、投影する画像を補正する必要がある。そのため、マルチベースラインステレオ法を用いてブックスクリーン表面の構造情報を求め、投影する画像の補正に必要なパラメータを取得する。ページ認識処理では、バーコードを利用して、ブックスクリーン上のページ番号を認識する。画像表示処理では、ページ認識結果からページに対応する画像を選択し、構造情報をもとに補正画像を生成する。そして、補正画像をブックスクリーンに投影する。さらに、この拡張現実感システムを用いて実験を行うことにより、本手法の有効性を示す。

第6章では、本論文において提案した各手法、それらの実験結果をまとめ、今後の課題について述べる。

第2章

実シーンからの構造情報の取得・記述・表現

本章では、CV、CGの研究分野における実世界の構造情報を取得・記述・表現する従来研究について述べ、本研究で提案する構造情報の取得・記述方法について述べる。次に、MR、HIの研究分野における実世界、仮想世界、利用者の相互関係を考察する。そして、これらの相互関係に基づき拡張現実感システムの構築方法を提案する。

2.1 はじめに

実世界の環境を仮想世界に入力し、それを応用する研究が盛んに行われている。実世界の環境は複雑であるため、一般的に実世界のあらゆる環境を仮想世界に入力することは、環境の取得方法、計算量、記憶容量などの点において難しいと言える。また、仮想世界に入力された情報から利用者が必要とする情報を抽出し、実世界にどのようにして表現するのかという問題もある。このような問題に対処するには、実世界、仮想世界、利用者の相互関係を考察する必要がある。

本研究の目的は、実世界から環境(ここでは構造情報と呼ぶ)を取得・記述し、仮想世界に3次元形状モデルを生成し表現することである。また、取得した構造情報を実世界に表現する拡張現実感システムを構築することである。本研究における構造情報とは、実世界の3次元情景(ここでは実シーンと呼ぶ)の幾何学的な構造であり、実シー

ンに関しての形状情報や位置情報を表したものである。また、仮想世界とは計算機内で構築される空間のことである。

2.2 シーン構造情報の取得

ここでは、実シーンから構造情報を取得する方法について述べる。実シーンの構造情報を取得するには、一般的に画像特徴や構造化光を利用する方法がある。

(1) 画像特徴の利用

実世界に存在する3次元シーンの光学像が、2次元平面に写像されたもの(2次元画像)として扱い、直接2次元画像から構造情報を取得するものである。これは、シーンの構造情報が撮影した画像情報から明確に取得できる場合に有効である。

(2) 構造化光 (Structured Light) の利用

直接2次元画像から構造情報を取得するのではなく、撮影装置、照明位置、方向、配置などを制御し、シーンの構造情報が明確に取得できるようにシーンに対して、構造化光を投射するものである。これは画像特徴を利用して構造情報を取得できない場合に有効である。

以下では上記に挙げた2つの手法の特徴について述べる。

2.2.1 画像特徴を用いたシーン構造情報の取得

実世界から時系列画像やステレオ画像を取得し、それらの画像から特徴を抽出し、シーン構造情報を取得する方法がある。以下に、これらの手法の概要と特徴について述べる。

[1] 時系列画像の利用

時系列画像を用いてシーン構造情報を取得する方法は、Shape from Motion と呼ばれ、その代表的な手法に、時系列画像中の特徴点を追跡し、因子分解 (Factorization)⁸⁾

を利用してシーンの構造情報とカメラ運動を取得する方法がある。これは以下のような特徴をもつ。

1. 運動や形状に対する知識や条件が不要
2. 数値計算的に解が安定
3. 復元精度が良い

因子分解法による構造情報の取得方法は、安定な構造情報を取得するために、カメラの投影モデルとして正射影モデルを仮定していた。しかしながら、正射影モデルは、シーン中の物体が遠くにあり、かつ奥行きがほぼ一定の場合には、精度は良いが、逆にシーン中の物体が近くにあり、かつ奥行きが大きく変化する場合には、正射影モデルは成り立たないという欠点がある。そこで、より精度の高い射影近似モデルを用いた構造情報の取得方法が提案された⁹⁾。また、上述の手法は、すべての特徴点追跡を終えた後バッチ処理で行われるが、実時間処理を目的とした研究が行われている¹⁰⁾。

また、これまでシーン中の物体の数は1個と仮定してきた。しかしながら、複数の対象物体が、画像中に含まれ、それぞれ異なる運動をしている場合でも、事前に物体の個数を知らなくても、個々の物体を分離し同時にそれぞれのカメラ運動と物体の構造情報を取得できる手法が提案されている¹¹⁾。また、建物の外観などの構造情報を取得するには点特徴だけでなく線分を特徴として利用することにより、歪みの少ない形状が取得できるため、線特徴に拡張した手法も提案されている¹²⁾。

[2] ステレオ画像の利用

ステレオ視の基本原理は三角測量である。異なる視点から同一の対象を観察し、それぞれの画像上への投影位置の違いからシーンの構造情報を取得するものであり、Shape from Stereoと呼ばれる。

ステレオ法では画像間において対応する点や特徴を求める必要がある。画像間で対応関係を求める方法は、以下に示すように領域ベース法と特徴ベース法の2つがある。

- 領域ベース法

領域ベース法^{15), 16), 17), 13), 14)}では、一方の画像の点の対応を他方の画像から探索

する場合に、その点の周りの局所的な濃度パターンを手がかりに探索するものである。具体的には、一方の画像から注目点の周りに小領域（ウィンドウ）を抽出し、それをテンプレートとして他方の画像のエピポーラ線上のあらかじめ設定された探索範囲内で対応づけを行う方法である。領域ベース法で問題となるのが、最適なウィンドウサイズの選択にある。ウィンドウサイズが小さい場合には、対応する領域が多く、また、ウィンドウサイズが大きい場合には分解能の低下や計算コストが増える¹⁶⁾。この問題を改善するために、対象物体の特徴に基づいてウィンドウの最適形状を求める手法も提案されている¹⁷⁾。

- 特徴ベース法

特徴ベース法^{18), 19), 20), 21), 22)}では、両方の画像からそれぞれ濃淡エッジ等の特徴を抽出した後、両画像間において対応関係を求める方法である。一般的にはエッジを画素単位で対応づけることが多いが、さらにエッジ点の集合をエッジセグメントととらえ、エッジセグメント同士の対応づけを行う手法も提案されている^{21), 22)}。

領域ベース法、特徴ベース法、いずれの手法においても、画像上のすべての対応を正確に求めることは、非常に困難である。なぜなら、画像間の対応付けには本質的にあいまいさが存在するからである。撮影する対象にもよるが画像中には、類似するパターンが複数存在する可能性がある。

次節では、類似するパターンが存在するシーンに対しても、有効的に構造情報を取得することができる構造光を用いた方法について述べる。

2.2.2 構造化光を用いたシーン構造情報の取得

ここでは、シーンに対して構造化光（スリット光やグレイコードパターンなど）を投射することにより、シーンの構造情報を取得する方法について述べる。

[1] スリット光

対象物体にスリット光を投影して、物体の構造情報を取得する手法は、光切断法と呼ばれている。この手法では、レーザ光源を半円筒レンズにより扇状に拡大したスリッ

ト光を投射し、カメラでその像を撮影するものである²⁴⁾。次に、撮影されたスリット光の検出を行い、その位置(画像座標値)を、あらかじめキャリブレーションしたカメラパラメータ、スリット光パラメータを用いて、三角測量の原理によって3次元座標値を求めるものである。

[2] 空間パターン光

空間パターン光では、投影回数において、上述の光切断法を用いる方法に比べて大幅に少ない。 n 本のスリット光に対応する分解能が、 $\log_2 n$ 回のパターン光投影で得られる。例えば128本のスリット画像と同じ分解能を7枚のパターン光投影で実現が可能である。

空間パターン光を用いる代表的な手法の1つに、グレイコードパターン投影法がある⁷¹⁾。これは、白と黒とを交互に配置したパターンを複数回にわたって徐々に細分化して投影することによって輝度を2進数としてとらえ、スリット光同士を判別する方法である。

[3] 強度変調パターン光

強度変調パターン光の代表的な手法として、正弦波格子位相シフト法²⁶⁾がある。これは、正弦波状の光強度パターンを時系列で位相変化させながら対象物体に投射するものである。それから、観測される時系列の光強度パターンの変化からパターンの位相値を求め、この位相値を用いて空間をコード化する。この手法は、原理的に最低3枚の画像を用いることによって画素ごとに物体の構造情報の取得が可能であり、グレイコードパターンを用いる手法よりも効率的である。

2.3 シーン構造情報の記述と表現

ここまでは、実シーンから構造情報を取得する方法について述べてきた。ここでは、CVとCGの研究分野において、入力された構造情報を仮想世界にどのように記述し表現するのかという問題について考察する。

2.3.1 構造情報の記述

CVにおいては、すでに述べたように、実シーンの光学像が、2次元画像に投影されたものとして扱い、実シーンの構造情報を取得し、仮想世界内に入力することを目的としていた。ここでは、仮想世界に入力された構造情報をどのようにして認識・理解するのかという問題に対して、CVの研究分野から考察する。CVの研究では、仮想世界に入力された構造情報を認識・理解するために、その処理過程で、構造情報の記述を行う必要がある。構造情報を記述する方法として、代表的なものに、領域記述法、体積記述法、論理構造記述法の3つがある。以下に、それぞれの手法の特徴について述べる。

[1] 領域記述法

領域記述法は、シーン中の3次元物体を見かけの2次元形状の集合として記述する方法であり、例えば、図2.1(a)に示す物体を図2.1(b)のように

小円 + 小円筒面 + 大円の一部 + 大円筒面

として記述する。この領域記述法は、画像中より対応する情報が直接得られるので、照合に関しては容易である。しかしながら、3次元物体は見る方向によって形状が変わるので、視点変化にどのように対応していくのかが問題となる。

領域記述法には、面関係グラフ法²⁷⁾やそれを改良したアスペクト法²⁸⁾、拡張ガウス像法²⁹⁾などがある。

[2] 体積記述法

体積記述法は3次元物体の中に座標系があると考え、この座標系をもとにして物体を記述する方法である。例えば、図2.1(a)の物体では図2.1(c)のような座標系を考え、 z 軸に沿っての半径 r の変化でこれを記述する。この方法では、1つの3次元物体に対して、1つのパラメータで記述することができる。しかしながら、画像中より物体固有の座標系をどのようにして得るかが問題となる。

体積記述法の最も代表的なものとして一般化円筒法³⁰⁾がある。一般化円筒法では物体中に1本の軸を仮定する。この軸の空間中における変化を1つのパラメータとする。

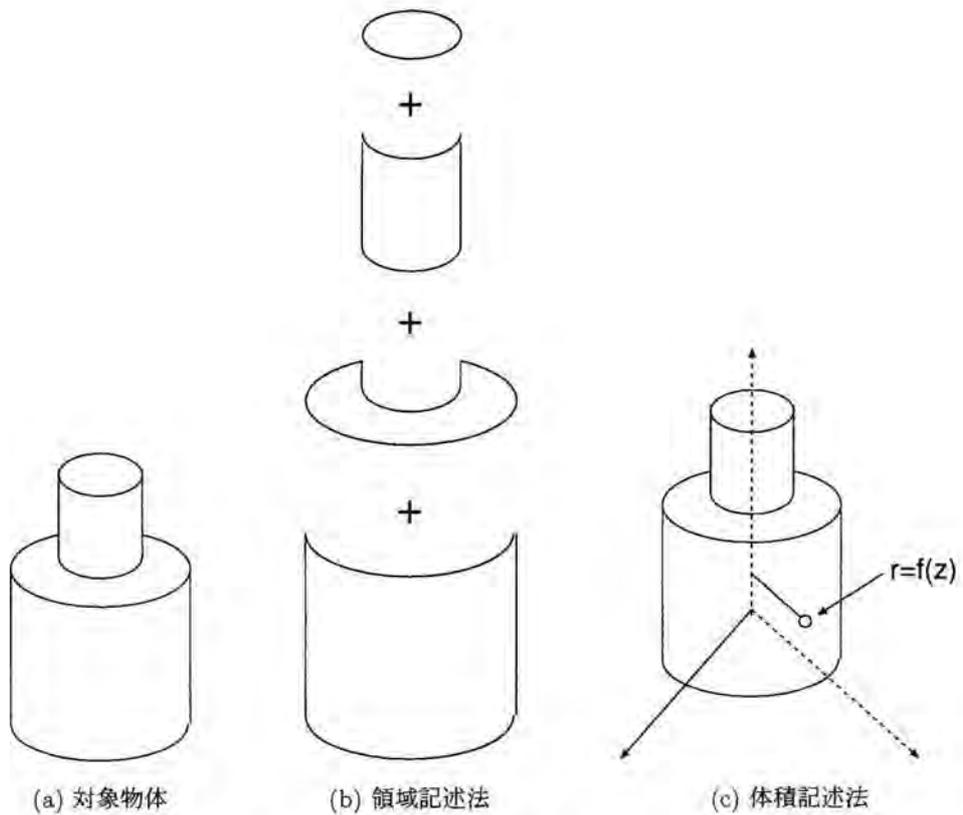


図 2.1 構造情報の記述法

一方、この軸の回りにおける物体の太さの変化をもう1つのパラメータとする。基本的には、この2つのパラメータで物体形状を記述するものである。

実シーンに対しては、レーザ・レンジファイダで得られる距離画像から一般化円筒を抽出する試みも行われている³¹⁾。

[3] 論理構造記述

シーンを論理構造記述するシーングラフについて述べる。シーングラフは、MPEG4システム (ISO/IEC Committee Draft 14496-1) においてサポートされている。これは、ビデオ、オーディオ規格に従って生成されるメディアオブジェクトを複数個まとめて

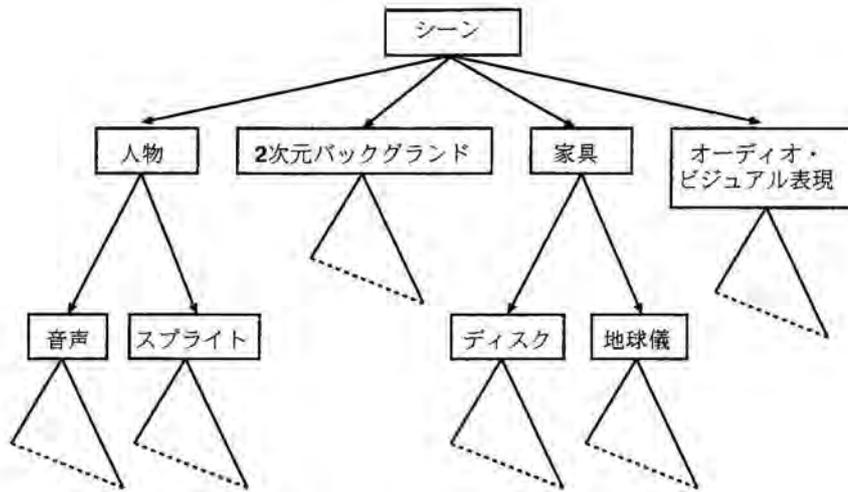


図 2.2 シーングラフ

空間的・時間的に配置するものである。そのために用いられるシーン記述フォーマットは、BIFS(Binary Format for Scene description)と呼ばれ、VRML(Virtual Reality Modeling Language)をもとにしている。

VRMLは、HTML(Hyper Text Markup Language)を3次元グラフィックスが扱えるように拡張したものとも言える。シーングラフの基本概念を図2.2に示す。シーングラフでは、ノードと呼ばれる構成単位を階層的に組み合わせることで1つのシーンを構成する。ノードとは、光源、形状、材質、色、座標など、シーンを構成する構成単位であり、さらにそれが子供のノードを持つことでシーン記述の階層化を実現している。シーングラフを用いてシーン記述を行う利点は、シーンの探索や編集(ノードの追加、削除)などが容易にできる点である。しかしながら、実シーンの構造情報をシーングラフで記述するためには次のような問題点がある。

- 実シーンから構造情報を精度良く取得する必要がある。
- 取得された構造情報から物体を認識する必要がある。

2.3.2 構造情報の生成と表現

ここでは、CGの研究分野において、実シーンから取得された構造情報を仮想世界内にどのようにして生成し表現するのかという問題について考察する。

実シーンから取得された構造情報を生成し表現する方法として、

(1) Model Based Rendering

(2) Image Based Rendering

の2つがある。

Model Based Rendering は、実シーンから取得された構造情報をもとにして、仮想世界の中で3次元形状モデルを生成し表現するものである。

一方、Image Based Rendering は、実シーンを撮影した2次元画像をもとにして、仮想世界の中で新たな2次元画像を生成し表現するものである。

以下に、Model Based Rendering 手法と Image Based Rendering 手法の特徴について述べる。

(1) Model Based Rendering

3次元形状モデルを生成し表現するには、形状表現の選択(ワイヤフレーム、サーフェスモデルなど)、隠面処理(Zソート法、Zバッファ法など)、陰影処理(レイトレーシング、ラジオシティ法など)が必要となる。しかしながら、仮想世界の中に、写実性の高い3次元形状モデルを生成し表現するには、厳密な物理法則を利用したり、膨大な計算量が必要となる。このような問題を回避するための手法として、テクスチャマッピング法やバンプマッピング法がある。この2つの手法は、計算量を抑え、厳密な物理法則を適用することなしに写実性の高い3次元形状モデルを表現することが可能である。以下に、この2つの手法の特徴について述べる。

[テクスチャマッピング法]

テクスチャマッピング法は、3次元形状モデルの形状表面に2次元画像を貼りつける方法である。この画像はテクスチャと呼ばれ、この方法によって、計算量を少なくし、写実性の高い3次元形状モデルを表現することが可能である。

[バンプマッピング法]

物体表面の凹凸を擬似的に表現する方法として、バンプマッピング法がある。これは、テクスチャマッピングがスカラマッピングであるのに対し、物体表面の法線ベクトルをマッピングし、物体上の小さな凹凸を表現するものである。バンプマッピング法は、小さな凹凸を実際にモデリングするのに比べ、モデルのデータ量を削減でき、かつほぼ同じ効果を得ることができる利点がある。なお、法線ベクトルは乱数、連続関数、または濃淡画像の濃淡値を微分したものなどを利用する。

(2) Image Based Rendering

実シーンを撮影した2次元画像を用いて、シーンを表現する方法にパノラマ画像を利用したものや、光線空間の原理を用いたものがある。パノラマ画像を利用したものは、広角なシーンを表現することができる。また、光線空間を利用したものは、利用者の視点変化に対応したシーンの表現が可能である。以下に、その2つの特徴について述べる。

[パノラマ画像の利用]

利用者の視野角は限られているため、実世界の広範囲なシーンを同時に把握することができないという問題点がある。これを改善する方法の1つがパノラマ画像を生成することである。パノラマ画像は、実シーンを撮影した複数の画像をつなぎ合わせ、1枚の広範囲な画像(パノラマ画像)を生成する。このため、生成したパノラマ画像を表現することによって、利用者はシーンの全体像を同時に把握することが可能である。

パノラマ画像の生成には様々な手法が提案がされており、大別すると静止画像を入力とする方法³²⁾と時系列画像を入力とする方法³³⁾がある。

(1) 静止画像を用いたパノラマ画像の生成

静止画像を入力とした場合には、撮影する画像の枚数は、少なくても済み、高解像度のパノラマ画像の生成が可能となる。しかしながら、画像間の歪みが問題となり、画像をつなぎ合わせる前に入力画像に対してなんらかの変換処理が必要となる。撮影されたシーンを平面仮定し、平面射影変換を用いることにより、画像間の歪みを改善させる手法が提案されている³²⁾。



図 2.3 原画像とパノラマ画像生成結果



図 2.4 投影法によるパノラマ画像の生成

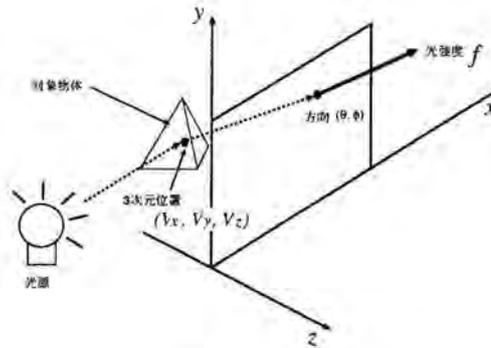


図 2.5 光線空間

図 2.3 に異なる視点から撮影された画像を示し、平面射影変換を用いてこれらの画像をつなぎ合わせた結果を図 2.3(f) に示す。

(2) 時系列画像を用いたパノラマ画像の生成

時系列画像を用いたパノラマ画像の生成手法は、カメラワークによって生じる動き情報を抽出して、画像の合成を行うものである。投影法を用いたパノラマ画像の生成³³⁾では、カメラワークをフレーム間の変化としてとらえ、そのフレームの変化量に基づいて画像の合成を行っている。そのため、カメラの撮影条件（レンズの焦点距離、カメラの視点位置など）は、考慮に入れる必要がない。

図 2.4 に、投影法を用いて、パノラマ画像を生成した結果を示す。

[光線空間の利用]

実シーンにおける対象物体の見え方を空間を伝搬する光線の形で記述し、それをもとにして別の視点からの見え方を生成する手法が提案されている²³⁾³⁴⁾。この手法によって記述される空間は、光線情報自体を多次元空間に記述することから光線空間と呼ばれる。

図 2.5 に示すように、空間中の点 (V_x, V_y, V_z) を時刻 t に (θ, ϕ) の方向に通過する波長 λ の光の強度 f は、

$$f(\theta, \phi, \lambda, V_x, V_y, V_z, t) \quad (2.1)$$

のような関数として表わすことができる。この関数の値を一般的に求めることができれば、利用者の視点位置に対応した画像を生成することができる。しかしながら、実際に7次元の光線情報を記述することは困難である。そのため、一般的にはすべての波長 λ に対して値を保持するのではなく、RGB値のみを持ち時間 t を省略して静止画として扱う。このことによって $(\theta, \phi, V_x, V_y, V_z)$ の5次元の光線空間として記述されることが多い²³⁾。

上記に述べた光線空間の原理を用いて、シーンを表現する方法にQuickTimeVR³⁵⁾、Lumigraph³⁶⁾、Light Field³⁷⁾、同心モザイク³⁸⁾などがある。QuickTimeVRでは、カメラを回転して画像を撮影するために視点の平行移動ができない。しかしながら、LumigraphやLight Field、同心モザイクではある平面の2次元配列にカメラを密に配置し、それぞれのカメラ（光軸が平面と直交）から画像を取得する。このため、同じ平面内の位置なら視点の平行移動が可能となる。

2.4 抽象モデルを用いたシーン構造情報の取得・記述・表現

実シーンから構造情報の取得方法においては、画像特徴を利用する手法と構造化光パターンを利用する方法があった。画像特徴を利用する手法では、撮影を簡易に行うことができるが、画像特徴が明確に取得できない実シーンに対しては、安定した構造情報を取得することが難しいと言える。一方、構造化光パターンを利用する手法は、実シーンから安定した構造情報を取得することが可能であるが、対象物体の大きさや、可搬性、撮影環境などの点に制限があった。

シーンの構造記述方法においては、領域記述法、体積記述法、論理記述法などがあった。領域記述法では視点変化に対応することが難しく、体積記述法では、物体固有の座標系の設定が難しいといった問題があった。また、論理構造記述では、実シーンから取得された構造情報に対して高次の物体認識処理が必要であった。

さらに、実シーンの構造情報の取得と記述はそれぞれ独立した手法であり、統合的に扱われることが少なかった。

そこで本研究では、実シーンの構造情報を取得するための手がかりとなる「抽象モデル」を導入し、実シーンから得られる構造情報や画像情報と抽象モデルから得られ

る構造情報を統合することによって、実シーンの構造情報の取得と記述を行う。抽象モデルを導入することで実シーンから安定した構造情報を取得することが可能となる。以下にその方法について述べる。

2.4.1 抽象モデル

人間が実シーンの構造情報を理解することができるのは、網膜から視覚情報を取り入れる機能、得られた視覚情報から特徴を抽出する機能、様々な経験や知識を用いて解釈する機能のような視覚機能があるからである。さらに、人間は理解したシーンの構造情報を図面、記号、文字、音声等に変換し、コミュニケーションの手段として利用することができる。そこで、本研究では、シーン構造情報の手がかりとして、図面に着目し、実シーンから構造情報を取得する手がかりとして、「抽象モデル」を導入し、構造情報を取得する手法を提案する。

本研究における抽象モデル (Abstract Model) とは、人間がシーンの構造情報を理解した内容が記述されたのもであり人間のもつ内部表現であるとする。モデル内部は、「実シーンを抽象化した構造情報が記述されているもの」と定義し、正確な構造情報を持たないが、対象に関しての相対的な位置関係が保たれているものとする。ここで抽象化とは、対象のもつ特性を調べ、特定の性質だけを抽出し他の特性は棄却し、これによって対象のもつ複雑さを軽減することである³⁹⁾。また、抽象モデルを図式化したものが図2.6に示す立面図、平面図のようなものを指し、抽象モデル、図面と実シーンの関係を図2.7に示す。

図2.6に示したような図面は、人工的に意味をもたせた画像であり、図面からは、それを書いた人間の意図を読みとることが可能である。すなわち、図面を利用することによって、抽象モデル内部に記述されている構造情報(抽象モデルに記述された構造情報の一部分(図2.7中の斜線部分))を抽出できると考えられる。しかしながら、このような抽象モデルを用いて、シーンを生成することが難しい。この理由は、抽象モデルを図式化した図面から得られる構造情報は、実シーンの構造情報の一部分に過ぎないからである。例えば、図面として平面図を与えた場合、部屋全体の構造情報を得ることができるが、部屋にある物体の構造情報が取得できない。本研究では、実

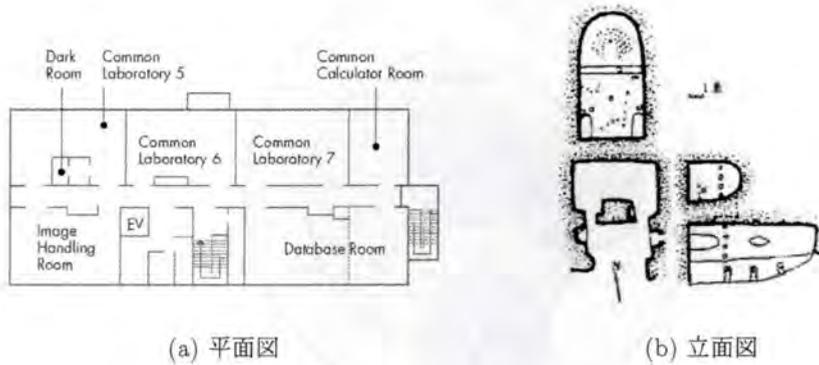


図 2.6 抽象モデルの図式化例

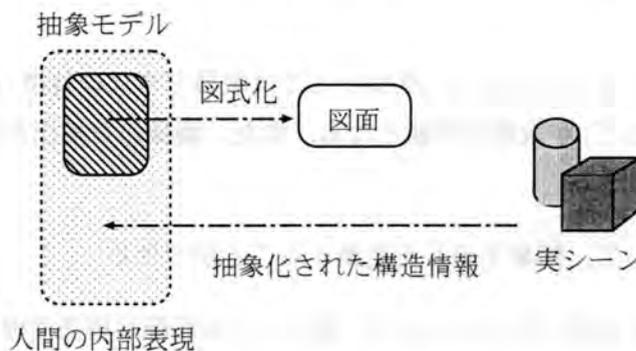


図 2.7 抽象モデル

シーンにおいて、画像特徴と抽象モデルの特徴(図面を利用して得られる特徴)は、図 2.8 のように表せるものとする。図 2.8 では実シーンの中に、画像特徴と抽象モデルが含まれている。また、抽象モデルは画像特徴の内外に複数存在するものである。この理由は、画像から取得される特徴(エッジ特徴、点特徴、面特徴など)は実シーン全体を考えた場合、その一部分にしか過ぎないからであり、抽象モデルは画像特徴以外の情報(例えば、相対的な位置関係、接続関係)も含んでいるからである。

以上のことから、画像特徴と抽象モデルの特徴の相互関係を用いることによって、以下のようなことが実現可能であると考えられる。

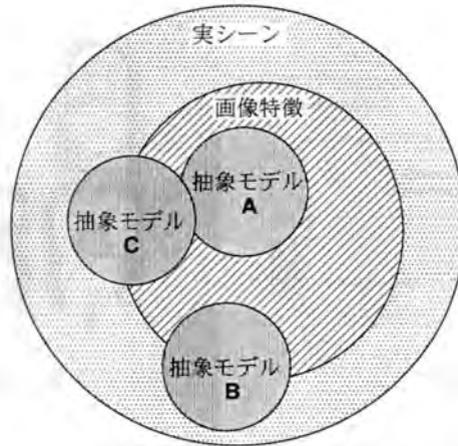


図 2.8 実シーンにおける画像特徴と抽象モデルの特徴

- (1) 抽象モデルを用いることにより、画像からでは取得できない特徴(相対的な位置関係や接続関係など)が取得が可能となる。また、画像から抽出が難しい特徴も取得することができる。
- (2) 画像特徴を利用して、抽象モデルを更新することができる。
- (3) 複数の抽象モデルを用いることにより、実シーンの詳細な構造情報を取得することができる。

本研究では上記(1)(2)に関して、図2.6に示したような抽象モデルを図式化したものから得られる構造情報と、画像から得られる構造情報を併用することにより、実シーンから構造情報が可能であることを示す。さらに、取得した構造情報を用いて抽象モデルを更新する方法について述べる。

2.4.2 構造情報の取得と記述

実シーンを撮影した画像からは、シーン(対象物体)に関する構造情報を取得することができるが、シーン(対象物体)に関する構造情報(相対的な位置関係)を取得することは難しいと言える。また、抽象モデルを図式化した図面では、構造情報(相対的な位

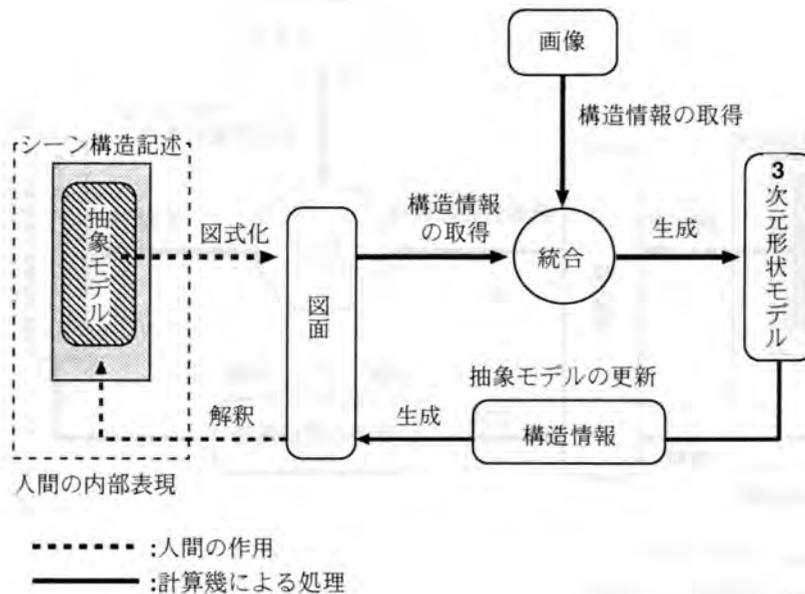


図 2.9 抽象モデルを用いた構造情報の取得と記述

置関係)を取得することができるが、シーン(対象物体)に関する詳細な構造情報を取得することは難しい。

そこで、本研究では、図 2.9 に示すように、上述した抽象モデルを用いて、実シーンから構造情報の取得と記述し、抽象モデルの更新を行う。まず、抽象モデルを図式化した図面を入力とする。次に、画像から取得される構造情報と図面から取得される構造情報を統合することにより、実シーンの構造情報を取得する。そして、取得された構造情報をもとにして、3次元形状モデルを生成する。このとき生成された3次元形状モデルのデータ構造は、以下のように記述される。

- 3次元座標点
- 3次元座標点に対応する画像の2次元座標
- 上記の座標点間の接続リスト(接続関係リスト)

この構造記述は、シーン構造情報を物理的に記述したものである。ただし、3次元形状

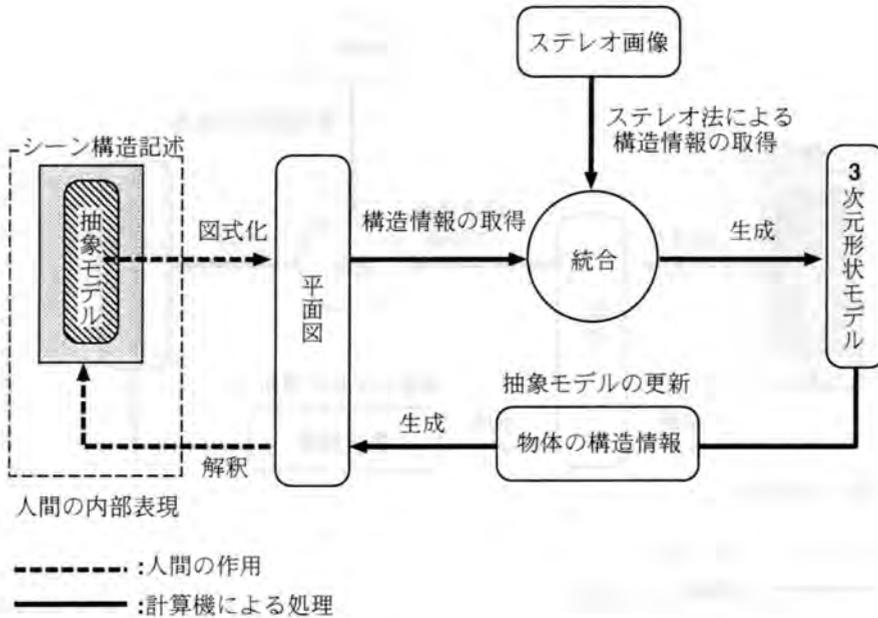


図 2.10 ステレオ画像と平面図を用いた構造情報の取得と記述

モデルの生成・表現に関しては次節で述べることにする。

さらに、抽象モデルの更新に関しては、生成された3次元形状モデルを用いて、新たに図面を生成する。生成した図面は抽象モデルを図式化したものと考えることができ、人間は生成された図面から、最初に入力とした与えた図面の情報以外の情報を読みとることができる。つまり、本研究では、人間の内部に更新された抽象モデルが形成され、シーンの構造情報が記述されたものとして考える。

具体的には、抽象モデルを図式化した平面図と立面図に対して、図 2.10 と図 4.2 に示す処理手順により、実シーンの構造情報の取得、記述、および3次元形状モデルの生成・表現を行う。ただし、入力された画像および図式化された抽象モデルの違いにより、図式化された抽象モデルを基準に考える場合と、入力された画像を基準に考える場合の二通りがあり、どちらを基準に3次元形状モデルを生成するかによって処理手順が異なる。図 2.10 は、図式化された抽象モデルを基準に考える場合の処理であり、図 4.2 は、入力される画像を基準に考える場合の処理である。

形状モデルの補正を行う。このとき、シーン構造情報の物理的な記述が行われる。それから、補正された3次元形状モデルの構造情報を用いて、新たに立面図を生成することにより、抽象モデルを更新する。これにより、シーンの構造情報を記述することができる。上記で述べた手法を用いた実験と実験結果については第4章で述べる。

2.4.3 3次元形状モデルの生成と表現

ここでは、取得された構造情報をもとに3次元形状モデルを生成し、表現する方法について述べる。

3次元形状モデルを生成・表現する方法には、先に述べたように Model Based Rendering(MBR)手法がある。MBR法において、写実性の高い3次元形状モデルを表現するためには、テクスチャマッピング法が有効である。本研究では、取得された構造情報をもとに3次元座標を求め、三角パッチによるポリゴン近似を行い、シーンの3次元形状モデルを生成する。生成した3次元形状モデルに、実画像をテクスチャとしてマッピングすることにより、3次元形状モデルを表現する。

図2.12は、西本願寺書院内の雁の間、菊の間の二間（国宝）を3次元形状モデルで表現したものである⁴⁰⁾。図2.12(a)は、抽象モデルを図式化した平面図である。まず、この平面図からコーナ特徴点を抽出し、平面図から構造情報を取得し、3次元形状モデルを生成する。さらに、実画像をテクスチャとして、生成した3次元形状モデルにマッピングしたものを図(b)に示す。

また、同様にして3次元形状モデルを生成した結果を図2.13に示す。これは、敦煌西北のトルファン地区にあるベゼクリク千仏洞の第9号窟を3次元形状モデルで表現したものである⁴¹⁾。

2.5 拡張現実感システムのためのシーン構造情報の表現

ここまでは、抽象モデルを用いた実シーンの構造情報の取得と記述方法について述べ、3次元形状モデルの生成と表現方法について述べた。

ここでは、実世界と仮想世界、仮想世界と利用者の相互関係について述べ、実世界、

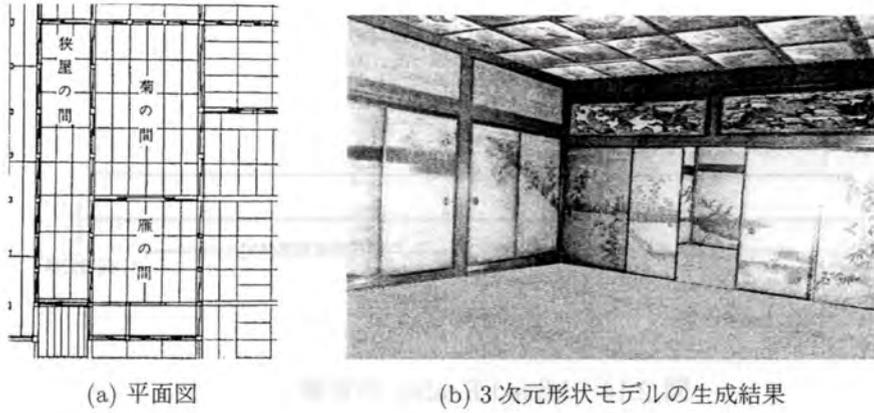


図 2.12 西本願寺の書院

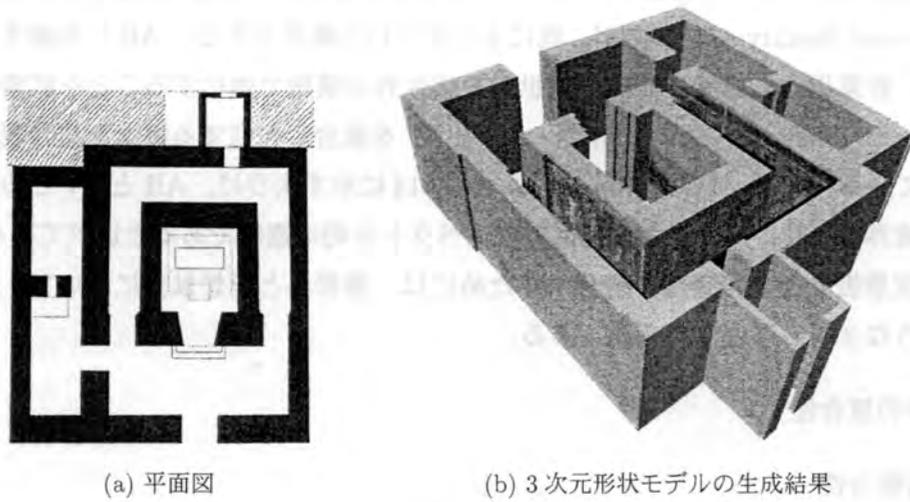


図 2.13 ベゼクリク第9号窟

仮想世界，利用者の相互関係を考察する．さらに，仮想世界に入力された情報の中から，利用者の必要な情報を抽出し，抽出された情報を実世界に表現する方法について述べる．

2.5.1 複合現実感

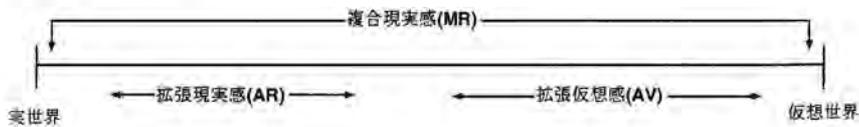


図 2.14 Mixed Reality の定義

実世界と仮想世界の相互関係は，Paul Milgram によって，提案された複合現実感 (Mixed Reality:MR)⁶⁾ の中で明示されている．実世界を電子的に増強・拡張する拡張現実感 (Augmented Reality:AR) の方が，既によく知られた概念である．AR に対置する概念として，計算機内に構築される仮想世界を実世界の情報で強化することを拡張仮想感 (Augmented Virtuality:AV) と呼び，AR と AV を統合し包含する概念を複合現実感と定義している．また，Paul Milgram は，図 2.14 に示すように，AR と AV との間には明確な境界はなく，実世界と仮想世界はスペクトル的に連続であると述べている．

実世界と仮想世界を違和感なく融合するためには，実世界と仮想世界において，以下に示すような 3 つの整合性が必要である．

- (1) 幾何学的整合性
- (2) 光学的整合性
- (3) 時間的整合性

幾何学的整合性では，利用者の視点からみて，幾何的に正しい仮想物体像を生成することと，実世界の正しい位置に位置ずれを生じないように仮想物体像を重ねること

が必要である。また、光学的整合性では実世界像と仮想物体像の画質の違いによる問題、陰影の問題、隠れの問題を解決する必要がある。さらに時間的整合性では利用者が移動したり視線を変化させたときに、実世界像の変化に対する仮想物体像の追従の際に、時間的な遅れが生じる問題である。

2.5.2 ヒューマンインタフェース

ここでは、利用者と実世界、利用者と仮想世界の相互関係について考察する。

利用者と実世界の相互関係については、利用者の振舞いや行為は、なんらかの作用を実世界に与えているものとする。また、実世界は、なんらかの刺激を利用者に与えているものとする。

次に、利用者と仮想世界の相互関係は、利用者と仮想世界の“橋渡し”の立場からインタフェースについて考察する。現在、インタフェースの入力部分に関しては、キーボード、マウス、タブレットなどに加え、タッチパネルやデータグローブなどの入力デバイスが、一般的によく用いられている。これらは、計算機側の視点から見た入力方法である。

一方、インタフェースは、人間側の視点から見た場合、ヒューマンインタフェース (Human Interface: HI) と呼ばれている。人間側の視点から見た入力方法については、文字・音声などの言語的な情報の入力や、人間の動作や表情といった非言語的 (ノンバーバル) な情報を利用する研究が盛んに行われており、そこでは、複合現実感の概念や技術が応用されている。

現在、HIの研究分野では、様々なものが開発されており、紙と電子情報を統合した研究に注目してみると、Interactive Desk⁴²⁾ や Digital Desk⁴³⁾、Enhanced Desk⁷⁾ などがある。

Interactive Desk では、1次元バーコードを用いて、実世界の紙フォルダと関連する電子メールやネット上の Web ページへのリンクを実装している。また、Digital Desk では、プリンタで印刷された Web ページとネット上の Web ページとのリンクを実現している。これらのシステムにおいて、電子情報の操作は、キーボードやマウスを用いて実現しており、計算機側の視点から見た入力方法である。

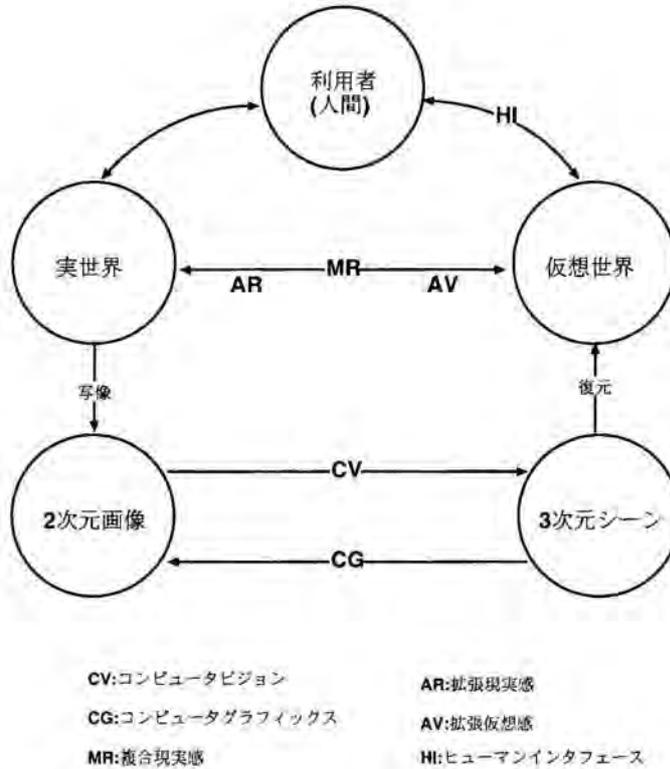


図 2.15 実世界，仮想世界，利用者の相互関係

一方，Enhanced Deskでは，書籍(教科書)に2次元バーコードを貼りつけ，それに関連した情報を机上に投影するシステムであり，電子情報の操作は，利用者の手指の動作を認識することにより実現しており，人間側の視点からみた入力方法である。

2.5.3 シーン構造情報の表現

以上のことから，実世界，仮想世界，利用者の相互関係は，図 2.15 に示すような関係にまとめることができる。ここでは，図 2.15 に示す相互関係に基づいた拡張現実感システムの構築を目的とする。ここでは，利用者がページをめくる動作をもとに仮想世界の環境を操作し，書籍画像，プロジェクタ，投影スクリーンに本を利用した拡張現

実感システムを構築する。仮想世界の情報と実世界のシーンを融合するためには、先に述べた幾何学的整合性、光学的整合性、時間的整合性を考慮する必要がある。このシステムでは、3つの整合性のうち幾何学的整合性を解決するものである。システム構築の方法と実験結果については、第5章で述べる。

2.6 まとめ

この章では、実シーンから構造情報を取得・記述・表現に関する従来研究について述べ、以下のような問題点を明らかにした。

実シーンから構造情報を取得する方法において、画像特徴を利用する手法では、画像特徴が明確に取得できない実シーンに対しては、安定した構造情報を取得することが難しく、構造化光パターンを利用する手法では、対象物体の大きさや、可搬性、撮影環境などの点に制限があった。シーンの構造記述方法において、領域記述法では視点変化に対応することが難しく、体積記述法では物体固有の座標系の設定が難しいといった問題があった。また、論理構造記述では、実シーンから取得された構造情報に対して高次の物体認識処理が必要であった。

そこで、シーンから構造情報を取得する際の手がかりとなる抽象モデルを用いた実シーンに対する構造情報の取得と記述方法について提案した。

また、CV、CG、HIの研究分野から、実世界、仮想世界、利用者の相互関係を考察した。さらに、これらの相互関係に基づいた拡張実感システムの構築法について提案した。次章以降では、本研究において提案した手法を用いて、実験を行い手法の有効性を示す。

第3章

構造情報と抽象モデルを用いた3次元形状モデルの生成と記述

本章では、実シーンから構造情報を取得するために、2種類の画像を用いる。1つは、実シーンを撮影したステレオ画像であり、もう1つは、抽象モデルを図式化した平面図である。これらの2種類の画像から取得される構造情報を統合し、室内シーンの3次元形状モデルを生成する。また、生成した3次元形状モデルを用いて、抽象モデルを更新することによって、実シーンの構造情報を記述することを目的とする。

3.1 はじめに

ここで用いる抽象モデルは、部屋全体の形状に関する構造情報が記述されているが、室内にある机や計算機などの物体に関する構造情報が記述されていないものとする。このような抽象モデルを図式化した平面図(以降、平面図と呼ぶ)を入力とするため、実シーンから物体に関する構造情報を取得する必要がある。

図3.1に全体の処理の流れを示す。図中で実線は計算機による処理であり、点線は人間の作用を表す。まず、室内にある複数の物体に対しての構造情報は、ステレオ画像からステレオ法を用いることにより取得する。しかしながら、ステレオ法では壁のように特徴がない部分については、ステレオ画像の対応点探索が難しく構造情報の取得が困難である。このため、壁に関しては平面図から得られる部屋全体の構造情報を用

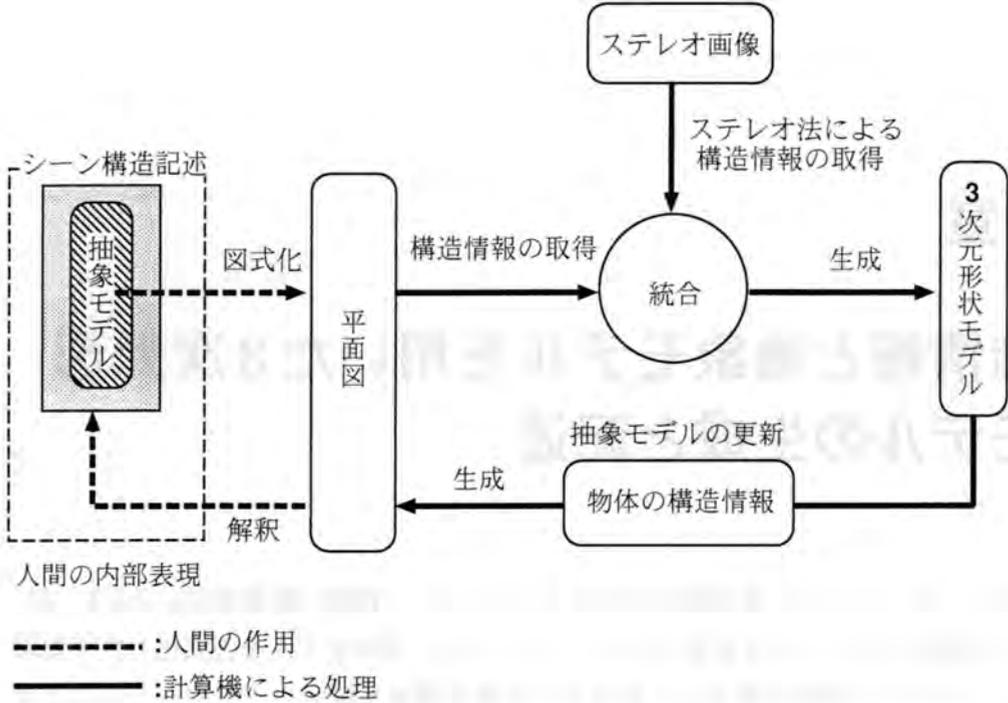


図 3.1 処理の流れ

いる。次に、これらの構造情報を統合することにより室内全体の構造情報を取得し、3次元形状モデルを生成し表現する。このとき、3次元形状モデルのデータ構造は、3次元座標点、その点に対応する画像の2次元座標点、座標点の接続リストで記述されており、これがシーン構造情報の物理的な記述である。

また、生成された3次元形状モデルを用いて、入力として与えた図式化された抽象モデル(平面図)に室内にある物体の構造情報を補うことにより平面図を生成する。生成した平面図は、抽象モデルを図式化したものと考えることができ、人間は生成した平面図から室内にある物体の構造情報を読みとることができる。つまり、人間の内部に更新された抽象モデルが形成され、シーンの構造情報が記述されたものとする。以下に、具体的な方法について述べる。

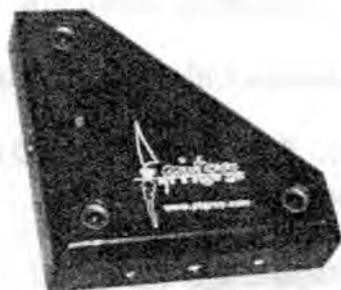


図 3.2 3眼式ステレオビジョンシステム

3.2 実シーンからの構造情報の取得

実シーンから3眼カメラを用いて、マルチベースラインステレオ法¹⁴⁾により、室内の物体に関する構造情報(ここでは、視差画像のことを指す)とテクスチャ画像を同時に取得する。

3.2.1 3眼式マルチベースラインシステムの特徴

ここで用いた3眼式のマルチベースラインシステムを図3.2に示す。また、3眼カメラモジュールの仕様を表3.1に、インターフェースの仕様を表3.2に示し、その特徴は、以下の通りである。

1. 光学歪、位置ズレは、システムとして校正済みである。

表 3.1 3眼カメラの仕様

イメージセンサー	1/3 インチ CCD
イメージセンサーの解像度	512×492
焦点距離	$f = 3.8\text{mm}$
水平画角	71.5°
ベースライン距離	100.0mm

2. リアルタイムに近いスピードで視差画像の取得が可能である.
3. 相関演算は, SAD(Sum of Absolute Differences Correlation) によるものである.
4. 物体表面にテクスチャがなく, 画像マッチングが行うことができない場合には, その領域を取り除くことが可能である.
5. 視差画像とテクスチャ画像が同時に撮影可能である.

上記の3. のSADアルゴリズムの原理は次の通りである. 左カメラで撮影した画像上での画素 (x, y) における輝度値を $I_l(x, y)$, 右カメラで撮影した画像上での画素 (x, y) における輝度値を $I_r(x, y)$ とし, 右カメラで撮影した画像上の画素 (x, y) について, 左カメラで撮影した画像上でのずれ(視差)を d とすると, SAD(Sum of Absolute Difference)の相関演算は, 式(3.1)で表される.

$$C'_{err}(x, y, d) = \sum_{i,j} |I_l(x + d + i, y + j) - I_r(x + i, y + j)| \quad (3.1)$$

SAD形式の局所相関演算法を利用すると, 各画素に対する視差は式(3.2)により計算される.

$$d(x, y) = \min_d C'_{err}(x, y, d) \quad (3.2)$$

これにより, 高速な演算が可能である.

また, ここで用いたシステムの奥行分解能を表3.3に示す. 表3.3から明らかなように, 室内のような狭い空間においては, 物体表面にテクスチャがある場合には, 高い精度で対象物体の構造情報を取得することができる.

表 3.2 インタフェースの仕様

インタフェース	Matrox Meteor フレームグラバーボード
フレームグラバーボードの解像度	640x480



図 3.3 パノラマ画像

3.2.2 全周のパノラマ画像の取得

本手法では、部屋全周の視差画像を取得するために、3眼カメラの位置を固定して、3眼カメラを回転させながら撮影する。以下に、撮影条件を述べる。

1. 室内における3眼カメラの位置は、既知とする。
2. 隣合う画像間での上下のずれが生じないように、3眼カメラを三脚に固定して、水平に回転させる。
3. 3眼カメラの回転角度を15度とし、15度ずつカメラを回転させながら24枚の視差画像と24枚のテクスチャ画像を取得する。

上記の条件によって取得された複数枚の視差画像とテクスチャ画像からそれぞれパノラマ画像を生成する。厳密には、撮影時にカメラのレンズ中心と回転軸を一致させる必要がある。しかし、ここでは、複数の画像を貼り合わせて、パノラマ画像を生成するという目的において、上記の撮影条件で十分である。

視差画像からパノラマ画像を生成する理由は、複数枚の視差画像を1枚のパノラマ画像として扱うことにより、構造情報の統合処理が簡単に行うことができるからである。また、複数枚のテクスチャ画像からパノラマ画像を生成するのは、部屋の3次元形状モデルにテクスチャとして、パノラマ画像をマッピングするためである。

パノラマ画像を生成するには、それぞれの画像間の対応づけが必要である。この画像間に対応づけには、Lucas-Kanade法⁴⁴⁾を用いる。この手法の特徴として、処理が高速で、画像間の対応付けの結果が信頼度を持つことである。図3.3にパノラマ画像を生成したものを示す。また、図3.4に複数枚の視差画像を貼り合わせて生成した部屋全周の視差画像を示す。ただし、図3.4の視差画像は、物体表面のテクスチャがないた



図 3.4 実シーンから取得された視差画像

め、画像のマッチングが行うことができなかった領域を視差値 0 で表している。図 3.4 の視差画像では、画像マッチングが行われていない領域は、全画像 (840x90pixel) 中の 24.5% であり、部屋の壁の領域に多く見られる。この画像マッチングが行われていない部屋の壁の領域を、平面図を用いることにより補う。次節では、平面図から構造情報を取得する方法について説明する。

3.3 平面図の利用

3.3.1 平面図から構造情報の取得

平面図から構造情報を取得するために、図 3.5 (a) に示す平面図のコーナーと平面図に記載されているスケールをもとにして、実世界と平面図のスケールを合わせたのち、対象となる部屋の高さ情報 (実測による取得) を入力する。これにより図 3.5 (b) に示す 3 次元室内モデルを生成する。この 3 次元形状モデルを用いて壁の視差画像を生成する。視差画像を生成する理由は、3.4 節に述べる構造情報の統合処理において、実シーンから取得された視差画像と平面図から生成した視差画像の比較が容易に行うことができるからである。

表 3.3 奥行分解能

計測範囲 (m)	計測範囲に対する計測誤差の比率
14.7	20%
7.3	10%
3.6	5%
0.7	1%

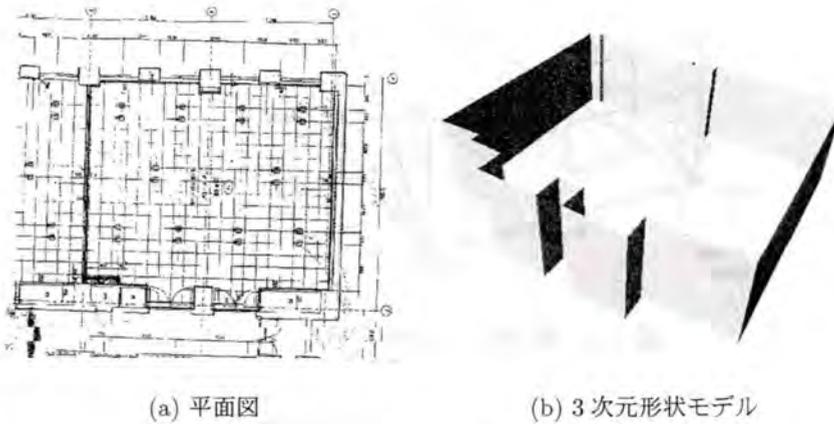


図 3.5 平面図と 3次元形状モデル

平面図から視差画像を生成するには、標準ステレオモデルと平面図より生成した 3次元形状モデルを用いる。具体的な方法を以下に示す。

図 3.6 に示すような標準ステレオモデルの場合には、3次元座標は、次の式で求められる。

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{ub}{d} \\ \frac{vb}{d} \\ \frac{fb}{d} \end{bmatrix}$$

$$d = u - u' \quad (3.3)$$

ここで、 (X, Y, Z) はカメラ座標系、 (u, v) は左画像の座標、 (u', v') は右画像の座標、 b は 2 台のカメラ間の基線長、 f はカメラの焦点距離、 d は視差である。

図 3.7 の位置 P_1 に、カメラを置いたと仮定すると、カメラと壁との実際の距離 L_i は、次のように表せる。

$$L_i = scale \times \sqrt{(x_i - x_{p_n})^2 + (z_i - z_{p_n})^2} \quad (3.4)$$

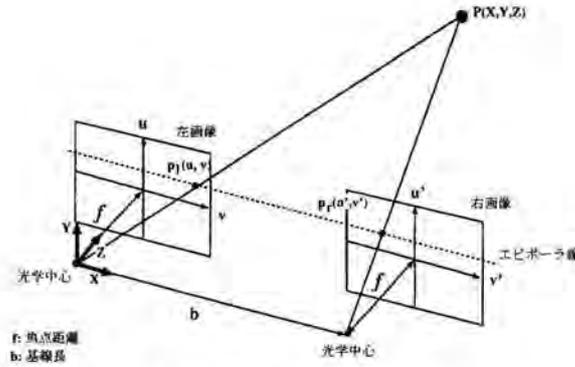


図 3.6 標準ステレオモデル

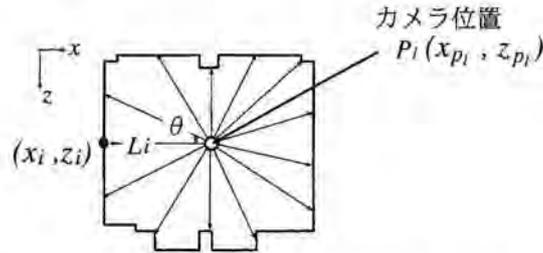


図 3.7 平面図からの構造情報の取得

ここで、 $scale$ は1ピクセルあたりの実際の距離、 (x_i, z_i) は、平面図から生成した3次元形状モデルにおける壁の座標、 (x_{p_n}, z_{p_n}) は、3次元形状モデルにおけるカメラ位置の座標である。壁の3次元座標は、垂直投影を仮定すると以下のようにして求められる。

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} L_i \sin \theta \\ y_j \\ L_i \cos \theta \end{bmatrix} \quad (3.5)$$

ここで、 θ ($0 \leq \theta \leq 2\pi$) は視線方向の角度、 y_j ($j = 0, \dots, height$) は、投影される画像の垂直解像度 $height$ である。したがって、壁の視差値は次のように計算することがで



図 3.8 平面図から取得された視差画像



図 3.9 キャリブレーションパターン

きる.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{d}{b} L_i \sin\theta \\ \frac{d}{b} y_j \end{bmatrix}$$

$$d = \frac{fb}{L_i \cos\theta} \quad (3.6)$$

式 (3.6) を用いて, 取得した壁の視差画像を図 3.8 に示す.

3.3.2 平面図におけるカメラ位置の決定

図 3.8 に示す壁の視差画像を取得するためには, 実際の部屋における 3 眼カメラの位置と平面図上における仮想的なカメラ位置との位置合わせが問題となる. 稲葉⁴⁵⁾らは, パノラマ画像と地図情報に基づき, 撮影点の位置と姿勢を計算する方法を提案している. また, リモートセンシング分野においては, 衛星画像と地図との位置合わせを行うために線情報を用いている. これは, まず画像から連続した長い線を抽出し, それらの形状を利用して地図中の対応する線を探査するものである^{46), 47)}.

しかし, いずれの手法においても, 平面図の仮想的なカメラ位置と実際に撮影した 3 眼カメラの位置を正確に一致させることは, 非常に困難である.

そこで本研究では、以下のようにして、平面図上のカメラ位置を推定することにより、実際に撮影したカメラ位置との位置合わせを行う。

Step1 : 対応点が確実に求められるキャリブレーションパターンを用いる。(図3.9参照)

Step2 : 実際の部屋において、3眼カメラから壁に貼られたパターンを撮影する。3眼カメラの位置からキャリブレーションパターンまでの距離をマルチベースラインステレオ法により推定する。

Step3 : 平面図上における仮想的なカメラ位置(図3.7の P_1)から壁まで距離を求める。

Step4 : **Step2** と **Step3** で求められた距離の差が最小となる点を平面図上のカメラ位置とする。

Step5 : **Step1** から **Step4** までの処理を壁に数ヶ所貼られたパターンに対して行い、平面図上のカメラ位置を推定する。

3.4 構造情報の統合

ここでは、平面図から得られる構造情報と実シーンから得られる構造情報を統合する方法について述べる。図3.4に示すように、実シーンから取得された視差画像は、いくつかの部分において精度が良くない。なぜなら、壁のようにテクスチャが一様な部分は、対応点の探索が難しいからである。

視差画像の補正は、平面図から取得された視差画像と、マルチベースラインステレオ法により取得された視差画像の視差値を比較することにより行う。

視差画像の視差値の比較方法は、図3.10に示すように、平面図から取得された壁の視差値を正しい視差値であると仮定し、視距離(視点と対象物体の距離)が近い方の視差値を用いる。この理由は、室内にある対象物体は、必ず壁より手前にあるからであり、壁より奥にある場合(図3.10の斜線部分)には、ステレオ法の誤対応とみなせるからである。具体的な比較方法は、以下のようなになる。

$r_d(u, v)$ をマルチベースラインステレオ法より取得された視差画像の (u, v) 座標における視差値、 $v_d(u, v)$ を平面図から取得された視差画像の (u, v) 座標における視差値

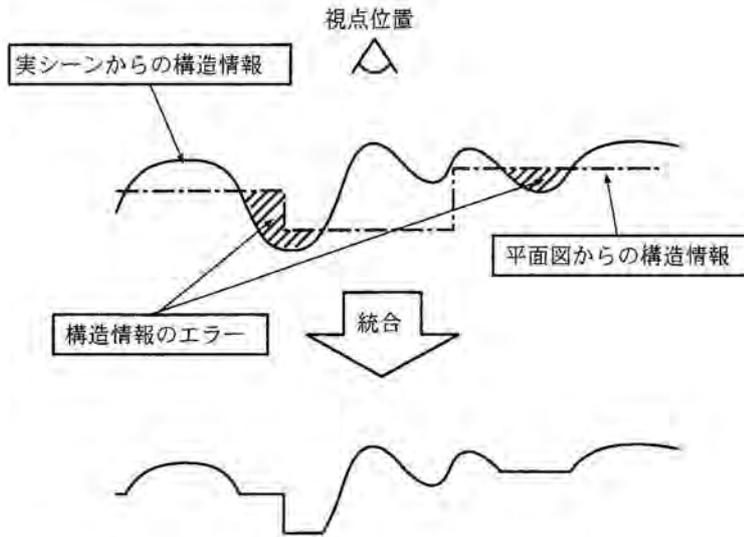


図 3.10 構造情報の統合

とすると、補正された視差値 $c.d(u, v)$ は式 (3.7) により算出される。これにより、部屋全周の視差画像を取得する。

$$c.d(u, v) = \begin{cases} v.d(u, v) & r.d(u, v) < v.d(u, v) \\ r.d(u, v) & \text{otherwise} \end{cases} \quad (3.7)$$

式 (3.7) を用いることで、部屋の壁に関して、ステレオ対応点探索の誤対応を補正することができる。しかし、壁より手前の物体に関しては、誤対応を補正することができない。そこで、手前の対象物体に関する補正手順を以下に示す。

Step1 : 実際の室内において、カメラ位置から最も近い物体までの距離を実測する。

Step2 : 実測された距離をもとに視差値を求め、それを閾値 th とする。

Step3 : (u, v) における補正された視差値 $c.d(u, v)$ が閾値 th 以上であれば、 (u, v) の周囲 8 近傍の平均視差値を (u, v) 座標における視差値とする。

Step4 : Step1 から Step3 の処理を閾値 th 以下になるまで繰り返す。

3.5 3次元形状モデルの生成と構造情報の記述

ここでは、上記の方法により統合された構造情報をもとにして、3次元形状モデルを生成し表現する方法について述べる。また、生成された3次元形状モデルを用いて、入力として与えた抽象モデルを更新することによりシーンの構造情報を記述する方法について述べる。

まず、3次元形状モデルの生成に関しては、上記で統合された視差値から3次元座標を求める必要がある。それには、カメラを回転させて360度撮影しているため、式(3.3)に回転行列を掛け合わせるにより求められ、次式ようになる。

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \frac{ub}{d} \\ \frac{vb}{d} \\ \frac{fb}{d} \end{bmatrix} \quad (3.8)$$

ただし、 $0 \leq \theta < 2\pi$ である。次に、得られた3次元座標点をもとに、三角パッチによるポリゴン近似を行い、室内シーンの3次元形状モデルを生成する。また、3次元形状モデルの表現に関しては、3眼カメラの左画像を貼り合わせて生成したパノラマ画像をテクスチャとして、3次元形状モデルにマッピングすることにより、室内シーンの3次元形状モデルを表現する。また、生成された3次元形状モデルには、部屋全体(壁および対象物体)の構造情報が含まれ、データ構造は、3次元座標点、その点に対応する画像の2次元座標点、および接続リスト(座標点間の接続関係)として記述されており、これがシーン構造情報の物理的な記述である。

抽象モデルの更新に関しては、図3.11に示すように、3次元形状モデルを平面に正射影投影し、投影された平面から輪郭抽出を行うことにより、対象物体の構造情報を含んだ平面図を生成する。この方法により、生成した3次元形状モデルの構造情報を抽象化する。人間は、このようにして生成された平面図からでも部屋の形状とその中にある物体の構造情報を読みとることができる。すなわち、生成された平面図は、抽象モデルを図式化したものであり、人間の内部に更新された抽象モデルが形成され、シーンの構造情報が記述されたものと考えられる。

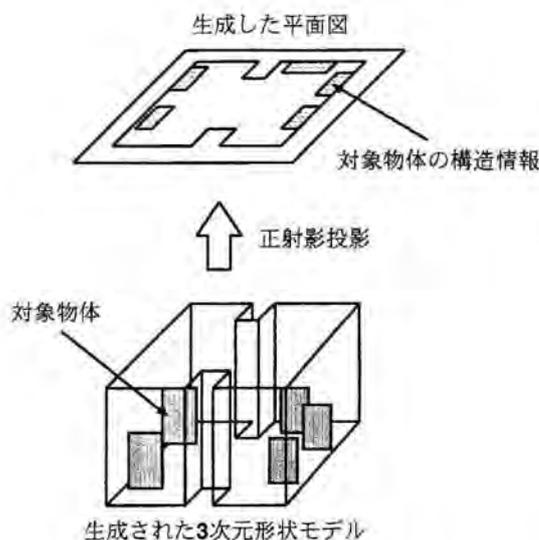


図 3.11 図式化された抽象モデル(平面図)の生成

3.6 実験結果と考察

ここでは、平面図とステレオ法を併用して取得された構造情報の有効性を確認する。まず、本手法で取得された視差画像の $y = 45$ ライン目をもとに、カメラから対象物体まで距離を推定し、その x 方向と実測値を比較したものを図 3.12 に示す。ここでの実測値とは、カメラ位置から対象物体までを実測した値である。また、マルチベースラインステレオ法 (MBS 法) と本手法との精度を比較するために、実測値との平均相対誤差を表 3.4 に示す。図 3.12 においては、本手法による推定値が、ほぼ実測値上にプロットされていることがわかる。また、表 3.4 においては、マルチベースラインステレオ法

表 3.4 マルチベースラインステレオ法と本手法の精度

手法	計測誤差
MBS 法	38.5%
本手法	4.5%

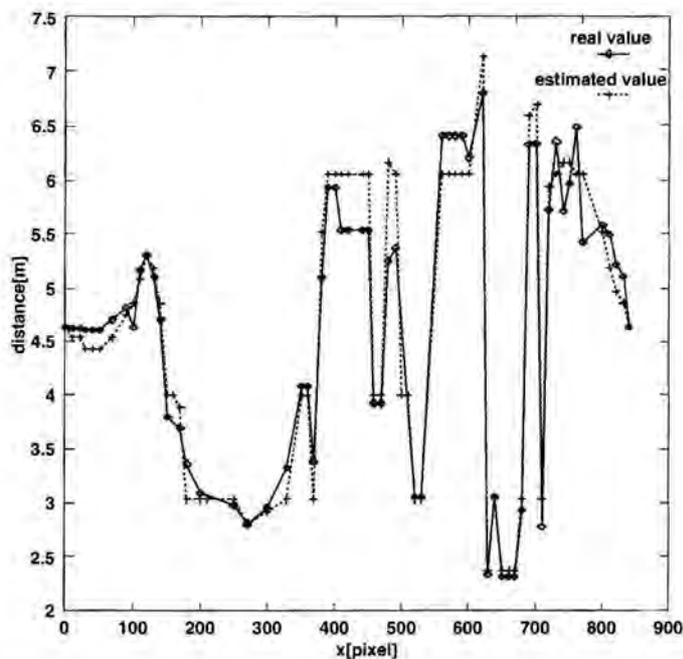


図 3.12 本手法における推定値と実測値 (line 45)

と比較して、平均相対誤差が減少し、本手法が有効であったことがわかる。

図 3.13(a)~(d) に本研究で取得された構造情報をもとに、部屋全体の3次元形状モデルを生成した結果を示す。図 3.13 では、部屋全体の形状および、部屋に配置されている対象物体の形状が取得されていることがわかる。

最後に、生成した3次元形状モデルを用いて、3.5節に述べた方法により、抽象モデルを更新し、抽象モデルを図式化したものを図 3.13(e) に示す。この図式化により、平面図だけでは、室内にある物体の構造情報が欠落していたが、生成した3次元形状モデルを用いることによって、平面図に室内にある物体の構造情報を補うことができた。生成した平面図は、図式化した抽象モデルが更新されたものと考えられ、人間の内部に更新された抽象モデルが形成され、シーンの構造情報が記述できたものと考えられる。

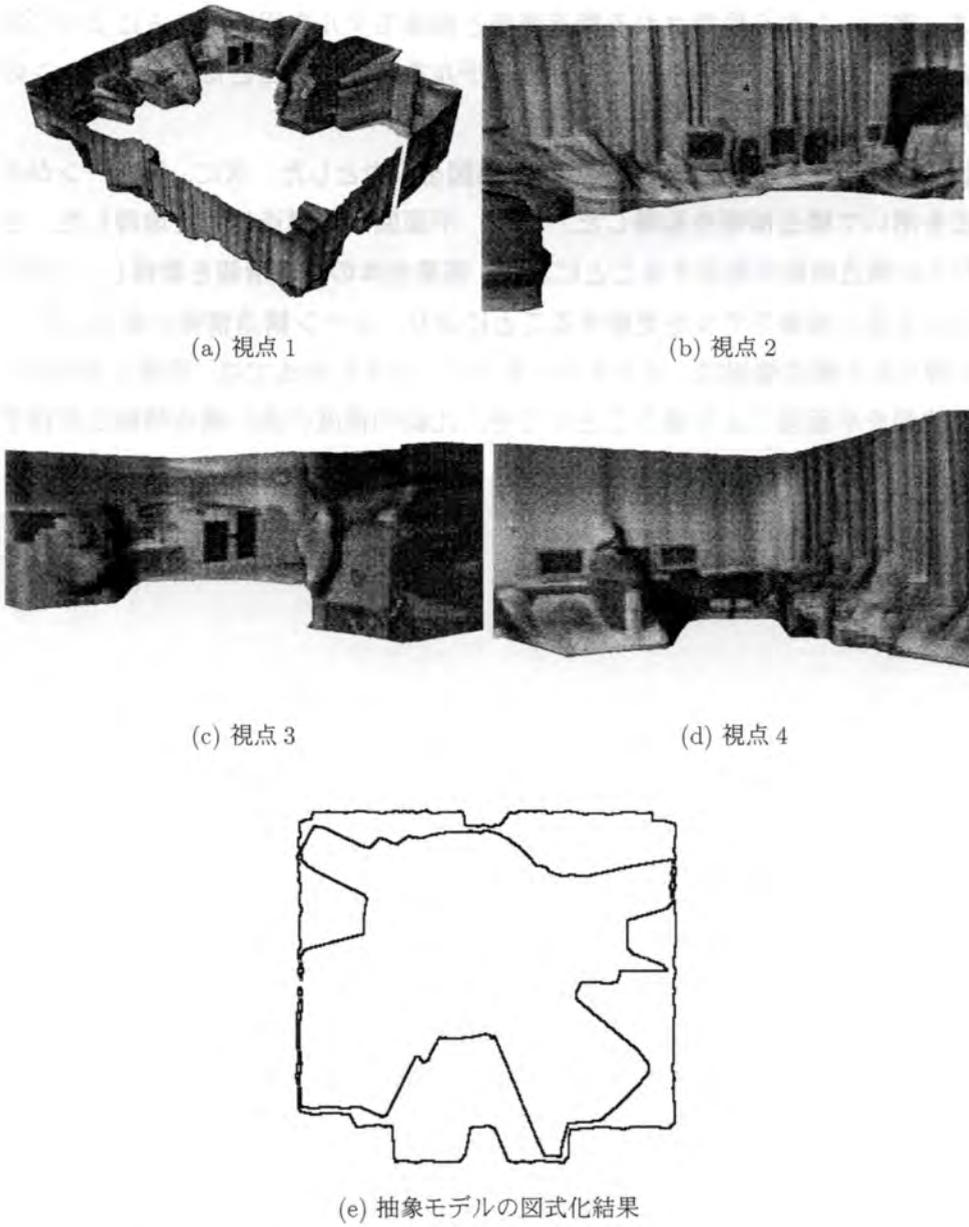


図 3.13 3次元形状モデルの生成と抽象モデルの図式化結果

3.7 まとめ

本章では、実シーンから取得される構造情報と抽象モデルを用いることによって室内シーンの3次元形状モデルを生成し、抽象モデルを更新することにより、シーンの構造情報の記述を行った。

ここでは、まず、抽象モデルを図式化した平面図を入力とした。次に、実シーンからステレオ法を用いて構造情報を取得した。また、平面図から構造情報を取得した。さらに、これらの構造情報を統合することにより、部屋全体の構造情報を取得し、3次元形状モデルの生成と抽象モデルを更新することにより、シーン構造情報を記述した。

本手法で得られた構造情報は、マルチベースラインステレオ法では、取得できなかった壁の構造情報を平面図により補うことができ、比較的精度の高い構造情報を取得することが可能であること示した。

第4章

画像情報と抽象モデルを用いた3次元形状モデルの生成と記述

本章では、抽象化モデルを図式化した手書き立面図(以降、手書き立面図と呼ぶ)と、実シーンを撮影した画像を用いて、実シーンから構造情報を取得し、3次元形状モデルを生成する。また、生成した3次元形状モデルを用いて、抽象モデルを更新することによって、シーンの構造情報を記述することを目的とする。ここでは、石窟寺院、仏塔、寺院を対象とした電子記録化方法への応用例を示す。

4.1 はじめに

近年、計算機の飛躍的な3次元表示能力の向上により、3次元コンピュータグラフィックスを利用した応用分野がますます広がっている。その1つに、3次元CGを用いた仮想考古学(Virtual Archaeology)⁴⁸⁾と呼ばれる研究がある。従来、研究者間の遺跡や遺構等に関する情報を伝達する手段として、文書、画像、スケッチ等が用いられてきたが、このような2次元的な情報では、直感的に遺跡や遺構等の形状を理解することは難しいものである。しかしながら、近年では、レーザレンジファインダを用いて、大規模な遺物を対象とした3次元形状計測が行われてきている⁴⁹⁾。

仮想考古学の発展は、遺跡や遺構に関しての新たな発見や認識を研究者間で共有したり、研究者と一般の人との共有な空間を可能とするものである。また、仮想考古学

は、我々に新たな知見を与えることが可能となる。例えば、光源と遺跡の配置関係について考えた場合、光源の位置・種類が、遺跡に与える影響について、計算機上で知ることができる。

また、現存する遺跡・遺構の保護・保存には多大なコストがかかる。例えば、中国新疆ウイグル自治区には、遺跡・遺構が数多く存在しているが、すべての遺跡・遺構を保護、保存することは現状では難しい。その一例として、図4.1に示すトユク石窟は、トルファン市街から東へ60kmほど離れた郊外にある石窟であり、大小合わせて46窟ある。図に示す1号窟は、2001年の夏に倒壊したものである。トユク石窟は、岩盤が脆く、崩れやすいにも関わらず、現在、石窟の補修および保護がされていない。このような状況から、早急に石窟に関して、詳細な電子記録が必要となってくる。特に、石窟を電子記録する場合には、石窟内にどのような壁画がどの場所に描かれていたのか、その石窟の主題を探るための非常に重要な手がかりとなる。そのため、石窟の形状および壁画の位置関係を記録する必要がある。現在、新疆ウイグル自治区にある石窟に関する電子記録(出版物)は、文献⁵⁰⁾のように、石窟内を撮影した壁画の画像と手書きで描かれた石窟の立面図のみである。このような電子記録では、壁画が石窟内のどの場所に描かれていたのか、直感的に把握することが難しい。このような問題を解決するためには、石窟の3次元形状モデルを生成し、壁画の位置関係を明示する必要がある。

ここでは、石窟や仏塔の3次元形状モデルの生成に関して、対象が膨大に存在するため、処理コストや機動性を重視し、撮影環境が簡易であり、屋内外に適用しうる汎用的な手法の確立を目的とする。その一例として、石窟を対象とし、手書き立面図からの3次元形状モデルを生成する手法と壁画の位置関係の記述(テキストチャマッピング)方法⁵¹⁾について述べる。また、撮影環境を十分に整えることができない屋外にある仏塔においても、本手法を適用することによって、3次元形状モデルの生成が可能であることを示す。本手法の特徴は、遺跡・遺構の概略として利用される立面図(正・平・側面図)を用いて3次元形状モデルを生成することにある。この利点は、すでに出版されている壁画の画像と立面図を再利用することができ、崩壊した石窟の場合でも、出版されたものが存在すれば、3次元形状モデルの生成が可能となる点にある。

また、屋外にある仏塔に関して立面図がない場合には、数枚の画像から正投影図(正射影変換された画像)を生成することにより、3次元形状モデルの生成が可能となる点



図 4.1 トユク 1 号窟の倒壊例 (2001 年 11 月撮影)

である。

4.2 概要

立面図から 3 次元形状モデルを生成する代表的な方法として、以下の 2 つに分類できる。1 つは、タブレットを用いて対話的に人が立面図 (三面図) を描きながら、逐次的に 3 次元形状モデルを生成する方法⁵²⁾である。もう 1 つは、立面図の稜線や頂点を用いて、自動的に 3 次元形状モデルを生成する方法⁵³⁾である。これらの研究で対象としている立面図は、機械・機構などの精密なものであり、2 次元 CAD により生成されたものである。しかし、ここで扱う立面図は、遺跡・遺構の概略表記として利用されるものであり、正確に描かれたものではない。そのため、正確な稜線や頂点情報に基づく従来法では、3 次元形状モデルを生成することは困難である。

図 4.2 に全体の処理の流れを示す。図中で実線は計算機による処理であり、点線は人間の作用を示す。ここでは、直接、対象に関する構造情報を取得することが難しい画像を入力とするため、まず、抽象モデルを図式化した手書き立面図から 3 次元形状モデルを生成する。それには、手書き立面図 (正面、平面、側面図) をそれぞれ 1 つの領域 (対象物体のシルエット) として扱い、視体積交差法⁵⁴⁾を適用し 3 次元形状モデルを生成する。生成した 3 次元形状モデルは、実シーンの形状と必ずしも一致することはない。なぜならば、手書き立面図を用いて、3 次元形状モデルを生成しているから

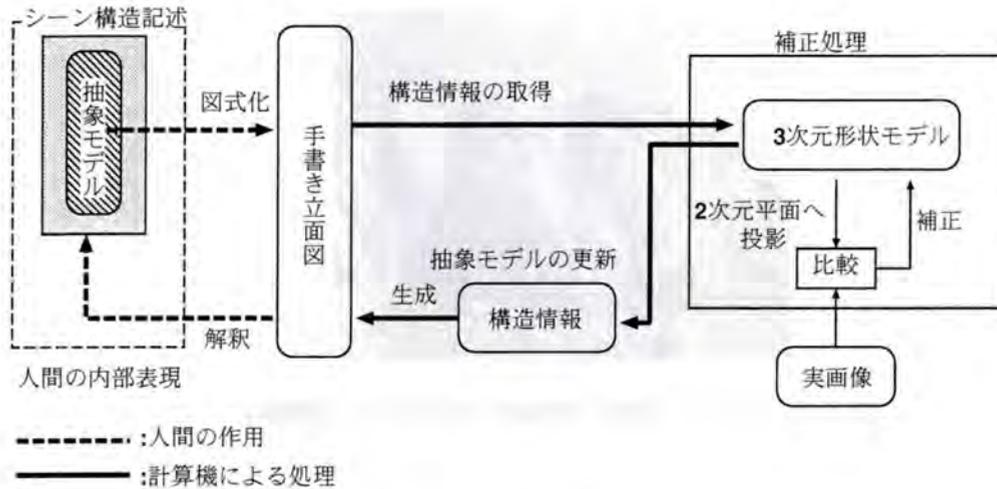


図 4.2 処理の流れ

である。また、3次元形状モデルに実画像をテクスチャマッピングするためには、生成した3次元形状モデルと実画像のエッジ特徴を用いて、形状マッチングを行う必要があるが、生成された3次元形状モデルは実シーンの形状と異なるため、従来法⁵⁵⁾⁵⁶⁾⁵⁷⁾の適用は難しい。そこで、生成された3次元形状モデルを2次元平面に投影し実画像と比較しながら、対話処理を利用して3次元形状モデルを補正する。次に、補正された3次元形状モデルと実画像の対応関係を求め、3次元形状モデルに実画像をマッピングし、3次元形状モデルを表現する。このとき、3次元形状モデルのデータ構造は、3次元座標点、その点に対応する画像の2次元座標点、座標点の接続リストで記述されており、これがシーン構造情報の物理的な記述である。さらに、生成された3次元形状モデルの輪廓形状を抽出し、立面図を生成することで、立面図を補正することにより、抽象モデルを更新し、シーンの構造情報を記述する。

また、ここでは、図式化された抽象モデルが存在しない対象に関して、4方向(対象物体の側面方向)から撮影したカメラパラメータが未知な実画像と上記の方法を用いて、3次元形状モデルを生成し記述する。具体的には、先に述べた手書き立面図から3次元形状モデルを生成する方法を実画像に適用するため、実画像から必要なパラメータ(移動量、回転角、拡大率)を求める。次に、得られたパラメータをもとに、画像を剛体変

換およびスケール変換を行うことにより，対象物体の正投影図(4つの側面図)を生成する．得られた正投影図を用いて，視体積交差法を適用し，3次元形状モデルを生成する．さらに，実画像をテクスチャとして，3次元形状モデルにマッピングを行う．これにより，上記と同様に対象の構造情報の記述が可能となる．以下に，具体的な方法について述べる．

4.3 対話的処理による3次元形状モデルの生成と記述

ここでは，立面図から3次元形状モデルを生成する方法について述べる．

4.3.1 前処理

3次元形状モデルを生成する前処理として，まず，図4.3(b)に示す立面図を2値化する(ここでは，対象領域を白とする)．次に，1枚の立面図に，正面図，側面図，平面図が描かれているために，これを3枚の画像(正面画像，側面画像，平面図画像)に分離する．それには，図4.4に示すように，画像の x 軸， y 軸に対して，それぞれ白画素の分布を調べ，画像の x 軸， y 軸に対する各ヒストグラムにおいて，谷となる点を対象領域の境界として，立面図の分離を行う．

4.3.2 視体積交差法

上記で，分離された正面図，側面図，平面図の3枚の画像に対して，視体積交差法を用いて3次元形状モデルの生成を行う．

視体積交差法⁵⁴⁾は，図4.5に示すように，実空間内に複数設置したカメラで撮影した画像から物体のシルエットを抽出し，あらかじめカメラキャリブレーションにより求まるカメラパラメータを用いて，空間に逆投影し，シルエットの交わりを計算することによって，3次元形状モデルを求める手法である．

ここで使用する立面図(正・平・側面図)は，画像面に垂直な軸に沿って投影する正射影図法によるものであり，奥行情報のみが失われる．このため，視体積交差法を用

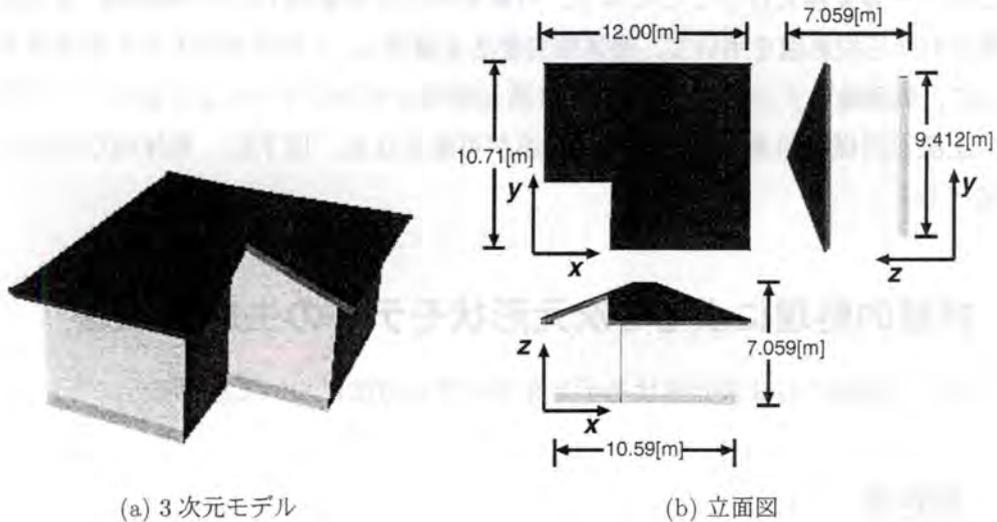


図 4.3 立面図

いて、3次元形状モデルを生成する場合に、必要とされるカメラパラメータを求める必要がない。

4.3.3 立面図からの3次元形状モデルの生成

以下に立面図から3次元形状モデルを生成する手順を述べる。

Step1 形状を構成する3次元空間(ボクセル空間)を立方体格子(ボクセル)に分割。

Step2 各ボクセルに対して、正射影による逆投影を行い、立面図(正面図・平面図・側面図)上におけるシルエットの存在有無の判定。

Step3 最終的に存在するボクセルデータを3次元形状モデルとみなす。

Step4 内部のボクセルを削除し、サーフェスモデルを生成する。

ただし、Step2で使用するシルエット画像は、立面図から生成したものである。また、Step4の処理は、3次元形状の表面にテクスチャマッピングを行うための前処理であり、

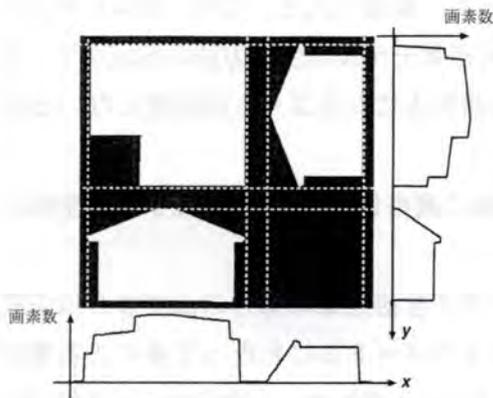


図 4.4 立面図の分割

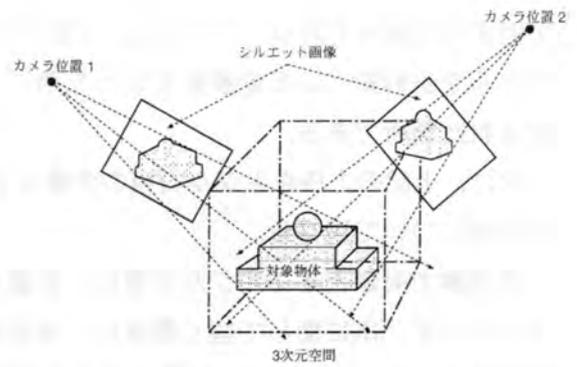


図 4.5 視体積交差法

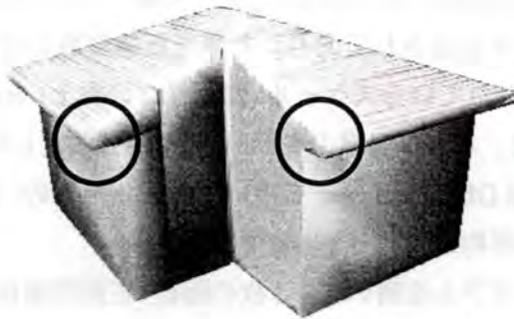


図 4.6 3次元形状モデルの生成結果

ボクセルデータからサーフェスモデルへの変換には、マーチング・キューブ法⁵⁹⁾を適用した。

上記の方法により、図 4.3(b) から生成した 3 次元形状モデルを図 4.6 に示す。このとき、ボクセル空間の大きさは、 $12.59[\text{m}] \times 12.15[\text{m}] \times 7.365[\text{m}]$ であり、各ボクセルの大きさは、 $3.200 \cdot 10^{-2}[\text{m}] \times 3.200 \cdot 10^{-2}[\text{m}] \times 3.200 \cdot 10^{-2}[\text{m}]$ である。

生成した 3 次元形状モデルの精度は、図 4.3(a) と図 4.6 の 3 次元形状モデルに対して、各辺の長さの誤差を求め、評価を行った。その結果、最小誤差は、 0.003m であり、最

大誤差は、0.453m、平均誤差は、0.077mとなった。誤差の大きい辺は、図4.6内の円で囲まれた部分である。これらは、上記の3次元形状モデルの生成方法のStep2で、ボクセルを正射影による逆投影を行った際に、立面図上にシルエットが存在しないと判定された箇所である。

次に、上記の方法をキジル石窟の手書き立面図に適用し、3次元形状モデルの生成する方法について述べる。

本実験で対象とするキジル石窟は、新疆ウイグル自治区庫車郊外のムザルト川の北岸にあって、南に面して続く断崖に、東西約2キロメートルにわたって多くの石窟が存在する。キジル石窟は、新疆の三大文化遺跡の1つと言われ、3世紀から9世紀ごろまで造営され、石窟の総数は、237窟あり、石窟の半数以上は、僧の住居や瞑想などに使われた僧房窟であるが、壁画や彫像で飾られた礼拝用の祠堂窟も50以上存在する。

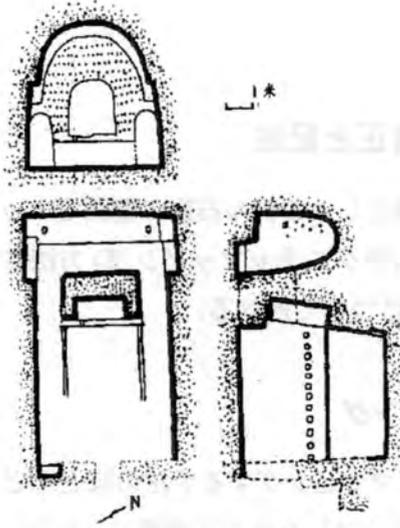
まず、入力として用いる立面図(図4.7(a)(c))は、文献⁶⁰⁾から引用したものである。この立面図の特徴は、手書きで描かれており、対象を正確に表したものではない点にある。実際に、立面図に描かれているスケール(図4.7(a)の右上)を基に、8号窟内の5箇所(前室の奥行、幅、高さ、および奥室の奥行、高さ)から測定した実測値(レーザー距離計 Leica Geosystems 社製 DISTO classic を用いて計測したもの)と比較してみると、図4.7(a)に示す正投影図の平均誤差は0.366mであった。

次に、立面図からヒストグラムを用いて、3枚の画像(正面図画像、平面図画像、側面図画像)に分割する。

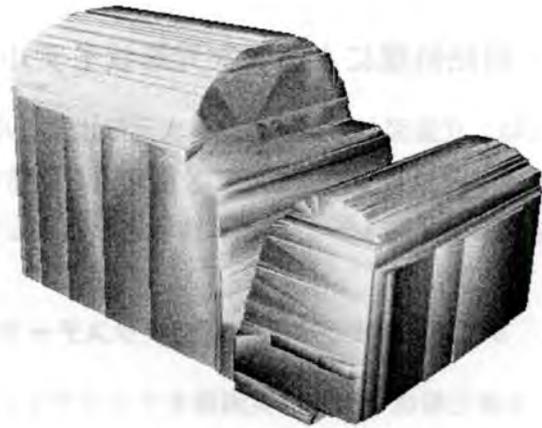
さらに、各画像に対して、輪郭抽出を行ったあと、輪郭形状を基にして、シルエット画像を生成する。

最後に、シルエット画像を用いて、視体積交差法に基づき、生成した3次元形状モデルを図4.7(b)(d)に示す。ただし、図4.7(b)に関しては、ボクセル空間の大きさは、5.828[m] × 8.862[m] × 5.276[m]であり、各ボクセルの大きさは、 $3.400 \cdot 10^{-2}$ [m] × $3.400 \cdot 10^{-2}$ [m] × $3.400 \cdot 10^{-2}$ [m]である。また、図4.7(d)に関しては、ボクセル空間の大きさは、5.588[m] × 6.706[m] × 10.41[m]であり、各ボクセルの大きさは、 $5.800 \cdot 10^{-2}$ [m] × $5.800 \cdot 10^{-2}$ [m] × $5.800 \cdot 10^{-2}$ [m]である。

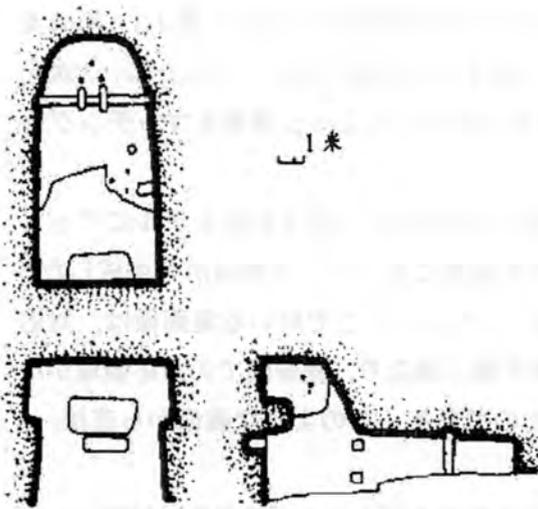
上記の実験結果より、手書きで描かれた立面図から稜線や頂点の情報を用いる手法では、3次元形状モデルの生成が困難であったが、立面図(正面図、平面図、側面図)を



(a) キジル8号窟の立面図



(b) キジル8号窟の3次元形状モデル生成結果



(c) キジル139号窟の立面図



(d) キジル139号窟の3次元形状モデル生成結果

図 4.7 キジル石窟の3次元モデルの生成結果

それぞれ1つの領域として扱うことによって、3次元形状モデルの生成が可能となり、上記の手法が有効であることが確認された。

4.3.4 対話処理による3次元形状モデルの補正と記述

ここでは、立面図から得られた3次元形状モデルを補正し、キジル石窟の壁画をマッピングすることによって、壁画の位置関係を記述する(テクスチャマッピング)方法について述べ、さらにシーンの構造情報を記述する方法について述べる。

4.3.4.1 3次元形状モデルの補正とテクスチャマッピング

従来、3次元形状モデルに実画像をテクスチャとしてマッピングする手法のほとんどが、レンジセンサを用いて得られる距離画像(もしくはリフレクタンス画像)とカラー画像からそれぞれのエッジ情報を抽出し、各エッジを比較することで、位置合わせを行うものである⁵⁵⁾⁵⁶⁾⁵⁷⁾。

しかしながら、ここで扱う立面図から生成した3次元形状モデルは、第4.3.3節の実験結果が示すように、必ずしも実際の対象と一致するとは限らない。そのため、3次元形状モデルから抽出したエッジ情報と実画像から抽出したエッジ情報をマッチングさせる従来手法の適用は難しい。

また、実画像からテクスチャとして立面図から生成した3次元形状モデルにマッピングするためには、実画像から取得した3次元情報によって、立面図から生成した3次元形状モデルを補正する方法が考えられる。しかし、ここで用いる実画像は、対応点が求められない画像(同じ石窟内でも、撮影位置が異なり、画像間での対応領域がないもの)であり、カメラパラメータが未知なものである。このような画像から直接、3次元情報を取得することが困難である。

そこで、ここでは、立面図から生成した3次元形状モデルを2次元画像に投影し、実画像と比較しながら、対話的に3次元形状モデルを補正し、壁画のテクスチャマッピングを行う。図4.8に処理概要を示し、以下に各方法について述べる。

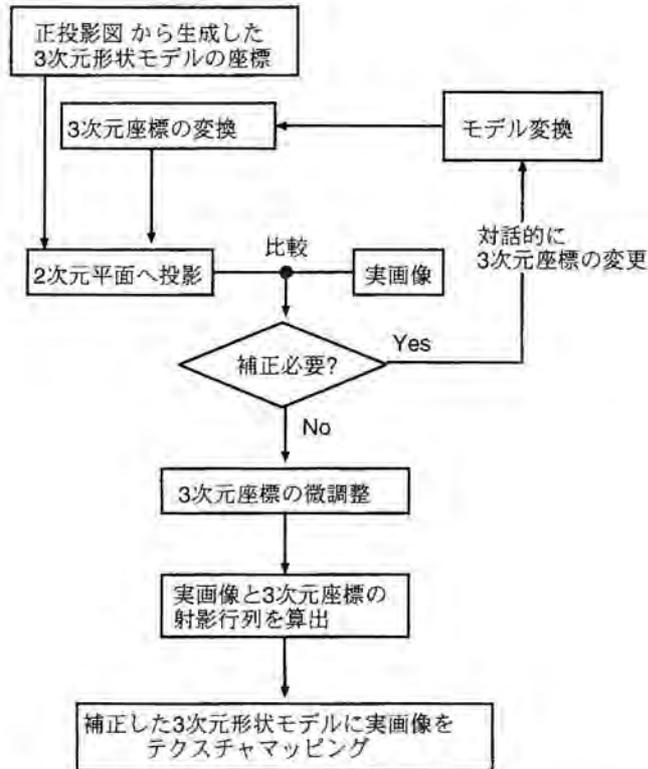


図 4.8 対話的な補正処理によるテクスチャマッピング

Step1:2 次元平面への投影

実画像の撮影したカメラは、透視投影カメラモデルであると仮定し、立面図から生成した3次元モデルから2次元画像への投影は透視投影とする。透視投影変換は、3次元グラフィックスライブラリの1つである OpenGL⁶¹⁾ の `gluPerspective` 関数を用いた。この関数に必要なパラメータである縦方向の画角 (fovy) の値は、60 に設定した。また、カメラ位置と方向に関しては、撮影された実画像のカメラ位置が未知であり、また1枚の画像からでは、カメラ位置を推定することは難しく、視点を設定することができない。そこで、ここでは、視点をある任意の1点に固定し、3次元形状モデルに対して、剛体変換・スケール変換 (モデル変換) を行い、実画像との形状マッチングを行う。

Step2:3次元形状モデルと実画像の比較

実画像と2次元平面に投影された3次元形状モデルを比較し、3次元形状モデルに補正が必要な場合、3次元形状モデルに対して、モデル変換を行う。

3次元形状モデルの変換は、平行移動 T 、スケーリング変換 S 、および回転変換 R である。具体的には、変換前の3次元座標を (X, Y, Z) とし、変換後の3次元座標を (X', Y', Z') とすると、次のように 4×4 の行列で表され、

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.1)$$

平行移動の場合、 4×4 の行列は、

$$T = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.2)$$

で表せる。

また、スケーリング変換の場合、 4×4 の行列は、

$$S = \begin{bmatrix} x & 0 & 0 & 0 \\ 0 & y & 0 & 0 \\ 0 & 0 & z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.3)$$

で表せる。

さらに、回転変換では、 x 軸回りの回転変換を R_x 、 y 軸回りの回転変換を R_y 、 z 軸回りの回転変換を R_z とすると次のように表される。

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

$$R_y = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.5)$$

$$R_z = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.6)$$

Step3: Step1 から Step2 までの処理を繰り返し、3次元形状モデルの補正を行う。

Step4:3次元形状モデルの微調整

上記の方法により、実画像と3次元形状モデルとの大まかな形状マッチングを行った後、3次元形状モデルの3次元座標点の微調整を対話的に行う。

Step5:実画像と3次元座標の対応

対話的に補正された3次元形状モデルに対して、テクスチャマッピングを行うために、補正された3次元形状モデルの3次元座標点と実画像の2次元座標点の関係を求める必要がある。ここでは、射影行列を求めることにより、2次元座標と3次元座標の関係を求める。それには、射影行列を

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} P_1^T & p_{14} \\ P_2^T & p_{24} \\ P_3^T & p_{34} \end{bmatrix} \quad (4.7)$$

とし、補正された3次元形状モデルの3次元座標点を M 、その点を画像上に投影した2次元座標点を m とする。ここで、

$$m = [u, v, 1]^T, M = [X, Y, Z, 1]^T \quad (4.8)$$

とすると、次の関係が成り立つ。

$$sm = PM \quad (4.9)$$

射影行列 P の算出は、立面図から生成される3次元形状モデルにおける3次元座標点 M と、その点を画像に投影した2次元座標点 m を与え、6点以上の対応する点を用いて、式(4.9)から得られた解を初期値として、対応する点を用いて、式(4.10)の C を最小にする評価関数を非線形最小2乗法によって解くことが可能である。

$$C = \sum_{i=1}^N \left(\frac{P_1^T M_i + p_{14}}{P_3^T M_i + p_{34}} - u_i \right)^2 + \sum_{i=1}^N \left(\frac{P_2^T M_i + p_{24}}{P_3^T M_i + p_{34}} - v_i \right)^2 \quad (4.10)$$

4.3.4.2 構造情報の記述

上記で3次元形状モデルの3次元座標と画像との対応関係が求まっているため、シーン構造情報の物理的な記述データは、3次元座標点、その点に対応する画像の2次元座標、座標点の接続リストである。また、抽象モデルの更新に関しては、生成した3次元形状モデルをXY平面、YZ平面、ZX平面にそれぞれ正射影投影を行い、投影された各平面から対象の輪郭抽出を行うことにより、立面図を生成する。この方法により3次元形状モデルの構造情報を抽象化する。生成された立面図は、入力とした与えた立面図を実画像を使った修正されたものであり、更新された抽象モデルを図式化したものであり、シーン構造情報が記述されたものと考えられる。

4.3.5 実験結果と考察

上記の方法を用いて、キジル8号窟の3次元形状モデルに図4.9に示すような壁画をテクスチャとして、マッピングした結果を図4.10に示す。ただし、図中で、テクスチャマッピングがされていない領域は、本実験で使用した画像内に、3次元形状モデルに対応する壁画が撮影されていないためである。

図4.10(a)(b)(c)に示すように、立面図から生成した3次元形状モデルを補正することにより、石窟内に壁画の位置関係を記述することができ、視覚的に壁画の位置を理解することができる3次元形状モデルの生成が可能となった。また、補正した3次元形状モデルに関しての精度評価に関しては、実際の石窟の計測データが少ないため、厳密に評価することが難しいが、8号窟内の5箇所から計測した実測値に対して、表4.1に示すように、3次元形状モデルの補正前と補正後の精度比較を行う。表4.1からは、

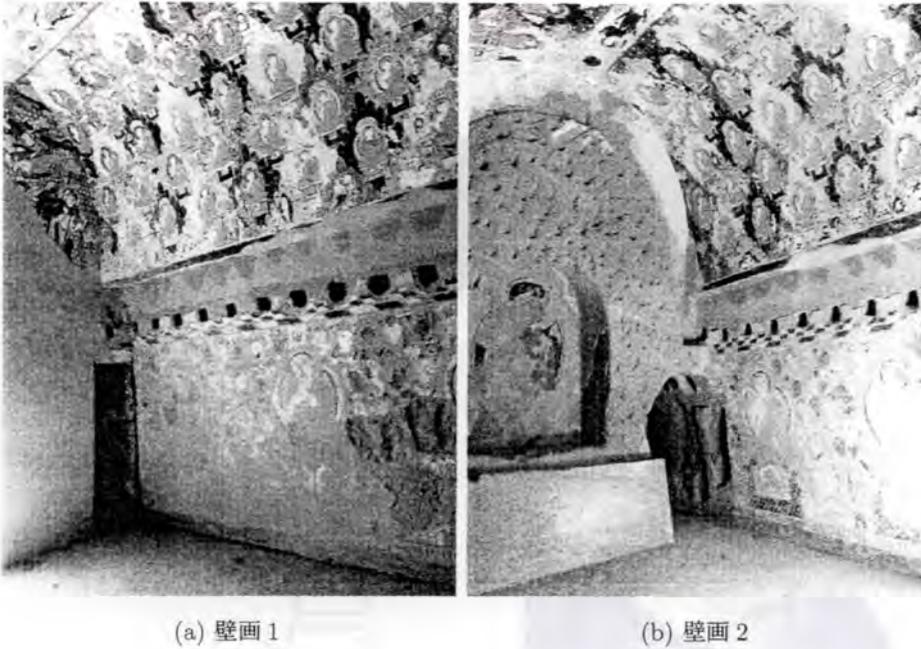


図 4.9 キジル 8 号窟内の前室部分の壁画 (平凡社 1983)

補正前と比べて、平均誤差が減少し、補正後の3次元形状モデルの精度が向上したことがわかる。

また、補正した3次元形状モデルを用いて、4.3.4.2節に述べた方法により、抽象モデルを更新し、抽象モデルの図式化を行った結果を図4.10(d)に示す。この図式化により、手書き立面図をより実シーンに近い立面図に修正することが可能となり、抽象モデルが更新され、シーンの構造情報が記述できたものと考えられる。

表 4.1 キジル 8 号窟の精度比較

3次元形状モデル	平均誤差 (m)
補正前	0.46
補正後	0.19



図 4.10 キジル 8 号窟の 3 次元形状モデルの生成と抽象モデルの図式化結果

4.4 実画像からの 3 次元形状モデルの生成と記述

ここでは、抽象モデルを図式化したものが存在しない場合に、実画像から正投影図を生成し、3次元形状モデルを生成する方法について述べる。また、生成した3次元形状モデルから抽象モデルを図式化したもの(正投影図)を生成する。ただし、ここでは対象を撮影した画像を正射影変換した画像を用いるために、立面図とは呼ばず、正投影図と呼ぶことにする。

4.4.1 実画像から正投影図の生成

上述では、抽象モデルを図式化した手書き立面図を用いて、対象に関して、3次元形状モデルの生成、および構造記述を行ってきた。ここでは、屋外にある対象に関して、抽象モデルを図式化したものが存在しない場合に、実画像から正投影図(抽象モデルを図式化したもの)を生成し、これを用いて、4.3節で述べた方法により、3次元形状モデルの生成を行う。一般的に、実画像から視体積交差法を用いて、直接3次元形状モデルを生成することが可能である。しかし、カメラキャリブレーション、対象物体の大きさ、撮影環境に制限があり、屋外環境への適用が難しい。そこで、ここでは、第4.3節で述べた方法を屋外環境に適用するため、入力画像に対して、剛体変換・スケール変換を行うことにより、実画像から正投影図を生成し、これに対して視体積交差法を適用し、3次元形状モデルを生成する。

まず、実画像から正投影図を生成するために、対象物体に関する以下のような条件設定を行う。

- (1) カメラモデルとして弱透視投影を仮定する。
- (2) 線対称の物体である。
- (3) 対象は、正面、背面、右側面、左側面の4方向から撮影されたものとする。

なお、上記の条件(2)に関しては、ここでは、対象物体の中心軸を基にして、正投影図の生成を行っている。このため、3次元形状モデルを生成する対象は、線対称であるものが望ましい。ここで扱う仏塔の形状は、厳密ではないが、線対称であるものが数多く存在する。このような対象に対して、後述する環状ウィンドウのパラメータを設定することによって、中心軸を求めることが可能となる。

上記の条件の下、実画像から対象物体を抽出し、移動量、回転角度、拡大率を求めて、剛体変換・スケール変換を行うことにより、実画像から正投影図を生成する。以下に、その方法について述べる。

Step1 弱透視投影カメラモデル

透視投影は、実際のカメラによる投影に近いものであるが、非線形であり扱いに

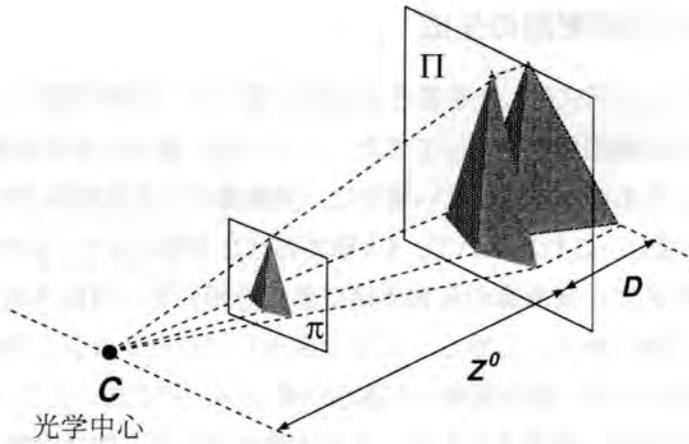


図 4.11 弱透視投影

くい。一方、正射影は、線形であり、正投影図を生成する場合に関しては、最適であるが、実際のカメラによる投影とはかなり異なる。そこで、弱透視投影カメラモデル⁶²⁾を適用する。これは、正射影と透視投影を組み合わせた次のような投影である。

図4.11に示すように、まず対象物体を画像面 π に平行な平面 Π に正射影し、その投影像をさらに画像面 π へ透視投影する。これが、透視投影である。最初の投影は、正射影であるから線形である。これは、対象物体上の各点の X 座標、 Y 座標を保存したまま Z 座標のみ一定の値、すなわち平面 P_i の Z 座標である Z^0 に変更することに等しい。したがって、この投影は式(4.11)のように表せる。

$$X' = X \quad Y' = Y \quad Z' = Z^0 \quad (4.11)$$

次の投影は、平面からそれに平行な平面への投影であるから、式(4.12)のように表せる。

$$x = \frac{X'}{Z'} = \frac{X}{Z^0} \quad y = \frac{Y'}{Z'} = \frac{Y}{Z^0} \quad (4.12)$$

Z^0 は定数であるから、これらの式は X と Y に関して線形である。したがって、これら2つの組み合わせた弱透視投影は線形であることがわかる。

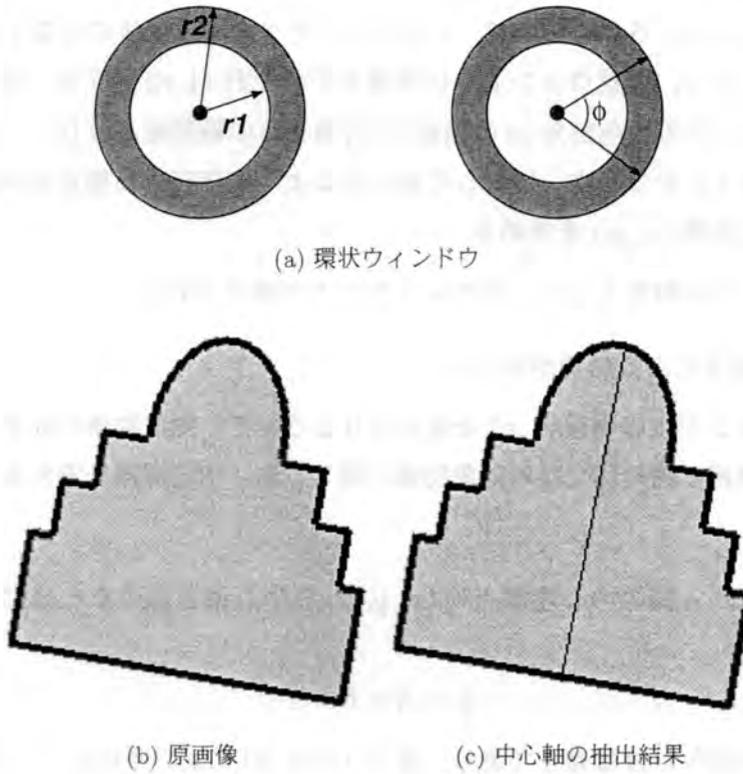


図 4.12 中心軸の抽出

また、弱透視投影が成り立つ条件は、図 4.11 において、一般的に、対象物体までの距離 Z と対象物体の Z 軸方向の厚み D との比が 10:1 以上である。

Step2 対象物体の輪郭抽出

対象物体と背景領域を分離するために、RGB 色空間から HSV 色空間へ変換し、明度 (V) をもとに、閾値処理により行う。次に、閾値処理により得られた画像から動的輪郭モデル Snakes⁶³⁾ を用いて輪郭抽出を行い、対象物体を抽出する。

Step3 対象物体の中心軸抽出

輪郭抽出された画像から正投影図を生成するために、移動量、拡大率、回転角度を求め、画像に対して、剛体変換・スケール変換を行う必要がある。そのため

に、まず対象物体の中心軸を求める。対象物体の中心軸抽出は、図4.12(a)に示すような環状ウィンドウを用いて、このウィンドウを画像全体に走査することにより行う。ここで、環状ウィンドウの半径をそれぞれ r_1, r_2 とする。図(a)に示す環状ウィンドウの灰色部分(ϕ の範囲)に対象物体の輪郭線が存在し、かつ、その輪郭線がウィンドウの中心に対して対になるように存在する場合のみ環状ウィンドウの中心座標 (x_i, y_i) を求める。

環状ウィンドウの特徴として、次のようなことが挙げられる。

- 画像の傾きによる影響が少ない
- 環状ウィンドウの半径 r_1, r_2 を変化させることで、対象物体の大きさに依存せず、厳密に線対称でない対象物体に関しても、中心座標を求めることができる。

次に、得られた n 個の中心座標点列 (x_i, y_i) から中心軸を求めるために、回帰直線を求める。

$$y = A + Bx \quad (4.13)$$

ここで、 A は切片で B は傾きであり、式(4.14)(4.15)により求めることができる。

$$B = \frac{n \sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)}{n \sum_i x_i^2 - (\sum_i x_i)^2} \quad (4.14)$$

$$A = \frac{\sum_i y_i - B \sum_i x_i}{n} \quad (4.15)$$

また、環状ウィンドウによって求めた中心座標 (x_i, y_i) が、式(4.13)をどの程度、適合しているか調べるには、次式の R^2 を求めることにより知ることができ、 R^2 の値が1に近いほどは適合している。

$$R^2 = \frac{A \sum_i y_i + B \sum_i x_i y_i - (\sum_i y_i)^2/n}{\sum_i y_i^2 - (\sum_i y_i)^2/n} \quad (4.16)$$

Step4 中心軸を用いた画像の剛体変換・スケール変換

ここでは、上式で得られた中心軸を用いて、4方向から撮影された画像に対して、並行移動、回転、拡大縮小を行う。

平行移動に関しては、中心軸の中心座標が、画像の中心座標になるように、画像を平行移動させる。

回転角度に関しては、中心軸と x 軸の正方向との傾きを ϕ とすると、補正する角度 θ は、 $\theta = \frac{\pi}{2} - \phi$ となる。

拡大率に関しては、正面画像の対象物体に対する中心軸の長さ L とし、それ以外の画像に対する中心軸の長さを L_i とすると、拡大率は、 $s = \frac{L}{L_i}$ により求める。したがって、正投影図 (4 方向からの側面図) は、原画像の座標を (x, y) 、変換後の画像を (X, Y) とすると、次式に示すアフィン変換により生成することができる。ただし、回転の中心は、平行移動後の画像の中心座標とする。

$$\begin{bmatrix} X \\ Y \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.17)$$

4.4.2 実験結果と考察

上述の方法により、環状ウィンドウを用いて、図 4.12(b) から中心軸を求めたものを図 (c) に示す。このときの A, B, R^2 の各係数は、表 4.2 のようになった。

次に、実画像から正投影図を生成し、これを用いて仏塔の 3 次元形状モデルの生成する。実験対象となるモール仏塔は、新疆ウイグル自治区カシュガル郊外にある仏教遺跡であり、中国最西端にある仏教遺跡として非常に価値のあるものである。モール仏塔は、高さが 14.5m あり、約 1000 年前に建造され、西域仏教の典型的な様式である。以下に、モール仏塔の 3 次元形状モデルを生成する方法について述べる。

まず、図 4.13(a)~(d) に示す画像に対して、第 4.4.1 節で述べたとおり、対象物体の輪郭を抽出し対象物体の中心軸を求め、移動量、回転角度、拡大率を計算する。求めたパラメータを表 4.3 に示す。これらのパラメータをもとに画像を剛体変換・スケール

表 4.2 中心軸の直線パラメータ

B	A	R^2
-5.25	1002.14	0.75

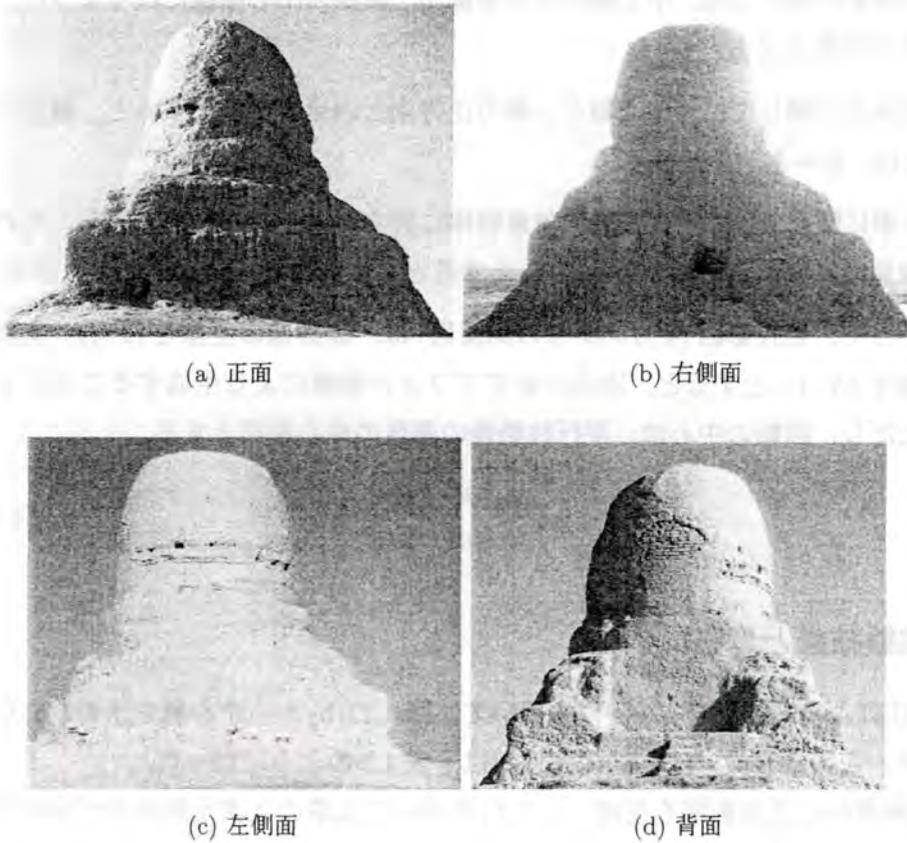


図 4.13 モール仏塔

変換を行い、正投影図を生成する。

次に、第4.3.3節で述べた方法により、視体積交差法に基づいて3次元形状モデルを生成する。このときのボクセル空間の大きさは、 $20.49[\text{m}] \times 21.15[\text{m}] \times 15.29[\text{m}]$ であり、各ボクセルの大きさは、 $9.400 \cdot 10^{-2}[\text{m}] \times 9.400 \cdot 10^{-2}[\text{m}] \times 9.400 \cdot 10^{-2}[\text{m}]$ である。

さらに、図4.13(a)~(d)に示すような4枚の実画像を使って3次元形状モデルにテクスチャマッピングを行う。ただし、この場合のテクスチャマッピングは、第4.3.4節で述べた対話的な処理を必要としない。なぜなら、正投影図は実画像を使って生成しているため、生成された3次元形状モデルを補正する必要がなく、3次元形状モデルを正

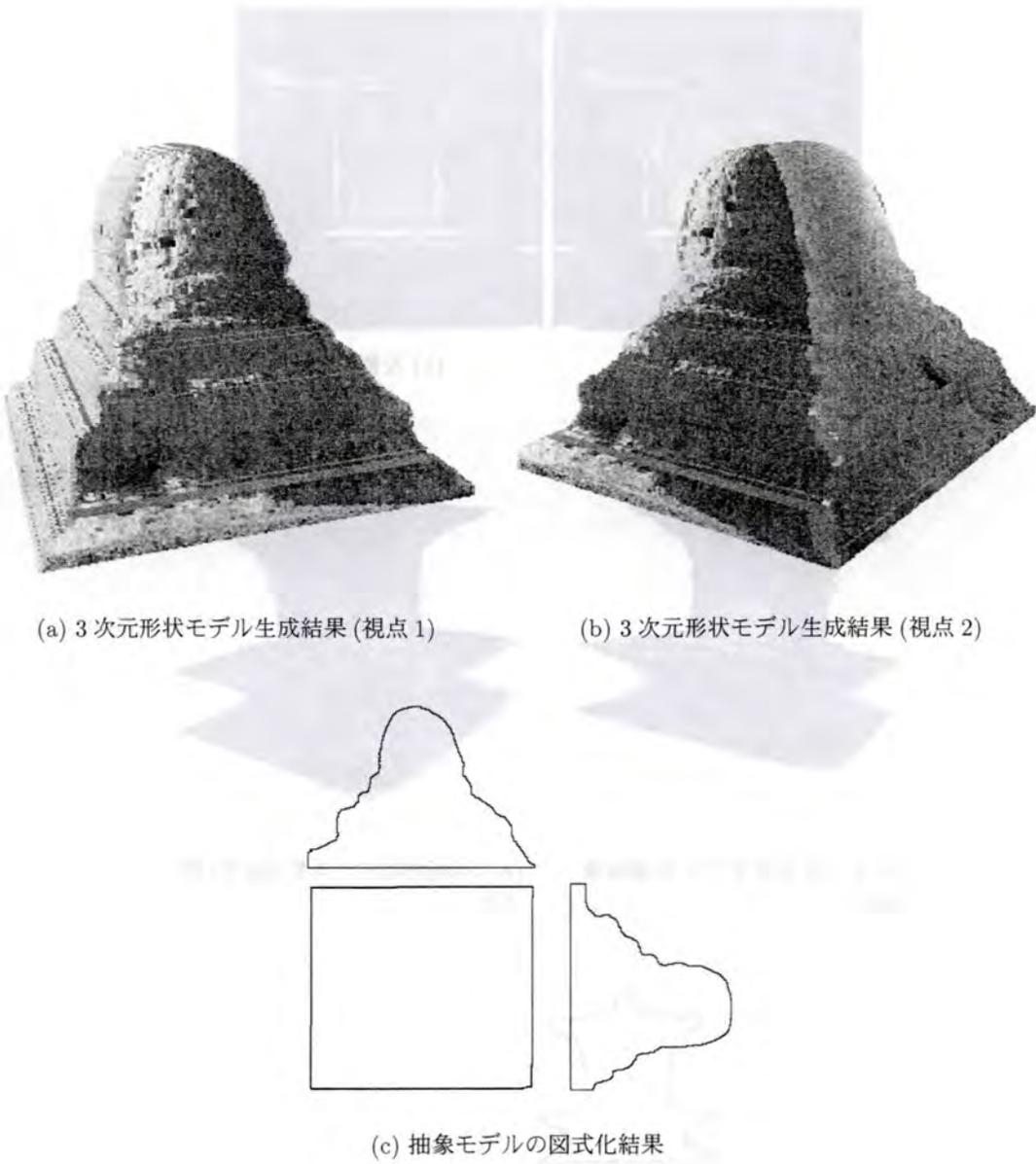
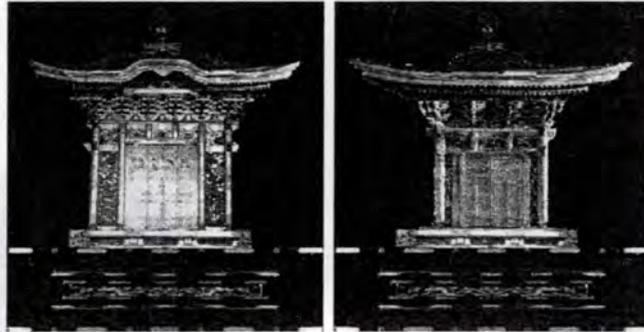
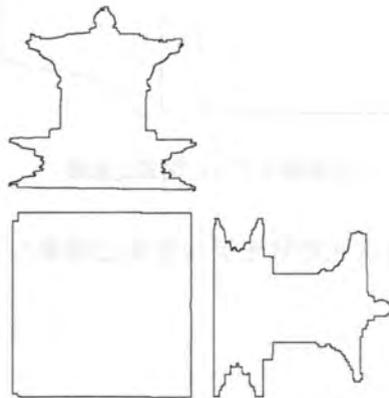


図 4.14 モール仏塔の3次元形状モデル生成と抽象モデルの図式化結果



(a) 正面

(b) 左側面

(c) 3次元形状モデル生成結果
(視点1)(d) 3次元形状モデル生成結果(視
点2)

(e) 抽象モデルの図式化結果

図 4.15 西本願寺御影堂内陣の須弥壇と厨子の3次元形状モデルの生成と抽象モデルの図式化結果

投影図に逆投影すれば、テクスチャマッピングが可能である。

上記の方法により、生成されたモール仏塔の3次元形状モデルを図4.13(a)(b)に示し、生成した3次元形状モデルから抽象モデルを図式化したもの(正投影図)を生成した結果を図4.14(c)に示す。

また、上記の手法を用いて、西本願寺御影堂内陣にある親鸞上人の木像を安置する須弥壇と厨子を対象として3次元形状モデルを生成した結果を図に4.15(c)(d)に示し、抽象モデルを図式化したもの(正投影図)を生成した結果を図4.15(e)に示す。

最後に、本手法と市販されている簡易写真測量システムとの比較を行う。Photomodeler や Image Modeler などの簡易写真測量システムでは、撮影された画像から対応する点を手動で取得する必要がある。また、住宅やビルなど多面体で近似できる対象に関しては、簡易に3次元形状モデルを生成することが可能である。しかし、対象物体の形状が複雑になるほど、精度良く3次元形状モデルを生成するためには、システムに与える対応点の数を増やす必要があり、全周の3次元形状モデルを生成するには、全周から撮影した多数の画像が必要である。

しかしながら、本手法では、画像間の対応点問題を扱う必要がなく、半自動的に3次元形状モデルの生成が可能である。また、全周の3次元形状モデルを生成するには、最低4方向から撮影した4枚の画像があれば可能である。

以上の結果より、屋外の対象に関して実画像から生成された正投影図を用いて、3次元形状モデルの生成が簡易に行うことができることが確認された。

表 4.3 剛体変換・スケール変換のパラメータ

画像	移動量 (x, y)	回転角度 (deg)	拡大率
正面画像	(5, 7)	0.78	1.00
右側面画像	(2, 4)	0.11	0.98
左側面画像	(34, 6)	-0.01	0.99
背面画像	(-2, 9)	0.50	1.02

4.5 まとめ

本章では、石窟寺院や仏塔、寺院を対象とした電子記録化方法への適用例について述べた。

ここでは、まず、抽象モデルを図式化した手書き立面図に視体積交差法を適用し、3次元形状モデルを生成することが可能であることを示した。また、画像情報と生成した3次元形状モデルを用いて、対話処理より3次元形状モデルを補正した。補正した3次元形状モデルと補正前の3次元形状モデルの精度比較を行い、本手法の有効性を確認した。それから、補正した3次元形状モデルを用いて、抽象モデルを更新することにより、シーン構造記述を行った。

さらに、抽象モデルを図式化したものが存在しない対象に関しても、実画像から正投影図(立面図)を生成する手法について述べ、生成した正投影図から3次元形状モデルを生成し、3次元形状モデルを用いて、抽象モデルを図式化した正投影図の生成を試みた。

第5章

シーン構造記述に基づく拡張現実感システムの構築

これまでは、実シーンの構造情報を仮想世界内に入力し、3次元形状モデルとしてシーンを表現する方法について述べてきた。本章では、仮想世界に入力された構造情報を実世界で表現する方法の一例として、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムの構築を目的とする。

5.1 はじめに

現在、博物館などにおいては、古典籍などの書籍を展示する場合には、書籍の保護や保存目的から、あるページを見開いた状態、もしくは表紙のみを展示している場合が多い。そのため、利用者が自由にページをめくって内容を閲覧することは難しい。このような問題に対しては、すでにデジタル化された書籍を計算機上で閲覧することにより解決する。しかし、一般的な文書ブラウザ(Acrobat Readerなど)では、書籍の内容を閲覧することは可能であるが、臨場感や材質感までを利用者に伝えることが難しい。

上記の問題を解決するために、利用者に対して、より自然でより直感的な書籍表示インタフェースの開発が必要となり、その解決法の1つとして、拡張現実感の技術を応用する。拡張現実感とは、実世界に対して、仮想世界の情報を融合するものであり、実

世界と仮想世界をシームレスに融合するために、幾何学的整合性⁶⁴⁾⁶⁵⁾や光学的整合性⁶⁶⁾⁶⁷⁾を解決する研究がなされている。拡張現実感を実現するための表示デバイスとして、HMD(Head Mounted Display)が、一般的に利用されているが、実世界に対して、仮想物体を表示する際の時間的なずれや、表示画面が小さく、解像度が低いため、利用者に違和感を与えるなどの問題点がある。このため、近年では、プロジェクタを用いて、拡張現実感を実現する研究⁶⁸⁾⁶⁹⁾が進められている。

ここでは、拡張現実感の技術における幾何学整合性を解決し、プロジェクタと、投影面として和紙で作られた本を利用して、書籍表示インタフェースを構築する方法について述べる。

5.2 概要

本システムは、キャリブレーション処理、幾何補正処理、ページ認識処理、画像表示処理の4つの処理で構成されている。キャリブレーション処理では、プロジェクタ座標系、画像座標系、世界座標系の3つの座標系のキャリブレーションを行う。幾何補正処理では、投影面が非平面であるため、キャリブレーション処理で得られたパラメータと投影面の3次元座標を取得し、投影される画像を補正するための構造情報を求める。ページ認識処理では、投影面に貼られたバーコードを利用して、ページ番号を認識する。画像表示処理では、ページ認識結果と得られた構造情報をもとに、補正画像を生成し、それを投影面に表示する。本システムの特徴は、実世界の本をめくるように、デジタル化された書籍を閲覧することが可能となり、また、本実験で用いた書籍画像は、実際には和紙に描かれているため、その材質感も表現することが可能となることである。

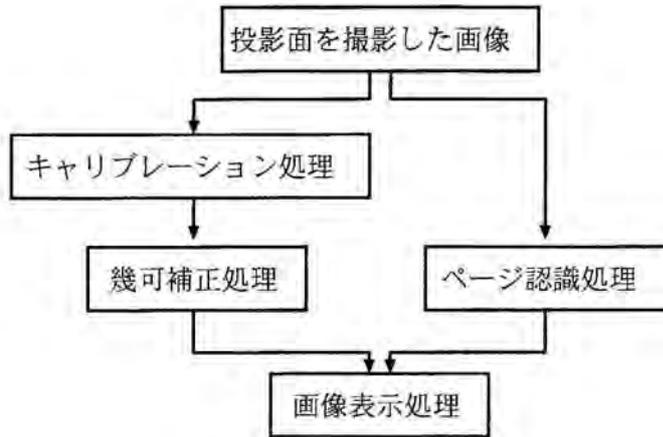


図 5.1 処理の流れ

5.3 書籍表示インタフェースの構築

5.3.1 本システムの条件設定と処理概要

本システムは、投影面として、白紙の書籍を利用する。さらに、書籍の各ページに、あらかじめページ番号(ページ番号を表すバーコード)を割り当てておく。利用者が、書籍のページをめくると、蓄積している書籍画像からそのページ番号と一致する画像を書籍に表示する。このシステムを実現するために、次のような条件設定を行う。

条件 1: 投影面である書籍の位置は固定。

条件 2: プロジェクタとカメラの位置は固定。

条件 1は、画像を投影するためのプロジェクタと、投影面である書籍の投影距離が、変化した場合に、書籍に投影した画像がぼけることを防ぐためである。

条件 2は、本システムでは、後述するキャリブレーション処理を行う必要がある。そのため、プロジェクタとカメラの位置を固定し、キャリブレーション処理を一度だけで済ませる。

本システムは、図 5.1 に示すように、以下の 4 つの処理により構成されている。

- (1) キャリブレーション処理
- (2) 幾何補正処理
- (3) ページ認識処理
- (4) 画像表示処理

上記(2)の幾何補正に関しては、投影面として、スケッチブックやリングファイル式の本を用いた場合には、投影面は平らな形状であるため、投影する書籍画像を補正する必要はない。

しかしながら、本システムでは、書籍画像を実世界に近い形で表現することを目的としている。そのために、白紙の書籍を投影面(以後、ブックスクリーンと呼ぶことにする)として用いる。ブックスクリーンは非平面であるため、投影する書籍画像の幾何補正を行う必要がある。下記に各処理部分について述べる。

5.3.2 キャリブレーション処理とブックスクリーン表面の3次元座標の取得

本システムでは、プロジェクタ座標、画像座標、および世界座標系のキャリブレーションと、ブックスクリーン表面の3次元座標を求める必要がある。ブックスクリーン表面の3次元座標を取得する方法として、陰影情報⁷⁰⁾やグレイコードパターン⁷¹⁾を利用したものがある。文献⁷⁰⁾では、イメージスキャナーの特性を利用しているため、本システムに適用することは難しい。また、文献⁷¹⁾では、表面の3次元座標を取得するまでに時間を要するため、実時間性を重視する本システムでは適用が難しい。

そこで、本システムでは図5.2のキャリブレーションパターンを用いて、以下に示す方法により、キャリブレーション、およびブックスクリーン表面の3次元座標を取得する。

Step1: ステレオ法を適用するため、テクスチャの無いブックスクリーンに図5.2のキャリブレーションパターンを投影する。ここでは、3眼式カメラを用いたマルチベースラインステレオ法¹³⁾¹⁴⁾を適用し、3次元座標を取得する。本システムで用いた

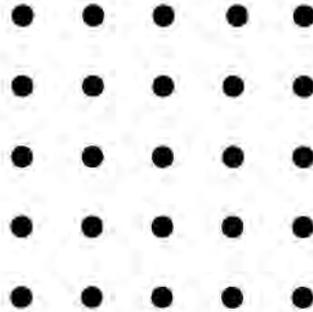


図 5.2 キャリブレーションパターン

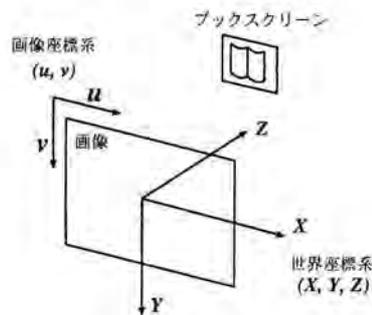


図 5.3 本システムにおける画像座標系と世界座標系の関係

マルチベースラインステレオシステムの特徴は、光学歪みの補正はされており、図 5.3 に示すように画像座標系 (u, v) と世界座標系 (X, Y, Z) も、システムとしてキャリブレーション済みである。また、入力される画像解像度が 320×240 の場合、約 10fps の速度で距離画像の生成が可能である。

Step2: Step1 では、ブックスクリーン表面の 3次元座標を取得する際に、すでに画像座標系と世界座標系のキャリブレーションが行われている。このため、ここでは、プロジェクタ座標系と世界座標系のキャリブレーションを行う。それには、プロジェクタ座標 (u_p, v_p) と、キャリブレーションパターンがブックスクリーン上に投影された点の 3次元座標 (X, Y, Z) との対応関係を求める。この対応関係を式

(5.1) に示す射影行列 P を用いて求める.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^T & p_{14} \\ \mathbf{p}_2^T & p_{24} \\ \mathbf{p}_3^T & p_{34} \end{bmatrix} \quad (5.1)$$

$$m = [u_p, v_p, 1]^T, \mathbf{M} = [X, Y, Z, 1]^T \quad (5.2)$$

とすると、次の関係が成り立つ.

$$sm = PM \quad (5.3)$$

ただし、 s はスカラーである. 式 (5.3) を展開して、 P の要素に関する2つの線形方程式

$$\begin{cases} \mathbf{p}_1^T M_i - u_{p_i} \mathbf{p}_3^T M_i + p_{14} - u_{p_i} p_{34} = 0 \\ \mathbf{p}_2^T M_i - v_{p_i} \mathbf{p}_3^T M_i + p_{24} - v_{p_i} p_{34} = 0 \end{cases} \quad (5.4)$$

が成立する. 射影行列 P の算出は、キャリブレーションパターンが、ブックスクリン上に投影された点の3次元座標 $M_i = [X, Y, Z]^T (i = 1, \dots, n)$ と、プロジェクタ座標 $m_i = [u_{p_i}, v_{p_i}]^T (i = 1, \dots, n)$ の n 個の対応する点 (ただし、 $n \geq 6$) を用いて、式 (5.4) を解く. 得られた解を初期値として、式 (5.5) の E を最小にする評価関数を非線形最小2乗法によって解くことが可能である.

$$E = \sum_{i=1}^n \left(\frac{\mathbf{p}_1^T M_i + p_{14}}{\mathbf{p}_3^T M_i + p_{34}} - u_{p_i} \right)^2 + \sum_{i=1}^n \left(\frac{\mathbf{p}_2^T M_i + p_{24}}{\mathbf{p}_3^T M_i + p_{34}} - v_{p_i} \right)^2 \quad (5.5)$$

プロジェクタ、および3眼式カメラの位置は、固定されているため、上記のキャリブレーション処理は一度だけ行えばよい.

5.3.3 幾何補正処理

投影面であるブックスクリンが、ページをめくる度に形状が変化する. そのため、ページをめくる度に、投影する画像の幾何補正を行う必要がある. すでに、5.3.2節で示したように、プロジェクタ座標系と画像座標系、および世界座標系のキャリブレーション処理は終了している. ここでは、5.3.2節で求めたパラメータ P を用いて、投影する画像の幾何補正を行う方法について述べる.

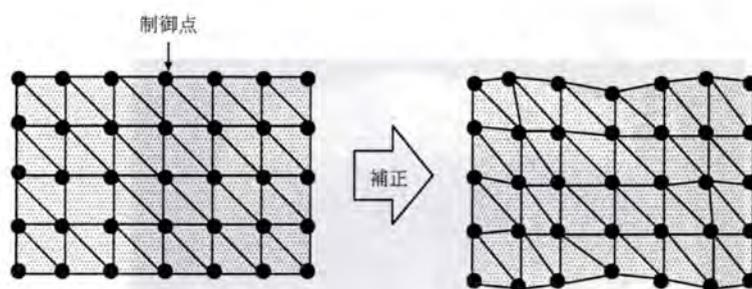


図 5.4 画像の幾何補正

Step1: 投影する画像において、図 5.4 に示すような制御点 (Control point) を設定し、ポリゴン近似を行う。この制御点の座標を変更し、各ポリゴンにテクスチャマップピングを行うことにより、投影する画像を補正する。

Step2: ブックスクリーンは、各ページが綴じられている。そのため、それらを外し、歪みのない状態にしてから、各ページに画像の歪み補正用のマーカを複数つける。マーカのつけ方は、下記の通りである。まず、ブックスクリーンの2ページ分(見開いた状態)を投影基準面に対して、水平に置く。次に、画像上の制御点をブックスクリーンの2ページ分に投影する。それから、投影された制御点上に、マーカをつけていく。これは、ブックスクリーンが、投影基準面に対して、水平に置かれているため、書籍画像をブックスクリーンに投影しても歪みが生じていない状態だと考えられる。これにより、画像上の制御点が、正しくブックスクリーン上に投影されるべき位置を、マーカによって知ることができる。実際の幾何補正は、ブックスクリーン上のマーカに、制御点が投影するように、制御点の座標を決定すればよいことになる。

上記のマーカづけをブックスクリーンの各ページに対して行い、再度、各ページを綴じて、それをブックスクリーンとする。実際に用いたブックスクリーンを図5に示す。

Step3: ブックスクリーンに配置された各マーカの3次元座標 (X, Y, Z) を 5.3.2 節で述べた方法により求める。次に、キャリブレーション処理で求めたパラメータ P を

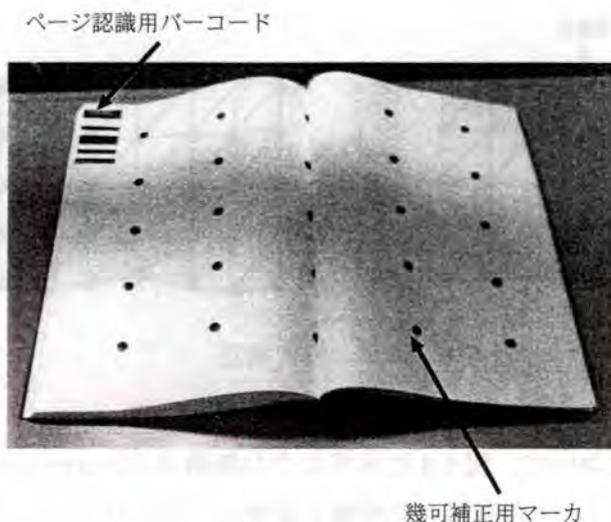


図 5.5 ブックスクリーン

用いて、式(5.3)よりプロジェクタ座標 (u_p, v_p) を求め、この座標を画像上の制御点座標とする。

Step4: Step1 から Step2 までの処理は、プロジェクタおよび3眼式カメラが固定されているため、一度行えばよい。実際の幾何補正処理では、ページをめくる度に、Step3を繰り返すことより実現が可能である。

マーカを用いた部分計測により、投影される画像の幾何補正を行う理由は以下の通りである。

ブックスクリーンを全点計測し、幾何補正を行う場合には、より正確な補正が可能となる。しかしながら、3次元座標の取得、および補正画像の生成には、処理コストがかかり、実用的ではないため、ブックスクリーン表面を部分計測する必要がある。

部分計測を行う方法の1つに、マルチスリットパターンを投影する方法が考えられる。マルチスリットパターンを用いれば、高速でかつ正確な3次元計測が可能である。しかしながら、本システムは、利用者に実物の本を見ているかのように書籍画像を表示することを目的としている。本システムで、マルチスリットパターンを使って、安

定した3次元計測を行う場合には、書籍画像の投影とパターン光の投影を交互に表示を繰り返す必要がある。このため、利用者に違和感を与えてしまうという問題がある。

一方、マーカを用いて、ステレオ法により3次元計測する方法が考えられる。この場合には、マーカが妨げになり、書籍画像の一部を隠蔽してしまうという問題がある。しかしながら、マーカを工夫(例えば、近赤外線にのみ、反射特性をもつ顔料の使用など)することにより、この問題は技術的に改善させることが可能である。このため、本システムではマーカとステレオ法を用いた3次元計測方法を適用した。

5.3.4 ページ認識処理

本システムではブックスクリーンを用いて、ページをめくるごとにページに対応した画像を表示させる必要がある。このため、ブックスクリーンにページ番号(もしくは、ページ番号を表現する図形)をつける必要がある。

図形を使ってページ番号を表現する方法として、バーコードを用いたページ認識方法⁷⁾がある。使用される代表的なバーコードとして、物流商品用にJIS化されているITF(Interleaved Two of Five)や2次元マトリックス⁷²⁾がある。バーコードは、認識率が高く、認識処理が速く、角度変化に柔軟に対応が可能である。2次元マトリクスなどの2次元バーコードの特徴は、ITFなどの1次元バーコードと比べて、より多くの情報量を表現することができる。また、2次元バーコードは、バーコードの大きさを縮小することができ、認識誤りのチェックや全方位からの読み取りが可能である。

本システムでは、ページ認識のために2つの理由から1次元バーコードを採用する。1つは、ページ認識のみにバーコードを使用するため、バーコードに表現する情報は少なくても済むからである。もう1つは、バーコードを撮影するためのCCDカメラは固定されているため、バーコードの変形は少ないからである。ここで用いるバーコードの例を図5.6に示す。図中でdata codeは8ビット分あり、各ビットに黒("1")か白("0")を与える。これにより、0から255までのページを表現することが可能となる。図中では、"45"の値が表現されている。また、投影する対象が255ページを超えるものに対しては、data codeのビット数を増やすことにより、容易に対応することができる。このパターンの特徴は、パターンの作成が容易であり、ビットパターン自体がページ番号



図 5.7 システム構成

5.4 実装と実験結果

5.4.1 システム環境と実装

書籍表示インタフェースを実現するために、図5.7のように液晶プロジェクタ1台(表5.1参照)、CCDカメラ1台、3眼式カメラ1台からなるシステムを構築した。また、表5.2に示すようなページ認識・画像表示するための計算機と、表5.3に示すようなブックスクリーン表面の構造情報を取得するための計算機、合計2台のPCを用いて分散処理を行った。

投影するブックスクリーンは、図5.5に示すような白色の和紙で作成されたものを用いた。これは、本システムを用いて投影する書籍画像の原本も和紙で描かれたものであり、より実世界に近い表示を行うためである。

また、3眼式カメラ、およびプロジェクタとブックスクリーンの距離は、0.7mに設定した。これは、本システムで使用するマルチベースラインステレオ法の距離計測分解(表5.4参照)とプロジェクタの投影距離を考慮したためである。

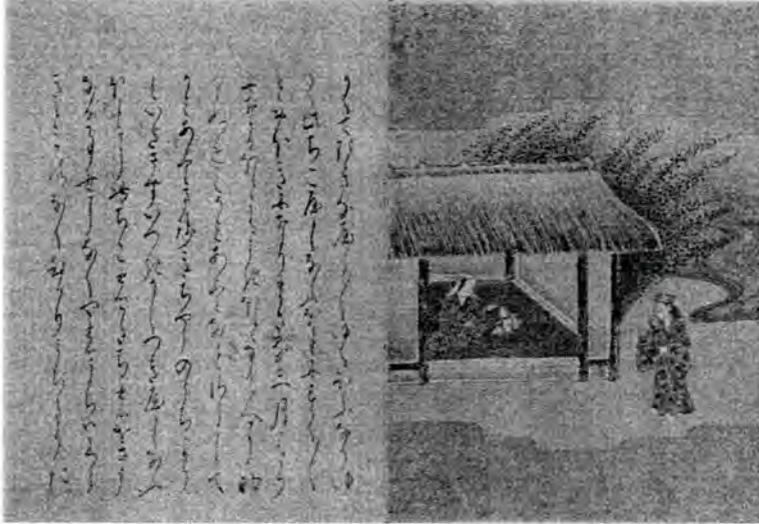


図 5.8 写字台文庫の一例 (竹取物語抄)

本システムを用いて、図5.8に示すような古典籍の表示を行った。本システムの実証例として用いた古典籍は、歴代の西本願寺歴代宗主が収集・伝持してきた写字台文庫と呼ばれる書籍である。龍谷大学大宮図書館では、現在、約3万冊の写字台文庫が保管され、そのデジタル化を進めている。写字台文庫のデジタル化の際に、見開いた状態で撮影されているため、1枚の画像に文庫の2ページ分が収録されている。そのため、ここでは文庫の2ページ分を1ページと数え、ページ認識のためのバーコードは、ブックスクリーンを見開いた状態で左上に貼りつけることとする。

写字台文庫の一例として、図5.8に示す文庫は、1783年に出版された「竹取物語抄」であり、最古の伝奇物語である「竹取物語」の注釈書である。

表 5.1 液晶プロジェクタ

型名	NEC ViewLight LT155
最大表示解像度	UXGA 1600 × 1200
輝度	1600ANSI ルーメン

本システムでは、まず、利用者が蓄積されている写字台文庫のタイトルを選択する。次に、システムは選択された写字台文庫をメモリに蓄積する。そして、利用者が、バーコードが貼られたブックスクリーンのページをめくると、ページ認識・画像表示用PCが、ブックスクリーンに貼られたバーコードからページを認識する。さらに、構造情報の取得用PCは、ページ認識・画像表示用PCからページ認識結果を受信し、書籍画像の幾何補正を行うかどうかを判定し、幾何補正を行う場合には、ブックスクリーンに配置されたマーカを検出し、各マーカの3次元座標を取得する。取得された各マーカの3次元座標点とキャリブレーション処理により求めたパラメータを用いて、マーカ上に投影されるべき画像中の制御点の座標(プロジェクタ座標)を求める。求められた制御点の座標値は、ページ認識・画像表示用PCに伝送される。最後に、ページ認識・画像表示用PCは、ページ認識結果に対応する画像を選択し、幾何情報の取得用PCから伝送された制御点座標をもとにして、画像の幾何補正を行い、ブックスクリーンに画像を投影する。

以下に、ページ認識、画像の幾何補正、および画像表示のそれぞれの結果を示す。

5.4.2 ページ認識結果

ページ認識に要した時間は、平均0.13秒であった。また、ページ認識の評価は、次のようにして行った。ページが正しく認識された場合には1、バーコードが検出できなかったため、認識できなかった場合には0、誤認識された場合には-1とする。利用者

表 5.2 ページ認識・画像表示用PC

OS	Windows2000
CPU	PentiumIII 1GHz
Memory	512MB
Capture Board	Matrox MeteorII
Graphics Board	3Dlabs OxygenVX1 32MB
Camera	SONY CCM-DS250
Network	100base-TX

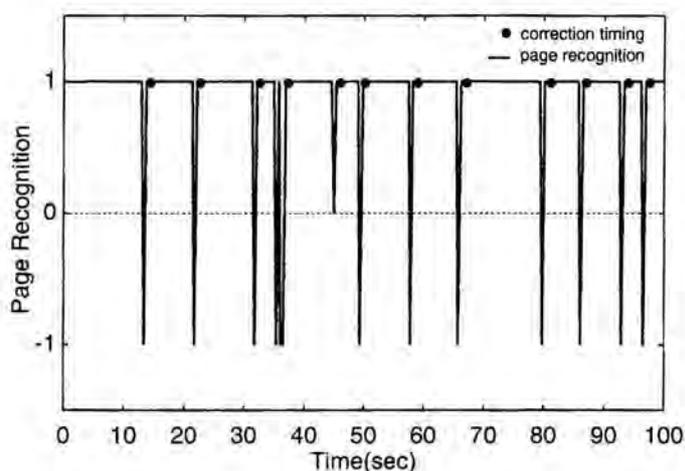


図 5.9 時間に対するページ認識結果と幾何補正のタイミング

が、100秒間、書籍のページをめくったときのページ認識結果を図5.9に示す。この図では、利用者がページをめくる過程(利用者が、ブックスクリーンに手を掛けてからめくり終わった直後までの一連の動作)において、誤認識、もしくはバーコードが検出できなかった場合が生じた。しかしながら、それ以外の場合には、ほぼ正しくページが認識されていることがわかる。

表 5.3 構造情報の取得用 PC

OS	Windows98
CPU	PentiumIII 500MHz
Memory	768 MB
Capture Board	Matrox Meteor
Camera	PointGrey Triclops
Network	100base-TX

5.4.3 幾何補正結果

幾何補正処理から得られた制御点の補正情報をもとに、補正画像を生成した結果を図 5.10 に示す。図 5.10(b) は、図 5.10(a) のテストパターンを補正した画像である。また、補正した画像を実際にブックスクリーンに投影した結果を図 5.10(c) に示す。この図からは、投影画像の歪みが、本の中央部分よりも、むしろ、本の周辺部分の方に歪みが多く生じていることが確認できる。また、本の中央部分の歪みは、プロジェクタに対して、ブックスクリーンは垂直に設置しているため、グリッドの横幅を変化させることで補正が可能であることを示している。

ここでは、画像上の制御点およびブックスクリーンのマーカの数、それぞれ 25 点とした。マーカ数に関しては、マーカ数が増加するにつれ、投影画像の補正がより正確に行うことができる反面、処理コスト、および投影画像が見づらくなるという問題がある。また、投影画像を均一に補正するために、左右対象となるようにマーカを配置する必要がある。ここでは、ブックスクリーンの縦じ部分を中心として、ブックスクリーンの左右のページに均等にマーカを配置する。そのためには、 $(2n + 1)^2$ 個のマーカが必要となる。そこで、 n の値を変化させ、試行実験を行ったところ、処理コスト、投影画像の補正結果、投影画像の見やすさの面を考慮に入れると、 $n = 2$ ときが良好な結果が得られたため、本実験ではマーカ数を 25 点に設定した。

図 (c) からわかるように、ブックスクリーンに配置したマーカ上に正しく、テストパターンが投影されていることが確認できる。また、1 枚の補正画像を生成するのに要した時間は平均 0.63 秒であった。

また、ページが誤認識、もしくはバーコードが検出できなかった場合には、ブック

表 5.4 距離計測分解能

計測範囲 (meters)	計測誤差
14.7	20%
7.3	10%
3.6	5%
0.7	1%

スクリーンに配置されたマーカの検出、3次元座標の取得、および画像の幾何補正は行わないものとした。この理由は、誤認識およびバーコードが検出できなかった際には、利用者がページをめくっている過程であることが考えられる。この場合には、マーカの検出が難しく、3次元座標の取得できないからである。

しかしながら、ページが誤認識もしくはバーコードが正しく検出されなかったかどうかをシステムが自動的に判別することは難しい。そこで、本システムでは、5.4.2節に示したページ認識結果から、誤認識もしくはバーコードが検出できない場合が頻繁に起こるのは、利用者がページをめくっている過程で生じていることに着目する。利用者が、ページをめくる動作を行っているかを自動的に判別できれば、正しくページが認識されていなかった箇所をある程度特定することが可能となる。そのため、以下のような処理を行っている。利用者が、ページをめくっている動作を検出するためにフレーム間差分を用いている。フレーム間差分において、ある一定以上の動きがあった場合には、利用者がページをめくっているとシステムが判断する。ただし、ページをめくっている過程において、利用者は、ページをめくり始めてからページをめくり終るまでの一連の動作を静止しないものとする。このとき、ページの誤認識もしくはバーコードの検出が難しいとシステムが判断し、マーカの検出、3次元座標の取得、画像の幾何補正処理は行わないものとする。

幾何補正処理のタイミングに関しては、ユーザがページをめくり終った直後(図5.9の●箇所)に一度だけ行うものとした。これは、本システムでは、幾何補正処理を行っている最中は、画像投影が行うことができないためである。この理由については、5.5節で述べる。

5.4.4 画像表示結果

最後に、本システムを用いて、写字台文庫を表示した結果を図5.11に示す。この図に示すように、ブックスクリーンの形状に合わせて投影される画像が、幾何補正されており、ページ番号に対応した画像が、正しく表示されていることが確認できた。

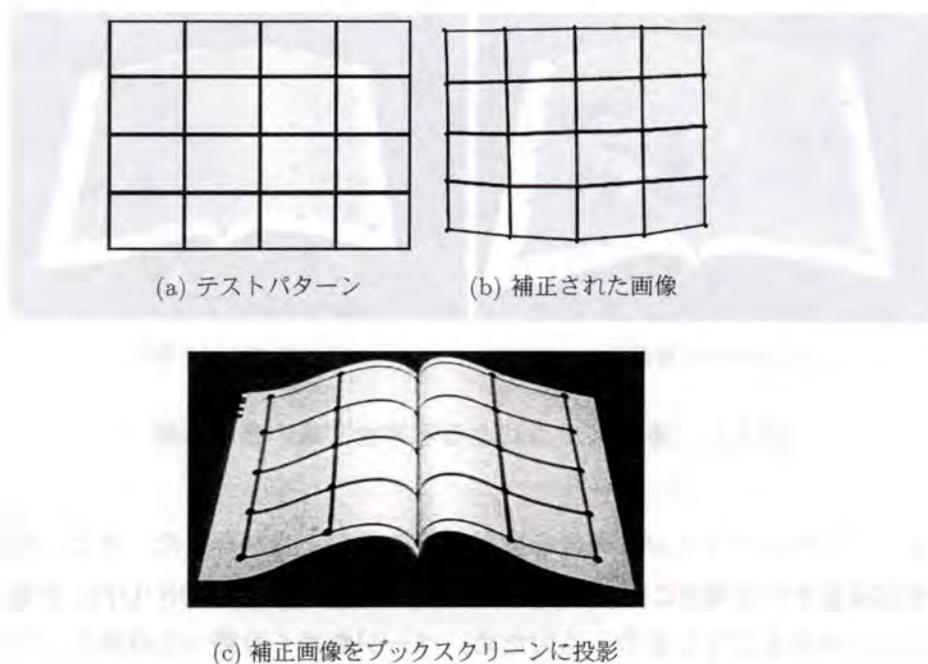
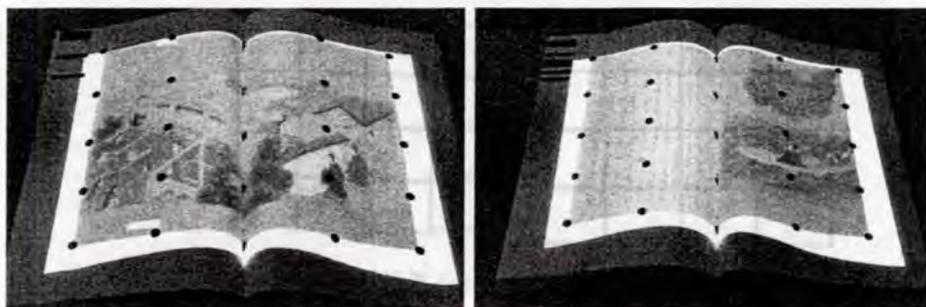


図 5.10 画像の幾何補正結果

5.5 まとめ

ここでは、仮想世界に入力された構造情報を実世界で表現する方法の一例として、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムの構築を行った。本システムでは、拡張現実感における幾何学的整合性を解決し、プロジェクタを用いて、書籍表示インタフェースを構築し、古典籍画像の表示を行った。投影面として和紙で作られた本を利用することによって、古典籍画像を実世界に近い形で表示することが可能となった。また、試作システムを構築することによって、本システムの有効性を示した。以下に、試作システムで明らかになった問題点とその改善方法について検討を行う。

マーカの3次元座標を取得する際に、画像の表示を行うとブックスクリーンに配置したマーカを安定して検出することができなかった。そのため、幾何補正処理が終了



(a) 8 ページを表示

(b) 45 ページを表示

図 5.11 本システムによる写字台文庫の表示結果

するまで、ブックスクリーンに画像を投影することができなかった。また、連続して画像の幾何補正を行う場合には、古典籍画像の切替(画像の表示 ON/OFF)を繰り返すため、ちらつきが生じてしまう。このため、ページをめくり終わった直後に一度だけ画像の幾何補正を行うことにした。これは、今回の実験において、マーカ検出で用いた3眼式カメラがモノクロカメラであったためである。カラー画像が取得できるカメラを利用すれば、色情報を用いて、マーカと投影された画像との分離が実現可能となる。

また、図 5.11 の実験結果に示す通り、ブックスクリーンに配置されたマーカにより、投影画像の一部分が隠蔽されてしまうという問題が生じる。これに関しては、近赤外線にのみ、反射特性をもつ顔料のマーカを利用⁷⁴⁾することにより、実現が可能である。

第6章

結論

本章では、本研究で得られた研究成果をまとめ、今後の課題を示す。

6.1 まとめ

本論文では、実シーンを対象とした構造情報の取得と記述、および表現に関する研究を行った。本研究のアプローチは、以下の通りである。まず、実シーンから取得される構造情報と抽象モデルから取得される構造情報を統合することにより、3次元形状モデルの生成を行った。また、生成した3次元形状モデルを用いて抽象モデルを更新することにより、シーンの構造記述を行った。さらに、実シーンから取得された構造情報を表現するために、実世界、仮想世界、利用者の相互関係に基づいた拡張現実感システムの構築を行った。本論文で得られた研究成果を各章ごとに順にまとめると以下ようになる。

第2章では、実シーンから構造情報を取得・記述・表現するための一般的な手法について述べ、本研究で提案する構造情報の取得と記述方法について述べた。さらに、CV、CG、HI、MRの研究分野において、実世界、仮想世界、利用者の相互関係を考察し、これらの相互関係に基づく拡張現実感システムの構築方法について述べた。

第3章では、抽象モデルを図式化した平面図と、実シーンから得られる構造情報をもとに、室内シーンの構造情報の記述と3次元形状モデルの生成方法について述べた。平面図からでは、室内の形状を捉うことが可能であるが、室内にある対象物体の構造

情報を得ることができない。そこで、ステレオ法を用いて室内にある複数の対象物体の構造情報を取得した。そして、平面図から得られた構造情報と実シーンから得られた構造情報を統合することにより、室内全体の3次元形状モデルを生成した。また、生成した3次元形状モデルを用いて、室内にある対象物体の構造情報を抽象モデルに加えることで、抽象モデルを更新し、シーン構造記述を行った。

第4章では、抽象モデルを図式化した手書き立面図と実画像をもとに、石窟や仏塔の3次元形状モデルを生成し、構造記述を行った。ここでは、立面図の各面を1つの領域として扱い、視体積交差法を導入することにより、立面図の歪みに関して柔軟に対応できる3次元形状モデルの生成方法について述べた。また、手書き立面図から生成した3次元形状モデルは必ずしも実世界の対象とは一致しないという問題点があった。この問題を解決するために、立面図から生成した3次元形状モデルと実画像を用いて、対話的な処理により、3次元形状モデルを補正した。そして、実世界に近い3次元形状モデルの生成が可能であることを示した。また、補正した3次元形状モデルを用いて抽象モデルを更新し、シーン構造記述を行った。

第5章では、実シーンから得られる構造情報と画像情報を用いて、実シーンの構造情報を表現する方法について述べた。その一例として、プロジェクタと書籍画像、ブックスクリーンを利用した書籍画像表示インタフェースシステムの構築方法について述べた。開発したシステムを用いて、実世界の本をめくるように、デジタル化された書籍画像を閲覧することが可能となった。

以上のことから、抽象モデルを導入することにより、複雑な実シーンに対する構造情報の取得と記述ができ、3次元形状モデルを生成することが可能であることを示した。

さらに、シーン構造情報の表現については、実世界、仮想世界、利用者の相互関係に基づいたヒューマンインタフェースの構築を行った。ここでは、拡張現実感の技術を応用し、利用者に対して、より自然で直感的な画像表示インタフェースを構築し、実世界と仮想世界を融合するときの技術的課題の1つである幾何学的整合性の問題を改善した。

6.2 今後の課題

本論文で提案した抽象モデルは、対象に関する構造情報は保存されているが、実シーンから取得される構造情報と抽象モデルから取得される構造情報は一致するとは限らないという問題に対して、本研究では、以下に示すアプローチを行った。平面図を用いた実験において、実世界のカメラ位置と平面図上の仮想的なカメラ位置を一致させる必要があった。そのため、平面図と実世界のスケーリングを行い、実世界のカメラ位置から壁までの距離を実測し、この実測値をもとにして、平面図上に仮想的なカメラ位置を決定した。手書き立面図を用いた実験においては、手書き立面図から生成した3次元形状モデルを2次元平面に投影し、投影画像と実画像を比較し、対話処理を利用して、3次元形状モデルを補正した。しかしながら、いずれの手法を用いた場合にもこれらの構造情報を完全に一致させるまでには至っていない。

また、本研究では抽象モデルを図式化した平面図や手書き立面図に対して、実シーンから構造情報を取得・記述し、3次元形状モデルの生成・表現を行ってきたが、さらに、複数の抽象モデルを図式化したもの(鳥瞰図、市街図など)を利用し、実シーンから対象物体の認識や理解する研究に発展させていく必要がある。

シーン構造情報を用いた拡張現実感システムの構築においては、ブックスクリーンに和紙を用いたが、その他の材質(絹、石、木等)のものに関しては、ブックスクリーンを交換し、再度、キャリブレーションを行う必要がある。投影面の材質に関しての反射特性を考慮に入れる必要がある。また、実世界の本の良さは、利用者が、実際に本を手にとって、内容を手軽に閲覧することであるが、現在のシステムでは、各装置のシステムの制約上、本を固定しているため、このようなことを実現するのが難しい。今後は、以上で挙げた課題についての検討をすすめていく。

謝 辞

本研究を遂行するにあたり、終始懇切なご指導を賜った龍谷大学理工学部 岡田 至弘教授に心から感謝致します。

また、本論文の作成にあたり貴重なご助言を頂きました龍谷大学理工学部 小淵 洋一教授、龍谷大学理工学部 阿部 宏尹教授、龍谷大学理工学部 有木 康雄教授、龍谷大学理工学部 杉田 繁治教授に厚くお礼を申し上げます。

本研究を遂行するにあたり、有意義なご意見を頂きました龍谷大学文学部 百済康義教授、龍谷大学理工学部 渡辺 靖彦講師、藤井 大輔実験講師に感謝致します。

また、本研究のために、貴重なご意見、ご配慮、ご支援を頂きました龍谷大学理工学研究科 岩成 英一氏、村山 健二氏をはじめ、岡田研究室の皆様に感謝します。

最後に、今日に至るまで、日々励まし、支援してくれた父母、友人に感謝いたします。

参考文献

- 1) デビッド・マー, “ビジョン-視覚の計算理論と脳内表現”, 産業図書, 1990.
- 2) Mansssh Agrawala, Ravi Ramamoorthi, Alan Heirich, Laurent Moll, “Efficient Image-Based Methods for Rendering Soft Shadows”, In SIGGRAPH2000, pp.375-384.
- 3) Tom Lokovic, “Deep Shadow Maps”, In SIGGRAPH2000, pp.385-392.
- 4) T. Kanade, P.J. Narayanan, and P. Rander, “Virtualized Reality: Concepts and Early Results”, IEEE Workshop on the Representation of Visual Scenes, pp.69-76, June, 1995.
- 5) I.Kitahara, H.Saito, S.Akimichi, T.Onno, Y.Ohta, T.Kanade, “Large-scale Virtualized Reality”, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2001) -Technical Sketches-, (2001)
- 6) Milgram, P. and Kishino, F., “A Taxonomy of Mixed Reality Visual Displays”, IEICE Transactions on Information Systems, Vol E77-D, No 12, December 1994.
- 7) 小池 英樹, 小林 貴訓, 佐藤 洋一: “机型実世界指向システムにおける紙と電子情報の統合および手指による実時間インタラクションの実現”, 情報処理学会論文誌, Vol.42, No.3, pp.577-585, 2001.

- 8) C.Tomasi and T.Kanade, "Shape and motion from image streams under orthography: a factorization method", *International Journal of Computer Vision*, vol.9, No.2, pp.137-154, 1992.
- 9) 金出 武雄, C ポールマン, 森田 俊彦, "因子分解法による物体形状とカメラ運動の復元", *信学論 D-II*, vol.J76-D-II, No.8, pp.1497-1505, 1993
- 10) T. Morita and T. Kanade, "A Sequential factorization method for recovering shape and motion from image streams", *Technical Report, CMU-CS-94-158*, Carnegie Mellon University, 1994
- 11) J.Costeria and T.Kanade, "A Multi-body factorization method for motion analysis", *International Conference on Computer Vision*, pp.1071-1076, 1995.
- 12) 千葉 直樹, 金出 武雄, "ビデオ映像からの3次元モデル化 -線特徴への拡張-", 第3回知能情報メディアシンポジウム, pp.41-46, 1997.
- 13) 奥富 正敏, 金出 武雄, "複数の基線長を利用したステレオマッチング", *信学論 (D-II)*, vol.J75, no.8, pp.1317-1327, 1992.
- 14) Don Murray and Cullen Jennings, "Stereo vision based mapping for a mobile robot, " *In Proc. IEEE Conf. on Robotics and Automation*, 1997, May
- 15) 服部 寛, "複数視点画像からの3次元形状復元 -局所アフィン変形ウィンドウによるステレオマッチング-, " *情報処理学会*, pp.33-40, 1998.
- 16) M. Okutomi, T. Kanade, "A stereo matching algorithm with an adaptive window: theory and experiment," *PAMI*, vol.16, no.9, pp.920-932, 1995.
- 17) 山口 証, 高地 和隆, 井口 征士, "石像の3次元計測のためのステレオ対応付け, " *電子情報通信学会論文誌, D-II* Vol.J81, No.4, pp.716-725, 1998.
- 18) Stan Birchfield, Carlo Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *Processings of the IEEE International Conference on Computer Vision*, Bombay, India, 1998.

- 19) 内藤 貴志, 木村 好克, “ステレオビジョンの開発,” 豊田中央研究所 R&D レビュー Vol.31, No.2, pp.39-49, 1996.
- 20) Ingemar J. Cox, Sunita Hingorani, Bruce M. Maggs and Satish B. Rao, “Stereo without regularization,” NEC Research Insitute 4 Independence Way Princeton, 1992.
- 21) 関 靖夫, “セグメントの形状を拘束するステレオ対応づけ”, 信学論 D-II Vol.J71, No.11, pp.2357-2369, 1988.
- 22) 河合 良浩, 植芝 俊夫, 石山 豊, 角 保志, 富田 文明, “セグメントベースステレオにおける連結性と対応評価,” 信学技法, PRMU96-135, 1997
- 23) L.McMillan and G.Bishop, “Plenoptic modeling: An image-based rendering system,” In SIGGRAPH95, pp.39-46,1995.
- 24) X.-X.Cheng, X.-Y.Su, and L.-R.Guo, “Automated measurement method for 360 Profilmetry of 3-D diffuse objects”, Appl. Opt., vol.30, no.10, pp.1274-1278, 1991.
- 25) 井口征士, 佐藤 宏介, “三次元画像計測”, 昭晃堂, 1990.
- 26) M.Halioua, H.C.Lui, “Optical three-dimensional sensing by phase measuring profilometry”, Optics and Lasers in Engineering, Vol.11, pp.185-215, 1989.
- 27) 大島正毅, 白井良明, “三次元情報を用いた物体認識”, 信学論, J65-D, No.5, pp.629-636, 1982.
- 28) 池内克史, 越川和忠, “幾何モデルより導出された解釈木によるピンピッキングタスク中の対象物の位置姿勢決定”, 信学論, J70-D, no.1, pp.127-138, 1987.
- 29) 池内克史, “拡張ガウス像に基づく被写体いがり表現像から観測方向決定のための一手法”, 信学論, J66-D, no.5, pp.463-470, 1983.
- 30) Binford, T.O., “Visual perception by computer”, Proc.IEEE Systems Science and Cybernetics Conf., 1971.

- 31) Agin, G.J. and Binford T.O, "Computer description of curved objects", Intern, Joint Conf. on Artificial Intelligence, pp.629-640, 1973.
- 32) 千葉直樹, 蚊野浩, 美濃導彦, 安田昌司, "画像特徴に基づくイメージモザイクング", 信学論DII, Vol.J82-D-II, No.10, pp.1581-1589,1999.
- 33) 阿久津明人, 外村佳伸, "投影法を用いた映像の解析方法と映像ハンドリングへの応用", 信学論DII, Vol.J79-D-II, No.5, pp.675-686,1996.
- 34) 苗村健, 柳澤健之, 金子正秀, 原島博, "光線情報による3次元実空間の効率的記述へ向けた光線空間射影法," 信学技報, IE95-119, pp.49-56, Feb, 1996.
- 35) Shenchang Eric Chen, "QuickTime VR - An Image-Based Approach to Virtual Environment Navigation", SIGGRAPH'95, pp. 29-38, 1995.
- 36) Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. "The Lumigraph",SIGGRAPH'96 ,pp. 43-54, 1996.
- 37) Marc Levoy and Pat Hanrahan, "Light Field Rendering",SIGGRAPH'96, pp. 31-42, 1996.
- 38) Heung-Yeung Shum, Li-wei He, "Rendering with Concentric Mosaics", SIGGRAPH'99, pp.299-306, 1999.
- 39) 岡本茂, 松山泰男, 大島邦夫編: "精鋭 コンピュータ理工学辞典", 共立出版株式会社, 1997.
- 40) 岡田至弘, 正司哲朗, 村山健二, 岩成英一, "複合メディア情報の解析に基づく仮想実空間の生成", 情報処理学会 人文科学とコンピュータシンポジウム「じんもんこん2000」, pp.139-146, (2000).
- 41) 正司哲朗, 村山健二, 岡田至弘, 百済康義, "簡易3次元モデルの解析による石窟寺院のVR復元", 情報報処理学会 人文科学とコンピュータシンポジウム「じんもんこん2001」, pp.197-204, (2001).

- 42) Arai, T., Machii, K. and Kuzunuki, S.: "Retrieving electronic documents with real-world objects on InteractiveDesk", Proc. ACM Symposium on UserInterface Software and Technology(UIST'95), pp.37-38, 1995.
- 43) Robinson, P.: "Animated paper documents", Proc. HCI'97(21B), pp.655-658, 1997.
- 44) B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," In Seventh International Joint Conference on Artificial Intelligence(IJCAI-81), pp.674-679, 1981.
- 45) 稲葉 俊, 本多 庸悟, 金子 俊一, "パノラマ撮像と地図情報に基づく位置同定法," 情報処理学会第 49 回全国大会 (分冊 2), pp.139-140, 平成 6 年度後期
- 46) H. G. Barrow et al., "Parametric correspondence and chamfer matching," in proc. of 5th Int. Joint Conf. on Artificial Intelligence, pp.659-663, 1977.
- 47) G. Medioni and R. Nevatia, "Matching images using linear features," IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, no.6, pp.675-685, 1984.
- 48) Dennis R. Holloway, "Native American Virtual Reality Archaeology: An Architect's Perspective" VIRTUAL REALITY IN ARCHAEOLOGY, Archo Press, London, England, 2000
- 49) 佐川 立昌, 西野 恒, 倉爪 亮, 池内 克史, "大規模観測対象のための幾何形状および光学情報統合システム", 画像の認識・理解シンポジウム (MIRU2002), Vol.I, pp.11-20, 2002
- 50) 新疆ウイグル自治区文物管理委員会, 拜城県キジル千仏洞文物保, 管所編, "中国石窟 キジル石窟", 平凡社, 1983
- 51) 正司 哲朗, 岡田 至弘, "立面図を用いたキジル石窟の 3 次元形状モデルの生成", 情報処理学会 第 54 回 人文科学とコンピュータ研究会, 2002-CH-54, pp.59-66, 2002

- 52) Kalay Y. E, "Worldview: An Integrated Geometric Modeling/Drafting System" IEEE CG & A, pp.36-46, 1987
- 53) 田村, 狩野, 西原, "省略のある板金三面図からの3次元モデルの復元", 情報処理学会論文誌, Vol.38, No.4, pp.771-778, 1997
- 54) P.Srivasan, P.Liang, and S.Hackwood, "Computational geometric method in volumetric intersections for 3D reconstruction" Pattern Recognition, 23(8),pp.843-857, 1990
- 55) 倉爪 亮, M.D.Wheler, 池内 克史, "リフレクタンスエッジと濃淡エッジを用いたテクスチャのアライメント", コンピュータビジョンとイメージメディア研究報告 (CVIM), 2001-CVIM-126, pp.41-48, 2001
- 56) P.Viola, W.M. Wells III, "Alignment by maximization of mutual information" International Journal of Computer Vision, Vol.24, No.2,pp.137-154, 1997
- 57) Ioannis Stamos, Peter K. Allen, "Integration of Range and Image Sensing for Photorealistic 3D modeling" Proceedings of the 2000 IEEE International Conference on Robotics and Automation, pp.1435-1440, 2000
- 58) William.E.Lorensen and Harvey.E.Cline, "Marching Cubes: A High Resolution 3D Surface Construction Algorithm", SIGGRAPH'87, Vol.21, pp.163-169, 1987.
- 59) W.E.Lorensen, H.E.Cline, "Marching cubes: A high-resolution 3d surface construction algorithm" Computer Graphics (SIGGRAPH '87), Vol.21, No.4, pp.163-169, 1987
- 60) 新疆龜茲石窟研究所編著, "克孜爾石窟内容総録", 新疆美術攝影出版社, 2000
- 61) Jackie Neider, Tom Davis, Mason Woo, "OpenGL Programming Guide(日本語版) The Official Reference Document for OpenGL, Release 1" アジソンウェスレイ, 1993

- 62) L.G.Roberts, "Machine perception of three-dimensional solids" In J.T.Tippet,editor,Optital and Electro-optical Information Processing. MIT Press, 1965
- 63) M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models" In First International Conferenceon Computer Vision, pp.259-268, 1987
- 64) 神原 誠之, 大隈 隆史, 竹村 治雄, 横矢 直和: "ビデオシースルー型拡張現実感のための実時間ステレオ画像合成", 信学論, Vol.J82-DII, No.10, pp.1775-1783, 1999.
- 65) 佐藤 清秀, 山本 裕之, 田村 秀行: "現実空間と仮想空間の位置合わせ手法-ステレオカメラと3次元センサの組み合わせ", 画像の認識理解シンポジウム (MIRU98), Vol.1, pp.7-12, 1998.
- 66) 宮間 美奈子, 小澤 庄次, 近藤 邦雄: "撮影条件の推定法を用いた実写画像とCG画像の合成", 第2回知能情報メディアシンポジウム, pp.113-118, 1996.
- 67) 佐藤 いまり, 林田 守広, 甲斐 郁代, 佐藤 洋一, 池内 克史: "複合現実感における光学的整合性の実現:基礎画像の線形和による高速レンダリング手法", 画像の認識・理解シンポジウム (MIRU2000), vol.1, pp.107-112, 2000.
- 68) 日浦 慎作, 東城 賢司, 稗田 洋也, 森谷 貴行, 井口 征士: "プロジェクタを用いた3次元遠隔指示インタフェースの構築", 画像の認識・理解シンポジウム (MIRU2002), vol.1, pp.29-38, 2002.
- 69) 向川 康博, 西山 正志, 尺長 健: "スクリーン物体への光学パターン投影による仮想光学環境の実現", 信学論, Vol.J84-DII, No.7, pp.1448-1455, 2001.
- 70) 浮田 浩行, 小西 克信, 和田 俊和, 松山 隆司: "固有空間法を用いた陰影情報からの書籍表面の3次元形状復元", 信学論, Vol.J83-D-II, No.12, pp.2610-2621, 2000.
- 71) 井口 征士, 佐藤 宏介: "三次元画像計測", 昭晃堂, 1990.

- 72) Rekimoto, J.: "Matrix: A realtime object identification and registration method for augmented reality", Proc.1998 Asia Pacific Computer Human Interaction (APCHI'98), pp.63-68, (1998).
- 73) 大津 展之: "判別および最小2乗規準に基づく自動しきい値選定法", 信学論, Vol.J63-D, 4, pp.349-356, 1980.
- 74) 牧野秀夫, 森下文仁, 阿部好夫, 山宮士郎, 長谷川勝, 石山郁夫, 中静 真: "非可視型バーコードを用いた視覚障害者用物体案内方式の研究", 電子情報通信学論, Vol.J80-DII, No.11, pp.3094-3100, 1997.

関連論文および口頭発表

論文誌

- (1) 正司 哲朗, 岡田 至弘: 平面図とステレオ法を併用した奥行き情報の獲得, 画像電子学会, Vol.29, No.4, pp.312-319, (2000).
- (2) 正司 哲朗, 百済 康義, 岡田 至弘: 正投影図を用いた視体積交差法に基づくキジル石窟およびモール仏塔の3次元形状モデリング, 情報考古学会誌, Vol.8, No.2, pp.11-21, (2003).
- (3) 正司 哲朗, 岡田 至弘: プロジェクタを用いた書籍表示インタフェースの構築, 画像電子学会誌, Vol.32, No.4, pp.397-406, (2003).
- (4) 村山健二, 正司哲朗, 岡田至弘, 平林哲: 分割K-平均法による領域分割結果を用いた航空写真画像からの道路抽出, 日本写真測量学会, 写真測量とリモートセンシング, Vol.42, No.2 pp.4-11, (2003).

国際会議

- (1) Tetsuo Shoji, Kogi Kudara, Kenji Murayama, Yoshihiro Okada: 3-D modeling and Coloring Wall Paintings for the Bezeklil caves, 5th International Dunhuang Project (IDP) Conference, Stockholm, October 2002.
- (2) Tetsuo Shoji, Yoshihiro Okada, Kogi Kudara, Takashi Irisawa and Yoshihisa Oda, Development of 3-D Contents Management System for Digital Archives, Proceedings of PNC Annual Conference and Joint Meetings 2002, (掲載予定)

- (3) T.Shoji, E.Iwanari, and Y.Okada: 3-D Objects Mapping Using Model and Image Based Approach, IAPR Workshop on Machine Vision Application MVA'98, pp.523-526, (1998)

研究会・シンポジウム

- (1) 正司 哲朗, 郡司 尚樹, 澤 弘義, 岡田 至弘: パノラマ画像を用いた仮想実空間の構築と呈示, 第3回知能情報メディアシンポジウム, pp.111-116, (1997).
- (2) 岡田 至弘, 正司 哲朗, 村山 健二, 岩成 英一: 複合メディア情報の解析に基づく仮想実空間の生成, 情報処理学会 人文科学とコンピュータシンポジウム「じんもんこん2000」, pp.139-146, (2000).
- (3) 正司 哲朗, 安田 一樹, 岡田 至弘: 3視点パノラマ画像による情景モデルの構築, MURU2000, Vol.II, pp.91-95, (2000).
- (4) 正司哲朗, 田辺 ゆかり, 安田 一樹, 岡田 至弘: 多視点画像を用いたモデル統合処理に基づく空間モデリング, 情報処理学会 CVIM-126-8, pp.57-64,(2001).
- (5) 正司 哲朗, 村山 健二, 岡田 至弘, 百済 康義: 簡易3次元モデルの解析による石窟寺院のVR復元, 情報処理学会 人文科学とコンピュータシンポジウム「じんもんこん2001」, pp.197-204, (2001).
- (6) 正司 哲朗, 岡田 至弘: 立面図を用いたキジル石窟の3次元形状モデルの生成, 情報処理学会 第54回 人文科学とコンピュータ研究会, 2002-CH-54, pp.59-66,(2002)

全国大会

- (1) 正司 哲朗, 村山 健二, 岩成 英一, 岡田 至弘: SNAKESによる対象物体の抽出に基づく画像補間, 電子情報通信学会総合大会講演論文集, Vol. 1997年. 情報システム, No. 2, p.358, 1997.
- (2) 正司哲朗, 田辺ゆかり, 安田一樹, 岡田至弘: 航空写真とステレオパノラマ画像を用いた景観モデルの生成, 電子情報通信学会 2001年総合大会, 基礎・境界, A-16-24, p.347.

- (3) 安田 一樹, 正司 哲朗, 岡田 至弘: Layered Model による広域な仮想空間表現, 電子情報通信学会 2001 年総合大会, 基礎・境界, A-16-23, p.346.
- (4) 正司哲朗, 岡田至弘: 立面図を用いた視体積交差法に基づく 3 次元モデリング, 電子情報通信学会 2002 年総合大会, 画像工学 D
- (5) 加藤 雅人, 正司 哲朗, 村山 健二, 岡田 至弘, 江南 和幸, 池田 和彦, 坂田 雅之: 簀の目の測定法の開発および応用例, 文化財保存修復学会, 第 25 回大会研究発表要旨集, pp.70-71, 2003.

解説

- (1) 正司哲朗, 岡田 至弘: 平面図とステレオ法を併用した奥行情報の推定, 画像ラボ, Vol.12, No.1, pp.31-34, (2001).