

NAIST-IS-MT0751049

修士論文

光線方向を考慮した奥行き情報の選択による 全方位動画像からの自由視点画像生成

越澤 広幸

2009年2月5日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

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

越澤 広幸

審査委員：

横矢 直和 教授 (主指導教員)

加藤 博一 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

佐藤 智和 助教 (副指導教員)

光線方向を考慮した奥行き情報の選択による 全方位動画像からの自由視点画像生成*

越澤 広幸

内容梗概

屋外環境を仮想化し計算機上に再現する技術は、遠隔テレプレゼンスや景観のデジタルアーカイブ、エンターテインメント等、様々な分野への応用が期待されている。現在、屋外環境を仮想空間内に再現するための手段として、人手によってCADデータを作成した上で実写によるテクスチャマッピングを行う手法が広く用いられている。しかし、広範囲の屋外環境を対象とした場合には、多くの人的コストがかかることが問題となっている。そこで、予め取得した画像群を利用して屋外環境の仮想化を自動化する試みとして、仮想化の対象となる環境を単一の三次元モデルとして再現する Model-Based Rendering(MBR) 法や、取得した画像群から任意の視点の画像を直接生成する Image-Based Rendering(IBR) 法が提案されてきた。しかし、前者の手法では、単一の三次元モデルとして再現することが難しい樹木のような複雑な形状を持つ物体の幾何情報を正確に推定することが困難であり、それに加えてオクルージョンにより遮蔽される領域に欠損が生じるため、写実性が損なわれるという問題がある。また、後者の手法では、入力画像を変形させることで新たな視点の画像を生成するため、前者に比べて視点の移動に伴って発生する幾何学的な歪みが大きく、入力として用いる画像の撮影範囲に対して、高品位な画質を保ちながら移動できる範囲が狭いという問題がある。このような問題に対して、本研究では任意に設定した視点位置に応じて適切な形状に変形する三次元メッシュモデルを用いて自由視点画像を生成する手法を提案

* 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0751049, 2009年2月5日.

する．提案手法では，仮想化現実空間を幾何学的に正しく描画可能である MBR 法を基礎とすることで，IBR 法の問題であった視点移動に伴う生成画像上での幾何学的な歪みを抑制する．また，MBR 法のように単一のモデルを生成せず，視点位置に応じて変形する三次元形状モデルを逐次生成することで，MBR 法の問題であったモデル上の欠損を解消し，画質の向上を図る．実験では，全方位マルチカメラシステムによって実環境を撮影した全方位動画像を入力として用い，自由視点画像生成を行うことにより，撮影範囲に対して比較的広範囲な視点移動が可能な仮想環境内でのウォークスルーを実現できることを示す．

キーワード

自由視点画像生成, 全方位動画像, 奥行き画像

Novel View Generation from Omni-Directional Videos Using View-Dependent Deformable 3-D Mesh Model*

Hiroyuki Koshizawa

Abstract

This thesis describes a method for novel view generation using deformable 3-D mesh model. For novel view generation, there are two kinds of approaches; Model-Based Rendering(MBR) and Image-Based Rendering(IBR). In the MBR, free viewpoint images are generated by using a single unified 3-D model of a target environment. On the other hand, in the IBR, free viewpoint images are generated by directly deforming captured images taken at several different view points. However, these methods have some problems. In the former approach, constructed 3-D models often have holes that are caused by occlusions and errors in 3-D reconstruction. In the latter approach, free viewpoint images are easily distorted when the viewpoint is set far from the capturing positions of the input images. In order to solve these problems, we use a deformable 3-D mesh model and omni-directional images to render the free viewpoint images. In the proposed method, the 3-D mesh is deformed by using the pre-estimated omni-directional depth maps that are selected based on the position and posture of the free viewpoint. Our approach is basically based on the MBR approach that can render the geometrically correct virtualized world. However, we don't use a single unified 3-D model to avoid the hole problem. In the experiment,

* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0751049, February 5, 2009.

free viewpoint images are generated from omni-directional images captured by an omni-directional multi-camera system to show the feasibility of the proposed method for the walk-through applications in the virtualized environment.

Keywords:

Novel view generation, Omni-directional video, Depth map

目次

1. はじめに	1
2. 自由視点画像生成の関連研究と本研究の位置付け	3
2.1 自由視点画像生成に関する従来研究	3
2.1.1 Model-Based Rendering	3
2.1.2 Image-Based Rendering	8
2.1.3 Hybrid Rendering	10
2.2 本研究の位置付けと方針	12
3. 光線方向を考慮した奥行き情報の選択による自由視点画像生成	14
3.1 自由視点画像生成手法の概要	14
3.2 入力情報の取得	16
3.2.1 カメラ位置・姿勢情報の推定	16
3.2.2 奥行き情報の推定	17
3.3 エネルギー最小化による視点位置に応じた三次元メッシュモデルの生成	18
3.3.1 エネルギー関数の定義	19
3.3.2 奥行き画像の選択	21
3.3.3 メッシュモデルの初期化	21
3.3.4 エネルギー最小化によるメッシュモデルの変形	22
3.4 メッシュモデルへのテクスチャのマッピング	24
4. 実験	26
4.1 全方位動画からの自由視点画像生成実験	26
4.1.1 入力情報の取得	26
4.1.2 自由視点画像生成	30
4.2 考察	36
5. まとめ	37

謝辭	38
參考文獻	39

目 次

1	レーザレンジファインダを用いた三次元モデルの復元結果 [3]	5
2	不特定多数の人物が撮影した画像群からのステレオ法 [15]	7
3	View Morphing [24]	9
4	平面の集合によって近似的に三次元形状を復元する手法 [31]	11
5	本手法による自由視点画像生成の流れ	15
6	三次元メッシュモデルの初期設定	18
7	メッシュモデルの各頂点の座標	19
8	エネルギーの算出に用いる変数	20
9	各頂点に対する光線と撮影カメラ位置からの距離	22
10	視点移動後のメッシュモデルの変形	23
11	視点移動後の奥行き d の探索範囲	24
12	メッシュモデルの三角形パッチ Ω と三頂点	25
13	全方位型マルチカメラシステム Ladybug	27
14	入力として用いた全方位動画の 1 フレーム	27
15	SFM 法によって推定されたカメラ位置・姿勢	28
16	マルチベースラインステレオ法により推定された全方位奥行き画 像の一例	29
17	ドローネの三角形分割法により補間された全方位奥行き画像の一例	29
18	自由視点画像の生成視点位置・視線方向	30
19	ルート A における自由視点画像	32
20	ルート B における自由視点画像	33
21	ルート A における自由視点画像に対応する奥行き画像	34
22	ルート B における自由視点画像に対応する奥行き画像	35

表 目 次

1	自由視点画像の生成手法の比較	12
2	本実験のパラメータ	31
3	使用した計算機のスペック	31
4	自由視点画像の生成にかかる処理時間	31

1. はじめに

実環境を仮想化し計算機上に再現する技術は仮想化現実 (Augmented Virtuality) と呼ばれ、遠隔テレプレゼンスや景観のデジタルアーカイブ、エンターテインメント等の様々な分野で応用が期待されている。このような応用分野において、ユーザは仮想空間内で自由に視点位置・視線方向を変えることによって、計算機上に再現された実環境を体験することができる。従来から、実環境を仮想空間内に再現するための手段として、人手によって建物の三次元形状を CAD データとして作成した上で、実写によるテクスチャマッピングを行う手法が最も広く用いられてきた。しかし、広範囲の実環境を対象とした場合には、このようなモデリングの作業に多くの人的コストがかかることが問題となっている。そこで、予め取得した画像群を利用して屋外環境の仮想化を自動化する試みが多くなされている。代表的な手法として、仮想化の対象となる実環境を多数の地点で撮影し、取得された画像群から三次元復元によって仮想環境を再現する Model-Based Rendering 法 (以下、MBR 法) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22] や、取得された画像群から撮影環境を三次元復元することなく、画像の変形・合成により任意の視点の画像を直接生成する Image-Based Rendering 法 (以下、IBR 法)[23, 24, 25, 26, 27, 28, 29, 30] が研究されてきた。しかし、MBR 法では、復元した三次元空間内で視点・視線方向を自由に設定することができる反面、単一の三次元モデルとして再現することが難しい樹木のような複雑な形状を持つ物体の幾何情報を正確に再現できず、それに加えてオクルージョンにより遮蔽される領域にモデルの欠損が生じるため、写実性が損なわれるという問題がある。これに対して、IBR 法では幾何情報を直接用いないため、複雑な形状の物体に対しても実物体に近い見え方を再現することが可能であるが、入力画像を変形させることで新たな視点の画像を生成するため、MBR 法に比べて視点の移動に伴って発生する幾何学的な歪みが大きく、写実性を保ったまま自由に広範囲を移動するアプリケーションに応用することは難しい。また、MBR 法と IBR 法を共に利用する Hybrid Rendering 法 [31, 32, 33] も提案されているが、各手法を線形的に結合したものが多く、根本的な問題の解決には至っていない。

このような問題に対して、本研究では任意に設定した視点位置に応じて適切な

形状に変形する三次元メッシュモデルを用いて自由視点画像を生成する手法を提案する．提案手法では，仮想化現実空間を幾何学的に正しく描画可能である MBR 法を基礎とすることで，IBR 法の問題であった視点移動に伴う生成画像上での幾何学的な歪みを抑制する．また，MBR 法のように単一のモデルを生成せず，視点位置に応じて変形する三次元形状モデルを逐次生成することで，MBR 法の問題であったモデル上の欠損を解消し，画質の向上を図る．具体的には，入力として用いる複数の画像に対してあらかじめ推定された奥行き情報の整合性に関するエネルギー関数を設定し，視点の前面に設定したメッシュモデルの各頂点を，エネルギーが最小となる地点に移動させることで，視点位置に応じた最適なモデルを作成する．最後に，生成したメッシュモデルに対して，視線方向を表す直線に最も近い地点からの画像をテクスチャマッピングすることで自由視点画像を生成する．ただし，本研究では仮想空間内での任意の見回しと視点位置に応じた最適な三次元モデルの生成を実現するために，実環境を全方位型マルチカメラシステムを用いて撮影し，全方位カメラの運動パラメータと全方位画像に対する奥行き画像を既存の手法によって推定した後に入力として用いる．実際には，撮影した画像群に対して Structure from Motion 法 [19] を適用することで全方位カメラの運動パラメータを推定し，更にマルチベースラインステレオ法 [13] を用いて各撮影画像に対応する奥行き画像を生成する．

以下，第 2 章では，自由視点画像の生成手法に関する従来研究と本研究の位置付けについて述べる．第 3 章では，視点位置に応じて適切に変化するメッシュモデルを用いた全方位動画からの自由視点画像生成の手法について詳述する．第 4 章では，提案手法を用いて，屋外環境において任意の視点の見回し・移動を行った際の自由視点画像を生成した結果を示す．最後に第 5 章で，まとめと今後の展望を述べる．

2. 自由視点画像生成の関連研究と本研究の位置付け

本章では、仮想空間内でユーザが任意に設定した視点位置・視線方向に対応する画像を生成する自由視点画像生成手法の関連研究について概観するとともに、本研究の位置付けを述べる。

2.1 自由視点画像生成に関する従来研究

現在までに提案されている自由視点画像の生成手法は、対象環境の三次元形状を用いることで自由な視点位置・視線方向の画像を生成する MBR 法、対象環境を明示的に三次元復元することなく画像の変形・合成により自由視点画像を生成する IBR 法、それらの手法を組み合わせた Hybrid Rendering 法に大別される。以下では、これら 3 つの手法について詳述する。

2.1.1 Model-Based Rendering

MBR 法では、自由視点画像を生成するために対象環境の三次元形状を直接用いる。この手法では、対象環境の幾何学的な構造が再現されるため、視点・視線方向の設定を 6 自由度 (位置 3 自由度・姿勢 3 自由度) で設定でき、また視点移動に伴う物体の奥行き関係の再現が可能である。また、仮想化された環境内に他の物体を追加表示したり、環境内の物体の配置を変更することが容易であるという特長がある。ただし、MBR 法で高品質な自由視点画像を生成するためには、対象環境の三次元形状を精度よく計測する必要がある。仮想化の対象となる環境の三次元形状を取得するための方法としては、撮影対象に積極的に電磁波を照射することで対象の形状を計測する能動的な方法と、対象環境への働きかけは行わず、撮影した画像群から画像処理によって三次元形状を推定する受動的な方法の 2 つが存在する。どちらの手法においても、一度に三次元復元可能な範囲は限られているため、1 地点のみから対象環境を完全に計測することはできない。従って、一般的には対象環境全体の三次元形状を取得するために、部分的に計測された三次元形状の位置合わせを行う必要がある。以下では、対象環境の部分形状に関する能動的、受動的な三次元復元手法およびその位置合わせの手法について述べる。

[能動的な三次元復元手法]

能動的な三次元復元手法は，スリット光を用いた三角測量による手法と，光レーザの飛行時間測定による手法に大別される．

スリット光の三角測量による手法 スリット光の三角測量による手法では，プロジェクタ等によってスリット光を対象に照射し，近接する位置に設置したカメラでスリット光の位置を検出し，三角測量の原理によって対象環境の三次元形状を奥行き画像として取得する [1, 2]．三角測量による計測では，スリット光を照射するプロジェクタとカメラの間の距離 (ベースライン) と，カメラと対象までの間の距離の比によって計測精度が決まり，対象までの距離に対して比較的容易にベースラインを長く設定できる小物体に対しては高精度でその三次元形状を推定できるという特長がある．その反面，屋外等の広域環境を対象とした場合には，ベースラインを十分に確保することが困難であり，また屋外では日照条件によってスリット光の観測が困難となる問題も加わって，対象環境の三次元形状を高精度に推定することは難しい．

光レーザの飛行時間測定による手法 光レーザの飛行時間測定による手法では，対象環境に光レーザを照射することによって，光が反射して計測されるまでの時間 (time of flight) を計測することでレンジデータを取得する．このような光レーザを二次元走査することで，レンジデータを奥行き画像として取得することが可能である．一般的に，レーザの反射強度が一定以上であれば屋外のような広域環境においても高精度なレンジデータを取得することが可能である．一般に取得されたレンジデータは対象環境の部分形状であるため，複数地点で取得されたレンジデータの幾何学的な位置合わせを行う手法が多数提案されている [3, 4, 5]．これらの手法では，レンジデータの位置合わせを行うために重複して計測された領域が必要となるため，計測地点を密にする必要がある．しかし，一般的なレンジファインダでは，各地点の計測において数分程度の計測時間が必要となるため，計測コストを考慮すると計測地点を密にすることは難しい．このため，オクルージョンが発生しやすい複雑な環境では，計測地点の粗さによる計測漏れが発生しやすく，図 1 のようにモデルの欠損が生じてしまうという問題を抱えている．

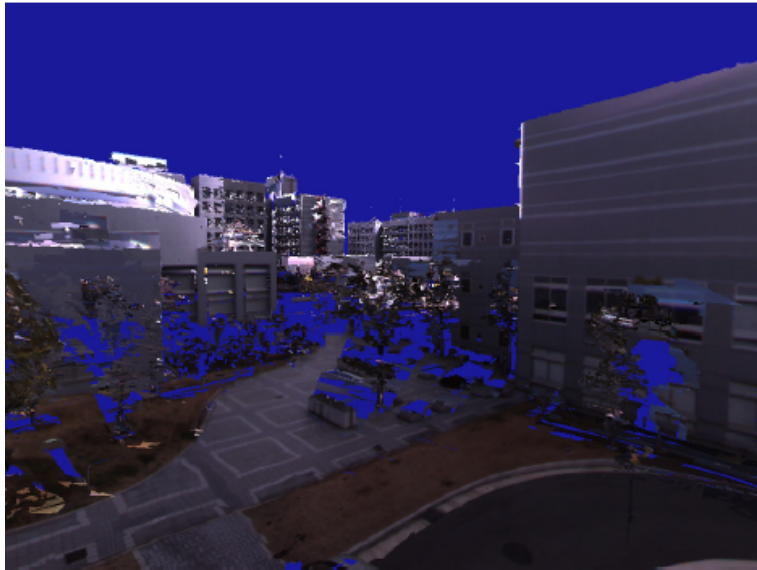


図 1 レーザレンジファインダを用いた三次元モデルの復元結果 [3]

[受動的な三次元復元手法]

受動的な三次元復元手法は、対象環境を撮影した一枚または複数枚の画像を用いて、画像処理によって三次元形状の復元を行う手法である。この手法は、一般的なカメラの他に、一度に全方位の画像を取得できる全方位カメラを入力機器として用いることもできるため、比較的短時間で入力データを取得できるという特長がある。受動的な三次元復元手法としては、陰影による手法、視体積交差法、ステレオ法、レンズ焦点法、特徴点の追跡による手法等がある。以下では、これらの各計測手法について述べる。

陰影による手法 陰影による手法は、撮影対象の陰影画像を取得し、撮影対象の反射特性や光源情報を用いて三次元形状の復元を行う手法である [6, 7]。この手法では、任意の点の奥行きを与えることで、画像上のすべての点の奥行き値を得ることができる。また、対象環境に関する滑らかさの制約条件などを用いることで比較的精度の良い復元を行えるという特長がある。しかし、陰影による手法は、暗室のような光源状況を任意に設定できる環境でしか利用できないため、光源環

境が複雑な屋外環境の三次元復元に用いることは困難である。

視体積交差法 視体積交差法は、対象の輪郭情報を用いる手法である。この手法では、撮影時のカメラ位置・姿勢を既知とし、輪郭情報を空間に投影して得られる錐体(視体積)の内側に対象が含まれるという仮定の下に対象物の三次元形状を取得する[8, 9, 10]。視体積交差法では、視体積の交差する領域を用いて三次元形状の復元を行うため、物体の凹み部分など、輪郭として現れない形状を復元することはできない。そのため、屋外環境のように形状が複雑で様々な方向からの輪郭情報を取得することが困難な環境に対してこの手法を適用することは難しい。

ステレオ法 ステレオ法は、位置・姿勢が既知の二台以上のカメラによって撮影された画像に対し、撮影画像上で各画素の対応を決定することで、三角測量の原理により奥行き画像を推定する手法である。三台のカメラを用いることで、対応点のあいまいさを減少させる三眼ステレオ法[11, 12]や、多数の画像を同時に利用するマルチベースラインステレオ法[13]も提案されている。一般的に、ステレオ法ではカメラ位置関係の把握が重要であり、高精度なキャリブレーションを実現することが必要となる。このようなステレオカメラのキャリブレーション手法として、三次元位置が既知のマーカを用いる手法[14]が代表的である。しかし、屋外環境を対象とした場合には、画角内に十分な数のマーカを配置することが困難であるため、後述する特徴点の追跡による手法によって得られるカメラ位置・姿勢の推定結果を用いてステレオを行う手法が提案されている[13]。また近年は、広域を対象としたステレオ法も提案されており、Goeseleら[15]は共有写真データベースなどを対象として、不特定多数の人物が撮影した観光地等の画像群からステレオを行う手法を提案している。しかし、図2に示すように、複雑な環境を単一の三次元モデルで再現するのは困難であり、モデルに欠損が生じるという問題がある。

レンズ焦点法 レンズ焦点法は、レンズの焦点深度を変化させた複数の画像において、被写界深度の範囲外に写る対象に画像上のボケが発生するという光学的な特徴を利用して奥行きを求める手法である[16, 17]。この手法は、対象環境に対

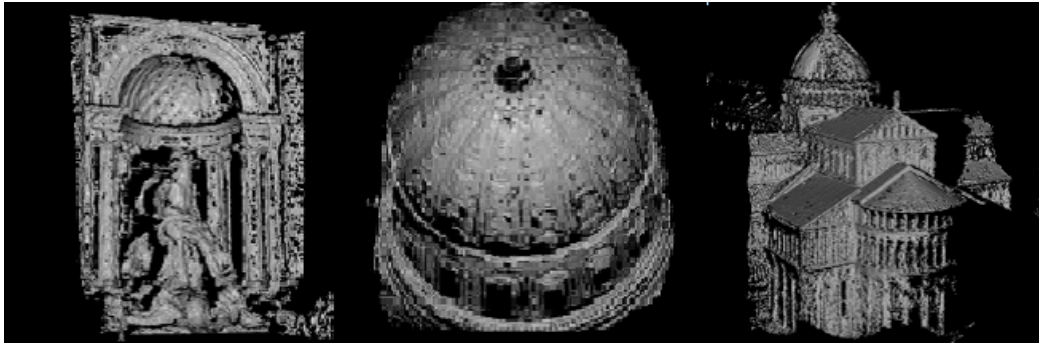


図 2 不特定多数の人物が撮影した画像群からのステレオ法 [15]

する事前知識や制約条件が必要でないという特長がある。しかし、レンズの焦点を微小に変化させながら計測を行う必要があるため、一枚の奥行き画像を推定するのに時間がかかるという問題がある。また、屋外環境においては、被写界深度が広がるため高精度な三次元形状の復元は困難である。

特徴点の追跡による手法 特徴点の追跡による手法は、複数の静止画像または動画像を入力とし、画像上の特徴点を複数の画像間で対応付ける（追跡する）ことによって、特徴点の三次元位置とカメラの移動パラメータを取得する手法である [18, 19]。この手法は Structure from Motion 法（以下、SFM 法）と呼ばれ、任意の移動を伴った画像系列から特徴点の三次元位置を取得できる。また、対象環境や撮影時のカメラ位置・姿勢に対する制約が少なく、容易に屋外環境の三次元形状を取得できる。しかし、一般的に動画像上の特徴点のみを用いる手法では、撮影カメラと特徴点の相対的な位置関係のみが推定され、絶対的な位置・姿勢情報を取得するためには何らかの外部指標を与えなければならない。このための外部指標として、GPS やジャイロなどのセンサ類を利用する手法 [20, 21] や、三次元位置が既知の基準点を用いる手法 [19] が提案されている。これらの手法を用いることで、現実世界との位置関係を取得することができる。また、近年は一般的な単眼カメラ以外に、全方位カメラによって撮影された動画像に対して SFM 法を適用する手法が提案されている [19, 22]。全方位カメラを用いることで対象環境の情報を効率的に取得できるという利点があるが、追跡された特徴点以外の点に対

する三次元情報が得られないため、復元される情報の密度がモデリング等に利用するためには十分でないという問題がある。

このように、特徴点の追跡による手法は、対象環境と撮影カメラの位置・姿勢に対する制約が少なく、自由に移動するカメラの位置・姿勢を推定するのに適しているが、三次元情報は疎にしか得られないため、三次元モデルを構築するためにはステレオ法などの手法と組み合わせて利用する必要がある。

2.1.2 Image-Based Rendering

IBR 法は対象環境の三次元形状を直接用いずに、物体やシーンの見え方そのものを入力画像として記録しておき、これらを変形・合成することで新しい視点の画像を生成する手法である。この手法では、対象物体の形状や反射特性を直接利用しないため、MBR 法では適用が難しい複雑な環境に対しても欠損なしに自由視点の映像を生成することができるという特長を持つ。以下では IBR 法を、蓄積した画像群を直接変形・合成することで自由視点画像を生成する手法と、空間中を伝搬する光線情報を用いる手法に大別し、それぞれの特徴を述べる。

[画像群の変形による手法]

蓄積された複数の画像群を直接変形することにより自由視点画像を生成する手法として、ワーピングやモーフィングと呼ばれる手法 [23, 24] が提案されている。これらの手法は、入力となる複数の画像間に対応点を与え、中間視点位置において対応点がどの位置に現れるかを決定することで、入力画像を変形させる。具体的には、入力画像を対応点によってあらかじめメッシュ状に分割しておき、メッシュ内のテクスチャを対応点の動きに応じて変形させた上で、複数の入力画像のテクスチャをメッシュ内でブレンディングする。また、その中間視点における対応点の RGB 値を入力画像の色のブレンディングによって決めることで自由視点画像を生成する (図 3)。この手法は、比較的単純であり手動で対応点を与えるだけで簡単に 2 視点間の中間視点画像を生成できるという特長があるが、これらの手法では、環境の形状に関する情報を用いないため、物体の隠蔽関係を正しく再現することが困難であり、そのような領域では不自然な画像が生成されるという

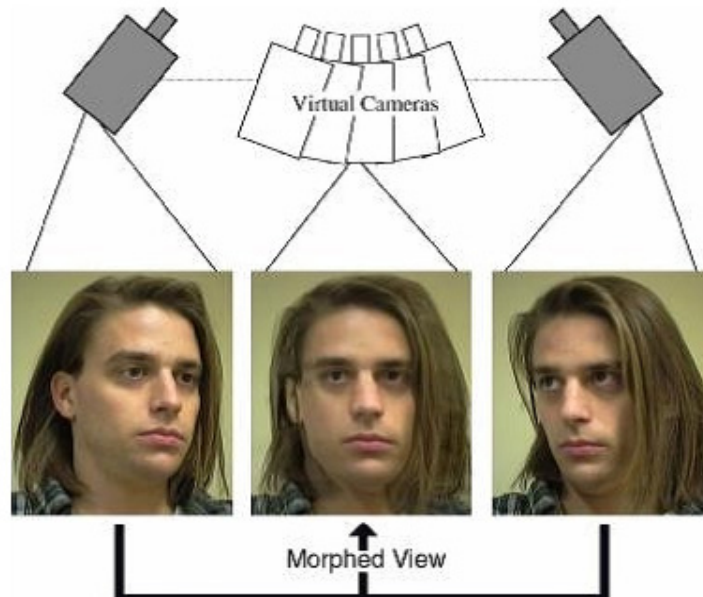


図 3 View Morphing[24]

問題がある．また，入力画像から直接自由視点画像を生成するため，視点移動に伴って生成画像上で幾何的な歪みが発生しやすいという問題がある．

[光線情報を用いた手法]

光線情報を用いた手法として，Plenoptic Function[25]，光線空間法 [26]，Light Field Rendering[27]，Lumigraph[28] 等が提案されている．Adelson ら [25] は Plenoptic Function を提案し，空間中に存在する光線は 7 次元ですべて表現することが出来るとした．すなわち，視点の座標 (x, y, z) ，光線の向き (θ, ϕ) ，波長 λ および時間 t の 7 次元のデータを用いることで空間中のすべての光線を記述する手法を提案している．しかし，7 次元のデータを全て記録するためには膨大な量の画像が必要となり，現実的に全てのデータを収集することは難しい．これに対して 苗村ら [26] は，より低い次元の情報で空間を記述することで任意視点の画像を効率的に再構築する手法として，光線空間法を提案している．光線空間法は，空間を伝搬する光線は減衰・変移しないものとする事で，Plenoptic Function の 7 次

元の光線情報から z 座標，波長 および時間 t を減らし，情報量を 4 次元とすることで取得すべき画像の量を大幅に減らした．また，Levoy ら [27] および Gortler ら [28] は同時期に Light Field Rendering，Lumigraph という手法を提案した．これらの手法では，光線空間法が平面 (x, y) および光線の向き (θ_x, θ_y) で 4 次元空間を表現したのに対して，この 4 次元を 2 つの平行した平面を横切る座標 (u, v) および (s, t) で表現することで，効率的なレンダリングを可能にした．

また，全方位カメラを用いることで，一度に多数の方向の光線情報を収集する手法も提案されている [29, 30]．山口ら [29] は，全方位画像センサ HyperOmni Vision を用いて直線上を移動しながら環境を撮影し，任意の視線方向に対してステレオ視可能な画像を生成する手法を提案した．また，高橋ら [30] は同様に全方位画像センサを搭載した自動車で環境中を走行し，その移動経路付近の自由視点画像を生成する手法を提案している．これらの手法では，全方位カメラを用いることで取得すべき画像の枚数を削減しているが，広域な環境を自由に移動するといった用途で利用する場合には依然として膨大な量の画像が必要となるという問題がある．

2.1.3 Hybrid Rendering

MBR 法と IBR 法とを組み合わせたハイブリッドレンダリングによる手法が提案されている．Kitahara ら [31] は平面の集合によって近似的に対象物体の三次元形状を表現することで自由視点画像を生成する手法を提案している (図 4)．この手法では，ボクセルやポリゴン等の幾何モデルによって物体を表現するよりも計算コストが少なく高速に自由視点映像を生成できるという特長がある．しかし，この手法は，サッカーを対象とした手法であり，対象環境の大部分が平面でかつ単色であるという前提を用いているため，一般的な環境を対象として利用できるものではない．

冨手ら [32] は，多地点で全方位画像を撮影し，それらの画像間に対応点を与えることで対象環境の疎な三次元形状を復元し，その形状に対応するテクスチャをモーフィングによって生成する手法を提案している．この手法は，あらかじめ複数の対応点を手動で与える必要があり，複雑で広域な屋外環境を対象として用い

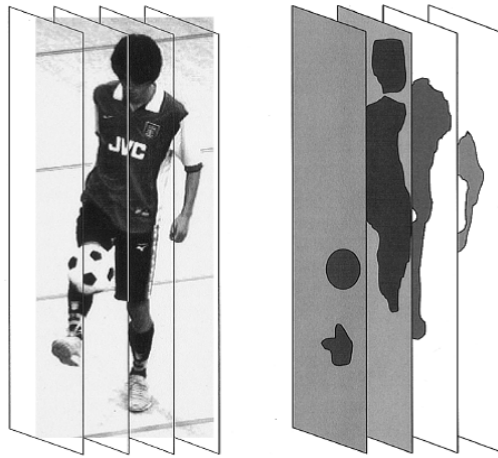


図 4 平面の集合によって近似的に三次元形状を復元する手法 [31]

ることは難しい。

篠原ら [33] は対象環境の複雑さに応じて、IBR 法と MBR 法を切り替えて利用する手法を提案している。具体的には、三次元モデルの生成が可能なシーンに対しては MBR 法、不可能なシーンに対しては IBR 法を用いて自由視点画像を生成することで、広範囲での自由視点画像生成を実現している。しかし、IBR 法・MBR 法のそれぞれの問題点は根本的には解決されておらず、また生成される画像の品質が利用する手法の切り替えによって不連続に変化してしまうという問題がある。

表 1 自由視点画像の生成手法の比較

	長所	短所
MBR 法	<ul style="list-style-type: none"> ・仮想化現実空間を幾何学的に正しく描画可能 	<ul style="list-style-type: none"> ・複雑な三次元形状を再現することが難しい ・生成される画像上に欠損が生じる
IBR 法	<ul style="list-style-type: none"> ・生成される画像上に欠損が生じない ・一定の移動範囲内では写実的な画像を生成できる 	<ul style="list-style-type: none"> ・視点移動に伴う生成画像上の歪みが大きい ・入力として膨大なデータ量が必要

2.2 本研究の位置付けと方針

本研究は、屋外環境のような複雑な環境下における自由視点画像の生成を目的としている。既に述べた通り、自由視点画像の生成手法には MBR 法、IBR 法、Hybrid Rendering 法の三つの手法が存在する。

表 1 に MBR 法、IBR 法の長所と短所をそれぞれ示す。MBR 法は対象環境を単一の三次元モデルとして復元しようとするため、樹木のような自然物等の複雑な形状を持つ物体の幾何情報を再現することが難しく、それに加えてオクルージョンにより遮蔽される領域にモデルの欠損が生じるため、写実性が損なわれるという問題を抱えている。IBR 法は対象環境の複雑さに依存せずに自由視点画像を生成することができるが、入力画像を変形することで新しい視点の画像を生成するため MBR 法と比べて視点移動に伴う生成画像上の歪みが大きく、写実性を保ったまま自由に広範囲を移動するためには対象環境を密に撮影する必要がある。しかし、現実的には広域環境全てを密に撮影することは難しく、広域での自由な視点移動を実現することは難しい。MBR 法と IBR 法を組み合わせた Hybrid Rendering 法も提案されているが、各手法を線形的に結合したものが多く、根本的な問題の解決には至っていない。

そこで本研究では、仮想化現実空間を幾何学的に正しく描画可能である MBR 法を基礎とすることで、IBR 法の問題であった視点移動に伴う生成画像上での幾何学的な歪みを抑制する。また、MBR 法のように単一の三次元モデルを生成せず、視点位置に応じて適切な形状に変形する三次元メッシュモデルを逐次生成することで、MBR 法の問題であったモデル上の欠損を解消し、画質の向上を図る。本研究では、仮想空間内での任意な見回しと視点位置に応じた三次元モデルの生成を実現するために、一度に全方向の情報を効率的に取得できる全方位型マルチカメラシステムを用いて入力画像を撮影する。また、撮影時の全方位カメラの運動パラメータを全方位動画像に対する SFM 法 [19] を用いて推定し、マルチベースラインステレオ法 [13] を用いて全方位画像群に対する全方位奥行き画像を生成する。これら全方位画像群、カメラ位置・姿勢、全方位奥行き画像群を入力として用い、視点の前面に生成した三次元メッシュモデルに対して適切な全方位画像をテクスチャマッピングすることで自由視点画像を生成する。

3. 光線方向を考慮した奥行き情報の選択による自由視点画像生成

本章では、まず初めに自由視点画像の生成手法の概要を述べる。次に、本研究で用いる入力情報の取得方法について述べる。最後に、エネルギー最小化を用いた視点依存三次元メッシュモデルの生成手法とメッシュモデルへのテクスチャマッピング法について詳述する。

3.1 自由視点画像生成手法の概要

図 5 に提案手法による自由視点画像生成の処理の流れを示す。提案手法では、まず対象環境の情報を取得し、次に取得した情報を入力として、ユーザが設定した視点位置・視線方向に対する自由視点画像を生成する。入力情報の取得は三つのステップから成る。まず最初に、全方位カメラを用いて移動しながら対象環境を撮影する。次に、撮影された動画像に対して SFM 法 [19] によって撮影時のカメラ位置・姿勢を計測する。最後に、マルチベースラインステレオ法 [13] により、対象環境の撮影画像上の特徴点の三次元位置を計測する。自由視点画像生成の処理は、三次元メッシュモデルの生成フェーズとテクスチャのマッピングフェーズから成る。三次元メッシュモデルの生成フェーズでは、まずメッシュモデルの各頂点に対し、頂点と視点位置を結ぶ光線と撮影地点との距離が最も近い順に複数の撮影地点を選択する。次に、初期値として任意に設定した視点の前面にメッシュモデルを生成する。最後に、選択された複数の撮影地点において推定された奥行き画像群を利用し、メッシュモデルの各頂点を光線上で視点位置から最適な距離に移動させる。テクスチャのマッピングフェーズでは、まず複数の撮影画像の中から、テクスチャマッピングに利用する画像を選択し、各メッシュへのテクスチャのマッピングを行う。本手法では、ユーザが視点位置・視線方向の変更を行う度に、奥行き画像の選択以降の処理を繰り返すことで自由視点画像を生成する。以上の手順により、視点位置に応じて最適な形状に変形する三次元メッシュモデルを用いた自由視点画像を生成する。

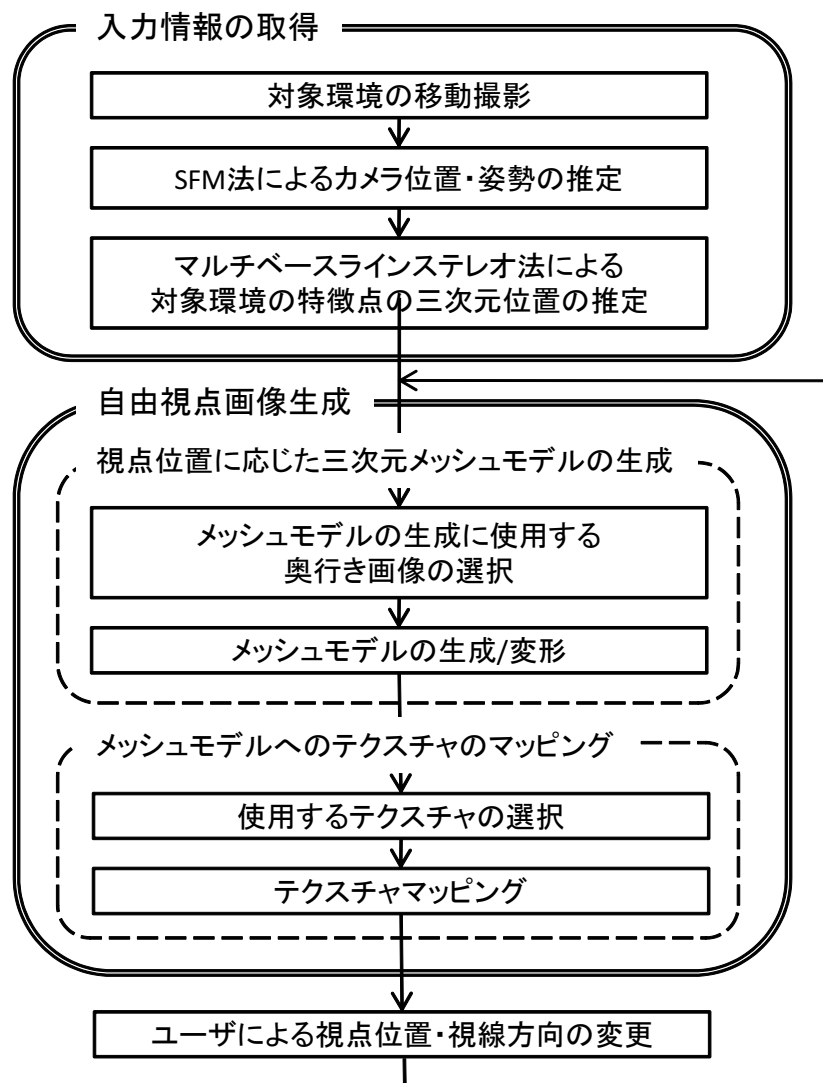


図 5 本手法による自由視点画像生成の流れ

3.2 入力情報の取得

本研究では、自由視点画像生成の対象となるシーンを、全方位の情報を効率的に一度に取得可能な全方位カメラを用いて移動しながら撮影する。本節では、全方位画像取得時の撮影カメラ位置・姿勢の推定方法、また取得した各全方位画像に対応する全方位奥行き画像の推定について述べる。

3.2.1 カメラ位置・姿勢情報の推定

ここでは、撮影した全方位動画像に対する SFM 法 [19] を用いて、取得した全方位動画像から特徴点の三次元位置と撮影時のカメラ位置・姿勢を推定する。次に、動画像中の各フレームにおいて以下に示す処理を行い、これらを最終フレームまで繰り返すことで、カメラ位置・姿勢と特徴点の三次元位置を逐次推定する(詳細は文献 [19] を参照)。

特徴点の追跡 テンプレートマッチングと Harris オペレータにより前のフレームに存在する各特徴点について、現フレームにおける対応点を探索する。さらに、RANSAC アルゴリズムにより、特徴点の誤対応を排除する。

カメラ位置・姿勢の推定 追跡された特徴点の画像上の座標と三次元座標から、カメラの位置・姿勢を推定する。ここでは、検出された全て特徴点の三次元座標を画像に再投影した際の再投影誤差の二乗和が最小になるようにカメラ位置・姿勢を推定する。なお、特徴点の三次元座標は、前フレームまでに推定されている値を用いる。

特徴点の三次元位置推定 現フレームまで追跡されている全ての特徴点について、特徴点ごとに再投影誤差の二乗和を最小化することで三次元座標を算出し、フレーム毎に特徴点の三次元座標を更新する。

特徴点の追加と削除 複数の評価尺度を用いて，カメラ位置・姿勢推定に用いる信頼度の高い特徴点の組をフレームごとに自動更新する．ここで用いる信頼度は，各特徴点に対する再投影誤差の分散の逆数として定義される．

最後に，動画像全体での最適化処理 (bundle adjustment) を行うことで推定誤差を最小にする．具体的には，動画像全体に対する再投影誤差の総和を最小化することで，撮影時のカメラ位置・姿勢と特徴点の三次元位置を最適化する．

3.2.2 奥行き情報の推定

上記の SFM 法により全方位動画像中の特徴点の三次元位置を推定することができるが，SFM 法により得られる特徴点の三次元位置は非常に疎なものである．そこで，SFM 法により求めたカメラ位置・姿勢を用い，特徴点の数え上げによるマルチベースラインステレオ法 [13] を適用することで，より多くの特徴点に対して奥行き画像を生成する．具体的には，撮影された動画像に対して以下の処理を行うことで，全方位動画像の各フレームに対応した奥行き画像を生成する (詳細は文献 [13] を参照) ．

画像特徴点の検出 画像の拡大・縮小や回転に対して頑健に画像特徴点を抽出可能な Harris オペレータを用いて，各入力画像上の特徴点を全て検出する．

特徴点の奥行き推定 三次元空間中に特徴点が存在する場合には，その特徴点の画像上への投影位置にも特徴点を検出されるという仮定の下，特徴点の奥行きに対応した三次元位置の画像上への投影座標付近に最も多くの特徴点を検出される奥行きを探索し，全ての特徴点に対する奥行き画像を算出する．

誤推定結果の排除 各入力画像を基準として算出された奥行き推定結果の整合性を検証することによって誤推定結果の排除を行う．各フレームの奥行き推定に用いた全ての画像を用い，奥行き推定値の信頼度を評価する．

ただし，特徴点の数え上げによるマルチベースラインステレオ法 [13] では，奥

行き推定の評価尺度として，画像特徴点の座標からなる評価尺度を用いることで，奥行き推定のロバスト性の向上と計算コストの低減を図るため，画像特徴点以外の画素に対する奥行きは算出されない．そのため本手法では，奥行きが推定された全ての特徴点に対して，ドロネーの三角形分割法を用いて領域分割を行い，分割された領域内の奥行き値を線形補間することで密な奥行き画像を生成する．

3.3 エネルギー最小化による視点位置に応じた三次元メッシュモデルの生成

本節では，視点位置に応じて適切な形状に変形する視点依存の三次元メッシュモデルを用いた全方位動画からの自由視点画像の生成手法について述べる．本手法では，図6に示すように，まず仮想視点位置の前面に三次元メッシュモデルを設定する．次に，メッシュモデルの各頂点に対して，モデルの変形に用いる複数枚の奥行き画像を光線方向を考慮してそれぞれ選択する．最後に，入力として用いる複数の奥行き情報の整合性に関するエネルギー関数を設定し，設定されたメッシュモデルの各頂点をエネルギーが最小となる地点に移動させることで，視点位置に応じた最適なモデルを生成する．以下ではまず，メッシュモデルの生成に用いるエネルギー関数を定義し，モデルの変形に用いる奥行き画像の選択について述べ，次にメッシュモデルの初期化，エネルギー最小化を用いたメッシュモデルの変形手法について詳述する．

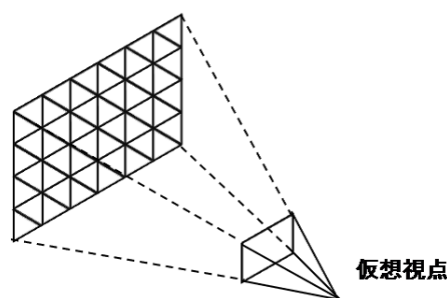


図6 三次元メッシュモデルの初期設定

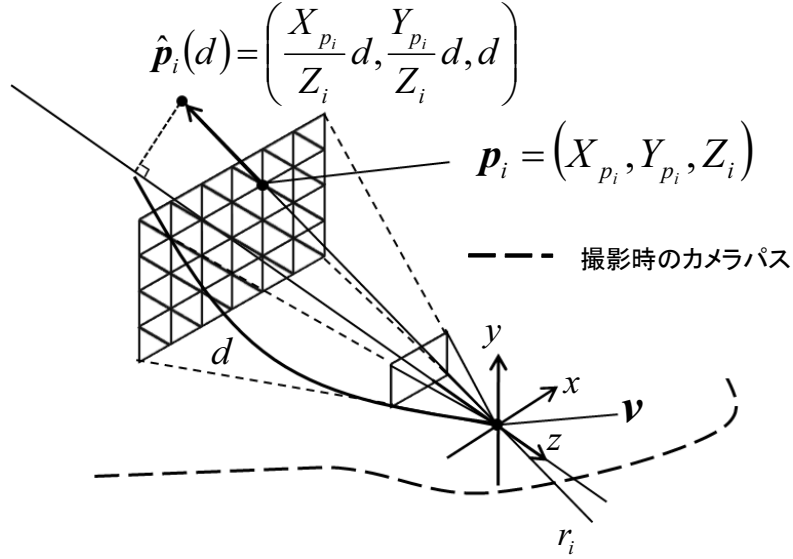


図 7 メッシュモデルの各頂点の座標

3.3.1 エネルギー関数の定義

本研究では，図 7 に示す仮想視点位置 v に対して設定されたメッシュモデル上の i 番目の頂点 p_i (カメラ座標系での三次元位置: (X_{p_i}, Y_{p_i}, Z_i)) に対して，式 (1) に示す奥行き画像の整合性に関するエネルギー $E_i(d)$ を定義し， p_i と v を結ぶ直線 r_i 上でエネルギー $E_i(d)$ が最小となる位置 $\hat{p}_i(d) = (\frac{X_{p_i}}{Z_i}d, \frac{Y_{p_i}}{Z_i}d, d)$ に頂点 p_i を移動させることで，メッシュモデルの変形を行う．

$$E_i(d) = \frac{\sum_{j=1}^N w_j(\hat{p}_i(d)) D_{ij}(d)}{\sum_{j=1}^N w_j(\hat{p}_i(d))} \quad (1)$$

$$D_{ij}(d) = \begin{cases} \{e(\mathbf{c}_{ij}, \hat{\mathbf{p}}_i(d)) - |\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}|\}^2 & ; (e(\mathbf{c}_{ij}, \hat{\mathbf{p}}_i(d)) > |\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}|) \\ 0 & ; otherwise \end{cases} \quad (2)$$

$$w(\hat{\mathbf{p}}_i(d)) = \begin{cases} \theta_{ij}(d)^{-1} = \arccos\left(\frac{(\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}) \cdot (\hat{\mathbf{p}}_i(d) - \mathbf{v})}{|\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}| |\hat{\mathbf{p}}_i(d) - \mathbf{v}|}\right)^{-1} & ; (D_{ij}(d) \neq 0) \\ 0 & ; otherwise \end{cases} \quad (3)$$

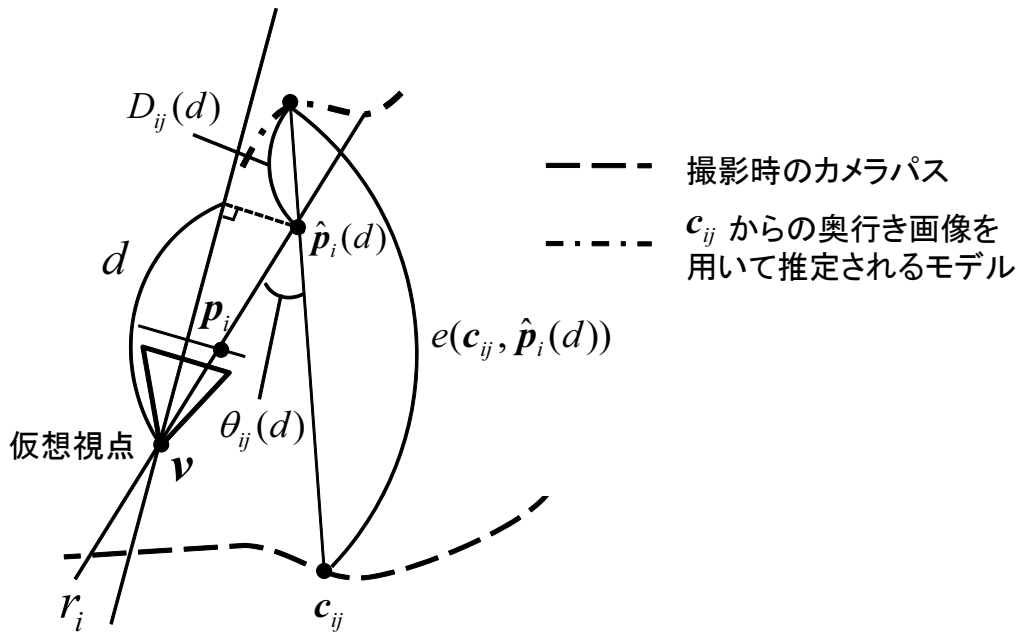


図 8 エネルギーの算出に用いる変数

ここで、 N は奥行き整合性の判定に用いる奥行き画像 ($DM_{i1}, DM_{i2}, \dots, DM_{iN}$) に対応する全方位画像の撮影地点 ($c_{i1}, c_{i2}, \dots, c_{iN}$) の数である。 $e(c_{ij}, \hat{p}_i(d))$ は図 8 に示すように、 c_{ij} を中心とする奥行き画像上における三次元位置 $\hat{p}_i(d)$ の方向に対応する奥行き値を表す関数である。 また、 $w(\hat{p}_i(d))$ は頂点 p_i に対して選択された撮影地点 c_{ij} に関する重みであり、 c_{ij} と $\hat{p}_i(d)$ を結ぶ直線と光線 r_i の成す角 $\theta_{ij}(d)$ の逆数として定義する。 ここで、 $(e(c_{ij}, \hat{p}_i(d)) > |\hat{p}_i(d) - c_{ij}|)$ の時、すなわち、撮影地点 c_{ij} から見てオクルージョンによって $\hat{p}_i(d)$ が遮蔽されている可能性がある場合には、地点 c_{ij} の奥行きを整合性の判定に用いず、 c_{ij} によるエネルギーを 0 として計算する。 ただし、整合性の判定に用いる c_{ij} の数が一定数 N_t 以下となる奥行き値 d に対しては、奥行き整合性を正しく判定できないため、エネルギー最小化による奥行き値の探索範囲から除外する。

式 (1) で定義したエネルギー $E_i(d)$ はオクルージョンとなる地点を除いた Q 地点 ($Q \leq N$) の撮影地点 ($c_{i1}, c_{i2}, \dots, c_{iQ}$) に対応する奥行き画像 ($DM_{i1}, DM_{i2}, \dots, DM_{iQ}$) から得られる三次元モデルと頂点 p_i の移動先 $\hat{p}_i(d)$ との距離の総和が最も小さい

時に最小となる。従って、メッシュモデルの各頂点 p_i に対してエネルギーを最小化する地点 $\hat{p}_i(d)$ に頂点を移動することで、各撮影地点から推定された奥行き情報に対して幾何学的な整合性が取れたメッシュモデルを生成することができる。

ここで、エネルギーが最小となる奥行き値 d を決定するために、 d が取り得る全ての探索範囲を対象としてエネルギー $E_i(d)$ を計算するには多くの計算コストが必要となるため、以下では全探索によって一旦最適なメッシュを作成し、以降はそれを初期値として探索範囲を制限することで、ユーザによる視点移動時のメッシュ変形にかかる計算コストを削減する。

3.3.2 奥行き画像の選択

ここでは、メッシュモデルの各頂点に対するエネルギー $E_i(d)$ の算出に用いる奥行き画像 $(DM_{i1}, DM_{i2}, \dots, DM_{iN})$ と、それに対応する全方位画像の撮影地点 $(c_{i1}, c_{i2}, \dots, c_{iN})$ の選択方法について述べる。本手法では、メッシュモデルの各頂点 p_i に対し、移動先となる座標 $\hat{p}_i(d)$ の奥行き値 d を決定する上で、 $\hat{p}_i(d)$ に対する奥行き情報の信頼度の高い複数の撮影地点を選択する。ここでは、頂点 p_i と仮想視点位置 v を結ぶ光線 r_i との距離 l が近い撮影地点ほど他の物体等の遮蔽による影響を受けにくいいため、信頼度が高いものとする。具体的には、図9に示すような、メッシュモデルの頂点 p_i に関する光線 r_i と、SFM法によって推定された各カメラ位置 c_{ij} との距離 l_{ij} を算出し、 l_{ij} が小さい順に N 個の撮影カメラ位置 $(c_{i1}, c_{i2}, \dots, c_{iN})$ とそれに対応する奥行き画像 $(DM_{i1}, DM_{i2}, \dots, DM_{iN})$ を選択する。選択された N 個の撮影地点は、奥行き値 d の探索時において、光線方向 r_i と奥行き情報に対する光線との角度差が少ないため、遮蔽の影響が小さく、後述するテクスチャマッピングにおいても仮想視点位置と見え方が近いテクスチャを利用することができる。

3.3.3 メッシュモデルの初期化

自由視点画像生成の初回には、初期値として利用できるメッシュモデルが存在しないため、奥行き値 d の全探索によりメッシュモデルの初期化を行う。まず、

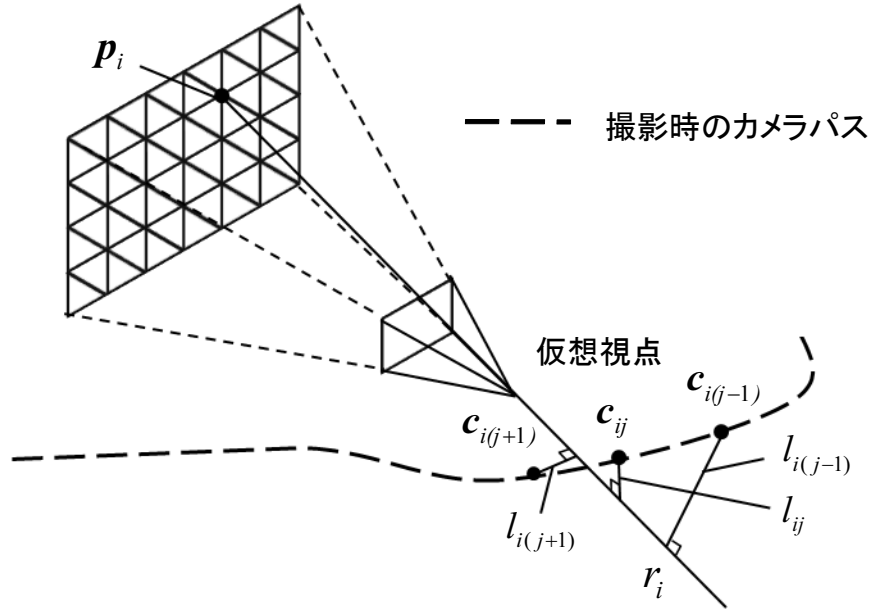


図9 各頂点に対する光線と撮影カメラ位置からの距離

メッシュモデル初期化の前準備として，図6に示すようにユーザが設定した初期仮想視点位置の前面の一定の奥行きのところ平面状の三次元メッシュモデルを設定する．次に，各頂点 p_i をそれぞれエネルギー $E_i(d)$ が最小となる座標へ移動させる．具体的には， p_i に対して光線 r_i 上で仮想視点位置からの距離 d を d_{min} から d_{max} の範囲内で変化させ，エネルギー $E_i(d)$ が最小となる距離 d_{EMIN} を決定する．すなわち，各頂点に対して以下の式を満たす d_{EMIN} を決定し， p_i を $(\frac{X_{p_i}}{Z_i}d_{EMIN}, \frac{Y_{p_i}}{Z_i}d_{EMIN}, d_{EMIN})$ に移動することでメッシュを変形する．

$$d_{EMIN} = \operatorname{argmin}_{d_{min} < d < d_{max}} E_i(d) \quad (4)$$

3.3.4 エネルギー最小化によるメッシュモデルの変形

ユーザの操作による視点移動・視線方向の変化量は生成されるフレーム間で微小であるという仮定の下で，移動前の仮想視点に対して生成されたメッシュモデルを初期値として用い，限定された探索範囲内でのエネルギー最小化処理を行う

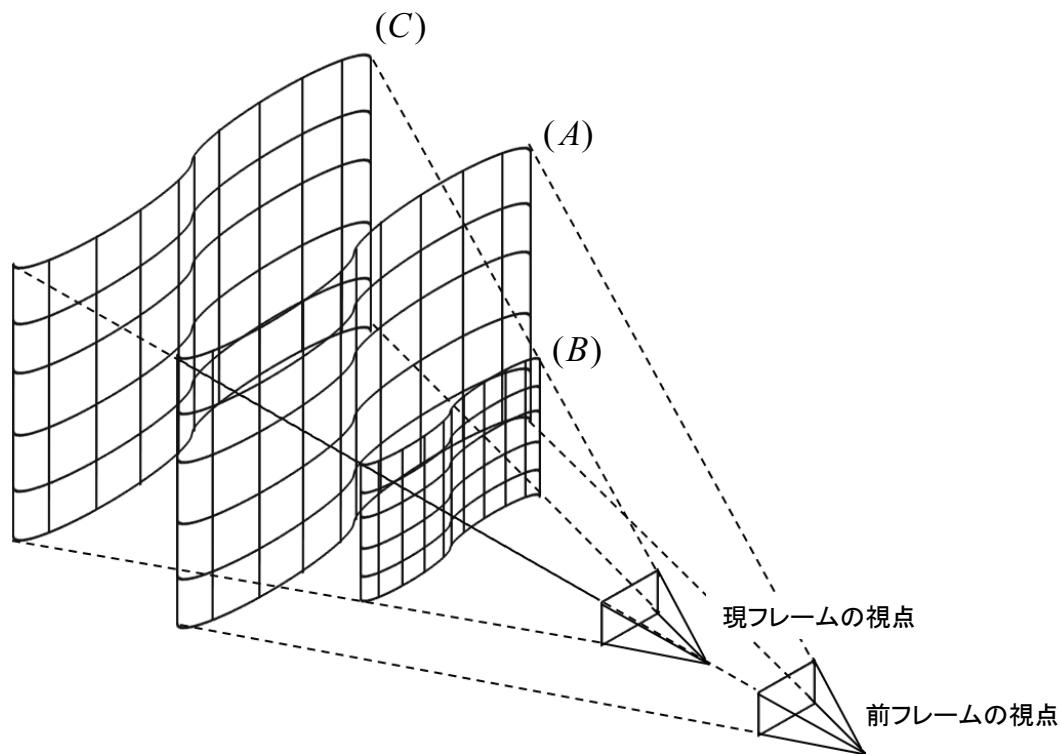


図 10 視点移動後のメッシュモデルの変形

ことで、視点移動時のメッシュモデルの生成処理の高速化および視点移動・視線方向の変化に伴う急激なモデルの変化を抑制する。

図 10 はユーザの操作による仮想視点移動時の探索範囲を表している。図 10 において (A) を視点移動前のメッシュモデルとすると、視点移動後のモデル生成時の d の探索範囲を (A) のモデルを中心とした (B) から (C) までの一定範囲内に制限する。以下では、探索範囲の具体的な決定方法について述べる。図 11 に示すように、まず現フレームの仮想視点の前面に設定した平面上の頂点 p_i に対して、頂点 p_i に対応する光線 r_i と前フレームで生成したモデルとの交点 \tilde{p}_i を求める。次に、現フレームの視点位置から交点 \tilde{p}_i までの奥行き値 \tilde{d} に対する一定の割合 $R_d (0 < R_d < 1)$ の範囲を探索範囲としてエネルギーが最小となる奥行き値 d_{EMIN}

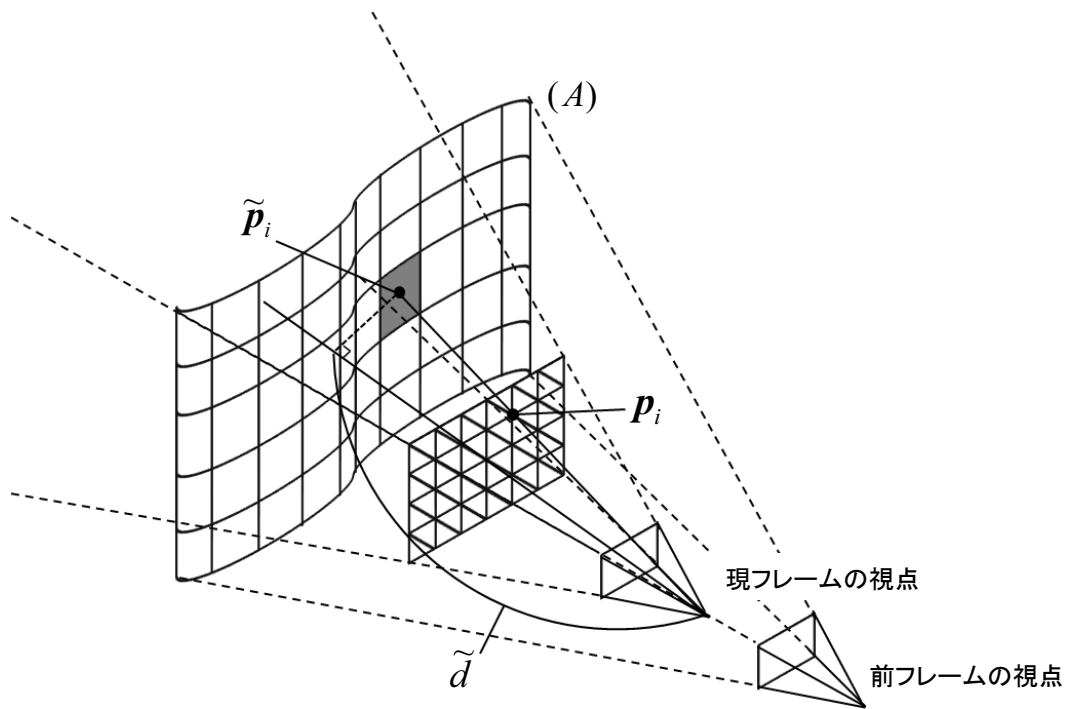


図 11 視点移動後の奥行き d の探索範囲

を探索する．

$$d_{E_{MIN}} = \underset{(1-R_d)\tilde{d} < d < (1+R_d)\tilde{d}}{\operatorname{argmin}} E_i(d) \quad (5)$$

ただし，上記の限定された探索範囲では，画像上にジャンプエッジが存在する場合には，正しく奥行き値を決定できない場合がある．従って，本研究では制限した探索範囲内においてエネルギーが事前に設定した閾値以下に収まらない場合には，その頂点に関して奥行き値 d の全探索を行い，エネルギーが最小となる距離 $d_{E_{MIN}}$ を探索する．

3.4 メッシュモデルへのテクスチャのマッピング

本節では，生成されたメッシュモデルへのテクスチャのマッピングについて述べる．ここではまず，3.3.2項で述べた手法により選択した全方位画像の撮影位置

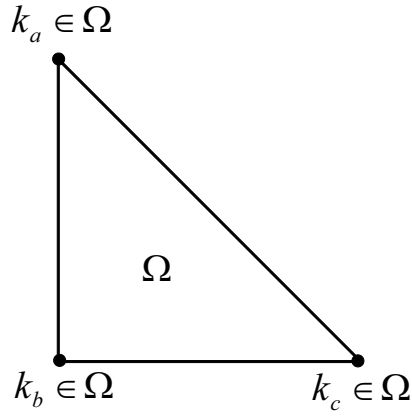


図 12 メッシュモデルの三角形パッチ Ω と三頂点

c_{ij} および各頂点においてエネルギーが最小となった時の重み $w(\hat{p}_i(d))$ を用いて、テクスチャマッピングに用いる画像の候補を選択する．次に、候補の中から最も優先度の高い画像を選択し、テクスチャマッピングを行う．具体的には、まず注目する三角形パッチ Ω の各頂点 $k(k \in \Omega)$ に対して、奥行き値 d の決定時に N 枚ずつ選択された撮影地点に対応する合計 $3N$ 枚の入力画像を選択する．ここで、選択された $3N$ 枚の画像のフレーム番号を $\phi = (f_1, f_2, \dots, f_{3N})$ とおく．これらの画像に対して、式 (6) に示すテクスチャの優先度 R_f を算出した上で、優先度 R_f の最も高いテクスチャを選択し、パッチ Ω にそのテクスチャをマッピングする．

$$R_f = \sum_{n=1}^{3N} \begin{cases} w_{f_n}(\hat{p}_i(d)) & ; (f_n = f) \\ 0 & ; otherwise \end{cases} \quad (6)$$

$3N$ 枚の画像中には、同一の撮影地点から撮影された画像が含まれる場合があり、そのようなテクスチャについては優先度 R_f は高くなる．また、 f_n フレームにおける撮影地点 c_{ij} の重み $w_{f_n}(\hat{p}_i(d))$ が大きいほどテクスチャの優先度 R_f が高く設定される．

4. 実験

4.1 全方位動画像からの自由視点画像生成実験

提案手法により，視点位置に応じて適切に変化する三次元メッシュモデルを用いて，全方位動画像から自由視点画像を生成できることを検証するために，屋外実環境を対象とした自由視点画像生成実験を行った．

4.1.1 入力情報の取得

全方位動画像を取得するカメラとして，図 13 に示す複数のカメラを外向きに配置した Point Grey Research 社製の全方位型マルチカメラシステム Ladybug を用いた．Ladybug では横 768 画素，縦 1024 画素の画像を水平方向 5 台および真上方向 1 台の計 6 台の CCD カメラで同期して取得する．また，全方向の約 75 % 以上の領域を 30fps の動画像として取得することが可能である．

本実験では，Ladybug を用いて移動しながら屋外環境を撮影した全方位動画像 (500 フレーム，3000 枚) を入力として用いた．図 14 は本実験で入力として用いる全方位動画像の 1 フレームである．ここではまず，この全方位動画像に対して 3.2.1 項で述べた SFM 法を適用することで，図 15 に示す全方位カメラのカメラ位置・姿勢を推定した．次に，3.2.2 項で述べたマルチベースラインステレオ法を適用することで，図 16 に示す特徴点の奥行き値を算出した．図 17 は，図 16 に対してドロネーの三角形分割法を用いた上で特徴点の間の奥行き値を補間した奥行き画像である．以下では，これら全方位動画像，それに対応する全方位奥行き動画像，カメラ位置・姿勢を入力として用い，自由視点画像生成を行う．



図 13 全方位型マルチカメラシステム Ladybug

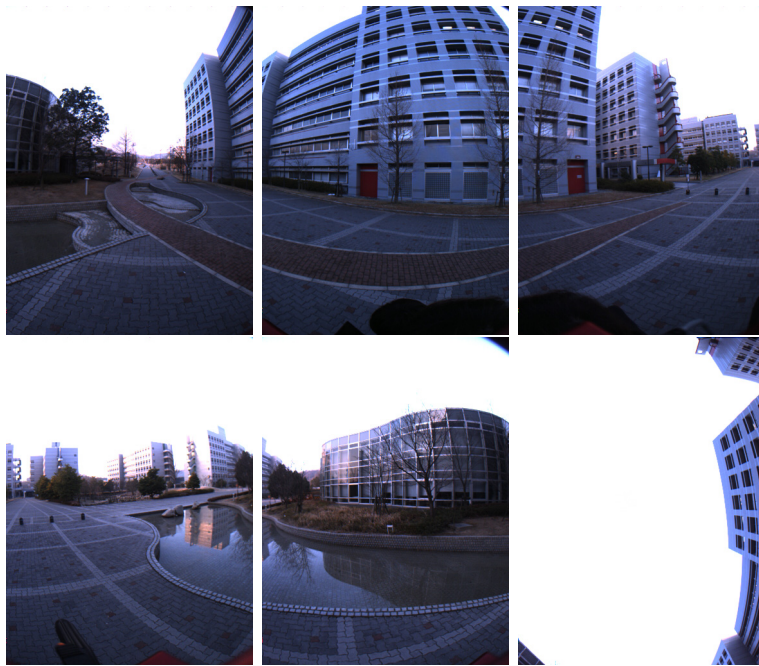
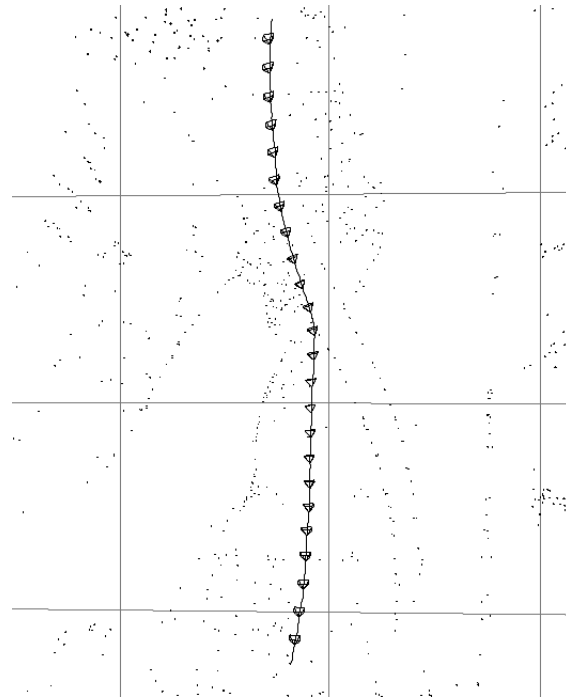
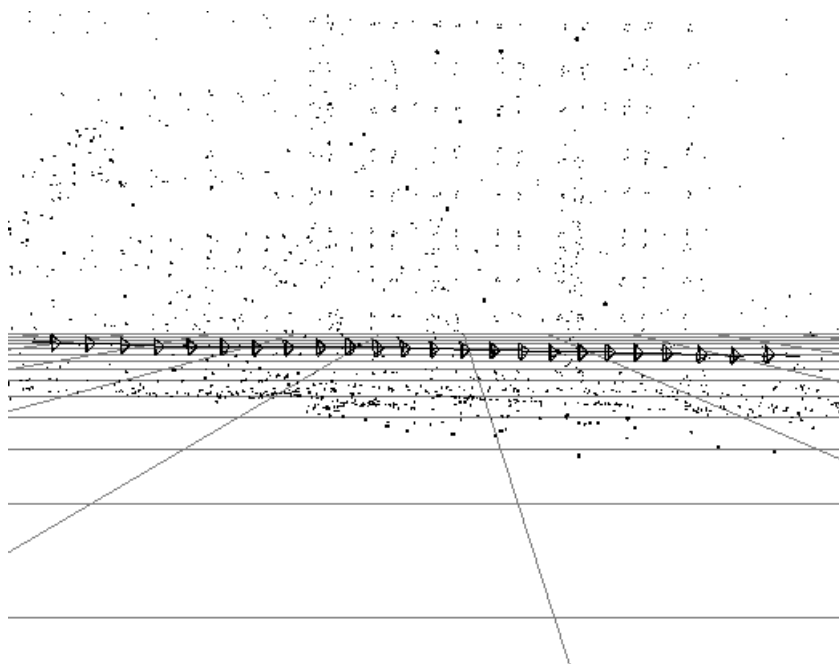


図 14 入力として用いた全方位動画像の1フレーム



(a) 上空から見たカメラ位置・姿勢



(b) 地上から見たカメラ位置・姿勢

図 15 SFM 法によって推定されたカメラ位置・姿勢

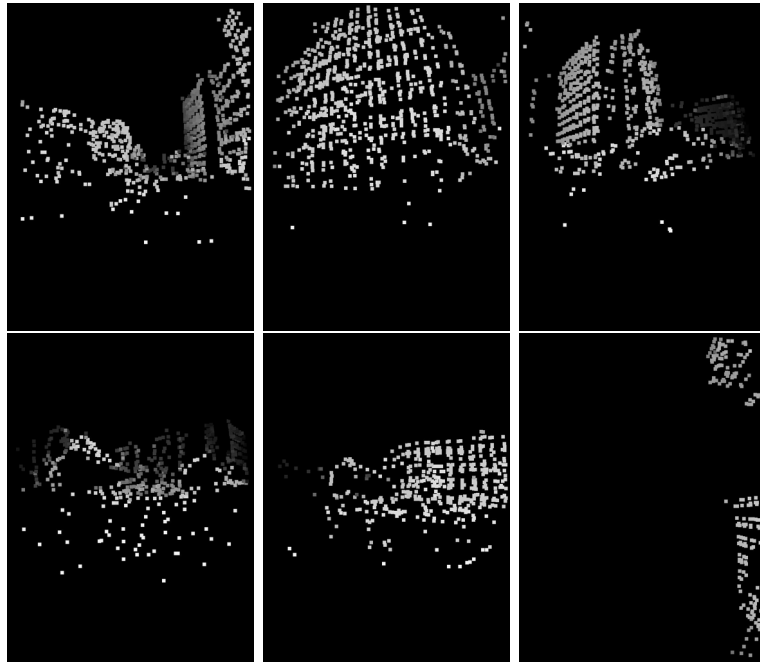


図 16 マルチベースラインステレオ法により推定された全方位奥行き画像の一例

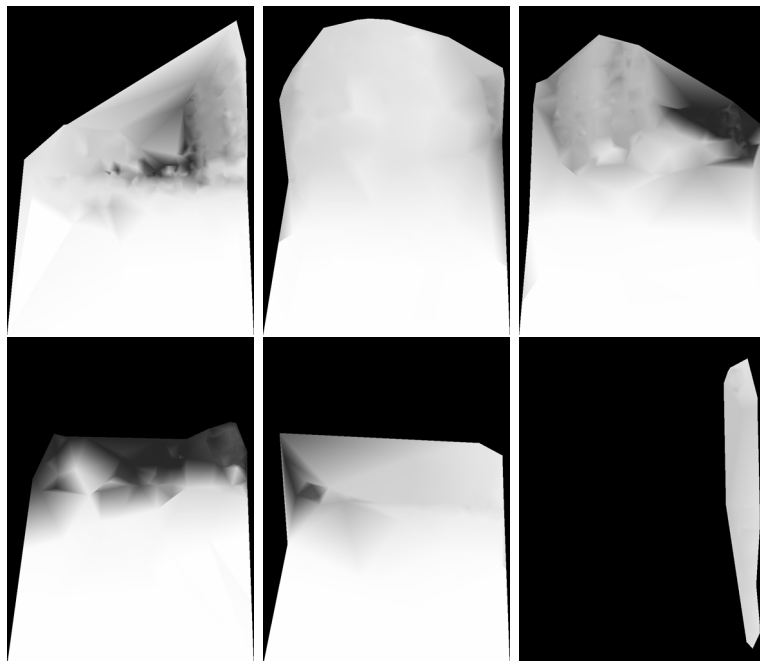


図 17 ドロネーの三角形分割法により補間された全方位奥行き画像の一例

4.1.2 自由視点画像生成

4.1.1 項で示した全方位動画像と推定されたカメラパラメータ，全方位動画像に対応する全方位奥行き画像群を入力として，表 2 に示すパラメータを用いて自由視点画像を生成した．ここでは，図 18 に示すように，入力画像取得時のカメラパスに沿って移動する A ルートと，そのカメラパスから離れるように移動する B ルートの二つのルートに対して，それぞれ地点 S から 2[m] おきに進行方向 (b) と進行方向に対して 30°視線方向を回転させた方向 (a)(c) の三方向の自由視点画像画像を生成した．

図 19，図 20 にルート A，B に対応する自由視点画像の生成結果を示す．また，図 21，図 22 にルート A，B の自由視点画像の生成時に用いられたメッシュモデルに対応する奥行き画像を示す．また，表 3 に使用した計算機のスペックを，表 4 に自由視点画像の生成にかかった平均処理時間を示す．

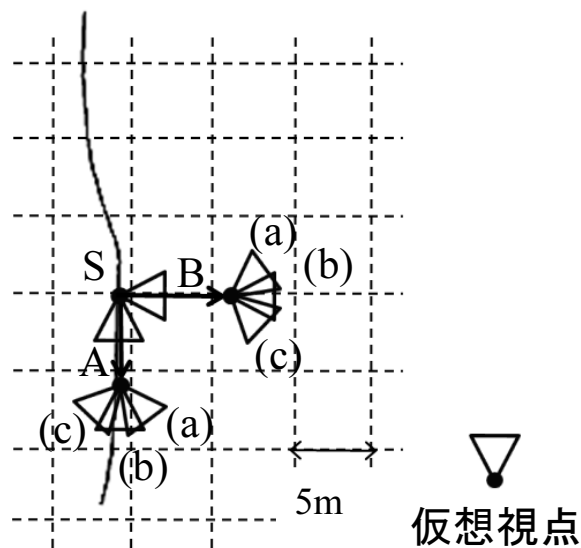


図 18 自由視点画像の生成視点位置・視線方向

表 2 本実験のパラメータ

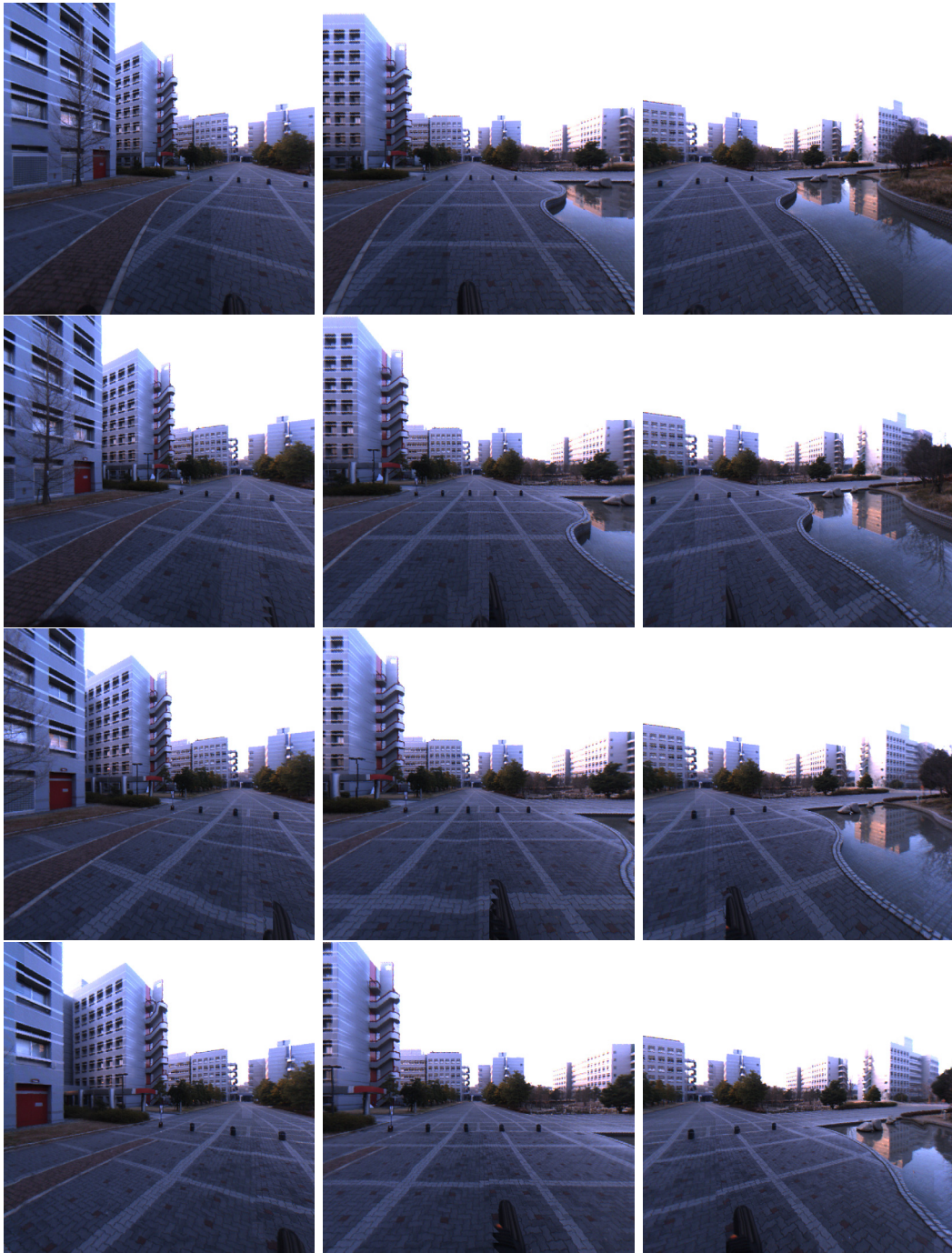
自由視点画像の解像度	横 800[pixel] 縦 800[pixel]
メッシュモデルの頂点数	961 点 (横 31 点 × 縦 31 点)
奥行き画像の選択枚数 N	5 枚
奥行き値の整合性の判定に必要な 奥行き画像の最低枚数 N_t	2 枚
探索範囲の限定に用いる 奥行き値の初期値 \tilde{d} に対する割合 R_d	0.5

表 3 使用した計算機のスペック

CPU	Intel(R) Core(TM)2 Duo E8600 3.33 GHz
Memory	16 GB
グラフィックスボード	NVIDIA GeForce 9800GT VRAM: 512 MB

表 4 自由視点画像の生成にかかる処理時間

	平均処理時間 [sec]
奥行き画像の選択	0.086
三次元メッシュモデルの初期化	4.190
三次元メッシュモデルの変形	1.183
テクスチャマッピング	5.340



方向 (a)

方向 (b)

方向 (c)

図 19 ルート A における自由視点画像
 (上から順に地点 S から 0m, 2m, 4m, 6m 離れた地点)

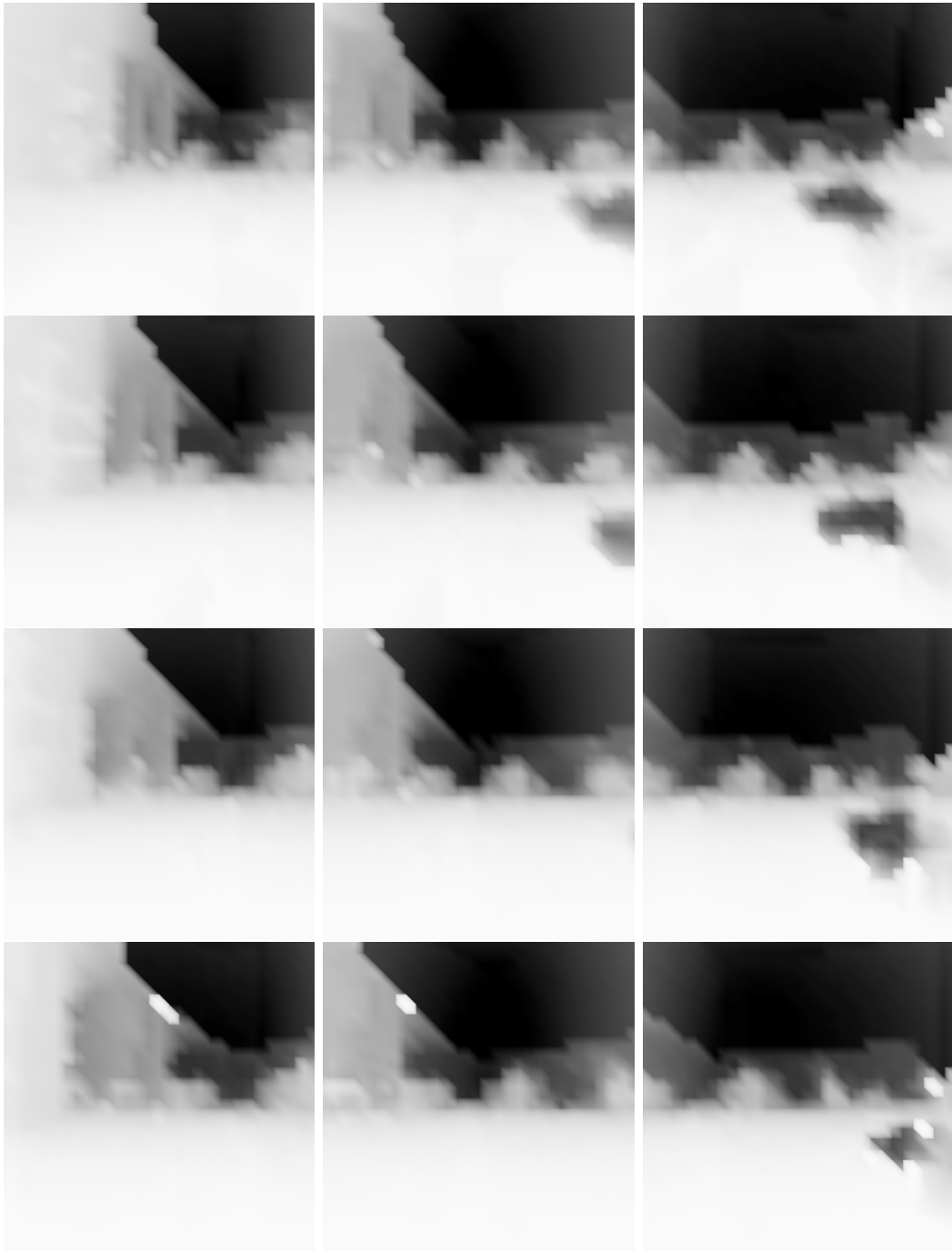


方向 (a)

方向 (b)

方向 (c)

図 20 ルート B における自由視点画像
 (上から順に地点 S から 0m, 2m, 4m, 6m 離れた地点)



方向 (a) 方向 (b) 方向 (c)

図 21 ルート A における自由視点画像に対応する奥行き画像
(上から順に地点 S から 0m, 2m, 4m, 6m 離れた地点)



方向 (a) 方向 (b) 方向 (c)
図 22 ルート B における自由視点画像に対応する奥行き画像
(上から順に地点 S から 0m, 2m, 4m, 6m 離れた地点)

4.2 考察

図 19, 図 20 より, 提案手法を用いることで欠損のない自由視点画像が生成できていることが確認できる。また, ルート A に対する生成結果 (図 19) より, 入力画像を取得した経路に沿って自由視点画像を生成した場合には, 品質の高い画像を生成できていることが分かる。しかし, 一部の地点において地面のテクスチャに不連続な部分が見られる。この原因としては, 生成されたメッシュモデル上で隣り合う三角形パッチにマッピングされるテクスチャの取得フレームが異なるためであると考えられる。これを解決するためには, 複数フレームのテクスチャをブレンディングして用いるなどの工夫が必要である。また, ルート B に対する生成結果 (図 20) より, この実験環境では撮影地点から 2m 以内の範囲であれば, 大きな違和感のない画像が生成できていることが分かる。従って, 提案手法は撮影地点付近での自由視点画像生成に対しては有効な画像を生成できると考えられる。しかし, 撮影地点から 4m 以上離れた場合には, 仮想視点が建物に近づくにつれて, 樹木がある領域の歪みが大きくなっている。これは, 図 22 から分かるように, 生成に用いたメッシュモデルの分割数が少ないことに起因して, 建物のテクスチャが樹木に対応する奥行きの方にマッピングされているためである。これを解決するためには, 対象環境の複雑さを考慮してメッシュの分割数を適応的に変化させることが考えられる。

一方, 処理時間について, ユーザによる仮想空間内での自由な視点移動, 見回しを実現するためには実時間処理が必要であるが, 現時点ではメッシュモデルの変形に約 1.2 秒, また, テクスチャマッピングに約 5.3 秒を要している。モデルの生成に関して, マルチスレッド処理や GPU を用いた実装上の工夫を行うことで高速化を行い, 最終的には実時間処理を実現する必要がある。また, テクスチャのマッピング処理についても, 予め VRAM 上に使用するテクスチャを全て読み込んでおくなど実装上の工夫による大幅な処理速度向上の余地が残されている。

5. まとめ

本論文では，単一の三次元モデルを生成せず，視点位置・視線方向に応じて適切な三次元メッシュモデルを逐次生成することで，自由視点画像を生成する手法を提案した．提案手法では，全方位カメラにより対象環境を移動しながら撮影し，撮影した画像群に対して Structure from Motion 法 [19] を適用することで全方位カメラの運動パラメータを推定し，更にマルチベースラインステレオ法 [13] を用いて各撮影画像に対応する奥行き画像を生成する．以上を入力として，仮想視点の前面に設定したメッシュモデルを，奥行き情報の整合性に関するエネルギーを最小化することで適切な形状に変形させ，信頼度の高いテクスチャをマッピングすることによって自由視点画像を生成する．

実験では，本学敷地内を対象として自由視点画像の生成を行うことにより，撮影範囲に対して比較的広範囲な視点移動が可能な仮想環境内でのウォークスルーが実現できることを示した．生成された自由視点画像は撮影地点付近では写実的に生成されており，撮影地点付近での自由視点画像生成には有効であると考えられる．しかし，撮影地点から遠く離れた場合には，生成画像上に歪みが生じることから，より広範囲の移動に対応可能なウォークスルーシステムを実現するためには，今後さらなる改善が必要である．

今後の課題として，生成される自由視点画像の品質を向上させるために，環境の複雑さを考慮してメッシュの分割数を適応的に変化させることや，テクスチャのブレンディングを行うことなどが必要である．また，処理速度の向上のために，マルチスレッド処理や使用するテクスチャを予め読み込んでおくなどの実装上の工夫が必要である．

謝辞

本研究を進めるにあたり，その全過程において細やかな御指導，御鞭撻を頂いた視覚情報メディア講座 横矢 直和 教授に心より感謝申し上げます．また，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いたインタラクティブメディア設計学講座 加藤 博一 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，始終暖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 准教授に深く感謝申し上げます．さらに，研究活動の全過程を通して多大なる御助言，御指導を頂いた視覚情報メディア講座 佐藤 智和 助教に心より感謝致します．特に，佐藤 智和 助教には本研究のテーマの設定から本論文の執筆，その他発表論文の添削，発表練習に至るまで細やかな御指導，御鞭撻を頂き深く感謝致します．また，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感謝いたします．特に，本研究の遂行にあたり，多大なる御助力を頂いた視覚情報メディア講座 伊吹 拓也 氏に深く御礼申し上げます．最後に，研究室での生活を支えていただいた視覚情報メディア講座 高橋 美央 女史に心より感謝申し上げます．

参考文献

- [1] 佐藤宏介, 井口征士: “空間コード化による距離画像入力”, 電子情報通信学会論文誌 (D), Vol. J68-D, No. 3, pp. 369-375, 1985.
- [2] 塚本壮輔, 呉本堯, 古賀和利, 三池秀敏: “階層化位相シフト法による高精度な奥行き計測”, 画像の認識・理解シンポジウム 講演論文集, Vol. I, pp. 185-190, 2000.
- [3] 浅井俊弘, 神原誠之, 横矢直和, “全方位距離画像と全方位カラー画像の統合による屋外環境の三次元モデル化”, 画像電子学会誌, Vol. 34, No. 5, pp. 529-538, 2005.
- [4] H. Zhao and R. Shibasaki: “Reconstruction of Textured Urban 3D Model by Fusing Ground-Based Laser Range and CCD Images”, IEICE Trans. on Information and Systems, Vol. 2, pp. 562-569, 2003.
- [5] C. Frueh and A. Zakhor: “Constructing 3D City Models by Merging Ground-Based and Airborne Views”, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 2, pp. 562-569, 2003.
- [6] R. J. Woodham, “Photometric Method for Determining Surface Orientation from Multiple Images”, Journal of Optical Engineering, Vol. 19, No. 1, pp. 139-144, 1980.
- [7] K. Ikuchi: “Determining Surface Orientation of Specular Surfaces by Using The Photometric Stereo Method”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 3, No. 6, pp. 661-669, 1981.
- [8] J. Y. Zheng: “Acquiring 3-D Models from Sequences of Contours”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 16, No. 2, pp. 163-178, 1994.

- [9] W. Niem and J. Wingbermuehle: “Automatic Reconstruction of 3D Objects using a Mobile Monoscopic Camera”, Proc. Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, pp. 173-180, 1997.
- [10] H. Baker, “Three-Dimensional Modeling”, Proc. 5th Int. Joint Conf. on Artificial Intelligence, pp. 649-655, 1977.
- [11] V. J. Milenkovic and T. Kanade: “Trinocular Vision Using Photometric and Edge Orientation Constraints”, Proc. DARPA Image Understanding Workshop, pp. 163-175, 1985.
- [12] N. Ayache and F. Lustman: “Fast and Reliable Passive Trinocular Stereo Vision”, Proc. Int. Conf. on Computer Vision, Vol. 1, pp. 422-426, 1987.
- [13] 佐藤智和, 横矢直和: “画像特徴点の数え上げによるマルチベースラインステレオ法”, 情報処理学会論文誌, Vol. 48, No.SIG16, pp. 25-37, 2007.
- [14] R. Y. Tsai: “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision”, Proc. Computer Vision and Pattern Recognition, pp. 364-374, 1986.
- [15] M. Goesele, N. Snavely, B. Curless, H. Hoppe and S. M. Seitz “Multi-View Stereo for Community Photo Collections”, Proc. Int. Conf. on Computer Vision, 2007.
- [16] H. N. Nair and C. V. Stewart: “Robust Focus Ranging”, Proc. Computer Vision and Pattern Recognition, pp. 309-314, 1992.
- [17] S. K. Nayar: “Shape from Focus System”, Proc. Computer Vision and Pattern Recognition, pp. 302-308, 1992.
- [18] M. Pollefeys, R. Koch, M. Vergauwn, A. A. Deknuydt and L. J. V. Gool: “Three-Dimensional Scene Reconstruction from Images”, Proc. SPIE, Vol. 3958, pp. 215-226, 2000.

- [19] 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”, 信学論, Vol. J88-D-II, No. 2, pp. 189-196, 2005.
- [20] T. Mukai and N. Ohnishi: “The Recovery of Object hap and Camra Motion Using a Sensing system with a Video Camera and a Gyro Sensor”, Proc. I Int. Conf. on Computer Vision, Vol. 1, pp. 411-147, 1999.
- [21] 上原将文, 塩崎剛志, 全炳東: “数値地図と動画像解析による都市空間モデリング”, 画像の認識・理解シンポジウム論文講演集, Vol. 2, pp. 373-378, 2000.
- [22] 山下淳, 林本悦一, 金子透: “全方位カメラ搭載移動ロボットによる Structure from Motion を用いた 3次元環境モデリング”, 日本機械学会論文集, Vol. 73, No. 726, pp. 512-519, 2007.
- [23] S. E. Chen and L. Williams: “View Interpolation for Image Synthesis”, Proc. SIGGRAPH'93, pp. 279-288, 1993.
- [24] S. M. Seitz and C. R. Dyer: “View Morphing”, Proc. SIGGRAPH'96, pp. 21-30, 1996.
- [25] E. H. Adelson and J. Bergen: “The Plenoptic Function and the Elements of Early Vision”, MIT Press Cambridge, MA, pp. 3-20, 1991.
- [26] 苗村健, 原島博: “Ray-Based Creation of Photo-Realistic Virtual World”, Virtual Reality and MultiMedia, pp. 59-68, 1997.
- [27] Marc Levoy and P. Hanranhan: “Light Field Rendering”, Proc. SIGGRAPH'96, pp. 31-42, 1996.
- [28] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen: “The Lumigraph”, Proc. SIGGRAPH'96, pp. 43-54, 1996.

- [29] 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和: “全方位動画像を用いた両眼ステレオ画像の実時間生成によるテレプレゼンス”, 電子情報通信学会論文誌, Vol. 84, No. 9, pp. 2048-2057, 2001.
- [30] T. Takahashi, H. Kawasaki, K. Ikeuchi and M. Sakauchi: “Arbitrary View Position and Direction Rendering for Large-Scale Scenes”, Proc. Computer Vision and Pattern Recognition, Vol. 2, pp. 296-303, 2000.
- [31] I. Kitahara and Y. Ohta, “Scalable 3D Representation for 3D Video Display in a Large-Scale Space”, Proc. Virtual Reality'03, pp. 45-52, 2003.
- [32] 冨手要, 山澤一誠, 横矢直和: “複数の全方位画像を用いた広範囲なウォークスルーの実現”, 画像の認識・理解シンポジウム 講演論文集, Vol. II, pp. 353-358, 2002.
- [33] 篠原秀俊, 八木康史, 谷内田正彦: “IBR と MBR を併用した空間モデリングとウォークスルー環境”, 情報処理学会 研究報告, No. 2, pp. 147-154, 2003.