

修士論文

視点位置に依存して変形する三次元メッシュモデル
を利用した自由視点画像生成における違和感の低減

伊吹 拓也

2009年2月4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。

伊吹 拓也

審査委員：

横矢 直和 教授 (主指導教員)

千原 國宏 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

佐藤 智和 助教 (副指導教員)

視点位置に依存して変形する三次元メッシュモデル を利用した自由視点画像生成における違和感の低減*

伊吹 拓也

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など、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、人手によるモデリング作業には限界があり、広範囲の屋外環境を対象とすることは困難である。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、Model-Based Rendering(MBR)法や、Image-Based Rendering(IBR)法、MBR法とIBR法を組み合わせたHybrid Rendering法が提案されている。しかし、MBR法では複雑な形状を持つ物体の幾何形状を正確に推定することが困難であり、また、オクルージョンにより遮蔽される領域に欠損が生じるという問題がある。IBR法では、生成される画像上に現れる幾何学的な歪みが大きいという問題がある。一方、Hybrid Rendering法では、視点位置に依存した幾何形状を用いて自由視点画像を生成することで、MBR法の問題であるモデルの欠損を解消し、比較的広範囲における写実的な自由視点画像生成を可能にしている。しかし、従来手法では、幾何学的な整合性のみを考慮した画像合成を行っているため、用いられる幾何情報の精度が十分でない場合、テクスチャの不連続な変形や位置ずれに起因する歪みが生じ、合成画像上に違和感が生じていた。本研究では、従来のHybrid Rendering法が用いてきた幾何学

* 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-MT0851011, 2009年2月4日.

的な整合性に加えて、人間が知覚しやすい違和感の要因に着目し、視点位置に依存して変形する幾何モデルの時間的・空間的な連続性およびテクスチャの幾何学的・光学的な連続性を考慮した自由視点画像生成を行う。これにより、合成映像上における違和感の要因を取り除き、品質の高い自由視点画像を生成する。具体的には、幾何モデルの時間的・空間的な連続性を保つために、幾何モデルの変形に用いるエネルギー関数に、モデルの連続性に関する制約項を新たに追加することで、幾何モデルの不連続な形状変形を抑制する。また、画像上で隣り合うパッチ間に対して、より近い位置で撮影された画像がテクスチャとして用いられるようテクスチャの選択処理を改良することで、合成画像上におけるテクスチャの幾何学的・光学的な不連続を抑制する。実験では全方位マルチカメラ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、生成した自由視点画像を従来手法と比較することで、提案手法の有効性を示す。

キーワード

自由視点画像生成, 全方位動画像, 奥行き画像

Reduction of Unnatural Feeling in Free-viewpoint Rendering Using View-Dependent Deformable 3-D Mesh Model*

Takuya Ibuki

Abstract

This thesis describes a method to reduce an unnatural feeling in free viewpoint images generated by using a deformable 3-D mesh model. For free viewpoint rendering, three kinds of approaches have already been investigated: Model-based rendering (MBR), image-based rendering (IBR) and hybrid rendering that is a hybrid approach of IBR and MBR. However, it is still difficult for conventional methods to generate natural images for wide and complex outdoor environments. In MBR, constructed 3-D models often have holes that are caused by occlusions and errors in 3-D reconstruction. In IBR, rendered images are easily distorted when the viewpoint is set far from the capturing positions of the input images. Although hybrid rendering relaxes such problems using view-dependent deformable 3-D models, discontinuous transformation and misalignment of textures have been still observed in generated images. In this thesis, in order to reduce unnatural feeling in generated images, we employ not only the geometric consistency but also smoothness constraints on spatial and temporal mesh deformation. Additionally, generation of discontinuous textures are suppressed by improving the strategy of texture selection for each patch in 3-D model. In experiments, the

* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0851011, February 4, 2009.

effectiveness of the proposed method is demonstrated by comparing with previous method.

Keywords:

Free-viewpoint Rendering, 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	4
2.1 自由視点画像生成に関する従来研究	4
2.1.1 Model-Based Rendering	4
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	13
3. 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からの自由視点画像生成	15
3.1 全方位動画からの自由視点画像生成手法の概要	15
3.2 入力情報の取得	17
3.2.1 カメラ位置・姿勢情報の推定	17
3.2.2 奥行き情報の推定	18
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	19
3.3.1 複数の奥行き情報の整合性に関するエネルギー関数の定義	21
3.3.2 メッシュモデルの形状変形における時空間的な連続性を考慮したエネルギー関数の拡張	23
3.3.3 奥行き画像の選択	25
3.3.4 メッシュモデルの生成・変形	26
3.4 メッシュモデルへのテクスチャマッピング	27
3.4.1 優先度を用いたテクスチャの決定	27
3.4.2 隣り合うパッチに対するテクスチャの連続化	29
3.4.3 GPUを用いたテクスチャマッピングの実装	30
4. 実験	32
4.1 全方位動画からの自由視点画像生成実験	32

4.1.1	入力情報の取得	32
4.1.2	提案手法による全方位動画からの自由視点画像の生成と 従来手法との比較	36
4.2	提案手法の従来手法との比較による考察	49
5.	まとめ	50
	謝辞	51
	参考文献	52

目 次

1	レーザレンジファインダを用いた三次元モデルにおける欠損の例 [3]	6
2	Merrell らの手法 [16]	8
3	Seitz らの手法 [19]	9
4	Irani らの手法 [28]	12
5	越澤らの手法 [30] における合成画像上の不連続な歪み ((a) は t フレーム目における生成画像, (b)~(d) は $t \sim t+3$ フレーム目の連続する生成画像上の (a) の赤枠部分を拡大した画像)	12
6	越澤らの手法 [30] におけるテクスチャの幾何学的・光学的な不連続	14
7	本手法による自由視点画像生成の流れ	16
8	三次元メッシュモデルの初期設定	20
9	メッシュモデルの各頂点の座標	20
10	複数の奥行き情報の整合性に関するエネルギーの算出に用いられる変数	22
11	前フレームで生成したメッシュモデルと仮想視点の位置関係	24
12	各頂点に対する光線と撮影カメラ位置からの距離	25
13	奥行き値の更新時における各頂点のグループ	27
14	メッシュモデルの三角形パッチ Ω と三頂点	28
15	テクスチャ決定時の光線とカメラパスの位置関係	29
16	GPU を用いたテクスチャマッピングの処理の流れ	31
17	全方位型マルチカメラシステム Ladybug	33
18	入力として用いた全方位動画の 1 フレーム	33
19	SFM 法によって推定されたカメラ位置・姿勢	34
20	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	35
21	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	35
22	自由視点画像の生成視点位置・視線方向	37
23	(a) 地点 (75 フレーム目) における自由視点画像	39
24	(a) 地点 (75 フレーム目) における自由視点画像の拡大画像	40

25	(b) 地点 (190 フレーム目) における自由視点画像	41
26	(b) 地点 (190 フレーム目) における自由視点画像の拡大画像 . . .	42
27	(c) 地点 (234 フレーム目) における自由視点画像	43
28	(c) 地点 (234 フレーム目) における自由視点画像の拡大画像 1 . .	44
29	(c) 地点 (234 フレーム目) における自由視点画像の拡大画像 2 . .	45
30	(a) 地点 (75 フレーム目) における自由視点画像に対応する奥行き 画像	46
31	(b) 地点 (190 フレーム目) における自由視点画像に対応する奥行 き画像	47
32	(c) 地点 (234 フレーム目) における自由視点画像に対応する奥行き 画像	48

表 目 次

1	本実験のパラメータ	37
2	使用した計算機のスペック	38
3	自由視点画像の生成にかかる平均処理時間	38

1. はじめに

近年, Microsoft 社の Virtual Earth, Google 社の Google Earth に代表される, 現実環境を計算機上に再現し, 再現された環境を仮想的に体験することが可能なサービスが注目を集めている. このような, 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への応用が期待されている. 現在, 実環境を仮想空間内に再現する手段として, モデリングソフトを用いて人手により作成された CAD データに対し, 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, 広域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や, 人手による作成では再現できる景観の写実性に限界があることが問題とされている. そこで, 予め取得した画像群を利用して, 屋外環境を自動的に仮想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) [1–17] 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) [18–25] 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. しかし, 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. そこで, MBR 法と IBR 法を組み合わせた Hybrid Rendering 法 [26–30] が提案されている. これらの手法は, MBR 法におけるモデルの欠損の問題を解消しつつ, 一定の品質を保ったままでの大幅な視点移動を可能にしている. しかし, 従来の Hybrid Rendering 法においては, フレームごとに独立に画像を生成するため, 動

画として生成された画像を再生した場合において、違和感が生じるという問題があった。加えて、複雑なシーンを対象とした場合において、合成に用いられる幾何情報を正しく推定することが難しく、このような箇所においてテクスチャの歪みが生じていた。

このように、従来、仮想環境を再現するための様々な手法が提案されているが、複雑な広域屋外環境を対象とした場合において、現実環境を高品位に再現することは困難であり、合成画像上に歪みや欠損などの違和感が生じているのが現状である。そこで、本論文では、従来の Hybrid Rendering 法が用いてきた幾何学的な整合性に加えて、人間が知覚しやすい違和感の要因に着目し、それらを取り除くことで、生成される自由視点画像に生じる違和感を低減する手法を提案する。提案手法は、比較的広範囲の自由視点画像を高速にレンダリング可能な越澤らの手法 [30] を基礎とし、メッシュモデルの時間的・空間的な連続性とテクスチャの幾何学的・光学的な連続性を考慮することで、合成画像上に生じる違和感を低減させ、より高品位な自由視点画像を生成する。具体的には、幾何モデルの時間的・空間的な連続性を保つために、幾何モデルの変形に用いるエネルギー関数にモデルの連続性に関する制約項を新たに追加することで、幾何モデルの不連続な形状変形を抑制する。提案手法では、まず、仮想視点に対する視線方向から、入力として用いる複数の奥行き画像を決定する。次に、決定された複数の奥行き画像から得られる奥行き情報を用い、視点の前面に設定したメッシュモデルの各頂点を、エネルギー関数に基づいて最適な位置に移動させることで、視点位置に応じたメッシュモデルを作成する。最後に、生成されたメッシュモデルに対して、視線方向およびパッチ間で利用される画像の撮影地点の位置関係を考慮して決定したテクスチャをマッピングすることで、合成画像上におけるテクスチャの幾何学的・光学的な不連続を抑制し、高品位な自由視点画像を生成する。

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づけについて述べる。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し、従来手法との比較を行う。最後に第

5章で、まとめと今後の展望について述べる.

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、あらかじめ対象とするシーンを撮影した動画像を基に、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、(1) 対象環境の三次元形状を用いることで自由視点画像を生成する MBR 法 [1–17] や、(2) 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 [18–25]、(3) MBR 法と IBR 法を組み合わせた Hybrid Rendering 法 [26–30] の 3 つに分類できる。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。環境の形状を用いるため、視点移動に関する自由度や、視点移動に伴う物体の隠蔽関係の表現に優れている。また、仮想化された環境内に他の物体を追加表示することや、環境内の物体の配置を変更することが容易である。ただし、MBR 法で高品位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高精度に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精度良く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手法と、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手法の 2 つに大別される。いずれの手法においても、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出来ない。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い、統合する必要がある。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及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として、スリット光投影法や光レーザの飛行時間測定による手法が提案されている。スリット光投影法では、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出来る小物体に対しては、高精度な計測が可能である。しかし、屋外環境などの広域環境に対しては、ベースラインを十分に確保することが難しく、また日照条件によってスリット光の観測が困難となる。そのため、屋外広域環境に対してスリット光投影法を適用することは難しい。一方、光レーザの飛行時間測定による手法では、対象環境を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操作することで、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。このような手法で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、複数地点で撮影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 [3-5] が多数提案されている。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のようなモデルの欠損が生じてしまうという問題を抱えている。

[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、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



図1 レーザレンジファインダを用いた三次元モデルにおける欠損の例 [3]

なカメラの他に、一度に全方位の画像を取得出来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出来るという特徴がある。受動的な三次元復元手法としては、ステレオ法、特徴点の追跡による手法などがある。

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決定する手法である。三台のカメラを用いることで、対応点の曖昧さを減少させる三眼ステレオ法 [6, 7] 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 [8] が提案されている。一般的に、ステレオ法では各カメラの位置関係を把握することが重要であり、高精度なカメラ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 [9] が代表的であるが、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難しいという問題がある。この問題を回避するために、以下に述べる動画像上の特徴点を追跡する手法を用いることで、カメラキャリブレーションを行う手法が近年用いられている。

特徴点の追跡による手法は、Structure from Motion 法 (以下, SFM 法) と呼ば

れ、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ことによって、カメラの移動パラメータに加えて、特徴点の三次元位置を推定する。ただし、一般的に動画像上の特徴点のみを用いる手法では、撮影に用いた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が必要となる。そこで、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 [11,12] や、三次元位置が既知の基準点を用いる手法 [13] 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可能である。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 [13,14] が提案されている。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出来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、対象環境の撮影カメラの位置・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、三次元情報は粗にしか得られない。

特徴点の追跡による手法によって得られるカメラ位置・姿勢の推定結果を利用してステレオ法による三次元復元を行う手法として、広域な屋外環境に適用可能な手法が複数提案されている [15,16]。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の画像群からステレオを行う手法を提案している。また、Merrellら [16] は車載カメラを用いて撮影された画像群からステレオを行い、GPUを用いて高速に奥行き情報を統合することで、広域屋外環境の三次元形状をリアルタイムで推定する手法を提案している (図2)。これらの手法は、比較的広域な環境を対象に密な三次元復元を行っているが、能動的な手法と同様に、複雑な形状に対する推定誤差やオクルージョンに起因するモデル欠損の問題は解決されていない。モデルの推定誤差に起因する欠損を減少させるために、環境に対する強い仮定を用いる手法が提案されている。Furukawaら [17] は、対象となるシーンが世界座標系における $x-y, x-z, y-z$ 平面のいずれかに平行な平面の集合で構成されているという仮定をおくことで、室内空間等の人



図 2 Merrell らの手法 [16]

工的なシーンに対する高品位な三次元モデルの生成を実現した。しかし、このような仮定は木などの自然物を含んだ一般的なシーンに適用することは難しい。

このように、屋外広域環境のように複雑な形状を持つ物体を含む環境を欠損なくモデル化することは困難であり、そのような環境に対して本アプローチにより自由視点画像を生成することは難しい。

2.1.2 Image-Based Rendering

イメージベースドレンダリングは、対象環境の三次元形状を明示的に復元することなく、物体やシーンを撮影した画像群を入力とし、これらを変形・合成することにより自由視点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法では正確な幾何形状の推定が難しい複雑な形状を持つ木などの物体に対しても、欠損の生じない自由視点画像を生成可能である。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、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 [18,19] 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で対応点を与え、中間視点位置において対

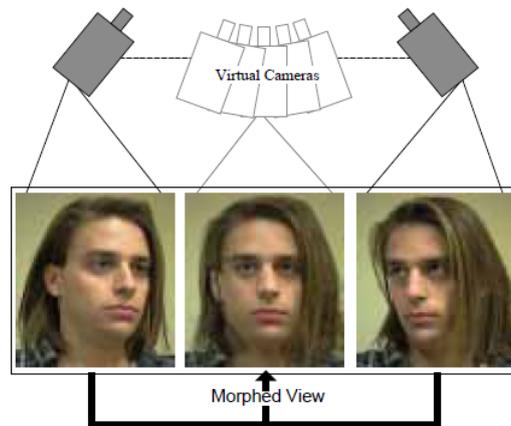


図 3 Seitz らの手法 [19]

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。さらに、中間視点位置における各画素の色を、対応点の位置関係より入力画像の色をブレンディングすることで求め、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出来るという特徴があるが、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学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]

対応点を必要としない手法として、Plenoptic Function [20], 光線空間法 [21], Light Field Rendering [22], Lumigraph [23] 等の光線情報を用いた手法が提案されている。これらの手法は、空間中を伝搬する光線を記録し、自由視点位置において視界画像面を通過する光線をリサンプリングすることで自由視点画像の生成を行う。Adelson ら [20] は Plenoptic Function を提案し、視点の座標 (x, y, z) , 光線の向き (θ_x, θ_y) , 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。ただし、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 [21] 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光線空間法と呼ばれる手法を提案している。光線空間法は、空間を伝搬する光線は減衰・変移しないと仮定することで、Plenoptic Function の 7 次元の光線情報から z 座標、波長 λ 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。Levoy ら [22] および Gortler ら [23] 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 [24, 25]。山口ら [24] 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 [25] 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り、一般的な計算機で処理を行うことは困難である。また、これらの光線情報を用いた手法においても、画像群の変形による手法と同様に、入力画像を撮影した地点から離れることにより発生する幾何学的な歪みを避けることは出来ない。

このように、IBR 法を用いた自由視点画像生成では、視点位置の移動に伴う幾何学的な歪みが大きく、IBR 手法単体において広域屋外環境を対象とした高品位な自由視点映像を実現することは難しい。

2.1.3 Hybrid Rendering

MBR 法と IBR 法を組み合わせた Hybrid Rendering による手法も多数提案されている。篠原ら [26] 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を生成している. この手法では, 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, また利用する手法が切り替わる箇所において, 生成される画像の品質が不連続に変化してしまうという問題がある. 冨手ら [27] は, 多地点で取得した全方位画像を用いて, それらの画像間の対応を与えることで, 対象環境の粗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. この手法は高速に自由視点画像を生成することが可能であるが, 予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ることは難しい.

一方, 視点位置に依存した形状モデルを利用する手法が提案されている. Irani ら [28] は, 自由視点からの光線を設定した上で, その光線上の点を複数の画像に投影した際に最も色が一致する点を探索し, その色をその光線の色とすることで自由視点画像を生成している (図 4). この手法は, 対象環境を密に撮影することなく, 比較的広範囲にわたり写実的な自由視点画像を生成可能という特徴がある. しかし, 自由視点決定後に, 全ての画素, 奥行きに対する photo consistency を算出する必要があるため, 計算コストが高く, 高速な処理を前提としているテレプレゼンスのようなアプリケーションには適用が難しい. 高橋ら [29] は, 二次元カメラアレイを用いて取得した多眼画像を元に, 対象環境の奥行き値を複数枚の平面で仮定した奥行きレイヤモデルを用いて推定することで, 自由視点画像の実時間生成を可能とする手法を提案している. ただし, 奥行き値の推定精度がレイヤモデルの分割数に依存するため, 奥行き方向の分解能が低いという問題がある. 越澤ら [30] は, 光線情報を考慮して決定された複数の撮影地点と, それらの撮影地点に対して予め算出済みの奥行き情報を用いることで, 最適な幾何形状とテクスチャを求める手法を提案している. この手法は奥行き値の分解能を低下させることなく, 予め求めておいた複数の奥行き情報を融合することで, 比較的高速なレンダリングを実現している. これらの視点位置に依存した形状モデルを利用する手法では, 視点位置が移動した際, 逐次的にモデルを生成しなおす必要があるが, 従来手法は, フレーム毎に独立に画像を生成するため, 図 5 に示すように動

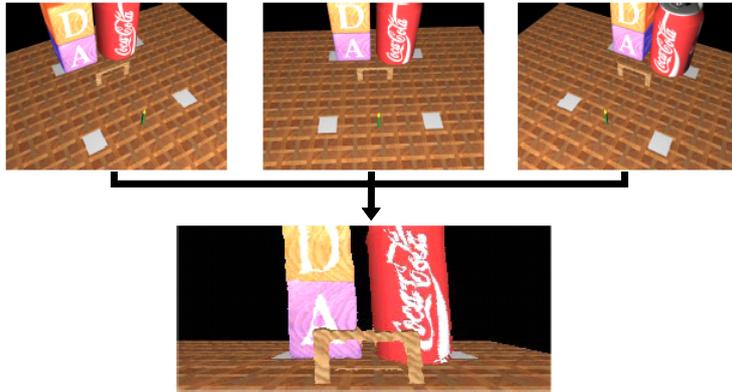


図 4 Irani らの手法 [28]

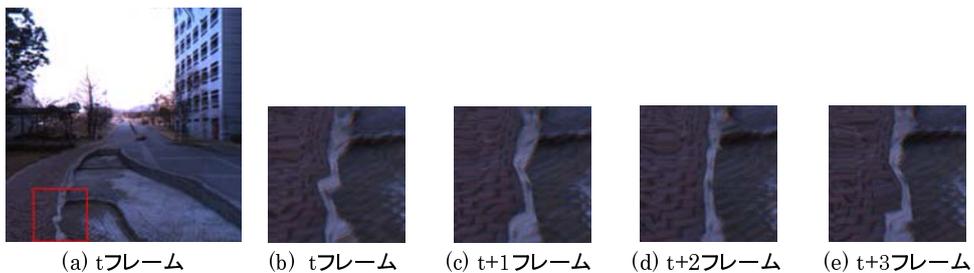


図 5 越澤らの手法 [30] における合成画像上の不連続な歪み ((a) は t フレーム目における生成画像, (b)~(d) は $t \sim t+3$ フレーム目の連続する生成画像上の (a) の赤枠部分を拡大した画像)

画として生成された画像を再生した場合において違和感が生じる場合がある。また、複雑なシーンを対象とした場合において、合成に用いられる幾何情報を正しく推定することが難しく、このような箇所においてテクスチャの歪みが生じるという問題も存在する。

このように、Hybrid Rendering 法を用いた場合においては、用いられる幾何情報の精度が十分でない場合、テクスチャの不連続な変形や位置ずれに起因する歪みが生じ、合成画像上に違和感が生じるという問題が残されている。

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高品位な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 3つの手法が存在する。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がある。IBR 法は対象環境の複雑さに依存せずに欠損のない自由視点画像を生成することが可能であるが、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、現実には広域環境全てを密に撮影することは難しく、IBR 法を用いて広域での自由視点画像を生成するのは難しい。MBR 法と IBR 法を組み合わせた Hybrid Rendering 法は様々な手法が提案されているが、その中でも、視点位置に依存した幾何モデルを利用する手法は、対象環境を密に撮影することなく、比較的広範囲にわたり写実的な自由視点画像の生成を可能としている。そのため、MBR 法や IBR 法や他の Hybrid Rendering 法と比べて、広域な環境を対象とした自由視点画像生成に適していると考えられる。ただし、従来手法はフレーム毎に独立に画像を生成するため、動画として生成された画像を再生した場合において、違和感が生じるという問題がある。また、複雑なシーンを対象とした場合において、合成に用いられる幾何情報を正しく推定することが難しく、このような箇所においてテクスチャの歪みが生じている。

以上のように、広範囲な視点移動を前提とした自由視点画像生成では、比較的テクスチャの歪みが小さく、欠損の生じない Hybrid Rendering による方法が望ましい。しかし、対象環境の幾何形状を正確に再現することは難しく、いずれの手法においても違和感のない自由視点画像を生成することは困難である。そこで、本研究では、従来手法のように仮想視点位置における正しい幾何形状の推定を行うことだけにとらわれず、各視点位置において視点移動に伴って生じる違和感が小さくなるような幾何形状を生成することで、合成画像上に生じる違和感を低減



図 6 越澤らの手法 [30] におけるテクスチャの幾何学的・光学的な不連続

させ、高品位な自由視点画像の生成を目指す。具体的には、視点位置に依存した幾何モデルを利用する手法の中でも、奥行き推定の分解能が高く、高速な処理を実現している越澤ら [30] の手法を基礎とし、従来用いられてきた幾何学的な整合性に加えて、メッシュモデルの時間的な連続性を考慮することで視点位置の変更に伴ってメッシュモデルが極端に変形することを抑制し、空間的な連続性を考慮することで、局所的な変形を抑制する。また、視線方向およびパッチ間で利用される画像の撮影地点の位置関係を考慮して決定したテクスチャをマッピングすることで、合成画像上におけるテクスチャの幾何学的・光学的な不連続 (図 6) を抑制する。なお、本研究では越澤らの手法と同様に、一度に全方向の情報を効率的に取得出来る全方位マルチカメラシステムを用いて対象環境を撮影し、入力画像とする。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 [13] を用いて推定し、マルチベースラインステレオ法 [8] を用いて全方位動画像に対する全方位奥行き画像生成する。これら全方位動画像、カメラ位置・姿勢、全方位奥行き動画像を入力として用い、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全方位 動画像からの自由視点画像生成

本章では、まずはじめに提案手法における自由視点画像の生成手法の概要を述べる。提案手法は、文献 [30] の手法を基礎とし、視点位置に応じた三次元メッシュモデルの生成時に用いる幾何形状の変形に関するエネルギー関数を拡張する。また、メッシュモデルにマッピングするテクスチャの決定手法に改良を加える。

以下ではまず、提案手法の処理の流れを概説し、本手法で用いる入力情報の取得方法について述べる。次に、提案手法が用いるエネルギー関数およびそれを用いたメッシュモデルの生成手法について述べ、最後にメッシュモデルへのテクスチャマッピング手法について詳述する。

3.1 全方位動画像からの自由視点画像生成手法の概要

図7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、まず対象環境および撮影時の情報などの入力情報を予め取得しておく。次に、これらの入力情報を用い、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の生成を行う。ユーザは視点位置・視線方向を操作することで、連続的に自由視点画像を生成する。

入力情報の取得 (A) は四つのステップから成る。まず、自由視点画像生成の対象環境を全方位カメラを用いて移動撮影する (a-1)。次に、移動撮影により得られた全方位動画像に対して SFM 法 [13] を適用することで、撮影時のカメラ位置・姿勢を取得し (a-2)、マルチベースラインステレオ法 [8] を適用することで、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 (a-3)。最後に、得られた特徴点の三次元位置を用い、奥行き情報を線形補間することで、対象環境の奥行き動画像を得る (a-4)。

自由視点画像生成処理 (B) 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 (b-1)。次に、ユーザが任意に

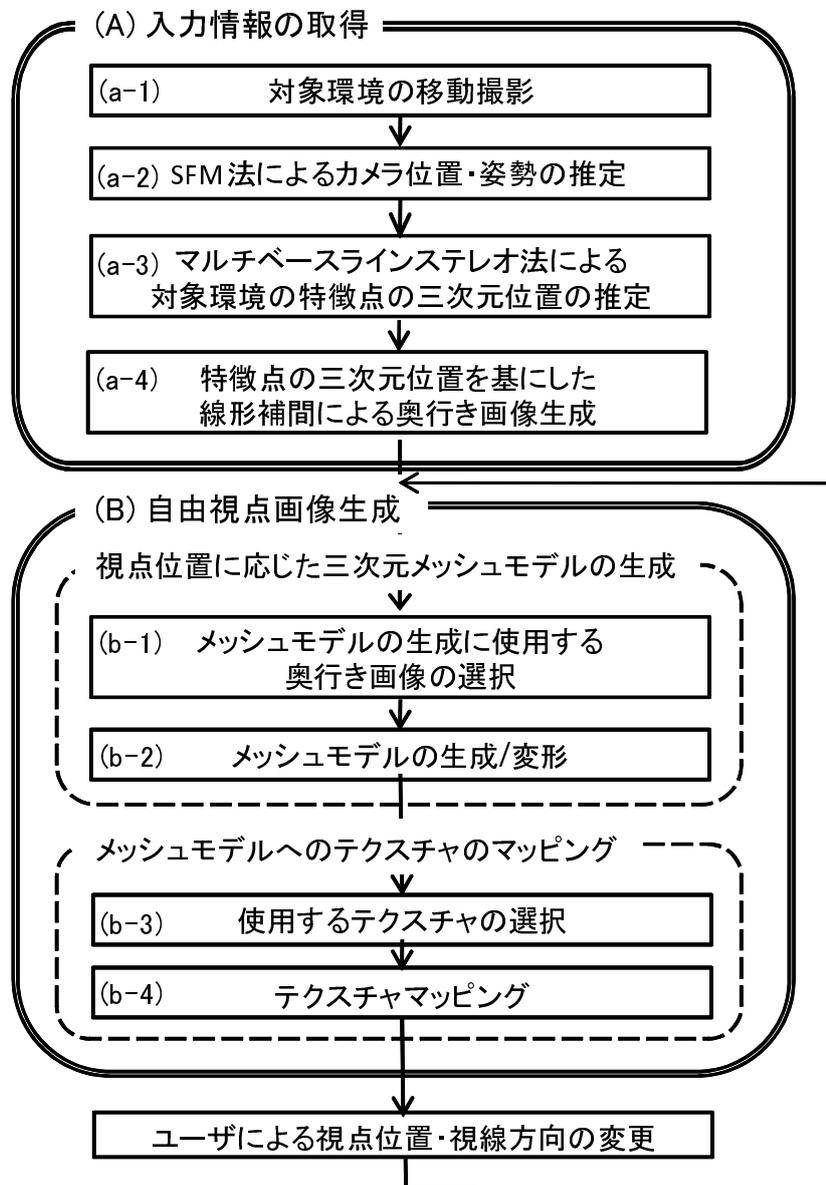


図 7 本手法による自由視点画像生成の流れ

設定した視点の前面にメッシュモデルを設定し、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て、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位置に移動させることで、三次元メッシュモデルの生成を行う (b-2). テクスチャのマッピングフェーズでは、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 (b-3), 各メッシュへのテクスチャマッピングを行う (b-4). 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、連続的に自由視点画像を生成する.

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 [13] 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 [13] 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 [8] 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、撮影された動画に対して以下の処理を行うことで、全方位動画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 [8] 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が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 [8] 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三次元メッシュモデルを利用した全方位動画画像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 8 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。具体的には、図 9 に示す仮想視点位置 \mathbf{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 \mathbf{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、各頂点におけるメッシュモデルの形状変形に関するエネルギー $E_i(d)$ を定義し、 \mathbf{p}_i と \mathbf{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 $\hat{\mathbf{p}}_i(d) = (\frac{X_{p_i}}{Z_i}d, \frac{Y_{p_i}}{Z_i}d, d)$ に頂点 \mathbf{p}_i を移動させる。以上の処理を全ての頂点に対して繰り返すことで、メッシュモデル全体のエネルギー E が最小となるメッ

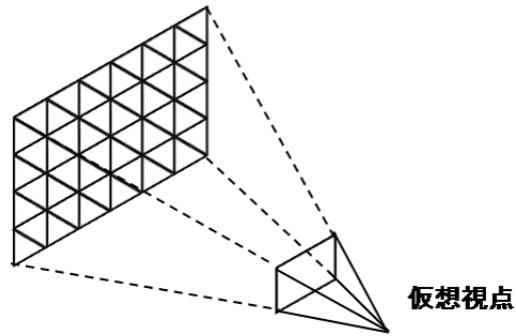


図 8 三次元メッシュモデルの初期設定

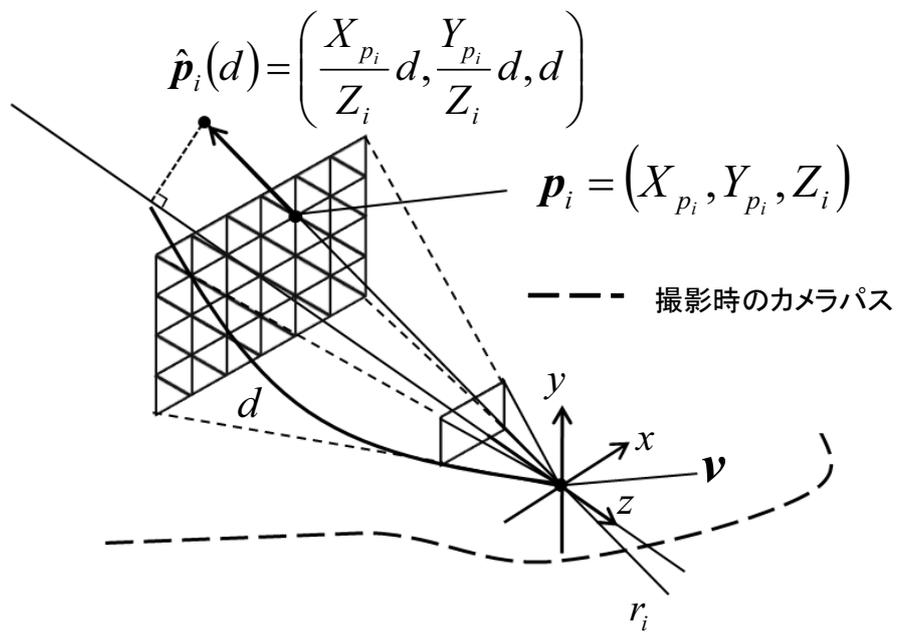


図 9 メッシュモデルの各頂点の座標

シュモデルを生成する。メッシュモデル全体のエネルギー E と各頂点での要素エネルギー $E_i(d)$ の関係は式1のように表される。ここで、 M はメッシュモデルの全ての頂点を示す集合とする。

$$E = \sum_{i \in M} E_i(d) \quad (1)$$

以下では、まず既存手法におけるメッシュモデルの形状変形に関するエネルギー関数を概説した上で、本手法におけるエネルギー関数の拡張について詳述する。次に、モデルの変形に用いる奥行き画像の選択について述べ、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。

3.3.1 複数の奥行き情報の整合性に関するエネルギー関数の定義

本節では、越澤ら [30] が提案した、複数の奥行き情報の整合性に関するエネルギー関数について概説する。文献 [30] では、メッシュモデルの形状変形に関するエネルギー関数を、複数の撮影地点から得られる奥行き情報の整合性に基づいて、以下のように定義している。

$$E_i(d) = E_g(d) = \frac{\sum_{j=1}^N w_j(\hat{\mathbf{p}}_i(d)) D_{ij}(d)}{\sum_{j=1}^N w_j(\hat{\mathbf{p}}_i(d))} \quad (2)$$

$$D_{ij}(d) = \begin{cases} \{e(\mathbf{c}_{ij}, \hat{\mathbf{p}}_i(d)) - |\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}|\}^2 & ; (e(\mathbf{c}_{ij}, \hat{\mathbf{p}}_i(d)) > |\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}|) \\ 0 & ; otherwise \end{cases} \quad (3)$$

$$w(\hat{\mathbf{p}}_i(d)) = \begin{cases} \theta_{ij}(d)^{-1} = \arccos\left(\frac{(\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}) \cdot (\hat{\mathbf{p}}_i(d) - \mathbf{v})}{|\hat{\mathbf{p}}_i(d) - \mathbf{c}_{ij}| |\hat{\mathbf{p}}_i(d) - \mathbf{v}|}\right)^{-1} & ; (D_{ij}(d) \neq 0) \\ 0 & ; otherwise \end{cases} \quad (4)$$

エネルギー関数内の変数の関係を図10に示す。ここで、 N は奥行き情報の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 ($\mathbf{c}_{i1}, \mathbf{c}_{i2}, \dots, \mathbf{c}_{iN}$) の数である。 $e(\mathbf{c}_{ij}, \hat{\mathbf{p}}_i(d))$ は図10に示すように、 \mathbf{c}_{ij} を中心とする奥行き画像上における三次元位置 $\hat{\mathbf{p}}_i(d)$ の方向に対応する奥行き値を表す関数である。また、 $w(\hat{\mathbf{p}}_i(d))$ は頂点 \mathbf{p}_i に対して選択された撮影地点 \mathbf{c}_{ij} に関する

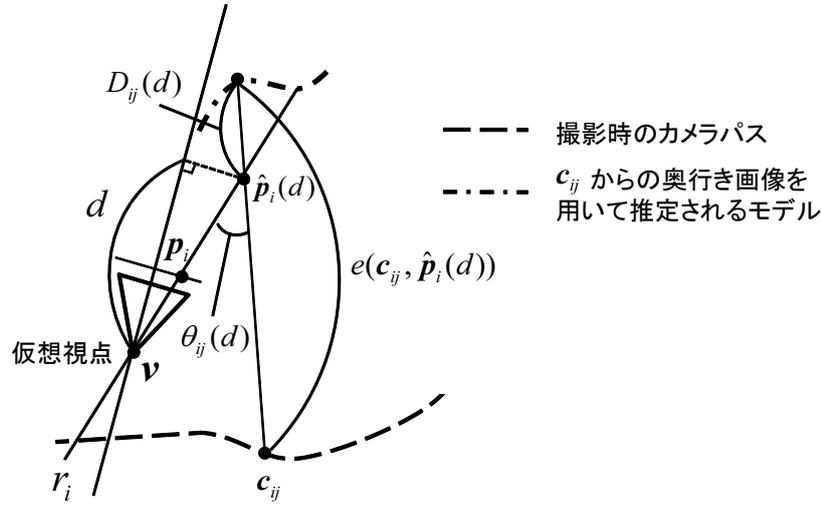


図 10 複数の奥行き情報の整合性に関するエネルギーの算出に用いられる変数

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，奥行きの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 (2) 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 箇所の撮影地点 ($c_{i1}, c_{i2}, \dots, c_{iQ}$) 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 ($D_{ij}, D_{i2}, \dots, D_{iQ}$) 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 に対して本項によって与えられるエネルギーが最小となる地点 $\hat{p}_i(d)$ 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．

3.3.2 メッシュモデルの形状変形における時空間的な連続性を考慮したエネルギー関数の拡張

本研究では、従来手法で定義された式 (2) による奥行き情報の整合性に関するエネルギー関数に、メッシュモデルの形状変形における時間的・空間的な連続性を考慮し下項を追加する。具体的には、メッシュモデルの時間的な連続性を考慮することで、メッシュモデルがフレーム間において極端に変形することを抑制し、空間的な連続性を考慮することで、不自然なテクスチャの変形を抑制する。具体的には、エネルギー関数を次のように定義する。

$$E_i(d) = E_g(d) + k_1 E_{st}(d) + k_2 E_{ss}(d) \quad (5)$$

ここで、 $E_{st}(d)$ はメッシュモデルの形状変形における時間的な連続性に関する制約項、 $E_{ss}(d)$ はメッシュモデルの形状変形における空間的な連続性に関する制約項であり、 k_1 および k_2 は各項の第一項に対する重みである。以下、 $E_{st}(d)$ 、 $E_{ss}(d)$ について詳述する。

[メッシュモデルの形状変形における時間的な連続性]

メッシュモデルの形状変形に対する時間的な連続性制約に関するエネルギー項を以下のように定義する。

$$E_{st}(d) = \alpha(d - \hat{d})^2 \quad (6)$$

$$\alpha = \begin{cases} 1 & ; d_{MIN} < \hat{d} < d_{MAX} \\ 0 & ; otherwise \end{cases} \quad (7)$$

ここで、図 11 に示すように、 \hat{d} は現フレームの仮想視点位置と \mathbf{p}_i を結ぶ直線と、前フレームで生成されたメッシュモデルとの交点に対する奥行き値を表す。ただし、視線方向の回転や視点の移動によって、直線と前フレームで生成されたメッシュモデルとの交点が存在しない場合や、 \hat{d} が奥行き値 d の探索範囲である d_{MIN} から d_{MAX} の間に存在しない場合には、係数 α を 0 に設定し、本項によって与えられるエネルギーを 0 とする。本項によって与えられるエネルギーは前フレーム

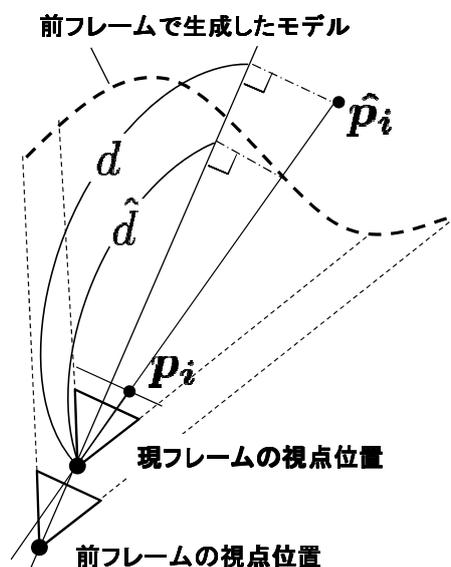


図 11 前フレームで生成したメッシュモデルと仮想視点の位置関係

で生成されたモデルと現フレームで生成されたモデルの位置が近いほど小さくなる。従って、本エネルギー項を用いることで、連続するフレーム間でメッシュモデルが大きく変形することを抑止する。

[メッシュモデルの形状変形における空間的な滑らかさ]

メッシュモデルの形状に対する空間的な滑らかさ制約に関するエネルギー項を以下のように定義する。

$$E_{ss} = \left(\frac{\partial^2 d}{\partial x^2} + \frac{\partial^2 d}{\partial y^2} \right)^2 \quad (8)$$

本項は、注目する頂点の奥行きに対するラプラシアン²の二乗で表現されるたわみエネルギー (2 乗曲率) である。つまり、本項によって与えられるエネルギーは、各頂点と隣接する頂点との間における奥行き値の勾配の変化量が小さいほど小さくなる。従って、本エネルギー項を用いることで、メッシュモデルの形状が局所的に大きく変形することを抑止する。

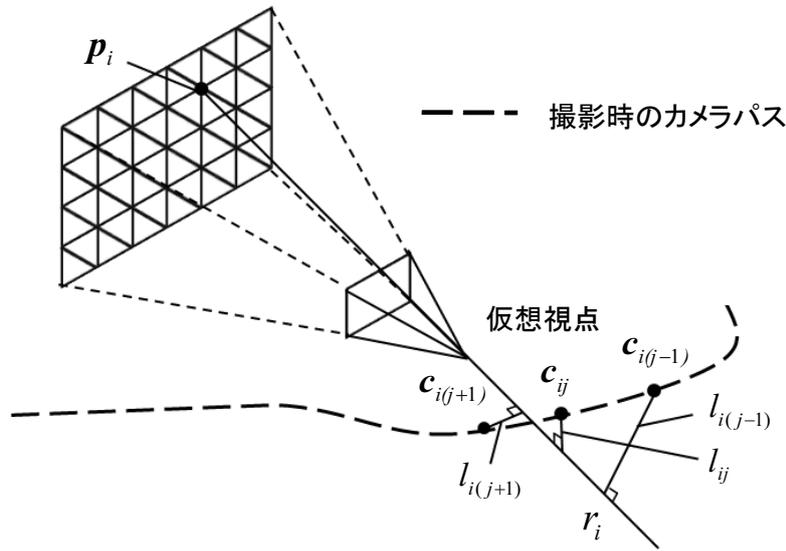


図 12 各頂点に対する光線と撮影カメラ位置からの距離

3.3.3 奥行き画像の選択

本項では、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を決定するために、それらの奥行き画像に対応する全方位画像の撮影地点 $(c_{i1}, c_{i2}, \dots, c_{iN})$ を選択する方法 (処理 (b-1)) 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 と仮想視点位置 v の位置関係から、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 12 に示すように、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 $(c_{i1}, c_{i2}, \dots, c_{iN})$ を選択する。これらの撮影地点 $(c_{i1}, c_{i2}, \dots, c_{iN})$ に対応する奥行き画像 $(DM_{i1}, DM_{i2}, \dots, DM_{iN})$ を、メッシュモデルの各頂点に対するエネルギーの算出に利用する。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ことから遮蔽の影響が小さく、後述するテクスチャマッピング処理においても利用する。

3.3.4 メッシュモデルの生成・変形

本研究では、図9に示したようなあらかじめ格子状に配置された頂点を、視点と各頂点を結ぶ光線上で移動させることでメッシュモデルをフレームごとに変形する。ここでは、式(5)で定義したエネルギー関数 $E_i(d)$ を用いて、メッシュモデルの生成・変形を行う方法(処理(b-2))について述べる。本処理は、前フレームまでに変形済のメッシュモデルを用いて各頂点の奥行き値を初期化する処理と、勾配法によって $E_i(d)$ が最小となる奥行きを探索する処理から成る。

[各頂点の奥行き値の初期値決定]

まず、対象とする頂点毎に、前フレームの仮想視点に対して生成されたメッシュモデルの情報を、初期値として用いられるかどうかを判定する。すなわち、前フレームの仮想視点位置に対して生成されたメッシュモデルと、視点と各頂点を結ぶ光線に交点が存在する場合、その交点の奥行き値 \hat{d}_i を初期値とする。初期フレームの場合や、仮想カメラのパンニングによって交点が存在しない場合には、光線上で仮想視点位置からの距離 d_i を d_{min} から d_{max} の範囲内で変化させ、全探索を行うことでエネルギー $E_g(d)$ が最小となる距離を決定し、初期値とする。メッシュ上の全ての点に対して初期値を算出した上で、以下に述べる探索処理を行う。

[勾配法による奥行き値探索]

得られた初期値を元に、勾配法を用いて E_i を最小化することで、メッシュモデル全体のエネルギーである E を最小化し、現在の視点位置に対する最適なメッシュモデルを生成する。具体的には、以下の式(9)を全頂点に対して逐次的に繰り返すことで、各頂点のエネルギー E_i を最小化し、その総和である E を最小化する。

$$d_i \leftarrow d_i - l_{d_i} \frac{\delta E_i}{\delta d_i} \quad (9)$$

ただし、矢印は更新を意味する。勾配法のスケール係数である l_{d_i} は、奥行き値の更新時点で、 E_i が最小となるように直線探索を行い決定する。式(9)に基づき、全頂点の更新を E が収束するまで繰り返し行うことで、現在の視点位置に対する最適なメッシュモデルを生成する。なお、本処理では図13に示すように、頂点同

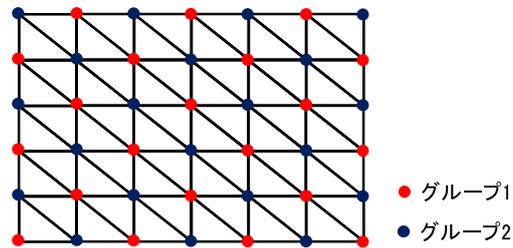


図 13 奥行き値の更新時における各頂点のグループ

士が4近傍隣接しない2つの頂点グループを用いて，グループ毎の更新を交互に行う．このような順番での更新を行うことで，全体のエネルギーが高速に収束することが知られている [31].

3.4 メッシュモデルへのテクスチャマッピング

本節では，生成されたメッシュモデルマッピングするテクスチャの決定 (処理 (b-3)) と，選択されたテクスチャのマッピングについて述べる (処理 (b-4)).

3.4.1 優先度を用いたテクスチャの決定

ここでは，3.3.3節で述べた手法により選択した全方位画像の撮影位置 \mathbf{c}_{ij} および各頂点においてエネルギーが最小となった奥行きに対応する重み $w(\hat{\mathbf{p}}_i(d))$ を用いて，テクスチャマッピングに用いる画像の撮影地点を決定する．三角形パッチ Ω とその三頂点の関係を，図 14 に示す．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 点ずつ選択された撮影地点に対応する合計 $3N$ 点の撮影地点を選択する．ここで，選択された $3N$ 点の撮影地点に対応する入力画像のフレーム番号をインデックスとして利用し， $\phi = (f_1, f_2, \dots, f_{3N})$ とおく．これらの撮影地点に対して，式 (10) 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 の最も高い撮影地点を決定する．

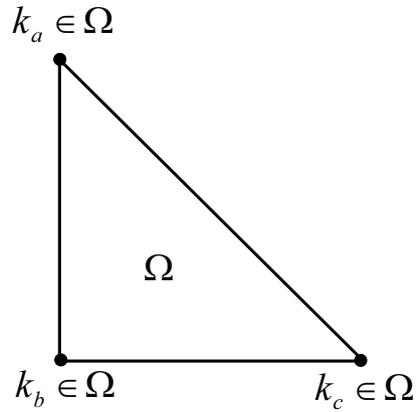


図 14 メッシュモデルの三角形パッチ Ω と三頂点

$$R_f = \sum_{n=1}^{3N} \begin{cases} w_{f_n}(\hat{\mathbf{p}}_i(d)) & ; (f_n = f) \\ 0 & ; otherwise \end{cases} \quad (10)$$

$3N$ 点の中に同一の撮影地点が含まれる場合、そのような撮影地点については優先度 R_f は高くなる。また、 f_n フレームにおける撮影地点 \mathbf{c}_{ij} の重み $w_{f_n}(\hat{\mathbf{p}}_i(d))$ が大きいほど、その撮影地点の優先度 R_f が高く設定される。以上より、各三角形パッチ毎に最も優先度の高い撮影地点を決定し、テクスチャマッピングを行う際に利用する撮影地点として仮決定する。

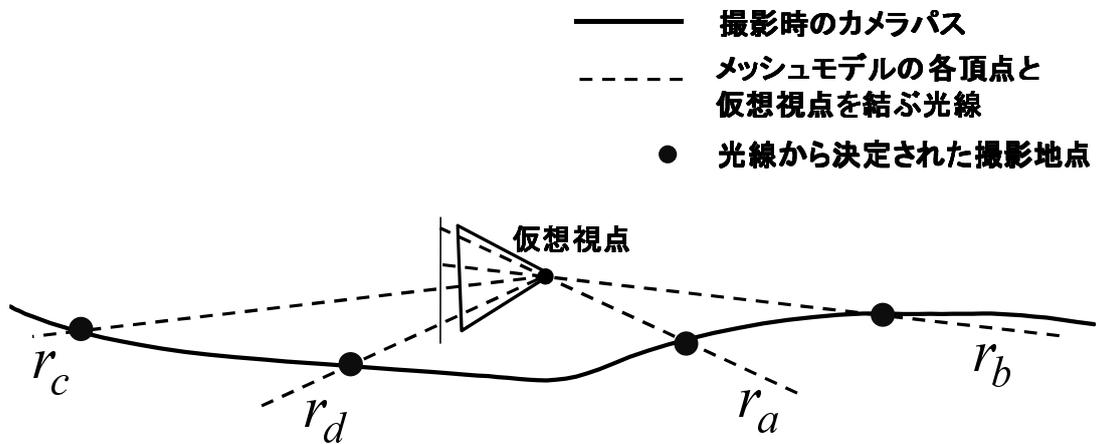


図 15 テクスチャ決定時の光線とカメラパスの位置関係

3.4.2 隣り合うパッチに対するテクスチャの連続化

前節で決定された各三角形パッチに対応する撮影地点は、隣り合うパッチ間で大きく離れている場合があり、テクスチャマッピングを行った際、そのような箇所では撮影地点が大きく異なることに起因して、図6に示したような、テクスチャの幾何学的・光学的な不連続が発生する可能性がある。具体的には、図15のようなカメラパスと仮想視点が設定されていた場合、光線 r_b と r_c を通る頂点の周辺において、画像上での位置が隣接しているにもかかわらず、利用されるテクスチャの撮影地点が大きく離れてしまう問題が発生する。そこで、隣り合うパッチ間なるべく近い地点の画像が利用されるように後処理を行うことで、そのような不連続を抑制する。具体的には、全三角形パッチを対象とし、一定サイズ $V \times V$ の平滑化フィルタを M 回適用することで平滑化する。次に、決定された各三角形パッチに対応する撮影地点から得られる入力画像を用い、次項で述べる各三角形パッチに対するテクスチャマッピング処理を行う。

3.4.3 GPUを用いたテクスチャマッピングの実装

本研究では、レンダリング処理高速化のために、GPUを用いたテクスチャマッピングの実装を行う。GPUを用いることで、各ピクセル毎の並列処理を行うことが可能であり、CPUを用いた場合より高速なレンダリングを実現できる。本レンダリング処理において、GPUが用いるデータには、視点位置が変更される毎に更新が必要なものと、更新が必要ないものに分類できる。更新が必要なデータは、メッシュモデルの全頂点の三次元座標、各三角形パッチに対応する撮影地点のインデックスである。これらのデータは、図16に示すように、レンダリング毎にCPUからGPUへ転送される。視点位置の変更によらず更新しないデータは、全方位画像、カメラ位置・姿勢、カメラパラメータである。これらのデータは、予めGPUに読み込んでおくことでデータ転送のオーバーヘッドをなくし、処理の高速化を図る。上記のデータを用いて、各三角形パッチに対してテクスチャマッピング処理を行い、自由視点画像をレンダリングする。

具体的な処理の流れを図16に示す。本レンダリングでは、自由視点画像の全てのピクセルに対して、以下の処理を並列的に行う。

- (1) 出力画像上で対象とするピクセルについて、自由視点からの光線を設定する
- (2) その光線とメッシュモデルとの交点を求め、交点を対応する撮影地点から得られる複数の画像に再投影することで、対応する撮影カメラおよびテクスチャ座標を計算する
- (3) その光線の色を、対応するテクスチャ座標の色とすることで、出力画像におけるその点の色を決定する

以上の手順を、出力画像上の全てのピクセルに対して行うことで、全てのピクセルの色を求め、それらを統合することで自由視点画像をレンダリングする。

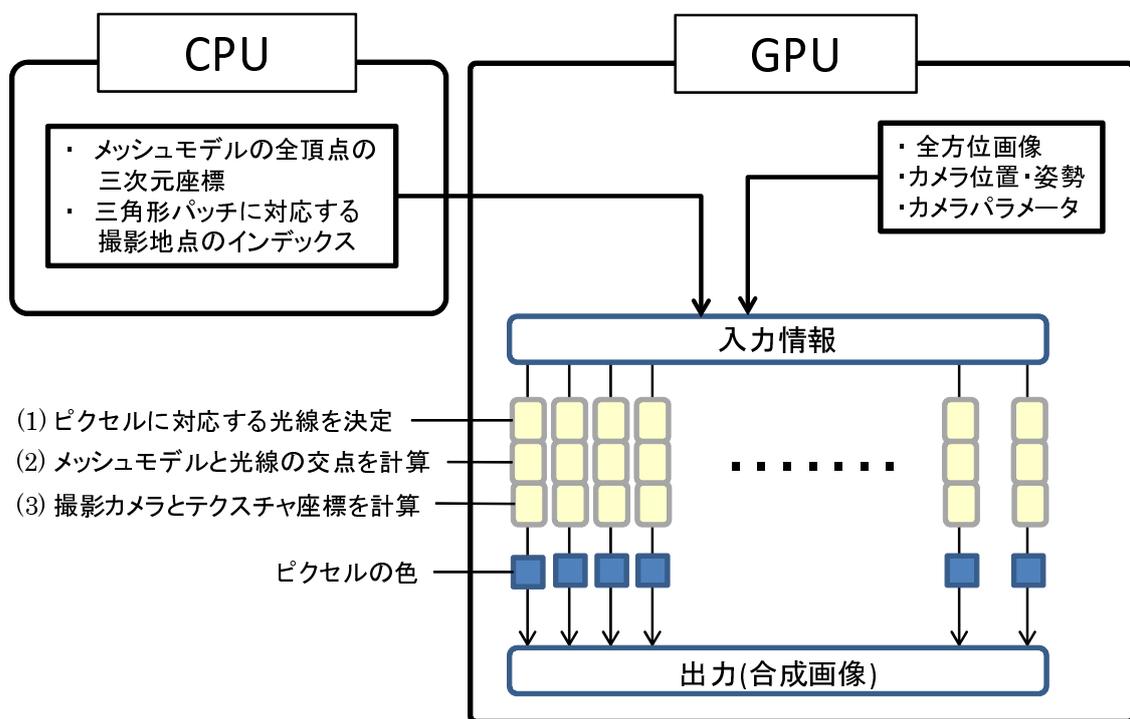


図 16 GPU を用いたテクスチャマッピングの処理の流れ

4. 実験

4.1 全方位動画像からの自由視点画像生成実験

提案手法の有効性を示すために、屋外で撮影された全方位動画像 (500 フレーム, 3000 枚) を用い、従来手法による生成結果との比較実験を行った。

4.1.1 入力情報の取得

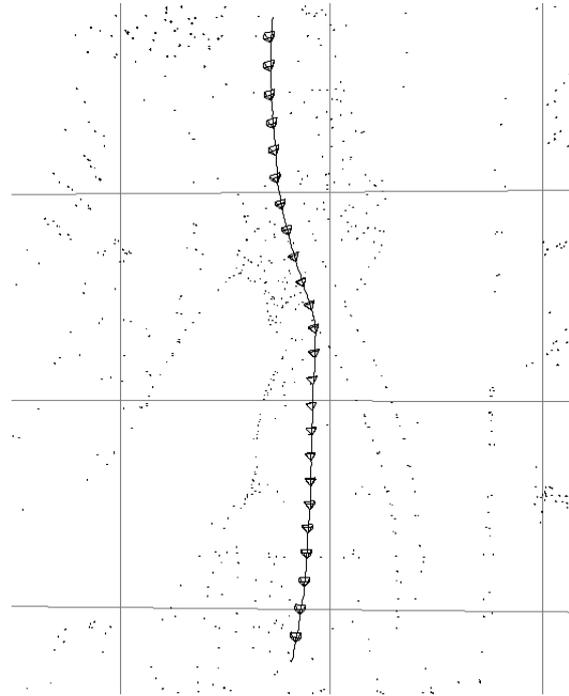
図 17 に示す複数のカメラを外向きに配置した Point Grey Research 社製の全方位型マルチカメラシステム Ladybug を用いて対象環境の全方位動画像の撮影を行った。本カメラシステムでは、横 768 画素、縦 1024 画素の画像を水平方向 5 台および真上方向 1 台の計 6 台の CCD カメラで同期して取得できる。また、全方向の約 75 % 以上の領域を 30fps の動画像として取得することが可能である。図 18 に本実験で入力として用いる全方位動画像の 1 フレームを示す。ここではまず、この全方位動画像に対して 3.2.1 項で述べた SFM 法を適用することで、図 19 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 20 に示す特徴点の奥行き値を算出した。図 21 は、図 20 に対してドローネーの三角形分割法を用いることで、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。



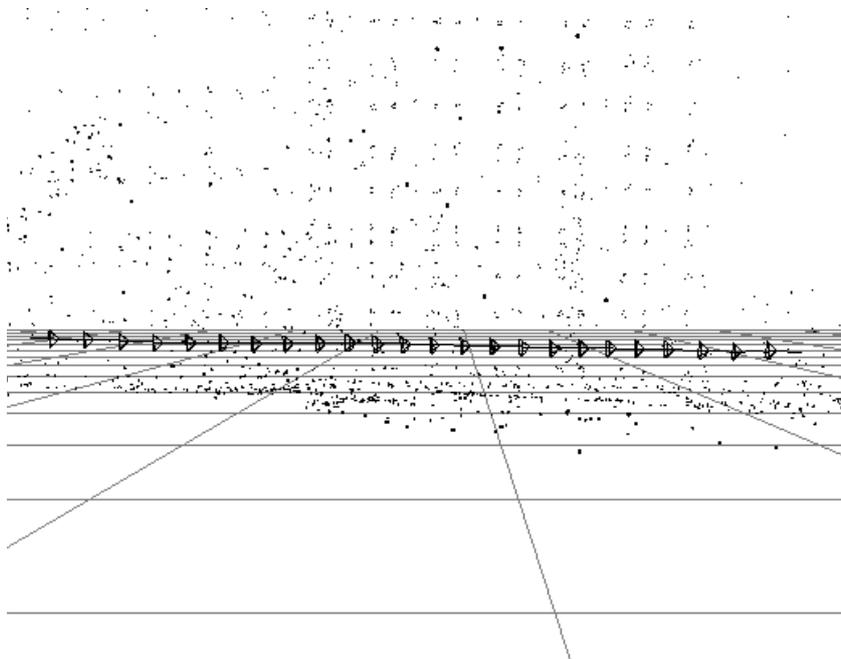
図 17 全方位型マルチカメラシステム Ladybug



図 18 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 19 SFM 法によって推定されたカメラ位置・姿勢

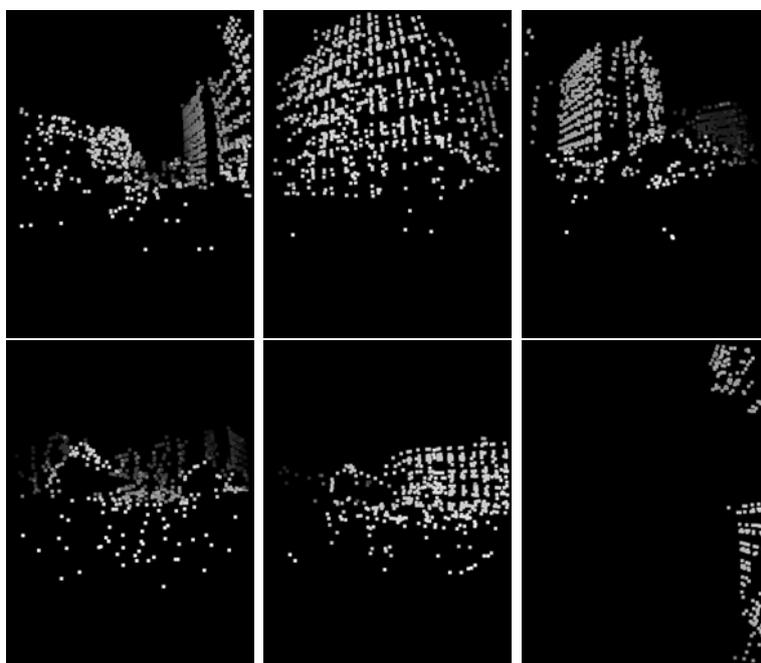


図 20 マルチベースラインステレオ法により推定された全方位奥行き画像の一例

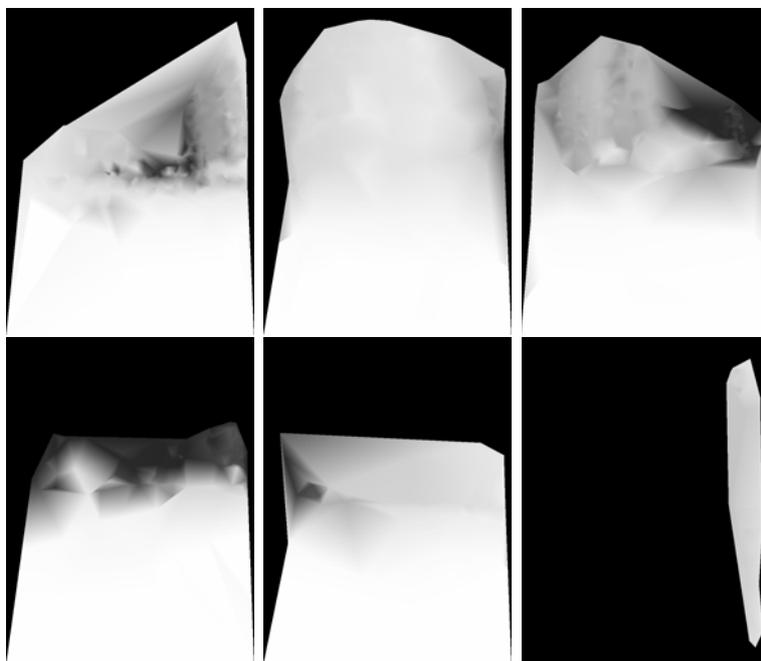


図 21 ドロネーの三角形分割法により補間された全方位奥行き画像の一例

4.1.2 提案手法による全方位動画像からの自由視点画像の生成と従来手法との比較

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, 表 1 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 22 に示すように, 全方位カメラのカメラパス周辺を移動する楕円状のルートに対して以下の 3 つの手法を用いて, それぞれ 300 フレームの自由視点画像を生成した.

- 従来手法: 文献 [30] で提案された手法
- 提案手法 A: 従来手法に加えて, メッシュモデルの時間的な連続性及び空間的な滑らかさの考慮のみを行った (テクスチャの連続性を考慮しない) 場合
- 提案手法 B: 提案手法 A に加えてテクスチャの連続性を考慮した場合

ただし, 従来手法におけるテクスチャマッピング処理は, 文献 [30] に示されている越澤らの手法における実装である三角形パッチ毎のレンダリングではなく, 提案手法と同様のピクセル毎に独立したテクスチャの選択処理を, GPU ではなく CPU によって行った. 図 23 に地点 (a) を起点とする 5 フレームに対応する自由視点画像の生成結果を示し, 図 24 にその一部の領域を拡大した画像を示す. 同様に, 図 25 および図 26 に地点 (b), 図 27~図 29 に地点 (c) の生成結果および拡大画像を示す. また, 図 30~図 32 に地点 (a), (b), (c) 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それぞれ示す. また, 表 2 に本研究で使用した計算機のスペックを, 表 3 に自由視点画像の生成にかかった平均処理時間を示す.

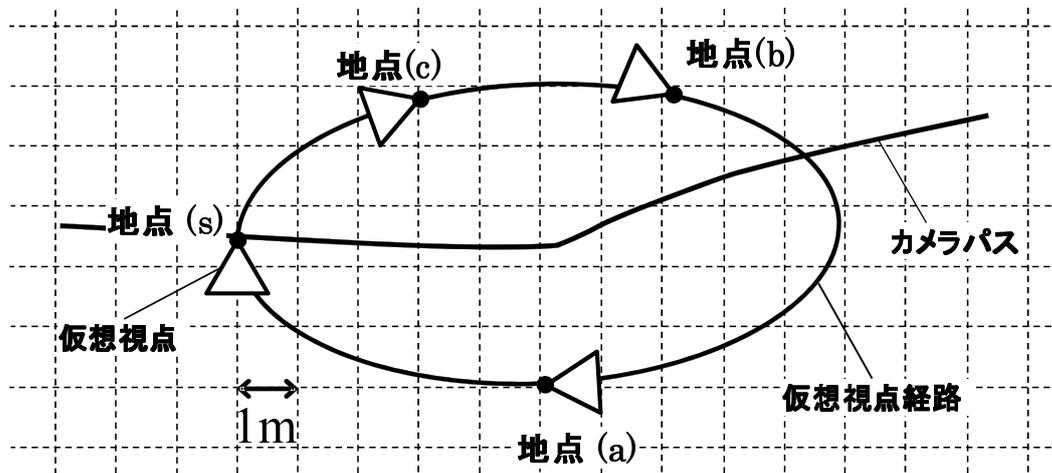


図 22 自由視点画像の生成視点位置・視線方向

表 1 本実験のパラメータ

自由視点画像の解像度	横 780[pixel] 縦 780[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	3 枚
第 2 項の係数 k_1	2.0
第 3 項の係数 k_2	2.0
平滑化フィルタのサイズ $V \times V$	3×3
平滑化フィルタの適用回数 M	3 回

表 2 使用した計算機のスペック

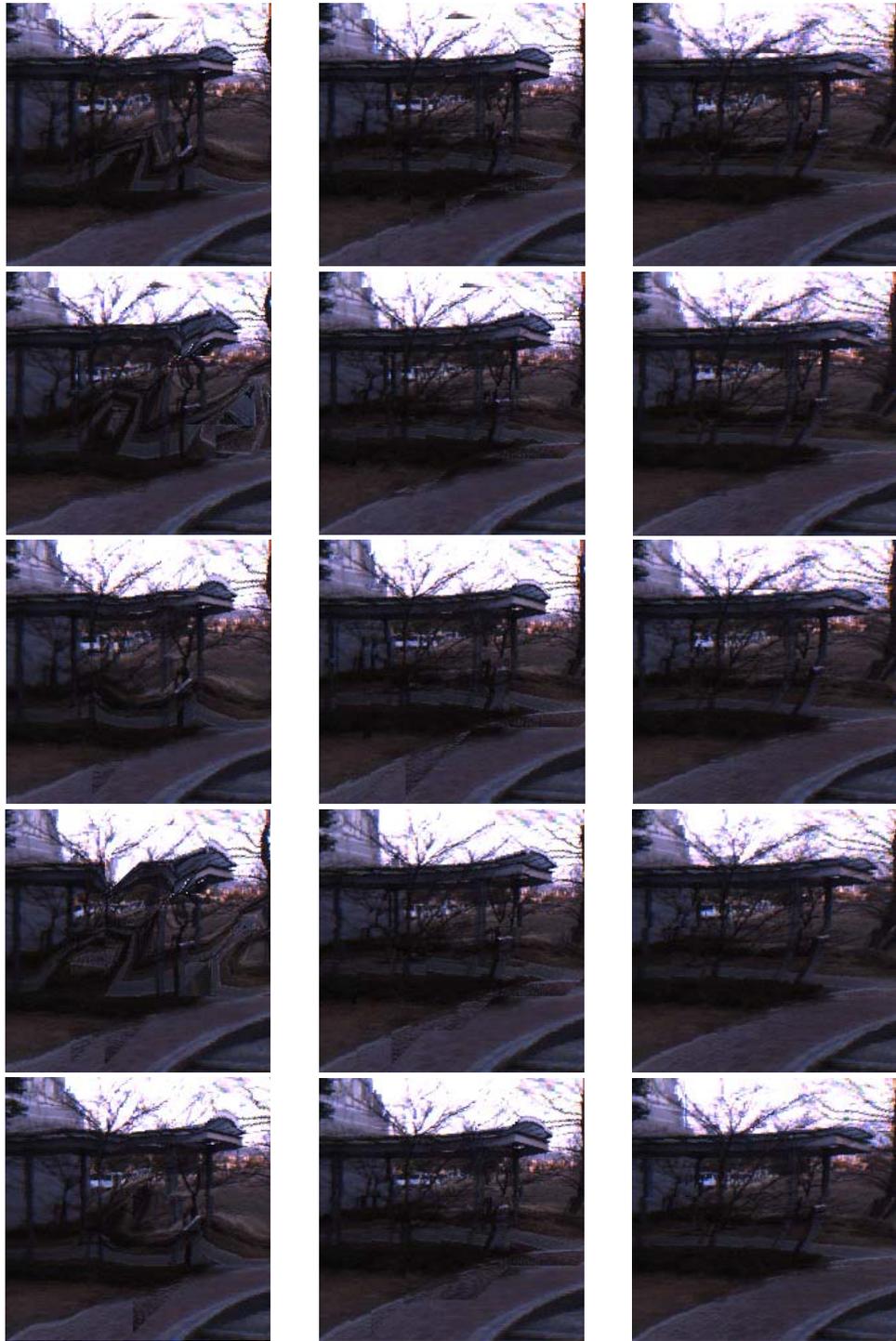
CPU	Intel(R) Core(TM)2 Duo E8600 3.33 GHz
Memory	16 GB
GPU	NVIDIA GeForce GTX285 VRAM: 2 GB

表 3 自由視点画像の生成にかかる平均処理時間

	従来手法	提案手法
奥行き画像の選択	0.067[sec]	0.067[sec]
三次元メッシュモデルの初期化	7.380[sec]	7.832[sec]
三次元メッシュモデルの変形	2.458[sec]	2.363[sec]
テクスチャマッピング	56.682[sec]	0.027[sec]



従来手法 提案手法 A 提案手法 B
 図 23 (a) 地点 (75 フレーム目) における自由視点画像
 (上から順に 75,76,77,78,79 フレーム目の生成画像)

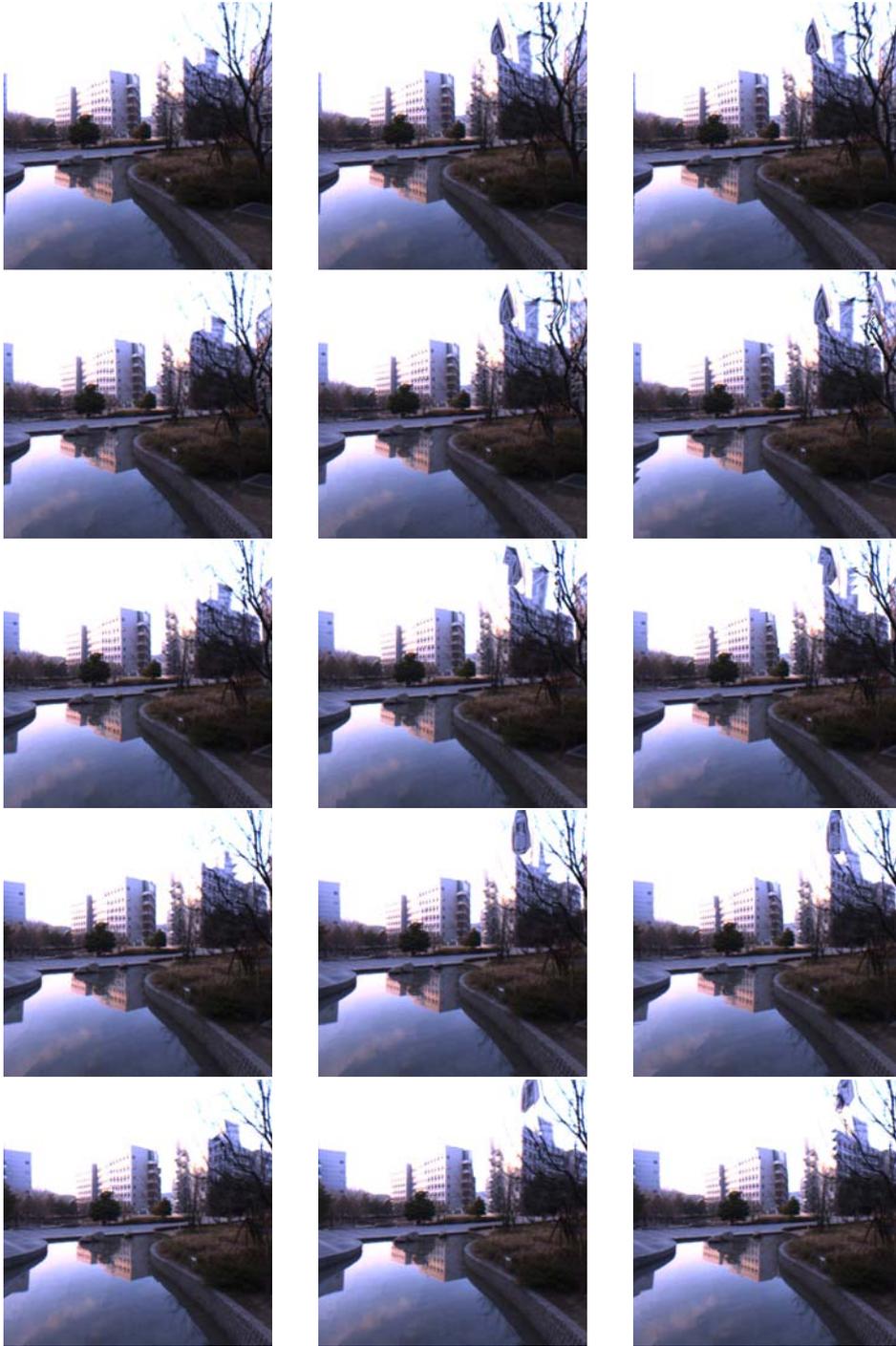


従来手法

提案手法 A

提案手法 B

図 24 (a) 地点 (75 フレーム目) における自由視点画像の拡大画像
(上から順に 75,76,77,78,79 フレーム目の生成画像)



従来手法

提案手法 A

提案手法 B

図 25 (b) 地点 (190 フレーム目) における自由視点画像
 (上から順に 190,191,192,193,194 フレーム目の生成画像)

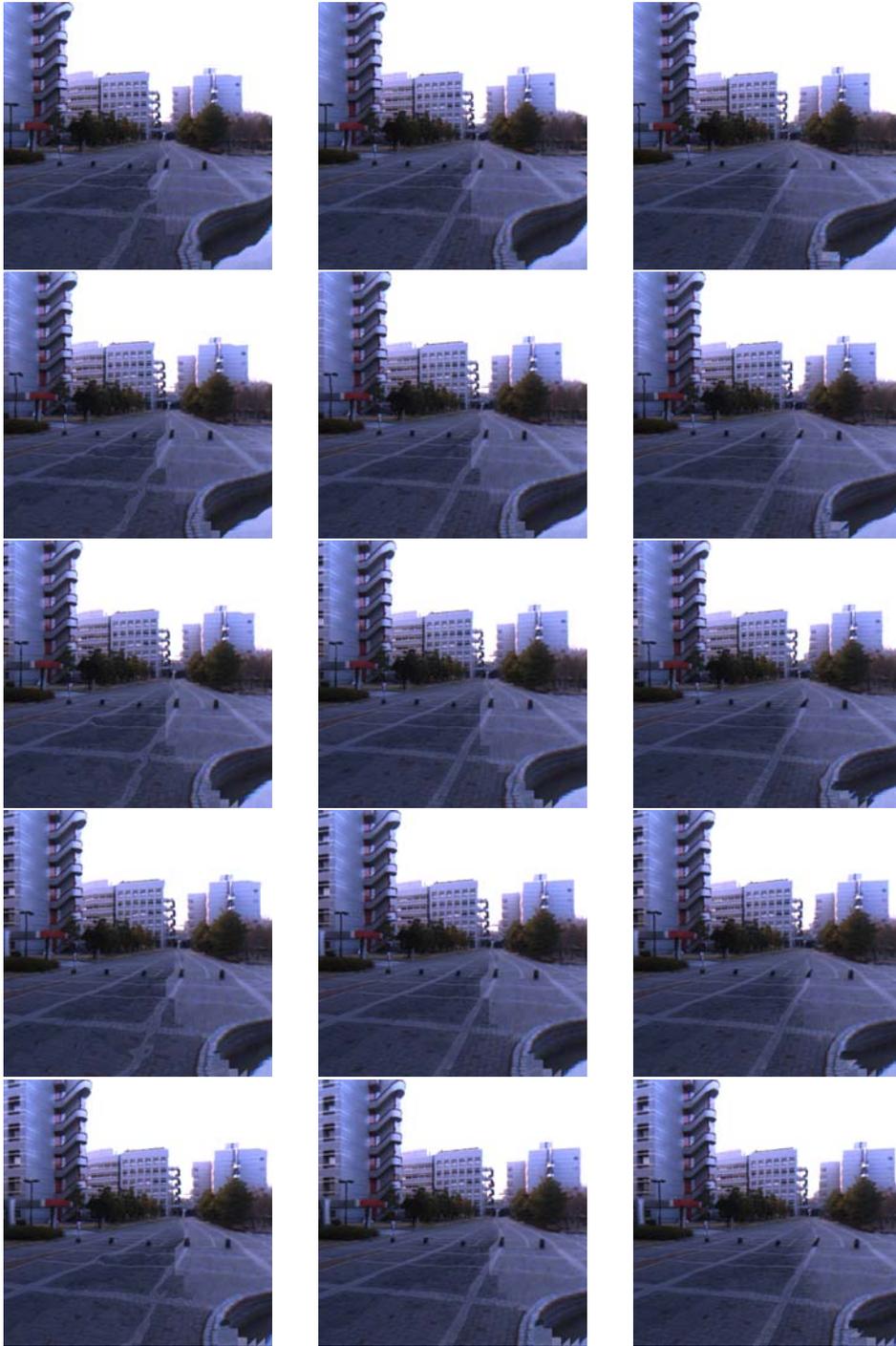


従来手法

提案手法 A

提案手法 B

図 26 (b) 地点 (190 フレーム目) における自由視点画像の拡大画像
(上から順に 190,191,192,193,194 フレーム目の生成画像)

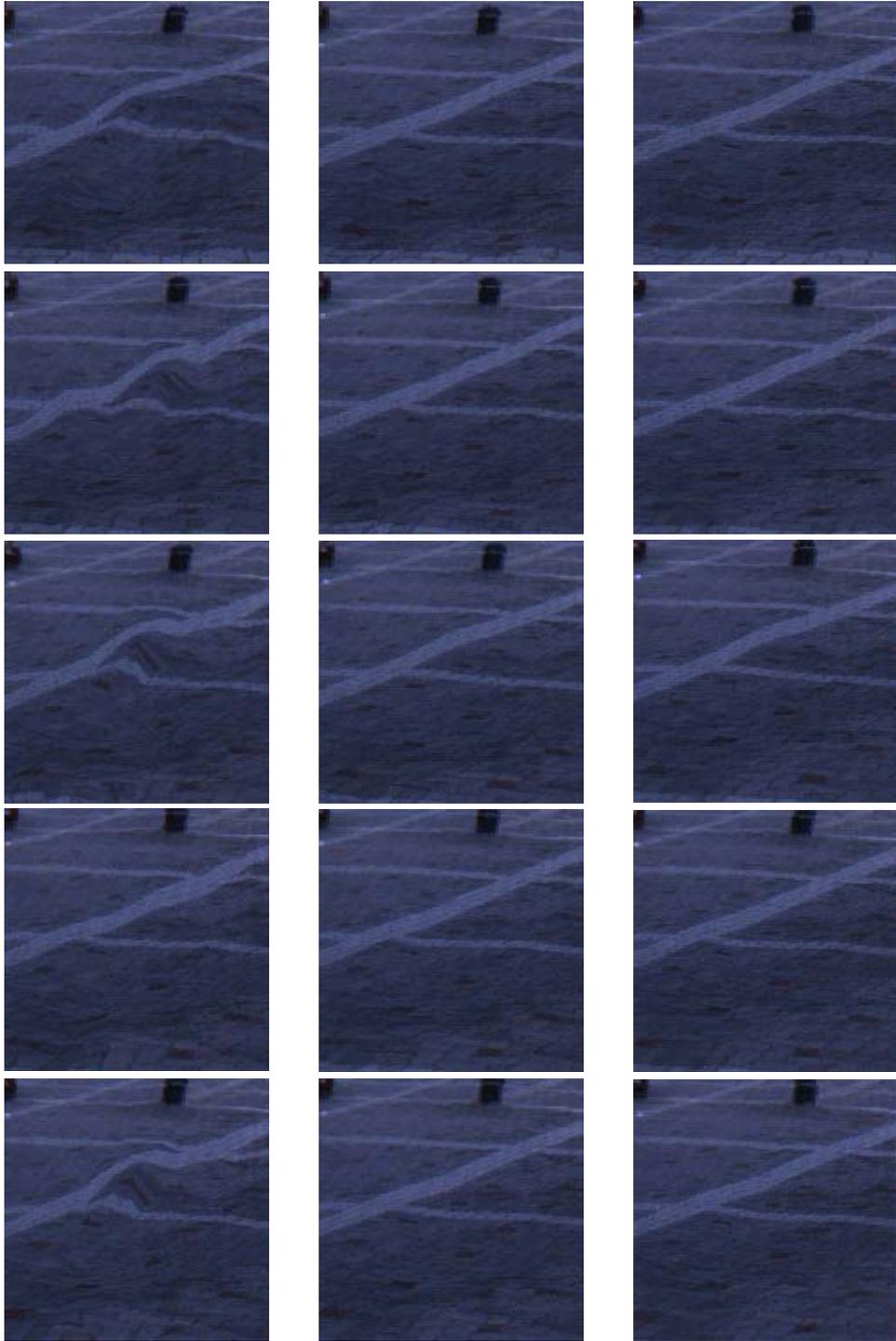


従来手法

提案手法 A

提案手法 B

図 27 (c) 地点 (234 フレーム目) における自由視点画像
 (上から順に 234,235,236,237,238 フレーム目の生成画像)

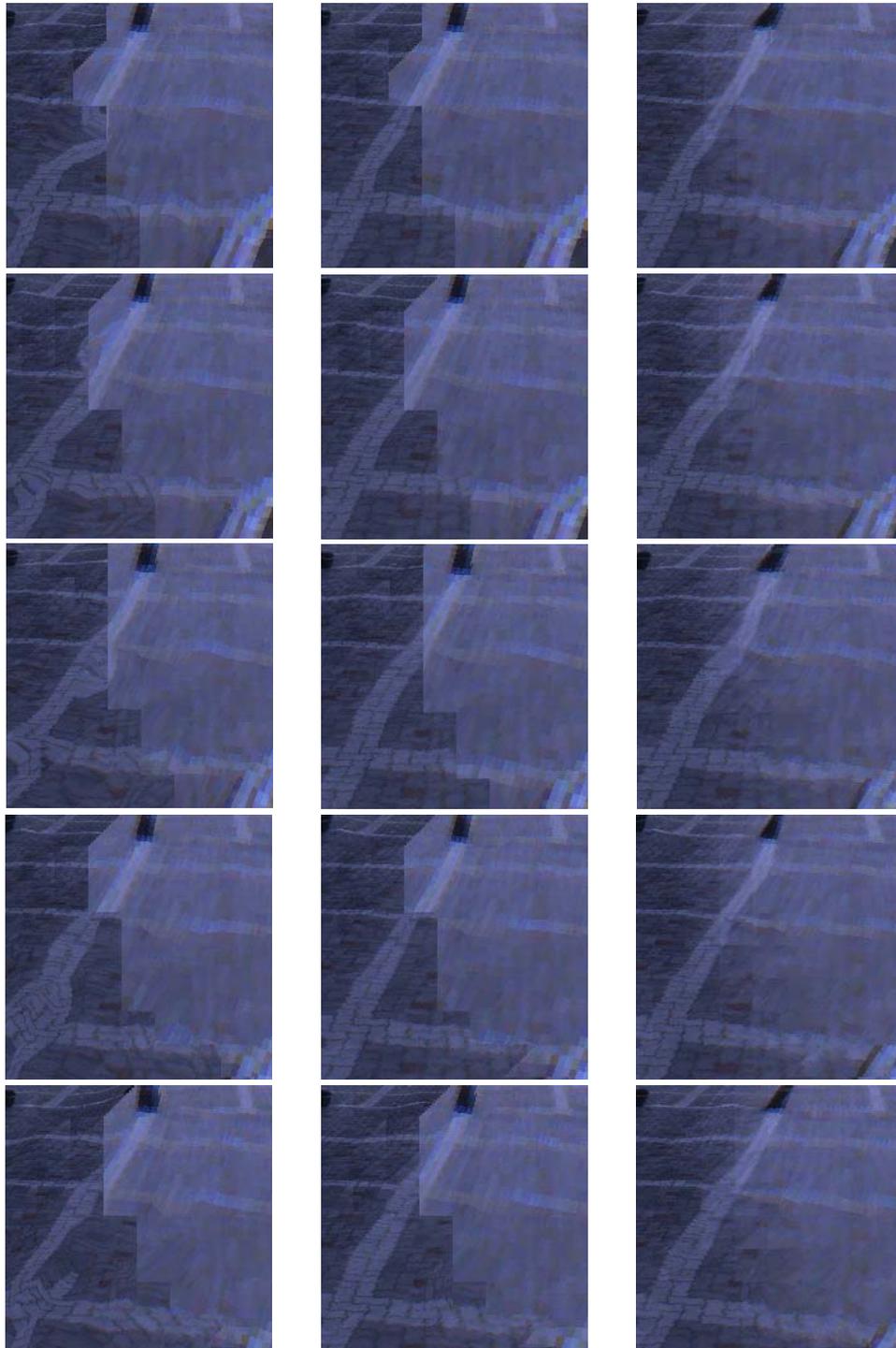


従来手法

提案手法 A

提案手法 B

図 28 (c) 地点 (234 フレーム目) における自由視点画像の拡大画像 1
(上から順に 234,235,236,237,238 フレーム目の生成画像)



従来手法

提案手法 A

提案手法 B

図 29 (c) 地点 (234 フレーム目) における自由視点画像の拡大画像 2
 (上から順に 234,235,236,237,238 フレーム目の生成画像)

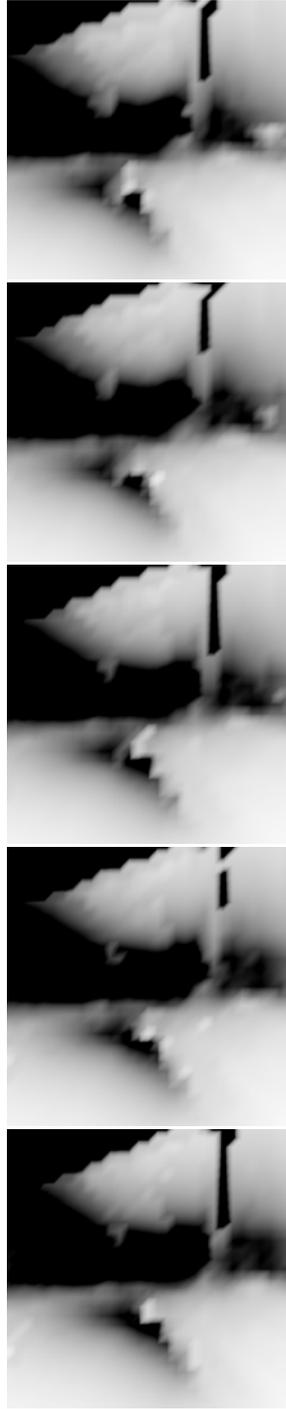


従来手法



提案手法 A および B

図 30 (a) 地点 (75 フレーム目) 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 75,76,77,78,79 フレーム目の生成画像)

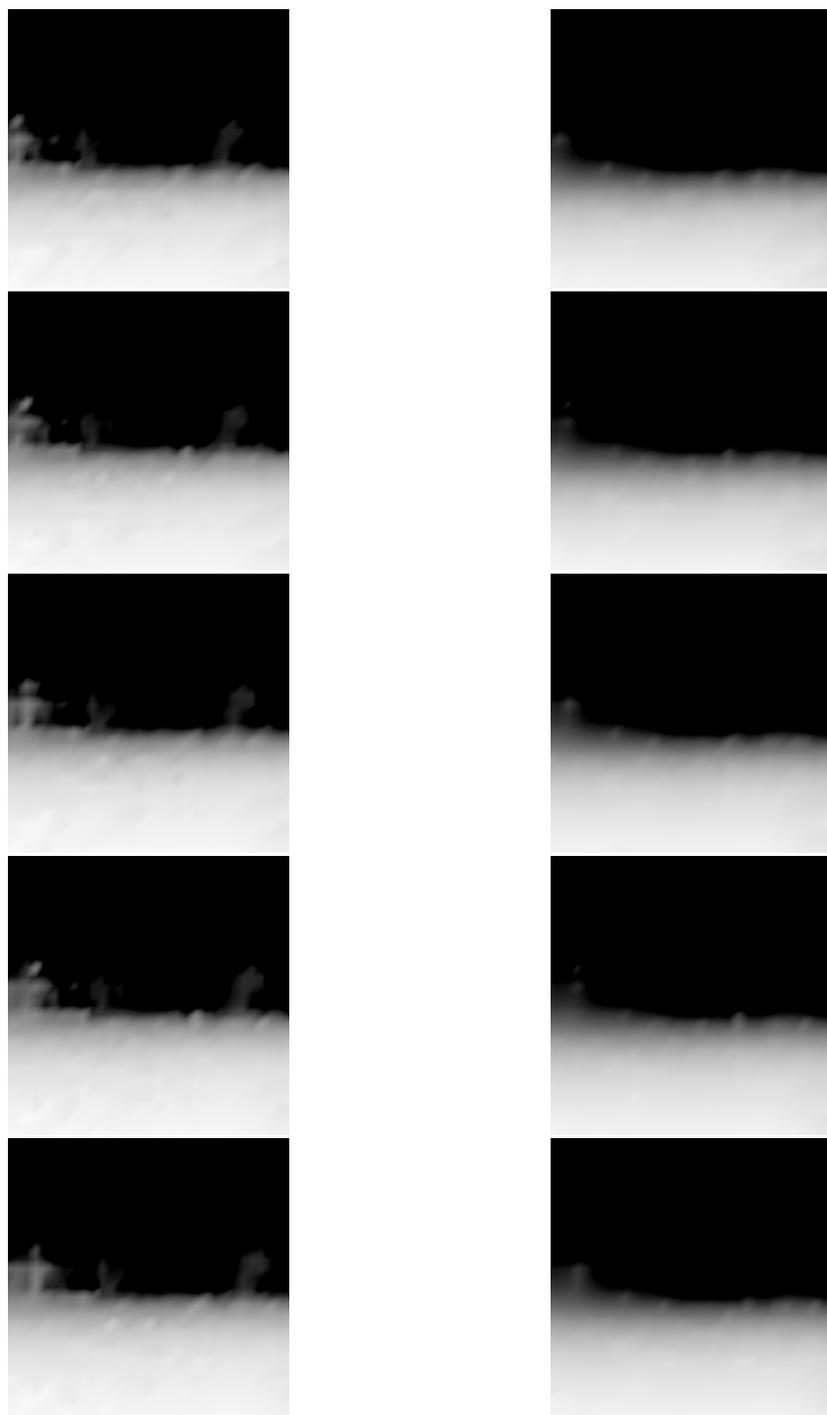


従来手法



提案手法 A および B

図 31 (b) 地点 (190 フレーム目) 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 190,191,192,193,194 フレーム目の生成画像)



従来手法

提案手法 A および B

図 32 (c) 地点 (234 フレーム目) 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 234,235,236,237,238 フレーム目の生成画像)

4.2 提案手法の従来手法との比較による考察

地点 (a) に関する生成画像である図 23 および図 24 より、従来手法では形状が時間的に不連続に変化していた箇所において、提案手法では安定した形状が保たれており、テクスチャの不連続な歪みが抑制されていることが確認出来る。地点 (c) に関する生成画像である図 27 および図 28, 図 29 でも同様に、地面に生じていた不連続な歪みが抑制されていることが確認出来る。特に、図 29 における比較では、3.4.2 項で述べた、隣り合うテクスチャの連続化処理を行うことで、生成画像上に生じている光学的な不連続が大幅に改善されている。ただし、図 25 および図 26 に示した地点 (b) の画像においては、従来手法では比較的写実的な画像が生成されているが、提案手法では合成画像上に不自然なテクスチャが発生しており違和感のある画像となっている。この原因として、従来手法では奥行き値の探索に一定区間の全探索を用いていたため、局所解からの脱出が比較的容易であったのに対し、提案手法ではエネルギー関数の第二項および第三項の性質上、頂点の奥行き値が局所解に陥ってしまった場合において、以降のフレームではその頂点および近傍の頂点が局所解に引きつけられてしまうことが考えられる。そのため、エネルギー探索時に局所解に陥ったかどうかを判定する何らかのアプローチを組み込み、局所解に陥っていると判定された場合には全探索を行い、局所解から抜け出す仕組みが必要であると考えられる。

一方、表 3 より、提案手法では GPU を用いることでテクスチャマッピングの処理時間を大幅に短縮していることがわかる。ただし、現時点ではメッシュモデルの変形に約 2.4 秒かかるため、1 フレーム辺りのレンダリングに約 2.5 秒を要しており、現在の実装において自由視点画像を実時間で生成することは難しい。今後、メッシュモデルの変形に必要な探索処理をマルチスレッド処理に切り替えることや、初期値決定時の探索手法をより高速に大局解へ収束するような手法へ切り替えることで、変形処理の高速化を行う必要がある。

5. まとめ

本論文では、高品位な自由視点画像を生成するため、メッシュモデルの時間的・空間的な連続性とテクスチャの幾何学的・光学的な連続性を考慮することで、合成画像上に生じる違和感を低減させる手法を提案した。提案手法では、まず、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 [13] 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 [8] 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、幾何形状の変形に関するエネルギー関数に基づいて適切な形状に変形させ、テクスチャマッピングを行うことで自由視点画像を生成する。提案手法は、幾何形状の変形時に、従来用いられてきた幾何学的な整合性に加えて、メッシュモデルの時間的な連続性を考慮することで、視点位置の変更に伴ってメッシュモデルが極端に変形することを抑制し、空間的な連続性を考慮することで、局所的な変形を抑制する。また、テクスチャマッピングにおいて、視線方向およびパッチ間で利用される画像の撮影地点の位置関係を考慮して決定したテクスチャをマッピングすることで、合成画像上におけるテクスチャの幾何学的・光学的な不連続を抑制する。

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い、違和感の少ない自由視点画像が生成できることを示した。また、GPU を用いたテクスチャマッピングの実装を行うことにより、従来手法と比較して大幅に高速なレンダリングを実現できることを示した。ただし、頂点の奥行き値が局所解に陥ってしまう箇所も存在し、そのような箇所においては合成画像上に違和感が生じているため課題が残る。

今後、合成画像上に生じる歪みを解決するためには、奥行き値が局所解に陥った際の判定処理を組み込むことが必要である。また、探索手法の改良や、マルチスレッド処理による処理速度の高速化を行うことで、実時間処理を達成することが必要である。

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 教授に心より感謝致します。そして，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いた像情報処理学講座 千原 國宏 教授に厚く御礼申し上げます。さらに，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。また，公私にわたり大変お世話になり，研究に関する的確な御助言を頂いた視覚情報メディア講座神原誠之助教に熱く御礼申し上げます。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，御鞭撻を頂きましたことを，重ねて御礼申し上げます。そして，研究の方向性について無数の御助言，御指導を賜った Microsoft Research Asia 松下 康之氏に心より深く感謝致します。また，研究室での生活を支えて頂いた視覚情報メディア講座事務補佐員 高橋 美央 女史に心より感謝致します。さらに，研究活動だけでなく日々の生活においても大変お世話になった視覚情報メディア講座の皆さまに深く感謝致します。最後に，両親をはじめ，私の二年間の大学院生活に関わった全ての方々に感謝の意を表します。

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] V. J. Milenkovic and T. Kanade: “Trinocular Vision Using Photometric and Edge Orientation Constraints”, Proc. DARPA Image Understanding Workshop, pp. 163-175, 1985.
- [7] N. Ayache and F. Lustman: “Fast and Reliable Passive Trinocular Stereo Vision”, Proc. Int. Conf. on Computer Vision, Vol. 1, pp. 422-426, 1987.
- [8] 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.
- [9] R. Y. Tsai: “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision”, Proc. Computer Vision and Pattern Recognition, pp. 364-374, 1986.

- [10] 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.
- [11] T. Mukai and N. Ohnishi: “The Recovery of Object hap and Camera Motion Using a Sensing system with a Video Camera and a Gyro Sensor”, Proc. I Int. Conf. on Computer Vision, Vol. 1, pp. 411-447, 1999.
- [12] 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.
- [13] 佐藤智和, 池田聖, 横矢直和: “複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定”, 電子情報通信学会論文誌, Vol. J88-D-II, No. 2, pp. 347-357, 2005.
- [14] 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.
- [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] P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J-M. Frahm, R. Yang, D. Nister and M. Pollefeys “Real-Time Visibility-Based Fusion of Depth Maps ”, Proc. Int. Conf. on Computer Vision, 2007.
- [17] Y. Furukawa, B. Curless, S. M. Seitz and R. Szeliski: “Reconstructing Building Interiors from Images”, Proc. Int. Conf. on Computer Vision, pp. 80-87, 2009.
- [18] S. E. Chen and L. Williams: “View Interpolation for Image Synthesis”, Proc. SIGGRAPH’93, pp. 279-288, 1993.

- [19] S. M. Seitz and C. R. Dyer: “View Morphing”, Proc. SIGGRAPH’96, pp. 21-30, 1996.
- [20] E. H. Adelson and J. Bergen: “The Plenoptic Function and the Elements of Early Vision”, MIT Press Cambridge, MA, pp. 3-20, 1991.
- [21] 苗村健, 原島博: “Ray-Based Creation of Photo-Realistic Virtual World”, Virtual Reality and MultiMedia, pp. 59-68, 1997.
- [22] Marc Levoy and P. Hanrahan: “Light Field Rendering”, Proc. SIGGRAPH’96, pp. 31-42, 1996.
- [23] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen: “The Lumigraph”, Proc. SIGGRAPH’96, pp. 43-54, 1996.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] M. Irani, T. Hassner and P. Anandan “What Does the Scene Look Like from a Scene Point?”, Proc. European Conf. on Computer Vision, vol. 2, pp. 883-893, 2002.

- [29] 高橋桂太, 苗村健: “視点依存奥行きマップ実時間推定に基づく多眼画像からの自由視点画像合成”, 映像情報メディア学会誌, Vol. 60, No. 10, pp. 1611-1622, 2006.
- [30] 越澤広幸, 佐藤智和, 横矢直和: “視点位置に応じた3次元メッシュモデルの逐次生成による全方位動画からの自由視点画像生成”, 画像の認識・理解シンポジウム 講演論文集, pp. 1586-1593, 2009.
- [31] Besag, J. : “Spatial Interaction and the Statistical Analysis of Lattice Systems”, Journal of the Royal Statistical Society, Series B, Vol. 36, No. 2, 192-236, 1974.