

NAIST-IS-MT1351050

## 修士論文

線分の直線性および画像間の対応点を考慮した画像変形  
に基づく自由視点画像における歪みの抑制

篠本 渉

2015年3月11日

奈良先端科学技術大学院大学  
情報科学研究科 情報科学専攻

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。

篠本 渉

審査委員：

横矢 直和 教授	(主指導教員)
向川 康博 教授	(副指導教員)
佐藤 智和 准教授	(副指導教員)
河合 紀彦 助教	(副指導教員)
中島 悠太 助教	(副指導教員)

# 線分の直線性および画像間の対応点を考慮した画像変形 に基づく自由視点画像における歪みの抑制\*

篠本 渉

## 内容梗概

自由視点画像生成は、テレプレゼンスやデジタルアーカイブの閲覧などのアプリケーションにおいて、カメラで撮影された視点以外の任意視点からの画像をユーザに提示し、高い臨場感を与えることができるため近年注目を集めている。これを実現する一手法として、撮影した画像群から対象となるシーンをあらかじめ三次元復元した上で、仮想視点に依存した適切なテクスチャを選択的に用いる視点依存テクスチャマッピングに基づく手法が有名であるが、三次元復元の精度が低い場合に生成画像に歪みが生じるという問題がある。一方、画像中の線分の直線性が保たれるように入力画像を変形することで歪みの少ない自由視点画像を生成する手法が近年提案されているが、視差により画像上での前景と背景の位置関係が変わるような状況には対応できず、設定する視点位置は入力画像のものと十分近くなければならない。本研究では、これらの問題を解決し、入力画像の視点位置から大きく離れた視点位置においても歪みのない自由視点画像を生成することを目的として、三次元形状を用いる視点依存テクスチャマッピングと画像の変形手法の組み合わせに基づく自由視点画像生成手法を提案する。提案手法では、まず三次元復元を行った上で、奥行き値を手がかりにした視点依存テクスチャマッピングにより各入力画像に対応した自由視点画像群を生成する。次に、入力画像中の線分を用いた直線性の補償および複数画像間の切り替わり位置におけるテクスチャの連続性を考慮した画像変形により自由視点画像の歪みを抑制する。最後

---

\*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 修士論文, NAIST-IS-MT1351050, 2015年3月11日.

に，照明条件の異なる画像からマッピングされたテクスチャ間の色の不連続をポアソン方程式を用いた色調補正により低減する．実験では，屋外環境を対象とした自由視点画像を生成し，従来手法と定性的に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, 三次元形状復元, ワーピング, 最適化

# Distortion Compensation for Novel View Synthesis Based on Image Deformation Considering Linearity of Line Segments and Correspondences between Multiple Images\*

Wataru Sasamoto

## Abstract

Novel view synthesis has recently attracted much attention because it can present users an image from an arbitrary viewpoint different from the ones of real cameras and offer them a sense of immersion. One of the conventional approaches reconstructs 3D shapes of a scene in advance and determines the color of a target image using view-dependent texture mapping. However, it often produces undesirable texture distortions due to inaccurately reconstructed shapes. Meanwhile, one of the state-of-the-art methods creates a novel view by image warping while preserving the linearity of line segments in the scene. This can generate a preferable image with the aid of content-preserving warps, but there are strict constraints on the virtual camera pose for a novel view because this method cannot deal with occlusions or object displacements. In order to overcome this problem, this thesis proposes a method for synthesizing a novel view based on the combination of the view-dependent texture mapping and the image warping methods. We first reconstruct 3D shapes of a scene and generate a novel view at a target viewpoint from each input image using depth information as a

---

\*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1351050, March 11, 2015.

cue. The generated images are synthesized into a novel view using the combination of view-dependent texture mapping and image warping while preserving the linearity of line segments and texture continuity around borders between textures mapped from different input images. This enables view synthesis at a viewpoint far from the ones of real cameras, which is difficult for methods using no 3D shapes, and reduces distortions, which often occur in methods using 3D shapes. Finally, we apply color correction using Poisson's equation for the areas where images that have different lighting conditions are mapped. Our results demonstrate novel view synthesis for outdoor environments, showing visual validity compared with conventional methods.

**Keywords:**

novel view synthesis, 3D reconstruction, image warping, optimization

# 目次

1. 緒言	1
2. 自由視点画像生成の関連研究	4
2.1 三次元形状を用いる手法	4
2.1.1 三次元形状復元手法	4
2.1.2 三次元形状を用いた自由視点画像生成手法	7
2.2 三次元形状を用いない手法	10
2.3 本研究の位置付けと本論文の構成	13
3. 線分および対応点を考慮した自由視点画像の歪みの抑制	14
3.1 提案手法の概要	14
3.2 仮想視点に対応する投影データの生成	16
3.2.1 画像群の投影	16
3.2.2 カメラマップの生成	19
3.2.3 線分の投影	20
3.3 境界領域における対応点探索	21
3.4 線分および対応点を考慮した画像変形	22
3.4.1 エネルギー関数の定義	22
3.4.2 エネルギー関数の最小化	25
3.5 ポアソン方程式を用いた色調補正	27
4. 実験	30
4.1 実験条件	30
4.2 主観評価実験	36
4.3 考察	57
5. 結言	59
謝辞	60



## 目 次

1	視点依存テクスチャマッピングで生じる歪みの例 . . . . .	2
2	Furukawa ら [19] の手法によるレンダリング結果 . . . . .	7
3	視点依存テクスチャマッピング [1] によるレンダリング . . . . .	8
4	伊吹ら [2] の手法における自由視点画像上の歪み . . . . .	10
5	Huang ら [4] による結果の比較 . . . . .	12
6	提案手法の概要 . . . . .	15
7	カメラマップの例 . . . . .	16
8	投影方法の種類 . . . . .	17
9	境界領域における対応点探索 . . . . .	21
10	エネルギー関数に関する諸定義 . . . . .	24
11	線分要素の投影 . . . . .	24
12	点 $s_i$ のバイリニア補間 . . . . .	24
13	疎行列 $Q$ のブロック構造 . . . . .	26
14	自由視点画像中に生じる色の不連続 . . . . .	28
15	未知領域と既知領域についての表記 . . . . .	28
16	ポアソン方程式を用いた色調補正結果 . . . . .	29
17	提案手法と比較手法の処理の流れ . . . . .	32
18	各データセットの入力画像例 . . . . .	33
19	復元した三次元モデル . . . . .	34
20	各データセットで生成する自由視点位置 (赤) および実視点位置 (青)	35
21	データセット (i) の仮想視点 A における自由視点画像 . . . . .	38
22	データセット (i) の仮想視点 A における自由視点画像の一部拡大画像	39
23	データセット (i) の仮想視点 B における自由視点画像 . . . . .	40
24	データセット (i) の仮想視点 B における自由視点画像 . . . . .	41
25	データセット (i) の仮想視点 C における自由視点画像 . . . . .	42
26	データセット (i) の仮想視点 C における線分情報の可視化結果 . .	43
27	データセット (ii) の仮想視点 A における自由視点画像 . . . . .	44

28	データセット (ii) の仮想視点 A における自由視点画像の一部拡大 画像 . . . . .	45
29	データセット (ii) の仮想視点 B における自由視点画像 . . . . .	46
30	データセット (ii) の仮想視点 B における自由視点画像の一部拡大画像	47
31	データセット (iii) の仮想視点 A における自由視点画像 . . . . .	48
32	データセット (iii) の仮想視点 A における自由視点画像の一部拡大 画像 . . . . .	49
33	データセット (iii) の仮想視点 B における自由視点画像 . . . . .	50
34	データセット (iii) の仮想視点 B における自由視点画像の一部拡大 画像 . . . . .	51
35	データセット (iv) の仮想視点 A における自由視点画像 . . . . .	52
36	データセット (iv) の仮想視点 A における自由視点画像の一部拡大 画像 1 . . . . .	53
37	データセット (iv) の仮想視点 A における自由視点画像の一部拡大 画像 2 . . . . .	54
38	データセット (iv) の仮想視点 B における自由視点画像 . . . . .	55
39	データセット (iv) の仮想視点 B における自由視点画像の一部拡大 画像 . . . . .	56

## 表 目 次

1	データセットの詳細 . . . . .	31
2	各データセットに用いたパラメータ . . . . .	31
3	投影に用いられた入力画像の枚数 . . . . .	37
4	計算機の構成 . . . . .	58

## 1. 緒言

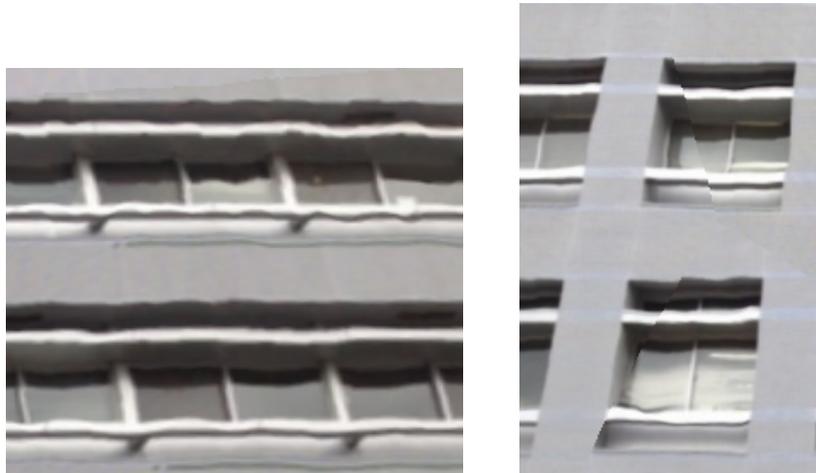
近年, Google Earth<sup>1</sup> に代表されるテレプレゼンスやデジタルアーカイブなどの仮想化現実空間を閲覧するアプリケーションが普及し, 一般に利用されている. このようなアプリケーションにおいて自由視点画像生成技術を用いれば, カメラで撮影された視点以外の任意視点からの画像をユーザに提示し, 高い臨場感を与えることが可能となるため注目を集めている. 任意視点における見えを再現する単純な方法として, テクスチャマッピングされた三次元モデルを手動で作成し用いることが考えられるが, モデルの作成には多くのコストを要するため, 近年, 撮影した画像群から自動的に自由視点画像を生成する手法が多く提案されている. これら自由視点画像生成に関する手法は三次元形状を用いる手法と用いない手法に大別される.

三次元形状を用いる手法は, 撮影された画像群に対して Muti-view Stereo (MVS) 法などを適用することで得られる三次元形状にテクスチャをマッピングし, これを設定した視点に応じてレンダリングすることで自由視点画像生成を行う. この手法では, 三次元復元の推定誤差に起因するテクスチャの膨張・欠損が生じるという問題があるため, 三次元形状とカメラ位置姿勢を利用し, 対象となる自由視点画像上で適切なテクスチャを画素毎に選択的に用いる, 視点依存テクスチャマッピングに基づく手法が近年多数提案されている [1, 2, 3]. この方法を用いることで, 三次元復元の精度が低い場合に生じるテクスチャの膨張・欠損はある程度改善されるが, 幾何学的に正しくないテクスチャの変形や複数画像からマッピングするテクスチャ間の不連続といった問題が依然として残っている (図 1). 本論文では図 1 に見られるような現象を歪みと呼び, 補償対象とする.

三次元形状を用いない手法の 1 つとして, 入力画像を格子状に区切りかつ入力画像上の線分を検出した上で, 線分の直線性が保たれるように格子を変形し自由視点画像を生成する手法 [4] が近年提案されている. この手法では, 入力画像中の線分の直線性を保持した画像変形により歪みの少ない画像を生成できるものの, 視差により画像上での前景と背景の位置関係が変わるような状況には対応できず, 設定する視点位置は入力画像のものと十分近くなければならない.

---

<sup>1</sup><https://www.google.co.jp/intl/en/earth/>



(a) 本来まっすぐな人工物の揺らぎ (b) テクスチャ間の不連続

図 1: 視点依存テクスチャマッピングで生じる歪みの例

このように、自由視点画像における違和感の低減を目指した手法が多数提案されているが、生成画像の品質または設定できる視点位置に限界がある。そこで、本研究では、入力視点から大きく離れた視点にも対応でき、三次元形状復元の精度の低さに起因する歪みを抑制した自由視点画像生成を目的として、三次元形状を用いる視点依存テクスチャマッピング [1] と線分の直線性を保持する画像の変形手法 [4] を組み合わせ、これに加えてテクスチャの繋ぎ目における連続性を考慮した自由視点画像生成手法を開発する。提案手法では、まず撮影された画像群を入力としてカメラ位置姿勢推定および対象シーンの三次元形状復元を行った上で、各入力視点および仮想視点に対応するデプスマップを生成する。次に、奥行き値を手がかりにした視点依存テクスチャマッピングにより各入力画像に対応した自由視点画像群を生成すると同時に、各入力画像から線分を検出し、自由視点画像上に投影する。次に、投影した線分の直線性および複数画像間の切り替わり位置におけるテクスチャの連続性を考慮したエネルギー関数最小化に基づく画像変形により自由視点画像の歪みを抑制する。最後に、視点依存マッピングされたテクスチャ間の光源の違いによる光学的な不整合を解決するため、ポアソン方程式を用いた色調補正を行う。

以下、第 2 章では自由視点画像生成に関する従来の取り組みを概観し、本研究

の位置付けについて述べる。第3章では、線分の直線性および画像間の対応点を用いたテクスチャの連続性を考慮した自由視点画像生成手法について詳述する。第4章では、屋外環境を撮影した画像群を入力として提案手法を適用した実験結果を示し、従来手法との比較を行う。

## 2. 自由視点画像生成の関連研究

本章では，自由視点画像生成に関する従来研究を，三次元形状を用いる手法と三次元形状を用いない手法に大別し，従来の取り組みを概観するとともに，本研究の位置付けについて述べる．

### 2.1 三次元形状を用いる手法

三次元形状を用いる自由視点画像生成手法は，まず撮影対象環境の三次元形状を事前に復元しておき，それと入力画像群を用いて任意の視点における画像を生成する．以下では，まず三次元形状復元手法を概観し，次に，従来の形状を用いた自由視点画像生成手法について述べる．

#### 2.1.1 三次元形状復元手法

仮想現実空間の構築，拡張現実感，デザイン設計などのアプリケーションにおいて，三次元モデルの作成に関する需要は高いが，人手での三次元モデルの作成には多くの手間や時間を要するため，カメラなどのセンサを利用して環境を計測した情報から三次元形状を自動的に復元する手法が多く提案されている．自動で三次元形状を復元する手法は，対象環境にレーザや投影光を照射して形状を計測する能動的な手法と，画像処理のみにより復元を行う受動的な手法に大別される．以下では，三次元形状復元手法を能動的な手法と受動的な手法に分類し，それぞれの手法の特徴について述べる．

#### 能動的な三次元形状復元手法

能動的な三次元形状復元手法は，レーザや投影光を対象環境に照射し，三次元形状の計測を行う．一般的な環境で用いることができる従来手法として，プロジェクタを用いたアクティブステレオ法 [5, 6] と，レーザレンジファインダを用いて照射したレーザが反射してセンサに届くまでの時間を計測する time-of-flight 方式の手法 [7] が挙げられる．アクティブステレオ法は，既知のパターンを投影することにより対象シーンの照明を能動的に制御することで，後述するステレオ

法における対応点の曖昧性に関する問題を緩和する。塚本ら [5] の手法ではプロジェクタによってスリット光を対象環境に照射し、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デプスマップとして取得する。Kawasaki ら [6] は、このような手法で煩雑となる事前のキャリブレーション処理を省略しつつ、コード化したパターン光投影により三次元形状を高精度に復元する手法を提案している。これらの手法では、プロジェクタとカメラ間の距離（ベースライン）と、カメラと対象物間の距離の比によって計測精度が決まり、後者に対してベースラインを比較的長く設定可能な小物体に対しては、高精度な計測が可能である。しかし、屋外などの広域環境を対象とする場合には、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照射光の観測が困難となる問題がある。一方、time-of-flight 方式 [7] では、撮影地点ごとに各画素で対象物までの距離を保持した距離画像が 1 視点での計測において得られる。レーザーの反射強度が一定以上あれば、高精度な距離画像の獲得が可能であり、屋外のような広域環境に対しても計測を行うことができる。Time-of-flight 方式の問題点として、例えば RIEGL 社の LMS-Z360 を用いて各地点において広範囲かつ空間的に密に計測を行うためには、計測に数分程度の時間がかかる。ゆえに、街全体などの広範囲な領域を計測するためには、後述する受動的な三次元形状復元手法よりも多くの時間を要する。

一般に、アクティブステレオ法または time-of-flight 方式のいずれの手法においても、1 視点での計測可能な範囲は限られるため、対象シーン全体の三次元形状を復元するためには、複数地点で得た三次元形状を統合する必要がある。統合における位置合わせに際して、計測した各三次元形状間である程度重複した共通領域が必要となるため、計測地点は密にしておく必要がある。三次元形状の統合手法として、Zhao ら [8] はデプスマップ上での二次元的な対応点探索による位置合わせを、Früh ら [9] は航空写真を併用した位置合わせを行うことで、三次元形状情報の統合し、対象環境の三次元モデルを構築している。

### 受動的な三次元形状復元手法

受動的な三次元形状復元手法は、対象環境を撮影した画像群を用いて画像処理により三次元形状を復元する。主な従来手法としては視体積交差法やステレオ法

が挙げられる。これらの手法では撮影に用いたカメラの各視点における位置姿勢が既知であるという前提を用いる。カメラの位置姿勢を求める手段として、カメラに搭載されたGPS、ジャイロセンサ、方位センサなどの外部指標を組み合わせる方法があるが、GPSは数メートル範囲の測位誤差が生じ、またGPSの性質上使用可能な環境は屋外に限定される。また、ジャイロセンサにおいても蓄積誤差の問題があり、視体積交差法やステレオ法に用いるには精度が不十分である。これに対して、撮影された複数枚の画像間で特徴点の対応付けを行うことでカメラ位置姿勢を高精度に推定するStructure from Motion (SfM)法が提案され、これに基づくBundler [10, 11] や VisualSFM [12, 13] などのソフトウェアがウェブ上で公開されており、利用することができる。

視体積交差法 [14, 15] は、撮影空間をボクセルの集合として表現し、そのボクセル空間内で多視点から撮影された物体のシルエットとカメラの光学中心からなる視体積の積集合空間 (Visual Hull [16]) を計算することで三次元モデルを生成する。この手法では画像間の対応付けが必要でないため、テクスチャのない物体でも計測が可能である。しかし、シルエットを抽出した際に輪郭として現れない形状は推定できないため、原理的に物体の凹みを復元できないという制約がある。

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を用い、画像中の各画素の対応付けを行うことで、三角測量の原理によりデプスマップを生成する。ステレオ法の中でも、多数の画像を同時に利用することで対応点の曖昧性を軽減したMVS法は二眼ステレオ法と比較して精度が高いため、近年多くの手法が提案されている [17, 18, 19, 20, 21]。Goeseleら [18] や Furukawaら [19] は、Flickr<sup>2</sup> に代表される写真共有サイトにアップロードされた膨大な写真データベースを利用して、広域環境を対象とした密な三次元形状を復元する手法を提案している。Goeseleら [18] の手法では、光源、天候、解像度など様々な条件の異なる膨大な画像群を、見た目や解像度の項目ごとに自動でサブセット分割することで、画像間の対応付けを簡単化している。Furukawaら [19] の手法では、複数画像を画像間の対応付けのために重複領域を持つようにクラスタリングし、各クラスタの画像から投影した三次元点について他のクラスタの投影点との色の一致度

---

<sup>2</sup><https://www.flickr.com/>

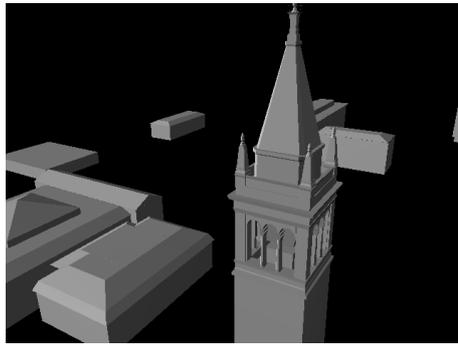


図 2: Furukawa ら [19] の手法によるレンダリング結果

を測りフィルタリングすることで外れ値を弾き，誤差の少ない三次元復元を行っている．これらの手法では，テクスチャの少ない領域などで対応付けがうまくいかず，形状の推定精度が悪くなるため，生成した三次元形状に欠損が生じる場合がある（図 2）．この問題に対して，Jancosek ら [20] の手法では，SfM 法などで推定された三次元点群において点群が疎に存在し復元が困難な領域を対象に面を貼付け，Visual Hull を用いて再構成することにより密な三次元復元を行う．これにより，欠損の少ない三次元モデルを生成することができる．

### 2.1.2 三次元形状を用いた自由視点画像生成手法

従来から行われている三次元形状を用いる最も単純な方法として，復元した三次元形状にあらかじめテクスチャマッピングしておき，それを設定した仮想視点に投影することで自由視点画像を生成する手法がある．このような手法では，対象物体を複数の異なる地点から撮影した画像群から画像を選択し，またブレンディングすることで，各部分形状に対するテクスチャを決定している．例えば，Waechter ら [22] は，写実性の高いテクスチャをマッピングするために，入力画像群と三次元モデルに貼られた面のテクスチャの色的一致度が向上するようなエネルギー関数を定義し最適化することで，三次元モデルのテクスチャを決定する手法を提案している．しかし，これらの手法は復元される三次元モデルの精度が高



(a) 三次元形状



(b) レンダリング結果

図 3: 視点依存テクスチャマッピング [1] によるレンダリング

いことを前提としており，復元精度が悪い場合には，テクスチャに歪みが生じる．

この問題に対し，形状の復元精度があまり良くない場合であっても，仮想視点の各画素に対して入力画像群からテクスチャを適応的に選択することで，テクスチャの歪みを低減しかつ写実性の向上を図る視点依存テクスチャマッピング [1, 23] が提案されている．Debevec ら [1] は，三次元形状に対して仮想視点位置と角度的に近い入力画像の視点位置の画像を画素毎に選択して色をマッピングする手法を提案している（図 3）．小甲ら [23] は，角度だけでなく対象物体とカメラの距離についても考慮し，より対象物体と距離の近いカメラを採用することで，解像度の高いテクスチャを選択するような枠組みを提案している．これらの手法では，仮想視点と見えの近い画像を選択するため，三次元形状復元の精度があまり良くない場合においても，単純なテクスチャマッピングよりは歪みの少ない画像を生成できるが，仮想視点が入力画像群の撮影視点から離れた場合には，図 1 に示すような歪みが生じてしまう問題が残されている．

これに対して，テクスチャだけでなく視点に応じて適応的に形状を変形させることで見えを改善する視点依存ジオメトリに基づく自由視点画像生成手法が提案されている [2, 3, 24, 25, 26]．Rademacher [24] は，アニメーションにおけるキャラクターの任意の視点における見えの再現を目的として，キーとなる視点での三次元モデルの変形（キー変形）を与えておき，設定する仮想視点の位置に応じてキー変形を補間して三次元モデルをレンダリングする手法を提案している．

この手法ではあらかじめ三次元モデルに対してキー変形を与えておく必要がある。Nakashimaら [25] は、拡張現実感で利用する仮想物体において、視点依存テクスチャマッピングを用いてレンダリングする場合の境界付近の欠損に着目し、仮想視点のデプスマップにおける対象物体の輪郭領域に対して膨張処理を適用し、かつ入力画像中の対象物体の前景抽出を行うことで、境界に欠損のない仮想物体の自由視点画像生成を実現している。この手法では、事前に前景抽出をする必要があるため、背景あるいは前景が単純なテクスチャで簡単に分離出来る場合には GrabCut [27] などを用いて前景・背景を領域分割する。Rademacher [24] や Nakashimaら [25] の手法は、特定物体を対象として視点依存の形状を用いることで高品質な自由視点画像生成を可能としており、映像制作や拡張現実感に応用される。

一方、Evers-Senneら [26] は、広域かつ複雑なシーンを対象とした、仮想視点に応じて変形する三次元メッシュモデルを用いた自由視点画像生成手法を提案している。この手法では、まず設定した仮想視点に対する入力画像のカメラ位置・姿勢および視野に基づき、テクスチャマッピングに利用する画像の優先度を決定しておき、その優先度および各入力画像のデプスマップに基づき仮想視点でのメッシュモデルを変形し、これを用いてレンダリングを行う。また、伊吹ら [2] は、空間的整合性だけでなく、時間的整合性も考慮したメッシュ変形とテクスチャ選択により、仮想視点の移動における違和感も考慮した自由視点画像生成手法を提案している。しかし、これらの手法では、画像の解像度に対して低い解像度の三次元メッシュモデルを用いているため、画像中の形状の周波数が高い領域において、形状が正しく再現されず、テクスチャに歪みが生じる場合がある (図4)。Chaurasiaら [3] の手法では、事前の三次元復元において草木のような画像間の対応付けが困難なテクスチャ領域では奥行き値がうまく得られないことに着目し、事前処理として画像を色の近い小領域 (superpixel [28]) に分割した上で、奥行き値を持たない superpixel に対して、画像の色が類似しかつ距離が近い他の superpixel を用いることで、奥行き値の補間を行う。次に、オンライン処理では、superpixel ごとにメッシュを設定し、復元または補間された奥行きを用いて変形することで、自由視点画像を生成する。これにより、従来奥行きがうまく得られないために復

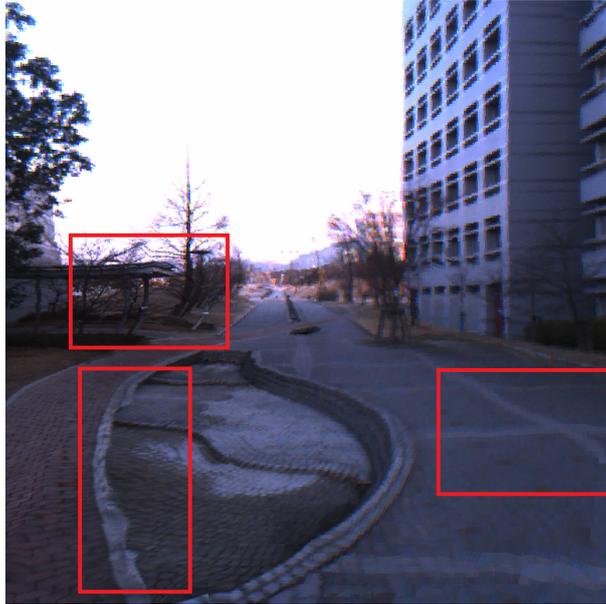


図 4: 伊吹ら [2] の手法における自由視点画像上の歪み

元が困難であった複雑形状をもつ物体に対しても違和感のない自由視点画像を生成している。しかし、全く奥行き値が得られないかつ似た色・テクスチャを持つ物体が隣接している場合や、あまりに細かい構造で superpixel に分割できない物体が存在する場合には、アーティファクトが生じることが重要な制限として示されている。

このように、三次元形状を用いる自由視点画像生成手法では、広域環境を対象とした手法が多く提案されており、三次元形状の復元精度があまり良くない場合においても、視点依存テクスチャ・ジオメトリを利用することで写実性の向上を実現している。一方で、設定する仮想視点位置を入力画像の視点位置と大きく離れた場合には、三次元形状の復元精度の低さに起因するテクスチャの歪みが生じるという問題が残されている。

## 2.2 三次元形状を用いない手法

三次元形状を用いない自由視点画像生成手法は、対象環境の三次元形状を明示的に復元することなく、撮影され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。三次元形状を用いる手法では正確な幾何形状の推定が困難な複雑形状をもつ物体に対しても欠損を生じず、写実的な自由視点画像生成が可能である。以下では、三次元形状を用いない自由視点画像生成手法を、光線情報を用いる手法と、画像群の変形・合成による手法に分類し、それぞれの手法について述べる。

光線情報を用いる手法として、Plenoptic Function [29], 光線空間法 [30], Light Field Rendering [31], Lumigraph [32] などがある。これらの手法では、空間内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を生成する。光線を記録する手法として、Adelson ら [29] は、光線の通過位置  $(x, y, z)$ , 光線方向  $(\theta, \phi)$ , 波長  $\lambda$  および時間  $t$  の 7次元データを用いて空間内のすべての光線を記述する Plenoptic Function を提案している。この手法において、7次元のデータをすべて記録するためには膨大な量の画像が必要となり、またすべてのデータをサンプリングすることは現実的に難しい。これに対して Naemura ら [30] 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を提案している。光線空間法は、光線が空間内を干渉や減衰などの影響を受けずに直進するものと仮定することにより、7次元の Plenoptic Function から  $z$  座標、波長  $\lambda$  および時間  $t$  を除外し、光線空間を 4次元にすることで、取得すべき画像枚数を大幅に減少させている。光線空間法と原理的に等価な手法として、Levoy ら [31] および Gortler ら [32] は、同時期にそれぞれ Light Field Rendering, Lumigraph を提案している。これらの手法では、光線空間法が 4次元の光線空間を光線の通過位置座標  $(p, q)$  および光線方向  $(\theta, \phi)$  で表現したのに対して、この 4次元を 2つの並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とする。これらのアプローチは、大量の画像を用意すれば高品質な自由視点画像生成が可能となるが、必要となるメモリ量、計算量も膨大となる。十分な画像枚数がない場合には、視点移動に伴うテクスチャの幾何学的な歪みが生じるという問題がある。

画像群を直接変形・合成して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 [33, 34] が提案されている。これらの手法では、入



図 5: Huang ら [4] による結果の比較

力の複数画像間で対応点を与え、自由視点位置においてその対応点がどの位置に現れるかを計算し、入力画像を変形させることで自由視点画像を生成する。手動で対応点を与えるだけで簡単に自由視点画像を生成できるという特徴があるが、対象の三次元形状を考慮しないため、視点移動に伴いテクスチャの幾何学的な歪みが生じ、また、対応点の数に結果が大きく依存してしまうという問題がある。この問題に対して、Wang ら [35] は、2 視点の中間視点を生成する際に結果が対応点数に大きく依存することを防ぐ目的で、画像から検出した線分情報を併用し、線分の直線性を保持するモーフィング手法を提案している。類似したアイデアの手法として、Huang ら [4] は、入力画像中の線分の直線性を自由視点画像中でも保つような制約を加えることで自由視点画像に生じる幾何学的な歪みを抑えるワーピングベースの自由視点画像生成手法を提案している（図 5）。

これらの手法を用いれば、三次元形状を用いる手法では正しく形状を得られないためにレンダリングが困難な複雑形状を持つ物体も写実的にレンダリングすることが可能であるが、三次元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はできない。さらに、入力画像の変形により自由視点画像を生成するため、入力画像から視点が大きく変わる場合には、テクスチャの幾何学的な歪みが大きくなる傾向がある。

## 2.3 本研究の位置付けと本論文の構成

既に述べたように，自由視点画像生成手法は数多く提案されている．三次元形状を用いた手法は，物体の隠蔽関係の表現に優れ，設定できる視点位置が広範囲であるが，三次元形状の推定誤差に起因するテクスチャの歪みが生じやすい．三次元形状を用いない手法では，写実的な表現が可能であるが，高品質な画像を生成するためには設定可能な視点位置に大きな制約がある．本研究は，上記の問題を解決し，入力画像の視点位置から大きく離れた仮想視点位置においても三次元形状復元の精度の低さに起因する歪みを抑制した自由視点画像を生成することを目的としている．そのために，三次元形状を用いた視点依存テクスチャマッピングと三次元形状を用いないワーピング手法の組み合わせに基づく自由視点画像生成を行う．入力画像の視点位置から大きく離れた視点位置における画像生成を実現するために，三次元形状復元と，奥行き値を参照した自由視点への投影を行う．このとき，生成した画像上に本来直線であるはずの線分に歪みが生じるため，従来研究でも行われている画像中の線分の直線性を保持したワーピングベースの手法を応用し，エネルギー関数の最小化により自由視点画像上で歪みが生じる線分要素を補正する．加えて，複数画像間から視点依存テクスチャマッピングした際に生じるテクスチャ間の不連続を抑制する新たな指標を導入する．

### 3. 線分および対応点を考慮した自由視点画像の歪みの抑制

本章では、撮影した画像群を入力とした線分の直線性および複数画像からマッピングされるテクスチャ間の連続性を保持する自由視点画像生成手法について詳述する。以下では、まず提案手法の概要について述べ、次に、仮想視点への画像群・線分の投影とカメラマップの生成、テクスチャの不連続箇所での対応点探索、エネルギー最小化による画像変形および後処理として行う色調補正について述べる。

#### 3.1 提案手法の概要

提案手法の流れを図6に示す。提案手法では、まず VisualSfM [12, 13] 等の SfM 法を用いてカメラ位置姿勢を推定した上で、CMPMVS [20] 等の MVS 法を用いてシーンの三次元形状を復元し (a)、各入力視点および仮想視点のデプスマップを生成する (b)。これを用いて各入力画像を仮想視点に投影することで、自由視点画像生成に利用する参照画像群を生成するとともに、各入力画像から抽出した線分を仮想視点に投影する (以降、投影された線分を投影線分と呼ぶ)。同時に、視点依存テクスチャマッピング (VDTM) [1] の基準により画素毎に投影に採用されるカメラのインデックスを保存した画像 (以降、カメラマップと呼ぶ) (図7) を作成する (c)。また、カメラマップ中の異なるインデックスが隣り合う境界線をまたぐ一定の幅を持つ領域 (以降、境界領域と呼ぶ) を設定し、境界領域においてその境界を挟む2枚の参照画像間で対応点を求める (d)。以上で求めた情報を用いてエネルギー関数を最小化することで、各参照画像をそれぞれ変形し (e)、これらを合成することで、1枚の自由視点画像を生成する。その際、ポアソン方程式を用いた色調補正により、テクスチャ間の色の不連続を補正する (f)。

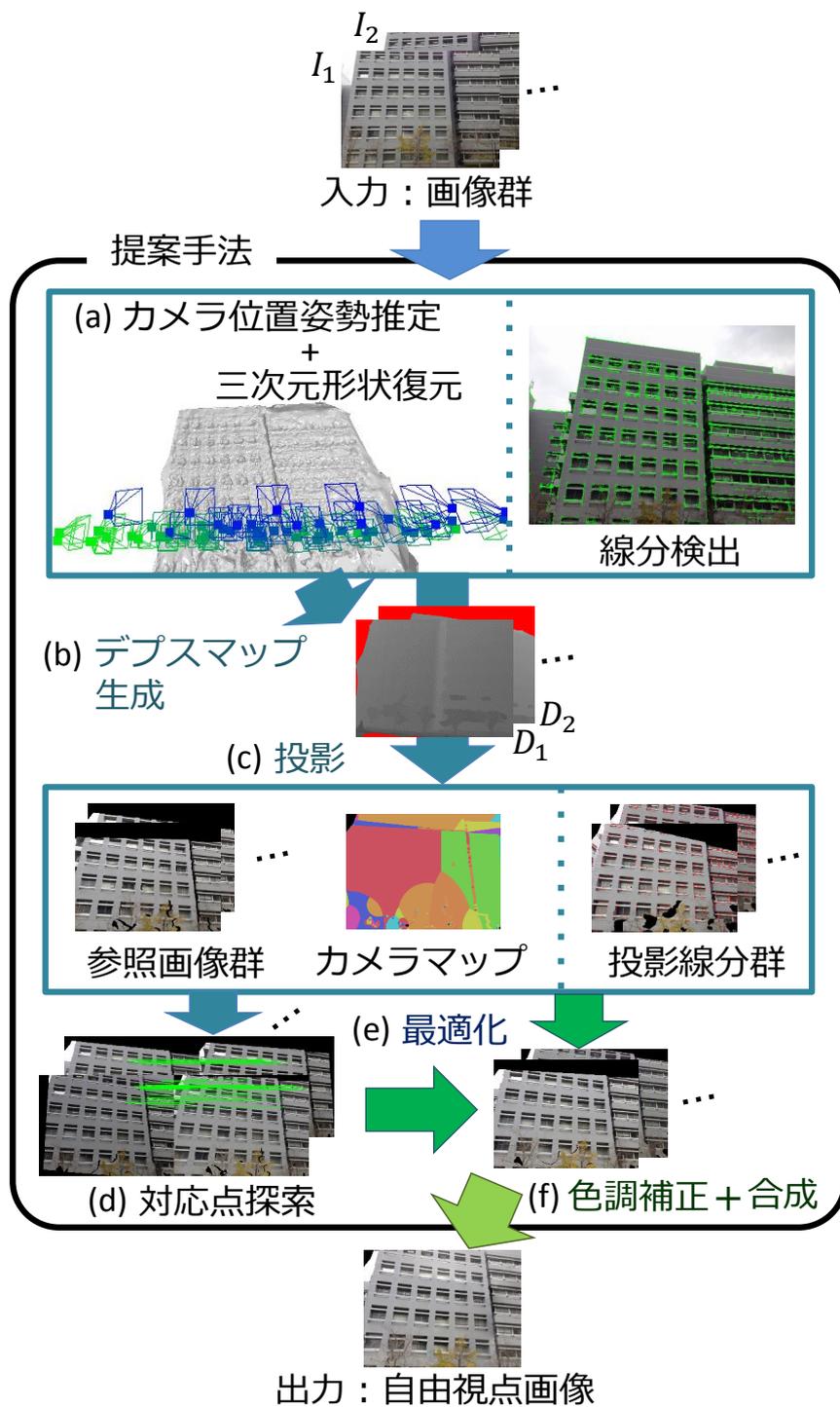


図 6: 提案手法の概要



(a) 視点依存テクスチャマッピングにより生成した自由視点画像 (b) (a) に対応するカメラマップ

図 7: カメラマップの例

### 3.2 仮想視点に対応する投影データの生成

本節では、各視点のデプスマップを利用した VDTM による仮想視点に対応する参照画像群、投影線分群およびカメラマップの生成方法について詳述する。VDTM による仮想視点への投影処理には入力画像視点のデプスマップを参照し投影する順変換と、仮想視点のデプスマップを参照し投影を行う逆変換があり (図 8), 本研究ではこれらを画像の投影と線分の投影プロセスで使い分ける。以下、画像の投影、カメラマップの生成および線分の投影について順に述べる。

#### 3.2.1 画像群の投影

実画像群から仮想視点に対応した参照画像群を逆変換により生成する。  $k$  番目のカメラ  $C_k$  に対応する入力画像を  $I_k$ , 仮想視点のカメラ  $C_V$  に対応する参照画像を  $I_V$ , 画像  $I_k$  の画素位置  $(u, v)$  における画素値を  $I_k(u, v)$  と表すものとする。このとき、参照画像中のある画素  $n \in I_V$  の画素値の算出方法を以下に説明する。

画素  $n$  の画像座標を同次座標で  $\mathbf{p}_{(V,n)} = (u, v, 1)^\top$  と表し、デプスマップにおける画素  $n$  の奥行き値が  $d_{(V,n)}$  であるとする。  $C_V$  の光学中心と  $C_V$  の画像平面上の画素  $n$  を結んだ直線が三次元形状と交わる世界座標系における同次座標で表した三次元点  $\mathbf{P}_n \in \mathbb{R}^4$  は、カメラ  $C_V$  の内部パラメータ  $\mathbf{A}_V \in \mathbb{R}^{3 \times 3}$  および

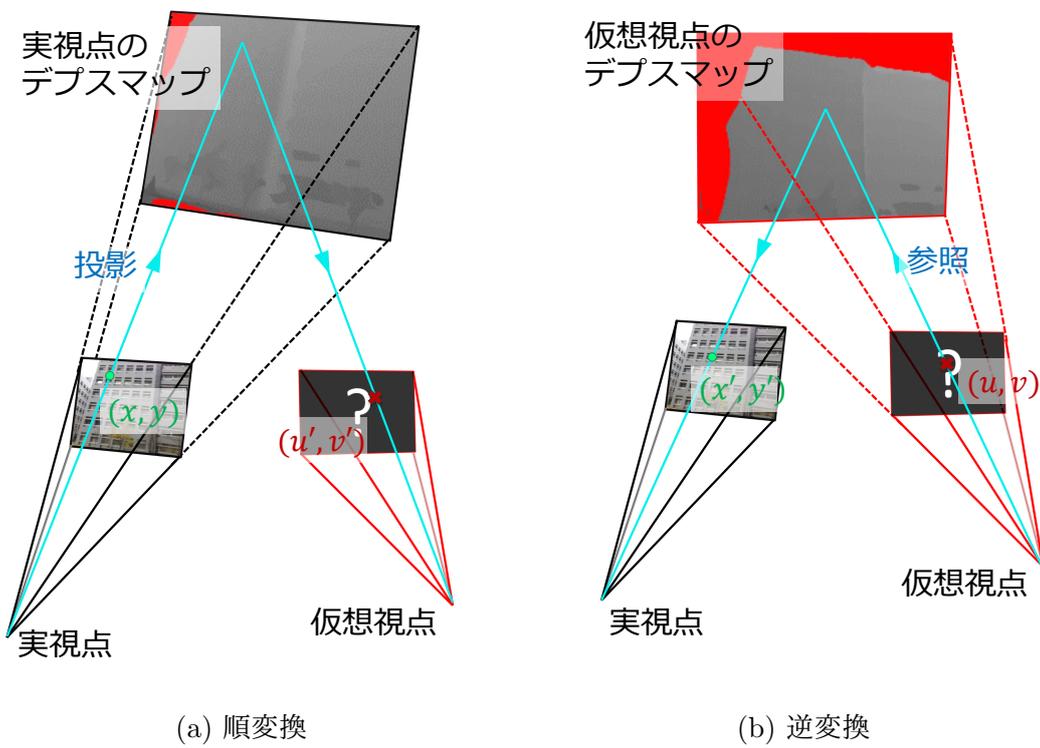


図 8: 投影方法の種類

外部パラメータ  $[\mathbf{R}_V \in \mathbb{R}^{3 \times 3}, \mathbf{t}_V \in \mathbb{R}^3]$  を用いて次のように表すことができる。

$$\mathbf{P}_n = d_{(V,n)} \begin{bmatrix} \mathbf{R}_V^\top & -\mathbf{R}_V^\top \mathbf{t}_V \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}_V^{-1} \mathbf{p}_{(V,n)} \\ 1 \end{bmatrix} \quad (1)$$

$\mathbf{P}_n$  を入力画像視点  $C_k$  に投影し、画像座標位置  $\mathbf{p}_{(k,n')}$  とその座標における奥行き値  $z_{(k,n')}$  を以下のように求める。

$$\hat{\mathbf{p}}_{(k,n')} = \begin{pmatrix} x_{(k,n')} \\ y_{(k,n')} \\ z_{(k,n')} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R}_k & \mathbf{t}_k \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{P}_n \quad (2)$$

$$\mathbf{p}_{(k,n')} = \begin{pmatrix} \frac{x_{(k,n')}}{z_{(k,n')}} \\ \frac{y_{(k,n')}}{z_{(k,n')}} \\ z_{(k,n')} \\ 1 \end{pmatrix} \quad (3)$$

ただし、投影位置はサブピクセル精度で求められるため、 $n'$  は最近傍の画素を表す。次に、カメラ  $C_k$  から三次元点  $\mathbf{P}_n$  が可視であるかどうかをデプステストにより判定する。デプステストでは、投影した奥行き値  $z_{(k,n')}$  と入力画像視点のデプスマップから得られる投影位置の画素  $n'$  の奥行き値  $d_{(k,n')}$  を次式により比較し、これを満たす場合に  $\mathbf{P}_n$  は入力画像視点  $C_k$  から可視であるとする。

$$z_{(k,n')} < d_{(k,n')} + e \quad (4)$$

ただし、 $e$  はデプステストにおける許容誤差である。デプステストの結果可視であると判定された場合、決定した画素位置  $\mathbf{p}_{(k,n')} = (u', v', 1)^\top$  から、参照画像の画素値  $I_V(u, v)$  を各カラーチャンネルについてバイリニア補間により以下のように算出する。

$$\begin{aligned} & I_V(u, v) \\ &= \begin{pmatrix} [v'] + 1 - v' & v' - [v'] \end{pmatrix} \begin{pmatrix} I_k([u'], [v']) & I_k([u'] + 1, [v']) \\ I_k([u'], [v'] + 1) & I_k([u'] + 1, [v'] + 1) \end{pmatrix} \begin{pmatrix} [u'] + 1 - u' \\ u' - [u'] \end{pmatrix} \end{aligned} \quad (5)$$

### 3.2.2 カメラマップの生成

VDTM の基準により，仮想視点画像での各画素に対して利用する入力画像のインデックスを表すカメラマップを生成する．3.2.1 項で求めた三次元点  $\mathbf{P}_n$  を用い，カメラマップの画素  $n$  に採用する入力画像のインデックスを，以下のアルゴリズムにより決定する．

**Step 1.** カメラ  $C_k$  の光学中心と三次元点  $\mathbf{P}_n$  を結ぶ直線およびカメラ  $C_V$  の光学中心と三次元点  $\mathbf{P}_n$  を結ぶ直線のなす角を求める．

**Step 2.** Step 1. をすべてのカメラについて計算する．

**Step 3.** 角度の小さい順にカメラのインデックスをソートする．

**Step 4.** ソートされたインデックスの先頭から，カメラから三次元点  $\mathbf{P}_n$  が可視であるかどうかをデプステストにより判定する．

**Step 5.** デプステストの結果，不可視であれば，次のインデックスに移り Step 4. へ．

**Step 6.** デプステストの結果，可視であれば，そのインデックスのカメラを採用する．

Step 1. において，カメラ  $C_V$  の光学中心の世界座標系での三次元点  $\mathbf{O}_V$  は  $\mathbf{P}_n$  の算出と同様に，以下のように表される．

$$\mathbf{O}_V = \begin{bmatrix} \mathbf{R}_V^\top & -\mathbf{R}_V^\top \mathbf{t}_V \\ \mathbf{0}^\top & 1 \end{bmatrix} \hat{\mathbf{o}}_V = -\mathbf{R}_V^\top \mathbf{t}_V \quad (6)$$

ただし， $\hat{\mathbf{o}}_V = (0, 0, 0, 1)^\top$  である．このとき，カメラ  $C_k$  の光学中心と三次元点  $\mathbf{P}_n$  を結ぶ直線およびカメラ  $C_V$  の光学中心と三次元点  $\mathbf{P}_n$  を結ぶ直線のなす角  $\theta_{(k,n)}$  に対応する単純類似度は以下のように求められる．

$$\cos \theta_{(k,n)} = \frac{(\mathbf{P}_n + \mathbf{O}_k) \cdot (\mathbf{P}_n + \mathbf{O}_v)}{|\mathbf{P}_n + \mathbf{O}_k| |\mathbf{P}_n + \mathbf{O}_v|} \quad (7)$$

求めた単純類似度  $\cos \theta_{(k,n)} \in [-1, 1]$  の値が大きいほど，角度は小さくなるため，単純類似度を角度の代わりに用いる．

### 3.2.3 線分の投影

入力画像上の線分について，自由視点画像上においてもその直線性を保持するため，入力画像  $I_k$  から線分を検出し，参照画像  $I_V$  上に順変換を用いて投影する．提案手法では，ゼロ交差法で検出したエッジを用いて，確率的ハフ変換 [36] により線分を検出する．次に，検出した線分  $l$  をエイリアシングの影響が生じないよう等間隔かつサブピクセル精度で標本化する．

$$l = \{ \mathbf{s}_{(k,0)}, \mathbf{s}_{(k,1)}, \dots, \mathbf{s}_{(k,i)}, \dots \} \quad (8)$$

ここで， $\mathbf{s}_{(k,i)}$  は線分  $l$  上の  $i$  番目の標本点 (以下，線分要素と呼ぶ) の二次元座標を表す．このとき， $\mathbf{s}_{(k,i)}$  に対応する三次元形状の世界座標系における三次元点  $\mathbf{P}_i$  を求める．ただし，入力画像  $I_k$  に対応するデプスマップの奥行き値はサブピクセル精度では得られないため，本論文では最近傍の奥行き値  $d_{NN}$  を採用する．

$$\mathbf{P}_i = d_{NN} \begin{bmatrix} \mathbf{R}_k^\top & -\mathbf{R}_k^\top \mathbf{t}_k \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}_k^{-1} \mathbf{s}_{(k,i)} \\ 1 \end{bmatrix} \quad (9)$$

これを仮想視点の画像座標上に投影することで，各線分要素の位置を保存する．

$$\hat{\mathbf{s}}_{(V,i')} = \begin{pmatrix} x_{(V,i')} \\ y_{(V,i')} \\ z_{(V,i')} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R}_V & \mathbf{t}_V \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{P}_i \quad (10)$$

$$\mathbf{s}_{(V,i')} = \begin{pmatrix} \frac{x_{(V,i')}}{z_{(V,i')}} \\ \frac{y_{(V,i')}}{z_{(V,i')}} \\ \frac{z_{(V,i')}}{z_{(V,i')}} \\ 1 \end{pmatrix} \quad (11)$$

なお，仮想視点上においてもエネルギー関数最小化の際のエイリアシングを防ぐために，線分要素の投影位置はサブピクセル精度で保存しておく．

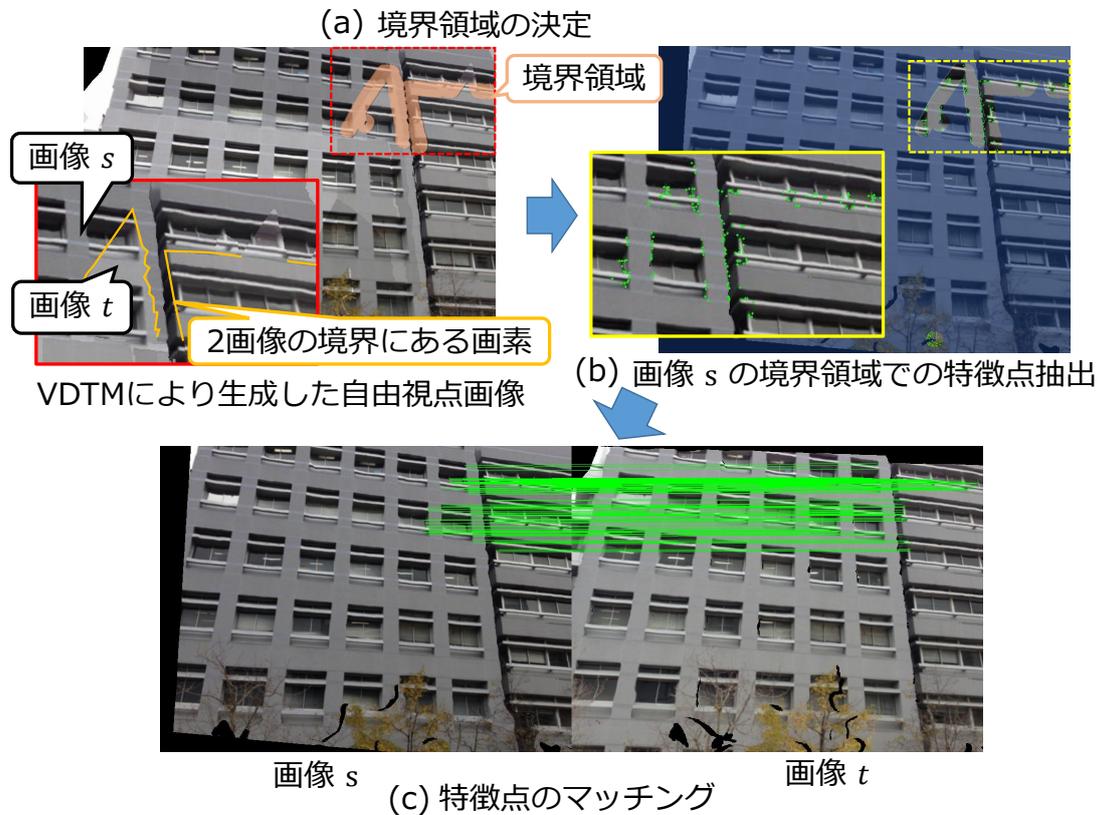


図 9: 境界領域における対応点探索

### 3.3 境界領域における対応点探索

前節で述べた各入力画像からの仮想視点における参照画像生成後、カメラマップの境界領域において対応点探索を行う。図 9 に境界領域における対応点探索の概要を示す。画像  $s$  および画像  $t$  が隣り合う領域にマッピングされたとき、その境界領域は、2 画像の境界に存在する画素に対してあらかじめ設定した回数の膨張処理を施すことで決定する (a)。次に、境界領域において Harris のコーナー検出 [37] により一方の画像で特徴点を抽出し (b)、もう一方の画像において各特徴点を中心とする一定領域内でゼロ平均正規化相互相関を評価したパターンマッチングにより対応点を探索する (c)。

### 3.4 線分および対応点を考慮した画像変形

本節では、投影した線分および生成した複数の参照画像間の対応点を利用した画像変形手法について詳述する。

#### 3.4.1 エネルギー関数の定義

本項では、Huang ら [4] の手法をベースとした、参照画像群をワーピングする際に用いるエネルギー関数の定義について述べる。提案手法では、線分の直線性に加え、マッピングする画像間のテクスチャの連続性を考慮したエネルギー関数  $E$  を以下のように定義する。

$$E = \sum_{k=1}^K \{E_{F_k} + \alpha E_{C_k} + \beta E_{L_k}\} + \gamma \sum_{(s,t) \in \mathcal{S}} E_{M_{(s,t)}} \quad (12)$$

ここで、 $K$  は投影に用いる画像枚数、 $\mathcal{S}$  は隣り合う領域に投影される画像インデックスのペアの集合、 $\alpha$ 、 $\beta$  および  $\gamma$  は重み係数である。Huang ら [4] の手法では、疎な特徴点位置の補間が変形後の画素位置の基準の一つとして用いられるため、メッシュを用いて画像を変形しているが、提案手法では三次元形状を用いたワーピング位置を変形後の各画素の位置の基準として用いることができる。従って、本最適化においてはメッシュを用いず、VDTM により色の投影されない画素を除く全ての画素位置を最適化のパラメータとする。以下、 $E_{F_k}$ 、 $E_{C_k}$ 、 $E_{L_k}$ 、 $E_{M_{(s,t)}}$  について述べる。

参照画像  $k$  の画素  $n$  は三次元形状を用いた投影位置が信頼できるものとし、投影位置から大きく移動しないための制約項  $E_{F_k}$  を以下のように定義する。

$$E_{F_k} = \sum_{n \in \mathcal{X}} \|f(\mathbf{x}_n^k) - \mathbf{x}_n^k\|^2 \quad (13)$$

ここで、 $\mathcal{X}$  は画素のインデックスの集合、 $\mathbf{x}_n^k$  は画像  $k$  における画素  $n$  の変形前の画素位置、 $f(\mathbf{x}_n^k)$  は変形後の画素位置を表す。ただし、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  は画素毎に独立である。

テクスチャ構造を保つために、画素の位置関係の変化を抑制する制約項  $E_{C_k}$  を以下のように定義する.

$$E_{C_k} = \sum_{(j_1, j_2) \in \mathcal{N}} \|(f(\mathbf{x}_{j_1}^k) - f(\mathbf{x}_{j_2}^k)) - (\mathbf{x}_{j_1}^k - \mathbf{x}_{j_2}^k)\|^2 \quad (14)$$

ここで、 $\mathcal{N}$  は隣り合う画素のインデックスのペアの集合である. これにより、図 10(a) のように、隣接画素を結ぶベクトルの回転や大きさの変化を抑制している.

$E_{L_k}$  項は、仮想視点に投影した線分要素がデプスマップの精度の低さに起因して歪む現象 (図 11) を抑制する目的をもつ. 3.2.3 項で述べたように、投影された線分の要素はサブピクセル精度で保存されるため、パラメータとして直接参照することができない. そこで図 12 のように、線分要素の位置をそれを囲む 4 画素の位置を用いて表す. 任意の線分要素  $\mathbf{s}_i$  は、画像変形後の画素位置群を  $\mathbf{V}$ 、線分要素  $\mathbf{s}_i$  に関するバイリニア補間の重みベクトルを  $\Phi_{\mathbf{s}_i}$  としたとき、 $\mathbf{s}_i = \Phi_{\mathbf{s}_i}^\top \mathbf{V}$  と表される. なお、 $\Phi_{\mathbf{s}_i}$  は 4 つの非零要素を持ったベクトルであり、図 12 に示す面積比を用い、以下のように表される.

$$\Phi_{\mathbf{s}_i} = (\dots, (1-c)(1-d), \dots, c(1-d), \dots, (1-c)d, \dots, cd, \dots)^\top \quad (15)$$

以上の関係を用いて、 $E_{L_k}$  は以下のように定義される.

$$E_{L_k} = \sum_{l \in \mathcal{L}_k} \omega_l \sum_{i=1}^m \|\Phi_{\mathbf{s}_i} \mathbf{V} - (1-h)\Phi_{\mathbf{s}_{i-1}} \mathbf{V} - h\Phi_{\mathbf{s}_{i+1}} \mathbf{V}\|^2 \quad (16)$$

ここで、 $\mathcal{L}_k$  は画像  $k$  の投影線分の集合、 $\mathbf{s}_i$  は投影線分  $l$  を構成する一つの線分要素の位置、 $\omega_l$  は入力画像中で検出した線分の長さに比例した重みを表す. 入力画像上では線分を等間隔にサンプリングしたが、投影する際に物体の隠蔽関係により見えなくなる、あるいはデプステストで弾かれることがある. そこで、端点  $\mathbf{s}_{i-1}$ 、 $\mathbf{s}_{i+1}$  を通る直線に点  $\mathbf{s}_i$  から下ろした垂線の足に点  $\mathbf{s}_i$  が移動するようにする (図 10(b)).

投影された画像の隣り合う領域の画素の動きを制約する項  $E_{M(s,t)}$  を以下のように定義する.

$$E_{M(s,t)} = \sum_{(a,b) \in \mathcal{M}(s,t)} \sum_{\mathbf{w} \in \mathcal{W}} g(\mathbf{w}) \|f(\mathbf{x}_a^s + \mathbf{w}) - f(\mathbf{x}_b^t + \mathbf{w})\|^2 \quad (17)$$

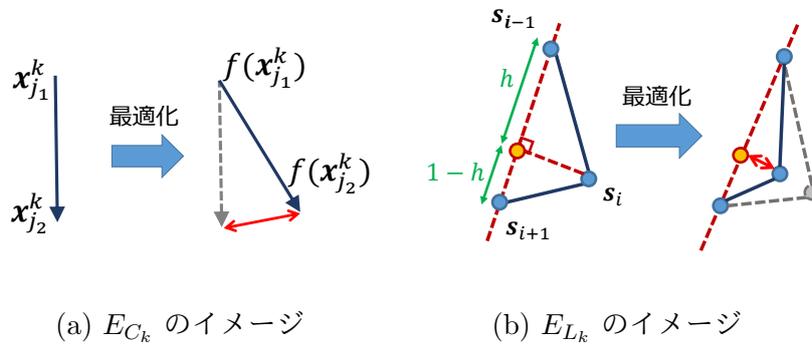


図 10: エネルギー関数に関する諸定義

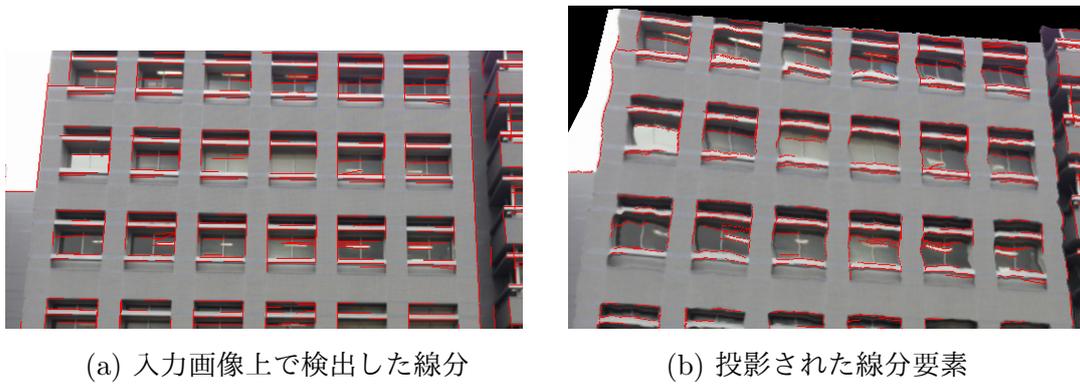


図 11: 線分要素の投影

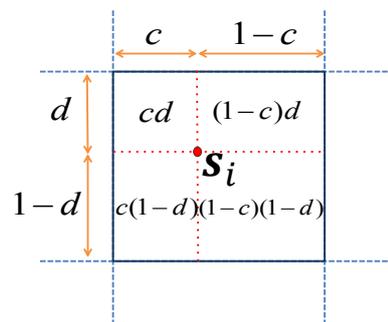


図 12: 点  $s_i$  のバイリニア補間

ここで、 $\mathcal{M}(s, t)$  は画像  $s$  と画像  $t$  の参照画像上での境界領域における対応点のインデックスの集合、 $\mathcal{W}$  は中心が  $(0, 0)$  の二次元位置を要素とする矩形ウィンドウ領域である。対応点のみを用いた場合は十分な変形が行われなため、対応点周りに次のような分布をもつウィンドウ  $\mathcal{W}$  を設定し、周辺画素に対してもウィンドウ内の位置  $\mathbf{w} \in \mathcal{W}$  に応じた重み  $g(\mathbf{w})$  を与える。

$$g(\mathbf{w}) = \exp\left(-\frac{\|\mathbf{w}\|^2}{2\sigma^2}\right) \quad (18)$$

ここで、 $\sigma^2$  は分布の分散である。

### 3.4.2 エネルギー関数の最小化

3.4.1 項で定義した写像  $f(\mathbf{x}_n^k)$  を一列に並べたベクトルを  $\mathbf{f}$  とする。

$$\mathbf{f} = (f(\mathbf{x}_0^0), f(\mathbf{x}_1^0), \dots, f(\mathbf{x}_0^1), f(\mathbf{x}_1^1), \dots)^\top \quad (19)$$

このとき、式 (12) に示すエネルギー関数は写像ベクトル  $\mathbf{f}$  を用いて以下のように表すことができる。

$$E = \mathbf{f}^\top \mathbf{Q} \mathbf{f} + \mathbf{r}^\top \mathbf{f} + Z \quad (20)$$

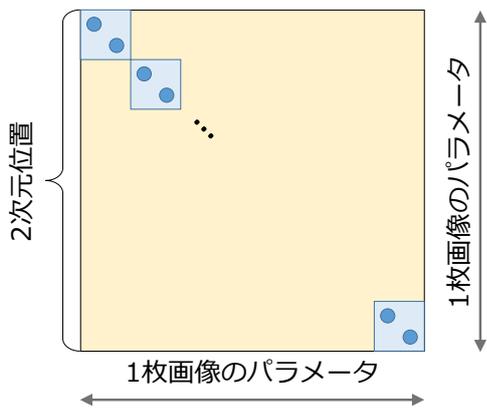
ただし、 $\mathbf{Q}$ ,  $\mathbf{r}$ ,  $Z$  はそれぞれ  $\mathbf{f}$  の二次の項の係数を要素にもつパラメータ数  $\times$  パラメータ数の行列、 $\mathbf{f}$  の一次の項の係数を要素にもつパラメータ数のベクトル、定数である。これを微分すると以下のように表される。

$$\frac{\partial E}{\partial \mathbf{f}} = 2\mathbf{Q}\mathbf{f} + \mathbf{r} \quad (21)$$

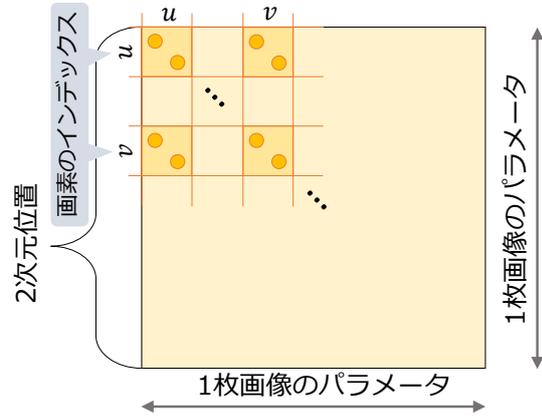
$E$  は下に凸の二次関数であるので、これが 0 となるとき、エネルギーは最小となり、そのときの  $\mathbf{f}$  は以下のように算出される。

$$\mathbf{f} = -\frac{1}{2}\mathbf{Q}^{-1}\mathbf{r} \quad (22)$$

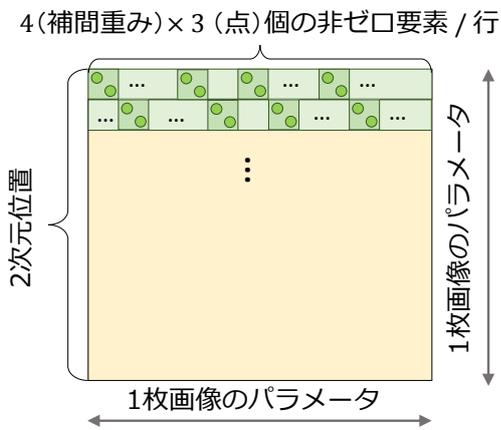
$\mathbf{Q}$  は図 13(d) に示すように疎行列となるため、標準的な sparse linear solver[38] を用いて効率的に解を求めることができる。このように得られた各画素の変形後の位置に基づき、OpenGL を用いた色のメッシュ補間により変形後の参照画像群を生成する。



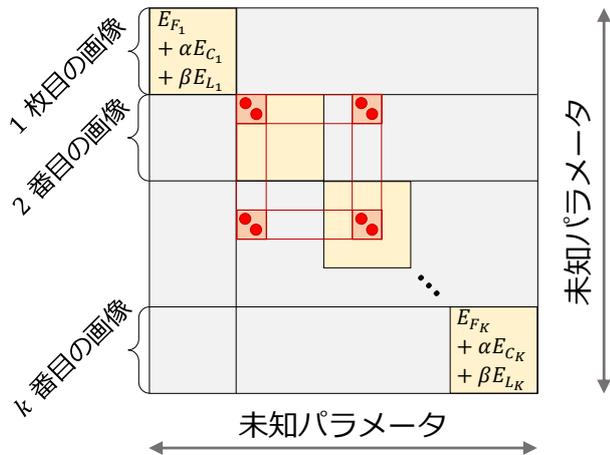
(a) 1枚画像のパラメータ行列における  $E_F$  の配置



(b) 1枚画像のパラメータ行列における  $E_C$  の配置



(c) 1枚画像のパラメータ行列における  $E_L$  の配置



(d) 疎行列  $Q$  における  $E_M(s,t)$  の配置

図 13: 疎行列  $Q$  のブロック構造

### 3.5 ポアソン方程式を用いた色調補正

変形した参照画像群をカメラマップに基づき単純にコピーした場合，図 14 の赤枠内に示すようにテクスチャの境界で色の不連続が生じる．この問題を解決するために，Poisson blending [39] を利用し，参照画像群の合成と同時に色調補正を行う．具体的には，カメラマップ上で最も多くの画素が投影される入力画像の画素を既知領域  $S$  とし，領域  $S$  外でかつ不可視領域を含まないすべての領域  $\Omega$  を未知領域として，その領域に含まれる画素の画素値を未知数のパラメータとみなす．ここで，不可視領域とは，三次元形状がないために仮想視点のデプスマップ上で値を持たず，自由視点画像上でどのカメラからも色が投影されない画素を指す．また，図 15 のように，画素  $p \in \Omega$  の 4 近傍の集合を  $N_p = \{p' \in S \cup \Omega\}$  として， $\Omega$  の外側境界領域を  $\partial\Omega = \{p' \in S \mid N_p \cap \Omega \neq \emptyset\}$  と表す．

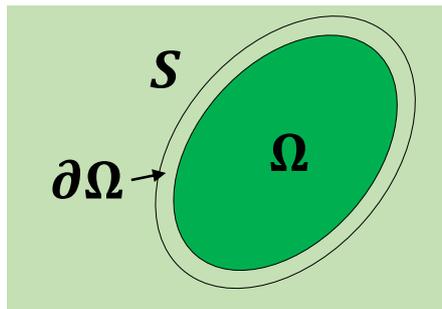
Poisson blending [39] に倣い，すべての最適化する画素  $p \in \Omega$  について，

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p \cap \partial\Omega} f_q^* + \sum_{q \in N_p} v_{pq} \quad (23)$$

を満たすようにエネルギー関数を解き，画素値を決定することで色調についてシームレスな画像を生成する (図 16)．ここで， $f_p$  は画素  $p$  の未知の画素値， $f_q^*$  は画素  $q$  の既知の画素値， $v_{pq}$  は画素  $p$  と画素  $q$  の勾配を示す．式 (23) は，提案手法で定義したエネルギー関数と同様に  $f_p$  の二次形式で表される問題であるため，標準的な sparse linear solver を用いて一意かつ効率的に解を求めることができる．



図 14: 自由視点画像中に生じる色の不連続



$f^*$ : 既知関数  
  $f$ : 未知関数

図 15: 未知領域と既知領域についての表記



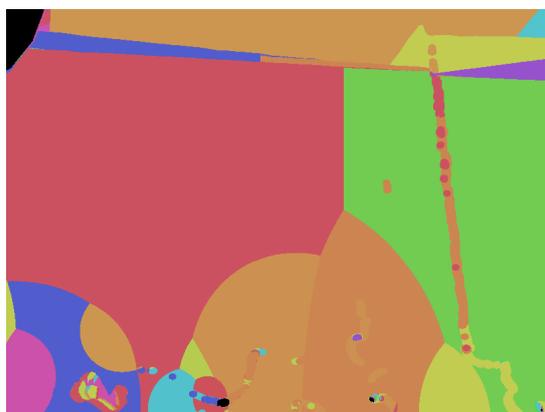
(a) 色調補正前



(b) 色調補正後



(c) マスク画像 (赤: 未知領域, 黒: 既知領域)



(d) カメラマップ

図 16: ポアソン方程式を用いた色調補正結果

## 4. 実験

本章では，屋外環境で撮影された画像群を入力とした自由視点画像生成実験について述べる．本実験では，提案手法と比較手法により自由視点画像を生成し，それらを比較することで提案手法による自由視点画像中に生じる歪みの抑制効果を検証することを目的とする．なお，評価では特に生成画像上の線分の歪みが目立たなくなっているか，複数画像から視点依存テクスチャマッピングされるテクスチャ間の不連続が目立たなくなっているかに注目する．以下ではまず，実験の目的と条件について述べ，次に，提案手法および比較手法を定性的に比較することで提案手法の有効性を示し，最後に考察する．

### 4.1 実験条件

本実験では，図 18 に示す屋外で撮影された 4 種類の画像データセット (i)～(iv) を入力として用いた．表 1 にデータセットの詳細を示す．また，カメラ位置姿勢推定・三次元復元には VisualSFM [12, 13] および CMPMVS [20] を用いた．各データセットに対する三次元復元およびカメラの位置姿勢推定結果を図 19, 20 に示す．生成する自由視点画像における仮想視点の位置姿勢は各データセットで図 20 のように設定し，解像度は入力画像と同一とした．各データセットに対して用いたエネルギー関数のパラメータを表 2 に示す．なお，パラメータの値は実験の際に概ね良好な結果が得られた  $\alpha = 100$ ,  $\beta = 3000$ ,  $\gamma = 5$  を基準として，変形が大きすぎる場合には  $\alpha$  の値を大きくするなど経験的に決定した．また，比較手法としては，図 17 に示すように，視点依存テクスチャマッピング [1] と Poisson blending[39] のみを行い，画像変形を行わない方法を採用した．

表 1: データセットの詳細

データセット	入力画像枚数	入力画像解像度
(i) ビル	63 枚	816 × 612 pix.
(ii) 神社	69 枚	816 × 612 pix.
(iii) 歩道	65 枚	612 × 816 pix.
(iv) モニュメント	133 枚	960 × 540 pix.

表 2: 各データセットに用いたパラメータ

データセット	仮想視点	$\alpha$	$\beta$	$\gamma$
(i) ビル	A	100	3000	5
	B	100	3000	5
	C	100	3000	3
(ii) 神社	A	100	3000	5
	B	200	3000	3
(iii) 歩道	A	100	3000	5
	B	50	2000	5
(iv) モニュメント	A	300	3000	5
	B	100	3000	5

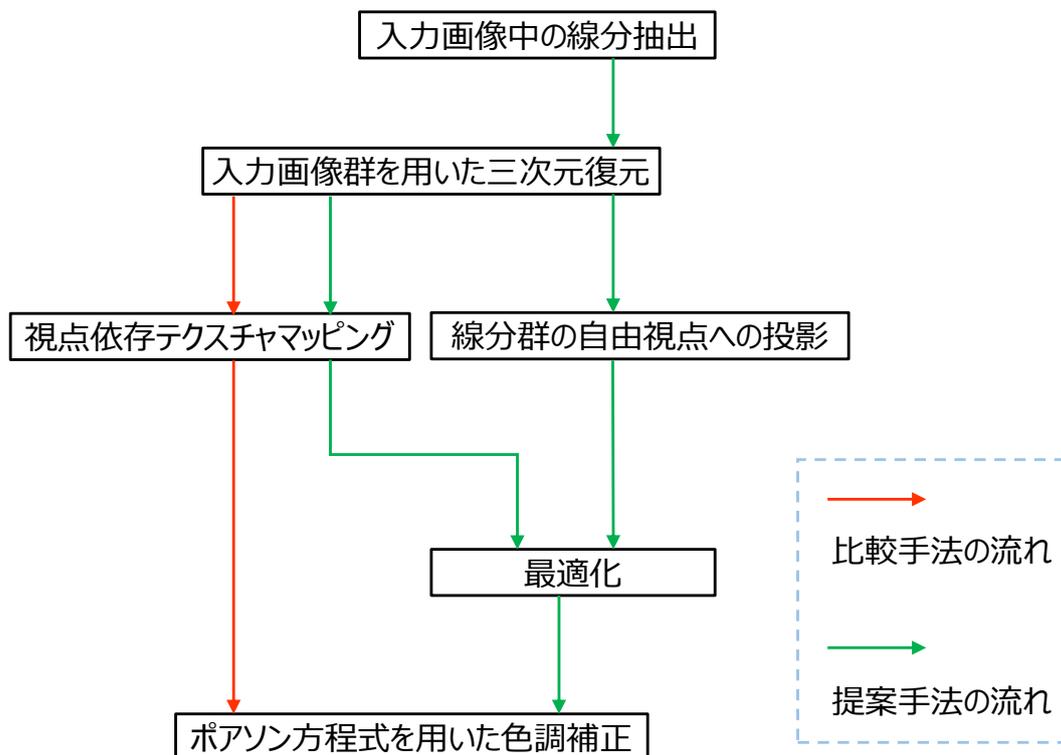


図 17: 提案手法と比較手法の処理の流れ



(i) ビルの入力画像例



(ii) 神社の入力画像例

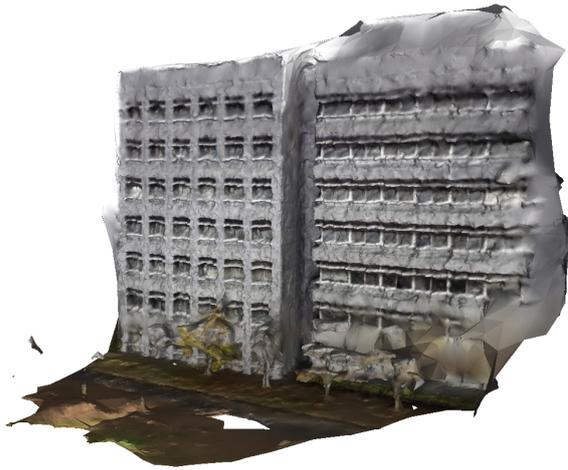


(iii) 歩道の入力画像例



(iv) モニュメントの入力画像例

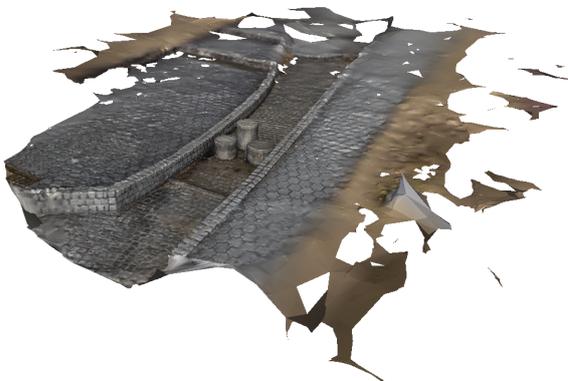
図 18: 各データセットの入力画像例



(i) ビル



(ii) 神社

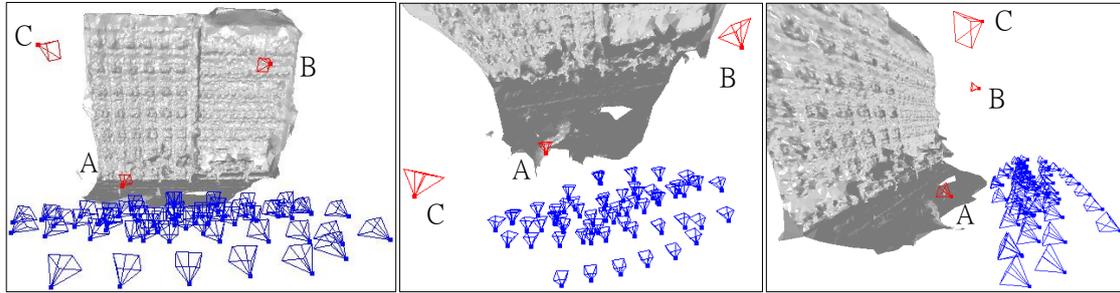


(iii) 歩道

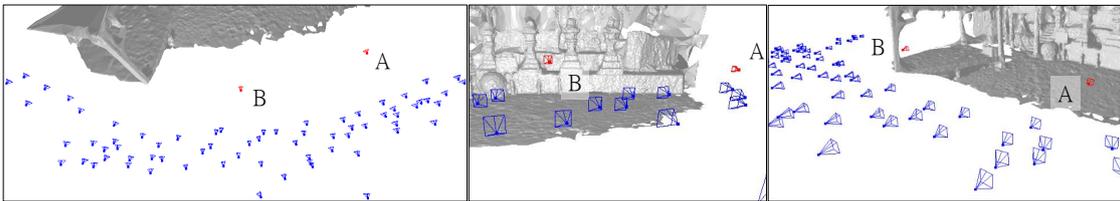


(iv) モニュメント

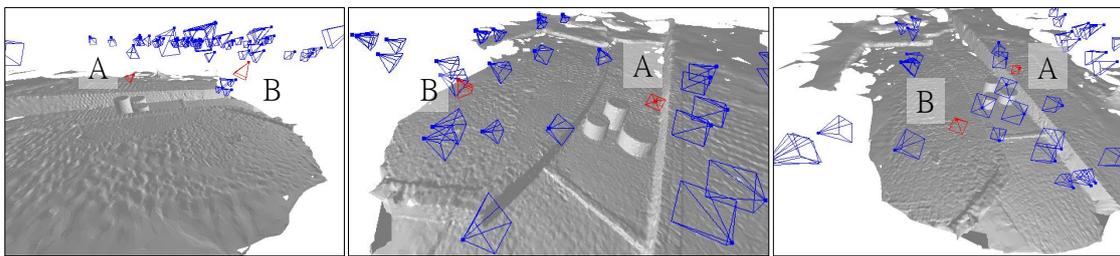
図 19: 復元した三次元モデル



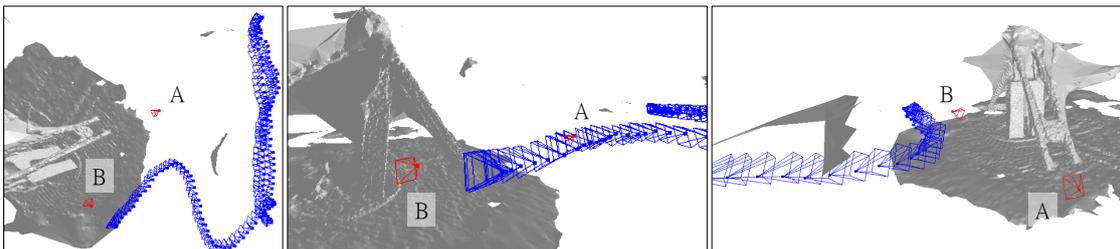
(i) ビルにおけるカメラ位置および三次元形状の位置関係



(ii) 神社におけるカメラ位置および三次元形状の位置関係



(iii) 歩道におけるカメラ位置および三次元形状の位置関係



(iv) モニュメントにおけるカメラ位置および三次元形状の位置関係

図 20: 各データセットで生成する自由視点位置 (赤) および実視点位置 (青)

## 4.2 主観評価実験

前節で示した画像データセットおよび設定した仮想視点位置を用いて、比較手法および提案手法により自由視点画像を生成した。ビルのシーン (i) に対する結果を図 21 から図 26 に、神社のシーン (ii) に対する結果を図 27 から図 30 に、歩道のシーン (iii) に対する結果を図 31 から図 34 に、モニュメントのシーン (iv) に対する結果を図 35 から図 39 に示す。また、視点依存テクスチャマッピングにおいて、投影に用いられた入力画像の枚数について表 3 にまとめる。なお、自由視点画像およびカメラマップ中の黒い領域は、三次元形状がないために奥行きが得られなかったあるいはデプステストにより対応する画像が入力画像群に存在しなかったために、テクスチャが投影されなかった領域を示す。また、投影線分の可視化において、各線分を見分けやすくするため 2 色を用いている。以下では、各データセット・視点に対する実験結果を示し、比較を行う。

データセット (i) に対する仮想視点 A での結果を図 21 および図 22 に、仮想視点 B での結果を図 23 および図 24 に、仮想視点 C での結果を図 25 および図 26 に示す。このデータセットでは、いずれの仮想視点も入力画像の視点位置から離れており、ビルに近い視点となっている。このため、比較手法によるマッピングでは、三次元形状の誤差の影響で窓や外壁の境界などの直線が大きく波打っていることが確認できる。これに対して、提案手法では歪みが抑制され直線性が保持されていることが確認できる。特に、図 25(a) および図 25(b) を比較したとき、画像中右にあるビルの直線構造においてこのことが顕著に現れている。図 22(a) および図 22(b) を図 22(c) のカメラマップと照らし合わせて比較することにより、従来の視点依存テクスチャマッピングにおいて生じたテクスチャが切り替わる境界におけるテクスチャの不連続についても、提案手法では抑制されていることがわかる。

データセット (ii) に対する仮想視点 A での結果を図 27 および図 28 に、仮想視点 B での結果を図 29 および図 30 に示す。図 28(a) および図 28(b) の比較において、比較手法では看板や黒い柵に生じている歪みが提案手法では低減されている。図 30(a) および図 30(b) の比較において、階段に生じているテクスチャの不連続は解消されているが、画像中の左上にある手すりの形状が提案手法では歪んでし

まっていることが確認できる。

データセット (iii) に対する仮想視点 A での結果を図 31 および図 32 に，仮想視点 B での結果を図 33 および図 34 に示す。図 32(a) および図 32(b)，図 34(a) および図 34(b) の比較において，比較手法で生じている地面のテクスチャの不連続が提案手法では解消されていることが確認できる。

データセット (iv) に対する仮想視点 A での結果を図 35，図 36，図 37 に，仮想視点 B での結果を図 38 および図 39 に示す。図 36(a) および図 36(b) の赤枠内において，比較手法では朱色の柱の輪郭付近にボケが生じているが，提案手法ではそれが改善し，柱の輪郭が鮮鋭化されている。一方，図 37(a) および図 37(b) の赤枠内の比較において，提案手法では地面のテクスチャに不連続が生じ，図 39(a) および図 39(b) の赤枠内の比較において，提案手法ではレンガの柱の輪郭に大きな屈曲が生じている。

表 3: 投影に用いられた入力画像の枚数

データセット	仮想視点	画像枚数
(i) ビル	A	34 枚
	B	11 枚
	C	25 枚
(ii) 神社	A	28 枚
	B	43 枚
(iii) 歩道	A	41 枚
	B	5 枚
(iv) モニュメント	A	72 枚
	B	9 枚



(a) 比較手法



(b) 提案手法

図 21: データセット (i) の仮想視点 A における自由視点画像



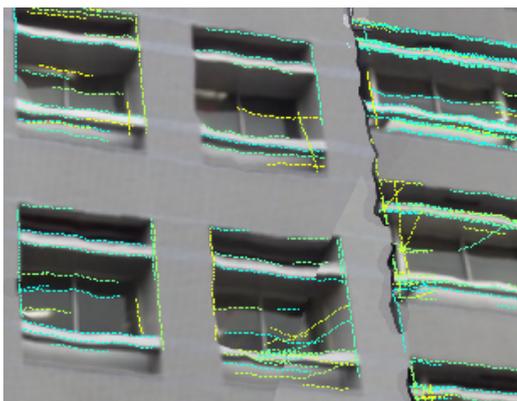
(a) 図 21(a) の一部拡大画像



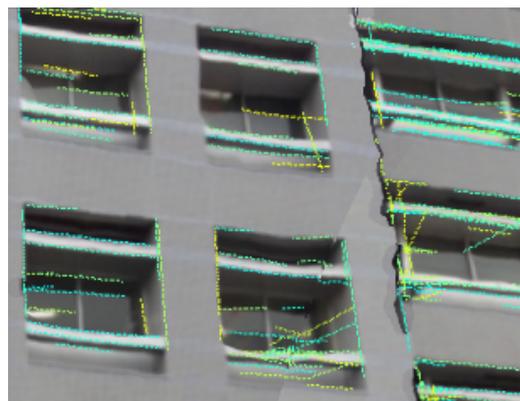
(b) 図 21(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ

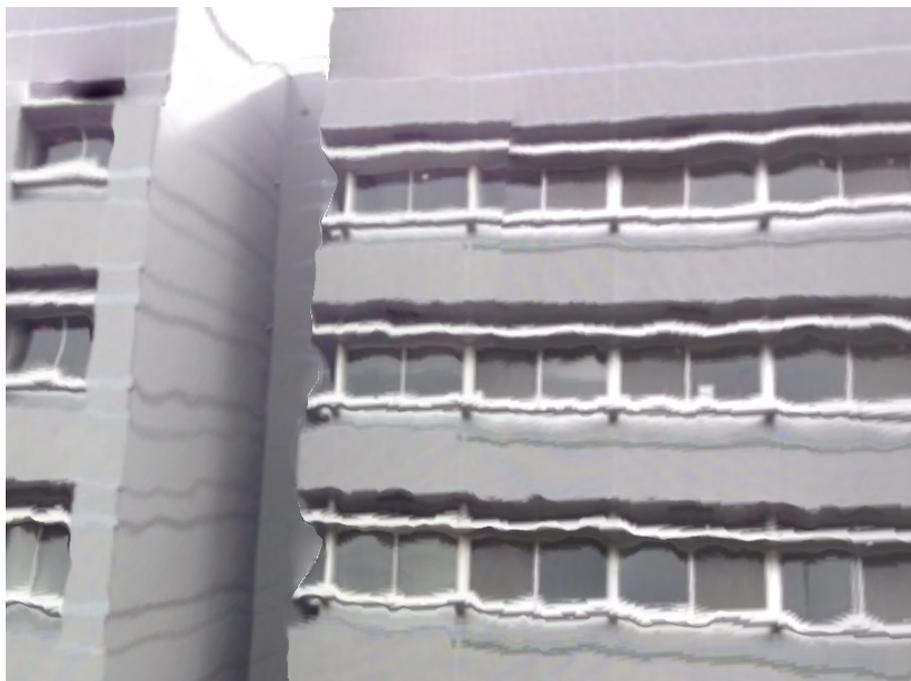


(d) (a) の領域における投影線分



(e) (b) の領域における投影線分

図 22: データセット (i) の仮想視点 A における自由視点画像の一部拡大画像

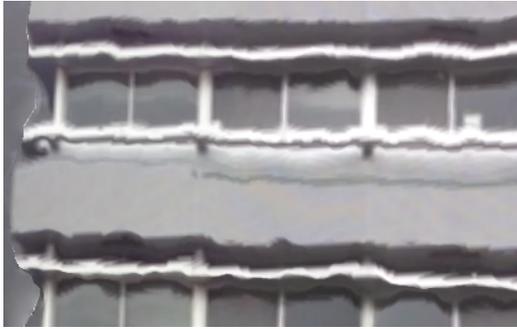


(a) 比較手法



(b) 提案手法

図 23: データセット (i) の仮想視点 B における自由視点画像



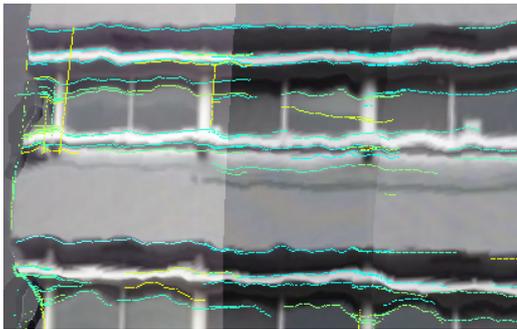
(a) 図 23(a) の一部拡大画像



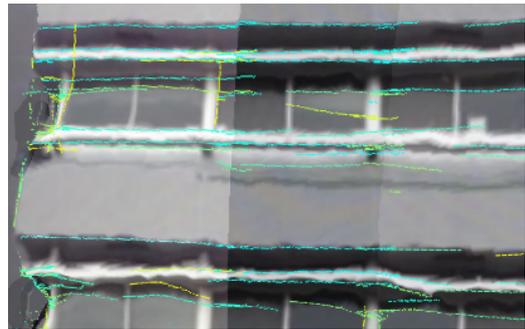
(b) 図 23(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ



(d) (a) の領域における投影線分

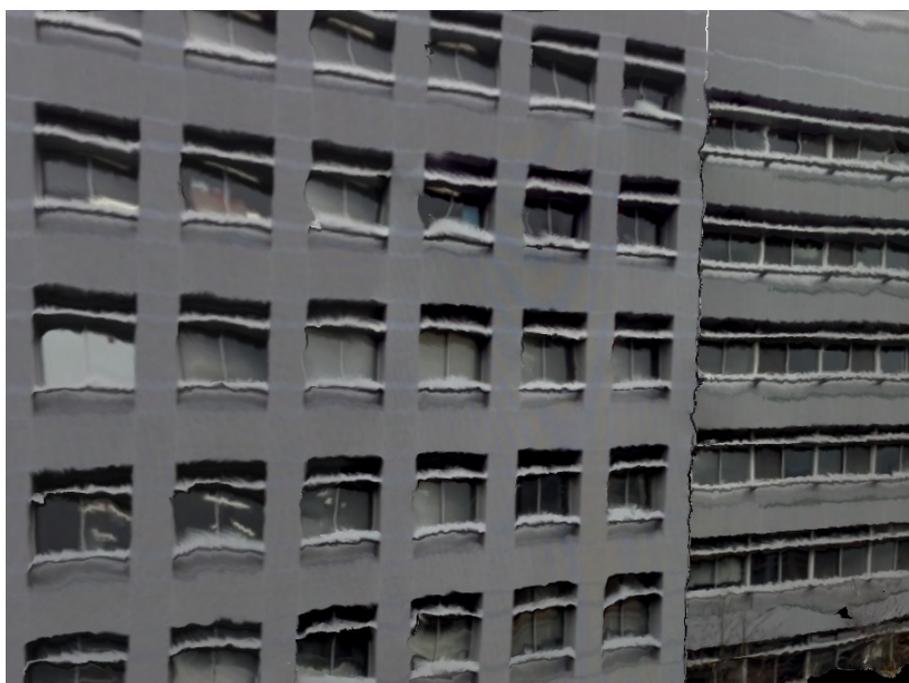


(e) (b) の領域における投影線分

図 24: データセット (i) の仮想視点 B における自由視点画像

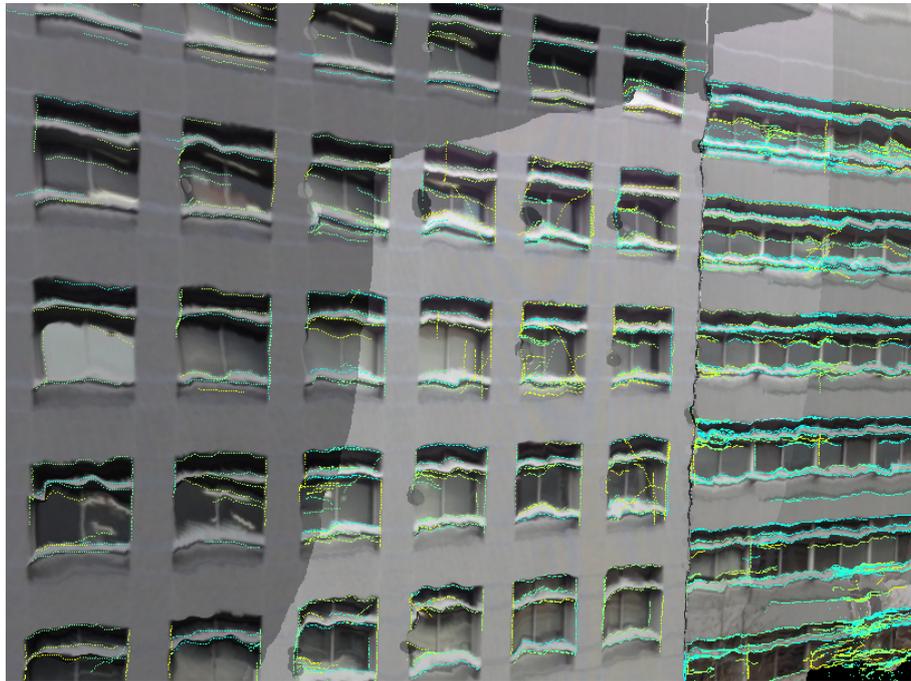


(a) 比較手法

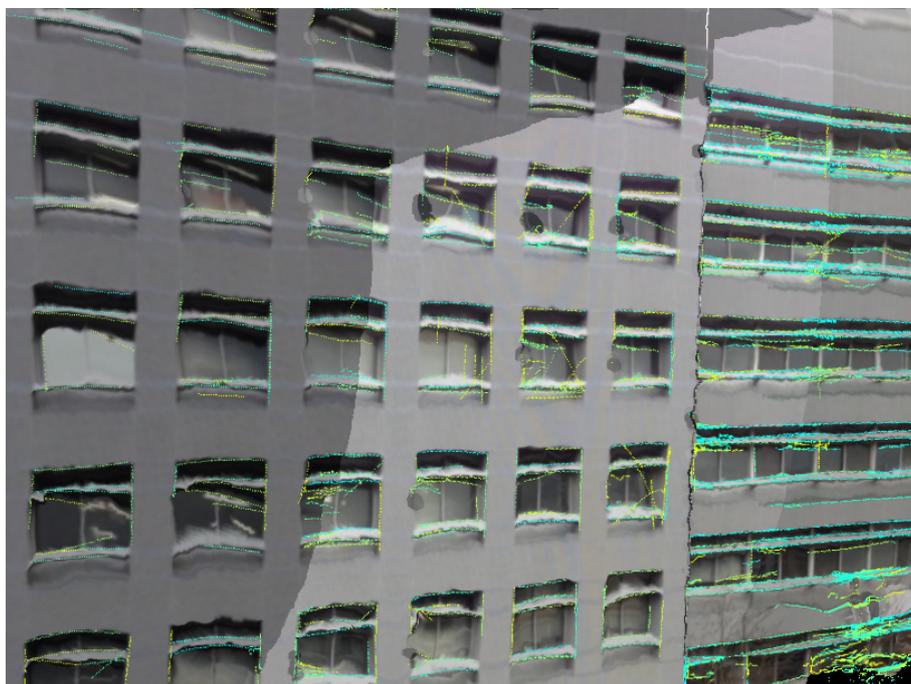


(b) 提案手法

図 25: データセット (i) の仮想視点 C における自由視点画像



(a) 比較手法での投影線分



(b) 提案手法での投影線分

図 26: データセット (i) の仮想視点 C における線分情報の可視化結果



(a) 比較手法



(b) 提案手法

図 27: データセット (ii) の仮想視点 A における自由視点画像



(a) 図 27(a) の一部拡大画像



(b) 図 27(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ



(d) (a) の領域における投影線分



(e) (b) の領域における投影線分

図 28: データセット (ii) の仮想視点 A における自由視点画像の一部拡大画像



(a) 比較手法



(b) 提案手法

図 29: データセット (ii) の仮想視点 B における自由視点画像



(a) 図 29(a) の一部拡大画像



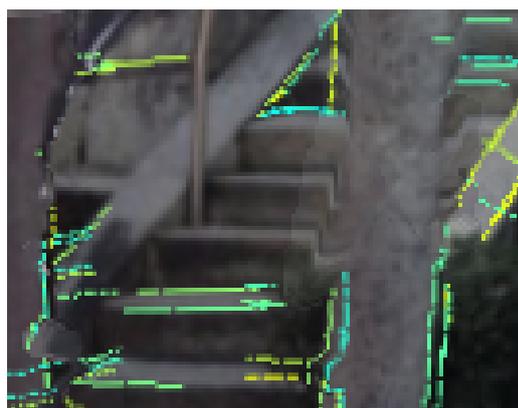
(b) 図 29(a) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ



(d) (a) の領域における投影線分



(e) (b) の領域における投影線分

図 30: データセット (ii) の仮想視点 B における自由視点画像の一部拡大画像



(a) 比較手法



(b) 提案手法

図 31: データセット (iii) の仮想視点 A における自由視点画像



(a) 図 31(a) の一部拡大画像



(b) 図 31(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ



(d) (a) の領域における投影線分

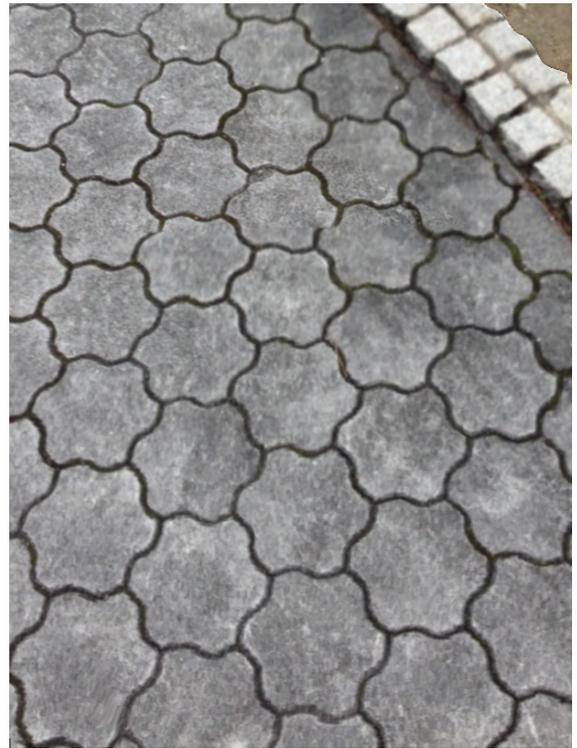


(e) (b) の領域における投影線分

図 32: データセット (iii) の仮想視点 A における自由視点画像の一部拡大画像

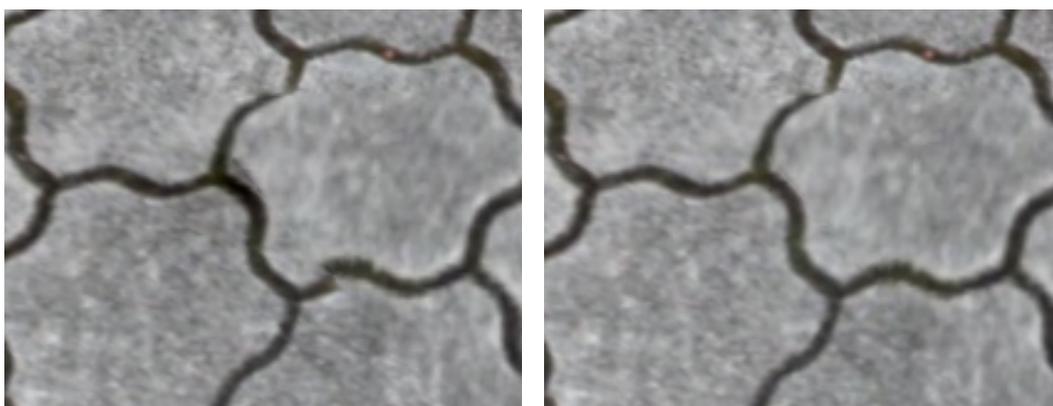


(a) 比較手法



(b) 提案手法

図 33: データセット (iii) の仮想視点 B における自由視点画像



(a) 図 33(a) の一部拡大画像

(b) 図 33(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ

図 34: データセット (iii) の仮想視点 B における自由視点画像の一部拡大画像



(a) 比較手法



(b) 提案手法

図 35: データセット (iv) の仮想視点 A における自由視点画像



(a) 図 35(a) の一部拡大画像

(b) 図 35(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ

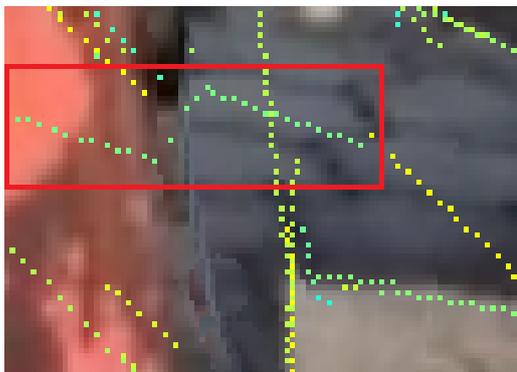
図 36: データセット (iv) の仮想視点 A における自由視点画像の一部拡大画像 1



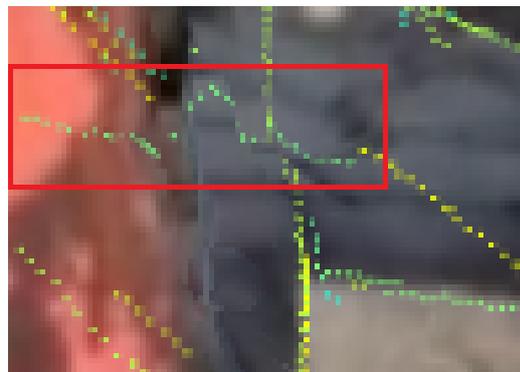
(a) 図 35(a) の一部拡大画像



(b) 図 35(b) の一部拡大画像



(c) (a) の領域における投影線分



(d) (b) の領域における投影線分

図 37: データセット (iv) の仮想視点 A における自由視点画像の一部拡大画像 2



(a) 比較手法



(b) 提案手法

図 38: データセット (iv) の仮想視点 B における自由視点画像



(a) 図 38(a) の一部拡大画像



(b) 図 38(b) の一部拡大画像



(c) 一部拡大画像におけるカメラマップ



(d) (a) の領域における投影線分



(e) (b) の領域における投影線分

図 39: データセット (iv) の仮想視点 B における自由視点画像の一部拡大画像

### 4.3 考察

前節で示した実験結果において，提案手法では比較手法と比較して仮想視点の位置を入力画像の視点位置から大きく離しても，テクスチャの歪みの少ない自由視点画像生成が可能であることを確認した．特に，直線性の保持効果については図 25 に，テクスチャの不連続の抑制効果については図 24，図 30，図 34 に顕著に現れている．ただし，本来直線的な構造であるはずのテクスチャが歪むという問題も見られた．その原因として考えられることを以下に列挙する．

1. 線分要素がテクスチャの連続性を考慮する項  $E_{M(s,t)}$  に引っ張られる．
2. 線分が近い領域で複数検出され，直線性を保持する項  $E_{L_k}$  により一方に引っ張られる．
3. 奥行き値の大きく異なる領域が隣接した境界で直線が検出され，変形の結果奥行き値が大きい方の領域が大きく移動する．

1 は，図 30(d) の画像中右下の鳥居の柱にある線分が，柱の輪郭の不連続箇所を補正する効果に引っ張られ，図 30(e) のように一部分だけ大きく変形していることから推測できる．同様に，図 32(d) および図 32(e) の赤枠内の比較において，境界領域の不連続箇所の補正のために，線分が屈曲する現象も確認している．これは，不連続の補正と直線性の保持が同時に成り立たない場合もあることを示す．2 については，図 39(d) の赤枠内においてレンガ上に存在する線分が，図 39(e) の赤枠内では他の線分の影響により潰れてテクスチャの歪みとなっていることが確認できる．また，この歪みが  $E_{M(s,t)}$  の影響によるものでないことも図 39(c) に示すカメラマップからわかる．このように複数の線分が密集している領域では，それぞれの直線性を保持しようとする変形が互いに干渉し，線分の直線性が保持されない場合がある．これを防ぐためには，線分検出の際に，視認できるかつ保持したい線分を手動で選択するなどの方法が挙げられるが，多くの手間や時間を要する．3 については，図 37(c) を見たときに，赤枠で囲ったカギ型に屈曲した線分要素は，変形後に図 37(d) に示すように大きく歪んでいる．その結果，図 37(b) に見られるように，奥の地面のテクスチャが変形しているのが確認できる．奥行

き値の差が大きいほどこのような変形は大きくなり歪みとなって現れる。これを解決するためには、奥行き値を基に変形量を決めるような枠組みが必要であると考えられるが、そもそもの三次元形状に誤差が多く含まれるために困難である。

また、今回歩道やモニュメントのように本来直線的でないテクスチャ構造のタイルやレンガまでもが直線的になる現象を確認した。このようなシーンにおいては直線性を保持する項  $E_{L_k}$  の重み  $\beta$  を小さくするなど、対象シーンによってエネルギー関数の重みを調整する必要があることがわかった。さらに、ビルのシーンにおいて、本来水平方向に並列した窓の水平方向の直線はすべて平行となるべきであるが、図 25 に見られるように、線分の直線構造自体は保持されるものの、その方向は補正されず、角度が不揃いになっている。ビルのような平面で構成される人工物に対しては、平面仮定を用いることで、これらの現象に対処できると考えられる。

提案手法全体の処理における計算時間は、表 4 に示す計算機環境で 3 分程度から 30 分程度までばらつきがあり、その多くは最適化処理および色調補正処理に要している。最適化処理について、計算時間は投影に利用される画像枚数および画像解像度に大きく依存し、色調補正処理については、未知領域の大きさに依存する。高速化を図るには、カメラの光学中心間の距離やカメラの光軸同士の角度の差を評価尺度として投影に用いる入力画像の枚数を制限するなどが挙げられる。

表 4: 計算機の構成

CPU	Intel Core i7-4770 3.40GHz
メモリ	16.0 GB
OS	Windows 7 Professional 64bit

## 5. 結言

本論文では、自由視点画像生成において、仮想視点位置を入力画像の取得位置から大きく離れた場合に生じる歪みを抑止することを目的として、視点依存テクスチャマッピングと画像変形手法の組み合わせに基づく新たな自由視点画像生成手法を提案した。提案手法では、撮影された画像群を入力として、まず三次元形状復元を行い、三次元形状を参照して入力画像群およびそれらから抽出した線分を仮想視点に投影し、参照画像群および投影線分群を生成すると同時に視点依存テクスチャマッピングの基準によりカメラマップを生成する。投影された参照画像群に対してカメラマップの境界領域における対応点探索を行った上で、線分の直線性およびテクスチャの切り替わり位置における連続性を考慮した画像群のワーピングを行い、変形した画像群を色調補正・合成することでテクスチャの歪みを抑制した自由視点画像を生成する。

実験では、屋外環境を対象とした自由視点画像生成を行い、提案手法を用いることで線分の直線性およびテクスチャの連続性が保持され、従来手法と比較して歪みの少ない画像生成が可能であることを示した。一方、エネルギー関数の項同士の干渉や奥行き値の差が大きい領域における線分の変形などの要因により、提案手法では歪みが生じる場合があることを確認した。これらの歪みはエネルギー関数の重みの決め方である程度抑制することが可能であるため、対象となるシーンに合わせた重みの調整が重要である。

今後、対象シーンに対して安定に良好な結果を得るためには、デプスマップを参照し、奥行き値の差の大きい領域において奥行き値の大きい方の領域の変形量を小さくする、あるいは、VDTMにより生成した自由視点画像についての顕著性マップを生成し、顕著性の高い領域では変形量を小さくするなど、エネルギー関数の重みパラメータを画像内の領域毎に適応的に決定する枠組みが必要である。また、処理の高速化を行うためには、カメラの位置姿勢や画角を考慮するなどして、投影に用いる入力画像の枚数を制限する手法が求められる。

## 謝辞

本研究の全過程を通して、懇切なる御指導、御鞭撻を賜りました視覚情報メディア研究室 横矢直和教授に心より感謝致します。本研究の遂行にあたり、有益な御助言を頂きました光メディアインタフェース研究室 向川康博教授に厚く御礼申し上げます。本研究を進めるにあたり、終始温かい御指導をして頂きました視覚情報メディア研究室 佐藤智和准教授に深く感謝申し上げます。そして、本研究を行うにあたり、多大なる御助言、ご鞭撻を賜りました視覚情報メディア研究室 河合紀彦助教、中島悠太助教に心より感謝致します。特に、河合紀彦助教には、本研究のテーマ設定から本論文の執筆に至るまで細やかな御指導を頂きました。また、研究室での生活を支えて頂きました視覚情報メディア研究室 石谷由美女史に心より感謝いたします。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りました視覚情報メディア研究室の皆様にも深く感謝致します。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] P. E. Debevec, Y. Yu, and G. Borshukov. Efficient view-dependent image-based rendering with projective texture-mapping. In *Proc. 9th Eurographics Workshop on Rendering (EGWR 1998)*, pp. 105–116, 1998.
- [2] 伊吹拓也, 佐藤智和, 松下康之, 横矢直和. 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. 電子情報通信学会 技術研究報告, Vol. 109, No. 470, pp. 437–442, 2010.
- [3] G. Chaurasia, S. Duchêne, O. Sorkine-Hornung, and G. Drettakis. Depth synthesis and local warps for plausible image-based navigation. *ACM Trans. Graphics*, Vol. 32, No. 3, pp. 30:1–30:12, 2013.
- [4] Y.-H. Huang, T.-K. Huang, Y.-H. Huang, W.-C. Chen, and Y.-Y. Chuang. Warping-based novel view synthesis from a binocular image for autostereoscopic displays. In *Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME 2012)*, pp. 302–307, 2012.
- [5] 塚本壮輔, 呉本亮, 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計測. 電子情報通信学会論文誌. D-II, Vol. J83-D2, No. 9, pp. 1962–1965, 2000.
- [6] H. Kawasaki, Y. Ohsawa, R. Furukawa, and Y. Nakamura. Dense 3D reconstruction with an uncalibrated active stereo system. In *Proc. 7th Asian Conf. on Computer Vision (ACCV 2006)*, pp. 882–891, 2006.
- [7] S. Kawahito, I. Halin, T. Ushinaga, T. Sawada, M. Homma, and Y. Maeda. A CMOS time-of-flight range image sensor with gates-on-field-oxide structure. *IEEE Sensors Journal*, Vol. 7, No. 12, pp. 1578–1586, 2007.
- [8] 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. E83-D, No. 7, pp. 1429–1440, 2000.

- [9] C. Früh and A. Zakhor. Constructing 3D city models by merging aerial and ground views. *IEEE Computer Graphics and Applications*, Vol. 23, No. 6, pp. 52–61, 2003.
- [10] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. In *Proc. ACM SIGGRAPH 2006*, pp. 835–846, 2006.
- [11] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *Int'l Journal of Computer Vision*, Vol. 80, No. 2, pp. 189–210, 2008.
- [12] C. Wu. Towards linear-time incremental structure from motion. In *Proc. Int'l Conf. on 3D Vision (3DV 2013)*, pp. 127–134, 2013.
- [13] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz. Multicore bundle adjustment. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2011)*, pp. 3057–3064, 2011.
- [14] H. Baker. Three-dimensional modelling. In *Proc. Int'l Joint Conf. on Artificial Intelligence*, pp. 649–655, 1977.
- [15] W. N. Martin and J. K. Aggarwal. Volumetric descriptions of objects from multiple views. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 5, No. 2, pp. 150–158, 1983.
- [16] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 2, pp. 150–162, 1994.
- [17] M. Goesele, B. Curless, and S. M. Seitz. Multi-view stereo revisited. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2006)*, pp. 2402–2409, 2006.

- [18] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz. Multi-view stereo for community photo collections. In *Proc. 11th Int'l Conf. on Computer Vision (ICCV 2007)*, 8 pages, 2007.
- [19] Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski. Towards internet-scale multi-view stereo. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2010)*, pp. 1434–1441, 2010.
- [20] M. Jancosek and T. Pajdla. Multi-view reconstruction preserving weakly-supported surfaces. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2011)*, pp. 3121–3128, 2011.
- [21] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2006)*, pp. 519–526, 2006.
- [22] M. Waechter, N. Moehrle, and M. Goesele. Let there be color! — Large-scale texturing of 3D reconstructions. In *Proc. European Conf. on Computer Vision (ECCV 2014)*, pp. 836–850, 2014.
- [23] 小甲啓隆, 糸秀行, 佐藤智和, 町田貴史, 佐野範佳, 横矢直和. 車載カメラ画像処理アルゴリズム評価のための自由視点画像生成. 情報科学技術フォーラム講演論文集, Vol. 12, No. 3, pp. 141–142, 2013.
- [24] P. Rademacher. View-dependent geometry. In *Proc. ACM SIGGRAPH 1999*, pp. 439–446, 1999.
- [25] Y. Nakashima, T. Sato, Y. Uno, N. Yokoya, and N. Kawai. Augmented reality image generation with virtualized real objects using view-dependent texture and geometry. In *Proc. IEEE Int'l Symp. on Mixed and Augmented Reality (ISMAR 2013) Works-in-Progress*, 6 pages, 2013.

- [26] J.-F. Evers-Senne and R. Koch. Image based interactive rendering with view dependent geometry. *Computer Graphics Forum*, Vol. 22, No. 3, pp. 573–582, 2003.
- [27] C. Rother, V. Kolmogorov, and A. Blake. “GrabCut”: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graphics*, Vol. 23, No. 3, pp. 309–314, 2004.
- [28] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 11, pp. 2274–2282, 2012.
- [29] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, Vol. 1, No. 2, pp. 3–20, 1991.
- [30] T. Naemura, T. Takano, M. Kaneko, and H. Harashima. Ray-based creation of photo-realistic virtual world. In *Proc. Int’l Conf. on Virtual Systems and MultiMedia (VSMM 1997)*, pp. 59–68, 1997.
- [31] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. ACM SIGGRAPH 1996*, pp. 31–42, 1996.
- [32] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proc. ACM SIGGRAPH 1996*, pp. 43–54, 1996.
- [33] S. E. Chen and L. Williams. View interpolation for image synthesis. In *Proc. ACM SIGGRAPH 1993*, pp. 279–288, 1993.
- [34] S. M. Seitz and C. R. Dyer. View morphing. In *Proc. ACM SIGGRAPH 1996*, pp. 21–30, 1996.

- [35] H. Wang, M. Sun, and R. Yang. Space-time light field rendering. *IEEE Trans. on Visualization and Computer Graphics*, Vol. 13, No. 4, pp. 697–710, 2007.
- [36] J. Matas, C. Galambos, and J. Kittler. Robust detection of lines using the progressive probabilistic hough transform. *Computer Vision and Image Understanding*, Vol. 78, No. 1, pp. 119–137, 2000.
- [37] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. 4th Alvey Vision Conf.*, pp. 147–151, 1988.
- [38] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2nd edition, 2003.
- [39] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. In *Proc. ACM SIGGRAPH 2003*, pp. 313–318, 2003.