

NAIST-IS-MT1351030

修士論文

視点依存テクスチャマッピングにおけるテクスチャの 連続性を考慮した自由視点画像生成

片桐 敬太

2015年3月12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。

片桐 敬太

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)
中島 悠太 助教	(副指導教員)
河合 紀彦 助教	(副指導教員)

視点依存テクスチャマッピングにおけるテクスチャの 連続性を考慮した自由視点画像生成*

片桐 敬太

内容梗概

シーンを複数の地点から撮影した画像群を入力として，任意の視点 (仮想視点) からの見えを再現する自由視点画像生成技術は，遠隔地のシーンを仮想的に体感できるテレプレゼンスなどのアプリケーションでの活用が期待されている．従来の自由視点画像生成手法として，入力画像群から Multi-view Stereo (MVS) などを用いて復元されたシーンの3次元形状を利用する手法がある．しかし，この手法では高精度な3次元形状が得られない場合，生成された自由視点画像上に歪みや欠損が生じるという問題がある．この問題を低減するために，3次元形状に対応するテクスチャとして仮想視点に最も近い入力視点の画像を適応的に貼り付けることで3次元復元精度が不十分な場合においても詳細な形状を視覚的に再現する視点依存テクスチャマッピングが用いられる．この手法では，入力画像間の色の差異や3次元形状とカメラパラメータの誤差に起因して入力画像の画素値にばらつきが生じる場合，異なるテクスチャの境界においてテクスチャの不連続が生じるため，境界領域で滑らかにテクスチャが変化するように入力画像をブレンディングすることで不連続を抑止している．しかし，3次元形状やカメラパラメータの誤差が大きい場合には，複数の入力画像のテクスチャの投影位置に位置ずれが生じるため，これらがブレンディングされることで物体表面のテクスチャがぼけるなどの問題がある．本研究では，視点依存テクスチャマッピングにおけるテクスチャ選択処理において，テクスチャの連続性を考慮することで，ブレンディン

*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-MT1351030, 2015年3月12日.

グすることなくテクスチャの不連続を抑止する手法を提案する。提案手法ではテクスチャ選択の基準として、従来から利用されてきた仮想視点と入力視点の見えの近さ、仮想視点から見た対象が各入力視点から見えているかを判定する可視判定の結果に加えて、テクスチャの連続性を考慮するために、生成される自由視点画像内の隣接する画素間の画素値の差分を、また材質による光の当たり方は形状の変化が大きい領域で切り替わり易いと考え、MVSを用いて復元された対象の3次元形状の変化の大きさを利用する。画素値の差分と3次元形状の変化は隣接する画素に対して定義されることから、これらに関するエネルギー関数を設計し、この関数をグラフカットを用いて最小化することにより各画素でテクスチャとして利用する入力画像を選択する。最後に、各画素に対して局所的な輝度調整を行い、異なるテクスチャの境界付近に生じる輝度の不連続性を低減する。実験では、様々なパラメータで設定したエネルギー関数において、提案手法で生成した自由視点画像の結果を比較し、各基準の有効性を検証した上で輝度調整の効果を確認する。

キーワード

自由視点画像生成, 視点依存テクスチャマッピング, テクスチャの連続性, 可視判定, グラフカット

Novel View Synthesis with View-dependent Texture Mapping Considering Continuity in Texture Boundaries*

Keita Katagiri

Abstract

Novel view synthesis (NVS) is a technique for synthesizing an image from an arbitrary viewpoint given by a user using multiple input images of a scene. This technique can be used in such applications as telepresence that makes users virtually experience a remote scene. A conventional NVS method uses a 3D geometry of a scene reconstructed from input images using a multi-view stereo (MVS) technique. However, this technique often suffers from a problem that textures in a synthesized image produce visual artifacts if accurate 3D geometry is not obtained. One approach to alleviate this problem is to use view-dependent texture mapping (VDTM). It reproduces the visual details on the 3D geometry, which are lost during the 3D reconstruction process, by applying the image captured from the viewpoint closest to the given viewpoint as a texture. In this approach, discontinuities in color may occur at boundaries of different textures. It is possible to suppress such discontinuities by blending the textures around their boundaries so as to make the textures change smoothly. However, when the 3D geometry is erroneous, inconsistent textures from multiple input images may be blended, resulting in blurring. In this study, we propose a new method based on view-dependent texture mapping that suppresses discontinuities in texture without

*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1351030, March 12, 2015.

blending, considering the continuity in the synthesized image in texture selection. We use four criteria in texture selection. The first is the proximity of the positions of given viewpoint and input viewpoint, which is used in VDTM. The second is the result of visibility test that verifies whether a certain point on the target visible from the given viewpoint is also visible from input viewpoint. The third is the color difference between adjacent pixels in the synthesized image. The fourth is the edges of a 3D geometry reconstructed using MVS, considering that edges in 3D geometry coincide texture boundaries. The last two criteria encode the continuity of color around texture boundaries. We design an energy function based on the four criteria described above, minimizing it with the graph cut algorithm, to determine the input image that is used as a texture for each pixel. Finally, the proposed method locally adjusts the brightness of the pixels to prevent discontinuities around texture boundaries. In the experiment, in order to verify the image quality improvement by the proposed method, we evaluate the method by comparing images synthesized by the proposed method with various parameters.

Keywords:

novel view synthesis, view-dependent texture mapping, continuity in texture, visibility test, graph cut

目次

1. はじめに	1
2. 関連研究と本研究の位置づけ	3
2.1 Model-based Rendering	3
2.2 Image-based Rendering	7
2.3 3次元形状を援用する Image-based Rendering	10
2.4 本研究の位置づけ	12
3. テクスチャの連続性を考慮した自由視点画像生成	13
3.1 提案手法の概要	13
3.2 エネルギー関数のデータ項	17
3.2.1 仮想視点と入力視点の見え方の近さに関する項 $f_i(x_i)$	17
3.2.2 可視判定に関する重み $\alpha_i(x_i)$	19
3.3 エネルギー関数の平滑化項	20
3.3.1 画素値の連続性に関する項 $g_{ij}(x_i, x_j)$	20
3.3.2 3次元形状の変化の大きさに関する重み β_{ij}	21
3.4 輝度の調整	25
4. 実験	27
4.1 実験の概要	27
4.2 小物体を対象とした平滑化項の効果の比較と考察	30
4.3 輝度の調整に関する比較と考察	44
5. まとめ	46
謝辞	47
参考文献	48

目 次

1	ToF 方式 ([5] より)	4
2	光切断方式 ([6] より)	4
3	視体積交差法 ([9] より)	5
4	Space Carving([9] より)	6
5	PMVS で復元した 3 次元形状を基に生成した自由視点画像例 ([12] より)	6
6	モーフィング ([14] より)	8
7	ライトフィールド ([23] より)	8
8	ライトフィールドレンダリングで生成した自由視点画像例 ([22] より)	9
9	Debevec らの視点依存テクスチャマッピング ([24] より)	11
10	画素値の不連続が生じた自由視点画像例 ([26] より)	11
11	提案手法の概要	14
12	エネルギー関数における各項の関係	17
13	仮想視点と入力視点の見え方の近さに基づくエネルギー	18
14	可視判定の例	20
15	不連続が目立つ自由視点画像例	21
16	3 次元形状の大きな変化と材質による光の当たり方の変化	22
17	ポリゴンの法線取得	23
18	法線マップと横・縦方向の形状の変化	24
19	不連続が残った自由視点画像	26
20	輝度調整の例	26
21	対象物体 A の入力画像例	29
22	対象物体 B の入力画像例	29
23	対象物体 C の入力画像例	29
24	各対象を CMPMVS で生成した結果	31
25	各対象の法線マップ	31
26	各対象における形状の変化の大きさ (黒 : 変化大, 白 : 変化小) . .	33
27	自由視点画像生成結果 (対象物体 A)	34

28	入力視点のラベル画像 (対象物体 A)	35
29	自由視点画像生成結果 (対象物体 B)	36
30	入力視点のラベル画像 (対象物体 B)	37
31	自由視点画像生成結果 (対象物体 C)	38
32	入力視点のラベル画像 (対象物体 C)	39
33	拡大画像 (対象物体 A) 1/2	40
34	拡大画像 (対象物体 A) 2/2	41
35	拡大画像 (対象物体 B)	42
36	拡大画像 (対象物体 C)	43
37	輝度調整	45
38	テクスチャの位置ずれが生じた画像	45

表 目 次

1	入力データの詳細	28
2	実験で用いた PC の仕様	28
3	実験で設定した固定のパラメータ	28

1. はじめに

近年、実写画像群に基づいた自由視点画像生成の研究が盛んに行われている [1-15, 17-20, 22-29]。実写画像群に基づいた自由視点画像生成とは、現実空間のあるシーンを複数の地点から撮影した画像群を入力として、任意の視点(仮想視点)からの見えを再現した画像を生成する技術である。自由視点画像ではユーザが自由に視点位置や視線方向を選択できることから、特に遠隔地のシーンを仮想的に体感できるテレプレゼンス [1] や文化遺産のデジタル化 [2] などのアプリケーションでの活用が期待される。

自由視点画像生成の従来手法として、シーンの3次元形状を利用する Model-based Rendering (MBR) [3-13] や、入力視点の画像群を利用する Image-based Rendering (IBR) [14, 15, 17-20, 22, 23] がある。MBR では、対象シーンの3次元形状を手作業または自動で復元し、仮想視点からのシーンの見えを画像として生成する。この手法では、手作業によるシーンの3次元形状復元に専門技術を要することから、Multi-view Stereo (MVS) [12, 13] などを用いて自動で3次元形状を復元する手法が広く研究されている。しかし、高精度な3次元形状の復元には限界があり、シーン内の物体表面の細かな凹凸が失われる可能性がある。一方、IBR は対象のシーンを撮影した入力画像群を変形・合成することによって仮想視点からの見えを再現する技術である。この手法は対象シーンの複雑な形状をテクスチャとして再現した写実性が高い画像を生成できるが、任意の視点からの見えを忠実に再現するためには多数の画像が必要となる。

そこで、少ない入力画像枚数で細かな形状を再現するための手法として、3次元形状を援用した IBR [24-28] が提案されている。その一手法である視点依存テクスチャマッピングでは、MVS 等で復元した3次元形状に対するテクスチャとして入力視点の画像群を適応的に貼り付けることで3次元復元精度が不十分な場合においても詳細な形状を視覚的に再現する [24-28]。この手法では、撮影された地点による光源環境の差異や3次元形状とカメラパラメータの誤差に起因して入力画像の画素値にばらつきが生じるため、テクスチャの境界において画素値の不連続が生じる可能性がある。この問題に対して、境界領域で滑らかにテクスチャが変化するように入力画像をブレンディングすることで不連続を抑制することが可

能である。しかし、3次元形状やカメラパラメータの誤差が大きい場合には、複数の入力画像のテクスチャの投影位置が大きくずれるため、これらがブレンディングされることで物体表面のテクスチャがぼけるという問題が生じる。

そこで、本研究では3次元形状が高精度に復元できない場合でも、入力画像をテクスチャとして用いることでシーンの見えを写實的に再現することが可能な視点依存テクスチャマッピングをベースとし、ブレンディングによらずに実写画像を適切に切り替えることで視覚的に自然な自由視点画像を生成する手法を提案する。具体的には、テクスチャ選択において出力画像上でのテクスチャの連続性を考慮することで、不連続が目立ちにくい領域でテクスチャを切り替えることを目標に、生成される自由視点画像内の隣接する画素間の画素値の差分や3次元形状を考慮したエネルギー関数を設計し、この関数をグラフカットを用いて最小化することにより各画素でテクスチャとして利用する入力画像を選択する。また、各画素に対して局所的な輝度調整を行い、エネルギー関数の最小化で除去できなかった輝度の不連続性を低減する。実験では、様々なパラメータで設定したエネルギー関数で生成した自由視点画像を比較することで提案手法の有効性を検証する。

本論文では、2章で自由視点画像生成の従来手法を紹介し、それらの問題点を踏まえて本研究の位置づけについて述べる。3章では、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。4章では、様々なパラメータで設定したエネルギー関数において、提案手法で生成した自由視点画像の結果を比較することで各基準の有効性を検証した上で輝度調整の効果を確認する。最後の5章でまとめと今後の展望について述べる。

2. 関連研究と本研究の位置づけ

本章では，自由視点画像生成の従来手法を，シーンの3次元形状を利用する Model-based Rendering (MBR) [3-13]，入力視点の画像群を利用する Image-based Rendering (IBR) [14, 15, 17-20, 22, 23]，3次元形状を援用する IBR [24-28] の3つに大別し，それぞれについて詳述するとともに，これらに対する本研究の位置づけについて述べる．

2.1 Model-based Rendering

Model-based Rendering (MBR) とは，手作業，もしくはレーザレンジファインダ [3-6] などのアクティブ計測や視体積交差法 [7-9]，Multi-view Stereo (MVS) [12, 13] などの3次元形状復元手法により得られた対象シーンの3次元形状を利用し，仮想視点からのシーンの見えを画像として生成する技術である．以下では代表的な3次元形状復元手法として，レーザレンジファインダなどを用いるアクティブ計測法および，対象のシーンを撮影した画像群を入力とするパッシブ計測法である視体積交差法と MVS の手法を紹介する．

アクティブ計測法 ToF 方式のレーザレンジファインダ [3] を用いる手法では，光源から投射されたレーザが測定対象で反射して戻るまでの時間から求められた距離を基に対象の3次元位置を算出する (図1)．一方，光切断方式 [4] では，光源から投射されたレーザの観測対象表面上の反射光をカメラで観測し，光源と反射光位置，カメラからなる3点に対して三角測量の原理で光が反射した点の3次元位置を求める (図2)．これらの手法はシーンの3次元形状を高精度に取得できるが，光を吸収したり透過したりする物体では反射したレーザ光を観測することができない．また，レーザレンジファインダなどの機器は高価で入手が容易ではないため，手軽にシーンの3次元形状を復元することはできない．

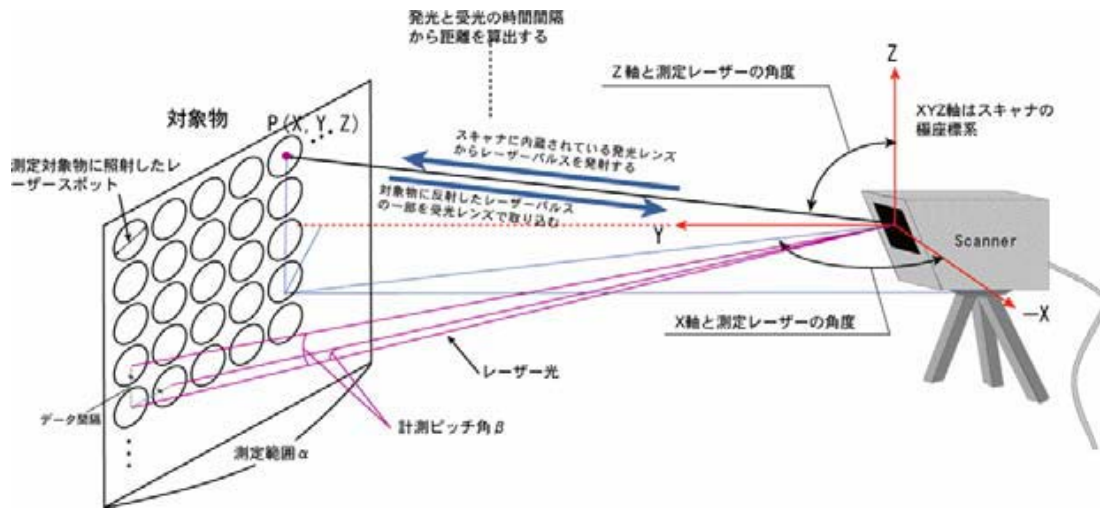


図 1 ToF 方式 ([5] より)

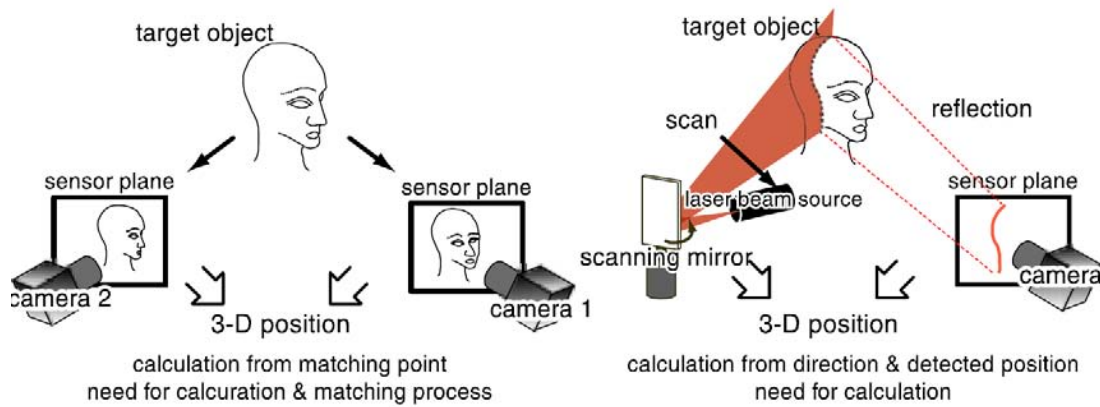


図 2 光切断方式 ([6] より)

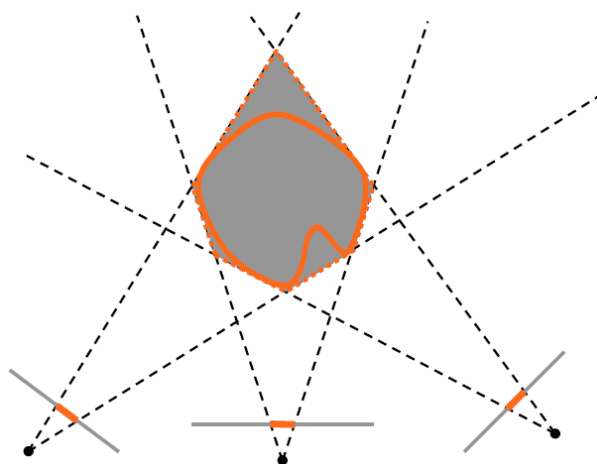


図 3 視体積交差法 ([9] より)

視体積交差法 視体積交差法は、複数の入力画像群から抽出された対象の輪郭を基に 3次元形状を復元する手法である [7-9]。視体積とは、入力視点を頂点とし、対象の輪郭で囲まれた領域を断面とする錐体である。Baker らは、図 3 のように撮影された対象の輪郭を 3次元空間に逆投影することで視体積を求め、異なる複数の入力視点から観測された視体積の交差部分を計算することで対象の 3次元形状を復元する手法を提案した [7]。この手法では、対象の輪郭を利用することから後述する MVS などの手法と比較するとアルゴリズムが単純で高速に計算可能である。また、テクスチャが無い対象でも輪郭さえ抽出できれば復元可能で、色や照明による誤差が小さい。しかし、視体積交差法では原理的に凹んだ形状を復元できず、毛や植物などの細かい形状の復元精度は低い。

MVS MVS では、まず Structure from Motion (SfM) [10] などで推定した入力画像群のカメラパラメータを基に密な 3次元形状を取得する。SfM は対象となるシーンを複数の地点から撮影した画像群を入力として、そのシーンの 3次元形状と入力視点のカメラパラメータを同時に復元する技術である。MVS では SfM で得られた入力視点のカメラパラメータを利用し、入力画像間の画素の対応に基づいて 3次元形状を復元する。Kutulakos らは Space Carving を提案した [11]。この

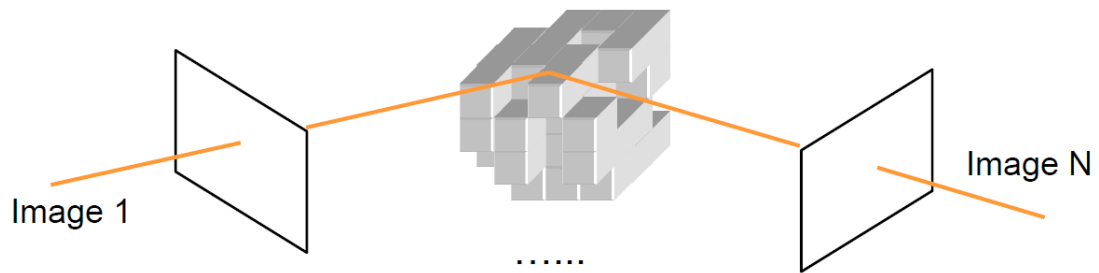


図 4 Space Carving([9] より)



図 5 PMVS で復元した 3 次元形状を基に生成した自由視点画像例 ([12] より)

手法ではまず，対象全体を内包するボクセルボリュームを用意する．続いて，図 4 のようにボクセルボリュームの表面上のボクセルを各入力視点の画像へ投影し，投影箇所画素値が異なればそのボクセルを削除する．この操作を収束するまで繰り返すことで，凹形状の表現も可能にした．MVS の一手法である Patch-based Multi-view Stereo (PMVS) では，画像群から法線付きの 3 次元点を復元し，それをメッシュモデルに変換することで，詳細な 3 次元形状を取得する [12]．しかし，テクスチャの無い領域では画像間の対応が取れないため，3 次元形状を推定できない領域に欠損を生じる (図 5)．Jancosek らが提案した CMPMVS では，視体積

交差法の考え方を取り入れることで、テクスチャの無い領域の3次元形状復元を可能にしている [13].

MBR では、植物などの複雑で細かい形状を持つ対象を高精度に復元することや、シーンの鏡面反射を再現することが困難である.

2.2 Image-based Rendering

Image-based Rendering (IBR) とは、対象のシーンを撮影した入力画像群を変形・合成することによって仮想視点からの見えを再現した自由視点画像を生成する技術である [14, 15, 17-20, 22, 23]. 以下では、入力画像を変形・合成するモーフィング [14, 15] と、入力画像群からライトフィールドを取得するライトフィールドレンダリング [16-23] について概説する.

モーフィング モーフィングとは、2つの画像間で指定された対応点と各画像の混合比を基に、一方の画像からもう一方の画像へと滑らかに変化させる手法である [14, 15] (図 6). 2つの異なる視点の画像を入力とするモーフィングで自由視点画像を生成する場合、単に混合比を変化させるだけではシーンの幾何学的な構造を再現できない. そこで、Seitz らは各入力視点の画像を平行化した上でモーフィングを行い、再び元の画像面に再変換することで、入力画像の視点間を繋ぐ直線上での仮想視点において幾何学的な構造を一定に保つことを可能にした [14]. Xiao らは3つの視点の画像を用いる手法を提案したが、仮想視点の位置はこれらに囲まれた範囲内に限定されるため、その適用範囲には限界がある [15].

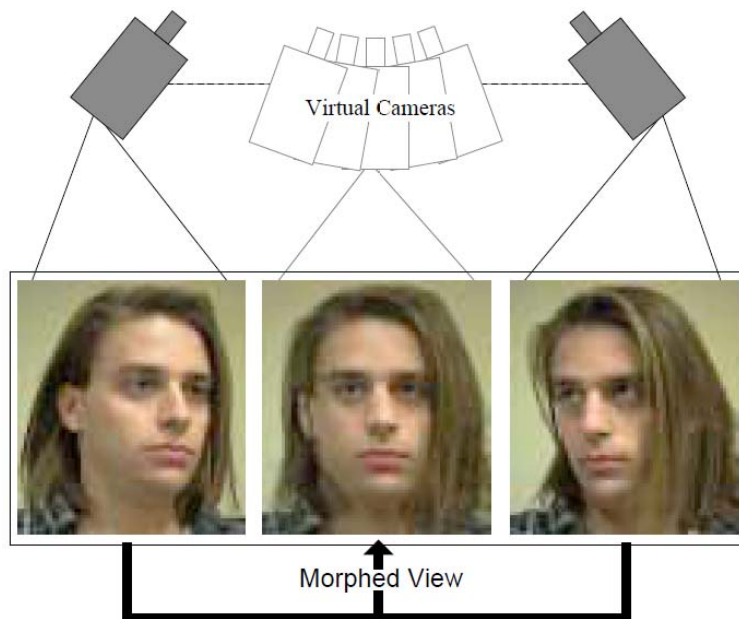


図 6 モーフィング ([14] より)

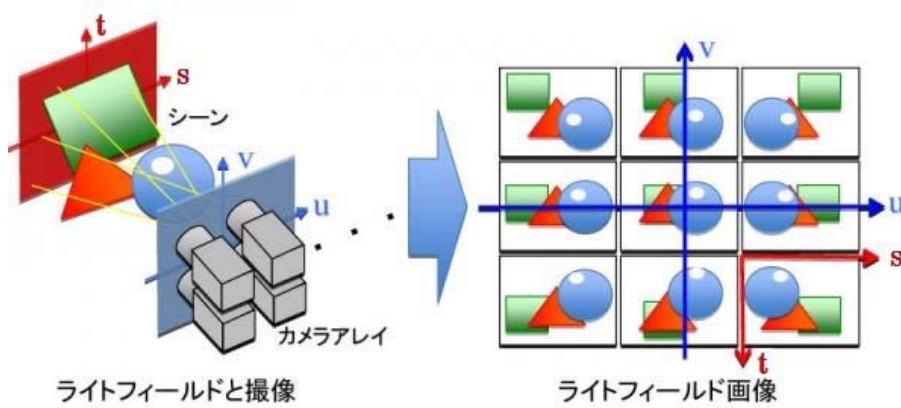


図 7 ライトフィールド ([23] より)



図 8 ライトフィールドレンダリングで生成した自由視点画像例 ([22] より)

ライトフィールドレンダリング ライトフィールドとは3次元空間内の全ての光線場を指すものであり，視点の位置 (x, y, z) と光線の向き (θ, ϕ) ，光線の波長 (λ) ，時間 (t) の計7次元で定義される Plenoptic 関数 [16] で表現される．自由視点画像生成において，対象シーンのライトフィールドが取得できれば，それは本来の3次元空間における全ての見えの情報を持つことに等しいため，原理的には任意の仮想視点を通る光線を取り出すことで自由視点画像を生成することが可能である．しかし，あるシーンのライトフィールドを取得する際，7次元の光線情報を密に取得するためには膨大な入力画像が必要になる．そこで，McMillan らは7次元で表現されるライトフィールドの次元数を減らすために，光線の波長 (λ) を RGB 色空間などの画素値で代替し，時間軸を無視した静止シーンを仮定することで，視点の位置 (x, y, z) と光線の向き (θ, ϕ) のみで定義した 5D Plenoptic 関数を提案している [17]．また，Naemura らは光線の減衰は無いと仮定した場合，2次元の視点位置 (x, y, z) と光線の向き (θ, ϕ) の4次元で表現できることを利用し，ライトフィールドを取得するために必要な画像枚数を大幅に削減した [18]．更に，Levoy ら [19] や Gortler ら [20] はそれぞれ Light Field Rendering, Lumigraph を提案し，前述した4次元のライトフィールドを直線で表した光線が2つの平行した平面を通るときの交点 (u, v) ， (s, t) の組で定義することで，より効率的なレンダリングを可能にした (図 7)．

ライトフィールドの取得に単眼カメラを用いる場合、各入力視点の相対的な位置を知るために複数のカメラを規則的に整列させるか、ロボットアームを用いてカメラを移動させる必要がある。しかし、複数のカメラを整列させたり、ロボットアームを用意したりすることは容易ではない。そこで、カメラの撮像面に複数のマイクロレンズを規則的に整列させることにより一台のカメラで撮影できるライトフィールドカメラが提案されている [21]。また、Davis らは携帯可能な単眼カメラを用いてライトフィールドを取得するために、最適な撮影位置姿勢をユーザに提示するシステムを提案した [22]。

IBR では、自由視点画像を生成可能な視点の位置や生成される画像の質が入力画像数に大きく依存し、少数の画像から高品質な画像を合成することは困難である (図 8)。

2.3 3次元形状を援用する Image-based Rendering

3次元形状を援用する IBR では、対象シーンの3次元形状に入力視点の画像を適応的に貼り付けることで、3次元復元精度が不十分な場合においても詳細な形状を視覚的に再現する [24-28]。以下では、この手法で広く用いられている視点依存テクスチャマッピングとその拡張について説明する。

視点依存テクスチャマッピングとは、MVSなどで復元した3次元形状に対するテクスチャとして適切な入力視点の画像を選択的に用いる手法である。仮想視点における対象物体の見えに近い入力画像を決定する基準として、Debevec ら [24] は図 9 に示すように、仮想視点から注目画素を見たベクトルと各入力視点から注目画素を見た方向ベクトルの成す角を採用している。源田ら [25] は Debevec らの基準に加え、物体表面上の法線情報も考慮することで、生成する画像の質を向上させた。また、高井ら [26] は高精細な自由視点画像を生成するために、高ズームで撮影された入力視点の画像を優先的に採用することで高精細なテクスチャを生成する手法を提案した。

しかし、撮影された地点による光源環境の差異に起因して入力画像の画素値にばらつきが生じるため、図 10 に示すようにテクスチャの境界において画素値の不連続が生じる場合がある。この問題に対して、Debevec ら [27] は境界領域で滑

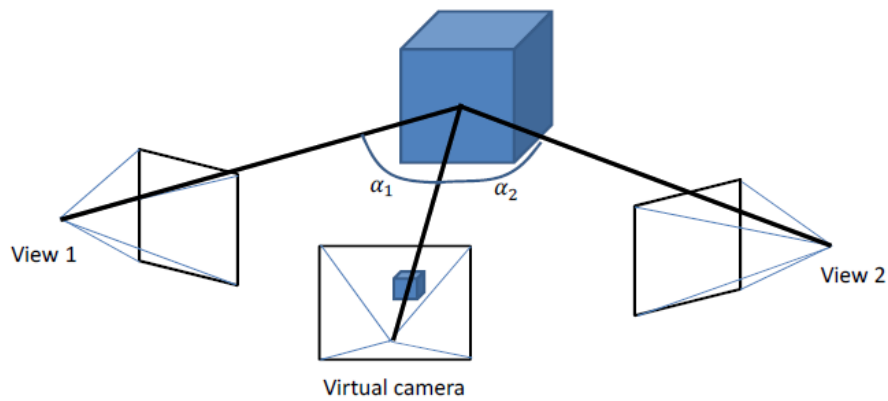


図 9 Debevec らの視点依存テクスチャマッピング ([24] より)



図 10 画素値の不連続が生じた自由視点画像例 ([26] より)

らかにテクスチャが変化するように複数の入力画像をブレンディングすることで不連続を抑制している。しかし、3次元形状やカメラパラメータに誤差がある場合には、複数の入力画像のテクスチャの投影位置に位置ずれが生じるため、これらがブレンディングされることでテクスチャのぼけが生じる。これらの問題を低減する手法として、Kopfら [28] はテクスチャの画素値を考慮したエネルギー関数をグラフカットで最小化し、テクスチャとして利用する入力視点を決定することで、不連続などが生じにくい画像を生成する手法を提案した。しかし、この手

法では時間軸を考慮するために仮想視点の経路を事前に指定する必要があり，また仮想視点の位置を入力視点の周辺に限定しているため，汎用的に用いることが困難である．

2.4 本研究の位置づけ

本研究では，視点依存テクスチャマッピングをベースとし，テクスチャに不連続などが生じない視覚的に自然な自由視点画像を生成する手法を提案する．視点依存テクスチャマッピングをベースとする理由は，3次元形状が高精度に復元できない場合でも，入力画像をテクスチャとして用いることでシーンの見えを写実的に再現することが可能なためである．提案手法は，Kopfら [28] の手法と同様に，テクスチャとして用いる画像に関するエネルギー関数を定義し，グラフカットアルゴリズム [30] を用いてこれを最小化することによってテクスチャを選択する．提案手法では，Kopfらのエネルギー関数で用いられるテクスチャの連続性に加え，物体の3次元形状が大きく変化する部分では物体の画素値が大きく変化すると考え，3次元形状の連続性に関する基準をエネルギー関数に導入する．これにより，形状の変化が大きい領域でのテクスチャの切り替えを許容し，形状の変化が小さい領域での不連続を低減する．更に，エネルギー最小化に基づく手法では，光源環境の変化などによる入力画像の画素値の変化による不連続を解消しきれない場合があることから，提案手法では局所的な輝度の調整により更に不連続を低減する．

3. テクスチャの連続性を考慮した自由視点画像生成

本章では、Debevecら [24] の視点依存テクスチャマッピングに対して、テクスチャの連続性を考慮したテクスチャ選択を導入することにより、ブレンディングすることなくテクスチャの不連続を抑止する自由視点画像の生成手法について述べる。以下では、まず提案手法について概説する。次に、提案するテクスチャ選択処理におけるテクスチャの連続性を考慮した目的関数について詳述する。最後に、目的関数を用いた処理で解消しきれないテクスチャの不連続を抑止するための輝度調整について述べる。

3.1 提案手法の概要

提案手法では、実空間の物体を対象として、その対象を複数の地点(入力視点)から撮影した画像群を入力とし、ユーザが指定した視点(仮想視点)からの対象の見えを画像(自由視点画像)として出力する。処理の概要を図11に示す。

オフライン処理では、対象を撮影した画像群から、既存の手法により対象の3次元形状、および入力視点のカメラの内部・外部パラメータを推定する。オンライン処理では、生成される自由視点画像内の隣接する画素間の画素値の差分や3次元形状などを考慮したエネルギー関数を設定し、グラフカットアルゴリズムによりその関数を最小化することで、自由視点画像の各画素に対してテクスチャとして利用する入力画像を選択する。最後に、局所的な輝度調整によって異なるテクスチャの境界で生じた画素値の不連続を更に低減し、最終的な自由視点画像としてユーザへ提示する。以下、オフライン処理とオンライン処理について述べる。

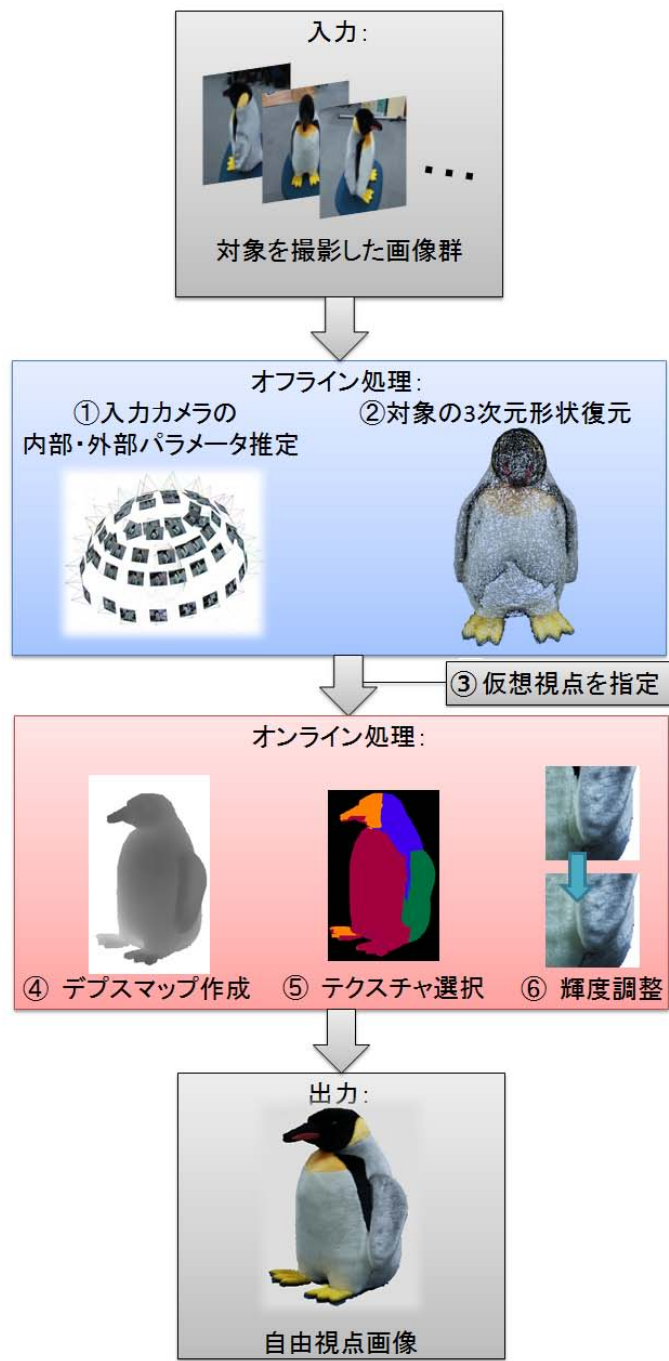


図 11 提案手法の概要

オフライン処理 オフライン処理で用いる3次元復元手法として、2.1節で紹介したMVSの一手法であるCMPMVS [13]を採用する。CMPMVSで密な3次元形状を求めるための前処理として、まず対象を複数の地点で撮影した画像 I_n からなる入力画像群 $S = \{I_n | n = 1, \dots, N\}$ から、SIFT特徴点 [31]を抽出し、SfMにより画像 I_n を撮影したカメラの内部パラメータ \mathbf{K}_n 、および任意の座標系における外部パラメータ (回転行列 \mathbf{R}_n 、並進ベクトル \mathbf{t}_n) を推定する。本研究では、SfMで得られたカメラの外部パラメータの座標系をワールド座標系とする。内部パラメータ \mathbf{K}_n は

$$\mathbf{K}_n = \begin{pmatrix} f_n & 0 & c_n^x & 0 \\ 0 & f_n & c_n^y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (1)$$

として表され、焦点距離 f_n と光学中心 (c_n^x, c_n^y) で構成される。提案手法では、SfMとしてVisual SFM [32, 33]を利用する。続いて、入力画像群とSfMによって得られたカメラの内部・外部パラメータからCMPMVS [13]によって対象の3次元形状を復元する。CMPMVSでは、3次元形状はメッシュモデルにより表される。

オンライン処理 オンライン処理では、2.3節で紹介したDebevecらの手法 [24]を基に、まずメッシュモデル M とユーザに指定された仮想視点 (内部パラメータ \mathbf{K}_V 、外部パラメータ \mathbf{R}_V 、 \mathbf{t}_V) における自由視点画像に対応するデプスマップを生成する。続いて自由視点画像中の画素 i に対応するデプスマップの値 d_i を利用して入力画像群 S 内の画像中の対応する画素を決定し、これらの画素値に基づいて自由視点画像の画素値を決定することで自由視点画像を生成する。提案手法では、複数の画像から得られた画素値のブレンディングによる問題を低減するために、画素 i に対して S 内の単一の画像を選択し、その画像から対応する画素の画素値を割り当てることで、画素 i の画素値を決定する。このとき、画像の選択のための基準として、下記4点の条件を考慮する。

- (1) 3次元形状は視点位置によってその見えが異なることから、自由視点画像の画素 i に対して割り当てる画素値として、入力画像群の中でも仮想視点の位

置に近い位置で撮影された画像上の画素 i に対応する画素値が望ましい。

- (2) 3次元形状上のある点は、自己遮蔽などにより入力画像群中のいくつかの画像では観測されない可能性がある。このため、自由視点画像の画素 i に対して画素値を割り当てる際には、入力画像上で画素 i に対応する3次元点が観測されているかを判定(可視判定)する必要がある。
- (3) 提案手法では、自由視点画像の各画素に対して適切な画像を S から選択するため、画像の切り替わりで画素値の不連続が生じる可能性がある。この問題を低減するために、自由視点画像中の画素 i と隣接する画素 j で異なる画像が選択される場合には、画素 i と画素 j の画素値が類似する必要がある。
- (4) 3次元形状の大きな変化は、その物体の構造などの切り替わりに対応すると考えられることから、画素 i と隣接する画素 j に対応する3次元形状上の点の法線が大きく変化する場合には、条件(3)は緩和される。

提案手法では、各画素で用いる入力画像を決定する問題を、上記を考慮したエネルギーの最小化問題として定式化する。具体的には、 $X = \{x_i | i = 1, \dots, N\}$ の要素 x_i を i の画素値の決定に利用される画像のラベル $x_i \in \{1, \dots, N\}$ として、下記のエネルギー関数を定義する(図12を参照)。

$$J(X) = \sum_{i \in A} \alpha_i(x_i) f_i(x_i) + \sum_{(i,j) \in B} (K_\beta \beta_{ij} + K_g) g_{ij}(x_i, x_j) \quad (2)$$

ここで、データ項の $f_i(x_i)$ は、条件(1)に対応しており、仮想視点位置と x_i に対応する画像を撮影したカメラの位置が離れている場合に大きな値を与える。また、 $\alpha_i(x_i)$ は条件(2)に対応しており、画素 i が x_i に対応する画像から観測される場合には1を、そうでない場合は大きな値を与える。平滑化項では、 $g_{ij}(x_i, x_j)$ によって条件(3)を表し、 β_{ij} によって条件(4)に対応させる。ただし、 A は自由視点画像で画素値を割り当てるすべての画素のインデックスの集合、 B は隣接する画素のインデックスの集合を表す。式(2)は、データ項と平滑化項で構成されるエネルギー関数であり、グラフカットアルゴリズムによって最小化することができる。ここで、 K_β , K_g は重みパラメータを表す。

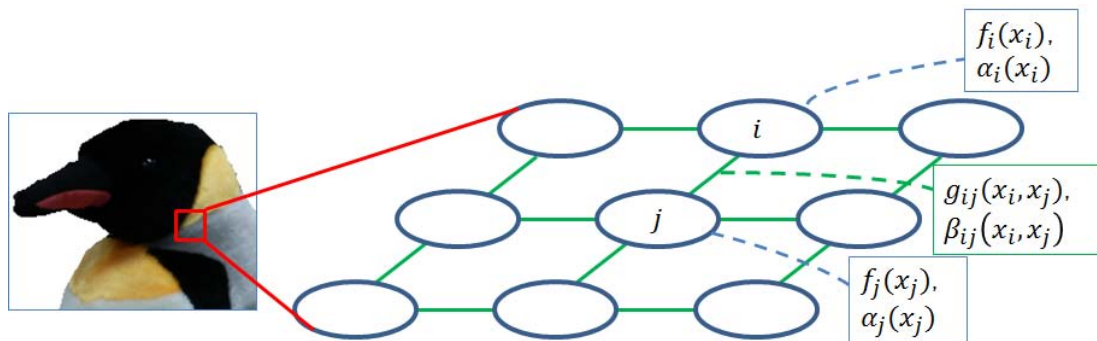


図 12 エネルギー関数における各項の関係

提案手法では，入力画像群から得られる画素値を直接利用することから，撮影時の照明条件の変化などに起因する画像間の画素値の変化に対応できず，自由視点画像中に生じる不連続を抑止しきれない可能性がある．提案手法では，この画素値の変化が輝度の変動のみであるものと仮定し，エネルギー最小化問題によって各画素の画素値として利用する画像を選択した後，局所的な輝度調整を行うことで，不連続な画素値の変化を抑制する．

以降の節では，エネルギー関数のデータ項と平滑化項についてそれぞれ詳述した後，局所的な輝度調整について説明する．

3.2 エネルギー関数のデータ項

この節では，データ項を構成する $f_i(x_i)$ ，および $\alpha_i(x_i)$ について述べる．

3.2.1 仮想視点と入力視点の見え方の近さに関する項 $f_i(x_i)$

視点依存テクスチャマッピングでは，3次元形状に対する適切な入力画像の選択基準として，仮想視点と入力画像を撮影したカメラ（入力視点）の位置の類似度が広く利用されている [24, 25, 27, 28]．代表的なものとして，2.3節で紹介した Debevec ら [24] の手法がある．この手法では，図 13 に示すように，仮想視点から

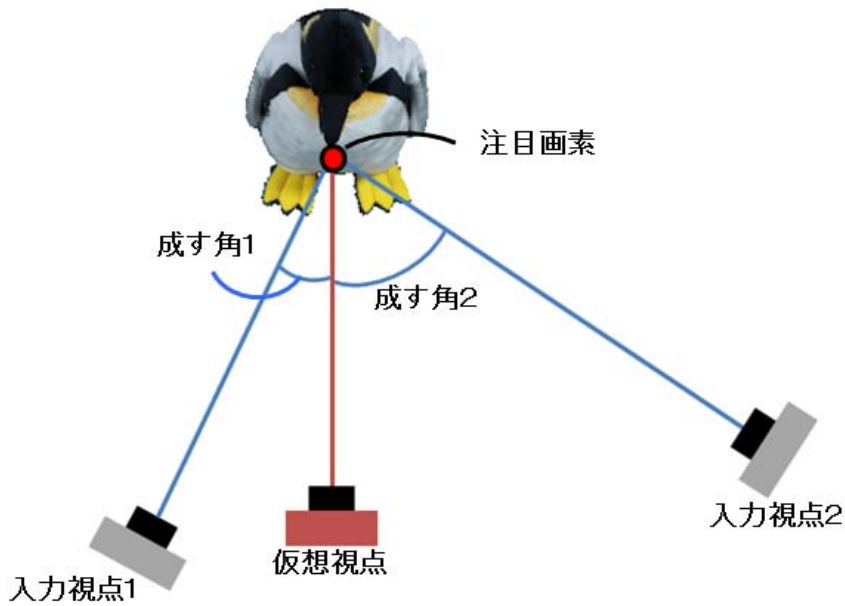


図 13 仮想視点と入力視点の見え方の近さに基づくエネルギー

注目画素に対応する 3次元形状上の点に向かうベクトルと各入力視点から同じ点に向かうベクトルの成す角の大きさを選択基準として利用している。

提案手法でも同様に、これらのベクトルの成す角の大きさを $f_i(x_i)$ として利用する。具体的には、 I_n に対応する入力視点位置 $\mathbf{t}'_n = -\mathbf{R}_n^\top \mathbf{t}_n$ 、仮想視点位置 $\mathbf{t}'_V = -\mathbf{R}_V^\top \mathbf{t}_V$ 、および画素 i に対応するデプス値 d_i から仮想視点のカメラパラメータを使って得られるワールド座標系における対象物体の 3次元形状上の点 \mathbf{p}_i を使って、これらの成す角 θ_{in} は次式により得られる。

$$\theta_{in} = \cos^{-1} \frac{(\mathbf{t}'_V - \mathbf{p}_i) \cdot (\mathbf{t}'_n - \mathbf{p}_i)}{\|\mathbf{t}'_V - \mathbf{p}_i\| \|\mathbf{t}'_n - \mathbf{p}_i\|} \quad (3)$$

この成す角を用いて、データ項 $f_i(x_i)$ を次式により与える。

$$f_i(x_i) = \theta_{ix_i} \quad (4)$$

この項は画素 i において θ_{ix_i} が小さい場合に小さな値となり、 x_i に対応する画像が選択され易くなる。

3.2.2 可視判定に関する重み $\alpha_i(x_i)$

可視判定では，自由視点画像中の各画素に対応する物体上の点 \mathbf{p}_i が各入力画像 I_n から観測されているかを判定する．図 14 中の赤色の領域は，入力視点から観測できない領域を例示している．可視判定では，まずオフライン処理として3次元形状に基づき各入力視点に対応するデプスマップを作成する．次に，オンライン処理において仮想視点のデプスマップから前節で述べた手法と同様にして自由視点画像の画素 i に対応する3次元形状上の点 \mathbf{p}_i を求め，各入力視点のデプスマップへと投影する．点 \mathbf{p}_i が入力画像 I_n のデプスマップの画素 k に投影されたものとし，そのときのデプス値を d'_{in} ，また入力視点のデプスマップのデプス値を e_{kn} とすると，これらのデプス値の差が小さい場合，点 \mathbf{p}_i は I_n 上で観測されたと考えることができる．一方，差が大きい場合には点 \mathbf{p}_i より入力視点に近い位置に遮蔽物があると考えられる．提案手法では I_{x_i} における \mathbf{p}_i の可視判定に対応する重み $\alpha_i(x_i)$ を次式により定義する．

$$\alpha_i(x_i) = \begin{cases} 1 & \text{if } |d'_{ix_i} - e_{kx_i}| < \theta_d \\ K_\alpha & \text{otherwise} \end{cases} \quad (5)$$

ただし， θ_d は閾値であり， K_α は \mathbf{p}_i が I_{x_i} において観測されていない場合のペナルティである．可視判定の結果， \mathbf{p}_i は I_{x_i} で観測されないと判定された場合， I_{x_i} の画素値は利用できないため，本研究では K_α に十分に大きな値を設定する．



図 14 可視判定の例

3.3 エネルギー関数の平滑化項

この節では、自由視点画像の各画素に割り当てられた入力画像ラベルの境界における画素値の連続性と、メッシュモデル M の法線に基づいた形状の変化の大きさに基づく平滑化項について詳述する。

3.3.1 画素値の連続性に関する項 $g_{ij}(x_i, x_j)$

テクスチャの不連続は、図 15 に示すように隣接する画素間において、異なる入力画像ラベルが割り当てられた際に、入力画像上に対応する画素の画素値が大きく異なる場合に生じる。そこで、この項では生成する自由視点画像上で隣接する画素の画素値の不連続を抑止する。画素値の連続性に関する項 $g_{ij}(x_i, x_j)$ を次式で定義する。

$$g_{ij}(x_i, x_j) = \max(\|\mathbf{c}_{ix_i} - \mathbf{c}_{ix_j}\|_\infty, \|\mathbf{c}_{jx_i} - \mathbf{c}_{jx_j}\|_\infty) \quad (6)$$

ただし、 \mathbf{c}_{ix_i} は自由視点画像上の画素 i に対応する入力画像 I_n 上の点における画素値を表す。 $\|\cdot\|_\infty$ は上限ノルムであり、ベクトルの要素の最大値を与える。こ

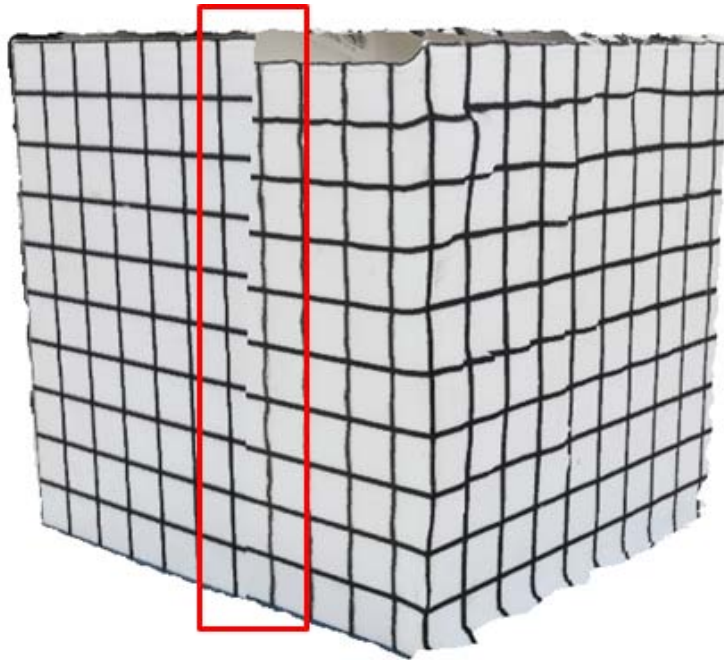


図 15 不連続が目立つ自由視点画像例

の項は，隣接する画素が異なる入力画像ラベルを割り当てられ，その画素値が大きく異なる場合にペナルティを与える．

3.3.2 3次元形状の変化の大きさに関する重み β_{ij}

図 16 に示すように，対象の 3次元形状が大きく変化する部分では，その表面を構成する材質や光の当たり方などが異なると考えられる．そこで提案手法では，3次元形状が大きく変化する部分で画素値が大きく変化するような入力画像ラベルの切り替えを許容することで，画素値の不連続を抑制することにより生じる歪みを緩和する．

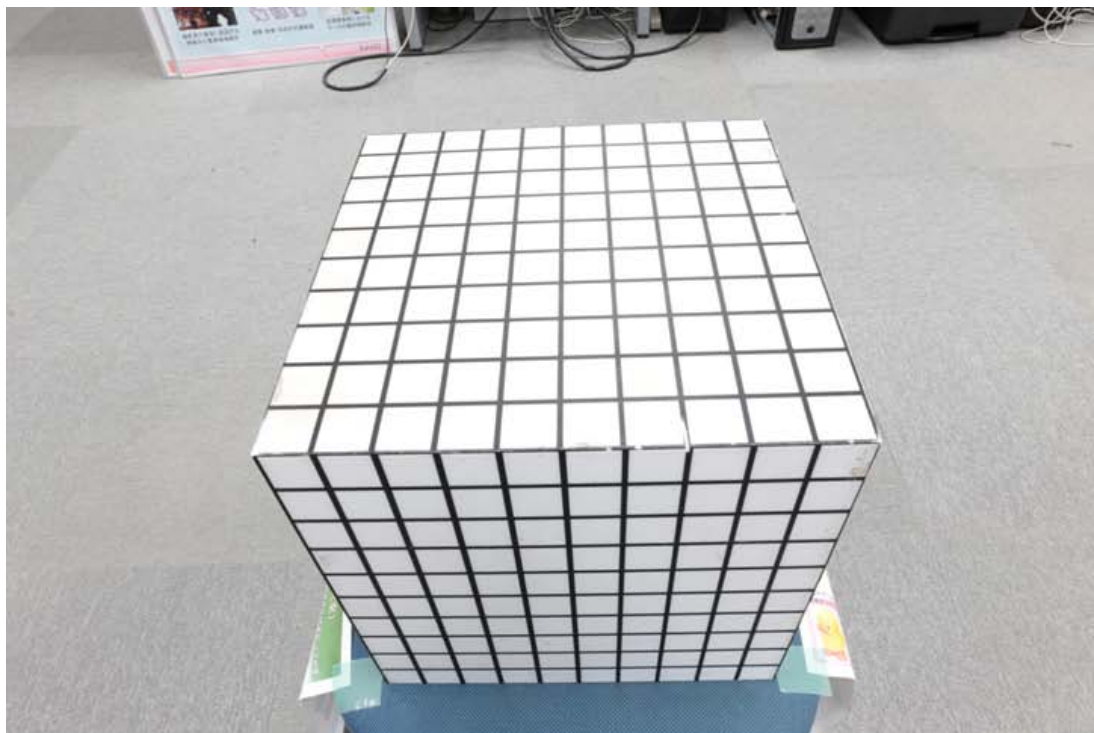


図 16 3次元形状の大きな変化と材質による光の当たり方の変化

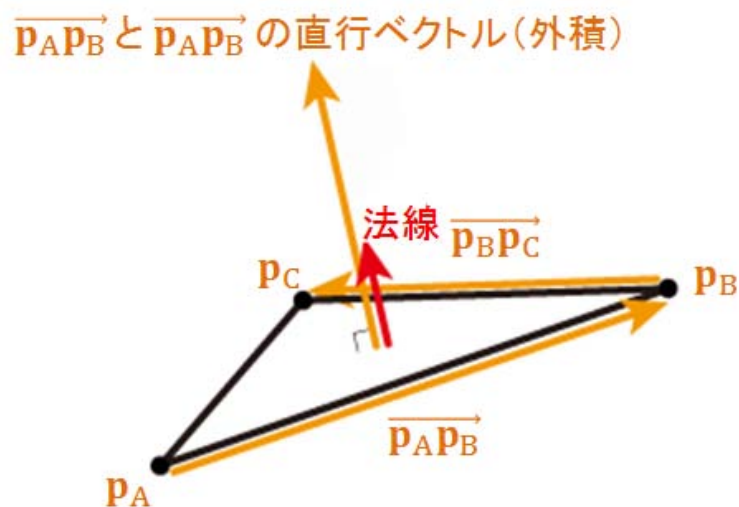
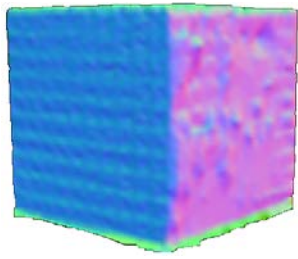


図 17 ポリゴンの法線取得

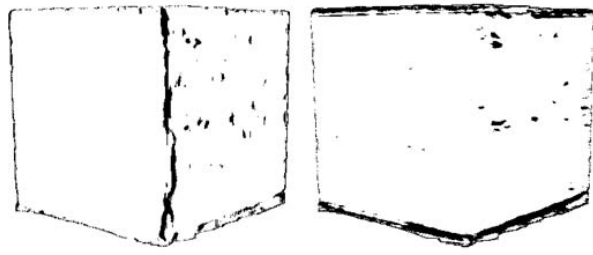
提案手法では、3次元形状の変化の大きさを表す量として、自由視点画像の各画素に対して割り当てられた法線の向き之差を利用する。まず、CMPMVS で得られた対象のメッシュモデル M から、オフライン処理として M を構成する面の法線を求める。ある面を構成する頂点を p_A , p_B , p_C とし、これらの頂点は面の表側から見たときにこの順番で反時計回りになるものとする、この面の法線 \mathbf{n} は次式により求められる (図 17 参照)。

$$\mathbf{n} = (\mathbf{p}_B - \mathbf{p}_A) \times (\mathbf{p}_C - \mathbf{p}_A) \quad (7)$$



法線マップ

$$(x, y, z) = (R, G, B)$$



形状の変化の大きさ (黒：変化大，白：変化小)

左：横方向の変化，右：縦方向の変化

図 18 法線マップと横・縦方向の形状の変化

オンライン処理では，仮想視点の画像平面上に投影されたメッシュモデル M を構成する面に基づいて自由視点画像上の各画素に法線を割り当てる．これにより得られた法線は， M の形状復元の誤差などにより，実際の対象が滑らかな場合でも方向の変動が生じるため，自由視点画像に対する法線マップを生成し，このマップに対してガウシアンフィルタを適用する (図 18 参照)．提案手法では，3次元形状の変化に関する重みとして，次式で定義されるものを利用する．

$$\beta_{ij} = \begin{cases} 1 & \text{if } \cos^{-1} \frac{\mathbf{n}_i \cdot \mathbf{n}_j}{\|\mathbf{n}_i\| \|\mathbf{n}_j\|} < \theta_n \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

ただし， \mathbf{n}_i ，および \mathbf{n}_j はそれぞれ画素 i ，および j に割り当てられたガウシアンフィルタ適用後の法線であり，大きさが 1 になるよう正規化されたものとする．この重みは，閾値 θ_n に基づいて，隣接する画素に割り当てられた法線の成す角が大きければ 3次元形状の変化が大きいと判断し，入力画像ラベルの切り替えを許容するために図 18 に示すような黒色の領域で重みを 0 とする．

3.4 輝度の調整

上記のエネルギー関数のグラフカットアルゴリズム [30] による最小化により自由視点画像内の不連続は低減されると考えられる。しかし、図 19 に示すように撮影時の照明環境の変化などにより、不連続を抑止しきれない領域が存在する可能性がある。これに対して、2.3 節で紹介した Debevec ら [24] の手法のように異なる入力画像から得られた画素値をブレンディングする手法があるが、3次元形状やカメラパラメータ推定に誤差があり、異なる入力画像上の点が対象の同一の点に対応しない場合に、表面の構造やテクスチャが二重になるなどの問題が生じる。そこで、提案手法では注目画素周辺の局所的な輝度を調整することによって、この不連続性の問題を解消する。

具体的には、まず注目画素 i に対してウィンドウを設定し、そのウィンドウ内に含まれる画素が持つ入力画像ラベルの集合を Q_i とし、入力画像ラベル $n \in Q_i$ について、ウィンドウ内で n が割り当てられた画素の割合を γ_{in} とする。また、全ての画素の入力画像ラベルを n として生成した自由視点画像の同じウィンドウ内における平均輝度を \bar{I}_{in} とする。注目画素 i の入力画像ラベルを x_i とすると、輝度調整後の画素値 c'_i は、テクスチャ選択により決定された画素値 c_i を使って

$$c'_i = \frac{c_i}{\bar{I}_{ix_i}} \sum_{n \in Q_i} \gamma_{in} \bar{I}_{in} \quad (9)$$

により与えられる (図 20 参照)。本処理においては、ウィンドウ内の全ての画素が同一のラベルに属する場合には、注目画素の画素値がそのまま使われる。

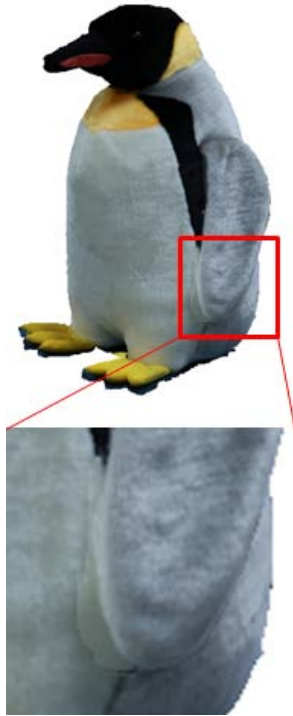


図 19 不連続が残った自由視点画像

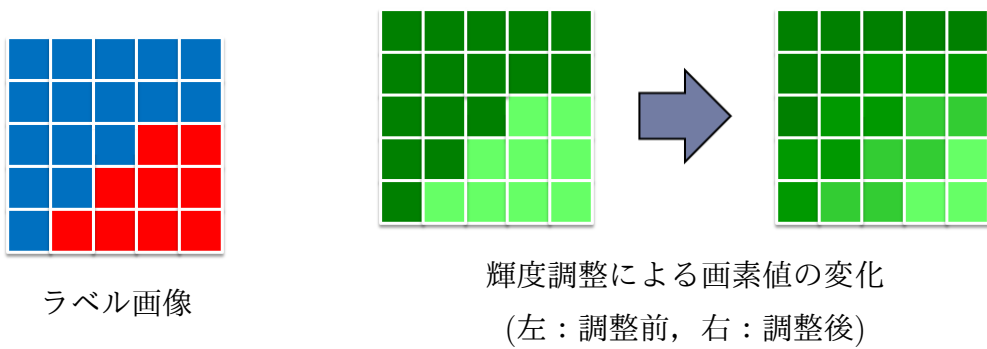


図 20 輝度調整の例

4. 実験

提案手法の有効性を示すために、現実空間の小物体を対象とした比較実験を行う。以下では、実験方法を概説し、様々なパラメータにおいて提案手法で生成した自由視点画像の生成結果を比較し、各基準の有効性を検証した上で、輝度調整の効果を確認する。

4.1 実験の概要

本実験では、3つの小物体を複数の地点(入力視点)から撮影した画像を入力として用いる。入力画像群の詳細、および実験で用いたPCの仕様をそれぞれ表1、表2に、またそれぞれの画像例を図21-23に示す。オフライン処理では、まず入力画像群に対してVisual SFM [32, 33]を適用し、入力画像群の各画像についてそのカメラの内部・外部パラメータを推定した。続いて、CMPMVS [13]を利用してそれぞれの小物体の3次元形状を復元する。オンライン処理では、Boykovらのグラフカットアルゴリズム [30]を用いて提案したエネルギー関数を最小化することで、生成される自由視点画像上の各画素の画素値を決定するための入力画像のラベルを決定する。最後に、各画素の画素値を得られたラベルに対応する入力画像から取得することで自由視点画像を生成する。

本実験では、エネルギー関数における各基準の有効性を示すために、 β_{ij} の重みパラメータ K_β と $g_{ij}(x_i, x_j)$ の重みパラメータ K_g をそれぞれ変化させて、入力視点を表すラベルと対応する自由視点画像がどのように変化するかを比較する。ここで、 β_{ij} と $g_{ij}(x_i, x_j)$ の重みパラメータを0にする場合、Debevecら [24]と同様な自由視点画像を得られるため、これをベースラインとする。また、表3に示すようにエネルギー関数の他のパラメータを経験的に設定して実験を行った。

4.2節では、輝度の調整を行わずに画像を生成することで平滑化項の効果を明らかにする。4.3節では、輝度の調整によりエネルギー関数の最小化のみの場合に生じる自由視点画像中の不連続を低減できることを示す。

表 1 入力データの詳細

対象	画像サイズ (画素)	画像枚数
対象物体 A	1696 × 1280	61
対象物体 B	1920 × 1280	52
対象物体 C	1920 × 1280	51

表 2 実験で用いた PC の仕様

項目	詳細
CPU	Intel Corei7-4770 3.40GHz
メモリ	RAM 16.0GB
GPU	NVIDIA GeForce GTX 780
OS	Windows 8.1
グラフィック API	OpenGL

表 3 実験で設定した固定のパラメータ

パラメータ	値	式番号
θ_d	0.1	(5)
K_α	3000	(5)
θ_n	5°	(8)
輝度を調整するウィンドウサイズ	51 × 51	(9)



図 21 対象物体 A の入力画像例



図 22 対象物体 B の入力画像例



図 23 対象物体 C の入力画像例

4.2 小物体を対象とした平滑化項の効果の比較と考察

図 24, 図 25, および図 26 に, 各対象物体における CMPMVS [13] で生成した結果, 法線マップ, および形状の変化の大きさを表した画像を示す. また, 平滑化項の重みパラメータを様々に変化させたときの対象物体 A, B, C に対する自由視点画像生成結果をそれぞれ図 27-28, 図 29-30, 図 31-32 に示す.

対象物体 A では, $K_\beta=0$, $K_g=0$, すなわち, Debevec ら [24] と同様に仮想視点の入力視点の成す角のみに応じて入力画像のラベルを決定した場合, 入力視点の画素値や対象物体の形状とは無関係に直線的なラベル境界が得られた. そのため, 対象物体 A の首付近の黒と黄色の画素値が切り替わる領域や, 腹部の画素値が均一な領域で画素値の不連続が顕著である. 黒と黄色の画素値が切り替わる領域の不連続については, 入力視点のカメラパラメータ, もしくは対象物体の形状が正確に推定されないことにより生じたものである. また, 腹部の不連続は各入力視点で撮影したときの光源環境, もしくはカメラのパラメータの違いにより生じたものとする (図 33 参照).

次に, 提案手法において画素値の連続性を考慮する $g_{ij}(x_i, x_j)$ のみ用いた $K_g = 200, K_\beta = 0$ の場合では, 黒と黄色の画素値の切り替わり領域におけるずれや腹部の不連続が抑止された. 一方で, 対象物体の羽の付け根付近で羽と体の境界を横切るようなラベルの切り替わりによる画素値の不連続が新たに生じていることがわかる. 更に $g_{ij}(x_i, x_j)$ の効果を強めた $K_g = 400, K_\beta = 0$ では, 画素値の不連続な変化はほとんど確認できない. これらの入力視点のラベル画像から, ラベルの境界でも不連続が目立ちにくいと考えられる黒色の領域の内部や入力画像内において画素値が大きく変化する境界付近にラベル境界が存在する. 以上より, 自由視点画像内の隣接する画素における画素値の差に関する $g_{ij}(x_i, x_j)$ の効果が確認されたと考える (図 34 参照).

平滑化項 β_{ij} を導入した $K_g = 0, K_\beta = 200$, $K_g = 0, K_\beta = 400$, $K_g = 0, K_\beta = 600$ では, $g_{ij}(x_i, x_j)$ のみの場合と似たラベルが得られており, 模様や腹部の不連続は抑止された. 平滑化項のそれぞれの基準単独での効果が類似する理由として, 対象物体 A では羽と体の境界や首の周辺のように, 形状の変化と画素値の変化が同時に生じるためであると考えられる. 従って, これらの基準を組み合わせた場合に



図 24 各対象を CMPMVS で生成した結果

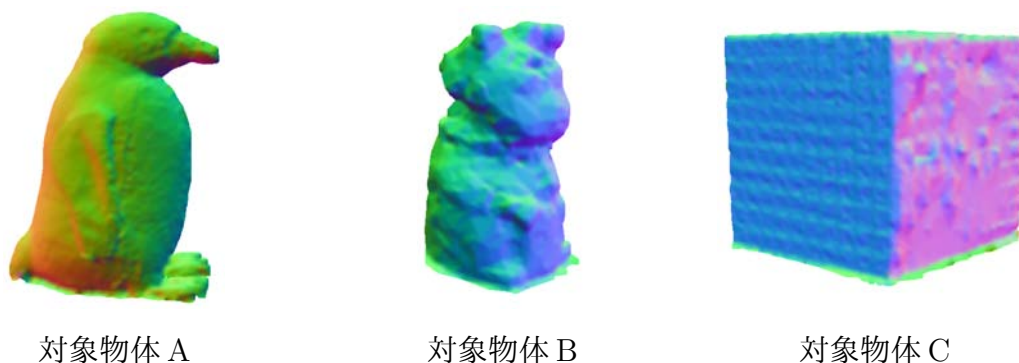


図 25 各対象の法線マップ

についても、それぞれが単独の場合と同様の結果となった (図 33 参照).

対象物体 B は鏡面反射のために Visual SFM [32, 33], と CMPMVS [13] でのカメラパラメータ推定, および形状復元の精度が低く, 平滑化項を用いない $K_g = 0, K_\beta = 0$ の場合に生成される自由視点画像では, 画素値の不連続が顕著である. $K_g = 0, K_\beta = 50$ や $K_g = 50, K_\beta = 0$ では平滑化項の効果が大きくなく, 平滑化項を用いない場合に準じるラベルが得られた. また, 対象物体 A と同様に二つの平滑化項は同じような働きをしており, 特に対象物体 B の場合, カメラパラメータ推定や形状復元の精度, さらに鏡面反射の問題から, 平滑化項はどちらも重みパラメータを大きくすると単一の入力画像を選択する傾向にある (図 35 参照). 対象物体 C は, 特にカメラのパラメータ推定に比較的大きな誤差がある入力視点付近の仮想視点において自由視点画像を生成した. 対象物体 C は格子模様のテク

スチャを持つため，平滑化項の重みパラメータをともに0にした場合にはカメラパラメータの誤差が自由視点画像に顕著に現れた．このとき， $g_{ij}(x_i, x_j)$ は格子模様が連続的に見えるようなラベルが選択された．一方， β_{ij} については，立方体の面の境界でテクスチャが切り替わった．この例では，カメラパラメータの誤差が大きいため，得られた自由視点画像の質は高くないものの，特に対象物体Cのような形状の物体に対する平滑化項の効果としては望ましいと考える (図 36 参照)．

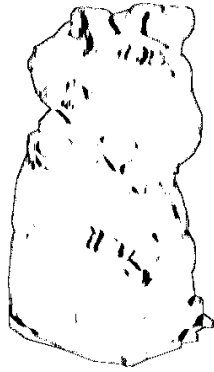
システム全体の計算時間は，各対象において出力画像を1枚生成するのに1分から2分程度掛かった．その計算時間の大部分 (約1分前後) を占める処理はオンライン処理中のグラフカットによるテクスチャ選択であるため，実時間でユーザへ出力画像を提示することは困難である．



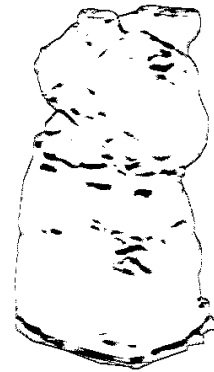
横方向の変化 (対象物体 A)



縦方向の変化 (対象物体 A)



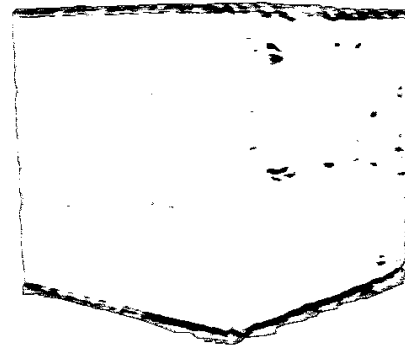
横方向の変化 (対象物体 B)



縦方向の変化 (対象物体 B)



横方向の変化 (対象物体 C)



縦方向の変化 (対象物体 C)

図 26 各対象における形状の変化の大きさ (黒 : 変化大, 白 : 変化小)

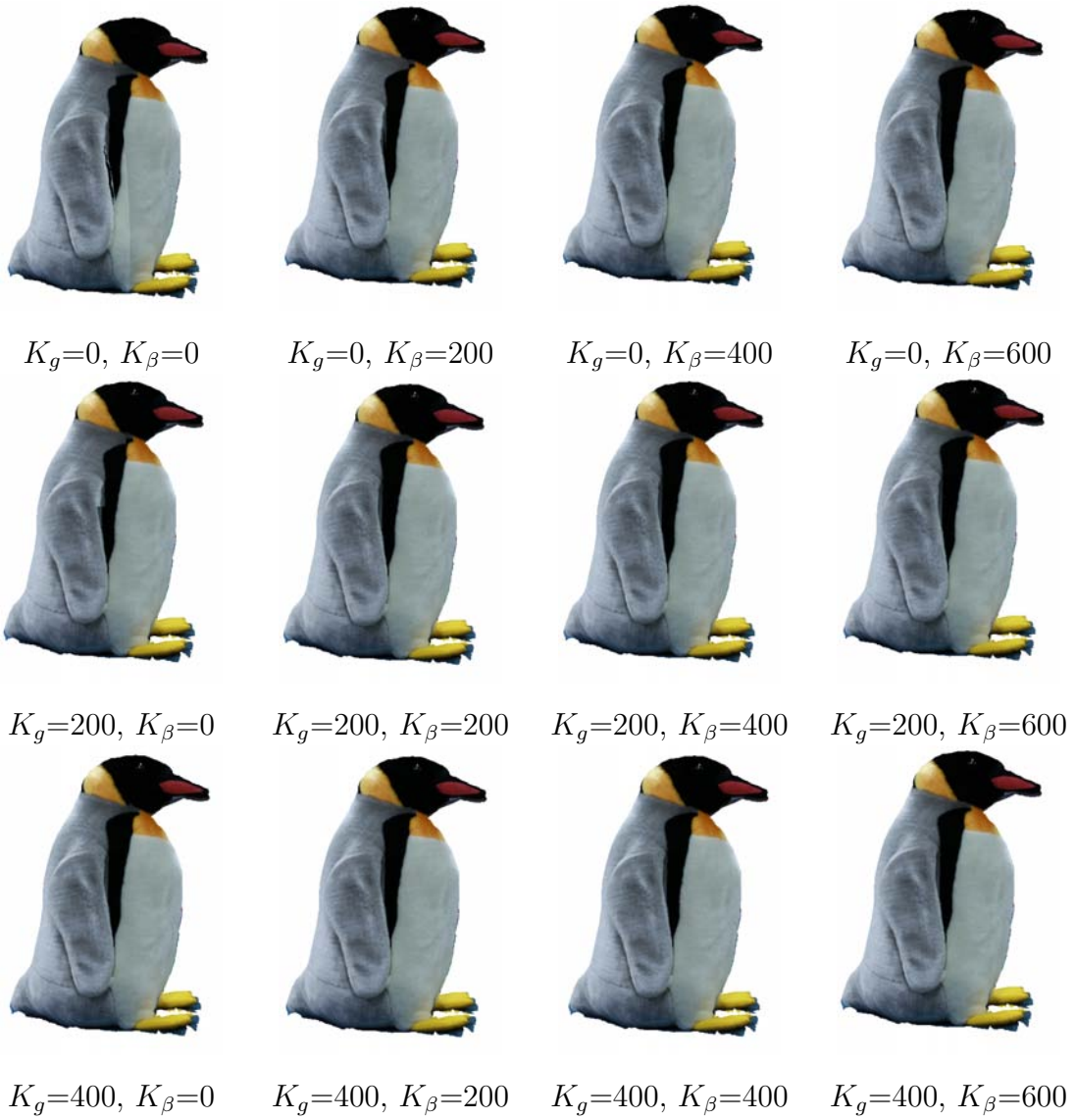


图 27 自由视点画像生成结果 (对象物体 A)

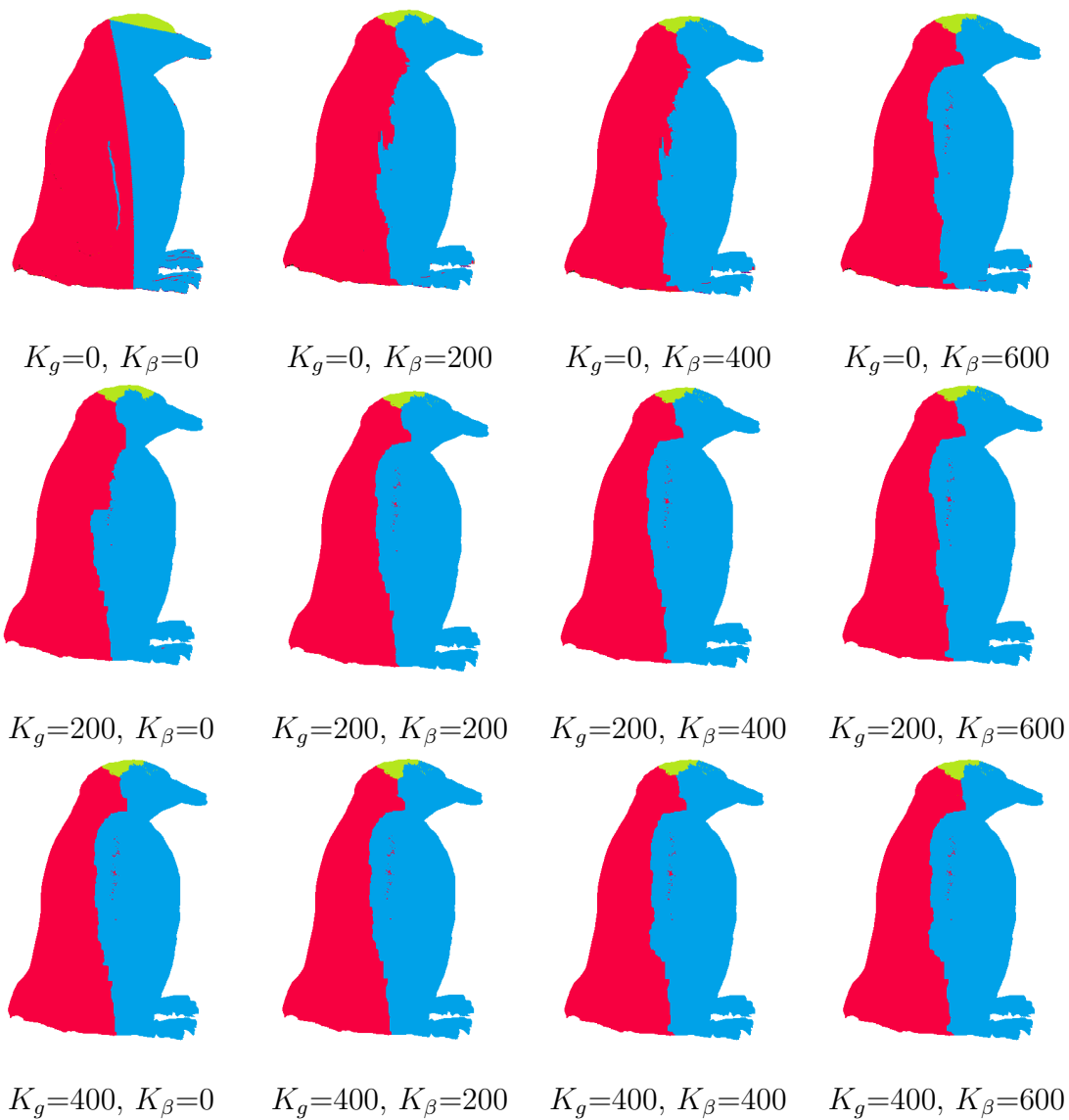


図 28 入力視点のラベル画像 (対象物体 A)



图 29 自由视点画像生成结果 (对象物体 B)

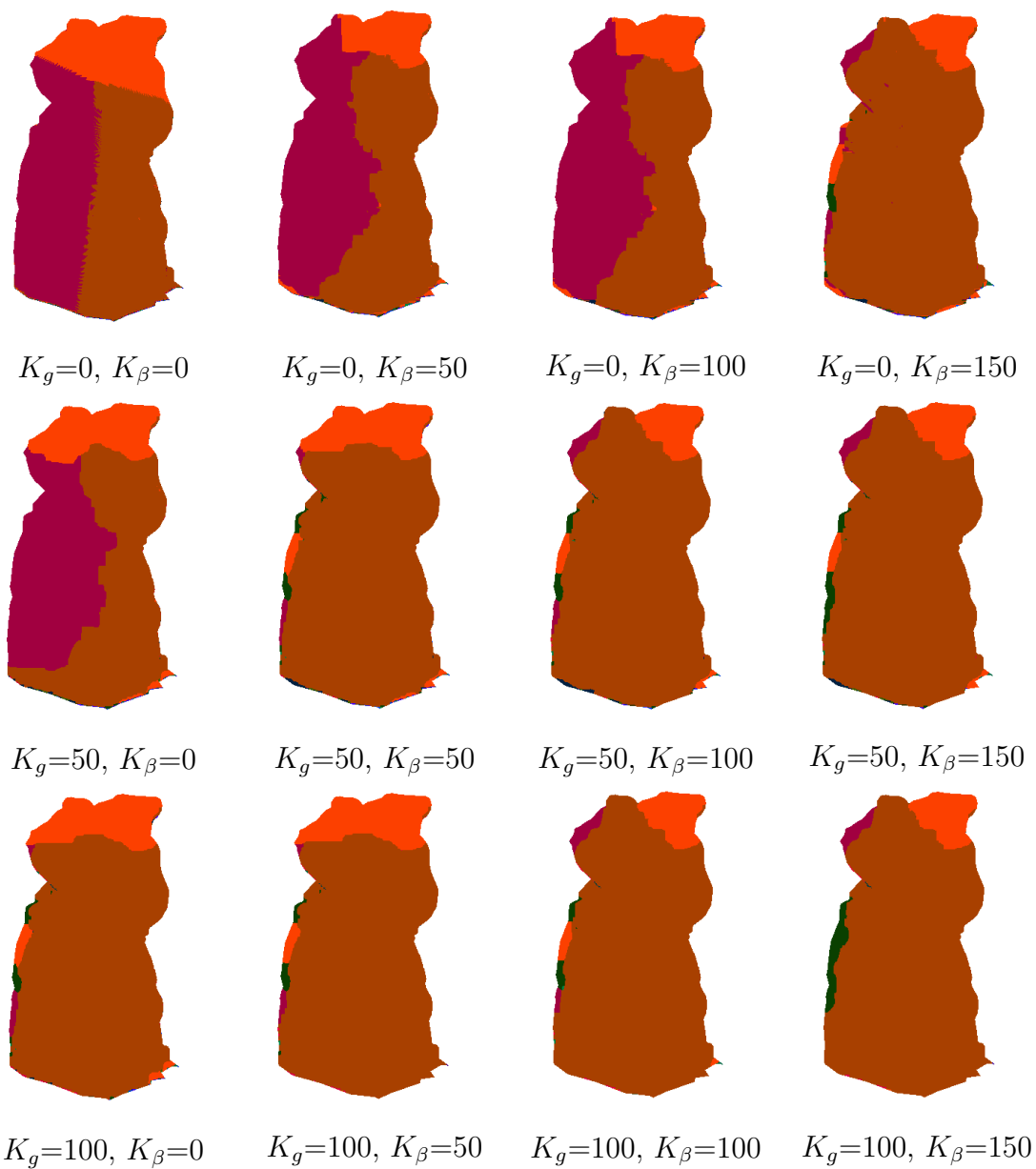


図 30 入力視点のラベル画像 (対象物体 B)

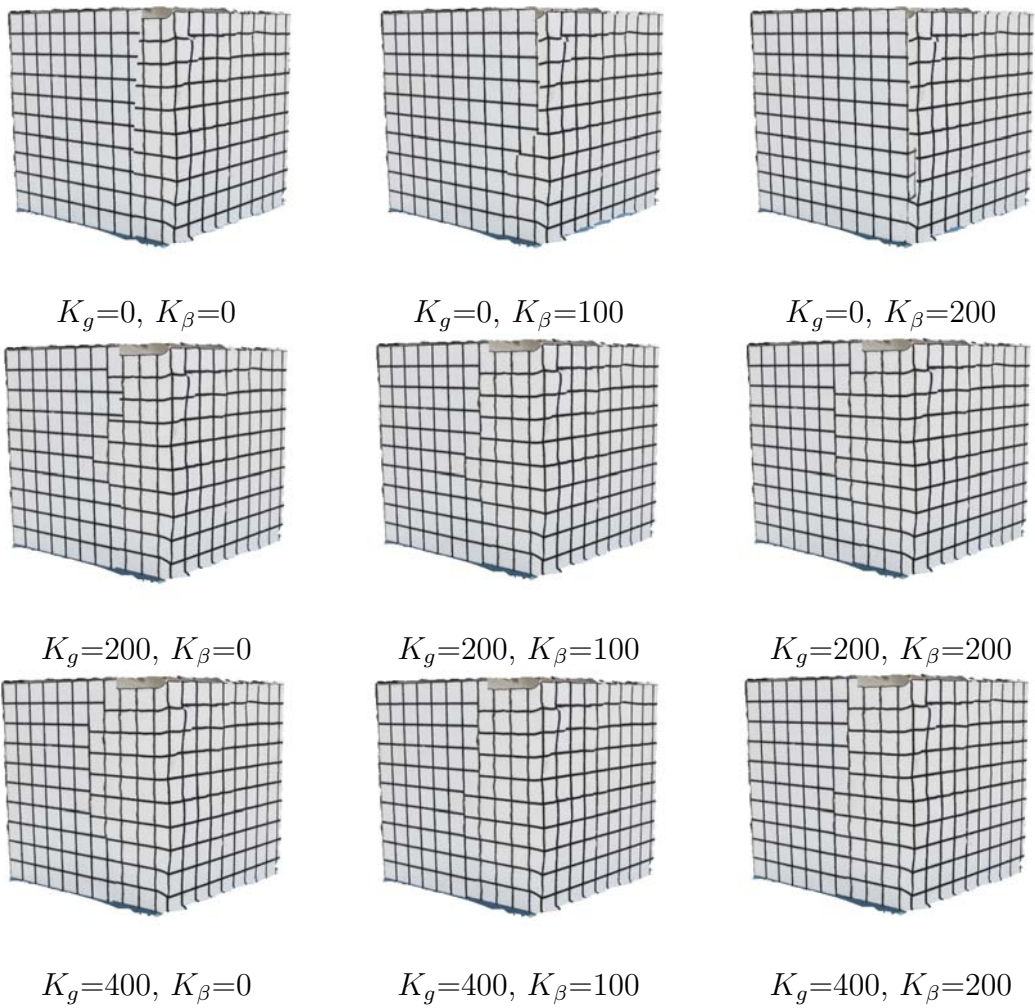


图 31 自由视点画像生成结果 (对象物体 C)

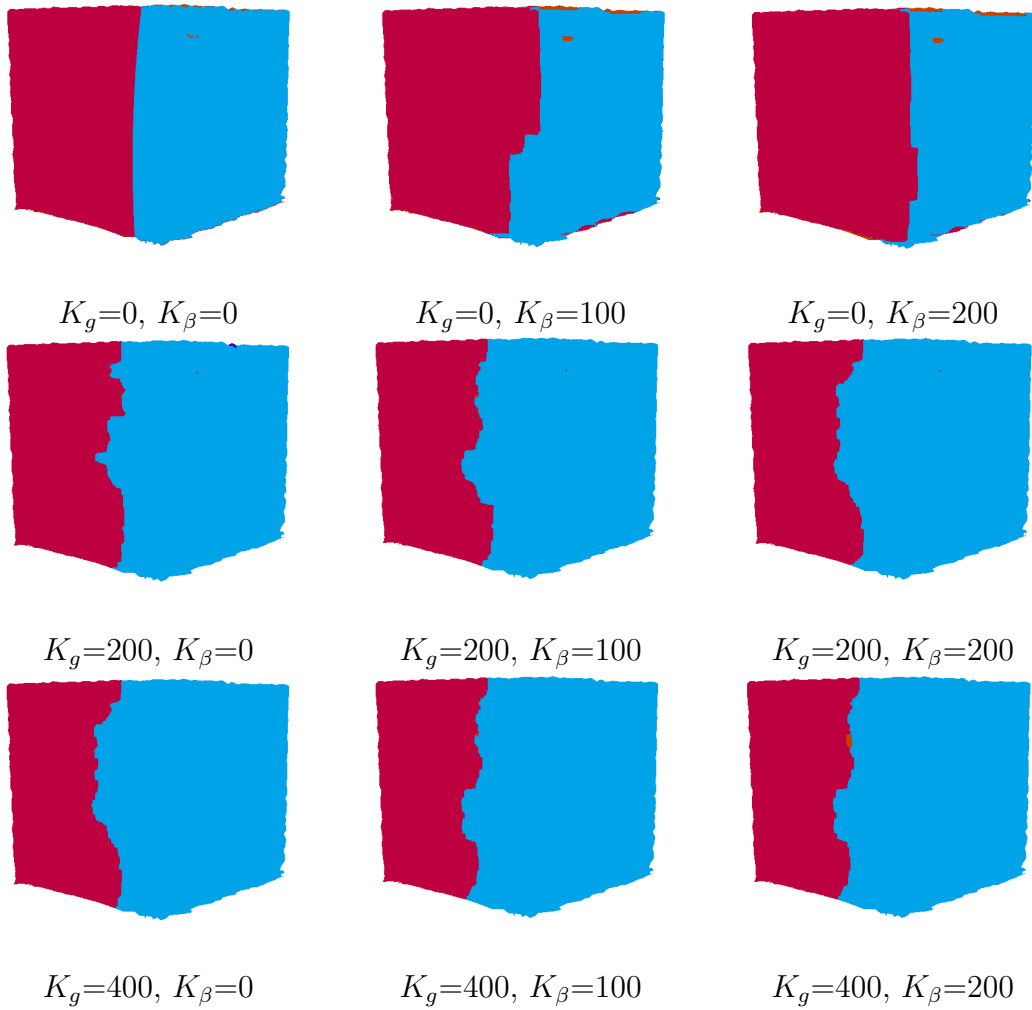


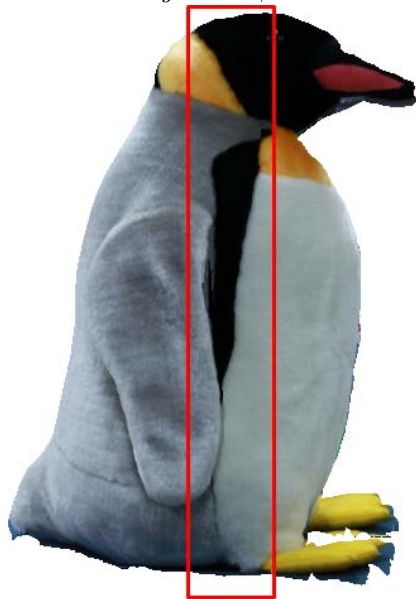
図 32 入力視点のラベル画像 (対象物体 C)



$K_g=0, K_\beta=0$



$K_g=0, K_\beta=200$



$K_g=0, K_\beta=400$



$K_g=0, K_\beta=600$

图 33 扩大画像 (对象物体 A) 1/2



$K_g=200, K_\beta=0$



$K_g=400, K_\beta=0$

图 34 扩大画像 (对象物体 A) 2/2



$K_g=0, K_\beta=0$



$K_g=0, K_\beta=50$



$K_g=50, K_\beta=0$



$K_g=100, K_\beta=150$

图 35 扩大画像 (对象物体 B)

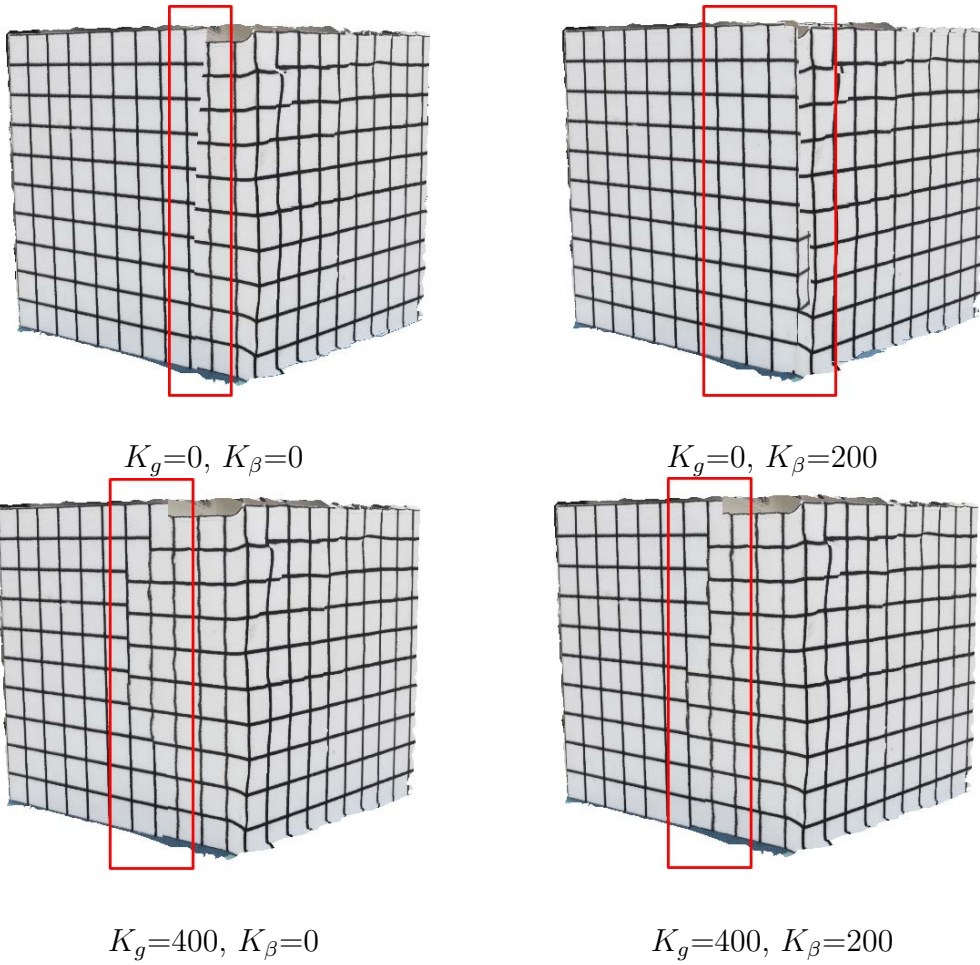


图 36 扩大画像 (对象物体 C)

4.3 輝度の調整に関する比較と考察

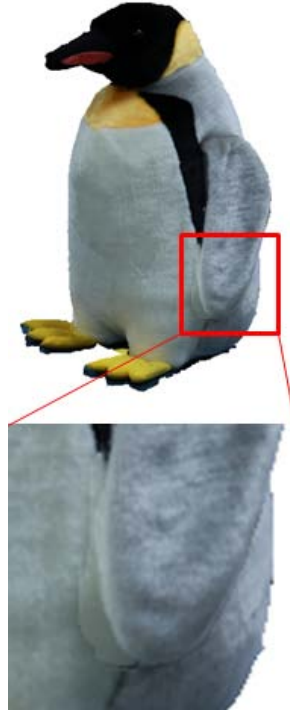
前節の結果から，例えば対象物体 A において光源環境などの違いによる不連続が生じることがわかる．この不連続性は，仮想視点の周囲の入力画像に依存するものの，平滑化項では抑止しきれない可能性がある．ここでは，提案手法の輝度調整がこの問題に有効であるかを検証する．

図 37 は対象物体 A の自由視点画像生成結果である．図 37(b) に示す画像では，輝度を調整しておらず，羽の下部の影の部分に不連続な画素値の変化が見られた．この原因として，緑と赤のラベルの入力画像が撮影された光源環境の違いから輝度の不連続が生じていると考えられる．また，黒色の領域がある羽の上部と比較して下部では羽と体の画素値の差が大きくないため，画素値の連続性を考慮した項 $g_{ij}(x_i, x_j)$ の効果が小さく不連続を抑止することができなかったと考えられる．

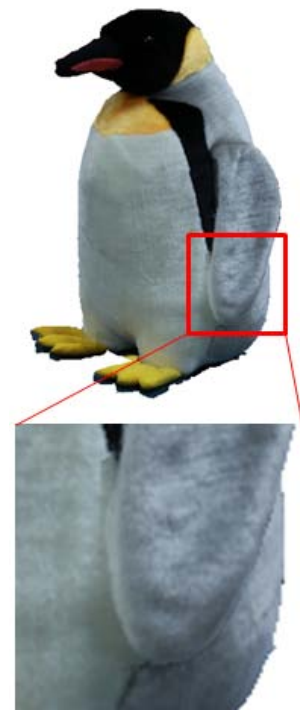
一方，図 37(c) に示す画像では，羽の下部の領域においてそれぞれのラベルの輝度値を基に，赤のラベルの輝度値を上げて緑のラベルの輝度値を下げることで，輝度が滑らかに変化するように不連続を解消した．しかし，拡大画像中の黒模様の下領域では，輝度調整することで黒いぼけが生じた．これは 3 次元形状やカメラパラメータに誤差があるため，図 38 に示すように青のラベルと緑のラベルの入力画像を投影したときにテクスチャの位置ずれが生じて，青と緑のラベルにおいて輝度調整する注目画素周辺の黒色の領域の割合が異なり，黒いぼけが生じていると考えられる．



(a) 入力視点を表すラベル画像

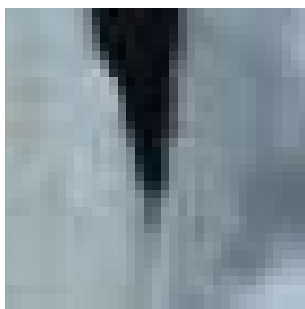


(b) 輝度調整無し



(c) 輝度調整有り

図 37 輝度調整



青のラベルのみ投影した画像



緑のラベルのみ投影した画像

図 38 テクスチャの位置ずれが生じた画像

5. まとめ

本研究では画素値の連続性を考慮した入力画像の選択により，異なる画像をブレンディングすることなく画素値の不連続を抑止する自由視点画像生成手法を提案した．提案手法では，生成される自由視点画像内の隣接する画素間の画素値の差分や3次元形状を考慮したエネルギー関数を設計し，この関数をグラフカットを用いて最小化することにより各画素に対して画素値を取得する入力画像を選択する．また，各画素に対して局所的な輝度調整を行い，エネルギー関数の最小化では除去できなかった輝度の不連続性を低減する．

実験では，現実空間の小物体を対象に提案手法の有効性を示すために，設計したエネルギー関数において，平滑化項の様々な重みパラメータで自由視点画像を生成してそれぞれを比較した．結果として，形状の復元やカメラパラメータ推定に大きな誤差がある場合には提案手法で生成される自由視点画像の質は高くはないものの，小さな誤差によるテクスチャのずれなどに対しては提案手法が有効であることが確認された．

提案手法では自然な自由視点画像の生成が可能である一方，1枚の自由視点画像の生成に数分かかることから，拡張現実感技術など実時間性が要求されるアプリケーションへの適用は困難である．そこで，今後の課題として処理時間の大部分を占めるグラフカットアルゴリズムによるエネルギー関数の最小化の高速化が挙げられる．

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ご指導，ご鞭撻頂い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。また，本学の研究におけるあらゆる面で，多大なご助言，ご協力を賜りましたことを深く感謝致します。そして，本研究の遂行にあたり，有益なご助言ご鞭撻を頂いた光メディアインタフェース研究室 向川康博 教授に厚くお礼申し上げます。更に，本研究の全過程を通して，始終温かいご指導をして頂い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。本研究でのご助言，ご協力を頂いた視覚情報メディア研究室 中島悠太 助教，河合紀彦 助教に深く感謝致します。また，研究室での生活を支えて頂いた視覚情報メディア研究室 石谷由美 女史に心より感謝致します。更に，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] 大倉史生, 神原誠之, 横矢直和. 空撮画像を用いた拡張テレプレゼンス-無人飛行船の自動操縦と全方位カメラによる AR システム-. 画像の認識・理解シンポジウム (MIRU2010), pp. 1183-1189, 2010.
- [2] M. Dellepiane, R. Marroquim, M. Callieri, P. Cignoni, and R. Scopigno. Flow-based local optimization for image-to-geometry projection. *IEEE Trans. on Visualization and Computer Graphics*, Vol. 18, No. 3, pp. 463-474, 2011.
- [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] 大池祐輔, 新宅宏彰, 池田誠, 浅田邦博. 光切断法による高精細・実時間3次元撮像システム. 映像情報メディア学会誌, Vol. 57, No. 9, pp. 1149-1151, 2003.
- [5] 西村正三. 3D レーザスキャナ. 文部省科学研究費補助金 研究成果報告書「光学的全視野計測技術による建設構造物の劣化・変状メカニズムの解明と診断法の開発」(代表: 松田浩, 課題番号: 17360217), 2007
- [6] 新宅宏彰. ハードウェア制御による高速三次元計測システムの構築と評価. 東京大学大学院 工学系研究科 電子工学専攻 修士論文, 2006.
- [7] H. Baker. Three-dimensional modeling. In *Proc. Int. Joint Conf. on Artificial Intelligence*, Vol. 2, pp. 649-655, 1977.
- [8] K. Kolev, T. Brox, and D. Cremers. Fast joint estimation of silhouettes and dense 3D geometry from multiple images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 3, pp. 493-505, 2012.
- [9] 池田聖. Shape from X. 奈良先端科学技術大学院大学 情報科学研究科インタラクティブメディア設計学研究室 計測情報処理 講義資料.

http://www.aist-nara.ac.jp/~takafumit/lecture/sip2_2011_ikeda_13.pdf

- [10] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *Int. Journal of Computer Vision*, Vol. 80, No. 2, pp. 189-210, 2008.
- [11] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. *Int. Journal of Computer Vision*, Vol. 38, No. 3, pp. 199-218, 2000.
- [12] Y. Furukawa and J. Ponce. Accurate, dense, and robust multi-view stereopsis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 8, pp. 1362-1376, 2009.
- [13] M. Jancosek and T. Pajdla. Multi-view reconstruction preserving weakly supported surfaces. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 3121-3128, 2011.
- [14] S. M. Seitz and C. R. Dyer. View morphing. In *Proc. ACM SIGGRAPH*, pp. 21-30, 1996.
- [15] J. Xiao and M. Shah. From images to video: view morphing of three images. In *Proc. Vision, Modeling and Visualization*, pp. 495-502, 2003.
- [16] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, M. Landy and J. A. Movshon, eds., MIT Press, pp. 3-20, 1991.
- [17] L. McMillan and G. Bishop. Plenoptic modeling: an image-based rendering system. In *Proc. ACM SIGGRAPH*, pp. 39-46, 1995.
- [18] T. Naemura, T. Takano, M. Kaneko, and H. Harashima. Ray-based creation of photo-realistic virtual world. In *Proc. Int. Conf. on Virtual Systems and MultiMedia*, pp. 59-68, 1997.

- [19] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. ACM SIGGRAPH*, pp. 31-42, 1996.
- [20] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proc. ACM SIGGRAPH*, pp. 43-54, 1996.
- [21] 蚊野浩. ライトフィールドカメラ Lytro の動作原理とアルゴリズム. 光アイアンス, Vol. 24, No.7, pp. 1-6, 2013.
- [22] A. Davis, M. Levoy, and F. Durand. Unstructured light fields. *Journal of Computer Graphics Forum*, Vol. 31, No. 2, pp. 305-314, 2012.
- [23] 九州大学 大学院システム情報科学研究所 情報知能工学部門 イメージ・メディア理解研究室. ライトフィールドビジョン.
<http://limu.ait.kyushu-u.ac.jp/project/project003.html>
- [24] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: a hybrid geometry-and image-based approach. In *Proc. ACM SIGGRAPH*, pp. 11-20, 1996.
- [25] 源田大輔, 向川康博, 尺長健. 視点と法線を組み合わせた任意視点映像の生成法. 情報処理学会研究報告, CVIM 2003-137-7, pp. 53-60, 2003.
- [26] 高井勇志, 延原章平, 田辺太一, 松山隆司, ウ小軍. 視点依存形状最適化による高精細自由視点画像生成. 画像の認識・理解シンポジウム (MIRU2012), 8 pages, 2012.
- [27] P. E. Debevec, Y. Yu, and G. Borshukov. Efficient view-dependent image-based rendering with projective texture-mapping. In *Proc. Eurographics Rendering Workshop*, pp. 105-116, 1998.
- [28] J. Kopf, M. F. Cohen, and R. Szeliski. First-person hyper-lapse videos. In *Proc. ACM Trans. on Graphics*, Vol. 33, No. 4, pp. 78:1-78:10, 2014.

- [29] G. Chaurasia, S. Duchne, O. S. Hornung, and G. Drettakis. Depth synthesis and local warps for plausible image-based navigation. *ACM Trans. on Graphics*, Vol. 32, No. 3, pp. 30:1-30:12, 2013.
- [30] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 11, pp. 1222-1239, 2001.
- [31] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [32] C. Wu. VisualSFM: A visual structure from motion system, 2011.
<http://www.cs.washington.edu/homes/ccwu/vsfm/>.
- [33] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz. Multicore bundle adjustment. *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 3057-3064, 2011.