

自由視点画像の合成画質を考慮した Visual Hull 高精度化方式

三功 浩嗣[†] 内藤 整[†]

本稿では、自由視点画像の合成画質向上を目指し、多視点画像から視体積交差法により復元される Visual Hull を高精度化する手法を提案する。視体積交差法は、高速かつ安定的に形状復元が可能である等の利点を持つが、被写体の凹領域を復元できないという原理的課題がある。本課題に対して、Visual Hull の表面上に存在するボクセルの視点間 Photo Consistency に基づき、形状の整形を行う space carving を始めとした Visual Hull 高精度化手法がいくつか提案されている。しかしながら、多くの手法は、各ボクセル単体の Photo Consistency のみを考慮して、当該ボクセルを除去するか否かを決定するため、最終的に獲得される Visual Hull の復元精度は、被写体のテクスチャ状態に大きく依存するという問題があった。以上の問題点を踏まえ、提案手法では、ボクセル空間における幾何的連結性を考慮するとともに、各カメラ視点において、整形過程の3次元形状モデルを用いて合成される自由視点画像と撮影画像間の差分を最小化する枠組みを導入する。提案手法の Visual Hull 高精度化に関する有効性を検証する目的で実施した、CG、および実写の多視点画像を用いた実験結果より、いずれのデータにおいても、従来手法に比べて、Visual Hull の凹領域をより適切に復元できることを確認した。

Precise Visual Hull Reconstruction Scheme Based on Quality Analysis of Free-viewpoint Image

Hiroshi Sankoh[†] and Sei Naito[†]

In this paper, we propose a refinement process for a visual hull reconstructed by shape from silhouettes algorithm, as the approach to improve the image quality of free viewpoint video. Shape from silhouettes algorithm reconstructs 3-dimensional model as a visual hull in a voxel space while maintaining a reasonable precision and a low complexity. However, it cannot reconstruct concavities. About this challenging problem, several methods which aim to refine a visual hull by considering photo consistency of each voxel locating close to the surface are proposed. However, most of them decide invalid voxels based on only the photo consistency, therefore the accuracy is highly dependent on texture of each viewpoint. To overcome the problem, we propose the refinement scheme for the visual hull reconstruction by utilizing geometric consistency as well as photo consistency. Moreover, the refinement framework of free viewpoint image based on the minimization of difference between synthesized image and captured

image for each viewpoint is introduced. Experimental results show the effectiveness of the proposed method from the comparison with the result by accurate 3d-model.

1. はじめに

近年、自由視点画像の合成技術が、超臨場感メディアの基盤技術として注目されている[1]。自由視点画像では、複数台のカメラで撮影された実写の多視点画像をもとに、単純にカメラ映像を切り替えるだけでなく、カメラを配置不可能な視点も含め、3次元空間中の任意視点での映像を再現・表示することを特徴とする。我々は特に、人物のように複雑な被写体が複数存在する場面を全周囲から撮影した多視点画像を対象に、通常はカメラを配置不可能な3次元空間中に入り込み、被写体群の間を通り抜けるウォークスルーや、上方から見下ろしたり下方から見上げたりといった、高さを自在に制御可能なフライスルーと呼ばれる視聴体験を映像合成によって実現することを目指している[2]。このような機能実現により、例えば、スポーツ中継において、特定の選手や審判の視点を再現することが可能となり、新たな視聴体験の提供につながると期待される。

自由視点画像の代表的な合成方式として、モデルベース方式[3]と画像ベース方式[4]の二つがあるが、後者の方式は主にカメラ間の内挿画像の合成を目的としたものであり、上記アプリケーションの実現を考えた場合は、一般的に前者の方式が用いられる。モデルベースの合成方式は、被写体の3次元形状モデルを復元し、実写の多視点画像をテクスチャとして、視点の位置に応じて重ね合わせの割合を調節することで、視点の移動に伴う見え方の変化を再現することが可能である。最終的に生成される自由視点画像の合成精度は、多視点の画像情報をもとに復元される被写体3次元形状モデルの復元精度に大きく左右される。

被写体の3次元形状モデルを復元する代表的な手法として、視体積交差法が広く用いられている[5]。視体積交差法は、多視点画像の各カメラにおいて、カメラの光学中心と被写体シルエット領域から形成される錐体領域を求め、全視点で得られる錐体の積集合領域を、3次元ボクセル空間中の Visual Hull として獲得する手法である。ステレオマッチング等の、カメラ間での対応点探索をもとに奥行きを推定する手法に比べ、被写体テクスチャの影響を受けにくい点や、アルゴリズムの高速性がメリットである。しかしながら、視体積交差法では、物体の曲面を正確に復元できないという原理的課題がある。これは、全視点において極めて正確なシルエット画像を抽出できたとしても、それらシルエット画像群に視体積交差法を適用した結果、復元される Visual Hull は真の被写体領域に外接する凸立体であり、背景領域を内包していることを意味する。

[†] 株式会社 KDDI 研究所
KDDI R&D Laboratories Inc.

上記課題に対して、Visual Hull の表面上に存在するボクセルの、視点間での Photo Consistency に基づき形状補正を行う space carving[6]を始め、厳密な 3 次元形状の復元を試みる手法がいくつか提案されている。これらの手法は、ボクセルを削るか否かの判定基準として、単一のボクセルや画素を用いており、ボクセル空間や撮影カメラ画像中での連続性を考慮していないため、表面形状の凹凸が目立つ等、復元精度の限界があった。

以上を踏まえ、本稿では、被写体の真の 3 次元形状復元を目的として、視体積交差法により復元される Visual Hull を、ボクセル空間中の連続性や、撮影カメラ画像中でのテクスチャ状態を十分に考慮した上で高精度化する手法を提案する。具体的には、Visual Hull 表面近傍のボクセルに対して、最寄りのカメラで取得される色情報をもとに、オブジェクトらしさを算出し、3 次元空間中でのエネルギー最小化を目的としたグラフカットを繰り返し適用する。さらに、グラフカットを繰り返し適用した結果得られる 3 次元形状モデルをもとに、各カメラ位置に対応する合成画像と、当該カメラ位置で撮影される実写画像をもとに、画像中でグラフカットを適用し、除去すべきと判断される画素の光線を探索し、3 次元形状モデルとの交点ボクセルを除去することで、高精度化を行う。

以下、2 章で関連研究とその問題点について述べ、3 章では Visual Hull 高精度化のための提案手法について詳細を述べる。4 章で CG、および実写の多視点画像を用いた実験結果を踏まえ、関連研究との性能比較を行い、5 章でまとめと今後の課題について記述する。

2. 関連研究

Visual Hull 高精度化に関する研究として、Photo Consistency のみならず、視体積交差法のシルエット拘束や、Visual Hull 表面形状の安定性を考慮したエネルギー関数を定義し、最適化の枠組みにより形状を補正する手法が提案されている[7][8]。文献[7]では、Visual Hull を離散マーチングキューブ法によりメッシュデータに変換し、表面モデル化することで、形状補正をモデルフィッティングによる復元として扱っている。各頂点の Photo Consistency に基づく外力だけでなく、変形モデルが各視点のシルエットに一致するよう補正するシルエット保持力を導入することで、復元形状の安定性確保と高精度化の両立に成功している。しかしながら、変形過程の 3 次元形状モデルのテクスチャ状態を、各視点の撮影画像と比較して評価する処理が存在しないため、最終的に復元される 3 次元形状モデルをもとに合成される自由視点画像に十分な画質が保証されないという問題が残っている。一方、文献[8]では、ステレオマッチングによる整形手法が提案されている。当該手法は、Photo Consistency によりカメラから被写体までの距離値を Visual Hull 内側へ押し込むための外力と、

Visual Hull 表面の局所的な形状特徴を維持するための内力の線形結合により評価関数を定義し、最小化する距離値を求めることで、被写体の表面形状を決定する。Space carving 等の単一ボクセルに注目する手法に比べて、表面上に不自然な凹凸が発生することが少なく、滑らかな形状を復元できる結果が示されている。しかしながら、最終的な復元結果は、ステレオマッチングの探索精度に大きく依存するため、被写体のテクスチャ変化が少ない領域では十分な復元精度が得られないという問題が残っている。

一方、ボクセル空間中での隣接関係を考慮したグラフカットにより Visual Hull の整形を行う手法が提案されている[9]。当該手法では、Visual Hull 表面近傍のボクセルに関して、ボクセル単体の Photo Consistency に基づくエネルギーと、ボクセル間の隣接関係に基づくエネルギー関数を定義し、グラフカットを適用することで、滑らかな形状を復元できることを示している。さらに、Visual Hull の内側に、被写体が確実に存在すると仮定する Core 領域を設け、グラフカットの適用範囲を Visual Hull 表面と Core 領域間に限定することで、被写体領域が誤って削られる可能性を低減することに成功している。しかしながら、ボクセル間の隣接関係に基づくエネルギー関数が、単純にボクセル単体のエネルギー値の平均で定義されているため、ボクセル空間中の隣接関係を十分に考慮できていないとは言えない。また、グラフカットの適用が一回に限定されているのに加え、適用結果を評価するプロセスが存在しないため、最終的に復元される 3 次元形状モデルに、背景領域が含まれる可能性が高いと考えられる。

3. 提案手法

本稿では、特に文献[9]の課題に注目し、ボクセル空間中での連続性を十分に考慮するとともに、整形過程の 3 次元形状モデルをもとに、実カメラ視点で生成される自由視点画像の画質を評価するプロセスを導入することで、3 次元形状モデルのテクスチャ状態を考慮した Visual Hull 高精度化手法を提案する。

提案手法は、ボクセル空間中での連続性を考慮した整形、および各視点での自由視点画像の画質を考慮した整形を特徴とする。提案手法のプロチャートを図 1 に示す。ある時刻 1 フレーム分の多視点画像と、各視点のカメラパラメータ、および視体積交差法で復元される Visual Hull を入力として、最終的に整形済みの 3 次元形状モデルを出力し、処理を終了する。

3.1 各ボクセルのオブジェクトらしさ

視体積交差法で復元される Visual Hull 表面近傍の各ボクセル v_i について、Photo Consistency に基づくオブジェクトらしさを、カメラ間における画素値の分散として導入する。 v_i の各視点 n ($n = 1, \dots, N$) への投影画素値を RGB 色空間中の多次元ベ

クトル $\mathbf{x}(v_i^{(n)})$ として扱い、 v_i を観測している撮影視点の集合を $n_{vis}(v_i)$ とする。このとき、 v_i の色彩に関する平均ベクトル $\mathbf{u}(v_i)$ と分散ベクトル $\sigma_c^2(v_i)$ をそれぞれ式(1)、および式(2)により算出する。ここで、 $\#n_{vis}(v_i)$ と $c \in (r, g, b)$ はそれぞれ、 v_i の観測視点数、および色成分を表す。また、 $\mathbf{x}_c(v_i^{(n)})$ 、 $\mathbf{u}_c(v_i)$ 、 $\sigma_c(v_i)$ はいずれも色成分ごとのスカラー値を与えるものとする。

$$\mathbf{u}(v_i) = \frac{1}{\#n_{vis}(v_i)} \sum_{n \in n_{vis}(v_i)} \mathbf{x}(v_i^{(n)}). \quad (1)$$

$$\sigma_c^2(v_i) = \frac{1}{\#n_{vis}(v_i)} \sum_{n \in n_{vis}(v_i)} (\mathbf{x}_c(v_i^{(n)}) - \mathbf{u}_c(v_i))^2. \quad (2)$$

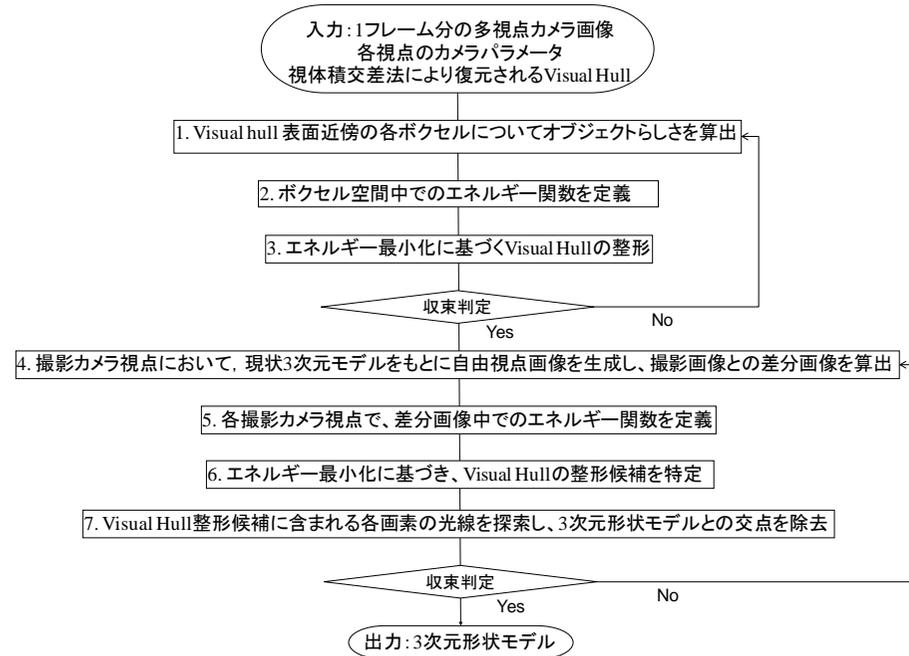


図 1. 提案手法のプロチャート

3.2 ボクセル空間中でのエネルギー関数

視体積交差法で復元される Visual Hull の高精度化は、Visual Hull 表面近傍に存在する各ボクセルを、そのオブジェクトらしさに基づき、背景領域または被写体領域に再分類する 2 値化問題であると捉えることができる。本稿では、グラフカットを用いるエネルギー最小化の枠組みにより 2 値化を行う方法について述べる。

ボクセル空間中でのエネルギー関数を、各ボクセルのオブジェクトらしさに基づき、定数 λ_v を用いて以下のように定義する。

$$E(\mathbf{v}; \mathbf{a}_v) = U(\mathbf{v}; \mathbf{a}_v) + \lambda_v V(\mathbf{v}; \mathbf{a}_v). \quad (3)$$

ここで、 $\mathbf{v} = (v_1, \dots, v_i, \dots)$ はボクセル空間中の全ボクセルを表し、 $\mathbf{a}_v = (\alpha_{v_1}, \dots, \alpha_{v_i}, \dots)$ は各ボクセルを背景領域または被写体領域のいずれに割り当てるかを 0 または 1 で表している。また、 $U(\mathbf{v}; \mathbf{a}_v)$ は各ボクセル単体のオブジェクトらしさに依存するデータ項であり、式(4)、および式(5)により算出される。

$$U(\mathbf{v}; \mathbf{a}_v) = \sum_i U(v_i; \alpha_{v_i}) \quad (4)$$

$$U(v_i; \alpha_{v_i}) = \begin{cases} \min_{c \in (r, g, b)} [\max[(th_{v(c)} - \sigma_c^2(v_i)), 0]] & (\alpha_{v_i} = 0) \\ \min_{c \in (r, g, b)} [\sigma_c^2(v_i)] & (\alpha_{v_i} = 1). \end{cases} \quad (5)$$

ここで、 $th_{v(c)}$ は定数であり、ボクセル単体を処理対象とした場合の閾値に相当する。一方、ボクセル空間中の Visual Hull 表面近傍の外側 (out_v)、および内側 (in_v) に存在するボクセルについては、それぞれ背景領域、および被写体領域とみなすための強制項を以下のように導入する。ただし、Visual Hull 表面近傍は、表面上ボクセル群、および表面から深さ D に存在するボクセル群で囲まれる閉領域として定義する。

$$U(v_i \in out_v; \alpha_{v_i}) = \begin{cases} 0 & (\alpha_{v_i} = 0) \\ \infty & (\alpha_{v_i} = 1). \end{cases} \quad (6)$$

$$U(v_i \in in_v; \alpha_{v_i}) = \begin{cases} \infty & (\alpha_{v_i} = 0) \\ 0 & (\alpha_{v_i} = 1). \end{cases} \quad (7)$$

一方、 $V(\mathbf{v}; \mathbf{a}_v)$ は隣接ボクセル間における平均色ベクトルの差に依存する平滑化項であり、隣接ボクセル間で Gibbs モデルを仮定し、式(8)、および式(9)により算出される。

$$V(\mathbf{v}; \boldsymbol{\alpha}_v) = \sum_{(i,j) \in N_v} V(v_i, v_j; \alpha_{v_i}, \alpha_{v_j}) \quad (8)$$

$$V(v_i, v_j; \alpha_{v_i}, \alpha_{v_j}) = \begin{cases} \frac{\exp(-\kappa_v |\mathbf{u}(v_i) - \mathbf{u}(v_j)|^2)}{\text{dis}_v(i, j)} & (\alpha_{v_i} \neq \alpha_{v_j}) \\ 0 & (\alpha_{v_i} = \alpha_{v_j}). \end{cases} \quad (9)$$

ここで、 N_v は、ボクセル空間において、近傍 26 ボクセルでの隣接関係をもつ全組合せであり、 $\text{dis}_v(\cdot)$ は隣接ボクセル間のユークリッド距離を表す。 κ_v は定数であり、 N_v に関する期待値の演算子 $\langle \cdot \rangle$ を用いて以下のように算出する。

$$\kappa_v = (2 \langle |\mathbf{u}(v_i) - \mathbf{u}(v_j)|^2 \rangle)^{-1}. \quad (10)$$

エネルギー関数 $E(\mathbf{v}; \boldsymbol{\alpha}_v)$ をグラフカットにより最小化することで、Visual Hull 表面近傍を整形する。ここで、Visual Hull 表面近傍を、表面からの深さに基づき限定することで、グラフカットの適用により削られる領域を制限することができる。その結果、整形を一回のみ適用するのではなく、整形後の 3 次元形状に基づきエネルギー関数 $E(\mathbf{v}; \boldsymbol{\alpha}_v)$ を再計算し、グラフカットによる最小化を実施する過程を繰り返し適用することで、より高精度な 3 次元形状の復元が可能になると考えられる。

3.3 撮影カメラ視点でのエネルギー関数

ボクセル空間中での連続性を考慮して整形された 3 次元形状モデルを用いて、実カメラ視点における自由視点画像を生成し、撮影画像との差分に基づき、3 次元形状モデルを再度整形する手法について述べる。実カメラ視点の中心射影行列をもとに、各画素の光線を探査し、3 次元形状モデルとの交点を探索する。当該交点を、観測している撮影視点のうち、左右それぞれで最も近い視点に投影し、投影画素ベクトルを成分ごとにブレンドすることで、合成画像を生成する。各カメラ視点において、合成画像と、撮影画像との差分画像を生成し、画像中の隣接関係を考慮したエネルギー関数を、ボクセル空間中でのそれと同様に、各画素単体の差分値に依存するデータ項 $U(\mathbf{p}; \boldsymbol{\alpha}_p)$ 、隣接画素間における差分ベクトルの差に依存する平滑化項 $V(\mathbf{p}; \boldsymbol{\alpha}_p)$ 、および定数 λ_p により、以下のように定義する。

$$E(\mathbf{p}; \boldsymbol{\alpha}_p) = U(\mathbf{p}; \boldsymbol{\alpha}_p) + \lambda_p V(\mathbf{p}; \boldsymbol{\alpha}_p). \quad (11)$$

$$U(\mathbf{p}; \boldsymbol{\alpha}_p) = \sum_i U(p_i; \alpha_{p_i}) \quad (12)$$

$$U(p_i; \alpha_{p_i}) = \begin{cases} \min_{c \in (r, g, b)} [\max[(th_{p(c)} - \mathbf{x}_c(p_i)), 0]] & (\alpha_{p_i} = 0) \\ \min_{c \in (r, g, b)} [\mathbf{x}_c(p_i)] & (\alpha_{p_i} = 1). \end{cases} \quad (13)$$

$$V(\mathbf{p}; \boldsymbol{\alpha}_p) = \sum_{(i,j) \in N_p} V(p_i, p_j; \alpha_{p_i}, \alpha_{p_j}) \quad (14)$$

$$V(p_i, p_j; \alpha_{p_i}, \alpha_{p_j}) = \begin{cases} \frac{\exp(-\kappa_p |\mathbf{x}(p_i) - \mathbf{x}(p_j)|^2)}{\text{dis}_p(i, j)} & (\alpha_{p_i} \neq \alpha_{p_j}) \\ 0 & (\alpha_{p_i} = \alpha_{p_j}). \end{cases} \quad (15)$$

ここで、 $\mathbf{p} = (p_1, \dots, p_i, \dots)$ は画像中の全画素を表し、 $\boldsymbol{\alpha}_p = (\alpha_{p_1}, \dots, \alpha_{p_i}, \dots)$ は各画素を背景領域または被写体領域のいずれに割り当てるかを 0 または 1 で表している。また、式(13)の $th_{p(c)}$ は定数であり、画素単体を処理対象とした場合の閾値に相当する。式(14)の N_p は、画像中において、近傍 8 画素での隣接関係をもつ全組合せであり、式(15)の $\text{dis}_p(\cdot)$ は隣接画素間のユークリッド距離を表す。 κ_p は定数であり、 N_p に関する期待値の演算子 $\langle \cdot \rangle$ を用いて以下のように算出する。

$$\kappa_p = (2 \langle |\mathbf{x}(p_i) - \mathbf{x}(p_j)|^2 \rangle)^{-1}. \quad (16)$$

一方、画像中で 3 次元形状モデル投影領域の外側 (out_p) に存在する画素について、背景領域とみなすための強制項を以下のように導入する。

$$U(p_i \in out_p; \alpha_{p_i}) = \begin{cases} 0 & (\alpha_{p_i} = 0) \\ \infty & (\alpha_{p_i} = 1). \end{cases} \quad (17)$$

各撮影視点において、エネルギー関数 $E(\mathbf{p}; \boldsymbol{\alpha}_p)$ をグラフカットにより最小化した結果、背景領域と見なされる画像領域を、3 次元形状モデルの整形候補とする。当該画像領域に含まれる各画素の光線と、3 次元形状モデルとの交点を探索し、当

該交点に相当するボクセルを除去することで、3次元形状モデルの整形を行う。各視点での整形結果をもとに、別視点でのエネルギー関数を再度計算し、グラフカットによる最小化を実施する過程を繰り返し適用することで、自由視点画質を考慮した、高精度な3次元形状の復元が可能になると考えられる。

4. 実験結果

提案手法の Visual Hull 高精度化に関する有効性を検証する目的で、CG、および実写の多視点画像を対象に実験を行った。CG データの入力について、立方体の各面に凹領域を含む3次元形状モデルにランダムカラーのテクスチャをマッピングしたモデル情報を構築し、様々なアングルを含む23視点で取得される多視点画像、およびカメラパラメータを用意した。実写データの入力については、水平方向11台のカメラで同期撮影した多視点映像の1フレームについて、CG データと同様の形式で用意した。実写データにおけるカメラパラメータは、全てのカメラで共通のキャリブレーションパターン(特徴点を配置したチェッカーボード)を撮影しておき、Zhang のアルゴリズムを用いて算出した。各入力データについて、ある視点での撮影画像、および視体積交差法を適用した結果、復元される Visual Hull をある視点から観測した画像のセットを図2、および図3に示す。図2、図3はそれぞれCG データ、実写データに対応する。いずれのデータにおいても、復元される Visual Hull が凹領域を再現できていないことがわかる。

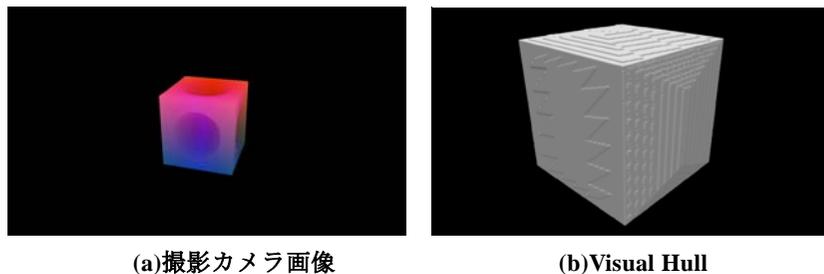
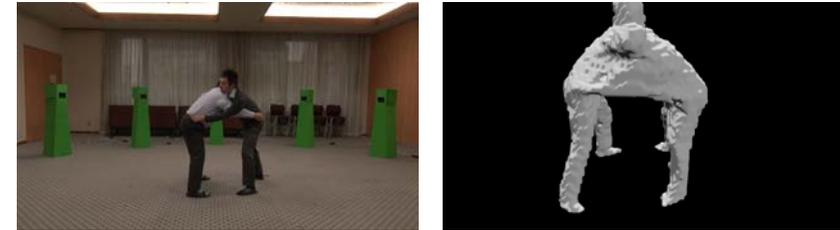


図2. 多視点画像と Visual Hull の例 (CG データ)

Visual Hull 高精度化の比較方式として、ボクセル単体のオブジェクトらしさに基づき整形する space carving[6]、およびボクセル空間中での隣接関係を考慮して整形を行う Hisatomi らの方式[9]に注目し、同様の実験を実施した。



(a)撮影カメラ画像 (b)Visual Hull

図3. 多視点画像と Visual Hull の例 (実写データ)

4.1 CGデータに対する性能評価

CG データを対象に提案手法を適用した結果、復元された3次元形状を、ある視点から観測した画像、および当該視点で合成される自由視点画像のセットを図4に示す。ボクセル空間中でのエネルギー関数、および撮影カメラ視点でのエネルギー関数に関わるパラメータは、予備実験の結果に基づき、データ項と平滑化項のバランス等を考慮して、表1のように設定した。

同様に、space carving、および Hisatomi らの方式を適用した結果を、それぞれ図5、図6に示す。また、オリジナル CG モデルを同じ位置から観測した画像を図7に示す。以上の結果より、提案手法の結果は、表面上の凹凸が含まれるものの、オリジナル CG モデルに最も近いことが確認できた。Space carving では、ボクセル単体のオブジェクトらしさに基づき、閾値により削るか否かを判定するため、被写体領域が誤って削られる領域や、背景領域が削られずに残る領域が、少なからず存在することを確認した。Hisatomi らの方式では、ボクセル空間中での隣接関係を考慮しているものの、グラフカットによる整形処理を一度適用するのみであり、整形後の3次元形状を評価するプロセスが存在しないことから、十分な精度が得られていないと考えられる。

さらに各方式により得られる3次元形状モデルについて、オリジナル CD モデルの結果との比較を行い、Precision/Recall/F-measure の算出を行った。結果を表2に示す。提案手法では、Recall を保持しつつ、高い Precision を達成しており、安定して3次元形状の整形ができたことを示している。



(a)3次元形状モデル

(b)自由視点画像

図 4. 提案手法を適用した結果 (CG データ)

表 1. 提案手法のパラメータ (CG データ)

λ_v	$th_{v(c)}$	D	λ_p	$Th_{p(c)}$
0.05	0.01	3	0.1	0.1



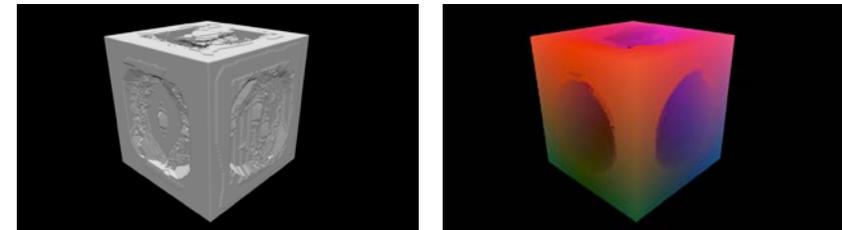
(a)3次元形状モデル

(b)自由視点画像

図 5. Space carving を適用した結果 (CG データ)

表 2. 定量評価の比較

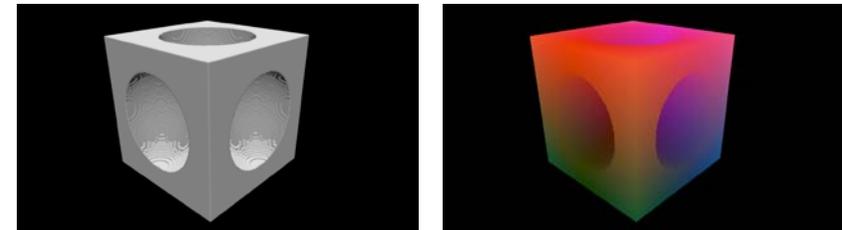
	Recall	Precision	F-measure
提案手法	0.992735	0.855730	0.919155
Space carving	0.889837	0.887012	0.888422
Hisatomi らの方式	0.976019	0.673336	0.796904



(a)3次元形状モデル

(b)自由視点画像

図 6. Hisatomi らの方式を適用した結果 (CG データ)



(a)3次元形状モデル

(b)自由視点画像

図 7. オリジナル CG モデル

4.2 実写データに対する性能評価

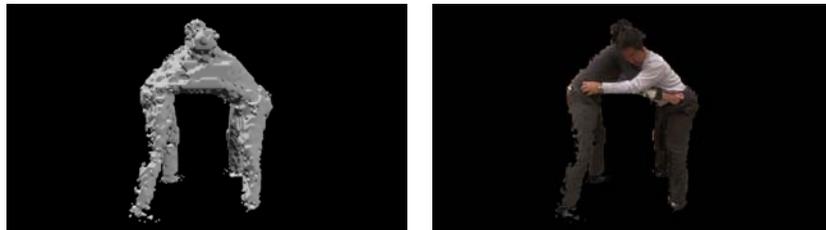
実写データを対象に提案手法を適用した結果、復元された 3 次元形状を、ある視点から観測した画像、および当該視点で合成される自由視点画像のセットを図 8 に示す。ボクセル空間中でのエネルギー関数、および撮影カメラ視点でのエネルギー関数に関わるパラメータは、予備実験の結果に基づき、データ項と平滑化項のパラメータ等を考慮して、表 3 のように設定した。同様に、space carving、および Hisatomi らの方式を適用した結果を、それぞれ図 9、図 10 に示す。図 8 より、提案手法では、被写体領域が削られることなく、被写体の腕が重なる凹領域を復元できていることが確認できる。図 9、および図 10 より、従来手法では、被写体が誤って削られる領域が存在すること、復元できない凹領域が残っていることを確認した。



(a)3次元形状モデル

(b)自由視点画像

図 8. 提案手法を適用した結果 (実写データ)



(a)3次元形状モデル

(b)自由視点画像

図 9. Space carving を適用した結果 (実写データ)



(a)3次元形状モデル

(b)自由視点画像

図 10. Hisatomi らの方式を適用した結果 (実写データ)

表 3. 提案手法のパラメータ (実写データ)

λ_v	$th_{v(c)}$	D	λ_p	$Th_{p(c)}$
0.1	0.1	1	0.1	0.1

4.3 考察

提案手法が、従来手法に比べて、被写体領域が誤って削られる問題の低減、および被写体凹領域の復元精度向上を実現できた要因として、ボクセル空間中でのエネルギー関数最小化に基づく整形処理と、撮影カメラ視点でのエネルギー関数最小化に基づく整形処理の二つの処理が有効に働いた結果であると考えられる。

前者の処理では、ボクセル空間中での隣接関係を考慮し、ボクセルを削り込む領域を限定しながら整形処理を繰り返し適用することで、被写体領域が誤って削られる可能性を低減しつつ、被写体凹領域の復元を進めることができると考えられる。しかしながら、3次元ボクセル空間の解像度が撮影画像の解像度に比べて十分ではないこと等から、各ボクセルのオブジェクトらしさを正確に算出することは難しい。特に、実写データでは、各撮影カメラ視点のカメラパラメータ算出や、撮影カメラ間の色補正等を厳密に実施することが困難であり、カメラ間での Photo Consistency に基づくオブジェクトらしさを正しく算出することは現実的ではない。そのため、ボクセル空間中での整形処理のみでは、復元精度の向上に限界がある。

この問題に対して、後者の処理では、各撮影カメラ視点において、整形過程の3次元形状モデルをもとに生成される自由視点画像と撮影カメラ画像の差分に注目し、ボクセル空間に比べて、より解像度の高い画像空間中でのエネルギー関数を定義している。この結果、オブジェクトらしさの不正確さに起因する復元精度の限界を改善することに成功していると考えられる。

5. むすびと今後の課題

本稿では、自由視点映像の合成画質向上を目標に、視体積交差法で復元される Visual Hull を高精度化する手法について提案した。提案手法は、ボクセル空間中での隣接関係だけでなく、整形過程の3次元形状モデルにおけるテクスチャ状態を考慮している点を最大の特徴としている。具体的なアプローチとしては、ボクセル空間中でのエネルギー関数と、撮影カメラ視点でのエネルギー関数を定義し、削るべきボクセルを、最小化問題の枠組みで決定することで、ボクセルの高精度化を図るものである。

CG、および実写の多視点映像を用いて、従来手法との比較実験を行った結果、Visual Hull 中の被写体領域を誤って削る問題を低減しつつ、凹領域を高精度に復元可能であることを確認した。今後の課題として、提案手法の自由視点合成画質に基づく処理を適用した結果において、形状の滑らかさを向上させる手法の確立が挙げられる。

参考文献

- 1) 石川彰夫, メヒルダドバナヒプルテヘラニ, 内藤整, 酒澤茂之, 小池淳: ウォークスルーを実現するための自由視点映像合成方式, 信学論, J92-D, 6, pp.854-867 (2009).
- 2) 三功浩嗣, 石川彰夫, 内藤整, 酒澤茂之: 被写体 3 次元形状モデル投影型背景分離方式, 映像情報メディア学会誌, 64, 11, pp. 1685-1697 (2010).
- 3) T. Kanade, P. W. Rander, and P. J. Narayanan: Virtualized reality: Constructing virtual worlds from real scenes, IEEE Multimedia, 4, 1, pp. 34-37 (1997).
- 4) N. Inamoto and H. Saito: Virtual Viewpoint Replay for a Soccer Match by View Interpolation From Multiple Cameras, IEEE trans. Multimedia, 9, 6, pp.1155-1166 (2007).
- 5) A. Laurentini: How far 3d shapes can be understood from 2d silhouettes, IEEE trans. Pattern Analysis and Machine Intelligence, 17, 2, pp. 188-195 (1995).
- 6) K. N. Kutulakos and S. M. Seitz: A theory of shape by space carving, Int. J. Comput. Vis., 38, 3, pp. 192-218 (2000).
- 7) 延原章平, 和田俊和, 松山隆司: 弾性メッシュモデルを用いた多視点画像からの高精度 3 次元形状復元, 情報処理学会論文誌. 43, SIG 11 (CVIM 5), pp. 53-63 (2002).
- 8) 富山仁博, 片山美和, 折原豊, 岩館祐一: 局所的形状特徴に拘束された 3 次元形状復元手法とそのリアルタイム動画表示, 映像情報メディア学会誌, 61, 4, pp. 471-481 (2007)
- 9) K. Hisatomi, K. Tomiyama, M. Katayama, Y. Iwadate: Method of 3D reconstruction using graph cuts, and its application to preserving intangible cultural heritage, IEEE conference on ICCV, pp. 923 – 930 (2009).