

3次元ビデオの高精度・多彩な映像化

高井 勇志 松山 隆司

京都大学大学院情報学研究科

〒 606-8501 京都市左京区吉田本町

{takesi-t, tm}@vision.kuee.kyoto-u.ac.jp

3次元ビデオは、実世界での人や動物などの形、動き、色を余すところなく記録した新しい映像メディアであり、教育、医療、娯楽、デジタルアーカイブなど様々な分野での利用が期待されている。本論文は、この3次元ビデオの映像化の方法として、撮影画像に対し忠実且つ高精度に映像化する方法（高精度映像化）、及び、幾何的、光学的編集を行うことによって新たな映像効果を加え映像化する方法（多彩な映像化）についてまとめたものである。

High Fidelity and Versatile Visualization of 3D Video

Takeshi Takai and Takashi Matsuyama

Graduate School of Informatics, Kyoto University

Yoshida, Sakyo-ku, Kyoto, 606-8501

{takesi-t, tm}@vision.kuee.kyoto-u.ac.jp

3D video is a real movie recording dynamic visual events in the real world as they are: time-varying 3D object shapes with a high-fidelity surface texture. This paper presents two different types of 3D video visualization that we consider; “high fidelity visualization” that enables us to visualize free-viewpoint object images that reflect input multi-view images with high fidelity, and “versatile visualization” that enables us to visualize free-viewpoint object images with versatile effects, i.e., geometric directing by utilizing a shape and/or motion, and photometric directing by estimating reflectance parameters and lighting environment.

1 序論

1.1 はじめに

多視点映像から自由視点映像を生成する研究は盛んに行われているが、そのアプローチによって以下のように分類することができる。

まず、対象の形状については、

- 3次元形状を復元しない方法
appearance の復元
- 3次元形状を復元する方法
generic property の復元

また、表面テクスチャ情報については、

- 撮影した照明環境における見え方を復元する方法
appearance の復元
- 照明環境に不変な表面情報を復元する方法
generic property の復元

となる。まとめると、表 1 のようになる。

表 1: 任意視点映像生成法の分類。

形状	app	app	gen	gen
テクスチャ	app	gen	app	gen
Type	I	II	III	IV

Type I (形状:app & テクスチャ:app) 陽に 3次元形状を復元することなく、任意視点から見た映像を生成する方法である。このタイプの典型的な例はイメージベースドレンダリング (IBR) と呼ばれる、非常に

多くの多視点映像から指定された視点・視線に近い映像を合成する方法であり、映画 [5] やスポーツ中継 [11] など一般的に用いられるようになっている。また、藤井らは自由視点テレビと呼ぶ光線空間法をリアルタイムで行うシステムを提案している [6]。ほかにも、中間表現としては対象の 3次元形状を復元しつつも、最終的にはこの形状データを含まない方法 [16]、また、中間表現として 2次元平面の組み合わせを用いる方法 [13] などが研究されている。しかし、動作認識やアーカイブとしての利用、また表面の反射特性の推定には 3次元形状が不可欠なため、このタイプの方法の応用は限られている。

Type II (形状:app & テクスチャ:gen) 一般に、照明に不変な対象表面の反射特性を推定するには、対象の 3次元形状を復元し反射モデルに基づいた解析を行う必要があり、このタイプの方法の実現は難しい。一方、文献 [19] や [29] では、反射モデルを利用することなく、多様な照明下で撮られた多数枚の画像に対して固有画像解析を施すことによって、任意照明下でのテクスチャの復元が行えることを示している。

Type III (形状:gen & テクスチャ:app) 対象の 3次元形状を復元し、表面情報は見え方をベースに復元する方法である。この方法では、対象の 3次元動的動作に同期したカメラワークなど、見え方のみの方法とは異なる映像表現が可能となる。ただし、表面テクスチャは撮影時の照明環境下でのものが使われるため、照明条件を変更した映像は生成できない。

Type IV (形状:gen & テクスチャ:gen) 対象の 3次元形状の復元と共に、撮影されたビデオ映像から撮影時の

照明環境に依存する要素を取り除き、対象表面の反射特性を推定する方法である。このタイプでは、任意の光源下での見え方の復元といった、他のタイプでは表現不可能な映像を生成することが可能となるが、対象の反射モデルの推定には、正確な3次元形状情報の取得や撮影時の照明環境の推定が必要となる。

本論文では、3次元ビデオ[15]の映像化を、type III と type IV の2つの観点から検討を行い、**高精度映像化と多彩な映像化**という2つの映像化法について述べていく。

1.2 3次元ビデオの高精度映像化

形状を復元し、テクスチャは撮影した多視点画像をそのまま用いる映像化方法として、我々は視点依存頂点ベーステクスチャマッピングを提案する。本方法の特徴は、映像化する際の仮想視点に基づいてテクスチャを生成する画像を動的に選択し、頂点に色を与えそれらを補間することでテクスチャを生成することであり、完全には正確とはいえない形状に対しても高精度なテクスチャマッピングを実現することができる。本論文では、第2節において本方法の有効性を各種の評価実験によって定量的に示す。

1.3 3次元ビデオの多彩な映像化

本論文では、多彩な映像化を、幾何的編集と光学的編集という2つの観点から捉え、それぞれについて映像化方法を示す。

光学的編集に必要となる照明環境推定方法として、スケルトンキューブと呼ぶ中空の立方体を参照物体として用いた照明環境推定法と、対象世界に置かれた複数の参照球上の陰影パターンに基づいて、光源の種類、位置、放射強度を推定する方法について第3節、第4節で述べる。

照明環境推定については、数多くの研究がなされている[20, 8, 9, 23, 24, 7, 22, 4, 18, 17, 31, 3, 14, 25, 32, 27, 28, 30]。しかし、その多くは、無限遠光源を用いたり少数に限った光源群のパラメータを求めているため、室内の自然な照明環境を推定・表現することは容易ではない。そこで、我々は近接光源を含めた照明環境を推定する方法を考案し、それぞれの方法の有効性を実画像実験によって評価した。

第5節では、光学的編集として、上記の方法によって推定された照明環境を用いて、撮影対象の反射パラメータを推定し、その後、推定された照明環境と反射パラメータを利用して、光源の影響の除去や新たな光源の照射環境下における対象の映像化について述べている。一方、幾何的編集では、シーンの構図を指定することで対象の形状に基づいてカメラパラメータを計算する、自動カメラワーク生成を提案している。本論文では、実画像を用いた実験により、これら多彩な映像化の効果を示す。

2 高精度テクスチャマッピング法

2.1 はじめに

3次元ビデオの映像化における一つの問題は、完全に正確とはいえない形状に対してテクスチャマッピングを

行うことである。本節では、まず単純な方法である視点独立面ベース法とその問題点、それらを解決するために本研究で考案した視点依存頂点ベース法について述べる。なお、テクスチャ生成は以下のデータを用いて行った。

- 時系列3次元メッシュデータ

以下の2種類のメッシュデータを用いた。

- **Mesh:** 各フレームのボクセルデータに離散マーチングキューブ法[12]を施し、その結果得られたメッシュデータ。
- **D-Mesh:** 各フレームの**Mesh**に対しフレーム内変形を施したメッシュデータ。前章の最後で述べたように、現状のメッシュ変形アルゴリズムはトポロジーの変化には対応していないため、フレーム間変形は行っていない。以下で述べるテクスチャマッピング法については、各フレーム独立して行われるため、フレーム毎にトポロジーが変化しても問題はない。

- 時系列多視点ビデオ映像
- カメラキャリブレーションデータ

2.2 3次元ビデオ撮影システム

図1に、3次元ビデオ撮影システムの概略図を示す。また、本システムの構成は以下の通りである。このシステムにより撮影した画像及び復元した形状に対して、テクスチャマッピングを行う。

- **PC クラスタ:** Myrinetにより接続された16台のPC。
- **カメラシステム:** 12台の視点固定型パン・チルトカメラ。

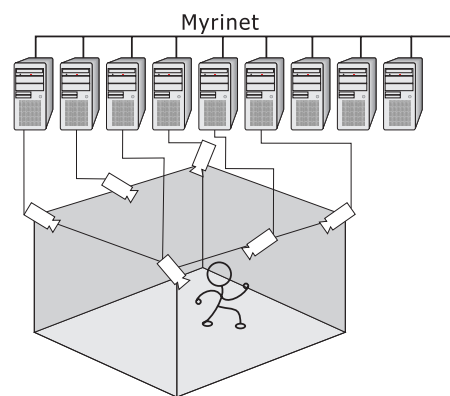


図1: 3次元ビデオ撮影システム。

2.3 視点独立面ベーステクスチャマッピング法

テクスチャマッピングアルゴリズム(図2)

- step 1** テクスチャマッピングを行う面 p_i とその面に隣接する全ての面の法線の平均ベクトル(局所平均法線ベクトル V_{lmn})を求める。
- step 2** V_{lmn} とカメラ c_j から面 p_i に向かうベクトル V_{c_j} のなす角が最大となるカメラ c_j を選ぶ。
- step 3** カメラ c_j により撮影された映像を用いて面パッチ p_i のテクスチャを生成する。

しかし、この方法では以下のような問題点が生じてしまう。

1. カメラ間の撮影映像の差異によるちらつき

マーチングキューブ法により生成される面は滑らかではないため、隣接する面パッチ間で法線の向きが大きく変化する。そのため、テクスチャマッピングに使用するカメラ c_j が隣接パッチで異なってしまう。つまり、異なったカメラで撮られた画像が対象表面の隣接パッチに張られ、映像化された対象表面にちらつきが生じてしまう。(視点独立法の問題点)

2. 隣接面パッチ間でのテクスチャの不整合によるちらつき

同じカメラで撮影された映像を基にしてテクスチャマッピングを行った場合でも、テクスチャは面パッチ毎に独立して生成されるため、各面パッチの継ぎ目において不整合が生じることがある。このことによっても映像化された対象表面にちらつきが生じてしまう。(面ベース法の問題点)

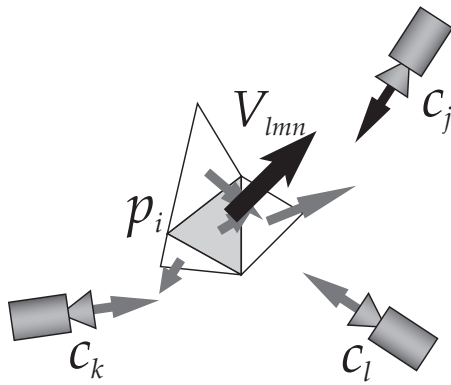


図 2: 視点独立面ベースのテクスチャ生成法.

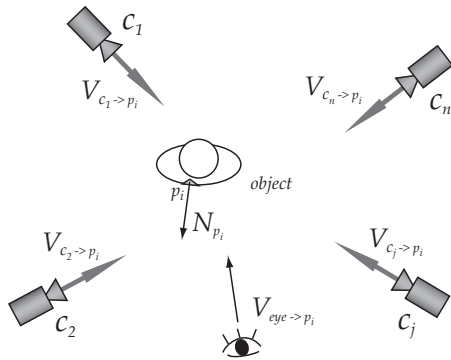


図 3: 視点, カメラ視点.

2.4 視点依存頂点ベーステクスチャマッピング法

上記の問題を解決し、滑らかなテクスチャマッピングを実現するために、視点依存頂点ベース法を考えた。なお、頂点ベース法では、3次元形状の解像度が問題となるが、本研究で用いる3次元形状は1cm ボクセルサイズで表現されているため、対象が人などである場合は十分な解像度が得られていると考えてよい。

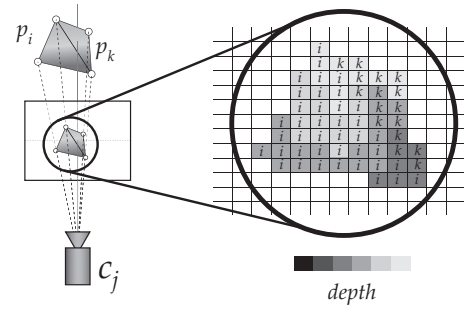


図 4: デプスバッファ.

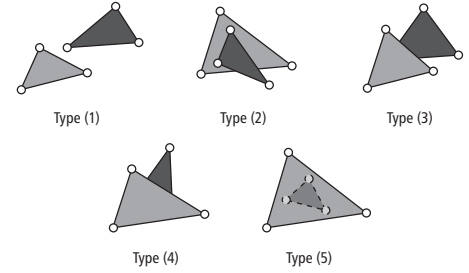


図 5: 面パッチ頂点の関係.

2.4.1 語句と記号の定義

まず、以下のように語句と記号を定義する(図 3).

- 撮影カメラの集合: $C = \{c_1, c_2, \dots, c_n\}$
 - 表示視点: **eye**
 - 離散マーチングキューブ法で得られた多面体モデルを構成する面パッチの集合: $P = \{p_1, p_2, \dots, p_m\}$
 - 面パッチ p_i の法線: N_{p_i}
 - 表示視線: $V_{eye \rightarrow p_i}$ (表示視点から注目する面パッチ p_i の重心へのベクトル)
 - カメラ c_j の視線: $V_{c_j \rightarrow p_i}$ (カメラ c_j から注目する面パッチ p_i の重心へのベクトル)
- (それぞれのベクトルは正規化されているものとする。以降も同様。)

- 面パッチ p_i の頂点: $\mathbf{v}_{p_i}^k$ ($k = 1, 2, 3$)
- カメラ c_j からの可視頂点: $\mathbf{v}_{p_i, c_j}^k$ (後述)
- 可視頂点 $\mathbf{v}_{p_i, c_j}^k$ の RGB 値: $I(\mathbf{v}_{p_i, c_j}^k)$ (後述)
- カメラ c_j のデプスバッファ(図 4): B_{c_j} (カメラ c_j により撮影された映像と同じ座標系を持つ平面で、各画素の値は以下のようにして求める。まず面パッチ集合をカメラの投影中心を中心とした中心投影によって平面上に投影する。そして各画素の位置に投影されたパッチの内、カメラの投影中心からの距離が最小であるパッチを求め、そのパッチ ID と距離値を画素の値とする。また、パッチの頂点が投影された画素にはその頂点 ID も記録する。)

2.4.2 カメラ c_j からの可視頂点

あるカメラ c_j から見えている面パッチ p_i の可視頂点 $\mathbf{v}_{p_i, c_j}^k$ は以下のように定義できる。

1. 法線 N_{p_i} とカメラ視線 $V_{c_j \rightarrow p_i}$ の間に式 (1) が成り立つ。

$$N_{p_i} \cdot V_{c_j \rightarrow p_i} < 0 \quad (1)$$

2. 他の面パッチ $p_k \in P$ により隠蔽されていない。

$\mathbf{v}_{p_i, c_j}^k$ を求める処理は以下になる。まず与えられたカメラ視線 $V_{c_j \rightarrow p_i}$ に対して式 (1) が成り立つ全ての面 $p_i \in P$ をデプスバッファ B_{c_j} に投影する。このとき、デプスバッファに記録されている面パッチ群とそれらの頂点の位置関係は図 5 のように分類でき、他の面によって隠蔽されていない頂点をカメラ c_j からの面パッチ p_i 上の可視頂点 $\mathbf{v}_{p_i, c_j}^k$ とする。つまり、type(1) と type(2) は全ての頂点が可視頂点であり、type(3) と type(4) においては他の面に隠蔽されていない頂点が可視頂点である。また、type(5) のように、他の面パッチに完全に隠蔽された面パッチ (図中の破線の三角形) には可視頂点は存在しない。

可視頂点 $\mathbf{v}_{p_i, c_j}^k$ の RGB ベクトル $I(\mathbf{v}_{p_i, c_j}^k)$ は以下のように求める。

$$I(\mathbf{v}_{p_i, c_j}^k) = I_{c_j}(\hat{\mathbf{v}}_{p_i, c_j}^k) \quad (2)$$

但し、 $I_{c_j}(\mathbf{v})$ はカメラ c_j で撮影された画像の画素位置 $\mathbf{v}$ での RGB 値を表し、 $\hat{\mathbf{v}}_{p_i, c_j}^k$ は可視頂点 $\mathbf{v}_{p_i, c_j}^k$ をその画像上に投影したときの画素位置を表す。

2.4.3 テクスチャマッピングアルゴリズム

step 1 全てのカメラ $C = \{c_1, \dots, c_n\}$ に対して、全ての頂点 $\mathbf{v}_{p_i}^k$ の可視・不可視判定を行い、可視頂点の RGB 値を予め決定しておく。また、各頂点は各カメラからの可視情報を持っている。

step 2 表示視点 **eye** を決定する。

step 3 全ての面パッチ $p_i \in P$ について step.4-9 の処理を行う。また、表示用として使う面パッチ p_i の各頂点の RGB 値 $I(\mathbf{v}_{p_i}^k) = \mathbf{0}$ ($k = 1, 2, 3$) と初期化しておく。

step 4 $V_{eye \rightarrow p_i} \cdot N_{p_i} < 0$ が成り立つ場合のみ、step.5-9 の処理を行う。(その面パッチ p_i は視点 **eye** から見て表向きである。)

step 5 全てのカメラ $c_j \in C$ について、重み係数 $w_{c_j} = (V_{c_j \rightarrow p_i} \cdot V_{eye \rightarrow p_i})^m$ を計算する。ここで m は重み付け係数¹である。また、 $w_{c_j} < 0$ となる場合は $w_{c_j} = 0$ とする。

step 6 面パッチ p_i の全ての頂点 $\mathbf{v}_{p_i}^k$ について step.7-8 の処理を行う。

step 7 頂点 $\mathbf{v}_{p_i}^k$ がカメラ c_j より可視である場合は $w_{c_j}^k = w_{c_j}$ 、不可視である場合は $w_{c_j}^k = 0$ とし、式 (3) を用いて重みの正規化を行う。

$$\bar{w}_{c_j}^k = \frac{w_{c_j}^k}{\sum_l w_{c_l}^k} \quad (3)$$

step 8 頂点 $\mathbf{v}_{p_i}^k$ の RGB 値 $I(\mathbf{v}_{p_i}^k)$ を式 (4) より求める。

$$I(\mathbf{v}_{p_i}^k) = \sum_{j=1}^n \bar{w}_{c_j}^k I(\mathbf{v}_{p_i, c_j}^k) \quad (4)$$

¹この m の値を大きくとれば、表示視線 ($V_{eye \rightarrow p_i}$) により近い視線 ($V_{c_j \rightarrow p_i}$) を持つカメラ c_j の重み w_{c_j} が相対的に大きくなる。つまり、表示視線に近い視線を持つカメラ c_j の影響を大きくすることで、カメラ c_j により撮影された画像の見え方に近いテクスチャ生成が期待できる。

step 9 頂点の RGB 値を用いた線形補間によって面パッチ p_i 上のテクスチャを生成する。即ち、各カメラから可視である頂点の数に従い、以下の処理を行って面パッチ p_i のテクスチャを生成する。

- **3 箇の場合:** 3 頂点の RGB 値の線形補間によって三角形内の各点の RGB 値を求める。
- **2 箇の場合:** 残りの 1 点の RGB 値を他の 2 点の RGB 値の平均値として与え、線形補間をする。
- **1 箇の場合:** 残りの 2 点の RGB 値はこの 1 点の値と同値とする。つまり生成されるテクスチャは 1 色となる。
- **0 箇の場合:** この面パッチにはテクスチャは生成されない。

RGB 値を持たない頂点の色付けは大雑把な推定となるが、面パッチの大きさが対象に比べて充分小さい場合には妥当な値が得られるものと考えられる。

以上の処理の結果、1 フレーム分の実写 3 次元像が生成される。

2.5 性能評価

本論文で提案した視点依存頂点ベース法 (以下方法 (I) と呼ぶ) の精度およびメッシュ変形の有効性を示すために、方法 (I) と視点独立面ベース法 (方法 (II)) を **Mesh** と **D-Mesh** に対して行い、生成した画像を評価した。

図 6 と図 7 に、**Mesh** および **D-Mesh** の同フレームに対して方法 (i) と方法 (II) をそれぞれ用いて生成した画像を示す。これらの画像から以下のことが分かる。

- 方法 (I) と方法 (II) によって生成された画像を比較すると、後者に多く見られた画像のちらつきが前者では大幅に改善されている。
- **Mesh** と **D-Mesh** を用いて生成された画像を比較すると、
 - **D-Mesh** と方法 (II) を用いてテクスチャ生成した場合、依然として多くのちらつきが発生しているが、詳細なテクスチャが生成されている。これは、メッシュ変形の結果、面の法線が平滑化され正確な形状に近付いたためだと考えられる。
 - 一方、**D-Mesh** と方法 (I) を用いてテクスチャ生成した場合、画質は改善されておらず、ボケが発生しているように見える。この結果は、視点に依存してテクスチャ生成することが、対象形状の正確さより重要であることを表している。つまり、視点依存頂点ベース法が、必ずしも正確とはいえない形状に対しても有効であることを示している。

また、図 8 に **D-Mesh** の 2 つのフレームを方法 (I) と方法 (II) によって映像化した画像を示す。この図からも方法 (I) の有効性が分かる。

次に、**Mesh** と **D-Mesh** を、方法 (I) および方法 (II) により映像化し、定量的な評価を行った。つまり、あるカメラ c_j の位置・視線方向を表示視点・視線とし、方法 (I) と手法 (II) によりそれぞれレンダリングを行い、生成された画像シーケンスの RGB 値とカメラ c_j によって撮影された生画像の RGB 値との 2 乗平均平方根誤差を求めた。

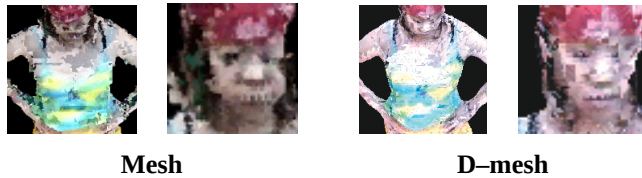


図 6: 視点独立面ベース法によって生成された画像。

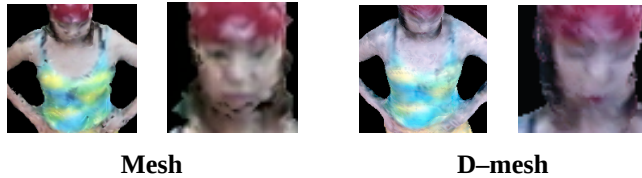


図 7: 視点依存頂点ベース法によって生成された画像。

なお、方法 (I)-1 はカメラ c_j によって撮影された画像を含めて映像化した場合、方法 (I)-2 はカメラ c_j による画像は含めずに映像化した場合である。実験環境は、以下の通りである。

- カメラ設定：図 9
- カメラ画像：640×480[pixel] RGB 各 8bit
- 復元視点：カメラ 5
- 重み付け係数： $m = 5$

図 10 より以下のことが分かる。

- **Mesh** および **D-Mesh** のどちらの場合も、方法 (I)-1、方法 (I)-2 は方法 (II) に比べ大幅に誤差を軽減している。
- **Mesh** と **D-Mesh** を比較すると、
 - **D-Mesh** を方法 (II) によって映像化すると大幅な改善が見られる。
 - **D-Mesh** を方法 (I)-2 によって映像化した場合、10%の誤差の減少が見られる。
 - 一方、**D-Mesh** を方法 (I)-1 によって映像化した場合は、誤差が増加している。これは前述したように、ボケが生じたためである。

最後に、元の面パッチを図 11 のように 3 分割・6 分割と細分化し、提案方法によるさらなる高精細表示の有効性の検証を行った。この実験はメッシュデータは **D-Mesh** のみを用いて行った。

まず、細分化前、3 分割、6 分割によって得られた各面パッチを、撮影に用いた各カメラの画像平面上に投影し、画像上における面パッチの 1 辺の長さの平均を求めた (図 12)。縦軸は投影後の面パッチ 1 辺の平均値 (pixel)、横軸はカメラ番号を表しており、表示しているデータは上から、元の面パッチ、3 分割した面パッチ、6 分割した面パッチを用いた結果を示している。カメラ 9 は対象の真上となる位置に取り付けたカメラで (図 9)、対象までの距離が他のカメラより近いいため解像度が高く、撮影画像上に対象が大きく写っている。このため、他のカメラの値より大きな値が結果として表れている。

図 12 が示すように 6 分割した場合の面パッチの 1 辺は、ほぼ 1 ピクセルとなっており、これが本撮影システムでの解像度の限界であると言える。この 6 分割した面パッチの実空間での解像度は 1 辺約 5mm である。

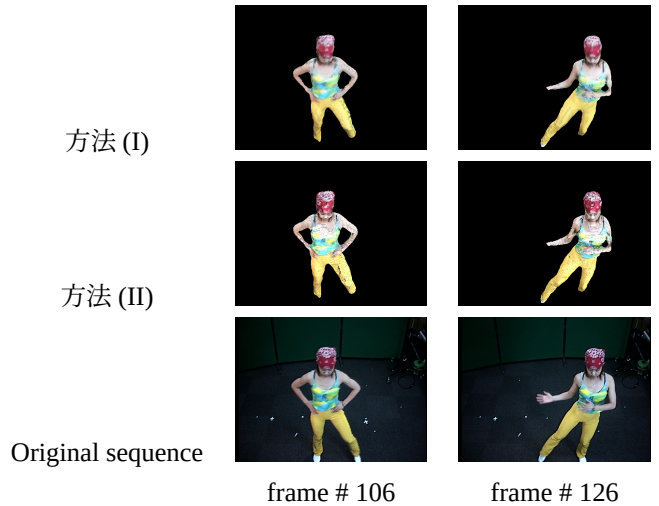


図 8: D-mesh を映像化した例。

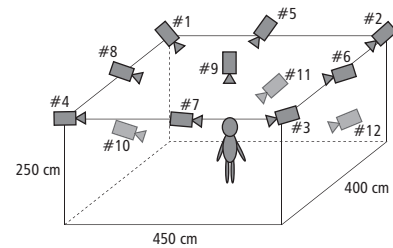


図 9: 撮影環境。

次に、分割した面パッチを用いた映像化の定量的な評価を行うため、方法 (I)-1 によって映像化した映像 (元面パッチ、3 分割、6 分割) を用いて、図 10 と同様にして 2 乗平均平方根誤差を求めた結果を図 13 に示す。

この図によると、面パッチの分割を行っても誤差値は殆ど変わらないという結果になった。しかし、生成された映像には明らかな違いが現れている (図 15)。この誤差の主なものは、対象の輪郭部分やテクスチャの境界部分 (肌と服との境界など) に現れている (図 14)。これらは撮影カメラ間の時間同期が完全には取れていないことやモーションブレンダーが原因となって生じているため、面パッチの分割数を多くしても減少させることは困難である。この結果、局所的に誤差が減少している個所があっても、全体的には誤差は減少しないことになり、図 13 のような結果になったと考えられる。

最後に、6 分割した面パッチを用いてカメラ 5 と 11 お

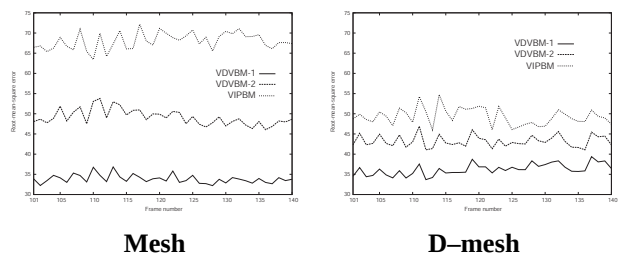


図 10: RGB 値の 2 乗平均平方根誤差 (1)。図中、VDVBM, VIPBM はそれぞれ方法 (I), 方法 (II) を示す。

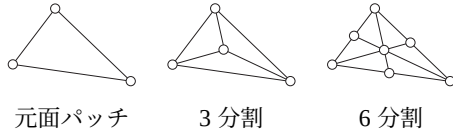


図 11: 面パッチの細分化.

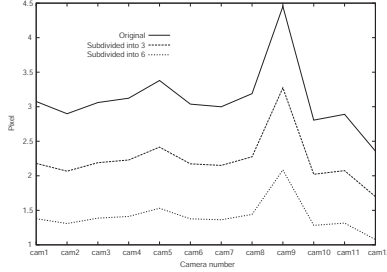


図 12: 元画像上における面パッチ 1 辺の平均値 (pixel).

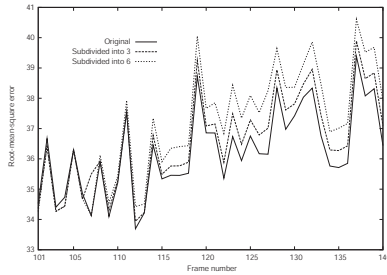


図 13: RGB 値の 2 乗平均平方根誤差 (2).



図 14: 元画像と生成画像 (6 分割パッチ) の差分画像 (frame #106): 元画像と生成画像 (6 分割パッチ) において、対象が映っている個所の RGB 値の差の絶対値の平均を輝度値として生成した。なお、結果を見やすくするために輝度値を反転させている。

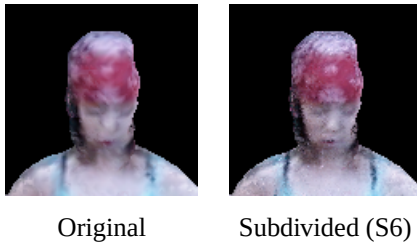


図 15: 元面パッチと 6 分割面パッチを用いて映像化した例 (frame #103).



図 16: 細分化した面パッチを用いた映像化 (frame #103).

よびその中間位置を表示視点として方法 (I)-1 によって映像化した例を図 16 に示す。これにより、任意の表示視点においても、ほぼ撮影された映像と同等の映像が生成できることが分かる。

3 スケルトンキューブを用いた照明環境推定

3.1 照明環境モデル

本論文では、照明環境を空間中に分布した光源群で表現しており、これらの光源は、種類、位置、方向、放射強度といったパラメタを持っている。特に本論文では、点光源のみでなく、自然な照明環境の表現が可能な面光源を扱っている。本節では、この 2 種類の光源のモデル化や、照明環境を表現するための光源の分布について述べる。

3.1.1 光源モデル

本論文では、点光源、面光源共に近接光源として扱っている。すなわち、その照明効果は、照らされる物体の位置や、距離によって変化する。我々はスタジオのような室内空間の照明環境の推定を目指しているため、このように近接光源として光源を扱うことが不可欠である。

点光源 点光源は空間中の 1 点から全ての方向に照射する光源である。図 17 は、点光源とそれによって照射される微小面 dS の関係を示している。点光源が距離 r 離れた位置からその面を照射しているとき、その放射照度は以下の式で表される。

$$E(\lambda) = \frac{d\Phi(\lambda)}{dS} = \frac{d\omega}{dS} L(\lambda) = \frac{\cos \theta}{r^2} L(\lambda), \quad (5)$$

ここで、 $\Phi(\lambda)$, $L(\lambda)$ はそれぞれ放射束、放射強度を示している。また、立体角については以下の関係を用いている。

$$d\omega = \frac{dS \cos \theta}{r^2}, \quad (6)$$

ここで、 θ は面 dS の法線と、その面から点光源へ向かう直線とがなす角を表している。

面光源 本研究では、面光源を、その面の法線方向に一樣な放射束を照射する光源と定義する。また、モデルを単純化するために、面光源から照射される面の距離を、その面光源の中心から照射される面の距離によって表すこ

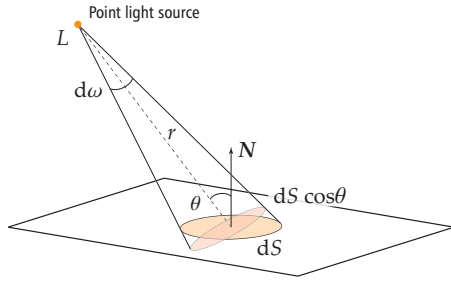


図 17: 点光源と照射される面の関係. N は面 dS の法線を表す.

とにする. 図 18 は, 上記の定義に基づく面光源とそれによって照射される微小面 dS の関係を示している. 面光源の面積を S_L , 距離を r とすると, その放射照度は以下の式で表される.

$$E(\lambda) = \frac{d\Phi(\lambda)}{dS} = \frac{\cos \theta}{r^2} \int_{S_L} L'(\mathbf{x}, \lambda) d\mathbf{x}, \quad (7)$$

ここで, $L'(\mathbf{x}, \lambda)$ は面光源上の点 $\mathbf{x}$ での単位面積当たりの放射強度を示しており, θ は面 dS の法線と, その面から面光源の中心へ向かう直線とがなす角である. 面光源は一様な放射強度を持つので,

$$\int_{S_L} L'(\mathbf{x}, \lambda) d\mathbf{x} = L'(\lambda) S_L \cos \varphi = L(\lambda) \cos \varphi, \quad (8)$$

と書ける. ここで, φ 面光源の向きと面光源の中心から面 dS へ向かう直線とがなす角である. 最終的に放射照度は以下の式により得られる.

$$E(\lambda) = \frac{\cos \theta}{r^2} L(\lambda) \cos \varphi. \quad (9)$$

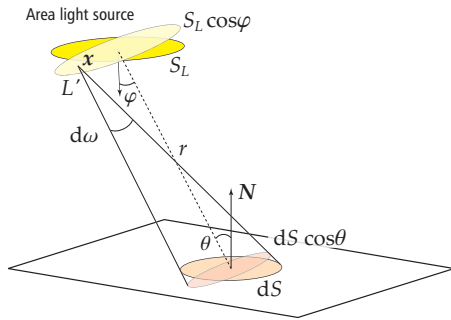


図 18: 面光源と照射される面の関係.

3.1.2 光源群の配置

様々な照明環境を表現するためには, 大量の光源を配置することが理想的である. また, 室内の照明環境を表現するために, 壁などからの反射も無視することはできず, 更に, スポットライトや蛍光灯なども考慮する必要がある. しかしながら, 計算コストの観点では光源数を最小限とする必要もある. 上記のような要求を鑑み, 本論文では, 以下の2種類からなる照明環境を考案した. すなわち, **背景照明**: スタジオの構造(壁, 床, 天井)を表す

照明, **前景照明**: 空間中の光源を表す照明である. この2種類の照明を用いることにより, 一般性を大きく失うことなく光源数を限定することができる.

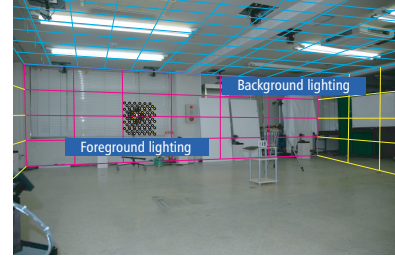


図 19: 背景照明と前景照明の例.

背景照明 本論文で用いる背景照明は, 壁や天井のような撮影空間を覆うものとしている. そこで, 撮影空間を形作るように面光源を配置することによって背景照明を表現する. どこまで詳細に表現するかは要求されるレベルによって異なるが, 大抵の場合は直方体上に面光源を分布させることで良い近似となる. 例えば, 蛍光灯の付いた天井は面光源を適切に分布させることによって表現できるように, 任意の背景照明が面光源によって表現できる. 更に, オープンスペースの照明環境は半球上に配置した面光源によって表現可能である.

前景照明 本論文で用いる前景照明は, スポットライトのような一般的な光源を点光源によって表している. これらの光源が存在する可能性のある空間は予め知ることができるため, 必要最小限の光源を分布させることができる.

3.2 スケルトンキューブ

3.2.1 構造的特徴

本研究では, 参照物体の陰影のみではなく, 参照物体が自分自身に落す影(**セルフシャドウ**)の情報を利用して, 照明環境を構成する光源のパラメタ推定を行う. この目的を達成するための参照物体には, 以下の2つの相反する性質が要求される. すなわち,

- **単純性**: 参照物体上の多くの点を計測するためには, より広い面が観測できることが重要である. また, 計算コストにおいても, 単純な形状が有利である.
- **複雑性**: 多様な照明環境下でセルフシャドウが起こるためには, ある程度複雑な形状が必要である.

の, 2点である.

以上の要求を満たすため, 我々は**スケルトンキューブ**と呼ぶ, 中空の立方体を照明環境推定の参照物体として提案する(図 20). スケルトンキューブは, どのような方向から照射する光源によっても内壁にセルフシャドウを生ずることができ, また, その表面が観測し易いという特徴がある. ある特別な視点と光源位置の組合せ²によっては, セルフシャドウを観測することが困難な場合もあるが, そのようなときには, 視点を変えるかスケルトン

²特に, スケルトンキューブ内部に光源がある場合は内壁にセルフシャドウが生じないが, この場合は, 単にスケルトンキューブの位置を変えれば良い.

キューブの位置や方向を変えることで、照明環境推定に必要な情報を得ることができる。つまり、複数視点からスケルトンキューブを撮影すれば、理論上では撮影シーンの完全な照明環境を推定することが可能である。

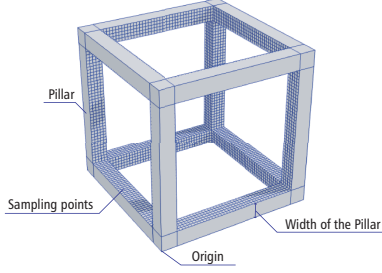


図 20: スケルトンキューブ。サンプル点はスケルトンキューブ内壁の格子点により表されている。

3.2.2 光学的特徴

スケルトンキューブ表面の反射特性は、Torrance-Sparrow モデル [26] を簡略化したモデル [10] で記述できると仮定する。Torrance-Sparrow モデルは、拡散反射と鏡面反射の両方を表現しており、特に、鏡面反射については反射面の物理的な意味に基づいて表現されているため、経験から導き出された Phong モデル [21] よりも、現実の物体の反射をより正確に表現することができる。また、本研究では、スケルトンキューブ自身の相互反射は無視できるものとし、1 次反射のみを扱って照明環境推定を行う。

視点 V と光源 L が与えられたとき、物体表面の点 x の放射輝度は以下のように書ける。

$$I(x) = (k_d R_d + k_s R_s) L_{\mathcal{L}} \quad (10)$$

ここで、 L は光源 $\mathcal{L}$ の放射強度である。また、 R_d と R_s はそれぞれ双方向反射分布関数 (BRDF) の拡散反射成分と鏡面反射成分を示しており、 k_d と k_s はそれぞれの係数である。更に、 R_d と R_s は以下の式で表される。 $I(x)$ と L は波長の関数であるが、本論文では、赤、緑、青の3種のみを扱う。また、以下では記述を簡略にするため波長については明示しない。

拡散反射を Lambertian と仮定すると、拡散反射成分 $R_d = N \cdot L / r^2$ が得られる。ここで、 r は光源 $\mathcal{L}$ と点 x との距離である。また、鏡面反射成分は以下のように書ける。

$$R_s = \frac{1}{N \cdot V} \exp \left[-\frac{(\cos^{-1}(N \cdot H))^2}{2\sigma^2} \right] \quad (11)$$

ここで、 V, H, σ はそれぞれ、視線ベクトル、 L と V の中間ベクトル、物体表面の荒さを示している。

3.2.3 スケルトンキューブの放射輝度

スケルトンキューブ上の点 x における放射輝度は以下のように書ける。

$$I(x) = \sum_{i=1}^N \mathcal{M}(x, \mathcal{L}_i) (k_d R_d + k_s R_s) L_{\mathcal{L}_i}, \quad (12)$$

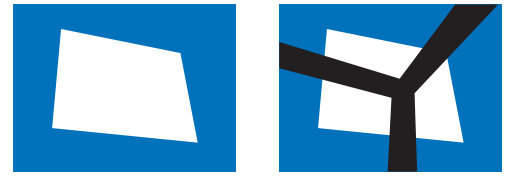
ここで、 N は光源数、 $\mathcal{M}(x, \mathcal{L}_i)$ はセルフシャドウを表す隠蔽項である。隠蔽項は光源の種類によって異なり以下のように表される。

点光源 $\mathcal{M}(x, \mathcal{L}_i)$ は、光源 $\mathcal{L}_i$ が点 x を照射しているか否かの二値で表される。すなわち、 $\mathcal{M}(x, \mathcal{L}_i) = 1$ 或いは $\mathcal{M}(x, \mathcal{L}_i) = 0$ となる。

面光源 $\mathcal{M}(x, \mathcal{L}_i)$ は、光源の隠蔽率を示す $[0, 1]$ の連続値で表される。すなわち、面光源 $\mathcal{L}_i$ の全面が点 x を照射している場合は $\mathcal{M}(x, \mathcal{L}_i) = 1$ となり、セルフシャドウにより全く照射しない場合は $\mathcal{M}(x, \mathcal{L}_i) = 0$ となる。隠蔽項の計算は以下のように行われる。

- 1° スケルトンキューブが**無い**状態で、点 x から見える面光源 $\mathcal{L}_i$ を描画し、その面積 S を計算する (図 21a)。
- 2° スケルトンキューブが**在る**状態で、点 x から見える面光源 $\mathcal{L}_i$ を描画し、その面積 S' を計算する (図 21b)。
- 3° 隠蔽項を以下の式により計算する。

$$\mathcal{M}(\mathcal{L}_i, x) = S' / S. \quad (13)$$



(a) 面光源 (b) 隠蔽された面光源

図 21: スケルトンキューブの内壁より見える面光源。白い四角形は面光源を、黒い多角形はスケルトンキューブを表している。

3.3 照明環境推定

3.3.1 実験環境

本研究では、下記的前提の下、スケルトンキューブの内壁の輝度をサンプリングすることで照明環境推定を行う。

- **座標系:** シーン座標系の原点とスケルトンキューブの中心を一致させ、シーン座標系の各座標軸とスケルトンキューブの各面をそれぞれ平行になるようにする。また、照明環境や撮影に用いるカメラも、このシーン座標系に配置する。
- **照明環境:** 近接点光源の分布は前節によるモデルを用いる。
- **スケルトンキューブ:** 大きさと表面の反射特性は予め計測しておく。
- **カメラ:** シーンを撮影するカメラはスケルトンキューブを用いて校正されるため、撮影画像とシーン内の対応する点の関係は既知である。

3.3.2 定式化

スケルトンキューブ内壁の M 点の放射輝度値をサンプリングすると、式 (12) を基に以下の行列が得られる。

$$\begin{bmatrix} I(x_1) \\ I(x_2) \\ \vdots \\ I(x_M) \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{M1} & K_{M2} & \cdots & K_{MN} \end{bmatrix} \begin{bmatrix} L_{\mathcal{L}_1} \\ L_{\mathcal{L}_2} \\ \vdots \\ L_{\mathcal{L}_N} \end{bmatrix}, \quad (14)$$

ここで、点光源の場合は

$$K_{mn} = \mathcal{M}(x_m, L_n) (k_d R_d + k_s R_s) \quad (15)$$

であり、面光源の場合は

$$K_{mn} = \mathcal{M}(x_m, L_n) (k_d R_d + k_s R_s) \cos \varphi \quad (16)$$

である。式 (14) を以下のように書き換える。

$$I = KL, \quad (17)$$

ここで、 $K = (K_{mn})$ である。充分な数のサンプルが得られると (i.e. $M \gg N$)、式 (17) を L について解くことが出来る。つまり、

$$L = K^+ I, \quad (18)$$

となる。ここで、 K^+ は K の擬似逆行列である。

3.4 性能評価

実画像に対する照明環境推定の結果を示す。実験は以下の条件に基づいて行った。

- 照明環境: 前景照明として 125 個の点光源を、背景光源として 200 個の面光源を配置する (図 22)。
- スケルトンキューブ: 大きさや反射パラメタは予め計測しておく。
- カメラ: カメラパラメタは予め推定しておく。また、画像は RGB 各 12bit で記録されている。

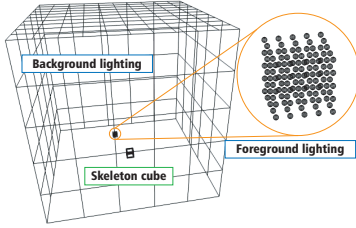


図 22: 照明環境モデル。

図 23a に入力画像を示す。この画像中のスケルトンキューブの陰影を解析することにより、照明環境を推定する。また、推定した照明環境を用いて実画像中に CG 物体（ティーポット）を合成した例を図 23c に示す。更に、比較のために無限遠光源を半球上に分布させることによって表現された照明環境を生成し、同様に推定・合成を行った結果を図 23b に示す。

図 23c を見ると陰影や机に落とす影など、以下の点において自然な表現ができていることがわかる。1) 近接光源の効果により、ティーポットが机に落とす影の方向がそれぞれ異なっている。2) ティーポット表面の輝度が電球からの距離に応じて異なっている。3) 鏡面反射の位置がティーポットの位置に応じて異なっており、また、画像右上にあるビリヤード球と矛盾無く現れている。しかしながら、上記の特徴は図 23b には当てはまらない。例えば、ティーポット表面の明るさは位置によらず一様で、影は同方向に落ちていることがわかる。これらの原因は、明らかに無限遠光源の仮定から生じているため、図 23b で用いたような参照物体の中心での放射照度のみを表現する照明モデルは、室内の照明環境には不十分であるといえる。対照的に、我々の提案モデルは近接光源を用いているため、効果的に室内の照明環境を表現できている。

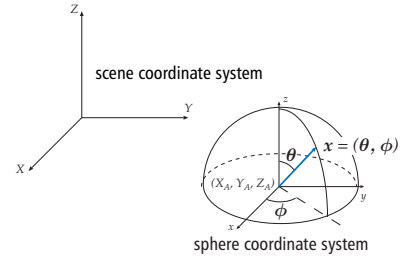


図 24: 座標系。

4 差分球：近接光源推定へのアプローチ

4.1 モデルの定義

4.1.1 仮定

本研究では、以下の仮定を満たしているものとして、光源推定を行う。

- 参照物体として、参照球と呼ばれる既知の形状と完全拡散反射特性を持った球を用いる。
- 画像を撮影するカメラは校正されており、参照球との位置関係は既知である。
- 画像輝度は対象表面の反射輝度と正比例関係にあるとし、画像輝度がそのまま対象表面の反射輝度を表すもの³とみなす。
- これらの参照球は、撮影画像上で互いに隠蔽し合わず、また、互いに影を落とし合わない位置に配置されている。
- 相互反射は考慮しない。

4.1.2 座標系

図 24 に示すように、シーン座標系 (X, Y, Z) 、参照球座標系 (x, y, z) を定義する。参照球 A が与えられたとき、その参照球座標系の各軸 x_A, y_A, z_A は、シーン座標系の各軸 X, Y, Z にそれぞれ互いに並行であるとする。また、参照球上に、極座標で表される参照球表面座標系 (θ, ϕ) を定義する。

4.1.3 光源モデル

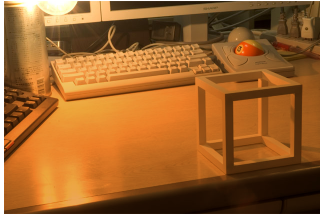
前述したように、本研究では、近接点光源、無限遠光源、環境光の 3 種類の光源を扱う。本節では、それぞれの光源と参照球の反射輝度の関係について述べる。なお、参照球の反射係数は η であるとする。

- **近接点光源** 参照球上の点 X において、法線と光源方向とがなす角を $\alpha_p(X, P)$ とすると、反射輝度 $I_p(X; P)$ は、

$$I_p(X; P) = \eta L_p \cos \alpha_p(X, P) / (D(X, P))^2, \quad (19)$$

と書ける。ここで、 L_p 、 P 、 $D(X, P)$ は、それぞれ、光源の放射強度、光源位置、光源と点 X の距離を表す。

³言い換えると、反射輝度と画像輝度の関係は線形であり、その係数は 1 である。



(a) 入力画像



(b) 合成画像例
(無限遠光源による照明環境モデル)



(c) 合成画像例
(提案照明環境モデル)

図 23: 実験結果.

- **無限遠光源** 参照球上の点 X において、法線と光源の照射方向とがなす角を $\alpha_d(X)$ とすると、反射輝度 $I_d(X)$ は、

$$I_d(X) = \eta L_d \cos \alpha_d(X). \quad (20)$$

と書ける．ここで、 L_d は光源の放射強度を表す．

- **環境光**: 環境光はバイアスとして扱う．放射強度が L_a で表される環境光のみによって、シーンが照射されているとすると、参照球の反射輝度は、

$$I_a = \eta L_a. \quad (21)$$

と書ける．

4.1.4 放射輝度

上記の光源によって照射されている参照球の放射輝度は、シーン座標系の下、以下のように記述できる．

$$I(X) = \sum_{i=1}^s I_p^{[i]}(X; P[i]) + \sum_{j=1}^t I_d^{[j]}(X) + I_a, \quad (22)$$

ここで、 s と t はそれぞれ近接点光源と無限遠光源の数 (未知数) を表し、 i と j はそれぞれ近接点光源と無限遠光源のラベルを表す．また、 $P[i]$ は光源 i のシーン座標系での 3 次元位置を表す．

4.2 参照球に現れる特徴

4.2.1 放射輝度

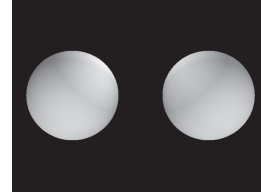
単独球 点 x_A を参照球表面座標系にある単独球 A 上の点とし、 $\mathcal{F}_A(x_A)$ をシーン座標系で表した点 x_A の 3 次元位置とする．点 x_A の反射輝度 $I_A(x_A)$ は、式 (22) において以下の I_p , I_d , I_a を用いることで記述できる．

$$I_p^{[i]} = \eta_A L_p^{[i]} \cos \alpha_p^{[i]}(\mathcal{F}_A(x_A), P[i]) / (D(\mathcal{F}_A(x_A), P[i]))^2, \quad (23)$$

$$I_d^{[j]} = \eta_A L_d^{[j]} \cos \alpha_d^{[j]}(\mathcal{F}_A(x_A)), \quad (24)$$

$$I_a = \eta_A L_a, \quad (25)$$

ここで、 η_A は拡散反射係数であり、 $L_p^{[i]}$ と $L_d^{[j]}$ はそれぞれ近接点光源 i と無限遠光源 j の放射強度を示す．



Negative light source Positive light source

図 25: 単独球と差分球．赤が正の輝度値、緑が負の輝度値を表している．

差分球 $I_A(\theta, \phi)$, $I_B(\theta, \phi)$ をそれぞれ参照球 A , B の反射輝度とすると、差分球 $A-B$ の反射輝度は以下の式で表される⁴．

$$I_{A-B}(\theta, \phi) = I_A(\theta, \phi) - I_B(\theta, \phi). \quad (26)$$

したがって、式 (22)-(25) および (26) より、差分球 $A-B$ 上の点 x_{A-B} における反射輝度 $I_{A-B}(x_{A-B})$ は以下のように書ける．

$$I_{A-B}(x_{A-B}) = I_A(x_A) - I_B(x_B) = \sum_{i=1}^s I_p^{[i]}(\mathcal{F}_A(x_A); P[i]) - \sum_{i=1}^s I_p^{[i]}(\mathcal{F}_B(x_B); P[i]), \quad (27)$$

$$x_{A-B} = x_A = x_B = (\theta, \phi). \quad (28)$$

上記の式によって示されるように、差分球の反射輝度値は、全ての無限遠光源と環境光による影響が相殺されるため、近接点光源の情報のみを持つことが分かる．

ここで、差分処理によって生じる負の反射輝度を扱うために、通常の光源である**正光源**と、照射面の反射輝度値を減少させる光を発する**負光源**の 2 種類の光源を導入する．つまり差分球は、正光源と負光源の 2 種類の光源によって照射されていると言える (図 25)．

この 2 種類の光源は、式 (27) より分かるように、近接点光源と同じ性質を持っている．光源推定の際には、この正負のどちらかの光源のパラメタを推定すれば、実際にこの参照球を照射する近接点光源の放射強度が推定されたことになる．また、光源位置については、差分球の定義より、正光源を用いた場合は参照球 A の座標系で表され、負光源を用いた場合は参照球 B の座標系で表される．

ここで、推定の際の正光源あるいは負光源の選択は特に重要な問題ではなく、一般には撮影画像上でより大きな領域を照射している光源を選択する．

⁴ 差分球 $A-B$ の反射輝度 $I_{A-B}(\theta, \phi)$ は、参照球 A あるいは B 上の点 (θ, ϕ) が撮影カメラから不可視である場合は定義されない．

4.2.2 照射領域解析

参照球表面の分類 参照球の表面は、その領域を照射する光源の数や種類によって以下の3通りの領域に分けることができる。

- **S 面:** 単一の光源 (と環境光) によって照射された面。特に、近接点光源によって照射されている面を S_p 面、無限遠光源によって照射されている面を S_d 面と呼ぶ。
- **M 面:** 複数の近接点光源や無限遠光源によって照射されている面。
- **A 面:** 環境光のみに照射されている面。

本研究で提案する光源推定方法では、**S 面の持つ特徴を利用して光源のパラメタを推定する**。すなわち、S 面の持つ以下の特徴、

- **特徴 1:** S 面上の同一輝度点は、3次元空間中の平面上の円弧として表される。この平面を**特徴面平面**と呼ぶ。
- **特徴 2:** 特徴面平面の参照球の中心に向かう法線ベクトルが、その S 面を照射する光源の照射方向と一致する。
- **特徴 3:** S 面の特徴面平面群の**輝度変化率**は、光源の照射輝度によって一意に決まる。

を利用する。

以下では、 S_p 面と S_d 面各々の場合における上記の性質を詳しく述べ、差分球を利用した光源推定法のメカニズムについて述べていく。

ここで、以下の説明で利用する記号についてまとめる。

- $O = (O_X, O_Y, O_Z)$: 参照球の中心
- $P = (P_X, P_Y, P_Z)$: 光源位置
- $l = (l_X, l_Y, l_Z)$: 光源の照射方向
- R : 参照球の半径

S_d 面の特徴 図 26 の太線は、輝度値 $I(x)$ である点の集合を表している。図から明らかなように、この点の集合は特徴面平面上の円弧を示している。[特徴 1]

また、図 26 より、特徴面平面の法線は光源の照射方向と一致することが分かる。この特徴面平面は以下の式で表される。

$$\begin{aligned} & l_X(X - (O_X - l_X R \cos \beta_d(x))) \\ & + l_Y(Y - (O_Y - l_Y R \cos \beta_d(x))) \\ & + l_Z(Z - (O_Z - l_Z R \cos \beta_d(x))) = 0, \end{aligned} \quad (29)$$

ここで、 $(l_X, l_Y, l_Z) = (O_X - P_X, O_Y - P_Y, O_Z - P_Z)$ であり、 $\beta_p(x, l)$ は l と O から x へ向かうベクトルがなす角を表している。また、式 (29) により、 S_d 面に定義される全ての特徴面平面は互いに並行であることが示されている。

[特徴 2]

更に、異なる輝度値を持つ特徴面平面群は、 S_d 面を照射する無限遠光源の放射強度の特徴を示している。また、 $\beta_d(x)$ の変化に伴って特徴面平面の輝度値が変化することから、以下では、 S_d 面の輝度値を $\alpha_d(x)$ ではなく $\beta_d(x)$ を用いて記述する。すなわち、 S_d 面上の点 x における反射輝度 $I(x)$ は、

$$I(x) = \eta L_d \cos \alpha_d(x) + \eta L_a, \quad (30)$$

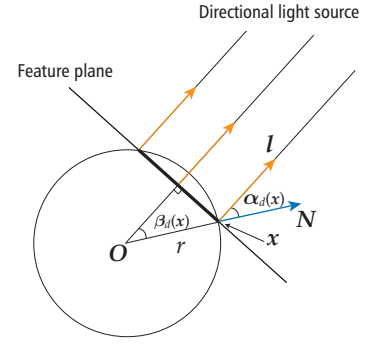


図 26: 無限遠光源と参照球との関係。

と書ける。ここで、この式を $\beta_d(x)$ を用いて以下のように書き換える。

$$I(x) = \eta L_d \cos \beta_d(x) + \eta L_a, \quad (31)$$

ここで、 $\beta_d(x) = \alpha_d(x)$ である。任意の 3 特徴面平面の輝度を用いると、 S_d 面の輝度変化率は、

$$\frac{I(x_1) - I(x_2)}{I(x_1) - I(x_3)} = \frac{I'(x_1) - I'(x_2)}{I'(x_1) - I'(x_3)} \quad (32)$$

と書ける。ここで、 $I'(x) = \cos \beta_d(x)$ である。[特徴 3]

以上の S 面の分析結果をまとめると、ある S 面の輝度変化率が式 (32) で与えられるとき、その S 面は S_d 面であり、その S_d 面を照射する無限遠光源及び環境光のパラメタを推定することができる、と言える。

4.2.3 S_p 面の特徴

図 27 より、 S_d 面の場合と同様に、 S_p 面の**特徴 1**および**特徴 2**を示すことができる。このとき、特徴面平面は、

$$\begin{aligned} & l_X(X - (O_X - l_X R \cos \beta_p(x, l))) \\ & + l_Y(Y - (O_Y - l_Y R \cos \beta_p(x, l))) \\ & + l_Z(Z - (O_Z - l_Z R \cos \beta_p(x, l))) = 0, \end{aligned} \quad (33)$$

と表される。

特徴面平面上の点 x における反射輝度 $I(x)$ は、

$$I(x) = \eta L_p \cos \alpha_p(\mathcal{F}(x), P) / (D(\mathcal{F}(x), P))^2 + \eta L_a, \quad (34)$$

と書ける。ここで、 $\cos \alpha_p(\mathcal{F}(x), P)$ および $D(\mathcal{F}(x), P)$ は以下の通りである。

$$\cos \alpha_p(\mathcal{F}(x), P) = \frac{D_p^2 - R^2 - (D(\mathcal{F}(x), P))^2}{2RD(\mathcal{F}(x), P)}, \quad (35)$$

$$D(\mathcal{F}(x), P) = \sqrt{D_p^2 + R^2 - 2RD_p \cos \beta_p(x, l)}. \quad (36)$$

ここで、 D_p は O と P との距離である。上式を式 (34) へ代入すると、

$$I(x) = \eta L_p \frac{D_p \cos \beta_p(x, l) - R}{(D_p^2 + R^2 - 2RD_p \cos \beta_p(x, l))^{\frac{3}{2}}} + \eta L_a. \quad (37)$$

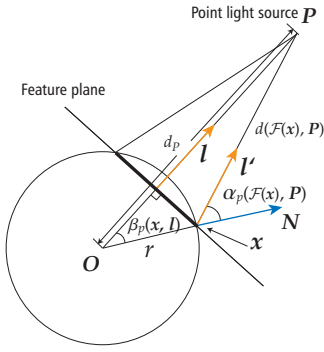


図 27: 近接点光源と参照球の関係。

が得られる。式 (37) は、反射輝度が未知数である L_p , L_a および D_p によって求まることを示している。ここで、任意の 2 特徴面平面の輝度値の差を取ることで環境光の項を消去すると、 S_p 面の輝度変化率は以下のように書ける。

$$\frac{I(\mathbf{x}_1) - I(\mathbf{x}_2)}{I(\mathbf{x}_1) - I(\mathbf{x}_3)} = \frac{I'(\mathbf{x}_1) - I'(\mathbf{x}_2)}{I'(\mathbf{x}_1) - I'(\mathbf{x}_3)}, \quad (38)$$

ここで、

$$I'(\mathbf{x}) = \frac{D_p \cos \beta_p(\mathbf{x}, \mathbf{l}) - R}{(D_p^2 + R^2 - 2RD_p \cos \beta_p(\mathbf{x}, \mathbf{l}))^{\frac{3}{2}}}.$$

である。

式 (38) を解くことで D_p を求めることができ、更に、式 (37) を解くことで、 L_p および L_a が得られる。[特徴 3] しかしながら、式 (38) を解くことは容易ではないため、単独球のみで光源のパラメタを推定することは困難である。

差分球の 1 光源照射領域の特徴 前節で述べたように、単独球のみから近接点光源のパラメタを求めることは容易ではない。式 (38) を解くことを困難にしている原因は環境光の項であるが、差分球においては、その定義から分かるように環境光の影響を除去することができる。つまり、差分球上の S 面は S_p 面と同様の性質を持つが、環境光の項を取り除いた式となる。すなわち、差分球上の点 $\mathbf{x}$ における反射輝度 $I(\mathbf{x})$ は以下のように書ける。

$$I(\mathbf{x}) = \eta L_p^* \cos \alpha_p(\mathcal{F}(\mathbf{x}), \mathbf{P}) / (D(\mathcal{F}(\mathbf{x}), \mathbf{P}))^2, \quad (39)$$

ここで、 L_p^* は正光源あるいは負光源の放射強度を表す。したがって、輝度変化率は以下ようになる。

$$\frac{I(\mathbf{x}_1)}{I(\mathbf{x}_2)} = \frac{I'(\mathbf{x}_1)}{I'(\mathbf{x}_2)}, \quad (40)$$

ここで、

$$I'(\mathbf{x}) = \frac{D_p \cos \beta_p(\mathbf{x}, \mathbf{l}) - R}{(D_p^2 + R^2 - 2RD_p \cos \beta_p(\mathbf{x}, \mathbf{l}))^{\frac{3}{2}}}.$$

である。

式 (40) を解くことで D_p を求めることができ、式 (39) により L_p^* を得ることができる。式 (40) により、差分球を用いることで輝度変化率の式がはるかに単純になり、近接点光源のパラメタが式 (38) と比べて容易に求まることが示されている。

4.3 光源推定アルゴリズム

前節までの記述に基づいた、光源推定のアルゴリズムの概略を以下に示す。

step 0 画像の撮影。

step 1 差分球画像の生成。

step 1-1 近接点光源パラメタの推定。

step 2 推定した近接点光源の影響を取り除いた画像の生成。

step 2-1 無限遠光源・環境光のパラメタの推定。

上記のサブステップ (1-1, 2-1) において、光源のパラメタを反復処理によって推定する。すなわち、入力画像から推定された光源候補の影響を一つずつ取り除き、残った領域を解析することで、その光源候補の検証を行いパラメタを推定する (図 28)。その処理は以下の通りである。

i. 画像を濃度等高線画像へ変換。 (4.3.1 節)

ii. 光源候補パラメタの推定。 (4.3.2 節)

1° 等高線を等高線線分に分割。

2° S 面候補の抽出。

3° S 面候補を解析し、光源候補パラメタを推定。

iii. 光源候補の検証。 (4.3.3 節)

1° 入力画像から、光源候補の影響を取り除いた画像を生成し、新たな入力画像を生成。

2° 画像中の残値を検証。

– もし、残値が一樣な値であれば、その画像を生成した光源候補群は正解であるとみなし、処理を終える。

– もし、残値が負の値を持つ領域（負領域）があれば、その光源候補は不正解であると判断する。

– 上記以外の場合は、iii-1° で生成した画像を新たな入力画像とし i に戻る。

以下では、上記の i から iii の処理について詳しく述べていく。

4.3.1 濃度等高線表現を用いた領域分割

濃度等高線表現は、画素の輝度値をその地点の標高とみなした地形図と似た表現方法である [1]。本研究では、入力画像を濃度等高線表現に変換し、それを利用して S 面の抽出を行う。濃度等高線表現は、画像中の対象の陰影の幾何的な情報を反映しており、更に画素単位のノイズが発生しても等高線全体としての性質は大きく変化しない。したがって本研究では、 S 面の解析に、通常の輝度値の配列で表現された画像より適していると考えられる濃度等高線表現を用いる。

4.3.2 光源パラメタの推定

複数の光源によって照射されている参照球上の面では、等高線は、大抵、複数の等高線線分によって成り立っている。そのため、まず、全ての等高線を円弧として表現できる等高線線分に切り分ける必要がある。この等高線線分への切り分けは、本研究の光源推定の基本となる S 面抽出の手がかりとなるため、重要な処理である。本研究では、等高線線分が示す平面の法線を調べ、同じ法線

を持つ平面上の等高線線分をまとめることにより、参照球上の S 面を抽出する。

ここで、前章で得られた S 面の特徴 1 を、濃度等高線を用いた表現に書き換えると、以下のようになる。

特徴 1' : S 面上の等高線線分は 3 次元空間中の円弧として表される。この平面を特徴面平面と呼ぶ。

以上より、濃度等高線表現を用いた光源パラメタの推定は以下のように書ける。

- 1° **特徴 1'** に基づいて、等高線を等高線線分に区切る。
- 2° 特徴面平面の法線が同一方向となる等高線線分をまとめ、 S 面を抽出する。
- 3° **特徴 2** および **特徴 3** に基づいて、その S 面を照射する光源のパラメタを推定する。

4.3.3 光源候補の検証

光源候補の検証は、入力画像からその光源の影響を取り除いた画像（残値）を解析することで行われる。つまり、もし残値が負となる領域があれば⁵、その光源は不正解であるといえる。

ここで、参照球が 3 つの無限遠光源で照射されている場合を考える（図 28）。入力画像中には 4 つの S 面候補があるが、正しい S 面は C_2 と C_4 であり、その他は M 面である。これらの候補の中から、同一輝度値の等高線線分を解析することだけによって、正しい S 面を見出すことは不可能であるが、光源候補の影響を取り除いて得られる残値を解析すれば、 M 面を見出すことは可能である。すなわち、図 28 中 1st verification の左から 3 つ目において負領域が現れているため、 C_3 は M 面であると判断できる。以降、画像を更新しながら処理を続けていくことで、正しい推定を行っている 3 つの経路を得ることができる。これらの経路は同一の光源群を推定しており、結果として照明環境が正しく推定されたこととなる。

一般的な照明環境下においても、少なくとも 1 つ正しい S 面が存在すれば、その面に応じた光源を推定し、その影響を入力画像から取り除くことにより、照明環境推定を継続することができる。

上に述べたように、本方法は、可能性のある経路を全て探索し正しい光源群を示す経路を決定している。つまり、試行錯誤を繰り返す処理であるため、誤差の蓄積や経路の選択といった問題が生じるが、正しい経路中にある光源の組み合わせを評価することにより、最適な光源群を見出している。

4.4 性能評価

実画像に対する本方法の有効性を示すため、2 つの参照球を撮影しその空間を照射する光源の推定を行った。図 29 に、光源推定の結果と、推定光源を用いた仮想物体と実画像との合成例を示す。図中では、推定された光源と、それによって照射された仮想物体（ティーポット）が (b)-(d))、また、(e) には推定された 3 光源によって照射された仮想物体が示されている。更に、実画像と仮想物体の、光源の影響を考慮した合成例を図 29f に示す。この

⁵step 1 では残値は零となり、step 2 では残値は一律な値、つまり環境光による輝度となる。

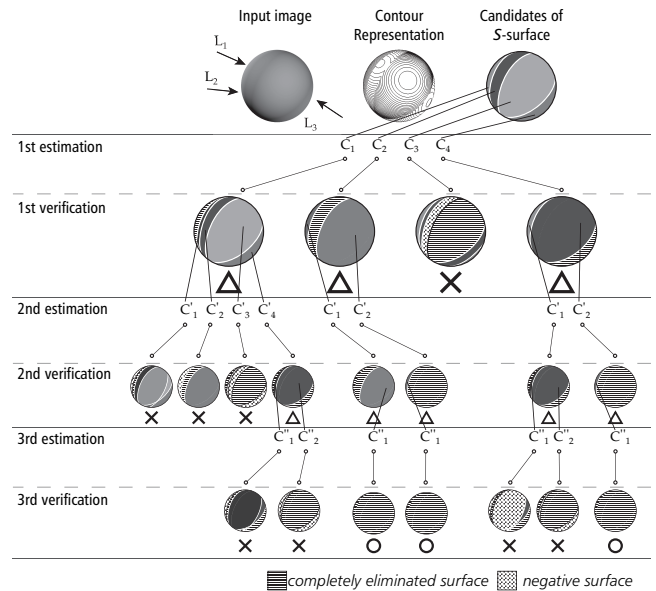


図 28: 推定・検証の流れ。

‘L’, ‘C’ はそれぞれ光源、 S 面候補を示している。また、○, × は正解、不正解を、△ は推定途中であることを示画像では、入力画像から 2 つの参照球を取り除かれ、その結果、参照球の背後にあった近接点光源が現れている。また、2 つの人物が仮想物体との比較のために加えられている。光源推定の結果、近接点光源が実際より床に近い位置に推定されたため、ティーポットの手前の床に落ちる影が若干不自然であるが、ティーポットの陰影などは実物体と違和感なく合成されていることが分かる。

5 3次元ビデオの多彩な映像化

5.1 はじめに

本節では、3次元ビデオの多彩な映像化として以下の2通りの編集方法について述べる。

- **幾何的編集** 撮影対象の形状や動きを利用したカメラワーク設計。
- **光学的編集** 撮影対象の反射パラメタを推定し新たな照明効果の付加。

3次元ビデオは対象の形状を復元し任意視点映像を生成しているため、イメージベースレンダリングのような見えのみを再現する方法では困難である、これらの映像効果を与えることができる。以下では、上記の2通りの編集について述べる。

5.2 幾何的編集

5.2.1 モデルの定義

幾何的編集は以下の要素を用いて行う。

- **撮影対象** 3次元ビデオシステムによって撮影された対象。
- **背景** 半球状の形状に対し、全方位画像をマッピングしたもの。

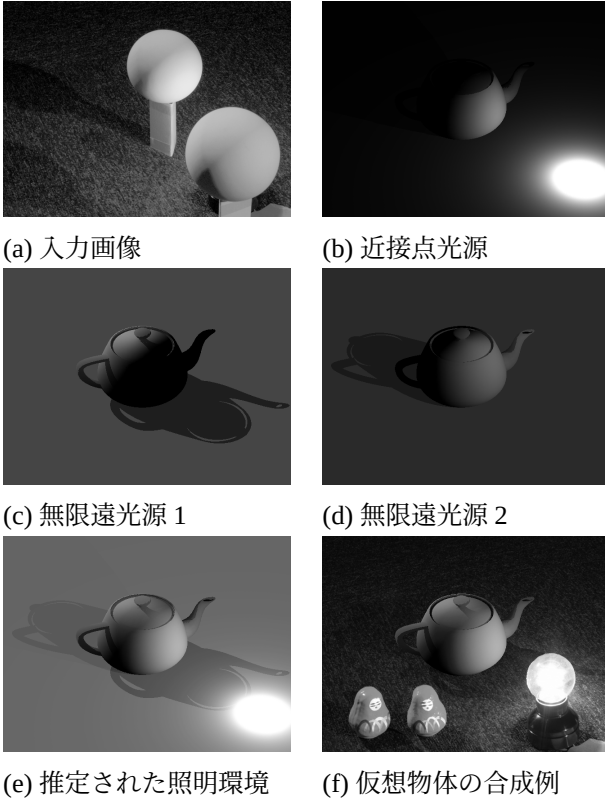


図 29: 実画像に対する光源推定と、推定された光源を用いた仮想物体の合成例。画像 (b), (c), (d) は、推定された光源によって照射される仮想物体を表している。また、画像 (e) は、推定された全ての光源によって照射された仮想物体を表している。

- **仮想カメラ** 仮想空間中に配置し、その視点から見える映像を表示する。本論文におけるカメラワーク設定とは、このカメラのパラメタ設定を指す。

以下では、カメラワークとして、対象の形状を用いて画像の構図（クローズアップ、ミドルショットなど）を決定するアルゴリズムについて述べる。

5.2.2 カメラワーク生成アルゴリズム

カメラワークは以下の処理によって生成される。

- step 1** 最初のフレーム（key frame 0）において構図を設定する。
 - step 2** その構図に応じたカメラパラメタを計算する。
 - step 3** 任意のフレーム（key frame i）において構図を設定する。
 - step 4** その構図に応じたカメラパラメタを計算し、前キーフレーム（key frame i-1）間でパラメタを線形補間する。
 - step 5** 最終フレームに達したら処理を終える。そうでない場合は、step 3 へ。
- 構図に対応するカメラパラメタ計算については以下の節で紹介する。

カメラパラメタ計算 画面構成の指定に用いるカメラパラメタは以下のように表される。

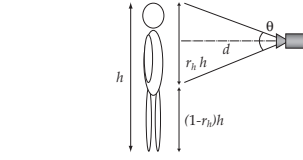


図 30: カメラと対象の位置関係 (I).

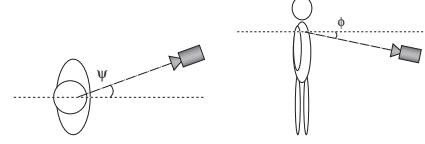


図 31: カメラと対象の位置関係 (II).

◇ 対象までの距離 d

$$d = \frac{r \cdot h}{2 \tan \frac{\theta}{2}} \quad (41)$$

ここで、 h は対象の高さである。また、 r は表示画面に対する対象の大きさの割合であり、以下のように定義する [2]。カメラの画角 θ は任意に設定する (図 30)。

- $r = 1.0$ フルショット
- $r = 0.7$ ニーショット
- $r = 0.5$ ミドルショット
- $r = 0.3$ クローズショット
- $r = 0.2$ クローズアップ

◇ カメラの位置 (x_c, y_c, z_c)

$$\begin{cases} x_c = d \cos \phi \cos(\psi + \delta) + x_p \\ y_c = d \cos \phi \sin(\psi + \delta) + y_p \\ z_c = d \sin \phi + z_p \end{cases} \quad (42)$$

ここで、 δ は対象の正面方向とシーン座標系の x 軸とがなす角である (図 32)。 $\psi \cdot \phi$ は任意に与えるものとし (図 31)、 d は式 (41) より得られる値である。また、点 (x_p, y_p, z_p) はカメラの注視点である。注視点の x_p, y_p は対象の位置の x 座標、 y 座標と一致させ、 z_p については画面構成に応じて変化させる (式 (43))。例えば、クローズアップ ($r = 0.2$) の場合は $z_p = 0.9h$ となり、注視点は対象の顔の辺りとなる。

$$z_p = h - \frac{h \cdot r}{2} \quad (0 \leq r \leq 1) \quad (43)$$

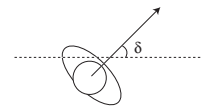


図 32: 対象とシーン座標系の関係。

5.2.3 映像生成結果

以下の構図設定 (表 2) に基づくカメラワーク生成の結果を図 33 に示す。なお、常に対象の正面を取り続けるよう設定してある。図 33 が示すように、指定した構図が表現されており、キーフレーム間で滑らかにカメラパラメタが変化していることがわかる。



図 33: 幾何的映像編集例.

表 2: 構図設定.

Frame Number	Composition
0	フルショット
119	フルショット
237	ニーショット
299	ミドルショット
333	フルショット

5.3 光学的編集

5.3.1 照明環境推定

第3節, 第4節で2種類の照明環境推定法について述べてきたが, 本節では, これら2つの方法を組み合わせ, より効果的な照明環境推定について述べる. スケルトンキューブを用いた方法には, 予め光源分布の仮定が必要であること, また, 参照球を用いた方法には, 複雑な照明環境の推定が困難であるという問題があった. そこで, この2つの方法の組み合わせとして以下の処理を行うことにより, 相補的に照明環境を推定する方法を紹介する.

step 1 照明環境を大まかにモデル化する (図).

step 2 差分球を用いて前景照明のパラメタを推定する (図).

step 3 前段階で求めたパラメタを用いて照明環境を再モデル化し, スケルトンキューブを用いて推定する (図).

この処理のキーポイントは, 差分球を用いて前景照明の正確なパラメタを求めた後に, 壁からの反射など複雑な照明環境はスケルトンキューブを用いることである. 差分球は球の大きさに対してある程度離れた位置にある光源の影響を除去することができる (つまり, 無限遠光源とみなせる光源は相殺される) ので, 前景光源のみのパラメタを求めることが可能である. 一方, 複雑な照明環境については, 光源分布の仮定の下, スケルトンキューブによって安定した推定が可能である. このようにして, 室内環境の照明を近接光源として扱いながら実践的な推定が可能となる.

5.3.2 反射パラメタ推定

3次元ビデオシステムにより生成される形状は, 対象の概形であるため精密な反射パラメタ推定は困難である. そこで, 本論文では, 対象の反射特性は完全拡散反射であると仮定してパラメタ推定を行う. このとき, 対象上の点 x の放射輝度は以下のように書ける.

$$I(x) = k(x) \sum_{i=1}^n \frac{\mathcal{M}(x, L_i)}{d(x, L_i)^2} N(x) \cdot U(x, L_i), \quad (44)$$

ここで, それぞれの変数は以下のように定義される.

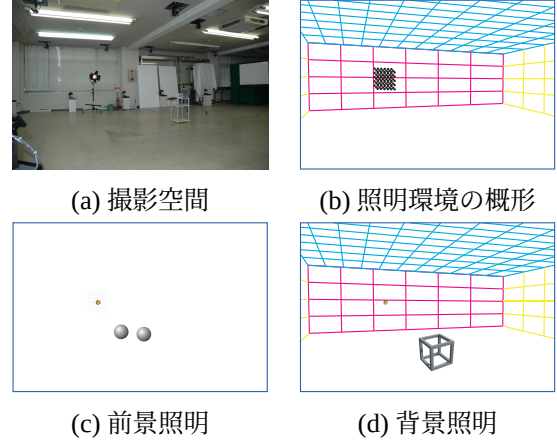


図 34: スケルトンキューブと差分球の組み合わせ.

- $k(x)$: x での反射パラメタ,
- L_i : 光源 i ,
- $d(x, L_i)$: x と L_i の距離,
- $\mathcal{M}(x, L_i)$: 隠蔽項,
- $N(x)$: x の法線ベクトル
- $U(i, x)$: x から L_i へ向かう単位ベクトル.

この式を k について解くことにより反射パラメタを得る. なお, 3次元ビデオの対象は, 各頂点について複数の放射輝度値を持っているが, 本節ではその中央値をとることによって, 対象の放射輝度値とする.

5.3.3 映像生成結果

3次元ビデオの照明効果の編集として, 撮影時の照明環境に対し照明効果の除去及び追加を行った. 図 35a は撮影画像の一例である. 撮影対象 (人) の正面にスポットライトが, 天井には複数の蛍光灯が配置されている. このスポットライトを前景照明, 壁・天井を背景照明とし, 照明環境の推定を行った.

図 35b は取得した3次元ビデオの例である. 照明環境から前景照明を除去して再描画した結果である, 図 35c が示すように, スポットライトの影響が除去されていることがわかる. また, スポットライトに換えて仮想の緑色のライトを配置し再描画した結果を図 35d に示す. 上記の結果により, 3次元ビデオの照明効果の編集が容易且つ効果的に行えることが示された.

6 結論

本論文では, 3次元ビデオの映像化の方法として, 高精度映像化法と多彩な映像化法の2つの観点から検討を行った. 高精度映像化法として視点依存頂点ベーステクスチャマッピング法を提案し, 完全に正確とはいえない

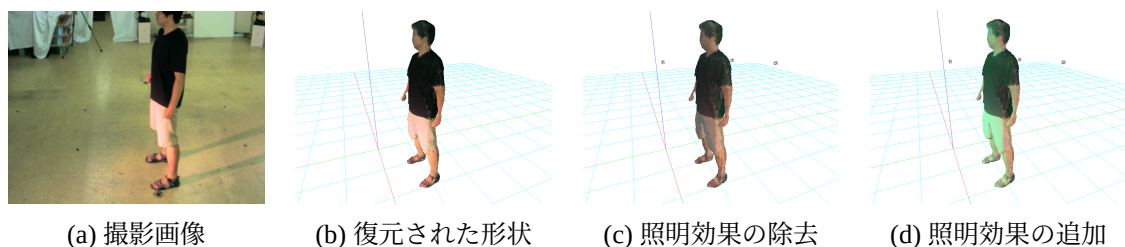


図 35: 光学的編集例.

形状に対してもちらつきの少ない映像生成を行うことが可能となった. 一方, 多彩な映像化に必要な照明環境推定法を提案し, 実環境に対する実験によりその有効性を示した. また, 幾何的・光学的編集を行うことで, 3次元ビデオがより効果的に映像化できることを示した.

今後の課題としては以下の点が挙げられる.

- **高精度映像化**
面ベーステクスチャへの対応.
- **多彩な映像化**
 - 幾何的編集
対象の動作を認識しそれに応じたカメラワーク生成.
 - 光学的編集
照明環境の設計法. 鏡面反射を含めた反射パラメータ推定.

参考文献

- [1] T. Asano and S. Kimura. Contour representation of an image with applications. In *IPSJ Signotes Algorithms*, number 058-009, pages 65–70, 1997.
- [2] Daniel Arijon. *The Grammar of the Film Language*. Focal Press, 1976.
- [3] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):281–233, 2003.
- [4] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *ACM SIGGRAPH*, pages 189–198, 1998.
- [5] P. Debevec. The bullet time, 1999. <http://www.debevec.org/Campanile/>.
- [6] T. Fujii and M. Tanimoto. Free-viewpoint tv (ftv) system. *PCM 2004*, pages 497–504, 2004.
- [7] K. Hara, K. Nishino, and K. Ikeuchi. Determining reflectance and light position from a single image without distant illumination assumption. In *IEEE International Conference on Computer Vision*, pages 560–567, 2003.
- [8] B. K. P. Horn. Image intensity understanding. *Artificial Intelligence*, 8:201–231, 1977.
- [9] D. R. Houge and M. Ahuja. Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In *IEEE International Conference on Computer Vision*, pages 148–155, 1993.
- [10] K. Ikeuchi and K. Sato. Determining reflectance properties of an object using range and brightness images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11):1139–1153, 1991.
- [11] T. Kanade. The eyevision, 2001. <http://www-2.cs.cmu.edu/afs/cs/project/VirtualizedR/www/explain.html>.
- [12] Y. Kenmochi, K. Kotani, and A. Imiya. Marching cubes method with connectivity consideration. In *PRMU-98-218*, pages 297–204, 1999.
- [13] I. Kitahara, H. Saito, S. Akimichi, T. Ono, Y. Ohta, and T. Kanade. Large-scale virtualized reality. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
- [14] S. R. Marschner and D. P. Greenberg. Inverse lighting for photography. In *Fifth Color Imaging Conference*, pages 262–265, 1997.
- [15] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara. Real-time 3d shape reconstruction, dynamic 3d mesh deformation, and high fidelity visualization for 3d video. 96(2004):393–434, 2004.
- [16] Y. Mukaigawa, Y. Nakamura, and Y. Ohta. Face synthesis with arbitrary pose and expression from several images—an integration of image-based and model-based approach—. In *Asian Conference on Computer Vision*, volume 1, pages 680–687, 1998.
- [17] P. Nillius and J.-O. Eklundh. Automatic estimation of the projected light source direction. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume I, pages 1076–1083, 2001.
- [18] P. Nillius and J.-O. Eklundh. Low-dimensional representation of shaded surfaces under varying illumination. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume II, pages 185–192, 2003.
- [19] K. Nishino, Y. Sato, and K. Ikeuchi. Eigen-texture method: Appearance compression and synthesis based on a 3d model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1257–1265, 2001.
- [20] A. P. Pentland. Finding the illumination direction. *Journal of Optical Society of America*, 72(4):448–455, 1982.
- [21] B. Phong. Illumination for computer generated pictures. In *Communications of the ACM*, pages 311–317, 1975.
- [22] M. W. Powell, S. Sarkar, and D. Goldgof. A simple strategy for calibrating the geometry of light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1022–1027, 2001.
- [23] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. In *IEEE International Conference on Computer Vision*, pages 875–882, 1999.
- [24] I. Sato, Y. Sato, and K. Ikeuchi. Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE Transactions on Visualization and Computer Graphics*, 5(1):1–12, 1999.
- [25] T. Matsuyama and T. Takai. Generation, visualization, and editing of 3d video. In *The 1st Symposium on 3D Data Processing Visualization and Transmission*, pages 234–245, 2002.
- [26] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughness surface. *Journal of the Optical Society of America*, 57:1105–1114, 1967.
- [27] Y. Wang and D. Samaras. Estimation of multiple illuminants from a single image of arbitrary known geometry. In *IEEE European Conference on Computer Vision*, pages 272–288, 2002.
- [28] Y. Yang and A. Yuille. Sources from shading. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 534–539, 1991.
- [29] R. Yumiba, W. Xiaojun, S. Tokai, and T. Matsuyama. Natural scene visualization using active camera—seamless composition of central camera image and peripheral panorama—. In *CVIM*, pages 15–19, 1999.
- [30] Y. Zhang and Y. Yang. Multiple illuminant direction detection with application to image system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:915–920, 2001.
- [31] Q. Zheng and E. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):680–702, 1991.
- [32] W. Zhou and C. Kambhampettu. Estimation of illuminant direction and intensity of multiple light sources. In *IEEE European Conference on Computer Vision*, pages 206–220, 2002.