

# 信頼度マッピング法による自由視点映像の実時間生成

國田 豊<sup>†</sup>

<sup>†</sup> 日本電信電話 (株) NTT サイバースペース研究所

多視点画像からの自由視点画像生成においては、被写体の幾何情報をどの程度まで陽に求めるかによって、光線ベースの方法から幾何ベースの方法まで、幅広い手法が提案されている。本発表の手法は、信頼度マップと呼ばれる多層平面上での被写体の存在確率情報を求めることにより、奥行き方向に被写体の多義性を許容する表現をとることが特徴である。このことにより、テクスチャが一様の場合など、ステレオマッチングがとりにくい場合においても、際だった雑音のない画像の生成が可能である。本稿では、信頼度マッピング法の基本的な原理を定式化したあと、複数の視点位置からの信頼度マップを統合して、オクルージョン (遮蔽) に対してロバストな視点依存信頼度マッピング法を記述する。また、計算コストの優位性を実証するため、実時間処理システムについて紹介する。

## Real-Time Rendering Using Probability Mapping Method

Yutaka KUNITA<sup>†</sup>

<sup>†</sup> NTT Cyber Space Laboratories, NTT Corporation

We propose an image-based rendering method called "layered probability mapping." The algorithm requires modest computer power but produces high-quality output images even from textureless or noisy input images. This accommodates ambiguities regarding object depth. Because the layers are locally view dependent, we can save the computing time to acquire globally consistent precise models. Also, by integrating the layers estimated from plural reference viewpoints, the rendered image that include occlusion regions are much improved, since the region with low fidelity estimated from one viewpoint is complemented by high fidelity from other viewpoints. The algorithm that implements the features above is computationally modest and straight forward, without including conditional branching and optimization of non-linear equations. Actually, by adding some limitations, real-time rendering is possible by current commodity hardware. Also, we show the image synthesis results to prove our method can generate higher quality images compared to some conventional methods.

### 1 はじめに

複数の視点位置で撮影された画像をもとにして、新たな視点から見た画像 (自由視点画像) や、3次元ディスプレイ用の画像を生成する技術が活発に研究されている。これらの技術の代表的なものとして、被写体の形状を推定し、所望の視点位置からレンダリングする幾何ベースの方法と、被写体から飛来する光線を所望の視点位置で再サンプリングする光線ベースの方法がある。

幾何ベースの方法では、被写体の幾何形状を推定するために画像間のステレオ対応を用いる方法<sup>1, 2)</sup> や、視体積交差法を用いる方法<sup>3, 4)</sup> がある。さらに両者を組み合わせ、視体積交差法によりまず大まかな形状を求め、ステレオ法により詳細な形状を求める方法も提案されている<sup>5, 6, 7, 8)</sup>。

また光線ベースの方法は、光線の強度を位置、方向、時間、波長に対する関数である plenoptic function<sup>9)</sup> により記述されることに基づき、その次元を落とし、ある視点位置を与えたときにそこを通る光線群を再サンプリングするのに適した表現方法として、いくつかの方式が提案されている<sup>10, 11, 12, 13, 14)</sup>。

両者には一長一短があり、幾何ベースの方法では、2次元の画像列から3次元の形状を一意に求めるのは困難であり、それが最終的な画質に影響する。一方、光線ベースの方法では、被写体が存在すると仮定した付近でしか、正しく画像生成できないことが知られている<sup>15, 16, 17)</sup>。

そのため近年では、純粋な幾何/光線ベースの方法が用いられることはまれで、何らかの方法 (ステ

レオ法, 視体積交差法など)により被写体の形状を求め, 視点依存テクスチャマッピング<sup>18)</sup>で光線方向に応じた表面色の重み付けをするハイブリッド的な手法が使われることが多い。

一方, 被写体の幾何情報を利用しつつも, 陽に推定しない方法が提案されている<sup>19, 20, 21, 22)</sup>。本稿で紹介する信頼度マッピング法も, ステレオマッチングを利用するが, 被写体の形状を陽に求めず, 確率的な表現を用いている。そして, 生成画像の各画像の色を, ステレオ拘束条件から対応づけられる元画像中の色の期待値として求める。

そのようにすることで, 対応点マッチングがとりにくい場合でも, 形状を一意に求めることをしなくてよい。そのため, 計算コストを抑えたり, 推定誤りによる画質の劣化を低減することが期待できる。また, 複数の位置で推定した形状情報を統合して, 新たな視点位置からの画像を生成する場合<sup>23, 2)</sup>でも, 推定の曖昧さを考慮して複数の視点位置の統合ができることが特徴となっている。

本稿では, 信頼度マッピング法の基本的な原理を定式化したあと, 複数の視点位置からの信頼度マップを統合して, オクルージョン(遮蔽)に対してロバストな視点依存信頼度マッピング法<sup>24)</sup>を記述する。また, 計算コストの優位性を実証するため, 実時間処理システムについて紹介する。

## 2 信頼度マップ

### 2.1 原理

まず, 最も基本的な画像生成の原理は次のような式で表される。

$$c_V = \sum_{c \in \Gamma} p(c)c. \quad (1)$$

ここで,  $c_V$  は目的となる画素の色ベクトルであり,  $[r, g, b]^T$  で表される。  $\Gamma$  は  $c_V$  がとりうる色ベクトルの集合であり,  $p(c)$  は  $c_V = c$  となる確率を示す。すなわち式(1)は, 目的となる色が, とりうる色の集合における期待値により表されることを示している。ここで  $\Gamma$  は

$$\Gamma_{\text{all}} = \{[r, g, b]^T | 0 \leq r \leq 1, 0 \leq g \leq 1, 0 \leq b \leq 1\}. \quad (2)$$

のようにすべての色空間とすることもできるが, 以下で述べるようにカメラの幾何的な拘束を利用すると, より限定することができる。

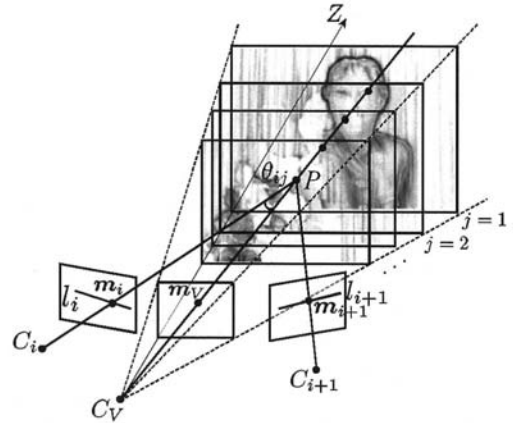


Fig. 1 提案手法における座標系の設定

図1に示すように, 元となるカメラを  $\{C_i | i \in N\}$ , その撮影画像を表す色ベクトルを  $I_i(m)$  とする。  $m \equiv [u, v]^T$  はデジタル座標で,  $\tilde{m} \equiv [u, v, 1]^T$  はその拡張ベクトルである。そして,  $u = u'/w', v = v'/w'$  として  $I_i(m)$  と  $I_i(\tilde{m})$  とは同じとみなす。同様に, 目的となる仮想カメラを  $C_V$ , 注目する画素を  $m_V$ , 生成する画像を  $I_V(m) (= I_V(\tilde{m}))$  とする。被写体上の点  $P$  を撮影している  $m_V$  の対応点は  $C_i (i \in N)$  におけるエピポーラ線上にあり, これを  $l_i (i \in N)$  とすると色の候補はこの線上にあると考えられる。

すると, 式(1)における  $\Gamma$  は仮想カメラの位置  $m_V$  の画素に対応するエピポーラ線上の色の集合に限定でき,

$$\Gamma = \{I_i(m_i) | m_i \in l_i, i \in N\}. \quad (3)$$

さらに, 被写体上の点は, 離散的な平面  $\{L_j | j \in M\}$  上にあると仮定する。すると対応点が特定でき,

$$\Gamma = \{I_i(m_i) | \tilde{m}_i = H_j^{V_i} \tilde{m}_V, i \in N, j \in M\}. \quad (4)$$

ただし  $H_j^{V_i}$  はレイヤ  $L_j$  により規定されるホモグラフィ行列であり,  $m_V$  を  $m_i$  に写す。

一般に, レイヤ  $L_j$  を介してカメラ  $C_a$  の点  $m_a$  をカメラ  $C_b$  の点  $m_b$  に写すホモグラフィ行列は, 以下のように与えられる。

$$H_j^{ab} = A_b \left( R + \frac{tn^T}{d_j} \right) A_a^{-1}. \quad (5)$$

ここで、カメラ  $C_a$  の座標系において、レイヤ  $L_j$  の正規ベクトルを  $\mathbf{n}$ 、原点からの距離を  $d_j$  で表している。  $A_a$  と  $A_b$  はそれぞれカメラ  $C_a, C_b$  の内部行列である。  $\mathbf{R}, \mathbf{t}$  はカメラ  $C_a$  の座標系における点  $\mathbf{X}_a$  をカメラ  $C_b$  の座標系における点  $\mathbf{X}_b$  に変換し、  $\mathbf{X}_b = \mathbf{R}\mathbf{X}_a + \mathbf{t}$  なる関係がある。

ここで、レイヤ  $L_j$  を介した  $\mathbf{m}_V$  のカメラ  $C_i$  における対応点を  $\mathbf{c}_{ij}$  とすると、  $\mathbf{c}_{ij} = \mathbf{I}_i(\mathbf{H}_{ij}^V \mathbf{m}_V)$ 。そして、  $\mathbf{m}_V$  における色が  $\mathbf{c}_{ij}$  である確率  $p(\mathbf{c}_{ij})$  は、被写体がレイヤ  $L_j$  上にある確率  $p(j)$  と、被写体がレイヤ  $L_j$  上にある条件でのカメラ  $C_i$  の寄与率  $p(i|j)$  の積として以下のように表される。

$$p(\mathbf{c}_{ij}) = p(j)p(i|j). \quad (6)$$

すると、式 (1) は以下のように表される。

$$\mathbf{c}_V = \sum_{j \in M} \sum_{i \in N} p(j)p(i|j)\mathbf{c}_{ij}. \quad (7)$$

さらに、式 (7) における画素単位の表現を書き直し、画像単位の表現とすることで、以下の式を得る。

$$\mathbf{I}_V(\mathbf{m}) = \sum_{j \in M} \sum_{i \in N} q_j^V(\mathbf{m})w_{ij}^V(\mathbf{m})\mathbf{I}_{ij}^V(\mathbf{m}). \quad (8)$$

ここで式 (7) 中の  $p(j)$ ,  $p(i|j)$ ,  $\mathbf{c}_{ij}$  は式 (8) 中ではそれぞれ  $q_j^V(\mathbf{m})$ ,  $w_{ij}^V(\mathbf{m})$ ,  $\mathbf{I}_{ij}^V(\mathbf{m})$  に対応する。またこのとき  $\mathbf{I}_{ij}^V$  はレイヤ  $L_j$  に投影され、ターゲットカメラ  $C_V$  により撮影されたカメラ  $C_i$  の画像であり、

$$\mathbf{I}_{ij}^V(\mathbf{m}) = \mathbf{I}_i(\mathbf{m}_i), \quad \tilde{\mathbf{m}}_i = \mathbf{H}_j^V \tilde{\mathbf{m}}. \quad (9)$$

なお、  $\mathbf{I}_{ij}^V(\mathbf{m})$ ,  $q_j^V(\mathbf{m})$ ,  $w_{ij}^V(\mathbf{m})$  の右上の添字 "V" はカメラ  $C_V$  の画像座標系における関数であることを示している。

このように、色と確率が世界座標系で統一の表現を持たず、視点依存性を持つことが、本手法の特徴の 1 つとなっている。

## 2.2 信頼度マップとカメラ寄与率の推定

$q_j^V(\mathbf{m})$  はレイヤ  $L_j$  上に被写体が存在する確率を表す信頼度マップである。この値は、ソースカメラの対応点の一致度により推定される。

まず、  $[r, g, b]^T$  のすべての色要素における分散ベクトルを以下のように計算する。

$$\delta_{ij} = \mathbf{I}_{ij}^V(\mathbf{m}) - \frac{1}{\text{Num}(N)} \sum_{i \in N} \mathbf{I}_{ij}^V(\mathbf{m}). \quad (10)$$

そして、要素の合計を計算し、以下のスカラー値を得る。

$$\sigma_j^2 = \frac{1}{\text{Num}(N)} \sum_{i \in N} \delta_{ij}^T \cdot \delta_{ij}. \quad (11)$$

ここで、  $\text{Num}(N)$  は集合  $N$  に含まれる要素数を表すものとする。

値を平滑化し、ノイズの影響を除去するために、以下のフィルタ処理をする。

$$\bar{\sigma}_j^2 = \frac{1}{\text{Num}(B)} \sum_{\mathbf{m}' \in B} \sigma_j^2(\mathbf{m} + \mathbf{m}'), \quad (12)$$

ここで  $B$  はフィルタのブロック領域である。  $\bar{\sigma}_j^2$  はカメラの対応点の色が一致するほど小さい値をとる。

$q_j(\mathbf{m})$  は一致度が高いほど大きな値をとり、その値は 0 から 1 までの値であり、さらに重なり合う点の合計値は 1 となる ( $\sum_j q_j(\mathbf{m}) = 1$ )。このような  $q_j(\mathbf{m})$  は、例えば以下のように得ることができる。

$$q_j(\mathbf{m}) = \text{norm} \frac{1}{j (\bar{\sigma}_j^2)^n + \epsilon} \quad (13)$$

ここで  $\text{norm}$  は正規化演算を表し、  $\text{norm}_j f_j(x) = f_j(x) / \sum_j f_j(x)$ 。  $n$  は  $q_j(\mathbf{m})$  の急峻さを決定する調整用パラメータであり、  $\epsilon$  は 0 で除算することを防ぐための微小値である。

次に、  $w_{ij}^V(\mathbf{m})$  を推定する。  $w_{ij}^V(\mathbf{m})$  は、視点依存テクスチャマッピング<sup>18)</sup>のカメラ重みパラメータに相当する。近接する光線ほどその色情報が近いという前提により、  $PC_V$  と  $PC_i$  のなす角度  $\theta_{ij}$  により重みが決定し、例えば

$$w_{ij}^V(\mathbf{m}) = \text{norm}_i \exp(-\tau \theta_{ij}^V(\mathbf{m})), \quad (14)$$

となる。ここで  $\tau$  は調整用パラメータである。

## 3 オクルージョンの対策

### 3.1 方針

ここまで、すべてのカメラ  $C_i (i \in N)$  を用いて画像を生成する方法について述べた。しかし、シーンが不連続な奥行きを持つ場合、あるカメラからは撮影できて、あるカメラからは撮影できない箇所がある。このような現象は一般にオクルージョンと呼ばれ、例えば図 2 において、カメラ集合  $\{C_1, C_2, C_3, C_4\}$  を用いてある 1 つの奥行きを特定する奥行きマップを用いる場合、奥行きを推定を失敗する可能性がある。

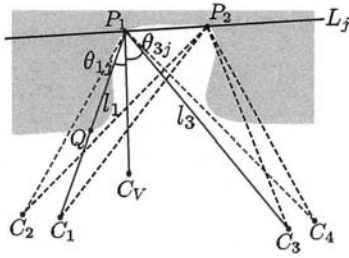


Fig. 2 オクルージョンを含むシーン

そのような場合でも、奥行きが多義性を許す信頼度マップを用いると、奥行きマップを用いる場合よりは滑らかな画像が生成されることが期待できる。しかしそれでも、奥行きが不連続な箇所では生成画像にぼけが生じ、画像が劣化する。

また、たとえ正しい奥行きが推定できた場合であっても、 $P_1$  における色を決定する際に、カメラ  $C_1, C_2$  における対応点の色を用いると、色の再現性が悪くなる。

そこで、カメラを部分集合に分け、それぞれの部分集合ごとに視点依存の推定を行い、最終的に仮想視点で統合することでより信頼性の高い推定を行うことを提案する。

例えば図 2 において、 $P_1$  はカメラ部分集合  $\{C_1, C_2\}$  を用いて推定すると信頼度が低い、カメラ部分集合  $\{C_3, C_4\}$  を用いると信頼度が高まる。逆に、 $P_2$  はカメラ部分集合  $\{C_1, C_2\}$  を用いて推定すると信頼度が高く、カメラ部分集合  $\{C_3, C_4\}$  を用いると信頼度が低い。そして、仮想視点  $V$  からの信頼度に統合し、それぞれのカメラ部分集合の推定のうち、信頼度の高い部分の影響が大きくなるようにする。

### 3.2 部分信頼度マップの推定と統合

以下では、カメラ部分集合ごとの信頼度の推定法について述べる。カメラ部分集合はそのメンバの中に基準となるカメラを持ち、基準カメラの座標系においてレイヤ  $L_j$  における信頼度  $q_j^k(\mathbf{m})$  の推定を行う。 $C_k$  を基準とするカメラ部分集合は一意で、これを  $\Xi_k = \{C_i | i \in N_k\}$  とおく。

図 2 の場合、 $\Xi_1 = \{C_1, C_2\}$  と  $\Xi_3 = \{C_3, C_4\}$  なる 2 つのカメラ部分集合があり、 $C_1$  と  $C_3$  がそれぞれ基準カメラとなる。すべての部分集合における信頼度マップは共通のレイヤ  $L_j$  上で推定さ

れる。

まず、カメラ部分集合  $\Xi_k$  における基準カメラと他のカメラとの不一致度は、SSAD(Sum of Square Absolute Differences) により、以下のように計算される。

$$\text{SSAD} = \sum_{i \in N_k} \sum_{\mathbf{m}' \in B} |I_{ij}^k(\mathbf{m} + \mathbf{m}') - I_k(\mathbf{m} + \mathbf{m}')|. \quad (15)$$

ここで、 $I_k$  は基準カメラ  $C_k$  における画像、 $I_{ij}^k$  は平面  $L_j$  を投影面としてカメラ  $C_i$  の画像を基準カメラ  $C_k$  の座標系に投影変換した画像を示す。すなわち、

$$I_{ij}^k(\mathbf{m}) = I_i(\mathbf{m}_i), \quad \tilde{\mathbf{m}}_i = \mathbf{H}_j^{ki} \tilde{\mathbf{m}}. \quad (16)$$

そして、カメラ部分集合  $\Xi_k$  により推定される基準カメラ  $C_k$  における信頼度  $q_j^k$  は、

$$q_j^k(\mathbf{m}) = \text{norm}_j \frac{1}{(\text{SSAD})^n + \epsilon}. \quad (17)$$

ここで、 $\|\mathbf{x}\|$  は  $\mathbf{x}$  の L1 ノルムをあらわす。また、 $n$  急峻さを調整するパラメータであり、 $\epsilon$  は 0 で除算するのを防止する微小値である。

すべてのカメラ部分集合  $\Xi_k, k \in K$  において信頼度マップ  $q_j^k$  を求めたら、次に仮想カメラ  $C_V$  における統合処理を行う。F まず、仮想視点への座標変換を以下に行い、仮想カメラ  $C_V$  から見たカメラ部分集合  $\Xi_k$  の信頼度マップ  $q_{kj}^V$  を得る。

$$q_{kj}^V(\mathbf{m}) = q_j^k(\mathbf{m}_k), \quad \tilde{\mathbf{m}}_k = \mathbf{H}_j^{V_k} \tilde{\mathbf{m}}. \quad (18)$$

次に、図 3 に示すように、全ての部分集合の信頼度マップ  $q_{kj}^V$  を掛け合わせて、さらにすべての投影面について正規化処理を行い、統合的な信頼度マップ  $q_j^V$  を得る。すなわち、

$$q_j^V(\mathbf{m}) = \text{norm}_j \prod_{k \in K} \{q_{kj}^V(\mathbf{m}) + \text{bias}\}. \quad (19)$$

ここで  $\text{bias}$  は  $\{q_{kj}^V | k \in K\}$  のどれかが極めて 0 に近い数である場合に、掛け算結果が 0 に丸まらないためのバイアスに相当する値である。

### 3.3 カメラ寄与率マップの決定

このようにして  $q_j^V(\mathbf{m})$  を求めたあとは、カメラ寄与率マップ  $w_{ij}(\mathbf{m})$  を求める。式 (14) のように、光線の角度をもとにして求めることも可能ではあるが、オクルージョンのあるシーンの場合、2 つの

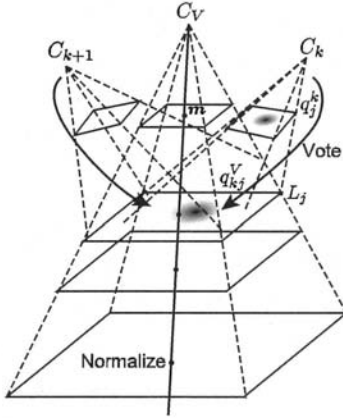


Fig. 3 視点依存信頼度の投票・正規化処理

光線  $C_iP$  と  $C_vP$  のなす角度が小さくても、2つの光線の間には遮蔽物がない保証はない. 図2に示すように、 $\theta_{1j} < \theta_{3j}$  であるが、 $C_1$  からは  $P$  は  $Q$  に遮られており、画像を撮影できていないため、 $w_{1j}$  を大きくするのは誤りである. そこで、着目する点において、高い信頼度を持つカメラ部分集合の基準カメラほど高い寄与率を割り振る.

$$w_{ij}^V(m) = \begin{cases} \text{norm}_k q_{kj}^V(m) | k = i, & \text{if } i \in K \\ 0, & \text{if } i \notin K \end{cases} \quad (20)$$

式(20)について更に論ずる. ある点において信頼度が高いことはその点が存在し、かつ遮蔽されずに撮影している可能性が高いことを示している. また逆に、信頼度が低いことは、その点が遮蔽されているかそもそも点が存在しない可能性が高い. 図2において、 $q_j^3(m_3)$  は光線  $l_3$  上にて正規化されており、 $P$  においては極大値をとることが予想される. 一方で  $q_j^1(m_1)$  は光線  $l_1$  上にて正規化されており、 $Q$  において極大値を持ち、 $P$  においては小さな値をとることが予想される. よって、式(20)によって  $P$  における重みは  $C_3$  が大きく、 $C_1$  が小さく割り振られる. つまり、 $C_1$  について言えば、手前の  $Q$  における大きな値の存在により  $P$  における重み  $w$  が抑制されていると言え、これは一種の可視性チェックとして働く.

### 3.4 レンダリング

このように、 $q_j^V(m)$  および  $w_{ij}^V$  が求めれば、式(8)により仮想視点  $V$  からの画像が生成できる.

式(21)のように重みと射影画像を掛け合わせることで、

$$T_j^V(m) \equiv \sum_{i \in N} w_{ij}^V(m) I_{ij}^V(m), \quad (21)$$

すると、式(8)は以下ようになる.

$$I_V(m) = \sum_{j \in M} q_j^V(m) T_j^V(m). \quad (22)$$

ここで、 $T_j^V(m)$  はレイヤ  $L_j$  の色を表すと考えられる.

## 4 実時間システム

本稿で提案した手法により、撮像から提示まで実時間で処理することができることを実証するため、試作システムを製作した.

システムは9眼のカメラヘッドと13台のPC(撮像用3台、信頼度マップ計算用9台、画像生成・提示用1台)および2次元もしくは3次元ディスプレイで構成される. カメラはIEEE1394によりPCに接続され、PCはお互いがギガビット・イーサにより接続されている.

図5にそのブロック図を示す. システム全体は大きく分けて撮像ユニット、計算ユニット、画像生成ユニットの3つのユニットにより構成される. 撮像ユニットでは同期がとられた9台のカメラからの映像を3台のPCで取り込み、計算ユニットで9台のPCにより分散して信頼度マップを計算し、画像生成ユニットにて任意視点画像もしくはDFD(Depth-Fused 3D)ディスプレイ<sup>25)</sup>へ立体画像を提示する. そして、そのすべての処理はオフライン処理を経ず、実時間で処理される.

## 5 結果

### 5.1 実時間システム

試作機の各ユニットでの所要時間を実測したものを表1に示す. 表中のDFDモードとは、信頼度マップを計算ユニットでDFDの輝度配分データに変換してから画像生成ユニットに送信する場合を示す. ブロックサイズが大きいほど、また、送信するレイヤの枚数が多いほどフレームレートが減少する傾向が見られるが、5~15FPS程度の実時間処理を実現できた.

### 5.2 視点依存信頼度マッピング法による画像生成

半径12cmの円周上に配置した8台のカメラにより、中心の位置から見た画像を生成した. 図6上



パラメータ設定	撮影ユニット		計算ユニット		画像生成ユニット		フレームレート
	撮影	収差補正	評価値計算	信頼度計算	合成	画像生成	
ブロックサイズ 3 × 3, レイヤ数 20, DFD モード	32ms	25ms	10ms	9ms	20ms	17ms	15.0 FPS
ブロックサイズ 3 × 3, レイヤ数 20	32ms	25ms	20ms	9ms	22ms	100ms	6.0 FPS
ブロックサイズ 15 × 15, レイヤ数 30, DFD モード	32ms	25ms	88ms	6ms	21ms	16ms	7.5 FPS
ブロックサイズ 15 × 15, レイヤ数 30	32ms	25ms	88ms	10ms	23ms	148ms	4.0 FPS

Table 1 各処理の所要時間.

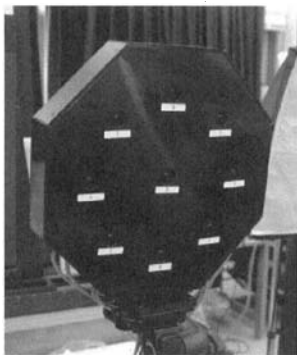


Fig. 4 9眼カメラヘッドの概観.

列にその結果を示す. シーン中, 人物の手前に紙風船が存在し, すべてのカメラから人物の顔が撮影できているわけではない. そのため, すべてのカメラをいちどに用い, 奥行きマップ(左)や信頼度マップ(中央)を求めた場合は画像に欠損が生じていることがわかる. 一方で, 近接する5つのカメラをカメラ部分集合として本稿で提案した手法を用いると, オクルージョン領域においても非常に高い再現性で画像生成できることを確認した.

次に, 筑波大学による多視点画像データセット"ohta"を用い, 定量的な比較を行った. 上下左右4つの元画像から中心の画像を合成し, 実際の中心画像との差をPSNRで評価した. なお, 視点依存信頼度マッピングにおいては, 近接する3つのカメラを部分集合として用いている. その結果, 奥行きマップを用いた方法が31.9dB, 従来の信頼度マッピング法が32.2dBに対し, 本稿で提案した方法は36.9dBと, 大幅な画質の改善を達成した.

## 6 まとめと将来展望

本稿では, 奥行き推定の多義性をレンダリングに利用した新たな画像生成法について述べ, 試作機により5~15FPS程度の3次元画像を実時間で生成できることを実証した.

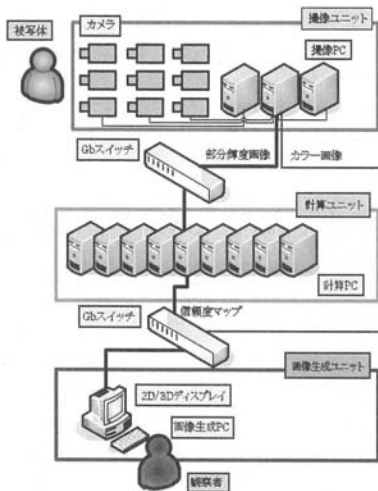


Fig. 5 システムブロック図.

また, 視点依存信頼度マッピング法により, 複数の視点位置から推定した信頼度マップを統合することにより, オクルージョンが生じる環境でも, 高品位な自由視点画像が生成できることを示した.

現在の試作機は, ネットワーク帯域の制限から色情報を中心のカメラのものしか用いておらず, システム構成も大がかりであるため, 今後はより制限が少なく, 小型のシステムを構築したい.

また, 符号化技術<sup>26)</sup>も視野に入れたデータの表現形式の検討も課題である.

## 参考文献

- 1) T. Kanade, P. J. Narayanan, and P. W. Rander: "Virtualized Reality: Concepts and Early Results", In *Proc. of IEEE Workshop on the Representation of Visual Scenes*, pp.69-76 (1995)
- 2) C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szaliski: "High-Quality

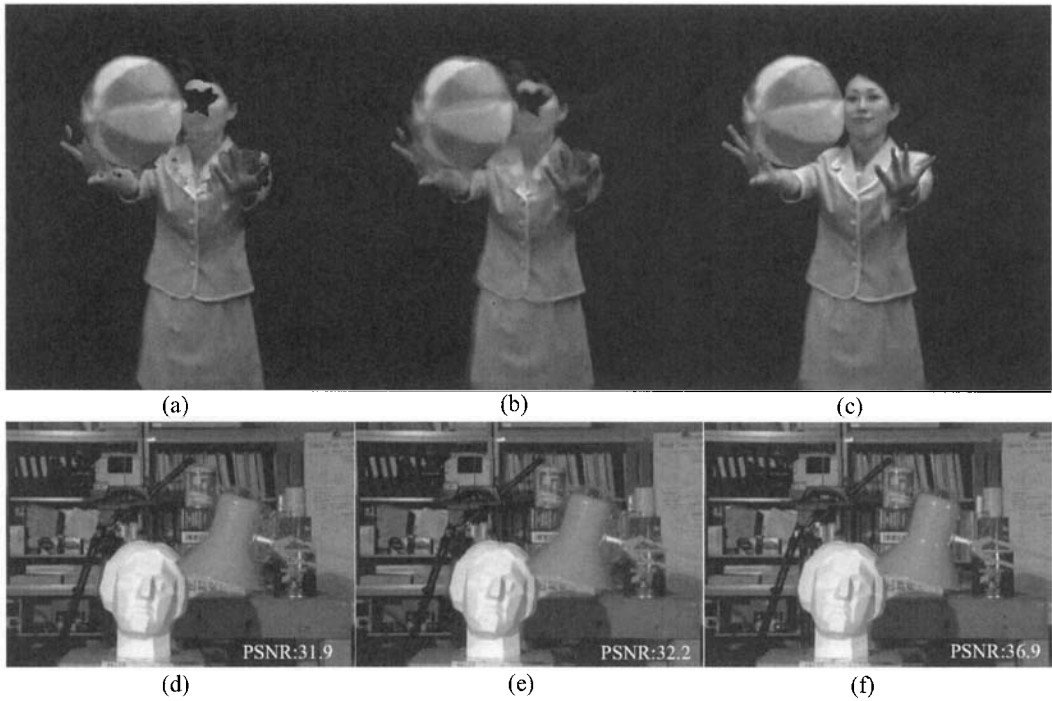


Fig. 6 画像生成結果. 左列：すべてのカメラをいちどに用いて奥行きマップを求めた場合. 中央列：すべてのカメラをいちどに用いて信頼度マップを求めた場合. 右列：視点依存信頼度マッピング法 (提案手法)

- Video View Interpolation Using a Layered Representation”, In *Proc. of SIGGRAPH 2004*, 600–608 (2004)
- 3) W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan: “Image-Based Visual Hulls”, In *Proc. of SIGGRAPH 2000*, pp.369–374 (2000)
  - 4) M. Li, M. Magnor, and H.-P. Seidel: “Hardware-Accelerated Visual Hull Reconstruction and Rendering”, In *Proc. of Graphics Interface 2003*, 65–71 (2003)
  - 5) 松山, 高井, ウ, 延原: ”3次元ビデオ映像の撮影・編集・表示”, 日本バーチャルリアリティ学会論文誌, Vol.7, No.4, pp.521-532 (2002)
  - 6) 富山, 片山, 岩館, 今泉: ”視体積交差法とステレオマッチング法を用いた多視点画像からの3次元動オブジェクト生成手法”, 映情学誌, **58**, 6, pp.797-806 (2004)
  - 7) G. G. Slabaugh, R. W. Schafer, and M. C. Hans: “Image-Based Photo Hulls for Fast and Photo-Realistic New View Synthesis”, *Real-Time Imaging (RTI)*, **9**, 5, 347–360 (2003)
  - 8) M. Li, M. Magnor, and H.-P. Seidel: “Hardware-Accelerated Rendering of Photo Hulls”, *Computer Graphics Forum (special issue Eurographics’04)*, **23**, 3, 635–642 (2004)
  - 9) E. H. Adelson and J. Bergen: “The plenoptic function and the elements of early vision,” *Computational Models of Visual Processing*, M. Landy and J. A. Movshon, Ed., pp. 3-20, MIT Press, Cambridge (1991)

- 10) M. Levoy and P. Hanrahan: "Light Field Rendering" In *Proc. of SIGGRAPH '96*, pp. 31-41 (1996)
- 11) S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen: "The Lumigraph", In *Proc. of SIGGRAPH '96*, pp.43-54 (1996)
- 12) 苗村, 金子, 原島: "光線情報の補間に基づく Ray-Based Rendering", 3次元画像コンファレンス'96, pp.7-12 (1996)
- 13) P. N. Banchang, M. P. Tehrani, T. Fujii, and M. Tanimoto: "Realtime System of Free Viewpoint Television", 映情学誌, **59**, 8, 1191-1198 (2005)
- 14) 石川, 酒澤, 小池: "局所領域分割型光線空間を用いたウォークスルー映像の生成", IMPS 2006, pp.19-20 (2006)
- 15) Y. Kunita, M. Inami, T. Maeda, and S. Tachi: "Realtime Rendering System of Moving Objects", In *Proc. of the 1999 IEEE Workshop on Multi-View Modeling & Analysis of Visual Scenes (MVIEW'99)*, pp.81-88 (1999)
- 16) J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum: "Plenoptic Sampling", In *Proc. of SIGGRAPH 2000*, pp.307-318 (2000)
- 17) 高橋, 苗村, 原島: "自由視点画像合成における光線情報のサンプリングと被写界深度に関する理論", 映情学誌, **57**, 10, pp.1292-1299 (2003)
- 18) P. E. Debevec, C. J. Taylor, and J. Malik: "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach", In *Proc. of SIGGRAPH '96*, pp.11-20 (1996)
- 19) 山本, 奥富: "ボリウムレンダリングを利用したイメージベーストレンダリングの試み", 画像電子学会誌, Vol.30, No.4, pp.371-378, 2001
- 20) 國田, 橋本, 木村, 中沢: "信頼度を利用した多層平面レンダリング法の提案", 3次元画像コンファレンス 2004, pp.135-138 (2004)
- 21) 國田, 上野, 田中: "多層信頼度マップを用いた3次元映像の実時間生成システム", 映情学誌, **60**, 7, pp.1102-1110 (2006)
- 22) 久保田, 児玉, 羽鳥: "スペクトル分解に基づいたシーン非依存型の視点内挿フィルタ", In *Proc. of IMPS 2007*, pp.117-118 (2007)
- 23) J. Shade, S. Gortler, L.-W. He, and R. Szeliski: "Layered Depth Images", In *Proc. of SIGGRAPH '98*, pp.231-242 (1998)
- 24) 國田, 上野, 田中: "遮蔽領域の再現性を向上させた視点依存信頼度マッピング法" 第5回情報科学技術フォーラム講演論文集, pp.163-166 (2006)
- 25) S. Suyama, S. Otsuka, H. Takada, K. Uehira, and Sakai S.: "Apparent 3-D Image Perceived from Luminance-modulated Two 2-D Images Displayed at Different Depths", *Vision Research*, **44**, pp.785-793 (2004)
- 26) 木全, 北原, 上倉, 八島, 藤井, 谷本: "自由視点映像通信のための多視点符号化", 信学論D, vol.J89-D, no.1, pp.40-55 (2006)