

複合現実感のための3次元モデルに基づく見えの圧縮

西野 恒 佐藤 洋一 池内 克史

東京大学 生産技術研究所
東京大学大学院 工学系研究科 電子情報工学専攻
{kon,ysato,ki}@cvl.iis.u-tokyo.ac.jp

あらまし: 現在まで複合現実感システムの研究において, 物体の幾何形状を表す3次元モデルと物体表面反射モデルを用いた解析的な手法や, 実画像を圧縮して蓄積する Imaged-based Rendering の手法などが提案されてきた. しかし, これらの手法は適用できる物体表面に強い制約があったり, 物体の精確な3次元モデルを持たないために正しい陰影をつけられないなど, 現実画像への重ね合わせが困難であるなどの問題を持つ. 本論文ではこれらの問題点を踏まえ, 新たな手法として物体上で定義された座標空間において実画像の圧縮をおこない, 仮想物体画像の合成をおこなう手法を提案する. また, 本研究で提案する手法を用いて実物体を用いた実験をおこない, 提案する手法が有効であることを確かめる.

キーワード: イメージベースドレンダリング, 固有空間法, 複合現実感

Appearance Compression based on 3D Model for Mixed Reality

Ko Nishino Yoichi Sato Katsushi Ikeuchi

Institute of Industrial Science
University of Tokyo
Dept. of Information and Communication Engineering
Graduate School of the University of Tokyo
{kon,ysato,ki}@cvl.iis.u-tokyo.ac.jp

Abstract: Recent investigations of mixed reality systems have proposed image-based rendering approaches and techniques that use 3D models and reflection models. However, few image-based rendering approaches use a 3D model, which is necessary to produce a correct shadow; also, there are strong constraints on most of the analytic approach applications. Considering these problems, we propose a new method which compresses the appearance of a real object, using a set of real images and a 3D model, and then uses the compressed data to reconstruct the image of a virtual object. To demonstrate that our method is effective, we cite an experiment.

Keyword: image-based rendering, eigenspace representation, mixed reality

1 はじめに

近年、複合現実感などの言葉に代表されるように、電子的に構築された仮想空間と実世界空間の融合に関する研究が盛んにおこなわれている。この際、仮想空間と実世界空間を境目なく融合させ、より現実感を高めることが大きな目的のひとつになる。

現実感を高める要素としてはいくつか考えられるが、特に物体表面における光の反射は、合成された仮想物体画像の見える現実感の高さに重要な影響を及ぼす。このような物体表面における特性を数値的に解析することはせず、仮想物体画像を合成する際にあらかじめ撮像された実画像を用いる Imaged-based Rendering の手法はこれまで盛んに研究がおこなわれてきた。これは実物体をさまざまな光源状況下や姿勢で撮像した実画像をデータベースとして保持しておき、特定の姿勢や光源状況下の仮想物体画像を合成する際にこのデータベースから正しい画像を取り出したり、補間により合成するという手法である [9][10]。

しかし、IBR の手法は元来物体形状の 3 次元モデルを積極的に利用せずに仮想空間を構築する目的ですすめられていることもあり、より現実感の高い仮想空間と実世界空間の融合をおこなうためには必要となる、精確な物体の 3 次元モデルを利用する提案は報告例が少ない。そのため、松井らにより IBR の手法による照明環境を考慮した仮想空間と実世界空間の融合の試みなどの研究 [14] がなされているものの、依然 3 次元モデルを持たないという点で、複合現実感システムにおいて IBR の手法を用いるには正しい陰影がつけられないなどの問題が残る。

一方、物体表面における光の反射を解析し、その数学的モデルを構築する研究は以前から盛んにおこなわれてきた。Shafer により提案された二色性反射モデル Dichromatic Reflection Model [8] は物体表面反射が body reflection と surface reflection の 2 つの反射成分により構成されることを示すものであり、その後の物体表面反射モデルの基礎をなす重要なモデルとなっている。このような物体表面反射モデルに基づき、上述の 2 つの反射成分

の分離、また各反射成分の数値を推定する研究が進められた [1][2][12]。

物体表面反射特性に加え、仮想物体画像の現実感を高めるためには仮想物体に実世界空間の光源状況に応じた精確な陰影をつけることも重要な要素となる。このためには物体の 3 次元モデルを利用することが必要になってくる。Sato らはカメラに対し物体を回転させて撮像した複数枚の入力画像を物体の 3 次元モデルに対応付け、物体表面の各点における反射成分のパラメータを物体表面反射モデルを用いて推定する研究をおこなった [13]。これにより物体表面のすべての点において反射成分のパラメータを保持し、さらに物体の幾何形状も保持しているため、非常に精度の高い仮想空間と実世界空間の融合が可能となっている。

しかし、物体表面反射モデルを用いて反射成分のパラメータを推定する手法は、物体表面の各点における面法線を必要とするため、ざらざらしているような粗い表面を持つ物体に対して適用するのが困難である。さらに、Sato らの研究では Dichromatic Reflection Model を用いているため、照明光と同一の色を body reflection として持つ物体表面の領域に関しては正しく反射成分の分離と反射成分のパラメータを推定できない。また、入力画像となる画像に非常に明るいハイライトが入ることによりカメラの特性の限界を越えた画素が存在すると、正しく反射成分のパラメータを推定できないという問題点も持つ。

本研究ではこれらの IBR の手法や物体表面反射モデルを用いた手法の問題点を踏まえ、これらの手法に代わる新たな手法として 3 次元モデルに基づく物体の見えるの圧縮をおこなうことにより、蓄積された実画像を効率よく利用して仮想物体画像を合成する手法の提案をおこなう。

以降、2 章で本研究で提案する手法の入力となる距離画像と明るさ画像列の撮像をおこなうシステムについて述べ、3 章で本研究において提案する「3 次元モデルに基づく見えるの圧縮」の具体的手法について述べる。続いて 4 章で今回提案する手法を用いて実物体を用いておこなった実験について述べ、本手法の有効性を議論する。最後に全体を総括し、まとめとする。

2 3次元モデルと明るさ画像列

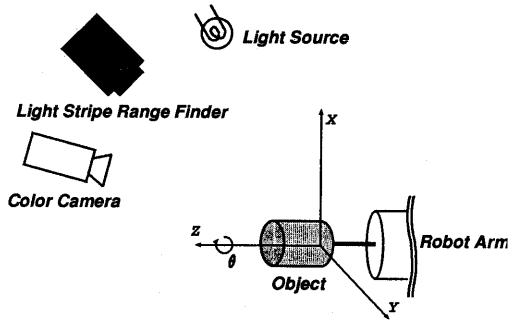


図 1: 実験装置

本研究で提案する手法の入力となる実画像である明るさ画像列の撮像法、物体の幾何形状を表す3次元モデルの作成法について述べる。

図1に示す実験装置を用いて、物体形状モデルを構築するための一連の距離画像と明るさ画像の連続画像列を撮像する。物体は5自由度のロボットの手の先に取りつけられ、Z軸を中心に θ の1自由度で回転させる。光源は単一の点光源であり、物体の座標系に対して固定している。

物体の3次元モデルを作成するための一連の距離画像は図1のカラーカメラおよび液晶シャッターで構成される光線パターン投影型レンジセンサを用い[5]、ロボットの1つの軸を用いてカメラに対し θ の角度を順次変えながら撮像をおこなう。後述の実験では θ を $0^\circ \sim 360^\circ$ の範囲で 30° ずつ回転させて撮像した12枚の距離画像をそれぞれ三角パッチの集合であるメッシュに変換し、すべてを統合した上で物体の幾何形状をそこなわない程度にメッシュを単純化したものを物体の3次元モデルとして使用した(図2)[3][6]。

明るさ画像は距離画像と同様に、カメラに対しロボットの1つの軸を用いて θ の角度を順次変えることにより物体の姿勢を変化させて撮像する。この際、本研究で提案する手法では物体表面の見えそのものを圧縮して仮想物体画像の合成に用いるため、サチュレーションを気にする必要はない。後述の実験では θ を $0^\circ \sim 360^\circ$ の範囲で 3° ずつ回転させて撮像しており、わざとハイライトにおいて

サチュレーションを起こすように光源を明るくしてある(図3)。

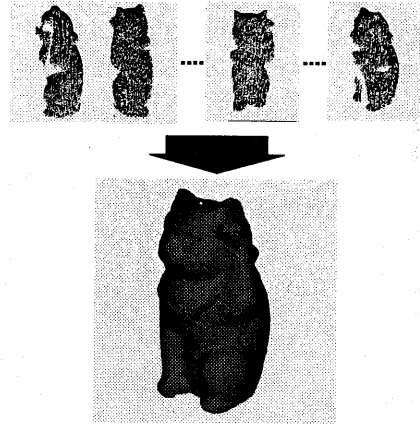


図 2: 距離画像セットと3次元モデル

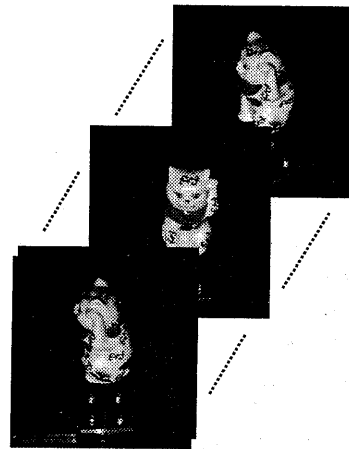


図 3: 明るさ画像列

図2に示す物体形状モデルに対し、各姿勢の明るさ画像を貼り付けることにより物体形状モデルの各三角パッチと明るさ画像内の小領域の対応が取れ、図4に示すように仮想物体画像を合成することができる。しかし、各姿勢の見えを再現するためには各姿勢の明るさ画像をすべて保持する必要がある、そのデータ量は膨大になる。 3° ずつ姿勢を変化させた場合で 480×512 画素の画像を120

枚, 約 90MB のデータ量となる. より細かく姿勢を変化させた見えを再現する場合にはこれ以上のデータを保存する必要がある, おのずとこの手法には限界があり, 保持する画像データをなんらかの手法により圧縮する必要があることがわかる.

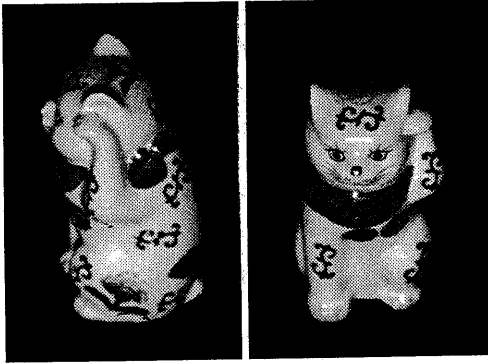


図 4: 実連続画像セットを用いて合成された仮想物体画像

3 3次元モデル上での見えの圧縮

本研究で提案する, 入力となる実画像の見えを 3次元モデルに基づいて圧縮して扱い, 仮想物体画像を合成する手法について述べる.

前節で述べた実験装置を用いて撮像された実物体の各姿勢の明るさ画像をデータとして保持し, 別個に撮像された距離画像により作成した物体の幾何形状を表す 3次元モデルに適宜対応づけることにより 3次元の仮想物体画像を合成することが可能である (図 4). しかし, この際データとして保持しておかなければならない実画像のデータは膨大な量になる. そこで, 本研究で提案する手法ではまず入力となる実画像を 3次元モデルと対応づけた上で, 保持すべきデータ量を減らすべく圧縮をおこなう. 具体的には, まず前節で述べた実験装置により撮像された各姿勢の明るさ画像を 3次元モデルと対応を取り, 各三角パッチに対応する小領域に明るさ画像を分割する (以後この小領域をセルと呼ぶ). 次に各セルを同じ大きさに正規化した上で, 各姿勢に対応した明るさ画像をすべて

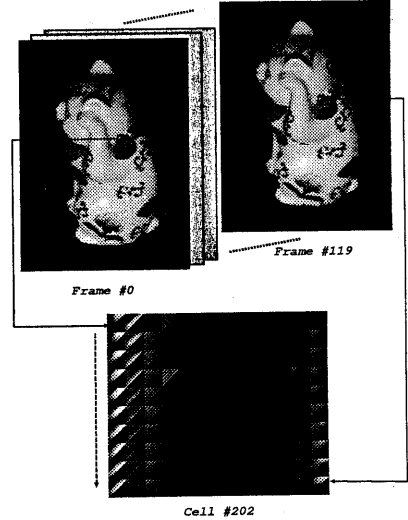


図 5: セル画像

各セル毎の画像に並べ換える. その結果図 5に示すように, 各姿勢に対応するセルの見えを並べた各セルの画像を得る. このセル画像に対し固有空間法を用いて圧縮をおこなう. 圧縮をおこなう対象は見えの変化が明るさの変化だけであるセル画像であり, セル画像内の各姿勢のセルは非常に相関が高いため, 固有空間法による圧縮により非常に高い圧縮率が得られることが予想される. また, 固有空間においてさまざまな補間がおこなえるなどの点により圧縮法として固有空間法を用いた.

各セル画像に対する固有空間法による圧縮は以下の手順でおこなう. まず, RGB 各バンド毎にセル画像の 1つの姿勢に対応するセルの画素値をラスタスキャンし, 式 (1) に示す $1 \times N$ 行ベクトルを得る. これらを行方向に並べた式 (2) に示す $M \times N$ 行列 (M : 1つのセルが持つ姿勢の数, N : 1つの姿勢に対応するセル内の画素数) が図 5に示した 1枚のセル画像に対応する.

$$\mathbf{X}_m^{(c)} = [x_{m,1}^{(c)} \quad x_{m,2}^{(c)} \quad \dots \quad x_{m,N}^{(c)}] \quad (1)$$

$$\mathbf{X}^{(c)} = [\mathbf{X}_1^{(c)} \quad \mathbf{X}_2^{(c)} \quad \dots \quad \mathbf{X}_M^{(c)}]^T \quad (2)$$

ただし, $c = R, G, B$

ここで, 最大の固有値を持つ固有ベクトルが画

像の変化が相対的に最大になる固有空間の次元を表すことを保証するために、すべての姿勢に対応するセルのすべての画素値の平均値 $E^{(c)}$ を各画素値から引いた行列 $\mathbf{X}_{ave}^{(c)}$ を作る (式 (3)).

$$\mathbf{X}_{ave}^{(c)} = \begin{bmatrix} x_{1,1}^{(c)} - E^{(c)} & x_{1,2}^{(c)} - E^{(c)} & \dots & x_{1,N}^{(c)} - E^{(c)} \\ x_{2,1}^{(c)} - E^{(c)} & x_{2,2}^{(c)} - E^{(c)} & \dots & x_{2,N}^{(c)} - E^{(c)} \\ \dots & \dots & \dots & \dots \\ x_{M,1}^{(c)} - E^{(c)} & x_{M,2}^{(c)} - E^{(c)} & \dots & x_{M,N}^{(c)} - E^{(c)} \end{bmatrix} \quad (3)$$

$$E^{(c)} = \frac{1}{MN} \sum_{i=1, j=1}^{M, N} x_{i,j}^{(c)}$$

この $M \times N$ 行列 $\mathbf{X}_{ave}^{(c)}$ を用いて、式 (4) に示す新たな $N \times N$ 行列 $\mathbf{Q}^{(c)}$ を定義する.

$$\mathbf{Q}^{(c)} = \mathbf{X}_{ave}^{(c)T} \mathbf{X}_{ave}^{(c)} \quad (4)$$

固有値 λ_i , 固有ベクトル \mathbf{e}_i はこの $N \times N$ 行列 $\mathbf{Q}^{(c)}$ について求める.

$$\lambda_i^{(c)} \mathbf{e}_i^{(c)} = \mathbf{Q}^{(c)} \mathbf{e}_i^{(c)} \quad (5)$$

この際、行列 $\mathbf{Q}^{(c)}$ の固有空間は N 次元の多次元空間となる。そこで行列 $\mathbf{Q}^{(c)}$ の特徴を適度に正確に表わす k ($k \leq N$) 本の固有ベクトルを抽出して用いることにより実質的な圧縮をおこなう。このような固有ベクトルを抽出する方法の1つとして行列 $\mathbf{Q}^{(c)}$ の固有ベクトルのうち対応する固有値の大きいものから順に k 個選択することが考えられる。行列 $\mathbf{Q}^{(c)}$ の各要素は非負であるため、その固有値はすべて非負である。そのため固有値を降順に並べ換え、寄与率を計算することにより (式 (6)) セル画像の最も重要な特徴の変化を表わす k 個の固有値、それらの固有値に対応する固有ベクトルを抽出することができる。

$$\frac{\sum_{i=1}^k \lambda_i^{(c)}}{\sum_{i=1}^N \lambda_i^{(c)}} \geq T \quad (6)$$

ただし、 $T \leq 1$

以上の手順により得られた行列 $\mathbf{Q}^{(c)}$ の上から k 個の固有値 $\{\lambda_i^{(c)} \mid i = 1, 2, \dots, k\}$ ($\lambda_1 \geq \lambda_2 \dots \geq$

λ_k) に対応する k 本の固有ベクトル $\{\mathbf{e}_i^{(c)} \mid i = 1, 2, \dots, k\}$ (ただし、 \mathbf{e}_i は $N \times 1$ の列ベクトルとする) を用いて、行列 $\mathbf{X}_{ave}^{(c)}$ を固有空間に射影することにより、セル画像を固有空間に射影した $M \times k$ 行列 $\mathbf{G}^{(c)}$ を得ることができる (式 (7)).

$$\mathbf{G}^{(c)} = \mathbf{X}_{ave}^{(c)} \times \mathbf{V}^{(c)} \quad (7)$$

ただし、 $\mathbf{V}^{(c)} = [\mathbf{e}_1^{(c)} \quad \mathbf{e}_2^{(c)} \quad \dots \quad \mathbf{e}_k^{(c)}]$

入力となる実画像の連続画像セットは、以上のようにセル画像の形式に変換し、 $M \times k$ 行列 \mathbf{G} と $N \times k$ 行列 \mathbf{V} の形で蓄積する。これらの行列は1枚のセル画像に対し RGB 各バンドについて保持する必要があるが、情報量としては約 k/N の圧縮率となる (一般的に $M \ll N$ であるため \mathbf{G} の大きさは微小である)。

各セル画像を上記の k 本の固有ベクトルを用いて復元した仮想物体画像は、固有空間への射影を表す $M \times k$ 行列 $\mathbf{G}^{(c)}$ と固有ベクトルを表す $N \times k$ 行列 $\mathbf{V}^{(c)}$ を用いた計算により式 (8) に示す $M \times N$ 行列 $\mathbf{R}^{(c)}$ として求めることができる。

$$\mathbf{R}_m^{(c)} = \sum_{i=1}^k g_{m,i} \mathbf{e}_i^{(c)T} + E^{(c)}$$

$$\mathbf{R}^{(c)} = [\mathbf{R}_1^{(c)} \quad \mathbf{R}_1^{(c)} \quad \dots \quad \mathbf{R}_M^{(c)}]^T \quad (8)$$

4 実験

本研究で提案する手法の有効性を示すために実物体を用いて実際に3次元モデルに基づく見えの圧縮をおこない、圧縮された画像を用いて仮想物体画像の合成をおこなった。その際の圧縮率や、固有空間法を用いる際の次元の設定について述べる。

実験には図 2,3 に示した伊万里焼きの招き猫を使用した。距離画像は図 1 の角度 θ を $0^\circ \sim 360^\circ$ の範囲で 30° ずつ回転させて撮像し、明るさ画像は θ を $0^\circ \sim 360^\circ$ の範囲で 3° ずつ回転させて撮像した。また、明るさ画像を撮像する際にはわざと照明光の明るさを大きくし、カメラのダイナミックレンジの範囲を越える明るさを持つ画素ができるようにした。

4.1 固定された次元

RGB 各バンドについてそれぞれ前章で示した方法により求めた固有ベクトルを用い、固有空間の次元数を固定して(すべてのセルについて同じ本数の固有ベクトルを用いて)画像を合成し3次元モデルに投影した結果を図6,7に示す。

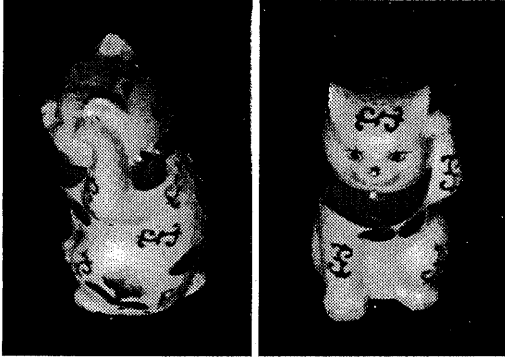


図 6: 次元数 1

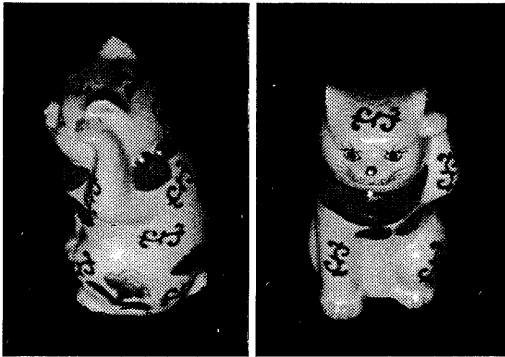


図 7: 次元数 3

固有空間の次元を1次元に固定して合成した画像は、対応する固有値が最も大きい固有ベクトルだけを用いて仮想物体画像の合成をおこなっているため、各セルのおおまかな特徴は表現できているものの、入力画像である実画像(図4)には程遠い精度でしか復元がおこなえていない。

一方、固有空間の次元を3次元に固定して画像を合成した場合、ハイライトの入り方、細かなテク

スチャの見え方等々、原画像(図4)とほぼ見分けのつかない程度に精度良く画像が合成されていることがわかる。本手法では、特徴の変化が明るさの変化のみに限定される非常に相関の高いセル画像に対し実画像の圧縮をおこなっている。そのため、この結果からも高々2.5%の非常に高い圧縮率で実画像と見分けのつかない程度に精度の高い仮想物体画像の合成がおこなえることが示された。

4.2 次元の最適化

前節では各セルを各バンドに対し、すべて同じ次元数を用いて固有空間法による圧縮をおこなう手法について述べた。しかし、実際には各セルの画像はテクスチャの色や細かさ、ハイライトの入り方等々の要因により、精度良く再生するために必要な次元数が各バンド毎に異なり、各セル毎にも異なる。そこで次元数を各セル各バンドに対して固定する代わりに、各セルの各バンドがそれぞれ一定の精度で合成画像の復元がおこなえるような次元の最適化をおこなった。すなわち、各セルの各バンドに対して求められる固有値が、式(6)によって計算される寄与率が特定の閾値を越えるものに対応する固有ベクトルを用いて仮想物体画像の合成をおこなった。

この手法を取ることで、精度良く合成するために必要な固有空間の次元数の高いセルに関しては次元数を増やし、反対に低次元で精度の良い合成をおこなえるセルに関しては次元数を減らすという、各セル各バンドに応じた固有空間の次元の最適化がおこなえる。このため、入力画像となる実画像の効率的な圧縮と蓄積がおこなえる。これは、蓄積すべき実画像をセルという小さな領域に分割して取り扱っているためにおこなえることであり、本手法の特徴の一つである。

各セルの各バンドの k 個の固有値を用いた寄与率が0.99以上になるように次元数 k を定め、各々に対応する k 本の固有ベクトルを用いて再生画像を合成し3次元モデルに投影した結果を図8に示す。寄与率0.99以上を満たす次元数は表1のようになる。

すべてのセルが寄与率0.99以上で復元されてい

るため、合成された仮想物体画像は入力となる実画像(図4)と見分けのつかない程度に精度良く合成されている。また、この際の圧縮に用いる固有空間の次元数は全セルの全バンドについての平均が約3.5次元となり、全体の圧縮率は約3%となる。このことより、本研究で提案した手法により、実画像を非常に高い圧縮率で蓄積し、そのデータを用いて精度の高い仮想物体画像の合成がおこなえることが示された。

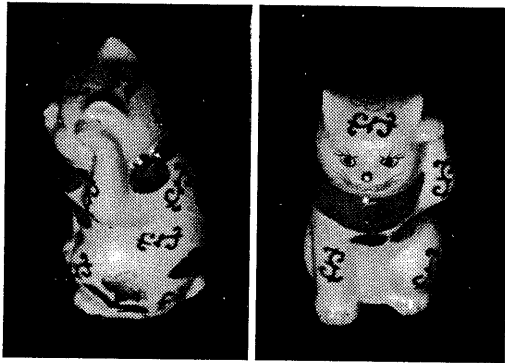


図 8: 次元を最適化した再生画像

	R	G	B	全バンド
平均次元数	2.92	3.29	4.42	3.54
平均圧縮率 (%)	2.43	2.74	3.68	2.95

表 1: 平均次元数と平均圧縮率

5 おわりに

本研究では、従来実物体の実画像を入力として用い仮想物体画像を合成する際に用いられていた、物体表面反射モデルを用いた手法や Imaged-based Rendering の手法の問題点を考慮し、物体の幾何形状を表す3次元モデルを積極的に利用し、3次元モデルに基づく物体の見えの圧縮をおこなうことにより、入力となる実画像を効率よく利用して仮想物体画像を合成する手法を新たに提案した。また、実物体を用いた実験を示すことにより本手法

が有効であることを示した。

特に今回示した実験では、実物体としてカメラの特性の限界を越える明るさのハイライトが入り、かつ物体の body reflection の色が照明光の色とほぼ同一であるような物体を用いることにより、従来の物体表面反射モデルを用いて数値的解析をおこなう手法では正しく仮想物体画像を合成をおこなえない実物体についても本手法が有効であることを示した。

今後は、従来の手法では仮想物体画像の合成が難しかったざらざらな表面を持つ実物体などのさまざまな実物体に対して本手法を適用し、その有効性を議論したい。また、実物体の姿勢だけではなく光源状況の変化も入力画像として用い、実空間の光源状況に応じた仮想物体画像と実画像の重ね込みの研究を進める予定である。

参考文献

- [1] G.J.Klinker, S.A.Shafer, and T.Kanade. The measurement of highlights in color images. *International Journal of Computer Vision*, Vol. 2, pp. 7-32, 1988.
- [2] G.J.Klinker, S.A.Shafer, and T.Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, Vol. 4, pp. 7-38, 1990.
- [3] G.Turk and M.Levoy. Zippered polygon meshes from range images. *Computer Graphics Proceedings, Annual Conference Series 1994, ACM SIGGRAPH 94*, pp. 311-318, 1994.
- [4] H.Murase and S.K.Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, Vol. 14, pp. 5-24, 1995.
- [5] K.Sato and S.Inokuchi. Range-imaging system utilizing nematic liquid crystal mask. *First International Conference on Computer Vision, IEEE*, pp. 657-661, 1987.

- [6] M.D.Wheeler, Y.Sato, and K.Ikeuchi. Consensus surfaces for modeling 3D objects from multiply range images. *Sixth International Conference on Computer Vision, IEEE*, pp. 917-924, 1998.
- [7] M.Levoy and P.Hanrahan. Light field rendering. *Computer Graphics Proceedings, Annual Conference Series 1996, ACM SIGGRAPH 96*, pp. 31-42, 1996.
- [8] S.A.Shafer. Using color to separate reflection components. *COLOR Research and Application*, Vol. 10, No. 4, pp. 210-218, 1985.
- [9] S.E.Chen. Quicktime VR: An image-based approach to virtual environment navigation. *Computer Graphics Proceedings, Annual Conference Series 1995, ACM SIGGRAPH 95*, pp. 29-38, 1995.
- [10] S.E.Chen and L.Williams. View interpolation for image synthesis. *Computer Graphics Proceedings, Annual Conference Series 1993, ACM SIGGRAPH 93*, pp. 279-288, 1993.
- [11] S.J.Gortler, R.Grzeszczuk, R.Szeliski, and M.F.Cohen. The lumigraph. *Computer Graphics Proceedings, Annual Conference Series 1996, ACM SIGGRAPH 96*, pp. 43-54, 1996.
- [12] Y.Sato and K.Ikeuchi. Temporal-color space analysis of reflection. *Journal of the Optical Society of America*, Vol. 11, No. 11, pp. 2990-3002, 1994.
- [13] Y.Sato, M.D.Wheeler, and K.Ikeuchi. Object shape and reflectance modeling from observation. *Computer Graphics Proceedings, Annual Conference Series 1997, ACM SIGGRAPH 97*, pp. 379-387, 1997.
- [14] 松井明, 佐藤宏介, 千原國宏. 多光源画像のKL展開と環境観測に基づく実照明と仮想照明の複合. 電子情報通信学会技術研究報告, Vol. 97, No. 324, pp. 29-36, 10 1997.