

視体積交差法を用いた3Dモデルに基づく任意視点カラー画像生成

秦, 真也
九州大学システム情報科学研究院知能システム学部門

有田, 大作
九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎
九州大学システム情報科学研究院知能システム学部門

<https://hdl.handle.net/2324/5913>

出版情報 : 火の国情報シンポジウム, pp.68-75, 2003-03
バージョン :
権利関係 :

視体積交差法を用いた 3 D モデルに基づく

任意視点カラー画像生成

秦 真也, 有田 大作, 谷口 倫一郎

九州大学大学院システム情報科学府

〒 816-8580 福岡県春日市春日公園 6-1

TEL: 092-583-7618

E-mail: {hata,arita,rin}@limu.is.kyushu-u.ac.jp

あらまし コンピュータビジョンの分野において、多視点画像から任意視点画像を生成することで、教育、医療、娯楽などの様々な分野への応用が可能となる。本稿では、利用価値の高いシステムを実現するために、人体のように複雑な形状を持つモデルを対象として、多視点画像から精度の高い任意視点カラー画像を生成する手法について述べる。提案手法では、まず各カメラにおいて撮影された視点において、影領域の切り出しを目的とするロバストなシルエット画像の抽出を行う。得られたシルエット画像から、視体積交差法を用いて 3 D モデル形状を仮想空間上に復元する。さらに、3 D モデルを基に任意視点カラー画像を生成し、その精度を評価することで本手法の有効性を示す。

キーワード 視体積交差法, 3 D モデル復元, 任意視点カラー画像

Arbitrary Viewpoint Color Image Generation Based on 3D Model Using Visual Cone Intersection

Shinya Hata, Daisaku Arita and Rin-ichiro Taniguchi

Graduate School of Information Science and Electrical Engineering, Kyushu University

6-1, Kasuga-kouen, Kasuga, Fukuoka, 816-8580 Japan

TEL : 092-583-7618

E-mail: {hata,arita,rin}@limu.is.kyushu-u.ac.jp

Abstract In the field of computer vision, generating an arbitrary viewpoint image from multi viewpoint images enables various applications such as education, medical treatment, and amusement. In this paper, we propose a method to generate an arbitrary view color image for a complicated shape object such as a human body from multi-viewpoint images. Since our method is able to detect shadow region, an object silhouette is robustly extracted from each input image. And, a 3D model is recovered in virtual space from silhouette images, using visual cone intersection. Furthermore, we generate a arbitrary viewpoint color image based on the 3D model, and we show the validity of our method by evaluating the color image.

key words visual cone intersection, 3D model recovery, arbitrary viewpoint color image

1. はじめに

近年、コンピュータビジョンの分野において、計算機の処理能力の向上に伴い、複数のカメラによって撮影された多視点画像を解析する研究が盛んに行われている。中でも、多視点入力画像から精度の高い任意視点画像を生成する研究は、教育、医療、娯楽など、幅広い分野への応用を可能とする。

多視点入力画像から任意視点画像を生成する手法としては、一般的に二つのアプローチがある。一つは多視点入力画像から対象物の3D座標を求めることで、3Dモデルを仮想空間上に復元して、その3Dモデルを基に任意視点画像を生成するモデルベースの手法である。もう一つは、View Morphing^[1]に代表されるように入力画像から対象物の3D座標を実際には求めることなく、直接的に任意視点画像を生成するイメージベースの手法である。両手法の特徴を比較すると、イメージベースの手法は、出力する視点についての制約が大きいという問題点がある。それに対してモデルベースの手法は、3D座標を獲得しているため、あらゆる位置の視点から出力画像を表示することが可能である。したがって、本手法ではモデルベースの手法を採用する。

現在、PCクラスタを用いた並列処理を行なうことで、視体積交差法を用いた3Dモデルの復元、任意視点画像の生成を実時間で実現することは可能であり、高速化については様々な研究がなされている^[2]。しかし、復元された3Dモデルは、形状の精度が低いという欠点があり、たとえ視点数を増やしてもモデルの形状を向上させることができなかった。このように、モデル形状の精度に対して悪影響を及ぼす原因として、各視点ごとに行われるシルエット抽出処理の精度が低い、カメラキャリブレーションの精度が低いという点が挙げられる。

そこで本稿ではまず、単純な背景差分を行なってシルエットを抽出するのではなく、明るさ情報を考慮することで、精度の高いシルエットをロバストに抽出する手法を提案する。そして弱校正を用いたシステム上で、人体のように複雑な形状を持つモデルを対象として、多視点画像

から3Dモデル形状を仮想空間上に復元する。その上で本稿では、復元された3Dモデル形状を基にして任意視点画像への投影を行なう。このとき、カメラにより撮影された入力画像の色情報を効果的に利用することによって、任意視点からの出力画像に色情報を付加する。

本稿の構成は以下の通りである。まず、2章で明るさを考慮したロバストなシルエットの抽出法について述べる。次に3章で、多視点カメラから得られた画像の情報のみを用いて画像間の幾何関係を求める、弱校正による3Dモデル復元システムについて述べ、実装法を示す。さらに4章では、3Dモデル形状と入力画像の色情報を用いて、効果的に任意視点画像のカラー表示を行なうアルゴリズムについて述べる。そして5章において、人体モデルを対象とした多視点実画像に本手法を適用して実験を行ない、その結果を示すことで本手法の有効性を評価する。

2. シルエットの抽出法

シルエット抽出は、一般に背景差分によって行われている。背景差分とは、対象物を含む入力画像と、あらかじめ用意された背景画像との差分を計算することで、対象物のシルエットのみを抽出する処理である。ただし、従来の単純な背景差分によりシルエット抽出を行なうと、対象物の色や照明条件によって、精度の高いシルエットが安定して抽出できないという問題点がある。また、影のように背景と比べて暗くなる部分が、誤ってシルエットに含まれることが多いという問題点もある。

Davisらは、これらの問題点に対してロバストな背景差分や影部分の検出を行った^[3]。具体的には、まずシルエット抽出の前処理として、各視点ごとに背景画像を複数枚獲得することで、それぞれの画素についての平均、標準偏差を獲得する。さらに、平均、標準偏差を基に明るさの変化度、色あいの変化度を定義し、変化度の分布から閾値を動的に求める。そして、それらを利用した背景差分により影部分の切り出しを行ない、精度の高いシルエット抽出を実現した。

ただし、Davisらは屋外のような背景の変化が大きい環境を想定してシルエット抽出を行っているのに対して、我々は、対象物が多数の力

メラに囲まれているような屋内環境に限定してシルエット抽出を行っている。屋内環境における背景の変化は小さく、背景の平均、標準偏差を用いなくても精度の高いシルエット抽出は可能である。そこで我々は、Davisらの手法を応用した、単純かつ効果的な背景差分のアルゴリズムを提案し、目的とする影部分の正確な判別を行い、高い精度のシルエット抽出を実現する。

提案手法の具体的なアプローチは以下の通りである。RGB空間において、入力画像のRGBベクトルを $I = [I_r, I_g, I_b]$ 、背景画像のRGBベクトルを $E = [E_r, E_g, E_b]$ 、とすると、背景に対する明るさの変化度 α 、RGBバランス(色あい)の変化度 C は図1のように表される。したがっ

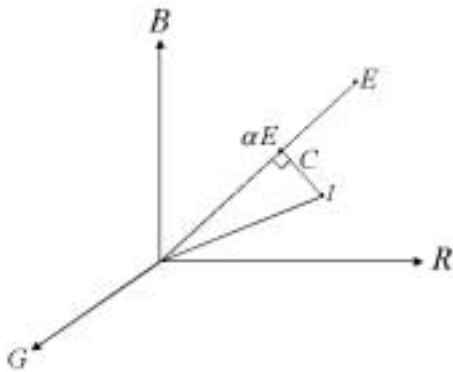


図 1: 明るさと色あいの変化度

て、二つの変化度を以下のように定式化することができる。

$$\alpha = \underset{\alpha}{\operatorname{argmin}} \left[\sum_{c=r,g,b} (I_c - \alpha E_c)^2 \right]$$

$$= \frac{I_r E_r + I_g E_g + I_b E_b}{E_r^2 + E_g^2 + E_b^2}$$

$$C = \|I - \alpha E\|$$

$$= \sqrt{\sum_{c=r,g,b} (I_c - \alpha E_c)^2}$$

入力画像と背景画像の明るさの変化が全くないときの α は 1 であり、 α の値が 1 より大きい場合には入力画像が背景画像に比べて明るい。また、 C の値が大きいほど入力画像と背景画像の色あいの変化も大きい。

α と C についての閾値処理によって、全画素を以下の 4 種類に分類する。

前景部分 $C > \theta_C$ 又は $\alpha < \theta_{\alpha 1}$

背景部分 上記以外で、 $\theta_{\alpha 2} < \alpha < \theta_{\alpha 3}$

影部分(暗くなる部分) 上記以外で、 $\alpha < 1$

明るくなる部分 上記以外

ここで、 θ_C 、 $\theta_{\alpha 1}$ 、 $\theta_{\alpha 2}$ 、 $\theta_{\alpha 3}$ は閾値であり、 $\theta_{\alpha 1} \sim \theta_{\alpha 3}$ が $0 < \theta_{\alpha 1} < \theta_{\alpha 2} < 1 < \theta_{\alpha 3}$ を満たすように設定する。上記 4 種類のうち、前景部分と、明るくなる部分を対象物のシルエット部分とする。上記のような処理を全画素に対して行ない、シルエットを抽出する。

3. 弱校正を用いた 3D モデル復元

3.1 弱校正

カメラキャリブレーションとは、実空間と仮想空間との対応づけのことであり、コンピュータビジョンの世界においては基本的な技術であるが、決定的な手法は開発されておらず現在も様々な研究がなされている^[4]。

1992年、Faugeras および Hartley らは、キャリブレーションされていないカメラによって撮影された画像間において、特徴点を求めることでカメラ間の幾何関係や実空間との対応を求める手法を発表した^{[5][6]}。このような手法は弱校正 (weak calibration) と呼ばれる。

3.2 レンズの幾何学的歪み

Shih らは、レンズの幾何学的歪みがキャリブレーションの精度に与える影響を考察しており^[7]、本手法では、画像間の特徴点を獲得するために撮影された画像、3Dモデル復元に必要なシルエット画像に対して、歪み補正処理を行なう。

3.3 射影グリッド空間

矢口らは、弱校正を用いたシステムによって良好な任意視点画像を生成しており^[8]、本手法においても同様のシステムを用いて 3Dモデル復元を行なう。

まず、数台あるカメラのうちの任意の 2 台を基底カメラ 1、2 とする。そして、この 2 台のカメラそれぞれの視点から、中心投影によって図 2 のような 3D 空間を定義する。ここで、空

間中のボクセル $A(p, q, r)$ は、基底カメラ 1 から撮影された画像 1 上の点 $a_1(p, q)$ へ、基底カメラ 2 から撮影された画像 2 上の点 $a_2(r, s)$ へ投影されるものとする。

このように定義された 3D 空間は射影グリッド空間 (Projective Grid Space : PGS) と呼ばれる。

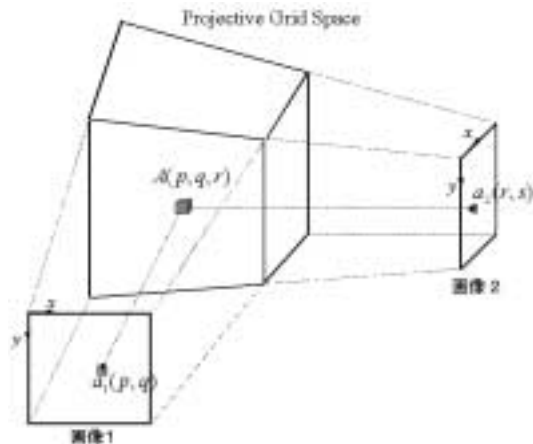


図 2: 射影グリッド空間

3.4 3Dモデル復元

本手法において、3Dモデルは視体積交差法の理論に基づいて復元される。視体積とは、視点を頂点、シルエットを断面とする錐体のことであり、すべての視点における視体積の共通部分を求めることにより、対象物の形状を復元する手法を視体積交差法という。視体積交差法は、高速処理が可能であるという長所を持つ反面、モデルの形状を正確には復元できないという短所を持つ。交差計算をする際には、画像間の幾何関係や、カメラ座標と空間座標との対応関係が必要であり、それらはF行列を算出することで既知となる。理論的に、F行列は2画像間の9点以上の点対応によって決定することができる。本手法では、F行列を利用して以下のような投影を行なう(図3)

まず、空間上の $A(p, q, r)$ に対する画像 1 上の投影点を a_1 については、PGSの定義より $a_1(p, q)$ に投影される。

次に画像 2 上の投影点 a_2 については、F行列の定義から F_{21} を用いて画像 2 にエピポーラ線 l_{21} として投影すると、 a_2 は l_{21} 上に存在するた

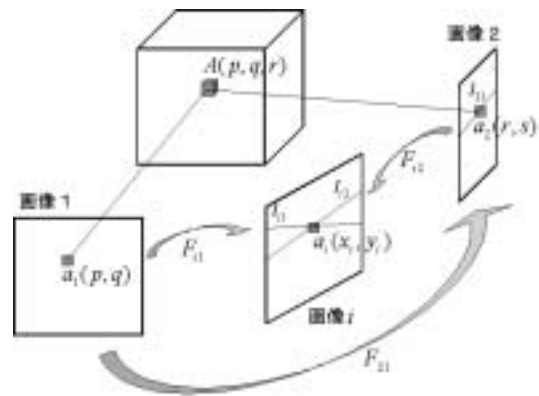


図 3: 各視点画像への投影法

め、直線 l_{21} は次式のように定義される。

$$l_{21} = F_{21} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix}$$

a_2 の x 座標値は PGS の定義より r であるから、 y 座標値 s も定まる。

そして、基底カメラ以外のカメラから撮影される画像 i に対する投影点の座標値 x_i, y_i は、次のようにして定まる。基底カメラ 2 への投影と同様に、 F_{i1} を用いて a_1 を画像 i 上に直線 l_{i1} として投影する。また F_{i2} を用いて、 a_2 を画像 i 上に直線 l_{i2} として投影する。2本のエピポーラ線 l_{i1}, l_{i2} の交点が画像 i の投影点の座標である。このような処理を、全視点の画像に対して行なうことで、注目ボクセルに対する全視点の画像の座標値を求めることができ、正確な 3Dモデル復元が可能となる。

交差計算による 3Dモデル復元の手順は、以下の通りである。定義された PGS 上で、その空間に含まれるボクセルを一枚のシルエット上に投影し、そのシルエット上にないボクセルを全て削除し、次のシルエット画像に投影する。この処理を、基底カメラ 1、基底カメラ 2、その他のカメラの順に行なう。全ての入力視点画像のシルエットに含まれるボクセルだけを「存在」とみなし、3Dモデルを復元する。

4. 任意視点カラー画像生成

本章では、復元された 3Dモデル形状と入力画像の色情報を基に、任意視点画像のカラー表

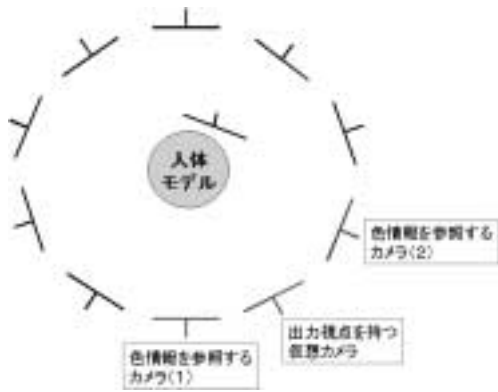


図 4: 入出力視点

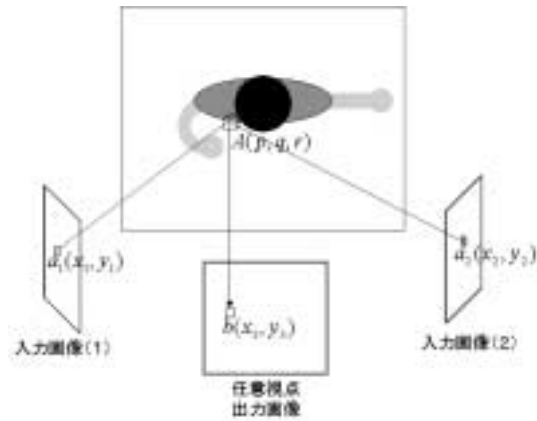


図 5: オクルージョンの影響

示を行なう手法について述べる。

4.1 色情報を参照する入力画像の選択

任意視点画像をカラー表示出力する際に、どの入力画像の色情報を参照するかについて議論する必要がある。一般に、出力する視点位置から離れた入力画像の色情報は、実際に出力する視点から見た色とは異なることが多い。そのため Matusik らは、入力視点からの中心射影と出力視点からの中心射影とのなす角度を評価し、もっとも角度の小さい入力画像の色情報だけを用いて任意視点カラー画像を生成する手法を提案している^[9]。本手法では、できるだけ無駄な色情報を用いずに処理を行なうために、図 4 のようにカメラ配置を仮定して、出力画像の視点を挟んでもっとも距離の近い 2 視点の画像だけを色情報参照画像とする。

4.2 オクルージョンの影響

多視点入力画像から任意視点カラー画像を出力する際には、オクルージョンの影響を考慮する必要がある。

図 5 のように、色情報を参照する入力画像中の対象画素を $a_1(x_1, y_1)$ 、 $a_2(x_2, y_2)$ 、任意視点における出力画像中の対象画素を $b(x_3, y_3)$ 、空間上の対象ボクセルを $A(p, q, r)$ とする。 $a_1(x_1, y_1)$ と $a_2(x_2, y_2)$ の RGB 値が近ければ、いずれか一方の RGB 値、または二つの RGB 値の平均値を $b(x_3, y_3)$ の RGB 値とすればよい。しかし、オクルージョンがある場合、 $a_1(x_1, y_1)$ と $a_2(x_2, y_2)$ の RGB 値は大きく異なっており、 $b(x_3, y_3)$ の RGB 値を決定することができない。

この問題に対する一般的な解決手段として、Z-buffer 法などを用いて入力視点と対象ボクセルとの距離情報を格納する手段がある。しかし、このような手法には計算量が大きくなるという問題点がある。

4.3 任意視点画像への色づけ法

そこで、我々は上記の問題に対して、全ボクセルに対する走査を一度だけ行ない、任意視点画像への RGB 値を決定する手法を提案する。具体的には、視体積交差計算により「存在」と判別された全てのボクセルに対して、図 6 の手順のように、任意視点画像への色づけ法を 2 種類提案する。どちらの手法においても、 $a_1(x_1, y_1)$ と $a_2(x_2, y_2)$ の RGB 値が大きく異なっている場合のみ、近傍 6 ボクセルの色情報を参照する。その後、提案手法 1 は、6 ボクセルの RGB 平均値 m を任意視点画像に出力する。提案手法 2 は、 $a_1(x_1, y_1)$ と $a_2(x_2, y_2)$ の 2 つの RGB 値のうち、 m に近い方の RGB 値を出力する。

各ボクセルから任意視点画像に色づけを行なう際には、3D モデル復元と同様に F 行列を用いて、対象ボクセルの座標値を任意視点画像に投影することで、出力する任意視点画像の座標値を得ることができる。

このとき、任意視点画像の各画素に対応する PGS 上のボクセルは、任意視点画像の視点位置から PGS に向けた放射線上に複数個存在するという問題がある。この問題に対して我々は、交差計算を行なう PGS の走査順を、あらかじめ任意視点画像から遠い順に定めておくことで、走

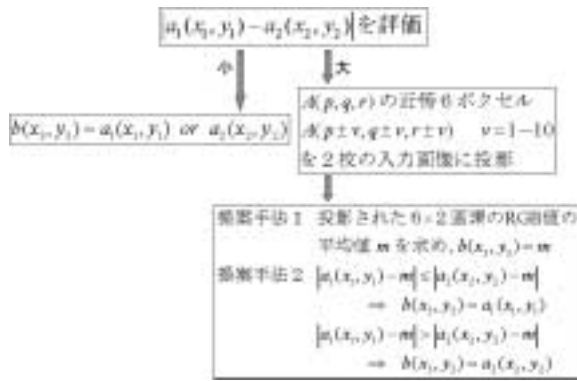


図 6: 任意視点画像への色づけ法

査が終了したときには任意視点画像にもっとも近いボクセルの色が任意視点画像の色となるように処理を行なう。

5. 実験

10台のカメラを用いたシステムにより、任意視点カラー画像を生成した。実験に用いられる画像サイズは全て640×480画素である。

5.1 シルエット抽出精度の評価

本手法によるシルエット抽出結果を図7に示す。図中の灰色部分は、シルエットに含まれる部分(前景部分+明るくなる部分)である。シルエット抽出の課題であった、影部分の切り出しに成功したことが確認できる。シルエットに雑音混ざっているが、視体積交差法による3Dモデル形状を復元する際の交差処理により雑音は消失するため、雑音除去を行なう必要はない。

5.2 任意視点カラー画像の評価

オクルージョンが起こるモデル形状を対象に、10台のカメラシステムのうち9台を入力視点として、視体積交差法により3Dモデル形状を復元し、任意視点画像を生成した。図8は、それぞれの視点において抽出された9枚のシルエット画像のうち、色情報を参照する2枚のシルエット画像である。また、入力視点に使われていない残りの1台のカメラが持つ視点を、出力する任意視点としてカラー画像生成を行う。図9に、その視点から撮影されたシルエット画像を示す。そして、出力視点のシルエット画像と、本手法によって生成された任意視点画像とのRGB二乗平均誤差(RMS)を求めることで、任意視点カ



(a) 入力画像

(b) 背景画像



(c) 前景部分



(d) 背景部分



(e) 影部分



(f) 明るくなる部分

図 7: シルエット抽出結果

ラー画像の客観的な精度評価を行なった。図10～図12は以下のような比較手法を用いてカラー表示した結果であり、図13、図14は、2つの提案手法を実装してカラー表示した結果である。

比較手法 1 $b(x_3, y_3) = a_1(x_1, y_1)$

比較手法 2 $b(x_3, y_3) = a_2(x_2, y_2)$

比較手法 3

$$b(x_3, y_3) = (a_1(x_1, y_1) + a_2(x_2, y_2))/2$$

また、()内の数値は二乗平均誤差の値である。

任意視点画像における人体の形状は良好であった。任意視点画像のカラー表示については、提案手法2の結果によって、比較手法1、2に現れるようなオクルージョンの影響を多少緩和できることが確認できた。提案手法1、比較手法3は他の手法に比べてRMSが小さいが、これらの手法では平均値を出力しているため、見た目が少しぼやけてしまう。また任意視点画像において、モデルの輪郭付近にモデル以外の色が付

加されているのは、色情報を参照したシルエット画像にモデルの輪郭付近の背景色が付加されているためである。したがって、シルエット抽出精度を高めることで、任意視点カラー画像の精度を高めることができる。



入力画像(1) 入力画像(2)

図 8: 色情報を参照するシルエット画像



図 9: 任意視点画像と同視点のシルエット画像



図 10: 任意視点画像：比較手法 1 (52.6)



図 11: 任意視点画像：比較手法 2 (54.0)



図 12: 任意視点画像：比較手法 3 (39.4)



図 13: 任意視点画像：提案手法 1 (38.0)



図 14: 任意視点画像：提案手法 2 (45.1)

6. おわりに

本稿では、人体のように複雑な形状を持つモデルを対象として、多視点画像から仮想空間に 3D モデルを復元し、復元された 3D モデル形状を基にして任意視点画像の生成する手法を提案した。カメラにより撮影された入力画像の RGB 値を効果的に利用することによって、任意視点からの出力画像に色情報を付加することができた。今後の課題としては、任意視点カラー画像を実時間処理できるように、高速化することが挙げられる。また、出力視点をユーザーが任意に切り替えられるような、自由度の高いシステムを構築したい。

参考文献

- [1] S.M.Seitz, and C.R.Cyer : “View Morphing”, Proc.SIGGRAPH '96, pp.21-30,1996.
- [2] Y.Yagi, D.Arita, R.Taniguchi : “Real-Time 3D Shape Modeling from Multi-View Image”, Proc.8th Joint Workshop –Frontiers of Computer Vision–, pp.55-61, 2002.
- [3] T.Horprasert, D.Harwood, and L.S.Davis : “A Robust Background Subtraction and Shadow Detection”, Proc.ACCV'2000, pp.983-988, 2000.
- [4] R.Y.Tsai : “An efficient and accurate camera calibration technique for 3D machine vision”, CVPR, pp.364-374, 1986.
- [5] O.D.Faugeras : “What can be seen in three dimensions with an uncalibrated stereo rig?”, ECCV, pp.563-578, 1992.
- [6] R.Hartley, R.Gupta and T.Chang : “Stereo from uncalibrated cameras”, CVPR, pp.761-764, 1992.
- [7] S.W.Shih, Y.P.Hung and W.S.Lin : “When should we consider lens distortion in camera calibration”, PR, Vol.28, pp.447-461 , 1995.
- [8] 矢口悟志 , 斎藤英雄 : “Projective Grid Space における多視点 Silhouette 画像からの自由視点画像生成”, 電子情報通信学会技術報告 PRMU2000-26, pp.23-28, 2000.
- [9] W.Matusik, C.Buehler, R.Raskar, S.J.Cortler, L.McMillan : “Image-Based Visual Hulls”, Proc.SIGGRAPH '2000, pp.21-30, 2000.