

実光源下における巨大仮想物体の
実画像への重ね込み

2002年度

慶應義塾大学大学院 理工学研究科
開放環境科学専攻

指導教授：原田賢一 教授

学籍番号:80126944

吉藤 伸幸

目次

第1章	序論	1
1.1	はじめに	1
1.2	本論文の目的	1
1.3	本研究の概要	2
1.4	本論文の構成	3
第2章	画像生成の原理	4
2.1	Radiance について	4
2.2	レンダリング	4
2.3	仮想物体の実画像への重ね込み	5
2.3.1	仮想物体	5
2.3.2	影をつけたい部分	5
2.3.3	背景	6
2.4	結果	6
2.5	実画像中の影の削除	10
2.5.1	レンダリング	10
2.5.2	実画像中の影の削除	11
第3章	ハイダイナミックレンジ	15
3.1	ダイナミックレンジの不足による情報の消失	15
3.2	ハイダイナミックレンジ	17
3.3	問題点	20
3.4	改良点と結果	20
第4章	カメラパラメータの推定	27
4.1	世界座標と画像座標の関係	28
4.2	パラメータの決定	30
4.3	関連研究におけるカメラキャリブレーション	33
4.4	本研究におけるカメラキャリブレーション	33

第5章 実験	35
5.1 幾何的整合性の実現	35
5.1.1 3次元モデルの作成	35
5.1.2 実験結果	35
5.2 光学的整合性の実現	35
5.2.1 光源画像の作成	38
5.2.2 画像生成	38
第6章 考察	47
6.1 光学的整合性について	47
6.2 幾何的整合性について	47
6.3 実画像中の影の削除	48
第7章 結論と今後の展望	50
7.1 結論	50
7.2 今後の展望	50
7.2.1 ハイダイナミックレンジのさらなる改良	50
7.2.2 実画像中の影の削除	51
7.2.3 他の建造物への応用	51
謝辞	52
参考文献	53

第1章 序論

1.1 はじめに

電子的に構築された仮想世界を現実世界に融合させる技術は様々な産業の分野で、今後を担う基礎技術として注目を集めている。この技術により実現される複合現実と呼ばれる融合の世界は、実世界の常識を越えた事象を可能とし、人間の思考空間や生活空間を拡大する。

実世界と仮想世界を違和感なく融合させるためには、実世界と仮想世界における以下の3つの整合性を実現しなければならない。

- (1) 幾何的整合性 (正しい配置での融合)
- (2) 光学的整合性 (違和感ない陰影での融合)
- (3) 時間的整合性 (時の流れの一致)

これらの整合性を実現するためには、まず、実世界を理解する必要がある。つまり、実世界の事象を理解し、仮想空間との関係を求めることにより、2つの世界の融合は、初めて可能となるのである。その点で、融合の技術は、実世界の理解を目的とするコンピュータビジョン技術と密接な関係を持ち始めた [1]。

1.2 本論文の目的

本研究では、これまで行なわれていない、巨大建造物を仮想物体として、現実世界に融合させることを目的とする。具体的には、現在の東大寺大仏殿の位置に仮想物体である創建時の大仏殿を復活させた画像を作成する。

その際の問題点として、

- 太陽を含む光源を正確に光源画像として用意できないこと。
- 対象物が巨大で形状が複雑であり、カメラと物体の距離が大きく、また、対応点を多く取れないため、正確にキャリブレーションがで

きないこと。

が挙げられる。

そこで、本研究は、仮想物体の影を実世界の光源環境に基づいて実画像上に生成するために、ハイダイナミックレンジの手法の改良を行ない、太陽光の影響を受けた、より正確な陰影で融合させることを目的とする。

また、キャリブレーションを工夫して行なうことで、より正確に配置することを目的とする。

創建時の大仏殿のモデル(図 1.1)は、創建当時の復元模型と唐招提寺金堂から測定したモデルを用いて復元されたものである。

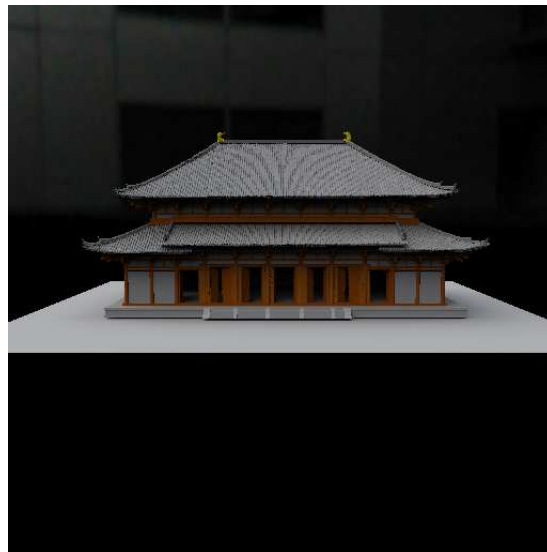


図 1.1: 創建時の大仏殿のモデル [2]

1.3 本研究の概要

本研究では、以下のような手順で仮想物体を実画像に違和感なく重ね込む。

(1) 1枚の写真と撮影された物体の3次元モデルからカメラパラメータを推定し、その写真を撮影したカメラの位置、カメラの向いていた方向、水平方向、垂直方向それぞれの画角を求める。

(2) 魚眼レンズで実世界の光源環境をシャッタースピードを変え数枚撮影し、それらから正確に光源環境を表現している画像を作成する。

(3) (1) で求めたそれぞれの情報, (2) で作成した画像から仮想物体を実画像に重ね込んだ画像を作成する。

1.4 本論文の構成

本論文の構成を以下に示す。

2章: 仮想物体を実画像に重ね込む理論について述べる。

3章: ハイダイナミックレンジについて述べる。

4章: カメラパラメータ推定に関する理論, また, 関連研究におけるカメラパラメータ推定方法, 本研究におけるカメラパラメータ推定方法について述べる。

5章: 実験結果について述べる。

6章: 考察を述べる。

7章: 結論と今後の展望について述べる。

第2章 画像生成の原理

光学的整合性を実現するためには，違和感ない陰影で融合する必要がある．本研究では，Radiance というソフトを用いる．このことにより佐藤らの手法 [3] に比べ，より正確に影を計算できるだけでなく，2.3 節に述べるような，完成画像の各ピクセルを3つのグループに分ける際もマスク画像を作成することにより，より容易に行なうことができる．本章では，仮想物体を実画像へ重ね込んだ画像を生成する原理について説明する．

2.1 Radiance について

Radiance は照明の分析，視覚化のためのパッケージソフトである．

入力データは3次元座標，材質，光源，時間，日付，(日光の計算用の)空の状態である．輝度と色を示すスペクトル放射輝度，照度と色を示す放射照度，眩しさ指数を計算する．シミュレーション結果は画像，数値で表される．

簡単に照明状態の計算やレンダリングを行なうことが出来る．また，シミュレーションを行なう際に，物体の形状や材質に制限はないといったことも大きな利点である [4][5]．

2.2 レンダリング

Radiance を用いてレンダリングを行なう．仮想物体に関しては創建時の大仏殿 (図 2.1) を，光源環境の画像には図 2.2 を用いる．また，画像を作成する際の視点は，光学的整合性のみの実現のため，適当なものを選んだ．これらの条件で，仮想物体の部分のマスク画像 (図 2.3)，地面の部分のマスク画像 (図 2.4)，さらに，光源環境を考慮した，仮想物体がない場合の地面の照度を示す画像 (図 2.5)，仮想物体がある場合の地面の照度と仮想物体を示す画像 (図 2.6) の計 4 枚の画像を作成する．なお，本研究

では，仮想物体を重ね込むことにより影がつくのは地面のみとする．

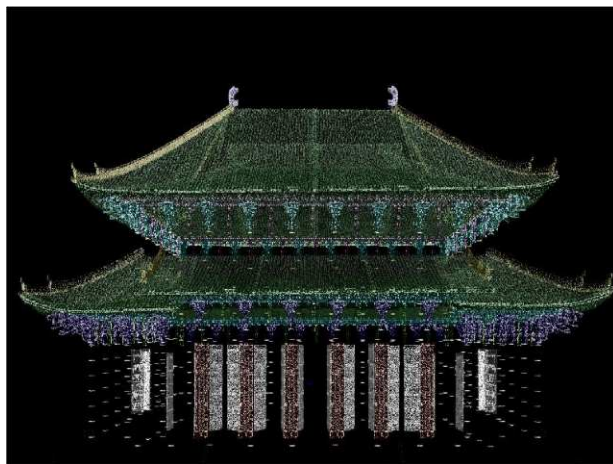


図 2.1: 創建時の大仏殿の 3 次元モデル (仮想物体)

2.3 仮想物体の実画像への重ね込み

マスク画像を用いて，出力する画像の各ピクセルを仮想物体，影をつけたい部分，背景の 3 つのグループに分類する．

2.3.1 仮想物体

図 2.3 において白の部分が出力画像における仮想物体となる．この部分に関しては，図 2.6 における仮想物体の部分を使う．

2.3.2 影をつけたい部分

図 2.4 において白の部分が出力画像における影をつけたい部分となる．この部分に関しては，図 2.5，図 2.6，式 (2.1)[3] を用いて仮想物体を重ね込んだ後の画像の色を求める．

$$M_m = I_m \frac{E_m^2}{E_m^1} \quad m = R, G, B \quad (2.1)$$



図 2.2: 光源画像

- E_m^1 : 仮想物体が存在しない場合の照度
- E_m^2 : 仮想物体が存在する場合の照度
- I_m : 実画像の色
- M_m : 仮想物体を重ね込んだ後の画像の色

2.3.3 背景

仮想物体，影をつけたい部分のいずれにも属さない部分は出力画像における背景となる．この部分に関しては，実画像(図2.7)をそのまま使う．

2.4 結果

以上の手法により，出力された結果は図2.8となる．

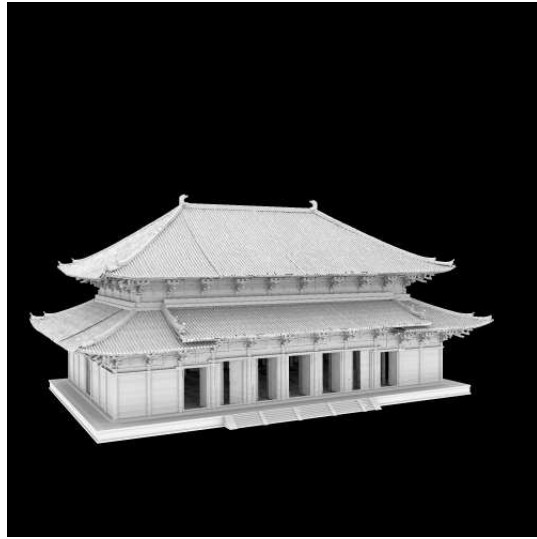


図 2.3: 仮想物体のマスク



図 2.4: 地面のマスク

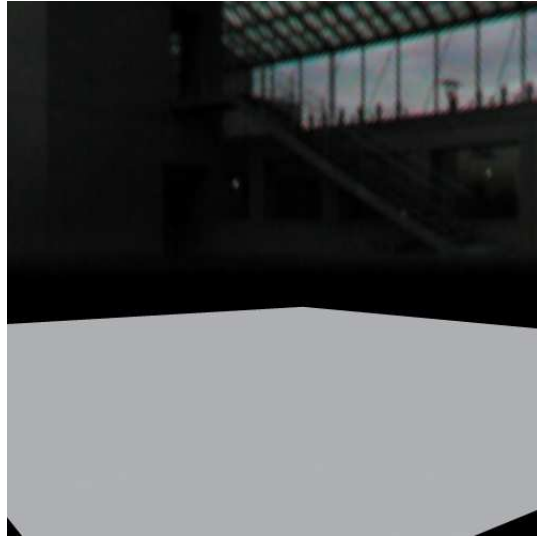


図 2.5: 仮想物体がない場合の地面の照度

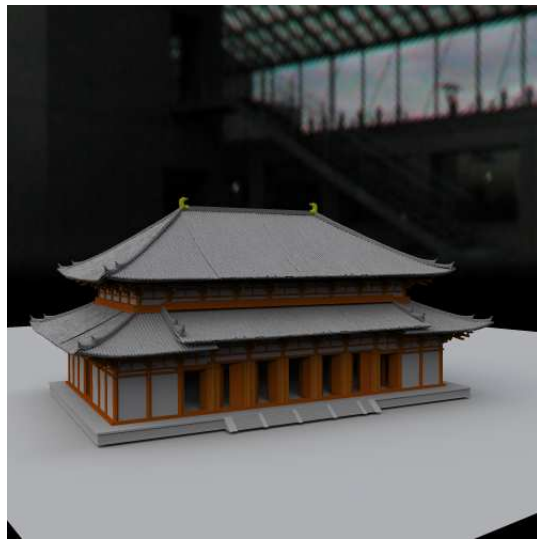


図 2.6: 仮想物体がある場合の地面の照度と仮想物体



図 2.7: 実画像



図 2.8: 仮想物体を重ね込んだ結果

2.5 実画像中の影の削除

本研究のように，仮想物体を実画像中の物体の位置に重ね込むという作業を行なう際，実画像中の物体の影響による影を削除した，新しい画像を用意し，その画像に重ね込むという作業を行なった方が，更に自然な画像を作成できると考えられる．そこで本節では，2.3節で説明した影をつける手法の逆の作業を行なうことにより，影を削除する方法を説明する．

2.5.1 レンダリング

2.2節と同様に，Radianceを用いてレンダリングを行なう．物体に関しては図2.9,2.10の様な一辺12cmの立方体を用意した，光源環境の画像には図2.11を用いる．また，画像を作成する際の視点は，物体に関してキャリブレーションを行い，カメラの位置，カメラが向いていた方向，画角をそれぞれ求めた．これらの条件で，物体の部分のマスク画像(図2.12)，地面の部分のマスク画像(図2.13)，さらに，光源環境を考慮した，物体がない場合の地面の照度を示す画像(図2.14)，物体がある場合の地面の照度と仮想物体を示す画像(図2.15)の計4枚の画像を作成する．なお，本節では，物体により影がつくのは地面のみとする．



図 2.9: 物体

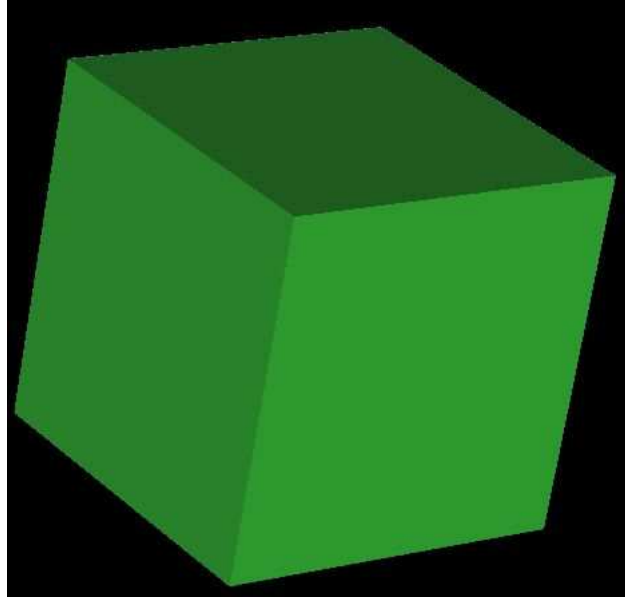


図 2.10: 物体の 3 次元モデル

2.5.2 実画像中の影の削除

マスク画像を用いて、出力する完成画像の各ピクセルを物体、影を削除したい部分、背景の 3 つのグループに分類する。

物体

図 2.12 において白の部分が完成画像における物体となる。この部分に関しては、図 2.9 における物体の部分を使う。

影を削除したい部分

図 2.13 において白の部分が完成画像における影を削除したい部分となる。この部分に関しては、図 2.14, 図 2.15, 式 (2.1) を I_m に関して解いた式 (2.2) を用いて影がないときの画像の色を求める。

$$I_m = M_m \frac{E_m^1}{E_m^2} \quad m = R, G, B \quad (2.2)$$

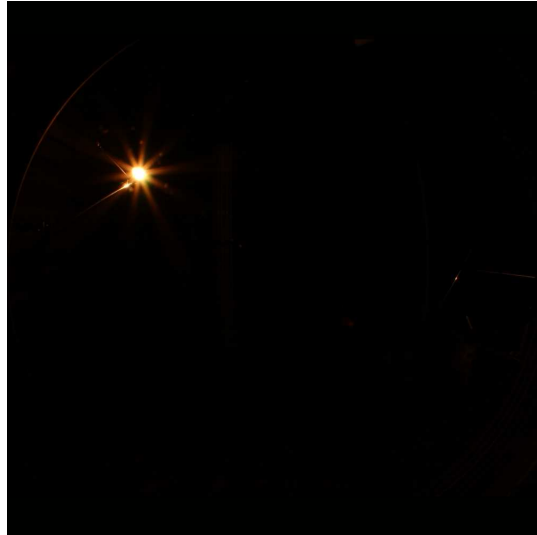


図 2.11: 光源画像

- E_m^1 : 物体が存在しない場合の照度
- E_m^2 : 物体が存在する場合の照度
- I_m : 実画像の色
- M_m : 物体があるときの画像の色

背景

仮想物体，影を削除したい部分のいずれにも属さない部分は，完成画像における背景となる．この部分に関しては，元の画像 (図 2.9) をそのまま使う．

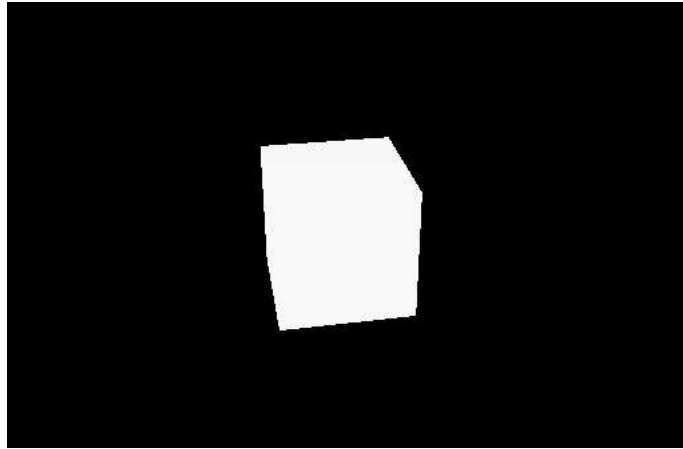


図 2.12: 物体のマスク



図 2.13: 地面のマスク

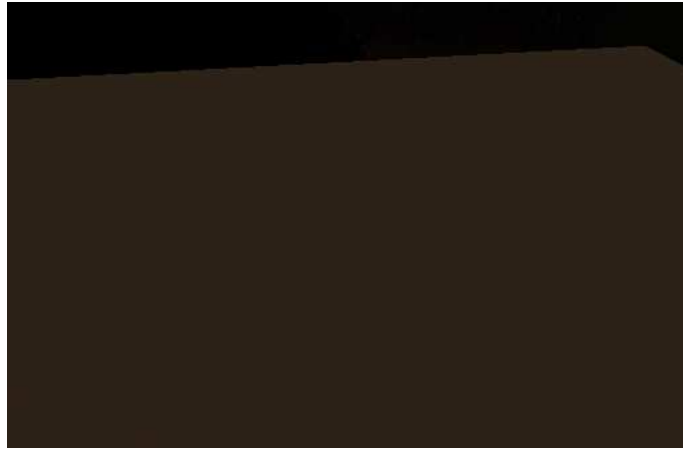


図 2.14: 物体がない場合の地面の照度

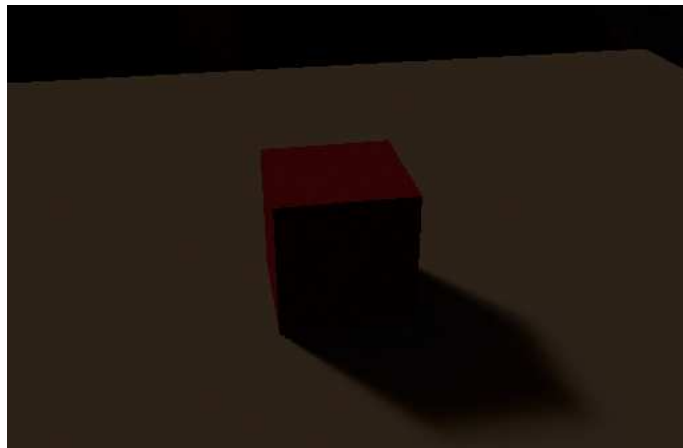


図 2.15: 物体がある場合の地面の照度と物体

第3章 ハイダイナミックレンジ

本章では，本研究において，光源画像を作成するために用いるハイダイナミックレンジについて説明する．

3.1 ダイナミックレンジの不足による情報の消失

撮像装置，表示装置による，ダイナミックレンジの圧縮は，以下のよ
うなメカニズムによるものである．

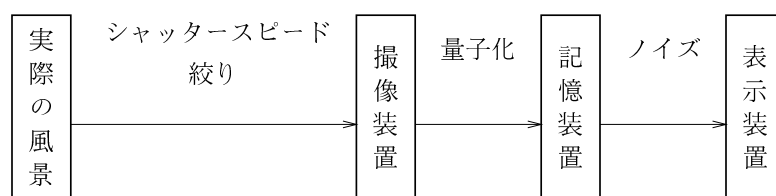


図 3.1: ダイナミックレンジの圧縮

実際に風景の持つダイナミックレンジに比べ，CCD カメラなどの撮像素子の持つダイナミックレンジは狭いため，撮像装置のシャッタースピード，絞りにより，決定される条件により，一部は線形変換に，その他の部分は圧縮され，非線形変換となる．

撮像装置の撮像素子から出力された電子信号は，記憶装置に記録されるため，量子化される．

表示装置は，一般にノイズや装置固有の特性により，入力されたデータをそのまま表示できるわけではない．すなわち，記録されているデータは，微細なデータが残っていて，表示される場合は，その微細なデータがノイズに埋もれてしまい，認識できなくなってしまうため，表示できる階調数は，記録されているデータに比べ，少なくなり，ダイナミックレンジは狭くなってしまう．

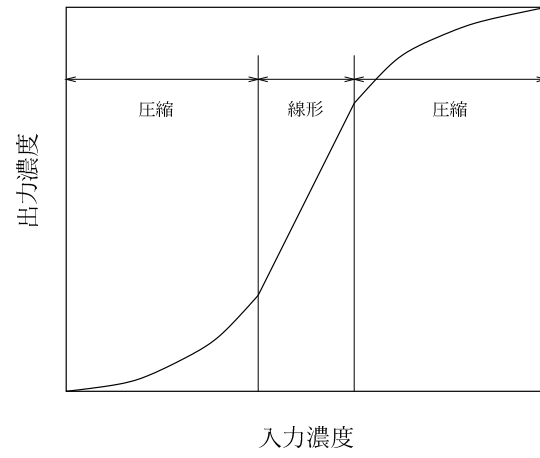


図 3.2: 撮像装置によるダイナミックレンジの圧縮

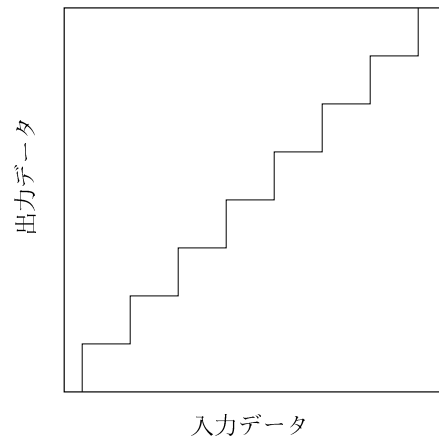


図 3.3: 記録データの量子化

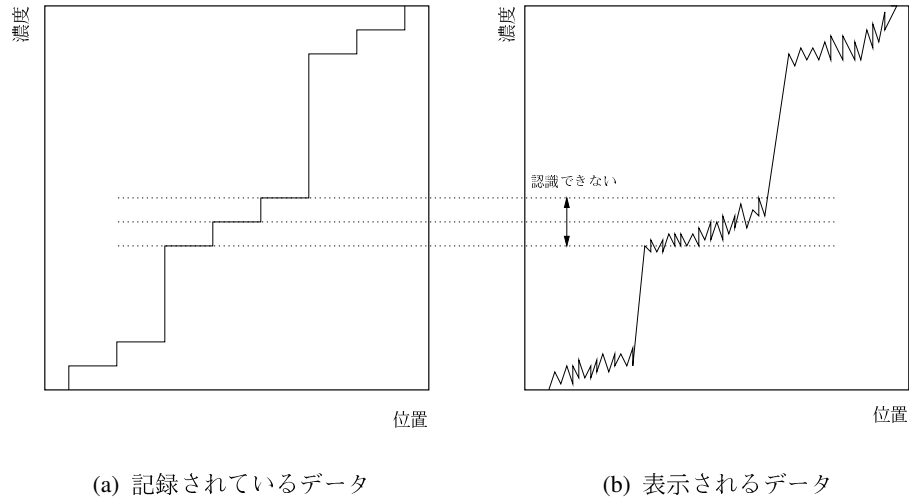


図 3.4: ノイズによるダイナミックレンジの低減

これらの過程を経て、実際の風景に比べ、表示される画像はダイナミックレンジが狭く、人間が目で見ているときに見えていた細部情報が失われてしまう。

3.2 ハイダイナミックレンジ

シャッタースピードや絞りを変えた写真を数枚撮影し、それらから新しい画像を作成する手法である。図 3.5 のようにシャッタースピードを変化させて撮影された写真から図 3.6 のような画像を生成する [6].

ある点からの光量を E_i 、シャッタースピードを Δt_j とすると、画像中のピクセルの値 Z は関数 f を用いて

$$Z_{ij} = f(E_i \Delta t_j) \quad (3.1)$$

と表すことが出来る。関数 f は単調増加であり、逆関数をもつと考えられるので、

$$f^{-1}(Z_{ij}) = E_i \Delta t_j \quad (3.2)$$

となる。そこで、両辺の \log をとると、

$$\ln f^{-1}(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (3.3)$$

と変換でき，さらにここで $\ln f^{-1}(Z_{ij})$ を関数 $g(Z_{ij})$ と置き換えると，

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (3.4)$$

となる．まず，この関数 g を最小自乗法を用いて求める．
また， E_i は 3.4 を変形した式

$$\ln E_i = g(Z_{ij}) - \ln \Delta t_j \quad (3.5)$$

より求めることが出来る．

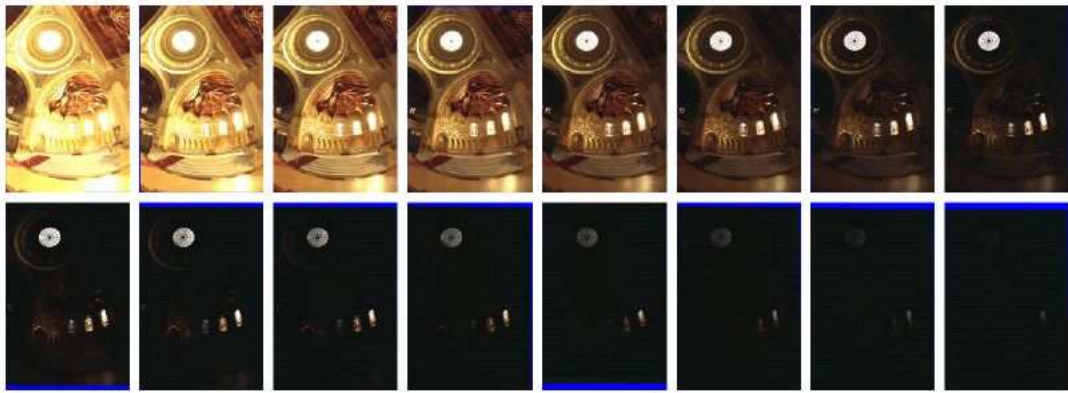


図 3.5: シャッタースピードを 30 秒から 1000 分の 1 秒の間で変化させて撮影された 16 枚の写真

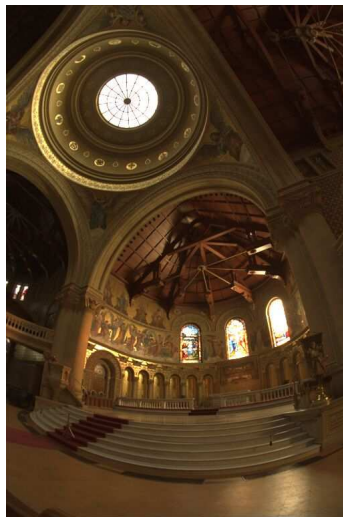


図 3.6: 図 3.5 の写真から作成したハイダイナミックレンジを実現した画像

3.3 問題点

本研究において、ハイダイナミックレンジの手法は光源画像に対して用いる。3.1節で説明したように、光源画像として画像を1枚だけ撮影したのでは、情報を消失してしまっているからである。

しかし、本研究で光源画像として用いる画像を作成するために、ハイダイナミックレンジの手法を用いる際に問題がある。本研究では、屋外で撮影した実画像に仮想物体を重ね込むために、当然光源画像には太陽が存在する。太陽というのは、非常に明るいため、既存の手法では、太陽の部分でダイナミックレンジを超えてしまい、その部分の情報がなくなってしまう(サチュレーション)(図3.7,図3.8)。つまり、どの写真においても太陽の部分のピクセルの値が255になってしまい、実際の明るさが分からないのである。そのため、本研究で使用するには改良を行う必要がある。

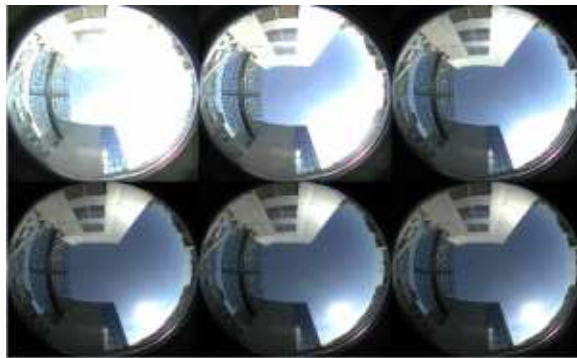


図 3.7: シャッタースピードを30分の1秒から1000分の1秒の間で変化させて撮影された6枚の写真

3.4 改良点と結果

サチュレーションを起こさない部分では既存の手法でピクセルの値を求め、サチュレーションを起こしているピクセルにはその境界のピクセルの値よりも大きな値を入れるという方法で画像を作成した(図3.9)。その値はサチュレーションを起こしている中心に大きな値を代入し、そのまわりには光の強さは光源からの距離の2乗に反比例するということを



図 3.8: 図 3.7 の写真から既存の手法で作成した画像



図 3.9: 図 3.7 の写真から作成したハイダイナミックレンジを実現した画像

用いて求めた値を代入した。

この画像がどれだけ正確に実光源を再現しているかを確かめるために、次のような実験を行った。なお、以下で使用した画像の大きさはすべて1000[pix]×750[pix]である。

まず、地面に箱をおき撮影した(図3.10)。その画像に様々な光源下で同じ大きさの箱を仮想物体として重ね込んだ際の影を実際の箱の影と比較した。また、図3.11は図3.10における画像中のRGBの変化を示している。

まず、光源を点光源として仮想物体に影をつけたものが図3.12である。図3.13を見ても明らかなように、影がついている部分とそうでない部分の境界がはっきりしすぎている。また、地面が真っ黒になってしまっている、つまり、ピクセルの値が0もしくはそれに近い値になってしまっている。このことから不自然な影である。

また、同様に同じ仮想物体に対して、図3.7の中の1枚の写真を光源画像として用い影をつけた場合が図3.14であるが、図3.11と図3.15を比べると僅かに影がついており、また、点光源の場合とは異なり、実際の影のように滑らかに変化していることは分かるが、値を比較した場合、実際の影とは大きな違いがある。

最後に図3.9の写真を光源画像として用い影をつけた場合が図3.16であるが、図3.17を見ても分かるように先の2つの光源環境に比べ実際の影をうまく再現できているといえよう。

よって、本研究では以上の手法で光源画像を作成し用いることとする。



図 3.10: 箱を撮影した画像

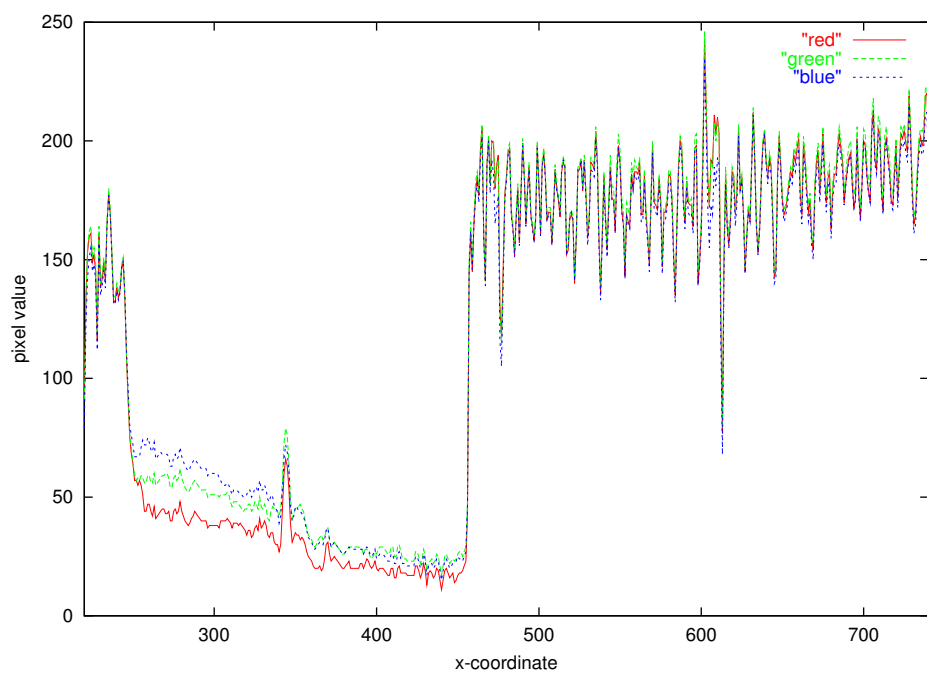


図 3.11: 図 3.10 の $220 \leq x \leq 740, y = 600$ における RGB の値の変化

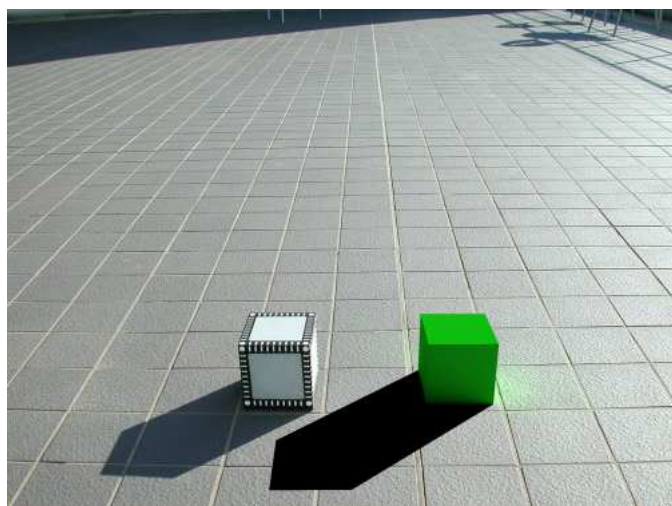


図 3.12: 仮想物体に点光源があたっている場合

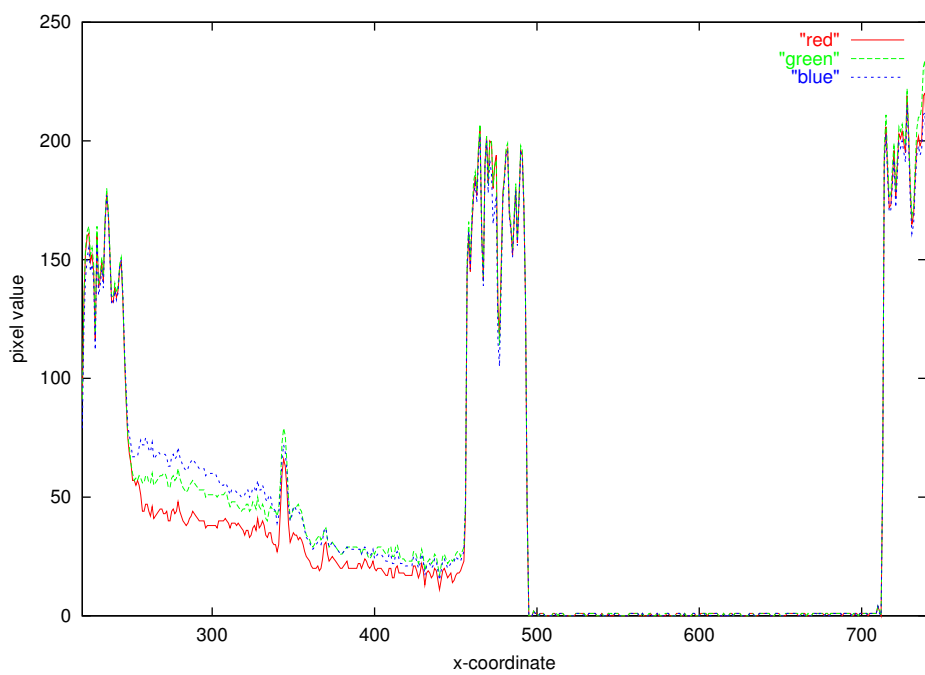


図 3.13: 図 3.12 の $220 \leq x \leq 740, y = 600$ における RGB の値の変化

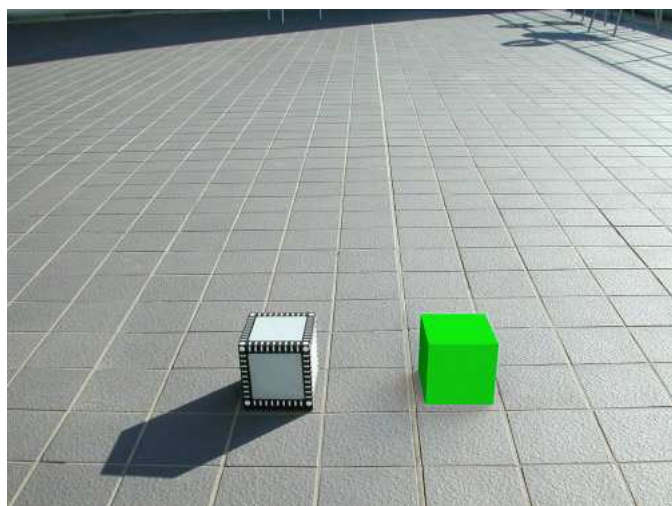


図 3.14: 図 3.7 中の 1 枚の写真光源画像として用い影をつけた場合

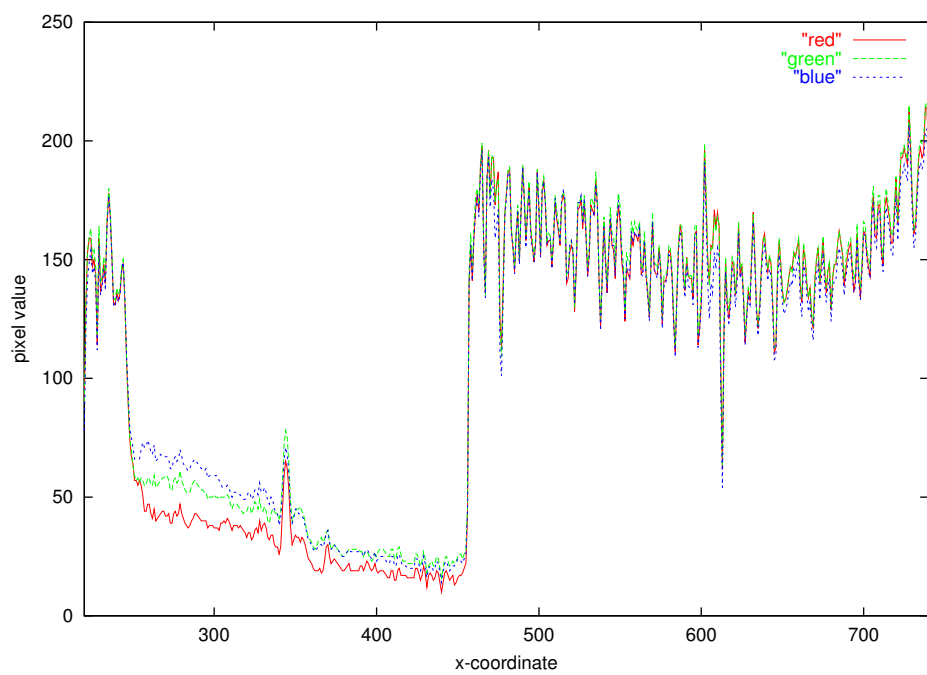


図 3.15: 図 3.14 の $220 \leq x \leq 740, y = 600$ における RGB の値の変化

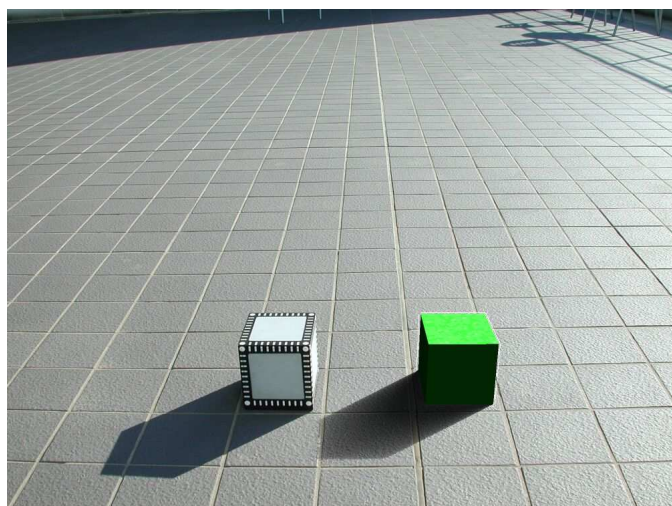


図 3.16: 図 3.9 の写真を光源画像として用い影をつけた場合

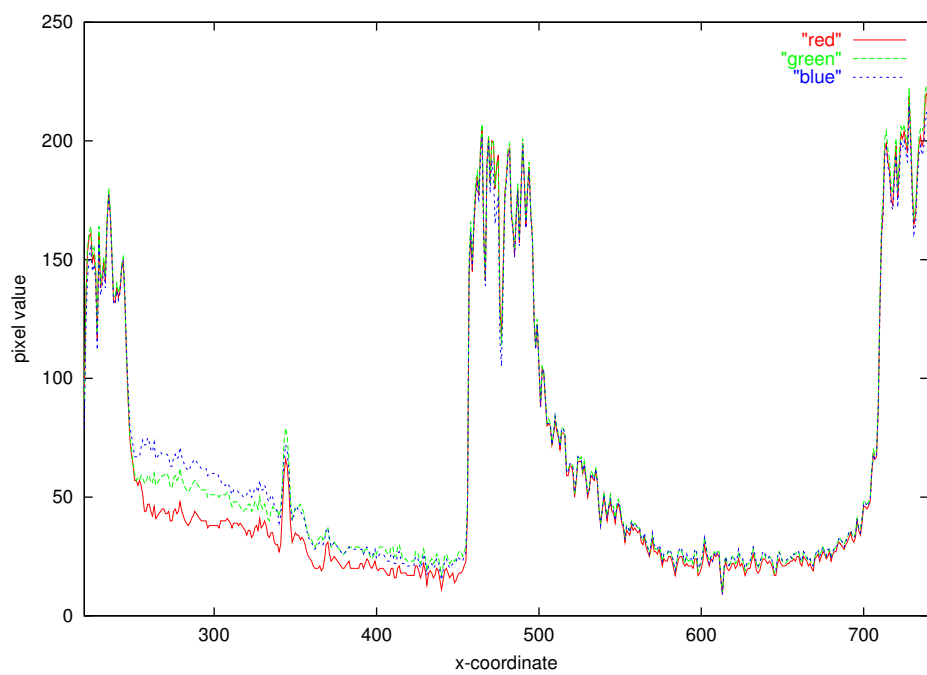


図 3.17: 図 3.16 の $220 \leq x \leq 740, y = 600$ における RGB の値の変化

第4章 カメラパラメータの推定

本研究で用いるカメラパラメータとは以下の 11 のパラメータで構成される。

- 外部パラメータ
 - x 軸基準の回転角 (radian)
 - y 軸基準の回転角 (radian)
 - z 軸基準の回転角 (radian)
 - 平行移動ベクトル T の x 成分
 - 平行移動ベクトル T の y 成分
 - 平行移動ベクトル T の z 成分
- 内部パラメータ
 - 焦点距離
 - 画像原点 C_x
 - 画像原点 C_y
 - レンズひずみ係数
 - スケール係数

本章では、まずカメラパラメータを推定する手法として、Tsai の手法について説明する [7][8][9]。

次に、関連研究において、カメラキャリブレーションがどのように行なわれているか簡単に述べ、最後に、本研究においてカメラパラメータをどのように求め、さらに、その求められたパラメータをどのように用いるかについて述べる。

4.1 世界座標と画像座標の関係

まず準備として、点 P の世界座標系における3次元座標 (x_w, y_w, z_w) とコンピュータ上の画像座標 (X_f, Y_f) の関係を考える。そのために以下の4段階の変換を行う。

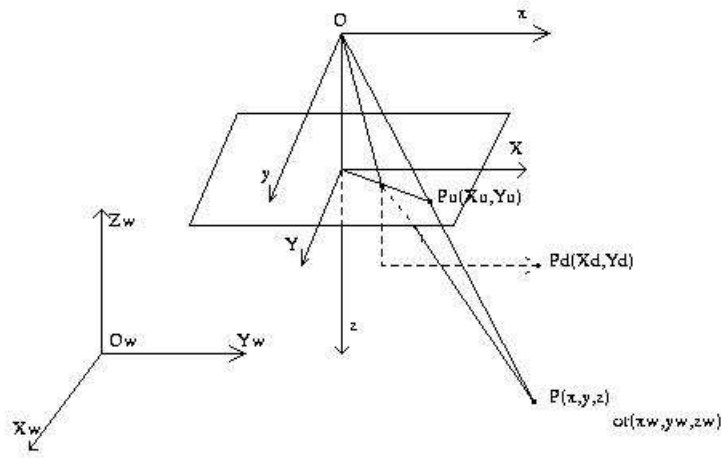


図 4.1: 世界座標と座像座標の関係

Step1

(x_w, y_w, z_w) から (x, y, z) への変換: 回転行列 R , 平行移動ベクトル T を用いて表す。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad (4.1)$$

ただし,

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}, T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (4.2)$$

Step2

透視変換による (x,y,z) から (X_u,Y_u) への変換: 焦点距離 f を用いて表す.

$$X_u = f \frac{x}{z}, Y_u = f \frac{y}{z} \quad (4.3)$$

Step3

(X_u,Y_u) から (X_d,Y_d) への変換:

$$X_d + D_x = X_u, Y_d + D_y = Y_u \quad (4.4)$$

ただし, D_x, D_y はレンズの半径方向のひずみ係数 k_1, k_2 を用いて次のように表わす.

$$D_x = X_d(k_1 r^2 + k_2 r^4) \quad (4.5)$$

$$D_y = Y_d(k_1 r^2 + k_2 r^4) \quad (4.6)$$

$$r = \sqrt{X_d^2 + Y_d^2} \quad (4.7)$$

Step4

(X_d,Y_d) から (X_f,Y_f) への変換:

$$X_f = s_x d'_x{}^{-1} X_d + C_x \quad (4.8)$$

$$Y_f = d'_y{}^{-1} Y_d + C_y \quad (4.9)$$

ただし, s_x はスケール係数, (C_x, C_y) はデジタル画面上の原点座標, d_x, d_y はそれぞれ X 方向 Y 方向の CCD 素子の間隔を表す. なお, d'_x は X 方向の CCD 素子数 N_{cx} と, 1 走査線のサンプル数 N_{fx} を用いて d_x を補正したものである. スケール係数 s_x とする.

$$d'_x = d_x \frac{N_{cx}}{N_{fx}} \quad (4.10)$$

以上の関係を整理すると,

$$X = X_f - C_x, Y = Y_f - C_y \quad (4.11)$$

を用いて

$$s_x^{-1}d'_x X + s_x^{-1}d'_x X(k_1r^2 + k_2r^4) = f \frac{r_1x_w + r_2y_w + r_3z_w + T_x}{r_7x_w + r_8y_w + r_9z_w + T_z} \quad (4.12)$$

$$d_y Y + d_y Y(k_1r^2 + k_2r^4) = f \frac{r_4x_w + r_5y_w + r_6z_w + T_y}{r_7x_w + r_8y_w + r_9z_w + T_z} \quad (4.13)$$

が得られる。ただし、

$$r = \sqrt{(s_x^{-1}d'_x X)^2 + (d_y Y)^2} \quad (4.14)$$

である。

4.2 パラメータの決定

(a) ひずみのある画像座標 (X_d, Y_d) の計算

- 1 それぞれの特徴点 i の座標を求め、 (X_{fi}, Y_{fi}) とする。
- 2 カメラおよび A/D 変換器の仕様から $N_{cx}, N_{fx}, d'_x, d_y$ を求める。
- 3 画像の中央を画像原点 (C_x, C_y) とする。
- 4 N 点の (X_{di}, Y_{di}) を求める。

$$X_{di} = s_x^{-1}d'_x(X_{fi} - C_x) \quad (4.15)$$

$$Y_{di} = d_y(Y_{fi} - C_y) \quad (4.16)$$

(b) 5つの未知数を計算する。

(x_{wi}, y_{wi}, z_{wi}) と (X_{di}, Y_{di}) の組から $T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x, T_y^{-1}r_4, T_y^{-1}r_5$ を未知数とする線形方程式を解く。

$$\begin{bmatrix} Y_{di}x_{wi} \\ Y_{di}y_{wi} \\ Y_{di} \\ -X_{di}x_{wi} \\ -X_{di}y_{wi} \end{bmatrix}^t \begin{bmatrix} T_y^{-1}r_1 \\ T_y^{-1}r_2 \\ T_y^{-1}T_x \\ T_y^{-1}r_4 \\ T_y^{-1}r_5 \end{bmatrix} = X_{di} \quad (4.17)$$

(c) $(T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x, T_y^{-1}r_4, T_y^{-1}r_5)$ から $(r_1, \dots, r_9, T_x, T_y)$ を計算する.

1. $T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x, T_y^{-1}r_4, T_y^{-1}r_5$ から $|T_y|$ を計算する.

$T_y^{-1}r_1, T_y^{-1}r_2, T_y^{-1}T_x, T_y^{-1}r_4, T_y^{-1}r_5$ がすべて 0 でない場合は, 以下のよう T_y^2 は求まる.

$$T_y^2 = \frac{S_r - \sqrt{S_r^2 - 4(r_1'r_5' - r_4'r_2')^2}}{2(r_1'r_5' - r_4'r_2')} \quad (4.18)$$

ただし, $S_r = r_1'^2 + r_2'^2 + r_4'^2 + r_5'^2$ とする.

そうでないときは,

$$T_y^2 = (r_i'^2 + r_j'^2)^{-1} \quad (4.19)$$

となる. r_i', r_j' は $r_i (i=1,2,4,5)$ のなかで 0 でないものとする.

2. まず, T_y の符号を正とする.

3. 以下の計算をする.

$$r_1 = (T_y^{-1}r_1)T_y, \quad r_2 = (T_y^{-1}r_2)T_y$$

$$r_4 = (T_y^{-1}r_4)T_y, \quad r_5 = (T_y^{-1}r_5)T_y$$

$$T_x = (T_y^{-1}T_x)T_y$$

$$x = x_1x_w + r_2y_w + T_x, \quad y = x_4x_w + r_5y_w + T_y$$

4. x と X の符号が等しくかつ y と Y の符号が等しいなら T_y の符号を正, それ以外では T_y の符号を負とする.

5. 以下の式より, R を計算する.

$$R = \begin{bmatrix} r_1 & r_2 & \sqrt{1 - r_1^2 - r_2^2} \\ r_4 & r_5 & s\sqrt{1 - r_4^2 - r_5^2} \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (4.20)$$

r_7, r_8, r_9 は R が直交行列であるという性質を用いて求める.

6. このあと求める焦点距離 f が負となる場合は R を以下の式を用いる.

$$R = \begin{bmatrix} r_1 & r_2 & -\sqrt{1 - r_1^2 - r_2^2} \\ r_4 & r_5 & -s\sqrt{1 - r_4^2 - r_5^2} \\ -r_7 & -r_8 & r_9 \end{bmatrix} \quad (4.21)$$

ただし, $s = -\text{sgn}(r_1 r_4 + r_2 r_3)$ とする.

(d) レンズひずみを見無視して f, T_z の初期値を求める

それぞれの点 i について, f, T_z を未知数とする以下の1次方程式を解く.

$$\begin{bmatrix} y_i & -d_y Y_i \end{bmatrix} \begin{bmatrix} f \\ T_z \end{bmatrix} = w_i d_y Y_i \quad (4.22)$$

ただし,

$$y_i = r_4 x_{wi} + r_5 y_{wi} + T_y$$

$$w_i = r_7 x_{wi} + r_8 y_{wi}$$

とする.

(e) f, T_z, k_1 の正確な答えを求める

(d) で求めた f, T_z および $k_1 = k_2 = 0$ を初期値として, (4.13) においてもっとも誤差が少なくなるようにそれぞれの値を決定する.

以上の手法により, 空間中のある点における世界座標とそれに対応する画像座標が分かれば, カメラパラメータを決定できることが分かる.

4.3 関連研究におけるカメラキャリブレーション

佐藤らはキャリブレーションボード(図4.2)を用いてキャリブレーションを行った[3]. しかし, この方法には, 対象物が小さいときにしか用いることができないという問題点がある.

また, DebevecらはCANOMA(図4.3)というツールを作成し用いた[7]. しかし, このツールは画像中に直方体のようなものがないと, 正確にカメラパラメータを求めることができない, という問題点がある.

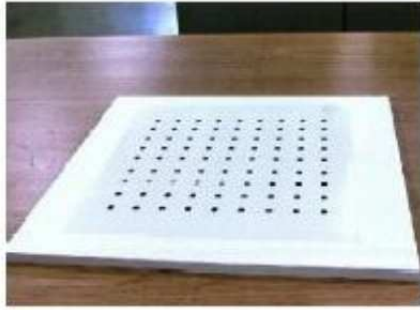


図 4.2: キャリブレーションボード

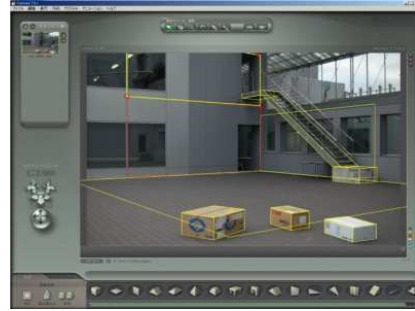


図 4.3: CANOMA

4.4 本研究におけるカメラキャリブレーション

本研究においては, 特徴点の世界座標とそれに対応する画像座標の組からキャリブレーションを行なう. その際, 対象物が巨大で複雑な形をしており, カメラと対象物との距離も大きく, また, 対応点を多くはとれないため, 4.2節で説明した方法では, 内部パラメータの誤差が大きい. そのため, 本研究では, 一部の内部パラメータを固定して, その他のパラメータを求めるという方法を用いる.

次に, 求められたカメラパラメータから, カメラの位置, カメラの向いていた方向, 水平方向, 垂直方向の画角を求める方法について説明する.

カメラの位置は, カメラ座標系における原点を世界座標系に変換したときのベクトルであると考えることが出来るので,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad (4.23)$$

を解くことによって求めることができる。

次に、カメラの向いている方向は、カメラ座標系において z 軸の方向であるので、(4.23) で求めたベクトルと (4.23) と同様にカメラ座標系の $(0, 0, -1)$ を世界座標系に変換したベクトルの差と考えることができる。

最後に画角であるが、まず水平方向に関しては図 4.4 より、

$$\tan \frac{\theta}{2} = \frac{\text{水平方向の CCD の大きさ} / \text{水平方向の画素数} / 2}{f} \quad (4.24)$$

であるから、 θ について解くと、

$$\theta = 2 \times \text{atan} \left(\frac{\text{水平方向の CCD の大きさ} / \text{水平方向の画素数} / 2}{f} \right) \times \frac{180}{\pi} \quad (4.25)$$

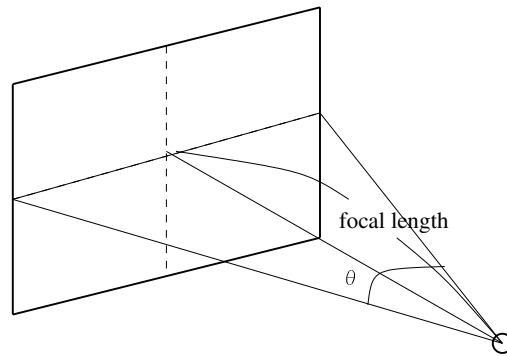


図 4.4: 画角について

として求める。同様に垂直方向の画角も求めることができる。

第5章 実験

本章では，本研究において行なった実験とその結果について述べる．また，本章で使用した画像の大きさはすべて720[pix]×480[pix]である．

5.1 幾何的整合性の実現

現在の大仏殿の写真と3次元モデルから，その写真を撮影した際のカメラの位置やカメラの向いていた方向などを求めた．以下，その方法について述べる．

5.1.1 3次元モデルの作成

現在の大仏殿の測量図[10](図5.1, 図5.2)からキャリブレーションを行なう際に必要となる大仏殿の3次元モデル(図5.3)を作成した．

5.1.2 実験結果

現在の大仏殿の3次元モデルと画像(図5.4)を用いてキャリブレーションを行なった結果が表5.1である．さらに，その結果からカメラの位置，カメラの向いていた方向，画角を求めたものが表5.2である．

また，このパラメータを用いて特徴点の世界座標から画像座標を求めた際の誤差は，最大11.320[pix]，平均5.330[pix]であった．

5.2 光学的整合性の実現

第3章で説明した手法で，複数の写真から光源画像を作成し，それをもとに第2章で説明した手法で違和感のない陰影で仮想物体を実画像に重ね込んだ．以下で，その方法について述べる．

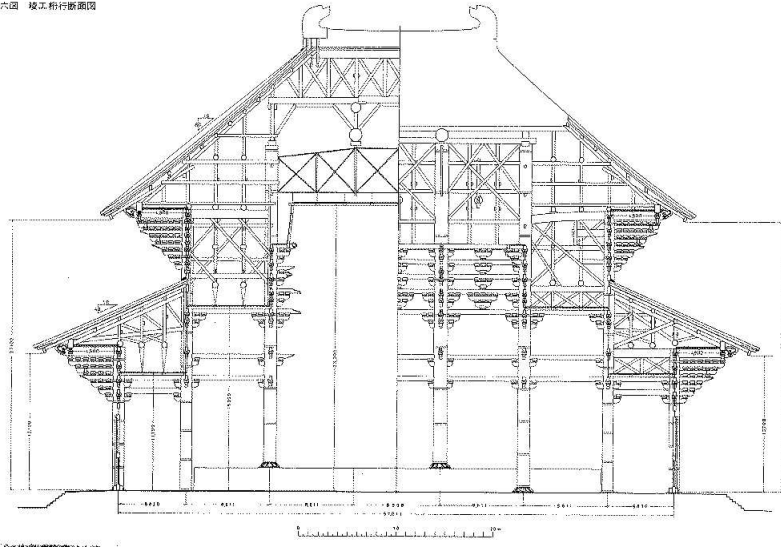


图 5.1: 测量图 1(视点:正面)

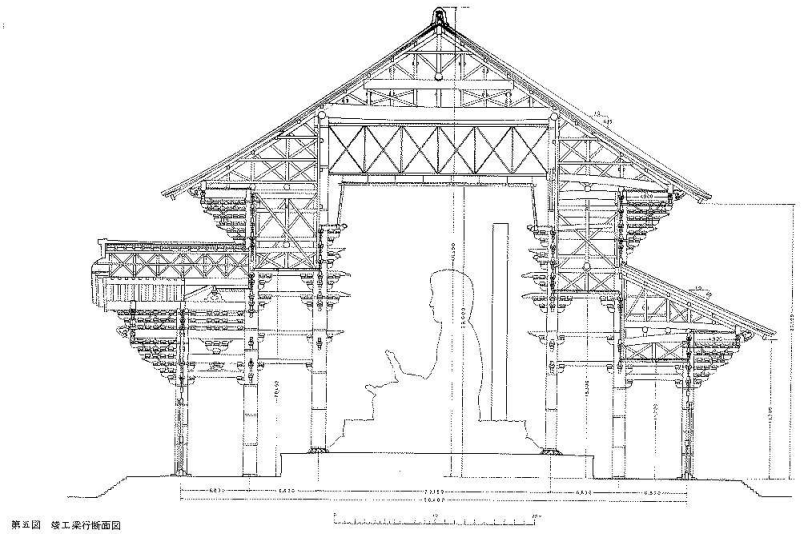


图 5.2: 测量图 2(视点:右)

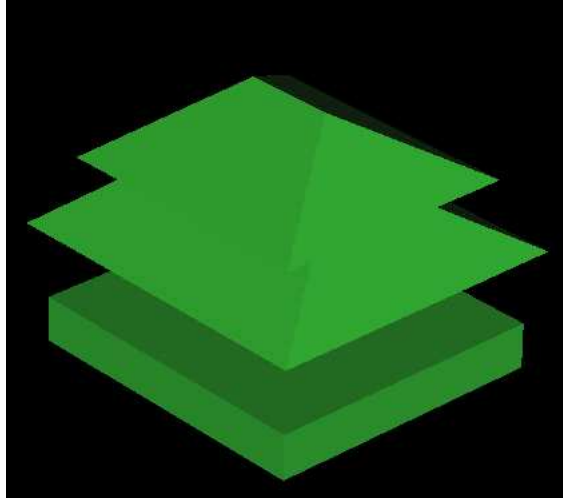


図 5.3: 現在の大仏殿の3次元モデル



図 5.4: 現在の大仏殿

表 5.1: キャリブレーションの結果

x 軸基準の回転角 (deg)	-136.667
y 軸基準の回転角 (deg)	54.415
z 軸基準の回転角 (deg)	142.391
平行移動ベクトル T の x 成分 (m)	-4.595
平行移動ベクトル T の y 成分 (m)	13.480
平行移動ベクトル T の z 成分 (m)	1361.208
水平方向の CCD の大きさ/水平方向の画素数を 0.72mm とした場合の焦点距離 (m)	7.497

表 5.2: 撮影時のカメラの状態

カメラの位置 x 方向 (m)	1100.111
カメラの位置 y 方向 (m)	544.447
カメラの位置 z 方向 (m)	588.576
カメラの向きの方向ベクトル x 成分	-0.813
カメラの向きの方向ベクトル y 成分	-0.399
カメラの向きの方向ベクトル z 成分	-0.423
水平方向の画角 (deg)	5.498
垂直方向の画角 (deg)	3.667

5.2.1 光源画像の作成

図 5.5 の 20 枚の写真から、第 3 章で説明した手法を用いて作成した光源画像 1 が図 5.6, 同様に、図 5.7 の 25 枚の写真から作成した光源画像 2 が図 5.8 である.

5.2.2 画像生成

第 2 章で説明した手法を用いて画像を生成する.

まず、表 5.2 のパラメータを用いて、仮想物体の部分のマスク画像 (図 5.9), 地面の部分のマスク画像 (図 5.10), さらに、5.2.1 で作成した光源

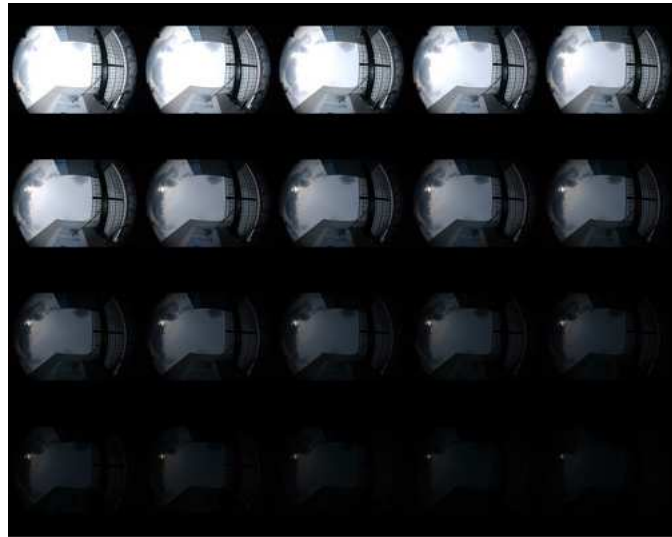


図 5.5: シャッタースピードを 125 分の 1 秒から 16000 分の 1 秒の間で変化させて撮影された 20 枚の写真



図 5.6: 光源画像 1

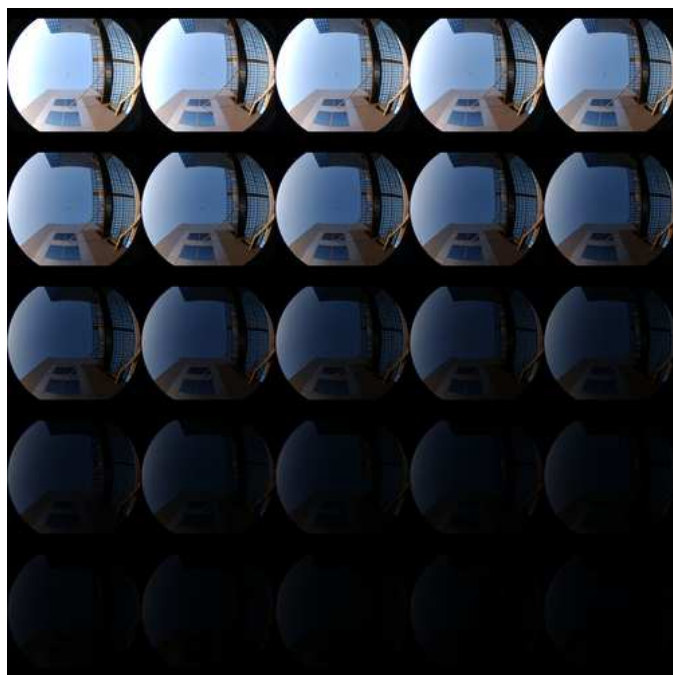


図 5.7: シャッタースピードを3分の1秒から800分の1秒の間で変化させて撮影された25枚の写真



図 5.8: 光源画像2

画像 (図 5.6, 図 5.8) を用い, 光源環境を考慮した, 仮想物体がない場合の地面の照度を示す画像 (図 5.11, 図 5.15), 仮想物体がある場合の地面の照度と仮想物体を示す画像 (図 5.12, 図 5.16) をそれぞれ作成する.

以上の画像と, 実画像から 2.3 節で説明した方法で仮想物体を実画像に重ね込んだ結果が図 5.14, 図 5.18 である. また, 図 5.13, 図 5.17 は陰影を考慮せずに重ね込んだ結果である.

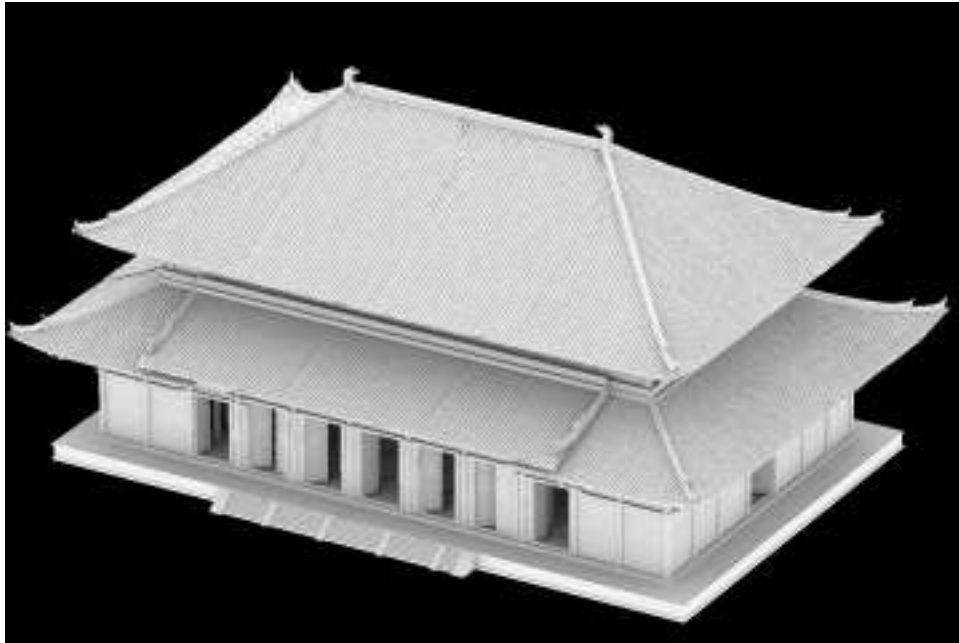


図 5.9: 仮想物体のマスク画像

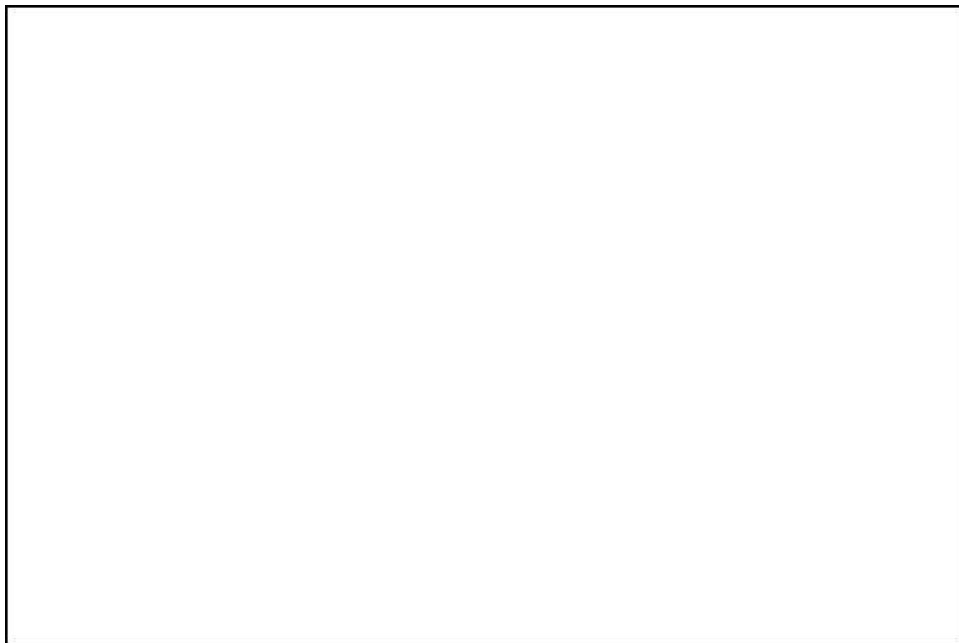


図 5.10: 地面のマスク

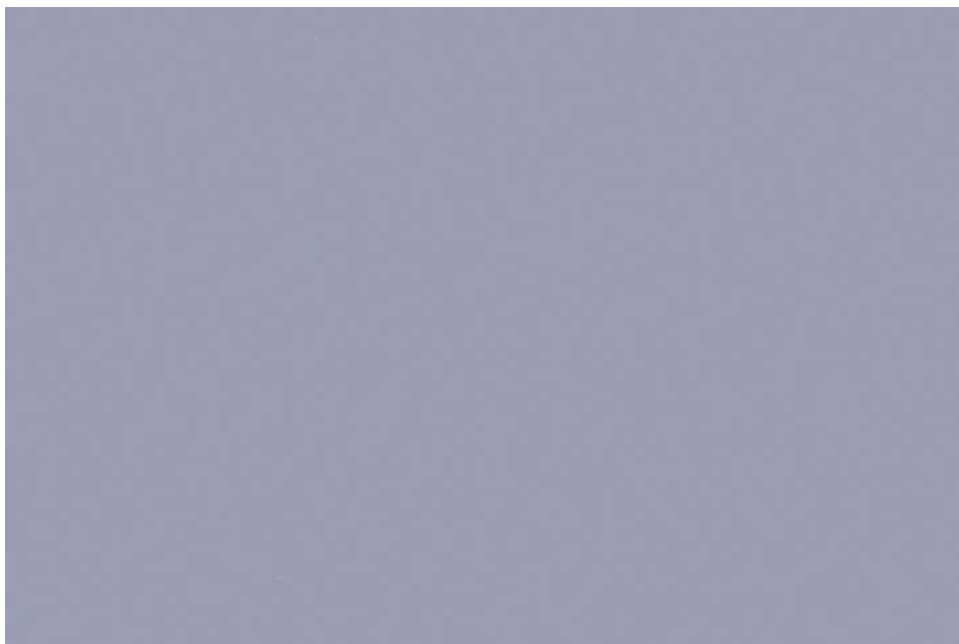


図 5.11: 光源画像 1 を用いた仮想物体がないときの地面の照度

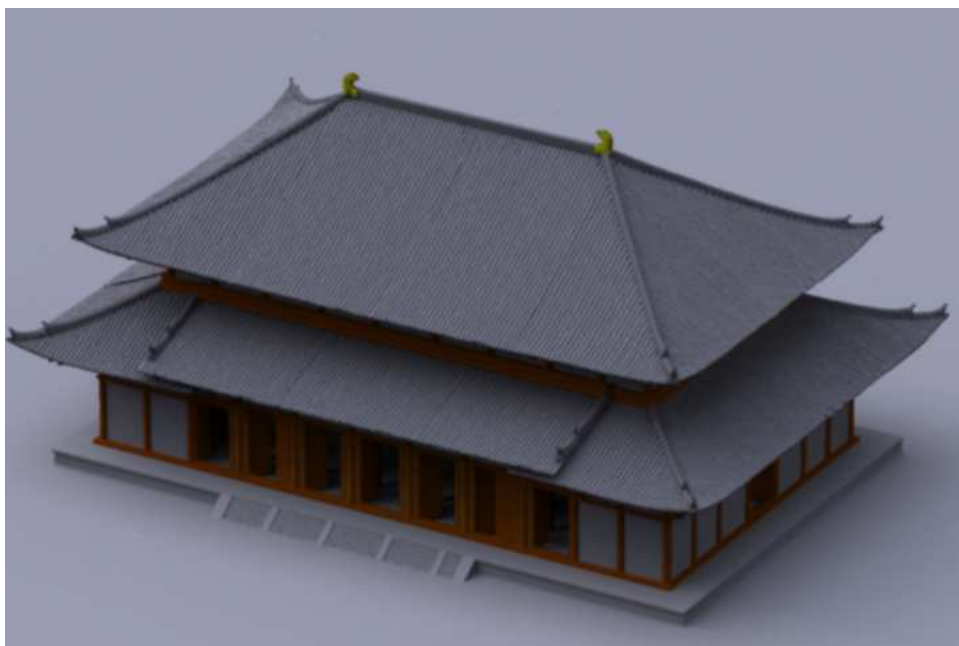


図 5.12: 光源画像 1 を用いた仮想物体があるときの地面の照度と仮想物体



図 5.13: 光源画像 1 を用いて実画像に仮想物体を重ね込んだ結果(影無し)



図 5.14: 光源画像 1 を用いて実画像に仮想物体を重ね込んだ結果(影あり)



図 5.15: 光源画像 2 を用いた仮想物体がないときの地面の照度

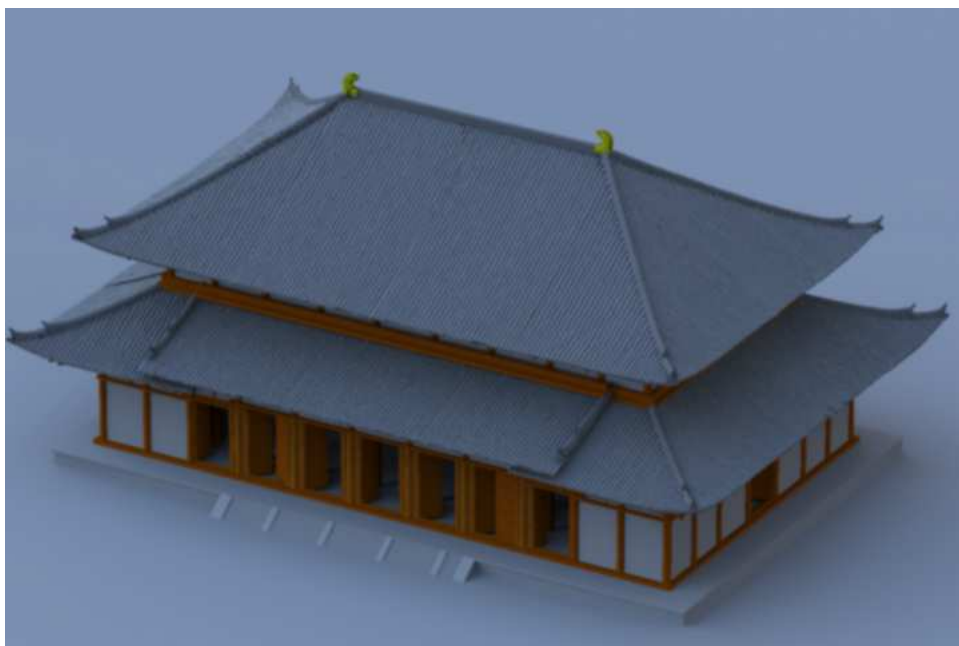


図 5.16: 光源画像 2 を用いた仮想物体があるときの地面の照度と仮想物体



図 5.17: 光源画像 2 を用いて実画像に仮想物体を重ね込んだ結果 (影無し)



図 5.18: 光源画像 2 を用いて実画像に仮想物体を重ね込んだ結果 (影あり)

第6章 考察

6.1 光学的整合性について

本研究では，創建時に図5.5, 図5.7のような天候であったとき，大仏殿の影響でどのような影がついていたかを再現した．しかし，実際のそのような画像がない以上その比較は当然不可能である．しかし，図5.11, 図5.12または図5.17, 図5.18をそれぞれ見比べると，影があるために現実感が増していることが分かる．

また，ハイダイナミックレンジを実現した画像を光源画像として用いることの有効性については，第3章で示したとおりである．

6.2 幾何的整合性について

本研究では，違和感のない位置に仮想物体を重ね込むことで幾何的整合性の実現を試みた．違和感のない位置というものが創建時に大仏殿があった位置と考えると，この2つを比較することは当然できない．しかし，画像を見る限り，ほぼ違和感のないものが作成できている．

また，5.1節で述べたような誤差が生じる原因としては，キャリブレーションを行なう際に用いた対応点の少なさによることが考えられる．対応点を更に減らしてキャリブレーションを行ない，その結果求められたカメラの位置，カメラの方向，画角を用いた結果が，図6.1である．大仏殿の前面に関しては正しいが，側面に関しては地面とあっていないため全体として違和感を感じる．

このように，対応点は多ければ多い程良いと考えられるが，図5.3のように作成した現在の大仏殿の3次元モデルが，測量図をもとにしている以上これ以上増やすのは難しい．



図 6.1: 違和感のある配置

6.3 実画像中の影の削除

実画像中の影の削除は、本研究のように仮想物体を建造物がある位置に重ね込む場合必要な作業である。手法としては、2.5節で説明した方法で行なうことができる。しかし、正確に行なうためには、光源画像が正確なものではなくてはならない。以下の図は、2.5で説明した手法で影を削除しようとした例である。

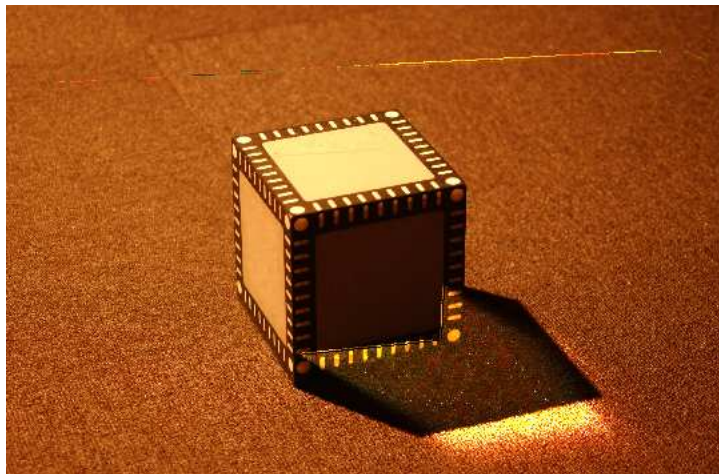


図 6.2: 影を画像中から削除しようとして失敗した例

図6.2を見て分かるように、光源画像が実際の光源環境とは異なるために、実際の影が残ってしまっているうえに、影が無い部分が逆に明るくなっている。

第7章 結論と今後の展望

7.1 結論

現在の東大寺大仏殿の位置に仮想物体である創建時の大仏殿を復活させた画像を作成した。巨大建造物を仮想物体として、現実世界に融合させることができたといえる。

光学的整合性に関しては、仮想物体の影を実世界光源環境に基づいて実画像上に生成するために、radianceを用い、さらに、ハイダイナミックレンジの手法の改良を行なったことで、これまでの手法に比べ、より正確な陰影での融合を行なうことができ、実現している。

また、幾何的整合性に関しても、キャリブレーションを工夫して行なうことで、より正確に仮想物体を配置することができており、実現しているといえる。

7.2 今後の展望

7.2.1 ハイダイナミックレンジのさらなる改良

本研究では、ハイダイナミックレンジを実現するために、既存の手法を用いると、サチュレーションを起こして、ピクセルの値が0になってしまう部分に、そのあとに値を代入するという方法を用いたが、そのほかにも、フィルターを用いて、カメラに入って来る光の量を減少させるという方法が考えられる。また、技術の向上により、カメラのシャッタースピードを更に速くすることができれば、そもそもサチュレーションを起こさなくなることも考えられる。

7.2.2 実画像中の影の削除

6.3節でも述べたように，ハイダイナミックレンジが更に改良され，光源画像が実際の光源環境を完全に再現することができたなら，2.5節の手法を用いることにより，実画像中の影を削除が可能になると考えられる．

7.2.3 他の建造物への応用

本研究では，東大寺の大仏殿を巨大建造物の例として用いたが，他の建造物であっても，過去と現在のそれぞれの建造物の3次元モデルと現在の写真があれば，当然過去の3次元モデルを仮想物体として実画像に重ね込むことはできる．また，現在の建造物の精密な3次元モデルがあれば，キャリブレーションの精度も上がり，さらに現実感の増した画像を作成することも可能だろう．

謝辞

本論文の執筆にあたり原田 賢一教授をはじめ、多くの方々に多大なる御指導及び御助言を頂いたことをこの場を借りて深く御礼申し上げます。

特に、私の我儘により東京大学生産技術研究所池内研究室での研究を薦めて頂いた斎藤博昭専任講師に深く感謝致します。おかげで素晴らしい研究生活だけでなく、新たな多くの出会いを経験することができました。これらの出会いは今後の私の人生においてかけがえの無い財産になると思います。

東京大学生産技術研究所池内研究室においては、十分すぎるほどの研究環境を提供して頂くとともに、お忙しい日々を送っていらっしゃる中で時間を割いて様々な助言を与えて下さった池内克史教授に心から感謝致します。また、日頃から私に様々な手助け、温かい励ましを頂いた池内研究室の諸先輩方にはこの場をかりて深く御礼申し上げます。

斎藤研究室に顔を出す機会は少し減ってしまいましたが、随所で色々助けてもらいました。本当にありがとうございました。

また、遠くから常に私の健康を気遣ってくれた両親、大きな心の支えとなってくれた素晴らしい友人達にこの場を借りて御礼を言いたいと思います。

2002年2月

参考文献

- [1] 田村 秀行 “複合現実感研究プロジェクト “ 信学ソ大, ISS-1-4, 1997
- [2] 大石 岳史, “文化財のデジタル保存と復元”, 東京大学修士論文, 2002
- [3] Imari Sato, Yoichi Sato, Katsushi Ikeuchi “Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene” IEEE Transactions on Visualization and Computer Graphics, Vol.5, No.1, pp. 1-12, January-March 1999
- [4] Pual E. Debevec, Jitendra Malik “Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography” , Proceeding of SIGGRAPH 98, Computer Graphics Proceedings, Annual Conference Series, pp. 189-198
- [5] <http://radsite.lbl.gov/radiance/>
- [6] Pual E. Debevec, Jitendra Malik “Recovering High Dynamic Range Radiance Maps from Photographs” Proceeding of SIGGRAPH 97, Computer Graphics Proceedings, Annual Conference Series, pp. 369-378
- [7] 浅田 尚紀 “カメラキャリブレーション “ コンピュータビジョン 技術 評論と将来展望 第3章, 新技術コミュニケーションズ, 1998
- [8] 徐 剛, 辻 三郎 “3次元ビジョン”, 共立出版, 1998
- [9] Tsai, R.Y. “A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses.” IEEE Journal of Robotics and Automation 3(4):pp.323-344, August 1987

- [10] 東大寺大仏殿昭和大修理修理委員会 “国宝東大寺金堂 (大仏殿) 修理工事報告書” 奈良県教育委員会事務局奈良県文化財保存事務所編, 1980.9