

修士論文

写実的なレンダリングのための偏光により分離された反射成分別圧縮と補完

Compression and Complementation of
Reflection Components Separated with
Polarization technique for
Photorealistic Rendering

指導教官 池内 克史 教授

東京大学大学院情報理工学系研究科
電子情報学専攻

T

氏名 46415 柴田 卓司

提出日 平成 18 年 2 月 3 日

概要

近年，コンピュータ性能の飛躍的な進歩により三次元物体を扱うことが容易になり，医療，教育，娯楽，芸術，デジタルアーカイブ等様々な分野において複合現実感の技術が利用されている．この複合現実感とは実世界空間と仮想空間が融合された空間における現実感のことである．このような背景のもとに，現実感の高い写実的な三次元物体画像の需要が増えている．現在，人の手によってモデリングを行うのはコストが高く，時間もかかる．そこでモデリングを自動的に行うために，現実物体を観測することにより，仮想現実感システム用のモデルを生成する研究が盛んに行われている．明るさ情報はカラーセンサーにより得られるが，視線方向や光源状況によって変化する物体の見えを生成することは重要な研究テーマである．そこで本研究では，複合現実感において，物体の見えを構成する明るさ情報に着目し，任意の状況下での写実的なレンダリング手法を提案する．

本研究では写実性において重要な表面反射が正確に再現できること，大規模物体を対象とした場合に膨大となるデータ量の削減を重視したシステムを目指す．表面反射の正確な再現において，鏡面反射成分と拡散反射成分の二つの反射成分の性質は大きく異なるため，それぞれを個別に扱う必要がある．よって光の反射を二色性反射モデルを用いて数学的に定式化し，その反射パラメータを推定することで厳密に求める．また反射成分の分離には円偏光の性質をも利用した分離法を提案し，よりロバストな反射成分分離を導入した．さらに提案手法では反射モデルで表現しきれない成分についても再現するため，幾何モデルを用いて三角パッチ状に展開した成分ごとの入力画像と反射モデルによるレンダリング画像の差分成分を保持し，圧縮した後に画像合成に用いることでレンダリング精度を高めている．

目次

第 1 章	序論	1
1.1	研究背景	2
1.2	関連研究	3
1.3	本研究の目的	4
1.4	本論文の構成	6
第 2 章	物体表面反射	7
2.1	鏡面反射	8
2.2	拡散反射	9
2.3	二色性反射モデル	10
2.4	様々な反射モデル	12
2.4.1	Lambertian モデル	12
2.4.2	Phong モデル	12
2.4.3	Torrance-Sparrow モデル	13
第 3 章	偏光	17
3.1	偏光	18
3.1.1	直線偏光	18
3.1.2	円偏光	21
3.2	偏光による反射成分分離	22
第 4 章	幾何情報	27
4.1	距離画像取得	28
4.2	アライメント	29
4.3	マーキング	29
4.4	表面法線	30

第5章	カメラキャリブレーション	32
5.1	カメラキャリブレーション	33
5.1.1	カメラパラメータ	33
5.1.2	世界座標系とカメラ座標系	34
5.1.3	同次座標	35
5.1.4	三次元アフィン変換	35
5.1.5	透視変換	36
第6章	反射パラメータ推定	39
6.1	拡散反射パラメータ推定	40
6.2	鏡面反射パラメータ推定	41
第7章	ウェーブレット	42
7.1	基礎理論	43
7.1.1	連続ウェーブレット変換	43
7.1.2	離散ウェーブレット変換	45
7.2	二次元離散ウェーブレット変換	50
7.3	三次元離散ウェーブレット変換	53
7.4	離散ウェーブレット変換による情報圧縮・再構成	55
7.4.1	画像圧縮：二次元離散ウェーブレット変換による情報圧縮・再構成	56
7.4.2	画像列（動画）圧縮：三次元離散ウェーブレット変換による情報圧縮・再構成	59
第8章	仮想現実感モデル生成システム	60
8.1	光学情報	61
8.2	幾何情報	65
8.3	キャリブレーション	65
8.4	反射パラメータ推定	66
8.5	Torrance-Sparrow レンダリング	67
8.6	差分画像	68
8.7	ドベシウエーブレットによる差分画像圧縮	70
8.8	差分圧縮画像補完付きレンダリング	71

第9章	実験・評価	73
9.1	データ取得	74
9.2	実験	74
9.3	評価	81
9.3.1	データ量	81
9.3.2	レンダリング精度	90
第10章	結論	96
10.1	結論	97
10.2	今後の課題	98
	謝辞	99

目次

2.1	物体表面における光の反射	9
2.2	二色性反射モデル	11
2.3	Phong モデル	14
2.4	Torrance-Sparrow モデル	16
3.1	直線偏光	19
3.2	直線偏光の重ね合わせ	20
3.3	直線偏光の性質	20
3.4	直線偏光板による光の遮断	21
3.5	円偏光	22
3.6	円偏光の性質	23
3.7	円偏光板による光の遮断	24
3.8	円偏光板による反射性成分の分離	25
3.9	拡散反射成分と鏡面反射成分 (明るさは実際の 4 倍)	26
4.1	距離画像	28
4.2	アライメント	29
4.3	PCA による法線計算	31
5.1	世界座標系とカメラ座標系	34
5.2	透視変換	37
7.1	ウェーブレット	44
7.2	ドベシウーブレット	45
7.3	ハールウェーブレット	46
7.4	ガボールウェーブレット	46
7.5	スケーリング関数の分解	50
7.6	スケーリング関数の再構成	50
7.7	横軸方向の二次元ウェーブレット変換	51

7.8	縦軸方向の二次元ウェーブレット変換	52
7.9	二次元の多重解像度分解	53
7.10	三次元ウェーブレット変換	54
7.11	三次元の多重解像度分解	55
7.12	画像の分解・再構成過程	56
7.13	Mandrill 画像の圧縮	57
7.14	Parrots 画像の圧縮	58
8.1	フローチャート：圧縮過程	62
8.2	フローチャート：再構成過程	63
8.3	上：全周囲の画像，下：反射成分分離	64
8.4	距離画像から 3D モデルの生成	65
8.5	上：2D-3D 対応，下：幾何モデルの簡略化	66
8.6	左：拡散反射パラメータ，右：鏡面反射パラメータ	67
8.7	Torance-Sparrow レンダリング画像	68
8.8	パッチ形式差分画像	69
8.9	分解過程	70
8.10	再構成過程	71
8.11	パッチごとの画像圧縮	72
8.12	パッチ形式差分圧縮画像補完レンダリング画像	72
9.1	データ取得システム	74
9.2	対象物体	75
9.3	(上) fish の入力画像 (偏光板無) (下) dinosaur,hyottoko,bear の入力画像 (偏光板無)	76
9.4	(上) 拡散反射成分画像・鏡面反射成分画像 (下) 三次元幾何モデル	77
9.5	fish の拡散反射パラメータ	79
9.6	bear,hyottoko,dinosaur の拡散反射パラメータ	80
9.7	拡散成分差分画像	81
9.8	結果画像	82
9.9	fish の結果画像列	83
9.10	展開係数の使用比率と再構成誤差	84
9.11	hyottoko 展開係数の使用比率ごとの再構成画像	85

9.12 Eigen-Texture 法との比較:(a),(d) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分非分離) , (b),(e) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分分離) (c),(f) はそれぞれ係数使用比率 26 %,2 %の提案手法	87
9.13 Eigen-Texture 法との比較 (差分画像) : (a),(d) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分非分離) , (b),(e) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分分離) (c),(f) はそれぞれ係数使用比率 26 %,2 %の提案手法	88
9.14 bear の切断面における輝度比較 (ウェーブレット圧縮率別)	91
9.15 hyottoko の切断面における輝度比較	94
9.16 bear の切断面における輝度比較 bear の切断面 (レンダリング手法別)	95

表 目 次

1.1	他手法との比較	5
7.1	$N = 2$ のドベシイの数列 p_n	48
9.1	西野らの手法と提案手法における RMSE と展開係数使用比率の関係	86

第1章

序論

1.1 研究背景

近年，コンピュータ性能の飛躍的な進歩により三次元物体を扱うことが容易になり，医療，教育，娯楽，芸術，デジタルアーカイブ等様々な分野において複合現実感の技術が利用されている．この複合現実感とは実世界空間と仮想空間が融合された空間における現実感のことである．このような背景のもとに，現実感の高い写実的な三次元物体画像の需要が増えている．現在，人の手によってモデリングを行うのはコストが高く，時間もかかる．そこでモデリングを自動的に行うために，現実物体を観測することにより，仮想現実感システム用のモデルを生成する研究が盛んに行われている．複合現実感に関する研究において重要となるのは，

- 物体の立体的な情報を表す幾何情報
- 物体表面の見え (色) を表す明るさ情報
- 周囲の環境や状況を表現する環境情報

の三つである．幾何情報は現実物体をレーザーセンサーで計測することにより容易に得ることができる．明るさ情報はカラーセンサーにより得られる．しかし視線方向や光源状況によって変化する物体の見えを生成することは重要な研究テーマである．そこで本研究では，複合現実感において，物体の見えを構成する明るさ情報に着目し，任意の状況下での写実的なレンダリング手法を提案する．

この問題に対するアプローチの一つに，光の反射モデルを導入して物体表面の反射を数学的に定式化し，物体表面におけるパラメータを推定することにより任意の状況下での仮想物体画像を合成するモデルベースレンダリング (MBR:Model Based Rendering) というコンピュータビジョンのテクニックがある．この手法は三次元幾何モデルと反射モデルを用いて物体表面の各点における反射パラメータを推定しているため，少ないデータ量で任意の状況下での仮想物体画像の合成が精度良く行えるという利点がある．しかし表面反射モデルは適用できる物体に制約があり，物体表面の形状が複雑で反射を定式化できない場合には精度のよい仮想現実物体画像が得られないという欠点がある．また表面がふさふさしているぬいぐるみのような物体ではそもそも厳密な三次元幾何モデルを得ることが難しく，反射パラメータ推定で必要になる法線が正確に得られないためうまくいかない．

また別のアプローチに，撮像された実画像の集合から仮想物体画像を合成するイメージベースレンダリング (IBR:Image Based Rendering) というコンピュータ

ラフィックスのテクニックがある．物体を様々な光源状況や視線方向の下で撮像した実画像をデータベースとして保持しておき，特定の状況下の仮想物体画像を得る際に，このデータベースからテクスチャを取り出したり補完により合成するため，物体個々の反射特性に制限を受けず，どのような物体にも適用できるという特徴を持つ．一方，反射モデルを用いずに仮想物体画像を合成するため光源状況の変化に合わせた正確な表面反射を再現することは難しく，複合現実感システムにおいて仮想空間と実世界空間の融合を現実感高く行うことに適していない．また，この手法は大量の実画像をデータベースとして保持する必要があり，データ量が大きくなってしまう．

1.2 関連研究

Gortler らによる Lumigraph[1] では，任意の光源方向からの光線を表す Lumigraph という四次元光学関数を用いて入力となる複数枚の画像より物体形状を復元して画像合成を行う．しかしこの物体形状はあくまでも光線空間における基底関数の補正を行うための大まかな距離値を求めるための非常に粗い三次元幾何モデルであるため，光源状況を考慮した正確な陰影を付けられない．Levoy らによる Light Field Rendering[2] では，Light Field と呼ばれる四次元の光学関数の二次元スライスを入力画像として，物体形状を用いずに任意の視点からの画像を合成する．しかしこの手法は光源が固定されており，またサンプル間隔の密な画像が必要なためデータ量が大きくなってしまう．

Wood らによる Surface Light Field[3] は精密な三次元幾何モデルを導入し，また Lumisphere と呼ばれる物体表面から放射される光線におけるポイントベースの色情報を用いている．さらにデータ量削減のために主成分分析 PCA とベクトル量子化を行っている．Miller ら [4] はこの Surface Light Field のアイデアを推し進めて，ブロックごとの圧縮を行っている．また Chen ら [5] は Surface Light Field をグラフィックハードウェアによりリアルタイムレンダリングできるよう高速化した．

西野らによる Eigen-Texture 法 [6] は複数枚の入力画像を精密な三次元幾何モデルの三角パッチの形状に Eigen-Texture として展開し，これを PCA で圧縮してデータ量を削減し，任意の状況下のレンダリングを行う．古川ら [7] は BTF:Bidirectional Texture Functions[8] として任意の状況下の画像をテンソル積展開により圧縮し，レ

レンダリングを行った．また物体形状の変形にも適応している．Debevec ら [9][10][11] は密にサンプリングした画像から reflectance field を獲得して任意の光源下の人間の顔のモデルをつくった．

佐藤ら [12] は拡散反射成分と鏡面反射成分に分離された明るさ情報から二色性反射モデルのパラメータを推定し，画像を合成した．この分離操作は色情報による．また高橋ら [13] は反射成分の分離において，カラーセンサーと光源の前に直線偏光板を置き，偏光の性質を利用してよりロバストに分離を行った．Ma ら [14] は入力画像の BTF を多重ラプラス変換と PCA により圧縮し，レンダリングにおいて再構成された BTF と反射モデルにより画像合成を行った．

1.3 本研究の目的

本研究の目的は任意の状況下での写実的なレンダリングをおこなうことにある．関連研究で述べたように任意の状況下での仮想物体画像を合成する手法には様々な手法がある．それぞれに長所と短所があり，目的によってアプローチが異なってくる．本研究では写実性において重要な表面反射が正確に再現できること，大規模物体を対象とした場合に膨大となるデータ量の削減を重視したシステムを目指す．表面反射の正確な再現において，鏡面反射成分と拡散反射成分の二つの反射成分の性質は大きく異なるため，それぞれを個別に扱う必要がある．よって光の反射を二色性反射モデルを用いて数学的に定式化し，その反射パラメータを推定することで厳密に求める．また反射成分の分離には円偏光の性質をも利用した分離法を提案し，よりロバストな反射成分分離を導入した．

しかし，反射モデルは現実の反射を簡略化した表現で，また前述したように反射モデルには制限があるためにいつも正確な反射を表現できるとは限らない．そこで提案手法では反射モデルで表現しきれない成分についても再現するため，幾何モデルを用いて三角パッチ状に展開した成分ごとの入力画像と反射モデルによるレンダリング画像の差分成分を保持し，画像合成に用いることでレンダリング精度を高めている．

さらにデータ量の観点から，この差分成分の集合を圧縮する必要がある．圧縮は再構成画像の精度をできるだけ落とさずおこなう必要があるが，圧縮においても性質が異なる二つの成分を一緒に扱うのではなく，個別に圧縮することで画像の合成を精度良くできると考える．本研究では三次元離散ウェーブレット変換による多重

解像度解析で圧縮を行う．圧縮対象は異なる状況下で撮像された二次元の固有な三角パッチ画像を並べた三次元配列である．三角パッチ画像の二次元の相関に加えて，もう一次元においも固有な三角パッチであるため相関性が高く，また三角パッチごとに固有に圧縮率を変えることで効率的な圧縮ができる．レンダリングにおいては再構成された反射成分ごとの三角パッチ差分画像列を，反射モデルによるレンダリングに補完して合成画像とする．

関連研究で述べた手法のうち，本研究が取り組んだ，圧縮の特化したものと反射モデルをもちいているものを選びその特徴をまとめたのが表 1.1 である．圧縮法についての比較をおこなうと，PCA とテンソル積展開は係数に加えて基底も保存しないといけなないので十分な圧縮ができない．ブロック DCT(FFT)，ラプラス変換，ウェーブレット変換は基底がそれぞれ余弦関数，指数関数，ウェーブレットであり，係数のみの保存となるが，この中でもウェーブレットの精度が最もよいことが知られている．また，MPEG では動き補償をしており，隣り合ったフレーム同士の差分画像を個別に圧縮しているが，提案手法では入力画像と反射モデルによるレンダリング画像の差分画像をまとめて圧縮しているという違いがある．この表より提案手法はレンダリング精度・データ量の両面において優れた手法であるといえる．また近年，携帯電話を始めモバイル機器の高性能化に伴い，モバイル機器上でも二次元コンテンツに取って代わって三次元コンテンツの需要が拡大すると予想されるが，本研究の提案手法は高い圧縮を実現できることにより，これらの分野でも有効である．

表 1.1: 他手法との比較

	圧縮法	拡散反射成分	鏡面反射成分	分離
MPEG	ブロック DCT	-	-	-
Eigen-Texture 法	PCA	-	-	-
古川	テンソル積展開	-	-	-
Ma	ラプラス変換 + PCA	Lambertian	Phong	-
佐藤	-	Lambertian	Torrance-Sparrow	色
高橋	-	Lambertian	Torrance-Sparrow	直線偏光
提案手法	ウェーブレット変換	Lambertian	Torrance-Sparrow	円偏光

1.4 本論文の構成

本論部の章立てを以下に示す。

第 2 章では、光の物体表面での反射のメカニズムについて述べ、この光の反射を数学的に定式化した反射モデルの説明をする。

第 3 章では、反射成分を分離するための偏光について述べる。偏光の種類・性質などを述べた後、偏光による反射性分離の操作を説明する。

第 4 章では、物体の形状を表す三次元幾何モデルについて、そのモデル構築法や様々な原理について説明する。

第 5 章では、世界座標系とカメラ座標系という異なる二つの座標系について説明し、その二つを結びつけるカメラパラメータの性質・変換などを述べた後、カメラパラメータを求めるカメラキャリブレーションについて説明する。

第 6 章では、第二章で説明した表面反射モデルにおける本研究での具体的なパラメータ推定のアルゴリズムを説明する。

第 7 章では、ウェーブレットの基礎理論に始まり、様々なウェーブレットの種類、変換の種類を述べ、ウェーブレットと変換を用いた信号圧縮として、画像圧縮と動画圧縮について説明する。

第 8 章では、提案手法の画像合成システムを述べる。ここではこれまでの章で述べてきた様々な技術や理論を統合し、入力データからいかにして画像合成を行うかを一連のシステムとして説明する。

第 9 章では、提案手法のシステムを用いて実物体を対象行った実験について述べる。ここではデータ量・レンダリング精度の両面から評価を行い、また他手法との比較を行うことにより、提案手法の有効性を検証する。

第 10 章では、本論文の結論をまとめる。

第2章

物体表面反射

本章では物体表面の光の反射について説明する。カメラを用いて物体を撮像すると、物体表面への入射光は物体表面で反射され、その反射光のうちカメラに到達したものが撮像された画像の各画素値になる。したがって画像を解析する際にはカメラの特性のみならず、物体表面における光の反射特性を解析することが重要である。

実世界に存在する物質には光学的に均質な物体と不均質な物質がある。光学的に均質な物質とは金属や水晶などに代表されるように材料がミクロに均質な物質のことであり、これらの物体では物体表面への入射光は拡散せず、フレネルの法則に従って完全鏡面反射方向に光が反射する。これに対し、光学的に不均質な物質は図 2.1 に示すように物体表面、中間物質、顔料で形成されている。一般的な物体はこのような不均質な物質で形成されているため、2.3 で説明する物体表面反射モデルは物体が光学的に不均質な物質で形成されていることを仮定している。これを二色性反射モデル dichromatic reflection model という。二色性反射モデルによると、物体表面の光の反射により二つの成分が観測されることが知られている。一つが鏡面反射成分 specular reflection components, もう一つが拡散反射成分 diffuse reflection components である。

以下では鏡面反射成分、拡散反射成分、二色性反射モデル、幾つかの反射モデルについて述べる。

2.1 鏡面反射

鏡面反射とは物体外部の物質（大気）と物体の屈折率の違いにより物体表面において起こる反射である。鏡面反射は以下の性質を持つ。

1. 表面法線に対して完全鏡面反射方向に存在。
2. 反射される光の量はフレネルの法則に従う。

物体表面が完全に滑らかであれば 1 が成り立つが、物体表面に粗さがあると鏡面反射は完全鏡面反射方向付近で拡散して放射される。また 2 より、反射される光の量は入射角、物体の屈折率、入射光の偏光に比例するが、屈折率は波長に比例するため、鏡面反射は波長に比例することになる。しかし、一般的に可視光の範囲では波長に対する屈折率の変化は一定と近似できるほどに微小であるため、結局鏡面反射は波長に対して一定と近似することができ、したがって鏡面反射光の色は入射光の色と等しいとすることができる。

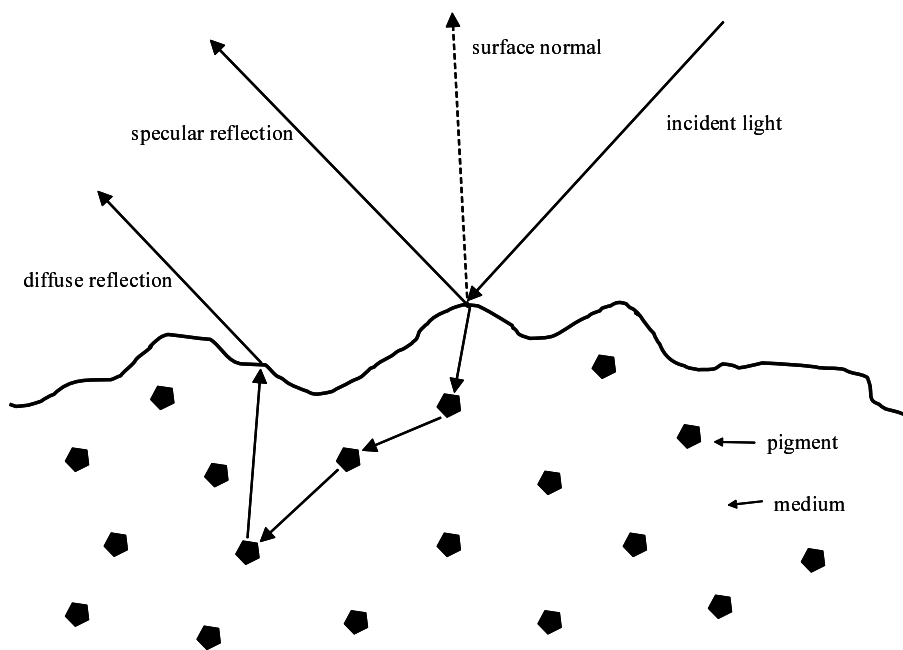


図 2.1: 物体表面における光の反射

2.2 拡散反射

物体表面で反射されなかった入射光は物体の内部に透過する。物体の内部に入った光は顔料にぶつかりながら中間物質を進むが、この光の一部は再び物体表面を通過して外部に放射される。これを拡散反射という。拡散反射には以下の要素の影響を受ける。

1. 中間物質の伝送特性
2. 顔料の拡散と吸収
3. 顔料の形状と分布

一般的に中間物質は透明であると近似でき、光の色や方向に影響を与えずに伝送すると考えてよいので 1 の影響は無視できる。鏡面反射の色は 2 に挙げたように、光が顔料にぶつかった際に波長ごとに光が吸収されるので入射光の色とは異なる色になる。すなわち物体の色を決定しているのはこの顔料の吸収特性であるといえる。また、拡散反射の方向は 3 の顔料の分布により決まるが、一般的に顔料は一様に分

布していると近似できるため，ランバートの法則に従うと考えてよい．次節で述べる物体表面反射モデルの多くは拡散反射にランバーシアンモデルを用いて表現している．

2.3 二色性反射モデル

二色性反射モデルとは，S.A.Shafer によって提案された反射モデルであり [15]，多くの反射モデルの基礎となっている．二色性反射モデルでは物体表面における反射は図 2.2 に示す面法線中心の座標系を用いると式 (2.1)(2.2) で表される．

$$I_c = I_{d,c} + I_{s,c} \quad (2.1)$$

$$= m_{d,c}c_{d,c} + m_{s,c}c_{s,c} \quad (c = R, G, B) \quad (2.2)$$

ただし，

- I_c : 反射光全体の放射輝度
- $I_{d,c}$: 拡散反射の放射輝度
- $I_{s,c}$: 鏡面反射の放射輝度
- $m_{d,c}, m_{s,c}$: 幾何学的強度 ($0 \leq m_{d,c}, m_{s,c} \leq 1$)
- $c_{d,c}, c_{s,c}$: スペクトルパワー分布

すなわち，反射光の放射輝度は拡散反射と鏡面反射の放射輝度の線形和で表すことができる．この二色性反射モデルが成り立つ条件として S.A.Shafer は以下の仮定をしている．

- 物体表面は不均質な物質でできており，かつ不透明である．また物体内部の顔料は一様に分布しており，物体表面から自ら光を放射していない．
- 物体表面からの反射光は面法線に対して等方性を持つ
- 光源のスペクトルパワー分布はシーン全体にわたって一定である．
- 相互反射は無視する．

これらの仮定により二色性反射モデルの一般性が失われることはなく、以下にあげる点において実世界に即した反射モデルであるといえる。

- 正射影や透視射影を用いることを仮定していないので、曲面や平面や巨視的に粗い面に対しても等しく適用することができる。
- $m_{d,c}$, $m_{s,c}$, $c_{d,c}$, $c_{s,c}$ として特定の関数を仮定しておらず、特にハイライトに関する特別な幾何学的モデルやハイライトが照明光と同色であるという仮定、鏡面反射が等方性を持つという仮定をしていない。そのため点光源、面光源、無限遠に対して等しく適用することができる。
- 実世界の光源はシーンに対して光量が一定ではないが、シーンのあらゆる場所で照明光量が等しいとは仮定しておらず、スペクトルパワー分布が全体に渡って一定であるだけ仮定しているため、実世界の光源に無理なく適用することができる。

図 2.2 に二色性反射モデルの様子を示す。

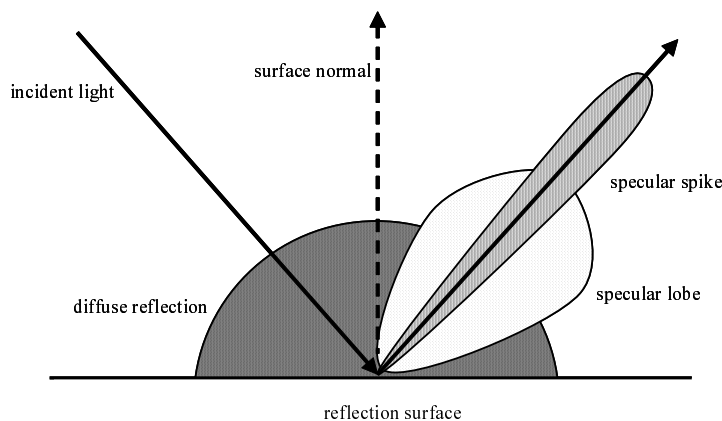


図 2.2: 二色性反射モデル

2.4 様々な反射モデル

2.4.1 Lambertian モデル

ランバートは拡散反射を式 (2.3)(2.4) のような簡潔な形で表した .

$$I_d = K_d \vec{N} \cdot \vec{L} \quad (2.3)$$

$$= K_d \cos \theta_i \quad (2.4)$$

ただし ,

- I_d : 拡散反射の放射輝度
- K_d : 拡散反射パラメータ
- \vec{N} : 物体表面の法線の単位ベクトル
- \vec{L} : 光源方向の単位ベクトル
- θ_i : 物体表面の法線と光源方向との間の角度

これより拡散反射の放射輝度は光の反射角度には依存せず , 光の入射角度に依存しているので , いずれの方向にも同じ強さで反射することがわかる .

2.4.2 Phong モデル

Phong モデルは二色性反射モデルを数式で表した最初のモデルである . この数式表現 (2.6) は非常に簡潔であるが , コンピュータグラフィックス・コンピュータビジョンにおいて広く受け入れられている .

$$I = I_d + I_s + I_a \quad (2.5)$$

$$= K_d \cos \theta_i + I_i W \cos^n \gamma_i + I_a K_a \quad (2.6)$$

ただし ,

- I : 反射光の放射輝度

- I_d : 拡散反射の放射輝度
- I_s : 鏡面反射の放射輝度
- I_a : 環境光の放射輝度
- $K_d \cos \theta_i$: ランバートの拡散反射モデル
- I_i : 入射光の放射輝度
- W : 鏡面反射率
- n : 物体表面の粗さ定数, ハイライトの特性を示す
- γ_i : 視線方向と正反射方向との間の角度
- I_a : 環境光の強さ
- K_a : 環境光定数

鏡面反射の項は, 視線方向と正反射方向のずれ角 γ_i による鏡面反射成分の減少の割合を $\cos \gamma_i$ で近似する方法である.

2.4.3 Torrance-Sparrow モデル

Torrance-Sparrow モデルは Torrance と Sparrow[16] が提案した物体表面モデルを基にした反射モデルであり, Blinn により提案された. 反射パラメータや幾何学的要因による光の減衰といったパラメータも入っており, Phong モデルよりも正確に鏡面反射を表現している [17]. また, Torrance-Sparrow モデルは粗い表面に対応しており, 滑らかな表面にも対応する Beckman-Spizzichino モデルもある. しかし Beckman-Spizzichino モデルは非常に複雑な式で表現されており, 計算量が膨大になる. そこで本研究では適当な反射モデルとして Torrance-Sparrow モデルを導入する.

Torrance-Sparrow モデルでは, まず物体表面を構成する微小面の分布を式 (2.7) で示すように正規確率分布により近似する.

$$P(\alpha) = b \exp\left(-\frac{\alpha^2}{2\sigma^2}\right) \quad (2.7)$$

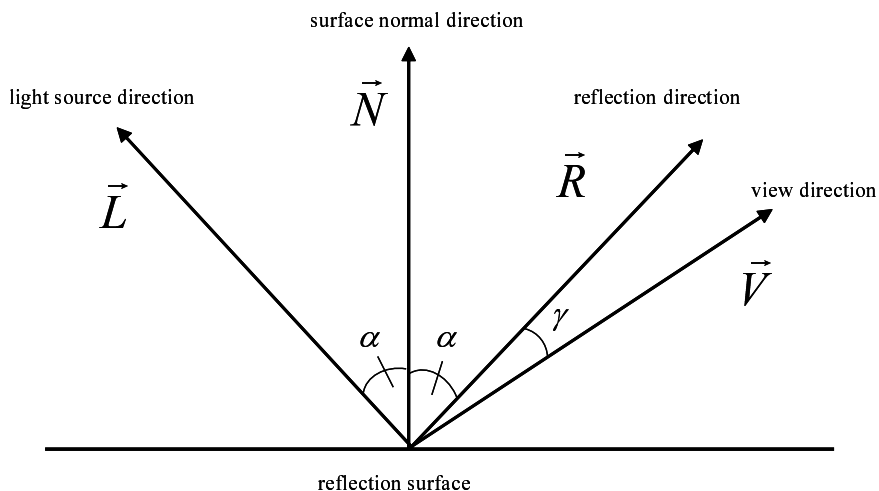


図 2.3: Phong モデル

ただし,

- $\alpha = \cos^{-1} \vec{N} \cdot \vec{H}$
- b : 定数
- σ : 物体表面の粗さを表すパラメータ
- \vec{N} : 物体表面の法線ベクトル
- \vec{H} : 視線方向と光源方向の二等分線を表すベクトル

この式は、鏡面反射は完全鏡面反射方向の周りで広がり、その分布の幅は物体表面の粗さによることを示している。

さらに幾何学的要因（微小面による遮蔽や陰影等）による光の減衰は式 2.8 により表される。

$$G = \min \left\{ 1, \frac{2(\vec{N} \cdot \vec{H})(\vec{N} \cdot \vec{V})}{(\vec{V} \cdot \vec{H})}, \frac{2(\vec{N} \cdot \vec{H})(\vec{N} \cdot \vec{L})}{(\vec{V} \cdot \vec{H})} \right\} \quad (2.8)$$

ただし,

- \vec{V} : 視線方向ベクトル

- \vec{L} :光源方向ベクトル

最後に，反射光自体の減衰は式 $F(\theta, \eta, \lambda)$ である．ただし，

- $\theta = \cos^{-1}(\vec{N} \cdot \vec{V})$:入射角
- η :屈折率
- λ :波長

以上の各要素により，Torrance-Sparrow モデルの鏡面反射成分は式 (2.9) のように記述される．

$$R(\vec{N}, \vec{V}, \vec{L}, \sigma, \eta, \lambda) = \frac{F(\vec{N}, \vec{L}, \eta, \lambda)P(\vec{N}, \vec{V}, \vec{L}, \sigma)G(\vec{N}, \vec{V}, \vec{L})}{(\vec{N} \cdot \vec{V})} \quad (2.9)$$

ここで反射パラメータ F は入射角がほぼ一定であればほとんど定数と見ることができ，さらに幾何学的要因による光の減衰 G は平面と仮定できる板などの物体においては 1 と仮定できる．

以上の鏡面反射項とランバートの拡散反射項により，Torrance-Sparrow モデルは式 (8.1) のように表現できる．

$$I = I_d + K_s \frac{\exp(-\frac{\alpha^2}{2\sigma^2})}{\cos\theta} \quad (2.10)$$

ただし，

- I_d :拡散反射の放射輝度
- K_s :鏡面反射パラメータ
- σ :物体表面の粗さを表すパラメータ
- α :視線方向ベクトルと光源方向ベクトルの間の角度
- θ :入射角

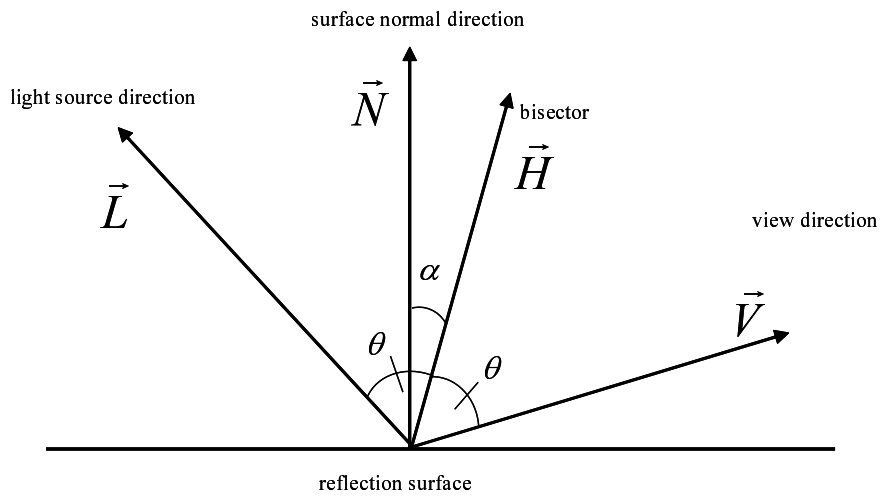


図 2.4: Torrance-Sparrow モデル

第3章

偏光

この章では偏光について説明する．偏光はリモートセンシングの研究分野で用いられてきた．Wolff ら [18] は直線偏光を解析し，ハイライト成分を除去するアルゴリズムを提案した．また Boulton ら [19] は偏光の性質に基づいてシーンのエッジを検出した．Schechner ら [20] は透過シートによって画像を分類する方法を提案している．

物体表面の反射モデルは鏡面反射成分と拡散反射成分の線形和として表現されるので，それぞれの反射率は独立して求められる．この二種類の反射率を解析する際，鏡面反射成分の画像と拡散反射成分の画像が必要となる．ここで，一般的に撮像した画像はこの両成分を含むので，通常の画像からそれぞれの反射成分を分離しなければならない．

偏光を用いると，拡散反射成分と鏡面反射成分を容易に分離することができる．以下では偏光の性質と，偏光による反射成分分離の過程を述べる．

3.1 偏光

光は電磁波であり横波の一種である．自然光は通常，振幅の大きさも方向もばらばらであるが，振幅の方向がある規則に従う光を偏光という．偏光には，振幅の方向が一定面内にある「直線偏光」と，振幅の方向が時間の経過で円を描く「円偏光」などがある．

3.1.1 直線偏光

直線偏光とは，電界の振動方向が特定の面内にある偏光のことである．自然光をガラス板で反射させたり，ニコルプリズム，ポーラロイド板（偏光サングラス）などを通すと，電界ベクトルの方向がある特定の方向に偏る．

図 3.1(a) に振動方向が水平（青），図 3.1(b) に垂直（赤）方向にある直線偏光を示す．振動面を進行方向から見たときに直線上を振動していることがわかる．

次にこの二つの直線偏光の重ね合わせを考えてみる．図 3.2(c) が上の二つの直線偏光を重ねて描いた物で，図 3.2(d) は進行方向の軸の同じ位置での振幅をベクトル的に足し合わせたものである．

新たにできた緑色の振動も直線偏光で，その振動方向は垂直・水平軸から 45 度

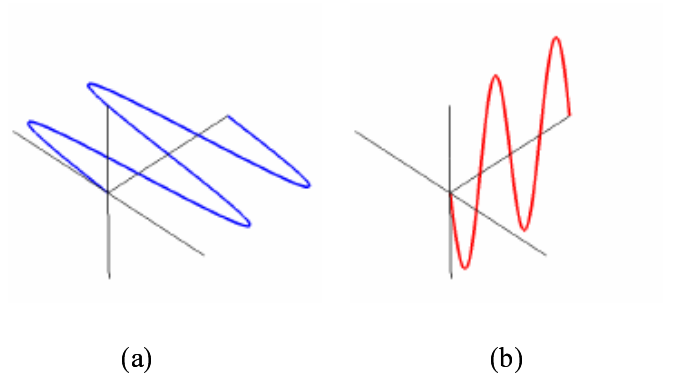


図 3.1: 直線偏光

傾いている．図 3.2(d) では，足し合わせる二つの波の位相が同じである．

直線偏光板は，自然光から振幅の向きが一定面内にある光だけを透過させるフィルターである．Takahashi らは二枚の直線偏光板を用いて光を遮断する方法 [13] を用いている．直線偏光板は式 (5.2) のように光の透過度 T が定義される．

$$T(\theta) = k_1 \cos^2 \theta + k_2 \sin^2 \theta \quad (3.1)$$

ただし， θ は直線偏光板への入射光の振動面と直線偏光板の透過軸との間の角度である．理想的な直線偏光板ならば， k_1 は 1， k_2 は 0 となる．しかし実際には k_1 は 1 より少し小さい値， k_2 は 0 より少し大きい値となる．

次に二枚の直線偏光板を用いての光の透過について考える．ここでは θ を二枚の直線偏光板同士の透過軸の間の角度として，式 (3.2) のように透過度 T_{pair} が定義される．

$$T_{pair}(\theta) = k_1 k_2 \sin^2 \theta + \frac{1}{2}(k_1^2 + k_2^2) \cos^2 \theta \quad (3.2)$$

ここで， $\theta = 90^\circ$ と $\theta = 0^\circ$ のとき，それぞれ式 (3.3) と式 (3.4) のようになる．

$$H_{90} = T_{pair}(90^\circ) = k_1 k_2 \quad (3.3)$$

$$H_0 = T_{pair}(0^\circ) = \frac{1}{2}(k_1^2 + k_2^2) \quad (3.4)$$

式 (3.2) に式 (3.3) と式 (3.4) を代入すると以下のようになる．

$$T_{pair}(\theta) = H_{90} \sin^2 \theta + H_0 \cos^2 \theta \quad (3.5)$$

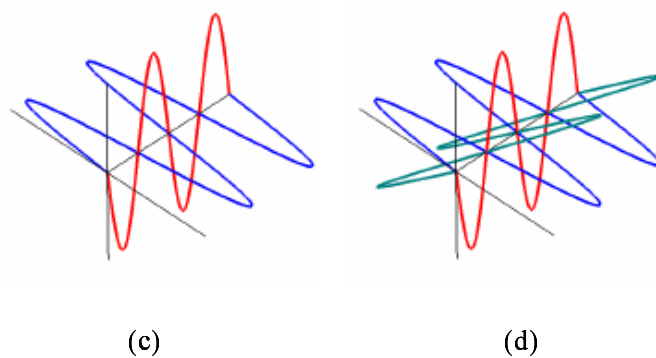


図 3.2: 直線偏光の重ね合わせ

$$= H_{90} + (H_0 - H_{90}) \cos^2 \theta \quad (3.6)$$

H_{90} は close transmissivity , H_0 は open transmissivity と呼ばれており , open transmissivity は close transmissivity に対して非常に大きな値である . よって二枚の直線偏光板の間の透過度はそれぞれの透過軸同士の間角度の余弦関数で近似される . θ が 90° のときと , 0° のときの様子を図 3.4 に示す .

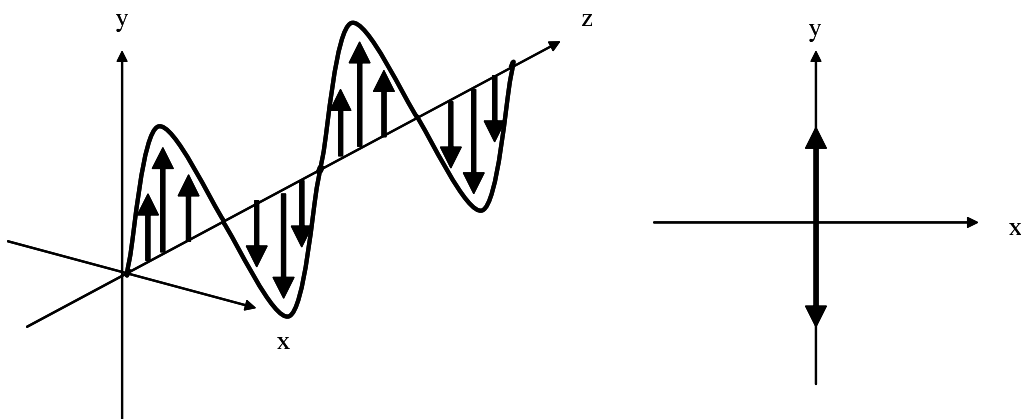


図 3.3: 直線偏光の性質

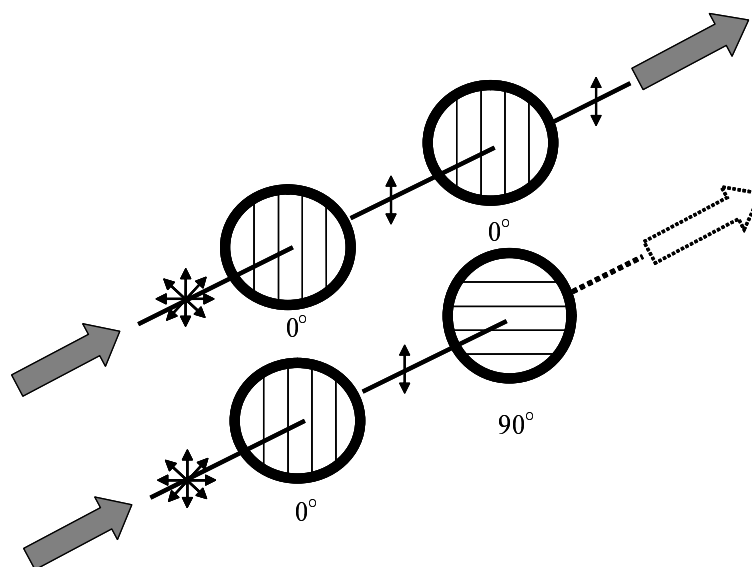


図 3.4: 直線偏光板による光の遮断

3.1.2 円偏光

円偏光とは、電界ベクトルが光の進行方向を軸として時間とともに回転するよような偏光のことである。

ここでもう一度二つの直線偏光の重ね合わせを考える。図 3.2(d) は、二つの波の位相が同じ場合であった。しかし二つの波の位相が異なると足し合わせた結果はもはや直線にはならない。図 3.5(e) では赤い波の位相を青い波に対して 90° 変えてある。

この二つの波を足し合わせると図 3.5(e) の緑の線のように振動は螺旋状になる。この螺旋を進行方向から見ると円となるので円偏光というわけである。円偏光には右円偏光と左円偏光の二種類がある。進む方向に向かって時計回りの円偏光を左円偏光、進む方向に向かって反時計回りの偏光のことを右円偏光という。右円偏光と左円偏光では螺旋の掌性が異なっている（図 3.5(f)）。

足し合わせる 2 つの波の位相が 90° 以外の場合（や位相差は 90° だけれど振幅が異なっている場合）は、振動の軌跡は進行方向からみて楕円となるので楕円偏

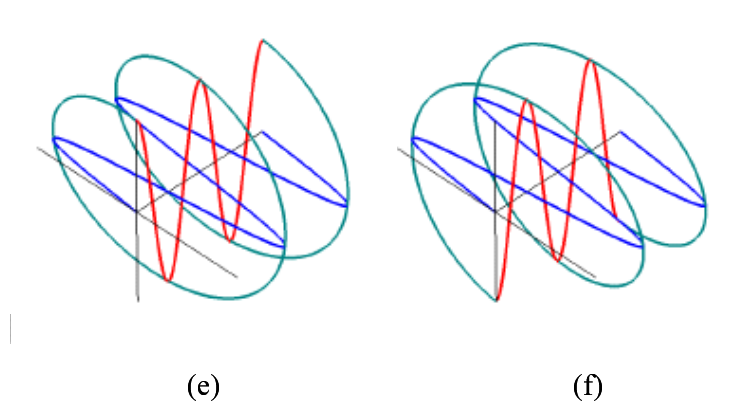


図 3.5: 円偏光

光と呼ばれている．直線偏光と円偏光は楕円偏光の特別な場合である．

円偏光板は，自然光から振幅の向きが時間とともに右（または左）に回転する光だけを透過させるフィルターである．直線偏光板と同様，円偏光板も二枚を組み合わせることにより光を遮断できる．3.1.1 で述べたように直線偏光板の場合は二枚の偏光板の透過軸の間の角度を適切に設定する必要がある．円偏光板には右円偏光板と左円偏光板の二種類があり，自然光が透過するとそれぞれ反時計回りの偏光（右円偏光）と時計回りの偏光（左円偏光）となる．そして左円偏光板は右円偏光を遮断し，右円偏光板は左円偏光を遮断する性質があるので図 3.7 のように単純に左右それぞれの円偏光板を組み合わせるだけで光を遮断することができる．直線偏光板は角度を適切に設定してやる誤差を含むので，円偏光板のほうがロバストであると言える．

3.2 偏光による反射成分分離

ここでは偏光による反射成分分離操作について説明する．図 3.8 は二枚の円偏光板を用いた反射成分分離のメカニズムを表している．3.1.2 で説明した円偏光板による光の遮断の原理が利用されている．

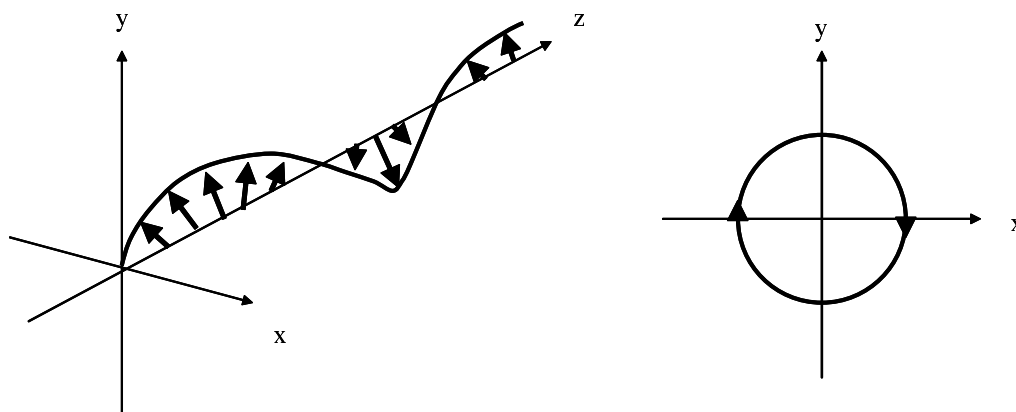


図 3.6: 円偏光の性質

光源と物体を撮像するカメラの前にそれぞれ逆回転の円偏光板を設置する．まず光源から放射された光は円偏光板を透過して円偏光となり物体表面に入射する．物体表面で直接反射する円偏光はその性質を保ったまま反射するので，鏡面反射成分は円偏光である．一方，内部に透過した円偏光は内部拡散を経て円偏光性は失われるので，拡散反射成分は非偏光な光である．ここでカメラの前の円偏光板に両反射成分が入射すると，円偏光の鏡面反射成分は遮断され，非偏光な拡散反射成分は円偏光となり透過する．そしてカメラには拡散反射成分のみが到達する．つまり鏡面反射成分は打ち消されたことになる．よってこのメカニズムにより拡散反射画像を得ることができる．

ただし，偏光板を通して物体を撮像すると位置と輝度値がずれる．偏光板有り画像の位置を偏光板無し画像の位置に合わせるためには，両者のエッジを計算してアライメントすればよい．また輝度値が減衰する問題については，基準白色を偏光板有り・無しで撮像することにより調整する．

明るさの比を式で表してみると，円偏光板を設置せずにカメラによって観測された輝度を I_{off} ，円偏光板を設置してカメラによって観測された輝度を I_{on} ，拡散反射成分の輝度を I_d ，鏡面反射成分の輝度を I_s として， I_{on} と I_{off} は

$$I_{on} = I_d \quad (3.7)$$

$$I_{off} = I_d + I_s \quad (3.8)$$

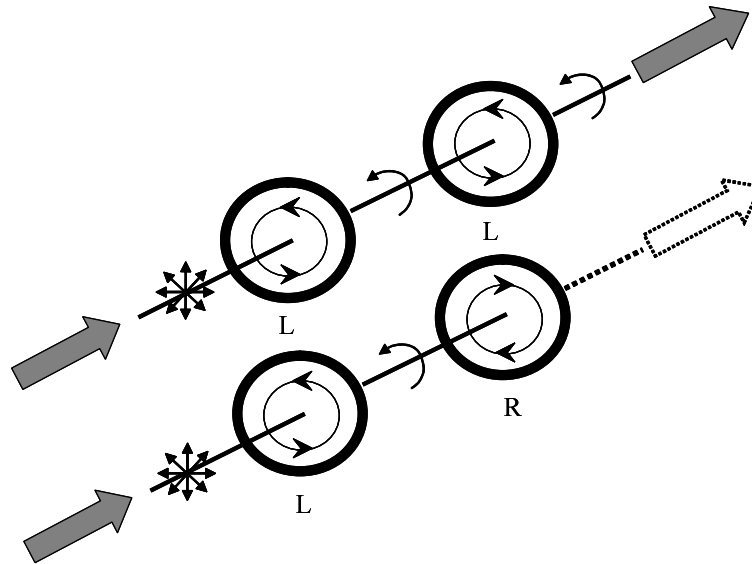


図 3.7: 円偏光板による光の遮断

つまり，拡散反射性分 I_d と鏡面反射性分 I_s は以下のようになる．

$$I_d = I_{on} \quad (3.9)$$

$$I_s = I_{off} - I_{on} \quad (3.10)$$

同様のことが直線偏光板によっても可能である．しかし視点が 90° に近い時，拡散反射成分は部分的に直線偏光となることが知られているので，直線偏光板により拡散反射成分はその分だけ減衰しており，代わりにその差分が鏡面反射成分として計算されることになるので誤差を含むことになる．円偏光板による反射成分分離メカニズムはこの問題も解決してくれる．円偏光板は逆方向の円偏光のみを遮断するので，部分的に直線偏光された拡散反射成分は完全に透過する．

以上より，直線偏光板ではなく円偏光板を用いることにより，偏光板の透過軸に影響されないこと，視点が 90° に近い時に起こる拡散反射成分の部分的な直線偏光化の影響を受けないことの二つの利点がある．図 3.9 はこの操作により分離された拡散反射成分，鏡面反射成分を表す．なお，ここでは鏡面反射成分画像の明るさを 4 倍にしてある．

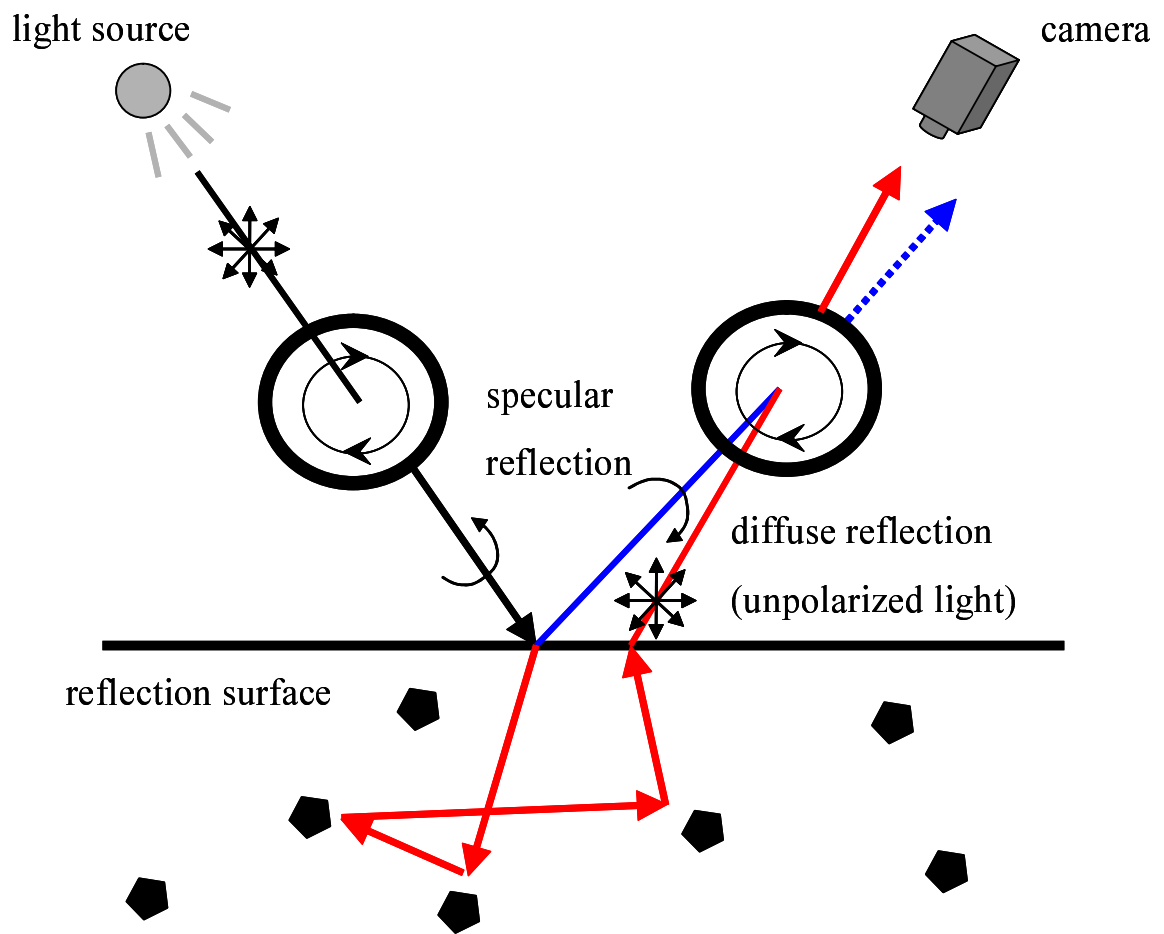


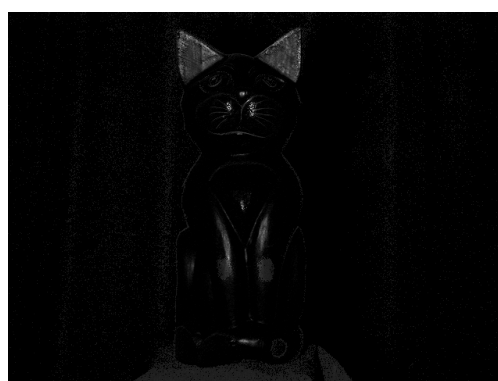
図 3.8: 円偏光板による反射性分の分離



original image



diffuse image



specular image

図 3.9: 拡散反射成分と鏡面反射成分 (明るさは実際の 4 倍)

第4章

幾何情報

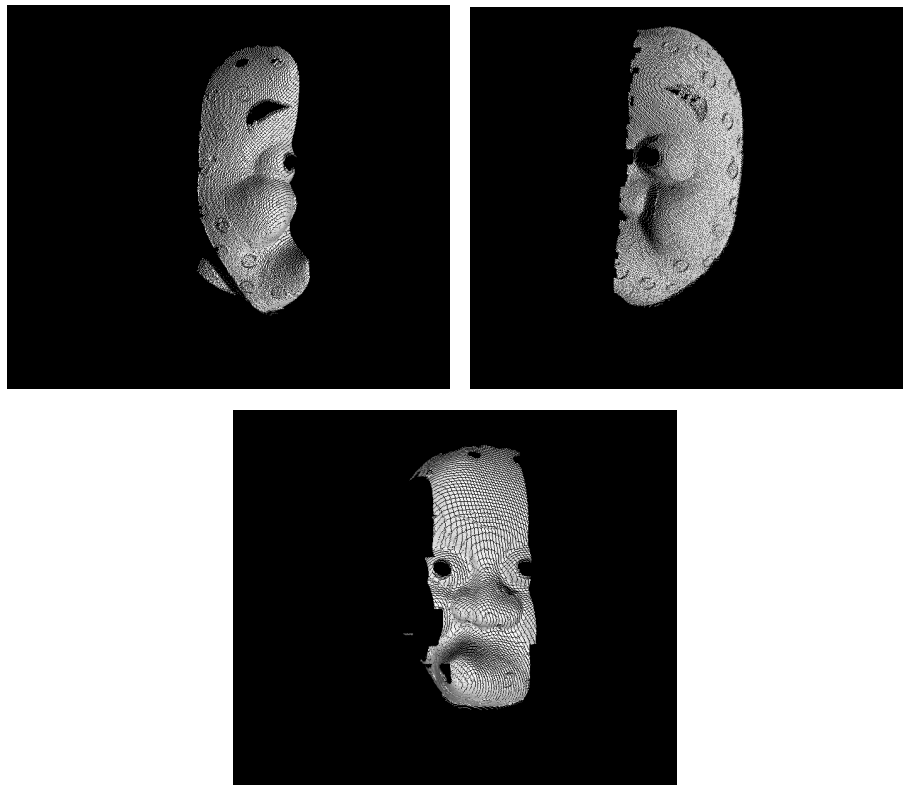


図 4.1: 距離画像

この章では幾何データについて説明する．物体の幾何モデルはスキャニング，アライメント，マージングの三つの行程からなる．

4.1 距離画像取得

対象物体の幾何モデルを形成するために，まずはじめにレーザーレンジファインダにより距離画像を得る．距離画像のピクセルは物体表面の (X, Y, Z) 座標を表す．三次元環境全ての位置が得られるように複数の観測点から撮影を行う必要がある．こうして対象物体の全周囲を覆うのに十分な距離画像列が用意できる．距離画像の例を図 4.1 に示す．

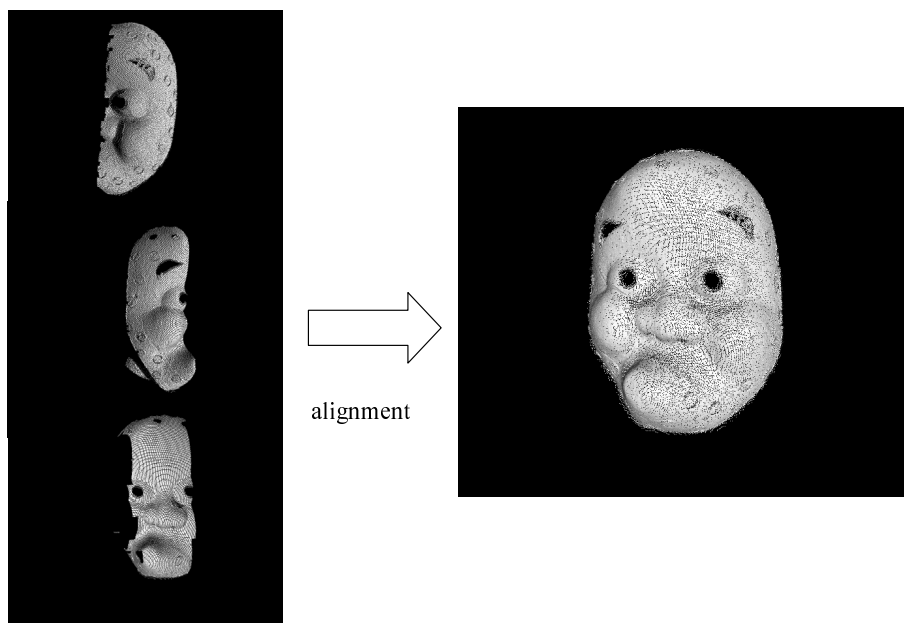


図 4.2: アライメント

4.2 アライメント

距離画像はそれぞれの観測点のレンジセンサーの座標系上のデータとして得られるので、それらの距離画像を一つの座標系上に合わせる必要がある。この操作をアライメントという。距離画像のアライメント手法の最も基本的なものとして、Iterative Closest Point(ICP) アルゴリズムがある [21]。これは距離画像間の位置を合わせるために、それぞれの距離画像の対応点間の距離の総和を誤差関数とし、繰り返し計算により最小化するアルゴリズムである。本研究では大石らによるアライメント手法を用いている [22]。アライメントの様子を図 4.2 に示す。

4.3 マージング

全ての距離画像を一つの座標系にアライメントした後、距離画像のマージングを行う。距離画像をアライメントしただけでは、ただ単に複数の距離画像が重なって置かれているだけであるため、複数の距離画像をつなぎ合わせ、無駄な重なりを省

き，一つのデータに統合しなければならない．この操作をマーキングという．本研究では Wheeler らによる手法を用いている [23]．まず全ての距離画像をボリューム表現を用いて表現する．空間を格子状に分割し，各ボクセルには陰表面をサンプリングして格納する．すなわち，ボクセルの中心 x から最も近い物体の表面までの符号付距離 $f(x)$ を計算する． $f(x)$ が正の場合，ボクセルは表面の外側にあるということを表し，負の場合は内側にあることを表す．一続きの距離画像の三角メッシュを構成するために Marching Cube Algorithm[24] を用いている．距離画像のマーキングとは，複数の距離画像の陰表面関数が与えられたときに，この関数が 0 となる同位面を抽出することである．

4.4 表面法線

マーキングされた物体幾何モデル上の三角メッシュから，物体の表面法線の近似が求まる．この法線は三角パッチ一つにつき一つ求められ，三角メッシュの各辺ベクトルの外積により計算される．もし物体表面が滑らかで，かつ大きな曲率で曲がっているところがあれば，現実の表面法線をうまく近似できるが，現実の物体はそういった形ばかりとも限らない．物体の反射率を求めるためには三つの方向ベクトル：(1) 視線方向，(2) 光源方向，(3) 表面法線が必要である．視線方向，光源方向は既知という条件だから，現実の物体において，精度の高いの反射率を得るためには表面法線をより正確に求めなければならない．

より正確な法線を求めるために，主成分分析 Principal Component Analysis(PCA) を用いる．PCA とは多変量データを統合し，新たな統合指標を作り出すための手法である．物体のある点の法線を求める場合，法線を求めたい点と，その近傍点 N 個の点郡を PCA する．点 $[\bar{X}, \bar{Y}, \bar{Z}]$ を中心とし，その近傍点の座標を $[X_l, Y_l, Z_l]$ とすると， n 次の共分散行列は式 4.1 のように表現される．

$$C = \sum_{l=1}^n \begin{bmatrix} (X_l - \bar{X}) \\ (Y_l - \bar{Y}) \\ (Z_l - \bar{Z}) \end{bmatrix} \begin{bmatrix} (X_l - \bar{X}) & (Y_l - \bar{Y}) & (Z_l - \bar{Z}) \end{bmatrix} \quad (4.1)$$

この共分散行列の固有値問題を解き，最小固有値に対応する固有ベクトルが法線となる (図 4.3)．全ての点についてこの方法で法線を求める．ここで物体表面がたくさん小さな面で構成されている場合，それぞれの面で法線が異なってくる．そこで，幾何モデル上の頂点ごとに PCA による表面法線を求めた後，その頂点の張る

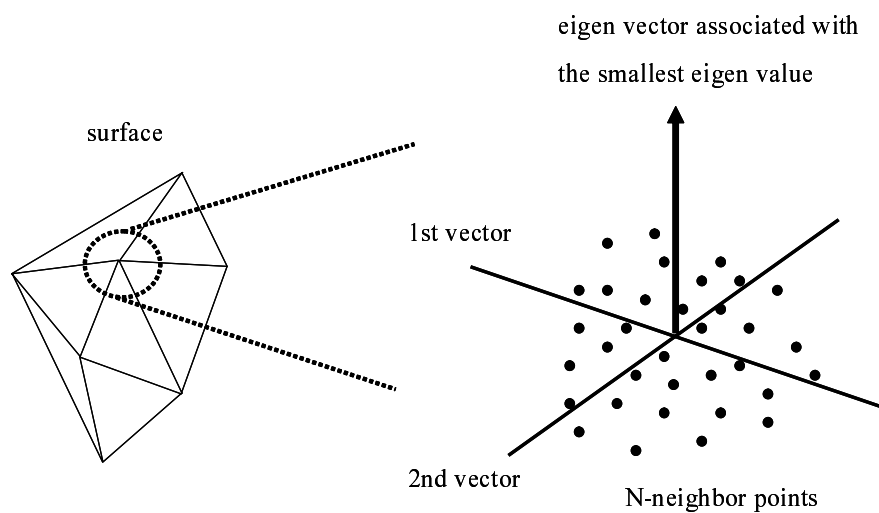


図 4.3: PCA による法線計算

三角メッシュ内では三頂点の法線を補間する必要がある．特に鏡面反射は法線の方に大きな影響を受ける．この補間を行うと，メッシュ間で法線が急に変化せず滑らかな法線が求まる．

第5章

カメラキャリブレーション

物体のある点の反射パラメータを推定するためには、法線情報と色情報が必要なので二次元の光学データと三次元の幾何データの位置を合わせる必要がある。本章では二次元の光学データと三次元の幾何モデルの対応をとるカメラキャリブレーションについて述べる。画像と幾何モデルマッピングを正確におこなうためには、二次元の画像を三次元空間上でどの位置からどのように撮影したものかというカメラパラメータを求める必要がある。また、本研究では対象物体の全周囲からの見えを生成するために、対象物体を回転テーブルに載せて、一定角度ごとに様々な方向から画像を取得する。これはある回転軸を中心とした回転であるので、カメラキャリブレーションをする際にこのことを考慮して精度を上げる。

5.1 カメラキャリブレーション

5.1.1 カメラパラメータ

カメラパラメータとは位置、姿勢、画角などのカメラに関するデータが含まれている。このカメラパラメータは内部パラメータ、外部パラメータで構成されており、以下のように内部パラメータが 5 個、外部パラメータが 6 個の計 11 個のパラメータよりなる。

- 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)
 - x 軸基準の回転角 (radian)
 - 平行移動ベクトル t の x 成分 : t_x
 - 平行移動ベクトル t の y 成分 : t_y
 - 平行移動ベクトル t の z 成分 : t_z
- 内部パラメータ
 - 焦点距離
 - 画像中心 C_x

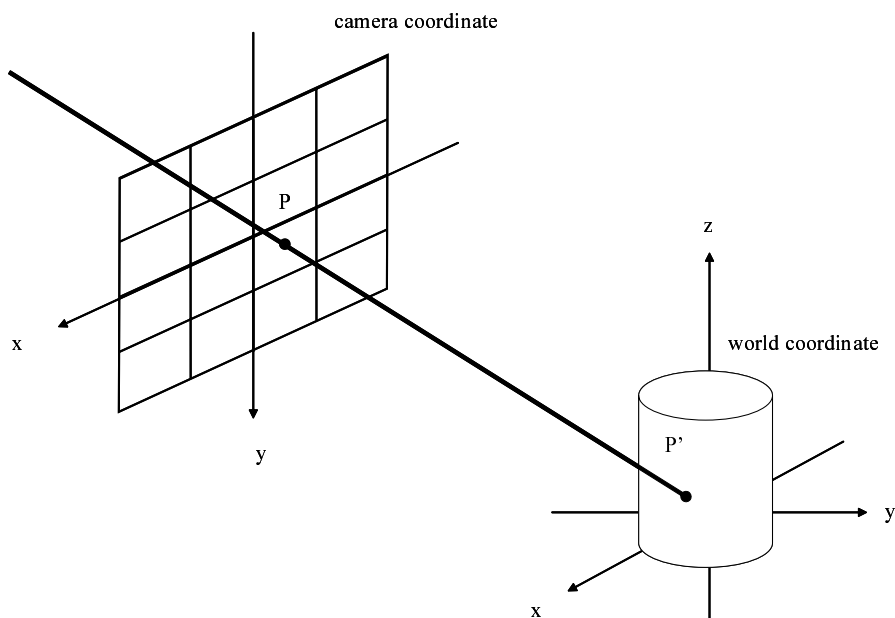


図 5.1: 世界座標系とカメラ座標系

- 画像中心 C_y
- レンズ歪み係数
- アスペクト比

内部パラメータとはカメラによる撮像ごとに設定するパラメータである．また外部パラメータとは対象物体の変換に用いるパラメータである．

5.1.2 世界座標系とカメラ座標系

カメラにおいて，三次元空間全体の中で三次元物体が定義されている座標と，三次元物体ごとに三次元空間が定義されている座標がある．世界座標とは三次元空間の位置を定義する座標である．一方，カメラ座標とは世界座標に対し，「視点」を置く位置を原点とした座標である．このとき，視線は z 軸方向であり，下方向が y 軸方向，右方向が x 軸方向となる．図 5.1 にこの関係を示す．オブジェクトと空間の間で定義する座標が異なるため，座標変換を行う必要がある．

5.1.3 同次座標

同次座標とは三次元上の点を一次元増やして表現した座標である。三次元上の点 $P(x, y, z)$ は同次座標では $P(x, y, z, 1)$ となる。また、同次座標で点 $P(x, y, z, w)$ は普通の座標系では $P(\frac{x}{w}, \frac{y}{w}, \frac{z}{w})$ となる点を表しているものとする。この w は倍率（重み）の役割をしている。 $w = 0$ なら無限遠点を表現することになる。この同次座標系を用いると、5.1.4 で述べる「平行移動」、「回転移動」、「拡大・縮小」のようなアフィン変換を一まとめに表現できる。

5.1.4 三次元アフィン変換

図形処理の最も基本的な処理には、「平行移動」、「回転移動」、「拡大・縮小」などがあり、この移動のための変換がアフィン変換である。ここで、点 $P(x, y, z)$ を点 $P'(x', y', z')$ に移動させるとき、三次元空間上のアフィン変換は回転行列を R 、平行移動ベクトルを \mathbf{t} とすると、回転移動が式 (5.1)、平行移動が式 (5.2) と表現できる。

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \mathbf{R} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.1)$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = [\mathbf{0} \mid \mathbf{t}] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.2)$$

ただし、

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, [\mathbf{0} \mid \mathbf{t}] = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.3)$$

この二つを一つにまとめると式 (5.4) となる。

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = [\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.4)$$

ただし,

$$[\mathbf{R} \mid \mathbf{t}] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.5)$$

5.1.5 透視変換

図 5.2 にピンホールカメラに基づいた透視変換モデルを示す。カメラを理想化したこの透視変換モデルは、レンズの中心にピンホールカメラが開いたものであり、視線は一個の点を通る一本の直線として定義される。ガラスレンズを用いた一般の結像系も、歪曲収差が無視できるほど小さければ、この簡易モデルで表現できる。実際のカメラは対象 レンズ 結像面と配置されるが、これでは像が逆転してわかりにくいため、仮想的に結像面をレンズの前におき、対象 結像面 レンズと配置している。図中の F はレンズの中心、 f はレンズの焦点距離、 I は結像面である。結像面をカメラに固定した座標系の基準に考えているため、 I の中心を座標軸の原点にしている。空間中にある一点 $P'(x', y', z')$ 、すなわち測定点 $P(x, y, z)$ に向かってる視線と結像面との交点は式 (5.6) で与えられる。

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -f \end{bmatrix} + \alpha \begin{bmatrix} x \\ y \\ f+z \end{bmatrix} \quad \text{かつ } z = 0 \quad (5.6)$$

また、式 (5.7) とも表現できる。

$$x' = fx/(f+z) \quad (5.7)$$

$$y' = fy/(f+z) \quad (5.8)$$

$$z' = 0 \quad (5.9)$$

三次元の点 (x, y, z) を、式 (5.10) のように w_h を媒介とする四次元の点 (x_h, y_h, z_h, w_h) で表現する。

$$x = x_h/x_h \quad (5.10)$$

$$y = y_h/y_h \quad (5.11)$$

$$z = z_h/z_h \quad (5.12)$$

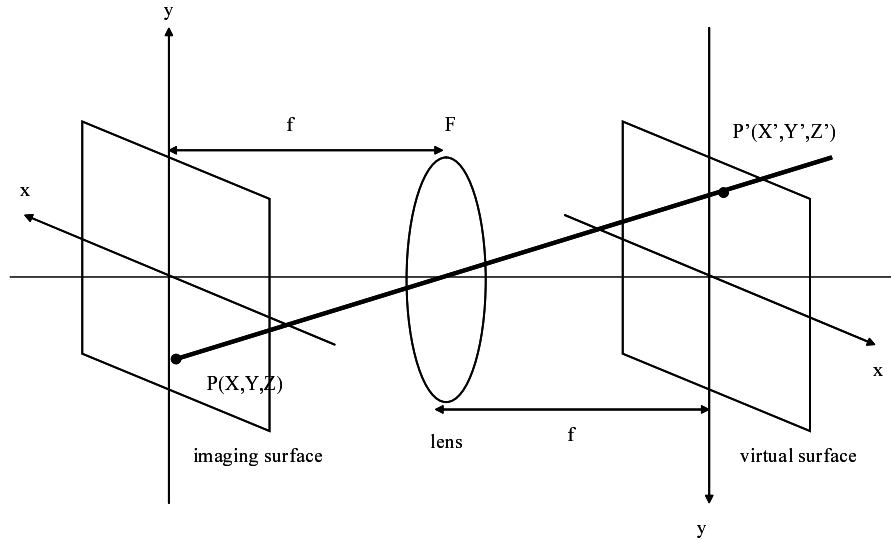


図 5.2: 透視変換

この同次座標により，透視変換は式 (5.13) の 4×4 の行列演算で記述することができる．

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w_h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.13)$$

ここで $[x, y, z, 1]^T$ と $[x_h, y_h, z_h, w_h]^T$ は同次座標系表現による点 P と点 P' の座標である． 図5.2の透視変換は点 P と点 P' が共にカメラに固定した座標系で表現されているときに適用できる．世界座標系での点 P からカメラ座標系での点 P' への変換は式 (5.14) となる．

$$\begin{bmatrix} x'_h \\ y'_h \\ z'_h \\ w'_h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \quad (5.14)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} [\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.15)$$

$$= \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ 0 & 0 & 0 & 0 \\ r_{31}/f & r_{32}/f & r_{33}/f & t_z/f + 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.16)$$

第6章

反射パラメータ推定

本章では反射パラメータの推定について述べる．第二章で，物体の表面反射における反射モデルは拡散反射成分と鏡面反射成分の線形和で表現できること説明した．この反射モデルの式のそれぞれの項の反射パラメータを推定することにより任意の角度の拡散反射・鏡面反射を再現することができる．任意の視点からのレンダリングを行うためには，対象物体の全ての点における反射パラメータを推定する必要がある．そのためには対象物体を回転テーブルに載せて回転させながら撮像した，一定角度ごとの画像を用いる．

Torrance-Sparrow モデルによると，それぞれの項は各反射成分の輝度（光学情報）を表し，それは各反射パラメータと，視線方向，法線方向等（幾何情報）より成り立っている．つまり反射パラメータを推定するためには対応の取れた光学情報と幾何情報が必要となる．まず第三章で説明した円偏光を利用した分離操作により，原画像から拡散反射画像と鏡面反射画像に分離する．次に第五章で述べた方法により，各成分の画像と幾何モデルの対応を取る．そうして得られた対応する光学情報と幾何情報により Torrance-Sparrow モデルをフィッティングしてやることにより反射パラメータを得る．

以下では各成分ごとに述べる．

6.1 拡散反射パラメータ推定

第二章で述べたように Torrance-Sparrow モデルの拡散反射成分の項は式 (6.1)(6.2) のように表すことができる．

$$I_{d,c} = K_{d,c} \vec{N} \cdot \vec{L} \quad (6.1)$$

$$= K_{d,c} \cos\theta_i \quad (c = R, G, B) \quad (6.2)$$

ただし，

- I_d : 拡散反射の放射輝度
- K_d : 拡散反射係数
- \vec{N} : 物体表面の法線の単位ベクトル
- \vec{L} : 光源方向の単位ベクトル

- θ_i : 物体表面の法線と光源方向との間の角度

これより推定すべき拡散反射パラメータは $K_{d,R}, K_{d,G}, K_{d,B}$ の三つである．これらを分離された拡散反射成分画像と，幾何モデルから計算される法線方向，既知の光源方向，視線方向を用いて求める．

反射パラメータは全ての点について求める．ある点に着目すると，複数枚の拡散反射成分画像よりこの点の様々な輝度値が求まるので，その全てについての式 6.2 を最小二乗法でフィッティングする．その後各画像の同一点ごとに法線方向と視線方向の間の角度の余弦関数により重み付けをして推定する．

6.2 鏡面反射パラメータ推定

一方鏡面反射成分の項は式 (6.3) のように表すことができる．

$$I_{s,c} = K_{s,c} \frac{\exp\left(-\frac{\alpha^2}{2\sigma^2}\right)}{\cos\theta} \quad (c = R, G, B) \quad (6.3)$$

ただし，

- $I_{s,c}$: 鏡面反射輝度
- $K_{s,c}$: 鏡面反射パラメータ
- σ : 物体表面の粗さを表すパラメータ
- α : 視線方向ベクトルと光源方向ベクトルの間の角度
- θ : 入射角

これより推定すべき鏡面反射パラメータは $K_{s,R}, K_{s,G}, K_{s,B}, \Sigma$ の四つである．しかし，鏡面反射パラメータの推定は，拡散反射パラメータの推定のようにはいかない．拡散反射は光が当たった全ての範囲で観測されるが，鏡面反射は限られた視点から，狭い範囲でのみ観測されるというのが原因である．だから，鏡面反射パラメータの推定にはサンプルする画素を注意深く選定する必要がある．本研究では線形最小二乗法と Levenberg-Marquardt 法を組み合わせた佐藤らによる手法を用いている [12]．

第7章

ウェーブレット

7.1 基礎理論

ウェーブレット変換とは、信号を空間的な周波数成分に変換する直行変換の一種であり、信号を有限の幅を持つウェーブレットの重ね合わせで表現するものである [25] .

7.1.1 連続ウェーブレット変換

実数 \mathbf{R} 上で定義された関数 f のうち、次の条件を満たすものを二乗可積分関数と呼ぶ .

$$\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty \quad (7.1)$$

つまり、 $f(t)$ は無限遠で 0 に収束する必要がある . 二乗可積分関数の空間を $L^2(\mathbf{R})$ で表す . フーリエ変換の積分核 $e^{-i\omega t} = \cos -i \sin$ は $L^2(\mathbf{R})$ には属さない . これに対し、ウェーブレット変換は $L^2(\mathbf{R})$ に属する関数 $\psi(t)$ を基底として、 $L^2(\mathbf{R})$ に属する任意の信号を表現する手段である . ここで図 7.1 に示すような、平均値が 0 で、原点 $t = 0$ の周りに局在する関数 ψ を考える . この関数 ψ をアナライジングウェーブレット analyzing wavelet または、単にウェーブレット wavelet と呼ぶ .

この $\psi(t)$ を t 軸上でシフト、あるいは拡大縮小して図 7.1 のような基底 $\psi_{a,b}(t)$ を生成する .

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \quad (7.2)$$

ここで、 b はシフト、 $a > 0$ は拡大縮小のパラメータであり、スケールと呼ばれる . この $\psi_{a,b}(t)$ と信号 $f(t)$ との内積がウェーブレット変換 wavelet transformation である .

$$(W_{\psi}f)(b, a) = \frac{1}{\sqrt{a}} \int_{\mathbf{R}} f(t) \overline{\psi \left(\frac{t-b}{a} \right)} dt \quad (7.3)$$

ここで、 $\overline{\psi(\cdot)}$ は $\psi(\cdot)$ の複素共役である . 時間軸に局在した与関数 $f(t)$ によく似た時間軸に局在したウェーブレットを用いれば、時間軸上に局在する波形を適切に検出できることがウェーブレット変換の特徴である .

次にウェーブレット逆変換を考える . まず、 C_{ψ} を以下のように定義する .

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega \quad (7.4)$$

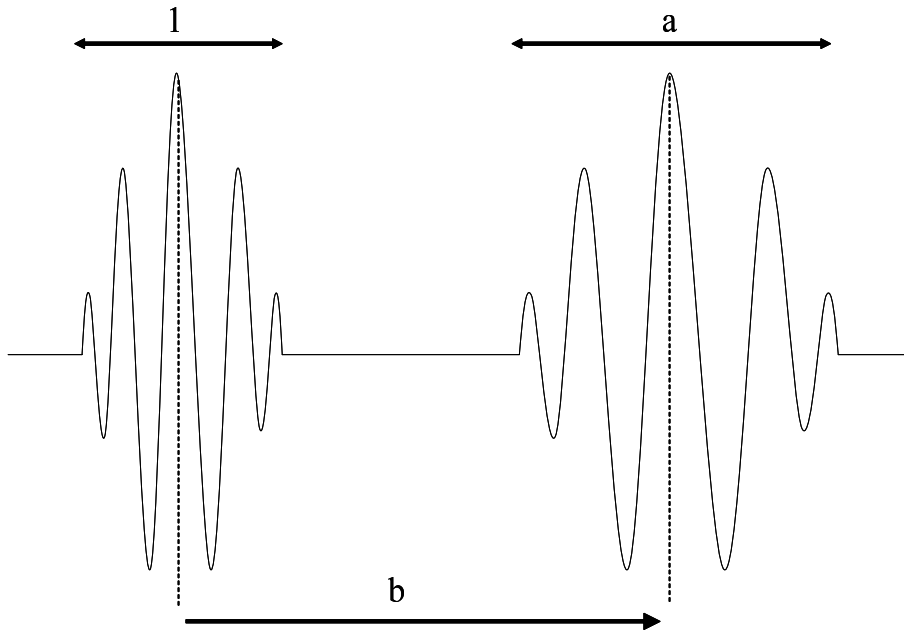


図 7.1: ウェーブレット

ここで, $\hat{\psi}(\omega)$ は $\psi(\omega)$ のフーリエ変換であり,

$$\hat{\psi}(\omega) = \int_{-\infty}^{\infty} \psi(\omega) e^{-i\omega t} dt \quad (7.5)$$

である.

ウェーブレットは逆変換の存在を確保するために必要な, 次の条件を満たすことが知られている.

$$\int_0^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = \int_0^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = \frac{1}{2} C_{\psi} < \infty \quad (7.6)$$

このとき, $\hat{\psi}(\omega)$ は有界連続であるから, $C_{\psi} < \infty$ が成立するためには $\hat{\psi}(0) = 0$ であることが必要である. つまり $\psi(t)$ の直流成分が 0, すなわち平均値が 0 でなければならない.

この条件を満たすとき, 次の逆変換が成立する.

$$f(t) = \frac{2}{C_{\psi}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} (W_{\psi} f)(b, a) \psi_{a,b}(t) db \right] \frac{da}{a^2} \quad (7.7)$$

つまり，基底 $\psi_{a,b}(t)$ により信号 $f(t)$ を表現することができる．しかしこの表現では，信号の再構成のために a, b による無限積分が必要であり現実的でない．数値計算により信号のウェーブレット変換及び逆変換を行うときは 7.1.2 の離散ウェーブレット変換を使う．

7.1.2 離散ウェーブレット変換

離散ウェーブレットとは，連続ウェーブレットを二進分割して得られるウェーブレットである．式 (7.3) のウェーブレット変換の離散化にあたり二進分割をとり， $a = 2^j, b = 2^j k$ とすると，式 (7.2) の連続ウェーブレットは次のように離散化される．

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad (7.8)$$

この離散ウェーブレット表現をドベシィウェーブレット daubechies wavelet という．ドベシィウェーブレットは直交ウェーブレットとなるよう，適当に ψ を選んでいる．ドベシィのウェーブレット及びそのスケーリング関数は形が複雑で，既知の関数では表現できない．ウェーブレット数列の長さの半分を N とし， $N = 20$ のときのスケーリング関数とウェーブレットを図 7.2 に示す．

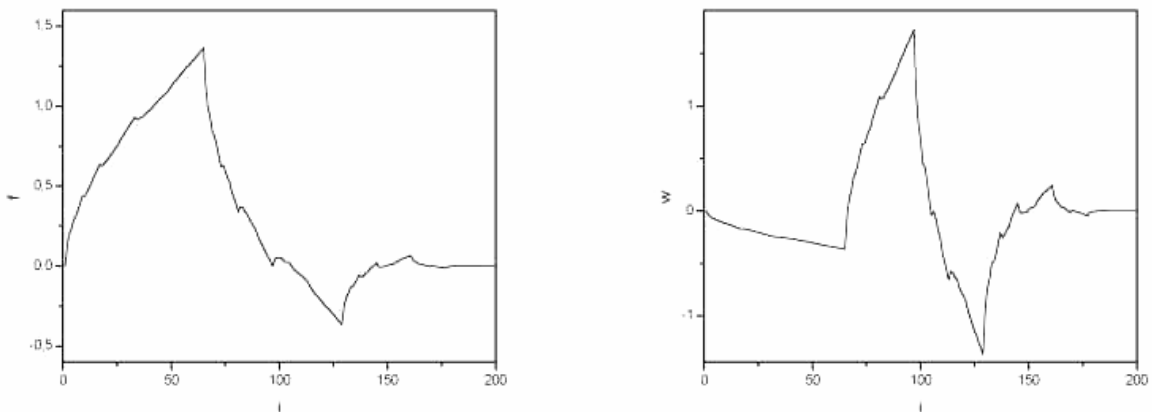


図 7.2: ドベシィウェーブレット

このほかにもウェーブレットには様々な形があり，最もシンプルなハールウェーブレット Haar wavelet (図 7.3) や，ガボールウェーブレット Gabor wavelet (図 7.4) 等があるが，本研究では画像再構成で他のウェーブレットとより優れているドベシウェーブレットを用いている．

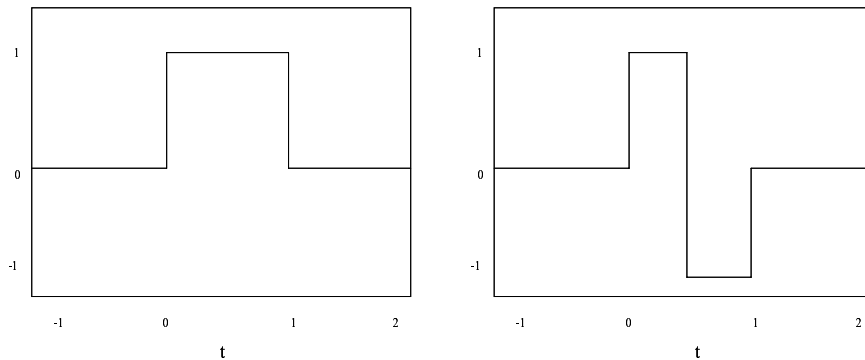


図 7.3: ハールウェーブレット

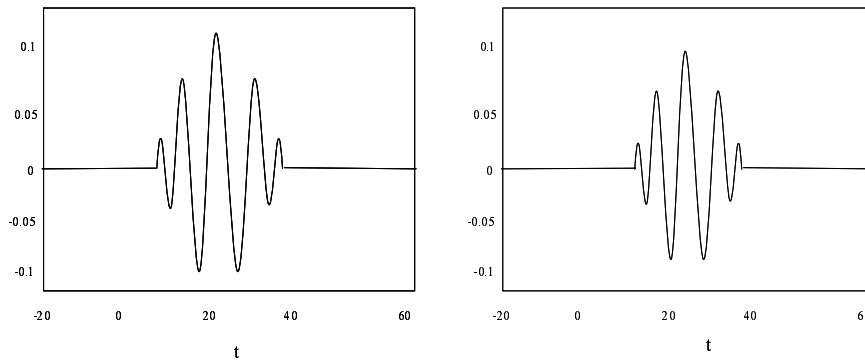


図 7.4: ガボールウェーブレット

直交条件を満たすとき，信号 $f(t)$ は次のようにウェーブレットを基底とした級数に展開される．

$$f(t) = \sum_j \sum_k w_k^{(j)} \psi_{j,k}(t) \quad (7.9)$$

その展開係数 $w_k^{(j)}$ は正規直行であれば次のように内積の形で与えられる .

$$w_k^{(j)} = \int_{-\infty}^{\infty} f(t) \overline{\psi_{j,k}(t)} dt = \langle f, \psi_{j,k} \rangle \quad (7.10)$$

7.1.2.1 スケーリング関数と近似関数

信号 $f(t)$ は , 信号を観測するための尺度であるスケーリング関数 $\phi_{j,k}$ の一次結合で近似する . これを近似関数という . また近似の精度をレベルと呼び , レベル 0 が最も精度の高い近似となる . このレベルを導入して , レベル j の近似関数は $j(t)$ と表記して , 以下のように定義する .

$$j(t) = \sum_k s_k^{(j)} \phi_{j,k}(t) \quad (7.11)$$

ここで , $s_k^{(j)}$ はスケーリング係数と呼び , 以下のような式で定義する .

$$s_k^{(j)} = \int_{-\infty}^{\infty} f(t) \overline{\phi_{j,k}(t)} dt \quad (7.12)$$

また , 任意のレベル j の近似関数 $j(t)$ と , それより 1 レベル粗い $j+1$ レベルの近似関数 $j+1(t)$ との差分はウェーブレットの一次結合である . つまり , 信号は , レベル 1 から , 任意のレベルまでのウェーブレットと , スケーリング関数との一次結合との和で表される .

7.1.2.2 多重解像度解析

$f_j(t)$ は $f_{j-1}(t)$ より信号 $f(t)$ の粗い近似であり , 情報が欠落しているので , 欠落分 $g_j(t)$ を $f_j(t)$ に補うことにより $f_{j-1}(t)$ が復元される .

$$f_{j-1}(t) = f_j(t) + g_j(t) \quad (7.13)$$

ここで ,

$$f_j(t) = \sum_k s_k^{(j)} \phi_{j,k}(t) \quad (7.14)$$

$$g_j(t) = \sum_k w_k^{(j)} \psi_{j,k}(t) \quad (7.15)$$

$$(7.16)$$

この関数 $g_j(t)$ をレベル j のウェーブレット成分という．この関係式を繰り返し用いることにより，以下の式が得られる $f_0(t) = g_1(t) + g_2(t) + \dots + g_J(t) + f_J(t) = \sum_{j=1}^J g_j(t) + f_J(t)$ (7.17) これより，信号 $f_0(t)$ を，任意の荒っぽさのレベルの近似関数 $f_J(t)$ と，レベル 0 からレベル J に至るまでのウェーブレット成分の和とで表現できる．このことより，信号 $f_0(t)$ を 1 から J レベルまでの J 個の解像度，すなわち多重の解像度を持つウェーブレットを用いて表すことができる．この解析を多重解像度解析という．

7.1.2.3 ツースケール関係

多重解像度解析のためには，レベル j のスケーリング関数 $\phi_{j,k}$ がレベル $j-1$ のスケーリング関数 $\phi_{j-1,k}$ で展開できなければならない．ツースケール関係とはこの条件を満たす，レベル j と $j-1$ とのスケール関数の関係をあらわす式 (equ:p) と，レベル j と $j-1$ とのスケール関数とウェーブレットとの関係をあらわす式 (7.19) よりなる二つの関係式のことである．

$$\phi_{j,k}(t) = \sum_n p_{n-2k} \phi_{j-1,n}(t) \quad (7.18)$$

$$\psi_{j,k}(t) = \sum_n q_{n-2k} \phi_{j-1,n}(t) \quad (7.19)$$

ここで， p_n と q_n はウェーブレットを表す数列である．本研究では $N = 2$ の場合のドベシの数列を使っており，以下ようになる（ただし， $q_n = (-1)^n p_{-n}$ ）

表 7.1: $N = 2$ のドベシの数列 p_n

p_0	0.482962913145
p_1	0.836516303738
p_2	0.224143868042
p_3	-0.129409522551

7.1.2.4 分解・再構成

ここまで直行ウェーブレット及びそのスケーリング関数は離散数列で表されることがわかった．この項では，離散数列 $f(n)$ を， $s_k^{(0)}$ とみなすマラーの方法に従う．式 (7.12) の中のスケーリング関数 $\phi_{j,k}(t)$ を，ツースケール関係式 (7.18) を利用して $j-1$ レベルのスケーリング関数 $\phi_{j-1,n}(t)$ で表すと次の関係式を得る．

$$s_k^{(j)} = \int_{-\infty}^{\infty} f(t) \sum_n \overline{p_{n-2k} \phi_{j-1,n}(t)} dt \quad (7.20)$$

$$= \sum_n \overline{p_{n-2k}} \int_{-\infty}^{\infty} f(t) \overline{\phi_{j-1,n}(t)} dt \quad (7.21)$$

$$= \sum_n \overline{p_{n-2k}} s_n^{(j-1)} \quad (7.22)$$

この式より，レベル 0 のスケーリング係数 $s_k^{(0)}$ から逐次レベル数が大きいスケーリング係数を求めることができる．

また，式 (7.10) 中のウェーブレット $\psi_{j,k}^{(t)}$ を，ツースケール関係式 (7.19) を利用して $j-1$ レベルのスケーリング関数 $\phi_{j-1,n}(t)$ で表すと次の関係式を得る．

$$w_k^{(j)} = \int_{-\infty}^{\infty} f(t) \sum_n \overline{q_{n-2k} \phi_{j-1,n}(t)} dt \quad (7.23)$$

$$= \sum_n \overline{q_{n-2k}} \int_{-\infty}^{\infty} f(t) \overline{\phi_{j-1,n}(t)} dt \quad (7.24)$$

$$= \sum_n \overline{q_{n-2k}} s_n^{(j-1)} \quad (7.25)$$

この式より，レベル 0 のスケーリング係数 $s_k^{(0)}$ から逐次レベル数が大きいウェーブレット展開係数を求めることができる．

以上より，スケーリング係数が，次々に精度の低いスケーリング係数とウェーブレット展開係数とに分解されることがわかる (図 7.5) ．

一方，再構成は以下の式による．

$$s_n^{(j-1)} = \sum_k [p_{n-2k} s_k^{(j)} + q_{n-2k} w_k^{(j)}] \quad (7.26)$$

この式より，精度の高いレベルのスケーリング係数が求められる (図??) ．

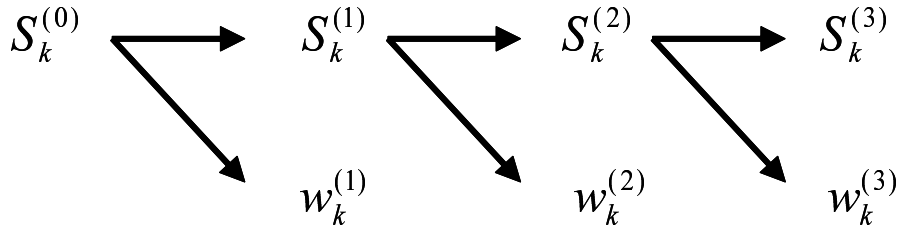


図 7.5: スケーリング関数の分解

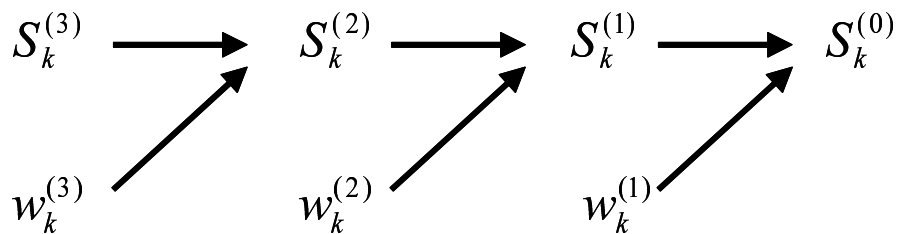


図 7.6: スケーリング関数の再構成

7.2 二次元離散ウェーブレット変換

ここでは7.1.2を拡張して二次元信号のウェーブレット変換について説明する．二次元信号を $f(x, y)$ とすると，まず横軸方向に離散ウェーブレット変換を行い，その係数に対して縦軸方向に離散ウェーブレット変換を行う．7.1.2 同様，二次元信号 $f(x, y)$ をレベル0のスケーリング係数 $s_{x,y}^{(0)}$ とみなす．

まず横軸方向に離散ウェーブレット変換を行う．

$$s_{x,y}^{(j+1,m)} = \sum_k \overline{p_{k-2x}} s_{k,y}^{(j)} \quad (7.27)$$

$$w_{x,y}^{(j+1,m)} = \sum_k \overline{q_{k-2x}} s_{k,y}^{(j)} \quad (7.28)$$

ただし，

- $s_{x,y}^{(j+1,x)}$: 横軸方向のスケーリング係数
- $w_{x,y}^{(j+1,x)}$: 横軸方向のウェーブレット展開係数

$j = 0$ の場合を図 7.7 に示す .

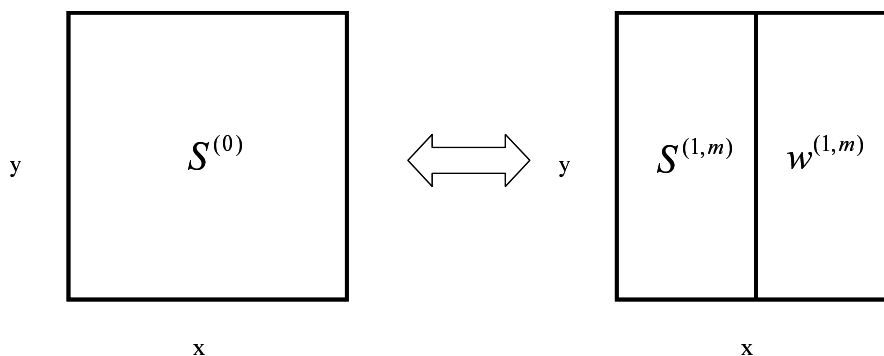


図 7.7: 横軸方向の二次元ウェーブレット変換

次にそれぞれの係数に対して、縦軸方向に離散ウェーブレット変換を行う .

$$s_{x,y}^{(j+1)} = \sum_l \overline{p_{l-2y}} s_{x,l}^{(j+1,m)} \quad (7.29)$$

$$w_{x,y}^{(j+1,a)} = \sum_l \overline{q_{l-2y}} s_{x,l}^{(j+1,m)} \quad (7.30)$$

$$w_{x,y}^{(j+1,b)} = \sum_l \overline{p_{l-2y}} w_{x,l}^{(j+1,m)} \quad (7.31)$$

$$w_{x,y}^{(j+1,c)} = \sum_l \overline{q_{l-2y}} w_{x,l}^{(j+1,m)} \quad (7.32)$$

ただし ,

- $s_{x,y}^{(j+1)}$: 横軸方向 , 縦軸方向共にスケーリング関数を実作用させた係数
- $w_{x,y}^{(j+1,a)}$: 横軸方向にスケーリング関数 , 縦軸方向にウェーブレットを実作用させた係数
- $w_{x,y}^{(j+1,b)}$: 横軸方向にウェーブレット , 縦軸方向にスケーリング関数を実作用させた係数

- $w_{x,y}^{(j+1,c)}$: 横軸方向, 縦軸方向共にウェーブレットを作用させた係数

$j = 1$ の場合を図 7.8 に示す .

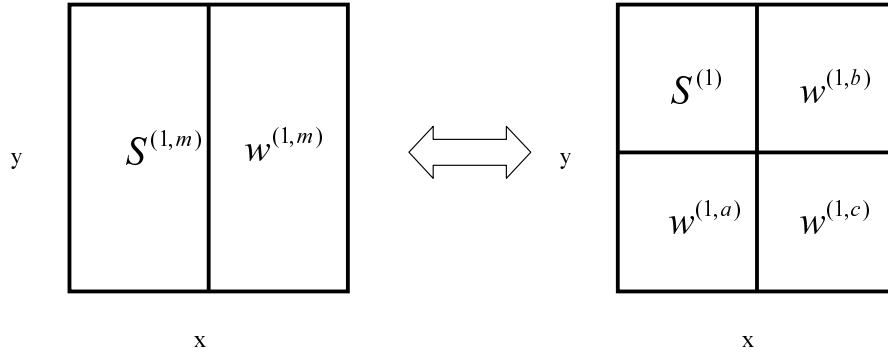


図 7.8: 縦軸方向の二次元ウェーブレット変換

以上をまとめると次式のようになる .

$$s_{x,y}^{(j+1)} = \sum_l \sum_k \overline{p_{k-2x} p_{l-2y}} s_{k,l}^{(j)} \quad (7.33)$$

$$w_{x,y}^{(j+1,a)} = \sum_l \sum_k \overline{p_{k-2x} q_{l-2y}} s_{k,l}^{(j)} \quad (7.34)$$

$$w_{x,y}^{(j+1,b)} = \sum_l \sum_k \overline{q_{k-2x} p_{l-2y}} s_{k,l}^{(j)} \quad (7.35)$$

$$w_{x,y}^{(j+1,c)} = \sum_l \sum_k \overline{q_{k-2x} q_{l-2y}} s_{k,l}^{(j)} \quad (7.36)$$

この再構成は, 一次元の場合の再構成と同様, 次式により行う .

$$s_{x,y}^{(j)} = \sum_k \sum_l \left[p_{x-2k} p_{y-2l} s_{k,l}^{(j+1)} + p_{x-2k} q_{y-2l} w_{k,l}^{(j+1,a)} + q_{x-2k} p_{y-2l} w_{k,l}^{(j+1,b)} + q_{x-2k} q_{y-2l} w_{k,l}^{(j+1,c)} \right] \quad (7.37)$$

上式において, $s_{x,y}^{(j+1)}$ のみをさらに四つの成分に分解することを繰り返すことにより, 二次元の多重解像度分解を行う (図 7.9) .

二次元信号 $f(x, y)$ を二次元画像とすると, 二次元ウェーブレット変換により画像の分解・再構成を扱うことができる .

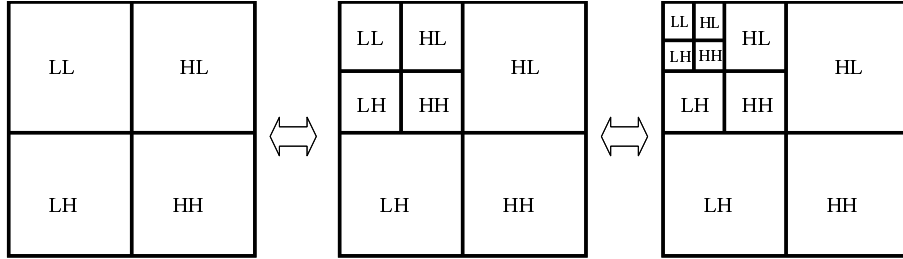


図 7.9: 二次元の多重解像度分解

7.3 三次元離散ウェーブレット変換

??と同様，三次元信号のウェーブレット変換について説明する．三次元信号を $f(x, y, z)$ とすると，まず横軸方向に離散ウェーブレット変換を行い，その係数に対して縦軸方向に離散ウェーブレット変換を行い，さらにその係数に対しておく方向に離散ウェーブレット変換を行う（図 7.10）．

分解の式は以下ようになる．

$$s_{x,y,z}^{(j+1)} = \sum_m \sum_l \sum_k \overline{p_{k-2x} p_{l-2y} p_{m-2z}} s_{k,l,m}^{(j)} \quad (7.38)$$

$$w_{x,y,z}^{(j+1,a)} = \sum_m \sum_l \sum_k \overline{p_{k-2x} p_{l-2y} p_{m-2z}} s_{k,l,m}^{(j)} \quad (7.39)$$

$$w_{x,y,z}^{(j+1,b)} = \sum_m \sum_l \sum_k \overline{p_{k-2x} q_{l-2y} p_{m-2z}} s_{k,l,m}^{(j)} \quad (7.40)$$

$$w_{x,y,z}^{(j+1,c)} = \sum_m \sum_l \sum_k \overline{p_{k-2x} q_{l-2y} q_{m-2z}} s_{k,l,m}^{(j)} \quad (7.41)$$

$$w_{x,y,z}^{(j+1,d)} = \sum_m \sum_l \sum_k \overline{q_{k-2x} p_{l-2y} p_{m-2z}} s_{k,l,m}^{(j)} \quad (7.42)$$

$$w_{x,y,z}^{(j+1,e)} = \sum_m \sum_l \sum_k \overline{q_{k-2x} p_{l-2y} q_{m-2z}} s_{k,l,m}^{(j)} \quad (7.43)$$

$$w_{x,y,z}^{(j+1,f)} = \sum_m \sum_l \sum_k \overline{q_{k-2x} q_{l-2y} p_{m-2z}} s_{k,l,m}^{(j)} \quad (7.44)$$

$$w_{x,y,z}^{(j+1,g)} = \sum_m \sum_l \sum_k \overline{q_{k-2x} q_{l-2y} q_{m-2z}} s_{k,l,m}^{(j)} \quad (7.45)$$

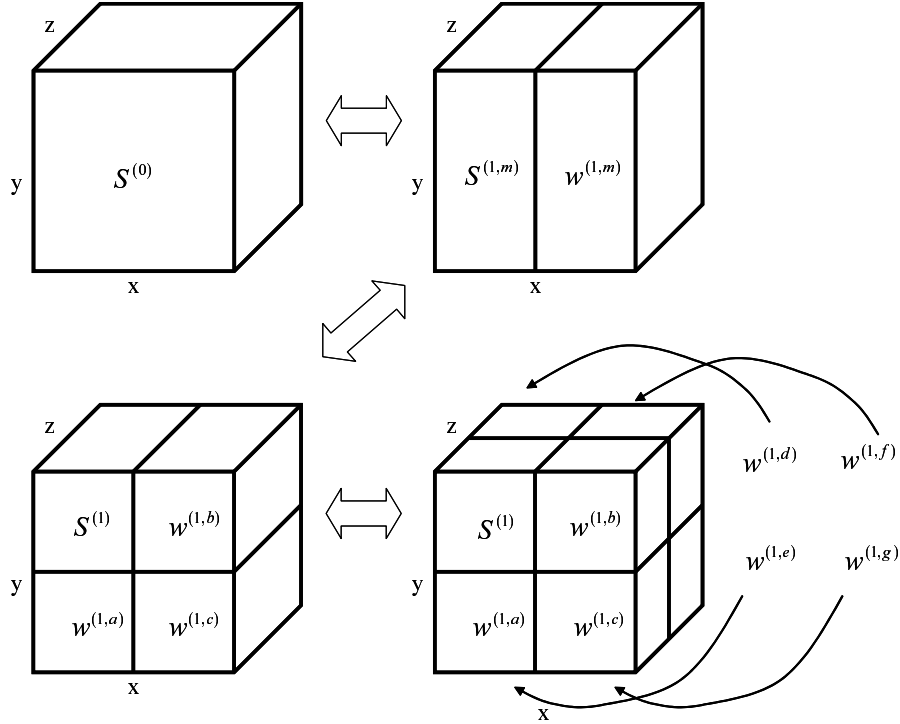


図 7.10: 三次元ウェーブレット変換

(7.46)

この再構成は次式により行う。

$$\begin{aligned}
 s_{x,y,z}^{(j)} = \sum_k \sum_l \sum_m & \left[p_{x-2k} p_{y-2l} p_{z-2m} s_{k,l,m}^{(j+1)} + p_{x-2k} p_{y-2l} q_{z-2m} w_{k,l,m}^{(j+1,a)} + \right. \\
 & p_{x-2k} q_{y-2l} p_{z-2m} w_{k,l,m}^{(j+1,b)} + p_{x-2k} q_{y-2l} q_{z-2m} w_{k,l,m}^{(j+1,c)} + \\
 & q_{x-2k} q_{y-2l} p_{z-2m} w_{k,l,m}^{(j+1,d)} + q_{x-2k} q_{y-2l} q_{z-2m} w_{k,l,m}^{(j+1,e)} + \\
 & \left. q_{x-2k} q_{y-2l} p_{z-2m} w_{k,l,m}^{(j+1,f)} + q_{x-2k} q_{y-2l} q_{z-2m} w_{k,l,m}^{(j+1,g)} \right]
 \end{aligned}$$

上式において、 $s_{x,y,z}^{(j+1)}$ のみをさらに八つの成分に分解することを繰り返すことにより、三次元の多重解像度分解を行う（図 7.11）。

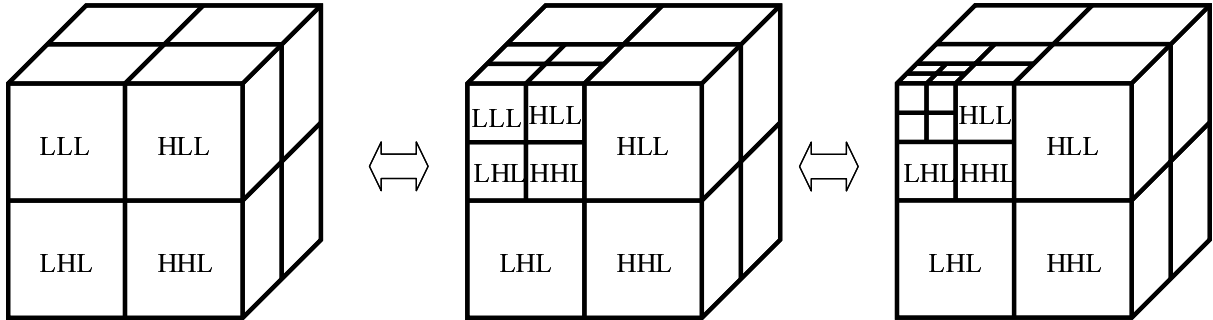


図 7.11: 三次元の多重解像度分解

三次元信号 $f(x, y)$ を動画像（画像列）とすると，三次元ウェーブレット変換により動画像（画像列）の分解・再構成を扱うことができる．

7.4 離散ウェーブレット変換による情報圧縮・再構成

離散ウェーブレット変換による分解・再構成は情報の欠落のない変換であり，絶対値の大きい順に上位数パーセントの展開係数にエネルギーが集中していた．ここで，絶対値の大きい展開係数のみ残し，残りの係数は 0 にして再構成すれば，信号の特徴を保存したまま少ないデータ量で信号が再構成される．この性質を利用すれば，離散ウェーブレット変換による情報圧縮が可能となる．

7.1.2 で述べた多重解像度解析により信号を分解して， $j = 0$ から $j = J - 1$ までの展開係数を求める．この J 個の展開係数を絶対値の大きい順に並べ，絶対値が大きい順に上位 α % の展開係数のみを残し，その他の展開係数を 0 にする．この展開係数を使って再構成された信号は圧縮率 α % で圧縮されている．この圧縮は不可逆圧縮である．

7.4.1 画像圧縮：二次元離散ウェーブレット変換による情報圧縮・再構成

ウェーブレット変換による情報圧縮を二次元信号，つまり画像について行うことで画像圧縮が可能となる．画像圧縮の目的は，見た目を損なわないでデータ量を低減させることである．圧縮率は再構成に使用する展開係数の比率，復元性能の数値的尺度は平均誤差 RMSE を用いる．また見た目には，目障りなノイズが現れないか，輪郭の再現性，高い周波数成分を含む画像パターンの再現性等を評価基準とする．図 7.12 に二次元離散ウェーブレット変換による画像の分解・再構成過程を示す．ただし，この場合はレベル 3 までの分解である．

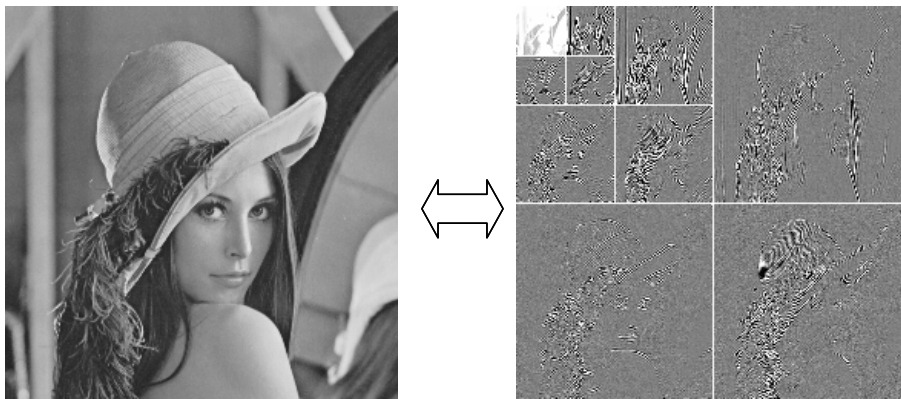


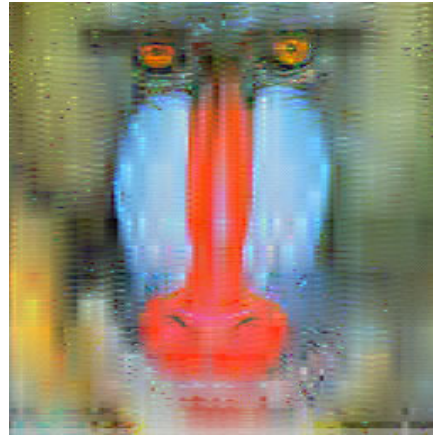
図 7.12: 画像の分解・再構成過程

このように分解した展開係数をいくつかの比率で切り上げ，再構成した画像を図 7.13，図 7.14 に示す．

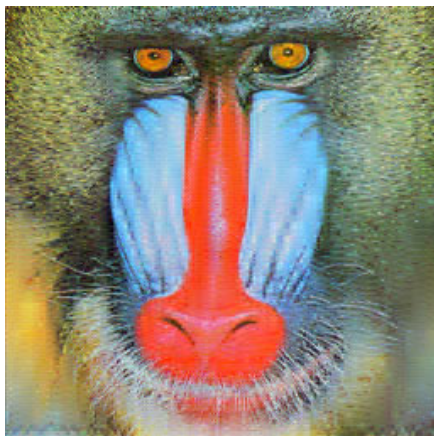
画像データ圧縮の標準規格であり，現在主流となっている JPEG では離散コサイン変換 discrete cosine transform:DCT が使用されている．JPEG では画像を 8×8 画素のブロックに分割して DCT をして，ブロック単位で展開係数を求めている．そのため，高圧縮率の画像ではブロック状のノイズが目立つ．離散コサイン変換はコサイン関数のみを基底として使用し，信号と正弦波または余弦波の相関を測るものである．しかし，正弦波または余弦波は無限に続く関数であり，信号が時間軸上に



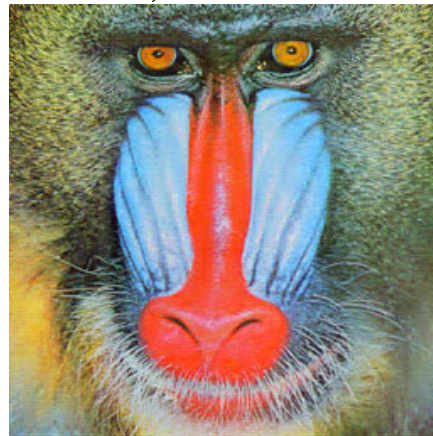
original



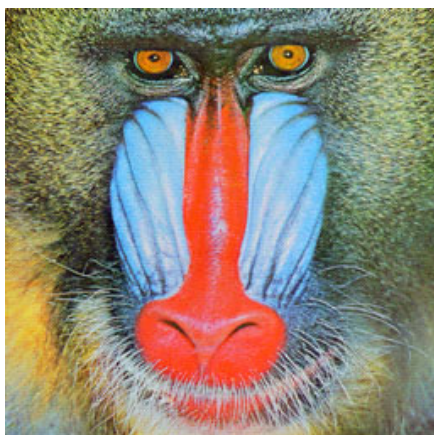
1%, RSME-38



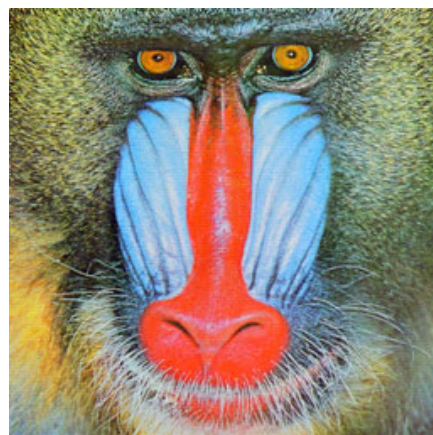
2%, RSME-13



3%, RSME-3.8



4%, RSME-0.8



5%, RSME-0.14

図 7.13: Mandrill 画像の圧縮



original



1%, RSME-7.4



2%, RSME-1.8



3%, RSME-0.16



4%, RSME-0.06



5%, RSME-0.03

図 7.14: Parrots 画像の圧縮

局在したものであるならば，正弦波または余弦波のように無限に続く関数にはあまりにない可能性がある．それに対して離散ウェーブレット変換は適切な直交離散ウェーブレットを基底として使用し，信号とウェーブレットの相関を測るものである．信号が時間軸に局在したものである場合，信号によく似たウェーブレットを用いたならば適切に検出できることがウェーブレットの特徴である．なお従来の JPEG よりも高圧縮，高品質な画像圧縮が行なえる新規格の JPEG2000 ではウェーブレット変換を用いており，ブロック状のノイズは解消されており，また様々な解像度の画像への変換ができるようになっている．

7.4.2 画像列（動画）圧縮：三次元離散ウェーブレット変換による情報圧縮・再構成

7.4.1 と同様にして，ウェーブレット変換による情報圧縮を三次元信号，つまり画像列（動画）について行うことで画像列圧縮が可能となる．本研究の圧縮・再構成はこの画像列圧縮を用いて，大幅な情報量の削減を行っている．ここで用いるウェーブレットとは先述したようにハールウェーブレットやガボールウェーブレットとより画像再構成において優れたドベシウエーブレットである．

JPEG と同じように現在主流の圧縮形式として MPEG があげられる．MPEG は画像列の各フレームごとに独立に二次元 DCT をおこなっているのに対し，三次元離散ウェーブレット変換は各フレームごとではなく，画像列全体で三次元変換を行うので MPEG と比べても効果的な圧縮ができるといえる．

第8章

仮想現実感モデル生成システム

本章では、これまでに述べてきた諸要素を統合した画像合成システムについて説明する。流れとしては、まず対象物体入力データを準備する。入力データはレンジセンサにより取得された幾何情報と、カラーセンサにより取得された光学情報の二種類である。この幾何情報から物体の幾何モデルの取得、円偏光板による分離操作により、光学情報から鏡面反射画像・拡散反射画像の取得、そしてその両者の二次元三次元対応をとるキャリブレーションを行った後、反射性分別の Torrance-Sparrow モデルに則した反射パラメータを推定する。

次に、Torrance-Sparrow モデルによるレンダリングを行う。物体の幾何モデルと鏡面・拡散反射パラメータを保存しておけばレンダリングを行えるというのが Torrance-Sparrow モデルによるレンダリングのアイディアである。しかし Torrance-Sparrow モデルには、相互反射を考慮していないことや、物体表面が粗い面で構成されている必要がある等の制約条件があるため、Torrance-Sparrow モデルによるレンダリングでは汎用性が低い。

そこで、Torrance-Sparrow モデルによるレンダリングでは再現しきれない情報をも保存してレンダリングを行うというのが本研究のアイディアである。この情報というのは、原画像と Torrance-Sparrow モデルによるレンダリング画像との差分画像である。しかしこの複数の差分画像は情報量が大きいのでウェーブレットにより三次元圧縮を施し、レンダリングにおいてはウェーブレットによる再構成を行った差分画像を、Torrance-Sparrow モデルによるレンダリングに補完する。

この一連の情報圧縮・再構成（レンダリング）の流れを図 8.1, 8.2 に示す。図 8.1 は圧縮過程を示しており、入力情報の光学情報、幾何情報から拡散反射パラメータ、鏡面反射パラメータ、対応点情報、圧縮差分画像を得るまでの過程、図 8.2 は図 8.1 の過程で得た情報より、画像合成をする過程を示す。以下の節ではそれぞれの要素を詳細に説明する。

8.1 光学情報

入力データの一つ、光学情報は、デジタルカメラのようなカラーセンサで取得された対象物体の画像である。本研究では第 2 章で説明した Torrance-Sparrow の二色性反射モデルを用いるので、拡散反射成分画像と鏡面反射成分画像を用意する。ここで、第 3 章で説明したように円偏光板を通して撮像した画像と、通さないで撮像

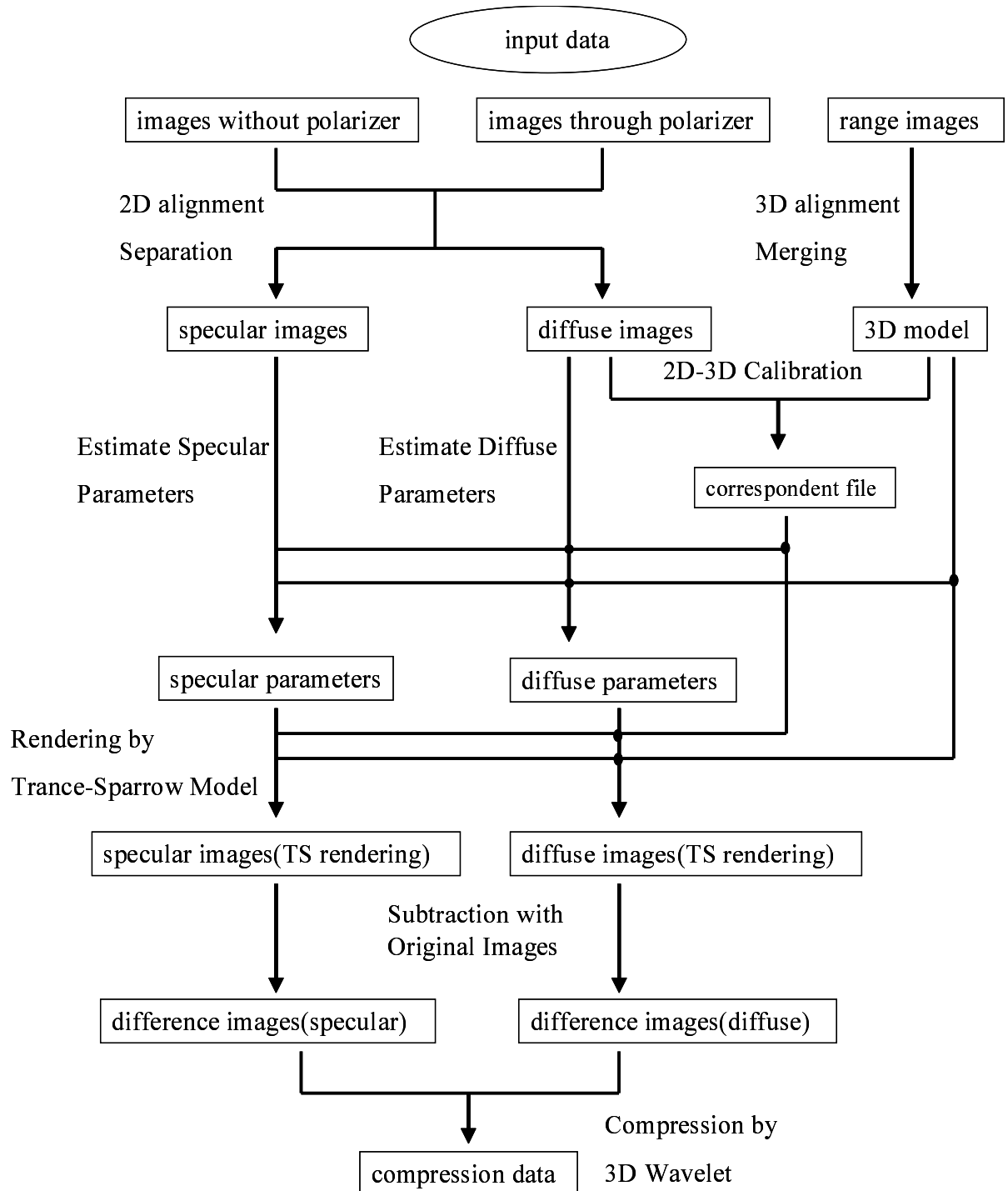


図 8.1: フローチャート : 圧縮過程

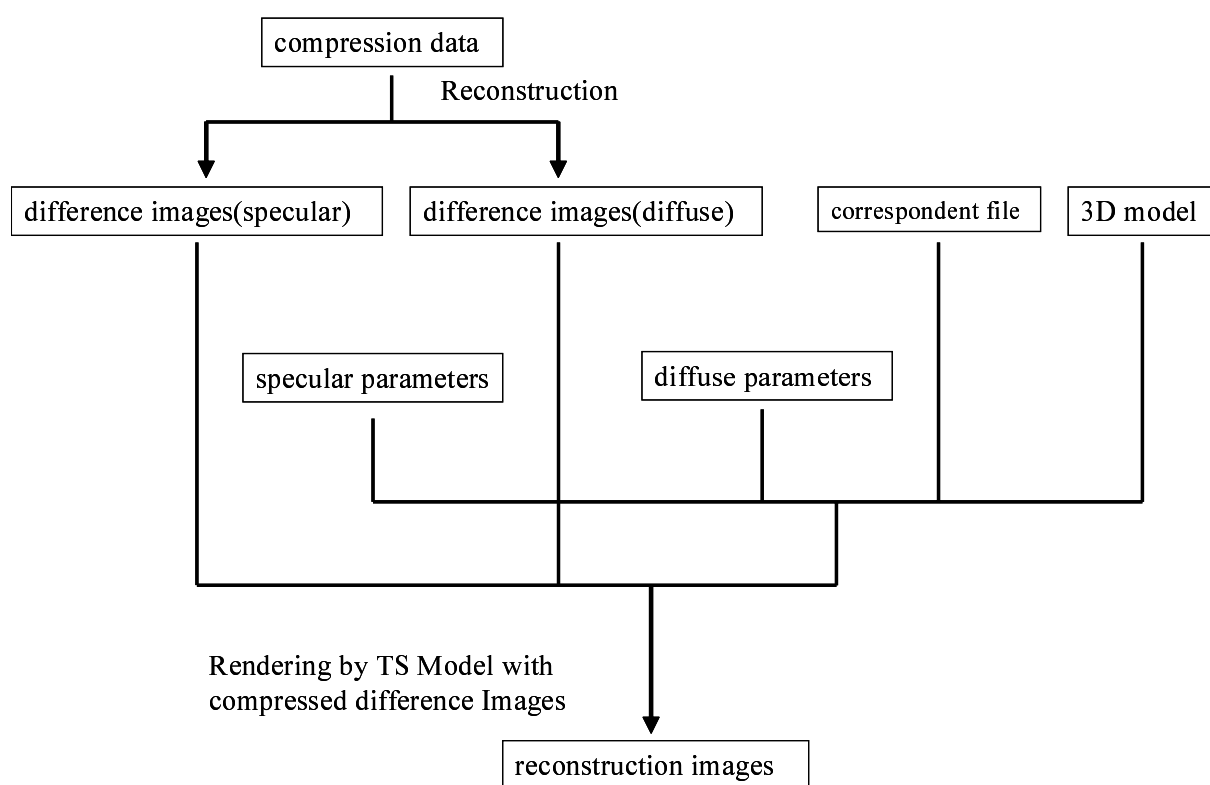


図 8.2: フローチャート：再構成過程

した画像より，拡散反射成分画像と鏡面反射成分画像を求めるのだが，円偏光板を通すことにより位置のずれが生じるので，補正をしてやる必要がある．対象物体全体の反射パラメータを推定するために，ある回転軸を中心に一定角度ごとに回転させた画像列を用いる．光学情報の例を図 8.3 に示す．

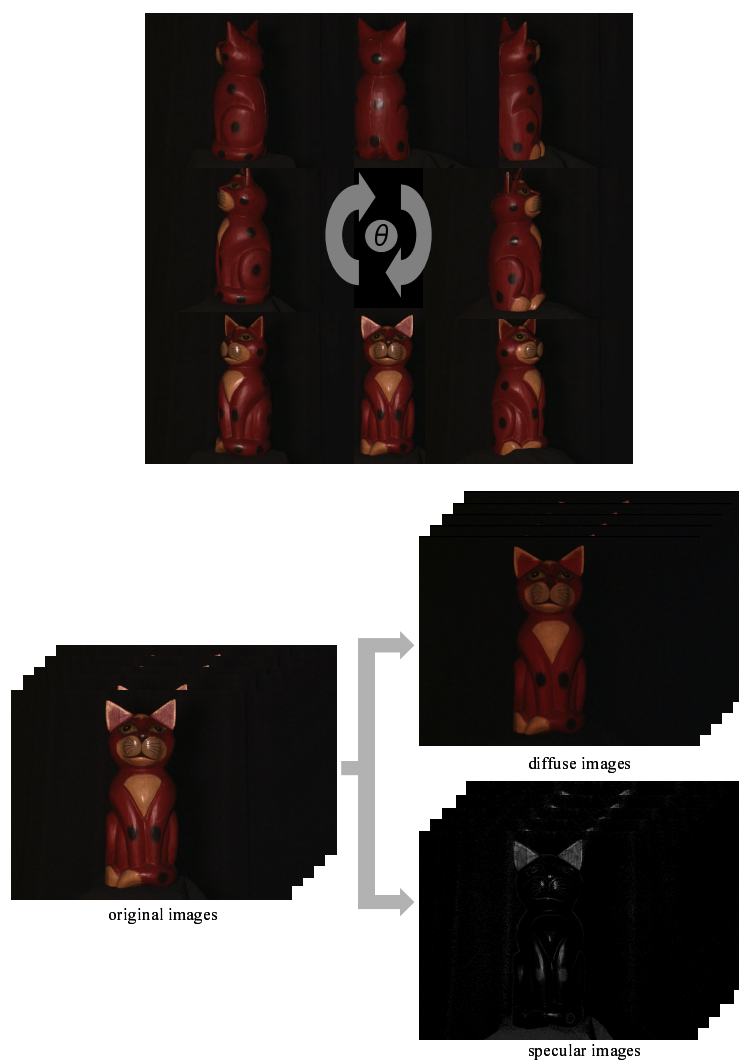


図 8.3: 上：全周囲の画像，下：反射成分分離

8.2 幾何情報

もう一つの入力データ，幾何情報は，対象物体の幾何モデルである．これは第 4 章で説明したように，レンジセンサにより様々な方向より得られた距離画像を，アライメント，マージングすることにより得られる．幾何情報取得の例を図 8.4 に示す．

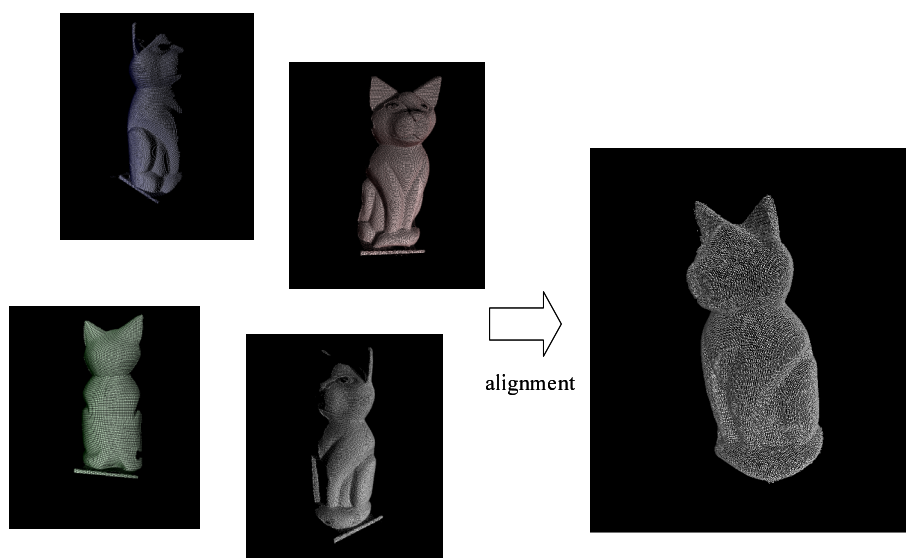


図 8.4: 距離画像から 3D モデルの生成

8.3 キャリブレーション

以上の二つの入力データについて，一段階目の処理としてキャリブレーションを行う．第五章で説明したように，カラーセンサーの置かれているカメラ座標と，レンジセンサーの置かれている世界座標との対応を取り，三次元モデルの元の位置から対応する二次元画像への移動行列を求めるのがこの操作である．このキャリブレーションを全ての入力画像について行う（図 8.5 上）．

なお本論文で提案する手法は，差分画像を補完して反射モデルでは表現しきれない部分についても再現すること，情報圧縮も目的の一つとしていることより，図 8.5 下のように 3D モデルの面数が 4000 程度となるように簡易化している．

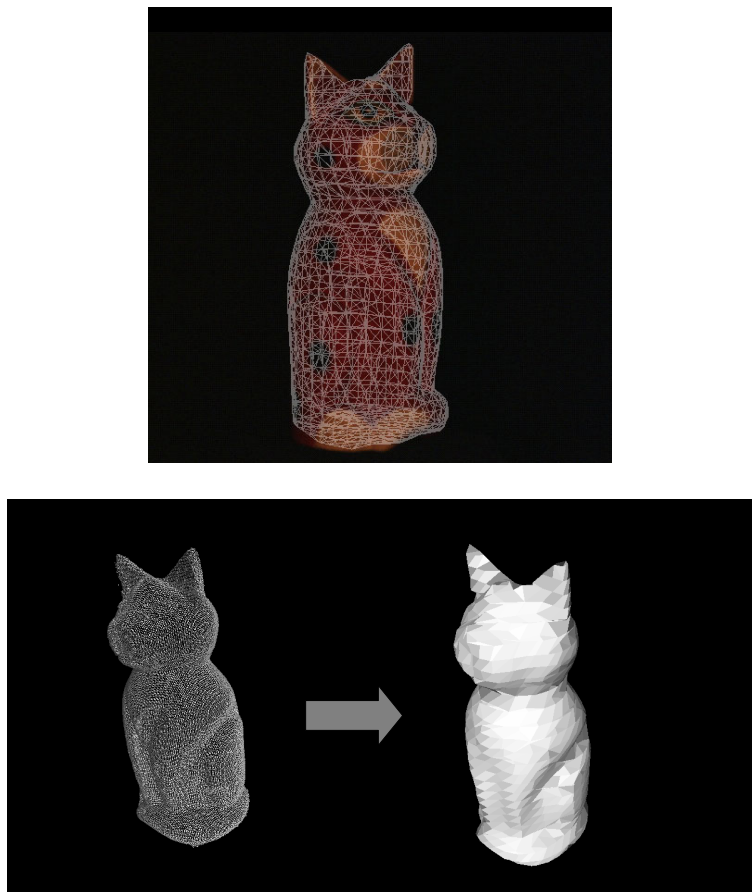


図 8.5: 上：2D-3D 対応，下：幾何モデルの簡略化

8.4 反射パラメータ推定

第六章で説明したように，拡散反射パラメータ，鏡面反射パラメータをそれぞれ求める．

推定して求めた拡散反射パラメータは幾何モデルの一点につき $K_{d,R}$, $K_{d,G}$, $K_{d,B}$ の三つである．拡散反射パラメータとは光源によるシェーディングの影響を受けていない，物体本来の色であるから，頂点ごとのデータではなく，面ごとのデータとして保存できる．この三角形の面ごとの領域に画像を分割し，正規化して，図 8.6 左のような三角パッチの集合画像として保存する．この三角パッチの並びは幾何モデルの面番号に対応しており，左上から右方向に向かって並んでいる．また，三角

パッチの各頂点は幾何モデルの頂点に対応している．第九章に記載する実験結果は 1024×1024 サイズの画像に格納しており，一つの三角パッチのサイズは 16×16 である．

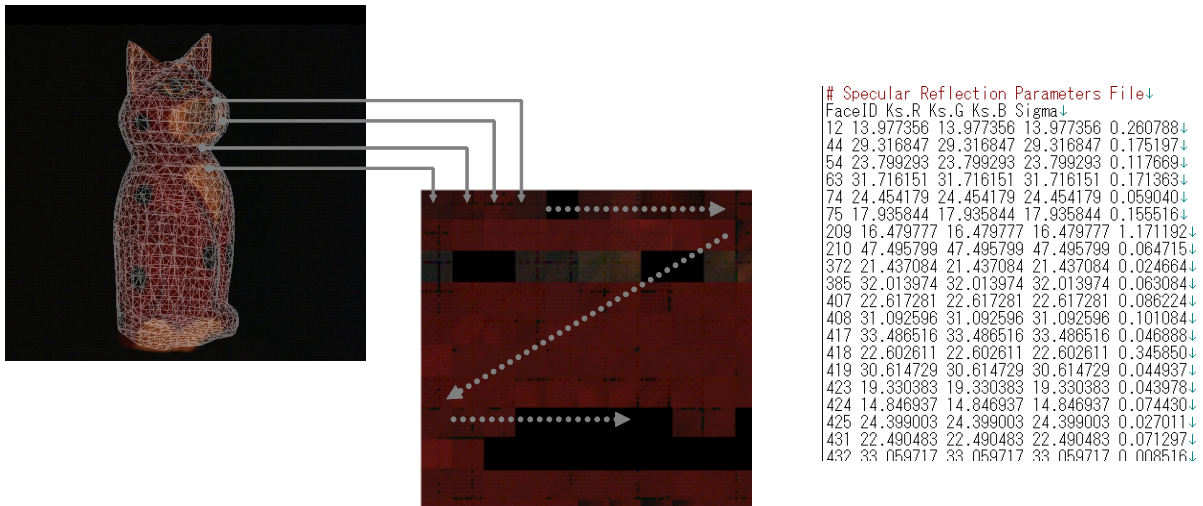


図 8.6: 左：拡散反射パラメータ，右：鏡面反射パラメータ

また，鏡面反射パラメータは幾何モデルの一点につき $K_{s,R}$, $K_{s,G}$, $K_{s,B}$, Σ の四つである．鏡面反射パラメータは頂点ごとに求められ，図 8.6 右のような数値データとして保存される．しかし，実際にレンダリングで用いる際は，面全体でレンダリングをする必要があるので，面を構成する三頂点より補完を行い，拡散反射パラメータのように面全体のデータに拡張する．

8.5 Torrance-Sparrow レンダリング

次に，推定した反射パラメータより，Torrance-Sparrow モデルのレンダリングを行う．Torrance-Sparrow の式は以下のものであった．

$$I = K_d \cos\theta + K_s \frac{\exp(-\frac{\alpha^2}{2\sigma^2})}{\cos\theta} \quad (8.1)$$

レンダリング結果の例を図 8.7 に示す．

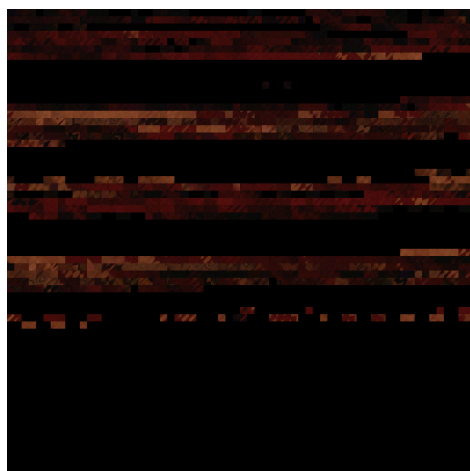


図 8.7: Torrance-Sparrow レンダリング画像

8.6 差分画像

8.5 の Torrance-Sparrow レンダリングの結果は入力画像との誤差が大きい。この誤差差分をもレンダリングのための情報として保存するのが本研究の趣旨である。差分画像は 8.5 のように二色性反射モデルを用いているので、拡散反射成分・鏡面反射成分それぞれ成分別に作成する。

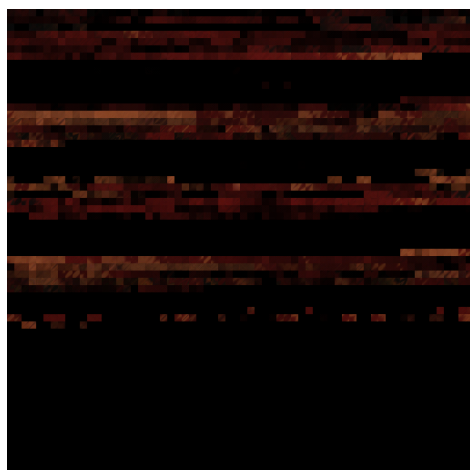
まず、元データとしての入力画像を、8.4 のように三角パッチの集合画像に展開する (図 8.8(a),(b))。次に、推定された各反射パラメータを用いて、入力画像に対応した視点と光源条件のもとで三角パッチの集合画像にレンダリングを行う (図 8.8(c),(d))。これらの差をとったものを差分画像とする。ここで画像のピクセル値は符号無整数型で 0 ~ 255 の値でデータ保持されているが、負の値をとりうること、また次の過程のウェーブレット変換においては小数との積計算を行うことより差分画像は浮動点小数方の配列にデータを格納する。図 8.8(e),(f) は差分画像であるが、負の値も含むので視覚化のために、絶対値を取り 256 段階の整数としてある。この操作を入力画像の枚数分行う。



(a) diffuse patch(original)



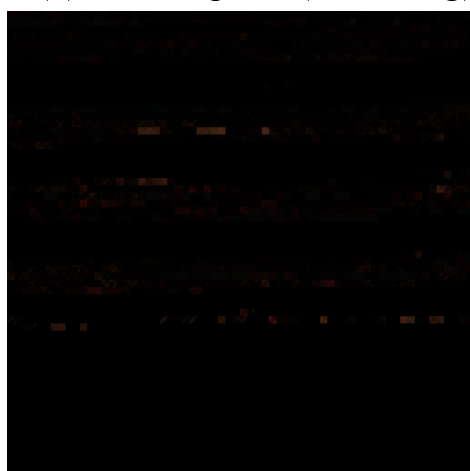
(b) specular patch(original)



(c) diffuse patch(rendering)



(d) specular patch(rendering)



(e) difference patch(diffuse)



(f) difference patch(specular)

8.7 ドベシウェーブレットによる差分画像圧縮

8.6 で入力画像の枚数分の差分画像ができるが、これはデータ量が非常に大きい。そこでこの差分画像列を第七章で述べた、三次元離散ウェーブレット変換による情報圧縮を行う。本研究では差分画像列全体で三次元離散ウェーブレット変換を行うのではなく、差分画像列の三角パッチ画像列ごとに三次元離散ウェーブレット変換を行う。何故ならば、この三角パッチ画像列は幾何モデルの同一面の、違う視点条件の画像により構成されているので、三次元空間上の奥行き方向の相関が非常に高く、高い圧縮率が得られるからである。また三角パッチごとに圧縮率を変化させることで、差分画像列全体で圧縮を行う場合に比べてより効率的な圧縮が実現できる。画像からパッチごとにパッチ画像列を作り、ウェーブレット変換により段階的に分解していく様子を図 8.9 に示す。

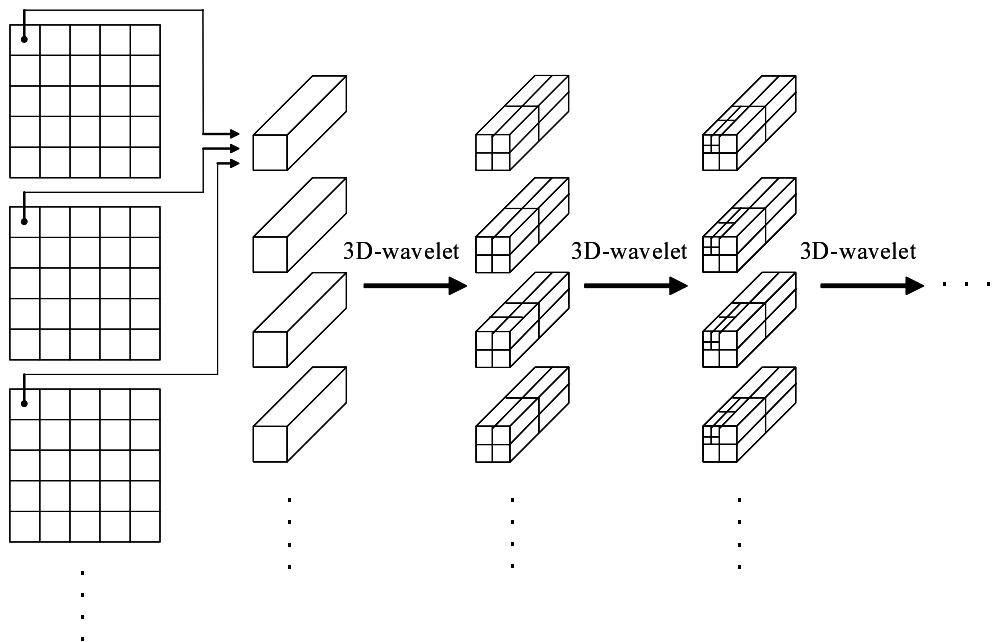


図 8.9: 分解過程

また分解したパッチ画像列をウェーブレット変換により段階的に再構成していき、元の画像を構成する様子を図 8.10 に示す。

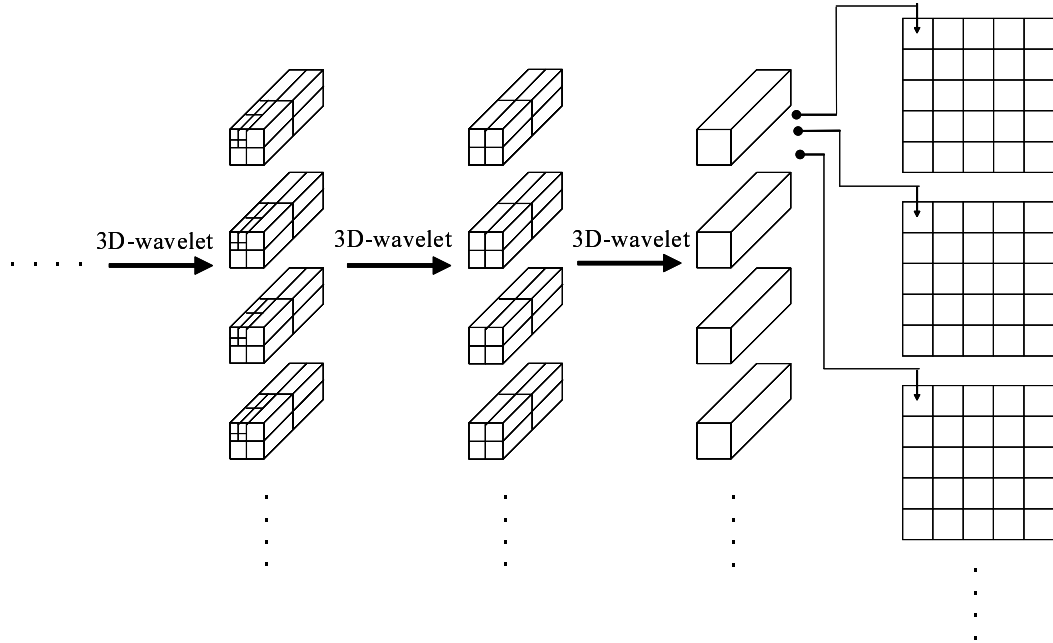


図 8.10: 再構成過程

この分解・再構成の過程において，パッチ画像列を分解した後，絶対値が大きい順に上位 α % の展開係数のみを残し，その他の展開係数を 0 にして再構成を行うことにより情報量を圧縮する．パッチごとの画像圧縮の様子を図 8.11 に示す．

8.8 差分圧縮画像補完付きレンダリング

最終的にレンダリングに用いるデータは，幾何モデル，鏡面反射パラメータ，拡散反射パラメータ，差分圧縮画像列である．

まず，幾何モデル，鏡面反射パラメータ，拡散反射パラメータにより 8.5 と同様にして Torrance-Sparrow レンダリングを行う．次に，8.7 により圧縮された差分画像列を再構成して，各 Torrance-Sparrow レンダリング画像に補完してレンダリングを行う．この差分画像補完を行うことにより Torrance-Sparrow モデルでは再現し切れなかった情報をもレンダリングできることがわかる（図 8.12）．

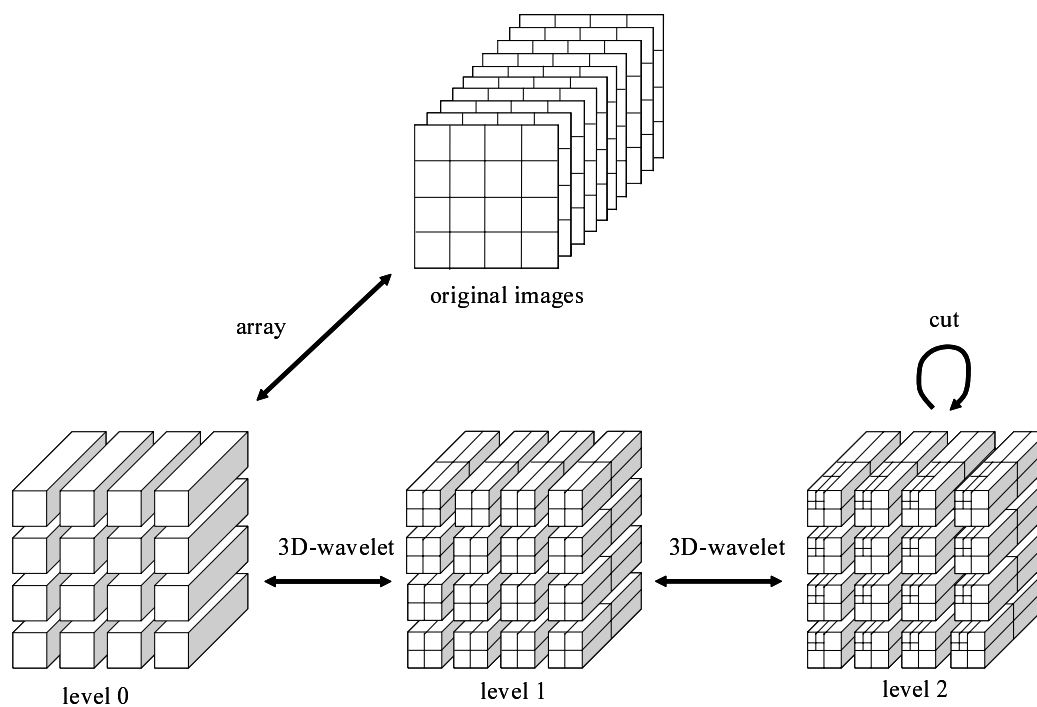


図 8.11: パッチごとの画像圧縮

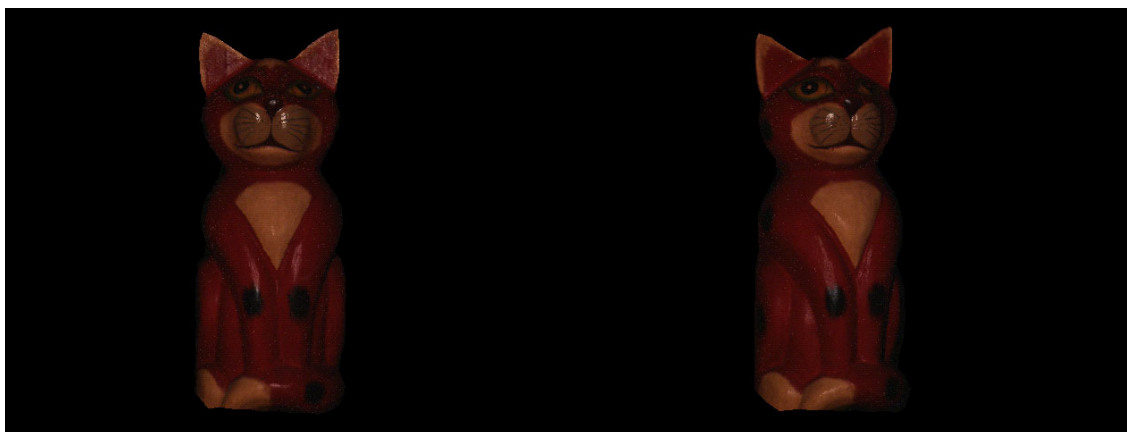


図 8.12: パッチ形式差分圧縮画像補完レンダリング画像

第9章

実験・評価

9.1 データ取得

本研究では図 9.1 のようなデータ取得システムを用いている。モデル化される対象物体を回転テーブルの上に置き、光源、カラーセンサ、レンジセンサ、偏光板を用意する。幾何データは回転テーブルを一定角度ずつ回転させながら取得する。

また、それぞれの回転角ごとにカラーセンサーにより光学データも同時に取得する。この時、二種類の光学データを得る。一つはカラーセンサと光源の前に偏光板を設置して撮像したデータ、もう一つは二枚の偏光板なしで撮像したデータである。

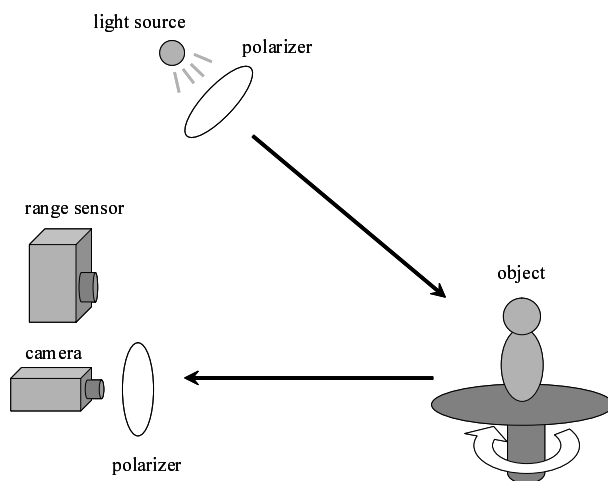


図 9.1: データ取得システム

ランプは蛍光ランプを用いている。本研究ではランプの大きさは小さく、対象物体から十分離れた位置に設置するので、点光源と仮定する。偏光を対象物体にあてるために、ランプの前に偏光板をおく。レンジセンサは Konica-Minolta VIVID 910, カラーセンサには Sony9000 video camera を用いている。

9.2 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三次元モデルに基づく見えの圧縮を行い、圧縮された情報を用いて仮想物体画像の合成を行った。本研究では図 9.2 のような対象物体を用いた。これらの物体は、第 1 章で紹介した関連

研究との比較において提案手法が優れていることを確認できるような物体を選んである。

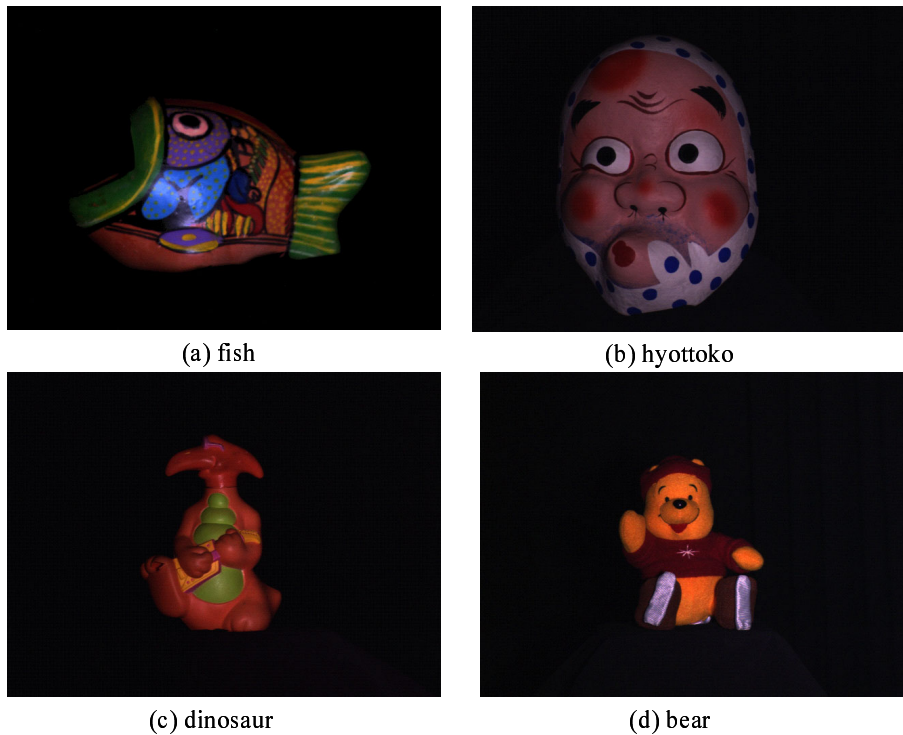


図 9.2: 対象物体

各物体は図 9.1 に示した撮像系において、視線方向を 12° ずつ回転させて全周囲の画像を撮像し、円偏光板あり・なし合わせて各 30 枚の画像を用意する (図 9.3)。

第 8 章で述べたシステムの流れのように、まず円偏光板あり・なし画像列から拡散反射成分画像列、鏡面反射成分画像列を作成する (図 9.4 上)。ただし対象物体 bear は鏡面反射成分を持たない。

各画像のサイズは 640×480 で、パッチ形式の画像に展開するさいは 1024×1024 の大きさの画像に各パッチ 16×16 のサイズとなるように正規化を行った。このパッチ形式画像一枚に格納できるパッチ数は $64 \times 64 = 4096$ だから、三次元幾何モデルの面数もそれ以下となるようにデータ処理している (図 9.4 下)。

次にレンジセンサーとカラーセンサーのカメラキャリブレーションを行い、そのカメラパラメータを基に Torrane-Sparrow モデルの拡散反射パラメータ、鏡面反射

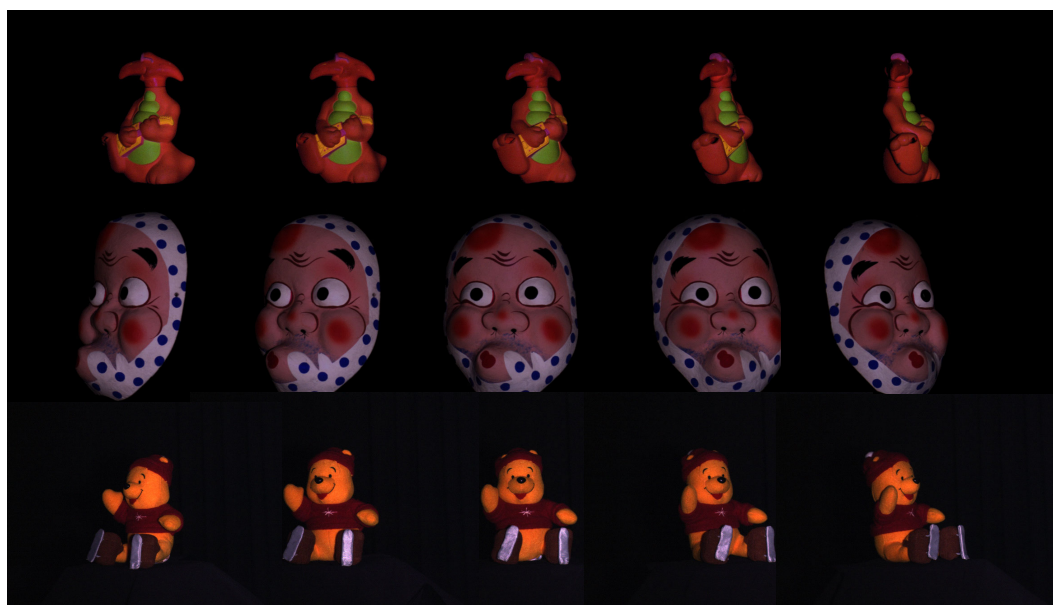
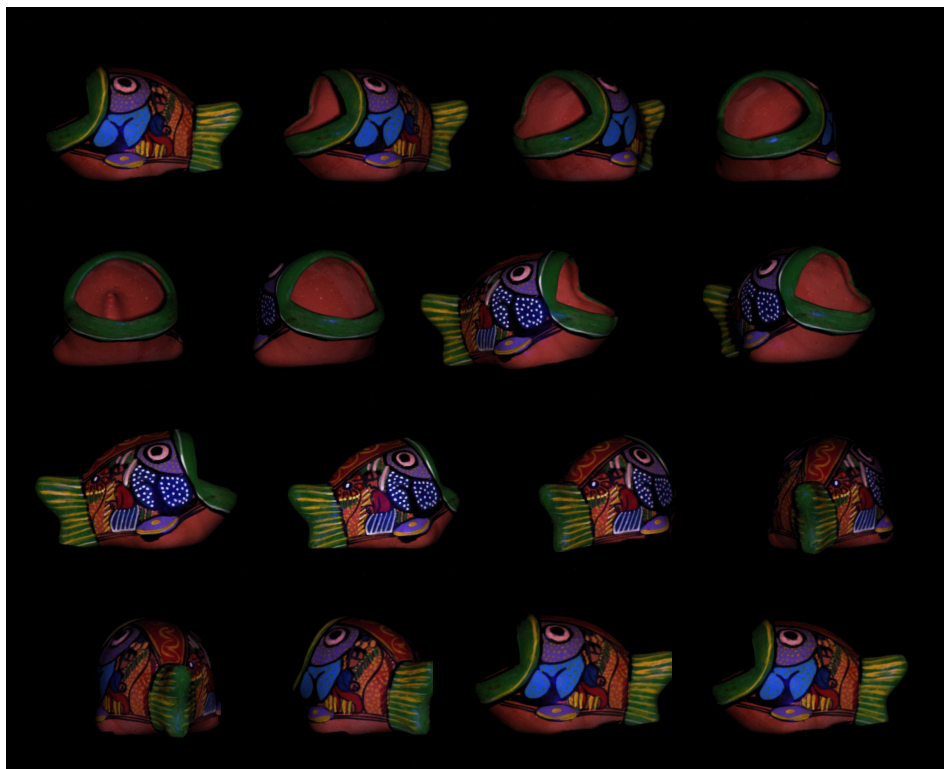


図 9.3: (上) fish の入力画像 (偏光板無) (下) dinosaur, hyottoko, bear の入力画像 (偏光板無)

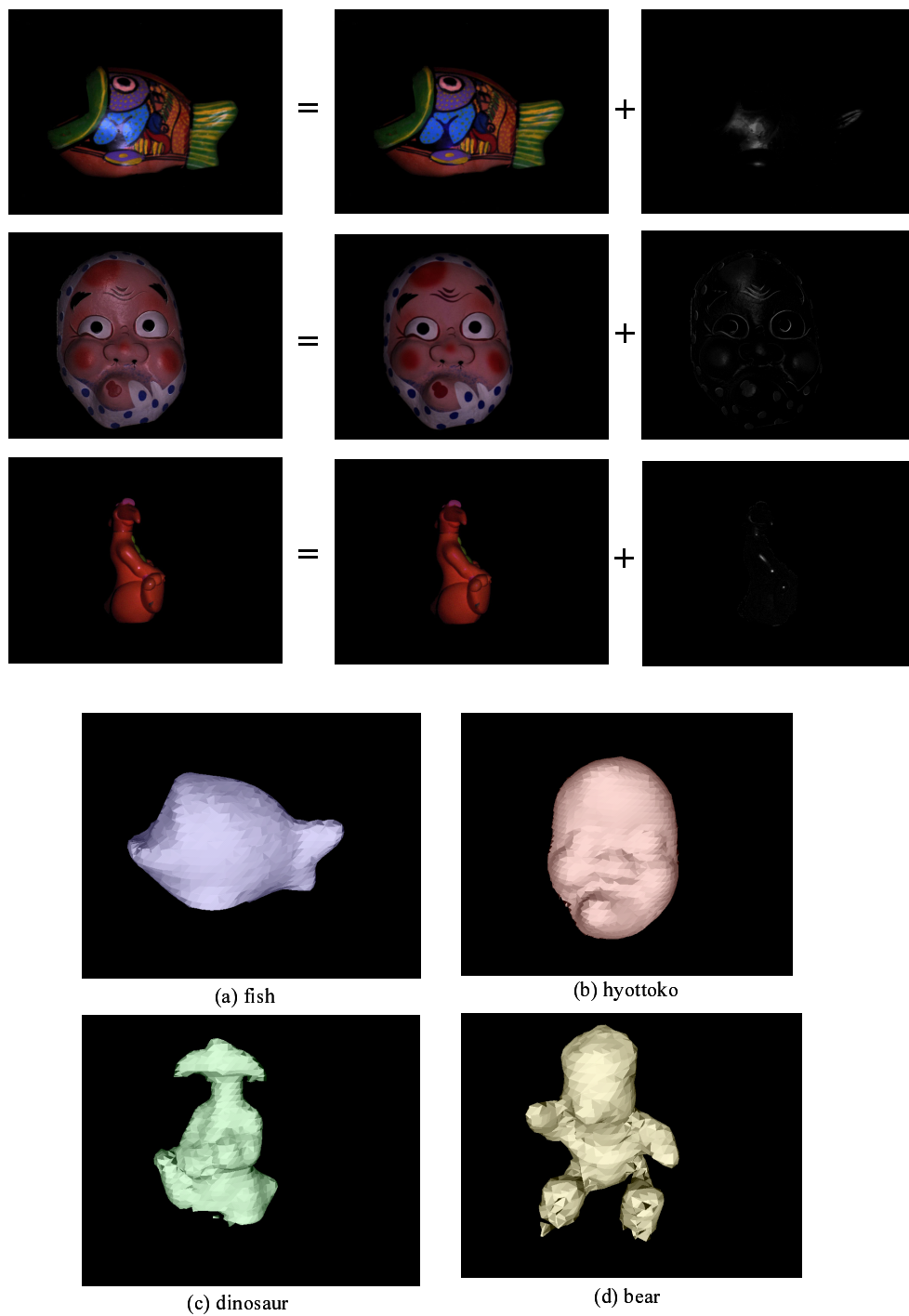


図 9.4: (上) 拡散反射成分画像・鏡面反射成分画像 (下) 三次元幾何モデル

パラメータを推定する (図 9.5) (図 9.6)。

これらの両反射パラメータから, 入力画像に対応した視点と光源条件のもとでパッチ形式画像にレンダリングを行う。また, もともとの拡散反射成分画像・鏡面反射成分画像も三次元モデルとカメラパラメータを用いてパッチ形式画像に展開し, 差分画像列を作る。図 9.7 は拡散反射成分についてこの差分を図示したものである。パッチ形式差分画像は絶対値をとって視覚化してあり, また差分を三次元モデルに貼り付けた図も示す。このパッチ形式差分画像を入力画像の枚数分用意する。鏡面反射成分についても同様のパッチ形式差分画像列を作成する。

以上の操作より得られた 30 枚分の 1024×1024 サイズの拡散反射成分差分画像列・鏡面反射成分画像列をパッチごとの $16 \times 16 \times 30$ の三次元配列に分割し, 三次元離散ウェーブレット変換を施しレベル 3 まで分解する。分解後の係数を昇順に並べ, 圧縮率に応じて下位数パーセントの係数を 0 として情報圧縮を行う。

次にレンダリング過程であるが, 圧縮されたパッチ形式差分画像列の係数から三次元離散ウェーブレット逆変換を施しレベル 3 まで再構成する。そして三次元モデル, 拡散反射パラメータ, 鏡面反射パラメータ, カメラパラメータと共に画像をレンダリングする。図 9.8 は提案手法によるレンダリング画像と Torrance-Sparrow モデルによるレンダリング画像, 図 9.9 は複数視線方向のレンダリング画像列である。

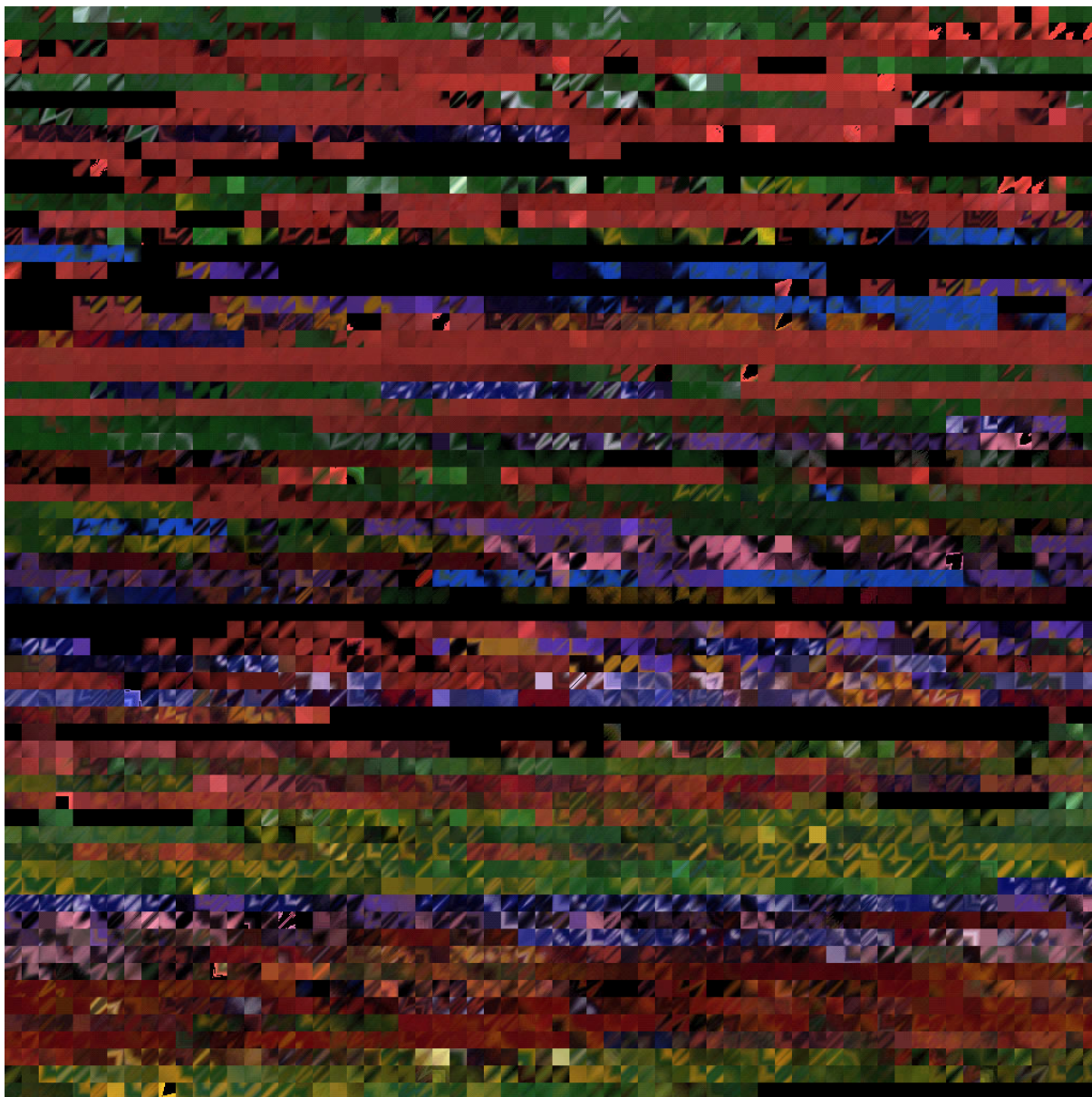


図 9.5: fish の拡散反射パラメータ

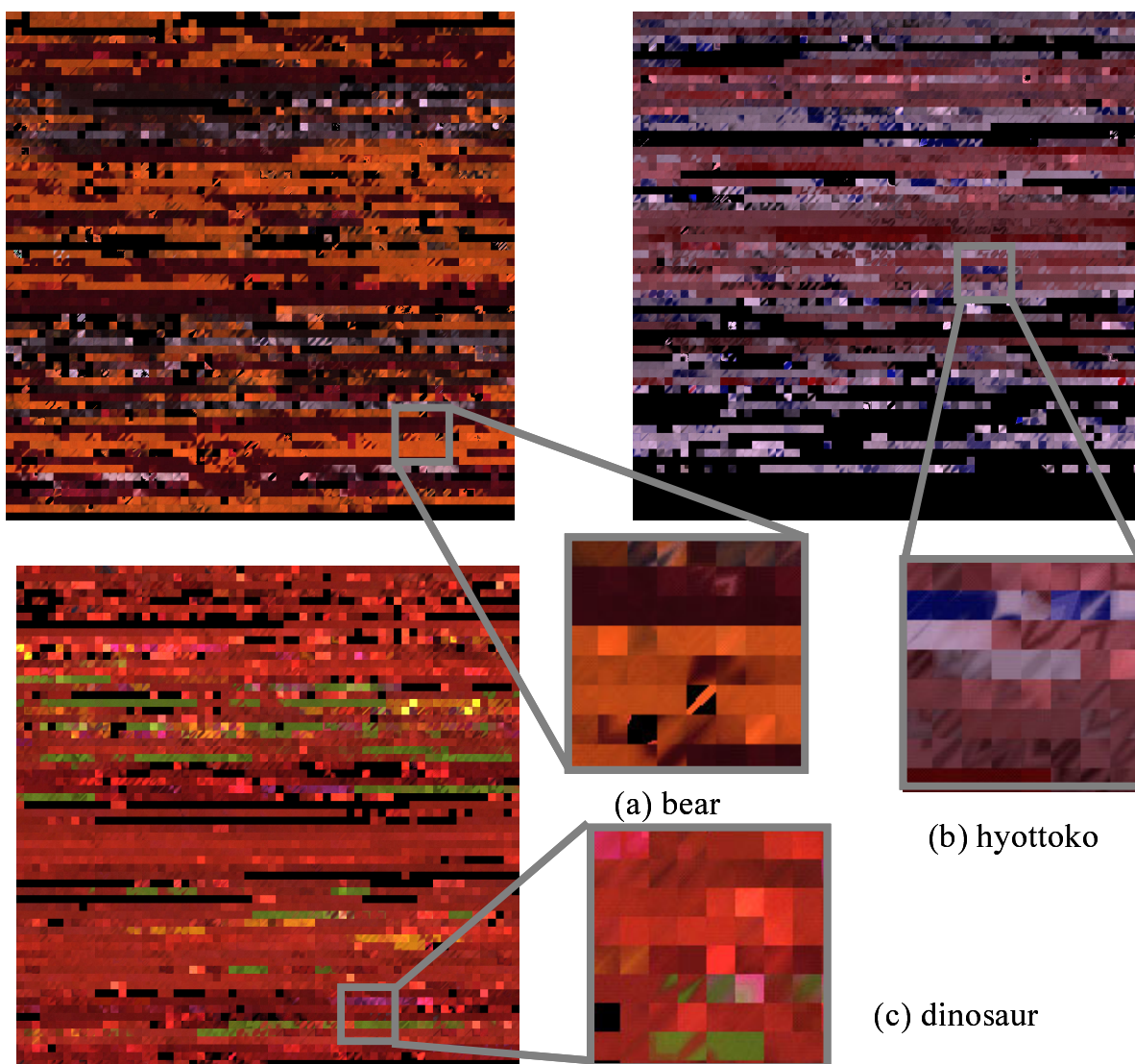


図 9.6: bear,hyottoko,dinosaur の拡散反射パラメータ

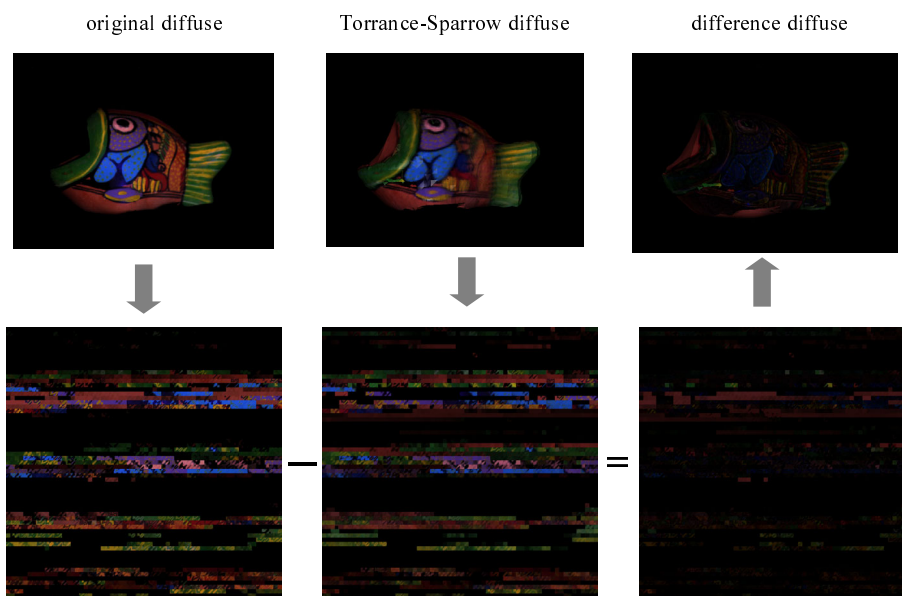


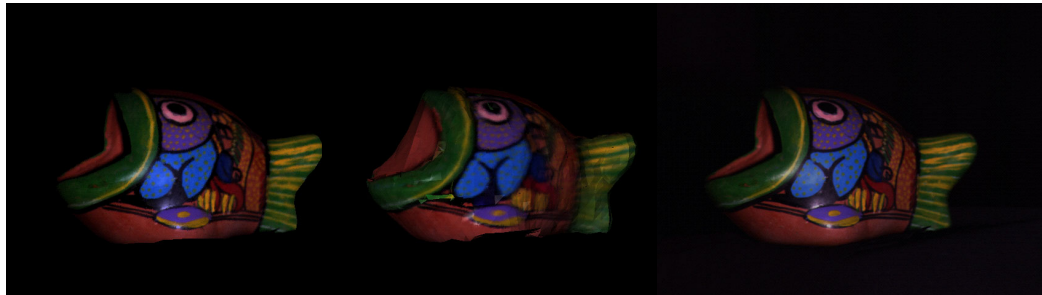
図 9.7: 拡散成分差分画像

9.3 評価

本節では提案手法によるレンダリング精度，ウェーブレット変換による情報圧縮のデータ量の評価を従来法との比較も含めた上で行う．ここでは従来法を，元のテクスチャ情報を用いて情報圧縮を行う手法と，元のテクスチャ情報を用いずに反射モデルによりレンダリングを行う手法に分ける．前者については Eigen-Texture 法，古川らの手法，Ma らの手法としてで，後者については佐藤らの手法，高橋らの手法としてで比較を行う．

9.3.1 データ量

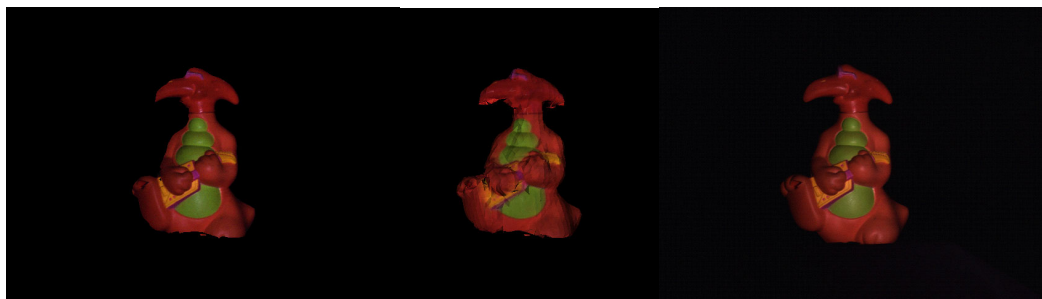
反射モデルだけでなく元のテクスチャ情報を用いることでレンダリング精度は大きく向上する．しかし多視点方向からのテクスチャ情報のデータ量は膨大である．そこで，膨大なテクスチャ情報をいかに効率よく圧縮するかが問題となってくる．本研究では反射モデルと元のテクスチャ情報とを組み合わせしており，圧縮においては単にテクスチャ情報を圧縮するのではなく，入力画像と反射モデルによるレンダリ



Rendering with difference interpolation (9%)

Torrance-Sparrow Rendering

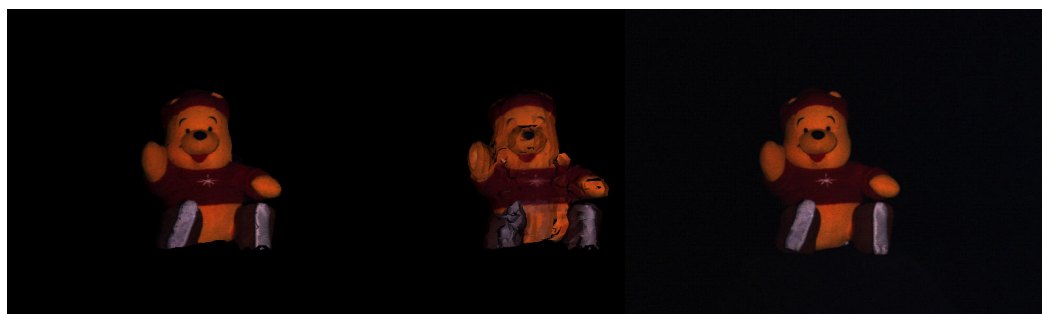
original



Rendering with difference interpolation (9%)

Torrance-Sparrow Rendering

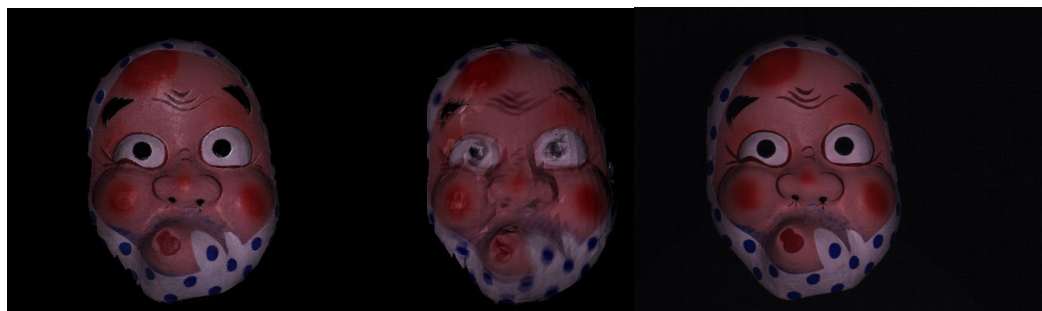
original



Rendering with difference interpolation (3.6%)

Torrance-Sparrow Rendering

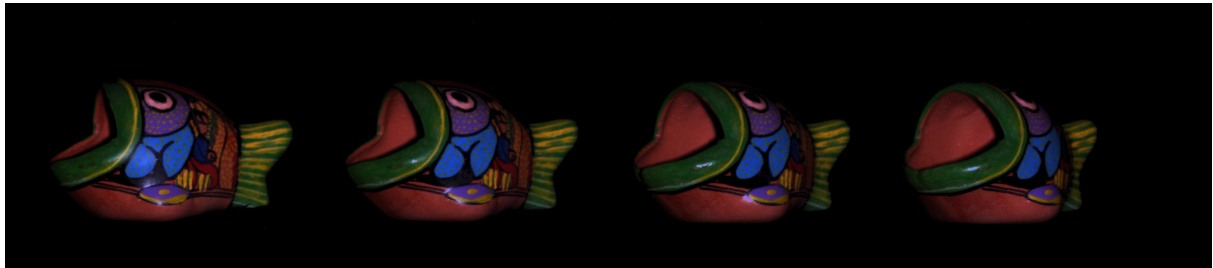
original



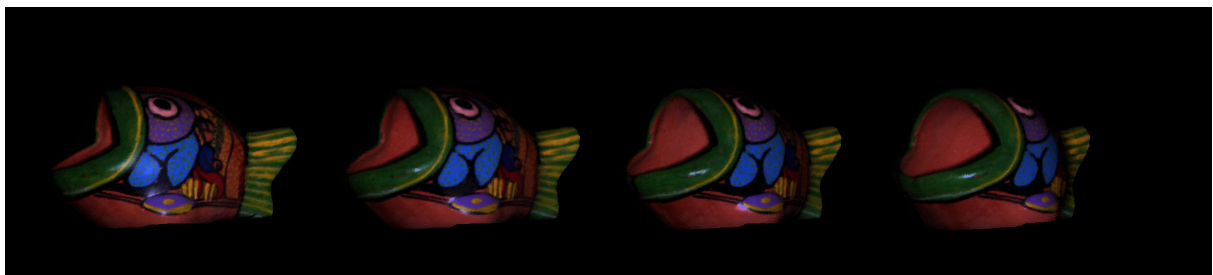
Rendering with difference interpolation (2%)

Torrance-Sparrow Rendering

original



original



Rendering with Difference Interpolation(1.5%)

図 9.9: fish の結果画像列

ング画像との差分成分を圧縮している．このことにより反射モデルで表現できる部分については圧縮対象のデータから除外することができるのでより効率的なデータ量削減ができると考えられる．そこで本節ではデータ量の観点から評価を行う．

9.3.1.1 ウェーブレットによる圧縮率と誤差

本研究ではデータ量を小さく保ちつつ，光学反射モデルによるレンダリングでは表現しきれない部分についても再現するために，入力画像と光学反射モデルによるレンダリング画像との差分成分を三次元離散ウェーブレット変換による情報量の圧縮を行っている．第 8 章で説明したように本システムでは三次元幾何モデルに対応する三角パッチごとに圧縮をするが，これは各パッチごとに視線方向ごとの相関が違っていたり，ハイライトの有無や self shadow があったりするから圧縮効率は三角パッチによるからである．圧縮比率は自由に決めることができ，誤差とのトレードオフとなる．三次元離散ウェーブレット変換により画像列を段階的に分解して展開係数とする．それぞれの対象物体について，拡散反射の差分成分と鏡面反射の差分

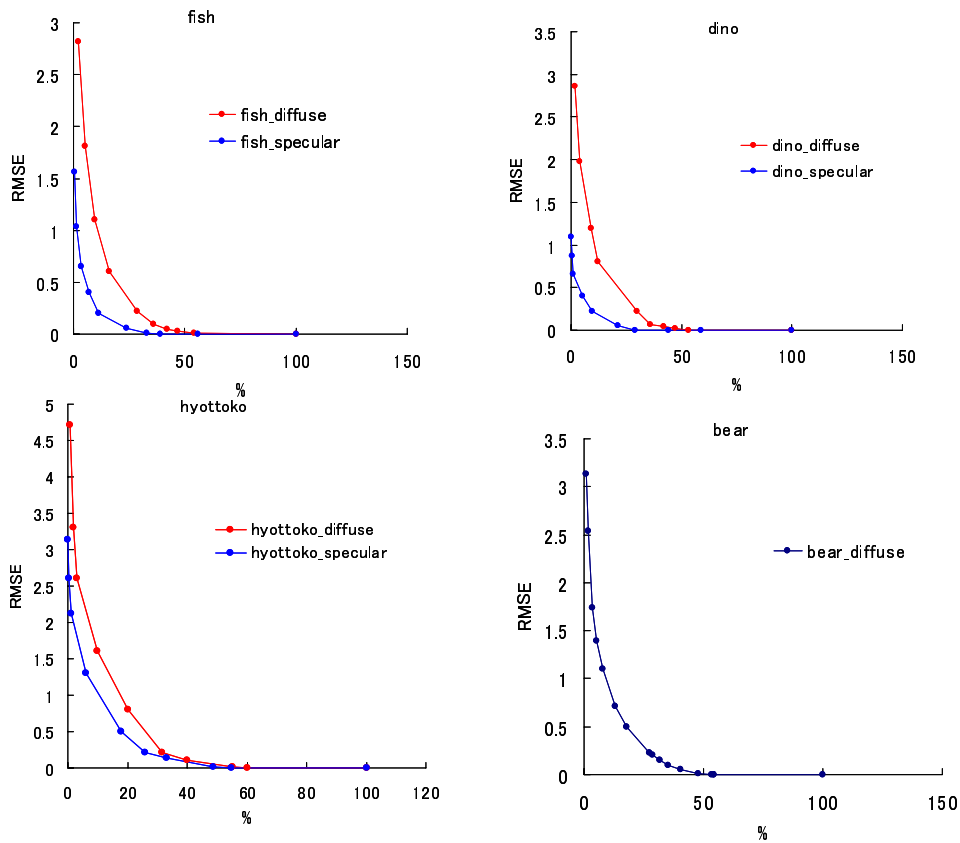


図 9.10: 展開係数の使用比率と再構成誤差

成分の係数の使用比率と画素ごとの誤差 (RMSE:Root Mean Square Error) との相関関係を図 9.10 に示す。なお、対象物体 bear は鏡面反射成分を持たない。

対象物体 hyottoko, bear, dinosaur について、同一圧縮率では拡散反射の差分成分のほうが鏡面反射の差分成分より誤差が大きくなっている。これは、鏡面反射成分は入射光を直接反射している成分であり、光源の単色となるのに比べて、拡散反射成分は入射光が反射表面から物体内部に透過し、内部で透過・拡散を繰り返した後に放射された光であり、物体の色を表す成分であるから今回のような様々な色を持つ物体では複雑な輝度値となることに起因する。全ての対象物体について、係数の使用比率が 50 % 以上の場合には誤差がほぼ 0 となることがわかる。係数の使用比率

が 50 % 以下の場合には圧縮率が 0 に近づいていくに従って、誤差がなだらかに大きくなってゆき、25 % のあたりを境に急峻になる。

次に再構成信号の誤差が提案手法においてどのように左右してくるかを、本システムによるレンダリング画像として視覚的に捉えることによって検証する。図 9.11 は対象物体 hyottoko の圧縮率ごとのレンダリング画像である。

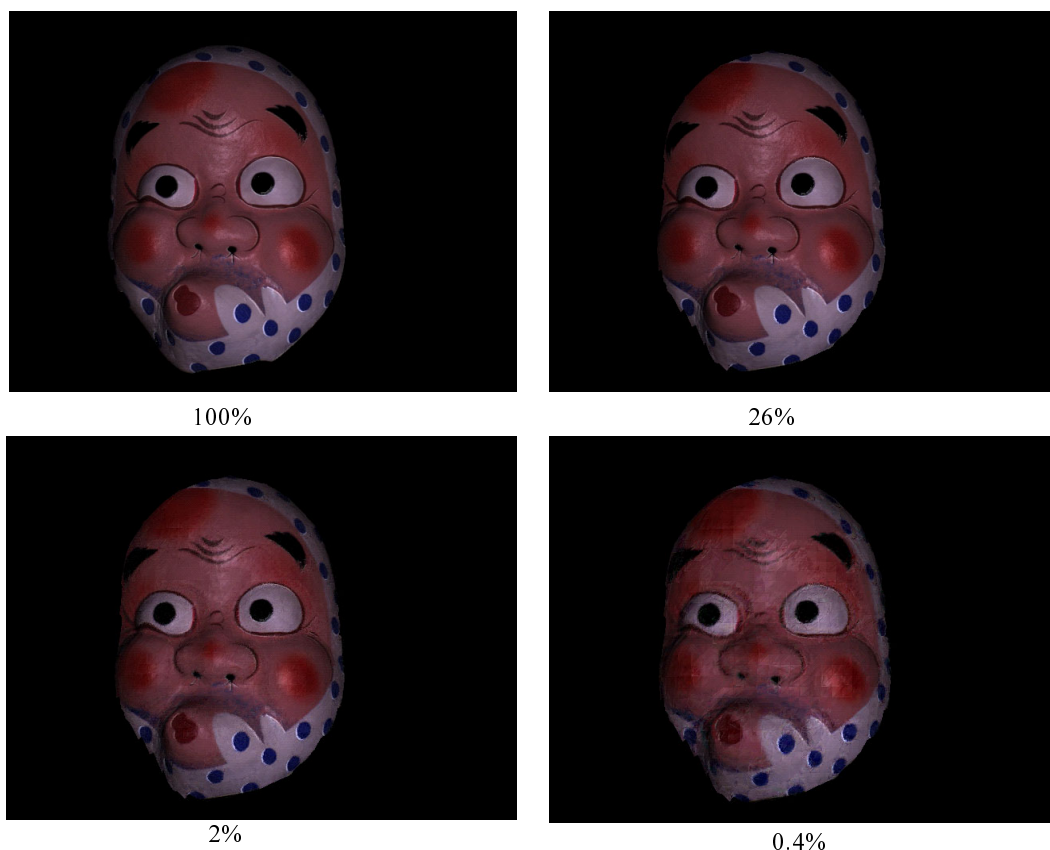


図 9.11: hyottoko 展開係数の使用比率ごとの再構成画像

100 % というのは無圧縮，つまり入力画像である．誤差と係数の使用比率の相関図より 50 % 程度までが誤差が非常に 0 に近い場合であったが，26 % の場合についても見た目には劣化が確認されない．前述したように 26 % とは誤差の曲線の傾きが大きく変化するあたりである．この場合の誤差は図 9.10 より拡散反射成分，鏡面反射成分共に 1 以下である．

使用比率を下げていき，2 % の場合の画像になると誤差が視覚的に確認できるようにはなるが，2 % という使用比率を考えると十分な品質であるといえる．さらに使用比率を下げた 0.4 % の場合の画像は拡散・鏡面反射成分の誤差が 2 ~ 3 となり，三次元幾何モデルの三角パッチがわかるようになり，品質的に問題がある．

9.3.1.2 従来法との比較: 西野らの手法，古川らの手法、Ma らの手法

ここでは提案手法を画像圧縮法という観点で捉え，西野らによる Eigen-Texture 法 [6] との比較を行う．Eigen-Texture 法とは三次元幾何モデルを用いて入力画像をパッチ形式画像に展開して主成分分析 PCA によりパッチごとに圧縮を行う手法である．なお提案手法との比較にあたり，Eigen-Texture 法の変換を PCA ではなく現在最も画像の再構成に優れているといわれるドベシウウェーブレットを使った．図 9.12 に同一圧縮率における西野らの手法と提案手法のレンダリング画像を示す．なお，Eigen-Texture 法は反射成分の分離を行わずに，拡散反射成分と鏡面反射成分を併せて扱っているが，更に反射成分分離を行って反射成分ごとに Eigen-Texture 法を行ったレンダリング画像も提案手法との比較としてあげてある．それぞれの誤差と展開係数使用比率の関係を表 9.1 に示す．また図 9.12 と表 9.1 だけではわかりづらい誤差を図 9.13 の差分画像として視覚化した．

表 9.1: 西野らの手法と提案手法における RMSE と展開係数使用比率の関係

		Eigen-Texture (not separation)		Eigen-Texture (separation)		difference interpolation
2 %	R	5.581256	R	4.449565	R	3.347976
	G	5.455502	G	4.282305	G	3.223461
	B	5.530034	B	4.410084	B	3.322767
26 %	R	1.234328	R	1.204168	R	0.710579
	G	1.293849	G	1.202791	G	0.71287
	B	1.239905	B	1.205595	B	0.710267

Eigen-Texture 法において，反射成分を分離した場合と分離しない場合では，分離したほうが誤差が小さくなることがわかる．これは第二章で説明したように，拡散

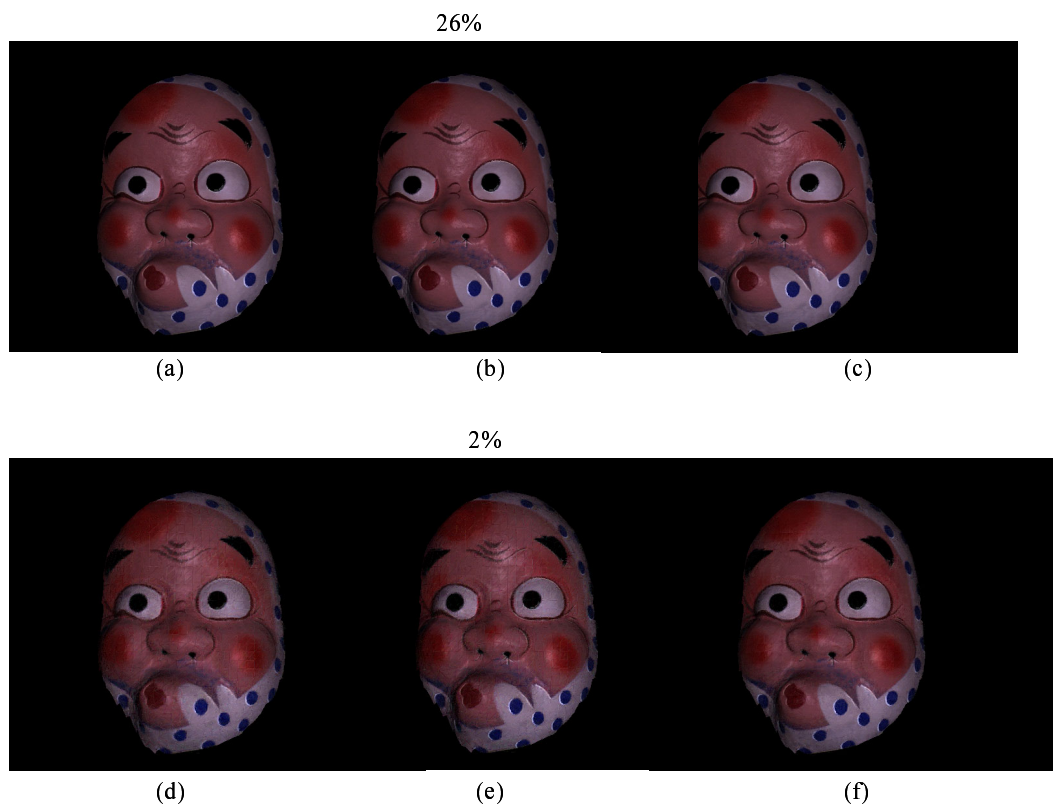


図 9.12: Eigen-Texture 法との比較:(a),(d) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分非分離), (b),(e) はそれぞれ係数使用比率 26 %,2 %の Eigen-Texture 法 (反射成分分離) (c),(f) はそれぞれ係数使用比率 26 %,2 %の提案手法

反射成分は視線方向によらず光源方向と法線方向に影響されるが、鏡面反射成分は視線方向の変化に非常に敏感であり、輝度値が大きく変化するという性質を持っており、その両者の性質は大きく異なることによると考えられる。データ量の観点では、反射成分を分離することにより画像が二枚になり大きくなる。

次に西野らの手法と提案手法を比べると、誤差の面では提案手法が優れていることがわかる。これは提案手法が現時点では最も画像再構成精度が良いと考えられるウェーブレット変換を使っていることに加えて、Eigen-Texture 法が画像そのままをパッチ形式に展開して圧縮を行っているのに対して、提案手法では画像をパッチ形式と Torrance-Sparrow モデルによるパッチ形式レンダリング結果との差分成分を圧

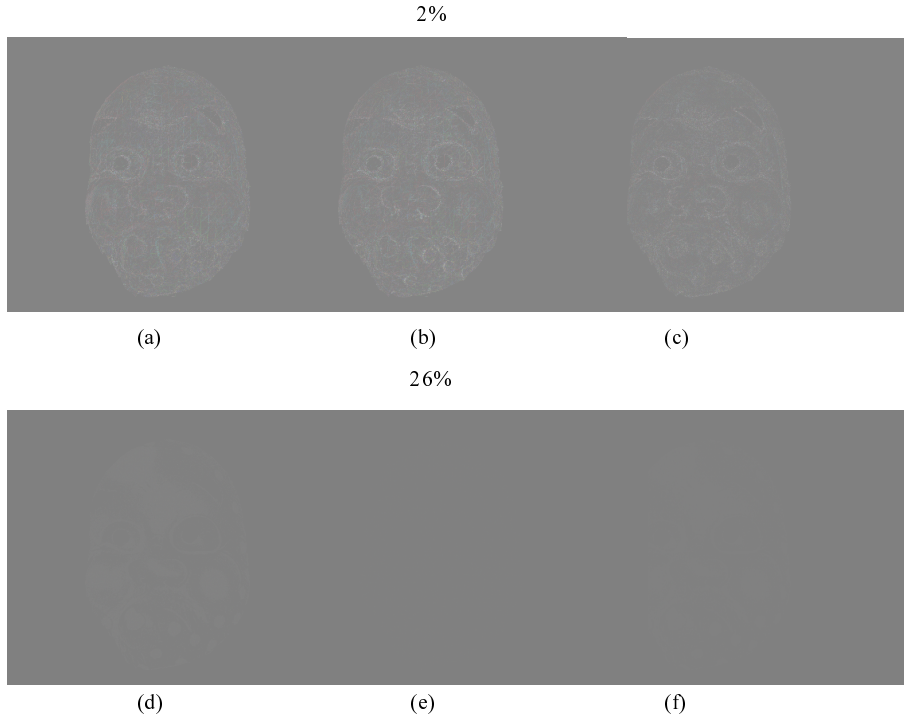


図 9.13: Eigen-Texture 法との比較 (差分画像) : (a), (d) はそれぞれ係数使用比率 26 %, 2 % の Eigen-Texture 法 (反射成分非分離), (b), (e) はそれぞれ係数使用比率 26 %, 2 % の Eigen-Texture 法 (反射成分分離) (c), (f) はそれぞれ係数使用比率 26 %, 2 % の提案手法

縮していることに起因する。またデータ量の観点でも差分成分を圧縮するということが効いてくる。

また、全周囲のレンダリング画像列全体を動画として捉えた場合、二色性反射モデルを考慮していない Eigen-Texture 法は視線方向に敏感な鏡面反射成分が画像ごとに滑らかではなく不連続で、ジャンプしてしまう。一方提案手法は鏡面反射成分を視線方向を含む Torrance-Sparrow モデルとして表現しているため、視線方向の移動に対しても滑らかな動画となる。

古川らの手法 [7] は三次元幾何モデルとテクスチャ画像列を用いて様々な状況下でのレンダリングを行う。様々な状況下とは視線方向・光源方向共に二次元の自由度を持つ状況であり、これらすべての状況で撮像したテクスチャ画像を保持するには

データ量が非常に大きくなってしまいますので、この画像列をテンソル積展開により圧縮を行う。しかし提案手法で用いているウェーブレットと変換による圧縮はテンソル積展開による圧縮より効率よく圧縮できる。また前項で述べたように反射モデルの導入により古川らの手法に比べて提案手法は動画としたときにも滑らかに反射成分を表現できる。一方本研究では光源については固定、視線方向については次元の自由度で実験を行っているが、光源方向と視線方向の自由度の増加にも容易に対応できる。

Ma らの手法 [14] は物体のテクスチャ画像列をラプラス変換により多重解像度解析し、さらにそれらの段階別の展開係数を主成分分析して固有ベクトルとして圧縮を行う。この際に多重解像度解析を行った後の係数を Phong モデルとして近似する。レンダリングにおいては Phong モデルによる係数から段階的にラプラス変換で再構成する。

まず用いる反射モデルだが、Ma らは Phong モデルを用いているのに対し、提案手法では Torrance-Sparrow モデルを用いている。第二章で述べたように Phong モデルは非常に簡潔な表現であり、厳密な鏡面反射成分の再現には適していない。Torrance-Sparrow モデルも同じく現実の反射を簡略化したものであるが鏡面反射成分の再現において Phong モデルより優れていることが知られている。

また圧縮についてだが Ma らはラプラス変換を用いて多重解像度解析を行っているのに対し、提案手法ではより画像再構成精度の高いウェーブレット変換による多重解像度解析を行っている。以上よりデータ量とレンダリング精度では提案手法が優れていることがわかる。しかし計算速度の観点から見ると、ウェーブレット変換はラプラス変換より計算コストが高く、また Torrance-Sparrow モデルは Phong モデルより複雑な表現であるため、結果としてレンダリング速度では W.Ma らの手法が優れている。

9.3.2 レンダリング精度

データ量の観点からは元のテクスチャ情報を用いないほうが優れているのは当然であるが、反射モデルだけでのレンダリングでは十分な精度を得られなかったり、制約条件があったりする。そこで本研究では反射モデルと元のテクスチャ情報とを組み合わせている。その中でウェーブレットによる差分成分の効率的な圧縮を行うことによりレンダリング精度とデータ量の問題を共に向上させている。そこで本節ではレンダリング精度の観点から評価を行う。

9.3.2.1 圧縮率ごとのレンダリング精度比較

レンダリング画像の精度を検証するため、入力画像とレンダリング画像の比較を行った。図 9.14 は赤い線を切断面とした、それぞれの画像の RGB 成分ごとの輝度の比較を表している。なお、横軸は 0 から大きくなるにしたがって赤い線の上端から下端への推移、縦軸は輝度値とする。

グラフは入力画像と展開係数使用比率 1%、12%、31% のレンダリング画像の比較である。局所的に見ると 1% の場合は誤差が大きくなることも多々あるが、大局的に見ると概ね同じ曲線上にあるといえる。31% の場合はかなり誤差が小さいことがわかる。12% においてもグラフから明らかな誤差はほとんど確認できない。前述したように圧縮率と誤差のトレードオフで画質を決めることになるが、展開係数が元の 1% 以上あれば明らかな劣化画像とは見受けられない。

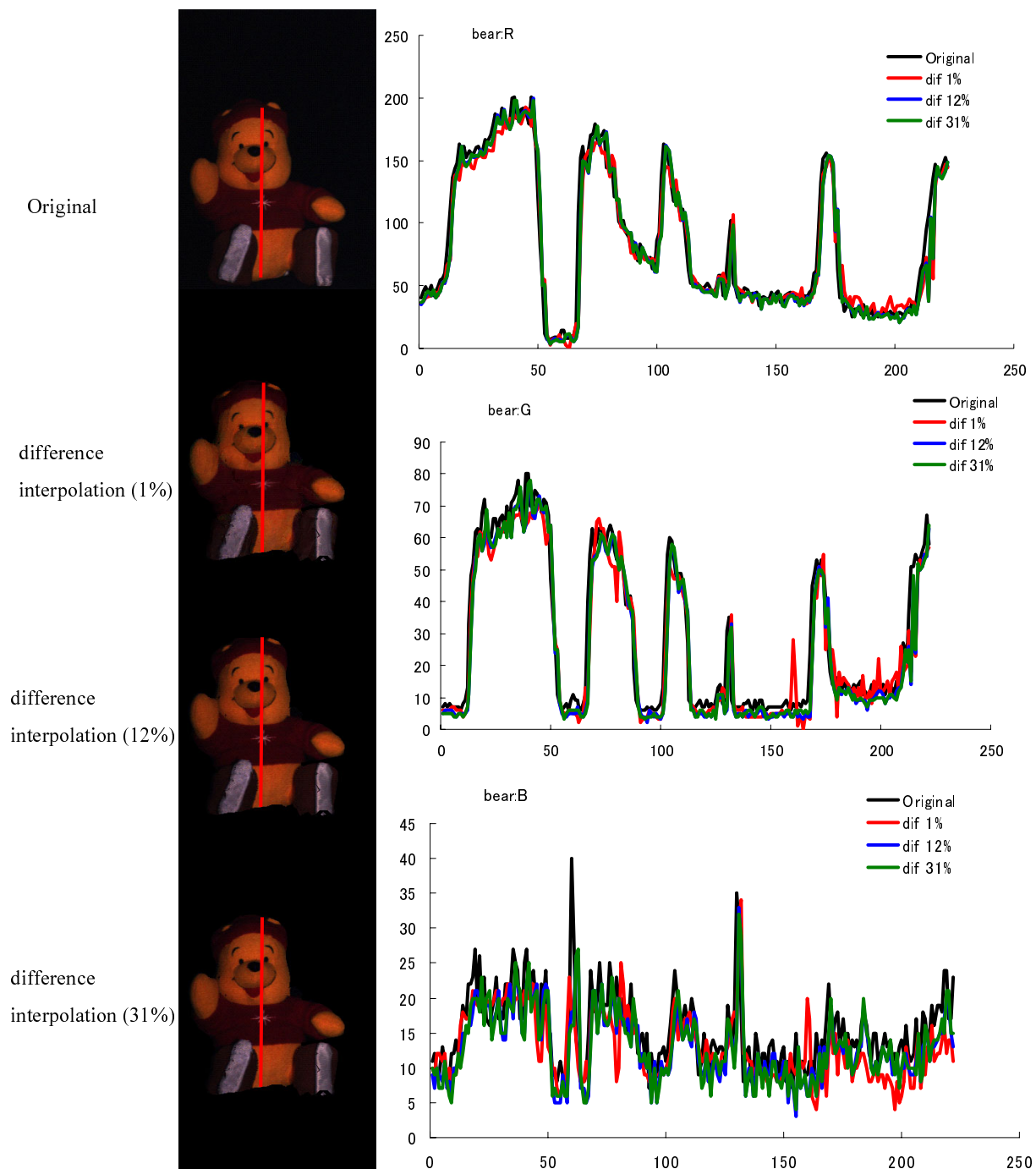


図 9.14: bear の切断面における輝度比較 (ウェーブレット圧縮率別)

9.3.2.2 従来法との比較: 佐藤ら, 高橋らの手法

ここでは提案手法が広範で有効であることを示す。佐藤らの手法 [12] では二色性反射モデルを用いてレンダリングを行っている。そのための反射成分分離は輝度値によって行う。しかし, 対象物体 *hyottoko* のように光源色と物体表面色が同じ場合は反射成分分離がうまくいかない。一方提案手法では第 3 章で説明したように反射成分の分離を偏光の性質を利用して行っているため光源色や物体表面色の拘束を受けない。このことを対象物体 *hyottoko* を使って示す。図 9.15 赤い線を切断面とした, それぞれの画像の RGB 成分ごとの輝度の比較を表している。ここでは併せて Torrance-Sparrow モデルによるレンダリング画像との比較も行っている。なお, 横軸は 0 から大きくなるにしたがって赤い線の左端から右端への推移, 縦軸は輝度値とする。

図 9.15 より提案手法によるレンダリング画像は入力画像と比べて細かい範囲では誤差は含むものに, RGB 全てにおいて概ね近い値となることがわかる。比較として掲載した Torrance-Sparrow モデルによるレンダリング画像が三次元幾何モデルの三角パッチの影響により入力画像からのずれが大きくなるのに対して精度が大きく向上しているところがわかる。以上より, 提案手法が光源色と物体表面色に拘束を受けないことがわかる。

また, 高橋らの手法 [13] では直線偏光により分離された反射成分より Torrance-Sparrow モデルの反射パラメータを推定してレンダリングを行っている。第 2 章で説明したように Torrance-Sparrow モデルは現実の反射を簡略化して数式で表現したものであるため様々な制約があり, 相互反射がおこる物体や表面がふさふさした物体には適さない。つまり, 対象物体 *bear* のような布地の物体や, 対象物体 *dinosaur* のような複雑な形状で相互反射が起こる物体には不適切である。一方, 提案手法は Torrance-Sparrow モデルによるレンダリングに加えて入力画像との差分成分を補完してレンダリングを行っているため, 光学反射モデルで表現しきれない成分ももらすことなく再現できる。ここでは対象物体 *bear* を用いて各レンダリング画像と入力画像の誤差を解析することによって検証する。

図 9.16 は赤い線を切断面とした, それぞれの画像の RGB 成分ごとの輝度の比較を表している。なお, 横軸は 0 から大きくなるにしたがって赤い線の上端から下端への推移, 縦軸は輝度値とする。

レンダリング画像を見ても一目瞭然だが, RGB ごとのグラフ全てにおいて提案手法は入力画像の良い近似であると言える。Torrance-Sparrow レンダリング画像の輝

度値は入力画像の輝度値から著しく外れて誤差が大きくなる部分があることがわかる。

さらに、この場合のようにレンジセンサーの問題で三次元幾何モデルの精度が悪い場合でも提案手法の有効性がわかる。また、高橋らの手法では反射成分分離を直線偏光板により行っているが、第三章で説明したように、偏光板の透過軸を適切に設定する必要があることと、視点が 90° に近いときに不具合が発生するなどの問題があった。これらの問題は前述したように偏光板による分離で解決できるので、反射成分分離においても提案手法が優れていることがわかる。

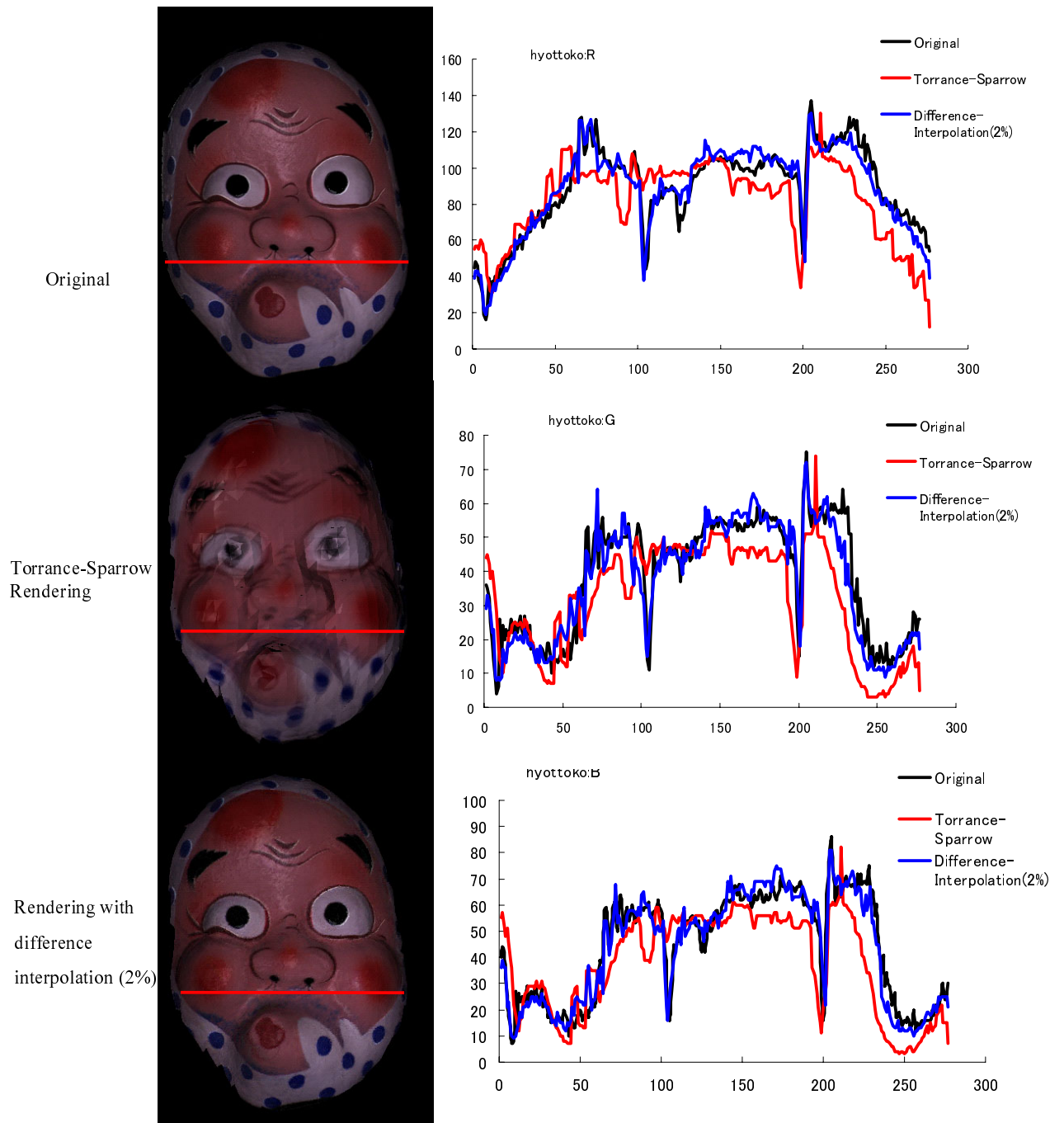


図 9.15: hyottoko の切断面における輝度比較

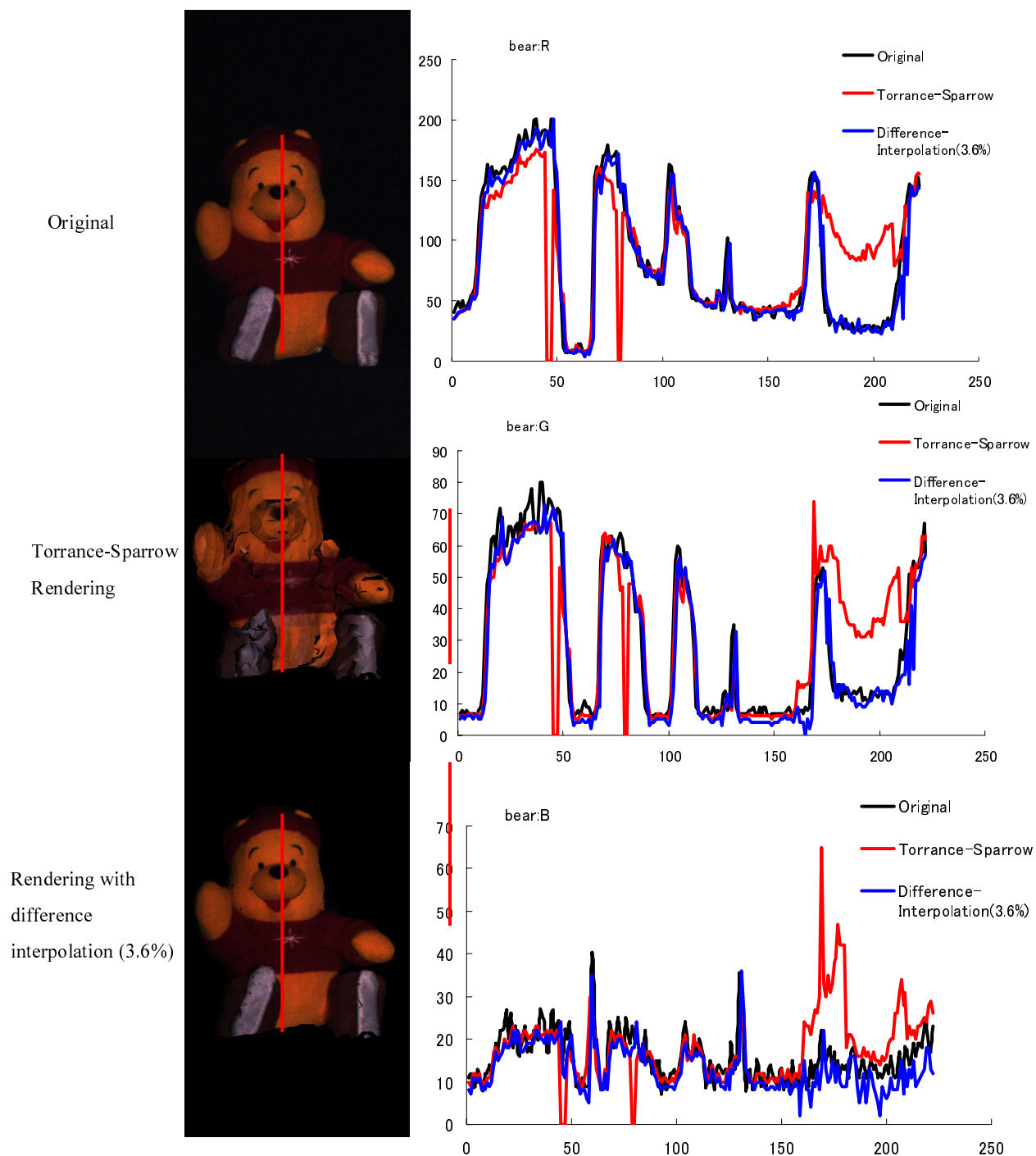


図 9.16: bear の切断面における輝度比較 bear の切断面 (レンダリング手法別)

第10章

結論

10.1 結論

本論分では複合現実感における重要な要素のうち、明るさ情報に着目し、様々な状況下での写実的なレンダリング方法を提案した。この手法において核となるのは二色性反射モデルの利用，円偏光による反射成分分離，差分画像の補完，ウェーブレットによるデータ圧縮である。提案手法により写実性において重要な表面反射が正確に再現でき，また大規模物体を対象とした場合に膨大となるデータ量の効率的な圧縮が可能となった。

仮想物体画像の合成手法は様々なものが提案されており，大まかに分類すると反射モデルを用いる MBR と複数枚の入力画像から IBR に分けられる。本研究では写実的な表面反射を再現するため，表面反射を数学的に定式化して扱う反射モデルを用いている。反射成分は拡散反射成分と鏡面反射成分という性質の異なる二つの成分に分離でき，これらを別々に扱うことが重要である。この反射成分の分離において本研究では円偏光の資質を利用したロバストな分離方法を提案しており，分離した成分それぞれについて反射モデルの数式の反射パラメータを推定してレンダリングを行っている。しかし反射モデルを適応できる物体には制限があり，反射モデルでは十分に表現できない物体についてのレンダリングに問題がある。つまり実際の画像と反射モデルによるレンダリング画像には差がある。そこで本研究では入力画像と反射モデルによるレンダリング画像の差分成分を補完して画像合成を行っている。しかしこの差分画像を保持することでデータ量が大きくなってしまっているので圧縮の必要性が出てくる。本研究ではウェーブレット変換による圧縮を行っているが，前述したように二つの反射成分の性質は異なるため，差分成分も成分別に扱うことが重要である。さらに幾何モデルに対応する三角パッチごとについてもそれぞれ固有の圧縮率で圧縮してやることで効果的な圧縮ができる。そこで提案手法では成分別の差分成分を三角パッチごとに並べた三次元配列を三次元ウェーブレット変換により圧縮してやることにより三次元の相関性を考慮した圧縮を行っている。このレンダリング手法により MBR の欠点を補うことができ，さらにデータ量の問題も解決されている。

第 9 章では圧縮率を変えて実験を行い，そのレンダリング精度とデータ量の評価を行った。圧縮率が 20 % 程度までは誤差は大して大きくなり，高精細なレンダリングが実現できる。さらに圧縮率を 2 % 程度まで上げても視覚的には十分な精度の合成画像を得ることができることが確認できた。

また従来手法との比較では、データ量の観点では IBR の手法との比較を行い、レンダリング精度の観点では MBR の手法との比較を行った。提案手法は現在最も画像再構成精度に優れているウェーブレットをベースにした圧縮を行っており、PCA、ラプラス変換等を用いている方法より効果的な圧縮が実現できていると考えられる。また反射モデルをもちいたレンダリング手法との比較では反射モデルでは扱えない物体についてのレンダリングを行えることを示し、提案手法の有効性を確認した。

10.2 今後の課題

提案手法は効果的なデータ圧縮と写実的なレンダリングを実現したが、計算速度の面では十分とはいえない。これはウェーブレット変換自体の計算量に起因する問題であり、リアルタイムでのレンダリングを実現するという改善点がある。提案手法は高い圧縮率を実現しており、今後益々需要が増える携帯機器向けの三次元コンテンツにも有効であると考えられる。さらに、本研究では光源方向は固定、視線方向の自由度が 1 という条件で実験を行ったが、様々な状況でのレンダリングは光源方向の自由度 2、視線方向の自由度 2 で行う必要がある。しかし提案手法はこの状況でも対応できるものであり、本研究の圧縮方法を用いることでデータ量を小さく抑えたレンダリングができる。

謝辞

本研究を進めるにあたり，忙しい御身にも関わらず，毎週ミーティングを行なっていたいただいて的確なアドバイスをしてくださり，またすばらしい研究環境を整えてくださった池内克史教授には心から御礼申し上げます。

また研究を進める上で数々の助言をいただいた宮崎大輔さんには大変お世話になりました。研究のアイデア，疑問点，論文の書き方，プログラム等，広範にわたって多くのことを御指導いただき，ここに修士論文として研究成果をまとめることができました。さらに同じ光学グループのメンバーであり，いつも理解の至らない私の質問を快く受けてくださり，活発に議論して頂いた猪狩壮文さん，川上玲さんにもこの場を借りて深く御礼申し上げます。より身近な存在として親身に相談に乗っていただいたき，様々な面で助けて頂きましてありがとうございます。

また同じ時期に研究室に入り，二年間を共にした岡本泰英さん，松井健さん，李暁路さん，鎌倉真音さんには日常のことから研究のことまで，多くの面でお世話になりました。

最後に，これまで様々な面から支援していただいた家族と友人たちに，心から感謝します。

平成 18 年 2 月 3 日

参考文献

- [1] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen: “The lumigraph,” Computer Graphics Proceedings, ACM SIGGRAPH’96, pp. 43–54 (1996).
- [2] M. Levoy and P. Hanrahan: “Light field rendering,” Computer Graphics Proceedings, ACM SIGGRAPH’96, pp. 31–42 (1996).
- [3] D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. H. Salesin and W. Stuetzle: “Surface light fields for 3d photography,” Computer Graphics Proceedings, ACM SIGGRAPH’00 (2000).
- [4] G. Miller, S. Rubin and D. Ponceleon: “Lazy decompression of surface light fields for precomputed global illumination,” Rendering Techniques (Eurographics ’98 Proceedings), pp. 281–292 (1998).
- [5] W. C. Chen, R. Grzeszczuk and J. Y. Bouguet: “Light field mapping: Hardware-accelerated visualization of surface light fields,” ACM SIGGRAPH 2002 (2002).
- [6] 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, pp. 1257–1265 (2001).
- [7] R. Furukawa, H. Kawasaki, K. Ikeuchi and M. Sakauchi: “Appearance based object modeling using texture database: Acquisition compression and rendering,” Proceedings of Eurographics Workshop on Rendering, pp. 257–266 (2002).

- [8] K. Dana, B. van Ginneken, S. Nayar and J. Koenderink: “Reflectance and texture of real-world surfaces,” *ACM Transactions on Graphics* 18, pp. 1–34 (1999).
- [9] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin and M. Sagar: “Acquiring the reflectance field of a human face,” *Computer Graphics Proceedings, ACM SIGGRAPH’2000* (2000).
- [10] P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese and T. Hawkins: “A lighting reproduction approach to live-action compositing,” *Computer Graphics Proceedings, ACM SIGGRAPH’2002* (2002).
- [11] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins and P. Debevec: “A lighting reproduction approach to live-action compositing,” *Computer Graphics Proceedings, ACM SIGGRAPH’2005* (2005).
- [12] Y. Sato, M. D. Wheeler and K. Ikeuchi: “Object shape and reflectance modeling from observation,” *Computer Graphics Proceedings, ACM SIGGRAPH’97*, pp. 379–387 (1997).
- [13] T. Takahashi, Y. Sato and K. Ikeuchi: “Separating the reflection components with the use of polarization and determining the reflection parameters,” *CVIM-2000-124-3*, pp. 17–24 (2000).
- [14] W.-C. Ma, S.-H. Chao, Y.-Y. Chuang, C.-F. Chang, B.-Y. Chen and M. Ouhyoung: “Level-of-detail representation of bidirectional texture functions for real-time rendering,” *ACM SIGGRAPH 2005 Symposium on Interactive 3D Graphics and Games*, pp. 187–194 (2005).
- [15] S. Shafer: “Using color to separate reflection components,” *Color Research and Applications*, **10**, pp. 210–218 (1985).
- [16] K. E. Torrance and E. M. Sparrow: “Theory of off-specular reflection from roughened surfaces,” *Journal of the Optical Society of America*, **57**, pp. 1105–1114 (1967).

- [17] 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 (1991).
- [18] L. B. Wolff and T. E. Boult: “Constraining object features using a polarization reflectance model,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **13**, 7, pp. 635–657 (1991).
- [19] T. E. Boult and L. B. Wolff: “Physically based edge labeling,” *Proceedings of Computer Vision and Pattern Recognition*, pp. 656–663 (1991).
- [20] Y. Y. Schechner, J. Shamir and N. Kiryati: “Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface,” *Proceedings of International Conference on Computer Vision '99* (1999).
- [21] P. J. Besl and N. D. McKay: “A method for registration of 3-d shapes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, 2, pp. 239–256 (1992).
- [22] T. Oishi, R. Sagawa, A. Nakazawa, R. Kurazume and K. Ikeuchi: “Parallel alignment of a large number of range images,” *The 4th International Conference on 3D Digital Imaging and Modeling (3DIM 2003)*, pp. 195–202 (2003).
- [23] M. Wheeler, Y. Sato and K. Ikeuchi: “Consensus surfaces for modeling 3d objects from multiple range images,” *ICCV*, pp. 917–923 (1997).
- [24] W. Lorensen and H. Cline: “Marching cubes: a high resolution 3d surface construction algorithm,” *SIGGRAPH '87*, pp. 163–170 (1987).
- [25] 中野, 山本, 吉田 : “ウェーブレットによる信号処理と画像処理,” .