

論文 / 著書情報
Article / Book Information

題目(和文)	屈折光の波長解析による実世界理解
Title(English)	
著者(和文)	石原慎
Author(English)	Shin Ishihara
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12084号, 授与年月日:2021年9月24日, 学位の種別:課程博士, 審査員:金子 寛彦,佐藤 いまり,山口 雅浩,熊澤 逸夫,渡辺 義浩,長谷川 晶 —
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12084号, Conferred date:2021/9/24, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

博士論文

屈折光の波長解析による実世界理解



東京工業大学
工学院
情報通信系
情報通信コース

石原 慎

主指導教員
副指導教員

金子 寛彦 教授
佐藤 いまり 教授

概要

実世界理解はコンピュータビジョンにおける重要なタスクの一つであり、カメラから得られた画像または動画を用いて、物体の追跡、認識、三次元形状推定といった要素技術から構成される。これらの要素技術に関する研究は盛んに行われており、近年では車の自動運転制御や運転者補助システム、ラインでの自動製品検査、CG化のための自然物の立体形状の取得等にも活用されている。実世界理解のための重要な性質の一つとして、対象物体の分光反射率（色情報）が挙げられる。近年では、対象物の詳細な色情報を取得する分光イメージングに関する技術が急速に進化してきており、高速に分光画像を撮影することのできるカメラも普及段階に入っている。分光イメージングによって得られる詳細な波長情報を用いることによって、自然物や美術品等の高精細な質感情報の解析が実現されてきており、コンピュータによる実世界理解に大きく貢献している。本研究では、実世界センシングのための新たな枠組みとして、分光画像中における屈折光の解析による「三次元分光カメラシステム」を提案する。豊富な色情報を含む分光画像に三次元形状の情報を付加することが可能となり、三次元情報と分光情報を組み合わせた実世界理解への応用が可能となる。

第一に、「レンズ内屈折における波長依存性（色収差）の解析による分光画像からの距離推定手法」を提案する。レンズによって集光される光はレンズを通過する際に屈折を生じるが、その際の屈折率が光の波長によって異なるため、分光画像として観測すると画像の焦点面および焦点ボケの様子が波長毎に変化する。本研究では、空間方向のボケと波長方向のボケを双方向に解析することによって対象物体までの距離を推定する手法を実現した。過去にも色収差を距離推定に応用した手法はいくつか存在するが、本研究では特殊なフィルタやレンズ設計を必要とせず、かつ連続的な距離を取得可能である点で優位である。また、従来法では市販のRGBカメラのような広帯域のセンサを用いているが、本研究では市販のRGBセンサに加えて、狭帯域のRGBセンサおよび5波長のマルチスペクトルセンサを持つカメラを用いた場合の距離推定精度を比較し、狭帯域、多波長センサの優位性を検証した。

第二に、「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」を提案する。空気中から水中の物体を撮影した場合、レンズを通過する光の屈折と同様に、水面を通過する光も屈折する。正投影のスペクトルカメラを用いて水中を観察すると、屈折率が観測波長によって異なることから、波長画像間に深度依存のシフト（視差）が生じる。本研究では、正投影の多波長スペクトルカメラによって観測される波長間の視差の関係を解析することによって、水面（屈折面）の法線方向および水中物体の水面からの距離を同時に推定する手法を実現した。これにより、従来では実現されていなかった、単一視点のカメラ映像からの水面形状および水中シーンの推定手法を実現することに成功した。

目次

図目次	7
表目次	8
第1章 序論	9
1.1 研究背景	9
1.2 研究目的	11
1.3 論文構成	12
第2章 関連研究	13
2.1 分光を用いた質感解析手法	13
2.2 主要な三次元深度推定手法	14
2.3 Depth-from-Defocus 法	16
2.4 水中シーンの三次元再構成	17
2.5 関連研究の総括および本研究の意義	20
第3章 光の屈折における波長依存性	22
3.1 レンズ内屈折における波長依存性	22
3.2 水面での屈折における波長依存性	25
第4章 レンズ内屈折解析による分光画像からの距離推定	28
4.1 分光画像からの距離推定のモデル	28
4.1.1 分光画像に含まれる焦点ボケのモデル化	28
4.1.2 焦点ボケからの深度の定式化	29
4.1.3 空間微分および波長微分の計算	31
4.1.4 波長の選択	35
4.2 評価実験	36
4.2.1 実験環境およびパラメータキャリブレーション	36
4.2.2 定量評価	37

4.2.3	複雑なシーンに対する定性評価	42
4.2.4	ノイズを含むシーンに対する定性評価	50
4.3	小括	54
4.3.1	本提案のまとめ	54
4.3.2	単色収差の影響	54
第5章	水面の屈折解析による分光画像からの水中シーン再構成	57
5.1	多波長の光線による水中シーンのモデル化	57
5.1.1	概要	57
5.1.2	水面における屈折と水面法線・水中シーンとの関係	60
5.1.3	既知の水面における水中シーンの再構成	61
5.1.4	未知の水面における水中シーンの再構成	62
5.2	改良アルゴリズムの導入	65
5.2.1	ルックアップテーブルによるパラメータ推定	65
5.2.2	多波長画像を用いたマッチング精度改善	67
5.2.3	変形を伴う画像の密なマッチング	67
5.3	実験	69
5.3.1	シミュレーションによる実験	69
5.3.2	実環境での実験	75
5.4	小括	80
第6章	結論	81
6.1	今後の課題と展望	81
6.1.1	高度な実世界理解	81
6.1.2	波長依存性の活用についての展望	81
6.1.3	システムの実応用についての課題と展望	82
6.2	総括	83

目 次

1	光の屈折における波長依存性	11
2	色収差の影響による焦点ボケの変化.	22
3	可視光域におけるガラスの屈折率の波長依存性. (出典: < https://www.tokaioptical.com/jp/technology04/ >)	23
4	単レンズ光学系における焦点ボケに関連する物理パラメータ.	24
5	水面での屈折における波長依存性.	25
6	水の屈折率の波長依存性.	26
7	水面での屈折と水面法線および水深との関係.	27
8	単レンズ光学系における焦点ボケに関連する物理パラメータ.	29
9	空間ラプラシアンフィルタを適用した様子.	32
10	隣接波長に対する波長微分を適用した様子.	34
11	第一, 第二波長として選択された波長のマップ.	35
12	カメラ, 光源, 被写体等の実験環境.	36
13	広帯域 RGB カメラと狭帯域マルチバンドカメラ, およびそれらの波長 応答.	37
14	定量評価実験に用いたテクスチャパターンの様子.	38
15	推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ (単位は cm).	40
16	推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ (単位は cm).	41
17	Wood toy (積み木) の例.	43
18	Concrete chair (椅子) の例.	44
19	Flowers (花) の例.	45
20	Llama (ラマ) の例.	46
21	Christmas (クリスマス) の例.	47
22	Tiger (トラ) の例.	48
23	焦点ボケを含む入力画像と, 推定された全焦点画像の比較.	49
24	Wood toy (積み木) の例 ($\sigma_n = 0.001$).	50

25	Wood toy (積み木) の例 ($\sigma_n = 0.01$)	50
26	Concrete chair (椅子) の例 ($\sigma_n = 0.001$)	51
27	Concrete chair (椅子) の例 ($\sigma_n = 0.01$)	51
28	Llama (ラマ) の例 ($\sigma_n = 0.001$)	51
29	Llama (ラマ) の例 ($\sigma_n = 0.01$)	52
30	Christmas (クリスマス) の例 ($\sigma_n = 0.001$)	52
31	Christmas (クリスマス) の例 ($\sigma_n = 0.01$)	52
32	Tiger (トラ) の例 ($\sigma_n = 0.001$)	53
33	Tiger (トラ) の例 ($\sigma_n = 0.01$)	53
34	単色収差 (ザイデル五収差) の様子. (出典: < https://ameblo.jp/stereo-gn-ryoma/entry-12559502437.html >)	56
35	水面での屈折によるシーンの深度依存の分光画像のズレ.	58
36	提案手法の概要.	59
37	水面における屈折の波長依存性と, 水面法線および水中の三次元点との関係.	60
38	入射光, 屈折光および屈折面の法線ベクトルのなす共平面.	63
39	提案システムの試作機.	66
40	カメラからの水面の入射角と, ある波長ペアにおける視差の比率の関係.	66
41	シミュレートされたシーンに対する拡散画像マッチングによる密な点群.	68
42	カメラと水面の入射角の増加による波長画像間の視差との関係. . . .	70
43	シミュレーションによる波状水面および水中の平面における再構成結果.	73
44	シミュレーションによる水平水面および水中の球体における再構成結果.	74
45	提案システムの試作機 (再掲).	75
46	試作システムによる湾曲した屈折水面および水中平面の再構成結果. .	78
47	試作システムによる水平水面および三角形のオブジェクトの再構成結果.	79

表 目 次

1	推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ の, 計測区間全域における 平均値 (単位は cm)	39
2	推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ の, 計測区間全域における 平均値 (単位は cm)	39
3	試作した提案システム (図 39) において抽出された視差の平均値と標 準偏差.	67
4	25mm の水深における最適化法による推定結果.	71
5	25mm の水深における LUT 法による推定結果.	71
6	50mm の水深における最適化法による推定結果.	72
7	50mm の水深における LUT 法による推定結果.	72
8	25mm の水深における LUT 法による推定結果 (実環境).	76
9	50mm の水深における LUT 法による推定結果 (実環境).	76

第1章 序論

本章では、本論文にて扱う実世界理解に関連する分光情報や三次元深度情報について、産業応用および研究分野における背景について述べる。

1.1 研究背景

コンピュータによる実世界理解は、コンピュータビジョン研究における重要な課題の一つである。実世界理解におけるタスクは、カメラから得られた画像または動画からの物体追跡または識別、三次元形状推定といった要素技術から構成される。車の自動運転制御や運転者補助システム、工場のラインでの自動製品検査、CG化のための自然物の立体形状および質感情報の取得等への応用が注目されている。これらのタスクの実現は、物理モデリングに基づく手法と、機械学習を活用する手法に大別される。物理モデリングに基づく手法は、現象の物理的な性質や光学的な振る舞いを定式化し、そのモデル式に基づく物理解を求めることによって推定を行うものである。機械学習に基づく手法は近年目覚ましい発展を遂げており、多くの学習データを効率よく学習することで精度の高い推定が可能になってきている。学習による推定は統計的に高い精度を実現できる反面、適用できるタスクが限定的であったり、学習されていない事象には対応できないといった課題が残る。

画像からの三次元情報の推定は、従来より多くの研究者が取り組んでいる問題であり、これまでに様々な研究が提案されている。物理モデリングに基づく手法としては、Structure-from-Motion (SfM) および Depth-from-Defocus (DfD) とよばれる手法が代表的である。SfM法では実際の測量法に基づいて複数視点のカメラ画像から対象点までの距離を推定し、DfD法では画像に含まれるボケ現象が対象点の距離に依存することを利用する。どちらの手法についても精度よく三次元情報を推定するためには、複数台のカメラを用いるか、カメラのパラメータ（撮影位置やピント位置など）を変えて複数撮影する必要がある。機材コスト、システムの大きさ、撮影の手間といった、実応用上の制約が少なからずある。そのため近年では、機械学習を用いて三次元情報の推定精度を改善しようとする試みもみられる。

実世界理解のための重要な性質の一つとして、対象物体の分光反射率（色情報）が挙げられる。ヒトの目の細胞には赤・緑・青の三種類の光の波長に反応する錐体細胞があることが知られている。ヒトの見る世界と同様な情報を撮影するために、三種類の錐体細胞に対応する受光素子を持つ RGB カメラが一般に普及している。他方で、蜂や蝶などの虫は紫外線に反応する錐体細胞を持ち、花の蜜のある場所を見分けることができる。このように、自然界の生物は必ずしもヒトと同じ波長を観測しているということではなく、カメラで観測する波長が RGB である必要もない。近年では、従来の RGB 情報では困難であった食品の腐敗検知などへの応用が有効であることから、シーンの詳細な色情報を取得する分光イメージングに関する技術が急速に進化してきており、高速に分光画像を撮影することのできるカメラも普及段階に入っている。RGB カメラの波長分解能が 100nm 程度であるのに対して、分光カメラの波長分解能は 10nm～1nm を実現しており、従来の 10 倍～100 倍の豊富な波長情報を扱うことが可能となっている。分光カメラによって得られる詳細な波長情報を用いることによって、コンピュータビジョン研究の分野では、例としてツヤ、散乱、濡れなどの物体表面の状態解析への応用がなされており、自然物や美術品等の高精細な質感情報の解析が実現されてきている。これらの高度な質感情報の取得はコンピュータによる実世界理解において欠かせない要素の一つであり、対象物体のもつ色情報は重要な性質の一つであるといえる。

本研究では、実世界センシングのための新たな枠組みとして、シーンの分光情報から三次元深度情報の推定を行う「三次元分光カメラシステム」を提案する。分光情報から三次元情報を推定するための手掛かりとして、光の屈折に着目する。図 1 に示すように光の屈折率には波長依存性があるため、分光カメラを用いると異なる波長の光は異なる経路の光として観測される。異なる経路を通る光を観測することは、物理モデリングによる三次元情報推定において、カメラのパラメータを変えた異なる撮影条件でシーンを観測することに相当する。すなわち、分光画像に含まれる屈折の波長依存性によって、複数回の撮影を行った場合と同等な三次元情報への手がかりを、一度の撮影で得ることができる。屈折を含む撮影システムにおける分光画像を、光の波長依存性に基づく屈折モデルを用いて解析することによって、三次元情報の推定と、屈折による画像劣化の復元を行う。シーンの分光画像と高精度な三次元情報を同時に取得することにより、三次元情報と分光情報を組み合わせた実世界理解への応用システムを実現する。

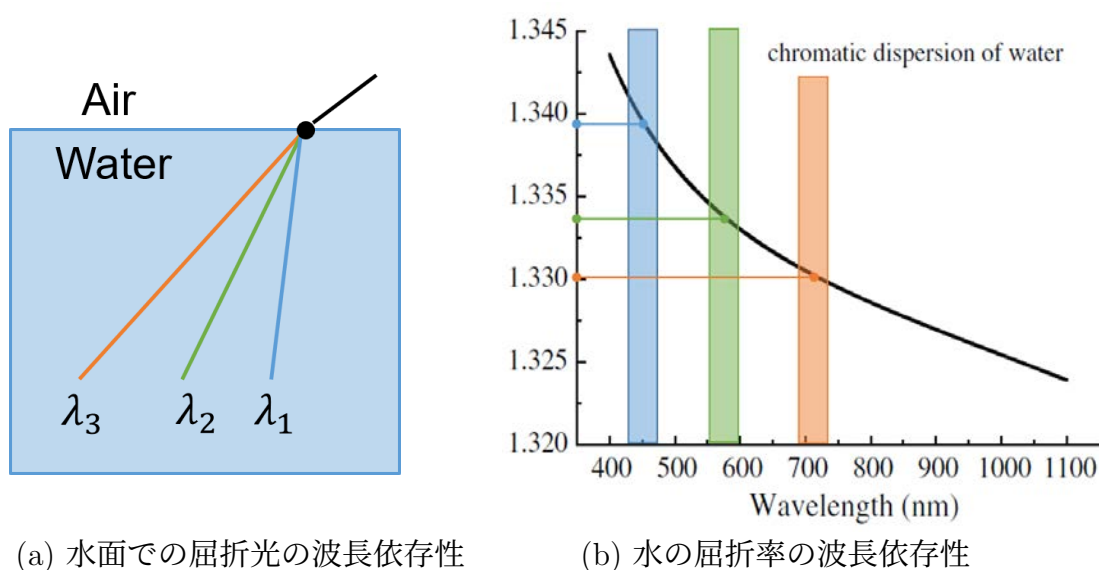


図 1: 光の屈折における波長依存性

1.2 研究目的

本論文では、屈折光の波長依存性に基づくモデルを用いて分光画像による解析を行うことにより、分光情報にシーンの三次元情報を付加する手法の実現を目指す。第一に、「レンズ内屈折における波長依存性（色収差）の解析による分光画像からの距離推定手法」を提案する。レンズによって集光される光はレンズを通過する際に屈折を生じるが、その際の屈折率が光の波長によって異なるため、分光画像として観測すると画像のフォーカス面およびボケ方が波長毎に変化する。また一般に、画像に生じるボケは対象物体までの距離にも依存する。本研究では、空間方向のボケと波長方向のボケを双方向に解析することによって対象物体までの距離を推定する手法を実現する。さらに、多波長センサを用いたプロトタイプシステムを使用して、本手法の有効性を検証する。第二に、「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」を提案する。空気中から水中の物体を撮影した場合、レンズを通過する光の屈折と同様に、水面を通過する光も屈折する。正投影のスペクトルカメラを用いて水中を観察すると、屈折率が観測波長によって異なることから、波長画像間に深度依存のシフト（視差）が生じる。本研究では、正投影の多波長スペクトルカメラによって観測される波長間の視差の関係を解析することによって、水面（屈折面）の法線方向および水中物体の水面からの距離を同時に推定する手法を実現する。シミュレーションおよびプロトタイプシステムによる実験を通して、本手法の有効性を検証する。

1.3 論文構成

本論文の構成を以下に示す.

第1章では, 本論文のコンピュータビジョン分野における研究背景と研究目的を述べた.

第2章では, 光の物理特性を用いた三次元形状推定手法や, 水中シーンの三次元再構成手法について関連研究を紹介し, 本研究の意義について議論する.

第3章では, 本論文にて扱う光の屈折における波長依存性について原理や特性を紹介する.

第4章では, 本論文にて提案する「レンズ内屈折における波長依存性(色収差)の解析による分光画像からの距離推定手法」について, 分光を用いた距離推定のモデルおよび推定アルゴリズムの詳細を述べる.

第5章では, 本論文にて提案する「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」について, 分光を用いた水面形状および水中シーンの再構成モデルおよび推定アルゴリズムの詳細を述べる.

第6章では, 本論文を通して達成された事項について総括としてまとめ, 今後の課題について述べる.

第2章 関連研究

本章では、コンピュータによる実世界理解に関連して、分光画像を用いた質感解析手法や、三次元深度推定手法に関する研究を紹介する。第一の目的である「レンズ内屈折における波長依存性の解析による分光画像からの距離推定手法」に関連して、コンピュータビジョンにおける主要な距離推定手法と、画像に含まれるボケを利用した距離推定手法について紹介する。第二の目的である「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」に関連して、水中シーンの深度解析手法に関する研究について紹介する。これらの提案を通して、本論文では実世界センシングのための新たな枠組みとして、シーンの分光情報から三次元深度情報の推定を行う「三次元分光カメラシステム」を提案する。

2.1 分光を用いた質感解析手法

ヒトは赤・緑・青の錐体細胞による反応に基づいてシーンを捉え、様々な質感を認識しているが、そのメカニズムは未だに明らかになっていない。そのため、RGBカメラによる映像をもとにしたコンピュータによる質感認識は非常に困難であることが知られており、現段階における実現可能性は高くない。分光イメージングによって、従来のRGBカメラよりも豊富な波長情報を取得できることは先にも述べた。シーンの波長情報が増えることの利点は単に色の区別が付きやすくなるというだけではなく、光の波長に基づく物理モデルを適用することで物体の高度な質感解析に应用することが可能となる点が挙げられる。ここでは、分光イメージングを用いて高度な質感解析を実現した研究事例について紹介する。

Zheng ら [68] は、物体表面における反射スペクトルが光源スペクトルによって変化する問題に対し、分光画像の照明・反射スペクトル分離問題 (IRSS) を一般的な照明下で解くための研究に取り組んだ。IRSS 問題を低ランク行列因数分解にモデル化することによって、反射スペクトルがあまり複雑でないという仮定の下で、一意に分離可能であるという理論が示された。これにより、様々な照明下においてもその影響を受けずに物体の反射スペクトルを取得することが可能となっている。

Okawa ら [36] は、物体表面における反射スペクトルが相互反射によって変化するという問題に対し、分光画像を用いて相互反射の影響を除去する研究に取り組んだ。均質な反射特性を持つ表面において、相互反射を繰り返して観測された光はそのスペクトルが先鋭になるという性質に着目し、シーンの各点における相互反射回数を非負値行列因子分解問題として分離することを実現した。これにより、相互反射によって物体表面の色の見えが変わってしまうシーンにおいても、その影響を除去して物体本来の色の見えを観測することができる。

Shimano ら [48] は、表面が濡れていると感じるヒトの質感認識について、光学的なモデルを用いて分光画像から表面の濡れ度合いを推定する手法を提案した。表面の濡れによって光の前方散乱性が強くなる性質を利用して、表面の濡れ度合いによる輝度や反射スペクトルの変化をモデル化した。さらに、二種類の異なるテクスチャの境界についてこの理論を拡張し、より一般的なシーンにおいても、分光画像による濡れ度合いの推定が適用可能であることを示した。

2.2 主要な三次元深度推定手法

コンピュータビジョンの分野では、二次元画像から三次元深度を推定するための様々な研究が数十年に渡って提案されてきた。コンピュータビジョンにおける主要な三次元深度推定手法をその原理に基づいて分類すると、Triangulation, Time-of-Flight (ToF), Shape-from-X (X の例としては Motion, Shading, Texture, Focus など) が典型) とよばれる手法が代表的である。様々なアルゴリズムが存在するために、これらの手法は、センサまたは光源の種類によって、能動的 (アクティブ) センシングと受動的 (パッシブ) センシングの二つのグループに大別される。Triangulation 法や Shape-from-Focus 法は一般的に受動的な光源下で適用可能であるのに対し、ToF 法や Photometric Stereo (PS) 法は能動的な光源を原理とした手法である。ここでは、主要な三次元深度推定手法について関連研究を紹介する。

Triangulation 法

Stereo 法 [33, 4, 22, 42, 51, 21, 64] は、既知の2つの視点から撮影された画像を用いて、各視点から対象物上の各点に対応する画素を探索し、各点について三角測量を行うことによって、対象物体の深度を推定する手法である。Triangulation 法を拡張し、複数台のカメラあるいは1台のカメラを移動させながら取得した画像群から、各視点におけるカメラの姿勢と並進移動の方向を推定する Structure-from-Motion (SfM) 法

[56, 54, 58, 45, 37] とよばれる手法も存在する。いずれの手法においても、異なるカメラ視点における物体の位置を特定するためには、各視点において共通の特徴点を求める必要がある。Triangulation 法は局所特徴量を用いて画像の対応点を確立するために、対象物体表面に十分なテクスチャが含まれていることを仮定する。

Structured-Active-Light 法 [7, 65, 43, 41] は、可視光あるいは赤外光のパターンを対象表面に照射し、パターンの変化を解析することによって表面形状や深度を推定する。能動的にパターンを照射するため、Stereo 法や SfM 法と異なり対象表面にテクスチャが十分に含まれている必要がない。しかしながら、パターンを照射する必要がある点でコストが高く、屋外などの環境光が強い場合には光学的な干渉の影響を受けやすい。

レーザー三角測量 [63] は、対象物体に照射されたレーザー点をカメラで観測し、画素位置および光源位置から三角測量により深度を推定する。レーザー光源が必要であることと、キャリブレーションが容易ではないというコスト面の欠点があるが、高精度な距離計測が可能である。また、レーザーの特性により長距離の測定よりもむしろ、近距離の小規模な対象物の測定に適している。

Time-of-Flight 法

Time-of-Flight 法 [59, 8, 11, 19] は、パルス光を計測対象に照射し、物体表面で反射した光が再び計測装置に戻るまでの時間や位相差を直接エンコードすることで、深度を推定する手法である。単発パルスの飛行時間を測定する直接 Time-of-Flight 法と、連続波の光の位相差を測定する間接 Time-of-Flight 法が存在する。直接 Time-of-Flight 法は、高速な光の飛行時間を計測する必要があるため、時間分解能の高いセンサが必要となる。動作範囲としては長距離の対象についても測距可能なため、広範囲に及ぶ屋外の建物等の深度推定に適している。間接 Time-of-Flight 法は、センサと同位相の光パルス照射し、対象表面にて反射して戻ってきた拡散反射成分の位相差を計測する。直接 Time-of-Flight 法と比較すると動作範囲は短く、近年では Microsoft 社製 Kinect 赤外線光源を用いた間接 Time-of-Flight 法を採用した。

Shape-from-X

Shape-from-X と呼ばれる深度推定手法は、対象表面における特定の性質をもとにして、表面の三次元情報を推定する手法である。Photometric Stereo (PS) [60, 26, 6, 5, 6] は、能動的な光源を用いる代表的な手法の一つであり、画像内の陰影の緩やかな変化から表面形状を推定する手法である。PS の基本は、様々な光源位置での撮影画像におけ

る表面の明るさの変化から、表面の法線を推定することである。すなわち、PS問題は撮影された画像の輝度に基づいて、光源位置と表面法線を推定する。PSと同様に物体表面の陰影情報を活用する手法として、Shape-from-Shading (SfS) [27, 16, 39, 40, 67]がある。PS手法は一般に能動的な光源を用いて、複数枚の画像を入力として与えるのに対し、SfS手法の一部は自然光環境下において撮影された単一の画像のみを用いて表面形状を推定する[23]。これらの方法の多くは、対象表面の反射についてランバート反射モデルを仮定していることが多く、シーンには強い陰影の変化をもたらすテクスチャが含まれていないことが重要となる。したがって、Triangulation法とは対照的に、テクスチャや非ランバート面はこれらの方法の精度を低下させる可能性があることがわかっている。また、環境を制御できない状況下で能動光源を用いた方法を適用することは、環境光の影響によりこれらの手法の前提が成立しない場合がある。

Shape-from-Focus または Depth-from-Focus 法 [35, 13] は、画像の焦点特性から深度を推定する初期の手法である。この手法は本研究において提案する「レンズ内屈折における波長依存性の解析による分光画像からの距離推定手法」と深く関係しているため、詳細は後述する。

2.3 Depth-from-Defocus 法

Depth-from-Focus (DfF) 法は、画像の焦点特性から深度を推定する初期の手法である。この方法では、キャリブレーション済みの様々な焦点距離で撮影された画像群から、最も焦点の合う画像を探索することにより、三次元深度を求める手法である。DfF法は信頼性の高い深度を取得することができるが、撮影時間や撮影枚数の観点ではコストが大きい。実用的な撮影コストの制約から撮影画像を数枚に限定する場合、焦点面 (Focus) ではなく焦点ボケ (Defocus) からの深度推定アルゴリズムを考えることが有効である。そこで、Depth-from-Defocus (DfD) [57, 28, 14, 15] と呼ばれる手法が提案されている。DfD法は自然光の下で機能する代表的な深度推定手法の一つであり、画像に含まれる焦点ボケを手掛かりにシーンの深度を推定する手法である。典型的な DfD 法は、少なくとも二種類の異なるボケカーネルを持つ画像を必要とする。ボケカーネルは様々な方法で変化させることができるが、レンズの絞りや焦点距離などのパラメータを操作して変化させることが代表的である。ボケカーネルの制御方法によって、計算コストや深度推定の精度に影響することが知られている。近年では、センサとレンズ間の距離やレンズの焦点距離を高速に変化させることができる固有の装置を用いて、ボケの時間変化を利用して DfD 法における距離推定の精度を向上させた例がいくつか存在する。Alexander ら [2] は、同じ光軸上のレンズとセンサの距離を

制御して、異なるカメラパラメータで複数の画像を撮影するシステムを提案した。この研究では、対象物体までの距離と撮像画像の空間的・時間的導関数の関係を定式化し、デフォーカスからの高速かつ正確な距離推定を可能にした。Guo ら [18] は、液晶レンズと呼ばれる高速にその焦点距離を変えられるレンズを用いて、焦点距離の異なる画像を連続して撮影できるシステムを提案した。彼らのシステムは、Alexander らと類似したモデルに従った画像を取得し、高精度な深度推定を実現できることを示した。DfD 法は二枚以上の画像を使用してシーンの奥行きを推定するという点や、シーンに十分なテクスチャが含まれることを仮定している点などで Triangulation 法と共通しており、しばしば両者のメリット・デメリットが議論されている [44]。

他方、一度の撮影による画像のみを用いてシーンの奥行きを推定する DfD 法もいくつか提案されている。Zhuo ら [69] らは、画像のエッジ位置で空間的に変化するデフォーカスボケの強度に着目し、シーン全体の奥行き関係をデフォーカスマップとして推定する手法を提案した。彼らの手法はシンプルで効果的だが、入力画像はシーンの最前面の物体に焦点を合わせなければならず、相対的な奥行きのみ取得可能であるといった制限がある。さらに、RGB の各チャンネルでボケカーネルの異なる RGB 画像を取得して、一枚の RGB 画像を用いた DfD により深度推定を行う方法もいくつか提案されている。Tang ら [52] は、色収差による RGB の各チャンネルでのボケの変化をデフォーカスマップ推定に導入し、精度を改善することに成功した。Chakrabarti ら [9] は、カラーフィルタによる開口部を用いたシステムを提案した。符号化開口 (Coded Aperture) とよばれる技術を応用して、RGB 各チャンネルの画像の被写界深度を制御し、単一の RGB 画像からシーンの奥行きを推定することに成功した。Trouve ら [55] は、レンズ群を特別に設計することにより、波長ごとの焦点距離の変化を制御するシステムを提案した。特殊なレンズ群を用いた RGB カメラで撮影した画像の各チャンネルにおいて、ボケカーネルから得られる PSF (Point-Spread Function; 点広がり関数) を距離情報と対応付けることにより、PSF 推定によってシーン深度を求めることに成功した。Chakrabarti らの手法と Trouve らの手法の共通の特徴として、システムの実現には特別な設計が必要であることと、デフォーカスボケと深度の対応が離散的な関係としてのみ得られることが挙げられる。

2.4 水中シーンの三次元再構成

ここでは、本研究において提案する「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」に関連する、水中を対象としたシーンの三次元再構成手法に取り組んだ研究について紹介する。一般に、水面での光の反射や屈折、水

中での光の散乱や減衰などにより、2.2 節で紹介した主要な深度推定手法を水中シーンに適用することは困難な問題である。このような水の光学特性を考慮し、水面が平坦である場合や流体である場合などについて、水中の三次元シーンを再構成するための様々な手法が提案されている。以下では、これまでの水中シーンの三次元再構成に関する研究について、三種類に分類して紹介する。

水中ハウジング内での Triangulation 法におけるカメラジオメトリ較正

水中シーンを対象として従来の三次元推定手法を適用させるために、カメラハウジングと水面での光の屈折による幾何学的歪みを補正する方法が初めに提案された。Chari と Sturm[10] は、水中ハウジング内のステレオカメラシステムのキャリブレーションを初期に提案した。彼らは、左右のカメラで共通の屈折面における屈折がカメラジオメトリに与える影響を行列形式で表現し、両眼ステレオや SfM などの従来の 3D 再構成方法を水中シナリオに適用することを可能にした。水中ステレオシステムには複数の屈折面がある（つまり、水とハウジングの境界およびハウジングと空気の境界）ことに着目して、Agrawal ら [1] は複数の屈折面のもたらすジオメトリ歪みの補正を提案した。彼らは、カメラに入射する光線の複数の平らな平行平面での屈折の影響が、線形制約と行列計算を使用することによって、キャリブレーションパターンから補正できることを示した。Sedlazeck と Koch[29] は、進化的アルゴリズムと組み合わせた最適化により、色のにじみを引き起こす水中散乱の影響を検討し、Agrawal の方法 [1] を改良した。彼らのアプローチは、カメラとガラスの間の距離とガラス表面の法線を与えます。Dolereit ら [12] は、水中ステレオの屈折パラメータをキャリブレーションする方法を提案し、精度を向上させた。

Yau ら [62] は、既存のフレームワークを拡張して、波長依存の屈折モデルを含め、複数のカメラジオメトリのキャリブレーションで使用するモデルパラメーターに新しい制約を導き出した。Kawahara ら [31] は、バリフォーカルカメラモデルによるガラス面での屈折プロセスを導入した。さらに、彼らはアクティブステレオ用のプロジェクタに屈折モデルを導入し、屈折の影響をピクセルごとの仮想カメラ用の仮想プロジェクタとして変換できることを示した。Ichimaru ら [24] は、ライトフィールドカメラを使用して水面の屈折の影響を補正する方法を提案した。彼らは、カメラが水中ハウジング内にあり、システムを適用できる場所を制限するいくつかの平面での光の屈折のみを考慮している。

平らな水面を仮定した水中シーン再構成

水面が平坦であることを仮定して, Gedge ら [17] は, 両眼視差が位置するエピポーラ線が水面での屈折により曲線になることを示した. 導き出されたエピポーラ曲線に基づいて, 従来の両眼ステレオを水中シーンに適用できることを示した. 彼らは, 屈折の影響を補償するマッチングコストを考慮することにより, 水中ステレオマッチングの精度を向上させた. Sedlazeck と Koch[46, 30] は, 水中シナリオでの屈折のために従来のピンホールカメラモデルが成立しないという事実に焦点を当て, SfM メソッドで屈折の効果を明示的に導入することを提案しました. 水中シーンについては, 光の屈折を考慮せずに透視投影のカメラモデルを想定して精度評価を行い, 透視投影モデルがもたらす誤差が測定精度を損なうことを示した.

Sorensen ら [49] は, 平らな水面を想定して, 水中シーンを再構築するためのレイトレーシングベースの方法を提案した. 彼らの方法は水面の屈折を考慮に入れているが, 2 台のカメラを必要とする. Yamashita ら [61] は, 1 台のカメラで撮影した画像から水中シーンを再構築するために, 水槽の隅で観察される 2 つの屈折面を使用することを提案した. Asano ら [3] は, 2 つまたは 3 つの近赤外波長間の水による光の吸収の差を利用して水中の深度を推定した. 彼らは, 観測された 2 つの波長でのすべての物体の反射率が一定であり, 水面が平坦であると仮定している.

流体の水面における水中シーン再構成

Kutulakos ら [32] と Morri ら [34] は, 複数のカメラで撮影した画像を使用して水面の形状を再構築することを提案しました. しかしながら, 水中シーンの再構築は達成されていません. 近年, 屈折面が平坦でない流体表面である場合においても, 水中シーンを再構築する手法がいくつか提案されている. Zhang ら [66] は, DfD にもとづく距離推定を使用して, 水面と水中 3D シーンの形状を同時に再構築することを可能にした. ただし, 彼らの方法では, 水中シーンが静止している必要があり, 水面が平面の場合は水中シーンの追加画像が必要となる. Ichimaru ら [25] は, 1 台のカメラでキャプチャされた複数の画像を使用して水面がわずかに波打つ状況に適用できるように SfM 技術を改善した. 彼らの方法は, 水中シーンを再構築するために多くの入力画像を必要とし, 表面の法線が単一の画像内で一定であると仮定している. 一方, 多視点ステレオの分野では, 両眼ステレオの拡張として, Qian ら [38] によって水面の形状と水中の深さを同時に取得できる方法が提案された. 水面の近傍点を 2 次関数で近似し, その関数の導関数から計算される法線と, 多視点ステレオとスネルの法則に基づくレイトレーシングから計算された水面の法線について, 両者の誤差を最小化する問題によって実現される.

2.5 関連研究の総括および本研究の意義

本章では、コンピュータによる実世界理解に関連する研究事例として、はじめに分光画像解析による質感解析手法について紹介した。分光画像はシーンの分光反射率を詳細に取得できるというだけでなく、高度な実世界理解を実現するために有用であることを示した。続いて、実世界理解における重要なタスクの一つとして、三次元深度情報の取得に関する手法を紹介した。

「レンズ内屈折における波長依存性の解析による分光画像からの距離推定手法」に関連して、画像に含まれる焦点ボケを利用した距離推定手法について紹介した。色収差を利用した既存研究では波長幅の広いRGBセンサを利用しているために色収差の詳細な観測が難しく、連続的な距離を正確に推定するためのモデルが提案されておらず、離散的な距離推定のみ留まっていた。本研究では分光カメラを用いて詳細な色収差を観測し、分光画像に含まれる焦点ボケから物理モデルに基づいて連続的な距離を推定する手法を提案する。

「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」に関連して、水中シーンの深度解析手法に関する研究について紹介した。要約すると、従来のアプローチは水中シーンの再構成に関して、平坦な水面、平坦な水中シーン、静的な水中シーンいずれかを仮定しているものが多く、水面の形状が不明な場合には多くの画像を必要としている。すなわち、水面の形状が不明な水中シーンを一枚の画像から再構成することはこれまで実現されていなかった。本研究では、水面の法線が不明な場合でもアクティブな照明を必要としない、一枚の分光画像を用いた水中シーンの三次元情報の推定手法を提案する。

本論文で提案する分光画像解析による距離推定手法はこれまでの深度推定において困難とされていた課題の解決への手掛かりとなるだけでなく、2.1節にて述べたような対象の分光情報を利用した質感解析と三次元情報を同時に扱うことができる点にも意義がある。三次元深度推定の代表的な応用である物体の3Dスキャニングでは、物体の形状情報を取得することが主な目的とされてきた。近年では文化遺産等のデジタルアーカイブ化を目的として三次元情報と色情報の同時記録に取り組んだ事例もあるが、対象表面の状態や繊細な見えの変化をもたらす高度な質感情報を記録するという取り組みは達成されていない。本研究で提案する分光情報からの三次元情報推定を通して、対象の三次元情報と質感情報の同時記録による高度な実世界理解システムの実現への足がかりとなることを目指す。

以降では、本研究で扱う光の屈折における波長依存性についてその原理や特性について述べ、一般シーンおよび水中シーンを対象とした分光情報からの三次元深度推定

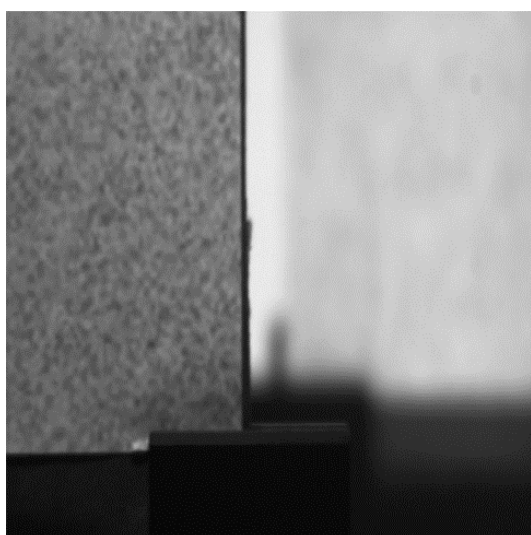
手法について物理モデルや推定アルゴリズムの詳細を述べる．総括として，本論文にて達成された三次元深度推定に関する事項についてまとめ，質感情報との統合に関する今後の課題について述べる．

第3章 光の屈折における波長依存性

本章では，レンズ内屈折における波長依存性と，水面での屈折における波長依存性について，本研究で扱うモデルに沿って説明する．

3.1 レンズ内屈折における波長依存性

本節では，レンズ内屈折における波長依存性に関連する事項について述べる．図 2 に，レンズを通過する光の屈折によって画像に生じる焦点ボケの例を示す．カメラとレンズのパラメータは固定されており，レンズ内を通過する光が屈折する際の屈折率が異なる．焦点距離の波長依存性によって，左右の画像で異なる焦点ボケが生じる．左の画像は 450nm の観測波長で撮影された画像であり，画像の左側ではピントが合っているが，右側では焦点ボケが見られる．右の画像は 650nm の観測波長で撮影された画像であり，450nm 画像とは対照的に画像の右側ではピントが合っているが，左側では焦点ボケが見られる．



(a) 450nm 画像



(b) 650nm 画像

図 2: 色収差の影響による焦点ボケの変化.

レンズの一般的な材質であるガラスの屈折率における、可視光域での観測波長との関係を図3に示す。縦軸がガラスの屈折率、横軸が観測波長を表している。図から、可視光域においては屈折率の変化は単調であり、観測する光の波長が短いほど実効的な屈折率が高くなり、長いほど屈折率が低くなることがわかる。観測波長ごとにレンズの実効屈折率が異なるふるまいをすることによって、図2に示したような焦点ボケの変化が生じる。本研究では、波長依存の屈折率の変化に対する焦点ボケの変化を画像から取得し、距離推定に適用する。

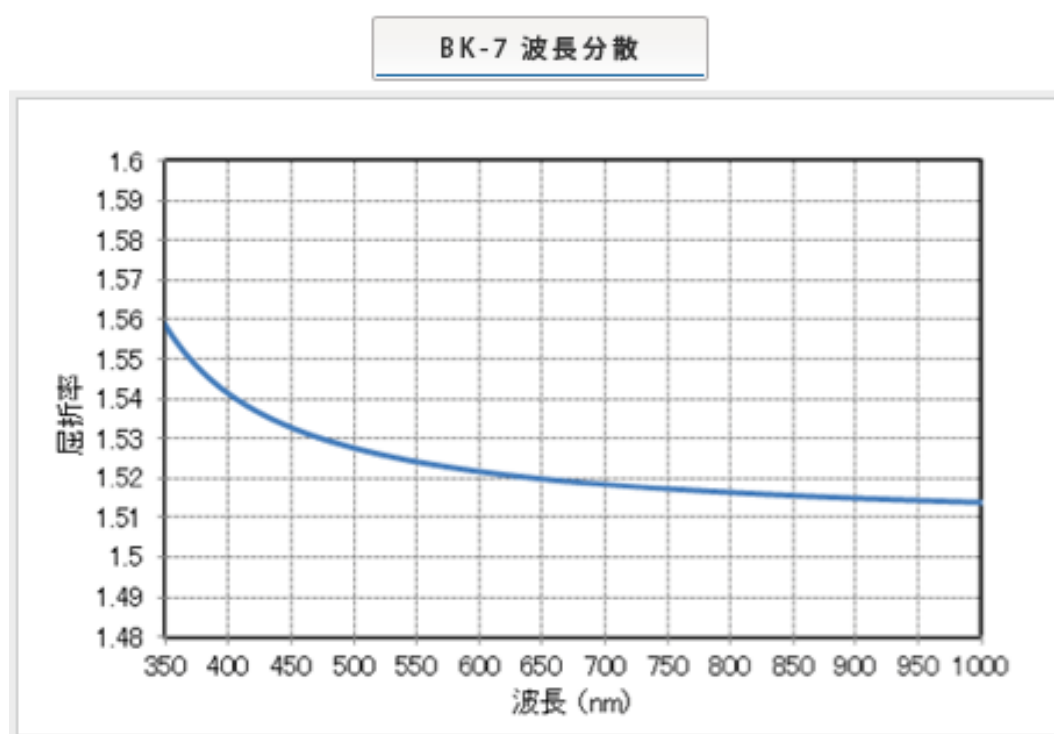
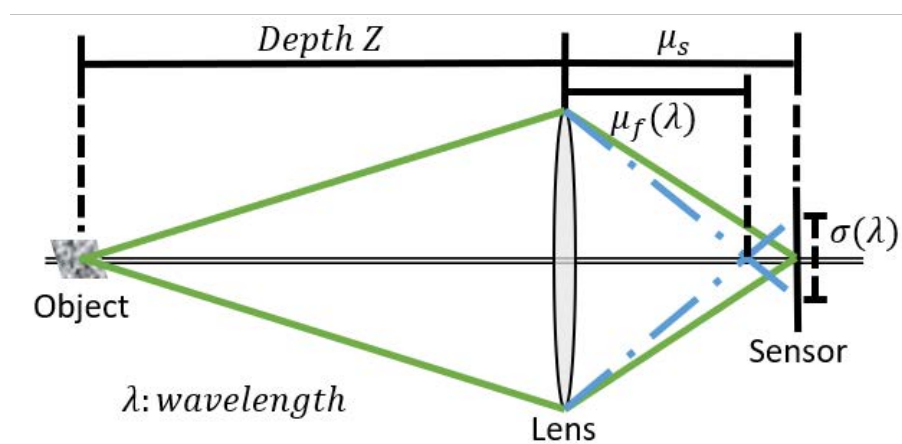
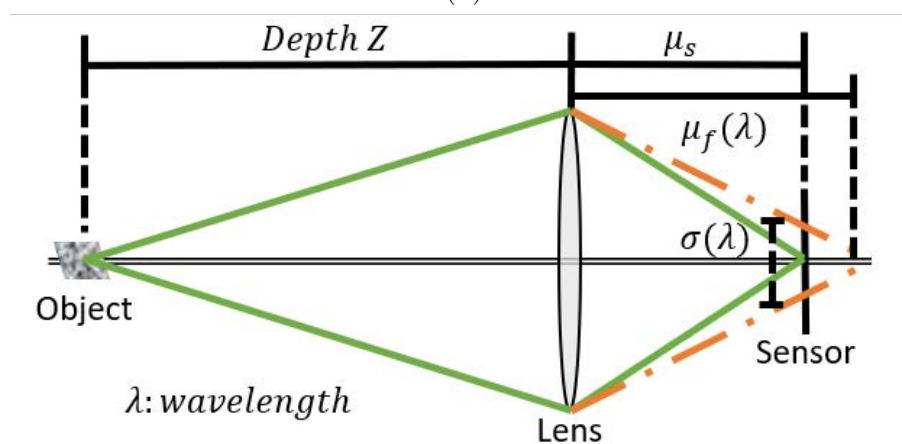


図 3: 可視光域におけるガラスの屈折率の波長依存性. (出典: <<https://www.tokaioptical.com/jp/technology04/>>)

このとき生じる焦点ボケとレンズ光学系との関係を図4に示す。ここで、対象物とセンサ間の距離を Z 、観測波長 λ における光線の焦点位置を $\mu_f(\lambda)$ 、レンズとセンサ間の距離を μ_s とする。 μ_s からの $\mu_f(\lambda)$ のずれは焦点ボケを引き起こし、レンズの色収差は光を波長依存的に屈折させ、レンズの焦点距離 $f(\lambda)$ の変化をもたらす。基準光線（緑で描かれたもの）よりも短い波長（青で描かれたもの）では、より短い焦点距離による焦点ボケを生じる。基準光線よりも長い波長（橙で描かれたもの）では、より長い焦点距離による焦点ボケを生じることがわかる。本研究では、この波長依存の焦点ボケの変化と対象点までの距離との関係を定式化し、(分光) 画像の各点について距離推定を実行することにより、一枚の分光画像からシーンの三次元深度を推定する。推定手法の詳細については後述する。



(a)



(b)

図4: 単レンズ光学系における焦点ボケに関連する物理パラメータ。

3.2 水面での屈折における波長依存性

本節では、水面での屈折における波長依存性に関連する事項について述べる。図5に、画像の観測波長によって異なる屈折の大きさが水深と関係することを表す例を示す。正投影カメラによる450nmの画像（マゼンタ）と700nmの画像（緑）の屈折率における画像をCGシミュレーションにより生成している。画像の上部は水深が浅く、下部は水深が深いため、青と黄色の領域で波長画像間のズレの大きさに違いが生じていることがわかる。

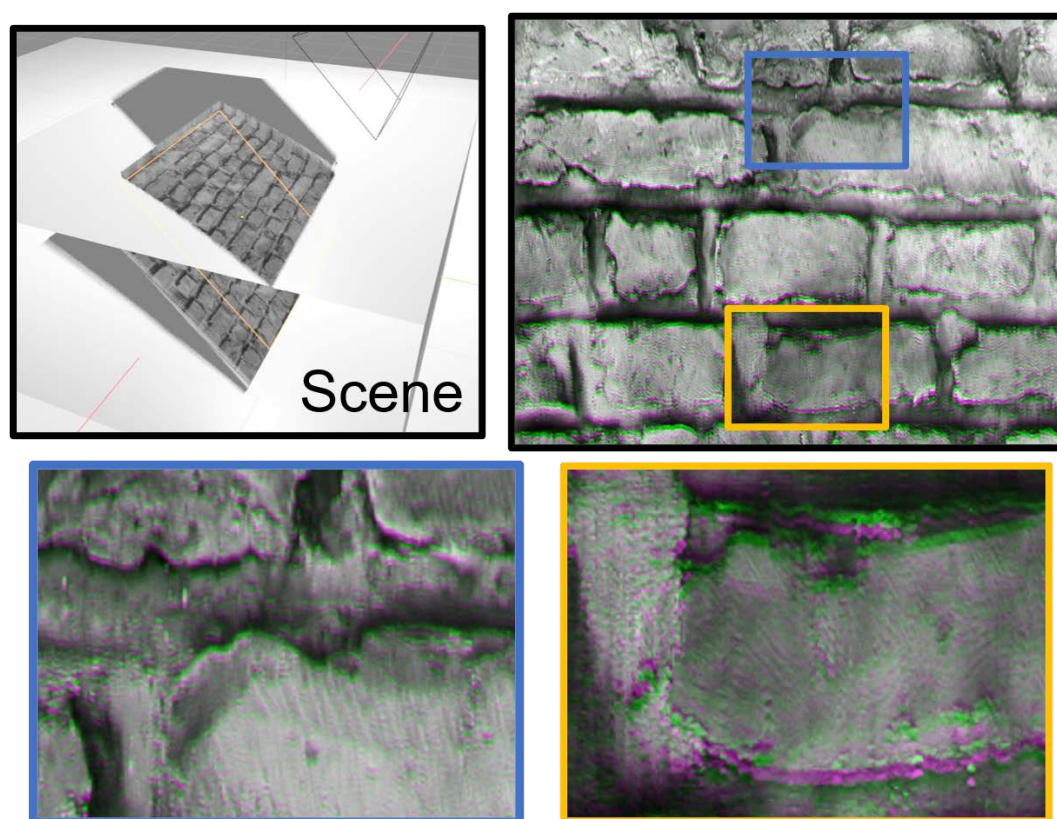
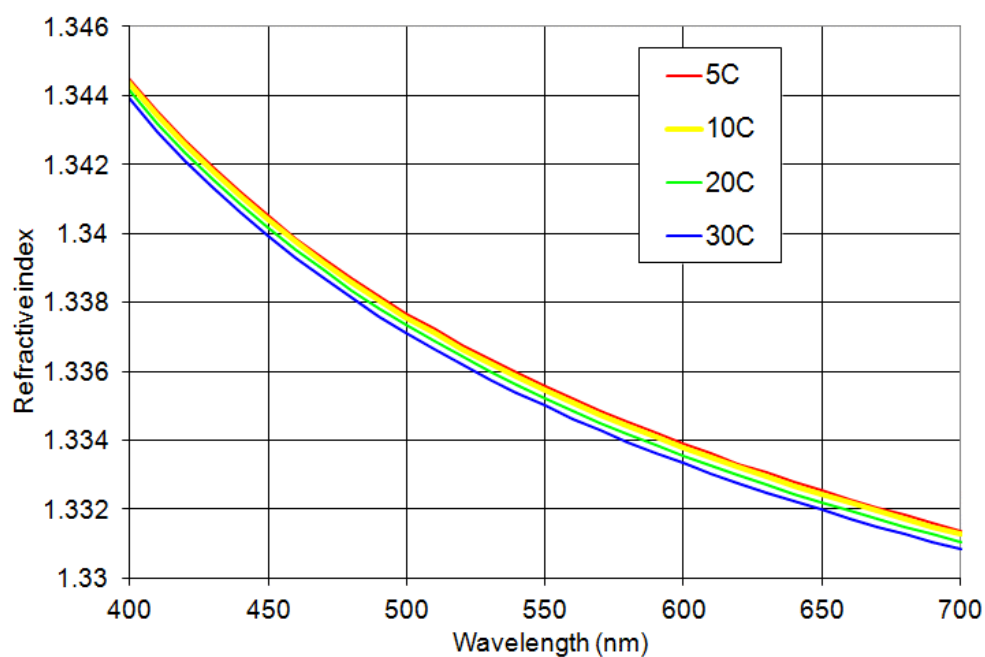
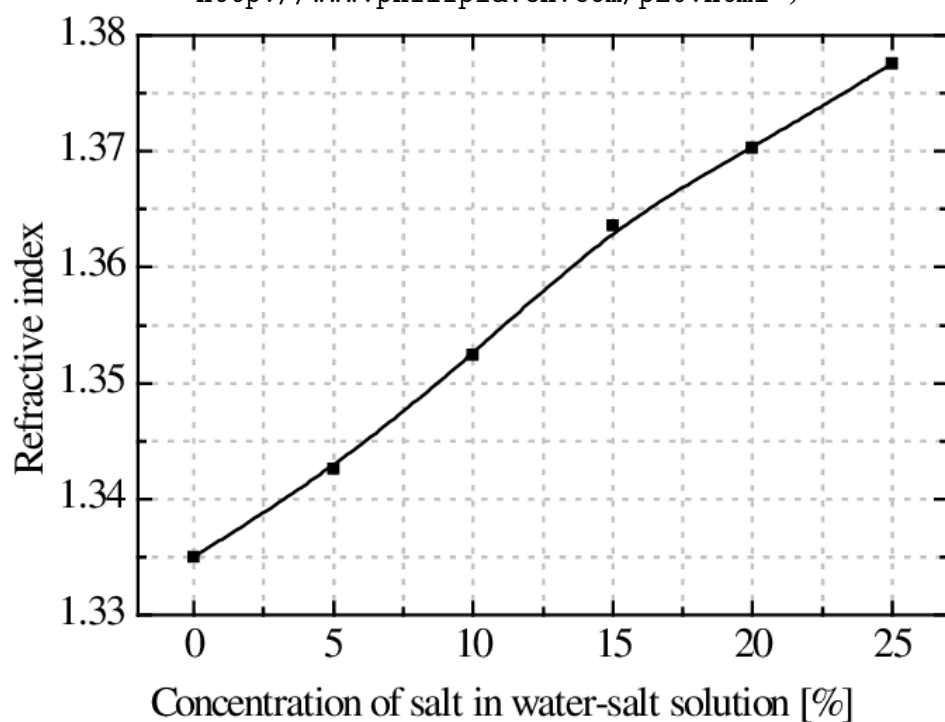


図 5: 水面での屈折における波長依存性.

水の屈折率における、可視光域での観測波長との関係を図6に示す。図6(a)では温度変化による屈折率の変化、図6(b)では混合物による屈折率の変化を示している。可視光全域(400nm – 700nm)を解析に利用できる環境では、波長による屈折率の変化は温度による変化と比較しても有効であることがわかる。混合物によって水の屈折率は変化するため、媒質の濃度が不明である場合には事前に媒質の屈折率を計測する必要がある。



(a) 温度による変化を含む. (出典: Refractive index
<<http://www.philiplaven.com/p20.html>>)



(b) 混合物による変化を含む. (出典: Stupar らの論文 [50])

図 6: 水の屈折率の波長依存性.

図5に見られるような、異なる波長で観測された画像のズレと、水面法線および水深との関係を図7に示す。実際は3次元空間上のモデルとして考えるが、ここでは提案モデルの物理的な設定を示す2次元図を用いて説明する。姿勢ベクトル c の方向を向いた正投影マルチスペクトルカメラを用いて、水中の対象物上の点 w からカメラに到達する光を波長 $\lambda_{1,2,3}$ にて観測する。水中光線 $l_{1,2,3}$ は空気と水の境界での屈折の影響を受けてカメラに入射する光線（屈折率 k は観測波長に依存する）であり、その方向ベクトル $v_{1,2,3}$ は水面法線 n と観測波長に依存する屈折率 $k(\lambda)$ で与えられる。本研究では、波長依存の画像のズレ $p_2 - p_1$ と水面法線および水中対象物上の点との関係を定式化し、（分光）画像の各点について推定を実行することにより、一枚の分光画像から水面および水中シーンを再構成する。推定手法の詳細については後述する。

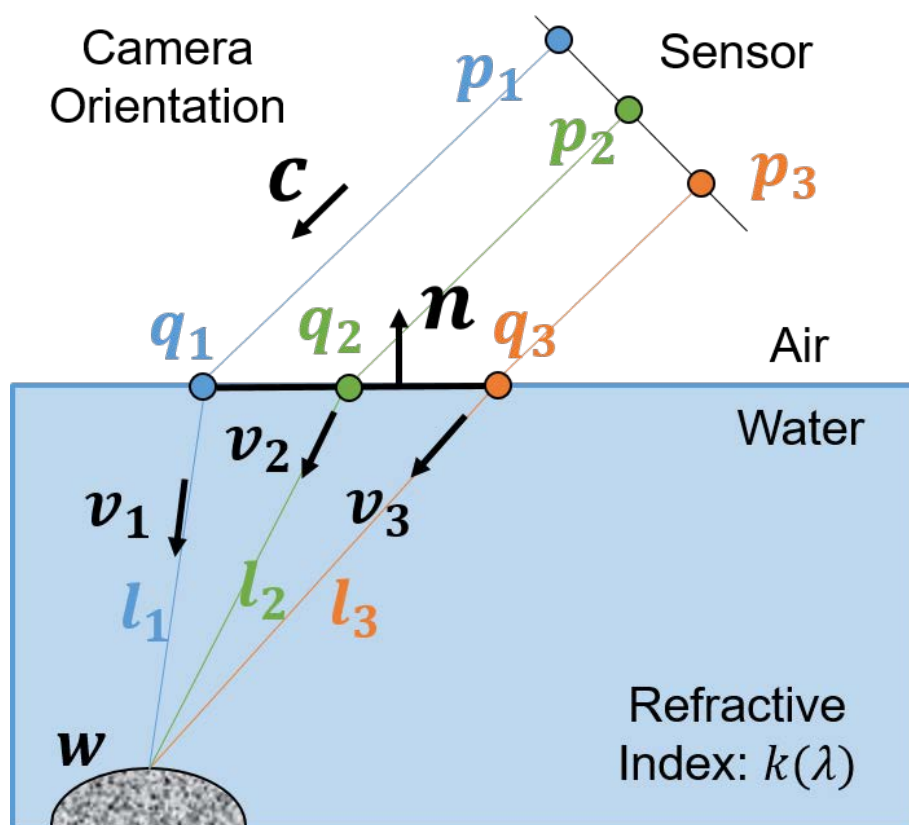


図7: 水面での屈折と水面法線および水深との関係。

第4章 レンズ内屈折解析による分光画像からの距離推定

本章では、本論文にて提案する「レンズ内屈折における波長依存性（色収差）の解析による分光画像からの距離推定手法」について、分光を用いた距離推定のモデルおよびアルゴリズムの詳細を述べる．また、定量評価および定性評価を通して、本手法による距離推定の有効性を検証する．

4.1 分光画像からの距離推定のモデル

4.1.1 分光画像に含まれる焦点ボケのモデル化

本手法では色収差と呼ばれる、レンズに入射する光の屈折方向がその波長ごとに僅かに変化する現象に着目する．屈折方向の変化はレンズ後方の異なる位置に光線を生じるため、レンズの焦点距離は波長の関数として表される．本手法では、一台のマルチスペクトルカメラを用いて多波長の画像を同時に記録し、一枚の多波長画像から得られる焦点距離および焦点ボケの波長依存性と対象の距離との関係を定式化する．

提案システムにて撮影される分光（マルチスペクトル）画像 $I \in \mathbb{R}^{M,N,L}$ （ここで、 M, N は画像の垂直および水平方向の画素数、 L は記録される波長数を表す）は次式のようにモデル化される：

$$I(x, y, \lambda) = k(x, y, \sigma(\lambda)) * P(x, y, \lambda). \quad (1)$$

ただし、 x, y は画像の空間的な座標、 λ は観測波長のチャンネル、 k は焦点ボケを表す空間方向のボケフィルタ、 P は対象のテクスチャである．式 1 は焦点ボケ k とテクスチャ P の両方が観測波長 λ に依存して変化することを表しており、 k に含まれる $\sigma(\lambda)$ はボケの強さを表す項である．

焦点ボケのモデルとしてガウシアンボケモデルを仮定すると、 k は次式のように表すことができる：

$$k(x, y, \sigma(\lambda)) = \frac{1}{\sigma(\lambda)} \exp\left(-\frac{x^2 + y^2}{2\gamma^2\sigma(\lambda)^2}\right). \quad (2)$$

$\gamma\sigma(\lambda)$ はガウス関数の分散であり，ガウシアンボケの強さを表す．

ここで，前章にて既に紹介したが，色収差によって生じる焦点ボケとレンズ光学系との関係を図8に再掲する．対象物とセンサ間の距離を Z ，観測波長 λ における光線の焦点位置を $\mu_f(\lambda)$ ，レンズとセンサ間の距離を μ_s としている． μ_s からの $\mu_f(\lambda)$ のズレは焦点ボケを引き起こし，レンズの色収差は光を波長依存的に屈折させ，レンズの焦点距離 $f(\lambda)$ の変化をもたらす．図8の光学系について，薄レンズモデルにもとづくレンズの公式により，次式が与えられる：

$$\frac{1}{Z} + \frac{1}{\mu_f} = \frac{1}{f}. \quad (3)$$

さらに，[2, 18] と同様に，センサ面におけるボケの半径を $\sigma(\lambda)$ で近似することにより，相似の関係から $\sigma(\lambda)$ とシーンの深度 Z ，レンズとセンサ間の距離 μ_s ，波長依存を含むレンズの焦点距離 $f(\lambda)$ との関係が次式のように表される：

$$\sigma(\lambda) = \pm \left(\frac{1}{Z} - \frac{1}{f(\lambda)} \right) \mu_s \pm 1, \quad (4)$$

ここで，正の符号は基準波長よりも短い波長，負の符号は基準波長よりも長い波長を表す．

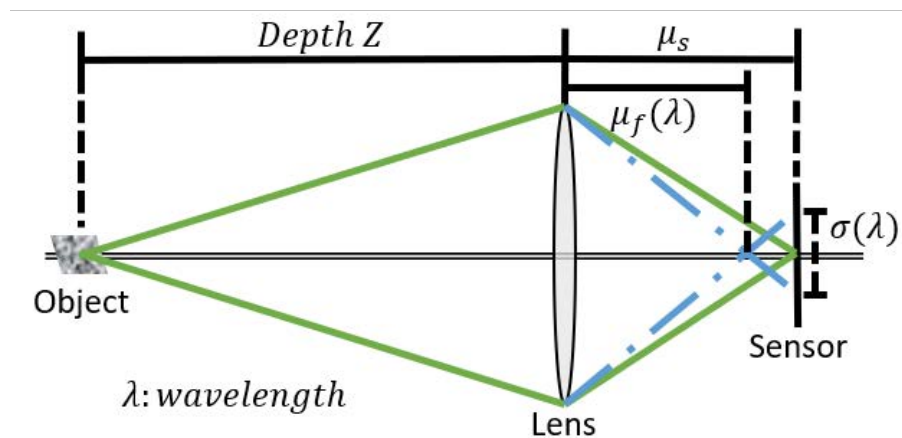


図8: 単レンズ光学系における焦点ボケに関連する物理パラメータ．

4.1.2 焦点ボケからの深度の定式化

続いて，マルチスペクトル画像の異なる波長で焦点ボケとして見える焦点距離 f の波長依存性の変化を利用して，マルチスペクトル画像 I の空間微分 ΔI とスペクトル

微分 $\partial_\lambda I$ から深度 Z の関係を導出し、深度に対する閉形式の非反復解を提示する．スペクトル次元に関するマルチスペクトル画像の微分は次のようになる：

$$\partial_\lambda I = \partial_\lambda (k * P) = \partial_\lambda k * P. \quad (5)$$

ここで、簡単のために、マルチスペクトル画像の全てのチャンネルで反射率が一定である場合、たとえば白い表面を対象とした場合を考える（この制限は後ほど緩和される）．このとき、ガウシアンボケの仮定は、マルチスペクトル画像のスペクトル方向の微分 $\partial_\lambda I$ がテクスチャ P とは無関係に空間微分 ΔI に比例することを意味する．すなわち、ラプラシアン演算子 Δ は空間成分のラプラシアンを表し、空間微分 ΔI に関する以下の式が得られる：

$$\Delta I = \Delta k * P = \frac{\partial_\lambda k}{\gamma^2 \sigma \partial_\lambda \sigma} * P = \frac{\partial_\lambda I}{\gamma^2 \sigma \partial_\lambda \sigma}. \quad (6)$$

ここで、 $\partial_\lambda \sigma$ は $\sigma(\lambda)$ のスペクトル方向の微分である．式 4 から、 $\sigma(\lambda)$ のスペクトル方向の微分は次のように表される：

$$\partial_\lambda \sigma(\lambda) = \pm \frac{\mu_s}{(\partial_\lambda f(\lambda))^2}. \quad (7)$$

式 6 について、式 4 から得られる σ と式 7 から得られる $\partial_\lambda \sigma$ を代入し、深度 Z について解くことによって、深度と画像微分との関係式を得ることができる：

$$\begin{aligned} Z(\alpha(\lambda), \beta(\lambda)) &= \frac{\alpha(\lambda) \Delta I}{\beta(\lambda) \Delta I - \partial_\lambda I}, \\ &\text{ただし} \\ \alpha(\lambda) &= \frac{\gamma^2 \mu_s^2}{(\partial_\lambda f(\lambda))^2}, \\ \beta(\lambda) &= \left(1 - \frac{\mu_s}{\partial_\lambda f(\lambda)}\right) \frac{\gamma^2 \mu_s}{(\partial_\lambda f(\lambda))^2}. \end{aligned} \quad (8)$$

式 23 は深度 Z と、入力マルチスペクトル画像 I の空間微分 ΔI および波長微分 $\partial_\lambda I$ との関係を表している．ここで、 $\alpha(\lambda)$ と $\beta(\lambda)$ は観測波長ごとのレンズパラメータであり、本手法ではキャリブレーションによってこれらのパラメータを決定する．

4.1.3 空間微分および波長微分の計算

式 23 において、深度と焦点ボケの関係は、空間次元とスペクトル次元に沿った微分が得られることを前提としている。本節では、空間微分および波長微分を入力画像から計算するための手法について述べる。

空間微分、すなわちラプラシアン計算は、ラプラシアンフィルタとしてよく知られるフィルタを入力画像に畳み込むことによって計算される。水平方向および垂直方向の画素値の一次微分 I_x, I_y は、次式で表される：

$$\begin{aligned} I_x(x, y) &= I(x+1, y) - I(x, y), \\ I_y(x, y) &= I(x, y+1) - I(x, y). \end{aligned} \quad (9)$$

同様に、二次微分は次のように表される：

$$\begin{aligned} I_{xx}(x, y) &= \{I(x+1, y) - I(x, y)\} - \{I(x, y) - I(x-1, y)\}, \\ I_{yy}(x, y) &= \{I(x, y+1) - I(x, y)\} - \{I(x, y) - I(x, y-1)\}. \end{aligned} \quad (10)$$

これを用いて、入力画像のラプラシアン $\Delta I(x, y)$ は次のように表される：

$$\begin{aligned} \Delta I(x, y) &= I_{xx}(x, y) + I_{yy}(x, y), \\ &= I(x+1, y) + I(x-1, y) + I(x, y+1) + I(x, y-1) - 4I(x, y). \end{aligned} \quad (11)$$

これを画像フィルタのカーネル k_l として記述すると次のように表される：

$$k_l = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad (12)$$

図 9 に示すように、ラプラシアンフィルタをマルチスペクトル画像 I の各波長画像に適用することにより、チャンネル数分の $\Delta I(x, y, \lambda)$ を得る。詳細は後述するが、距離推定に用いる波長ペア (λ_1, λ_2) を選択し、その 2 波長画像の空間微分の平均値を ΔI として用いる。

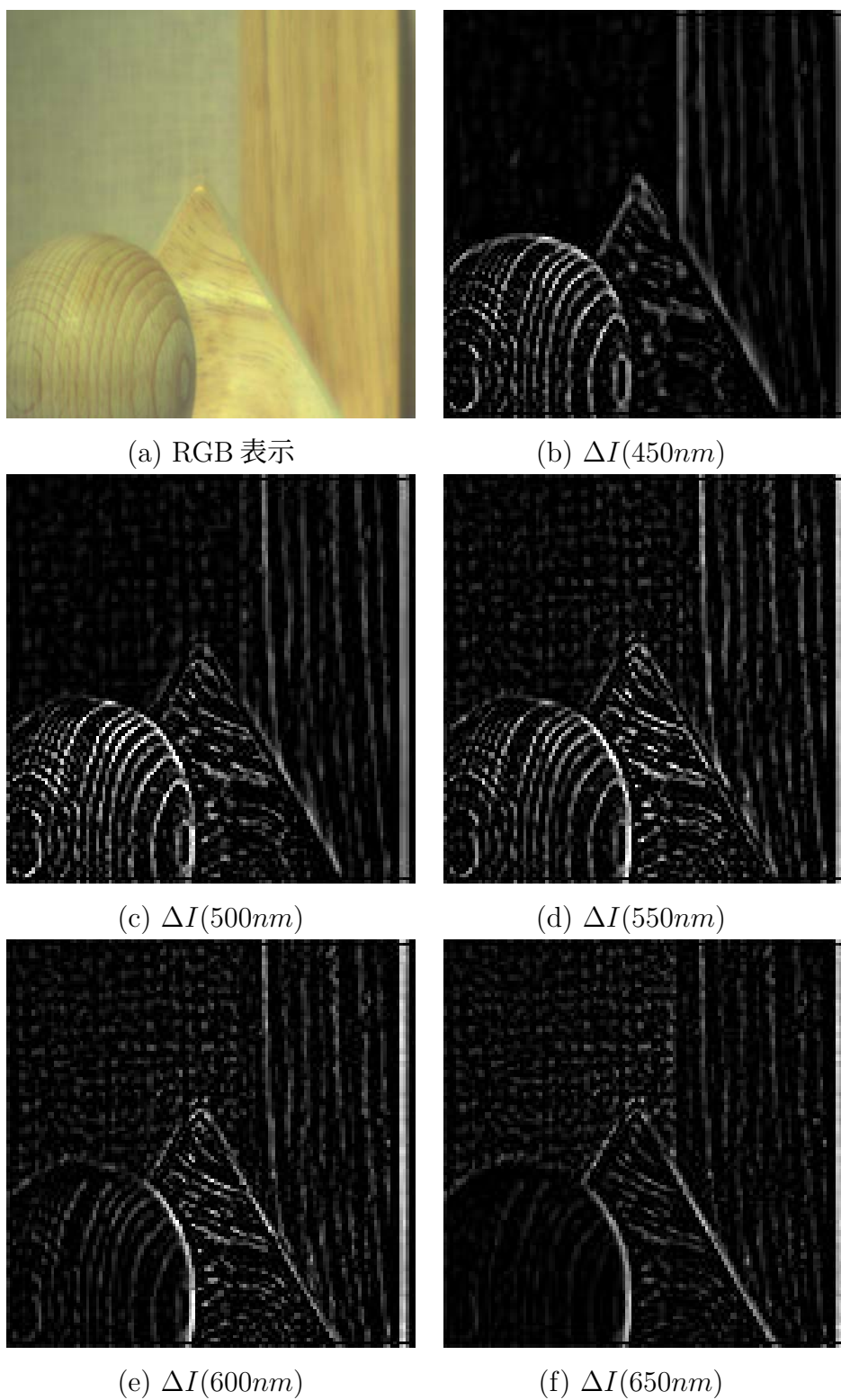


図 9: 空間ラプラシアンフィルタを適用した様子.

他方、分光反射率が可変の場合、テクスチャ P は波長依存となる。さらに式 5 は、入力画像 I のスペクトル方向に関する微分が波長依存のテクスチャ P に依存することを意味する。ここでは、最新の多焦点画像の合成手法 [20] を用いて、入力画像の波長微分 $\partial_\lambda I$ と全焦点マルチスペクトル画像 $\hat{P}$ を近似する方法を示す。

はじめに、波長方向の次元を持たない全焦点画像 $\hat{P}_m(x, y)$ を考える。我々はこの全焦点画像をシャープネスにもとづく各波長画像の重みづけ和として計算する。

$$\hat{P}_m(x, y) = \sum_{\lambda \in L} w(x, y, \lambda) I(x, y, \lambda). \quad (13)$$

ここで、重み w は正規化されたシャープネスの尺度であり、ある位置 (x, y) とスペクトル成分 λ のシャープネス $S(x, y, \lambda)$ は、空間方向の画像勾配 I_x と I_y を用いて、入力画像の変動の総和として計算される。

$$S(x, y, \lambda) = \sqrt{I_x^2(x, y, \lambda) + I_y^2(x, y, \lambda)}. \quad (14)$$

$w(x, y, \lambda)$ は空間平滑化フィルタ $C(x, y)$ にて平滑化され、正規化されたシャープネスとして計算される：

$$w(x, y, \lambda) = \frac{S(x, y, \lambda)}{\sum_{\lambda \in L} S(x, y, \lambda)} * C(x, y). \quad (15)$$

続いて、マルチスペクトルの全焦点画像 $\hat{P}(x, y, \lambda)$ をグレースケールの全焦点画像 $\hat{P}_m(x, y)$ から近似することを考える。表面の模様は波長に依存しないことを仮定して、全焦点画像の各波長成分は、グレースケールの $\hat{P}_m(x, y)$ と、 $a(x, y, \lambda)$ で重み付けされた各波長成分のそれぞれの寄与の組み合わせとして近似する：

$$\hat{P}(x, y, \lambda) = a(x, y, \lambda) \hat{P}_m(x, y), \quad \text{ただし} \quad (16)$$

$$a(x, y, \lambda) = \sum_{m, n \in W(x, y)} \frac{I(m, n, \lambda)}{\hat{P}_m(m, n)}.$$

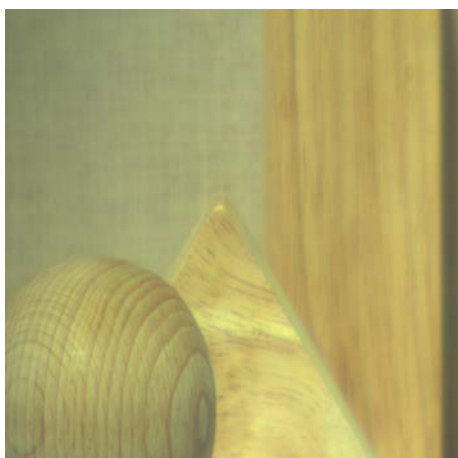
係数 $a(x, y, \lambda)$ は、与えられたスペクトル成分の寄与が $\hat{P}_m$ に対して小さな長方形の近傍 W に存在しているかを表す。全焦点画像 $\hat{P}$ から、元の画像 I と推定された $\hat{P}$ との差として、与えられた波長でのボケ量を計算する：

$$B(x, y, \lambda) = I(x, y, \lambda) - \hat{P}(x, y, \lambda). \quad (17)$$

最終的に、波長方向のボケの微分 $\partial_\lambda I$ は、次のように近似される：

$$\partial_\lambda I(x, y, (\lambda_i, \lambda_j)) \approx B(x, y, \lambda_i) - B(x, y, \lambda_j). \quad (18)$$

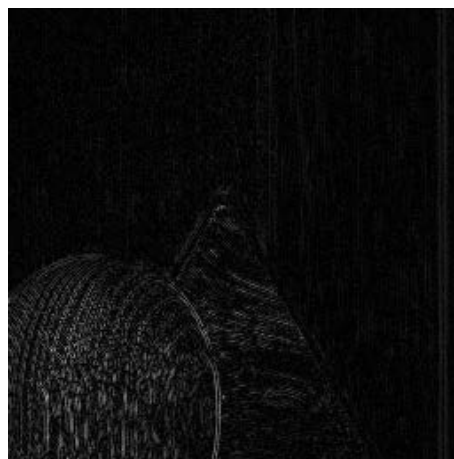
近似された波長微分の様子を図 10 に示す。



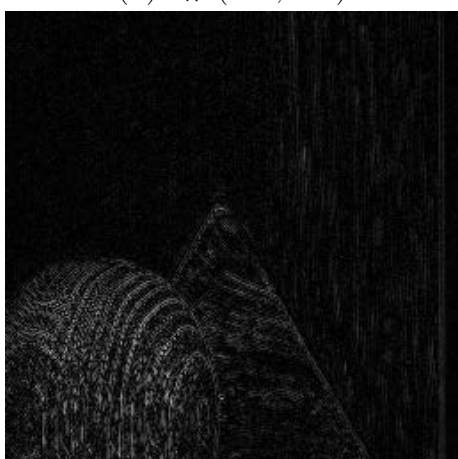
(a) RGB 表示



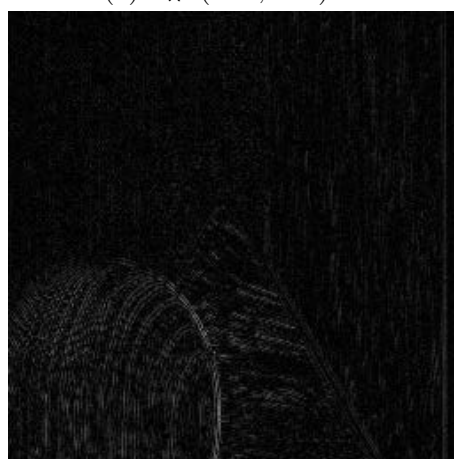
(b) $\partial_{\lambda}I(450, 500)$



(c) $\partial_{\lambda}I(500, 550)$



(d) $\partial_{\lambda}I(550, 600)$



(e) $\partial_{\lambda}I(600, 650)$

図 10: 隣接波長に対する波長微分を適用した様子.

4.1.4 波長の選択

深度推定の精度を向上させるためには，入力のマルチスペクトル画像から適切な波長ペア (λ_i, λ_j) を選択することが重要である．前章にて示したように，波長の変化に対するガラスの屈折率変化は可視光域では単調であることが知られている．波長の近いペアでは，わずかなボケの変化を観察することができるため，深度分解能の向上が期待できる．反対に，波長差の大きいペアを使用すると，ボケの違いを観察できる深度範囲が広がり，作動範囲の拡大につながる．

3つ以上の波長画像が利用可能な場合，シーンの小領域に対して，最も焦点の合った2つの波長画像を適応的に選択することで，広い作動範囲と正確な深度推定の両方の利点を得ることを提案する．これには，小領域の空間的に最もシャープな画像を持つ波長を第一の波長として選択することが合理的である．そこで，2つの隣り合う波長画像のシャープネスが最も高いものと，2番目に高いものを波長ペア (λ_i, λ_j) として選択する．すなわち，最も高い重みのスコア $w(x, y, \lambda_i)$ を持つ波長と，2番目に高いスコア $w(x, y, \lambda_j)$ を持つ波長を (x, y) ごとに決定する．

図 11 に，シーンに対して選択された第一，第二の波長を表すマップを示す．これから，シーンの小領域ごとに用いられる波長が適応的に選択されていることがわかる．

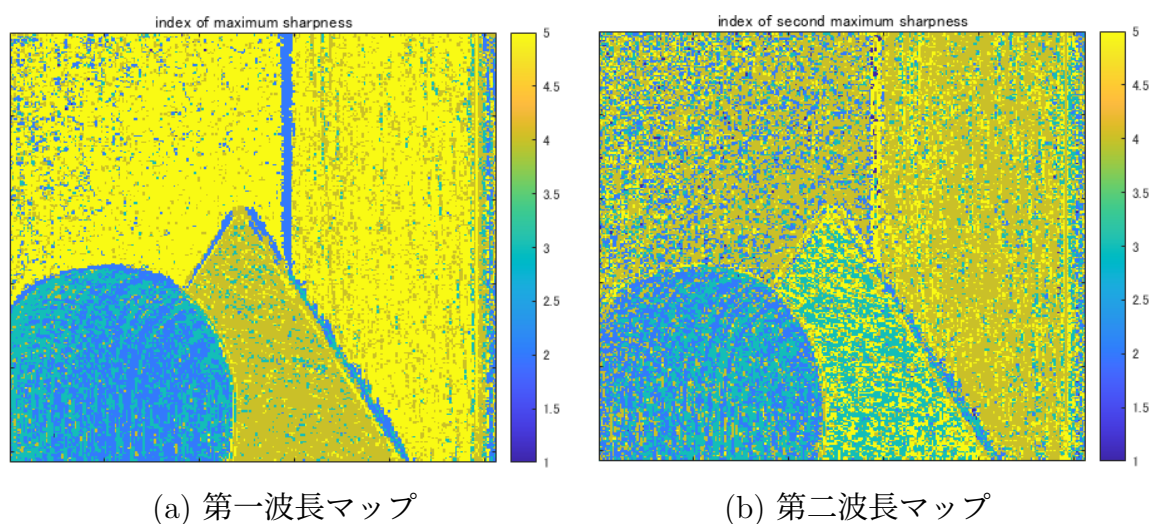


図 11: 第一，第二波長として選択された波長のマップ。

4.2 評価実験

4.2.1 実験環境およびパラメータキャリブレーション

実験には、市販のバンドパスフィルタを複数搭載可能なマルチスペクトルカメラ (EbaJapan SI6) と、Edmund Optics DCX-25-X-100-VIS-NIR-CTD レンズを用いて、図 12 のように実験を行った。レンズの性能はメーカー表示で波長 587nm における焦点距離が 100mm であり、 f 値は 4 である。レンズは C マウントチューブを使用してカメラに装着している。マルチスペクトルカメラは、450nm から 650nm の波長範囲の光を 50nm 間隔で 512×512 ピクセルの解像度で撮影する。小領域 W の値は 15×15 ピクセルに設定している。

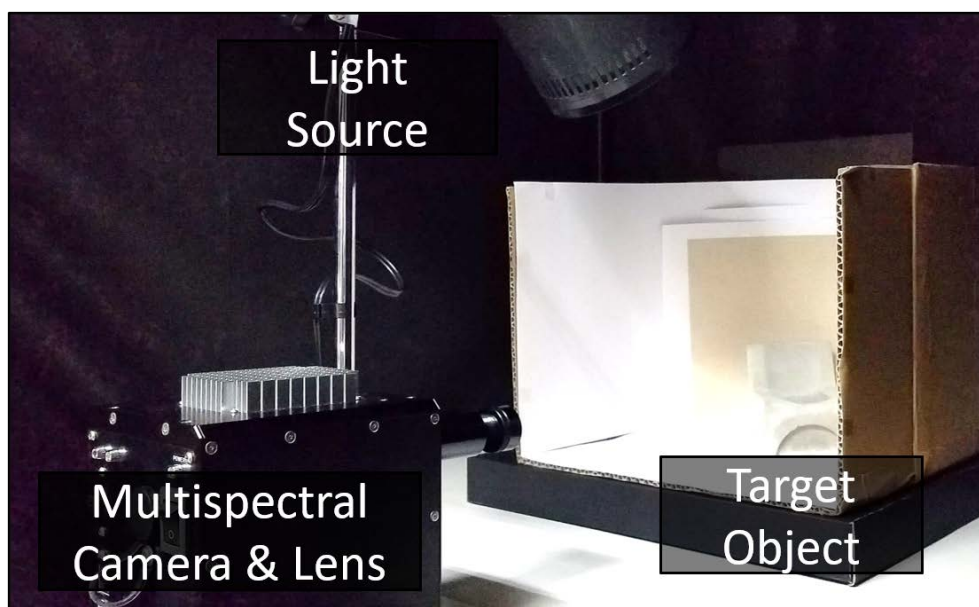


図 12: カメラ，光源，被写体等の実験環境.

実験において、パラメータ $\{\alpha(\lambda), \beta(\lambda)\}$ の値はレンズ依存であり、各波長で固有の値を持つ。我々は、これらのパラメータの値を深度の真値 Z^* を用いて事前に推定する。テクスチャを持つ平面の板を 90cm から 110cm 離れた場所に 2cm 間隔で設置して撮影し、実際の推定過程を適用して推定値が真値に近づくようにパラメータを調整する。実際には、それぞれの波長 $\lambda \in L$ ごとに、 $\|Z^* - Z(\alpha, \beta)\|_2^2$ を最小化するパラメータ $\{\alpha(\lambda), \beta(\lambda)\}$ を最適化によって推定する。

4.2.2 定量評価

定量評価では初めに，カラーセンサの分光帯域の幅が深度推定精度にどのように影響するかを，広帯域 RGB センサと狭帯域 RGB センサを用いて比較する．広帯域の RGB カメラとして，スナップショットマルチスペクトル（RGB + 近赤外）カメラ *Spectral Devices Inc. MSC-RGBN-1-A* の RGB チャンネルのみをデモザイク処理を行わずに用いる．広帯域 RGB カメラの感度は、図 13(c) に示すものである．近赤外の感度については本実験では扱わない．狭帯域 RGB カメラは，図 13(d) に示すバンドパスフィルタの内，第一波長（450nm），第三波長（550nm），第五波長（650nm）を使用することで実現している．ここで，広帯域 RGB カメラと狭帯域 RGB カメラの各チャンネルの中心波長はほぼ同一である．

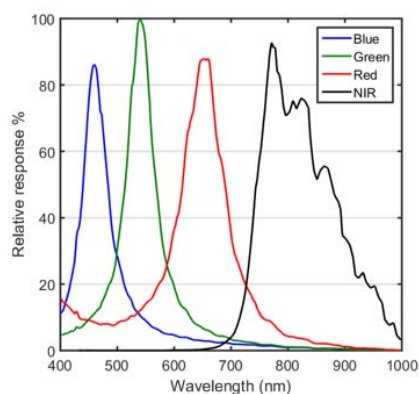
さらに，3バンドと5バンドのマルチスペクトルセンサを用いて，前項で述べた適応的な波長選択手法の効果を検証した．3バンドマルチスペクトルセンサは狭帯域 RGB カメラと同じものを使用し，5バンドマルチスペクトルセンサは図 13(d) に示す番号付けされているバンドパスフィルタを全て搭載することで実現した．



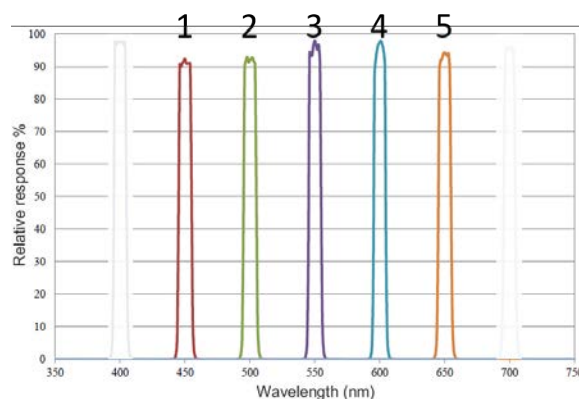
(a) *Spectral Devices Inc. MSC-RGBN-1-A*



(b) *EbaJapan SI6*



(c) *MSC-RGBN-1-A* センサ



(d) *SI6* カメラに搭載のバンドパスフィルタ

図 13: 広帯域 RGB カメラと狭帯域マルチバンドカメラ，およびそれらの波長応答．

定量評価に用いるシーンは、キャリブレーションと同様に、90cm から 110cm までの距離を 2cm 間隔で変化させた複数の平面的なテクスチャパターンで構成される。また、キャリブレーションと同様の手順で、図 14 で示した平面テクスチャパターンをカメラから少しずつ移動させて計測を行った。Pattern 1 はパラメータキャリブレーションにも用いている。Pattern 1 および Pattern 2 は不規則なパターンであり、Pattern 3 および Pattern 4 は規則的に近いパターンである。本実験では広帯域センサと狭帯域センサの帯域幅が異なることによるセンサの受光量にもとづくノイズの影響を考慮し、広帯域センサの画像では SNR が約 40[db] となるように弱ノイズを付与し、狭帯域センサの画像では SNR が約 30[db] となるよう強ノイズを付与して定量評価を行った。

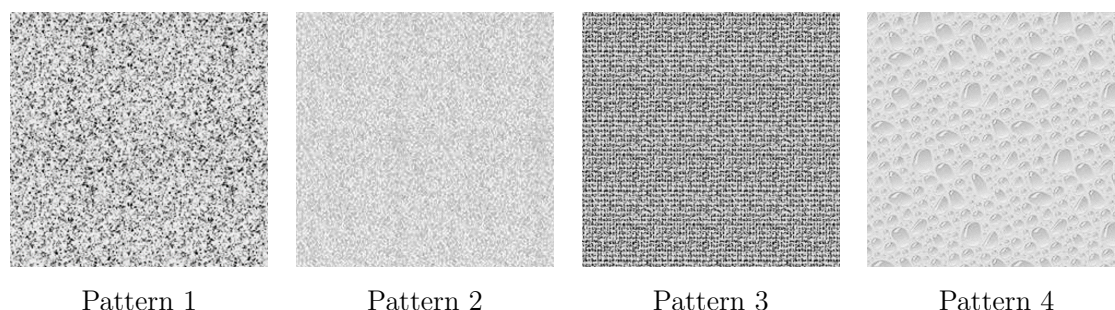


図 14: 定量評価実験に用いたテクスチャパターンの様子。

表 1 と図 15 に、各テクスチャについてのカメラ毎の深度推定結果を示す。表 1 には、計測範囲全域の平均二乗誤差と標準偏差の平均値を示している。図 15(a) は、市販の RGB カメラと同等である広帯域の RGB 感度を持つセンサに対して提案手法を適用した結果である。(a) の結果から、広帯域の RGB センサを用いた場合の推定値はおおむね真値に沿っているが、誤差がパターンに依存することが観察された。図 15(b) は、狭帯域の RGB 感度を持つセンサに対して提案手法を適用した結果である。(b) の結果から、狭帯域 RGB 感度のカメラを使用した場合、広帯域 RGB 感度のカメラを使用した場合よりも良好な推定値が得られることがわかる。また、狭帯域の RGB センサを使用する利点として、推定値の標準偏差が広帯域の RGB センサに比べてはるかに小さいということが挙げられる。図 15(c) は、狭帯域の 5 波長のチャンネル（狭帯域 RGB およびその中間の 2 波長）のマルチスペクトルカメラを使用した場合の推定結果を示す。(c) の結果から、狭帯域のマルチバンドカメラを使用した場合、推定値はより長い範囲で真値に沿っており、他の 2 種類の結果と比較して、より正確な結果が得られることがわかる。また、推定値の標準偏差は狭帯域 RGB カメラの結果と同程度に小さく、標準偏差の観点でも良好な結果を示している。表 1 の結果から、誤差、標準偏差、パターン依存性について、5 チャンネルのマルチスペクトルセンサ

を使用した場合に最も小さくなることが確認されている。以上の定量評価より、狭帯域の感度を持つマルチバンドカメラの使用が提案手法に最適であると結論づけた。

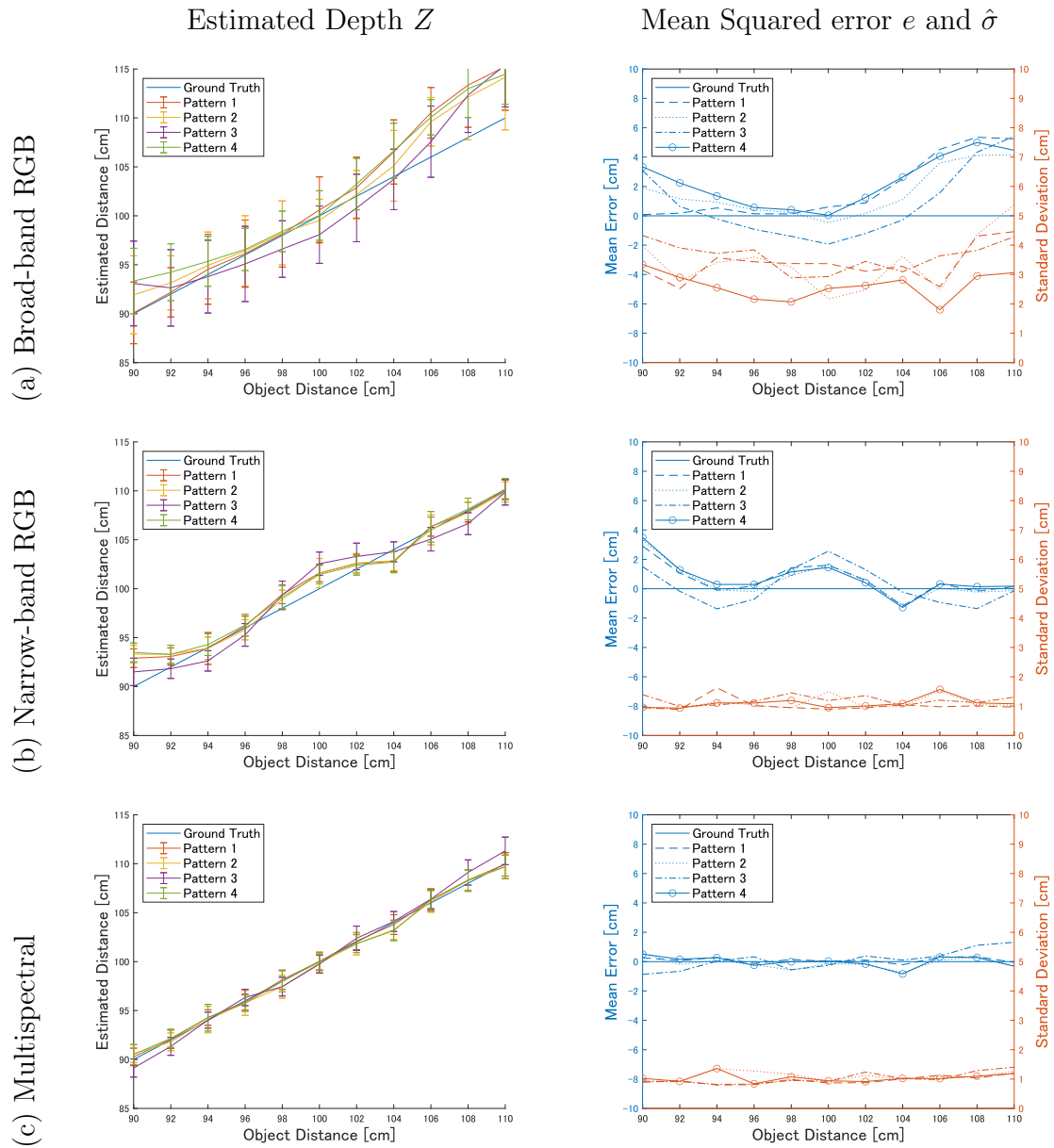
さらに、表2と図16にてシーンの色を変化させた際の距離推定への影響を調査した結果を示す。評価に用いたシーンの色は図??(a)に示したものである。また、シーン色の影響を調査した際のセンサは5chのマルチバンドセンサである。表2と図16を見ると、彩度の高いシーンであっても色による距離推定への影響は僅かであるといえる。一般に彩度の高いシーンを撮影したマルチスペクトル画像はチャンネル間の輝度変化が大きいことが経験的に知られているが、このような輝度変化をもたらすシーンにおいても提案手法は安定に動作することが示された。

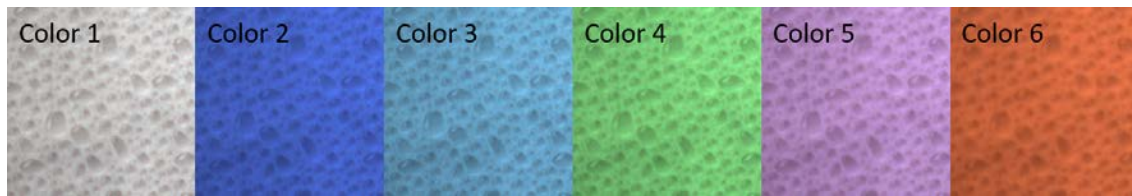
Sensor Type	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Broad-band RGB	1.835 ± 2.771	1.661 ± 2.232	1.907 ± 2.496	2.304 ± 2.834
Narrow-band RGB	1.890 ± 0.454	1.826 ± 0.527	1.702 ± 0.700	2.696 ± 0.642
Multispectral	0.108 ± 0.418	0.223 ± 0.409	0.757 ± 0.541	0.346 ± 0.514

表1: 推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ の、計測区間全域における平均値 (単位は cm) .

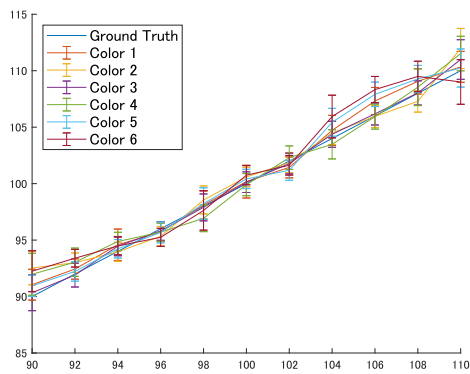
	Color 1	Color 2	Color 3	Color 4	Color 5	Color 6
Absolute mean error e	0.605	0.792	0.283	0.755	0.732	1.193
Standard deviation $\hat{\sigma}$	0.711	1.093	0.382	0.952	0.923	1.382

表2: 推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ の、計測区間全域における平均値 (単位は cm) .

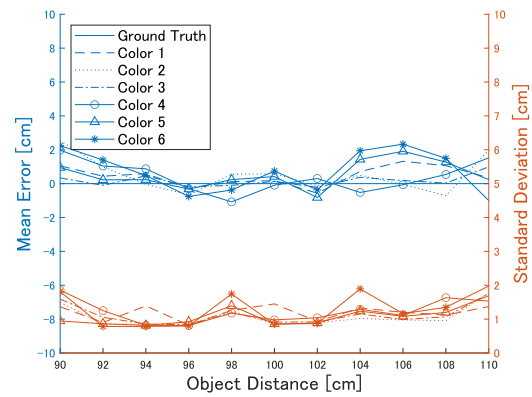
図 15: 推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ (単位は cm) .



(a) 評価実験に使用した色付きシーン（テクスチャは図 14 中の Pattern 4 を使用）.



(b)



(c)

図 16: 推定深度の平均値誤差 e および標準偏差 $\hat{\sigma}$ （単位は cm ）.

4.2.3 複雑なシーンに対する定性評価

本節では、より複雑なシーンを対象とした定性評価を行うことで、提案手法の有効性を検証する。定量評価の結果から、定性評価には5チャンネルのマルチスペクトルセンサを用いた場合の結果で検証を行う。定性評価では安定性を高めるために、各局所領域の最も焦点の合う3つの波長画像から得られる距離の平均値を使用している。

複雑な非平面シーンに対する深度推定の結果を図17から22に示す。*Concrete chair*の結果は、単純なテクスチャのみを含む平滑面に対して有効であることを示している。椅子の側面と背もたれに注目すると、前から後ろに向かって奥行きが深くなっていることがわかる。*Wood toy*の結果は、似た色の物体に対して正しい奥行き関係を推定できることを示している。球形のブロックに注目すると、推定された深さが中心から端まで正しく変化していることがわかる。*Flowers*は、色の異なるオブジェクトに対する深さ推定結果である。緑色の花が手前にあることが正しく推定されている。

さらに、*Llama*, *Christmas*「トラ」, *Tiger*に、より色や質感が複雑なシーンに対する距離推定結果を示す。*Llama*の結果では、手前の動物と奥の動物の深度が異なっていることが見て取れる。*Christmas*の結果では、トナカイのオブジェクトがツリーよりも手前であることがわかる。しかしながら、ツリーの影になっている部分では輝度が不十分であるため、推定深度にエラーが生じている。*Tiger*の結果は、対象表面に多くの斑点が見られるが、我々の手法ではその影響を受けずに深さを推定できていることがわかる。

提案手法がボケを含む入力マルチスペクトル画像から、鮮明なマルチスペクトル全焦点画像 $\hat{P}$ を推定できることを図23に示す。*Wood toy*, *Concrete chair*, *Flowers*の局所領域に着目して、元画像と推定された全焦点画像とを比較する。*Wood toy*の結果では、木目のテクスチャが鮮明になっていることが確認できる。*Concrete chair*の結果では、背景との境界線や、椅子側面のテクスチャが鮮明になっていることが確認できる。*Flowers*の結果では、花の輪郭や背景のテクスチャが鮮明になっていることが確認できる。

続いて、*Llama*, *Christmas*, *Tiger*の顔領域に着目して、元画像と推定された全焦点画像とを比較する。*Llama*の結果では、頭の模様や輪郭がはっきりしていることが確認できる。*Christmas*の結果では、トナカイの顔とその後ろの木が鮮明になっていることが確認できる。*Tiger*の結果では、体表の毛がより鮮明になっていることが確認できる。

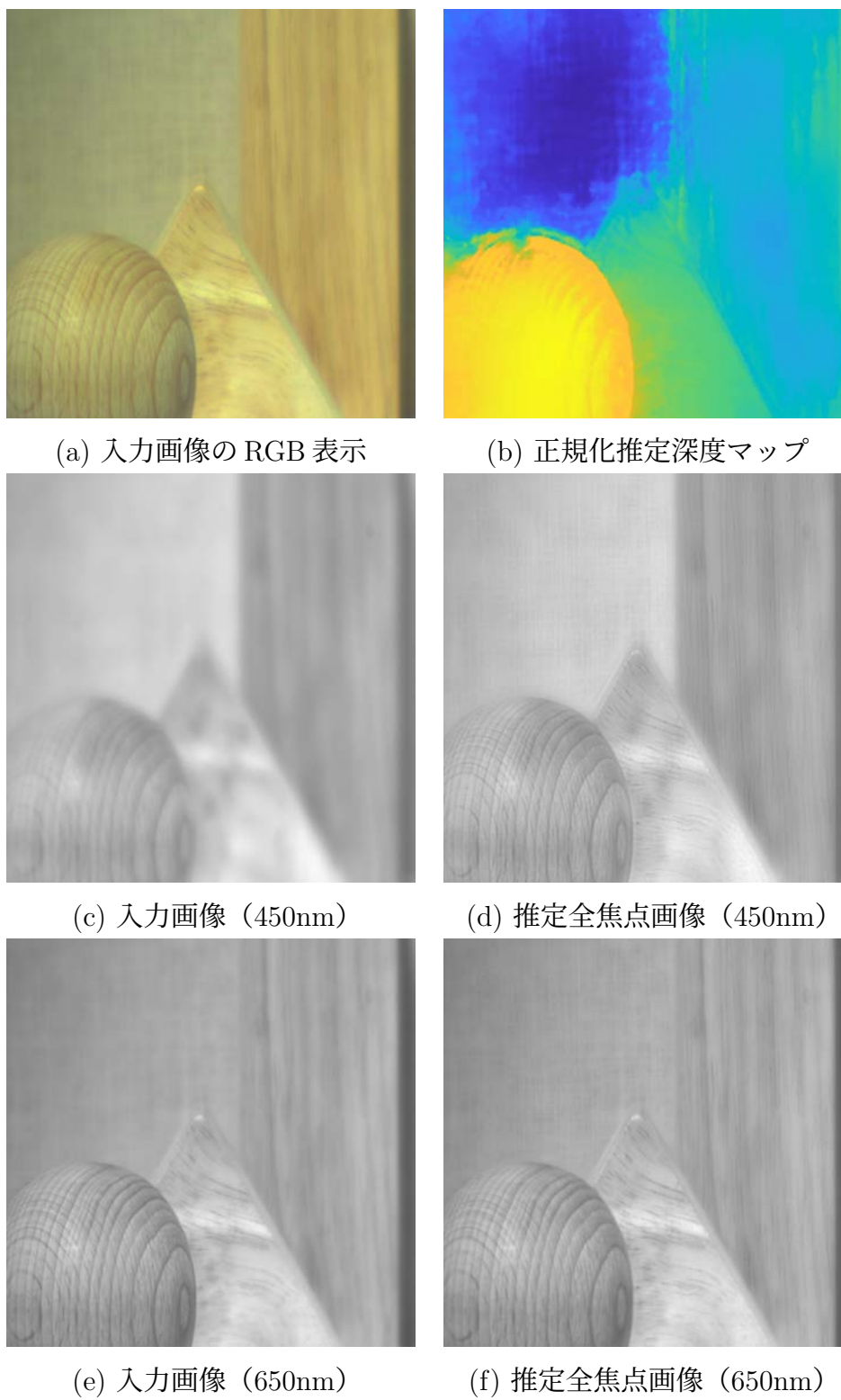


図 17: Wood toy (積み木) の例.

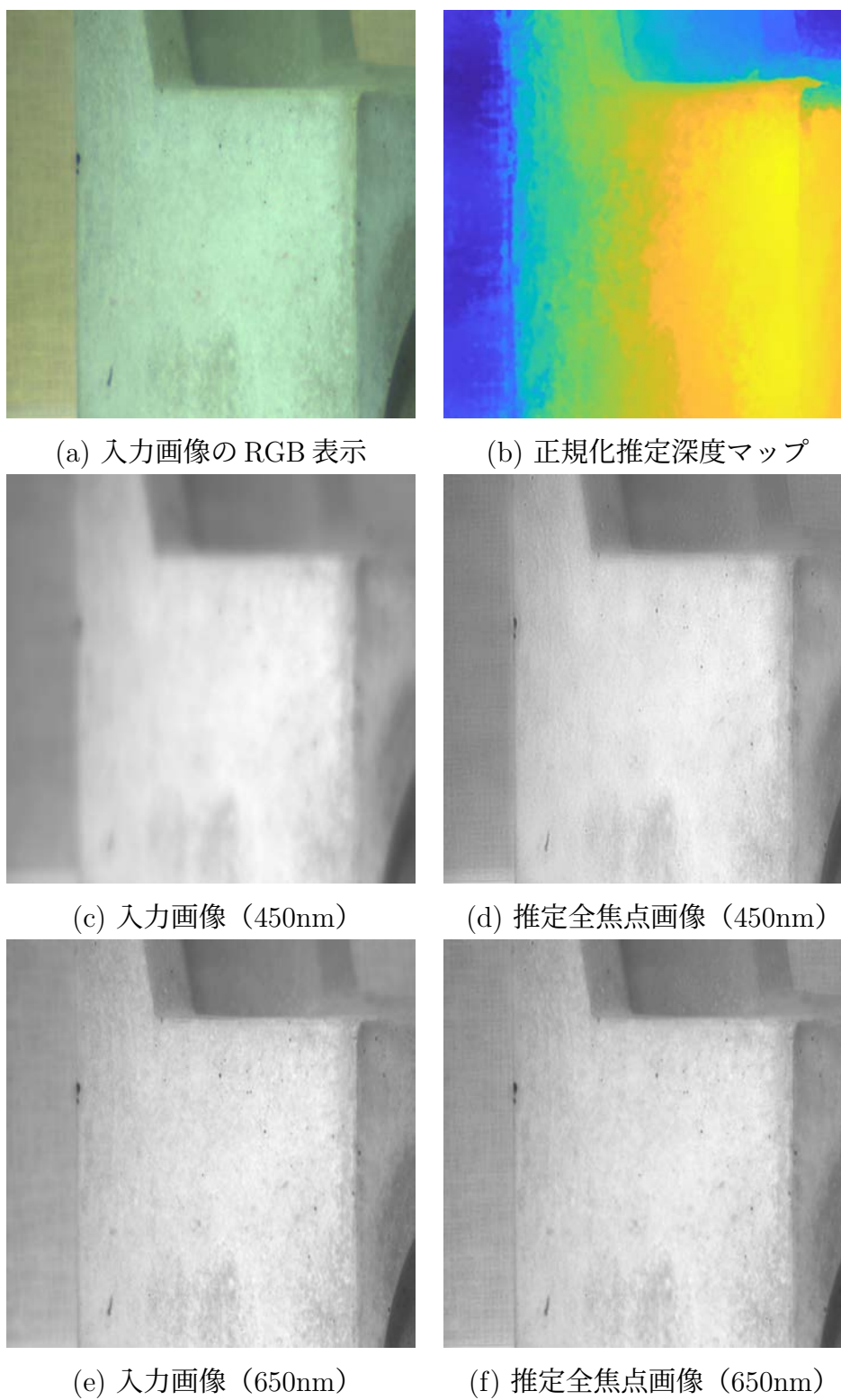


図 18: Concrete chair (椅子) の例.

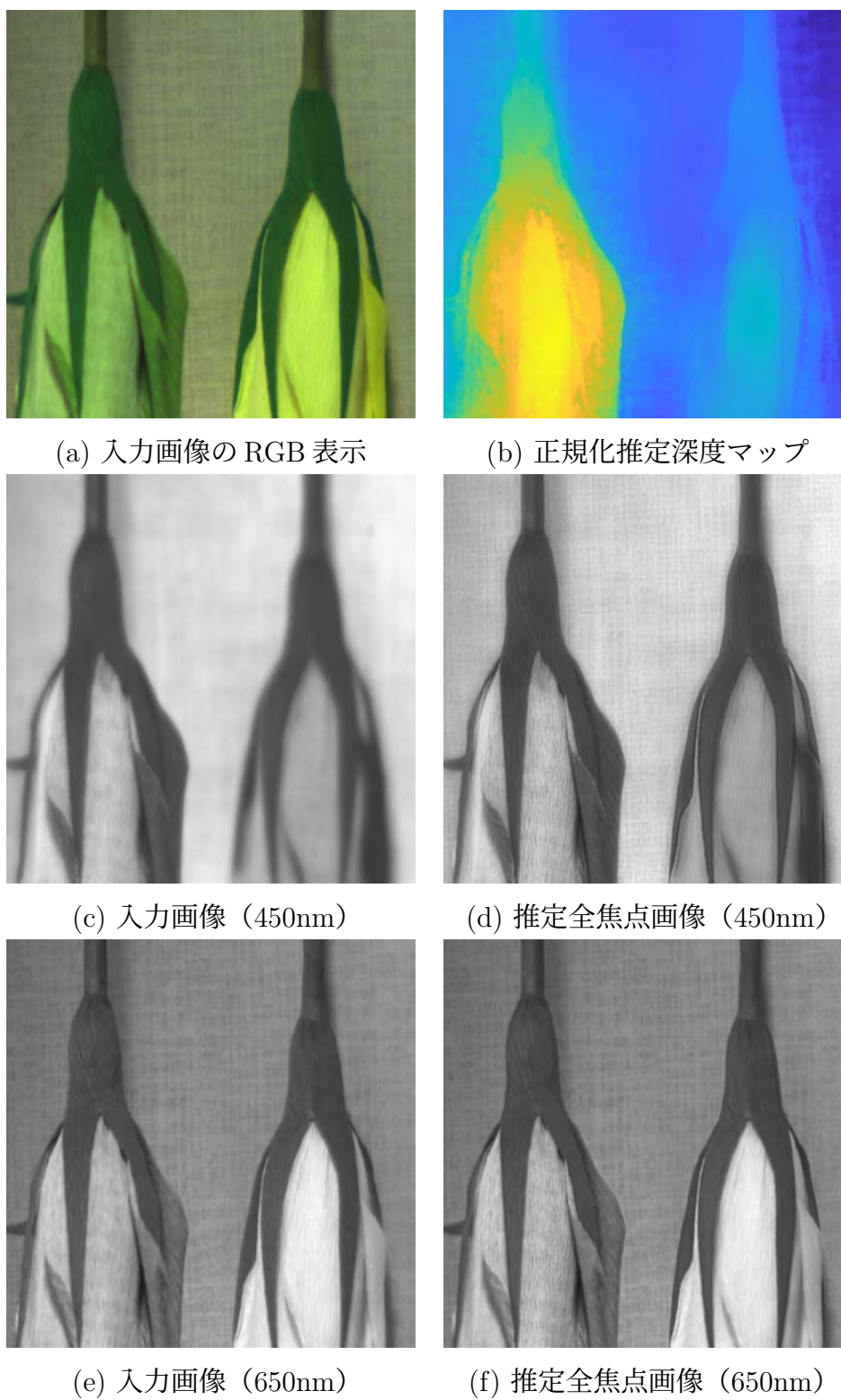


図 19: Flowers (花) の例.

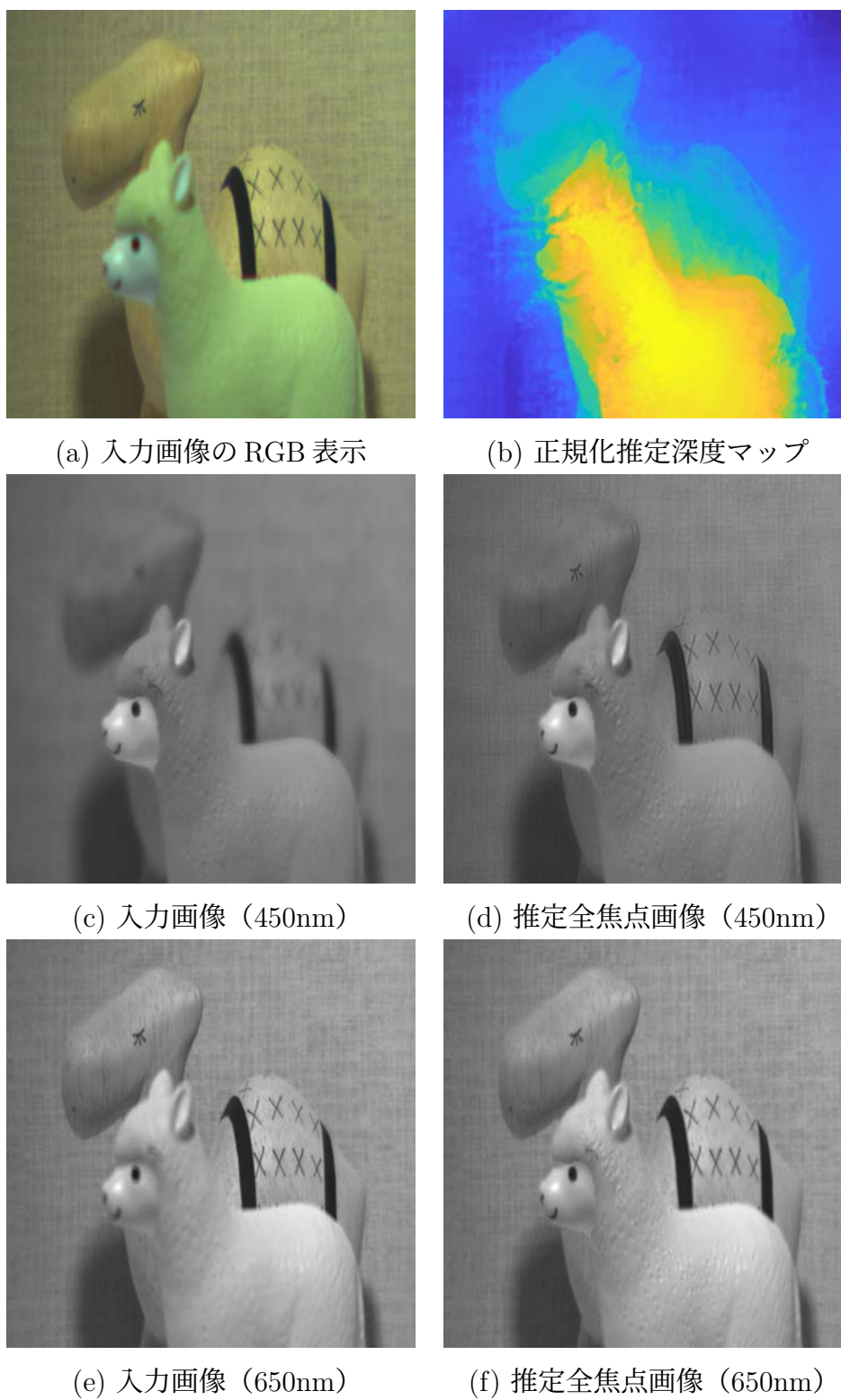


図 20: Llama (ラマ) の例.

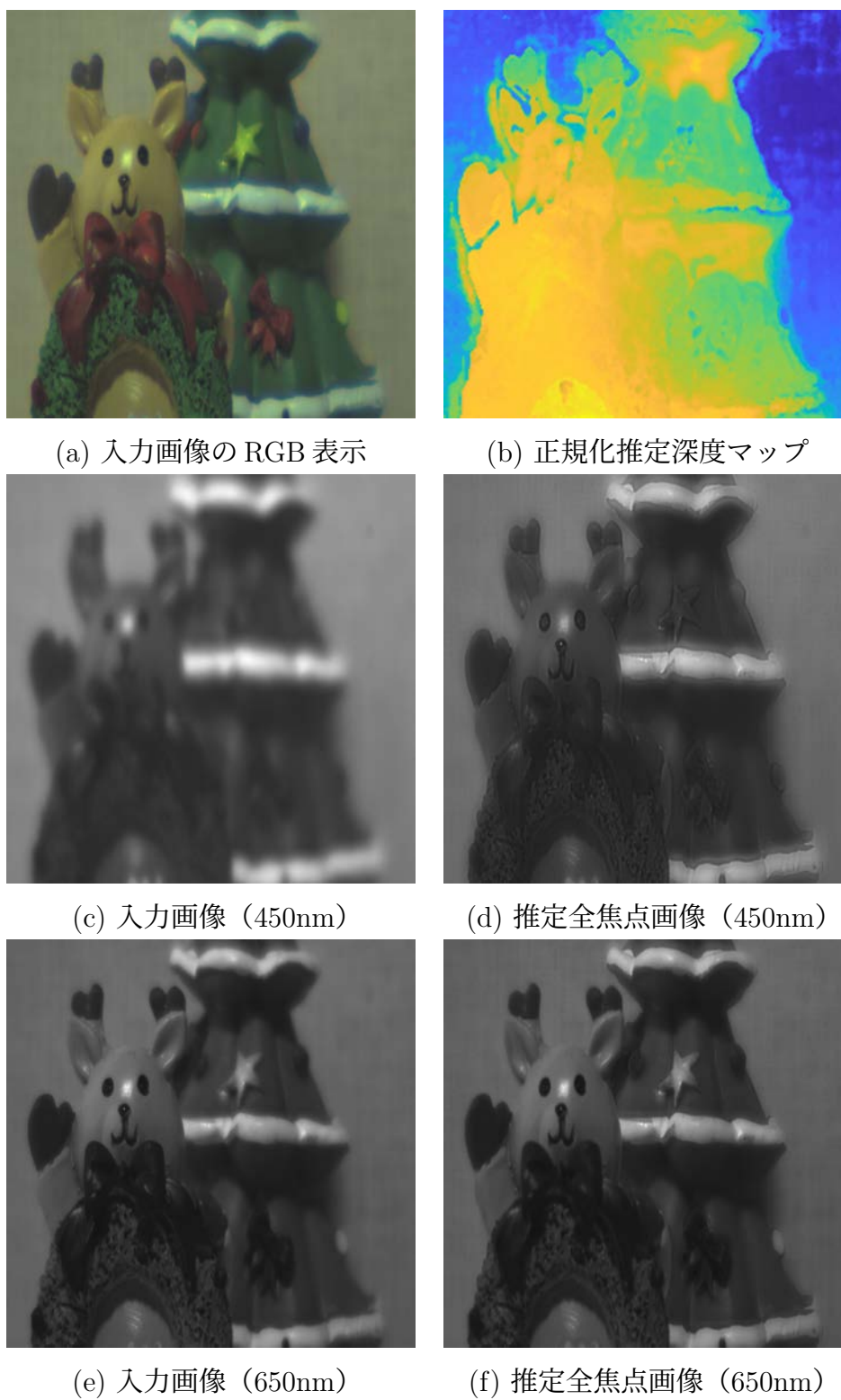


図 21: Christmas (クリスマス) の例.

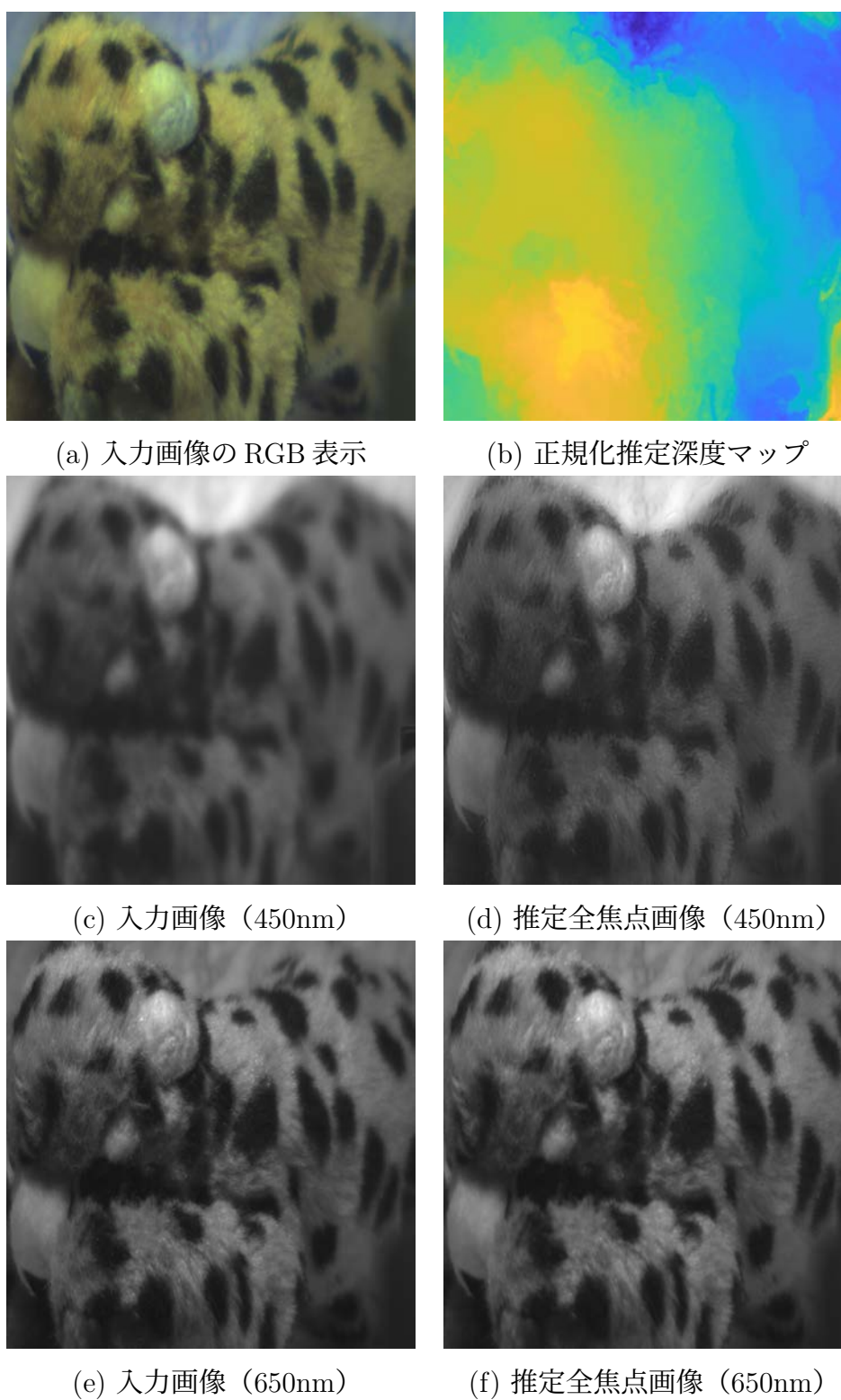


図 22: Tiger (トラ) の例.

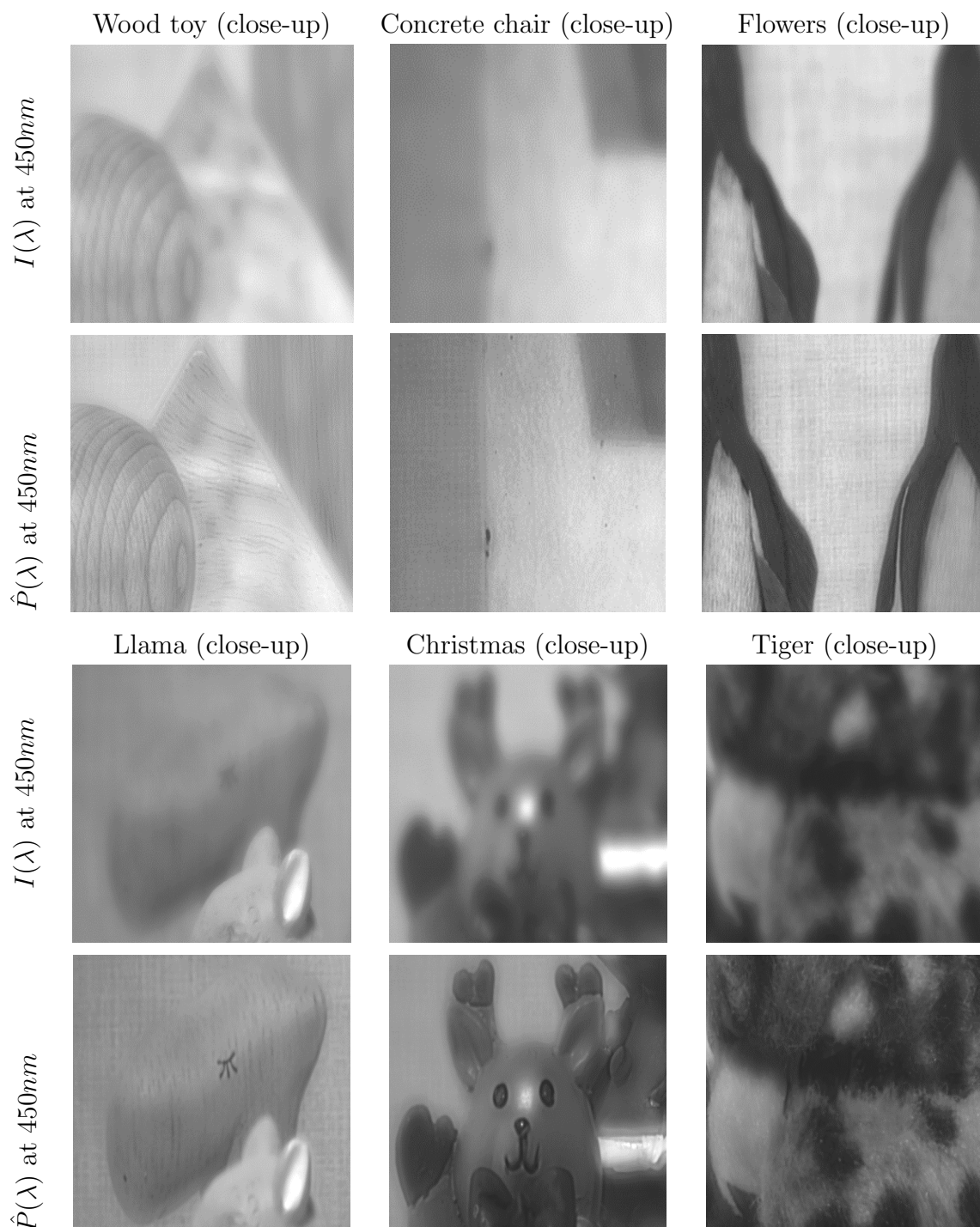


図 23: 焦点ボケを含む入力画像と、推定された全焦点画像の比較.

4.2.4 ノイズを含むシーンに対する定性評価

本節では、定性評価で用いたシーンに対してノイズを付与することで、提案手法の耐ノイズ性を検証する．図 24 から図 33 に、様々なシーンに対してノイズを付与した際の推定全焦点画像および推定深度マップを示す．

推定深度マップに着目すると、全体を通して $\sigma_n = 0.001$ 程度のノイズであれば推定深度への影響が少ないが、 $\sigma_n = 0.01$ では多くの例において推定深度への影響が強いといえる．推定全焦点画像に着目すると、入力画像に付与したノイズ自体が全焦点画像では見えづらくなっている例が見られた．これにより、入力画像と全焦点画像の差分として計算していた波長微分の項にノイズが影響し、推定深度への影響が強くなっていると考えられる．図 33 の *Tiger* の例では強さが $\sigma_n = 0.01$ であるにもかかわらずノイズの影響が少ない．このことから、動物の毛のように非常に細かいテクスチャを有する対象ではノイズの影響が比較的少ないことが考えられる．

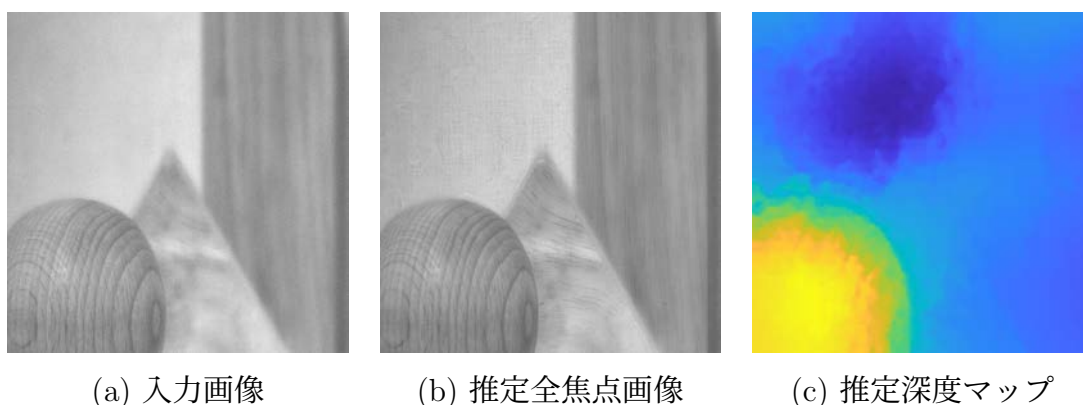


図 24: Wood toy (積み木) の例 ($\sigma_n = 0.001$) .

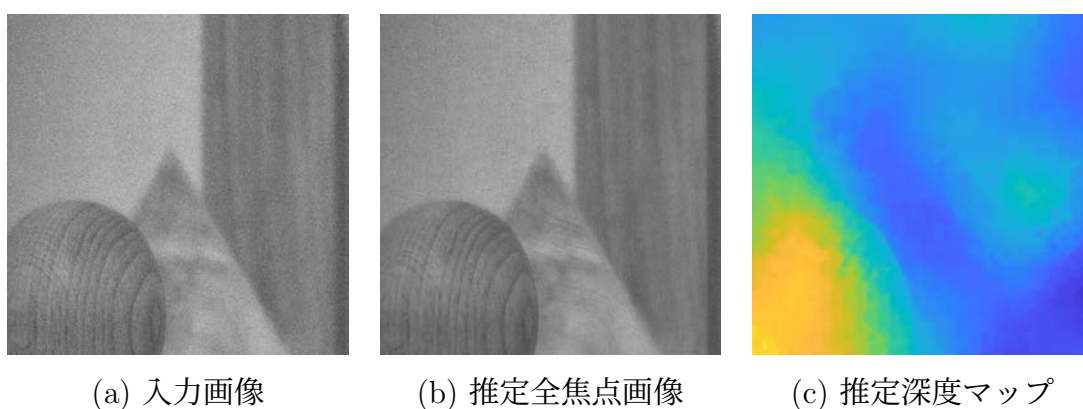
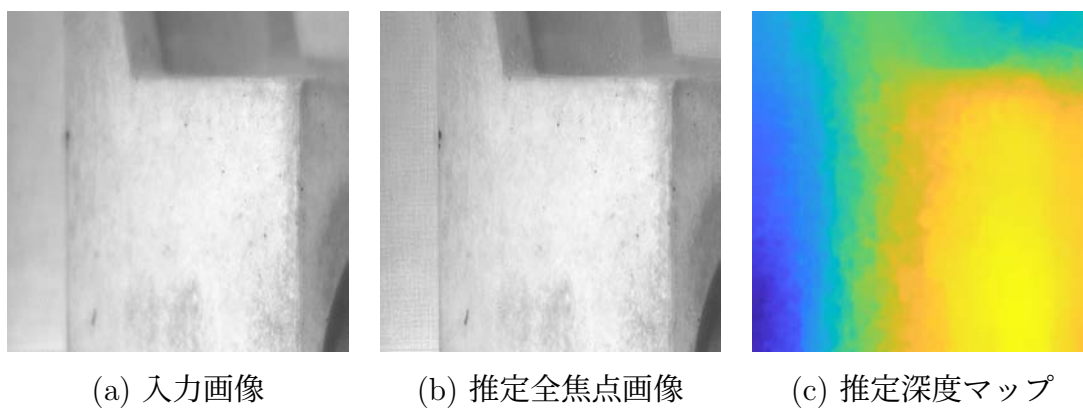
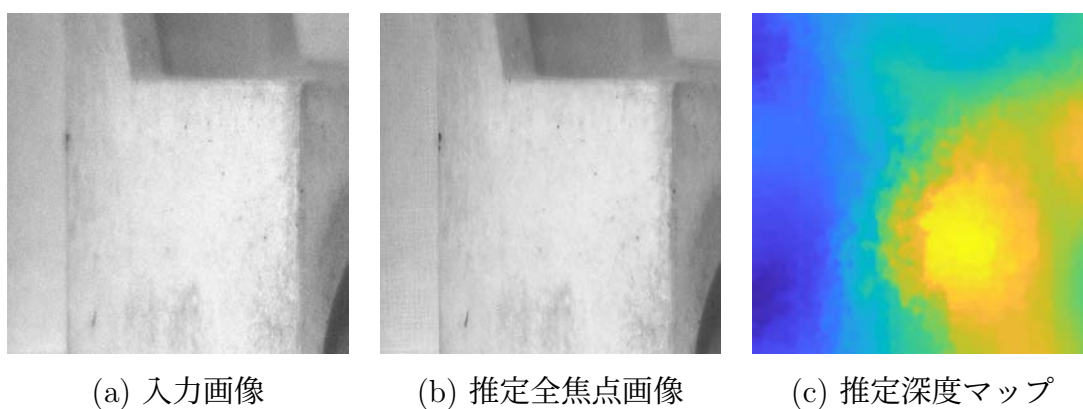
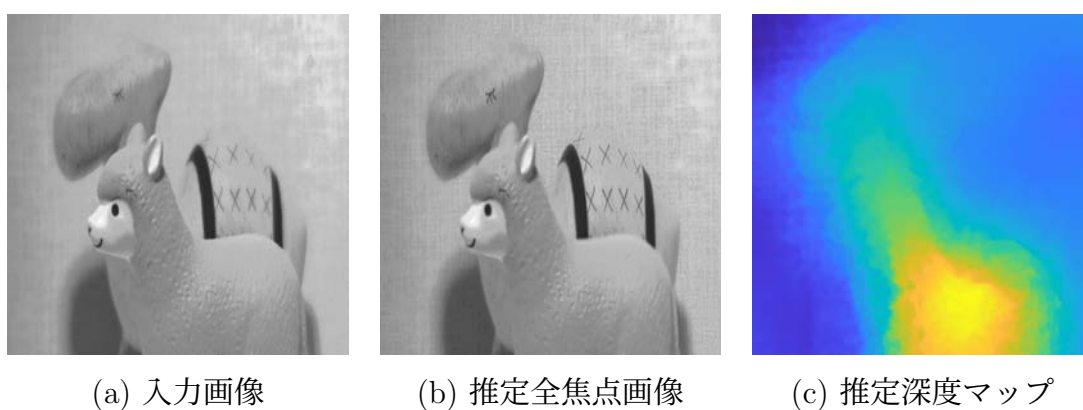


図 25: Wood toy (積み木) の例 ($\sigma_n = 0.01$) .

図 26: Concrete chair (椅子) の例 ($\sigma_n = 0.001$) .図 27: Concrete chair (椅子) の例 ($\sigma_n = 0.01$) .図 28: Llama (ラマ) の例 ($\sigma_n = 0.001$) .

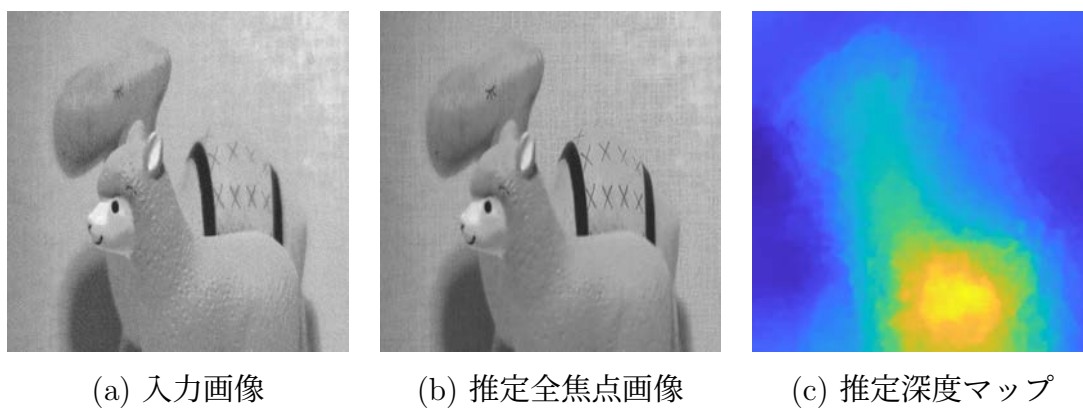


図 29: Llama (ラマ) の例 ($\sigma_n = 0.01$) .

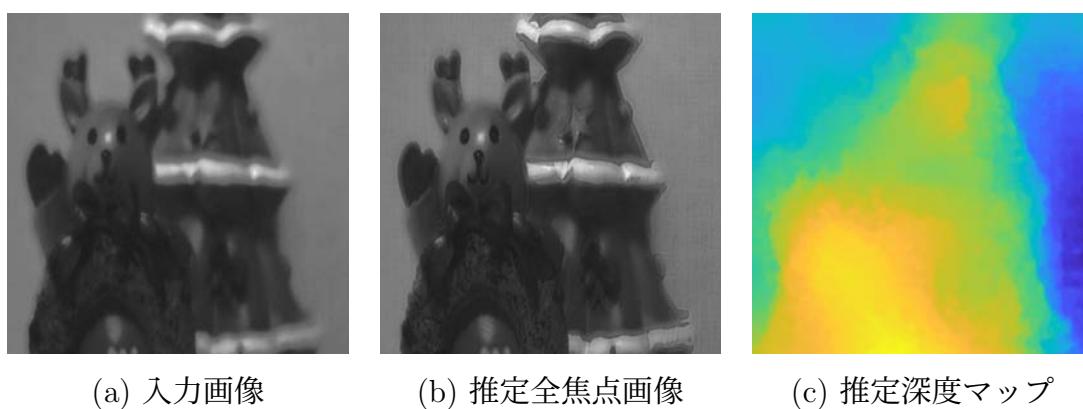


図 30: Christmas (クリスマス) の例 ($\sigma_n = 0.001$) .

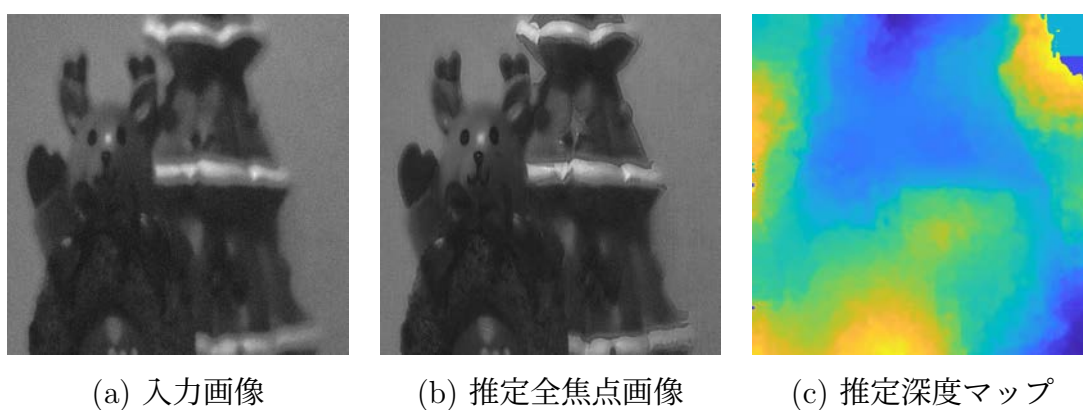


図 31: Christmas (クリスマス) の例 ($\sigma_n = 0.01$) .

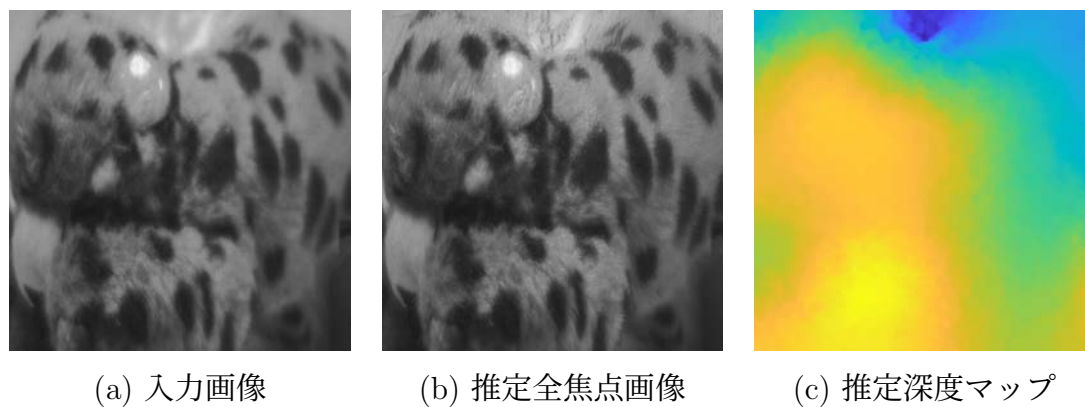


図 32: Tiger (トラ) の例 ($\sigma_n = 0.001$) .

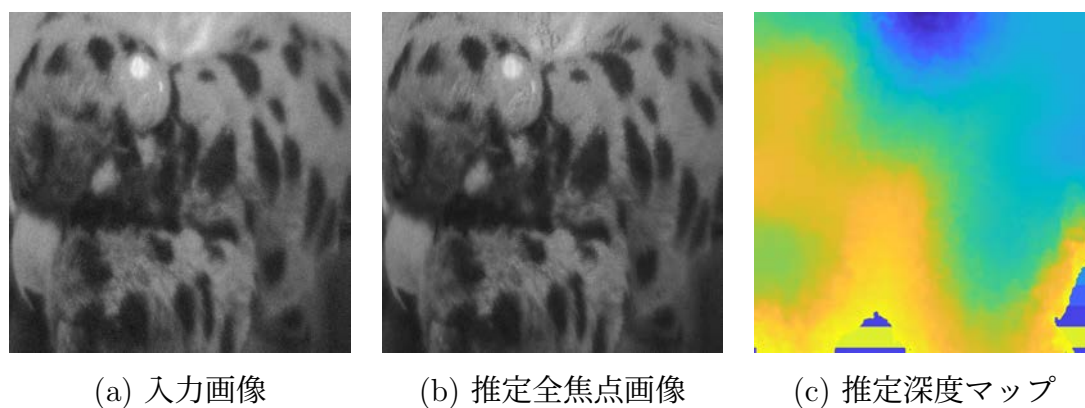


図 33: Tiger (トラ) の例 ($\sigma_n = 0.01$) .

4.3 小括

4.3.1 本提案のまとめ

本章では、色収差と呼ばれる標準的な広角レンズの屈折特性を利用して、一枚のマルチスペクトル画像からシーンの奥行きを推定するアプローチを紹介した。レンズを通過した光は、波長に対して角度が変化し、波長ごとに焦点距離が変化し、焦点ボケが発生することを利用している。多くの DfD 法とは異なり、我々の手法は複数のスペクトル成分を用いて複数の焦点距離を計算して深度推定を行うため、焦点ボケが焦点面の前後どちらの点から生じているかという曖昧さの問題を含まない。実験を通して、マルチスペクトル画像における色収差の影響が、深度を回復するために十分な情報を提供することを示した。様々な感度曲線を持つセンサの距離推定精度を比較し、波長帯域が狭く、チャンネル数が多い方が有利であることを示した。さらに、マルチスペクトルの全焦点画像を効率的に復元する手法を提案した。提案手法の有効性を実世界の様々なシーンを対象とした実験で検証し、良好な全焦点画像が得られ、部分的な鏡面反射が見られるシーンの奥行きについても安定に推定できることを確認した。

4.3.2 単色収差の影響

本提案ではレンズにおける波長依存の性質である色収差に着目して、レンズの色収差を利用した深度推定手法を提案した。レンズにおける代表的な収差として色収差と単色収差が存在するが、これまでの議論では単色収差に関する影響について考慮していない。本節では、本提案における単色収差の影響と、その対応策について議論する。

単色収差はザイデル五収差とも呼ばれ、以下に示す五種類の収差が存在する。

- 球面収差

球面収差とは図 34 (b) のように、点光源からの光線がレンズの焦点に収束せず、光線がまばらになってしまう現象を指す。この現象が生じている場合、焦点が合わないために画像の解像度が低下する。

- コマ収差

コマ収差とは図 34 (c) のように、光軸と平行でない光がセンサ面の一点に収束しない現象を指す。その性質から望遠よりも広角な焦点距離を持つレンズにて生じやすく、また画像の周辺部において生じやすい。

- 非点収差

非点収差とは図 34 (d) のように、光軸と平行でない光がレンズに入射する際に、レンズの水平方向（赤線）と鉛直方向（青線）で見かけの直径が異なるために実効的な焦点距離が変化し、光が一点に収束しない現象を指す。コマ収差と同様にその性質から望遠よりも広角な焦点距離を持つレンズにて生じやすく、また画像の周辺部において生じやすい。

- 像面収差

像面収差とは図 34 (e) のように、レンズが湾曲しているためにレンズの結像面が平面ではなく球面状になってしまう現象を指す。光軸と垂直な平面を撮影したときに画像の中央部でピントが合っている場合、画像の周辺部ではピントが合わずボケが生じる。

- 歪曲収差

歪曲収差とは図 34 (f) のように、レンズが湾曲しているためにセンサ面に結像する像に歪みが生じ、方眼を撮影した場合に正しい方眼の画像が得られない現象を指す。一般に、広角レンズでは左のような樽型収差を生じ、望遠レンズでは糸巻型収差を生じる。

ここで、これらの単色収差と提案手法との関係を議論する。球面収差・コマ収差・非点収差はモデルに含まれない焦点ボケを生じるために画像の解像度を低下させることに繋がる。これにより観測できる焦点ボケの変化度合が少なくなり、深度推定の分解能が低下する可能性がある。コマ収差および非点収差についてはレンズの焦点距離が広角であるほど生じやすい収差であるため、標準的または望遠に近い焦点距離のレンズを選択することで影響を抑えることができる。像面収差は画像の中央部と周辺部で深度に対する焦点ボケが変化するため、平面を撮影した場合でも中央部と周辺部で推定深度が変化することが予測される。これは、提案手法におけるレンズパラメータ α, β を画像の空間依存のパラメータとして扱い、キャリブレーション段階でパラメータの空間分布を推定するようにアルゴリズムを改善することで対応可能であると考えられる。歪曲収差はカメラシステムが固定されている場合、キャリブレーションパターンを用いた事前補正が可能であるため、事前補正を行うことにより改善されることが予想される。

事前補正が難しい単色収差についても絞りを使用するなどして改善することが可能であるため、Trouve ら [55] の使用したレンズのように色収差の影響を強める光学系を使用することにより、単色収差の影響を弱めつつ色収差による波長依存のボケを観測する方法も考えられる。

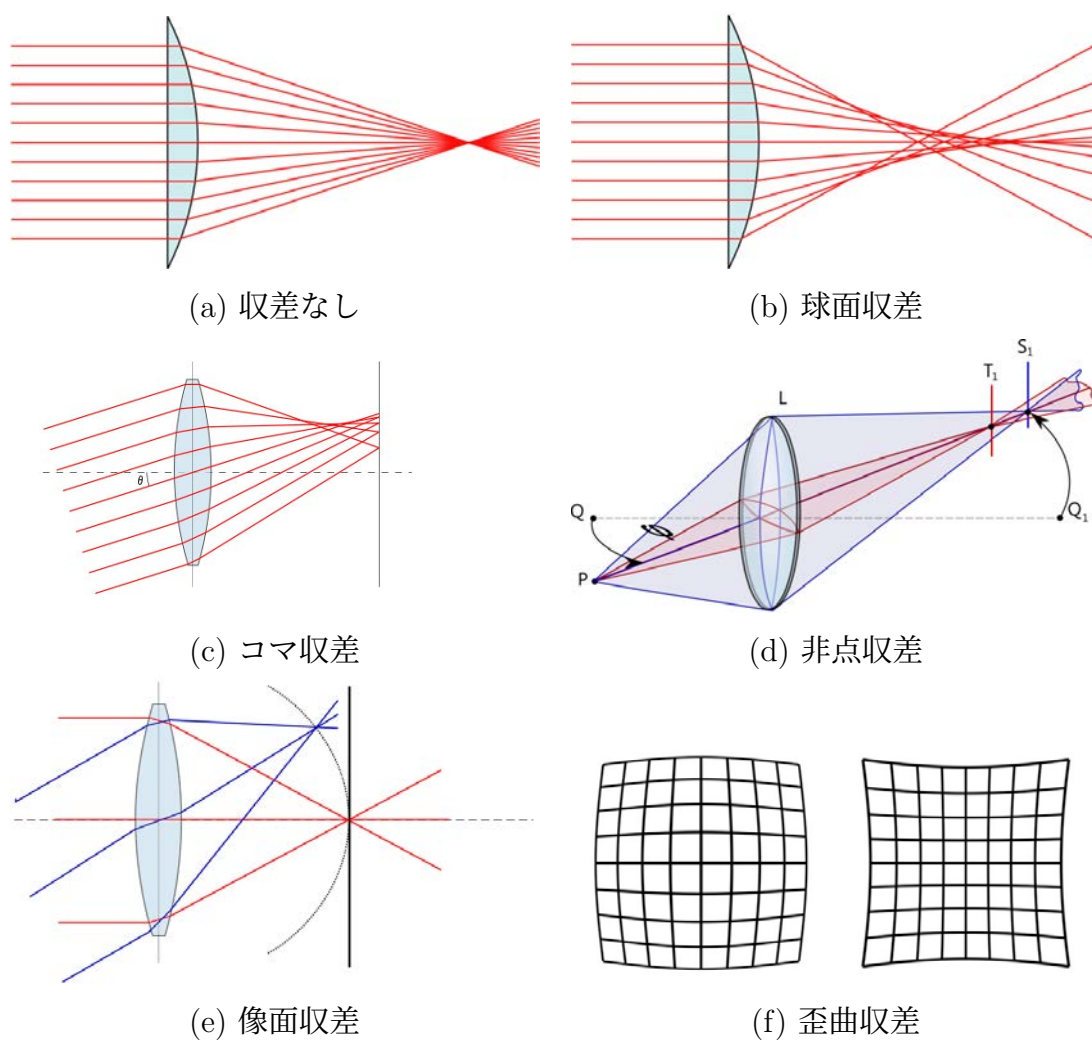


図 34: 単色収差（ザイデル五収差）の様子. (出典: <<https://ameblo.jp/stereo-gn-ryoma/entry-12559502437.html>>)

第5章 水面の屈折解析による分光画像からの水中シーン再構成

本章では、本論文にて提案する「水面の屈折における波長依存性の解析による分光画像からの水中シーン再構成」について、分光を用いた水面形状および水中シーンの再構成モデルおよびアルゴリズムの詳細を述べる。実験を通して、本手法による水中シーン再構成の有効性を検証する。

5.1 多波長の光線による水中シーンのモデル化

本節では、分光モデル用いた水面における屈折光解析による、水面形状および水中シーンの再構成モデルおよびアルゴリズムについて詳細を述べる。

5.1.1 概要

初めに、本提案にて考慮する物理現象およびアルゴリズムについて、概要を説明する。水面に入射した光は屈折し、その方向はスネルの法則により水の屈折率と水面と光との角度（入射角）によって決定する。屈折率は波長に依存しているため、正投影の分光カメラシステムにおいて異なる波長で観測された画像には、図 35 に示すように屈折の強さが異なることに起因するズレが生じる。対象物が十分なテクスチャを持っていると仮定して、異なる波長で撮影された画像の特徴点の視差を計算することで、対応する特徴点の視差を得ることができる。従来のステレオ方式と同様に、異なる波長で撮影された画像間に生じる視差は、水面から水中物体までの距離が長くなるにつれて大きくなる。本研究では、異なる波長で撮影された画像の視差から、水面の形状が既知でない場合にも適用可能な水中シーン再構成手法を提案する。

図 36 に提案手法の概要図を示す。提案手法において、水面が未知の水中シーンを再構成するための過程：(a) 水中を撮影した正投影分光カメラの画像を入力とする。水中物体を撮影した多波長の画像群は、水面の法線方向に依存した相対的な視差の方向と、観測波長と入射角および水中物体の深度に依存する視差の大きさに関する情報

を持つ. (b) 特徴点マッチングまたは画像マッチング法により, 異なる波長の画像間に共通の対応点を求め, 各対応点における視差方向と視差の大きさをの情報を取得する. (c) 3つの波長から得られる視差情報の関係を利用して, 水面法線および水中シーンと密接に関係するパラメータを推定し, 深度係数と角度係数を推定する. 水面法線および水中シーンのパラメータ化について, 本節にて詳細を述べる.

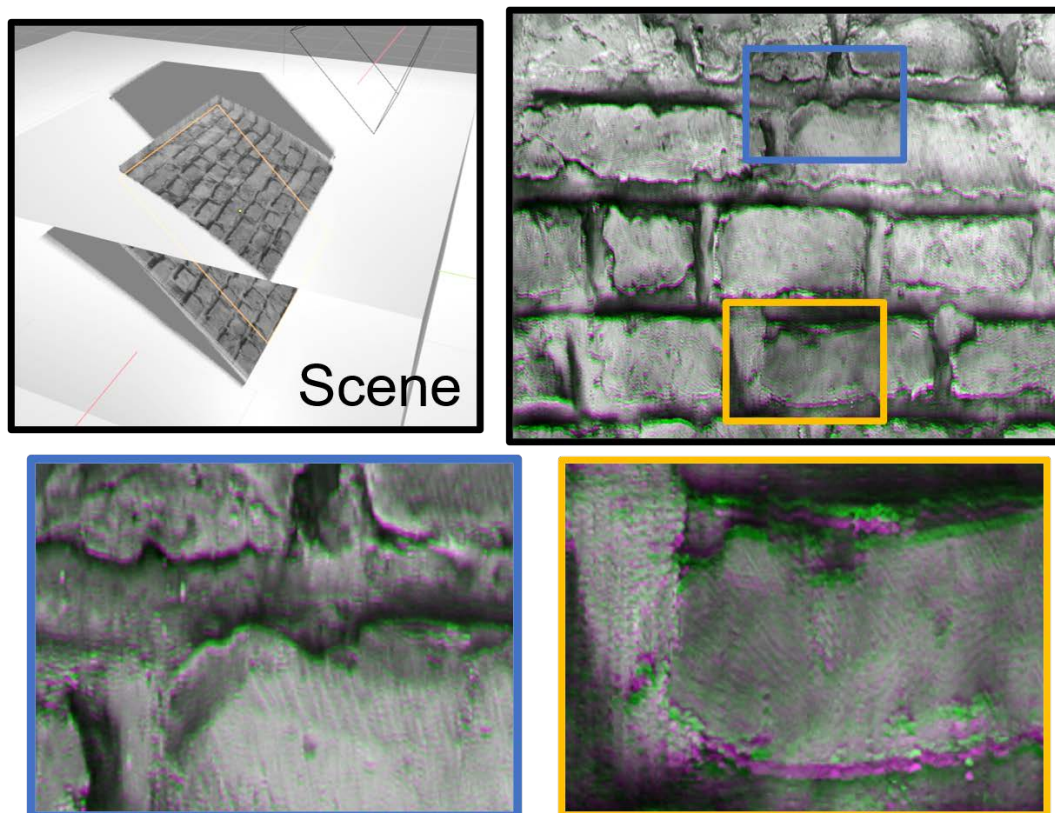


図 35: 水面での屈折によるシーンの深度依存の分光画像のズレ.

以下, 太字の文字は三次元ベクトルを表し, 演算子 $(\cdot)$ は二つのベクトルの内積を表す. また, 本提案にて扱う三次元座標系は, 以下のように定義する. z 軸は水面が水平であるときの法線方向に対応する. x 軸と y 軸は, カメラの光軸を z 軸に合わせたときのカメラの画像座標とする. カメラの姿勢ベクトル $\mathbf{c}$ と, 波長 λ における媒質の屈折率 $k(\lambda)$ は既知であると仮定する.

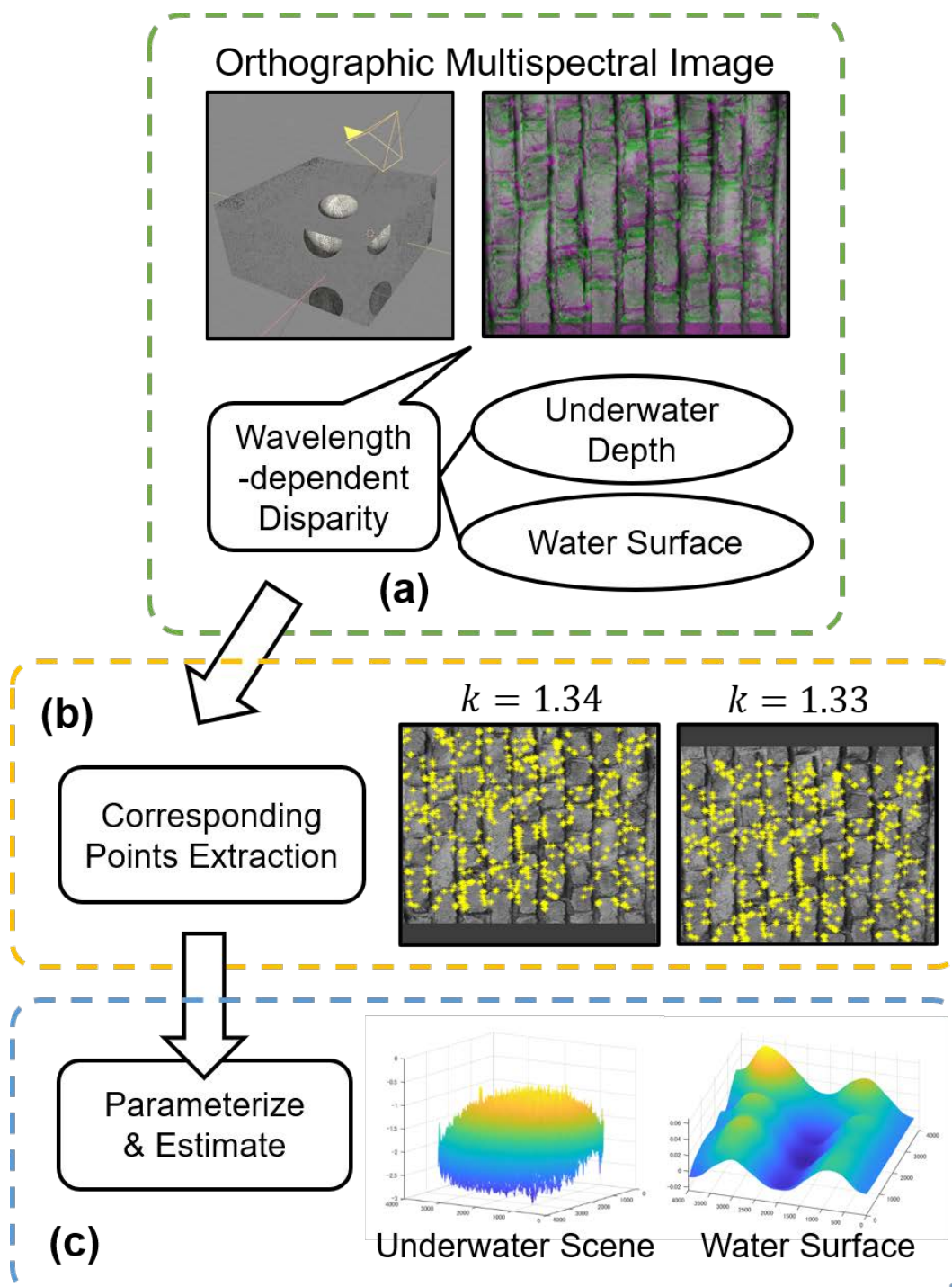


図 36: 提案手法の概要.

5.1.2 水面における屈折と水面法線・水中シーンとの関係

初めに、水面における屈折と、水面法線および水中シーンとの関係をモデル化する。図 37 にカメラに並行に入射する多波長の光線と、水面における屈折を介する水面法線および水中シーンとの関係を表す平面図を示す。ここで、水面の法線を n とし、観測波長 $\lambda_1, \lambda_2, \lambda_3$ におけるそれぞれの分光画像中の対応点を $p_{1,2,3}$ とする。すなわち、 $p_{1,2,3}$ は水中物体の表面上における同一点を捉えている。さらに、 $l_{1,2,3}$ はこれらの対応点に対応する水中光線の三次元直線、 $v_{1,2,3}$ はこれらの水中光線の正規化された方向ベクトル、 $q_{1,2,3}$ はこれらの水中光線と水面との交点、 w は対象とする水中物体の表面点の三次元位置ベクトルをそれぞれ表す。

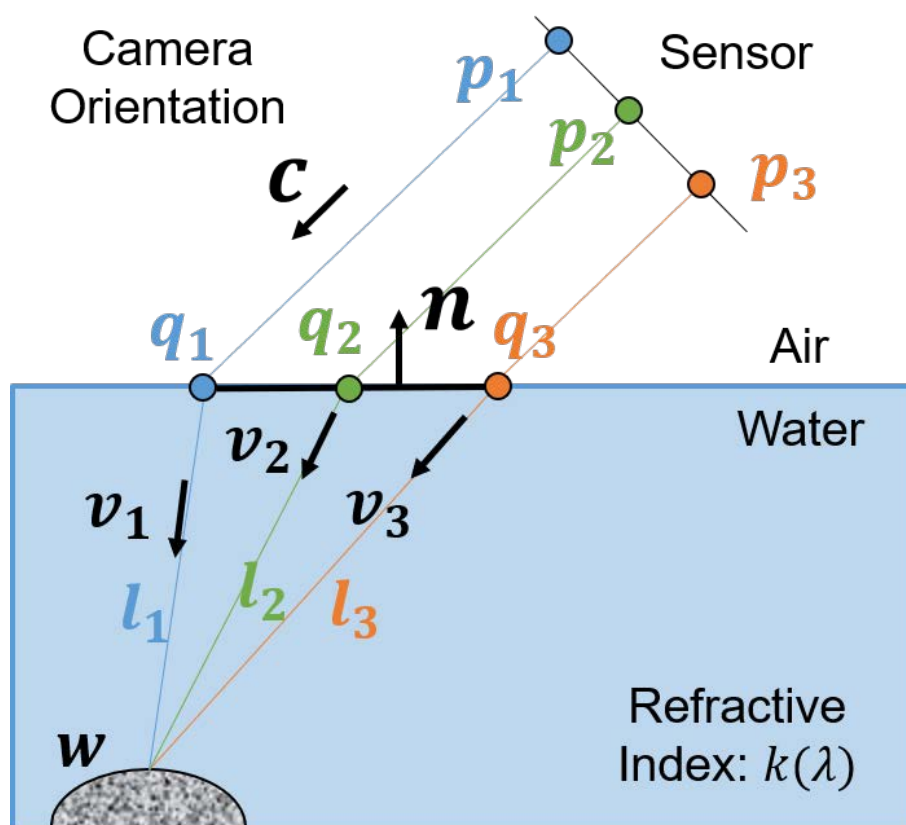


図 37: 水面における屈折の波長依存性と、水面法線および水中の三次元点との関係。

本節では、システムにより観測される波長画像間の視差を利用して、水面法線および水中の三次元点を推定する手法について述べる。第一に、水面の法線が既知である場合について、水中の三次元点を得るためのモデルについて説明する。第二に、水面の法線が未知である場合について、水面の法線を推定するためのモデルについて説明

し、最適化およびルックアップテーブルを用いた推定アルゴリズムについて議論する。

5.1.3 既知の水面における水中シーンの再構成

ここでは、水面の法線が既知である場合の水中物体の一点 $\mathbf{w}$ の深さを取得するモデルについて考える。水中を通過する二つの波長 λ_1, λ_2 に対応する光線 $\mathbf{l}_{1,2}$ は、直線の式から以下のように与えられる：

$$\begin{aligned} \mathbf{l}_1(s) &= \mathbf{q}_1 + s\mathbf{v}_1, \\ \mathbf{l}_2(t) &= \mathbf{q}_2 + t\mathbf{v}_2. \end{aligned} \quad (19)$$

ここで、 s, t は直線のパラメータである。波長 λ_1 中の画像上の点 $\mathbf{p}_1$ と波長 λ_2 中の画像上の点 $\mathbf{p}_2$ を表す特徴点ペアがシーン内の同じ三次元点に対応する。一般に三次元直線は必ずしも交差しないが、これらの三次元直線の交点は同一点 $\mathbf{w}$ を表しており、 $\mathbf{l}_1(s)$ と $\mathbf{l}_2(t)$ が交差することが保証されている。水面法線とカメラの姿勢ベクトルから、光が水中を通り、水面で屈折して $\mathbf{p}_{1,2}$ に向かう水面上の点 $\mathbf{q}_{1,2}$ を計算することができる。さらに、相対位置ベクトル $\mathbf{r}_{12} = \mathbf{q}_1 - \mathbf{q}_2$ を定義することにより、以下のよう $\mathbf{l}_1(s) = \mathbf{l}_2(t)$ であることを用いて s と t の交点を推定することができる：

$$\begin{aligned} \mathbf{l}_1(s_{12}) &= \mathbf{l}_2(t_{12}) = \mathbf{w} \text{ のとき,} \\ s_{12} &= \frac{\mathbf{r}_{12} \cdot (\mathbf{v}_1 - \alpha_{12}\mathbf{v}_2)}{\alpha_{12}^2 - 1}, \\ t_{12} &= \frac{\mathbf{r}_{12} \cdot (\alpha_{12}\mathbf{v}_1 - \mathbf{v}_2)}{\alpha_{12}^2 - 1}, \\ \text{ただし } \mathbf{r}_{12} &= \mathbf{q}_1 - \mathbf{q}_2, \alpha_{12} = \mathbf{v}_1 \cdot \mathbf{v}_2. \end{aligned} \quad (20)$$

光線の方法ベクトル $\mathbf{v}_{1,2}$ は、カメラの姿勢ベクトル $\mathbf{c}$ 、水面法線 $\mathbf{n}$ および屈折率の逆数 $e_i = 1/k(\lambda_i)$ から、ベクトル形式のスネルの法則に基づいて計算することができる。

$$\begin{aligned} \mathbf{v}_i &= e_i \mathbf{c} + \left(e_i \gamma - \sqrt{1 - e_i^2(1 - \gamma^2)} \right) \mathbf{n}, \\ i &\in \{1, 2\}, \text{ where } \gamma = -\mathbf{n} \cdot \mathbf{c}. \end{aligned} \quad (21)$$

式 21 を $\alpha_{12} = \mathbf{v}_1 \cdot \mathbf{v}_2$ に代入することにより、パラメータ α_{12} は、 γ と $e_{1,2}$ から計算することができる。

$$\begin{aligned} \alpha_{12} &= \sqrt{1 - (e_1^2 + e_2^2)(1 - \gamma^2) + e_1^2 e_2^2 (1 - \gamma^2)^2} \\ &+ e_1 e_2 (1 - \gamma^2). \end{aligned} \quad (22)$$

続いて、水面からの水中物体の相対位置ベクトル $\hat{\mathbf{w}}$ を考える． $\hat{\mathbf{w}}$ は対象物表面上の点であり、 r_{12} , s_{12} , $\mathbf{v}_1$ を式 19 に代入することで与えられる：

$$\begin{aligned}\hat{\mathbf{w}} &= \mathbf{w} - \mathbf{q}_1 \\ &= s_{12}\mathbf{v}_1 \\ &= \frac{(\alpha_{12}e_1 - e_2)}{\alpha_{12}^2 - 1}\beta_{12}\{e_2\mathbf{c} + (e_2\gamma - \sqrt{1 - e_2^2(1 - \gamma^2)})\mathbf{n}\},\end{aligned}\quad (23)$$

ただし $\beta_{12} = \mathbf{r}_{12} \cdot \mathbf{c}$.

式 23 について、水面法線 $\mathbf{n}$ が既知である場合、 β_{12} は $\mathbf{r}_{12}$ とカメラの姿勢ベクトル $\mathbf{c}$ から決定される．視差ベクトル $\mathbf{d}_{12} = \mathbf{p}_1 - \mathbf{p}_2$ を考えると、二直線の水面上における相対位置ベクトル $\mathbf{r}_{12}$ の大きさは、カメラセンサ上の $\mathbf{d}_{12}$ の大きさと入射角を用いて以下のように表現される：

$$|\mathbf{r}_{12}| = \frac{|\mathbf{d}_{12}|}{\cos \theta_i} = \frac{|\mathbf{d}_{12}|}{\gamma}. \quad (24)$$

図 38 に示すように、三つのベクトル $\mathbf{n}$, $\mathbf{r}_{12}$, $\mathbf{d}_{12}$ は同じ屈折面内にあり、 $\mathbf{n}$ と $\mathbf{r}_{12}$ は直交しているため、 β_{12} は次のように計算できる．

$$\left(\frac{\mathbf{r}_{12}}{|\mathbf{r}_{12}|} \cdot \mathbf{c}\right)^2 + (\mathbf{n} \cdot \mathbf{c})^2 = \frac{\beta_{12}^2}{|\mathbf{r}_{12}|^2} + \gamma^2 = 1, \quad (25)$$

$$\beta_{12} = |\mathbf{r}_{12}|\sqrt{1 - \gamma^2} = \frac{|\mathbf{d}_{12}|}{\gamma}\sqrt{1 - \gamma^2}.$$

ここまでの議論から、水面の法線 $\mathbf{n}$ がわかっている場合には、二種類の波長情報を用いて水中対象上の点 $\hat{\mathbf{w}}$ を求められる．

5.1.4 未知の水面における水中シーンの再構成

前節では、水面の法線 $\mathbf{n}$ が与えられている場合に、水中対象上の点 $\hat{\mathbf{w}}$ を求める手法を説明した．本節では、表面法線が未知であり非平坦な水面の場合にも水中シーンの再構成を適用できるように、第三の波長情報を用いて水面の法線ベクトルを推定する手法を導入する．ここで、式 19 と 20 から、次の式が成り立つ：

$$\mathbf{l}_1(s_{12}) - \mathbf{l}_2(t_{12}) = \mathbf{r}_{12} + s_{12}\mathbf{v}_1 - t_{12}\mathbf{v}_2 = \mathbf{0}. \quad (26)$$

式 26 は、三ベクトル $\mathbf{r}_{12}$, $\mathbf{c}$, および $\mathbf{n}$ が互いに共有する二次元平面が存在することを示している．ここで、 $\mathbf{r}_{12}$ と $\mathbf{d}_{12}$ が同一平面に位置することは自明であることを用

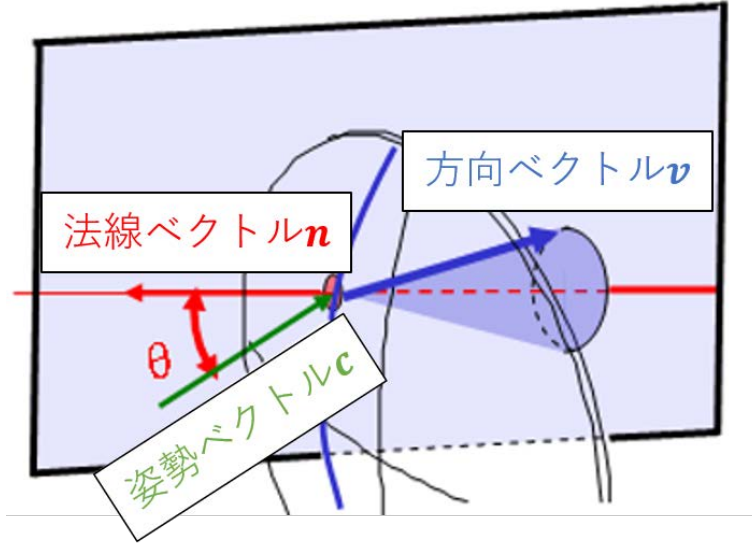


図 38: 入射光, 屈折光および屈折面の法線ベクトルのなす共平面.

いると, 水面の法線ベクトル $\mathbf{n}$ は, 互いに直交する二つのベクトル $\mathbf{d}_{12}$, $\mathbf{c}$ と同一平面上にあり, これらの直交ベクトルと自身との内積を用いて表すことができる. したがって, カメラの姿勢ベクトル $\mathbf{c}$ と正規化視差ベクトル $\mathbf{d}_n$ を用いて, 水面の法線ベクトル $\mathbf{n}$ を式 27 として導出することができる.

$$\begin{aligned}\mathbf{n} &= (\mathbf{n} \cdot \mathbf{c})\mathbf{c} + (\mathbf{n} \cdot \mathbf{d}_n)\mathbf{d}_n, \\ &= -\gamma\mathbf{c} + \sqrt{1 - \gamma^2}\mathbf{d}_n, \\ &\quad \text{ただし} \\ \gamma &= -\mathbf{n} \cdot \mathbf{c}, \mathbf{d}_n = \frac{\mathbf{d}_{12}}{|\mathbf{d}_{12}|}.\end{aligned}\tag{27}$$

この式から, 二つの異なる波長で観測された対応点間の視差 $\mathbf{d}_{12}$ とカメラの姿勢ベクトル $\mathbf{c}$ およびパラメータ γ を用いて, 水面の法線ベクトル $\mathbf{n}$ を表すことができる. ただし, パラメータ γ は法線ベクトル $\mathbf{n}$ に依存するパラメータであり, 水面の法線が未知である場合には直接計算できない. 前節の議論では, カメラの姿勢ベクトル $\mathbf{c}$ が既知の場合には, カメラの姿勢ベクトルおよび二つの波長画像上の視差から水深が推定できることを示してきた. 以下では, 第三の波長 λ_3 に対応する追加の光線を導入することで, γ を推定するためのモデルを提案する.

図 37 に示すような第三波長の水中光線に対応する第三の直線 $\mathbf{l}_3$ を以下のように導入する:

$$\mathbf{l}_3(u) = \mathbf{q}_3 + u\mathbf{v}_3,\tag{28}$$

前の二波長の光線と同様に, $\mathbf{q}_3$ は $\mathbf{l}_3$ と水面との交点であり, $\mathbf{v}_3$ は $\mathbf{l}_3$ の正規化方向ベ

クトルである．二直線の交点の線形パラメータは， l_3 と l_2 にもとづく連立方程式から計算される：

$$\begin{aligned} l_2(t_{23}) &= l_3(u_{23}) = \mathbf{w} \text{ のとき,} \\ t_{23} &= \frac{\mathbf{r}_{23} \cdot (\mathbf{v}_2 - \alpha_{23}\mathbf{v}_3)}{\alpha_{23}^2 - 1}, \\ u_{23} &= \frac{\mathbf{r}_{23} \cdot (\alpha_{23}\mathbf{v}_2 - \mathbf{v}_3)}{\alpha_{23}^2 - 1}, \end{aligned} \quad (29)$$

ただし $\mathbf{r}_{23} = \mathbf{q}_2 - \mathbf{q}_3$, $\alpha_{23} = \mathbf{v}_2 \cdot \mathbf{v}_3$.

ここでは，三直線の水面上の点間の距離 $|\mathbf{r}_{12}|$ と $|\mathbf{r}_{23}|$ が十分に小さいとき，水面の法線が一様であるという仮定を導入していることに注意する．

式 22 および式 25 と同様に，第二，第三の直線に係るパラメータ α_{23} と β_{23} は，次のように計算することができる．

$$\begin{aligned} \alpha_{23} &= \sqrt{1 - (e_2^2 + e_3^2)(1 - \gamma^2) + e_2^2 e_3^2 (1 - \gamma^2)^2} \\ &\quad + e_2 e_3 (1 - \gamma^2), \end{aligned} \quad (30)$$

$$\beta_{23} = \frac{|\mathbf{d}_{23}|}{\gamma} \sqrt{1 - \gamma^2}.$$

水中の三直線は同一点 $\mathbf{w}$ で交差するため，第二の直線に係る交点のパラメータ t_{12} と t_{23} は同一の値をとることは明らかである．これから， t_{12} と t_{23} の商は 1 になる：

$$\begin{aligned} \frac{t_{12}}{t_{23}} &= \frac{\mathbf{r}_{12} \cdot (\alpha_{12}\mathbf{v}_1 - \mathbf{v}_2) (\alpha_{23}^2 - 1)}{\mathbf{r}_{23} \cdot (\mathbf{v}_2 - \alpha_{23}\mathbf{v}_3) (\alpha_{12}^2 - 1)} \\ &= \frac{(\alpha_{12}e_1 - e_2) (\alpha_{23}^2 - 1) \beta_{12}}{(e_2 - \alpha_{23}e_3) (\alpha_{12}^2 - 1) \beta_{23}} = 1. \end{aligned} \quad (31)$$

β_{12} と β_{23} を式 31 に代入し展開することにより，陰関数 $f(\gamma)$ を得ることができる：

$$f(\gamma) = \frac{|\mathbf{d}_{12}|}{|\mathbf{d}_{23}|} - \frac{(e_2 - e_3\alpha_{23}(\gamma)) (\alpha_{12}^2(\gamma) - 1)}{(e_1\alpha_{12}(\gamma) - e_2) (\alpha_{23}^2(\gamma) - 1)} = 0. \quad (32)$$

γ の値は二つの単位ベクトル $\mathbf{n}$ (水面の法線ベクトル) および $\mathbf{c}$ (カメラの姿勢ベクトル) の内積であり，一般に正の値をとることができるため， $f(\gamma)$ を目的関数として適当な範囲で最小化する最適化手法を用いて， γ の値を推定することができる：

$$\arg \min_{\gamma} \{|f(\gamma)|, \gamma \in [0, 1]\}. \quad (33)$$

ここまでの議論により，未知の水面法線ベクトル $\mathbf{n}$ は， γ を推定することにより求められることが示された．さらに，水中対象物上の点を決定するための二つのパラ

メータ $(\alpha_{12}, \beta_{12})$ も γ を用いて、それぞれ式 22 と式 25 のように決定される。したがって対象物の水深は、前節と同様に $\hat{w}$ のように与えられる。 γ の値は、 $[0, 1]$ の範囲内にあることが一般的であるため、本提案ではこの範囲内の式 32 の関数の最小値を求める形の最適化によるパラメータ推定を導入する。

5.2 改良アルゴリズムの導入

5.2.1 ルックアップテーブルによるパラメータ推定

提案システムにおいて、図 39 に示すような試作機によって撮影された画像は、提案モデルに含まれていないレンズの歪みなどの要因によって、最適化の推定精度に影響を及ぼす可能性があることが分かっている。

推定の複雑さを軽減してこの問題に対処するため、ルックアップテーブル (LUT) を用いてパラメータ γ を推定する手法を提案する。式 32 より、パラメータ γ の値は基本的に視差の比率 d_{12}/d_{23} に依存して変化することがわかる。図 40 はカメラからの水面の入射角と、ある波長ペアにおける視差の比率 d_{12}/d_{23} の関係を示している。図 40 から、入射角が大きくなるにつれて、二つの波長画像に含まれる視差の比率がほぼ線形に推移することを示している。また、このプロットにおける観測では、入射角が大きくなるにつれて分散の範囲は小さくなることがわかる。

この方法では、テーブルを生成するための学習データとして、入射角が既知である状態で水中物体を撮影した分光画像を複数用意する。各画像について、特徴点マッチングを用いて特定の波長ペアの視差の比率を計算し、その平均値を使用して入射角と視差の比率の関係を図 40 のようにマッピングする。実際の入力画像については、特徴点ごとにこのテーブルを参照して、水面とカメラの間の角度を推定する。ここで、テーブルにより得られる入射角 θ_i の余弦が以下の関係になることを利用して、パラメータ γ を求めることができる：

$$\gamma = -\mathbf{n} \cdot \mathbf{c} = \cos \theta_i \quad (34)$$

前節で述べたように、二つのパラメータ $(\alpha_{12}, \beta_{12})$ と水面の法線ベクトル $\mathbf{n}$ は、 γ の関数で記述されている。従って、テーブルを用いてパラメータ γ を求めることでこれらの値を得ることができる。水中対象物上の点 $\hat{w}$ についても同様である。

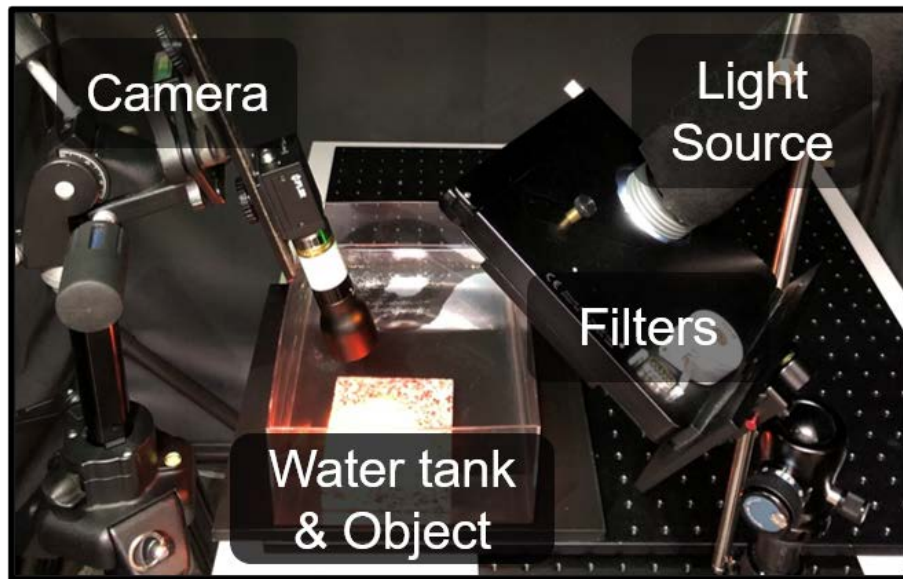


図 39: 提案システムの試作機.

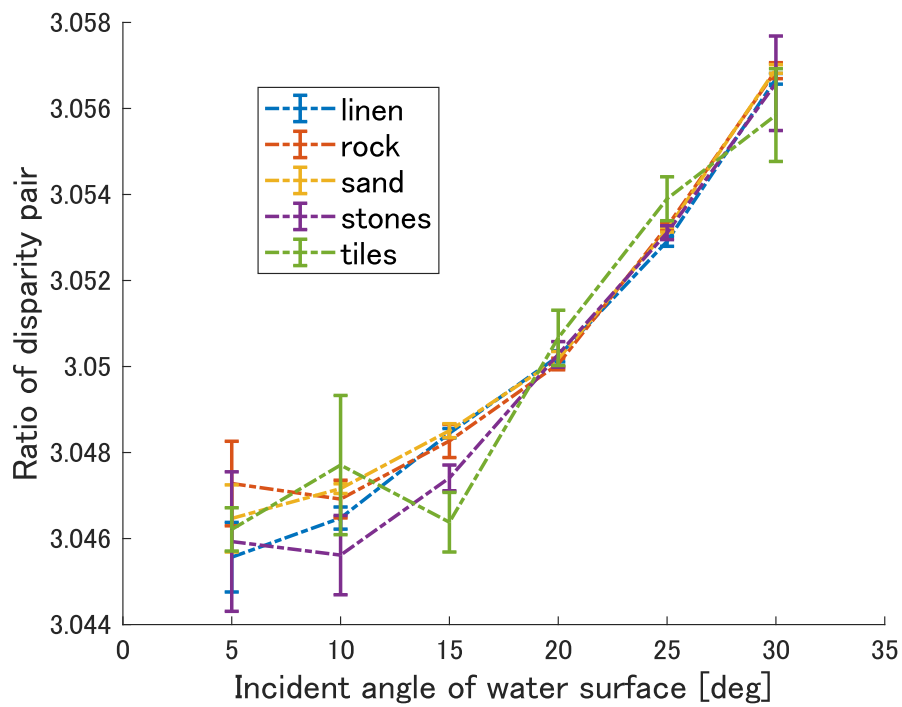


図 40: カメラからの水面の入射角と, ある波長ペアにおける視差の比率の関係.

5.2.2 多波長画像を用いたマッチング精度改善

提案手法では三つ以上の波長で観測された特徴点の集合を水面および水中シーンの再構成に使用している．本提案を含む特徴点マッチングを利用した手法では，そのマッチング精度が手法の精度と密接に関係することは明らかである．本節では，最大で五つの波長画像を用いて，特徴点マッチングに用いる波長画像数とマッチング精度との関係を調査する．

表 3 に，平坦な水面下にある平面の対象を撮影した場合の，特徴点マッチングによってされる視差の平均値と標準偏差を示す．水面および水中の対象が両方とも平坦であるため，視差の大きさは対象の水深に依存し，視差の標準偏差は理想的には 0 となる．しかしながら，画像に含まれる歪みやノイズ等の影響により，視差の大きさにばらつきが生じる．三波長の画像間でマッチングをとった場合と五波長の画像間でマッチングをとった場合を比較すると，五波長を使用した場合に標準偏差は平均で約 10% の改善，最大で約 20% の改善が見られることが分かった．

Incident angle [deg]	Used wavelength images		
	3	4	5
5	2.22 ± 0.65	2.10 ± 0.67	2.11 ± 0.66
10	3.33 ± 0.54	3.33 ± 0.52	3.33 ± 0.51
15	4.29 ± 0.75	4.29 ± 0.72	4.28 ± 0.67
20	6.31 ± 0.73	6.31 ± 0.70	6.31 ± 0.68
25	7.49 ± 0.90	7.48 ± 0.82	7.47 ± 0.74
30	9.49 ± 1.08	9.47 ± 1.02	9.49 ± 0.99

表 3: 試作した提案システム（図 39）において抽出された視差の平均値と標準偏差．

5.2.3 変形を伴う画像の密なマッチング

テクスチャがあまり鮮明でないシーンについては，密な特徴点を得ることは一般に難しく，特徴点にもとづいてシーン全域を再構成することは容易ではない．本節ではこのようなシーンに対応するために，密な対応関係を得るための画像マッチング技術を導入し，特徴の少ない領域における精度を向上させることを試みる．本システムで観測される波長画像ペアには，対象物の深さ，水面の法線，および観測波長に依存する屈折率と密接に関係した画像の変形が含まれている．そこで，本提案では，拡散モデル [53] に基づく波長画像間のマッチングアルゴリズムを導入して，シーン全域の密

な対応関係を得ることを提案する．拡散モデルに基づくアプローチは，動きのある被写体を対象とした非剛体マッチングにも適用可能であり，我々の場合は水の屈折による変形を扱うことに適している．

図 41 は，テクスチャのある球体と平坦な水面からなるシミュレートされたシーンについて，拡散モデルによる画像マッチングによって得られた二つの波長画像間の対応点群の視差を示している．図 41 (b) では，拡散画像マッチングにより，与えられた二種類の波長画像間から密な視差マップが得られることがわかる．図 41 (c) は，(b) のマップにおける橙色の線に沿う輝度値のプロットを図 41 の中で示している．プロットの縦軸は対応点から得られる視差の大きさ，横軸は橙の線に沿うピクセルの座標を表している．図 41 に示すプロット (c) から，拡散画像マッチングによって得られる視差は大域的には正しいように見えるが，プロットがなだらかでないことから，不正確な視差を含んでいることがわかる．

従って，経験的に，拡散画像マッチングは密な視差を提供することが知られているが，その精度は特徴点マッチングよりも劣る場合があるといえる．本提案では，特徴点マッチングによって得られる精度の高い点群と，拡散画像マッチングによって得られる密な点群の両方を活用する，統合的なアプローチを提案します．具体的には，初めに特徴点マッチングにより得られる精度の高い視差から水面の表面法線を推定する．次に，水面が比較的滑らかであると仮定して，推定された表面法線をシーン全域に対して補間する．最後に，補間された法線マップを式 23 の $\mathbf{n}$ とし，拡散画像マッチングによって得られた点群の視差を式 25 の $\mathbf{d}_{12}$ として与えることで，密な三次元の水に対象物上の点を再構成する．

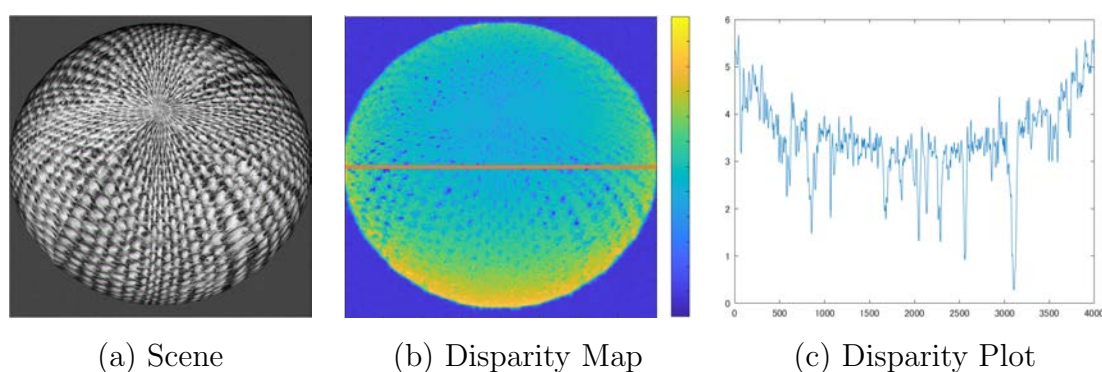


図 41: シミュレートされたシーンに対する拡散画像マッチングによる密な点群．

5.3 実験

本節では、シミュレーションによる実験と試作機を用いた実環境における実験の両方を通して、提案した水面および水中シーンの再構成手法の有効性を確認する。

定量的な評価のための水面法線誤差と深度誤差については、以下の評価指標を用いる：

$$\begin{aligned} E_{angle} &= \arccos(\bar{\mathbf{n}} \cdot \mathbf{n}), \\ E_{depth} &= \bar{w}_z - \hat{w}_z, \end{aligned} \quad (35)$$

$\bar{\mathbf{n}}$ と $\bar{w}_z$ はそれぞれ、水面法線 $\mathbf{n}$ と水中対象物の表面点から水面までの距離 w_z の真値を表す。ここで、 w の添え字 z は、その z 軸に沿う座標値を表すものとする。抽出された全ての特徴点に対してこれらの誤差を計算し、シーン全域の平均誤差と標準偏差を評価する。

定性評価では、平らでない形状をもつ屈折面や水中物体を含むシーンについて、水面法線および水中対象物の表面点の再構成結果を図示する。定性評価における水中対象物の表面点を示す三次元のプロットでは、水面を基準とした推定水中深度を負の値として表示する。

5.3.1 シミュレーションによる実験

ここでは、シミュレーションによる理想的な場合について、いくつかシーンに対する定量評価および定性評価を通して提案手法の有効性を検証する。レンダリングソフト *Blender 2.91* を用いて、屈折率を $k_{1,2,3} = \{1.34, 1.33, 1.32\}$ とした媒体中の対象物を正投影カメラで撮影したシーンをシミュレーションしている。ここで、屈折率の値は可視光域の水中の屈折率に基づいて設定され、三つの異なる波長で撮影されたシーンを仮想的にレンダリングするために使用される。

初めに、深さの異なる様々なテクスチャの平面物体からなるシーンについて、水平な水面を通して観測されたシーンをカメラと水面との入射角を様々に変化させた場合の再構成精度を評価する。表4と表5には、最適化法とルックアップテーブル法の場合について、それぞれ E_{angle} と E_{depth} の平均誤差と標準偏差を示す。さらに、より深部に位置する平面物体からなるシーンについての結果を表6と、表7に示す。全ての結果に対する共通の傾向として、推定された水面法線の精度は、カメラと水面の入射角の増加に伴って向上することがわかる。これは、図42に示すように、入射角が大きいほど水面での屈折の影響が強くなり、波長画像間の視差が大きくなるためである。波長画像間の視差が大きくなるほど、特徴点マッチングの誤差が減少し、全体の推定

精度が向上するといえる．同様の理由で，水中の深い位置に配置した平面物体に対する推定結果に着目すると，異なる波長画像間の視差が大きくなるために推定精度がよくなることが考えられる．実際に表 6 と 表 7 の結果を比較すると，水深の深い場所に物体がある場合の推定精度が向上していることがわかる．他方，表 4 と 表 6 について平均誤差に着目すると，最適化法の再構成精度は，ほとんどの場合で LUT 法よりも優れているという結果が観測される．対照的に，標準偏差について着目すると，LUT 法が最適化法に対して優れており，ほとんどの場合で安定しているといえる．

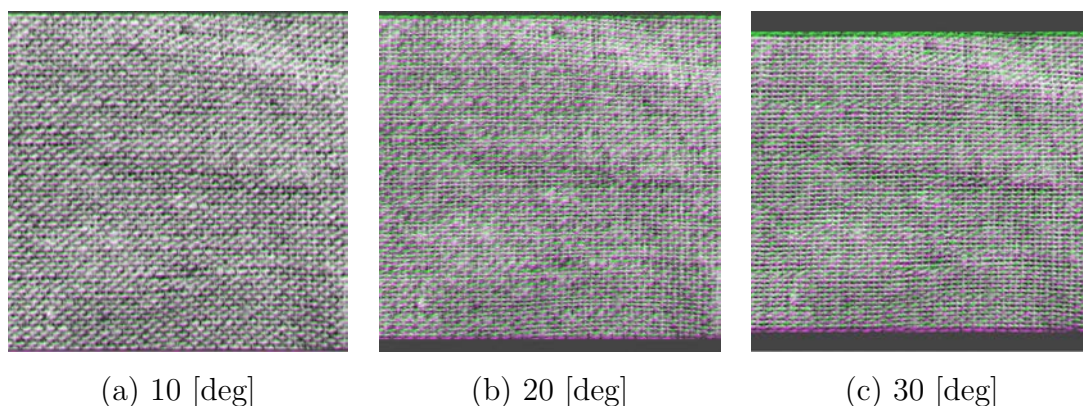


図 42: カメラと水面の入射角の増加による波長画像間の視差との関係．

続いて，シミュレーションを用いたより複雑な水面および水中対象物を含むシーンに対する提案手法の有効性を検証する．波状の屈折面と平らな水中物体からなるシーンに対する再構成結果を図 43 に示す．(a) は二種類の異なる波長画像に抽出された特徴点を重畳表示した入力画像である．(b,c) は (a) における赤く囲われた部分と緑で囲われた部分の拡大画像である．(d) はレンダリングソフト上で作成したシーンであり，カメラ，水面，水中対象物の位置関係を表している．(e,f) では法線ベクトルの x, y, z 成分を r, g, b 色にマッピングしたカラーマップを用いて，水面の法線を可視化している．正解の法線マップ (e) と推定された法線マップ (f) とが類似していることが確認できる．また，推定された法線マップから，水面形状 (g) も正しく再構成されていることがわかる．さらに，水中の平面物体については，特徴点マッチング (h) と拡散画像マッチング (i) の両方の手法に基づいて，水中光路の推定から水中平面が適切に再構成されていることがわかる．

図 44 は平坦な水面と水中の球体からなるシーンの再構成結果を表示している．特徴点マッチング (g) と拡散画像マッチング (h) で得られた再構成結果を比較すると，拡散画像マッチング法で水中物体表面の密な点が効果的に取得できていることがわかる．

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	16.4 ± 4.03	9.10 ± 3.45	4.09 ± 2.42	2.35 ± 1.72	2.81 ± 1.93	2.50 ± 1.58
rock	17.0 ± 3.86	8.69 ± 3.45	2.70 ± 2.28	2.50 ± 1.56	1.80 ± 1.35	2.56 ± 1.63
sand	16.8 ± 4.36	7.47 ± 2.73	3.64 ± 2.21	2.38 ± 1.73	3.19 ± 2.07	2.46 ± 1.66
stones	14.8 ± 3.83	5.81 ± 3.06	2.79 ± 2.11	1.92 ± 1.62	2.75 ± 1.70	2.49 ± 1.50
tiles	15.3 ± 4.28	9.25 ± 3.44	4.38 ± 2.55	2.39 ± 1.65	2.68 ± 1.90	2.71 ± 1.88

E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	18.8 ± 1.21	11.2 ± 2.64	4.59 ± 2.99	1.76 ± 3.90	2.85 ± 2.70	1.90 ± 1.55
rock	19.0 ± 1.06	10.9 ± 2.67	2.52 ± 3.45	2.62 ± 3.44	1.19 ± 2.02	1.73 ± 1.82
sand	18.9 ± 1.20	10.1 ± 2.38	3.95 ± 3.16	1.95 ± 3.79	3.38 ± 2.96	1.76 ± 1.74
stones	18.3 ± 1.25	16.8 ± 6.37	0.22 ± 5.36	1.45 ± 3.52	2.34 ± 2.70	1.85 ± 1.51
tiles	18.4 ± 1.37	11.3 ± 4.70	4.80 ± 4.30	0.42 ± 3.47	2.38 ± 2.88	2.19 ± 4.30

表 4: 25mm の水深における最適化法による推定結果.

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	15.2 ± 1.04	9.55 ± 1.12	5.35 ± 0.99	0.99 ± 0.68	2.88 ± 0.85	4.95 ± 0.64
rock	15.0 ± 1.12	9.79 ± 0.97	4.66 ± 1.20	0.89 ± 0.82	1.64 ± 0.78	4.84 ± 0.74
sand	15.5 ± 1.13	10.1 ± 0.99	4.92 ± 1.09	1.07 ± 0.72	2.57 ± 0.83	4.72 ± 0.67
stones	15.0 ± 0.90	9.24 ± 0.97	4.60 ± 1.19	0.75 ± 0.58	3.03 ± 0.71	4.71 ± 0.71
tiles	15.9 ± 1.22	9.75 ± 1.18	5.46 ± 1.13	1.24 ± 0.86	2.51 ± 0.79	4.90 ± 0.72

E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	18.7 ± 0.46	11.9 ± 1.03	6.23 ± 1.29	0.43 ± 1.82	3.02 ± 1.38	4.26 ± 0.97
rock	18.6 ± 0.49	12.1 ± 0.89	5.63 ± 1.44	0.21 ± 1.74	1.67 ± 1.16	4.15 ± 1.06
sand	18.7 ± 0.46	12.2 ± 0.90	5.83 ± 1.52	0.50 ± 1.79	2.65 ± 1.39	4.03 ± 1.00
stones	18.6 ± 0.41	11.7 ± 0.97	5.57 ± 1.59	0.17 ± 1.49	3.16 ± 1.23	4.07 ± 1.15
tiles	18.8 ± 0.51	12.0 ± 3.02	6.27 ± 3.35	0.71 ± 1.87	2.57 ± 1.33	4.32 ± 4.02

表 5: 25mm の水深における LUT 法による推定結果.

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	13.4 ± 3.20	4.76 ± 2.30	1.72 ± 1.28	2.46 ± 1.54	2.24 ± 1.47	0.99 ± 0.78
rock	12.5 ± 3.05	4.66 ± 2.28	1.63 ± 1.16	2.12 ± 1.46	1.65 ± 1.05	1.08 ± 0.69
sand	14.6 ± 3.57	5.10 ± 2.12	1.72 ± 1.36	2.29 ± 1.54	2.15 ± 1.41	0.88 ± 0.66
stones	13.4 ± 2.80	5.81 ± 3.06	1.89 ± 1.20	2.27 ± 1.31	2.79 ± 1.48	1.43 ± 0.98
tiles	14.1 ± 3.24	5.19 ± 2.34	1.76 ± 1.33	2.04 ± 1.48	2.17 ± 1.40	1.52 ± 1.09

E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen	35.8 ± 2.62	15.1 ± 5.49	1.08 ± 6.48	6.88 ± 5.42	4.37 ± 3.76	1.24 ± 1.46
rock	35.1 ± 2.70	14.7 ± 5.62	1.09 ± 7.31	5.42 ± 5.37	3.10 ± 2.53	1.44 ± 1.28
sand	36.5 ± 2.59	15.9 ± 4.76	0.77 ± 6.60	5.88 ± 5.74	4.12 ± 3.73	0.83 ± 1.41
stones	35.9 ± 2.37	16.8 ± 6.37	1.10 ± 7.84	5.80 ± 4.84	5.89 ± 3.57	2.27 ± 6.10
tiles	36.3 ± 2.35	15.8 ± 5.88	0.28 ± 8.31	4.14 ± 6.16	4.27 ± 3.88	2.08 ± 2.06

表 6: 50mm の水深における最適化法による推定結果.

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen*	14.1 ± 1.15	8.33 ± 1.15	3.82 ± 0.83	0.62 ± 0.49	1.52 ± 0.79	2.86 ± 0.39
rock	14.2 ± 3.05	8.62 ± 0.85	3.57 ± 0.98	0.79 ± 0.69	1.00 ± 0.55	2.47 ± 0.42
sand	14.4 ± 1.13	8.77 ± 1.15	3.91 ± 0.99	0.73 ± 0.56	1.39 ± 0.69	2.49 ± 0.37
stones	14.2 ± 0.86	8.66 ± 1.27	3.18 ± 1.02	0.75 ± 0.65	1.56 ± 0.64	3.42 ± 0.54
tiles	14.6 ± 0.90	8.42 ± 1.24	3.78 ± 0.93	0.96 ± 0.68	1.55 ± 0.71	3.24 ± 0.67

E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
linen*	36.6 ± 1.09	22.1 ± 2.41	9.65 ± 2.64	0.17 ± 2.63	3.05 ± 2.37	4.51 ± 1.11
rock	36.7 ± 2.70	22.6 ± 1.94	9.13 ± 2.84	0.20 ± 3.12	1.90 ± 1.63	3.82 ± 0.98
sand	36.9 ± 1.03	22.8 ± 2.18	9.77 ± 2.98	0.01 ± 3.11	2.70 ± 2.12	3.86 ± 0.98
stones	36.7 ± 0.89	22.6 ± 2.46	8.21 ± 3.18	0.27 ± 3.00	3.09 ± 1.93	5.70 ± 5.94
tiles	37.0 ± 0.90	22.3 ± 2.48	9.42 ± 4.12	0.72 ± 3.43	3.13 ± 2.67	5.18 ± 1.57

表 7: 50mm の水深における LUT 法による推定結果.

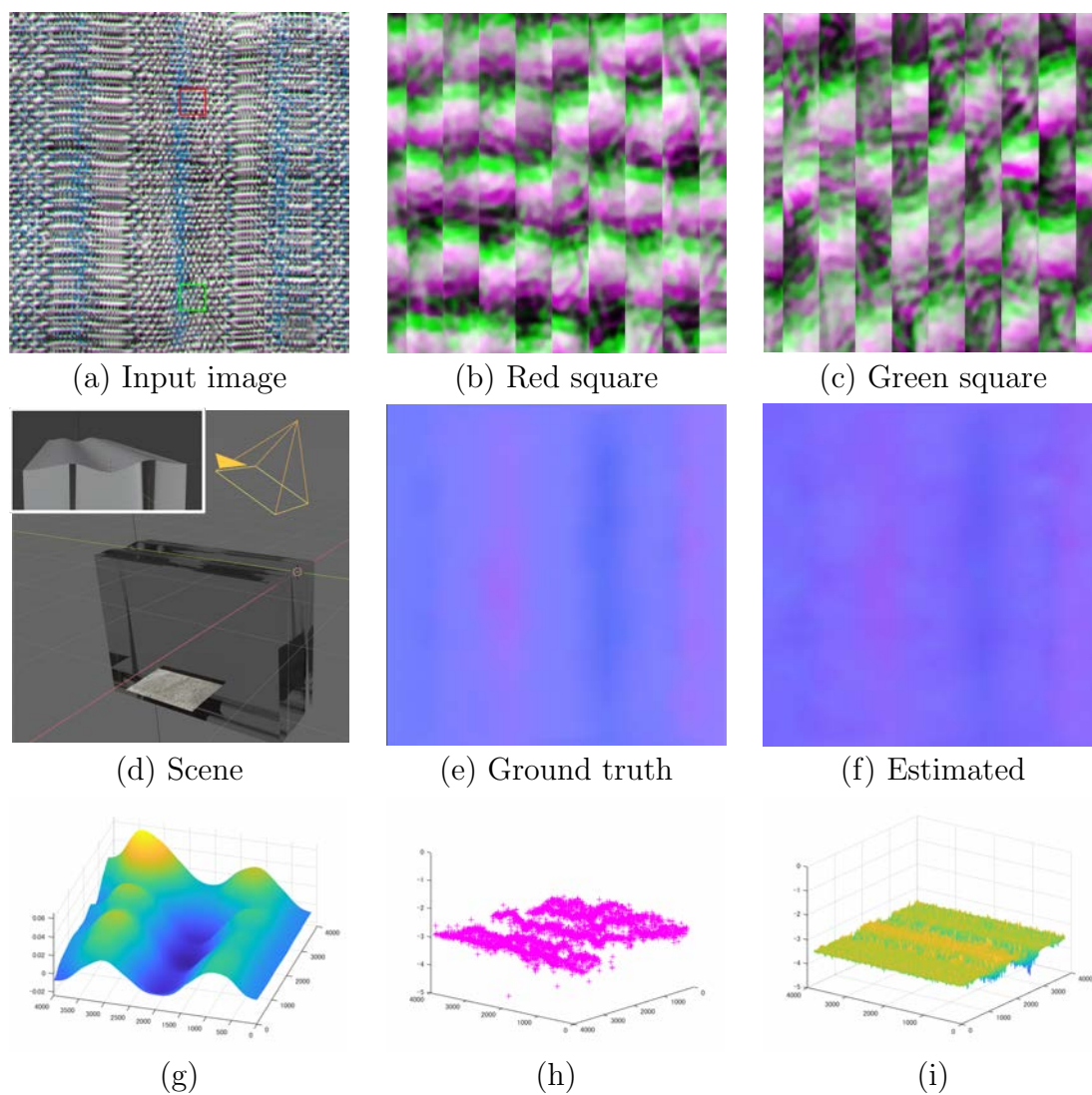


図 43: シミュレーションによる波状水面および水中の平面における再構成結果.

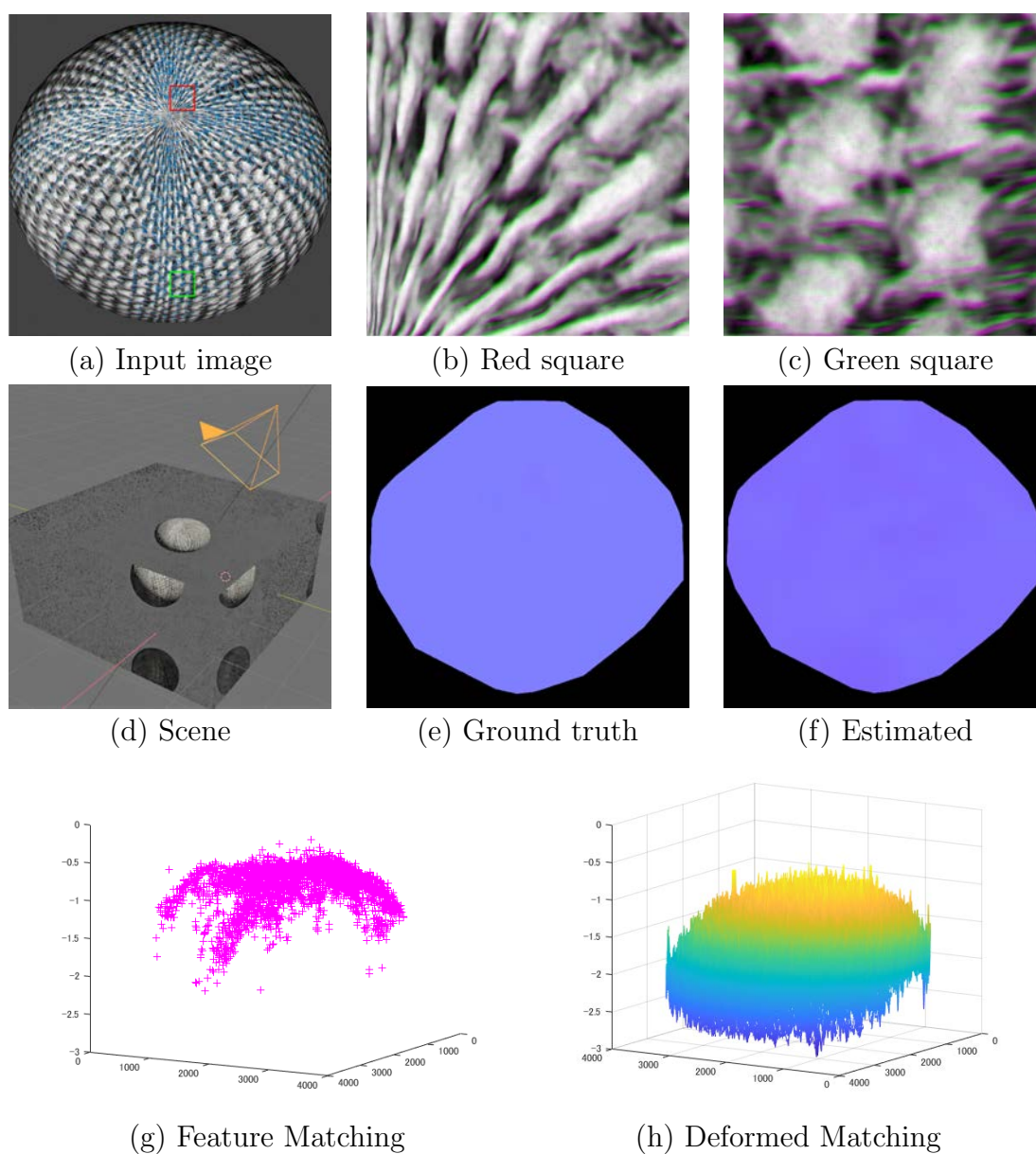


図 44: シミュレーションによる水平水面および水中の球体における再構成結果.

5.3.2 実環境での実験

実環境における実験では提案システムの試作機として、図 39 に示すグレースケールカメラ *FLIR GS3-U3-123S6* とテレセントリックレンズ *Edmund Optics 1.0X SilverTL* を用いて、広帯域の可視光と可視光源をそれぞれ5種類の狭帯域フィルタを組み合わせ、 $\lambda_i = \{450, 500, 550, 600, 700\}nm$ を中心波長とする分光画像を撮影する。5つのチャンネル間の画像アライメントは、データ取得前に事前のキャリブレーションを行っている。また、5つの波長画像について、強度ヒストグラムが互いに近くなるように全体の画像コントラストを調整する後処理を行う。

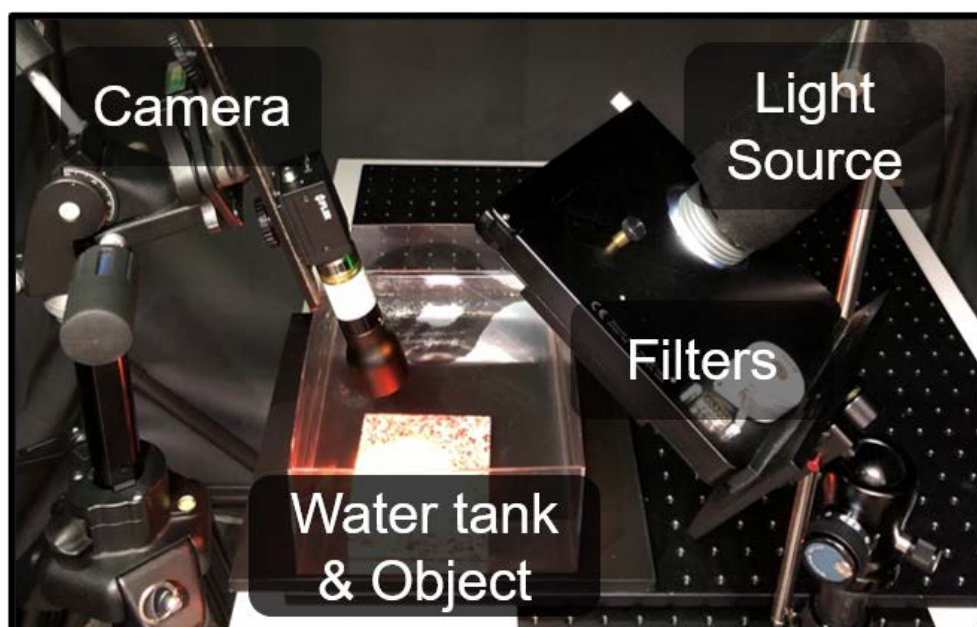


図 45: 提案システムの試作機（再掲）。

試作システムでは画像の歪みやノイズによって、最適化によるパラメータ推定を行うことが困難であるため、本節ではLUT手法を用いたパラメータ推定のみを行っている。シミュレーションの場合と同様に、深さの異なる様々なテクスチャの平面物体からなるシーンについて、水平な水面を通して観測されたシーンをカメラと水面との入射角を様々に変化させた場合の再構成精度を評価する。表 8 と表 9 は、対象物を浅い深さ (25mm) と深い深さ (50mm) に配置した場合の誤差 E_{angle} と E_{depth} とその標準偏差をそれぞれ示している。表 8 と表 9 ではシミュレーションの場合と同様に、入射角が大きいほど、また水深が深いほど、ほとんどのケースで推定精度が向上していることが確認できる。

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
spray	9.63 ± 2.26	7.14 ± 2.02	4.16 ± 1.31	3.22 ± 0.89	3.58 ± 0.77	6.40 ± 1.01
smooth cloth	14.8 ± 1.77	7.21 ± 1.80	8.29 ± 2.14	4.49 ± 0.88	2.38 ± 0.47	4.26 ± 0.47
mid. smooth	16.0 ± 1.68	7.21 ± 1.80	5.46 ± 1.44	3.21 ± 1.12	3.43 ± 0.91	10.3 ± 1.04
mid. rough	10.3 ± 2.19	10.7 ± 2.04	6.83 ± 1.47	4.33 ± 1.31	3.30 ± 0.74	4.90 ± 1.03
rough cloth	11.6 ± 1.61	7.43 ± 1.91	6.28 ± 1.35	5.56 ± 0.86	4.00 ± 1.31	7.33 ± 0.83
E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
spray	15.9 ± 2.62	7.06 ± 2.59	3.44 ± 2.33	0.14 ± 2.21	4.40 ± 2.11	7.00 ± 2.00
smooth cloth	14.6 ± 2.01	6.34 ± 3.05	0.68 ± 13.4	0.13 ± 2.19	1.75 ± 2.00	6.31 ± 2.00
mid. smooth	16.8 ± 1.77	6.34 ± 3.05	5.02 ± 2.60	0.25 ± 3.24	0.29 ± 2.50	9.33 ± 4.50
mid. rough	14.9 ± 2.75	10.5 ± 2.12	4.40 ± 2.02	0.97 ± 2.21	3.16 ± 2.14	6.12 ± 2.35
rough cloth	11.1 ± 3.53	0.08 ± 4.98	5.01 ± 3.08	13.1 ± 4.76	5.52 ± 4.28	8.94 ± 4.41

表 8: 25mm の水深における LUT 法による推定結果（実環境）。

E_{normal} : Error of normal estimation [deg]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
spray*	1.08 ± 0.97	2.83 ± 1.62	3.80 ± 2.05	2.55 ± 1.42	1.49 ± 0.85	3.30 ± 2.05
smooth cloth	5.24 ± 1.43	4.42 ± 1.40	1.21 ± 0.62	4.00 ± 0.71	1.59 ± 0.76	2.83 ± 0.89
mid. smooth	4.48 ± 2.51	2.71 ± 1.77	1.96 ± 1.32	2.59 ± 1.51	2.38 ± 1.05	3.10 ± 1.42
mid. rough	6.51 ± 2.92	2.29 ± 1.20	1.92 ± 1.11	1.67 ± 1.04	4.76 ± 1.52	5.41 ± 3.55
rough cloth	5.76 ± 0.82	4.17 ± 1.35	2.38 ± 1.35	2.55 ± 1.57	2.35 ± 1.46	9.06 ± 3.96
E_{depth} : Error of water depth estimation [mm]						
Pattern \ Angle	$\theta = 5$ [deg]	10	15	20	25	30
spray*	7.85 ± 4.91	8.36 ± 4.05	5.93 ± 4.35	1.86 ± 3.33	2.53 ± 1.64	11.8 ± 4.21
smooth cloth	20.0 ± 3.17	13.6 ± 2.25	3.98 ± 2.40	4.09 ± 0.91	4.63 ± 1.12	11.6 ± 1.39
mid. smooth	17.8 ± 5.21	6.53 ± 6.86	0.88 ± 3.79	2.20 ± 4.73	1.83 ± 1.63	11.1 ± 2.22
mid. rough	22.9 ± 4.91	9.49 ± 2.77	3.09 ± 1.90	1.42 ± 2.12	10.3 ± 2.31	16.1 ± 5.02
rough cloth	21.8 ± 1.75	13.1 ± 2.47	3.61 ± 4.25	4.25 ± 3.17	9.57 ± 3.63	23.9 ± 13.7

表 9: 50mm の水深における LUT 法による推定結果（実環境）。

続いて、実環境におけるより複雑な水面および水中対象物を含むシーンに対する提案手法の有効性を検証する。湾曲する水面の屈折面と、水中平面を含むシーンの再構成結果を図 46 に示す。シミュレーションと同様、(a) は二種類の異なる波長画像に抽出された特徴点を重畳表示した入力画像、(b,c) は (a) における赤く囲われた部分と緑で囲われた部分の拡大画像である。(d) は撮影に用いているシーンを一般的な RGB カメラで撮影した画像であり、水面と水中対象物の様子や位置関係を表している。(e,f) では法線ベクトルの x, y, z 成分を r, g, b 色にマッピングしたカラーマップを用いて、水面の法線を可視化している。シミュレーションと同様に、正解の法線マップ (e) と推定された法線マップ (f) とが類似していることが確認できる。また、推定された法線マップから、水面形状 (g) も正しく再構成されていることがわかる。さらに、水中の平面物体については、特徴点マッチング (h) と拡散画像マッチング (i) の両方の手法に基づいて、水中光路の推定から水中平面が適切に再構成されていることがわかる。

図 47 に平らでない水中物体を含むシーンの水深推定結果を示す。特徴点マッチング (g) 法と拡散画像マッチング (h) 法を用いて再構成した水中シーンにおいて、三角形の物体の傾きを復元することができている。

これらの結果から、試作システムによる実環境における水面や水中シーンの再構成が、シミュレーションと同様に機能することが示された。

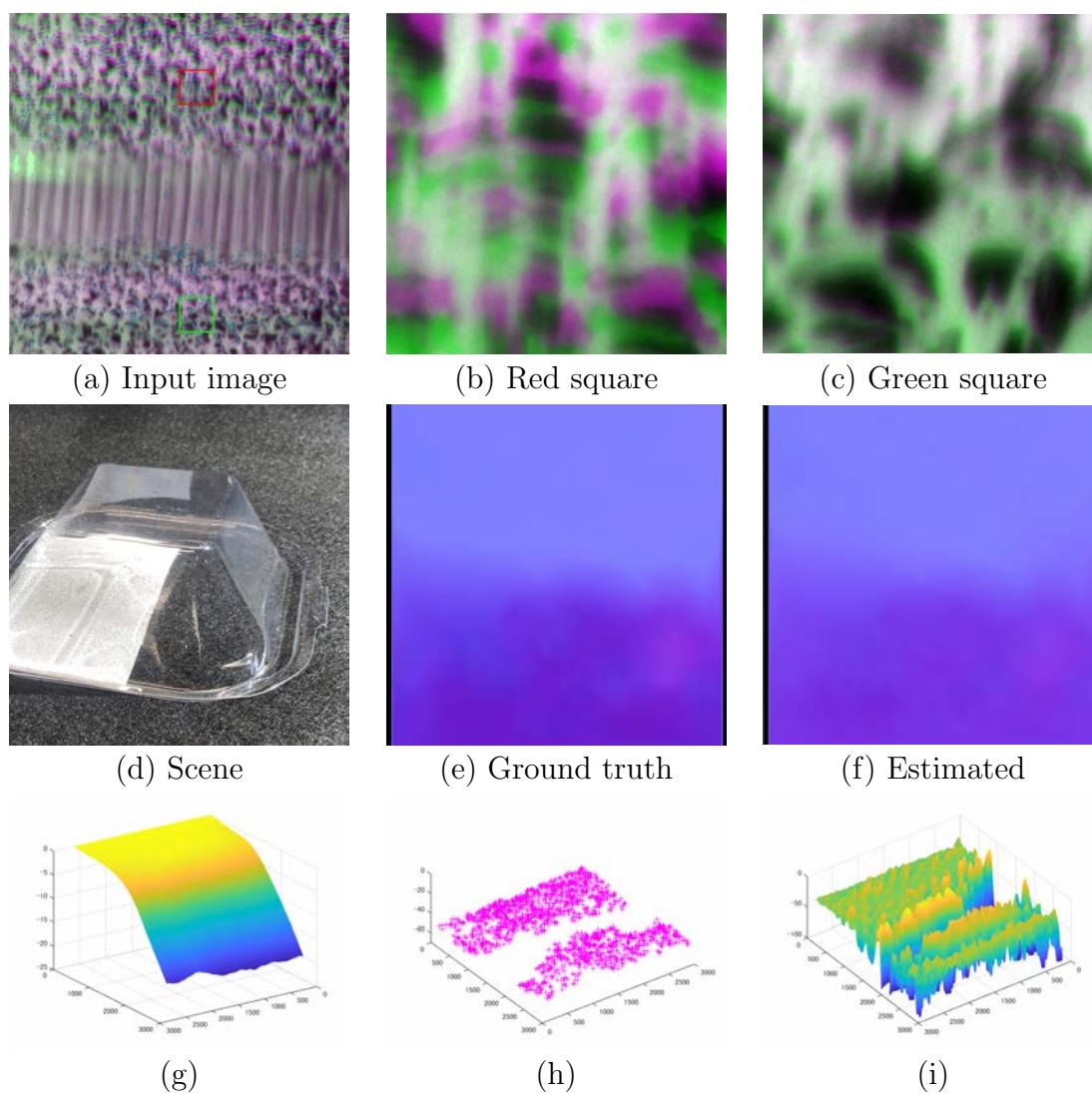


図 46: 試作システムによる湾曲した屈折水面および水中平面の再構成結果.

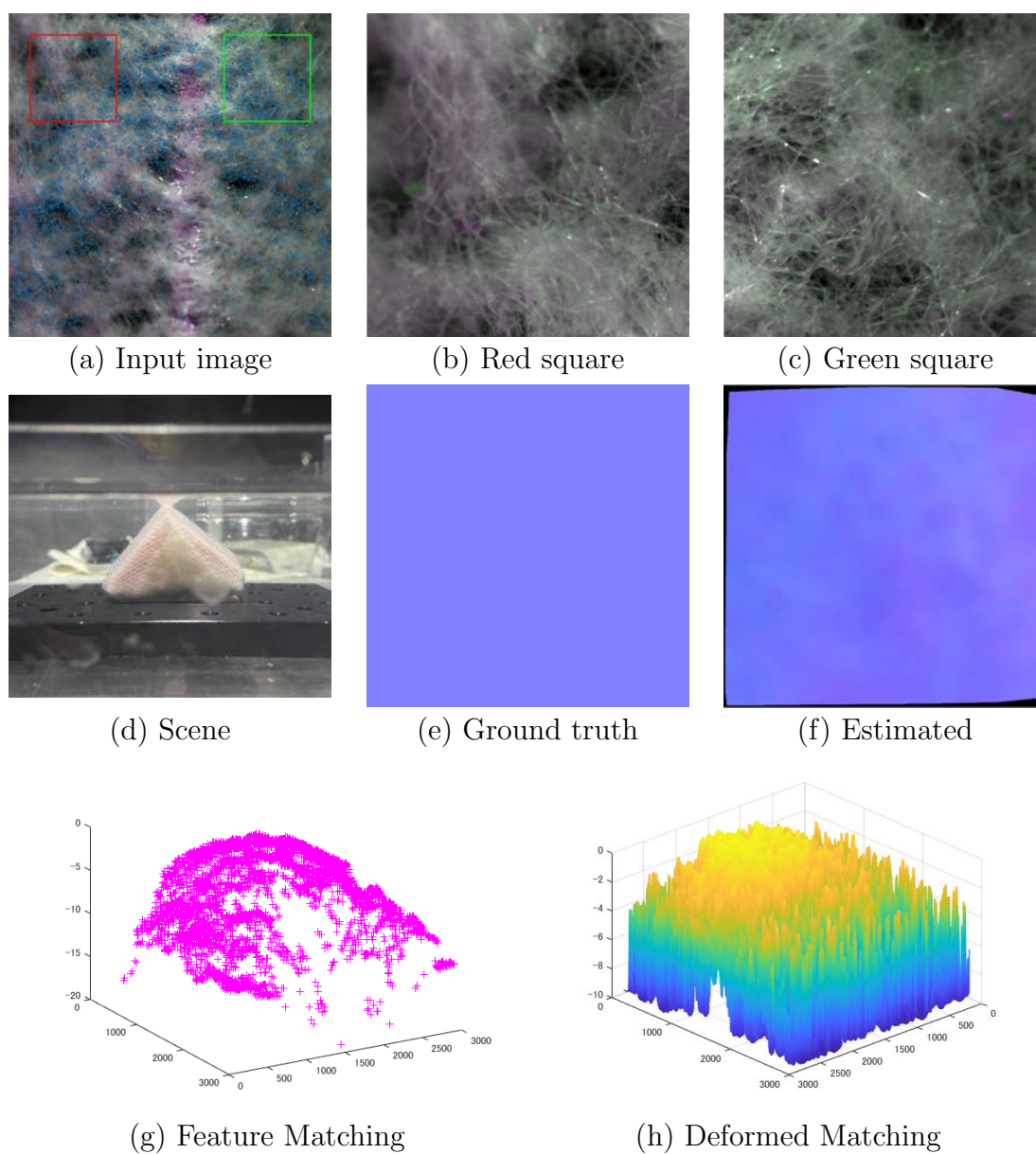


図 47: 試作システムによる水平水面および三角形のオブジェクトの再構成結果.

5.4 小括

水の波長依存性屈折率を利用して、マルチスペクトル画像から水面が不明な水中シーンを取得する手法を発表した。スネルの法則により、入射した光は水面に入射すると屈折し、その方向は屈折率と水面と光との角度によって決まります。屈折率は波長に依存するため、異なる波長で撮影された画像間には多少の画像格差が生じます。

この画像の視差と水面の表面法線、水中物体の形状との関係を丁寧に調べました。本研究では、マルチスペクトル画像から視差の原因となる成分を分解し、水面上の屈折の法線成分と深さ成分を別々に推定することが可能であることを明らかにした。

また、実際に試作した画像処理装置を用いたシミュレーションと実実験により、本手法の有効性を様々なシーンで実証することができました。プロトタイプシステムを用いた実験の結果、本手法が様々なテクスチャを持つシーンに適用可能であることが示された。

第6章 結論

本章では、本論文を通して達成された事項および今後の課題と展望についてまとめ、最後に総括を述べる。

6.1 今後の課題と展望

6.1.1 高度な実世界理解

本提案の発展的な目標は分光情報を利用した質感解析技術との融合による高度な実世界理解の実現である。本論文で扱ってきた対象物は一般的な反射特性を有する物体が殆どであり、質感解析の対象となるような特殊な反射特性を有する物体を対象とした精度評価等を行えていない。従って、様々な反射特性を有する物体を対象とした提案手法の精度評価を行い、質感解析技術と組み合わせたシステムの実現可能性について検討していく。

本提案にて扱ってきた分光画像は主に可視光域（400nm-700nm）で、50nm 間隔、10nm の半値幅を持つ数枚の画像群である。一般に取得可能な分光画像の種類は多岐にわたり、近紫外（400nm 以下）や近赤外（700nm 以上）とよばれる波長域の光や、より狭い波長間隔や半値幅を持つ画像を扱うことも可能である。環境によって用いることのできる光の種類が異なる場合があるため、様々な光の種類を用いた場合についても精度評価を行っていくことで、提案手法の応用可能性を広げることができる。

6.1.2 波長依存性の活用についての展望

本提案では、光の屈折における波長依存性に着目し、レンズ内部で生じる屈折と水面で生じる屈折の2種類を活用した深度推定手法を提案してきた。本研究を応用した研究の展望としては、シーンの深度ではなく屈折率の絶対値を推定することなどが考えられる。媒質の屈折率を対象とする場合、容器の形状が既知であれば本論文で提案した理論を応用することで簡易に屈折率を計測でき、媒質の種類の特特定や溶質の濃度推定等への活用が期待できる。

また発展として、屈折率だけでなく散乱係数などの物理パラメータにおける波長依存性を活用することが考えられる。特に散乱現象はその係数のみでなく粒子径や表面の状態（乾いている、湿っている等）によっても振る舞いが増えるため、コンピュータビジョン研究では解析が非常に困難な対象として知られている。このような複雑な光学現象を解析するための手がかりとして、分光画像による波長依存性を利用した解析手法が一助となることを期待し、研究課題として取り組んでいく。

6.1.3 システムの実応用についての課題と展望

本研究の実応用に向けた課題として、システムの動作範囲（ワーキングレンジ）を広げることが必要である。レンズの屈折を利用した距離推定については、より広範囲の波長域の分光画像を使用することによって屈折率の変化を大きくすることと、光学系のF値を大きくして被写界深度を深くすることでシステムのフォーカスする範囲を広げることが考えられる。被写界深度を深くすることにより波長画像間におけるボケの差が減少し精度が低下することが予想されるが、推定に使用する波長画像を増やすことによって距離分解能を確保することが可能である。水面の屈折を利用した再構成については、テレセントリック光学系の被写界深度が非常に浅いことが現実的な問題としてあげられる。先行研究として、液体レンズを利用して高速に焦点面を移動させながら露光すること（フォーカススweep）によってテレセントリック光学系の被写界深度を拡張する例が存在する[47]ため、これを利用してシステムの十分な被写界深度を確保することが必要である。

本研究の産業的な展望としては、生産コストや撮影コストの削減およびシステムの小型化等を目指すことが考えられる。これらを同時に達成するための方策として、機械学習による多波長重畳画像の分離を検討する。ここで考える多波長重畳画像とは、複数の波長画像が1つのチャンネルに重畳された画像を差し、例として複数のピークを持つ帯域透過フィルタとモノクロセンサにより実現される。本研究で提案したシステムにおける分光画像では波長毎に空間的なボケの様子または空間方向のシフト量が異なるため、シーン色が複雑でない場合に限定するか、照明条件を制限するなどして重畳された波長画像を分離することが十分可能であると考えている。多波長重畳画像の分離により、回転フィルタ等による多波長撮影の撮影コストが低減され、システムの小型化が実現できる。

6.2 総括

本論文では、光の屈折における波長依存性に着目し、レンズ内屈折における波長依存の焦点ボケの変化と、水面での屈折における波長依存性によって生じる画像の変形を利用した三次元情報の再構成手法を提案した。

レンズ内屈折における波長依存の焦点ボケの変化を利用した深度推定手法では、レンズを形成するガラスの実効的な屈折率が光の観測波長によって変化する性質を利用した。従来より焦点ボケの強さは対象点の距離に応じて変化することが知られており、カメラのパラメータを変化させることで異なる焦点ボケを持つ画像を取得することで深度推定を行う DfD 手法が提案されている。本論文では、対象点の深度および画像の観測波長の両方に依存する焦点ボケについて、分光モデルを用いた解析による深度推定手法を提案した。また、狭帯域の波長画像を観測できるマルチバンドカメラを用いたシステムを実装し、様々なシーンに対する定量評価および定性評価を通して本手法の有効性を示した。色収差を用いた従来手法では対象シーンの離散的な深度のみ推定可能であったのに対し、本提案ではステレオや DfD 手法と同様にシーンの連続的な距離を推定できることを示した。

水中シーンの再構成手法では、正投影の分光カメラ光学系において空気中から水中の対象を観測した場合に、水面での屈折における波長依存性によって異なる観測波長の光線間に視差が生じる現象を利用した。このとき生じる視差の大きさはカメラと水面の入射角および水中対象点の深度に依存し、視差の方向は水面の法線方向に依存する。これらの現象を分光カメラを用いた解析システムとしてモデル化し、一枚の分光画像から水面形状および水中の三次元シーンを再構成するための手法を提案した。シミュレーションおよび実環境での試作システムによる実験を通して、本手法の有効性を検証した。単一視点の一枚の画像のみから水面形状と水中の三次元シーンを同時に再構成する手法は過去になく、本提案がコンピュータビジョン研究における水中シーンの解析に貢献できることが期待される。

本論文では、上記二つの提案を通して、屈折光の波長解析による分光画像からの三次元情報推定を実現してきた。分光情報を手掛かりとする利点の一つとして、一台のカメラから得られる一枚の画像情報から三次元情報を推定できることが挙げられる。従来手法では複数のカメラを用いたシステムや条件を変えて複数回の撮影を必要とするシステムが一般的であったが、これらの手法では一台のカメラのほかに必要な装置があるため、小型化が難しくやキャリブレーションが必要であるという点でコストが大きい。提案手法のように一台のカメラから得られる一枚の画像情報のみを用いることは、深度推定のタスクにおいてこれらの点で優位な面をもつ。また、第2章の関連研

究の項目でも挙げたように，分光情報はコンピュータビジョンの分野において高度な実世界理解のための質感解析技術の実現にも用いられている．本提案では実世界理解のための三次元情報の推定に着目した手法を実現してきたが，さらなる発展として質感情報の取得と組み合わせたより高度な実世界理解を実現していくことが考えられる．

謝辞

本博士論文は，筆者が東京工業大学工学院情報通信系情報通信コース博士課程在学中に，佐藤いまり研究室において行った研究をまとめたものです．本研究に関してご指導いただきました佐藤いまり教授ならびに金子寛彦教授に心より感謝いたします．本論文をご精読いただき，ご意見を頂きました本学の金子寛彦教授，山口雅浩教授，熊澤逸夫教授，渡辺義浩准教授，長谷川晶一准教授に深謝いたします．

最後に，様々なアドバイスを頂いた研究室の皆様，支えとなっていた家族や友人に感謝いたします．

参考文献

- [1] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3346–3353. IEEE, 2012.
- [2] E. Alexander, Q. Guo, S. Koppal, S. Gortler, and T. Zickler. Focal flow: Measuring distance and velocity with defocus and differential motion. In *European conference on computer vision*, pages 667–682. Springer, 2016.
- [3] Y. Asano, Y. Zheng, K. Nishino, and I. Sato. Shape from water: Bispectral light absorption for depth recovery. In *European Conference on Computer Vision*, pages 635–649. Springer, 2016.
- [4] S. T. Barnard and M. A. Fischler. Computational stereo. Technical report, SRI INTERNATIONAL MENLO PARK CA ARTIFICIAL INTELLIGENCE CENTER, 1982.
- [5] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.
- [6] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007.
- [7] J. Batlle, E. Mouaddib, and J. Salvi. Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern recognition*, 31(7):963–982, 1998.
- [8] G. Bergmann. Weak localization in thin films: a time-of-flight experiment with conduction electrons. *Physics Reports*, 107(1):1–58, 1984.

- [9] A. Chakrabarti and T. Zickler. Depth and deblurring from a spectrally-varying depth-of-field. In *European Conference on Computer Vision*, pages 648–661. Springer, 2012.
- [10] V. Chari and P. Sturm. Multiple-view geometry of the refractive plane. In *British Machine Vision Conference*, pages 1–11, 2009.
- [11] D. T. Delpy, M. Cope, P. van der Zee, S. Arridge, S. Wray, and J. Wyatt. Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine & Biology*, 33(12):1433, 1988.
- [12] T. Dolereit and U. F. von Lukas. Calibration of shared flat refractive stereo systems. In *International Conference on Image Analysis and Recognition*, pages 433–442. Springer, 2016.
- [13] P. Favaro. Depth from focus/defocus. *document dated Jun, 25, 2002*.
- [14] P. Favaro and S. Soatto. A geometric approach to shape from defocus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):406–417, 2005.
- [15] P. Favaro, S. Soatto, M. Burger, and S. J. Osher. Shape from defocus via diffusion. *IEEE transactions on pattern analysis and machine intelligence*, 30(3):518–531, 2008.
- [16] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, 10(4):439–451, 1988.
- [17] J. Gedge, M. Gong, and Y.-H. Yang. Refractive epipolar geometry for underwater stereo matching. In *2011 Canadian Conference on Computer and Robot Vision*, pages 146–152. IEEE, 2011.
- [18] Q. Guo, E. Alexander, and T. E. Zickler. Focal track: Depth and accommodation with oscillating lens deformation. In *ICCV*, pages 966–974, 2017.
- [19] M. Hansard, S. Lee, O. Choi, and R. P. Horaud. *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media, 2012.

- [20] H. Hariharan, A. Koschan, and M. Abidi. An adaptive focal connectivity algorithm for multifocus fusion. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–6. IEEE, 2007.
- [21] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [22] R. I. Hartley and P. Sturm. Triangulation. *Computer vision and image understanding*, 68(2):146–157, 1997.
- [23] L. Hashemi, A. Azizi, and M. Hashemi. Implementation of a single photo shape from shading method for the automatic dtm generation. *INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY REMOTE SENSING AND SPATIAL INFORMATION SCIENCES*, 34(3/B):71–74, 2002.
- [24] K. Ichimaru and H. Kawasaki. Underwater stereo using refraction-free image synthesized from light field camera. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1039–1043. IEEE, 2019.
- [25] K. Ichimaru, Y. Taguchi, and H. Kawasaki. Unified underwater structure-from-motion. In *2019 International Conference on 3D Vision (3DV)*, pages 524–532. IEEE, 2019.
- [26] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):661–669, 1981.
- [27] K. Ikeuchi and B. K. Horn. Numerical shape from shading and occluding boundaries. *Artificial intelligence*, 17(1-3):141–184, 1981.
- [28] H. Jin and P. Favaro. A variational approach to shape from defocus. In *European Conference on Computer Vision*, pages 18–30. Springer, 2002.
- [29] A. Jordt-Sedlazeck and R. Koch. Refractive calibration of underwater cameras. In *European conference on computer vision*, pages 846–859. Springer, 2012.
- [30] A. Jordt-Sedlazeck and R. Koch. Refractive structure-from-motion on underwater images. In *Proceedings of the IEEE international Conference on Computer Vision*, pages 57–64, 2013.

- [31] R. Kawahara, S. Nobuhara, and T. Matsuyama. Dynamic 3d capture of swimming fish by underwater active stereo. *Methods in Oceanography*, 17:118–137, 2016.
- [32] K. N. Kutulakos and E. Steger. A theory of refractive and specular 3d shape by light-path triangulation. *International Journal of Computer Vision*, 76(1):13–29, 2008.
- [33] D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science*, 194(4262):283–287, 1976.
- [34] N. J. Morris and K. N. Kutulakos. Dynamic refraction stereo. *IEEE transactions on pattern analysis and machine intelligence*, 33(8):1518–1531, 2011.
- [35] S. K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern analysis and machine intelligence*, 16(8):824–831, 1994.
- [36] H. Okawa, Y. Zheng, A. Lam, and I. Sato. Spectral reflectance recovery with interreflection using a hyperspectral image. In *Asian Conference on Computer Vision*, pages 52–67. Springer, 2016.
- [37] O. Özyeşil, V. Voroninski, R. Basri, and A. Singer. A survey of structure from motion*. *Acta Numerica*, 26:305–364, 2017.
- [38] Y. Qian, Y. Zheng, M. Gong, and Y.-H. Yang. Simultaneous 3d reconstruction for water surface and underwater scene. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 754–770, 2018.
- [39] V. S. Ramachandran. Perception of shape from shading. *Nature*, 331(6152):163, 1988.
- [40] E. Rouy and A. Tourin. A viscosity solutions approach to shape-from-shading. *SIAM Journal on Numerical Analysis*, 29(3):867–884, 1992.
- [41] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern recognition*, 37(4):827–849, 2004.
- [42] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1):7–42, 2002.

- [43] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2003.
- [44] Y. Y. Schechner and N. Kiryati. Depth from defocus vs. stereo: How different really are they? *International Journal of Computer Vision*, 39(2):141–162, 2000.
- [45] J. L. Schonberger and J.-M. Frahm. Structure-from-motion revisited. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4104–4113, 2016.
- [46] A. Sedlazeck and R. Koch. Perspective and non-perspective camera models in underwater imaging—overview and error analysis. In *Outdoor and Large-Scale Real-World Scene Analysis*, pages 212–242. Springer, 2012.
- [47] S. W. Seo, S. Han, J. H. Seo, W. B. Choi, and M. Y. Sung. Liquid lens module with wide field-of-view and variable focal length. *Electronic Materials Letters*, 6(4):141–144, 2010.
- [48] M. Shimano, H. Okawa, Y. Asano, R. Bise, K. Nishino, and I. Sato. Wetness and color from a single multispectral image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3967–3975, 2017.
- [49] S. Sorensen, A. Kolagunda, P. Saponaro, and C. Kambhamettu. Refractive stereo ray tracing for reconstructing underwater structures. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 1712–1716. IEEE, 2015.
- [50] D. Z. Stupar, J. S. Bajić, A. V. Joža, B. M. Dakić, M. P. Slankamenac, M. B. Živanov, and E. Cibula. Remote monitoring of water salinity by using side-polished fiber-optic u-shaped sensor. In *2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC)*, pages LS4c–4. IEEE, 2012.
- [51] J. Sun, N.-N. Zheng, and H.-Y. Shum. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (7):787–800, 2003.
- [52] H. Tang, S. Cohen, B. L. Price, S. Schiller, and K. N. Kutulakos. Depth from defocus in the wild. In *CVPR*, pages 4773–4781, 2017.

- [53] J.-P. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. 1998.
- [54] P. Tresadern and I. Reid. Articulated structure from motion by factorization. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 1110–1115. IEEE, 2005.
- [55] P. Trouvé, F. Champagnat, G. Le Besnerais, J. Sabater, T. Avignon, and J. Idier. Passive depth estimation using chromatic aberration and a depth from defocus approach. *Applied optics*, 52(29):7152–7164, 2013.
- [56] S. Ullman. The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 203(1153):405–426, 1979.
- [57] M. Watanabe and S. K. Nayar. Rational filters for passive depth from defocus. *International Journal of Computer Vision*, 27(3):203–225, 1998.
- [58] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds. ‘ structure-from-motion ’ photogrammetry: A low-cost, effective tool for geo-science applications. *Geomorphology*, 179:300–314, 2012.
- [59] W. Wiley and I. H. McLaren. Time-of-flight mass spectrometer with improved resolution. *Review of scientific instruments*, 26(12):1150–1157, 1955.
- [60] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139, 1980.
- [61] A. Yamashita, Y. Shirane, and T. Kaneko. Monocular underwater stereo-3d measurement using difference of appearance depending on optical paths. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3652–3657. IEEE, 2010.
- [62] T. Yau, M. Gong, and Y.-H. Yang. Underwater camera calibration using wavelength triangulation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2499–2506, 2013.
- [63] K. Yoshida and S. Hirose. Laser triangulation range finder available under direct sunlight. *Transactions of the Society of Instrument and Control Engineers*, 24(5):445–451, 1988.

- [64] D. C. Yuen and B. A. MacDonald. Vision-based localization algorithm based on landmark matching, triangulation, reconstruction, and comparison. *IEEE Transactions on robotics*, 21(2):217–226, 2005.
- [65] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*, pages 24–36. IEEE, 2002.
- [66] M. Zhang, X. Lin, M. Gupta, J. Suo, and Q. Dai. Recovering scene geometry under wavy fluid via distortion and defocus analysis. In *European Conference on Computer Vision*, pages 234–250. Springer, 2014.
- [67] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shape-from-shading: a survey. *IEEE transactions on pattern analysis and machine intelligence*, 21(8):690–706, 1999.
- [68] Y. Zheng, I. Sato, and Y. Sato. Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1779–1787, 2015.
- [69] S. Zhuo and T. Sim. Defocus map estimation from a single image. *Pattern Recognition*, 44(9):1852–1858, 2011.

研究業績リスト

博士論文の研究業績

原著論文（査読あり）

1. Shin Ishihara, Antonin Sulc, and Imari Sato, "Depth estimation using spectrally varying defocus blur," J. Opt. Soc. Am. A 38, 1140-1149 (2021)

国際会議（査読あり）

1. S. Ishihara, A. Sulc and I. Sato, "Depth from Spectral Defocus Blur," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 1980-1984.
2. S. Ishihara, Y. Asano, Y. Zheng and I. Sato, "Underwater Scene Recovery Using Wavelength-Dependent Refraction of Light," 2020 International Conference on 3D Vision (3DV), 2020, pp. 32-40.

博士論文以外の研究業績

国際会議（査読あり）

1. Mihoko Shimano, Yuta Asano, Shin Ishihara, Ryoma Bise, Imari Sato: Imaging Scattering Characteristics of Tissue in Transmitted Microscopy. International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 236 - 245 2020.