

多視点マルチバンド画像からの3次元復元と色再現に関する研究

著者	伊藤 秀哉
学位授与機関	Tohoku University
URL	http://hdl.handle.net/10097/00121251

修士論文

多視点マルチバンド画像からの
3次元復元と色再現に関する研究

情報基礎科学専攻

伊藤 秀哉

目次

第 1 章	緒言	3
第 2 章	多視点 3 次元復元とマルチバンド画像を用いた色再現に関する基礎的考察	6
2.1	まえがき	6
2.2	多視点 3 次元復元の概要	7
2.3	多視点ステレオ手法の分類と特徴	10
2.4	マルチバンド画像を用いた色再現の概要	13
2.5	むすび	19
第 3 章	多視点マルチバンド画像を用いた 3 次元復元	20
3.1	まえがき	20
3.2	位相限定相関法	20
3.3	多視点マルチバンド画像からの 3 次元復元	25
3.4	むすび	34
第 4 章	多視点マルチバンド画像を用いた色再現	35
4.1	まえがき	35
4.2	色収差の補正	35
4.3	適切な画像の選択	39
4.4	分光反射率の推定	40
4.5	むすび	41
第 5 章	性能評価実験	42
5.1	まえがき	42
5.2	データセット作成	42
5.3	評価実験	43
5.4	むすび	45

第 6 章 結言	48
参考文献	49
謝辞	55

第 1 章

緒言

近年，文化資源の保護・保全やインターネットを用いた資料の配布などを目的として，文化財や美術作品などをデジタルデータで保存するデジタルアーカイビングが博物館，美術館，研究機関などで進められている [1, 2, 3]．アーカイビングの対象となる文化財は，建築物や人形などの有形文化財と伝統芸や工芸技術などの無形文化財の 2 つに分けられる．伝統芸や工芸技術は，モーションキャプチャを用いて，演者あるいは職人の動作が記録される [1, 4]．建築物や人形は，カメラやレーザスキャナなどを用いて 3 次元形状や色が記録される．本論文では，文化財の形状と色の記録に着目し，多視点マルチバンド画像を用いた物体の形状および色を計測する手法を提案する．

文化財の 3 次元形状を高精度に計測するために，レーザスキャナに代表される能動型 3 次元計測が用いられている．能動型 3 次元計測は，レーザや構造光を用いた計測手法であり，精度と安定性が優れていることから，3 次元計測の主流となっている．一方で，プロジェクタやレーザなどの投影装置が必要となるため，計測機器が大型かつ高価になりがちで，撮影環境や利用者が限定される．レーザスキャナなどが利用できない場合に，カメラを用いた受動型 3 次元計測が用いられている．受動型の 3 次元計測の中でも，特に，異なる視点から撮影された複数枚の画像を用いて物体の 3 次元形状を復元する技術である多視点ステレオ (Multi-View Stereo: MVS) [5, 6, 7, 8] が用いられている．MVS では，画像間の対応関係とカメラパラメータを用いて，三角測量の原理から対象物体の 3 次元形状を復元する．能動型の 3 次元復元に比べて精度と安定性が低いですが，復元対象や撮影環境に合わせて，レンズやカメラなどの撮影装置や撮影方法を変更することが可能であり，復元対象や撮影環境に対する制約が少ない．一方で，画像間の対応付けの精度が 3 次元復元の精度を決める重要な要素になる．

文化財の色を高精度に計測するために，任意の照明環境下での物体の色を再現する技術であるマルチバンド画像を用いた色再現 [9, 3, 10, 11] が用いられている．マルチバンド画像は，回折格子やカラーフィルタなどを用いて，カメラが観測する光を複数の波長帯域

に分割して撮影された画像である。通常のカメラで撮影される RGB 画像は、被写体表面で反射された可視光波長域の光を赤・緑・青の 3 領域に分割して色信号として記録しているため、被写体の色が撮影時の照明環境に依存する。計測時の環境に依存せずに物体の色を計測するためには、計測対象全体の分光反射率を計測する必要がある。分光反射率の計測には、一般に、分光光度計や分光放射輝度計が利用される。これらの機器は、計測対象に単色光を照射して感光センサで観測する、あるいは、物体が反射した光を単色光に分解して感光センサで観測する。物体のある 1 点のみの分光反射率しか得ることができないため、対象全体の分光反射率を得るためには、膨大な時間と量力が必要であり、現実的ではない。そこで、対象全体の分光反射率を計測するために、マルチバンド画像を用いた手法が提案されている [9, 3, 10, 11]。カメラ感度、撮影時の照明光スペクトル、マルチバンド画像から各ピクセルの分光反射率を推定することができる。また、推定した分光反射率と任意の照明光スペクトルを用いることで、任意の照明環境下での物体の色を再現することが可能である。

これまで述べてきたように、計測対象の形状は多視点画像を用いて復元することができ、計測対象の色はマルチバンド画像を用いて再現することができる。多視点画像もマルチバンド画像もカメラで撮影できるため、複数視点からマルチバンド画像を撮影することで、計測対象の形状と色を同時に求めることが可能である。そこで、本論文では、多視点マルチバンド画像を用いた 3 次元復元および色再現手法を提案する。3 次元形状の復元には、位相限定相関法 (Phase-Only Correlation: POC) を用いた 3 次元復元手法 [12] を用いる。POC は、画像を離散フーリエ変換 (Discrete Fourier Transform: DFT) したときに得られる位相情報に着目した画像マッチング手法 [13, 14] である。マルチバンド画像を用いることで POC のマッチング精度が向上するため、以下の処理で使用する画像対応付けでは、マルチバンド画像を用いた POC [15] を使用する。まず、Structure from Motion (SfM) [5, 16] を用いてカメラ運動を推定するとともに、物体の疎な 3 次元形状を復元する。次に、疎な 3 次元点群から三角形メッシュを生成する。そして、各メッシュに対して画像対応付けに適切なステレオカメラペアを選択し、密な画像対応付けにより物体の 3 次元形状を復元する。色の再現には、復元された 3 次元形状を利用することで、物体の局所的な形状を考慮した色再現手法を用いる。メッシュごとに色再現に適したマルチバンド画像を選択することで、計測対象全体の色を正確に再現することが可能となる。また、高精度な色再現のために、マルチバンド画像のためのキャリブレーション手法を提案する。光の波長ごとにレンズの屈折率が異なるため、各波長帯域ごとに、像の結像位置がことなり、色収差と呼ばれる色ズレが生じる。カメラのパラメータ推定に利用するチェッカーボードの画像を用いてマルチバンド画像に生じる色収差を補正する。提案手法の有効性を定量的に評価するために、被写体の全周囲のマルチバンド画像を撮影するシステムを構築し、精度評価用のデータセットを作成する。作成したデータセットを用いた精度評価

実験を通して、提案手法が高精度に被写体の形状および色を計測できることを示す。

本論文は、以上の内容をまとめたものであり、多視点マルチバンド画像を用いた 3 次元復元および色再現手法を提案する。以下に、本論文の構成を示す。

第 1 章は、本論文の背景と目的、およびその概要を述べたものである。

第 2 章においては、多視点 3 次元復元とマルチバンド画像を用いた色再現に関する基礎的考察を行う。まず、多視点画像からの 3 次元復元における 3 つのステップについて述べる。次に、詳細な 3 次元形状の復元に用いられる MVS 手法の分類とそれぞれの特徴について述べる。また、マルチバンド画像を用いた色再現の概要について述べる。次に、マルチバンド画像を用いた色再現における 4 つのステップについて述べる。

第 3 章においては、多視点マルチバンド画像を用いた 3 次元復元について述べる。まず、多視点画像間のマッチングに用いる位相限定相関法について述べる。次に、最適な画像選択に基づくマルチバンド画像からの 3 次元モデル生成の流れと、その具体的な処理について述べる。

第 4 章においては、多視点マルチバンド画像を用いた色再現について述べる。まず、マルチバンド画像における色収差の補正法について述べる。次に、色再現を行う際の適切な画像選択について述べ、マルチバンド画像を用いた分光反射率の推定法について述べる。

第 5 章においては、マルチバンド画像を用いて実験を行い、提案手法の有効性を実証する。まず、性能評価実験に用いるマルチバンド画像のデータセットの作成法について述べる。次に、実際に撮影したマルチバンド画像を用いた性能評価実験を行い、提案手法の有効性を実証する。

第 6 章は結言である。

以上、本論文の企図するところを概説した。

第 2 章

多視点 3 次元復元とマルチバンド画像を用いた色再現に関する基礎的考察

2.1 まえがき

多視点 3 次元復元は、異なる視点から撮影された複数枚の画像を用いて物体の 3 次元形状や表面の色彩情報を計算機上で生成する技術である [5, 6, 7, 8]。多視点 3 次元復元は、レーザや構造光の投影が必要な能動型の 3 次元復元手法に比べ、小型かつ安価なシステムで実現可能である。また、撮影機材がカメラのみであり、撮影における制限が少ないといった利点がある。このような背景から、多視点画像からの 3 次元復元は、産業・医療・芸術などのさまざまな分野において、能動型の 3 次元復元生成手法に代わる手法として実用化が進んでいる。

マルチバンド画像を用いた色再現は、複数の異なる波長帯域で撮影した画像を用いて物体の分光反射率を推定し、任意の照明環境下での物体の色を再現する技術である [10, 9, 11, 3]。対象を撮影した画像の各ピクセルにおける分光反射率を推定可能であり、一度の撮影で広範囲の分光反射率を推定可能である。また、物体の分光反射率情報を得ることができるため、撮影時の照明環境に依存しない物体の色情報を取得可能である。

本章では、多視点 3 次元復元に関する基礎的考察を行う。まず、多視点 3 次元復元における 3 つのステップについて述べる。さらに、密な 3 次元形状の復元に用いられる多視点ステレオ (Multi-View Stereo: MVS) 手法の分類とそれぞれの特徴について述べる。次に、マルチバンド画像を用いた色再現に関する基礎的考察を行う。まず、マルチバンド画像を用いた色再現の概要について述べる。そして、マルチバンド画像を用いた分光反射率の推定手法について述べる。さらに、マルチバンド画像の撮影法について述べる。

2.2 多視点3次元復元の概要

多視点画像からの3次元復元は、一般に、多視点画像の取得、カメラパラメータ推定、密な3次元形状の復元の3つのステップから構成される [8]。本節では、各ステップの概要を述べる。

2.2.1 多視点画像の取得

多視点画像の撮影には、固定された多視点カメラを用いる方法、ロボットアームや回転テーブルを用いる方法、単眼カメラを用いた自由移動による撮影などが用いられる [8]。それぞれの手法について以下にまとめる。

固定された多視点カメラを用いる方法

固定された多視点カメラを用いる場合には、対象の物体を中心に複数のカメラを固定して設置し、多視点画像を撮影する。この方法は、カメラを固定することで、事前にカメラキャリブレーションを行うことが可能であり、カメラの内部パラメータおよび外部パラメータを高精度かつ安定に求められる。また、カメラを同期させて撮影することで、動物体の3次元復元が可能である。一方で、物体を囲むようにカメラを設置するため、対象は比較的小さな物体に限られ、建築物などの大きな物体を撮影することは困難である。また、多くのカメラを必要とするため、撮影機器にかかる費用が多くなるといった問題がある。

ロボットアームや回転テーブルを用いる方法

ロボットアームや回転テーブルを用いる方法では、カメラはロボットアームの先端に固定され、回転テーブルおよびロボットアームを動かすことにより、多視点画像を撮影する [8, 6]。また、ロボットアームを用いることが困難な場合、カメラは三脚などに固定される。この方法は、多視点カメラを用いる場合と同様に、事前にカメラキャリブレーションを行うことができるが、ロボットアームや回転テーブルの大きさから、撮影対象は比較的小さな物体に限られる。一方で、この手法では、多視点画像を同時に取得することができないため、動物体の撮影は不可能である。

単眼カメラによる自由な移動撮影

撮影の制限が少ない撮影方法として、単眼カメラによる自由な移動撮影がある [8, 7, 17, 18]。カメラの移動撮影では、撮影者が自由に移動しながら同一の対象を撮影することで、多視点画像を取得する。また、近年では、1人の撮影者が撮影した画像ではなく、インターネットから同一の対象を撮影した画像を収集することで、多視点画像を取得し、3次元復元を行う方法も提案されている [19, 20, 21]。この方法では、多視点カメラやロボッ

トアームなどを用いる手法に比べて、撮影の制限が少なく、テーブルの上に乗るくらいの小さな物体から、建物全体といった大きな物体、さらには、都市のような大規模な 3 次元復元まで生成可能である。一方で、カメラの移動撮影では、事前にキャリブレーションを行うことが困難であり、撮影された多視点画像から直接カメラパラメータを推定する必要がある。

現在、多視点画像からの 3 次元復元において、その撮影の自由度から、カメラの移動撮影やインターネット上の画像収集による多視点画像の取得が特に注目されている。この背景として、SIFT (Scale-Invariant Feature Transform)[22] や SURF (Speeded-Up Robust Features)[23] などの画像の回転、拡大縮小にロバストな特徴ベースの画像マッチングにより、安定して多視点画像間の画像マッチングが可能になったことや、バンドル調整 [24, 25, 26, 5] により、多視点画像間の画像対応付けの結果を用いてパラメータ最適化し、高精度なカメラパラメータ推定が可能になったことがあげられる。

2.2.2 カメラパラメータ推定

カメラパラメータ推定は、各視点の画像におけるカメラの内部パラメータおよび外部パラメータを求める [5, 27, 16]。内部パラメータは、焦点距離や画像中心、撮像素子のサイズと画像サイズの関係などの情報を含み、カメラ座標系の 3 次元座標とデジタル画像座標系の 2 次元座標の関係を表す。外部パラメータは、カメラの回転と並進移動の情報を含み、世界座標系の 3 次元座標におけるカメラの位置および姿勢を表す。

これらのパラメータは、画像中の特徴点が他の画像中のどの点に対応するかという対応点情報から推定できる。いま、図 2.1 に示すように、三角錐の頂点 X_1, X_2 が 3 枚の画像 I_1, I_2, I_3 上で特徴点として検出されたとする。このとき、カメラパラメータが正しければ、対応する頂点の視線ベクトル (図 2.1 中の赤、オレンジのベクトル) はそれぞれ一点で交わるはずである。したがって、カメラパラメータ推定は、画像中で特徴点を抽出し、画像間で対応を取ったのち、すべての対応する視線ベクトルが、3 次元上の 1 点で交わるようなカメラパラメータを求めるという問題に帰着する。カメラパラメータ推定の出力としては、カメラパラメータの他に、各視線ベクトルが交わる 3 次元点 X_1, X_2 の座標も得られる。推定手法によって大きく異なるのは、画像間の対応点情報をいかに計算するかという点である。以下に 2 種類のカメラパラメータ推定手法をまとめる。

既知形状物体の撮影によるカメラキャリブレーション

既知形状物体の撮影によるカメラキャリブレーションは、形状が既知である平板や直方体を複数回撮影し、世界座標系における平板や直方体の座標とデジタル画像座標系における 2 次元座標を対応付けることによりカメラパラメータを推定する [5, 16]。このとき、平板や直方体の表面のテクスチャとして格子パターンやチェッカーパターン、整列された

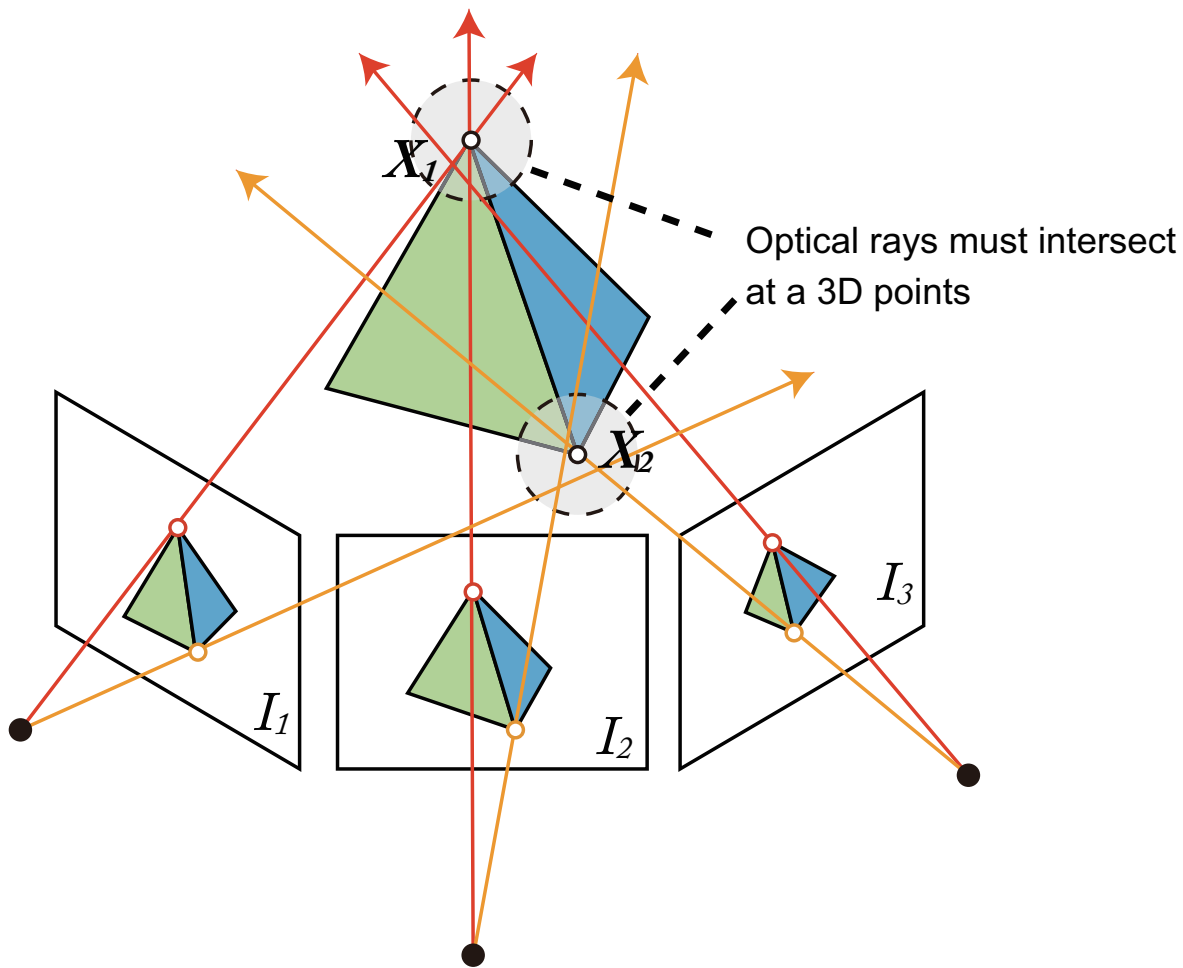


図 2.1 カメラパラメータ推定の原理 [8]：正しいカメラパラメータがならば，対応する頂点の視線ベクトルが 1 点で交わる

円などを用いることで，コーナー検出や円検出により，平板や直方体上の 3 次元座標とカメラ画像の画像対応付けを行う．この手法は，後述する Structure from Motion に比べて，安定して高精度なカメラパラメータ推定が可能である．一方で，カメラパラメータは，既知形状物体を撮影した状態で最適化されるため，カメラキャリブレーション時における多視点カメラの設定と，モデル生成対象の撮影時におけるカメラの設定を変更することができない．そのため，カメラの移動撮影のように多視点画像を同一の条件で複数回撮影できない場合には，この手法を適用することが困難である．

Structure from Motion (SfM)

Structure from Motion では，モデル生成対象を撮影した多視点画像間の対応付け結果から直接カメラパラメータを推定する [5, 16]．多視点画像間の画像対応付けには，ステレオ画像間の画像変形にロバストな手法が必要であるため，SIFT や SURF などの特徴ベースの手法 [22, 23] が用いられる．また，カメラパラメータの推定には，特異値分解に

基づく手法や、Direct Linear Transformation (DLT) 法、バンドル調整などが用いられる [5, 16, 24, 25]。近年では、5点アルゴリズム [28] や P3P 問題の解法 [29] により、カメラパラメータを安定かつ高精度に求めて、それを初期値としてバンドル調整を実行することで、パラメータを最適化する方法が良く用いられる。Structure from Motion では、3次元モデル生成対象を撮影した多視点画像からカメラパラメータを直接求めるため、カメラの移動撮影のように、事前のカメラキャリブレーションが難しい場合においても、カメラパラメータ推定を行うことが可能である。一方で、多視点画像間の画像対応付けにおいて誤対応が含まれると、カメラパラメータの推定精度が著しく低下するといった問題がある。そのため、RANSAC (RANdom SAmple Consensus) [30] などを用いて誤対応を正確に除去することが重要となる。また、Structure from Motion によるカメラパラメータ推定では、推定結果にスケールの不定性が原理的に含まれ、生成された3次元モデルの絶対的な大きさを一意に決定することができない。

2.2.3 密な3次元形状の復元

Structure from Motion では、カメラパラメータ推定と同時に3次元復元を行う。しかし、Structure from Motion で得られる3次元点は、特徴点のみに対応するため、点数が少なくノイズを多く含む。したがって、多視点画像からの3次元モデル生成の最後のステップとして、密でノイズの少ない3次元復元を行う必要がある。これには、多視点ステレオ (Multi-View Stereo: MVS) と呼ばれる手法が用いられる。

多視点ステレオでは、一般に photometric consistency, もしくは photo-consistency と呼ばれる関数を定義し、その関数の値によって物体の表面を探索する問題と定式化される。いま、図 2.2 に示すように、ある3次元点 X_i が与えられたとき、 X_i を複数の画像に投影することを考える。もし、 X_i が物体表面に存在するのであれば、投影されるピクセルの周囲のテクスチャーには一貫性があるはずである (図 2.2 の X_1)。一方、そうでない場合には、投影されるピクセルの周囲のテクスチャーは異なっているはずである (図 2.2 の X_2)。このように、 X_i が物体表面に存在するとき最大 (もしくは最小) となることが、photo-consistency に期待される (図 2.2 の右図)。

2.3 多視点ステレオ手法の分類と特徴

多視点ステレオの手法は、3次元ボリュームの最適化に基づく手法、メッシュの最適化に基づく手法、デプスマップ統合に基づく手法、特徴領域拡張に基づく手法の4つに大きく分類される [6]。それぞれの手法について、以下にまとめる。

3次元ボリュームの最適化に基づく手法

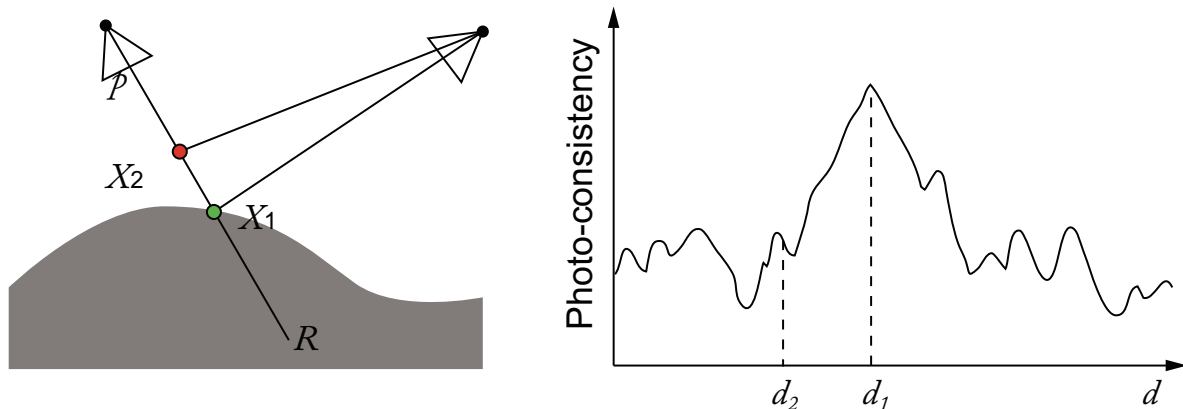


図 2.2 Photo-consistency を用いた物体表面の探索 [8]: 3次元点 X_i が物体表面にあるとき, 各カメラで観測される投影点の周辺テクスチャには一貫性があり, Photo-consistency の値は最大 (または最小) となる.

3次元ボリュームの最適化に基づく手法では, 多視点画像と3次元ボリュームからエネルギー関数を定義し, そのエネルギー関数を最小化することで3次元復元を行う [31, 32, 33]. 一般に, 自由な3次元ボリュームの表現を定義することは難しく, また, 複雑な3次元ボリュームの表現はエネルギー関数を複雑にする. そこで, 多くの場合, 3次元ボリュームをボクセルによって表現し, グラフカットなどの最適化により, 各ボクセルが3次元復元対象の内部と外部のどちらにあるかを求める.

一般に, 3次元ボリュームの最適化に基づく手法では, 特徴領域拡張に基づく手法やデプスマップ統合に基づく手法に比べ, 3次元復元精度が高いといった特徴がある. 一方で, 複雑な最適化を行うため, 初期値によって問題が収束しなかったり, 最適化の際の処理時間が膨大になったりするといった問題がある. また, 3次元復元精度と処理時間はボクセルの分解能に依存するため, 高精度な3次元復元を行うためにはボクセルを細かく設定する必要があり, 処理時間が増加する. これらの特徴から, 3次元ボリュームの最適化に基づく手法は, 視体積交差法 [34] により得られた3次元ボリュームや, 特徴領域拡張に基づく手法またはデプスマップ統合に基づく手法により得られた3次元復元結果を初期値として用いることが多く, 他の手法の3次元復元結果の高精度化に適しているといえる.

メッシュの最適化に基づく手法

メッシュの最適化に基づく手法では, 多視点画像と3次元メッシュモデルからエネルギー関数を定義し, そのエネルギー関数を最小化することで3次元復元を行う [35, 36, 37]. 一般に, 自由な3次元メッシュモデルの表現を定義することは難しく, また, 複雑な3次元メッシュモデルの表現は最小化するエネルギー関数を複雑にする. そこで, 多くの場合, 3次元メッシュモデルを多角形メッシュによって表現し, 多角形の頂点を移動させた

り，多角形を分割したりすることで最適化を行う。

一般に，メッシュの最適化に基づく手法では，特徴領域拡張に基づく手法やデプスマップ統合に基づく手法に比べ，3 次元復元精度が高いといった特徴がある。一方で，3 次元ボリュームの最適化に基づく手法と同様に，初期値によって問題が収束しなかったり，最適化の際の処理時間が膨大になったりするという問題がある。また，局所的な最適解に収束しやすいため，初期値のメッシュモデルを高精度に求める必要がある。これらの特徴から，メッシュの最適化に基づく手法は，視体積交差法により得られた 3 次元ボリュームや，特徴領域拡張に基づく手法またはデプスマップ統合に基づく手法により得られた 3 次元復元結果を初期値として用いることが多く，他の手法の 3 次元復元結果の高精度化に適しているといえる。

デプスマップ統合に基づく手法

デプスマップ統合に基づく手法では，(a) 多視点画像から複数のデプスマップを生成し，(b) 複数のデプスマップを統合することで 3 次元復元を行う [38, 39, 40, 41, 42, 43]。一般に，デプスマップの生成には，ウィンドウマッチングとプレーンスイーピングに基づく手法や，エネルギー関数の最小化に基づく手法がある。また，デプスマップの統合には，能動型 3 次元復元手法の復元結果の統合に用いられるボリュームベースの手法 [44] や，単純な座標統合により得られる 3 次元点群からの 3 次元メッシュモデル生成手法 [45] などが適用される。

デプスマップ統合に基づく 3 次元復元におけるデプスマップ生成手法として，プレーンスイーピングに基づくデプスマップ生成手法が良く知られている [5]。プレーンスイーピングに基づくデプスマップ生成では，ある参照視点における視線上で 3 次元点の奥行きを変化させながら，その 3 次元点を近傍の視点に投影し，投影された座標を中心として参照視点と近傍視点間のウィンドウマッチングを行うことで，その 3 次元点の photo-consistency を計算する。そして，奥行きと photo-consistency の関係から参照視点における奥行きを決定することで，デプスマップを生成する。多くの場合，参照視点と近傍視点間のウィンドウマッチングには NCC が用いられる。

プレーンスイーピングに基づくデプスマップ生成は，複雑な最適化を必要としない。そのため，3 次元ボリュームの最適化に基づく手法やメッシュの最適化に基づく手法と比べて，局所的な最適解に問題が収束するといった問題がない。また，視体積交差法の 3 次元復元結果などの初期値が必須でない，処理の並列化による高速化が可能である，といった利点がある。一方で，画像間の幾何的な変形を考慮せずにウィンドウマッチングを行うため，カメラの基線長が長い場合，正確なデプスマップの推定が困難である。また，奥行きの探索範囲を設定する必要があるが，一般に屋外環境などでは探索範囲を決定できない。さらに，プレーンスイーピングに基づくデプスマップ生成は，3 次元点の奥行き変化の刻み幅が，3 次元復元精度と処理時間に大きく影響を与える。そのため，高精度な 3 次元復

元を行うためには、奥行き変化の刻み幅を非常に小さくする必要があり、マッチング回数が膨大になるという問題がある。

特徴領域拡張に基づく手法

特徴領域拡張に基づく手法では、(a) 特徴点の3次元復元を行い、(b) 復元結果を周囲の点に繰り返し伝搬することで3次元復元を行う [46, 47, 48]。特徴点の3次元復元には、SIFT や SURF のような特徴ベースの画像対応付けや、NCC などの領域ベースのウィンドウマッチングを用いたパッチベースの画像対応付けが用いられる。

一般に、特徴領域拡張に基づく手法は、3次元ボリュームの最適化に基づく手法やメッシュの最適化に基づく手法に比べて、処理時間が短く、また、視体積交差法の3次元復元結果などの初期値が必須でないといった利点がある。一方で、3次元復元を特徴点から伝搬させるため、3次元復元対象によっては、特徴点の数が少なく物体全体が復元できなかったり、特徴点の少ない領域で局所的に3次元復元精度が低下するといった問題がある。また、特徴点の3次元復元結果には誤対応点が含まれるため、安定した3次元復元を行うためには、誤対応除去を正確に行うことが重要である。

特徴領域拡張に基づく手法の中でも、Furukawa らのパッチベースの3次元復元手法 [48] は良く知られている。Furukawa らの手法では、DoG (Difference-of-Gaussian) と Harris のコーナー検出により特徴点を検出し、エピポーラ線上で NCC を用いたパッチベースのマッチングを行うことで、特徴点の3次元復元を行う。その後、誤対応除去と3次元復元結果の伝搬を繰り返し行うことで、密な3次元形状を復元する。パッチベースの画像対応付けは、物体表面の法線情報を利用し、マッチングウィンドウを設定する。このため、カメラ間の基線長が比較的長い場合でも、画像間の幾何的な変形にロバストなマッチングが可能である。一方で、この画像対応付けは参照画像の近傍のすべての画像に対して行われ、カメラの基線長やオクルージョンが考慮されない。このため、入力画像によっては、3次元復元精度が低下したり、復元結果の一部が欠損するといった問題がある。

2.4 マルチバンド画像を用いた色再現の概要

人間が観測する物体の色は、物体の分光反射率、照明光スペクトル、それを観察する人間の目の分光感度（等色関数）の3つの特性によって決まる（図 2.3,(a)）。それに対して通常の RGB カメラで記録される色は、物体の分光反射率、照明光スペクトル、カメラの感度の3つの特性によって決まる（図 2.3,(b)）。人間の目の分光感度とカメラの感度の違いが、実際に目で見た色と画像で見た色に違いが生じる主な原因となる。また、画像撮影時と異なる照明環境で物体を見た場合も、それぞれの照明光スペクトルが異なるため、実際に見た色と画像で見た色に違いが生じる。デジタルアーカイビングにおいては、この色の違いが大きな問題となる。それに対してマルチバンド画像を用いた色再現

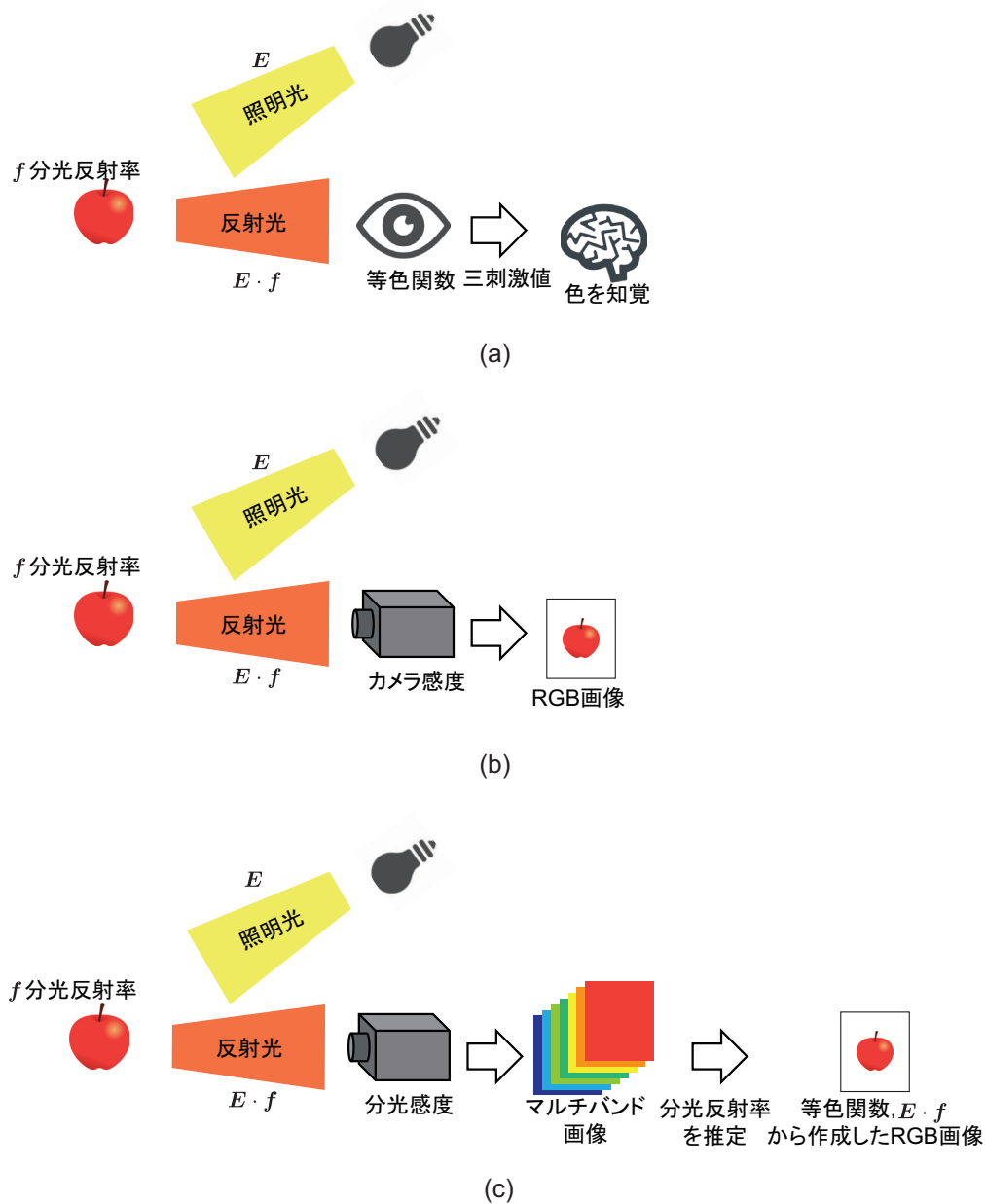


図 2.3 観測する色の違い：(a) 人間が観測する色，(b) RGB カメラが観測する色，(c) マルチバンド画像から再現される色

(図 2.3,(c)) では，予め計測したカメラ感度，照明光スペクトル，マルチバンド画像から物体の分光反射率を推定し，照明光スペクトル，等色関数，ディスプレイ感度などを用いて実際に人が知覚する色と同等の色を再現する．マルチバンド画像からの色再現は大きく分けて，マルチバンド画像の撮影，分光反射率の推定，三刺激値への変換，表示画像信号の算出の 4 ステップに分けられる．本節では，各ステップの概要を述べる．

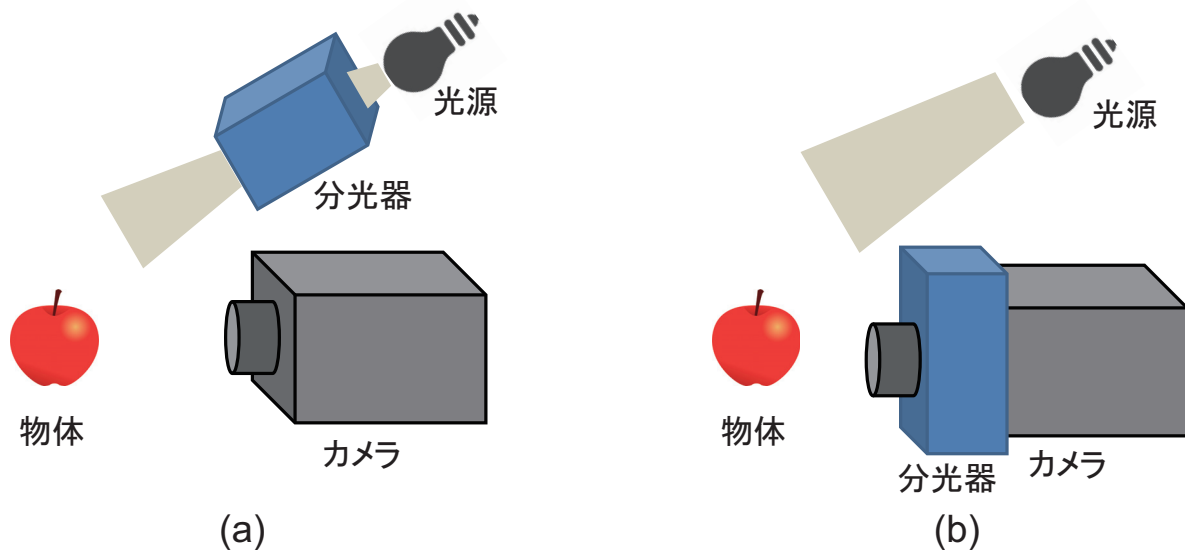


図 2.4 マルチバンド画像の撮影における分光光学系の違い：(a) 光源側での分光 (b) 結像系での分光 [49]

2.4.1 マルチバンド画像の撮影

マルチバンド画像の撮影方式は図 2.4 に示すように、分光器などの分光光学系の挿入位置によって、2 つの撮影法に分けられる [49]。照明光スペクトルを変化させる場合は図 2.4(a)、分光感度を変化させる場合は図 2.4(b) の撮影方式となる。分光器には、回折格子や、干渉フィルタ、カラーフィルターなどが用いられる。以下に、それぞれの詳細を述べる。

回折格子を用いる場合

分光器に回折格子を用いる場合は、図 2.5 に示すような画像入力方法となる。物体の像をスリット上に結像させ、このスリットを透過した光が凹面鏡により平行光となり、回折格子に入射する。回折格子が垂直方向に光を分散させ、2 つ目の凹面鏡にて分散した光が波長毎に 2 次元センサ上の異なる高さに集まる。原理的に高い波長分解能を有するが、この方法によってセンサから得られる情報は、スリットを透過した線状の領域のみとなるため、2 次元のマルチバンド画像を取得するためには物体を垂直方向に走査する必要がある。

干渉フィルタを用いる場合

干渉フィルタはガラス上に光学薄膜が多重にコーティングされたものである。境界面での多重反射により、狭帯域の分光透過特性を持つため、高い波長分解能を有する。マルチ

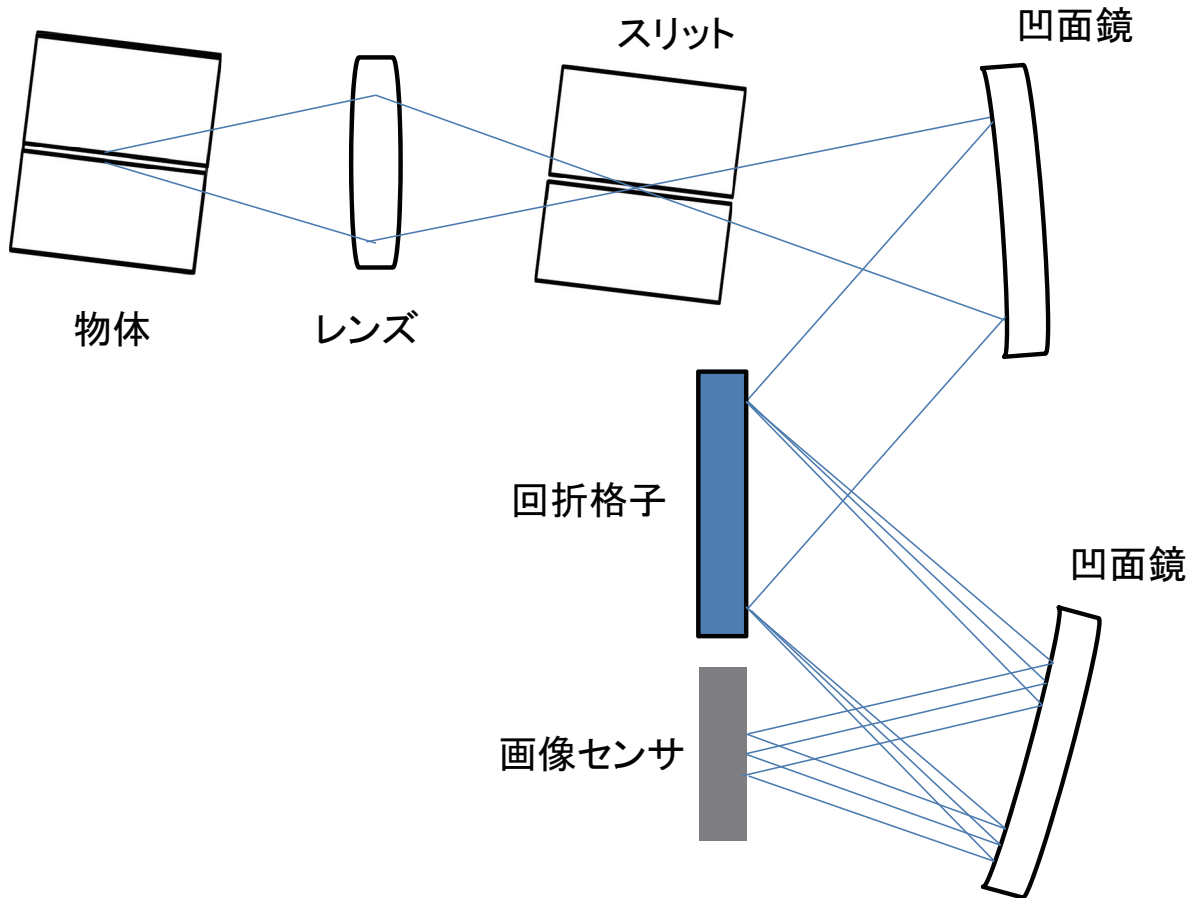


図 2.5 回折格子を用いたマルチバンド画像撮影装置 [49]

バンド画像を取得するためには、異なる特性を持つ干渉フィルタを交換しつつ撮影する必要がある。そのため、図 2.6 に示すような回転するフィルタホイールに多数の干渉フィルタを装着する方式となる。

カラーフィルタを用いる場合

カラーフィルタは特定の波長域の光を透過し、特定の波長域の光を吸収するという広帯域の分光透過特性を持つため、高い波長分解能を得ることは困難である。しかし、対象物の分光反射率の統計的な性質を利用したウィナー推定などを用いることで、6バンドほどのマルチバンド画像から高精度に分光反射率を推定することが可能である [11]。干渉フィルタのように、フィルタホイールを用いる方式や、CCD カメラのセンサ部分に多色のカラーフィルタを設置して撮影する方式 [50] がある。通常カラー画像は、CCD カメラのセンサの前にモザイク状の3色のカラーフィルタ図 2.7 を設置して撮影した画像である。また、干渉フィルタを組み合わせることで波長分解能を上げることが可能であり、Tsuchida らは、通常のデジタルカメラ2台と干渉フィルタを用いることで、6バンドの



図 2.6 干渉フィルタを用いたマルチバンド画像撮影装置 [49]

マルチバンド画像を撮影した [3] .

2.4.2 分光反射率の推定

マルチバンド画像を用いた分光反射率の推定手法には、主に主成分分析を用いた手法 [51] や、最小二乗誤差規範に基づく手法（ウィナー推定） [52, 53, 3] などがある．以下にそれぞれの特長を述べる．

主成分分析を用いた手法

分光反射率は比較的滑らかな特性を持っているため、少ない本数の基底の線形和で表現できる．主成分分析を用いた手法では、分光反射率のサンプルデータに対して主成分分析を行う．マルチバンド画像のバンド数が基底のベクトルの数以上であれば、分光反射率の基底ベクトルに主成分を用いることで、最小二乗法により分光反射率を推定することができる．分光反射率のサンプルデータが必要となるため、サンプルデータと大きく異なる分光反射率を持つ物体の計測には不向きであり、撮影対象が限定される．

最小二乗誤差規範に基づく手法（ウィナー推定）

最小二乗誤差規範に基づく手法は、推定対象のサンプル集合を用いて、真の分光反射率と推定分光反射率との二乗誤差の集合平均が最小になるように、分光反射率の相関行列を求める手法である．例えば、絵画を対象としたとき、絵の具のサンプルの分光反射率を計測することで、絵の具のサンプルのマルチバンド画像から、分光反射率の相関行列を推定

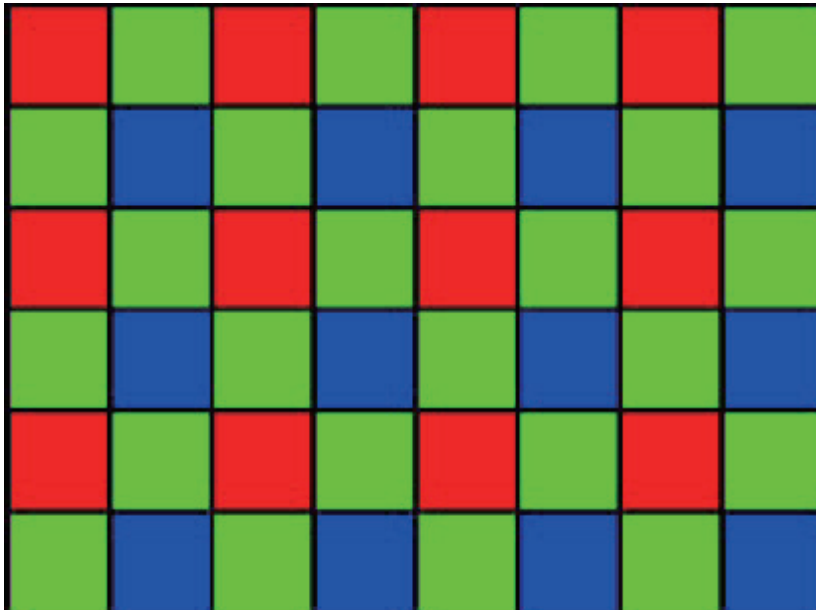


図 2.7 CCD カメラのカラーフィルタの配置

することができる．このとき推定した相関行列を用いて，実際の絵画のマルチバンド画像から絵画の分光反射率を推定することができる．一般的に，推定対象のサンプルを多数集めることは困難であるため，分光反射率を1次マルコフモデルと仮定し，分光反射率の相関行列を定義することで，安定して分光反射率の推定ができる．6バンドほどの少ないバンド数でも高精度に分光反射率が推定出来ることが確認されている [11]

2.4.3 三刺激値の算出

推定した分光反射率，色再現したい照明光スペクトル，等色関数（図 2.8）を用いて， RGB に対応する三刺激値 XYZ を算出する．等色関数は以下に説明する等色実験から求めた関数で，人の目の分光感度に相当する．色光の3原色 RGB としてそれぞれ $700.0[\text{nm}]$ ， $546.1[\text{nm}]$ ， $435.8[\text{nm}]$ の3種の波長の単色光を採用し，これらの三原色を加法混色した結果と，様々な波長の単色光の色とを比較し，同じ色に見える混色比を求める．一般的には，CIE（国際照明委員会）などで採用されている等色関数 (<http://cvrl.ioo.ucl.ac.uk/cmfs.htm>) が用いられる．

2.4.4 表示画像信号の算出

算出した三刺激値から，モニター感度やトーンカーブを用いて，表示信号を算出する．モニター感度は，表示したいモニターに実際に $RGB = (255, 0, 0), (0, 255, 0), (0, 0, 255)$

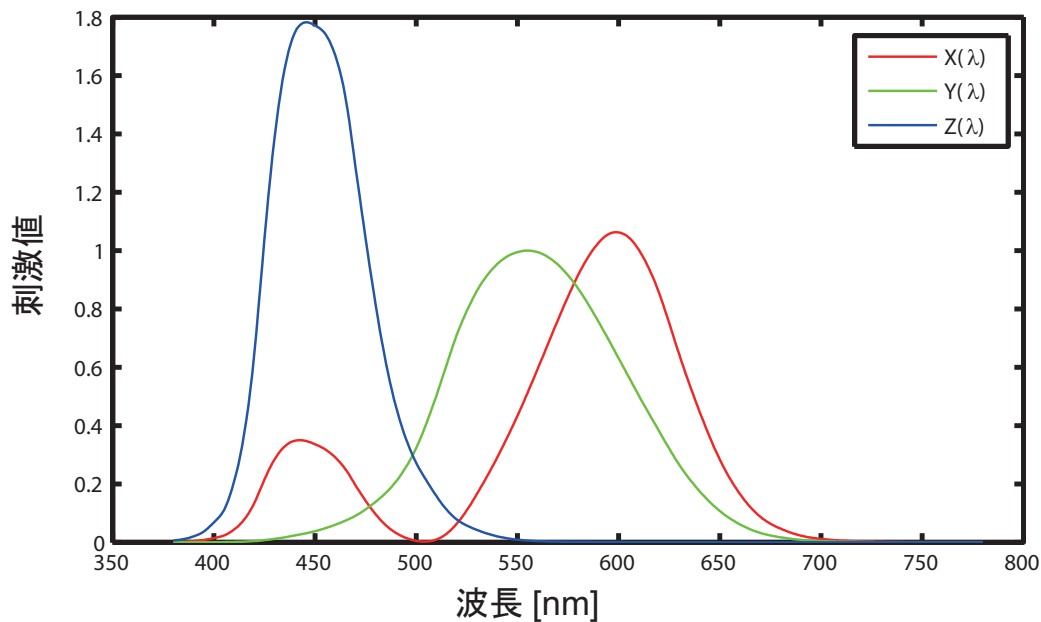


図 2.8 CIE で採用されている等色関数

などの原色を表示し，分光放射計などでモニターが出す光のスペクトルを計測することで得られる．トーンカーブについては，モニターに $RGB = (0, 0, 0), (1, 0, 0), (2, 0, 0) \dots (255, 0, 0)$ などを表示し， RGB それぞれについて， RGB の入力に対してモニターが出す光の入出力特性を計測することで得られる．

2.5 むすび

本章ではまず，多視点3次元復元に関する基礎的考察を行った．始めに，多視点画像からの3次元復元における3つのステップについて述べた．さらに，密な3次元形状の復元に用いられる多視点ステレオ (Multi-View Stereo: MVS) 手法の分類とそれぞれの特徴について述べた．次に，マルチバンド画像を用いた色再現に関する基礎的考察を行った．まず，マルチバンド画像を用いた色再現の概要について述べた．次に，マルチバンド画像を用いた色再現4つのステップについて述べた．

第 3 章

多視点マルチバンド画像を用いた 3 次元復元

3.1 まえがき

従来の多視点 3 次元復元の多くは, NCC に基づく画像対応付けを行い, 画像対応付けのスコアが高い画像ペアの結果のみを 3 次元復元に利用することで, 画像変形の影響を低減し, 3 次元復元の精度を高めている [54, 55, 56, 38, 48]. これらの手法は, 画像ペアの基線長やオクルージョンを考慮せずに, 基準となる画像との画像対応付けをすべての近傍画像に対して行うため, 3 次元復元精度が低下したり, 3 次元モデルの一部に欠損が生じたりすることがある.

これに対して, 山尾らは, 復元した物体の疎な 3 次元形状をもとに, 基線長やオクルージョンを考慮した画像ペアを選択する 3 次元復元手法を提案している [12]. 本論文では, 山尾らの手法とマルチバンド画像を用いた POC に基づく画像対応付けとを組み合わせることにより, より高精度な 3 次元復元を行う.

本章では, 始めに POC に基づく画像対応付けについて述べる [13, 14]. さらに, マルチバンド画像を用いた対応付けの高精度化手法について述べる. 次に, 多視点マルチバンド画像からの 3 次元復元手法について述べる [15]. まず, 復元手法の大まかな流れを説明した後, それぞれの処理について詳しく述べる.

3.2 位相限定相関法

3.2.1 位相限定相関関数の定義

画像のサイズが $N_1 \times N_2$ ピクセルの 2 枚の画像を $f(n_1, n_2), g(n_1, n_2)$ とする. 定式化の便宜上, 離散空間のインデックスを $n_1 = -M_1, \dots, M_1$ および $n_2 = -M_2, \dots, M_2$ とし, 画像サイズを $N_1 = 2M_1 + 1$ ピクセル, $N_2 = 2M_2 + 1$ ピクセルとする. ここでは説明を簡単にするために離散空間のインデックスを正負対称にとり, 2 次元画像信号の

大きさ N_1, N_2 を奇数としているが、これは必須ではない。

画像 $f(n_1, n_2), g(n_1, n_2)$ の2次元離散フーリエ変換 (2D DFT) をそれぞれ $F(k_1, k_2), G(k_1, k_2)$ として次式で与える。

$$\begin{aligned} F(k_1, k_2) &= \sum_{n_1 n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)} \end{aligned} \quad (3.1)$$

$$\begin{aligned} G(k_1, k_2) &= \sum_{n_1 n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)} \end{aligned} \quad (3.2)$$

ここで、 $k_1 = -M_1, \dots, M_1$ および、 $k_2 = -M_2, \dots, M_2$ 、 $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$ および、 $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$ であり、 $\sum_{n_1 n_2}$ は、 $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2}$ を表す。また、 $A_F(k_1, k_2), A_G(k_1, k_2)$ は振幅スペクトルであり、 $\theta_F(k_1, k_2), \theta_G(k_1, k_2)$ は位相スペクトルである。このとき、 $F(k_1, k_2)$ と $G(k_1, k_2)$ の正規化相互パワースペクトルを次式で与える。

$$\begin{aligned} R(k_1, k_2) &= \frac{F(k_1, k_2) \overline{G(k_1, k_2)}}{|F(k_1, k_2) G(k_1, k_2)|} \\ &= e^{j\theta(k_1, k_2)} \end{aligned} \quad (3.3)$$

ここで、 $\overline{G(k_1, k_2)}$ は、 $G(k_1, k_2)$ の複素共役を示す。また、 $\theta(k_1, k_2) = \theta_F(k_1, k_2) - \theta_G(k_1, k_2)$ である。

位相限定相関関数 (POC 関数) $r(n_1, n_2)$ は、 $R(k_1, k_2)$ の2次元逆離散フーリエ変換 (2D IDFT) であり、次式で定義される。

$$r(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1 k_2} R(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \quad (3.4)$$

ここで、 $\sum_{k_1 k_2}$ は、 $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2}$ を表す。二つの画像が類似している場合、POC 関数は、デルタ関数に近いきわめて鋭いピークを有する。この相関ピークの座標を求めることで、画像間の平行移動量を検出することができる。以下では、微小な平行移動量 (δ_1, δ_2) を考えた場合の POC 関数、平行移動量の推定について述べる。

3.2.2 サブピクセル精度の平行移動量推定

本節では、サブピクセルレベルの画像間の平行移動量を検出する手法について述べる。まず、画像間にサブピクセルレベルの平行移動がある場合の POC 関数の相関ピークモデルを導出する。次に、導出された相関ピークモデルを用いたサブピクセル移動量推定手法について述べる。

連続空間で定義された2次元信号 $f_c(x_1, x_2)$ を考える。 x_1, x_2 は実数である。 δ_1, δ_2 をそれぞれ x_1, x_2 方向に関するサブピクセルの移動量を表す実数とすると、連続空間で $f_c(x_1, x_2)$ を (δ_1, δ_2) だけ移動させた信号は、 $f_c(x_1 - \delta_1, x_2 - \delta_2)$ と表せる。これらの連続空間画像 $f_c(x_1, x_2), f_c(x_1 - \delta_1, x_2 - \delta_2)$ を標本化間隔 T_1, T_2 で標本化した2次元離散空間信号（画像）をそれぞれ $f(n_1, n_2)$ と $g(n_1, n_2)$ とし、次式で定義する。

$$\begin{aligned} f(n_1, n_2) &= f_c(x_1, x_2)|_{x_1=n_1T_1, x_2=n_2T_2} \\ g(n_1, n_2) &= f_c(x_1 - \delta_1, x_2 - \delta_2)|_{x_1=n_1T_1, x_2=n_2T_2} \end{aligned}$$

ただし、 $n_1 = -M_1, \dots, M_1, n_2 = -M_2, \dots, M_2$ とする。このとき、離散空間で定義される画像 $f(n_1, n_2), g(n_1, n_2)$ に関するPOC関数を用いて、連続空間での微小移動量 (δ_1, δ_2) を推定する問題を考える。ただし、一般的には δ_1, δ_2 は、離散空間においてサブピクセルレベルの移動量に対応するものとする。まず、画像 $f(n_1, n_2), g(n_1, n_2)$ の2D DFTをそれぞれ $F(k_1, k_2), G(k_1, k_2)$ とすると、これらの間には次のような近似が成り立つ。

$$G(k_1, k_2) \simeq F(k_1, k_2) \cdot e^{-j\frac{2\pi}{N_1}k_1\delta_1} e^{-j\frac{2\pi}{N_2}k_2\delta_2} \quad (3.5)$$

上式が近似であるのは、連続空間画像と離散空間画像に対するフーリエ変換の性質の違いに起因する。

$F(k_1, k_2), G(k_1, k_2)$ の正規化相互パワースペクトル $R(k_1, k_2)$ は、次式で与えられる。

$$\begin{aligned} R(k_1, k_2) &= \frac{F(k_1, k_2)\overline{G(k_1, k_2)}}{|F(k_1, k_2)\overline{G(k_1, k_2)}|} \\ &\simeq e^{j\frac{2\pi}{N_1}k_1\delta_1} e^{j\frac{2\pi}{N_2}k_2\delta_2} \end{aligned} \quad (3.6)$$

POC関数 $r(n_1, n_2)$ は、 $R(k_1, k_2)$ の2D IDFTとして、次のように表される。

$$\begin{aligned} r(n_1, n_2) &= \frac{1}{N_1 N_2} \sum_{k_1 k_2} R(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \\ &\simeq \frac{1}{N_1 N_2} \sum_{k_1 k_2} e^{j\frac{2\pi}{N_1}k_1\delta_1} e^{j\frac{2\pi}{N_2}k_2\delta_2} W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \\ &= \frac{1}{N_1 N_2} \sum_{k_1} e^{j\frac{2\pi}{N_1}k_1(n_1+\delta_1)} \sum_{k_2} e^{j\frac{2\pi}{N_2}k_2(n_2+\delta_2)} \\ &= \frac{1}{N_1 N_2} \frac{\sin\{\pi(n_1+\delta_1)\}}{\sin\{\frac{\pi}{N_1}(n_1+\delta_1)\}} \frac{\sin\{\pi(n_2+\delta_2)\}}{\sin\{\frac{\pi}{N_2}(n_2+\delta_2)\}} \end{aligned} \quad (3.7)$$

上式は、画像間の微小移動量が (δ_1, δ_2) である場合のPOC関数の一般形を示している。式(3.7)において、 δ_1, δ_2 が整数の場合、POC関数 $r(n_1, n_2)$ は $(n_1, n_2) = (\delta_1, \delta_2)$ のみ

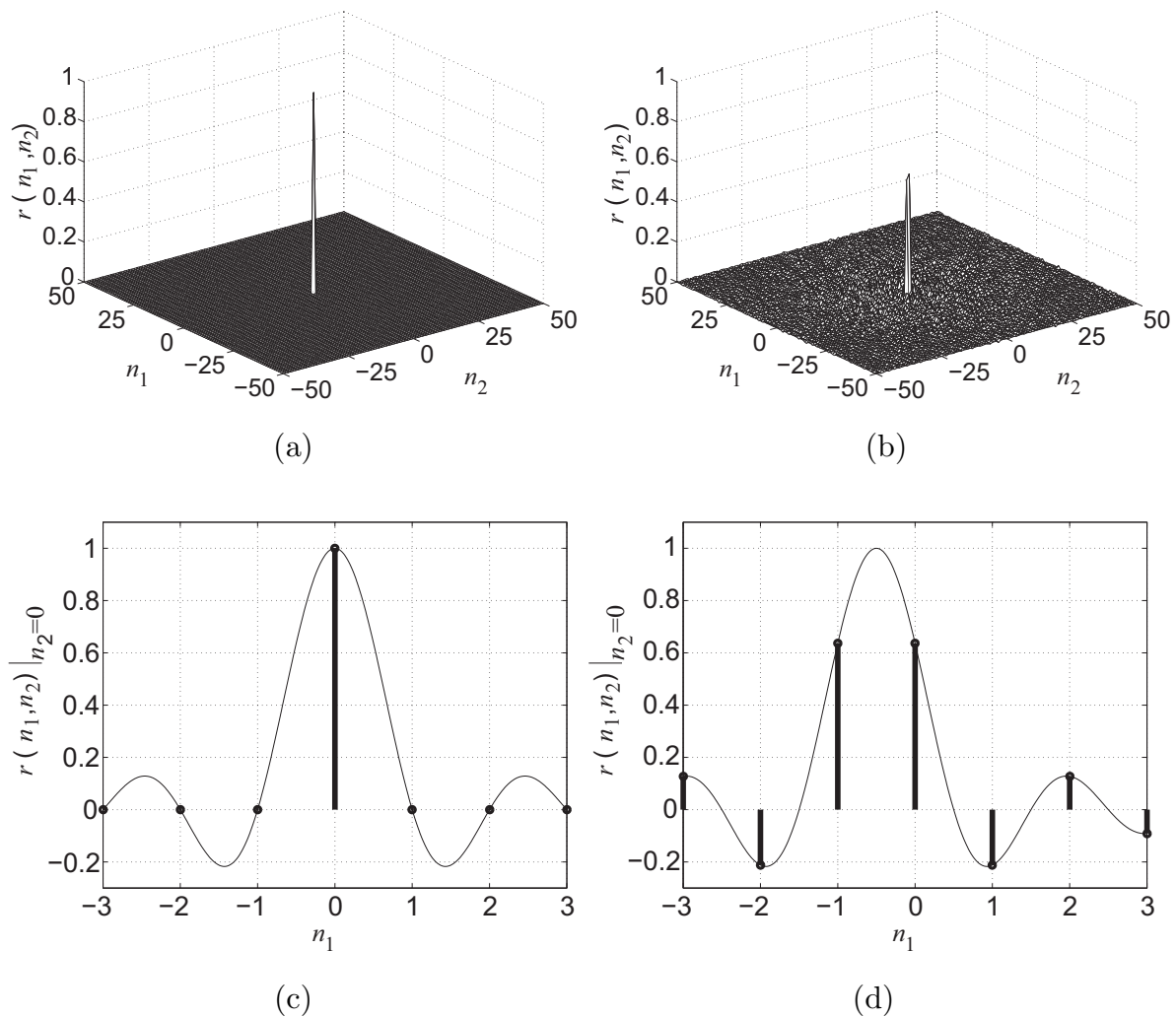


図 3.1 POC 関数 $r(n_1, n_2)$ の 3 次元および 2 次元プロット: (a) $(\delta_1, \delta_2) = (0, 0)$, (b) $(\delta_1, \delta_2) = (0.5, 0)$, (c) 上図 (a) のピーク近傍の拡大図 ($n_2 = 0$), (d) 上図 (b) のピーク近傍の拡大図 ($n_2 = 0$)

で 1 となり, その他では 0 となるような, いわゆる Kronecker のデルタ関数になる*1. 例えば, $(\delta_1, \delta_2) = (0, 0)$ の場合, POC 関数は次式で表せる (図 3.1 (a),(c)).

$$r(n_1, n_2) = \begin{cases} 1 & n_1 = n_2 = 0 \text{ のとき} \\ 0 & \text{その他のとき} \end{cases} \quad (3.8)$$

2 枚の画像が同一の時, POC 関数 $r(n_1, n_2)$ のピーク値は, 常に 1 になる. また, このピーク座標は, 2 枚の画像間の平行移動量 (δ_1, δ_2) に対応している. よって, $r(n_1, n_2)$ のピーク座標を検出することで, 2 枚の画像間の平行移動量を推定することが可能である.

*1 実際には, 画像の平行移動とともに画像端に一致しない画像が現れ, この画像端部分がノイズとして働くために $r(n_1, n_2)$ のピークの高さは減少することになる. ここでは, この減少が問題にならないような微小移動を想定している.

しかし, δ_1 あるいは δ_2 が非整数の場合, $r(n_1, n_2)$ のピーク座標はピクセル間に存在するため, 画像間の平行移動量を正確に推定することが困難になる. ここで, POC 関数の一般式が式 (3.7) で表されることを考慮すると, 実際に計算された POC 関数に式 (3.7) を数値的にフィッティングすることにより, $r(n_1, n_2)$ のピーク座標を高精度に求めることが可能であると考えられる.

実際には, 種々の外的要因が存在するため, POC 関数 $r(n_1, n_2)$ のピークの高さの変化を表すパラメータを導入し, 2枚の画像に微小変化と平行移動 (δ_1, δ_2) が同時に存在する場合の $r(n_1, n_2)$ のモデルとして次式を仮定する.

$$r(n_1, n_2) \simeq \frac{\alpha}{N_1 N_2} \frac{\sin\{\pi(n_1 + \delta_1)\}}{\sin\{\frac{\pi}{N_1}(n_1 + \delta_1)\}} \frac{\sin\{\pi(n_2 + \delta_2)\}}{\sin\{\frac{\pi}{N_2}(n_2 + \delta_2)\}} \quad (3.9)$$

ここで, α は, $r(n_1, n_2)$ のピークの高さを表現するために導入されたパラメータである. ノイズ, 画像の撮影条件などのさまざまな外的要因による変化が加わるため, $\alpha \leq 1$ となることが実験を通して確認されている. $\alpha, \delta_1, \delta_2$ をパラメータとして, 上式の相関ピークのモデル式を実際に計算された POC 関数の計算値にフィッティングすることにより, 高精度な平行移動量推定が可能である.

3.2.3 マルチバンド画像を用いた POC

POC はマルチバンド画像を用いることでグレースケール画像を用いるよりも高精度な画像マッチングが可能である [15]. 異なる波長帯域を持つ複数枚の 2次元画像信号 $f_i(n_1, n_2), g_i(n_1, n_2)$ ($i = 1, 2, \dots, N$) から, 1つの POC 関数 $\hat{r}(n_1, n_2)$ を計算する. ここで i はそれぞれの波長帯域を表す添字であり, N は波長帯域数を表す (赤青緑の3つの波長帯域の画像の場合 $N = 3$ である). この時の平均正規化相互パワースペクトル $\hat{R}(k_1, k_2)$ を, 各波長帯域における正規化相互パワースペクトル $R_i(k_1, k_2)$ の重み付き平均で与える.

$$\hat{R}(k_1, k_2) = \frac{\sum_{i=1}^N W_i(k_1, k_2) R_i(k_1, k_2)}{\sum_{i=1}^N W_i(k_1, k_2)} \quad (3.10)$$

ここで $W_i(k_1, k_2)$ は各周波数インデックスにおける波長帯域ごとの重み関数である. POC 関数 $\hat{r}(n_1, n_2)$ は, 平均正規化相互パワースペクトル $\hat{R}(k_1, k_2)$ の IDFT で与えられる. 本論文では, 各周波数インデックスにおける画像信号のエネルギーによって重みを決定する. このとき, 重み関数 $W_i(k_1, k_2)$ は,

$$W_i(k_1, k_2) = \left| F(k_1, k_2) \overline{G(k_1, k_2)} \right| \quad (3.11)$$

となる. この重み関数は, 各周波数インデックスにおける波長帯域ごとのエネルギーの比を表しており, ある周波数インデックスにおいて, SN 比の大きい波長帯域の情報ほど大

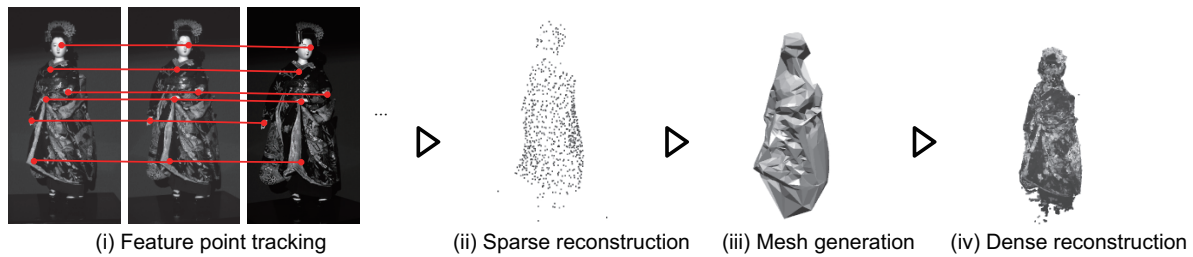


図 3.2 多視点マルチバンド画像を用いた3次元復元: (i) マルチバンド画像を用いた POC に基づく画像のトラッキング, (ii) SfM に基づく疎な3次元復元, (iii) 粗い3次元メッシュモデルの生成, (iv) 最適なフレーム選択に基づく密な3次元復元

大きく影響する関数となっている。

3.3 多視点マルチバンド画像からの3次元復元

提案手法は, 図 3.2 のように, (i) マルチバンド画像を用いた POC に基づく特徴点のトラッキング, (ii) SfM に基づく疎な3次元復元, (iii) 粗い3次元メッシュモデルの生成, (iv) 最適な画像選択に基づく密な3次元復元の4つのステップから構成される. 本論文では, 処理の簡単化のために, 計測対象の全周囲を連続で撮影された画像シーケンスを入力画像群とする. まず, 入力された画像から特徴点の検出と, マルチバンド画像を用いた POC に基づく画像対応付けによる特徴点のトラッキングを行い, 隣接する画像間での特徴点の対応を得る. 次に, 画像間の特徴点の対応から, SfM を用いたカメラパラメータ推定と, 疎な3次元点群の復元を行う. そして, グラフカットに基づくエネルギー最小化 [57, 58] により疎な3次元点群から, 物体の粗い3次元メッシュモデルを生成する. さらに, 粗い3次元メッシュモデルの各メッシュについて, メッシュの位置・法線, カメラの位置・姿勢, 画像間の基線長, オクルージョンを考慮して, 画像群から基準画像と対応付けを行う近傍画像を選択する. 最後に, 選択した画像間でマルチバンド画像を用いた POC に基づく画像対応付けを行い, 密な3次元点群を復元する. 以下では, 各ステップの処理について詳細に述べる. (i) 特徴点のトラッキング

特徴点のトラッキングでは, 先頭の画像シーケンスから順番に特徴点の検出とトラッキングを行うことで, 画像間の特徴点の対応を求める [5]. このとき, 後段のメッシュモデル生成で, 各メッシュを比較的均一な大きさするために, 物体の表面に均一な密度で分布するような特徴点をトラッキングする. 以下では, K 視点のマルチバンド画像から構成される画像の i 番目の画像を I_i ($i = 1, 2, \dots, K$), 特徴点の密度を制御するためのパラ

メータを D とする．画像 I_i における特徴点の検出とトラッキングの概略を図 3.3 に示し，具体的な処理を説明する．

まず，画像 I_i から特徴点を検出する．特徴点の検出に利用する画像はカメラ感度の高い波長の画像を利用する．本論文では，Shi らの最小固有値法 [59] を用いた特徴点の検出を行う．Shi らの手法は，Harris らのコーナー点検出手法 [60] や，Rosten らの局所輝度比較法 [61] に比べて，多くの特徴点を偏りなく検出可能である．Shi らの手法を用いて， I_i の各画素 (u, v) における特徴量 $h(u, v)$ は以下のように表せる．

$$h(u, v) = \min(\lambda_1, \lambda_2) \quad (3.12)$$

λ_1 , λ_2 は，次式で与えられる自己相関行列 $A(u, v)$ の2つの固有値である．

$$A(u, v) = \begin{pmatrix} \sum_{x,y} w(x, y) I_u^2 & \sum_{x,y} w(x, y) I_u I_v \\ \sum_{x,y} w(x, y) I_u I_v & \sum_{x,y} w(x, y) I_v^2 \end{pmatrix} \quad (3.13)$$

ここで， I_u, I_v は I_i の u, v 方向の1次微分， $w(x, y)$ は画素 (u, v) を中心とする2次元のガウシアン型窓関数である．特徴量 $h(u, v)$ が閾値以上である画像領域において， $h(u, v)$ が極大値となる画素 (u, v) を特徴点とし，その特徴点の集合を $F_{\text{detected}, i}$ とする（図 3.3(a)）．このとき，ポアソンディスクサンプリング [62] を組み合わせることで，画像上でのユークリッド距離が少なくとも D ピクセルだけ離れるように特徴点を検出し， I_i で検出される特徴点の密度を制御する．次に， I_i で検出した特徴点の集合 $F_{\text{detected}, i}$ と，1フレーム前の画像 I_{i-1} からトラッキングした特徴点の集合 $F_{\text{tracked}, i}$ を統合し， I_i における特徴点の集合 F_i を得る．このとき，特徴点の密度を均一に保つために， $F_{\text{tracked}, i}$ と近い領域で検出された特徴点を $F_{\text{detected}, i}$ から削除する．特徴点 $m_j \in F_{\text{tracked}, i}$ を中心として半径 D ピクセルの局所的な円形領域を D_j とするとき，次式により F_i を求める（図 3.3(b)(c)）．

$$F_i = F_{\text{tracked}, i} \cup (F_{\text{detected}, i} \setminus D) \quad (3.14)$$

ただし， $D = \bigcup_j D_j$ である．なお， $F_{\text{tracked}, 1} = \emptyset$ と定義する．

最後に，画像 I_i と，その次の画像 I_{i+1} の間で画像対応付けを行い， $F_{\text{tracked}, i+1}$ を求める（図 3.3(d)）．このときの特徴点の対応関係をもとに，次のステップでカメラ位置を推定し，疎な3次元復元を行うため，特徴点の対応付けの精度が復元された3次元モデルの精度に大きく影響する．本論文では，高精度な対応付けが可能なマルチバンド画像を用いたPOCに基づく画像対応付けを行う [15]．このとき，POC関数の相関ピーク値が閾値以下である点は，誤対応点として除去する．

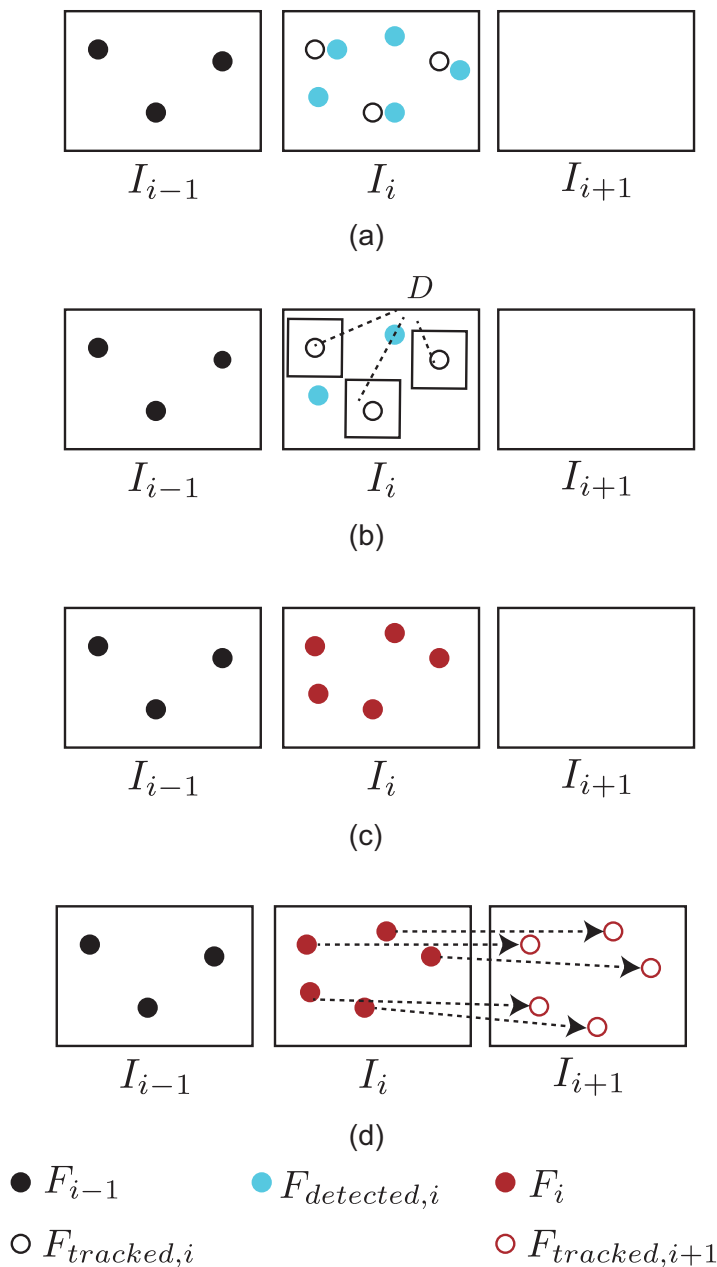


図 3.3 画像 I_i における特徴点の検出とトラッキング: (a) I_i に含まれる特徴点の検出, (b) 直前の画像 I_{i-1} からトラッキングした特徴点と重複する特徴点の削除, (c) I_i における特徴点の決定, (d) 次の画像 I_{i+1} への特徴点のトラッキング

(ii) 疎な3次元復元

画像シーケンスのトラッキングで得られた結果から, SfM を用いて, カメラの外部パラメータの推定と, 物体の疎な3次元点群の復元を逐次的に実行する [16, 5]. SfM を用

いたカメラの外部パラメータ推定と疎な3次元点群の復元は，初期復元と逐次復元の2ステップで構成される．さらに，バンドル調整を定期的に行うことで，推定した外部パラメータ，3次元点群の座標を最適化する [24, 25]．以下では，初期復元と逐次復元，バンドル調整について具体的な処理を説明する．カメラの内部パラメータ K_i は事前のキャリブレーションにより求めておくものとする．

初期復元

2枚の画像 I_1, I_α に対応するカメラ C_1, C_α について，それぞれの外部パラメータ R_1, t_1 および R_α, t_α を推定し， C_1 と C_α の外部パラメータと特徴点の対応から3次元点群を復元する．まず， C_1 の回転行列 R_1 を 3×3 の単位行列 I ，並進ベクトル t_1 を 3×1 の零ベクトルとし， C_1 の座標系を世界座標系とする． C_1 と C_α の間の対応点の關係を用いて， C_1 に対する C_α の相対的なカメラ位置・姿勢を求める．本論文では，正規化5点アルゴリズム [28] を用いて， C_1 と C_α の間の基本行列 E を求める．正規化5点アルゴリズムは，8点アルゴリズムや7点アルゴリズム [16]，6点アルゴリズム [63] に比べて，対応点ペアのノイズに対して頑健で，安定したパラメータ推定が可能である．また，RANSAC [30] を組み合わせることで，誤対応点に対してロバストに基本行列 E を推定可能である．RANSAC では，ランダムで選んだ様々な対応点ペアの組み合わせを用いてパラメータの推定と評価を行い，なるべく多くの対応点ペアが幾何的に整合するようなパラメータを求める．本論文では，正規化画像座標の対応点ペア x, x' に関する基本行列 E の幾何的な整合性を，以下に示す Sampson 誤差 ϵ を用いて評価する．

$$\epsilon = \frac{\tilde{x}'^T E \tilde{x}}{l_1^2 + l_2^2 + l_1'^2 + l_2'^2} \quad (3.15)$$

ただし， \tilde{x} は x の同次座標表現， \tilde{x}'^T は \tilde{x}' の転置であり，

$$\begin{aligned} E \tilde{x} &= [l_1, l_2, l_3]^T \\ E^T \tilde{x}' &= [l_1', l_2', l_3']^T \end{aligned}$$

である．もし，Sampson 誤差 ϵ が閾値以下であれば， E は x, x' に関して幾何的な整合性を満たす．そして，推定した基本行列 E を分解し，カメラ C_α の回転行列 R_α と並進ベクトル t_α を求める [16]．最後に， C_1 と C_α の間で，三角測量による特徴点の3次元復元を行う． C_1 の特徴点のデジタル画像座標 $m = (u, v)$ と， m に対応する C_α の特徴点のデジタル画像座標 $m' = (u', v')$ から，特徴点の3次元座標 M を次式で計算する．

$$M = B^+ b \quad (3.16)$$

ただし

$$B = \begin{pmatrix} up_{31} - p_{11} & up_{32} - p_{12} & up_{33} - p_{13} \\ vp_{31} - p_{21} & vp_{32} - p_{22} & vp_{33} - p_{23} \\ u'p'_{31} - p'_{11} & u'p'_{32} - p'_{12} & u'p'_{33} - p'_{13} \\ v'p'_{31} - p'_{21} & v'p'_{32} - p'_{22} & v'p'_{33} - p'_{23} \end{pmatrix}$$

$$b = \begin{pmatrix} p_{14} - up_{34} \\ p_{24} - vp_{34} \\ p'_{14} - u'p'_{34} \\ p'_{24} - v'p'_{34} \end{pmatrix}$$

である． B^+ は B の擬似逆行列である． p_{nm} , p'_{nm} は，それぞれ， C_1 , C_α の射影行列 $P_1 = K_1[R_1 t_1]$, $P_\alpha = K_\alpha[R_\alpha t_\alpha]$ における n 行 m 列の要素である．また，復元した3次元点群のなかで，再投影誤差が閾値よりも大きい3次元点を誤対応点として除去する．

逐次復元

逐次復元では，追加された画像 I_i に対応するカメラ C_i の外部パラメータ R_i , t_i を推定し，3次元点群を追加復元する．まず，追加された画像 I_i の特徴点とすでに復元されている3次元点群との対応関係から，カメラ C_i の外部パラメータ R_i , t_i を求める．本論文では，Kneip らの手法 [29] を用いる．Kneip らの手法では，特徴点と3次元点との3組の対応点ペアから記述される P3P (Perspective-Three-Point) 問題を解くことで， R_i , t_i を推定する．また，初期復元と同様に，RANSAC を組み合わせることで，誤対応点に対してロバストにパラメータを推定する． R_i , t_i の幾何的な整合性は，それらを用いて得られる再投影誤差を用いて評価する．そして，カメラ C_i の外部パラメータを求めたら，まだ復元されていない C_i の特徴点の中から，外部パラメータが既知のカメラと対応する特徴点を探す．得られた特徴点ペアから，式 (3.16) を用いて3次元点を復元し，3次元点群に追加する．

バンドル調整

バンドル調整を用いたパラメータ最適化では，非線形最適化に基づく再投影誤差の最小化により，カメラの外部パラメータと3次元点群を最適化する [16, 5, 24]．本論文では，最適化の対象範囲の異なる2種類のバンドル調整として，大域バンドル調整と局所バンドル調整を考える．大域バンドル調整では，先頭のカメラ C_1 を除く全てのカメラの外部パラメータと，全ての3次元点を最適化の対象とする．カメラ C_i ($i = 2, 3, \dots, N_p$) の外部パラメータ p_i の集合を P , 復元した3次元点 q_j ($j = 1, 2, \dots, N_q$) の集合を Q とする．ここで， N_p はカメラの数， N_q は3次元点数である．大域バンドル調整で最適化する再

投影誤差に基づくコスト関数 $E_g(P, Q)$ を次式で定義する．

$$E_g(P, Q) = \frac{1}{2} \sum_i \sum_j |m_{i,j} - m_{\text{rep}}(p_i, q_j)|^2 \quad (3.17)$$

ここで， $m_{i,j}$ は，カメラ C_i における，3次元点 q_j と対応する特徴点のデジタル画像座標である．また， $m_{\text{rep}}(p_i, q_j)$ は，外部パラメータ p_i を用いてカメラ C_i に3次元点 q_j を投影したときの，投影点のデジタル画像座標である．大域バンドル調整は，全3次元点の再投影誤差の Root Mean Square (RMS) が閾値を超えた場合，もしくは，追加した画像のフレームが一定数を超えた場合に行う．

局所バンドル調整では，カメラ C_i の外部パラメータ p_i を最適化の対象とする． C_i の特徴点と対応する3次元点を q'_j ($j = 1, 2, \dots, N_{q'}$) とする．ただし， $N_{q'}$ は， C_i で観測される3次元点群の点数である．局所バンドル調整で最適化するコスト関数 $E_l(p_i)$ を次式で定義する．

$$E_l(p_i) = \frac{1}{2} \sum_j |m_{i,j} - m_{\text{rep}}(p_i, q'_j)|^2 \quad (3.18)$$

局所バンドル調整は，カメラ C_i の外部パラメータを推定した直後に実行する．

また，本論文では，コスト関数の最小化に，非線形最小二乗アルゴリズムの1つである Levenberg-Marquardt (LM) 法を用いる．

(iii) 粗い3次元メッシュモデルの生成

SfM で復元した疎な3次元点群から，Labatut らの手法 [57, 58] を用いて，物体の粗い3次元メッシュモデルを生成する．まず，3次元ドロネー四面体分割により，疎な3次元点群を頂点とする四面体群を生成する．生成した四面体群から，四面体がノード，四面体の隣接面がエッジに対応するグラフを作成する．作成したグラフに対し，グラフカットに基づくエネルギー最小化を行うことで，各四面体が物体の内部にあるか外部にあるかを示す2値のラベリングを行う．グラフカットで最小化するエネルギー関数 E_m は次式で定義される．

$$E_m(S) = E_{\text{vis}}(S) + \lambda_{\text{photo}} E_{\text{photo}}(S) + \lambda_{\text{area}} E_{\text{area}}(S) \quad (3.19)$$

ここで， S は四面体を構成する面の集合であり， λ_{photo} と λ_{area} は正の重みである．また， $E_{\text{vis}}(S)$ ， $E_{\text{photo}}(S)$ ， $E_{\text{area}}(S)$ は，それぞれ可視性，Photo-consistency，滑らかさを表す項である．これらの項に関する詳細な定義は文献 [57, 58] を参照されたい．最後に，異なるラベルの四面体を結ぶエッジの集合，すなわち，物体内部と物体外部の境界に相当する面の集合を抽出することで，物体の粗い3次元メッシュモデルを得る．

(iv) 最適な画像選択に基づく密な3次元復元

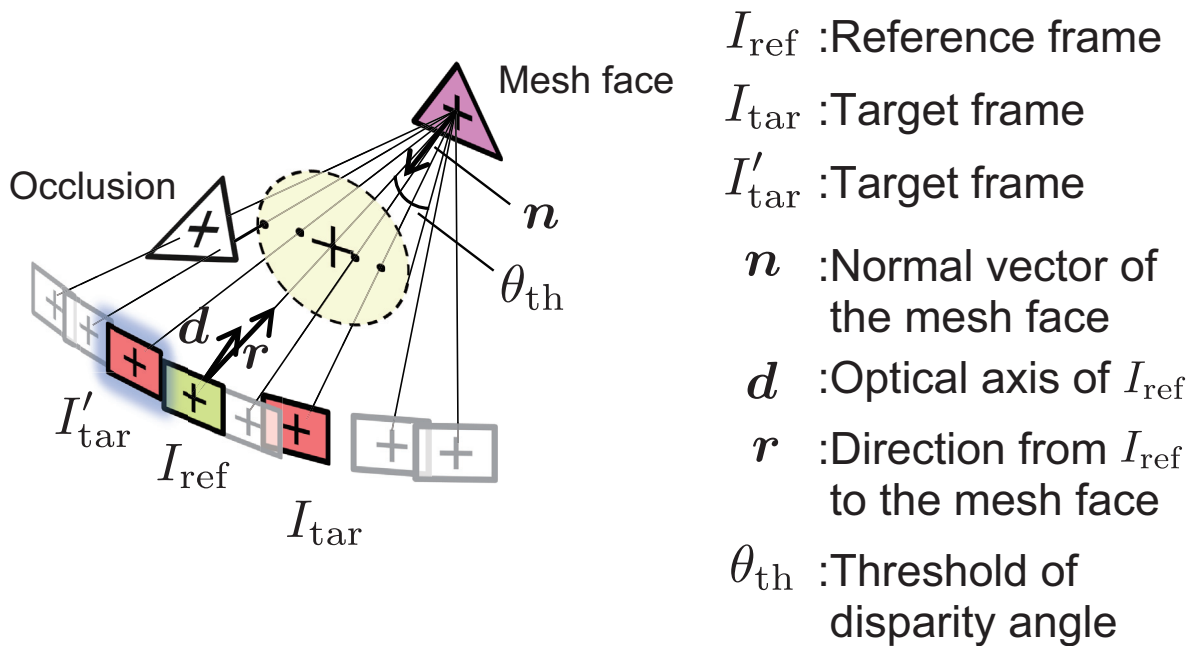


図 3.4 局所領域の3次元復元に最適な画像の選択

粗い3次元メッシュモデルの各メッシュごとに3次元復元に最適な画像を選択し、選択した画像間の密な画像対応付けで得られる対応点から、メッシュ面の領域における3次元点群を復元する。

基線長の短い画像ペアを用いて3次元復元を行う場合、画像変形が小さいため安定した画像対応付けが可能であるが、画像対応付けの誤差が3次元復元結果に与える影響が大きくなるため、3次元復元精度が低下する。一方で、基線長の長い画像ペアを用いて3次元復元を行う場合、画像変形が大きくなったり、オクルージョンが発生しやすくなったりするため、画像対応付けの誤差が増加し、3次元復元精度が低下する。

そこで、図3.4に示すように、粗い3次元メッシュモデルのメッシュ面ごとに、オクルージョンや基線長を考慮し、3次元復元に最適な画像を選択する。まず、オクルージョンが発生せず、メッシュ面と最も正対するカメラの画像を基準画像 I_{ref} として選択する。次に、基準画像 I_{ref} と正確な画像対応付けが可能な範囲で、できるだけ長い基線長をとる画像を近傍画像 I_{tar} として選択する。また、基準画像 I_{ref} を中心として近傍画像 I_{tar} となるべく対称の位置にある画像を2枚目の近傍画像 I'_{tar} として選択する。そして、2組の画像ペア $(I_{\text{ref}}, I_{\text{tar}})$ 、 $(I_{\text{ref}}, I'_{\text{tar}})$ のそれぞれにおいて、メッシュ面に相当する画像領域を対象として、マルチバンド画像を用いたPOCに基づく密な画像対応付けを行う。最後に、式(3.16)により、近傍画像のペア $(I_{\text{tar}}, I'_{\text{tar}})$ の間の対応点から密な3次元点群を復元する。

POC に基づく画像対応付けは、画像間の変形に平行移動のみを仮定しているため、画像間の射影的な幾何変形が大きい場合、画像対応付けが困難となる。これに対して、本論文では、Ishii らの手法 [64] を用いて、射影変換により画像の局所的な変形を補正することで、画像変形が大きい場合においても正確に画像対応付けを行う。メッシュ面 S に関する、カメラ C からカメラ C' への射影変換行列 H を次式で計算する。

$$H = K' \left(R + \frac{tn^T}{d} \right) K^{-1} \quad (3.20)$$

ここで、 n はメッシュ面の法線ベクトル、 d はカメラ C からメッシュ面への距離である。 R 、 t は、それぞれ C から C' への回転行列と並進ベクトルである。 K 、 K' は、それぞれ C と C' の内部パラメータである。 n^T は n の転置である。

あるメッシュ面における最適なフレーム選択と密な3次元復元の具体的な処理手順を以下に述べる。

ステップ1 (メッシュ面が観測可能なカメラ群の抽出):

全てのカメラ C の中から、メッシュ面を観測できる位置にあり、かつ、オクルージョンが生じないカメラの集合 C_{vis} を次式により抽出する。

$$C_{\text{vis}} = \{C \in C \mid (n \cdot d_C < 0) \wedge (n \cdot r_C < 0) \wedge P_{\text{vis}}(C)\} \quad (3.21)$$

ここで、 d_C 、 n 、 r_C は、それぞれ、カメラ C の光軸方向の単位ベクトル、メッシュ面の単位法線ベクトル、カメラ C からメッシュ面の重心へ向かう単位ベクトルである。 $P_{\text{vis}}(C)$ は、カメラ C とメッシュ面の上に別のメッシュ面が存在しない場合に真となる論理式である。本論文では、 $P_{\text{vis}}(C)$ の計算に Havel らの手法 [65] を用いる。

ステップ2 (基準フレームの選択):

メッシュ面と最も正対するカメラを基準カメラ C_{ref} として次のように選択する。

$$C_{\text{ref}} = \arg \max_{C \in C_{\text{vis}}} ((n \cdot d_C)(n \cdot r_C)) \quad (3.22)$$

そして、 $C_{\text{vis}} \leftarrow C_{\text{vis}} \setminus \{C_{\text{ref}}\}$ とする。選択した基準カメラ C_{ref} に対応するフレームを基準フレーム I_{ref} とする。

ステップ3 (輻輳角に基づくカメラ群の抽出):

基準カメラ C_{ref} との輻輳角が θ_{th} 以下であるカメラの集合 C_{match} を、 C_{vis} から次のように抽出する。

$$C_{\text{match}} = \{C \in C_{\text{vis}} \mid r_C \cdot r_{C_{\text{ref}}} \geq \cos \theta_{\text{th}}\} \quad (3.23)$$

θ_{th} は、正確な画像対応付けが可能な輻輳角の最大値として設定されるパラメータである。

ステップ4 (2枚の近傍フレームの選択):

C_{match} のなかで, C_{ref} との輻輳角が最大となるカメラを1つ目の近傍カメラ C_{tar} として次のように選択する.

$$C_{\text{tar}} = \arg \min_{C \in C_{\text{match}}} (r_C \cdot r_{C_{\text{ref}}}) \quad (3.24)$$

そして, $C_{\text{match}} \leftarrow C_{\text{match}} \setminus \{C_{\text{tar}}\}$ とする. また, C_{ref} を中心として C_{tar} と対称の位置にあり, C_{tar} との輻輳角が最大となるカメラを2つ目の近傍カメラ C'_{tar} として以下のように選択する.

$$C'_{\text{tar}} = \arg \min_{C \in C_{\text{sym}}} (r_C \cdot r_{C_{\text{tar}}}) \quad (3.25)$$

ここで

$$C_{\text{sym}} = \{C \in C_{\text{match}} \mid (r_{C_{\text{ref}}} \times r_{C_{\text{tar}}}) \cdot (r_{C_{\text{ref}}} \times r_C) > 0\} \quad (3.26)$$

である. そして, $C_{\text{match}} \leftarrow C_{\text{match}} \setminus \{C'_{\text{tar}}\}$ とする. 選択した近傍カメラ $C_{\text{tar}}, C'_{\text{tar}}$ に対応するフレームを, それぞれ近傍フレーム $I_{\text{tar}}, I'_{\text{tar}}$ とする.

ステップ5 (画像対応付けの基準点の取得):

メッシュ面上に均一な密度で配置した3次元点を I_{ref} に投影し, 投影点のデジタル画像座標を画像対応付けの基準点とする.

ステップ6 (画像対応付け):

2組のフレームペア $(I_{\text{ref}}, I_{\text{tar}}), (I_{\text{ref}}, I'_{\text{tar}})$ のそれぞれにおいて, 式(3.20)の射影変換行列を用いて局所的な画像変形を補正し, POCに基づく画像対応付けにより, 基準点と対応する点を取得する. このとき, POC関数の相関ピーク値が閾値以下である対応点を誤対応点とする. 誤対応点が存在する場合はステップ7に進み, 誤対応とされた基準点ごとに, より適切な近傍フレームを選択する. 一方で, 全ての対応点の相関ピーク値が閾値よりも大きい場合や, C_{match} が空集合である場合はステップ8へ進む.

ステップ7 (近傍フレームの再選択):

フレームペア $(I_{\text{ref}}, I_{\text{tar}})$ の間で誤対応点が存在する場合, C_{match} のなかで, C'_{tar} との輻輳角が最大となるカメラを C_{tar} として次のように再選択する.

$$C_{\text{tar}} = \arg \min_{C \in C_{\text{match}}} (r_C \cdot r_{C'_{\text{tar}}}) \quad (3.27)$$

そして, $C_{\text{match}} \leftarrow C_{\text{match}} \setminus \{C_{\text{tar}}\}$ とする. 同様に, フレームペア $(I_{\text{ref}}, I'_{\text{tar}})$ の間で誤対応点が存在する場合, C_{match} のなかで, C_{tar} との輻輳角が最大となるカメラを C'_{tar}

として次のように再選択する．

$$C'_{\text{tar}} = \arg \min_{C \in C_{\text{match}}} (r_C \cdot r_{C_{\text{tar}}}) \quad (3.28)$$

そして， $C_{\text{match}} \leftarrow C_{\text{match}} \setminus \{C'_{\text{tar}}\}$ とする．再選択した近傍カメラ C_{tar} ， C'_{tar} に対応するフレームを，それぞれ近傍フレーム I_{tar} ， I'_{tar} として更新する．そして，誤対応とされた基準点のみを対象に，ステップ 6 に戻り画像対応付けを実行しなおす．

ステップ 8 (3次元復元)：

近傍フレームのペア $(I_{\text{tar}}, I'_{\text{tar}})$ の間の対応から，メッシュ面に相当する領域の3次元点群を式 (3.16) により復元する．

3.4 むすび

本章では，多視点マルチバンド画像を用いた3次元復元手法について述べた．始めに POC に基づく画像対応付けについて述べた．さらに，マルチバンド画像を用いた POC に基づく画像対応付けの高精度化手法について述べた．次に，マルチバンド画像を用いた POC を適用した多視点マルチバンド画像からの3次元復元手法について述べた．

第 4 章

多視点マルチバンド画像を用いた色再現

4.1 まえがき

色収差とは，レンズで像をつくる際，像の色ズレとして現れる収差であり，マルチバンド画像を用いて色再現を行った際に，見た目に大きな影響を及ぼす．また，多視点のマルチバンド画像から物体の全周囲の色再現を行うには，物体の形状やカメラ位置を考慮して，画像を選択する必要がある．

本章では，多視点のマルチバンド画像から画像上の各ピクセルにおける分光反射率を推定する手法を述べる．まず，マルチバンド画像の色収差の補正法について述べる．そして，色再現を行う際の適切な画像選択について述べ，マルチバンド画像を用いた分光反射率の推定法について述べる．

4.2 色収差の補正

色収差は，軸上色収差と倍率色収差の 2 つに分けられる．図 4.1 にそれぞれの色収差の概略図を示す．軸上色収差（図 4.1(a)）は，光の波長によって像の結像位置が変わる収差である．波長によっては画像平面にピンぼけした像が投影されるため，色のにじみとして観測される．レンズの端を通る光ほど，結像位置に差がでるため，図 4.1(b) のようにレンズの絞りによって収差の影響を低減できる．倍率色収差（図 4.1(c)）は，光の波長によって像の倍率が変わる収差である．波長ごとに，異なる画像座標上に投影されるため，色ズレとして観測される．軸上色収差とは異なり，レンズの中心を通る光でも生じる収差のため，図 4.1(d) のようにレンズの絞りをを用いても，影響を低減することはできない．本論文では，軸上色収差はレンズ絞りによって軽減されているものと考え，倍率色収差の補正を行う．倍率色収差の補正は，(i) 固定したチェッカーボードのマルチバンド画像を撮影，(ii) カメラの内部パラメータ推定（コーナー点検出），(iii) 倍率色収差のパラメータを推定，(iv) 推定したパラメータから倍率色収差の補正，の 4 ステップから構成される．以下

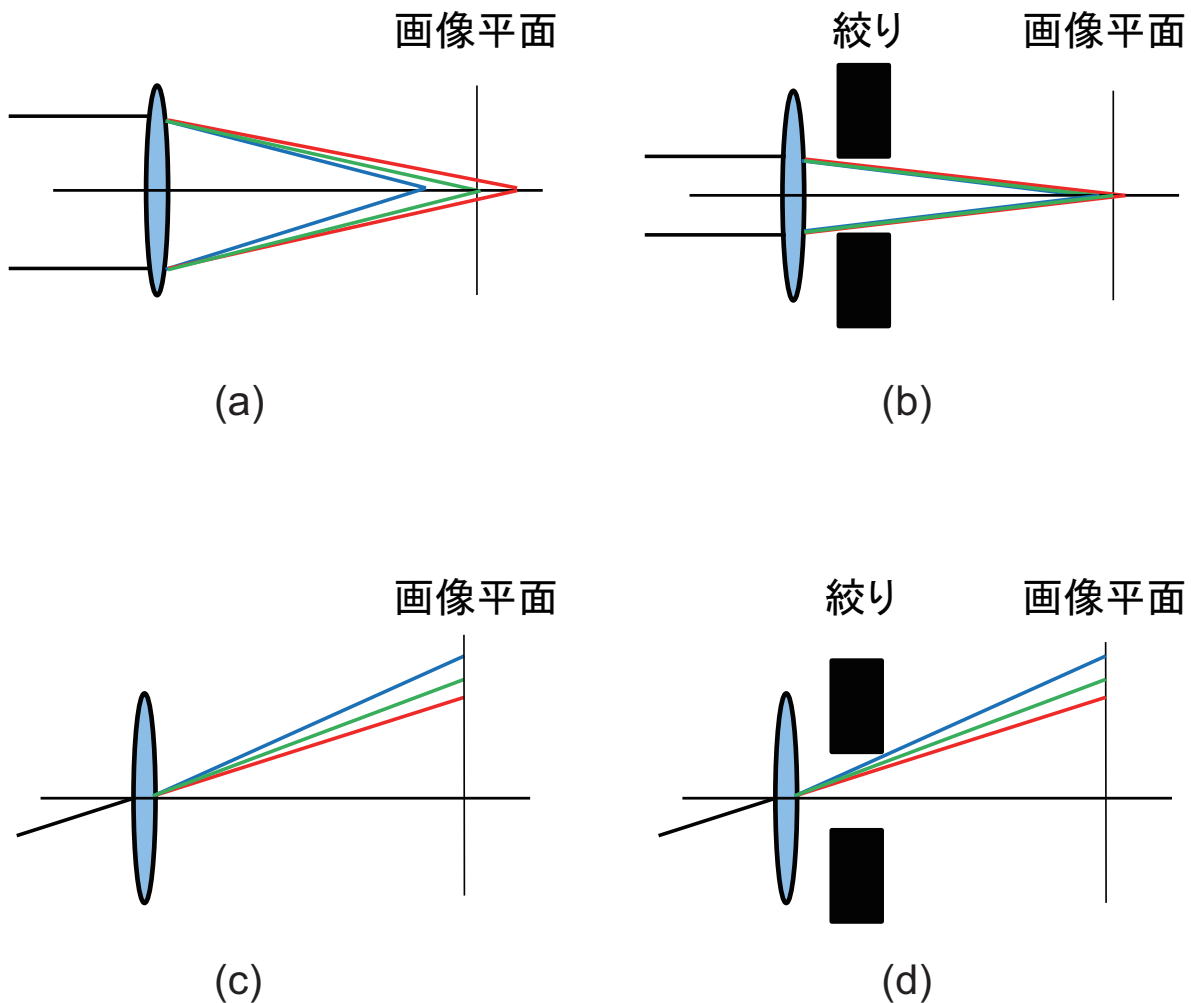


図 4.1 色収差の概略図：(a) 軸上色収差，(b) 軸上色収差（レンズ絞りあり），
(c) 倍率色収差，(d) 倍率色収差（レンズ絞りあり）

に詳細を述べる．

(i) チェッカーボードの撮影

チェッカーボードのマルチバンド画像を撮影する．この時，検出コーナー点の座標を用いて補正を行うので，チェッカーボードを完全に固定して撮影を行う．また，チェッカーボードの画像はカメラの内部パラメータ推定に用いるため，チェッカーボードの位置や角度を変えて複数枚撮影する．本論文では 20 視点 × 8 バンド分の画像を撮影した．

(ii) カメラの内部パラメータ推定（コーナー点検出）

撮影したマルチバンド画像を用いて各バンドごとのカメラの内部パラメータの推定を行う．本論文では Camera Calibration Toolbox for Matlab ^{*1} によるキャリブレーション

^{*1} http://www.vision.caltech.edu/bouguetj/calib_doc/

を行った．

(iii) 倍率色収差のパラメータの推定

図 4.3 にレンズにおける光の屈折を簡略化した倍率色収差の概略図を示す．図 4.3 では，レンズの厚みを無視し，レンズの垂直軸を屈折の境界面としている．2つの異なる種類の媒質（屈折率 n_1, n_2 ）があるとき，その境界面に入射する光の角度 θ_1 と，射出される光の角度 θ_2 の関係は，スネルの法則によって以下のように表される．

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (4.1)$$

入射角 θ で入射する2つの光（波長 r ，波長 b ）がレンズを通過し，出射角 θ_r, θ_b で出射する時，出射角 θ_r, θ_b の関係は，それぞれの波長におけるレンズの屈折率 n_r, n_b を用いて以下のように表せる [66]．

$$\begin{aligned} n_r \sin \theta_r &= n_b \sin \theta_b \\ n_r \sin \theta_r / \cos \theta_b &= n_b \tan \theta_b \end{aligned}$$

ここで $\cos \theta_b \simeq \cos \theta_r$ として，入射する2つの光がそれぞれ画像の x 軸座標 x_r, x_b に投影される時，

$$\begin{aligned} n_r \tan \theta_r &\simeq n_b \tan \theta_b \\ n_r x_r / f &\simeq n_b x_b / f \\ x_r &\simeq \alpha x_b \end{aligned} \quad (4.2)$$

ただし， $\alpha = n_b / n_r$ である．上式から，画像の x 軸方向において，波長 r の画像に対する波長 b の画像の倍率が α であることがわかる．この α を倍率色収差のパラメータとして推定する．ある波長帯域 n の画像で検出したコーナー点の座標と，他の波長帯域 m の画像で検出したコーナー点の座標は，倍率色収差とレンズ歪によって異なる値となる．各波長帯域ごとに推定したカメラの内部パラメータを用いて，レンズ歪を除去する．

波長帯域 n の画像で検出したコーナー点の座標は，波長帯域 m の画像で検出したコーナー点の座標と倍率色収差を表す倍率 α_{nm} を用いて，式 4.2 と同様に以下のように表せる．

$$x_n \simeq \alpha_{nm} x_m \quad (4.3)$$

上式より，以下の式の E が最小となるような α_{nm} を，最小二乗法により求めることで倍率色収差のパラメータを推定する．

$$x_n - \alpha_{nm} x_m = E \quad (4.4)$$

また，画像の y 軸方向について同様のモデルで表すことが可能である．

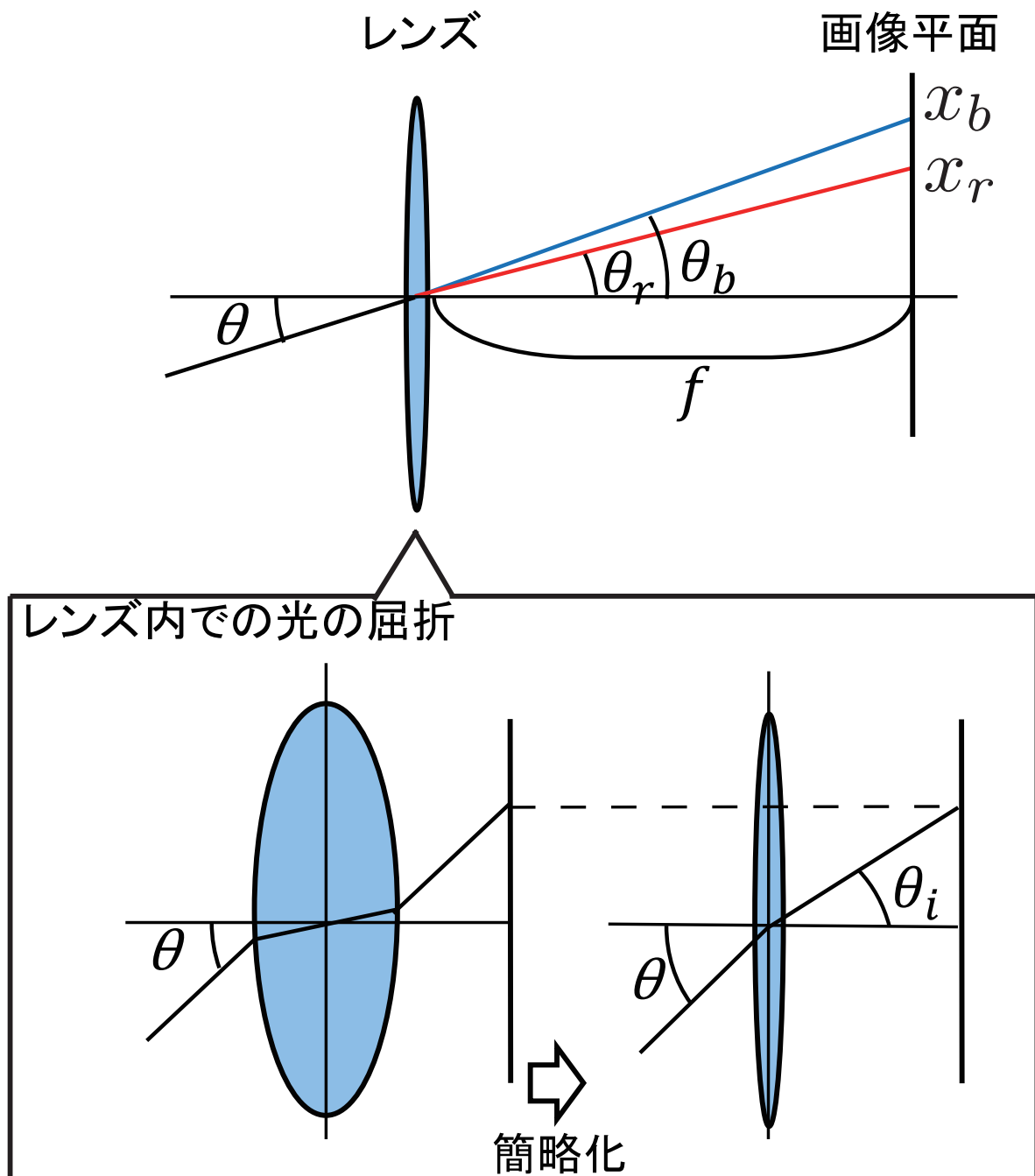


図 4.2 レンズ内での光の屈折を簡略化した倍率色収差の概略図

(iv) 倍率色収差の補正 ステップ (iii) により求めた α_{nm} を用いることで、波長帯域 n の画像を基準画像として考え、波長帯域 m の画像を拡大・縮小の画像変形を用いて正規化することができる。本論文では、拡大・縮小による画像変形を小さくするため、基準画像を中心の波長帯域の画像とした。実際に撮影したチェッカーボードのマルチバンド画像を用いて色収差を補正した結果を図 4.3 に示す。補正前に見えていた、赤や青のラインが消

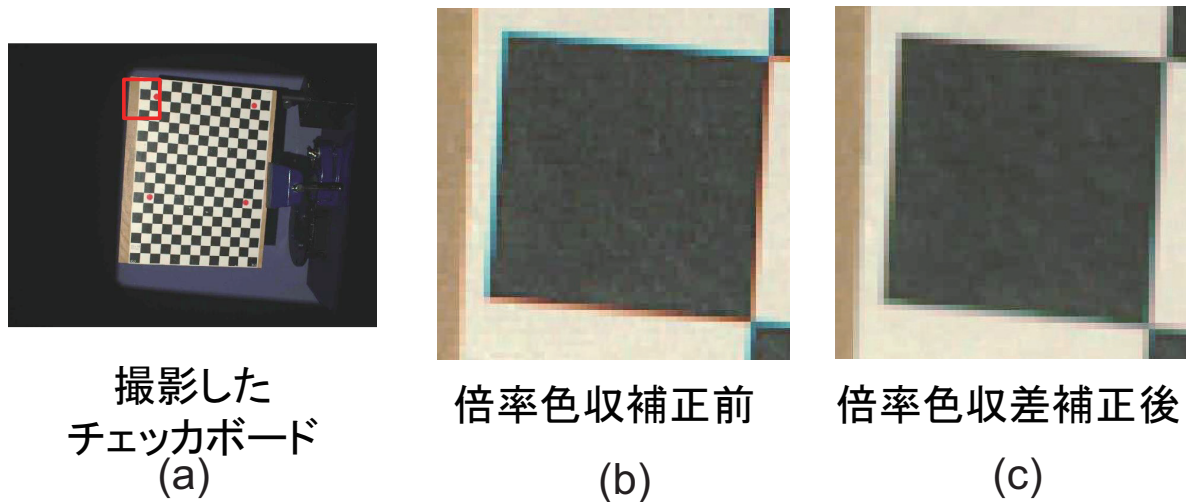


図 4.3 倍率色収差の補正結果：(a) 撮影したマルチバンド画像（色再現結果），
 (b) 色収差補正前（(a) の赤枠部），(c) 色収差補正後（(a) の赤枠部）

え，色ズレが低減されたのが確認できる．

4.3 適切な画像の選択

本節では，物体の形状にあわせて，色再現に適切な画像を選ぶ方法を述べる．ランバートモデルは，物体の拡散反射を 3 次元的な光反射として記述したモデルである [67]．ランバートモデルによれば，光が物体表面で完全に拡散しているとするとき，物体の反射光の強さは，入射光の強さと光の入射角の 2 つの要素で決まる．物体にある光を照射し，その反射光をカメラで観測した場合，分光反射率が同じ値であっても，物体表面の法線の向きが異なれば，カメラで記録される信号値の大きさや，そこから推定される分光反射率に差が生じる．そのため，物体の形状を考慮して色再現に用いる画像を選択する必要がある．本論文では，物体の 3 次元形状を同時に復元するため，カメラの位置や物体の大まかな形状が既知となっている．また，本論文における物体の全周囲のマルチバンド画像は，カメラ，照明の位置を変えずに，対象物を回転させながら撮影される．そのため，物体のある表面の色再現を考えたとき，その物体表面の法線とカメラの視線ベクトルが，なるべく正対した画像を選択して色再現に用いることで，形状を考慮した色再現を行う．

本論文では，復元した密な 3 次元点群の色を再現する際，3 章の (iv) 最適な画像選択に基づく密な 3 次元復元のステップで求めた基準画像 I_{ref} を利用する．復元した点を，それぞれの 3 次元復元に用いられた基準画像 I_{ref} に投影して，各マルチバンド画像の投影点の輝度値を用いることで，色再現を行う．

4.4 分光反射率の推定

本節では、マルチバンド画像を用いた分光反射率推定手法について述べる。本論文では Tsuchida らの手法 [3] を用いる。照明光のスペクトルを $E(\lambda)$ 、物体の分光反射率を $f(\lambda)$ 、物体を撮影するカメラのバンド数を N 、イメージセンサの第 k 番目の分光感度を $S_k(\lambda)$ とすると、得られる画像の信号値 g_k は、次式で表される。

$$g_k = \int S_k(\lambda)E(\lambda)f(\lambda)d\lambda \quad (4.5)$$

$$k = 1, \dots, N$$

上式では、画像上のある 1 点のみを考えるものとして、画像座標を省略している。 $S_k(\lambda)E(\lambda) = h_k(\lambda)$ とすると、式 (4.5) は、以下のように表すことができる。

$$g_k = \int h_k(\lambda)f(\lambda)d\lambda \quad (4.6)$$

$h_k(\lambda)$ および $f(\lambda)$ を細かい波長間隔で離散化すると、式 (4.6) は、以下のようなベクトル表現になる。

$$\mathbf{g} = \mathbf{H}\mathbf{f} \quad (4.7)$$

ここで、 \mathbf{H} は、 $[H_{ki}] = h_k(\lambda_i)$ を要素とする $N \times L$ の行列であり、 $\mathbf{f} = (f(\lambda_1), \dots, f(\lambda_i), \dots, f(\lambda_L))$ 、 $\mathbf{g} = (g_1, \dots, g_N)^t$ である。 i は波長を表す添字であり、 L は波長方向のサンプリング数を表す (380nm~780nm を 1nm 刻みでサンプリングしたとき、 $L = 401$ となる)。分光反射率 f の次元数は 401 次元となるのに対して、画像信号 g の次元は、マルチバンド画像のバンド数となるため、3~20 次元程度となる。そのため、一般的に画像信号 g の次元よりも、分光反射率 f の次元数の方が多くなる。このとき、4.7 から分光反射率 f を推定しようとするとき、解が一意に定まらない状態となる。そこで、解の性質を表す情報 (先見情報) を利用し、何らかの拘束条件を設定することで、解を 1 つに絞る。Wiener 推定 [52] では、分光反射率 f を確率関数として扱い、推定解の誤差のノルムの期待値

$$\Phi = \langle \|\mathbf{f} - \hat{\mathbf{f}}\|^2 \rangle \quad (4.8)$$

が最小となるように解を求める。解は次式で求められる。

$$\hat{\mathbf{f}} = \mathbf{M}\mathbf{g} \quad (4.9)$$

$$\mathbf{M} = \mathbf{R}\mathbf{H}^t (\mathbf{H}\mathbf{R}\mathbf{H}^t)^{-1} \quad (4.10)$$

ここで、 M は H から得られる Wiener 推定行列であり、 R は物体の分光反射率 f の相関行列である。 R を正確に求めることは困難であるため、 f を一次マルコフモデルと仮定して、相関行列 R を次式で定義する。

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{L-1} \\ \rho & 1 & \rho & \dots & \rho^{L-2} \\ \rho^2 & \rho & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{L-1} & \rho^{L-2} & \dots & \dots & 1 \end{bmatrix} \quad (4.11)$$

上式で表された R を用いることで物体の分光反射率 f が推定できる。 ρ は、隣合う波長の要素間の相関（分光反射率特性の滑らかさに相当） $\langle f(\lambda_j), f(\lambda_{j+1}) \rangle$ を表し、本論文では、 $\rho = 0.999$ を用いた。また、推定された分光反射率、観察用の照明のスペクトルパワー分布、トーン曲線およびディスプレイモニタの原色の色度値を用いて、色再現された RGB 信号を得ることができる。

4.5 むすび

本章では、多視点のマルチバンド画像から画像上の各ピクセルにおける分光反射率を推定する手法を述べた。まず、マルチバンド画像の色収差の補正法について述べた。そして、色再現を行う際の適切な画像選択について述べ、マルチバンド画像を用いた分光反射率の推定法について述べた。

第 5 章

性能評価実験

5.1 まえがき

本章では、まず、性能評価に用いる多視点マルチバンド画像の撮影法について述べる。多視点マルチバンド画像を用いた 3 次元復元の性能評価実験を通して、3,4 章で述べた提案手法の有効性を確認する。

5.2 データセット作成

提案手法の有効性を確認するために、図 5.1 に示すような物体の全周囲のマルチバンド画像を撮影するシステムを構築した。撮影は他の光源の影響をなくすため、暗室の内部で行う。本システムは、回転テーブル、光源、カメラの制御を PC で行うため、自動撮影が可能である。被写体を回転テーブルに設置し、1 度回転するごとに光源から射出する光の波長を変えて画像を撮影する。本実験では、光の中心波長を 450, 490, 530, 570, 610, 650, 690, 730nm に変えて 8 バンドの画像を撮影した。照明光スペクトルは、分光放射照度計 CL-500A (コニカミノルタ) を用いて計測した。また、カメラ感度はモノクロメータ、標準白色板、分光放射計 SR-3 (TOPCON) を用いて計測した。さらに、3 次元形状計測の精度を評価するために、対象物体の 3 次元形状を 3 次元スキャナ VIVID910 (コニカミノルタ) で計測し、真値の 3 次元メッシュモデルとした。撮影した全周囲の多視点マルチバンド画像と、計測した真値モデルをあわせてデータセットとした。従来手法による比較実験を行うために、射出する光を白色光にして RGB 画像も同様に撮影した。カメラは、マルチバンド画像撮影には Point gray 社製の FL3-U3-13E4C-C, RGB 画像撮影には Point Grey Research 社製の FL3-U3-13Y3C-C を使用した。レンズには μ -TRON 社製の FV1022 を使用した。撮影された画像の大きさは、1,280 × 1,024 pixel である。撮影対象として、鮮やかな着物を着た 2 体の人形と猫の置物を選択した。撮影対象と VIVID910 で計測した真値の 3 次元メッシュモデルを図 5.2 に示す。撮影したマル

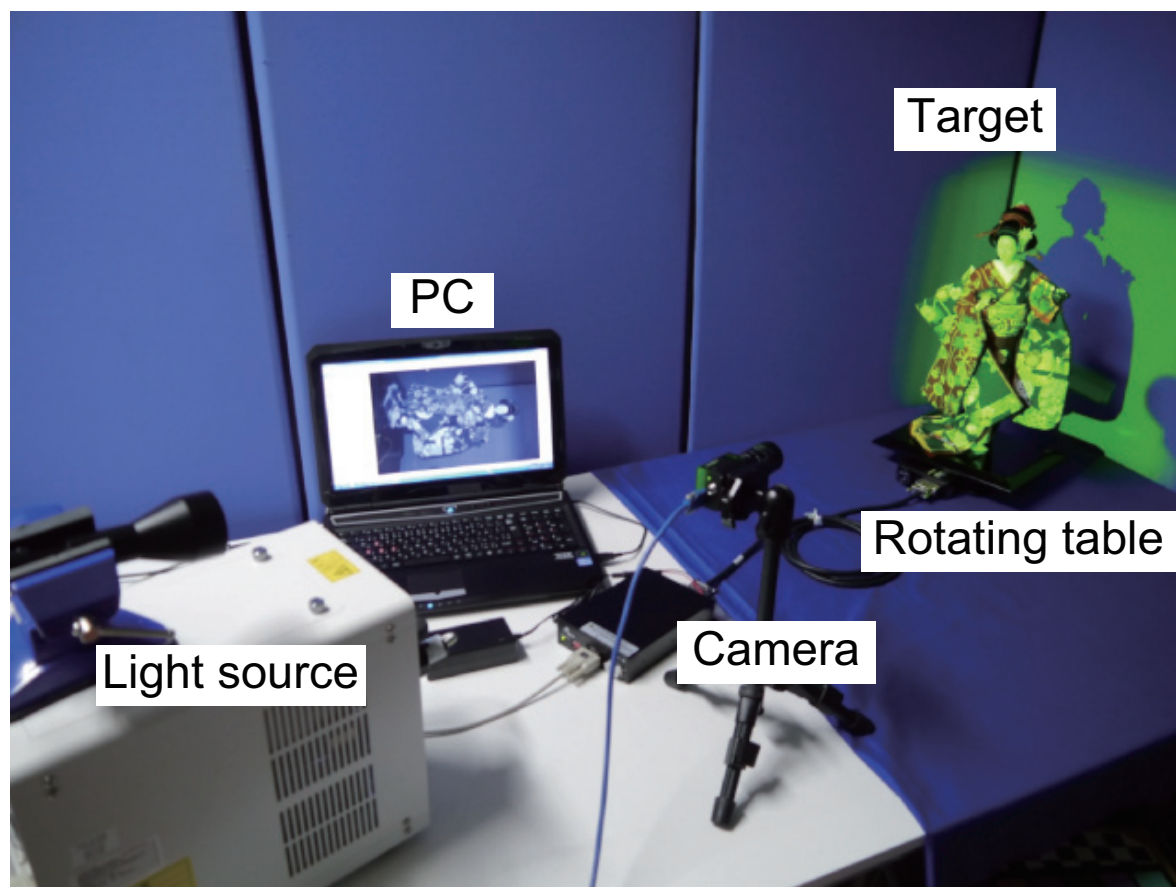


図 5.1 構築した全周囲マルチバンド画像撮影システム



図 5.2 撮影対象と計測した真値の 3 次元モデル

チバンド画像の例を図 5.3 に示す。

5.3 評価実験

従来の多視点画像からの 3 次元復元との比較を通して、提案手法で復元される 3 次元点群の精度を評価する。本論文では、従来手法として、Scale Invariant Feature

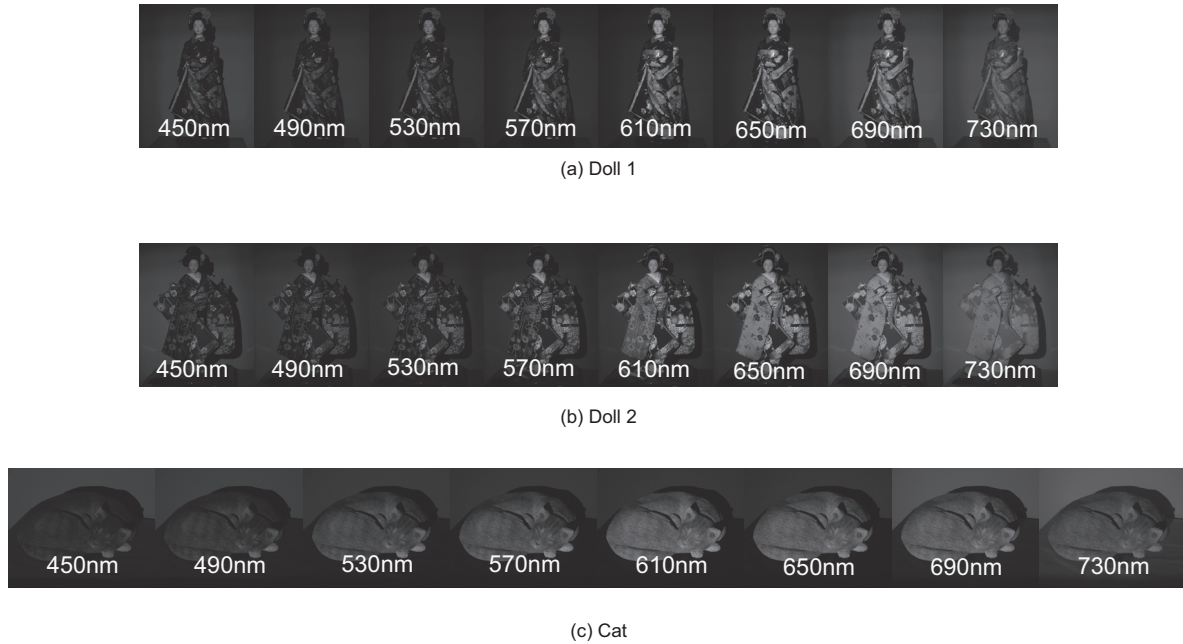


図 5.3 撮影したマルチバンド画像：(a) Doll 1, (b) Doll 2, (c) Cat

Transform(SIFT) [68] に基づく SfM によるカメラ位置推定と，Furukawa らの Patch-based Multi-view Stereo(PMVS)[48] に基づく密な復元を行う．さらにもうひとつの従来手法として，山尾らの手法 [12] による密な復元を行う．従来手法では，撮影したカラー画像を用いて復元を行い，提案手法では，撮影したマルチバンド画像を用いて復元を行う．まず，復元精度の評価を行う．復元精度を評価する指標として，文献 [69] で用いられている $D_{accuracy}$ と $D_{complete}$ を用いる． $D_{accuracy}$ は，3次元計測点群 P_{data} の各3次元点から最近傍の真値の3次元点までの距離集合で定義され，復元点の精度の指標である． $D_{complete}$ は，真値の3次元点群 P_{truth} の各3次元点から最近傍の3次元計測点までの距離集合で定義され，復元された範囲の指標である．なお，本実験では，復元誤差が 5mm 以下の点のみを対象として $D_{accuracy}$ と $D_{complete}$ を計算した．図 5.4 に，それぞれの手法で復元された結果を示す．表 5.1 に復元結果から求めた $D_{accuracy}$ と $D_{complete}$ の RMS [mm] と，復元された3次元点の数を示す．表 5.1 より，従来手法である VisualSfM+PMVS の精度が最も性能が低い．また，山尾らの手法と提案手法を比較すると，提案手法の方が高精度な復元ができているのがわかる．復元された3次元点群の数は同程度であった．これは，マルチバンド画像のための POC を用いることで，特徴点の追跡や密な復元の精度が向上したためである．

次に，分光反射率の推定精度の評価を行う．提案手法による画像選択の有効性を確認するために，(i) 提案手法で選択したマルチバンド画像 I_{ref} ，提案手法で選択した RGB 画像

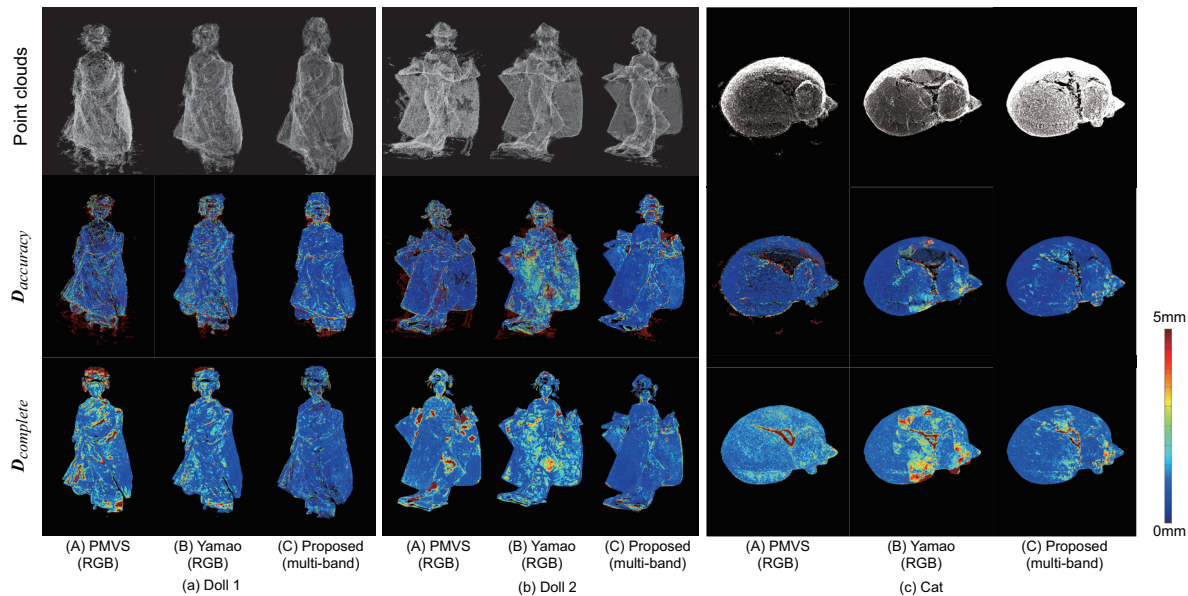


図 5.4 復元した 3 次元点群 (上段) と $D_{accuracy}$ (中段) と $D_{complete}$ (下段): (a)Doll 1, (b)Doll 2, (c)Cat

画像 $I_{ref}RGB$ (3 バンドのマルチバンド画像とみなす), (ii) 画像 I_{ref} から 50 フレーム後のマルチバンド画像, (iii) 画像 I_{ref} より 50 フレーム前のマルチバンド画像, の 4 パターンの画像を用いて, 分光反射率を推定する. 真値として, 分光放射計 SR-3 (TOPCON) により実際に人形を計測した値を用いる. 図 5.5 に, 実際の比較を行う人形の部位を示す. 図 5.6 に, s 推定した分光反射率と計測した分光反射率を示す. 図 5.6 から, 提案手法により選択した画像を用いることで, 高精度に分光反射率が推定できているのが確認できる. またマルチバンド画像と RGB 画像を比較しても, マルチバンド画像が高精度に分光反射率が推定できているのが確認できる. また, 図 5.7 に提案手法により色再現した 3 次元点群を示す.

5.4 むすび

本章では, まず, 性能評価に用いる多視点マルチバンド画像の撮影法について述べた. 多視点マルチバンド画像を用いた 3 次元復元の性能評価実験を通して, 3,4 章で述べた提案手法の有効性を確認した.

表 5.1 $D_{accuracy}$ と $D_{complete}$ の RMS [mm] と復元点数 : (A) VisualSFM+PMVS , (B) 山尾らの手法 , (C) 提案手法

		(A)	(B)	(C)
Doll 1	$D_{accuracy}$	1.58	1.53	1.30
	$D_{complete}$	1.61	1.22	1.17
	# of 3D points	78,731	236,769	214,119
Doll 2	$D_{accuracy}$	1.37	1.74	1.32
	$D_{complete}$	1.90	1.34	1.15
	# of 3D points	109,909	308,441	232,417
Cat	$D_{accuracy}$	0.80	0.76	0.70
	$D_{complete}$	0.99	1.13	0.84
	# of 3D points	82,821	89,349	220,899



(a) Doll1

(b) Doll2

図 5.5 計測を行った部位 (赤枠内): (a) Doll 1 , (b) Doll 2

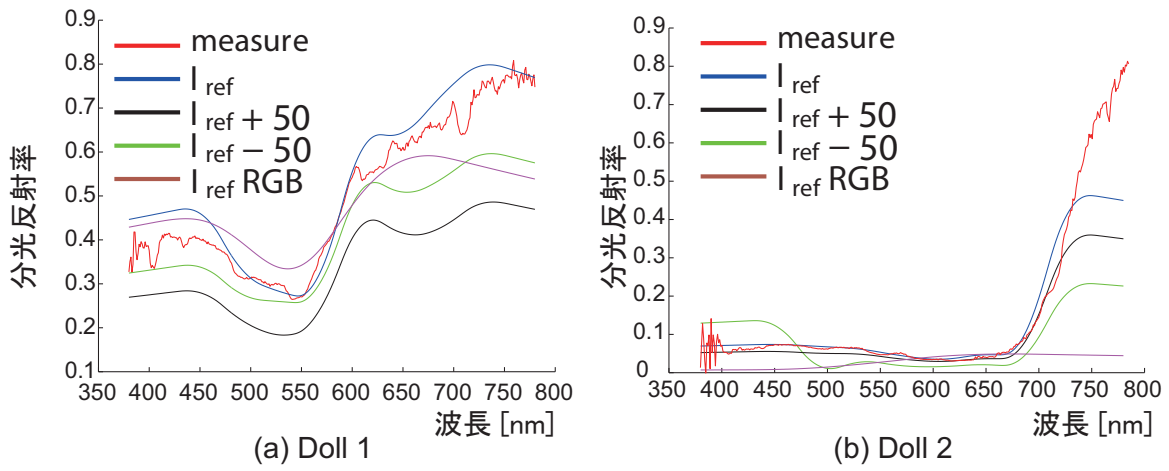


図 5.6 推定した分光反射率と計測した分光反射率：(a) Doll 1 , (b) Doll 2



図 5.7 提案手法により色再現した 3 次元点群：(a) Doll 1 , (b) Doll 2, (c)Cat

第 6 章

結言

本論文では，多視点マルチバンド画像を用いた 3 次元復元および色再現手法を提案した．実際に撮影した多視点マルチバンド画像を用いた性能評価実験を通して，提案手法は，従来手法に比べて高精度な 3 次元復元が可能であることを示した．

以下に各章の内容を要約して述べる．

第 1 章においては，本論文の背景と目的，およびその概要を述べた．

第 2 章においては，多視点 3 次元復元とマルチバンド画像を用いた色再現に関する基礎的考察を行った．まず，多視点画像からの 3 次元復元における 3 つのステップについて述べた．次に，詳細な 3 次元形状の復元に用いられる MVS 手法の分類とそれぞれの特徴について述べた．また，マルチバンド画像を用いた色再現の概要について述べた．次に，マルチバンド画像を用いた色再現 4 つのステップについて述べた．

第 3 章においては，多視点マルチバンド画像を用いた 3 次元復元について述べた．まず，多視点画像間のマッチングに用いる位相限定相関法について述べた．次に，最適な画像選択に基づく多視点マルチバンド画像からの 3 次元復元の流れと，その具体的な処理について述べた．

第 4 章においては，多視点マルチバンド画像を用いた色再現について述べた．まず，マルチバンド画像における色収差の補正法について述べた．次に，色再現を行う際の適切な画像選択について述べ，マルチバンド画像を用いた分光反射率の推定法について述べた．

第 5 章においては，マルチバンド画像を用いて性能評価実験を行い，提案手法の有効性を実証した．まず，性能評価実験に用いるマルチバンド画像のデータセットの作成法について述べた．次に，実際に撮影したマルチバンド画像を用いた性能評価実験を行い，提案手法の有効性を実証した．

今後の展望としては，カメラに取り付けたフィルタを変えつつ移動撮影したマルチバンド画像からの計測システム・アルゴリズムの開発が考えられる．

参考文献

- [1] Katsushi Ikeuchi, Takeshi Oishi, Jun Takamatsu, Ryusuke Sagawa, Atsushi Nakazawa, Ryo Kurazume, Ko Nishino, Mawo Kamakura, and Yasuhide Okamoto, “The great buddha project: Digitally archiving, restoring, and analyzing cultural heritage objects,” *International Journal of Computer Vision*, Vol. 75, No. 1, pp. 189–208, 2007.
- [2] Anna Pelagotti, Andrea Del Mastio, Alessia De Rosa, and Alessandro Piva, “Multispectral imaging of paintings,” *IEEE Signal Processing Magazine*, Vol. 25, No. 4, pp. 27–36, 2008.
- [3] M. Tsuchida, S. Sakai, M. Miura, K. Ito, K. Kawanishi, K. Kashino, J. Yamato, and T. Aoki, “Stereo one-shot six-band camera system for accurate color reproduction,” *Journal of Electronic Imaging*, Vol. 23, No. 3, pp. 033025–1–033025–12, Sept. 2013.
- [4] 三浦武, 海賀孝明, 柴田傑, 桂博章, 田島克文, 玉本英夫, *et al.*, 秋田県の民俗芸能研究におけるモーションキャプチャデータの活用. PhD thesis, Muroran Institute of Technology, 2015.
- [5] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer-Verlag New York Inc., 2010.
- [6] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, “A comparison and evaluation of multi-views stereo reconstruction algorithms,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 519–528, June 2006.
- [7] C. Strecha, W. von Hansen, L. V. Gool, P. Fua, and U. Thoennessen, “On benchmarking camera calibration and multi-view stereo for high resolution imagery,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
- [8] Y. Furukawa, “Multi-view 3D reconstruction techniques,” *研究報告コンピュータビジョンとイメージメディア*, Vol. 176, No. 12, pp. 1–9, Mar. 2011.
- [9] Norimichi Tsumura, Hideaki Haneishi, and Yoichi Miyake, “Estimation of spectral reflectance from multi-band images by multiple regression analysis,”

- Japanese Journal of Optics*, Vol. 27, No. 7, pp. 384–391, 1998.
- [10] Peter D Burns and Roy S Berns, “Analysis multispectral image capture,” in *Color and Imaging Conference*, vol. 1996, pp. 19–22, Society for Imaging Science and Technology, 1996.
- [11] Hideaki Haneishi, Takayuki Hasegawa, Asako Hosoi, Yasuaki Yokoyama, Norimichi Tsumura, and Yoichi Miyake, “System design for accurately estimating the spectral reflectance of art paintings,” *Applied Optics*, Vol. 39, No. 35, pp. 6621–6632, 2000.
- [12] 山尾創輔, 酒井修二, 伊藤康一, 青木孝文, “動画像からの高精度・高密度な3次元点群の復元に関する検討,” 情報処理学会 CVIM 研究会, Vol. 187, No. 17, pp. 1–8, 5 2013.
- [13] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, “High-accuracy subpixel image registration based on phase-only correlation,” *IEICE Trans. Fundamentals*, Vol. E86-A, No. 8, pp. 1925–1934, Aug. 2003.
- [14] K. Takita, M. A. Muquit, T. Aoki, and T. Higuchi, “A sub-pixel correspondence search for computer vision applications,” *IEICE Trans. Fundamentals*, Vol. E87-A, No. 8, pp. 1913–1923, Aug. 2004.
- [15] M. Tsuchida, et al, “Efficient poc-based correspondence detection method for multi-channel images,” *Proc. Color and Imaging Conference*, pp. 113–118, 2014.
- [16] R. Hartley and A. Zisserman, *Multiple View Geometry*. Cambridge University Press, 2004.
- [17] 酒井修二, 高橋徹, 伊藤康一, 青木孝文, 運天弘樹, “汎用デジタルカメラを用いた2視点からの3次元復元,” 研究報告コンピュータビジョンとイメージメディア, Vol. 176, No. 4, pp. 1–8, Mar. 2011.
- [18] 三浦衛, 酒井修二, 石井純平, 伊藤康一, 青木孝文, “汎用デジタルカメラを用いた2視点からの3次元復元システム,” 第14回画像の認識・理解シンポジウム, Vol. 14, No. DS-6, pp. 1707–1708, July 2011.
- [19] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, “Multi-view stereo for community photo collections,” *Proc. Int’l Conf. Computer Vision*, pp. 1–8, Oct. 2007.
- [20] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski, “Building Rome in a day,” *Proc. Int’l Conf. Computer Vision*, pp. 72–79, Oct. 2009.
- [21] Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski, “Towards internet-scale multi-view stereo,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 1434–1441, June 2010.

- [22] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int’l J. Comput. Vision*, Vol. 60, No. 2, pp. 91–110, Nov. 2004.
- [23] H. Bay, T. Tuytelaars, and L. V. Gool, “SURF: Speeded up robust features,” *Proc. European Conf. Computer Vision*, pp. 404–417, May 2006.
- [24] 岡谷貴之, “バンドルアジャストメント,” 研究報告コンピュータビジョンとイメージメディア, Vol. 167, No. 37, pp. 1–16, June 2009.
- [25] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon, “Bundle adjustment a modern synthesis,” in *Vision algorithms: theory and practice*, pp. 298–372, Springer, 2000.
- [26] M. I. A. Lourakis and A. A. Argyros, “SBA: A software package for generic sparse bundle adjustment,” *ACM Trans. Math. Software*, Vol. 36, No. 1, pp. 1–30, Mar. 2009.
- [27] 徐剛 辻三郎, 3次元ビジョン. 共立出版, 1998.
- [28] D. Nistér, “An efficient solution to the five-point relative pose problem,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 756–770, 2004.
- [29] L. Kneip, D. Scaramuzza, and R. Siegwart, “A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 2969–2976, 2011.
- [30] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. ACM*, Vol. 24, No. 6, pp. 381–395, 1981.
- [31] S. Tran and L. Davis, “3D surface reconstruction using graph cuts with surface constraints,” *Proc. European Conf. Computer Vision*, Vol. 2, pp. 219–231, May 2006.
- [32] G. Vogiatzis, P. H. Torr, and R. Cipolla, “Multi-view stereo via volumetric graph-cuts and occlusion robust photo-consistency,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 29, No. 12, pp. 2241–2246, Dec. 2007.
- [33] S. N. Sinha, P. Mordohai, and M. Pollefeys, “Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh,” *Proc. Int’l Conf. Computer Vision*, pp. 1–8, Oct. 2007.
- [34] A Laurentini, “The visual hull concept for silhouette-based image understanding,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 2, pp. 150–162, Feb. 1994.

- [35] A. Zaharescu, E. Boyer, and R. Horaud, “Transformesh: A topology-adaptive mesh-based approach to surface evolution,” *Proc. Asian Conf. Computer Vision*, pp. 166–175, Nov. 2007.
- [36] A. Delaunoy, E. Praos, P. Gargallo, J. P. Pons, and P. Sturm, “Minimizing the multi-view stereo reprojection error for triangular surface meshes,” *Proc. British Machine Vision Conf.*, Sept. 2008.
- [37] V. H. Hiep, R. Keriven, P. Labatut, and J. P. Pons, “Towards high-resolution large-scale multi-view stereo,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 1430–1437, June 2009.
- [38] M. Goesele, B. Curless, and S. M. Seitz, “Multi-view stereo revisited,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 2402–2409, June 2006.
- [39] C. Strecha, R. Fransens, and L. V. Gool, “Combined depth and outlier estimation in multi-view stereo,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 2394–2401, June 2006.
- [40] D. Bradley, T. Boubekeur, and W. Heidrich, “Accurate multi-view reconstruction using robust binocular stereo and surface meshing,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
- [41] N. D. F. Campbell, G. Vogiatzis, C. Hernández, and R. Cipolla, “Using multiple hypotheses to improve depth-maps for multi-view stereo,” *Proc. European Conf. Computer Vision*, pp. 766–779, Oct. 2008.
- [42] J. Li, E. Li, Y. Chen, L. Xu, and Y. Zhang, “Bundle depth-map merging for multi-view stereo,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 2569–2776, June 2010.
- [43] E. Tola, C. Strecha, and P. Fua, “Efficient large-scale multi-view stereo for ultra high-resolution image sets,” *Machine Vision and Applications*, Vol. 23, No. 5, pp. 903–920, Sept. 2012.
- [44] B. Curless and M. Levoy, “A volumetric method for building complex models from range images,” *Proc. SIGGRAPH ’96*, pp. 303–312, Aug. 1996.
- [45] M. Kazhdan, M. Bolitho, and H. Hoppe, “Poisson surface reconstruction,” *Proc. Symp. Geometry Processing*, pp. 61–70, June 2006.
- [46] M. Lhuillier and L. Quan, “A quasi-dense approach to surface reconstruction from uncalibrated images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, No. 3, pp. 418–433, Mar. 2005.
- [47] M. Habbeck and L. Kobbelt, “A surface-growing approach to multi-view stereo reconstruction,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*,

- pp. 1–8, June 2007.
- [48] Y. Furukawa and J. Ponce, “Accurate, dense, and robust multiview stereopsis,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 32, No. 8, pp. 1362–1376, Aug. 2010.
- [49] 三宅洋一, 富永昌治, 羽石秀昭, 山口雅浩, 津村徳道, “分光画像処理入門,” 分光画像処理入門, pp. 95–166, 2006.
- [50] 加藤直哉, “4色ccdカメラおよびその色再現・色管理技術,” 日本写真学会誌, Vol. 67, No. Supplement1, pp. 14–16, 2004.
- [51] Michael J Vrhel and HJ Trussell, “Color correction using principal components,” *Color Research & Application*, Vol. 17, No. 5, pp. 328–338, 1992.
- [52] William K Pratt and Clanton E Mancill, “Spectral estimation techniques for the spectral calibration of a color image scanner,” *Applied Optics*, Vol. 15, No. 1, pp. 73–75, 1976.
- [53] Hideaki Haneishi, “Design of color filters for recording art works,” in *Proceedings of IS&T’s 50th Annual Conference, 1997*, pp. 360–372, IS&T, 1997.
- [54] A. Wendel, M. Maurer, G. Graber, T. Pock, and H. Bischof, “Dense reconstruction on-the-fly,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 1450–1457, 2012.
- [55] M. Pollefeys, D. Nistér, J. M. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S. J. Kim, P. Merrell, C. others Salmi, S. Sinha, B. Talton, L. Wang, Q. Yang, H. Stewénus, R. Yang, G. Welch, and H. Towles, “Detailed real-time urban 3d reconstruction from video,” *Int’l Journal of Computer Vision*, Vol. 78, No. 2-3, pp. 143–167, 2008.
- [56] G. Vogiatzis and C. Hernández, “Video-based, real-time multi-view stereo,” *Image and Vision Computing*, Vol. 29, No. 7, pp. 434–441, 2011.
- [57] P. Labatut, J.P. Pons, and R. Keriven, “Efficient multi-view reconstruction of large-scale scenes using interest points, delaunay triangulation and graph cuts,” in *Proc. Int’l Conf. Computer Vision*, pp. 1–8, 2007.
- [58] P. Labatut, J.P. Pons, and R. Keriven, “Robust and efficient surface reconstruction from range data,” in *Computer Graphics Forum*, vol. 28, pp. 2275–2290, Wiley Online Library, 2009.
- [59] J. Shi and C. Tomasi, “Good features to track,” in *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 593–600, 1994.
- [60] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Proc. Alvey Vision Conf.*, vol. 15, p. 50, 1988.

- [61] E. Rosten and T. Drummond, “Machine learning for high-speed corner detection,” *Proc. European Conf. Computer Vision*, pp. 430–443, 2006.
- [62] R.L. Cook, “Stochastic sampling in computer graphics,” *ACM Trans. Graphics*, Vol. 5, No. 1, pp. 51–72, Jan. 1986.
- [63] Oscar Pizarro, Ryan Eustice, and Hanumant Singh, “Relative pose estimation for instrumented, calibrated imaging platforms,” in *DICTA*, pp. 601–612, Citeseer, 2003.
- [64] J. Ishii, S. Sakai, K. Ito, and T. Aoki, “Wide-baseline stereo matching using ASIFT and POC,” *Proc. IEEE Int’l Conf. Image Processing*, pp. 2977–2980, Sept. 2012.
- [65] J. Havel and A. Herout, “Yet faster ray-triangle intersection (using SSE4),” *Trans. Visualization and Computer Graphics*, Vol. 16, No. 3, pp. 434–438, 2010.
- [66] Micah K Johnson and Hany Farid, “Exposing digital forgeries through chromatic aberration,” in *Proceedings of the 8th workshop on Multimedia and security*, pp. 48–55, ACM, 2006.
- [67] Alan Watt, *Fundamentals of three-dimensional computer graphics*. Addison-Wesley, 1989.
- [68] D. G. Lowe, “Demo Software: SIFT Keypoint Detector.” <http://www.cs.ubc.ca/~lowe/keypoints/>.
- [69] R. Jensen, A. Dahl, G. Vogiatzis, E. Tola, and H. Anæs, “Large scale multi-view stereopsis evaluation,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 406–413, June 2014.

謝辞

本論文は、著者が東北大学 大学院情報科学研究科 情報基礎科学専攻 計算機構論分野（青木研究室）において行った研究を取りまとめたものであります。本研究を推し進めるにあたり、多くの方々からご協力とご助言を頂きました。

恩師青木孝文教授には、研究に対する懇切なるご指導と終始変わらぬ励ましを頂きました。研究だけでなく、何事に対しても情熱的に取り組む先生のご姿勢から、非常に多くのことを学ばせて頂きました。先生のもとで研究を行えたことは、私にとって大変な幸運であり、今後の人生の大きな糧になると確信しております。ここに改めて、深く感謝の意を表します。

本論文をまとめるにあたり、鈴木陽一教授ならびに大町真一郎教授より、それぞれの専門の立場から有意義なご意見を賜りました。ここに深く感謝いたします。

NTT コミュニケーション基礎科学研究所の土田勝博士には、マルチバンド画像からの色再現に関するご討論とご助言を頂きました。ここに改めて深く感謝いたします。

コニカミノルタの近藤敏志博士ならびに大宮淳氏には、カメラ位置推定ならびに、3次元超音波像の構成に関するご討論とご助言を頂きました。ここに改めて深く感謝いたします。

パナソニックの日下博也博士ならびに松延徹氏には、多視点画像を用いた屋外環境の3次元復元に関するご討論とご助言を頂きました。ここに改めて深く感謝いたします。

伊藤康一助教には、研究全般に関する直接的なご指導と活発なご討論を頂くとともに、本論文の執筆においても様々なご助言を頂きました。ここに改めて深く感謝いたします。

本間尚文教授には、本研究を進めるにあたり、様々なご討論とご助言を頂きました。ここに改めて深く感謝いたします。

研究室諸氏には、研究室生活において様々な面でご協力頂くとともに、本研究に関する様々なご意見やご協力をいただきました。皆様には改めて感謝の意を表すとともに、以下に名前を挙げさせていただきます。Luis Rafael Marval Pérez 氏、立見駿介氏、石幡大輔氏、河井航氏、鈴木遼雅氏、李星氏、菱沼将太氏、藤本竜一氏。

このように、本研究を進めるにあたり、多くの方々からご協力とご助言を頂きました。私一人の力では、到底このような形で研究を取りまとめ、論文として著すことはできませ

んでした。諸氏に重ねて厚くお礼申し上げます。

2017年2月10日