

# 位相限定相関法に基づく多視点画像からの3次元復元に関する研究

著者	酒井 修二
学位授与機関	Tohoku University
URL	<a href="http://hdl.handle.net/10097/53959">http://hdl.handle.net/10097/53959</a>

# 修士論文

## 位相限定相関法に基づく多視点画像からの 3次元復元に関する研究

情報基礎科学専攻

酒井 修二

3D Reconstruction from Multi-View Stereo  
Based on Phase-Only Correlation

by

Shuji SAKAI

Graduate School of Information Sciences, Tohoku University, 2012

Copyright © 2012 Shuji SAKAI,

All Rights Reserved.

# 目次

第 1 章	緒言	3
第 2 章	多視点画像からの 3 次元復元に関する基礎的考察	7
2.1	まえがき	7
2.2	多視点画像の撮影とカメラキャリブレーション	7
2.2.1	多視点画像の撮影	8
2.2.2	カメラキャリブレーション	9
2.3	多視点画像からの 3 次元復元の分類と特徴	10
2.4	デプスマップ統合に基づく 3 次元復元	13
2.5	NCC を用いたプレーンスイーピングに基づくデプスマップ生成	15
2.6	むすび	16
第 3 章	1 次元位相限定相関法に基づく画像マッチング	17
3.1	まえがき	17
3.2	1 次元 POC に基づく画像マッチング	17
3.3	1 次元 POC に基づく画像マッチングの高精度化	20
3.3.1	窓関数の適用	20
3.3.2	スペクトルの重み付け	20
3.3.3	ピークモデルの関数フィッティングによるパラメータ推定	21
3.3.4	複数の 1 次元画像信号の利用	21
3.3.5	ウィンドウの拡大縮小によるステレオ画像間の画像変形の軽減	21
3.4	むすび	22
第 4 章	位相限定相関法に基づく多視点画像からの 3 次元復元	23
4.1	まえがき	23
4.2	POC に基づく多視点画像からのデプスマップ生成	24
4.2.1	視差の正規化	24

## 目次

4.2.2	正規化視差に基づく POC 関数の統合 . . . . .	26
4.2.3	POC に基づく多視点画像からのデプスマップ生成手法 . . . . .	29
4.3	提案手法の性能評価 . . . . .	31
4.3.1	実験方法 . . . . .	32
4.3.2	3次元復元精度の評価 . . . . .	34
4.3.3	計算コストの評価 . . . . .	44
4.3.4	むすび . . . . .	47
第5章	結言 . . . . .	48
付録 A	Scaled Window-POC に基づく画像変形にロバストな高精度対応付け手法 . . . . .	50
A.1	まえがき . . . . .	50
A.2	SW-POC に基づく対応付け手法 . . . . .	51
A.2.1	Scaled Window-POC (SW-POC) . . . . .	51
A.2.2	拡大縮小率の推定 . . . . .	52
A.2.3	粗密戦略を用いた計算コストの削減 . . . . .	54
A.2.4	物体の形状を考慮した平均 POC 関数の計算 . . . . .	54
A.3	性能評価実験 . . . . .	55
A.3.1	実験方法 . . . . .	56
A.3.2	3次元復元精度の評価 . . . . .	58
A.3.3	計算コストの評価 . . . . .	60
A.4	むすび . . . . .	61
謝辞	. . . . .	67

## 第1章

### 緒言

3次元復元技術は、物体の形状や空間の構造を計算機上で復元する技術である。近年、3次元復元技術に対する需要は、産業、医療、芸術などのさまざまな分野において高まりつつある [1], [2]。産業の分野では、エレクトロニクス製品や自動車などの3次元情報による製品検査に用いられる。医療の分野では、外科や歯科において、人体の形状の解析・診断に用いられている。芸術の分野では、文化遺産の保護を目的として、美術品や遺跡などの文化財のデジタルアーカイブに用いられている。

一般に、3次元復元手法は、能動型の手法と受動型の手法の2つに分類することができる [1], [3], [4]。能動型の手法は、物体にレーザや構造光を投影し、その挙動を観察することで物体の3次元形状を取得する [1]。一般に、能動型の手法は、受動型の手法に比べ精度や安定性に優れている。しかし、レーザなどを物体に投影するためのシステムが複雑で高価であることや、撮影装置が大型になりやすく、撮影環境が限定されるといった問題が存在する。特に、大規模な復元を行う場合、大型の撮影装置により撮影を繰り返し行う必要があるため、撮影にかかる負担は大きい。

一方、受動型の3次元復元手法は、カメラの位置や焦点距離などの撮影条件を変化させながら複数枚の画像を撮影し、その挙動を観察することで物体の3次元形状を取得する。受動型の3次元復元手法のなかでも、異なる視点から同一の対象を撮影することで3次元復元を行う多視点画像からの3次元復元は、自由度の高い撮影で高精度な3次元復元が可能である [3]–[7]。多視点画像からの3次元復元では、カメラによる画像取得という非常に簡便な作業により3次元復元が行えるため、撮影にかかる負担が少なく、撮影環境も限定されにくい。また、近距離用のレンズや遠距離用のレンズ、工業用カメラや汎用デジタルカメラなど、さまざまな撮影装置に対して適用可能である。そのため、多視点画像からの3次元復元は、産業における小さな工業部品の3次元復元から、文化財のデジタルアーカイブにおける建築物や都市全体といった大規模な復元まで、広範囲な応用に用いることが

## 第 1 章 緒言

できる．近年，計算機や撮影装置の技術的発展や，3次元復元手法の進歩により，受動型の手法においても，能動型の3次元復元手法に匹敵する高精度な3次元復元が行えるようになった．しかし，安定性や計算コストにおいて，多視点画像からの3次元復元手法の実用化への課題は多い．特に，高精度な3次元復元のために，高解像度の画像や非常に多くの枚数の多視点画像を用いる場合，計算コストの多い手法では，膨大な処理時間を必要とする．これらの課題に対して，本研究では，カメラ画像のみから，少ない計算コストで高精度かつロバストな3次元復元を行うことを目標とする．自由度の高い撮影，および，短い処理時間で，高精度な3次元復元が可能になることにより，多視点画像からの3次元復元は，能動型の3次元復元手法と比べて，より広い範囲の応用における実用化が期待できる．

多視点画像からの3次元復元手法は，(i) 3次元ボリュームの最適化に基づく手法 [8]–[10]，(ii) メッシュの最適化に基づく手法 [11]–[13]，(iii) 特徴領域拡張に基づく手法 [14]–[16]，(iv) デプスマップ統合に基づく手法 [17]–[22] の4つに分類される [5]．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は，高精度な3次元復元が可能である一方で，非常に複雑な最適化を行うため，計算コストが大きく，また，初期値によって問題が収束しないといった問題がある．これに対して，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は，複雑な最適化が必須ではないため，特別な初期値を用いることなく，カメラ画像のみから3次元復元が可能である．中でも，デプスマップ統合に基づく手法は，従来のステレオマッチングの技術を適用することが容易であること，および，復元対象によって復元される点群が特徴点周辺に偏ったりすることがないことから，実用性が高い．本論文では，このデプスマップ統合に基づく手法に着目する．

デプスマップ統合に基づく手法は，(i) 多視点画像から複数のデプスマップを生成し，(ii) 複数のデプスマップからメッシュモデルを生成する2つのステップで構成される．ここで，(i) のステップで生成されるデプスマップの精度は，最終的に生成されるメッシュモデルの精度に大きく影響する．そのため，高精度な3次元復元を行うためにはデプスマップを高精度に求めることが重要である．最も基本的なデプスマップ生成手法として，プレーンスweepingが有名であり，多くのデプスマップ統合に基づく3次元復元手法に適用されている [3], [17]．プレーンスweepingに基づくデプスマップ生成手法は，3次元点の奥行きを変化させながらウィンドウマッチングを繰り返し適用することで，その3次元点の奥行きを決定する．プレーンスweepingに基づくデプスマップ生成手法では，デプスマップの精度が奥行き変化の刻み幅に依存し，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細かくする必要がある．しかし，奥行き変化の刻み幅を細かくすると，マッチング回数が増加するため，処理時間が膨大になるという問題がある．これは，多くのプレーンスweepingに基づく手法がウィンドウマッチングに正規化相互相関 (Normalized

Cross-Correlation: NCC) [23] を用いているため、NCC の相関値を、ある仮定した 3 次元点の奥行きが尤もらしいかどうかの判断にしか利用できていないことに起因する。

そこで、本論文では、位相限定相関法 (Phase-Only Correlation: POC) に基づく多視点画像からのデプスマップ生成手法を提案する [24]。POC は、画像の位相成分に着目した画像マッチング手法であり、そのマッチング精度がきわめて高く、さまざまな応用で有効である [25]–[27]。また、2 つの画像信号から 1 つの相関値を計算する NCC とは異なり、POC は 2 つの画像信号から 1 つの相関関数 (POC 関数) を計算することが特徴としてあげられる。この POC 関数は、理論的なピーク形状を有しており、POC 関数のピークの高さが 2 つの画像信号間の類似度に、POC 関数のピーク座標が 2 つの画像信号間の平行移動量に相当する。ステレオ画像間のウィンドウマッチングのように、ウィンドウ間の平行移動が 1 次元方向に限定される場合、1 次元 POC に基づく画像マッチングが有効である。この 1 次元 POC を多視点画像間のウィンドウマッチングに用いることで、ある仮定した 3 次元点の奥行きが、最も尤もらしい奥行きからどれくらい離れているかを推定することが可能である。提案手法では、ウィンドウマッチングにより得られる 1 次元 POC 関数のピーク座標から直接 3 次元点の奥行きを算出するため、1 回のマッチングでサブピクセルレベルの高精度な奥行きを求めることができる。さらに、(i) 複数のステレオペアから計算される POC 関数の統合、および、(ii) 画像ピラミッドを用いた階層的探索を組み合わせることで、非常に少ないマッチング回数で高精度かつロバストなデプスマップ生成を実現する。これにより、従来の NCC を用いたプレーンスイーピングに比べて、少ない計算コストで高精度な 3 次元復元が行える。本論文では、実際の多視点画像を用いた実験を通して、3 次元復元精度および計算コストの観点から、従来手法と提案手法の比較を行う。

本論文は以上の内容をまとめたものであり、位相限定相関法に基づく多視点画像からの 3 次元復元手法を提案する。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、多視点画像からの 3 次元復元の基礎的な考察を行う。はじめに、多視点画像からの 3 次元復元に用いる多視点画像とカメラパラメータの取得方法について述べる。次に、多視点画像からの 3 次元復元の分類とそれぞれの手法の特徴について述べる。そして、多視点画像からの 3 次元復元手法の 1 つであるデプスマップ統合に基づく手法について説明する。さらに、従来のデプスマップ生成によく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手法の原理を説明する。

第 3 章においては、1 次元 POC に基づく画像マッチング手法について述べる。まず、1 次元 POC の原理について説明する。そして、1 次元 POC 関数を用いて平行移動量推定を行う際に重要となる各種高精度化手法について述べる。



## 第 1 章 緒言

第 4 章においては、位相限定相関法に基づく多視点画像からの 3 次元復元手法を提案する。まず、提案手法で重要となる正規化視差の概念について説明し、正規化視差に基づく POC 関数の統合について述べる。そして、1 次元 POC を用いた画像マッチングに基づく奥行き推定手法と画像ピラミッドを用いた階層的探索を組み合わせることで、POC に基づく多視点画像からのデプスマップ生成手法を提案する。さらに、性能評価実験を通して、提案手法により少ない計算コストで高精度な 3 次元復元が可能であることを示す。

第 5 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

# 多視点画像からの3次元復元に関する基礎的考察

## 2.1 まえがき

多視点画像からの3次元復元は，異なる視点から撮影された複数枚のカメラ画像を用いて物体の表面形状を復元する技術であり，コンピュータビジョンの分野において最も注目されている研究の1つである [3], [5]–[7]．多視点画像からの3次元復元は，レーザや構造光を投影する能動型の3次元復元手法に比べ，安価で簡便なシステムを構築可能である．また，撮影機材がカメラのみであるので，撮影の際の制限が少ないといった利点がある．近年，デジタル画像技術の発展や，デジタルカメラの高精度化，および，さまざまな3次元復元手法の進歩に伴い，カメラ画像のみを用いて高品質な3次元復元が行えるようになった．これにより，多視点画像からの3次元復元は，コンピュータビジョンの研究分野だけでなく，産業や医療，芸術などさまざまな分野において，能動型の3次元復元手法に代わる手法として注目されるようになった．

本章では，多視点画像からの3次元復元に関する基礎的考察を行う．まず，多視点画像の撮影とカメラキャリブレーションについて述べる．次に，多視点画像からの3次元復元の分類とそれぞれの特徴について述べる．そして，多視点画像からの3次元復元手法の1つであるデプスマップ統合に基づく3次元復元について述べる．さらに，最も基本的なデプスマップ生成手法の1つであるプレーンスイーピングに基づくデプスマップ生成について述べる．

## 2.2 多視点画像の撮影とカメラキャリブレーション

本論文では，多視点画像からの3次元復元の入力データとして，異なる視点から同一の物体を撮影した多視点画像と，それら多視点画像におけるカメラパラメータを想定する．

## 第 2 章 多視点画像からの 3 次元復元に関する基礎的考察

本節では、多視点画像からの 3 次元復元を行う前処理として、多視点画像の撮影方法と、カメラパラメータを得るためのカメラキャリブレーションについてまとめる。

### 2.2.1 多視点画像の撮影

多視点画像の撮影には、(i) 固定された多視点カメラを用いる方法、(ii) ロボットアームや回転テーブルを用いる方法、(iii) 単眼カメラによる自由な移動撮影などが用いられる [7]。それぞれの手法について以下にまとめる。

#### (i) 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、3 次元復元対象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 [7]。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、移動物体の 3 次元復元が可能である。一方で、撮影スタジオの大きさなどの理由から、3 次元復元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できない。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。

#### (ii) 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 [5], [7]。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、移動物体の撮影は不可能である。また、事前にキャリブレーションを行う場合、キャリブレーションを行ったときと同じ条件で 3 次元復元対象の撮影を行うために、ロボットアームや回転テーブルを高精度に動かす必要がある。

#### (iii) カメラの移動撮影

多視点カメラやロボットアームなどを用いる撮影方法に対して、より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 [6], [7], [28], [29]。カメ

## 2.2 多視点画像の撮影とカメラキャリブレーション

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 [30]–[32]。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) [33] やSURF (Speeded-Up Robust Features) [34] などの特徴ベースの対応付け手法により安定して多視点画像間の対応付けが可能になったこと、および、バンドルアジャストメント [3], [35], [36] により多視点画像間の対応付け結果のみから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 カメラキャリブレーション

カメラキャリブレーションとは、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作業である [3], [4], [37]。内部パラメータには、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が含まれ、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。外部パラメータには、カメラの回転と平行移動が含まれ、世界座標系の3次元座標におけるカメラの位置関係を表す。カメラキャリブレーションには、(i) 既知形状物体の撮影によるキャリブレーションと (ii) Structure from Motion (SfM) がある。それぞれの手法について以下にまとめる。

#### (i) 既知形状物体の撮影によるキャリブレーション

既知形状物体の撮影によるキャリブレーションでは、あらかじめ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キャリブレーションを行う [3], [37]。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、整列された円などを用いることで、コーナー検出や円検出により、平

## 第 2 章 多視点画像からの 3 次元復元に関する基礎的考察

板や直方体上の 3 次元座標とカメラ画像の対応付けを行う。この手法は、SfM と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、3 次元復元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。

### (ii) Structure from Motion (SfM)

SfM では、3 次元復元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 [3], [37]。多視点画像間の対応付けには、ステレオ画像間の画像変形にロバストな対応付け手法が必要なため、SIFT や SURF などの特徴ベースの対応付け手法 [33], [34] 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、バンドルアジャストメントなどが用いられる [3], [35], [37]。特に、近年では、高精度なカメラパラメータ推定が行えるため、バンドルアジャストメントがよく用いられる。SfM では、3 次元復元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、事前にカメラキャリブレーションが難しい場合においても、カメラキャリブレーションを行うことが可能である。一方で、多視点画像間の対応付けにおいて誤対応が含まれると、キャリブレーション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) [38] 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、SfM によるキャリブレーション結果では、原理的にスケールの不定性があり、3 次元復元結果について大きさを一意に決定することができない。

## 2.3 多視点画像からの 3 次元復元の分類と特徴

多視点画像からの 3 次元復元手法は、(i) 3 次元ボリュームの最適化に基づく手法、(ii) メッシュの最適化に基づく手法、(iii) 特徴領域拡張に基づく手法、(iv) デプスマップ統合に基づく手法の 4 つに大きく分類される [5]。それぞれの手法について、以下にまとめる。

### (i) 3 次元ボリュームの最適化に基づく手法

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

## 2.3 多視点画像からの 3 次元復元の分類と特徴

う [8]–[10] . 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 , あるボクセルが 3 次元復元対象内部かどうかを最適化により求める . また , ボクセルの最適化手法として , グラフカットなどが用いられる .

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 [39] 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 .

### (ii) 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 [11]–[13] . 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るといった問題がある . 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 .

## 第 2 章 多視点画像からの 3 次元復元に関する基礎的考察

### (iii) 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 [14]–[16]．特徴点の 3 次元復元には，SIFT または SURF のような特徴ベースの対応付け手法や，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対応付け手法が用いられる．中でも，Furukawa らのパッチベースの 3 次元復元手法 [16] はよく知られている．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，物体全体の 3 次元メッシュモデルを復元する．

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つたり，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．

### (iv) 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 [17]–[22]．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 [40] 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 [41] などが適用される．

デプスマップ統合に基づく手法の中でも，特に，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く，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．また，特徴領域拡張に基づく手法と比べて，3 次元復元対象によって 3 次元復元結果が特徴点周辺に限定されるといった問題もない．一方で，デプスマップ生成が 3 次元復元精度と処理時間に大きく影響し，高精度な 3 次

元復元を行うためには、計算コストが膨大になるといった問題がある。また、特徴領域拡張に基づく手法と同様に、多視点画像間の対応付けには誤対応が含まれるため、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。

## 2.4 デプスマップ統合に基づく 3 次元復元

2.3 節で述べた 4 つの 3 次元復元手法の中でも、入力する多視点画像によって問題が収束しなかったり、3 次元復元結果が偏ったりすることがないことから、デプスマップ統合に基づく 3 次元復元手法は実用性が高い。本論文では、このデプスマップ統合に基づく手法に着目し、本節において、デプスマップ統合に基づく手法についてより詳しく述べる。デプスマップ統合に基づく手法は、(i) 多視点画像からの複数のデプスマップを生成し、(ii) 複数のデプスマップからメッシュモデルを生成する 2 つのステップで構成される。(i) のステップでは、ウィンドウマッチングを用いた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により、多視点画像からある参照視点におけるデプスマップを生成する。(ii) のステップでは、(i) のステップで生成されたデプスマップを、3 次元ボリュームに基づく手法や 3 次元点群に基づく手法により統合し、3 次元メッシュモデルを生成する。ここで、(i) のステップで生成されるデプスマップの精度は、最終的に生成されるメッシュモデルの精度に大きく影響する。そのため、高精度な 3 次元復元を行うためには、デプスマップを高精度に求めることが重要である。以下に、各種デプスマップ生成手法の特徴についてまとめる。

デプスマップ統合に基づく 3 次元復元において、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よく知られている [3]。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 次元点の尤度を計算する。そして、奥行きと尤度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。Goesele らは、複数のステレオペア間で計算される NCC の相関値を統合することで、信頼性の高いデプスマップ生成を実現している [17]。また、Campbell らは、Goesele らの手法に対して、各座標における奥行きと NCC の相関値の関係から決定される MRF (Markov Random Field) に基づきデプスマップを生成することで、高精度かつ誤対応の少ないデプスマップ生成を行っている [20]。さらに、Bradley らは、Furukawa らの特徴領域拡張に基づく 3 次元復元手法と同様に、3 次元点の 3 次元座標だけでなく法線ベク



## 第 2 章 多視点画像からの 3 次元復元に関する基礎的考察

トルも考慮することで、画像間の変形が大きくなる基線長の長いステレオ画像にも対応し、より自由な撮影による 3 次元復元を実現している [19]。これらの手法では、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った。この問題に対して、多くの手法では、視体積交差法により生成される 3 次元ボリュームを用いて奥行きの探索範囲を限定したり、始めに疎に配置されたシード点の奥行きを決定し、そのシード点の奥行きから密な点の奥行きの探索範囲を限定したりすることで、処理時間の増加を抑えている [22], [30]。

プレーンスweepingに基づくデプスマップ生成手法の他に、多視点画像からの 3 次元復元に用いられるデプスマップ生成手法として、偏微分方程式に基づく最適化などのエネルギー関数の最小化に基づくデプスマップ生成手法がある [18], [42]。これらの手法では、奥行きの連続性や視点間の輝度の類似性からエネルギー関数を定義し、そのエネルギー関数が最小となるようにデプスマップを生成する。エネルギー関数の最小化には、最急降下法や確率伝搬法などが用いられる事が多い。このエネルギー最小化に基づくデプスマップ生成では、プレーンスweepingに基づく手法に比べ、誤対応が少なく、高精度なデプスマップ生成が可能である。しかし、複雑な最適化問題を解く必要があるため、3 次元ボリュームやメッシュの最適化に基づく 3 次元復元手法と同様に、局所的な最適解に収束しやすく、初期値のデプスマップを高精度に求める必要がある。そのため、エネルギー関数の最小化に基づく手法は、他の手法により生成されたデプスマップの高精度化の際に用いられることが多い。

また、Tola らや Li らは、NCC のような領域ベースのマッチング手法ではなく、特徴ベースのマッチング手法の 1 つである DAISY に基づく対応付け手法を用いてデプスマップを生成している [21], [22], [43]。DAISY に基づく対応付け手法は、画像のすべての座標について DAISY の特徴量記述子を求め、エピポーラ線上で DAISY の特徴量同士を比較することで、対応付けを行う。この手法は、SIFT や SURF などの他の特徴ベースの対応付け手法と同様に、画像の幾何的変形にロバストであり、画像変形の大きい基線長の長いステレオ画像においてもロバストな対応付けが可能である。また、グラフカットなどの最適化手法と組み合わせることで、誤対応の少ないデプスマップ生成が行える。しかし、DAISY の特徴量は、各ピクセルごとにしか求められないため、サブピクセルレベルの対応付けが行えず、デプスマップの精度は、他の手法に比べて低いといった問題がある。この問題に対して、Li らは、デプスマップの統合にバンドルアジャストメントを適用することで、精度の低いデプスマップから高精度な 3 次元復元を行っている [21]。また、Tola らは、デプスマップの統合の際に、基線長の長いステレオ画像の対応付け結果を優先することで、3

## 2.5 NCC を用いたプレーンスイーピングに基づくデプスマップ生成

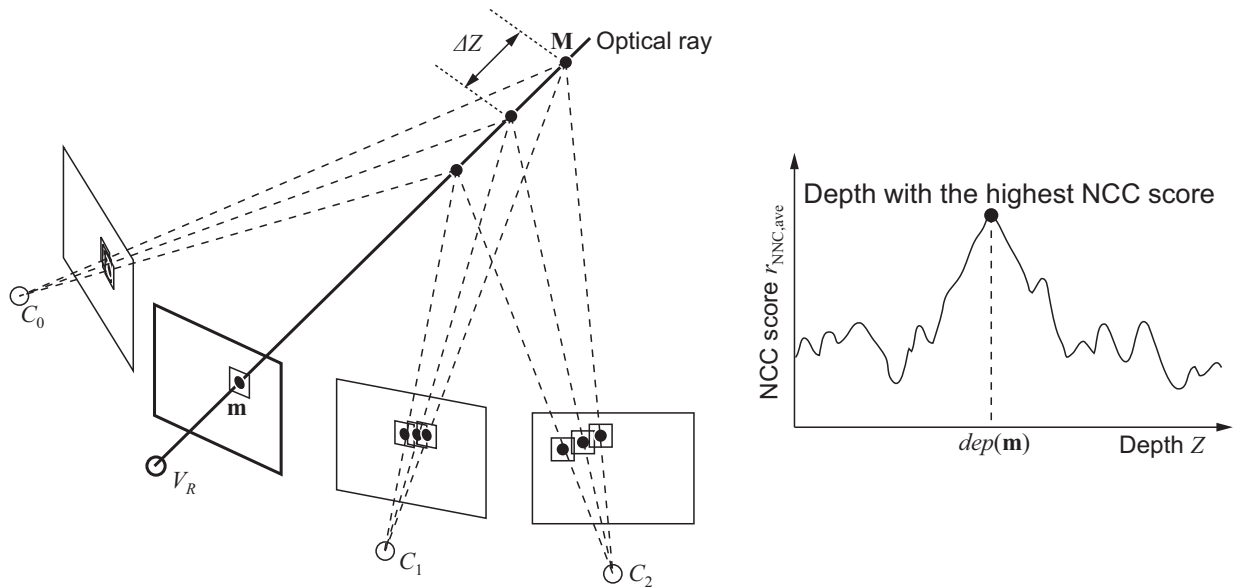


図 2.1 プレーンスイーピングに基づくデプスマップ生成

次元復元精度を向上させている [22] .

## 2.5 NCC を用いたプレーンスイーピングに基づくデプスマップ生成

プレーンスイーピングに基づくデプスマップ生成は、最も基本的なデプスマップ生成手法の 1 つである [3] . 本節では、NCC によるウィンドウマッチングを用いたプレーンスイーピングに基づくデプスマップ生成手法として、Goesele らの手法 [17] について具体的な処理を述べる .

Goesele らの手法では、カメラパラメータが既知の多視点画像  $V = \{V_0, \dots, V_{L-1}\}$  のうち、参照視点  $V_R \in V$  とその近傍視点  $C = \{C_0, \dots, C_{K-1}\} \subset V - \{V_R\}$  を入力として考える . ここで、 $L$  および  $K$  は、それぞれ多視点画像の数と近傍視点の数である . また、奥行き  $Z$  の探索範囲を  $Z_{\min}$  から  $Z_{\max}$  とし、奥行き変化の刻み幅を  $\Delta Z$  とする . 参照視点  $V_R$  上の任意の点  $m = (u, v)$  について以下の Step1 から Step5 の処理により奥行きおよび相関値、信頼値を決定し、 $dep(m)$ ,  $corr(m)$ ,  $conf(m)$  に書き込む . この  $m$  の座標を  $V_R$  上で変化させながら処理を繰り返すことで、デプスマップ  $dep$ 、相関値マップ  $corr$ 、信頼値マップ  $conf$  を決定する . 各点  $m$  における処理の流れを図 2.1 に示す .

## 第 2 章 多視点画像からの 3 次元復元に関する基礎的考察

Step1: 参照視点  $V_R$  において座標  $\mathbf{m}$  を通る視線上で奥行きが  $Z_{\min}$  である 3 次元点を  $\mathbf{M} = (X, Y, Z)$  の初期値とする .

Step2:  $i = 0, \dots, K-1$  について,  $\mathbf{M}$  を近傍視点  $C_i$  に投影した点を  $\mathbf{m}'_i$  とする . さらに, 参照視点  $V_R$  上で  $\mathbf{m}$  を中心に切り出された  $w \times w$  ピクセルのウィンドウ  $f$  と, 近傍視点  $C_i$  上で  $\mathbf{m}'_i$  を中心に切り出された  $w \times w$  ピクセルのウィンドウ  $g_i$  から NCC を用いて相関値  $r_{\text{NCC},i}$  を計算する .

Step3: NCC の相関値  $r_{\text{NCC},i}$  のうち,  $r_{\text{NCC},i} > th_{\text{corr}}$  であるものを平均化し,  $r_{\text{NCC,ave}}$  とする . ここで,  $th_{\text{corr}}$  は相関値の閾値であり,  $r_{\text{NCC},i} > th_{\text{corr}}$  となる NCC の相関値の個数を  $K'$  とする .

Step4:  $Z < Z_{\max}$  である限り, 3 次元点  $\mathbf{M}$  を視線上で  $Z \leftarrow Z + \Delta Z$  となるように動かし, Step2 から Step4 を繰り返す .

Step5:  $Z_{\min}$  から  $Z_{\max}$  において, 最も  $r_{\text{NCC,ave}}$  が高くなったときの奥行き  $Z$  と相関値  $r_{\text{NCC,ave}}$  から,  $dep(\mathbf{m})$ ,  $corr(\mathbf{m})$ ,  $conf(\mathbf{m})$  を以下のように定義する .

$$dep(\mathbf{m}) = Z \quad (2.1)$$

$$corr(\mathbf{m}) = r_{\text{NCC,ave}} \quad (2.2)$$

$$conf(\mathbf{m}) = \frac{K'}{K} \frac{(r_{\text{NCC,ave}} - th_{\text{corr}})}{(1 - th_{\text{corr}})} \quad (2.3)$$

以上の処理により決定されたデプスマップ  $dep$  について, 相関値マップ  $corr$ , 信頼値マップ  $conf$  が閾値以上の点のみを復元することにより, 信頼性の高い 3 次元点群を得ることが可能である .

## 2.6 むすび

本章では, 多視点画像からの 3 次元復元に関する基礎的考察を行った . まず, 多視点画像の撮影方法とカメラキャリブレーションについて述べた . 次に, 多視点画像からの 3 次元復元の分類とそれぞれの特徴について述べた . そして, 多視点画像からの 3 次元復元手法の 1 つであるデプスマップ統合に基づく 3 次元復元について述べた . さらに, 最も基本的なデプスマップ生成手法の 1 つであるプレーンスライピングに基づくデプスマップ生成について述べた .

## 第3章

# 1 次元位相限定相関法に基づく画像マッチング

## 3.1 まえがき

位相限定相関法 (POC) は画像の位相情報に着目した画像マッチング手法である。2つの画像信号から1つの相関値を計算するNCCのような画像マッチング手法と異なり、POCは2つの画像信号から1つの相関関数を計算するといった特徴がある。この相関関数から2つの画像信号の平行移動量をサブピクセル精度で求めることが可能であり、POCを用いることで非常に高精度な画像マッチングを行うことができる。

本章では、1次元位相限定相関法 (1次元 POC) に基づく画像マッチング手法について述べる [27]。まず、2つの1次元画像信号間の平行移動量を1次元 POC 関数を用いて算出する手法について述べる。次に、1次元 POC に基づく画像マッチングの各種高精度化手法について述べる。

## 3.2 1次元 POC に基づく画像マッチング

はじめに、2つの1次元画像信号間の平行移動量を1次元 POC 関数から算出する手法について述べる。一般に、ステレオ画像間の局所的な平行移動は2次元方向に生じるが、ステレオ画像のそれぞれのエピポーラ線が水平座標、または、垂直座標と平行になるようにステレオ平行化 [3], [4] を行うことで、局所的な平行移動を1次元方向に限定することができる。平行化されたステレオ画像のように、視差が1次元方向に限定される場合、1次元 POC 関数を用いることで、2つの1次元画像信号間の平行移動量を求めることが可能である。

2つの1次元画像信号を  $f(n)$  および  $g(n)$  とする。ここで、1次元画像信号の離散空間インデックスを便宜上、 $n = -M, \dots, M$  とする。ただし、 $M$  は正の整数であり、1次元

### 第 3 章 1 次元位相限定相関法に基づく画像マッチング

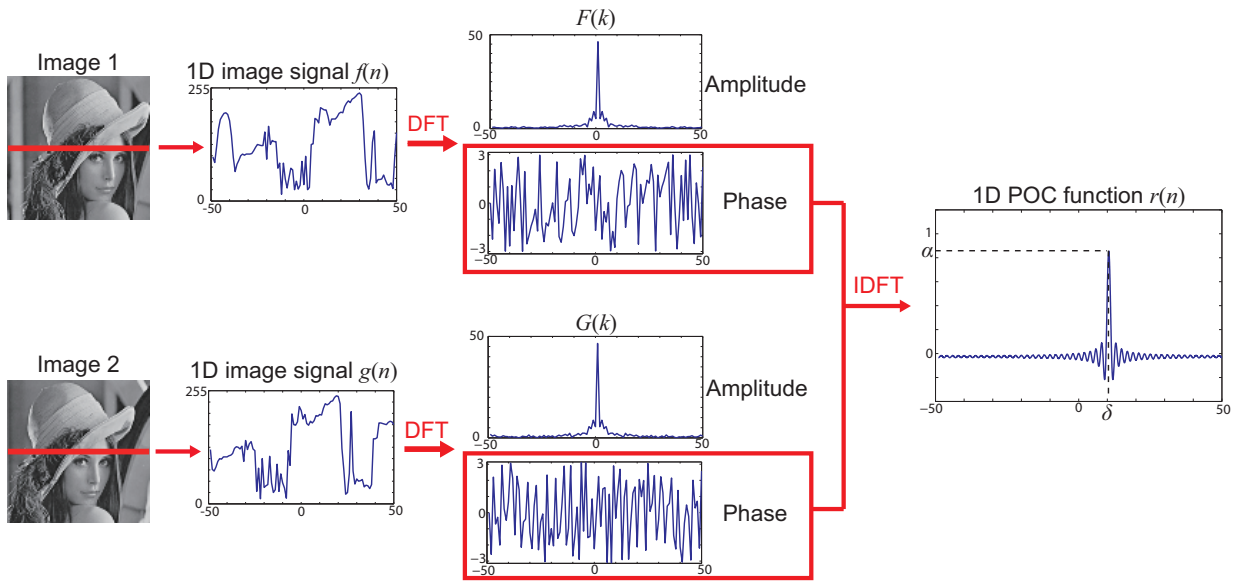


図 3.1 1次元位相限定相関法 (1次元 POC)

画像信号の長さ  $N$  は  $N = 2M + 1$  で与えられる．なお，ここでは説明を簡単にするために離散空間インデックスを正負対称にとり，かつ 1次元画像信号の長さを奇数にしているが，これは本手法の構成において必須ではない．すなわち，通常よく用いられるように非負の離散空間インデックスを用い，1次元画像信号の長さを任意の正の整数に設定するように一般化することが可能である．以下では，フーリエ変換の記述の簡便さのために，上記のような設定を用いる．1次元画像信号  $f(n)$  および  $g(n)$  から 1次元 POC 関数  $r(n)$  を計算する際の計算の流れを図 3.1 に示す．

1次元画像信号  $f(n)$  および  $g(n)$  の 1次元離散フーリエ変換 (Discrete Fourier Transform: DFT)  $F(k)$  および  $G(k)$  を次式で定義する．

$$F(k) = \sum_{n=-M}^M f(n)W_N^{kn} = A_F(k)e^{j\theta_F(k)} \quad (3.1)$$

$$G(k) = \sum_{n=-M}^M g(n)W_N^{kn} = A_G(k)e^{j\theta_G(k)} \quad (3.2)$$

ただし， $W_N = e^{-j\frac{2\pi}{N}}$  は回転因子を表す．ここで， $A_F(k)$  および  $A_G(k)$  はそれぞれ 1次元画像信号  $f(n)$  および  $g(n)$  の振幅成分であり， $\theta_F(k)$  および  $\theta_G(k)$  はそれぞれ  $f(n)$  および  $g(n)$  の位相成分である．また，離散周波数インデックスを  $k = -M, \dots, M$  とする．こ

のとき，正規化相互パワースペクトル  $R(k)$  を次式で与える．

$$R(k) = \frac{F(k)\overline{G(k)}}{|F(k)G(k)|} = e^{j(\theta_F(k)-\theta_G(k))} \quad (3.3)$$

ここで， $\overline{G(k)}$  は  $G(k)$  の複素共役を表す． $f(n)$  と  $g(n)$  の 1 次元 POC 関数  $r(n)$  を，正規化相互パワースペクトル  $R(k)$  の 1 次元逆離散フーリエ変換 (Inverse Discrete Fourier Transform: IDFT) として以下の式で定義する．

$$r(n) = \frac{1}{N} \sum_{k=-M}^M R(k)W_N^{-kn} \quad (3.4)$$

以下では，1 次元画像信号  $f(n)$  と  $g(n)$  が互いに微小量  $\delta$  だけ平行移動した関係にある場合を考える．つまり，この  $f(n)$  と  $g(n)$  に含まれる平行移動量  $\delta$  をサブピクセル分解能で推定する問題を考える． $\delta$  が小さい場合， $f(n)$  と  $g(n)$  の 1 次元 DFT  $F(k)$  および  $G(k)$  について次式の近似が成り立つ．

$$G(k) \simeq F(k)e^{-j\frac{2\pi}{N}k\delta} \quad (3.5)$$

このとき， $F(k)$  と  $G(k)$  の正規化相互パワースペクトル  $R(k)$  は次式で与えられる．

$$R(k) \simeq e^{j\frac{2\pi}{N}k\delta} \quad (3.6)$$

式 (3.4) から，1 次元 POC 関数  $r(n)$  は次式で与えられる．

$$r(n) = \frac{1}{N} \sum_{k=-M}^M R(k)W_N^{-kn} \quad (3.7)$$

$$\simeq \frac{\alpha \sin(\pi(n+\delta))}{N \sin\left(\frac{\pi}{N}(n+\delta)\right)} \quad (3.8)$$

上式は，1 次元画像信号  $f(n)$  と  $g(n)$  が  $\delta$  だけ微小に平行移動した場合の 1 次元 POC 関数の一般系を表している．ここで， $\alpha$  は相関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．画像に対して，無相関なノイズが加わると  $\alpha$  の値が減少するため，実際には  $\alpha \leq 1$  となる．このとき，1 次元 POC 関数のピークの高さ  $\alpha$  は 2 つの信号の類似度に，相関ピークの位置座標  $\delta$  は 2 つの信号の平行移動量に相当する．1 次元 POC 関数の計算値からピーク値  $\alpha$  とピーク座標  $\delta$  を推定することで，2 つの信号の類似度とサブピクセル精度の平行移動量を求めることができる．

### 3.3 1 次元 POC に基づく画像マッチングの高精度化

以下では, 1 次元 POC 関数を用いて平行移動量  $\delta$  を推定する際に重要となる各種高精度化手法について述べる [27].

#### 3.3.1 窓関数の適用

1 次元 DFT では, 信号が循環することを仮定するため, 端点での信号の不連続性が問題となる. この不連続性の影響を軽減するため, 入力された 1 次元画像信号  $f(n)$  および  $g(n)$  に対して窓関数を適用する. 本論文では, 窓関数として次式で表される 1 次元ハニング窓を用いる.

$$w(n) = \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\pi n}{M}\right) \quad (3.9)$$

1 次元ハニング窓では, 信号長の半分が半値幅となる. したがって, 窓関数を適用しないウィンドウマッチング手法と比較する場合, 窓関数の半値幅がウィンドウサイズに対応すると考えられる.

#### 3.3.2 スペクトルの重み付け

一般に, 自然画像のエネルギーは低周波数成分に集中し, 高周波数成分のエネルギーは相対的に小さいことが知られている. このため, エイリアシング, ノイズ, 歪みなどの外乱が加わると, 高周波数成分の SN 比が大幅に低下する. そこで, 信頼性の低い高周波数成分の影響を抑制するために, 正規化相互パワースペクトル  $R(k)$  の計算の際に, 低域通過型のスペクトル重み付け関数  $H(k)$  を適用する. 本論文では, 次式で表されるガウス関数を用いる.

$$H(k) = e^{-2\pi^2\sigma^2\frac{k^2}{N^2}} \quad (3.10)$$

ここで,  $\sigma$  はガウス関数の幅を示す定数であり, 本論文では  $\sigma = \sqrt{0.5}$  とする. このとき, POC 関数のピークモデルも以下の式で表されるガウス型になる.

$$r(n) = \frac{1}{N} \sum_{k=-M}^M H(k)R(k)W_N^{-kn} \quad (3.11)$$

$$\simeq \frac{\alpha}{\sqrt{2\pi\sigma}} e^{-\frac{(n+\delta)^2}{2\sigma^2}} \quad (3.12)$$

### 3.3.3 ピークモデルの関数フィッティングによるパラメータ推定

一般に、2つの1次元画像信号間の平行移動量  $\delta$  は実数値であり、1次元 POC 関数のピーク座標はサンプリング格子点の間に存在する。そこで、1次元 POC 関数のピークモデルが式 (3.12) で表されることを考慮し、実際に計算された POC 関数の数値データに対して式 (3.12) に基づく関数フィッティングを適用することで、画像のピクセル間に存在するピーク位置を推定する。このとき、 $\alpha$  および  $\delta$  がフィッティングパラメータとなる。本論文では、関数フィッティング手法として、非線型最小 2 乗問題の解法の 1 つである Levenberg-Marquardt 法を用いる。

### 3.3.4 複数の 1次元画像信号の利用

1次元 POC に基づく画像マッチング手法をステレオ画像間のウィンドウマッチングに適用する場合、注目点を中心にマッチングウィンドウを設定し、設定されたウィンドウ間で1次元 POC 関数を計算する。3.2 節では、1組の1次元画像信号を用いて、その1次元画像信号間の平行移動量を求める手法について説明した。しかし、実際に撮影された画像において、1組の1次元画像信号のみを用いて平行移動量を推定すると、ノイズなどの影響により誤差が増幅することを実験的に確認している。そこで、マッチングウィンドウを設定する際に、注目点の近傍から複数の1次元画像信号の組を抽出し、それらの1次元 POC 関数の平均を平均 1次元 POC 関数とする。図 3.2 に示すように、1組の1次元画像信号のみを用いて求めた1次元 POC 関数の Peak-to-Noise Ratio (PNR) が低い場合においても、近傍の複数の1次元 POC 関数を平均化することで PNR を向上させることが可能である。この平均 1次元 POC 関数を相関ピーク値  $\alpha$  および平行移動量  $\delta$  の推定に用いることでノイズの影響を抑えることが可能である。

### 3.3.5 ウィンドウの拡大縮小によるステレオ画像間の画像変形の軽減

1次元 POC に基づく画像マッチング手法をステレオ画像間のウィンドウマッチングに適用する場合、ステレオ画像間の局所的な画像変形を平行移動に近似し、その平行移動量を1次元 POC を用いて計算する。基線長の短いステレオ画像を用いる場合、ステレオ画像間の画像変形は小さく、局所的な画像変形は平行移動のみに近似することが可能である。しかし、基線長の長いステレオ画像を用いる場合、ステレオ画像間の画像変形が大きくなるため、局所的な画像変形を平行移動のみに近似できず、1次元 POC を用いて高精度な平行移動量を計算することが困難である。一方で、平行化後のステレオ画像の場合、その



### 第 3 章 1 次元位相限定相関法に基づく画像マッチング

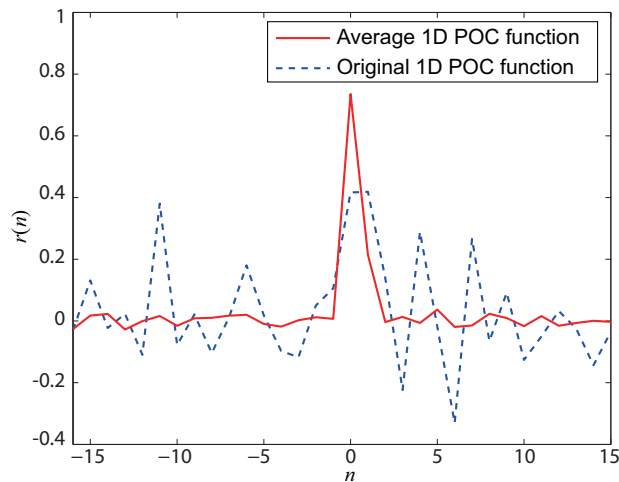


図 3.2 1 次元 POC 関数の平均化による Peak-to-Noise Ratio (PNR) の向上

画像変形はエピポーラ線上の 1 次元方向に限定される。つまり，エピポーラ線が水平座標軸と平行になるように平行化した場合，ステレオ画像間の画像変形は水平方向のみに，エピポーラ線が垂直座標軸と平行になるように平行化した場合，ステレオ画像間の画像変形は垂直方向のみに限定される。そこで，ステレオ画像間の局所的な画像変形をエピポーラ線方向の拡大縮小と平行移動に近似し，マッチングウィンドウのウィンドウサイズを拡大縮小することで，ステレオ画像間の画像変形を軽減する。これにより，基線長の長いステレオ画像のように，ステレオ画像間の画像変形が大きい場合においても，1 次元 POC に基づく画像マッチング手法を適用することが可能である [28], [44], [45]。詳細は付録 A を参照されたい。

## 3.4 むすび

本章では，1 次元位相限定相関法（1 次元 POC）に基づく画像マッチング手法について述べた。まず，2 つの 1 次元画像信号間の平行移動量を 1 次元 POC 関数を用いて算出する手法について述べた。次に，1 次元 POC に基づく画像マッチングの各種高精度化手法について述べた。

## 第4章

# 位相限定相関法に基づく多視点画像からの3次元復元

## 4.1 まえがき

従来のデプスマップ統合に基づく3次元復元手法では、多くの場合、NCCを用いたプレーンスイーピングに基づく手法によりデプスマップを生成する。プレーンスイーピングに基づくデプスマップ生成において、3次元復元精度は奥行き変化の刻み幅に依存する。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った。一方で、奥行き変化の刻み幅が非常に小さい場合、3次元点の移動幅がステレオ画像上でサブピクセルとなり、サブピクセルの座標の輝度値を補間により求めているため、信号処理の観点から、無駄な計算が多いといえる。これに対して、本論文では、ステレオ画像間のウィンドウマッチングに位相限定相関法を用い、1回のマッチングでサブピクセル精度の奥行きを決定することで、少ない計算コストで高精度な3次元復元を行う。さらに、複数のステレオ画像間で計算されるPOC関数を統合し、誤対応の少ないデプスマップ生成を実現する。

本章では、位相限定相関法に基づく多視点画像からの3次元復元について述べる。まず、提案手法で重要となる正規化視差について説明し、正規化視差に基づくPOC関数の統合について述べる。次に、POCに基づく多視点画像からのデプスマップ生成手法について述べる。そして、実際の多視点画像を用いた実験によって、提案手法の評価を行う。

## 4.2 POC に基づく多視点画像からのデプスマップ生成

本節では，POC に基づく多視点画像からのデプスマップ生成手法を提案する．提案手法では，複数のステレオペアにおける 1 次元 POC 関数を統合し，1 次元 POC 関数のピーク座標を用いて奥行きを決定する．しかし，異なるステレオペアから計算された 1 次元 POC 関数では，ステレオペアにおける視差の違いから，ピーク座標が異なるという問題がある．そこで，提案手法では，それぞれのステレオペアにおける視差を正規化することで，複数の 1 次元 POC 関数を統合する．従来の多視点画像間における相関関数の統合手法として Okutomi らの手法 [46] があるが，Okutomi らの手法では，すべてのカメラが同じ直線上にあることを仮定しており，一般的な多視点画像に適用できない．本節で述べる正規化視差に基づく相関関数の統合手法は，Okutomi らの手法をより一般化したものであり，カメラが同一の直線上にない場合においても相関関数を統合することが可能である．

提案手法では，カメラパラメータが既知の多視点画像  $V = \{V_0, \dots, V_{L-1}\}$  のうち，参照視点  $V_R \in V$  とその近傍視点  $C = \{C_0, \dots, C_{K-1}\} \subset V - \{V_R\}$  を入力として考える．ここで， $L$  および  $K$  は，それぞれ多視点画像の数と近傍視点の数である． $V_R$  および  $C$  から  $K$  組の平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  ( $i = 0, \dots, K-1$ ) を生成し [3], [4]，それぞれのペアで計算された 1 次元 POC 関数を正規化視差に基づき統合することで，1 次元 POC 関数のピーク座標から奥行きを決定する．この 1 次元 POC 関数による奥行き推定と画像ピラミッドを用いた階層的探索を組み合わせることで，従来手法と比べ，少ないマッチング回数で奥行き探索を実現することが可能である．

### 4.2.1 視差の正規化

平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における 3 次元点  $M = (X, Y, Z)$  の視差を考える [3], [4]．ここで， $C_i \in C$  である．また，世界座標と参照視点  $V_R$  のカメラ座標が一致すると仮定する．さらに，参照視点  $V_R$  と近傍視点  $C_i$  でステレオ平行化を行う際に，参照視点におけるカメラ座標の回転行列を次式のように  $R_i$  とする．

$$\mathbf{R}_i = \begin{bmatrix} R_{i11} & R_{i12} & R_{i13} \\ R_{i21} & R_{i22} & R_{i23} \\ R_{i31} & R_{i32} & R_{i33} \end{bmatrix} \quad (4.1)$$

このとき，世界座標における 3 次元点  $M$  と平行化後のカメラ座標における 3 次元点  $M_{\text{rect},i} = (X_i, Y_i, Z_i)$  の関係は次式で表される．

$$\mathbf{M} = \mathbf{R}_i \mathbf{M}_{\text{rect},i} \quad (4.2)$$

## 4.2 POC に基づく多視点画像からのデプスマップ生成

また，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における 3次元点  $M_{\text{rect},i}$  と視差  $d_i$  の関係は次式で表される [4] .

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} (u_i - u_{0i})B_i/d_i \\ (v_i - v_{0i})B_i/d_i \\ \beta_i B_i/d_i \end{bmatrix} \quad (4.3)$$

ここで， $(u_i, v_i)$  は平行化後の参照視点  $V_{\text{rect},i}$  における  $M_{\text{rect},i}$  の対応点のデジタル画像座標， $(u_{0i}, v_{0i})$  は平行化後の参照視点  $V_{\text{rect},i}$  における画像中心， $\beta_i$  は平行化後の焦点距離， $B_i$  はステレオペア間の基線長を表す．一般的に，焦点距離は  $f$  で表されるが，本論文では，マッチングを行う 1次元画像信号  $f$  との混同を避けるため， $\beta$  を用いる．式 (4.2) と式 (4.3) より，3次元点  $M$  と平行化後の視差  $d_i$  との関係は次式で表される．

$$\mathbf{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i/d_i \\ (v_i - v_{0i})B_i/d_i \\ \beta_i B_i/d_i \end{bmatrix} \quad (4.4)$$

同様に，参照視点  $V_R$  と近傍視点  $C_j \in C - \{C_i\}$  の間で次式が成り立つ．

$$\mathbf{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_j \begin{bmatrix} (u_j - u_{0j})B_j/d_j \\ (v_j - v_{0j})B_j/d_j \\ \beta_j B_j/d_j \end{bmatrix} \quad (4.5)$$

よって，式 (4.4) と式 (4.5) より，

$$\begin{aligned} X &= \frac{B_i}{d_i}(R_{i11}(u_i - u_{0i}) + R_{i12}(v_i - v_{0i}) + R_{i13}\beta_i) \\ &= \frac{B_j}{d_j}(R_{j11}(u_j - u_{0j}) + R_{j12}(v_j - v_{0j}) + R_{j13}\beta_j) \end{aligned} \quad (4.6)$$

$$\begin{aligned} Y &= \frac{B_i}{d_i}(R_{i21}(u_i - u_{0i}) + R_{i22}(v_i - v_{0i}) + R_{i23}\beta_i) \\ &= \frac{B_j}{d_j}(R_{j21}(u_j - u_{0j}) + R_{j22}(v_j - v_{0j}) + R_{j23}\beta_j) \end{aligned} \quad (4.7)$$

$$\begin{aligned} Z &= \frac{B_i}{d_i}(R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i) \\ &= \frac{B_j}{d_j}(R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j) \end{aligned} \quad (4.8)$$

となり，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における視差  $d_i$  と視点ペア  $V_{\text{rect},i} - C_{\text{rect},j}$  における視差  $d_j$  の間には次の関係が成り立つ．

$$d_i = \frac{R_{i11}(u_i - u_{0i}) + R_{i12}(v_i - v_{0i}) + R_{i13}\beta_i}{R_{j11}(u_j - u_{0j}) + R_{j12}(v_j - v_{0j}) + R_{j13}\beta_j} \frac{B_i}{B_j} d_j$$

## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

$$\begin{aligned}
 &= \frac{R_{i21}(u_i - u_{0i}) + R_{i22}(v_i - v_{0i}) + R_{i23}\beta_i \frac{B_i}{B_j} d_j}{R_{j21}(u_j - u_{0j}) + R_{j22}(v_j - v_{0j}) + R_{j23}\beta_j \frac{B_j}{B_j} d_j} \\
 &= \frac{R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i \frac{B_i}{B_j} d_j}{R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j \frac{B_j}{B_j} d_j} \quad (4.9)
 \end{aligned}$$

式 (4.9) で示したように，異なる平行ステレオペアにおける視差  $d_i$  と  $d_j$  はある倍率によって関係付けられ，その倍率はカメラパラメータと参照視点における 3 次元点の対応点のデジタル画像座標に依存する．以上より，各ステレオペアにおける視差の倍率を考慮することで，正規化視差  $d$  を定義する．平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  ( $i = 0, \dots, K-1$ ) が与えられたとき，各平行ステレオペアにおける視差  $d_i$  と正規化視差  $d$  の間には以下の関係が成り立つ．

$$d_i = s_i d \quad (4.10)$$

ここで， $s_i$  は視差の倍率である．平行化前の参照視点における注目点  $(u, v)$  について，平行化後の各平行ステレオペアにおける視差倍率  $s_i$  を式 (4.9) から次式のように定義する．

$$s_i = \frac{1}{|s|} (R_{i31}(u_i - u_{0i}) + R_{i32}(v_i - v_{0i}) + R_{i33}\beta_i) B_i \quad (4.11)$$

ここで， $|s|$  は

$$|s| = \frac{1}{K} \sum_{j=0}^{K-1} (R_{j31}(u_j - u_{0j}) + R_{j32}(v_j - v_{0j}) + R_{j33}\beta_j) B_j \quad (4.12)$$

である．このとき，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における 3 次元点  $M_{\text{rect},i}$  は次式で与えられる．

$$\mathbf{M}_{\text{rect},i} = \begin{bmatrix} (u_i - u_{0i})B_i / (s_i d) \\ (v_i - v_{0i})B_i / (s_i d) \\ \beta_i B_i / (s_i d) \end{bmatrix} \quad (4.13)$$

### 4.2.2 正規化視差に基づく POC 関数の統合

図 4.1 のように，3 次元点  $M$  に対して，視線を微小量  $\Delta M = (\Delta X, \Delta Y, \Delta Z)$  だけずれた点  $M' = M + \Delta M$  について，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  ( $i \in \{0, \dots, K-1\}$ ) における視差を考える．このとき，式 (4.4) より，3 次元点  $M'$  は，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における視差  $d'_i$  を用いて

$$\mathbf{M}' = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i / d'_i \\ (v_i - v_{0i})B_i / d'_i \\ \beta_i B_i / d'_i \end{bmatrix} \quad (4.14)$$

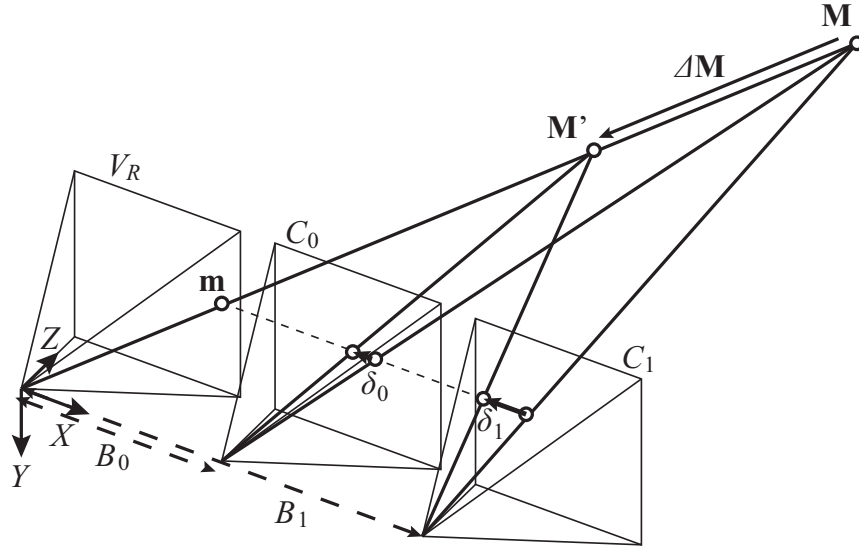


図 4.1 3次元点の変移と視差の変移の関係

と表される．ここで， $M$  を真の 3 次元点とすると，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  における視差に誤差  $\delta_i$  が生じるので， $M$  に対する視差  $d_i$  と  $M'$  に対する視差  $d'_i$  の関係は

$$d'_i = d_i + \delta_i \quad (4.15)$$

となる．このとき，局所的な画像変形を平行移動のみと仮定すると， $M'$  の対応点を中心に  $V_{\text{rect},i}$  から切り出したウィンドウ  $f_i$  と  $C_{\text{rect},i}$  から切り出したウィンドウ  $g_i$  の間には  $\delta_i$  の平行移動が生じていると考えられる．この平行移動量  $\delta_i$  は，3 章で述べたように， $f_i$  と  $g_i$  の 1 次元 POC 関数  $r_i$  を計算し，そのピーク座標を推定することで求められる．しかしながら，異なる平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  および  $V_{\text{rect},j} - C_{\text{rect},j}$  ( $j \in \{0, \dots, K-1\} - \{i\}$ ) において，平行移動量  $\delta_i$  と  $\delta_j$  は必ずしも一致せず，それぞれの平行ステレオペアから切り出されたウィンドウで計算される 1 次元 POC 関数  $r_i$  と  $r_j$  のピーク座標は異なる．例えば，図 4.1 では，ステレオペア  $V_R - C_0$  の基線長  $B_0$  よりも  $V_R - C_1$  の基線長  $B_1$  の方が長く，同じ奥行きの変化  $\Delta Z$  に対して  $V_R - C_1$  上の対応点の変動  $\delta_1$  の方が  $V_R - C_0$  上の対応点の変動  $\delta_0$  よりも大きくなる．そのため， $V_R - C_0$  から計算される POC 関数のピーク座標  $\delta_0$  と  $V_R - C_1$  から計算される POC 関数のピーク座標  $\delta_1$  は異なっており，このようなピーク座標が異なる POC 関数を統合することは困難である．

そこで，提案手法では，正規化視差に基づき各平行ステレオペアで切り出すウィンドウを拡大縮小することで，異なる平行ステレオペアで計算される 1 次元 POC 関数のピーク

#### 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

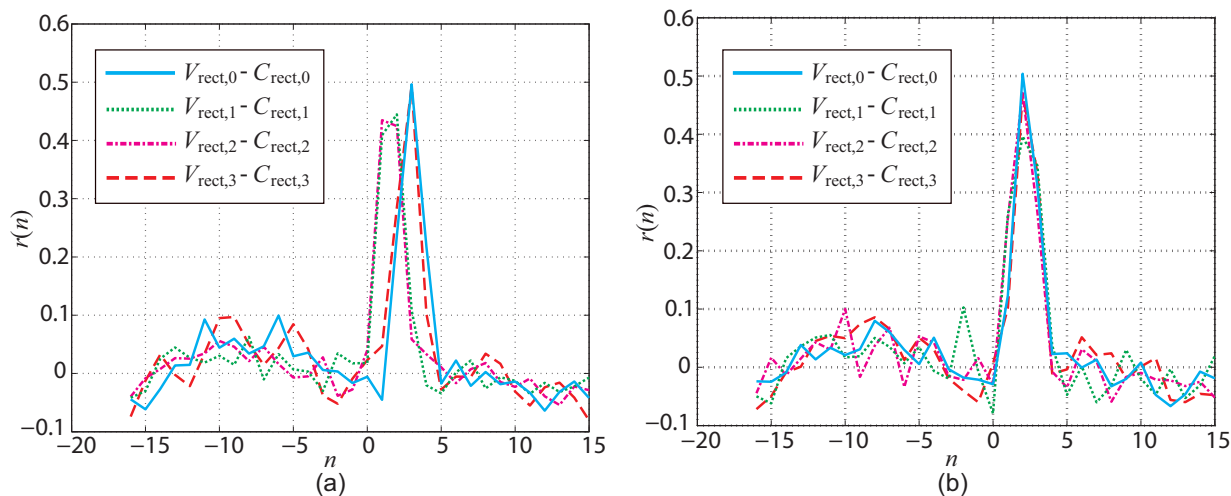


図 4.2 正規化視差に基づく相関関数の統合: (a) ウィンドウサイズ  $w$  で計算された POC 関数, (b) ウィンドウサイズ  $s_i w$  で計算された POC 関数

座標を一致させる． $M$  における正規化視差を  $d$ ， $M'$  における正規化視差を  $d' = d + \delta$  とすると式 (4.13) より  $M'$  は

$$M' = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i / (s_i(d + \delta)) \\ (v_i - v_{0i})B_i / (s_i(d + \delta)) \\ \beta_i B_i / (s_i(d + \delta)) \end{bmatrix} \quad (4.16)$$

と表される．よって，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  において対応点を中心に切り出したウィンドウ  $f_i$  と  $g_i$  間の平行移動量は  $s_i \delta$  となる．ここで，基準となるウィンドウサイズ  $w$  に対して， $f_i, g_i$  を切り出す際のウィンドウサイズを  $s_i w$  とする．この  $s_i w$  の大きさの 1 次元画像信号  $f_i, g_i$  を  $w$  の大きさにそろえるため， $1/s_i$  倍に拡大縮小した信号をそれぞれ  $f'_i, g'_i$  とする．このとき， $f'_i, g'_i$  間の平行移動量は  $\delta$  となり， $f'_i, g'_i$  から計算される 1 次元 POC 関数  $r_i$  のピーク座標は  $\delta$  となる．

ここで，平行ステレオペア  $V_{\text{rect},j} - C_{\text{rect},j}$  ( $j \in \{0, \dots, K-1\} - \{i\}$ ) について同様に考えると，切り出すウィンドウのサイズ  $s_j w$  は  $s_i w$  とは異なるが，1 次元 POC 関数  $r_j$  のピーク座標は  $\delta$  となり，平行ステレオペア  $V_{\text{rect},i} - C_{\text{rect},i}$  で計算される 1 次元 POC 関数  $r_i$  のピーク座標と一致する．実際に計算される POC 関数の例を図 4.2 に示す．図 4.2 (a) は，視差の倍率を考慮せず，すべての平行ステレオペアにおいて同じウィンドウサイズ  $w$  でウィンドウを切り出した場合の POC 関数である．この場合，ウィンドウ間の平行移動量  $\delta_i$  はそれぞれ異なり，計算される POC 関数のピーク座標は一致しない．これに対して，

図 4.2 (b) は、各ステレオペアにおいて、正規化視差に基づき切り出すウィンドウサイズを  $s_i w$  とした場合の POC 関数である。図 4.2 (b) より、視差の倍率を考慮することで、それぞれの POC 関数のピーク座標  $\delta$  が一致することがわかる。これにより、異なる平行ステレオペアで計算される 1 次元 POC 関数を統合することが可能となる。本論文では、複数の 1 次元 POC 関数の統合として、各平行ステレオペアで計算された 1 次元 POC 関数  $r_i$  ( $i = 0, \dots, K - 1$ ) を平均化し  $r_{ave}$  とする。統合された 1 次元 POC 関数  $r_{ave}$  のピーク位置  $\delta$  を推定することで真の 3 次元点  $M$  を以下の式より求めることができる。

$$\mathbf{M} = \mathbf{R}_i \begin{bmatrix} (u_i - u_{0i})B_i / (s_i(d' - \delta)) \\ (v_i - v_{0i})B_i / (s_i(d' - \delta)) \\ f_i B_i / (s_i(d' - \delta)) \end{bmatrix} \quad (4.17)$$

ここで、オクルージョンにより 3 次元点  $M$  が近傍視点  $C_i \in C$  の画像に写っていない場合や、物体境界において複数の視差を持つ領域でウィンドウを抽出した場合は、そのウィンドウから計算される 1 次元 POC 関数のピーク座標に非常に大きい誤差が生じることが予想される。一方で、マッチングウィンドウ間の画像変形が平行移動のみに近似できない場合は、そのウィンドウから計算される 1 次元 POC 関数のピーク値  $\alpha$  が低くなることが実験的に知られている。そこで、複数の 1 次元 POC 関数の平均化の際に、ピーク値  $\alpha$  が閾値  $th_{corr}$  を超えるもののみを用いて計算することで、オクルージョンや物体境界の影響などを抑えることが可能である。

### 4.2.3 POC に基づく多視点画像からのデプスマップ生成手法

本節では、POC に基づく多視点画像からのデプスマップ生成手法を提案する。4.2.2 節で述べた 1 次元 POC 関数のピーク座標を用いた奥行き推定では、マッチングウィンドウのウィンドウサイズの半分の視差に相当する奥行きのずれを求めることができる。そのため、奥行きの探索範囲を大きく設定する場合、ウィンドウサイズを大きくする必要がある。一方で、ウィンドウサイズが大きい場合、生成されるデプスマップが平滑化され、詳細な 3 次元形状を復元することが難しいといった問題がある。そこで、提案手法では、4.2.2 節で述べた POC 関数の統合による奥行き推定手法と、画像ピラミッドを用いた階層的探索 [3], [27] とを組み合わせる。以下に、処理手順を説明する。ここで、入力は参照視点  $V_R$  とその近傍視点  $C = \{C_0, \dots, C_{K-1}\}$  とし、出力はデプスマップ  $dep$  および相関値マップ  $corr$ 、信頼値マップ  $conf$  とする。また、各パラメータとして、相関ピーク値の閾値を  $th_{corr}$ 、画像ピラミッドの階層数を  $H$ 、基準となるウィンドウの大きさを  $w$ 、平均 1 次元 POC の計算に用いるライン数を  $l$  とする。



## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

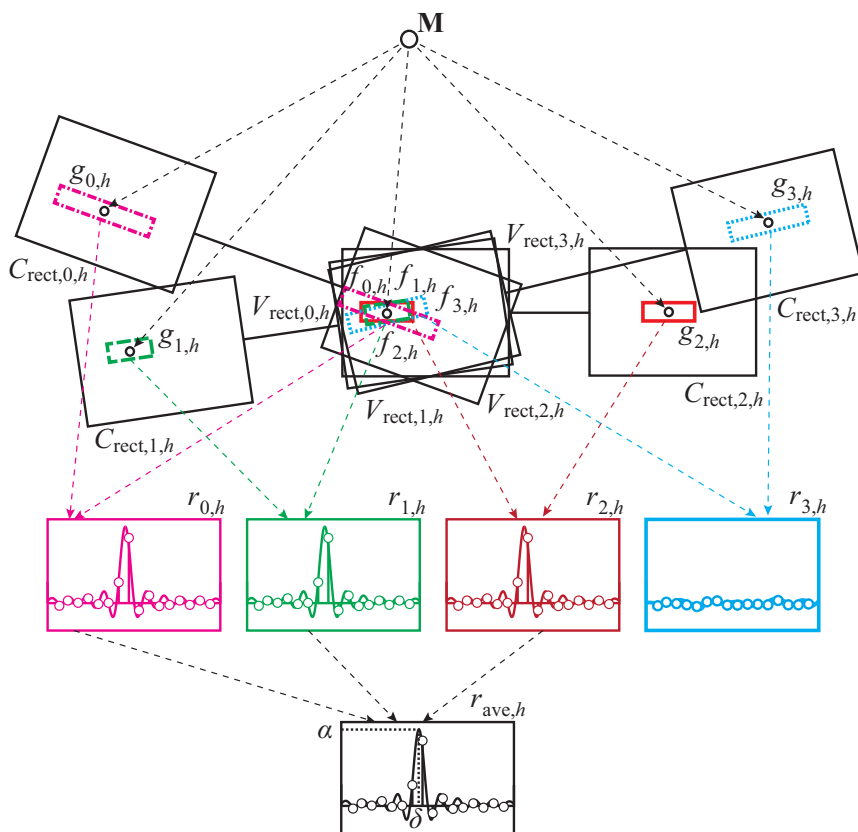


図 4.3 複数の POC 関数の統合による平行移動量推定

まず、参照視点  $V_R$  と近傍視点  $C$  から平行ステレオペアおよびその階層画像を生成する。  $i = 0, \dots, K-1$  について、ステレオペア  $V_R - C_i$  に対してステレオ平行化を行い、平行化後の画像をそれぞれ  $V_{rect,i,0}$ 、 $C_{rect,i,0}$  とする。そして、 $h = 1, \dots, H-1$  について、 $V_{rect,i,0}$  および  $C_{rect,i,0}$  を  $2^{-h}$  倍に縮小した画像を  $V_{rect,i,h}$  および  $C_{rect,i,h}$  とする。次に、画像ピラミッドの最上位層  $H-1$  において、画像全体のマッチングを行い、奥行き探索開始初期値  $Z_{init}$  を決定する。この  $Z_{init}$  は、 $i = 0, \dots, K-1$  について、 $V_{rect,i,H-1}$  と  $C_{rect,i,H-1}$  の POC 関数を計算し、そのピーク位置から画像全体の視差を求めることで決定する。そして、参照視点  $V_R$  上の任意の点  $m = (u, v)$  について、以下の Step1 から Step7 の処理により奥行きおよび相関値、信頼値を決定し、 $dep(m)$ 、 $corr(m)$ 、 $conf(m)$  に書き込む。この  $m$  の座標を  $V_R$  上で変化させながら処理を繰り返すことで、 $dep$ 、 $corr$ 、 $conf$  を決定する。画像ピラミッドの各階層における処理の流れを図 4.3 に示す。

Step1:  $i = 0, \dots, K-1$  について、 $m$  における視差倍率  $s_i$  を式 (4.11) から決定する。こ

のとき，平行化後の参照視点  $V_{\text{rect},i,0}$  における  $\mathbf{m}$  の対応点  $\mathbf{m}_i = (u_i, v_i)$  はステレオ平行化の変形に用いる射影行列から計算される．

Step2:  $h \leftarrow H - 1$  および 3 次元点  $\mathbf{M} \leftarrow \mathbf{M}_{\text{init}}$  として探索を開始する．この  $\mathbf{M}_{\text{init}}$  は，参照視点  $V_R$  において座標  $\mathbf{m}$  を通る視線上で奥行きが  $Z_{\text{init}}$  となる 3 次元点である．

Step3:  $i = 0, \dots, K - 1$  について， $\mathbf{M}$  をステレオペア  $V_{\text{rect},i,h} - C_{\text{rect},i,h}$  に投影した点をそれぞれ  $\mathbf{m}_{i,h} = (u_{i,h}, v_{i,h})$ ,  $\mathbf{m}'_{i,h} = (u'_{i,h}, v'_{i,h})$  とする．さらに， $V_{\text{rect},i,h}$  から  $\mathbf{m}_{i,h}$  を中心に  $s_i w \times l$  の大きさを切り出されたウィンドウ  $f_{i,h}$  と  $C_{\text{rect},i,h}$  から  $\mathbf{m}'_{i,h}$  を中心に  $s_i w \times l$  の大きさを切り出されたウィンドウ  $g_{i,h}$  から 1 次元 POC 関数  $r_{i,h}$  を計算する．

Step4:  $r_{i,h}$  のうち POC 関数のピーク値  $\alpha_i > th_{\text{corr}}$  であるものを平均化し， $r_{\text{ave},h}$  とする．ここで， $\alpha_i > th_{\text{corr}}$  となる POC 関数の個数を  $K'$  とする．

Step5: 1 次元 POC 関数  $r_{\text{ave},h}$  に対して関数フィッティングを行うことで，POC 関数のピーク値  $\alpha$  とピーク座標  $\delta$  を得る．この  $\delta$  より式 (4.17) を用いて 3 次元点  $\mathbf{M}$  を更新する．

Step6:  $h = 0$  となるまで  $h \leftarrow h - 1$  とし，Step3 から Step6 を繰り返す．

Step7:  $h = 0$  のときの 1 次元 POC 関数  $r_{\text{ave},0}$  と 3 次元点  $\mathbf{M} = (X, Y, Z)$  より， $dep(\mathbf{m})$ ,  $corr(\mathbf{m})$ ,  $conf(\mathbf{m})$  を以下のように定義する．

$$dep(\mathbf{m}) = Z \quad (4.18)$$

$$corr(\mathbf{m}) = \alpha \quad (4.19)$$

$$conf(\mathbf{m}) = \frac{K'(\alpha - th_{\text{corr}})}{K(1 - th_{\text{corr}})} \quad (4.20)$$

以上の処理により決定されたデプスマップ  $dep$  について，相関値マップ  $corr$ ，信頼値マップ  $conf$  が閾値以上の点のみを復元することにより，信頼性の高い 3 次元点群を得ることが可能である．

### 4.3 提案手法の性能評価

本論文で提案する POC に基づく多視点画像からのデプスマップ生成手法を用いることで，少ない計算コストで高精度かつロバストに 3 次元復元を行うことが可能である．本節

## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

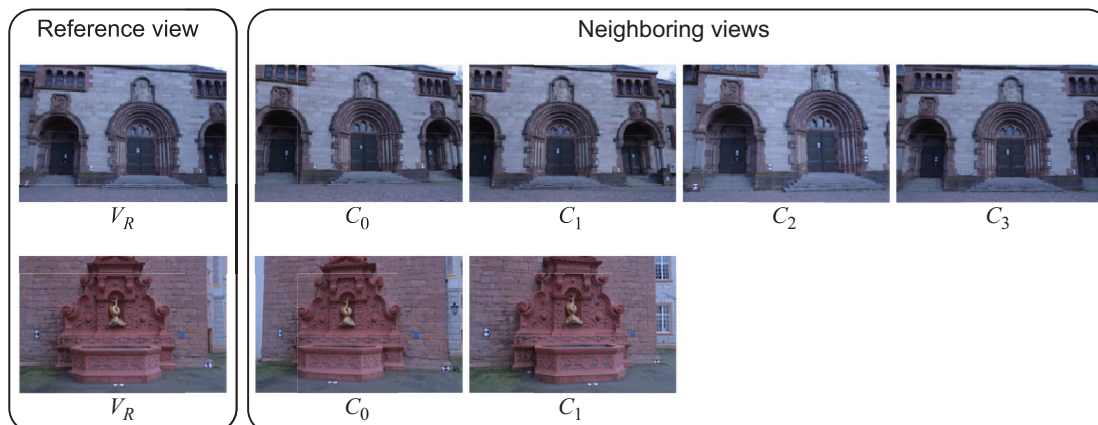


図 4.4 実験に用いる参照視点画像  $V_R$  および近傍視点画像  $C$  (上段: Herz-Jesu-P25, 下段: Fountain-P11)

では, データセット Herz-Jesu-P25 および Fountain-P11 [47] を用いた実験により, 従来手法および提案手法の 3 次元復元精度と処理時間を比較する.

### 4.3.1 実験方法

データセット Herz-Jesu-P25, Fountain-P11 には, 多視点画像とそのカメラパラメータ, および復元対象の真値のメッシュモデル, バウンディングボックスが含まれる [47]. ここで, データセットに含まれる多視点画像は, Herz-Jesu-P25 が 25 枚, Fountain-P11 が 11 枚である. それぞれのデータセットのうち 1 枚を参照視点  $V_R$  とし, 従来手法と提案手法を用いて  $V_R$  におけるデプスマップを生成する. このとき, Herz-Jesu-P25 では  $V_R$  の近傍 4 枚の画像を近傍視点  $C = \{C_0, C_1, C_2, C_3\}$  とし, Fountain-P11 では  $V_R$  の近傍 2 枚の画像を近傍視点  $C = \{C_0, C_1\}$  とする. Herz-Jesu-P25 および Fountain-P11 のうち実験に用いる参照視点画像  $V_R$  および近傍視点画像  $C$  を図 4.4 に示す. また, 各データセットに含まれる多視点画像は,  $3,072 \times 2,048$  ピクセルと高解像度の画像であるが, その他に,  $1,536 \times 1,024$  ピクセルおよび  $768 \times 512$  ピクセルに縮小した画像を生成し, 解像度の変化に対する比較も行う.

本実験では, 従来手法と提案手法を C 言語で実装し, 1 つの参照視点からのデプスマップ生成について比較する. 本論文では, 従来手法として, 最も基本的なプレーンスライピングに基づくデプスマップ生成手法の 1 つである Goesele らの手法を用いる. 提案手法は

4.2.3 節で述べた手法である．Goesele らの手法および提案手法を用いて生成されたデプスマップから 3 次元点群を復元し，復元された 3 次元点群と真値のメッシュモデルを比較することで，3 次元復元精度を評価する．さらに，Goesele らの手法と提案手法で，1 点の奥行き探索に必要な演算回数を比較することで，計算コストを評価する．また，実際に計算機を用いてデプスマップを生成するのに必要な処理時間を測定する．このとき，それぞれの解像度の画像において奥行きの探索を行う点数を同じにするため， $3,072 \times 2,048$  ピクセルの画像では 4 ピクセル間隔の格子状に配置した点について， $1,536 \times 1,024$  ピクセルの画像では 2 ピクセル間隔の格子状に配置した点について， $768 \times 512$  ピクセルの画像ではすべてのピクセルについてデプスマップを生成する．

まず，Goesele らの手法における各パラメータを以下のように設定する．Goesele らの手法では，奥行き変化の刻み幅  $\Delta Z$  が精度と処理時間に大きく影響する．そこで，本論文では，最も基線長の短いステレオ画像上で  $1, 1/2, 1/5, 1/10, 1/20$  ピクセルの視差の変化に相当するように，奥行きの刻み幅  $\Delta Z$  を設定し，この 5 つの  $\Delta Z$  について実験を行う．このとき，奥行きの探索範囲はデータセットに含まれるバウンディングボックス内とする．また，NCC のマッチングウィンドウの大きさは  $9 \times 9$  ピクセルとし，複数のステレオペアで計算される NCC の相関値を平均化する際の相関値の閾値  $th_{corr}$  は 0.6 とする．

次に，提案手法における各パラメータを以下のように設定する．階層的探索を行う際の画像ピラミッドの階層数  $H$  は， $3,072 \times 2,048$  ピクセルの画像で  $H = 4$ ， $1,536 \times 1,024$  ピクセルの画像で  $H = 3$ ， $768 \times 512$  ピクセルの画像で  $H = 2$  とする．画像ピラミッドの上位層  $h > 0$  では，相関ピークの閾値  $th_{corr}$  を 0.1，基準となるウィンドウの大きさ  $w$  を 32 ピクセルとする．画像ピラミッドの最下位層  $h = 0$  では，相関ピークの閾値  $th_{corr}$  を 0.6，基準となるウィンドウの大きさ  $w$  を 16 ピクセルとする．平均 POC の計算に用いるライン数  $l$  は，どの階層においても  $l = w/2 + 1$  とする．ここで，平均 1 次元 POC 関数の計算では，3.3.1 節で述べたようにハニング窓をかけるため，最下位層における  $16 \times 9$  ピクセルのウィンドウの情報量は，Goesele らの手法で用いる  $9 \times 9$  ピクセルのウィンドウの情報量に相当することに注意されたい．また，Fountain-P11 では，Herz-Jesu-P25 に比べて，復元対象の奥行きに対するステレオ画像間の基線長が長く，ステレオ画像間の画像変形が大きくなりやすい．ステレオ画像間の画像変形が大きくなると POC 関数の相関ピーク  $\alpha$  が低くなるため，Fountain-P11 では，最下位層における相関ピークの閾値  $th_{corr}$  を 0.3 とする．さらに，データセット Fountain-P11 では，復元対象が奥行き方向に広く配置されている．そのため，画像ピラミッド最上位層  $h = H - 1$  における 3 次元点の初期値  $M_{init}$  を決定する際に，視線上で奥行きが  $Z_{init}$  となる 3 次元点の他に，奥行きが  $Z_{init} \pm D$  となる 3 次元点についてもマッチングを行い，相関ピークの高さが最も高くなるものを選

## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

択する．このとき，奥行き探索幅  $D$  は最も基線長が短いステレオ画像上で  $w/4$  に相当するように決定する．

### 4.3.2 3 次元復元精度の評価

本節では，Goesele らの手法と提案手法を用いて生成されたデプスマップの 3 次元復元精度を評価する．このとき，3 次元復元精度として，次式で表される誤差率  $e$  を用いる．

$$e = \frac{|Z_{\text{calculated}} - Z_{\text{ground truth}}|}{Z_{\text{ground truth}}} \quad (4.21)$$

ここで， $Z_{\text{calculated}}$  は生成されたデプスマップの奥行きを， $Z_{\text{ground truth}}$  は参照視点  $V_R$  における真値のメッシュモデルの奥行きを表す．この誤差率を生成されたデプスマップの各ピクセルについて計算し，それぞれの手法における誤差率を比較する．

まず，データセット Herz-Jesu-P25 における 3 次元復元精度を比較する．Goesele らの手法および提案手法を用いて生成されたデプスマップより復元された 3 次元点群を図 4.5 から図 4.7 に示す．ここで，図 4.5，図 4.6，図 4.7 がそれぞれ Herz-Jesu-P25 の  $768 \times 512$ ， $1,536 \times 1,024$ ， $3,072 \times 2,048$  ピクセルの画像における 3 次元復元結果である．データセット Herz-Jesu-P25 において，誤差率の閾値とその閾値以下の誤差率をもつ 3 次元点数の関係を図 4.8 に，全 3 次元点のうち閾値以下の誤差率をもつ 3 次元点の割合（以下，正対応率と記す）の関係を図 4.9 に，それぞれの誤差率に対する 3 次元点の分布を図 4.10 に示す．また，画像サイズの変化に対する誤差率の平均値の変化を図 4.11 (a) に，正対応率の変化を図 4.11 (b) に示す．このとき，誤差率の閾値をそれぞれの画像サイズにおいて最も基線長の短いステレオ画像上で 1 ピクセルの誤差に相当する値に設定し，誤差率の平均値は誤差率が閾値以下である点のみを用いて計算した．図 4.11 (b) の正対応率についても同様に，誤差率の閾値は，それぞれの画像サイズにおいて最も基線長の短いステレオ画像上で 1 ピクセルの誤差に相当する誤差率とした．ここで，図 4.4 に示した参照視点と近傍視点において，最も基線長の短いステレオ画像における 1 ピクセルの誤差率は，画像サイズ  $768 \times 512$  ピクセルの場合に約 3.0 %，画像サイズ  $1,536 \times 1,024$  ピクセルの場合に約 1.5 %，画像サイズ  $3,072 \times 2,048$  ピクセルの場合に約 0.8 % である．

図 4.8 より，画像サイズが  $3,072 \times 2,048$  ピクセルの場合は，提案手法の 3 次元復元点数が Goesele らの手法に比べて多いことが確認できる．また，画像サイズが  $768 \times 512$  ピクセル，および， $1,536 \times 1,024$  ピクセルの場合においても，誤差の小さい 3 次元点数は提案手法の方が多い．図 4.9 より，正対応率はすべての画像サイズで提案手法が最も高く，誤対応点の少ない結果となった．また，図 4.10 の誤差率に対する 3 次元点の分布から，提

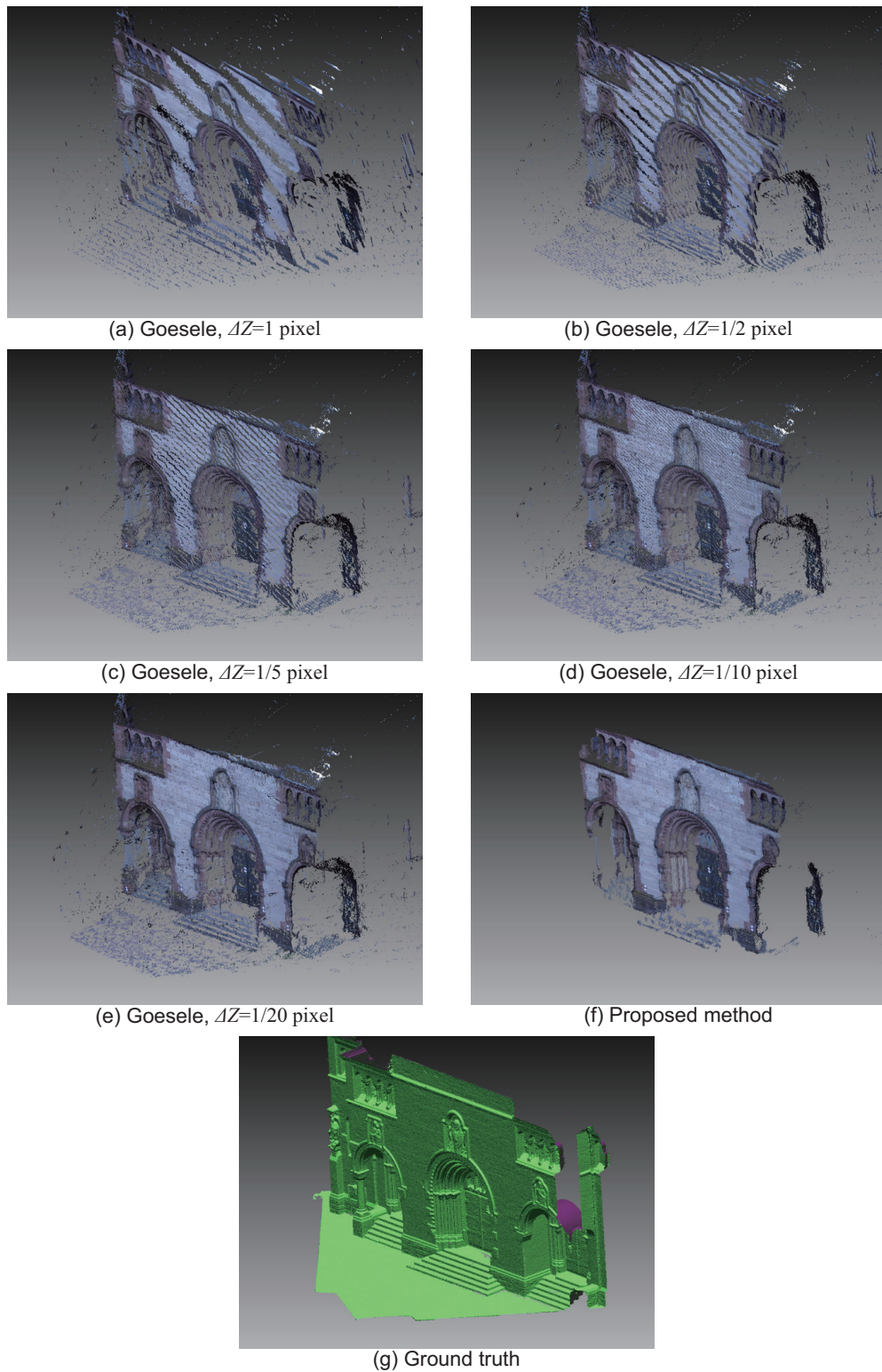


図 4.5 Herz-Jesu-P25 (768 × 512 ピクセル) における 3 次元復元結果: (a)–(e) Goesele からの手法, (f) 提案手法, (g) 真値のメッシュモデル

#### 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

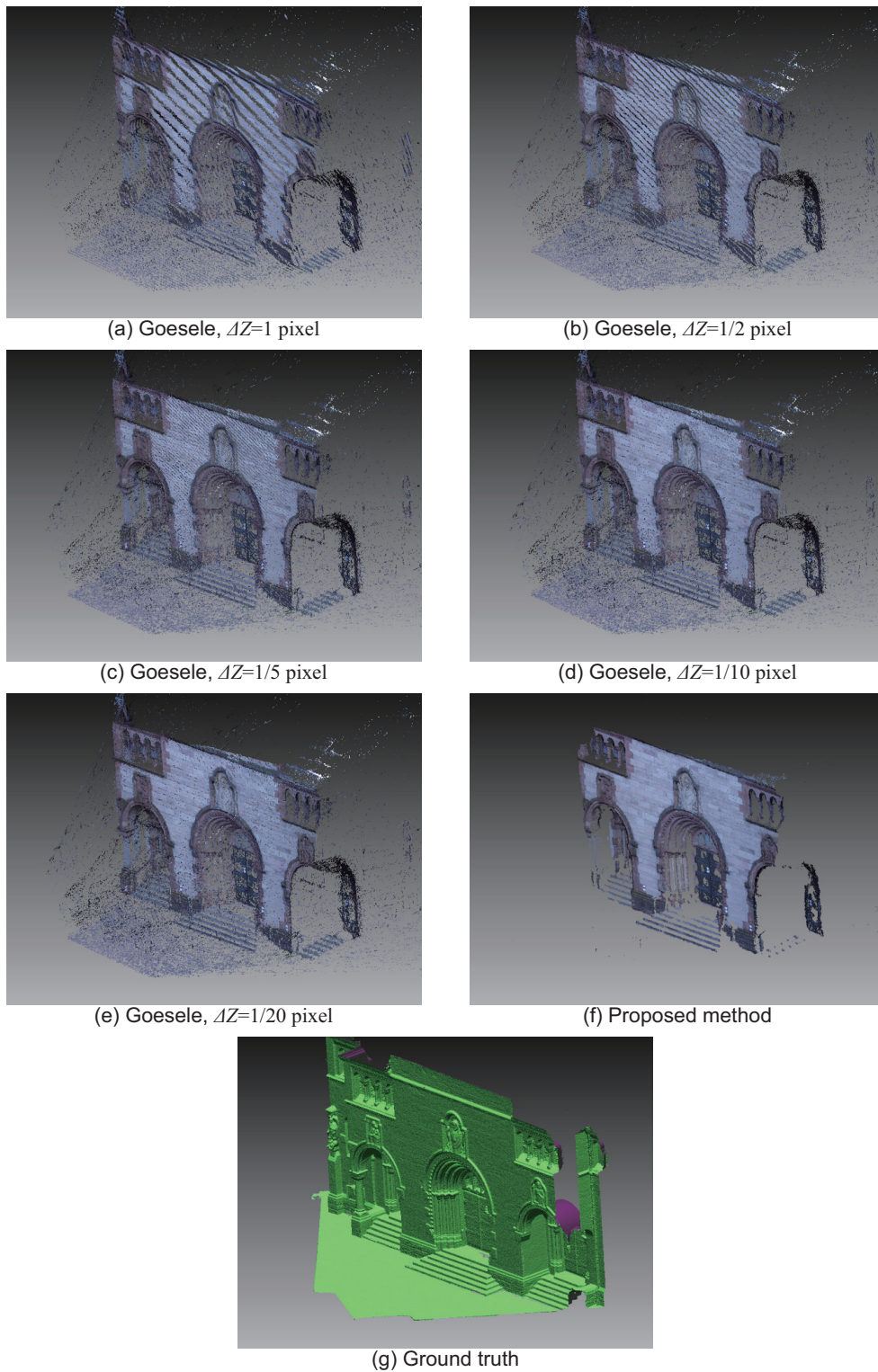


図 4.6 Herz-Jesu-P25 (1,536 × 1,024 ピクセル) における 3 次元復元結果: (a)–(e) Goesele らの手法, (f) 提案手法, (g) 真値のメッシュモデル

### 4.3 提案手法の性能評価

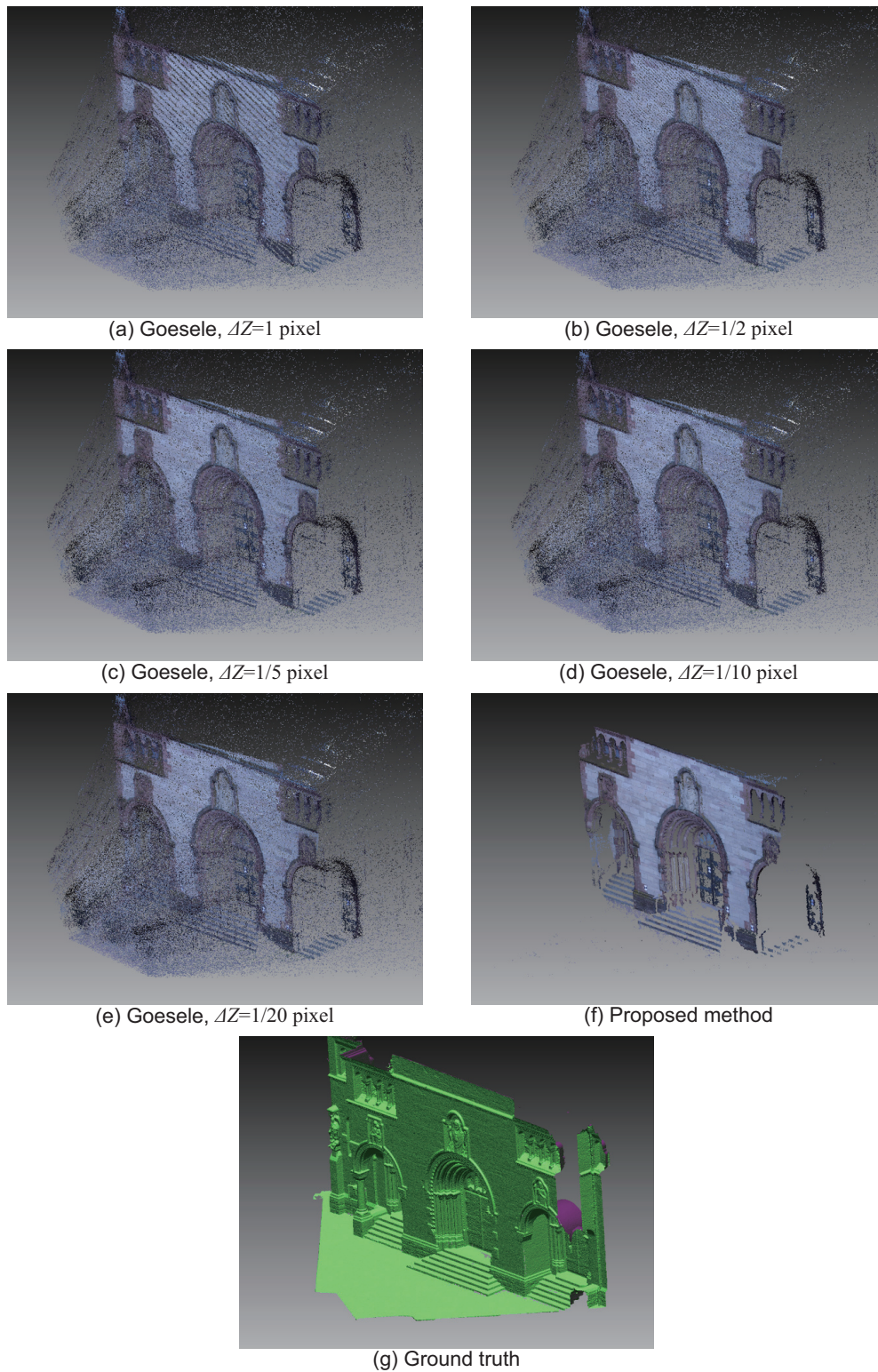


図 4.7 Herz-Jesu-P25 (3,072 × 2,048 ピクセル) における 3 次元復元結果: (a)–(e) Goesele らの手法, (f) 提案手法, (g) 真値のメッシュモデル



## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

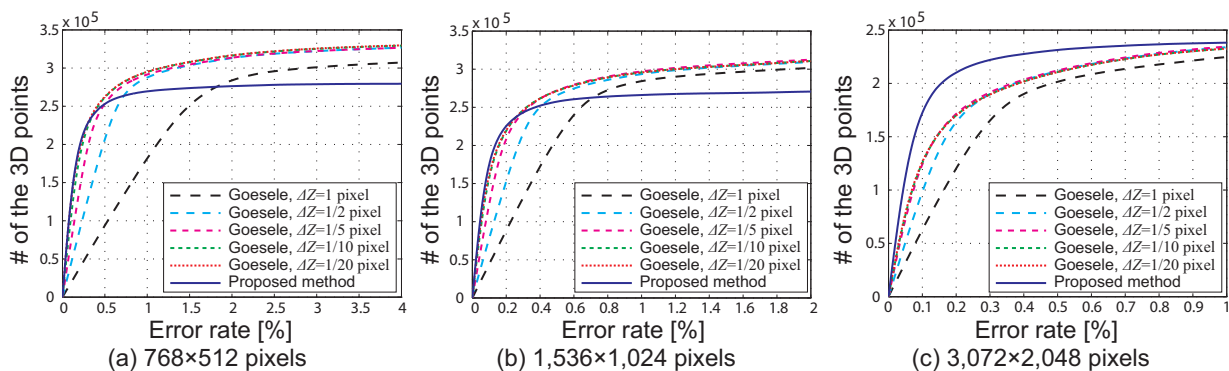


図 4.8 Herz-Jesu-P25 における誤差率と 3 次元点数の関係

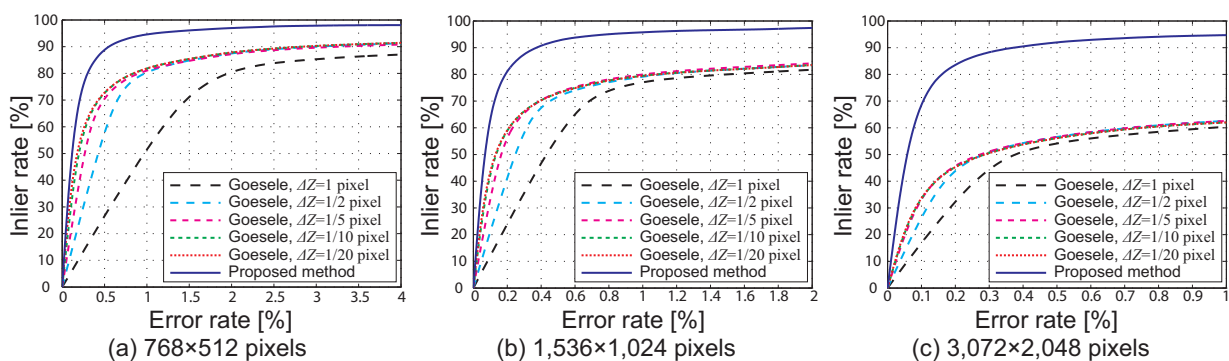


図 4.9 Herz-Jesu-P25 における誤差率と正対応率の関係

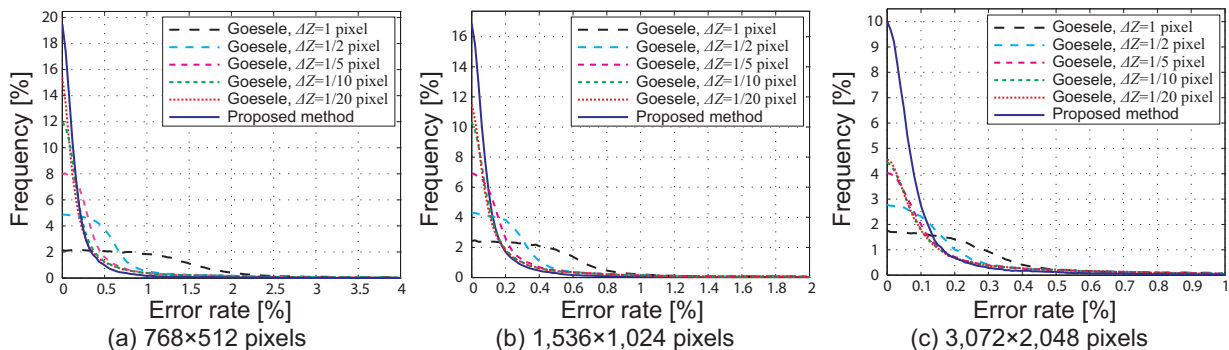


図 4.10 Herz-Jesu-P25 における誤差率に対する 3 次元復元点の分布

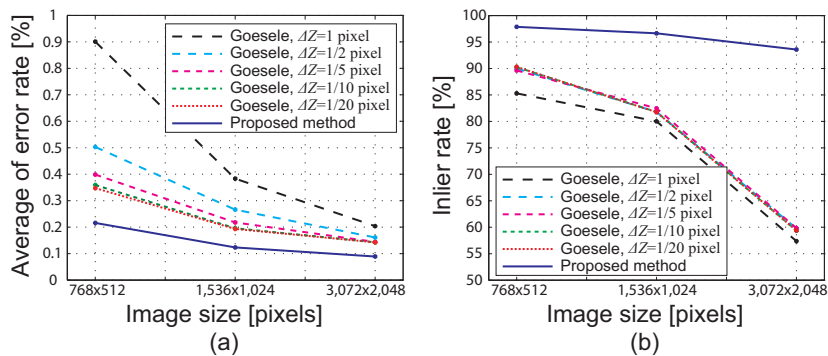


図 4.11 Herz-Jesu-P25 における画像サイズと 3 次元復元精度の関係: (a) 画像サイズに対する平均誤差率の変化, (b) 画像サイズに対する正対応率の変化

案手法は誤差の小さい領域に 3 次元点が集中しており、高精度な 3 次元復元が行えているといえる。そのため、図 4.11 においても、すべての画像サイズに対して、提案手法の誤差が最も小さく、提案手法の正対応率が最も大きくなっている。これらの結果は、図 4.5 から図 4.7 の 3 次元復元結果からも確認できる。Goesele らの手法において、奥行き刻み幅が大きいときは、奥行き分解能が粗く、3 次元復元結果に段差のような誤差がみられる。特に、画像サイズが小さい場合は、ステレオ画像上の 1 ピクセルの視差に相当する奥行きの変化が大きくなるため、この誤差が大きくなる。図 4.11 (a) においても、画像サイズが小さいほど、誤差の差が大きくなっていることが確認できる。また、画像サイズが大きい場合、相対的に画像サイズに対するウィンドウサイズが小さくなり、マッチングウィンドウ内のテクスチャが少なくなる。そのため、図 4.7 からも確認できるように、Goesele らの手法では誤対応が多く発生している。これに対して提案手法では、テクスチャが少ない領域においても高精度なマッチングが行えているといえる。一方で、図 4.5 から図 4.7 より、提案手法では、Herz-Jesu-P25 の地面の部分のように、ステレオ画像間の変形が大きい領域において 3 次元復元が行えていないことが確認できる。この原因として、ステレオ画像間にスキューのような変形が生じた場合、平均 1 次元 POC 関数の計算の際に異なるピーク座標をもつ POC 関数を平均化するため、POC 関数のピーク値が大幅に低くなったためであると考えられる。

次に、データセット Fountain-P11 における 3 次元復元精度を比較する。図 4.12, 図 4.13, 図 4.14 にそれぞれ Fountain-P11 の  $768 \times 512$ ,  $1,536 \times 1,024$ ,  $3,072 \times 2,048$  ピクセルの画像における 3 次元復元結果を示す。Fountain-P11 においても、Herz-Jesu-P25 と同様に、誤差率の閾値とその閾値以下の誤差率をもつ 3 次元点数の関係を図 4.15 に、正対応率の関係を図 4.16 に、誤差率に対する 3 次元点の分布を図 4.17 に示す。また、画像サイズの変化に対する誤差率の平均値の変化を図 4.18 (a) に、正対応率の変化を図 4.18 (b) に示す。このとき、誤差率の平均値および正対応率の計算の際に、誤差率の閾値を最も基線長の短いステレオ画像上で 1 ピクセルの誤差となるように設定した。ここで、図 4.4 に示した参照視点と近傍視点において、最も基線長の短いステレオ画像における 1 ピクセルの誤差率は、画像サイズ  $768 \times 512$  ピクセルの場合に約 1.2 %、画像サイズ  $1,536 \times 1,024$  ピクセルの場合に約 0.6 %、画像サイズ  $3,072 \times 2,048$  ピクセルの場合に約 0.3 % である。

図 4.15 から、画像サイズが  $3,072 \times 2,048$  ピクセルの場合は、Goesele らの手法よりも提案手法の方が 3 次元復元点数が多いが、画像サイズが  $768 \times 512$  ピクセルおよび  $1,536 \times 1,024$  ピクセルの場合は、Goesele らの手法の方が 3 次元点数が多くなった。これは、提案手法では、画像の平行化によって画像の端が復元できないこと、および、地面のようにステレオ画像間の変形が大きい領域において POC 関数のピークが低くなり奥行きが

#### 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

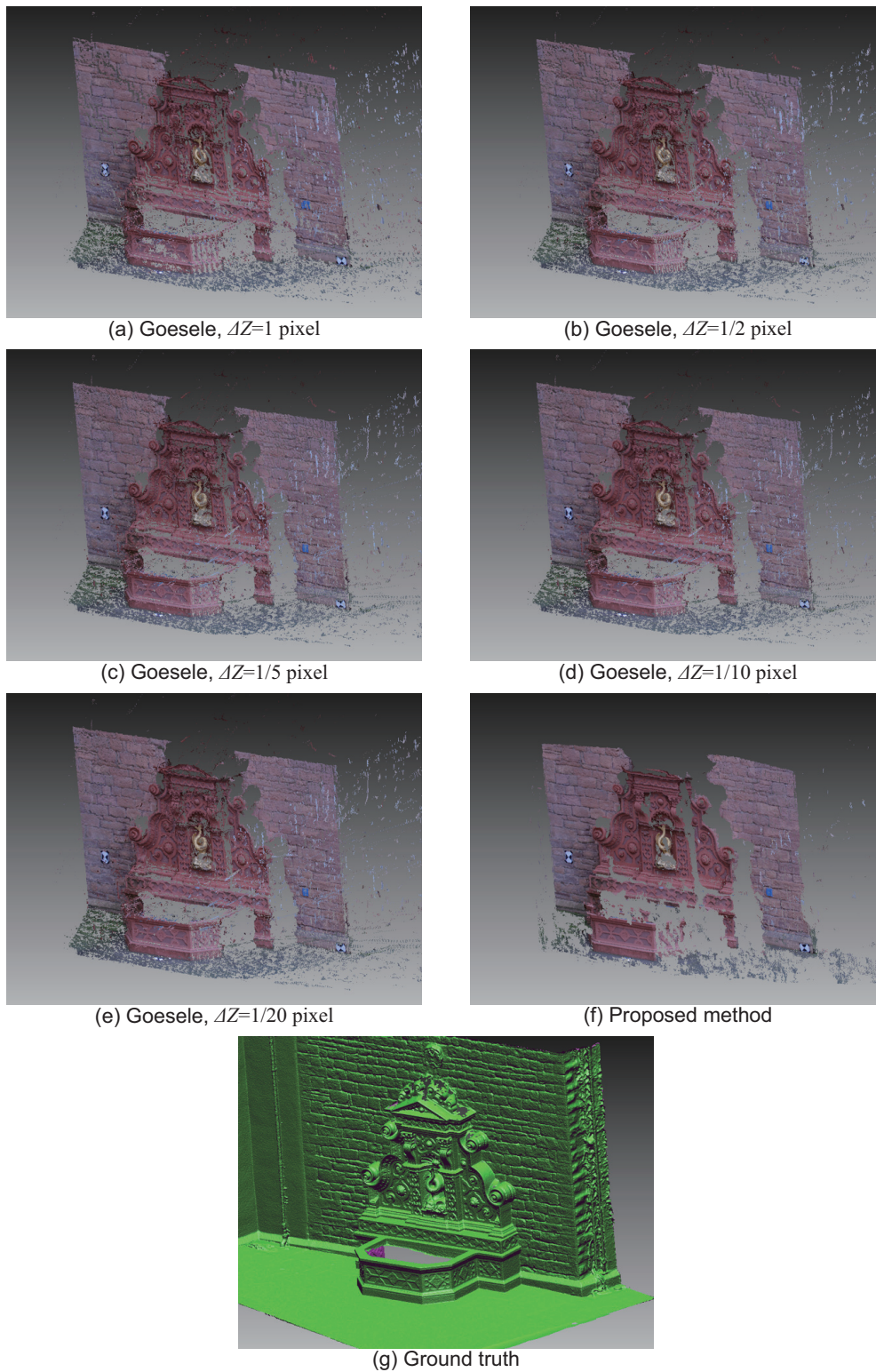


図 4.12 Fountain-P11 (768 × 512 ピクセル) における 3 次元復元結果: (a)–(e) Goesele からの手法, (f) 提案手法, (g) 真値のメッシュモデル

### 4.3 提案手法の性能評価

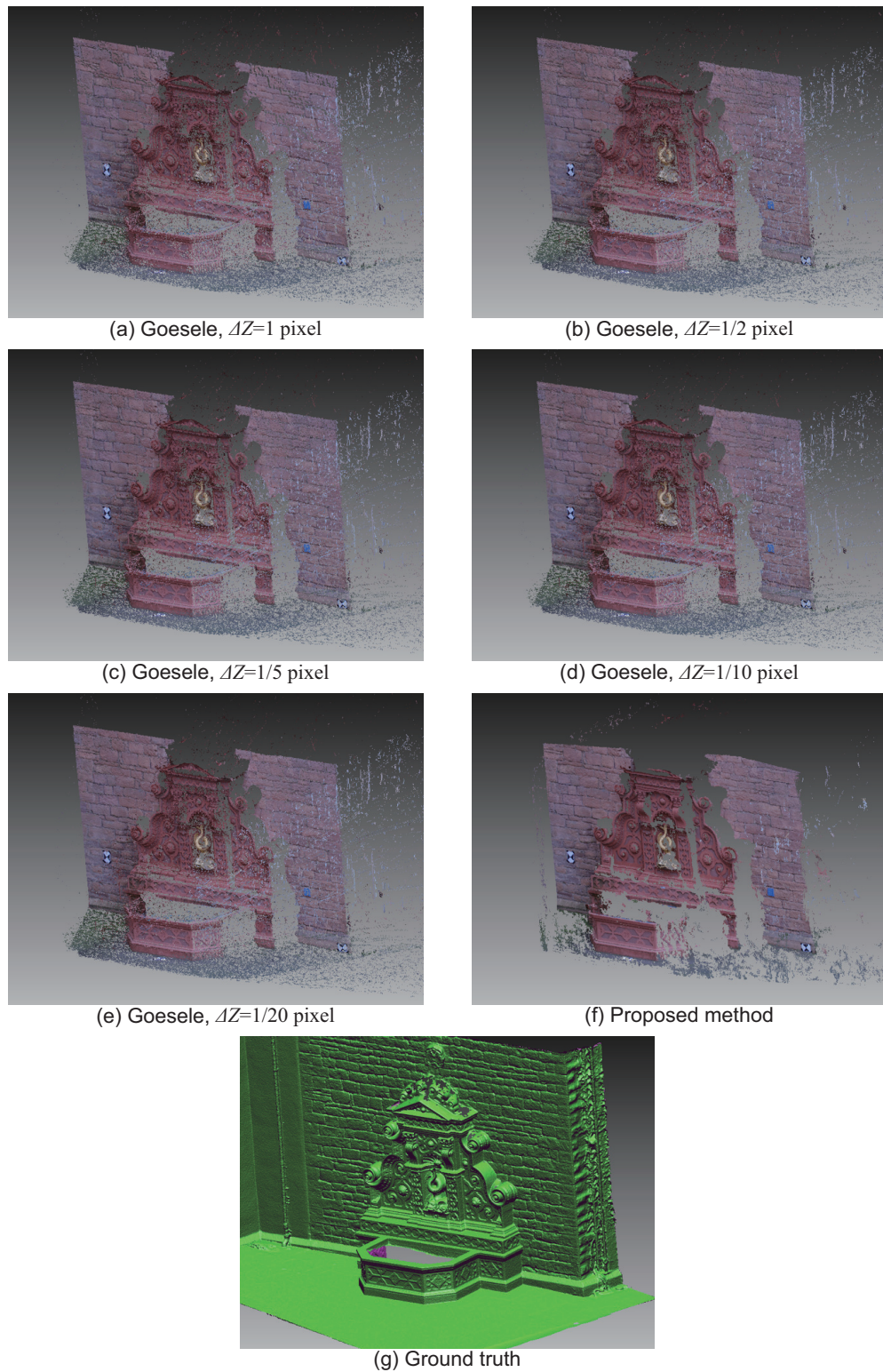


図 4.13 Fountain-P11 (1,536 × 1,024 ピクセル) における 3 次元復元結果: (a)–(e) Goesele らの手法, (f) 提案手法, (g) 真値のメッシュモデル

#### 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

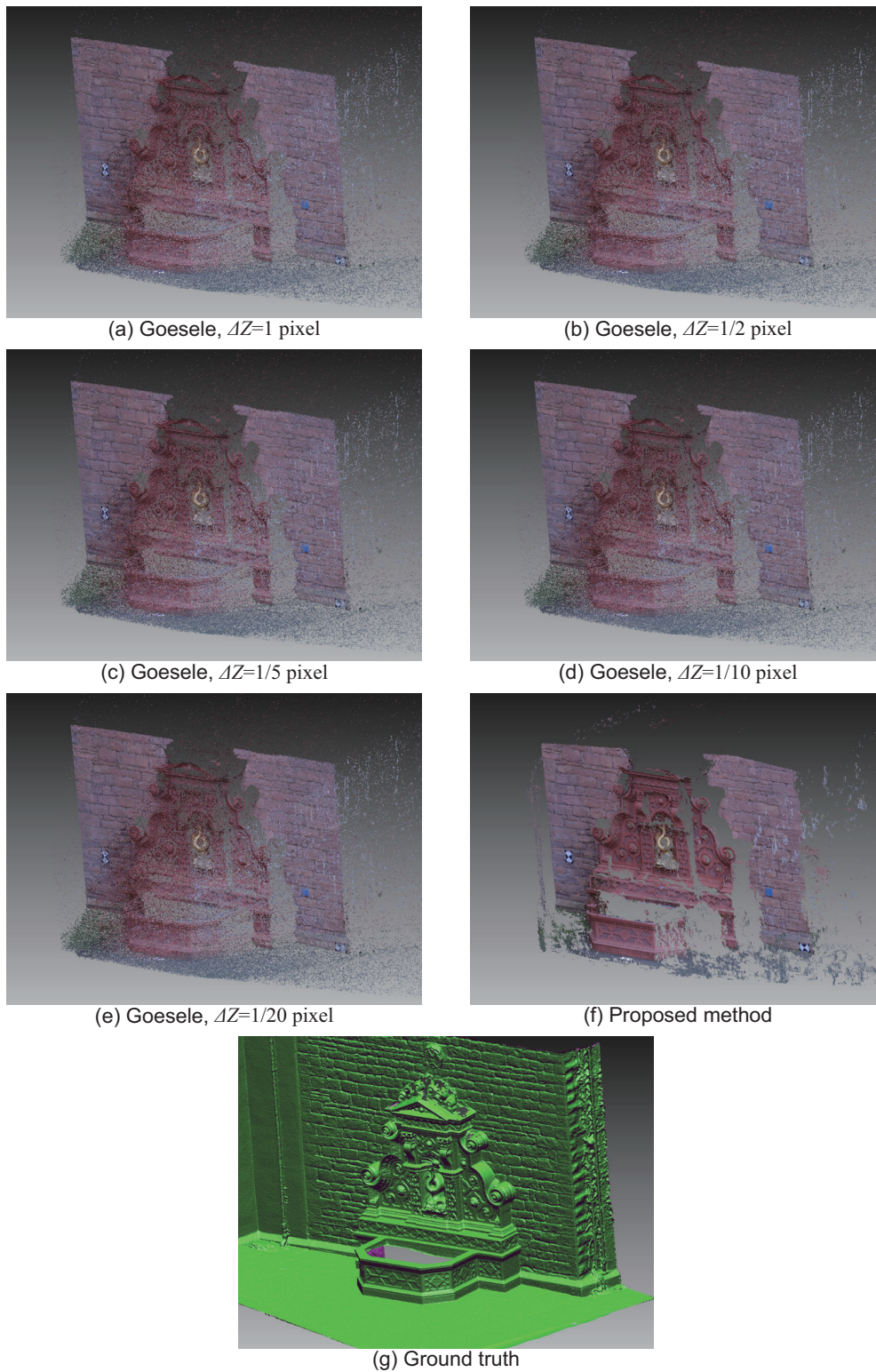


図 4.14 Fountain-P11 (3,072 × 2,048 ピクセル) における 3 次元復元結果: (a)–(e) Goesele らの手法, (f) 提案手法, (g) 真値のメッシュモデル

### 4.3 提案手法の性能評価

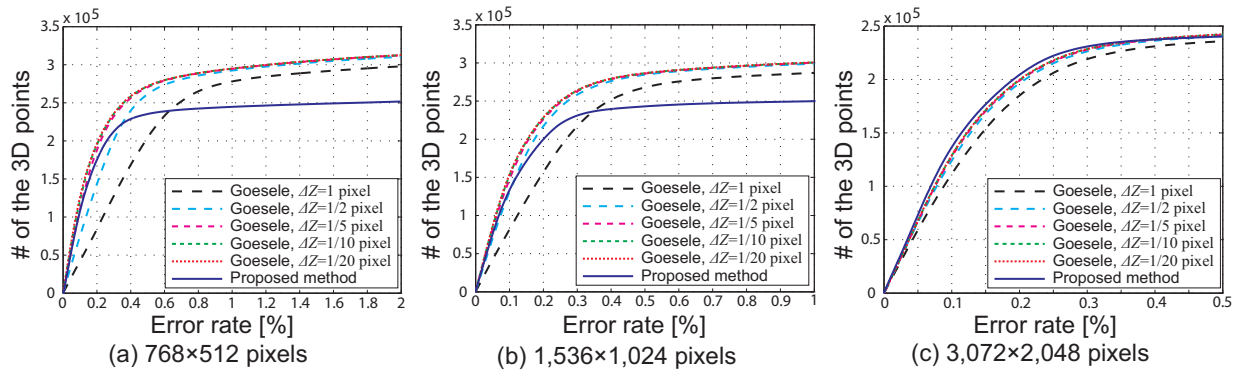


図 4.15 Fountain-P11 における誤差率と 3 次元点数の関係

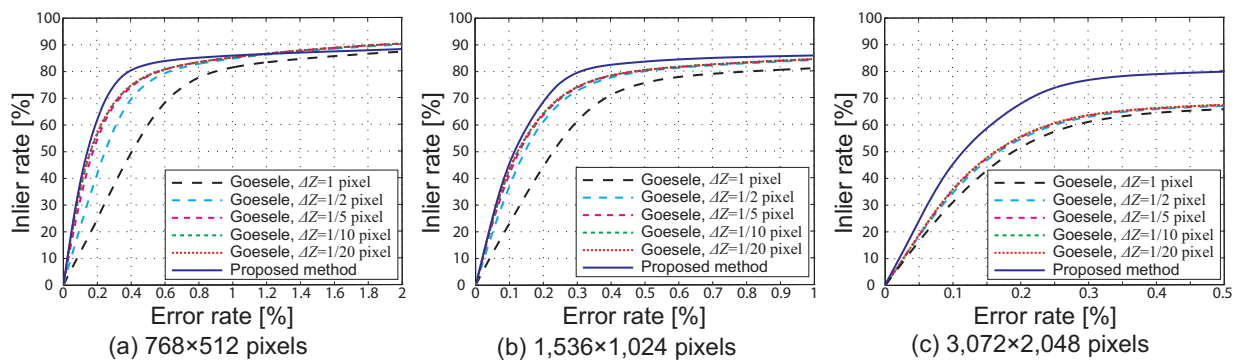


図 4.16 Fountain-P11 における誤差率と正対応率の関係

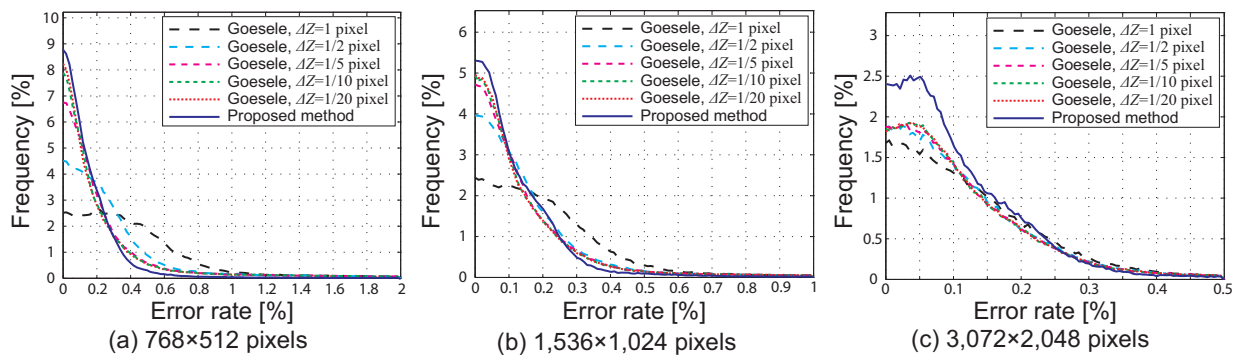


図 4.17 Fountain-P11 における誤差率に対する 3 次元復元点の分布

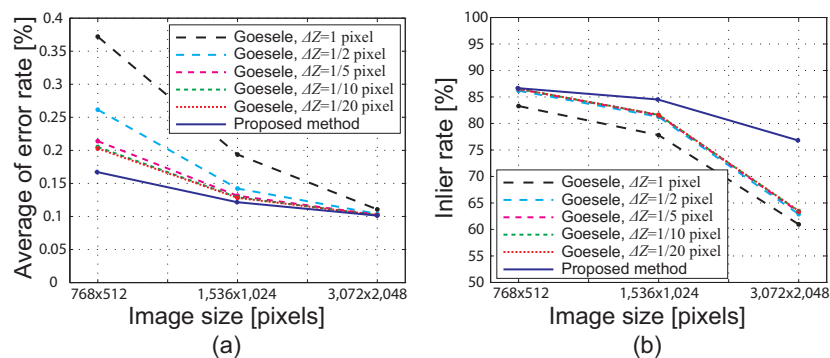


図 4.18 Fountain-P11 における画像サイズと 3 次元復元精度の関係: (a) 画像サイズに対する平均誤差率の変化, (b) 画像サイズに対する正対応率の変化

## 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

推定できなかったことが原因として考えられる。しかしながら，正対応率は，提案手法の方が Goesele らの手法に比べて高く，誤対応の少ない結果となっている。これは，図 4.12 から図 4.14 から確認できる。誤差に対する 3 次元点の分布に関しても，提案手法の方が誤差の小さい点が集中していることが確認できる。そのため，Herz-Jesu-P25 の場合と同様に，図 4.18 では，すべての画像サイズにおいて提案手法の誤差が最も小さく，また，正対応率も最も高くなっている。

以上の結果より，提案手法は，Goesele らの手法に比べて高精度な 3 次元復元が可能である。Goesele らの手法は，最も基本的なプレーンスイーピングに基づくデプスマップ生成手法であり，これに対して周囲の点の情報を用いた高精度化 [20] や法線ベクトルを用いた高精度化 [16], [19] を適用することで，より高精度化かつロバストな 3 次元復元が可能である。これらの高精度化は，提案手法においても適用でき，提案手法のさらなる高精度化が期待できる。特に，法線ベクトルを考慮することで，ステレオ画像間の画像変形が大きい領域においても高精度な復元が行える。これにより，Herz-Jesu-P25 や Fountain-P11 における地面のように，提案手法が復元できなかった領域において，大幅に精度が向上すると予想される。

### 4.3.3 計算コストの評価

まず，提案手法と Goesele らの手法について，参照視点上のある 1 点の奥行き探索に必要な計算コストを比較する。Goesele らの手法では，奥行きの探索範囲をバウンディングボックスにより決定するため，計算コストが入力データに依存する。そこで，Goesele らの手法および提案手法において，同じ探索範囲について奥行き探索をする場合について考える。提案手法では，大きさ  $w$  ピクセルのウィンドウを用いる場合，ステレオ画像上で  $\pm w/4$  ピクセルに相当する奥行きの範囲を 1 回のマッチングで探索することができる。Goesele らの手法においても，提案手法における  $\pm w/4$  ピクセルに相当する奥行きの探索範囲について，計算コストを求める。ここで， $w$  は，4.3.1 節でのべたように，16 ピクセルとする。提案手法と Goesele らの手法における計算コストを表 4.1 に示す。表 4.1 より，Goesele らの手法では，奥行き変化の刻み幅  $\Delta Z$  が細かいほど，演算回数が増加することが確認できる。これに対して，提案手法では，Goesele らの手法と同じ範囲の奥行き探索を行う場合，Goesele らの手法における  $\Delta Z = 1$  ピクセルから  $\Delta Z = 1/2$  ピクセルと同程度の演算回数が必要である。一方で，4.3.2 節で示したように，提案手法では，Goesele らの手法の  $\Delta Z = 1/20$  の場合に比べても，高精度な 3 次元復元が可能である。Goesele らの手法においても， $\Delta Z$  を大きく設定することで，計算コストの削減が可能であるが，こ

表 4.1 1 点の奥行き探索に必要な演算回数

	Additions	Multiplications	Divisions	Square roots
Goesele, $\Delta Z = 1$ pixel	10,692	4,878	162	162
Goesele, $\Delta Z = 1/2$ pixel	21,384	9,756	324	324
Goesele, $\Delta Z = 1/5$ pixel	53,460	24,390	810	810
Goesele, $\Delta Z = 1/10$ pixel	106,920	48,780	1,620	1,620
Goesele, $\Delta Z = 1/20$ pixel	213,840	97,560	3,240	3,240
Proposed method	8,960	8,096	576	288

表 4.2 実験に用いた計算機の仕様

CPU	Intel(R) Core(TM)2 Extreme CPU X9650 3.00 GHz
Memory	3 GB
OS	Vine Linux 4.2
C compiler	gcc version 3.3.6

の場合、3次元復元精度が低下する。これに対して、提案手法では、Goesele らの手法に比べて、3次元復元精度を低下させることなく、少ない計算コストの奥行き探索が行える。

次に、実際に Goesele らの手法と提案手法を用いて、デプスマップを生成する際にかかった処理時間を比較する。表 4.1 では、同じ奥行き探索に必要な計算コストを示したが、ここでは、デプスマップ生成の処理全体の時間を測定する。処理時間の測定に用いた計算機の仕様を表 4.2 に、プログラムの処理時間を表 4.3 および表 4.4 に示す。表 4.3 および表 4.4 より、提案手法の処理時間は、Herz-Jesu-P25 および Fountain-P11 のどちらのデータセットにおいても、すべての解像度で Goesele らの手法の処理時間よりも短いことが確認できる。特に、画像サイズが大きい場合、Goesele らの手法と提案手法の処理時間の差が大きい。実験結果をみると、提案手法を用いることで、Goesele らの手法の  $\Delta Z = 1/20$  ピクセルの場合と比べて、Herz-Jesu-P25 の  $3,072 \times 2,048$  ピクセルの場合に約 90 倍、Fountain-P11 の  $3,072 \times 2,048$  ピクセルの場合に約 180 倍の高速化を実現した。また、Goesele らの手法の  $\Delta Z = 1$  ピクセルの場合と比べても、Herz-Jesu-P25 の  $3,072 \times 2,048$  ピクセルの場合に約 4 倍、Fountain-P11 の  $3,072 \times 2,048$  ピクセルの場合に約 9 倍の高速化を実現し



#### 第 4 章 位相限定相関法に基づく多視点画像からの 3 次元復元

表 4.3 Herz-Jesu-P25 における処理時間 [sec]

Image size [pixels]	768 × 512	1,536 × 1,024	3,072 × 2,048
Goesele, $\Delta Z = 1$ pixel	524	1,307	2,634
Goesele, $\Delta Z = 1/2$ pixel	1,279	2,585	5,256
Goesele, $\Delta Z = 1/5$ pixel	2,557	5,214	13,024
Goesele, $\Delta Z = 1/10$ pixel	5,081	12,901	25,999
Goesele, $\Delta Z = 1/20$ pixel	12,694	25,946	52,376
Proposed method	369	475	574

表 4.4 Fountain-P11 における処理時間 [sec]

Image size [pixels]	768 × 512	1,536 × 1,024	3,072 × 2,048
Goesele, $\Delta Z = 1$ pixel	1,229	2,511	6,277
Goesele, $\Delta Z = 1/2$ pixel	2,457	6,290	12,636
Goesele, $\Delta Z = 1/5$ pixel	6,135	12,559	25,310
Goesele, $\Delta Z = 1/10$ pixel	12,292	24,806	62,973
Goesele, $\Delta Z = 1/20$ pixel	24,439	62,244	125,557
Proposed method	546	614	690

た．画像サイズが大きい場合，ステレオ画像上で 1 ピクセルの視差に相当する奥行きの変化が小さくなる．そのため，Goesele らの手法では，画像サイズに比例してマッチング回数が多くなり，処理時間が増加する．これに対して，提案手法では，画像ピラミッドを用いた階層的探索を適用することで，画像サイズに対する処理時間の増加を抑えている．

以上の結果より，提案手法は，Goesele らの手法に比べて少ない計算コストでデプスマップが生成可能である．Goesele らの手法は，最も基本的なプレーンスイーピングに基づくデプスマップ生成手法であり，これに対して粗密戦略による高速化 [22], [30] を適用することで，より実用的な時間で 3 次元復元を行える．この粗密戦略による高速化は，提案手法においても適用でき，提案手法と粗密戦略を組み合わせることで，より計算コストの少ないデプスマップ生成が行えると予想される．

### 4.3.4 むすび

本章では，位相限定相関法に基づく多視点画像からの 3 次元復元について述べた．まず，提案手法で重要となる正規化視差について説明し，正規化視差に基づく POC 関数の統合について述べた．次に，POC に基づく多視点画像からのデプスマップ生成手法について述べた．さらに，実際の多視点画像を用いた実験によって，提案手法の評価を行った．

## 第 5 章

### 結言

本論文では、位相限定相関法に基づく多視点画像からの 3 次元復元手法について述べた。まず、多視点画像からの 3 次元復元の分類とそれぞれの手法の特徴を説明した。次に、提案手法において重要となる 1 次元 POC を用いた画像マッチングとその高精度化手法について述べた。そして、POC を用いた多視点画像からの 3 次元復元手法を提案し、実験を通して、従来のプレーンスライピングに基づくデプスマップ生成に比べ、少ない計算コストで高精度な 3 次元復元が行えることを示した。

以下に各章の内容を要約して述べる。

第 1 章においては、研究の背景として、3 次元復元技術の概要およびその応用について述べた。そして、多視点画像からの 3 次元復元について、その特徴をまとめ、現在の多視点画像からの 3 次元復元手法における課題を述べた。

第 2 章においては、多視点画像からの 3 次元復元に関する基礎的考察を行った。はじめに、多視点画像からの 3 次元復元に用いる多視点画像とカメラパラメータの取得方法について述べた。次に、多視点画像からの 3 次元復元の分類とそれぞれの手法の特徴について述べた。そして、多視点画像からの 3 次元復元手法の 1 つであるデプスマップ統合に基づく手法について説明した。さらに、従来のデプスマップ生成に良く用いられるプレーンスライピングに基づくデプスマップ生成手法の原理を説明した。

第 3 章においては、1 次元 POC に基づく画像マッチング手法について述べた。まず、1 次元 POC の原理について説明した。そして、1 次元 POC 関数を用いて平行移動量推定を行う際に重要となる各種高精度化手法について述べた。

第 4 章においては、位相限定相関法に基づく多視点画像からの 3 次元復元手法を提案した。まず、提案手法で重要となる正規化視差の概念について説明し、正規化視差に基づく POC 関数の統合について述べた。そして、1 次元 POC を用いた画像マッチングに基づく奥行き推定手法と画像ピラミッドを用いた階層的探索を組み合わせることで、POC に

基づく多視点画像からのデプスマップ生成手法を提案した。さらに、性能評価実験を通して、提案手法により少ない計算コストで高精度な 3 次元復元が可能であることを示した。

本論文で提案した位相限定相関法を用いた多視点画像からの 3 次元復元手法は、最も基本的な 3 次元復元手法の 1 つであるプレーンスweepingに基づく手法と共通する部分が多い。一方で、プレーンスweepingに基づく手法に対して、多くの高精度化手法 [16], [19], [20] や高速化手法 [22], [30] が提案されている。これらの高精度化手法および高速化手法は、提案手法に対しても適用可能であり、提案手法のさらなる改善が期待できる。

今後は、上記のような高精度化や高速化の検討を行う予定である。具体的には、3 次元点の奥行きだけでなく法線ベクトルも考慮することで、より高精度かつロバストな 3 次元復元を実現する。このとき、法線ベクトルを考慮した POC に基づく画像マッチング手法として Scaled Window-POC (SW-POC) が非常に有効であると考えられる [28], [44], [45]。また、提案手法に対して粗密戦略を適用することで、より高速な 3 次元復元の実現が期待できる。このような高精度化および高速化により提案手法を改善することで、自由度の高い撮影で得られた多視点画像から高精度な 3 次元データを高速に生成することが可能となり、より実用的な 3 次元復元システムが構築できると考えている。今後の展望として、提案手法を用いた多視点画像からの 3 次元復元システムを構築し、文化財のデジタルアーカイブや工業製品の部品検査、医療分野での人体の 3 次元復元などの応用を目指す。特に、文化財のデジタルアーカイブでは、狭い部屋の中で撮影するなど、撮影環境が限定されたり、建築物などの比較的大きい対象を撮影したりするため、従来の能動型の 3 次元復元手法が適用困難な場合があり、撮影の制限が少ない多視点画像からの 3 次元復元は非常に有効であると考えられる。

## 付録 A

# Scaled Window-POC に基づく画像変形にロバストな高精度対応付け手法

## A.1 まえがき

本章では，Scaled Window-POC に基づく，2 視点ステレオ画像間の対応付け手法について述べる．画像の対応付けは，画像センシング，映像信号処理，コンピュータビジョンなどの分野で重要になる基本技術である．特に，ステレオビジョンに基づく 3 次元復元では，対応付け精度が 3 次元復元精度に影響するため，ステレオ画像間の高精度な対応付けが必要である [3], [4]．2 視点ステレオ画像間の密な対応付け手法は，動的計画法やグラフカットなどの最適化に基づく手法と，ウィンドウマッチングに基づく手法に分類される [3]．最適化に基づく手法は，画像間の対応関係と周辺画素の関係からコスト関数を定義し，そのコスト関数を最小化することで，画像間の対応付けを行う．この手法は，物体境界やオクルージョン領域の影響に対して安定であるが，複雑な最適化を行うため計算コストが大きいという欠点がある．また，グラフカットのラベリング処理などにより最適化を行うため，サブピクセル精度の対応付けに適用することが困難である．一方，ウィンドウマッチングに基づく手法は，対応候補点近傍から切り出した小領域のウィンドウに対して，領域ベースの画像マッチング手法を適用することで，画像間の対応付けを行う．この手法は，最適化に基づく手法に対して計算コストが少なく，また，サブピクセル精度の対応付けが容易に行えるといった利点がある．そのため，ウィンドウマッチングに基づく手法は，3 次元復元のようにサブピクセル精度の対応付けが重要となる応用に適している．

一般に，ウィンドウマッチングに基づく対応付け手法では，ウィンドウ間の変形を平行移動のみと仮定しているため，ステレオ画像間の変形が小さくなるように，基線長の短いステレオカメラを用いることを想定している．一方で，基線長の長いステレオカメラを

用いる場合、対応付け誤差が3次元復元結果に与える影響が小さくなり、3次元復元精度が向上するという利点がある [3], [4]。しかし、基線長が長くなると、ステレオカメラで撮影した左右の画像間で変形が大きくなり、正確な対応付けが困難になるという問題がある。従来の基線長が長いステレオ画像間の対応付け手法の1つとして Bradley らの手法 [19] がある。Bradley らの手法では、ウィンドウの拡大縮小により、ウィンドウ間に生じている画像変形を軽減することで、基線長の長いステレオ画像間の対応付けを行う。一方で、Bradley らの手法では、拡大縮小率や視差を変化させながらマッチングを繰り返すため、非常に計算コストが大きく、視体積交差法などを用いて、探索の範囲を限定する必要がある。そのため、一般的な2視点ステレオ画像に対しては、探索の範囲を限定することができず、Bradley らの手法を適用することが難しかった。また、サブピクセルレベルの対応付け精度と処理時間はトレードオフの関係にあり、高精度な対応付けを行うためには、さらに計算量が多くなるといった問題がある。

これに対して、本章では、Scaled Window-POC (SW-POC) に基づく対応付け手法を提案する。SW-POC に基づく対応付け手法では、Bradley らの Scaled Window Matching (SWM) の考え方を、1次元 POC に基づく対応付け手法 [27] に適用し、Bradley らの手法に比べ少ない計算コストで基線長の長いステレオ画像間の対応付けを行う。SW-POC に基づく対応付け手法は、基線長の長いステレオ画像間においても、高精度な対応付けを行えるため、基線長の拡大による3次元復元精度の向上が期待できる。実験を通して、SW-POC に基づく手法を用いることで、基線長の長いステレオ画像に対してもロバストに対応付けを行えることを示す。

## A.2 SW-POC に基づく対応付け手法

本節では、SW-POC に基づくステレオ画像間の変形にロバストな対応付け手法を提案する。提案手法では、ステレオ画像間に生じている変形を軽減するようにウィンドウを設定することで、基線長が長いステレオ画像間においてもロバストに対応付けを行う。以下では、平行化された2視点のステレオ画像を入力として、片方の画像上に設定された基準点に対応するもう一方の画像上の対応点を求める問題を考える。

### A.2.1 Scaled Window-POC (SW-POC)

SW-POC とは、復元する物体の表面形状に合わせてマッチングウィンドウを拡大縮小させることで、ステレオ画像間の変形に対応した1次元 POC に基づく画像マッチング手

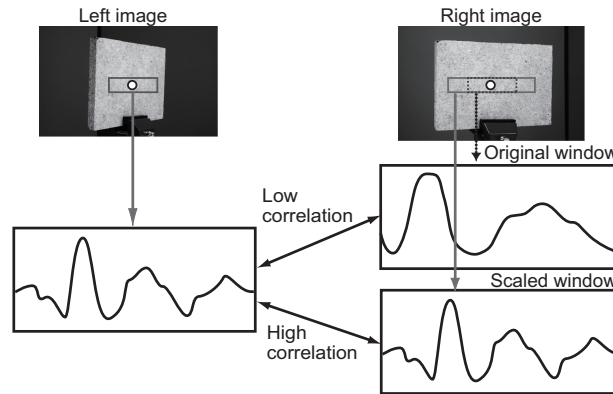


図 A.1 Scaled Window-POC

法である．ステレオ平行化後の画像では，基準点と対応点の垂直座標が等しくなる [3]．ここで，ある垂直座標上の基準点に注目すると，その対応点はすべて同じ垂直座標上に存在し，その垂直座標上の画像変形は水平方向のみであると考えられる．そこで，ステレオ画像間の変形を局所的なスケール変化に近似して画像マッチングを行う．図 A.1 に示すように，マッチングウィンドウを拡大縮小させることで，左右のマッチングウィンドウ間に生じている画像変形を軽減させる [19]．この拡大縮小したマッチングウィンドウに対して，1次元 POC を用いた画像マッチング手法 (3 章) を適用する．これにより，基線長の長いステレオ画像のように，ステレオ画像間の変形が大きい場合においても，高精度な平行移動量推定が可能になる．この SW-POC を用いた平行移動量推定手法と，画像ピラミッドを用いた階層的探索を組み合わせることで，ステレオ画像間の画像変形にロバストな対応付けを実現できる．

### A.2.2 拡大縮小率の推定

SW-POC において，左右のマッチングウィンドウ間の拡大縮小率  $s_{\text{win}}$  は，物体の形状 (法線ベクトル  $\mathbf{n}$ ) と物体までの距離に依存する．図 A.2 に示すように，ある 3 次元点  $M = (X, Y, Z)$  に注目したとき，その点の左カメラ画像におけるウィンドウサイズ  $w_1$  と右カメラ画像におけるウィンドウサイズ  $w_2$  は拡大縮小率  $s_{\text{win}}$  を用いて次式で表される [19]．

$$w_2 = s_{\text{win}} w_1 \quad (\text{A.1})$$

$$s_{\text{win}} = \frac{\cos \psi_1 \cos \phi_2}{\cos \psi_2 \cos \phi_1} \quad (\text{A.2})$$

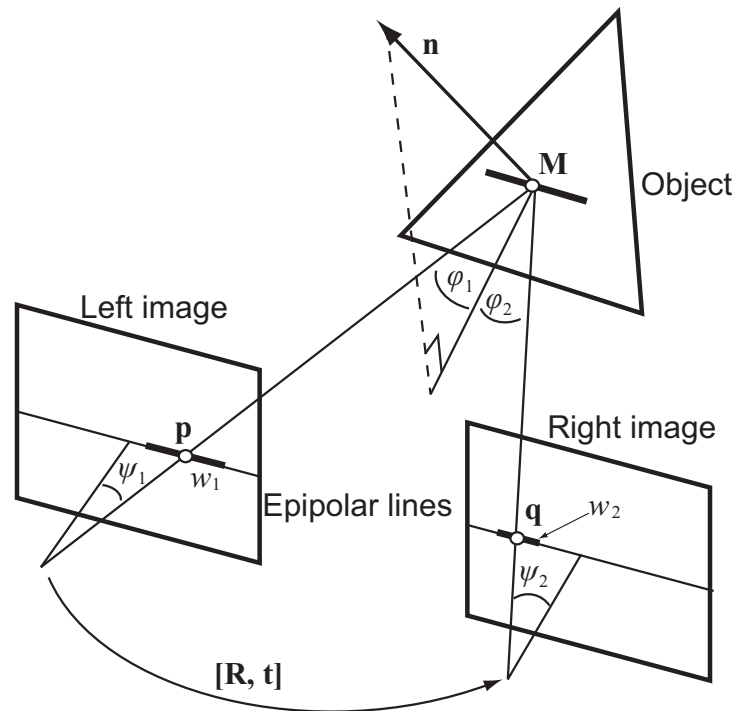


図 A.2 マッチングウィンドウ間の拡大縮小率

ここで、 $\psi_i$  は、カメラ  $i$  における光軸と視線のなす角を表し、カメラ  $i$  の内部パラメータと外部パラメータおよび 3 次元点  $M$  から求められる。また、 $\phi_i$  は、カメラ  $i$  における視線と法線ベクトル  $\mathbf{n}$  のエピポーラ平面への投影ベクトルのなす角を表し、2 つのカメラの外部パラメータと 3 次元点  $M$  およびその点における法線ベクトル  $\mathbf{n}$  から求められる。一般に、ステレオ画像から 3 次元復元を行う場合、復元する物体の形状は未知であることが多い。そのため、式 (A.2) のように物体の形状から拡大縮小率  $s_{\text{win}}$  を求めることはできない。そこで、SW-POC に基づく対応付け手法では、1 次元 POC 関数の相関ピークの高さが 2 つの信号の類似度に相当することを利用し、拡大縮小率  $s_{\text{win}}$  を推定する。 $s_{\text{win}}$  を  $s_{\text{min}}$  から  $s_{\text{max}}$  の範囲で  $\Delta s$  ずつ変化させながら SW-POC を用いてマッチングを繰り返し、最も相関ピークが高くなる時の対応点を真の対応点とする。これにより、物体の形状が未知の場合においても SW-POC を用いることが可能となる。一方で、この手法は、拡大縮小率の試行回数だけマッチングを繰り返すので、計算コストが増加する問題がある。そこで、計算コスト削減のため、拡大縮小率の推定は、対応点探索における画像ピラミッドの最上位層  $h_{\text{max}}$  のみで行う。

また、1 次元 POC に基づく対応付け手法 [27] では、画像ピラミッドの最上位層にお



いて基準点  $p_{h_{\max}}$  と対応点  $q_{h_{\max}}$  の水平座標が一致すると仮定している．しかし，基線長の長いステレオ画像では，ステレオ画像間の変形が大きくなるため，この仮定が必ずしも成り立たない．そこで，SW-POC を用いた対応付けでは，画像ピラミッドの最上位層において，拡大縮小率  $s_{\text{win}}$  だけでなく，初期対応点  $q_{h_{\max}}$  を  $p_{h_{\max}} \pm (D, 0)$  の範囲で変化させながらマッチングを行い，最も相関ピークの高い結果を，最上位層におけるマッチング結果とする．

### A.2.3 粗密戦略を用いた計算コストの削減

A.2.2 節の手法では，画像ピラミッドの最上位層において，拡大縮小率  $s_{\text{win}}$  と初期対応点  $q_{h_{\max}}$  を変化させるため，1次元 POC に基づく対応付け手法に比べ計算コストが大きくなるという問題がある．さらに，SW-POC では，拡大縮小率  $s_{\text{win}}$  が式 (A.2) で与えられる値に近いほど対応付け精度が向上するのに対して，A.2.2 節の手法では， $s_{\text{win}}$  を式 (A.2) で与えられる値に近づけるために，刻み幅  $\Delta s$  を非常に小さくする必要があり，計算コストが非常に大きくなってしまふ．そこで，A.2.2 節の手法と粗密戦略を組み合わせることで，計算コストを削減する．まず，疎な基準点群について A.2.2 節の手法を用いて対応付けを行う．次に，求めた対応点を用いて 3次元復元を行い，物体の粗い形状を求める．そして，密な基準点群について，それぞれの基準点における適切な拡大縮小率  $s_{\text{win}}$  と初期対応点  $q_{h_{\max}}$  を物体の粗い形状から計算し，この  $s_{\text{win}}$  と  $q_{h_{\max}}$  を用いて対応付けを行う．これにより，疎な点群の対応付けについては，1次元 POC に基づく手法に比べて計算コストが大きくなるが，密な点群については，1次元 POC に基づく手法と同じ計算コストで対応付けを行うことが可能である．

### A.2.4 物体の形状を考慮した平均 POC 関数の計算

SW-POC では，左右の画像から抽出した 2つの 1次元画像信号間の変形を拡大縮小と平行移動のみであると仮定している．平行化された画像において，エピポーラ線上の 1ラインのみに注目すると，この仮定は局所的に良い近似で成り立つ．しかし，実際は，3.2 節で述べたように，信頼性の高い相関ピーク座標の推定を行うために，1組の基準点と対応点ペアについて複数のラインで平均 1次元 POC 関数を計算する必要がある [27]．そのため，左右のウィンドウ間でスキューのような画像変形があると，ウィンドウ間の画像変形を拡大縮小と平行移動のみで近似できない．A.2.1 節から A.2.3 節で述べた手法では，ある程度のスキューに対してもロバストに対応点を求めることができるが，スキューが大きくなることで，対応付け誤差が増加することを実験的に確認している．これは，異なった

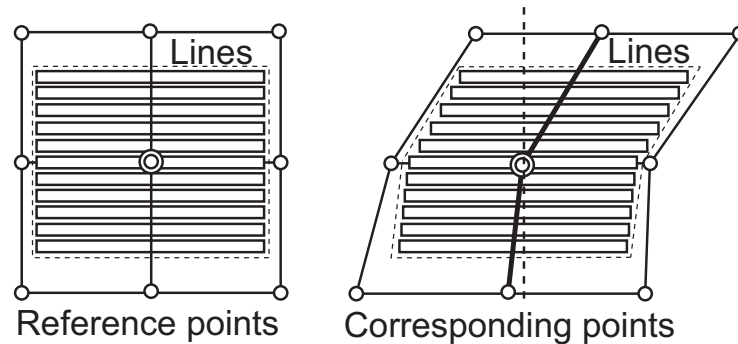


図 A.3 物体の形状を考慮したウィンドウ

ピーク座標を有する複数の 1 次元 POC 関数から平均 1 次元 POC 関数を計算することで、ピーク座標を用いた平行移動量推定に誤差が生じたためである。

そこで、A.2.1 節から A.2.3 節で述べた手法により一度対応点群を求めた後に、ラインごとのピーク座標のずれを軽減するように各ラインを平行移動させて平均 1 次元 POC 関数を計算し、対応点群を計算し直すことで誤差を減少させる。この時、各ラインごとの平行移動量は物体の形状から計算する。基準点が格子状に配置されている場合、対応点群は物体の形状を反映する形で求まる。よって、図 A.3 に示すように、現在注目している点の上下の対応点を直線で結び、その直線に合わせて各ラインを平行移動させることで、スキューのような画像変形に対してもロバストかつ高精度な対応付けを行うことが可能となる。一方で、注目点の周囲の情報を用いると、その周囲の点に誤対応があった場合、誤対応の影響により注目点の対応付け誤差が増加する。そこで、誤対応点の影響を少なくするため、対応点群の更新の際に、更新後の対応点の方が平均 1 次元 POC 関数の相関ピークが高くなる場合のみ対応点を更新する。

### A.3 性能評価実験

本節では、2 台のカメラからなるステレオビジョンシステムを構築し、実際に物体の 3 次元復元を行うことで、従来手法および SW-POC に基づく対応付け手法の 3 次元復元精度と計算コストを比較する。ここでは、従来手法として、1 次元 POC に基づく対応付け手法 [27] と Bradley らの手法 [19] を用いる。

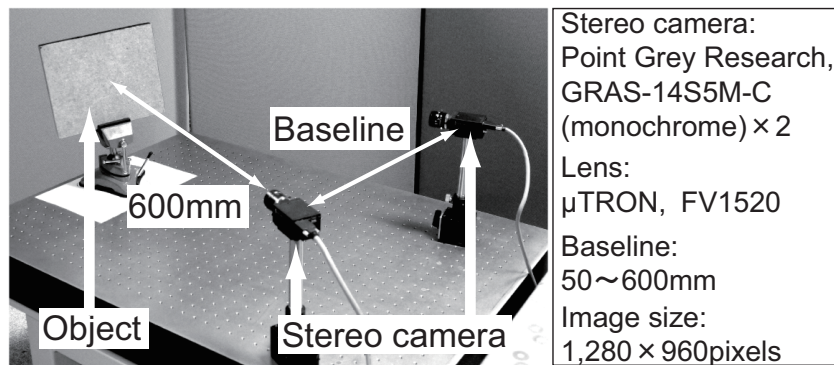


図 A.4 実験で用いたステレオビジョンシステム

### A.3.1 実験方法

実験に用いるステレオビジョンシステムを図 A.4 に示す．復元対象は，形状が既知である平板と球を用いる．この復元対象をカメラから約 600 mm の距離に配置し，2 台のカメラ間の基線長を変化させながら撮影を行う．このとき，2 台のカメラのうち，左カメラを固定し，右カメラのみを移動させ，基線長が 50 mm から 600 mm の間で変化するようにする．実際に撮影された画像を図 A.5 に示す．このステレオ画像について，それぞれの基線長で左右の画像を平行化し，左カメラ画像における復元対象上に配置された基準点に対して対応付けを行う．本実験では，基準点を復元対象上の 10 ピクセル間隔の格子状に配置した．このとき，SW-POC に基づく手法では，A.2.3 節で述べたように疎な基準点と密な基準点でそれぞれ対応付けを行う．疎な基準点は 30 ピクセル間隔の格子状に，密な基準点は 10 ピクセル間隔の格子状に配置した．

各手法における対応付けのパラメータを以下のように設定する．ウィンドウサイズは，Bradley らの手法において 16 ピクセル × 15 ライン，1 次元 POC および SW-POC に基づく手法において 32 ピクセル × 15 ラインとする．このとき，1 次元 POC および SW-POC に基づく手法では，3.3.1 節で述べたように，1 次元 POC 関数の計算の際にハニング窓をかけるため [27]，32 × 15 ピクセルのウィンドウの情報量は，Bradley らの手法で用いる 16 × 15 ピクセルのウィンドウの情報量に相当する．1 次元 POC および SW-POC に基づく手法において，画像ピラミッドの階層数は 3 とする．また，SW-POC について，画像ピラミッドの最上位層における初期対応点  $q_{h_{\max}}$  の探索範囲は水平座標方向に  $\pm 40$  ピクセルとする．Bradley らの手法では，対応点の探索範囲を提案手法と同じにするため，画像ピ

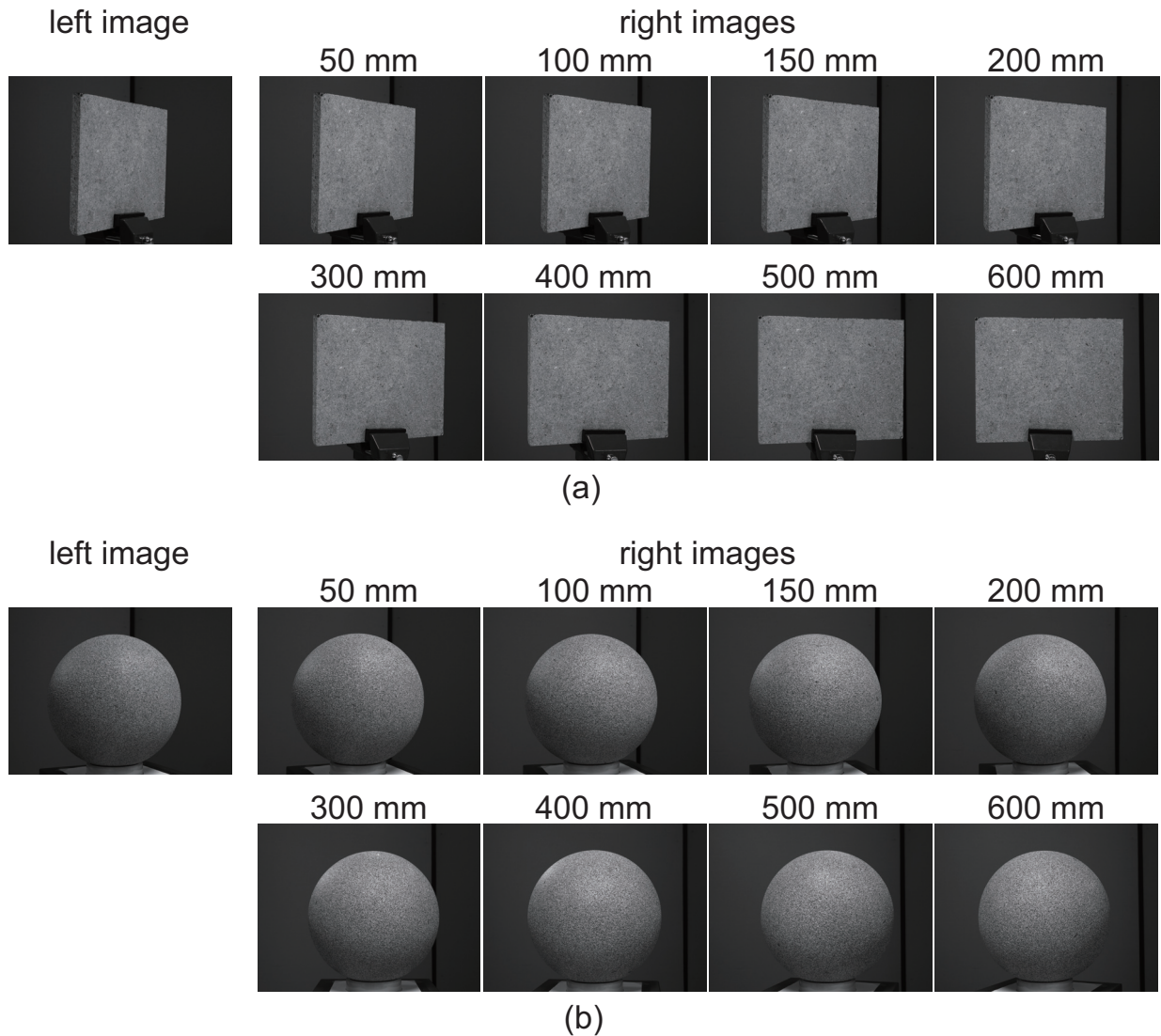


図 A.5 撮影されたステレオ画像: (a) 平板, (b) 球

ラミッドの最上層における  $\pm 40$  ピクセルに相当する  $\pm 160$  ピクセルについて探索を行う。また, Bradley らの手法において, サブピクセルレベルの対応付けは  $1/10$  ピクセルの分解能で行う。SW-POC に基づく手法および Bradley らの手法について, 拡大縮小率は  $1/2$ ,  $1/\sqrt{2}$ ,  $1$ ,  $\sqrt{2}$ ,  $2$  倍の 5 つとする。ここで, Bradley らの手法について, 拡大縮小率は, 本来  $1/\sqrt{2}$ ,  $1$ ,  $\sqrt{2}$  としているが [19], 本実験においては, それ以上の画像変形が生じることがあり, 3 つの拡大縮小率では不十分であったので, 5 つの拡大縮小率についてマッチングを行った。

### A.3.2 3次元復元精度の評価

まず、基線長の変化に対するロバスト性について評価する。それぞれの基線長において、提案手法および従来手法でステレオ画像間を対応付け、3次元復元を行う。この3次元復元結果に対して、平板または球の方程式を当てはめ、各3次元点における当てはめ誤差を算出する。このとき、当てはめ誤差がステレオ画像上で1ピクセル以上である点を誤対応点とし、3次元復元点群のうち誤対応点が占める割合を誤対応率とする。平板および球の復元に対する誤対応率を表 A.1 に示す。また、基線長 50 mm と 400 mm のときの平板および球の3次元復元点群を図 A.6 に示す。表 A.1 および図 A.6 より、ウィンドウの拡大縮小を行わない1次元 POC に基づく手法では、基線長の拡大に伴い、誤対応率が増加していることが確認できる。これは、基線長の拡大により、ステレオ画像間の変形が大きくなり、ウィンドウ間の画像変形を平行移動のみに近似できなくなったためである。これに対して、SW-POC に基づく手法および Bradley らの手法は、基線長が長いステレオ画像においても、誤対応が少ないことが確認できる。特に、SW-POC に基づく手法は、基線長が 300 mm 以下の場合にすべての基準点について対応付けが行えている。ただし、基線長 400 mm 以上の球については、ステレオ画像間の変形が非常に大きくなり、提案手法を用いた場合においても誤対応率が増加している。また、SW-POC に基づく手法と Bradley らの手法を比較すると、SW-POC に基づく手法では、Bradley らの手法に比べて基線長が長いときの誤対応率が小さいことがわかる。特に、球の結果では、それが顕著に表れている。球のような復元対象では、基線長が長くなるとステレオ画像間に様々な種類の変形が生じる。Bradley らの手法では、ウィンドウ間にスキューが生じた場合、画像変形を拡大縮小に近似できなくなり、誤対応が発生する。これに対して、SW-POC に基づく手法では、A.2.4 節で述べたようにスキューのような変形を軽減することで、ロバストな対応付けを実現している。

次に、サブピクセルレベルの対応付け精度について評価する。上記の実験のように、復元された3次元点群に対して、平板または球の方程式を当てはめ、誤差が1ピクセル以上の点を誤対応点として除去する。この誤対応点が除去された3次元復元点群に対して、方程式の当てはめ誤差の Root Mean Square (RMS) で評価する。表 A.2 に平板および球の3次元復元誤差を示す。表 A.2 より、SW-POC に基づく手法および Bradley らの手法では、基線長が長くなるにつれて3次元復元誤差が減少していることが確認できる。これは、基線長の拡大により、対応付け誤差が3次元復元結果に与える影響が小さくなったためであると考えられる。ただし、球の結果については、基線長が長くなり過ぎると、対応付け誤差が大きくなり、その結果、3次元復元誤差も増加している。3次元復元誤差につ

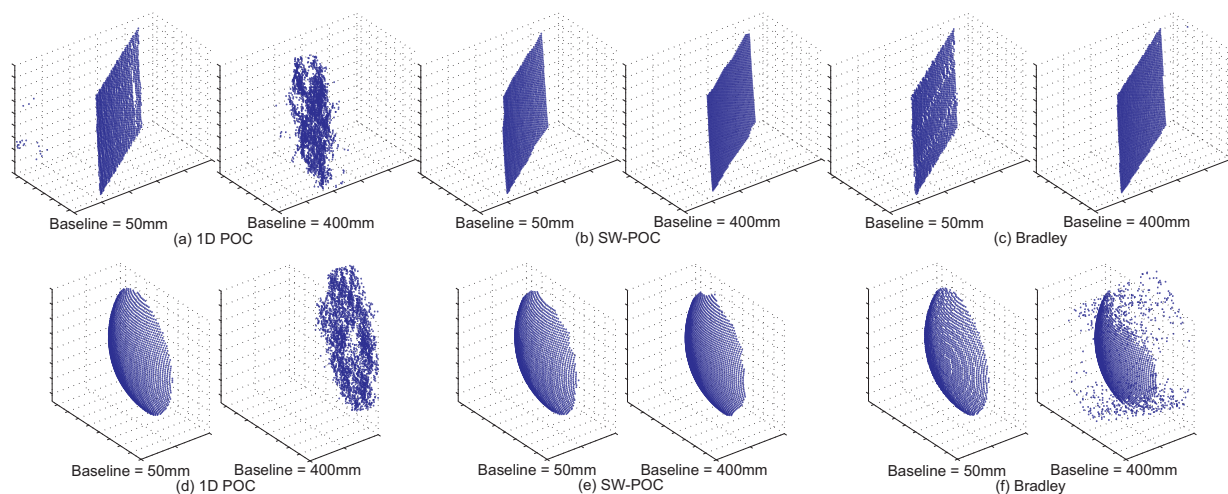


図 A.6 3次元復元結果: (a)–(c) 平板の復元結果, (d)–(f) 球の復元結果

表 A.1 平板および球の誤対応率 [%]

Baseline [mm]		50	100	150	200	300	400	500	600
Plane	1D POC	1.69	11.94	18.28	34.04	93.67	92.75	92.46	91.51
	SW-POC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Bradley	0.00	0.00	0.00	0.00	0.00	0.00	8.47	13.67
Sphere	1D POC	0.00	0.98	43.03	98.89	98.76	99.05	99.32	98.93
	SW-POC	0.00	0.00	0.00	0.00	0.00	0.03	13.32	36.21
	Bradley	0.00	0.00	0.00	0.18	6.04	18.63	35.06	57.76

いて, SW-POC に基づく手法と Bradley らの手法を比較すると, どちらの復元対象についても, SW-POC に基づく手法の方が 3次元復元誤差が小さいことが確認できる. また, このことは, 図 A.6 から確認できる. 特に, 平板において基線長が短いとき, および, 球において基線長が長いときは, 3次元復元精度の差が顕著に表れている. この要因として, SW-POC に基づく手法は, Bradley らの手法と比べてサブピクセルレベルの平行移動量推定精度が高いこと, および, スキューに対してもウィンドウ間の画像変形を軽減していることが考えられる.

表 A.2 平板および球の 3 次元復元誤差の RMS [mm]

Baseline [mm]		50	100	150	200	300	400	500	600
Plane	1D POC	0.3224	0.3298	0.3857	0.4730	0.5276	0.4153	0.2418	0.2753
	SW-POC	0.2010	0.1601	0.1343	0.1320	0.1130	0.1117	0.1096	0.1085
	Bradley	0.4513	0.3864	0.1889	0.1889	0.1271	0.1385	0.1519	0.1555
Sphere	1D POC	0.2240	0.1319	0.1083	0.6104	0.4636	0.3893	0.3330	0.2738
	SW-POC	0.2203	0.1119	0.0727	0.0688	0.0564	0.0588	0.0713	0.0716
	Bradley	0.3479	0.1998	0.1526	0.1455	0.1404	0.1211	0.1111	0.1139

表 A.3 1 点の対応付けに必要な演算回数

	Additions	Multiplications	Divisions	Square roots
1D POC	70,784	61,056	3,840	1,920
SW-POC	200,555	172,992	5,120	5,440
Bradley	3,109,300	1,302,200	25,500	25,500

### A.3.3 計算コストの評価

ここでは、1 点の対応付けに必要な計算コストの比較を行う。つまり、左カメラ画像上の基準点  $p$  に対応する右カメラ画像上の対応点  $q$  を 1 点だけ、サブピクセル精度で求める場合の計算コストを考える。ここで、SW-POC に基づく手法では、30 ピクセルごとに配置された基準点と 10 ピクセルごとに配置された基準点で拡大縮小率の推定手法の違いにより、計算コストも異なるため、平均の計算コストを考える。

表 A.3 にそれぞれの手法で必要となる加算、乗算、除算、平方根の演算回数を示す。表 A.3 から、Bradley らの手法では、探索範囲を  $\pm 160$  ピクセルに限定した場合においても、計算コストが大きいことがわかる。実際には、Bradley らの手法において、探索範囲は視体積交差法により決定される。よって、本実験に使用されたステレオ画像のようにシルエットが定義されていない場合、画像全体の探索を行うため、さらに計算コストが大きくなると考えられる。これに対して、SW-POC に基づく手法では、画像ピラミッドを用いた階層的探索を適用することで、マッチングの回数を削減しているため、Bradley らの手法に比べ計算コストが小さくなっている。

## A.4 むすび

本章では，SW-POC に基づくステレオ画像間の変形にロバストな対応付け手法について述べた．はじめに，SW-POC の原理について説明し，SW-POC をステレオ画像間の局所的なウィンドウマッチングに適用することで，SW-POC に基づく対応付け手法を提案した．そして，性能評価実験を通して，SW-POC に基づく手法は，基線長の長いステレオ画像においても，従来手法に比べ少ない計算コストで高精度な対応付けが可能であることを示した．SW-POC はステレオ画像間の基線長によらず，高精度なマッチングが可能であり，多視点画像からの 3 次元復元に応用することで，より自由な撮影による 3 次元復元の実現が期待できる．



## 参考文献

- [1] 吉澤徹, 最新光三次元計測. 朝倉書店, 2006.
- [2] 池内克史, 大石岳史, 3次元デジタルアーカイブ. 東京大学出版会, 2010.
- [3] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer-Verlag New York Inc., 2010.
- [4] 徐剛, 辻三郎, 3次元ビジョン. 共立出版, 1998.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.

- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] N. D. F. Campbell, G. Vogiatzis, C. Hernandez, and R. Cipolla, “Using multiple hypotheses to improve depth-maps for multi-view stereo,” *Proc. European Conf. Computer Vision*, pp. 766–779, Oct. 2008.

## 参考文献

- [21] 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.
- [22] E. Tola, C. Strecha, and P. Fua, “Efficient large-scale multi-view stereo for ultra high-resolution image sets,” *Machine Vision and Applications*, 2011 (Published Online).
- [23] 奥富正敏 (編), デジタル画像処理. CG-ARTS 協会, 2004.
- [24] 酒井修二, 伊藤康一, 青木孝文, 運天弘樹, “位相限定相関法に基づく多視点画像からのデプスマップ生成の検討,” 第 14 回画像の認識・理解シンポジウム, Vol. 14, No. IS4-10, pp. 1347–1354, July 2011.
- [25] 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.
- [26] 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.
- [27] 柴原琢磨, 沼徳仁, 長嶋聖, 青木孝文, 中島寛, 小林孝次, “1 次元位相限定相関法に基づくステレオ画像の高精度サブピクセル対応付け手法,” 電子情報通信学会論文誌 D, Vol. J91-D, No. 9, pp. 2343–2356, Sept. 2008.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.

- [32] 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.
- [33] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int’l J. Comput. Vision*, Vol. 60, No. 2, pp. 91–110, Nov. 2004.
- [34] H. Bay, T. Tuytelaars, and L. V. Gool, “SURF: Speeded up robust features,” *Proc. European Conf. Computer Vision*, pp. 404–417, May 2006.
- [35] 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.
- [36] 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.
- [37] R. Hartley and A. Zisserman, *Multiple View Geometry*. Cambridge University Press, 2004.
- [38] 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.
- [39] 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.
- [40] B. Curless and M. Levoy, “A volumetric method for building complex models from range images,” *Proc. SIGGRAPH ’96*, pp. 303–312, Aug. 1996.
- [41] M. Kazhdan, M. Bolitho, and H. Hoppe, “Poisson surface reconstruction,” *Proc. Symp. Geometry Processing*, pp. 61–70, June 2006.
- [42] Y. Liu, X. Cao, Q. Dai, and W. Xu, “Continuous depth estimation for multi-view stereo,” *Proc. Int’l Conf. Computer Vision and Pattern Recognition*, pp. 2121–2128, June 2009.

## 参考文献

- [43] E.Tola, V. Lepetit, and P. Fua, “DAISY: An efficient dense descriptor applied to wide-baseline stereo,” *Proc. IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 32, No. 5, pp. 815–830, May 2010.
- [44] 酒井修二, 高橋徹, 伊藤康一, 青木孝文, “位相限定相関法を用いた 3 次元計測の高精度化と性能評価,” *映像情報メディア学会技術報告*, Vol. 34, No. 34, pp. 43–46, Aug. 2010.
- [45] S. Sakai, K. Ito, T. Aoki, and H. Unten, “Accurate and dense wide-baseline stereo matching using SW-POC,” *Proc. Asian Conf. Pattern Recognition*, pp. 335–339, Nov. 2011.
- [46] M. Okutomi and T. Kanade, “A multiple-baseline stereo,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, pp. 353–363, Apr. 1993.
- [47] C. Strecha, “Multi-view evaluation.” <http://cvlab.epfl.ch/data/>.

## 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（孝）・本間（尚）研究室）において行った研究を取りまとめたものです。

本研究を推し進めるにあたり、恩師青木孝文教授には、筆者が学部4年次の分野配属以来、終始熱心な御指導と御鞭撻を頂きました。先生の研究・教育に対する真摯な御姿勢から多くを学んだことを銘記し、ここに改めて感謝の意を表します。

本論文をまとめるにあたり、本論文の審査員として御専門の立場から有意義な御意見・御批判を頂いた塩入諭教授ならびに大町真一郎教授に感謝いたします。

凸版印刷株式会社の運天弘樹氏には、共同研究を通して、3次元復元技術および文化財のデジタルアーカイブに関する有意義なご討論とご助言を頂きました。ここに、深く感謝いたします。

伊藤康一助教には、研究に対する懇切なる御指導と終始変わらぬ励ましを頂くとともに、本論文の執筆においても様々な御助言を賜りました。ここに改めて感謝の意を表します。

宮澤一之博士（現 三菱電機株式会社）には、位相限定相関法の基礎原理および、位相限定相関法に基づくステレオ画像間の対応付け手法について、有益なご議論とご指導を頂きました。ここに、改めて感謝の意を表します。また、高橋徹博士（現 日本電気株式会社）には、研究全般を通して多大なるご助力を頂きました。深く感謝いたします。

青木（孝）・本間（尚）研究室の石井純平氏には、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。ここに改めて感謝の意を表します。

最後に、日頃の研究室生活において様々な面で御協力頂いた本間尚文准教授をはじめとする研究室諸氏に心より御礼申し上げます。

2012年2月10日