



HOKKAIDO UNIVERSITY

Title	実環境システムノイズにロバストな三次元計測および認識に関する研究
Author(s)	住吉, 信一
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第14727号
Issue Date	2021-09-24
DOI	https://doi.org/10.14943/doctoral.k14727
Doc URL	https://hdl.handle.net/2115/83276
Type	doctoral thesis
File Information	Sumiyoshi_Shinichi.pdf



-博士論文-

実環境システムノイズにロバストな
三次元計測および認識に関する研究

Study on 3D Measurement and Recognition Robust
to Real Environmental System Noise

北海道大学大学院情報科学研究科
メディアネットワーク専攻

住吉 信一

令和3年8月

学位論文内容の要旨

計算機博士の専攻分野の名称 博士（情報科学） 氏名 住吉 信一

学位論文題名

実環境システムノイズにロバストな三次元計測および認識に関する研究
(Study on 3D Measurement and Recognition Robust
to Real Environmental System Noise)

近年、工業分野においてファクトリーオートメーションが進みつつある。例えば、ロボットアームとコンベアを用いたシステムを用い、コンベア上に載せたワークをロボットアームで自動でピッキングし、組み立てることで、目的の製品を作り上げるための自動生産システムがある。このようなシステムを作り上げるためには、さまざまなロボティクス、情報処理等の技術が必要となるが、中でも、三次元計測、および、三次元認識技術は欠かせない要素技術である。

三次元計測技術は、様々に存在し、用途に応じ選択し利用されるのが一般的であり、工業的にはプロジェクタカメラシステムが用いられることが多い。ただし、投影光が相互反射したり、透過したりするなどの多様な光学現象に対して高いロバスト性を持つ安価な計測器や安価に実現できる研究・開発技術は未だ登場していない。

三次元認識技術についても実用性が高い技術がいくつか発表されており、計測データにノイズがあっても頑健にマッチングする手法や、多少のデータ欠損があってもロバストにマッチングして対象の位置姿勢を推定できる技術が登場している。ただし、これらも工業的な応用にそのまま活用できるわけではなく、精度、速度面で実用に耐えるようにさらなる工夫をしなければならない。

このように、それぞれの技術分野において、実環境に導入して活用できる技術にはまだ成熟していない。このため、各技術分野において特に課題となる、A. 対象の物性によって三次元計測が困難となる課題、および、B. 対象の認識が困難となる環境での三次元認識課題、について取り組んだ。

本論文は以下の7の章から構成される。

1章にて、ファクトリーオートメーションに関する三次元計測および認識に関する研究背景・目的について述べる。

2章では、プロジェクタとカメラを用いた三次元計測技術で、金属物のような表面にて相互反射が起こる物体を計測する技術を示す。このような場合、物体表面で2回以上反射した間接反射光と、物体表面で一度だけ反射した直接反射光とを分離する必要がある。投影光にこれらの二つの光を分離できるようにするために、時空間的な符号を与え、プロジェクタとカメラの配置関係で決定されるエピポーラ制約を

用いることで、間接反射成分の多くを分離取得できる手法がある。しかし金属物で起こる鏡面相互反射現象で観測される成分は、既存法では分離できない点が発生する課題がある。そこで本研究では、形状を復元する奥行きを限定する幾何制約を既存法に加えることで、直接反射成分と間接反射成分の二つを分離可能な手法を提案する。さらに、この幾何制約でパターン枚数を低減させるアルゴリズムと一般のプロジェクタより 200 倍程度、高速にパターンを投影可能な DLP プロジェクタを取り入れることで、鏡面相互反射が起きる物体を既存法よりも高速に計測可能な手法を提案する。提案法は、鏡面相互反射が起きるシーンにおいて、122 fps で計測可能であることを確認した。

3章では、プロジェクタとカメラを用いたより光学的なアーティファクトに対して頑健な計測手法について示す。プロジェクタ投影光の変調方式に工夫を施すことで相互反射の影響を抑制し、金属物体の 3 次元計測の精度を向上させる手法が提案されているが、単純なグレイコード法に比べると投影枚数が非常に多く、計測に長時間を要するという課題があった。そこで本研究では、ランダムパターンの相関演算に基づく対応付けに幾何学的な制約を導入することで、高速化と高精度化を両立させた 3 次元計測法を提案する。また実運用における対象シーンと投影数による計測精度見積もりのために、投影パターンの観測モデルを作成し数値解析により対応付けの精度を見積もる。実験では投影枚数が 30 枚と少ない場合でも相互反射光の影響を抑制し、良好に形状を復元可能であることを確認した。^{*1}

4章では、生産現場などでコントロールされた動きをする対象物が、移動中でも 3 次元計測できる手法について示す。本研究では、(a) 再構成対象が金属部品のような非拡散物体であり、(b) 物体は既知の剛体運動の結果としてスライディングコンベアによって運ばれるという、特定のファクトリーオートメーションシナリオにおける 3 次元形状再構成の問題に取り組む。この問題を解決するために、上記のシナリオを考慮したカメラとプロジェクタ間の密なステレオ対応関係を得るための光学現象にロバストな手法を提案する。本研究ではこの条件下では自然に既知とできる対象物の動きが強い幾何制約条件として働くことに注目した。本研究の貢献は 2 つある。(a) 固定 2 次元ランダムパターンを用いて直接反射と間接反射に分離できること、(b) カメラとプロジェクタ間のエピポラ制約と物体の動きの事前知識の両方を用いて不正確なステレオ対応を除去すること、である。カメラの frame rate と物体の速度とが 6mm/frame の関係のときの実験において、同程度の精度要求時にスリット光投影法に比べ提案法は 54% 程度の計測密度向上ができることを確認した。

5章では、大きく遮蔽された物体の 3 次元認識技術について示す。三次元点群情報から特徴量をスパースに得ることで、セミリアルタイムな 3D 物体検出が可能な手法が提案されているが、強遮蔽された物体は検出が難しい。本研究では、検出対象物の周辺に、仮想的な補助点群を与えることで、強遮蔽された物体の検出が可能と

なる手法を提案する．この補助点群を，手前側の検出物をオクルーダとした時に推定される遮蔽空間の内部に限定して生成することで，処理効率と精度を向上させる．実環境データを用いた様々な強遮蔽シーンに対する実験を行い，提案法は平均約 0.5 sec で，強遮蔽された物体の検出が可能であることを確認した．*²

6章では，金属物の 3 次元的存在領域の推定技術について示す．廃棄物の自動分別処理のため，低コストセンサで高速に金属物を判別する必要がある．本研究ではプロジェクタカメラを用いて，対象シーンから金属物がある領域を推定する手法を提案する．提案法は金属物で発生する相互反射光と直接反射光とを分離取得できる手法を用い，相互反射が起きる領域を金属物の存在領域として検出する．ランバート反射物と金属物のあるシンプルなシーンで実験を行い，金属物を検出可能であることを確認した．*³

最後に前述した提案法に関するまとめを 7 章にて述べる．これらの結果から，産業分野において肝要と考えられるファクトリーオートメーションに関する課題で，A. 対象物性によって起こる光学現象の影響を抑制した計測が可能な三次元計測技術，および，B. 対象物の周辺物から受ける影響を抑制した三次元認識技術，に貢献する技術を示した．

目次

第 1 章	序論	1
1.1	はじめに	1
1.2	三次元計測とは	3
1.3	プロジェクタカメラ方式による三次元計測	6
1.4	三次元認識とは	8
1.5	三次元計測および認識技術の産業応用における課題	9
1.6	本論文の構成	10
第 2 章	提案手法 1 : DLP プロジェクタを用いた鏡面相互反射光の影響を抑制可能な高 速アクティブ三次元計測手法	13
2.1	研究背景・目的	13
2.2	関連研究	14
2.3	鏡面相互反射を抑制可能な高速三次元計測手法	16
2.3.1	投影パターンの設計	17
2.3.2	幾何制約による鏡面相互反射点の除去	19
2.4	実験と評価	21
2.4.1	実験条件	21
2.4.2	復号化情報の分離と三次元計測性能の評価	24
2.4.3	計測速度の評価	24
2.5	まとめ	25
第 3 章	提案手法 2 : ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速ア クティブ三次元計測法 ^{*1}	26
3.1	はじめに ^{*11}	26
3.2	関連研究 ^{*12}	28
3.2.1	空間周波数特性を利用した手法	28

3.2.2	エピソード面外からの伝播を抑制する手法	28
3.2.3	エピソード面内の符号の重畳を分離する手法	29
3.2.4	2次元符号化	30
3.2.5	偏光等を利用する手法	31
3.3	提案手法* ¹³	31
3.3.1	2次元ランダムパターンの投影による対応点の取得	31
3.3.2	幾何学的制約の導入	36
3.3.3	投影パターン数と相関値の分布の関係	37
	非対応点, 対応点の相関値の分布モデル	38
	数値解析による対応付け精度の見積もり	40
3.4	実験と評価* ¹⁴	41
3.4.1	奥行き制約の効果に関する実験	44
3.4.2	従来手法との比較	44
3.4.3	3次元計測精度の定量評価	44
3.4.4	計測時間	47
3.5	まとめ* ¹⁵	47
第 4 章	提案手法 3 : 動き制約と固定 2 次元ランダムパターン投影を用いた移動物体の密 なアクティブ計測法	49
4.1	はじめに	49
4.2	関連研究	51
4.3	提案法	53
4.3.1	システム概要	54
4.3.2	動き制約	57
4.3.3	動き制約と 2 次元ランダムパターンを用いた時空間符号	57
4.3.4	光学現象の影響の抑制	58
4.3.5	実装法	58
4.4	実験	59
4.4.1	計測密度と対象物のカメラフレームレートに対する移動速度	62
4.4.2	計測精度	62
4.4.3	光学現象の影響の抑制	66
4.5	まとめ	68
第 5 章	提案手法 4 : 遮蔽空間への補助点群を用いた強隠蔽物の 3 次元位置のセミリアル	

	タイム推定* ²	70
5.1	背景* ²¹	70
5.2	関連研究* ²²	72
5.2.1	テンプレートベースの手法	72
5.2.2	デンス情報ベースの手法	73
5.2.3	スパース情報ベースの手法	73
5.3	提案法* ²³	74
5.3.1	遮蔽空間の推定	76
5.3.2	遮蔽空間内での補助点群の生成	78
5.3.3	物体検出	82
5.4	実験と評価* ²⁴	82
5.4.1	実験条件	83
	実装	83
	精度の評価	84
	実験方法	84
5.4.2	評価	85
	補助点群生成手法の検証実験	85
	提案法の実験	85
5.5	まとめと今後の課題* ²⁵	92
第 6 章	提案手法 5 :	
	相互反射領域推定法を用いた高速な金属物の検出* ³	93
6.1	はじめに* ³¹	93
6.2	関連研究* ³²	94
6.3	金属物の形状復元法を用いた, 相互反射領域推定による高速な金属物の検出* ³³	95
6.3.1	少数投影で鏡面相互反射を抑制可能な 3 次元計測手法	96
6.3.2	相互反射領域推定による高速な金属物の検出	98
6.4	実験と評価* ³⁴	99
6.5	まとめ* ³⁵	100
第 7 章	結論	102
	謝辞	104
	参考文献	105

図目次

1.1	ファクトリーオートメーションのロボットピッキングシーンでの課題イメージ：三次元計測が困難なシーン．	2
1.2	ファクトリーオートメーションのロボットピッキングシーンでの課題イメージ：三次元認識が困難なシーン．	2
1.3	株式会社リコー製の産業用ステレオカメラ．	4
1.4	株式会社リコー製の産業用ステレオカメラで得られたデプスマップ．	4
1.5	対象物の材質によって起こる多様な光学現象例．	5
1.6	透過，半透過，相互反射の光学現象イメージ．	5
1.7	プロジェクタ・カメラ画像間の画素の対応付けのための投影・撮影方法のイメージ．	7
1.8	三角測量による対応点の三次元計測．	7
1.9	点群による物体検出イメージ．	9
1.10	実時間で高速に三次元計測をすることが難しい物体例．	10
1.11	大きく遮蔽された対象物の検出シーン．	11
1.12	定まった形状のない金属片の検出シーン．	11
2.1	直接反射成分と間接反射成分とが重畳したケース．	14
2.2	提案法の概要図．	17
2.3	DLP プロジェクタと高速度カメラによる符号化予備実験．	18
2.4	色を 9 種用いたプロジェクタからの投影パターン画像と，各色ごとに出力される座標符号化信号同士の NCC 値の混同行列．	19
2.5	各色の座標符号化信号のキャリブレーション結果．	20
2.6	本実験で用いるプロジェクタ，カメラ，計測対象物．	20
2.7	図 2.4 左のパターン投影時に発生した鏡面相互反射の様子．	21
2.8	復号化に用いる高速度カメラ画像列．	21
2.9	投影光の復号化結果．	22
2.10	奥行き制約の説明図．	23
2.11	提案法の形状復元のフローイメージ図．	23

2.12	提案法と既存法による三次元計測結果.	24
3.1	直接反射成分と鏡面相互反射成分とが重畳する 2 つの例.	27
3.2	鏡面相互反射成分がエピポーラ制約を満たさない場合と満たす場合.	29
3.3	提案法のシステムイメージ.	32
3.4	平面に対する復号化実験の系.	33
3.5	平面における投影枚数 N を変化させたときの復号化結果の変化.	34
3.6	平面における投影枚数 N を変化させたときの ZNCC の変化.	34
3.7	実験系と相互反射物の撮影画像.	35
3.8	相互反射物における投影枚数 N を変化させたときの復号化結果の 変化.	35
3.9	相互反射物における投影枚数 N を変化させたときの相関値の変化.	36
3.10	左：実験環境，右：エピポーラ制約と奥行き制約.	37
3.11	エピポーラ制約の効果を調べる実験系.	38
3.12	エピポーラ制約の相互反射光対応点除去化.	39
3.13	左：非対応点と対応点の相関値の確率分布から求めた ROC 曲線， 右：相関を演算する領域の画素数 M_c において，その領域内に対応 点が存在し，対応点の相関値が閾値 C_{th} を上回り，かつ， M_c の中 で最大となる確率 q	42
3.14	実シーンに対する実験における相関値の分布.	42
3.15	図 3.14 のシーンの相関値のヒストグラム.	43
3.16	GC，および， $N = 50$ の場合の提案法による対応付けの結果.	45
3.17	$N = 30$ の時の提案法による復号化の結果.	45
3.18	実験シーン（図 3.10 左，図 3.16(a)）の形状復元結果.	46
3.19	左：復元の成功率，右：計測点の奥行き方向（Z 軸）の平均絶対誤差.	47
4.1	提案計測システムの処理イメージ.	50
4.2	アルゴリズム概要.	55
4.3	実験システム.	56
4.4	カメラでの撮影画像.	56
4.5	動き制約で 1 つのカメラ画素で発生させた全仮説点を移動させなが らカメラ画像，プロジェクタ画像へ投影した図.	58
4.6	実験に用いた投影パターン.	59
4.7	Object Motion パラメタのキャリブレーションの系.	60
4.8	Object Motion パラメタの推定用のキャリブレーション画像.	61
4.9	検出したチェスボードのコーナー点の 3D プロットの結果.	61

4.10	各パターンで $N = 54(\text{move } 1\text{mm}/\text{frame})$ で得た復元結果とカメラフレームレートに対して高速な移動体を撮影したと仮定して撮影数 N を間引いた際の復元結果との点群量の比.	63
4.11	各投影パターンで得た左 : $N = 9(\text{move } 6\text{mm}/\text{frame})$, 右 : $N = 54(\text{move } 1\text{mm}/\text{frame})$ の復元結果.	64
4.12	平面の評価に用いた Resin Cube の部屋照明下と白投影時のカメラ画像.	64
4.13	平面の評価 : 最小二乗フィッティングの結果.	65
4.14	球面の評価に用いた Wood Ball の部屋照明下と白投影時のカメラ画像.	65
4.15	球面の評価 : 最小二乗フィッティングの結果.	66
4.16	各パターンでの計測結果を最小二乗法でフィッティングした平面あるいは球と計測データとの MAE と RMSE.	67
4.17	左 : ランダムドットパターン 100 枚で得た疑似真値と各パターンでの計測結果で共通する点の割合. 右 : 各パターンでの計測結果を疑似真値との MAE.	68
4.18	様々なシーンでの各投影パターンの形状復元結果.	69
5.1	機械知覚 & ロボティクスグループ中部大学のアマゾンピッキングチャレンジの様子.	71
5.2	作業現場にて, 強遮蔽された危険物があつた場合の, 危険物の検出イメージ.	72
5.3	危険物 (検出対象物) を検出した際の, HMD での表示イメージ.	72
5.4	遮蔽空間推定のフロー図.	77
5.5	本稿で用いる座標系.	77
5.6	入力シーンの例.	78
5.7	YOLO アルゴリズムを用いた, 手前側の物体 (オクルーダ候補) の検出.	79
5.8	ラップトップ PC の拡張エッジ領域の計算法.	79
5.9	3D SIFT の結果.	80
5.10	机上のラップトップ PC から推定された遮蔽空間.	80
5.11	遮蔽空間を考慮した補助点群の生成.	82
5.12	ラップトップ PC により推定された遮蔽空間内の補助点群.	83
5.13	RGBD センサ : Intel RealSense R200.	86
5.14	検出対象物と理想的な遮蔽空間データ.	86
5.15	遮蔽率を変化させた場合の, マッチング成功率.	87

5.16	実シーンの点群を用いて得られた, マッチング結果.	88
5.17	検出対象物のモデルの点群データベース.	88
5.18	FPFH のみでマッチングする既存法と, 提案法のマッチングプロセスに FPFH を用いる手法による, 様々なシーンでの検出対象物の位置推定結果.	90
5.19	PPF のみでマッチングする既存法と, 提案法のマッチングプロセスに PPF を用いる手法による, 検出対象物の位置推定結果. . . .	91
6.1	直接反射成分と間接反射成分とが重畳したケース.	94
6.2	左: 投影パターン画像と, 右: 実験環境.	97
6.3	ZNCC で得た最大相関 (左) と第 2 位の相関 (右) をマップした図. . . .	97
6.4	奥行き制約の説明図.	98
6.5	$N=30$ の時の提案法による対応付けの結果.	100
6.6	$N=30$ の時の実験シーン (図 6.2 右) の形状復元結果. 左: 既存法を用いた結果, 右: 3 章の手法を用いた結果.	100
6.7	$N=30$, 実験シーン (図 6.2 右) の, 左: 相互反射推定マップ, 右: 近傍ドット光混入マップ.	101
6.8	$N=30$, 実験シーン (図 6.2 右) の, 左: 金属物領域の一部と推測されたカメラ画像上での領域, 右: 全体の復元点 (白) に, 金属物領域の一部と推測された復元点 (緑) で重畳した結果.	101

表目次

5.1	図 5.18の各実験の結果.	90
5.2	図 5.19の各実験の結果.	91

第 1 章

序論

1.1 はじめに

近年，コンピュータビジョン技術の工業的な応用が活発になされており，特に自動生産分野における発展が目覚ましい．特に実応用を考える上で社会的に需要が高まりつつあるファクトリーオートメーションを実現するという観点から，本論文では三次元計測技術，および，三次元認識技術に関して重要と思われる事項について検討する．

ファクトリーオートメーションの代表的なシステムであるロボットピッキングシステムについて考える．ロボットピッキングは様々な利活用シーンが考えられるが，図 1.1や図 1.2のように，大量の物資をピッキングして梱包するようなロボットを扱いたい場合，対象物が液体であったり，透明なガラスコップ，あるいは，金属でできた物体であるなど，多様な材質物が存在する．このような場合，三次元計測が一般に困難な反射特性をもつ材質であることが多いと考えられる．またこのような場合，計測にノイズが多量にのってしまうが，これが原因で後段処理である認識にも誤差がのってしまう．さらに認識処理では箱詰めタスクのような奥側の物体検知や材質同定をする必要があり，これらの課題も周りの物体の影響を受け認識が困難となる．このため既存技術を用いても，現実にはロボットピッキングタスクが失敗してしまうケースが多く存在すると考えられる．このような課題を解消するため，本研究では三次元計測，および，三次元認識の分野において，それぞれの技術分野において特に重要な課題と考えられる，A. 対象の物性によって三次元計測が困難となる課題，および，B. 対象の認識が困難となる環境での三次元認識課題，について取り組む．

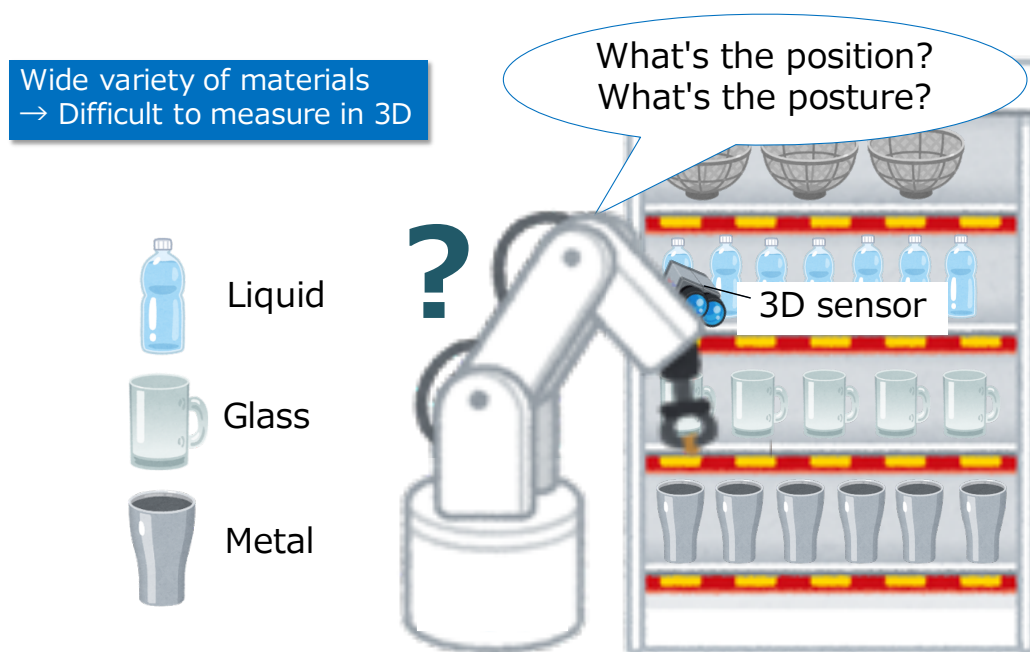


図1.1 ファクトリーオートメーションのロボットピッキングシーンでの課題イメージ：三次元計測が困難なシーン．

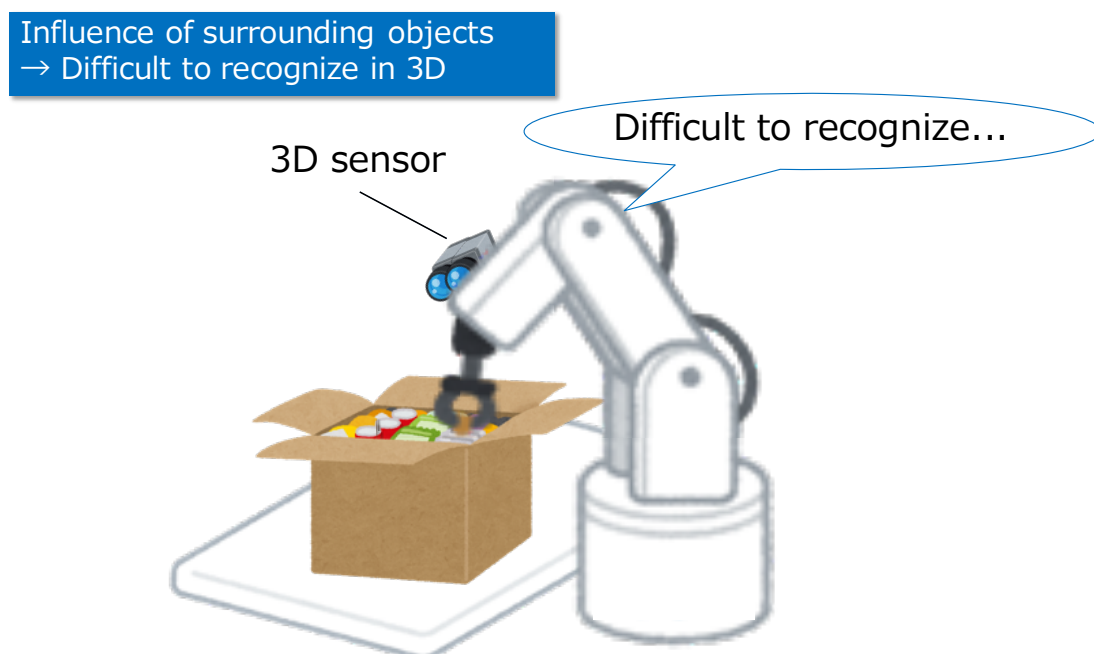


図1.2 ファクトリーオートメーションのロボットピッキングシーンでの課題イメージ：三次元認識が困難なシーン．

1.2 三次元計測とは

三次元計測技術は古くから研究が続けられており、以下のような様々な方式が研究開発されている。

- カンチレバー方式（接触型）
- マイケルソン光学方式（非接触型）
- ラインレーザー方式（非接触型）
- ステレオカメラ方式（非接触型）
- プロジェクタカメラ方式（非接触型）
- ToF 方式（非接触型）

一点物や抜き取り検査などのオフライン計測に有用な計測方式として、カンチレバー方式、マイケルソン光学方式などが知られている。カンチレバー方式は、様々な大きさ・物性の物体を精度よく測れる方式として古くから用いられている。ただし、この方式は計測に時間がかかり、また対象物に接触する方式であるために、一点物の計測には向いているが、工業的なラインでオンラインで用いるという活用法には不向きである。マイケルソン光学方式は、レンズ計測の主流な計測方法であり、台座の上においたレンズの透過して反射した光と直接反射した光の光干渉から形状を得ることができる。レンズのような透明な物体に適用できるものの、カンチレバー方式と同様に抜き取り検査などの一点物の計測には向いているがオンラインで利用すべき手法ではない。

よりリアルタイムにオンラインでの計測を目指した方式は、ラインレーザー方式、ステレオカメラ方式、プロジェクタカメラ方式、ToF 方式などが代表的である。ラインレーザー方式は、コンベア上を動く物体をリアルタイムに計測することに適した計測法である。ただし、静止した物体を計測することが難しく、ロボットアームと組み合わせることを考えた場合は対象が箱の中にばら積みされたパーツであったり、ワークが置かれたシーンであったりするために、汎用性には欠ける方式といえる。ステレオカメラ方式は、各社メーカーから衝突防止センサ等で販売もされているが、ファクトリーオートメーションのシナリオでも活用されつつある。例えば、株式会社リコー製の工業用ステレオカメラなどがある [1]。2つのカメラセンサにより三角測量の原理で奥行き値を得ることができる（図 1.3, 図 1.4）。ただしこの方式は対象にテクスチャが乏しい場合、レーザー照明などを用いて対象物にテクスチャ情報を与え、三次元計測をするシステムにする必要があり、装置が大きくなったりコストが高くなる傾向がある。プロジェクタカメラ方式は、高精度に短い時間で計測ができるため、多くの産業現場にて採用されている方式であり、さまざまなメーカ



図1.3 株式会社リコー製の産業用ステレオカメラ。(Copyright(C)2016 RICOH [1])

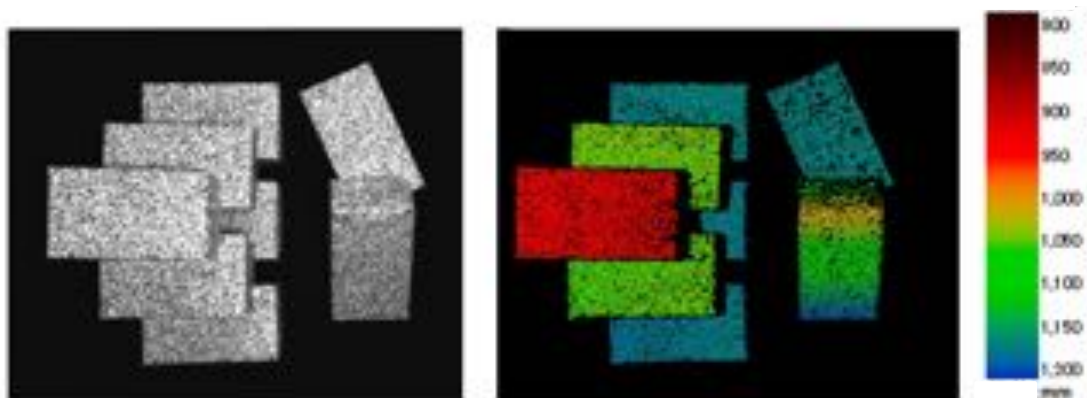


図1.4 株式会社リコー製の産業用ステレオカメラで得られたデプスマップ。(Copyright(C)2016 RICOH [1])

一がこの方式の製品を製造・販売している。この方式はさまざまなシーンで利用ができるが、対象の反射特性がランバート反射からかけ離れている場合は、計測が困難な課題が残っている。ToF (Time of Flight) 方式は、LiDAR (Light Detection and Ranging) が代表的な手法であり、システムで発生させた光を受光器で観測するまでの時間を用いて奥行きを推定する方式である。遠方の物体でも計測ができることが特徴であるため自動運転車の研究などで活用されているが、他の方式に比べ近傍物体に対する計測精度はやや低く、また計測密度も低いため、ファクトリーオ

オートメーションのシナリオではあまり採用されない方式である。

これらの方式の中でもプロジェクタカメラ方式の計測装置はファクトリーオートメーションと精度，速度の観点から相性が良いとされるが，安価なシステムにて対象シーンで発生する光学現象へのロバスト性を得る課題は解決できていない（図 1.5, 図 1.6）。

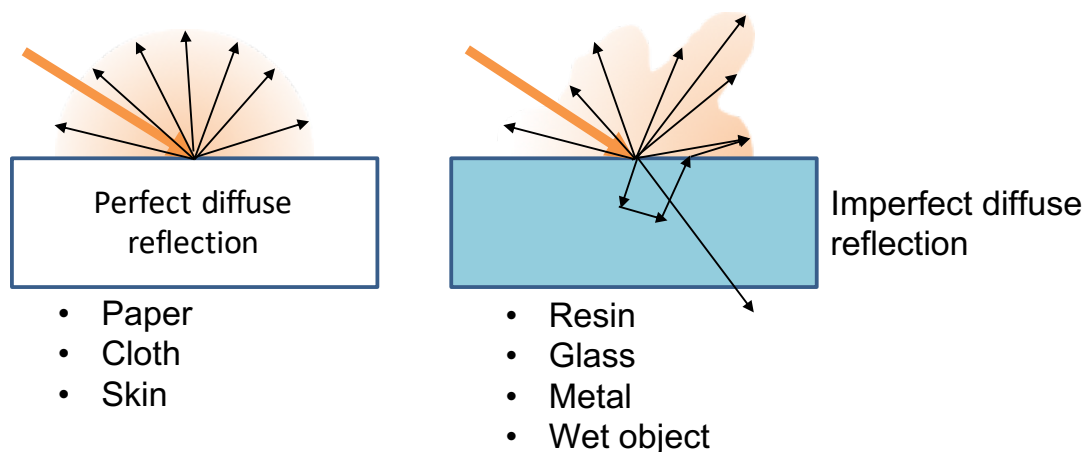


図1.5 対象物の材質によって起こる多様な光学現象例.

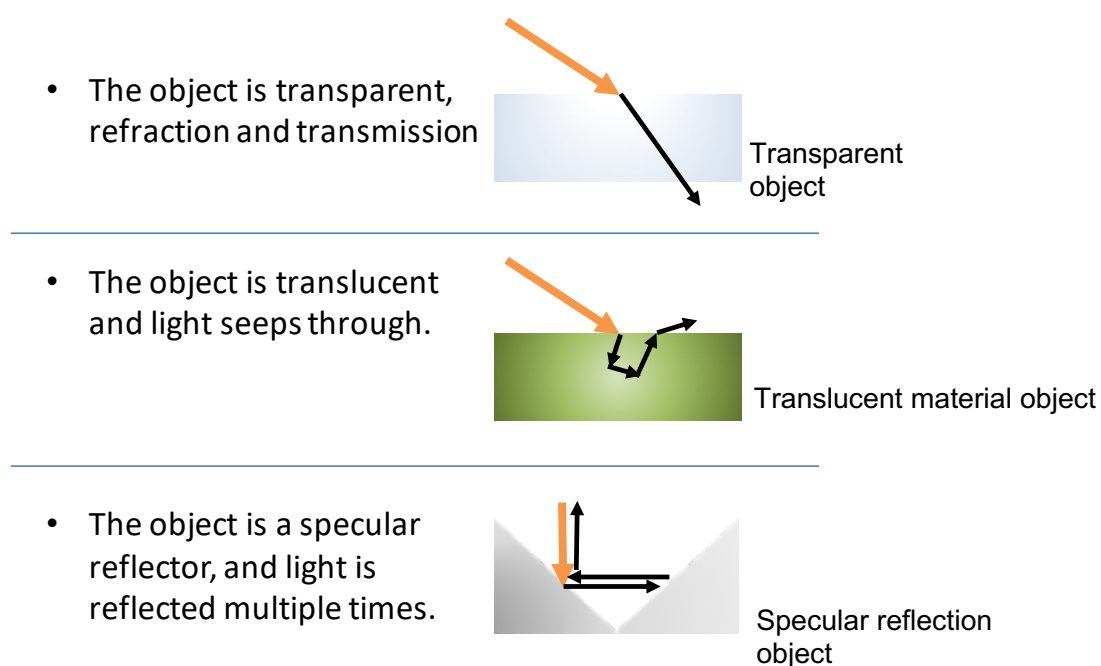


図1.6 透過，半透過，相互反射の光学現象イメージ.

前述したように市販製品や先行研究において活発に技術開発が進んでおり，さまざまなメーカー（GOM 社，カール・ツァイス社，Hexagon Metrology 社，Photoneo

社, LMI tech 社など) が参入し, 開発が進められている. しかし, これらの製品でも計測が困難な物体は未だ多く, 三次元計測の課題としては解消できていない部分が多いといえる. 例えば, 対象物の材質が金属のワークであった場合, 凹み部で相互反射現象が起きて計測が困難となってしまう. このような光学現象にロバストに対応でき, かつ, 安価な装置で実現できるようなシステムは, 既存技術では未だ困難といえ, 実応用の観点からも非常に需要が高く解決が必要な課題であるといえる.

1.3 プロジェクタカメラ方式による三次元計測

プロジェクタカメラ方式での三次元計測の基礎的な部分について説明する. 三次元計測の計算は以下の2つのステップで構成される.

- プロジェクタ・カメラ画像間の画素の対応付け
- 三角測量の原理による三次元計測

プロジェクタ・カメラ画像間の画素の対応付け

図 1.7はプロジェクタカメラシステムと, プロジェクタ・カメラ画像間の画素の対応付けのための撮影方法のイメージ図である. この図のように各投影座標にユニークな符号を埋め込むことができる投影パターンを複数投影・撮影することで, プロジェクタ画像の画素とカメラ画像の画素の1対1の対応点を得ることができる. ここでは広く活用されているグレイコード法について説明する [2]. グレイコード法は最小枚数で画像間の対応付けができる手法である. この手法では, N 回, 投影パターン画像を投影し, そのパターン画像が投影されたシーンを撮影する, という処理をして得た投影画像と撮影画像のセットを入力として用いる. グレイコード法の投影パターン画像は白黒2値で半分に塗り分けられており, 前時刻の投影画像で白領域だった箇所をさらに半分にして白黒2値に塗り分け, 黒領域に関しても同様の処理を行う. これを1画素単位での塗り分けになるまで行うことで, 各画素に固有の符号を与える投影パターン画像を生成することができる. このグレイコード法を用いた対応付けの方法は様々あるが, ベーシックには図 1.7にあるように, プロジェクタとカメラの配置が横であった場合は縦のストライプでパターン画像を生成することが多い. グレイコードパターン画像では線上で同一の符号を持つため, エピポラ幾何と呼ばれる幾何制約により得られるカメラ画像・プロジェクタ画像上のエピポラ線を用いて, グレイコード法のストライプ線とエピポラ線とが交差した点を対応点として決定することができる.

エピポラ幾何について説明する. 図 1.8のように三角測量の原理を用いて奥行き値を求めるステレオ視で三次元点の計算をする場合を考える. 対応点と測りたい物体上の三次元点の3点を結ぶと三角形ができる. この三角形がのっている平面はエ

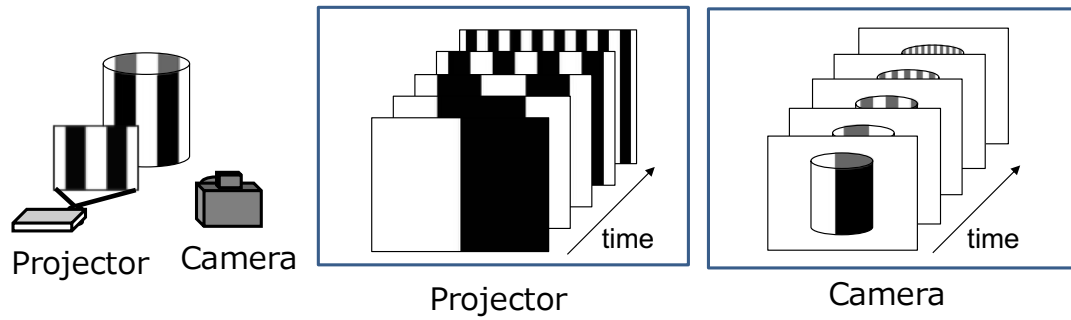


図1.7 プロジェクタ・カメラ画像間の画素の対応付けのための投影・撮影方法のイメージ。グレイコード法による対応付けのイメージ。左：計測シーンイメージ，中：プロジェクタの投影のイメージ，右：カメラの撮影のイメージ

エピポーラ平面と呼ばれる。このエピポーラ平面とプロジェクタ画像、および、カメラ画像とが交差してそれぞれの画像上に直線が得られる。この直線はエピポーラ線と呼ばれる。三角測量の原理からステレオ視で計測を行う場合、このエピポーラ線に対応点が存在するといえる。ゆえに前述したように、グレイコード法により符号化された画像上の1つの直線とエピポーラ線との交点を得ることで、対応点を求めることが可能である。

三角測量の原理による三次元計測

図 1.8はプロジェクタカメラシステムによる三角測量の原理を用いた三次元計測のイメージ図である。前述の方法でプロジェクタ・カメラ間の対応点が得られてい

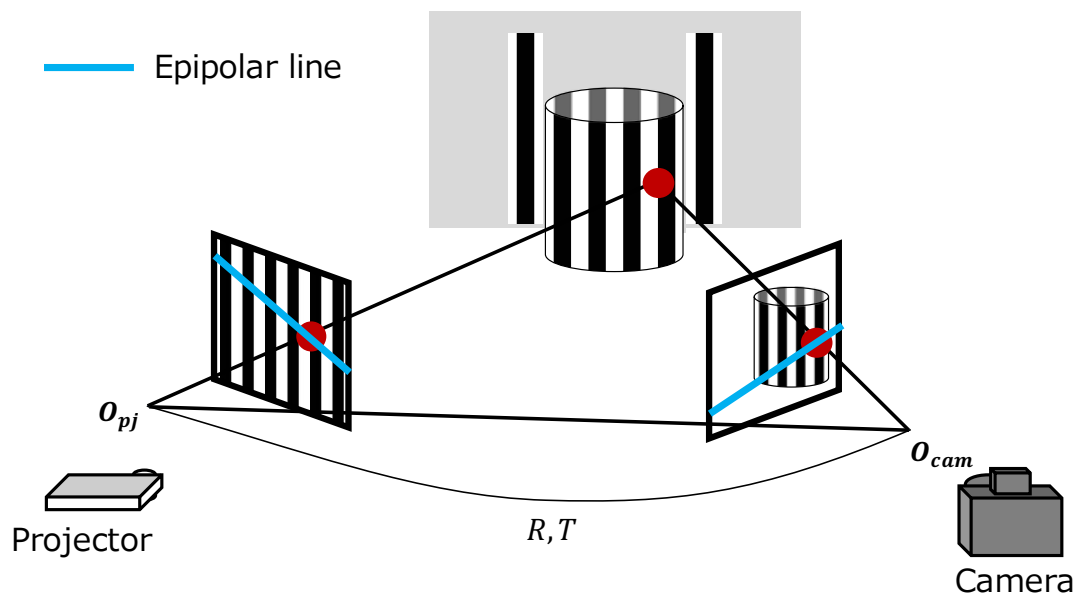


図1.8 三角測量による対応点の三次元計測。

る．この対応点を用いて，三次元計測の計算により対応点の三次元点を求める．ここではプロジェクタ，カメラそれぞれのデバイスが独自に持つ光学パラメタである内部パラメタ（センサ・レンズ間の光軸ずれ，レンズ収差），プロジェクタとカメラ間の位置関係パラメタである外部パラメタ（回転・並進パラメタ）については事前に較正済みで既知であるとする．対応点の三次元点を計算する上で用いる式は以下のようなになる．

$$m_p^T = \mathbf{P} M_p^T \quad (1.1)$$

$$\mathbf{P} = \mathbf{A}[\mathbf{R} \mid \mathbf{T}] \quad (1.2)$$

$$m_p = (u, v, 1)$$

$$M_p = (X, Y, Z, 1)$$

ここで $\mathbf{P}$ はカメラあるいはプロジェクタの透視投影行列である． $\mathbf{A}$ はカメラあるいはプロジェクタの内部パラメタ行列， $\mathbf{R}$ は回転行列， $\mathbf{T}$ は並進ベクトル， m_p はカメラ画像あるいはプロジェクタ画像上の対応点の座標値 (u, v) を表すベクトル， M_p は対応点の三次元点 (X, Y, Z) を表すベクトルである（図 1.8）．式 (1.1)，式 (1.2) がカメラ側，プロジェクタ側，それぞれで立てられる．2つの式 (1.1) を用いた連立方程式を解くことで，対応点の三次元点 M_p を求めることができる．

1.4 三次元認識とは

三次元認識技術も古くから研究が続けられており，三次元計測装置にて得られた対象シーンのデータ点の集まりをシーン点群データとして取得し，このシーン点群に対して CAD データなどから得たモデル点群をマッチングさせる方法で，所望の物体の認識をする手法が古くから研究されている（図 1.9）．さらに 5章でも後述するように，2次元画像で得られた知見をベースにしたキーポイントマッチング系の研究を応用した研究や，点群独自のデータ形式に特化した検出技術など，さまざまなアプローチで広く研究がなされてきている．また近年はディープラーニング技術を点群に応用し，特徴量抽出や点群マッチング，物体検出に活用している事例も増えてきている [3]．このように成熟しつつある技術群の一部をパッケージ化して，市販ライブラリとして MVTec 社の HALCON などが販売されており，ファクトリーオートメーションの分野でも三次元認識技術が活用されつつある．

しかし，工業的に活用するためには課題がまだ多く残されており，特にロボットシステム等の自動生産システムにてばら積みピッキングなどのタスクを実施したい場合に多くの工夫が必要とされる．例えば，ばら積みピッキングタスクの場合は対象物体の観測データの欠損が大きい傾向がある．このような場合にも安定的に対象物の位置，姿勢を把握することは実用上，クリティカルな課題となる．また，物体

でなく破片等の明確な形を成していない物の分類タスクも重要である。これは搬送中に破損した物品の破片などを回収し、材料ごとの自動分別処理を実現する上で不可欠である。このように実用を考える上では、多様な課題が残されているといえる。

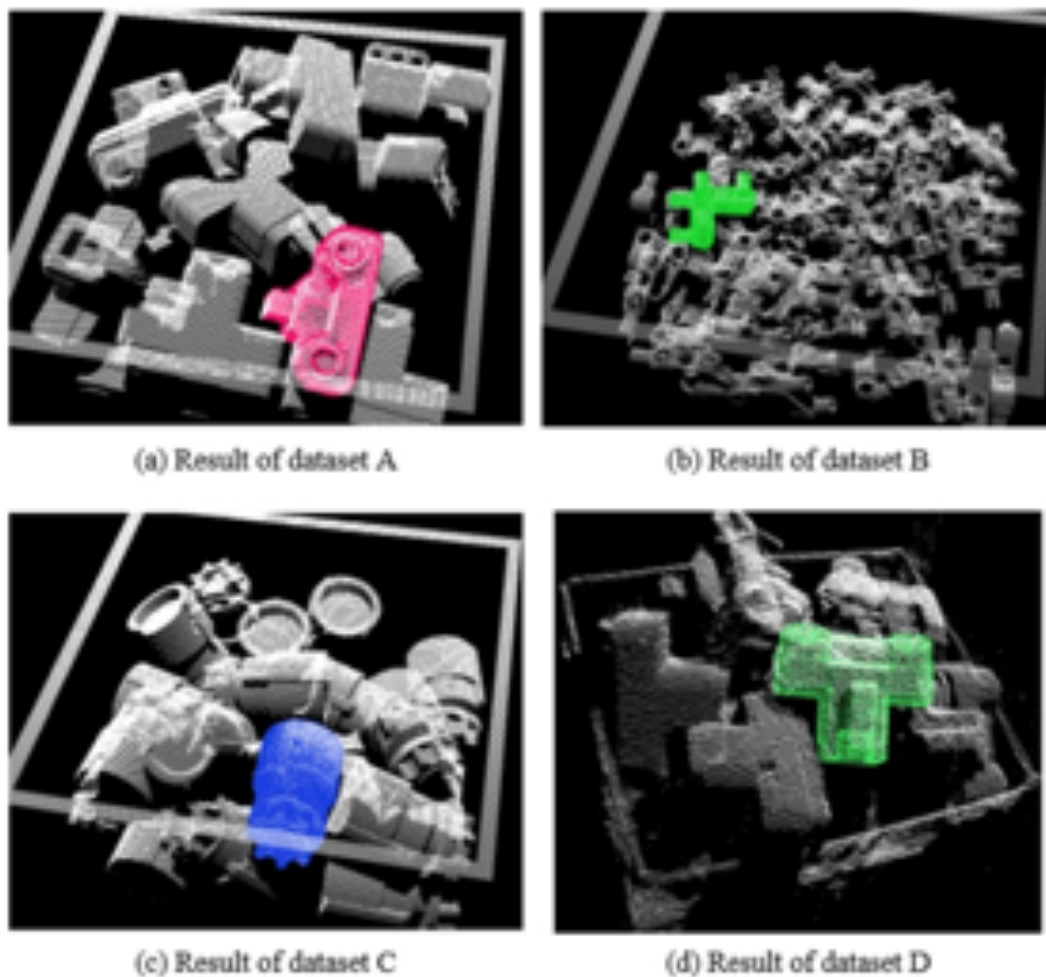


図1.9 点群による物体検出イメージ (Akizuki らの研究 [4] から引用)。

1.5 三次元計測および認識技術の産業応用における課題

三次元計測および三次元認識技術は、産業応用において広く活用されつつある技術である。しかし、前述したように、対象の物性や状態によっては計測や認識が困難なケースが多く存在するため、実用上は解決しなければならない課題が山積していると言える。そのため、本研究ではそのような課題の中で重要と思われるいくつかの課題について検討する。

本研究では、A. 対象の物性によって三次元計測が困難となる課題、および、B. 対象の認識が困難となる環境での三次元認識課題、について取り組む。

課題 A については，既存手法では金属物で起こる相互反射により計測が破綻する，あるいは，計測できるが非常に計測時間がかかり実時間で高速に計測することが困難な課題に取り組む．図 1.10 は計測が難しい物体の例である．これらの物体ではプロジェクタから投影したパターン画像が乱反射により崩れてしまい，既存法では実時間で対応点を得ることが困難となる課題がある．



図1.10 実時間で高速に三次元計測することが難しい物体例.

課題 B については，対象物に強い遮蔽が発生している場合に物体検出が困難となる課題（図 1.11），および，対象が物体特有の形状を成していない破片のような形状である場合に検出が困難である課題（図 1.12）について取り組む．前者は，検出対象物と三次元センサ（RGBD センサ）の間になんらかの物体があることで遮蔽されてしまい観測データに大きな欠測がある場合に，既存法では検出が困難となる課題がある．後者は，対象物体が特定の形状をもっていないために既存の物体認識システムを適用することが難しい場合にも検出を可能とする課題である．

1.6 本論文の構成

本論文の構成について述べる．1章にて，研究背景と目的について述べた．2章では，市販の DLP プロジェクタと高速度カメラのみを用いて，投影フレームレートよりも高速に計測ができ，かつ，対象で起こる光学現象のひとつである相互反射現象に頑健な手法を示す．3章では，市販のプロジェクタと市販のカメラにて実現可能な，2次元ランダムパターンを用いた高速かつ相互反射に高ロバストな計測手法に

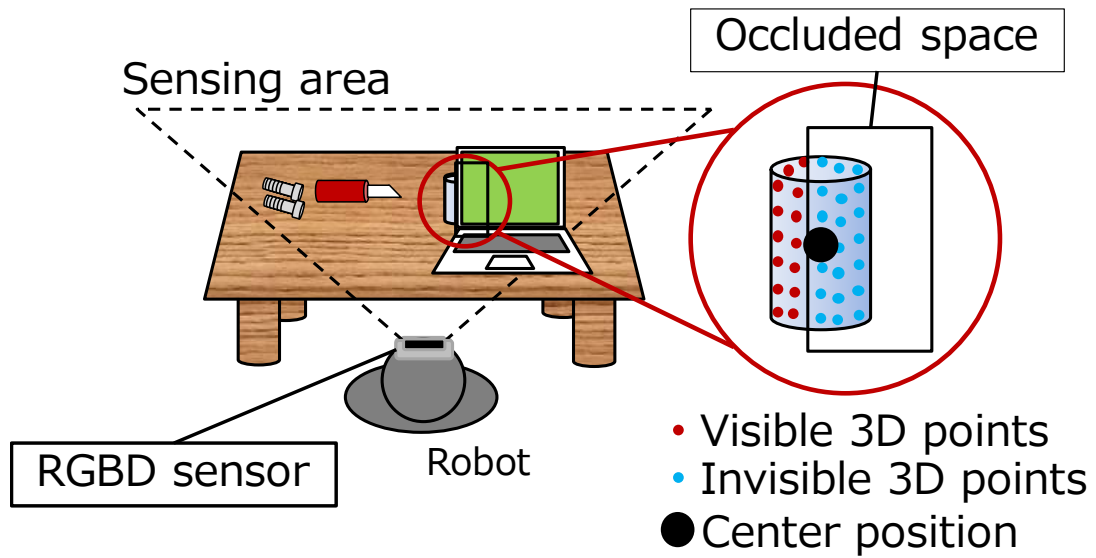


図1.11 大きく遮蔽された対象物の検出シーン.

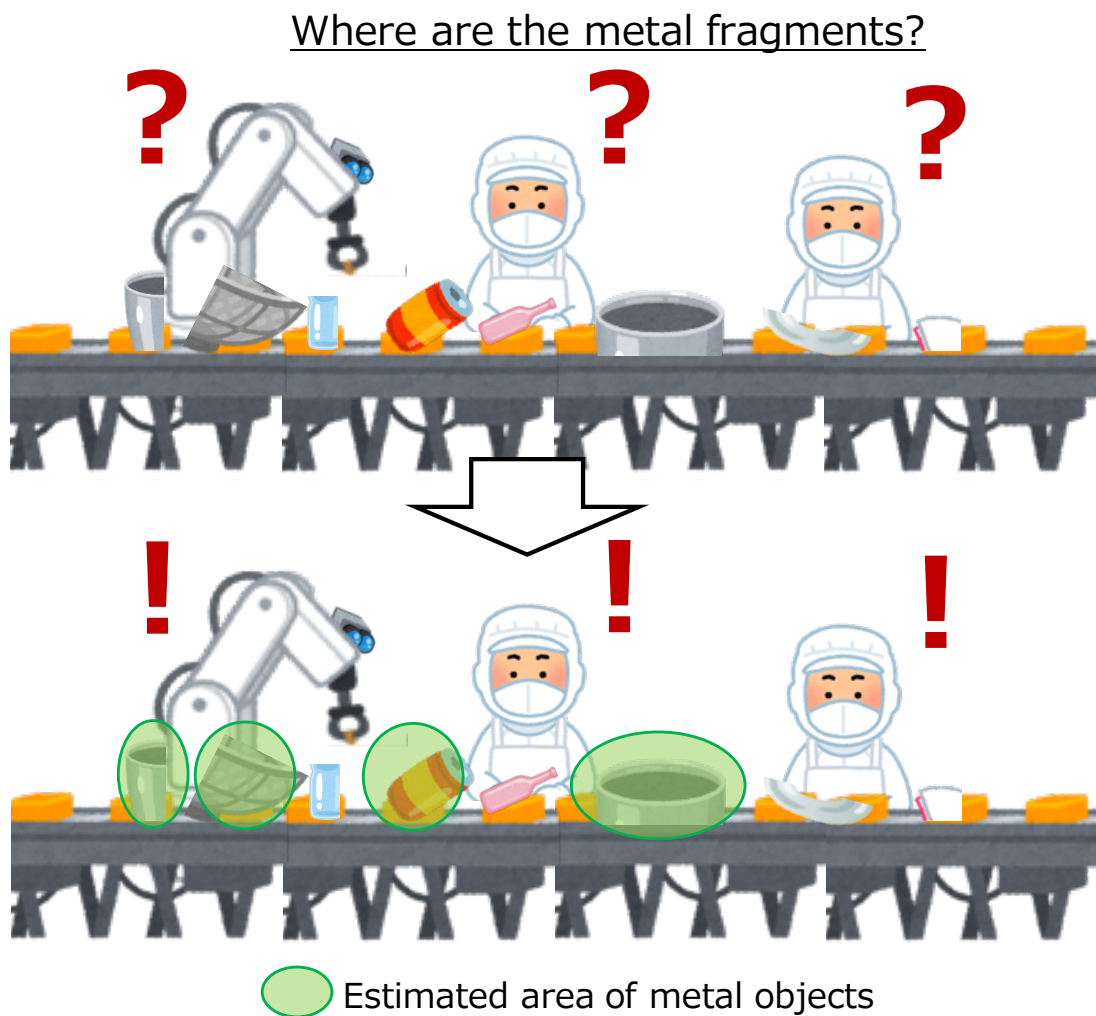


図1.12 定まった形状のない金属片の検出シーン.

ついて述べる．本章では実験的な精度評価の他，計測シーンにおける対象物の反射特性を考慮した観測モデルから，速度と精度のトレードオフの見積もりの議論も行う．4章では，工業的なシーンでは自然と既知とできる対象物の動きに着目し，対象物が既知の動きをする場合に市販の安価なプロジェクタと低FPSカメラを用いても，三次元計測を実現できる手法について述べる．5章では，大きく遮蔽された物体を3次元的に検出する手法について述べる．本手法は遮蔽を起こしている計測器に近い物体を検出し，その背後に仮想データを配置することで，ロバストな検出を可能とする．6章では，物体形状を成していない金属の破片のような物体がある領域を推定する認識技術について述べる．本手法は3章で述べる提案法の応用であり，金属物を精度良く計測しながら，どの領域に金属物があるかを検出できる手法である．最後にこれらの提案法に関するまとめを7章にて述べる．この章で提案法の結果から，A. 対象物性によって起こる光学現象の影響を抑制した計測が可能な三次元計測技術，および，B. 対象物の周辺物から受ける影響を抑制した三次元認識技術，に貢献する技術のまとめを示す．

第 2 章

提案手法 1：

DLP プロジェクタを用いた鏡面 相互反射光の影響を抑制可能な高 速アクティブ三次元計測手法

2.1 研究背景・目的

三次元計測技術は、生産現場において、ワークの傷や変形の検査やアームロボットでワークを把持するための基本的な技術である。中でもプロジェクタを用いた光計測技術は、装置系を小型に実現 [5–7] でき、計測時のワークへの侵襲が小さいという利点があり、工業的な応用価値が高い。プロジェクタカメラシステム（以下、ProCam）の技術として、グレイコード法と位相シフト法を組み合わせた手法 [2,8,9] や、特殊な投影パターンを用いた高速計測手法 [10] などが広く知られている。

ProCam を用いた一般的な計測法は、対象物の鏡面性が強い場合、対象物の凹部の三次元計測が難しい。鏡面性の強いワークは、プラスチックパーツや透明樹脂をコーティングした物体など様々あるが、ここでは工業的に広く扱われている金属ワークについて考える。[2,8,10] のような手法は、ワークの材質がランバート反射であるという仮定を置いている。金属はランバート反射ではなく鏡面性が強いので、図 2.1 のように、対象物の凹部では相互反射が生じる。結果、投影光の反射光がカメラとプロジェクタの配置関係からランバート反射物では観測されないはずの画素で観測されたり、カメラの一つの画素に複数の反射光が重なって観測されたりする。前者の場合、プロジェクタから投影した光の直接反射成分が観測されると推定される領域をカメラ画像のエピポーラ線上に限定できること（エピポーラ制約）を利用して、エピポーラ線外で観測される成分を間接反射成分として排除することで対応付

けができる手法が提案されている [11–13]. しかし後者の場合, 図 2.1 のようにプロジェクタ画像のエピポーラ線上の異なる画素からの投影光が, 鏡面相互反射によって一つのカメラ画素で観測されてしまうと, 前者のようにエピポーラ線を用いた間接反射成分の対応点の除去ができず, 原理的に既存法ではカメラ画像とプロジェクタ画像間の対応付けが難しい課題がある.

また, 既存法 [11–13] では光分離のロバスト性のために時間方向に多数のパターン画像の投影を必要とし, 光学的・撮影システム的に特殊な装置を用いなければ短い時間での計測が困難であることが課題である. ゆえに, 鏡面性が高い物体を ProCam で計測する場合, 光成分の分離がしやすい符号を少数の投影パターン数で実現することが課題となる.

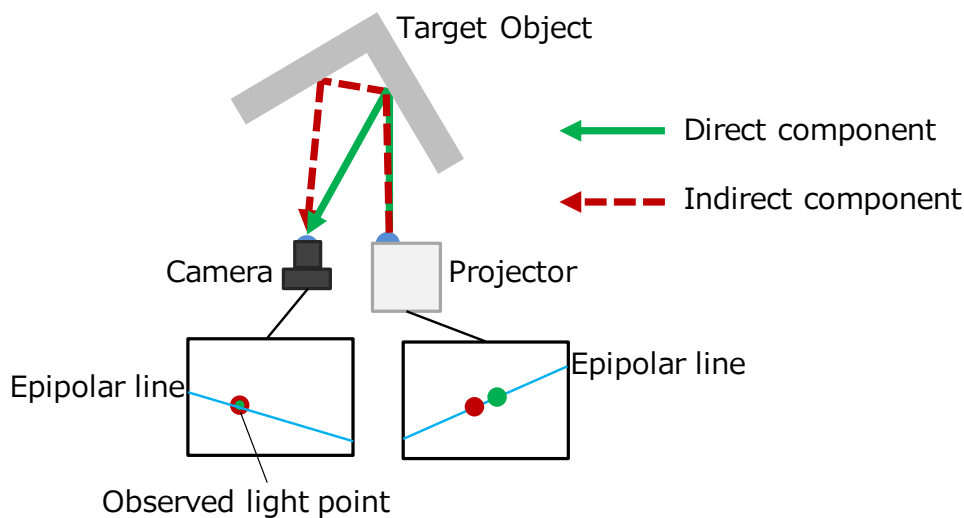


図2.1 直接反射成分と間接反射成分とが重疊したケース. プロジェクタ画像上のエピポーラ線上の2点から投影された光が, カメラ画像上の1点で観測されている.

2.2 関連研究

投影光に時空間方向の符号を与えることで, 直接反射成分と間接反射成分を分離する, あるいは, 分離して三次元計測する手法が提案されている. Nayar らは, ProCam を用いて高周波パターンを投影し, 直接反射成分と対象シーンで起こる拡散相互反射, 表面化散乱などの間接反射成分とに, 投影光の反射光を分離した画像を得る手法を提案している [11]. この手法は, 高周波成分を含み, 白黒それぞれを50%の割合で埋め込んだバイナリパターンとその反転パターンを用いる. 高周波パターンの中の白が投影された対象物体の点は, 直接反射成分と50%の間接反射成分との和の合成成分が観測されたと考えられる. また黒が投影された点は残りの50%

の間接反射成分が観測されると考えられる．これらの観測モデルをベースに複数枚の投影画像から計算を行い，観測画像を直接反射成分画像と間接反射成分画像に分離して取得できる．しかし，投影パターンが多数，必要であるため，時間コストが高い．Furuse らは [11] を応用し，エピポーラ制約と M 系列パターンとグレイコードパターンを合わせた画像により，相互反射を抑制可能な計測手法を提案している [12]．自己相関の高い M 系列を用いる，かつ，高周波成分を含ませることで直接反射成分か相互反射成分かの認識を可能としている．このため，直接反射成分の対応点を抽出でき，これを用いて三次元計測が可能である．M 系列は符号の重畳，その分離を考慮した系列であるため，相互反射を検知することができる．また，グレイコードパターンと M 系列パターンを組み合わせ，M 系列パターンのみでは多量の投影が必要なところを，小さな投影枚数で実現している．しかし，図 2.1 のようなケースは考慮しておらず対応できない．また，光分離を安定させるためには M 系列の信号長を長くするために投影枚数を増やす必要があり，高い分離性と高速性とを両立させることが難しい．また面計測への拡張が容易ではなく，スリット計測法では時間コストが大きい．O' Toole らは，[11] と同様の考えで，投影パターンをエピポーラ線でランダムにバイナリマスクしたパターンを，DMD で投影光をマスクできる特殊なシステムを用いて高速に多数，投影・撮影することで，直接反射成分と間接反射成分とに投影光を分離する手法を提案している [13]．このシステムで得た直接反射成分画像を用いて位相シフト法を適用し，対象シーンの三次元計測が可能であることを示している．より詳細に述べると，プロジェクタ画像上のひとつのエピポーラ線（プロジェクタエピポーラ線）と，この線と対応するカメラ画像上のエピポーラ線（カメラエピポーラ線）のペアがあり，プロジェクタエピポーラ線上の複数画素からの投影光がシーンを経てカメラエピポーラ線上で観測されるライトトランスポートを考えると，原理的に図 2.1 のケース以外の間接反射成分は観測されないことを用いて，直接反射成分を分離して観測している．しかし，[12] と同様に，図 2.1 のようなケースは考慮しておらず対応できない．また，直接反射と間接反射の分離性を上げるためには多くのパターン投影が必要であり，分離性と高速性のトレードオフの問題が残されている．これらの手法は直接反射成分を取得でき，かつ，高速化のための投影枚数低減のアルゴリズム，あるいは，特殊な装置を用いた高速投影・撮影システムによる工夫がなされている．しかし，鏡面性の強いシーンや図 2.1 のケースでは適用できない課題，および，投影光の分離をロバストにするためには符号長を長くするために投影枚数を増加させる必要があるため，高い符号の分離性と高速性とを兼ね備えることが難しい課題が残されている．

一方で，DLP プロジェクタを用いた ProCam による高速な符号化による対応付けが可能な手法が提案されている．Narasimhan らは，DLP プロジェクタ内部で用いられる DMD の高速な駆動特性に着目し，プロジェクタ画像とカメラ画像との座標

対応付けのための高速な符号化手法を提案している (Temporal Dithering Coding 法と呼ぶ) [14]. 一般にプロジェクタのフレームレートは 60 fps である. DMD の各ミラーはプロジェクタ画像の 1 画素に対応しており, それらが数千~数万 [Hz] で駆動し, プロジェクタ内のひとつ, あるいは, 複数の光源 (RGB など) の光のオンオフを制御して映像を生成する. このように制御することで, 投影光は人の目には多様な階調・色の画素として認識され, 投影像が入力映像と同じ画像として見える仕組みである. このように制御された投影光を高速度カメラで撮影すると, DMD の駆動によってごく短時間に明滅が繰り返される信号を観測できる. この信号は, DLP プロジェクタに入力された色 (RGB 値) により決定される仕組みをもっているため, [14] 手法ではこのルールをリバースエンジニアリングにより調べ上げ, 各色に対応する信号を取得し, この信号を座標符号化信号として符号化に用いることで, 高速な符号化を実現している.

また, Mirdehghan らは, シーンごとに動的に, 投影と撮影枚数を最小にできる投影パターンを生成することを目的とした研究に取り組んでいる [15]. この研究では, プロジェクタの投影範囲とカメラの撮影範囲が重なる計測可能領域の中で, アプリケーションにより決まるヒューリスティックに定めた復元範囲内での形状復元処理を許す制約を Geometry 制約と呼び, この制約を前提にシーンにおいて最適な最小の投影パターン生成について議論している. Geometry 制約を与えることで, 投影パターンとして周期パターンを採用することが可能となる. 周期パターンは, 投影画像全画素を符号化するよりも小さい符号化長で符号化ができる. ゆえに, パターン画像上の符号化領域を低減させ, 信号長を短くすることで, 投影・撮影枚数を最小化し, ProCam による計測の高速化を実現している.

既存の投影光を直接反射成分と間接反射成分に分離する手法は, 投影光に何らかの信号を時空間方向に与えることで光分離を実現している. 本研究では,

1. Temporal Dithering Coding 法を応用してリアルタイムレベルの高速な符号化を実現しつつ, 投影光の直接反射成分と間接反射成分の情報を分離取得できる
2. 復元範囲を限定する Geometry 制約を応用しこれを奥行き制約として用いることで, 周期パターンを用いても投影枚数を低減させつつ, 復元範囲外に復元される間接反射による復元点を除去して直接反射成分の復元点のみを残すことができるという二つの仮説を立てた.

2.3 鏡面相互反射を抑制可能な高速三次元計測手法

提案法は, DLP プロジェクタの高速なパターン投影を用いて, 鏡面相互反射が起こる物体の三次元計測を可能とする (図 2.2). 本手法は, 投影パターンに時空間的に埋め込んだ符号を用いて, 鏡面相互反射により直接反射成分と鏡面相互反射成分

とが重畳している点の復号化情報を分離することができる。ただし、鏡面相互反射により投影光が重畳した点において、復号化情報を分離できたとしても、図 2.1 のケースでは、エピポーラ制約のみでは直接反射成分か間接反射成分かを弁別することができない。図 2.1 のシーンのように、直接反射成分と相互反射成分の 2 つの復号化情報のうち、どちらが直接反射成分かの判別ができない点については、エピポーラ制約だけでなく、新たな奥行き制約という幾何制約を与えることで、直接反射成分の対応点を得る。この奥行き制約を課すことで、相互反射成分の復号化情報を排除し、直接反射成分のみの復号化情報を得ることが可能となる。このエピポーラ制約と奥行き制約の 2 つの幾何制約による弁別処理を全復号化点について行い、直接反射成分の対応点を得て形状復元をする。

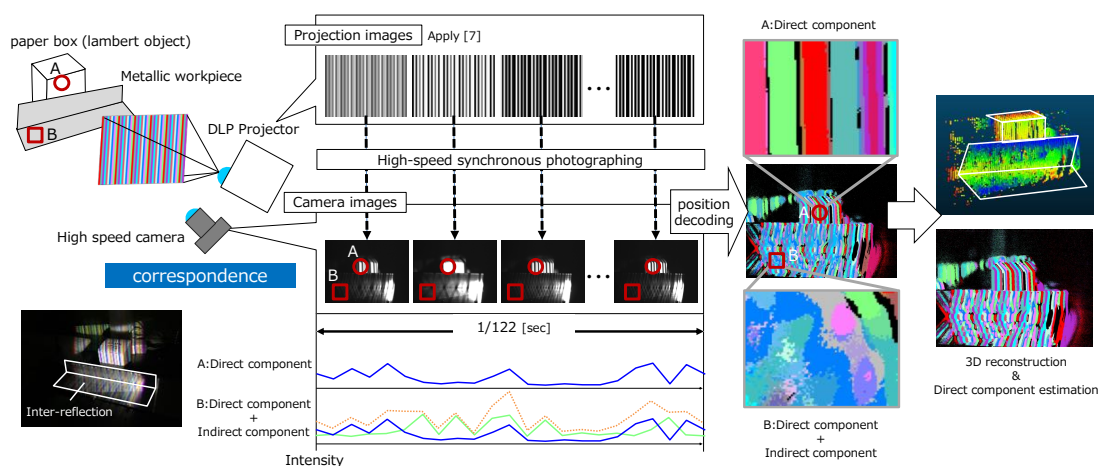


図2.2 提案法の概要図.

2.3.1 投影パターンの設計

Temporal Dithering Coding 法 [14] で用いる DLP プロジェクタの DMD の駆動パターンはメーカーが設計する投影信号パターンが組み込まれており、ユーザが設計できないため、Narasimhan ら [14] は正規化相互相関 (NCC : Normalized Cross Correlation) が相対的に低い座標符号化信号が発生する RGB 値を選別することで、頑健な符号化を実現している。予備実験から Narasimhan ら [14] が用いたプロジェクタとは異なる市販の DLP プロジェクタでも、符号化が可能であることを確認した (図 2.3)。この手法では対象物がランバート反射物であると仮定しており、基本的にこの条件のみで復号化が可能であると報告されている。ProCam を用いた反射光を用いた形状復元では、拡散反射成分の観測が必須であるが、本研究で扱う対象物の材質は金属であるため、鏡面反射方向以外の反射光の光量が少なく、拡散反射成分が相対的に低くなってしまう。そのため、投影するパターンは高い輝度値に設定す

る必要がある．一方で，投影光が相互反射等によって重畳することを想定するため，例えば2つの投影光が重なると考えて，2つの座標符号化信号の和の合成信号の最大値（最大輝度値）がカメラの感度限界を上回らないよう，設計する必要もある．

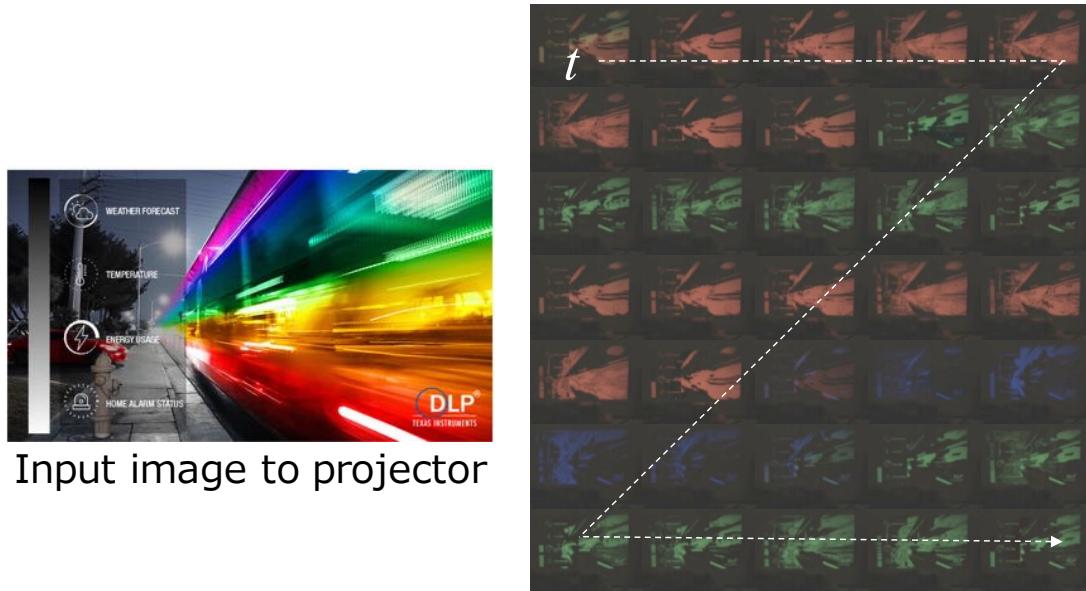


図2.3 DLP プロジェクタと高速度カメラによる符号化予備実験．4087 fps で DLP プロジェクタの映像を撮影した．35 フレームで 1 周期の輝度変化が生じ，この輝度変化を符号化に利用できるであろうことを確認した．また全画素の色を異なる色にすることで，全画素で符号化の可能性も確認できた．

そこで，本研究では投影パターンの設計指針を次のようにおき，投影パターンを生成した（図 2.4左）．次のような条件を課すものとする．

1. 各色の座標符号化信号の NCC 値が相対的に低い色を選別する（図 2.4右），
 2. 可能な限り，RGB の輝度値が高い色を採用する，
 3. 可能な限り，座標符号化信号の和の合成信号がカメラの感度限界を超えない，
2. は，本研究での計測対象物の材質が金属であり，反射特性として鏡面性の高い材質であるために考慮すべき条件である．カメラで観測される直接反射成分は，金属表面の鏡面性の強さのため，鏡面反射方向以外では小さいと仮定し，ここでは RGB の値のうち少なくとも 1 つが 180 以上となるランダムな色を生成し，上記指針に適合する色を選別するものとする．3. は，重畳点の投影光の復号化情報を，直接反射成分と相互反射成分の復号化情報に分離するために課した条件である．投影光の反射経路は物体形状や物体表面の反射特性に依存し，ひとつのカメラ画素に 2 重，3 重し得るが，本研究では光強度が強いと想定される 2 重までの重畳のみを仮定し，2 つの座標符号化信号が重畳した場合に，カメラの感度限界を超えないよう設計する．

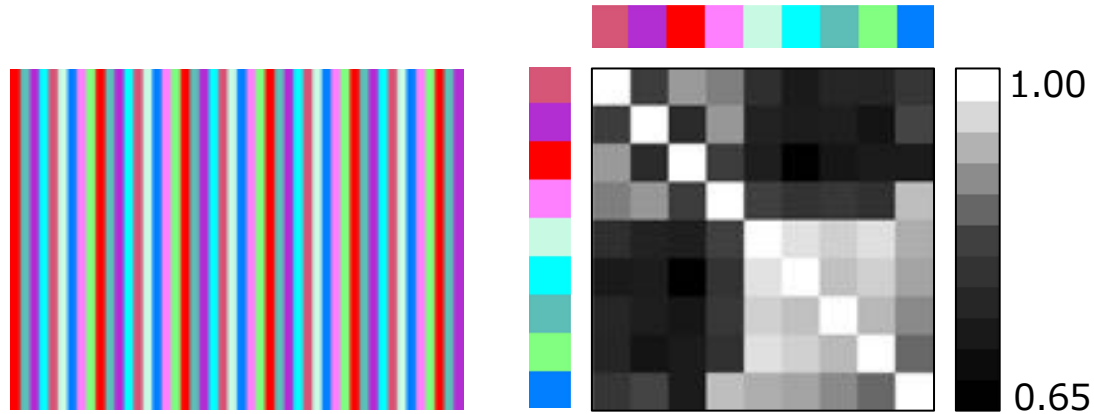


図2.4 色を9種用いたプロジェクタからの投影パターン画像と、各色ごとに出力される座標符号化信号同士の NCC 値の混同行列.

2.3.2 幾何制約による鏡面相互反射点の除去

このような設計指針に基づいて設計された RGB 値を用いたストライプパターン (図 2.4 左) を投影し、信号の分離度の高い座標符号化信号 (図 2.5) を発生させることで、カメラ画像中で直接反射成分と相互反射成分が重畳している点を認識することができる. 図 2.6 の系で図 2.7 のようにストライプパターンを投影して撮影された図 2.8 の入力画像列を用いて得られた、復号化結果を図 2.9 に示す. この図では直接反射の点には座標符号化信号を発生させた色をつけ、2 つの信号の和の合成信号が観測されている点はそれぞれの座標符号化信号を発生させた色の RGB 値それぞれを足して 2 で割った色 (混色) をつけて表示している. 計測シーンの対象物である L 字アルミアングルでは鏡面相互反射が発生しているため、直接反射したストライプと鏡面相互反射にて折り返して反射してきたストライプとが重なっている状況である (図 2.7). これを考慮して図 2.9 を見ると、2 つの色のストライプが重畳した箇所ですれら 2 つの色で構成された混色がある様子が見て取れる. このため提案法で正しく 2 つの座標符号化信号を分離して推定できていることがわかる. この分離された復号化結果とエピポラ幾何を用いてカメラ画像とプロジェクタ画像の画素の対応付けを行い三次元計測を行う.

このとき図 2.1 のような場合、エピポラ幾何だけでは、これら 2 つに分離された復号化情報のどちらが直接反射成分による復号化情報か不定となる. そのため、図 2.10 に示すように、三次元計測をする空間について奥行き方向に復元制限を与える奥行き制約を設ける. この制約により、鏡面相互反射成分の対応点から復元される復元領域外の三次元点を排除することができ、復元結果に直接反射成分の対応点を用いた三次元計測点のみを残すことができる. そのため提案法は最終的な結果とし

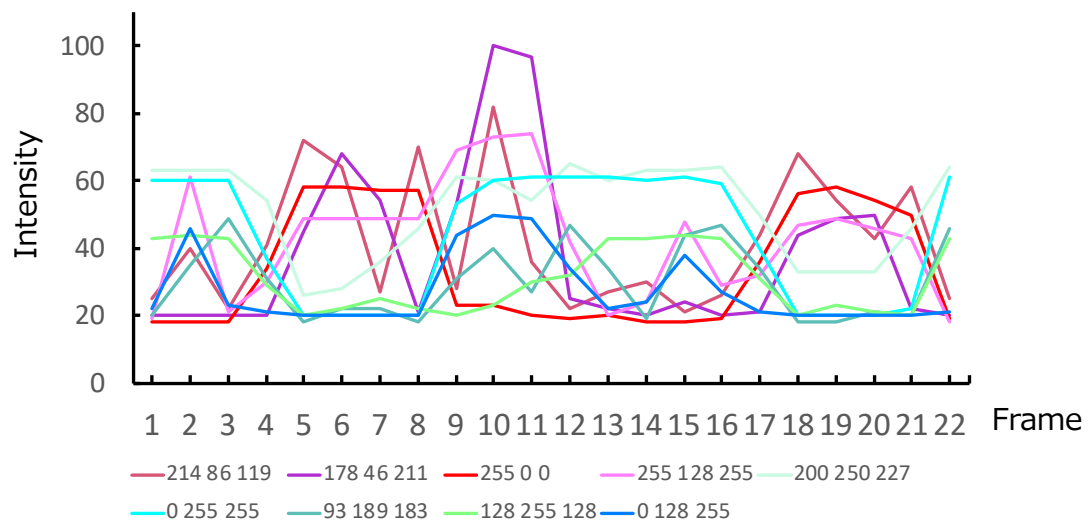


図2.5 各色の座標符号化信号のキャリブレーション結果. 本実験で用いたプロジェクタ : DLP2000 では, 2688 fps で撮影したカメラ画像のうち, 22 フレームで 1 周期する信号が出力される.

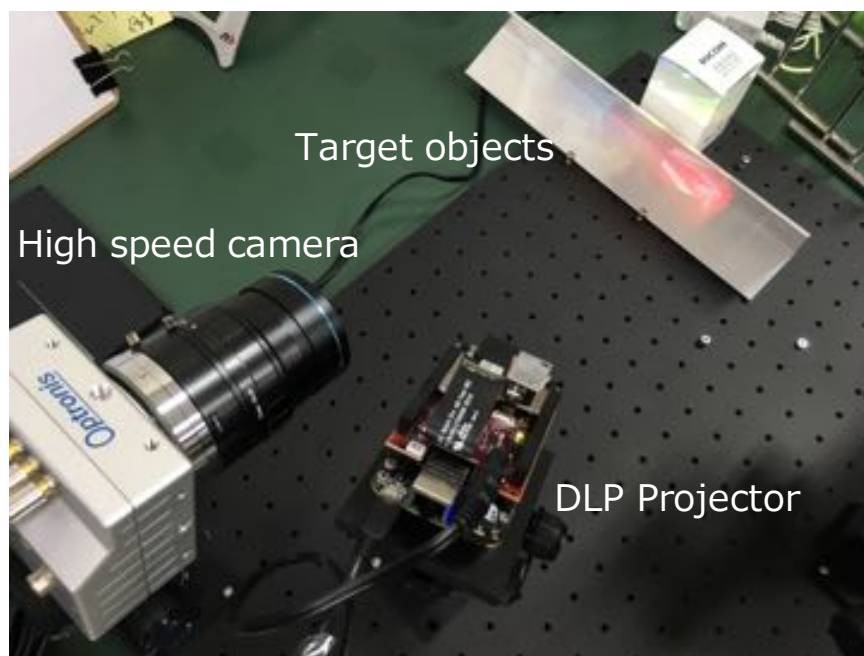


図2.6 本実験で用いるプロジェクタ, カメラ, 計測対象物.

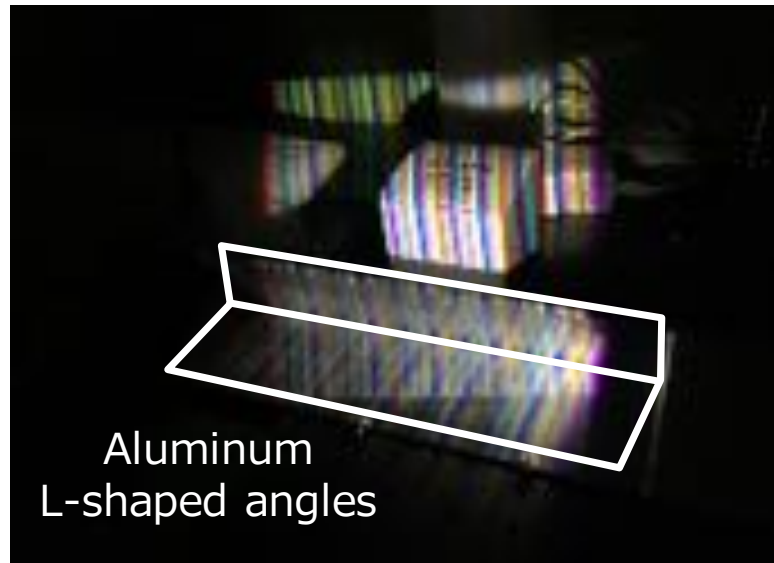


図2.7 図 2.4左のパターン投影時に発生した鏡面相互反射の様子. L 字アングルの上部に当たった光が，下部に反射して光が重畳している．

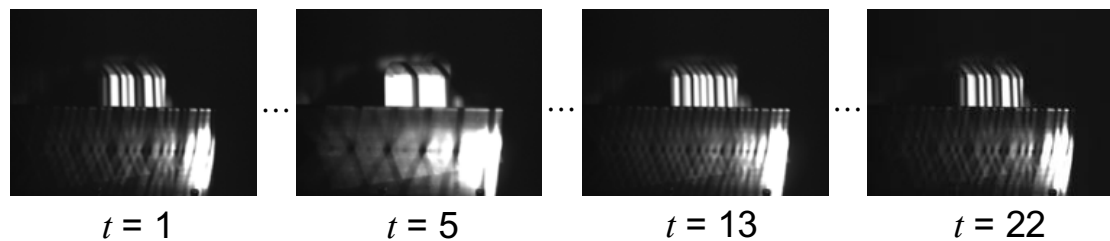


図2.8 復号化に用いる高速度カメラ画像列. t はフレーム番号で，フレームを抜粋して掲載．これらの画像は，対象物に図 2.4左の投影パターン光を当て，それを高速度カメラで撮影して得た画像列である．

て，直接反射成分のみの三次元計測結果を取得できる（図 2.11）．

2.4 実験と評価

提案法の有効性を確かめるため，復号化情報の分離性能の検証，および，既存法よりも高速に計測が可能かの検証をするため，直接反射と相互反射が起こる環境を想定し実験を行った．

2.4.1 実験条件

実験系は図 2.6のようにし，相互反射が起こる計測対象物としてアルミ製の L 字アングルと，ランバート反射物である白い紙箱を用いた（図 2.6右上）．高速度カメラは，Optronis 社のモノクロカメラ：CP70-1HS-M-1900 を用いた．解像度は 1024×768

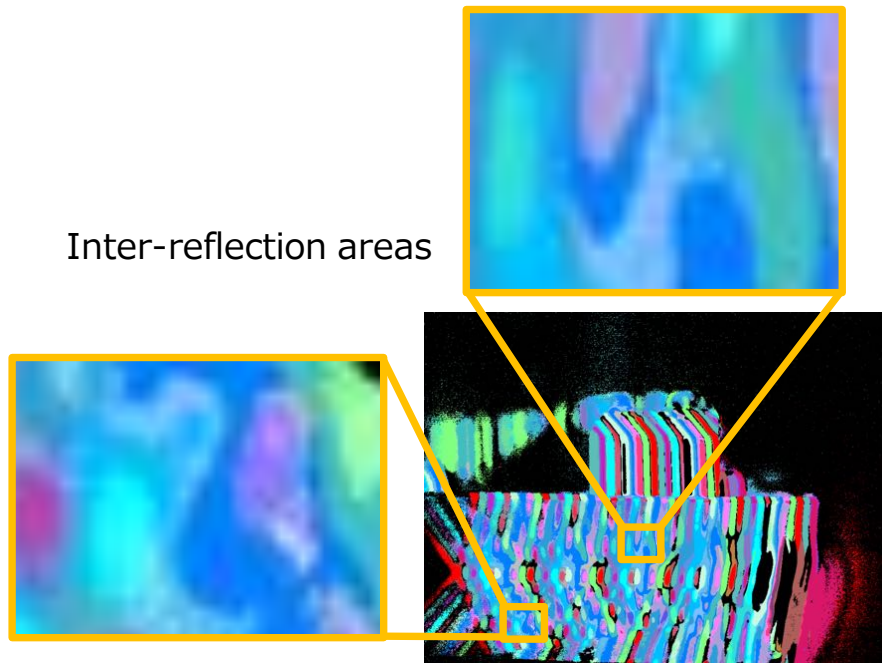


図2.9 投影光の復号化結果. 投影パターン（図 2.4左）の 9 色のいずれかの色の点は直接反射の点を表す. それら 9 色の中の 2 色が混色した点は, 相互反射で投影光が重畳した点, あるいは, 物体上の対象とする点の近傍の投影光が重畳した点を表す. 本実験系では前者の比率が大きい. 提案法では重畳点が 2 種の座標符号化信号の和の合成信号として認識され, 定性的に正しい復号化情報を推定している.

pixels, 撮影フレームレートは 2688 fps, 露光時間は $368 \mu\text{sec}$, カメラのレンズは F 値 : 0.9, の設定で用いた. プロジェクタは, TEXAS INSTRUMENTS 社の DLP2000 を用いた. 解像度は 480×360 pixels の設定で用いた. これらで ProCam システムを構成した. 幾何的なキャリブレーションは Zhang のカメラキャリブレーションを用い, 参照物体を用いてカメラとプロジェクタの内部パラメタ, および, 外部パラメタを求めた [16]. 座標符号化信号のキャリブレーションは, アルミ板にカラーパターンを投影し, 各色をプロジェクタへ入力した際に発生する座標符号化信号を取得し用いた (図 2.5). カメラでのラインの認識のしやすさのため, 投影パターン (図 2.4左) の 1 本のラインの太さは 10 pixels とし, 対応付けには左から 5 pixels の位置の点のみを用いた. 既存法は, 60 fps のフレームレートのプロジェクタとカメラを用いた, グレイコード法と位相シフト法を組み合わせるシステムとした. 投影パターン画像は縦縞のパターンを用い, グレイコード法で 9 枚, 位相シフト法で 4 枚, 計 13 枚の投影と撮影を行って復号化をし, ProCam 間の対応付けを行った. また実験の簡略化のため, 提案法, 既存法ともに, カメラ画像上の走査時, 等間隔に 5 pixels につき 1 点, 復号化と三次元計測処理を行った.

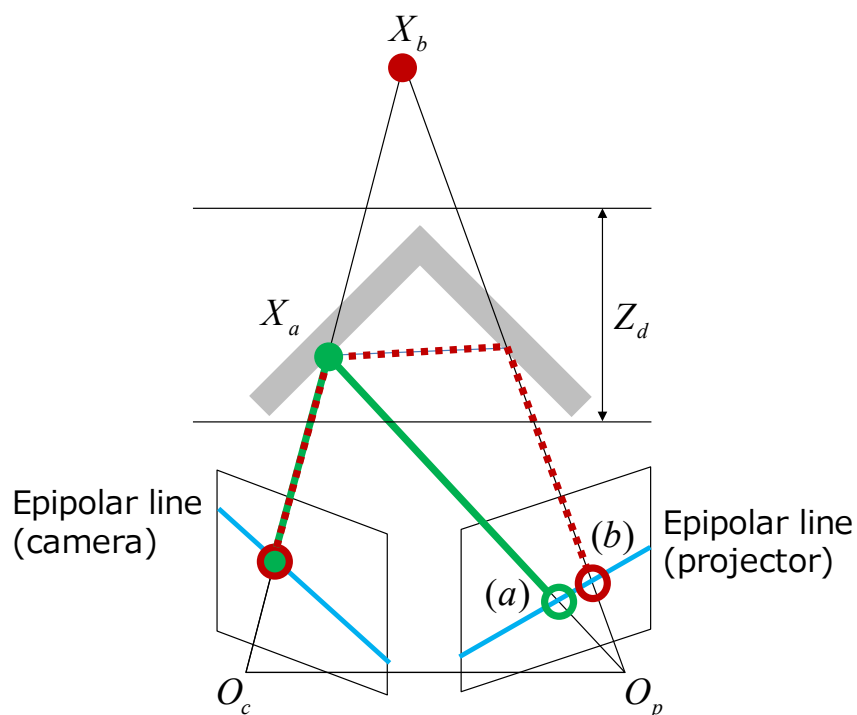


図2.10 奥行き制約の説明図. Z_d : 復元範囲 (奥行き方向), X_a : (a) の対応点を用いた復元点, X_b : (b) の対応点を用いた復元点, O_c : カメラ光学中心, O_p : プロジェクタ光学中心, である. 直接反射光の対応点 (a) , および, 鏡面相互反射光の対応点 (b) の2つの対応点のうち, (b) で3次元計測された点は, Z_d 外に復元されるため排除することができる.

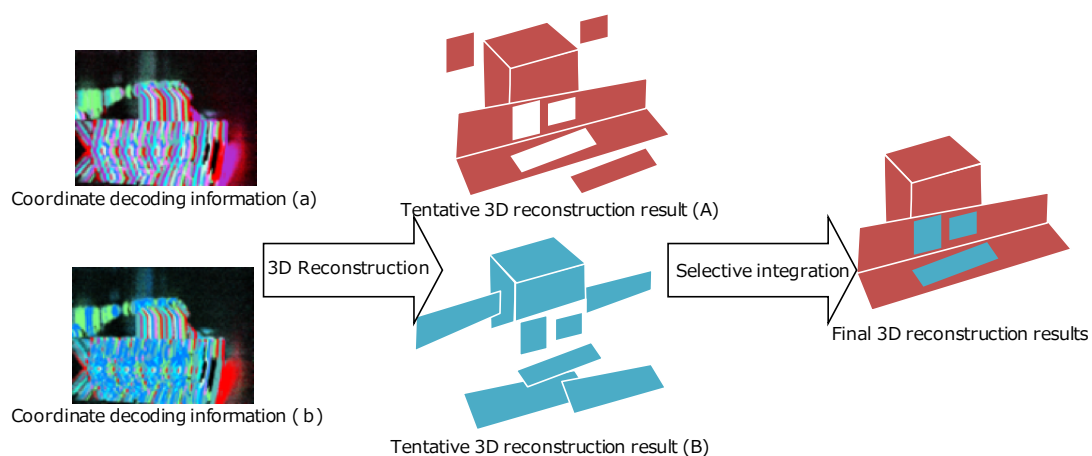


図2.11 提案法の形状復元のフローイメージ図.

2.4.2 復号化情報の分離と三次元計測性能の評価

DLP プロジェクタと高速度カメラにて得た画像列を用いて提案法による復号化を行い、三次元計測する実験を行った。復号化処理は、選別した 9 色から生成される 9 種の座標符号化信号（図 2.4, 2.5）と、それらの中からの 2 種の組み合わせ 36 種を合わせた 45 種の中から、最も NCC 値が高い信号を得ることで行う。本実験では、NCC の閾値を 0.6 とし、この値以上の点のみを復号化結果として用いた。三次元計測処理は、この復号化情報と、エピポーラ幾何を用いてプロジェクタ・カメラの画素を対応付け、奥行き制約により鏡面相互反射により生成される三次元計測点を排除しながら復元を行う。図 2.7 の投影シーンで撮影された入力画像列（図 2.8）を用いて得られた復号化結果を図 2.9 に示す。この結果から、提案法では定性的に正しい復号化情報の推定がなされていることがわかる。さらに、提案法と既存法の復号化結果を用いて得られる三次元計測結果を示す（図 2.12）。鏡面相互反射による重畳点（図 2.7）において、既存法では投影したパターン光の重畳を考慮していないため復号化が困難であり、三次元計測に失敗しているが、提案法では直接反射成分の対応点が見られるため、重畳点においても欠損の少ない計測が可能であることがわかる。

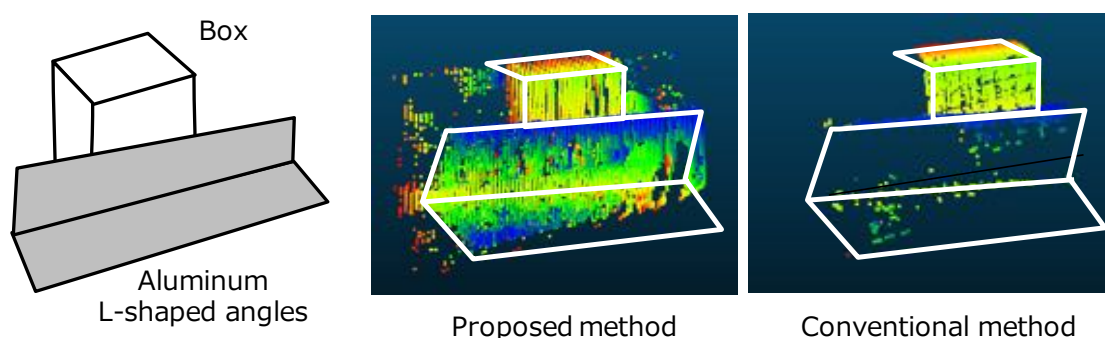


図2.12 提案法と既存法による三次元計測結果.

2.4.3 計測速度の評価

高速性に関して評価した。既存法では、カメラとプロジェクタともに 60 fps で、13 枚分の投影と撮影を行うため約 5 fps で観測データを取得し計測を行う。一方で提案法は [14] を応用し、DLP プロジェクタの DMD の駆動を座標符号化信号に応用しているため、約 122 fps で観測データを取得し計測が可能であり、既存法よりも高速に計測が可能であることを確認した。

2.5 まとめ

本研究では，ProCam を用いた高速な符号化手法を応用し，鏡面性の強い相互反射が発生する物体に有効な三次元計測手法を提案した．提案法は既存法よりも高速で計測が可能であり，さらに鏡面相互反射の影響を抑制して投影光の重畳点でも欠損の少ない三次元計測結果が得られることを確認した．今後の課題は，投影光の復号化情報の分離において，より複雑な投影光の重畳現象に強い符号化方式の開発，また他の材質物への適用の拡張であり，これらに取り組んでいく．

第 3 章

提案手法 2：

ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速アクティブ三次元計測法^{*1}

3.1 はじめに^{*11}

2章の背景・目的でも述べたように、ProCam を用いた計測法に共通する課題として、対象物の凹部において生じる相互反射の影響による誤計測の問題がある。工業的に用いられているワークには金型や切削加工品など、鏡面反射率が高くランバート反射とは大きくかけ離れた物体が多い。このようなワークでは凹み部で鏡面相互反射が生じ、これが誤計測の原因となる（図 3.1）。本論文で扱う鏡面相互反射現象の典型例について述べる。ひとつは図 3.1(a) に示すような、第 1 の反射面で生じる強い鏡面反射により、第 2 の反射面上で直接反射成分と相互反射成分が重畳する場合である。もうひとつは図 3.1(b) に示すように、第 2 の反射面が平滑であるために、拡散反射面である第 1 面に照射されたパターンが映り込む場合である。ただし図 3.1(b) の第 2 の反射面が完全な鏡面である場合のように、計測の対象となる面が拡散反射を全く生じない場合、直接反射成分がほとんど観測されず、三角測量に基づく形状計測は困難であるため本論文では対象外とする。

^{*1} 本章は、情報処理学会論文誌, Vol.62, No.8, Copyright(C)2021 IPSJ に掲載された同著者による”ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速アクティブ 3 次元計測法” [17] に基づいています。本章の素材は、情報処理学会論文誌, Vol.62, No.8 [17] で一部発表されたものであり、本章の一部の図は、情報処理学会誌の許可を得て、[17] から再利用されています。

^{*11} 本節の素材は、情報処理学会論文誌, Vol.62, No.8 [17] の 1 節に一部掲載されたものであり、本節の一部の図は、情報処理学会誌の許可を得て、[17] から再利用されています。

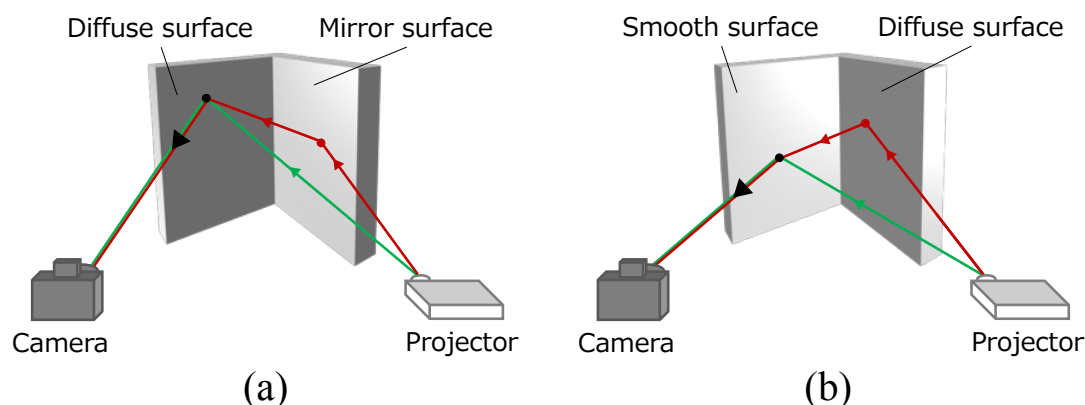


図3.1 直接反射成分と鏡面相互反射成分とが重畳する 2 つの例. (Copyright(C)2021 IPSJ, [17] 図 1)

第 1 の手法は、直接反射と間接反射の空間周波数特性の違いを利用する方法である．相互反射は間接反射の一種である．間接反射は観測点とは別の点に入射した光が回り込む現象であり、観測点の輝度は各点への入射光の強度の重み付き和で表される．よって多くの場合、間接反射は入射強度分布に対する平滑化フィルタと同様の効果を持ち、空間周波数が低いパターンが観測されやすい．この原理を用い、空間的に投影した高周波パターンが観測できるか否かで直接反射と間接反射を分離する手法 [11] と、これを 3 次元計測に応用した手法が提案されている．しかし、これらの手法は主に拡散反射物体の凹部で生じる拡散相互反射に対して有効であるが、図 3.1 に示すような、光路中に鏡面反射を含む鏡面相互反射現象の影響を十分に抑制することは困難である．

第 2 の手法として、重畳した投影光を、そのもととなるプロジェクタの各画素に分離する手法がある．一般に直接反射しか生じないシーンでは、カメラ画像上の各点に対しプロジェクタ画像上のただ 1 点だけが対応する．GC や PS はこのような性質を前提とすることで、少ない投影パターン数での 3 次元計測を実現している．それに対し鏡面相互反射が生じる物体では、カメラ画像上の 1 点に対し、プロジェクタ画像上の異なる 2 点以上の光が入射することがある（図 3.1）．よって第 2 の手法では、プロジェクタの各画素の輝度変化パターンに豊富な情報を埋め込むことで、反射点で投影光が重畳しカメラ画像上の 1 点で観測されたとしても分離ができ、直接反射の対応点を推定できる．しかし必然的に多くの投影パターン数を必要とし、計測に長時間を要したり、計測時間の短縮のために ProCam のフレームレートを高める必要が生じる．

そこで本研究では、鏡面相互反射成分を除去するために、独立な乱数系列は互いに無相関となることを利用し、投影された 2 次元ランダムパターンの相関演算に基

づき 3 次元計測を行う．ここで，相関演算を行う範囲に幾何学的な制約を導入することで投影パターン数を削減し，頑健かつ高速な 3 次元計測を実現する．さらに鏡面相互反射が起こるシーンでの提案法の投影パターンの観測モデルを作成する．このモデルを用いた数値解析から，実運用時に必要となる提案法の対象シーンと投影数による計測精度を見積もる．実験では，GC による投影枚数に近い 30 枚の投影パターン数でも鏡面相互反射の影響を除去しながら良好な 3 次元計測が行えることを示す．

3.2 関連研究^{*12}

ここでは ProCam に基づく 3 次元計測において，相互反射の影響を除去することで計測を頑健化する手法 [18] について述べる．簡単のため，以下ではプロジェクタとカメラが左右に並べられている（エピポーラ線がほぼ水平である）として記述する．

3.2.1 空間周波数特性を利用した手法

Nayar らは高周波パターンを複数投影し，反射光を直接反射成分と間接反射成分とに分離する手法を提案している [11]．この手法は 3 次元計測における相互反射の影響の抑制に応用できる．例えば Furuse らの研究 [12] ではスリット光投影法（以下，SLP）や GC の投影パターンの明部を，さらに M 系列で縦方向に符号化する手法を提案している．また同様に Gupta ら [19, 20] は PS や GC の投影パターンを高周波数化することで，同様の原理に基づき相互反射を抑制している．しかし，他の手法 [13, 21] も含め，これらの相互反射の空間周波数特性を利用した手法では，図 3.1 に示したようなシーンにおいて，鏡面相互反射の影響を十分に除去出来ない．なぜなら鏡面相互反射では非鏡面の間接反射と比較して影響するプロジェクタ画素数が少ないために平滑化フィルタのような効果が得られず，高周波パターンも観測されてしまうからである [11]．

3.2.2 エピポーラ面外からの伝播を抑制する手法

図 3.2(a) に示すように，直接反射成分は常にエピポーラ制約を満たすのに対し，鏡面相互反射成分は必ずしもエピポーラ制約を満たさない．この性質を用いることで，エピポーラ面外からの相互反射を抑制する手法が提案されている．例えば前述

^{*12} 本節の素材は，情報処理学会論文誌，Vol.62, No.8 [17] の 2 節に一部掲載されたものであり，本節の一部の図は，情報処理学会誌の許可を得て，[17] から再利用されています．

の Furuse ら [12] や O'Toole ら [22] の研究では、単純な高周波パターンの代わりに M 系列や乱数など、相異なる符号間が直交する、または直交することが期待できる符号を用いている。これは直交符号の自己相関がインパルス状となることを利用し、直接反射成分だけを選択的に抽出するためである。しかしこれらの手法ではいずれも、基本となる SLP や GC の各投影パターンをさらに縦方向に符号化するために、投影パターン数が増加し計測に時間を要するという問題を有する。例えば、Furuse らの研究 [12] では横 1024 画素のプロジェクタ画像を用いる場合、投影数 620 枚を要する。また図 3.2(b) に示すように、鏡面反射成分が偶然エピポーラ制約を満たす場合、これを除去することもできない。

O' Toole らは計測時間の問題を解決するため、CMOS 型画像センサのローリングシャッタ現象とレーザプロジェクタの走査を同期させた装置構成により、エピポーラ面外からの間接成分を除去する手法 EpiScan を提案した [22]。この手法ではフレームレートを低下させることがないという利点を持つが、プロジェクタ・カメラ双方の走査線をすべてエピポーラ線に一致させる必要があるなど精密な位置合わせが必要である、使用できる機器の種類が限定されるなどの問題がある。原理的には、プロジェクタからエピポーラ線に沿った水平線を投影し、その線上の画素値をカメラ画像から抽出する処理を繰り返すことと等価である。このため、投影パターン数が極めて多い計測法をハードウェアの工夫により高速化したものであると言える。

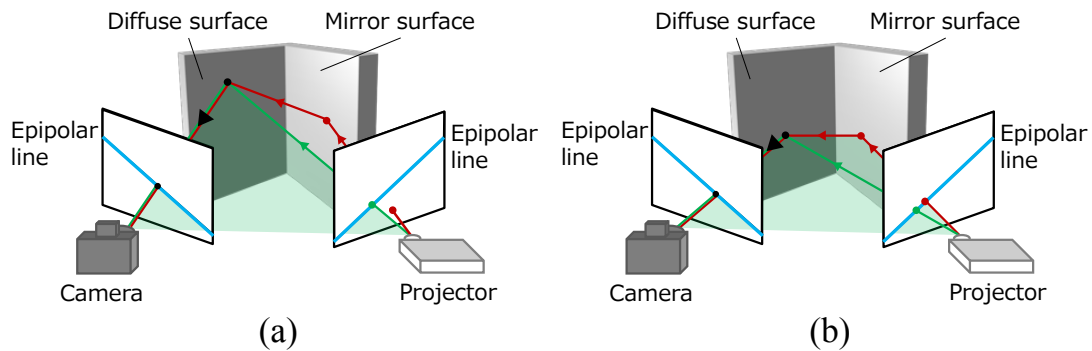


図3.2 鏡面相互反射成分がエピポーラ制約を満たさない場合と満たす場合。
(Copyright(C)2021 IPSJ, [17] 図 2)

3.2.3 エピポーラ面内の符号の重畳を分離する手法

PS や GC では、複数の符号が相互反射により重畳された部分では視差の復元が不安定になる。そこで、PS や GC と同様の縦縞状のパターンでありながら直交符号を横方向に配置し、複数の符号が重なった場合でも、それぞれの位置（視差）と強度が復元できる手法が提案されている [14, 15]。これらの手法では横 600 – 1000

画素程度のプロジェクタ画像を用いる場合、投影数に 100 枚程度を要する。

これらの手法は SLP の高速化手法とみなすことができる。SLP では、プロジェクタ画像の視差方向（横）の画素数を M_d とすると M_d 回の投影と撮影を行わなければならない（以下、これをブルートフォース法 (BF 法) と呼ぶこととする）。そのかわり、各カメラ画素では最も強い影響を受けるスリット番号を特定することが容易である。一方、GC は投影パターン数を $N_{GC} = \log_2 M_d$ に圧縮する手法であり、 M_d 通りの視差を決定するのに冗長性のない符号であるが、複数の符号が重畳されない、あるいは重畳があっても主たる符号の復号化が誤りなく行える程度に他の符号の影響が小さいことを前提としているために相互反射に対し脆弱である。直交符号に基づく手法はこれらの中間に位置する手法であるといえ、スリット光の重畳を許しつつも M_d よりも大幅に少ない投影パターン数での計測を可能にすることを志向した手法であるといえる。

3.2.4 2次元符号化

3.2.2節と 3.2.3節で述べた手法はそれぞれ、プロジェクタ画像の縦または横に直交性の高い符号を割り当てる方法であった。それらを組み合わせ、2次元的な符号化を行う手法が提案されている。

プロジェクタ画像とカメラ画像との対応関係を表すための表現法としてライトトランスポート（以下、LT）がある。LT は重ね合わせの原理に基づくシーンに対する光の入出力関係の記述である。ProCam において LT を取得するためには、プロジェクタの各画素を 1 つずつ点灯して、その都度、画像を撮影する方法がある。しかし、この手法（BF 法）には長時間を要するため、やはり符号化により高速化する手法が提案されている。Peers らは LT を、圧縮センシングアルゴリズムを用いて高速に取得・復元する手法を提案している [23]。また Chiba らは LT 行列が疎行列となる性質を用いて、ランダムパターンを複数回照射して得た観測行列から、投影光の直接反射成分を高速に抽出する手法を提案している [24, 25]。しかし 2 次元符号化では使用される符号の数が非常に多いため、1 次元的な符号化に比べ非常に多くの観測を必要とし、計測には長時間を要する。例えば、Chiba らの研究 [25] では 256×256 の解像度のカメラ・プロジェクタ画像を用いる場合、投影数 113 枚、投影・撮影および計算時間を合わせ 342 秒程度の時間（計算機のパフォーマンスにより変化する）を要する。また、相互反射が生じる場合は観測行列のスパース性が損なわれ、LT の復元が困難となる恐れがある。

Chiba らの研究 [25] と同様に、多数のランダムパターンを投影して得た観測画像に対し相関を求めることでロバストな対応点の取得と形状復元が可能であることが Zhang らによっても示されている [26]。また多くの圧縮センシングに関する研究で

も示されているように、ランダムパターンは M 系列などの数理的に定義された直交符号に比べ計測回数を柔軟に変更できることや、欠測に対して頑健であるなど様々な利点を有する。しかし前述のように、2次元符号化においては依然として多数の符号を用いる必要があり、従来用いられてきた3次元計測法と同等の速度で計測を行う手法はこれまでに示されていない。圧縮センシングを用いる手法では計測シーンの LT を復元するための最適化計算に大きな計算時間がかかる。またこの計算時、演算量の低減のために幾何制約を対応点候補数の削減に用いることも難しい。そこで本研究ではランダムパターンを用いた2次元符号化に幾何制約を組み合わせ、対応付けの候補数を削減することで、GC 法などに比べ遜色のない速度での計測を可能にしつつ、鏡面相互反射の影響も抑制できることを示す。提案法は図 3.1(b) のようなシーンで直接反射成分よりも相互反射成分が支配的となった場合にも、幾何制約を用いることで直接反射成分を同定しうる。また提案法は幾何制約を満たす領域内に限定して直接反射成分があるか否かを相関演算のみで求めるため、既存法よりも演算量を低く抑えることができる。

3.2.5 偏光等を利用する手法

3.2.2節で述べた手法の欠点であるエピソード面内の相互反射の問題を解決するため、前田らは EpiScan に偏光を組み合わせた手法を提案している [27]。このような光の物理的な性質に基づく手法は本論文が対象とする符号化に基づく手法とは排他的な関係になく、組み合わせることでより高度な除去を行うことができる。そのため、本論文ではこれ以上論じない。

3.3 提案手法^{*13}

本論文では 3.2.4節で述べた2次元符号化に基づく計測法に幾何学的な制約を組み合わせ、相関を演算する領域を限定することで、投影パターン数の削減による高速化と鏡面相互反射の抑制を両立させる手法を提案する。

3.3.1 2次元ランダムパターンの投影による対応点の取得

図 3.3は、2次元符号化と相関演算の組み合わせによる3次元計測の原理を示したものである。プロジェクタからは各画素が点灯または消灯の二値をとるランダムパターンを投影する。二値のランダムパターンを投影するごとにカメラで撮影を行

^{*13} 本節の素材は、情報処理学会論文誌, Vol.62, No.8 [17] の3節に一部掲載されたものであり、本節の一部の図は、情報処理学会誌の許可を得て、[17] から再利用されています。

い、得られた各画素の輝度変化パターンをプロジェクタ画像の各画素のパターンと相関演算により比較し、相関値の高い画素を対応点とする．物体の反射率や面の向きなどによる輝度変化の影響を抑制するために、相関演算には ZNCC (Zero-mean Normalized Cross-Correlation) を用いる．

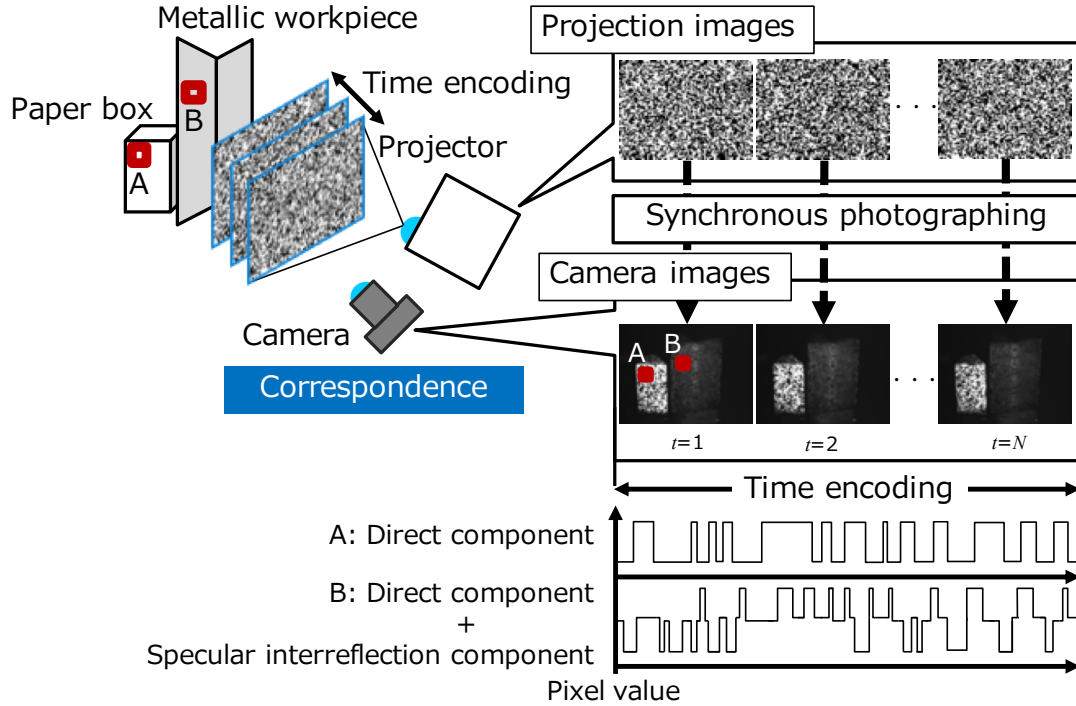


図3.3 提案法のシステムイメージ．2次元ランダムパターンを複数枚、投影・撮影し、時間符号化を行う．これにより鏡面相互反射で投影光同士が重畳した点の情報源分離を可能とする．（Copyright(C)2021 IPSJ, [17] 図3）

この手法では3.2.4節で述べたように、BF法に比べ少ない枚数でLTの近似値を得ることができる．しかし後述するように、相関値のノイズ（分散）と投影パターン数はトレードオフの関係にあり、十分な精度でLTを得るためには依然として多くの投影パターン数が必要である．一方、本論文の問題設定では対象シーンで起こる光学現象を表すLTを求める必要はなく、三次元計測を行う上で必須な情報である直接反射成分を有するかどうかの二値のみを推定すればよい．これを幾何学的制約を満たす範囲についてのみ取得すれば対応点を求めることができる．よって相関を求める範囲を限定することで、相互反射成分を効果的に除去しつつ、投影パターン数の削減を実現する．

ここでランダムパターンを用いる対応付け手法の符号化特性を調べるため、ランバート反射特性を持つ平面に対して予備実験をおこなった．図3.4の系を用いて、図3.5の復号化結果を得た．この結果は投影枚数 N を変化させた際の観測信号の復号化結果である．図3.5の左上の画像は各投影座標の符号化信号にグラデーションで

着色をした画像である，図 3.4の系で投影座標を符号化したパターン画像を複数投影し，投影に同期して投影されたシーンを複数撮影する．この複数の投影・撮影画像からカメラ画像上で投影座標を復号化する．図 3.4の系では投影対象物の平面をプロジェクタに近い視点に配置されたカメラで撮影しているため，カメラ画像上で正しく復号化がなされれば，左上の画像のようなグラデーションで復号化結果が得られる．ゆえに図 3.5の左上の画像のようなグラデーションが復号できた場合，投影パターンの符号化信号を正しく復号化できていると判断できる．次に，この復号化をした際の相関値を可視化した（図 3.6）．予備実験で用いた投影画像に用いた符号化信号数は $M_e = 6912$ であったため，理論的には $N = \log_2 M_e \simeq 13$ 以上であれば復号化可能である．図 3.6から，投影領域とそれ以外の領域との ZNCC の値の差が少なくなる $N=10-30$ の間に，高い信頼度で対応付けが可能な投影枚数の下限があることが推測された．



図3.4 平面に対する復号化実験の系．ProCam と計測対象物である平面があるシーン．複数のランダムパターンを投影&撮影する．右のような，撮影画像が得られる．

また相互反射が起こる物体に対しても同様の実験をおこなった．図 3.7のアルミ L 字アングルを計測する実験系を用いて，図 3.8の復号化結果を得た．図 3.5のようなランバート反射物の場合の結果とは異なり，相互反射する物体においては L 字アングルの凹み部で相互反射が起き，この領域で復号化が正しく行われない傾向があることがわかる．これは相互反射が起こる領域では複数の投影光が重なり合ってしまうために投影パターンの各画素に埋め込んだランダム符号が壊れてしまうためであり，互いの符号が無相関となるランダム符号であっても高い相関を得ることが難しいことがわかる．相関値のマップを見ても高い相関を得る点がないことからその傾向が見て取れる（図 3.9）．

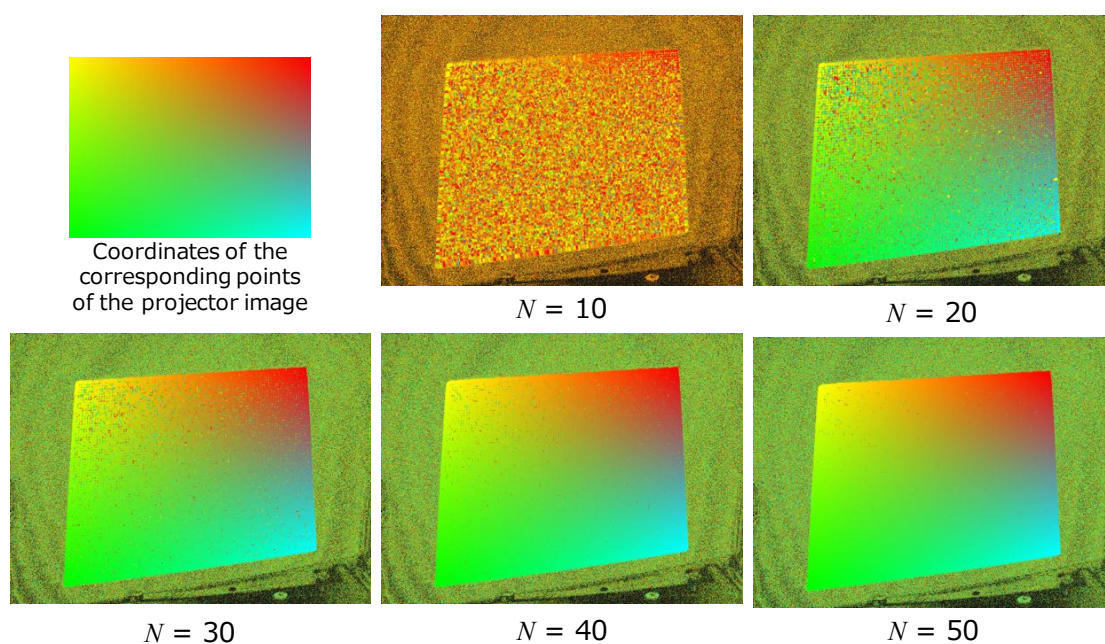


図3.5 平面における投影枚数 N を変化させたときの復号化結果の変化. N が 10 の場合, ドットの数に対し, 符号長 (投影枚数) が不足しているため, 復号化が破綻している. これ以外の N では, 概ね正しく復号化がなされている.

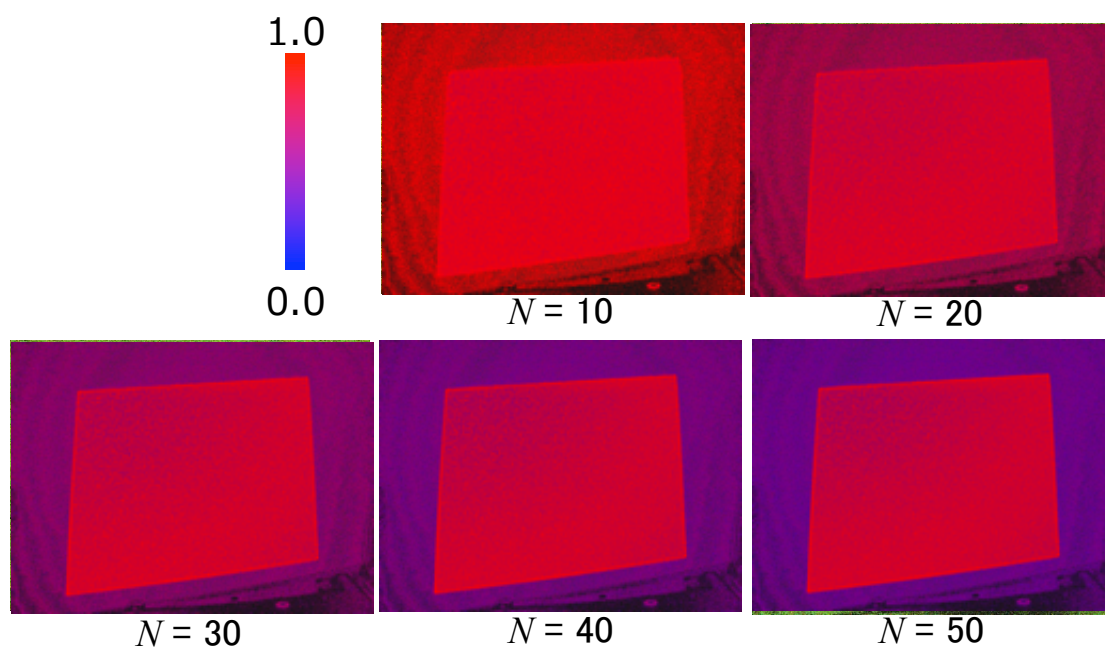


図3.6 平面における投影枚数 N を変化させたときの ZNCC の変化. N を増加させることで, 投影領域とそれ以外の領域との ZNCC の値の差が大きくなっていくことが分かる.

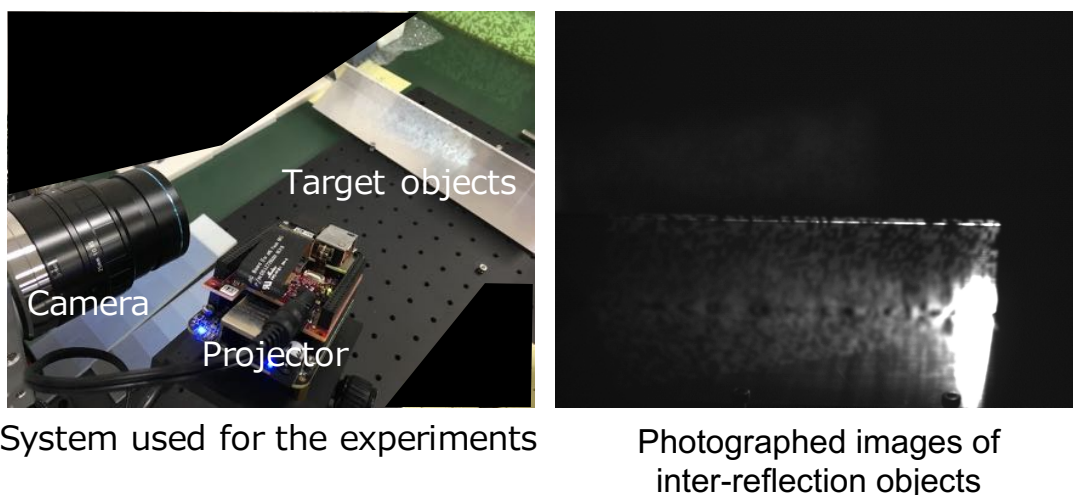


図3.7 実験系と相互反射物の撮影画像. 左：実験系，プロジェクタで複数のランダムパターンを対象物（L 字アルミアングル）に当て，そのシーンを同期撮影するシステムで構成される．右：カメラ画像のひとつ，画像下側に対象物である L 字アルミアングルが置かれており，それにランダムパターンが投影されているシーンを撮影している．

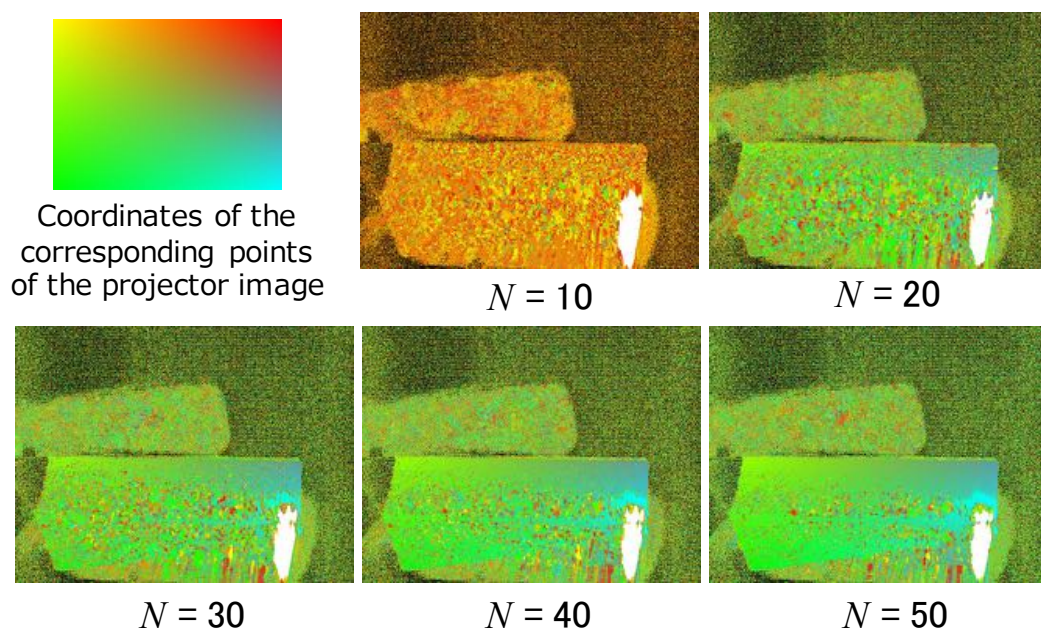


図3.8 相互反射物における投影枚数 N を変化させたときの復号化結果の変化. N が 10 の場合，ドットの数に対し，符号長（投影枚数）が不足しているため，復号化が破綻している．これ以外の N では，概ね正しく復号化がなされている．しかし，凹み部周辺ではグラデーションが乱れているため，復号化に失敗している．また画像右下にある白い領域は，カメラ撮影時に対象の鏡面反射による白飛びで，正しく符号化パターンを検出できなかった領域である．

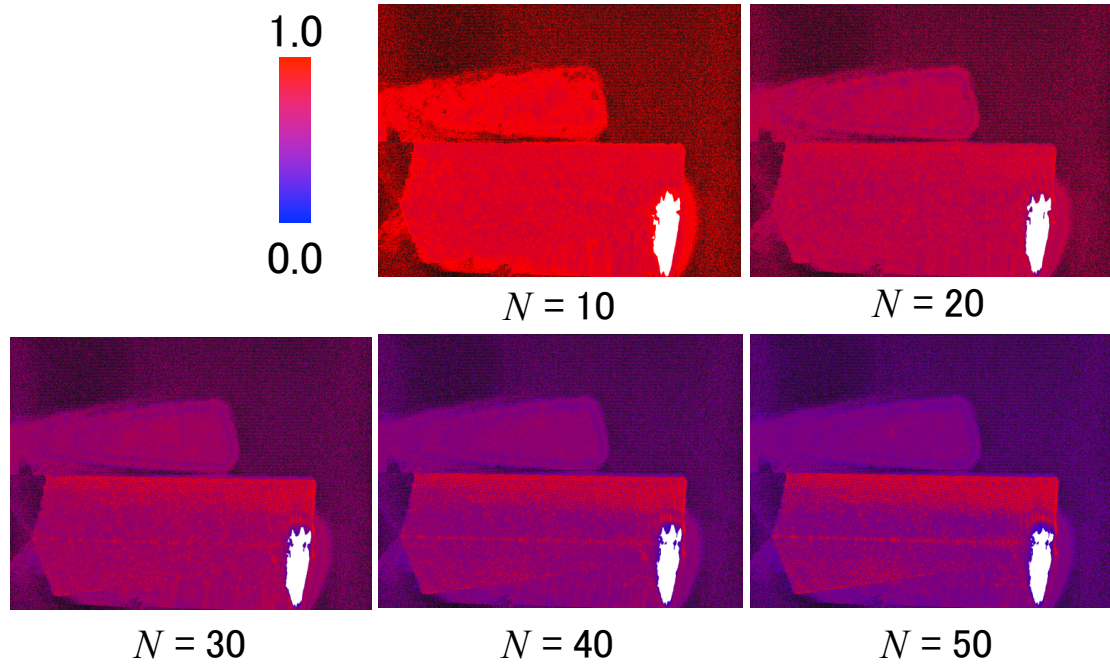


図3.9 相互反射物における投影枚数 N を変化させたときの相関値の変化. N を大きくすることで, 投影領域とそれ以外の領域との ZNCC の値の差が大きくなっていくことが分かる. ただし, 凹み部周辺では大きな相関値の差がない. また画像右下にある白い領域は, カメラ撮影時に対象の鏡面反射による白飛びで, 正しく符号化パターンを検出できなかった領域である.

3.3.2 幾何学的制約の導入

図 3.10は, 投影パターン数を削減するための幾何学的な制約について説明した図である. カメラで観測される光には直接反射光 (緑で表示) のほかに相互反射光 (赤で表示) が重畳することがある. まず, 直接反射光はエピポーラ線 (水色で表示) 上にしか対応点がないため, プロジェクタ画像上の画素のうち, エピポーラ線外の画素はすべて相関演算の対象から外することができる (これをエピポーラ制約と呼ぶ). さらに, 対象物体の存在し得る奥行き Z_d が限定されていれば, 点 (b) からの投影光は直接反射として弁別され得ない. 鏡面相互反射が問題となるような工業応用においては, 計測範囲全体に渡って同等の計測精度を得る必要があることから, 例えばカメラ画像の左端の画素に対してプロジェクタ画像の右端の画素が対応付けられるような条件での計測は行われることはない. よって必然的に, 正しい対応点となりうる点はエピポーラ線上でさらに限定される. これを奥行き制約と呼ぶ. これら 2 つの制約を用いることで, 例えばエピポーラ線長が 1024 画素のプロジェクタ画像の場合, 相関演算の対象となる画素数を数百画素程度に削減できることを予備実験

にて確認している.

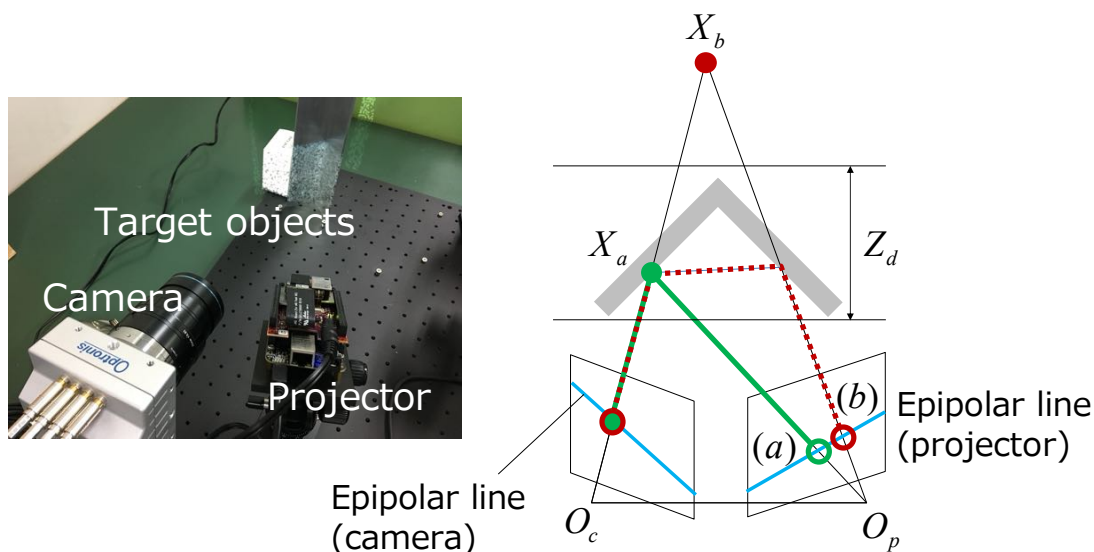


図3.10 左：実験環境，右：エピポーラ制約と奥行き制約. (Copyright(C)2021 IPSJ, [17] 図 4)

相関演算の対象となる画素数の削減は計測速度の向上と精度向上において重層的な効果を有する. まず, 相関演算の計算コストがそのまま削減される. さらに, 後で詳しく解析するように, 投影枚数が少ない場合, 符号長が短いため各投影画素のランダムパターンが偶然類似することがありえる. このような誤った対応付けの発生を抑制でき, これにより投影パターン数も大幅に削減できる. 投影パターン数の削減は計測に要する時間と相関演算に要する時間の双方の削減に寄与する.

ここで, 幾何制約を与えることで対応点候補を削減し, より対応点を発見しやすくなるかどうかを調べる予備実験を行った. 前述した図 3.7の実験系のようなシーンで, エピポーラ制約を与えることで, 対応点の取得率が上げられるかを調べた. 図 3.11のアルミ L 字アングルを計測する実験系を用いて, 図 3.12の復号化結果を得た. この結果から, エピポーラ制約を与えることで対応点の演算をする範囲を限定することができ, 正しい対応付けができる点が増加することを確認した. また演算領域が低減するため, 対応点探索の計算時間を低減できることも確認した. また後述する奥行き制約を幾何制約に加えると, さらに対応付けの精度が向上する.

3.3.3 投影パターン数と相関値の分布の関係

この節では, 投影パターン数と相関値の分布の関係を解析的に調べる. 続いて, 対応点の候補となる領域を限定することで, 誤対応の発生する確率がどの程度減少するのかについて論じる. その結果, 3.3.2節で述べた幾何学的制約を導入することで

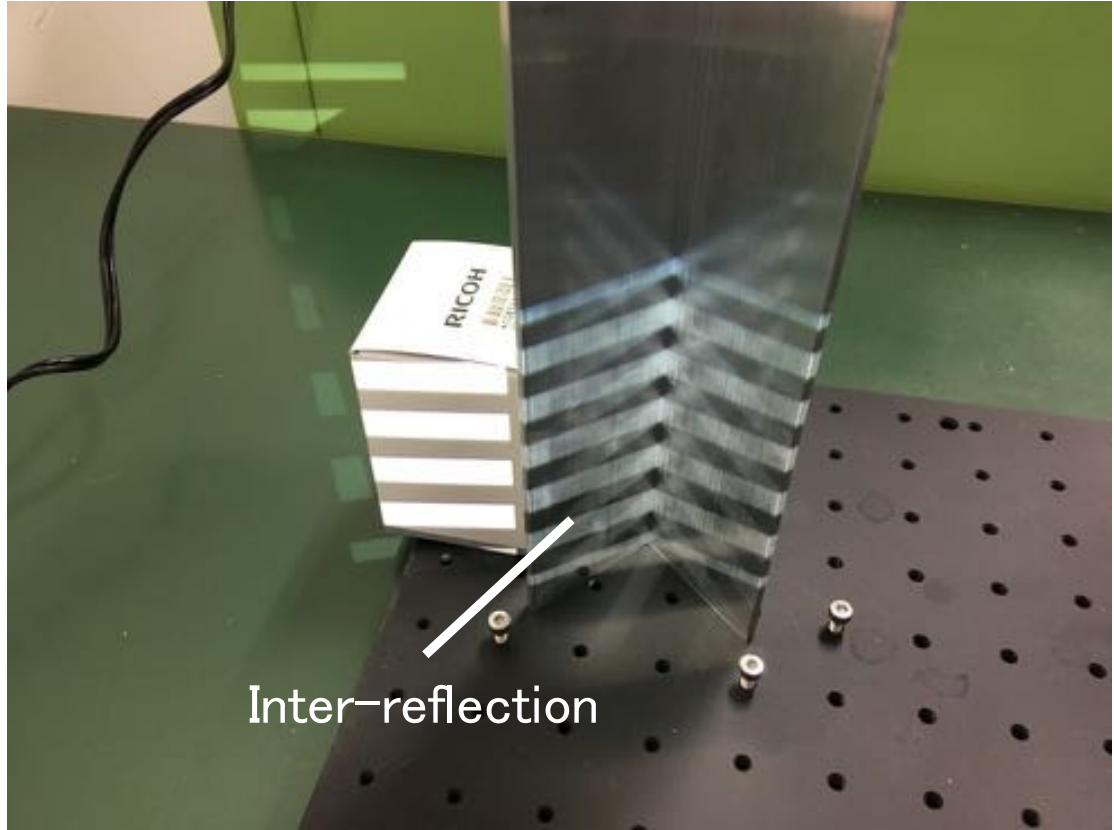


図3.11 エピボーラ制約の効果を調べる実験系．肉眼での見た様子 (投影画像は白黒の横縞ストライプ)．

投影パターン数を削減しても安定な対応付けが可能となることを示す．

非対応点，対応点の相関値の分布モデル

プロジェクタから各画素をランダムに点灯または消灯したパターンを N 枚，順次投影する．ここでは解析のため，点灯と消灯にそれぞれ 1 と -1 の符号を割り当てる． N 枚のうち 1 つのパターンに着目し，そのプロジェクタの各画素での投影信号を $S_p \in \{1, -1\}$ とする．カメラでの SN がある程度高く，輝度ノイズよりも投影信号 1 の観測値と -1 の観測値の差が十分に大きく分離可能な場合，例えば ZNCC の正規化処理などによって観測値は 1 か -1 の値をとるとみなせる．このためカメラでの各画素での観測信号を $S_o \in \{1, -1\}$ とすると，得られる相関値 c の確率分布は以下のように表せる．

$$P(c | \lambda) = \int P(c | \lambda, S_p, S_o) P(S_p, S_o | \lambda) dS_p dS_o \quad (3.1)$$

ここで， λ は非対応点 ($\lambda = F$) か対応点 ($\lambda = T$) かを表す記号である．また $P(c | \lambda, S_p, S_o)$ は投影信号と観測信号が与えられたときにとり得る相関値の確率， $P(S_p, S_o | \lambda)$ は投影信号と観測信号が生成される確率である．

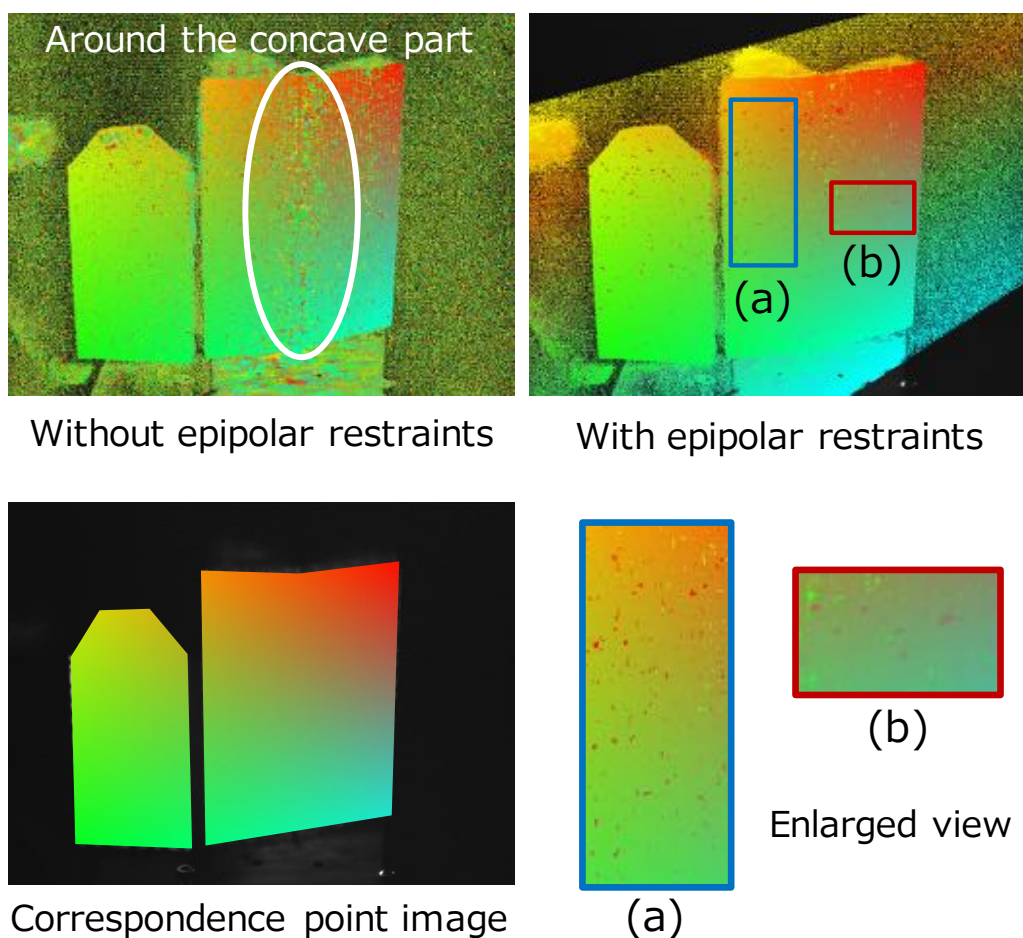


図3.12 エピポーラ制約の相互反射光対応点除去.

非対応点あるいは対応点で得る相関値の確率分布について考える．ランダムパターンの各画素は 0.5 の確率で点灯および消灯するものとする．この条件下では投影信号と観測信号の組み合わせは $(1,1)$, $(1,-1)$, $(-1,1)$, $(-1,-1)$ の 4 つがあり得る．このため相関値 c は 1 と -1 のいずれかを確率的に得る．この相関値を得る確率は式 (3.1) から以下のように算出できる．

$$\begin{aligned}
 P(c \mid \lambda) &= P(c \mid \lambda, S_p = 1, S_o = 1)P(S_p = 1, S_o = 1 \mid \lambda) \\
 &+ P(c \mid \lambda, S_p = 1, S_o = -1)P(S_p = 1, S_o = -1 \mid \lambda) \\
 &+ P(c \mid \lambda, S_p = -1, S_o = 1)P(S_p = -1, S_o = 1 \mid \lambda) \\
 &+ P(c \mid \lambda, S_p = -1, S_o = -1)P(S_p = -1, S_o = -1 \mid \lambda)
 \end{aligned}$$

対象シーンの光学現象により投影信号 1 が -1 と観測される確率を r_{+-} ，投影信号 -1 が 1 と観測される確率を r_{-+} とし、非対応点あるいは対応点で得る相関値の

平均 $\mathbb{E}_\lambda$ と分散 $\mathbb{V}_\lambda$ を求めると以下ようになる．

$$\begin{aligned}\mathbb{E}_F[c] &= (r_{+-} - r_{-+})^2 \\ \mathbb{V}_F[c] &= 1 - (r_{+-} - r_{-+})^4 \\ \mathbb{E}_T[c] &= 1 + 2(-r_{+-} + r_{+-}^2 - r_{-+} + r_{-+}^2) \\ \mathbb{V}_T[c] &= 1 - \{1 + 2(-r_{+-} + r_{+-}^2 - r_{-+} + r_{-+}^2)\}^2\end{aligned}$$

ここで N 枚のパターン画像すべての投影と観測について考えると，非対応点あるいは対応点の相関値の N 回の観測の平均の分布は， N が十分大きい時，中心極限定理により N の回数に漸近的に以下の分布に従う．

$$\mathcal{N}(c \mid \lambda) = \mathcal{N}(\mathbb{E}_\lambda[c], \frac{1}{N} \mathbb{V}_\lambda[c]) \quad (3.2)$$

これらの確率分布を用いて誤対応しない確率を見積もる．

数値解析による対応付け精度の見積もり

幾何学的制約により相関値を求める領域を小さくすることで，どの程度，対応付けの誤り率が減少するかについて考察する．相関の演算対象となる領域の画素数を M_c とする．式 (3.2) で求めた非対応点，対応点の 2 つの確率分布を用いて数値解析を行うことで， M_c の範囲の中で誤対応をしない確率を求める．

まず，さまざまな N ， r_{+-} ， r_{-+} で得た ROC 曲線（図 3.13 左）から各パラメタでの精度を見積もる．ここで前述したように投影信号と観測信号の組み合わせは 4 つであり，様々な光学現象が起こるシーンではそれら 4 つが等しい確率で起こり得る．このため r_{+-} ， r_{-+} は 0 ～ 0.25 の値をとり得る．投影信号 1 が -1 と観測されるケースは光の吸収が強い物体に投影した場合などが考えられる．これは本研究の想定シーンでは低い確率で起こると考えられるため $0 \leq r_{+-} \leq 0.01$ で検討する．また投影信号 -1 が 1 と観測されるケースは鏡面相互反射が起きた場合などが考えられる．これが本研究の想定シーンでは高い確率で起こり得ると考えられるため $0 \leq r_{-+} \leq 0.25$ で検討する．これらの範囲でパラメタ設定し得た ROC 曲線から， r_{-+} の値が高ければ高いほど主に鏡面相互反射の影響で精度が低下していくことが見積もられる．また N が大きければ大きいほど投影画像の各画素の信号の直交性が高くなるため対応付け精度も高くなると推測される．

次に相関を求める画素数 M_c と， N ， r_{+-} ， r_{-+} を変化させ，対応付けの精度を見積もる．誤対応しない確率を以下の 2 つの条件を満たす確率から求める．

1. 対応点の相関値が閾値 C_{th} を上回る
2. M_c 個の中に対応点が 1 つあるとし，対応点の相関値が $M_c - 1$ 個の非対応点の相関値をすべて上回る

この確率を求める式は以下ようになる．

$$q = \mathbb{E}_{c_t \sim \mathcal{N}(c|T), c_f^1, \dots, c_f^{M_c-1} \sim \mathcal{N}(c|F)} \left[\mathbb{1}_{\{c_t > C_{th}, c_f^1 \dots c_f^{M_c-1} < c_t\}} \right] \quad (3.3)$$

この式 (3.3) を解析的に求めることが難しいため以下の式を数値解析することで求める．

$$q = \mathbb{E}_{c_t \sim \mathcal{N}(c|T)} [P(c_t)]$$

$$P(c_t) = \begin{cases} \text{when } c_t \leq C_{th} & 0 \\ \text{otherwise} & [\int_{-\infty}^{c_t} c_f \sim \mathcal{N}(c|F) dc_f]^{M_c-1} \end{cases}$$

ここではモンテカルロ法にて解析をして得た M_c と q の関係を図 3.13右に示す．工業的に良好な運用が可能という観点から False positive rate が 0.005 となる C_{th} を求めて解析を行った．図 3.13右から投影パターン数 N が 30 と小さい場合でも相関を求める対象領域が小さい場合では誤対応の確率が十分小さくなると推測される．また N が 50 で r_{-+} が 0.2 以下の場合是对应付けの精度が極めて高いと見積もられる．誤対応はランダムに発生するため，belief propagation 法など周囲画素との関係を用いた正則化を施すことで容易に除去できる（本論文では提案手法そのものの特性を示すため，正則化によるノイズ除去は導入していない．）

実シーンに対する実験により相関値の分布を求めた例を図 3.14に示す．この図では各点の輝度は相関値を表しており，First は対象画素内で見つかった最大の相関値（正しい対応点である可能性が高い），Second は 2 番目の相関値である．ただし，この図は奥行き制約を課していない出力であることに注意する．図 3.14の相関値のヒストグラムを図 3.15に示す．図 3.14と図 3.15から，背景部を除き，First の相関値に違いが少ない一方で， $N = 20$ の場合の第 2 位の相関値は $N = 50$ の場合に対し大きくなっていることがわかる．

3.4 実験と評価^{*14}

提案法の有効性を確かめるため，鏡面相互反射が生じる環境を想定した実験を行った．実験系を図 3.10左に示す．図 3.16(a) 光沢があり等方性反射であるアルミ製 L 字アングルとランバート反射である紙箱を用いたシーン，弱い光沢があり表面化散乱も起こる樹脂模型（スタンフォードバニーの 3D プリンタ出力物）で，図 3.16(b) 不透明と図 3.16(c) 半透明の 2 種，アルミ製 L 字アングルより光沢が強く等方性反

^{*14} 本節の素材は，情報処理学会論文誌，Vol.62, No.8 [17] の 4 節に一部掲載されたものであり，本節の一部の図は，情報処理学会誌の許可を得て，[17] から再利用されています．

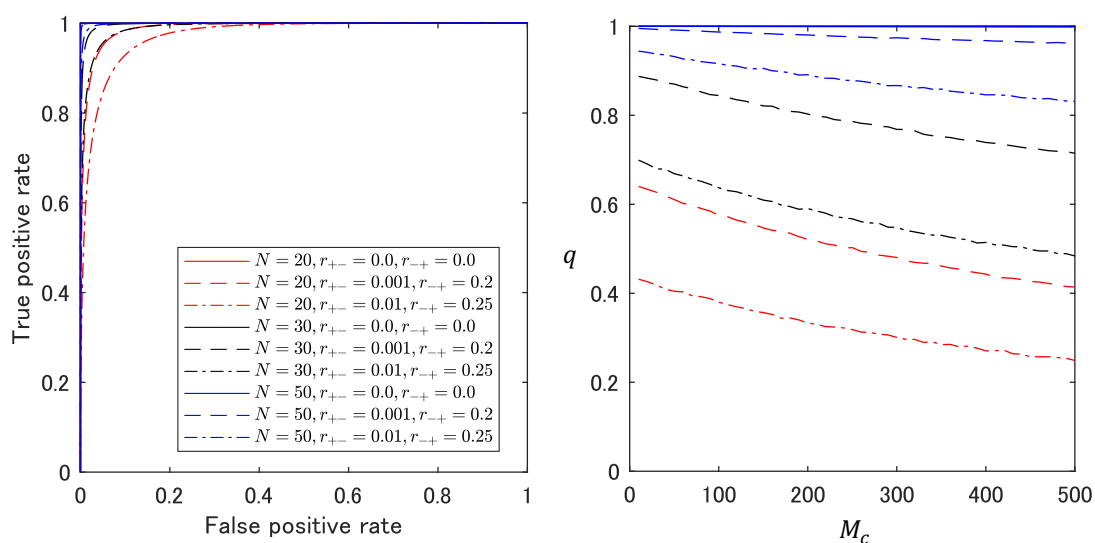


図3.13 左：非対応点と対応点の相関値の確率分布から求めた ROC 曲線, 右：相関を演算する領域の画素数 M_c において, その領域内に対応点が存在し, 対応点の相関値が閾値 C_{th} を上回り, かつ, M_c の中で最大となる確率 q . (Copyright(C)2021 IPSJ, [17] 図 5)

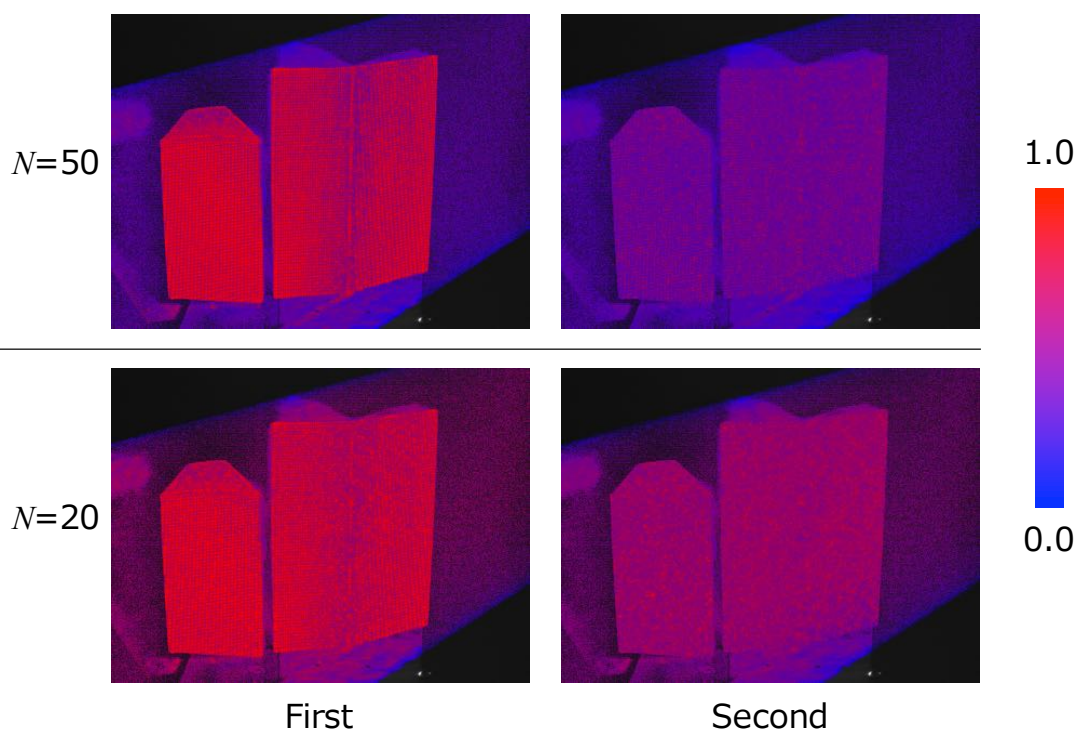


図3.14 実シーンに対する実験における相関値の分布. ZNCC で得た最大の相関値 (左) と 2 番目の相関値 (右) を表示している. (Copyright(C)2021 IPSJ, [17] 図 6)

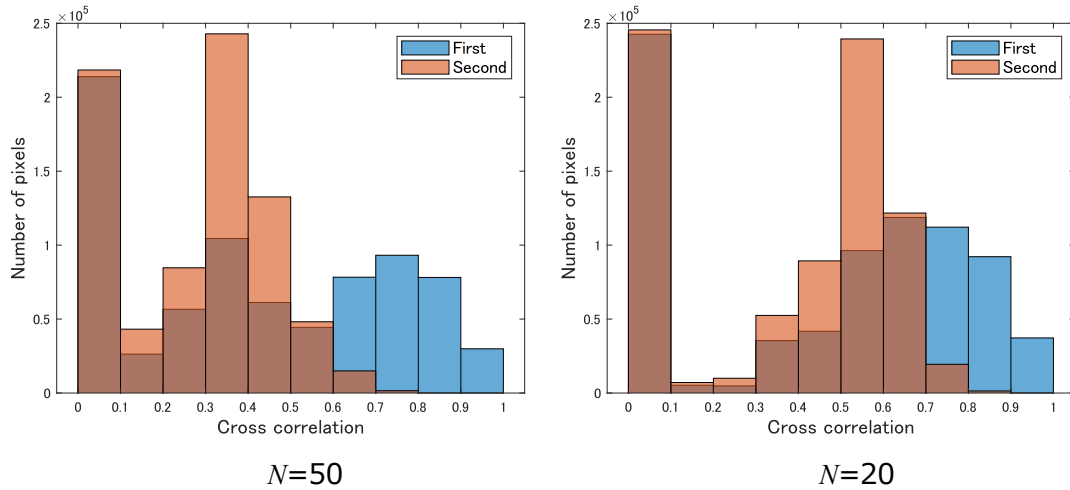


図3.15 図 3.14のシーンの相関値のヒストグラム. ZNCC で得た最大の相関値と 2 番目の相関値のそれぞれの画素数を示している. 0 に近い相関値が得られている領域は背景部である. (Copyright(C)2021 IPSJ, [17] 図 7)

射である金属製品（図 3.16(d) 金属製のボウルと図 3.16(e) 杯の 2 種）をそれぞれ設置したシーンを用いた.

カメラは Optronis 社のモノクロカメラ CP70-1HS-M-1900 を用い, 入力画像の解像度は 1024×768 画素とした. またプロジェクタは TEXAS INSTRUMENTS 社の DLP2000 を用い, プロジェクタ画像の解像度は 480×360 画素とした. これらについて Zhang のキャリブレーション法により参照物体に対する ProCam システムの内部・外部パラメタを求めた [16].

実験に用いた計測対象の多くでは間接反射成分の影響を非常に強く受けるため, SN 比を向上させるために, 提案手法の実装においてはプロジェクタ画像の 5×5 画素を 1 画素とみなし, 画素の中心を対応点とした. よって投影パターンに含まれるランダム符号の種類は $480 \times 360 / (5 \times 5) = 6912$ 種類であり, これらの生成にはメルセンヌツイスタを用いた [28].

提案手法を適用する工業的なタスクによって設定すべき False positive rate, C_{th} は変わる. 今回の分析では, ロボットによるワークの把持タスクにおいて妥当な対象物の姿勢推定精度が得られると考えられることを基準に, False positive rate を 0.005 とし, 相関値の閾値を求めた. 実験の対象物は拡散反射性があるため $r_{-+} = 0.25$ で想定する対象物よりも鏡面相互反射の発生率は低く, r_{-+} が 0.2 程度以下が妥当であると考えられる. ゆえに $r_{-+} = 0.25$ で求めた相関値の閾値 C_{th} よりも小さな 0.4 を閾値とし, この値以上の相関値が得られた画素を対応点とした.

比較のため, グレイコード法 (GC) と位相シフト法 (PS) を組み合わせた手法を実装した. 前述のように間接反射の影響が大きいため, GC のみで線幅を細くする

と計測が不安定となった．そこで GC では最小線幅を 30 画素とし，カメラ画像上の各画素に対しプロジェクタ画像上の対応点を 15 画素の分解能で得る．その上で正弦波の 1 周期を 30 画素とした PS を用いることで，より高分解能な形状が得られる設計とした．投影枚数は GC で 10 枚（反転二値化を用いるため，ネガ・ポジ画像それぞれ 5 枚），PS では投影枚数（ステップ数）として一般的に用いられている 4 枚とし，合計 14 枚である．

3.4.1 奥行き制約の効果に関する実験

まずこの実験では，奥行き制約の効果を確認する．図 3.17 はアルミ製 L 字アングル（図 3.16(a) の右側の対象物）に対し相関演算により得られた対応付け結果で，左はエピポーラ制約のみを用いた結果，右はそれに加えて奥行き制約を用いた結果である．本来はなめらかなグラデーション状に表示されるべき対応付け結果に現れるノイズの量から，エピポーラ制約に加え奥行き制約を用いることで，鏡面相互反射によって起こったとみられる対応付けの誤りが低減できることを確認した．

3.4.2 従来手法との比較

次に，様々な物体に対する対応付け結果を図 3.16 に示す．ここでは結果を比べやすいよう，カメラ画像の各画素に対応付けられたプロジェクタ画像に，15 画素幅の白黒の縞を配置した結果を示している．これはちょうど GC のみで算出した視差の分解能に対応している．

(a) の赤四角部はアルミ製 L 字アングルの凹みがあり，この領域で鏡面相互反射が生じる．GC ではその領域で大きな対応付けの誤りが見られるが，提案法では比較的良好な復号化がなされている．次に (b)，(c) の弱い光沢のある樹脂製の模型では，いずれも赤四角部に凹みがあり，鏡面相互反射が生じている．特に拡散反射率が低く，表面下散乱が生じる (c) では大きな乱れが見られるが，これに比べ提案法は良好である．

また，(d)，(e) はいずれも強い光沢を有する，金属製で凹部を持つ物体であり，どちらも赤四角部で強い鏡面相互反射が発生している．このような場合でも，GC に比べ提案法は対応付けを安定に行えることがわかる．

3.4.3 3 次元計測精度の定量評価

アルミ製 L 字アングルと紙箱が配置されたシーン（図 3.16(a)）について求めた奥行き値を図 3.18 に示す．ここでは，求められた奥行き値が実際の物体を含む範囲から逸脱した場合には復元が失敗したとみなし除外している．カメラから被写体まで

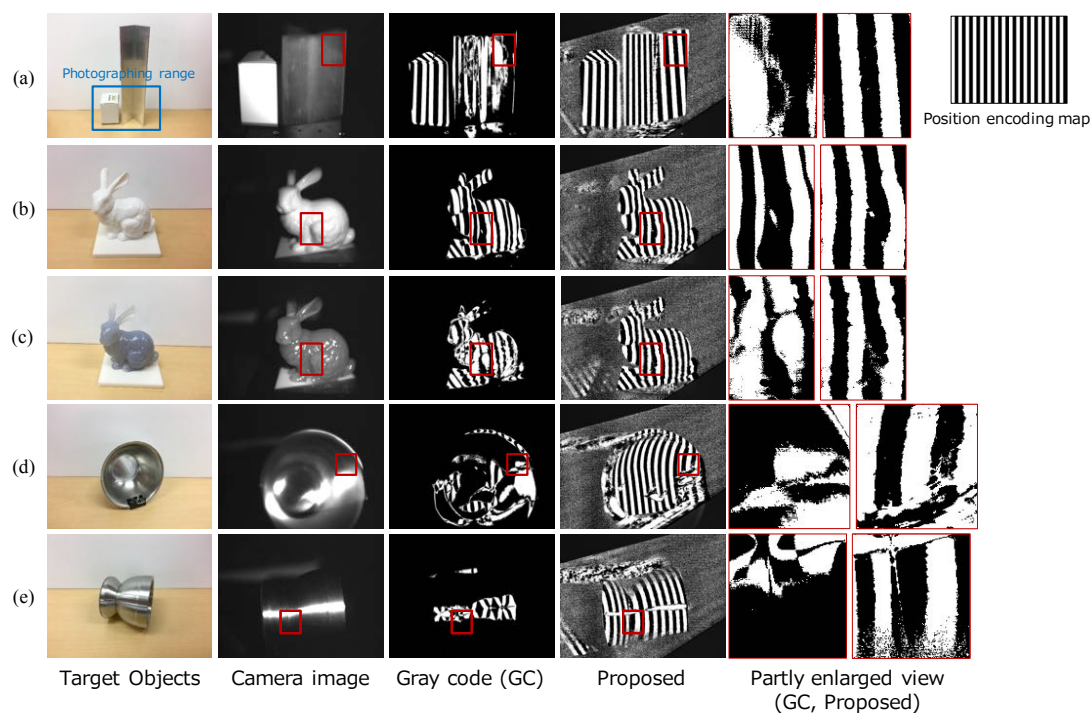


図3.16 GC, および, $N = 50$ の場合の提案法による対応付けの結果. 右上に示す座標符号化マップのように, 得られた視差値を周期的に白・黒に変換して図示している. (Copyright(C)2021 IPSJ, [17] 図 8)

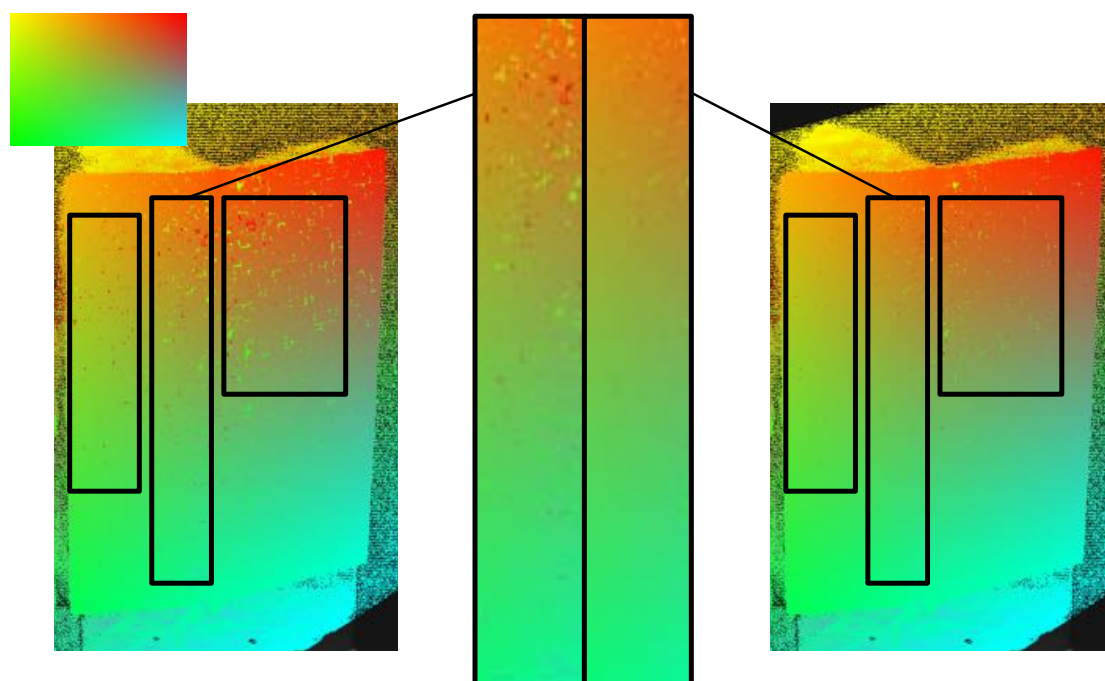


図3.17 $N = 30$ の時の提案法による復号化の結果. 左: エピポーラ制約を用いた結果, 中: 結果の一部拡大図, 右: エピポーラ制約と奥行き制約を用いた結果. 座標符号化マップは左上の図を参照. 特に黒枠部にて奥行き制約の有無による差異が見られる. (Copyright(C)2021 IPSJ, [17] 図 9)

の距離はおおよそ 300-400 mm で実験した. (Copyright(C)2021 IPSJ, [17] 図 10)

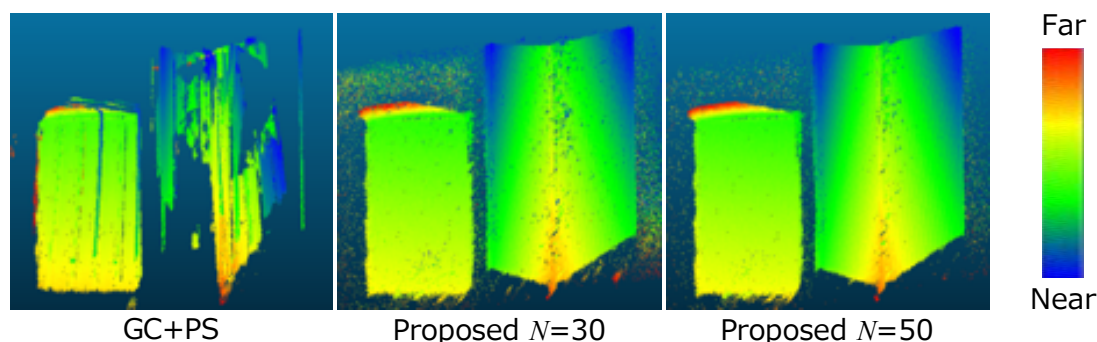


図3.18 実験シーン (図 3.10左, 図 3.16(a)) の形状復元結果. 左: 従来法, 中: 提案法 ($N = 30$), 右: 正解としたデータ (提案法 $N = 50$). (Copyright(C)2021 IPSJ, [17] 図 10)

従来法では GC と PS を併用して視差を求めているが, ランバート反射とみなすことができる紙箱部分は概ね良好に形状復元がなされている. しかし, アルミ製 L 字アングルはその凹部で鏡面相互反射が起こり, この影響で形状復元結果が大きく欠け, また, 復元された部分についても誤った奥行き値が多く含まれている. 一方で提案法 ($N = 30$) の復元結果では, いずれの物体でも良好に形状が復元されている.

このシーンでは形状の真値が与えられていない. このため, 同シーンに対し前述した妥当な r_{-+} の設定で得た対応付けの観測モデルから, 十分に良好な復元精度が得られると見積もられた投影枚数 ($N = 50$) で計測した結果を正解とした. これに対する画像全体での復元成功率, および, 奥行き値の平均絶対誤差を示す (図 3.19). ここで, 復元成功率 (図 3.19左) は前述の処理で除外されなかった画素の割合である. 結果から, 紙箱については従来法と提案法どちらの手法でも同程度の復元成功率であることがわかる. 一方アルミ製 L 字アングルについては, 従来法では復元成功率が約 39% であったのに対し, 提案法では $N = 30$ のときに復元成功率が約 95% であり, さらに投影パターン数を減じた $N = 20$ の場合でも大きく悪化しないことがわかる.

奥行き値の誤差についても同様に, 従来法・提案法ともに紙箱の部分はアルミ製 L 字アングルに比べ誤差が小さいことがわかる. しかし, アルミ製 L 字アングルの部分については従来法では誤差が約 0.87 mm であったのに対し提案法は $N = 30$ の場合に誤差が約 0.15 mm で, $N = 20$ の場合でも従来法よりも高精度であることがわかる.

ただし提案法には直接反射の対応点の弁別が難しいケースがある. 例えばアルミ製 L 字アングル (図 3.16(a) 右側) の凹みの中心付近では直接反射の対応点, 鏡面相互反射の対応点の双方がエピポーラ制約と奥行き制約の 2 つの幾何制約を満たし

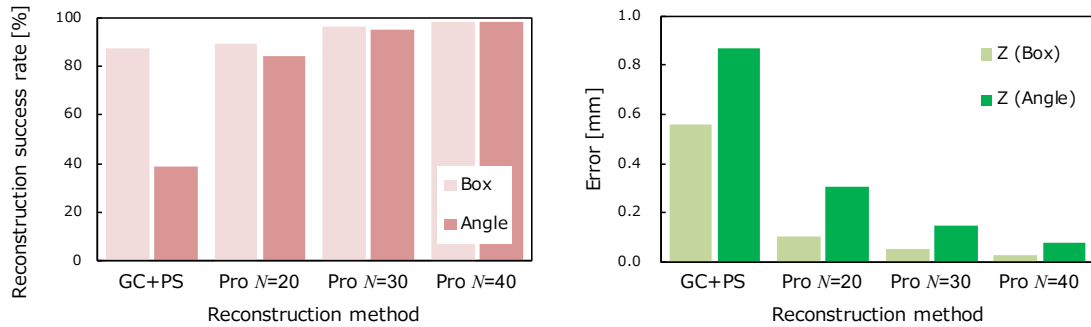


図3.19 左：復元の成功率，右：計測点の奥行き方向（Z 軸）の平均絶対誤差。
GC+PS：従来法，Pro：提案法．（Copyright(C)2021 IPSJ, [17] 図 11）

てしまうことがありうる．このような直接反射光と相互反射光の対応点が近い場合，誤った計測点が得られる可能性が相対的に上昇する．ただしこのような場合でも，誤って復元される点は正しい値から大きく離れてしまうことはなく，位置合わせ等の処理に対する影響は小さい．

3.4.4 計測時間

計測に要する時間は投影枚数に比例する．前述の評価実験では，従来法として GC と PS の併用により 14 枚の画像の投影により形状を計測した．GC 単独で提案法と同様の奥行き分解能を得るためには， $2^6 < 480/5 \leq 2^7$ より 14 枚（ネガポジ画像それぞれ 7 枚）で計測が可能である．一方で，提案法は実験により 20 ～ 30 枚程度の投影で良好な対応付けと復元ができることがわかった．よって，提案法は GC に比べ多くの投影枚数が必要ではあるものの，幾何学的制約の導入により過大な投影枚数の増加なしに相互反射の影響を抑制しつつ安定な形状復元が可能であることを確認した．また，投影枚数は対象物体の計測の困難さに応じて調整でき，枚数を増やすことで相互反射成分に比べ直接反射成分の SN 比を十分に高くできる反射点においては，より条件の悪い計測にも対応することができるという利点も有する．

3.5 まとめ^{*15}

ProCam を使い，鏡面相互反射が生じる物体の 3 次元形状を高速に計測する手法を提案した．提案法では 2 次元ランダムパターンを用いることで，鏡面相互反射現象により投影光が重畳しても，投影光の信号源を分離することができる．本論文ではさらに，エピポーラ制約と奥行き制約の 2 つの幾何学的制約を導入することで，

^{*15} 本節の素材は，情報処理学会論文誌，Vol.62, No.8 [17] の 5 節に一部掲載されたものであり，本節の一部の図は，情報処理学会誌の許可を得て，[17] から再利用されています．

相関値を求める領域を限定し，投影パターン数を削減し高速化するとともに計測を安定化する手法を提案した．さらに提案法の対応付けの観測モデルを作成し，提案手法の誤対応の発生確率を数値解析した．実シーンでの計測実験では，広く利用されているグレイコード法に近い 30 枚の投影・撮影で，強い鏡面相互反射が生じる物体を良好に形状復元することが可能であることを確認した．

第 4 章

提案手法 3 :

動き制約と固定 2 次元ランダムパターン投影を用いた移動物体の密なアクティブ計測法

4.1 はじめに

近年，生産現場では全自動化に向けて，ベルトコンベア上の動く物体を三次元計測しロボットアームで自動把持するための研究開発が進められている．動く物体をロボットアームで自動把持するためには，高密度かつ高精度，光学現象に高ロバストな計測が可能な 3 次元計測システムが欠かせない．中でもプロジェクタとカメラを組み合わせたシステムに基づくアクティブ計測法は，視野全体を同時に計測できるため計測時間が短い．また，レーザ光の走査や特殊な光学系を用いる計測装置に比べ小数の汎用部品で構成されているためにコストを下げやすく，量産性にも優れている．グレイコード法 [2] や位相シフト法 [8]，間接反射に頑健な手法 [11, 19] などが，プロジェクタとカメラを用いる代表的な手法として提案されている．しかし，そのような生産現場においては対象が動いていることを考慮した計測が求められるため，これらの手法は複数枚の投影と撮影を要してしまうために適さない．

フルオートメーションを目的とする生産現場では対象の動きを制御するため自然と対象物の動きを既知とできる．このため既知の動きをする動く物体計測に関する研究は，1980 年代からスリット光投影法 [29] を代表手法として発展してきた．この手法は対象物が投影されたスリットラインを通過する際の撮影画像から形状復元を行う手法であり，対象物が動く物体である場合に有効である．しかし低 frame rate のローコストカメラを用いる場合，対象物が撮影速度よりも高速に移動することが

ありえ、スリット光投影法が1次元のパターンを投影しているために計測の密度が低下してしまうことがある課題がある。また1ラインの情報しかなく相互反射などの光学現象の影響を受けてノイズが発生しやすいことや復元精度が低下するケースがありうる。

本研究ではスリット光投影法のような密な計測が難しい1次元パターンを用いる手法ではなく、密な計測が可能と予想される2次元パターンを用いる手法の可能性検討をする。高速パターン投影・撮影手法 [22,30–34] やワンショットパターン計測法 [10,35,36], また動き方向推定しながら計測をするというチャレンジングな課題に取り組んでいる手法 [37–41] などが提案されている。これらの手法は対象の動きが未知である場合に効果的だが、装置コストが高くなる, あるいは, 光学現象の影響を抑制した密な計測が難しいなどの課題がある。よって, 低 frame rate なローコストデバイスで高密度・高精度な計測ができる2次元パターンを用いる計測法であり, 対象物が既知の移動していてもストップレスで計測が可能なる方法を実現する必要があると考えられる。さらに生産現場においては計測対象物が金属物などのランバート反射な物体からかけ離れた物体も多いため, 相互反射や鏡面反射などの光学現象に高ロバストな計測法であることも要求される。

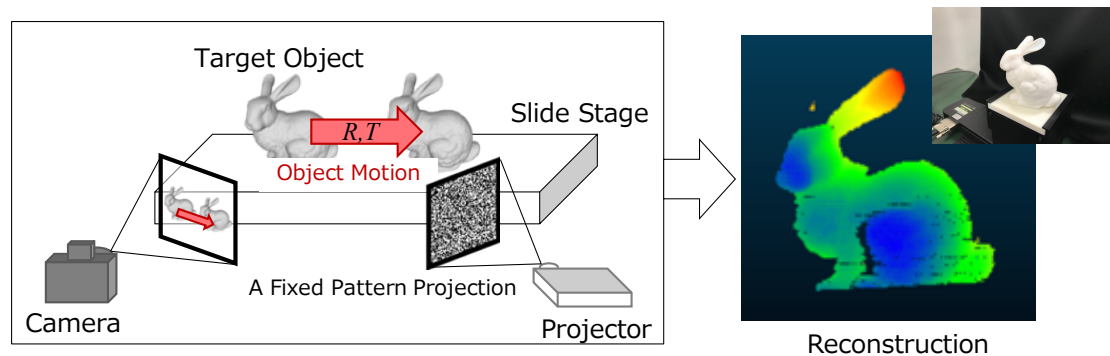


図4.1 提案計測システムの処理イメージ.

本研究では前述したフルオートメーションの生産現場という条件下においては容易に入手できる対象物の動き情報を幾何制約条件とし, 1つの固定2次元ランダムパターン投影することで, 動く対象物の高い密度での形状復元できる新しいアプローチを提案する (図 4.1). 本手法はエピポラ制約を用いてカメラ視線方向の空間に複数の仮説三次元点を生成する。さらにこれらの点が既知の移動をするという制約下で, 実物体表面に近いデプスの点を仮説点の中から選出することで形状復元する。このとき既知の動きを用いることで1枚のパターンで座標符号化ができる。このパターンに固定2Dランダムパターンを用いることでユニークな座標符号化を行うことができ, 光学現象にロバスト, かつ, 密な対応付けができる。実験から提案法は低コストな30FPS程度のカメラで撮影すること想定したとき, 180mm/secで移

動する物体を既存法であるスリット光投影法より 54% 程度高い密度で計測できる．
我々の貢献を以下にまとめる．

- エピポーラ制約に加えて物体の動きを幾何制約として用いることで良好な対応付けと復元ができる
- 固定 2 次元ランダムパターンを用いることで，固定 1 次元パターン計測法よりも少ない撮影枚数で高密度に形状復元でき，かつ，光学現象にロバストな計測ができる

4.2 関連研究

対象物が未知の動きをする場合にも計測を実現する手法や，スリット光投影法をベースとした既知の動きをする物体計測に関する研究など，これまで多くの研究がなされている．1D パターンの手法を参考に，スリットパターン投影法の密な計測が可能と思われる 2 次元パターンへの拡張について検討する．

装置工夫により高速投影&撮影により計測を行う方法

Konolige らはステレオカメラのウィンドウマッチングを補助する形でプロジェクションするパターンを提案している [30]．デブルーイン系列をベースにユニークなテクスチャが視差方向に配備されるよう工夫をしている．また同様のアプローチで Lohry らは，ランダムパターンと位相波パターンを合成したバイナリディザパターンを提案している [31]．ただしこれらの研究は相互反射など光学現象がおきたときに誤った推定が多くなると思われる．また，最小構成システムに加えてカメラが 1 個余分に必要で，コスト，装置配置制約がある．

Matsuda らや Leroux らは，様々な活用がなされつつあるイベントカメラを活用した手法を提案している [32,33]．イベントカメラの特性により高速な計測，外乱光に頑健な計測，低消費電力で処理を可能としている [34]．しかし，イベントカメラは特殊光学系を持つためコスト高となり，感度の問題でノイジーになりやすいため形状復元もノイジーになる傾向がある．

O' Toole らは CMOS 型画像センサのローリングシャッタ現象とレーザプロジェクタの走査を同期させた装置構成により，エピポーラ面外からの間接成分を除去しつつ高速な計測が可能な手法 EpiScan を提案している [22]．この手法ではフレームレートを低下させることがないという利点を持つ．一方で精密な位置合わせが必要で，使用できる機器の種類が限定される問題がある．

ワンショットで対応点の観測が可能なパターンを投影する手法

Kawasaki らはワンショット型計測が可能なグリッドパターンを提案しリアルタイムに計測を実現している [35]. Sagawa らはこの研究を発展させデブルーイン系列を用いて検出の頑健性を向上させた手法を提案している [10]. また Nagamatsu らは投影パターンにドットとラインを仕込み、それらを CNN で検出することでロボバストに計測できる手法を提案している [36]. これらの手法はワンショットで形状復元ができる利点がある. しかしグリッドパターンを用いているので, 計測データは全画素を符号化する手法に比べ計測密度が下がってしまう. また精度はワンショット方式であるために, 時間符号化方式よりも低くなる傾向がある.

動く物体を高速計測しつつ動き方向推定を行う手法

Zhang らは, ステレオカメラとプロジェクタという構成で対象になんらかの動きが発生することを考慮して時空間特徴を利用することで精度の高い復元手法を提案している [37]. 同様の構成で, Lim は視差方向にユニークなテクスチャを投影し動くシーンに対応する手法を提案している [38]. この研究では動的なシーンではランダムスリットパターンが有効であることにも触れられている. また Li らは, 速度に合わせて復号するパターンを変える複合パターンを用いて多状態の速度 (遅い, 速い) で動く物体のデプスを推定できる手法を提案している [42]. これらの研究は対象物の動きに対応できるものの, ステレオカメラを用いる必要があるため最小構成に対しコスト高になり, また 2 つ目のカメラの配置位置の制約が強い. また相互反射など光学現象がおきたときに対応付けが破綻する恐れがある.

Zhang らは, 単一カメラとプロジェクタ構成で動きありなしを自動判別でき, ランダムパターンと位相波パターンを混ぜてユニークなテクスチャが視差方向に配備されるパターンを提案している [39]. ただし相互反射など光学現象がおきたときに誤った復元が多くなると思われ, 動きありモードでの復元精度が低い. Taguchi らは動く物体でも空間分解能を下げずに形状復元することができるパターンと窓が自動生成されることにより, 動く物体, 静止物体, どちらも復号化することができる方式を提案している [40]. これらの手法は間接反射現象が起こるシーンでは対応付けが破綻する恐れがあり, 解像力が落ちてしまう課題がある. Shiba らは, 市販の低速 frame rate カメラと低速 frame rate プロジェクタを用いて, 時間方向に解像度を上げて形状復元, デプスと動きの両方を推定することができる手法を提案している [41]. ただし復元密度が低く, ブラーにより精度が低く, 間接反射に対するロボバスト性も低い.

スリット光投影法をベースとした計測法

Wang らはシリンドリカルレンズを搭載したラインスキャンカメラにて複数のエピポーラ平面で観測したとみなす計測系を実現し, 安価な高速計測系を実現してい

る [43]. 同様に光学系工夫のアプローチで, Sheinin らは, 特殊な拡散をする光学系とラインセンサの構成により非常に高速に撮影を行いつつ, 対象の光源の対応点座標を計算することができるシステムを提案している [44]. これらの研究は高密度な計測が可能なものの, 特殊光学系が必要になりコスト高になることや, カメラと光源の配置に制約がある. また低 frame rate カメラでの動く物体の計測も困難である.

スリット光投影法の 2 次元パターンへの拡張

前述の研究やスリット光投影法 [29] で用いられる固定 1 次元パターンを用いる手法では動く物体を, ローコストな低 frame rate カメラで, 密, かつ, 精度の高い, 光学現象にロバストな計測が難しいと思われる. このため, スリット光投影法を固定 2 次元パターンを用いる手法に拡張できないかを考える.

まず単純にスリットの本数を等間隔に増やした 2 次元パターンを用いる方法をとった場合を考える. 複数のスリットが同じパターンとして抽出されてしまうために, 誤った対応付けが多数発生してしまう. また相互反射等の光学現象がおきた時にはさらに多く誤った対応点が発生することが予想される. このため単純なスリットの拡張では良好な密度, 精度, 光学現象へのロバスト性で計測することは困難であることが予想される. ゆえにより座標符号が一意的に行えるパターンを考える必要がある.

次に, ユニークな 2 次元の座標符号が可能な固定パターンがないかを検討する. 代表的な符号化方式として, まずランダムパターンが考えられる. ランダムパターンは時空間方向にユニークな符号化をする方式として極めてリーズナブルであることが広く知られている. 先行研究においてもランダムスリットパターン, ランダムドットパターンが活用されており, 精度の高い, 光学現象にロバストな計測ができる知見がある [23, 25, 38]. しかし動く物体にランダムパターンを投影したシーンを撮影したとしても, エピポーラ制約のみでは十分な制約条件が得られないため, 対応点の探索は困難になる. ゆえに他の制約条件が必要となる.

そこで本研究ではユニークな座標符号が可能と思われる固定 2 次元ランダムパターンと, フルオートメーションシナリオのような対象の動きが既知である事前知識を合わせこむことで良い性能が得られるのでは, と考えた. 既知の動き情報を幾何制約として用いることで密度や精度, 光学現象へのロバスト性の向上が図れるのではないかという仮説を立てた.

4.3 提案法

2 次元ランダムパターンの持つ豊富な符号化情報と, エピポーラ制約に加え動き制約を用いることで, 良好な密度, 精度, 光学現象へのロバスト性で形状復元が可能

な手法を提案する．動き制約はエピソード制約で制約される空間とは異なる幾何制約を復元対象空間に課することができるため，対象物の 3D 位置推定が安定的に行えるようになり，高い密度での計測が可能となる．また 2 次元ランダムパターンを用いることでユニークな対応付けができ，光学現象が起きてもロバスト性が向上できる．また提案法は動き制約をもちいることで 2 次元パターンを複数枚用いても，1 枚の固定パターンでも，符号化ができるため，さまざまな計測システム設計ができる特徴をもつ．本稿では，スリット光投影法のようにコストが低くシステムが作成可能，かつ，既知の動きを用いることで座標符号が可能，であるため 1 つの固定パターンを用いることを想定し，以降説明していく．

4.3.1 システム概要

アルゴリズム概要を図 4.2 に示し，これにそって説明をしていく．はじめにカメラ画素 1 つに着目したとき，その視線方向に物体上の 1 点が含まれると考え，いくつかのデプスの仮説点を設置する（図 4.2 Step 1）．これらの仮説点の中から物体からの反射点と思われる点を選び出すことで推定形状を得る．

次に時刻 t のときの仮説点を時刻 $t+1$ の位置に移動させ，移動後の点をカメラ画像，プロジェクタ画像に透視投影する（図 4.2 Step 2,3）．ここで本稿で示すシステムは，物体がスライドステージに載せられて平行移動する（図 4.3）．その物体へプロジェクタでパターンを投影しカメラで撮影する（図 4.4）．このとき，スライドステージとカメラ，プロジェクタの関係は較正済みである．このため，仮説点をカメラ画像，プロジェクタ画像へ透視投影する処理は以下のように表せる．

$$q_c(d, t) = \mathbf{A}_c[\mathbf{R}_c | T_c]p(d, t) \quad (4.1)$$

$$q_p(d, t) = \mathbf{A}_p[\mathbf{R}_p | T_p]p(d, t) \quad (4.2)$$

ここで $p(d, t) = (X(d, t), Y(d, t), Z(d, t), 1) \in \mathbb{R}^4$ の第 1, 2, 3 成分は時刻 t のときの奥行き d の仮説点の 3 次元位置， $q_\rho(d, t) = (x_\rho(d, t), y_\rho(d, t), 1) \in \mathbb{R}^3$ の第 1, 2 成分は時刻 t のときの奥行き d の仮説点を透視投影した画素位置， $\mathbf{A}_\rho \in \mathbb{R}^{3 \times 3}$ は内部パラメタ行列， $\mathbf{R}_\rho \in SO(3)$ は回転行列， $T_\rho \in \mathbb{R}^3$ は並進ベクトルである（ $[\mathbf{R}_\rho | T_\rho] \in \mathbb{R}^{3 \times 4}$ ）． ρ は c であればカメラ，p であればプロジェクタ，s であれば対象物を表す記号である．

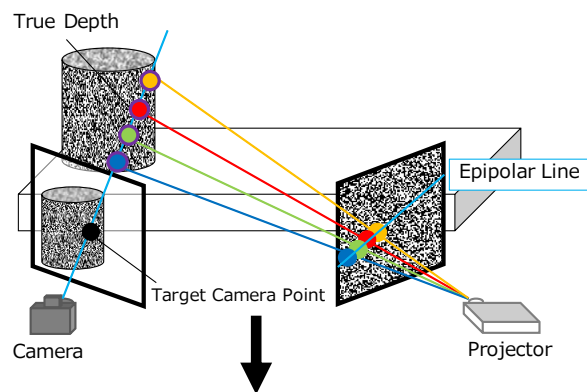
また仮説点を移動する操作は以下のように表せる．

$$p(d, t+1) = [\mathbf{R}_s^{t \rightarrow t+1} | T_s^{t \rightarrow t+1}]p(d, t) \quad (4.3)$$

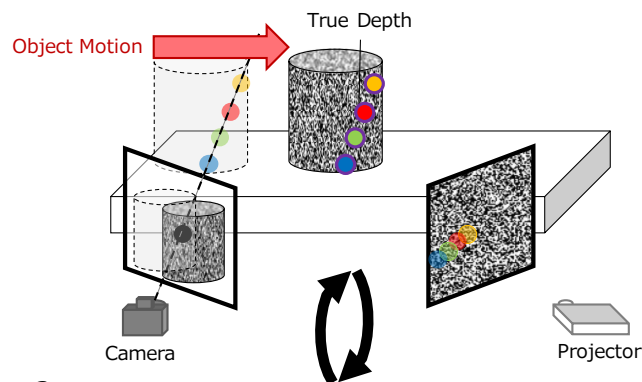
式 (4.3) を用いて奥行き d の仮説点 $p(d, t)$ を移動させ，式 (4.1)，式 (4.2) を用いてカメラ画像，プロジェクタ画像へ透視投影を行い，各画像で投影された画素の値を得ることができる（図 4.2 Step 2,3）．

Step 1

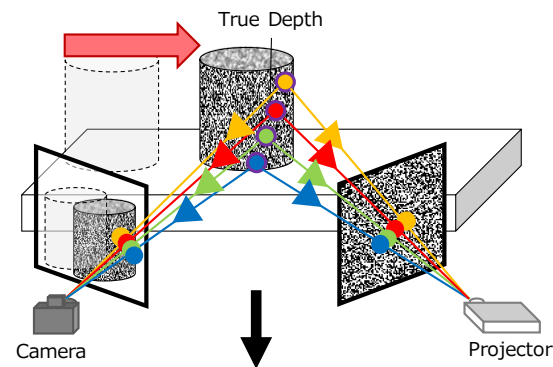
Hypothesis 3D Points 



Step 2



Step 3



Step 4

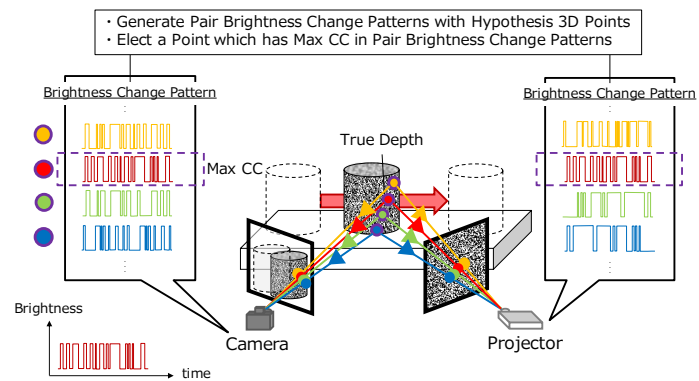


図4.2 アルゴリズム概要.

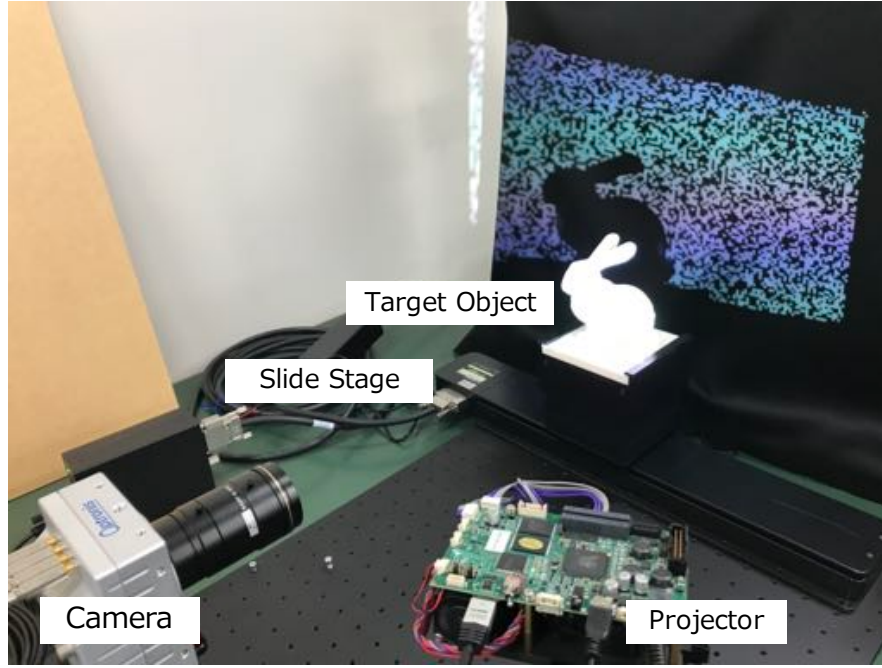


図4.3 実験システム.

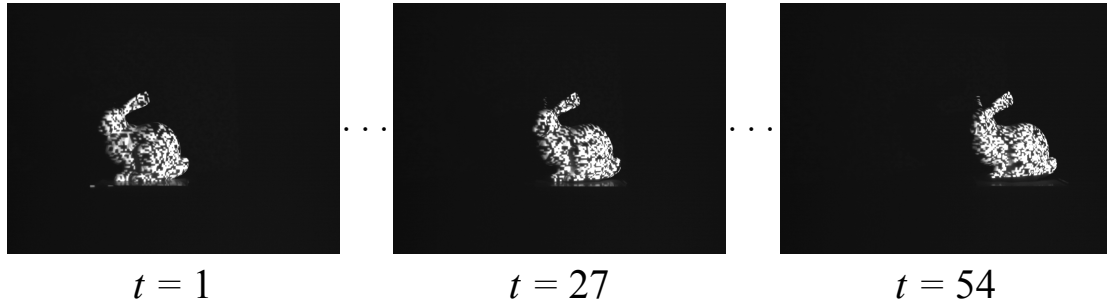


図4.4 カメラでの撮影画像.

この時系列に得た画素値がカメラ側，プロジェクタ側で輝度変化パターン f_c, f_p として得られ，式では以下のように表せる（図 4.2 Step 4）.

$$f_c(d, \mathcal{N}) = \{I_c(q_c(d, 1)) \dots I_c(q_c(d, N))\} \quad (4.4)$$

$$f_p(d, \mathcal{N}) = \{I_p(q_p(d, 1)) \dots I_p(q_p(d, N))\} \quad (4.5)$$

$$\mathcal{N} = \{1, 2, \dots, N\}$$

ここで $I_\rho(x, y)$ は画像上の座標 (x, y) にて得られる輝度値を， N は撮影枚数を表す．理想的には各時刻で $I_c(q_c(d, t)) = I_p(q_p(d, t))$ となる奥行き d の点を得られるが，実際は相互反射や光減衰の影響でこのように輝度値を得ることは難しい．そのため相関値を用いて復元すべき点を推定する．仮説した奥行き d の点で得られる2つの輝度変化パターン（式 (4.4)，式 (4.5)）のパターン同士の相関を得て，最大値を得た奥行き仮説点 $p(d, t)$ を復元点として採用する（図 4.2 Step 4）．この処理を式

で表すと以下のようになる.

$$p^*(d, t) = \arg \max_d CC(f_c(d, \mathcal{N}), f_p(d, \mathcal{N})) \quad (4.6)$$

$$d = d_{low}, \dots, d_{up} \quad \mathcal{N} = \{1, 2, \dots, N\}$$

ここで $p^*(d, t)$ は仮説点の中から選ばれた推定点であり, 時刻 t のときの撮影位置をベースに求められた復元点とする. また $CC()$ は相関演算を表し, d_{low}, d_{up} は仮説する奥行き値の下限と上限を表す. 全てのカメラ画素にて, 式 (4.6) を用いて仮説点の中から復元点を選出する処理を行い, 対象物の全体の形状復元を行う.

4.3.2 動き制約

提案法ではエピポーラ制約に加え, 新たな動き制約という幾何制約を与える. 特定のアプリケーションにおいては対象物の動きである Object Motion が得られるため, これを幾何制約として用いることで頑健に対象物を形状復元できる. 前述したように, 広く用いられているエピポーラ制約のみで動く物体を復元するアルゴリズムを考えると, デプスの変化に伴ってエピポーラ線上を移動する点が発生するが, これでは制約条件が不十分であり一意に復元点を決定することが難しい. そこで本研究では動き制約とエピポーラ制約は異なる空間制約を課す幾何制約条件として働くため, これら 2 つを満たす点が復元点となることを考えた. 図 4.5 はカメラ画素 1 つに対して発生させた全仮説点を移動させながらカメラ画像, プロジェクタ画像へ投影した図である. 図 4.5 の右のプロジェクタ画像の図に注目する. 左端の緑の斜線が, 復元領域として設定した奥行き範囲内で, 一つのカメラ画素の視線方向に設置した初期の全仮説点を表す. また赤い矢印が動き制約のかかる方向を表している. これらは独立した幾何制約であるので, 異なった方向で制約が課されている事がわかる.

4.3.3 動き制約と 2 次元ランダムパターンを用いた時空間符号

2 次元ランダムパターンを用いてカメラ画像, プロジェクタ画像の座標の時空間符号をすることで密な対応付けができる. 提案法は対象物に 1 枚の 2 次元ランダムパターンを投影し, 対象物が動く様を撮影することで符号化を行う. 関連研究で述べたように, シングルスリットの投影パターンはエピポーラ制約のみでは間接反射に頑健でないシーンが発生する恐れがある. またシングルスリットパターンではローコストカメラでの低 frame rate でも高密度な復元ができないが, 2 次元パターンであれば高密度に復元ができると思われる. このため前述したように 2 次元のランダムドットあるいはランダムスリットパターンを採用する. ランダムパターンは少ない撮影枚数で符号化できることが知られており, また欠測にも強いことが知られている [23, 25, 38].

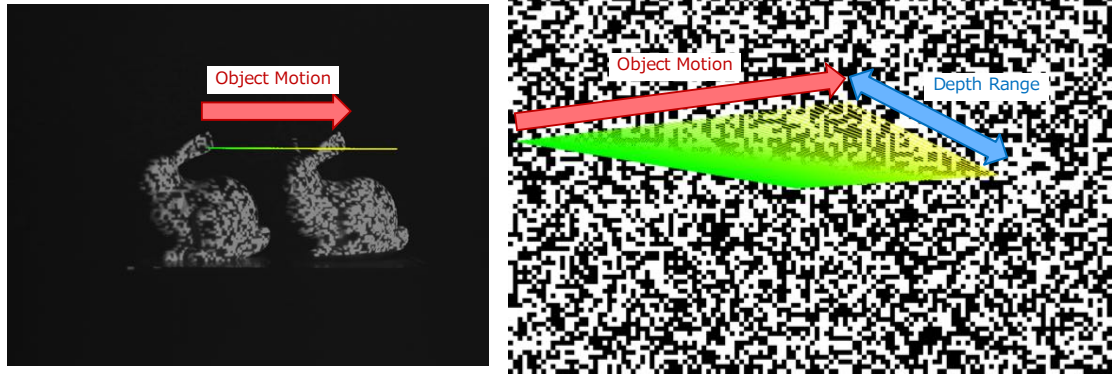


図4.5 動き制約で1つのカメラ画素で発生させた全仮説点を移動させながらカメラ画像，プロジェクタ画像へ投影した図．スタート位置に近いほど緑で，エンド位置に近いほど黄色で着色している．またバニーを撮影したスタートとエンドのフレームの画像をアルファブレンドしている．

4.3.4 光学現象の影響の抑制

動き制約とエピポーラ制約を用いること，2次元ランダムパターンを用いることで光学現象に頑健に高精度な復元が可能である．2次元ランダムパターンは，一部のフレームで白飛びが発生してしまうケースでは欠測に頑健であるために安定的に仮説点から復元点を選出できると考えられる．またこのパターンは，直接反射と間接反射が重畳してしまうようなケースも，ランダムパターンが独立性が高いため，ランダムスリットパターンの場合はエピポーラ線内部の，ランダムドットパターンの場合はエピポーラ線内部と外部の座標符号の分離ができる．この符号化パターンの構成とエピポーラ線との関係性については，3章の関連研究を参照されたい．さらに前述した2つの幾何制約によって頑健に直接反射の点を同定することができる．また鏡面反射による欠測の影響などは，対象が動くことで対象物と計測系との位置関係の変化することで抑制される．

4.3.5 実装法

本稿では議論を単純化するために，単一方向に等速で移動する物体を考える．初期撮影位置での剛体の対象物上の1点を推定するために生成する仮説点の移動について考える．時刻 t のとき対象物の移動速度を v とすると，小さな移動時間 Δt を考え，前述した仮説点の時刻 t の位置から時刻 $t+1$ の位置への移動 $[R_s^{t \rightarrow t+1} \mid T_s^{t \rightarrow t+1}]$ (式(4.3)) は以下のように表せる．

$$p(d, t+1) = p(d, t) + \Delta t v$$

ここで Δt は時刻 t で撮影されたフレームから次の時刻 $t + 1$ でフレームの撮影時間までの微小な時間を表す．本論文の実験ではこちらの仮説点の移動計算を用いる．また，前述した相関演算 $CC()$ には，物体の反射率や面の向きなどによる輝度変化の影響を抑制するために，ZNCC (Zero-mean Normalized Cross-Correlation) を用いる．

4.4 実験

平面や球などのプリミティブな形状での形状復元誤差の解析，フレーム数を下げた場合の復元密度の変化，様々な材質物の形状復元において起こる光学現象へのロバスト性，から本手法の有効性を評価する．

カメラは Optronis 社のモノクロカメラ CP70-1HS-M-1900 を用い，入力画像の解像度は 1024×768 画素とした．またプロジェクタは TEXAS INSTRUMENTS 社の DLP4500 を用い，プロジェクタ画像の解像度は 1280×800 画素とした．これらについて Zhang のキャリブレーション法により参照物体に対する内部・外部パラメタを求めた [16]．

すべての実験で，1枚の固定パターンのみを用いて計測をした．この条件下ではスリット光投影法とコストはさほど変わらない．投影パターンには図 4.6 のパターン画像 3 種を用いた．提案法ではランダムドットパターン，ランダムスリットパターン，既存法ではシングルスリットパターンを用いた．実験に用いた計測対象の多くでは間接反射成分の影響を非常に強く受けるため，SN 比を向上させるためにプロジェクタ画像の 10×10 画素でドット，あるいは，10 画素のラインを生成した．また後述するバニーの樹脂模型の疑似真値を求める際はこのドットの中心を対応点とした．固定ランダムスリットパターン，固定ランダムドットパターン，樹脂模型の疑似真値を求めるのために用いる複数のランダムドットパターンの生成にはメルセンヌツイスタを用いた [28]．



図4.6 実験に用いた投影パターン．左：ランダムドットパターン（提案法），中：ランダムスリットパターン（提案法），右：シングルスリットパターン（既存法）

対象物を平行移動させるためにシグマ光機社のスライドステージ OSMS26-

300(X) を用いた．1mm 移動したら撮影するシステムを作り，パターン投影された移動物体の撮影データを取得した．またスライドステージ上に参照物体を載せて移動させ，参照物体上の特徴点を用いて Object Motion のパラメタを得た．参照物体にはチェスボードを貼り付けた金属板を用いた（図 4.7）．様々な角度をつけた金属板をステージで移動させながら複数回，撮影を行い，それぞれの画像上でチェスボードのコーナー点を検出し，この点の三次元点を世界座標系の点として求めた（図 4.8，図 4.9）．これらの三次元点の移動量の平均値を求め，ステージ上の物体の Object Motion パラメタとして用いた．

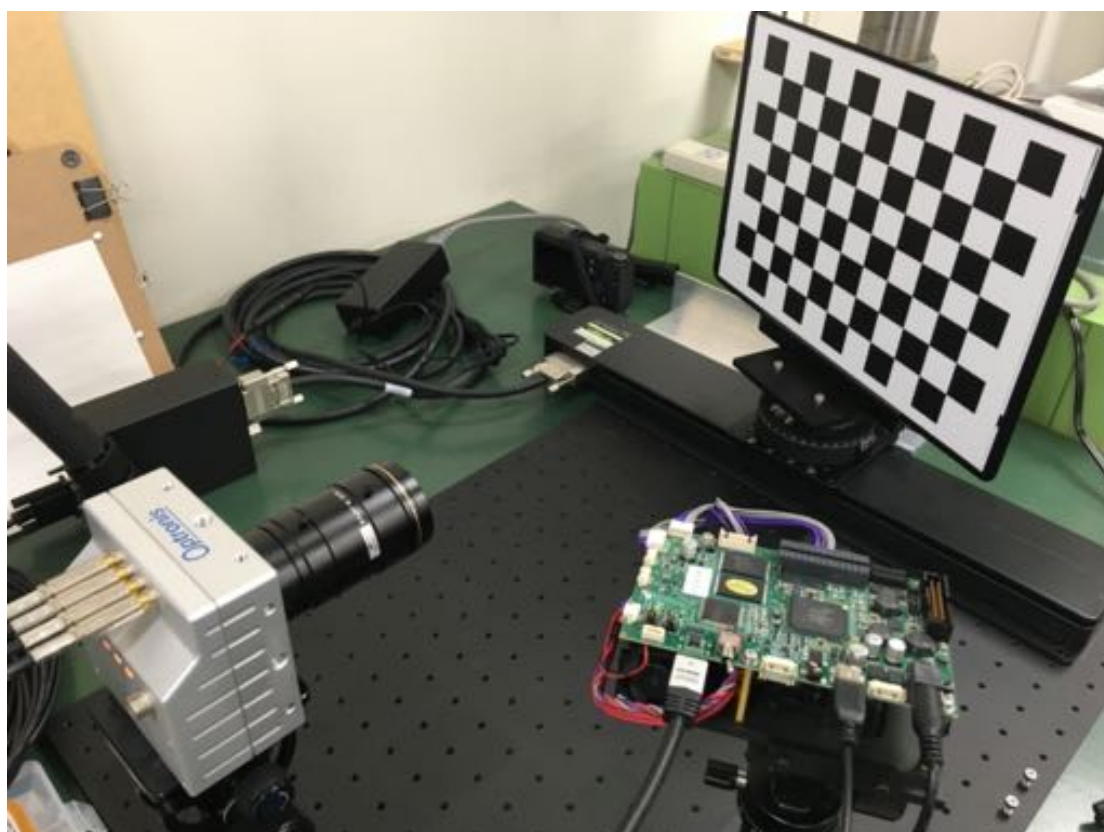


図4.7 Object Motion パラメタのキャリブレーションの系．チェスパターン上の世界座標系の点を用いて Object Motion パラメタを求め，カメラとプロジェクタに世界座標系の三次元点から投影ができるようになるためのキャリブレーションをする．

計測対象物は Gloss-Free Plastic Board(Lambert Board), Glossy Plastic Board, Wooden Stands, Aluminum L-shaped Angle, Resin Cube, Styrofoam Ball, Wood Ball, Ceramic Stands, Bunny's Resin Model である．また全撮影フレームにおいて直接反射成分が観測されない点，例えば移動中に投影範囲外に移動してしまった物体上の点や移動中に遮蔽された点は復元対象外として形状復元を行った．また鏡面性の強い Stainless Board, Stainless Ball 上の点や，光を吸収する Shading Plate

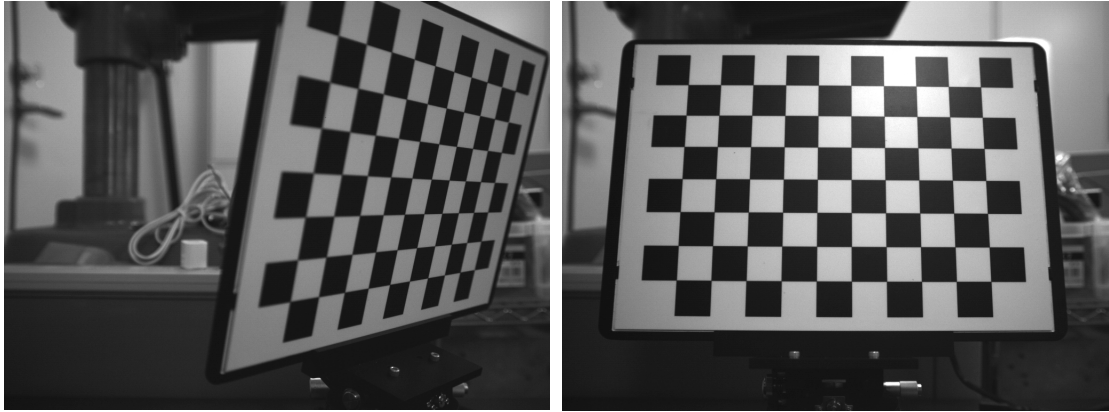


図4.8 Object Motion パラメタの推定用のキャリブレーション画像. 左：撮影画像 43, 右：撮影画像 134.

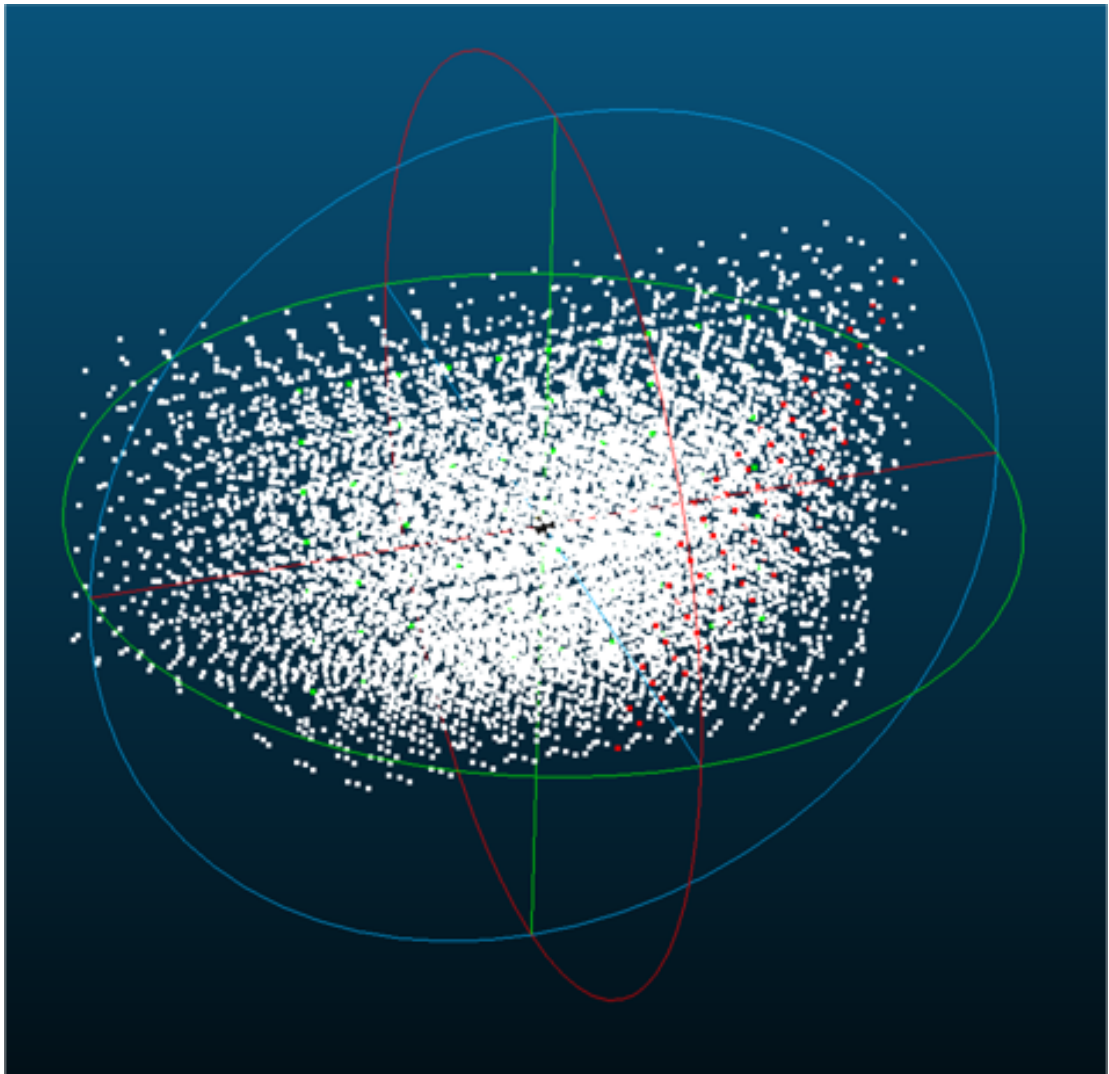


図4.9 検出したチェスボードのコーナー点の 3D プロットの結果. 天地と左右とも実世界と合っている. 確認用に 43, 134 の画像の点に色 (43 は赤, 134 は緑) を付けている.

上の点も復元対象外とし形状復元を行った。

4.4.1 計測密度と対象物のカメラフレームレートに対する移動速度

対象物が移動，かつ，低 frame rate のカメラにて対象物を撮影をして三次元計測をするシーンを想定した実験を行った．実験の撮影枚数はシングルスリットパターンにて対象物全体への投影・撮影が可能な枚数で実験している．また，これと同じ撮影方法で 4.4.2 節で撮影を行った．図 4.10 に各パターンで $N = 54(1\text{mm}/\text{frame})$ で得た復元結果とカメラフレームレートに対して高速な移動体を撮影したと仮定して撮影枚数 N を間引いた際の復元結果との点群量の比を示す．各パターンで得られる復元精度が同程度になるようパラメタ調整をし実験した．このグラフから 1 次元パターン計測手法のシングルスリットパターンの結果は当然ながらフレームが減るにつれて点群の量は減っていく．一方で 2 次元パターン計測手法であるランダムスリット，ランダムドットはどちら高い復元率を維持しており，低速度カメラで移動する対象物を撮影したとしても十分な密度の点群を得ることができる．図 4.11 に復元結果の定性的な結果を示す．こちらの結果でも対象物の移動速度に対して撮影フレームレートが小さい場合 ($N = 9(\text{move } 6\text{mm}/\text{frame})$) は，シングルスリットパターンではスパースな復元となり対象物の外形を得ることが難しい．一方で 2 次元パターン計測手法であるランダムスリット，ランダムドットはどちらも $N = 9$ の結果と撮影枚数が十分に大きい場合の結果 ($N = 54$) とがほぼ同等の密度で計測が行えているため，十分な密度で計測が可能であることがわかる．また撮影枚数が少ない場合 ($N = 9$) は座標符号化のユニークさが低下するためにノイズが多少増加するものの，ロボットアームでのピッキングタスク等を考慮した場合，対象物の外形の認識が可能なレベルで復元できていることがわかる．

4.4.2 計測精度

シングルスリットパターン，ランダムスリットパターン，ランダムドットパターンにて同条件で撮影をし，どの程度の計測誤差があるかを調べた．Resin Cube (図 4.18 (b)) の 1 面を平面として用い，Ceramic Stands に載せた Wood Ball (図 4.18 (c)) を球体として用いた．球の結果では球の台座の Ceramic Stands も復元されてしまうので，その台座領域はカットして球にフィッティングし誤差を求めた．また，シングルスリットパターンの計測結果のみノイズが多く発生したため，対象物の計測結果を包む Bounding Box を手動で指定しノイズをカットしてから MAE (Mean Absolute Error), RMSE (Root Mean Square Error) を求めた．ランダムスリットとランダムドットはノイズ除去処理をせず，そのままの生のデータで MAE, RMSE

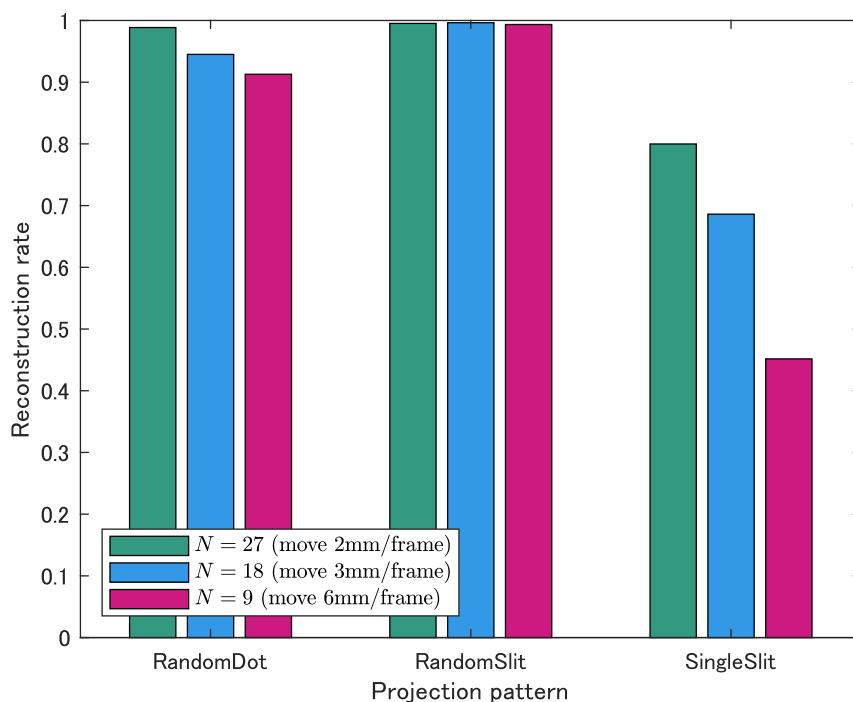


図4.10 各パターンで $N = 54$ (move 1mm/frame) で得た復元結果とカメラフレームレートに対して高速な移動体を撮影したと仮定して撮影数 N を間引いた際の復元結果との点群量の比.

を求めた. 平面と球面のフィッティング時に用いた物体 (Resin Cube, Wood Ball) の撮影画像と点群のフィッティングの様子をそれぞれ図で示す (図 4.12, 図 4.13, 図 4.14, 図 4.15). 結果, 平面と球, どちらもランダムスリットパターンが最小の誤差で形状復元をし, それに近い誤差量でランダムドットで計測したことを確認した. 一方で, シングルスリットパターンは白く撮影された幾何制約を満たす領域すべてで計測がなされてしまうためにノイズが発生してしまいランダムパターンに比べ誤差が大きくなった (図 4.16).

また, 静止したバニーの樹脂模型に対して 100 枚のランダムドットパターンを投影・撮影し形状復元した結果を疑似真値として, これと各パターンで得た復元結果とを比較して復元成功率と MAE を求めた. ここで復元成功率とは, 疑似真値が計測点を持つ座標に対して, 各パターンでの計測データが共通の座標をもつ割合である. また MAE は疑似真値と同座標をもつ計測データと疑似真値との差を用いて計算される. 結果は図 4.17 のようになった. 左は疑似真値と各パターンでの計測結果で共通する点の割合, 右は各パターンでの計測結果と疑似真値との MAE である. シングルスリットパターンは計測点は疑似真値とほぼ同等に得られるものの全体に

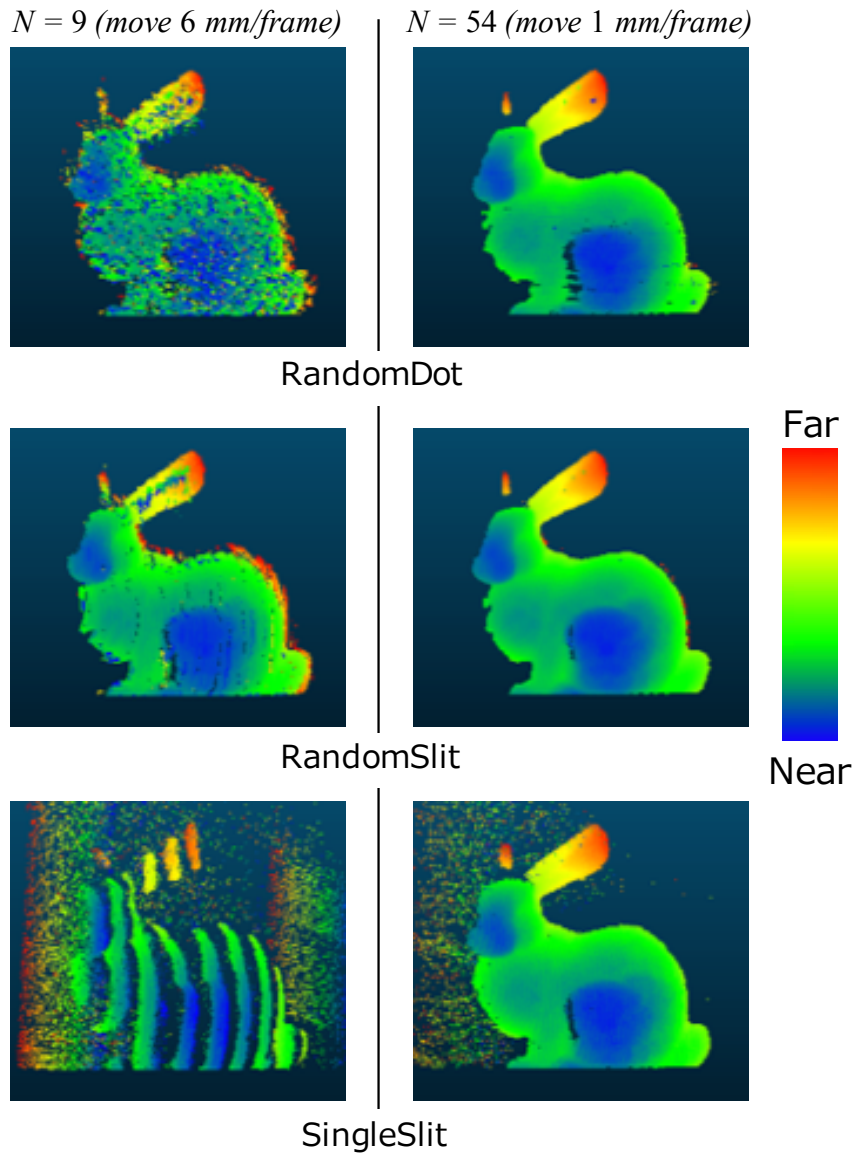


図4.11 各投影パターンで得た左 : $N = 9$ (move 6 mm/frame), 右 : $N = 54$ (move 1 mm/frame) の復元結果.



図4.12 平面の評価に用いた Resin Cube の部屋照明下と白投影時のカメラ画像.

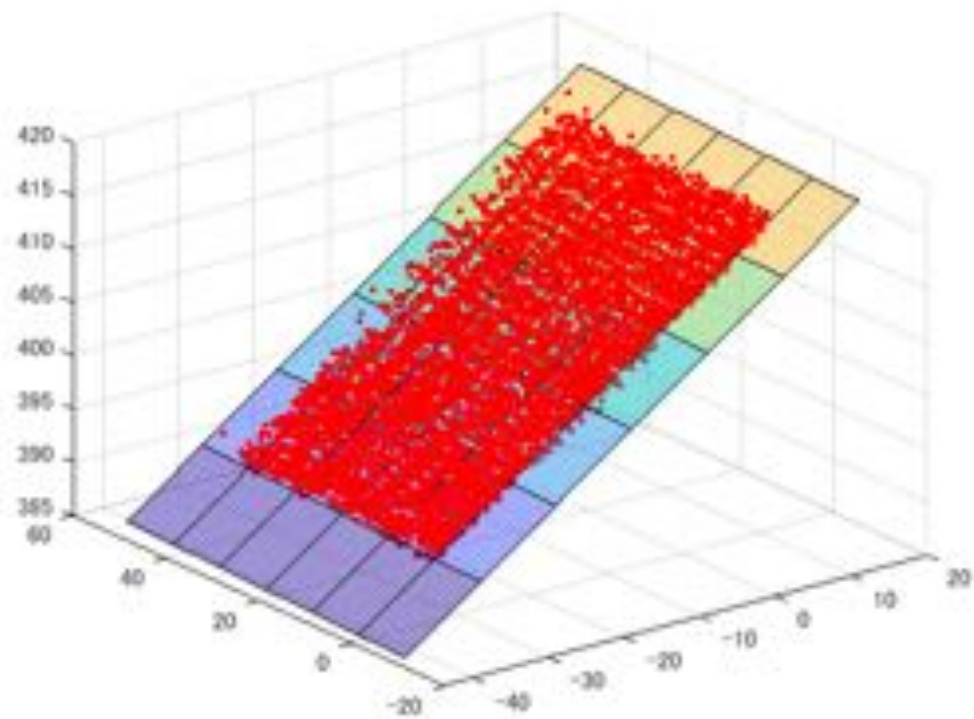


図4.13 平面の評価：最小二乗フィッティングの結果.



図4.14 球面の評価に用いた Wood Ball の部屋照明下と白投影時のカメラ画像.

ノイジーな復元結果となっているために精度はやや低い傾向にある．ランダムドット及びランダムスリットは同等の性能を示しており，ある程度の復元成功率をもちつつ高い精度で復元ができることがわかる．

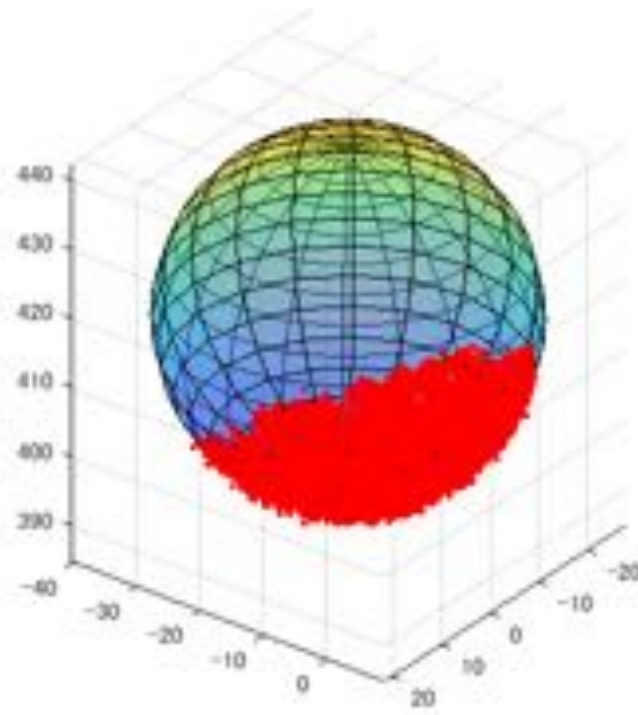


図4.15 球面の評価：最小二乗フィッティングの結果.

4.4.3 光学現象の影響の抑制

様々な形状，材質物の形状復元を行う実験をした．結果を図 4.18に示す．(a) では Crumpled Aluminum Foil から発せられる弱い投影光の相互反射光の影響を抑制できるかを調べている．復元の結果からどの手法でも弱い相互反射に対するロバスト性があることを確認した．また (b) では，Aluminum L-shaped Angle で起こる弱い相互反射，および，Glossy Plastic Board にて発生する Blown-out areas での観測データの白飛びによる欠測の影響を調べた．結果，(a) と同様にどの手法でも弱い相互反射に対するロバスト性があることを確認し，またランダムドット，および，ランダムスリットパターンにおいては白飛びによる欠測の影響も抑制できることを確認した．次に (c) では，Balls(Stainless, Styrofoam, Wood), Ceramic Stands を用いて，強い相互反射，弱い相互反射の影響を抑制できるかについて調べた．(c) の復元結果は上が正面から見た図，下が上から見た図である．結果，ランダムスリットパターンについては白矢印部にて相互反射により誤った形状が計測されており，またシングルスリットパターンにおいても Stainless Ball と Styrofoam Ball の間で相互反射の影響により形状計測が正しくできなかった．ランダムドットパターンにつ

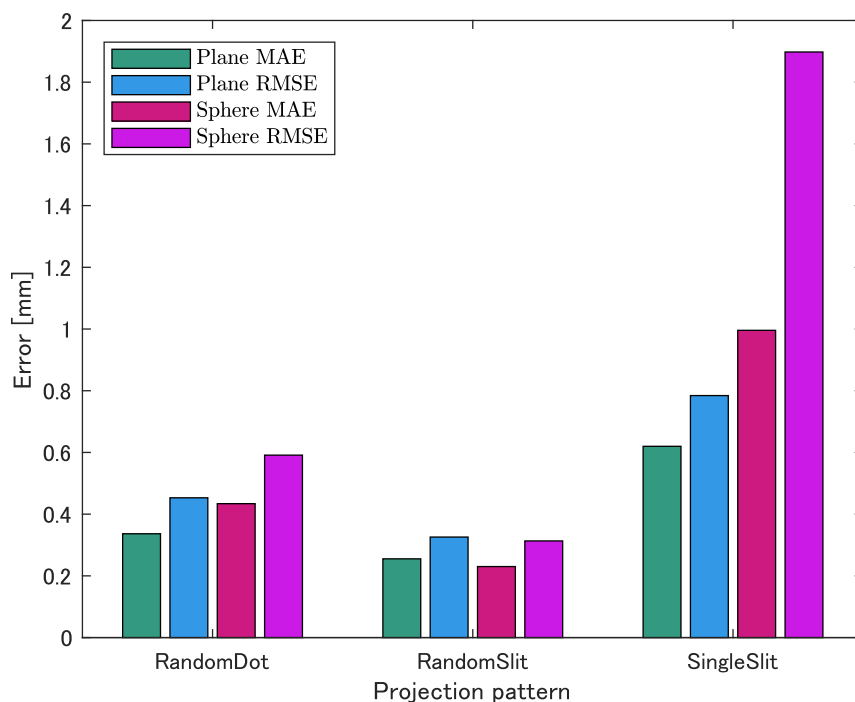


図4.16 各パターンでの計測結果を最小二乗法でフィッティングした平面あるいは球と計測データとの MAE と RMSE. 撮影枚数は Resin Cube : $N=29$, Wood Ball : $N=27$ である. 移動速度とカメラ frame rate の関係は $move\ 1mm/frame$ である.

いては全体に相互反射の影響を抑制し形状復元ができた. (d) と (e) は同一シーンに対し Shading Plate があるかないかの差を設け計測実験をしている. この Shading Plate により, Stainless Board により生じる相互反射の影響を調べることができる. ランダムスリットパターンにおいては白矢印部で誤った形状が得られてしまっている. またシングルスリットパターンも同様に白矢印部にて誤った形状が得られてしまっている. ランダムドットについては全体に良好に形状計測ができた.

ランダムスリットパターンはライン間の境界部が少なく隣のラインの影響を受けづらく光学現象が起きていないシーンにおいて特に良い形状復元の傾向が見られた. ただし, 相互反射によるエピポーラ線外の影響を受けやすく相互反射が発生するシーンでは誤計測してやや大きなノイズ塊が生まれやすい傾向が見られた. ランダムドットパターンは座標復号化時に隣接ドットの影響を受けやすく点が復元されづらい傾向はあるものの, 相互反射や白飛びによる欠測に対するロバスト性は高かった. このロバスト性の高さは2次元にランダム符号を配置しているため, エピポーラ線の内外の影響を抑制できたためであると考えられる.

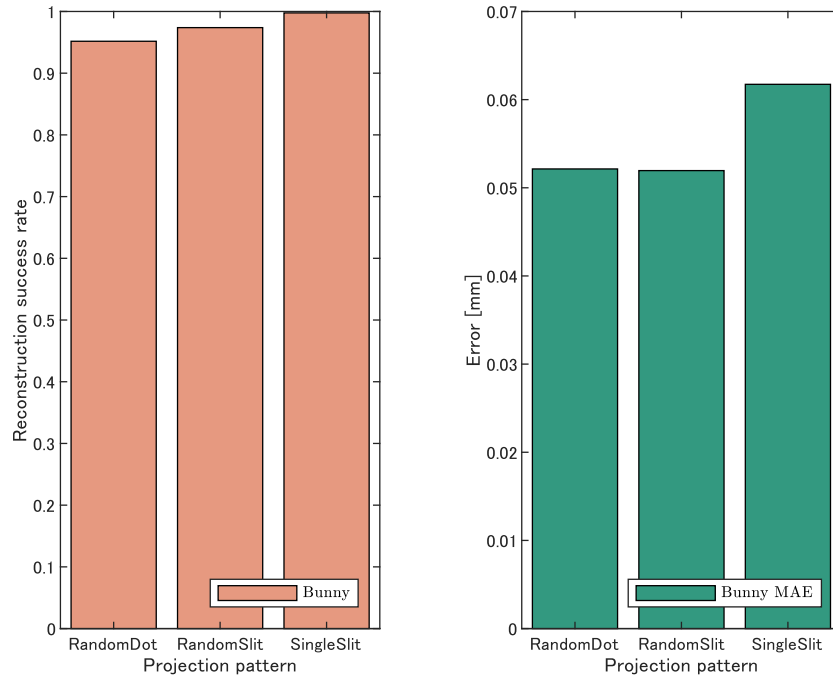


図4.17 左：ランダムドットパターン 100 枚で得た疑似真値と各パターンでの計測結果で共通する点の割合．右：各パターンでの計測結果を疑似真値との MAE. 撮影枚数は Bunny's Resin Model : $N=54$ である．移動速度とカメラ frame rate の関係は $move\ 1mm/frame$ である．

4.5 まとめ

本稿では，産業シーンで自然に知ることができる物体の動きを幾何学的制約として，2次元ランダムパターンを移動する物体に投影したシーンを撮影し，物体を形状復元をする手法を提案した．提案手法は，低コストのカメラで物体の速度に比べて低いフレームレートで少ない枚数の画像を撮影した場合でも，物体の形状を高密度かつ高精度に，また光学的なアーチファクトに対して高いロバスト性をもって計測することができる．実験では，提案手法が既存のシングルスリットライン投影法よりも少ない撮影枚数で高密度に形状復元できることを確認した．また，スリットライン投影法の結果と比較して，間接反射がほとんど発生しないシーンでは，ランダムスリットパターンを用いた提案手法の形状復元精度が高く，鏡面反射の白飛びや相互反射などの光学的なアーチファクトが発生するシーンでは，ランダムドットパターンを用いた提案手法のロバスト性が高いことがわかった．

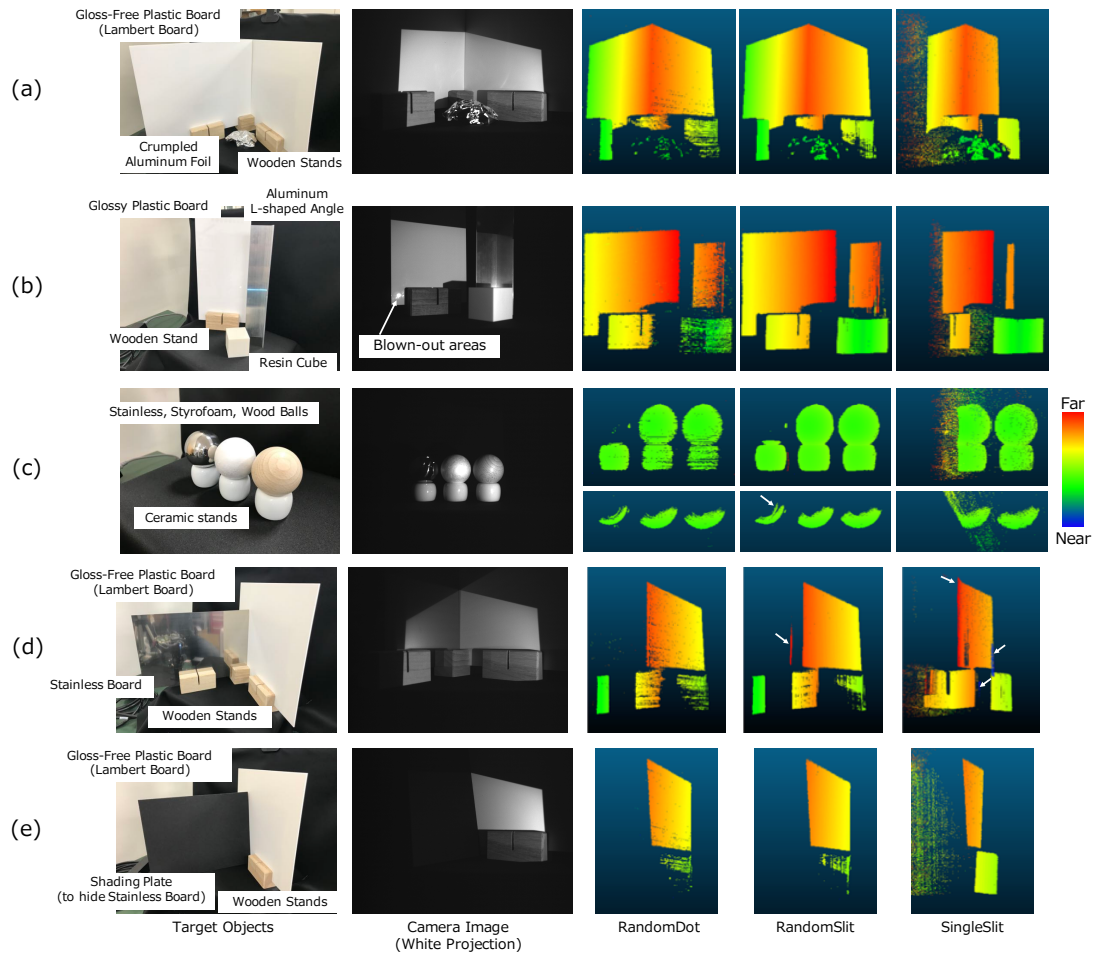


図4.18 様々なシーンでの各投影パターンでの形状復元結果. それぞれの撮影枚数は (a) $N=50$, (b) $N=50$, (c) $N=56$, (d) $N=50$, (e) $N=50$ である. 物体の移動速度とカメラ frame rate の関係は $move\ 1mm/frame$ である.

第 5 章

提案手法 4： 遮蔽空間への補助点群を用いた強 隠蔽物の 3 次元位置のセミリアル タイム推定^{*2}

5.1 背景^{*21}

本研究は、発見が困難な対象物を検出することを目的とした 3D 認識技術について検討する。このような課題は、安価なロボットアームと 3 次元センサで構成されるシステムを用いた自動物体認識システムにおいて存在する [46]。例えば、物流を自動化するロボットを製造する場合、さまざまな物体がばら積み、あるいは、箱に大量に整列された状態のシーンを認識する必要がある（図 5.1）。

同様の課題で、スタンドアロン HMD による作業支援において、作業者の安全性を向上させるアプリケーションの実現がある。ここで、スタンドアロン HMD は、スタンドアロンで動作する演算処理装置を持った、AR グラス、あるいは、VR ゴーグル＋カメラなどの Head Mount Display とする。屋外での飛行機などの大型機器の整備時、組み上げるパーツを認識し、指示書をベースに作業手順を指示するスタ

^{*2} 本章は、International Journal of Automation Technology Vol.13, No.4, pp.464-474, Copyright(C)2019 IJAT に掲載された同著者による”Estimating 3d position of strongly occluded object with semi-real time by using auxiliary 3d points in occluded space” [45] に基づいています。本章の素材は、International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] で一部発表されたものであり、本章の一部の図は、富士技術出版株式会社の許可を得て、[45] から再利用されています。

^{*21} 本節の素材は、International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] の 1 節に一部掲載されたものであり、本節の一部の図は、富士技術出版株式会社の許可を得て、[45] から再利用されています。

ンドアロン HMD を用いて、作業効率の向上を意図する研究がある [47–49]。このような作業環境においては、指示されたパーツ付近に危険な物体があり、それに誤って触った場合、怪我をするリスクがある。例えば、HMD に指定されたパーツ付近にあるラップトップ PC に隠れて見えづらい小刃が多量についたホールソーなどが付随したドリルがあった場合、このドリルで手に怪我をする危険性がある (図 5.2)。

このような課題について、本稿では HMD に内蔵された RGBD センサによって、危険性のある物体を検出し、提示することで解決する方法を一事例としてフォーカスし考える (図 5.3)。ユーザに、行動の危険性を速やかに伝える必要があるため、この機能には、セミリアルタイム性が必要である。ここでセミリアルタイムとは、1.0 sec 以下の処理速度であること、と定義する。一般に物体検出機能は処理量が大いだが、軽量化や省電力化のために HMD 内蔵の計算リソースはハイエンド PC よりも貧弱であるため、処理量を小さく抑えることができ、HMD の環境でもセミリアルタイムに実行できる手法が望まれる。また、HMD 内蔵センサを用いた物体検出機能においては、一人称視点でのセンシングとなるため、検出対象物が他の物体に強遮蔽されるケースが多い。ここで強遮蔽とは、物体の見える 50% 以上が遮蔽された状態であること、と定義する。この強遮蔽によって検出対象物のデータが欠損されてしまい、対象物の検出が困難となる。

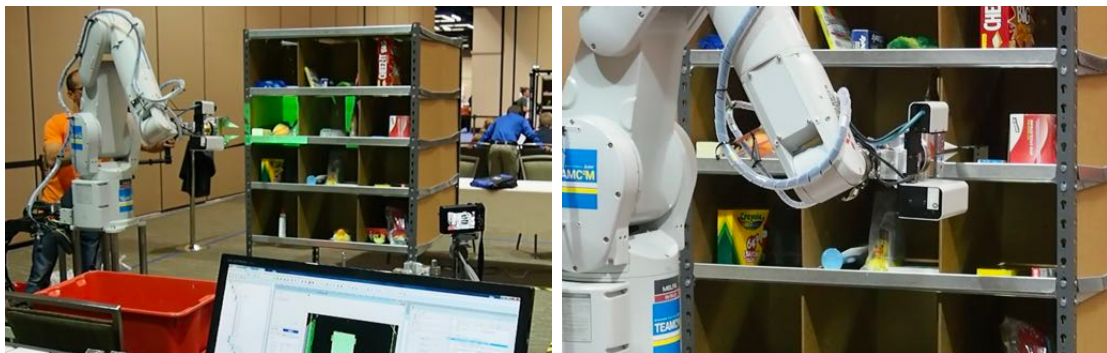


図5.1 機械知覚 & ロボティクスグループ中部大学のアマゾンピッキングチャレンジの様子。(Copyright(C)2015 機械知覚 & ロボティクスグループ中部大学, [46])

RGBD データのマッチングによる物体検出には、いくつかの手法がある。その中でも、デプスデータから得た三次元点群に対し、点同士の普遍かつ幾何的な性質を用いたスパースな特徴量を用いたマッチング手法 [50, 51] は、RGBD 画像の画素単位でデンスに情報を扱う手法 [52] に比べ、データ量が軽量であることから、高速に計算できる。しかし、この手法は、大部分が遮蔽された物体であると検出が難しい課題がある。

そこで、本研究ではスパース情報を用いた手法に対し、強遮蔽に対するロバスト

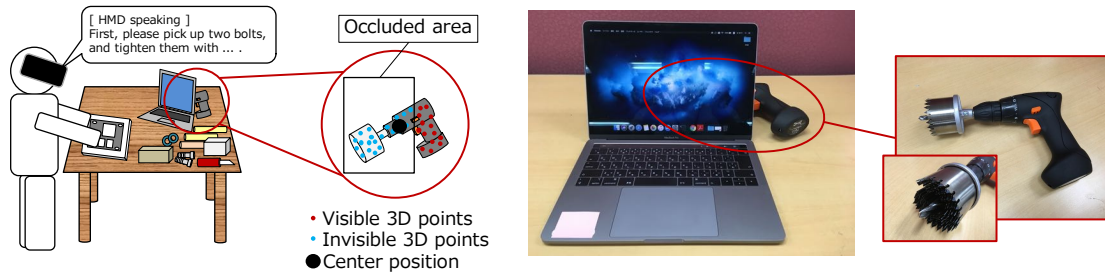


図5.2 作業現場にて，強遮蔽された危険物があった場合の，危険物の検出イメージ．（Copyright(C)2019 IJAT, [45] 図 1）



図5.3 危険物（検出対象物）を検出した際の，HMD での表示イメージ．（Copyright(C)2019 IJAT, [45] 図 2）

性を向上させる手法を提案する．提案法は，検出対象物に強遮蔽による欠損があった場合，遮蔽された空間に限定して欠損埋めの仮説をしきことで，検出の精度と処理効率を向上させることができる．

5.2 関連研究^{*22}

三次元点群を用いた物体検出手法を，テンプレートベース，デンス情報ベース，スパース情報ベース，3つのアプローチに大別し，説明する．

5.2.1 テンプレートベースの手法

はじめに，テンプレートベースのアプローチは，テンプレートとなる特定物体の3Dモデルから，視点依存しない頑健な特徴量を得ることで，高速かつ頑健な特定物

^{*22} 本節の素材は，International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] の2節に一部掲載されたものであり，本節の一部の図は，富士技術出版株式会社の許可を得て，[45] から再利用されています．

体検出を可能とする手法である．テンプレートベースアプローチの代表的な手法として，LINE-MOD [53] が挙げられる．この手法は，カラー画像の勾配と，デプス画像の法線ベクトルを特徴量として用い，シーン中でテンプレートと最も近い領域を検出結果とする．Stefan らは，LINE-MOD を応用し，対象物体モデルを仮想カメラで多方向から撮影した 2D 画像すべてをテンプレートとして用いることで，高速に処理が可能な手法を提案している [54]．これらのアプローチの手法は一般には高速に動作するため，セミリアルタイムに物体を検出可能だが，強遮蔽が原因でテンプレートとかけ離れた見えの物体の検出が難しく，検出率が下がってしまう傾向がある．また，強遮蔽に対応するためにはテンプレート量を大幅に増やす必要があるが，それでは処理速度が低下してしまい，セミリアルタイム性を期待できない．

5.2.2 デンス情報ベースの手法

次に，デンス情報ベースのアプローチは，デプス情報，近傍座標間の連続性，オブジェクト構成などの制約により，マッチング対象のモデルを仮に配置したシーンと観測シーンとの差を最小化する手法であり，カメラ画像とデプス画像の全画素を用いて，モデルの最適な位置姿勢推定を行う．Brachmann らは，デプス状態，観測座標状態，オブジェクト構成制約により，対象物の 6DoF (Degrees of Freedom) を最適化する手法を提案しており，RGBD 画像全体で最適な結果を出力をするため，ロバスト性が高い [55]．同様の方針で，CRF (Conditional Random Fields) を用いたシーン最適なモデルの 6DoF 推定をする手法 [52, 56] や，DNN を用いて物体の局所的な見えの情報を学習し高性能に物体検出を行う手法 [57] が提案されており，いずれも精度が高い．しかし，[52] に示されているように，このアプローチは全ての画素情報を特徴として扱うため，他のアプローチに比べ処理量が大きくなる傾向にある．例えば，11 万点程度の物体認識に，デンス情報ベースの手法の代表的なアプローチである Brachmann らの手法 [55] では 2 sec 程度が必要であるが，スパース情報ベースの代表的なアプローチである PPF (Point Pair Features) [51] では 0.4 sec 程度で処理を完了できる．このように，デンス情報ベースの手法は，ローレベルの計算リソース上でセミリアルタイム処理を実行することは困難である．

5.2.3 スパース情報ベースの手法

最後に，スパース情報ベースのアプローチは，点群中のある点と，その点の周辺の点群との相互の位置関係から特徴量を記述し，点群同士を対応付ける手法である．このアプローチは，扱う特徴を，RGBD センサの特性や，物体を観測する位置姿勢，物体が置かれている環境，などの要因にロバストと考えられる特徴量に限定す

るため、精度を維持しつつ、テンプレートベースの手法以上に高速な物体検出を行うことが可能であり、処理速度の面で考えると最もセミリアルタイムでの実行可能性の高い手法と考えられる。特徴量記述の代表的な手法として、Spin Image [58], FPFH (Fast Point Feature Histograms) [50], SHOT (Signature of Histograms of Orientations) [59,60], PPF [51] がある。Spin Image は、注目点の周囲の点群から法線ベクトルを割り出し、特徴記述のための局所座標系を定義して、回転不変な特徴量を記述する手法である。FPFH や SHOT は、注目点の周囲にサポート球を仮定し、サポート球を複数領域に分割し、各領域内にある点群量をヒストグラムとして特徴記述する手法で、RGBD センサのノイズが強いデータに対するマッチングのロバスト性が高いと考えられる。PPF は 2 点間の距離や角度の情報を特徴として記述する手法で、処理を高速にでき、またそれら 2 点の間の点が欠損しても特徴量は不変であるため、他のスパース情報ベースの手法に比べ、遮蔽のあるシーンに対するロバスト性が高いと考えられる。さらに、これらの手法をベースに発展させた様々な研究がなされている [61,62]。しかし、これらの手法では強遮蔽が発生しているシーンにおいては、強遮蔽部の特徴量が欠損してしまうため、高いマッチングスコアが得られず、検出が困難となる傾向がある。ここでマッチングスコアとは、すべての対応点同士の特徴量の距離の総和を表すスコアと定義する。また、強遮蔽を考慮した手法がいくつか提案されている。Johnson らは、Spin Image 特徴をベースに、構造が似ている局所形状を探索するアプローチで、遮蔽に頑健かつ高速なマッチング手法を提案している [63]。Akizuki らは、点群の構造に依存しづらい特徴点の局所座標系の記述法を提案しており、対応点の特徴記述を正規化し対応点同士のマッチング率を上げることで精度を向上させ、遮蔽に対する頑健性をも向上させている [64]。これらの手法では、強遮蔽が発生した際の認識率を向上させるための可視領域に対する処理法の提案がなされている。ただし、可視領域の特徴量が不足している場合は、マッチング対応点の数が不足してしまい、マッチングが失敗する恐れがある。

この課題を解決するため、本研究では、高速処理が可能なスパース情報ベースのアプローチを応用し、強遮蔽に対してロバストとなるように改善できる、遮蔽空間推定、および、その空間への補助点群生成の手法を提案する。

5.3 提案法^{*23}

^{*23} 本節の素材は、International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] の 3 節に一部掲載されたものであり、本節の一部の図は、富士技術出版株式会社の許可を得て、[45] から再利用されています。

強遮蔽された物体は一部しか観測されないため、あらかじめデータベースにもっている物体の点群モデルと、それが一致するかを調べるときに、全体的な点群間の誤差が大きくなる、あるいは、同じとみなされる対応点の数が小さくなる。本研究では、データベースに登録されている物体の点群（以降、参照点群）の一部（以降、部分参照点群）とシーン点群とをマッチングする手法と、シーン中で遮蔽されずに観測された対象物の一部（以降、部分観測点群）の欠損部を仮想的に補助してデータベースの物体点群とマッチングする手法、の二つの手法を考えた。前者の手法は、想定され得る部分参照点群をすべて作りシーン点群と部分参照点群とをマッチングする場合、参照点群とシーン点群とをマッチングする回数に比べ、マッチング回数が増加してしまい、高速な検出を期待することは難しい。後者の手法は、オクルーダによって不可視となった空間内に存在する対象物領域に、形状を仮想的に補助する点群（補助点群）を与えることで、遮蔽がない場合には観測できるはずであった部分観測点群の対応点が疑似的に得られる。この疑似対応点を用いることで、参照点群と部分観測点群との対応点の数を、背景やノイズなどの対象物の点群ではないその他の点群に比べ、大きく観測できると仮説を立てた。

この補助点群をシーン全体で生成させるとした場合、空間全体に一樣に大量の仮説をしく必要があるため、マッチング試行回数が爆発的に増加し、それによってもって計算時間が増加してしまう。また、疑似対応点が空間に一樣にばらまかれることになり、精度向上に寄与しない。そのため、センサの観測領域の中で（センサ近くに存在する）手前側の物体の背後に検出対象物が大きく隠れた場合を想定し、この背後の空間に限定して補助点群の仮説をしくことで、実時間で精度高く強遮蔽された物体の検出が可能となる、と期待できる。そのため、本研究では後者の手法を採用した提案を行う。

一方で、危険物の位置を知らせるという作業支援を対象とすると、ナイフやカッターなどの薄い物体や細い物体などのボリュームに乏しい物体が対象となりうる。一般に、スパース情報ベースの3D マッチング処理を用いた場合、ボリュームのある物体は検出が可能だが、薄い物体や細い物体などのボリュームに乏しい物体の検出が難しい。これは、ボリュームのある物体に比べ、ボリュームに乏しい物体の三次元計測点はそもそも数が少ないため、対象物とノイズなどの他の点群との区別が困難となり、マッチングが難しくなることに起因する。ゆえに、提案法の有効性を正しく検証するために、対象物はボリュームのある物体に限定し実験・評価していく。

スパース情報ベースのアプローチは、強遮蔽された物体については点群が欠損しているため検出ができない。この理由は、強遮蔽が発生している物体を検出する場合、可視領域のみでは対応点数が不足してマッチングスコアが低下するためであり、検出時に誤マッチングを引き起こす傾向が強い。そこで、提案法では、欠損部の形状を補助するための補助点群を生成して得られる疑似対応点と、可視領域で観測

された点群から得られる対応点とを用いて、マッチングスコアを向上させる。これにより、強遮蔽された物体を検出することができる。オクルーダは、入力として、RGB 画像と三次元点群が得られるため、既存の画像認識処理や点群処理により、検出することができる。最終的に、シーン点群と補助点群とを組み合わせ、対象物を 3D マッチングにより検出する。提案法は、

- 遮蔽空間の推定
- 遮蔽空間内での補助点群の生成
- 物体検出

の 3 つのステップに分けることができる。

5.3.1 遮蔽空間の推定

遮蔽領域の推定ステップのフローチャートを図 5.4 に示す。

はじめに、HMD 内蔵の RGBD センサから得られた RGBD データと検出対象物の 3D 点群モデルデータセットを入力とする。

この RGB 画像に対して、本手法では DNN を用いた物体検出手法の中でも、高速に処理が可能な YOLO [65] を用いて、HMD の近くにある手前側の物体を検出し、これをオクルーダとする。YOLO で検出された物体は、その物体を検出した検出枠の内部では、中心付近に存在していると考えられる。そのため、この物体検出枠に対して、グラフカット [66] を適用し、検出された物体のシルエットを入手する。

オクルーダのシルエットは、オクルーダによって遮蔽を生じている空間を決定するために用いられる。このオクルーダのシルエットのエッジを Canny エッジ検出器 [67] にて検出し、これを拡張する。このエッジを拡張した領域を拡張エッジ領域と呼ぶ。具体的な拡張エッジ領域の求め方を以下に述べる。膨張回数 D だけ、このエッジを周囲 8 近傍に膨張処理する。そして、この膨張させたエッジの内部、かつ、シルエットの外部、という条件を満たす領域を、拡張エッジ領域とする。

求めた拡張エッジ領域の 2D 座標に対応する、3D 特徴点を選別する。このようにすることで、オクルーダ周辺の 3D 特徴点のみを抽出できる。また、3D 特徴点を用いた遮蔽空間推定を行うことで、遮蔽空間にノイズ点群などによる余分な領域を含んでしまうことを防ぐことが期待できる。本手法では、3D 特徴点抽出に 3D SIFT (3D scale-invariant feature transform (SIFT)) [68,69] を用いた。

遮蔽空間は、このように抽出した、強遮蔽された物体、および、シーン中に存在する 3D 特徴点を包含する直方体として表現される。

ここでは、HMD の RGBD センサの光軸と机上は平行と仮定する (図 5.5)。遮蔽空間は、座標軸に平行または垂直に整列したエッジを有する直方体として近似

される．この直方体は，図 5.5に示すように，エッジの位置によって一意に定義され， $\min X_j, \max X_j$ ($j = 1, \dots, 3$) は， $\min X_j = \min_k X_j^k$, $\max X_j = \max_k X_j^k$ ($k = 1, \dots, N_p$) である． X_j^k が k 番目の点の X_j 軸成分を示しており， N_p はシーン中の点群の点の総数を指す．

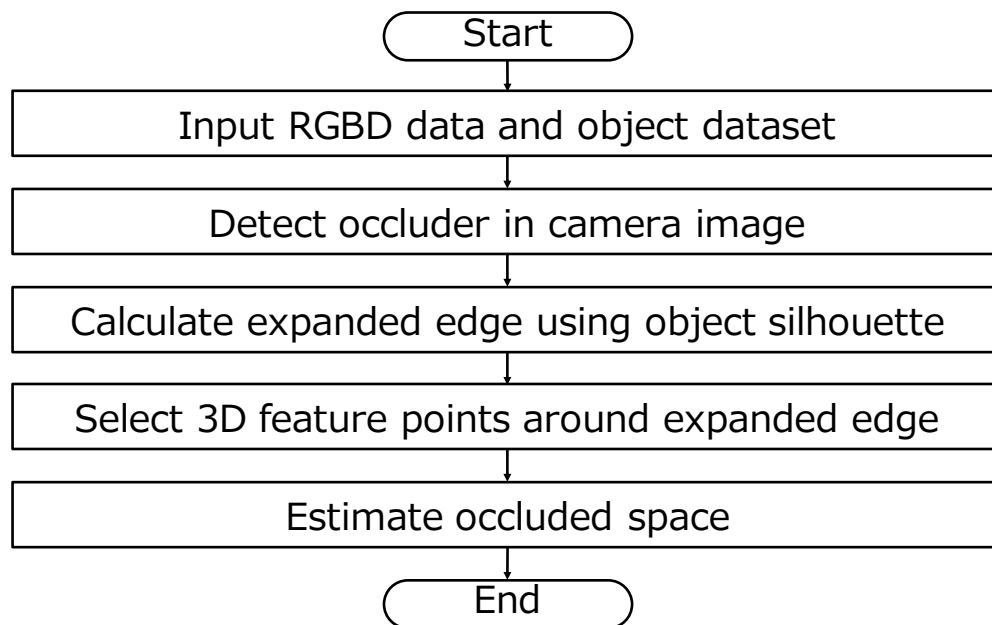


図5.4 遮蔽空間推定のフロー図．(Copyright(C)2019 IJAT, [45] 図 3)

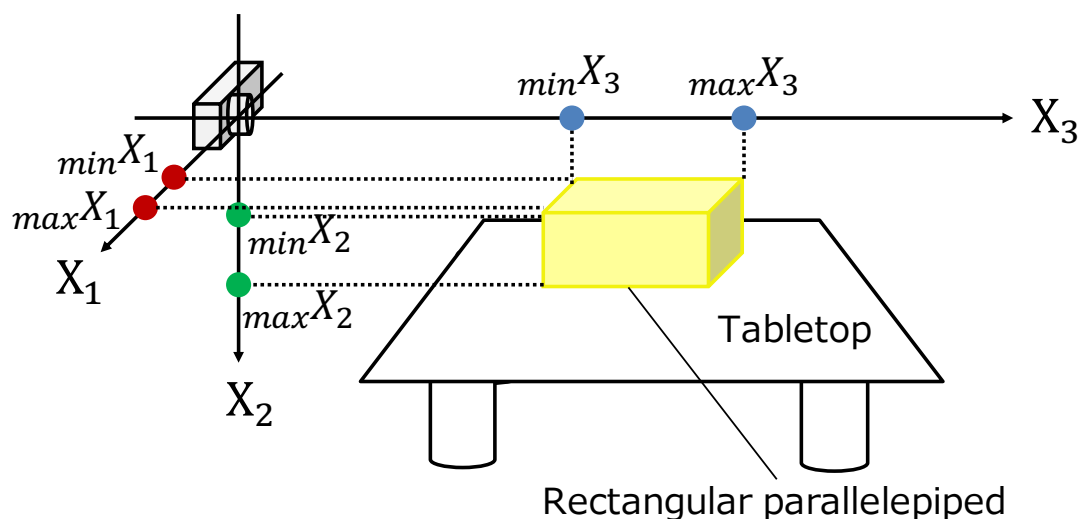


図5.5 本稿で用いる座標系．(Copyright(C)2019 IJAT, [45] 図 4)

ラップトップ PC（オクルーダ）の背後に検出対象物のコップがあるシーンで説明する（図 5.6）．最初に，カメラ画像上で物体検出アルゴリズムである YOLO を用いてオクルーダ候補を検出し，任意のオクルーダ（ここではラップトップ PC）を

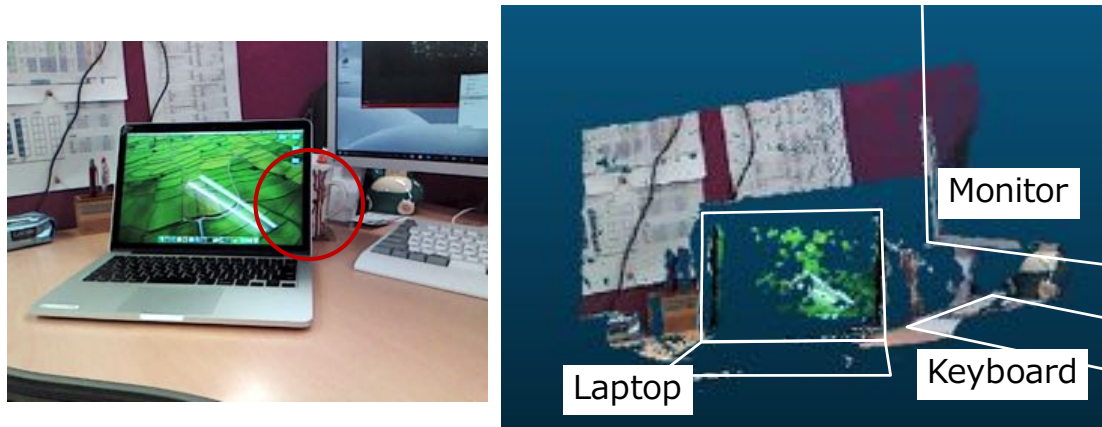


図5.6 入力シーンの例：カメラ画像，赤丸部分に検出対象物（コップ）がある（左），点群（右）．（Copyright(C)2019 IJAT, [45] 図 5）

選択する（図 5.7）．次に，そのラップトップ PC のシルエットを求め，そのエッジを拡張させた拡張エッジ領域を求める（図 5.8）．最後に，拡張エッジ領域内部の 2D 座標に対応する 3D 特徴点のみを抽出（図 5.9）し，それらの点で定義される直方体を遮蔽空間とする（図 5.10）．ここで，3D ノイズフィルタ後であってもシーンには残留してしまうノイズ点群が存在するため，このような点群を処理から排除するために，3D SIFT で得た特徴点を用いて，遮蔽空間を計算する．3D SIFT は，シーン中の特徴的な点を抽出するため，壁などの形状に特徴が少ないと考えられる背景よりも物体表面上で検出されやすいという特性を用いることができる．また，シーンに合わせたスケールパラメタで特徴点検出することで，ノイズ処理後に残ってしまう塊状のノイズ点群を処理から排除するために用いることができる．そのため，このように求めた 3D 特徴点を用いることで，遮蔽空間に不要な空間が含まれる可能性を減少させることが期待できる．具体的な処理内容としては，シーン中のメインを占めるサイズを持つ点群のみの特徴点を取得できるように，塊状のノイズ点群では特徴点が計算されない閾値（最小スケール： S_{SIFT} ，オクターブ数： O_{SIFT} ，単位オクターブあたりのスケール数： N_{SIFT} ，最小コントラスト： C_{SIFT} ）をヒューリスティックに設定し，3D SIFT 特徴点を計算する．

5.3.2 遮蔽空間内での補助点群の生成

前項で説明した遮蔽空間に対して，3D 補助点群の生成を行う．もしオクルーダが存在しないシーンであれば，遮蔽によって観測できない点群が発生しないため，対象物体を検出できた可能性が高い，と考えられる．そこで，補助点群は，オクルーダによって観測できなかった対象物体の一部の点群を表現するために生成する．以下に，この補助点群の生成法について説明していく．

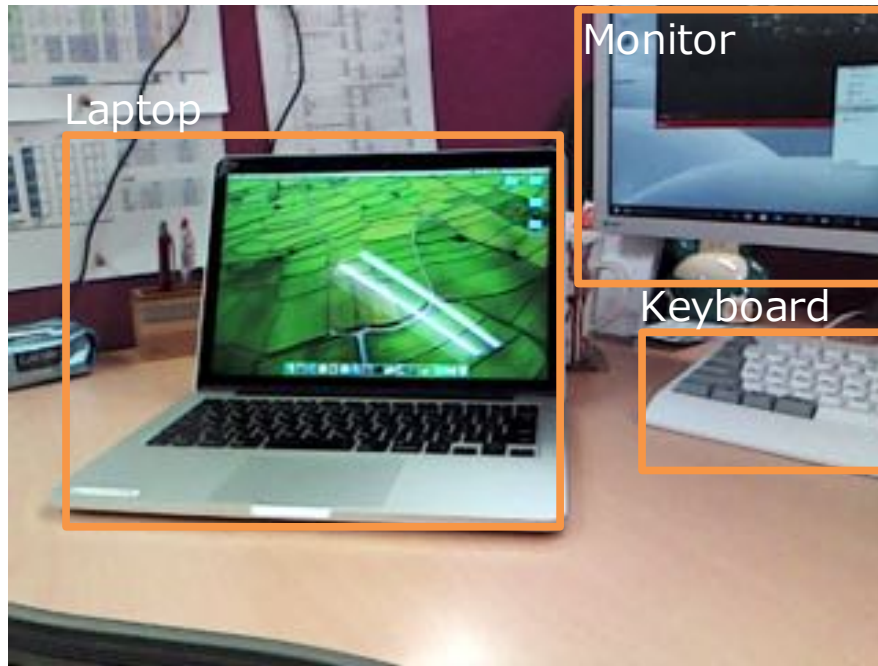


図5.7 YOLO アルゴリズムを用いた，手前側の物体（オクルーダ候補）の検出．
 (Copyright(C)2019 IJAT, [45] 図 6)



図5.8 ラップトップ PC の拡張エッジ領域の計算法：(a) は入力画像．(b) は YOLO を用いて検出されたオクルーダ（ラップトップ PC）のシルエット．(c) はシルエットの二値化画像．(d) はシルエットのエッジを拡張した領域である，拡張エッジ領域．(Copyright(C)2019 IJAT, [45] 図 7)

補助点群をランダムに発生させることが考えられるが，局所的に点群が偏る箇所では，意図せずマッチングしてしまう高い特徴量が得られる可能性がある．そのような原因での誤マッチングを防ぐため，補助点群単体では 3D マッチングの特徴量が高くないように設計することが必要となる．もし，強遮蔽された検出対象物が存在し，その物体がオクルーダと同程度の大きさであった場合，遮蔽空間の中心付近に検出対象物の一部が存在する可能性が高いと考えられる．そのため，サイズの異なる複数の楕円形状で構成された補助点群を生成することで，遮蔽空間の中心付近に物体の点群が観測されやすく，遮蔽空間の外側では観測されにくい，という性質を表現する（図 5.11，図 5.12）．このような補助点群が，オクルーダが原因で観

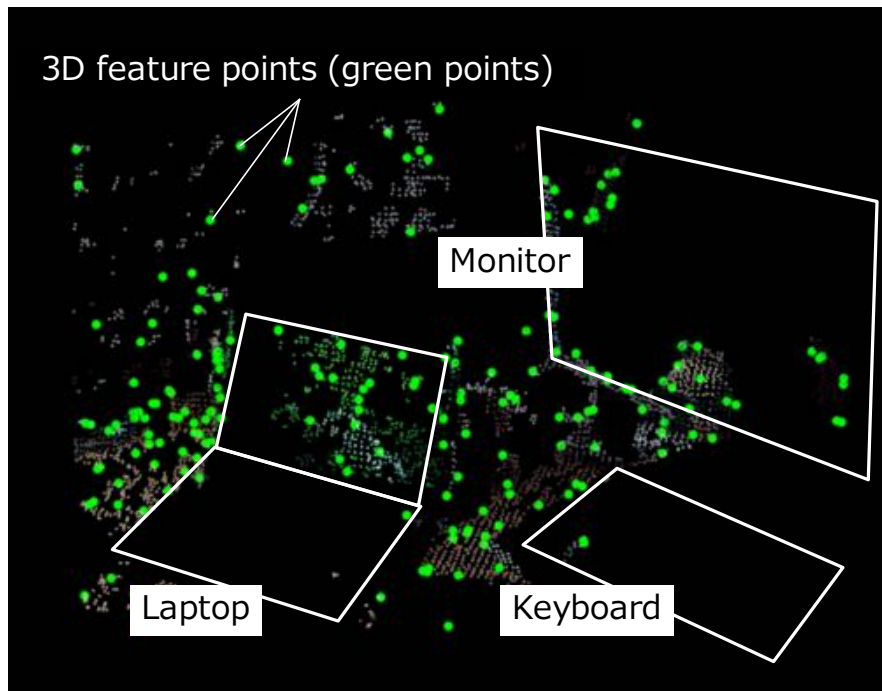


図5.9 3D SIFT の結果（ノイズ点群除去，平面除去，および，ダウンサンプリングの処理後に得られた点群を用いて算出）．（Copyright(C)2019 IJAT, [45] 図 8）

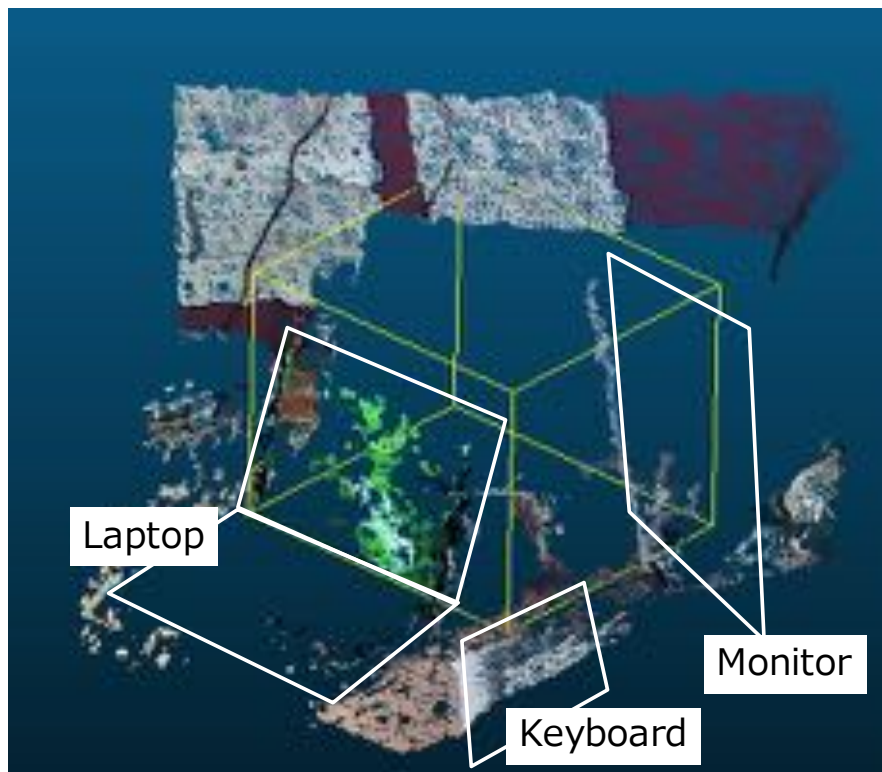


図5.10 机上のラップトップ PC から推定された遮蔽空間（黄色の直方体）．（Copyright(C)2019 IJAT, [45] 図 9）

測されなかった，検出対象物の物体表面を表現する．補助点群の構成を表す式は以下ようになる．

$$\begin{aligned} {}^hX &= (X_1^{k,m}, X_2^{k,m}, X_3^h), \\ h &= 1, \dots, H, k = 1, \dots, K, m = 1, \dots, M, \end{aligned} \quad (5.1)$$

ここで，

$$\begin{aligned} X_1^{k,m} &= \Delta X_1^m \cos \theta_k, \\ X_2^{k,m} &= \Delta X_2^m \sin \theta_k, \\ X_3^h &= ((H - h + 1) * \min X_3 + (h - 1) * \max X_3) / H, \\ \Delta X_1^m &= m(\max X_1 - \min X_1) / (2M), \\ \Delta X_2^m &= m(\max X_2 - \min X_2) / (2M), \\ \theta_k &= 2k\pi / K. \end{aligned}$$

である．

k は円周方向の位置， m は半径方向の位置， h は仮説を立てた平面の位置を表し，この平面は遮蔽空間内に等間隔で配置される．パラメタ K はひとつの楕円円周方向に沿って配置される補助点の数を，パラメタ M は半径方向に配置される異なる大きさの楕円の数を， H は深さ方向（図 5.5 X_3 ）における仮説平面の数を表す． K と M は，ダウンサンプリングされたシーン点群の一部として，補助点群を表現する役割を果たす．全ての補助点群の総数は， $H \times K \times M$ となる．

複数ある仮説は，遮蔽空間の直方体内部に深さ方向にしかれる（図 5.5 X_3 ）．補助点群は，ある深さで，2次元平面上に生成される．ここで，本手法では作業現場において扱う物体は単純形状物体が多数を占めると考え，検出対象物のある一部の形状は平面近似できるものと仮定し，同じ平面上に構成した点群を補助点群として用いる．補助点群は，各仮説において，同一の中心をもつ，異なった大きさの楕円上に存在する．

具体的な事例で，補助点群の機能について説明する．例えばコップを対象物とした場合，対象物は歪曲した物体表面をもっており，一方で補助点群は平面形である．提案法では，この歪曲表面と補助点群平面とが重なる領域では，実際には欠損してしまった形状付近の位置に補助点群があることとなる．マッチング時に，この補助点群を用いた疑似対応点が計算される．この疑似対応点と正対応点と合わせたものを全てをマッチング対応点として用いることで，補助点群を与えない場合に比べ，マッチングに寄与する対応点が仮想的に増加しているため，マッチングスコアが向上し検出が可能となる．このように，検出対象物の物体表面と平面上に発生させた補

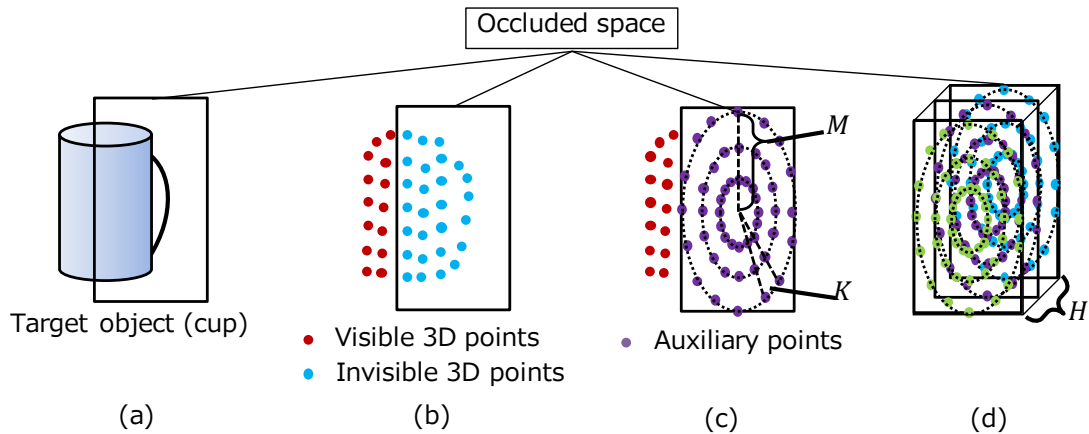


図5.11 遮蔽空間を考慮した補助点群の生成：(a) は強遮蔽された検出対象物。(b) はオクルーダが存在しない環境を想定した理想的な観測点群を表現している。(c) は平面上に配置された補助点群。(d) 生成する全補助点群。(Copyright(C)2019 IJAT, [45] 図 10)

助点群とが、完全には一致しなくとも、歪曲表面と補助点群平面とが重なる領域で生成される疑似対応点を用いることで、マッチングに関する対応点数が増えるため、補助点群が検出の補助として機能する。

5.3.3 物体検出

本手法では、3D マッチング手法の代表的な特徴記述法である、FPFH [50]、および、PPF [51] を用いた。また、これらを単体で用いる手法を既存法とした。マッチングスコアは、シーン点群と検出対象物モデル点群の各点で得られる FPFH および PPF 特徴ベクトルのベクトルペア間の内積値のモデル点数分の総和で得られる。既存法は、シーン点群と検出対象物モデルとを 3D マッチングし、最大マッチングスコアの結果を出力する。一方で、提案法では、オクルーダの近傍に推定した遮蔽空間内に生成した補助点群を加えた、シーン点群に対して、これらの手法を用いて 3D マッチングし、最大マッチングスコアの結果を出力する。このようにすることで、提案法では、検出に関する 妥当な疑似対応点の補填を行うことができ、強遮蔽された対象物を検出することが可能となる。

5.4 実験と評価^{*24}

^{*24} 本節の素材は、International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] の 4 節に一部掲載されたものであり、本節の一部の図は、富士技術出版株式会社の許可を得て、[45] から再利用されています。

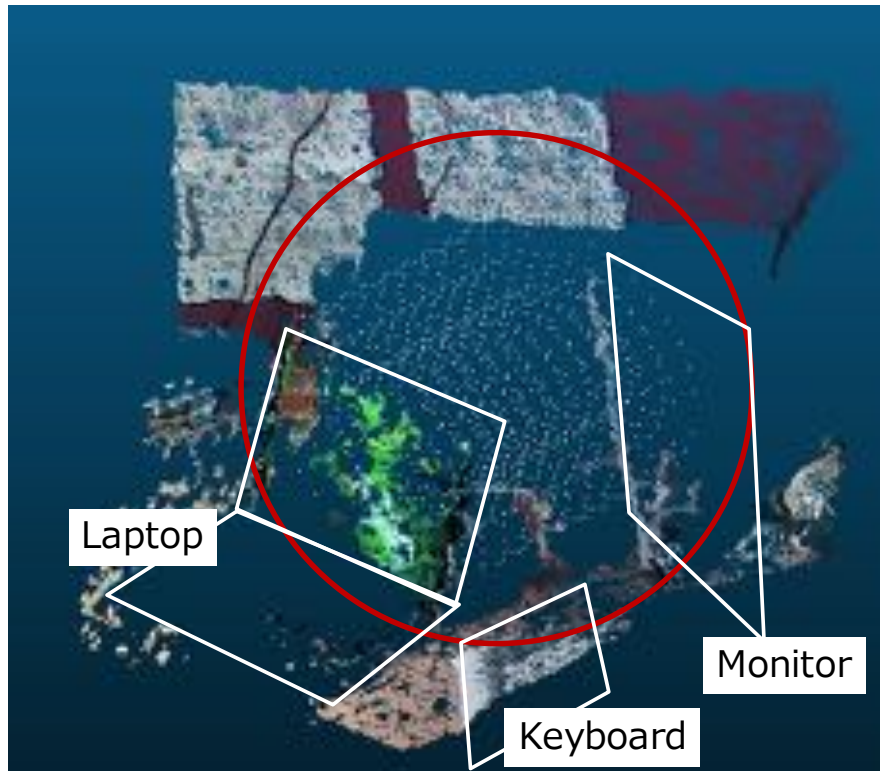


図5.12 ラップトップ PC により推定された遮蔽空間内の補助点群（赤円内）.
(Copyright(C)2019 IJAT, [45] 図 11)

提案法の有効性を検証するため、既存法との比較実験を行なった。まず、提案法における補助点群生成手法の有効性を検証するための実験を行なった。さらに、提案法全体を検証するため、実環境で作った強遮蔽された物体があるシーンを用いて、評価を行なった。

5.4.1 実験条件

実装

小型かつ軽量であり、想定するスタンドアロン HMD にもアタッチできるため、Intel RealSense R200 を、本実験での RGBD センサとして採用した（図 5.13）。検出対象物のサイズは約 6-30 cm であり、検出範囲は 50-150 cm とした。YOLO は、HMD の近くに存在し、遮蔽されていない物体、すなわちオクルーダ候補を検出するために用いる。YOLO 開発者によって配布されている学習済みの重みデータ（Tiny YOLO）を用いて得られる物体検出結果を、実験に使用した [70]。実験をシンプルにするため、任意の物体検出結果をひとつ選び、オクルーダとして用いた。3D 処理ライブラリとして PCL 1.8.0 を使用し、2D 画像処理ライブラリとして OpenCV の 3.1.0 を使用した。また計算機は、Intel Core i7 3.4-GHz の CPU を搭載した PC

を使用した。

精度の評価

本論文では、危険な物体が原因での怪我を未然に防ぐため、いち早く HMD を付けた作業者に危険物がある位置を示す、ということを目的としたアプリの要請から、検出結果の成否の判定方法は、対象物モデルの Bounding Box の最長辺の半分の長さを基準長 C と定義し、正解位置と推定位置のユークリッド距離が C を超えるか、とした。正解位置は、オクルーダを配置せずに対象物のみをシーン中に配置した環境で、3D マッチングを行って得られた推定結果の対象モデルの Bounding Box の中心位置、と定義する。推定位置は、正解位置を求めた対象物を固定し、この対象物の前に仮想的あるいは実際にオクルーダを配置して構成された強遮蔽シーンでの、マッチング推定結果の対象物モデルの Bounding Box の中心位置、と定義する。正解位置への近さは、物体の大きさに依存するため、この判定法を用いた。また提案法は、強遮蔽された検出対象物の近くにマーカを配置するアプリを実現することが目的であり、基準長 C を用いた判定法で十分な評価が可能と考える。

実験方法

本手法の強遮蔽に対する有効性を検証するため、ロバスト性を持たない既存法として、FPFH および PPF を選択した。3D マッチングの前処理として、シーン点群に対して、ノイズ点群除去と平面除去（テーブル面、壁、等）を行い、シーン点群、データセット点群、共にダウンサンプリング処理（0.8 cm のボクセルフィルタ）を実施する。また、FPFH、PPF ともに法線ベクトルを用いて計算されるため、法線ベクトルを求める。ある三次元点に注目したとき、その点の周辺に計測された k_n 個（ k 近傍）の三次元点を用いて最小二乗法により局所近似平面を求め、その平面の法線ベクトルを求める方法で算出する。また、全ての法線ベクトルを求めた後、これら法線ベクトルが、センサ側を正の方向となるように揃える処理をする。法線ベクトルを求める処理を、既存法ではシーン点群のみに対して、提案法ではシーン点群と補助点群の和を取った合成点群に対して行う。これらの処理後のシーン点群全てに対して、FPFH および PPF 特徴量を計算し、random sample consensus (RANSAC) を用いて最大マッチングスコアの検出結果を求める。既存法と提案法、どちらも同じ前処理、シーンで、実験を行った。拡張エッジ領域算出に関するパラメタ D は 20 とした。3D SIFT 算出に関するパラメタ S_{SIFT} , O_{SIFT} , N_{SIFT} , および C_{SIFT} は、それぞれ、0.01, 3, 4, 0.001 とした。補助点群生成に関するパラメタ H , K , および M は、それぞれ、10, 36, 3 とした。

5.4.2 評価

補助点群生成手法の検証実験

検出対象物に強遮蔽が起き、かつ、理想的な遮蔽空間と補助点群が与えられた場合、提案法の補助点群の有効性が得られるかを検証する。ここでは、FPFH のみでマッチングする既存法と、提案法のマッチングプロセスに FPFH を用いる手法とを比較する。理想的な遮蔽空間は、図 5.14 に示すように、任意の遮蔽率で対象物体をカットして得る。点群のノイズの強さと粗密を変化させるため、対象物体を、60-140 cm の間に 20 cm の間隔ごとの 5 か所に配置し実験を行った。また、センサの繰り返し計測によるノイズ変化を考慮し、各位置で 10 回の実験を実施した。よって、合計 50 のデータセットを得て、実験を行った。テンプレートとなるモデルデータは、センサから 60 cm の位置で得た RGBD データを用いた。

図 5.15 に示すように、検出対象物としてボール（上部）を用いた実験では、可視領域が 30% 未満のとき、有効性を確認できた。既存法では、ノイズ除去後に残った物体（机上のテープ片など）に誤マッチングする傾向があったが、提案法では誤マッチングを低減することができた。この結果から、理想的な遮蔽空間と補助点群は、平面近似可能な単純なオブジェクトに対して、有効であると推定される。ここで平面近似可能な単純なオブジェクトとは、定性的に、RGBD センサの深さ方向 X_3 ベクトル（図 5.5）と直交する仮想平面が、対象の物体表面と大きく交わる可能性が高く、凸凹が少ない物体、とする。対象物体を図 5.15 のコップ（下部）にして同様の実験を行い、5% 未満の可視領域の場合、提案法の有効性を確認できた。コップについては図 5.15 のグラフで示されるように、既存法はコップがボールよりも複雑な形状を持っていたため、ロバストに検出できる傾向が見られた。しかし、既存法は非常に可視領域が小さい場合、検出が難しい傾向が見られた。対照的に、提案法は理想的な遮蔽空間内の仮想特徴を用いてマッチング率を向上させることができた。

提案法の実験

単一物体があるシーンでの実験:

実データを用いた実験を行った。はじめに、単一の検出対象物に対して、手で遮蔽を作り出したシーンについて、検出を行う実験を行った。ここでは、FPFH のみでマッチングする既存法と、提案法のマッチングプロセスに FPFH を用いる手法とを比較する。検出対象物のモデルデータは、図 5.14 の左の三次元点群を用いた。既存法では検出対象物とは別の物体にマッチングしてしまうが、提案法では対象物体近傍にマッチングができ、その正解位置と推定位置とのユークリッド距離は 0.5 cm 程度であった（図 5.16）。



図5.13 RGBD センサ : Intel RealSense R200. (Copyright(C)2019 IJAT, [45] 図 12)

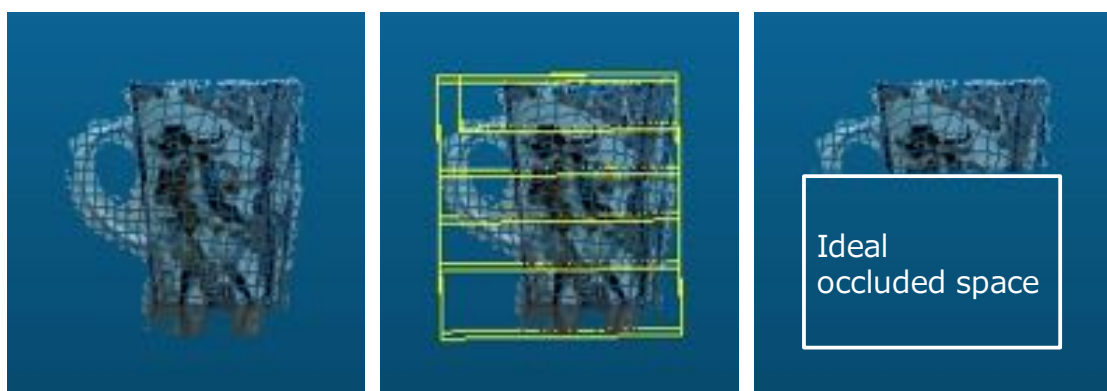


図5.14 検出対象物と理想的な遮蔽空間データ : (左) 検出対象物のコップ. (中央) 遮蔽率を変化させた際の, 理想的な遮蔽空間と検出対象物 (黄色ボックス内が可視領域を表す). (右) 60 % の遮蔽率の時の, 理想的な遮蔽空間と検出対象物. (Copyright(C)2019 IJAT, [45] 図 13)

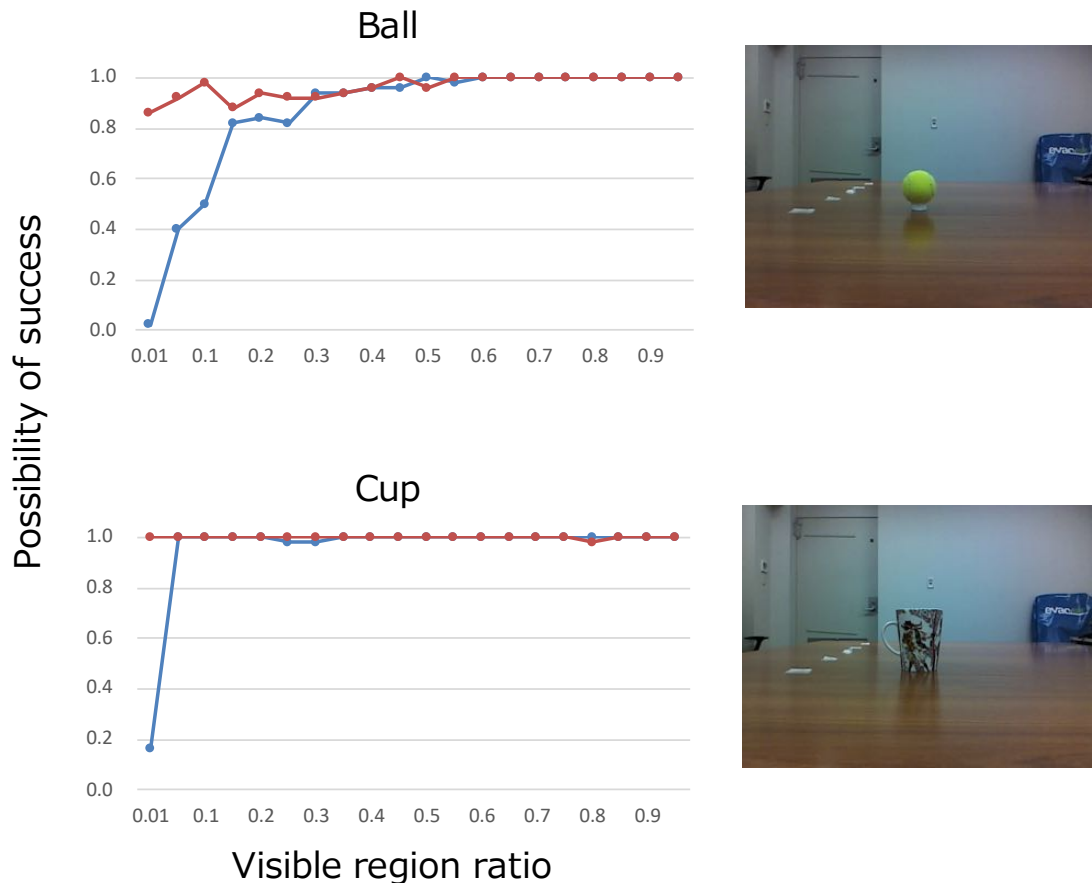


図5.15 遮蔽率を変化させた場合の、マッチング成功率（青：既存法，赤色：提案法）．カメラ画像は，2つの検出対象物，ボールとコップを示してる（センサからの距離が 60 cm の位置で撮影）．潜在的に誤マッチングする可能性のある物体として，物体左側に複数の白テープが存在する．（Copyright(C)2019 IJAT, [45] 図 14）

複数物体があるシーンでの実験:

次に，机の上に複数のオブジェクトがあるシーンにて，実験を行った．ここでは，FPFH のみでマッチングする既存法と，提案法のマッチングプロセスに FPFH を用いる手法とを比較する．また，これと同様に，PPF のみでマッチングする既存法と，提案法のマッチングプロセスに PPF を用いる手法とを比較する．検出対象物には，身のまわりの物体で，コップ，車のおもちゃ，植木の模型，石膏像，いくつかのポーズのデッサン人形とモデルハンド，を用い，それらの物体の全周囲モデルを 3D スキャンすることで得て，検出対象物のデータセットとした（図 5.17）．本研究では，提案法の評価シーンとして，これらの検出対象物がほぼ見えない強遮蔽の条件を再現したデータセットを作成した．検出対象物同士が RGBD センサから見て，重なるようにいくつかのパターンで配置し，50% 以上の遮蔽が起きた対象物のある

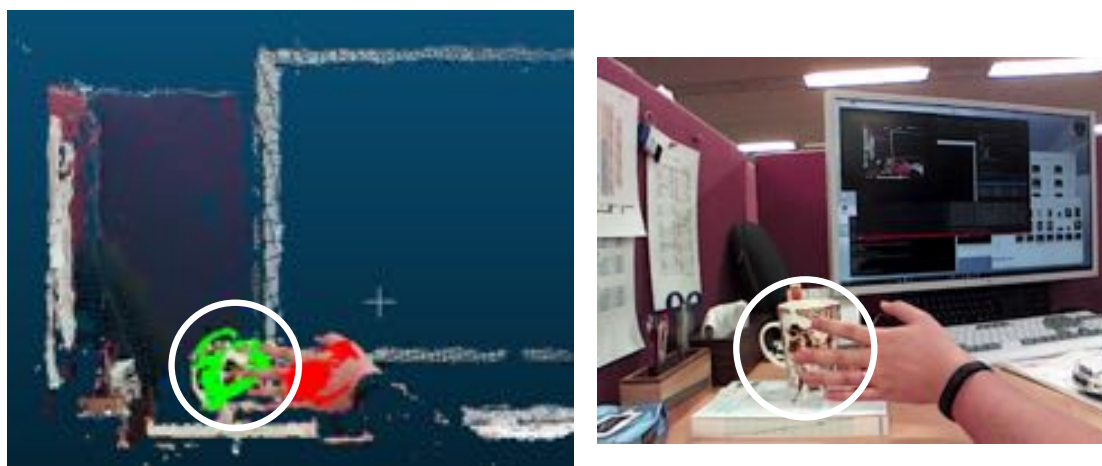


図5.16 実シーンの点群を用いて得られた，マッチング結果（左）：既存法（赤点）と提案法（緑点）と，実シーンのカメラ画像（右）．検出対象物はコップ，オクルーダは人の手，である． おおよその遮蔽率，既存法および提案法による正解位置と推定位置とのユークリッド距離（小数点第2位で四捨五入），閾値 C の数値（小数点第2位で四捨五入）はそれぞれ，約 50 %，13.3 cm，0.5 cm，6.0 cm である．（Copyright(C)2019 IJAT, [45] 図 15）



図5.17 検出対象物のモデルの点群データベース．（Copyright(C)2019 IJAT, [45] 図 16）

シーンを複数，作成した．オクルーダは，コップ，車のおもちゃ，植木の模型，石膏像，人の手，ラップトップ PC，の 6 種類を用い，RGB 画像とデプスデータをもとに生成した点群データのセットを合計 18 種，用意した．画像に対する物体検出アルゴリズム YOLO は，オクルーダ候補となる手前側にある物体（コップ，など）を高速に検出するために用いた．

- 検出口バスト性:

複数の検出対象物候補の中から、強遮蔽された物体の検出が可能かについて検証する。FPFH を用いた場合の実験結果例を、図 5.18, 表 5.1 に示す。既存法では、強遮蔽が原因で誤マッチングをしてしまうシーンでも、提案法では基準長 C 以内で、対象物を検出することができた。このとき、各結果のマッチングスコアは、モデルデータセットの中ではそれぞれ最大値となった。

また、PPF を用いた場合の実験結果例を図 5.19, 表 5.2 に示す。PPF を用いた提案法は、FPFH を用いた提案法と同様に強遮蔽された物体の検出ができ、基準長 C 以内で、対象物を検出することができた。FPFH を用いた実験と同様に、マッチングスコアについてもモデルデータセットの中ではそれぞれ最大値となった。また、PPF 手法が、他のスパース情報ベースの手法に比べ、遮蔽シーンに強いとされる手法であるため、PPF のみでマッチングする既存法でも基準長 C を満たすケースがあったが、提案法のマッチングプロセスに PPF を用いる手法では、正解位置と推定位置とのユークリッド距離がより小さい値で推定が可能であることを確認した。

これらの結果から、提案法は様々なシーン、形状の物体において、強遮蔽された物体の検出が可能であることが確認できた。さらに、提案法は、マッチング手法として用いるスパース情報ベースの手法に依らず、強遮蔽された物体の検出性能向上の効果が期待できることも確認できた。

しかし、提案法はいくつかのシーンでは効果がなかった。例えば、手前側の物体の周りに特徴的なテクスチャをもつ領域がない場合は、遮蔽空間を推定することができないため補助点群を与えることができず、提案法の効果を得ることができない。また、推定した遮蔽空間が、検出対象物よりもはるかに大きな空間であった場合、同様に提案法の効果を得ることが難しい。他に、検出対象物の形状が、平面仮定から大きく逸脱し、平面近似するには適さない形状の場合、補助点群による効果を得ることが難しい。

また、本実験での結果の中で図 5.18 の (c) と、図 5.19 の (a) の結果については、姿勢が 180 度フリップしたような結果が得られている。これらの結果については、厳密なマッチング結果としては誤マッチングであると考えられる。この結果の要因として、可視領域での対象物の特徴的な 3D 形状がある物体領域が欠損してしまい、姿勢を正しく推定することが困難となったため、と考えられる。ただし、シーン中の対象物に関する可視領域の対応点と不可視領域の擬似対応点が表す特徴量あるいはボリウムが対象物モデルと近い場合位置精度を満たす検出はなされた、と考える。

- 処理速度:

処理速度について検証する。3D マッチング処理以外の、提案法による処理時間の増分は、平均 0.3 sec 程度であった。これと 3D マッチング処理の処理時間とを合わせて 1.0 sec 未満で処理を完了できれば、セミリアルタイムな処理が期待できる。FPFH 単体の処理時間が平均約 3.6 sec であったの対し、PPF 単体の処理時間は平

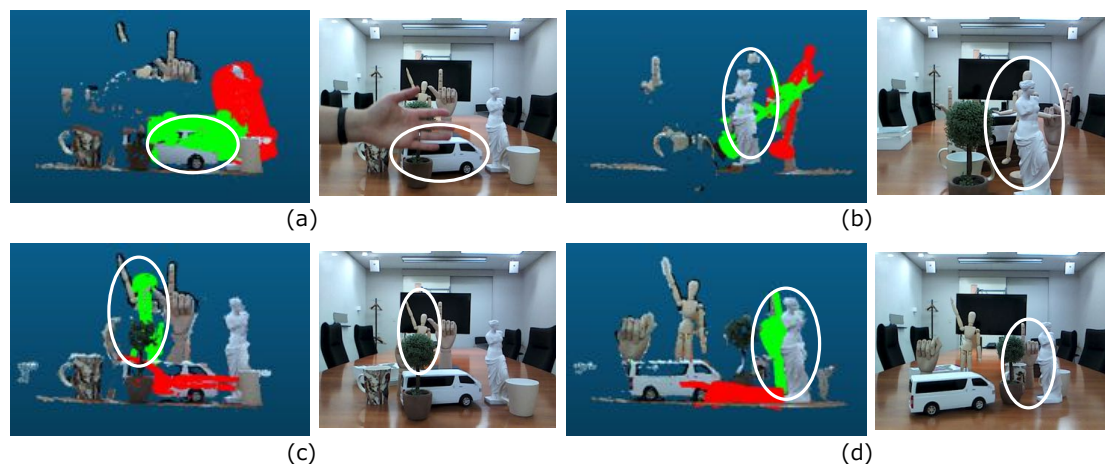


図5.18 FPFH のみでマッチングする既存法と，提案法のマッチングプロセスに FPFH を用いる手法による，様々なシーンでの検出対象物の位置推定結果．白楕円で，検出対象物のおおよその正しい位置を示す．左に，既存法（赤点）と提案法（緑点）の結果を示す．右に，これらの結果を得た際のカメラ画像を示す．
(Copyright(C)2019 IJAT, [45] 図 17)

均約 0.2 sec であった．そのため，提案法のマッチングプロセスに PPF を用いる手法であれば，遮蔽空間推定および補助点群生成に関するすべての処理を加えても，平均約 0.5 sec で処理完了させることができる．これらの結果から，提案法のマッチングプロセスに PPF を用いる手法では，低計算リソースのスタンドアロン HMD においても，セミリアルタイムに，強遮蔽された対象物の検出が期待できる．

表5.1 図 5.18の各実験の結果．既存法および提案法による正解位置と推定位置とのユークリッド距離は小数点第 2 位で四捨五入，閾値 C の数値も同様に小数点第 2 位で四捨五入している．(a) の遮蔽率計算の補足で，黒い窓部とタイヤ部はセンサで点群が計測できなかったため計算から除外している．(Copyright(C)2019 IJAT, [45] 図 17 キャプション)

検出対象・オクルーダ	遮蔽率	既存法の距離	提案法の距離	閾値 C
(a) 車のおもちゃ・人の手	約 50 %	21.5 cm	8.4 cm	13.9 cm
(b) デッサン人形・石膏像	約 75 %	17.0 cm	9.2 cm	16.0 cm
(c) デッサン人形・植木の模型	約 70 %	34.2 cm	2.1 cm	18.3 cm
(d) モデルハンド・植木の模型	約 50 %	29.8 cm	2.5 cm	14.4 cm

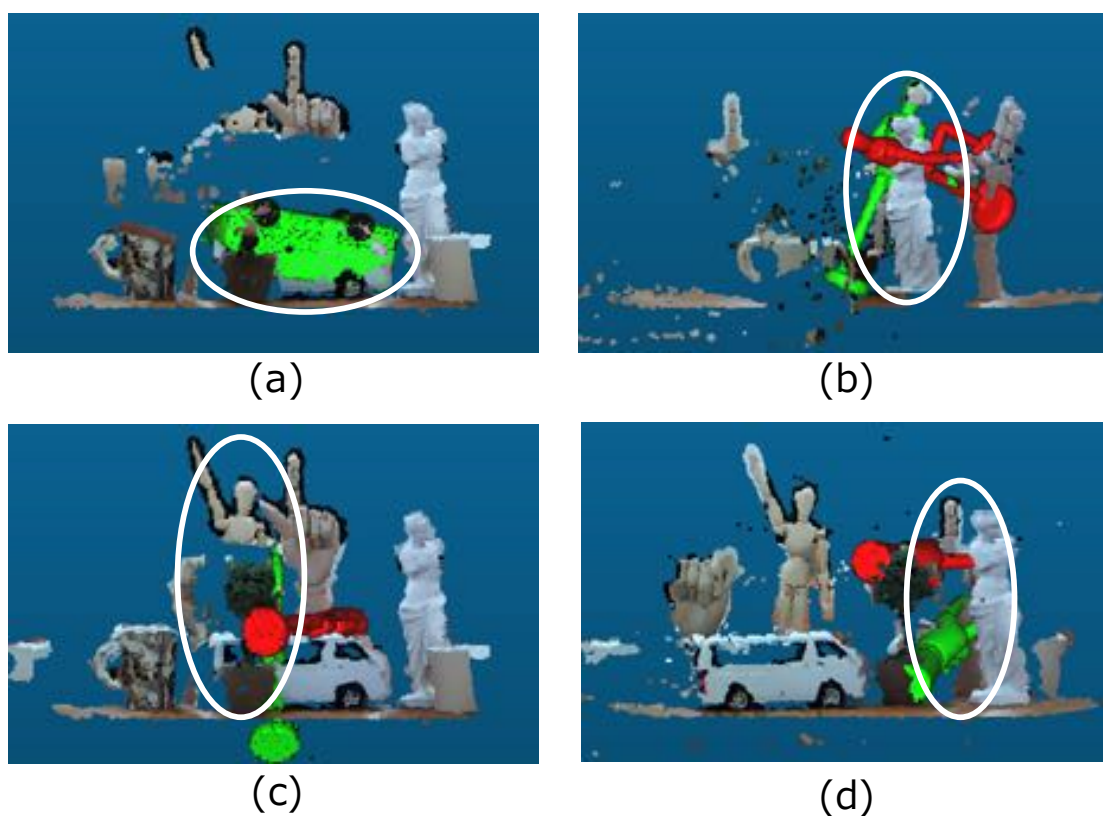


図5.19 図 5.18と同シーンでの、PPF のみでマッチングする既存法と、提案法のマッチングプロセスに PPF を用いる手法による、検出対象物の位置推定結果. 白楕円にて、検出対象物のおおよその正しい位置を示す. 既存法（赤点）と提案法（緑点）の結果を示す. また (a) については、提案法の結果と既存法の結果が同じとなり重なっている. (Copyright(C)2019 IJAT, [45] 図 18)

表5.2 図 5.19の各実験の結果. 既存法および提案法による正解位置と推定位置とのユークリッド距離は小数点第 2 位で四捨五入, 閾値 C の数値も同様に小数点第 2 位で四捨五入している. (a) の遮蔽率計算の補足で, 黒い窓部とタイヤ部はセンサーで点群が計測できなかったため計算から除外している. (Copyright(C)2019 IJAT, [45] 図 18 キャプション)

検出対象・オクルーダ	遮蔽率	既存法の距離	提案法の距離	閾値 C
(a) 車のおもちゃ・人の手	約 50 %	5.5 cm	5.5 cm	13.9 cm
(b) デッサン人形・石膏像	約 75 %	9.9 cm	2.7 cm	16.0 cm
(c) デッサン人形・植木の模型	約 70 %	27.0 cm	17.2 cm	18.3 cm
(d) モデルハンド・植木の模型	約 50 %	13.3 cm	12.7 cm	14.4 cm

5.5 まとめと今後の課題^{*25}

スタンドアロン HMD の低リソース計算機環境を想定し，セミリアルタイムでの処理が期待できる，一人称視点で発生しやすい強遮蔽された物体の検出法を提案した．現場で想定される，他の物体で検出対象物が強遮蔽されてしまうケースを想定した実験を行い，提案法の有効性を確認した．また，処理時間については，セミリアルタイムな処理が期待できることを確認した．

今後の課題は，対象の形状が平面近似できないケースでも検出ができるような改良や，姿勢推定手法とその評価指標について考えていく．また，整備場や工場などの作業現場にて，実環境での評価を実施していく．

^{*25} 本節の素材は，International Journal of Automation Technology Vol.13, No.4, pp.464-474 [45] の 5 節に一部掲載されたものであり，本節の一部の図は，富士技術出版株式会社の許可を得て，[45] から再利用されています．

第 6 章

提案手法 5： 相互反射領域推定法を用いた高速 な金属物の検出^{*3}

6.1 はじめに^{*31}

3次元計測技術は、廃棄物処理やリサイクル処理現場におけるアームロボットで乱雑に置かれた金属破片やプラスチック物、紙類などを自動で把持するために有用な技術である。このような材質物を3次元的に検知することができれば、ロボットで自動的にピッキングをし、材質ごとに自動分類を実施することが可能となる。

2章の背景・目的でも述べたように、工業的な応用価値が高い計測法であるProCamを用いた一般的な計測法は、対象物の鏡面性が高い場合、対象物の凹部の3次元計測が難しい(図6.1)。この課題を解消するため3章の手法では、金属物の高速な形状復元を可能とする手法を提案している。この手法では、ランダムパターンを用いて、直接反射光と相互反射光を分離取得し、複数の幾何制約を用いて金属物の高速かつ高精度な形状復元を実現している。ただし、金属物の検出までは手法が確立していない。

本研究では、対象物の反射特性、光学現象を考慮した制約を与え、直接反射光と

^{*3} 本章は、研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4, Copyright(C)2020 IEICE に掲載された同著者による”相互反射領域推定法を用いた高速な金属物の検出” [71] に基づいています。本章の素材は、研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4 [71] で一部発表されたものであり、本章の一部の図は、電子情報通信学会の許可を得て、[71] から再利用されています。

^{*31} 本節の素材は、研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4 [71] の1節に一部掲載されたものであり、本節の一部の図は、電子情報通信学会の許可を得て、[71] から再利用されています。

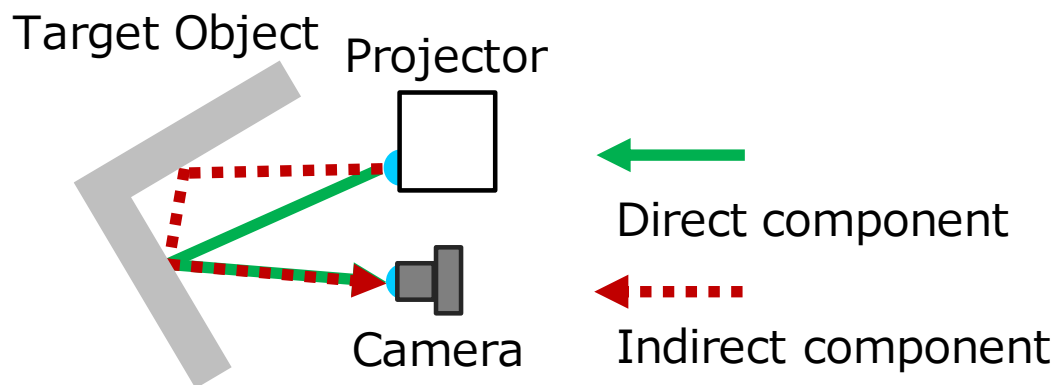


図6.1 直接反射成分と間接反射成分とが重畳したケース．(Copyright(C)2020 IEICE, [71] 図 1)

相互反射光を分離する 3 章の手法を用いて，相互反射領域の推定法による高速な金属物の 3 次元的な検出手法を提案する．投影パターン数が少数でも，分離性の高いランダムパターンを用いて符号化し，反射光を直接反射成分と相互反射成分とに分離取得する．ここで相互反射成分が多く得られる領域を検出することで，カメラ画像上で金属物の存在領域を推定する．そして，鏡面相互反射の影響を抑制し直接反射光による形状復元をして，金属物の存在領域推定結果と合わせることで，3 次元的な金属物の位置を検出可能とする．

6.2 関連研究^{*32}

投影光の変調方式を工夫することで，直接反射光と間接反射光を分離して 3 次元計測する手法が提案されている．Furuse らは [11] を応用し，エピポーラ制約と，M 系列パターンを埋め込んだグレイコード画像により，相互反射を抑制可能な計測を実現している [12]．高周波成分を含み，自己相関の高い M 系列を応用することで直接反射光領域か相互反射光領域かの認識が可能となっており，直接反射光の対応点を用いて 3 次元計測が可能である．M 系列は自己相関がインパルス状となる系列であるため，相互反射を検知することができる．また，グレイコードパターンと M 系列パターンを組み合わせることで，投影枚数の低減を実現している．O' Toole らは，[11] と同様の考えで，投影パターンをエピポーラ線に沿ってランダムにバイナリマスクしたパターンを，DMD の直接制御による高速投影・撮影システムを用いて高速に多数，投影・撮影することで，直接反射光と間接反射光とに反射光を分離

^{*32} 本節の素材は，研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4 [71] の 2 節に一部掲載されたものであり，本節の一部の図は，電子情報通信学会の許可を得て，[71] から再利用されています．

し、直接反射光画像を用いて位相シフト法を適用することで、対象の 3 次元計測が可能であることを示している [13]。また、O' Toole らは CMOS 型センサカメラのローリングシャッタ現象とレーザプロジェクタのスキャンを同期させた装置構成により、多数の投影・撮影をすることと等価な処理を短時間で実現し、エピポーラ制約を満たす対応点の高速な 3 次元計測手法を提案している [22]。

これらの手法は直接反射成分と間接反射成分とを分離して抽出でき、かつ、高速化のための投影数低減のアルゴリズム、あるいは、特殊な ProCam 装置を用いた高速投影・撮影可能なシステムによる工夫がなされている。しかし、反射分離を安定させるために投影枚数が多数、必要である課題があり、高い分離性と高速性とを両立させることが難しい。

一方で、Mirdehghan らは、シーンごとに動的に、投影と撮影枚数を最小にできる投影パターンを生成することを目的とした研究に取り組んでいる [15]。この研究では、プロジェクタの投影範囲とカメラの撮影範囲が重なる計測可能領域の中で、アプリケーションの要請などからヒューリスティックに定めた復元範囲内でのみ 3 次元計測処理を許す Geometry 制約を用いて、シーン最適な最小投影パターン生成について議論している。Geometry 制約で、復元範囲外に対応する画像領域の復号化情報は用いずともよいとため、重畳する可能性のある符号化信号の候補が減る。このため、符号化長を短くすることができ、投影・撮影枚数を最小化し、ProCam による計測の高速化を実現している。この手法は投影枚数を最小化できるものの、鏡面相互反射の影響を抑制することは難しい。

これら相互反射の抑制のために多くの投影枚数を要する課題、および、高速な手法では相互反射の影響の抑制が難しい課題に対し、3 章の手法はランダムパターンと奥行き制約を用いた高速かつ高精度な相互反射する物体の形状復元手法を提案している。しかし、この手法は対象物の形状復元はできるものの、相互反射領域の推定、および、金属物が存在する領域推定までは難しい課題が残されている。

本研究では、この鏡面相互反射の影響を抑制可能な形状復元手法（3 章の手法）で得られる反射光の分離情報を用いて、相互反射領域の推定と金属物の高速な検出ができる、という仮説を立て、検証を行う。

6.3 金属物の形状復元法を用いた、相互反射領域推定による高速な金属物の検出^{*33}

^{*33} 本節の素材は、研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4 [71] の 3 節に一部掲載されたものであり、本節の一部の図は、電子情報通信学会の許可を得て、[71] から再利用されています。

提案法は、3章の形状復元手法で得られる反射光の分離に関する相関情報を用いることで相互反射領域推定と3次元的な金属物の検出を可能とする。この形状復元手法は、少数の投影・撮影枚数で高速に、鏡面相互反射の影響を抑制した形状復元を可能とする手法である。はじめにこの手法の詳細を述べ、その後に相互反射領域推定による金属物の検出法について示す。

6.3.1 少数投影で鏡面相互反射を抑制可能な3次元計測手法

3章の手法は投影パターン画像の各画素に埋め込んだ符号化信号を用いて、鏡面相互反射により発生する直接反射光と間接反射光が重なった点の復号化情報を分離する。このような点を重畳点と定義する。直接反射光と間接反射光の2つの復号化情報のうち、エピソード制約のみでは、どちらが直接反射光の復号化情報かの判別ができない重畳点については、奥行き制約を用いることで直接反射光のみの復号化情報と3次元計測点を得ることができる。

また、この手法では白黒2値のランダムドットパターンを用いる。この理由は、他のパターン画像に比べ、各ドットの符号化信号の直交性を容易に高くでき、各ドットの信号が空間的に独立しているため投影枚数 N の増減を柔軟に行うことができるため、である。このパターン画像列を投影したシーンを撮影したカメラ画像列で座標復号化を行う(図 6.2)。復号化の相関演算には、ZNCC(Zero-mean Normalized Cross-Correlation)を用いる。各カメラ画素で得られる観測信号とエピソード制約を満たす全符号化信号とで ZNCC を算出し、最大の ZNCC 値 $zncc_1$ と第2位の ZNCC 値 $zncc_2$ の信号を求める。これらの比 $r_{zncc} = zncc_1 / zncc_2$ が閾値 th_{zncc} 以上の点を直接反射点、それ以外の点を重畳点として識別する(図 6.3)。

次に各重畳点において、どの二つの符号化信号の和の信号が最も高い ZNCC 値となるかを探索する。このとき、符号化信号の組み合わせ全てで ZNCC 値を求めると計算量が膨大となるため、直交する信号同士であれば式(6.1)が成り立つことを用いて重畳点の信号の ZNCC 値を求める。

$$CC(f(x) + g(x), h(x)) = \frac{1}{\sqrt{2}}(CC(f(x), h(x)) + CC(g(x), h(x))) \quad (6.1)$$

ここで、 CC は相関関数 (ZNCC)、 f, g は直接反射光、相互反射光の符号化信号、 h は重畳点の観測信号を表す。

重畳点に対し投影画像上で直接反射、相互反射の2つの対応点を得られた状態となるため、相互反射の対応点で復元された点については排除する。相互反射光で得られた対応点は図 6.4のように対象物の復元範囲から離れた位置に復元される可能性が高いため、奥行き制約でこれを排除できる。

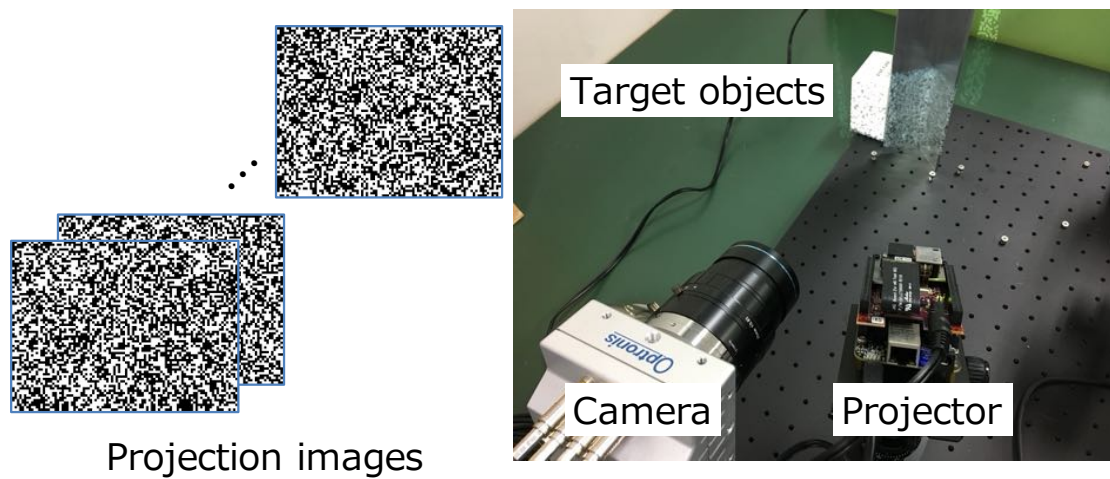


図6.2 左：投影パターン画像と，右：実験環境．左図のようなランダムドットパターンを複数枚 (最大 50 枚)，投影・撮影する． (Copyright(C)2020 IEICE, [71] 図 2)

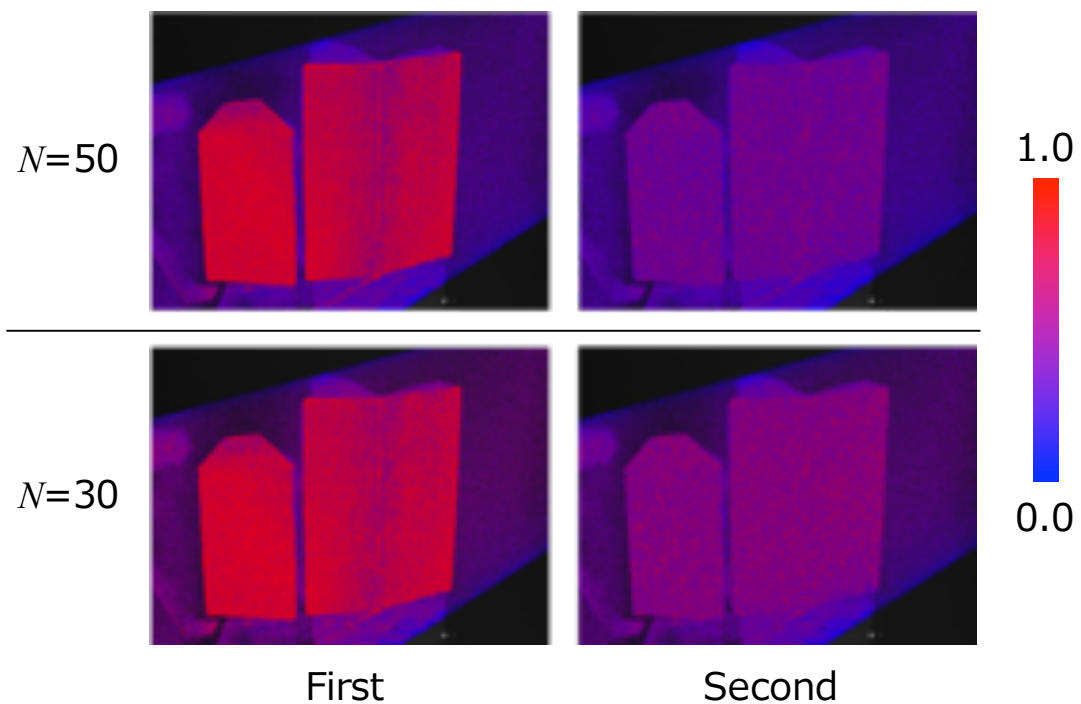


図6.3 ZNCC で得た最大相関 (左) と第 2 位の相関 (右) をマップした図．上段は $N=50$ ，下段は $N=30$ のマップである．投影光が当たる領域では， N が増加すると最大相関と第 2 位の相関の差が大きくなり，復号化の安定性が高くなる様子が確認できる． (Copyright(C)2020 IEICE, [71] 図 3)

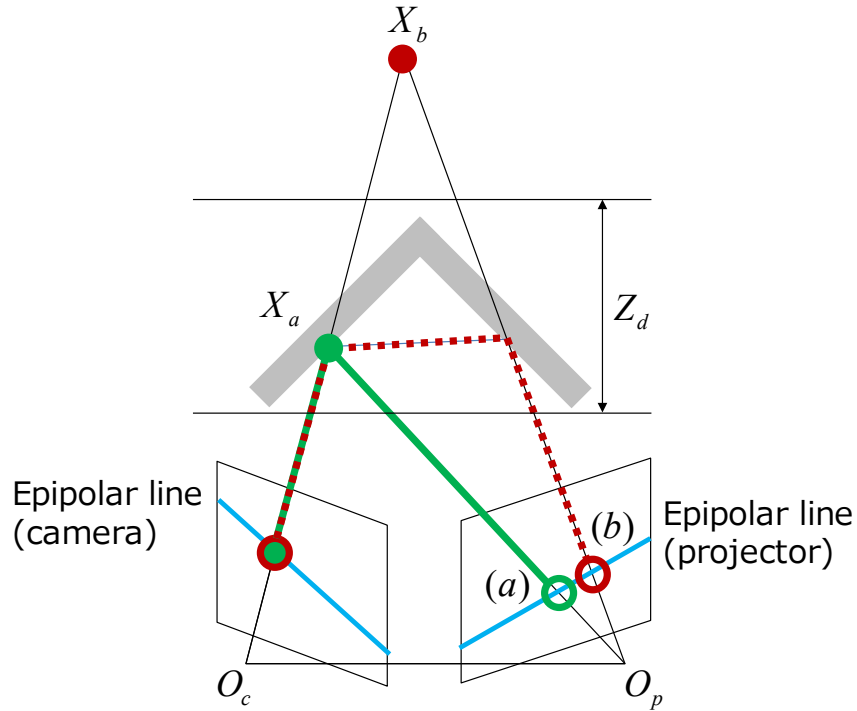


図6.4 奥行き制約の説明図. Z_d : 復元範囲 (奥行き方向), X_a : (a) の対応点を用いた復元点, X_b : (b) の対応点を用いた復元点, O_c : カメラ光学中心, O_p : プロジェクタ光学中心, である. 直接反射光の対応点 (a), および, 鏡面相互反射光の対応点 (b) の2つの対応点のうち, (b) で3次元計測された点は, Z_d 外に復元されるため排除することができる.

6.3.2 相互反射領域推定による高速な金属物の検出

鏡面相互反射する物体の形状復元処理 (3章の手法) で得られている直接反射光と相互反射光の ZNCC 値を用いて, 相互反射領域推定および金属物の検出を行う. 相互反射領域推定は, 形状復元法での直接反射と相互反射の分離を行う際の方法と同様の考え方で, 相互反射領域を推定する. 最大と第2位の ZNCC 値の比を求め, 閾値 th_{zncc}^{inref} 以下の点を重畳点, つまり, 相互反射が起きている点として推定する (図6.3). この推定結果を相互反射推定マップと呼ぶ. ここで, 形状復元の比とは別の閾値を用いる. この理由は, 直接反射を優先して得ることを目的とする形状復元法と, 相互反射を優先して得ることを目的とする相互反射領域推定法とでは, それぞれに優位な感度に変えることが有効と考えたためである.

相互反射推定マップには, 投影ドットの近傍ドットの信号が混入することで誤って相互反射領域と推定される点が発生してしまう. このような点を相互反射推定結果から低減することを目的に, 近傍ドットの信号の影響を受けていると思われる点の推定を行う. 投影ドットの信号, および, 近傍のエピポーラ制約を満たすドット

の信号の和の信号が最も高い ZNCC 値を得ている点を抽出する．この推定結果を近傍ドット光混入マップと呼ぶ．

これら 2 つのマップを用いて、相互反射推定マップの点、かつ、近傍ドット光混入マップの点ではない、という条件を満たす点を、相互反射が起きている点として推定する．この点を抽出したマップを補正相互反射推定マップと呼ぶ．補正相互反射推定マップで得られた点が集中している領域に、相互反射が起こりやすい金属物があると推定できる．このため、ウィンドウ内に存在する点の数の割合 r_{win}^{inref} が閾値 th_{win}^{inref} を超える領域を、画像中での金属物の存在領域として推定することができる．点の数の割合は、以下の式(6.2)で求められる．

$$r_{win}^{inref} = \frac{point^{inref}}{(size_{win})^2} \quad (6.2)$$

ここで、 $point^{inref}$ はウィンドウ内の補正相互反射推定マップの点数、 $size_{win}$ はウィンドウの縦横のサイズ（縦と横のサイズは同じ）、である．

6.4 実験と評価^{*34}

提案法の有効性を確かめるため、鏡面相互反射が起こる環境を想定し実験を行った．実験系は図 6.2 のようにし、計測対象物として鏡面相互反射が起こるアルミ製の L 字アングルと、ランバート反射である白箱を用いた．カメラは Optronis 社のモノクロカメラ：CP70-1HS-M-1900 を用い、解像度は 1024×768 pixels で用いた．プロジェクタは TEXAS INSTRUMENTS 社の DLP2000 を用い、解像度は 480×360 pixels で用いた．キャリブレーションは Zhang のカメラキャリブレーションを用い、ProCam の内部・外部パラメタを求めた [16]．観測の SN 比を上げるため、投影パターン (図 6.2 左) の 1 ドットのサイズを 5×5 pixels とし、高い相関値が得られた場合、その座標を 5×5 の中心とみなすものとした．符号化信号の数は 6912 とした． th_{zncc} は 2.0 とし、ZNCC 値が 0.5 以上の対応点を用いて復元を行った．相互反射領域推定および金属物の検出には、 th_{zncc}^{inref} は 1.1、 th_{win}^{inref} は 0.07、 $size_{win}$ は 128 pixels として実施した．既存法は同 ProCam を用いたグレイコード法と位相シフト法を組み合わせるシステムとし、縦横の縞パターンを、グレイコード法で 12 枚、位相シフト法で 8 枚、計 20 枚用いて対応付けを行い形状復元した．

カメラ画像列を用いて提案法による復号化と形状復元をする実験を行った．座標復号化結果を図 6.5 に示す．3 章の手法を用いた提案法で、エピポーラ制約および奥行き制約を用いることで、相互反射が起きている物体領域においても、良好な対応

^{*34} 本節の素材は、研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.1-4 [71] の 4 節に一部掲載されたものであり、本節の一部の図は、電子情報通信学会の許可を得て、[71] から再利用されています．

付けがなされていることがわかる．次に形状復元結果を図 6.6 に示す．既存法の結果 (左) では相互反射部を計測できないが，提案法の結果 (右) ではその部分が計測できる様子が確認できる．この復元結果に対し，図 6.7 の 2 つのマップから求めた金属物の存在領域を反映した結果を図 6.8 に示す．この結果から，金属物であるアルミ L 字アングルに対してのみ，その物体領域の 60 % 程度のアノテーション (図 6.8 緑部) がなされ，金属物の検出ができることを確認した．

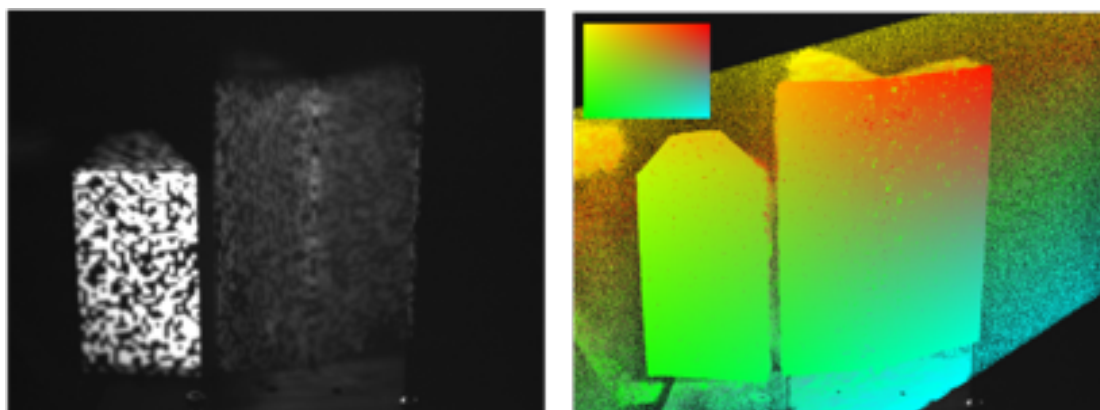
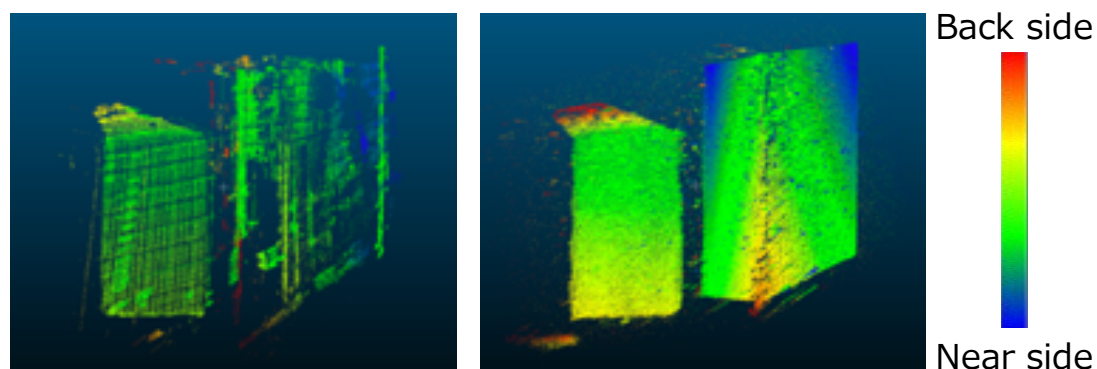


図6.5 $N=30$ の時の提案法による対応付けの結果．座標符号化情報は右画像の左上の図を参照．左：実験シーン (図 6.2 右) のカメラ撮影画像 (左：白箱，右：アルミ L 字アングル)，右：3 章の手法を用いた対応付け結果． (Copyright(C)2020 IEICE, [71] 図 5)



d

図6.6 $N=30$ の時の実験シーン (図 6.2 右) の形状復元結果．左：既存法を用いた結果，右：3 章の手法を用いた結果． (Copyright(C)2020 IEICE, [71] 図 6)

6.5 まとめ^{*35}

^{*35} 本節の素材は，研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9,

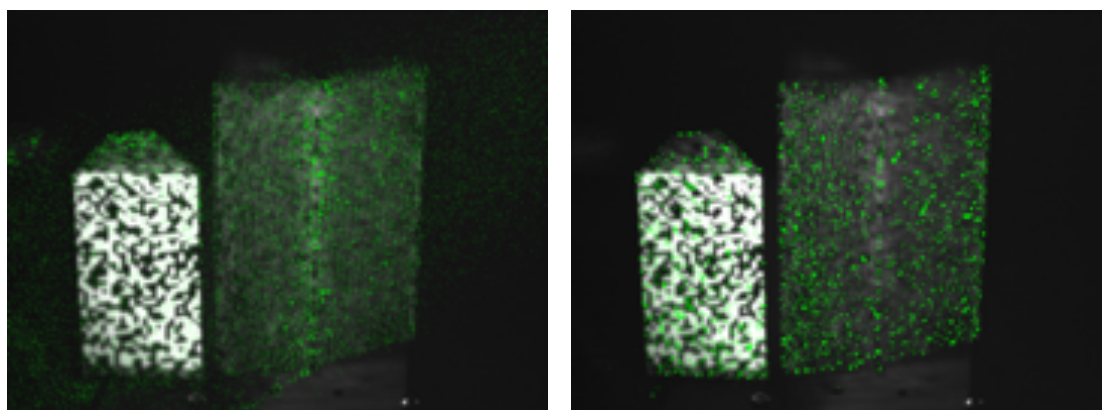


図6.7 $N=30$, 実験シーン (図 6.2右) の, 左: 相互反射推定マップ (相互反射が起きていると推定された画素 (緑点) のマップ), 右: 近傍ドット光混入マップ (近傍の投影ドットの漏れ光により相互反射とみなされた画素 (緑点) のマップ). (Copyright(C)2020 IEICE, [71] 図 7)

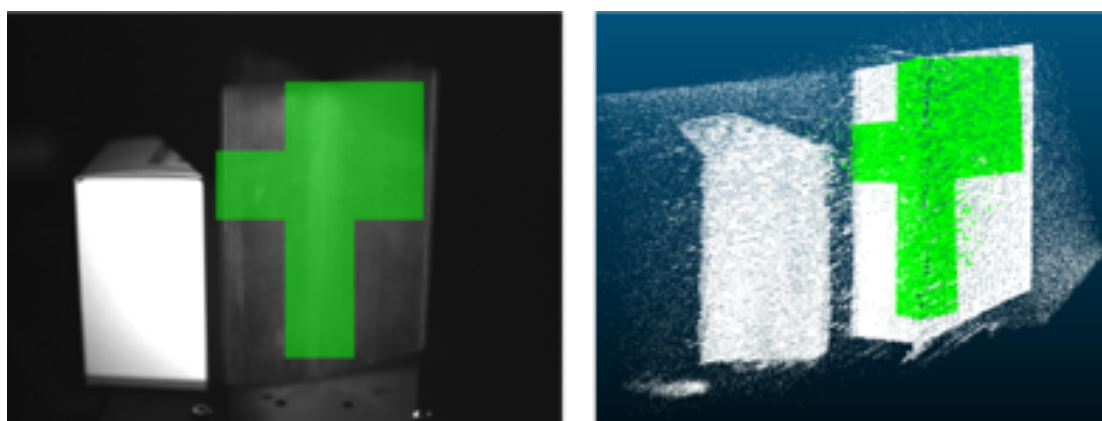


図6.8 $N=30$, 実験シーン (図 6.2右) の, 左: 金属物領域の一部と推測されたカメラ画像上での領域 (半透明の緑色つけた領域), 右: 全体の復元点 (白) に, 金属物領域の一部と推測された復元点 (緑) で重畳した結果. (Copyright(C)2020 IEICE, [71] 図 8)

本研究では, ProCam を用いた少数のパターン投影・撮影で鏡面相互反射が発生する物体を 3 次元計測可能, かつ, 検出が可能な手法を提案した. $N=30$ と少数の投影枚数でも高速に, 相互反射が発生している領域を推定する手法を提案し, 金属物の 3 次元的な検出が可能であることを確認した. 今後は, より多くの材質物の計測と検出が可能な手法について, 取り組んでいく.

pp.1-4 [71] の 5 節に一部掲載されたものであり, 本節の一部の図は, 電子情報通信学会の許可を得て, [71] から再利用されています.

第 7 章

結論

本稿では，コンピュータビジョン技術の応用による特に自動生産分野のファクトリーオートメーションを実現するための実応用の観点で，三次元計測技術，および，三次元認識技術に関する，重要と思われる事項について検討した。

三次元計測においては，計測が対象物性によっては困難な課題について検討した。本論文では，プロジェクタカメラシステムを用いる計測手法について取り組み，特に金属物で起こる鏡面相互反射現象による計測劣化問題について着目して取り組んだ。まず，市販の DLP プロジェクタと高速度カメラのみで，特別な制御システムを用いることなく 122fps 程度で計測が可能な手法を提案した（2章）。さらに，より安価な市販プロジェクタと市販カメラを用いることを想定し，ランダムパターンを用いることでより光学現象にロバストにしつつ，少ない投影枚数で高速に対象物の形状を得ることが可能な手法を示した（3章）。またより高速な対象物の計測を実現しつつ，対象でおこる光学アーチファクトにも対応できるようにするために，工業的には自然と対象の動きが既知であることを幾何制約として用いることで，高速に光学現象にロバストに計測が可能な手法を提案した（4章）。

三次元認識においては，対象の隠れが大きな場合の検出手法，および，金属物が存在すると思われる 3 次元領域の推定技術に取り組んだ。前者は大きな遮蔽によって対象の検出が困難な場合に，オクルーダを推定し，その背後に対象物のボリュームが隠されていると仮定することで検出性能の向上ができることを確認した（5章）。後者は前述した金属物の計測技術を応用し，カメラ画像上で相互反射が多量に発生している領域を推定しこれを 3 次元空間に射影することで，3 次的に金属物が存在する領域の推定が可能であることを確認した（6章）。

これらの研究はこれまで広く研究されている実験室環境では高精度，高速度，高ロバスト性が得られる研究とは志向性が異なり，低コスト，かつ，環境ノイズにロバストな計測・認識を実現した技術であり，実利用環境での計測・認識課題に投入できるレベルの技術である。これらの技術が解決する課題は，ファクトリーオート

メーションなどの産業分野での実用を考える上で必須かつ重要な課題である。また本論文では工場内での利用をメインに説明をしたが、農業、林業、水産業などの第一次産業における動植物の形状計測による生育状態の推定や、土木、建築などの第二次産業における道路のクラックの検査、建築物の歪み推定などにも応用可能であり、広範な課題に対して有効な手段となりうる。

本論文で一部解決している環境ノイズ課題は、工業的な応用を考える上で避けては通れない極めて困難かつ重要な課題である。環境ノイズは対象シーン依存で起こるノイズであり、計測時に起こる望ましくない光学現象や、認識時に障害となる遮蔽、特殊な物体形状などのシーンの物体の状態が関与している。このような実環境で起こるノイズに対し本論文で示した各手法で有効性があるため、現在、本研究の一部技術はすでに実用化開発を進めている段階である。また前述したノイズ以外にも自然界には様々なノイズが存在し、これらノイズに対してもロバストな技術は実環境での利用を目指した産業分野のような業界では解決が必須な課題であると言える。本論文の技術群はそのような耐環境ノイズロバスト性という大きな課題の解決の糸口になると思われる。

さらに本論文で貢献しているデバイス数を抑えコストを小さく抑えることができることは、デバイスの小型化にも繋がり、近年活用されているスマートフォンやドローン、これに続く新しい装置、ロボットへの搭載を可能とし、多様な技術展開のきっかけになりうる。例えば、昆虫や動物を模したロボットが研究開発されつつあるが、特に小さな昆虫であるハチやトンボなどを模したロボットには小型かつ省エネルギーで動作するセンサしか搭載できない。また人体への負荷を減らす目的で人体内部への入れる胃カメラのようなデバイスも活発に研究されつつある。このような小型デバイス、ロボットの発展は今後も続いていくと予想され、これらの新しいハードに付随できるセンサの研究開発は必須である。このため本論文で示した計測・認識技術は省エネルギーに稼働し、かつ、小型化を目指す際にも有効な方針であるため、このような技術発展に大きく貢献すると考えられる。

このような発展性の予想から、多様なデバイス、ロボットへの実装環境を想定し、実環境で耐えうるレベルの安価な装置構成で環境ノイズに対して高いロバスト性をもつ技術という観点で今後も研究に取り組んでいく。具体的には、より多様な反射特性や大きさ・構造をもつ物体があるシーンでも計測を可能とする、あるいは、より複雑に積み重なるように物体が配置されたシーンでも認識を可能とする、といった計測および認識が安価で小型なシステムにて可能となる技術について取り組んでいく。

謝辞

本論文をまとめるにあたり，本学情報科学研究院メディアネットワーク部門情報メディア学分野メディア創生学研究室坂本雄児教授には，本研究の機会を与えてくださり，学士の頃から長年に渡りご指導いただきました．社会人博士課程への入学と在学中の研究に関してもいつも相談にのってください，多大なサポートをいただき円滑に研究活動を行うことができました．ここに深く感謝の意を表します．並びに，貴重なご教示を賜りました本学情報科学研究院メディアネットワーク部門情報メディア学分野言語メディア学研究室荒木健治教授，メディアダイナミクス研究室長谷山美紀教授，情報メディア環境学研究室土橋宜典教授，メディア創生学研究室姜錫助教に深く感謝致します．またジャーナル，および，学会発表原稿について掲載を許可いただきました，情報処理学会誌，富士技術出版株式会社，電子情報通信学会に深く感謝致します．

最後に，本研究を進めるにあたり，多くの助言やご協力を頂いた本研究室の先輩，同輩，後輩の皆さんに深く感謝いたします．特に，社会人博士課程を進めるにあたってさまざまなアドバイスをくださった諏訪東京理科大の山口一弘先生，新社会人の頃から研究開発の進め方について多角的にアドバイスくださった前職株式会社リコーの山口高司さん，中重文宏さん，伊藤仁志さん，現職株式会社デンソーアイティーラボラトリで，メンターとして業務アドバイスや文書作成法について丁寧なご指導をいただきました吉田悠一さん，いつも研究についての鋭いご指摘をくださった安倍満さん，佐藤育郎さん，坂倉義明さん，石川康太さん，関川雄介さん，小澤圭右さん，関東コンピュータビジョン勉強会等にて専門的な知見をくださった木村誠さん，高い専門性から洞察の深いアドバイスをくださった兵庫県立大学の日浦慎作先生，また陰ながらいつも支えてくれた妻の香葉，子の柊洵に感謝致します．

参考文献

- [1] 株式会社リコー, “リコー製ステレオカメラ,”
https://jp.ricoh.com/technology/tech/045_stereocamera (2021 参照).
- [2] Kosuke Sato and Seiji Inokuchi, “Three-dimensional surface measurement by space encoding range imaging,” *Journal of Robotic Systems*, Vol. 2, pp. 27–39 (1985).
- [3] Gu Wang, Fabian Manhardt, Jianzhun Shao, Xiangyang Ji, Nassir Navab, and Federico Tombari, “Self6d:Self-supervised monocular 6d object pose estimation,” *European Conference on Computer Vision* Springer, Springer, pp. 108–125 (2020).
- [4] Shuichi Akizuki and Manabu Hashimoto, “Position and pose recognition of randomly stacked objects using highly observable 3d vector pairs,” *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*, IEEE, pp. 5266–5271 (2014).
- [5] Ioannis Gkioulekas, Anat Levin, Frédo Durand, and Todd Zickler, “Micron-scale light transport decomposition using interferometry,” *ACM Transactions on Graphics (ToG)*, Vol. 34, No. 4, pp. 37 (2015).
- [6] Matthew O’Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich and Kiriakos N Kutulakos, “Temporal frequency probing for 5D transient analysis of global light transport,” *ACM Transactions on Graphics (ToG)*, Vol. 33, No. 4, pp. 87 (2014).
- [7] Tatsunori Taniai and Takanori Maehara, “Neural Photometric Stereo Reconstruction for General Reflectance Surfaces,” *arXiv preprint arXiv:1802.10328* (2018).
- [8] Claus Brenner, Jan Boehm and Jens Guehring, “Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector,” *Videometrics VI*, Vol. 3641, International Society for Optics and Photonics, pp. 164–173 (1998).

- [9] Jens Guehring, Claus Brenner, Jan Böhm and Dieter Fritsch, “Data processing and calibration of a cross-pattern stripe projector,” *International Archives of Photogrammetry and Remote Sensing*, Vol. 33, No. B5/1; PART 5, pp. 327–338 (2000).
- [10] Ryusuke Sagawa, Ryo Furukawa and Hiroshi Kawasaki, “Dense 3D reconstruction from high frame-rate video using a static grid pattern,” *IEEE transactions on pattern analysis and machine intelligence*, Vol. 36, No. 9, pp. 1733–1747 (2014).
- [11] Shree K Nayar, Gurunandan Krishnan, Michael D Grossberg and Ramesh Raskar, “Fast separation of direct and global components of a scene using high frequency illumination,” *ACM Transactions on Graphics (TOG)*, Vol. 25, No. 3, pp. 935–944 (2006).
- [12] Tatsuhiko Furuse, Shinsaku Hiura and Kosuke Sato, “Synchronous detection for robust 3-d shape measurement against interreflection and subsurface scattering,” *International Conference on Image Analysis and Processing*, Springer, pp. 276–285 (2011).
- [13] Matthew O’Toole, John Mather and Kiriakos N Kutulakos, “3d shape and indirect appearance by structured light transport,” *Computer Vision and Pattern Recognition (CVPR)*, 2014 IEEE Conference on, IEEE, pp. 3246–3253 (2014).
- [14] Srinivasa G Narasimhan, Sanjeev J Koppal and Shuntaro Yamazaki, “Temporal dithering of illumination for fast active vision,” *European Conference on Computer Vision*, Springer, pp. 830–844 (2008).
- [15] Parsa Mirdehghan, Wenzheng Chen and Kiriakos N Kutulakos, “Optimal Structured Light a la Carte,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6248–6257 (2018).
- [16] Zhengyou Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, No. 11, pp. 1330–1334 (2000).
- [17] 住吉信一, 吉田悠一, 石川康太, 安倍満, 日浦慎作, “ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速アクティブ3次元計測法,” *情報処理学会論文誌*, Vol. 62, No. 8 (2021).
- [18] Esdras Medeiros, Harish Doraiswamy, Matthew Berger and Claudio T Silva, “Using physically based rendering to benchmark structured light scanners,” *Computer Graphics Forum*, Vol. 33, No. 7, Wiley Online Library, pp. 71–80 (2014).

- [19] Mohit Gupta and Shree K Nayar, “Micro phase shifting,” 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 813–820 (2012).
- [20] Mohit Gupta, Amit Agrawal, Ashok Veeraraghavan and Srinivasa G Narasimhan, “A practical approach to 3D scanning in the presence of interreflections, subsurface scattering and defocus,” International journal of computer vision, Vol. 102, No. 1-3, Springer, pp. 33–55 (2013).
- [21] Mohit Gupta and Nikhil Nakhate, “A geometric perspective on structured light coding,” Proceedings of the European Conference on Computer Vision (ECCV), pp. 87–102 (2018).
- [22] Matthew O’Toole, Supreeth Achar, Srinivasa G Narasimhan and Kiriakos N Kutulakos, “Homogeneous codes for energy-efficient illumination and imaging,” ACM Transactions on Graphics (ToG), Vol. 34, No. 4, pp. 35 (2015).
- [23] Pieter Peers, Dhruv K Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi and Paul Debevec, “Compressive light transport sensing,” ACM Transactions on Graphics (TOG), Vol. 28, No. 1, pp. 1–18 (2009).
- [24] Naoya Chiba and Koichi Hashimoto, “3D measurement by estimating homogeneous light transport (HLT) matrix,” 2017 IEEE International Conference on Mechatronics and Automation (ICMA), IEEE, pp. 1763–1768 (2017).
- [25] Naoya Chiba and Koichi Hashimoto, “Ultra-Fast Multi-Scale Shape Estimation of Light Transport Matrix for Complex Light Reflection Objects,” 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 1–5 (2018).
- [26] Qican Zhang, Renchao Xu, Yihang Liu and Zhenyi Chen, “3D Shape Measurement Based on Digital Speckle Projection and Spatio-Temporal Correlation,” Multidisciplinary Digital Publishing Institute Proceedings, Vol. 2, No. 8, pp. 552 (2018).
- [27] 前田涼汰, 山田憲, 久保尋之, 向川康博, “偏光 Episcan による金属物体の相互反射の除去,” 情報処理学会第 81 回全国大会, Vol. 1, pp. 03 (2019).
- [28] Makoto Matsumoto and Takuji Nishimura, “Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator,” ACM Transactions on Modeling and Computer Simulation (TOMACS), Vol. 8, No. 1, pp. 3–30 (1998).
- [29] Ray A Jarvis, “A perspective on range finding techniques for computer vision,” IEEE Transactions on Pattern Analysis and Machine Intelligence,

- Vol. , No. 2, pp. 122–139 (1983).
- [30] Kurt Konolige, “Projected texture stereo,” 2010 IEEE International Conference on Robotics and Automation, IEEE, pp. 148–155 (2010).
 - [31] William Lohry and Song Zhang, “High-speed absolute three-dimensional shape measurement using three binary dithered patterns,” *Optics express*, Vol. 22, No. 22, pp. 26752–26762 (2014).
 - [32] Nathan Matsuda, Oliver Cossairt and Mohit Gupta, “Mc3d: Motion contrast 3d scanning,” 2015 IEEE International Conference on Computational Photography (ICCP), IEEE, pp. 1–10 (2015).
 - [33] T Leroux, S-H Ieng and Ryad Benosman, “Event-based structured light for depth reconstruction using frequency tagged light patterns,” arXiv preprint arXiv:1811.10771, (2018).
 - [34] Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew Davison, Jörg Conradt, Kostas Daniilidis, et al., “Event-based vision: A survey,” arXiv preprint arXiv:1904.08405 (2019).
 - [35] Hiroshi Kawasaki, Ryo Furukawa, Ryusuke Sagawa and Yasushi Yagi, “Dynamic scene shape reconstruction using a single structured light pattern,” 2008 IEEE conference on computer vision and pattern recognition, Vol. , No. IEEE, pp. 1–8 (2008).
 - [36] Genki Nagamatsu, Ryo Furukawa, Ryusuke Sagawa and Hiroshi Kawasaki, “Single-wavelength and multi-parallel dotted-and solid-lines for dense and robust active 3D reconstruction,” 2019 16th International Conference on Machine Vision Applications (MVA), IEEE, pp. 1–6 (2019).
 - [37] Li Zhang, Brian Curless and Steven M Seitz, “Spacetime stereo: Shape recovery for dynamic scenes,” 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., IEEE, pp. II–367 (2003).
 - [38] Jongwoo Lim, “Optimized projection pattern supplementing stereo systems,” 2009 IEEE International Conference on Robotics and Automation, IEEE, pp. 2823–2829 (2009).
 - [39] Yueyi Zhang, Zhiwei Xiong, Zhe Yang and Feng Wu, “Real-time scalable depth sensing with hybrid structured light illumination,” *IEEE Transactions on Image Processing*, Vol. 23, No. 1, pp. 97–109 (2013).
 - [40] Yuichi Taguchi, Amit Agrawal and Oncel Tuzel, “Motion-aware structured light using spatio-temporal decodable patterns,” *European Conference on*

- Computer Vision, Springer, pp. 832–845 (2012).
- [41] Yuki Shiba, Satoshi Ono, Ryo Furukawa, Shinsaku Hiura and Hiroshi Kawasaki, “Temporal shape super-resolution by intra-frame motion encoding using high-fps structured light,” *Proceedings of the IEEE International Conference on Computer Vision*, pp. 115–123 (2017).
 - [42] Dongxue Li, Tong Jia, Chengdong Wu and Hongyu Wang, “Multi-state objects depth acquisition based on binocular structured light,” *Optics and Lasers in Engineering*, Vol. 121, pp. 521–528 (2019).
 - [43] Jian Wang, Aswin C Sankaranarayanan, Mohit Gupta and Srinivasa G Narasimhan, “Dual structured light 3d using a 1d sensor,” *European Conference on Computer Vision*, Springer, pp. 383–398 (2016).
 - [44] Mark Sheinin, Dinesh N Reddy, Matthew O’ Toole and Srinivasa G Narasimhan, “Diffraction Line Imaging,” *European Conference on Computer Vision*, Springer, pp. 1–16 (2020).
 - [45] Shinichi Sumiyoshi and Yuichi Yoshida, “Estimating 3D Position of Strongly Occluded Object with Semi-Real Time by Using Auxiliary 3D Points in Occluded Space,” *International Journal of Automation Technology*, Vol. 13, No. 4, pp. 464–474 (2019).
 - [46] 機械知覚 & ロボティクスグループ中部大学, “アマゾンピッキングチャレンジ,” http://mprg.jp/research/amazon_picking_challenge_j (2021 参照).
 - [47] Steven Henderson and Steven Feiner, “Exploring the benefits of augmented reality documentation for maintenance and repair,” *IEEE transactions on visualization and computer graphics*, Vol. 17, No. 10, pp. 1355–1368 (2011). <https://doi.org/10.1109/tvcg.2010.245>
 - [48] Anna Syberfeldt, Oscar Danielsson, Magnus Holm and Lihui Wang, “Visual assembling guidance using augmented reality,” *Procedia Manufacturing*, Vol. 1, No. , pp. 98–109 (2015). <https://doi.org/10.1016/j.promfg.2015.09.068>
 - [49] Ren-Jung Chang and Jia-Cheng Jau, “Augmented Reality in Peg-in-Hole Microassembly Operations,” *Int. J. of Automation Technology*, Vol. 10, No. 3 (2016). https://doi.org/10.1299/jsmelem.2015.8._1910-1_
 - [50] Radu Bogdan Rusu, Nico Blodow and Michael Beetz, “Fast point feature histograms (FPFH) for 3D registration,” *Robotics and Automation*, 2009. ICRA’09. IEEE International Conference on, IEEE, pp. 3212–3217 (2009). <https://doi.org/10.1109/robot.2009.5152473>
 - [51] Bertram Drost and Slobodan Ilic, “3d object detection and localization us-

- ing multimodal point pair features,” 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on, IEEE, pp. 9–16 (2012). <https://doi.org/10.1109/3dimpvt.2012.53>
- [52] Frank Michel, Alexander Kirillov, Eric Brachmann, Alexander Krull, Stefan Gumhold, Bogdan Savchynskyy and Carsten Rother, “Global hypothesis generation for 6D object pose estimation,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017). <https://doi.org/10.1109/cvpr.2017.20>
- [53] Stefan Hinterstoisser, Stefan Holzer, Cedric Cagniart, Slobodan Ilic, Kurt Konolige, Nassir Navab and Vincent Lepetit, “Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes,” 2011 international conference on computer vision, IEEE, pp. 858–865 (2011). <https://doi.org/10.1109/iccv.2011.6126326>
- [54] Stefan Hinterstoisser, Vincent Lepetit, Slobodan Ilic, Stefan Holzer, Gary Bradski, Kurt Konolige and Nassir Navab, “Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes,” Asian conference on computer vision, Springer, pp. 548–562 (2012). https://doi.org/10.1007/978-3-642-37331-2_42
- [55] Eric Brachmann, Alexander Krull, Frank Michel, Stefan Gumhold, Jamie Shotton and Carsten Rother, “Learning 6d object pose estimation using 3d object coordinates,” European conference on computer vision, Springer, pp. 536–551 (2014). https://doi.org/10.1007/978-3-319-10605-2_35
- [56] Derek Hoiem, Carsten Rother and John Winn, “3d layoutcrf for multi-view object class recognition and segmentation,” Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on, IEEE, pp. 1–8 (2007). <https://doi.org/10.1109/CVPR.2007.383045>
- [57] Wadim Kehl, Fausto Milletari, Federico Tombari, Slobodan Ilic and Nassir Navab, “Deep learning of local RGB-D patches for 3D object detection and 6D pose estimation,” European Conference on Computer Vision, Springer, pp. 205–220 (2016). https://doi.org/10.1007/978-3-319-46487-9_13
- [58] Andrew E Johnson, “Spin-images: a representation for 3-D surface matching,” (1997).
- [59] Federico Tombari, Samuele Salti and Luigi Di Stefano, “Unique signatures of histograms for local surface description,” European Conference on Computer Vision, Springer, pp. 356–369 (2010). <https://doi.org/10.1007/>

- [60] Federico Tombari, Samuele Salti and Luigi Di Stefano, “A combined texture-shape descriptor for enhanced 3D feature matching,” *Image Processing (ICIP)*, 2011 18th IEEE International Conference on, IEEE, pp. 809–812 (2011). <https://doi.org/10.1109/ICIP.2011.6116679>
- [61] Stefan Hinterstoisser, Vincent Lepetit, Naresh Rajkumar and Kurt Konolige, “Going further with point pair features,” *European Conference on Computer Vision*, Springer, pp. 834–848 (2016). https://doi.org/10.1007/978-3-319-46487-9_51
- [62] Haowen Deng, Tolga Birdal and Slobodan Ilic, “PPF-FoldNet: Unsupervised Learning of Rotation Invariant 3D Local Descriptors,” *European Conference on Computer Vision*, Springer, pp. 620–638 (2018).
- [63] Andrew E. Johnson and Martial Hebert, “Using spin images for efficient object recognition in cluttered 3D scenes,” *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 21, No. 5, pp. 433–449 (1999). <https://doi.org/10.1109/34.765655>
- [64] Shuichi Akizuki and Manabu Hashimoto, “DPN-LRF: A local reference frame for robustly handling density differences and partial occlusions,” *International Symposium on Visual Computing*, Springer, pp. 878–887 (2015). https://doi.org/10.1007/978-3-319-27857-5_78
- [65] Joseph Redmon and Ali Farhadi, “YOLO9000: Better, Faster, Stronger,” *Computer Vision and Pattern Recognition (CVPR)*, 2017 IEEE Conference on, IEEE, pp. 6517–6525 (2017). <https://doi.org/10.1109/CVPR.2017.690>
- [66] Carsten Rother, Vladimir Kolmogorov and Andrew Blake, “Grabcut: Interactive foreground extraction using iterated graph cuts,” *ACM transactions on graphics (TOG)*, Vol. 23, No. 3, ACM, pp. 309–314 (2004). <https://doi.org/10.1145/1015706.1015720>
- [67] John Canny, “A computational approach to edge detection,” *IEEE Transactions on pattern analysis and machine intelligence*, Vol. , No. 6, pp. 679–698 (1986). <https://doi.org/10.1109/TPAMI.1986.4767851>
- [68] David G Lowe, “Local feature view clustering for 3D object recognition,” *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 1, IEEE, pp. I–I (2001). <https://doi.org/10.1109/CVPR.2001.990541>
- [69] David G Lowe, “Distinctive image features from scale-invariant keypoints,”

- International journal of computer vision, Vol. 60, No. 2, pp. 91–110 (2004).
<https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [70] Joseph Redmon and Ali Farhadi, “YOLO: Real-Time Object Detection,” ,
<https://pjreddie.com/darknet/yolo/> (2017 参照).
- [71] 住吉信一, 坂本雄児, “相互反射領域推定法を用いた高速な金属物の検出,” 研究
報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp.
1–4 (2020).

研究業績

論文（主著）

- Shinichi Sumiyoshi and Yuichi Yoshida, “Estimating 3D Position of Strongly Occluded Object with Semi-Real Time by Using Auxiliary 3D Points in Occluded Space,” International Journal of Automation Technology, Vol. 13, No. 4, pp. 464–474 (2019).
- 住吉信一, 吉田悠一, 石川康太, 安倍満, 日浦慎作, “ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速アクティブ3次元計測法,” 情報処理学会論文誌, Vol. 62, No. 8, pp. 1435–1444 (2021).

論文（共著）

- 能勢将樹, 住吉信一, 長谷川史裕, 中重文宏, 内川恵二, “超短焦点プロジェクタによる映像歪みの画質定量化,” 映像情報メディア学会誌, Vol. 70 No. 2, pp. J56–J61 (2016).

国際学会（主著）

- Shinichi Sumiyoshi, “[POSTER] Position Estimation of a Strongly Occluded Object by Using an Auxiliary Point Cloud in Occluded Space,” 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), pp. 194–199 (2017).

国際学会（共著）

- Keisuke Ozawa, Shinichi Sumiyoshi, Yusuke Sekikawa, Keisuke Uto, Yuichi Yoshida, Mitsuru Ambai, “Snapshot multispectral image comple-

tion via self-dictionary transformed tensor nuclear norm minimization with total variation,” IEEE International Conference on Image Processing (ICIP) (2021).

国内学会（主著）

- 住吉信一, 坂本雄児, “部分主成分分析を用いた多視点画像からの任意視点画像合成,” 平成 20 年度電気・情報関係学会北海道支部連合大会, 東海大学札幌キャンパス (2008).
- 住吉信一, 坂本雄児, “凹凸形状を考慮した多視点画像からの任意視点画像合成 (感性とメディア及び一般),” 映像情報メディア学会技術報告 33.44. pp. 31–34 (2009).
- 住吉信一, 坂本雄児, “BSC 法を用いた多視点画像からの 3 次元モデル構築とその評価に関する一検討 (画像符号化, 通信・ストリーム技術, 一般),” 映像情報メディア学会技術報告 33.55, pp. 17–22 (2009).
- 住吉信一, 長谷川史裕, 内山幸央, “ロバスト平面検出法による任意平面へのプロジェクタ投影像の幾何補正 (D-11. 画像工学, 一般セッション),” 電子情報通信学会総合大会, Vol. 2015, No. 2, pp. 30 (2015).
- 住吉信一, 吉田悠一, 安倍満, 佐藤育郎, 石川康太, “鏡面性の強い間接反射光を抑制可能な高速アクティブ 3D 計測手法,” 第 21 回画像の認識・理解シンポジウム / The 21th MIRU(Meeting on Image Recognition and Understanding), PS2-43 (2018).
- 住吉信一 他, “[パネルディスカッション] PRMU 企画セッション: 企業ニーズセッション,” 電子情報通信学会パターン認識・メディア理解研究会 (PRMU 研究会) 12 月 (2018).
- 住吉信一, “[招待講演] デンズなアクティブ 3 次元計測 –基礎と研究開発事例–,” 精密工学会大規模環境の 3 次元計測と認識・モデル化技術専門委員会, 第 34 回定例研究会 4 月 (2019).
- 住吉信一, 吉田悠一, 安倍満, 佐藤育郎, 石川康太, 日浦慎作, “鏡面相互反射光を抑制可能な高速アクティブ 3D 計測手法,” 第 25 回画像センシングシンポジウム / The 25th SSII(Symposium on Sensing via Image Information), IS3-25 (2019).
- 住吉信一, 吉田悠一, 石川康太, 安倍満, 日浦慎作, “少数の投影・撮影で鏡面相互反射光を抑制可能なアクティブ 3 次元計測,” 第 22 回画像の認識・理解シンポジウム / The 22th MIRU(Meeting on Image Recognition and

Understanding), PS2-40 (2019).

- 住吉信一, 安倍満, 吉田悠一, 藤吉弘亘, 日浦慎作, 塚田明宏, “[企業企画イベント] MIRU2019 CollabTech Seminar,” 第 22 回画像の認識・理解シンポジウム / The 22th MIRU(Meeting on Image Recognition and Understanding) (2019).
- 住吉信一 他, “[パネルディスカッション] 企画セッション 1: 企業での研究の紹介, 企業研究者からの学生へのメッセージ,” 電子情報通信学会パターン認識・メディア理解研究会 (PRMU 研究会) 5 月 (2020).
- 住吉信一, “[解説記事] 5 分で分かる! ? 有名論文ナナメ読み: Yasutaka Furukawa and Jean Ponce: Accurate, Dense, and Robust Multi-View Stereopsis,” 情報処理 Vol. 61, No. 9, pp. 992–994 (2020).
- 住吉信一, 坂本雄児, “相互反射領域推定法を用いた高速な金属物の検出,” 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2020, No. 9, pp. 1–4 (2020).

国内学会（共著）

- Keisuke Ozawa, Yusuke Arai, Shinichi Sumiyoshi, Yusuke Sekikawa, Mitsuru Ambai, Akihiro Tsukada, “Multiview-stereo supervised monocular depth estimation,” 第 23 回画像の認識・理解シンポジウム / The 23th MIRU(Meeting on Image Recognition and Understanding), IS3-1-13 (2020).

公開特許（主著）

- Shinichi Sumiyoshi, “Image correction system, image correction method, and computer program product,” US Patent 9,443,294 (2016).
- Shinichi Sumiyoshi, Fumihiro Nakashige and Otoichi Nakata, “Image projection apparatus and presentation system,” US Patent 9,470,966 (2016).
- Shinichi Sumiyoshi, “Image correction system, image correction method, and computer program product,” US Patent 9,443,294 (2016).
- Shinichi Sumiyoshi, Fumihiro Nakashige and Otoichi Nakata, “Information processing device, image projecting system, and computer program,” US Patent App. 14/904,643 (2016).
- Shinichi Sumiyoshi, “Image processing apparatus, image processing

- system, image processing method, and computer program,” US Patent 9,787,960 (2017).
- Shinichi Sumiyoshi and Fumihiko Hasegawa, “Image processing apparatus, projector and projector system including image processing apparatus, image processing method,” US Patent 9,759,548 (2017).
 - Shinichi Sumiyoshi, Naoki Motohashi, Toshifumi Yamaai and Sadao Takahashi, “Image processing apparatus, moving body device control system, computer-readable medium, and image processing method,” US Patent 10,810,749 (2020).
 - Shinichi Sumiyoshi, “Image processing apparatus, imaging device, device control system, frequency distribution image generation method, and recording medium,” US Patent 10,580,155 (2020).

公開特許（共著）

- Fumihiko Hasegawa and Shinichi Sumiyoshi, “Image projecting device, image processing device, image projecting method, and computer-readable recording medium,” US Patent 8,950,872 (2015).
- Fumihiko Hasegawa and Shinichi Sumiyoshi, “Projector and control method thereof,” US Patent 8,953,101 (2015).
- Otoichi Nakata, Fumihiko Nakashige and Shinichi Sumiyoshi, “Image projection device,” US Patent 9,524,696 (2016).
- Otoichi Nakata, Fumihiko Nakashige and Shinichi Sumiyoshi, “Image projection device,” US Patent 9,401,129 (2016).
- Fumihiko Hasegawa and Shinichi Sumiyoshi, “Image processing apparatus, projector and image processing method,” US Patent 9,696,613 (2017).
- Otoichi Nakata, Naoki Motohashi and Shinichi Sumiyoshi, “Image processing apparatus, imaging device, moving object device control system, and image processing method,” US Patent 9,524,696 (2017).
- Sadao Takahashi, Seiya Amano, Hiroyoshi Sekiguchi, Soichiro Yokota and Yuu Yamada, Shinichi Sumiyoshi, “Image processing apparatus, device control system, imaging apparatus, image processing method, and recording medium,” US Patent App. 16/122,425 (2019).
- Sadao Takahashi, Yuji Takahashi, Shinichi Sumiyoshi, Yuu Yamada, Seiya Amano, Soichiro Yokota and Hiroyoshi Sekiguchi, “Object detection de-

- vice, device control system, and medium,” US Patent 10,853,963 (2020).
- Otoichi Nakata, Shinichi Sumiyoshi, Naoki Motohashi, Sadao Takahashi, Soichiro Yokota and Toshifumi Yamaai, “Image processing apparatus, imaging apparatus, and device control system,” US Patent 10,755,116 (2020).
 - Yuu Yamada, Soichiro Yokota, Sadao Takahashi and Shinichi Sumiyoshi, “Processing device, object recognition apparatus, device control system, processing method, and computer-readable recording medium,” US Patent 10,748,014 (2020).
 - Yuu Yamada, Soichiro Yokota, Sadao Takahashi and Shinichi Sumiyoshi, “Image processing device, image processing method, computer-readable recording medium, object recognition device, device control system, and movable body,” US Patent 10,733,459 (2020).
 - Otoichi Nakata, Naoki Motohashi and Shinichi Sumiyoshi, “Image processing apparatus, imaging device, moving object device control system, and image processing method,” US Patent 10,650,271 (2020).

受賞

- 論文：“ランダムパターンを用いた鏡面相互反射の影響を抑制可能な高速アクティブ 3 次元計測法” が、情報処理学会論文誌ジャーナル/JIP 特選論文に選定（2021 年 08 月）

学会活動

- 精密工学会 大規模環境の 3 次元計測と認識・モデル化技術専門委員会 企業委員（2015 年 04 月～）
- PRMU 研究会 専門委員（2018 年 04 月～）
- SSII2021 オーガナイズドセッション部会委員（2020 年 10 月～2021 年 06 月）
- MIRU2022 運営副委員長（2021 年 04 月～）