

2016 年度

博 士 論 文

形状的整合性および配置実現性に基づく
3次元物体認識に関する研究

2016 年 6 月

中京大学 情報科学研究科 情報認知科学専攻

学籍番号 H11401D

秋月 秀一

目次

第 1 章	緒言	7
1.1	研究の背景	7
1.2	研究の目的	9
1.3	本論文の構成	10
第 2 章	ベクトルペアマッチング法に基づく位置姿勢認識	13
2.1	はじめに	13
2.2	ベクトルペア特徴	15
2.3	ベクトルペアを用いた位置姿勢認識	20
2.4	可観測性を考慮したベクトルペア選択による認識の高信頼化	29
2.5	実験結果と考察	33
2.6	ICP アルゴリズムによる位置姿勢推定精度の改善	43
2.7	結論	48
第 3 章	照合戦略の適応的選択に基づく対象物の凹凸形状に依存しない物体仮説生成	51
3.1	はじめに	51
3.2	平面的な見えに対応した認識手法の重要性	52
3.3	Global Reference Frame を用いた姿勢仮説の生成	53
3.4	照合戦略の切り替えを導入した位置姿勢認識	56
3.5	実験と考察	58
3.6	結論	62
第 4 章	点群密度の変化に頑健な物体仮説生成のための 3D キーポイントマッチング	65
4.1	はじめに	65
4.2	従来の LRF とその問題点	66
4.3	DPN-LRF : 点群密度と欠落に頑健な LRF	67
4.4	実験結果と考察	70
4.5	結論	78
第 5 章	形状的整合性と配置実現性に基づく仮説検証型物体認識アルゴリズム	81
5.1	はじめに	81
5.2	形状的整合性と配置実現性に基づく仮説検証	82
5.3	実験結果と考察	85
5.4	結論	90
第 6 章	結言	91
6.1	本論文の成果の要約	91
6.2	今後の展望	92

参考文献

93

发表論文

99

博士論文要旨

形状的整合性および配置実現性に基づく 3次元物体認識に関する研究

学籍番号：H11401D

氏名：秋月秀一

物体の3次元的位置姿勢の認識は、ロボットによる特定物体のハンドリングやトラッキング、AR/MR等、様々な応用先のある基本的な技術である。本論文では、シーン中に存在する複数物体の識別と位置姿勢認識の問題を同時に解くための手法を提案した。

複数の物体が密に配置されている入力シーンに対して、一般的なモデルマッチング法、すなわち、物体の3Dモデルと入力シーンデータを照合させることによって位置姿勢を計算する手法を適用すると認識性能が低下する。これは、正しい位置における一致度合い（スコア）が、隠れ等の影響を受け、他の位置におけるスコアよりも低下することが原因である。従来のモデルマッチング法では、物体モデルと入力シーンの一部、すなわち局所的な情報のみを使ってスコアを算出していたために、この現象を誤認識と判断することができなかった。

そこで、本論文ではこのような偽の位置への誤認識の抑制機能を持つ物体認識アルゴリズムとして、入力シーン中に存在する全ての形状データを同時に解釈することのできる大局的な整合性によるスコア算出手法を提案する。このために、入力シーン中に存在する各物体の見え方（物体仮説）をモデルマッチング法によって算出し、これらを組み合わせることによって仮のシーンの解釈結果（シーン仮説）を生成し、入力シーンとのスコアを算出する。このスコアを最大化する物体仮説を選択する試行を組み合わせ最適化問題として定式化し、仮説検証法のフレームワークとして物体仮説の組み合わせを探索する。このとき、入力シーンとシーン仮説の一致度だけを評価するだけでなく、シーン仮説中の物体仮説の配置の実現可能性も考慮することによって、多数の物体が密に配置されたシーンにおいても、高信頼にそれぞれの物体を認識できることを示す。

仮説検証法では、モデルマッチングによって得られた物体候補を取捨選択しながら仮説シーンを生成するため、誤った候補を棄却することはできるが、未検出の候補を新たに生成する機能が無い。すなわち、モデルマッチングの時点で確実にすべての物体候補を生成しておくことが重要である。

このために、本論文では入力シーンデータの性質（物体の凹凸密度や、外乱）に合わせた新たなモデルマッチング法についても提案した。凹凸部分の多い物体、平面的な形状の物体に有効なモデルマッチング法であるVPM法、GRF法を提案した。さらに、これらの相補的な性質を持ったマッチング方式を適応的に組み合わせることによって、凹凸密度に依存しない物体認識を実現した。

以上のモデルマッチング法を物体候補検出器とし、これを提案した仮説検証型アルゴリズムの前段に配置することによって、複数物体の位置姿勢を同時認識するためのフレームワークを作成した。さらに、3D物体認識のための公開データベースを利用して提案手法と従来の仮説検証型アルゴリズムの性能を比較し、認識信頼性および処理時間の両観点において優位であることを実証した。

第1章 緒言

1.1 研究の背景

1.1.1 製造分野における3次元物体認識

物体の3次元的位置姿勢の認識は、ロボットによる特定物体のハンドリングやトラッキング、AR/MR等、様々な応用先のある基本的な技術である。

産業分野における主要な応用事例としては、ロボットアームによるばら積み部品のピッキングシステムがある。ばら積み状態で供給された部品取り出しの自動化は、製造のほとんどの工程が自動化された現在においても、最後に残る課題とされているため、この技術の完成が生産ラインの完全自動化の鍵を握ると考えられている。このシステムでは、ロボットアームによって部品を取り出すために、対象の部品の位置姿勢をセンシングする必要がある。一般的には、ロボットアームの手先やシステム上部に設置されたレンジセンサによってばら積み状態の部品を撮影し、画像処理技術によってそれぞれの位置姿勢を認識するという方法が採られる。

物体の位置姿勢は、CAD等の3Dデータとしてモデリングした物体モデル（本論文では、このデータを物体モデル、あるいは単にモデルと呼ぶ。）と、レンジセンサによって得られた距離データとを照合する、いわゆるモデルマッチング法によって認識される。したがって、このためには実物の対象物と物体モデルが必要である。部品の多くはCADソフトによって設計されるため、製造分野においては、この条件が成り立っていることが一般的である。このことから、モデルマッチング法の適用先として、ばら積み部品の認識システムは非常に相性が良く、これまでに多くの手法 [34, 27, 61, 9] が研究されてきた。

他の位置姿勢認識のアプローチとして、モデルベース法と対極の概念として語られることの多い、アピランスペース手法 [31] がある。この手法では、通常の濃淡カメラによって撮影した対象物の見かけの画像を物体の姿勢と関連付けた状態で多数用意しておき、入力画像とマッチングする。対象物の位置姿勢認識の問題を画像照合問題として置き換えていることが特徴であり、レンジファインダを用意しなくてもよいという利点がある。認識精度を高めるためには必然的に膨大な「見え」のデータを保持しておく必要があるため、画像データを低次元の特徴量に変換しておくことが一般的である。しかしながら、現在においてはレンジファインダの低価格化や3Dデータの操作手段の整備によって、モデルマッチング法の方が一般的に用いられるようになってきた。

モデルマッチング法は主流の物体認識手法ではあるが、実用化のためには超えるべきハードルが今なお残されている。

- (1) ロボットの動作スピードに認識の処理時間が追いついていない。
- (2) 信頼性良く認識することのできる物体形状に制約がある。

(1) に関しては、部品のピック&プレースのロボットシステムを例に挙げると、ロボットの動作サイクルはおよそ2秒である。ロボットの動作効率を考慮すると、この時間以内の認識が目標値であるといえる

が、コンスタントにこの速度を達成できるアルゴリズムは存在していない。

(2) に関しては、多くのモデルマッチング法は物体から特徴的な形状（またはその代表点としてのキーポイント）を検出し、マッチングに利用している。しかしながら、工業部品は人工物であるため、平面的な形状で構成されていることが多く、このような形状の物体に対しては認識に十分なキーポイントが得られず、結果として誤認識を引き起こすという問題点がある。すなわち、物体の凹凸密度に対する汎用性が十分でないと考えられる。

1.1.2 サービス分野における 3 次元物体認識

近年では、3次元データの入出力装置としてのレンジセンサや 3D プリンタの低価格化に伴い、これらのデバイスが普及してきた。これにより、日用品をスキャンすることによって 3D モデルとしてデータ化することや [54, 50]、あるいは反対に、3D モデルとして存在する物体を実物体としてプリントアウトすることが容易になってきた。したがって、身の回りにある物体はすべて 3次元モデルを持っているという仮定は現実になりつつあり、モデルマッチング法の応用先は今後さらに大きく広がると考えられる。

実際に、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) [58] が公表したロボット産業の将来市場予測では、2015 年が 1.6 兆円の市場規模に対して、2035 年には 9.7 兆円まで成長することが予想されている。この資料では、ロボット産業を 4つの分野（サービス分野、農林水産分野、ロボットテクノロジー製品、製造分野）に分けているが、特筆すべきことは、現在では製造分野が中心の市場であることに対して、10年後にはサービス分野が最も大きな市場に成長するという点である。サービス分野におけるロボットは、物流自動化ロボットや生活支援ロボットがある。今後のロボット産業の市場の転換に則して、これらのロボットのためのビジョンシステムを開発することは非常に有意義であり、さらなる市場の活性化に寄与できるものと考えられる。

ところで、サービス分野のためのロボットビジョンシステムの技術開発の盛り上がりは、わが国だけでなく世界的な潮流となっている。

物流に関していえば、米国 Amazon Robotics 社、インドの GreyOrange 社、そしてノルウェーの Hatteland 社は、倉庫内を自由に動き回ることのできるロボットを開発し、商品ピッキングを担当する作業員の目の前まで商品の入った箱を持ってくるシステムを実用化している。ここでは、作業員による商品ピッキングが残された人手作業となっており、これを自動化するロボットシステムが望まれている。米国 Amazon.com, Inc は、この物流ロボットシステムのコンペティション (Amazon Picking Challenge) [7] を開催した。Amazon Picking Challenge では、Amazon が倉庫で実際に利用している棚から、指示された商品を正確に取り出すためのロボットシステムの性能が競われる。第一回大会は IEEE ICRA2015 で開催され、今年 2016 年にも RoboCup にて第二回大会が開催されることが決まっており、継続的な技術開発が進められている。

生活支援に関していえば、Willow Garage が開発したパートナーロボット PR2 は非常に有名である。他にも同様のロボットは Fraunhofer IPA が開発した Care-O-bot 4 や、日本においてはソフトバンクロボティクス (株) の Pepper や、トヨタ自動車 (株) の HSR 等、生活支援を目的としたさまざまなロボットが開発されている。学術分野では、RoboCup@Home[37] と呼ばれる生活支援タスクをこなすロボットのコンペティションが継続して開催されている。このことからロボットによる生活支援は注目度の高いタスクといえる。

物流用途や生活支援用途を含むサービス分野における物体認識タスクの難しさとしては、認識環境を事前に整備しにくいということが挙げられる。例えば、部品ピッキングの例を考えてみると、ワーキング

ディスタンスはあらかじめ固定して設定しておくことが可能であるし、部品の配置状態をある程度認識しやすいうようにコントロールすることが可能である。サービス分野への応用時には、これらの条件を仮定しないビジョンシステムが必要であると考えられる。したがって、解決しなければならない課題は以下の2つであると考えられる。

- (3) 形状データ（3次元点群）の密度変化に対する頑健性が不十分。
- (4) 多種類の物体が密に配置されている場合には認識性能が低下する。

(3) に関しては、ワーキングディスタンスが固定できないことが前提となっている状況における認識では、あらかじめ用意しておいた物体モデルのデータ密度と入力シーン中の計測データの密度が異なっているということが想定される。このような状況では、従来のモデルマッチング法では局所参照座標系（LRF）と呼ばれるキーポイントを原点とした3次元の局所的な座標系の方向の再現性が劣化する。多くの3次元特徴量はLRFを基準として記述されるため、特徴量の再現性も同時に劣化し、結果として誤認識を引き起こす。

(4) に関しては、物流ロボットや生活支援ロボットでは、多種類の物体を取り扱うことが望まれる。従来のモデルマッチング法では、物体モデルと距離データとして得られた入力シーン間の形状的な一致度合い（スコア）を基準に物体を検出するが、物体の配置状態によっては、対象物が他の物体に隠されることによって一部しか見えていない時に得られるスコアよりも、複数の対象物以外の物体によって構成される偽の面がもつスコアの方が高くなり、誤認識を引き起こす場合がある。モデル単体としてのスコアが高い以上、モデルマッチング法にはこの誤認識を検知することができないため、このことはモデルマッチング法のアルゴリズムの限界といえる。

1.2 研究の目的

本研究の目的は、製造業、および物流・生活支援を志向した物体認識システムのためのモデルマッチング法の機能強化である。具体的には、前節までに説明した、

- (1) ロボットの動作スピードに認識の処理時間が追いついていない。
- (2) 信頼性良く認識することのできる物体形状に制約がある。
- (3) 形状データ（3次元点群）の密度変化に対する頑健性が不十分。
- (4) 多種類の物体が密に配置されている場合には認識性能が低下する。

上記4つのモデルマッチング法が抱える問題点を解決することである。

(1), (2) は製造分野における実問題である。ロボットの動作サイクル以内の物体認識を実現するための高速モデルマッチング法を提案する。また、物体形状の凹凸密度に依存しないモデルマッチング手法を開発することにより、認識可能な形状に対する制約を取り去る。

(3), (4) はロボット産業がサービス分野に展開したときに起こりうる問題である。

(3) の状況のように、マッチング対象の点群間の密度が異なっている場合には、これらの外乱に対する頑健性が重要である。そこで、局所参照座標系 LRF の生成過程において、上記の外乱に対する補正機能を追加することによって、点群密度変化とオクルージョンによる欠落に対する頑健性を向上させる。提案した LRF を認識に利用することによって、キーポイントマッチングの信頼性が向上することを示す。

(4) においては、従来のモデルマッチング法は物体モデルと入力シーンの一部、すなわち局所的な情報のみを使って一致度合い（スコア）を算出していたために、入力シーン中に多種類の物体が密に配置され

ている時には、誤った領域とのスコアの方が高くなる場合があり、誤認識が誘発されることが問題点であった。そこで、本論文では大局的な整合性、すなわち入力シーン中に存在する全ての形状データを同時に解釈することのできるスコア計算方法を導入する。このために、入力シーン中に存在する各物体の見え方（物体仮説）をモデルマッチング法によって算出し、これらを組み合わせることによって仮のシーンの解釈結果（シーン仮説）を生成し、入力シーンとのスコアを算出する。このスコアを最大化する物体仮説を選択する試行を組み合わせ最適化問題として定式化し、最適な物体仮説の組み合わせを探索する。このとき、入力シーンとシーン仮説の一致度だけを評価するだけでなく、シーン仮説中の物体仮説の配置の実現可能性も考慮することによって、多数の物体が密に配置されたシーンにおいても、信頼性高く各々の物体を認識することができることを示す。

1.3 本論文の構成

本論文の構成を Fig.1.1 に示す。2 章から 5 章まではそれぞれ、上記に示した 4 つの問題に対するソリューションである。

第 2 章では、ロボットの動作サイクル時間内での対象物の位置姿勢を認識する手法である、ベクトルペアマッチング (VPM) 法について述べる。VPM 法のコンセプトは、マッチングが高速な低次元特徴量をごく少数に厳選し、認識に利用することである。まず、提案するベクトルペア特徴について説明し、認識率・処理速度を向上させるための戦略的ベクトルペア選択について述べる。

第 3 章では平面的な形状の物体の姿勢認識に特化した Global Reference Frame (GRF) について述べる。従来のキーポイントベース認識における平面的な形状の物体の認識の難しさについて述べ、これを解決するための姿勢認識アルゴリズムを説明する。さらに、入力シーンデータの凹凸形状に合わせて、VPM 法と GRF を適応的に切り替える、照合戦略選択による姿勢認識手法について述べる。

第 4 章では照合対象の点群同士の密度が異なっていたとしても頑健にマッチング可能な 3D キーポイントマッチング手法に関して説明する。キーポイントマッチングの性能は局所参照座標系の再現性に依存することを説明し、点群密度変化に頑健な局所参照座標系の生成アルゴリズムを説明する。

第 5 章では、モデルマッチングによって得られた物体仮説を用いる複数物体の同時認識アルゴリズムに関して説明する。まず、局所的な情報のみをスコア計算に利用するモデルマッチング法の問題点を明らかにし、大局的な情報を利用することの妥当性を説明する。そして、認識信頼性を向上させるための物体配置の実現性を定義し、実験によりその有効性を示す。

第 6 章にて本研究成果をまとめる。

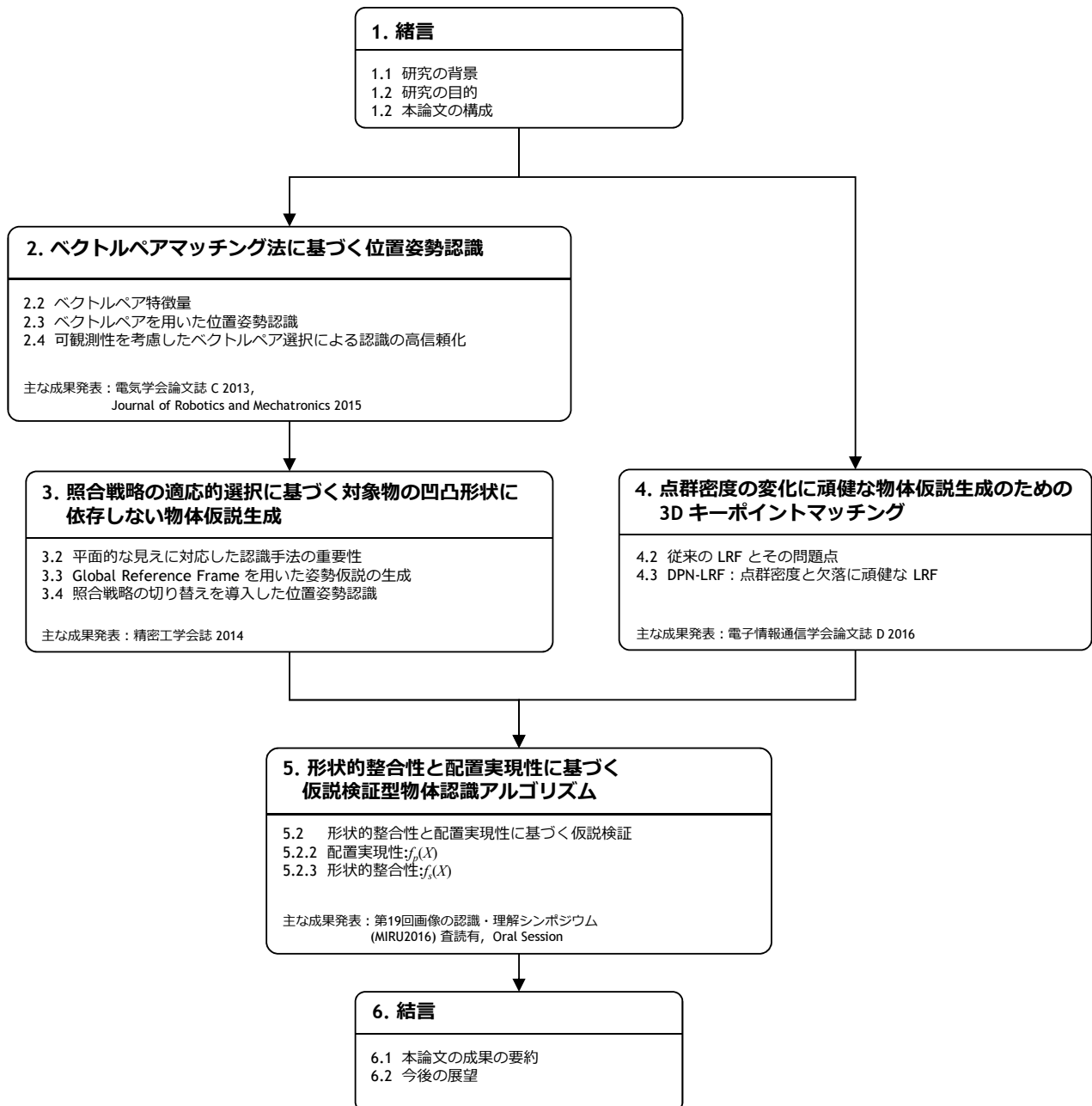


Fig.1.1 Structure of this thesis.

第2章 ベクトルペアマッチング法に基づく位置姿勢認識

シーン中に存在する物体の位置姿勢の候補（物体候補）を高速かつ信頼性高く検出することはモデルマッチング法の実用化のための基本的な性能目標である。本章では、物体の凹凸部分を手掛かりとして、高速に物体候補を生成するためのベクトルペアマッチング（VPM）法に関して述べる。なお、VPM法はモデルマッチング法的一种であるため、物体の識別機能を有さない。したがって、VPMの性能実証方法として、識別の必要がない、単一部分のばら積みピッキングをケーススタディとし、認識性能を評価する。さらに、実際の産業用ロボットアームを用いて、認識した物体をビンから取り出すことが可能なことを示す。

本論文では、「物体モデル」という用語を多用する。これは、Fig. 2.1 に示すような認識対象物の形状データのことを指す。(a) は対象物の3次元形状データの外観である。見やすさのために、面を表示している。(b) は(a)の点群データの例である。実際の認識処理には、このように物体モデルの表面に配置された各点に3次元の座標群（点群データ）を扱う。(c) は物体モデルに対して法線ベクトルを白い線として表示した例である。離散点群からの法線ベクトルの算出方法に関しては本章にて述べる。

2.1 はじめに

物体の位置姿勢は、自由度が3である。標定座標系は通常は地上座標系。物体モデルの姿勢は、自由度が6である。

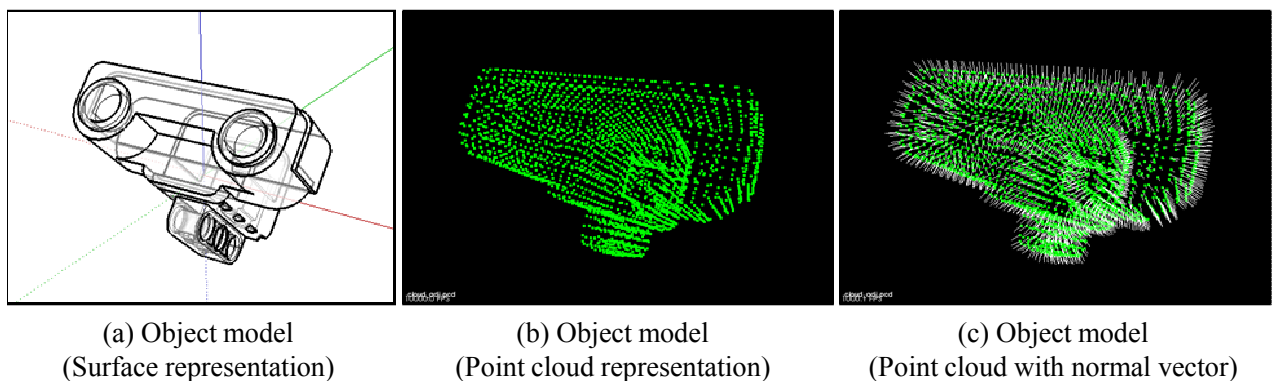


Fig.2.1 Representation of the term 'object model'.

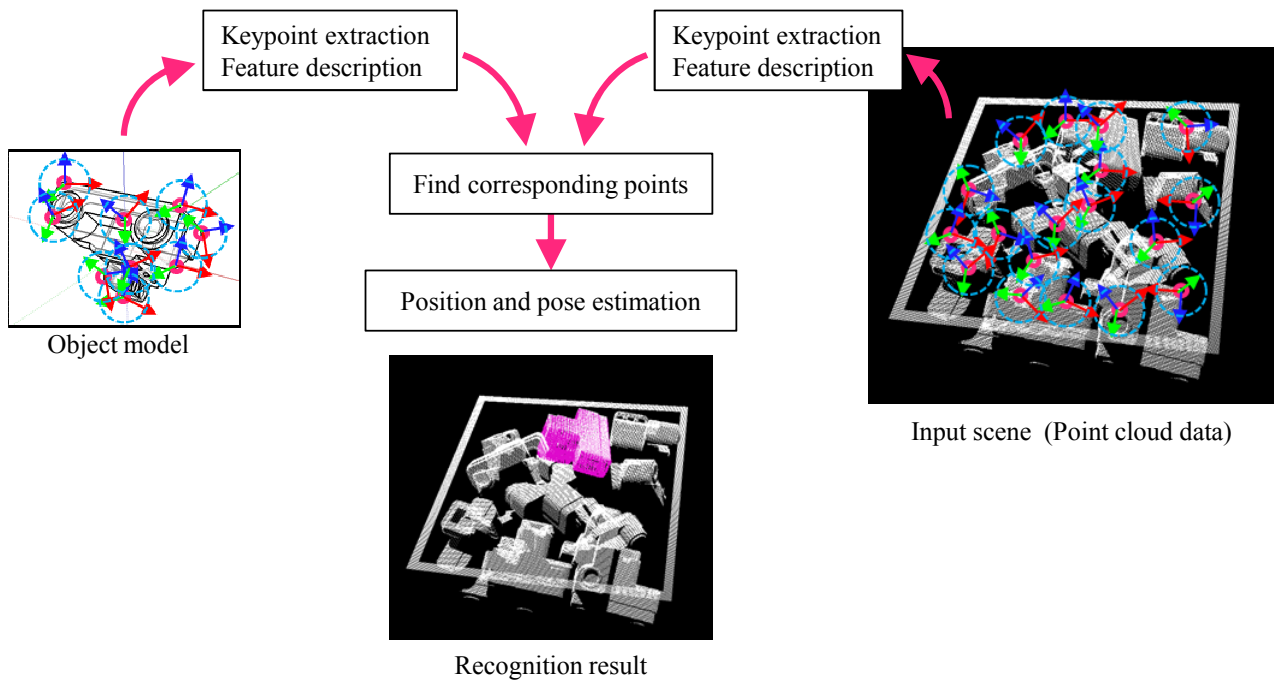


Fig.2.2 Basic pipeline of the model matching method.

れた対応点間の幾何学的関係を利用して、入力シーンにおける物体モデルの位置姿勢を算出する。

この一連の流れにおいて最も重要なのは、一番はじめのモジュールである、特徴点検出と特徴量算出である。このモジュールで記述される特徴が十分に良く設計されていなければ、後段のマッチング処理や姿勢計算で誤認識が発生する。したがって、これまでに特徴点の算出方法や特徴量の記述方法は、様々な手法 [23, 20, 59, 60, 48, 49, 14, 26, 9, 13, 40, 39] が提案されてきた。

これらの特徴量には、視点の変化に対する不変量を用いる。代表的なものに Spin Image[23] がある。Spin Image は物体表面の法線方向と、法線方向と直交した 2 次元投影面を用意し、法線方向を軸として投影面を回転させ、物体モデルを構成するデータ点をプロットした画像特徴量である。

Local Feature Histograms[20] では、不変量として物体モデルのデータ点、または三角パッチ間の距離、法線方向や、Shape Index 値 [11] を利用する。物体モデルから得られたこれらの不変量を軸とした頻度分布を特徴量とする手法である。

さらに、曲率の大きな形状の代表点を特徴点とし、特徴点のみから特徴記述することで照合を高速化する Depth Aspect Image (DAI) 法 [59, 60] や、Local Surface Patch (LSP) 法 [8] もある。DAI 特徴量は仮想的な視点から見た低解像度の距離画像であり、物体モデルから得られた全ての DAI をハッシュテーブルに記憶しておくことによって高速な対応点探索を実現する。LSP 特徴量は特徴点の法線と特徴点周りのデータ点の法線方向の差分値と Shape Index 値を計算し、これら 2 種類のデータを軸とする頻度分布である。

また、近年では局所領域内の 3 次元点群の法線方向分布を記述した Signature of Histograms of Orientations (SHOT) 特徴量 [48] や、SHOT 特徴量にテクスチャ情報を統合した Color SHOT 特徴量 [49] を用いた投票型照合手法 [47] が提案された。SHOT 特徴量では特徴点周りに複数に分割された特徴量の記述領域を設定する。それぞれの領域から特徴点の法線方向と他のデータ点の法線方向との差分値をヒストグラム化し、それらを結合することによって特徴ベクトルを生成する。Color SHOT 特徴量は、

SHOT 特徴量に加えてカラーカメラから得られるテクスチャ情報を付加した特徴量である。

別の高速化アプローチとしては、さまざまな見えの物体モデルから生成された輪郭データや、距離画像をデータベースに登録し、GPU を用いることを前提として類似度計算を並列化する Park らの認識手法 [34] がある。

これらの局所特徴量は特徴量としてのアイデンティティを高めるために、一定の空間的な広がりを持つ記述領域を設定している。しかしながら、物体同士が密に接触するばら積みシーンにおいては、特徴量の記述範囲に物体の接触に起因した外乱データを多く含みやすいため、照合性能の低下が懸念される。さらに、特徴次元数が高く対応点探索のコストが高くなりやすいといった問題点もある。

したがって、高速性が求められるピンピッキング用途においては、低次元な特徴量を採用することが望ましいと考えられる。そのような特徴量としては、3次元点の点对特徴である Surflet Pair [52] や、Point Pair Feature (PPF) [55] がある。文献 [52] の手法では、物体モデルから検出され得る全ての Surflet Pair を用いてヒストグラムを生成し、3D モデルの検索に利用した。文献 [55] の手法では、サーフェスレジストレーションに応用された。Surflet Pair と PPF は、点間の変位ベクトルと、2点の法線ベクトルの幾何関係のみで記述された4次元特徴量である。このため、記述領域の空間的な広がりが小さく、物体同士の接触による照合性能の低下の影響を受けにくい利点がある。しかしながら、特徴量自体のアイデンティティが低く、誤った対応点が発生しやすいため、安定した位置姿勢認識のためには大量の PPF を用いる必要があると考えられる。

このため、Drost らは物体モデルから検出された PPF をハッシュテーブルに保存したデータベースを用いて対応点探索を効率化し、さらに、剛体変換の計算を低コスト化することによって処理速度を向上させた [14]。また、Choi らは点对のペアリングを稜線同士や稜線と面との組合せに限定することによって、データベースに登録する PPF 数を削減した [9]。認識信頼性の向上に関しては、PPF に新たな属性として、Visibility Context を加える方法 [26] がある。3次元空間を Visible space, Surface, Invisible space の3種類の属性に分類し、照合に用いる PPF にこれらの属性値を付加することで、PPF 同士の対応付けを高信頼化した。さらに、Kinect のようなカラーカメラ付きのレンジファインダの使用を前提とし、PPF にテクスチャ情報を付加した Multimodal PPF [13] によって PPF 自体のアイデンティティを高める手法がある。しかしながら、いずれの手法も安定した照合のためには多くの PPF が必要であり、処理コストは依然として高いままであった。

2.2 ベクトルペア特徴

物体の3次元的な位置姿勢の決定には最低三つの3次元点が必要である。提案手法では高速な位置姿勢推定を実現するためにこれらの三点を、共通した一つの始点と、二つの独立した終点からなる3D ベクトルペアとして表現する。提案手法では3D ベクトルペアが照合の単位であり、複数のベクトルペアを用いることで、物体の位置姿勢を認識する。

ベクトルペアは点 P と P から変位ベクトル $\mathbf{d}_{q1}, \mathbf{d}_{q2}$ 離れた点 Q_1, Q_2 で構成される。また、それぞれの変位ベクトルの長さは l_1, l_2 、変位ベクトル間の角度を θ と表す。したがって、ベクトルの長さとその間の角度はベクトルペアを規定するパラメータである。このパラメータは実際に数値を変動させながら、認識性能が高くなるように設定する。パラメータと認識性能の関係については2.5.4節にて述べる。Fig.2.3に3D ベクトルペアの構成を示す。

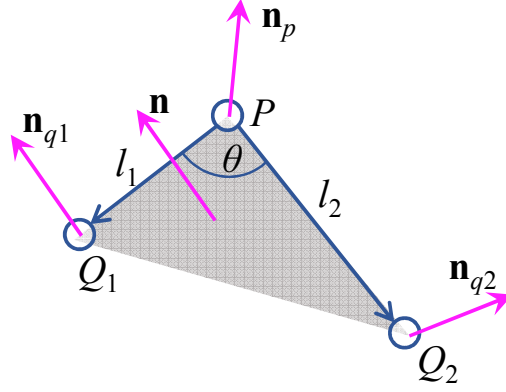


Fig.2.3 Structure of the 3D vector pair.

ベクトルペアは特徴量として、 $f(s_p, s_{q1}, s_{q2})$ を持つ。 s_p, s_{q1}, s_{q2} はそれぞれ P, Q_1, Q_2 の法線ベクトル $\mathbf{n}_p, \mathbf{n}_{q1}, \mathbf{n}_{q2}$ と、 $\triangle PQ_1Q_2$ の法線ベクトル \mathbf{n} を用いて式 (2.1) によって計算される。このとき、各法線は単位ベクトルとする。

$$s_p = \cos(\mathbf{n} \cdot \mathbf{n}_p)^{-1}, \quad s_{q1} = \cos(\mathbf{n} \cdot \mathbf{n}_{q1})^{-1}, \quad s_{q2} = \cos(\mathbf{n} \cdot \mathbf{n}_{q2})^{-1} \quad (2.1)$$

s_p, s_{q1}, s_{q2} は法線ベクトルのなす角を示しており、姿勢変動に対する不変量である。

2.2.1 離散化された 3 次元点群からの法線ベクトルの算出方法

提案手法では、照合対象とする入力データにはレンジファインダで取得された離散的な 3 次元点群を想定している。このため、物体モデルも同一のデータ形式である離散化された 3 次元点群を扱う。

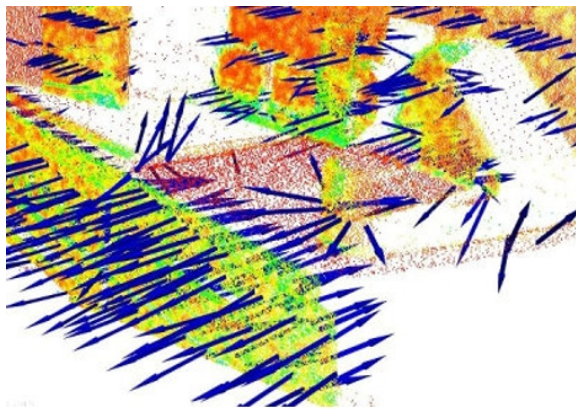
離散化された 3 次元点群からの法線ベクトルの計算は文献 [38] で提案された注目周りの複数のデータ点を利用する手法を採用する。複数の計測データを用いて法線を算出するため、三角パッチの法線をそのまま割り当てる方法よりも計測ノイズに頑健な利点がある。

まず、注目点周りの 3 次元点群を用いて式 (2.2) によって共分散行列 \mathbf{C} を作成し、固有値 $\lambda_0, \lambda_1, \lambda_2$ (ただし、 $\lambda_0 \leq \lambda_1 \leq \lambda_2$) を算出する。

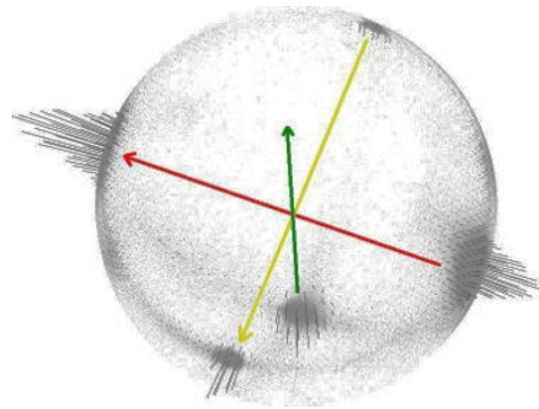
$$\mathbf{C} = \frac{1}{k} \sum_{i=1}^K (\mathbf{p}_i - \bar{\mathbf{p}}) \cdot (\mathbf{p}_i - \bar{\mathbf{p}})^T \quad (2.2)$$

\mathbf{p}_i は 3 次元点、 $\bar{\mathbf{p}}$ は \mathbf{p}_i の平均値を示す。 \mathbf{C} の最小固有値 λ_0 の固有ベクトルが $+\mathbf{n}$ または $-\mathbf{n}$ を示す。 Fig.2.4 に算出された法線ベクトルの例を示す。 Fig.2.4(a) は推定された法線ベクトルを青矢印で示している。また、 Fig.2.4(b) は Extended Gaussian Image (EGI) [21] 上に法線ベクトルをマッピングした結果である。同一平面上の法線であっても、正負の向きを持つため、EGI の全体に法線が分布している。このように、式 (2.2) によって算出した法線には向きの曖昧性があり、固有ベクトルの向きの正負を数学的に一意に決定することは困難である。そこで、視点の位置ベクトル \mathbf{v}_p を定め、全ての点に対して式 (2.3) を満たす法線方向を \mathbf{p}_i の法線方向とする。

$$\mathbf{n}_i \cdot (\mathbf{v}_p - \mathbf{p}_i) > 0 \quad (2.3)$$



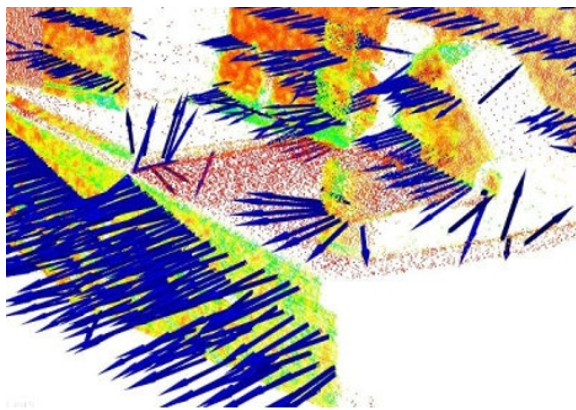
(a) Estimated normal vectors. (unflipped)



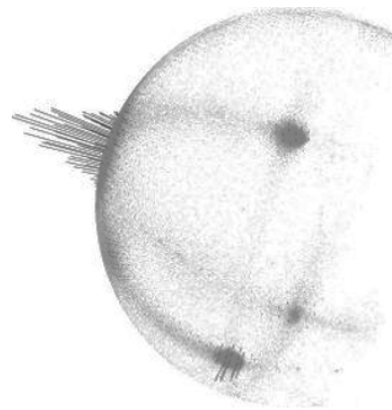
(b) Extended Gaussian Image of normal vectors of (a).

(a), (b) reference: http://pointclouds.org/documentation/tutorials/normal_estimation.php#normal-estimation

Fig.2.4 Distribution of estimated normal vectors. (a) Estimated normal vectors having orientation inconsistency. (b) Extended Gaussian Image (EGI) of the orientation of all normal vectors of (a).



(a) Estimated normal vectors. (flipped)



(b) Extended Gaussian Image of normal vectors of (c).

(a), (b) reference: http://pointclouds.org/documentation/tutorials/normal_estimation.php#normal-estimation

Fig.2.5 Distribution of flipped estimated normal vectors. (a) Overview of the flipped normal vectors. (b) Extended Gaussian Image (EGI) of the orientation of all normal vectors of (a).

Fig.2.5 に式 (2.3) を用いて法線の向きを修正した例を示す. Fig.2.5(a) では, Fig.2.4(a) と同様に推定された法線ベクトルを青矢印で示している. 同一平面上の法線方向が一定の向きに修正されていることがわかる. また, Fig.2.5(b) は Fig.2.4(b) と同様に推定された法線ベクトルを EGI 上にマッピングした例である. 法線推定をおこなった点群はセンサから得られた 2.5 次元データであるため, 法線は EGI 上で半球状に分布している.

3 次元点群の共分散行列から求めた固有ベクトルと, 視点位置を設定することによる固有ベクトルの向きの修正によって, 離散化された 3 次元点群に対して安定的に法線ベクトルを算出することができることを確認した.

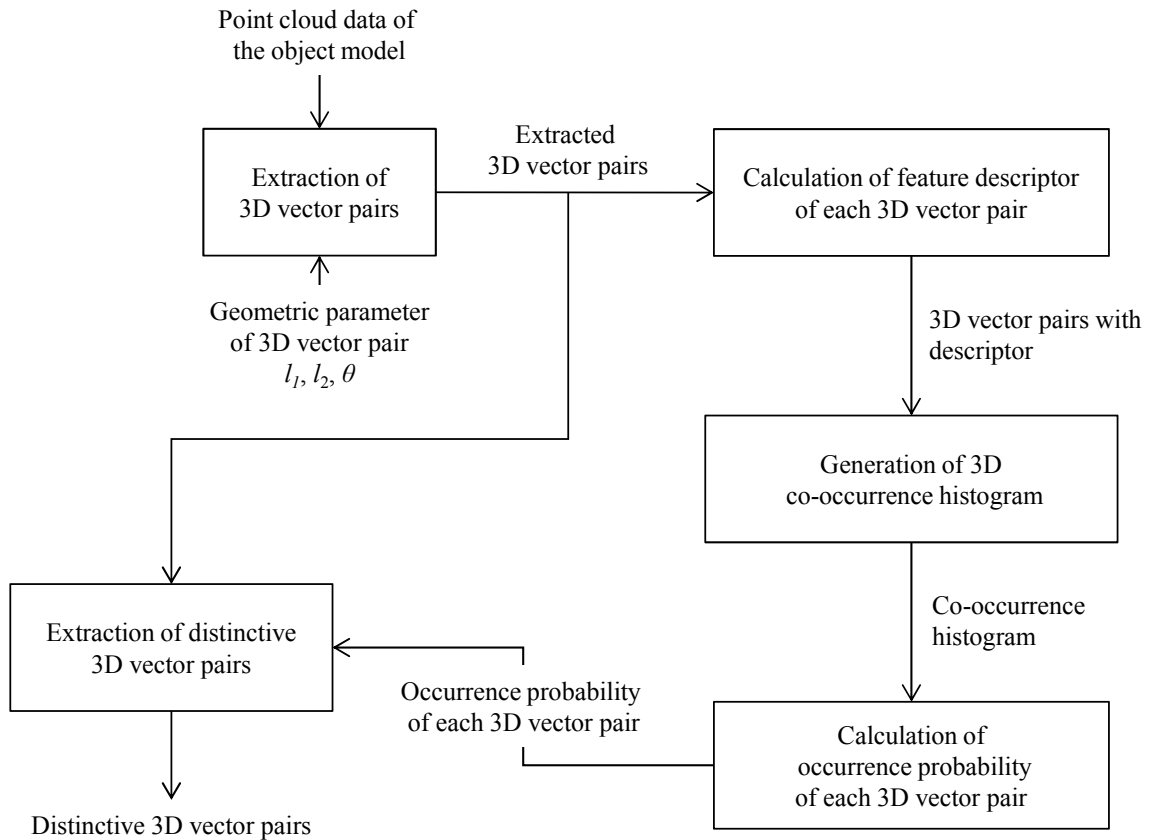


Fig.2.6 Overview of algorithm for extracting distinctive vector pairs.

2.2.2 発生確率分析に基づくベクトルペア選択

3次元物体モデルから特徴的ベクトルペアを選択する手法を説明する。概要を Fig.2.6 に示す。

Fig.2.6 において、太字は入出力データを示す。

[Extraction of 3D vector pairs] では、離散化 3次元点群表現された物体モデルとベクトルペアを規定するパラメータ l_1, l_2, θ を入力し、3D ベクトルペアを検出する。

[Calculation of feature descriptor of each 3D vector pair] は、検出されたベクトルペアすべてに対して特徴量 $f(s_p, s_{q1}, s_{q2})$ を算出である。

[Generation of 3D co-occurrence histogram] は、特徴量付き 3D ベクトルペア群からベクトルペア特徴の各成分を軸とした 3次元共起ヒストグラムを作成する。

[Calculation of occurrence probability of each 3D vector pair] は、共起ヒストグラムの投票データから、各ベクトルペアごとに発生確率を算出する。

[Extraction of distinctive 3D vector pairs] は、全てのベクトルペアから低発生確率のベクトルペアを決定し、データベースに登録する。登録されたものが特徴的ベクトルペアであり、認識に用いられる。

以下に、アルゴリズムを示す。まず、計算の高速化のために離散点群に前処理としてダウンサンプリングをおこなう。これは、候補となるベクトルペア数を抑制するための処理であり、物体モデルの凹凸が損なわれない程度にサンプリングをおこなう。また、平面上に存在する点は高精度な位置姿勢推定に不適と考えられるため削除しておく。

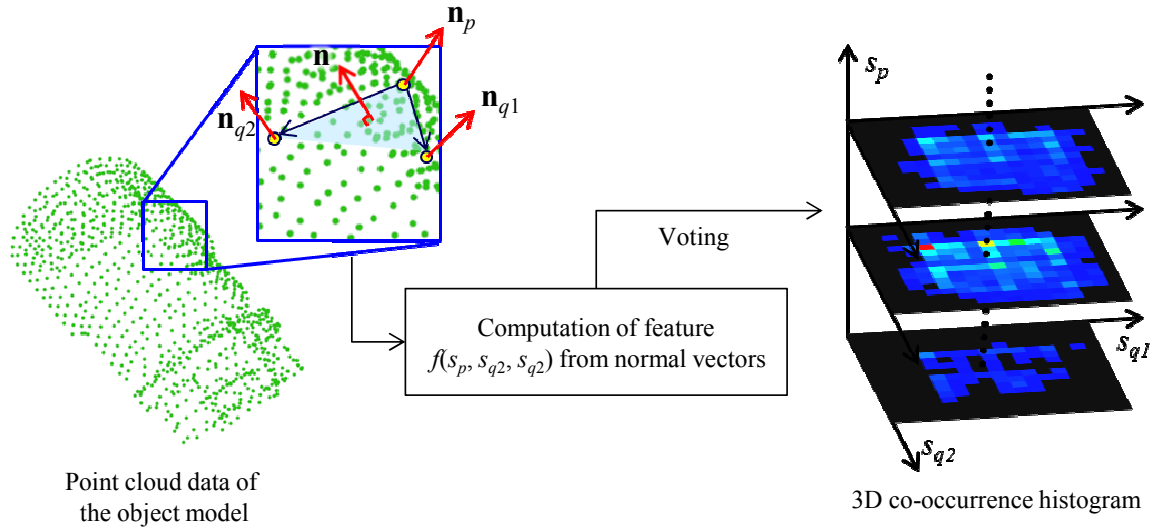


Fig.2.7 Overview of generating of the 3D co-occurrence histogram.

次に離散点群に対して、式 (2.4) を満たす 3 つの 3 次元点をベクトルペア $\{\mathbf{v}_n|0, \dots, N-1\}$ として抽出する。 L_1, L_2, θ は抽出するパラメータであり、 th_θ, th_l は許容誤差である。

$$\begin{cases} |\cos(\mathbf{d}_{q1} \cdot \mathbf{d}_{q2})^{-1} - \theta| \leq th_\theta \\ |l_1 - L_1| \leq th_l \\ |l_2 - L_2| \leq th_l \end{cases} \quad (2.4)$$

式 (2.4) の 1 行目は 2 ベクトルのなす角に関する条件であり、下 2 行は 2 ベクトルそれぞれの長さに関する条件である。 th_l は扱うデータ点群の解像度の 2 倍程度に決定し、 th_θ は th_l が許容するずれの範囲に決定する。

さらに、抽出された全てのベクトルペアに対して特徴量 f を算出する。そして、式 (2.5)(2.6) を適用することによって s_p, s_{q1}, s_{q2} をインデクスとした 3 次元共起ヒストグラム h を生成する。

$$h(s_p, s_{q1}, s_{q2}) = \sum_{n=0}^{N-1} \delta(\mathbf{v}_n) \quad (2.5)$$

$$\begin{cases} \delta = 1 & \text{when } \{\mathbf{v}_n(s_p) = s_p\} \cap \{\mathbf{v}_n(s_{q1}) = s_{q1}\} \cap \{\mathbf{v}_n(s_{q2}) = s_{q2}\} \\ \delta = 0 & \text{otherwise} \end{cases} \quad (2.6)$$

ヒストグラムの各ビンには物体モデルにおける特徴量 $f(s_p, s_{q1}, s_{q2})$ の含有量を示す。 Fig.2.7 に 3 次元共起ヒストグラムの生成の様子を示す。この処理では、以下の 3 ステップを物体モデルに含まれるベクトルペアの数だけ繰り返す。

1. 物体モデルからベクトルペアをサンプリング
2. 特徴量 $f(s_p, s_{q1}, s_{q2})$ の算出
3. $f(s_p, s_{q1}, s_{q2})$ が示すビンの値をインクリメント

Fig.2.7 の共起ヒストグラムは 3 次元であるが、便宜上 Fig.2.7 では特定の s_p による 2 次元共起ヒストグラムの重なりとして表示した。また、各ヒストグラムは得票数の少ないものから順に青～緑～赤で色付けしてある。

さらに、共起ヒストグラムの総和が1になるように式(2.7)で正規化することによって、各ベクトルペアの発生確率 Ph を求める。

$$Ph(s_p, s_{q1}, s_{q2}) = \frac{h(s_p, s_{q1}, s_{q2})}{\sum_{s_p=1}^L \sum_{s_{q1}=1}^L \sum_{s_{q2}=1}^L h(s_p, s_{q1}, s_{q2})} \quad (2.7)$$

ここで、 L は共起ヒストグラムの階調数を示す。この階調は認識システムに採用するレンジファインダの空間分解能を基準に設定すべきであるが、実験的に16とした。また、 $Ph(s_p, s_{q1}, s_{q2})$ は特徴量 $f(s_p, s_{q1}, s_{q2})$ をもつベクトルペアの発生確率を示している。

最後に発生確率の低いベクトルペアを物体モデルを構成する点の数%~10%程度選択し、データベースに登録し、認識に利用する。このとき、選択されたベクトルペアが一ヶ所に集中した場合、オクルージョンに対する頑健性が低下する。そこで、これを避けるために、ベクトルペアをある程度空間的に分散させながら選択する。発生確率の低いベクトルペアから順にデータベースへ登録するが、このとき、候補となるベクトルペアの周り $th_{extract}[\text{mm}]$ 以内にすでにデータベースに登録されたベクトルペアが存在する場合には、登録をスキップする。

2.3 ベクトルペアを用いた位置姿勢認識

2.3.1 アルゴリズム概要

選択されたベクトルペア群を用いた物体の位置姿勢推定モジュールについて説明する。提案手法では部分的なオクルージョンに対応するために、一般化ハフ変換 [4] ベースのアルゴリズムを用いて、部分的な情報から得られた姿勢仮説を統合することによって、最終的な位置姿勢を生成する。位置姿勢推定の流れを Fig.2.8 に示す。

Fig.2.8において、太字で表された **Vector pairs of an object model**, **Point cloud data of an input scene**, **Point cloud data of an object model** は入力データであり、**The most consistent (R, t)** と **Refined (R, t)** は出力データを表す。また、**[Checking geometric consistency between the object model and the input point cloud data]** は、多くの投票値を得られたいくつかの姿勢仮説から、最も信頼性の高いものを決定するモジュールである。VPM法の認識アルゴリズムとしては、ここで最も誤差値が少なかった姿勢仮説を最終的な認識結果とするが、図に示す通り、後段にICPアルゴリズムを追加することによって、認識精度を向上させることもできる。また、以下にこのアルゴリズムを各ステップごとに示す。

3D ベクトルペアによる位置姿勢認識

1. 前処理
2. ベクトルペア検出

入力3次元点群からすべてのベクトルペア $V_s = \{\mathbf{v}_{sm} | m = 0, \dots, M-1\}$ を検出。

[Repeat M times]

3. 姿勢仮説生成と投票処理

(a) 対応点 $c_k = (\mathbf{v}_m, \mathbf{v}_s)$ の決定

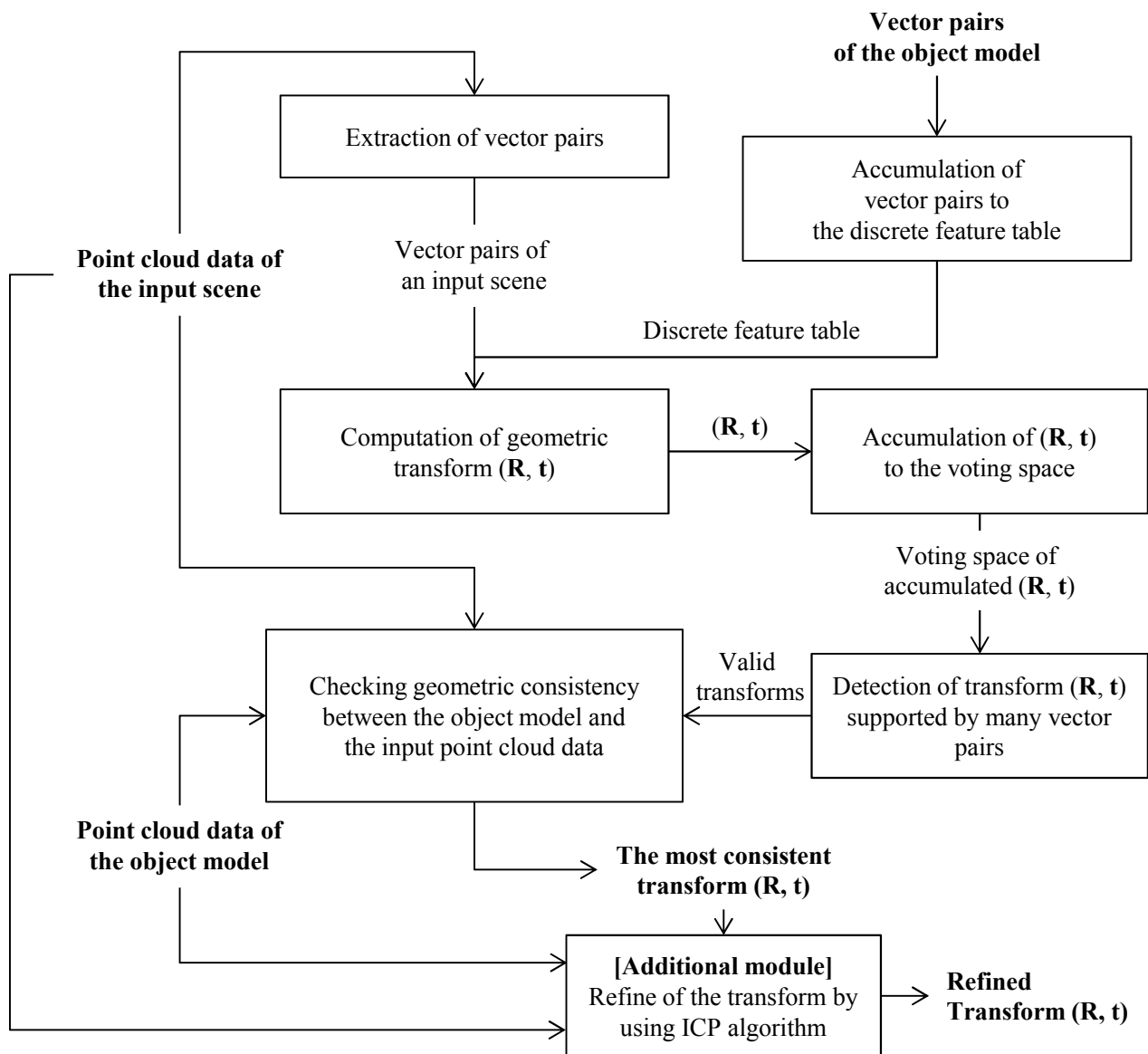


Fig.2.8 Overview of algorithm for recognizing position and pose of the object using vector pairs.

(b) 姿勢仮説の T_k の算出

(c) 投票

[End repeat]

4. 姿勢仮説のクラスタリング

5. モデル照合

6. 追加モジュール：位置姿勢の修正

1. では、前処理として物体モデルから選択されたベクトルペア群をテーブル化する。これは対応点探索を高速化するための工夫である。2. では入力3次元点群からのベクトルペア検出をおこなう。検出されたベクトルペアとテーブル化された特徴的ベクトルペアを用いて、3. では姿勢仮説を生成する。生成された姿勢仮説は投票空間へ蓄積される。4. にて投票空間から有効な得票数を持つ座標を検出し、そこ

に投票された姿勢仮説を統合する．5. では，統合された姿勢仮説を用いて物体モデルを剛体変換し入力シーンとの誤差計算をおこない，誤差の小さいものを最終的な認識結果とする．6. は追加のモジュールである．6. ではICP アルゴリズムを用いて推定された位置姿勢を修正し，高精度化する．

2.3.2 前処理・ベクトルペア検出

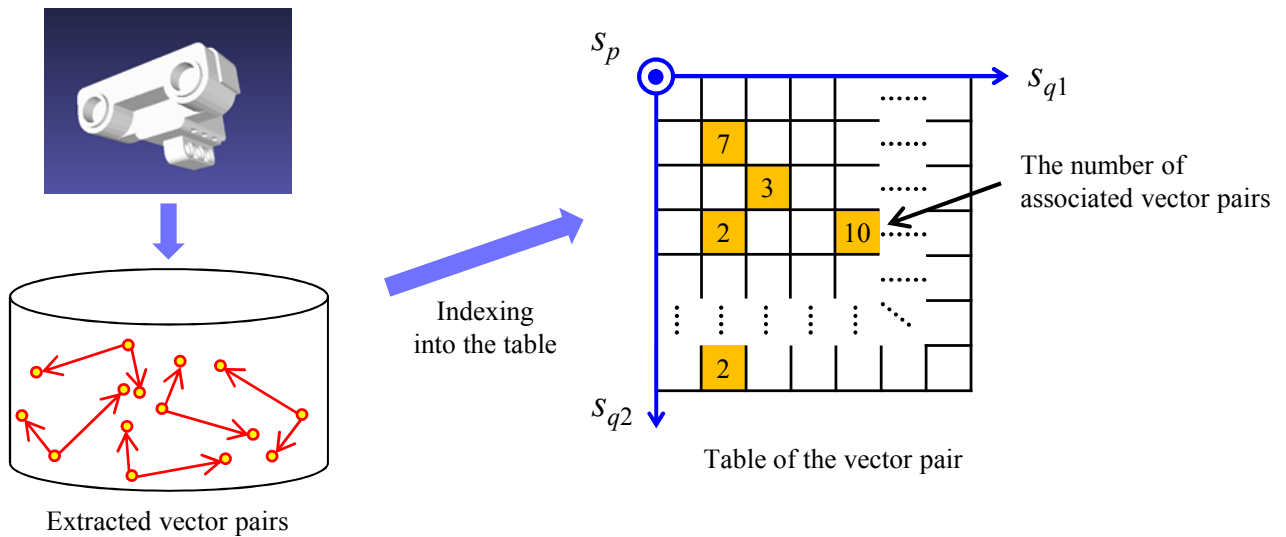


Fig.2.9 Overview of the indexing vector pairs into the table.

入力 3 次元点群 S に一様ダウンサンプリングと平面部分に属する点群を除去することにより， S^* を得る．平面除去は各データ点をもつ曲率に着目することで実行する． S に属する点のうち，曲率の小さな点を平面に属するとみなすことで除去する点を決定する．曲率は式 (2.2) の共分散行列の固有値 $\lambda_0, \lambda_1, \lambda_2$ (ただし， $\lambda_0 \leq \lambda_1 \leq \lambda_2$) を用いて式 (2.8) によって計算される．

$$\frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \quad (2.8)$$

ダウンサンプリング，平面除去後の 3 次元点群を Fig.2.10 に示す．白点は S を表し，赤点は S^* を示す．

前処理が終了したら， S^* から L_1, L_2, θ をもつベクトルペアをすべて検出する．これには， L_1, L_2 の距離を持つ点对を検出し，点对のペアのなす角が θ である時にベクトルペアとして保持する，という処

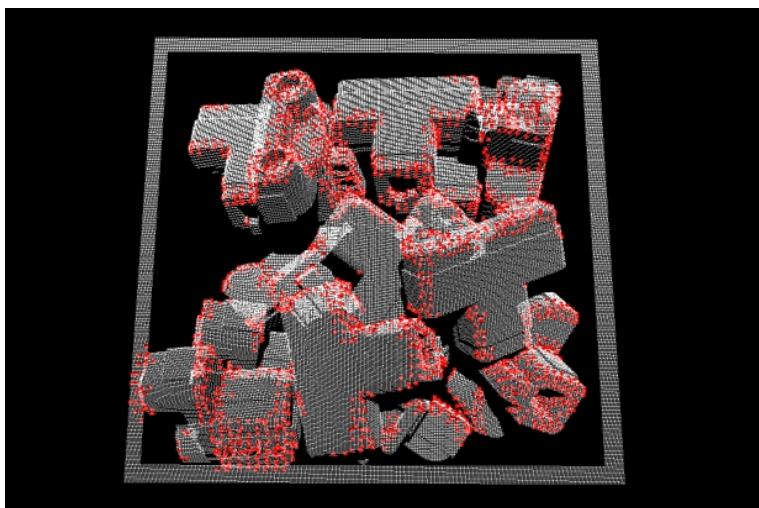


Fig.2.10 Overview of the reduced point cloud data. White dots mean original point cloud S . Red dots mean reduced point cloud S^* which has large curvature.

理を繰り返す。この探索には、膨大な回数の点同士の距離計算をおこなう必要がある。そこで、 S^* を kd 木構造に変換した上で探索をおこなうことによって高速化する。ここで検出されたベクトルペアを $V_s = \{\mathbf{v}_{sm} | m = 0, \dots, M-1\}$ とする。

2.3.3 姿勢仮説生成と投票処理

物体モデルと入力シーンとの対応点を決定する。これには、テーブルに登録された V_m と S^* から検出された V_s を用いる。以下、(a)~(c) を手順を説明する。

対応点の決定：対応点 $c_k = (\mathbf{v}_m, \mathbf{v}_s)$ は \mathbf{v}_s をテーブルにアクセスすることによって決定される。アクセスしたテーブルのビン内に登録されているすべてのベクトルペアを対応点群 C に追加する。この操作では特徴ベクトルのノルム計算が不要であるため、非常に高速に対応点を生成することができる。

姿勢仮説の算出：生成されたすべての対応点 c に対して、 \mathbf{v}_m を \mathbf{v}_s へ位置合わせする剛体変換 T を計算する。この T が姿勢仮説である。a) の処理では多くの対応点が生成されるため、剛体変換の算出は低コストであることが望ましい。そこで、式 (2.9) に示す単純な行列演算のみで高速に剛体変換の回転成分 \mathbf{R} を計算する。

$$\mathbf{R} = \mathbf{M}_m^T \mathbf{M}_s \quad (2.9)$$

このとき、 \mathbf{M}_m と \mathbf{M}_s はベクトルペアから計算される 3×3 行列であり、それぞれの成分はベクトルペアを構成する 3 点の位置ベクトル $\mathbf{p}, \mathbf{q}_1, \mathbf{q}_2$ を用いて式 (2.10) で表す。

$$\mathbf{M} = \begin{bmatrix} \frac{\mathbf{q}_1 - \mathbf{p}}{\|\mathbf{q}_1 - \mathbf{p}\|} & \frac{\mathbf{q}_2 - \mathbf{p}}{\|\mathbf{q}_2 - \mathbf{p}\|} & \frac{(\mathbf{q}_1 - \mathbf{p}) \times (\mathbf{q}_2 - \mathbf{p})}{\|(\mathbf{q}_1 - \mathbf{p}) \times (\mathbf{q}_2 - \mathbf{p})\|} \end{bmatrix} \quad (2.10)$$

\mathbf{M}_m と \mathbf{M}_s の関係は剛体変換パラメータを用いて $\mathbf{M}_m \mathbf{R} = \mathbf{M}_s$ によって表現される。このため、 \mathbf{R} の計算には \mathbf{M}_m^{-1} の計算が必要であるが、ベクトルペアの幾何パラメータ θ を 90 度に設定することと、抽出されたベクトルペアに対して、グラム・シュミットの正規直交化法を適用することによって、 \mathbf{M}_m と \mathbf{M}_s

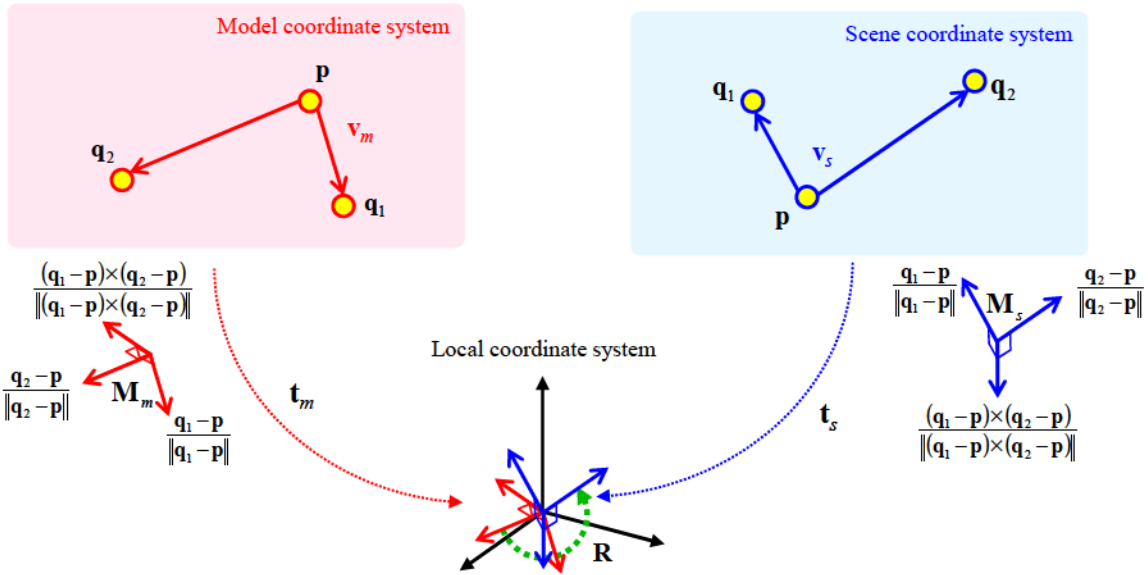


Fig.2.11 Overview of calculation of the rigid transformation.

Fig.2.11 では、まず剛体変換を算出する物体モデル座標系で表された \mathbf{v}_m , シーン座標系で表された \mathbf{v}_s をそれぞれローカル座標系で表し、点 P を原点に移す。原点に移した際の変位ベクトルを $\mathbf{t}_m, \mathbf{t}_s$ とする。このとき、二つのベクトルペアを $\mathbf{M}_m, \mathbf{M}_s$ に変換する。そして、 $\mathbf{M}_m, \mathbf{M}_s$ を位置合わせする \mathbf{R} を計算する。

投票：XYZ 軸によって構成された投票空間に対して投票を行う。この空間には物体重心の位置を投票する。したがって、多くの得票を得た座標は物体の重心位置である可能性が高いことを意味する。投票座標は姿勢仮説 T と T を生成したベクトルペアのオフセットベクトル \mathbf{d}_c を用いて式 (2.11) によって計算され、投票空間における E に対応する座標の投票値をインクリメントする。

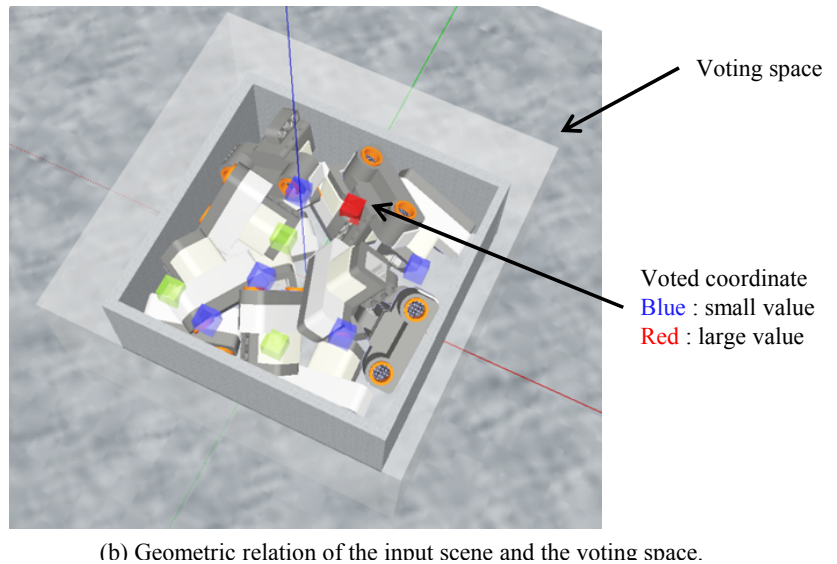
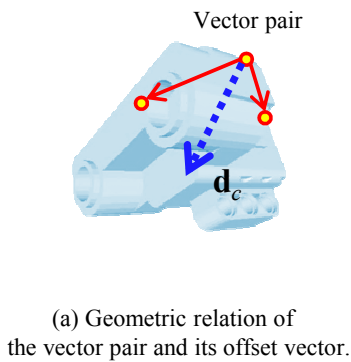
$$E = (\mathbf{d}_c - \mathbf{t}_m)\mathbf{R} + \mathbf{t}_s \quad (2.11)$$

このとき、同一の \mathbf{v}_m によって得られた複数の姿勢仮説が同一の E を指す場合は、重複してインクリメントしない。これは、物体の位置姿勢において \mathbf{v}_m と \mathbf{v}_s は一対一対応であるということに基づいている。つまり、シーン中のある点に対応付けられる物体モデルの点は一点のみであるべきである。投票の様子を Fig.2.12 に示す。

Fig.2.12(a) はベクトルペアとオフセットベクトルの関係を示しており、Fig.2.12(b) は投票空間への姿勢仮説の蓄積の様子を示している。投票の少ない座標から青～赤になるように着色した。

2.3.4 姿勢仮説のクラスタリング

投票処理が終了したら、しきい値以上の得票を得た投票空間の座標を確認し、その座標への投票に使われた T を平均化することで T' を得る。投票された座標には、誤った姿勢仮説が投票されている可能性が



ある.
平均

を

T_0
T_1
\vdots
T_x

Pose hypothesis which are voted into same coordinate of voting space

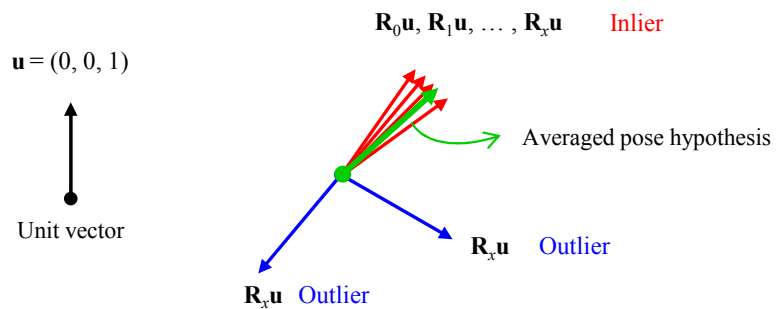


Fig.2.13 Overview of the clustering method of the pose hypothesis.

まず、同一の投票座標に関連付けられた T の回転変換パラメータ \mathbf{R} と単位方向ベクトル \mathbf{u} を用いて $\mathbf{R}\mathbf{u}$ を計算する。多くの回転変換パラメータはほぼ同一の変換を示すため、 $\mathbf{R}\mathbf{u}$ は類似した方向を向く。一方で、誤った回転変換パラメータは孤立した方向を向く。 $\mathbf{R}\mathbf{u}$ 群に対してクラスタリングを行い、最も要素数の多いクラスに対して、 T の平均化をおこなう。クラスタリングには以下の手順をおこなう。

1. $\mathbf{R}\mathbf{u}$ 群に対して kd 木を生成する。
2. kd 木に対して $\mathbf{R}\mathbf{u}_x$ をクエリとして th_{rot} の探索範囲で要素数を調べる。
3. th_{rot} の範囲で最大の要素数となった T を記憶する。(インライアとする。)

th_{rot} は平均化に用いる姿勢パラメータの要求精度である。実験的にクエリの $\mathbf{R}\mathbf{u}_x$ に対して他の $\mathbf{R}\mathbf{u}_x$

が5度以内をインライアとなるように設定した．平均化に際しては， T のうち，回転成分を表す \mathbf{R} は回転行列であるため，各成分を単純平均することができない．そこで， \mathbf{R} をXYZ系オイラー角に変換し，それぞれの軸ごとに平均化し，回転行列に再変換する方法を採る．回転行列 \mathbf{R} を式(2.12)としたとき， \mathbf{R} とXYZ系オイラー角の相互変換は式(2.13)(2.14)の通りである．

$$\mathbf{R} = \begin{bmatrix} m_{00} & m_{01} & m_{02} \\ m_{10} & m_{11} & m_{12} \\ m_{20} & m_{21} & m_{22} \end{bmatrix} \quad (2.12)$$

$$x = \tan^{-1}(m_{21}/m_{22}), \quad y = \sin^{-1}(-m_{20}), \quad z = \tan^{-1}(m_{10}/m_{00}) \quad (2.13)$$

$$\mathbf{R} = \begin{bmatrix} \cos y \cos z & \sin x \sin y \cos z - \cos x \sin z & \cos x \sin y \cos z + \sin x \sin z \\ \cos y \sin z & \sin x \sin y \sin z + \cos x \cos z & \cos x \sin y \sin z - \sin x \cos z \\ -\sin y & \sin x \cos y & \cos x \cos y \end{bmatrix} \quad (2.14)$$

最大要素数のクラスに属する T のうち，平行移動成分を表す \mathbf{t} に関しては単純平均する．以上の操作によって T' が計算される．

2.3.5 モデル照合

算出された姿勢仮説が入力シーン中の対象物の位置姿勢を正しく表しているとき，その姿勢仮説を用いて剛体変換した物体モデルと入力3次元点群の誤差は非常に少なくなる．そこで，算出された複数の T' から最も信頼性の高い姿勢仮説を決定するために，姿勢変換した物体モデルと入力3次元点群との誤差計算をおこない，最も誤差の少ない姿勢仮説を最終的な認識結果とする．誤差計算には式(2.15)を用いる．

$$Error = \arg \min_{T'(\mathbf{R}, \mathbf{t})} \left\{ \sum_{m=0}^M \|\mathbf{R}\mathbf{p}_m + \mathbf{t}_s - \mathbf{q}\| \right\} \quad (2.15)$$

M は計算に用いた物体モデルの点数である． \mathbf{p}_m は物体モデルの m 番目の3次元点の位置ベクトルであり， \mathbf{q} は \mathbf{p}_m の最近傍点である．この誤差計算に用いられる物体モデルは全周囲のデータをもつ3次元点群であることに対して，入力シーンはレンジファインダで取得された1視点からの3次元点群データである．このため，入力シーンで観測される点群データは1物体につき片面のみの最大50%である．したがって，式(2.15)の計算時に観測されるはずの無い点に関する誤差まで蓄積されるため，見かけ以上に誤差が大きくなる場合がある．

そこで，式(2.15)において，剛体変換した物体モデル点の中で，レンジファインダから観測されないと予想されるデータ点を削除した上で誤差計算をおこなう．これには式(2.16)を用いて削除するデータを判断する．

$$\begin{cases} \delta(\mathbf{R}\mathbf{n}_m, \mathbf{v}) = 1 & \text{when } \mathbf{R}\mathbf{n}_m \cdot \mathbf{v} < 0 \\ \delta(\mathbf{R}\mathbf{n}_m, \mathbf{v}) = 0 & \text{otherwise} \end{cases} \quad (2.16)$$

\mathbf{n}_m は物体モデルの m 番目の3次元点の法線ベクトルである． \mathbf{v} は視線方向ベクトルであり，レンジファインダの視線と一致するように設定する．式(2.16)において，関数 δ の出力を0にするデータ点を削除対象とする．すなわち，姿勢変換後の物体モデルの点のうち，視線方向との内積が0より大きくなる法線ベクトルをもつ点を削除対象とした．

提案手法における物体モデルと入力シーンとの誤差計算は、式 (2.15) と式 (2.16) を統合した式 (2.17) とした。

$$Error = \arg \min_{T'(\mathbf{R}, \mathbf{t})} \left\{ \sum_{m=0}^M \delta(\mathbf{n}_m \mathbf{R}, \mathbf{v}) \|\mathbf{R} \mathbf{p}_m + \mathbf{t}_s - \mathbf{q}\| \right\} \quad (2.17)$$

誤差計算に用いた点を Fig.2.14 に示す。

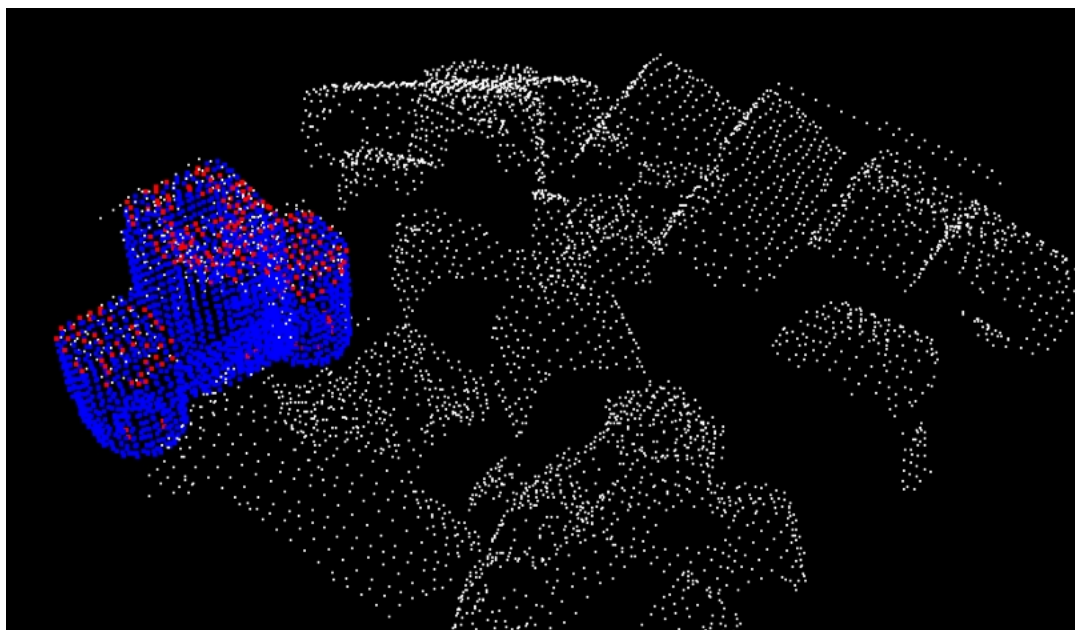


Fig.2.14 Overview of the result of the model matching module. White dots represent the input scene. Blue dots represent a transformed object model. Red dots are points which are used for calculation of alignment error.

Fig.2.14 において、白点は入力 3 次元点群、青点が姿勢変換後の物体モデルを表し、物体モデルの点の中で、誤差計算に用いられたものを赤点で表す。視線ベクトルと正対する点のみが誤差計算に用いられたことがわかる。

さらに誤差計算の高速化のために式 (2.17) の早期打ち切りもおこなう。物体モデルの点数の 10% おきに蓄積されている誤差値をチェックし、すでに保存されている T' よりも誤差が大きい場合に計算を打ち切る。以上の処理によって入力 3 次元点群との誤差を最小化する T' を認識結果とする。

2.3.6 追加モジュール：位置姿勢の修正

本モジュールは、アプリケーションが要求する位置姿勢推定精度に応じて追加するモジュールである。前段のモジュールまでで推定された位置姿勢に基づいて剛体変換した物体モデルの 3 次元点群と入力 3 次元点群を用いて、ICP アルゴリズム [5] を適用することによって剛体変換の推定精度を向上する。

ICP アルゴリズムは大まかに位置合わせされた位置姿勢を初期値として、「データ点群の最近傍点の対応付け」と「対応点からの剛体変換の推定」を繰り返すことによって、剛体変換を修正し、位置合わせ誤差を削減するアルゴリズムである。この繰り返し演算は山登り探索の性質を持っており、局所解に収束す

るため、初期値として精度の高い剛体変換を与える必要がある。

ICP アルゴリズムにて実行される処理は以下の通りである。

ICP アルゴリズムの処理の概要

[Repeat K times]

1. 最近傍点の対応付け $Y_k = C(P_k, Q)$
2. 対応点からの剛体変換の推定 $(T'_{k+1}, d_{k+1}) = X(P, Y_k)$
3. 剛体変換の適用 $P_{k+1} = T'_{k+1}P$
4. 収束判定 $d_k - d_{k+1} < \tau$

[Repeat end]

最近傍点の対応付け： $Y_k = C(P_k, Q)$ は関数 C によって物体モデル 3 次元点群の最近傍点を入力 3 次元点群から割り当てるモジュールである。 P_k, Q はそれぞれ、 k 番目の変換を適用した物体モデルの 3 次元点群と入力 3 次元点群を示す。 Y_k は P_k の最近傍点の集合である。 Y の要素である物体モデルの点 \mathbf{p} に対する入力 3 次元点群の最近傍点は式 (2.18) を用いて決定する。

$$\mathbf{y} = \arg \min_{q \in Q} \|\mathbf{q} - \mathbf{p}\| \quad (2.18)$$

式 (2.18) を物体モデルを構成する全ての点に対して適用し、最近傍点の集合 Y を求める操作を式 (2.19) で表す。

$$Y = C(P, Q) \quad (2.19)$$

対応点からの剛体変換の推定： $(T'_{k+1}, d_{k+1}) = X(P, Y_k)$ において、関数 X は点群 P, Y_k の位置合わせ誤差を最小化する剛体変換 T'_{k+1} を求める。 d_{k+1} はこのときの位置合わせ誤差である。

関数 X は式 (2.20) によって表される。

$$X(\Sigma_{pq}) = \begin{bmatrix} \text{tr}(\Sigma_{pq}) & \Delta^t \\ \Delta & \Sigma_{pq} + \Sigma_{pq}^T - \text{tr}(\Sigma_{pq})\mathbf{I}_3 \end{bmatrix} \quad (2.20)$$

ここで、 Σ_{pq} は下記の通りである。

$$\Sigma_{pq} = \frac{1}{N_p} \sum_{i=1}^{N_p} [\mathbf{p}_i - \boldsymbol{\mu}_p][\mathbf{y}_i - \boldsymbol{\mu}_y] \quad (2.21)$$

$\boldsymbol{\mu}_p, \boldsymbol{\mu}_y$ は、 P と Y それぞれの重心を示す。 また、 Δ は $\mathbf{A}_{ij} = (\Sigma_{pq} - \Sigma_{pq}^T)_{ij}$ の反対称成分から作成された列ベクトルであり、 $\Delta = [A_{23} \ A_{31} \ A_{12}]$ で表す。 なお、 \mathbf{I}_3 は 3x3 の恒等変換である。

$X(\Sigma_{pq})$ の最大固有ベクトルが最適な回転を表す単位四元数 $\mathbf{x}_R = [x_0 \ x_1 \ x_2 \ x_3]^T$ である。 さらに、式 (2.22) によって単位四元数を回転行列 \mathbf{R} に変換する。

$$\mathbf{R} = \begin{bmatrix} x_0^2 + x_1^2 - x_2^2 - x_3^2 & 2(x_1x_2 - x_0x_3) & 2(x_1x_3 + x_0x_2) \\ 2(x_1x_2 + x_0x_3) & x_0^2 - x_1^2 + x_2^2 - x_3^2 & 2(x_2x_3 - x_0x_1) \\ 2(x_1x_3 - x_0x_2) & 2(x_2x_3 + x_0x_1) & x_0^2 - x_1^2 - x_2^2 + x_3^2 \end{bmatrix} \quad (2.22)$$

平行移動成分 \mathbf{x}_t は式 (2.23) の通りである。

$$\mathbf{x}_t = \boldsymbol{\mu}_y - \mathbf{R}\boldsymbol{\mu}_p \quad (2.23)$$

式 (2.22) (2.23) によって得られた \mathbf{R}, \mathbf{x}_t が $k+1$ 番目の剛体変換の修正結果 T'_{k+1} である。式 (2.24) によって、 T'_{k+1} を適用した物体モデルの 3 次元点群と入力 3 次元点群の誤差 d を計算する。

$$d = \frac{1}{N_p} \sum_{i=0}^{N_p} \|\mathbf{y}_i - \mathbf{R}\mathbf{p}_i - \mathbf{x}_t\|^2 \quad (2.24)$$

剛体変換の適用：ここでは、式 (2.25) を適用する。式 (2.25) では、最新の剛体変換 T' を用いて物体モデルの 3 次元点群を剛体変換する。

$$P_{k+1} = T'_{k+1}P \quad (2.25)$$

収束判定：上記までの処理の最後に、ここまでの繰り返し演算の終了判定をおこなう。 $d_k - d_{k+1} < \tau$ によって、誤差の改善がしきい値 τ 未満となったら収束したとみなし、 T'_{k+1} を最終的な解とする。

2.4 可観測性を考慮したベクトルペア選択による認識の高信頼化

本節では、データ点の「可観測率」という新たな評価指標を特徴点選択に加えることによって、高信頼な対応点探索を実現する手法について説明する。

認識に用いる特徴点の厳選手法 [59, 60, 8, 48, 40, 39] はデータベースの削減と対応付けの信頼性の向上に大きく貢献する。しかしながら、選択された特徴点は必ずしも入力シーンで観測されるとは限らない。これらの特徴点選択手法には、認識時に特徴点が隠れている場合は対応点が得られず、認識の信頼性が低下するという共通的な問題がある。

そこで本研究では、物体モデルを構成するデータ点の入力シーンにおける観測可能性を事前にシミュレートし、さまざまな視点から安定して観測されうる点を決定する。

そして、「可観測性」と「独自性」を統合し、2つの指標を両立するベクトルペアを選択することによって認識性能を高性能化する。

2.4.1 可観測性の定義

対応点探索時に特徴点が入力シーンで観測されない場合としては、以下の2つが考えられる。

1. 他の物体によって隠されている場合
2. 物体自身によって隠されている場合

Fig.2.15 にばら積みシーンにおいて発生するオクルージョンの例を示す。この図では枠線で囲われた3つの物体に絞って説明する。Fig.2.15 では、物体によって隠された部分を点線で示しており、Object A と Object B が Object C を隠していることがわかる。このとき、Object C を認識できるようにしたとしても、Object A, Object B への干渉なしに拾い上げることが困難である。干渉によってハンドを傷める場合や、部品を傷つけることが想定される。このように他の物体によって隠された Object C のような

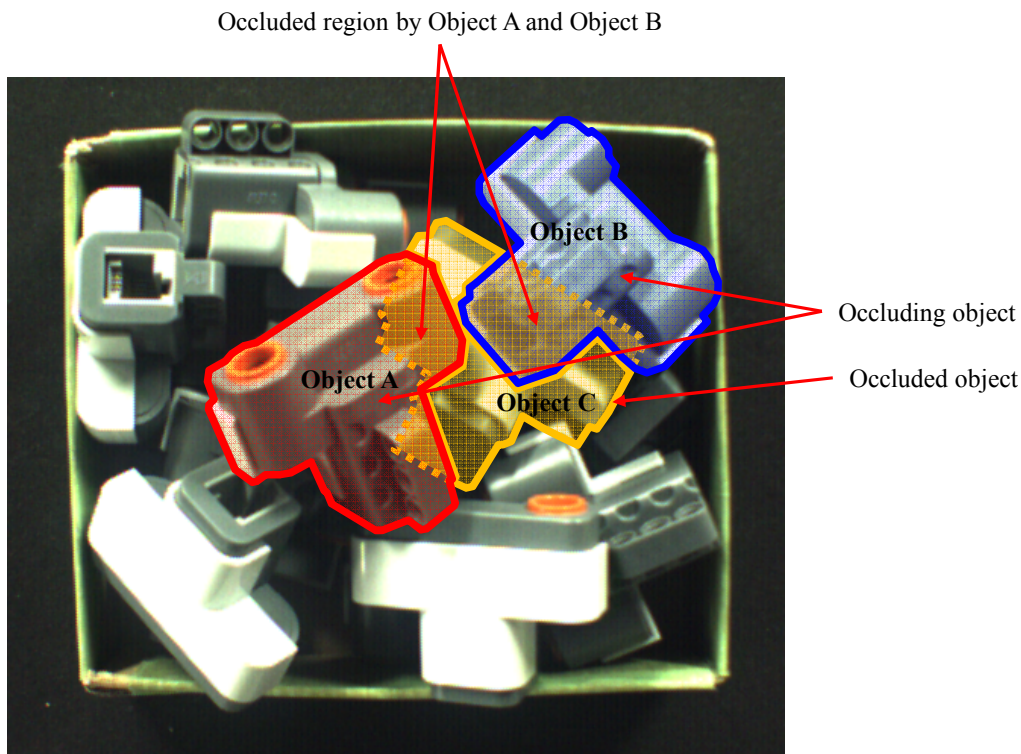


Fig.2.15 An example of occlusion of randomly stacked objects in the scene. Object C is occluded by Object A and Object B.

物体の認識は重要でないと考えられる。したがって、1. の条件によって観測されない特徴点を持つ物体を認識することはピッキングシステムの実用上の利点は高くないと考えられる。

一方で、Object A や Object B のように、一番上に配置された物体は他の物体との干渉なしに拾い上げることでできる可能性が高いため、そのような物体の位置姿勢認識の信頼性向上はピッキングシステムの性能向上において有用である。しかしながら、シーン中でばら積み状態の部品群の一番上に配置された物体であっても、2. が原因で観測されない特徴点が存在する。これは主に物体自身の形状等によって引き起こされるセルフオクルージョンである。

したがって、本研究では2. に関するオクルージョンを解決することによって対応点探索を高信頼化する。これには、2. をシミュレートすることによって、物体モデルを構成するデータ点の観測可能性を計算することによって、セルフオクルージョンのリスクの低い、安定して観測されるデータ点から特徴点を選択する方法を採る。以下、可観測点のシミュレートの方法を述べる。

提案手法で離散点群表現された物体の各点に注目し、特定の視点から観測可能な点を可観測な点とする。これには、一般的な隠面除去をおこなう。すなわち、視点方向を法線ベクトルとした射影面に物体モデルの表面データを投影した際に、投影面の各 (i, j) 座標において最も射影面に近いデータ点を可観測と判定する。Fig.2.16 にある視点に対する可観測な点群の例を示す。

Fig.2.16(a) は物体モデルと視点の位置関係を示す。視線方向は視点位置から物体モデル重心を向を向いている。また、視線ベクトルを法線とした射影面が設定されており、この面に物体の表面データを投影する。Fig.2.16(b) は (a) の視点から得られた可観測データ点群である。側面や、裏面に属するデータ点が削除されていることがわかる。

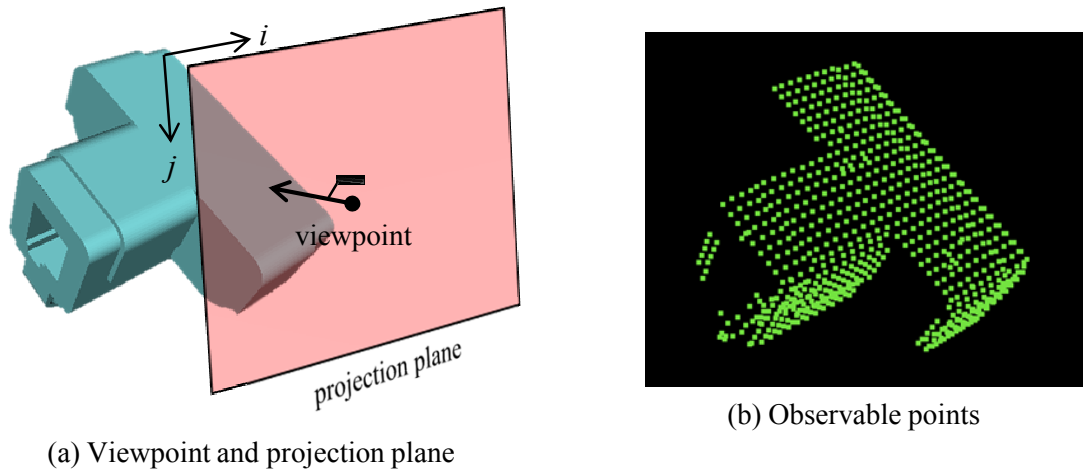


Fig.2.16 An example of projection plane and its observable points. (a) Geometric relation of the projection plane and the object model. (b) Observable points of viewpoint (a).

2.4.2 可観測率の算出

物体モデルの可観測率は、物体モデルの周囲に K 点の観測点をランダムに設置し、各観測点から物体モデルの各点の観測状態を分析することによって物体モデル 1 点ごとに確率として算出する。点 p の可観測率は式 (2.26) によって計算される。

$$Obs(p) = \frac{1}{K} \sum_{k=0}^{K-1} \delta(p, \mathbf{o}_k) \quad (2.26)$$

ここで、 \mathbf{o}_n は観測方向ベクトル、 δ は \mathbf{o}_n を法線とする平面に物体モデルを投影したときに p が観測可能な場合に 1 を出力する関数である。500 点の観測点から生成された物体モデルの可観測率をカラーマップで表示した例を Fig.2.17 に示す。

Fig.2.17(a)(c) は物体モデル外観、(b)(d) は対応する可観測率マップである。可観測率が高くなるにつれて青～緑～黄～赤になるように着色されている。物体の外輪郭部分の可観測率が高く、窪み部分での可観測率が低くなる結果が得られた。また、物体モデルには同様の窪み形状であっても、多くの視点から観測可能な部分とそうでない部分が存在する。そのような違いも Fig.2.17(d) の凹形状部分の可観測率に反映されることがわかる。

ベクトルペア \mathbf{v} の可観測率 $Obs(\mathbf{v})$ の計算は、物体モデルの可観測率と同様のアプローチで計算する。ある視点からベクトルペアが観測可能な状態は、ベクトルペアを構成する P, Q_1, Q_2 が 3 点とも可観測であるときである。すなわち、 $Obs(\mathbf{v})$ は $\delta(P, \mathbf{o}), \delta(Q_1, \mathbf{o}), \delta(Q_2, \mathbf{o})$ の同時生起確率として表現できる。したがって、式 (2.26) を拡張した式 (2.27) によって計算される。

$$Obs(\mathbf{v}) = \frac{1}{K} \sum_{k=0}^{K-1} \delta(P, \mathbf{o}_k) \delta(Q_1, \mathbf{o}_k) \delta(Q_2, \mathbf{o}_k) \quad (2.27)$$

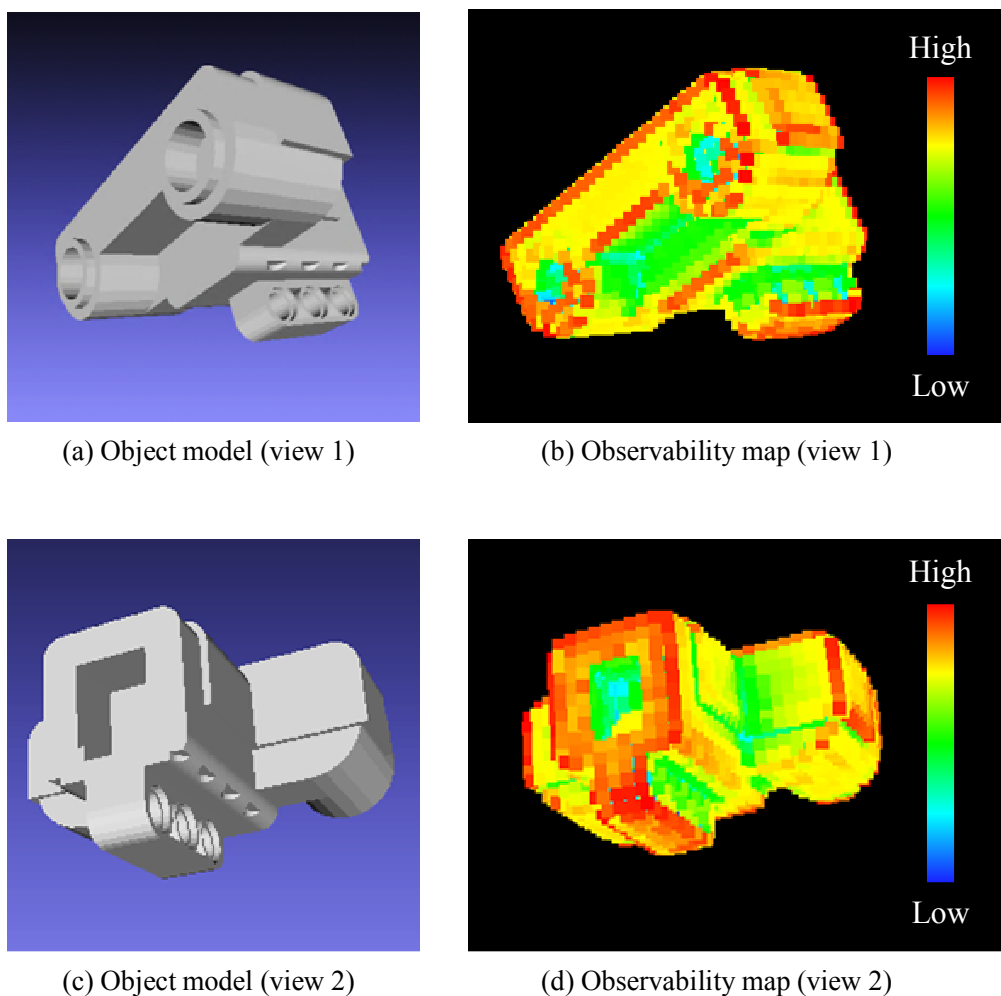


Fig.2.17 Object model and observability map. (blue: low observability; red: high observability)

2.4.3 可観測性と独自性を統合したベクトルペア選択

物体モデルから算出された各ベクトルペアの可観測率 $Obs(\mathbf{v})$ と発生確率 $Ph(\mathbf{v})$ を統合することによって、2つの指標を両立した、認識に有効なベクトルペアを選択する。ベクトルペアの統合指標の計算方法を式 (2.28) に示す。

$$I(\mathbf{v}) = \omega Obs(\mathbf{v}) + (1 - \omega)(1 - Ph(\mathbf{v})) \quad (2.28)$$

式 (2.28) では、可観測性重視型とするか、独自性重視型とするかを調節できるように、ベクトルペアを2つの指標の重み付き線形和として評価することとした。 ω は統合時の重みパラメータである。それぞれのパラメータのバランスに関する考察は 2.5.6 にて述べる。また、 $Obs(\mathbf{v}), Ph(\mathbf{v})$ は取りうる値の範囲が大きく異なるため、それぞれ事前に最大値で除しておく。

統合指標 $I(\mathbf{v})$ の値の高いベクトルペアを順にデータベースに登録する。ここでも、2.2.2 にて説明した方法と同様に、候補となるベクトルペアの周り $th_{extract}[\text{mm}]$ 以内にすでにデータベースに登録されたベクトルペアが存在する場合に、登録をスキップすることによって、選択されるベクトルペアを空間的に分散させる。

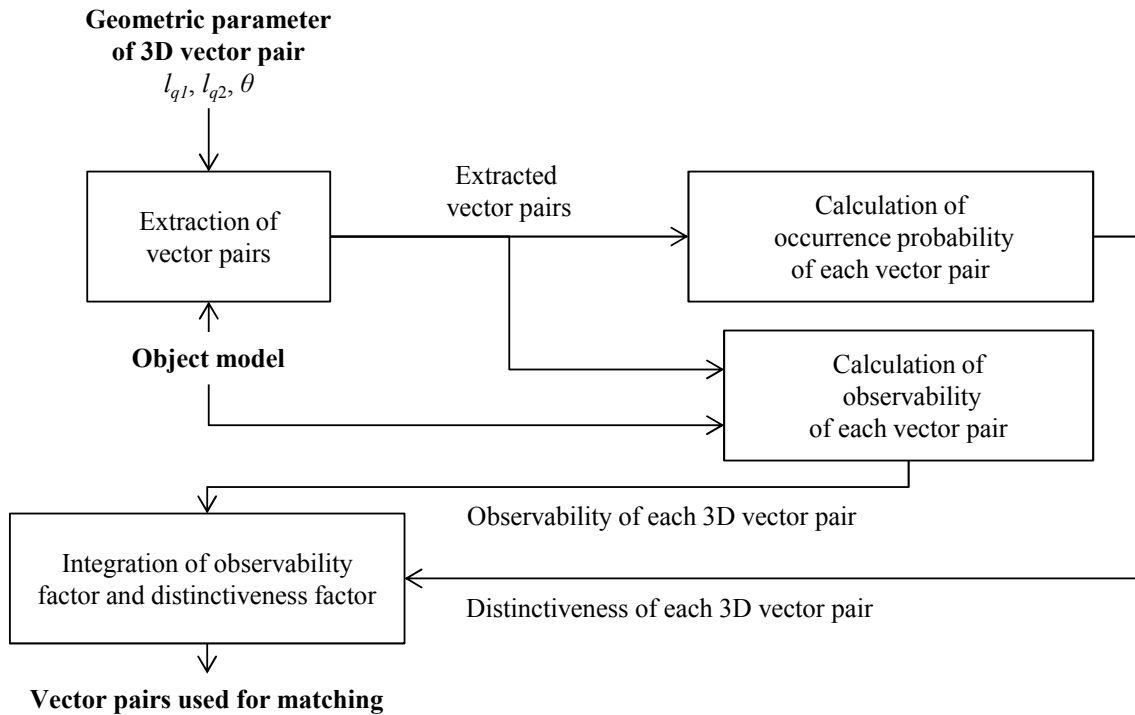


Fig.2.18 Overview of algorithm for extracting 3D vector pairs by integrating observability factor and distinctiveness factor.

Fig.2.18 において、太字は入出力データである。

[**Extraction of vector pairs**] は、設定されたベクトルペアを規定する幾何パラメータに一致するベクトルペアを物体モデルから検出するモジュールである。

[**Calculation of occurrence probability of each vector pair**] は、検出されたベクトルペア特徴の発生確率を算出するモジュールである。

[**Calculation of observability of each vector pair**] は、検出されたベクトルペアの可観測率を算出するモジュールである。

[**Integration of observability factor and distinctiveness factor**] は、ベクトルペアの可観測性と特徴量の発生確率を統合することによって、認識に用いるベクトルペアを選択するモジュールである。

2.5 実験結果と考察

2.5.1 物体モデル

物体モデルとして、レンジファインダで撮影された離散的な 3 次元点群を模擬した、密度一定の物体の全周囲を網羅した 3 次元点群を用意した。このデータは一般的な 3 次元モデルのデータフォーマットである、STL や PLY 形式のデータの三角パッチ内に点群を発生させることによって生成した。

サーフェスモデルの全てのパッチに対して式 (2.29) を満たす 3 次元点 \mathbf{x} をランダムに発生させる。発生させる 3 次元点の数はパッチの面積に応じて、離散点群の密度が一定になるように設定する。

a, b, c
(ただし, (c)
には, 対応
群化の結果

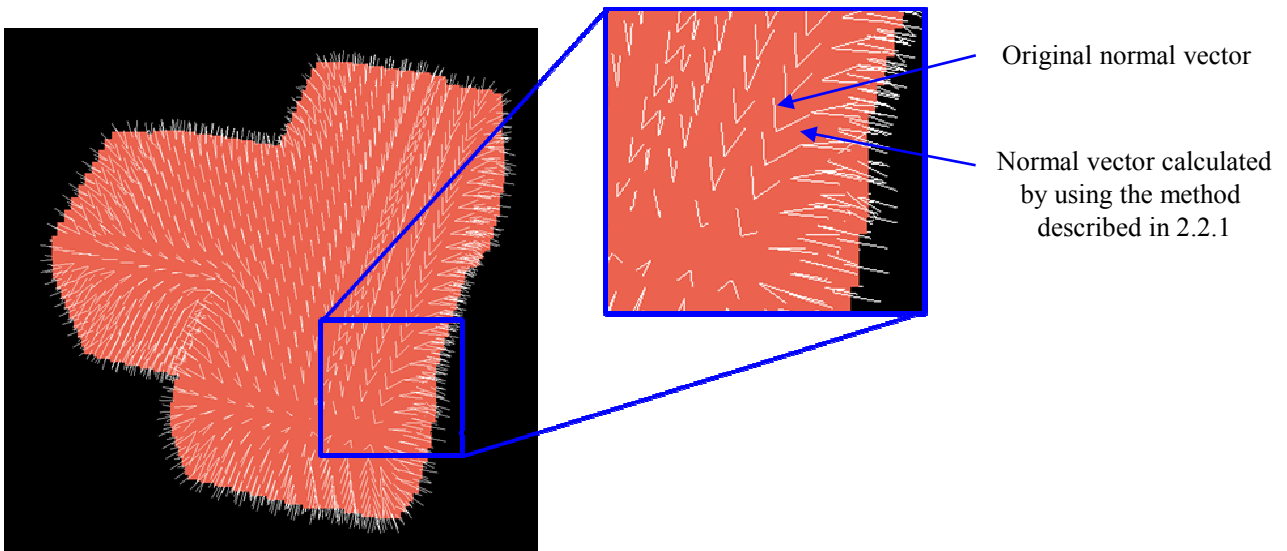
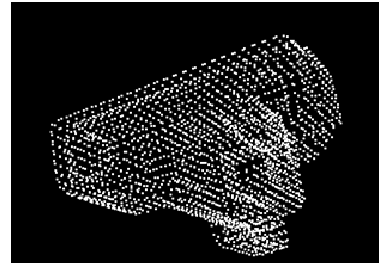


Fig.2.20 Difference of surface normals direction between the original normal vectors and the calculated normal vectors by using the method described in chapter 2.1.1.

物体の縁の部分では, 三角パッチの法線は稜線をまたいで直角に分布している. 2.2.1 の方法を用いて法線ベクトルを計算すると, 法線方向は滑らかに分布する. 物体の表面に対する法線ベクトルとしては前者が正しいが, 物体認識用途においてはモデルシーン間における法線方向の同一性が重要である. 後者

の法線ベクトルは前者の法線ベクトルよりも多くのデータ点を用いて算出するため、ノイズに対する頑健性が高いと考えられる。そこで、物体モデルの法線方向を 2.2.1 の方法の方法を使った法線ベクトルに置き換える。

3 次元点群の共分散行列の固有ベクトルは方向の曖昧性がある。2.2.1 では、視点方向に正対する方向に向きを修正するが、これはレンジセンサ等で取得した片面データに対してのみ有効な方法である。このため、全周囲モデルにおいて視点方向を適切に設定することは困難である。そこで、元のデータが持っている法線ベクトルの向きを正とすることによって、すべてのデータ点に対する法線方向を正しく設定した。

また、STL や PLY 形式の物体モデルデータは稀に法線方向が物体の裏側を向いている場合がある。この場合は上記の方法で法線方向を修正不可能であるため、あらかじめ 3D モデリングソフトを用いて法線方向を反転させる必要がある。本研究ではモデリングソフトとして、Blender[6] を用いた。アプリケーションの外観を Fig.2.21 に示す。Fig.2.21 では、物体モデルを構成する三角パッチの法線が水色で表示されている。Blender では、三角パッチ単位で法線方向を修正することができる。

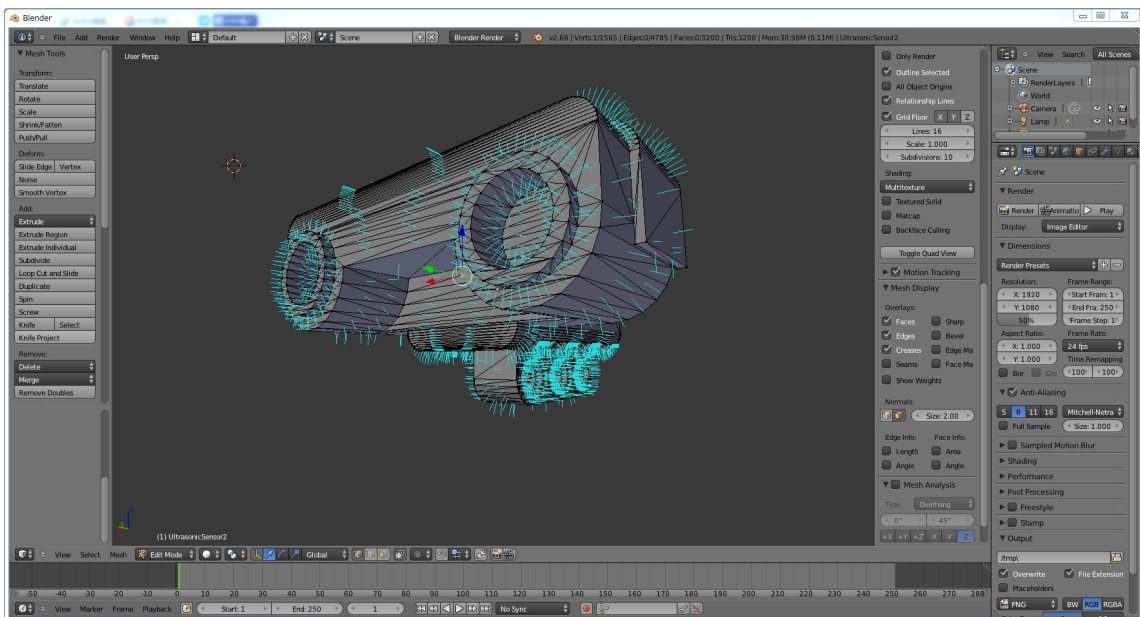


Fig.2.21 Overview of Blender.

2.5.2 入力データ

人工的に作成した同種部品のばら積みシーンと、レンジファインダを用いて実際のばら積みシーンを撮影した 3 次元点群を用いた。入力データは各物体ごとに約 140 例作成した。

人工データ：3D-CG のモデリングソフトである Google SketchUp[43] と、その物理エンジンのプラグインである sketchyphysics[44] を用いて生成した。

SketchUp では .skp という独自のファイル形式を用いて 3D-CAD を作成するが、プラグインの導入により、作成したモデルファイルを STL 形式で出力することが可能である。SketchUp の基本機能は 3D-CG の作成であるが、sketchyphysics の導入により、作成した 3D モデルを使って、重力や物体同士の接触等を考慮した物理シミュレーションを実行することができる。Fig.2.22 に Google SketchUp によ

る箱への部品落下のシミュレーションの様子を示す。

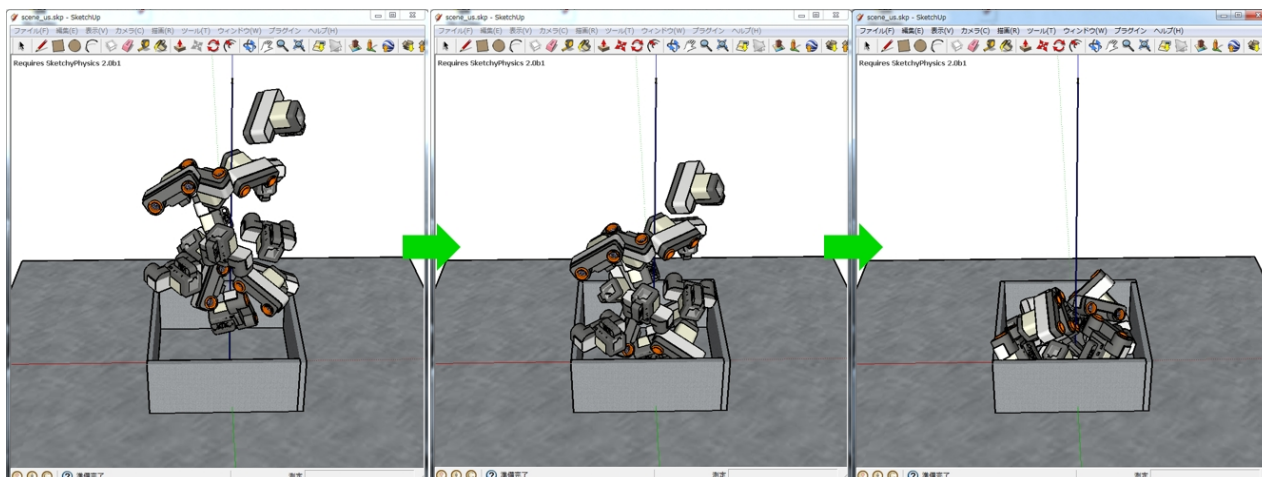


Fig.2.22 Overview of physical simulation by Google SketchUp.

Fig.2.22 では、箱の上方に複数の物体を設置し、物体を落とす様子である。物体同士や、箱の側面と物体が干渉しながらばら積みシーンを形成することができる。

人工データは以下の手順で生成した。

1. モデリングソフトと物理エンジンによる物体モデルのばら積みシーンの再現 (STL ファイル)
2. 重力方向を視線方向ベクトルとした隠面除去による距離画像生成 (画像データ)
3. 距離画像の離散点群化 (0.4[mm] ピッチの 3 次元点群データ)

生成されたばら積みシーン、距離画像、3次元点群データを Fig.2.23 に示す。

実データ：空間コード化レンジファインダである、(株)スペースビジョンの Cartesia 3D Handy Scanner HS-01 (以下、単にレンジファインダと記す)を用いた。このレンジファインダの測定精度は水平垂直 0.3[mm] である。計測された 3次元点群データの例を Fig.2.24 に示す。Fig.2.24(a) は入力シーンの外観であり、Fig.2.24(b) は計測された 3次元点群である。

2.5.3 選択されたベクトルペア

提案手法によって選択されたベクトルペアを Fig.2.25 に示す。本来は点同士を結ぶことでベクトルペアを表現することが望ましいが、見やすさのため点のみの表示とした。比較手法は以下の 3 手法である。

1. 独自性のみのベクトルペア選択手法
2. 可観測性のみのベクトルペア選択手法
3. 可観測性と独自性の統合によるベクトルペア選択手法

独自性だけに注目してベクトルペアを選択する手法では、Fig.2.25(a) の下段に示すように、くぼみ部分から多くのベクトルペアが選択された。これは実験に用いた物体モデルにおいて、くぼみ形状の含有率が低く、そこから抽出されるベクトルペアの特徴量の発生確率が低かったことに起因する。反対に、可観

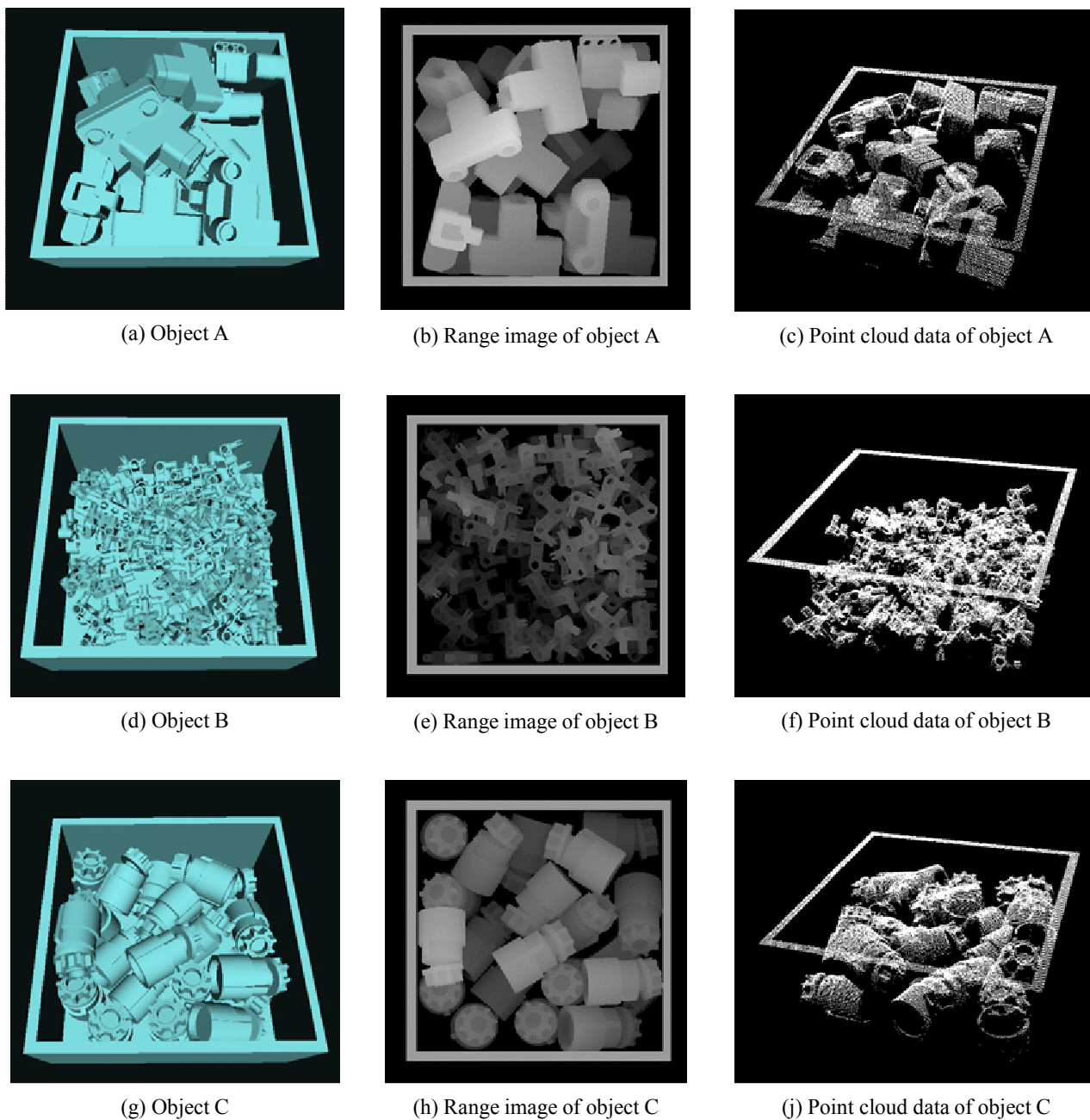


Fig.2.23 Overview of input dataset. (a)(d)(g) Synthetic scenes. (b)(e)(h) Range image of input scenes. (c)(f)(j) Point cloud data of input scenes.

測性のみに注目してベクトルペアを選択する手法では Fig.2.25(b) に示すように、可観測率の低いくぼみ部分からのベクトルペア選択が抑制された。可観測性と独自性の統合によるベクトルペア選択手法では可観測性のみに注目してベクトルペアを選択する手法と同様の選択傾向であり、独自性のみに注目して選択する手法に対して、可観測率の高い物体の縁の付近に多くのベクトルペアが選択された。

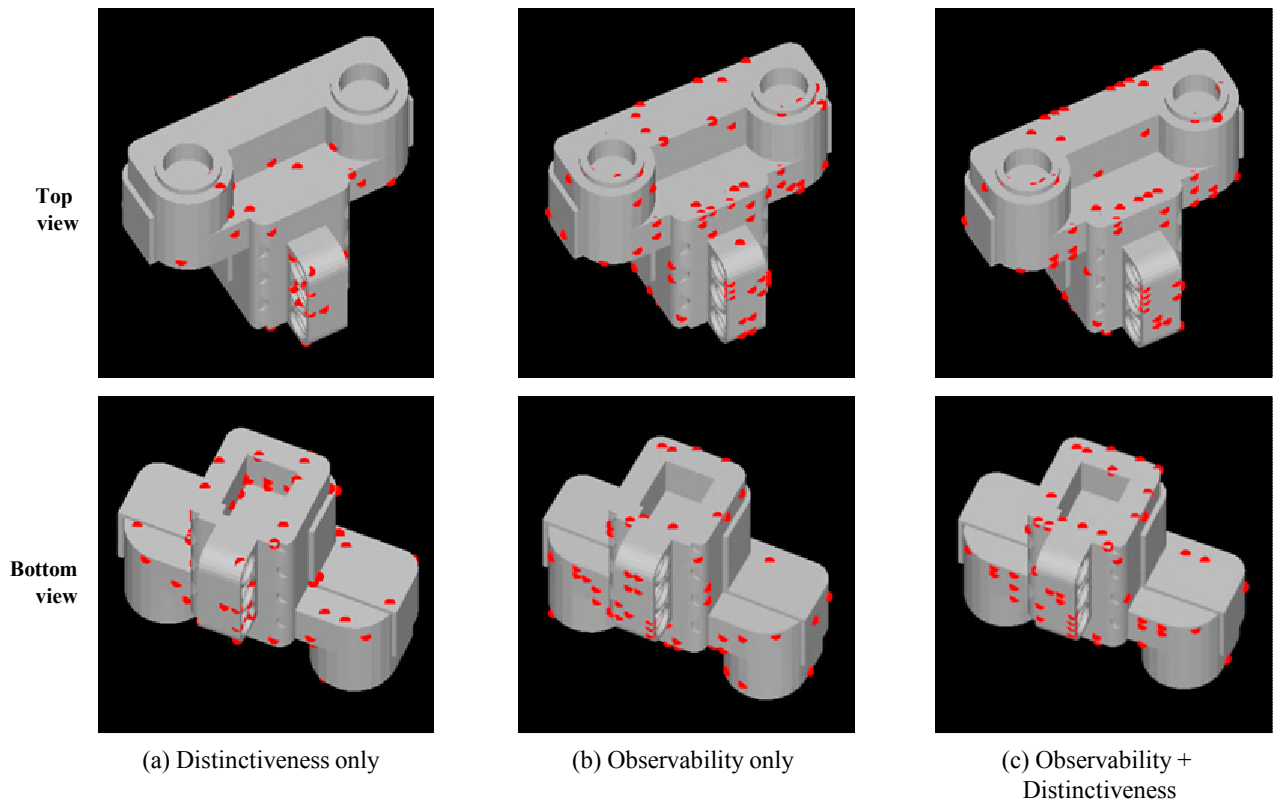
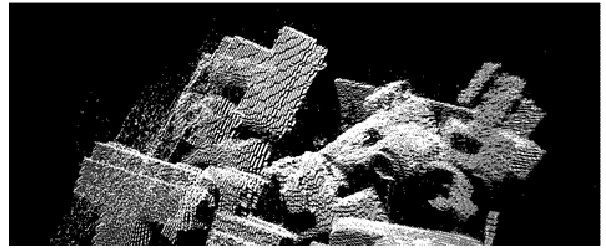


Fig.2.25 Extracted vector pairs by three methods, distinctiveness only, observability only and integration of distinctiveness and observability.

2.5.4 ベクトルペアを規定するパラメータと認識性能の関係

ベクトルペアを規定するパラメータと認識性能の関係を明らかにするために、始点と終点間の距離 l_1 , l_2 を変化させながら認識をおこなった。実験に用いたモデルデータは Fig.2.23 の Object C である。認識は 3D プリンタにて生成した Object C をレンジファインダで撮影した実距離データに対しておこなった。

実験に用いた距離・
 ルペア数は 80 とし
 示す。図中の横軸は
 誤差の平均値であ

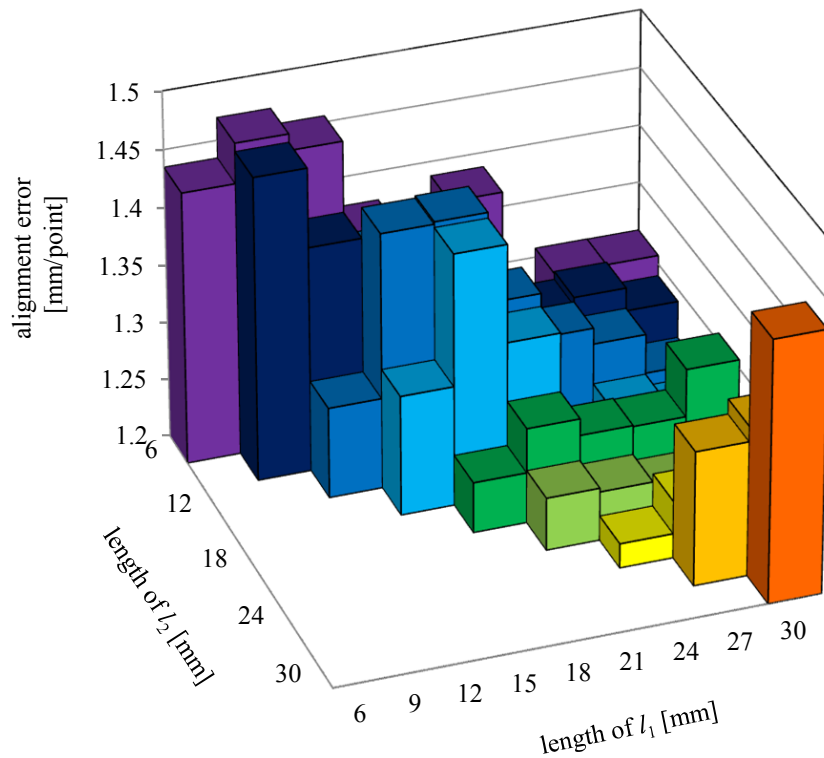


Fig.2.26 Relationship between the length of displacement vector of vector pair and the accuracy of estimated pose.

l_1 , l_2 が短いときは位置合わせ誤差が大きく，長くなるにつれて位置合わせ誤差が小さくなる傾向があった。しかしながら，27mm を越えると位置合わせ誤差が大きくなった。なお，物体モデルは直径約 40mm の球体に内接する大きさである。Fig.2.27 に誤差の大きかったパラメータのベクトルペアと誤差の小さかったパラメータのベクトルペアの例を示す。

Fig.2.27(a) に示すベクトルペアは物体の上端部分と下端部分を同時に選択している。ベクトルペアを構成する点間の距離が大きくなりすぎたことに起因して，全て点が同時に観測しにくくなったことが精度低下の原因と考えられる。このため，ベクトルペアを構成する 3 点が同時に観測されうる範囲で l_1 , l_2 をできるだけ大きい値に設定することで信頼性の高い照合が実現することが分かった。最も誤差の少なかったパラメータは Fig.2.27(b) に示す $l_1 = 24$ [mm], $l_2 = 15$ [mm] であり，位置合わせ誤差平均値は 1.22[mm/point] であった。この時の認識成功率は 90.7% であった。

2.5.5 認識性能の比較

2.5.2 にて作成したデータを用いて認識性能の比較実験を行った。比較手法は以下の 5 手法である。

1. SHOT 特徴量 [48] + Correspondence Grouping 法 [47]

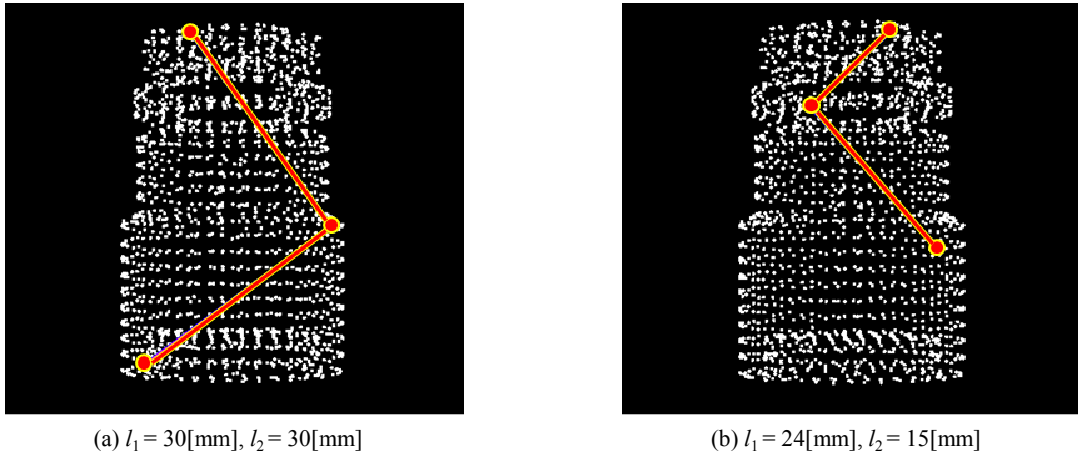


Fig.2.27 Two examples of extracted vector pair. (a) Vector pair with large alignment error. (b) Vector pair with small alignment error.

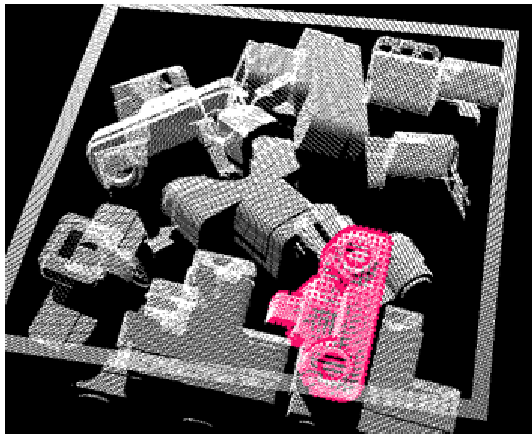
2. FPFH 特徴量 [39] + RANSAC ベースモデルマッチング [33]
3. ベクトルペア (独自性のみ) + ベクトルペアマッチング (以下, VPM と記す.)
4. ベクトルペア (可観測性のみ) + VPM
5. ベクトルペア (可観測性+独自性) + VPM

これらの手法は Point Cloud Library (PCL) [41] を用いて実装した. VPM 法による認識結果例を Fig.2.28 に示す. 白点は入力 3 次元点群を示し, 色つきの物体モデルを認識結果の剛体変換パラメータに基づいて姿勢変換させて入力 3 次元点群に重ね合わせた. また, 認識率と処理時間の関係を表 2.1 に示す. 実験は CPU : Intel®CORE™i7, システムメモリ : 4GB で構成されるシステムでおこなった.

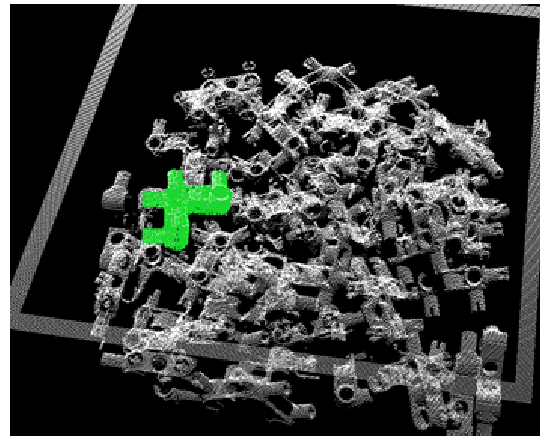
Table 2.1 Relationship between recognition rate Pr [%] and processing time T [sec].

Feature	Matching method		Object A	Object B	Object C	Object D
SHOT[48]	Correspondence Grouping[47]	Pr	51.3	77.9	33.6	20.3
		T	27.78	19.46	33.69	22.48
FPFH[39]	RANSAC[33]	Pr	54.9	11.5	73.5	31.0
		T	0.67	3.97	1.07	0.91
Vector pair (Distinctiveness)	VPM	Pr	40.7	31.0	47.8	33.6
		T	0.27	1.41	0.39	0.72
Vector pair (Observability)		Pr	97.3	100.0	94.7	85.0
		T	0.47	1.76	1.76	1.99
Vector pair (Integration)		Pr	99.1	95.6	92.0	85.8
		T	0.44	1.76	1.42	1.81

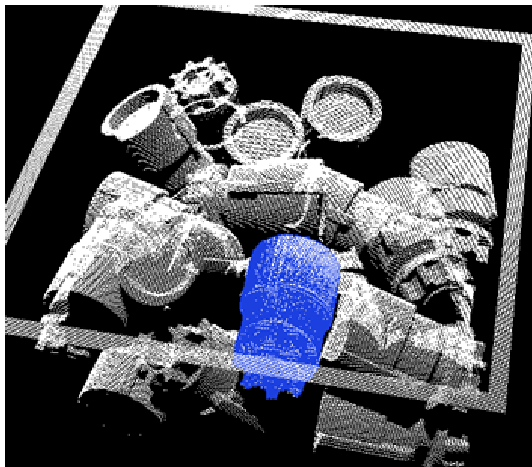
SHOT 特徴量 + Correspondence Grouping 法 : SHOT 特徴量は 352 次元であるため, ベクトルペア特徴と比較すると対応点探索の計算コストが高い. 加えて, 密に物体が配置されたばら積みシーンでは, 特徴量の記述領域に複数物体の距離データを含みやすいため, 正しい対応点が得られにくかった. 結果として, 認識率が他の手法に対して低かった.



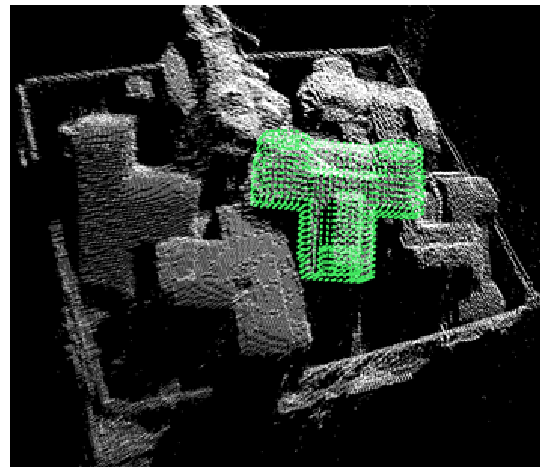
(a) Result of object A



(b) Result of object B



(c) Result of object C



(d) Result of object D

Fig.2.28 Example recognition results for four datasets (superimposed on input point cloud).

FPFH 特徴量 + RANSAC ベースモデルマッチング : この照合手法は数点の対応点のみから姿勢仮説を生成し、RANSACによって精度の高い姿勢仮説を推定結果とする。マージン内に入った点群数をカウントするだけでよいので、高速な認識が可能である。FPFH 特徴量は SHOT 特徴量に比べて特徴量の記述領域の空間的な広がり小さい。しかしながら、数十点のデータ点を特徴記述に必要とするため、注目物体と接触した他の物体の点群データを外乱として含みやすかった。このため、SHOT 特徴量と同様に対応点が得られにくく、認識率が低下したと考えられる。

VPM 法 : VPM 法は他手法に比べて高速に認識することができた。これは、SHOT 特徴量や FPFH 特徴量が 352 次元、16 次元であることに対して、ベクトルペアは特徴次元数が 3 であり、次元数が低いということ、対応点探索は実数ベクトル同士のノルム計算が必要なく、特徴点をテーブル化したデータベースへのアクセスのみで対応点を検出することができるからであると考えられる。独自性のみ注目してベクトルペアを選択する手法 (Distinctiveness) では、物体モデルにおいて含有率の低い局所形状からベクトルペアを選択するため、偽の対応点が削減され、検証が必要な姿勢仮説数を減らすことができた。しかしながら、観測不可能なベクトルペアを多く選択していたことが原因で認識率が低下した。ベクトルペアの選択基準に可観測性を用いた場合 (Observability, Integration) では、独自性のみ注目した選択手法

よりも安定して対応点を検出できるため、認識率が向上した。2手法の認識性能はほぼ同等であった。しかしながら、可観測率のみを利用したベクトルペア選択は、物体モデル中の高可観測率の形状が繰り返しパターンである等のありふれた形状であった場合には、ベクトルペア間の誤対応による不要な投票が増加し、処理時間の増加や認識率の低下を引き起こすと考えられる。したがって、相対手法の運用時は両指標を統一

2.5.1

べ
られ
Fig.2

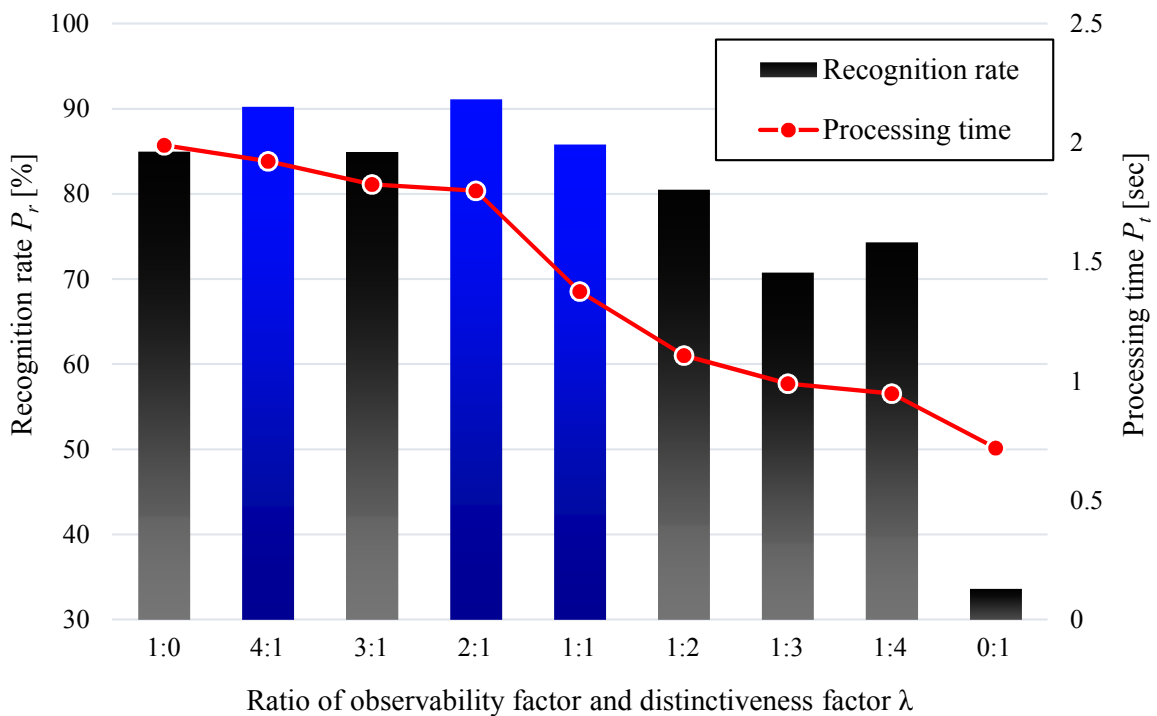


Fig.2.29 Relationship between recognition rate and processing time in the case of various value λ .

可観測性の比率が高い場合は認識成功率が高くなるが、処理時間が増大した。これは独自性を無視してベクトルペアを選択した結果、ありふれた特徴ベクトルを持つベクトルペアが選択され、膨大な姿勢仮説が生成されたことが原因である。反対に、独自性の比率が高い場合には、処理時間は短縮されたが、対応点がシーン中に表れにくく、認識率が低下した。一般的な小型ロボットアームの動作サイクルが2秒程度であり、レンジファインダの撮像速度が0.7秒程度であることから、処理時間は1.3秒以内が望ましい。この条件に当てはまるものは、比率1:1~1:4, 0:1であった。この中で最も認識率が高いものは比率が1:1のときであった。したがって、可観測性と独自性のバランスは1:1のときが実利用上で最もパフォーマンスが高くなると言える。

認識率は可観測率の比率が高い場合に向上する傾向があった。しかしながら、発生確率の低い非常にユニークなベクトルペアがシーン中で観測できた場合は、姿勢決定のための有効性が非常に高いと考えられ

る。この考え方の適用方法として、マッチング時の投票値を調整することが挙げられる。VPM法では、すべてのベクトルペアが可観測率・発生確率に関係なく投票値を1としている。非常にユニークな特徴量を持つベクトルペアによる投票がおこなわれた場合は、投票重みを大きくすることによって、認識結果への影響度を音図的に高めることができ、認識率を向上させることが可能になると考えられる。また、結果



Fig
ICP:

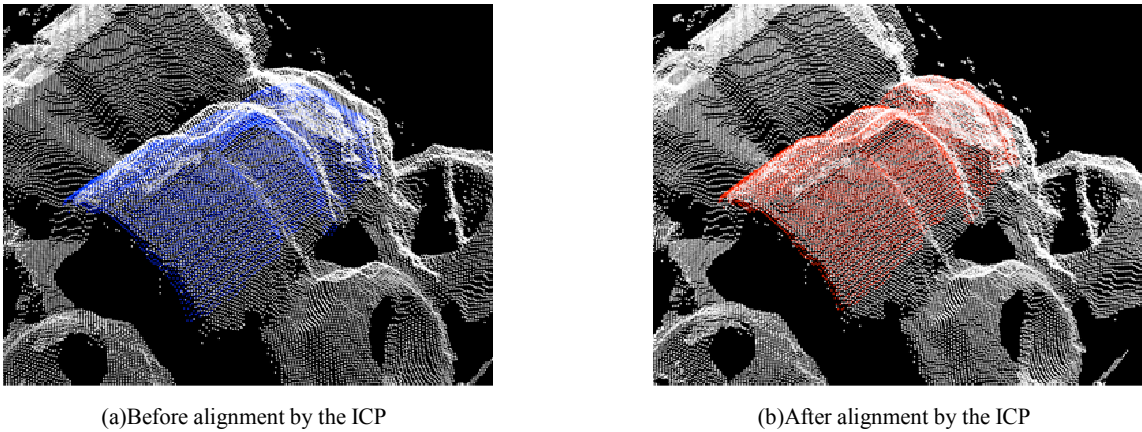


Fig.2.31 Results of pose estimation. (a) Before alignment by the ICP. (b) After alignment by the ICP.

Fig.2.32 において、青は提案手法による誤差平均値であり、赤はICP アルゴリズムの適用後の誤差平均値を示す。ICP 前は重ね合わせ誤差が $0.84[\text{mm/point}]$ であったのに対し、ICP 後は $0.6[\text{mm/point}]$ であり、大きな精度向上は起こらなかった。この実験に用いた距離データの分解能は約 0.4mm であり、センサの計測ノイズを考慮すると、位置合わせ可能な精度は $0.6[\text{mm/point}]$ で頭打ちとなったと考えら

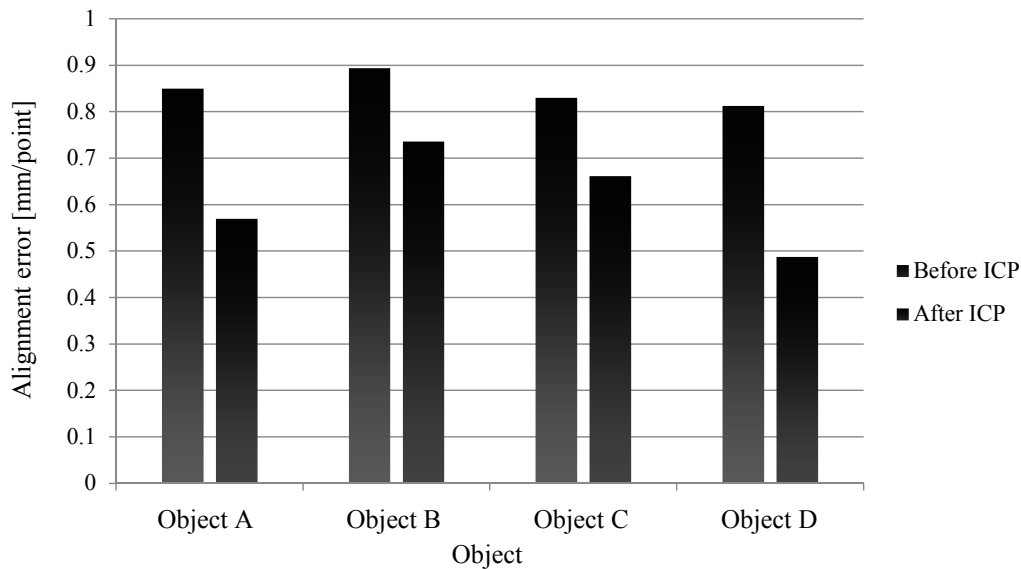


Fig.2.32 Alignment accuracy for each dataset.

れる。

ICP アルゴリズム自体は 0.5~1 秒と比較的処理時間が掛かる処理であるため、特に精度の必要なアプリケーションでなければ必須のモジュールではないと考えられる。

2.6.1 産業用ロボットアームを用いたビンピッキングシステムの構築

レンジファインダとロボットアームを用いて、提案手法を実装したビンピッキングシステムを構築した。詳細を以下に示す。

2.6.2 システムの構成

ビンピッキングシステムの構成は以下の通りである。

1. レンジファインダ：Cartesia 3D Handy Scanner HS-01（株式会社スペースビジョン）
2. ロボットアーム：MELFA RV-1A（三菱電機株式会社製）
3. ロボットアーム制御用コントローラ：CR1-571
4. 制御用 PC：レンジファインダ・ロボットアーム制御用 PC，物体認識用 PC
5. ピッキング対象：LEGO Mindstorms 超音波センサ部品 18 個

システムの構成を Fig.2.33 に示す。

本システムにおける 3 次元点群の取得からピッキングまでの流れを説明する。まず、レンジファインダにてシーンの 3 次元点群を取得する。この 3 次元点群データをレンジファインダ・ロボットアーム制御用 PC から、物体認識用 PC へ送信し、物体の位置姿勢を認識する。次に、物体の位置姿勢パラメータをレンジファインダ・ロボットアーム制御用 PC へ送信し、ロボットアームがアプローチすべき手先の位置姿勢パラメータに変換し、ロボットアーム制御用コントローラへ送信する。最後にロボットアームは指令通りに物体へアプローチし、物体をピッキングする。

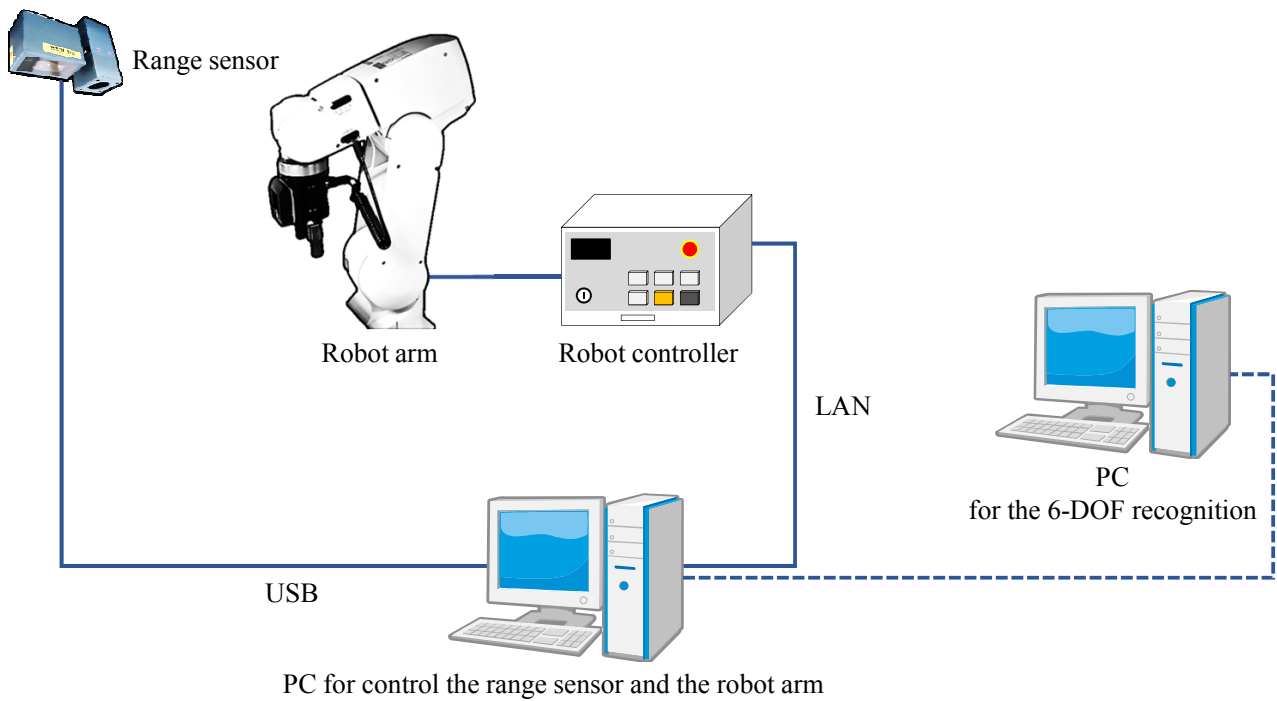


Fig.2.33 Configuration of the bin-picking system.

ピンピック
設置した。

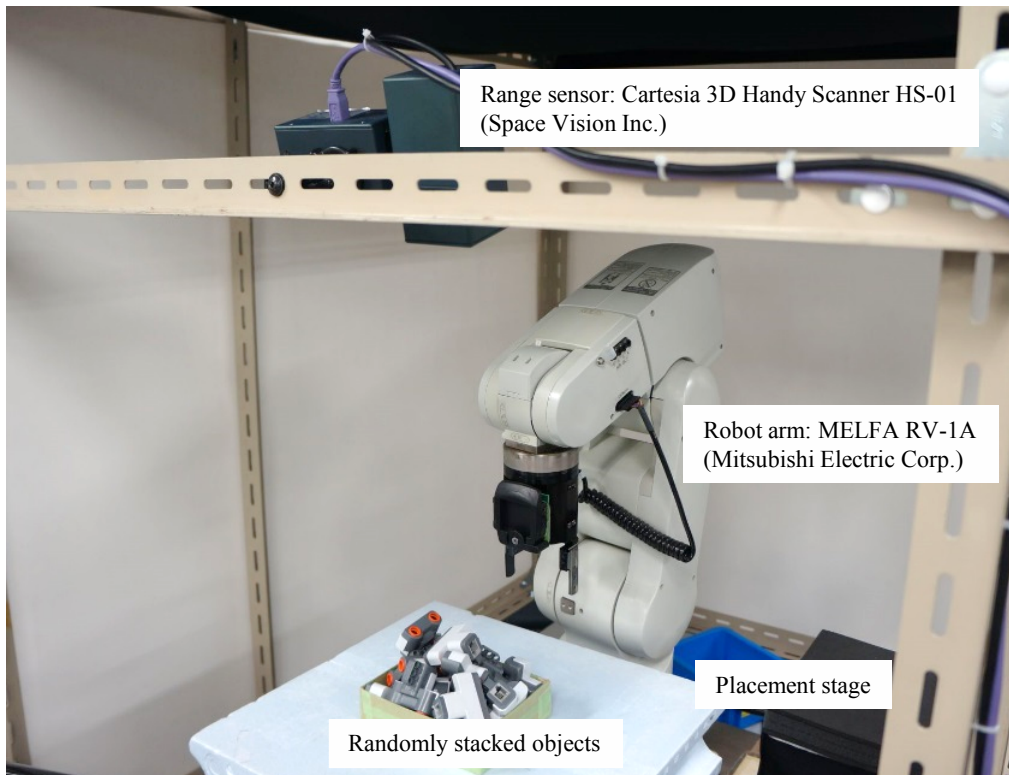


Fig.2.34 Overview of the bin-picking system.

2.6.3 レンジファインダとロボット間のキャリブレーション

レンジファインダのカメラ座標系とロボット座標系のキャリブレーション方法を説明する。レンジファインダのカメラ座標系とロボット座標系のキャリブレーション方法を説明する。レンジファインダのカメラ座標系とロボット座標系のキャリブレーション方法を説明する。

Fig. 2.35 なる補群データその高い

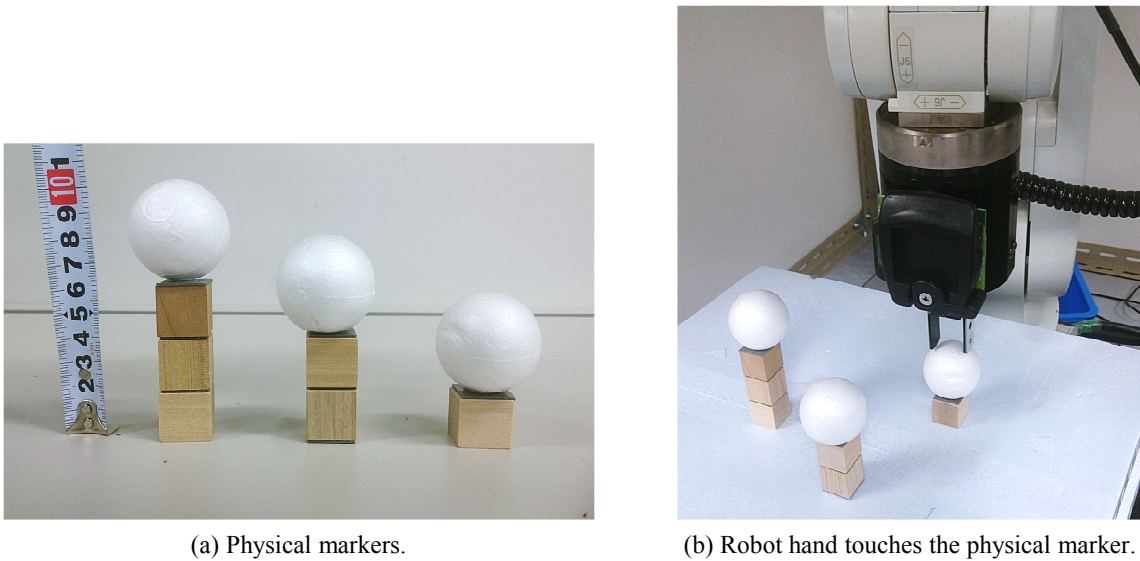


Fig.2.35 Overview of how to calibrate the robot coordinate system and the range sensor coordinate system. (a) Physical markers which are used for calibration. (b) Robot hand touches the physical marker.

取得した 3 次元座標をそれぞれ行列として，式 (2.30) で表す。

$$\mathbf{A}_{Cam} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}, \quad \mathbf{A}_{Robo} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \quad (2.30)$$

$\mathbf{T}_{Cam2Robo}$ は \mathbf{A}_{Cam} の疑似逆行列を用いて，式 (2.31) によって算出した。このキャリブレーションでは，対応点を 12 点用いた。

$$\mathbf{T}_{Cam2Robo} = \mathbf{A}_{Cam}^{-1} \mathbf{A}_{Robo} \quad (2.31)$$

認識された部品の位置を \mathbf{t}_{Cam} とすると，ハンドがアプローチすべきロボット座標系での位置 \mathbf{t}_{Robo} は， $\mathbf{t}_{Robo} = \mathbf{T}_{Cam2Robo} \mathbf{t}_{Cam}$ である。手先姿勢はロール・ピッチ・ヨーで指定する必要がある。このため，認識された部品の回転成分 \mathbf{R}_{Robo} を式 (2.32) (2.33) によって変換する。

$$\mathbf{R}_{Robo} = \mathbf{T}_{Cam2Robo} \mathbf{R}_{Cam} = \begin{bmatrix} r_{00} & r_{10} & r_{20} \\ r_{01} & r_{11} & r_{21} \\ r_{02} & r_{12} & r_{22} \end{bmatrix} \quad (2.32)$$

$$\begin{aligned} roll &= \tan^{-1}(r_{12}/r_{22}) \\ pitch &= \tan^{-1}(-r_{02}/\sqrt{r_{12}r_{12} + r_{22}r_{22}}) \\ yaw &= \tan^{-1}(r_{01}/r_{00}) \end{aligned} \quad (2.33)$$

以上によって計算された手先位置・姿勢パラメータをロボットコントローラへ送ることで部品ピッキングを実現した。

2.6.4 部品のピッキング

ロボットアームによる部品ピッキングの様子を Fig.2.36 を用いて説明する。

Fig.2.36 では、ピッキング動画において、主要なフレームを切り出して並べてある。各フレームの右上にレンジファインダで取得した 3 次元点群を表示した。白点が入力 3 次元点群であり、色付きの点で認識結果を重畳した。

Fig.2.36(a) は初期状態であり、レンジファインダによって入力 3 次元点群を取得した状態である。

Fig.2.36(b) は認識結果を示している。得られた姿勢仮説のずれが 1 点当たり 4.5[mm] 未満である場合に色付き点群での認識結果の重畳をおこなった。この入力 3 次元点群では 2 つの物体が認識された。認識時間はおよそ 2 秒であった。

Fig.2.36(c) は認識結果に基づいて、ロボットアームが物体に対してアプローチを開始したフレームである。赤点で示した認識結果に対してアプローチしている。

Fig.2.36(d) は認識した物体の直上へのアプローチが完了したフレームである。直接把持する位置まで手先を動かすと、ほかの部品へ干渉する恐れがあるため、物体直上 50[mm] の位置にアプローチすることとした。この際、認識結果の回転成分 \mathbf{R} を用いてロボットアームの手先の姿勢（ロール・ピッチ・ヨー）を算出している。

Fig.2.36(e) では (d) のポジションから物体直上から 50[mm] 接近している。この部品は斜め向きに置かれているが、ハンドもその向きに合わせてアプローチできていることがわかる。

Fig.2.36(f) ではハンドを閉じ、物体を把持している。精度よく物体の位置姿勢を認識できたため、ハンドを閉じた時に物体自身や、他の物体を動かすことがなかった。

Fig.2.36(g)~(j) は、把持した物体を持ち上げ、隣の台に移している様子である。ここについても大きな問題なく実行することができた。

Fig.2.36(k) はロボットアームが次の部品をピッキングするために、初期位置に戻る様子である。

Fig.2.36(l) では再度レンジファインダでシーンのデータを取得している。ピッキングした部品の点群データが無くなっていることがわかる。

以上のように、提案手法を用いて認識した部品の位置姿勢は、実際のロボットアーム物体をピッキングできる程度に精度が高いということが確認できた。

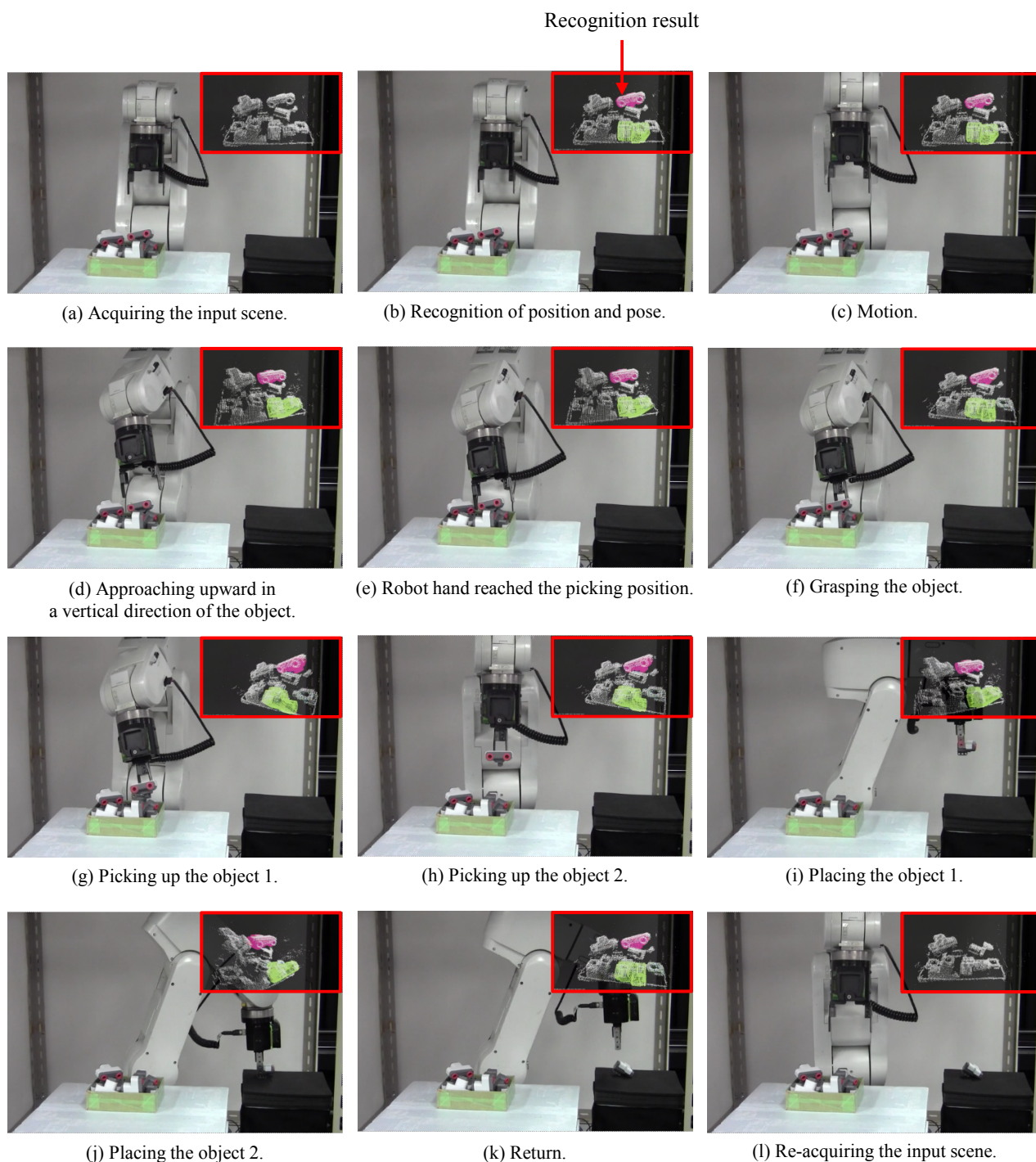


Fig.2.36 Overview of the bin-picking by using the robot arm. Principal frames are displayed. Point cloud data are attached upper right side of each frame. Recognition results are displayed by colored points.

2.7 結論

本章では、凹凸部分を有する物体に対して高速に3次元位置姿勢を認識するためのモデルマッチング法を提案した。3次元位置姿勢を認識するための最小データは3点の3次元点の組であることを説明し、これをベクトルペアと呼ぶ新たな特徴量の表現方法として提案した。さらに、物体モデルを事前分析するこ

とによって選択した、特徴量の独自性と可観測性を備えた少数のベクトルペアを用いることによって、高速な認識が可能なことを示した。

以下に、本章で明らかになった結論を要約する。

- (1) 物体の3次元的位置姿勢の決定のための最小データセットは3点の3次元点であること、およびこの3点を共通した一つの始点と、二つの独立した終点からなるベクトルペアとして表現することを説明した。また、ベクトルペアを照合単位とすることを説明した。照合に用いるベクトルペア数を数%~10%程度に厳選することによって、ピンピッキングの一般的なタクトタイム(2秒)以内での認識が可能なことを示した。このときのベクトルペア数の選択基準として、独自性と呼ぶ指標を導入し、物体モデルにおけるベクトルペア特徴から、低頻度(低発生確率)な特徴量を持つベクトルペアは誤認識のリスクが低いことを説明した。
- (2) ベクトルペアを用いた照合方法として、一般化ハフ変換ベースの方法を採用することによって、部分的なオクルージョンに頑健な照合が可能であることを示した。対応点探索時に事前に物体モデルから選択されたベクトルペアをテーブル化しておくことによって、ノルム計算することなく高速に対応点を生成可能なことを説明した。また、ベクトルペアを直交に設計しておくことで、姿勢仮説の生成コストを低減する計算方法について説明した。
- (3) 物体モデル入力シーン間において、信頼性高く対応点を生成するためのベクトルペア選択基準として、可観測性と呼ぶ指標を提案した。さまざまな視点からの物体モデル表面の観測可能性を計算することによって安定して観測可能な特徴点を選択する計算方法を提案し、可観測性と独自性の2つの指標を組み合わせたベクトルペアの選択方法を説明した。可観測性と独自性の高さを両方兼ね備えたベクトルペアが、独自性のみを基準として選択したベクトルペアによる照合よりも、認識信頼性が高くなることを示した。
- (4) 4種類のデータセットを用いて従来の3次元位置姿勢認識手法である、SHOT特徴量やFPFH特徴量を用いた手法と比較したところ、認識率が45.8%から93.1%に向上することが確認された。さらに、処理時間は平均1.36秒であり、ロボットアームの動作サイクルと比較して十分高速であることを確認した。可観測性と独自性の統合重みに関しては、可観測性を重視した際は認識率が向上し、独自性を重視した場合は処理速度が向上することが分かった。この重みは1:1のときに認識率と処理時間のバランスが最も高いことを明らかにした。提案手法による位置姿勢認識後に、ICPアルゴリズムを適用することで、推定精度が約0.2[mm/point]改善し、0.6[mm/point]を達成することが確認された。6軸小型産業用ロボットアームと空間コード化法によるレンジファインダを用いて、ピンピッキングシステムを試作し、提案手法にて認識したばら積みされた部品が実際にピンピッキング可能であることを実証した。

本章に係る成果は、以下の2件の学術論文として公表済みである。

1. 秋月秀一, 橋本学, “特徴的 3-D ベクトルペアを用いたばら積み部品の高速位置姿勢認識”, 電気学会論文誌 C, Vol.133, No.9, pp.1853-1854, 2013.
2. Shuichi Akizuki, Manabu Hashimoto, “Stable Position and Pose Estimation of Industrial Parts using Evaluation of Observability of 3D Vector Pairs”, *Journal of Robotics and Mechatronics (Special Issue on Vision and Motion Control)*, Vol.27, No.2, pp.174-181, 2015.

第3章 照合戦略の適応的選択に基づく対象物の凹凸形状に依存しない物体仮説生成

3.1 はじめに

ビンピッキングシステムにおいて、連続した Pick and Place (Insert) 動作を効率よく実行するためには、システムの上部に固定されたレンジファインダから得られたワンショットの距離データのみを用いて、部品の位置姿勢を認識することが望ましい。部品の距離データは形状や姿勢によって、凹凸を多く含む場合や、反対に平面的になる等、さまざまなパターンが想定される。したがって、これらに依存しない安定的な認識手法が必要である。これまでに3次元の位置姿勢認識手法として、認識対象物のCADモデルから選択した局所特徴量を用いた認識手法である、局所特徴ベースの照合法が数多く提案されてきた [8, 29, 39, 40, 48, 56, 59].

これらの手法は、物体モデルの局所的な形状を分析することによって得られる少数の特徴点のみを照合に用いるため、処理効率が高く、高速である。特徴点の選択基準としては、曲率の大きな点を採用する Depth Aspect Image 法 [59] や、Shape Index[11] の極値を利用する Local Surface Patches 法 [8] などの、形状情報を特徴点決定の基準とする手法がある。また、物体モデルにおいてアウトライアとなる特徴量を算出する点を特徴点とすることによって、誤対応を低減する PFH[40], FPFH[39] や、特徴量の再現性が高くなる点を特徴点とする SHOT 法 [48] がある。

しかしながら、これらの特徴点検出に依存した局所特徴ベース照合法では、レンジファインダから得られた認識対象物の距離データ（以下、単に見えと呼ぶ）が形状的な起伏を持たない場合には、特徴点検出が困難になることに起因して、認識信頼性が低下するという問題がある。一般的な製造ラインで扱われる工業部品では、平面的な面を持つ物体が少なくなく、このような見えに対する照合性能を無視することはできない。

また、SHOT 法 [48] に代表される形状記述子には、特徴量の記述領域の3次元的なオリエンテーションである、Local Reference Frame(LRF) が関連付けてあり、入力シーンとの対応点の LRF との変位量を解消する剛体変換パラメータを、認識対象物の位置姿勢としている。LRF は xyz の3軸の独立した方向ベクトルを局所形状に割り当てることによって決定されるが、安定的な認識のためには、これらの方向ベクトルは物体同士の姿勢が異なる場合においても再現性が高いことが期待される。

LRF の生成手法には、局所領域内の点群の主成分分析から得られる固有ベクトルを利用する手法 [29, 48] や、 z 軸を面法線とし、 y 軸には SIFT[28] のように勾配方向の支配的なものを割り当てる手法 [56] 等がある。また、 x 軸は、 z 軸と y 軸の外積ベクトルとして算出される。しかしながら、これらの LRF の生成手法は、領域内の点群の分布が3次元的な広がりを持っていることを仮定している。例えば、領域内の点群の分布が平面的な場合には、 z 軸の接平面方向の点群分布の偏りが少ないことから、LRF を再現性よく決定することは困難である。

本研究の目的は、平面のような特徴点を検出することが困難であったり、LRF の再現性が低くなる見えを持つ物体に対しても実用的な認識性能を有する手法を提案することである。

認識対象物の見えが平面的であったとしても、外輪郭の形状には、姿勢同定のための情報が存在すると

考えられる。すなわち、見え全体を捉える輪郭を形成する点群の分布には、特定方向の偏りがある。そこで、本研究では大域的な情報を用いて、新たに平面的な見えのための3次元なオリエンテーションである Global Reference Frame (GRF) を算出し、位置姿勢認識に利用する。GRF の z 軸は、平面的な見えが持つ法線方向のうち、支配的なものを割り当てる。 y 軸は、距離データを z 軸の接平面に投影した距離画像を生成し、外輪郭を短い線分（線素）群で近似した際の方向分布の支配的なものとする。これにより、認識対象物の見えが平面的な場合においても、外形の情報を利用することによって安定的な RF を生成することが可能である。

なお、認識対象物の見えが特徴点を検出できる程度の凹凸を持つ場合は、従来通り、処理効率の高い局所特徴ベース照合法が適用可能と考えられる。そこで、提案手法では、入力シーンにおける処理対象領域を「平面」、「非平面」の2カテゴリに自動分類し、これらの領域に適した相補関係にある照合方式を切り替えることによって、結果として見えに依存しない位置姿勢認識を実現する。GRF 照合と組み合わせる局所特徴ベース照合法には、高速な VPM 法を用いる。

3.2 平面的な見えに対応した認識手法の重要性

工業部品は複数種類の製品に共通的に使うために、単純な形状によって構成されることが多い。したがって、それらをレンジファインダで撮影した場合は、平面的な見えになることが少なくないと推測される。そこで、本章ではこのような見えに対応した照合方式の重要性を確認するために、一般的に用いられる工業部品の見えと、平面との類似性の関係を調査した。

そこでレンジファインダで撮影した距離データを想定し、物体の CAD モデルに対してランダムに設定した視点からの距離データを人工的に生成した。そして、それぞれの距離データに対して RANSAC によって平面フィッティングをおこない、平面との類似度 S_p を式 (3.1) として定義し、 S_p の分布を調べた。

$$S_p = (\text{平面上の点数}) / (\text{距離データの全点数}) \quad (3.1)$$

(平面上の点数)は RANSAC によって得られた平面パラメータのインライア点数である。例えば $S_p = 0.5$ は、見えの距離データの 50% が同一平面上に存在することを表している。また、各 S_p 値を持つ距離データに対して、局所特徴ベース照合である VPM 法で用いる特徴点検出を適用し、検出される平均特徴点数を調べた。

多種類の部品を用いて実験をおこなうために、本実験では膨大な現実の工業部品を集めた CAD モデルデータセットである、Engineering Shape Benchmark (ESB)[22] から対象部品を選定した。ESB は、“Clips”, “T shaped parts”, “Bolt like parts” のように、形状ごとに分けられた 45 クラスの CAD モデル合計 867 個で構成される。クラスごとに登録モデル数が異なるため、1 クラスから 3 モデルをランダムに選び、135 モデル \times 100 視点の距離データを生成した。実験に用いた ESB の CAD モデルの例を Fig.3.1 に示す。Object A~C は丸い穴を有するため、形状の一部が円筒によって構成されている。この形状から得られる距離データは、平面による近似が困難であるため、 S_p が低い非平面的な形状にカテゴライズされると考えられる。また、Object B の両端や Object E は平面的な形状によって構成されている。したがって、これらの形状から得られる距離データは、 S_p の高い平面的な形状にカテゴライズされると考えられる。

生成した距離データに対する S_p の分布と特徴点数の関係を Fig.3.2 に示す。棒グラフが S_p に対応する距離データの頻度、折れ線グラフが各 S_p における平均特徴点数を示している。Fig.3.2 に付属されている距離画像は、Fig.3.1 の CAD モデルから生成した距離データと対応している。

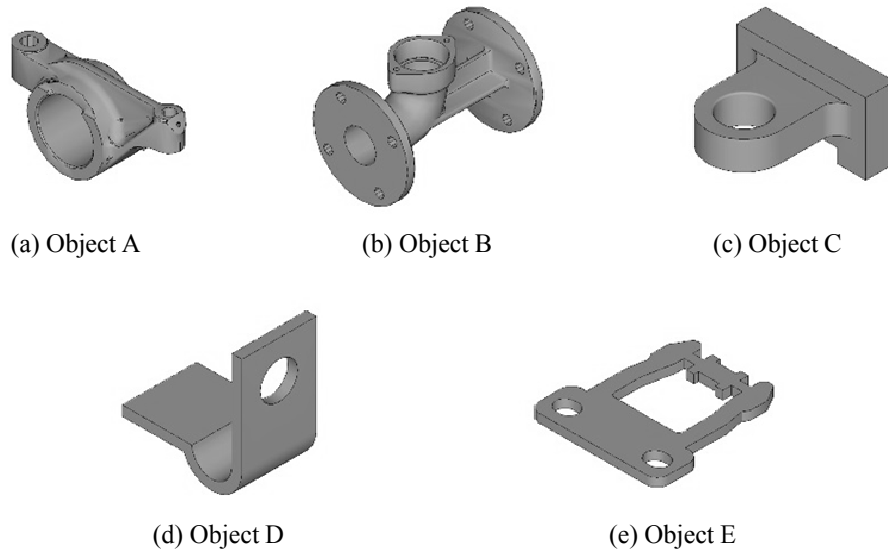


Fig.3.1 Examples of object models of the ESB

Object A は 100 視点の平均 S_p が 0.22 であった。2 割程度の距離データのみが同一平面上に分布していたことから、いずれの視点において非平面的といえる。

一方で、もともと平面的である Object E は平均 S_p が 0.72 であり、多くの視点において平面的な見えであった。また、 S_p が 0.11~0.81 である Object B のように、視点位置によって S_p が大きく変化する部品も存在することが分かった。さらに、全体として 10.1% の距離データが $0.8 < S_p$ であった。

結論として、一般的な工業部品の位置姿勢認識において、平面的な見えがレンジファインダから得られる状況は無視できないほど多いといえる。また、検出され得る特徴点数は、 S_p が高くなるにつれて減少した。これには、完全な平面や若干の凹凸情報を有する距離データが該当する。一般的に、入力データから得られる特徴点数が減少すれば、期待される認識率も低下するため、距離データが平面的である場合は、局所特徴ベース照合では認識信頼性が低下するといえる。

以上の一連の実験およびその結果の考察から、実工業部品を想定した位置姿勢認識アルゴリズムは、平面的な見えに対しても頑健な照合機能を有することがきわめて重要であることが明らかになった。次節以降に、この機能を実現する新たな GRF 照合および、これを統合した照合戦略切り換え型位置姿勢認識手法を説明する。

3.3 Global Reference Frame を用いた姿勢仮説の生成

これまでの代表的な 3 次元オリエンテーションである Local Reference Frame は、3 軸の直交した (x, y, z) の方向ベクトルによって構成されているが、サポート内 (Local) の点群の分布が平面のような 2 次元的な広がりしか持たない場合には、面法線 (z 軸) 以外の方向ベクトルを規定できない。したがって、このような場合には 3 次元的な姿勢推定が困難である。一方、物体の外形のような大域的な形状データの分布には、特定方向の偏りが存在する (長手方向が存在する) と考えられる。本論文では、このような距離データの大域的な (Global) 情報を用いた 3 次元的なオリエンテーションである Global Reference Frame (GRF) を提案する。

GRF は処理対象領域の代表法線と、その接平面における支配的な方向ベクトルによって構成されてい

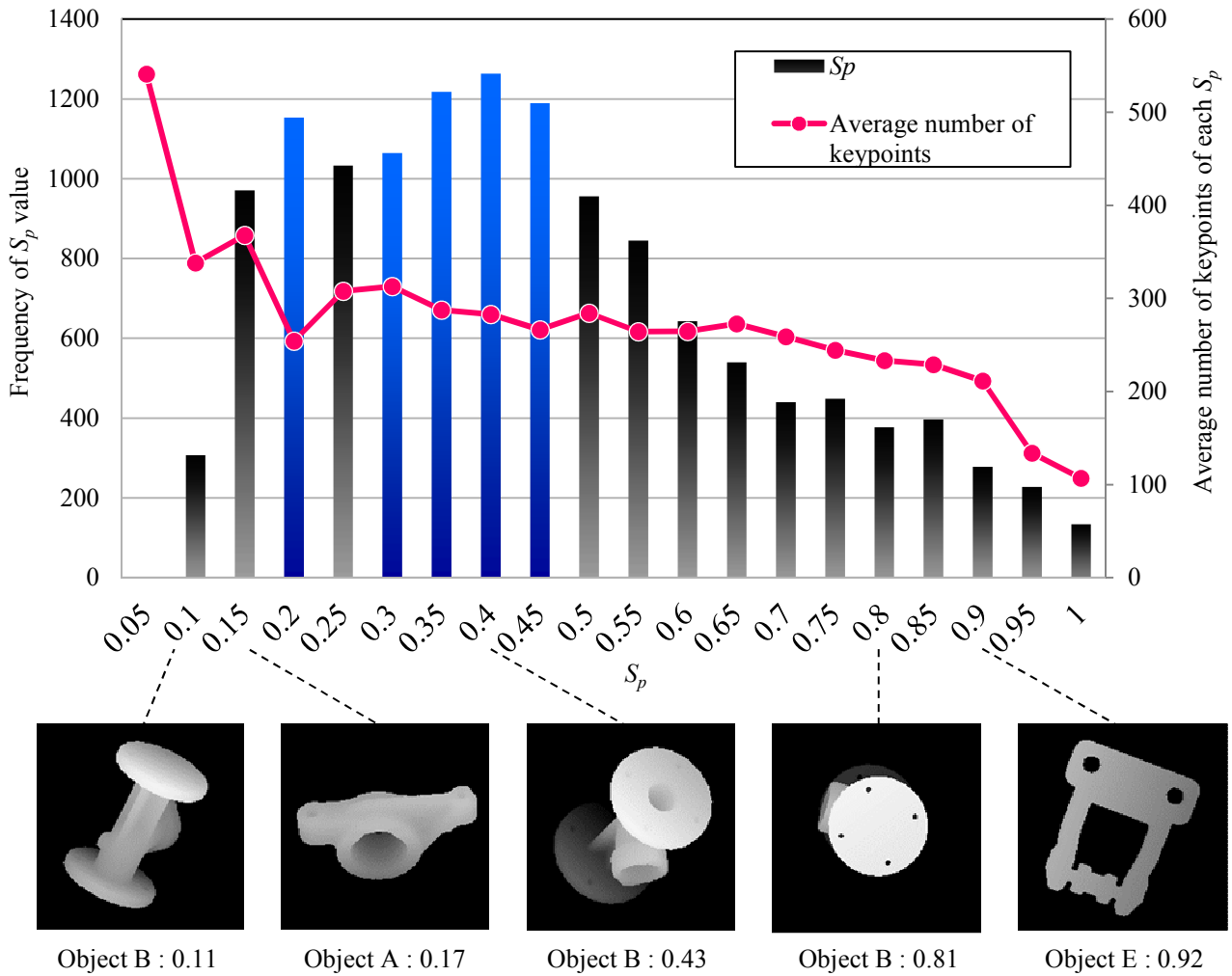


Fig.3.2 Distribution of S_p values and the number of keypoints which are extracted from range images

る。物体モデルと入力シーンから得られた GRF 間の変位量である剛体変換パラメータを姿勢仮説とする。以下に、GRF の生成方法と姿勢仮説の算出手法について述べる。

3.3.1 GRF の生成

GRF を構成する軸の生成方法を説明する。まず、距離データ各点の法線を EGI (Extended Gaussian Image) [21] 上にプロットする。次に EGI 上のプロットの集中性を評価する。EGI 上で最も集中しているプロットに対応する法線がその距離データにおいて支配的であるとみなし、代表法線とする。これが、GRF における z 軸である。

Fig.3.3 に同一の物体を 2 種類の視点から観測した際の距離データと、それぞれから算出された代表法線を示す。青線が各点の法線であり、赤矢印が代表法線を示す。(a), (b) は姿勢の異なる距離データであるため、欠損している形状データが異なっているが、支配的な平面は共通であるため、どちらの距離データからも共通した代表法線を算出することが可能である。

つぎに、 z 軸が持つ接平面上に距離データを射影し、距離画像 f を得る。 f は代表法線から観測した物体の外形を表しており、特定方向のデータの偏りを持つと考えられる。この接平面上における、データの偏りの方向 $o(f)$ が GRF の y 軸である。

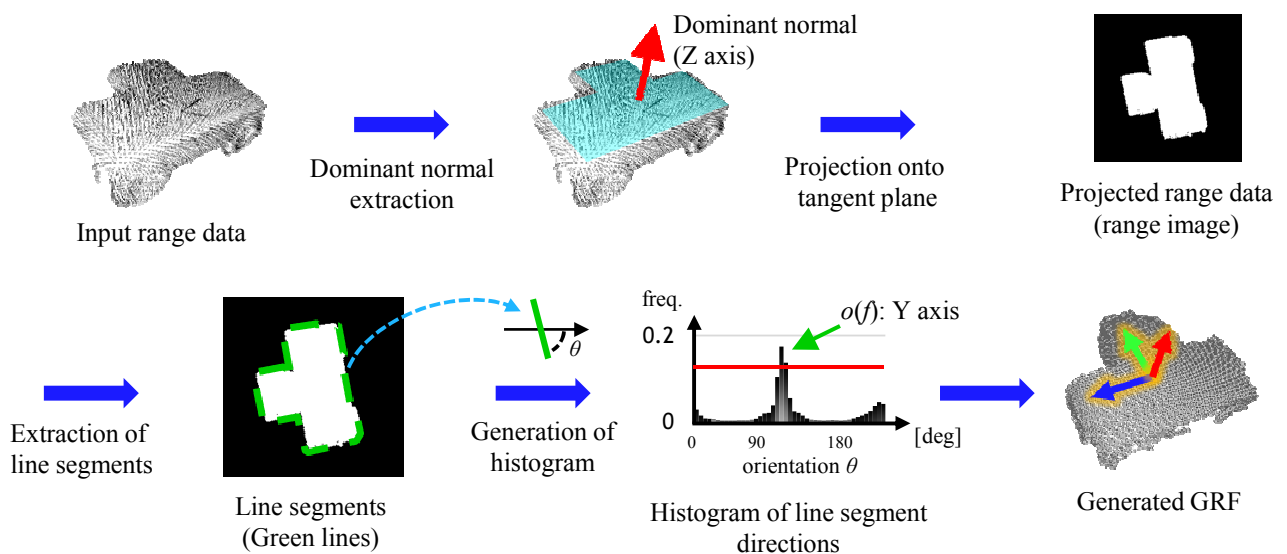


Fig.3.4 Overview of calculation of the $o(f)$

Fig.3.4に $o(f)$ の算出の概要を示す。 $o(f)$ の算出には、距離データの外輪郭から得られる線素群の方向分布を用いる。まず、距離画像 f から輪郭を検出する。そして輪郭点を数画素おきにサンプリングしながら線素を抽出し、画像平面上の傾き θ を計算し、 θ の頻度分布を作成する。このとき、ノイズの低減のために、ヒストグラムには $\mu = \theta$ を中心としたガウス分布を投票し、しきい値以上の得票の角度を $o(f)$ とする。しきい値を越える角度が複数存在するときは、それらすべてを $o(f)$ として登録する。さらに、ヒストグラムのピークを補完することによって $o(f)$ を高精度化する。

以上の処理により、平面的な距離データに対しても、3次元オリエンテーションを算出することができる。また、GRFを構成する方向ベクトルは、統計的なデータ分布解析によって得られた、形状データの支配的な法線方向と外形の偏り方向から構成される。したがって、これらの方向ベクトルはオクルージョン等による形状データの微小な欠けに対して頑健である。

3.3.2 物体モデルの形状特徴に合わせた GRF 数の自動決定

信頼性の高い認識のために、物体モデルから複数のGRFを事前に生成しておき、認識に用いる。これには、GRFを設定するために必要な平面形状の面積をあらかじめ設定しておくことによって、その個数

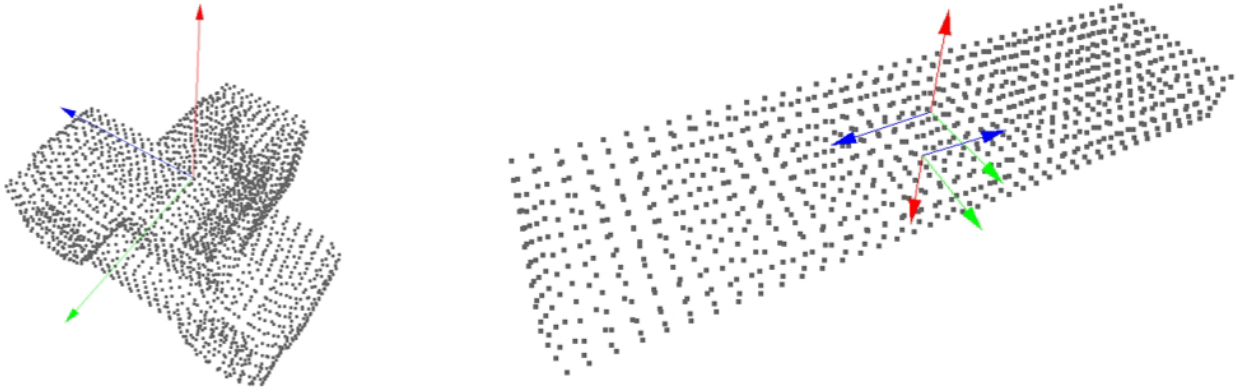


Fig.3.5 Examples of GRF for two objects.

を自動的に決定する。Fig.3.5 に 2 種類の物体モデルに対して設定した GRF の例を示す。

左は物体の上面に支配的な平面領域が存在する物体における GRF である。この平面領域以外は、凹凸形状が優勢であるため、他の GRF は生成されなかった。右は直方体の物体モデルにおける GRF である。この物体は上面と底面が支配的な平面であるため、それぞれに一つずつ GRF が設定された。以上のように、提案手法では、物体の形状に合わせて自動的に GRF の数を決定できる機能を有するため、認識の前にあらかじめ複数の GRF を用意することは容易であるといえる。

3.3.3 姿勢仮説の生成

GRF を用いた認識対象物の位置姿勢の認識方法について説明する。まず、オフライン処理として物体モデルの代表法線から複数の GRF を作成しておく。

物体モデルと入力距離データの GRF を一致させる変換行列 $\mathbf{R}_{m \rightarrow s}$ は、式 (3.2) によって計算される。

$$\mathbf{R}_{m \rightarrow s} = \mathbf{R}_{f \rightarrow s} \mathbf{R}_r \mathbf{R}_{m \rightarrow f} \quad (3.2)$$

$\mathbf{R}_{f \rightarrow s}$, $\mathbf{R}_{m \rightarrow f}$ はそれぞれ、代表法線の接平面（距離画像）からシーン座標系への変換、物体モデル座標系から距離画像への変換行列である。 \mathbf{R}_r は GRF における接平面方向の差分値 $o(\mathbf{f}_s) - o(\mathbf{f}_m)$ の回転をおこなう回転行列である。

平行移動成分は、 $\mathbf{R}_{m \rightarrow s}$ を適用した物体モデルと入力距離データの誤差が最小となる平行移動ベクトルとする。

3.4 照合戦略の切り替えを導入した位置姿勢認識

3.4.1 概要

提案手法では、GRF を用いた照合方式と、局所特徴ベース照合の切り替えを導入することによって、認識対象物の見えに依存しない位置姿勢認識を実現する。提案手法の概要を Fig.3.6 に示す。GRF は単一の物体の見えに対して算出される。したがって処理対象領域が物体ごとに分割されている必要がある。そこで、入力距離データにセグメンテーションをおこない、処理対象領域を物体ごとに分割する。分割することによって、これまでの LRF では利用できなかった物体の外輪郭を用いた Reference Frame を算出することができる。次に、各セグメントに対して平面/非平面の分類をおこなう。平面的なセグメントからは特徴点が検出されにくいと推察されるため、大域特徴である GRF を用いた照合をおこなうセグメン

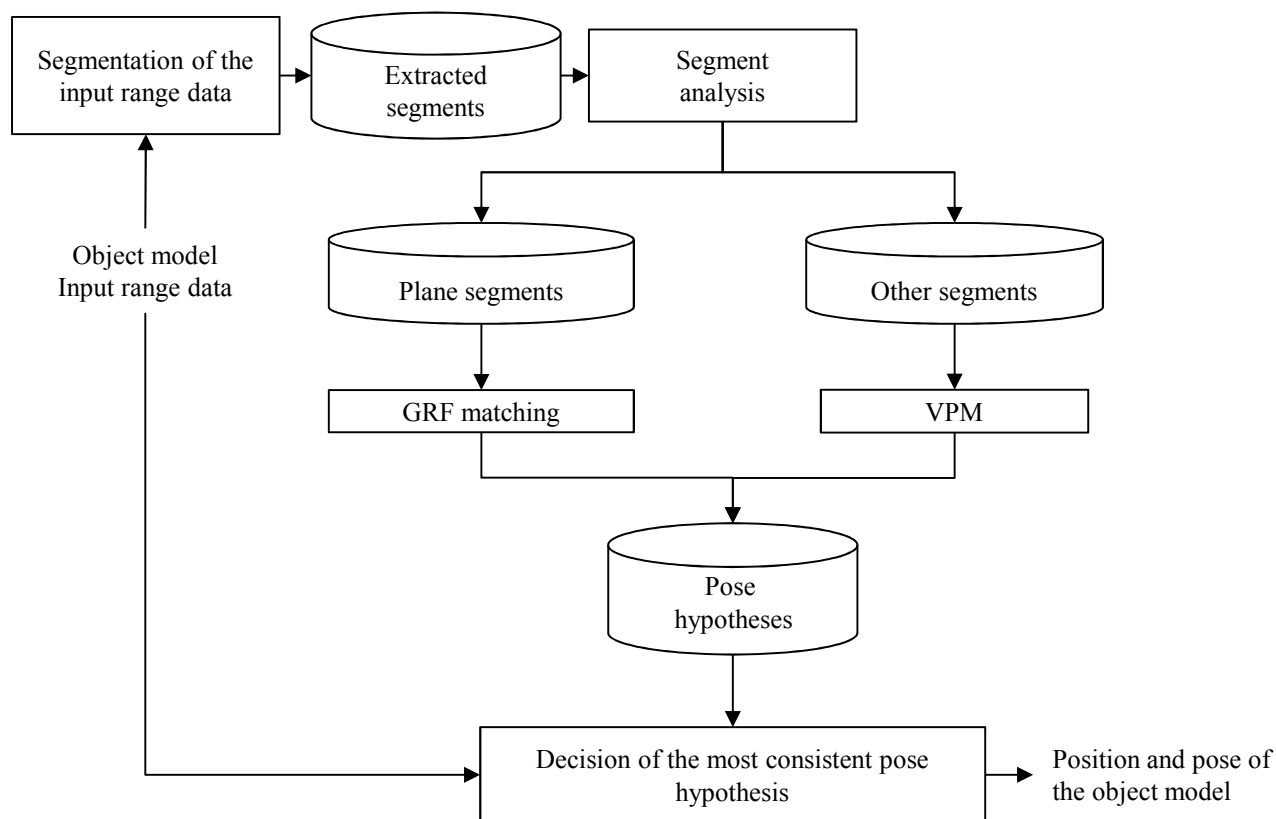


Fig.3.6 Overview of the procedure of the proposed method

トとする。そして、それ以外のセグメントに対しては、局所特徴ベース照合である VPM 法を適用する領域として判定し、それぞれのアルゴリズムを用いて姿勢仮説を生成する。姿勢仮説は複数生成されるため、尤もらしいものを選択する。そこで、物体モデルを姿勢仮説によって姿勢変換し、入力距離データとの誤差を計算する。そして、誤差が最も小さいものを最終的な認識結果とする。

3.4.2 距離データのセグメンテーションと平面/非平面分類

入力距離データをレンジファインダの空間解像度程度に離散化した画像平面に投影する。そして、事前に設定したしきい値以内の近接関係を持つ距離データをグルーピングすることによって、入力距離データをセグメンテーションする。距離データごとの近接関係にしきい値を持たせることによって、画像上で隣り合う距離データであっても高さが異なる場合は別セグメントとして出力することができる。

セグメンテーションの結果を Fig.3.7 に示す。Fig.3.7 では、セグメントごとに異なる色で着色してある。

次に、照合アルゴリズムの切り替えのために、法線方向分布を分析することによって、各セグメントを平面/非平面に分類する。3.3.1 で説明したアルゴリズムと同様に、全ての法線を EGI 上にプロットし、法線の支配率を評価する。全体のプロット数に対してしきい値以上のプロットが EGI 上で集中している場合は、特定の法線がセグメントを支配しているとみなし、平面に分類する。反対に、プロットの集中性がしきい値未満の場合は非平面とみなす。例として、曲面に代表される法線が放射状に分布する形状や、凹凸形状などは非平面に分類される。法線の支配率に関するしきい値の設定に関しては 3.5.1 にて述べる。



Fig.3.7 1
segments

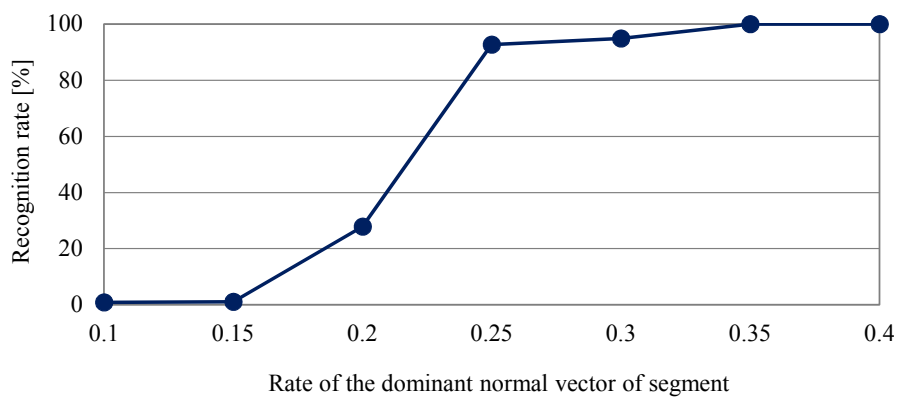


Fig.3.8 Relationship between dominant rate of normal vector of the object model and recognition rate

3.5 実験と考察

3.5.1 平面/非平面分類のためのしきい値設定

セグメント分析モジュールにおいて処理対象領域の平面/非平面の正確な分類は、認識性能を安定させる目的において非常に重要である。そこで、セグメントにおける法線の支配率と、GRF 照合の認識率の関係を調べることによって、セグメントを平面と判断するための適切なしきい値を決定する。

ランダムに生成した 1000 セグメントに対する認識結果を Fig.3.8 に示す。主法線の支配率が 2 割以下の場合では認識率が 27.9% であり、それ以上の支配率を持つセグメントに対しては高い認識率を記録した。したがって、支配率 0.2 をしきい値として、セグメント分類をおこなうこととした。

3.5.2 姿勢認識の相補性能の評価

認識に用いる 2 つの照合戦略は、相補性能が高いことが期待される。物体をさまざまな視点から観測した距離データに対して、2 手法によって姿勢推定をおこない、相補性能を評価した。

認識に用いた距離データは、ランダムに生成した 1000 個の視点から観測した物体モデルの距離データを用いた。実験に用いた物体モデルと生成された距離データを Fig.3.9 に示す。位置合わせ誤差が 3.0[mm/point] 未満となった結果を認識成功とし、(認識成功数) / (入力距離データ数) を認識率とした。

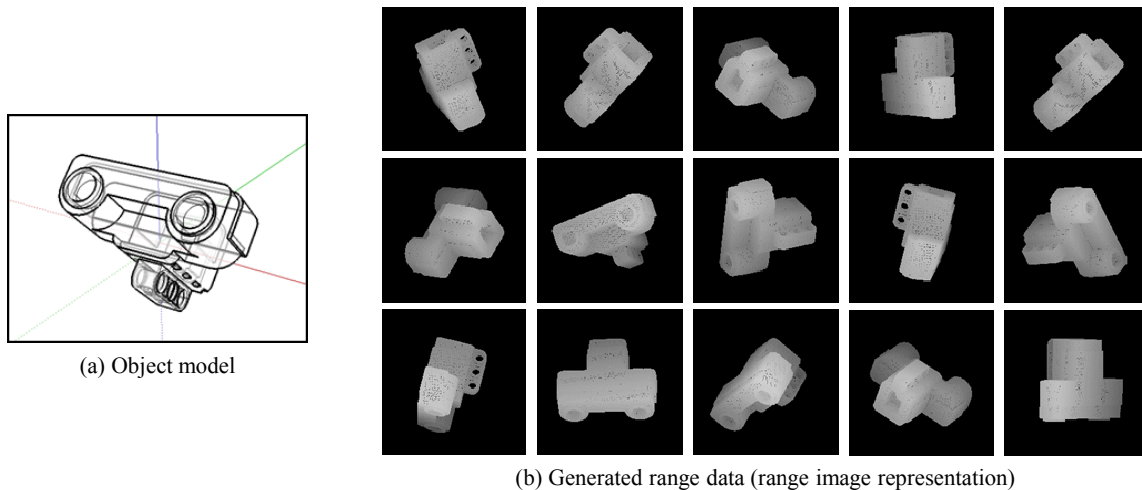


Fig.3.9 Object model and generated range data (depth image representation)

それぞれの手法が認識に成功した視点と物体モデルの位置関係を Fig.3.10 に示す．同図は結果の 2 面図であり，(a) は上からの見え方，(b) は横からの見え方である．物体はシルエットとして表示してあり，赤点は VPM 法にて認識に成功した視点，青点は GRF 照合にて認識に成功した視点である．両者の色が重なる場合は両手法で認識に成功したことを示す．多くの点が全体にわたってプロットされれば，さまざまな姿勢に対して認識可能であることを示す．結果として，全体に視点が表示されているため，本手法は物体の姿勢変化に対して頑健に照合可能なことを確認した．また，2 色のプロットの分布が分かれていることがわかる．GRF 照合は (b) Side view の上半分にプロットが集中している．これらの視点からの見えは平面的であったからである．反対に，VPM 法は (b) Side view の下半分にプロットが集中している．これらの視点からの見えは物体の物理稜線等の凹凸を含む形状であったからである．

Table 3.1 に凹凸の多い見えと，平面的な見えのそれぞれに対する認識率を示す．なお，1000 視点中で凹凸の多い見えは 611 視点，平面的な見えは 389 視点であった．

凹凸の多い見えは法線方向のばらつきが多く，安定的な代表法線を決定することが困難である．したがって，GRF 照合では認識性能が低かった．しかしながら，平面的な見えに対しては，局所特徴ベース照合の VPM 法に対して認識率が 25.7% 向上し，94.6% を記録した．平面的な見えに対して VPM 法の認識率が低かった原因は特徴点数が十分でなかったことに起因する．実際に 2 種類の形状に対して得られた平均特徴点数を比較したところ，平面的な見えから検出された特徴点数は凹凸の多い見えに対して 65% 程度であった．見えが平面的な場合は凹凸がある場合と比べて検出可能な特徴点数が減少するという傾向は 2 章の実験と一致していた．一方で，GRF 照合で極端に認識率が低かった凹凸の多い見えでは VPM 法の認識率は 94.8% を記録した．以上の結果をまとめると，認識に用いる 2 つの照合戦略は認識を得意とする形状が反対の関係にあるため，相補的であるといえる．

また，1000 視点のうちで VPM 法で認識できなかった視点の約 30% の視点を GRF 照合によって認識することができた．この結果は 2 種類の照合戦略を統合することによって位置姿勢認識の信頼性を向上させることが可能なことを示唆していると考えられる．本実験では，全体として 962 視点に対して認識に成功したことを確認した．

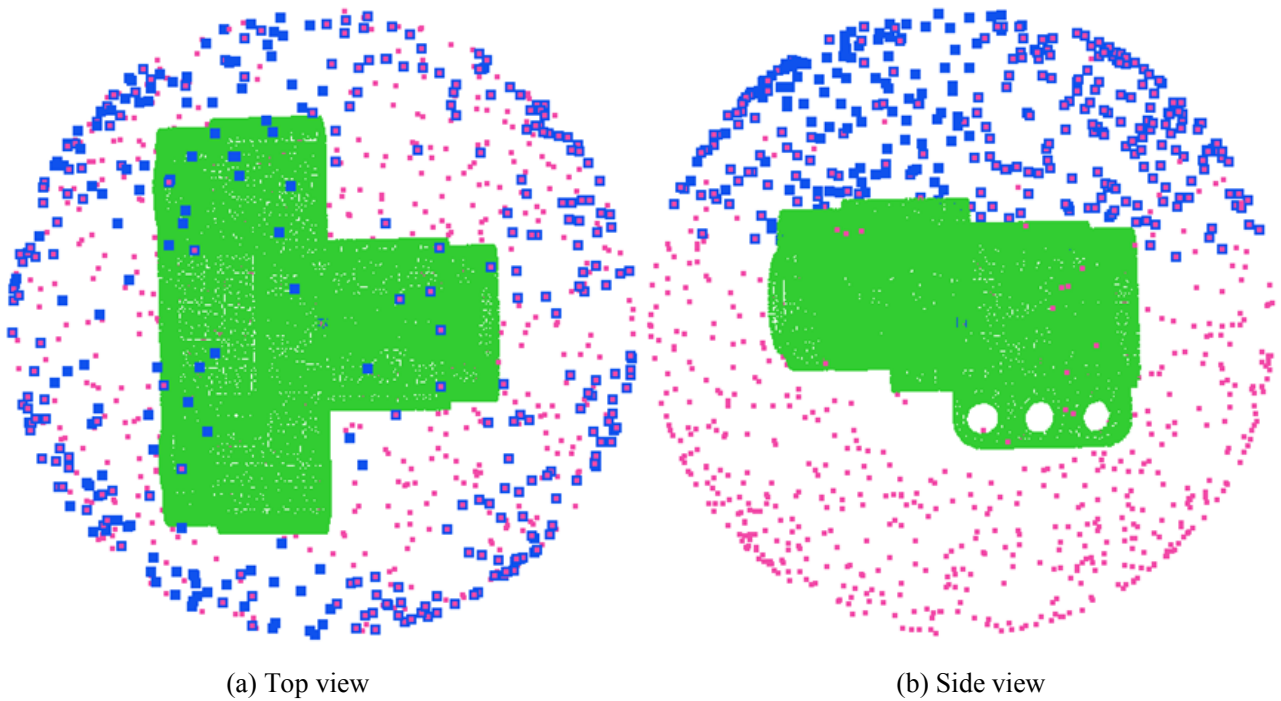


Fig.3.10 Relationship between viewpoints of correct matching and the object model. (Red dot: viewpoint of correct matching by the VPM. Blue dot: viewpoint of correct matching by the GRF matching.)

Table 3.1 Recognition rate % of two method. One is the GRF matching proposed in this paper, the other is the VPM

	Aspect of appearance	
	Indented shape	Plane like shape
GRF matching (proposed)	2.8 %	94.6 %
VPM	94.8 %	68.9 %

3.5.3 ばら積みシーンに対する位置姿勢認識の性能評価

提案手法による認識性能を評価するために、実際のばら積みシーンに対する位置姿勢認識をおこなった。入力距離データは実物の物体モデルをばら積みしたシーンをレンジファインダで撮影した 113 シーンである。実験には、平面/非平面の両方の見えを持つ、先の実験にて用いた物体モデルを選定した。比較手法は提案手法 (Strategy switching) に加え、SHOT (SHOT 特徴量 [48] + 3D hough voting[47]), FPFH (FPFH 特徴量 [39] + Papazov らの手法 [33]), VPM 法, GRF 照合 (本章で提案した平面形状の姿勢認識のための照合方法) である。

従来法は、Point Cloud Library[41] に用意されている関数を用いて実装した。認識率 P_r [%] と処理時間 T [sec] の比較結果を Table 3.2 に示す。実験は CPU : Intel®CORE™i7 860, システムメモリ : 12GB で構成されるシステムでおこなった。

VPM 法による認識率が 84.1% であったが、セグメント分析を取り入れた上で GRF 照合と統合するこ

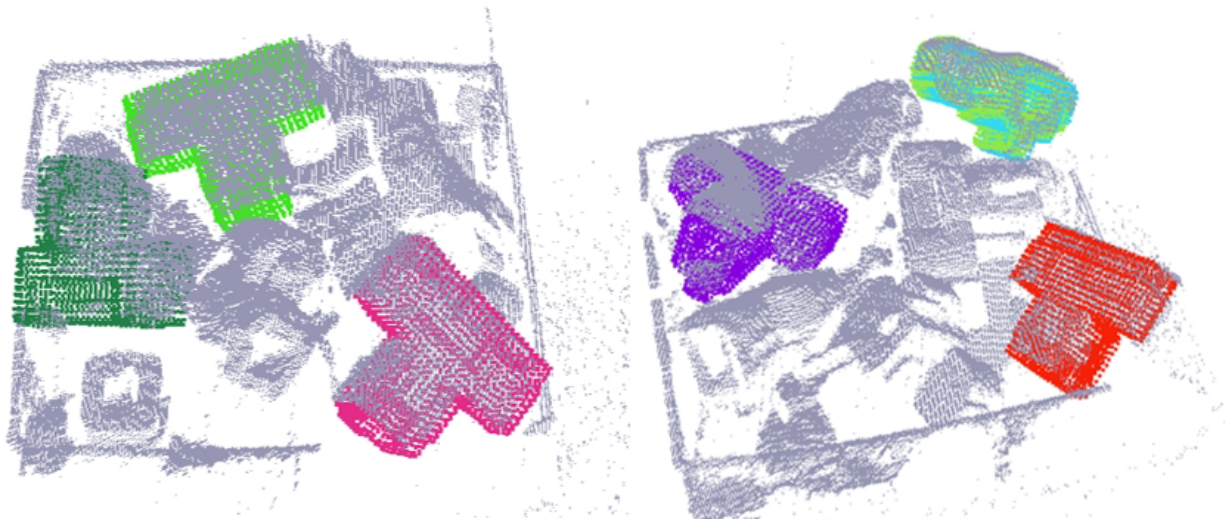


Fig.3.11 Examples of recognition results

とによって、認識率が 94.7% に向上したことを確認した。また、SHOT 特徴、FPFH 特徴を用いた照合方法よりも、VPM 法が優位であった。したがって、GRF 照合と組み合わせる局所特徴ベース照合としては VPM 法が適当であると考えられる。

2 例の認識結果を Fig.3.11 に示す。この図では、入力距離データが灰色で表されており、提案手法によって得られた姿勢仮説のうちで入力距離データとの誤差が 2.5[mm/point] のものを色付きの点群で表示している。物体モデルと入力シーンの点群が精度よく位置合わせされていることがわかる。

Table 3.2 Relationship between recognition rate Pr [%] and processing time T [sec]

Feature	Matching method	Pr	T
SHOT[48]	3D hough voting[47]	20.3	22.48
FPFH[39]	Buch[33]	31.0	0.91
Vector pair	VPM	84.1	2.72
GRF	GRF matching	64.6	0.50
GRF+Vector pair	Strategy switching	94.7	3.06

3.5.4 認識失敗例に関する考察

前節までの実験により、章の主たる提案である GRF による平面的な見えに対する位置姿勢認識手法は従来の局所特徴ベース照合よりも認識信頼性が高いことを確認できた。本節では GRF を用いた認識の失敗例に関して考察する。Fig.3.12 は典型的な認識失敗例である。(a) は入力距離データ、(b) はセグメンテーション結果を示している。(a) において、円で囲われている領域は 2 物体の境界部分であり、計測された距離データ同士が接近している。したがって、この距離データに帰属する 2 つの物体が同一セグメントに分類された。(c) は同一セグメントと判定された 2 つの物体から構成されたセグメントである。

この例のように、正しくセグメンテーションされなかった場合は、セグメント内の主法線方向が不安定になることや、主法線の接平面への投影像が物体モデルと異なりやすくなる。このため、GRF を構成す

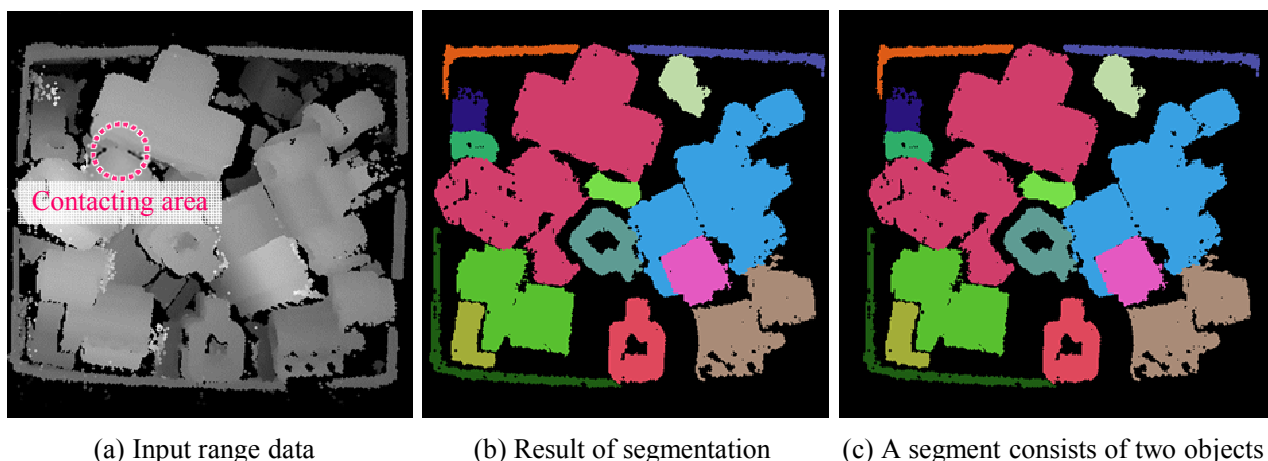


Fig.3.12 Typical example of false case. (a) Input range data. (b) Result of segmentation. (c) Mis-detected segment consists of two object's range data

る各軸の方向を安定して算出することができず、結果として正しい姿勢仮説が得られなかった。

そこで、セグメンテーション時の距離データの接続関係のしきい値を調整したうえで、再度セグメンテーションしたところ、正しく分割でき、認識にも成功することを確認した。したがって、セグメンテーションのしきい値最適化、または、文献 [51][36] 等のセグメンテーションアルゴリズムの適用によって、認識率を向上させることが可能と考えられる。

3.6 結論

本章では、物体形状の凹凸密度に依存しないモデルマッチング法を提案した。提案手法を構築するにあたり、平面的な見え方の形状データに特化した姿勢表現方法である Global Reference Frame (GRF) を提案した。さらに、物体形状の平面性を評価し、認識に用いるモデルマッチングを自動で切り替える照合戦略選択手法を提案し、前章で提案した VPM 法と、本章で提案した GRF を適応的に切り替えながらマッチングをおこなった。提案手法により、従来の VPM 法と比較して、認識性能が向上したことを実験的に示した。

以下に、本章で明らかになった結論を要約する。

- (1) 物体の見え方が平面的である場合は得られる特徴点数が減少するため、特徴点ベースの従来の 3 次元物体認識アルゴリズムの認識信頼性が低下することを説明した。工業部品を集めた CAD モデルデータセットである、Engineering Shape Benchmark を用いて見え方の平面性を分析し、全体の 10.1% が 8 割以上のデータ点が同一平面上に分布することを示し、平面的な見え方の物体に特化した位置姿勢認識手法の重要性を説明した。
- (2) 物体の平面的な見え方が平面的であっても、その外輪角の方向には支配的な方向が存在することを説明し、この情報を用いた姿勢計算のための新たな特徴表現である Global Reference Frame (GRF) を提案した。
- (3) 入力距離データをセグメンテーションし、各セグメントの平面性を分析することによって、その凹凸量に合わせて照合戦略を切り替える、照合選択切り替え式の物体認識手法を提案した。このときに重要なことは、照合戦略を切り替えるためのセグメントの平面性に関するしきい値の最適値を決

定することであるということを説明し、セグメントの平面支配率が0.2以上の場合にGRF照合を用いることが望ましいということを説明した。

- (4) 認識に用いる2つの照合戦略は、相補性能が高いことが重要であるということを説明し、さまざまな見え方の距離データに対してGRF照合とVPM法による認識実験をおこない、平面的な見え方の形状に対してはGRFが認識率94.6%(VPMは68.9%)、凹凸が多い形状に対してはGRFが認識率2.8%(VPMは94.8%)であることを確認し、これらの2手法の相補性能が高いことを実証した。
- (5) ばら積みシーンの認識実験において、提案した照合戦略選択式マッチングは従来のVPM法と比較して認識率が84.1%から94.7%に改善されたことを示した。また、典型的な誤認識例としては、入力シーンのセグメンテーションの失敗により、形状データが未分割となった時に起きるということを示した。これに対しては、他のセグメンテーション法の適用によって改善されることを説明した。

本章に係る成果は、以下の学術論文として公表済みである。

1. 秋月秀一, 橋本学, “Global Reference Frame を用いた対象物の見えに依存しない3次元位置姿勢認識”, 精密工学会誌, Vol.80, No.12, pp.1176-1181, 2014.

第4章 点群密度の変化に頑健な物体仮説生成のための3D キーポイントマッチング

本章では、キーポイントマッチングによる物体仮説生成法に焦点を当てる。キーポイント周りから記述される特徴量は、ベクトルペア特徴と比べると多くの点群データを使用するため、アイデンティティを確保しやすい。しかしながら、3次元特徴量は一般的に局所参照座標系 (LRF) と呼ばれるキーポイントごとに設定された局所参照座標をもとに記述されるため、LRF の再現性が低い場合には、特徴量の再現性も連動して低下するという性質がある。本章では、安定した3D キーポイントマッチングのために、生活支援ロボットや物流ロボット等の移動することが前提となっているロボットが撮影する距離データ特有の外乱である、点群密度の変化と欠落に頑健な LRF を提案する。従来の LRF を提案する LRF (DPN-LRF) に差し替えるだけで、キーポイントマッチングの信頼性が大幅に向上することを示す。

4.1 はじめに

家庭用ロボットなどの移動ロボットにおいて、物体の3次元位置姿勢の認識は基本的なタスクである。この状況における主な外乱としては、視点の移動が原因となって、予め点群として用意した物体モデルと、計測した物体表面の点群密度が異なりやすいということ、オクルージョンに起因した点群欠落が発生するということの2点が挙げられる。

一般に3次元物体の位置姿勢認識は、a) キーポイント検出、b) 局所参照座標系 Local Reference Frame (LRF) の生成、c) 特徴量記述、d) 対応点探索・姿勢推定、の4ステップによって構成される。

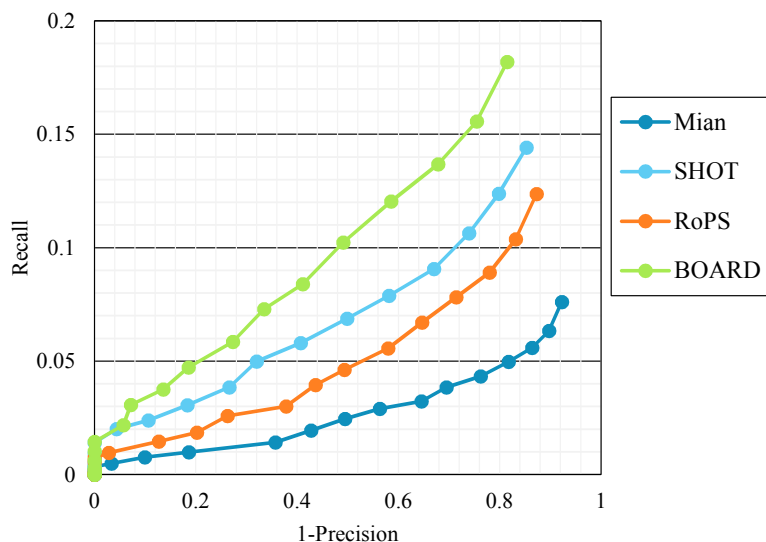
ここで、b) で生成される LRF とは、各キーポイントがもつ局所的な3次元の座標系であり、処理 c)、d) の性能向上において、非常に重要な次の2つの役割を担っている。

1. 3次元特徴量の多く [10, 15, 18, 30, 32, 48, 42, 45, 46, 56, 57] は、キーポイント周りの領域を複数のセルに分割し、セル内の点群データを特徴量化するが、このときの分割方向を定める。
2. 対応点が得られた後に物体の姿勢推定をおこなう際、LRF 間の相対的な姿勢変換パラメータが物体の姿勢となる。

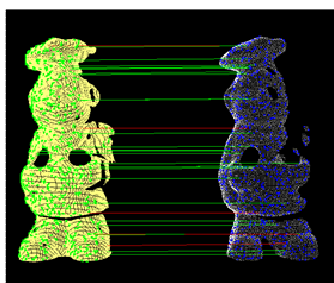
物体モデル-シーン間で正しい対応点の位置にキーポイントが検出されたとしても、LRF の設定が不安定だと、記述される特徴ベクトルも安定しない。さらに、姿勢変換パラメータも誤った値が算出されることになる。したがって、LRF を構成する軸方向の繰り返し算出性能 (再現性) がキーポイントマッチングの信頼性を大きく左右するといえる。

Fig.4.1 は、特徴量を SHOT 法 [48] に固定した上で、LRF のみを差し換えてキーポイントマッチングをおこなった時の Recall - 1-Precision curve である。比較に用いた LRF は現在の主要な手法である。それぞれの手法のアルゴリズムに関しては後述する。LRF を差し替えるのみで、キーポイントマッチングの信頼性が大きく変動することがわかる。したがって、LRF はセンシングシステムの適用環境で想定される外乱に対して頑健になるように設計するべきである。

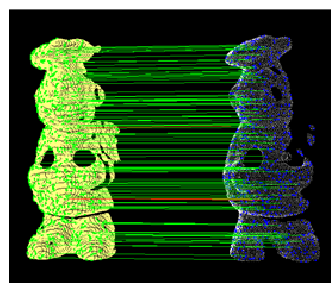
本研究の目的は、移動ロボットの実利用シーンで発生しやすい、照合対象の点群同士の密度の異なりや、点群の欠落に頑健な LRF を提案し、キーポイントマッチングの信頼性を向上させることである。



(a) Recall vs. 1-Precision curve



(b) Matching result of the SHOT-LRF and the SHOT descriptor



(c) Matching result of the proposed-LRF and the SHOT descriptor

Fig.4.1 3D keypoint matching performance. (a) Recall vs. 1-precision curve. Keypoints were matched to each other by using a SHOT descriptor and four types of LRFs. (b), (c) Example matching results for SHOT-LRF and the proposed LRF with the SHOT descriptor.

以下、4.2 では、従来の LRF の算出方式を概観し、それらの問題点について言及する。4.3 では、提案手法 (DPN, Fast DPN) のアイデアと算出方法を説明する。4.4 では、実験結果について説明し、4.5 にて、本研究の成果をまとめる。

4.2 従来の LRF とその問題点

LRF は、その算出手順の観点から、a) 一括算出型 LRF、b) 個別算出型 LRF の 2 つのタイプに分類することができる。以下、それぞれの手法に関して説明する。

a) 一括算出型 **LRF** : LRF を構成する 3 つの軸を同時に算出する方式である。キーポイント周りに設定した、球領域 (以下、単にサポートと呼ぶ。) の点群の分布から共分散行列を算出し、その固有ベクトルを用いて LRF を算出する [30, 57, 48, 42, 18, 12].

最もシンプルな手法は Mian[30], ISS[57] であり、固有ベクトルをそのまま LRF に割り当てる。しかしながら、固有ベクトルの方向には、正負の曖昧性が存在するため、いくつかの LRF はマッチング時に向きが反転する。このため、ISS では、あらかじめ一つのキーポイントに反転を考慮して複数の LRF を用意した。しかしながら、この方法では冗長なキーポイントが増え、処理時間が増加するだけでなく、誤

認識も誘発する可能性があった。これに対して、SHOT[48]では、象限ごとに分割した局所領域内の点数の優勢方向を決定し、軸方向の曖昧性を解消した。

さらに、DosSantos[12], 3D-Div[42], RoPS[18]では、点群の密度の違いの影響を解消するために、局所領域内の三角パッチの総面積によって正規化された共分散行列を生成した。しかしながら、上記の一括算出型 LRF では、サポート内の点群が部分的に欠落した場合には、分布の広がり方が変化するため、固有ベクトルの向きが不安定となり、再現性が低下するという問題点がある。

b) 個別算出型 LRF：このタイプは LRF の 3 軸をそれぞれ独立した計算方法によって算出する手法である [10, 46, 32, 56, 35]。キーポイント周りの局所形状の面法線方向は安定しやすいことから、これを z 軸とすることが共通しており、 x 軸の算出手法に特色がある。

古典的な手法としては、 z 軸の接平面に任意の点を射影することによって x 軸を生成する、Point Signature(PS)[10], 3D Point's Fingerprint[46]があるが、射影するべき点が安定しにくいいため、 x 軸の一意性が高くないという問題点があった。そこで、BOARD[35]では、キーポイント周りの点群の法線のうちで、 z 軸に対して最も傾いたものを検出し、これを直交化することによって x 軸を生成する方法を採った。さらに、点群の欠落に対応するために、欠落部分に存在する法線方向を推定することによって、LRF の再現性を向上させた。一方、Mesh HoG[56]では、キーポイントを原点とする z 軸の接平面上に規定される極座標系において、勾配方向分布から主方向を統計的に推定する方法を採った。これにより、点群の欠落に対する頑健性の高い x 軸を決定することが可能となった。

BOARD, Mesh HOG で重要なことは、 x 軸の方向を決定するために、特定の方向を最も支持するデータのみを取り出していることである。全ての点群を利用することなく計算するため、部分的な欠落に頑健な性質を持つ。しかしながら、これらの LRF は点群間の密度が異なる場合には再現性が低下するという問題点があり、文献 [18] によって指摘されている。

以上の a), b) の両タイプの LRF は、点群密度の変化と部分的な欠落の両方に対処可能とは言い難い。

ところで、LRF とよく似たカテゴリとして、局所領域ではなく、領域分割後のセグメントのような比較的大域的な情報を使って座標系 Reference Frame (RF) を生成する手法も存在する [3]。前章で説明した GRF もこのカテゴリに当てはまる。RF の生成にセグメント全体のデータを利用するため、対象物の凹凸が少ない場合に有効であるが、セグメンテーション自体が正確にできない場合には RF の再現性が低下しやすいという欠点がある。したがって、これらの手法は 3D キーポイントマッチングによる位置姿勢認識と併用して使われることが多い [2]。本章の話題の中心は 3D キーポイントマッチングのための LRF であるため、このカテゴリの RF は比較対象から除外することとする。

4.3 DPN-LRF : 点群密度と欠落に頑健な LRF

4.3.1 基本アイデア

点群密度の変化と欠落に対処するために、提案手法では 2 つの戦略を採る。ひとつめは、密度変化の影響を解消するために、RoPS で提案されたサポート内の局所的な面の面積の影響を正規化する重み係数を導入することである。もうひとつは、欠落に頑健な x 軸を得るために、サポート内の形状を代表する法線方向を用いることである。代表的な法線を z 軸の接平面に対して正射影することによって x 軸を生成するプロセスを採ることから、提案手法を” Dominant Projected Normal (DPN)”と呼ぶ。

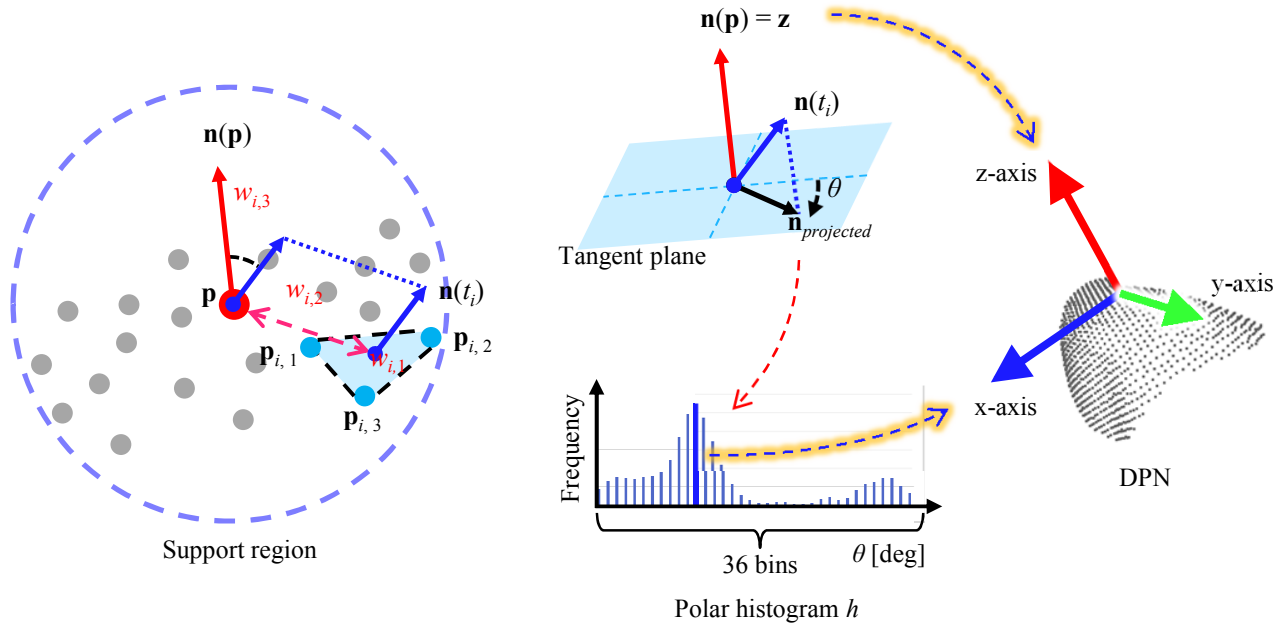


Fig.4.2 Overview of the method used to generate DPN. Left indicates a point cloud within the support region centered on the keypoint (red point). Center top indicates the tangent plane of the z-axis of the DPN-LRF; θ shows orientation of the projected normal vector. Center bottom shows the polar histogram. Right shows the DPN-LRF on the point cloud within the support region.

4.3.2 各軸の算出方法

DPN-LRF は、キーポイント \mathbf{p} を中心とした半径 r のサポート内に存在する点群 $P = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_m\}$, P によって構成される三角パッチ $T = \{t_0, t_1, \dots, t_l\}$ を利用して算出される。Fig.4.2 に提案する LRF の算出方法の概要を示す。以下に、各軸の算出方法を述べる。

z 軸：キーポイント \mathbf{p} が属する、局所面の法線とする。これには、個別算出型 LRF と同様に、キーポイントの法線 $\mathbf{n}(\mathbf{p})$ を割り当てる。 $\mathbf{n}(\mathbf{p})$ は \mathbf{p} を中心とした局所的な点群（総数= k ）を利用して算出される共分散行列 \mathbf{C} の最小固有値に対する固有ベクトルとする。

$$\mathbf{C} = \frac{1}{k} \sum_{i=0}^k (\mathbf{p}_i - \bar{\mathbf{p}})(\mathbf{p}_i - \bar{\mathbf{p}})^T \quad (4.1)$$

ここで、 $\bar{\mathbf{p}}$ は点群の平均値を示す。固有ベクトルは、向きに関して正負の曖昧性が存在する。これを解消するために、視点位置に正対する方向を正の向きとする。

従来の z 軸の推定方法としては、BOARD が採用しているキーポイントに近い点群を平面フィッティングし、その法線を割り当てる方法がある。後述する z 軸の精度に関する実験 (4.4.1) では、提案手法が採用した z 軸算出方法とほぼ同程度の精度であることが確認された。BOARD 法では、式 (4.1) のような法線算出に追加して、キーポイントごとに平面フィッティングが必要であるため、処理効率の観点から、法線をそのまま z 軸に割り当てる方法を採用した。

x 軸：この軸は、z 軸と直交した方向に設定した方向ベクトルである。提案手法では、点群の欠落に対する頑健性を高めるために、投票方式の軸方向の決定方法を取る。まず、原点を \mathbf{p} に一致させた T の法

線群を、 z 軸の接平面に正射影し、射影面における極形式の角度 $\Theta = \{\theta_0, \theta_1, \dots, \theta_l\}$ を計算する。これらを式 (4.2)(4.3) で定義される θ を軸としたヒストグラム h に投票することによって、 x 軸の方向候補の分布を作成する。

$$h(\theta) = \sum_{i=0}^l \delta(\theta, \theta_i) w_{i,1} w_{i,2} w_{i,3} \quad (4.2)$$

$$\delta(\theta, \theta_i) = \begin{cases} 1 & \text{if } |\theta - \theta_i| < th \\ 0 & \text{else} \end{cases} \quad (4.3)$$

th はヒストグラムのビン数を決定するためのパラメータである。 θ の頻度分布におけるピークを x 軸とする。頻度分布を利用するため、点群が一部欠落していたとしても、 x 軸は安定しやすいという利点がある。頻度分布を作成する際には、3 つの重み係数 w_1, w_2, w_3 を考慮した、重み付き投票をおこなう。

w_1 は式 (4.4) によって定義される点群の密度の違いの影響を抑制するための係数である。正しい対応点間において、サポート内に存在する物体の面積は、点群の密度によらずに一定であると考えられる。そこで、投票重みの総数を面積で正規化する。あらかじめサポート内の点群を三角パッチで表現しておき、パッチごとにヒストグラムへの投票をおこなう。

$$w_{i,1} = \frac{(\mathbf{p}_{i,2} - \mathbf{p}_{i,1}) \times (\mathbf{p}_{i,3} - \mathbf{p}_{i,1})}{\sum_{i=0}^l (\mathbf{p}_{i,2} - \mathbf{p}_{i,1}) \times (\mathbf{p}_{i,3} - \mathbf{p}_{i,1})} \quad (4.4)$$

ここで、 $\mathbf{p}_{i,1}, \mathbf{p}_{i,2}, \mathbf{p}_{i,3}$ は i 番目のパッチを構成する点である。パッチ面積が大きいほど、サポート内の形状を代表していると考えられるため、重みを大きくする。

w_2 は \mathbf{p} とサポート内の三角パッチ t_i の距離に関する係数であり、式 (4.5) によって定義される。

$$w_{i,2} = r - \left\| \frac{(\mathbf{p}_{i,1} + \mathbf{p}_{i,2} + \mathbf{p}_{i,3})}{3} - \mathbf{p} \right\| \quad (4.5)$$

r はサポート半径を示す。キーポイントから離れた点群は他の物体の混入の影響を受けやすい。したがって、このような点を持つ法線の投票重みを下げることによって、外乱の影響を抑制する。

w_3 は θ の安定性に関する係数であり、式 (4.6) によって定義される。

$$w_{i,3} = 1 - \mathbf{n}(\mathbf{p}) \cdot \mathbf{n}(t_i) \quad (4.6)$$

t_i の法線 $\mathbf{n}(t_i) = (\mathbf{n}(\mathbf{p}_{i,1}) + \mathbf{n}(\mathbf{p}_{i,2}) + \mathbf{n}(\mathbf{p}_{i,3}))/3$ が $\mathbf{n}(\mathbf{p})$ に対して傾いているほど、 θ が安定すると考えられる。したがって、 $\mathbf{n}(\mathbf{p})$ と t_i の法線のなす角度が大きいほど、 w_3 を大きくする。

ヒストグラム h の生成が完了したら、最も高い投票値を得たビンに対応する軸を LRF の x 軸として生成する。このとき、バイリニア補間によって、生成される x 軸の方向の精度を高める。

サポート内の点群がシンメトリーであった場合、ヒストグラムが双峰性の分布を持つことになる。この場合では、 x 軸の候補が 2 つ検出されたことになるので、2 つの LRF を生成し、位置が同一の別キーポイントとして扱うことができる。

y 軸： z , x 軸が決まっていれば、最後の軸は一意に決定できる。ここでは、 y 軸は $\mathbf{y} = \mathbf{z} \times \mathbf{x}$ として算出される。

4.3.3 重み係数の算出の高速化

上記の方法によって三角パッチを処理することによって、提案する LRF は点群の密度の変化と欠落に対する頑健性を得ることができる。入力データが三角パッチのようにメッシュ表現されている場合は、これを前提とした特徴量である、Mesh HoG 特徴量、3D-Div 特徴量等と相性が良い。しかしながら、入力データは必ずしも三角パッチで表現されておらず、むしろ整列されていない、離散的な点群であることが多い。したがって、離散点群から DPN を算出するためには、前処理として、三角パッチ生成処理が必要となる。

本節では、三角パッチなしで、DPN を算出する方法を説明する。具体的には、前節まで利用していた三角パッチを利用した重み係数算出式を、式 (4.7)–(4.9) に置き換える。

$$w'_{i,1} = \frac{\|\mathbf{p}_{nearest} - \mathbf{p}_i\|}{\sum_{i=0}^m \|\mathbf{p}_{nearest} - \mathbf{p}_i\|} \quad (4.7)$$

$\mathbf{p}_{nearest}$ は、注目点 \mathbf{p}_i の最近傍点である。すなわち、上式では、三角パッチの面積を最近傍点までの距離に置き換えている。

$$w'_{i,2} = r - \|\mathbf{p}_i - \mathbf{p}\| \quad (4.8)$$

$$w'_{i,3} = 1 - \mathbf{n}(\mathbf{p}) \cdot \mathbf{n}(\mathbf{p}_i) \quad (4.9)$$

式 (4.7)–(4.9) は三角パッチ生成が不要ため、式 (4.4)–(4.6) と比べて、計算コストが低い。以降、この方法で計算された LRF を “Fast DPN” と呼ぶ。

4.4 実験結果と考察

4.4.1 LRF の軸方向の再現性評価

実験方法

提案する LRF の性能を比較するために、対応点間における x, y, z 軸の方向の再現性を評価した。実験に用いた点群は、文献 [30] で公開された Mian らの点群データである。このデータセットには、全周囲がレジストレーション済みの物体が登録されている。5 種類登録されているモデルデータの内で、比較的計測が安定した 4 種類を選定した。実験に用いた点群を Fig.4.3 に示す。

これらの点群に対して、外乱として、1) 密度変化、2) 欠落を発生させた。Fig.4.4 に Fig.4.3 のモデル (a) に対して、それぞれの外乱を与えた場合の例を示す。同図 (a) は外乱無し (面表示)、(b) は外乱としてガウシアンノイズ ($\sigma = 0.1 \times mr$) を加えた例、(c) は (b) に対して密度を減少させた例 (ダウンサンプリングレートは $5.0 \times mr$)、(d) は (b) に対して欠落を発生させた例である。それぞれの外乱の発生方法は後述する。

モデルデータ M からランダムに 1000 点を選択し、キーポイントとする。また、各キーポイントとの最近傍点を、外乱を受けたモデルデータ S (これをマッチング対象とする) から探索し、 $\|\mathbf{p}_S - \mathbf{p}_M\| < 2.5 \times mr$ を満たす点対を対応点 $C(\mathbf{p}_M, \mathbf{p}_S)$ として登録する。 mr (mesh resolution) とは、点群の密度を示す数値である。扱う点群内に存在する全ての点の最近傍点までの距離の平均値として算出される。このとき、 \mathbf{p}_M は M のキーポイント、 \mathbf{p}_S は S のキーポイントである。対応点における LRF をそれぞれ 3×3 行列 $\mathbf{L}_M, \mathbf{L}_S$ とし、それらの誤差 e を式 (4.10) によって計算する。

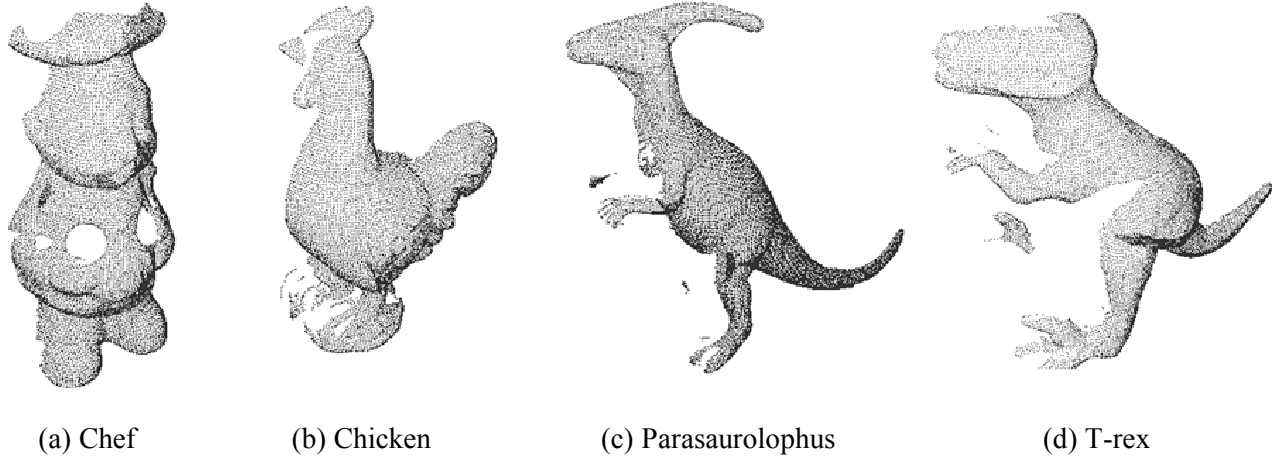


Fig.4.3 Point clouds used for experiments.

$$e = \arccos \left(\frac{\text{trace}(\mathbf{L}_S \mathbf{L}_M^{-1}) - 1}{2} \right) \frac{180}{\pi} \quad (4.10)$$

式 (4.10) は文献 [18] によって定義された LRF の誤差の計算方法である。値 e が 0 に近いほど、対となった LRF が一致していることを示す。

比較手法は、一括算出型 LRF として、Mian[30], SHOT[48], RoPS[18], 個別算出型 LRF として、PS[10], BOARD[35], 提案手法として、DPN, Fast DPN, さらに、提案する重み係数の効果を確認するために、 $w_1 \sim w_3$ の値をすべて 1 に変更した上で式 (4.2) を計算する DPN(b) の合計 8 手法である。これらの手法を Point Cloud Library[41] を用いて実装した。

点群密度変化に対する再現性

点群の密度変化に対する LRF の頑健性を評価した。シーンデータ S にガウシアンノイズ ($\sigma = 0.1 \times mr$) を付加したのち、任意のダウンサンプリングレート ($1.0 \times mr \sim 5.0 \times mr$) で密度を減少させたデータを作成した。これには、PCL に登録されている API である Voxel Grid Filter を用いた。空間中にボクセルを設定し、その内部の点群を重心点に置き換えることによって、点群が疎になる。ダウンサンプリングレートはボクセルのサイズに対応しており、この値が大きいほど、外乱を受けたモデルデータ S が低密度であることを意味する。

各手法における、 $e < 10$ [deg] の対応点の割合と S の密度 [mr] の関係を Fig.4.5 に示す。右側の数値はダウンサンプリングレートが $1.0 \sim 5.0[mr]$ に低下した場合における性能の低下率である。実験の公平性のために、LRF 算出のパラメータであるサポート半径 r に関しては、(5, 10, 15, 20)[mr] を試し、最も性能の高いものを Fig.4.5 に示した。結果として、すべての LRF において $r = 20[mr]$ のときに、最も性能が高かった。

サポートは大きい方が軸の生成に利用できる点数が増加するため、統計量としての信頼性が高く、結果として高精度化したと考えられる。しかしながら、物体辺縁部でのキーポイントでは、対象物以外の他の物体上の点群がサポート内に入ることや、欠落の影響を受けることを考慮しなければならない。本論文の実験で精度の良かった $20[mr]$ が物体のサイズのおよそ 5% あり、この値が目安の一つと考えられる。

一括算出型 **LRF** のふるまい：このカテゴリの LRF である、Mian, SHOT, RoPS は、他の LRF と

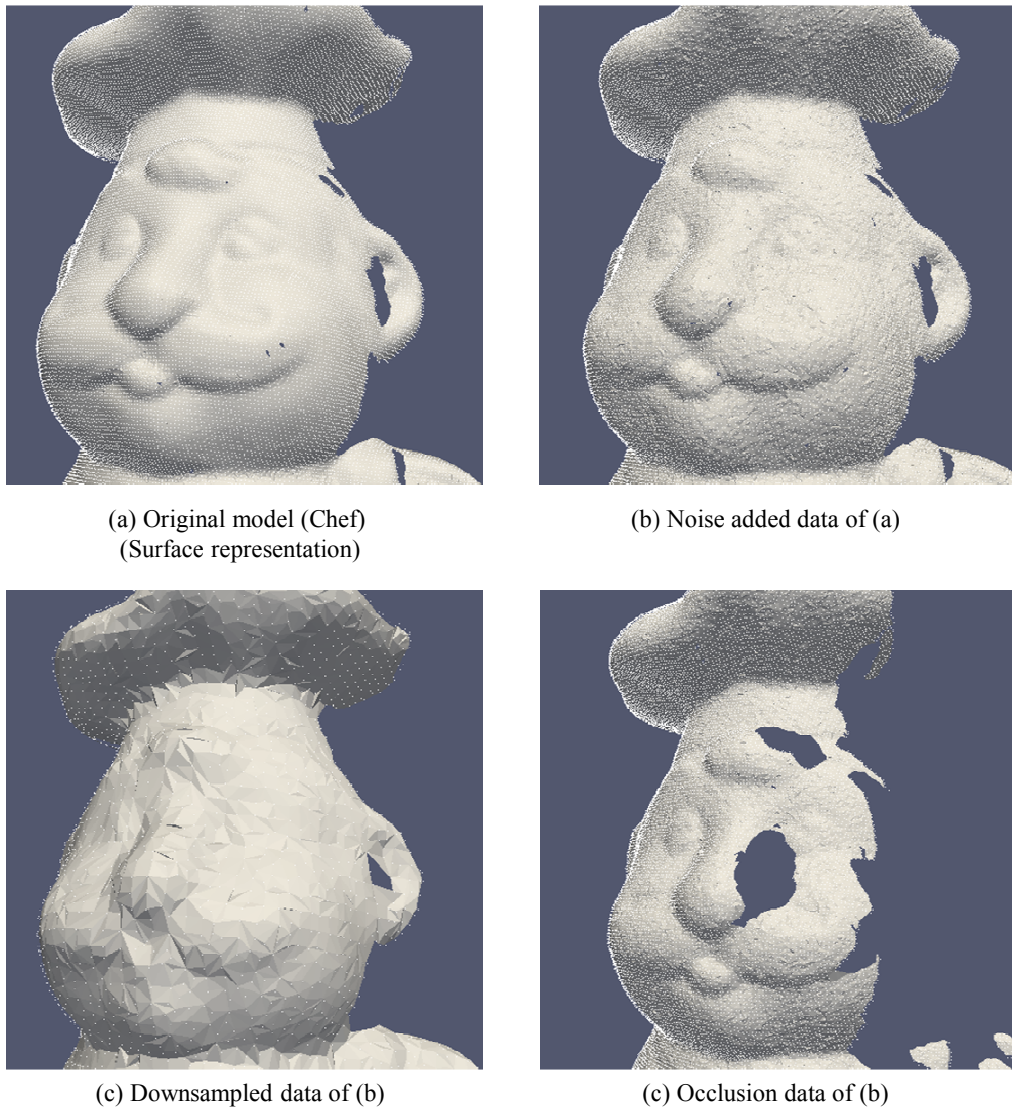


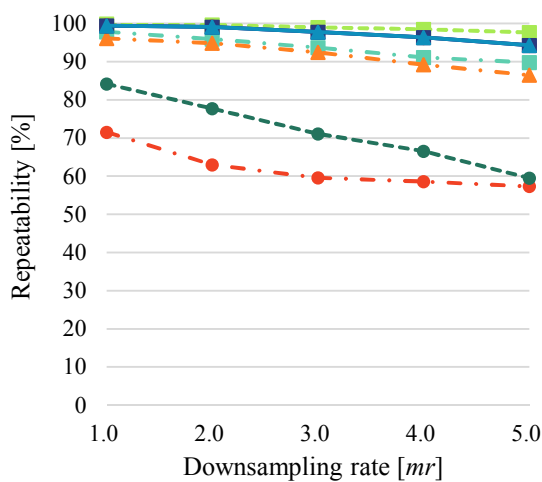
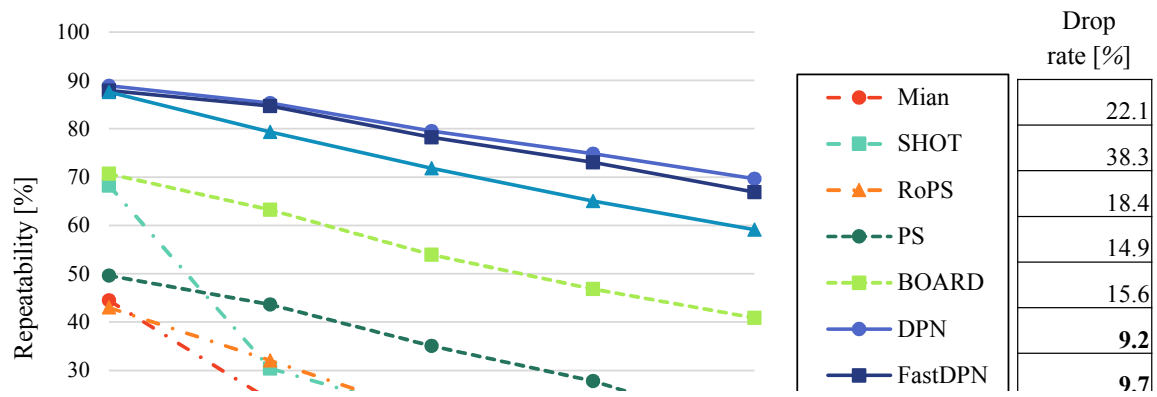
Fig.4.4 Point cloud (Chef) and its disturbed versions.

比べると性能の低下率が高いことが確認された。しかしながら、RoPS では、点群密度の変化の影響を正規化する機能を持っているため、Mian, SHOT と比べると低下率が低くなった。

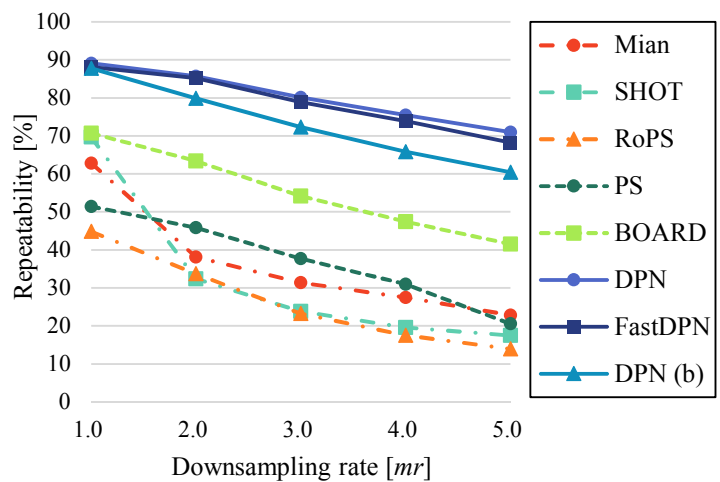
個別算出型 LRF のふるまい: このカテゴリの LRF である, PS, BOARD, DPN, Fast DPN, DPN(b) は、一括算出型 LRF と比べると低下率が低い傾向があった。PS, BOARD, DPN(b) は点群密度の変化の影響を抑制する処理が実装されておらず、ほぼ同程度の低下率であった。一方で、DPN, Fast DPN は RoPS と同様に、密度変化の影響を抑制するための重み係数が考慮されているため、最も低下率を抑制することができた。

さらに、本実験では LRF の各軸ごとの推定性能を評価するために、z 軸, x 軸の誤差を計測した。結果を Fig.4.6 に示す。この実験に関しても、縦軸は誤差が 10[deg] となった LRF の割合、横軸がダウンサンプリングレートである。

z 軸の推定性能: 多くの手法で、z 軸は安定的に推定可能な面法線を割り当てるため、点群が粗くなった場合でも推定精度の劣化が少なかった。しかし、Mian 法には、面法線の曖昧性を解消する処理が実装さ



(a) Repeatability of the z-axis



(b) Repeatability of the x-axis

Fig.4.6 The relationship between the estimation performance of the z-axis and x-axis of LRF, and the point cloud density.

れていないため、z 軸の正負の反転が発生し、推定性能が低下した。このとき、最も推定精度が高かった手法は BOARD である。点群密度が 1.0[*mr*] から 5.0[*mr*] に低下した場合に、推定精度の劣化は 2.3% であった。提案手法である DPN と Fast DPN は z 軸の推定方法が共通である。したがって、z 軸の推定性能の劣化も共通であり、5.2% であった。この結果は、他の手法の性能劣化の割合と比べると相対的に小さい。したがって、BOARD と提案手法 (DPN, Fast DPN) の z 軸の推定性能はほぼ同等といえる。

x 軸の推定性能：密度の低下に対する x 軸の推定性能に関しては、従来法では大きく劣化することが確認された。点群密度が 1.0[*mr*] から 5.0[*mr*] に低下した場合に、BOARD, SHOT の劣化はそれぞれ、29.2%, 52.2% であった。一方で、提案手法 (DPN, Fast DPN) の劣化はそれぞれ、18.2%, 19.9% であ

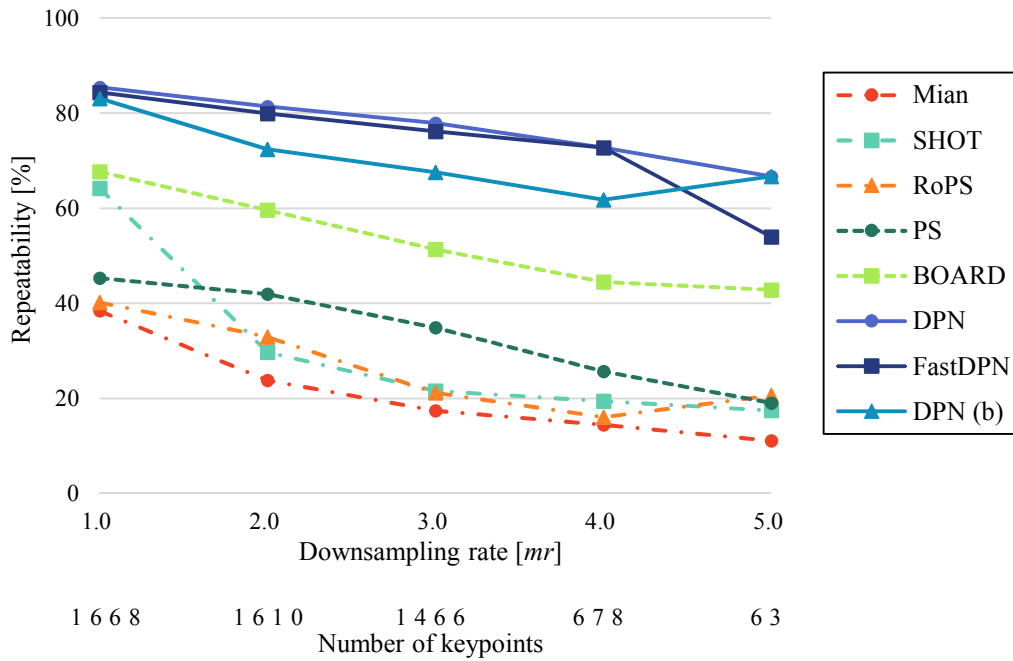


Fig.4.7 Relationship between the performance of LRF on the ISS keypoint and the downsampling rate.

り，他手法と比べて劣化が少ないことが確認された．さらに，点群の密度差が $5.0[mr]$ のとき，重み係数を考慮した DPN の性能は，考慮しない DPN(b) と比べて 10.6% 高かった．したがって，提案した重み係数を考慮したヒストグラム生成が， x 軸の生成精度の向上のために効果的であったことを確認した．

ISS キーポイント上の LRF の性能：これまでに提案された 3D キーポイントマッチング法には，キーポイント選択に特定の検出器を利用する手法もある．そこで，検出器として ISS detector[57] を利用した場合の LRF の再現性についても確認した．結果を Fig.4.7 に示す．図の下部に検出された対応点数を記載した．ダウンサンプリングレートが高くなるにつれて，キーポイント検出器の位置再現性能が低くなり，実験に利用できる対応点数が減少した．LRF の再現性としては，提案手法が他手法に対して優位であることを確認した．また，Fig.4.5 と比べると，結果にほとんど差がなかったため，LRF はキーポイント検出器に依存しにくいことが確認された．

点群欠落に対する再現性

点群の欠落率に対する LRF の頑健性を評価した．点群に部分的な欠損を発生させるために，以下の処理をおこなった．

- 1) 全周囲の物体モデルを $0 \sim 90[\text{deg}]$ の範囲で $10[\text{deg}]$ おきに回転させた点群を生成する．
- 2) 各点群に対して，単一視点からの計測を模擬して，Hidden Point Removal operator[24] による隠面除去を適用する．
- 3) 回転なしの場合の点群をモデルデータ M とし，それ以外の点群をマッチング対象 S とする．
- 4) S にガウシアンノイズ ($\sigma = 0.1 \times mr$) を付加する．

さらに，対応点 $C(\mathbf{p}_M, \mathbf{p}_S)$ を 4.4.1 と同様の方法で生成した．このとき， S は M に対して回転しているので， $|\mathbf{p}_m - \mathbf{R}\mathbf{p}_s| < 2.5[mr]$ を満たす点を対応点とした． \mathbf{R} は処理 1) で利用した回転行列である．また，対応点間の欠落率を $missing\ rate = |N_M - N_S| / \max(N_M, N_S)$ として算出した．ここで， N_M, N_S

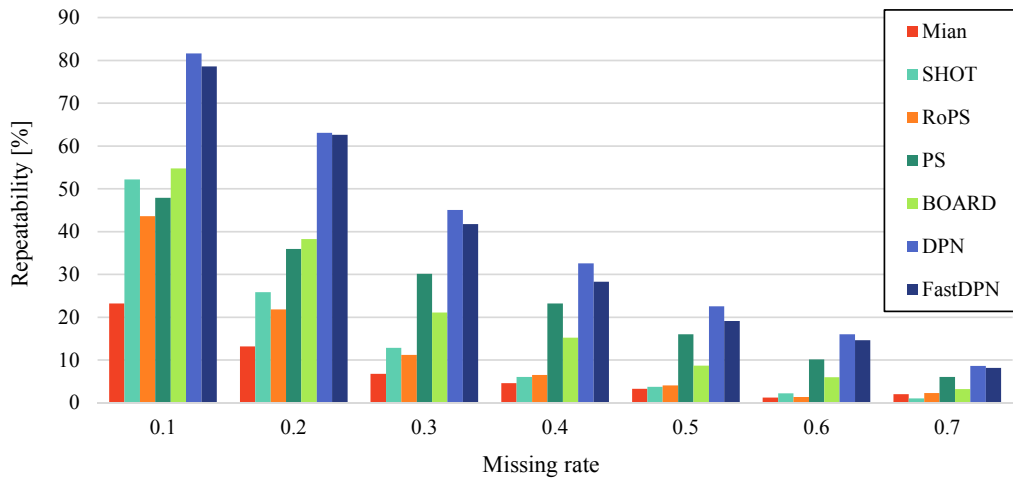


Fig.4.8 Relationship between the missing rate and the angular error of LRF. Vertical axis represents percentage of LRF has error within 10 degree.

はそれぞれ、サポート内に存在する点群の個数である。この実験では、 $missing\ rate < 0.7$ の対応点を実験に用いることとした。このとき生成された対応点は、合計で約 20000 点である。

上記の処理によって、実計測データで発生する点群欠落を模擬した対応点が生成される。Fig.4.8 に $e < 10[deg]$ の対応点の割合と欠落率の関係を示す。

4.4.1 と同様に、サポート半径 r に関しては、(5, 10, 15, 20)[mr] を試し、最も性能の高いものを Fig.4.8 に示した。RoPS は 15[mr]、他の手法は 20[mr] のとき、最も性能が高かった。

一括算出型の LRF である、Mian, SHOT, RoPS は点群に欠損を含む場合には、軸方向の安定性が低下しやすい。したがって、欠落率が高くなると、軸方向の信頼性が大きく低下したことを確認した。反対に、個別算出型である PS, BOARD は点群の欠落率の増加に対する性能の劣化が一括算出型 LRF と比べて、相対的に小さいことが確認された。提案手法では、すべての欠損率において、他手法よりも軸の推定性能が高いことを確認した。

ISS キーポイント上の LRF の性能：本実験においても、前節と同様に ISS キーポイント上の LRF の再現性を比較した。結果を Fig.4.9 に示す。この実験では約 10000 点の対応点が検出された。LRF の再現性としては、提案手法が他手法に対して優位であった。また、Fig.4.8 の結果を比べると、ISS detector を利用した方が提案手法の再現性の低下が抑えられることが確認された。ISS detector によって、凹凸部分をより多く選択することができたため、欠落が起こったとしても再現性が低下しにくかったと考えられる。

4.4.2 キーポイントマッチングの性能評価

これまでの実験の結果、提案する LRF は他の従来の LRF と比較して高い再現性を持つことを確認した。したがって、従来のキーポイントマッチングにおいて、LRF の算出手法を提案 LRF に差し替えることによって、特徴量の記述再現性が高くなり、結果として対応点探索の信頼性が向上することが期待される。

本節では、文献 [25] によって提案された比較方法を用いてキーポイントマッチングの性能を比較した。比較方法は Recall-1-Precision curve である。Recall = (正しい対応点) / (キーポイント数) であり、

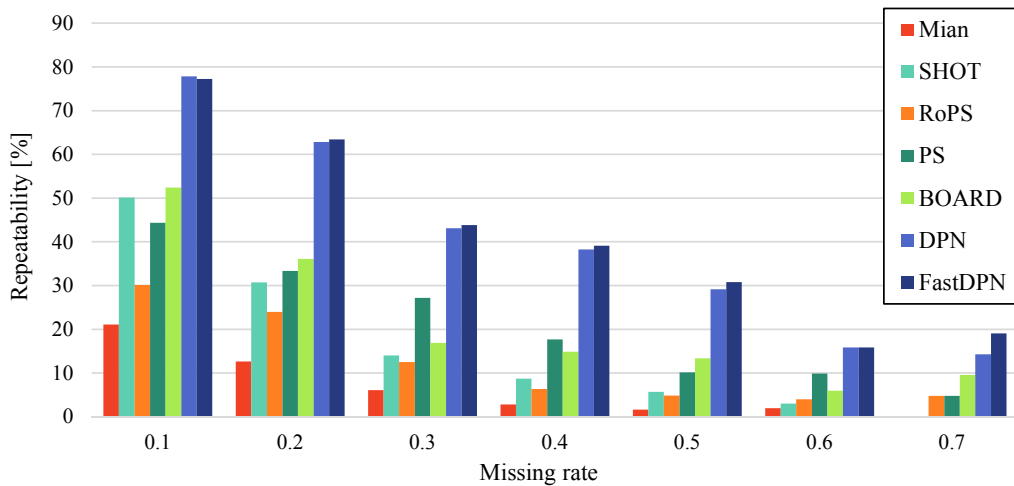


Fig.4.9 Relationship between the performance of LRF on the ISS keypoint and the missing rate.

1-Precision = (検出された対応点数 - 正しい対応点数) / (正しい対応点数) である。媒介変数は特徴量のノルムにおける 1 位と 2 位の比である。実験には SHOT 特徴 [48] を利用した。

実験条件は以下の 4 種類である。

- (a) 0.1[*mr*] ガウシアンノイズ
- (b) 0.1[*mr*] ガウシアンノイズ + 3.0[*mr*] のダウンサンプリング
- (c) 0.1[*mr*] ガウシアンノイズ + 部分的な欠損
- (d) 0.1[*mr*] ガウシアンノイズ + 3.0[*mr*] のダウンサンプリング + 部分的な欠損

キーポイント抽出器の性質の影響を除外するために、ランダムに生成した 1000 点 × 4 物体の対応点に対して実験をおこなった。なお、部分的な欠損の発生方法は、4.4.1 の方法と同様である。結果を Fig.4.10 に示す。

条件 (a) において、最も性能が高かったのは、文献 [48] で提案された、SHOT 特徴とその LRF の組み合わせであった。SHOT 特徴オリジナルの LRF は固有ベクトルをベースとして算出するため、外乱がノイズのみであるときの再現性が高かったと考えられる。また、BOARD, DPN, Fast DPN は SHOT 特徴オリジナルの LRF に次いで性能が高いことが確認された。

条件 (b) において、SHOT の性能が大幅に低下した。この実験条件では、ノイズに加えて、3.0[*mr*] のダウンサンプリングが付加されている。この条件のとき、4.4.1 の実験では、SHOT の推定性能が 25% 以下に低下したため、これが原因と考えられる。反対に、提案手法 (DPN, FastDPN) と BOARD を LRF として利用した時のマッチング性能は、他の LRF の時と比べて相対的に高いことを確認した。

条件 (c), (d) のときは、(a), (b) のときと比べて、すべての手法で、マッチングの信頼性が低下した。しかしながら、提案手法を利用した時のマッチングの信頼性は、他の LRF と比べて高かった。

全ての実験条件において、DPN と FastDPN の性能はほぼ同等であった。したがって、重み係数の近似計算法である式 (4.4)–(4.6) は、三角パッチが用意されていない点群に対して、有効であると考えられる。

条件 (a) と (b),(c),(d) の結果を比較すると Recall 値が低下している。一般的にはキーポイントマッチングの後で、物体の位置姿勢を算出する処理が控えている。ここでは、数点~10 数点の正しい対応点が得られていれば、正しい姿勢認識をおこなうことができるため、実用上の問題点は無いと考えられる。

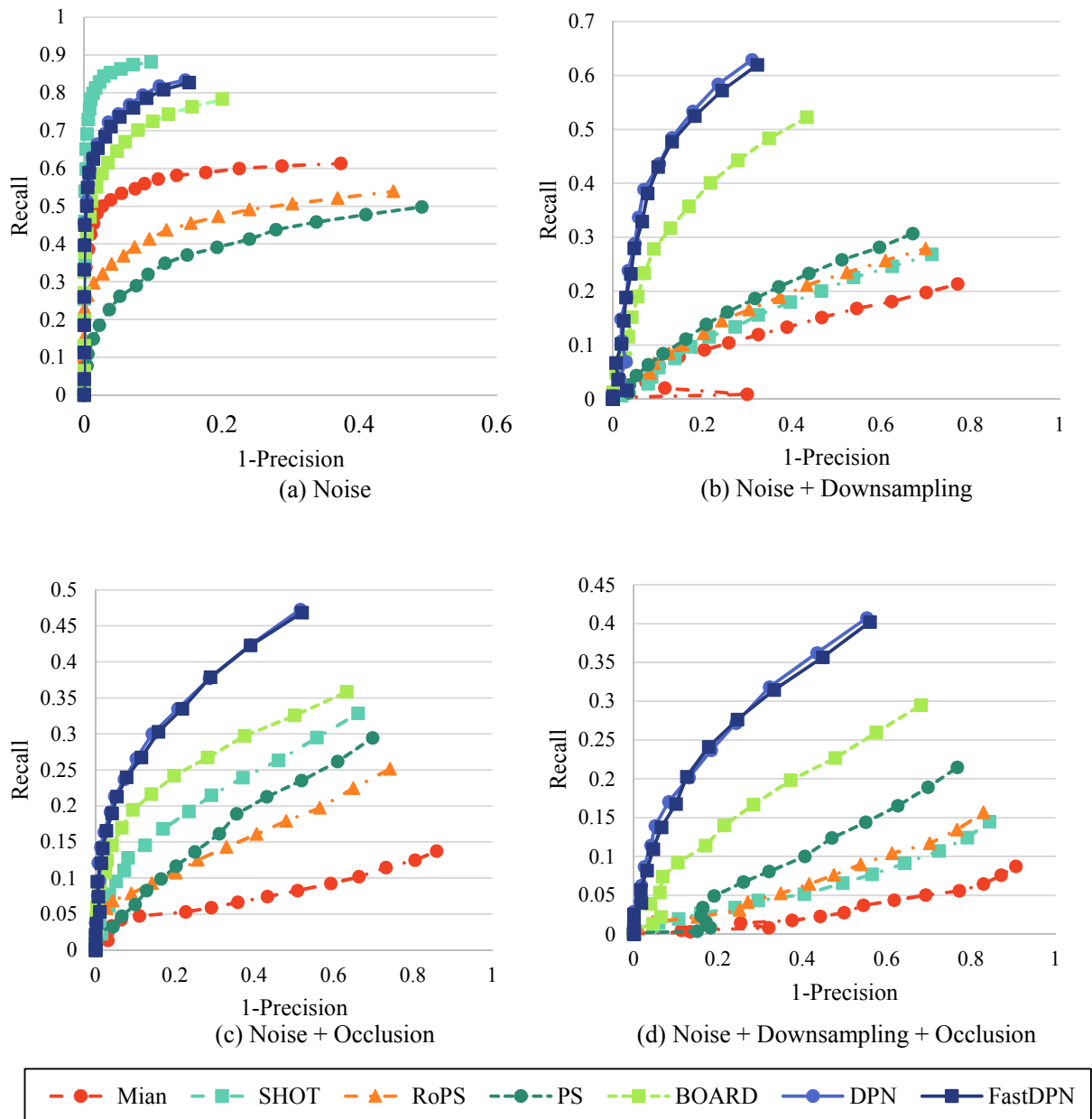


Fig.4.10 Recall vs. 1-precision curves for four conditions.

4.4.3 処理時間の比較

LRF の算出に必要な処理時間を計測した。この実験では、4 物体 × 1000 個のキーポイントから LRF を算出し、1 つ当たりの処理時間を計測した。なお、全ての手法において、共通の半径 $r = 15[mr]$ を用いて LRF を算出した。また、RoPS, DPN に関しては、前処理として点群の三角パッチ化が必要であるが、その時間を含めていない。実験に用いた PC のスペックは CPU: Intel CORE™i7 860, RAM: 12GB である。それぞれの処理は並列化を用いていない。結果を Fig.4.11 に示す。

提案手法は、これまでの実験で性能の高かった BOARD 等の手法と比べて、1.4 倍程度掛かることが確認された。しかしながら、提案手法を用いてマッチングをおこなうことで、対応点の信頼性が高くなるため、結果として後段処理の姿勢推定や、レジストレーションの性能が高くなりやすい。したがって、従来

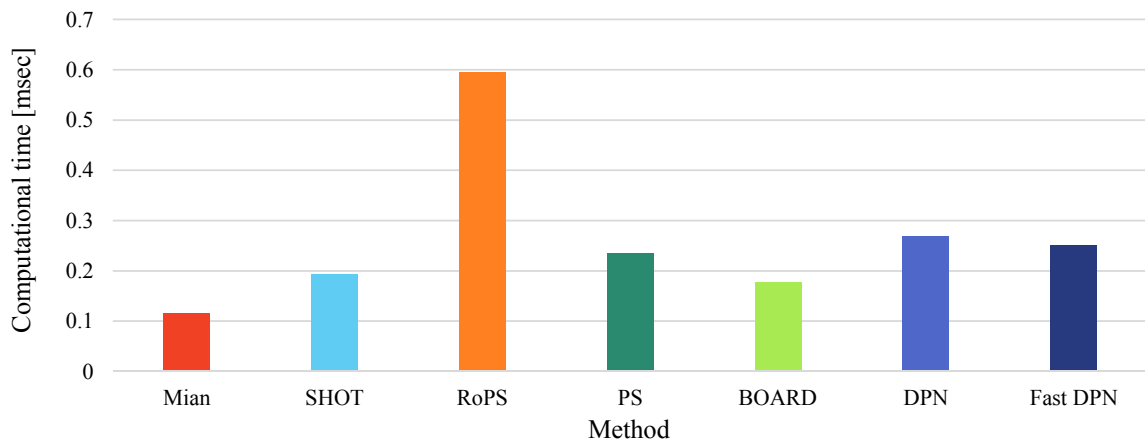


Fig.4.11 Computational time of each LRF.

の LRF を提案手法に差し替える利点は十分にあると考えられる。

4.5 結論

本章では、物体モデルと入力シーン中の対象物間の点群密度が大きく異なっている場合や、点群の欠落に頑健な 3D キーポイントマッチング法を提案した。3D キーポイントマッチングの信頼性を高めるためには局所参照座標系 (LRF) を再現性高く算出することが重要であることを説明し、上記の条件においても再現性が高い LRF である、DPN を提案した。従来の LRF を DPN に差し替えることによって、対応点探索の信頼性が向上することを示した。

以下に、本章で明らかになった結論を要約する。

- (1) 3D キーポイントマッチングにおける対応点探索の信頼性は、局所参照座標系 (LRF) の再現性に依存することを説明した。照合対象の点群に外乱として密度の変化と欠落を与えた場合のキーポイントマッチングにおいて、特徴量を SHOT[48] とし、LRF のみを Mian-LRF[30], SHOT-LRF[48], RoPS-LRF[18], BOARD[35] に切り替えて照合したところ、マッチングの性能が大きく変動することを示した。この結果を受けて、入力シーンの外乱の性質に適した LRF を設計することが重要であることを説明した。
- (2) LRF を構成する軸方向を算出する時に、局所的な点群密度を正規化項として加えることによって点群密度の変化に頑健性を向上させること、および軸方向候補を表す頻度分布から最頻値に対応する方向を LRF の軸方向とすることによって、欠落への頑健性を向上させるアイデアを説明した。これを実装した基本手法として DPN を提案し、局所密度計算を簡略化することによって処理コストを削減する、FastDPN を提案した。
- (3) 提案手法は点群の密度変化や欠落がある状況において、従来法と比較して LRF を構成する各軸を再現性高く推定することが可能であることを示した。また、キーポイント検出方法として一般的に利用されている ISS 法 [57] を使用したとしても、同様の結果が得られることを示した。さらに、3D キーポイントマッチングにおいて、LRF を提案手法に差し替えることによって、対応点探索の信頼性が向上することを確認した。

本章に係る成果は、以下の学術論文誌に採録が決定済みである。

1. 秋月秀一, 橋本学, “3次元キーポイントマッチングのための点群密度変化および欠落に頑健な Local Reference Frame”, 電子情報通信学会 D, Vol.J99-D, No.8, 2016. (採録決定)

第5章 形状的整合性と配置実現性に基づく仮説検証型物体認識アルゴリズム

本章では、複数の3次元物体の位置姿勢を同時に認識するための仮説検証型アルゴリズムについて述べる。提案手法では、前章までに説明したモデルマッチング法の結果として出力される物体の位置姿勢を「物体候補」として入力し、これらを取捨選択しながら、入力シーンと最もよく一致する物体候補の組み合わせを算出する。このとき、従来の仮説検証型アルゴリズムでは導入されていなかった物体の干渉モデルを利用し、形状的にも、物理的にも整合性のとれた解を効率よく探索する。

5.1 はじめに

近年、ロボットピッキングのためのばら積み物体の認識等の用途のために醸成されてきた3次元物体の位置姿勢認識アルゴリズムが、物流分野 [16] や生活支援ロボットのアプリケーション開発の分野への展開を始めている。部品ピッキングにおける認識タスクに対する、物流用途や生活支援用途における物体認識タスクの難しさとしては、多種類の物体が存在するシーンから、目的の物体のみを見つけ出すという識別問題に新たに取り組まなければならないということが挙げられる。一般に、部品ピッキングでは一つのビンに入っている部品は一種類のみであるため、隠れやノイズの影響が最も少ない一番上に置かれている物体を認識すればよい。しかしながら、物流用途ではシーン中の物体の配置状態に関係なく対象物を見つけなければならない。前章までに説明したモデルマッチング法では、物体モデルと距離データとして得られた入力シーン間の形状的な一致度合い（スコア）を基準に物体を検出するが、物体の配置状態によっては、対象物が他の物体に隠されることによって一部しか見えていない時のスコアよりも、複数の対象物以外の物体によって構成される偽の面がもつスコアの方が高くなり、誤認識を引き起こす場合がある。モデル単体としてのスコアが高い以上、モデルマッチング法にはこの誤認識を検知することができないため、このことはモデルマッチング法のアルゴリズムの限界といえる。

偽の面への誤認識は、入力シーンの一部から一致度合いが高くなるような物体モデルの姿勢が入力シーンの一部から見つかることが原因である。このモデルマッチング法によって発生する誤認識の解決方法としては、大局的データを用いた仮説検証型アルゴリズムの適用が有効である。仮説検証型アルゴリズムでは、モデルシーン間でスコアを定義するのではなく、個々の物体の位置姿勢候補（物体候補）の組み合わせとして表現されるシーン仮説と入力シーン間でスコアを定義する。すなわち、シーン中に存在するデータすべてをスコア計算に寄与させる。シーン全体を矛盾なく解釈することのできる物体候補の組み合わせを探索することによって、モデルマッチング法によって発生した誤認識を棄却することができる。

仮説検証型アルゴリズムで最も重要なことは、どのような情報を使ってスコアを算出するかということである。文献 [19] の手法では、モデルマッチング法によって得られた姿勢仮説群からシーン仮説を距離画像としてレンダリングして、入力シーン（距離画像）との画像整合性を最大化することによって、シーンを解釈する方法をとった。文献 [1] では、モデルとシーンの両データを3次元点群として扱い、これらの形状データとしての一致度と法線方向の一致度を利用して、シーンを解釈した。さらに、距離データだけ

でなく色情報も利用できる場合は、色の類似性を取り扱ったスコア設計により、認識信頼性を向上することができることが報告されている [2].

しかしながら、上述した従来の仮説検証型アルゴリズムは入力シーンとシーン仮説の間の一致性のみをスコアとしていたため、しばしば物体仮説同士が空間的にオーバーラップしたシーン仮説が生成され、誤認識を引き起こす要因となっていた。

本研究では、シーン仮説を構成する物体候補同士の配置の実現可能性に着目し、これをスコアに取り入れることによって、物体配置の物理的な自然さと形状的な整合性の両方を満たすシーン仮説を探索する仮説検証型アルゴリズムを提案する。

5.2 形状的整合性と配置実現性に基づく仮説検証

5.2.1 大局的情報利用の効果

本節では、提案する物体認識フレームワークの妥当性を説明する。従来のモデルマッチング法では、局所的な情報、すなわち単一の物体モデルと入力シーンの一部（物体モデルと重なる部分）の一致度合いを認識の基準としていた。この方法では、物体モデルと重ならなかった入力シーンの余り部分に関しては解釈が与えられることなく認識結果が出力されていた。したがって、もしマッチングスコアの高い誤認識結果が発生したとしても、入力シーンにおける余り部分を説明することのできる物体モデルが存在しないことを証明することができれば、その結果を棄却することができるはずである。したがって、提案するマッチング方式では大局的情報、すなわちシーン仮説と入力シーン全体の一致度合いをスコアとすることは、正確に対象物を認識するための方策として妥当といえる。

5.2.2 提案手法の概要

提案手法は、仮説生成モジュールと検証モジュールで構成されている。提案手法の処理の流れを Fig.5.1 に示す。

仮説生成モジュールについて説明する。まずはじめに、Model matching モジュールが動作する。モデルマッチング法によって、物体候補群 H を生成する。ここでは、モデルライブラリに登録されている各物体 $M = \{M_1, \dots, M_m\}$ を入力シーンの 3 次元点群 S に対してマッチングさせ、物体候補群 $H = \{h_1, \dots, h_n\}$ を生成する。 h_i はペア (M_i, T_i) で構成されている。 T_i は物体候補の 6 自由度の位置姿勢であり、物体モデル M_i を入力シーン S へ位置合わせするための姿勢変換パラメータである。次に Scene hypothesis generator では、いくつかの物体候補を組み合わせることによってシーン仮説 $hs(X)$ を生成する。 X は、長さ n のビットストリング $X = \{x_1, \dots, x_n\}$ であり、 $x_i = 1$ のときに、物体候補 h_i がシーン仮説 $hs(X)$ の構成要素として選択されたことを示す。

検証モジュールは、 S と $hs(X)$ の整合性を計算するジュールである。提案手法では、 X 中の任意のビットを選択することによって、 S と最も類似度の高いシーン仮説を選択する組み合わせ最適化を解く。したがって、解くべき解空間のサイズは 2^n である。提案手法では、配置実現性と形状的整合によって構成された評価値を最適化する。次節からは、評価値を構成する各成分について説明する。

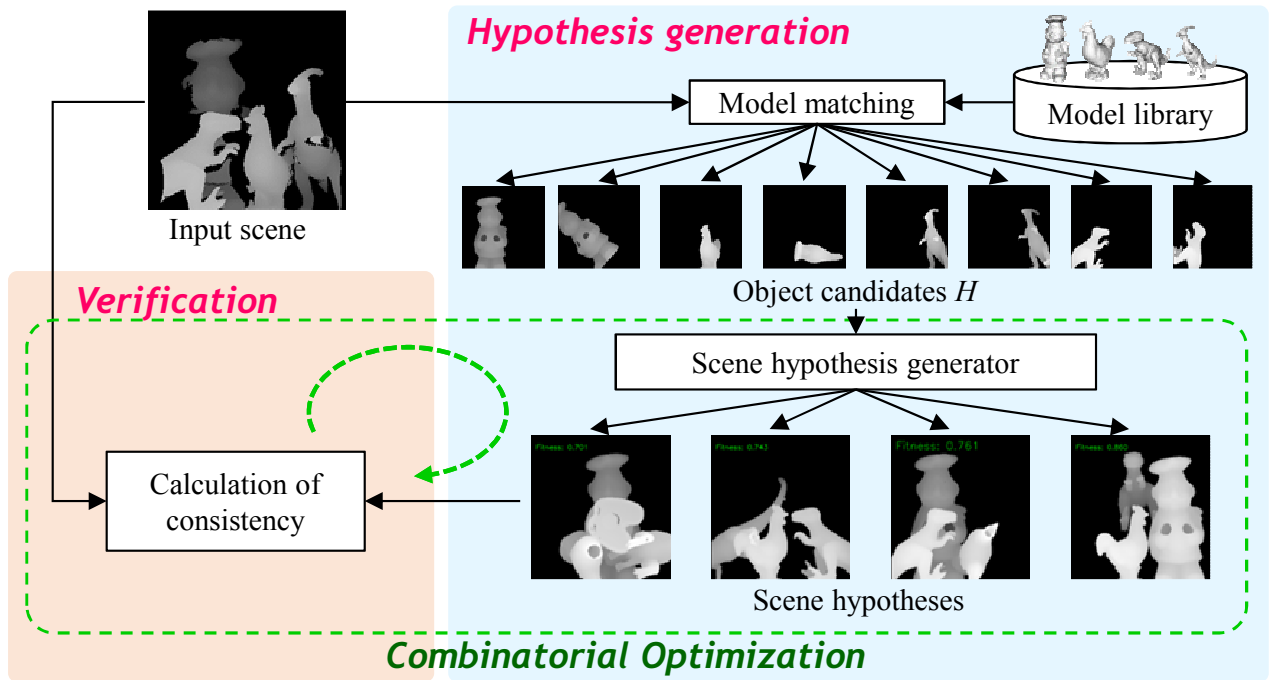


Fig.5.1 The overview of framework of the proposed hypothesis verification algorithm.

5.2.3 配置実現性: $f_P(X)$

この項は、仮説シーンの物理的な自然さを評価する。具体的には、仮説シーン中の物体候補同士が互いに空間的に重なりあっているかどうかを判断する。現実には物体同士を空間的に重ねた状態で配置することは不可能であるため、このような場合に配置実現性 $f_P(X)$ の値を低下させる。 $f_P(X)$ は式 (5.1) によって定義される。

$$f_P(X) = 1 - \frac{1}{R} \sum_{i,j \in R} C(h_i, h_j) \quad (5.1)$$

$C(h_i, h_j)$ は干渉判定器である。物体仮説 h_i と h_j が空間的に重なり合っている場合に 1 を返す関数である。 R は $hs(X)$ 中で有効となっている物体候補の中から、自分自身を除いた全てのペアを示す。

物体候補同士の重なり合いの判定には、各物体の干渉モデルを使用する。 Fig.5.2 に提案する干渉モデルを示す。

干渉モデルは以下の 2 種類の球によって構成される。

1. Outer sphere (Red) : 物体の外接球。
2. Inner spheres (Gray) : 物体の形状を近似した、複数の球体の集合。

球は物体の干渉を判定するための最も単純なプリミティブ形状である。したがって、高速な干渉の判定が可能である。 2 つの球が干渉している場合、次の条件が成り立つ ; (二つの球の中心間の距離) < (二つの球の直径の合計)。

提案手法における干渉判定は、以下の 2 段階の判定によって処理を高速化した。

1. Outer sphere 同士の干渉判定

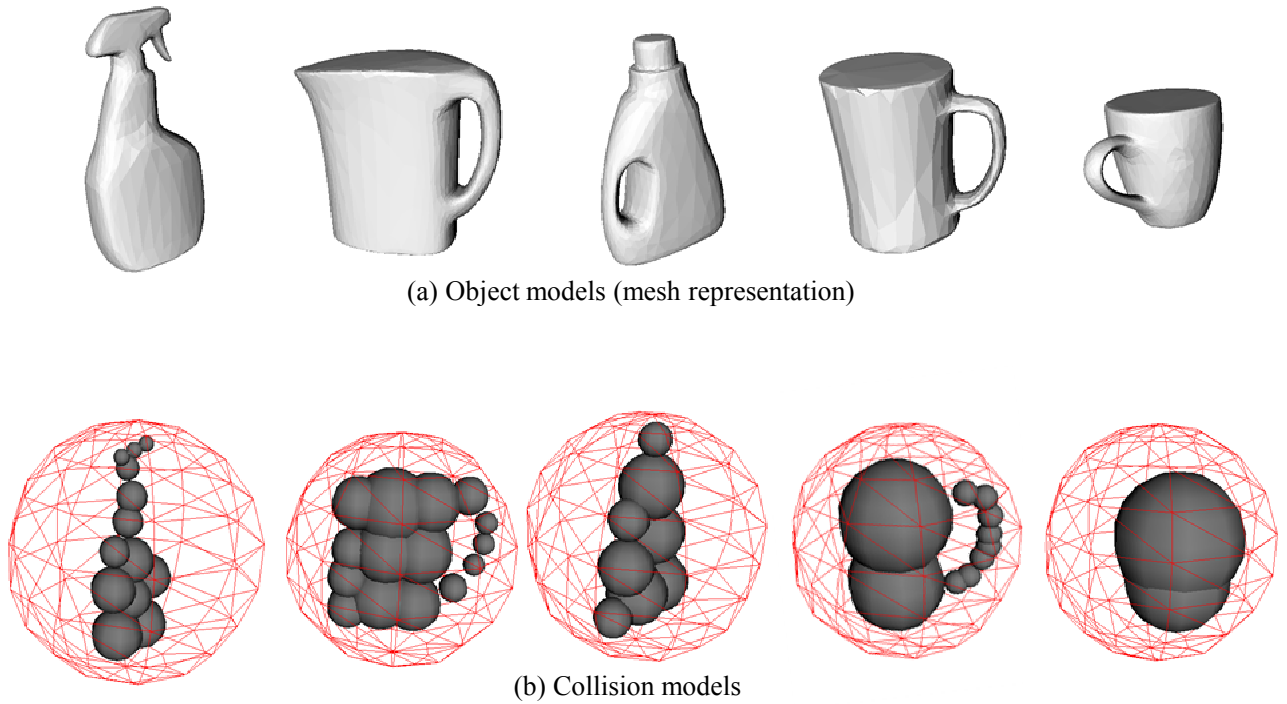


Fig.5.2 (a) Object model. (b)Collision model of (a). Red sphere is a circumscribe sphere of (a). Gray spheres are approximating object shape.

2. Inner spheres 同士の干渉判定

もし、物体同士が十分に離れている場合は、処理 1 によって干渉していないことを判断することができる。処理 1 で干渉が起きていた場合、処理 2 に移る。1 度だけでも干渉が起きた場合は、物体候補同士は空間的な重なりを起こしていたとみなす。

Fig.5.3 はシーン仮説における干渉判定の結果の一例である。(a) は入力シーンの概観である。12 個の物体が密に配置されているシーンである。(b) は干渉モデルとして表現されたシーン仮説である。このシーン仮説は 13 個の物体候補によって構成されている。白い干渉モデルは干渉を起こしていない物体候補を示し、色付きの干渉モデルは干渉を起こしている物体候補を示す。このケースでは、奥の物体候補が干渉を起こしているため、 $f_P(X)$ の値が小さくなる。(c) はこのシーン仮説を構成する物体候補をレンダリングした結果である。緑の数字は各物体候補のインデックスを示す。

Inner Spheres を生成するための方法としては、物体形状を球の集合として近似する文献 [53] の手法の適用が考えられるが、本研究では手動で割り当てることとする。

5.2.4 形状的整合性: $f_S(X)$

この項は、シーン仮説 $h_s(X)$ の形状的整合性を評価する。具体的には、距離画像としてレンダリングした入力シーンと $h_s(X)$ 間の画像類似度を式 (5.2) によって計算する。

$$f_S(X) = \frac{1}{N} \sum_{i=1}^N Sim(I_S(i), I_{h_s}(i)) \quad (5.2)$$

$I_S(i)$ と $I_{h_s}(i)$ は S の $h_s(X)$ 距離画像における i 番目の画素値である。 N は $I_S(i)$ と $I_{h_s}(i)$ のどちら

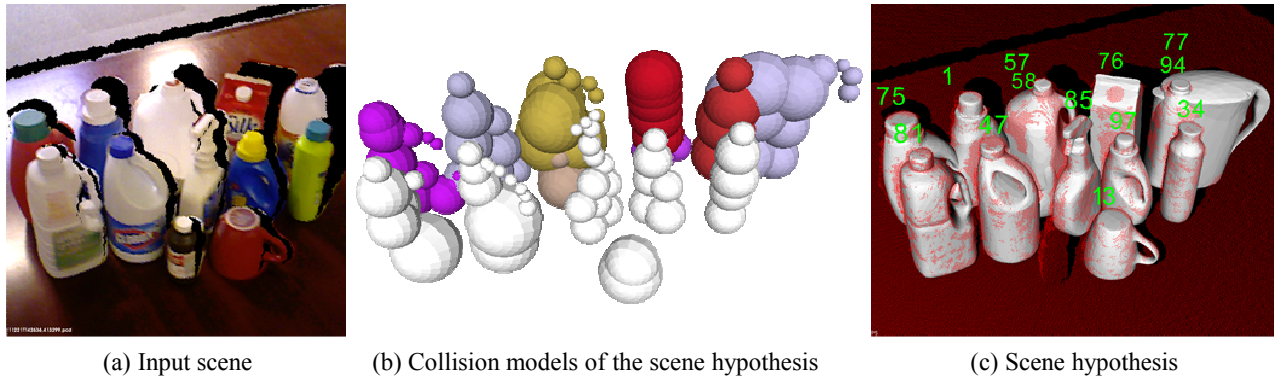


Fig.5.3 Example of the collision detection. (a) Input scene. (b) Collision models of the scene hypothesis. White models have not occur spatial intersection. But colored models have occur spatial intersection. (c) Scene hypothesis of (b).

かが距離値をもつ画素数の合計を示す。関数 $Sim(I_S(i), I_{hs}(i))$ は式 (5.3) として定義される。

$$\begin{cases} 1 & \text{if } |I_S(i) - I_{hs}(i)| < th_d \\ 0 & \text{else.} \end{cases} \quad (5.3)$$

th_d は入力シーン中の距離値の計測精度に合わせて調整する、距離画像の類似度に関するしきい値である。

5.2.5 最適化処理

提案手法の検証モジュールでは、式 (5.4) によって定義される関数 $F(X)$ を最大化するビット列 X を探索する、組み合わせ最適問題を解く。

$$\tilde{X} = \arg \max\{F(X) = wf_P(X) + (1 - w)f_S(X)\} \quad (5.4)$$

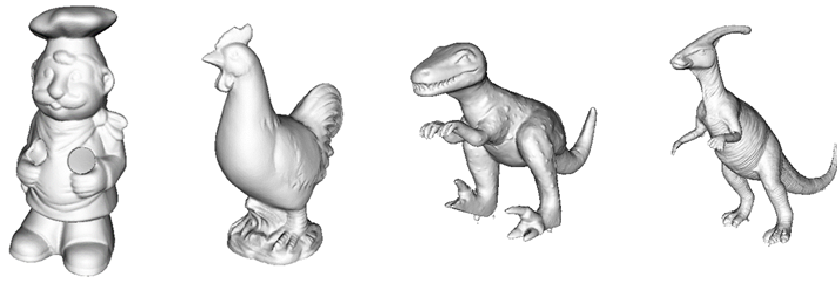
w は2つの指標の重み係数を示す。この値の最適値に関する考察は後述する。提案手法では、最適化を遺伝的アルゴリズム [17] よって解く。染色体はビットストリングそのものとした。

5.3 実験結果と考察

5.3.1 データセット

提案手法の性能を評価するために、公開データベースを利用して認識性能を比較した。物体認識の汎用性の評価するために、品質の異なる距離データによって構成された2つのデータセット (Laser Scanner dataset, Kinect dataset) を利用した。以下にそれぞれの内容を示す。

Laser Scanner dataset[29]: 高精度なレーザスキャナで撮影された50シーンと、5つのメッシュ表現された物体モデルで構成されるデータセットである。シーン中に存在する各物体には、正解データとして姿勢変換行列が割り当てられている。5つの物体モデルのうち、1つ (rhino) は完全な全周囲モデルになっていなかったため、実験には用いなかった。実験に用いたシーンデータの例と4つ物体モデルの例を Fig.5.4 に示す。



(a) Object models (mesh representation)



(a) Object models (mesh representation)



(b) Input scene (point cloud representation)

Fig.5.5 (a) Examples of the Object model (mesh representation). (b) input scene (point cloud representation).

Kinect dataset[1]: Kinect センサで取得されたシーンデータと 35 種類の物体モデルで構成されている。Laser Scanner データセットと同様に、各シーン中の物体に対して姿勢変換行列が割り当てられている。本実験では、最もに複雑な 12 個の物体が密接して置かれているシーンを利用した。実験に用いたシーンデータと物体モデルの例を Fig.5.5 に示す。

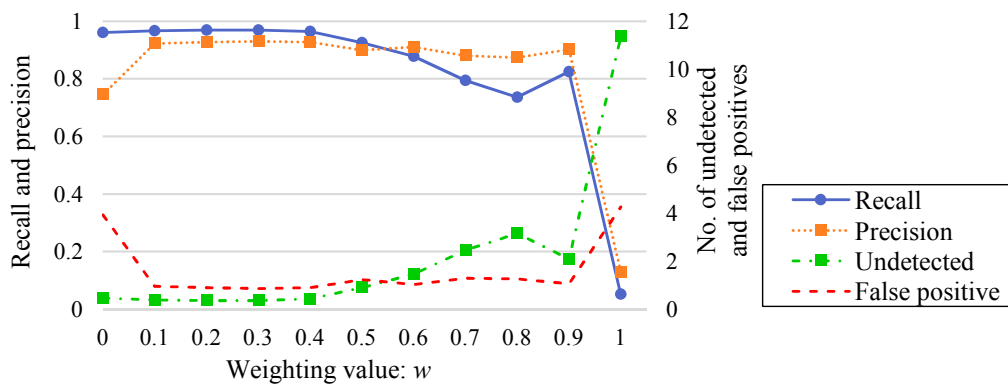


Fig.5.6 Relationship between recognition performance and the value w .

認識実験に際しては、シーンデータに割り当てられている正解データの姿勢に対して、回転と平行移動成分に外乱を加えた新たな姿勢仮説を生成し、認識実験用データセットを作成した。各シーンデータに対して合計 100 個の姿勢仮説を用意したため、組み合わせ最適化で解く探索範囲は 2^{100} である。

5.3.2 配置実現性の有効性

重み係数 w は提案手法の認識性能を左右する最も重要なパラメータである。本実験では、 w の値を範囲を $[0, 1]$ で変化させながら認識性能の変化を調べた。データセットは Kinect を用いた。結果を Fig.5.6 に示す。図において、縦軸は Recall, Precision の値、横軸は w の値を示す。

$w = 0$ のとき: w が 0 の時は、配置実現性を使用せずに、形状の整合性のみを用いて最適化することを示す。このときは、物体同士の空間的な重なりを考慮しないため、過検出が多く、Precision が低くなった。12 個の物体が置いてあるシーン (30 シーン) において、平均 15.5 個の物体を検出した。

$0 < w$ のとき: 配置実現性を考慮することによって、1 シーン当たりの過検出数が 1 以下に削減されることを確認した。したがって、認識率の向上のためには、物体配置の実現性を考慮することがよいと考えられる。しかしながら、 w が大きくなりすぎると Recall が低下した。

このときの典型的な認識結果を Fig.5.7 に示す。(a) は入力シーンの距離画像、(b) が認識結果、(c) は正解のシーン仮説である。(b) において、点線で囲んだ 2 つの物体が検出できていなかった。2 つの物体に共通することは、距離画像中に占める面積が小さいということである。これは、そのような物体が正しく検出できたとしても、形状の整合性の数値の上昇が少なく、最終的な適応度が大きく向上しないことを意味しており、結果として未検出の状態のまま解が収束したと考えられる。

最も性能が高かったのは、 $w = 0.3$ のときである。以降の実験では、この値を用いて認識性能を比較する。

5.3.3 認識性能の比較

提案手法の有効性を定量的に評価するために、2 種類のデータセットに対して認識実験を実施した。比較手法は以下の 3 手法である。

1. GoHV [1]
2. Proposed HV (S)
3. Proposed HV (S+P)

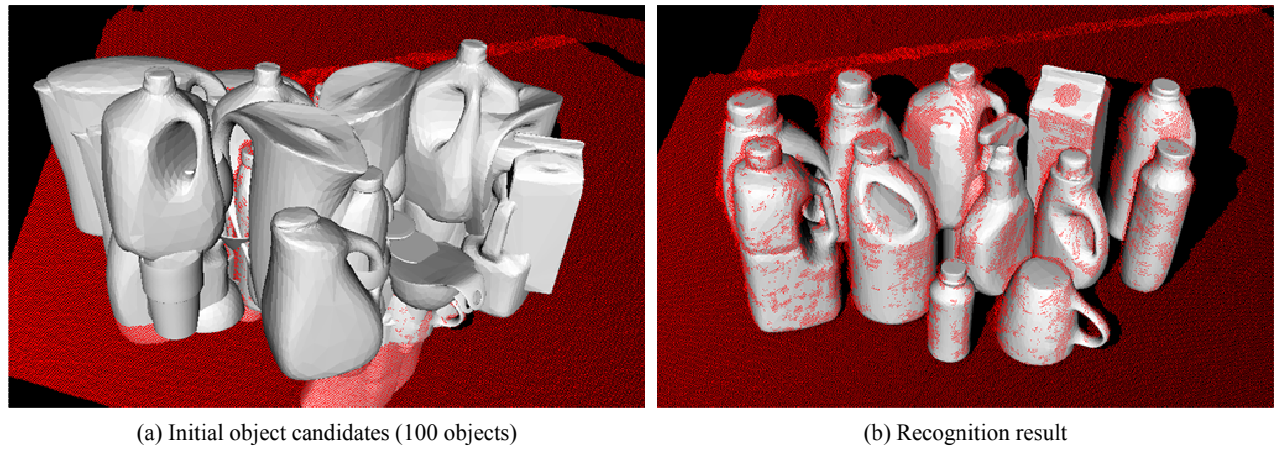
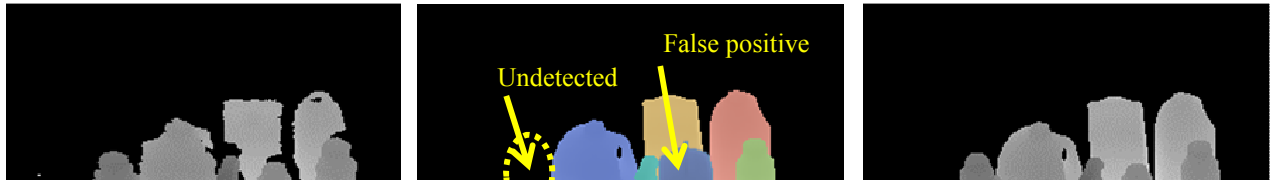


Fig.5.8 Recognition result. (a) Initial object hypotheses (surface representation). (b) Final recognition result.

GoHV は Aldoma らによって提案された仮説検証アルゴリズムである。Proposed HV(S) は提案した仮説検証型アルゴリズムにおいて、形状の整合性のみを適応度とした手法である Proposed HV(S+P) は、提案した仮説検証型アルゴリズムにおいて、形状的整合性と配置実現性を同時に利用した手法である。配置実現性と形状的整合の結合重み $w = 0.3$ である。いずれの手法も C++ 言語で実装しており、Point Cloud Library[41] を用いた。実験に用いた PC のスペックは CPU : Intel Core i7-6700 3.40GHz, RAM : 16GB である。

提案手法による認識結果例を Fig.5.8 に示す。(a) は Fig.5.3(a) の入力シーンに設定した物体候補群を同時に表示した結果である。合計 100 個の物体候補がレンダリングされており、正解を 12 個、誤認識結果を 88 個含めている。物体仮説をメッシュデータとしてレンダリングし、入力シーンは赤色の点群としてレンダリングした。(b) は提案手法による認識の結果、残った物体仮説である。物体の上部に表示されている数字は物体候補のインデクスである。12 個の物体仮説が残ったが、これらはすべて正しい物体仮説であった。

Table 5.1 Recognition performance for Laser Scanner dataset. (T : processing time)

Method	Recall	Precision	F-measure	T [msec]
--------	--------	-----------	-----------	------------

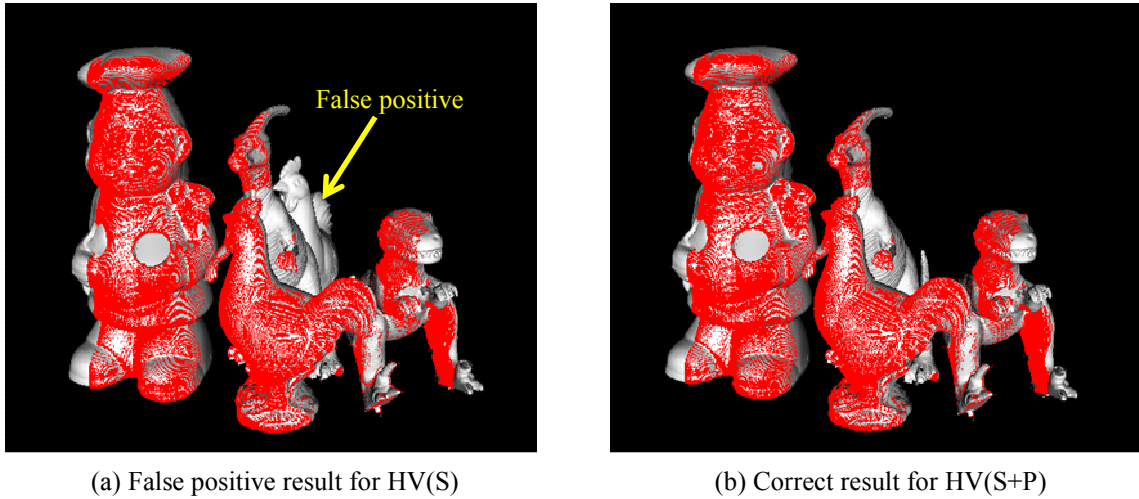


Fig.5.9 Example of the false positive of the HV(S) and the correct result of the HV(S+P).

Table 5.2 Recognition performance for Kinect dataset. (T : processing time)

Method	Recall	Precision	F-measure	T [msec]
GoHV[1]	0.91	0.94	0.93	2559
HV (S)	0.96	0.75	0.84	713
HV (S+P)	0.97	0.93	0.95	613

表 5.1 は各手法における Laser scanner dataset に対する認識性能である。手法 HV (S) の結果を見ると、Precision が Recall よりも低い結果となった。これは、この手法が過検出を多く起こしやすかったことに起因する。過検出された物体候補は Fig.5.9(a) に示すように、空間的に重なっており、現実には再現しえない物体配置となっていた。提案手法 (HV(S+P)) では配置実現性を利用するので、適応度が低下させることができ、この問題を解決することができた (Fig.5.9(b))。したがって、形状的整合性のみを用いる HV (S) と比較して、Precision が改善された。

表 5.2 は各手法における Kinect dataset に対する認識性能である。このデータセットはシーン中の物体数やノイズ量多いことと、点群密度が低い、さらに考慮しなければならない物体モデル数が多い (35 種類) ため、Laser scanner dataset よりも難易度が高いといえる。このデータセットに対しては、GoHV の Precision が最も高かったが、提案手法もほぼ同等の結果であった。提案手法は Recall も同時が高かったため、比較手法の中で最も認識信頼性が高かったといえる。処理時間の観点から比較すると、提案手法は最も高速であり、従来の GoHV 法のおよそ 4 倍の処理速度であることを確認した。

5.4 結論

本章では、大局的な情報を用いた仮説検証型の物体認識手法を提案した。提案手法では、入力シーン全体とシーン中の各物体の姿勢候補（物体候補）の組み合わせによって構成されるシーン仮説を照合する。入力シーン中に存在するすべての物体に対して解釈を与えるため、局所的なミスマッチングを棄却することができる。この手法では、物体候補を事前に生成することが必要であるが、前章までに説明したモデルマッチング法を利用することが可能である。得られた物体候補を組み合わせ最適化のプロセスによって取捨選択することによって、入力シーンと一致性の高いシーン仮説を生成し、複数の物体の配置を同時に認識する。

提案手法は、入力シーンとシーン仮説の形状的整合性に加えて、シーン仮説を構成する各物体の配置実現性を検証することが特徴である。これには、高速に干渉判定が可能な干渉モデルを提案し、物体の空間的な重なりを評価することが可能なことを説明した。有名な公開データベースを用いた実験の結果、信頼性・処理速度の両観点において従来の仮説検証型アルゴリズムと比較して優位であることを実証した。

以下に、本章で明らかになった結論を要約する。

- (1) シーン仮説の配置実現性を検証する機能をための方法として、物体モデルごとに干渉モデルを利用することを説明した。物体同士の干渉をチェックするための最も簡単なプリミティブ形状は球であることから、干渉モデルを複数の球の集合として表現することを提案した。
- (2) 提案する仮説検証型アルゴリズムでは、入力シーンとシーン仮説の形状的整合性とシーン仮説の配置実現性の2つの指標を評価関数とし、これを最大化する。このとき、2つの指標の重み係数が認識性能を左右することを説明し、この比率を配置実現性を0.3、形状的整合性を0.7に設定した時に認識性能が最も高くなることを示した。
- (3) 3D物体認識のベンチマークとして有名なデータセットを2種類（Laser Scanner dataset, Kinect dataset）用いて、認識性能を比較した。提案手法は、従来の仮説検証型アルゴリズム [1] と比較して認識信頼性が向上したことを実証した。また、処理速度においては従来手法と比較して4倍高速であることを実証した。

本章に係る成果のベースとなる技術は以下の1.の国際会議にて発表済みであり、最新の成果は第19回画像の認識・理解シンポジウム（MIRU）のオーラル発表（査読有）に採択済みである。

1. Shuichi Akizuki and Manabu Hashimoto, “Multiple 3D Object Recognition using RGB-D Data and Physical Consistency for Automated Warehousing Robots”, *11th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2016.
2. Shuichi Akizuki and Manabu Hashimoto, “Multiple 3D Object Recognition using Shape Consistency and Physical Possibility”, 第19回画像の認識・理解シンポジウム（MIRU）, 2016. (査読有, オーラル)

第6章 結言

6.1 本論文の成果の要約

本論文では、製造業、および物流・生活支援応用を志向した物体認識システムのためのモデルマッチング法の機能強化手法を提案した。以下に、各分野で駆動されるロボットビジョンシステムへの本研究の貢献を総括する。

6.1.1 製造分野における貢献

製造分野において、ロボットアームによるばら積み部品ピッキングシステムの実用化のためには、以下の2つの問題が存在していた。

- (1) ロボットの動作スピードに認識の処理時間が追いついていない。
- (2) 信頼性良く認識することのできる物体形状に制約がある。

これに対して、(1)に対する解決策として高速なモデルマッチングである、ベクトルペアマッチング(VPM)法を提案した。VPM法の処理速度は平均1.36秒であった。一般的なロボットアームの動作サイクルタイムの2秒であるため、十分に高速であることを実証した。さらに、4種類のデータセットを用いて従来の3次元位置姿勢認識手法であるSHOT特徴量やFPFH特徴量を用いた手法と比較し、処理時間と認識率の両観点において優位であることを確認した。

(2)に対する解決策として、照合戦略切り替え型モデルマッチング法を提案した。従来のモデルマッチング法は、平面的な形状の物体に対する認識性能が低いことが問題点であった。平面形状の認識に特化した特徴表現であるGlobal Reference Frame(GRF)を提案し、VPM法とGRFを適応的に切り替えながらマッチングすることによって、対象物の凹凸密度に依存しない位置姿勢認識を実現した。

6.1.2 サービス分野における貢献

物流・生活支援を志向したロボットビジョンの実用化のためには、移動のための地図生成技術、人物認識技術、障害物回避技術、特定物体認識技術等、さまざまな技術のブレイクスルーが必要である。その中で、本研究では指示された物体を確実に認識するためのビジョン技術を開発した。

実用的な認識性能を得るためには、従来のモデルマッチング法は以下の2つの問題点を解決しなければならないと考えられる。(3)形状データ(3次元点群)の密度変化に対する頑健性が不十分。(4)多種類の物体が密に配置されている場合には認識性能が低下する。

(3)に関しては、この分野への用途を前提とした場合には、ワーキングディスタンスを固定しにくいという状況が発生する。この場合には計測データの密度が物体モデルと異なる場合が多いため、局所参照座標系(LRF)の再現性が低下することが誤認識の原因になっていた。本研究では、このような状況においても再現性の高いLRFを提案し、モデルマッチングの性能向上を実現した。マッチング対象の点密度が異なることは、本用途へビジョン技術を転用する際に発生する基本的な外乱である。したがって、本論文にて提案したキーポイントマッチング法はこの用途における利便性が極めて高いといえる。

(4)に関しては、従来のモデルマッチング法は物体モデルと入力シーンの一部、すなわち局所的な情報のみを使って一致度合い(スコア)を算出していたために、入力シーン中に多種類の物体が密に配置され

ている時には、誤った領域とのスコアの方が高くなる場合があり、誤認識が誘発されることが問題点であった。本研究では、モデルマッチング法によって得られた、個々の物体候補を複数組み合わせることによって生成したシーン仮説と、入力シーン全体、すなわち大局的情報を用いた類似度指標を定義することによって、局所的な誤認識を棄却する枠組みを提案した。このとき、物体候補同士の配置の実現可能性も検証することによって、従来の仮説検証法と比較して、認識信頼性が優位であることを確認した。提案手法は物体モデルが 35 種類登録されているデータセットに対して、F 値 0.95 を達成している。実応用上はロボットに搭載されている他のセンサの併用や、ロボットとの対話的処理によっても認識性能の向上が見込めるため、提案手法の認識性能は実用性能を達成していると考えられる。

6.2 今後の展望

本論文で提案した物体認識手法は、認識対象の物体が剛体の 3D 形状モデルとして表現可能であることが前提となっている。しかしながら、物流ロボットや生活支援ロボットで扱うべき対象物は、衣服や袋もの等のいわゆる非剛体物体も多い。これらの物体は剛体の形状モデルを定義できないため、モデルマッチング法によって位置姿勢の候補を生成することができない。

提案した仮説検証型アルゴリズムで利用している情報は、物体の距離画像としての見え方と干渉モデルの位置情報である。すなわち、もし対象物の位置姿勢が認識できなかったとしても、距離画像における対象物の領域をセグメンテーションし、その領域が占有しているボリュームを知覚することができれば、これまでと同一のフォーマットとして物体仮説を生成することができるため、提案手法の枠組みで認識することが可能である。したがって、今後の展望としては、提案した仮説検証型アルゴリズムが取り扱うことのできる物体の材質的な制約を解消するために、非剛体の物体に対する物体候補生成手法の確立が挙げられる。

また、認識した物体をロボットによってハンドリングするという観点においては、物体の掴むべき位置を、後に続く動作に応じて適応的に判断しなければならない。例えば、物流ロボットが物体を棚から取り出して箱に詰めるだけならば、対象物を持ち上げることさえできればいいので、掴み位置に対する制約が少ない。しかしながら、生活支援ロボットがコップを持ち上げて中に液体を注ぐ場合には、取っ手を掴んでこぼれない向きに保持しなければならない。したがって、今後はタスクを作用させる位置を認識することにも着手する必要があると考えられる。

参考文献

- [1] Aitor Aldoma, Federico Tombari, Luigi di Stefano, and Markus Vincze. A global hypotheses verification method for 3d object recognition. In *Proceedings of the 12th European Conference on Computer Vision (ECCV)*, pp. 511–524, 2012.
- [2] Aitor Aldoma, Federico Tombari, Johann Prankl, Andreas Richtsfeld, Luigi di Stefano, and Markus Vincze. Multimodal cue integration through hypotheses verification for RGB-D object recognition and 6dof pose estimation. In *IEEE International Conference on Robotics and Automation ICRA*, pp. 2104–2111, 2013.
- [3] Aitor Aldoma, Federico Tombari, Radu Bogdan Rusu, and Markus Vincze. OUR-CVFH - oriented, unique and repeatable clustered viewpoint feature histogram for object recognition and 6dof pose estimation. In *Proceedings of the Joint Pattern Recognition Symposium 34th DAGM, 36th OAGM*, pp. 113–122, 2012.
- [4] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, Vol. 13, No. 2, pp. 111–122, 1981.
- [5] Paul Besl and Neil McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 14, No. 2, pp. 239–256, 1992.
- [6] Blender. <http://blender.jp/>.
- [7] Amazon Picking Challenge. <http://amazonpickingchallenge.org/>.
- [8] Hui Chen and Bir Bhanu. 3D free-form object recognition in range images using local surface patches. *Pattern Recognition Letters*, Vol. 28, No. 10, pp. 1252–1262, 2007.
- [9] Changhyun Choi, Yuichi Taguchi, Oncel Tuzel, Ming-Yu Liu, and Srikumar Ramalingam. Voting-based pose estimation for robotic assembly using a 3D sensor. *IEEE International Conference on Robotics and Automation ICRA*, pp. 1724–1731, 2012.
- [10] Chin-Seng Chua and Ray Jarvis. Point signatures: A new representation for 3d object recognition. *International Journal of Computer Vision (IJCV)*, Vol. 25, No. 1, pp. 63–85, 1997.
- [11] Chitra Dorai and Anil K. Jain. COSMOS - A representation scheme for 3d free-form objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 19, No. 10, pp. 1115–1130, 1997.
- [12] Thiago R. dos Santos, Alfred M. Franz, Hans-Peter Meinzer, and Lena Maier-Hein. Robust multi-modal surface matching for intra-operative registration. In *Proceedings of the 24th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 1–6, 2011.
- [13] Bertram Drost and Slobodan Ilic. 3d object detection and localization using multimodal point pair features. In *Proceedings of the Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission (3DIMPVT)*, pp. 9–16, 2012.
- [14] Bertram Drost, Markus Ulrich, Nassir Navab, and Slobodan Ilic. Model globally, match locally: Efficient and robust 3d object recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 998–1005, 2010.

- [15] Andrea Frome, Daniel Huber, Ravi Kolluri, Thomas Bülow, and Jitendra Malik. Recognizing objects in range data using regional point descriptors. In *Proceedings of the 8th European Conference on Computer Vision (ECCV)*, pp. 224–237, 2004.
- [16] Taiki Fuji, Nobutaka Kimura, and Kiyoto Ito. Architecture for recognizing stacked box objects for automated warehousing robot system. In *Proceedings of the Irish Machine Vision and Image*, pp. 51–58, 2015.
- [17] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989.
- [18] Yulan Guo, Ferdous Ahmed Sohel, Mohammed Bennamoun, Min Lu, and Jianwei Wan. Rotational projection statistics for 3D local surface description and object recognition. *International Journal of Computer Vision (IJCV)*, Vol. 105, No. 1, pp. 63–86, 2013.
- [19] Manabu Hashimoto, Kazuhiko Sumi, and Teruo Usami. Recognition of multiple objects based on global image consistency. In *Proceedings of the British Machine Vision Conference (BMVC) 1999, Nottingham, 13-16 September 1999*, pp. 1–10, 1999.
- [20] G Hetzel, B Leibe, P Levi, and B Schiele. 3D object recognition from range images using local feature histograms. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. II–394–II–399, 2001.
- [21] B. K. P. Horn. Extended gaussian images. *Proceedings of the IEEE*, Vol. 72, No. 12, pp. 1671–1686, 1984.
- [22] Subramaniam Jayanti, Yagnanarayanan Kalyanaraman, Natraj Iyer, and Karthik Ramani. Developing an engineering shape benchmark for CAD models. *Computer-Aided Design*, Vol. 38, No. 9, pp. 939–953, 2006.
- [23] 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 (PAMI)*, Vol. 21, No. 5, pp. 433–449, 1999.
- [24] Sagi Katz, Ayellet Tal, and Ronen Basri. Direct visibility of point sets. *ACM Transactions on Graphics (TOG) - Proceedings of ACM*, Vol. 26, No. 3, p. 24, 2007.
- [25] Yan Ke and Rahul Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 506–513, 2004.
- [26] Eunyoung Kim and Gérard G. Medioni. 3d object recognition in range images using visibility context. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3800–3807, 2011.
- [27] Ming-Yu Liu, Oncel Tuzel, Ashok Veeraraghavan, Rama Chellappa, Amit K. Agrawal, and Haruhisa Okuda. Pose estimation in heavy clutter using a multi-flash camera. In *IEEE International Conference on Robotics and Automation ICRA*, pp. 2028–2035, 2010.
- [28] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV)*, Vol. 60, No. 2, pp. 91–110, 2004.
- [29] Ajmal S. Mian, Mohammed Bennamoun, and Robyn Owens. Three-dimensional model-based

- object recognition and segmentation in cluttered scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 28, No. 10, pp. 1584–1601, 2006.
- [30] Ajmal S. Mian, Mohammed Bennamoun, and Robyn A. Owens. On the repeatability and quality of keypoints for local feature-based 3d object retrieval from cluttered scenes. *International Journal of Computer Vision (IJCV)*, Vol. 89, No. 2-3, pp. 348–361, 2010.
- [31] Hiroshi Murase and Shree K. Nayar. Visual learning and recognition of 3-d objects from appearance. *International Journal of Computer Vision (IJCV)*, Vol. 14, No. 1, pp. 5–24, 1995.
- [32] John Novatnack and Ko Nishino. Scale-dependent / invariant local 3d shape descriptors for fully automatic registration of multiple sets of range images. In *Proceedings of the 10th European Conference on Computer Vision (ECCV)*, pp. 440–453, 2008.
- [33] Chavdar Papazov and Darius Burschka. An efficient RANSAC for 3D object recognition in noisy and occluded scenes. *Proceedings of the 10th Asian Conference on Computer Vision (ACCV)*, Vol. 6492 LNCS, No. PART 1, pp. 135–148, 2010.
- [34] In Kyu Park, Marcel Germann, Michael D. Breitenstein, and Hanspeter Pfister. Fast and automatic object pose estimation for range images on the GPU. *Proceedings of the IAPR International Conference on Machine Vision Applications (MVA)*, Vol. 21, No. 5, pp. 749–766, 2010.
- [35] Alioscia Petrelli and Luigi di Stefano. On the repeatability of the local reference frame for partial shape matching. In *Proceedings of the International Conference on Computer Vision (ICCV)*, pp. 2244–2251, 2011.
- [36] Andreas Richtsfeld, Thomas Morwald, Johann Prankl, Michael Zillich, and Markus Vincze. Segmentation of unknown objects in indoor environments. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4791–4796, 2012.
- [37] RoboCup@Home. <http://www.robocupathome.org/>.
- [38] Radu Bogdan Rusu. [phd thesis]semantic 3d object maps for everyday manipulation in human living environments. 2010.
- [39] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (FPFH) for 3d registration. In *IEEE International Conference on Robotics and Automation ICRA*, pp. 3212–3217, 2009.
- [40] Radu Bogdan Rusu, Nico Blodow, Zoltan Csaba Marton, and Michael Beetz. Aligning point cloud views using persistent feature histograms. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3384–3391, 2008.
- [41] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (PCL). In *IEEE International Conference on Robotics and Automation ICRA*, pp. 1–4, 2011.
- [42] S. A. A. Shah, M. Bennamoun, F. Boussaid, and A. A. El-Sallam. A novel local surface description for automatic 3d object recognition in low resolution cluttered scenes. In *Proceedings of the International Conference on Computer Vision (ICCV) Workshops*, pp. 638–643, 2013.
- [43] SketchUp. <http://www.sketchup.com/>.
- [44] Sketchyphysics. <https://code.google.com/p/sketchyphysics/>.

- [45] Fridtjof Stein and Gérard G. Medioni. Structural indexing: Efficient 3-d object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 14, No. 2, pp. 125–145, 1992.
- [46] Yiyong Sun and Mongi A. Abidi. Surface matching by 3d point’s fingerprint. In *Proceedings of the International Conference on Computer Vision (ICCV)*, pp. 263–269, 2001.
- [47] Federico Tombari and Luigi Di Stefano. Object recognition in 3d scenes with occlusions and clutter by hough voting. In *Proceedings of the 4th Pacific-Rim Symposium on image and Video Technology (PSIVT)*, pp. 349–355, 2010.
- [48] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *Proceedings of the 11th European Conference on Computer Vision (ECCV)*, pp. 356–369, 2010.
- [49] Federico Tombari, Samuele Salti, and Luigi di Stefano. A combined texture-shape descriptor for enhanced 3d feature matching. In *18th IEEE International Conference on Image Processing (ICIP)*, pp. 809–812, 2011.
- [50] Dimitrios Tzionas and Juergen Gall. 3d object reconstruction from hand-object interactions. In *Proceedings of the International Conference on Computer Vision (ICCV)*, pp. 729–737, 2015.
- [51] Andre Uckermann, Christof Elbrechter, Robert Haschke, and Helge J. Ritter. 3d scene segmentation for autonomous robot grasping. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1734–1740, 2012.
- [52] E. Wahl, U. Hillenbrand, and G. Hirzinger. Surflet-pair-relation histograms: a statistical 3d-shape representation for rapid classification. In *Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling (3DIM)*, pp. 474–481, Oct 2003.
- [53] Rui Wang, Kun Zhou, John Snyder, Xinguo Liu, Hujun Bao, Qunsheng Peng, and Baining Guo. Variational sphere set approximation for solid objects. *The Visual Computer*, Vol. 22, No. 9-11, pp. 612–621, 2006.
- [54] Thibaut Weise, Bastian Leibe, and Luc J. Van Gool. Accurate and robust registration for in-hand modeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [55] Simon Winkelbach, Sven Molkenstruck, and Friedrich M. Wahl. Low-cost laser range scanner and fast surface registration approach. In *Proceedings of the 28th Conference on Pattern Recognition (DAGM)*, DAGM’06, pp. 718–728, Berlin, Heidelberg, 2006. Springer-Verlag.
- [56] Andrei Zaharescu, Edmond Boyer, Kiran Varanasi, and Radu Horaud. Surface feature detection and description with applications to mesh matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 373–380, 2009.
- [57] Yu Zhong. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In *Proceedings of the International Conference on Computer Vision (ICCV) Workshops*, pp. 689–696, 2009.
- [58] 国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) . <http://www.nedo.go.jp/content/100080673.pdf>.

- [59] 武口智行, 金子俊一. 距離アスペクト画像照合による高効率なロバスト探索. 電気学会論文誌 C, Vol. 124, No. 3, pp. 637–646, mar 2004.
- [60] 北明靖雄, 奥田晴久, 橋本学, 金子俊一. 距離アスペクト画像照合と階層化 icp 照合による 3 次元物体認識. 電気学会論文誌 C, Vol. 127, No. 4, pp. 615–622, 2007.
- [61] 立野圭祐, 小竹大輔, 内山晋二. ビンピッキングのための距離・濃淡画像を最ゆうに統合する高精度高安定なモデルフィッティング手法. 電子情報通信学会論文誌 D, Vol. 94, No. 8, pp. 1410–1422, aug 2011.

謝 辞

本研究を遂行するにあたり、多くのご教授を受け賜りました橋本学教授に厚く御礼申し上げます。私が橋本研究室の第1期生として2009年に配属されてから7年間、終始熱心にご指導いただきました。

また、半期短縮での学位取得であったため、タイトな審査スケジュールになっていたかと思います。それにもかかわらずご多忙の中、副査を引き受けてくださった輿水大和教授、長谷川純一教授、目加田慶人教授、青木公也教授に感謝いたします。

ロボットピッキングシステムの構築にあたっては、橋本研究室OBで現日産自動車(株)の永瀬誠信氏と武井翔一氏の協力を頂きました。感謝いたします。さらに、日頃から熱心に討論してくださった橋本研究室の諸兄諸氏に感謝いたします。

本研究の成果の一部は特別研究員奨励費(16J09262)の助成を受けたものである。

本論文に関する発表論文

学術論文

- (1) 秋月秀一, 橋本学, “3次元キーポイントマッチングのための点群密度変化および欠落に頑健な Local Reference Frame”, 電子情報通信学会 D, Vol.J99-D, No.8, 2016. (採録決定)
- (2) Shuichi Akizuki, Manabu Hashimoto, “Stable Position and Pose Estimation of Industrial Parts using Evaluation of Observability of 3D Vector Pairs”, *Journal of Robotics and Mechatronics (Special Issue on Vision and Motion Control)*, Vol.27, No.2, pp.174-181, 2015.
- (3) 秋月秀一, 橋本学, “Global Reference Frame を用いた対象物の見えに依存しない 3次元位置姿勢認識”, 精密工学会誌, Vol.80, No.12, pp.1176-1181, 2014.
- (4) 秋月秀一, 橋本学, “特徴的 3-D ベクトルペアを用いたばら積み部品的高速位置姿勢認識”, 電気学会論文誌 C, Vol.133, No.9, pp.1853-1854, 2013.

国際会議

- (1) Shuichi Akizuki and Manabu Hashimoto, “Multiple 3D Object Recognition using RGB-D Data and Physical Consistency for Automated Warehousing Robots”, *11th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2016.
- (2) Shuichi Akizuki and Manabu Hashimoto, “Relative Point Density (RPD) Feature for Object Recognition Independent of Point Cloud Sparseness”, *The Korea-Japan joint workshop on Frontiers of Computer Vision (FCV)*, 2016.
- (3) Shuichi Akizuki and Manabu Hashimoto, “DPN-LRF: A Local Reference Frame for Robustly Handling Density Differences and Partial Occlusions”, *11th International Symposium on Visual Computing (ISVC)*, Part I, LNCS 9474, pp.878-887, 2015.
- (4) Shuichi Akizuki and Manabu Hashimoto, “A Proposal of the Global Reference Frame for Surface Flatness-independent 3D Object Detection”, *Proc. Joint Conference of the International Workshop on Advanced Image Technology (IWAIT) and the International Forum on Medical Imaging in Asia (IFMIA)*, OS. 27, 2015.
- (5) Shuichi Akizuki and Manabu Hashimoto, “Position and Pose Recognition of Randomly Stacked Objects using Highly Observable 3D Vector Pairs”, *Proc. the 40th Annual Conference of the IEEE Industrial Electronics Society*, pp.5266-5271, Oct. 2014.
- (6) Shuichi Akizuki and Manabu Hashimoto, “Fast and Reliable 3-D Object Recognition based on Surface Normal Distributions”, *Proceeding of International Symposium on Optomechatronic Technologies (ISOT)*, T1A-1, pp.1-9, Oct. 2013.
- (7) Shuichi Akizuki and Manabu Hashimoto, “High-speed and Reliable Object Recognition using Distinctive 3-D Vector Pairs in a Range Image”, *Proceeding of IEEE International Symposium on Optomechatronic Technologies (ISOT)*, pp.1-6, Oct. 2012.

国内シンポジウム

- (1) Shuichi Akizuki and Manabu Hashimoto, “Multiple 3D Object Recognition using Shape Con-

- sistency and Physical Possibility”, 第 19 回画像の認識・理解シンポジウム (MIRU), 2016. (オーラル, 採録決定)
- (2) 秋月秀一, 橋本学, “RGB-D 整合性と物理的整合性評価に基づく 3 次元多品種物品の識別と姿勢認識”, ビジョン技術の実利用ワークショップ (ViEW), OS2-H4(IS1-18), pp.78-81, Dec. 2015.
 - (3) 秋月秀一, 橋本学, “点群の密度変化と欠落に頑健な Local Reference Frame の提案”, 平成 26 年度電気・電子・情報関係学会東海支部連合大会, 2015. 【連合大会奨励賞】
 - (4) 秋月秀一, 橋本学, “3 次元キーポイントマッチングのための点群密度変化と欠落に頑健な Local Reference Frame”, 第 18 回画像の認識・理解シンポジウム (MIRU), SS5-36, pp.1 - 2, Jul. 2015. 【インタラクティブ発表賞】
 - (5) 秋月秀一, 橋本学, “安定的な 3D 特徴量記述のための点群の密度変化にロバストな Local Reference Frame”, ビジョン技術の実利用ワークショップ (ViEW), IS1-31, pp.1-4, 2014.
 - (6) 秋月秀一, 橋本学, “凹凸密度に依存しない 3D 姿勢推定のための Global Reference Frame の提案”, 平成 26 年度電気・電子・情報関係学会東海支部連合大会, H2-2, Sept. 2014..
 - (7) 秋月秀一, 橋本学, “Global Reference Frame を用いた形状データの凹凸に依存しない 3 次元位置姿勢認識”, 第 17 回画像の認識・理解シンポジウム (MIRU), SS2-27, pp.1 - 2, Jul. 2014.
 - (8) 秋月秀一, 橋本学, “Global Reference Frame:平面的な形状のための 3D オリエンテーションを用いた表面凹凸密度に依存しない位置姿勢推定”, 第 19 回知能メカトロニクスワークショップ (iMec2014), A2-4, pp.82-86, Jul. 2014. 【優秀講演賞】
 - (9) 秋月秀一, 橋本学, “照合戦略選択に基づく対象物の見えの変動にロバストな 3 次元位置姿勢認識”, 第 19 回画像センシングシンポジウム (SSII), pp.IS2-03-1-5, Jun. 2014.
 - (10) 秋月秀一, 橋本学, “可観測性を考慮した 3D ベクトルペア選択によるばら積み部品の位置姿勢認識”, ビジョン技術の実利用ワークショップ (ViEW), OS5-O2, Dec. 2013.
 - (11) 秋月秀一, 橋本学, “局所法線分布を記述子とする高信頼な 3 次元物体検出”, 精密工学会・画像応用技術専門委員会サマーセミナー 2013, Vol.37, No.36, pp.27-30, Aug. 2013.
 - (12) 秋月秀一, 橋本学, “サーフェスモデルから選択した特徴的 3D ベクトルペアによる高信頼な位置姿勢認識”, 画像の認識・理解シンポジウム (MIRU), SS5-8, Jul. 2013.
 - (13) 秋月秀一, 武井翔一, 橋本学, “特徴的ベクトルペアによる高速・高信頼なばら積み物体認識”, 画像応用技術専門委員会報告, Vol.28, No.2, pp.24-29, Jul. 2013.
 - (14) 秋月秀一, 橋本学, “全周囲物体認識のためのサーフェスモデルからの高独自性ベクトルペア選択”, 第 19 回画像センシングシンポジウム (SSII2013), pp.IS2-24-1-6, Jun. 2013.
 - (15) 秋月秀一, 橋本学, “3-D ベクトルペアを用いた複雑シーンからの高速物体検出”, 精密工学会・画像応用技術専門委員会サマーセミナー, Vol.21, pp.9-12, Sep. 2012. 【優秀発表賞】
 - (16) 秋月秀一, 橋本学, “特徴的 3-D ベクトルペアを用いたばら積み部品の高速位置姿勢認識”, 画像の認識・理解シンポジウム (MIRU), IS1-65, Aug. 2012.
 - (17) 秋月秀一, 橋本学, “特徴的 3-D ベクトルペアマッチングによる高速位置姿勢認識”, 画像センシングシンポジウム (SSII), IS4-15, Jun. 2012. 【優秀学術賞】
 - (18) 秋月秀一, 橋本学, “特徴的 3-D ベクトルペアマッチングによるばら積み部品の高速認識”, 情報処理学会コンピュータビジョンとイメージメディア研究会 (CVIM), Vol.2012-CVIM-182, No.15, pp.1-7, May 2012. 【最優秀賞】

受賞

- (1) 平成 26 年度電気・電子・情報関係学会東海支部連合大会 連合大会奨励賞, 2016/1/20.
- (2) 第 18 回画像の認識・理解シンポジウム (MIRU2015) インタラクティブ発表賞, 2015/7/30.
- (3) IEEE ICRA2015 Amazon Picking Challenge Travel Reimbursement Awards Team C^2M (チームメンバーとしての受賞), 2015/3/26.
- (4) 第 19 回知能メカトロニクスワークショップ (iMec2014) 優秀講演賞, 2014/7/13.
- (5) 第 18 回画像センシングシンポジウム (SSII2012) 優秀学術賞, 2013/6/13.
- (6) 精密工学会画像応用技術専門委員会・映像情報メディア学会メディア工学研究委員会合同サマーセミナー 優秀発表賞, 2012/9/4.
- (7) 情報処理学会コンピュータビジョンとイメージメディア研究会 (CVIM) 卒論セッション 最優秀賞, 2012/5/24.

その他の発表論文

学術論文

- (1) 武井翔一, 秋月秀一, 橋本学, “マルチスケールシエル領域の点群占有率に基づく 3 次元特徴量の提案”, 電気学会論文誌 C, Vol. 136, No. 8, 2016. (採録決定)
- (2) 橋本学, 秋月秀一, 武井翔一, “物体認識のための 3 次元特徴量の研究動向”, 電気学会論文誌 C, Vol. 136, No. 8, 2016. (採録決定)
- (3) 櫻本泰憲, 兼松裕一, 大野広揮, 秋月秀一, 橋本学, 渡邊清高, 関真規人, “Coded Multi Flush Imaging による 3 次元凹凸構造特徴マッチング”, 精密工学会誌, Vol.81, No.10, pp.944-950, 2015.
- (4) 武井翔一, 秋月秀一, 橋本学, “識別性能の予測に基づく選択的特徴量を用いたばら積み部品の認識”, 精密工学会誌, Vol.81, No.4, pp.363-367, 2015.
- (5) 永瀬誠信, 秋月秀一, 橋本学, “CCDoN: ばら積み部品の高速・高信頼な 6 自由度位置姿勢推定のための局所特徴量”, 精密工学会誌, Vol.80, No.12, pp.1138-1143, 2014.
- (6) 永瀬誠信, 秋月秀一, 橋本学, “誤照合を最小化する 3-D 特徴点を用いた高信頼な物体認識手法”, 精密工学会誌, Vol.79, No.11, pp.1058-1062, 2013.
- (7) 秋月秀一, 橋本学, “最適配置された画素群の濃度共起発生確率に着目した画像のテクスチャ量にロバストな照合手法”, 電気学会論文誌 C, Vol.133, No.10, pp.1943-1949, 2013.

国際会議

- (1) K.Tobitani, S.Akizuki, K.Katahira, M.Hashimoto, N.Nagata, “A Comparison Study on 3D Features in Term of Effective Representation for Impression of Shape”, The 2nd International Conference on DigitalFabrication (ICDF), No.22, 2016.
- (2) Shoichi Takei, Shuichi Akizuki, Manabu Hashimoto, “SHORT: A Fast 3D Feature Description based on Estimating Occupancy in Spherical Shell Regions”, *Proceedings of the 30th International Conference on Image and Vision Computing New Zealand(IVCNZ)*, Nov. 2015.
- (3) Masanobu Nagase, Shuichi Akizuki, Manabu Hashimoto, “High-speed and Reliable Object Recognition based on Low-dimensional Local Shape Features”, *Proc. the 13th International Conference on Control, Automation, Robotics and Vision(ICARCV)*, pp.82-87, Dec. 2014.
- (4) Shoichi Takei, Shuichi Akizuki, Manabu Hashimoto, “3D Object Recognition using Effective Features Selected by Evaluating Performance of Discrimination”, *Proc. the 13th International Conference on Control, Automation, Robotics and Vision(ICARCV)*, pp.70-75, Dec. 2014.
- (5) Masanobu Nagase, Shuichi Akizuki, Manabu Hashimoto, “3-D Feature Point Matching for Object Recognition Based on Estimation of Local Shape Distinctiveness”, *15th International Conference on Computer Analysis of Images and Patterns(CAIP)*, Part I, LNCS, Vol.8047, pp.473-481, Aug. 2013.
- (6) Shuichi Akizuki and Manabu Hashimoto, “Robust Matching for Low-texture Images based on Co-occurrence of Geometry-optimized Pixel Patterns”, *Proceeding of IEEE International Conference on Quality Control by Artificial Vision(QCAV)*, pp.113-116, May 2013.

国内シンポジウム

- (1) 飯塚正樹, 秋月秀一, 橋本学, “物体がもつ共通的功能に着目したテーブルトップ物体認識”, 第 22 回画像センシングシンポジウム (SSII2016), IS1-22, pp.1-6, Jun. 2016.
- (2) 田口皓一, 秋月秀一, 橋本学, 飛谷謙介, 長田典子, “物体形状に対する感性指標を評価するための 3 次元特徴量の提案”, 第 22 回画像センシングシンポジウム (SSII2016), IS3-34, pp.1-5, Jun. 2016.
- (3) 鳥居拓耶, 秋月秀一, 橋本学, “局所形状の幾何学的関係に基づく大域特徴量を用いた直方体検出”, 第 22 回画像センシングシンポジウム (SSII2016), IS2-20, pp.1-6, Jun. 2016.
- (4) 飯塚正樹, 武井翔一, 秋月秀一, 橋本学, “日用品認識のためのアフォーダンス特徴量の提案～物体が有する機能に着目したカテゴリレベル物体認識～”, 動的画像処理実利用化ワークショップ (DIA2016), OS3-2, pp.30-31, Mar. 2016. 【研究奨励賞】
- (5) 川口雅浩, 秋月秀一, 佐藤吉将, 渡邊瞭太, 橋本学, 長田典子, “一人称視点映像のみを用いた対象物体上におけるユーザの 3D 注目領域推定”, 動的画像処理実利用化ワークショップ (DIA2016), IS2-A1, pp.116-117, Mar. 2016.
- (6) 田口皓一, 秋月秀一, 渡邊瞭太, 橋本学, 飛谷謙介, 長田典子, “対象物を感性評価するための 3 次元マルチスケール曲率特徴量の提案”, 動的画像処理実利用化ワークショップ (DIA2016), IS1-C4, pp.70-71, Mar. 2016.
- (7) 間瀬基之, 武井翔一, 秋月秀一, 橋本学, “nDCG 尺度とクラス分離度を考慮した特徴点の自動選択に基づく高精度 3D モデル検索”, 動的画像処理実利用化ワークショップ (DIA2016), IS2-B1, pp.138-139, Mar. 2016.
- (8) 武井翔一, 秋月秀一, 橋本学, “3 次元点群の空間的広がりに着目した高速キーポイント検出”, 第 21 回画像センシングシンポジウム (SSII), pp.IS1-18-1-5, Jun. 2015. 【オーディエンス賞】
- (9) 橋祥平, 武井翔一, 永瀬誠信, 秋月秀一, 橋本学, “点群の粗さに依存しない物体認識のための RPD (Relative Point Density) 特徴量の提案”, 動的画像処理実利用化ワークショップ (DIA2015), OS3-1, pp.1-5, Mar. 2015. 【研究奨励賞】
- (10) 朝倉茉理, 武井翔一, 永瀬誠信, 秋月秀一, 橋本学, “手持ち物体の 3 次元モデリングのための時系列距離データ合成～オクルージョン領域の自動検出と抑制～”, 動的画像処理実利用化ワークショップ (DIA2015), IS1-D4, pp.1-6, Mar. 2015.
- (11) 飛谷謙介, 武藤和仁, 李奈栄, 片平建史, 白岩史, 中島加恵, 長田典子, 岸野文郎, 山本倫也, 秋月秀一, 橋本学, 河崎圭吾, 荷方邦夫, 浅野隆, “パーソナルファブリケーションを促進する感性指標化技術～3次元形状の感性評価因子と物理特徴量との関係～”, ビジョン技術の実利用ワークショップ (ViEW), IS2-16, pp.1-6, Dec. 2014.
- (12) 櫻本泰憲, 兼松裕一, 秋月秀一, 橋本学, 渡邊清高, 関真規人, “Coded Multi Flash Imaging による 3 次元凹凸構造特徴マッチング”, ビジョン技術の実利用ワークショップ (ViEW), OS4-H1 (IS2-3), pp.1-7, 2014.
- (13) 武井翔一, 秋月秀一, 橋本学, “識別性能の予測評価に基づく選択的特徴量を用いた物体認識”, 第 19 回画像センシングシンポジウム (SSII), pp.IS2-06-1-6, Jun. 2014.
- (14) 永瀬誠信, 秋月秀一, 橋本学, “法線差分ベクトルに基づく低次元特徴量を用いた高速 3 次元物体認識”, 電気学会研究会 (知覚情報/次世代産業システム合同研究会), PI-14-2 IIS-14-47, pp.7-12,

Mar. 2014.

- (15) 武井翔一, 秋月秀一, 橋本学, “特徴空間における識別性能が高いキーポイントを用いた 3 次元物体認識”, 電気学会研究会 (知覚情報/次世代産業システム合同研究会), PI-14-1 IIS-14-46, pp.1-6, Mar. 2014.
- (16) 永瀬誠信, 秋月秀一, 橋本学, “局所形状の独自性に着目した物体認識に有効な 3-D 特徴点の自動抽出”, 情報処理学会コンピュータビジョンとイメージメディア研究会 (CVIM2013), Vol.2013-CVIM-187, No.26, pp.1-8, May 2013.
- (17) 武井翔一, 秋月秀一, 橋本学, “準リアルシーンをを用いた 3 次元特徴量の生成学習に基づくばら積み部品の認識”, 第 19 回画像センシングシンポジウム (SSII), pp.IS3-29-1-6, Jun. 2013.
- (18) 橋本定佳, 柴田悠太郎, 秋月秀一, 橋本学, “ステレオ画像マッチングにおけるブロックサイズの適応的制御手法”, 電子情報通信学会 2013 年総合大会, ISS-SP-366, p.208, Mar. 2013.
- (19) 武井翔一, 永瀬誠信, 秋月秀一, 橋本学, “準リアルシーンの生成学習とベクトルペア最適化に基づくばら積み部品の認識”, 動的画像処理実利用化ワークショップ (DIA), O4-2, pp.244-249, Mar. 2013. 【研究奨励賞】
- (20) 永瀬誠信, 秋月秀一, 橋本学, “誤対応確率を最小化する 3-D 特徴点を用いた高速ロボットビジョン”, ビジョン技術の実利用ワークショップ (ViEW2012), pp.IS2-B5, Dec. 2012.
- (21) 永瀬誠信, 秋月秀一, 柴田悠太郎, 橋本学, “局所形状の独自性に着目した物体認識のための 3 次元特徴点の自動抽出”, 第 18 回画像センシングシンポジウム (SSII), pp.IS2-17-1-7, Jun. 2012.
- (22) 永瀬誠信, 秋月秀一, 柴田悠太郎, 橋本学, “局所法線分布分析に基づく 3 次元物体認識に有効な高独自性領域抽出手法”, 情報処理学会第 74 回全国大会, pp.5T-4 2-511-512, Mar. 2012.
- (23) 橋本学, 斎藤正孝, 秋月秀一, “参照画素削減による超高速テンプレートマッチングと部品認識への応用”, 精密工学会秋季大会, pp. 926-927, Sep. 2011.
- (24) 秋月秀一, 橋本学, “最適配置された画素群の濃度共起発生確率に着目した距離画像の高速マッチング手法”, 画像の認識・理解シンポジウム (MIRU), IS2-53, pp.860-867, Jul. 2011.
- (25) 秋月秀一, 奥田晴久, 鷲見和彦, 橋本学, “相対濃度の共起発生確率に着目したテクスチャレス画像の高速マッチング”, 動的画像処理実利用化ワークショップ (DIA), O2-1, pp.15-20, Mar. 2011. 【研究奨励賞】

受賞

- (1) MIRU2014 若手プログラム Honorable Mention Wakate Presentation Award, 2014/8/1.
- (2) 第 16 回パターン認識・メディア理解研究会 (PRMU) アルゴリズムコンテスト最優秀賞, 2012/9/4.
- (3) 動的画像処理実利用化ワークショップ (DIA2011) 研究奨励賞, 2011/3/3.
- (4) 第 14 回パターン認識・メディア理解研究会 (PRMU) アルゴリズムコンテスト優秀賞, 2010/9/7.