チュートリアル資料

物体認識の基礎と研究事例(3D 的認識)

Version 2.01

中京大学 工学部 橋本 学

2021年1月5日

目次

1	はし	じめに	3
2	ロオ	ドットビジョン課題と構成	4
2.1	1 主要	£3分野におけるロボットとビジョン課題傾向	4
2.2	2 口才	ミットビジョンの基本課題と構成	6
3	3次	ス元センサの原理と現況	8
3.1	1 3次	、元センサの分類	8
3.2	2 主な	♀パッシブセンサの基本原理	10
	3.2.1	2 眼ステレオ	10
:	3.2.2	マルチベースラインステレオ	12
:	3.2.3	その他のパッシブセンサ	12
3.3	3 主な	♀アクティブセンサの基本原理	13
:	3.3.1	スポット光投影法................................	13
:	3.3.2	スリット法投影法	13
:	3.3.3	空間コード化法	14
:	3.3.4	位相シフト法	16
:	3.3.5	Time of Flight 法	17
	3.3.6	ランダムドット投光併用ステレオ............................	17
	3.3.7	照度差ステレオ	18
:	3.3.8	その他のアクティブセンサ	19
3.4	4 主な	よ市販センサの現況	19
4	物位	\$認識アルゴリズム	22
4.1	1 物位	*認識の分類	22
4.2	2 P E	ピアランスベース物体認識	23
	4.2.1	アピアランスベース物体認識の考え方	23
	4.2.2	パラメトリック固有空間法	23
	4.2.3	その他の手法	25
4.5		デルベース物体認識	25
	4.3.1	キーポイントマッチングの原理	25
	4.3.2	3 次元特徴量の分類	$\frac{-\circ}{28}$
	4.3.3	タイプ A の特徴量(キーポイント周辺のデータをもとに算出)	$\frac{-5}{28}$
	4.3.3	1 概要	$\frac{-5}{28}$
	4.3.3.	2 SHOT 特徴量に基づくキーポイントマッチング	30

4.3.3	.3 PFH 特徴量に基づくキーポイントマッチング.........	31
4.3.4	タイプ B の特徴量(複数のキーポイント間の関係を特徴量化)	32
4.3.4	.1 概要	32
4.3.4	.2 PPF 特徴量に基づくキーポイントマッチング.........	33
4.3.5	局所参照座標系と主要技術............................	33
4.3.5	.1 局所参照座標系の概要と分類	33
4.3.5	.2 Mian \mathcal{O} LRF	36
4.3.5	.3 Mian の LRF の改良	36
4.3.5	.4 BOARD-LRF	36
4.3.6	その他のモデルベース物体認識.............................	37
4.3.6	.1 EGI を用いた物体認識	37
4.3.6	.2 Spin Image	38
4.3.6	.3 DAI を用いた物体認識	38
4.3.7	ICP による高精度位置合わせ............................	39
5 33	R元物体認識に関する王な研究事例 5.4 - 4	40
5.1 対象	象物の位置姿勢認識に関する研究事例	40
5.1.1	可観測性を考慮したベクトルペアマッチング............	40
5.1.2	入力シーン予測に基づく特徴量選択	41
5.1.3	CCDoN 特徴量に基づくキーポイントマッチング	41
5.1.4	SHORT 特徴量に基づくキーポイントマッチング	42
5.2 D:	ギット動作生成のための物体認識に関する研究車例	43
591	小り下動下土成のための物件砲艇に因りる明九事内	
0.2.1	プリミティブ近似によるモデルレス把持位置推定	43
5.2.1 5.2.2	プリミティブ近似によるモデルレス把持位置推定 ピッキングリスク最小化のためのロボット動作生成	43 45
5.2.1 5.2.2 5.2.3	プリミティブ近似によるモデルレス把持位置推定 ピッキングリスク最小化のためのロボット動作生成 日用品の機能認識とその応用	43 45 46
5.2.1 5.2.2 5.2.3 5.2.4	ペット動作主成のための体認識に因うる研究事例	43 45 46 49
$5.2.1 \\ 5.2.2 \\ 5.2.3 \\ 5.2.4 \\ 5.2.5$	 プリミティブ近似によるモデルレス把持位置推定 ピッキングリスク最小化のためのロボット動作生成 日用品の機能認識とその応用 部品の機能部位を利用した点と動作の転移 低能属性ラベル付き全周囲データセット 	43 45 46 49 51
5.2.1 5.2.2 5.2.3 5.2.4 5.2.5 5.3	 ペット動作主成のための体認識に関する研究事例 プリミティブ近似によるモデルレス把持位置推定 ピッキングリスク最小化のためのロボット動作生成 日用品の機能認識とその応用 こ ご ご<td>43 45 46 49 51 52</td>	43 45 46 49 51 52
5.2.1 5.2.2 5.2.3 5.2.4 5.2.5 5.3	 ペット動作主成のための体認識に関する研究事例 プリミティブ近似によるモデルレス把持位置推定 ピッキングリスク最小化のためのロボット動作生成 日用品の機能認識とその応用 こ 部品の機能部位を利用した点と動作の転移 こ ひ他の 3D 認識に関する研究事例 こ Coded Multi Flash Imaging による疑似 3D マッチング 	43 45 46 49 51 52 52

6 おわりに

1 はじめに

物体認識は、工場の生産ラインにおけるロボット視覚や屋内外の移動ロボット視覚など、幅 広い分野で実用化が望まれている.特に近年では、小型で使いやすいセンサやオープンソース の認識ソフトウェアや機械学習用のデータセットなどの研究環境が充実しており、ロボット知 能化のための3次元物体認識技術が着実に身近になってきていることは間違いない.実際、筆 者の研究室に寄せられる生産・物流分野における新ビジネスの創出に関する相談は増加してお り、この分野への期待感は大きい.ただ一方では、センサと認識アルゴリズムのいずれについ ても、現時点でできていることとそうでないことがあるのも事実である.

そこで,この技術の実応用を目指すうえで,現時点で何が足らないのか,今後どのような技術開発が必要なのかについて,筆者自身があらためて考え直すために,このタイミングで3次 元物体認識の基礎的,代表的な技術を整理しなおすことにした.

本稿は,筆者がここ数年の間に,学会や民間企業主催のセミナーやチュートリアル講演の機 会を通じて語らせていただいた内容 [72],および大学院生向けの講義にて解説してきた内容を もとに,重要部分をまとめたものである.ただ執筆時間が十分に確保できないこともあり,3次 元センサと認識アルゴリズム,および研究事例について,比較的汎用性が高そうな部分を抜粋 した内容となっている.そのため,初学者・ベテランいずれの読者にとっても中途半端になっ てしまった.この点については,どうかご寛恕を賜りたい.

なお,生産ラインや物流・配送センターで扱われる対象物は,(どんなに薄いものであって も)基本的には3次元物体であるが,これを認識する手法としては,2次元的アプローチと3 次元的アプローチの2種類がある.本稿では後者に関して解説したが,筆者の経験上は,3次 元的アプローチでの実現が困難な認識タスクについては,いまいちど2次元的アプローチに立 ち返って実用解を追求するほうがよい場合が多いと感じている.この点については現在いくつ かの事例を通じて検討中であり,折をみて別稿にてまとめたいと考えている.

以下,第2章では産業用途のロボットの一般的な課題を整理するとともに,特にロボットビジョンについては重要な構成要素を明確にする.続く第3章では基本的な3次元センサの仕組みと市販品の状況についてまとめる.また第4章ではセンサによってデータ化されたシーンをもとに物体認識を実現するためのアルゴリズムについて概説する.さらに第5章では当研究室における研究事例のなかから代表的なものを抜粋して紹介する.

2 ロボットビジョン課題と構成

2.1 主要3分野におけるロボットとビジョン課題傾向

近年,知能ロボットの普及が進んでいる.最も古くから取り組まれているのは,いうまでも なく工場における生産システムの自動化であるが,生活様式の変化や新型ウィルス感染症の影 響もあり,電子商取引のさらなる活発化を背景にして,物流用途のロボットマーケットも活気 づいている.

図1に示したように、家庭からの通販の注文を受信した配送センターは、工場で生産された 商品を各家庭に送り届ける.将来は生活支援ロボットが普及し、人間に代わってロボットが日 用品を扱う場面も増えると予想される.



図1 ロボットの利用分野

生産,物流,家庭,この3つが,当面の知能ロボットの主たる活躍場所であると考えられ, それぞれにおいてロボットが知能を持つための必須技術として,ロボットビジョンの高度化が 望まれている.

図 2 に示したのは、これら主要 3 分野におけるロボットとビジョンのタスクである.ただし、これはあくまで一例である.

ロボットが行うべき作業のうちピッキング作業を例に取ると、一般には、Pick and Place 作 業と Pick and Operation 作業に大別されると考えられる. 前者は対象物をハンドで取り上げ (Pick),目的の場所で手を離す (Place)という作業であり、これに対して後者では、ロボット がハンドで取り上げた対象物を道具として扱うなど「その後の」作業が行われる.



図2 主要3分野におけるロボットとビジョンの課題傾向

生産現場の中でも、比較的初期工程である部品投入工程では、いわゆるばら積みピッキング という作業が代表例であるが、現状ではほとんどの場合、Pick and Place タスクとなる. この タスクではロボットにとっては Pick が重要であり、ばら積みされている対象物を如何にして 的確に Pick するかが技術的な目標になる.物流工程としての配送センターにおけるタスクも 同様であり、Pick and Place が主流である.

一方,家庭における生活支援ロボットについては,Pick したものが日用品や道具であること が多いことから,Pick した対象物をなんらかの作業に用いる,Pick and Operation タスクが 主になると考えられる.この場合は,対象物を単にPick するだけでは不足であり,対象物のど こをどのように把持できたかをロボット自身が把握していないことには後続の Operation 作業 を円滑におこなうことはできない.すなわち,Pick 作業に課せられた難易度は,2種のタスク で自ずと異なってくる.

次に、ロボットの動作パターン種類数について考えると、Pick and Place タスクの場合は、 毎回の Pick 作業をおこなうための対象物の位置は少しずつ揺らぐとしても、基本的には Pick したあとはハンドを離すだけでよいから、動作種類数も限られ、しかも事前に生成しておくこ とも容易である.これに対して、生活支援の場では、動作パターンは状況や目的によって多種 多様に変化する可能性があり、その場での臨機応変な動作生成が重要な技術項目となってくる.

また,物体モデルの利用可否という観点では,何を製造しているのかをシステムが把握でき ている生産現場においては,ロボット操作対象の物品に関する形状データや外観データを保有 していることが期待できることから,物体モデルの利用は可能であると考えられる.一般家庭 では,現時点では多様な対象物のモデルを利用できる可能性は低い.品種数も同様であり,工 場(生産)では複数ではあるが基本的には品種数は限定され,既知である場合も多い.一方, 家庭の場面では品種は未知であることが多い.

以上,主要3分野におけるロボットとビジョンがおかれた状況とその課題傾向を概観した.

ここでもういちど,3つの分野に関連して画像認識の課題は何かについて整理する.生産工場 や物流工程では対象物の名称は事前に把握されていることが多く,特に生産ラインでは対象物 の形状モデルを利用することも難しくない.ゆえにロボットがビジョンによってそのつど得た い情報は,位置や姿勢に関するものである.我々はこれを特定物体認識と呼んでいる.

一方,家庭および物流の一部工程では,把持対象に関する形状モデルなどを保持しておくこ とは難しく,そもそも,目の前にある対象物は何か?を知る必要もある.我々はこれを一般物 体認識と呼び,特定物体認識とは区別している.なお,物流工程の代表例である配送センター を想定すると,例えば大手通販企業が持つ巨大な配送センターの場合は,1日あたり5000~1 万アイテムが新入荷するとも言われており,そのつどアイテムに関する個々のモデルを生成し たり入手することは現実的には難しいことから,ビジョンとしては,これらのモデルを利用で きないとの前提をおくべきである.

このように、生産、物流、家庭という3つの現場ではロボットに課された主たるタスクや、 利用可能な情報などの環境条件が異なることから、それぞれの分野に適したロボットビジョン を開発必要がある.ただし、たとえば生産ラインでは将来的には加工や組立工程を含めた全自 動化を目指していることを考えても、ロボットに対しては、人間のようにドライバやスパナな どの工具を自由に扱ったり、ケーブル先端のコネクタを嵌合するなど、Pick and Operation タ スクへの期待は高まる一方である.

2.2 ロボットビジョンの基本課題と構成

前節で述べたように、物体認識(ロボットビジョン)の課題はシチュエーションによって異 なるが、対象物を掴むという基本タスクに対するビジョンの課題をおおざっぱに述べるならば、 基本的には図3に示す3つが重要である.すなわち、どこに、なにがあるか、またそれをロ ボットハンドでどう掴むかという課題である.



図3 ロボットビジョンの基本課題

物体認識の代表的なアプリケーションは、図4に示すようなビンピッキング、すなわち、ば ら積みされた対象物体それぞれの位置と姿勢を認識し、ロボットハンドによって把持(ピッキ ング)するシステムである.ロボットの前に乱雑に積まれた対象物体群に関して、3次元セン サから取得されたレンジデータ(深度データ)は、ポイントクラウド、点群データなどと呼ば れ、これをもとにして個々の物体の位置と姿勢パラメータ(回転行列 R および並進ベクトル t)が計算される.このデータがロボットに送られ、対象物が把持される.図からわかるように、 物体認識には大きく2つの課題がある.一つ目は,3次元的な対象物を3Dセンサによってで きるかぎり正確にデータ化するための3次元計測に関する課題,二つ目はデータをもとにして 対象物の位置や姿勢,種類などを正しく認識するための物体認識アルゴリズムとしての課題で ある.



図4 ロボットビジョンの基本構成

以下の章では、3D センサ、物体認識アルゴリズムそれぞれについて、主要技術を紹介する.

3 3次元センサの原理と現況

3次元の対象物を、ダイレクトに3次元的に計測するための3次元センサの研究の歴史は長 く、光学式を中心としてじつにさまざまな方式が生み出されている.次章では、3次元センサ を前提とした物体認識手法だけでなく、一般的な RGB カメラを前提とした手法についても述 べるが、近年のトレンドはあきらかに前者に偏っているようにみえる.3次元センサが必ずし も万能でないことについては後述するが、少なくとも研究用途においては手軽に利用できるセ ンサの品種も増え、3次元センサは既に特殊なデバイスではなくなっているのは事実である. すなわち、多くの研究者や技術者にとっては、3次元センサはもはや研究、開発するものでは なく、購入して利用するものになっている.このことから、本章では3次元センサを正しく選 択するための情報を提供する目的で、3次元センサの原理と現況について概観する.

3.1 3次元センサの分類

図5に, 普通の RGB カメラと, 3次元センサの実例を示す. 木製の積み木を並べて同一視 点から撮影した画像をみるとわかるとおり, RGB カメラでは識別しにくい遠近情報や物体の 境界情報が, 3次元センサではわかりやすくなっている. なお, ここで 3次元センサといえど も, 実際にはセンサから見える範囲の奥行き分布のみが計測されていることについてあらため て意識しておく必要がある. 3次元センサを深度センサ, デプスセンサ (D センサ) などとも 呼称する所以である.



図5 RGB カメラと3次元センサの例

光学式の3次元センサの簡単な分類を、図6に示す.この表では、三角測量の原理に基づく ものと、それ以外の代表例として同軸測量形式のものに2分し、さらにパッシブ、アクティブ という観点から分類している.

三角測量と同軸測量の違いについて、おおざっぱに図示したのが図7である.三角測量は、



いうまでもなく三角形という形状が一定の条件下でただ一つに決まるという性質に基づいた測定原理となっている. この図に示した三角形 PQR については,Qが測定したい遠く離れた点であり,PとRがセンサ側のいわば計測者に近い点である. PとR はたとえばそれぞれがカメラであり,遠方の点Qを観測している. このとき,もし仮にカメラ Pとカメラ Qの間の間隔 L が既知であり,なおかつ角度 α と β がわかったならば,三角形 PQR は直ちに同定され,辺 d の長さや点 Q の座標を計算によって求めることができる. 実際には L はカメラキャリブレーション作業によって求められることが多く, α と β は画像処理によって点Q を計測することによって求められる.

同図右側の同軸測量については,センサ側の基点は点 P のみである. 点 P と計測点 Q を結 ぶ線は1本のみであることから同軸測量と呼ばれる. この図から容易に推察されるように,同 軸測量の場合は P から Q が見えれば基本的には距離を求めることができるが,三角測量の場 合は,点 Q の位置によっては P からは見えるが R からは見えない可能性がある. このことは 三角測量方式の原理的な問題点でもある.

次に,パッシブセンサとアクティブセンサの違いを模式的に表現したのが図8である.パッ シブ方式とは,文字通り(光)エネルギーをセンサが受け取ることによって計測する形式であ り,たとえば2眼ステレオ視覚はその代表例である.パッシブといっても当然ながらエネル ギー源は必要であり,たとえば窓から入射する太陽光や,室内の天井照明装置がその役目を果 たしている.

一方,アクティブ方式は同図右のように,センサそのものがエネルギー発生源としての光源 を持ち,光パターンを対象物に照射する.むろん照射されたエネルギーがカメラ等のセンサ上



図8 パッシブセンサとアクティブセンサ

の素子に帰還することが前提であり、それによって距離計測が実現する.この図では違いを明 確にするためにパッシブ、アクティブともに三角測量方式として描いた.ロボットのための距 離センサとしては、近接距離を測るための超音波センサなど光以外のエネルギーを利用するも のもあるが、本稿では実用性の高い光利用センサに限定して説明を続ける.

3.2 主なパッシブセンサの基本原理

3.2.1 2 **眼ステレオ**

パッシブセンサの例として,最もポピュラーな2眼ステレオ視覚について説明する.これは 図9に示したように三角測量の原理そのものでもある.センサ側のパラメータとしての基線長 (baseline 長) *L* と,画像処理によって求めた左右カメラから対象点 P を見たときの視線ベク トル角度 α と β をもとに、P の座標が計算される.



図9 2眼ステレオ

ステレオ視の原理は単純であるが,カメラを購入して自身で実装しようとすると,さまざま なノウハウも必要である.まず点 P を左右カメラで対応づける必要が生じるが,通常は図 10 に示したエピポーラ拘束を利用する.いま左カメラ L によって,求めたい遠方のある点が,画 像上に点 d として計測されたと仮定する.この事実だけは遠方点を一つに同定することはでき ないが,図からわかるように,左カメラの光学中心 L と撮像素子上の点 d を結んだ直線上の どこかに遠方点があるはずである.それはたとえば,P かもしれないし Q か R かもしれない. つまり,左カメラで得られた情報から対象の 3 次元点は空間上のある直線上に拘束されたこと になる.

次に、この直線を右カメラで撮影することを想像すると、この光学系が線形写像を構成する ならば、図のように画像上でも直線として存在するはずである.これをエピポーラ線という. このとき、もし左カメラと右カメラの幾何学的関係が既知ならば、dが計測された時点でエピ ポーラ線 E を計算することも可能である.すなわち、左カメラにおいて d として撮影された 3 次元点を右カメラ画像上で探索する場合、我々は画像全体を探す必要はなく、エピポーラ線上 のみを探索すれば足りる.このことは、処理時間の高速化に寄与するのみならず、対応点探索 の信頼性を向上させるという利点がある.



図 10 エピポーラ拘束を利用したステレオ対応付け

ここで、もし2つのカメラの光軸が平行であれば(平行カメラ)、同図右のようにエピポーラ 線は画像のi軸に平行になる.しかし、ステレオ視ではそもそも計測対象物が左右両方のカメ ラの視野内に存在し、同時に撮影されていることが前提であるから、共有視野を広げるという 観点からは2つの光軸を互いに内側を向ける非平行ステレオを構成する場合が多く、エピポー ラ線は撮像素子の水平軸とは平行にならない.このことは、たとえばステレオ対応点探索を電 子回路等に組み込もうとする場合には、エピポーラ線に相当する直線状のアドレス発生が若干 面倒になることを示唆している.

画像処理による対応点探索では、1 画素単位の対応付けは画素がもつアイデンティティが低 いことに起因して難しく、実際には、たとえば 3x3 や 5x5 などのブロック領域を単位として 左右画像を対応づけることが多い.すなわち、左の画像を適当な大きさでブロック分割し、ブ ロック単位で右画像における対応箇所を検出するブロックマッチングが用いられる.このとき ブロックのサイズが重要である.もし一つあたりのブロックサイズが大きい場合には、たとえ ば対象物の辺縁などの物理的な境界部分を撮影した場合には、対象物(前景)と背景という距 離が異なる部分を同時に同じブロックに含めることを意味し、左右カメラの間の見えの違いが 対応探索の妨げとなりかねない.一方ブロックサイズが小さい場合にはブロック内の画素情報 が少ないことに起因して、画像小領域の情報にアイデンティティが少なく、対応探索時に正し く照合されにくいという懸念が生じることになるので、注意が必要である.

3.2.2 マルチベースラインステレオ

ステレオ視においては、2 つのカメラ間の間隔(ベースライン長)が距離計測精度に大きな 影響を与える.これが長い方が精度が高いが、その分、左右カメラによる対象点の見えの違い も大きくなるため、画像処理による対応点探索は困難になる.この問題を改善するために提案 された手法が図 11 に示すマルチベースラインステレオである.これは、ベースライン長が異 なる複数のカメラを準備することによって、対応点探索の確実性と計測精度の両立を狙った方 法である.

なおステレオ視には、単眼カメラを移動させることによって連続的に対象物を撮影し、画像 間での対応付けをおこなう移動ステレオ(モーションステレオ)という考え方があり、カメラ 内部パラメータのキャリブレーションが一つで済むなどの利点もあるが、前述の可変ベースラ イン長の考え方を利用できるという利点もある.

3.2.3 その他のパッシブセンサ

現時点では詳細な説明を割愛するが、その他のパッシブ方式センサとして、ロボット用セン サとしての顕著な適用事例は見当たらないが、対象物までの距離によって発生する焦点ぼけを 利用する方式や、人体計測に用いられる多カメラ利用による視体積交差法も、それぞれ同軸測 量、三角測量のパッシブ方式と言える.



図11 マルチベースラインステレオ

3.3 主なアクティブセンサの基本原理

本節ではアクティブセンサについて説明する.パッシブ方式と異なり,センサ自らがさまざ まな光パターンを投射し,これを積極的に利用する.空間的になんらかの模様を形成する構造 化光 (structured light) や,光強度に時間的な変化を与えたり波長を変更するなど,さまざま なアイデアがある.本稿では主要な技術に絞り,投光パターンの単純さをもとに準備技術を整 理していく.

3.3.1 スポット光投影法

図 12 のようにスポット光線を投射する方式であり、ステレオ視覚の 2 つのカメラのうち片 方を投光装置に変更したものと考えると理解しやすい.対象物に投射した光の軌跡を知ってお けば、三角測量の原理そのものが適用できる.図のような構成の場合は 1 回の計測(1 枚の画 像撮影)で 1 点の計測しかできないが、光源にレーザを利用する場合には遠方まで光を到達さ せることが期待できるため、土木分野での岩盤計測などに利用されることもある.

スポット光照射装置を機械的に走査し,その光線の方向を常に把握できるならば,時間はか かるものの広範囲の距離分布を得ることもできる.レーザレーダとして前方を走行する車両の 計測や,巨大建造物の継続にも適用できる.ただし走査には一定の時間がかかるために,計測 対象物が移動している場合には注意が必要となる.

3.3.2 スリット法投影法

1回の撮影で1点のみの距離計測では時間がかかるという問題を解決したい場合には,図13 のようなスリット光を投影する方式も選択肢になる.この場合は,縦スリット光を投射し,投 光装置と水平方向に離れたカメラにて画像を撮影することによって,1回の撮影で1ライン分 の距離計測が可能になる.スリット光線が対象物を平面で切るようなイメージから,光切断法 とも呼ばれる.この方式で全画面を計測しようとする場合には,スリット光を平板ミラーや多



図 12 スポット光投影法

角柱状のミラーで左右に走査すればよく,三角測量の原理が成立する.光切断法の計測装置の 中には,光源の走査ではなく,対象物を機械的に移動させる方式もある.

スリット光線の生成方法としては、レーザスポット光を円筒形のレンズ(シリンドリカルレ ンズ)を通してスリット光に変換する方法がよく用いられる.手軽であるが、光エネルギー密 度が下がることから、スポット光投影法と比較すると遠方の計測が難しい、あるいは外乱光の 影響を受けやすいというデメリットがある.



図 13 スリット光投影法

3.3.3 空間コード化法

前項のスリット光の発展として,複数のスリット光線を同時に投光し,つまり縦縞を投光す る方式が考えられる.ただし,この方法を三角測量方式として成立させるためには,画像とし て撮影した縦縞についてどのスリットが何番目のものであるかという縞同定の問題を解決す る必要がある.このことから,複数の縞パターンになんらかのマークをつけておき,撮影した 画像からそのマークを検出して縞番号を同定するという研究例がある.たとえば縦スリットに ランダムな途切れを含めておく方法やスリット毎に色を変える方法などがある.文献は省略 する.

このような方法の中でとりわけ実用性が高く,広く普及した手法が空間コード化法である. 原理を図 14 に示す.元になる原理は米国で開発されたが,わが国で実用レベルの装置が開発



図 14 空間コード化法(3ビットの例)

図の例では,縦8本の短冊状の白黒2値の光パターンをあるルールに基づいて3回,時間的 に切り替えて投光し,それぞれのパターンが照射された画像を合計3枚取得する.これを解析 することによって,各画素に照射されているのが何番目の縞であるかを同定するという仕組み である.ルールとしては,図のように0~7までの2進符号を用いる.たとえば1回目の撮影 では黒,2回目は白,3回目は再び黒として撮影された画素があるならば,黒→白→黒と推移し たことになるから,その画素には左から3番目のスリットが照射されたと判定でき,あとは三 角測量の原理に基づいて距離を計算すればよいことになる.

すなわちこの手法では,図のように空間を時間的に白または黒の光パターンでコード化し, それを撮影した画像群からパターンをデコードする点がアイデアとなっている.

実際の装置としては、8枚の縞パターンによる256本の縦スリットが用いられることが多い. また2進符号としても図のような単純な2進符号ではなく交番2進符号(グレイコード)を用 いることによってパターン間のハミング距離を1に均一化している.さらに、対象物自体の明 るさの影響を受けない画像処理を行うため、通常の8パターンの投光ではなく白黒反転させた 相補パターンをペアで投射し、ペア画像間の差分をとることによってコントラストを2倍に拡 大する(ただし投光パターン数が2倍になるので処理時間も2倍かかる)など、さまざまな工 夫がなされている.

なお投光装置としては,液晶や DMD による PC プロジェクタが用いられたり,前項で説明 した縦スリット化されたレーザ光線を回転する多面体ミラーで走査しながら光源を ON/OFF 制御することによってコード化光線を作り出す手法などが開発された.また,コントラスト改 善の目的では,相補パターン投光ではなく白黒の中間調のパターンを1枚追加投光し,これと 8枚のパターンとの差分をとる方法も提案されている.

このように,空間コード化法は原理がシンプルで効果が高い手法であることから,いまなお 多くのメーカによってセンサとして実用化がなされている.筆者らのこれまでの経験からも, 高精度が実現できるセンサとして実用的との印象がある.ただしこの方法も三角測量方式であ るため,照射したパターン光がカメラで撮影できない幾何関係は存在する,距離計測精度は基 本的にはプロジェクタとカメラ間のベースライン長に依存するためにセンササイズの小型化に は限界がある,さらに複数の画像撮影に基づくことから移動体の計測には不向きであるという 原理的な問題点は残っている.

3.3.4 位相シフト法

位相シフト法も、空間コード化法と同じように、画像上の各画素がプロジェクタから投光さ れたどの位置に相当するかを知ることに基づいて距離計算をおこなう方法である.図15に位 相シフト法の模式図を示す.対象物に対して、水平方向に正弦波的に明暗が変化するパターン を投光し、画像を取得する.このとき投光する正弦波パターンは1種類ではなく、位相をずら しながら3枚(または4枚)を投光し、投光するたびに画像を取得し、連続画像を得る.3枚 の画像の同一画素における濃淡値をもとに、投光した正弦波の位相を知ることができるので、 結果としてプロジェクタのどの画素が撮影されたのかがわかり、三角測量の原理が成立する.

1枚の画像のみの場合は、ある画素の明るさがわかったとしても正弦波のどの位置に相当す るかはわからないが、3枚の画像から得られた明るさ情報をもとの正弦波に当てはめることに よって、位相ずれを検出できる.位相がずれている複数の正弦波パターンを投光することは一 般的には簡単ではないが、近年、PC プロジェクタ等の任意パターン投光システムが発展した ことから、濃淡値が一定の分布を持つパターンであっても、比較的容易に投光できるように なった.

なお,正弦波パターンは周期パターンなので,まるまる n 周期ずれたパターンを計測しても 同じ明るさが得られることから,区別ができなくなるという問題がある.ただし,投光する正 弦波の周期を調整し,得られた距離分布には一定の滑らかさが存在するとの仮定を利用するこ とによって,現実的な距離分布を得ることができる.



図 15 位相シフト法

3.3.5 Time of Flight 法

Time of Flight 法(飛行時間法)の原理は単純であり,基本的には光を照射して帰還するま での時間差を計測し,光の進行速度を用いて距離に換算する.ただ,光の速度は非常に速いた め,超音波センサのようなパルス光を照射したとしてもダイレクトに時間差を測るのは難しい ことから,正弦波状に変調した光パターンを照射し,帰還光との位相差を計測することによっ て時間差を割り出すことが多い.図 16 は原理を説明するための模式図である.



図 16 Time of Flight 法

TOF センサは,帰還光の明るさを計測することによって正弦波の形を同定し,位相ずれを計 算しているため,時間のずれを明るさの変化として捉えているともいえる.明るさ計測を時間 ずれに敏感にすることによって距離推定精度が高まるが,そのためには,たとえば正弦波の周 期が短くすることが考えられる.ただしこの場合は,帰還光が1周期以上ずれたかどうかの区 別ができないことに起因して計測可能範囲が浅くなるという問題が生じる.

TOF センサは,一時期は三角測量方式と比較して高価であることや,高輝度 LED の発熱や それにともなう受光センサの感度補正など,実用面ではやや敬遠されていた傾向がみられたが, 最近では低価格化もすすみ,同軸測距であることやほぼリアルタイムの計測が可能という利点 が見直され,徐々にユーザが増えている.

3.3.6 ランダムドット投光併用ステレオ

この方式は基本的には2眼ステレオであるが,図17のようにパターン投光を併用すること によって,計測対象物にテクスチャがなくても安定した距離計測ができる,またパターン投光 装置をステレオビジョンのオプションとして追加的に利用できることなどから,選択肢の一つ として位置づけられている.センサシステムとしては,ハードウエアとしてのランダムドット 投光装置が簡便に実現できるかどうかが重要であろう.



図 17 ランダムドットステレオ

3.3.7 照度差ステレオ

照度差ステレオの基本的な考え方は,画像の明るさは照明光と対象物の形状によって決まる というモデルを利用することである.図18のように,同じ対象物に対して3方向から照明を投 射して同一視点から3枚の画像を取得すれば,各画素ごとに照明方向と明るさの関係を導くこ とができる.これを連立させて解くことによって,共通項である対象物の表明形状を推定する.

対象物の表面材質にスペキュラーな要素が大きい場合には,明るさ情報が画像に反映されに くいので誤差を生む要因になる.また,この方式では対象物の表面座標が得られるのではなく, あくまで表面の法線ベクトル方向が得られることになるので,これを形状として復元するため には表面が滑らかであるなどの一定の制約が必要になる.



図 18 照度差ステレオ法

3.3.8 その他のアクティブセンサ

縞パターンを投光して受光時に別の縞と干渉させるモアレ(Moire)法は,人体計測などの分 野で長い歴史をもつ手法である.ロボット用センサとしてはあまり利用例をみないが,計測装 置としては小さな凹凸も計測できるなど現在でも利用されている.本稿では詳細説明は割愛さ せていただく.

3.4 主な市販センサの現況

本節では,前節で説明したいくつかの原理に基づく市販の3次元センサについて,現況を概 説する.図19は,市販されているセンサについて,当研究室にて使用実績があるものを中心 に,現時点では入手できない旧製品を含めてまとめたものである.アクティブ方式・三角測量 方式にカテゴライズされる製品が多いことがわかる.



図 19 市販 3 次元センサの分類

また、図 20 はカタログ値、公称値としての代表的センサのスペックである.

実際の性能についてはカタログ値では不明であることが多いことから,7種のセンサについ ては,当研究室内で実測実験をおこなった.その結果を,図21および図22に示す.撮影対象 物は,図のように,①金属光沢を有するステンレス製コップ,②拡散表面的な冷間圧延鋼板素 材の機械部品,③白色(つやけし)石膏像,④黒色(つやけし)石膏像,⑤透明のガラス瓶, ⑥小型プラスチック製の玩具用ギア,⑦M8 ボルトの7種類である.撮影条件を一定にして, KinectV2, ENSENSO, R200, ASTRA, SR300, D415, D435の各センサにて撮影した.実 際には動画として記録しているが,予想通り,万能のセンサはなく,どの製品も一長一短であ る.たとえば, ENSENSO センサは床面(平面)計測精度が高いが黒色石膏像は情報欠落が多

m	Kinect V2	ENSENSO	R200	Astra	SR300	D415	D435
		000		0, 00			
メーカ名	Microsoft	iDS	Intel	Orbbec	Intel	Intel	Intel
型番	GT3-00005	N35-804-16-IR	82634DSB2P	Orbbec Astra	82535IVCQSPL04N	82635ASRCDVKHV	82635AWGDVKPRQ
販売時期	2014	2015	2015	2016	2016	2018	2018
概略価格	2万1 578 円	約80万円	約1万 9000 円 (開発ツール セット料金)	約1万7000円	約1万8000円	約2万2677円	約2万4400円
デプス解像度	512x424	1280x1024	628x468	640x480	640x480	1280x720	1280x720
撮影範囲 (角度・距離)	水平70°× 垂直60° 0.5~4.5m	0.3~3.0m	0.5~3.5m	水平60°x 垂直49.5° 0.6~8.0m (最適0.6~5.0m)	0.2m∼1.5m	水平69.4°x 垂直425.°x 斜め77° 0.3m~10m	水平91.2°x 垂直65.5°x 斜め100.6° 0.2m~10m
計測形式	Time of Flight	Stereo + Light Coding	Stereo + Light Coding	ランダムドット パターン (Light Coding)	ランダムドット パターン (Light Coding)	Stereo + Light Coding	Stereo + Light Coding
インター フェイス	USB 3.0	Ethernet	USB 3.0	USB 2.0	USB 3.0	USB 3.0	USB 3.0
電源	0	0	×	×	×	×	×

図 20 主な市販 3 次元センサの性能比較

い. 一方 D415 は平面計測結果がかなり荒れているが, 黒色石膏像も写っている.

全体的な印象としては,計測精度としては ENSENSO の性能が高いように見受けられた. また,価格面も考慮した上でのバランスのよいセンサとしては,SR300(または SR305)が候 補になり得るとの印象を持った.

ただし、本質的にデータ化しにくい対象物が以下のように4種類あると思われる. 1~3 は材 質起因、4 は大きさに起因したものである.光学センサである以上、原理的に計測が難しいも のばかりであるが、これらについては、今後のセンサ開発に期待したい.

- 1. 鏡面反射的な金属物体
- 2. 黒色物体
- 3. 透明物体
- 4. 小型部品

これらのセンサから得られるデータは,図 23 のような点群(ポイントクラウドデータ)である. 最近では ROS 経由でセンサ出力を手軽に受け取れるようになったのは利点であるが,上述のようにデータ化しにくい対象物が存在することは事実である.





(b) 対象物





図 22 3 次元センサの比較結果



図 23 ポイントクラウドデータの例

4 物体認識アルゴリズム

本章では,図4で説明した後半部分として,3次元データ化された入力シーンを元にした特 定物体認識のアルゴリズムについて解説する.

4.1 物体認識の分類

2 にて説明したように、ロボットビジョンとしての 3 次元物体認識は 2 種類にわけられる. センサから入力されるデータは通常 3 次元または 2 次元であるが、本稿では 3 次元データとし て入力されることを前提とする.このとき、対象物の種類名(クラス名)は既知であり、位置 と姿勢を認識する技術を特定物体認識という.また種類名そのものを同定する技術を一般物体 認識と呼んでいる.通常、ロボットの知能化のためには、これら 2 つの物体認識が必要となる. とりわけ生産ラインや配送センターにおいては特定物体認識が用いられることが多く、家庭用 ロボットについては一般物体認識の必要性が高い.

どちらの認識技術もまだ未完成であるが,着実に実用域に達しつつある.一般物体認識については,深層学習ベースの手法の性能が向上しており,適切な機械学習のためにはデータの確保が簡単ではない場合があったり,工場の現場では GPU 付きマシンを使いにくいなどの課題は残っているものの,学習さえさせれば,よほどクラス数が多くないかぎり,未学習の対象物であってもほぼ正しいクラス名を出力しうるシステムは存在している.また特定物体認識についても,前節で述べたように材質や大きさの点で現在の3次元センサではデータしにくい対象物があることは事実であるが,データ化できれば,ばら積みピッキングのようなタスクを実行できる環境は整いつつある.

以上のことから、本稿では3次元データをもとにした特定物体認識の諸技術に関して解説す る.まず、特定物体認識は、図24のように、アピアランスベース認識と、モデルベース認識の 2つに分けられる.

アピアランスベース認識は、対象物に関する見え(Appearance)の情報をもとにした手法で ある.一方、モデルベース認識は、3次元物体モデルの利用を前提とし、計算機内に格納され たモデルと3次元の外界を写し取った入力データ(の一部)との整合性を評価することによっ て対象物体を認識する.

ロボットビジョンの黎明期では,センサとしてはモノクロカメラまたはカラーカメラが用い られていたため,3次元対象物を2次元に投影した2次元画像をベースにしたアピアランス ベース認識が主流であった.その後,徐々に3次元センサが使用されるようになり,2010年代 初頭に,Microsoft 社が Kinect センサを発売したことが一つの契機となり,モデルベース物体 認識への関心が一挙に高まった.

ただし,第3章でも述べたように,現状の3次元センサはまだ完全ではなく,正しくデータ 化できない対象物が存在することから,これらに対しては2次元認識的なアプローチが有効で あることが少なくない.また,アピアランスベース認識は,撮影デバイスとしてカメラ一つで



図 24 アピアランスベース認識とモデルベース認識

済むが,近年のカメラの小型化や高性能化の利点を鑑みると,やはり,まだなおアピアランス ベース認識の利点は継続しているとみるべきである.たとえば,ベルトコンベヤ上の対象物を すばやくピッキングするなどの用途については高速処理の観点から,また対象物を多方向から 撮影する場合に有効なハンドアイへの適用容易さの観点からは,単眼カメラによるアピアラン スベース認識が有利である場面が多いであろう.

4.2 アピアランスベース物体認識

4.2.1 アピアランスベース物体認識の考え方

アピアランスベース物体認識の代表的な考え方を,図 25 に沿って説明する.これは,モデル 物体をさまざまな方向から撮影した多数の画像群を,撮影時の視線情報とともにデータベース 化しておき,未知の入力データがその画像群のどれと最も近いかを判定することによって対象 物の位置姿勢を推定するものである.

画像群の撮影方法としては、図のようにモデル物体をターンテーブルのような装置に載せ、 鉛直軸周りの回転 θ と、仰角 φ をパラメータとする視点からモデルを撮影する. これにより、 辞書データベースを作成する.

当然ながら、大量の画像データの保管と効率的な検索が実用上のボトルネックとなるため、 見えを構造化する APO[12] や、画像群を固有空間においてパラメトリックに表現された多様 体(Manifold)として扱うパラメトリック固有空間法 [13, 14, 15, 16] などが開発された.

4.2.2 パラメトリック固有空間法

図 26 に、アピアランスベース認識の代表例であるパラメトリック固有空間法の概念を示す.



図 25 アピアランスベース物体認識の考え方



図 26 パラメトリック固有空間法

1 枚の画像は、画素数分の要素からなる多次元空間上の点に相当するが、多視点画像群にお いては、近接する視点から撮影した画像群は、この多次元空間上でも近接した位置にプロッ トされるはずである.この考え方を延長すると、ターンテーブル上の対象物を鉛直軸周りに 回転させながら撮影した画像群は、多次元空間上で閉じた軌跡(trajectory)を描くと考え られる.これを Manifold と呼ぶ.このとき、この画像群は、主成分分析(PCA; Principal Components Analysis)を使って、適切な次元圧縮をおこなったとしても、やはり多次元空間 において Manifold を形成すると考えられる.この Manifold を記録すれば、データ圧縮の効 果が期待できる.また、入力画像との照合においても、入力画像を PCA により同じ次元に低 次元化した後に Manifold との距離を測り、最短距離になる Manifold 上の点を決定すれば、そ れに相当する辞書データベース上の画像に付与されているカメラ視線情報を引き出すことがで き、結果として、対象物の姿勢を認識したことになる.

このパラメトリック固有空間法の考え方は非常に示唆的であり、仰角を考慮しても同様のス

キームでデータ圧縮や照合が可能になる.ただし,辞書データベース作成時の対象物画像は, 通常,背景が単色かつ無地であることが多いのに対して,入力画像には対象物以外の周辺物体 が映り込んでいることが多いため,これらの差が PCA 空間上での最短距離計算の際の誤差に なる点には注意が必要である [17].

4.2.3 その他の手法

アピアランスベース認識の著名な方法としては、シーン理解、物体認識の古典的な手法とし ての線画解釈や 3DPO[20], APO (Assembly Plan for Observation) [12], ステレオ視によっ て得られた 3 次元線分を利用する VVV[21, 22] などがあるが、現時点では、説明を割愛させて いただく.

4.3 モデルベース物体認識

モデルベース物体認識は,図 27 に示したように、3 次元モデル(通常は CAD データから得 られる点群データや、三角パッチ集合体)と入力シーンとの照合に基づき、3 次元モデルをど のように幾何移動させれば入力シーンとの整合性が高くなるかを探索する方式である.2 次元 画像照合方式としてよく利用されるテンプレートマッチングやキーポイントマッチングとのア ナロジーを考えると理解しやすいであろう.



図 27 モデルベース物体認識

キーポイントマッチングは,代表的なモデルベース手法の一つであり,さまざまな特徴量の 開発と共に発展してきた.詳細は 28 以降に解説する.また本稿では扱わないが,ハードウエ アに立脚したアルゴリズム提案についてもいくつかの研究例があり [23, 24, 25],機会を改めて 紹介したい.

4.3.1 キーポイントマッチングの原理

本節では、3次元点群を前提としたキーポイントマッチングの原理、および3次元特徴量の 分類について解説する.例題として、図28のように、コンピュータに格納された3次元モデ ルと、3次元センサから得られた入力シーンとのマッチング問題を考える. まず最初に、3次元モデル上にいくつかのキーポイントを設定する.ここでは3次元モデル は点群として表現されていることを前提とする.実際のCADモデルの場合は、ソリッドモデ ル、サーフェスモデルを利用することが多いが、この場合も、マッチングに先だって、点群デー タの形に変換しておくことが必要となる.点群の密度はマッチング性能に影響を与えるので、 センサから得られた入力シーンにおける点群密度をもとに、これとほぼ同等の密度にてモデル 側の点群化をおこなっておくことが好ましい.またキーポイントの設定については、一般には、 表面形状としての不連続点箇所として、凸部分に設定されることが多いが、形状的不連続が少 ない対象物の場合は、ランダムに設定されることもある.

次に,設定したモデルのキーポイントについて,特徴量と呼ばれる計量を付与する.通常は, そのキーポイントを特徴付けるデータが使われる.たとえば,キーポイントやその周辺の法線 ベクトルの分布などである.特徴量は通常は複数の数値からなることから,特徴ベクトルと呼 ばれることもある.

以上のような3次元モデルに対しておこなった一連の処理を,センサ入力データに対しても 同様におこなう.すなわち,キーポイントを設定し,特徴量を求めてキーポイントごとに付与 する.

すなわち,この時点で、3 次元モデルと入力シーン双方について、キーポイントの設定と特 徴量の付与が完了したことになる.ここまでの処理で、前処理が終了したので、次にこれらの データを照合する.基本的には、モデル上の N 点と、入力シーン上の M 点のキーポイントを すべて総当たりで対応させる.図からは、モデル上の点 a, d が、それぞれ入力シーン上の点 4、13 と対応づけられることが示されている.このとき、各キーポイントの 3 次元座標は用い られず、特徴ベクトルのみを手がかりにした照合がおこなわる.このような点間の対応付けを おこなった後に、LMed や M 推定などによる外れ値処理がおこなわれ、モデルと入力シーンと の間における回転行列 R と並進ベクトル t を求める.

以上のように,キーポイントマッチングでは,点群データ全体ではなく,選ばれたキーポイントにおける特徴量同士の比較によって点間の対応付けがなされることから,特徴量としてどのようなものを設計するか,あるいは選択するかは,物体認識の性能を大きく左右することになる.

図 29 に示したのは,特徴量やキーポイントが,認識成功率に大きな影響をあたえるという実 験結果である.同図 (a) は,特徴量の個数 (次元数) と認識成功率の関係を示している.次元数 が増えると認識成功率も高まるが,認識に必要な時間もまた増加することになる.図 (b) はモ デル上に設定されたキーポイントの数と認識成功率との関係であるが,やはり予想通り,キー ポイントが多い方が認識性能は高くなる.ただしこの実験でも,キーポイント数が多いほど処 理時間はかかることになるから,次元数やキーポイント数は認識性能と処理時間の trade-off を 考慮して,慎重に設計する必要があることがわかる.図 (c) はキーポイントの設定方法と認識 成功率との関係である.後述する SHOT 法ではランダム選択が用いられているが,キーポイ ントが平面あるいは平面に近いところに設定されてしまうと,類似の特徴量をもつ点が周囲に も存在するという危険性があることから,凸部に設定するほうがユニークな特徴量になるとい

27



図 28 キーポイントマッチングの原理

う利点がある.しかし一方では,凸部形状が存在しにくい対象物も少なくないことから,これ についても慎重な検討が必要になる.



図 29 特徴量の重要性

4.3.2 3次元特徴量の分類

以上のことから,キーポイントマッチングにおける特徴量の研究例は多いが,大きく分ける と,図 30 のような 2 つのタイプに分けられる.なお,3 次元特徴量に関する詳細なサーベイは 文献 [1] にある.



図 30 2つのタイプの 3次元特徴量

ひとつめのタイプ (A) は、キーポイントひとつひとつに対して、そのキーポイントの周辺に 存在する点群のみから特徴量を計算、記述するタイプである.ふたつめのタイプ (B) は、モデ ル物体上に設定した 2 つまたは 3 つのキーポイント間の幾何学的関係そのものを特徴量として 扱うタイプである.

いずれのタイプについても非常に多くの研究例があるが,上記の2つのタイプに分類した結果を,図31に示す.図中で,SHOT,PFHなど略号で記載されているものが,特徴量である. 図の下部にフルスペルを示す.

4.3.3 タイプ A の特徴量(キーポイント周辺のデータをもとに算出)

4.3.3.1 概要

このタイプに属する特徴量の記述の形態としては、計算された情報をそのまま保持するもの もあるが、これをヒストグラム化して統計量として扱うこともある.また、記述領域を事前に 複数のサブ領域に分割し、サブ領域毎に算出された特徴量の分布を保持することもある.これ は、特徴量におおまかな位置情報を含んでいることに相当する.

従来から比較的よく用いられてきた例として, Spin Image 法 [26, 27, 28] がある. これは, キーポイントの法線に垂直な面を設定し,法線方向βとそれに直交する方向αに関して, βを 軸とする円柱面に周辺の点座標を投票し,これを画像的に表現した特徴量である. また武口ら によって提案された Depth Aspect Image (DAI) 法 [29, 30, 31, 32] は,物体から特徴的な 3

キーポイントベース手法						
(a) +-:	ポイント周りの	D情報記述	(b) 複数点間の関係記述			
座標データ利用	座標データ利用 法線ベクトル 利用		2点間 or 3点間			
SHOT, PFH, CSHO	DT, HONV	LSP, CCDoN, NARF,	PPF, ER, VC-PPF, MPPF,			
SI, SSI, DAI, 3DSC, SDLSD	FPFH, DoN	MHOG, RoPS	PPF B2B or S2B or L2L, VPM			
 PFH : Point Feature Hist SHOT : Signature of Histo of OrienTations CSHOT : Color SHOT HONV : Histogram of Oriented Normal N SI : Spin Image SSI : Spherical Spin Ima DAI : Depth Aspect Imag 3DSC : 3D Shape Context 	ogram SDLSD grams FPFH DoN LSP /ector CCDoN ge NARF e MHOG s RoPS	 Scale-Dependent Local Shape Descriptor Fast PFH Difference of Normals Local Surface Patches Combination of Curvatures and Difference of Normals Normal Aligned Radial Feature Mesh HOG Rotational Projection Statistics 	PPF : Point Pair Feature ER : Efficient Ransac VC-PPF : Visibility Context PPF MPPF : Multimodal PPF PPF B2B or S2B or L2L : PPF Boundary-to-Boundary or Surface to Boundary or Line to Line VPM : Vector Pair Matching			

図 31 3 次元特徴量の分類と特徴

点(曲率の大きな点)を抽出して ACF(Aspect Coordinate Frame)と呼ぶ平面を構成し、こ れに周辺座標を投影した画像(DAI)を特徴量としている. どちらも画像情報を特徴量として いるために扱いやすいが、データ容量や計算コストの点では改善が必要であった.

これらに対し、キーポイント周辺の記述領域から得られた情報を、画像ではなく特徴ベクトルとして記述する方法があり、タイプAの一連の特徴量の源流になっている. 3D Shape Contexts 法は、キーポイント周りの領域を緯度、経度、直径方向に分割し、領域内の点群の個数をヒストグラム表現した特徴量である.また後述する Point Feature Histogram (PFH) [33] は、キーポイントと球領域内の k 個の近傍点から多数の 2 点組を選択し、各点における法線ベクトルから計算される角度等のパラメータをヒストグラム特徴としている.

さらに、Tombari らによって提案された著名な Signature of Histograms of OrienTations (SHOT) 特徴量 [35, 36] やその関連 [45] もまた、法線ベクトルを利用したものである. SHOT 特徴量はキーポイント周辺の球状の領域を 32 分割し、それぞれにおける法線ベクトル群とキー ポイント法線との関係を内積値としてヒストグラム化し、352 次元の特徴量として算出してい る. このほか、Chen らによって提案された Local Surface Patches (LSP) [37] は、その 3 次 元点の局所的な凹凸形状を表す Shape Index 値と呼ばれる指標を利用している.

Aldoma らが提案した OUR-CVFH (Oriented, Unique and Repeatable Clustered Viewpoint Feature Histogram) 特徴量 [42] は,従来の特徴量と大きく異なる.従来のキーポイン トがすべてのポイントクラウドデータ,あるいは形状的な特徴(例えば,凸部の先端付近など) を持つ点であったのに対し,モデル表面を曲面領域に分割し,領域ごとに特徴量を記述してい る点である.点群から抽出されたなめらかな連結領域(クラスタ)をもとに共分散行列によっ て局所参照座標系(LRF)を設定し,これに基づいて当該領域から 303 次元の特徴ベクトルを 生成している.

このタイプの特徴量としては、上記の他に HONV 特徴量 [43], 3D-Div 特徴量 [58], RoPS 特徴量 [59] などがある.

4.3.3.2 SHOT 特徴量に基づくキーポイントマッチング

SHOT(Signature of Histograms of OrienTations)特徴量 [36] は,2010 年に Tombari ら によって提案された特徴量である.万能の特徴量ではないが,その考え方や特徴量を求める際 の基本的な考え方には学ぶべき点が多く,ある意味では 3 次元特徴量のデファクトスタンダー ドと言っても差し支えない.そのため,筆者は当研究室に特定物体認識の共同研究テーマが持 ち込まれた際には,まずはオープンソースソフトウェアなどを利用して SHOT 特徴量を用い た認識実験をおこない [56],その振る舞いを詳細に分析することを推奨している.この作業に よって特徴量そのものの改良を行うべきかどうか等の開発方針を効率よく検討できるからで ある.

SHOT 特徴量の求め方を図 32 を用いて説明する.まず最初にモデル物体からキーポイント が抽出される.SHOT の原著論文ではランダムに配置することになっている.図中の赤丸が キーポイントである.このキーポイントに対してキーポイント周辺の点群情報だけを用いて特 徴量を作っていく.



図 32 SHOT 特徴量

最初にすべきことは座標系(局所参照座標系)を定めることである.詳細は後に 4.3.5 にて 解説するが,図のようにキーポイントを原点として,法線方向に Z 軸がくるように直交座標系 を設定する.次に,キーポイントを中心とする仮想的な球(サポート球)を設定する.球の半 径は重要であるが,たとえば,半径 1~2cm 程度を想定すればよい.今後の処理では,このサ ポート球の内部に存在する点群データのみが処理対象になる.

次にこのサポート球を分割する.まず最初に XY 平面で上下に分割する.サポート球を地球 にたとえると,赤道面にて北半球と南半球に分割するようなイメージである.次に,さきほど 設定したサポート球の内部に,同じキーポイントを中心とする半径の小さい2つめの球を設定 し,その内部と外部に2分割する.最後に,Z軸を含む面で分割する.サポート球を地球にた とえると,経度 45 度おきに地球を8 個に縦割りにするようなイメージとなる.ここまでの処 理で,2x2x8 の合計 32 分割したことになる.

この 32 個の小片領域については,各小片内部に点群が含まれている(点群数がゼロの場合もあり得る). 32 個の小片領域の1つについて説明を続ける.この小片領域に含まれる複数の3次元点は,それぞれ法線ベクトルを持っていると考える.これを**n**_iと書くことにする.正確に

言えば、3 次元点一つから法線ベクトルを求めることはできないが、その点の近傍に存在する 数点を用いて複数の小さな面を作り、その面の法線をもって注目点の法線と定義するなどの方 法により、便宜上、点ごとに法線ベクトルを付与することとする.一方、最初に設定したキー ポイントにも法線ベクトルがあると考える.これを r と書く.ここで、r と n_i の内積 cos θ を 計算する.小片に含まれる 3 次元点は複数個存在するから、この内積も複数個計算されること になる.この複数個の内積を 11 ビンのヒストグラムとして表現する.これが一つの小片に関 する特徴量となる.

小片は合計 32 個存在するから,この 11 ビンのヒストグラムは 32 個作られることになる. これをすべて一列に並べると 352 個の数字になる.これが SHOT 特徴量である.SHOT 特徴 量は,注目しているキーポイントの周辺部位を粗く領域分割し,領域毎の法線方向の分布を Z 軸を基準とした数値で表現しなおしたものになっている.したがって,たとえば領域毎の支配 的な法線ベクトル方向が,どの向きであるかという大局的な法線方向分布情報が内包されて いる.

2 次元キーポイントマッチングの技法として SIFT 特徴量があるが, SIFT も注目キーポイント周辺の濃度勾配方向の分布を表現しているので, SHOT と SIFT との間には一種のアナロジーを感じる.また,2次元画像処理の分野では HOG 特徴量も, SIFT と同様に濃度勾配方向の分布に関係した特徴量になっているので,こちらとの類似性も感じられる.

いずれにしても,SHOT 特徴量は次元数が 352 次元と大きいために特徴量の計算には時間が かかるが,それだけにアイデンティティが強い特徴量でもある.SHOT 特徴量が似ている 2 つ のキーポイントがあるとき、それらが正しい対応点同士である可能性は高いといえる.

なお,SHOT 特徴量を用いたマッチングの方法としては,センサから入力されたシーン点群 についてもモデル同様にキーポイント設定と,キーポイント毎の特徴量記述をおこなった後, モデルとシーンとの間で特徴量が似ているキーポイント同士を結びつけ,必要に応じて偽の キーポイント対応を減らすためのロバスト推定を経たうえで,最終的にはモデル物体をシーン に移動させて一致させるための回転行列 *R* と並進ベクトル *t* を計算することになる.

4.3.3.3 PFH 特徴量に基づくキーポイントマッチング

SHOT と同様にタイプ A に属する特徴量として、PFH(Point Feature Histgrams)特徴量 がある [33]. この特徴量も図 33 のように注目キーポイントを中心とするサポート球を想定し、 その中に含まれる点群のみから特徴量を計算する. PFH においてはサポート球内のすべての 3 次元点について 2 点を選択し、 α 、 θ 、 ϕ 、 $\mathbf{p_t} - \mathbf{p_s}$ のパラメータを算出する. これをヒストグ ラム化して 125 次元の特徴量として定義する.

PFH 特徴量の変形版として FPFH (Fast Point Feature Histgrams) 特徴量がある [34]. こ れは図 34 に示すように図 33 と似ているが, PFH がサポート球内の点群からさまざまな 2 点 ペアを取り出して特徴量にしたのに対し, FPFH においてはペア点群の片方を必ず注目キーポ イントに固定したものになっている. すなわち, 2 点ペアといっても片方だけがサポート球内 の任意の点として扱われることによってペア数が減り, 結果として特徴量記述が高速化される という効果がある.



図 33 PFH 特徴量



図 34 FPFH 特徴量

4.3.4 タイプBの特徴量(複数のキーポイント間の関係を特徴量化)

4.3.4.1 概要

このタイプの特徴量は、2点または3点の幾何学的な関係を記述している.したがって、モ デル表面のどの点を組み合わせるか、またその組み合わせにおけるどのような幾何パラメータ を用いるかの2つが主要な設計要素となる.

この手法の代表例は,後述する Drost らの Point Pair Feature (PPF) [46] である. これは, 物体モデル上の全ての 3 次元点から 2 点(点対)を選択して組み合わせて 4 つのパラメータを 4 次元特徴量 F として定義したものである. PPF は Choi らによって改良されており,また 我々もベクトルペア特徴量を提案している. タイプ A と比較すると,一定の距離を離れた点対 を利用することから,大局的な特徴であり,凸部が少ない対象物に向いているが,距離は慎重 に検討する必要がある. 短ければタイプ B の効果が小さくなるし,長ければ照合時に同一物体 上の 2 点でなければならないという条件が満たされなくなる可能性が増す.

その他の手法として,点群密度への依存性を低減した特徴量(RPD)[54,55] や VFH 特徴 量などがある [18]. RPD は生産現場ではセンサと対象物との距離が変化する可能性は低いが, AGV のための物体認識などでは対象物までの距離が一定しないことに起因してセンサで得ら れる点群密度が変化しやすい場合もある.このようなときに有効な特徴量である.

4.3.4.2 PPF 特徴量に基づくキーポイントマッチング

PPF (Point Pair Feature) は Drost によって提案されたものであり [46],図 35 のように物体モデル上の全ての 3 次元点から 2 点(点対)を選択して組み合わせて 4 つのパラメータを 4 次元特徴量 F として定義したものである。物体モデル上の全ての 3 次元点から 2 点(点対)を 選択して組み合わせ、図のように 2 点間の距離 (F1)、2 点を結ぶ線分と各点の法線方向との角度 (F2,F3)、2 点の法線間の角度 (F4)、の計 4 つのパラメータを 4 次元特徴量 F として定義したものである。



図 35 PPF 特徴量に基づくキーポイントマッチング

また,対応づけ後には点対間の剛体変換パラメータを求めることになる.このパラメータを 投票空間に投票すれば,高得票のパラメータは多くの点対が共通の剛体変換パラメータを支持 していることを意味するので,多少の誤照合が発生したとしても,モデルが存在する位置と姿 勢を正しく求めることができるという利点がある.

なお, PPF を発展させた手法として, 点対の選択方法に工夫を凝らした Choi らの手法 [48] がある. これは図 36 に示すように, 点対をある一定の条件を満たす部位から選択するもので あり, 部位としては表面 (Surface), 稜線 (Boundary), 線セグメント (Line Segment) が提 案されている. 図中で, たとえば S2S は 2 つの点を両方とも表面から選択する場合, B2B は 両方とも稜線から選択することを意味している.

4.3.5 局所参照座標系と主要技術

本節では、3 次元特徴量を語るときに避けて通れない概念である、局所参照座標系(LRF: Local Reference Frame)について、その意味と分類および主要な研究動向を概説する.

4.3.5.1 局所参照座標系の概要と分類

局所参照座標系(LRF: Local Reference Frame)とは、図 37 に示すようなキーポイントご とに設定される座標系のことである.この座標系の最も大きな役割は、特徴量の定義である. すなわち、前章で述べたさまざまな特徴量はこの LRF に基づいて数値化されるため、LRF の



図 36 PPF 特徴量マッチングの改良

安定性は特徴量の安定性に直結する.また LRF は対応付け後の2つのキーポイントの幾何学 的関係を表現しているので,物体の姿勢推定に用いられることもある.



図 37 局所参照座標系

LRF の性能がマッチング性能そのものに大きな影響を与えるというエビデンスを,図 38 に 示す.

この実験では特徴量としては SHOT を用い,LRF の算出方法を3種切り替えて認識性能を 検証した.図からわかるように,SHOT 特徴量についてはこの論文で提案されている LRF と の相性がよく,よい性能になっているが,もし仮に LRF のみを後述する Mian 法に差し替え



図 38 局所参照座標系の種類と認識性能

たところ,性能が下がっていることがわかる.なお2つの性能グラフ(ROC 曲線)は,左側 が点群密度が高い場合,右側が低い場合である.絶対的な性能数値には差があるが,LRF を差 し替えれば性能が変化するという意味では傾向は同じである.

通常,LRFは3次元直交座標系である.まず,第1軸(z軸)は注目キーポイント周囲の局 所的な面の法線ベクトルであることが多く,比較的安定に求めることができる.一方,第2軸 (x 軸)はこれと直交するベクトルである.第3軸(y 軸)は,第1軸および第2軸をもとに外 積ベクトルとして算出されることになる.すなわち,LRFの設定における現実の課題として は,第2軸(x 軸)の設定が最も重要である.LRFの研究事例は,特徴量の事例に比べれば多 くないが,この状況を図 39 に示す.

固有ベクトル利用型	法線ベクトル利用型		
キーポイント周りの点群の座標から共分散行列を計算し, その固有ベクトル群を LRF とする.	3軸をそれぞれ独立に算出する. z 軸=キーポイント周りの <u>法線ベクトル</u> x 軸=さまざまな工夫		
	x 軸の決定方法		
Mian, SHOT,	最大值選択	方向分布分析	
RoPS, EM	BOARD PS	MH DPN, GRF	
Mian : Mian-LRF SHOT : Signature of Histograms of OrienTations RoPS : Rotational Projection Statistics EM : Exponential Map	BOARD : BOrder Aware PS : Point Signatu DPN : Dominant Pro GRF : Global Refere MH : Mesh HoG	e Repeatable Directions re jected Normals ence Frame	

図 39 LRF の 2 つのタイプ

固有ベクトル等を利用して3つの軸を一括的に算出するタイプと,法線などを基準にして一 つずつ個別に設定するものの2つのタイプに分けることができる.次項より,いくつかの主要 な LRF 算出手法を具体的に説明するが,総じて2つめのタイプのほうが性能が高いことがわ かっている.

4.3.5.2 Mian *Φ* LRF

Mian らが提案した手法 [57] であり,固有ベクトル利用型に属する.図 40 に示したように, 注目点を中心に予め設定した半径 R のサポート球を考え,その後はこのサポート球内部の点群 のみを計算に利用する.



図 40 Mian の LRF

サポート球内部の k 個の点群 \mathbf{p}_i を考えると、これらはすべて x, y, z の座標である. k 個の 点群の平均ベクトル $\hat{\mathbf{p}}$ を計算し、これを使ってさらに次式によって分散共分散行列 \sum を求 める.

$$\hat{\mathbf{p}} = \frac{1}{k} \sum_{i=0}^{k} \mathbf{p}_i \tag{1}$$

$$\boldsymbol{\Sigma} = \frac{1}{k} \sum_{i=0}^{k} \left(\mathbf{p}_{i} - \hat{\mathbf{p}} \right) \left(\mathbf{p}_{i} - \hat{\mathbf{p}} \right)^{T}$$
(2)

分散共分散行列をもとに主成分分析を行うことができ,得られた3次元の主成分ベクトルを そのまま LRF として採用するという考え方である.主成分ベクトルは直交基底を張ることが わかっているので,その点でも手軽な手法である.

4.3.5.3 Mian の LRF の改良

SHOT 法の論文において, Tombari らは, Mian の LRF を改良した方法を用いている. 図 41 のように, 注目点すなわちキーポイントから各点までの距離 *d_i* を算出し, 重みファクター として利用している. これにより, キーポイントに近い点は重く, 遠い点は軽く扱われる.

$$\boldsymbol{\Sigma} = \frac{1}{k} \sum_{i=0}^{k} (R - d_i) (\mathbf{p}_i - \hat{\mathbf{p}}) (\mathbf{p}_i - \hat{\mathbf{p}})^T$$
(3)

4.3.5.4 BOARD-LRF

BOARD (BOrder Aware Repeatable Directions) は,法線ベクトル利用型に属する [60]. 図 42 のようにして 3 つの軸を定める.



図 41 Mian の LRF の改良



図 42 BOARD-LRF

まずキーポイントを中心にサポート球を設定するところは Mian の方法と同様である.サ ポート球内の点群を平面にフィッティングし,その法線方向を Z 軸とする.次にサポート球内 の点群それぞれについて法線ベクトルを決定し,そのベクトル群の中から Z 軸ベクトルとの内 積が最も小さい,すなわち最も直交に近いベクトルを選び出す.このベクトルをさきほどの平 面に投影したものを X 軸とする.このとき,3次元センサの状態によっては点群の部分的な欠 損(情報欠落)があり得るので,周辺の点群からの補完法も提案されている.ここまでで X 軸 と Z 軸が定まったので,残りの Y 軸は XZ 平面に直交するベクトルとして,外積計算により求 まる.なお,ある点の法線を求める際には,実際には面を仮定する必要があることから,たと えば点の周囲の点も利用していくつかの三角パッチを形成し,それから計算される法線ベクト ルの統計的処理によって法線を求めるなどの工夫が必要である.

当研究室におけるこれまでの経験では,この BOARD 法の安定性が高いことがわかっている.

4.3.6 その他のモデルベース物体認識

4.3.6.1 EGI を用いた物体認識

これまでに説明した 3 次元特徴量は,ほとんどの場合,法線情報を利用している.これはモ デルや対象物の面の方向が,形状を表現する手がかりとして有効であるとの考え方に基づいて いる.その意味では,EGI (Extended Gaussian Images)も同様である [19].

これは図 refExtendedGaussianImage に示すように,対象物の法線方向分布を統計的に扱う 枠組みである.図(a)は立方体の例であるが立方体は6個の面があり,それぞれの面ごとに法



⊠ 43 ExtendedGaussianImage

線ベクトルが一つずつ決まる.この法線ベクトル群の始点をすべて原点にそろえ,しかも法線 ベクトルの長さを1に正規化して表示すると,同図右のように半径1の球の表面上にベクトル の終点が集中する仮称が6個できる.言い換えると,法線ベクトルの分布をみれば,集中箇所 が6カ所あることからこの物体には平面が6個あることや,2面ずつ正負逆の法線が存在する ことから立方体または直方体であることなどが読み取れる.

同図 (b) は円柱の例であるが, (a) と同様に上面と底面は逆向きの法線ベクトルがペアを作っており, 側面については法線ベクトルの軌跡が図のように円環状になることがわかる. この EGI 表現を用いることによって, 対象物の形状や向きを推定することができる.

4.3.6.2 Spin Image

Johnson らによって提案された手法 [26, 27, 28] であり,点の法線をベースにした特徴量を 画像表現している.局所的な面の状態を特徴量として捉えるという意味では先駆的な研究であ るが,処理時間がかかることや画像表現としての特徴量の記録容量が大きいなど,実用上の課 題もある.本稿においては詳細説明は省略する.

4.3.6.3 DAI を用いた物体認識

Spin Image と類似した方法として, 武口らによって提案された Depth Aspect Image (DAI) 法 [29, 30, 31, 32] がある. これは, 図 44 のように物体から特徴的な 3 点を選択して ACF (Aspect Coordinate Frame) と呼ばれる平面を算出し, この面に対して周辺の 3 次元データ を座標を投影した画像 (DAI = Depth Aspect Image) を特徴量としている. いわば, Spin Image の改良版に相当するような考え方である.



🗵 44 Depth Aspect Image

4.3.7 ICP による高精度位置合わせ

局所特徴量ベースのキーポイントマッチングとは異なるが,モデルベース認識としては, ICP について触れないわけにはいかない. この手法は,図 45 のように,モデル X と入力データ P の 2 つの点群を精密に位置合わせするための技術である.



図 45 ICP

図中の式のように、入力データ点 $\mathbf{P} = \vec{P_i}$ に対して回転行列 Rと並進ベクトル tを作用させた点と、モデルデータ $\mathbf{X} = \vec{X_i}$ のユークリッドノルムの平均値を誤差と定義し、これを最小化するような R と tを求める. 誤差最小化問題を解析的に解く場合は、局所解に陥らないように事前の位置合わせ(初期位置の設定)が重要である. また、計測データ側には誤差が乗ったデータがあり得るため、ロバスト推定を組み込んで外れ値除去の機能を有する方法など、ICPには多くの発展手法 [2, 3, 4, 5, 6, 7, 8, 9, 10, 11] がある.

なお,これは ICP に限らず点群位置合わせには共通の問題点であるが,モデル側とデータ側 ではそもそも対象物上で物理的に同一の点がサンプリングされているという保証がないことに は注意しておく必要がある.すなわち,一定の高密度でサンプリングされているとしても,点 群間の距離は理論的最小値であるゼロを取る可能性はほとんどない.ただし,ICP アルゴリズ ムについてはかなり実用的な性能を達成できるソフトウェアが存在し,初期位置合わせの精度 がよければ,問題なく利用できるレベルになっている.

5 3次元物体認識に関する主な研究事例

本章では,当研究室における研究事例を中心に,対象物の位置姿勢認識に関する研究,およ びロボットとの接点を意識した研究について主要なものを紹介する.

5.1 対象物の位置姿勢認識に関する研究事例

5.1.1 可観測性を考慮したベクトルペアマッチング

4.3.3 に属する特徴量として,図 46 に示すようなベクトルペア (Vector pair) 特徴量 [49,50] がある. これは 3 次元姿勢を同定するための最小データセットである 3 つの 3 次元点を,始点 を共有する 2 本の空間ベクトルとして表現し,3 つの端点にさまざまな特徴記述子を設定する という概念である.2 次元画像照合としてのテンプレートマッチングの高速化,高信頼化手法 としてユニークな画素のみを選択する手法があるが [63],3D マッチングにおいてもユニーク な特徴量が好ましいという考え方は妥当であり,たとえば TriSI 特徴量もその一つである [44]. この考え方を踏襲して,モデル内の発生確率を分析してユニークなベクトルペアを選択するこ とによって誤照合を低減している.



図 46 3D ベクトルペア

ベクトルペア特徴量を用いたモデル照合手法 (VPM: Vector Pair Matching) では,ユニー クなベクトルペアであることを優先したため,凹部のような外部から観測されにくい点が選択 されてしまうという問題があった.マッチングにあたっては,当然ながら観測可能な特徴量の ほうが好ましいため [47],図 47 に示すような可観測な点を優先して照合に用いる手法 [51] が 開発された.

モデル物体をさまざまな方向から観測するというシミュレーションを実施し,多くの方向から観測可能であるほど適切な点であるとの判断基準を加えた.実験によれば,この効果は大きく,ユニークな点を選択するよりも効果が高いことがわかっている.



図 47 キーポイントの可観測性の利用

5.1.2 入力シーン予測に基づく特徴量選択

特徴量を設計する際に,そもそも,モデルを考慮するだけでよいのか?という基本的な問題 がある.よい特徴量とは,モデルのみから設計できるものではなく,入力シーンデータにも依 存して設計されるべきものであろう.この考え方に立脚し,3次元 CG と物理シミュレータを 利用して,特徴空間内での照合性能を事前予測する手法 [?,53] が提案された.

図 48 に示すように、ランダムにばらまいた特徴量群に対して、物理シミュレータ(Sketchey Phisics)と 3D-CG を用いてばら積みシーンを人工的に生成し、これをもとにどの特徴量が正 しくマッチングに寄与する可能性が高いかを予想し、性能の高い特徴量を優先して選び出すこ とに成功した.

ここでは、よい特徴量の定義として、その特徴量がポジティブサンプル群(正しくマッチさ せたい正解サンプル群)との類似性が高く、ネガティブサンプル群(マッチさせたくない正解 ではないサンプル群)とは類似性が低いという指標を用いている.

5.1.3 CCDoN 特徴量に基づくキーポイントマッチング

一般的に、キーポイントマッチングにおける特徴量は、高次元であるほどアイデンティティ が高いが処理時間がかかり、低次元であればその逆の性質を持っており、いわば trade-off の



図 48 入力シーン予測に基づく特徴量選択

関係にある. これに対して永瀬らは, Ioannou らによる DoN (Difference of Normals) 特徴 量 [38] を改良し, DoN に対して大小 2 つの球状領域で算出した曲率値 (Curvature) を追加 した 3 次元特徴量 CCDoN 特徴量 [39, 40] を提案した. 概要を図 49 に示す. CCDoN 特徴量 は, わずか 3 次元という低次元数でありながら一定のアイデンティティを表現する能力を持っ ている.



図 49 CCDoN 特徴量に基づくキーポイントマッチング

5.1.4 SHORT 特徴量に基づくキーポイントマッチング

一般に3次元特徴量は、面方向などの局所形状を反映していることが望ましく、点群情報を 用いてできるかぎり正確で安定な形状情報を得る必要がある.このとき、個々の3次元点だけ では面素の「方向」情報が得られにくいことから、多くの特徴量は、キーポイント検出、特徴 量記述の両方において、複数の3次元点から計算される法線ベクトルを用いて面方向情報を表 現している.しかし法線ベクトルを算出するためには、たとえば3点からなる面素(三角パッ チ)を定義する必要があることから、法線の利用は一般的には処理時間の増大を招きやすい.

この問題に対して,武井らが提案した Shell Histograms and Occupancy from Radial Transform (SHORT) 特徴量 [52, 53] では,キーポイントの設定および特徴量の記述それぞれ について高速化の工夫がなされている.

まずキーポイントの設定については,図 50(a)のように,注目点を中心に設定された5層からなるマルチスケールシェルを考え,各層に含まれる点群の数をカウントする.これは,点群の空間占有率に相当するが,なだらかな表面上の点では大きく,形状的に不連続な部分では小さくなる傾向があることから、これを用いて高速にキーポイントを検出することができる.

次に,特徴量の算出については,同じくマルチスケールシェルを用いて,各シェル内に存在 する3次元点と,別途算出した主方向ベクトルとの内積値のヒストグラムを特徴量としている. この処理も,内積計算のみであり,きわめて高速に実行できるという特徴がある.

SHORT 特徴量は,当研究室において最後に開発された3次元局所特徴量であり,高速性と 認識性能については,いまのところこの特徴量の性能が最も高い.



図 50 SHORT 特徴量に基づくキーポイントマッチング

5.2 ロボット動作生成のための物体認識に関する研究事例

本稿の主たる狙いは、特定物体認識に関する諸技術について、比較的問い合わせの多い内容 を概括的に紹介することであるが、ロボットビジョンとしては特定物体認識以外の認識タスク も多い. 2.2 で述べたように、さまざまなシーンで対象物を掴むというタスクにおけるビジョ ンの基本課題は、どこに、なにがあって、どう掴むか、という3点になる.本節では、3つめ の「どう掴むか」に関連する研究事例を紹介する.

5.2.1 プリミティブ近似によるモデルレス把持位置推定

認識した対象物は、掴まなければ意味が無い.対象物の位置と姿勢が正確に把握できたとし ても、ロボットハンドがそれを掴み上げるためには、ハンドが対象物のどこを狙ってアプロー チすべきか、把持点とアプローチの方向などの情報が必要になる.対象物の3次元形状モデル の利用を前提にできる場合は、たとえば形状モデルに、把持すべき位置などの情報を埋め込む ことも考えられる.むろん、そのような情報を埋め込んだとしても実際にその点が見えている かどうかは保証されない.対象物のモデルを持たない場合には、なおさら、把持位置の推定ア ルゴリズムが必要になる.

実際,2015 年から3 年間にわたって開催された,物流ロボットの国際大会である Amazon Robotics Challenge では,2017 年大会では,対象物に関するモデル無しに品種を同定し,移 載する必要があった.現実の物流・配送センターでも,倉庫に格納されているアイテムについ て,形状モデルを利用できないというシチュエーションは十分にあり得る.このことから,モ デルレスでの把持位置推定技術は、ロボットピッキングにおける重要な課題であるといえる.

これに関連して,我々が開発したプリミティブ近似方式 [64] について,図 51 を用いて説明 する.この方法は,いわば brute-force の典型的な方法である.要は,目の前にあるすべての ピッキング対象物を,事前に設定されたいくつかのプリミティブ形状モデルのどれか一つに無 理矢理あてはめてしまうという考え方である.具体的には,直方体,円柱,球の3つのプリミ ティブ形状が設定され,対象の点群シーンをこれらの3つのどれかに当てはめることにした. プリミティブ形状の数を限定することによって、プリミティブ物体毎に適切な把持パターンを 事前に検討して辞書データベースとして格納しておくことができる.たとえば、平面のみから なる直方体に当てはめることができた場合は、吸着ハンドであれば上面や側面から面の中央部 にアプローチすればよいし、2指のグリッパの場合は上面の辺の両脇から挟み込むようなハン ド接近動作をさせればよいなど、アプローチの候補を読み出すことができる.実際の把持にあ たっては、上面の傾きや周辺の間隙の状況などを把握し、候補動作のなかから最も適切な動作 を選択することになる.



図 51 プリミティブ近似によるモデルレス把持位置推定

この手法の概要を,図 52 に示す.3 次元センサから入力された3 次元点群データに対して, 面の候補を切り出すセグメンテーション処理がおこなわれ,その結果(セグメント)に対して 3D-DNN による面プリミティブ当てはめがなされる.ここで面プリミティブとは,3 次元空間 上に存在する面という意味である.図の例では,対象物の上面について長方形らしさが 0.32, 円盤らしさが 0.21,円柱側面らしさが 0.27 などの尤度分布が出力される.側面についても同 様である.これらの尤度分布を事前に設定したルール群を用いて統合することによって,面プ リミティブラベルがついた複数の面が認識され,この場合は最終的に直方体という立体プリミ ティブが当てはまることになる.

この図では、物体一つ一つが離れて配置されているシーンを対象にした実験結果が示されて いるが、異種物体が混載されたばら積みシーンを上方から撮影した深度画像に対しても適用可 能である.ただし、モデルレスでの認識であることから、どの面プリミティブが一つの立体を 構成しているのかは不明であり、隣の別の物体の面セグメントを考慮してしまうことによるエ ラーは起こりえる.

また,世の中のさまざまな形状の対象物をわずか3種のプリミティブに当てはめることが妥 当かという疑問が生じるが,これに対しては一般的な日用品の3次元モデルデータベース約 600個を用いた実験により,全体の約85%の形状に対してはロボット把持が可能という意味で この3種のプリミティブへの当てはめが妥当であるとの予備実験結果を得ている.



図 52 プリミティブ近似手法

5.2.2 ピッキングリスク最小化のためのロボット動作生成

この例では,図 53 のような異種物体が近接して置かれているシーンに対するロボット動作 生成について,物体認識がどのように寄与できるか,一例を示す [65]. ロボットが把持目的の 対象物を掴もうとしてハンドを近づける際に,図のように対象物の周辺に障害物が存在する場 合がある.円滑な把持動作のためには,できるかぎり周辺障害物への干渉を避けた動作パスの 生成が好ましい.



把持余裕度推定モデル p:把持パラメータ(把持位置,姿勢,開口幅)



図 53 ピッキングリスク最小化のためのロボット動作生成

提案手法においては、まずシーンの構成要素として、ロボットハンド自身、把持目的の対象

物,それに周辺障害物の3種があると仮定した.ここで3種の要素に対する2者間の関係を整 理すると,以下の3つになる.

1. ロボットハンドと対象物の関係(掴みやすさファクター)

- 2. ロボットハンドと障害物の関係(動的なクリアランスファクター)
- 3. 対象物と障害物の関係(静的なクリアランスファクター)

1 については当然ながら対象物を安定に掴めるような把持位置やアプローチが好ましいし、2 については障害物との接近を回避することが望ましい.3 については静止物同士の関係である が対象物の把持位置としては障害物とのクリアランスが大きい箇所が優先されるべきである. このような3種の関係性をビジョンにより計測し、次式による把持余裕度 P_s を計算する.こ こで P_{H-O} は動的クリアランスファクター、 P_{H-T} は掴みやすさファクター、 P_{T-O} は静的な クリアランスファクターを示す. α 、 β は、調整用の重み係数である.

$$P_s(\mathbf{p}) = P_{H-O}(\mathbf{p}) \{ \alpha P_{H-T}(\mathbf{p}) + \beta P_{T-O}(\mathbf{p}) \}$$
(4)

ここで,**p**はロボットの動作パラメータを表すベクトルであり,具体的には把持位置座標, 把持アプローチの姿勢パラメータ,および2指ハンドの開き幅パラメータから構成されている. この式は値が大きいほどその動作パラメータに余裕があり円滑に把持できる可能性が高いこと を示している.したがってこの式を最大化するようなパラメータ**p**を求める次式のような最適 化問題となる.なお,3つのファクターが単純な重み付き線形和のような形式になっていない 理由は,3つのファクターのうち特に動的クリアランスファクターを優先したためである.

$$\hat{\mathbf{p}} = argmax P_s(\mathbf{p}) \tag{5}$$

5.2.3 日用品の機能認識とその応用

これまでの事例では、Pick and Place タスクを想定していた. ここでは、Pick and Operation タスクのためのロボット動作に対して、ロボットビジョンがどのように貢献できるかについて 検討した事例を紹介する [66, 67, 68, 69, 70, 71]. まず最初に問題設定を説明する. 図 54 に示 したのは、人間による点茶動作を細かく分解した結果である. お茶を点てる動作をごくシンプ ルに言うと、お茶を茶碗に入れる、お湯を茶碗に入れる、さいごにそれらをかきまぜるという 3 つの部分からなる. もしこれをロボット動作として実現しようとすると、実際にはかなり複 雑なサブ動作が必要になる. たとえばお茶を茶碗に入れるという動作についても、まず最初に 茶さじ(スプーン)を探し、次にその持ち方を決定し、実際に持ち、お茶が入っている棗(な つめ)を探し、茶さじによるすくい方を決定し、実際にお茶をすくい、さらに茶碗を探し、お 茶の入れ方を決定し、実際にお茶を茶碗に入れる動作を実行する、という具合である.

ここで我々が注目したのは,茶さじを持つという動作である.もしこれが物流現場での Pick and Place タスクであれば,ロボットは茶さじのどこを掴んでもよいが,お茶会ロボットにとっ ては,茶さじを掴む動作は Pick and Operation の一環であり,その先にお茶をすくうという



図 54 お茶たて動作の困難性

動作が待っていることを想定した上で適切な箇所を掴む必要がある. Pick and Place と Pick and Operation の違いはここにある.

ロボットにこのような把持をさせるためには、そもそもスプーン(茶さじ)のどこをつかん でどこでお茶をすくえばよいかという知識を持っておく必要がある.そこで、我々は日用品が 持っている「機能」に着目した.図 55 に示したように、一般的な日用品は使用目的がかなり明 確に想定されているため、その目的を達成するために必要な「機能」が、部分形状に作り込ま れていることが多い.たとえばコップは水を蓄えるための大きなくぼみがあり、人間が手で持 つためのとってがついている.スプーンは何かをすくうための道具であるから、小さなくぼみ が作られており、人間が持ちやすいような柄の部分もある.もし、このような部分形状が持っ ている機能を認識することができれば、ロボットはその機能情報を利用して把持のあとに続く Operation 動作を作り出すことができるはずである.

対象物の形に機能が備わっているという考え方は、いわゆるアフォーダンス(Affordance) として研究されてきたもの[67]に近いが、一般的なアフォーダンス研究ではモノの形状そのも のが人間に対して機能を誘発すると考えているのに対して、我々が対象としている日用品は、 そもそも設計時からその機能を有するように人為的に形が定められたものであること、またロ ボットにとっては与えられたタスクに対して、利用可能なモノを使うのではなく利用すべき道 具を適切に使うことが求められているという立場から、我々はあえてアフォーダンスという用 語ではなく、機能あるいは機能属性という用語を使用している.

また,日用品などの道具に関する機能はモノの部分形状だけに起因するものではなく,コッ プの場合は重力方向との関係性,フォークの場合は動作方向と速度など,形状以外の条件が満 足される必要がある.我々はこのことを理解しながらも,本研究の初期段階としては形状情報 に着目し,形状と機能とのマッピングをベースに,ロボット動作生成への応用を試みている.



図 55 日用品が形に内包している機能

対象物の部分形状をもとに機能ラベルというクラス分類を実現するためには、3 次元形 状とクラス名をマッピングすればよい. 我々はこれを、図 56 のような機械学習ベースの方 法で実現した.入力データとしては 3 次元点群または RGB 画像、識別器としては SegNet, PointNet, RandomForest などさまざま手法を検討した結果、現時点では 3 次元点群をもと にした PoineNet が最も高性能であると判断している. 実際には図のような結果を得るため に、機械学習による仮ラベル付与の後にラベル修正のプロセスを追加した. たとえば「蓄える (Contain)」というラベルの領域内に「つまむ (Grasp)」というラベルが混入していた場合に はそれを Contain に変更するメカニズムを組み込んでいる. 図 57 に認識実験の結果例を示す.



図 56 機械学習による機能認識

このような機能情報が獲得できた場合には,ロボット動作生成において,どのような利用が 考えられるであろうか.図 58 に 2 例を示す.図(a)は適切な把持位置決定に応用した事例で ある. 左端のような RGB-D データが入力されたとき,2 指ハンドで把持できそうな箇所が候 補として自動生成される.これに機能情報を追加することによって,ロボットは把持可能候補



図 57 機能認識の結果

の中からさらに把持すべき箇所,把持に相応しい箇所を決定することができる.図(b)につい てはモデルレスにて,たとえば YOLO のようなアルゴリズムを利用してスプーンの概略位置 がわかったとき,我々の提案手法を適用することによってスプーンのどの部分にどの機能があ るのかを認識することができる.そこで grasp 領域の重心をロボット把持位置とし,さらに scoop 領域の先端を,棗(なつめ)に差し込んでお茶をすくうという動作を生成することがで きる.これにより,正確な大きさや形が未知のスプーンに対しても,適切なロボット動作軌跡 を生成することができ,結果としてスプーン種類によらず一定量のお茶をすくうことが可能に なった.

5.2.4 部品の機能部位を利用した点と動作の転移

前項で紹介したお茶会ロボットへの応用に加えて,我々は機能情報を生産ロボットに適用し ようとしている.ここでは,工場で製造される部品に関するロボット動作教示の簡素化のため の技術にこの機能情報を利用した事例を紹介する [71].一般に製造工場では,単一の部品名称 が与えられているが実際にはサイズや形が若干異なる部品群はきわめて多い.たとえばコネク ティングロッド,ギア,ベアリング,コネクタなどのカテゴリ名はその内部に多くのサイズ違 い,形状違いの部品群を含んでいる.カテゴリ名はそのような部品群の総称であるにすぎない.

たとえばロボットによる組み立て作業を考えた場合,現実の工場では,同一カテゴリに属す る対象物であってもサイズや形が異なれば部品毎にロボット動作教示作業を行われている.変 種変量生産が望まれる中で,部品毎のロボット教示は手間がかかるためにロボット導入を阻害 する一因にもなりかねない.

ところで,異種部品であっても同じカテゴリ名で呼称される理由を考えると,サイズや形状が異なっている反面,なにかの共通性があるからこそ同一カテゴリに属していると考えられる. 我々は,この共通性は前節で説明した日用品の部分形状がもつ機能と同質であると着眼した. 図 59 に提案手法の概念を示す.ここでは,コネクティングロッドという種類の部品を例に考え



図 58 機能情報の利用例

る.多種のコネクティングロッドがあるとき,その一つについては人間がロボット動作を教示 したと仮定する.もしこのロボット動作情報をサイズや形状が異なる他のコネクティングロッ ドに自動的に転移することができれば,現場でのロボット教示作業を低減させることができる.

ここで転移させるべき具体的な情報は、ロボットによる部品の把持位置、組み立てを実現す るための作用位置、それにロボット動作軌跡の3つである.提案手法では、図のような部品 A と部品 B(ともにコネクティングロッドという名称で呼ばれている)には、部分的な形状に共 通性があることを利用する.具体的には、すべてのコネクティングロッドには図のような把持 点を内包する把持領域や作用点を含む作用領域が存在し、部分形状が異なっていても共通の事 実であると考える.

そこで, 部品 A と部品 B に対して機能認識をおこない, 把持領域と作用領域を対応づける. この対応関係を利用して, 部品 A の把持位置を把持領域における位置として表現し, 部品 B の 把持領域を用いて部品 A の把持位置を転移させる. 作用位置も同様に転移させる. 把持位置と 作用位置の転移に成功したならば, それらの情報を活用してロボット動作軌跡の転移について



図 59 部品の機能部位を利用した点と動作の転移

も実現できる.

5.2.5 機能属性ラベル付き全周囲データセット

当研究室では,機能認識に関する過去の研究において構築した,機能属性ラベル付きの全周 囲データセットを公開している.機械学習による機能認識の十ヶに利用することができる.公 開中のデータセットの概要を,図 60 に示す.また,本データセットにアクセスするために必要 な URL を図 61 に示す.





図 60 機能属性ラベル付き全周囲データセット



図 61 機能認識実験用データ公開 QR コード

5.3 その他の 3D 認識に関する研究事例

本節では、3次元認識に関するその他の話題を提供する.

5.3.1 Coded Multi Flash Imaging による疑似 3D マッチング

特殊な照明装置とそれに連動して動作するカメラ [61] との組み合わせを利用して,対象物 に関する 3 次元的な凹凸情報を 2 次元画像データの形式に巧みに埋め込み, 2 次元的な画像 処理技術を適用した事例 [62] を紹介する.図 62 に示したのは,Coded Multi Flash Imaging (MFC)と呼んでいる特殊な画像撮影デバイスである.リング状に配置された 8 個の高輝度 LED 素子の中央にカメラが配置されている.この LED 素子は,図のように順次ひとつずつ点 灯して対象物を照明する.このとき各 LED の発行と同期して 1 枚の画像が撮影される.すな わち,8 方向それぞれからの照明が当たった対象物の画像が 8 枚セットで入力される.



図 62 Coded Multi Flash Imaging による疑似 3D マッチング

実際に撮影された 8 枚の画像を,図 63 に例示する.図の上部に並んでいる 8 枚の画像群が入力画像である.対象物はごくわずかな凹凸を有する小型のプラスチック製部品である.ここで重要なことは,各画像撮影時にどの方向から LED が照射されているかについて,システム自身が把握できているということである.

この8枚の画像は、8ビットの濃淡画像である.提案手法ではこれをもとに濃度勾配画像を 生成する.すなわち、各画素の値が、明るさではなく濃度勾配値であるような画像を生成する. 次に一つの画素に注目して説明する.8枚の画像群において同じ画素位置の勾配画像を読み取 ると、各画像におけるその画素の濃度勾配の大きさがわかる.図に例示するが、この濃度勾配



図 63 MFC 画像群の 8 ビット画像への集約

値が大きな値を示している画素は、その方向から照明が当たったときに他の方向からの照明時 よりも勾配が大きい、すなわちその方向からの照明に対して敏感に反応した画素であることを 意味している. もしこのときの照明方向が右方からのものであるならば、右方からの照明に対 して敏感に反応するということになり、その画素は縦方向の稜線あるいはそれに類する形状的 不連続を有しているのではないかと類推できる. すなわち、3 次元的な凹凸情報に関するある 種の手がかりを得たことになる. 図のようにこの濃度勾配値の大きさを1 位から 8 位までの順 位として表現すれば、上位(1 位~4 位)にランクされたビット(照明の方向)はその方向から の照明に敏感であることを示し、逆に下位ランク(5 位~8 位)についてはその照明方向に鈍感 なことを示している.

この上位ランク4ビットに値1を与え、下位ランクビットには0を与えると、結果としてその画素が図のように8ビットにコード化されたされたことになる.このような処理をすべての 画素についておこなう.得られた8ビット画像は、どの方向からの照明に敏感か鈍感かを表現 しているので、8ビット画像ながも対象物に関する3次元的な情報を含んでいる.この特殊画 像を用いて、たとえばテンプレートマッチングを適用して、対象物を検出することができる.

6 おわりに

本稿では,3次元的な対象物を3次元的なアプローチによって認識する基本的な手法について,センサとアルゴリズムの両面から概説を試みた.

3次元センサについては市販品を手軽に入手できる時代になり,認識アルゴリズムについて も3次元特徴量ベースのキーポイントマッチングの普及も進んでいる.しかし一方では,これ らの技術を現実の生産分野,物流分野に適用しようとすると,工業製品が緩やかな曲面で構成 されていたり,稜線にも面取りが施されているなど,これらの技術が期待している前提条件を 満たしていない場合もある.また,金属光沢や小型部品など,そもそも一般的な3次元センサ では良好にデータ化できない対象物もある.すなわち,人間と同等あるいはそれ以上の作業効 率や品質が要求される現実のシチュエーションでは,残念ながら現状の技術を組み合わせるだ けでは対応できない課題が意外に多いことに気づかされる.

むろん,深層学習などの新技術を投入して認識アルゴリズムを改良していくことは重要であ るが,おそらくそれだけでは不足であろう.センサそのものの改良や,ロボット動作との組合 せによるシステム全体のソリューションを追求するほうが,効率よく実用解に到達することが あり得ると思われる.このような状況のもとで,いま何が不足しているか,今後どのような研 究開発が必要かを考える上で,本稿が多少とも貢献できれば幸いである.併せて,拙文をお読 みくださった方からの忌憚のないご意見,ご指導をいただきたいと願う.

謝辞

本稿で紹介した当研究室の研究成果については,卒業生を含む関係学生によっておこなわれ たものである.ここに心より感謝申し上げる.

参考文献

- Manabu Hashimoto, Shuichi Akizuki, Shoichi Takei, A Survey and Technology Trends of 3D Features for Object Recognition, Electronics and Communications in Japan, Vol.100, Issue 11, pp.31-42, 2017.
- [2] P. J. Besl and N. D. McKay: "A Method For Registration Of 3-D Shapes", IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), Vol.14, No.2, pp. 239-256 (1992)
- [3] S. Granger, X. Pennec: "Multi-scale EM-ICP: A Fast and Robust Approach for Surface Registration", European Conference on Computer Vision, Vol. 2353, pp.418-432 (2002)
- [4] D.Chetverikov, D. Svirko, D. Stepanov, P. Krsek: "The Trimmed Iterative Closest Point Algorithm", In Proc. International Conference on Pattern Recognition, Vol.3,

pp.545-548 (2002)

- [5] T.ZinBer, J.Schmidt, H.Niemann: "A Refind ICP Algorithm for Robust 3-D Correspondence Estimation", In Proc. International Conference on Image Processing, Vol.2, pp.II-695-8 vol.3 (2003)
- [6] S. Kaneko, T. Kondo, A. Miyamoto: "Robust Matching of 3D Contours using Iterative Closest Point Algorithm Improved by M-estimation, Pattern Recognition, Vol.36, Issue 9, pp.2041-2047 (2003)
- [7] A. W. Fitzgibbon: "Robust Registration of 2D and 3D points sets", Image and Vision Computing, Vol.21, pp.1145-1153 (2003)
- [8] J. M. Phillps, R. Liu, C. Tomasi: "Outlier Robust ICP for Minimizing Fractional RMSD", International Conference on 3-D Digital Imaging and Modeling, pp.427-434 (2007)
- [9] A. Nuchter, K. Lingemann, J. Hertzberg: "Cached K-d Tree Search for ICP Algorithms", International Conference on 3-D Digital Imaging and Modeling, pp.419-426 (2007)
- [10] 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", 信学論 D, Vol.J94-D, No. 8, pp.1410-1422 (2011)
- [11] K. Tateno, D. Kotake, S. Uchiyama, "A Model Fitting Method Using Intensity and Range Images for Bin-Picking Applications", IEICE Trans. Inf. & Syst., Vol.J94-D, No. 8, pp.1410-1422 (2011)
- [12] K. Ikeuchi, S. B. Kang: "Assembly Plan from Observation", AAAI Technical Report FS-93-04, pp.115-119 (1993)
- [13] 村瀬洋, シュリーナイヤー: "2次元照合による3次元物体認識-パラメトリック固有空間法-",電学論 D-II, Vol.J77-D-II, No.11, pp.2179-2187(1994)
- [14] H. Murase, S. K. Nayar, "3D Object Recognition from Appearance Parametric Eigenspace Method – ", IEICE Trans. Inf. & Syst., Vol.J77-D-II, No.11, pp.2179-2187(1994)
- [15] 安藤慎吾, 草地良規, 鈴木章, 荒川賢一: "サポートベクトル回帰を用いた三次元物体の姿勢推定法", 信学論 D, Vol. J89-D, No. 8, pp.1840-1847(2006)
- [16] S. Ando, Y. Kusachi, A. Suzuki, K. Arakawa, "Pose Estimation of 3D Object Using Support Vector Regression", IEICE Trans. Inf. & Syst., Vol. J89-D, No. 8, pp.1840-1847(2006)
- [17] Y. Shibata, M. Hashimoto: "An Extended Method of the Parametric Eigenspace Method by Automatic Background Elimination, In Proc. Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp.246-249, 2013.
- [18] R. B. Rusu, G. Bradski, R. Thibaux, J. Hsu: "Fast 3D Recognition and Pose Using

the Viewpoint Feature Histogram" , International Conference on Intelligent Robots and System, pp.2155-2162 (2010)

- [19] B. K. P. Horn: "Extended Gaussian Images", In Proc. IEEE, Vol.72, Issue 12, pp.1671-1686 (1984)
- [20] R. Bolles, R. Horaud: "3DPO: A Three-dimensional Part Orientation System", International Journal of Robotics Research, Vol.5, Issue 3, pp.3-26 (1986)
- [21] 角保志, 富田文明: "ステレオビジョンによる 3 次元物体の認識", 信学論 D-II, Vol.J80-D-II, No.5, pp.1105-1112 (1997)
- [22] Y. Sumi, F. Tomita, "Three-Dimensional Object Recognition Using Stereo Vision", IEICE Trans. Inf. & Syst., Vol.J80-D-II, No.5, pp.1105-1112 (1997)
- [23] M. Germann, M. D. Breitenstein, I. K. Park, H. Pfister: "Automatic Pose Estimation for Range Images on the GPU, International Conference on 3-D Digital Imaging and Modeling, pp.81-90 (2007)
- [24] I. K. Park, M. Germann, M. D. Breitenstein, H. Pfister: "Fast and Automatic Object Pose Estimation for Range Images on the GPU", Machine Vision and Applications, Vol.21, No,5, pp.749-766 (2010)
- [25] M. Liu, O. Tuzel, A. Veeraraghavan, R. Chellappa, A. Agrawal, H. Okuda: "Pose Estimation in Heavy Clutter using a Multi-Flash Camera", IEEE International Conference on Robotics and Automation, pp.2028-2035 (2010)
- [26] A. E. Johnson, M. Hebert: "Surface Registration by Matching Oriented Points", In Proc. International Conference on Recent Advances in 3-D Digital Imaging and Modeling, pp.121-128 (1997)
- [27] A. E. Johnson, M. Hebert: "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes, IEEE Trans. on Pattern Analysis and Machine Intelligence, pp433-449 (1999)
- [28] S. R-Correa, L. G. Shapiro, M. Melia: "A New Signature-Based Method for Efficient 3-D Object Recognition", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.I-769-I-776 (2001)
- [29] 武口智行, 金子俊一, 近藤司, 五十嵐悟: "距離アスペクト画像の 2 次元照合によるロバスト物体認識", 信学論 D-II, Vol.J84-D-II, No.8, pp.1710-1721 (2001)
- [30] T. Takeguchi, S. Kaneko, T. Kondo, S. Igarashi, "Robust Object Recognition Based on Depth Aspect Image Matching", IEICE Trans. Inf. & Syst., Vol.J84-D-II, No.8, pp.1710-1721 (2001)
- [31] 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",電学論 (C), Vol.127, No.4, pp.615-622 (2007)
- [32] Y. Kitaaki, H. Okuda, M. Hashimoto, S. Kaneko, "Object Recognition with Depth Aspect Image and Hierarchical M-ICP", IEEJ Transactions on Electronics, Infor-

mation and Systems, Vol.127, No.4, pp.615-622 (2007)

- [33] R. B. Rusu, N. Blodow, Z. C. Marton, M. Beetz: "Aligning Point Cloud Views using Persistent Feature Histograms", IEEE International Conference on Intelligent Robots and System, pp.22-26 (2008)
- [34] R. B. Rusu, N. Blodow, M. Beetz: "Fast Point Feature Histgrams (FPFH) for 3D Registration", In Proc. IEEE International Conference on Robotics and Automation, pp3212-3217 (2009)
- [35] F. Tombari, S. Salti, L. D. Stefano: "Unique Shape Context for 3D Data Description", In Proc. the ACM workshop on 3D object retrieval, pp.57-62, 2010.
- [36] F. Tombari, S. Salti, L. D. Stefano: "Unique Signatures of Histograms for Local Surface Description", European Conference on Computer Vision, pp.356-369 (2010)
- [37] H. Chen, B. Bhanu: "3D Free-Form Object Recognition in Range Images using Local Surface Patches", Pattern Recognition Letters, Vol.28, Issue 10, pp.1252-1262 (2007)
- [38] Y. Ioannou, B. Taati, R. Harrap, M. Greenspan: "Difference of Normals as a Multiscale Operator in Unorganized Point Clouds", International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, pp.501-508 (2012)
- [39] 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)
- [40] M. Nagase, S. Akizuki, M. Hashimoto, "CCDoN: Local Features for High-speed and Reliable 6-DOF Pose Estimation of Randomly Stacked Objects", Journal of the JSPE, Vol.80, No.12, pp.1138-1143 (2014)
- [41] S.Takei, S.Akizuki and M.Hashimoto: "SHORT: A Fast 3D Feature Description based on Estimating Occupancy in Spherical Shell Regions", International Conference on Image and Vision Computing New Zealand, (2015)
- [42] A. Aldoma, F. Tombari, R. B. Rusu, M. Vincze: "OUR-CVFH Oriented, Unique and Repeatable Clustered Viewpoint Feature Histogram for Object Recognition and 6DOF Pose Estimation", DAGM-OAGMPRS, pp.113-122 (2012)
- [43] S. Tang, X. Wang, X. Lv, T. X. Han, J. Keller, Z. He, M. Skubic, S. Lao: "Histogram of Oriented Normal Vectors for Object Recognition with a Depth Sensor", Asian Conference on Computer Vision, pp.525-538 (2012)
- [44] Y. Guo, F. Sohel, M. Bennamoun, J. Wan, M. Lu: "A Novel Local Surface Feature for 3D Object Recognition under Clutter and Occlusion", International Journal of Computer Vision, Vol. 293, Issue 1, pp. 196-213 (2015)
- [45] S. M. Prakhya, B. Liu, W. Lin: "B-SHOT: A Binary Feature Descriptor for Fast and Efficient Keypoint Matching on 3D Point Clouds", IEEE International Conference on Intelligent Robots and Systems, (2015)

- [46] B. Drost, M. Ulrich, N. Navab, S. Ilic: "Model Globally, Match Locally: Efficient and Robust 3D Object Recognition", IEEE Computer Vision and Pattern Recognition, pp.998-1005 (2010)
- [47] E. Kim, G. Medioni: "3D Object Recognition in Range Images using Visibility Context", IEEE International Conference on Intelligent Robots and System, pp.3800-3807 (2011)
- [48] C. Choi, Y. Taguchi, O. Tuzel, M. Liu, S. Ramalingam: "Voting-Based Pose Estimation for Robotic Assembly Using a 3D Sensor", IEEE International Conference on Robotics and Automation, pp.1724-1731 (2012)
- [49] 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) 研究開発レター, Vol.133, No.9, pp.1853-1854 (2013)
- [50] S. Akizuki, M. Hashimoto, "High-speed Object Recognition using Distinctive 3-D Vector Paris", IEEJ Transactions on Electronics, Information and Systems, Vol.133, No.9, pp.1853-1854 (2013)
- [51] S. Akizuki, M. 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)
- [52] 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)
- [53] S. Takei, S. Akizuki, M. Hashimoto, "Object Recognition using Selective Features Extracted by Predicting Performance of Discrimination", Journal of the JSPE, Vol.81, No.4, pp.363-367 (2015)
- [54] 高橋祥平, 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 (2015)
- [55] S. Takahashi, S. Takei, M. Nagase, S. Akizuki, M. Hashimoto, "A Proposal of Relative Point Density Features for Object Recognition Independent on Point Density", Proc. of Dynamic Image Processing for Real Application Workshop, OS3-1, pp.1-5 (2015)
- [56] R. B. Rusu, and S. Cousins: "3D is here: Point Cloud Library (PCL)", In Proc. IEEE International Conference on Robotics and Automation, pp.1-4 (2011)
- [57] A. Mian, M. Bennamoun, R. Owens: "On the Repeatability and Quality of Keypoints for Local Feature-based 3D Object Retrieval from Cluttered Scenes", International Journal of Computer Vision, Vol.89, Issue 2-3, pp.348-361 (2010)
- [58] S. A. A. Shah, M. Bennamoun, F. Boussaid, A. A. El-Sallam: "A novel Local Surface Description for Automatic 3D Object Recognition in Low Resolution Cluttered Scenes, IEEE International Conference on Computer Vision Workshops, pp.638-643,

2013.

- [59] Y. Guo, F. Sohei, M. Bennamoun, M. Lu, J. Wan: "Rotational Projection Statistics for 3D Local Surface Description and Object Recognition", International Journal of Computer Vision, Vol.105, Issue 1, pp.63-86 (2013)
- [60] A. Petrelli and L. D. Stefano: "On the Repeatability of the Local Reference Frame for Partial Shape Matching", IEEE International Conference on Computer Vision, pp.2244-2251 (2011)
- [61] R.Raskar et al., "Non-photorealistic camera:Depth edge detection and stylized rendering using multi-flash imaging", 2004.
- [62] Yasunori Sakuramoto, Yuichi Kanematsu, Shuichi Akizuki, Manabu Hashimoto, Kiyotaka Watanabe, and Makito Seki, Reliable Image Matching using Binarized Gradient Features Obtained with Multi-flash Camera, VISAPP2015, pp.260-pp.264, 2015.
- [63] Manabu Hashimoto, Takayuki Fujiwara, Hiroyasu Koshimizu, Haruhisa Okuda, Kazuhiko Sumi,"Extraction of Unique Pixels based on Co-occurrence Probability for High-speed Template Matching", ISOT2010, MVI-3, 2010/10/26.
- [64] 鳥居拓耶,橋本学,モデルレス把持パラメータ決定のための3次元プリミティブ近似手法,計測自動制御学会論文誌, Vol.55, No.1, pp.35-41, 2019.
- [65] Kentaro Kozai, Manabu Hashimoto, Determining Robot Grasping-Parameters by Estimating "Picking Risk", IWAIT2018, A3-96, 2018.
- [66] 飯塚正樹,秋月秀一,橋本学,物体形状を考慮した denseCRF による機能属性認識の高 精度化,電気学会論文誌 C, Vol.138, No.9, pp.1088-1093, 2018.
- [67] 山野辺夏樹,万偉偉,ラミレス・イクシェル,プティ・ダミアン,辻徳生,秋月秀一,橋本学,永田和之,原田研介,ロボットマニピュレーション研究におけるアフォーダンスに関するレビュー,日本ロボット学会誌, Vol.36, No.5, pp.327-337, 2018.
- [68] Masaki Iizuka, Shuichi Akizuki, Manabu Hashimoto, Affordance-based 3D Feature for Generic Object Recognition, QCAV2017, Vol.10338, 103380W-1-6, DOI: 10.1117/12.2266917, 2017.
- [69] Masaki Iizuka, Manabu Hashimoto, Detection of Semantic Grasping-Parameter using Part-Affordance Recognition, REM2018, pp.136-140, DOI: 10.1109 / REM.2018.8421780, 2018.
- [70] Shuichi Akizuki, Manabu Hashimoto, Semi-automatic training data generation for semantic segmentation using 6DoF pose estimation, VISAPP2019, pp.607-613, 2019.
- [71] Takahiro Suzuki, Kentaro Kozai, Shuichi Akizuki, Manabu Hashimoto, Method for Recognizing Objects of Unknown Size Using Surface Primitives, IWAIT2020, 5A-48, 2020.
- [72] http://isl.sist.chukyo-u.ac.jp/Archives/