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*

永瀬 誠 信** 秋月 秀 一*** 橋 本 学†

Reliable Object Recognition Using 3-D Feature Points for Minimizing its Mismatching

Masanobu NAGASE, Shuichi AKIZUKI and Manabu HASHIMOTO

This paper introduces a reliable 3-D object recognition method which can statistically minimize its mismatching. This method basically uses a 3-D object model which is represented as a set of feature points with 3-D coordinate. Each feature point also has an attribute value about local shape around the point. The attribute value is represented as an orientation histogram of normal vector, which can be calculated by using several neighboring feature points around each point. Here, this attribute value means its local shape. By estimating relative similarity of two points of all possible combination in the model, the distinctiveness of each point is defined. In the proposed method, only a small number of distinctive feature points are selected and used for matching with all feature points extracted from an acquired range image. Finally, position and pose of the target object can be estimated from some correctly matched points. Experimental results using real scene have proved that recognition rate of this method is 93.8%, which is 42.2% higher than the conventional Spin Image method. Also, computing time is about nine times faster than that.

Key words: object recognition, 3-D feature point matching, robot vision, point cloud data, 3-D descriptor, bin-picking

1. 序 論

セル生産など多品種少量生産を必要とする製品の製造過程においては、ロボットによる自動化のニーズは高い。なかでもピンビッキングシステムは重要な要素技術であり、物体の3次元的な位置姿勢を高信頼かつ高速に認識することが求められている。

3次元物体認識の研究事例として、事前に対象物の3次元モデルを入力した距離画像と照合する、モデルベース物体認識手法がある。代表的な手法として、Spin Image法¹⁾がある。この手法は注目点の法線ベクトルを基準とし、注目点周辺のモデル点との相対的な位置関係を記述した位置姿勢に不変な特徴量を画像として表現し、物体モデルと入力距離画像を照合する手法である。視点に依存しない照合が可能であるが、物体モデル全点から算出した特徴量を照合に用いる必要があるため、処理コストがかかる。また、モデル全点ではなく、距離情報を保持したエッジを利用する手法²⁾³⁾がある。物体モデルの局所的な情報のみを照合に用いるため、モデル全点を用いる手法よりも高速な認識が可能であるが、ばら積みシーンでは、物体同士の重なりによって発生する疑似エッジが原因となり、誤照合が増加する。

また、モデルベースの他のアプローチとしては、特徴点のみを照合に用いる高速な手法として、距離データのある視点方向から見た仮説的な距離画像である、距離アスペクト画像(DAI)を作成してマッチングする手法⁴⁾⁵⁾や、法線情報とShape Index⁶⁾によって記述されたLocal Surface Patch⁷⁾と呼ばれる局所特徴量を照合に用いる手法がある。曲率の大きな形状の代表点のみを特徴点とするが、これらの点は必ずしも認識に有効とは限らない。つまり、モデル内で曲率の大きな形状を有する特徴点が

多く存在する場合には、誤照合が起こりやすくなる。

さらに近年、3次元点群データから構成される物体モデルに対して、基準座標系(Reference Frame)の再現性をもつ点の特徴点とする、SHOT特徴量⁸⁾⁹⁾が提案された。この手法は特徴点の周辺点からReference Frameを作成するため、特徴点1点同士の対応で高速な姿勢推定が可能である。しかしながら、ばら積みシーンでは、Reference Frameの作成時に複数の物体領域を含んでしまうため、位置姿勢パラメータを正しく求めることができない。また、特徴点となりうる形状の増加に伴い、処理コストの増加も問題となる。

以上のように、従来の3次元物体認識手法は、信頼性や高速性に問題があった。そこで本研究では、処理効率の観点から、点群データを利用した特徴点マッチングをベースに、認識の高信頼性と高速性を両立する物体認識システムの構築を目的とする。

本研究で扱う物体モデルは3次元座標を持つ点群モデルであり、これに注目点周辺の局所形状を表現する属性値を付加した特徴点から構成されている。属性値は各特徴点まわりの法線方向分布ヒストグラムで記述され、ヒストグラム同士の相違度を計算することで、各特徴点の独自性を算出し、独自性の高いものから優先的に照合に用いる。このような独自性の高い少数の特徴点のみを入力距離画像との照合に用いることで誤照合の低減と処理の高速化を図る。

本手法では、物体モデルの1点の特徴量を、その周辺の点群情報を用いて表現することにより、入力データに生じた誤差を許容でき、安定した特徴記述を可能にしている。また、特徴量を法線方向分布のヒストグラムとして記述することにより、物体の位置姿勢変化の影響を受けにくくなるという利点がある。

以下、第2章では本研究における基本アイデアと特徴点の独自性について説明し、第3章では提案アルゴリズムの概要と特徴点の独自性の算出方法について説明する。第4章では3次元物体モデルから選択された特徴点分布の分析とばら積みされた

* 原稿受付 平成25年4月1日

** 学生会員 中京大学大学院(愛知県名古屋市中昭和区八事本町101-2)

*** 中京大学大学院

† 正 会 員 中京大学大学院

Fig. 1 Distinctive feature points and non-distinctive feature points

多種類の対象物に対する認識性能評価について述べ、第 5 章にて本研究の成果をまとめる。

2. 基本アイデア

本研究では、物体認識の高信頼性と高速性を両立するために、以下に示す 2 つの基本アイデアを導入する。

1 つ目のアイデアは物体モデル各特徴点の独自性を数値化することである。注目特徴点周辺の局所形状が他の特徴点の周辺形状と類似していなければ独自性が高いと考えられる。物体モデル内で独自性の高い特徴点ほど、入力距離画像内において一意に対応する確率が高まるため、誤照合の低減が可能である。

2 つ目のアイデアは、照合に用いる特徴点を厳選することである。これには算出した各特徴点の独自性の値が大きい少数の特徴点を選択する方法を採る。マッチングに用いるデータ量の削減により、高速な物体認識が可能となる。

図 1 に物体モデルの独自性が高い特徴点と低い特徴点の例を示す。○印で示した特徴点は、直線的に大きな曲率が続く形状部分であり、物体モデルを特徴づけているため独自性が高い。また、×印で示した特徴点は、物体モデルの大部分を占めるなめらかな形状の一部であるため、独自性が低い。本研究では、このような独自性の高い少数の特徴点を認識に用いることで高信頼性と高速性を図る。

3. 提案アルゴリズム

3.1 概要

提案アルゴリズムの流れを図 2 に示す。提案アルゴリズムは、モデル分析モジュールと認識モジュールから構成される。本研究では、物体モデルとして 3 次元座標を持つ特徴点から構成される 3 次元点群いわゆるポイントクラウドデータを仮定する。

モデル分析モジュールでは、物体モデル各特徴点の属性値をその周辺の特徴点の法線ベクトルを用いたヒストグラムで記述する。そして、注目特徴点とそれ以外のすべての特徴点の法線方向分布ヒストグラム同士の相違度を比較することにより、各特徴点の独自性を算出する。最後に独自性の高い点から優先的に照合に用いる特徴点として抽出する。

認識モジュールでは、実際に 3 次元センサで撮影した距離画像を入力して、モデル分析モジュールと同様に特徴を記述し、抽出した独自性の高い特徴点の法線方向分布ヒストグラムと照合する。最後に物体モデルと入力距離画像の対応点の組を用いて対象物の位置姿勢を推定する。

3.2 法線方向分布ヒストグラム

提案手法における特徴記述表現である、法線方向分布ヒストグラムの作成方法を図 3 を用いて説明する。

まず、3 次元点群から構成される物体モデルの注目点 n を中

Fig. 2 Schematic block diagram of proposed algorithm

Fig. 3 Method for creating normal distribution histogram

心とする半径 r の球領域を設定し、 n と球領域内に含まれる点群 m_t の法線ベクトル $\mathbf{N}_n$ と $\mathbf{N}_{m_t}$ のなす角度 θ を求める。次に、縦軸を頻度値、横軸を角度 θ とする法線方向分布ヒストグラムを作成する。これは注目特徴点周辺の局所形状を表現している。注目特徴点の特徴量を球領域内に含まれる点群情報を用いて記述することにより、多少の誤差がデータに生じても許容することができ、安定した特徴記述が可能である。また、ヒストグラム形式にすることにより、モデルの姿勢変化によらず不変な頻度分布となり、同一の周辺形状を保持する点は同じ特徴量となる。物体モデルを構成するすべての点でこの処理を適用し、法線方向分布ヒストグラムを作成する。

3.3 独自性指標の算出

物体モデル点の独自性算出方法について説明する。まず、3.2 節で作成された物体モデルの注目特徴点と注目特徴点以外の法線方向分布ヒストグラム同士の相違度 B を Bhattacharyya 係数により算出する。

$$B(P, Q) = 1 - \sum_{u=1}^U \sqrt{P_u Q_u} \quad (1)$$

P, Q は法線方向分布ヒストグラム、 U はヒストグラムのビン数、 u は注目ビンである。相違度 B は 0 から 1 までの値で算出され、1 に近いほどヒストグラム同士の相関が低いことを表現している。次に、式 (2) より注目特徴点の独自性指標 S_n を算出する。

$$S_n = \frac{1}{T} \sum_{t=1}^T B(p_n, q_t) \quad (2)$$

ここで、 p, q は法線方向分布ヒストグラム、 T はモデル点数、 n は注目特徴点、 t はその他の物体モデル点である。独自性指標 S_n は 0 から 1 までの値を算出し、1 に近いほど注目特徴点の独自性が高いことを示している。本研究では、この独自性指標をもとに物体モデル各特徴点の独自性を算出し、認識に用いる

Fig. 4 Geometric consistency of an object model and the range image

Fig. 5 Matching by the voting process

特徴点を抽出する。

3.4 特徴点マッチングを用いた物体の位置姿勢推定

特徴点を用いたマッチングによる物体モデルと入力距離画像の対応点抽出手法について説明する。まず、物体モデルから抽出した独自性の高い特徴点と入力距離画像から算出した法線方向分布ヒストグラム同士の類似度を Bhattacharyya 係数により計算し、物体モデルと入力距離画像の対応点を決定する。次に、図 4 に示すように物体モデル間の特徴点 m_1, m_2 、入力距離画像間の点 s_1, s_2 をそれぞれグルーピングする。そして、式 (3) を満たすような物体モデルと入力距離画像の対応点組を用いて、物体モデルの位置姿勢パラメータを算出する。

$$(|d_{s_1, s_2} - d_{m_1, m_2}| < th_d) \wedge (|\theta_{s_1} - \theta_{m_1}| < th_t) \wedge (|\theta_{s_2} - \theta_{m_2}| < th_t) \quad (3)$$

$d_{m_1, m_2}, d_{s_1, s_2}$ は 2 点間のユークリッド距離、 $\theta_{m_1}, \theta_{m_2}, \theta_{s_1}, \theta_{s_2}$ は 2 点間の距離と法線ベクトルのなす角度、 th_d, th_t はそれぞれ 2 点間の距離と角度のしきい値である。

あらかじめ物体モデルの注目点 m_1 には物体モデル重心へ伸びるベクトルを関連づけておき、算出された位置姿勢パラメータを適用する。そして入力距離画像に投票空間を設定し、ベクトルの先端の座標に姿勢パラメータを投票する。図 5 に投票処理によるマッチングの模式図を示す。

多くの投票がある座標は、多くの対応点によって支持されたことを表すため、対象物が存在する可能性が高い候補とみなすことができる。しきい値以上の投票数のある位置姿勢パラメータ候補を用いて、式 (4)(5) により姿勢変換した物体モデルと入力距離画像の整合性を確認し、最終的に最も整合性の高い位置姿勢パラメータを決定する。

$$R_l = \frac{1}{T} \sum_{t=1}^T D(m_t) \quad (4)$$

Fig. 6 Calculation result of distinctiveness

$$D(m_t) = \begin{cases} c & \text{if } f(m_t(x), m_t(y))=0 \\ |m_t(z) - f(i, j)| & \text{otherwise} \end{cases} \quad (5)$$

R_l は l 番目の物体モデルと入力距離画像の整合性であり、この値が低いほど整合性が高くなる。 m_t は T 点で構成される物体モデルの t 番目のデータ点を示す。 $D(m_t)$ は物体モデル点と対応する入力距離画像との整合性を示している。関数 f は距離画像中の画素値を示し、 $m_t(x), m_t(y), m_t(z)$ は物体モデル上の 3 次元座標を表す。上段の条件式は姿勢変換した物体モデルに対応する入力距離画像の画素がデータ欠落している場合の処理である。なおペナルティ c は実験的に決定する。

4. 実験結果と考察

4.1 独自性分布の評価実験と特徴点選択

3 次元点群データから構成される物体モデルを用いて各特徴点の独自性を算出した結果を示す。図 6(a) に示すような多面体の物体モデルを用いて、モデル各特徴点の独自性を算出した。図 6(b) にモデル各特徴点の独自性算出結果を表示し、図 6(c) に横軸を独自性指標、縦軸を頻度値とした独自性分布を示す。今回実験に用いた面が直交する物体モデルでは、法線方向分布ヒストグラムが 0 度や 90 度付近に偏る傾向がある。そのため、微小な角度の差も検出できるようにヒストグラムの分解能を 2 度と細かく設定した。この物体モデルは 1.0mm ピッチで作成され、特徴記述する球領域内に一つの頂点のみ含まれるように半径 r を 5.0mm とした。

多面体モデルにおいて、平面付近より稜線付近、稜線付近より頂点付近の独自性が高く算出されたことを確認した。また、合計 14 点ある頂点の中でも、凹部分の点を中心とする形状は 2 点のみであり、多面体モデルを特徴づけている。そのため、凹部分の頂点の独自性が最も高く算出された。一般的に頂点の独自性が高いのは明らかであるが、提案手法を用いることによって、特徴点周りの局所形状が他の特徴点周りの局所形状と類似していない特徴点から順に高い独自性を算出していることが確認できた。

提案手法とランダムに特徴点を選択する手法 (ランダム法)、曲率の大きな点から順に特徴点を選択する手法 (曲率法) を用いて特徴点の分布を比較した。図 7(a), (b) は実験に用いた物体モ

Fig. 7 Calculation result of distinctiveness by the industrial parts

Fig. 8 Distribution of the selected distinctive feature points

デル外観とモデル各特徴点の独自性を算出した結果である。この物体モデルはレンジセンサで撮影した 0.4mm ピッチのモデルであり、算出した法線には誤差が含まれている可能性があるため、ヒストグラムの分解能を 5 度と粗く設定した。局所領域を表現するため半径 r は 2.0mm とした。図 8(a), (b), (c) にランダム法と曲率法および提案手法によって選択された特徴点の分布を示す。物体モデルからそれぞれ 30 点の特徴点を選択した。

ランダム法によって選択された特徴点は、物体モデルの大部分を占める、なめらかな形状部分の特徴点も選択されていることが確認できた。また、曲率法によって選択された特徴点は、曲率の大きさのみを考慮しているため、曲率の大きな窪んだ形状部分の点が特徴点として選択されていることが確認できた。曲率の大きな窪んだ形状部分の特徴点は対象物が姿勢変化した場合に、隠れてしまいやすく正しく照合できないと考えられる。しかしながら、提案手法によって選択された特徴点は、モデルを特徴づける大きな曲率部が直線的に続く部分の点や中心部の平面的な形状部分の点などが選択されているため、対象物が姿勢変動した場合にも正しく照合できると考えられる。

4.2 認識性能評価実験

4.2.1 ばら積み物体に対する認識結果

提案手法の有効性を確認するために、ばら積みされた対象物に対する認識性能を評価した。Spin Image 法¹⁾、Point Cloud Library¹⁰⁾ (PCL) の認識モジュールに登録されている Correspondence Grouping 法¹¹⁾、ランダム法、曲率法を用いて認識成功率と処理時間を比較した。本研究では処理の高速化を図るため、PCL に登録されている Voxel Grid Filter を用いて 3 次元点群モデルをダウンサンプリングしたものから独自性の高い特徴点を抽出する。ダウンサンプリングした物体モデルと点数、各物体がばら積みされたシーンを撮影した入力距離画像の例を図 9 に示す。それぞれ、凹凸の多い物体 A、同一形状の繰り返しが存在する物体 B、凹凸の少ない物体 C、曲率の大きな部分と小さな部分を含む物体 D である。実験には、各物体約 130 枚の距離画像を用いた。

表 1 に認識成功率 P_r [%] と処理時間 T [sec], 特徴点数 N [point] を示す。なお本研究では、姿勢推定したパラメータを物体モデルに適用し、入力した距離画像と重畳させた時の位置

Fig. 9 Object model and range image database

Table 1 Recognition success rate and processing time

		ObjectA	ObjectB	ObjectC	ObjectD	Mean
Spin Image ¹⁾	P_r	49.6	50.8	35.2	70.6	51.6
	T	24.95	55.34	31.21	20.68	33.05
Correspondence Grouping ¹⁰⁾¹¹⁾	P_r	74.4	96.2	65.6	84.9	80.3
	T	45.41	52.85	29.26	26.10	38.41
Random method	N	70	50	30	30	-
	P_r	89.1	96.9	85.2	69.8	85.3
	T	5.87	4.77	2.06	1.74	3.61
Curvature method	N	70	50	30	30	-
	P_r	86.8	76.9	18.0	54.8	59.1
	T	6.01	4.94	2.21	1.97	3.78
Proposed method	N	70	50	30	30	-
	P_r	89.1	97.7	94.5	93.7	93.8
	T	5.83	4.76	2.05	1.76	3.60

合わせ誤差が 1.5mm 以内の場合に認識成功とした。図 10(a), (c), (e), (g) は入力シーンの画像外観であり、図 10(b), (d), (f), (h) に認識結果の位置姿勢パラメータを適用した物体モデルを入力シーンに重畳した結果を示す。実験は CPU : Intel®CORE™i7-3.40GHz, システムメモリ : 8GB で構成されるシステムでおこなった。なお、投票処理におけるパラメータは予備実験により経験的に $th_d=1.2\text{mm}$, $th_t=5$ 度, ペナルティ $c=30$, 投票数のしきい値は 3 とした。

4.2.2 従来手法との認識性能比較と考察

提案手法は、従来の Spin Image 法に対して、物体モデルの 1% から 5% 程度のごく少数の特徴点数で、4 種類の A から D の物体に対する平均認識成功率が、51.6% から 93.8% に向上したことを確認した。処理時間に関しては約 9 倍以上の高速化となった。Spin Image 法はモデル全点を用いて照合しているのに対し、提案手法は独自性の高い、ごく少数の厳選された特徴点のみを照合に用いていることがその要因である。また、Spin Image 法は注目点の法線ベクトル方向と、それに直交する軸を基準とした 2 次元空間にモデル点の相対的な位置を投票した画像を照合に用いる手法である。そのため注目点から遠い点ほど法線ベクトルの揺らぎに敏感であり、安定した特徴記述ができず、認識成功率が低かった。提案手法も法線ベクトルを使用しているが、点群を利用した法線ベクトル同士のなす角度を軸としたヒスト

Fig. 10 Example recognition results

グラムで表現しているため、法線ベクトルに多少の揺らぎがあったとしても、安定した特徴記述ができた。また、Spin Image 法は入力距離画像からランダムに点を選択して、Spin Image を作成し物体モデル全点と照合する。ランダムに選択した点が物体モデル内の同じ点と対応していなければ正しく照合することができない。そのことも認識成功率が低かった理由の一つである。Correspondence Grouping 法の認識成功率が低かった理由としては、モデル点をダウンサンプルした時の点を認識に用いていることが原因である。また、正しい姿勢パラメータを算出するには半径の大きな Reference Frame を作成する必要がある、その処理に時間がかかっている。ランダム法が提案手法より認識成功率が低かった理由としては、特徴点をランダムに選択しているため物体モデル内で同じような形状部分を保持する特徴点を選択されており、入力距離画像との照合の際に誤照合が増加したためである。曲率法に関しては、曲率の大きな点のみを特徴点として選択しているため、入力距離画像中に曲率の大きな点が多く存在する場合に認識成功率が低下してしまった。また曲率法で選択した特徴点は窪んだ曲率部の点も選択しているため、入力距離画像中の対象物の姿勢によっては対応点が隠れてしまう。そのため認識成功率が低かった。

以上より、Spin Image 法、PCL の Correspondence Grouping 法、ランダム法、曲率法に対して、提案手法は平均認識成功率が向上し、処理時間に関しても同等以上であることを確認した。

なお、図 11 に認識に失敗した例を示す。本研究では、姿勢パラメータを適用した物体モデルと入力距離画像との誤差を計算し、一番小さい誤差の位置姿勢パラメータを認識結果としている。図 11 のような平面的な物体の場合では、姿勢変換したモデルが 2 つの対象物上に重畳された場合でも、位置合わせ誤差が

Fig. 11 Recognition failure cases

小さくなり認識に失敗した。この例において、抽出した特徴点数 70 点をデフォルトとして、独自性の高い点から順に 30 点増やした方法と独自性の低い点から順に 30 点増やした方法で認識比較したところ、独自性の高い点を増やした方法は、認識に成功したが、独自性の低い点を増やした方法は、認識に失敗した。この結果より、独自性の高い特徴点の中に認識に有効な特徴点があり、その特徴点を用いることにより認識に成功したと考えられる。最適な特徴点数の自動決定と投票処理を考慮した特徴点の選択は今後の課題である。

5. 結 論

本研究では、事前に物体モデルを分析し、照合に有効な独自性の高いごく少数の特徴点を用いることにより、高信頼性と高速性を両立する物体認識システムを提案した。従来手法である Spin Image 法に対し、認識成功率が 51.6% から 93.8% に向上し、約 9 倍以上の処理の高速化を実現した。また、PCL の Correspondence Grouping 法やランダム法、曲率法に対しても認識成功率が向上し、提案手法によって抽出した独自性の高い特徴点が認識に有効であることを確認した。今後は、さらなる処理の高速化、各種パラメータの最適化、提案手法を実装したピンピッキングシステムを構築する予定である。

謝 辞

本研究の一部は中京大学特定研究助成の成果である。

参 考 文 献

- 1) A. E. Johnson and M. Hebert: Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes, Trans. IEEE Pattern Analysis and Machine Intelligence, **21**, 5, (1999) 433.
- 2) 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), **J80-D-II**, 5, (1997) 1105.
- 3) B. Steder, R. B. Rusu, K. Konolige and W. Burgard: Point Feature Extraction on 3D Range Scans Taking into Account Object Boundaries, IEEE International Conference on Robotics and Automation, (2011) 2601.
- 4) 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, 電学論 (C), **124**, 3, (2004) 637.
- 5) 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), **127**, 4, (2007) 615.
- 6) C. Dorai and A. K. Jain: COSMOS-A Representation Scheme for 3D Free-form Objects, Trans. IEEE Pattern Analysis and Machine Intelligence, **19**, 10, (1997) 1115.
- 7) H. Chen and B. Bhanu: 3D Free-form Object Recognition in Range Images Using Local Surface Patches, Pattern Recognition Letters, **28**, 10, (2007) 1252.
- 8) F. Tombari, S. Salti, and L. D. Stefano: Unique Signatures of Histograms for Local Surface Description, European Conference on Computer Vision, **6313**, (2010) 356.
- 9) F. Tombari, S. Salti, and L. D. Stefano: A Combined Texture-Shape Descriptor for Enhanced 3D Feature Matching, IEEE International Conference on Image Processing, (2011) 809.
- 10) R. B. Rusu and S. Cousins: 3D is here: Point Cloud Library (PCL), IEEE International Conference on Robotics and Automation, (2011) 1.
- 11) F. Tombari and L. D. Stefano: Object Recognition on 3D Scenes with Occlusions and Clutter by Hough Voting, IEEE Proc. on 4th Pacific-Rim Symposium on Image and Video Technology, (2010) 349.