

距離情報を用いた 3 次元物体認識

水谷了† 中村友昭† Muhammad Attamimi† 長井隆行†
船越孝太郎‡ 中野幹生‡

†電気通信大学大学院電気通信学研究科 ‡(株)ホンダ・リサーチ・インスティテュート・ジャパン

1 はじめに

近年、人間の生活する実環境において人間を支援するロボットへの期待が高まっている。そのような環境でロボットが動作する際に重要となるのが、そのロボットの安全性・柔軟性である。これらを実現するためには、ロボットが自身の置かれた環境を把握する必要がある。得られた情報から環境を理解するために、周りに存在する物体を正確に認識することが重要となる。

家庭環境のような実環境で働くロボットに搭載する物体認識アルゴリズムには、環境の変化に対するロバスト性が要求される。しかし、一般的な色やテクスチャを特徴量とした認識では、照明条件の大きな変化に対応することが難しい。そこで本稿では、赤外線 TOF カメラによって取得する 3 次元情報をベースとした認識手法を提案する。提案手法では、奥行き情報をヒストグラムとする Histogram of Depth(HOD) を特徴量とする。また、従来 3 次元モデルの検索などに利用される Shape Distribution(SD) を併用することでより認識精度を向上させることを図る。これらの特徴量は、スケールや回転に不変な性質を持っており、特に HOD は高速に計算が可能である。さらに、提案する物体認識システムでは、赤外線 TOF(Time of flight) カメラと CCD カメラを組み合わせているため、色情報を同時に得ることができ、これにより、色やテクスチャ情報を組み合わせた物体認識が可能となる。

提案手法は、ヒストグラム間の距離をベースとしているため、非常に高速に認識を行うことができる。また、ロボットへの適用を想定しているため、ロボットが動くことで複数視点の情報を容易に得ることができる。こうした特性を利用し、複数フレームの情報を統合することで、さらに認識精度を向上させることが可能となる利点もある。

関連研究としては、3 次元情報を用いた物体認識手法 [2] などが挙げられるが、回転に対する不変性がなく、類似度計算でこの点を考慮する必要がある。また、色情報を組み込んでいるため、照明条件に対する依存性を避けることができない。

2 提案手法

2.1 センサの構成

本稿では、視覚センサを搭載したロボットが家庭環境で物体認識することを想定する。視覚センサとしては、図 1 に示すような赤外線 TOF カメラと左右に設置された CCD カメラを用いる [1]。これらのセンサ間でキャリブレーションを行い、3 次元情報と色情報の対応関係を求めることで、それぞれの情報を統合することができる。また、各カメラの設置位置のずれによって生じるオクルージョンの問題は、左右 2 台の CCD カメ

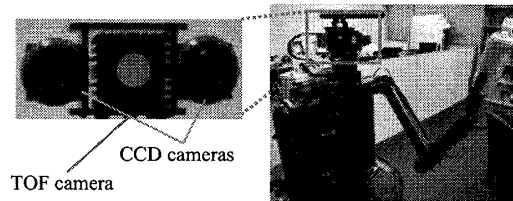


図 1 センサの構成 (左) とロボット (右)。

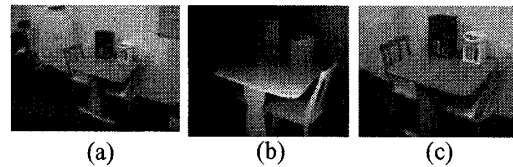


図 2 カメラキャリブレーション, (a) カラー画像, (b) 距離画像, (c) キャリブレーション結果

ラを使い欠落した情報を補い合うことで低減することができる。これにより、色情報と高精度な 3 次元情報を同時かつリアルタイムで得ることが可能となる。視覚センサによって得られる情報の例を図 2 に示す。

2.2 提案手法の流れ

物体を認識するために、ロボットは事前に物体を学習する必要がある。ここでは、人がロボットに物体を様々な方向から見せることで物体の情報を学習させる。物体の切り出しは、文献 [1] で提案した手法を用いる。個々の物体について、様々な方向から複数フレーム検出を行い、物体領域中の特徴量を計算し蓄積することで、データベースを作成する。認識時には、学習時と同様に人がロボットに物体を提示することで物体の切り出しを行う。もしくは、認識対象が机などの平面上に置いてある場合は、平面検出を行うことで容易に物体のみを切り出すことができる。また、ロボット自身が候補物体を把持し観測することも可能である。その後、対象物体の特徴量を抽出し学習済みの物体と照合することで認識を行う。特徴量と認識については、以下で説明する。

2.3 特徴量

ここでの特徴量に対する要件は、スケールと回転、シフトに対する不変性である。回転・シフト不変性に関しては、ヒストグラムをベースとした特徴量を用いることで容易に実現可能である。また、スケール不変性に関しても、3 次元情報による正規化により容易に保証できる。視点変化に対する不変性は、特徴量のレベルで実現することが困難であるため、学習時に複数の視点から物体を観測しそれら全ての情報との照合によって解決する。以下ではまず、3 次元情報として二つの異なる特徴量を検討する。

2.3.1 奥行きヒストグラム

3 次元情報の特徴量の一つとして、物体の奥行きヒストグラム HOD を提案する。これは、カメラから物体の各画素までの距離をヒストグラムとするものであるが、

Three-Dimensional Object Recognition Using Depth Information

†Akira MIZUTANI, †Tomoaki NAKAMURA,

†Muhammad ATTAMIMI, †Takayuki NAGAI,

‡Kotaro FUNAKOSHI, and ‡Mikio NAKANO

†The University of Electro-Communications

‡Honda Research Institute Japan Co., Ltd.

物体までの距離によって値が変化してしまうため、物体領域までの奥行きを平均値を基準として正規化を行う。赤外線 TOF カメラの距離計測の精度を考慮して、ピンのサイズを 5mm とする。HOD は非常に高速に計算でき、スケール、回転、シフトに関して不変性を有する。ただし、形状を識別するための表現力は必ずしも十分ではないため、もう一つの情報と組み合わせることを考える。

2.3.2 Shape Distribution

二つ目の特徴量として、SD を用いる。これは、物体領域中の頂点間の距離を計算し、それヒストグラムをとするものである。SD は、従来 3 次元モデル同士の類似度を計るために提案されたものであるが、3 次元モデルでは視点に依存しない特徴であるのに対し、ここでの情報は 2.5 次元であるため視点依存となる。また、提案システムにおいて 3 次元情報は点群として取得されるため、3 次元メッシュからランダムに点群を生成する必要はない。SD も HOD と同様にスケール、回転、シフトに関して不変性を有する。ただし、SD は HOD とは異なり、物体の大きさの情報を持っているため、これらを組み合わせることで認識精度が向上することが期待できる。

2.3.3 色・テクスチャ情報

3 次元情報だけでは、形状が全く同じ物体を識別することはできない。照明条件が良い場合には、色やテクスチャ情報を用いることで認識精度を向上することができる。ここでは、色情報として HSV 表色系の H(色相) と S(彩度) のヒストグラムを用いる。また、局所特徴である SIFT(Scale Invariant Feature Transform) 記述子をベクトル量子化し、その頻度ヒストグラムをテクスチャ情報として用いる。

2.4 物体の認識

学習フェーズで作成されたデータベースと、新たに入力された物体の情報から得られた特徴量を比較することで認識を行う。各特徴量に関してヒストグラム間の距離を計算し、尤度を算出することで確率として物体らしさを表す。全ての特徴量について尤度の計算を行い、それらを掛け合わせることで物体の確率とするが、全ての特徴量を同等に扱うのではなく、認識対象の特徴によって、用いる特徴量は選択されるべきである。そこで、認識時に入力されたそれぞれの特徴から各特徴の重みを決定し、使い分けて認識を行う。また、認識時の照明条件によっては色情報やテクスチャ情報の信頼性が低くなることが考えられる。ここでは、特徴量の全てのデータに対する尤度中の最大値をその特徴量の信頼度と考え、最大値が低い場合は、その特徴量を利用しない。

本稿では、ロボットによる物体認識を考えているため、認識結果の信頼度が低い場合は、視点を変えて複数フレーム物体を観測することも可能である。特に 3 次元情報は視点を変えることにより識別により有用な情報を取得できる可能性がある。このように取得した複数フレームの情報を同様に統合することでより認識精度を向上させることが可能である。

3 実験

実験は、16 種類の物体を用いて行った。この中には、形状が様々なぬいぐるみ、形状が全く一緒で表面の色・テクスチャが異なるコップ、無色でテクスチャのない食器類などが含まれている。学習フェーズでは、ユーザがロボットに対して様々な角度で物体を教示し、ロ



図 3 照明条件の異なる実験環境

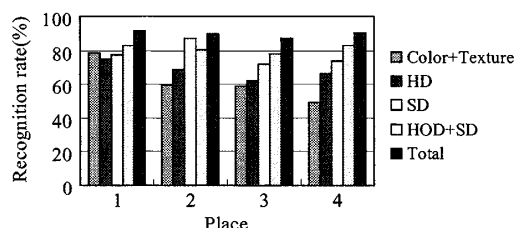


図 4 各場所における認識率の比較

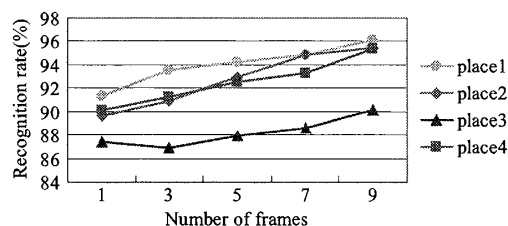


図 5 使用フレーム数による認識率の変化

ボットは各物体 50 フレーム分のデータからデータベースを作成した。認識フェーズでは、照明条件の異なる 4ヶ所(図 3)で 1ヶ所につき 50 フレーム、計 200 フレームの認識を各物体ごとに行った。

実験結果を図 4 に示す。縦軸は認識率の平均、横軸は実験を行った場所を表す。場所番号は図 3 の左から順に 1,2,3,4 と対応する。また、認識については 5 つの手法で行った。色・テクスチャ情報のみを用いた認識では、色味がなくテクスチャの少ない物体に対しての認識率が低く、また照明変化による変動が大きい。距離情報に関しては HOD と SD の 2 つの手法を組み合わせることで認識率が向上した。この場合、照明変化による影響は受けないが、人が提示する際の物体の見せ方による認識率の変動があった。また距離情報のみでは形と同じ物体は認識できなかったが、色・テクスチャ・距離情報の全てを統合することで、それらの物体も区別できるようになり、より安定した認識を行うことができた。ただし、距離情報を用いた場合と同様に物体を見る角度によって誤認識することがあった。これはデータベースを作成する際に、様々な角度から均一にデータを蓄えることができなかったことが原因として考えられる。次に、異なる角度での情報を複数取得し、それらを用いて認識を行うことができる。より大規模なデータセットでの手法の評価が今後の課題である。

4 まとめ

本稿では、3 次元情報を利用した物体認識手法を提案した。3 次元情報を利用することで、照明条件にロバストな物体認識が可能であることを示した。また、最終的には色情報やテクスチャ情報と組み合わせることで、高精度に物体認識を行うことができる。より大規模なデータセットでの手法の評価が今後の課題である。

参考文献

- [1] 水谷ほか, “赤外線 TOF カメラと CCD カメラのキャリブレーションによる 3 次元センサの実現と画像処理への応用”, 電気学会研究会資料 IM-09-60~74, pp.69-74, 2009
- [2] 原田ほか, “三次元環境地図からの物体探索タスク応用を目指したカラー立体高次局所自己相関特徴の開発”, 日本ロボット学会誌, 27(7), pp.749-758, 2009