

協働ロボットによる作業者補助を実現する操作システムの開発（第4報）

—ToF カメラを活用したオブジェクト認識技術の開発—
坂東直行*、安部貴大*、渡辺博己**

Development of support system for the operators using collaborative robots (IV) - Development of object recognition technology utilizing ToF camera - BANDO Naoyuki*, ABE Takahiro* and WATANABE Hiroki**

本研究では、オブジェクトを3次元で認識するため ToF カメラを採用することで、深度イメージから 3D 点群情報を取得し、オブジェクト認識によってワークの位置・姿勢を推定できることを示した。その結果、ToF カメラから取得する深度イメージ画像から空間的にゆがみの少ない 3D 点群情報を取得できること、点群情報を使ったオブジェクト認識により、ワークの位置・姿勢を推定でき、対象シーン点群を絞ることで高精度に姿勢を推定できること、本技術はピンピッキングに活用できることを確認した。

1 はじめに

協働ロボットは、PC 等の外部制御装置（以下、外部コントローラ）で動作させることができる。これは、基本的に動作中は外部コントローラからの操作を受け付けない産業用ロボットに対する協働ロボットの優れた特徴で、これにより画一的な動作しかできない産業用ロボットに対し、協働ロボットは高い柔軟性をもった動作ができる。この特徴を活かせば、ロボットの新しい活用方法を検討できるようになる。

例えば、ロボット周辺の環境情報をセンシングし、そこからロボット動作に必要な情報を抽出して、動作プランニングに反映させる機能を外部コントローラが備えれば、産業用ロボットと比較して環境変化に強靱なロボットシステムを実現できる。

そこで本研究ではこれまでに、作業エリア上部に固定した Web カメラを使い、ワークの画像をリアルタイムに確認しながら、GUI 操作でティーチングするシステムを構築した¹⁾。しかし対象を真上から見下ろした 2D 画像を用いるため、ワークの高さ情報が予め既知であることを前提としていること、画面の端になるほど位置の歪みが大きくなることが課題として残った。

これらの課題解決のため、複数の視点からカメラで撮影したイメージ画像をもとに、SfM (Structure from Motion) 技術を用いてワークの 3D 点群情報を取得し、点群処理技術によってワークの位置・姿勢を推定する技術を開発した²⁾。これにより、高さ情報・位置の歪みの問題は解決したが、ワークが存在するシーンの点群情報を取得するのに数十秒程度の時間が必要で、リアルタイム性に乏しいという課題が残された。

そこで本研究では、シーンの点群情報取得に後述する ToF カメラを採用することで課題解決を試みたので報告する。

2 ToF センサを用いたロボットシステム

図1に本研究で提案するロボットシステムの処理過程を示す。

本システムでは、2 台の協働ロボットと ToF カメラによって構成される。ToF カメラはロボットアームのエンドエフェクタに取り付けられ、ワークの状態に応じて撮影位置・姿勢を変更できる。ロボットの作業対象は ToF カメラにより深度画像として取得され、3D 点群情報に変換される。この点群と予め用意してあるワーク点群をマッチングすることでオブジェクト（ワーク）が認識され、同時にワークの位置・姿勢が推定される。これによりワークの基本情報（隅、辺、重心等）の座標が求まる。これら基本情報をもとにロボットを動作させることで、ワークの状態に応じたロボット動作が可能になる。

本報では、本システムにおいて主要部となる ToF カメラ情報を用いたオブジェクト認識技術について述べる。

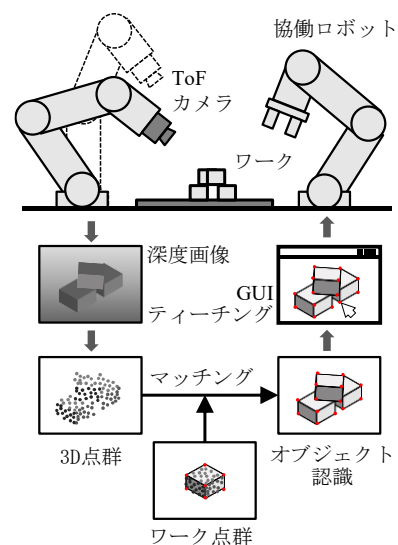


図1 提案するロボットシステム

* 生産システム部

** 情報技術部

3 ToF カメラデータの処理

3.1 ToF カメラ

本研究では、オブジェクトを含むロボット周辺の 3D データを取得するのに、Analog Devices 社の EVAL-ADTF3175D を用いた。これは、PC と接続することで、動作を PC で制御することができる ToF カメラである。外観図を図 2 に、諸元を表 1 に示す。ToF とは Time of Flight の略称で、光を照射してから返ってくるまでの時間を計測し、光の速さと往復にかかった時間から奥行きを求める計測手法である。なお、奥行き情報を取得するセンサには、この他にステレオカメラ方式がある。これは 2 つの光学系における視差から奥行きを計算するものだが、光の反射が理想的であることを前提にしているため、金属部品の計測が難しく、一方のカメラで死角になる領域はデータが得られず、測定データのゆらぎ・ノイズが大きいなど、不都合な点が多い。そこで、それらの問題に対し比較的優れている ToF カメラを本研究では採用した。

3.2 データの校正

ToF カメラで図 3 および図 4 に示すシーンを撮影した結果を図 5 に示す。このように ToF カメラでは深度情報をイメージ画像として取得でき、各画素の値は被写体までの距離を表す。これを 3D 点群情報に置き換えてプロットしたものを図 6 に示す。

ここから、イメージ情報を 3D 点群としてプロットすると、イメージ画像の中心から端に向かうほど空間的な歪みが大きくなることが分かる。そこでピンホールカメラモデルを用いて、深度情報および上下左右方向の距離情報を校正した。結果を図 7 に示す。ここから、適切に



図 2 ToF カメラ外観

表 1 ToF カメラ諸元

項目	値
型式	Analog Devices EVAL-ADTF3175
サイズ	66mm×68mm×59mm
解像度	1024×1024 pixel
画角	75° ×75°
精度	±5mm
フレームレート	10fps (1024×1024pix)

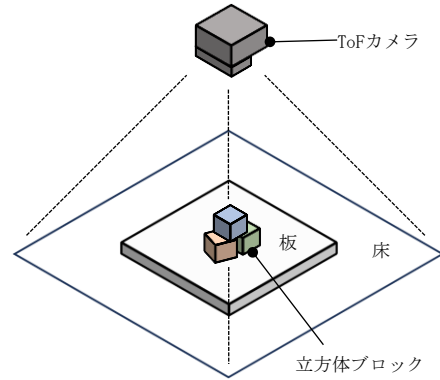


図 3 立方体ブロックの撮影

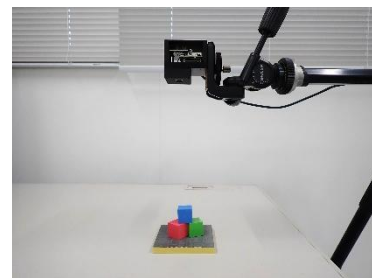


図 4 撮影風景

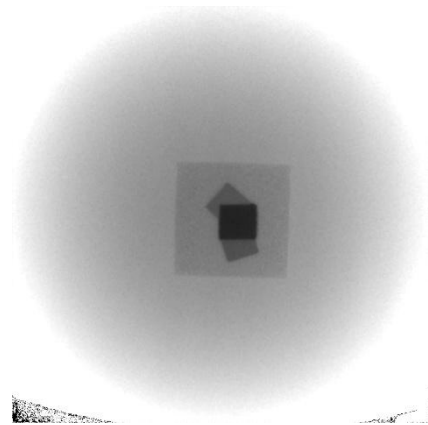


図 5 撮影結果

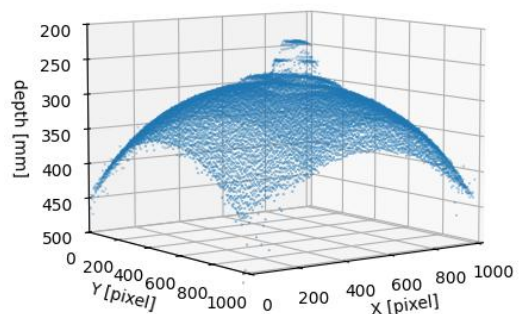


図 6 3D プロット結果

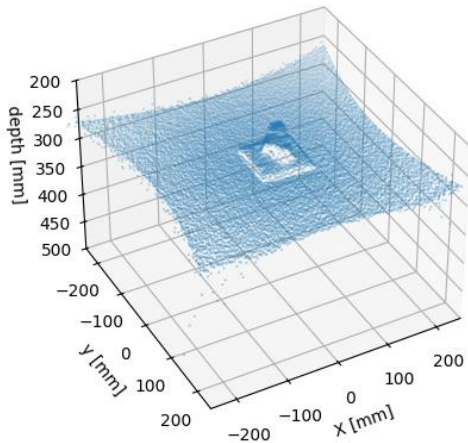


図7 校正した点群の3Dプロット結果

校正することで深度イメージ画像から空間的に歪みの少ない3D点群情報を取得できることが分かる。

なお、校正済みシーン点群を取得するのに要した時間はIntel Corei7 (2.90GHz) において1024×1024pixelの深度イメージ画像に対し14ミリ秒であり、SfMと比較し短い処理時間でシーン点群が得られた。

4 3D シーン点群からのオブジェクト認識

次に、3D シーン点群からオブジェクト認識を行った。

ここでは、図7に示すシーン点群から、一辺25mmの立方体オブジェクトの認識を試みた。

まず、立方体オブジェクトのCADデータを作成し、そこからオブジェクト表面の点群データを生成した。

次に、立方体オブジェクト点群および3D シーン点群に対し、ダウンサンプリングを行い、特徴点検出を行った。その後、特徴点毎に特徴量を求め、2つの点群から得た特徴量のノルムを計算して最も小さい特徴点のペアを求める対応点検索を行った。そしてRANSAC

(Random Sample Consensus) によって、立方体オブジェクト点群を3D シーン点群にマッチングさせる位置・姿勢変換行列を推定した。なお、これらの点群処理にはOpen3Dを用いた³⁾。

こうして得られた姿勢変換行列によって、立方体オブジェクト点群をシーン点群に位置合わせした結果を図8に示す。この結果から分かるように、シーン点群全体を対象にオブジェクト認識を行った場合、位置・姿勢変換行列の推定精度が十分ではなかった。そこで、シーン点群において注目領域 (ROI : Region Of Interest) を立方体オブジェクトが存在する範囲に設定し、ROI内の点群を使ってオブジェクト認識した結果を図9に示す。ROIを設定した場合、オブジェクト認識によって位置・姿勢がほぼ推定できている。ここから、点群情報を使ったオブジェクト認識において、対象シーン点群を絞ることで

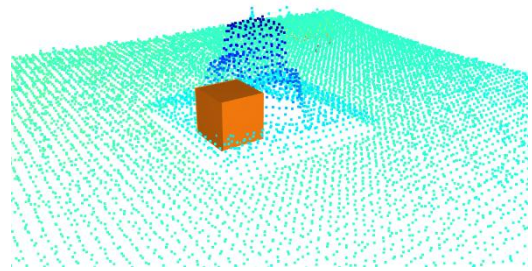


図8 シーン点群全体に位置合わせした結果

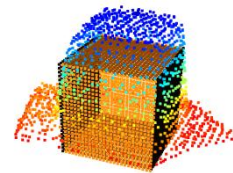


図9 ROI内のシーン点群に位置合わせした結果

ワークの位置・姿勢を推定できることがわかる。

5 ビンピッキング実験

本研究で開発したシステムを活用して実際にロボットによるバラ積みピッキング (ビンピッキング) 実験を行った。

実験では、一辺25mmの立方体をワークに見立て、複数のワークが無造作に山積みされた状況において、ワークを認識してピッキングすることを試みた。

なお、本システムにおいて、協働ロボットおよびToFカメラは各々独自の座標系を持つ。ToFカメラで取得するデータは、ToFカメラに設定された座標系における値であるため、静止した物体を撮影した場合であっても、ロボットの動作によりToFカメラ位置が変われば値が変わる。そこでToFカメラを取り付けたロボットの座標系を基準 (以下、基準座標系) とし、ロボットのエンドエフェクタの位置姿勢情報を用いてToFカメラの座標系を基準座標系に一致させた。また、他方のロボットの座標系も、基準座標系と一致するようにキャリブレーションを行った。これにより、協働ロボットおよびToFカメラを同一の座標系で扱うことができる。

システムは以下のように動作させた

1. ToFカメラをワーク上部に移動させワークを撮影する。
2. 取得した深度イメージを3Dシーン点群に変換する。

3. 基準座標系における X、Y 値（床面の幅と奥行きに相当）が予め設定した範囲内にある点群を抽出する。これによりワークが存在しないエリアの点群が除かれる。
 4. 基準座標系における Z 値（高さに相当）が予め設定した閾値以上のデータを抽出する。これにより背景情報である床の点群が除かれる。
 5. シーン点群の中から、基準座標系の Z 値（高さに相当）の最大値を求め、そこから既定の距離の深さまでの点群を抽出する。ここで既定の距離はオブジェクトに応じて設定するパラメータであり、ここでは立方体の対角距離（立方体の一辺の長さを a としたとき、 $\sqrt{3}a$ ）を設定した。これによりシーン点群はほぼワークに相当するものだけが残る。
 6. シーン点群に対してオブジェクト認識を行い、ワークの位置・姿勢を推定する。
 7. 推定したワークの位置から、ワーク中心を原点とし、ワークの傾きと連動するワーク座標系を設定する。
 8. 基準座標系の Z 軸となす角が最も小さくなるワーク座標系の軸を求め、その軸がワーク座標系の Z 軸となるように姿勢変換する。
 9. ワーク座標系の原点から Z 軸方向に既定の距離（ここでは 5 cm）離れた場所に、ワーク座標系の姿勢にあわせてロボットのグリッパーを移動する。この位置・姿勢を P とする。
 10. ロボットのグリッパー先端をワーク座標系原点の位置まで移動し、ワークをピックアップする。
 11. グリッパーを P に移動する。
- 以上により、ビンピッキングが可能になる。
結果を図 10 に示す。

この結果から、本システムはビンピッキングに対して有効であることを確認した。

なお、ワークの位置・姿勢を推定する点群処理においては、ダウンサンプリングサイズ、特徴量の計算範囲等、処理パラメータが存在し、これを対象に応じて調整する必要があり、本実験では試行錯誤で調整した。誤差を抑えつつ安定的に姿勢を推定するためのパラメータチューニング方法については検討が必要である。

また、今回の実験では、ワーク同士が密着すると、オブジェクト認識の精度が下がることも確認した。ビンピッキングの完遂確度を高めるには、ワーク同士に密着が認められた場合、ワークを適度に散らす動きをロボットに組み込むなど、更なる改良が必要である。

6 まとめ

本研究では、オブジェクトを 3 次元で認識するため ToF カメラを採用することで、深度イメージから 3D 点群情報を取得し、オブジェクト認識によってワークの位

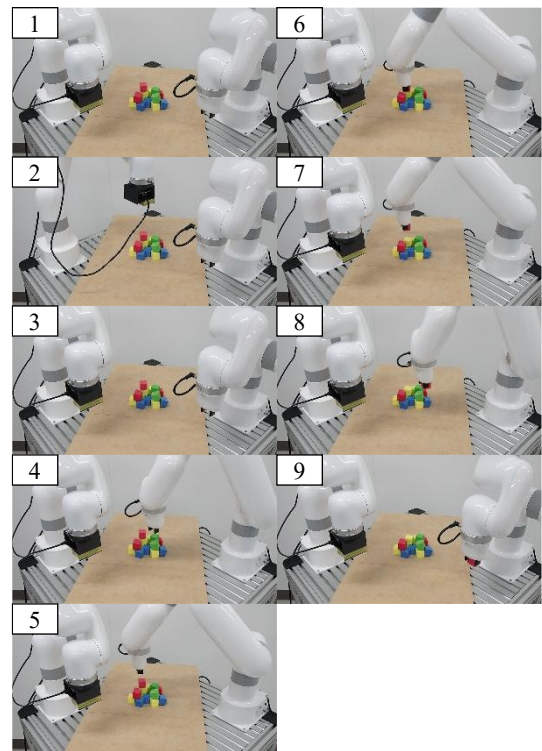


図 10 ビンピッキング実験結果

置・姿勢を推定できることを示した。結果を以下に示す。

- ・ ToF カメラから取得する深度イメージ画像から空間的に歪みの少ない 3D 点群情報を取得できる。
- ・ 点群情報を使ったオブジェクト認識により、ワークの位置・姿勢を推定でき、対象シーン点群を絞ることで高精度に姿勢を推定できる。
- ・ 本技術はビンピッキングに活用できる。一方でビンピッキングの完遂確度を高めるためには更なる工夫が必要である。

【参考文献】

- 1) 坂東ら,岐阜県産業技術総合センター研究報告 No.3, pp7-8,2022
- 2) 坂東ら,岐阜県産業技術総合センター研究報告 No.4, pp85-88,2023
- 3) 金崎ら,詳解 3 次元点群処理,講談社,2022