

欠損復元 AutoEncoder による欠損を含む物体の姿勢推定の検討

A Study on the Pose Estimation of an Occluded Object by a De-occluding AutoEncoder

立道 大樹¹ 川西 康友¹ 出口 大輔¹ 井手 一郎¹ 村瀬 洋¹ 安間 絢子²
 Hiroki Tatemichi Yasutomo Kawanishi Daisuke Deguchi Ichiro Ide Hiroshi Murase Ayako Anma

名古屋大学¹
 Nagoya University

トヨタ自動車株式会社²
 TOYOTA MOTOR CORPORATION

1 はじめに

近年、日常生活の支援のために家庭環境にロボットが導入されつつある。ロボットの基本動作として、物体の持ち運びが挙げられる。物体を持ち運ぶためには、対象物体の認識だけでなく、その姿勢を推定する必要がある。特に、机上など物体が密集して置かれている状況でも、安定して物体の姿勢を推定できることが望ましい。

物体の色や照明の影響を受けにくい観測をするため、ロボットには距離画像センサが搭載されることが多い。距離画像を用いる姿勢推定手法として、CNN に基づく回帰モデルを用いる手法 [1] がある。この手法は、物体全体が観測できていることを仮定しているため、物体が密集して存在する場合、他の物体に遮蔽されることで物体画像に欠損が生じ、姿勢推定が困難である。一方、欠損物体を対象とした手法 [2] では、欠損を復元する AutoEncoder を学習し、Encoder 部分から得られる特徴ベクトルを姿勢推定に用いるが、この手法は欠損による物体中心の位置ずれを考慮していないため、欠損が大きいほど精度が低い。

そこで本発表では、欠損を復元できるように学習した AutoEncoder を用い、欠損復元と位置ずれの修正をしつつ姿勢推定をする手法を提案する。

2 提案手法

学習段階ではまず、姿勢が既知の物体の全体が見える学習用距離画像とそれを欠損させたものを用意し、欠損を復元する AutoEncoder を学習する。次に、その AutoEncoder の Encoder 部を用いて、欠損させた学習用距離画像から特徴ベクトルを抽出する。最後に、それに対して主成分分析 (PCA) による次元削減を行う。

推定段階ではまず、姿勢が未知の物体の距離画像に対し、同様の特征抽出と次元削減を行う。そして、姿勢が既知の物体の特徴ベクトルのうち最も似たものを最近傍法により探索し、その姿勢を出力する。

図 1 に AutoEncoder による特徴抽出の処理手順を示す。特徴抽出は、まず AutoEncoder により欠損部分を復元した画像を生成する。次に、復元後の画像中の物体が中心にくるように位置合わせを行い、再度同じ AutoEncoder に入力する。このとき、Encoder 部分から得られる特徴ベクトルを姿勢推定に用いる。

3 実験

提案手法の有効性を確認するために、物体の姿勢推定精度の評価実験を行った。AutoEncoder を 1 回のみ使用し、得られる特徴ベクトルを用いて姿勢推定する手法

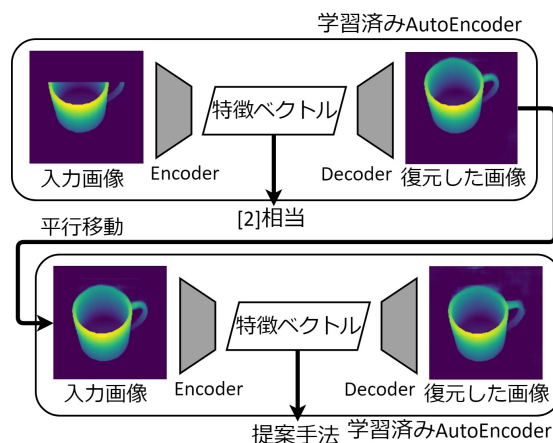


図 1 AutoEncoder による特徴抽出の処理手順

([2] 相当) と、AutoEncoder を 2 回用いる手法 (提案手法) とを比較した。

学習データには、ShapeNet[3] の 100 種類のマグカップを物体の鉛直軸まわりに回転させた距離画像計 3,600 枚、評価データには学習データとは異なる 35 種類のマグカップを同様に回転させた距離画像計 1,260 枚を使用した。画像は上下左右いずれかをランダムに 0~50% 欠損させた。PCA を用いた特徴ベクトルの次元数は 400 次元とした。10 度刻みの計 36 クラスに分類したときの推定結果と真値との平均角度誤差を評価したところ、[2] 相当の手法では 14.26 度、提案手法では 10.96 度であった。この結果から、学習済み AutoEncoder を 2 回使用する提案手法による精度の向上を確認した。

4 むすび

欠損復元 AutoEncoder を用いた欠損を含む物体の姿勢推定手法について検討した。今後は、特徴抽出手法の改善、多様体構築による未知姿勢推定について検討する。

謝辞 本研究の一部は科学研究費補助金による。

参考文献

- [1] H. Ninomiya, et al., "Deep Manifold Embedding for 3D Object Pose Estimation," VISIGRAPP, vol.5, pp.173-178, March 2017.
- [2] M. Sundermeyer, et al., "Implicit 3D Orientation Learning for 6D Object Detection from RGB Images," ECCV, pp.699-715, Sept. 2018.
- [3] A.X. Chang, et al., "ShapeNet: An Information-Rich 3D Model Repository," arXiv preprint arXiv:1512.03012, Dec. 2015.