

深層学習による物体検出のための三次元点群による物体形状の再構成

田中 靖浩[†] 西口 敏司^{††}

[†] 大阪工業大学 大学院情報科学研究科

^{††} 大阪工業大学

1. はじめに

画像に含まれる物体の領域と種類を判定する物体検出において、深層学習が広く利用されるようになってきた。深層学習では学習データとして識別対象の物体を撮影した画像と、その領域に関するアノテーション情報を大量に用意する必要があり、非常に手間のかかる作業であることが知られている。この問題に対処するために、少ない学習データを水増しするデータ拡張や、既存の学習モデルを追加学習するための転移学習やファインチューニングなどの技術が利用されており、一定の効果が得られている。

一方、学習データを水増しする方法の一つとして、道路標識の検出を目的として、3DCG 技術を用いて学習データを生成する枠組みが提案されている[1]。この枠組みによって識別に必要なかつ十分な学習データが生成できれば、実世界で稼働するロボットの視覚など、環境に応じた適切な学習データを効率的に生成できる可能性がある。

そこで本研究では、距離センサを用いて識別対象の形状を獲得し、学習データを生成する枠組みについて検討する。

2. 提案手法

2.1 三次元点群の獲得

識別したい実物体の三次元形状を、深度センサを用いて三次元点群として獲得する。三次元点群とは、三次元空間上に表現された xyz 座標情報と RGB 情報を持つ三次元点の集合で構成される点群である。実物体の例を図 1(a)に、また、深度センサを用いて獲得した三次元点群をある視点から可視化した例を図 1(b)に示す。この図に示すように、深度センサの視点から獲得された三次元点群は、センサの反対側の形状は観測できないため欠損が存在する。

2.2 複数視点の三次元点群の獲得

2.1 に示した欠損を補うために、識別対象の実物体をターンテーブルに乗せて 90 度ずつ回転させながら、計 4 回、視線(z 軸)を回転中心に向けた距離センサで三次元点群を獲得する。

2.3 物体形状の再構成

2.2 で獲得した複数の三次元点群について、ターンテーブルを含む平面の法線ベクトルおよび、ターンテーブルの中心までの距離、およびターンテーブルの回転角度を利用して、各三次元点群を平行移動および回転することで位置合わせし、三次元物体形状を再構成する。

2.4 学習用画像の獲得

再構成した三次元物体形状にポリゴン情報を付与した上で 3DCG 空間に配置し、環境光や照明光、および、撮影カメラの位置や向きを様々に変えた上で、物体を撮影し、学習用画像として獲得する。

3. 三次元物体形状再構成実験

三次元点群の獲得と再構成には、Point Cloud Library(PCL)を使用した。複数視点からの三次元点群を用いて物体形状を再構成した例を図 1(c)に示す。



(a) 撮影した物体



(b) ある視点から撮影した三次元点群の例



(c) 複数視点からの三次元点群を合成した結果の例

図1. 三次元点群による物体形状の再構成の例

4. まとめ

本研究では、物体検出のための深層学習に用いる大量の画像を生成するための手法の一つとして、三次元形状データを用いる手法を検討した。今後は獲得した三次元形状データにポリゴン情報を付与することで 3DCG 化し、実際に多数の学習用画像を生成したデータを用いて学習した結果がどの程度の識別性能を持つかを評価・検討していく予定である。

参考文献

- [1] Ryuto Kato, Satoshi Nishiguchi, Wataru Hashimoto, Yasuharu Mizutani, "Generating Training Images Using a 3D City Model for Road Sign Detection," 20th International Conference, HCI International 2018, Proceedings, Part III, Vol. 852, pp.381-386, 2018-07.