

Object Detection

ER17076 安井 理

目的

- ・実際に実装することができるか.
検出・推定・について理解が薄いので,そこを調べる事
- ・検出部について調べるために、論文中のobject detectionを調査
- ・今後の方針を決めるため

概要

- ECCV2018のBest paperに選ばれた6次元物体検出の論文
Implicit 3D Orientation Learning for 6D Object Detection from RGB Images
- 物体検出と6次元ポーズ推定のためのRGBベースの手法を提案
- CADモデル(3Dモデル)だけで6DoF物体検出の学習ができる
- Domain Randomizationを活用したDenoising Autoencoderを学習

概要

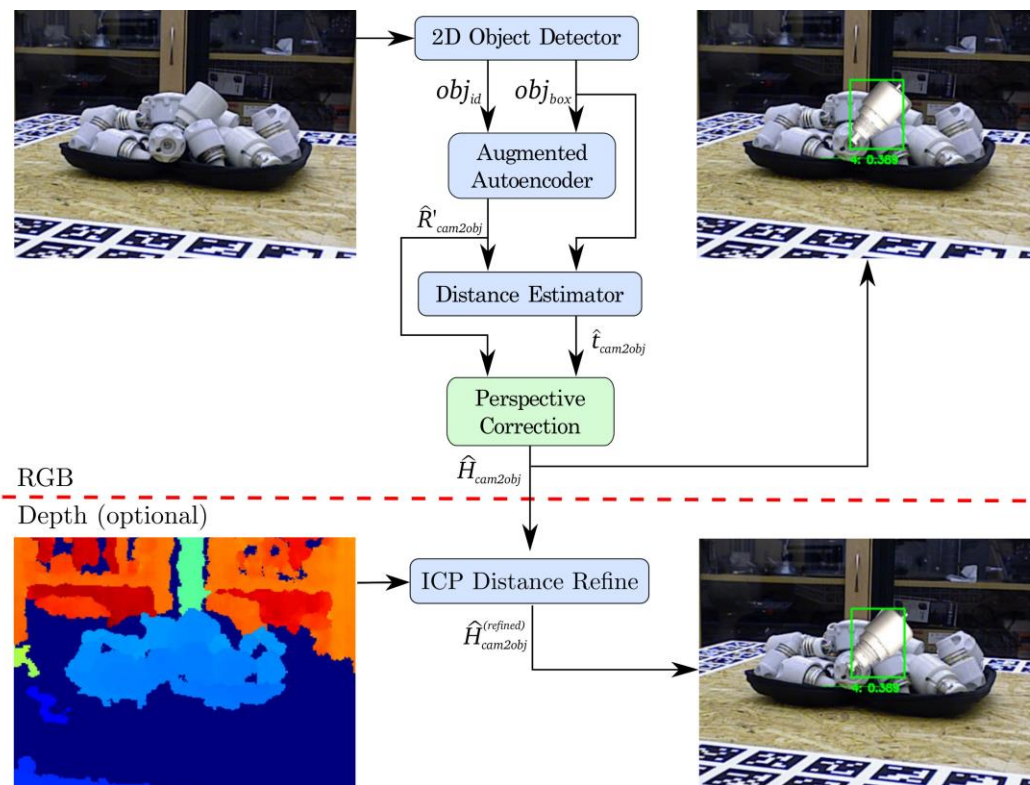
・トレーニング

- ・あらかじめAugumented Autencoder(AAE)を使い物体を様々な視点でトレーニング
- ・物体IDとバウンディングボックスの長さを蓄積

・テスト

- ・特徴検出にはVGG-16ベースのSSDを使用
- ・検出された物体IDとバウンディングの情報をAAEに通し蓄積されたデータと合わせ姿勢推定

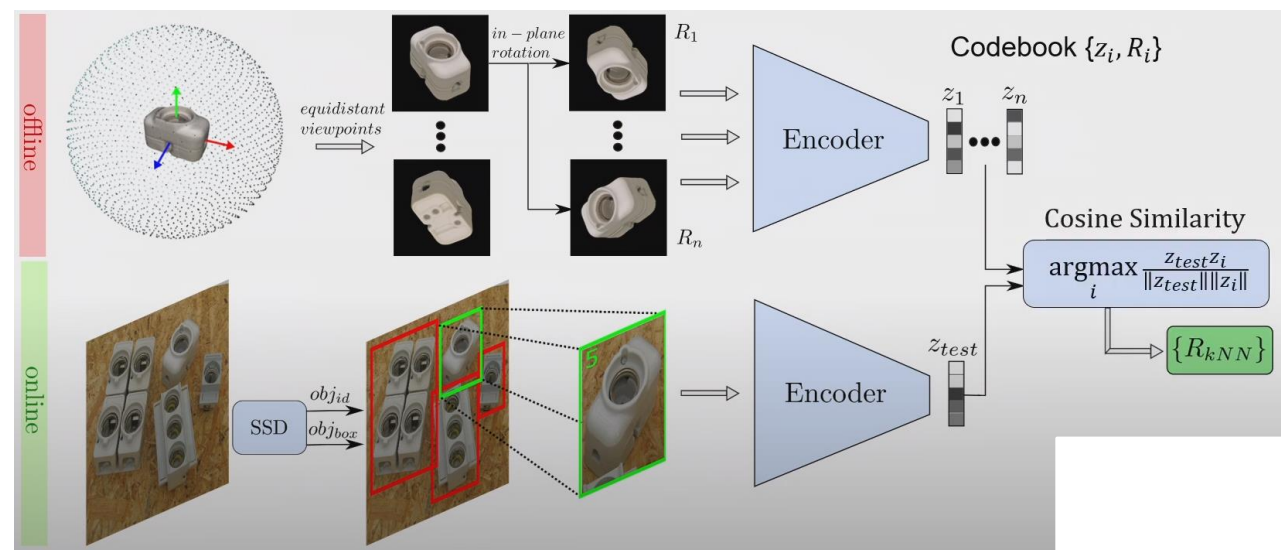
推定の流れ



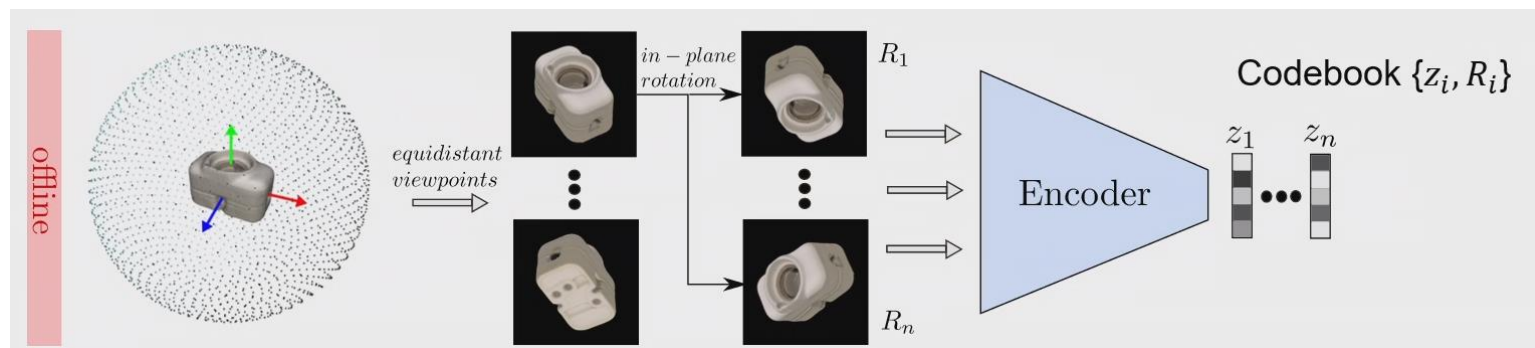
- ・赤色の切り取り線の上までがRGBデータのみでのプロセス
- ・深度センサーがあればより精度が上がる
- ・AAEは3Dモデルを用いて学習

検出と推定

- ・AAEを使い3Dモデルの色々な角度の潜在表現を計算
- ・テスト画像が入力されSSDを通したらAAEに通し潜在表現を計算
- ・計算済みの潜在表現に類似したものを
cosin距離に基づいて探し出力



検出と推定

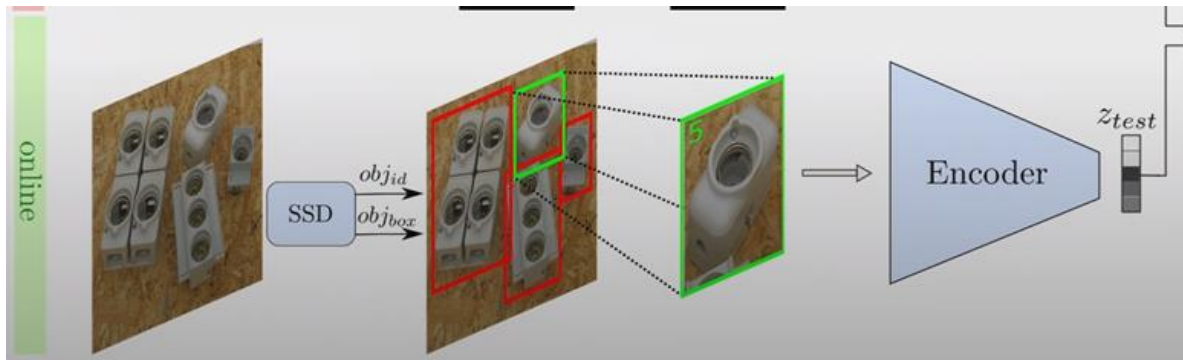


オフライン: AAEのトレーニング

- ・合成物体を用意
- ・等間隔に物体を回転させ、すべての角度の視点を用意
- ・AAEにかけ潜在変数 $z \in \mathbb{R}^{128}$ を生成して (z_i, R_i) を記録 (92232)

R_i : (x、y、z) 方向の情報

検出と推定



オンライン: SSDによる検出

- ・テスト時には、最初にRGBシーンで対象となる物体が検出
- ・バウンディングボックスは二次に切り取られ、エンコーダの入力サイズにリサイズ
- ・エンコーダーで潜在変数 $z_{test} \in \mathbb{R}^{128}$ を得る

検出と推定

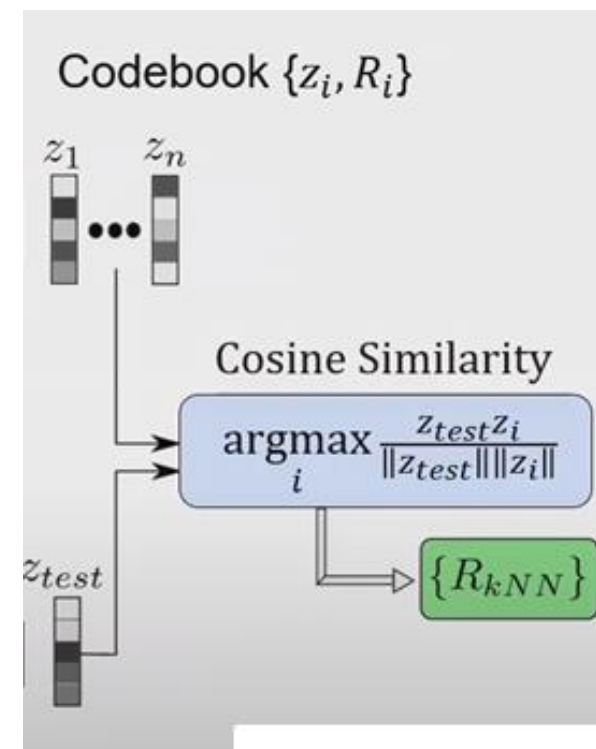
推定

- 蓄積されたデータと類似度の高いものをk近傍法で決定

$$\text{COS}_i = \frac{z_i \times z_{t_{est}}}{|z_i| |z_{t_{est}}|} \quad \text{この計算で一番大きい値を選択}$$

- 最も近かったデータから

回転行列 (R_{kNN}) が推定値として返される



検出と推定

推定

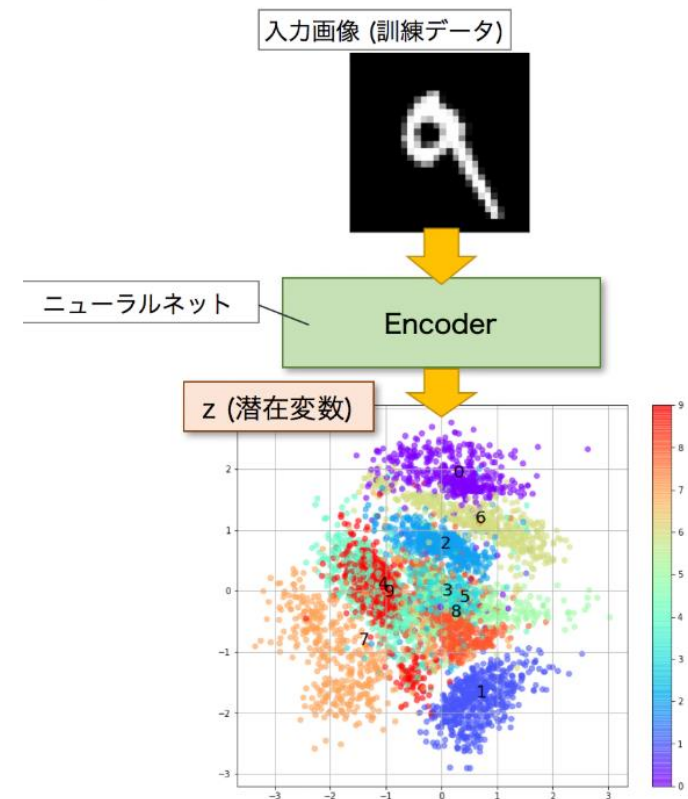
$$\text{COS}_i = \frac{z_i \times z_{t_{est}}}{|z_i| |z_{t_{est}}|}$$

この計算で一番大きい値を選択

z の値は潜在変数

AAEのコード内で推定まで含まれている.

Encoder部分



実装について

- ・主にAugumented Autencoderのトレーニングコード,テストコードは有
- ・SSD検出部分は他のgitを参考にしながら実装が可能.
- ・推定はAugumented Autencoderのテスト部に含まれている

・ 今後

- ・まず、実際に動作できるかわからないので動作確認を行う,
- ・できなそうな場合、報告しSSDの拡張を考える

参考文献

6次元物体検出の論文

http://openaccess.thecvf.com/content_ECCV_2018/papers/Martin_Sundermeyer_Implicit_3D_Orientation_ECCV_2018_paper.pdf

- git-hub <https://github.com/DLR-RM/AugmentedAutoencoder#testing>
- 論文読み: <https://qiita.com/ttyszk/items/c58250e1bbf9983d0cf3>
- 論文要約: <http://toaruharunohi.hatenablog.com/entry/2018/09/14/120238>