

Augmented Augmented

ER17076 安井 理

目的

- ・SSD-6Dの論文を実装するのが難しい
- ・もう一つのRGB画像からの方向推定の論文を読んでみる
- ・とりあえず一週間しかないので調べてみる

概要

- ECVV2018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
- 特に方向推定の部分に長けている
- 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

推定の流れ

RGB画像をSSDなどの既存の物体検出手法に入力

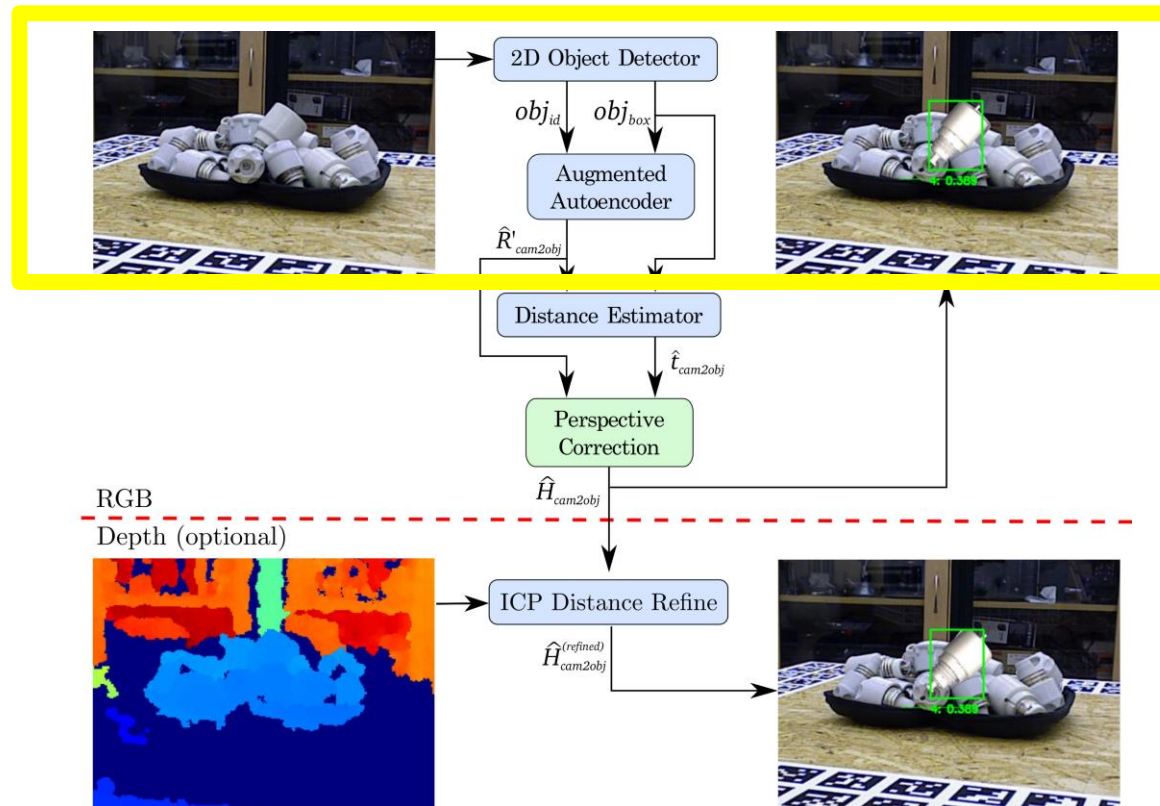
→ (x, y) 情報を持ったバウンディングボックスを検出

検出された画像をAugmented (拡張) Autencoderに入力3次元方向推定

アウトプット情報を使い奥行き z を計算することで6次元の推定が行える

ここで言われる6次元は (x, y, z) 座標に加えて (x, y, z) の回転方向の推定

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
- ・Augmented Autencoderは3Dモデルを用いて学習

Domain Randomization

- ・3Dデータで学習し、現実の物を推定するための手法
- ・3Dモデルに様々な環境のパラメータを追加して学習
- ・物体と背景の対称性を明確化
- ・テスト時に現実物でも推定可能になる



Fig. 1. Illustration of our approach. An object detector is trained on hundreds of thousands of low-fidelity rendered images with random camera positions, lighting conditions, object positions, and non-realistic textures. At test time, the same detector is used in the real world with no additional training.

Augumented Autencoder (AAE)

- Autencoder

教師なし学習・データを表現する特徴を獲得するネットワーク

- Denoise Autencoder

ノイズありの画像を入力.ノイズなしの画像を出力するように学習
ノイズによらない本質的な潜在表現を得る

- Augumented Autencoder

ノイズ以外の画像加工を加え学習する

Denoise Autencoderの応用

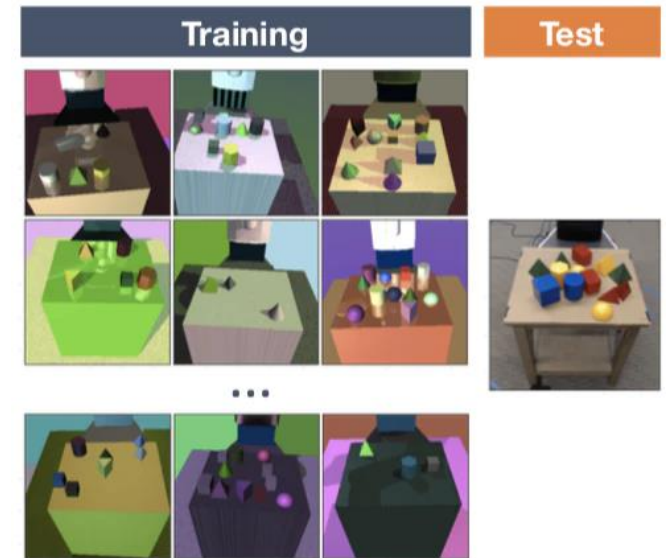


Fig. 1. Illustration of our approach. An object detector is trained on hundreds of thousands of low-fidelity rendered images with random camera positions, lighting conditions, object positions, and non-realistic textures. At test time, the same detector is used in the real world with no additional training.

Augumented Autencoder(AAE)

- Augumented Autencoderを示す数式

$$x'' = (\psi * \Phi * f_{augm})(x) = (\psi * \Phi)(x') = \psi(z')$$

- 損失関数 ℓ_2 ノルム(差を最小にして本物に近づける)

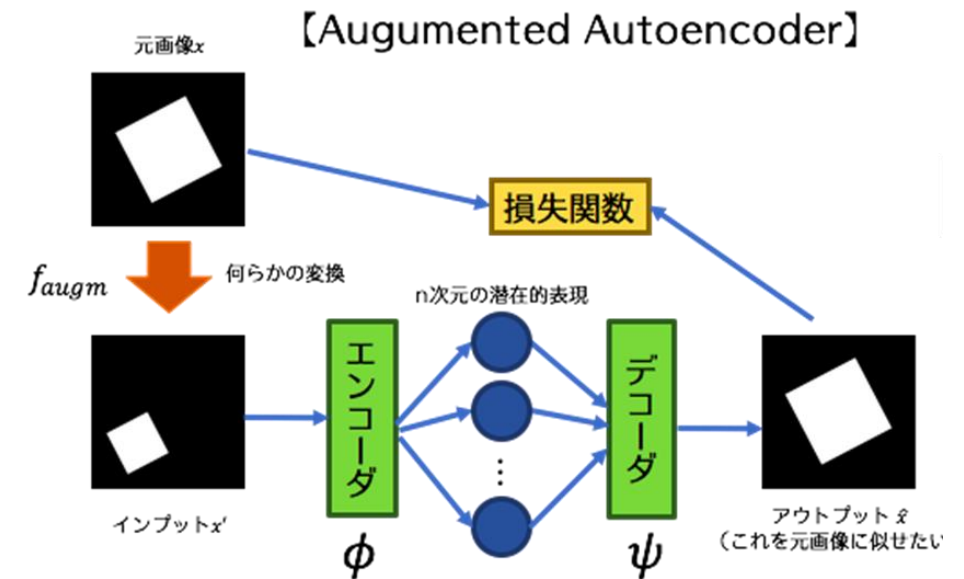
$$\ell_2 = \sum_{i \in D} ||x(i) - x''(i)||_2$$

x : 入力(インプット) f_{augm} : 変換(加工)の関数

x' : インプット画像を変換(加工)したもの

x'' : アウトプット ϕ : エンコーダー

ψ : デコーダー



Augmented Autencoder (AAE)

通常AutencoderとAugmented Autencoderの比較

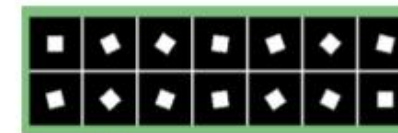
緑枠の(a)をベース画像として

1: (a) → (a) 2: (d) → (d) 3: (d) → (a) の入出力で学習

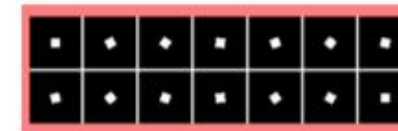
結果として

通常のAutencoder1・2は応用があまり利かない

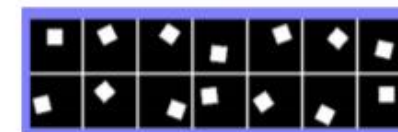
Augmented Autencoderは四角形の潜在表現を反映できる



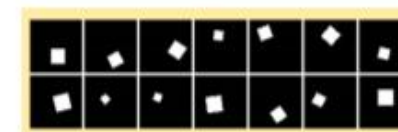
(a) $X_s=1.0, t_{xy}=0.0, r \in [0, 2\pi]$



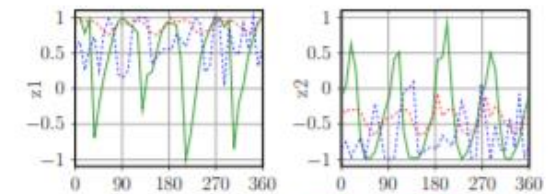
(b) $X_s=0.6, t_{xy}=0.0, r \in [0, 2\pi]$



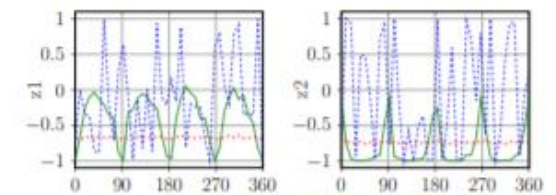
(c) $X_s=1.0, t_{xy} \sim \mathcal{U}(-1, 1), r \in [0, 2\pi]$



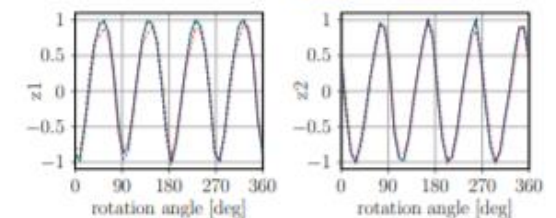
(d) $X_s \sim \mathcal{U}(0.5, 1), t_{xy} \sim \mathcal{U}(-1, 1), r \in [0, 2\pi]$



(1) Autoencoder (a) → (a)



(2) Autoencoder (d) → (d)



(3) Augmented Autoencoder (d) → (a)

Augumented Autencoder (AAE)

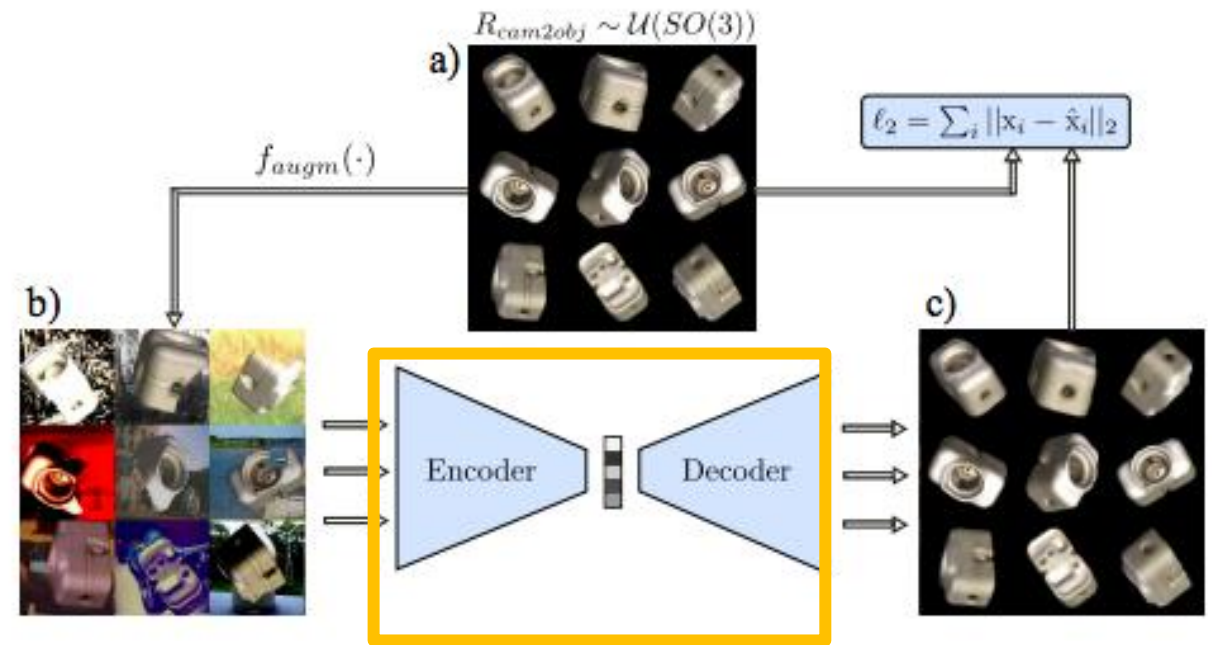
3Dモデルデータに対して学習するとき

a)学習したいもの

b)背景・光・遮蔽物などを追加

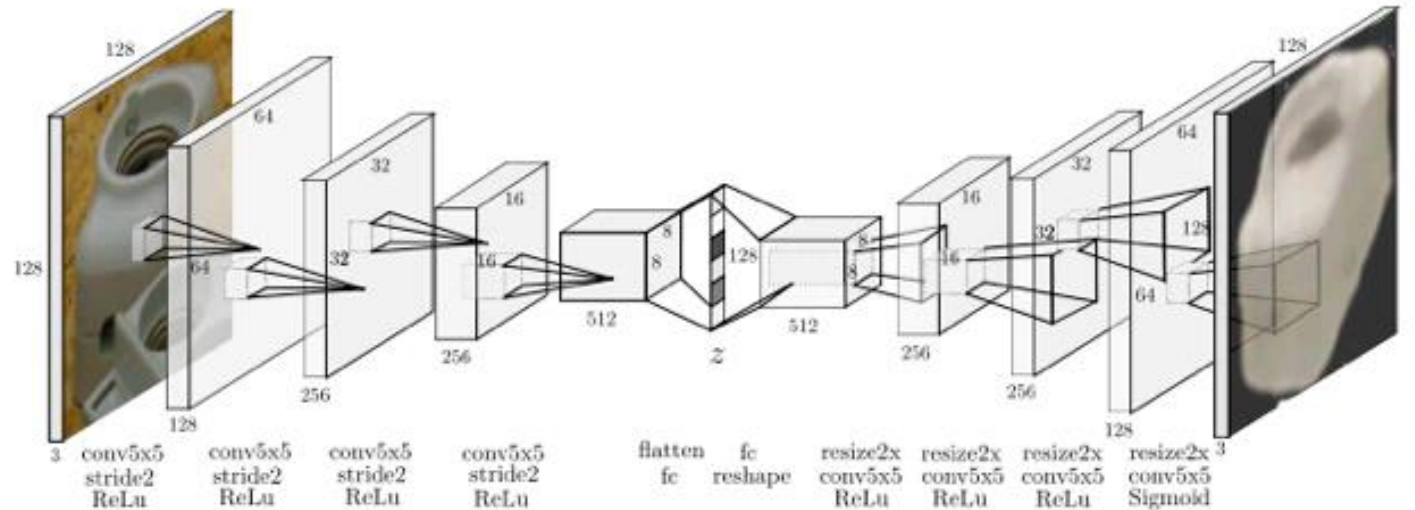
c)潜在表現の出力

損失関数[(a)と(b)の差]を最小になるように学習



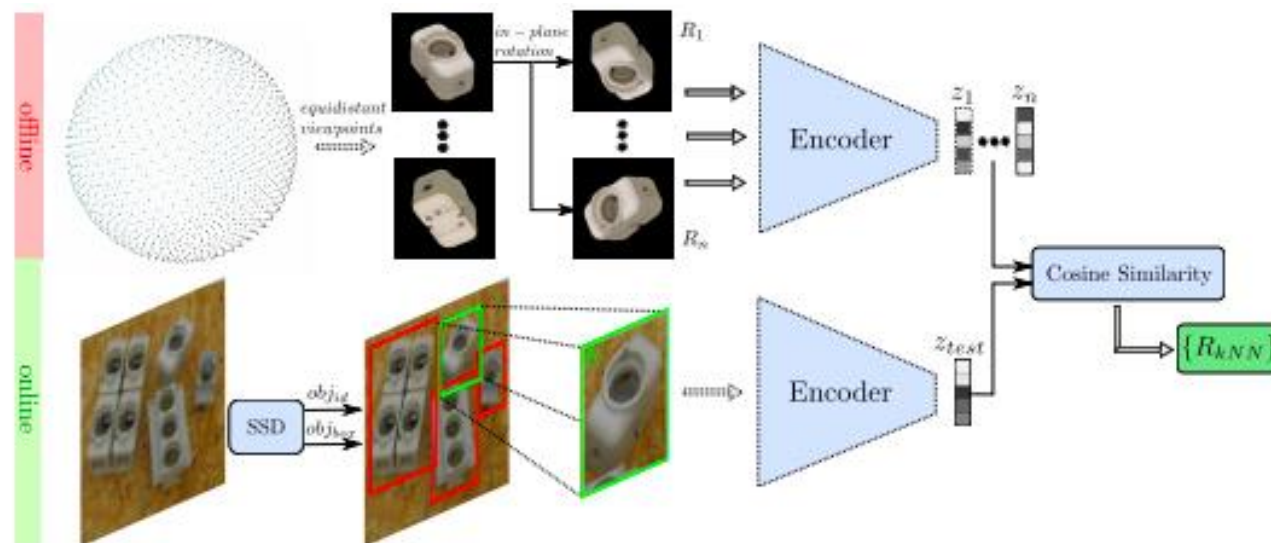
Augumented Autencoder(AAE)

構造は 128×128 の画像を 5×5 の畳み込みを4回
conv 5×5 stride2 ReLu
全結合層を1回行って潜在表現を得る



検出

- Augmented Autencoder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Augmented Autencoderに通り潜在表現を計算
- 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



・ 今後

- ・ 検出部の座標、深度の求め方について調べる
- ・ 実際に動かしてみる

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

- ・論文読み: <https://qiita.com/ttyszk/items/c58250e1bbf9983d0cf3>
- ・論文要約: <http://toaruharunohi.hatenablog.com/entry/2018/09/14/120238>