

第 8 月 進捗確認会報告資料

変形 AR マーカの検出・推定

安井 理

令和 2 年 10 月 06 日

1 はじめに

キャッシュレス決済や物品管理, 広告, ロボットの認識機能等の分野で QR コードや AR マーカに代表される 2 次元コードが利用されている. 2 次元コードには, 数百から数千バイトの情報を埋め込むことができ, シンボルと呼ばれ特殊なパターンによって視点が変化しても高精度な検出が可能である. さらに, 2 次元コードの大きさを事前に定義すればカメラの位置・姿勢を推定することができる. しかし, 2 次元コードは平面に貼ることを前提条件としており, 曲面に貼られた 2 次元コードは歪みによる見えの変化を引き起こすため, 認識精度が低下する問題を抱えている. そこで, 前任者の鈴木さんの研究では機械学習により変形した AR マーカを認識する手法を提案する. 変形した画像を FasterRCNN[1] により学習することで, 歪みを含む画像を正確に認識し, ID, 座標, 大きさ, 変形度合いを推定を行った, 今回行っている研究では FasterRCNN では座標, 大きさ, 変形度合いの推定が明確に行うことができないため, Augmented Autencoder[2] を用いて正確に姿勢情報を推定し, 推定した情報から歪みを取り除き正面から観測した AR マーカ画像に変換する.

2 進捗報告

今回の進捗報告の内容を以下に示す.

- Tensorflow-gpu の動作確認
- AAE のトレーニング

2.0.1 Tensorflow の概要・特徴

2.0.1.1 Tensorflow の概要

機械学習の分野で使用するための OSS (オープンソフトウェアライブラリ) であり, 開発元が Google Brain チームである.

対応 OS : Linux, macOS, Windows, Android, iOS...

対応言語 : C 言語, C++, Python, Java, Go などと機械学習の分野で幅広く使用されるものである.

主に以下の特徴が挙げられる

- ・データの読み込み、前処理、計算、状態、出力などを多次元的に処理する
- ・分散処理が可能なためビッグデータ（大量のデータ）も扱える

2.0.1.2 OSS

ソースコードの学習や変更・配布することが認められているライブラリ

2.0.1.3 Tensorflow 最終環境

CUDA9.0, cuDNN7.0, Tensorflow-gpu1.12.0 が現在使っている PC で Tensorflow-gpu を動作させられる version である.

2.0.1.4 GPU 確認コマンド

GPU 確認コマンドで確かめた際, gpu がうまく動作したときと動作しなかったときの出力の違いを以下に示す.

GPU が動作しなかったときの出力

```
physical_device_desc: "device: XLA_GPU device"
```

おそらくオンボードの GPU が反応していまっていたのではないかと考えられる.

GPU が動作したときの出力

```
physical_device_desc: "device: 0, name: GeForce GTX 1080, pci bus id: 0000:01:00.0, compute capability: 6.1"
```

device が GeForce GTX1080 と GPU の名前に代わっており学習に GPU が使用されることも確認できた. 変更点としては CUDA9.0 に対応する Tensorflow のバージョンを色々試し, 最終的に最新の 1.12.0 がうまくいった.

2.0.2 AAE のおさらい

2.0.2.1 AAE の概要

AAE は implicit 3d orientation learning for 6d object detection from rgb images[2] の論文で提案された手法であり, ECCV2018 の Best paper に選ばれた 6 次元物体検出の論文である.

6D 物体検出は、3次元空間座標だけでなく3方向の向き姿勢情報も含んだ検出問題であり、高速に推定を行えることに加え、6D のラベル付き教師データがなくても学習可能な手法である、6D ラベル付き教師データの代わりに、検出対象となる物体の3D CAD データが必要となる．全体の処理の流れ図1としては、まず入力となる RGB 画像に対して SSD を用いて対象物体の Bounding Box を推定、その後、推定された Bounding Box 領域から物体の姿勢情報を推定するという処理を行う、後半の Bounding Box 領域から物体の姿勢情報を推定する部分が今回自分の研究に取り入れようと考えている方法で、AAE という手法が用いられている．全体の流れを図1 図2 に示す．

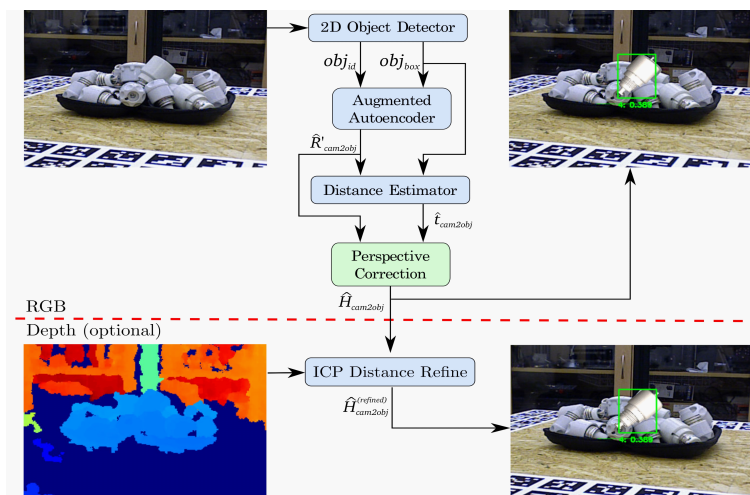


図 1: 全体の流れ.

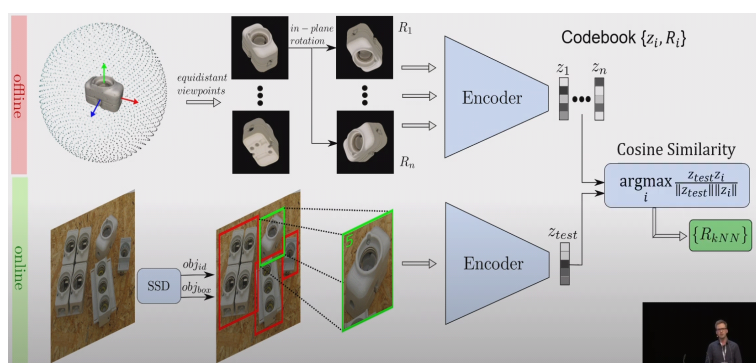


図 2: 全体の流れ 2.

2.1 AAE のトレーニング

今回はトレーニングを行う段階まで進められた．まずトレーニングは学習したい物体の様々な視点から見た画像を計 92232 通り用意しそれを環境画像（VOC のデータセット [3][4]）に張り付けノイズを生成したものを用意する．準備した画像を図 3 に示す．

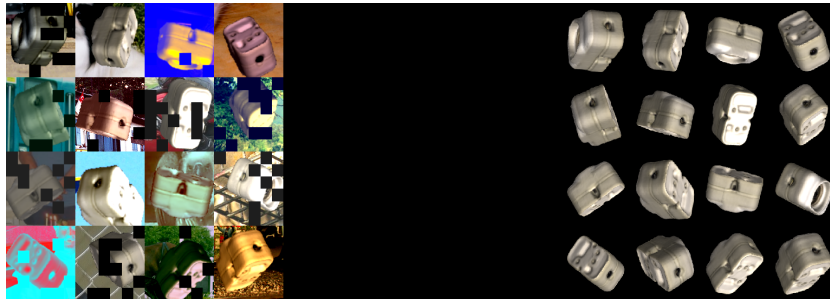


図 3: トレーニングイメージ.

3 おわりに

今回まで Tensorflow のバージョン合わせに時間を取られ実際の研究を進めることができず，一か月の内容としても進捗状況も進みが悪くなっていたが，やることとしては明確になっているので，ペースアップをしていきたいと思っている．これまではトレーニング部分を完了したので次にやることとしては姿勢推定を行うテスト，自分の研究の AR マーカーモデルの作成，これを用いたテスト，二次元の検出からの AAE を用いた姿勢推定までの流れを行うという順番で研究を進めていきたいと考えている．

参考文献

- [1] S.Ren *et al.* : “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, Proc. of NIPS, 2015.
- [2] Martin Sundermeyer Zoltan-Csaba Marton Maximilian Durner Rudolph Triebel , 2018
<https://arxiv.org/pdf/1902.01275.pdf>
- [3] “The PASCAL Visual Object Classes Challenge 2007”. <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html>
- [4] “Visual Object Classes Challenge 2012”, <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html>
- [5] Karen Simonyan, Andrew Zisserman, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION”, 2015.