

深度カメラを活用した RGB画像からの奥行推定

平石研究室 S19096 小島 龍輝



はじめに

テーマ

自動運転技術の発展に伴い、距離センサーとしてLiDARやステレオカメラを活用するケースが増えている。

本研究では、機械学習を用いて単眼カメラによる通常のRGB画像から深度画像を生成する試みを行う。

目指すもの

人間が視覚的に推論している物体の距離感というものを、機械学習によって獲得できないか。



研究概要

機械学習には、敵対的生成ネットワーク (GAN)の一種である、Pix2Pix手法を使用した。

使用機材

- Intel RealSense D435 (図 1)

使用ソフト・ライブラリ

- Python 3.7.9
- OpenCV 4.5.1.48
- TensorFlow-GPU 1.15.4
- Pix2Pix-tensorflow (注1)



図 1 RealSense D435

注1 pix2pix-tensorflow, <https://github.com/affinelayer/pix2pix-tensorflow>

研究手法 -深度カメラ特性-

学習データを作るにあたり，深度カメラのIRレーザーには以下の特性がある．

- 乱反射
- IR光の吸収

図2のような情報の欠損が起こることがわかった．

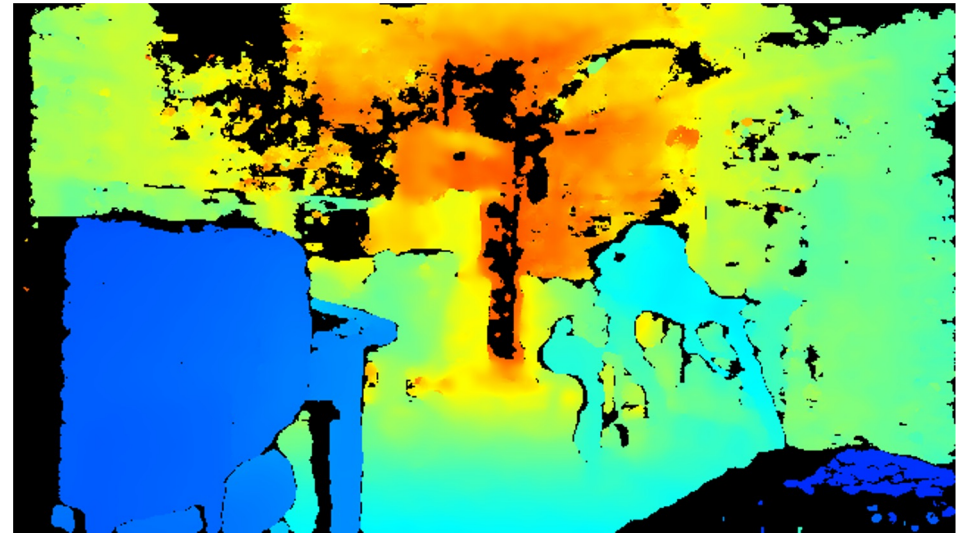


図2 深度カメラ画像

研究手法 -学習データ作成-

そこで、OpenCVを用いて欠損した情報の補完を行った。(図3)

- A) メディアンフィルタによる小さな誤り深度情報をノイズとして除去
- B) 欠損部分のマスク画像を作成
- C) AとBを元にした未反射部分のFast Marchingのアルゴリズムによる色情報の補完を行う

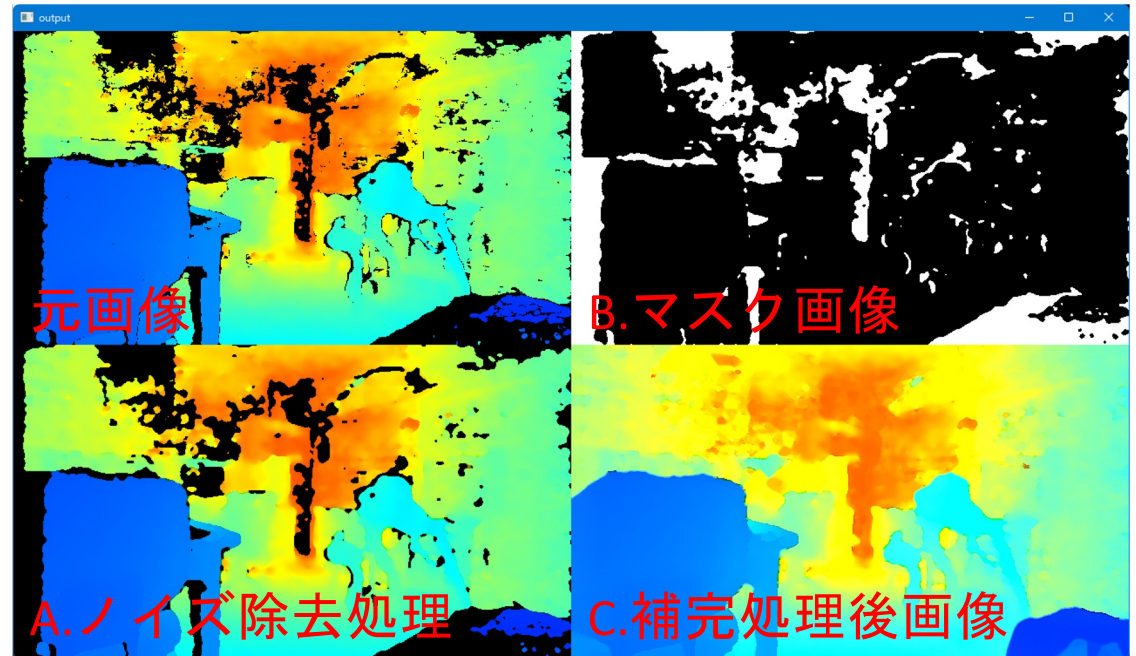


図3 学習データの補完処理例

研究手法 -学習-

学習に使用したPix2Pix手法とは、GANの一種で画像を別の画像へ変換するというものである。

図4上のように画風変換が得意なネットワークとなっている。

前項で触れた学習データの例が図4下である

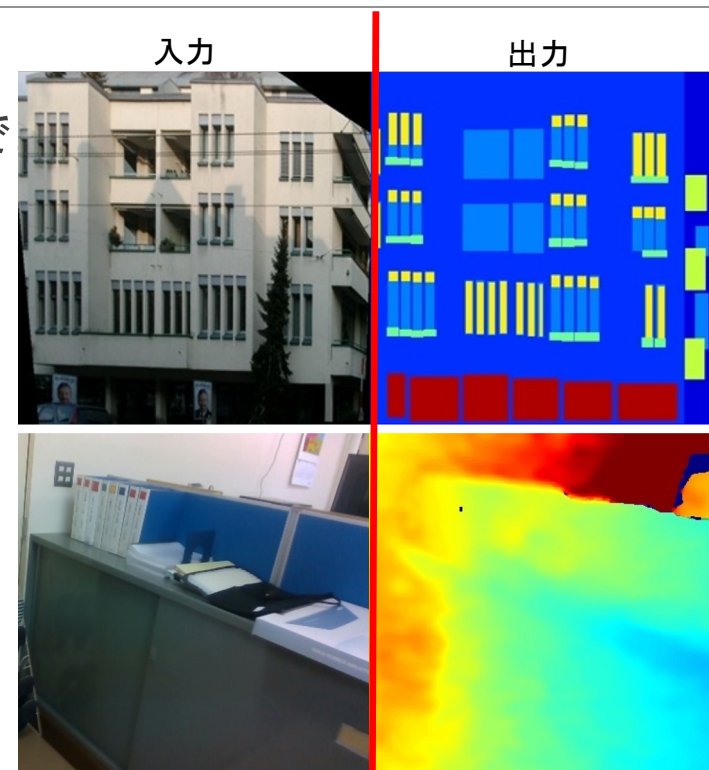


図4 画像変換例 上：注2より引用 下：学習データ例

注2 pix2pix-tensorflow, <https://github.com/affinelayer/pix2pix-tensorflow>

実験結果

学習を行ったネットワークに評価データを入力した例（図5）。

成功例の評価

- 大まかな色の分布
- 冷蔵庫に対する着色の正確性

失敗例の評価

- オブジェクトに対する不正確な着色が目立つ

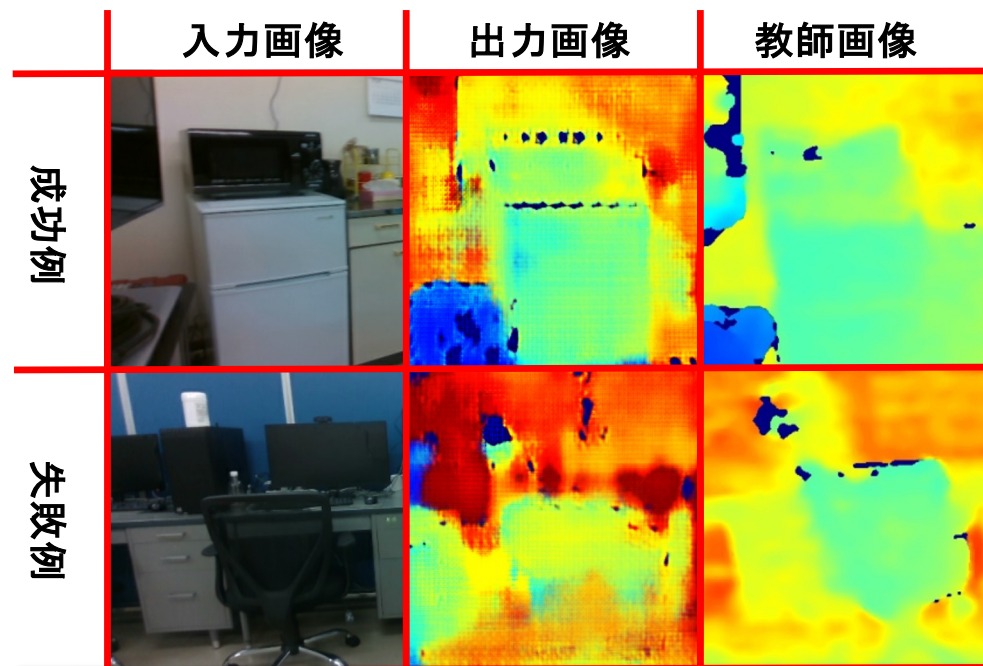


図5 学習評価例

おわりに

前期の本段階では、公開されていたPix2Pix手法を活用した機械学習を行い、ある程度の精度を獲得することができた。

後期では、本研究の発展として以下の改良を行うつもりである。

- Attention機構の取り入れなどによる機械学習ネットワークの改良
- 学習データの拡充による精度の向上