

複数の補助情報を活用した 浸透学習法による単眼深度推定

岩上 友哉[†]長尾 智晴^{††}[†] 横浜国立大学 理工学部^{††} 横浜国立大学 大学院環境情報研究院

1. はじめに

近年、自動運転の研究が盛んに行われており、深度推定に対するニーズが高まっている。深度推定には LIDAR センサやステレオカメラを用いることが一般的だが、単眼の画像から深度を推定することで低コストに実現可能となる。しかし、単眼深度推定は不良設定問題であり、既存手法では動画の前フレームなどの補助情報を用いて学習を行うことで精度の向上を実証している。

また、一般に複数の情報を使用することが精度向上に有用である。しかし、テスト時も使用可能な補助情報を複数用意することは高コストである。そこで、学習時にのみ補助情報を用いる浸透学習法[1]が提案されている。

本研究では、浸透学習法を拡張し、複数の補助情報を利用することで高精度に深度を推定する手法を提案し、その有効性を検証する。

2. 浸透学習法

浸透学習法では、学習時とテスト時両方において使用可能な Main データと、学習時にのみ使用可能な Aux データの 2 種類を使用する。学習は事前学習と浸透学習の 2 段階に分かれている。はじめに、事前学習では Main データと Aux データ全てを用いて目的のタスクを解く学習を行い、その特徴量を抽出する。浸透学習では、Aux Network の出力に掛ける decay を徐々に 0 に近づけながら、事前学習で得られた特徴量を Main データのみから再現するように学習を行う。これによって、Main データのみから Aux データの特徴量も活用した学習を行うことができる。

3. 提案手法

図 1 に提案手法のモデル構造を示す。本手法では、浸透学習法における Aux Network を複数に拡張する。Aux データそれぞれに対して特徴量を抽出するネットワークを用意することで各データから有効な特徴量を獲得できる。その特徴と Main データの特徴を結合して 1 つの Percolation Network に入力し、浸透学習を行う。

4. 実験設定

本研究では、SYNTHIA-SF DATASET[2]という CG で作成された画像を用いた。データセットを学習画像 1761 枚、テスト画像 463 枚に分割して使用した。入力としては、Main データとして単眼 RGB 画像。Aux データとしてセグメンテーションマップとエッジ画像を用いた。教師画像である深度画像は 100m 以上の値を切り捨てた。モデル構造につい

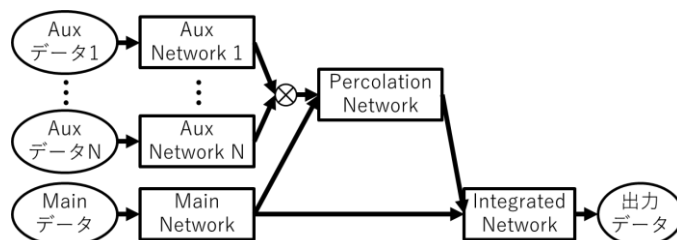


図1. モデル構造

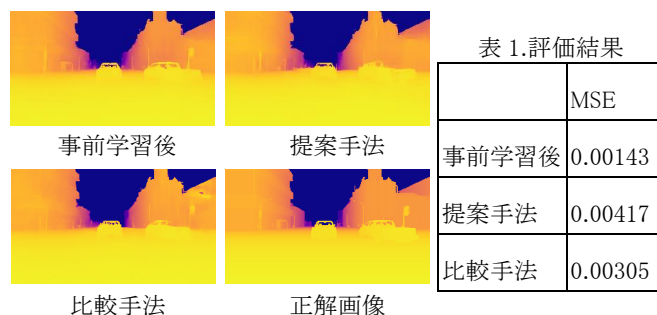


図 2. 実験結果

ては U-Net[3] をベースとし、Percolation Network は 7 層の CNN とした。decay は 0.95^{epoch} に設定した。また、比較手法としては Aux データをゼロ行列にしたものとした。

5. 実験結果

図 2 に実験結果を示す。図 2 や表 1 から、複数の補助情報を用いた事前学習後の方が、単眼 RGB 画像のみで推定した画像よりも高精度であることがわかる。しかし、提案手法の精度は比較手法に及ばなかった。これは事前学習時に精度の向上に寄与する特徴を獲得できたが、浸透学習によって、その特徴を再現することができなかつたためと考えられる。

6. まとめ

本稿では、浸透学習法において複数の補助情報を用いた単眼深度推定の手法を提案した。今後は他の補助情報を検討するとともに、その情報の組み合わせについても検討していく。

参考文献

- [1] 柳元美玖, 長尾智晴. “学習時にのみ使用可能な情報を浸透させるニューラルネット”. 研究報告数理モデル化と問題解決 (MPS), vol. 2017.6, pp. 1-6, 2017.
- [2] D. Hernandez-Juarez, et al. “Slanted Stixels: Representing San Francisco’s Steepest Streets,” In Proc. BMVC, 2017
- [3] Ronneberger, et al. “U-net: Convolutional networks for biomedical image segmentation.” In MICCAI, 2015.