

浸透学習法を用いた単眼画像からの距離画像推定

井上 多門[†]

† 横浜国立大学 理工学部

長尾 智晴^{††}

†† 横浜国立大学 大学院環境情報研究院

1. はじめに

単眼画像のみからの距離画像推定問題は一意な解をもたない不良設定問題である。これまでの距離画像推定を行う研究では、付加情報として両眼視差やセグメンテーション画像を利用する。しかし、正解画像そのものに正解画像を推定するために有効な特徴量が含まれていると考えられる。そこで本稿では、正解画像の緩和を用いた距離画像推定を提案する。学習時、正解画像の緩和を使用するために浸透学習法[1]を用いる。

2. 浸透学習法

浸透学習法は、学習時、テスト時両方で利用可能なデータ(主データ)の他に、学習時のみ使用できるデータ(補助データ)を利用する手法である。図1にモデル構造を示す。図1中の decay は、補助データの入力値に乘算される 0 以上 1 以下の係数である。学習は事前学習と浸透学習の2段階で行う。事前学習では、主データと補助データ両方を使用してモデル全体を学習する。テスト時の誤差が収束した段階で、浸透学習に移行する。浸透学習では、decay を指数関数的に下げ、Percolation Network の出力値を教師データとして Percolation Network のみを学習する。この手法によって、主データのみで補助データの特徴量を再現する。

3. 提案手法

本稿では、浸透学習法による正解画像の緩和を用いた単眼画像からの距離画像推定を提案する。正解画像の緩和とは正解画像にマスク処理を施した画像である。学習には 3DCG 画像を使用し、特に補助データは正解画像の緩和を使用する。これらにより、カメラと被写体の遠近に依らず、画像内全体の詳細な距離情報を学習時に考慮することができる。正解画像の一部にマスクをかけた画像を使用することで、Aux Network の情報に依存しすぎない学習を行うことができる。[1]では、モデルの構造に全結合層のみを使用していたが、提案手法では、特徴抽出のために Convolutional Neural Network(CNN)を使用する。

4. 実験設定

本実験では 3DCG 画像で構成される Virtual KITTI Dataset[2]のうち、1669 枚を学習用に、447 枚をテスト用に使用した。主データは単眼画像、教師データは距離画像、補助データは距離画像にマスク処理を施した画像を用いた。誤差関数は平均二乗誤差、最適化法は Adam を使用した。また、epoch 数は事前学習を 100、浸透学習を 500 と

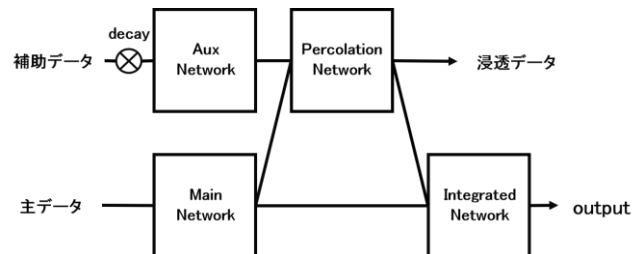


図1. モデル構造

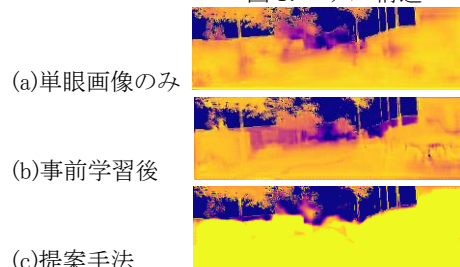


図2. 実験結果

した。本実験では $\text{decay} = 0.95^{\text{epoch}}$ とした。また、単眼画像のみで学習した場合を比較手法とした。

5. 実験結果

図2に実験結果を示す。この結果より、事前学習後の距離画像の精度は、単眼画像のみの距離画像よりも良いことがわかる。また、提案手法の距離画像の精度は、事前学習後の距離画像に及ばなかった。これは、補助データによって得られた特徴量を浸透学習で十分に再現できなかったためであると考えられる。

6. まとめ

本稿では、浸透学習法による正解画像の緩和を用いた単眼画像からの距離画像推定を提案した。提案手法は、浸透学習法によって正解画像の緩和を用いた単眼画像からの距離画像推定を提案した。実験より、学習時に正解画像の緩和を使用することで、精度は向上するが、有効な特徴量の再現までには至らなかった。今後、精度の向上に寄与する他の正解画像の緩和を検討する必要がある。

参考文献

- [1] 柳元美玖ほか. “学習時のみに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] A. Gaidon, *et al*, “Virtual worlds as proxy for multi-object tracking analysis.”, In Proc. CVPR, pp. 4340-4349, 2016.