In-depth Evaluation of Convolutional Spatial Propagation Networks

Paula López Diez, Kim Tang, Luis Samuel Patzwald

1. Introduction

Depth estimation from 2D images combined with distance measuring sensors is used in a variety of applications e.g. parking assistance, self driving cars, augmented reality and robotics. Past researches focus on improving accuracy, reducing necessary computing power and making it an application that is usable in real time. In order to create a reliable 3-dimensional view out of an image and selected given depths, deep fully convolutional neural networks were trained with large amounts of data in recent efforts. This approach results in accurate estimation of global scene layout and scales with advanced networks like VGG and ResNet and a better local structure recovery. Since the predicted depths and the real depths still do not align very well, there is space for improvement. Further studies lead by Liu et al. [reference?] propose to directly learn the image-dependent affinity through a deep CNN with spatial propagation networks (SPN). Although this leads to better results regarding the per-pixel distance prediction, either a scan-line or scan-column propagation is performed. This means that a value for the top left pixel can only be updated after the information from the bottom left pixel, which can lead to inaccurate distance predictions because the to pixels do not necessarily have any relation to each other. This research paper proposes to only use local surrounding pixels in order to predict values. Convolutional spatial propagation networks (CSPN), where all depths are updated simultaneously are used, which leads to improvements in speed and quality when compared to spatial propagation networks (SPN).

2. Proposed Method

In our final project we want to focus on two parts with our research. In the first part we will implement the code of the presented Convolutional Spatial Propagation Network (CSPN) on our own systems. To test reproducibility, different hyper parameters will be considered for evaluation on the New York University Depth Version 2 ("NYU Depth V2") dataset. Moreover, we plan to test the CSPN also on another dataset, namely the "NYU Depth V1" dataset and evaluate the performance of the network by comparing it with the performance on the NYU Depth V2 dataset. This allows us to evaluate whether the original model is overfitted to the dataset it was trained on and if it can perform well on other unseen data. The performance will be then evaluated with statistical metrics that were also used in the original paper, such as root-mean squared deviation (RSME).

We will also research other state of the art implementations as depth estimation with given sparse samples or learning affinity for spatial diffusion. We consider this will give us a better perspective about the method described in the paper and give us the possibility of comparing different approaches. By analyzing other implementations we expect to be able to derive some meaningful discussion that will arise to our final conclusions.

References

[1] X. Cheng, P. Wang, R. Yang, "Depth Estimation via Affinity Learned with Convolutional Spatial Propagation Network", ECCV, pages 103-119, 2018