Image Guided Depth Enhancement via Deep Fusion and Local Linear Regularization

Jiang Zhu Jing Zhang Yang Cao Zengfu Wang

Department of Automation, University of Science and Technology of China

Abstract

Depth maps captured by RGB-D cameras are often noisy and incomplete at edge regions. In this paper, we propose a deep residual network for image guided depth enhancement. The proposed deep fusion scheme can effectively extract the correlation between depth map and color image in the deep feature space. In addition, a specific layer of network is presented to introduce a local linear regularization constraint on the output depth. Experiments on various applications, including depth denoising, super-resolution and inpainting, demonstrate the effectiveness and reliability of our proposed approach.

Introduction

In this paper, we propose a CNN-based framework to learn the underlying correlation between depth maps and color images, and then to use the correlation to enhance the quality of depth map. The challenges and contributions are summarized as follows:

Challenges:

(1) The pooling layer, which is the common solution to wide the receptive field, will cause the blurry edges in the output depth-map.

(2) One possible solution to improve the performance of deep CNN is to increase the network depth by adding convolutional layers. However, this will introduce more parameters and increase the risk of overfitting.

(3) Depth map and color image have different noise level and data distribution. It does not work well for directly using depth map and color image as the input of the CNN. A feature level fusion strategy is required for the learning process.

Contributions:

(1) To preserve the sharp edges in depth map, a deep residual convolutional neural network is proposed for image guided depth enhancement.

(2) To avoid overfitting, a specific layer which introduces a local linear regularization constraint on the output depth is presented.

(3) A deep fusion strategy is proposed to extract the correlation between depth map and color image in the deep feature space.

Network Architecture

We propose a deep residual network (DRECNN) which employs a single residual unit to learn the residual image, as shown in Fig. 1. Our network consists of three components: depth branch, intensity branch and deep feature fusion part. Moreover, to achieve a wide receptive field, we set the number of convolution layers to 20 for all of the depth enhancement applications.

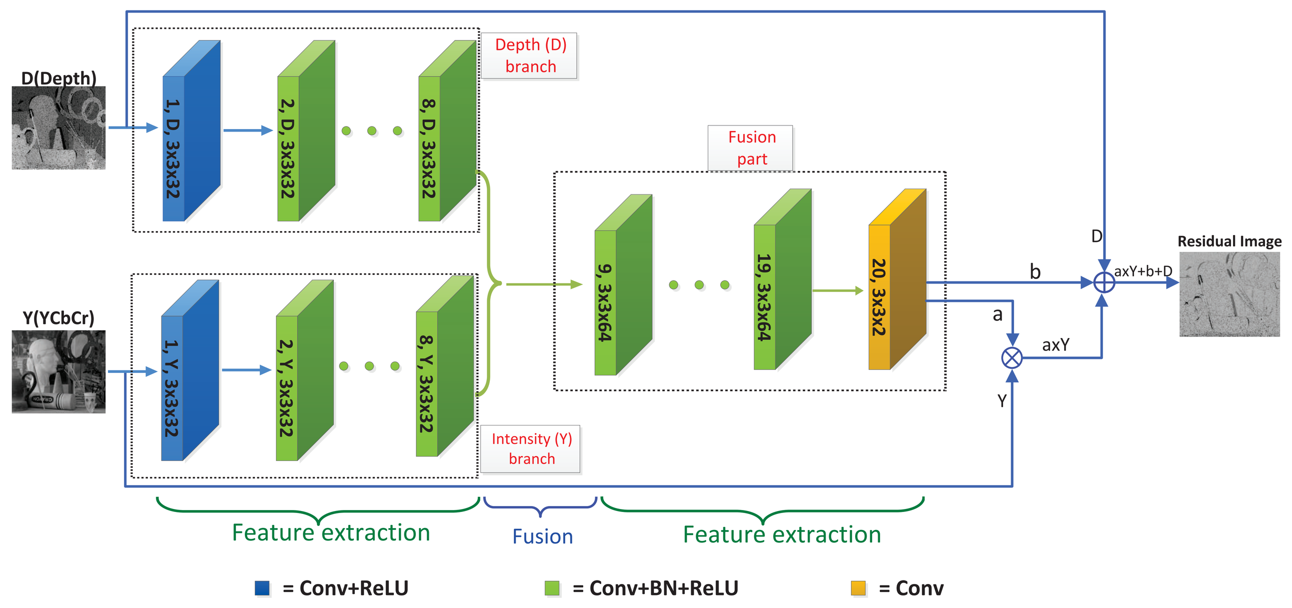

Figure: 1. Overview of our proposed network architecture.

A Local Linear Regularization Constraint

The image guided filter starts from a local linear model as follows:

$$q_i = a_k I_i + b_k, \forall i \in \omega_k. \tag{1}$$

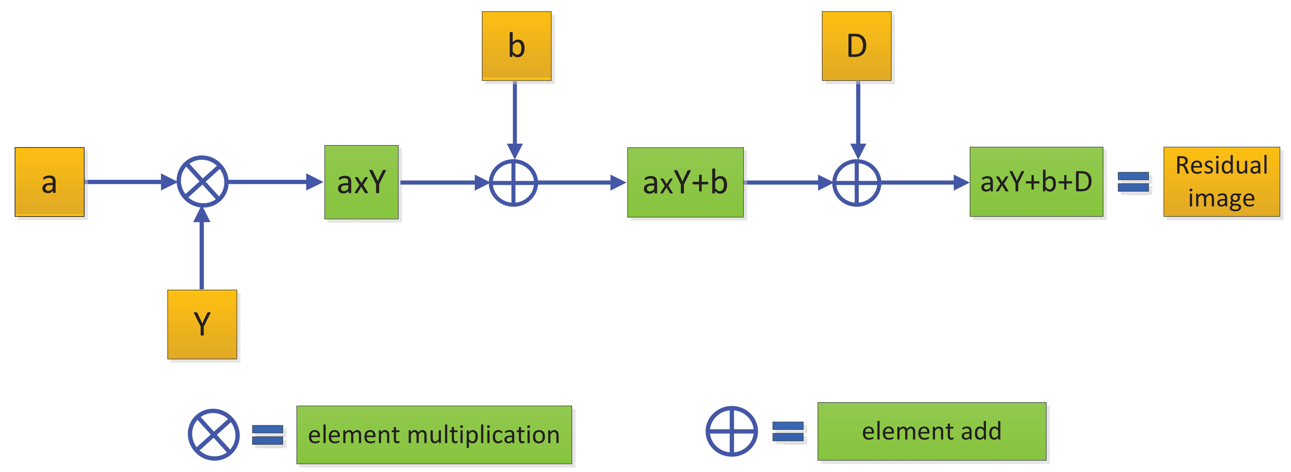

Figure: 2. The structure of the local linear regularization constraint.

Inspired by the local linear model, we propose a specific regularization (SR) layer which is shown as Fig. 2. Specifically, we first insert a convolutional layer to output the linear coefficient map *a* and *b*, and then compute $aY+b$ as the results of linear model. Finally, the residual map $aY+b-D$ is computed and supervised by the ground truth label ($aY+b+D$ is an equivalent expression as $aY+b-D$ as DRECNN learns opposite coefficients).

Applications and Experimental Results

(1) Performance Analysis on Regularized Layer.

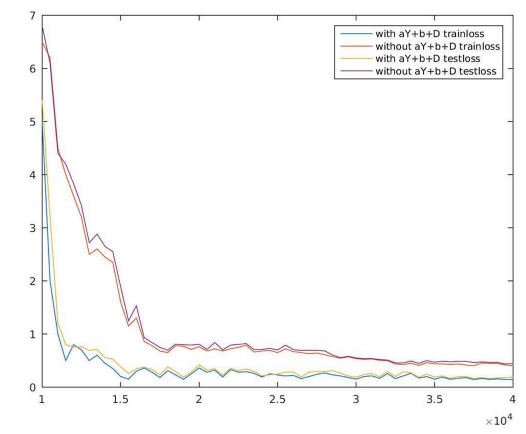

Figure: 3. The training and test loss of two different methods.

(2) Depth Blind Denosing.

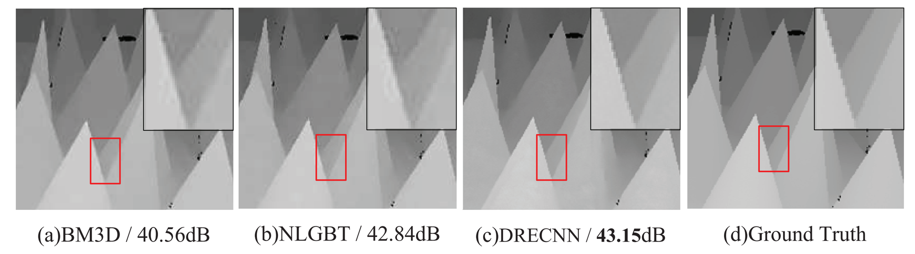

Figure: 4. Denoising results with noise level 10.

(3) Depth Super-resolution.

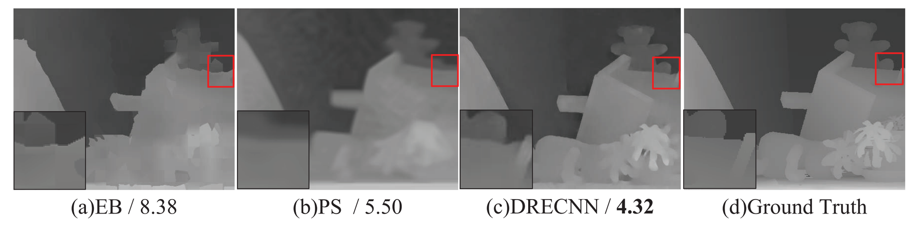

Figure: 5. Depth super-resolution results with upscaling factor 8.

(4) Depth Inpainting.

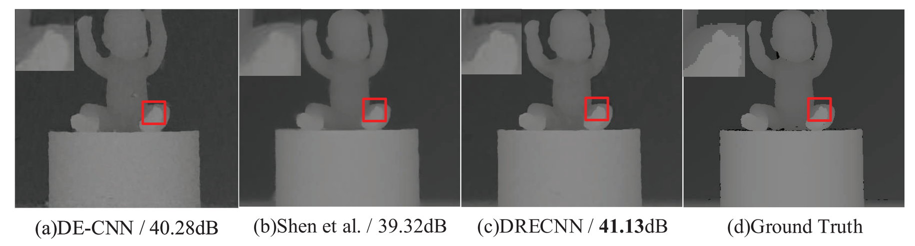

Figure: 6. Depth map inpainting results.

Table: Depth inpainting comparison in terms of PSNR (dB)

Methods	Art	Dolls	Reindeer	Laundry	Baby	Wood1	Teddy
DE-CNN	33.84	40.67	35.36	38.11	40.28	40.84	39.77
Shen et al.	33.42	**42.89**	37.19	38.71	39.32	41.54	40.89
DRECNN	**34.02**	41.59	**38.18**	**39.17**	**41.13**	**41.74**	**41.21**

Conclusion and Future Work

- A local linear regularization constraint is presented.
- Experiments demonstrated effectiveness of DRECNN.
- Future Work:Exploiting the structural correlation of RGB-D images more accurately to further improve the quality of depth map.

zj130129@mail.ustc.edu.cn, forrest@ustc.edu.cn