

Supplementary Material: Deep Hyperspectral-Depth Reconstruction Using Single Color-Dot Projection

Chunyu Li, Yusuke Monno, and Masatoshi Okutomi
Tokyo Institute of Technology, Tokyo, Japan

{lchunyu, ymonno}@ok.sc.e.titech.ac.jp, mxo@ctrl.titech.ac.jp

Network Architecture

We designed the architectures of the disparity estimation network and the spectral reconstruction network based on the Disparity Decoder presented in the previous work [1]. The detailed architectures are shown in Fig. 1 of the next page. Each network consists of a contractive part and an expanding part with long-range links between them. The contracting part contains convolution layers with the strides of 2, resulting in a total downsampling factor of 128. The expanding part of the network then gradually and nonlinearly upsamples the feature maps by also taking into account the features from the contractive part. Convolutional filter sizes decrease towards deeper layers of the network: 7×7 for the first and second layers, 5×5 for the following two layers and 3×3 for the layers after the fifth layer. In total, each network has 32 convolution layers and each of them is followed by ReLU. The final layer is followed by a scaled sigmoid non-linearity which constrains the output to the range between 0 and the maximum.

References

- [1] Gernot Riegler, Yiyi Liao, Simon Donne, Vladlen Koltun, and Andreas Geiger. Connecting the dots: Learning representations for active monocular depth estimation. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 7624–7633, 2019. 1

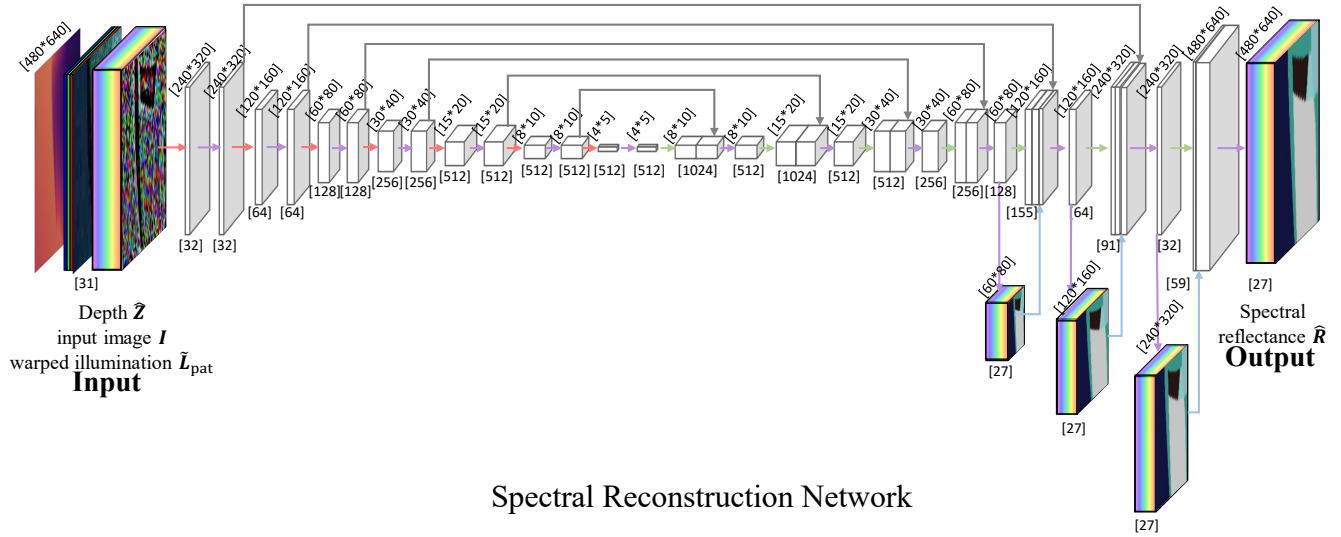
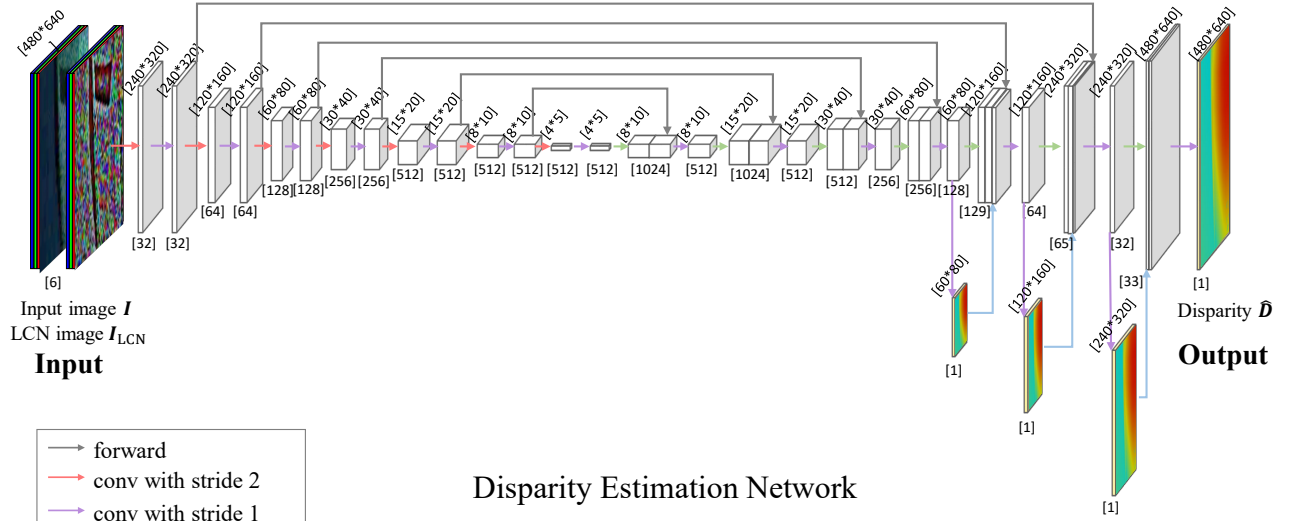


Figure 1. Detailed architectures of the disparity estimation network (top) and the spectral reconstruction network (bottom).