Multispectral illumination estimation using deep unrolling network

Yuqi Li Qiang Fu Wolfgang Heidrich
{yuqi.li, qiang.fu, wolfgang.heidrich}@kaust.edu.sa
KAUST, Saudi Arabia

1. Dataset

Our dataset contains hundreds of reflectance images. We capture various scenes containing a whiteboard with flat spectral reflectance as shown in Fig. 1, and calculate the reflectance spectra. The whiteboard also helps us get the ground truth illumination spectra in real experiments. We ensure a single, uniform illumination during capturing so that the calculated reflectance is reliable.

Figure 1. Left: A captured scene of our dataset. Right: The spectral reflectance function of the white reference board.

2. Training

In the first stage of the network, to ensure the initial reflectance $\mathbf{R}^{(0)} \leq 1$, we initialize $\mathbf{L}^{(0)}$ to:

$$\mathbf{L}^{(0)}(x,y,\lambda) = \max_{(x',y')\in\mathcal{N}(x,y)} \mathbf{I}^{(0)}(x',y',\lambda), \tag{1}$$

where $\mathcal{N}(x,y)$ is a neighborhood region of the pixel at coordinate (x,y). It is also a good choice to apply local Max-Spectra or GrayEdge algorithm [1] to initialize $\mathbf{L}^{(0)}$.

To reduce memory consumption and improve the inference speed, we downsample the illumination and reflectance images into $4 \times 4 = 16$ subimages in spatial dimension with a scale factor of 4, feed the subimages to the regularization blocks, and recover the original spatial resolution from the 16 updated subimages. The processing of a single $256 \times 256 \times 31$ image can be accomplished in three seconds. Limited by the page length, only two results of the real images are shown in the paper, here Figure 2 gives two more results, and we believe that they are sufficient to prove the effectiveness of our methods on the real scenes with a large variety.

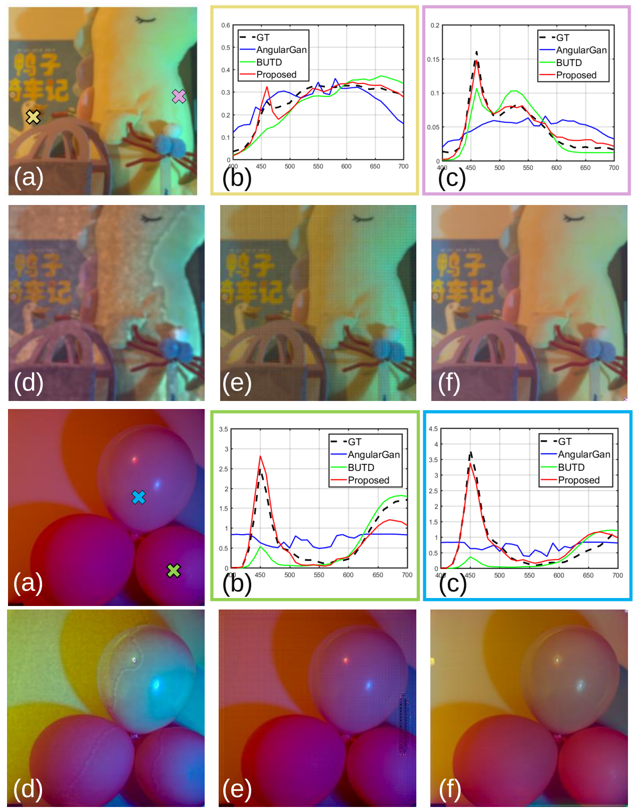

Figure 2. Two results on real images. (a) The RGB image rendered from captured multispectral images. (b-c) The comparisons of the estimated spectra of the three methods at the positions of two crosses in (a). (d-f) The comparison of the RGB images rendered from the estimated reflectance images of the three methods(from left to right: BUTD, AngularGAN, and ours).

3. Discussion

Compared with previous methods Although previous methods either make assumptions on smooth illumination, amount of light sources, or similarity of spectral reflectances, as far as we know, none of them embed these constraints into an optimization model as priors to handle single/multiple illumination spectra estimation. Our network is flexible in generalizing to spectra that are not trained on, it can learn representative spatial-spectral features for illumination separation and an optimal combination of parameters for an efficient estimation. Through observing the simulation and experiments, we argue that the proposed architecture is more suitable to the illumination spectral estimation problem when compared to the CNN-based methods. The improvement brought by each prior is presented in the ablation study. Beyond the model, we built an unrolling network mainly for three advantages, and the network is not equivalent to the basic CNMF model anymore. Firstly, the non-local denoising subnetwork enhances the representation capacity of the model to deal with illumination separation; Next, relaxing the variables accelerate the convergence of the method and can reduce the amount of iterations. Finally, the optimal parameters help us understand the real contribution of each subnetwork in each iteration.

Why model-based rather than CNN-based Even with more than 10M parameters (more than ten times of ours), an Alexnet-like CNN method(PWIR) and a GAN-based method(AngularGAN) both show poor generalization ability in simulations and real experiments. CNN specializes in object recognition but requires a huge amount of data for feature extraction. The lack of multispectral reflectance data and the high dimensionality of illumination spectra limit the application of CNN on the illumination spectra estimation problem. While trained with the same dataset, our network can provide much more accurate and robust results.

References

[1] Haris Ahmad Khan, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Olivier Laligant. Illuminant estimation in multispectral imaging. *JOSA A*, 34(7):1085–1098, 2017. 1