AN OPTICAL PHYSICS INSPIRED CNN APPROACH FOR INTRINSIC IMAGE DECOMPOSITION

Harshana Weligampola[1] Gihan Jayatilaka[2] Suren Sritharan[1] Parakrama Ekanayake[2] Roshan Ragel[2] Vijitha Herath[2] Roshan Godaliyadda[2]
[1]Faculty of IT and Computing, Sri Lanka Technological Campus, Padukka [10500], Sri Lanka
[2]Faculty of Engineering, University of Peradeniya [20400], Sri Lanka

Overview

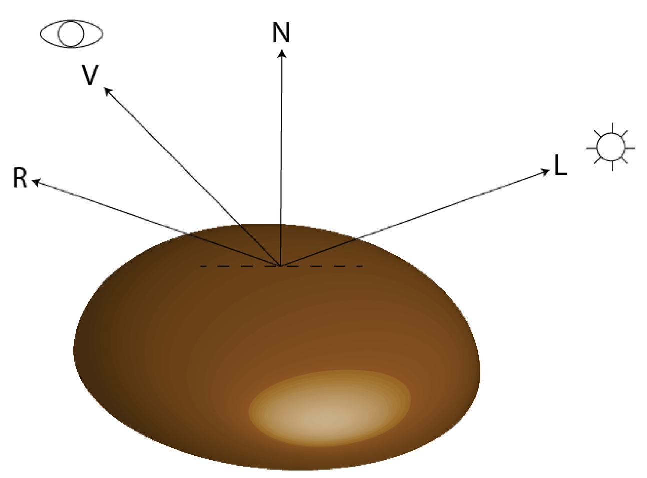
Generating reflectance and shading from a single image is a challenging task when there is **no ground truth**. We propose a novel reflectance approximation map to train the neural network and a **physics-based loss function** to learn intrinsic properties in an image. Through numerical evaluation metrics, we show that the proposed model performs consistently well with different datasets consist of variety of scenes. There is room for improvement in regards to the color leakage problem in the shading map.

Problem

Intrinsic Image Decomposition (IID) is the problem of decomposing an image into its constituents

Reflectance   Original image   Shading

Image Interpretation



The perceived pixel intensity of an image is given by the Phong's model. Main components of an image: Ambient, Diffused, Specular

$$I_p = \int_{\lambda} k_a r_p(\lambda) i_a(\lambda) + \sum_{\hat{L}^{(n)} \in \mathbf{L}} \{k_d r_p(\lambda)[\hat{L}^{(n)}.\, \hat{N}_p] i_d^{(n)}(\lambda) + k_s s_p(\lambda)[\hat{R}^{(n)}.\, \hat{V}]^{\gamma} i_s^{(n)}(\lambda)\} d\lambda$$

Consider a narrow band and assuming that ambient illumination is constant, only one light source exists, and specular term is negligible,

$$I_p(\lambda_c) = r_p(\lambda_c)[\hat{L}.\, \hat{N}_p] i_d(\lambda_c)$$

Derivations from the image

Reflectance Ratio Gradient (RRG)

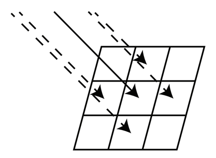
◆ *Identify the boundaries of the uniform reflectance in an image.*

The log ratio of pixel intensity between 2 narrow band wavelengths is,

$$\mathcal{J}_p(\lambda_a, \lambda_b) = log\left(\frac{I_p(\lambda_a)}{I_p(\lambda_b)}\right) = log\left(\frac{r_p(\lambda_a) i_d(\lambda_a)}{r_p(\lambda_b) i_d(\lambda_b)}\right)$$

Taking the gradient assuming that single wave-length intensity of adjacent pixels are equal

$$\nabla \mathcal{J}(\lambda_a, \lambda_b) = \nabla log\left(\frac{r(\lambda_a)}{r(\lambda_b)}\right)$$



Shading Gradient (SG)

◆ *Gradient of the shading map of an image.*

Consider the natural logarithm gradient of the narrow band wavelength of a pixel. Assuming that the neighbouring pixels have same reflectance,

$$\mathcal{K}_p(\lambda_a) = log\left(I_p(\lambda_a)\right)$$

$$\nabla \mathcal{K}(\lambda_a) = \mathcal{K}_{p_1}(\lambda_a) - \mathcal{K}_{p_2}(\lambda_a) = \nabla log\left([\hat{L}.\, \hat{N}]\right)$$

Map of invalid pixels can be derived from RRG as,

$$M_{RRG} = \left[m_p^{(c)}\right] = \begin{cases} (\nabla \mathcal{J}(\lambda_R, \lambda_G) + \nabla \mathcal{J}(\lambda_R, \lambda_B))/2 & \text{if } c = R \\ (\nabla \mathcal{J}(\lambda_G, \lambda_B) + \nabla \mathcal{J}(\lambda_G, \lambda_R))/2 & \text{if } c = G \\ (\nabla \mathcal{J}(\lambda_B, \lambda_G) + \nabla \mathcal{J}(\lambda_B, \lambda_R))/2 & \text{if } c = B \end{cases}$$

SG is given by,

$$\nabla \mathcal{K}(\lambda_c) = \begin{cases} \nabla log\left([\hat{L}.\, \hat{N}]\right) & \text{if } M_{RRG}^{(c)} < 0.1 \\ 0 & \text{otherwise} \end{cases}$$

Reflection Approximation Map (RAM)

◆ *Approximate reflectance of an image.*

if $I_p(\lambda_a) \approx I_p(\lambda_b)$ then $\mathcal{J}_p(\lambda_a, \lambda_b) = 0$

if $I_p(\lambda_a) \gg I_p(\lambda_b)$ then $\mathcal{J}_p(\lambda_a, \lambda_b) \gg 0$

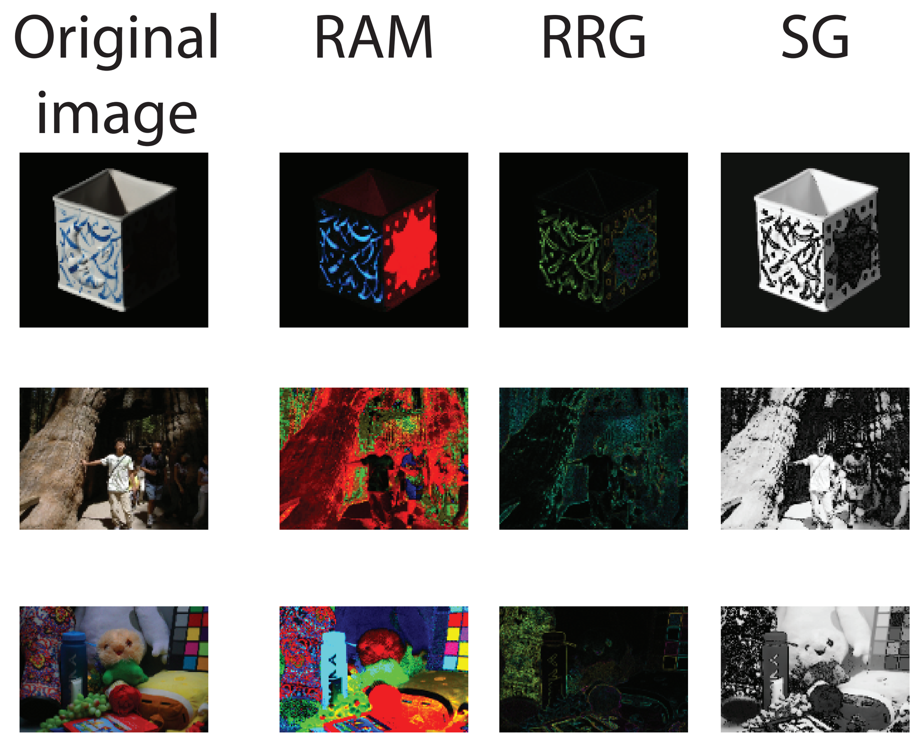
Based on this we can approximate the reflectance for red channel as follows,

$$RAM_R = \frac{\bar{\mathcal{J}}_p(\lambda_R, \lambda_G) + \bar{\mathcal{J}}_p(\lambda_R, \lambda_B)}{2}$$

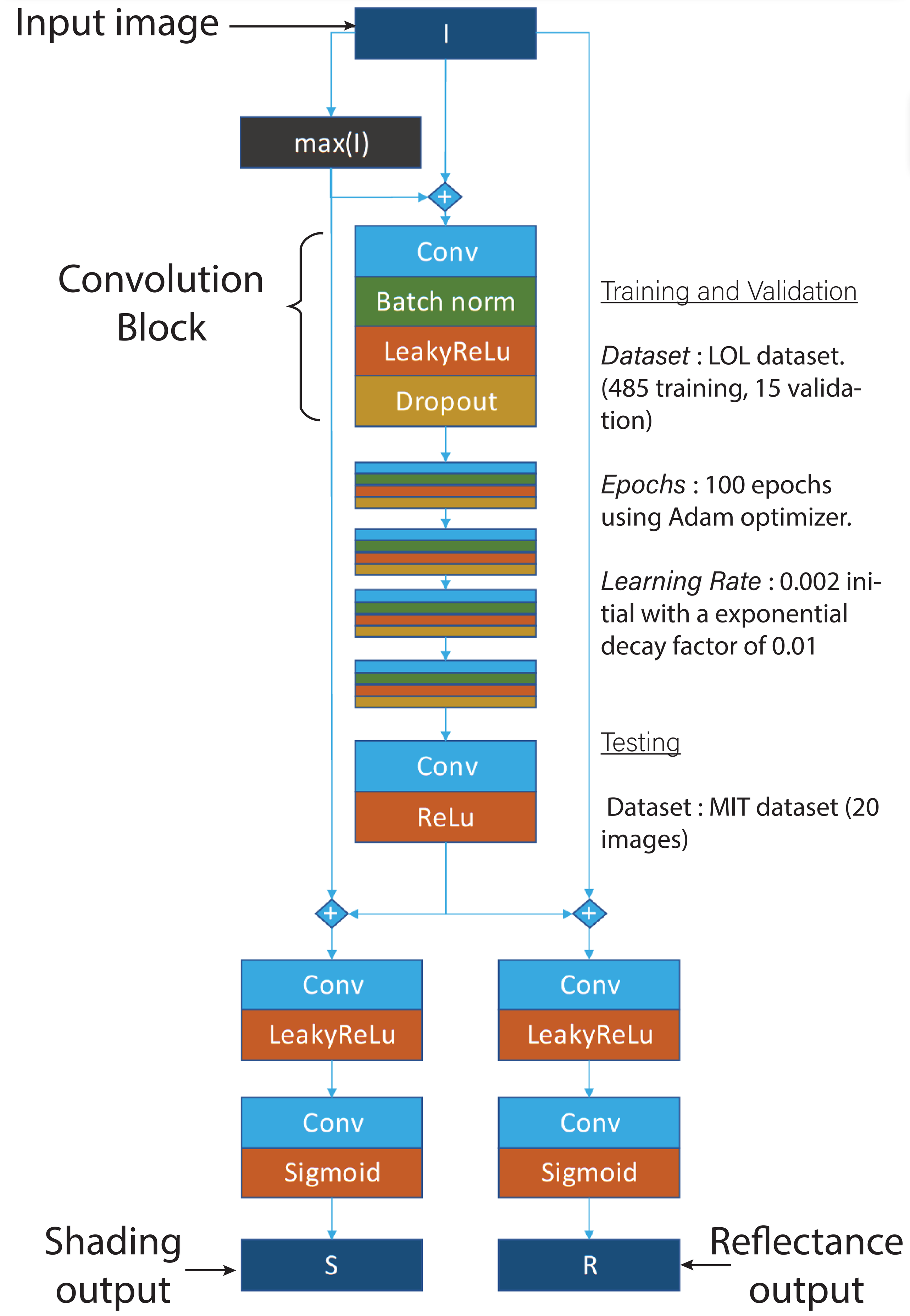
Similarly we can get the RAM for blue and green channels as well.

$$M_{RAM} = \left[m_p^{(c)}\right] = \begin{cases} (\overline{\mathcal{J}}_p(\lambda_R, \lambda_G) + \overline{\mathcal{J}}_p(\lambda_R, \lambda_B))/2 & \text{if } c = R \\ (\overline{\mathcal{J}}_p(\lambda_G, \lambda_R) + \overline{\mathcal{J}}_p(\lambda_G, \lambda_B))/2 & \text{if } c = G \\ (\overline{\mathcal{J}}_p(\lambda_B, \lambda_G) + \overline{\mathcal{J}}_p(\lambda_B, \lambda_R))/2 & \text{if } c = B \end{cases}$$

RRG, SG, RAM



Network architecture



Training and Validation

Dataset : LOL dataset. (485 training, 15 validation)

Epochs : 100 epochs using Adam optimizer.

Learning Rate : 0.002 initial with a exponential decay factor of 0.01

Testing

Dataset : MIT dataset (20 images)

Loss functions

$$\mathcal{L} = \alpha_1 \mathcal{L}_{recon} + \alpha_2 \mathcal{L}_{ss} + \alpha_3 \mathcal{L}_{rrg} + \alpha_4 \mathcal{L}_{sg} + \alpha_5 \mathcal{L}_{ram}$$

Reconstruction loss

$$\mathcal{L}_{recon} = ||\mathbf{R}_i \mathbf{S}_i - \mathbf{I}_i||_1$$

Shading smoothness loss

$$\mathcal{L}_{ss} = ||\nabla \mathbf{S}_i \exp(-10 f_{RRG}(i))||_1$$

Reflectance Ratio Gradient (RRG) loss

$$\mathcal{L}_{rrg} = ||f_{RRG}(\mathbf{R}_i) - f_{RRG}(i)||_1$$

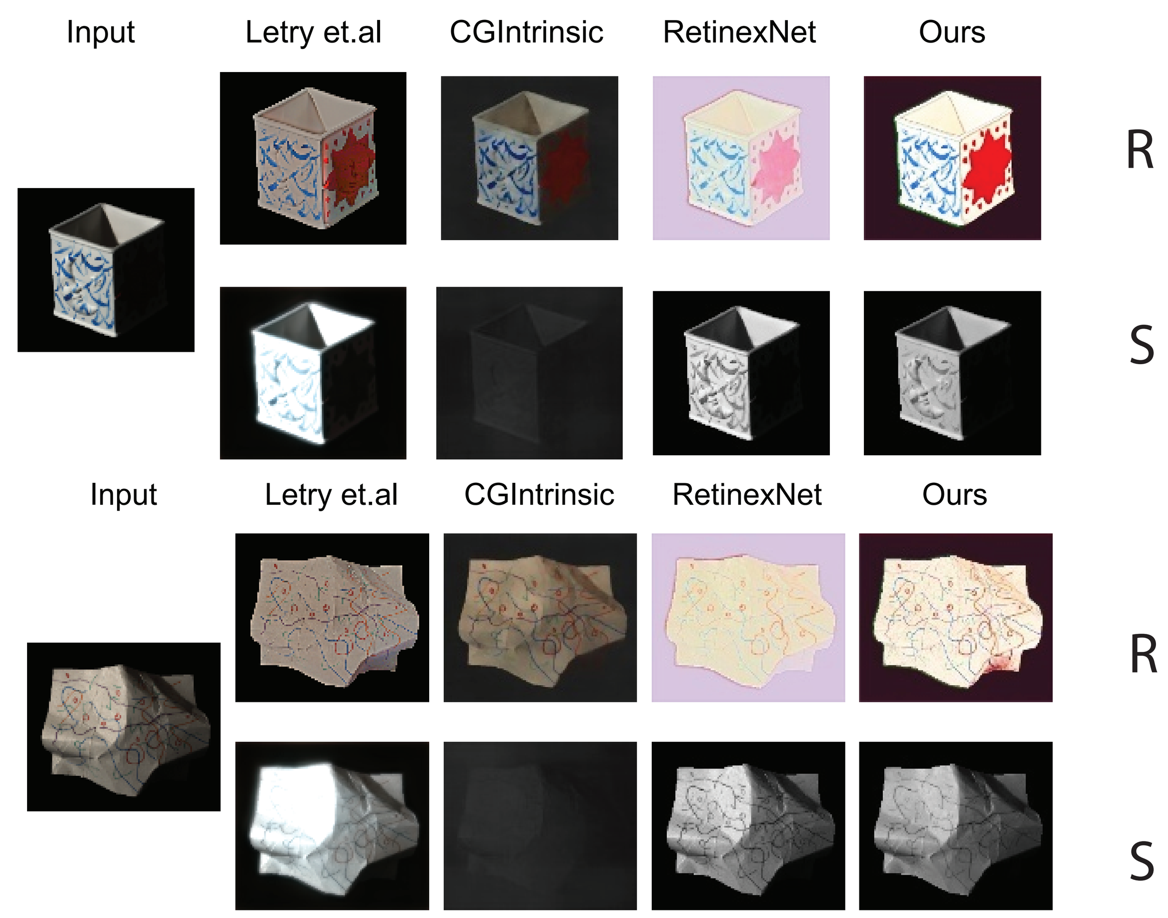
Shading Gradient (SG) loss

$$\mathcal{L}_{sg} = ||(\nabla \log(\mathbf{S_i}) - f'_{SG}(i)) \times f'_{SG}(i)||_1$$

Reflectance Approximation Map (RAM)lss

$$\mathcal{L}_{ram} = ||(\mathbf{R}_i - f_{RAM}(i)) \times f_{RAM}(i)||_1$$

Results



Method \ Metric	LOL images			
	RMSE	PSNR	SSIM	NIQE
Letry et.al	21.87	35.28	0.96	7.55
CGIntrinsic	63.28	18.95	0.36	14.78
Retinex-net	6.88	34.64	0.90	7.63
Ours	2.00	43.12	0.95	7.63

Method \ Metric	MIT Images				MIT (R)		MIT (S)	
	RMSE	PSNR	SSIM	NIQE	RMSE	PSNR	RMSE	PSNR
Letry et.al	6.67	39.26	**0.99**	12.06	**41.91**	**16.58**	40.88	16.46
CGIntrinsic	40.95	17.36	0.11	**17.47**	48.47	16.28	59.62	12.99
Retinex-net	3.77	37.85	0.95	14.02	67.39	13.48	37.97	18.54
Ours	**1.04**	**41.66**	0.96	14.02	45.90	15.82	**30.54**	**20.14**